Vergleich von statistischen Tests im verbundenen und unabhängigen Stichprobenfall

Von der Fakultät für Mathematik und Physik
der Gottfried Wilhelm Leibniz Universität Hannover

zur Erlangung des Grades
Doktor der Naturwissenschaften
Dr. rer. nat.

genehmigte Dissertation
von

Daniel Gaigall, M.Sc.

2016

Referent: Prof. Dr. Ludwig Baringhaus

Korreferent: Prof. Dr. Norbert Henze

Tag der Promotion: 09.02.2016

Zusammenfassung

Es werden Effizienzbegriffe zum Vergleich von statistischen Tests basierend auf verschiedenen statistischen Experimenten eingeführt. Dabei handelt es sich um die schon aus dem Vergleich von statistischen Tests in je demselben Modell bekannten asymptotischen relativen Effizienzen wie die Hodges-Lehmann-Effizienz, die Bahadur-Effizienz und die Pitman-Effizienz sowie um Kriterien basierend auf Volumina von Konfidenzbereichen. Effizienzaussagen werden unter anderem für Likelihood-Quotienten-Tests und Waldsche Tests im Rahmen eines allgemeinen multivariaten parametrischen Modells erhalten. Statistische Tests zur Prüfung von Hypothesen über die relative Wirksamkeit zweier Experimente werden vorgeschlagen. Auf der Grundlage der erhaltenen Ergebnisse erfolgt ein Vergleich der Wirksamkeit von korrespondierenden Verfahren bei verbundener Stichprobenerhebung und unabhängiger Stichprobenerhebung. Die Rolle der Kovarianzmatrix bei verbundener Stichprobenerhebung wird insbesondere unter der Annahme, dass die zugrunde liegenden Verteilungen durch k-parametrische Exponentialfamilien modellierbar sind, herausgearbeitet. Verbindungen zu Effizienzbegriffen bei Punkt- und Konfidenzbereichsschätzverfahren werden aufgezeigt. Ausführlichere Untersuchungen betreffen die korrespondierenden Hotellingschen T^2-Tests im multivariaten Normalverteilungsfall, die klassischen Homogenitätstests bei $k \times k$-Kontingenztafeln und die Wilcoxon Tests in nichtparametrischen Lagealternativmodellen.

Schlagworte: Vergleich von Experimenten, Hypothesentests, Effizienz

Abstract

Efficiency notions for the comparison of statistical tests based on different statistical experiments are introduced. The notions under consideration are those already known for the comparisons of procedures in the same underlying statistical model, i.e., the Hodges-Lehmann efficiency, the Bahadur efficiency and the Pitman efficiency, as well as criteria based on volumes of confidence sets. Efficiency results are presented for likelihood-ratio tests and Wald tests in a general multivariate parametric setting. Statistical tests for testing hypotheses on the relative efficiency of two experiments are suggested. Based on the results obtained, a comparison is done for corresponding procedures in paired samples and independent samples. The role of the covariance matrix in the paired sample case is discussed, especially for the k-parametric exponential family model. Connections to efficiency concepts in point estimation and set estimation are shown. Detailed analysis is done for the T^2-tests of Hotelling in the multivariate normal case, the classical homogeneity tests in $k \times k$-contingency tables and the Wilcoxon tests in nonparametric location models.

Keywords: comparison of experiments, testing hypotheses, efficiency

Danksagung

Mein besonderer Dank gilt Prof. Dr. Ludwig Baringhaus für seine Anregungen, seine Beratung und seine Betreuung bei der Entstehung dieser Arbeit. Ebenso bin ich meiner Lebensgefährtin Ina Lewandrowski zu großem Dank verpflichtet, die mir in meinem Promotionsvorhaben stets zur Seite stand. Bei der Hans-Böckler-Stiftung bedanke ich mich für die Förderung meiner Promotion. Prof. Dr. Stefan Tappe danke ich für die anregenden Gespräche, seine Unterstützung sowie für seine Mitarbeit in der Promotionskommission. Herzlich bedanke ich mich auch bei Prof. Dr. Norbert Henze für das Korreferat und bei Prof. Dr. Marc Steinbach für die Übernahme des Vorsitzes der Promotionskommission. Prof. Dr. Arnold Janssen danke ich sehr für sein Entgegenkommen und seine Unterstützung beim Verfassen dieser Arbeit.

Inhaltsverzeichnis

1 Einleitung

1.1 Problemstellung

Wir erläutern die zu behandelnde Thematik anhand eines einführenden Anwendungsbeispiels. Die Konstrukteure eines Autoherstellers vermuten, dass der Reifenabrieb eines neuen Automodells mit Frontantrieb bei den Vorderreifen und Hinterreifen verschieden ist. Unabhängige Messungen an $2n$ Fahrzeugen des Modells liefern die Beobachtungswerte

$$(x_i^\ell, x_i^r, y_i^\ell, y_i^r),\ i = 1, \ldots, 2n,$$

mit x_i^ℓ bzw. x_i^r als Reifenabrieb beim linken bzw. rechten Vorderreifen des i-ten Fahrzeugs und y_i^ℓ bzw. y_i^r als Reifenabrieb beim linken bzw. rechten Hinterreifen des i-ten Fahrzeugs. Mit begleitender Dokumentation zum Messverfahren werden einem Statistiker A die ermittelten Werte

$$x_i = (x_i^\ell, x_i^r),\ i = 1, \ldots, n, \quad \text{und} \quad y_i = (y_i^\ell, y_i^r),\ i = 1, \ldots, n, \tag{1.1.1}$$

und einem weiteren Statistiker B die ermittelten Werte

$$x_i' = (x_i^\ell, x_i^r),\ i = 1, \ldots, n, \quad \text{und} \quad y_i' = (y_{n+i}^\ell, y_{n+i}^r),\ i = 1, \ldots, n, \tag{1.1.2}$$

zur Verfügung gestellt, jeweils mit der Bitte, auf der Basis der erhaltenen Werte Stellung zum vermuteten unterschiedlichen Abrieb bei Vorderreifen und Hinterreifen zu nehmen.

Wissend, dass x_i^ℓ, x_i^r und y_i^ℓ, y_i^r bei demselben Fahrzeug i ermittelt wurden, wird der Statistiker A eine Abhängigkeit zwischen x_i und y_i nicht ausschließen und seine Aussage auf der Grundlage eines diese Abhängigkeit berücksichtigenden statistischen Verfahrens treffen. In der Terminologie der Statistik handelt es sich bei (1.1.1) um eine verbundene Stichprobe. Ein auf einer verbundenen Stichprobe fußendes statistisches Verfahren ist ein Verfahren bei verbundener Stichprobenerhebung bzw. beim verbundenen Stichprobenfall. Messungen an verschiedenen Fahrzeugen vor Augen, wird der Statistiker B nicht von Abhängigkeiten zwischen den ermittelten Werten x_i' und y_i' für die Vorderreifen und den ermittelten Werten für die Hinterreifen ausgehen und seine Stellungnahme auf der Basis eines die Unabhängigkeit der Messwerte gründenden statistischen Verfahren formulieren. In der Sprache der Statistik handelt es sich bei (1.1.2) um zwei unabhängige Stichproben. Ein auf einer Erhebung von zwei unabhängigen Stichproben aufsetzendes statistisches Verfahren ist ein Verfahren bei unabhängiger Stichprobenerhebung bzw. beim unabhängigen Stichprobenfall.

Ausgehend davon, dass die von den Statistikern einzusetzenden adäquaten statistischen Verfahren je vom selben Typ (statistischer Test, Punktschätzverfahren oder Konfidenzbereichsschätzverfahren) sind, ist ein Vergleich ihrer Wirksamkeit naheliegend. Im günstigen Fall kann als Folgerung aus einem solchen Vergleich eine Aussage zur Überlegenheit einer der beiden Stichprobenerhebungsarten getroffen werden.

Der Fragestellung entsprechend bringt hier jeder der beiden Statistiker einen statistischen Test an. Die $(x_1, y_1), \ldots, (x_{2n}, y_{2n})$ werden aufgefasst als Beobachtungen von unabhängigen und identisch verteilten Tupeln $(X_1, Y_1), \ldots, (X_{2n}, Y_{2n})$ von zweidimensionalen Zufallsvektoren, deren unbekannte gemeinsame Verteilung $\mathcal{L}(X_1, Y_1)$ Element einer Verteilungsfamilie $\mathcal{F}$ ist. Die Behandlung der Frage, ob die Verteilung $\mathcal{L}(X_1)$ von X_1 gleich der Verteilung $\mathcal{L}(Y_1)$ von Y_1 ist (Hypothese H), oder nicht (Alternative K), in der Sprache der Statistik, die Behandlung des statistischen Testproblems

$$\mathrm{H}: \ \mathcal{L}(X_1) = \mathcal{L}(Y_1), \ \mathrm{K}: \ \mathcal{L}(X_1) \neq \mathcal{L}(Y_1), \tag{1.1.3}$$

erfolgt durch den Statistiker A bzw. den Statistiker B durch Bereitstellung eines auf den verbundenen Stichprobenvariablen $(X_1, Y_1), \ldots, (X_n, Y_n)$ bzw. auf den unabhängigen Blöcken $X_1, \ldots, X_n$ und $Y_{n+1}, \ldots, Y_{2n}$ von Stichprobenvariablen basierenden statistischen Tests. Das Hauptaugenmerk dieser Arbeit ist gerichtet auf einen Wirksamkeitsvergleich dieser Tests, allgemeiner solcher für Testprobleme, die ähnlich zu (1.1.3) gelagert sind. In der Regel reichen elementare Überlegungen nicht aus, damit ein zufriedenstellender Vergleich gelingt.

1.1.1 Beispiel: Im Rahmen einer verbundenen Stichprobenerhebung gehen wir aus von unabhängigen 2-dimensionalen Zufallsvektoren $\binom{X_1}{Y_1}, \ldots, \binom{X_n}{Y_n}$, die je dieselbe bivariate Normalverteilung

$$\mathrm{N}_2\left(\begin{pmatrix} a \\ b \end{pmatrix}, \begin{pmatrix} \sigma^2 & \sigma^2\rho \\ \sigma^2\rho & \sigma^2 \end{pmatrix} \right)$$

haben. Es seien $\sigma^2 > 0$ und $\rho \in (-1, 1)$ bekannt und $a, b \in \mathbb{R}$ unbekannt. Zur Behandlung des Testproblems 1.1.3, welches hier äquivalent ist zum Testproblem

$$\mathrm{H}: \ a = b, \ \mathrm{K}: \ a \neq b,$$

schlagen wir bei gegebenem Testniveau α den auf der Testgröße

$$T_n = \frac{\sqrt{n}|\overline{X}_n - \overline{Y}_n|}{\sqrt{2\sigma^2(1-\rho)}},$$

mit

$$\overline{X}_n = \frac{1}{n}\sum_{j=1}^{n} X_j \ \text{ und } \ \overline{Y}_n = \frac{1}{n}\sum_{j=1}^{n} Y_j,$$

basierenden statistischen Test

$$\text{„Verwirf H, falls } T_n \geq u_{1-\alpha/2}\text{“}$$

vor. Hierbei ist $u_{1-\alpha/2}$ das $(1-\alpha/2)$-Quantil der Standardnormalverteilung. Bei Gültigkeit der Hypothese wird für jeden zugrunde liegenden Parameter $a = b$ das Testniveau α exakt eingehalten. Wir definieren für $y > 0$ die Abbildung

$$\chi_y(x) := \Phi(y - x) - \Phi(-y - x),\ x \in \mathbb{R},$$

mit Φ als Verteilungsfunktion der Standardnormalverteilung. Bei Gültigkeit der Alternative mit zugrunde liegenden Parametern $a, b \in \mathbb{R}$ mit $\mu := a - b \neq 0$ ist die Wahrscheinlichkeit für einen Fehler 2. Art

$$\begin{aligned}
\beta_n :&= P_{a,b}\Bigg(- u_{1-\alpha/2} \leq \frac{\sqrt{n}(\overline{X}_n - \overline{Y}_n)}{\sqrt{2\sigma^2(1-\rho)}} \leq u_{1-\alpha/2}\Bigg) \\
&= \Phi\Bigg(u_{1-\alpha/2} - \frac{\sqrt{n}}{\sqrt{2\sigma^2(1-\rho)}}\mu\Bigg) - \Phi\Bigg(- u_{1-\alpha/2} - \frac{\sqrt{n}}{\sqrt{2\sigma^2(1-\rho)}}\mu\Bigg) \\
&= \chi_{u_{1-\alpha/2}}\Bigg(\frac{\sqrt{n}}{\sqrt{2\sigma^2(1-\rho)}}\mu\Bigg).
\end{aligned}$$

Im unabhängigen Stichprobenfall gehen wir aus von unabhängigen 2-dimensionalen Zufallsvektoren $\binom{X_1'}{Y_1'}, \dots, \binom{X_n'}{Y_n'}$, die je dieselbe bivariate Normalverteilung

$$\mathrm{N}_2\left(\begin{pmatrix} a \\ b \end{pmatrix}, \begin{pmatrix} \sigma^2 & 0 \\ 0 & \sigma^2 \end{pmatrix}\right)$$

haben. Wir betrachten

$$T_n' = \frac{\sqrt{n}|\overline{X_n'} - \overline{Y_n'}|}{\sqrt{2\sigma^2}},$$

mit

$$\overline{X_n'} = \frac{1}{n}\sum_{j=1}^{n} X_j' \quad \text{und} \quad \overline{Y_n'} = \frac{1}{n}\sum_{j=1}^{n} Y_j',$$

und schlagen bei gegebenem Testniveau $\alpha \in (0, 1)$ zur Behandlung des Testproblems

$$\mathrm{H}:\ a = b,\ \mathrm{K}:\ a \neq b$$

den zweiseitigen Zweistichproben-Gauß-Test

$$\text{„Verwirf H, falls } T_n' \geq u_{1-\alpha/2}\text{“}$$

vor. Auch dieser Test hält bei Gültigkeit der Hypothese für jeden zugrunde liegenden Parameter $a = b$ das Testniveau α exakt ein. Bei Gültigkeit der Alternative mit zugrunde liegenden Parametern $a, b \in \mathbb{R}$ mit $\mu = a - b \neq 0$ ist die Wahrscheinlichkeit für einen Fehler 2. Art

$$\begin{aligned} \beta_n' &:= P_\mu\bigg(-u_{1-\alpha/2} \leq \frac{\sqrt{n}(\overline{X_n'} - \overline{Y_n'})}{\sqrt{2\sigma^2}} \leq u_{1-\alpha/2} \bigg) \\ &= \Phi\bigg(u_{1-\alpha/2} - \frac{\sqrt{n}}{\sqrt{2\sigma^2}}\mu \bigg) - \Phi\bigg(-u_{1-\alpha/2} - \frac{\sqrt{n}}{\sqrt{2\sigma^2}}\mu \bigg) \\ &= \chi_{u_{1-\alpha/2}}\bigg(\frac{\sqrt{n}}{\sqrt{2\sigma^2}}\mu \bigg). \end{aligned}$$

Für die Ableitung der Abbildung χ_y gilt für jedes $y > 0$

$$\chi_y'(x) = -\varphi(y-x) + \varphi(-y-x) = -\varphi(y-x) + \varphi(y+x) \begin{cases} < 0, & x > 0 \\ > 0, & x < 0 \end{cases} ,$$

mit φ als Dichtefunktion der Standardnormalverteilung. Daher ist die Abbildung χ_y für jedes $y > 0$ streng monoton fallend auf $(0, \infty)$ und streng monoton wachsend auf $(-\infty, 0)$. Wir stellen für $\mu > 0$

$$\beta_n \left\{ \begin{matrix} < \\ = \\ > \end{matrix} \right\} \beta_n' \ \forall\, n \in \mathbb{N} \ \Leftrightarrow \ \frac{\sqrt{n}}{\sqrt{2\sigma^2(1-\rho)}}\mu \left\{ \begin{matrix} > \\ = \\ < \end{matrix} \right\} \frac{\sqrt{n}}{\sqrt{2\sigma^2}}\mu \ \forall\, n \in \mathbb{N} \ \Leftrightarrow \ \rho \left\{ \begin{matrix} > \\ = \\ < \end{matrix} \right\} 0,$$

und für $\mu < 0$

$$\beta_n \left\{ \begin{matrix} < \\ = \\ > \end{matrix} \right\} \beta_n' \ \forall\, n \in \mathbb{N} \ \Leftrightarrow \ \frac{\sqrt{n}}{\sqrt{2\sigma^2(1-\rho)}}\mu \left\{ \begin{matrix} < \\ = \\ > \end{matrix} \right\} \frac{\sqrt{n}}{\sqrt{2\sigma^2}}\mu \ \forall\, n \in \mathbb{N} \ \Leftrightarrow \ \rho \left\{ \begin{matrix} > \\ = \\ < \end{matrix} \right\} 0$$

fest. Legen wir die Wahrscheinlichkeit für einen Fehler 2. Art als Qualitätsmaßstab an, ziehen wir die verbundene Stichprobenerhebung der unabhängigen Stichprobenerhebung vor, wenn die Korrelation ρ zwischen X_1 und Y_1 positiv ist; dagegen ziehen wir die unabhängige Stichprobenerhebung der verbundenen Stichprobenerhebung vor, wenn die Korrelation negativ ist. Verschwindet die Korrelation, so sind beide Verfahren gleichwertig.

Im Allgemeinen sind Wahrscheinlichkeiten für einen Fehler 2. Art schwer zu bestimmen, sodass ein wie im vorhergehenden Beispiel vorgenommener einfacher Vergleich nicht möglich ist. Der sich anbietende Ausweg, auf asymptotischen Aussagen für Folgen von Test- oder Schätzgrößen fußende Vergleiche durchzuführen, wird in dieser Arbeit durchweg beschritten. Wir starten mit einer Vorstellung der wichtigsten Effizienzbegriffe für Folgen von statistischen Tests. Es werden Modifikationen und Ergänzungen der bekannten Ansätze und Ergebnisse vorgenommen. Ein Effizienzbegriff basierend auf dem Volumen von Konfidenzbereichen wird vorgestellt, Zusammenhänge mit den traditionellen Effizienzbegriffen werden aufgezeigt.

Die hier betrachteten statistischen Tests bei verbundener Stichprobenerhebung einerseits und bei unabhängiger Stichprobenerhebung andererseits sind auffassbar als statistische Verfahren in verschiedenen Experimenten, wobei diese Experimente ihrerseits auf einem gemeinsamen Grundexperiment aufsetzen. Losgelöst von diesen beiden speziellen Experimenten stellen wir im vierten Kapitel Wirksamkeitsvergleiche von Experimenten mit zugrunde liegendem gemeinsamen Grundexperiment, einem statistischen Raum $(\Omega, \mathfrak{A}, \mathfrak{P})$, $\mathfrak{P} = \{P_\vartheta; \vartheta \in \Theta\}$ eine durch den Parameter ϑ aus der Parametermenge Θ indizierte Familie von Wahrscheinlichkeitsmaßen auf dem Messraum $(\Omega, \mathfrak{A})$, an.

Lässt sich eine Teilmenge Θ^* von Θ angegeben mit der Eigenschaft, dass bei zwei im Vergleich stehenden Experimenten E_1 und E_2 bei zugrunde liegendem Parameter $\vartheta \in \Theta^*$ das Experiment E_1 wirksamer ist als das Experiment E_2, so ist die (Vorab-)Anbringung eines statistischen Tests zur Behandlung des Testproblems $H : \vartheta \in \Theta^*,\ K : \vartheta \in \Theta \setminus \Theta^*$ naheliegend. Das Ergebnis des Tests kann zur Entscheidungsfindung, ob das Experiment E_1 oder das Experiment E_2 durchgeführt werden soll, beitragen. Im fünften Kapitel geht es um die Ermittlung derartiger Teilmengen Θ^* sowie die Bereitstellung und Diskussion der zugehörigen statistischen Tests.

Für die Ausgangsfrage der Bevorzugung einer Stichprobenerhebungsart wird durch die (zumindest in speziellen Fällen mögliche) Identifizierung von solchen Teilmengen Θ^* auf der Basis der Kreuzkovarianzmatrix von X und Y eine eingängige Antwort gegeben (Kapitel 6). Effizienzbetrachtungen bei Punktschätzverfahren bzw. Bereichsschätzverfahren sind Gegenstand des nachfolgenden Abschnitts. Zusammenhänge mit Effizienzen bei statistischen Tests werden aufgezeigt. Ergebnisse, die über die bis dahin erhaltenen Resultate hinausgehen, sind unter der Annahme, dass X und Y d-dimensionale Zufallsvektoren mit einer gemeinsamen $2d$-dimensionalen Normalverteilung sind, möglich und werden im achten Kapitel erarbeitet. Im neunten Kapitel finden die vorher erarbeiteten Resultate im Zusammenhang mit $d \times d$-Kontingenztafeln Anwendung. Schließlich findet im zehnten Kapitel ein eigenständiger Vergleich der Stichprobenerhebungsarten in einem Lagealternativmodell statt.

Aus der Literatur sind nur wenige theoretisch untermauerte Aussagen zum Vergleich der Wirksamkeit von statistischen Tests bei verbundener bzw. unabhängiger Stichprobenerhebung bekannt. Knüsel ([20]) gelingt auf der Basis der Pitman-Effizienz bei 2×2-Kontingenztafeln eine solche Aussage für den einseitigen Symmetrietest von McNemar (verbundene Stichprobenerhebung) und den einseitigen exakten Zweistichprobentest von Fisher (unabhängige Stichprobenerhebung). Frischmuth ([7]) nimmt den Vergleich dieser Tests sowie den Vergleich des einseitigen Einstichproben-t-Tests (verbundene Stichprobenerhebung) mit dem einseitigen Zweistichproben-t-Test (unabhängige Stichprobenerhebung) anhand der Hodges-Lehmann-Effizienz vor. Eigenschaften der Gütefunktionen bei endlichem Stichprobenumfang erarbeitend, kommen Pollak und Cohen ([31]) zu einer vergleichenden Bewertung dieser t-Tests.

1.2 Mathematischer Rahmen

Wir gehen aus von einem zugrunde liegenden statistischen Raum $(\Omega, \mathfrak{A}, \mathfrak{P})$. Die Familie $\mathfrak{P}$ von Wahrscheinlichkeitsmaßen auf dem Messraum $(\Omega, \mathfrak{A})$ denken wir uns indiziert durch einen Parameter ϑ aus einer nicht leeren Parametermenge Θ,

$$\mathfrak{P} = \{P_\vartheta; \vartheta \in \Theta\}.$$

Es seien

$$\emptyset \neq \Theta_0 \subsetneq \Theta$$

und

$$\Theta_1 := \Theta \setminus \Theta_0$$

Teilmengen von Θ mit der Eigenschaft, dass die zugehörigen Teilmengen

$$\mathfrak{P}_0 := \{P_\vartheta; \vartheta \in \Theta_0\} \text{ und } \mathfrak{P}_1 := \{P_\vartheta; \vartheta \in \Theta_1\}$$

von $\mathfrak{P}$ disjunkt sind. Die Frage, ob das zugrunde liegende Wahrscheinlichkeitsmaß P_ϑ Element von $\mathfrak{P}_0$ oder Element von $\mathfrak{P}_1$ ist, lässt sich so als Testproblem

$$\mathrm{H} : \vartheta \in \Theta_0, \ \mathrm{K} : \vartheta \in \Theta_1 \tag{1.2.1}$$

schreiben. Zur Behandlung dieses Testproblems betrachten wir ausschließlich nichtrandomisierte statistische Tests, die über reelle Testgrößen $T : (\Omega, \mathfrak{A}) \to (\mathbb{R}, \mathfrak{B})$ und einen kritischen Wert $c \in \mathbb{R}$ verbal durch

„Verwirf die Hypothese H, falls $T \geq c$“

oder in der üblichen Weise als Indikatorvariablen $\mathrm{I}(T \geq c)$ formulierbar sind. Der Abhängigkeit vom kritischen Wert c bei einem solchen Test bewusst, sprechen wir gleichwohl kurz auch vom statistischen Test T.

Unsere anzustellenden asymptotischen Betrachtungen gründen auf Folgen $T = (T_n)_{n\in\mathbb{N}}$ und $T' = (T'_n)_{n\in\mathbb{N}}$ von statistischen Tests T_n und T'_n, $n \in \mathbb{N}$, für das Testproblem (1.2.1). Die in der Regel getroffene Grundannahme ist, dass gegeben sind

- zwei Messräume $(R, \mathfrak{S})$ und $(R', \mathfrak{S}')$,
- Zufallsvariablen $\zeta_i : (\Omega, \mathfrak{A}) \to (R, \mathfrak{S}),\ i \in \mathbb{N}$, die unter jedem $P_\vartheta \in \mathfrak{P}$ unabhängig und identisch verteilt sind, und Zufallsvariablen $\zeta'_i : (\Omega, \mathfrak{A}) \to (R', \mathfrak{S}'),\ i \in \mathbb{N}$, die unter jedem $P_\vartheta \in \mathfrak{P}$ unabhängig und identisch verteilt sind,
- reellwertige Statistiken

  $$t_n : (R^n, \mathfrak{S}^n) \to (\mathbb{R}, \mathfrak{B}) \text{ und } t'_n : \big((R')^n, (\mathfrak{S}')^n\big) \to (\mathbb{R}, \mathfrak{B}),\ n \in \mathbb{N},$$

  mit $(R^n, \mathfrak{S}^n)$ bzw. $((R')^n, (\mathfrak{S}')^n)$ als n-facher Produktraum von $(R, \mathfrak{S})$ bzw. $(R', \mathfrak{S}')$, und dass

$$T_n = t_n(\zeta_1, \dots, \zeta_n) \text{ und } T'_n = t'_n(\zeta'_1, \dots, \zeta'_n) \text{ für jedes } n \in \mathbb{N}$$

ist. In der Literatur behandelte Vergleiche von Testfolgen $T = (T_n)_{n\in\mathbb{N}}$ und $T' = (T'_n)_{n\in\mathbb{N}}$ betreffen üblicherweise den Spezialfall, dass $(R, \mathfrak{S}) = (R', \mathfrak{S}')$ und $\zeta_i = \zeta'_i$ für alle $i \in \mathbb{N}$ ist, also auch

$$T_n = t_n(\zeta_1, \dots, \zeta_n), \quad T'_n = t'_n(\zeta_1, \dots, \zeta_n) \text{ für jedes } n \in \mathbb{N}$$

gilt. Für jedes $n \in \mathbb{N}$ fußen hier die Tests T_n und T'_n auf den Beobachtungen derselben Zufallsvariablen $\zeta_1, \dots, \zeta_n$, also der Beobachtung von

$$\zeta^{(n)} := (\zeta_1, \dots, \zeta_n) \; n \in \mathbb{N};$$

es basieren T und T' auf derselben Folge

$$\Big(\big(R^n, \mathfrak{S}^n, \{P_\vartheta^{\zeta^{(n)}}; \vartheta \in \Theta\}\big), n \geq 1\Big)$$

von statistischen Experimenten. Wir betrachten dagegen den allgemeineren Fall, dass mit

$$\zeta'^{(n)} := (\zeta'_1, \dots, \zeta'_n), \; n \in \mathbb{N},$$

die Folgen von Tests T und T' auf den (im Allgemeinen verschiedenen) Folgen von Experimenten

$$\Big(\big(R^n, \mathfrak{S}^n, \{P_\vartheta^{\zeta^{(n)}}; \vartheta \in \Theta\}\big), n \geq 1\Big) \text{ und } \Big(\big((R')^n, (\mathfrak{S}')^n, \{P_\vartheta^{\zeta'^{(n)}}; \vartheta \in \Theta\}\big), n \geq 1\Big)$$

gründen. Die Behandelbarkeit des Testproblems (1.2.1) bei diesen Folgen von Experimenten setzt voraus, dass für jedes $n \in \mathbb{N}$ die Mengen

$$\mathfrak{P}_0^{\zeta^{(n)}} := \{P_\vartheta^{\zeta^{(n)}}; \vartheta \in \Theta_0\} \text{ und } \mathfrak{P}_1^{\zeta^{(n)}} := \{P_\vartheta^{\zeta^{(n)}}; \vartheta \in \Theta_1\}$$

bzw. die Mengen

$$\mathfrak{P}_0^{\zeta'^{(n)}} := \{P_\vartheta^{\zeta'^{(n)}}; \vartheta \in \Theta_0\} \text{ und } \mathfrak{P}_1^{\zeta'^{(n)}} := \{P_\vartheta^{\zeta'^{(n)}}; \vartheta \in \Theta_1\}$$

disjunkt sind. Wir werden im Verlauf der Arbeit Strukturannahmen an $\mathfrak{P}$ und Θ formulieren, die dieser Minimalforderung Rechnung tragen.

Bezugnehmend auf die Ausführungen in Abschnitt (1.1) stellen wir uns

$$(R, \mathfrak{S}) = (R', \mathfrak{S}') = (S \times S, \mathfrak{T} \otimes \mathfrak{T})$$

als Produktraum eines Messraums $(S, \mathfrak{T})$ und eine Folge von unter jedem $P_\vartheta \in \mathfrak{P}$ unabhängigen und identisch verteilten Zufallsvariablen $(X_i, Y_i) : (\Omega, \mathfrak{A}) \to (S \times S, \mathfrak{T} \otimes \mathfrak{T})$, $i \in \mathbb{N}$, vor. Unter der Voraussetzung, dass für jedes $n \in \mathbb{N}$ Beobachtungen entweder nur von

$(X_1,,Y_1),\dots,(X_n,Y_n)$ oder nur von $(X_1,Y_{n+1}),\dots,(X_n,Y_{2n})$ möglich sind, gehen wir aus von einer Folge von Zufallsvariablen $\zeta_i,\ i\in\mathbb{N}$, die unter jedem $P_\vartheta\in\Theta$ unabhängig sind und je dieselbe Verteilung

$$P_\vartheta^{\zeta_i}=P_\vartheta^{(X_1,Y_1)}$$

haben sowie von einer Folge von Zufallsvariablen $\zeta_i',\ i\in\mathbb{N}$, die unter jedem $P_\vartheta\in\Theta$ unabhängig sind und je dieselbe Verteilung

$$P_\vartheta^{\zeta_i'}=P_\vartheta^{X_1}\otimes P_\vartheta^{Y_1}$$

haben.

Der hier formulierte mathematische Rahmen ist auch für Folgen von Punktschätzern oder Folgen von Konfidenzbereichsschätzern für (Funktionen von) ϑ passend. Genauere Ausführungen hierzu werden wir im Zusammenhang mit Aussagen zur Effizienz von solchen Folgen geben.

Die Beantwortung der interessierenden Frage, welche Art der Stichprobenerhebung vorzuziehen ist, geschieht auf der Basis von Effizienzkriterien für Folgen von statistischen Tests, Punktschätzern oder Konfidenzbereichsschätzern.

2 Relative Effizienzen

2.1 Definition

Es sei Q ein (zunächst nicht näher formuliertes) Qualitätskriterium für Tests für das Testproblem (1.2.1). Für eine Folge von Tests $T = (T_n)_{n\in\mathbb{N}}$ sei für jedes $n \in \mathbb{N}$ bekannt, ob T_n dieses Qualitätskriterium erfüllt oder nicht. Dann sei

$$N_T := \inf\{n \in \mathbb{N}; T_m \text{ erfüllt } Q \;\forall\; m \geq n\},$$

mit $\inf \emptyset := \infty$. Die relative Effizienz einer Folge von Tests $T' = (T'_n)_{n\in\mathbb{N}}$ bezüglich einer Folge von Tests $T = (T_n)_{n\in\mathbb{N}}$ für das Testproblem (1.2.1) ist im Fall $N_T < \infty$ und $N_{T'} < \infty$ definiert als der Quotient

$$e_{T',T} := N_T/N_{T'}.$$

Mittels $e_{T',T}$ ist ein Qualitätsvergleich der Testfolgen T und T' möglich. Allerdings ist $e_{T',T}$ in der Regel schwer zu bestimmen.

2.2 Hodges-Lehmann-Effizienz

Für ein vorgegebenes Testniveau $\alpha \in (0,1)$ sei eine Folge von kritischen Werten $c_n := c_n(\alpha) \in \mathbb{R}$ so gewählt, dass

$$\sup\{P_{\vartheta_0}(T_n \geq c_n); \vartheta_0 \in \Theta_0\} \leq \alpha \;\forall\; n \in \mathbb{N}$$

erfüllt ist, d.h. es liege eine Folge von Tests zum Testniveau α vor. Es sei $\vartheta \in \Theta_1$ fest gewählt. Dann ist

$$\beta_n := \beta_n(\alpha, \vartheta) := P_\vartheta(T_n < c_n)$$

die Wahrscheinlichkeit für den zugehörigen Fehler 2. Art. Weiter sei für ein vorgegebenes $\beta \in (0,1)$

$$N_T := N_T(\alpha, \beta, \vartheta) := \inf\{n \in \mathbb{N}; \beta_m \leq \beta \;\forall\; m \geq n\}.$$

Dann heißt der Grenzwert

$$e^{\mathrm{HL}}_{T',T} := e^{\mathrm{HL}}_{T',T}(\alpha, \vartheta) := \lim_{\beta \downarrow 0} \frac{N_T(\alpha, \beta, \vartheta)}{N_{T'}(\alpha, \beta, \vartheta)}$$

asymptotische relative Effizienz im Hodges-Lehmann Sinn der Testgrößen T und T', falls $N_T < \infty$ und $N_{T'} < \infty$ gilt, und dieser Grenzwert existiert. In klassischen Anwendungsfällen hängt diese Größe nicht mehr vom vorgegebenen Testniveau α ab.

Falls ein $d_T(\vartheta)$ mit $0 < d_T(\vartheta) < \infty$ existiert, sodass

$$-d_T(\vartheta) = \lim_{n\to\infty} \frac{1}{n} \log \beta_n(\alpha, \vartheta) \tag{2.2.1}$$

gilt, so heißt $d_T := d_T(\vartheta)$ Hodges-Lehmann-Index von T. Der Index d_T ist etwa die Rate der exponentiellen Konvergenzgeschwindigkeit der Wahrscheinlichkeit des Fehlers 2. Art bei wachsendem Stichprobenumfang.

Existiert der Hodges-Lehmann-Index, so gilt wegen (2.2.1) auch

$$\beta_n > 0 \text{ für genügend großes } n \in \mathbb{N} \text{ und } \lim_{n\to\infty} \beta_n(\alpha, \vartheta) = 0. \tag{2.2.2}$$

Daher gilt offensichtlich

$$N_T(\alpha, \beta, \vartheta) < \infty \text{ für jedes } \beta \in (0, 1).$$

Darüber hinaus gilt in diesem Fall

$$\lim_{\beta\downarrow 0} N_T(\alpha, \beta, \vartheta) = \infty,$$

was man wie folgt sieht: Wäre dies nicht der Fall, so gäbe es ein $\tilde{n} \in \mathbb{N}$ und eine Folge von Zahlen $\gamma_k \in (0, 1)$, $k \in \mathbb{N}$, mit $\lim_{k\to\infty} \gamma_k = 0$, sodass

$$\forall\ k \in \mathbb{N}:\ N_T(\alpha, \gamma_k, \vartheta) \leq \tilde{n}.$$

Dann wäre für alle $n \geq \tilde{n}$ die Ungleichung $\beta_n(\alpha, \vartheta) \leq \gamma_k$ für jedes $k \in \mathbb{N}$ erfüllt und folglich $\beta_n = 0$ für alle $n \geq \tilde{n}$, was im Widerspruch zu (2.2.2) steht.

Weiter gilt nach Definition von N_T

$$\beta_{N_T} \leq \beta \leq \beta_{N_T - 1},$$

mit $\beta_0 := 1$. Angenommen der Hodges-Lehmann-Index von T existiert. Dann folgt

$$\underbrace{\frac{1}{N_T} \log \beta_{N_T}(\alpha, \vartheta)}_{\to -d_T(\vartheta)} \leq \frac{1}{N_T} \log \beta \leq \frac{1}{N_T} \log \beta_{N_T - 1}(\alpha, \vartheta) \sim \underbrace{\frac{1}{N_T - 1} \log \beta_{N_T - 1}(\alpha, \vartheta)}_{\to -d_T(\vartheta)} \text{ für } \beta \downarrow 0,$$

und somit

$$\lim_{\beta\downarrow 0} \frac{1}{N_T} \log \beta = -d_T(\vartheta), \tag{2.2.3}$$

falls der Hodges-Lehmann-Index für T existiert.

Deshalb gilt wegen (2.2.3) für zwei Folgen von Testgrößen T und T', deren Hodges-Lehmann-Indizes existieren,

$$d_{T'}/d_T = e^{\mathrm{HL}}_{T',T}.$$

Die Ermittlung von Hodges-Lehmann-Effzienzen kommt demnach auf die Bestimmung der exponentiellen Konvergenzraten von Wahrscheinlichkeiten für den Fehler 2. Art unter festen Alternativen $\vartheta \in \Theta_1$ hinaus.

Gilt $T_n = t_n(\zeta_1, ..., \zeta_n)$, so lässt sich für den Hodges-Lehmann-Index der Folge von Testgrößen T mit Hilfe der Verteilung von ζ_1 eine untere Schranke angeben. Es gilt für jedes $\vartheta \in \Theta_1$

$$\liminf_{n\to\infty} \frac{1}{n} \log \beta_n(\alpha, \vartheta) \geq -K(\Theta_0, \vartheta).$$

Hierbei sei für $A \subset \Theta$ $K(A, \vartheta) := \inf\{K(\tilde{\vartheta}, \vartheta), \tilde{\vartheta} \in A\}$, mit $K(\vartheta, \vartheta')$ als Kullback-Leibler-Information, d.h.

$$K(\vartheta, \vartheta') := \int \log \frac{\mathrm{d}P_{\vartheta}^{\zeta_1}}{\mathrm{d}P_{\vartheta'}^{\zeta_1}} \mathrm{d}P_{\vartheta}^{\zeta_1},$$

falls $P_{\vartheta}^{\zeta_1} \ll P_{\vartheta'}^{\zeta_1}$, und $K(\vartheta, \vartheta') := \infty$ sonst. Zum Beweis dieser Schranke siehe Theorem 2.1 aus [22].

Lässt sich $\limsup_{n\to\infty} \log \beta_n(\alpha, \vartheta)/n \leq -K(\Theta_0, \vartheta)$ zeigen, so folgt bei existierendem Hodges-Lehmann-Index die Gleichheit $d_T(\vartheta) = K(\Theta_0, \vartheta)$. Man spricht dann von der Hodges-Lehmann-Optimalität der Folge von Testgrößen T.

Es sei nun Θ ein metrischer Raum und[1] $\partial\Theta_0 \subset \Theta_0$. Weiter sei $\vartheta_0 \in \partial\Theta_0$ sowie $(\vartheta_i)_{i\in\mathbb{N}}$ eine Folge mit $\vartheta_i \in \Theta_1$ für alle $i \in \mathbb{N}$ und $\lim_{i\to\infty} \vartheta_i = \vartheta_0$. Falls der Grenzwert

$$e_{T',T}^{\mathrm{LHL}} := e_{T',T}^{\mathrm{LHL}}\big(\alpha, \vartheta_0, (\vartheta_i)_{i\in\mathbb{N}}\big) := \lim_{i\to\infty} e_{T',T}^{\mathrm{HL}}(\vartheta_i)$$

existiert, so heißt dieser lokale asymptotische relative Effizienz im Hodges-Lehmann Sinn der Testgrößen T und T'.

2.2.1 Bemerkung: In der Arbeit von Groeneboom und Oosterhoff ([10]) sowie bei Nikitin ([29]) wird eine andere Definition der Hodges-Lehmann-Effizienz gegeben. Dabei wird für ein $\vartheta \in \Theta_1$ und für ein vorgegebenes $\beta \in (0, 1)$ eine Folge von kritischen Werten $c_n = c_n(\vartheta, \beta) \in \mathbb{R}$ so gewählt, dass

$$P_\vartheta(T_n < c_n) \leq \beta \leq P_\vartheta(T_n \leq c_n) \ \forall\ n \in \mathbb{N}$$

gilt. Dann wird für ein $0 < \alpha < 1 - \beta$ mit

$$N_T = \min\{n \in \mathbb{N}; \sup\{P_{\vartheta_0}(T_n \geq c_n); \vartheta_0 \in \Theta_0\} \leq \alpha \ \forall\ m \geq n\}$$

die Hodges-Lehmann-Effizienz zweier Folgen von Tests T und T' als der Grenzwert $\lim_{\beta\downarrow 0} N_T/N_{T'}$ definiert.

[1] Man beachte, dass im Fall $\Theta \subset \mathbb{R}^d$ für ein $d \in \mathbb{N}$ $\partial\Theta_0$ der Rand von Θ_0 bezüglich der Relativtopologie auf Θ ist.

2.3 Bahadur-Effizienz

Für ein vorgegebenes β und ein fest vorgegebenes $\vartheta \in \Theta_1$ sei eine Folge $c_n := c_n(\beta, \vartheta) \in \mathbb{R}$ so gewählt, dass

$$P_\vartheta(T_n < c_n) \leq \beta \; \forall \; n \in \mathbb{N}$$

erfüllt ist. Es sei

$$\alpha_n := \alpha_n(\beta, \vartheta) := \sup\{P_{\vartheta_0}(T_n \geq c_n); \vartheta_0 \in \Theta_0\}.$$

Weiter sei für ein Testniveau $\alpha \in (0, 1 - \beta)$

$$N_T := N_T(\alpha, \beta, \vartheta) := \inf\{n \in \mathbb{N}; \alpha_m \leq \alpha \; \forall \; m \geq n\}.$$

Für die zwei Folgen von Tests T und T' heißt der Grenzwert

$$e^{\mathrm{B}}_{T',T} := e^{\mathrm{B}}_{T',T}(\beta, \vartheta) := \lim_{\alpha \downarrow 0} \frac{N_T(\alpha, \beta, \vartheta)}{N_{T'}(\alpha, \beta, \vartheta)}$$

asymptotische relative Effizienz im Bahadur Sinn der Testgrößen T und T', falls $N_T < \infty$ und $N_{T'} < \infty$ gilt, und dieser Grenzwert existiert. In klassischen Anwendungsfällen hängt diese Effizienz nicht mehr von β ab. Zu diesem Konzept der Bahadur-Effizienz siehe z.B. Nikitin ([29]) oder Groeneboom und Oosterhoff ([10]).

Falls ein $c_T(\vartheta)$ mit $0 < c_T(\vartheta) < \infty$ existiert, sodass

$$-c_T(\vartheta) = \lim_{n \to \infty} \frac{1}{n} \log \alpha_n(\beta, \vartheta) \tag{2.3.1}$$

gilt, so heißt $c_T := c_T(\vartheta)$ Bahadur-Steigung von T.

Existiert die Bahadur-Steigung, so gilt wegen (2.3.1)

$$\alpha_n > 0 \text{ für genügend großes } n \in \mathbb{N} \text{ und } \lim_{n \to \infty} \alpha_n(\beta, \vartheta) = 0. \tag{2.3.2}$$

Daher gilt offensichtlich auch

$$N_T(\alpha, \beta, \vartheta) < \infty \text{ für jedes } \alpha \in (0, 1).$$

Darüber hinaus gilt in diesem Fall

$$\lim_{\alpha \downarrow 0} N_T(\alpha, \beta, \vartheta) = \infty,$$

was man wie folgt sieht: Wäre dies nicht der Fall, so gäbe es ein $\tilde{n} \in \mathbb{N}$ und eine Folge von Zahlen $\gamma_k \in (0, 1)$, $k \in \mathbb{N}$, mit $\lim_{k \to \infty} \gamma_k = 0$, sodass

$$\forall \; k \in \mathbb{N} : \; N_V(\gamma_k, \beta, \vartheta) \leq \tilde{n}.$$

Dann wäre für alle $n \geq \tilde{n}$ die Ungleichung $\alpha_n(\beta, \vartheta) \leq \gamma_k$ für jedes $k \in \mathbb{N}$ erfüllt und folglich $\alpha_n = 0$ für alle $n \geq \tilde{n}$, was im Widerspruch zu (2.3.2) steht.

Weiter gilt nach Definition von N_T

$$\alpha_{N_T} \leq \alpha \leq \alpha_{N_T - 1},$$

mit $\alpha_0 := 1$. Angenommen, die Bahadur-Steigung von T existiert. Dann folgt

$$\underbrace{\frac{1}{N_T} \log \alpha_{N_T}(\beta, \vartheta)}_{\to -c_T(\vartheta)} \leq \frac{1}{N_T} \log \alpha \leq \frac{1}{N_T} \log \alpha_{N_T - 1}(\beta, \vartheta) \sim \underbrace{\frac{1}{N_T - 1} \log \alpha_{N_T - 1}(\beta, \vartheta)}_{\to -c_T(\vartheta)} \text{ für } \alpha \downarrow 0,$$

und somit

$$\lim_{\alpha \downarrow 0} \frac{1}{N_T} \log \alpha = -c_T(\vartheta), \tag{2.3.3}$$

falls die Bahadur-Steigung für T existiert.

Deshalb gilt wegen (2.3.3) für zwei Folgen von Testgrößen T und T', deren Bahadur-Steigungen existieren,

$$c_{T'}/c_T = e^{\mathrm{B}}_{T',T}.$$

Gilt $T_n = t_n(\zeta_1, ..., \zeta_n)$, so lässt sich für die Bahadur-Steigung der Folge von Testgrößen T mit Hilfe der Verteilung von ζ_1 und ebenfalls mit Hilfe der Kullback-Leibler-Information eine untere Schranke angeben. Es gilt für jedes $\vartheta \in \Theta_1$

$$\liminf_{n \to \infty} \frac{1}{n} \log \alpha_n(\beta, \vartheta) \geq -K(\vartheta, \Theta_0),$$

mit $K(\vartheta, A) := \inf\{K(\vartheta, \tilde{\vartheta}), \tilde{\vartheta} \in A\}$ für $A \subset \Theta$. Der Beweis für das Bestehen dieser Schranke funktioniert analog zum Beweis von Theorem 2.1 aus [22].

Gilt darüber hinaus $\limsup_{n \to \infty} \log \alpha_n(\beta, \vartheta)/n \leq -K(\vartheta, \Theta_0)$, d.h. $c_T(\vartheta) = K(\vartheta, \Theta_0)$, so heißt die Folge von Testgrößen T optimal im Bahadur Sinn.

Für eine Folge von Tests T sei

$$G_n(t) := \sup\{P_{\vartheta_0}(T_n \geq t); \vartheta_0 \in \Theta_0\},\ t \in \mathbb{R}.$$

Dann ist

$$L_n := G_n(T_n)$$

der P-Wert der Testgröße T_n. Hierbei handelt es sich um eine Zufallsvariable. Gilt nun für $\vartheta \in \Theta_1$

$$\lim_{n\to\infty} \frac{1}{n} \log L_n = -c_T(\vartheta) \text{ in } P_\vartheta\text{-Wahrscheinlichkeit}$$

mit $0 < c_T(\vartheta) < \infty$, so ist c_T die Bahadur-Steigung von T. Vergleiche hierzu Theorem 2.1 aus [10]. Auch hier lässt sich im Fall $T_n = t_n(\zeta_1, ..., \zeta_n)$ eine untere Schranke mit Hilfe der Kullback-Leibler-Information angeben. Es gilt

$$\liminf_{n\to\infty} \frac{1}{n} \log L_n \geq -K(\vartheta, \Theta_0) \ P_\vartheta\text{-f.s.}$$

für $\vartheta \in \Theta_1$; siehe zum Beweis z.B. Theorem 1.2.3 aus [29].

2.3.1 Bemerkung: Gilt im Fall $T_n = t_n(\zeta_1, ..., \zeta_n)$

$$\forall\ \varepsilon > 0 : \lim_{n\to\infty} P_\vartheta\Big(\frac{1}{n} \log L_n + K(\vartheta, \Theta_0) > \varepsilon\Big) = 0,$$

so ist die Folge von Testgrößen T optimal im Bahadur Sinn. Dies sieht man wie folgt: Es ist

$$\begin{aligned}
&\Big\{\liminf_{n\to\infty} \frac{1}{n} \log L_n \geq -K(\vartheta, \Theta_0)\Big\} = \Big\{\lim_{n\to\infty} \inf_{m\geq n} \frac{1}{m} \log L_m \geq -K(\vartheta, \Theta_0)\Big\} \\
=&\Big\{\forall\ \varepsilon > 0\ \exists\ k \in \mathbb{N}\ \forall\ n \geq k : \inf_{m\geq n} \frac{1}{m} \log L_m + K(\vartheta, \Theta_0) \geq -\varepsilon\Big\} \\
=&\bigcap_{\varepsilon>0} \bigcup_{k\in\mathbb{N}} \bigcap_{n\geq k} \Big\{\inf_{m\geq n} \frac{1}{m} \log L_m + K(\vartheta, \Theta_0) \geq -\varepsilon\Big\},
\end{aligned}$$

und daher

$$0 = P_\vartheta\bigg(\bigcup_{\varepsilon>0} \bigcap_{k\in\mathbb{N}} \bigcup_{n\geq k} \Big\{\inf_{m\geq n} \frac{1}{m} \log L_m + K(\vartheta, \Theta_0) < -\varepsilon\Big\}\bigg).$$

Dies impliziert

$$\begin{aligned}
\forall\ \varepsilon > 0 :\ 0 &= P_\vartheta\bigg(\bigcap_{k\in\mathbb{N}} \bigcup_{n\geq k} \Big\{\inf_{m\geq n} \frac{1}{m} \log L_m + K(\vartheta, \Theta_0) < -\varepsilon\Big\}\bigg) \\
&= \lim_{k\to\infty} P_\vartheta\bigg(\bigcup_{n\geq k} \Big\{\inf_{m\geq n} \frac{1}{m} \log L_m + K(\vartheta, \Theta_0) < -\varepsilon\Big\}\bigg) \\
&\geq \limsup_{n\to\infty} P_\vartheta\Big(\inf_{m\geq n} \frac{1}{m} \log L_m + K(\vartheta, \Theta_0) < -\varepsilon\Big),
\end{aligned}$$

also

$$\forall\ \varepsilon > 0 : \lim_{n\to\infty} P_\vartheta\Big(\inf_{m\geq n} \frac{1}{m} \log L_m + K(\vartheta, \Theta_0) < -\varepsilon\Big) = 0.$$

Insgesamt folgt dann

$$\begin{aligned}
\forall\, \varepsilon > 0 : P_\vartheta&\Big(\Big|\frac{1}{n}\log L_n + K(\vartheta,\Theta_0)\Big| > \varepsilon\Big)\\
=&P_\vartheta\Big(\frac{1}{n}\log L_n + K(\vartheta,\Theta_0) > \varepsilon\Big) + P_\vartheta\Big(\frac{1}{n}\log L_n + K(\vartheta,\Theta_0) < -\varepsilon\Big)\\
\le&P_\vartheta\Big(\frac{1}{n}\log L_n + K(\vartheta,\Theta_0) > \varepsilon\Big) + P_\vartheta\Big(\inf_{m\ge n}\frac{1}{m}\log L_m + K(\vartheta,\Theta_0) < -\varepsilon\Big)\\
&\longrightarrow 0 \text{ für } n\to\infty.
\end{aligned}$$

In manchen Fällen kann relativ leicht die Bahadur-Steigung bestimmt werden. Existiert nämlich eine Funktion $b : \Theta_1 \to \mathbb{R}$, sodass

$$\lim_{n\to\infty} T_n = b(\vartheta) \text{ in } P_\vartheta\text{-Wahrscheinlichkeit } \forall\, \vartheta \in \Theta_1$$

gilt, und existiert eine stetige Funktion $g : I \to \mathbb{R}$, mit einem offenen Intervall $I \supset \{b(\vartheta); \vartheta \in \Theta_1\}$, sodass

$$\lim_{n\to\infty}\frac{1}{n}\log G_n(t) = -g(t)\ \forall\, t \in I$$

gilt, so existiert die Bahadur-Steigung für alle $\vartheta \in \Theta_1$, es ist

$$\lim_{n\to\infty}\frac{1}{n}\log L_n = -g\big(b(\vartheta)\big) \text{ in } P_\vartheta\text{-Wahrscheinlichkeit,}$$

und daher

$$c_T = g \circ b.$$

Siehe hierzu z.B. Theorem 7.2 aus [4]. Zur Berechnung der Bahadur-Steigung, und zum Nachweis der Optimalität im Bahadur Sinn einer Folge von Testgrößen T für das oben genannte Testproblem kann folgendes Lemma nützlich sein:

2.3.2 Lemma: *Es existiere eine Funktion $b : \Theta_1 \to \mathbb{R}$, sodass*

$$\lim_{n\to\infty} T_n = b(\vartheta) \textit{ in } P_\vartheta\textit{-Wahrscheinlichkeit } \forall\, \vartheta \in \Theta_1$$

gilt, und es existiere eine linksseitig stetige Funktion $g : I \to \mathbb{R}$, mit einem offenen Intervall $I \supset \{b(\vartheta); \vartheta \in \Theta_1\}$, sodass

$$\limsup_{n\to\infty}\frac{1}{n}\log G_n(t) \le -g(t)\ \forall\, t \in I$$

gilt. Dann folgt

$$\forall\, \varepsilon > 0 : \lim_{n\to\infty} P_\vartheta\Big(\frac{1}{n}\log L_n + g\big(b(\vartheta)\big) > \varepsilon\Big) = 0\ \forall\, \vartheta \in \Theta_1.$$

Gilt im Fall $T_n = t_n(\zeta_1, ..., \zeta_n)$ zusätzlich $g \circ b = K(\cdot, \Theta_0)$, so ist T optimal im Bahadur Sinn für alle $\vartheta \in \Theta_1$.

Beweis: Es sei $\vartheta \in \Theta_1$, und $\varepsilon > 0$ beliebig. Wegen der linksseitigen Stetigkeit von g lässt sich ein $\delta > 0$ wählen, sodass

$$\big(b(\vartheta) - \delta\big) \in I \text{ und } \tilde{\varepsilon} := \varepsilon - \big|g\big(b(\vartheta) - \delta\big) - g\big(b(\vartheta)\big)\big| > 0$$

gilt. Da G_n monoton fallend ist, folgt

$$L_n \leq G_n\big(b(\vartheta) - \delta\big) \text{ auf } \{T_n > b(\vartheta) - \delta\}.$$

Ferner gilt nach Voraussetzung für $t \in I$

$$\limsup_{n \to \infty} \frac{1}{n} \log G_n(t) \leq -g(t),$$

woraus

$$\exists\, k \in \mathbb{N}\ \forall\, n \geq k : \sup_{m \geq n} \frac{1}{m} \log G_m\big(b(\vartheta) - \delta\big) + g\big(b(\vartheta) - \delta\big) \leq \tilde{\varepsilon}$$

folgt. Dies impliziert

$$\exists\, k \in \mathbb{N}\ \forall\, n \geq k : \frac{1}{n} \log G_n\big(b(\vartheta) - \delta\big) + g\big(b(\vartheta) - \delta\big) \leq \tilde{\varepsilon},$$

woraus

$$\begin{aligned}
\exists\, k \in \mathbb{N}\ \forall\, n \geq k : \varepsilon &\geq \frac{1}{n} \log G_n\big(b(\vartheta) - \delta\big) + g\big(b(\vartheta) - \delta\big) + \big|g\big(b(\vartheta) - \delta\big) - g\big(b(\vartheta)\big)\big| \\
&\geq \frac{1}{n} \log G_n\big(b(\vartheta) - \delta\big) + g\big(b(\vartheta) - \delta\big) - g\big(b(\vartheta) - \delta\big) + g\big(b(\vartheta)\big) \\
&= \frac{1}{n} \log G_n\big(b(\vartheta) - \delta\big) + g\big(b(\vartheta)\big)
\end{aligned}$$

folgt. Insgesamt folgt für $n \geq k$

$$\begin{aligned}
&P_\vartheta\Big(\frac{1}{n} \log L_n + g\big(b(\vartheta)\big) > \varepsilon\Big) \\
=&P_\vartheta\Big(\frac{1}{n} \log L_n + g\big(b(\vartheta)\big) > \varepsilon, T_n > b(\vartheta) - \delta\Big) + P_\vartheta\Big(\frac{1}{n} \log L_n + g\big(b(\vartheta)\big) > \varepsilon, T_n \leq b(\vartheta) - \delta\Big) \\
\leq&P_\vartheta\Big(\frac{1}{n} \log G_n\big(b(\vartheta) - \delta\big) + g\big(b(\vartheta)\big) > \varepsilon, T_n > b(\vartheta) - \delta\Big) \\
&+ P_\vartheta\Big(\frac{1}{n} \log L_n + g\big(b(\vartheta)\big) > \varepsilon, T_n \leq b(\vartheta) - \delta\Big) \\
=&P_\vartheta\Big(\frac{1}{n} \log L_n + g\big(b(\vartheta)\big) > \varepsilon, T_n \leq b(\vartheta) - \delta\Big) \\
\leq&P_\vartheta\big(T_n \leq b(\vartheta) - \delta\big) \longrightarrow 0 \text{ für } n \to \infty \text{ für alle } \vartheta \in \Theta_1.
\end{aligned}$$

Gilt im Fall $T_n = t_n(\zeta_1, ..., \zeta_n)$ nun zusätzlich $g \circ b = K(\cdot, \Theta_0)$, so folgt die Optimalität im Bahadur Sinn von T für alle $\vartheta \in \Theta_1$ aus Bemerkung 2.3.1. $\square$

Das folgende Lemma kann ebenfalls bei der Berechnung von Bahadur-Effizienzen nützlich sein. Es ist etwas allgemeiner als die oben erwähnte klassische Aussage in Theorem 7.2 aus [4].

2.3.3 Lemma: *Es sei $U \subset \Theta_1$ eine offene Teilmenge. Weiter sei $(B_n)_{n\in\mathbb{N}}$ eine Folge von reellen Zufallsvariablen auf dem zugrunde liegenden statistischen Raum, sodass eine Funktion $b : U \to \mathbb{R}$ existiere mit*

$$\lim_{n\to\infty} B_n = b(\vartheta) \text{ in } P_\vartheta\text{-Wahrscheinlichkeit } \forall\ \vartheta \in U.$$

gilt. Ferner existiere eine stetige Funktion $h : I \to \mathbb{R}$, mit einem offenen Intervall $I \supset \{b(\vartheta); \vartheta \in U\}$, sodass

$$\lim_{n\to\infty} \frac{1}{n} \log G_n(t) = -h(t)\ \forall\ t \in I$$

gilt. Dann folgt

$$\lim_{n\to\infty} \frac{1}{n} \log G_n(B_n) = -h\big(b(\vartheta)\big) \text{ in } P_\vartheta\text{-Wahrscheinlichkeit } \forall\ \vartheta \in U.$$

Beweis: Es sei $\vartheta \in U$ und $\varepsilon > 0$ beliebig. Wegen der Stetigkeit von h gilt für genügend kleines $\delta > 0$

$$\big[b(\vartheta) - \delta, b(\vartheta) + \delta\big] \subset I$$

sowie

$$\varepsilon_- := \varepsilon - \big|h\big(b(\vartheta) - \delta\big) - h\big(b(\vartheta)\big)\big| > 0$$

und

$$\varepsilon_+ := \varepsilon - \big|h\big(b(\vartheta) + \delta\big) - h\big(b(\vartheta)\big)\big| > 0,$$

Da G_n monoton fallend ist, gilt außerdem für jedes $\delta > 0$

$$G_n\big(b(\vartheta) + \delta\big) \leq G_n(B_n) \leq G_n\big(b(\vartheta) - \delta\big) \text{ auf } \big\{B_n \in \big[b(\vartheta) - \delta, b(\vartheta) + \delta\big]\big\}.$$

Da nach Voraussetzung

$$\lim_{n\to\infty} \frac{1}{n} \log G_n(t) = -h(t),\ t \in I,$$

erfüllt ist, gilt

$$\begin{aligned}\exists\, k \in \mathbb{N}\ \forall\, n \geq k : &\frac{1}{n}\log G_n\big(b(\vartheta)-\delta\big) + h\big(b(\vartheta)-\delta\big) \leq \varepsilon_-,\\ &\frac{1}{n}\log G_n\big(b(\vartheta)+\delta\big) + h\big(b(\vartheta)+\delta\big) \geq -\varepsilon_+.\end{aligned}$$

Dies impliziert

$$\begin{aligned}\exists\, k \in \mathbb{N}\ \forall\, n \geq k : \varepsilon &\geq \frac{1}{n}\log G_n\big(b(\vartheta)-\delta\big) + h\big(b(\vartheta)-\delta\big) + \Big|h\big(b(\vartheta)-\delta\big) - h\big(b(\vartheta)\big)\Big|\\ &\geq \frac{1}{n}\log G_n\big(b(\vartheta)-\delta\big) + h\big(b(\vartheta)-\delta\big) - h\big(b(\vartheta)-\delta\big) + h\big(b(\vartheta)\big)\\ &= \frac{1}{n}\log G_n\big(b(\vartheta)-\delta\big) + h\big(b(\vartheta)\big),\\ -\varepsilon &\leq \frac{1}{n}\log G_n\big(b(\vartheta)+\delta\big) + h\big(b(\vartheta)+\delta\big) - \Big|h\big(b(\vartheta)+\delta\big) - h\big(b(\vartheta)\big)\Big|\\ &\leq \frac{1}{n}\log G_n\big(b(\vartheta)+\delta\big) + h\big(b(\vartheta)+\delta\big) - h\big(b(\vartheta)+\delta\big) + h\big(b(\vartheta)\big)\\ &= \frac{1}{n}\log G_n\big(b(\vartheta)+\delta\big) + h\big(b(\vartheta)\big).\end{aligned}$$

Dann folgt insgesamt für $n \geq k$

$$\begin{aligned}&P_\vartheta\Big(\Big|\frac{1}{n}\log G_n(B_n) - h\big(b(\vartheta)\big)\Big| > \varepsilon\Big)\\ =&P_\vartheta\Big(\frac{1}{n}\log G_n(B_n) - h\big(b(\vartheta)\big) > \varepsilon\Big) + P_\vartheta\Big(\frac{1}{n}\log G_n(B_n) - h\big(b(\vartheta)\big) < -\varepsilon\Big)\\ =&P_\vartheta\Big(\frac{1}{n}\log G_n(B_n) - h\big(b(\vartheta)\big) > \varepsilon, B_n > b(\vartheta)-\delta\Big)\\ &+ P_\vartheta\Big(\frac{1}{n}\log G_n(B_n) - h\big(b(\vartheta)\big) > \varepsilon, B_n \leq b(\vartheta)-\delta\Big)\\ &+ P_\vartheta\Big(\frac{1}{n}\log G_n(B_n) - h\big(b(\vartheta)\big) < -\varepsilon, B_n < b(\vartheta)+\delta\Big)\\ &+ P_\vartheta\Big(\frac{1}{n}\log G_n(B_n) - h\big(b(\vartheta)\big) < -\varepsilon, B_n \geq b(\vartheta)+\delta\Big)\\ \leq&P_\vartheta\Big(\frac{1}{n}\log G_n\big(b(\vartheta)-\delta\big) - h\big(b(\vartheta)\big) > \varepsilon, B_n > b(\vartheta)-\delta\Big)\\ &+ P_\vartheta\Big(\frac{1}{n}\log G_n(B_n) - h\big(b(\vartheta)\big) > \varepsilon, B_n \leq b(\vartheta)-\delta\Big)\\ &+ P_\vartheta\Big(\frac{1}{n}\log G_n\big(b(\vartheta)+\delta\big) - h\big(b(\vartheta)\big) < -\varepsilon, B_n < b(\vartheta)+\delta\Big)\\ &+ P_\vartheta\Big(\frac{1}{n}\log G_n(B_n) - h\big(b(\vartheta)\big) < -\varepsilon, B_n \geq b(\vartheta)+\delta\Big)\end{aligned}$$

$$
\begin{aligned}
=&P_\vartheta\Big(\frac{1}{n}\log G_n(B_n) - h\big(b(\vartheta)\big) > \varepsilon, B_n \le b(\vartheta) - \delta\Big)\\
&+ P_\vartheta\Big(\frac{1}{n}\log G_n(B_n) - h\big(b(\vartheta)\big) < -\varepsilon, B_n \ge b(\vartheta) + \delta\Big)\\
\le&P_\vartheta\big(B_n \le b(\vartheta) - \delta\big) + P_\vartheta\big(B_n \ge b(\vartheta) + \delta\big) \longrightarrow 0 \text{ für } n \to \infty.
\end{aligned}
$$

Somit folgt die Behauptung. □

Es sei nun Θ ein metrischer Raum und $\partial\Theta_0 \subset \Theta_0$. Weiter sei $\vartheta_0 \in \partial\Theta_0$ sowie $(\vartheta_i)_{i\in\mathbb{N}}$ eine Folge mit $\vartheta_i \in \Theta_1$ für alle $i \in \mathbb{N}$ und $\lim_{i\to\infty} \vartheta_i = \vartheta_0$. Falls der Grenzwert

$$
e^{\mathrm{LB}}_{T',T} := e^{\mathrm{LB}}_{T',T}\big(\beta, \vartheta_0, (\vartheta_i)_{i\in\mathbb{N}}\big) := \lim_{i\to\infty} e^{\mathrm{B}}_{T',T}(\beta, \vartheta_i)
$$

existiert, so heißt dieser lokale asymptotische relative Effizienz im Bahadur Sinn der Testgrößen T und T'.

2.4 Pitman-Effizienz

Es sei im Zusammenhang mit der Pitman-Effizienz Θ stets ein metrischer Raum und stets $\partial\Theta_0 \subset \Theta_0$. Ferner sei $\vartheta_0 \in \partial\Theta_0$ fest gewählt. Für ein vorgegebenes Testniveau $\alpha \in (0,1)$ lasse sich eine Folge von kritischen Werten $c_n := c_n(\alpha, \vartheta_0) \in \mathbb{R}$ so wählen, dass

$$
P_{\vartheta_0}(T_n \ge c_n) \le \alpha \tag{2.4.1}
$$

sowie

$$
\lim_{n\to\infty} P_{\vartheta_0}(T_n \ge c_n) = \alpha \tag{2.4.2}
$$

erfüllt ist, d.h. es liege im Fall der einfachen Hypothese gegeben durch ϑ_0 eine Folge von Tests zum Testniveau α vor, die asymptotisch das Testniveau exakt einhalten. Es sei $(\vartheta_i)_{i\in\mathbb{N}}$ eine Folge mit $\vartheta_i \in \Theta_1$ für alle $i \in \mathbb{N}$ und $\lim_{i\to\infty} \vartheta_i = \vartheta_0$. Dann ist

$$
\beta_{n,i} := \beta_n(\alpha, \vartheta_i) := P_{\vartheta_i}(T_n < c_n),\ i \in \mathbb{N},
$$

der zu ϑ_i zugehörige Fehler 2. Art, $i \in \mathbb{N}$. Weiter sei für eine vorgegebene Folge $(\beta_i^T)_{i\in\mathbb{N}}$ mit $\beta_i^T \in (0,1)$ für alle $i \in \mathbb{N}$ und $\lim_{i\to\infty} \beta_i^T = \beta \in (0, 1-\alpha)$

$$
N_{T,i} := N_{T,i}\big(\alpha, \beta, \vartheta_0, (\vartheta_i)_{i\in\mathbb{N}}, (\beta_i^T)_{i\in\mathbb{N}}\big) := \inf\{n \in \mathbb{N}; \beta_{m,i} \le \beta_i^T\ \forall\ m \ge n\},\ i \in \mathbb{N}. \tag{2.4.3}
$$

Wir nehmen ab sofort an, dass im Zusammenhang mit der Pitman-Effizienz für jedes $\vartheta \in \Theta_1$ $\lim_{n\to\infty} \beta_n(\alpha, \vartheta) = 0$ gilt, d.h. es liegt eine Folge von konsistenten Tests vor. Dann gilt offensichtlich

$$
N_{T,i} < \infty \text{ für alle } i \in \mathbb{N}. \tag{2.4.4}
$$

Zudem nehmen wir ab sofort an, dass im Zusammenhang mit der Pitman-Effizienz für jedes $n \in \mathbb{N}$ die Abbildung $\vartheta \mapsto \beta_n(\alpha, \vartheta)$ stets stetig in ϑ_0 ist. Dann gilt

$$\lim_{i \to \infty} N_{T,i} = \infty, \tag{2.4.5}$$

was man wie folgt sieht: Angenommen, dies gilt nicht. Dann existiert eine beschränkte Teilfolge $(N_{T,i_k})_{k \in \mathbb{N}}$ von $(N_{T,i})_{i \in \mathbb{N}}$ und da die Folgeglieder dieser Teilfolge nur endlich viele Werte annehmen können, existiert eine weitere Teilfolge $(N_{T,i_{k_l}})_{l \in \mathbb{N}}$ von $(N_{T,i_k})_{k \in \mathbb{N}}$, die konstant ist. Damit folgt wegen der Stetigkeit der Abbildung $\vartheta \mapsto \beta_n(\alpha, \vartheta)$ in ϑ_0 und der Definition von $(N_{T,i})_{i \in \mathbb{N}}$

$$1 - \alpha > \beta = \lim_{l \to \infty} \beta^T_{i_{k_l}} \geq \lim_{l \to \infty} \beta_{N_{T,i_{k_l}}, i_{k_l}} = \lim_{l \to \infty} \beta_{N_{T,i_{k_1}}, i_{k_l}} = P_{\vartheta_0}(T_{N_{T,i_{k_1}}} < c_{N_{T,i_{k_1}}}) \geq 1 - \alpha,$$

also ein Widerspruch, und damit die Behauptung. Also gilt $\lim_{i \to \infty} N_{T,i} = \infty$.

Wir verlangen zudem, dass

$$\lim_{i \to \infty} \beta_{N_{T,i}, i} = \beta \tag{2.4.6}$$

erfüllt ist.

Erfüllen nun zwei Folgen von Tests T und T' für identische ϑ_0, β und identische Folgen $(\vartheta_i)_{i \in \mathbb{N}}$ diese Eigenschaften, so heißt der Grenzwert

$$e^{\mathrm{P}}_{T',T} := e^{\mathrm{P}}_{T',T}\big(\alpha, \beta, \vartheta_0, (\vartheta_i)_{i \in \mathbb{N}}, (\beta^T_i)_{i \in \mathbb{N}}, (\beta^{T'}_i)_{i \in \mathbb{N}}\big) := \lim_{i \to \infty} N_{T,i}/N_{T',i} \tag{2.4.7}$$

asymptotische relative Effizienz im Pitman Sinn der Testgrößen T und T', falls dieser existiert. In klassischen Anwendungsfällen hängt dieser Grenzwert nicht von α, β, $(\vartheta_i)_{i \in \mathbb{N}}$, $(\beta^T_i)_{i \in \mathbb{N}}$ und $(\beta^{T'}_i)$ ab.

Eine sehr ähnliche Definition der Pitman-Effizienz wird von Wieand ([40]) verwendet. Konkret definiert Wieand seine Pitman-Effizienz genauso, bis auf die Tatsache, dass er auf die Forderungen verzichtet, dass für jedes $n \in \mathbb{N}$ die Abbildung $\vartheta \mapsto \beta_n(\alpha, \vartheta)$ stetig in ϑ_0 ist, und dass die Konvergenz $\lim_{i \to \infty} \beta_{N_{T,i}, i} = \beta$ vorliegt, und stattdessen verlangt, dass die Pitman-Effizienz nicht mehr von $(\vartheta_i)_{i \in \mathbb{N}}$, $(\beta^T_i)_{i \in \mathbb{N}}$ und $(\beta^{T'}_i)$ abhängt. Wie allerdings die Ausführungen im Bemerkung 2.4.1 zeigen, ist es sinnvoll, diese Stetigkeits- und Konvergenzforderungen bei der Definition der Pitman-Effizienz zu stellen.

Auch van der Vaart ([38]) wählt ein ähnliches Vorgehen, um die Pitman-Effizienz zu definieren. Einen ähnlichen Ansatz bei der Definition von relativen Effizienzen, also auch bei der Definiton der Pitman-Effizienz, kann ebenfalls z.B. bei Serfling ([36]) nachgelesen werden.

2.4.1 Bemerkung: In der Literatur wird häufig eine andere Definition der Pitman-Effizienz verwendet. Dabei werden vom Stichprobenumfang $n \in \mathbb{N}$ abhängige Folgen von Alternativen $(\vartheta_n)_{n\in\mathbb{N}}$ mit $\vartheta_n \in \Theta_1$ für alle $n \in \mathbb{N}$ und $\lim_{n\to\infty} \vartheta_n = \vartheta_0$ betrachtet. Im Fall, dass dann für eine Abbildung $h : \mathbb{N} \to \mathbb{N}$

$$\lim_{n\to\infty} P_{\vartheta_n}(T_n < c_n) = \lim_{n\to\infty} P_{\vartheta_n}\big(T'_{h(n)} < c'_{h(n)}\big) \in (0, 1-\alpha), \tag{2.4.8}$$

sowie

$$\lim_{n\to\infty} \frac{n}{h(n)} \in (0, \infty) \tag{2.4.9}$$

gilt, wird die Pitman-Effizienz als der Grenzwert $\lim_{n\to\infty} n/h(n)$ definiert. Vergleiche hierzu z.B. mit Puri und Sen ([32]).

Nehmen wir an, dass der Fehler 2. Art der betrachteten Tests für alle festen Alternativen steng monoton fallend im Stichprobenumfang ist, also für alle $\vartheta \in \Theta_1$ die Abbildungen

$$n \mapsto P_\vartheta(T_n < c_n) \text{ und } n \mapsto P_\vartheta(T'_n < c'_n)$$

streng monoton fallend sind. Existiert dann die Pitman-Effizienz in dem Sinn, so wie sie von Puri und Sen gemäß (2.4.8) und (2.4.9) definiert wird, so existiert die Pitman-Effizienz auch in dem Sinn, wie sie von uns gemäß (2.4.7) definiert ist, und beide Effizienzen sind identisch. Dies sieht man wie folgt:

Gehen wir davon aus, dass die Pitman-Effizienz im Sinne von Puri und Sen existiert. Wir setzen

$$\beta_i^T := P_{\vartheta_i}(T_i < c_i),\ \beta_i^{T'} := P_{\vartheta_i}(T'_{h(i)} < c'_{h(i)}),\ i \in \mathbb{N}.$$

Dann existiert nach Annahme ein $\beta \in (0, 1-\alpha)$ mit

$$\lim_{i\to\infty} \beta_i^T = \lim_{i\to\infty} \beta_i^{T'} = \beta.$$

Zudem gilt wegen der geforderten Monotonie der Fehler 2. Art für jedes $i \geq 2$

$$P_{\vartheta_i}(T_m < c_m) \leq \beta_i^T \text{ für alle } m \geq i \text{ und } P_{\vartheta_i}(T_{i-1} < c_{i-1}) > \beta_i^T,$$

und genauso für jedes $i \in \mathbb{N}$ mit $h(i) \geq 2$

$$P_{\vartheta_i}(T'_m < c'_m) \leq \beta_i^{T'} \text{ für alle } m \geq h(i) \text{ und } P_{\vartheta_i}(T'_{h(i)-1} < c'_{h(i)-1}) > \beta_i^{T'}.$$

Dies bedeutet

$$N_{T,i} = i \text{ und } N_{T',i} = h(i).$$

Damit erhält man

$$\lim_{i\to\infty} P_{\vartheta_i}(T_{N_{T,i}} < c_{N_{T,i}}) = \lim_{i\to\infty} P_{\vartheta_i}(T_i < c_i) = \beta,$$

und genauso

$$\lim_{i\to\infty} P_{\vartheta_i}(T_{N'_{T',i}} < c'_{N_{T',i}}) = \lim_{i\to\infty} P_{\vartheta_i}(T'_{h(i)} < c'_{h(i)}) = \beta.$$

Darüber hinaus gilt

$$\lim_{i\to\infty} \frac{N_{T,i}}{N_{T',i}} = \lim_{i\to\infty} \frac{i}{h(i)},$$

d.h. die Pitman-Effizienz existiert in unserem Sinn, und die Effizienzen stimmen überein.

Nehmen wir umgekehrt an, dass die Pitman-Effizienz in unserem Sinn existiert, so erhält man unter gewissen Voraussetzungen zumindest noch, dass die in der Definition der Pitman-Effizienz im Sinn von Puri und Sen auftauchenden Konvergenzen (2.4.8) und (2.4.9) längs einer Teilfolge zutreffen, und der entsprechende Grenzwert (2.4.9) längs dieser Teilfolge mit der Pitman-Effizienz in unserem Sinn gemäß (2.4.7) übereinstimmt, wie man wie folgt sieht:

Es existiere die Pitman-Effizienz nach unserer Definition der Pitman-Effizienz. Gehen wir erneut davon aus, dass der Fehler 2. Art der betrachteten Tests für alle festen Alternativen steng monoton fallend im Stichprobenumfang ist, und dass die Abbildung

$$j : \mathbb{N} \to \mathbb{N},\ j(i) := N_{T,i},$$

streng monoton wachsend[2] ist. Definiere dann eine Folge $(\vartheta_i^*)_{i\in\mathbb{N}}$ durch

$$\begin{aligned} \vartheta_1^*, \ldots, \vartheta_{N_{T,1}}^* &:= \vartheta_1, \\ \vartheta_{N_{T,1}+1}^*, \ldots, \vartheta_{N_{T,2}}^* &:= \vartheta_2, \\ \vartheta_{N_{T,2}+1}^*, \ldots, \vartheta_{N_{T,3}}^* &:= \vartheta_3, \\ &\vdots \end{aligned}$$

Dann ist $\vartheta_i^* \in \Theta_1$, $i \in \mathbb{N}$, $\lim_{i\to\infty} \vartheta_i^* = \vartheta_0$ und $\vartheta_{N_{T,i}}^* = \vartheta_i$, $i \in \mathbb{N}$. Es folgt wegen $\lim_{i\to\infty} \beta_{N_{T,i},i} = \beta$

$$\beta = \lim_{i\to\infty} P_{\vartheta_i}(T_{N_{T,i}} < c_{N_{T,i}}) = \lim_{i\to\infty} P_{\vartheta_{N_{T,i}}^*}(T_{N_{T,i}} < c_{N_{T,i}}).$$

Betrachten wir weiter die Abbildung

$$k : \mathbb{N} \to \mathbb{N},\ k(i) := N_{T',i}.$$

[2]Die Existenz von Sequenzen von Alternativen mit dieser Eigenschaft wird in Bemerkung 4.2.7 behandelt.

Es existiert eine Abbildung $j^{-1} : j(\mathbb{N}) \to \mathbb{N}$ mit $j^{-1}(j(i)) = i$, $i \in \mathbb{N}$. Mit $h := k \circ j^{-1}$ folgt $N_{T',i} = h(N_{T,i})$. Damit gilt erneut wegen $\lim_{i\to\infty} \beta_{N_{T',i},i} = \beta$

$$\beta = \lim_{i\to\infty} P_{\vartheta_i}(T'_{N_{T',i}} < c'_{N_{T',i}}) = \lim_{i\to\infty} P_{\vartheta^*_{N_{T,i}}}(T'_{N_{T',i}} < c'_{N_{T',i}}) = \lim_{i\to\infty} P_{\vartheta^*_{N_{T,i}}}(T'_{h(N_{T,i})} < c'_{h(N_{T,i})}).$$

Insgesamt folgt

$$\lim_{i\to\infty} P_{\vartheta^*_{N_{T,i}}}(T_{N_{T,i}} < c_{N_{T,i}}) = \lim_{i\to\infty} P_{\vartheta^*_{N_{T,i}}}(T'_{h(N_{T,i})} < c'_{h(N_{T,i})}) = \beta,$$

und auch

$$\lim_{i\to\infty} \frac{N_{T,i}}{h(N_{T,i})} = \lim_{i\to\infty} \frac{N_{T,i}}{N_{T',i}} = e^{\mathrm{P}}_{T',T}.$$

Deshalb liegen die in der Definition der Pitman-Effizienz im Sinne von Puri und Sen geforderten Konvergenzen (2.4.8) und (2.4.9) zumindest längs der Teilfolge $(N_{T,i})_{i\in\mathbb{N}}$ der Folge $(n)_{n\in\mathbb{N}}$ vor.

Ein bekanntes Resultat zur Pitman-Effizienz[3] lässt sich wie folgt formulieren: Es sei $\Theta = \mathbb{R}$ und $\Theta_0 = \{\vartheta_0\}$. Es existieren Funktionen $\mu_n : \Theta \to \mathbb{R}$ und $\sigma_n : \Theta \to [0, \infty)$ sowie eine stetige und streng monoton wachsende Verteilungsfunktion G, sodass σ_n für alle $n \in \mathbb{N}$ auf einer Umgebung $U \subset \Theta$ um ϑ_0 nicht verschwindet und

$$\lim_{n\to\infty} \sup_{\vartheta\in U} \sup_{t\in\mathbb{R}} \left| P_\vartheta\Big(\frac{T_n - \mu_n(\vartheta)}{\sigma_n(\vartheta)} \le t\Big) - G(t)\right| = 0$$

gilt. Ferner sei μ_n auf U differenzierbar mit $\mu'_n(\vartheta_0) > 0$. Zudem existiere eine Konstante $k_T > 0$, sodass

$$\sigma_n(\vartheta_0) \sim k_T \frac{\mu'_n(\vartheta_0)}{\sqrt{n}} \text{ für } n \to \infty$$

gilt und für Folgen $(\vartheta_n)_{n\in N}$ mit $\vartheta_n := \vartheta_0 + O(1/\sqrt{n})$ gilt

$$\mu'_n(\vartheta_n) \sim \mu'_n(\vartheta_0), \sigma_n(\vartheta_n) \sim \sigma_n(\vartheta_0) \text{ für } n \to \infty.$$

Erfüllen die Folgen von Testgrößen T und T' diese Bedingungen, so ist

$$e^{\mathrm{P}}_{T',T} = \Big(\frac{k_{T'}}{k_T}\Big)^2.$$

Zum Beweis siehe hierzu z.B. Abschnitt 10.2.1 aus [36].

[3] Beachte hierbei Bemerkung 2.4.1.

3 Volumen-Effizienz

3.1 Definition

Über eine gegebene geeignete Schar von statistischen Tests ist die Aufstellung von Konfidenzbereichen für passende Parameter möglich. Umgekehrt liefert ein Konfidenzbereich für einen Parameter eine Schar von statistischen Tests für adäquate Testprobleme. Dieses Korrespondenzprinzip legt die Frage nach Effizienzvergleichen von statistischen Tests auf der Basis der korrespondierenden Konfidenzbereiche nahe. Darüber hinaus ist ein Wirksamkeitsvergleich konkurrierender Konfidenzbereichsschätzverfahren selbst interessant. Für den angestrebten Vergleich führen wir den Begriff der Volumen-Effzienz ein. Ausgangspunkt ist das zu Beginn der Arbeit eingeführte mathematische Modell und eine Abbildung $\delta : \Theta \to \mathbb{R}^p$. Für jedes $n \in \mathbb{N}$ und jedes $\alpha \in (0,1)$ sei eine Abbildung $b_n(\cdot, \alpha) : R^n \to \mathfrak{B}^d$ gegeben mit den Eigenschaften

$$\forall\ \vartheta \in \Theta :\ \{(z_1, \ldots, z_n) \in R^n; \delta(\vartheta) \in b_n(z_1, \ldots, z_n, \alpha)\} \in \mathfrak{S}^n \tag{3.1.1}$$

und

$$\forall\ \vartheta \in \Theta :\ P_\vartheta^{(Z_1,\ldots,Z_n)}(\{z_1, \ldots, z_n) \in R^n; \delta(\vartheta) \in b_n(z_1, \ldots, z_n, \alpha)\}) \geq 1 - \alpha. \tag{3.1.2}$$

Die Eigenschaft (3.1.1) bringen wir kürzer auch durch die Schreibweise

$$\forall\ \vartheta \in \Theta :\ P_\vartheta(\delta(\vartheta) \in b_n(Z_1, \ldots, Z_n, \alpha)) \geq 1 - \alpha$$

zum Ausdruck. Es ist $(b_n(\cdot, \alpha))_{n \in \mathbb{N}}$ eine Folge von Konfidenzbereichen zum Konfidenzniveau $1 - \alpha$ für $\delta(\vartheta)$. Für jedes $\vartheta_0 \in \Theta$ liegt für jedes $n \in \mathbb{N}$ bei gegebenem $\alpha \in (0,1)$ mit dem verbal formulierten statistischen Test

„Verwirf H, falls $\delta(\vartheta_0) \notin b_n(Z_1, \ldots, Z_n, \alpha)$“,

ein Test zum Testniveau α für das Testproblem

$$\mathrm{H}(\vartheta_0) :\ \delta(\vartheta) = \delta(\vartheta_0),\ \mathrm{K}(\vartheta_0) :\ \delta(\vartheta) \neq \delta(\vartheta_0)$$

vor. Wir schreiben $T = (T_n)_{n \in \mathbb{N}}$ für diese Folge von Scharen von Tests und $B_{T,n} := B_{T,n}(\alpha) := b_n^T(Z_1, \ldots, Z_n, \alpha) := b_n(Z_1, \ldots, Z_n, \alpha)$. Der unbekannte Parameter $\vartheta \in \Theta$ wird mit einer Wahrscheinlichkeit, die größer oder gleich $1-\alpha$ ist, von der Menge $B_{T,n}$ überdeckt. Wir gehen in der Folge davon aus, dass für jedes $n \in \mathbb{N}$ und jedes $\alpha \in (0,1)$

$$\{(s, z) \in \mathbb{R}^p \times R^n; s \in b_n(z, \alpha)\} \in \mathfrak{B}^p \otimes \mathfrak{S}^n$$

ist. Diese Annahme garantiert für jedes $n \in \mathbb{N}$ und jedes $\alpha \in (0,1)$ die $(\mathfrak{S}^n, \overline{\mathfrak{B}})$-Messbarkeit der Abbildung

$$z \longmapsto \int_{b_n^T(z,\alpha)} \mathrm{d}x,\ z \in R^n,$$

sodass das zufällige Volumen

$$\mathrm{Vol}_{T,n} := \int\limits_{B_{T,n}} \mathrm{d}x$$

eine Zufallsvariable ist. Es sei

$$\mathrm{E}_\vartheta(|\mathrm{Vol}_{T,n}|) < \infty \quad \text{für jedes } \vartheta \in \Theta.$$

Für die durch den Konfidenzbereichsschätzer $B_{T,n}$ behauptete Lokalisierung von $\delta(\vartheta)$ liegt mit $\mathrm{E}_\vartheta(\mathrm{Vol}_{T,n}), \vartheta \in \Theta$, ein Qualitätskriterium vor. Für zwei auf den Folgen $(b_n^T(\cdot,\alpha))_{n\in\mathbb{N}}$ und $(b_n^{T'}(\cdot,\alpha))_{n\in\mathbb{N}}$ von Konfidenzbereichen basierenden Folgen $T = (T_n)_{n\in\mathbb{N}}$ und $T' = (T_n')_{n\in\mathbb{N}}$ sei

$$e^{\mathrm{vol}}_{T',T,n} := e^{\mathrm{vol}}_{T',T,n}(\alpha,\vartheta) := \big(\mathrm{E}_\vartheta(\mathrm{Vol}_{T,n})/\,\mathrm{E}_\vartheta(\mathrm{Vol}_{T',n})\big)^2$$

für $\mathrm{E}_\vartheta(\mathrm{Vol}_{T',n}) \neq 0$ die Volumen-Effizienz der Testscharen T_n und T_n'. Es sei

$$C_b(\mathbb{R}) := \{f : \mathbb{R} \to \mathbb{R};\, f \text{ ist gleichmäßig stetig und beschränkt}\}.$$

Existiert eine Folge $(a_n)_{n\in\mathbb{N}} \subset (0,\infty)$ und existieren reellwertige Zufallsvariablen A und B mit $0 < \mathrm{E}_\vartheta(A), \mathrm{E}_\vartheta(B) < \infty$, sodass

$$\lim_{n\to\infty} \mathrm{E}_\vartheta\big(f(a_n\mathrm{Vol}_{T,n})\big) = \mathrm{E}_\vartheta\big(f(A)\big),\ \lim_{n\to\infty} \mathrm{E}_\vartheta\big(f(a_n\mathrm{Vol}_{T',n})\big) = \mathrm{E}_\vartheta\big(f(B)\big)\ \forall\ f \in C_b(\mathbb{R})$$

gilt, also Verteilungskonvergenz von $a_n\mathrm{Vol}_{T,n}$ gegen A und $a_n\mathrm{Vol}_{T',n}$ gegen B vorliegt, so heißt

$$e^{\mathrm{Lvol}}_{T',T} := e^{\mathrm{Lvol}}_{T',T}(\alpha,\vartheta) := \big(\mathrm{E}_\vartheta(A)/\,\mathrm{E}_\vartheta(B)\big)^2$$

asymptotische Volumen-Effizienz von T und T'. In typischen Anwendungsfällen hängt $e^{\mathrm{vol}}_{T',T,n}$ nicht von $\vartheta \in \Theta$, $e^{\mathrm{Lvol}}_{T',T}$ weder von $\vartheta \in \Theta$, noch von α ab.

3.1.1 Bemerkung: Zur Motivation der Definition der asymptotischen Volumen-Effizienz betrachte man folgende Überlegung. Gilt

$$\lim_{n\to\infty} \mathrm{E}_\vartheta(a_n\mathrm{Vol}_{T,n}) = \mathrm{E}_\vartheta(A),\ \lim_{n\to\infty} \mathrm{E}_\vartheta(a_n\mathrm{Vol}_{T',n}) = \mathrm{E}_\vartheta(B),$$

so folgt

$$\begin{aligned}\lim_{n\to\infty} e^{\mathrm{vol}}_{T',T,n} &= \lim_{n\to\infty} \big(\mathrm{E}_\vartheta(\mathrm{Vol}_{T,n})/\,\mathrm{E}_\vartheta(\mathrm{Vol}_{T',n})\big)^2 \\ &= \lim_{n\to\infty} \big(\mathrm{E}_\vartheta(a_n\mathrm{Vol}_{T,n})/\,\mathrm{E}_\vartheta(a_n\mathrm{Vol}_{T',n})\big)^2 = \big(\mathrm{E}_\vartheta(A)/\,\mathrm{E}_\vartheta(B)\big)^2.\end{aligned}$$

Es handelt sich bei der asymptotischen Volumen-Effizienz der Tests T_n und T_n' also um eine Approximation der Volumen-Effizienz der Tests T_n und T_n' für für große $n \in \mathbb{N}$, wobei bei der Approximation unter der genannten Bedingung für $n \to \infty$ der Fehler sogar verschwindet.

3.1.2 Bemerkung: Die asymptotische Volumen-Effizienz der Tests T_n und T'_n ist wohldefiniert. Um dies zu sehen, betrachte man eine weitere Folge $(\tilde{a}_n)_{n\in\mathbb{N}} \subset (0,\infty)$ und weitere reellwertige Zufallsvariablen $\tilde{A}$ und $\tilde{B}$, mit $0 < \mathrm{E}_\vartheta(\tilde{A}), \mathrm{E}_\vartheta(\tilde{B}) < \infty$, sodass

$$\lim_{n\to\infty} \mathrm{E}_\vartheta\left(f(\tilde{a}_n \mathrm{Vol}_{T,n})\right) = \mathrm{E}_\vartheta\left(f(\tilde{A})\right),\ \lim_{n\to\infty} \mathrm{E}_\vartheta\left(f(\tilde{a}_n \mathrm{Vol}_{T',n})\right) = \mathrm{E}_\vartheta\left(f(\tilde{B})\right)\ \forall\ f \in C_b(\mathbb{R})$$

gilt. Es ist

$$\lim_{n\to\infty} \mathrm{E}_\vartheta\left(f\Big(\frac{\tilde{a}_n}{a_n}\underbrace{a_n \mathrm{Vol}_{T,n}}_{\overset{\mathrm{v}}{\to} A}\Big)\right) = \lim_{n\to\infty} \mathrm{E}_\vartheta\left(f(\tilde{a}_n \mathrm{Vol}_{T,n})\right) = \mathrm{E}_\vartheta\left(f(\tilde{A})\right)\ \forall\ f \in C_b(\mathbb{R}).$$

Dann ist die Folge $(\tilde{a}_n/a_n)_{n\in\mathbb{N}}$ beschränkt. Denn nehmen wir an, dass dies nicht gilt, so existiert eine Teilfolge $(\tilde{a}_{n_k}/a_{n_k})_{k\in\mathbb{N}}$ mit $\lim_{k\to\infty} \tilde{a}_{n_k}/a_{n_k} = +\infty$. Da wegen $\mathrm{E}_\vartheta(A) > 0$ auch A nicht P_ϑ-f.s. identisch 0 ist, kann wegen der Verteilungskonvergenz von $(a_n \mathrm{Vol}_{T,n})_{n\in\mathbb{N}}$ die Folge $(\tilde{a}_{n_k}\mathrm{Vol}_{T,n_k})_{k\in\mathbb{N}}$ nicht in Verteilung gegen die reelle Zufallsvariable $\tilde{A}$ konvergieren. Also ist die Folge $(\tilde{a}_n/a_n)_{n\in\mathbb{N}}$ beschränkt, und besitzt daher eine konvergente Teilfolge, die ebenfalls mit $(\tilde{a}_{n_k}/a_{n_k})_{k\in\mathbb{N}}$ bezeichnet werden soll. Damit folgt die Verteilungskonvergenz von $\tilde{a}_{n_k}\mathrm{Vol}_{T,n_k} = \frac{\tilde{a}_{n_k}}{a_{n_k}} a_{n_k}\mathrm{Vol}_{T,n_k}$ gegen cA, mit $\lim_{k\to\infty} \tilde{a}_{n_k}/a_{n_k} = c > 0$. Ferner ist $P_\vartheta^{\tilde{A}} = P_\vartheta^{cA}$. Überdies gilt

$$\lim_{k\to\infty} \mathrm{E}_\vartheta\left(f\Big(\frac{\tilde{a}_{n_k}}{a_{n_k}} a_{n_k}\mathrm{Vol}_{T',n_k}\Big)\right) = \lim_{k\to\infty} \mathrm{E}_\vartheta\left(f(\tilde{a}_{n_k}\mathrm{Vol}_{T',n_k})\right) = \mathrm{E}_\vartheta\left(f(\tilde{B})\right)\ \forall\ f \in C_b(\mathbb{R}),$$

woraus wegen

$$\lim_{k\to\infty} \mathrm{E}_\vartheta\left(f(a_{n_k}\mathrm{Vol}_{T',n_k})\right) = \mathrm{E}_\vartheta\left(f(B)\right)\ \forall\ f \in C_b(\mathbb{R})$$

auch $P_\vartheta^{\tilde{B}} = P_\vartheta^{cB}$ folgt, und damit insgesamt

$$\left(\mathrm{E}_\vartheta(\tilde{A})/\,\mathrm{E}_\vartheta(\tilde{B})\right)^2 = \left(\mathrm{E}_\vartheta(cA)/\,\mathrm{E}_\vartheta(cB)\right)^2 = \left(\mathrm{E}_\vartheta(A)/\,\mathrm{E}_\vartheta(B)\right)^2.$$

3.2 Zusammenhang mit der Pitman-Effizienz

Nun sei speziell $d = 1$, also $\Theta = \mathbb{R}$. Wir betrachten für $\vartheta_0 \in \mathbb{R}$ das Testproblem

$$\mathrm{H}:\ \vartheta = \vartheta_0,\ \mathrm{K}:\ \vartheta \neq \vartheta_0.$$

Für jedes $\vartheta_0 \in \mathbb{R}$ lasse sich das Testproblem anhand eines Konfidenzintervalls gemäß der Vorschrift

$$„\text{Verwirf H, falls } \vartheta_0 \notin [a_n^T, b_n^T)“$$

behandeln mit $a_n^T, b_n^T : (R^n, \mathfrak{S}^n) \to (\mathbb{R}, \mathfrak{B})$, $\mathrm{E}_\vartheta(|a_n^T|), \mathrm{E}_\vartheta(|b_n^T|) < \infty$ für alle $\vartheta \in \mathbb{R}$ und alle $n \in \mathbb{N}$, und $a_n^T < b_n^T$ für alle $\vartheta \in \mathbb{R}$ und alle $n \in \mathbb{N}$. Weiter gelte $P_{\vartheta_0}(\vartheta_0 \notin [a_n^T, b_n^T)) \leq \alpha$ für alle $\vartheta_0 \in \mathbb{R}$ und alle $n \in \mathbb{N}$ sowie

$$\lim_{n\to\infty} P_{\vartheta_0}(\vartheta_0 \notin [a_n^T, b_n^T)) = \alpha \text{ für alle } \vartheta_0 \in \mathbb{R}.$$

Es liege also mit $[a_n^T, b_n^T)$ eine Folge von Konfidenzintervallen zum Konfidenzniveau $1-\alpha$ vor, die das Konfidenzniveau asymptotisch exakt einhalten.

Um uns auf die bisherigen Ausführungen beziehen zu können, ohne neue Notationen einführen zu müssen, gehen wir davon aus, dass sich der Test mit Hilfe einer Testgröße äquivalent formulieren lasse. Konkret verlangen wir, dass

$$P_\vartheta\big(T_n(\vartheta_0) \geq c_n(\vartheta_0)\big) = P_\vartheta\big(\vartheta_0 \notin [a_n^T, b_n^T)\big) \text{ für alle } \vartheta, \vartheta_0 \in \mathbb{R}$$

gelte, mit einer Folge von Testgrößen $(T_n(\vartheta_0))_{n\in\mathbb{N}}$ und einer Folge von kritischen Werten $(c_n(\vartheta_0))_{n\in\mathbb{N}}$. Dies kann z.B. durch die Setzung $T_n(\vartheta_0) = \mathrm{I}(\vartheta_0 \notin [a_n^T, b_n^T))$ und $c_n(\vartheta_0) = 1/2$ erreicht werden.

Liegen nun zwei Folgen von Testgrößen T und T' mit diesen Eigenschaften vor, so ergibt sich unter bestimmten Voraussetzungen, dass die asymtotische Volumen-Effizienz der Tests mit der Pitman-Effizienz übereinstimmt. Um ein entsprechendes Resultat zu erhalten, werden wir zwei einfache Lemmata voranstellen.

3.2.1 Lemma: *Es sei $-\infty < a < c \leq \infty$ und für jedes $n \in \mathbb{N}$ $f_n : [a,c] \to \mathbb{R} \cup \{\infty\}$, $f_n|_{[a,c)} < \infty$, wobei $\int_a^c e^{-f_n(x)}\mathrm{d}x < \infty$ gelte und $f_n|_{(a,c)}$ differenzierbar sei mit $f_n' > 0$ und f_n' monoton wachsend. Weiter sei $(b_n)_{n\in\mathbb{N}}$ eine reelle Folge mit $0 < b_n < c - a$ für jedes $n \in \mathbb{N}$. Dann gilt*

$$\int\limits_{a+b_n}^{c} e^{-f_n(x)n}\mathrm{d}x \leq \frac{e^{-f_n(a+b_n)n} - e^{-f_n(c)n}}{e^{-f_n(a)n} - e^{-f_n(a+b_n)n}} \int\limits_{a}^{a+b_n} e^{-f_n(x)n}\mathrm{d}x.$$

Beweis: Es gilt einerseits

$$\begin{aligned}\int\limits_{a+b_n}^{c} e^{-f_n(x)n}\mathrm{d}x &= \int\limits_{a+b_n}^{c} \frac{1}{f_n'(x)} f_n'(x)e^{-f_n(x)n}\mathrm{d}x \leq \sup_{x\in(a+b_n,c)} \frac{1}{f_n'(x)} \int\limits_{a+b_n}^{c} f_n'(x)e^{-f_n(x)n}\mathrm{d}x \\ &= \frac{1}{f_n'(a+b_n)} \frac{1}{n}(e^{-f_n(a+b_n)n} - e^{-f_n(c)n}).\end{aligned}$$

Andererseits gilt

$$\begin{aligned}\int\limits_{a}^{a+b_n} e^{-f_n(x)n}\mathrm{d}x &= \int\limits_{a}^{a+b_n} \frac{1}{f_n'(x)} f_n'(x)e^{-f_n(x)n}\mathrm{d}x \geq \inf_{x\in(a,a+b_n)} \frac{1}{f_n'(x)} \int\limits_{a}^{a+b_n} f_n'(x)e^{-f_n(x)n}\mathrm{d}x \\ &= \frac{1}{f_n'(a+b_n)} \frac{1}{n}(e^{-f_n(a)n} - e^{-f_n(a+b_n)n}).\end{aligned}$$

Somit folgt insgesamt

$$\frac{\int\limits_{a+b_n}^{c} e^{-f_n(x)n}\mathrm{d}x}{\int\limits_{a}^{a+b_n} e^{-f_n(x)n}\mathrm{d}x} \leq \frac{e^{-f_n(a+b_n)n} - e^{-f_n(c)n}}{e^{-f_n(a)n} - e^{-f_n(a+b_n)n}}$$

und damit die Behauptung. $\square$

3.2.2 Lemma: *Es seien $(a_n)_{n\in\mathbb{N}}$ und $(b_n)_{n\in\mathbb{N}}$ reelle Folgen mit $a_n < b_n$ für alle $n \in \mathbb{N}$ und $f_n, g_n : (a_n, b_n) \to \mathbb{R}_{>0}$ mit $0 < \int_{a_n}^{b_n} f_n(x)\mathrm{d}x < \infty$ und $0 < \int_{a_n}^{b_n} g_n(x)\mathrm{d}x < \infty$ für alle $n \in \mathbb{N}$ gegeben. Ferner gelte $\lim_{n\to\infty} \sup_{x\in(a_n,b_n)} |f_n(x)/g_n(x) - 1| = 0$. Dann gilt*

$$\lim_{n\to\infty} \frac{\int\limits_{a_n}^{b_n} f_n(x)\mathrm{d}x}{\int\limits_{a_n}^{b_n} g_n(x)\mathrm{d}x} = 1.$$

Beweis: Es gilt

$$\begin{aligned}\left|\frac{\int\limits_{a_n}^{b_n} f_n(x)\mathrm{d}x}{\int\limits_{a_n}^{b_n} g_n(x)\mathrm{d}x} - 1\right| &\leq \frac{\int\limits_{a_n}^{b_n} |f_n(x) - g_n(x)|\mathrm{d}x}{\int\limits_{a_n}^{b_n} g_n(x)\mathrm{d}x} = \frac{\int\limits_{a_n}^{b_n} g_n(x)\left|\frac{f_n(x)}{g_n(x)} - 1\right|\mathrm{d}x}{\int\limits_{a_n}^{b_n} g_n(x)\mathrm{d}x} \\ &\leq \sup_{x\in(a_n,b_n)} \left|\frac{f_n(x)}{g_n(x)} - 1\right| \longrightarrow 0 \text{ für } n \to \infty,\end{aligned}$$

also die Behauptung. $\square$

Wir definieren für festgehaltenes $\vartheta^* \in \mathbb{R}$

$$\rho^T_{n,\vartheta^*}(\vartheta) := P_{\vartheta^*}\big(T_n(\vartheta) < c_n(\vartheta)\big),\ \vartheta \in \mathbb{R},$$

und gehen davon aus, dass für die betrachteten Tests stets $\rho^T_{n,\vartheta^*}(\vartheta) > 0$ für alle $n \in \mathbb{N}$ und alle $\vartheta \in \mathbb{R}$ gilt. Weiter sei

$$g^T_{n,\vartheta^*}(\vartheta) := -\frac{1}{n}\log\left(\rho^T_{n,\vartheta^*}(\vartheta)\right),\ \vartheta \in \mathbb{R}.$$

Es sei für $\vartheta \in \mathbb{R} \setminus \{\vartheta^*\}$ $d^T_{\vartheta^*}(\vartheta)$ der Hodges-Lehmann-Index von T an der Stelle ϑ^* für das Testproblem, ob der zugrunde liegende Parameter gleich ϑ ist oder nicht, falls dieser existiert. In der Notation von Unterabschnitt 2.2 bedeutet dies

$$d^T_{\vartheta^*} = d_T(\vartheta^*).$$

Weiter setzen wir $d^T_{\vartheta^*}(\vartheta^*) := 0$. Dann gilt

$$\lim_{n\to\infty} g^T_{n,\vartheta^*}(\vartheta) = d^T_{\vartheta^*}(\vartheta) \text{ für alle } \vartheta \in \mathbb{R},$$

sofern der beteiligte Hodges-Lehmann-Index existiert.

Ferner sei die Folge $(N_{T,i})_{i\in\mathbb{N}}$ entsprechend der Definition (2.4.3) im Zusammenhang mit der Definition der Pitman-Effizienz in Unterabschnitt 2.4 für das Testproblem, ob der zugrunde liegende Parameter ϑ^* ist oder nicht, definiert, und $(\vartheta_i)_{i\in\mathbb{N}}$ die zugehörige reelle Folge aus der Definition der $(N_{T,i})_{i\in\mathbb{N}}$ mit $\lim_{i\to\infty}\vartheta_i = \vartheta^*$. Ebenso sei $\beta \in (0, 1-\alpha)$ aus der Definition der $(N_{T,i})_{i\in\mathbb{N}}$. Zusammengefasst bedeutet dies, dass

$$N_{T,i} = \inf\left\{n \in \mathbb{N}; P_{\vartheta_i}\big(T_m(\vartheta^*) < c_m(\vartheta^*)\big) \leq \beta_i \ \forall\ m \geq n\right\},\ i \in \mathbb{N},$$

mit einer Folge $(\beta^T_i)_{i\in\mathbb{N}}$ mit $\beta^T_i \in (0,1)$ für alle $i \in \mathbb{N}$ und $\lim_{i\to\infty}\beta^T_i = \beta \in (0, 1-\alpha)$, und dass $\lim_{i\to\infty} N_{T,i} = \infty$ sowie $\lim_{i\to\infty}\beta_{N_{T,i},i} = \beta$ erfüllt sein soll; siehe Unterabschnitt 2.4.

Wir definieren nun zur Formulierung der nächsten Resultate folgende Bedingungen:

(A1) Es gelte $\lim_{n\to\infty} \mathrm{E}_{\vartheta^*}(\sqrt{n}(b^T_n - a^T_n)) = c^T_{\vartheta^*}$ mit $c^T_{\vartheta^*} \in (0,\infty)$.

(A2) Es sei $\lim_{\vartheta\to+\infty}\vartheta\rho^T_{n,\vartheta^*}(\vartheta) = \lim_{\vartheta\to-\infty}\vartheta\rho^T_{n,\vartheta^*}(\vartheta) = 0$ für genügend großes $n \in \mathbb{N}$.

(A3) Für genügend großes $n \in \mathbb{N}$ existiere eine Funktion $f^T_n : \mathbb{R}_{\geq 0} \to \mathbb{R}_{\geq 0}$, sodass $g^T_{n,\vartheta^*}(\vartheta) = f^T_n(|\vartheta - \vartheta^*|)$, $\vartheta \in \mathbb{R}$, gilt, und diese Funktion sei auf $[0,\infty)$ zweimal stetig differenzierbar, wobei ihre Ableitung auf $(0,\infty)$ monoton wachsend und strikt positiv sei. Zudem gelten für die zugehörige Taylorentwicklung von g^T_{n,ϑ^*} um ϑ^*

$$g^T_{n,\vartheta^*}(\vartheta) = a^T_{0,n} + a^T_{1,n}(\vartheta - \vartheta^*) + a^T_{2,n}(\vartheta - \vartheta^*)^2 + r^T_n(\vartheta),\ \vartheta \in \mathbb{R},$$

die Eigenschaften $\liminf_{n\to\infty} a^T_{2,n} > 0$ und $\lim_{i\to\infty}\sup_{n\in\mathbb{N}} |r^T_n(\tilde{\vartheta}_i)|/(\tilde{\vartheta}_i - \vartheta^*)^2 = 0$ für alle Folgen $(\tilde{\vartheta}_i)_{i\in\mathbb{N}}$ mit $\tilde{\vartheta}_i \neq \vartheta^*$ für alle $i \in \mathbb{N}$ und $\lim_{i\to\infty}\tilde{\vartheta}_i = \vartheta^*$.

3.2.3 Bemerkung:

a) Dass die Bedingung (A1) in konkreten Beispielen erfüllt ist, zeigen die Betrachtungen in Beispiel 3.2.9 und Unterabschnitt 8.10.

b) Ist (A3) erfüllt, so gilt $\rho^T_{n,\vartheta^*}(\vartheta) = \rho^T_{n,\vartheta}(\vartheta^*)$, $\vartheta \in \mathbb{R}$. Es ist (A2) erfüllt, wenn der Fehler 2. Art bei zugrunde liegendem Parameter ϑ^* für das Testproblem, ob der zugrunde liegende Parameter gleich ϑ ist oder nicht, hinreichend schnell gegen $0 \in \mathbb{R}$ für $|\vartheta| \to \infty$ geht. Konkret muss dieser Fehler 2. Art von der Größenordnung $o(1/|\vartheta|)$ für $|\vartheta| \to \infty$ sein.

c) Ist die Größe des Fehlers 2. Art des betrachteten Testproblems, ob der zugrunde liegende Parameter gleich ϑ ist oder nicht, nur abhängig vom Abstand zwischen der Hypothese ϑ und der Alternative ϑ^*, so ist die Existenz einer Funktion f_n^T im Sinne von (A3) erfüllt. Des Weiteren ist zu den in (A3) geforderten weiteren Eigenschaften der Funktion f_n^T zu bemerken, dass gewisse Monotonien des Fehlers 2. Art im Abstand zwischen der Hypothese und Alternative in der Regel vorliegen sollten.

d) Für den in (A3) genannten Koeffizienten $a_{2,n}^T$ gilt unter den in (A3) genannten Bedingungen an die Funktion f_n^T stets $a_{2,n}^T \geq 0$ für genügend große $n \in \mathbb{N}$, da es sich bei dem Koeffizienten um die zweite Ableitung der Funktion g_{n,ϑ^*}^T an der Stelle ϑ^* handelt. Ist die Funktion $d_{\vartheta^*}^T$ an der Stelle ϑ^* zweimal differenzierbar, wobei die zweite Ableitung an der Stelle ϑ^* nicht verschwindet, und konvergiert die zweite Ableitung der Funktion g_{n,ϑ^*}^T an der Stelle ϑ^* für $n \to \infty$ mit, so ist die Bedingung $\liminf_{n\to\infty} a_{2,n}^T > 0$ aus (A3) erfüllt.

e) Zur Bedingung $\lim_{i\to\infty} \sup_{n\in\mathbb{N}} |r_n^T(\tilde{\vartheta}_i)|/(\tilde{\vartheta}_i - \vartheta^*)^2 = 0$ für alle Folgen $(\tilde{\vartheta}_i)_{i\in\mathbb{N}}$ mit $\tilde{\vartheta}_i \neq \vartheta^*$ für alle $i \in \mathbb{N}$ und $\lim_{i\to\infty} \tilde{\vartheta}_i = \vartheta^*$ in (A3) soll bemerkt werden, dass in jedem Fall für alle festen $n \in \mathbb{N}$ $\lim_{i\to\infty} |r_n^T(\tilde{\vartheta}_i)|/(\tilde{\vartheta}_i - \vartheta^*)^2 = 0$ für alle Folgen $(\tilde{\vartheta}_i)_{i\in\mathbb{N}}$ mit $\tilde{\vartheta}_i \neq \vartheta^*$ für alle $i \in \mathbb{N}$ und $\lim_{i\to\infty} \tilde{\vartheta}_i = \vartheta^*$ gilt, da es sich bei dem Term r_n^T um den Restterm in der Taylorentwicklung der Funktion g_{n,ϑ^*}^T um die Stelle ϑ^* handelt. Ist g_{n,ϑ^*}^T in einer Umgebung U um ϑ^* dreimal differenzierbar, und ist die dritte Ableitung in U gleichmäßig beschränkt in dem Sinne, dass

$$\sup_{n\in\mathbb{N}} \sup_{\vartheta\in U} |(g_{n,\vartheta^*}^T)'''(\vartheta)| < \infty$$

gilt, so liefert für genügend großes $i \in \mathbb{N}$ die Lagrange-Darstellung des Restglieds die Existenz eines $\tilde{\tilde{\vartheta}}_i \in U$ mit

$$\frac{|r_n^T(\tilde{\vartheta}_i)|}{(\tilde{\vartheta}_i - \vartheta^*)^2} = \frac{|\frac{1}{6}(g_{n,\vartheta^*}^T)'''(\tilde{\tilde{\vartheta}}_i)(\tilde{\vartheta}_i - \vartheta^*)^3|}{(\tilde{\vartheta}_i - \vartheta^*)^2} = \frac{1}{6}|(g_{n,\vartheta^*}^T)'''(\tilde{\tilde{\vartheta}}_i)||\tilde{\vartheta}_i - \vartheta^*|.$$

Daraus folgt für genügend großes $i \in \mathbb{N}$

$$\sup_{n\in\mathbb{N}} \frac{|r_n^T(\tilde{\vartheta}_i)|}{(\tilde{\vartheta}_i - \vartheta^*)^2} \leq \frac{1}{6} \sup_{n\in\mathbb{N}} \sup_{\vartheta\in U} |(g_{n,\vartheta^*}^T)'''(\vartheta)||\tilde{\vartheta}_i - \vartheta^*| \to 0 \text{ für } i \to \infty,$$

sodass die Bedingung $\lim_{i\to\infty} \sup_{n\in\mathbb{N}} |r_n^T(\tilde{\vartheta}_i)|/(\tilde{\vartheta}_i - \vartheta^*)^2 = 0$ für alle Folgen $(\tilde{\vartheta}_i)_{i\in\mathbb{N}}$ mit $\tilde{\vartheta}_i \neq \vartheta^*$ für alle $i \in \mathbb{N}$ und $\lim_{i\to\infty} \tilde{\vartheta}_i = \vartheta^*$ aus (A3) erfüllt ist.

3.2.4 Lemma: *Die Tests T und T' erfüllen (A1) und (A2). Dann gilt*

$$\lim_{n\to\infty} e_{T',T,n}^{\mathrm{vol}}(\vartheta^*) = \lim_{n\to\infty} \left(\frac{\int e^{-g_{n,\vartheta^*}^T(\vartheta)n} \mathrm{d}\vartheta}{\int e^{-g_{n,\vartheta^*}^{T'}(\vartheta)n} \mathrm{d}\vartheta} \right)^2.$$

Existiert zusätzlich die Pitman-Effizienz $e^{\mathrm{P}}_{T',T}(\vartheta^*)$ und gilt $0 < e^{\mathrm{P}}_{T',T}(\vartheta^*) < \infty$, so gilt auch

$$\lim_{n\to\infty} e^{\mathrm{vol}}_{T',T,n}(\vartheta^*) = e^{\mathrm{P}}_{T',T}(\vartheta^*) \lim_{i\to\infty} \left(\frac{\int e^{-g^T_{N_{T,i},\vartheta^*}(\vartheta) N_{T,i}} \mathrm{d}\vartheta}{\int e^{-g^{T'}_{N_{T',i},\vartheta^*}(\vartheta) N_{T',i}} \mathrm{d}\vartheta} \right)^2.$$

Beweis: Zunächst ist festzustellen, dass aufgrund der Gestalt des vorliegenden Tests

$$P_{\vartheta^*}\big(T_n(\vartheta) \geq c_n(\vartheta)\big) = P_{\vartheta^*}(\vartheta < a^T_n) + P_{\vartheta'}(\vartheta \geq b^T_n),\ \vartheta \in \mathbb{R},$$

äquivalent

$$-\rho^T_{n,\vartheta^*}(\vartheta) = P_{\vartheta^*}(b^T_n \leq \vartheta) - P_{\vartheta^*}(a^T_n \leq \vartheta),\ \vartheta \in \mathbb{R},$$

gilt, und genauso

$$-\rho^{T'}_{n,\vartheta^*}(\vartheta) = P_{\vartheta^*}(b^{T'}_n \leq \vartheta) - P_{\vartheta^*}(a^{T'}_n \leq \vartheta),\ \vartheta \in \mathbb{R}.$$

Es ist[4]

$$\begin{aligned}\lim_{n\to\infty} e^{\mathrm{vol}}_{T',T,n}(\vartheta^*) &= \left(\frac{c^T_{\vartheta^*}}{c^{T'}_{\vartheta^*}}\right)^2 = \lim_{n\to\infty}\left(\frac{\mathrm{E}_{\vartheta^*}\left(\sqrt{n}(b^T_n - a^T_n)\right)}{\mathrm{E}_{\vartheta^*}\left(\sqrt{n}(b^{T'}_n - a^{T'}_n)\right)}\right)^2 \\ &= \lim_{n\to\infty}\left(\frac{\mathrm{E}_{\vartheta^*}(b^T_n - a^T_n)}{\mathrm{E}_{\vartheta^*}(b^{T'}_n - a^{T'}_n)}\right)^2 = \lim_{n\to\infty}\left(\frac{\int \vartheta \mathrm{d}\rho^T_{n,\vartheta^*}(\vartheta)}{\int \vartheta \mathrm{d}\rho^{T'}_{n,\vartheta^*}(\vartheta)}\right)^2.\end{aligned}$$

Es folgt mit Hilfe von partieller Integration

$$\lim_{n\to\infty} \frac{\int \vartheta \mathrm{d}\rho^T_{n,\vartheta^*}(\vartheta)}{\int \vartheta \mathrm{d}\rho^{T'}_{n,\vartheta^*}(\vartheta)} = \lim_{n\to\infty} \frac{\int \rho^T_{n,\vartheta^*}(\vartheta)\mathrm{d}\vartheta}{\int \rho^{T'}_{n,\vartheta^*}(\vartheta)\mathrm{d}\vartheta} = \lim_{n\to\infty} \frac{\int e^{-g^T_{n,\vartheta^*}(\vartheta)n}\mathrm{d}\vartheta}{\int e^{-g^{T'}_{n,\vartheta^*}(\vartheta)n}\mathrm{d}\vartheta},$$

und damit insgesamt

$$\lim_{n\to\infty} e^{\mathrm{vol}}_{T',T,n}(\vartheta^*) = \lim_{n\to\infty}\left(\frac{\int e^{-g^T_{n,\vartheta^*}(\vartheta)n}\mathrm{d}\vartheta}{\int e^{-g^{T'}_{n,\vartheta^*}(\vartheta)n}\mathrm{d}\vartheta}\right)^2,$$

also die erste Aussage. Nun seien $N_{T,i}$ und $N_{T',i}$ entsprechend des Unterabschnittes 2.4 gewählt. Aufgrund der generellen Voraussetzungen in Unterabschnitt 2.1 gilt $\lim_{i\to\infty} N_{T,i} = \lim_{i\to\infty} N_{T',i} = \infty$, und somit

$$\begin{aligned}\lim_{n\to\infty} e^{\mathrm{vol}}_{T',T,n}(\vartheta^*) &= \left(\frac{c^T_{\vartheta^*}}{c^{T'}_{\vartheta^*}}\right)^2 = \lim_{i\to\infty} \frac{N_{T,i}}{N_{T',i}}\left(\frac{\mathrm{E}_{\vartheta^*}(b^T_{N_{T,i}} - a^T_{N_{T,i}})}{\mathrm{E}_{\vartheta^*}(b^{T'}_{N_{T',i}} - a^{T'}_{N_{T',i}})}\right)^2 \\ &= e^P_{T',T}(\vartheta^*) \lim_{i\to\infty}\left(\frac{\int \vartheta \mathrm{d}\rho^T_{N_{T,i},\vartheta^*}(\vartheta)}{\int \vartheta \mathrm{d}\rho^{T'}_{N_{T',i},\vartheta^*}(\vartheta)}\right)^2.\end{aligned}$$

[4] Die hier genannte Integration ist in dem Sinne $\mathrm{d}\rho^T_{n,\vartheta^*}(\cdot) := \mathrm{d}P_{\vartheta^*}(b^T_n \leq \cdot) - \mathrm{d}P_{\vartheta^*}(a^T_n \leq \cdot)$ zu verstehen.

Erneut mit Hilfe von partieller Integration folgt

$$\lim_{i\to\infty}\frac{\int \vartheta \mathrm{d}\rho^{T}_{N_{T,i},\vartheta^*}(\vartheta)}{\int \vartheta \mathrm{d}\rho^{T'}_{N_{T',i},\vartheta^*}(\vartheta)} = \lim_{i\to\infty}\frac{\int \rho^{T}_{N_{T,i},\vartheta^*}(\vartheta)\mathrm{d}\vartheta}{\int \rho^{T'}_{N_{T',i},\vartheta^*}(\vartheta)\mathrm{d}\vartheta} = \lim_{i\to\infty}\frac{\int e^{-g^{T}_{N_{T,i},\vartheta^*}(\vartheta)N_{T,i}}\mathrm{d}\vartheta}{\int e^{-g^{V}_{N_{T',i},\vartheta^*}(\vartheta)N_{T',i}}\mathrm{d}\vartheta},$$

also insgesamt

$$\lim_{n\to\infty} e^{\mathrm{vol}}_{T',T,n}(\vartheta^*) = e^{P}_{T',T}(\vartheta^*)\lim_{i\to\infty}\left(\frac{\int e^{-g^{T}_{N_{T,i},\vartheta^*}(\vartheta)N_{T,i}}\mathrm{d}\vartheta}{\int e^{-g^{V}_{N_{T',i},\vartheta^*}(\vartheta)N_{T',i}}\mathrm{d}\vartheta}\right)^2$$

und so die Behauptung. □

3.2.5 Satz: *Die Tests T und T' erfüllen (A1) - (A3). Für ein $0 < \gamma < 1 - \alpha$ existiere für jedes $0 < \beta < \gamma$ eine Folge $(\vartheta_i)_{i\in\mathbb{N}}$, sodass die Pitman-Effizienz bei zugrunde liegender Folge $(\vartheta_i)_{i\in\mathbb{N}}$ und zugrunde liegendem β im Sinne der Definition aus Unterabschnitt 2.4 existiert mit $0 < e^{\mathrm{P}}_{T',T}(\vartheta^*) < \infty$. Zudem hänge $e^{\mathrm{P}}_{T',T}(\vartheta^*)$ nicht von $(\vartheta_i)_{i\in\mathbb{N}}$, $(\beta^{T}_i)_{i\in\mathbb{N}}$, $(\beta^{T'}_i)_{i\in\mathbb{N}}$ und $\beta \in (0,\gamma)$ ab. Dann gilt*

$$\lim_{n\to\infty} e^{\mathrm{vol}}_{T',T,n}(\vartheta^*) = e^{\mathrm{P}}_{T',T}(\vartheta^*).$$

Beweis: Nach Lemma 3.2.4 genügt es,

$$\lim_{i\to\infty}\frac{\int e^{-g^{T}_{N_{T,i},\vartheta^*}(\vartheta)N_{T,i}}\mathrm{d}\vartheta}{\int e^{-g^{T'}_{N_{T',i},\vartheta^*}(\vartheta)N_{T',i}}\mathrm{d}\vartheta} = 1$$

zu zeigen. Da die Pitman-Effizienz nicht mehr von der Folge $(\vartheta_i)_{i\in\mathbb{N}}$ abhängt, sowie wegen (A3), kann ohne Einschränkung $\vartheta_i > \vartheta^*$ für alle $i \in \mathbb{N}$ angenommen werden. Da dann wegen (A3) für genügend großes $i \in \mathbb{N}$ alle Bedingungen von Lemma 3.2.1 erfüllt sind, folgt mit Hilfe dieses Lemmas für genügend großes $i \in \mathbb{N}$

$$\frac{\int e^{-g^{T}_{N_{T,i},\vartheta^*}(\vartheta)N_{T,i}}\mathrm{d}\vartheta}{\int e^{-g^{T'}_{N_{T',i},\vartheta^*}(\vartheta)N_{T',i}}\mathrm{d}\vartheta} = \frac{\int\limits_{\vartheta^*}^{\infty} e^{-g^{T}_{N_{T,i},\vartheta^*}(\vartheta)N_{T,i}}\mathrm{d}\vartheta}{\int\limits_{\vartheta^*}^{\infty} e^{-g^{T'}_{N_{T',i},\vartheta^*}(\vartheta)N_{T',i}}\mathrm{d}\vartheta} \leq \frac{\int\limits_{\vartheta^*}^{\vartheta_i} e^{-g^{T}_{N_{T,i},\vartheta^*}(\vartheta)N_{T,i}}\mathrm{d}\vartheta + \int\limits_{\vartheta_i}^{\infty} e^{-g^{T}_{N_{T,i},\vartheta^*}(\vartheta)N_{T,i}}\mathrm{d}\vartheta}{\int\limits_{\vartheta^*}^{\vartheta_i} e^{-g^{T'}_{N_{T',i},\vartheta^*}(\vartheta)N_{T',i}}\mathrm{d}\vartheta}$$

$$\leq \frac{\int\limits_{\vartheta^*}^{\vartheta_i} e^{-g^{T}_{N_{T,i},\vartheta^*}(\vartheta)N_{T,i}}\mathrm{d}\vartheta + \frac{e^{-g^{T}_{N_{T,i},\vartheta^*}(\vartheta_i)N_{T,i}}}{e^{-g^{T}_{N_{T,i},\vartheta^*}(\vartheta^*)N_{T,i}} - e^{-g^{T}_{N_{T,i},\vartheta^*}(\vartheta_i)N_{T,i}}}\int\limits_{\vartheta^*}^{\vartheta_i} e^{-g^{T}_{N_{T,i},\vartheta^*}(\vartheta)N_{T,i}}\mathrm{d}\vartheta}{\int\limits_{\vartheta^*}^{\vartheta_i} e^{-g^{T'}_{N_{T',i},\vartheta^*}(\vartheta)N_{T',i}}\mathrm{d}\vartheta}$$

$$=\left(1+\frac{\rho^T_{N_{T,i},\vartheta^*}(\vartheta_i)}{\rho^T_{N_{T,i},\vartheta^*}(\vartheta^*)-\rho^T_{N_{T,i},\vartheta^*}(\vartheta_i)}\right)\frac{\int\limits_{\vartheta^*}^{\vartheta_i} e^{-g^T_{N_{T,i},\vartheta^*}(\vartheta)N_{T,i}}\mathrm{d}\vartheta}{\int\limits_{\vartheta^*}^{\vartheta_i} e^{-g^{T'}_{N_{T',i},\vartheta^*}(\vartheta)N_{T',i}}\mathrm{d}\vartheta}.$$

Wegen (A3) und $\lim_{i\to\infty} N_{T,i}=\infty$ gilt für genügend großes $i\in\mathbb{N}$ $\rho^T_{N_{T,i},\vartheta^*}(\vartheta_i)=\rho^T_{N_{T,i},\vartheta_i}(\vartheta^*)$, und daher aufgrund von (2.4.6) $\lim_{i\to\infty}\rho^T_{N_{T,i},\vartheta^*}(\vartheta_i)=\lim_{i\to\infty}\rho^T_{N_{T,i},\vartheta_i}(\vartheta^*)=\beta$. Da Folgen von Tests zum Testniveau α vorliegen, die asymptotisch das Testniveau exakt einhalten, gilt wegen $\lim_{i\to\infty} N_{T,i}=\infty$ außerdem $\lim_{i\to\infty}\rho^T_{N_{T,i},\vartheta^*}(\vartheta^*)=1-\alpha$, womit

$$\lim_{i\to\infty}\left(1+\frac{\rho^T_{N_{T,i},\vartheta^*}(\vartheta_i)}{\rho^T_{N_{T,i},\vartheta^*}(\vartheta^*)-\rho^T_{N_{T,i},\vartheta^*}(\vartheta_i)}\right)=1+\frac{\beta}{1-\alpha-\beta}=\frac{1-\alpha}{1-\alpha-\beta}$$

folgt.

Es lässt sich auch aufgrund von (A3) für die Funktionen g^T_{n,ϑ^*} und $g^{T'}_{n,\vartheta^*}$ eine Entwicklung um ϑ^* gemäß

$$\begin{aligned}
g^T_{n,\vartheta^*}(\vartheta)&=a^T_{0,n}+a^T_{1,n}(\vartheta-\vartheta^*)+a^T_{2,n}(\vartheta-\vartheta^*)^2+r^T_n(\vartheta),\ \vartheta\in[\vartheta^*,\infty),\\
g^{T'}_{n,\vartheta^*}(\vartheta)&=a^{T'}_{0,n}+a^{T'}_{1,n}(\vartheta-\vartheta^*)+a^{T'}_{2,n}(\vartheta-\vartheta^*)^2+r^{T'}_n(\vartheta),\ \vartheta\in[\vartheta^*,\infty),
\end{aligned}$$

anbringen. Wegen

$$g^T_{n,\vartheta^*}(\vartheta^*)n=-\log\left(\rho^T_{n,\vartheta^*}(\vartheta^*)\right),\ g^{T'}_{n,\vartheta^*}(\vartheta^*)n=-\log\left(\rho^{T'}_{n,\vartheta^*}(\vartheta^*)\right)$$

und da aufgrund von (A3) $a^T_{1,n}=a^{T'}_{1,n}=0$ für genügend große $n\in\mathbb{N}$ gelten muss, vereinfachen sich diese Darstellungen zu

$$\begin{aligned}
g^T_{n,\vartheta^*}(\vartheta)&=-\frac{\log\left(\rho^T_{n,\vartheta^*}(\vartheta^*)\right)}{n}+a^T_{2,n}(\vartheta-\vartheta^*)^2+r^T_n(\vartheta),\ \vartheta\in[\vartheta^*,\infty),\\
g^{T'}_{n,\vartheta^*}(\vartheta)&=-\frac{\log\left(\rho^{T'}_{n,\vartheta^*}(\vartheta^*)\right)}{n}+a^{T'}_{2,n}(\vartheta-\vartheta^*)^2+r^{T'}_n(\vartheta),\ \vartheta\in[\vartheta^*,\infty).
\end{aligned}$$

Wegen $\lim_{i\to\infty}\rho^T_{N_{T,i},\vartheta^*}(\vartheta_i)=\beta$ und $\lim_{i\to\infty}\rho^{T'}_{N_{T',i},\vartheta^*}(\vartheta_i)=\beta$ ist nach Definition

$$\lim_{i\to\infty} g^T_{N_{T,i},\vartheta^*}(\vartheta_i)N_{T,i}=\lim_{i\to\infty} g^{T'}_{N_{T',i},\vartheta^*}(\vartheta_i)N_{T',i}=-\log\beta<\infty.$$

Daher ist

$$\limsup_{i\to\infty}(\vartheta_i-\vartheta^*)^2N_{T,i}<\infty,\ \limsup_{i\to\infty}(\vartheta_i-\vartheta^*)^2N_{T',i}<\infty.$$

Dies sieht man wie folgt ein: Wäre z.B. $\limsup_{i\to\infty}(\vartheta_i-\vartheta^*)^2N_{T,i}=\infty$, so müsste wegen

$$\liminf_{i\to\infty} a^T_{2,N_{T,i}}>0 \text{ und } \lim_{i\to\infty}\frac{|r^T_{N_{T,i}}(\vartheta_i)|}{(\vartheta_i-\vartheta^*)^2}\le\lim_{i\to\infty}\frac{\sup_{n\in\mathbb{N}}|r^T_n(\vartheta_i)|}{(\vartheta_i-\vartheta^*)^2}=0,$$

was wegen (A3) gilt, auch

$$\begin{aligned}&\limsup_{i\to\infty}\left(a^T_{2,N_{T,i}}N_{T,i}(\vartheta_i-\vartheta^*)^2+N_{T,i}r^T_{N_{T,i}}(\vartheta_i)\right)\\=&\limsup_{i\to\infty}\left(N_{T,i}(\vartheta_i-\vartheta^*)^2\left(a^T_{2,N_{T,i}}+\frac{r^T_{N_{T,i}}(\vartheta_i)}{(\vartheta_i-\vartheta^*)^2}\right)\right)=\infty\end{aligned}$$

folgen, sodass insgesamt

$$\limsup_{i\to\infty} g^T_{N_{T,i},\vartheta^*}(\vartheta_i)N_{T,i}=\infty$$

gelten müsste, was aber wegen $\lim_{i\to\infty} g^T_{N_{T,i},\vartheta^*}(\vartheta_i)N_{T,i}=-\log\beta<\infty$ nicht sein kann.

Da Folgen von Tests zum Testniveau α vorliegen, die asymptotisch das Testniveau einhalten, gilt außerdem

$$\begin{aligned}\lim_{n\to\infty} g^T_{N_{T,i},\vartheta^*}(\vartheta^*)N_{T,i}&=\lim_{n\to\infty}-\log\left(\rho^T_{n,\vartheta^*}(\vartheta^*)\right)=-\log(1-\alpha),\\ \lim_{n\to\infty} g^{T'}_{N_{T',i},\vartheta^*}(\vartheta^*)N_{T',i}&=\lim_{n\to\infty}-\log\left(\rho^{T'}_{n,\vartheta^*}(\vartheta^*)\right)=-\log(1-\alpha).\end{aligned}$$

Nun sei $(\tilde\vartheta_i)_{i\in\mathbb{N}}$ eine beliebige Folge mit $\tilde\vartheta_i\in(\vartheta^*,\vartheta_i]$ für alle $i\in\mathbb{N}$. Dann ergibt sich durch einfache Umformung

$$\begin{aligned}&|g^T_{N_{T,i},\vartheta^*}(\tilde\vartheta_i)N_{T,i}-g^{T'}_{N_{T',i},\vartheta^*}(\tilde\vartheta_i)N_{T',i}|\\
=&\Big|-\log\left(\rho^T_{n,\vartheta^*}(\vartheta^*)\right)+a^T_{2,N_{T,i}}(\tilde\vartheta_i-\vartheta^*)^2N_{T,i}+N_{T,i}r^T_{N_{T,i}}(\tilde\vartheta_i)\\
&+\log\left(\rho^{T'}_{n,\vartheta^*}(\vartheta^*)\right)-a^{T'}_{2,N_{T',i}}(\tilde\vartheta_i-\vartheta^*)^2N_{T',i}-N_{T',i}r^{T'}_{N_{T',i}}(\tilde\vartheta_i)\Big|\\
=&\Bigg|\frac{(\tilde\vartheta_i-\vartheta^*)^2}{(\vartheta_i-\vartheta^*)^2}\Big(-\log\left(\rho^T_{n,\vartheta^*}(\vartheta^*)\right)+a^T_{2,N_{T,i}}(\vartheta_i-\vartheta^*)^2N_{T,i}+N_{T,i}r^T_{N_{T,i}}(\vartheta_i)\\
&+\log\left(\rho^{T'}_{n,\vartheta^*}(\vartheta^*)\right)-a^{T'}_{2,N_{T',i}}(\vartheta_i-\vartheta^*)^2N_{T',i}-N_{T',i}r^{T'}_{N_{T',i}}(\vartheta_i)\Big)\\
&+\frac{(\tilde\vartheta_i-\vartheta^*)^2}{(\vartheta_i-\vartheta^*)^2}\Big(\log\left(\rho^T_{n,\vartheta^*}(\vartheta^*)\right)-\log\left(\rho^{T'}_{n,\vartheta^*}(\vartheta^*)\right)\Big)-\log\left(\rho^T_{n,\vartheta^*}(\vartheta^*)\right)+\log\left(\rho^{T'}_{n,\vartheta^*}(\vartheta^*)\right)\\
&+\frac{(\tilde\vartheta_i-\vartheta^*)^2}{(\vartheta_i-\vartheta^*)^2}\Big(-N_{T,i}r^T_{N_{T,i}}(\vartheta_i)+N_{T',i}r^{T'}_{N_{T',i}}(\vartheta_i)\Big)+N_{T,i}r^T_{N_{T,i}}(\tilde\vartheta_i)-N_{T',i}r^{T'}_{N_{T',i}}(\tilde\vartheta_i)\Bigg|\\
=&\Bigg|\frac{(\tilde\vartheta_i-\vartheta^*)^2}{(\vartheta_i-\vartheta^*)^2}\left(g^T_{N_{T,i},\vartheta^*}(\vartheta_i)N_{T,i}-g^{T'}_{N_{T',i},\vartheta^*}(\vartheta_i)N_{T',i}\right)\\
&+\frac{(\tilde\vartheta_i-\vartheta^*)^2}{(\vartheta_i-\vartheta^*)^2}\Big(\log\left(\rho^T_{n,\vartheta^*}(\vartheta^*)\right)-\log\left(\rho^{T'}_{n,\vartheta^*}(\vartheta^*)\right)\Big)-\log\left(\rho^T_{n,\vartheta^*}(\vartheta^*)\right)+\log\left(\rho^{T'}_{n,\vartheta^*}(\vartheta^*)\right)\\
&-(\tilde\vartheta_i-\vartheta^*)^2N_{T,i}\frac{r^T_{N_{T,i}}(\vartheta_i)}{(\vartheta_i-\vartheta^*)^2}+(\tilde\vartheta_i-\vartheta^*)^2N_{T',i}\frac{r^{T'}_{N_{T',i}}(\vartheta_i)}{(\vartheta_i-\vartheta^*)^2}\\
&+(\tilde\vartheta_i-\vartheta^*)^2N_{T,i}\frac{r^T_{N_{T,i}}(\tilde\vartheta_i)}{(\tilde\vartheta_i-\vartheta^*)^2}-(\tilde\vartheta_i-\vartheta^*)^2N_{T',i}\frac{r^{T'}_{N_{T',i}}(\tilde\vartheta_i)}{(\tilde\vartheta_i-\vartheta^*)^2}\Bigg|,\end{aligned}$$

woraus mit Hilfe der bisherigen Überlegungen

$$
\begin{aligned}
&|g^T_{N_{T,i},\vartheta^*}(\tilde{\vartheta}_i)N_{T,i} - g^{T'}_{N_{T',i},\vartheta^*}(\tilde{\vartheta}_i)N_{T',i}| \\
&\leq \underbrace{\frac{(\tilde{\vartheta}_i-\vartheta^*)^2}{(\vartheta_i-\vartheta^*)^2}}_{\leq 1} \underbrace{|g^T_{N_{T,i},\vartheta^*}(\vartheta_i)N_{T,i} - g^{T'}_{N_{T',i},\vartheta^*}(\vartheta_i)N_{T',i}|}_{\to 0} \\
&\quad + \underbrace{\frac{(\tilde{\vartheta}_i-\vartheta^*)^2}{(\vartheta_i-\vartheta^*)^2}}_{\leq 1} \underbrace{\left|\log\left(\rho^T_{n,\vartheta^*}(\vartheta^*)\right) - \log\left(\rho^{T'}_{n,\vartheta^*}(\vartheta^*)\right)\right|}_{\to 0} + \underbrace{\left|-\log\left(\rho^T_{n,\vartheta^*}(\vartheta^*)\right) + \log\left(\rho^{T'}_{n,\vartheta^*}(\vartheta^*)\right)\right|}_{\to 0} \\
&\quad + \underbrace{(\tilde{\vartheta}_i-\vartheta^*)^2 N_{T,i}}_{\leq(\vartheta_i-\vartheta^*)^2 N_{T,i}} \underbrace{\frac{|r^T_{N_{T,i}}(\vartheta_i)|}{(\vartheta_i-\vartheta^*)^2}}_{\to 0} + \underbrace{(\tilde{\vartheta}_i-\vartheta^*)^2 N_{T',i}}_{\leq(\vartheta_i-\vartheta^*)^2 N_{T,'i}} \underbrace{\frac{|r^{T'}_{N_{T',i}}(\vartheta_i)|}{(\vartheta_i-\vartheta^*)^2}}_{\to 0} \\
&\quad + \underbrace{(\tilde{\vartheta}_i-\vartheta^*)^2 N_{T,i}}_{\leq(\vartheta_i-\vartheta^*)^2 N_{T,i}} \underbrace{\frac{|r^T_{N_{T,i}}(\tilde{\vartheta}_i)|}{(\tilde{\vartheta}_i-\vartheta^*)^2}}_{\to 0} + \underbrace{(\tilde{\vartheta}_i-\vartheta^*)^2 N_{T',i}}_{\leq(\vartheta_i-\vartheta^*)^2 N_{T',i}} \underbrace{\frac{|r^{T'}_{N_{T',i}}(\tilde{\vartheta}_i)|}{(\tilde{\vartheta}_i-\vartheta^*)^2}}_{\to 0} \longrightarrow 0 \text{ für } n\to\infty
\end{aligned}
$$

folgt.

Also gilt für beliebige Folgen $(\tilde{\vartheta}_i)_{i\in\mathbb{N}}$ mit $\tilde{\vartheta}_i \in (\vartheta^*, \vartheta_i]$ für alle $i \in \mathbb{N}$

$$\lim_{i\to\infty} |g^T_{N_{T,i},\vartheta^*}(\tilde{\vartheta}_i)N_{T,i} - g^{T'}_{N_{T',i},\vartheta^*}(\tilde{\vartheta}_i)N_{T',i}| = 0$$

und man erhält

$$\lim_{i\to\infty} \sup_{\vartheta\in(\vartheta^*,\vartheta_i]} \left| \frac{e^{-g^T_{N_{T,i},\vartheta^*}(\vartheta)N_{T,i}}}{e^{-g^{T'}_{N_{T',i},\vartheta^*}(\vartheta)N_{T',i}}} - 1 \right| = \lim_{i\to\infty} \sup_{\vartheta\in(\vartheta^*,\vartheta_i]} \left| e^{-g^T_{N_{T,i},\vartheta^*}(\vartheta)N_{T,i} + g^{T'}_{N_{T',i},\vartheta^*}(\vartheta)N_{T',i}} - 1 \right| = 0.$$

Somit sind die Bedingungen von Lemma 3.2.2 sind erfüllt und es ist daher

$$\lim_{i\to\infty} \frac{\int_{\vartheta^*}^{\vartheta_i} e^{-g^T_{N_{T,i},\vartheta^*}(\vartheta)N_{T,i}} \mathrm{d}\vartheta}{\int_{\vartheta^*}^{\vartheta_i} e^{-g^{T'}_{N_{T',i},\vartheta^*}(\vartheta)N_{T',i}} \mathrm{d}\vartheta} = 1.$$

Insgesamt resultiert einerseits

$$\lim_{i\to\infty} \frac{\int e^{-g^T_{N_{T,i},\vartheta^*}(\vartheta)N_{T,i}} \mathrm{d}\vartheta}{\int e^{-g^{T'}_{N_{T',i},\vartheta^*}(\vartheta)N_{T',i}} \mathrm{d}\vartheta} \leq \frac{1-\alpha}{1-\alpha-\beta}.$$

Durch die gleiche Behandlung des Kehrwertes des Integrals erhält man andererseits

$$\lim_{i\to\infty} \frac{\int e^{-g^T_{N_{T,i},\vartheta^*}(\vartheta)N_{T,i}} \mathrm{d}\vartheta}{\int e^{-g^{T'}_{N_{T',i},\vartheta^*}(\vartheta)N_{T',i}} \mathrm{d}\vartheta} \geq \frac{1-\alpha-\beta}{1-\alpha}.$$

Da $e^{\mathrm{P}}_{T',T}(\vartheta^*)$ nach Annahme nicht von β oder $(\vartheta_i)_{i\in\mathbb{N}}$ abhängt, genauso wie $\lim_{n\to\infty} e^{\mathrm{vol}}_{T',T,n}(\vartheta^*)$, folgt aufgrund der nach Lemma 3.2.4 gegebenen Darstellung

$$\lim_{n\to\infty} e^{\mathrm{vol}}_{T',T,n}(\vartheta^*) = e^{\mathrm{P}}_{T',T}(\vartheta^*) \lim_{i\to\infty} \left(\frac{\int e^{-g^{T}_{N_{T,i},\vartheta^*}(\vartheta)N_{T,i}} \mathrm{d}\vartheta}{\int e^{-g^{T'}_{N_{T',i},\vartheta^*}(\vartheta)N_{T',i}} \mathrm{d}\vartheta} \right)^2,$$

dass auch der Term

$$\lim_{i\to\infty} \frac{\int e^{-g^{T}_{N_{T,i},\vartheta^*}(\vartheta)N_{T,i}} \mathrm{d}\vartheta}{\int e^{-g^{T'}_{N_{T',i},\vartheta^*}(\vartheta)N_{T',i}} \mathrm{d}\vartheta}$$

nicht mehr von β oder $(\vartheta_i)_{i\in\mathbb{N}}$ abhängen kann.

Nach Annahme ist für ein $0 < \gamma < 1-\alpha$ für jedes $0 < \beta < \gamma$ die Existenz einer Folge $(\vartheta_i)_{i\in\mathbb{N}}$ gesichert, sodass die Pitman-Effizienz bei zugrunde liegender Folge $(\vartheta_i)_{i\in\mathbb{N}}$ und zugrunde liegendem β im Sinne der Definition aus Unterabschnitt 2.4 existiert. Daher können wir $(\vartheta_i)_{i\in\mathbb{N}}$ so wählen, dass $\beta > 0$ beliebig klein wird. Somit folgt für $\beta \downarrow 0$

$$\lim_{i\to\infty} \frac{\int e^{-g^{T}_{N_{T,i},\vartheta^*}(\vartheta)N_{T,i}} \mathrm{d}\vartheta}{\int e^{-g^{T'}_{N_{T',i},\vartheta^*}(\vartheta)N_{T',i}} \mathrm{d}\vartheta} = 1$$

und damit die Behauptung. □

3.2.6 Bemerkung: Im Fall dass es sich bei den betrachteten Tests jeweils um den Likelihood-Quotienten-Test oder den Waldschen Test handelt, so erhält man

$$\lim_{n\to\infty} e^{\mathrm{vol}}_{T',T,n}(\vartheta^*) = e^{\mathrm{P}}_{T',T}(\vartheta^*)$$

unter anderen Voraussetzungen. Betrachte hierzu die Unterabschnitte 4.2 und 4.4.

3.2.7 Bemerkung: Die in Satz 3.2.5 genannte Aussage für die Pitman-Effizienz lässt sich nicht für höherdimensionale Testprobleme in der Form $\mathrm{H}:\ \vartheta = \vartheta_0,\ \mathrm{K}:\ \vartheta \neq \vartheta_0$, mit $\vartheta_0 \in \mathbb{R}^d$, $d > 1$, verallgemeinern. Betrachte hierzu die Ausführungen in Unterabschnitt 4.4.

Zur Vorbereitung auf das kommende Beispiel formulieren wir eine Hilfsaussage.

3.2.8 Lemma: *Es sei $u \in (0,\infty)$ und die Funktion $h : (0,\infty) \to (0,\infty)$ sei definiert durch*

$$h(t) = \frac{\varphi(t) - \varphi(t+u)}{\Phi(t+u) - \Phi(t)},\ t \in (0,\infty),$$

mit φ als Dichtefunktion der Standardnormalverteilung bezüglich des Lebesgue-Borelschen Maßes und Φ als Verteilungsfunktion der Standardnormalverteilung. Dann ist die Funktion h streng monoton wachsend.

Beweis: Für $t \in (0, \infty)$ ist

$$\begin{aligned} h(t) &= \frac{\varphi(t)}{\Phi(t+u) - \Phi(t)} \left(1 - \frac{\varphi(t+u)}{\varphi(t)}\right) \\ &= \frac{\varphi(t)}{\Phi(t+u) - \Phi(t)} \left(1 - \exp\left(-ut - \frac{1}{2}u^2\right)\right). \end{aligned}$$

Die Behauptung ist bewiesen, wenn wir zeigen, dass die Funktion

$$w(t) = \frac{\varphi(t)}{\Phi(t+u) - \Phi(t)}, \ t \in (0, \infty),$$

streng monoton wachsend ist. Dazu stellen wir

$$\begin{aligned} w'(t) &= \frac{-t\varphi(t)\left(\Phi(t+u) - \Phi(t)\right) - \varphi(t)\left(\varphi(t+u) - \varphi(t)\right)}{\left(\Phi(t+u) - \Phi(t)\right)^2} \\ &= \frac{\varphi(t)}{\left(\Phi(t+u) - \Phi(t)\right)^2} z(t), \ t \in (0, \infty), \end{aligned}$$

mit

$$z(t) = \varphi(t) - \varphi(t+u) - t\left(\Phi(t+u) - \Phi(t)\right), \ t \in (0, \infty),$$

fest. Wir wollen beweisen, dass z positiv ist. Wegen

$$\begin{aligned} z'(t) &= -t\varphi(t) + (t+u)\varphi(t+u) - \left(\Phi(t+u) - \Phi(t)\right) - t\left(\varphi(t+u) - \varphi(t)\right) \\ &= u\varphi(t+u) - \left(\Phi(t+u) - \Phi(t)\right) \\ &= u\varphi(t+u) - \int_t^{t+u} \varphi(s)\, ds \\ &< u\varphi(t+u) - u\varphi(t+u) = 0 \end{aligned}$$

für $t \in (0, \infty)$ ist z streng monoton fallend. Da

$$\lim_{t \to \infty} z(t) = 0$$

ist, ist z positiv und damit die behauptete Aussage bewiesen. □

3.2.9 Beispiel: Wir betrachten ein Modell ähnlich dem Beispiel 1.1.1: Beim Vergleich von verbundener und unabhängiger Stichprobenerhebung sind im verbundenen Stichprobenfall mit $\binom{X_1}{Y_1}, \ldots, \binom{X_n}{Y_n}$ unabhängig und identisch verteilte $\mathbb{R}^2$-wertige Zufallsvektoren gegeben, die jeweils die gemeinsame bivariate Normalverteilung

$$\begin{pmatrix} X_1 \\ Y_1 \end{pmatrix} \sim \mathrm{N}_2\left(\begin{pmatrix} a \\ b \end{pmatrix}, \begin{pmatrix} \sigma^2 & \sigma^2\rho \\ \sigma^2\rho & \sigma^2 \end{pmatrix} \right)$$

besitzen, wobei $a, b \in \mathbb{R}$ unbekannt, und $\sigma^2 > 0$ und $\rho \in (-1, 1)$ bekannt seien. Es sei $\mu := a - b$. Für ein gegebenes $\mu_0 \in \mathbb{R}$ betrachten wir zur Behandlung des Testproblems

$$\mathrm{H}: \ \mu = \mu_0, \ \mathrm{K}: \ \mu \neq \mu_0$$

die Testgröße

$$T_n(\mu_0) = \frac{\sqrt{n}|\overline{X}_n - \overline{Y}_n - \mu_0|}{\sqrt{2\sigma^2(1-\rho)}},$$

und schlagen zur Behandlung des Testproblems für ein vorgegebenes Testniveau $\alpha \in (0, 1)$ den Test

$$\text{„Verwirf H, falls } T_n(\mu_0) \geq u_{1-\alpha/2}\text{“}$$

vor.

Im unabhängigen Stichprobenfall sollen dagegen mit $X'_1, \ldots, X'_n$ und $Y'_1, \ldots, Y'_n$ zwei voneinander unabhängige Stichproben unabhängiger und identisch verteilter $\mathbb{R}$-wertiger Zufallsvariablen vorliegen, die jeweils entsprechend

$$X'_1 \sim \mathrm{N}(a, \sigma^2), \ Y'_1 \sim \mathrm{N}(b, \sigma^2)$$

verteilt sind, wobei $a, b \in \mathbb{R}$ unbekannt seien, und $\sigma^2 > 0$ bekannt. Zur Behandlung des Testproblems

$$\mathrm{H}: \ \mu = \mu_0, \ \mathrm{K}: \ \mu \neq \mu_0$$

betrachten wir die Testgröße

$$T'_n(\mu_0) = \frac{\sqrt{n}|\overline{X'_n} - \overline{Y'_n} - \mu_0|}{\sqrt{2\sigma^2}},$$

und schlagen zur Behandlung des Testproblems für ein vorgegebenes Testniveau $\alpha \in (0, 1)$ den Test

$$\text{„Verwirf H, falls } T'_n(\mu_0) \geq u_{1-\alpha/2}\text{“}$$

vor.

Sowohl im verbundenen als auch im unabhängigen Stichprobenfall halten die Tests das Testniveau für alle $n \in \mathbb{N}$ exakt ein, was wie in Beispiel 1.1.1 begründet werden kann.

Im verbundenen Stichprobenfall kann der Test mit Hilfe eines Konfidenzintervalls gemäß

$$\text{„Verwirf H, falls}$$
$$0 \notin [\underbrace{\overline{X}_n - \overline{Y}_n - \mu_0 - \sqrt{2\sigma^2(1-\rho)}u_{1-\alpha/2}/\sqrt{n}}_{=a_n^T}, \underbrace{\overline{X}_n - \overline{Y}_n - \mu_0 + \sqrt{2\sigma^2(1-\rho)}u_{1-\alpha/2}/\sqrt{n}}_{=b_n^T})\text{“}$$

formuliert werden.

Im unabhängigen Stichprobenfall kann der zugehörige Test mit Hilfe eines Konfidenzintervalls gemäß

„Verwirf H, falls
$0 \notin [\underbrace{\overline{X'_n} - \overline{Y'_n} - \mu_0 - \sqrt{2\sigma^2}u_{1-\alpha/2}/\sqrt{n}}_{=a_n^{T'}}, \underbrace{\overline{X'_n} - \overline{Y'_n} - \mu_0 + \sqrt{2\sigma^2}u_{1-\alpha/2}/\sqrt{n}}_{=b_n^{T'}})$“

formuliert werden.

Es ist

$$\mathrm{E}_\mu\left(\sqrt{n}(b_n^T - a_n^T)\right) = 2\sqrt{2\sigma^2}(1-\rho)u_{1-\alpha/2},\ \mathrm{E}_\mu\left(\sqrt{n}(b_n^{T'} - a_n^{T'})\right) = 2\sqrt{2\sigma^2}u_{1-\alpha/2},$$

sodass sich für die zugehörigen Effizienzen basierend auf den Intervalllängen

$$e^{\mathrm{vol}}_{T',T,n} = \left(\frac{\mathrm{E}_\mu\left(\sqrt{n}(b_n^T - a_n^T)\right)}{\mathrm{E}_\mu\left(\sqrt{n}(b_n^{T'} - a_n^{T'})\right)}\right)^2 = \left(\frac{\mathrm{E}_\mu\left(\sqrt{n}(b_n^T - a_n^T)\right)}{\mathrm{E}_\mu\left(\sqrt{n}(b_n^{T'} - a_n^{T'})\right)}\right)^2 = 1-\rho = e^{\mathrm{Lvol}}_{T',T,n} = \lim_{n\to\infty} e^{\mathrm{vol}}_{T',T,n}$$

ergibt.

Im verbundenen Stichprobenfall handelt es sich um den Likelihood-Quotienten-Test, wenn als Beobachtungen die Zufallsvariablen $X_1 - Y_1, \ldots, X_n - Y_n$ zugrunde liegen. Die zugehörige Fisher-Information hat die Gestalt $i_{X_1-Y_1}(\mu_0) = 1/2\sigma^2(1-\rho)$.

Im unabhängigen Stichprobenfall handelt es sich um den Likelihood-Quotienten-Test, wenn als Beobachtungen die Zufallsvariablen $\binom{X'_1}{Y'_1}, \ldots, \binom{X'_n}{Y'_n}$ zugrunde liegen. Die zugehörige Fisher-Information hat die Gestalt $i_{(X'_1,Y'_1)}(\mu_0) = 1/2\sigma^2$.

Somit folgt mit Korollar 4.2.5 für die Pitman-Effizienz der Tests

$$e^P_{T',T} = \frac{2\sigma^2(1-\rho)}{2\sigma^2} = 1-\rho = \lim_{n\to\infty} e^{\mathrm{vol}}_{T',T,n}.$$

Diese Resultate möchten wir nun mit Hilfe von Lemma 3.2.4 und Satz 3.2.5 bestätigen. In der vorliegenden Situation ergibt eine kurze Rechnung, welche analog zu der in Beispiel 1.1.1 präsentierten Rechnung ist, dass sich für Alternativen der Form $\mu_n = \mu^* + k/\sqrt{n}$, $n \in \mathbb{N}$, mit einem $k \in \mathbb{R} \setminus \{0\}$, und für $h(n) := n/(1-\rho)$, $n \in \mathbb{N}$,

$$P_{\mu_n}\left(T_n(\mu^*) < u_{1-\alpha/2}\right) = P_{\mu_n}\left(T_{h(n)}(\mu^*) < u_{1-\alpha/2}\right) = \chi_{u_{1-\alpha/2}}\left(\frac{k}{\sqrt{2\sigma^2(1-\rho)}}\right)$$

ergibt, wobei wir für $y > 0$

$$\chi_y(x) := \Phi(y-x) - \Phi(-y-x),\ x \in \mathbb{R},$$

definieren.

Es ist $\lim_{|x|\to\infty} \chi_{u_{1-\alpha/2}}(x) = 0$, $\chi_{u_{1-\alpha/2}}(x) < 1-\alpha$ für alle $x \in \mathbb{R}\setminus\{0\}$ und $\lim_{x\to 0} \chi_{u_{1-\alpha/2}}(x) = \chi_{u_{1-\alpha/2}}(0) = 1-\alpha$. Also existiert für jedes $\beta \in (0, 1-\alpha)$ ein $k \in \mathbb{R}\setminus\{0\}$ mit

$$P_{\mu^*+k/\sqrt{n}}\big(T_n(\mu^*) < u_{1-\alpha/2}\big) = P_{\mu^*+k/\sqrt{n}}\big(T_{h(n)}(\mu^*) < u_{1-\alpha/2}\big) = \beta.$$

Für die Ableitung der Abbildung χ_y gilt für jedes $y > 0$

$$\chi_y'(x) = -\varphi(y-x) + \varphi(-y-x) = -\varphi(y-x) + \varphi(y+x) \begin{cases} < 0, & x > 0 \\ > 0, & x < 0 \end{cases},$$

mit φ als Dichtefunktion der Standardnormalverteilung. Zudem haben wir für $\mu \in \mathbb{R}\setminus\{\mu^*\}$

$$\begin{aligned} P_\mu\big(T_n(\mu^*) < u_{1-\alpha/2}\big) &= \chi_{u_{1-\alpha/2}}\left(\frac{\sqrt{n}}{\sqrt{2\sigma^2(1-\rho)}}(\mu-\mu^*)\right), \\ P_\mu\big(T_{h(n)}(\mu^*) < u_{1-\alpha/2}\big) &= \chi_{u_{1-\alpha/2}}\left(\frac{\sqrt{n}}{\sqrt{2\sigma^2}}(\mu-\mu^*)\right). \end{aligned}$$

Daher sind für alle $\mu \in \mathbb{R}\setminus\{\mu^*\}$ die Abbildungen

$$n \mapsto P_\mu\big(T_n(\mu^*) < u_{1-\alpha/2}\big) \text{ und } n \mapsto P_\mu\big(T_n(\mu^*) < u_{1-\alpha/2}\big)$$

streng monoton fallend. Unter Beachtung von Bemerkung 2.4.1 ist damit für jedes $0 < \beta < 1-\alpha$ die Existenz einer Folge $(\vartheta_i)_{i\in\mathbb{N}}$ sichergestellt, sodass die Pitman-Effizienz bei zugrunde liegender Folge $(\vartheta_i)_{i\in\mathbb{N}}$ und zugrunde liegendem β im Sinne der Definition aus Unterabschnitt 2.4 existiert.

Es ist also

$$\rho^T_{n,\mu^*}(\mu) = \chi_{u_{1-\alpha/2}}\left(\frac{\sqrt{n}}{\sqrt{2\sigma^2(1-\rho)}}(\mu^*-\mu)\right),\ \rho^{T'}_{n,\mu^*}(\mu) = \chi_{u_{1-\alpha/2}}\left(\frac{\sqrt{n}}{\sqrt{2\sigma^2}}(\mu^*-\mu)\right),$$

und daher

$$\begin{aligned} g^T_{n,\mu^*}(\mu) &= -\frac{1}{n}\log \chi_{u_{1-\alpha/2}}\left(\frac{\sqrt{n}}{\sqrt{2\sigma^2(1-\rho)}}(\mu^*-\mu)\right), \\ g^{T'}_{n,\mu^*}(\mu) &= -\frac{1}{n}\log \chi_{u_{1-\alpha/2}}\left(\frac{\sqrt{n}}{\sqrt{2\sigma^2}}(\mu^*-\mu)\right). \end{aligned}$$

Es ist wegen

$$\mathrm{E}_{\mu^*}\left(\sqrt{n}(b_n^T - a_n^T)\right) = 2\sqrt{2\sigma^2}(1-\rho)u_{1-\alpha/2},\ \mathrm{E}_{\mu^*}\left(\sqrt{n}(b_n^{T'} - a_n^{T'})\right) = 2\sqrt{2\sigma^2}u_{1-\alpha/2}$$

Bedingung (A1) erfüllt, und natürlich ist auch (A2) erfüllt. Dann ist wegen $\chi_y(x) = \chi_y(-x)$ für alle $x \in \mathbb{R}$ und alle $y > 0$ mit Satz 3.2.4

$$\lim_{n\to\infty} e^{\mathrm{vol}}_{T',T,n}(\mu^*) = \lim_{n\to\infty} \left(\frac{\int e^{-g^T_{n,\mu^*}(\mu)n}\mathrm{d}\mu}{\int e^{-g^{T'}_{n,\mu^*}(\mu)n}\mathrm{d}\mu}\right)^2 = \lim_{n\to\infty} \left(\frac{\int \rho^T_{n,\mu^*}(\mu)\mathrm{d}\mu}{\int \rho^{T'}_{n,\mu^*}(\mu)\mathrm{d}\mu}\right)^2$$

$$= \lim_{n\to\infty} \left(\frac{\int \chi_{u_{1-\alpha/2}}\big(\frac{\sqrt{n}}{\sqrt{2\sigma^2(1-\rho)}}(\mu^*-\mu)\big)\mathrm{d}\mu}{\int \chi_{u_{1-\alpha/2}}\big(\frac{\sqrt{n}}{\sqrt{2\sigma^2}}(\mu^*-\mu)\big)\mathrm{d}\mu}\right)^2 = \lim_{n\to\infty} \left(\frac{\int_{\mu^*}^{\infty} \chi_{u_{1-\alpha/2}}\big(\frac{\sqrt{n}}{\sqrt{2\sigma^2(1-\rho)}}(\mu^*-\mu)\big)\mathrm{d}\mu}{\int_{\mu^*}^{\infty} \chi_{u_{1-\alpha/2}}\big(\frac{\sqrt{n}}{\sqrt{2\sigma^2}}(\mu^*-\mu)\big)\mathrm{d}\mu}\right)^2$$

$$= \lim_{n\to\infty} \left(\frac{\int_0^{\infty} \chi_{u_{1-\alpha/2}}\big(\frac{\sqrt{n}}{\sqrt{2\sigma^2(1-\rho)}}\mu\big)\mathrm{d}\mu}{\int_0^{\infty} \chi_{u_{1-\alpha/2}}\big(\frac{\sqrt{n}}{\sqrt{2\sigma^2}}\mu\big)\mathrm{d}\mu}\right)^2 = \lim_{n\to\infty} \left(\frac{\frac{\sqrt{2\sigma^2(1-\rho)}}{\sqrt{n}}\int_0^{\infty} \chi_{u_{1-\alpha/2}}\big(\mu\big)\mathrm{d}\mu}{\frac{\sqrt{2\sigma^2}}{\sqrt{n}}\int_0^{\infty} \chi_{u_{1-\alpha/2}}\big(\mu\big)\mathrm{d}\mu}\right)^2 = 1-\rho.$$

Um die Bedingungen in (A3) zu überprüfen, setzen wir

$$f_n^T(x) := -\frac{1}{n}\log \chi_{u_{1-\alpha/2}}\left(\frac{\sqrt{n}}{\sqrt{2\sigma^2(1-\rho)}}x\right),\ x \in \mathbb{R},$$

$$f_n^{T'}(x) := -\frac{1}{n}\log \chi_{u_{1-\alpha/2}}\left(\frac{\sqrt{n}}{\sqrt{2\sigma^2}}x\right),\ x \in \mathbb{R}.$$

Es gilt wie oben erwähnt $\chi_y(x) = \chi_y(-x)$ für alle $x \in \mathbb{R}$ und alle $y > 0$, und deshalb

$$g^T_{n,\mu^*}(\mu) = f_n^T(|\mu^*-\mu|),\ g^{T'}_{n,\mu^*}(\mu) = f_n^{T'}(|\mu^*-\mu|).$$

Es ist die Funktion χ_y für alle $y > 0$ auf $[0,\infty]$zweimal stetig differenzierbar, und daher sind auch f_n^T und $f_n^{T'}$ auf $[0,\infty]$ zweimal stetig differenzierbar. Für die Ableitung dieser Funktionen gilt

$$(f_n^T)'(x) = -\frac{1}{n}\frac{\sqrt{n}}{\sqrt{2\sigma^2(1-\rho)}}\frac{\chi'_{u_{1-\alpha/2}}\left(\frac{\sqrt{n}}{\sqrt{2\sigma^2(1-\rho)}}x\right)}{\chi_{u_{1-\alpha/2}}\left(\frac{\sqrt{n}}{\sqrt{2\sigma^2(1-\rho)}}x\right)},\ x > 0,$$

$$(f_n^{T'})'(x) = -\frac{1}{n}\frac{\sqrt{n}}{\sqrt{2\sigma^2}}\frac{\chi'_{u_{1-\alpha/2}}\left(\frac{\sqrt{n}}{\sqrt{2\sigma^2}}x\right)}{\chi_{u_{1-\alpha/2}}\left(\frac{\sqrt{n}}{\sqrt{2\sigma^2}}x\right)},\ x > 0.$$

Es ist $\chi_y(x) > 0$ für alle $x > 0$ und $y > 0$. Für die Ableitung der Abbildung χ_y gilt zudem für jedes $y > 0$ $\chi'_y(x) < 0,\ x > 0$, was wir bereits oben sahen. Somit sind die Ableitungen der Funktionen f_n^T und $f_n^{T'}$ auf $(0,\infty)$ strikt positiv. In der Tat sind die Ableitungen der Funktionen f_n^T und $f_n^{T'}$ auf $(0,\infty)$ auch monoton wachsend; siehe Lemma 3.2.8.

Weiter ist für jedes $y > 0$

$$\begin{aligned}\chi''_y(x) &= \varphi'(y-x) - \varphi'(-y-x)\\ &= -\frac{1}{\sqrt{2\pi}}(y-x)e^{-\frac{1}{2}(y-x)^2} + \frac{1}{\sqrt{2\pi}}(-y-x)e^{-\frac{1}{2}(-y-x)^2},\ x \in \mathbb{R},\end{aligned}$$

also für jedes $y > 0$

$$\chi_y''(0) = -\frac{2}{\sqrt{2\pi}} y e^{-\frac{1}{2}y^2}.$$

Zudem ist

$$\begin{aligned}
(g_{n,\mu^*}^T)''(\mu) = &- \frac{1}{2\sigma^2(1-\rho)} \frac{\chi''_{u_{1-\alpha/2}}\left(\frac{\sqrt{n}}{\sqrt{2\sigma^2(1-\rho)}}(\mu^*-\mu)\right)}{\chi_{u_{1-\alpha/2}}\left(\frac{\sqrt{n}}{\sqrt{2\sigma^2(1-\rho)}}(\mu^*-\mu)\right)} \\
&+ \frac{1}{2\sigma^2(1-\rho)} \frac{\chi'_{u_{1-\alpha/2}}\left(\frac{\sqrt{n}}{\sqrt{2\sigma^2(1-\rho)}}(\mu^*-\mu)\right)^2}{\chi_{u_{1-\alpha/2}}\left(\frac{\sqrt{n}}{\sqrt{2\sigma^2(1-\rho)}}(\mu^*-\mu)\right)^2}, \\
(g_{n,\mu^*}^{T'})''(\mu) = &- \frac{1}{2\sigma^2} \frac{\chi''_{u_{1-\alpha/2}}\left(\frac{\sqrt{n}}{\sqrt{2\sigma^2}}(\mu^*-\mu)\right)}{\chi_{u_{1-\alpha/2}}\left(\frac{\sqrt{n}}{\sqrt{2\sigma^2}}(\mu^*-\mu)\right)} + \frac{1}{2\sigma^2} \frac{\chi'_{u_{1-\alpha/2}}\left(\frac{\sqrt{n}}{\sqrt{2\sigma^2}}(\mu^*-\mu)\right)^2}{\chi_{u_{1-\alpha/2}}\left(\frac{\sqrt{n}}{\sqrt{2\sigma^2}}(\mu^*-\mu)\right)^2}.
\end{aligned}$$

Wegen $\chi_y(0) = 2\Phi(y) - 1$ und $\chi_y'(0) = 0$ für jedes $y > 0$ folgt daher für jedes $y > 0$

$$\begin{aligned}
a_{2,n}^T &= (g_{n,\mu^*}^T)''(\mu^*) = -\frac{1}{2\sigma^2(1-\rho)} \frac{\chi''_{u_{1-\alpha/2}}(0)}{2\Phi(u_{1-\alpha/2}) - 1} = \frac{1}{2\sigma^2(1-\rho)} \frac{\frac{2}{\sqrt{2\pi}} u_{1-\alpha/2} e^{-\frac{1}{2}u_{1-\alpha/2}^2}}{2\Phi(u_{1-\alpha/2}) - 1} > 0, \\
a_{2,n}^{T'} &= (g_{n,\mu^*}^{T'})''(\mu^*) = -\frac{1}{2\sigma^2} \frac{\chi''_{u_{1-\alpha/2}}(0)}{2\Phi(u_{1-\alpha/2}) - 1} = \frac{1}{2\sigma^2} \frac{\frac{2}{\sqrt{2\pi}} u_{1-\alpha/2} e^{-\frac{1}{2}u_{1-\alpha/2}^2}}{2\Phi(u_{1-\alpha/2}) - 1} > 0.
\end{aligned}$$

Somit ist auch $\liminf_{n\to\infty} a_{2,n}^T > 0$ und $\liminf_{n\to\infty} a_{2,n}^{T'} > 0$ erfüllt. Ebenso existiert eine Umgebung U um μ^* mit

$$\sup_{n\in\mathbb{N}} \sup_{\mu\in U} |(g_{n,\mu^*}^T)'''(\mu)| < \infty \text{ und } \sup_{n\in\mathbb{N}} \sup_{\mu\in U} |(g_{n,\mu^*}^{T'})'''(\mu)| < \infty,$$

sodass wegen Bemerkung 3.2.3 Teil e) auch $\lim_{i\to\infty} \sup_{n\in\mathbb{N}} |r_n^V(\tilde{\vartheta}_i)|/(\tilde{\vartheta}_i - \vartheta^*)^2 = 0$ für alle Folgen $(\tilde{\vartheta}_i)_{i\in\mathbb{N}}$ mit $\tilde{\vartheta}_i \neq \vartheta^*$ für alle $i \in \mathbb{N}$ und $\lim_{i\to\infty} \tilde{\vartheta}_i = \vartheta^*$ erfüllt ist.

Also gilt neben (A1) und (A2) auch (A3) und es folgt mit Satz und 3.2.5

$$e_{T',T}^{\mathrm{P}} = \lim_{n\to\infty} e_{T',T,n}^{\mathrm{vol}} = 1 - \rho.$$

4 Effizienzen beim Vergleich von Experimenten

4.1 Modell

Für das zugrunde liegende Modell in diesem Abschnitt soll die Basis das in Unterabschnitt 1.2 formulierte Grundmodell mit einer wie folgt strukturierten Parametermenge Θ sein.

Für $k, k', d_1 \in \mathbb{N}$ mit $d_1 \leq k$ und $d_1 \leq k'$ sei

$$\emptyset \neq \Theta \subset \mathbb{R}^{d_1} \times \mathbb{R}^{k-d_1} \times \mathbb{R}^{k'-d_1}.$$

Elemente $\vartheta \in \Theta$ sollen in der Folge gemäß $\vartheta = (\vartheta_1, \vartheta_2, \vartheta_2') \in \mathbb{R}^{d_1} \times \mathbb{R}^{k-d_1} \times \mathbb{R}^{k'-d_1}$ geschrieben werden. Häufig ist es zweckmäßig, Θ als Teilmenge $\Theta \subset \mathbb{R}^{k+k'-d_1}$ und Elemente $\vartheta = (\vartheta_1, \vartheta_2, \vartheta_2') \in \Theta \subset \mathbb{R}^{d_1} \times \mathbb{R}^{k-d_1} \times \mathbb{R}^{k'-d_1}$ als Spaltenvektoren

$$\vartheta = \begin{pmatrix} \vartheta_1 \\ \vartheta_2 \\ \vartheta_2' \end{pmatrix} \in \Theta \subset \mathbb{R}^{k+k'-d_1}$$

aufzufassen, um wie gewohnt rechnen zu können. Wir erinnern an die in Abschnitt 1.2 eingeführten Folgen von je unabhängigen und identisch verteilten Zufallsvariablen $\zeta_j, j \in \mathbb{N}$, und $\zeta_j', j \in \mathbb{N}$, und verlangen, dass

$$P^{\zeta_1}_{(\vartheta_1,\vartheta_2,\vartheta_2')} = P^{\zeta_1}_{(\vartheta_1,\vartheta_2,\tilde{\vartheta}_2')} \ \forall\ (\vartheta_1, \vartheta_2, \vartheta_2'), (\vartheta_1, \vartheta_2, \tilde{\vartheta}_2') \in \Theta, \tag{4.1.1}$$

sowie

$$P^{\zeta_1'}_{(\vartheta_1,\vartheta_2,\vartheta_2')} = P^{\zeta_1'}_{(\vartheta_1,\tilde{\vartheta}_2,\vartheta_2')} \ \forall\ (\vartheta_1, \vartheta_2, \vartheta_2'), (\vartheta_1, \tilde{\vartheta}_2, \vartheta_2') \in \Theta \tag{4.1.2}$$

gilt. D.h. die Verteilung von ζ_1 unter P_ϑ hängt nicht von ϑ_2' ab, und die Verteilung von ζ_1' unter P_ϑ hängt nicht von ϑ_2 ab. Allgemein ist zu bemerken, dass im Fall $d_1 = k$ bzw. $d_1 = k'$ die Komponente ϑ_2 bzw. die Komponente ϑ_2' verschwindet.

Auf der zugrunde liegenden Parametermenge Θ ist für $(\vartheta_1, \vartheta_2, \vartheta_2'), (\tilde{\vartheta}_1, \tilde{\vartheta}_2, \tilde{\vartheta}_2') \in \Theta$ durch

$$(\vartheta_1, \vartheta_2, \vartheta_2') \sim (\tilde{\vartheta}_1, \tilde{\vartheta}_2, \tilde{\vartheta}_2') \ :\Leftrightarrow\ \vartheta_1 = \tilde{\vartheta}_1,\ \vartheta_2 = \tilde{\vartheta}_2,\ P^{\zeta_1}_{(\vartheta_1,\vartheta_2,\vartheta_2')} = P^{\zeta_1}_{(\tilde{\vartheta}_1,\tilde{\vartheta}_2,\tilde{\vartheta}_2')}$$

eine Äquivalenzrelation definiert. Genauso ist auf der zugrunde liegenden Parametermenge Θ für $(\vartheta_1, \vartheta_2, \vartheta_2'), (\tilde{\vartheta}_1, \tilde{\vartheta}_2, \tilde{\vartheta}_2') \in \Theta$ durch

$$(\vartheta_1, \vartheta_2, \vartheta_2') \sim' (\tilde{\vartheta}_1, \tilde{\vartheta}_2, \tilde{\vartheta}_2') \ :\Leftrightarrow\ \vartheta_1 = \tilde{\vartheta}_1,\ \vartheta_2' = \tilde{\vartheta}_2',\ P^{\zeta_1'}_{(\vartheta_1,\vartheta_2,\vartheta_2')} = P^{\zeta_1'}_{(\tilde{\vartheta}_1,\tilde{\vartheta}_2,\tilde{\vartheta}_2')}$$

eine Äquivalenzrelation definiert. Es seien $[\vartheta]$ und $[\vartheta]'$ die zu den Äquivalenzrelationen $\sim$ und $\sim'$ zugehörigen Äquivalenzklassen. Für je einen beliebigen Vertreter $\vartheta = (\vartheta_1, \vartheta_2, \vartheta_2') \in [\vartheta]$ der jeweiligen Äquivalenzklasse $[\vartheta]$ definieren wir Wahrscheinlichkeitsmaße

$$Q_{(\vartheta_1,\vartheta_2)} := P_{(\vartheta_1,\vartheta_2,\vartheta_2')},$$

und genauso definieren wir für je einen beliebigen Vertreter $\vartheta = (\vartheta_1, \vartheta_2, \vartheta_2') \in [\vartheta]'$ der jeweiligen Äquivalenzklasse $[\vartheta]'$ Wahrscheinlichkeitsmaße

$$M_{(\vartheta_1,\vartheta_2')} := P_{(\vartheta_1,\vartheta_2,\vartheta_2')}.$$

Selbstverständlich hängen die so definierten Wahrscheinlichkeitsmaße $Q_{(\vartheta_1,\vartheta_2)}$ noch von der Wahl der Vertreter, also von ϑ_2', ab, und genauso hängen die so definierten Wahrscheinlichkeitsmaße $M_{(\vartheta_1,\vartheta_2')}$ noch von der Wahl der Vertreter, also von ϑ_2, ab. Dennoch wollen wir die Abhängigkeit der Wahrscheinlichkeitsmaße $Q_{(\vartheta_1,\vartheta_2)}$ von ϑ_2' genauso wie die Abhängigkeit der Wahrscheinlichkeitsmaße $M_{(\vartheta_1,\vartheta_2')}$ von ϑ_2 in der Notation unterdrücken, da im Folgenden im Zusammenhang mit den Wahrscheinlichkeitsmaßen $Q_{(\vartheta_1,\vartheta_2)}$ ausschließlich Ereignisse basierend auf den Zufallsvariablen $\zeta_1, \ldots, \zeta_n$ betrachtet werden, und genauso im Zusammenhang mit den Wahrscheinlichkeitsmaßen $M_{(\vartheta_1,\vartheta_2')}$ ausschließlich Ereignisse basierend auf den Zufallsvariablen $\zeta_1', \ldots, \zeta_n'$ betrachtet werden. Daher handelt es sich bei den betrachteten Wahrscheinlichkeiten stets um Wahrscheinlichkeiten, die mit Hilfe der Bildmaße $Q^{\zeta_1}_{(\vartheta_1,\vartheta_2)}$ bzw. $M^{\zeta_1'}_{(\vartheta_1,\vartheta_2')}$ berechnet werden können, die wegen (4.1.1) bzw. (4.1.2) unabhängig von ϑ_2' bzw. ϑ_2 sind.

Wir definieren

$$\tilde{\Theta} := \{(\vartheta_1, \vartheta_2) \in \mathbb{R}^{d_1} \times \mathbb{R}^{k-d_1}; \exists\, \vartheta_2' \in \mathbb{R}^{k'-d_1} : (\vartheta_1, \vartheta_2, \vartheta_2') \in \Theta\},$$

und

$$\tilde{\Theta}' := \{(\vartheta_1, \vartheta_2') \in \mathbb{R}^{d_1} \times \mathbb{R}^{k'-d_1}; \exists\, \vartheta_2 \in \mathbb{R}^{k-d_1} : (\vartheta_1, \vartheta_2, \vartheta_2') \in \Theta\}.$$

Dann gilt wegen $\emptyset \neq \Theta$ auch $\emptyset \neq \tilde{\Theta}$ und $\emptyset \neq \tilde{\Theta}'$. Elemente $\vartheta \in \tilde{\Theta}$ sollen in der Folge gemäß $\vartheta = (\vartheta_1, \vartheta_2) \in \mathbb{R}^{d_1} \times \mathbb{R}^{k-d_1}$ geschrieben werden, und Elemente $\vartheta' \in \tilde{\Theta}'$ sollen in der Folge gemäß $\vartheta' = (\vartheta_1, \vartheta_2') \in \mathbb{R}^{d_1} \times \mathbb{R}^{k'-d_1}$ geschrieben werden. Auch hier soll noch einmal darauf hingewiesen werden, dass im Fall $d_1 = k$ bzw. $d_1 = k'$ die Komponente ϑ_2 bzw. die Komponente ϑ_2' verschwindet. Häufig ist es zweckmäßig, $\tilde{\Theta}$ bzw. $\tilde{\Theta}'$ als Teilmenge $\tilde{\Theta} \subset \mathbb{R}^k$ bzw. $\tilde{\Theta}' \subset \mathbb{R}^{k'}$ und Elemente $\vartheta = (\vartheta_1, \vartheta_2) \in \tilde{\Theta} \subset \mathbb{R}^{d_1} \times \mathbb{R}^{k-d_1}$ bzw. $\vartheta' = (\vartheta_1, \vartheta_2') \in \tilde{\Theta}' \subset \mathbb{R}^{d_1} \times \mathbb{R}^{k'-d_1}$ als Spaltenvektoren

$$\vartheta = \begin{pmatrix} \vartheta_1 \\ \vartheta_2 \end{pmatrix} \in \tilde{\Theta} \subset \mathbb{R}^k \text{ bzw. } \vartheta' = \begin{pmatrix} \vartheta_1 \\ \vartheta_2' \end{pmatrix} \in \tilde{\Theta}' \subset \mathbb{R}^{k'}$$

aufzufassen, um wie gewohnt rechnen zu können.

Es sind

$$(\Omega, \mathfrak{A}, \{Q_\vartheta; \vartheta \in \tilde{\Theta}\}) \text{ und } (\Omega, \mathfrak{A}, \{M_{\vartheta'}; \vartheta' \in \tilde{\Theta}'\})$$

statistische Räume und es gilt

$$\{P_\vartheta^{\zeta_1}; \vartheta \in \Theta\} = \{Q_\vartheta^{\zeta_1}; \vartheta \in \tilde{\Theta}\} \text{ und } \{P_\vartheta^{\zeta_1'}; \vartheta \in \Theta\} = \{M_{\vartheta'}^{\zeta_1'}; \vartheta' \in \tilde{\Theta}'\}.$$

Wir nehmen an, dass die Familien von Verteilungen

$$\{Q_{\vartheta}^{\zeta_1}; \vartheta \in \tilde{\Theta}\} \text{ und } \{M_{\vartheta'}^{\zeta_1'}; \vartheta' \in \tilde{\Theta}'\}$$

jeweils injektiv parametrisiert sind.

Weiter definieren wir

$$\Theta^1 := \{\vartheta_1 \in \mathbb{R}^{d_1}; \exists\ (\vartheta_2, \vartheta_2') \in \mathbb{R}^{k-d_1} \times \mathbb{R}^{k'-d_1} : (\vartheta_1, \vartheta_2, \vartheta_2') \in \Theta\}.$$

Dann gilt wegen $\emptyset \neq \Theta$ auch $\emptyset \neq \Theta^1$. Zudem gehen wir von der Existenz einer Menge $\emptyset \neq \Theta_0^1 \subsetneq \Theta^1$ mit

$$\Theta_0 = \{(\vartheta_1, \vartheta_2, \vartheta_2') \in \Theta; \vartheta_1 \in \Theta_0^1\}$$

aus. Es ist wegen $\emptyset \neq \Theta_0^1 \subsetneq \Theta^1$ auch $\emptyset \neq \Theta_0 \subsetneq \Theta$. Wir setzen $\Theta_1^1 := \Theta^1 \setminus \Theta_0^1$, und weiter

$$\tilde{\Theta}_0 := \{(\vartheta_1, \vartheta_2) \in \tilde{\Theta}; \vartheta_1 \in \Theta_0^1\},\ \tilde{\Theta}_1 := \tilde{\Theta} \setminus \tilde{\Theta}_0,$$

und

$$\tilde{\Theta}_0' := \{(\vartheta_1, \vartheta_2') \in \tilde{\Theta}'; \vartheta_1 \in \Theta_0^1\},\ \tilde{\Theta}_1' := \tilde{\Theta}' \setminus \tilde{\Theta}_0'.$$

Dann gilt

$$\Theta^1 = \{\vartheta_1 \in \mathbb{R}^{d_1}; \exists\ \vartheta_2 \in \mathbb{R}^{k-d_1} : (\vartheta_1, \vartheta_2) \in \tilde{\Theta}\} = \{\vartheta_1 \in \mathbb{R}^{d_1}; \exists\ \vartheta_2' \in \mathbb{R}^{k'-d_1} : (\vartheta_1, \vartheta_2') \in \tilde{\Theta}'\},$$

sowie

$$\Theta_0^1 = \{\vartheta_1 \in \mathbb{R}^{d_1}; \exists\ \vartheta_2 \in \mathbb{R}^{k-d_1} : (\vartheta_1, \vartheta_2) \in \tilde{\Theta}_0\} = \{\vartheta_1 \in \mathbb{R}^{d_1}; \exists\ \vartheta_2' \in \mathbb{R}^{k'-d_1} : (\vartheta_1, \vartheta_2') \in \tilde{\Theta}_0'\},$$

und

$$\Theta_1^1 = \{\vartheta_1 \in \mathbb{R}^{d_1}; \exists\ \vartheta_2 \in \mathbb{R}^{k-d_1} : (\vartheta_1, \vartheta_2) \in \tilde{\Theta}_1\} = \{\vartheta_1 \in \mathbb{R}^{d_1}; \exists\ \vartheta_2' \in \mathbb{R}^{k'-d_1} : (\vartheta_1, \vartheta_2') \in \tilde{\Theta}_1'\}.$$

Dies lässt erkennen, dass wegen $\emptyset \neq \Theta_0^1 \subsetneq \Theta^1$ auch $\emptyset \neq \tilde{\Theta}_0 \subsetneq \tilde{\Theta}$ sowie $\emptyset \neq \tilde{\Theta}_0' \subsetneq \tilde{\Theta}'$ gilt.

Nun kann mit $\vartheta = (\vartheta_1, \vartheta_2, \vartheta_2') \in \Theta$ das Testproblem (1.2.1),

$$\mathrm{H}:\ \vartheta \in \Theta_0,\ \mathrm{K}:\ \vartheta \in \Theta_1,$$

äquivalent gemäß

$$\mathrm{H}:\ \vartheta_1 \in \Theta_0^1,\ \mathrm{K}:\ \vartheta_1 \in \Theta_1^1,$$

bzw.

$$\mathrm{H}:\ (\vartheta_1, \vartheta_2) \in \tilde{\Theta}_0,\ \mathrm{K}:\ (\vartheta_1, \vartheta_2) \in \tilde{\Theta}_1,$$

bzw.

$$\mathrm{H}: \ (\vartheta_1, \vartheta_2') \in \tilde{\Theta}_0', \ \mathrm{K}: \ (\vartheta_1, \vartheta_2') \in \tilde{\Theta}_1',$$

formuliert werden.

Es seien die Mengen $\tilde{\Theta}$ und $\tilde{\Theta}'$ offen und konvex. Zudem gelte $\partial\Theta_0^1 \subset \Theta_0^1$, $\partial\tilde{\Theta}_0 \subset \tilde{\Theta}_0$, $\partial\tilde{\Theta}_0' \subset \tilde{\Theta}_0'$ und $\partial\Theta_0 \subset \Theta_0$, wobei $\partial\Theta_0^1$ bzw. $\partial\tilde{\Theta}_0$ bzw. $\partial\tilde{\Theta}_0'$ bzw. $\partial\Theta_0$ den Rand von Θ_0^1 bzw. $\tilde{\Theta}_0$ bzw. $\tilde{\Theta}_0'$ bzw. Θ_0 bezüglich der Relativtopologie auf Θ^1 bzw. $\tilde{\Theta}$ bzw. $\tilde{\Theta}'$ bzw. Θ bezeichne.

Im Zusammenhang mit lokalen relativen Effizienzen, also der lokalen Hodges-Lehmann-, lokalen Bahadur- und Pitman-Effizienz, sollten zur Effizienzuntersuchung sinnvollerweise Alternativen der Form

$$\begin{pmatrix} \vartheta_0^1 \\ \vartheta_2 \\ \vartheta_2' \end{pmatrix} + \begin{pmatrix} \vartheta_1 \\ 0 \\ 0 \end{pmatrix}$$

mit $\vartheta_1 \in \mathbb{R}^{d_1} \setminus \{0\}$, $(\vartheta_0^1, \vartheta_2, \vartheta_2') \in \partial\Theta_0$, betrachtet werden, die $(\vartheta_0^1 + \vartheta_1, \vartheta_2, \vartheta_2') \in \Theta_1$ erfüllen.

4.2 Pitman-Effizienz bei Likelihood-Quotienten-Tests und Waldschen Tests

Bewiesen werden soll ein allgemeines Resultat, das es erlaubt, die Pitman- Effizienz von zwei Tests explizit anzugeben, wenn es sich bei den Tests jeweils um den Likelihood-Quotienten-Test oder den Waldschen Test handelt. Zur Anwendung kann es in den Abschnitten 8 und 9 kommen.

Likelihood-Quotienten-Tests oder Waldsche Test besitzen unter Alternativen einer bestimmten Form asymptotisch eine nichtzentrale Chi-Quadrat- Verteilung; siehe hierzu z.B. Theorem 16.7 in [17]. Es ist bekannt, dass sich in vielen Fällen die Pitman-Effizienz von Tests, deren Grenzverteilung unter sogenannten Pitman-Alternativen eine nichtzentrale Chi-Quadrat-Verteilung ist, als Quotient der beteiligten Nichtzentralitätsparameter ergibt. Ein allgemeiner Ansatz zur Berechnung der Pitman-Effizienz in solchen Situationen ist bei Puri und Sen ([18]) oder bei Hannan ([26]) zu finden. Überlegungen dazu stellen auch Müller-Funk und Witting ([19]) an. Von einer einfachen Hypothese $\mathrm{H} : \vartheta = \vartheta_0$ mit bekanntem $\vartheta_0 \in \mathbb{R}^d$ ausgehend, betrachten diese Autoren vom Stichprobenumfang $n \in \mathbb{N}$ abhängige Alternativen der Form

$$\vartheta_n = \vartheta_0 + \frac{1}{\sqrt{n}}\vartheta$$

mit einem gegebenen $\vartheta \in \mathbb{R}^d$. Wir wählen eine der in Abschnitt 2.4 gegebenen Definition der Pitman-Effizienz gerecht werdende Herangehensweise.

Im Folgenden wollen wir Annahmen an die Zufallsvariablen ζ_1 und ζ_1' bei den zugrunde liegenden injektiven Parametrisierungen durch $\tilde{\Theta}$ und $\tilde{\Theta}'$ formulieren. Der Einfachheit halber formulieren wir diese Annahmen nur für ζ_1 und $\tilde{\Theta}$; sie sollen natürlich in analoger Weise auch für ζ_1' und $\tilde{\Theta}'$ gelten. Dabei werden die im Zusammenhang mit ζ_1 und $\tilde{\Theta}$ eingeführten Größen im Zusammenhang mit ζ_1' und $\tilde{\Theta}'$ mit einem Strich versehen.

Wir nehmen an, dass ein σ-endliches Maß μ auf $(R, \mathfrak{S})$ existiert, sodass für jedes $\vartheta \in \tilde{\Theta}$ $Q_\vartheta^{\zeta_1} \ll \mu$ gilt, und $f(\cdot;\vartheta)$ eine μ-Dichte von $Q_\vartheta^{\zeta_1}$ ist. Zur Vereinfachung der Notation definieren wir

$$L(z, \vartheta) := \log f(z;\vartheta),\ z \in R,\ \vartheta \in \tilde{\Theta},$$

mit $\log 0 := -\infty$. Wir setzen dabei generell voraus, dass $f(z;\vartheta) > 0$ für alle $\vartheta \in \tilde{\Theta}$ und alle $z \in R$.

Im Fall einer zusammengesetzten Hypothese $\tilde{\Theta}_0$ sei stets $k \geq 2$ und es sei $g(\Delta) = \tilde{\Theta}_0$ mit einer Menge $\Delta \subset \mathbb{R}^h$ offen und konvex, $1 \leq h < k$, und einer zweimal stetig differenzierbaren und injektiven Abbildung $g : \Delta \to \tilde{\Theta}$. Ferner besize die Funktionalmatrix von g an der Stelle $\eta \in \Delta$, $\mathcal{J}_g(\eta)$, für alle $\eta \in \Delta$ den Rang h.

Nun betrachten wir noch für $n \in \mathbb{N}$ die Schätzfunktionen $\hat{\vartheta}_n : (R^n, \mathfrak{S}^n) \to (\tilde{\Theta}, \mathfrak{B}^k_{\tilde{\Theta}})$ und, im Fall einer zusammengesetzten Hypothese, $\hat{\eta}_n : (R^n, \mathfrak{S}^n) \to (\Delta, \mathfrak{B}^h_\Delta)$, die in typischen Anwendungsfällen die Maximum-Likelihood-Schätzfunktionen sind. Um die Notation zu vereinfachen, schreiben wir oft auch einfach $\hat{\vartheta}_n = \hat{\vartheta}_n(\zeta_1, \ldots, \zeta_n)$ und, im Fall einer zusammengesetzten Hypothese, $\hat{\eta}_n = \hat{\eta}_n(\zeta_1, \ldots, \zeta_n)$.

Im weiteren Verlauf der Arbeit wird folgende Notation benutzt: Es sei eine Folge $(\vartheta_n)_{n\in\mathbb{N}}$ mit $\vartheta_n \in \tilde{\Theta}$ für alle $n \in \mathbb{N}$ gegeben. Für eine auf dem zugrunde liegendem statistischen Raum definierte Folge von reellwertigen Zufallsvariablen $(U_n)_{n\in\mathbb{N}}$ und eine reellwertige Zufallsvariable U gelte $\lim_{n\to\infty} U_n = U$ in Q_{ϑ_n}-Wahrscheinlichkeit, wenn für alle $\varepsilon > 0$ der Grenzwert $\lim_{n\to\infty} Q_{\vartheta_n}(|U_n - U| > \varepsilon)$ existiert und $\lim_{n\to\infty} Q_{\vartheta_n}(|U_n - U| > \varepsilon) = 0$ gilt. Dies soll auch äquivalent durch die Schreibweise $U_n \overset{Q_{\vartheta_n}}{\to} U$ für $n \to \infty$ ausgedrückt werden können. Ein Term, der mit $o_{Q_{\vartheta_n}}(1)$ bezeichnet wird, gehe in diesem Sinne auch in Q_{ϑ_n}-Wahrscheinlichkeit gegen $0 \in \mathbb{R}$.

Des Weiteren werden wir folgende Notation benutzen: Für ein $\vartheta \in \tilde{\Theta}$ sei eine Folge $(\vartheta_n)_{n\in\mathbb{N}}$ mit $\vartheta_n \in \tilde{\Theta}$ für alle $n \in \mathbb{N}$ und $\lim_{n\to\infty} \vartheta_n = \vartheta$ gegeben. Es sei $l \in \mathbb{N}$. Für eine auf dem zugrunde liegendem statistischen Raum definierte Folge von $\mathbb{R}^l$-wertigen Zufallsvariablen $(U_n)_{n\in\mathbb{N}}$ und eine $\mathbb{R}^l$-wertige Zufallsvariable U gelte $U_n \overset{\vartheta_n}{\rightsquigarrow} U$ für $n \to \infty$, wenn für alle stetigen und beschränkten Funktionen $h : \mathbb{R}^l \to \mathbb{R}$ $\lim_{n\to\infty} \int h \mathrm{d}Q_{\vartheta_n}^{U_n} = \int h \mathrm{d}Q_\vartheta^U$ zutrifft. Äquivalent soll hierfür auch $Q_{\vartheta_n}^{U_n} \overset{\vartheta_n}{\rightsquigarrow} Q_\vartheta^U$ für $n \to \infty$ geschrieben werden.

Betrachte nun die folgenden Bedingungen:

(B1) Es sei die Abbildung $\vartheta \mapsto f(z;\vartheta)$ für alle $z \in R$ zweimal stetig differenzierbar.

(B2) Für alle $1 \leq i,j \leq k$ und für alle $\vartheta \in \tilde{\Theta}$ gelte $\int \frac{\partial}{\partial \vartheta_i} f(z;\vartheta)\mathrm{d}\mu(z) = \int \frac{\partial^2}{\partial \vartheta_i \partial \vartheta_j} f(z;\vartheta)\mathrm{d}\mu(z) = 0$ und für alle $\vartheta \in \tilde{\Theta}$ existiere die Fisher-Informationsmatrix aus $Q_\vartheta^{\zeta_1}$, $i_{\zeta_1}(\vartheta)$, besitze endliche Einträge und sei positiv definit.

(B3) Es existiere für jedes $\vartheta \in \tilde{\Theta}$ und jedes $1 \leq i \leq k$ ein $\delta > 0$ sodass mit $U_\delta := \{\tilde{\vartheta} \in \mathbb{R}^k : |\tilde{\vartheta} - \vartheta| < \delta\} \subset \tilde{\Theta}$ und $M_\delta(z) := \sup_{\vartheta^* \in U_\delta} |\frac{\partial}{\partial \tilde{\vartheta}_i} L(z,\tilde{\vartheta})|_{\tilde{\vartheta}=\vartheta^*}| \sup_{\vartheta^* \in U_\delta} f(z;\vartheta^*)$, $z \in R$, $\int M_\delta(z)\mathrm{d}\mu(z) < \infty$ ist. Genauso existiere für jedes $\vartheta \in \tilde{\Theta}$ und jede $1 \leq i,j \leq k$ ein $\delta > 0$ sodass mit $\tilde{M}_\delta(z) := \sup_{\vartheta^* \in U_\delta} |\frac{\partial^2}{\partial \tilde{\vartheta}_i \partial \tilde{\vartheta}_j} L(z,\tilde{\vartheta})|_{\tilde{\vartheta}=\vartheta^*}| \sup_{\vartheta^* \in U_\delta} f(z;\vartheta^*)$, $z \in R$, $\int \tilde{M}_\delta(z)\mathrm{d}\mu(z) < \infty$ ist.

(B4) Es sei für $n \in \mathbb{N}$ $\hat{\vartheta}_n : (R^n, \mathfrak{S}^n) \to (\tilde{\Theta}, \mathfrak{B}^k_{\tilde{\Theta}})$, sodass für alle $\vartheta \in \tilde{\Theta}$ und alle Folgen $(\vartheta_n)_{n\in\mathbb{N}}$ mit $\vartheta_n \in \tilde{\Theta}$ für alle $n \in \mathbb{N}$ und $\lim_{n\to\infty} \vartheta_n = \vartheta$ $\lim_{n\to\infty} \hat{\vartheta}_n(\zeta_1, ..., \zeta_n) = \vartheta$ in Q_{ϑ_n}-Wahrscheinlichkeit, sowie für $1 \leq i \leq k$ $\lim_{n\to\infty} \frac{1}{\sqrt{n}} \sum_{j=0}^n \frac{\partial}{\partial \tilde{\vartheta}_i} L(\zeta_j, \tilde{\vartheta})|_{\tilde{\vartheta}=\hat{\vartheta}_n(\zeta_1,...,\zeta_n)} = 0$ in Q_{ϑ_n}-Wahrscheinlichkeit gilt.

(B5) Es sei für $n \in \mathbb{N}$ $\hat{\eta}_n : (R^n, \mathfrak{S}^n) \to (\Delta, \mathfrak{B}^h_\Delta)$, sodass für alle $\vartheta_0 \in \tilde{\Theta}_0$ und alle Folgen $(\vartheta_n)_{n\in\mathbb{N}}$ mit $\vartheta_n \in \tilde{\Theta}$ für alle $n \in \mathbb{N}$ und $\lim_{n\to\infty} \vartheta_n = \vartheta_0$ $\lim_{n\to\infty} \hat{\eta}_n(\zeta_1, ..., \zeta_n) = \eta_0$ in Q_{ϑ_n}-Wahrscheinlichkeit gilt, mit $\eta_0 := g^{-1}(\vartheta_0)$, sowie für $1 \leq i \leq h$ $\lim_{n\to\infty} \frac{1}{\sqrt{n}} \sum_{j=0}^n \frac{\partial}{\partial \tilde{\eta}_i} L(\zeta_j, g(\tilde{\eta}))|_{\tilde{\eta}=\hat{\eta}_n(\zeta_1,...,\zeta_n)} = 0$ in Q_{ϑ_n}-Wahrscheinlichkeit gilt.

Für $\vartheta \in \tilde{\Theta}$, mit $\vartheta = (\vartheta_1, \ldots, \vartheta_k)^t \in \mathbb{R}^k$, besitzt die Fisher-Informationsmatrix aus $Q_\vartheta^{\zeta_1}$ konkret die Gestalt

$$i_{\zeta_1}(\vartheta) = \left(\mathrm{E}_\vartheta \left(\frac{\partial}{\partial \vartheta_i} L(\zeta_1, \vartheta) \frac{\partial}{\partial \vartheta_j} L(\zeta_1, \vartheta) \right) \right)_{\substack{1\leq i\leq k\\ 1\leq j\leq k}}.$$

In diesem Unterabschnitt betrachten wir für das zugrunde liegende anfangs erwähnte Testproblem im Fall der einfachen Hypothese, d.h. im Fall $\tilde{\Theta}_0 = \{\vartheta_0\}$, $\vartheta_0 \in \tilde{\Theta}$, die Testgröße

$$T_n = 2\Big(\sum_{i=1}^n L\big(\zeta_i, \hat{\vartheta}_n(\zeta_1, \ldots, \zeta_n)\big) - \sum_{i=1}^n L(\zeta_i, \vartheta_0) \Big), \tag{4.2.1}$$

die in typischen Anwendungsfällen die Likelihood-Quotienten-Testgröße ist, oder die Testgröße

$$T_n = \sqrt{n}\big(\hat{\vartheta}_n(\zeta_1, \ldots, \zeta_n) - \vartheta_0\big)^t i_{\zeta_1}(\vartheta_0) \sqrt{n}\big(\hat{\vartheta}_n(\zeta_1, \ldots, \zeta_n) - \vartheta_0\big), \tag{4.2.2}$$

die in typischen Anwendungsfällen die Waldsche Testgröße ist. Im Fall einer zusammengesetzten Hypothese betrachten wir die Testgröße

$$T_n = 2\Bigg(\sum_{i=1}^{n} L\big(\zeta_i, \hat{\vartheta}_n(\zeta_1, \ldots, \zeta_n)\big) - \sum_{i=1}^{n} L\Big(\zeta_i, g\big(\hat{\eta}_n(\zeta_1, \ldots, \zeta_n)\big)\Big)\Bigg), \tag{4.2.3}$$

die in typischen Anwendungsfällen die Likelihood-Quotienten-Testgröße ist, oder die Testgröße

$$\begin{aligned} T_n =& \sqrt{n}\Big(\hat{\vartheta}_n(\zeta_1, \ldots, \zeta_n) - g\big(\hat{\eta}_n(\zeta_1, \ldots, \zeta_n)\big)\Big)^t i_{\zeta_1}\Big(g\big(\hat{\eta}_n(\zeta_1, \ldots, \zeta_n)\big)\Big) \\ & \sqrt{n}\Big(\hat{\vartheta}_n(\zeta_1, \ldots, \zeta_n) - g\big(\hat{\eta}_n(\zeta_1, \ldots, \zeta_n)\big)\Big), \end{aligned} \tag{4.2.4}$$

die in typischen Anwendungsfällen die Waldsche Testgröße ist. Der zum Testniveau $\alpha \in (0, 1)$ zugehörige Test lautet dann

„Verwirf die Hypothese H, falls $T_n \geq \chi^2_{d_1-1;1-\alpha}$“,

wobei $\chi^2_{d_1-1;1-\alpha}$ das $(1-\alpha)$-Quantil der Chi-Quadrat-Verteilung mit $d_1 - 1$ Freiheitsgraden sei. Zur Vorbereitung auf den Satz werden wir nun zunächst ein paar Lemmata beweisen.

4.2.1 Lemma: *Es sei eine Folge $(\vartheta_n)_{n\in\mathbb{N}}$ mit $\vartheta_n \in \tilde{\Theta}$ für alle $n \in \mathbb{N}$ und $\lim_{n\to\infty} \vartheta_n = \vartheta \in \tilde{\Theta}$ gegeben. Wir nehmen an, dass die Abbildung $\tilde{\vartheta} \mapsto f(z;\tilde{\vartheta})$ für alle $z \in R$ auf $\tilde{\Theta}$ stetig ist. Weiter sei eine Abbildung $k : R \times \tilde{\Theta} \to \mathbb{R}$ gegeben, sodass für alle $\tilde{\vartheta} \in \tilde{\Theta}$ $k(\cdot, \tilde{\vartheta}) : (R, \mathfrak{S}) \to (\mathbb{R}, \mathfrak{B})$ gilt, und für alle $z \in R$ die Abbildung $\tilde{\vartheta} \mapsto k(z, \tilde{\vartheta})$ stetig auf $\tilde{\Theta}$ ist. Zudem existiere ein $\delta > 0$ sodass mit $U_\delta = \{\tilde{\vartheta} \in \mathbb{R}^k : |\tilde{\vartheta} - \vartheta| < \delta\} \subset \tilde{\Theta}$, $M_\delta(z) := \sup_{\tilde{\vartheta}\in U_\delta} |k(z, \tilde{\vartheta})| \sup_{\tilde{\vartheta}\in U_\delta} f(z;\tilde{\vartheta})$, $z \in R$, $\int M_\delta(z)\mathrm{d}\mu(z) < \infty$ gilt. Ferner sei für $n \in \mathbb{N}$ $\hat{\vartheta}_n : (R^n, \mathfrak{S}^n) \to (\tilde{\Theta}, \mathfrak{B}^k_{\tilde{\Theta}})$, sodass $\lim_{n\to\infty} \hat{\vartheta}_n(\zeta_1, ..., \zeta_n) = \vartheta$ in Q_{ϑ_n}-Wahrscheinlichkeit gilt. Dann ist k $(\mathfrak{S} \otimes \mathfrak{B}^k_{\tilde{\Theta}}, \mathfrak{B})$-messbar und es gilt*

$$\lim_{n\to\infty} \frac{1}{n} \sum_{i=1}^{n} k\big(\zeta_i, \hat{\vartheta}_n(\zeta_1, ..., \zeta_n)\big) = \mathrm{E}_\vartheta\left(k(\zeta_1, \vartheta)\right) \text{ in } Q_{\vartheta_n}\text{-Wahrscheinlichkeit.}$$

Beweis: Es sei $\{\vartheta_1, \vartheta_2, ...\}$ eine abzählbare dichte Teilmenge von $\tilde{\Theta}$. Weiter sei $U_{n,i} = \{\tilde{\vartheta} \in \tilde{\Theta} : |\tilde{\vartheta} - \vartheta_i| < \frac{1}{n}\}$ und $V_{n,i} = U_{n,i} \cap (\cup_{j<i} U_{n,j})^c$ für $n, i \in \mathbb{N}$. Dann ist $(V_{n,i})_{i\in\mathbb{N}}$ für alle $n \in \mathbb{N}$ eine Partition von $\tilde{\Theta}$ und die Abbildung

$$k_n(z, \tilde{\vartheta}) := \sum_{i=1}^{\infty} k(z, \vartheta_i)\, \mathrm{I}_{V_{n,i}}(\tilde{\vartheta}),\ (z, \tilde{\vartheta}) \in R \times \tilde{\Theta},$$

ist als Reihe von $(\mathfrak{S} \otimes \mathfrak{B}^k_{\tilde{\Theta}}, \mathfrak{B})$-messbaren Abbildungen $(\mathfrak{S} \otimes \mathfrak{B}^k_{\tilde{\Theta}}, \mathfrak{B})$-messbar. Da sich k nach Konstruktion als Grenzfunktion der k_n, $n \in \mathbb{N}$, ergibt, folgt somit die gewünschte

Messbarkeit. Weiter ist

$$\begin{aligned}
&\left|\frac{1}{n}\sum_{i=1}^{n} k\big(\zeta_i, \hat{\vartheta}_n(\zeta_1,...,\zeta_n)\big) - \mathrm{E}_\vartheta\left(k(\zeta_1,\vartheta)\right)\right| \\
\leq &\left|\frac{1}{n}\sum_{i=1}^{n} k\big(\zeta_i, \hat{\vartheta}_n(\zeta_1,...,\zeta_n)\big) - \frac{1}{n}\sum_{i=1}^{n} k(\zeta_i,\vartheta)\right| + \left|\frac{1}{n}\sum_{i=1}^{n} k(\zeta_i,\vartheta) - \mathrm{E}_\vartheta\left(k(\zeta_1,\vartheta)\right)\right|. \qquad (4.2.5)
\end{aligned}$$

Nach Voraussetzung existiert ein $\delta > 0$ mit $\int M_\delta(z)\mathrm{d}\mu(z) < \infty$. Wegen

$$\lim_{n\to\infty} k(z,\vartheta)f(z;\vartheta_n) = k(z,\vartheta)f(z;\vartheta) \text{ für alle } z \in R,$$

$|k(z,\vartheta)|f(z;\vartheta_n) \leq M_\delta(z)$ für alle $z \in R$ für genügend großes $n \in \mathbb{N}$ und $\int M_\delta(z)\mathrm{d}\mu(z) < \infty$ gilt mit dem Satz von der majorisierten Konvergenz

$$\lim_{n\to\infty} \mathrm{E}_{\vartheta_n}\left(|k(\zeta_1,\vartheta)|\right) = E_\vartheta\big(|k(\zeta_1,\vartheta)|\big).$$

Wegen der Stetigkeit der Abbildung $\tilde{\vartheta} \mapsto f(z;\tilde{\vartheta})$ für alle $z \in R$ auf $\tilde{\Theta}$ folgt mit dem Satz von Scheffé für alle stetigen und beschränkten Funktionen $l : \mathbb{R} \to \mathbb{R}$

$$\lim_{n\to\infty} \mathrm{E}_{\vartheta_n}\Big(l\big(k(\zeta_1,\vartheta)\big)\Big) = E_\vartheta\Big(l\big(k(\zeta_1,\vartheta)\big)\Big),$$

also $Q_{\vartheta_n}^{k(\zeta_1,\vartheta)} \overset{\vartheta_n}{\rightsquigarrow} Q_\vartheta^{k(\zeta_1,\vartheta)}$ für $n \to \infty$. Somit folgt mit dem schwachen Gesetz der großen Zahlen für Dreiecksschemata (siehe Lemma A.1.2 im Anhang), dass der zweite Term in (4.2.5) in Q_{ϑ_n}-Wahrscheinlichkeit für $n \to \infty$ gegen $0 \in \mathbb{R}$ konvergiert. Daher muss nur noch der erste Term betrachtet werden.

Es existiert nach Voraussetzung ein $\delta > 0$, sodass $U_\delta \subset \tilde{\Theta}$. Wir definieren für beliebiges $\tilde{\delta}$ mit $0 < \tilde{\delta} \leq \delta$

$$A_{\tilde{\delta}}(z) := \sup_{\tilde{\vartheta}\in U_{\tilde{\delta}}} |k(z,\tilde{\vartheta}) - k(z,\vartheta)|, \; z \in R.$$

Wegen der Stetigkeit von k gilt für alle $z \in R$ $\lim_{\tilde{\delta}\to 0} A_{\tilde{\delta}}(z) = 0$. Weiter gilt $|A_{\tilde{\delta}}(z)f(z;\vartheta)| \leq 2M_\delta(z)$ für alle $z \in R$ und $\int M_\delta(z)\mathrm{d}\mu(z) < \infty$. Daher gilt mit dem Satz von der majorisierten Konvergenz

$$\lim_{\tilde{\delta}\to 0} \mathrm{E}_\vartheta\left(A_{\tilde{\delta}}(\zeta_1)\right) = 0.$$

Wegen $\lim_{n\to\infty} A_{\tilde{\delta}}(z)f(z;\vartheta_n) = A_{\tilde{\delta}}(z)f(z;\vartheta)$ für alle $z \in R$, $|A_{\tilde{\delta}}(z)|f(z;\vartheta_n) \leq 2M_\delta(z)$ für alle $z \in R$ für genügend großes $n \in \mathbb{N}$, und $\int M_\delta(z)\mathrm{d}\mu(z) < \infty$ gilt erneut mit dem Satz von der majorisierten Konvergenz

$$\lim_{n\to\infty} \mathrm{E}_{\vartheta_n}\left(A_{\tilde{\delta}}(\zeta_1)\right) = \mathrm{E}_\vartheta\left(A_{\tilde{\delta}}(\zeta_1)\right).$$

Wegen der Stetigkeit der Abbildung $\tilde{\vartheta} \mapsto f(z; \tilde{\vartheta})$ für alle $z \in R$ auf $\tilde{\Theta}$ folgt erneut mit dem Satz von Scheffé für alle stetigen und beschränkten Funktionen $l : \mathbb{R} \to \mathbb{R}$

$$\lim_{n\to\infty} \mathrm{E}_{\vartheta_n} \Big(l\big(A_{\tilde{\delta}}(\zeta_1)\big)\Big) = E_{\vartheta}\Big(l\big(A_{\tilde{\delta}}(\zeta_1)\big)\Big),$$

also $Q_{\vartheta_n}^{A_{\tilde{\delta}}(\zeta_1)} \overset{\vartheta_n}{\rightsquigarrow} Q_{\vartheta}^{A_{\tilde{\delta}}(\zeta_1)}$ für $n \to \infty$. Damit erhält man erneut mit Hilfe des schwachen Gesetzes der großen Zahlen für Dreiecksschemata (Lemma A.1.2 im Anhang)

$$\lim_{n\to\infty} \left| \frac{1}{n} \sum_{i=1}^{n} A_{\tilde{\delta}}(\zeta_i) - \mathrm{E}_{\vartheta}\left(A_{\tilde{\delta}}(\zeta_1)\right) \right| = 0 \text{ in } Q_{\vartheta_n}\text{-Wahrscheinlichkeit.}$$

Nun sei $\varepsilon > 0$ beliebig. Dann folgt mit der Markov-Ungleichung

$$\begin{aligned}
&Q_{\vartheta_n}\left(\left| \frac{1}{n}\sum_{i=1}^{n} k\big(\zeta_i, \hat{\vartheta}_n(\zeta_1,...,\zeta_n)\big) - \frac{1}{n}\sum_{i=1}^{n} k(\zeta_i, \vartheta) \right| \geq \varepsilon \right) \\
\leq &Q_{\vartheta_n}\left(\frac{1}{n}\sum_{i=1}^{n} \left| k\big(\zeta_i, \hat{\vartheta}_n(\zeta_1,...,\zeta_n)\big) - k(\zeta_i, \vartheta) \right| \geq \varepsilon \right) \\
= &Q_{\vartheta_n}\left(\frac{1}{n}\sum_{i=1}^{n} \left| k\big(\zeta_i, \hat{\vartheta}_n(\zeta_1,...,\zeta_n)\big) - k(\zeta_i, \vartheta) \right| \geq \varepsilon, |\vartheta_n(\zeta_1,...,\zeta_n) - \vartheta| < \tilde{\delta} \right) \\
&+ Q_{\vartheta_n}\left(\frac{1}{n}\sum_{i=1}^{n} \left| k\big(\zeta_i, \hat{\vartheta}_n(\zeta_1,...,\zeta_n)\big) - k(\zeta_i, \vartheta) \right| \geq \varepsilon, |\vartheta_n(\zeta_1,...,\zeta_n) - \vartheta| \geq \tilde{\delta} \right) \\
\leq &Q_{\vartheta_n}\left(\frac{1}{n}\sum_{i=1}^{n} A_{\tilde{\delta}}(\zeta_i) \geq \varepsilon, |\vartheta_n(\zeta_1,...,\zeta_n) - \vartheta| < \tilde{\delta} \right) \\
&+ Q_{\vartheta_n}\left(\frac{1}{n}\sum_{i=1}^{n} \left| k\big(\zeta_i, \hat{\vartheta}_n(\zeta_1,...,\zeta_n)\big) - k(\zeta_i, \vartheta) \right| \geq \varepsilon, |\vartheta_n(\zeta_1,...,\zeta_n) - \vartheta| \geq \tilde{\delta} \right) \\
\leq &Q_{\vartheta_n}\left(\frac{1}{n}\sum_{i=1}^{n} A_{\tilde{\delta}}(\zeta_i) \geq \varepsilon \right) + Q_{\vartheta_n}\big(|\vartheta_n(\zeta_1,...,\zeta_n) - \vartheta| \geq \tilde{\delta} \big) \\
\leq &\frac{1}{\varepsilon}\,\mathrm{E}_{\vartheta_n}\left(A_{\tilde{\delta}}(\zeta_1)\right) + Q_{\vartheta_n}\big(|\vartheta_n(\zeta_1,...,\zeta_n) - \vartheta| \geq \tilde{\delta} \big) \longrightarrow \frac{1}{\varepsilon}\,\mathrm{E}_{\vartheta}\left(A_{\tilde{\delta}}(\zeta_1)\right) \text{ für } n \to \infty.
\end{aligned}$$

Mit $\lim_{\tilde{\delta}\to 0} \mathrm{E}_{\vartheta}(A_{\tilde{\delta}}(\zeta_1)) = 0$ folgt die Behauptung. □

Mit $\vartheta = (\vartheta_1, \ldots, \vartheta_k)^t \in \tilde{\Theta}$ und $\eta = (\eta_1, \ldots, \eta_h)^t \in \Lambda$ definieren wir die Nabla-Operatoren $\nabla_k := (\frac{\partial}{\partial\vartheta_1}, \ldots, \frac{\partial}{\partial\vartheta_k})^t$ und $\nabla_h := (\frac{\partial}{\partial\eta_1}, \ldots, \frac{\partial}{\partial\eta_h})^t$.

Das folgende Lemma liefert Aussagen über das Grenzwertverhalten der Verteilungen der Schätzer $\hat{\vartheta}_n$ und $\hat{\eta}_n$ für $n \to \infty$, wenn als Parameter $\vartheta_n \in \tilde{\Theta}$ mit $\lim_{n\to\infty} \vartheta_n = \vartheta_0 \in \tilde{\Theta}_0$ zugrunde liegt. Für den Spezialfall $\vartheta_n = \vartheta_0 \in \tilde{\Theta}_0$ für alle $n \in \mathbb{N}$ sind solche Aussagen bei einer sehr ähnlichen Gesamtkonstruktion z.B. in dem Buch von Witting und Müller-Funk ([27]) zu finden; siehe Satz 6.35 und Satz 6.50 dort.

4.2.2 Lemma: *Es gelten die Bedingungen (B1) - (B5) und es sei* $(\vartheta_n)_{n\in\mathbb{N}}$ *eine Folge mit* $\vartheta_n \in \tilde{\Theta}$ *für alle* $n \in \mathbb{N}$ *und* $\lim_{n\to\infty} \vartheta_n = \vartheta_0 \in \tilde{\Theta}_0$. *Dann ist*

$$\begin{aligned}\sqrt{n}(\hat{\eta}_n - \eta_0) =& i^*_{\zeta_1}(\eta_0)^{-1} \mathcal{J}_g(\eta_0)^t i_{\zeta_1}(\vartheta_0)\sqrt{n}(\hat{\vartheta}_n - \vartheta_0) + o_{Q_{\vartheta_n}}(1)\sqrt{n}(\hat{\vartheta}_n - \vartheta_0)\\ &+ o_{Q_{\vartheta_n}}(1)\sqrt{n}(\hat{\eta}_n - \eta_0) + o_{Q_{\vartheta_n}}(1) \text{ für } n \to \infty,\end{aligned}$$

mit

$$i^*_{\zeta_1}(\eta_0) := \mathcal{J}_g(\eta_0)^t i_{\zeta_1}(\vartheta_0)\mathcal{J}_g(\eta_0),$$

und es gilt

$$\sqrt{n}(\hat{\vartheta}_n - \vartheta_n) \overset{\vartheta_n}{\rightsquigarrow} \mathrm{N}_k\left(0, i_{\zeta_1}(\vartheta_0)^{-1}\right) \text{ für } n \to \infty.$$

Beweis: Eine Taylorentwicklung der Funktion $\vartheta \mapsto \sum_{i=1}^n \frac{\partial}{\partial\vartheta_j} L(\zeta_i, \vartheta)$, $j = 1, \ldots, k$, liefert unter Beachtung von (B4)

$$\begin{aligned}&\frac{1}{\sqrt{n}} \sum_{i=1}^n \frac{\partial}{\partial\vartheta_j} L(\zeta_i, \vartheta)_{|\vartheta=\vartheta_0} + \frac{1}{n} \sum_{i=1}^n \nabla_k^t \frac{\partial}{\partial\vartheta_j} L(\zeta_i, \vartheta)_{|\vartheta=\overline{\vartheta}_{j,n}} \sqrt{n}(\hat{\vartheta}_n - \vartheta_0)\\ =& \frac{1}{\sqrt{n}} \sum_{i=1}^n \frac{\partial}{\partial\vartheta_j} L(\zeta_i, \vartheta)_{|\vartheta=\hat{\vartheta}_n} \xrightarrow{Q_{\vartheta_n}} 0 \text{ für } n \to \infty,\ j = 1, \ldots, k,\end{aligned}$$

mit einer Zwischenstelle $\overline{\vartheta}_{j,n}$ auf der Verbindungsstrecke zwischen $\hat{\vartheta}_n$ und ϑ_0, $j = 1, \ldots, k$. Siehe Satz 168.1 in [13]. Man beachte, dass die Zwischenstelle stets messbar gewählt werden kann. Dies folgt aus Hilfssatz 6.7 in [27]. Anders geschrieben erhält man

$$\nabla_k \frac{1}{\sqrt{n}} \sum_{i=1}^n L(\zeta_i, \vartheta)_{|\vartheta=\vartheta_0} + \begin{pmatrix} \frac{1}{n}\sum_{i=1}^n \nabla_k^t \frac{\partial}{\partial\vartheta_1} L(\zeta_i, \vartheta)_{|\vartheta=\overline{\vartheta}_{1,n}} \\ \vdots \\ \frac{1}{n}\sum_{i=1}^n \nabla_k^t \frac{\partial}{\partial\vartheta_k} L(\zeta_i, \vartheta)_{|\vartheta=\overline{\vartheta}_{k,n}} \end{pmatrix} \sqrt{n}(\hat{\vartheta}_n - \vartheta_0) = o_{Q_{\vartheta_n}}(1)$$

für $n \to \infty$.

Mit Lemma 4.2.1, (B2) und (B3) erhält man

$$\lim_{n\to\infty} \begin{pmatrix} \frac{1}{n}\sum_{i=1}^n \nabla_k^t \frac{\partial}{\partial\vartheta_1} L(\zeta_i, \vartheta)_{|\vartheta=\overline{\vartheta}_{1,n}} \\ \vdots \\ \frac{1}{n}\sum_{i=1}^n \nabla_k^t \frac{\partial}{\partial\vartheta_k} L(\zeta_i, \vartheta)_{|\vartheta=\overline{\vartheta}_{k,n}} \end{pmatrix} = -i_{\zeta_1}(\vartheta_0) \text{ in } Q_{\vartheta_n}\text{-Wahrscheinlichkeit.}$$

Somit folgt einerseits

$$\nabla_k \frac{1}{\sqrt{n}} \sum_{i=1}^n L(\zeta_i, \vartheta)_{|\vartheta=\vartheta_0} = i_{\zeta_1}(\vartheta_0)\sqrt{n}(\hat{\vartheta}_n - \vartheta_0) + o_{Q_{\vartheta_n}}(1)\sqrt{n}(\hat{\vartheta}_n - \vartheta_0) + o_{Q_{\vartheta_n}}(1)$$

für $n \to \infty$.

Andererseits liefert eine entsprechende Taylorentwicklung unter Beachtung von (B5)

$$\begin{aligned}
&\frac{1}{\sqrt{n}}\sum_{i=1}^{n}\frac{\partial}{\partial\eta_j}L\big(\zeta_i,g(\eta)\big)_{|\eta=\eta_0}+\frac{1}{n}\sum_{i=1}^{n}\nabla_h^t\frac{\partial}{\partial\eta_j}L\big(\zeta_i,g(\eta)\big)_{|\eta=\overline{\eta}_{j,n}}\sqrt{n}(\hat{\eta}_n-\eta_0)\\
=&\frac{1}{\sqrt{n}}\sum_{i=1}^{n}\frac{\partial}{\partial\eta_j}L(\zeta_i,g\big(\eta)\big)_{|\eta=\hat{\eta}_n}\xrightarrow{Q_{\vartheta_n}}0 \text{ für } n\to\infty,\ j=1,\dots,h,
\end{aligned}$$

mit Zwischenstellen $\overline{\eta}_{j,n}$ auf der Verbindungsstrecke zwischen $\hat{\eta}_n$ und $\eta_0,\ j=1,\dots,h$. Anders geschrieben erhält man wegen $g(\eta_0)=\vartheta_0$

$$\mathcal{J}_g(\eta_0)^t\nabla_k\frac{1}{\sqrt{n}}\sum_{i=1}^{n}L(\zeta_i,\vartheta)_{|\vartheta=\vartheta_0}+\begin{pmatrix}\frac{1}{n}\sum_{i=1}^{n}\nabla_h^t\frac{\partial}{\partial\eta_1}L\big(\zeta_i,g(\eta)\big)_{|\eta=\overline{\eta}_{1,n}}\\ \vdots\\ \frac{1}{n}\sum_{i=1}^{n}\nabla_h^t\frac{\partial}{\partial\eta_h}L\big(\zeta_i,g(\eta)\big)_{|\eta=\overline{\eta}_{h,n}}\end{pmatrix}\sqrt{n}(\hat{\eta}_n-\eta_0)=o_{Q_{\vartheta_n}}(1)$$

für $n\to\infty$.

Wegen

$$\begin{aligned}
&\frac{\partial}{\partial\eta_i}\frac{\partial}{\partial\eta_j}L\big(x,g(\eta)\big)\\
=&\sum_{s=1}^{k}\Big(\nabla_k^t\frac{\partial}{\partial\vartheta_s}L(x,\vartheta)_{|\vartheta=g(\eta)}\frac{\partial}{\partial\eta_i}g(\eta)\frac{\partial}{\partial\eta_j}g_s(\eta)+\frac{\partial}{\partial\vartheta_s}L(x,\vartheta)_{|\vartheta=g(\eta)}\frac{\partial}{\partial\eta_i}\frac{\partial}{\partial\eta_j}g_s(\eta)\Big),\\
&x\in R,\ \eta\in\Delta,\ i,j=1,\dots,h, \qquad (4.2.6)
\end{aligned}$$

$g(\eta_0)=\vartheta_0$ und der zweimaligen stetigen Differenzierbarkeit von g erhält man mit Lemma 4.2.1 und unter Heranziehung von (B2) und (B3)

$$\lim_{n\to\infty}\begin{pmatrix}\frac{1}{n}\sum_{i=1}^{n}\nabla_h^t\frac{\partial}{\partial\eta_1}L\big(\zeta_i,g(\eta)\big)_{|\eta=\overline{\eta}_{1,n}}\\ \vdots\\ \frac{1}{n}\sum_{i=1}^{n}\nabla_h^t\frac{\partial}{\partial\eta_h}L\big(\zeta_i,g(\eta)\big)_{|\eta=\overline{\eta}_{h,n}}\end{pmatrix}=-i_{\zeta_1}^*(\eta_0) \text{ in } Q_{\vartheta_n}\text{-Wahrscheinlichkeit},$$

wobei zu beachten ist, dass $i_{\zeta_1}^*(\eta_0)=\mathcal{J}_g(\eta_0)^t i_{\zeta_1}(\vartheta_0)\mathcal{J}_g(\eta_0)$ regulär ist. Somit folgt andererseits

$$\mathcal{J}_g(\eta_0)^t\nabla_k\frac{1}{\sqrt{n}}\sum_{i=1}^{n}L(\zeta_i,\vartheta)_{|\vartheta=\vartheta_0}=i_{\zeta_1}^*(\eta_0)\sqrt{n}(\hat{\eta}_n-\eta_0)+o_{Q_{\vartheta_n}}(1)\sqrt{n}(\hat{\eta}_n-\eta_0)+o_{Q_{\vartheta_n}}(1)$$

für $n\to\infty$,

und damit insgesamt

$$\begin{aligned}
\sqrt{n}(\hat{\eta}_n-\eta_0)=&i_{\zeta_1}^*(\eta_0)^{-1}\mathcal{J}_g(\eta_0)^t i_{\zeta_1}(\vartheta_0)\sqrt{n}(\hat{\vartheta}_n-\vartheta_0)+o_{Q_{\vartheta_n}}(1)\sqrt{n}(\hat{\vartheta}_n-\vartheta_0)\\
&+o_{Q_{\vartheta_n}}(1)\sqrt{n}(\hat{\eta}_n-\eta_0)+o_{Q_{\vartheta_n}}(1) \text{ für } n\to\infty,
\end{aligned}$$

also die erste Behauptung.

Zum Beweis der zweiten Behauptung gehen wir wie folgt vor: Eine erneute Taylorentwicklung der Funktion $\vartheta \mapsto \sum_{i=1}^n \frac{\partial}{\partial \vartheta_j} L(\zeta_i, \vartheta),\ j = 1, \ldots, k$, liefert

$$\begin{aligned}
&\frac{1}{\sqrt{n}} \sum_{i=1}^n \frac{\partial}{\partial \vartheta_j} L(\zeta_i, \vartheta)_{|\vartheta=\vartheta_n} + \frac{1}{n} \sum_{i=1}^n \nabla_k^t \frac{\partial}{\partial \vartheta_j} L(\zeta_i, \vartheta)_{|\vartheta=\overline{\overline{\vartheta}}_{j,n}} \sqrt{n}(\hat{\vartheta}_n - \vartheta_n) \\
=&\frac{1}{\sqrt{n}} \sum_{i=1}^n \frac{\partial}{\partial \vartheta_j} L(\zeta_i, \vartheta)_{|\vartheta=\hat{\vartheta}_n} \xrightarrow{Q_{\vartheta_n}} 0 \text{ für } n \to \infty,\ j = 1, \ldots, k,
\end{aligned}$$

mit Zwischenstellen $\overline{\overline{\vartheta}}_{j,n}$ auf der Verbindungsstrecke zwischen $\hat{\vartheta}_n$ und $\vartheta_n,\ j = 1, \ldots, k$. Setzen wir für $n \in \mathbb{N}$

$$G_n := \begin{pmatrix} \frac{1}{n} \sum_{i=1}^n \nabla_k^t \frac{\partial}{\partial \vartheta_1} L(\zeta_i, \vartheta)_{|\vartheta=\overline{\overline{\vartheta}}_{1,n}} \\ \vdots \\ \frac{1}{n} \sum_{i=1}^n \nabla_k^t \frac{\partial}{\partial \vartheta_k} L(\zeta_i, \vartheta)_{|\vartheta=\overline{\overline{\vartheta}}_{k,n}} \end{pmatrix},$$

so erhalten wir

$$\nabla_k \frac{1}{\sqrt{n}} \sum_{i=1}^n L(\zeta_i, \vartheta)_{|\vartheta=\vartheta_n} + G_n \sqrt{n}(\hat{\vartheta}_n - \vartheta_n) = o_{Q_{\vartheta_n}}(1) \text{ für } n \to \infty.$$

Zudem folgt erneut mit Lemma 4.2.1, (B2) und (B3)

$$\lim_{n \to \infty} G_n = -i_{\zeta_1}(\vartheta_0) \text{ in } Q_{\vartheta_n}\text{-Wahrscheinlichkeit.} \tag{4.2.7}$$

Nun betrachten wir für $n \in \mathbb{N}$ den Ausdruck

$$W_n := \nabla_k \frac{1}{\sqrt{n}} \sum_{i=1}^n L(\zeta_i, \vartheta)_{|\vartheta=\vartheta_n} - i_{\zeta_1}(\vartheta_0) \sqrt{n}(\hat{\vartheta}_n - \vartheta_n),$$

der sich auch schreiben lässt in der Form

$$W_n = R_n \, \mathrm{I}(G_n \text{ regulär}) + W_n \, \mathrm{I}(G_n \text{ singulär}),$$

mit

$$\begin{aligned}
R_n :=& - i_{\zeta_1}(\vartheta_0) \tilde{G}_n^{-1} \Big(G_n \sqrt{n}(\hat{\vartheta}_n - \vartheta_n) + \nabla_k \frac{1}{\sqrt{n}} \sum_{i=1}^n L(\zeta_i, \vartheta)_{|\vartheta=\vartheta_n} \Big) \\
&+ \big(1 + i_{\zeta_1}(\vartheta_0) \tilde{G}_n^{-1}\big) \nabla_k \frac{1}{\sqrt{n}} \sum_{i=1}^n L(\zeta_i, \vartheta)_{|\vartheta=\vartheta_n},
\end{aligned}$$

wobei

$$\tilde{G}_n^{-1} := \begin{cases} G_n^{-1} & , G_n \text{ regulär} \\ -i_{\zeta_1}(\vartheta)^{-1} & , G_n \text{ singulär} \end{cases}$$

gesetzt wird. Mit (4.2.7) gilt $\lim_{n\to\infty} W_n \,\mathrm{I}(G_n \text{ singulär}) = 0$ in Q_{ϑ_n}-Wahrscheinlichkeit. Es genügt also, nur noch $\lim_{n\to\infty} R_n \,\mathrm{I}(G_n \text{ regulär}) = 0$ in Q_{ϑ_n}-Wahrscheinlichkeit sowie die entsprechende Verteilungskonvergenz für $\nabla_k \frac{1}{\sqrt{n}} \sum_{i=1}^n L(\zeta_i, \vartheta)_{|\vartheta=\vartheta_n}$ zu zeigen. Hierzu soll die Lindeberg-Bedingung verifiziert werden.

Wir haben

$$\frac{\partial^2}{\partial\vartheta_i\partial\vartheta_j} L(z,\vartheta) = \left(\frac{\partial^2}{\partial\vartheta_i\partial\vartheta_j} f(z;\vartheta)\right)\frac{1}{f(z;\vartheta)} - \left(\frac{\partial}{\partial\vartheta_i} f(z;\vartheta)\right)\left(\frac{\partial}{\partial\vartheta_j} f(z;\vartheta)\right)\frac{1}{f(z;\vartheta)^2},$$
$$1 \le i,j \le k,\ z \in R,\ \vartheta \in \tilde{\Theta}.$$

Weiter ist wegen (B3) mit dem Satz von der majorisierten Konvergenz

$$\lim_{n\to\infty} \int \frac{\partial^2}{\partial\vartheta_i\partial\vartheta_j} L(z,\vartheta)_{|\vartheta=\vartheta_n} f(z;\vartheta_n)\mathrm{d}\mu(z) = \int \frac{\partial^2}{\partial\vartheta_i\partial\vartheta_j} L(z,\vartheta)_{|\vartheta=\vartheta_0} f(z;\vartheta_0)\mathrm{d}\mu(z),$$
$$1 \le i,j \le k,$$

was wegen (B2) äquivalent ist zu

$$\begin{aligned}
&\lim_{n\to\infty} \mathrm{E}_{\vartheta_n}\left(\frac{\partial}{\partial\vartheta_i} L(\zeta_1,\vartheta)_{|\vartheta=\vartheta_n} \frac{\partial}{\partial\vartheta_j} L(\zeta_1,\vartheta)_{|\vartheta=\vartheta_n}\right) \\
=&\lim_{n\to\infty} \int \frac{\partial}{\partial\vartheta_i} f(z;\vartheta)_{|\vartheta=\vartheta_n} \frac{\partial}{\partial\vartheta_j} f(z;\vartheta)_{|\vartheta=\vartheta_n} \frac{1}{f(z;\vartheta_n)}\mathrm{d}\mu(z) \\
=&\int \frac{\partial}{\partial\vartheta_i} f(z;\vartheta)_{|\vartheta=\vartheta_0} \frac{\partial}{\partial\vartheta_j} f(z;\vartheta)_{|\vartheta=\vartheta_0} \frac{1}{f(z;\vartheta_0)}\mathrm{d}\mu(z) \\
=&\mathrm{E}_{\vartheta_0}\left(\frac{\partial}{\partial\vartheta_i} L(\zeta_1,\vartheta)_{|\vartheta=\vartheta_0} \frac{\partial}{\partial\vartheta_j} L(\zeta_1,\vartheta)_{|\vartheta=\vartheta_0}\right) \\
=&\big(i_{\zeta_1}(\vartheta_0)\big)_{i,j},\ 1 \le i,j \le k.
\end{aligned}$$

Setzen wir für $n \in \mathbb{N}$ und $a := (a_1,\ldots,a_k)^t \in \mathbb{R}^k \setminus \{0\}$ beliebig

$$U_{i,n} := a^t \nabla_k L(\zeta_i,\vartheta)_{|\vartheta=\vartheta_n},\ i = 1,\ldots,n,\ S_n := \sum_{i=1}^n U_{i,n},\ s_n^2 := \mathrm{Var}_{\vartheta_n}(S_n),$$

so haben wir mit (B2)

$$\begin{aligned}
s_n^2 &= n\mathrm{Var}_{\vartheta_n}(U_{1,n}) \\
&= n \sum_{1\le i,j\le k} a_i a_j \mathrm{Cov}_{\vartheta_n}\left(\frac{\partial}{\partial\vartheta_i} L(\zeta_1,\vartheta)_{|\vartheta=\vartheta_n} \frac{\partial}{\partial\vartheta_j} L(\zeta_1,\vartheta)_{|\vartheta=\vartheta_n}\right) \\
&= n \sum_{1\le i,j\le k} a_i a_j \,\mathrm{E}_{\vartheta_n}\left(\frac{\partial}{\partial\vartheta_i} L(\zeta_1,\vartheta)_{|\vartheta=\vartheta_n} \frac{\partial}{\partial\vartheta_j} L(\zeta_1,\vartheta)_{|\vartheta=\vartheta_n}\right) \\
&\sim n a^t i_{\zeta_1}(\vartheta_0) a \text{ für } n \to \infty,
\end{aligned}$$

wobei zu bemerken ist, dass wegen (B2) $na^t i_{\zeta_1}(\vartheta_0)a > 0$ ist.

Es gilt also

$$\begin{aligned}
&\lim_{n\to\infty} \int |a^t \nabla_k L(z,\vartheta)_{|\vartheta=\vartheta_n}|^2 f(z;\vartheta_n)\mathrm{d}\mu(z) \\
&= \lim_{n\to\infty} \int \sum_{1\le i,j\le k} a_i a_j \frac{\partial}{\partial\vartheta_i} f(z;\vartheta)_{|\vartheta=\vartheta_n} \frac{\partial}{\partial\vartheta_j} f(z;\vartheta)_{|\vartheta=\vartheta_n} \frac{1}{f(z;\vartheta_n)} \mathrm{d}\mu(z) = a^t i_{\zeta_1}(\vartheta_0) a.
\end{aligned}$$

Ferner ist wegen (B1)

$$\lim_{n\to\infty} |a^t \nabla_k L(z,\vartheta)_{|\vartheta=\vartheta_n}|^2 f(z;\vartheta_n) = |a^t \nabla_k L(z,\vartheta_0)|^2 f(z;\vartheta_0), \ \ z \in R.$$

Ebenfalls wegen (B1) gilt

$$\forall\ t > 0 : \underbrace{|a^t \nabla_k L(z,\vartheta)_{|\vartheta=\vartheta_n}|^2 f(z;\vartheta_n)}_{\longrightarrow |a^t \nabla_k L(z,\vartheta)_{|\vartheta=\vartheta_0}|^2 f(z;\vartheta_0)} \underbrace{\mathrm{I}(|a^t \nabla_d L(z,\vartheta)_{|\vartheta=\vartheta_n}| > t\sqrt{s_n^2})}_{\longrightarrow 0} \longrightarrow 0 \text{ für } n \to \infty,$$

$z \in R$.

Deshalb folgt mit Pratts Version des Satzes von der majorisierten Konvergenz (siehe Lemma A.1.1 im Anhang)

$$\begin{aligned}
\forall\ t > 0 : &\lim_{n\to\infty} \frac{1}{s_n^2} \sum_{i=1}^n \mathrm{E}_{\vartheta_n}\left(|U_{i,n}|^2\, \mathrm{I}(|U_{i,n}| > t\sqrt{s_n^2})\right) \\
&= \lim_{n\to\infty} \frac{n}{s_n^2} \mathrm{E}_{\vartheta_n}\left(|U_{1,n}|^2\, \mathrm{I}(|U_{1,n}| > t\sqrt{s_n^2})\right) = 0,
\end{aligned}$$

also die Gültigkeit der Lindeberg-Bedingung, also auch

$$\frac{S_n}{\sqrt{s_n^2}} \overset{\vartheta_n}{\rightsquigarrow} \mathrm{N}(0,1) \text{ für } n \to \infty.$$

Somit folgt

$$a^t \nabla_k \frac{1}{\sqrt{n}} \sum_{i=1}^{n} L(\zeta_i, \vartheta)_{|\vartheta=\vartheta_n} \overset{\vartheta_n}{\rightsquigarrow} \mathrm{N}\left(0, a^t i_{\zeta_1}(\vartheta_0) a\right) \text{ für } n \to \infty,$$

und damit auch mit dem Satz von Cramér-Wold

$$\nabla_k \frac{1}{\sqrt{n}} \sum_{i=1}^{n} L(\zeta_i, \vartheta)_{|\vartheta=\vartheta_n} \overset{\vartheta_n}{\rightsquigarrow} \mathrm{N}_k\left(0, i_{\zeta_1}(\vartheta_0)\right) \text{ für } n \to \infty.$$

Insgesamt folgt mit dem Lemma von Slutsky $\lim_{n\to\infty} R_n = 0$ in Q_{ϑ_n}-Wahrscheinlichkeit, sodass insgesamt $\lim_{n\to\infty} W_n = 0$ in Q_{ϑ_n}-Wahrscheinlichkeit folgt. Damit folgt die Behauptung. □

4.2.3 Lemma: *Es seien zwei Folgen von reellwertigen Zufallsvariablen $(U_m)_{m\in\mathbb{N}}$ und $(W_m)_{m\in\mathbb{N}}$ gegeben, sodass*

$$\lim_{m\to\infty} W_m = 0 \textit{ in } P\textit{-Wahrscheinlichkeit}$$

erfüllt ist. Dann gilt:

$$\begin{aligned}
&\forall\, t \in \mathbb{R}: \ \forall\, \varepsilon > 0\ \exists\, m_\varepsilon \in \mathbb{N}\ \forall\, m \geq m_\varepsilon : \\
&P(U_m + \varepsilon < t) - \varepsilon \leq P(U_m + W_m < t) \leq P(U_m - \varepsilon < t) + \varepsilon.
\end{aligned}$$

Gilt darüber hinaus für ein fest gegebenes $t' \in \mathbb{R}$ und eine Konstante $c \geq 0$

$$\lim_{m\to\infty} P(U_m + W_m < t') = c,$$

so gilt zusätzlich

$$\forall\, \varepsilon > 0\ \exists\, m_\varepsilon \in \mathbb{N}\ \forall\, m \geq m_\varepsilon : \ P(U_m < t' - \varepsilon) \leq c + 2\varepsilon,\ P(U_m < t' + \varepsilon) \geq c - 2\varepsilon.$$

Beweis: Aufgrund der geforderten Konvergenzen gilt:

$$\forall\, \varepsilon > 0\ \exists\, m_\varepsilon \in \mathbb{N}\ \forall\, m \geq m_\varepsilon : \ P(|W_m| > \varepsilon) \leq \varepsilon.$$

Deshalb gilt für $m \geq m_\varepsilon$

$$\begin{aligned}
P(U_m + W_m < t) &= P(U_m + W_m < t, |W_m| \leq \varepsilon) + P(U_m + W_m < t, |W_m| > \varepsilon) \\
&\leq P(U_m - \varepsilon < t, |W_m| \leq \varepsilon) + P(U_m + W_m < t, |W_m| > \varepsilon) \\
&\leq P(U_m - \varepsilon < t) + P(|W_m| > \varepsilon) \\
&\leq P(U_m - \varepsilon < t) + \varepsilon,
\end{aligned}$$

sowie

$$\begin{aligned}P(U_m+W_m<t) &= P(U_m+W_m<t,|W_m|\le\varepsilon)+P(U_m+W_m<t,|W_m|>\varepsilon)\\ &\ge P(U_m+\varepsilon<t,|W_m|\le\varepsilon)\\ &= P(U_m+\varepsilon<t)-P(U_m+\varepsilon<t,|W_m|>\varepsilon)\\ &\ge P(U_m+\varepsilon<t)-P(|W_m|>\varepsilon)\\ &\ge P(U_m+\varepsilon<t)-\varepsilon,\end{aligned}$$

sodass die erste Aussage folgt. Also gilt gemäß der erweiterten Annahme

$\forall\, \varepsilon>0\ \exists\, m_\varepsilon\in\mathbb{N}\ \forall\, m\ge m_\varepsilon$:
$|P(U_m+W_m<t')-c|\le\varepsilon,\ P(U_m+\varepsilon<t')-\varepsilon\le P(U_m+W_m<t')\le P(U_m-\varepsilon<t')+\varepsilon.$

Daraus folgt für $m\ge m_\varepsilon$

$$\varepsilon\ge P(U_m+W_m<t')-c\ge P(U_m+\varepsilon<t')-\varepsilon-c,$$

und

$$\varepsilon\ge c-P(U_m+W_m<t')\ge c-P(U_m-\varepsilon<t')-\varepsilon.$$

Damit erhält man für $m\ge m_\varepsilon$

$$P(U_m<t'-\varepsilon)\le c+2\varepsilon,\ P(U_m<t'+\varepsilon)\ge c-2\varepsilon,$$

also die zweite Aussage. □

In Vorbereitung auf die kommenden zentralen Resultate dieses Unterabschnitts definieren wir

$$i^{**}_{\zeta_1}(\eta_0):=\big(\mathrm{I}_k-\mathcal{J}_g(\eta_0)i^{*}_{\zeta_1}(\eta_0)^{-1}\mathcal{J}_g(\eta_0)^t i_{\zeta_1}(\vartheta_0)\big)^t i_{\zeta_1}(\vartheta_0),$$

wobei $i^{*}_{\zeta_1}(\eta_0)$ wie in Lemma 4.2.2 definiert ist. Zudem sei die Matrix $A^{1,1}_{\zeta_1}(\vartheta_0)\in\mathbb{R}^{d_1\times d_1}$ durch

$$i^{**}_{\zeta_1}(\eta_0)=\begin{pmatrix}A^{1,1}_{\zeta_1}(\vartheta_0) & A^{1,2}_{\zeta_1}(\vartheta_0)\\ A^{2,1}_{\zeta_1}(\vartheta_0) & A^{2,2}_{\zeta_1}(\vartheta_0)\end{pmatrix}$$

festgelegt, $A^{1,2}_{\zeta_1}(\vartheta_0)\in\mathbb{R}^{d_1\times(k-d_1)}$, $A^{2,1}_{\zeta_1}(\vartheta_0)\in\mathbb{R}^{(k-d_1)\times d_1}$, $A^{2,2}_{\zeta_1}(\vartheta_0)\in\mathbb{R}^{(k-d_1)\times(k-d_1)}$.

Es sei für $m\in\mathbb{N}$ und $b\ge 0$ $G_{b,m}$ die Verteilungsfunktion der $\chi^2_m(b)$-Verteilung, und $\chi^2_{m;1-\alpha}$ das $(1-\alpha)$-Quantil der χ^2_m-Verteilung.

Im Fall, dass bei einem Test eine einfache Hypothese vorliegt, setzen wir der Kompaktheit der Notation wegen $\mathcal{J}_g(\eta_0)=0$ und $h=0$.

4.2.4 Satz: *Es sei die Folge von Testgrößen T gegeben durch (4.2.1), (4.2.2), (4.2.3) oder (4.2.4). Für die Verteilung von ζ_1 sowie für die zugehörigen Schätzfunktionen gelten die Bedingungen (B1) - (B5). Es sei eine reellwertige Folge $(\nu_i)_{i\in\mathbb{N}}$ mit $\nu_i \neq 0$ für alle $i \in \mathbb{N}$ und $\lim_{i\to\infty} \nu_i = 0$ gegeben. Für $i \in \mathbb{N}$ betrachten wir Alternativen der Form*

$$\vartheta^i := \begin{pmatrix} \vartheta_0^1 \\ \vartheta_2 \\ \vartheta_2' \end{pmatrix} + \nu_i \begin{pmatrix} \vartheta_1 \\ 0 \\ 0 \end{pmatrix}$$

mit $\vartheta_1 \in \mathbb{R}^{d_1} \setminus \{0\}$, $(\vartheta_0^1, \vartheta_2, \vartheta_2') \in \partial\Theta_0$, die $(\vartheta_0^1 + \nu_i\vartheta_1, \vartheta_2, \vartheta_2') \in \Theta_1$ erfüllen, wobei mit $\vartheta_0 := \left(\begin{smallmatrix} \vartheta_0^1 \\ \vartheta_2 \end{smallmatrix}\right)$

$$\vartheta_1^t A_{\zeta_1}^{1,1}(\vartheta_0)\vartheta_1 \neq 0$$

gelte. Zudem sei die Abbildung

$$i \mapsto N_{T,i}, \;\; i \in \mathbb{N},$$

streng monoton wachsend[5], wobei $N_{T,i}$ bei zugrunde liegender Folge von Alternativen $(\vartheta^i)_{i\in\mathbb{N}}$ und zugrunde liegendem $\beta \in (0, 1-\alpha)$ wie in (2.4.3) definiert sei. Dann gilt mit $\vartheta^0 := (\vartheta_0^1, \vartheta_2, \vartheta_2')$

$$\lim_{i\to\infty} (N_{T,i}\nu_i^2) = \frac{b(\alpha, \beta, k-h)}{\vartheta_1^t A_{\zeta_1}^{1,1}(\vartheta_0)\vartheta_1},$$

wobei für $m \in \mathbb{N}$ die positive reelle Zahl $b(\alpha, \beta, m) \in (0, \infty)$ eindeutig durch die Eigenschaft

$$G_{b(\alpha,\beta,m),m}(\chi^2_{m;1-\alpha}) = \beta$$

festgelegt ist.

Beweis: Der Beweis wird nur für den Fall geführt, dass es sich bei den Tests um Tests für eine zusammengesetzte Hypothese handelt; das Vorgehen bei einfacher Hypothese ist dann klar. Zudem liegen die Familien von Verteilungen von ζ_1 und ζ_1' in der Form $\{Q_\vartheta^{\zeta_1}; \vartheta \in \tilde{\Theta}\}$ und $\{M_{\vartheta'}^{\zeta_1}; \vartheta' \in \tilde{\Theta}'\}$ vor.

Zunächst handle es sich bei den Tests T um die Tests (4.2.3). Es sei $\vartheta_0 = \left(\begin{smallmatrix} \vartheta_0^1 \\ \vartheta_2 \end{smallmatrix}\right)$, ϑ_0^1 und ϑ_2 wie in der Voraussetzung.

Eine Taylorentwicklung der Funktion $\vartheta \mapsto \sum_{i=1}^n L(\zeta_i, \vartheta)$ liefert für eine beliebige Folge

[5] Die Existenz von Sequenzen von Alternativen mit dieser Eigenschaft wird in Bemerkung 4.2.7 behandelt.

$(\tilde{\vartheta}_n)_{n\in\mathbb{N}}$ mit $\tilde{\vartheta}_n \in \tilde{\Theta}$ für alle $n \in \mathbb{N}$ und $\lim_{n\to\infty} \tilde{\vartheta}_n = \vartheta_0$

$$\begin{aligned}\sum_{j=1}^{n} L(\zeta_j, \vartheta_0) =& \sum_{j=1}^{n} L(\zeta_j, \hat{\vartheta}_n) + \sqrt{n}(\vartheta_0 - \hat{\vartheta}_n)^t \frac{1}{\sqrt{n}} \sum_{j=1}^{n} \nabla_k L(\zeta_j, \vartheta)_{|_{\vartheta=\hat{\vartheta}_n}} \\ &+ \frac{1}{2}\sqrt{n}(\vartheta_0 - \hat{\vartheta}_n)^t \left(\frac{1}{n} \sum_{j=1}^{n} \frac{\partial^2}{\partial\vartheta_i \partial\vartheta_l} L(\zeta_j, \vartheta)_{|_{\vartheta=\overline{\vartheta}_n}} \right)_{\substack{1\leq i\leq k\\ 1\leq l\leq k}} \sqrt{n}(\vartheta_0 - \hat{\vartheta}_n),\end{aligned}$$

mit einer Zwischenstelle $\overline{\vartheta}_n$ auf der Verbindungsstrecke zwischen $\hat{\vartheta}_n$ und ϑ_0. Mit Lemma 4.2.1, (B2) und (B3) gilt

$$\lim_{n\to\infty} \left(\frac{1}{n} \sum_{j=1}^{n} \frac{\partial^2}{\partial\vartheta_i \partial\vartheta_l} L(\zeta_j, \vartheta)_{|_{\vartheta=\overline{\vartheta}_n}} \right)_{\substack{1\leq i\leq k\\ 1\leq l\leq k}} = -i_{\zeta_1}(\vartheta_0) \text{ in } Q_{\tilde{\vartheta}_n}\text{-Wahrscheinlichkeit.}$$

Damit ergibt sich

$$\begin{aligned}\sum_{j=1}^{n} L(\zeta_j, \hat{\vartheta}_n) - \sum_{j=1}^{n} L(\zeta_j, \vartheta_0) =& \sqrt{n}(\hat{\vartheta}_n - \vartheta_0)^t o_{Q_{\tilde{\vartheta}_n}}(1) + \frac{1}{2}\sqrt{n}(\hat{\vartheta}_n - \vartheta_0)^t i_{\zeta_1}(\vartheta_0)\sqrt{n}(\hat{\vartheta}_n - \vartheta_0) \\ &+ \sqrt{n}(\hat{\vartheta}_n - \vartheta_0)^t o_{Q_{\tilde{\vartheta}_n}}(1)\sqrt{n}(\hat{\vartheta}_n - \vartheta_0) \text{ für } n \to \infty.\end{aligned}$$

Andererseits liefert eine entsprechende Taylorentwicklung

$$\begin{aligned}\sum_{j=1}^{n} L\big(\zeta_j, g(\eta_0)\big) =& \sum_{j=1}^{n} L\big(\zeta_j, g(\hat{\eta}_n)\big) + \sqrt{n}(\eta_0 - \hat{\eta}_n)^t \frac{1}{\sqrt{n}} \sum_{j=1}^{n} \nabla_h L\big(\zeta_j, g(\eta)\big)_{|_{\eta=\hat{\eta}_n}} \\ &+ \frac{1}{2}\sqrt{n}(\eta_0 - \hat{\eta}_n)^t \left(\frac{1}{n} \sum_{j=1}^{n} \frac{\partial^2}{\partial\eta_i \partial\eta_l} L\big(\zeta_j, g(\eta)\big)_{|_{\eta=\overline{\eta}_n}} \right)_{\substack{1\leq i\leq h\\ 1\leq l\leq h}} \sqrt{n}(\eta_0 - \hat{\eta}_n),\end{aligned}$$

mit einer Zwischenstelle $\overline{\eta}_n$ auf der Verbindungsstrecke zwischen $\hat{\eta}_n$ und η_0. Wegen (4.2.6), $g(\eta_0) = \vartheta_0$ und der zweimaligen stetigen Differenzierbarkeit von g erhält man mit Lemma 4.2.1 und unter Heranziehung von (B2) und (B3)

$$\lim_{n\to\infty} \left(\frac{1}{n} \sum_{j=1}^{n} \frac{\partial^2}{\partial\eta_i \partial\eta_l} L\big(\zeta_j, g(\eta)\big)_{|_{\eta=\overline{\eta}_n}} \right)_{\substack{1\leq i\leq h\\ 1\leq l\leq h}} = -i^*_{\zeta_1}(\eta_0) \text{ in } Q_{\tilde{\vartheta}_n}\text{-Wahrscheinlichkeit,}$$

wobei zu beachten ist, dass $i^*_{\zeta_1}(\eta_0) = \mathcal{J}_g(\eta_0)^t i_{\zeta_1}(\vartheta_0)\mathcal{J}_g(\eta_0)$ regulär ist. Wegen (B4) und $g(\eta_0) = \vartheta_0$ lässt sich dies auch in der Form

$$\begin{aligned}\sum_{j=1}^{n} L\big(\zeta_j, g(\hat{\eta}_n)\big) - \sum_{j=1}^{n} L(\zeta_j, \vartheta_0) =& \sqrt{n}(\hat{\eta}_n - \eta_0)^t o_{Q_{\tilde{\vartheta}_n}}(1) + \frac{1}{2}\sqrt{n}(\hat{\eta}_n - \eta_0)^t i^*_{\zeta_1}(\eta_0)\sqrt{n}(\hat{\eta}_n - \eta_0) \\ &+ \sqrt{n}(\hat{\eta}_n - \eta_0)^t o_{Q_{\tilde{\vartheta}_n}}(1)\sqrt{n}(\hat{\eta}_n - \eta_0) \text{ für } n \to \infty\end{aligned}$$

schreiben.

Insgesamt erhalten wir damit für unsere Testgröße die Darstellung

$$
\begin{aligned}
T_n =& 2\Big(\sum_{i=1}^n L(\zeta_i, \hat{\vartheta}_n) - \sum_{i=1}^n L\big(\zeta_i, g(\hat{\eta}_n)\big)\Big) \\
=& 2\Big(\sum_{i=1}^n L(\zeta_i, \hat{\vartheta}_n) - \sum_{i=1}^n L(\zeta_i, \vartheta_0)\Big) - 2\Big(\sum_{i=1}^n L\big(\zeta_i, g(\hat{\eta}_n)\big) - \sum_{i=1}^n L(\zeta_i, \vartheta_0)\Big) \\
=& \sqrt{n}(\hat{\vartheta}_n - \vartheta_0)^t i_{\zeta_1}(\vartheta_0)\sqrt{n}(\hat{\vartheta}_n - \vartheta_0) - \sqrt{n}(\hat{\eta}_n - \eta_0)^t i^*_{\zeta_1}(\eta_0)\sqrt{n}(\hat{\eta}_n - \eta_0) \\
& + \sqrt{n}(\hat{\vartheta}_n - \vartheta_0)^t o_{Q_{\tilde{\vartheta}_n}}(1) - \sqrt{n}(\hat{\eta}_n - \eta_0)^t o_{Q_{\tilde{\vartheta}_n}}(1) \\
& + \sqrt{n}(\hat{\vartheta}_n - \vartheta_0)^t o_{Q_{\tilde{\vartheta}_n}}(1)\sqrt{n}(\hat{\vartheta}_n - \vartheta_0) - \sqrt{n}(\hat{\eta}_n - \eta_0)^t o_{Q_{\tilde{\vartheta}_n}}(1)\sqrt{n}(\hat{\eta}_n - \eta_0) \text{ für } n \to \infty.
\end{aligned}
$$

Mit Hilfe von Lemma 4.2.2 lässt sich dies schreiben als

$$T_n = \sqrt{n}(\hat{\vartheta}_n - \vartheta_0)^t i^{**}_{\zeta_1}(\eta_0)\sqrt{n}(\hat{\vartheta}_n - \vartheta_0) + R_n,$$

mit

$$
\begin{aligned}
R_n :=& -\Big(i^*_{\zeta_1}(\eta_0)^{-1}\mathcal{J}_g(\eta_0)^t i_{\zeta_1}(\vartheta_0)\sqrt{n}(\hat{\vartheta}_n - \vartheta_0)\Big)^t i^*_{\zeta_1}(\eta_0) \\
& \Big(o_{Q_{\tilde{\vartheta}_n}}(1)\sqrt{n}(\hat{\vartheta}_n - \vartheta_0) + o_{Q_{\tilde{\vartheta}_n}}(1)\sqrt{n}(\hat{\eta}_n - \eta_0) + o_{Q_{\tilde{\vartheta}_n}}(1)\Big) \\
& - \Big(o_{Q_{\tilde{\vartheta}_n}}(1)\sqrt{n}(\hat{\vartheta}_n - \vartheta_0) + o_{Q_{\tilde{\vartheta}_n}}(1)\sqrt{n}(\hat{\eta}_n - \eta_0) + o_{Q_{\tilde{\vartheta}_n}}(1)\Big)^t i^*_{\zeta_1}(\eta_0) \\
& \Big(i^*_{\zeta_1}(\eta_0)^{-1}\mathcal{J}_g(\eta_0)^t i_{\zeta_1}(\vartheta_0)\sqrt{n}(\hat{\vartheta}_n - \vartheta_0) + o_{Q_{\tilde{\vartheta}_n}}(1)\sqrt{n}(\hat{\vartheta}_n - \vartheta_0) \\
& + o_{Q_{\tilde{\vartheta}_n}}(1)\sqrt{n}(\hat{\eta}_n - \eta_0) + o_{Q_{\tilde{\vartheta}_n}}(1)\Big) \\
& + \sqrt{n}(\hat{\vartheta}_n - \vartheta_0)^t o_{Q_{\tilde{\vartheta}_n}}(1) - \sqrt{n}(\hat{\eta}_n - \eta_0)^t o_{Q_{\tilde{\vartheta}_n}}(1) \\
& + \sqrt{n}(\hat{\vartheta}_n - \vartheta_0)^t o_{Q_{\tilde{\vartheta}_n}}(1)\sqrt{n}(\hat{\vartheta}_n - \vartheta_0) - \sqrt{n}(\hat{\eta}_n - \eta_0)^t o_{Q_{\tilde{\vartheta}_n}}(1)\sqrt{n}(\hat{\eta}_n - \eta_0) \text{ für } n \to \infty.
\end{aligned}
$$

Nun sei die Folge von Alternativen $(\vartheta_i)_{i\in\mathbb{N}}$ durch $\vartheta_i := \left(\begin{smallmatrix}\vartheta_0^1 \\ \vartheta_2\end{smallmatrix}\right) + \nu_i\left(\begin{smallmatrix}\vartheta_1 \\ 0\end{smallmatrix}\right)$ mit $(\vartheta_0^1 + \nu_i\vartheta_1, \vartheta_2) \in \tilde{\Theta}_1$ gegeben, wobei ϑ_0^1, ϑ_1 und ϑ_2 und ν_i wie in der Voraussetzung gegeben seien.

An dieser Stelle soll darauf hingewiesen werden, dass der zweite Teil von Lemma 4.2.2 für $\sqrt{N_{T,i}}(\hat{\vartheta}_{N_{T,i}} - \vartheta_i)$, $i \in \mathbb{N}$ Gültigkeit hat, und daher

$$\sqrt{N_{T,i}}(\hat{\vartheta}_{N_{T,i}} - \vartheta_i) \overset{\vartheta_i}{\rightsquigarrow} \mathrm{N}_k\left(0, i_{\zeta_1}(\vartheta_0)^{-1}\right) \text{ für } i \to \infty.$$

gilt. Um dies zu sehen, definiere eine Folge $(\vartheta_i^*)_{i\in\mathbb{N}}$ durch

$$\begin{aligned}\vartheta_1^*,\dots,\vartheta_{N_{T,1}}^* &:= \vartheta_1,\\ \vartheta_{N_{T,1}+1}^*,\dots,\vartheta_{N_{T,2}}^* &:= \vartheta_2,\\ \vartheta_{N_{T,2}+1}^*,\dots,\vartheta_{N_{T,3}}^* &:= \vartheta_3,\\ &\vdots\end{aligned}$$

Die Definition der Folge $(\vartheta_i^*)_{i\in\mathbb{N}}$ in dieser Form ist möglich, da nach Annahme die Abbildung $i \mapsto N_{T,i}$ streng monoton wachsend ist.

Dann ist $\vartheta_i^* \in \tilde{\Theta}_1$, $i \in \mathbb{N}$, $\lim_{i\to\infty} \vartheta_i^* = \vartheta_0$ und $\vartheta_{N_{T,i}}^* = \vartheta_i$, $i \in \mathbb{N}$. Also gilt

$$\sqrt{N_{T,i}}(\hat{\vartheta}_{N_{T,i}} - \vartheta_i) = \sqrt{N_{T,i}}(\hat{\vartheta}_{N_{T,i}} - \vartheta_{N_{T,i}}^*),\ i \in \mathbb{N},$$

wobei $\vartheta_{N_{T,i}}^*$ der zugrunde liegende Parameter ist, $i \in \mathbb{N}$. Die Folge $\sqrt{N_{T,i}}(\hat{\vartheta}_{N_{T,i}} - \vartheta_{N_{T,i}}^*)$, $i \in \mathbb{N}$, kann als eine Teilfolge der Folge $\sqrt{n}(\hat{\vartheta}_n - \vartheta_n^*)$, $n \in \mathbb{N}$, aufgefasst werden, wobei ϑ_n^* der zugrunde liegende Parameter ist. Ferner gilt $\lim_{\to\infty} N_{T,i} = \infty$, siehe (2.4.5).

Analog kann argumentiert werden, um zu zeigen, dass der erste Teil von Lemma 4.2.2 für $\sqrt{N_{T,i}}(\hat{\eta}_{N_{T,i}} - \eta_0)$, $i \in \mathbb{N}$, Gültigkeit hat, und daher

$$\begin{aligned}\sqrt{N}_{T,i}(\hat{\eta}_{N_{T,i}} - \eta_0) =& i_{\zeta_1}^*(\eta_0)^{-1}\mathcal{J}_g(\eta_0)^t i_{\zeta_1}(\vartheta_0)\sqrt{N_{T,i}}(\hat{\vartheta}_{N_{T,i}} - \vartheta_0) + o_{Q_{\vartheta_i}}(1)\sqrt{N_{T,i}}(\hat{\vartheta}_{N_{T,i}} - \vartheta_0)\\ &+ o_{Q_{\vartheta_i}}(1)\sqrt{N_{T,i}}(\hat{\eta}_{N_{T,i}} - \eta_0) + o_{Q_{\vartheta_i}}(1) \text{ für } i \to \infty,\end{aligned}$$

gilt.

Aus der oben genannten Gestalt von R_n wird ersichtlich, dass

$$\limsup_{i\to\infty} \sqrt{N_{T,i}}|\nu_i| < \infty$$

gelten muss.

Dies sieht man wie folgt: Angenommen dies würde nicht gelten. Dann existiert eine Teilfolge $(\sqrt{N_{T,i_k}}|\nu_{i_k}|)_{k\in\mathbb{N}}$ von $(\sqrt{N_{T,i}}|\nu_i|)_{i\in\mathbb{N}}$ mit $\lim_{k\to\infty} \sqrt{N_{T,i_k}}|\nu_{i_k}| = +\infty$. Der Einfachheit halber sei diese Teilfolge die Folge selbst. Es sei also ϑ_i der zugrunde liegende Parameter. Dann ist wegen $\lim_{i\to\infty} N_{T,i} = +\infty$ (siehe (2.4.5)), mit $\vartheta := \binom{\vartheta_1}{0}$, ϑ_1 wie in der Voraussetzung,

$$\begin{aligned}T_{N_{T,i}} =& \sqrt{N_{T,i}}(\hat{\vartheta}_{N_{T,i}} - \vartheta_0)^t i_{\zeta_1}^{**}(\eta_0)\sqrt{N_{T,i}}(\hat{\vartheta}_{N_{T,i}} - \vartheta_0) + R_{N_{T,i}}\\ =& \sqrt{N_{T,i}}(\hat{\vartheta}_{N_{T,i}} - \vartheta_i)^t i_{\zeta_1}^{**}(\eta_0)\sqrt{N_{T,i}}(\hat{\vartheta}_{N_{T,i}} - \vartheta_i)\\ &+ \sqrt{N_{T,i}}\nu_i\sqrt{N_{T,i}}(\hat{\vartheta}_{N_{T,i}} - \vartheta_i)^t i_{\zeta_1}^{**}(\eta_0)\vartheta + \sqrt{N_{T,i}}\nu_i\vartheta^t i_{\zeta_1}^{**}(\eta_0)\sqrt{N_{T,i}}(\hat{\vartheta}_{N_{T,i}} - \vartheta_i)\\ &+ (\sqrt{N_{T,i}}\nu_i)^2\vartheta^t i_{\zeta_1}^{**}(\eta_0)\vartheta + R_{N_{T,i}}.\end{aligned}$$

Da $i^{**}_{\zeta_1}$ positiv definit ist, gilt nach Voraussetzung

$$\vartheta^t i^{**}_{\zeta_1}(\eta_0)\vartheta > 0,$$

und somit dominiert der Term $(\sqrt{N_{T,i}}\nu_i)^2\vartheta^t i^{**}_{\zeta_1}(\eta_0)\vartheta$, der seinerseits gegen $+\infty$ strebt, aufgrund von Lemma 4.2.2 das Grenzverhalten der Testgröße. Es sei die Folge von kritischen Werten $(c_n)_{n\in\mathbb{N}}$ gemäß (2.4.1) und (2.4.2) gegeben, und es sei für $m \in \mathbb{N}$ $\chi^2_{m;1-\alpha}$ das $(1-\alpha)$-Quantil der χ^2_m-Verteilung. Da die Folge von Testgrößen T bei Gültigkeit der Hypothese in Verteilung gegen eine zentrale Chi-Quadrat-Verteilung mit $k-h$ Freiheitsgeraden konvergiert (siehe hierzu z.B. Satz 6.50 in [27]), und deshalb die zugehörige Verteilungsfunktion auf $\mathbb{R}$ auch gleichmäßig gegen die Verteilungsfunktion der zentralen Chi-Quadrat-Verteilung mit $k-h$ Freiheitsgeraden konvergiert, die auf $(0,\infty)$ stetig und streng monoton wachsend ist, muss gemäß (2.4.2) für die Folge von kritischen Werten $(c_n)_{n\in\mathbb{N}}$

$$\lim_{i\to\infty} c_{N_{T,i}} = \chi^2_{k-h;1-\alpha}$$

gelten, da $\lim_{i\to\infty} N_{T,i} = \infty$. Somit folgt mit $Q^{\zeta_1}_{\vartheta_i} = P^{\zeta_1}_{\vartheta^i}$ wegen (2.4.6)

$$Q_{\vartheta_i}(T_{N_{T,i}} < c_{N_{T,i}}) = \underbrace{P_{\vartheta^i}(T_{N_{T,i}} < c_{N_{T,i}})}_{\to\beta} \longrightarrow 0 \text{ für } i \to \infty,$$

mit einem $\beta \in (0, 1-\alpha)$ aus der Definition von $N_{T,i}$, siehe (2.4.3), und man erhält einen Widerspruch. Somit folgt mit Lemma 4.2.2 auch

$$\lim_{i\to\infty} R_{N_{T,i}} = 0 \text{ in } Q_{\vartheta_i}\text{-Wahrscheinlichkeit.}$$

Sei nun $\varepsilon > 0$ beliebig. Dann folgt wegen $\lim_{i\to\infty} Q_{\vartheta_i}(T_{N_{T,i}} < c_{N_{T,i}}) = \beta$, was wie gesagt wegen $Q^{\zeta_1}_{\vartheta_i} = P^{\zeta_1}_{\vartheta^i}$ und (2.4.6) gilt, mit dem zweiten Teil von Lemma 4.2.3 für genügend großes $i \in \mathbb{N}$

$$\begin{aligned} Q_{\vartheta_i}\left(\sqrt{N_{T,i}}(\hat\vartheta_{N_{T,i}} - \vartheta_0)^t i^{**}_{\zeta_1}(\eta_0)\sqrt{N_{T,i}}(\hat\vartheta_{N_{T,i}} - \vartheta_0) < \chi^2_{k-h;1-\alpha} - \varepsilon\right) &\le \beta + 2\varepsilon, \\ Q_{\vartheta_i}\left(\sqrt{N_{T,i}}(\hat\vartheta_{N_{T,i}} - \vartheta_0)^t i^{**}_{\zeta_1}(\eta_0)\sqrt{N_{T,i}}(\hat\vartheta_{N_{T,i}} - \vartheta_0) < \chi^2_{k-h;1-\alpha} + \varepsilon\right) &\ge \beta - 2\varepsilon. \end{aligned}$$

Es existiert wegen Lemma 4.2.2 eine Folge von $\mathbb{R}^k$-wertigen Zufallsvektoren $(U_i)_{i\in\mathbb{N}}$ und ein $\mathbb{R}^k$-wertiger Zufallsvektor U, sodass

$$U_i \overset{\text{v}}{=} \sqrt{N_{T,i}}(\hat\vartheta_{N_{T,i}} - \vartheta_i),\ i \in \mathbb{N},$$

sowie

$$\lim_{i\to\infty} U_i = U \sim \mathrm{N}_k\left(0, i_{\zeta_1}(\vartheta_0)^{-1}\right) \text{ f.s.}$$

gilt. Dann folgt

$$\begin{aligned}
&\sqrt{N_{T,i}}(\hat{\vartheta}_{N_{T,i}}-\vartheta_0)^t i^{**}_{\zeta_1}(\eta_0)\sqrt{N_{T,i}}(\hat{\vartheta}_{N_{T,i}}-\vartheta_0)\\
=&\big(\sqrt{N_{T,i}}(\hat{\vartheta}_{N_{T,i}}-\vartheta_i)^t+\sqrt{N_{T,i}}\nu_i\vartheta^t\big)i^{**}_{\zeta_1}(\eta_0)\big(\sqrt{N_{T,i}}(\hat{\vartheta}_{N_{T,i}}-\vartheta_i)+\sqrt{N_{T,i}}\nu_i\vartheta\big)\\
\overset{\mathrm{v}}{=}&\big(U_i^t+\sqrt{N_{T,i}}\nu_i\vartheta^t\big)i^{**}_{\zeta_1}(\eta_0)\big(U_i+\sqrt{N_{T,i}}\nu_i\vartheta\big),\ i\in\mathbb{N},
\end{aligned}$$

und wegen $\limsup_{i\to\infty}\sqrt{N_{T,i}}|\nu_i|<\infty$

$$\begin{aligned}
&\big(U_i^t+\sqrt{N_{T,i}}\nu_i\vartheta^t\big)i^{**}_{\zeta_1}(\eta_0)\big(U_i+\sqrt{N_{T,i}}\nu_i\vartheta\big)\\
=&\big(U^t+\sqrt{N_{T,i}}\nu_i\vartheta^t\big)i^{**}_{\zeta_1}(\eta_0)\big(U+\sqrt{N_{T,i}}\nu_i\vartheta\big)+o_P(1)\text{ für } i\to\infty.
\end{aligned}$$

Also gilt mit dem ersten Teil von Lemma 4.2.3 für genügend großes $i\in\mathbb{N}$

$$\begin{aligned}
&P\Big(\big(U^t+\sqrt{N_{T,i}}\nu_i\vartheta^t\big)i^{**}_{\zeta_1}(\eta_0)\big(U+\sqrt{N_{T,i}}\nu_i\vartheta\big)+\varepsilon<\chi^2_{k-h;1-\alpha}-\varepsilon\Big)-\varepsilon\\
\le&P\big(\sqrt{N_{T,i}}(\hat{\vartheta}_{N_{T,i}}-\vartheta_0)^t i^{**}_{\zeta_1}(\eta_0)\sqrt{N_{T,i}}(\hat{\vartheta}_{N_{T,i}}-\vartheta_0)<\chi^2_{k-h;1-\alpha}-\varepsilon\big)\le\beta+2\varepsilon,
\end{aligned}$$

und auch

$$\begin{aligned}
&P\Big(\big(U^t+\sqrt{N_{T,i}}\nu_i\vartheta^t\big)i^{**}_{\zeta_1}(\eta_0)\big(U+\sqrt{N_{T,i}}\nu_i\vartheta\big)-\varepsilon<\chi^2_{k-h;1-\alpha}+\varepsilon\Big)+\varepsilon\\
\ge&P\big(\sqrt{N_{T,i}}(\hat{\vartheta}_{N_{T,i}}-\vartheta_0)^t i^{**}_{\zeta_1}(\eta_0)\sqrt{N_{T,i}}(\hat{\vartheta}_{N_{T,i}}-\vartheta_0)<\chi^2_{k-h;1-\alpha}+\varepsilon\big)\ge\beta-2\varepsilon.
\end{aligned}$$

Es ist für $m\in\mathbb{N}$ für jedes $x>0$ die Abbildung

$$b\longmapsto G_{b,m}(x),\ b\ge 0,$$

stetig und streng monoton fallend, mit

$$\lim_{b\to\infty}G_{b,m}(x)=0,$$

was aus Satz 2.36 in [41] folgt. Natürlich ist für $b\ge 0$ und $m\in\mathbb{N}$ auch die Abbildung

$$x\longmapsto G_{b,m}(x),\ x\ge 0,$$

stetig. Da $i^{**}_{\zeta_1}(\eta_0)i^{-1}_{\zeta_1}(\vartheta_0)i^{**}_{\zeta_1}(\eta_0)=i^{**}_{\zeta_1}(\eta_0)$ erfüllt ist, gilt mit Theorem 1.4.5 in [28]

$$\big(U^t+\sqrt{N_{T,i}}\nu_i\vartheta^t\big)i^{**}_{\zeta_1}(\eta_0)\big(U+\sqrt{N_{T,i}}\nu_i\vartheta\big)\sim\chi^2_{k-h}(b_i),$$

mit

$$b_i:=(\sqrt{N_{T,i}}\nu_i)^2\vartheta^t i^{**}_{\zeta_1}(\eta_0)\vartheta$$

für $i\in\mathbb{N}$. Deshalb gilt

$$\begin{aligned}
\beta+3\varepsilon\ge\limsup_{i\to\infty}G_{b_i,k-h}(\chi^2_{k-h;1-\alpha}-2\varepsilon)&=\lim_{i\to\infty}\lim_{l\to\infty}\sup_{l\ge j\ge i}G_{b_j,k-h}(\chi^2_{k-h;1-\alpha}-2\varepsilon)\\
&=\lim_{i\to\infty}\lim_{l\to\infty}G_{\inf_{l\ge j\ge i}b_j,k-h}(\chi^2_{k-h;1-\alpha}-2\varepsilon)=G_{\liminf_{i\to\infty}b_i,k-h}(\chi^2_{k-h;1-\alpha}-2\varepsilon),
\end{aligned}$$

und auch

$$\begin{aligned}\beta - 3\varepsilon &\leq \liminf_{i\to\infty} G_{b_i,k-h}(\chi^2_{k-h;1-\alpha} + 2\varepsilon) = \lim_{i\to\infty}\lim_{l\to\infty}\inf_{l\geq j\geq i} G_{b_j,k-h}(\chi^2_{k-h;1-\alpha} + 2\varepsilon)\\ &= \lim_{i\to\infty}\lim_{l\to\infty} G_{\sup_{l\geq j\geq i} b_j,k-h}(\chi^2_{k-h;1-\alpha} + 2\varepsilon) = G_{\limsup_{i\to\infty} b_i,k-h}(\chi^2_{k-h;1-\alpha} + 2\varepsilon),\end{aligned}$$

wobei wegen $\limsup_{i\to\infty}\sqrt{N_{T,i}}|\nu_i| < \infty$ auch $0 \leq \liminf_{i\to\infty} b_i \leq \limsup_{i\to\infty} b_i < \infty$ gilt. Für $\varepsilon \to 0$ folgt daraus

$$\beta \leq G_{\limsup_{i\to\infty} b_i,k-h}(\chi^2_{k-h;1-\alpha}) \leq G_{\liminf_{i\to\infty} b_i,k-h}(\chi^2_{k-h;1-\alpha}) \leq \beta,$$

und damit

$$\limsup_{i\to\infty} b_i = \liminf_{i\to\infty} b_i = \lim_{i\to\infty} b_i = b(\alpha,\beta,k-h) \in (0,\infty),$$

wobei $b(\alpha,\beta,k-h)$ eindeutig durch die Eigenschaft

$$G_{b(\alpha,\beta,k-h),k-h}(\chi^2_{k-h;1-\alpha}) = \beta$$

festgelegt ist. Ferner ist

$$\vartheta^t i^{**}_{\zeta_1}(\eta_0)\vartheta = (\vartheta_1,0)^t i^{**}_{\zeta_1}(\eta_0)(\vartheta_1,0) = \vartheta_1^t A^{1,1}_{\zeta_1}(\vartheta_0)\vartheta_1.$$

Nun handle es sich bei der Folge von Tests T um Tests der Form (4.2.4) und es sei $(\tilde{\vartheta}_n)_{n\in\mathbb{N}}$ eine beliebige Folge mit $\lim_{n\to\infty}\tilde{\vartheta}_n = \vartheta_0$, sodass $\tilde{\vartheta}_n$ der zugrunde liegende Parameter ist. Dann ergibt eine Taylorentwicklung der Funktion $\eta \mapsto g_i(\eta)$, $i = 1,\ldots,k$,

$$\sqrt{n}\big(g_i(\hat{\eta}_n) - g_i(\eta_0)\big) = \nabla_h^t g_i(\eta)_{|\eta=\overline{\eta}_{i,n}}\sqrt{n}(\hat{\eta}_n - \eta_0),\ i = 1,\ldots,k,$$

mit Zwischenstellen $\overline{\eta}_{i,n}$ auf der Verbindungsstrecke zwischen $\hat{\eta}_n$ und η_0, $i = 1,\ldots,k$. Dies lässt sich äquivalent in der Form

$$\sqrt{n}\big(g(\hat{\eta}_n) - g(\eta_0)\big) = \begin{pmatrix}\nabla_h^t g_1(\eta)_{|\eta=\overline{\eta}_{1,n}}\\ \vdots\\ \nabla_h^t g_d(\eta)_{|\eta=\overline{\eta}_{k,n}}\end{pmatrix}\sqrt{n}(\hat{\eta}_n - \eta_0)$$

schreiben. Weiter gilt mit der stetigen Differenzierbarkeit von g

$$\lim_{n\to\infty}\begin{pmatrix}\nabla_h^t g_1(\eta)_{|\eta=\overline{\eta}_{1,n}}\\ \vdots\\ \nabla_h^t g_d(\eta)_{|\eta=\overline{\eta}_{k,n}}\end{pmatrix} = \mathcal{J}_g(\eta_0) \text{ in } Q_{\tilde{\vartheta}_n}\text{-Wahrscheinlichkeit},$$

womit auch

$$\sqrt{n}\big(g(\hat{\eta}_n) - g(\eta_0)\big) = \mathcal{J}_g(\eta_0)\sqrt{n}(\hat{\eta}_n - \eta_0) + o_{Q_{\tilde{\vartheta}_n}}(1)\sqrt{n}(\hat{\eta}_n - \eta_0) \text{ für } n \to \infty$$

folgt. Damit erhält man wegen $g(\eta_0) = \vartheta_0$ und mit Hilfe von Lemma 4.2.2

$$\begin{aligned}
&\sqrt{n}\big(\hat{\vartheta}_n - g(\hat{\eta}_n)\big)\\
=&\sqrt{n}(\hat{\vartheta}_n - \vartheta_0) - \sqrt{n}\big(g(\hat{\eta}_n) - g(\eta_0)\big)\\
=&\sqrt{n}(\hat{\vartheta}_n - \vartheta_0) - \mathcal{J}_g(\eta_0)\sqrt{n}(\hat{\eta}_n - \eta_0) + o_{Q_{\tilde{\vartheta}_n}}\sqrt{n}(\hat{\eta}_n - \eta_0)\\
=&\big(\,\mathrm{I}_d - \mathcal{J}_g(\eta_0) i^*_{\zeta_1}(\eta_0)^{-1}\mathcal{J}_g(\eta_0)^t i_{\zeta_1}(\vartheta_0)\big)\sqrt{n}(\hat{\vartheta}_n - \vartheta_0) + o_{Q_{\tilde{\vartheta}_n}}(1)\sqrt{n}(\hat{\vartheta}_n - \vartheta_0)\\
&+ o_{Q_{\tilde{\vartheta}_n}}(1)\sqrt{n}(\hat{\eta}_n - \eta_0) + o_{Q_{\tilde{\vartheta}_n}}(1) + o_{Q_{\tilde{\vartheta}_n}}(1)\sqrt{n}(\hat{\eta}_n - \eta_0) \text{ für } n \to \infty,
\end{aligned}$$

und daraus wegen der Stetigkeit der Abbildung $\vartheta \mapsto i_{\zeta_1}(\vartheta)$, $\vartheta \in \tilde{\Theta}$, die aus den Annahmen folgt, schließlich

$$\begin{aligned}
T_n &= \sqrt{n}\big(\hat{\vartheta}_n - g(\hat{\eta}_n)\big)^t i_{\zeta_1}\big(g(\hat{\eta}_n)\big)\sqrt{n}\big(\hat{\vartheta}_n - g(\hat{\eta}_n)\big)\big)\\
&= \sqrt{n}(\hat{\vartheta}_n - \vartheta_0)^t i^{**}_{\zeta_1}(\eta_0)\sqrt{n}(\hat{\vartheta}_n - \vartheta_0) + R_n,
\end{aligned}$$

mit einem entsprechenden Restterm R_n. Der Rest folgt analog zu den oben dargestellten Ausführungen. □

4.2.5 Korollar: *Es seien die zwei Folgen von Testgrößen T und T' jeweils gegeben durch (4.2.1), (4.2.2), (4.2.3) oder (4.2.4). Für ζ_1 und ζ_1' sowie für die zugehörigen Schätzfunktionen gelten die Bedingungen (B1) - (B5). Es sei eine reellwertige Folge $(\nu_i)_{i\in\mathbb{N}}$ mit $\nu_i \neq 0$ für alle $i \in \mathbb{N}$ und $\lim_{i\to\infty} \nu_i = 0$ gegeben. Für $i \in \mathbb{N}$ betrachten wir Alternativen der Form*

$$\vartheta^i := \begin{pmatrix}\vartheta_0^1\\ \vartheta_2\\ \vartheta_2'\end{pmatrix} + \nu_i\begin{pmatrix}\vartheta_1\\ 0\\ 0\end{pmatrix}$$

mit $\vartheta_1 \in \mathbb{R}^{d_1}\setminus\{0\}$, $(\vartheta_0^1, \vartheta_2, \vartheta_2') \in \partial\Theta_0$, die $(\vartheta_0^1 + \nu_i\vartheta_1, \vartheta_2, \vartheta_2') \in \Theta_1$ erfüllen, wobei mit $\vartheta_0 := \binom{\vartheta_0^1}{\vartheta_2}$ und $\vartheta_0' := \binom{\vartheta_0^1}{\vartheta_2'}$

$$\vartheta_1^t A^{1,1}_{\zeta_1}(\vartheta_0)\vartheta_1 \neq 0 \text{ und } \vartheta_1^t A^{1,1}_{\zeta_1'}(\vartheta_0')\vartheta_1 \neq 0$$

gelte. Zudem seien die Abbildungen

$$i \mapsto N_{T,i},\ i \mapsto N_{T',i},\ i \in \mathbb{N},$$

streng monoton wachsend, wobei $N_{T,i}$ und $N_{T',i}$ bei zugrunde liegender Folge von Alternativen $(\vartheta^i)_{i\in\mathbb{N}}$ und zugrunde liegendem $\beta \in (0, 1-\alpha)$ wie in (2.4.3) definiert seien. Dann gilt mit $\vartheta^0 := (\vartheta_0^1, \vartheta_2, \vartheta_2')$ im Fall der Existenz der Pitman-Effizienz

$$e^{\mathrm{P}}_{T',T}\big(\alpha, \beta, \vartheta^0, \vartheta_1\big) = \frac{b(\alpha, \beta, k-h)}{b(\alpha, \beta, k'-h')}\frac{\vartheta_1^t A^{1,1}_{\zeta_1'}(\vartheta_0')\vartheta_1}{\vartheta_1^t A^{1,1}_{\zeta_1}(\vartheta_0)\vartheta_1},$$

wobei für $m \in \mathbb{N}$ das $b(\alpha, \beta, m) \in (0, \infty)$ wie in Satz 4.2.4 bestimmt ist.

Beweis: Nach Satz 4.2.4 gilt

$$\lim_{i \to \infty} (N_{T,i} \nu_i^2) = \frac{b(\alpha, \beta, k-h)}{\vartheta_1^t A_{\zeta_1}^{1,1}(\vartheta_0) \vartheta_1},$$

sowie

$$\lim_{i \to \infty} (N_{T',i} \nu_i^2) = \frac{b(\alpha, \beta, k'-h')}{\vartheta_1^t A_{\zeta_1'}^{1,1}(\vartheta_0') \vartheta_1}.$$

Da die Pitman-Effizienz nach Annahme existiert, folgt

$$e_{T',T}^{\mathrm{P}} = \lim_{i \to \infty} \frac{N_{T,i} \nu_i^2}{N_{T',i} \nu_i^2} = \frac{b(\alpha, \beta, k-h)}{b(\alpha, \beta, k'-h')} \frac{\vartheta_1^t A_{\zeta_1'}^{1,1}(\vartheta_0') \vartheta_1}{\vartheta_1^t A_{\zeta_1}^{1,1}(\vartheta_0) \vartheta_1},$$

womit alles gezeigt ist □

4.2.6 Bemerkung: Es lässt sich unter Beachtung von Bemerkung 2.4.1 mit den Überlegungen präsentiert in 3.8.3 in [32] zeigen, dass die Pitman-Effizienz für $\nu_i = 1/\sqrt{i}$, $i \in \mathbb{N}$, existiert.

4.2.7 Bemerkung: Die Existenz von Folgen von Alternativen mit der in Korollar 4.2.5 genannten Eigenschaft, dass die Abbildungen

$$i \mapsto N_{T,i},\ i \mapsto N_{T',i},\ i \in \mathbb{N},$$

streng monoton wachsend sind, ist stets gewährleistet. Zur Begründung sei für fest gewähltes $\vartheta^0 \in \partial\Theta_0$ die Folge $(\vartheta^i)_{i \in \mathbb{N}}$ mit $\vartheta^i \in \Theta_1$ für alle $i \in \mathbb{N}$ und $\lim_{i \to \infty} \vartheta^i = \vartheta^0$ gemäß der Definition der Pitman-Effizienz in Unterabschnitt 2.4 gegeben. Wegen $N_{T,i} < +\infty$ für alle $i \in \mathbb{N}$ und $\lim_{i \to \infty} N_{T,i} = +\infty$ (siehe (2.4.4) und (2.4.5)) existiert eine Teilfolge $(\vartheta^{i_k})_{k \in \mathbb{N}}$ sodass die Abbildung $k \mapsto N_{T,i_k}$, $k \in \mathbb{N}$, streng monoton wachsend ist. Natürlich erfüllt wegen $N_{T',i} < +\infty$ für alle $i \in \mathbb{N}$ und $\lim_{i \to \infty} N_{T',i} = +\infty$ (siehe (2.4.4) und (2.4.5)) die Teilfolge $(\vartheta^{i_k})_{k \in \mathbb{N}}$ auch $\mathbb{N}_{T',i_k} < +\infty$ für alle $i \in \mathbb{N}$ und $\lim_{k \to \infty} N_{T',i_k} = +\infty$. Dann existiert also eine Teilfolge der Teilfolge $(\vartheta^{i_{k_l}})_{l \in \mathbb{N}}$ sodass die Abbildung $l \mapsto N_{T',i_{k_l}}$, $l \in \mathbb{N}$, streng monoton wachsend ist. Also erfüllt die Folge von Alternativen $(\vartheta^{i_{k_l}})_{l \in \mathbb{N}}$ die Eigenschaft, dass die Abbildungen

$$l \mapsto N_{T,i_{k_l}},\ l \mapsto N_{T',i_{k_l}},\ l \in \mathbb{N},$$

streng monoton wachsend sind.

4.2.8 Bemerkung: Betrachten wir den Fall, dass für ein $\vartheta_0^1 \in \mathbb{R}^{d_1}$

$$\Theta_0^1 = \{\vartheta_0^1\}$$

gilt, d.h. die Hypothese lässt sich mit $(\vartheta_1, \vartheta_2, \vartheta_2') \in \Theta$, äquivalent gemäß

$$\mathrm{H}: \ \vartheta_1 = \vartheta_0^1$$

formulieren.

Da die Mengen $\tilde{\Theta}$ und $\tilde{\Theta}'$ offen sind, sind auch die Mengen

$$\{\vartheta_2 \in \mathbb{R}^{k-d_1}; (\vartheta_0^1, \vartheta_2) \in \tilde{\Theta}\} \text{ und } \{\vartheta_2' \in \mathbb{R}^{k'-d_1}; (\vartheta_0^1, \vartheta_2') \in \tilde{\Theta}'\}$$

offen. Dann sind die Mengen Δ und Δ' gegeben durch

$$\Delta = \{\vartheta_2 \in \mathbb{R}^{k-d_1}; (\vartheta_0^1, \vartheta_2) \in \tilde{\Theta}\} \text{ und } \Delta' = \{\vartheta_2' \in \mathbb{R}^{k'-d_1}; (\vartheta_0^1, \vartheta_2') \in \tilde{\Theta}'\}$$

und die Abbildungen g und g' sind gegeben durch $g : \Delta \to \tilde{\Theta}$, $g(\vartheta_2) = \binom{\vartheta_0^1}{\vartheta_2}$ und $g' : \Delta' \to \tilde{\Theta}'$, $g'(\vartheta_2') = \binom{\vartheta_0^1}{\vartheta_2'}$. Dann gilt

$$k - h = k - (k - d_1) = d_1 = k' - (k' - d_1) = k' - h'.$$

Zudem ist $\mathcal{J}_g(\eta) = (0_{(k-d_1)\times d_1}, \mathrm{I}_{k-d_1})^t$ und $\mathcal{J}_{g'}(\eta') = (0_{(k'-d_1)\times d_1}, \mathrm{I}_{k'-d_1})^t$, wobei für $m, l \in \mathbb{N}$ die Matrix $0_{m\times l} \in \mathbb{R}^{m\times l}$ die Nullmatrix sei. Es sei

$$i_{\zeta_1}(\vartheta_0) = \begin{pmatrix} i_{\zeta_1}^{1,1}(\vartheta_0) & i_{\zeta_1}^{1,2}(\vartheta_0) \\ i_{\zeta_1}^{2,1}(\vartheta_0) & i_{\zeta_1}^{2,2}(\vartheta_0) \end{pmatrix} \in \mathbb{R}^{k\times k},$$

mit $i_{\zeta_1}^{1,1}(\vartheta_0) \in \mathbb{R}^{d_1\times d_1}$, $i_{\zeta_1}^{1,2}(\vartheta_0) \in \mathbb{R}^{d_1\times(k-d_1)}$, $i_{\zeta_1}^{2,1}(\vartheta_0) \in \mathbb{R}^{(k-d_1)\times d_1}$, $i_{\zeta_1}^{2,2}(\vartheta_0) \in \mathbb{R}^{(k-d_1)\times(k-d_1)}$. Dann erhält man wegen

$$i_{\zeta_1}^*(\eta_0) = \mathcal{J}_g(\eta_0)^t i_{\zeta_1}(\vartheta_0)\mathcal{J}_g(\eta_0)$$

leicht $i_{\zeta_1}^*(\eta_0) = i_{\zeta_1}^{2,2}(\vartheta_0)$, und weiter

$$i_{\zeta_1}^{**}(\vartheta_0) = \begin{pmatrix} i_{\zeta_1}^{1,1}(\vartheta_0) - i_{\zeta_1}^{1,2}(\vartheta_0) i_{\zeta_1}^{2,2}(\vartheta_0)^{-1} i_{\zeta_1}^{2,1}(\vartheta_0) & 0_{d_1\times(k-d_1)} \\ 0_{(k-d_1)\times d_1} & 0_{(k-d_1)\times(k-d_1)} \end{pmatrix} \in \mathbb{R}^{k\times k}.$$

Daher handelt es sich bei

$$A_{\zeta_1}^{1,1}(\vartheta_0) = i_{\zeta_1}^{1,1}(\vartheta_0) - i_{\zeta_1}^{1,2}(\vartheta_0) i_{\zeta_1}^{2,2}(\vartheta_0)^{-1} i_{\zeta_1}^{2,1}(\vartheta_0)$$

um das Schurkomplement von $i_{\zeta_1}^{2,2}(\vartheta_0)$ in $i_{\zeta_1}(\vartheta_0)$. Analog erhält man mit

$$i_{\zeta_1'}(\vartheta_0') = \begin{pmatrix} i_{\zeta_1'}^{1,1}(\vartheta_0') & i_{\zeta_1'}^{1,2}(\vartheta_0') \\ i_{\zeta_1'}^{2,1}(\vartheta_0') & i_{\zeta_1'}^{2,2}(\vartheta_0') \end{pmatrix} \in \mathbb{R}^{k'\times k'},$$

mit $i^{1,1}_{\zeta_1'}(\vartheta_0') \in \mathbb{R}^{d_1\times d_1}$, $i^{1,2}_{\zeta_1'}(\vartheta_0') \in \mathbb{R}^{d_1\times(k'-d_1)}$, $i^{2,1}_{\zeta_1'}(\vartheta_0') \in \mathbb{R}^{(k'-d_1)\times d_1}$, $i^{2,2}_{\zeta_1'}(\vartheta_0') \in \mathbb{R}^{(k'-d_1)\times(k'-d_1)}$,

$$i^{**}_{\zeta_1'}(\vartheta_0') = \begin{pmatrix} i^{1,1}_{\zeta_1'}(\vartheta_0') - i^{1,2}_{\zeta_1'}(\vartheta_0') i^{2,2}_{\zeta_1'}(\vartheta_0')^{-1} i^{2,1}_{\zeta_1'}(\vartheta_0') & 0_{d_1\times(k'-d_1)} \\ 0_{(k'-d_1)\times d_1} & 0_{(k'-d_1)\times(k'-d_1)} \end{pmatrix} \in \mathbb{R}^{k'\times k'},$$

und daher

$$A^{1,1}_{\zeta_1'}(\vartheta_0') = i^{1,1}_{\zeta_1'}(\vartheta_0') - i^{1,2}_{\zeta_1'}(\vartheta_0') i^{2,2}_{\zeta_1'}(\vartheta_0')^{-1} i^{2,1}_{\zeta_1'}(\vartheta_0').$$

Die Pitman-Effizienz der Tests T und T' ergibt sich dann über

$$e^{\mathrm{P}}_{T',T}(\vartheta^0,\vartheta_1) = \frac{\vartheta_1^t\big(i^{1,1}_{\zeta_1'}(\vartheta_0') - i^{1,2}_{\zeta_1'}(\vartheta_0') i^{2,2}_{\zeta_1'}(\vartheta_0')^{-1} i^{2,1}_{\zeta_1'}(\vartheta_0')\big)\vartheta_1}{\vartheta_1^t\big(i^{1,1}_{\zeta_1}(\vartheta_0) - i^{1,2}_{\zeta_1}(\vartheta_0) i^{2,2}_{\zeta_1}(\vartheta_0)^{-1} i^{2,1}_{\zeta_1}(\vartheta_0)\big)\vartheta_1}.$$

Gilt zusätzlich $i^{1,2}_{\zeta_1}(\vartheta_0) = 0_{d_1\times(k-d_1)}$, so ist $A^{1,1}_{\zeta_1}(\vartheta_0) = i^{1,1}_{\zeta_1}(\vartheta_0)$ die Informationsmatrix aus ζ_1, wenn der Parameter ϑ^1_0 als einziger variabler Parameter, und ϑ_2 als fest gegeben betrachtet wird.

Gilt zusätzlich $i^{1,2}_{\zeta_1'}(\vartheta_0') = 0_{d_1\times(k'-d_1)}$, so ist $A^{1,1}_{\zeta_1'}(\vartheta_0') = i^{1,1}_{\zeta_1'}(\vartheta_0')$ die Informationsmatrix aus ζ_1', wenn der Parameter ϑ^1_0 als einziger variabler Parameter, und ϑ_2' als fest gegeben betrachtet wird.

4.2.9 Beispiel: Wir führen wie zu Beginn dieser Arbeit motiviert einen Vergleich von unabhängiger und verbundener Stichprobenerhebung durch, wobei eine Situation wie aus Abschnitt 8 vorliege:

Im unabhängigen Stichprobenfall seien mit $X_1',\dots,X_n'$ und $Y_1',\dots,Y_n'$ zwei voneinander unabhängige Stichproben von unabhängigen $\mathbb{R}$-wertigen Zufallsvariablen gegeben, wobei die $X_1',\dots,X_n'$ jeweils die gleiche $\mathrm{N}(a,\Sigma)$-Verteilung besitzen, und die $Y_1',\dots,Y_n'$ jeweils die gleiche $\mathrm{N}(b,\Sigma)$-Verteilung besitzen, mit $a,b\in\mathbb{R}$ und $\Sigma>0$ unbekannt, d.h. es gilt

$$\begin{pmatrix} X_1' \\ Y_1' \end{pmatrix} \sim \mathrm{N}_2\left(\begin{pmatrix} a \\ b \end{pmatrix}, \begin{pmatrix} \Sigma & 0 \\ 0 & \Sigma \end{pmatrix}\right).$$

Im verbundenen Stichprobenfall liege dagegen eine Stichprobe $\binom{X_1}{Y_1},\dots,\binom{X_n}{Y_n}$ von unabhängigen $\mathbb{R}^2$-wertigen Zufallsvektoren vor, die jeweils die gleiche Verteilung, konkret

$$\begin{pmatrix} X_1 \\ Y_1 \end{pmatrix} \sim \mathrm{N}_2\left(\begin{pmatrix} a \\ b \end{pmatrix}, \begin{pmatrix} \Sigma & \Pi \\ \Pi & \Sigma \end{pmatrix}\right)$$

besitzen, wobei $\Pi\in\mathbb{R}$ ebenfalls unbekannt sei, sodass die zugrunde liegende Kovarianzmatrix $\left(\begin{smallmatrix} \Sigma & \Pi \\ \Pi & \Sigma \end{smallmatrix}\right)$ positiv definit sei.

In beiden Fällen soll das zu Beginn dieser Arbeit genannte Testproblem betrachtet werden, ob die Randverteilungen identisch sind oder nicht, also

$$\mathrm{H}:\ a=b,\ \mathrm{K}:\ a\neq b.$$

Mit $\mu := a-b$ kann das Testproblem äquivalent gemäß

$$\mathrm{H}:\ \mu=0,\ \mathrm{K}:\ \mu\neq 0$$

formuliert werden.

Nehmen wir an, dass es sich bei der Sequenz von Tests T' um die Likelihood-Quotienten-Tests basierend auf den Beobachtungen $\zeta_1' = \left(\begin{smallmatrix} X_1' \\ Y_1' \end{smallmatrix}\right), \ldots, \zeta_n' = \left(\begin{smallmatrix} X_n' \\ Y_n' \end{smallmatrix}\right)$ für das Testproblem handelt (siehe Unterabschnitt 8.5), und dass es sich bei der Sequenz von Tests T um die Likelihood-Quotienten-Tests basierend auf den Differenzen der Beobachtungen $\zeta_1 = X_1 - Y_1, \ldots, \zeta_n = X_n - Y_n$ für das Testproblem handelt (siehe Unterabschnitt 8.2). Dann gilt

$$\begin{pmatrix} X_1' \\ Y_1' \end{pmatrix} \sim \mathrm{N}_2\left(\begin{pmatrix} \mu+b \\ b \end{pmatrix}, \begin{pmatrix} \Sigma & 0 \\ 0 & \Sigma \end{pmatrix}\right)$$

und

$$X_1 - Y_1 \sim \mathrm{N}(\mu, \Gamma),$$

mit $\Gamma := 2\Sigma - 2\Pi$. Es kann folgende Parametrisierung benutzt werden: Wir wählen als gemeinsame Parametermenge der Verteilungen von ζ_1 und ζ_1'

$$\Theta = \left\{\left(\mu, \Gamma, \left(\begin{smallmatrix} b \\ \Sigma \end{smallmatrix}\right)\right) \in \mathbb{R}\times\mathbb{R}\times\mathbb{R}^2; \Sigma, \Gamma > 0, 4\Sigma > \Gamma\right\}.$$

Dann ist $\Theta^1 = \mathbb{R}$, $\vartheta_0^1 = 0 \in \mathbb{R}$, und

$$\Theta_0 = \left\{\left(\mu, \Gamma, \left(\begin{smallmatrix} b \\ \Sigma \end{smallmatrix}\right)\right) \in \Theta; \mu = 0\right\}.$$

Die injektive Partametrisierung der Verteilung von ζ_1 bzw. ζ_1' ist gegeben durch

$$\tilde{\Theta} = \{(\mu, \Gamma) \in \mathbb{R}\times\mathbb{R}; \Gamma > 0\}$$

bzw.

$$\tilde{\Theta}' = \left\{\left(\mu, \left(\begin{smallmatrix} b \\ \Sigma \end{smallmatrix}\right)\right) \in \mathbb{R}\times\mathbb{R}^2; \Sigma > 0\right\},$$

also

$$\tilde{\Theta}_0 = \{(\mu, \Gamma) \in \tilde{\Theta}; \mu = 0\}$$

bzw.

$$\tilde{\Theta}'_0 = \big\{ \big(\mu, \left(\begin{smallmatrix} b \\ \Sigma \end{smallmatrix}\right)\big) \in \tilde{\Theta}'; \mu = 0 \big\}.$$

Im unabhängigen Stichprobenfall ergibt sich nach einer elementaren Rechnung die zugrunde liegende Informationsmatrix bei der genannten injektiven Parametrisierung

$$i_{(X'_1,Y'_1)}(\vartheta'_0) = \begin{pmatrix} 1/\Sigma & 1/\Sigma & 0 \\ 1/\Sigma & 2/\Sigma & 0 \\ 0 & 0 & 1/\Sigma^2 \end{pmatrix}.$$

Damit folgt

$$\begin{aligned} A^{1,1}_{(X'_1,Y'_1)}(\vartheta'_0) &= i^{1,1}_{(X'_1,Y'_1)}(\vartheta'_0) - i^{1,2}_{(X'_1,Y'_1)}(\vartheta'_0) i^{2,2}_{(X'_1,Y'_1)}(\vartheta'_0)^{-1} i^{2,1}_{(X'_1,Y'_1)}(\vartheta'_0) \\ &= 1/\Sigma - \begin{pmatrix} 1/\Sigma & 0 \end{pmatrix} \begin{pmatrix} \Sigma/2 & 0 \\ 0 & \Sigma^2 \end{pmatrix} \begin{pmatrix} 1/\Sigma \\ 0 \end{pmatrix} = 1/\Sigma - 1/(2\Sigma) = 1/(2\Sigma). \end{aligned}$$

Im verbundenen Stichprobenfall ergibt sich nach kurzer Rechnung bei der genannten injektiven Parametrisierung $i^{1,1}_{X_1-Y_1}(\vartheta_0) = 1/\Gamma$ und $i^{1,2}_{X_1-Y_1}(\vartheta_0) = 0$, sodass

$$A^{1,1}_{X_1-Y_1}(\vartheta_0) = i^{1,1}_{X_1-Y_1}(\vartheta_0) = 1/\Gamma$$

folgt. Also ergibt sich insgesamt für die Pitman-Effizienz bei den entsprechenden Folgen von Alternativen

$$e^{\mathrm{P}}_{T',T}(\vartheta^0) = \frac{\Gamma}{2\Sigma}.$$

Vergleiche mit den Ergebnissen in Unterabschnitt 8.9.

4.2.10 Beispiel: Es soll eine Situation ähnlich wie in Beispiel 4.2.9 betrachtet werden, wobei wieder ein Vergleich von unabhängiger und verbundener Stichprobenerhebung durchgeführt werden soll.

Im unabhängigen Stichprobenfall seien mit $X'_1, \ldots, X'_n$ und $Y'_1, \ldots, Y'_n$ zwei voneinander unabhängige Stichproben von unabhängigen $\mathbb{R}^{d_1}$-wertigen Zufallsvariablen gegeben, wobei die $X'_1, \ldots, X'_n$ jeweils die gleiche $\mathrm{N}_{d_1}(a, \Sigma)$-Verteilung besitzen, und die $Y'_1, \ldots, Y'_n$ jeweils die gleiche $\mathrm{N}_{d_1}(b, \Lambda)$-Verteilung besitzen, mit $a, b \in \mathbb{R}^{d_1}$ unbekannt, und $\Sigma, \Lambda \in \mathbb{R}^{d_1 \times d_1}$ symmetrisch positiv definit und unbekannt.

Im verbundenen Stichprobenfall liege dagegen eine Stichprobe $\left(\begin{smallmatrix} X_1 \\ Y_1 \end{smallmatrix}\right), \ldots, \left(\begin{smallmatrix} X_n \\ Y_n \end{smallmatrix}\right)$ von unabhängigen $\mathbb{R}^{2d_1}$-wertigen Zufallsvektoren vor, die jeweils die gleiche Verteilung besitzen, konkret

$$\begin{pmatrix} X_1 \\ Y_1 \end{pmatrix} \sim \mathrm{N}_{2d_1}\left(\begin{pmatrix} a \\ b \end{pmatrix}, \begin{pmatrix} \Sigma & \Pi \\ \Pi^t & \Lambda \end{pmatrix} \right),$$

wobei $\xi \in \mathbb{R}$ ebenfalls unbekannt sei, sodass die zugrunde liegende Kovarianzmatrix $\left(\begin{smallmatrix} \Sigma & \Pi \\ \Pi^t & \Lambda \end{smallmatrix}\right)$ positiv definit sei.

In beiden Fällen soll dass zu Beginn dieser Arbeit genannte Testproblem betrachtet werden, ob die Randverteilungen identisch sind oder nicht, also

$$\mathrm{H}: \ a = b, \ \mathrm{K}: \ a \neq b.$$

Es handelt sich also im unabhängigen Stichprobenfall um das klassische Behrens-Fisher-Problem. Mit $\mu := a - b$ kann das Testproblem äquivalent gemäß

$$\mathrm{H}: \ \mu = 0, \ \mathrm{K}: \ \mu \neq 0$$

formuliert werden.

Nehmen wir an, dass es sich bei der Sequenz von Tests T' um die Likelihood-Quotienten-Tests basierend auf den Differenzen der Beobachtungen $\zeta_1' = X_1' - Y_1', \ldots, \zeta_n' = X_n' - Y_n'$ für das Testproblem handelt (siehe Unterabschnitt 8.2), und dass es sich bei der Sequenz von Tests T um die Likelihood-Quotienten-Tests basierend auf den Differenzen der Beobachtungen $\zeta_1 = X_1 - Y_1, \ldots, \zeta_n = X_n - Y_n$ für das Testproblem handelt (siehe Unterabschnitt 8.2). Dann gilt

$$X_1' - Y_1' \sim \mathrm{N}(\mu, \Gamma_1),$$

mit $\Gamma_1 := \Sigma + \Lambda$, und

$$X_1 - Y_1 \sim \mathrm{N}(\mu, \Gamma_2),$$

mit $\Gamma_2 := \Sigma + \Lambda - (\Pi + \Pi^t)$. Es kann folgende Parametrisierung benutzt werden: Es sei für $l \in \mathbb{N}$ $\mathcal{P}_l := \{A \in \mathbb{R}^{l \times l}; A \text{ ist symmetrisch positiv definit}\}$. Wir wählen als gemeinsame Parametermenge der Verteilungen von ζ_1 und ζ_1'

$$\begin{aligned}\Theta =& \Big\{ \big(\mu, \mathrm{vech}(\Gamma_2), \mathrm{vech}(\Gamma_1)\big) \in \mathbb{R}^{d_1} \times \mathbb{R}^{d_1(d_1+1)/2} \times \mathbb{R}^{d_1(d_1+1)/2} \\ & \exists\, \Sigma, \Lambda \in \mathcal{S}_{d_1}, \ \Pi \in \mathbb{R}^{d_1 \times d_1} : \Gamma_1 = \Sigma + \Lambda, \ \Gamma_2 = \Sigma + \Lambda - (\Pi + \Pi^t), \left(\begin{smallmatrix} \Sigma & \Pi \\ \Pi^t & \Sigma \end{smallmatrix}\right) \in \mathcal{P}_{2d_1} \Big\}.\end{aligned}$$

Dabei ist für $\ell \in \mathbb{N}$ der vech-Operator definiert durch

$$\begin{aligned}&\mathrm{vech} : \mathbb{R}^{\ell \times \ell} \to \mathbb{R}^{\ell(\ell+1)/2}, \\ &\mathrm{vech}(M) := (m_{1,1}, \ldots, m_{\ell,1}, m_{2,2}, \ldots, m_{\ell,2}, \ldots, m_{\ell,\ell})^t, \ M = (m_{i,j})_{1 \le i,j \le \ell}.\end{aligned}$$

Dann ist $\Theta^1 = \mathbb{R}^{d_1}$, $\vartheta_0^1 = 0 \in \mathbb{R}^{d_1}$, und

$$\Theta_0 = \Big\{ \big(\mu, \mathrm{vech}(\Gamma_2), \mathrm{vech}(\Gamma_1)\big) \in \Theta; \mu = 0 \Big\}.$$

Die injektive Partametrisierung der Verteilung von ζ_1 bzw. ζ_1' ist dann gegeben durch

$$\tilde{\Theta} = \Big\{ \big(\mu, \mathrm{vech}(\Gamma_2)\big) \in \mathbb{R}^{d_1} \times \mathbb{R}^{d_1(d_1+1)/2}; \Gamma_2 \in \mathcal{P}_{d_1} \Big\}$$

bzw.

$$\tilde{\Theta}' = \big\{ \big(\mu, \operatorname{vech}(\Gamma_1)\big) \in \mathbb{R}^{d_1} \times \mathbb{R}^{d_1(d_1+1)/2}; \Gamma_1 \in \mathcal{P}_{d_1} \big\},$$

also

$$\tilde{\Theta}_0 = \big\{ \big(\mu, \operatorname{vech}(\Gamma_2)\big) \in \tilde{\Theta}; \mu = 0 \big\}$$

bzw.

$$\tilde{\Theta}'_0 = \big\{ \big(\mu, \operatorname{vech}(\Gamma_1)\big) \in \tilde{\Theta}'; \mu = 0 \big\}.$$

Im unabhängigen Stichprobenfall ergibt sich nach kurzer Rechnung bei der genannten injektiven Parametrisierung $i^{1,1}_{X'_1 - Y'_1}(\vartheta'_0) = \Gamma_1^{-1}$ und $i^{1,2}_{X'_1 - Y'_1}(\vartheta'_0) = 0$, sodass

$$A^{1,1}_{X'_1 - Y'_1}(\vartheta'_0) = i^{1,1}_{X'_1 - Y'_1}(\vartheta'_0) = \Gamma_1^{-1}$$

folgt.

Im verbundenen Stichprobenfall ergibt sich nach kurzer Rechnung bei der genannten injektiven Parametrisierung $i^{1,1}_{X_1 - Y_1}(\vartheta_0) = \gamma_2^{-1}$ und $i^{1,2}_{X_1 - Y_1}(\vartheta_0) = 0$, sodass

$$A^{1,1}_{X_1 - Y_1}(\vartheta_0) = i^{1,1}_{X_1 - Y_1}(\vartheta_0) = \Gamma_2^{-1}$$

folgt.

Also ergibt sich insgesamt für die Pitman-Effizienz bei den entsprechenden Folgen von Alternativen

$$e^{\mathrm{P}}_{T',T}(\vartheta^0, \vartheta_1) = \frac{\vartheta_1^t i^{1,1}_{X'_1 - Y'_1}(\vartheta'_0)\vartheta_1}{\vartheta_1^t i^{1,1}_{X_1 - Y_1}(\vartheta_0)\vartheta_1} = \frac{\vartheta_1^t \Gamma_1^{-1} \vartheta_1}{\vartheta_1^t \Gamma_2^{-1} \vartheta_1}.$$

4.3 Lokale Bahadur- und Hodges-Lehmann-Effizienz bei optimalen Tests

Handelt es sich bei den betrachteten Tests um optimale Tests im Bahadur- bzw. Hodges-Lehmann-Sinn, so besitzt die lokale Bahadur- bzw. Hodges-Lehmann-Effizienz unter gewissen Voraussetzungen die Gestalt der in Unterabschnitt 4.2 ermittelten Pitman-Effizienz. Dies soll in diesem Abschnitt gezeigt werden. Angewendet werden können die hier erarbeiteten Resultate in den Abschnitten 8 und 9.

Im Folgenden wollen wir Annahmen an die Zufallsvariablen ζ_1 und ζ'_1 bei den zugrunde liegenden injektiven Parametrisierungen durch $\tilde{\Theta}$ und $\tilde{\Theta}'$ formulieren. Der Einfachheit halber formulieren wir diese Annahmen nur für ζ_1 und $\tilde{\Theta}$; sie sollen natürlich in analoger

Weise auch für ζ_1' und $\tilde{\Theta}'$ gelten. Dabei werden die im Zusammenhang mit ζ_1 und $\tilde{\Theta}$ eingeführten Größen im Zusammenhang mit ζ_1' und $\tilde{\Theta}'$ mit einem Strich versehen.

Wir nehmen an, dass ein σ-endliches Maß μ auf $(R, \mathfrak{S})$ existiert, sodass für jedes $\vartheta \in \tilde{\Theta}$ $Q_\vartheta^{\zeta_1} \ll \mu$ gilt, und $f(\cdot;\vartheta)$ eine μ-Dichte von $Q_\vartheta^{\zeta_1}$ ist. Zur Vereinfachung der Notation definieren wir

$$L(z,\vartheta) := \log f(z;\vartheta),\ z \in R,\ \vartheta \in \tilde{\Theta},$$

mit $\log 0 := -\infty$. Wir setzen generell voraus, dass $f(z;\vartheta) > 0$ für alle $\vartheta \in \tilde{\Theta}$ und alle $z \in R$.

Im Fall einer zusammengesetzten Hypothese sei stets $k \geq 2$ und für $1 \leq h < k$ außerdem $\Delta \subset \mathbb{R}^h$ offen und konvex, und es existiere eine zweimal stetig differenzierbare und injektive Funktion $g : \Delta \to \tilde{\Theta}$, sodass $g(\Delta) = \tilde{\Theta}_0$ gilt, und dass die Funktionalmatrix von g an der Stelle $\eta \in \Delta$, $\mathcal{J}_g(\eta)$, für alle $\eta \in \Delta$ den Rang h habt.

Wir wollen im Folgenden ausschließlich den Fall betrachten, dass eine zusammengesetzte Hypothese vorliegt. Im Fall, dass bei einem Test eine einfache Hypothese vorliegt, kann entsprechend vorgegangen werden

Mit $\vartheta = (\vartheta_1, \ldots, \vartheta_k)^t \in \tilde{\Theta}$ und $\eta = (\eta_1, \ldots, \eta_h)^t \in \Lambda$ definieren wir die Nabla-Operatoren

$$\nabla_k := \Big(\frac{\partial}{\partial\vartheta_1}, \ldots, \frac{\partial}{\partial\vartheta_k}\Big)^t,\ \nabla_h := \Big(\frac{\partial}{\partial\eta_1}, \ldots, \frac{\partial}{\partial\eta_h}\Big)^t.$$

Zudem formulieren wir in Vorbereitung auf das Resultat zur lokalen Bahadur-Effizienz folgende Bedingungen:

(C1) Es sei die Abbildung $\vartheta \mapsto f(z;\vartheta)$ für alle $z \in R$ zweimal stetig differenzierbar.

(C2) Für alle $1 \leq i,j \leq k$ und für alle $\vartheta \in \tilde{\Theta}$ gelte $\int \frac{\partial}{\partial\vartheta_i} f(z;\vartheta)\mathrm{d}\mu(z) = \int \frac{\partial^2}{\partial\vartheta_i\partial\vartheta_j} f(z;\vartheta)\mathrm{d}\mu(z) = 0$ und für alle $\vartheta \in \tilde{\Theta}$ existiere die Fisher-Informationsmatrix aus $Q_\vartheta^{\zeta_1}$, $i_{\zeta_1}(\vartheta) = (\mathrm{E}_\vartheta(\frac{\partial}{\partial\vartheta_i} L(\zeta_1,\vartheta)\frac{\partial}{\partial\vartheta_j} L(\zeta_1,\vartheta)))_{\substack{1\leq i\leq k\\ 1\leq j\leq k}}$, besitze endliche Einträge und sei positiv definit.

(C3) Es existiere für jedes $\vartheta \in \tilde{\Theta}$ und jedes $1 \leq i \leq k$ ein $\delta > 0$ sodass mit $U_\delta := \{\tilde{\vartheta} \in \mathbb{R}^k : |\tilde{\vartheta} - \vartheta| < \delta\} \subset \tilde{\Theta}$ und $M_\delta(z) := \sup_{\vartheta^* \in U_\delta} |\frac{\partial}{\partial\tilde{\vartheta}_i} L(z,\tilde{\vartheta})|_{\tilde{\vartheta}=\vartheta^*}| \sup_{\vartheta^* \in U_\delta} f(z;\vartheta^*)$, $z \in R$, $\int M_\delta(z)\mathrm{d}\mu(z) < \infty$ ist. Genauso existiere für jedes $\vartheta \in \tilde{\Theta}$ und jede $1 \leq i,j \leq k$ ein $\delta > 0$ sodass mit $\tilde{M}_\delta(z) := \sup_{\vartheta^* \in U_\delta} |\frac{\partial^2}{\partial\tilde{\vartheta}_i\partial\tilde{\vartheta}_j} L(z,\tilde{\vartheta})|_{\tilde{\vartheta}=\vartheta^*}| \sup_{\vartheta^* \in U_\delta} f(z;\vartheta^*)$, $z \in R$, $\int \tilde{M}_\delta(z)\mathrm{d}\mu(z) < \infty$ ist.

(C4) Für alle $\vartheta, \tilde{\vartheta} \in \tilde{\Theta}$ existiere die Kullback-Leibler-Information $K(\vartheta, \tilde{\vartheta})$ aus ζ_1. Zudem existiere für jedes $\vartheta_0 \in \partial\tilde{\Theta}_0$ eine offene Umgebung U_{ϑ_0} um ϑ_0 und eine auf U_{ϑ_0} stetige Abbildung $\hat{\eta} : U_{\vartheta_0} \to \Delta$ mit den Eigenschaften $K(\vartheta, \tilde{\Theta}_0) = K(\vartheta, g(\hat{\eta}(\vartheta)))$ und $\int \nabla_h(\log \frac{f(z;\vartheta)}{f(z;g(\eta))} f(z;\vartheta))|_{\eta=\hat{\eta}(\vartheta)}\mathrm{d}\mu(z) = 0$ für alle $\vartheta \in U_{\vartheta_0}$.

Die Bedingungen (C1) - (C3) sind identisch mit den Bedingungen (B1) - (B3) aus Unterabschnitt 4.2. Zur Verifizierung der Bedingung (C4) kann das folgende Lemma nützlich sein.

4.3.1 Lemma: *Es gelte (C1) und (C2). Weiter sei $\vartheta_0 \in \partial\tilde{\Theta}_0$ fest gewählt und $\eta_0 := g^{-1}(\vartheta_0)$. Es existiere das Integral $\int \nabla_h^t(\log \frac{f(z;\vartheta)}{f(z;g(\eta))} f(z;\vartheta))\mathrm{d}\mu(z)$ für (ϑ,η) in einer offenen Umgebung um (ϑ_0,η_0) und es sei für (ϑ,η) in dieser Umgebung um (ϑ_0,η_0)*

$$F(\vartheta,\eta) := \int \nabla_h^t\Big(\log \frac{f(z;\vartheta)}{f\big(z;g(\eta)\big)} f(z;\vartheta)\Big)\mathrm{d}\mu(z),$$

wobei F auf dieser offenen Umgebung um (ϑ_0,η_0) stetig differenzierbar sei. Zudem gelte $\nabla_h F(\vartheta_0,\eta)|_{\eta=\eta_0} = \int \nabla_h \nabla_h^t(\log \frac{f(z;\vartheta_0)}{f(z;g(\eta))} f(z;\vartheta_0))|_{\eta=\eta_0}\mathrm{d}\mu(z)$. Dann existiert eine offene Umgebung U_{ϑ_0} um ϑ_0 und eine eindeutige auf U_{ϑ_0} stetige Abbildung $\hat{\eta} : U_{\vartheta_0} \to \Delta$ mit $F\big(\vartheta,\hat{\eta}(\vartheta)\big) = 0$ für alle $\vartheta \in U_{\vartheta_0}$.

Beweis: Es ist für (ϑ,η) in einer Umgebung um (ϑ_0,η_0)

$$F(\vartheta,\eta) = -\int \Big(\nabla_h^t f\big(z;g(\eta)\big)\Big)\frac{1}{f\big(z;g(\eta)\big)} f(z;\vartheta)\mathrm{d}\mu(z)$$

und daher wegen (C2) und $\eta_0 = g^{-1}(\vartheta_0)$ $F(\vartheta_0,\eta_0) = 0$. Weiter ist mit (C2) und $\eta_0 = g^{-1}(\vartheta_0)$

$$\begin{aligned}
\nabla_h F(\vartheta_0,\eta)_{|\eta=\eta_0} &= \nabla_h \int \nabla_h^t\Big(\log \frac{f(z;\vartheta_0)}{f\big(z;g(\eta)\big)} f(z;\vartheta_0)\Big)_{|\eta=\eta_0}\mathrm{d}\mu(z)\\
&= \int \nabla_h \nabla_h^t\Big(\log \frac{f(z;\vartheta_0)}{f\big(z;g(\eta)\big)} f(z;\vartheta_0)\Big)_{|\eta=\eta_0}\mathrm{d}\mu(z)\\
&= -\int \Bigg(\frac{1}{f\big(z;g(\eta_0)\big)}\nabla_h\nabla_h^t f\big(z;g(\eta)\big)_{|\eta=\eta_0} f(z;\vartheta_0)\\
&\quad - \frac{1}{f\big(z;g(\eta_0)\big)^2}\Big(\nabla_h f\big(z;g(\eta)\big)\Big)_{|\eta=\eta_0}\Big(\nabla_h^t f\big(z;g(\eta)\big)\Big)_{|\eta=\eta_0} f(z;\vartheta_0)\Bigg)\mathrm{d}\mu(z)\\
&= \int \frac{\nabla_h f\big(z;g(\eta)\big)_{|\eta=\eta_0}}{f\big(z;g(\eta_0)\big)}\frac{\nabla_h^t f\big(z;g(\eta)\big)_{|\eta=\eta_0}}{f\big(z;g(\eta_0)\big)} f(z;\vartheta_0)\mathrm{d}\mu(z) = i_{\zeta_1}^*(\eta_0),
\end{aligned}$$

mit

$$i_{\zeta_1}^*(\eta_0) := \mathcal{J}_g(\eta_0)^t i_{\zeta_1}(\vartheta_0)\mathcal{J}_g(\eta_0).$$

Wegen (C2) ist $\nabla_h F(\vartheta_0,\eta)_{|\eta=\eta_0} = i_{\zeta_1}^*(\eta_0)$ positiv definit. Nach dem Satz von der impliziten Funktion, siehe Satz 169.1 in [13], existiert eine offene Umgebung U_{ϑ_0} um ϑ_0 und eine eindeutige auf U_{ϑ_0} stetige Abbildung $\hat{\eta} : U_{\vartheta_0} \to \Delta$ mit $\hat{\eta}(\vartheta_0) = \eta_0$ und

$$0 = F\big(\vartheta,\hat{\eta}(\vartheta)\big) = \int \nabla_h^t\Big(\log \frac{f(z;\vartheta)}{f\big(z;g(\eta)\big)} f(z;\vartheta)\Big)_{|\eta=\hat{\eta}(\vartheta)}\mathrm{d}\mu(z) \text{ für alle } \vartheta \in U_{\vartheta_0}.$$

Damit ist alles gezeigt. $\square$

Nun werden zwei Lemmata betrachtet.

4.3.2 Lemma: *Es gelte (C1) und (C4). Zudem sei $\vartheta_0 \in \partial\tilde{\Theta}_0$ fest gewählt und $\eta_0 := g^{-1}(\vartheta_0)$. Dann gilt*

$$\lim_{\vartheta\to\vartheta_0} \hat{\eta}(\vartheta) = \hat{\eta}(\vartheta_0) = \eta_0.$$

Beweis: Da eine injektive Parametrisierung der Verteilung von ζ_1 durch die Elemente von $\tilde{\Theta}$ vorliegt, gilt für alle $\vartheta, \tilde{\vartheta} \in \tilde{\Theta}$ die Ungleichung $K(\vartheta, \tilde{\vartheta}) \geq 0$ sowie die Äquivalenz $K(\vartheta, \tilde{\vartheta}) = 0 \;\Leftrightarrow\; \vartheta = \tilde{\vartheta}$. Wegen der in (C4) geforderten Stetigkeit der Abbildung $\hat{\eta}$ in ϑ_0 gilt $\lim_{\vartheta\to\vartheta_0} \hat{\eta}(\vartheta) = \hat{\eta}(\vartheta_0)$. Weiter folgt aus

$$0 = K(\vartheta_0, \tilde{\Theta}_0) = K\Big(\vartheta_0, g\big(\hat{\eta}(\vartheta_0)\big)\Big)$$

sofort

$$g\big(\hat{\eta}(\vartheta_0)\big) = \vartheta_0,$$

äquivalent

$$\hat{\eta}(\vartheta_0) = g^{-1}(\vartheta_0) = \eta_0,$$

womit alles gezeigt ist. $\square$

4.3.3 Lemma: *Für ζ_1 gelten die Bedingungen (C1) - (C4). Zudem sei $\vartheta_0 \in \partial\tilde{\Theta}_0$ fest gewählt. Wir definieren $\eta_0 := g^{-1}(\vartheta_0)$. Dann gilt für ϑ in einer offenen Umgebung um ϑ_0*

$$(\hat{\eta} - \eta_0) = i^*_{\zeta_1}(\eta_0)^{-1} \mathcal{J}_g(\eta_0)^t i_{\zeta_1}(\vartheta_0)(\vartheta - \vartheta_0) + R_0(\vartheta, \vartheta_0)(\vartheta - \vartheta_0),$$

wobei wir $\hat{\eta} = \hat{\eta}(\vartheta)$ schreiben, mit

$$i^*_{\zeta_1}(\eta_0) := \mathcal{J}_g(\eta_0)^t i_{\zeta_1}(\vartheta_0) \mathcal{J}_g(\eta_0),$$

und einer Matrix $R_0(\vartheta, \vartheta_0) \in \mathbb{R}^{h\times k}$ mit $\lim_{\vartheta\to\vartheta_0} R_0(\vartheta, \vartheta_0) = 0$.

Beweis: Eine Taylorentwicklung der Funktion $\vartheta \mapsto \frac{\partial}{\partial\vartheta_j} L(z, \vartheta)$, $z \in R$, $j = 1, \ldots, k$, liefert

$$\frac{\partial}{\partial\vartheta_j} L(z, \vartheta) = \frac{\partial}{\partial\vartheta_j} L(z, \vartheta)_{|\vartheta=\vartheta_0} + \nabla^t_k \frac{\partial}{\partial\vartheta_j} L(z, \vartheta)_{|\vartheta=\overline{\vartheta}_{j,z}} (\vartheta - \vartheta_0), \; z \in R, \; j = 1, \ldots, k,$$

mit einer Zwischenstelle $\overline{\vartheta}_{j,z}$ auf der Verbindungsstrecke zwischen ϑ und ϑ_0, $z \in R$, $j = 1, \ldots, k$. Anders geschrieben erhält man

$$\nabla_k L(z, \vartheta)_{|\vartheta=\vartheta_0} = \nabla_k L(z, \vartheta) - \begin{pmatrix} \nabla^t_k \frac{\partial}{\partial\vartheta_1} L(z, \vartheta)_{|\vartheta=\overline{\vartheta}_{1,z}} \\ \vdots \\ \nabla^t_k \frac{\partial}{\partial\vartheta_k} L(z, \vartheta)_{|\vartheta=\overline{\vartheta}_{k,z}} \end{pmatrix} (\vartheta - \vartheta_0), \; z \in R.$$

Es ist mit (C2)

$$0=\int \nabla_k f(z;\vartheta)\mathrm{d}\mu(z)=\int \big(\nabla_k L(z,\vartheta)\big) f(z;\vartheta)\mathrm{d}\mu(z),$$

sodass durch Integrieren für ϑ in einer offenen Umgebung um ϑ_0

$$\int \big(\nabla_k L(z,\vartheta)\big)_{|\vartheta=\vartheta_0} f(z,\vartheta)\mathrm{d}\mu(z)=-\begin{pmatrix}\int \big(\nabla_k^t \frac{\partial}{\partial\vartheta_1} L(z,\vartheta)_{|\vartheta=\overline{\vartheta}_{1,z}}\big) f(z,\vartheta)\mathrm{d}\mu(z)\\ \vdots \\ \int \nabla\big(k^t \frac{\partial}{\partial\vartheta_k} L(z,\vartheta)_{|\vartheta=\overline{\vartheta}_{k,z}}\big) f(z,\vartheta)\mathrm{d}\mu(z)\end{pmatrix}(\vartheta-\vartheta_0) \tag{4.3.1}$$

folgt, wobei die Existenz der beteiligten Integrale durch (C3) für ϑ in einer offenen Umgebung um ϑ_0 gewährleistet werden kann.

Andererseits liefert eine entsprechende Taylorentwicklung

$$\begin{aligned}&\frac{\partial}{\partial\eta_j} L(z,g(\eta))_{|\eta=\hat\eta}=\frac{\partial}{\partial\eta_j} L\big(z,g(\eta)\big)_{|\eta=\eta_0}+\nabla_h^t\frac{\partial}{\partial\eta_j} L\big(z,g(\eta)\big)_{|\eta=\overline{\eta}_{j,z}}(\hat\eta-\eta_0),\\ &z\in R,\ j=1,\dots,h,\end{aligned}$$

mit einer Zwischenstelle $\overline{\eta}_{j,z}$ auf der Verbindungsstrecke zwischen $\hat\eta$ und $\eta_0,\ z\in R,\ j=1,\dots,h$. Anders geschrieben erhält man wegen $g(\eta_0)=\vartheta_0$

$$\mathcal{J}_g(\eta_0)^t\nabla_k L(z,\vartheta)_{|\vartheta=\vartheta_0}=\mathcal{J}_g(\hat\eta)^t\nabla_k L(z,\vartheta)_{|\vartheta=g(\hat\eta)}-\begin{pmatrix}\nabla_h^t\frac{\partial}{\partial\eta_1}L\big(z,g(\eta)\big)_{|\eta=\overline{\eta}_{1,z}}\\ \vdots\\ \nabla_h^t\frac{\partial}{\partial\eta_h}L\big(z,g(\eta)\big)_{|\eta=\overline{\eta}_{h,z}}\end{pmatrix}(\hat\eta-\eta_0),\ z\in R.$$

Es ist wegen (C4) für ϑ in einer offenen Umgebung um ϑ_0

$$\begin{aligned}0&=-\int\nabla_h\Big(\log\frac{f(z;\vartheta)}{f\big(z;g(\eta)\big)}f(z;\vartheta)\Big)_{|\eta=\hat\eta(\vartheta)}\mathrm{d}\mu(z)=\int\nabla_h\log f\big(z;g(\eta)\big)_{|\eta=\hat\eta(\vartheta)}f(z;\vartheta)\mathrm{d}\mu(z)\\ &=\mathcal{J}_g(\hat\eta)^t\int\big(\nabla_k L(z,\vartheta)\big)_{|\vartheta=g(\hat\eta)}f(z,\vartheta)\mathrm{d}\mu(z).\end{aligned}$$

Durch Integrieren folgt daraus für ϑ in einer offenen Umgebung um ϑ_0

$$\begin{aligned}&\mathcal{J}_g(\hat\eta)^t\int\big(\nabla_k L(z,\vartheta)\big)_{|\vartheta=\vartheta_0}f(z,\vartheta)\mathrm{d}\mu(z)\\ &=-\begin{pmatrix}\int\nabla_h^t\frac{\partial}{\partial\eta_1}L\big(z,g(\eta)\big)_{|\eta=\overline{\eta}_{1,z}}f(z,\vartheta)\mathrm{d}\mu(z)\\ \vdots\\ \int\nabla_h^t\frac{\partial}{\partial\eta_h}L\big(z,g(\eta)\big)_{|\eta=\overline{\eta}_{h,z}}f(z,\vartheta)\mathrm{d}\mu(z)\end{pmatrix}(\hat\eta-\eta_0),\end{aligned} \tag{4.3.2}$$

wobei die Existenz der beteiligten Integrale mit Hilfe von Lemma 4.3.2 durch (C4) für ϑ in einer offenen Umgebung um ϑ_0 gewährleistet werden kann.

Durch Einsetzen von (4.3.1) in (4.3.2) folgt für ϑ in einer offenen Umgebung um ϑ_0

$$\begin{aligned}
&\mathcal{J}_g(\hat{\eta})^t \begin{pmatrix} \int \left(\nabla_k^t \frac{\partial}{\partial \vartheta_1} L(z,\vartheta)_{|\vartheta=\overline{\vartheta}_{1,z}}\right) f(z,\vartheta)\mathrm{d}\mu(z) \\ \vdots \\ \int \left(\nabla_k^t \frac{\partial}{\partial \vartheta_k} L(z,\vartheta)_{|\vartheta=\overline{\vartheta}_{k,z}}\right) f(z,\vartheta)\mathrm{d}\mu(z) \end{pmatrix} (\vartheta - \vartheta_0) \\
=& \begin{pmatrix} \int \nabla_h^t \frac{\partial}{\partial \eta_1} L\big(z,g(\eta)\big)_{|\eta=\overline{\eta}_{1,z}} f(z,\vartheta)\mathrm{d}\mu(z) \\ \vdots \\ \int \nabla_h^t \frac{\partial}{\partial \eta_h} L\big(z,g(\eta)\big)_{|\eta=\overline{\eta}_{h,z}} f(z,\vartheta)\mathrm{d}\mu(z) \end{pmatrix} (\hat{\eta} - \eta_0).
\end{aligned}$$

Wegen (4.2.6), $g(\eta_0) = \vartheta_0$, der zweimaligen stetigen Differenzierbarkeit von g und mit Lemma 4.3.2 erhält man unter Heranziehung von (C1) -(C4) mit dem Satz von der majorisierten Konvergenz

$$\lim_{\vartheta \to \vartheta_0} \begin{pmatrix} \int \nabla_h^t \frac{\partial}{\partial \eta_1} L\big(z,g(\eta)\big)_{|\eta=\overline{\eta}_{1,z}} f(z,\vartheta)\mathrm{d}\mu(z) \\ \vdots \\ \int \nabla_h^t \frac{\partial}{\partial \eta_h} L\big(z,g(\eta)\big)_{|\eta=\overline{\eta}_{h,z}} f(z,\vartheta)\mathrm{d}\mu(z) \end{pmatrix} = -i^*_{\zeta_1}(\eta_0).$$

Wegen der Invertierbarkeit von $i^*_{\zeta_1}(\eta_0)$ ist daher für ϑ in einer offenen Umgebung um ϑ_0 die Matrix

$$\begin{pmatrix} \int \nabla_h^t \frac{\partial}{\partial \eta_1} L\big(z,g(\eta)\big)_{|\eta=\overline{\eta}_{1,z}} f(z,\vartheta)\mathrm{d}\mu(z) \\ \vdots \\ \int \nabla_h^t \frac{\partial}{\partial \eta_h} L\big(z,g(\eta)\big)_{|\eta=\overline{\eta}_{h,z}} f(z,\vartheta)\mathrm{d}\mu(z) \end{pmatrix}$$

invertierbar. Somit gilt für ϑ genügend nahe an ϑ_0

$$\begin{aligned}
(\hat{\eta} - \eta_0) = & \begin{pmatrix} \int \nabla_h^t \frac{\partial}{\partial \eta_1} L\big(z,g(\eta)\big)_{|\eta=\overline{\eta}_{1,z}} f(z,\vartheta)\mathrm{d}\mu(z) \\ \vdots \\ \int \nabla_h^t \frac{\partial}{\partial \eta_h} L\big(z,g(\eta)\big)_{|\eta=\overline{\eta}_{h,z}} f(z,\vartheta)\mathrm{d}\mu(z) \end{pmatrix}^{-1} \\
& \mathcal{J}_g(\hat{\eta})^t \begin{pmatrix} \int \left(\nabla_k^t \frac{\partial}{\partial \vartheta_1} L(z,\vartheta)_{|\vartheta=\overline{\vartheta}_{1,z}}\right) f(z,\vartheta)\mathrm{d}\mu(z) \\ \vdots \\ \int \left(\nabla_k^t \frac{\partial}{\partial \vartheta_k} L(z,\vartheta)_{|\vartheta=\overline{\vartheta}_{k,z}}\right) f(z,\vartheta)\mathrm{d}\mu(z) \end{pmatrix} (\vartheta - \vartheta_0)
\end{aligned}$$

Wegen (C1), (C2) und (C3) ist mit dem Satz von der majorisierten Konvergenz

$$\lim_{\vartheta\to\vartheta_0}\begin{pmatrix}\int\big(\nabla_k^t\frac{\partial}{\partial\vartheta_1}L(z,\vartheta)_{|_{\vartheta=\overline{\vartheta}_{1,z}}}\big)f(z,\vartheta)\mathrm{d}\mu(z)\\ \vdots\\ \int\big(\nabla_k^t\frac{\partial}{\partial\vartheta_k}L(z,\vartheta)_{|_{\vartheta=\overline{\vartheta}_{k,z}}}\big)f(z,\vartheta)\mathrm{d}\mu(z)\end{pmatrix}=-i_{\zeta_1}(\vartheta_0).$$

Damit folgt insgesamt für ϑ in einer offenen Umgebung um ϑ_0 wegen der stetigen Differenzierbarkeit von g

$$(\hat{\eta}-\eta_0)=i^*_{\zeta_1}(\eta_0)^{-1}\mathcal{J}_g(\eta_0)^t i_{\zeta_1}(\vartheta_0)(\vartheta-\vartheta_0)+R_0(\vartheta,\vartheta_0)(\vartheta-\vartheta_0),$$

mit einer Matrix $R_0(\vartheta,\vartheta_0)\in\mathbb{R}^{h\times k}$ mit $\lim_{\vartheta\to\vartheta_0}R_0(\vartheta,\vartheta_0)=0$, also die Behauptung. □

In Vorbereitung auf die nächsten Resultate definieren wir

$$i^{**}_{\zeta_1}(\eta_0):=\big(\mathrm{I}_d-\mathcal{J}_g(\eta_0)i^*_{\zeta_1}(\eta_0)^{-1}\mathcal{J}_g(\eta_0)^t i_{\zeta_1}(\vartheta_0)\big)^t i_{\zeta_1}(\vartheta_0),$$

wobei $i^*_{\zeta_1}(\eta_0)$ wie in Lemma 4.3.3 definiert ist. Im folgenden Satz wird das lokale Verhalten der Bahadur-Steigung einer Folge von Tests T, die optimal im Bahadur Sinn ist, angegeben:

4.3.4 Satz: *Es sei die Folge von Tests T optimal im Bahadur Sinn. Für ζ_1 gelten die Bedingungen (C1) - (C4). Zudem sei $\vartheta_0\in\partial\tilde{\Theta}_0$ fest gewählt und $\eta_0:=g^{-1}(\vartheta_0)$. Dann ist für ϑ in einer offenen Umgebung um ϑ_0 die Bahadur-Steigung von T gegeben durch*

$$K(\vartheta,\tilde{\Theta}_0)=\frac{1}{2}(\vartheta-\vartheta_0)^t i^{**}_{\zeta_1}(\eta_0)(\vartheta-\vartheta_0)+(\vartheta-\vartheta_0)^t R_3(\vartheta,\vartheta_0)(\vartheta-\vartheta_0),$$

mit einer Matrix $R_3(\vartheta,\vartheta_0)\in\mathbb{R}^{k\times k}$ mit $\lim_{\vartheta\to\vartheta_0}R_3(\vartheta,\vartheta_0)=0$.

Beweis: Es sei $\hat{\eta}$ gemäß (C4) gewählt. Wir definieren $\eta_0:=g^{-1}(\vartheta_0)$. Eine Taylorentwicklung der Funktion $\vartheta\mapsto L(z,\vartheta)$, $z\in R$, liefert

$$L(z,\vartheta)=L(z,\vartheta_0)+(\vartheta-\vartheta_0)^t\nabla_k L(z,\vartheta)-\frac{1}{2}(\vartheta-\vartheta_0)^t\left(\frac{\partial^2}{\partial\vartheta_i\partial\vartheta_l}L(z,\vartheta)_{|_{\vartheta=\overline{\vartheta}_z}}\right)_{\substack{1\le i\le k\\1\le l\le k}}(\vartheta-\vartheta_0),$$
$$z\in R, \tag{4.3.3}$$

mit einer Zwischenstelle $\overline{\vartheta}_z$ auf der Verbindungsstrecke zwischen ϑ und ϑ_0, $z\in R$.

Andererseits liefert eine entsprechende Taylorentwicklung

$$\begin{aligned}L\big(z,g(\hat{\eta})\big)=&L\big(z,g(\eta_0)\big)+(\hat{\eta}-\eta_0)^t\nabla_h L\big(z,g(\eta)\big)_{|_{\eta=\hat{\eta}}}\\&-\frac{1}{2}(\hat{\eta}-\eta_0)^t\left(\frac{\partial^2}{\partial\eta_i\partial\eta_l}L\big(z,g(\eta)\big)_{|_{\eta=\overline{\eta}_z}}\right)_{\substack{1\le i\le h\\1\le l\le h}}(\hat{\eta}-\eta_0),\ z\in R,\end{aligned}\tag{4.3.4}$$

mit einer Zwischenstelle $\overline{\eta}_z$ auf der Verbindungsstrecke zwischen $\hat{\eta}$ und $\eta_0,\ z \in R$. Somit folgt mit (4.3.3) und (4.3.4) für die Bahadur-Steigung von T wegen (C4) und $g(\eta_0) = \vartheta_0$ für ϑ in einer offenen Umgebung um ϑ_0

$$\begin{aligned}
K(\vartheta, \tilde{\Theta}_0) &= \int \log \frac{f(z;\vartheta)}{f\big(z; g(\hat{\eta})\big)} f(z;\vartheta)\mathrm{d}\mu(z) = \int \Big(L(z,\vartheta) - L\big(z, g(\hat{\eta})\big)\Big) f(z;\vartheta)\mathrm{d}\mu(z) \\
&= \int \Big(L(z,\vartheta) - L(z,\vartheta_0) + L(z,\vartheta_0) - L\big(z, g(\hat{\eta})\big)\Big) f(z;\vartheta)\mathrm{d}\mu(z) \\
&= \int \Big(L(z,\vartheta) - L(z,\vartheta_0) + L\big(z, g(\eta_0)\big) - L\big(z, g(\hat{\eta})\big)\Big) f(z;\vartheta)\mathrm{d}\mu(z) \\
&= \int \Bigg((\vartheta - \vartheta_0)^t \nabla_k L(z,\vartheta) - \frac{1}{2}(\vartheta - \vartheta_0)^t \bigg(\frac{\partial^2}{\partial \vartheta_i \partial \vartheta_l} L(z,\vartheta)_{|_{\vartheta = \overline{\vartheta}_z}} \bigg)_{\substack{1 \le i \le k \\ 1 \le l \le k}} (\vartheta - \vartheta_0) \\
&\quad - (\hat{\eta} - \eta_0)^t \nabla_h L\big(z, g(\eta)\big)_{|_{\eta = \hat{\eta}(\vartheta)}} \\
&\quad + \frac{1}{2}(\hat{\eta} - \eta_0)^t \bigg(\frac{\partial^2}{\partial \vartheta_i \partial \vartheta_l} L\big(z, g(\eta)\big)_{|_{\eta = \overline{\eta}_z}} \bigg)_{\substack{1 \le i \le h \\ 1 \le l \le h}} (\hat{\eta} - \eta_0) \Bigg) f(z;\vartheta)\mathrm{d}\mu(z),
\end{aligned}$$

wobei wir $\hat{\eta} = \hat{\eta}(\vartheta)$ schreiben. Es ist mit (C2)

$$0 = \int \nabla_k f(z;\vartheta)\mathrm{d}\mu(z) = \int \big(\nabla_k L(z,\vartheta)\big) f(z;\vartheta)\mathrm{d}\mu(z),$$

und wegen (C4) für ϑ in einer offenen Umgebung um ϑ_0

$$0 = -\int \nabla_h \Big(\log \frac{f(z;\vartheta)}{f\big(z; g(\eta)\big)} f(z;\vartheta) \Big)_{|_{\eta = \hat{\eta}(\vartheta)}} \mathrm{d}\mu(z) = \int \nabla_h L\big(z, g(\eta)\big)_{|_{\eta = \hat{\eta}(\vartheta)}} f(z;\vartheta)\mathrm{d}\mu(z).$$

Daher folgt für die Bahadur-Steigung von T für ϑ in einer offenen Umgebung um ϑ_0

$$\begin{aligned}
K(\vartheta, \tilde{\Theta}_0) = &-\frac{1}{2}(\vartheta - \vartheta_0)^t \bigg(\int \Big(\frac{\partial^2}{\partial \vartheta_i \partial \vartheta_l} L(z,\vartheta) \Big)_{|_{\vartheta = \overline{\vartheta}_z}} f(z;\vartheta)\mathrm{d}\mu(z) \bigg)_{\substack{1 \le i \le k \\ 1 \le l \le k}} (\vartheta - \vartheta_0) \\
&+ \frac{1}{2}(\hat{\eta} - \eta_0)^t \bigg(\int \Big(\frac{\partial^2}{\partial \eta_i \partial \eta_l} L\big(z, g(\eta)\big) \Big)_{|_{\eta = \overline{\eta}_z}} f(z;\vartheta)\mathrm{d}\mu(z) \bigg)_{\substack{1 \le i \le h \\ 1 \le l \le h}} (\hat{\eta} - \eta_0),
\end{aligned}$$

wobei die Existenz der beteiligten Integrale mit Hilfe von Lemma 4.3.2 durch (C3) und (C4) für ϑ in einer offenen Umgebung um ϑ_0 gewährleistet werden kann.

Wegen (C1), (C2) und (C3) gilt mit dem Satz von der majorisierten Konvergenz

$$\lim_{\vartheta \to \vartheta_0} \bigg(\int \Big(\frac{\partial^2}{\partial \vartheta_i \partial \vartheta_l} L(z,\vartheta) \Big)_{|_{\vartheta = \overline{\vartheta}_z}} f(z;\vartheta)\mathrm{d}\mu(z) \bigg)_{\substack{1 \le i \le k \\ 1 \le l \le k}} = -i_{\zeta_1}(\vartheta_0).$$

Mit (C1), (C2), (C4) und der zweimaligen stetigen Differenzierbarkeit von g erhält man mit Lemma 4.3.2 und dem Satz von der majorisierten Konvergenz

$$\lim_{\vartheta\to\vartheta_0}\left(\int\Big(\frac{\partial^2}{\partial\eta_i\partial\eta_l}L\big(z,g(\eta)\big)\Big)_{|\eta=\overline{\eta}_z} f(z;\vartheta)\mathrm{d}\mu(z)\right)_{\substack{1\le i\le h\\ 1\le l\le h}} = -i^*_{\zeta_1}(\eta_0).$$

Somit folgt für die Bahadur-Steigung von T für ϑ in einer offenen Umgebung um ϑ_0

$$\begin{aligned}K(\vartheta,\tilde\Theta_0) =&\frac{1}{2}(\vartheta-\vartheta_0)^t i_{\zeta_1}(\vartheta_0)(\vartheta-\vartheta_0) - \frac{1}{2}(\hat\eta-\eta_0)^t i^*_{\zeta_1}(\eta_0)(\hat\eta-\eta_0)\\ &+(\vartheta-\vartheta_0)^t R_1(\vartheta,\vartheta_0)(\vartheta-\vartheta_0) + (\hat\eta-\eta_0)^t R_2(\vartheta,\vartheta_0)(\hat\eta-\eta_0),\end{aligned}$$

mit Matrizen $R_1(\vartheta,\vartheta_0)\in\mathbb{R}^{k\times k}$ und $R_2(\vartheta,\vartheta_0)\in\mathbb{R}^{h\times h}$ mit $\lim_{\vartheta\to\vartheta_0}R_1(\vartheta,\vartheta_0)=0$ und $\lim_{\vartheta\to\vartheta_0}R_2(\vartheta,\vartheta_0)=0$. Mit Hilfe von Lemma 4.3.3 lässt sich dies für ϑ in einer offenen Umgebung um ϑ_0 gemäß

$$K(\vartheta,\tilde\Theta_0) = \frac{1}{2}(\vartheta-\vartheta_0)^t i^{**}_{\zeta_1}(\eta_0)(\vartheta-\vartheta_0) + (\vartheta-\vartheta_0)^t R_3(\vartheta,\vartheta_0)(\vartheta-\vartheta_0)$$

schreiben, mit einer Matrix $R_3(\vartheta,\vartheta_0)\in\mathbb{R}^{k\times k}$ mit $\lim_{\vartheta\to\vartheta_0}R_3(\vartheta,\vartheta_0)=0$. □

In Vorbereitung auf das nächste Resultat sei eine Matrix $A^{1,1}_{\zeta_1}(\vartheta_0)\in\mathbb{R}^{d_1\times d_1}$ durch

$$i^{**}_{\zeta_1}(\eta_0) = \begin{pmatrix}A^{1,1}_{\zeta_1}(\vartheta_0) & A^{1,2}_{\zeta_1}(\vartheta_0)\\ A^{2,1}_{\zeta_1}(\vartheta_0) & A^{2,2}_{\zeta_1}(\vartheta_0)\end{pmatrix}$$

festgelegt, $A^{1,2}_{\zeta_1}(\vartheta_0)\in\mathbb{R}^{d_1\times(k-d_1)}$, $A^{2,1}_{\zeta_1}(\vartheta_0)\in\mathbb{R}^{(k-d_1)\times d_1}$, $A^{2,2}_{\zeta_1}(\vartheta_0)\in\mathbb{R}^{(k-d_1)\times(k-d_1)}$. Nun können wir für zwei Folgen von Tests T und T', die jeweils optimal im Bahadur Sinn sind, folgendes Resultat für die lokale Bahadur-Effizienz formulieren:

4.3.5 Korollar: *Es seien die Folgen von Tests T und T' jeweils optimal im Bahadur Sinn. Für die Verteilungen von ζ_1 und ζ_1' gelten die Bedingungen (C1) - (C4). Es sei eine reellwertige Folge $(\nu_i)_{i\in\mathbb{N}}$ mit $\nu_i\neq 0$ für alle $i\in\mathbb{N}$ und $\lim_{i\to\infty}\nu_i=0$ gegeben. Für $i\in\mathbb{N}$ betrachten wir Alternativen der Form*

$$\vartheta^i := \begin{pmatrix}\vartheta_0^1\\ \vartheta_2\\ \vartheta_2'\end{pmatrix} + \nu_i\begin{pmatrix}\vartheta_1\\ 0\\ 0\end{pmatrix}$$

mit $\vartheta_1\in\mathbb{R}^{d_1}\setminus\{0\}$, $(\vartheta_0^1,\vartheta_2,\vartheta_2')\in\partial\Theta_0$, die $(\vartheta_0^1+\nu_i\vartheta_1,\vartheta_2,\vartheta_2')\in\Theta_1$ erfüllen, wobei mit $\vartheta_0 := \binom{\vartheta_0^1}{\vartheta_2}$ und $\vartheta_0' := \binom{\vartheta_0^1}{\vartheta_2'}$

$$\vartheta_1^t A^{1,1}_{\zeta_1}(\vartheta_0)\vartheta_1 \neq 0 \text{ und } \vartheta_1^t A^{1,1}_{\zeta_1'}(\vartheta_0')\vartheta_1\neq 0$$

gelte. Dann gilt mit $\vartheta^0 := (\vartheta_0^1,\vartheta_2,\vartheta_2')$

$$e^{\mathrm{LB}}_{T',T}\big(\vartheta^0,\vartheta_1\big) = \frac{\vartheta_1^t A^{1,1}_{\zeta_1'}(\vartheta_0')\vartheta_1}{\vartheta_1^t A^{1,1}_{\zeta_1}(\vartheta_0)\vartheta_1}.$$

Beweis: Zunächst betrachten wir nur die Folge von Tests T. Es sei $\vartheta_0 = \binom{\vartheta_0^1}{\vartheta_2}$, ϑ_0^1 und ϑ_2 wie in der Voraussetzung. Weiter sei $\vartheta \in \tilde{\Theta}_1$.

Mit Hilfe von Satz 4.3.4 lässt sich für ϑ in einer offenen Umgebung um ϑ_0 die Bahadur-Steigung von T gemäß

$$K(\vartheta, \tilde{\Theta}_0) = \frac{1}{2}(\vartheta - \vartheta_0)^t i_{\zeta_1}^{**}(\eta_0)(\vartheta - \vartheta_0) + (\vartheta - \vartheta_0)^t R_3(\vartheta, \vartheta_0)(\vartheta - \vartheta_0)$$

schreiben, mit einer Matrix $R_3(\vartheta, \vartheta_0) \in \mathbb{R}^{k \times k}$ mit $\lim_{\vartheta \to \vartheta_0} R_3(\vartheta, \vartheta_0) = 0$.

Somit erhät man für die Bahadur-Steigung der Folge von Tests T bei Folgen von Alternativen $(\vartheta^i)_{i \in \mathbb{N}}$ wie in der Voraussetzung beschrieben

$$c_T(\vartheta^i) = \frac{1}{2}\nu_i^2 \vartheta_1^t A_{\zeta_1}^{1,1}(\vartheta_0)\vartheta_1 + o(\nu_i^2) \text{ für } i \to \infty.$$

Analog erhält man für die Folge von Tests T'

$$c_{T'}(\vartheta^i) = \frac{1}{2}\nu_i^2 \vartheta_1^t A_{\zeta_1'}^{1,1}(\vartheta_0')\vartheta_1 + o(\nu_i^2) \text{ für } i \to \infty,$$

sodass insgesamt

$$e_{T',T}^{\mathrm{LB}}(\vartheta^0, \vartheta_1) = \lim_{i\to\infty} e_{T',T}^{\mathrm{B}}(\vartheta^i) = \lim_{i\to\infty} \frac{c_{T'}(\vartheta^i)}{c_T(\vartheta^i)} = \lim_{i\to\infty} \frac{\frac{1}{2}\nu_i^2 \vartheta_1^t A_{\zeta_1'}^{1,1}(\vartheta_0')\vartheta_1 + o(\nu_i^2)}{\frac{1}{2}\nu_i^2 \vartheta_1^t A_{\zeta_1}^{1,1}(\vartheta_0)\vartheta_1 + o(\nu_i^2)} = \frac{\vartheta_1^t A_{\zeta_1'}^{1,1}(\vartheta_0')\vartheta_1}{\vartheta_1^t A_{\zeta_1}^{1,1}(\vartheta_0)\vartheta_1}$$

folgt. □

In Vorbereitung auf das Resultat zur lokalen Hodges-Lehmann-Effizienz definieren wir folgende Bedingungen:

(C5) Es existiere für jedes $\vartheta \in \tilde{\Theta}$ und jede $1 \le i, j \le k$ ein $\delta > 0$ sodass mit $U_\delta := \{\tilde{\vartheta} \in \mathbb{R}^k : |\tilde{\vartheta} - \vartheta| < \delta\} \subset \tilde{\Theta}$, $M_\delta(z) := \sup_{\vartheta^* \in U_\delta} |\frac{\partial}{\partial \tilde{\vartheta}_i} L(z, \tilde{\vartheta})|_{\tilde{\vartheta}=\vartheta^*}| \sup_{\vartheta^* \in U_\delta} |\frac{\partial}{\partial \tilde{\vartheta}_j} L(z, \tilde{\vartheta})|_{\tilde{\vartheta}=\vartheta^*} f(z, \vartheta^*)|$, $z \in R$, $\int M_\delta(z) \mathrm{d}\mu(z) < \infty$ ist.

(C6) Es existiere für jedes $\vartheta_0 \in \partial\tilde{\Theta}_0$ und jede $1 \le i, j \le h$ ein $\delta > 0$ sodass mit $\eta_0 := g^{-1}(\vartheta_0)$, $\tilde{U}_\delta := \{\eta \in \Delta : |\eta - \eta_0| < \delta\} \subset \Delta$, $M_\delta(z) := \sup_{\eta^* \in \tilde{U}_\delta} |\frac{\partial}{\partial \eta_i} L(z, g(\eta))|_{\eta=\eta^*}| \sup_{\eta^* \in \tilde{U}_\delta} |\frac{\partial}{\partial \eta_j} L(z, g(\eta))|_{\eta=\eta^*} f(z, g(\eta^*))|$, $z \in R$, $\int M_\delta(z) \mathrm{d}\mu(z) < \infty$ ist. Zudem existiere für jedes $\vartheta_0 \in \partial\tilde{\Theta}_0$ und jede $1 \le i, j \le h$ ein $\delta > 0$ sodass mit $U_\delta := \{\vartheta \in \mathbb{R}^k : |\vartheta - \vartheta_0| < \delta\} \subset \tilde{\Theta}$, $\tilde{M}_\delta(z) := \sup_{\eta^* \in \tilde{U}_\delta} \sup_{\vartheta^* \in U_\delta} |\frac{\partial^2}{\partial \eta_i \partial \eta_j}((L(z, g(\eta)) - L(z, \vartheta^*))f(z; g(\eta)))|_{\eta=\eta^*}|$, $z \in R$, $\int \tilde{M}_\delta(z) \mathrm{d}\mu(z) < \infty$ ist.

(C7) Für alle $\vartheta, \tilde{\vartheta} \in \tilde{\Theta}$ existiere die Kullback-Leibler-Information $K(\vartheta, \tilde{\vartheta})$ aus ζ_1. Zudem existiere für jedes $\vartheta_0 \in \partial\tilde{\Theta}_0$ eine offene Umgebung U_{ϑ_0} um ϑ_0 und eine auf U_{ϑ_0} stetige Abbildung $\hat{\hat{\eta}} : U_{\vartheta_0} \to \Delta$ mit den Eigenschaften $K(\tilde{\Theta}_0, \vartheta) = K(g(\hat{\hat{\eta}}(\vartheta)), \vartheta)$ und $\int \nabla_h (\log \frac{f(z;g(\eta))}{f(z;\vartheta)} f(z;g(\eta)))|_{\eta=\hat{\hat{\eta}}(\vartheta)} \mathrm{d}\mu(z) = 0$ für alle $\vartheta \in U_{\vartheta_0}$.

Zur Verifizierung der Bedingung (C7) kann das folgende Lemma nützlich sein.

4.3.6 Lemma: *Es gelte (C1) und (C2). Weiter sei $\vartheta_0 \in \partial\tilde{\Theta}_0$ fest gewählt und $\eta_0 := g^{-1}(\vartheta_0)$. Es existiere das Integral $\int \nabla_h^t (\log \frac{f(z;g(\eta))}{f(z;\vartheta)} f(z;g(\eta))\mathrm{d}\mu(z)$ für (ϑ, η) in einer offenen Umgebung um (ϑ_0, η_0) und es sei für (ϑ, η) in dieser Umgebung um (ϑ_0, η_0)*

$$G(\vartheta, \eta) := \int \nabla_h^t \Big(\log \frac{f\big(z; g(\eta)\big)}{f(z;\vartheta)} f\big(z; g(\eta)\big)\Big) \mathrm{d}\mu(z),$$

wobei G stetig differenzierbar auf dieser offenen Umgebung um (ϑ_0, η_0) sei. Zudem gelte $\nabla_h G(\vartheta, \eta)|_{\eta=\eta_0} = \int \nabla_h \nabla_h^t (\log \frac{f(z;g(\eta))}{f(z;\vartheta_0)} f(z;g(\eta)))|_{\eta=\eta_0} \mathrm{d}\mu(z)$. Dann existiert eine offene Umgebung U_{ϑ_0} um ϑ_0 und eine eindeutige auf U_{ϑ_0} stetige Abbildung $\hat{\hat{\eta}} : U_{\vartheta_0} \to \Delta$ mit $G\big(\vartheta, \hat{\hat{\eta}}(\vartheta)\big) = 0$ für alle $\vartheta \in U_{\vartheta_0}$.

Beweis: Es ist für (ϑ, η) in einer Umgebung um (ϑ_0, η_0) mit (C2)

$$G(\vartheta, \eta) = \int \Big(\log f\big(z; g(\eta)\big) \nabla_h^t f\big(z; g(\eta)\big) - \log f(z;\vartheta) \nabla_h^t f\big(z; g(\eta)\big)\Big) \mathrm{d}\mu(z).$$

und daher wegen $\eta_0 = g^{-1}(\vartheta_0)$ $G(\vartheta_0, \eta_0) = 0$. Weiter ist mit $\eta_0 = g^{-1}(\vartheta_0)$

$$\begin{aligned}
\nabla_h G(\vartheta_0, \eta)_{|\eta=\eta_0} &= \nabla_h \int \nabla_h^t \Big(\log \frac{f\big(z; g(\eta)\big)}{f(z;\vartheta_0)} f\big(z; g(\eta)\big)\Big)_{|\eta=\eta_0} \mathrm{d}\mu(z) \\
&= \int \nabla_h \nabla_h^t \Big(\log \frac{f\big(z; g(\eta)\big)}{f(z;\vartheta_0)} f\big(z; g(\eta)\big)\Big)_{|\eta=\eta_0} \mathrm{d}\mu(z) \\
&= \int \frac{1}{f\big(z; g(\eta_0)\big)} \Big(\nabla_h f\big(z; g(\eta)\big)\Big)_{|\eta=\eta_0} \Big(\nabla_h^t f\big(z; g(\eta)\big)\Big)_{|\eta=\eta_0} \mathrm{d}\mu(z) \\
&= \int \frac{\nabla_h f\big(z; g(\eta)\big)_{|\eta=\eta_0}}{f\big(z; g(\eta_0)\big)} \frac{\nabla_h^t f\big(z; g(\eta)\big)_{|\eta=\eta_0}}{f\big(z; g(\eta_0)\big)} f(z;\vartheta_0) \mathrm{d}\mu(z) = i_{\zeta_1}^*(\eta_0).
\end{aligned}$$

Wegen (C2) ist $\nabla_h G(\vartheta_0, \eta)_{|\eta=\eta_0} = i_{\zeta_1}^*(\eta_0)$ positiv definit. Nach dem Satz von der impliziten Funktion existiert eine offene Umgebung U_{ϑ_0} um ϑ_0 und eine eindeutige auf U_{ϑ_0} stetige Abbildung $\hat{\hat{\eta}} : U_{\vartheta_0} \to \Delta$ mit $\hat{\hat{\eta}}(\vartheta_0) = \eta_0$ und

$$0 = G\big(\vartheta, \hat{\hat{\eta}}(\vartheta)\big) = \int \nabla_h^t \Big(\log \frac{f\big(z; g(\eta)\big)}{f(z;\vartheta)} f\big(z; g(\eta)\big)\Big)_{|\eta=\hat{\hat{\eta}}(\vartheta)} \mathrm{d}\mu(z) \text{ für alle } \vartheta \in U_{\vartheta_0}.$$

Damit ist alles gezeigt. □

Nun werden erneut zwei Lemmata betrachtet.

4.3.7 Lemma: *Es gelte (C1) und (C7). Zudem sei $\vartheta_0 \in \partial\tilde{\Theta}_0$ fest gewählt und $\eta_0 := g^{-1}(\vartheta_0)$. Dann gilt*

$$\lim_{\vartheta \to \vartheta_0} \hat{\hat{\eta}}(\vartheta) = \hat{\hat{\eta}}(\vartheta_0) = \eta_0.$$

Beweis: Da eine injektive Parametrisierung der Verteilung von ζ_1 durch die Elemente von $\tilde{\Theta}$ vorliegt, gilt für alle $\vartheta, \tilde{\vartheta} \in \tilde{\Theta}$ die Ungleichung $K(\vartheta, \tilde{\vartheta}) \geq 0$ sowie die Äquivalenz $K(\vartheta, \tilde{\vartheta}) = 0 \;\Leftrightarrow\; \vartheta = \tilde{\vartheta}$. Wegen der in (C7) geforderten Stetigkeit der Abbildung $\hat{\hat{\eta}}$ in ϑ_0 gilt $\lim_{\vartheta \to \vartheta_0} \hat{\hat{\eta}}(\vartheta) = \hat{\hat{\eta}}(\vartheta_0)$. Weiter folgt aus

$$0 = K(\tilde{\Theta}_0, \vartheta_0) = K\Big(g\big(\hat{\hat{\eta}}(\vartheta_0)\big), \vartheta_0\Big)$$

sofort

$$g\big(\hat{\hat{\eta}}(\vartheta_0)\big) = \vartheta_0,$$

äquivalent

$$\hat{\hat{\eta}}(\vartheta_0) = g^{-1}(\vartheta_0) = \eta_0,$$

womit alles gezeigt ist. $\square$

4.3.8 Lemma: *Für ζ_1 gelten die Bedingungen (C1),(C2),(C5),(C6) und (C7). Zudem sei $\vartheta_0 \in \partial\tilde{\Theta}_0$ fest gewählt. Wir definieren $\eta_0 := g^{-1}(\vartheta_0)$. Dann gilt für ϑ in einer offenen Umgebung um ϑ_0*

$$(\hat{\hat{\eta}} - \eta_0) = i^*_{\zeta_1}(\eta_0)^{-1} \mathcal{J}_g(\eta_0)^t i_{\zeta_1}(\vartheta_0)(\vartheta - \vartheta_0) + R_4(\vartheta, \vartheta_0)(\vartheta - \vartheta_0),$$

wobei wir $\hat{\hat{\eta}} = \hat{\hat{\eta}}(\vartheta)$ schreiben, mit einer Matrix $R_4(\vartheta, \vartheta_0) \in \mathbb{R}^{h \times k}$ mit $\lim_{\vartheta \to \vartheta_0} R_4(\vartheta, \vartheta_0) = 0$.

Beweis: Eine Taylorentwicklung der Funktion $\vartheta \mapsto L(z, \vartheta)$, $z \in R$, liefert

$$L(z, \vartheta) = L(z, \vartheta_0) + \nabla_k^t L(z, \vartheta)_{|\vartheta = \overline{\vartheta}_z} (\vartheta - \vartheta_0), \; z \in R,$$

mit einer Zwischenstelle $\overline{\vartheta}_z$ auf der Verbindungsstrecke zwischen ϑ und ϑ_0, $z \in R$.

Andererseits liefert eine entsprechende Taylorentwicklung wegen $g(\eta_0) = \vartheta_0$

$$\begin{aligned} L\big(z, g(\hat{\hat{\eta}})\big) &= L\big(z, g(\eta_0)\big) + \nabla_h^t L\big(z, g(\eta)\big)_{|\eta = \overline{\eta}_z} (\hat{\hat{\eta}} - \eta_0) \\ &= L(z, \vartheta_0) + \nabla_h^t L\big(z, g(\eta)\big)_{|\eta = \overline{\eta}_z} (\hat{\hat{\eta}} - \eta_0), \; z \in R, \end{aligned}$$

mit einer Zwischenstelle $\overline{\eta}_z$ auf der Verbindungsstrecke zwischen $\hat{\hat{\eta}}$ und η_0, $z \in R$. Durch Subtraktion der beiden Gleichungen folgt

$$L\big(z, g(\hat{\hat{\eta}})\big) - L(z, \vartheta) = \nabla_h^t L\big(z, g(\eta)\big)_{|\eta = \overline{\eta}_z} (\hat{\hat{\eta}} - \eta_0) - \nabla_k^t L(z, \vartheta)_{|\vartheta = \overline{\vartheta}_z} (\vartheta - \vartheta_0), \; z \in R,$$

äquivalent

$$\log \frac{f\big(z; g(\hat{\hat{\eta}})\big)}{f(z;\vartheta)} = \nabla_h^t L\big(z, g(\eta)\big)_{|_{\eta=\overline{\eta}_z}} (\hat{\hat{\eta}} - \eta_0) - \nabla_k^t L(z,\vartheta)_{|_{\vartheta=\overline{\vartheta}_z}} (\vartheta - \vartheta_0), \ z \in R.$$

Daraus folgt

$$\begin{aligned}\log \frac{f\big(z; g(\hat{\hat{\eta}})\big)}{f(z;\vartheta)} \nabla_h f\big(z; g(\eta)\big)_{|_{\eta=\hat{\hat{\eta}}}} =& \Big(\nabla_h f\big(z; g(\eta)\big)\Big)_{|_{\eta=\hat{\hat{\eta}}}} \nabla_h^t L\big(z, g(\eta)\big)_{|_{\eta=\overline{\eta}_z}} (\hat{\hat{\eta}} - \eta_0) \\ &- \Big(\nabla_h f\big(z; g(\eta)\big)\Big)_{|_{\eta=\hat{\hat{\eta}}}} \nabla_k^t L(z,\vartheta)_{|_{\vartheta=\overline{\vartheta}_z}} (\vartheta - \vartheta_0), \ z \in R.\end{aligned}$$

Es ist wegen (C2) für ϑ in einer offenen Umgebung um ϑ_0

$$\begin{aligned}0 &= \int \nabla_h \Big(\log \frac{f\big(z; g(\eta)\big)}{f(z;\vartheta)} f\big(z; g(\eta)\big)\Big)|_{\eta=\hat{\hat{\eta}}} \, \mathrm{d}\mu(z) \\ &= \int \log \frac{f\big(z; g(\hat{\hat{\eta}})\big)}{f(z;\vartheta)} \nabla_h f\big(z; g(\eta)\big)|_{\eta=\hat{\hat{\eta}}} \, \mathrm{d}\mu(z)\end{aligned}$$

sodass durch Integrieren für ϑ in einer offenen Umgebung um ϑ_0

$$\begin{aligned}&\int \Big(\nabla_h f\big(z; g(\eta)\big)\Big)_{|_{\eta=\hat{\hat{\eta}}}} \nabla_h^t L\big(z, g(\eta)\big)_{|_{\eta=\overline{\eta}_z}} \, \mathrm{d}\mu(z) (\hat{\hat{\eta}} - \eta_0) \\ =& \int \Big(\nabla_h f\big(z; g(\eta)\big)\Big)_{|_{\eta=\hat{\hat{\eta}}}} \nabla_k^t L(z,\vartheta)_{|_{\vartheta=\overline{\vartheta}_z}} \, \mathrm{d}\mu(z) (\vartheta - \vartheta_0),\end{aligned}$$

äquivalent

$$\begin{aligned}&\int \Big(\nabla_h L\big(z, g(\eta)\big)\Big)_{|_{\eta=\hat{\hat{\eta}}}} \nabla_h^t L\big(z, g(\eta)\big)_{|_{\eta=\overline{\eta}_z}} f\big(z; g(\hat{\hat{\eta}})\big) \, \mathrm{d}\mu(z) (\hat{\hat{\eta}} - \eta_0) \\ =& \mathcal{J}_g(\hat{\hat{\eta}})^t \int \Big(\nabla_k L(z,\vartheta)\Big)_{|_{\vartheta=g(\hat{\hat{\eta}})}} \nabla_k^t L(z,\vartheta)_{|_{\vartheta=\overline{\vartheta}_z}} f\big(z; g(\hat{\hat{\eta}})\big) \, \mathrm{d}\mu(z) (\vartheta - \vartheta_0),\end{aligned}$$

folgt, wobei die Existenz der beteiligten Integrale mit Lemma 4.3.7 durch (C5) und (C6) für ϑ in einer offenen Umgebung um ϑ_0 gewährleistet werden kann.

Mit Lemma 4.3.7 und (C6) gilt mit dem Satz von der majorisierten Konvergenz

$$\lim_{\vartheta \to \vartheta_0} \int \Big(\nabla_h L\big(z, g(\eta)\big)\Big)_{|_{\eta=\hat{\hat{\eta}}}} \nabla_h^t L\big(z, g(\eta)\big)_{|_{\eta=\overline{\eta}_z}} f\big(z; g(\hat{\hat{\eta}})\big) \, \mathrm{d}\mu(z) = i_{\zeta_1}^*(\eta_0).$$

Wegen der Invertierbarkeit von $i_{\zeta_1}^*(\eta_0)$ ist daher für ϑ in einer offenen Umgebung um ϑ_0 die Matrix

$$\int \Big(\nabla_h L\big(z, g(\eta)\big)\Big)_{|_{\eta=\hat{\hat{\eta}}}} \nabla_h^t L\big(z, g(\eta)\big)_{|_{\eta=\overline{\eta}_z}} f\big(z; g(\hat{\hat{\eta}})\big) \, \mathrm{d}\mu(z)$$

invertierbar. Somit gilt für ϑ genügend nahe an ϑ_0

$$(\hat{\eta}-\eta_0) = \Bigg(\int \Big(\nabla_h L\big(z,g(\eta)\big)\Big)_{|_{\eta=\hat{\hat{\eta}}}} \nabla_h^t L\big(z,g(\eta)\big)_{|_{\eta=\overline{\eta}_z}} f\big(z;g(\hat{\hat{\eta}})\big)\,\mathrm{d}\mu(z)\Bigg)^{-1}$$
$$\mathcal{J}_g(\hat{\hat{\eta}})^t \int \Big(\nabla_k L(z,\vartheta)\Big)_{|_{\vartheta=g(\hat{\hat{\eta}})}} \nabla_k^t L(z,\vartheta)_{|_{\vartheta=\overline{\vartheta}_z}} f\big(z;g(\hat{\hat{\eta}})\big)\,\mathrm{d}\mu(z)(\vartheta-\vartheta_0).$$

Wegen Lemma 4.3.7 und (C5) ist mit dem Satz von der majorisierten Konvergenz

$$\lim_{\vartheta\to\vartheta_0} \int \Big(\nabla_k L(z,\vartheta)\Big)_{|_{\vartheta=g(\hat{\hat{\eta}})}} \nabla_k^t L(z,\vartheta)_{|_{\vartheta=\overline{\vartheta}_z}} f\big(z;g(\hat{\hat{\eta}})\big)\,\mathrm{d}\mu(z) = i_{\zeta_1}(\vartheta_0).$$

Damit folgt insgesamt für ϑ in einer offenen Umgebung um ϑ_0 wegen der stetigen Differenzierbarkeit von g

$$(\hat{\eta}-\eta_0) = i_{\zeta_1}^*(\eta_0)^{-1}\mathcal{J}_g(\eta_0)^t i_{\zeta_1}(\vartheta_0)(\vartheta-\vartheta_0) + R_4(\vartheta,\vartheta_0)(\vartheta-\vartheta_0),$$

mit einer Matrix $R_4(\vartheta,\vartheta_0)\in\mathbb{R}^{h\times k}$ mit $\lim_{\vartheta\to\vartheta_0} R_4(\vartheta,\vartheta_0)=0$, also die Behauptung. $\square$

Im folgenden Satz wird das lokale Verhalten des Hodges-Lehmann-Index einer Folge von Tests T, die optimal im Hodges-Lehmann Sinn ist, angegeben.

4.3.9 Satz: *Es sei die Folge von Tests T optimal im Hodges-Lehmann Sinn. Für ζ_1 gelten die Bedingungen (C1),(C2),(C3),(C5),(C6) und (C7). Zudem sei $\vartheta_0\in\partial\tilde{\Theta}_0$ fest gewählt und $\eta_0 := g^{-1}(\vartheta_0)$. Dann ist für ϑ in einer offenen Umgebung um ϑ_0 der Hodges-Lehmann-Index von T gegeben durch*

$$K(\tilde{\Theta}_0,\vartheta) = \frac{1}{2}(\vartheta-\vartheta_0)^t i_{\zeta_1}^{**}(\eta_0)(\vartheta-\vartheta_0) + (\vartheta-\vartheta_0)^t R_7(\vartheta,\vartheta_0)(\vartheta-\vartheta_0),$$

mit einer Matrix $R_7(\vartheta,\vartheta_0)\in\mathbb{R}^{k\times k}$ mit $\lim_{\vartheta\to\vartheta_0} R_7(\vartheta,\vartheta_0)=0$.

Beweis: Es sei $\hat{\hat{\eta}}$ gemäß (C7) gewählt. Wir definieren $\eta_0 := g^{-1}(\vartheta_0)$. Eine Taylorentwicklung der Funktion $\eta\mapsto (L(z,g(\eta))-L(z,\vartheta))f(z;g(\eta))$, $z\in R$, liefert mit $g(\eta_0)=\vartheta_0$

$$\begin{aligned}
&\Big(L\big(z,g(\hat{\hat{\eta}})\big)-L(z,\vartheta)\Big)f\big(z;g(\hat{\hat{\eta}})\big)\\
=&\Big(L\big(z,g(\eta_0)\big)-L(z,\vartheta)\Big)f\big(z;g(\eta_0)\big) + (\hat{\hat{\eta}}-\eta_0)^t\nabla_h\bigg(\Big(L\big(z,g(\eta)\big)-L(z,\vartheta)\Big)f\big(z;g(\eta)\big)\bigg)_{|_{\eta=\hat{\hat{\eta}}}}\\
&-\frac{1}{2}(\hat{\hat{\eta}}-\eta_0)^t\nabla_h^t\nabla_h\bigg(\Big(L\big(z,g(\eta)\big)-L(z,\vartheta)\Big)f\big(z;g(\eta)\big)\bigg)_{|_{\eta=\overline{\eta}_z}}(\hat{\hat{\eta}}-\eta_0)\\
=&\big(L(z,\vartheta_0)-L(z,\vartheta)\big)f(z;\vartheta_0) + (\hat{\hat{\eta}}-\eta_0)^t\nabla_h\bigg(\Big(L\big(z,g(\eta)\big)-L(z,\vartheta)\Big)f\big(z;g(\eta)\big)\bigg)_{|_{\eta=\hat{\hat{\eta}}}}\\
&-\frac{1}{2}(\hat{\hat{\eta}}-\eta_0)^t\nabla_h^t\nabla_h\bigg(\Big(L\big(z,g(\eta)\big)-L(z,\vartheta)\Big)f\big(z;g(\eta)\big)\bigg)_{|_{\eta=\overline{\eta}_z}}(\hat{\hat{\eta}}-\eta_0),\ z\in R,
\end{aligned}$$

mit einer Zwischenstelle $\overline{\eta}_z$ auf der Verbindungsstrecke zwischen $\hat{\hat{\eta}}$ und η_0, $z \in R$, wobei wir $\hat{\hat{\eta}} = \hat{\hat{\eta}}(\vartheta)$ schreiben.

Andererseits liefert eine Taylorentwicklung der Funktion $\vartheta \mapsto L(z,\vartheta)$, $z \in R$,

$$\begin{aligned}L(z,\vartheta) =& L(z,\vartheta_0) + (\vartheta-\vartheta_0)^t \nabla_k L(z,\vartheta)_{|\vartheta=\vartheta_0}\\ &+ \frac{1}{2}(\vartheta-\vartheta_0)^t \left(\frac{\partial^2}{\partial\vartheta_i\partial\vartheta_l} L(z,\vartheta)_{|_{\vartheta=\overline{\vartheta}_z}}\right)_{\substack{1\le i\le k\\ 1\le l\le k}} (\vartheta-\vartheta_0),\ z \in R,\end{aligned}$$

mit einer Zwischenstelle $\overline{\vartheta}_z$ auf der Verbindungsstrecke zwischen ϑ und ϑ_0, $z \in R$. Dies liefert insgesamt

$$\begin{aligned}&\Big(L\big(z, g(\hat{\hat{\eta}})\big) - L(z,\vartheta)\Big) f\big(z; g(\hat{\hat{\eta}})\big)\\ =& -(\vartheta-\vartheta_0)^t \nabla_k f(z;\vartheta)|_{\vartheta=\vartheta_0} - \frac{1}{2}(\vartheta-\vartheta_0)^t \left(\frac{\partial^2}{\partial\vartheta_i\partial\vartheta_l} L(z,\vartheta)_{|_{\vartheta=\overline{\vartheta}_z}}\right)_{\substack{1\le i\le k\\ 1\le l\le k}} f(z;\vartheta_0)(\vartheta-\vartheta_0)\\ &+ (\hat{\hat{\eta}}-\eta_0)^t \nabla_h \bigg(\Big(L\big(z,g(\eta)\big) - L(z,\vartheta)\Big) f\big(z;g(\eta)\big)\bigg)_{|_{\eta=\hat{\hat{\eta}}}}\\ &- \frac{1}{2}(\hat{\hat{\eta}}-\eta_0)^t \nabla_h^t \nabla_h \bigg(\Big(L\big(z,g(\eta)\big) - L(z,\vartheta)\Big) f\big(z;g(\eta)\big)\bigg)_{|_{\eta=\overline{\eta}_z}} (\hat{\hat{\eta}}-\eta_0),\ z \in R, \qquad (4.3.5)\end{aligned}$$

Es gilt wegen (C7) für ϑ in einer offenen Umgebung um ϑ_0

$$0 = \int (\hat{\hat{\eta}}-\eta_0)^t \nabla_h \bigg(\Big(L\big(z,g(\eta)\big) - L(z,\vartheta)\Big) f\big(z;g(\eta)\big)\bigg)_{|_{\eta=\hat{\hat{\eta}}}} \mathrm{d}\mu(z),$$

und deshalb folgt aus (4.3.5) mit (C2) für den Hodges-Lehmann-Index von T für ϑ in einer offenen Umgebung um ϑ_0

$$\begin{aligned}K(\tilde{\Theta}_0,\vartheta) =& \int \Big(L\big(z,g(\hat{\hat{\eta}})\big) - L(z,\vartheta)\Big) f\big(z;g(\hat{\hat{\eta}})\big) \mathrm{d}\mu(z)\\ =& -\frac{1}{2}(\vartheta-\vartheta_0)^t \left(\int \frac{\partial^2}{\partial\vartheta_i\partial\vartheta_l} L(z,\vartheta)_{|_{\vartheta=\overline{\vartheta}_z}} f(z;\vartheta_0)\mathrm{d}\mu(z)\right)_{\substack{1\le i\le k\\ 1\le l\le k}} (\vartheta-\vartheta_0)\\ &- \frac{1}{2}(\hat{\hat{\eta}}-\eta_0)^t \int \nabla_h^t \nabla_h \bigg(\Big(L\big(z,g(\eta)\big) - L(z,\vartheta)\Big) f\big(z;g(\eta)\big)\bigg)_{|_{\eta=\overline{\eta}_z}} \mathrm{d}\mu(z)(\hat{\hat{\eta}}-\eta_0),\end{aligned}$$

wobei die Existenz der beteiligten Integrale mit Hilfe von Lemma 4.3.7 durch (C3) und (C5) für ϑ in einer offenen Umgebung um ϑ_0 gewährleistet werden kann.

Wegen (C1), (C2) und (C3) gilt mit dem Satz von der majorisierten Konvergenz

$$\lim_{\vartheta\to\vartheta_0} \left(\int \frac{\partial^2}{\partial\vartheta_i\partial\vartheta_l} L(z,\vartheta)_{|_{\vartheta=\overline{\vartheta}_z}} f(z;\vartheta_0)\mathrm{d}\mu(z)\right)_{\substack{1\le i\le k\\ 1\le l\le k}} = -i_{\zeta_1}(\vartheta_0).$$

Mit (C1), (C2), (C6) und der zweimaligen stetigen Differenzierbarkeit von g erhält man mit Lemma 4.3.7 und dem Satz von der majorisierten Konvergenz

$$\lim_{\vartheta\to\vartheta_0}\int \nabla_h^t\nabla_h\Big(\big(L\big(z,g(\eta)\big)-L(z,\vartheta)\big)f\big(z;g(\eta)\big)\Big)_{|\eta=\overline{\eta}_z}\,\mathrm{d}\mu(z)=i^*_{\zeta_1}(\eta_0).$$

Somit folgt für den Hodges-Lehmann-Index von T für ϑ in einer offenen Umgebung um ϑ_0

$$\begin{aligned}K(\tilde\Theta_0,\vartheta)=&\frac{1}{2}(\vartheta-\vartheta_0)^t i_{\zeta_1}(\vartheta_0)(\vartheta-\vartheta_0)-\frac{1}{2}(\hat\eta-\eta_0)^t i^*_{\zeta_1}(\eta_0)(\hat\eta-\eta_0)\\&+(\vartheta-\vartheta_0)^t R_5(\vartheta,\vartheta_0)(\vartheta-\vartheta_0)+(\hat\eta-\eta_0)^t R_6(\vartheta,\vartheta_0)(\hat\eta-\eta_0),\end{aligned}$$

mit Matrizen $R_5(\vartheta,\vartheta_0)\in\mathbb{R}^{k\times k}$ und $R_6(\vartheta,\vartheta_0)\in\mathbb{R}^{h\times h}$ mit $\lim_{\vartheta\to\vartheta_0}R_5(\vartheta,\vartheta_0)=0$ und $\lim_{\vartheta\to\vartheta_0}R_6(\vartheta,\vartheta_0)=0$. Mit Hilfe von Lemma 4.3.8 lässt sich dies für ϑ in einer offenen Umgebung um ϑ_0 gemäß

$$K(\tilde\Theta_0,\vartheta)=\frac{1}{2}(\vartheta-\vartheta_0)^t i^{**}_{\zeta_1}(\eta_0)(\vartheta-\vartheta_0)+(\vartheta-\vartheta_0)^t R_7(\vartheta,\vartheta_0)(\vartheta-\vartheta_0)$$

schreiben, mit einer Matrix $R_7(\vartheta,\vartheta_0)\in\mathbb{R}^{k\times k}$ mit $\lim_{\vartheta\to\vartheta_0}R_7(\vartheta,\vartheta_0)=0$. □

Nun können wir für zwei Folgen von Tests T und T', die jeweils optimal im Hodges-Lehmann Sinn sind, folgendes Resultat für die lokale Hodges-Lehmann-Effizienz formulieren:

4.3.10 Korollar: *Es seien die Folgen von Tests T und T' jeweils optimal im Hodges-Lehmann Sinn. Für ζ_1 und ζ_1' gelten die Bedingungen (C1),(C2),(C3),(C5),(C6) und (C7). Es sei eine reellwertige Folge $(\nu_i)_{i\in\mathbb{N}}$ mit $\nu_i\neq 0$ für alle $i\in\mathbb{N}$ und $\lim_{i\to\infty}\nu_i=0$ gegeben. Für $i\in\mathbb{N}$ betrachten wir Alternativen der Form*

$$\vartheta^i:=\begin{pmatrix}\vartheta_0^1\\ \vartheta_2\\ \vartheta_2'\end{pmatrix}+\nu_i\begin{pmatrix}\vartheta_1\\ 0\\ 0\end{pmatrix}$$

mit $\vartheta_1\in\mathbb{R}^{d_1}\setminus\{0\}$, $(\vartheta_0^1,\vartheta_2,\vartheta_2')\in\partial\Theta_0$, die $(\vartheta_0^1+\nu_i\vartheta_1,\vartheta_2,\vartheta_2')\in\Theta_1$ erfüllen, wobei mit $\vartheta_0:=\binom{\vartheta_0^1}{\vartheta_2}$ und $\vartheta_0':=\binom{\vartheta_0^1}{\vartheta_2'}$

$$\vartheta_1^t A^{1,1}_{\zeta_1}(\vartheta_0)\vartheta_1\neq 0 \text{ und } \vartheta_1^t A^{1,1}_{\zeta_1'}(\vartheta_0')\vartheta_1\neq 0$$

gelte. Dann gilt mit $\vartheta^0:=(\vartheta_0^1,\vartheta_2,\vartheta_2')$

$$e^{\mathrm{LHL}}_{T',T}\big(\vartheta^0,\vartheta_1\big)=\frac{\vartheta_1^t A^{1,1}_{\zeta_1'}(\vartheta_0')\vartheta_1}{\vartheta_1^t A^{1,1}_{\zeta_1}(\vartheta_0)\vartheta_1}.$$

Beweis: Zunächst betrachten wir nur die Folge von Tests T. Es sei $\vartheta_0=\binom{\vartheta_0^1}{\vartheta_2}$, ϑ_0^1 und ϑ_2

wie in der Voraussetzung. Weiter sei $\vartheta \in \tilde{\Theta}_1$.

Mit Hilfe von Satz 4.3.9 lässt sich für ϑ in einer offenen Umgebung um ϑ_0 der Hodges-Lehmann-Index von T gemäß

$$K(\tilde{\Theta}_0, \vartheta) = \frac{1}{2}(\vartheta - \vartheta_0)^t i_{\zeta_1}^{**}(\eta_0)(\vartheta - \vartheta_0) + (\vartheta - \vartheta_0)^t R_7(\vartheta, \vartheta_0)(\vartheta - \vartheta_0)$$

schreiben, mit einer Matrix $R_7(\vartheta, \vartheta_0) \in \mathbb{R}^{k \times k}$ mit $\lim_{\vartheta \to \vartheta_0} R_7(\vartheta, \vartheta_0) = 0$.

Somit erhät man für den Hodges-Lehmann-Index der Folge von Tests T bei Folgen von Alternativen $(\vartheta^i)_{i \in \mathbb{N}}$ wie in der Voraussetzung beschrieben

$$d_T(\vartheta^i) = \frac{1}{2}\nu_i^2 \vartheta_1^t A_{\zeta_1}^{1,1}(\vartheta_0)\vartheta_1 + o(\nu_i^2) \text{ für } i \to \infty.$$

Analog erhält man für die Folge von Tests T'

$$d_{T'}(\vartheta^i) = \frac{1}{2}\nu_i^2 \vartheta_1^t A_{\zeta_1'}^{1,1}(\vartheta_0')\vartheta_1 + o(\nu_i^2) \text{ für } i \to \infty,$$

sodass insgesamt

$$e_{T',T}^{\mathrm{LHL}}(\vartheta^0, \vartheta_1) = \lim_{i \to \infty} e_{T',T}^{\mathrm{HL}}(\vartheta^i) = \lim_{i \to \infty} \frac{d_{T'}(\vartheta^i)}{d_T(\vartheta^i)} = \lim_{i \to \infty} \frac{\frac{1}{2}\nu_i^2 \vartheta_1^t A_{\zeta_1'}^{1,1}(\vartheta_0')\vartheta_1 + o(\nu_i^2)}{\frac{1}{2}\nu_i^2 \vartheta_1^t A_{\zeta_1}^{1,1}(\vartheta_0)\vartheta_1 + o(\nu_i^2)} = \frac{\vartheta_1^t A_{\zeta_1'}^{1,1}(\vartheta_0')\vartheta_1}{\vartheta_1^t A_{\zeta_1}^{1,1}(\vartheta_0)\vartheta_1}$$

folgt. □

4.3.11 Bemerkung: Die in Korollar 4.2.5 gefundene Pitman-Effizienz stimmt im Wesentlichen mit der in Korollar 4.3.5 gefundenen lokalen Bahadur-Effizienz und mit der in Korollar 4.3.10 gefundenen lokalen Hodges-Lehmann-Effizienz überein. Das Resultat aus Korollar 4.2.5 hat Gültigkeit für Likelihood-Quotienten-Tests unter bestimmten Regularitätsannahmen. Die Resultate aus den Korollaren 4.3.5 und 4.3.10 haben dagegen Gültigkeit für optimale Tests im Bahadur- bzw. Hodges-Lehmann Sinn unter bestimmten Regularitätsannahmen.

In der Tat finden sich in der Literatur einige Arbeiten, die sich mit der Optimalität von Likelihood-Quotienten-Tests im Bahadur- bzw. Hodges-Lehmann beschäftigen. Dabei können die Arbeiten von Bahadur ([3]), Baringhaus ([5]), Kallenberg ([19]), Arcones ([2]) und Rublík ([35]) genannt werden.

4.3.12 Beispiel: Betrachten wir in der Situation aus Beispiel 4.2.9 anstelle von Likelihood-Quotienten-Tests optimale Tests im Bahadur Sinn bzw. Hodges-Lehmann Sinn, z.B. die in Abschnitt 8 betrachteten Tests, so gilt mit den dortigen Bezeichnungen

$$e_{T',T}^{\mathrm{LB}}(\vartheta^0) = \frac{\Gamma}{2\Sigma} \text{ bzw. } e_{T',T}^{\mathrm{LHL}}(\vartheta^0) = \frac{\Gamma}{2\Sigma}.$$

Betrachten wir im Modell aus Beispiel 4.2.10 anstelle von Likelihood-Quotienten-Tests optimale Tests im Bahadur Sinn bzw. Hodges-Lehmann Sinn, z.B. die in Abschnitt 8 betrachteten Tests, so gilt mit den dortigen Bezeichnungen

$$e^{\mathrm{LB}}_{T',T}(\vartheta^0, \vartheta_1) = \frac{\vartheta_1^t \Gamma_1^{-1} \vartheta_1}{\vartheta_1^t \Gamma_2^{-1} \vartheta_1} \text{ bzw. } e^{\mathrm{LHL}}_{T',T}(\vartheta^0, \vartheta_1) = \frac{\vartheta_1^t \Gamma_1^{-1} \vartheta_1}{\vartheta_1^t \Gamma_2^{-1} \vartheta_1}.$$

4.4 Volumen-Effizienz bei Likelihood-Quotienten-Tests und Waldschen Tests

Nun wird ein allgemeines Resultat erarbeitet das erlaubt, die Volumen- Effizienz von zwei Tests explizit anzugeben, wenn es sich bei den Tests jeweils um den Likelihood-Quotienten-Test oder den Waldschen Test handelt. Zur Anwendung kann es in den Abschnitten 8 und 9 kommen.

In diesem Unterabschnitt sei für ein gegebenes $\vartheta_0^1 \in \Theta^1$

$$\{\vartheta_0^1\} = \Theta_0^1.$$

Das in Unterabschnitt 3 vorgestellte Konzept der Volumen-Effizienz kann in offensichtlicher Weise auf die vorliegende Situation der (im Allgemeinen) zusammengesetzten Hypothese erweitert werden, da hier das Testproblem gemäß

$$\mathrm{H}: \ \vartheta_1 = \vartheta_0^1, \ \mathrm{K}: \ \vartheta_1 \in \Theta_1^1$$

äquivalent formuliert werden kann. Die Volumen-Effizienzen und die Betrachtungen in diesem Zusammenhang basieren dann auf Konfidenzbereichsschätzer für den unbekannten Parameter $\vartheta_1 \in \Theta^1$. Die weiteren Parameter ϑ_2 und ϑ_2' mit $(\vartheta_1, \vartheta_2, \vartheta_2') \in \Theta$ sind Nebenparameter.

Im Folgenden wollen wir Annahmen an die Zufallsvariablen ζ_1 und ζ_1' bei den zugrunde liegenden injektiven Parametrisierungen durch $\tilde{\Theta}$ und $\tilde{\Theta}'$ formulieren. Der Einfachheit halber formulieren wir diese Annahmen nur für ζ_1 und $\tilde{\Theta}$; sie sollen natürlich in analoger Weise auch für ζ_1' und $\tilde{\Theta}'$ gelten. Dabei werden die im Zusammenhang mit ζ_1 und $\tilde{\Theta}$ eingeführten Größen im Zusammenhang mit ζ_1' und $\tilde{\Theta}'$ mit einem Strich versehen.

Wir nehmen an, dass ein σ-endliches Maß μ auf $(R, \mathfrak{S})$ existiert, sodass für jedes $\vartheta \in \tilde{\Theta}$ $Q_\vartheta^{\zeta_1} \ll \mu$ gilt, und $f(\cdot;\vartheta)$ eine μ-Dichte von $Q_\vartheta^{\zeta_1}$ ist. Zur Vereinfachung der Notation definieren wir

$$L(z, \vartheta) := \log f(z; \vartheta), \ z \in R, \ \vartheta \in \tilde{\Theta},$$

mit $\log 0 := -\infty$. Wir setzen generell voraus, dass $f(z;\vartheta) > 0$ für alle $\vartheta \in \tilde{\Theta}$ und alle $z \in R$.

Dann betrachten wir für $n \in \mathbb{N}$ die Schätzfunktionen $\hat{\vartheta}_n : (R^n, \mathfrak{S}^n) \to (\tilde{\Theta}, \mathfrak{B}^k_{\tilde{\Theta}})$ und $\hat{\hat{\vartheta}}_n : (R^n \times \Theta^1, \mathfrak{S}^n \otimes \mathfrak{B}^{d_1}_{\Theta^1}) \to (\tilde{\Theta}, \mathfrak{B}^k_{\tilde{\Theta}})$, die in typischen Anwendungsfällen die Maximum-Likelihood-Schätzfunktionen sind. Um die Notation zu vereinfachen, schreiben wir oft $\hat{\vartheta}_n = \hat{\vartheta}_n(\zeta_1, \ldots, \zeta_n)$ und bei gegebenem $\vartheta_1^* \in \Theta^1$ auch $\hat{\hat{\vartheta}}_n^{\vartheta_1^*} = \hat{\hat{\vartheta}}_n(\zeta_1, \ldots, \zeta_n, \vartheta_1^*)$ sowie $\hat{\hat{\vartheta}}_n^{\vartheta_1^*} = (\vartheta_1^*, \hat{\hat{\vartheta}}_{2,n}^{\vartheta_1^*})$. In typischen Anwendungsfällen handelt es sich hierbei einerseits um den Maximum-Likelihood-Schätzer, andererseits um den Maximum-Likelihood-Schätzer unter der Hypothese, dass für den zugrunde liegenden Parameter $\vartheta = (\vartheta_1, \vartheta_2) \in \tilde{\Theta}$ die Gleichheit $\vartheta_1 = \vartheta_1^*$ gilt. Für $\vartheta_1^* = \vartheta_0^1$ handelt es sich also in typischen Anwendungsfällen um den Maximum-Likelihood-Schätzer unter der Hypothese H.

Betrachte nun die folgenden Bedingungen:

(D1) Es sei die Abbildung $\vartheta \mapsto f(z; \vartheta)$ für alle $z \in R$ zweimal stetig differenzierbar.

(D2) Für alle $1 \leq i, j \leq k$ und für alle $\vartheta \in \tilde{\Theta}$ gelte $\int \frac{\partial}{\partial \vartheta_i} f(z; \vartheta) \mathrm{d}\mu(z) = \int \frac{\partial^2}{\partial \vartheta_i \partial \vartheta_j} f(z; \vartheta) \mathrm{d}\mu(z) = 0$ und für alle $\vartheta \in \tilde{\Theta}$ existiere die Fisher-Informationsmatrix aus $Q_\vartheta^{\zeta_1}$, $i_{\zeta_1}(\vartheta)$, besitze endliche Einträge und sei positiv definit.

(D3) Es existiere für jedes $\vartheta \in \tilde{\Theta}$ und jedes $1 \leq i \leq k$ ein $\delta > 0$ sodass mit $U_\delta := \{\tilde{\vartheta} \in \mathbb{R}^k : |\tilde{\vartheta} - \vartheta| < \delta\} \subset \tilde{\Theta}$ und $M_\delta(z) := \sup_{\vartheta^* \in U_\delta} |\frac{\partial}{\partial \tilde{\vartheta}_i} L(z, \tilde{\vartheta})|_{\tilde{\vartheta} = \vartheta^*}| \sup_{\vartheta^* \in U_\delta} f(z; \vartheta^*)$, $z \in R$, $\int M_\delta(z) \mathrm{d}\mu(z) < \infty$ gilt. Genauso existiere für jedes $\vartheta \in \tilde{\Theta}$ und jede $1 \leq i, j \leq k$ ein $\delta > 0$ sodass mit $\tilde{M}_\delta(z) := \sup_{\vartheta^* \in U_\delta} |\frac{\partial^2}{\partial \tilde{\vartheta}_i \partial \tilde{\vartheta}_j} L(z, \tilde{\vartheta})|_{\tilde{\vartheta} = \vartheta^*}| \sup_{\vartheta^* \in U_\delta} f(z; \vartheta^*)$, $z \in R$, $\int \tilde{M}_\delta(z) \mathrm{d}\mu(z) < \infty$ gilt.

(D4) Es sei für $n \in \mathbb{N}$ $\hat{\vartheta}_n : (R^n, \mathfrak{S}^n) \to (\tilde{\Theta}, \mathfrak{B}^k_{\tilde{\Theta}})$, sodass für alle $\vartheta \in \tilde{\Theta}$ $\lim_{n\to\infty} \hat{\vartheta}_n(\zeta_1, ..., \zeta_n) = \vartheta$ in Q_ϑ-Wahrscheinlichkeit gilt, sowie für $1 \leq i \leq k$ auch $\lim_{n\to\infty} \frac{1}{\sqrt{n}} \sum_{j=0}^n \frac{\partial}{\partial \tilde{\vartheta}_i} L(\zeta_j, \tilde{\vartheta})|_{\tilde{\vartheta} = \hat{\vartheta}_n(\zeta_1, ..., \zeta_n)} = 0$ in Q_ϑ-Wahrscheinlichkeit gilt.

(D5) Es sei $\vartheta = (\vartheta_1, \vartheta_2) \in \tilde{\Theta}$. Dann sei für genügend großes $n \in \mathbb{N}$ $\hat{\hat{\vartheta}}_n : (R^n \times \Theta^1, \mathfrak{S}^n \otimes \mathfrak{B}^{d_1}_{\Theta^1}) \to (\tilde{\Theta}, \mathfrak{B}^k_{\tilde{\Theta}})$, sodass für alle $\vartheta_1^* \in \Theta^1$ $\lim_{n\to\infty} \hat{\hat{\vartheta}}_n^{\vartheta_1^n}(\zeta_1, ..., \zeta_n) = \vartheta$ in Q_ϑ-Wahrscheinlichkeit gilt, mit $\vartheta_1^n := \vartheta_1 + \vartheta_1^*/\sqrt{n}$, sowie für $1 \leq i \leq k - d_1$ $\lim_{n\to\infty} \frac{1}{\sqrt{n}} \sum_{j=0}^n \frac{\partial}{\partial \tilde{\vartheta}_{d_1+i}} L(\zeta_j, \tilde{\vartheta})\Big|_{\tilde{\vartheta} = \hat{\hat{\vartheta}}_n^{\vartheta_1^n}(\zeta_1, ..., \zeta_n)} = 0$ in Q_ϑ-Wahrscheinlichkeit.

Die Bedingungen (D1) - (D3) sind identisch mit den Bedingungen (B1) - (B3) aus Unterabschnitt 4.2 und den Bedingungen (C1) - (C3) aus Unterabschnitt 4.3. Bei der Bedingung (D4) bzw. (D5) handelt es sich um eine abgeschwächte bzw. modifizierte Version der Bedingung (B4) bzw. (B5) aus Unterabschnitt 4.2.

Es ist für $\vartheta \in \tilde{\Theta}$, mit $\vartheta = (\vartheta_1, \ldots, \vartheta_k)^t \in \mathbb{R}^k$, die Fisher-Informationsmatrix aus $Q_\vartheta^{\zeta_1}$

konkret gegeben durch

$$i_{\zeta_1}(\vartheta) = \left(\mathrm{E}_\vartheta \left(\frac{\partial}{\partial \vartheta_i} L(\zeta_1, \vartheta) \frac{\partial}{\partial \vartheta_j} L(\zeta_1, \vartheta) \right) \right)_{\substack{1 \le i \le k \\ 1 \le j \le k}}.$$

Wir wollen in diesem Unterabschnitt für $\vartheta_1^* \in \Theta^1$ die Testgröße

$$T_n(\vartheta_1^*) = 2\Big(\sum_{i=1}^n L\big(\zeta_i, \hat{\vartheta}_n(\zeta_1, \dots, \zeta_n)\big) - \sum_{i=1}^n L\big(\zeta_i, \hat{\hat{\vartheta}}_n^{\vartheta_1^*}(\zeta_1, \dots, \zeta_n)\big)\Big), \tag{4.4.1}$$

die in typischen Anwendungsfällen die Likelihood-Quotienten-Testgröße ist, oder die Testgröße

$$\begin{aligned} T_n(\vartheta_1^*) =& \sqrt{n}\big(\hat{\vartheta}_n(\zeta_1, \dots, \zeta_n) - \hat{\hat{\vartheta}}_n^{\vartheta_1^*}(\zeta_1, \dots, \zeta_n)\big)^t i_{\zeta_1}\big(\hat{\hat{\vartheta}}_n^{\vartheta_1^*}(\zeta_1, \dots, \zeta_n)\big) \\ & \sqrt{n}\big(\hat{\vartheta}_n(\zeta_1, \dots, \zeta_n) - \hat{\hat{\vartheta}}_n^{\vartheta_1^*}(\zeta_1, \dots, \zeta_n)\big), \end{aligned} \tag{4.4.2}$$

die in typischen Anwendungsfällen die Waldsche Testgröße ist, betrachten. Anschaulich handelt es sich hierbei um Testgrößen zur Verifizierung der Hypothese, dass für den zugrunde liegenden Parameter $\vartheta = (\vartheta_1, \vartheta_2) \in \tilde{\Theta}$ die Gleichheit $\vartheta_1 = \vartheta_1^*$ gilt. Der zum Testniveau $\alpha \in (0,1)$ zugehörige Test lautet

$$\text{„Verwirf die Hypothese, falls } T_n(\vartheta_1^*) \ge \chi^2_{d_1;1-\alpha}\text{“}, \tag{4.4.3}$$

wobei $\chi^2_{d_1;1-\alpha}$ das $(1-\alpha)$-Quantil der Chi-Quadrat-Verteilung mit d_1 Freiheitsgraden sei. Für $\vartheta_1^* = \vartheta_0^1$ handelt es sich also um Testgrößen für unser zugrunde liegendes Testproblem H gegen K.

4.4.1 Lemma: *Es sei $\vartheta \in \tilde{\Theta}$ gegeben. Wir nehmen an, dass die Abbildung $\tilde{\vartheta} \mapsto f(z; \tilde{\vartheta})$ für alle $z \in R$ auf $\tilde{\Theta}$ stetig ist. Weiter sei eine Abbildung $k : R \times \tilde{\Theta} \to \mathbb{R}$ gegeben, sodass für alle $\tilde{\vartheta} \in \tilde{\Theta}$ $k(\cdot, \tilde{\vartheta}) : (R, \mathfrak{S}) \to (\mathbb{R}, \mathfrak{B})$ gilt, und für alle $z \in R$ die Abbildung $\tilde{\vartheta} \mapsto k(z, \tilde{\vartheta})$ stetig auf $\tilde{\Theta}$ ist. Zudem existiere ein $\delta > 0$ sodass mit $U_\delta = \{\tilde{\vartheta} \in \mathbb{R}^k : |\tilde{\vartheta} - \vartheta| < \delta\} \subset \tilde{\Theta}$, $M_\delta(z, \vartheta) := \sup_{\tilde{\vartheta} \in U_\delta} |k(z, \tilde{\vartheta})| f(z; \vartheta)$, $z \in R$, $\int M_\delta(z, \vartheta) \mathrm{d}\mu(z) < \infty$ gilt. Ferner sei für $n \in \mathbb{N}$ $\hat{\vartheta}_n : (R^n, \mathfrak{S}^n) \to (\tilde{\Theta}, \mathfrak{B}^k_{\tilde{\Theta}})$, sodass $\lim_{n\to\infty} \hat{\vartheta}_n(\zeta_1, ..., \zeta_n) = \vartheta$ in Q_ϑ-Wahrscheinlichkeit gilt. Dann ist k $(\mathfrak{S} \otimes \mathfrak{B}^k_{\tilde{\Theta}}, \mathfrak{B})$-messbar und es gilt*

$$\lim_{n\to\infty} \frac{1}{n} \sum_{i=1}^n k\big(\zeta_i, \hat{\vartheta}_n(\zeta_1, ..., \zeta_n)\big) = \mathrm{E}_\vartheta\left(k(\zeta_1, \vartheta)\right) \text{ in } Q_\vartheta\text{-Wahrscheinlichkeit.}$$

Beweis: Analog zum Beweis von Lemma 4.2.1. □

In Vorbereitung auf die nachfolgenden Resultate sei

$$i_{\zeta_1}(\vartheta) = \begin{pmatrix} i^{1,1}_{\zeta_1}(\vartheta) & i^{1,2}_{\zeta_1}(\vartheta) \\ i^{2,1}_{\zeta_1}(\vartheta) & i^{2,2}_{\zeta_1}(\vartheta) \end{pmatrix} \in \mathbb{R}^{k \times k},$$

mit $i_{\zeta_1}^{1,1}(\vartheta) \in \mathbb{R}^{d_1\times d_1}$, $i_{\zeta_1}^{1,2}(\vartheta) \in \mathbb{R}^{d_1\times(k-d_1)}$, $i_{\zeta_1}^{2,1}(\vartheta) \in \mathbb{R}^{(k-d_1)\times d_1}$, $i_{\zeta_1}^{2,2}(\vartheta) \in \mathbb{R}^{(k-d_1)\times(k-d_1)}$. Zudem seien für $\tilde{\vartheta} = (\tilde{\vartheta}_1, \dots, \tilde{\vartheta}_k)^t \in \tilde{\Theta}$ Nabla-Operatoren gemäß

$$\nabla_k := \Big(\frac{\partial}{\partial\tilde{\vartheta}_1}, \dots, \frac{\partial}{\partial\tilde{\vartheta}_k}\Big)^t$$

und

$$\nabla_{k-d_1} := \Big(\frac{\partial}{\partial\tilde{\vartheta}_{d_1+1}}, \dots, \frac{\partial}{\partial\tilde{\vartheta}_k}\Big)^t$$

definiert.

4.4.2 Lemma: *Es gelten die Bedingungen (D1) bis (D5). Es sei $\vartheta = (\vartheta_1, \vartheta_2) \in \tilde{\Theta}$, $\vartheta_1^* \in \mathbb{R}^{d_1}$ beliebig und $\vartheta_1^n := \vartheta_1 + \vartheta_1^*/\sqrt{n}$, $n \in \mathbb{N}$. Weiter sei $\vartheta_n := (\vartheta_1^n, \vartheta_2)$. Dann ist*

$$\begin{aligned}\sqrt{n}(\hat{\hat{\vartheta}}_{2,n}^{\vartheta_1^n} - \vartheta_2) =& i_{\zeta_1}^{2,2}(\vartheta)^{-1}(0_{(k-d_1)\times d_1}, \mathrm{I}_{k-d_1}) i_{\zeta_1}(\vartheta)\sqrt{n}(\hat{\vartheta}_n - \vartheta_n) + o_{Q_\vartheta}(1)\sqrt{n}(\hat{\vartheta}_n - \vartheta_n)\\ &+ o_{Q_\vartheta}(1)\sqrt{n}(\hat{\hat{\vartheta}}_{2,n}^{\vartheta_1^n} - \vartheta_2) + o_{Q_\vartheta}(1) \text{ für } n \to \infty,\end{aligned}$$

und es gilt bei zugrunde liegendem Parameter ϑ

$$\sqrt{n}(\hat{\vartheta}_n - \vartheta_n) \xrightarrow{\mathrm{v}} \mathrm{N}_k\Big(-\left(\begin{smallmatrix}\vartheta_1^*\\ 0\end{smallmatrix}\right), i_{\zeta_1}(\vartheta)^{-1}\Big) \text{ für } n \to \infty.$$

Beweis: Eine Taylorentwicklung der Funktion $\tilde{\vartheta} \mapsto \sum_{i=1}^n \frac{\partial}{\partial\tilde{\vartheta}_j} L(\zeta_i, \tilde{\vartheta})$, $j = 1, \dots, k$, liefert

$$\begin{aligned}&\frac{1}{\sqrt{n}}\sum_{i=1}^n \frac{\partial}{\partial\tilde{\vartheta}_j} L(\zeta_i, \tilde{\vartheta})_{|_{\tilde{\vartheta}=\vartheta_n}} + \frac{1}{n}\sum_{i=1}^n \nabla_k^t \frac{\partial}{\partial\tilde{\vartheta}_j} L(\zeta_i, \tilde{\vartheta})_{|_{\tilde{\vartheta}=\overline{\vartheta}_{j,n}}} \sqrt{n}(\hat{\vartheta}_n - \vartheta_n)\\ =& \frac{1}{\sqrt{n}}\sum_{i=1}^n \frac{\partial}{\partial\tilde{\vartheta}_j} L(\zeta_i, \tilde{\vartheta})_{|_{\tilde{\vartheta}=\hat{\vartheta}_n}} \xrightarrow{Q_\vartheta} 0 \text{ für } n \to \infty,\ j = 1, \dots, k,\end{aligned}$$

mit einer Zwischenstelle $\overline{\vartheta}_{j,n}$ auf der Verbindungsstrecke zwischen $\hat{\vartheta}_n$ und ϑ_n, $j = 1, \dots, k$. Anders geschrieben erhält man

$$\nabla_k \frac{1}{\sqrt{n}}\sum_{i=1}^n L(\zeta_i, \tilde{\vartheta})_{|_{\tilde{\vartheta}=\vartheta_n}} + \begin{pmatrix}\frac{1}{n}\sum_{i=1}^n \nabla_k^t \frac{\partial}{\partial\tilde{\vartheta}_1} L(\zeta_i, \tilde{\vartheta})_{|_{\tilde{\vartheta}=\overline{\vartheta}_{1,n}}}\\ \vdots\\ \frac{1}{n}\sum_{i=1}^n \nabla_k^t \frac{\partial}{\partial\tilde{\vartheta}_k} L(\zeta_i, \tilde{\vartheta})_{|_{\tilde{\vartheta}=\overline{\vartheta}_{k,n}}}\end{pmatrix} \sqrt{n}(\hat{\vartheta}_n - \vartheta_n) = o_{Q_\vartheta}(1)$$

für $n \to \infty$.

Mit Lemma 4.4.1, (D2) und (D3) erhält man

$$\lim_{n\to\infty} \begin{pmatrix}\frac{1}{n}\sum_{i=1}^n \nabla_k^t \frac{\partial}{\partial\tilde{\vartheta}_1} L(\zeta_i, \tilde{\vartheta})_{|_{\tilde{\vartheta}=\overline{\vartheta}_{1,n}}}\\ \vdots\\ \frac{1}{n}\sum_{i=1}^n \nabla_k^t \frac{\partial}{\partial\tilde{\vartheta}_k} L(\zeta_i, \tilde{\vartheta})_{|_{\tilde{\vartheta}=\overline{\vartheta}_{k,n}}}\end{pmatrix} = -i_{\zeta_1}(\vartheta) \text{ in } Q_\vartheta\text{-Wahrscheinlichkeit.}$$

Somit folgt einerseits

$$\nabla_k \frac{1}{\sqrt{n}} \sum_{i=1}^n L(\zeta_i, \tilde{\vartheta})_{|\tilde{\vartheta}=\vartheta_n} = i_{\zeta_1}(\vartheta)\sqrt{n}(\hat{\vartheta}_n - \vartheta_n) + o_{Q_\vartheta}(1)\sqrt{n}(\hat{\vartheta}_n - \vartheta_n) + o_{Q_\vartheta}(1)$$

für $n \to \infty$.

Andererseits liefert mit $\tilde{\vartheta} = (\tilde{\vartheta}_1, \tilde{\vartheta}_2)^t$ eine Taylorentwicklung der Funktion $\tilde{\vartheta}_2 \mapsto \sum_{i=1}^n \frac{\partial}{\partial \tilde{\vartheta}_{d_1+j}} L(\zeta_i, \tilde{\vartheta})$, $j = 1, \ldots, k - d_1$,

$$\begin{aligned}
&\frac{1}{\sqrt{n}} \sum_{i=1}^n \frac{\partial}{\partial \tilde{\vartheta}_{d_1+j}} L(\zeta_i, \tilde{\vartheta})_{|\tilde{\vartheta}=\vartheta_n} + \frac{1}{n} \sum_{i=1}^n \nabla_{k-d_1}^t \frac{\partial}{\partial \tilde{\vartheta}_{d_1+j}} L(\zeta_i, \tilde{\vartheta})_{|_{\tilde{\vartheta}=\overline{\vartheta}^{\vartheta_1^n}_{j,n}}} \sqrt{n}(\hat{\hat{\vartheta}}{}^{\vartheta_1^n}_{2,n} - \vartheta_2) \\
=&\frac{1}{\sqrt{n}} \sum_{i=1}^n \frac{\partial}{\partial \tilde{\vartheta}_{d_1+j}} L(\zeta_i, \tilde{\vartheta})_{|_{\tilde{\vartheta}=\hat{\hat{\vartheta}}{}^{\vartheta_1^n}_n}} \xrightarrow{Q_\vartheta} 0 \text{ für } n \to \infty,\ j = 1, \ldots, k - d_1,
\end{aligned}$$

mit $\overline{\vartheta}^{\vartheta_1^n}_{j,n} = (\vartheta_1^n, \overline{\vartheta}^{\vartheta_1^n}_{2,j,n})^t$ und Zwischenstellen $\overline{\vartheta}^{\vartheta_1^n}_{2,j,n}$ auf der Verbindungsstrecke zwischen $\hat{\hat{\vartheta}}{}^{\vartheta_1^n}_{2,n}$ und ϑ_2, $j = 1, \ldots, k - d_1$. Anders geschrieben erhält man

$$\begin{aligned}
&(0_{(k-d_1)\times d_1}, \mathrm{I}_{k-d_1})\nabla_k \frac{1}{\sqrt{n}} \sum_{i=1}^n L(\zeta_i, \tilde{\vartheta})_{|\tilde{\vartheta}=\vartheta_n} \\
&+ \begin{pmatrix} \frac{1}{n}\sum_{i=1}^n \nabla_{k-d_1}^t \frac{\partial}{\partial \tilde{\vartheta}_{d_1+1}} L(\zeta_i, \tilde{\vartheta})_{|_{\tilde{\vartheta}=\overline{\vartheta}^{\vartheta_1^n}_{1,n}}} \\ \vdots \\ \frac{1}{n}\sum_{i=1}^n \nabla_{k-d_1}^t \frac{\partial}{\partial \tilde{\vartheta}_k} L(\zeta_i, \tilde{\vartheta})_{|_{\tilde{\vartheta}=\overline{\vartheta}^{\vartheta_1^n}_{k-d_1,n}}} \end{pmatrix} \sqrt{n}(\hat{\hat{\vartheta}}{}^{\vartheta_1^n}_{2,n} - \vartheta_2) = o_{Q_\vartheta}(1) \text{ für } n \to \infty.
\end{aligned}$$

Mit Lemma 4.4.1 erhält man wegen (D2) und (D3)

$$\lim_{n\to\infty} \begin{pmatrix} \frac{1}{n}\sum_{i=1}^n \nabla_{k-d_1}^t \frac{\partial}{\partial \tilde{\vartheta}_{d_1+1}} L(\zeta_i, \tilde{\vartheta})_{|_{\tilde{\vartheta}=\overline{\vartheta}^{\vartheta_1^n}_{1,n}}} \\ \vdots \\ \frac{1}{n}\sum_{i=1}^n \nabla_{k-d_1}^t \frac{\partial}{\partial \tilde{\vartheta}_k} L(\zeta_i, \tilde{\vartheta})_{|_{\tilde{\vartheta}=\overline{\vartheta}^{\vartheta_1^n}_{k-d_1,n}}} \end{pmatrix} = -i^{2,2}_{\zeta_1}(\vartheta) \text{ in } Q_\vartheta\text{-Wahrscheinlichkeit},$$

wobei zu beachten ist, dass $i^{2,2}_{\zeta_1}(\vartheta)$ regulär ist. Somit folgt andererseits

$$\begin{aligned}
&(0_{(k-d_1)\times d_1}, \mathrm{I}_{k-d_1})\nabla_k \frac{1}{\sqrt{n}} \sum_{i=1}^n L(\zeta_i, \tilde{\vartheta})_{|\tilde{\vartheta}=\vartheta_n} \\
=&i^{2,2}_{\zeta_1}(\vartheta)\sqrt{n}(\hat{\hat{\vartheta}}{}^{\vartheta_1^n}_{2,n} - \vartheta_2) + o_{Q_\vartheta}(1)\sqrt{n}(\hat{\hat{\vartheta}}{}^{\vartheta_1^n}_{2,n} - \vartheta_2) + o_{Q_\vartheta}(1) \text{ für } n \to \infty,
\end{aligned}$$

und damit insgesamt

$$\begin{aligned}
\sqrt{n}(\hat{\hat{\vartheta}}{}^{\vartheta_1^n}_{2,n} - \vartheta_2) =&i^{2,2}_{\zeta_1}(\vartheta)^{-1}(0_{(k-d_1)\times d_1}, \mathrm{I}_{k-d_1})i_{\zeta_1}(\vartheta)\sqrt{n}(\hat{\vartheta}_n - \vartheta_n) + o_{Q_\vartheta}(1)\sqrt{n}(\hat{\vartheta}_n - \vartheta_n) \\
&+ o_{Q_\vartheta}(1)\sqrt{n}(\hat{\hat{\vartheta}}{}^{\vartheta_1^n}_{2,n} - \vartheta_2) + o_{Q_\vartheta}(1) \text{ für } n \to \infty,
\end{aligned}$$

also die erste Behauptung.

Zum Beweis der zweiten Behauptung zeigen wir zunächst

$$\sqrt{n}(\hat{\vartheta}_n - \vartheta) \xrightarrow{\mathrm{v}} \mathrm{N}_k\left(0, i_{\zeta_1}(\vartheta)^{-1}\right) \text{ für } n \to \infty$$

bei zugrunde liegendem Parameter ϑ. Dafür gehen wir wie folgt vor: Eine erneute Taylorentwicklung der Funktion $\tilde{\vartheta} \mapsto \sum_{i=1}^n \frac{\partial}{\partial \tilde{\vartheta}_j} L(\zeta_i, \tilde{\vartheta}),\ j = 1, \ldots, k$, liefert

$$\begin{aligned} &\frac{1}{\sqrt{n}} \sum_{i=1}^n \frac{\partial}{\partial \tilde{\vartheta}_j} L(\zeta_i, \tilde{\vartheta})_{|_{\tilde{\vartheta}=\vartheta}} + \frac{1}{n} \sum_{i=1}^n \nabla_k^t \frac{\partial}{\partial \tilde{\vartheta}_j} L(\zeta_i, \tilde{\vartheta})_{|_{\tilde{\vartheta}=\overline{\overline{\vartheta}}_{j,n}}} \sqrt{n}(\hat{\vartheta}_n - \vartheta) \\ =&\frac{1}{\sqrt{n}} \sum_{i=1}^n \frac{\partial}{\partial \tilde{\vartheta}_j} L(\zeta_i, \tilde{\vartheta})_{|_{\tilde{\vartheta}=\hat{\vartheta}_n}} \xrightarrow{Q_\vartheta} 0 \text{ für } n \to \infty,\ j = 1, \ldots, k, \end{aligned}$$

mit Zwischenstellen $\overline{\overline{\vartheta}}_{j,n}$ auf der Verbindungsstrecke zwischen $\hat{\vartheta}_n$ und $\vartheta,\ j = 1, \ldots, k$. Setzen wir für $n \in \mathbb{N}$

$$G_n := \begin{pmatrix} \frac{1}{n} \sum_{i=1}^n \nabla_k^t \frac{\partial}{\partial \tilde{\vartheta}_1} L(\zeta_i, \tilde{\vartheta})_{|_{\tilde{\vartheta}=\overline{\overline{\vartheta}}_{1,n}}} \\ \vdots \\ \frac{1}{n} \sum_{i=1}^n \nabla_k^t \frac{\partial}{\partial \tilde{\vartheta}_k} L(\zeta_i, \tilde{\vartheta})_{|_{\tilde{\vartheta}=\overline{\overline{\vartheta}}_{k,n}}} \end{pmatrix},$$

so erhalten wir

$$\nabla_k \frac{1}{\sqrt{n}} \sum_{i=1}^n L(\zeta_i, \tilde{\vartheta})_{|_{\tilde{\vartheta}=\vartheta}} + G_n \sqrt{n}(\hat{\vartheta}_n - \vartheta) = o_{Q_\vartheta}(1) \text{ für } n \to \infty.$$

Zudem folgt erneut mit Lemma 4.4.1, (D2) und (D3)

$$\lim_{n \to \infty} G_n = -i_{\zeta_1}(\vartheta) \text{ in } Q_\vartheta\text{-Wahrscheinlichkeit.}$$

Nun betrachten wir für $n \in \mathbb{N}$ den Ausdruck

$$W_n := \nabla_k \frac{1}{\sqrt{n}} \sum_{i=1}^n L(\zeta_i, \tilde{\vartheta})_{|_{\tilde{\vartheta}=\vartheta}} - i_{\zeta_1}(\vartheta) \sqrt{n}(\hat{\vartheta}_n - \vartheta),$$

der sich auch schreiben lässt in der Form

$$W_n = R_n \,\mathrm{I}(G_n \text{ regulär}) + W_n \,\mathrm{I}(G_n \text{ singulär}),$$

mit

$$\begin{aligned} R_n :=& - i_{\zeta_1}(\vartheta) \tilde{G}_n^{-1} \left(G_n \sqrt{n}(\hat{\vartheta}_n - \vartheta) + \nabla_k \frac{1}{\sqrt{n}} \sum_{i=1}^n L(\zeta_i, \tilde{\vartheta})_{|_{\tilde{\vartheta}=\vartheta}} \right) \\ &+ \left(1 + i_{\zeta_1}(\vartheta) \tilde{G}_n^{-1}\right) \nabla_k \frac{1}{\sqrt{n}} \sum_{i=1}^n L(\zeta_i, \tilde{\vartheta})_{|_{\tilde{\vartheta}=\vartheta}}, \end{aligned}$$

wobei

$$\tilde{G}_n^{-1} := \begin{cases} G_n^{-1} & , G_n \text{ regulär} \\ -i_{\zeta_1}(\vartheta)^{-1} & , G_n \text{ singulär} \end{cases}$$

gesetzt wird. Nach Annahme gilt $\lim_{n\to\infty} W_n \, \mathrm{I}(G_n \text{ singulär}) = 0$ in Q_ϑ-Wahrscheinlichkeit. Es genügt also, nur noch $\lim_{n\to\infty} R_n \, \mathrm{I}(G_n \text{ regulär}) = 0$ in Q_ϑ-Wahrscheinlichkeit sowie die entsprechende Verteilungskonvergenz für $\nabla_k \frac{1}{\sqrt{n}} \sum_{i=1}^n L(\zeta_i, \tilde{\vartheta})_{|_{\tilde{\vartheta}=\vartheta}}$ zu zeigen.

Es gilt mit dem mehrdimensionalen zentralen Grenzwertsatz bei zugrunde liegendem Parameter ϑ

$$\nabla_k \frac{1}{\sqrt{n}} \sum_{i=1}^n L(\zeta_i, \tilde{\vartheta})_{|_{\tilde{\vartheta}=\vartheta}} \xrightarrow{\mathrm{v}} \mathrm{N}_k\left(0, i_{\zeta_1}(\vartheta)\right) \text{ für } n \to \infty.$$

Insgesamt folgt mit dem Lemma von Slutsky $\lim_{n\to\infty} R_n = 0$ in Q_ϑ-Wahrscheinlichkeit, sodass insgesamt $\lim_{n\to\infty} W_n = 0$ in Q_ϑ-Wahrscheinlichkeit folgt. Somit gilt bei zugrunde liegendem Parameter ϑ

$$\sqrt{n}(\hat{\vartheta}_n - \vartheta) \xrightarrow{\mathrm{v}} \mathrm{N}_k\left(0, i_{\zeta_1}(\vartheta)^{-1}\right) \text{ für } n \to \infty.$$

Weiter ist

$$\sqrt{n}(\hat{\vartheta}_n - \vartheta_n) = \sqrt{n}(\hat{\vartheta}_n - \vartheta) + \sqrt{n}(\vartheta - \vartheta_n) = \sqrt{n}(\hat{\vartheta}_n - \vartheta) - \left(\begin{smallmatrix} \vartheta_1^* \\ 0 \end{smallmatrix}\right).$$

Damit folgt die Behauptung. □

Es soll angemerkt werden, dass die zweite Aussage aus Lemma 4.4.2 auch aus klassischen Aussagen zur asymptotischen Verteilung von Maximum-Likelihood-Schätzern hergeleitet werden kann. Siehe hierzu etwa Satz 6.35 in [27].

4.4.3 Lemma: *Es sei T_n durch die Testgröße (4.4.1) oder durch die Testgröße (4.4.2) gegeben und es gelten die Bedingungen (D1) bis (D5). Weiter sei $\vartheta = (\vartheta_1, \vartheta_2)^t \in \tilde{\Theta}$, $\vartheta_1^* \in \mathbb{R}^{d_1}$ beliebig und $\vartheta_1^n := \vartheta_1 + \vartheta_1^*/\sqrt{n}$, $n \in \mathbb{N}$. Dann gilt bei zugrunde liegendem Parameter ϑ*

$$T_n(\vartheta_1^n) \xrightarrow{\mathrm{v}} \chi^2_{d_1}(\delta) \textit{ für } n \to \infty,$$

mit

$$\delta := (\vartheta_1^*)^t \left(i_{\zeta_1}^{1,1}(\vartheta) - i_{\zeta_1}^{1,2}(\vartheta) i_{\zeta_1}^{2,2}(\vartheta)^{-1} i_{\zeta_1}^{2,1}(\vartheta)\right) \vartheta_1^*.$$

Beweis: Zunächst handle es sich bei den Tests T um Tests der Form (4.4.1). Es sei $\vartheta_n := (\vartheta_1^n, \vartheta_2)$. Eine Taylorentwicklung der Funktion $\tilde{\vartheta} \mapsto \sum_{i=1}^n L(\zeta_i, \tilde{\vartheta})$ liefert für genügend großes $n \in \mathbb{N}$

$$\sum_{j=1}^{n} L(\zeta_j, \vartheta_n) = \sum_{j=1}^{n} L(\zeta_j, \hat{\vartheta}_n) + \sqrt{n}(\vartheta_n - \hat{\vartheta}_n)^t \frac{1}{\sqrt{n}} \sum_{j=1}^{n} \nabla_k L(\zeta_j, \tilde{\vartheta})_{|_{\tilde{\vartheta}=\hat{\vartheta}_n}}$$
$$+ \frac{1}{2}\sqrt{n}(\vartheta_n - \hat{\vartheta}_n)^t \left(\frac{1}{n} \sum_{j=1}^{n} \frac{\partial^2}{\partial\tilde{\vartheta}_i \partial\tilde{\vartheta}_l} L(\zeta_j, \tilde{\vartheta})_{|_{\tilde{\vartheta}=\overline{\vartheta}_n}} \right)_{\substack{1\le i\le k\\1\le l\le k}} \sqrt{n}(\vartheta_n - \hat{\vartheta}_n),$$

mit einer Zwischenstelle $\overline{\vartheta}_n$ auf der Verbindungsstrecke zwischen $\hat{\vartheta}_n$ und ϑ_n. Mit Lemma 4.4.1, (D2) und (D3) gilt

$$\lim_{n\to\infty} \left(\frac{1}{n} \sum_{j=1}^{n} \frac{\partial^2}{\partial\tilde{\vartheta}_i \partial\tilde{\vartheta}_l} L(\zeta_j, \tilde{\vartheta})_{|_{\tilde{\vartheta}=\overline{\vartheta}_n}} \right)_{\substack{1\le i\le k\\1\le l\le k}} = -i_{\zeta_1}(\vartheta) \text{ in } Q_\vartheta\text{-Wahrscheinlichkeit.}$$

Wegen (D4) lässt sich dies auch in der Form

$$\sum_{j=1}^{n} L(\zeta_j, \hat{\vartheta}_n) - \sum_{j=1}^{n} L(\zeta_j, \vartheta_n) = \sqrt{n}(\hat{\vartheta}_n - \vartheta_n)^t o_{Q_\vartheta}(1) + \frac{1}{2}\sqrt{n}(\hat{\vartheta}_n - \vartheta_n)^t i_{\zeta_1}(\vartheta)\sqrt{n}(\hat{\vartheta}_n - \vartheta_n)$$
$$+ \sqrt{n}(\hat{\vartheta}_n - \vartheta_n)^t o_{Q_\vartheta}(1)\sqrt{n}(\hat{\vartheta}_n - \vartheta_n) \text{ für } n \to \infty$$

schreiben.

Andererseits liefert eine entsprechende Taylorentwicklung für genügend großes $n \in \mathbb{N}$

$$\sum_{j=1}^{n} L\big(\zeta_j, \vartheta_n\big) = \sum_{j=1}^{n} L(\zeta_j, \hat{\hat{\vartheta}}{}_n^{\vartheta_1^n}) + \sqrt{n}(\vartheta_2 - \hat{\hat{\vartheta}}{}_{2,n}^{\vartheta_1^n})^t \frac{1}{\sqrt{n}} \sum_{j=1}^{n} \nabla_{k-d_1} L(\zeta_j, \tilde{\vartheta})_{|_{\tilde{\vartheta}=\hat{\hat{\vartheta}}{}_n^{\vartheta_1^n}}}$$
$$+ \frac{1}{2}\sqrt{n}(\vartheta_2 - \hat{\hat{\vartheta}}{}_{2,n}^{\vartheta_1^n})^t \left(\frac{1}{n} \sum_{j=1}^{n} \frac{\partial^2}{\partial\tilde{\vartheta}_i \partial\tilde{\vartheta}_l} L(\zeta_j, \tilde{\vartheta})_{|_{\tilde{\vartheta}=\overline{\vartheta}_n^{\vartheta_1^n}}} \right)_{\substack{k-d_1+1\le i\le h\\k-d_1+1\le l\le h}} \sqrt{n}(\vartheta_2 - \hat{\hat{\vartheta}}{}_{2,n}^{\vartheta_1^n}),$$

mit $\overline{\vartheta}_n^{\vartheta_1^n} = (\vartheta_1^n, \overline{\vartheta}_{2,n}^{\vartheta_1^n})^t$ und einer Zwischenstelle $\overline{\vartheta}_{2,n}^{\vartheta_1^n}$ auf der Verbindungsstrecke zwischen $\hat{\hat{\vartheta}}{}_{2,n}^{\vartheta_1^n}$ und ϑ_2. Mit Lemma 4.4.1, (D2) und (D3) gilt

$$\lim_{n\to\infty} \left(\frac{1}{n} \sum_{j=1}^{n} \frac{\partial^2}{\partial\tilde{\vartheta}_i \partial\tilde{\vartheta}_l} L(\zeta_j, \tilde{\vartheta})_{|_{\tilde{\vartheta}=\overline{\vartheta}_n^{\vartheta_1^n}}} \right)_{\substack{k-d_1+1\le i\le h\\k-d_1+1\le l\le h}} = -i_{\zeta_1}^{2,2}(\vartheta) \text{ in } Q_\vartheta\text{-Wahrscheinlichkeit,}$$

wobei zu beachten ist, dass $i_{\zeta_1}^{2,2}(\vartheta)$ regulär ist. Wegen (D4) lässt sich dies auch in der Form

$$\sum_{j=1}^{n} L(\zeta_j, \hat{\hat{\vartheta}}{}_n^{\vartheta_1^n}) - \sum_{j=1}^{n} L(\zeta_j, \vartheta_n) = \sqrt{n}(\hat{\hat{\vartheta}}{}_{2,n}^{\vartheta_1^n} - \vartheta_2)^t o_{Q_\vartheta}(1)$$
$$+ \frac{1}{2}\sqrt{n}(\hat{\hat{\vartheta}}{}_{2,n}^{\vartheta_1^n} - \vartheta_2)^t i_{\zeta_1}^{2,2}(\vartheta)\sqrt{n}(\hat{\hat{\vartheta}}{}_{2,n}^{\vartheta_1^n} - \vartheta_2)$$
$$+ \sqrt{n}(\hat{\hat{\vartheta}}{}_{2,n}^{\vartheta_1^n} - \vartheta_2)^t o_{Q_\vartheta}(1)\sqrt{n}(\hat{\hat{\vartheta}}{}_{2,n}^{\vartheta_1^n} - \vartheta_2) \text{ für } n \to \infty$$

schreiben.

Insgesamt erhalten wir für unsere Testgröße die Darstellung

$$\begin{aligned}
T_n(\vartheta_1^n) =& 2\Bigg(\sum_{i=1}^n L(\zeta_i, \hat{\vartheta}_n) - \sum_{j=1}^n L(\zeta_j, \hat{\hat{\vartheta}}_n^{\vartheta_1^n}) \Bigg) \\
=& 2\Bigg(\sum_{i=1}^n L(\zeta_i, \hat{\vartheta}_n) - \sum_{i=1}^n L(\zeta_i, \vartheta_n) \Bigg) - 2\Bigg(\sum_{j=1}^n L(\zeta_j, \hat{\hat{\vartheta}}_n^{\vartheta_1^n}) - \sum_{i=1}^n L(\zeta_i, \vartheta_n) \Bigg) \\
=& \sqrt{n}(\hat{\vartheta}_n - \vartheta_n)^t i_{\zeta_1}(\vartheta) \sqrt{n}(\hat{\vartheta}_n - \vartheta_n) - \sqrt{n}(\hat{\hat{\vartheta}}_{2,n}^{\vartheta_1^n} - \vartheta_2)^t i_{\zeta_1}^{2,2}(\vartheta) \sqrt{n}(\hat{\hat{\vartheta}}_{2,n}^{\vartheta_1^n} - \vartheta_2) \\
&+ \sqrt{n}(\hat{\vartheta}_n - \vartheta_n)^t o_{Q_\vartheta}(1) - \sqrt{n}(\hat{\hat{\vartheta}}_{2,n}^{\vartheta_1^n} - \vartheta_2)^t o_{Q_\vartheta}(1) \\
&+ \sqrt{n}(\hat{\vartheta}_n - \vartheta_n)^t o_{Q_\vartheta}(1) \sqrt{n}(\hat{\vartheta}_n - \vartheta_n) - \sqrt{n}(\hat{\hat{\vartheta}}_{2,n}^{\vartheta_1^n} - \vartheta_2)^t o_{Q_\vartheta}(1) \sqrt{n}(\hat{\hat{\vartheta}}_{2,n}^{\vartheta_1^n} - \vartheta_2) \\
&\text{für } n \to \infty.
\end{aligned}$$

Mit Hilfe von Lemma 4.4.2 lässt sich dies schreiben als

$$T_n(\vartheta_1^n) = \sqrt{n}(\hat{\vartheta}_n - \vartheta_n)^t i_{\zeta_1}^{**}(\vartheta) \sqrt{n}(\hat{\vartheta}_n - \vartheta_n) + o_{Q_\vartheta}(1) \text{ für } n \to \infty,$$

mit

$$i_{\zeta_1}^{**}(\vartheta) := \begin{pmatrix} i_{\zeta_1}^{1,1}(\vartheta) - i_{\zeta_1}^{1,2}(\vartheta) i_{\zeta_1}^{2,2}(\vartheta)^{-1} i_{\zeta_1}^{2,1}(\vartheta) & 0_{d_1 \times (k-d_1)} \\ 0_{(k-d_1) \times d_1} & 0_{(k-d_1) \times (k-d_1)} \end{pmatrix} \in \mathbb{R}^{k \times k}.$$

Mit Hilfe der zweiten Aussage von Lemma 4.4.2 sowie mit $i_{\zeta_1}^{**}(\tilde{\vartheta}) i_{\zeta_1}^{-1}(\tilde{\vartheta}) i_{\zeta_1}^{**}(\tilde{\vartheta}) = i_{\zeta_1}^{**}(\tilde{\vartheta})$, $\tilde{\vartheta} \in \tilde{\Theta}$, und Satz 1.4.5 in [28] folgt

$$T_n(\vartheta_1^n) \xrightarrow{\text{v}} \chi_{d_1}^2(\delta) \text{ für } n \to \infty.$$

Nun handle es sich bei der Folge von Tests T um Tests der Form (4.4.2). Man erhält mit Lemma 4.4.2

$$\begin{aligned}
\sqrt{n}(\hat{\vartheta}_n - \hat{\hat{\vartheta}}_n^{\vartheta_1^n}) =& \sqrt{n}(\hat{\vartheta}_n - \vartheta_n) - \sqrt{n}(\hat{\hat{\vartheta}}_n^{\vartheta_1^n} - \vartheta_n) \\
=& \sqrt{n}(\hat{\vartheta}_n - \vartheta_n) - \begin{pmatrix} 0_{d_1 \times (k-d_1)} \\ \mathrm{I}_{k-d_1} \end{pmatrix} \sqrt{n}(\hat{\hat{\vartheta}}_{2,n}^{\vartheta_1^n} - \vartheta_2) \\
=& \Bigg(\mathrm{I}_k - \begin{pmatrix} 0_{d_1 \times (k-d_1)} \\ \mathrm{I}_{k-d_1} \end{pmatrix} i_{\zeta_1}^{2,2}(\vartheta)^{-1} (0_{(k-d_1) \times d_1}, \mathrm{I}_{k-d_1}) i_{\zeta_1}(\vartheta) \Bigg) \sqrt{n}(\hat{\vartheta}_n - \vartheta_n) \\
&+ o_{Q_\vartheta}(1) \sqrt{n}(\hat{\vartheta}_n - \vartheta_n) + o_{Q_\vartheta}(1) \sqrt{n}(\hat{\hat{\vartheta}}_{2,n}^{\vartheta_1^n} - \vartheta_2) + o_{Q_\vartheta}(1) \text{ für } n \to \infty
\end{aligned}$$

und daraus mit $i_{\zeta_1}^{**}(\tilde{\vartheta}) i_{\zeta_1}^{-1}(\tilde{\vartheta}) i_{\zeta_1}^{**}(\tilde{\vartheta}) = i_{\zeta_1}^{**}(\tilde{\vartheta})$, $\tilde{\vartheta} \in \tilde{\Theta}$, sowie der Stetigkeit der Abbildung $\tilde{\vartheta} \mapsto i_{\zeta_1}(\tilde{\vartheta})$, $\tilde{\vartheta} \in \tilde{\Theta}$, welche aus den Annahmen folgt, schließlich

$$\begin{aligned}
T_n &= \sqrt{n}(\hat{\vartheta}_n - \hat{\hat{\vartheta}}_n^{\vartheta_1^n})^t i_{\zeta_1}(\hat{\hat{\vartheta}}_n^{\vartheta_1^n}) \sqrt{n}(\hat{\vartheta}_n - \hat{\hat{\vartheta}}_n^{\vartheta_1^n}) \\
&= \sqrt{n}(\hat{\vartheta}_n - \vartheta_n)^t i_{\zeta_1}^{**}(\vartheta) i_{\zeta_1}^{-1}(\vartheta) i_{\zeta_1}^{**}(\vartheta) \sqrt{n}(\hat{\vartheta}_n - \vartheta_n) + o_{Q_\vartheta}(1) \\
&= \sqrt{n}(\hat{\vartheta}_n - \vartheta_n)^t i_{\zeta_1}^{**}(\vartheta) \sqrt{n}(\hat{\vartheta}_n - \vartheta_n) + o_{Q_\vartheta}(1) \text{ für } n \to \infty.
\end{aligned}$$

Der Rest folgt analog zu den oben dargestellten Ausführungen. $\square$

Nach diesen Vorbereitungen lassen sich die zentralen Resultate in diesem Unterabschnitt formulieren.

4.4.4 Satz: *Es sei T_n durch die Testgröße (4.4.1) oder durch die Testgröße (4.4.2) gegeben und es gelten die Bedingungen (D1) bis (D5). Es sei $\vartheta = (\vartheta_1, \vartheta_2)^t \in \tilde{\Theta}$. Zudem existieren zu ϑ nicht negative messbare Abbildungen $m_n : \mathbb{R}^{d_1} \to \mathbb{R}$, $n \in \mathbb{N}$, und $m : \mathbb{R}^{d_1} \to \mathbb{R}$ mit*

$$Q_\vartheta\big(T_n(\vartheta_1 + \vartheta_1^*/\sqrt{n}) < \chi^2_{d_1;1-\alpha}\big) \le m_n(\vartheta_1^*) \text{ für } (\vartheta_1 + \vartheta_1^*/\sqrt{n}) \in \Theta^1 \tag{4.4.4}$$

und

$$\lim_{n\to\infty} m_n(\vartheta_1^*) = m(\vartheta_1^*) \text{ für alle } \vartheta_1^* \in \mathbb{R}^{d_1} \tag{4.4.5}$$

sowie

$$\int m_n(\vartheta_1^*)\mathrm{d}\vartheta_1^* < \infty,\ n \in \mathbb{N},\ \int m(\vartheta_1^*)\mathrm{d}\vartheta_1^* < \infty \text{ und } \lim_{n\to\infty}\int m_n(\vartheta_1^*)\mathrm{d}\vartheta_1^* = \int m(\vartheta_1^*)\mathrm{d}\vartheta_1^*. \tag{4.4.6}$$

Dann gilt für das erwartete Volumen des Konfidenzbereichs bzgl. des in (4.4.3) genannten Tests

$$\lim_{n\to\infty} n^{d_1/2}\,\mathrm{E}_\vartheta(\mathrm{vol}_{T,n}) = \frac{(\chi^2_{d_1;1-\alpha})^{d_1/2}\pi^{d_1/2}}{\Gamma(d_1/2+1)} \det\big(i^{1,1}_{\zeta_1}(\vartheta) - i^{1,2}_{\zeta_1}(\vartheta) i^{2,2}_{\zeta_1}(\vartheta)^{-1} i^{2,1}_{\zeta_1}(\vartheta)\big)^{-1/2}.$$

Beweis: Durch Substitution und mit dem Satz von Fubini erhält man

$$\begin{aligned} n^{d_1/2}\,\mathrm{E}_\vartheta(\mathrm{vol}_{T,n}) &= \mathrm{E}_\vartheta\left(n^{d_1/2}\int \mathrm{I}\left(T_n(\vartheta_1^*) < \chi^2_{d_1;1-\alpha}\right)\mathrm{I}(\vartheta_1^* \in \Theta^1)\mathrm{d}\vartheta_1^*\right) \\ &= \mathrm{E}_\vartheta\left(\int \mathrm{I}\left(T_n(\vartheta_1 + \vartheta_1^*/\sqrt{n}) < \chi^2_{d_1;1-\alpha}\right)\mathrm{I}\left((\vartheta_1 + \vartheta_1^*/\sqrt{n}) \in \Theta^1\right)\mathrm{d}\vartheta_1^*\right) \\ &= \int Q_\vartheta\big(T_n(\vartheta_1 + \vartheta_1^*/\sqrt{n}) < \chi^2_{d_1;1-\alpha}\big)\,\mathrm{I}\left((\vartheta_1 + \vartheta_1^*/\sqrt{n}) \in \Theta^1\right)\mathrm{d}\vartheta_1^*. \end{aligned}$$

Mit Lemma 4.4.3 gilt

$$T_n(\vartheta_1^n) \xrightarrow{\mathrm{v}} T(\vartheta_1^*) := (Z - \vartheta_1^*)^t\big(i^{1,1}_{\zeta_1}(\vartheta) - i^{1,2}_{\zeta_1}(\vartheta) i^{2,2}_{\zeta_1}(\vartheta)^{-1} i^{2,1}_{\zeta_1}(\vartheta)\big)(Z - \vartheta_1^*) \text{ für } n \to \infty,$$

mit einem ohne Einschränkung auf dem zugrunde liegendem Wahrscheinlichkeitsraum existierenden Zufallsvektor Z mit

$$Z \sim \mathrm{N}_{d_1}\big(0, i^{1,1}_{\zeta_1}(\vartheta) - i^{1,2}_{\zeta_1}(\vartheta) i^{2,2}_{\zeta_1}(\vartheta)^{-1} i^{2,1}_{\zeta_1}(\vartheta)\big)$$

bei zugrunde liegendem Parameter ϑ. Wegen der Stetigkeit der Verteilungsfunktion von $T(\vartheta_1^*)$ bei zugrunde liegendem Parameter ϑ gilt

$$\lim_{n\to\infty} Q_\vartheta\big(T_n(\vartheta_1+\vartheta_1^*/\sqrt{n}) < \chi^2_{d_1;1-\alpha}\big) = Q_\vartheta\big(T(\vartheta_1^*) < \chi^2_{d_1;1-\alpha}\big) \text{ für alle } \vartheta_1^* \in \mathbb{R}^{d_1}.$$

Dann gilt wegen $\lim_{n\to\infty} \mathrm{I}((\vartheta_1+\vartheta_1^*/\sqrt{n}) \in \Theta^1) = 1$ für alle $\vartheta_1^* \in \mathbb{R}^{d_1}$ auch

$$\lim_{n\to\infty} Q_\vartheta\big(T_n(\vartheta_1+\vartheta_1^*/\sqrt{n}) < \chi^2_{d_1;1-\alpha}\big)\,\mathrm{I}\left((\vartheta_1+\vartheta_1^*/\sqrt{n}) \in \Theta^1\right) = Q_\vartheta\big(T(\vartheta_1^*) < \chi^2_{d_1;1-\alpha}\big)$$
$$\text{für alle } \vartheta_1^* \in \mathbb{R}^{d_1}.$$

Nach Voraussetzung existieren nicht negative messbare Abbildungen $m_n : \mathbb{R}^{d_1} \to \mathbb{R}$, $n \in \mathbb{N}$, und $m : \mathbb{R}^{d_1} \to \mathbb{R}$ mit

$$Q_\vartheta\big(T_n(\vartheta_1+\vartheta_1^*/\sqrt{n}) < \chi^2_{d_1;1-\alpha}\big)\,\mathrm{I}\left((\vartheta_1+\vartheta_1^*/\sqrt{n}) \in \Theta^1\right) \le m_n(\vartheta_1^*)$$
$$\text{für alle } \vartheta_1^* \in \mathbb{R}^{d_1} \text{ und alle } n \in \mathbb{N},$$

die (4.4.4), (4.4.5) und (4.4.6) erfüllen. Damit folgt mit Pratts Version des Satzes von der majorisierten Konvergenz (siehe A.1.1 im Anhang) und erneut mit dem Satz von Fubini

$$\begin{aligned}
&\lim_{n\to\infty} n^{d_1/2}\,\mathrm{E}_\vartheta(\mathrm{vol}_{T,n}) = \int Q_\vartheta\big(T(\vartheta_1^*) < \chi^2_{d_1;1-\alpha}\big)\mathrm{d}\vartheta_1^* = \mathrm{E}_\vartheta\left(\int \mathrm{I}\left(T(\vartheta_1^*) < \chi^2_{d_1;1-\alpha}\right)\mathrm{d}\vartheta_1^*\right)\\
&= \mathrm{E}_\vartheta\left(\int \mathrm{I}\left((Z-\vartheta_1^*)^t\big(i^{1,1}_{\zeta_1}(\vartheta) - i^{1,2}_{\zeta_1}(\vartheta) i^{2,2}_{\zeta_1}(\vartheta)^{-1} i^{2,1}_{\zeta_1}(\vartheta)\big)(Z-\vartheta_1^*) < \chi^2_{d_1;1-\alpha}\right)\mathrm{d}\vartheta_1^*\right)\\
&= \mathrm{E}_\vartheta\left(\frac{(\chi^2_{d_1;1-\alpha})^{d_1/2}\pi^{d_1/2}}{\Gamma(d_1/2+1)} \det\left(i^{1,1}_{\zeta_1}(\vartheta) - i^{1,2}_{\zeta_1}(\vartheta) i^{2,2}_{\zeta_1}(\vartheta)^{-1} i^{2,1}_{\zeta_1}(\vartheta)\right)^{-1/2}\right)\\
&= \frac{(\chi^2_{d_1;1-\alpha})^{d_1/2}\pi^{d_1/2}}{\Gamma(d_1/2+1)} \det\left(i^{1,1}_{\zeta_1}(\vartheta) - i^{1,2}_{\zeta_1}(\vartheta) i^{2,2}_{\zeta_1}(\vartheta)^{-1} i^{2,1}_{\zeta_1}(\vartheta)\right)^{-1/2},
\end{aligned}$$

wobei hier die bekannte Formel zur Berechnung des Volumens für Ellipsoide, siehe Beispiel 1 in Abschnitt 9.1 in [21], benutzt wurde. □

Die beiden nachfolgenden Lemmata können zum Nachweis der Existenz von nicht negativen messbaren Abbildungen $m_n : \mathbb{R}^{d_1} \to \mathbb{R}$, $n \in \mathbb{N}$, und $m : \mathbb{R}^{d_1} \to \mathbb{R}$ mit (4.4.4), (4.4.5) und (4.4.6) nützlich sein.

4.4.5 Lemma: *Es sei T_n durch die Testgröße (4.4.2) gegeben und es gelten die Bedingungen (D1) bis (D5). Es sei $\vartheta = (\vartheta_1, \vartheta_2)^t \in \tilde{\Theta}$. Weiter existieren Mengen $K_1, K_2 \in \mathfrak{B}^k_{\tilde{\Theta}}$, $K_1 \cap K_2 = \emptyset$, $K_1 \cup K_2 = \tilde{\Theta}$, sodass eine reelle Zahl $\lambda_{\min} > 0$ existiert mit*

$$\inf_{\tilde{\vartheta}\in K_1} \lambda_{\tilde{\vartheta}} > \lambda_{\min},$$

mit $\lambda_{\tilde{\vartheta}}$ als kleinster Eigenwert der Matrix $i_{\zeta_1}(\tilde{\vartheta})$, $\tilde{\vartheta} \in K_1$, und dass

$$\lim_{n\to\infty} Q_\vartheta\big(T_n(\vartheta_1+\vartheta_1^*/\sqrt{n}) < \chi^2_{d_1;1-\alpha}, (\vartheta_1+\vartheta_1^*/\sqrt{n}) \in \Theta^1, \hat{\tilde{\vartheta}}_n^{\vartheta_1+\vartheta_1^*/\sqrt{n}} \in K_2\big) = 0$$
$$\textit{für alle } \vartheta_1^* \in \mathbb{R}^{d_1}$$

sowie

$$\lim_{n\to\infty}\int Q_\vartheta\big(T_n(\vartheta_1+\vartheta_1^*/\sqrt{n})<\chi^2_{d_1;1-\alpha},(\vartheta_1+\vartheta_1^*/\sqrt{n})\in\Theta^1,\hat{\hat{\vartheta}}_n^{\vartheta_1+\vartheta_1^*/\sqrt{n}}\in K_2\big)\mathrm{d}\vartheta_1^*=0$$

gilt. Dann existieren nicht negative messbare Abbildungen $m_n:\mathbb{R}^{d_1}\to\mathbb{R}$, $n\in\mathbb{N}$, *und* $m:\mathbb{R}^{d_1}\to\mathbb{R}$ *mit (4.4.4), (4.4.5) und (4.4.6).*

Beweis: Es sei $\hat{\vartheta}_n=(\hat{\vartheta}_{1,n}\hat{\vartheta}_{2,n})^t$. Weiter sei $\vartheta_1^*\in\mathbb{R}^{d_1}$ beliebig, $n\in\mathbb{N}$ und $\vartheta_1^n:=\vartheta_1+\vartheta_1^*/\sqrt{n}$. Dann ist

$$\begin{aligned}T_n(\vartheta_1^n)&=\sqrt{n}(\hat{\vartheta}_n-\hat{\hat{\vartheta}}_n^{\vartheta_1^n})^t i_{\zeta_1}(\hat{\hat{\vartheta}}_n^{\vartheta_1^n})\sqrt{n}(\hat{\vartheta}_n-\hat{\hat{\vartheta}}_n^{\vartheta_1^n})\geq\lambda_{\hat{\hat{\vartheta}}_n^{\vartheta_1^n}}\sqrt{n}(\hat{\vartheta}_n-\hat{\hat{\vartheta}}_n^{\vartheta_1^n})^t\sqrt{n}(\hat{\vartheta}_n-\hat{\hat{\vartheta}}_n^{\vartheta_1^n})\\&\geq\lambda_{\min}\sqrt{n}(\hat{\vartheta}_n-\hat{\hat{\vartheta}}_n^{\vartheta_1^n})^t\sqrt{n}(\hat{\vartheta}_n-\hat{\hat{\vartheta}}_n^{\vartheta_1^n})\geq\lambda_{\min}\sqrt{n}\big(\hat{\vartheta}_{1,n}-\vartheta_1^n\big)^t\sqrt{n}\big(\hat{\vartheta}_{1,n}-\vartheta_1^n\big)\\&\text{auf }\{\hat{\hat{\vartheta}}_n^{\vartheta_1^n}\in K_1\}\text{ für }\vartheta_1^n\in\Theta^1.\end{aligned}$$

Damit ist

$$\begin{aligned}\mathrm{I}\big(T_n(\vartheta_1^n)<\chi^2_{d_1;1-\alpha},\vartheta_1^n\in\Theta^1\big)=&\,\mathrm{I}\big(T_n(\vartheta_1^n)<\chi^2_{d_1;1-\alpha},\vartheta_1^n\in\Theta^1,\hat{\hat{\vartheta}}_n^{\vartheta_1^n}\in K_1\big)\\&+\mathrm{I}\big(T_n(\vartheta_1^n)<\chi^2_{d_1;1-\alpha},\vartheta_1^n\in\Theta^1,\hat{\hat{\vartheta}}_n^{\vartheta_1^n}\in K_2\big)\\\leq&\,\mathrm{I}\big(\lambda_{\min}\sqrt{n}(\hat{\vartheta}_{1,n}-\vartheta_1^n)^t\sqrt{n}(\hat{\vartheta}_{1,n}-\vartheta_1^n)<\chi^2_{d_1;1-\alpha}\big)\\&+\mathrm{I}\big(T_n(\vartheta_1^n)<\chi^2_{d_1;1-\alpha},\vartheta_1^n\in\Theta^1,\hat{\hat{\vartheta}}_n^{\vartheta_1^n}\in K_2\big),\ \vartheta_1^*\in\mathbb{R}^{d_1}.\end{aligned}$$

Es folgt mit der Monotonie des Erwartungswertes

$$\begin{aligned}\mathrm{E}_\vartheta\Big(\mathrm{I}\big(T_n(\vartheta_1^n)<\chi^2_{d_1;1-\alpha},\vartheta_1^n\in\Theta^1\big)\Big)\leq&\,\mathrm{E}_\vartheta\Big(\mathrm{I}\big(\lambda_{\min}\sqrt{n}(\hat{\vartheta}_{1,n}-\vartheta_1^n)^t\sqrt{n}(\hat{\vartheta}_{1,n}-\vartheta_1^n)<\chi^2_{d_1;1-\alpha}\big)\\&+\mathrm{I}\big(T_n(\vartheta_1^n)<\chi^2_{d_1;1-\alpha},\vartheta_1^n\in\Theta^1,\hat{\hat{\vartheta}}_n^{\vartheta_1^n}\in K_2\big)\Big),\\&\vartheta_1^*\in\mathbb{R}^{d_1},\end{aligned}$$

äquivalent

$$\begin{aligned}Q_\vartheta\big(T_n(\vartheta_1^n)<\chi^2_{d_1;1-\alpha},\vartheta_1^n\in\Theta^1\big)\leq&\,Q_\vartheta\big(\lambda_{\min}\sqrt{n}(\hat{\vartheta}_{1,n}-\vartheta_1^n)^t\sqrt{n}(\hat{\vartheta}_{1,n}-\vartheta_1^n)<\chi^2_{d_1;1-\alpha}\big)\\&+Q_\vartheta\big(T_n(\vartheta_1^n)<\chi^2_{d_1;1-\alpha},\vartheta_1^n\in\Theta^1,\hat{\hat{\vartheta}}_n^{\vartheta_1^n}\in K_2\big),\ \vartheta_1^*\in\mathbb{R}^{d_1}.\end{aligned}$$

Wir setzen

$$\begin{aligned}m_n(\vartheta_1^*):=&\,Q_\vartheta\big(\lambda_{\min}\sqrt{n}(\hat{\vartheta}_{1,n}-\vartheta_1^n)^t\sqrt{n}(\hat{\vartheta}_{1,n}-\vartheta_1^n)<\chi^2_{d_1;1-\alpha}\big)\\&+Q_\vartheta\big(T_n(\vartheta_1^n)<\chi^2_{d_1;1-\alpha},\vartheta_1^n\in\Theta^1,\hat{\hat{\vartheta}}_n^{\vartheta_1^n}\in K_2\big),\ \vartheta_1^*\in\mathbb{R}^{d_1},\ n\in\mathbb{N}.\end{aligned}$$

Dann erfüllt m_n (4.4.4). Weiter gilt mit Lemma 4.4.2

$$\sqrt{n}(\hat{\vartheta}_{1,n}-\vartheta_1^n)\sqrt{n}(\hat{\vartheta}_{1,n}-\vartheta_1^n)^t\xrightarrow{\mathrm{v}}(W-\vartheta_1^*)^t(W-\vartheta_1^*)\text{ für }n\to\infty,\ \vartheta_1^*\in\mathbb{R}^{d_1},$$

mit einem ohne Einschränkung auf dem zugrunde liegendem Wahrscheinlichkeitsraum existierenden Zufallsvektor W mit

$$W \sim \mathrm{N}_{d_1}\Big(0, \big(i_{\zeta_1}^{1,1}(\vartheta) - i_{\zeta_1}^{1,2}(\vartheta) i_{\zeta_1}^{2,2}(\vartheta)^{-1} i_{\zeta_1}^{2,1}(\vartheta)\big)^{-1}\Big)$$

bei zugrunde liegendem Parameter ϑ. Wegen der Stetigkeit der Verteilungsfunktion von $(W - \vartheta_1^*)^t (W - \vartheta_1^*)$ bei zugrunde liegendem Parameter ϑ gilt

$$\begin{aligned}&\lim_{n\to\infty} Q_\vartheta\big(\lambda_{\min}\sqrt{n}(\hat{\vartheta}_{1,n} - \vartheta_1^n)^t \sqrt{n}(\hat{\vartheta}_{1,n} - \vartheta_1^n) < \chi^2_{d_1;1-\alpha}\big)\\ =&Q_\vartheta\big(\lambda_{\min}(W - \vartheta_1^*)^t (W - \vartheta_1^*) < \chi^2_{d_1;1-\alpha}\big),\ \vartheta_1^* \in \mathbb{R}^{d_1}.\end{aligned}$$

Nach Voraussetzung gilt

$$\lim_{n\to\infty} Q_\vartheta\big(T_n(\vartheta_1^n) < \chi^2_{d_1;1-\alpha}, \vartheta_1^n \in \Theta^1, \hat{\hat{\vartheta}}_n^{\vartheta_1^n} \in K_2\big) = 0 \text{ für alle } \vartheta_1^* \in \mathbb{R}^{d_1}.$$

Setzen wir

$$m(\vartheta_1^*) := Q_\vartheta\big(\lambda_{\min}(W - \vartheta_1^*)^t (W - \vartheta_1^*) < \chi^2_{d_1;1-\alpha}\big),\ \vartheta_1^* \in \mathbb{R}^{d_1},$$

gilt also $\lim_{n\to\infty} m_n(\vartheta_1^*) = m(\vartheta_1^*)$ für alle $\vartheta_1^* \in \mathbb{R}^{d_1}$, d.h. (4.4.5). Es muss nur noch (4.4.6) nachgewiesen werden. Es gilt mit dem Satz von Fubini

$$\begin{aligned}&\int Q_\vartheta\big(\lambda_{\min}\sqrt{n}(\hat{\vartheta}_{1,n} - \vartheta_1^n)^t \sqrt{n}(\hat{\vartheta}_{1,n} - \vartheta_1^n) < \chi^2_{d_1;1-\alpha}\big) \mathrm{d}\vartheta_1^*\\ &= \mathrm{E}_\vartheta\left(\int \mathrm{I}\left(\lambda_{\min}\sqrt{n}(\hat{\vartheta}_{1,n} - \vartheta_1^n)^t \sqrt{n}(\hat{\vartheta}_{1,n} - \vartheta_1^n) < \chi^2_{d_1;1-\alpha}\right) \mathrm{d}\vartheta_1^*\right)\\ &= \mathrm{E}_\vartheta\left(\frac{(\chi^2_{d_1;1-\alpha})^{d_1/2}\pi^{d_1/2}}{\lambda_{\min}^{d_1/2}\Gamma(d_1/2+1)}\right) = \frac{(\chi^2_{d_1;1-\alpha})^{d_1/2}\pi^{d_1/2}}{\lambda_{\min}^{d_1/2}\Gamma(d_1/2+1)} < \infty,\ n \in \mathbb{N},\end{aligned}$$

wobei hier die bekannte Formel zur Berechnung des Volumens für Kugeln benutzt wurde. Damit folgt mit den Voraussetzungen $\int m_n(\vartheta_1^*)\mathrm{d}\vartheta_1^* < \infty$ für genügend großes $n \in \mathbb{N}$, was uns genügen soll. Erneut mit dem Satz von Fubini ist

$$\begin{aligned}&\int Q_\vartheta\big(\lambda_{\min}(W - \vartheta_1^*)^t (W - \vartheta_1^*) < \chi^2_{d_1;1-\alpha}\big) \mathrm{d}\vartheta_1^*\\ &= \mathrm{E}_\vartheta\left(\int \mathrm{I}\left(\lambda_{\min}(W - \vartheta_1^*)^t (W - \vartheta_1^*) < \chi^2_{d_1;1-\alpha}\right) \mathrm{d}\vartheta_1^*\right)\\ &= \mathrm{E}_\vartheta\left(\frac{(\chi^2_{d_1;1-\alpha})^{d_1/2}\pi^{d_1/2}}{\lambda_{\min}^{d_1/2}\Gamma(d_1/2+1)}\right) = \frac{(\chi^2_{d_1;1-\alpha})^{d_1/2}\pi^{d_1/2}}{\lambda_{\min}^{d_1/2}\Gamma(d_1/2+1)} < \infty.\end{aligned}$$

Damit folgt $\int m(\vartheta_1^*)\mathrm{d}\vartheta_1^* < \infty$. Wie wir also sehen, ist

$$\begin{aligned}&\int Q_\vartheta\big(\lambda_{\min}\sqrt{n}(\hat{\vartheta}_{1,n} - \vartheta_1^n)^t \sqrt{n}(\hat{\vartheta}_{1,n} - \vartheta_1^n) < \chi^2_{d_1;1-\alpha}\big) \mathrm{d}\vartheta_1^*\\ &= \int Q_\vartheta\big(\lambda_{\min}(Z - \vartheta_1^*)^t (Z - \vartheta_1^*) < \chi^2_{d_1;1-\alpha}\big) \mathrm{d}\vartheta_1^*,\ n \in \mathbb{N}.\end{aligned}$$

Damit ist mit den Voraussetzungen $\lim_{n\to\infty} \int m_n(\vartheta_1^*)\mathrm{d}\vartheta_1^* = \int m(\vartheta_1^*)\mathrm{d}\vartheta_1^*$ und insgesamt (4.4.6) gezeigt. □

4.4.6 Lemma: *Es sei T_n durch die Testgröße (4.4.1) gegeben und es gelten die Bedingungen (D1) bis (D5). Es sei $\vartheta = (\vartheta_1, \vartheta_2)^t \in \tilde{\Theta}$. Weiter existieren Mengen $W_{1,n}, W_{2,n} \subset \mathfrak{S}^n$, $W_{1,n} \cap W_{2,n} = \emptyset$, $W_{1,n} \cup W_{2,n} = R^n$, $n \in \mathbb{N}$, und $K_1, K_2 \in \mathfrak{B}^k_{\tilde{\Theta}}$, $K_1 \cap K_2 = \emptyset$, $K_1 \cup K_2 = \tilde{\Theta}$, K_1 konvex, sodass eine reelle Zahl $\lambda_{\min} > 0$ existiert mit*

$$\inf_{\tilde{\vartheta} \in K_1} \lambda_{n,\tilde{\vartheta}} > \lambda_{\min} \text{ auf } \{(\zeta_1, \dots, \zeta_n) \in W_{1,n}\} \text{ für alle } n \in \mathbb{N},$$

mit $\lambda_{n,\tilde{\vartheta}}$ als kleinster Eigenwert der Matrix $(-\frac{1}{n}\sum_{j=1}^n \frac{\partial^2}{\partial\tilde{\vartheta}_i \partial\tilde{\vartheta}_l} L(\zeta_j, \tilde{\vartheta}))_{\substack{1\le i\le k\\ 1\le l\le k}}$, $\tilde{\vartheta} \in K_1$, und dass

$$\lim_{n\to\infty} Q_\vartheta\big(T_n(\vartheta_1 + \vartheta_1^*/\sqrt{n}) < \chi^2_{d_1;1-\alpha}, (\vartheta_1 + \vartheta_1^*/\sqrt{n}) \in \Theta^1, (\zeta_1, \dots, \zeta_n) \in W_{1,n},$$
$$\hat{\vartheta}_n \in K_2 \text{ oder } \hat{\hat{\vartheta}}_n^{\vartheta_1 + \vartheta_1^*/\sqrt{n}} \in K_2\big) = 0 \text{ für alle } \vartheta_1^* \in \mathbb{R}^{d_1}$$

sowie

$$\lim_{n\to\infty} \int Q_\vartheta\big(T_n(\vartheta_1 + \vartheta_1^*/\sqrt{n}) < \chi^2_{d_1;1-\alpha}, (\vartheta_1 + \vartheta_1^*/\sqrt{n}) \in \Theta^1, (\zeta_1, \dots, \zeta_n) \in W_{1,n},$$
$$\hat{\vartheta}_n \in K_2 \text{ oder } \hat{\hat{\vartheta}}_n^{\vartheta_1 + \vartheta_1^*/\sqrt{n}} \in K_2\big) \mathrm{d}\vartheta_1^* = 0$$

gilt. Ferner gelte

$$\lim_{n\to\infty} Q_\vartheta\big(T_n(\vartheta_1 + \vartheta_1^*/\sqrt{n}) < \chi^2_{d_1;1-\alpha}, (\vartheta_1 + \vartheta_1^*/\sqrt{n}) \in \Theta^1, (\zeta_1, \dots, \zeta_n) \in W_{2,n}\big) = 0$$
$$\text{für alle } \vartheta_1^* \in \mathbb{R}^{d_1}$$

sowie

$$\lim_{n\to\infty} \int Q_\vartheta\big(T_n(\vartheta_1 + \vartheta_1^*/\sqrt{n}) < \chi^2_{d_1;1-\alpha}, (\vartheta_1 + \vartheta_1^*/\sqrt{n}) \in \Theta^1, (\zeta_1, \dots, \zeta_n) \in W_{2,n}\big) \mathrm{d}\vartheta_1^* = 0.$$

Außrdem sei

$$\sum_{j=0}^n \frac{\partial}{\partial\tilde{\vartheta}_i} L(\zeta_j, \tilde{\vartheta})_{|\tilde{\vartheta} = \hat{\vartheta}_n(\zeta_1, \dots, \zeta_n)} = 0 \text{ auf } \{(\zeta_1, \dots, \zeta_n) \in W_{1,n}\} \text{ für alle } n \in \mathbb{N}$$

erfüllt. Dann existieren nicht negative messbare Abbildungen $m_n : \mathbb{R}^{d_1} \to \mathbb{R}$, $n \in \mathbb{N}$, und $m : \mathbb{R}^{d_1} \to \mathbb{R}$ mit (4.4.4), (4.4.5) und (4.4.6).

Beweis: Es sei $\hat{\vartheta}_n = (\hat{\vartheta}_{1,n} \hat{\vartheta}_{2,n})^t$. Weiter sei $\vartheta_1^* \in \mathbb{R}^{d_1}$ beliebig, $n \in \mathbb{N}$ und $\vartheta_1^n := \vartheta_1 + \vartheta_1^*/\sqrt{n}$. Eine Taylorentwicklung der Funktion $\tilde{\vartheta} \mapsto \sum_{i=1}^n L(\zeta_i, \tilde{\vartheta})$ liefert für $\vartheta_1^n \in \Theta^1$

$$\begin{aligned}\sum_{j=1}^n L(\zeta_j, \hat{\hat{\vartheta}}_n^{\vartheta_1^n}) = &\sum_{j=1}^n L(\zeta_j, \hat{\vartheta}_n) + \sqrt{n}(\hat{\hat{\vartheta}}_n^{\vartheta_1^n} - \hat{\vartheta}_n)^t \frac{1}{\sqrt{n}} \sum_{j=1}^n \nabla_k L(\zeta_j, \tilde{\vartheta})_{|\tilde{\vartheta} = \hat{\vartheta}_n} \\ &+ \frac{1}{2}\sqrt{n}(\hat{\hat{\vartheta}}_n^{\vartheta_1^n} - \hat{\vartheta}_n)^t \left(\frac{1}{n}\sum_{j=1}^n \frac{\partial^2}{\partial\tilde{\vartheta}_i \partial\tilde{\vartheta}_l} L(\zeta_j, \tilde{\vartheta})_{|_{\tilde{\vartheta} = \overline{\vartheta}_n^{\vartheta_1^n}}}\right)_{\substack{1\le i\le k\\ 1\le l\le k}} \sqrt{n}(\hat{\hat{\vartheta}}_n^{\vartheta_1^n} - \hat{\vartheta}_n),\end{aligned}$$

mit einer Zwischenstelle $\overline{\vartheta}_n^{\vartheta_1^n}$ auf der Verbindungsstrecke zwischen $\hat{\hat{\vartheta}}_n^{\vartheta_1^n}$ und $\hat{\vartheta}_n$. Mit den Annahmen folgt daraus für $\vartheta_1^n \in \Theta^1$

$$T_n(\vartheta_1^n) = \sqrt{n}(\hat{\vartheta}_n - \hat{\hat{\vartheta}}_n^{\vartheta_1^n})^t \Big(-\frac{1}{n} \sum_{j=1}^n \frac{\partial^2}{\partial\tilde{\vartheta}_i \partial\tilde{\vartheta}_l} L(\zeta_j, \tilde{\vartheta})_{|_{\tilde{\vartheta}=\overline{\vartheta}_n^{\vartheta_1^n}}} \Big)_{\substack{1\le i\le k\\1\le l\le k}} \sqrt{n}(\hat{\vartheta}_n - \hat{\hat{\vartheta}}_n^{\vartheta_1^n})$$

auf $\{(\zeta_1, \ldots, \zeta_n) \in W_{1,n}\}$ für alle $n \in \mathbb{N}$.

Der Rest folgt analog zum Beweis von Lemma 4.4.5. □

Mit Hilfe der bisherigen Resultate können wir nun ein Ergebnis für die Volumen-Effizienz zweier Likelihood-Quotienten- oder Waldscher Tests angeben.

4.4.7 Korollar: *Es seien T_n oder T_n' jeweils durch die Testgröße (4.4.1) oder durch die Testgröße (4.4.2) gegeben, und es gelten jeweils die Bedingungen (D1) bis (D5). Es sei $\tilde{\vartheta} = (\vartheta_1, \vartheta_2, \vartheta_2')^t \in \Theta$, $\vartheta = (\vartheta_1, \vartheta_2)^t \in \tilde{\Theta}$ und $\vartheta' = (\vartheta_1, \vartheta_2')^t \in \tilde{\Theta}'$. Zudem existieren zu $\tilde{\vartheta}$ nicht negative messbare Abbildungen $m_n : \mathbb{R}^{d_1} \to \mathbb{R}$, $n \in \mathbb{N}$, und $m : \mathbb{R}^{d_1} \to \mathbb{R}$ mit*

$$Q_\vartheta\big(T_n(\vartheta_1 + \vartheta_1^*/\sqrt{n}) < \chi^2_{d_1;1-\alpha}\big) \le m_n(\vartheta_1^*) \text{ für } (\vartheta_1 + \vartheta_1^*/\sqrt{n}) \in \Theta^1$$

und

$$\lim_{n\to\infty} m_n(\vartheta_1^*) = m(\vartheta_1^*) \text{ für alle } \vartheta_1^* \in \mathbb{R}^{d_1}$$

sowie

$$\int m_n(\vartheta_1^*) \mathrm{d}\vartheta_1^* < \infty,\ n \in \mathbb{N},\ \int m(\vartheta_1^*) \mathrm{d}\vartheta_1^* < \infty \text{ und } \lim_{n\to\infty} \int m_n(\vartheta_1^*) \mathrm{d}\vartheta_1^* = \int m(\vartheta_1^*) \mathrm{d}\vartheta_1^*.$$

Ferner existieren nicht negative Abbildungen $m_n' : \mathbb{R}^{d_1} \to \mathbb{R}$, $n \in \mathbb{N}$, und $m' : \mathbb{R}^{d_1} \to \mathbb{R}$ mit

$$M_{\vartheta'}\big(T_n'(\vartheta_1 + \vartheta_1^*/\sqrt{n}) < \chi^2_{d_1;1-\alpha}\big) \le m_n'(\vartheta_1^*) \text{ für } (\vartheta_1 + \vartheta_1^*/\sqrt{n}) \in \Theta^1$$

und

$$\lim_{n\to\infty} m_n'(\vartheta_1^*) = m'(\vartheta_1^*) \text{ für alle } \vartheta_1^* \in \mathbb{R}$$

sowie

$$\int m_n'(\vartheta_1^*) \mathrm{d}\vartheta_1^* < \infty,\ n \in \mathbb{N},\ \int m'(\vartheta_1^*) \mathrm{d}\vartheta_1^* < \infty \text{ und } \lim_{n\to\infty} \int m_n'(\vartheta_1^*) \mathrm{d}\vartheta_1^* = \int m'(\vartheta_1^*) \mathrm{d}\vartheta_1^*.$$

Dann hat man

$$\lim_{n\to\infty} e^{\mathrm{vol}}_{T',T,n}(\tilde{\vartheta}) = \lim_{n\to\infty} \left(\frac{\mathrm{E}_\vartheta(\mathrm{Vol}_{T,n})}{\mathrm{E}_{\vartheta'}(\mathrm{Vol}_{T',n})}\right)^2 = \frac{\det\big(i^{1,1}_{\zeta_1'}(\vartheta') - i^{1,2}_{\zeta_1'}(\vartheta') i^{2,2}_{\zeta_1'}(\vartheta')^{-1} i^{2,1}_{\zeta_1'}(\vartheta')\big)}{\det\big(i^{1,1}_{\zeta_1}(\vartheta) - i^{1,2}_{\zeta_1}(\vartheta) i^{2,2}_{\zeta_1}(\vartheta)^{-1} i^{2,1}_{\zeta_1}(\vartheta)\big)}.$$

Beweis: Dies folgt unmittelbar aus Satz 4.4.4. □

4.4.8 Beispiel: In der Situation aus Beispiel 4.2.9 gilt mit den dortigen Bezeichnungen

$$\lim_{n\to\infty} e^{\mathrm{vol}}_{T',T,n}(\tilde{\vartheta}) = \frac{\Gamma}{2\Sigma} = e^{\mathrm{P}}_{T',T}(\tilde{\vartheta}).$$

Betrachten wir das Modell aus Beispiel 4.2.10, so gilt mit den dortigen Bezeichnungen

$$\lim_{n\to\infty} e^{\mathrm{vol}}_{T',T,n}(\tilde{\vartheta}) = \frac{\det \Gamma_2}{\det \Gamma_1}.$$

Dabei ist zu beachten, dass in diesen Beispielen die Likelihood-Quotienten-Tests jeweils mit den zugehörigen Waldschen Tests übereinstimmen.

4.5 Vergleichskriterien

Wie wir in den vorhergehenden Unterabschnitten sahen, ist die Effizienz zweier Tests oft durch den Quotienten entsprechender quadratischer Formen, genauer durch einen verallgemeinerten Rayleigh-Quotienten, gegeben. Diese Tatsache ist in der einschlägigen Literatur bekannt; siehe hierfür z.B. Puri und Sen ([32]) oder Müller-Funk und Witting ([27]).

Es lohnt zu überlegen, wie ein Qualitätsvergleich von Experimenten anhand solcher Effizienzen vorgenommen werden kann. Wir wollen deshalb in diesem Unterabschnitt davon ausgehen, dass eine Effizienz in der Form der Resultate von Korollar 4.2.5, Korollar 4.3.5 und Korollar 4.3.10 vorliegt.

Zusätzlich betrachten wir Effizienzen, die durch den Quotienten zweiter Determinanten von symmetrisch positv definiten Matrizen gegeben sind. Dies wird durch das Resultat von Korollar 4.4.7 motiviert.

Es sei $\Theta \subset \mathbb{R}^d$ und $\partial\Theta_0 \subset \Theta_0$. Weiter sei $d_1 \in \mathbb{N}$ mit $d_1 \leq d$ und $\emptyset \neq B \subset \mathbb{R}^{d_1} \setminus \{0\}$. Für $\vartheta_0 \in \partial\Theta_0$ und $\vartheta_1 \in B$ sei eine abstrakte Effizienz der Tests T und T' gegeben durch

$$\mathrm{eff}_{T',T}(\vartheta_0, \vartheta_1) := \frac{\vartheta_1^t A_{\zeta_1'}(\vartheta_0)\vartheta_1}{\vartheta_1^t A_{\zeta_1}(\vartheta_0)\vartheta_1}, \tag{4.5.1}$$

oder durch

$$\widetilde{\mathrm{eff}}_{T',T}(\vartheta_0) := \frac{\det A_{\zeta_1'}(\vartheta_0)}{\det A_{\zeta_1}(\vartheta_0)}, \tag{4.5.2}$$

wobei mit $A_{\zeta_1}(\vartheta_0) \in \mathbb{R}^{d_1 \times d_1}$ und $A_{\zeta_1'}(\vartheta_0) \in \mathbb{R}^{d_1 \times d_1}$ symmetrische positiv definite Matrizen gegeben seien, die vom Parameter $\vartheta_0 \in \partial\Theta_0$ abhängen.

Es ist zunächst von Interesse zu fragen, wie klein oder groß die Effizienz $\mathrm{eff}_{T',T}$ sein kann, wenn wir von einem gegebenen ϑ_0 ausgehen. Durch

$$\inf_{\vartheta_1 \in B} \frac{\vartheta_1^t A_{\zeta_1'}(\vartheta_0)\vartheta_1}{\vartheta_1^t A_{\zeta_1}(\vartheta_0)\vartheta_1} \leq \mathrm{eff}_{T',T}(\vartheta_0, \vartheta_1) \leq \sup_{\vartheta_1 \in B} \frac{\vartheta_1^t A_{\zeta_1'}(\vartheta_0)\vartheta_1}{\vartheta_1^t A_{\zeta_1}(\vartheta_0)\vartheta_1} \ \forall\, \vartheta_1 \in B$$

sind untere und obere Schranken für die Effizienz gegeben.

Wir definieren die Matrix

$$E(\vartheta_0) := A_{\zeta_1'}(\vartheta_0) A_{\zeta_1}(\vartheta_0)^{-1},$$

wobei die nach wachsender Größe geordneten Eigenwerte von $E(\vartheta_0)$ durch $0 \leq \lambda_1(\vartheta_0) \leq \cdots \leq \lambda_{d_1}(\vartheta_0)$ gegeben seien sollen.

4.5.1 Lemma: *Es ist*

$$\lambda_1(\vartheta_0) \leq \mathrm{eff}_{T',T}(\vartheta_0, \vartheta_1) \leq \lambda_{d_1}(\vartheta_0) \ \forall \ \vartheta_1 \in B,$$

wobei die angegebenen Schranken im Fall $B = \mathbb{R}^{d_1} \setminus \{0\}$ *angenommen werden.*

Beweis: Da $A_{\zeta_1}(\vartheta_0)$ symmetrisch und positiv definit ist, existiert eine Cholesky-Zerlegung von $A_{\zeta_1}(\vartheta_0)$ gemäß

$$A_{\zeta_1}(\vartheta_0) = C(\vartheta_0) C(\vartheta_0)^t,$$

mit einer regulären Matrix $C(\vartheta_0) \in \mathbb{R}^{d_1 \times d_1}$. Wir setzen

$$D(\vartheta_0) := C(\vartheta_0)^{-1} A_{\zeta_1'}(\vartheta_0) \big(C(\vartheta_0)^{-1} \big)^t$$

und für $\vartheta_1 \in B$

$$\nu := C(\vartheta_0)^t \vartheta_1.$$

Die Matrix $D(\vartheta_0)$ ist symmetrisch und besitzt ebenfalls die Eigenwerte $0 \leq \lambda_1(\vartheta_0) \leq \cdots \leq \lambda_{d_1}(\vartheta_0)$. Ferner ist

$$\mathrm{eff}_{T',T} = \mathrm{eff}_{T',T}(\vartheta_0, \vartheta) = \frac{\nu^t D(\vartheta_0) \nu}{\nu^t \nu}$$

der Rayleigh-Quotient der symmetrischen Matrix $D(\vartheta_0)$ an der Stelle ν. Bekanntlich ist

$$\lambda_1(\vartheta_0) \leq \frac{\nu^t D(\vartheta_0) \nu}{\nu^t \nu} \leq \lambda_{d_1}(\vartheta_0).$$

Im Fall $B = \mathbb{R}^{d_1} \setminus \{0\}$ wird für $\vartheta_1 = (C(\vartheta_0)^t)\nu$ mit ν als Eigenvektor zum Eigenwert $\lambda_1(\vartheta_0)$ bzw. $\lambda_{d_1}(\vartheta_0)$ der Matrix $D(\vartheta_0)$ die untere bzw. obere Schranke angenommen. $\square$

Mit der Matrix

$$\tilde{E}(\vartheta_0) := A_{\zeta_1'}(\vartheta_0) - A_{\zeta_1}(\vartheta_0),$$

und ihrer nach wachsender Größe geordneten Eigenwerten $\tilde{\lambda}_1(\vartheta_0) \leq \cdots \leq \tilde{\lambda}_{d_1}(\vartheta_0)$ gelingt eine weitere für einen Wirksamkeitsvergleich nützliche Feststellung.

4.5.2 Lemma: *Es treffen die Implikationen*

$$\tilde{\lambda}_{d_1}(\vartheta_0) \leq 0 \;\Rightarrow\; \sup_{\vartheta_1 \in B} \operatorname{eff}_{T',T}(\vartheta_0, \vartheta_1) \leq 1,$$

und

$$\tilde{\lambda}_1(\vartheta_0) \geq 0 \;\Rightarrow\; \inf_{\vartheta_1 \in B} \operatorname{eff}_{T',T}(\vartheta_0, \vartheta_1) \geq 1,$$

zu, wobei im Fall $B = \mathbb{R}^{d_1} \setminus \{0\}$ auch die Umkehrung jeder dieser beiden Aussagen gilt.

Beweis: Es ist

$$\begin{aligned}
\operatorname{eff}_{T',T}(\vartheta_0, \vartheta_1) \begin{Bmatrix} \leq \\ \geq \end{Bmatrix} 1 \;&\Leftrightarrow\; \frac{\vartheta_1^t A_{\zeta_1'}(\vartheta_0)\vartheta_1}{\vartheta_1^t A_{\zeta_1}(\vartheta_0)\vartheta_1} \begin{Bmatrix} \leq \\ \geq \end{Bmatrix} 1 \\
&\Leftrightarrow\; \vartheta_1^t A_{\zeta_1'}(\vartheta_0)\vartheta_1 \begin{Bmatrix} \leq \\ \geq \end{Bmatrix} \vartheta_1^t A_{\zeta_1}(\vartheta_0)\vartheta_1 \\
&\Leftrightarrow\; \vartheta_1^t \big(A_{\zeta_1'}(\vartheta_0) - A_{\zeta_1}(\vartheta_0)\big)\vartheta_1 \begin{Bmatrix} \leq \\ \geq \end{Bmatrix} 0 \\
&\Leftrightarrow\; \frac{\vartheta_1^t \tilde{E}(\vartheta_0)\vartheta_1}{\vartheta_1^t \vartheta_1} \begin{Bmatrix} \leq \\ \geq \end{Bmatrix} 0 \;\forall\; \vartheta_1 \in B.
\end{aligned}$$

Eine Argumentation wie im Beweis von Lemma 4.5.1 liefert

$$\tilde{\lambda}_1(\vartheta_0) \leq \frac{\vartheta_1^t \tilde{E}(\vartheta_0)\vartheta_1}{\vartheta_1^t \vartheta_1} \leq \tilde{\lambda}_{d_1}(\vartheta_0) \;\forall\; \vartheta_1 \in B.$$

Dies impliziert

$$\tilde{\lambda}_{d_1}(\vartheta_0) \leq 0 \;\Rightarrow\; \sup_{\vartheta_1 \in B} \operatorname{eff}_{T',T}(\vartheta_0, \vartheta_1) \leq 1$$

und

$$\tilde{\lambda}_1(\vartheta_0) \geq 0 \;\Rightarrow\; \inf_{\vartheta_1 \in B} \operatorname{eff}_{T',T}(\vartheta_0, \vartheta_1) \geq 1$$

sowie im Fall $B = \mathbb{R}^{d_1} \setminus \{0\}$ auch die Umkehrung jeder dieser beiden Aussagen. □

4.5.3 Bemerkung: Existiert eine offene Umgebung $U \subset \mathbb{R}^{d_1}$ um $0 \in \mathbb{R}^{d_1}$ mit $U \setminus \{0\} \subset B$, so kann bei der Betrachtung von Effizienzen der Form (4.5.1) ohne Einschränkung von $B = \mathbb{R}^{d_1} \setminus \{0\}$ ausgegangen werden. Für jedes $\vartheta_1 \in B$ und jedes $a \in \mathbb{R} \setminus \{0\}$ ist nämlich

$$\operatorname{eff}_{T',T}(\vartheta_0, \vartheta_1) = \frac{\vartheta_1^t A_{\zeta_1'}(\vartheta_0)\vartheta_1}{\vartheta_1^t A_{\zeta_1}(\vartheta_0)\vartheta_1} = \frac{a\vartheta_1^t A_{\zeta_1'}(\vartheta_0)a\vartheta_1}{a\vartheta_1^t A_{\zeta_1}(\vartheta_0)a\vartheta_1}.$$

Bei kanonischer Fortsetzung von $\mathrm{eff}_{T',T}(\vartheta_0, \cdot)$ auf $\tilde{B} := \{a\vartheta_1; a \in \mathbb{R} \setminus \{0\}, \vartheta_1 \in B\}$ finden wir

$$\{\mathrm{eff}_{T',T}(\vartheta_0, \vartheta_1); \vartheta_1 \in B\} = \{\mathrm{eff}_{T',T}(\vartheta_0, \vartheta_1); \vartheta_1 \in \tilde{B}\}.$$

Bei $U \setminus \{0\} \subset B$ ist $\tilde{B} = \mathbb{R}^{d_1} \setminus \{0\}$.

Um gewisse Aussagen im Zusammenhang mit der Effizienz $\widetilde{\mathrm{eff}}_{T',T}$ zu erhalten genügt es, die Effizienz $\mathrm{eff}_{T',T}$ zu betrachten, wie das folgende Lemma zeigt.

4.5.4 Lemma: Es liege die Situation aus Bemerkung 4.5.3 vor. Dann gelten folgende Implikationen:

a) $\sup_{\vartheta_1 \in \mathbb{R}^{d_1} \setminus \{0\}} \mathrm{eff}_{T',T}(\vartheta_0, \vartheta_1) < 1 \;\Rightarrow\; \widetilde{\mathrm{eff}}_{T',T}(\vartheta_0) < 1$.

b) $\sup_{\vartheta_1 \in \mathbb{R}^{d_1} \setminus \{0\}} \mathrm{eff}_{T',T}(\vartheta_0, \vartheta_1) \leq 1 \;\Rightarrow\; \widetilde{\mathrm{eff}}_{T',T}(\vartheta_0) \leq 1$.

c) $\mathrm{eff}_{T',T}(\vartheta_0, \vartheta_1) = 1 \;\forall\; \vartheta_1 \in \mathbb{R}^{d_1} \setminus \{0\} \;\Rightarrow\; \widetilde{\mathrm{eff}}_{T',T}(\vartheta_0) = 1$.

d) $\inf_{\vartheta_1 \in \mathbb{R}^{d_1} \setminus \{0\}} \mathrm{eff}_{T',T}(\vartheta_0, \vartheta_1) \geq 1 \;\Rightarrow\; \widetilde{\mathrm{eff}}_{T',T}(\vartheta_0) \geq 1$.

e) $\inf_{\vartheta_1 \in \mathbb{R}^{d_1} \setminus \{0\}} \mathrm{eff}_{T',T}(\vartheta_0, \vartheta_1) > 1 \;\Rightarrow\; \widetilde{\mathrm{eff}}_{T',T}(\vartheta_0) > 1$.

Beweis: Dies sieht man z.B. für Implikation a) wie folgt. Es seien $0 \leq \lambda_1(\vartheta_0) \leq \cdots \leq \lambda_{d_1}(\vartheta_0)$ die nach wachsender Größe geordneten Eigenwerte von $E(\vartheta_0) = A_{\zeta_1'}(\vartheta_0) A_{\zeta_1}(\vartheta_0)^{-1}$. Mit Lemma 4.5.1 folgt aus $\sup_{\vartheta_1 \in \mathbb{R}^{d_1} \setminus \{0\}} \mathrm{eff}_{T',T}(\vartheta_0, \vartheta_1) < 1$ auch $\lambda_{d_1}(\vartheta_0) < 1$. Damit ist

$$\widetilde{\mathrm{eff}}_{T',T}(\vartheta_0) = \prod_{i=1}^{d_1} \lambda_i(\vartheta_0) < 1.$$

Die anderen Implikationen folgen analog. □

Interessiert man sich also für die Frage, welches Experiment eine höhere Effizienz aufweist, so können die Eigenwerte der Matrix $E(\vartheta_0)$ oder $\tilde{E}(\vartheta_0)$ betrachtet werden.

Die beiden Aussagen von Lemma 4.5.2 können natürlich auch äquivalent mit Hilfe der Definitheit der Matrix $\tilde{E}(\vartheta_0)$ formuliert werden, da diese Matrix symmetrisch ist. Schließlich ist eine symmetrische Matrix genau dann positiv semidefinit (negativ semidefinit), wenn alle ihre Eigenwerte positiv oder 0 (negativ oder 0) sind. In diesem Zusammenhang ist auf folgende nützliche Tatsache hinzuweisen: Sind die Matrizen $A_{\zeta_1}(\vartheta_0)$ und $A_{\zeta_1'}(\vartheta_0)$ positiv definit, so ist $A_{\zeta_1'}(\vartheta_0) - A_{\zeta_1}(\vartheta_0)$ genau dann positiv semidefinit (negativ semidefinit), wenn $A_{\zeta_1'}(\vartheta_0)^{-1} - A_{\zeta_1}(\vartheta_0)^{-1}$ negativ semidefinit (positiv semidefinit) ist. Siehe hierzu Lemma A.1.3 im Anhang.

Eine naheliegende Frage ist in diesem Zusammenhang natürlich, wie sich die Wahl des

Parameters $\vartheta_0 \in \partial\Theta_0$ auf die Effizienz oder die genannten Schranken auswirkt. So soll z.B. im Hinblick auf die zu Beginn dieser Arbeit beschriebenen Fragestellung im weiteren Verlauf dieser Arbeit unter anderem der Frage nachgegangen werden, wie sich eine bestimmte Abhängigkeitsstruktur zwischen X_1 und Y_1, wobei (X_1, Y_1) gemäß der verbundenen Stichprobe verteilt sei, auf die Effizienz auswirkt. Diese Abhängigkeitsstruktur wird durch den Parameter $\vartheta_0 \in \partial\Theta_0$ festgelegt.

Es ist von Interesse, bei welcher zugrunde liegender Abhängigkeitsstruktur zwischen X_1 und Y_1 die verbundene Stichprobenerhebung basierend auf den Beobachtungen $\binom{X_1}{Y_1}, \dots, \binom{X_n}{Y_n}$ der unabhängigen Stichprobenerhebung basierend auf den Beobachtungen $\binom{X_1'}{Y_1'}, \dots, \binom{X_n'}{Y_n'}$ auf Grundlage der betrachteten Effizienz vorzuziehen ist.

5 Tests zur relativen Wirksamkeit

5.1 Motivation und Situation

Ein Anwender kann die Wahl eines Experiments von der Größe der Effizienz zweier verschiedener Tests abhängig machen. Wie wir in den vorangegangenen Unterabschnitten sahen, hängen solche Effizienzen allerdings oft von variablen Größen ab, die dem Anwender nicht bekannt sind.

Liegen dem Anwender z.B. Referenzdaten vor, anhand derer er auf die gesuchte Effizienz schließen kann, so befindet er sich erneut in einer statistischen Situation, die zeitlich vor dem eigentlichen Experiment liegt, und davon getrennt zu betrachten ist. Der Anwender kann zum Beispiel erneut mit Hilfe eines statistischen Tests die Frage angehen, ob die betrachtete Effizienz einen gewissen Wert über- oder unterschreitet.

In diesem Unterabschnitt sollen die Definitionen und Notationen aus Unterabschnitt 4.5 übernommen werden. Es sei also $\Theta \subset \mathbb{R}^d$ und $d_1 \in \mathbb{N}$ mit $d_1 \leq d$. Wir gehen davon aus, dass für die zu Beginn von Unterabschnitt 4.5 eingeführte Menge $B \subset \mathbb{R}^{d_1} \setminus \{0\}$ in der vorliegenden Situation $B = \mathbb{R}^{d_1} \setminus \{0\}$ erfüllt ist. Tatsächlich genügt die Existenz einer offenen Umgebung $U \subset \mathbb{R}^{d_1}$ um $0 \in \mathbb{R}^{d_1}$ mit $U \setminus \{0\} \subset B$, damit ohne Eischränkung von $B = \mathbb{R}^{d_1} \setminus \{0\}$ ausgegangen werden kann; siehe hierzu Bemerkung 4.5.3.

Es sei $\partial\Theta_0 \subset \Theta_0$. Für $\vartheta_0 \in \partial\Theta_0$ und $\vartheta_1 \in \mathbb{R}^{d_1} \setminus \{0\}$ sei eine Effizienz der Tests T und T' gegeben durch

$$\mathrm{eff}_{T',T}(\vartheta_0, \vartheta_1) := \frac{\vartheta_1^t A_{\zeta_1'}(\vartheta_0)\vartheta_1}{\vartheta_1^t A_{\zeta_1}(\vartheta_0)\vartheta_1},$$

oder durch

$$\tilde{\mathrm{eff}}_{T',T}(\vartheta_0) := \frac{\det A_{\zeta_1'}(\vartheta_0)}{\det A_{\zeta_1}(\vartheta_0)},$$

wobei mit $A_{\zeta_1}(\vartheta_0) \in \mathbb{R}^{d_1 \times d_1}$ und $A_{\zeta_1'}(\vartheta_0) \in \mathbb{R}^{d_1 \times d_1}$ symmetrische positiv definite Matrizen gegeben seien, die vom Parameter $\vartheta_0 \in \partial\Theta_0$ abhängen.

Die Betrachtung von Effizienzen in dieser Form wird durch die Resultate von Korollar 4.2.5, Korollar 4.3.5, Korollar 4.3.10 und Korollar 4.4.7 motiviert. Wie wir sahen, ergeben sich solche Effizienzen, wenn in beiden Experimenten ein optimales Verfahren in einem bestimmten Sinn gewählt wird, und gewisse Regularitätsvoraussetzungen erfüllt sind. Wie bereits in der Einleitung zu dieser Arbeit begründet, wird ein Anwender in der Regel in beiden Situationen ein optimales Verfahren wählen, wenn ein solches existiert. Daher ist es sinnvoll, solche Effizienzen weiter zu studieren.

In diesem Abschnitt wollen wir hauptsächlich Tests zur Verifizierung der folgenden Hypothesen betrachten: Zur Behandlung der Frage, ob die Effizienz zugehörig zum ersten Experiment stets höher ist als die Effizienz zugehörig zum zweiten Experiment betrachten wir die Hypothese

$$\mathrm{H}_1: \sup_{\vartheta_1 \in \mathbb{R}^{d_1} \setminus \{0\}} \mathrm{eff}_{T',T}(\vartheta_0, \vartheta_1) \leq 1.$$

Zur Behandlung der Frage, ob die Effizienz zugehörig zum zweiten Experiment stets höher ist als die Effizienz zugehörig zum ersten Experiment betrachten wir die Hypothese

$$\mathrm{H}_2: \inf_{\vartheta_1 \in \mathbb{R}^{d_1} \setminus \{0\}} \mathrm{eff}_{T',T}(\vartheta_0, \vartheta_1) \geq 1.$$

Zur Behandlung der Frage, ob die Effizienz zugehörig zum ersten Experiment stets gleich der Effizienz zugehörig zum zweiten Experiment ist betrachten wir die Hypothese

$$\mathrm{H}_3: \ \forall\, \vartheta_1 \in \mathbb{R}^{d_1} \setminus \{0\} : \mathrm{eff}_{T',T}(\vartheta_0, \vartheta_1) = 1.$$

Am Ende dieses Abschnitts werden wir im Rahmen einer Bemerkung konsistente Tests für die Testprobleme

$$\mathrm{H}_4: \ \widetilde{\mathrm{eff}}_{T',T}(\vartheta_0) \leq 1, \ \mathrm{K}_4: \ \widetilde{\mathrm{eff}}_{T',T}(\vartheta_0) > 1$$

sowie

$$\mathrm{H}_5: \ \widetilde{\mathrm{eff}}_{T',T}(\vartheta_0) \geq 1, \ \mathrm{K}_5: \ \widetilde{\mathrm{eff}}_{T',T}(\vartheta_0) < 1$$

und

$$\mathrm{H}_6: \ \widetilde{\mathrm{eff}}_{T',T}(\vartheta_0) = 1, \ \mathrm{K}_6: \ \widetilde{\mathrm{eff}}_{T',T}(\vartheta_0) \neq 1$$

behandeln. Angewendet werden können die erzielten Resultate in den Abschnitten 8 und 9.

5.2 Konsistente Tests

Konkret betrachten wir die Testprobleme mit generellen Alternativen, genauer

$$\mathrm{H}_1: \sup_{\vartheta_1 \in \mathbb{R}^{d_1} \setminus \{0\}} \mathrm{eff}_{T',T}(\vartheta_0, \vartheta_1) \leq 1, \ \mathrm{K}_1: \sup_{\vartheta_1 \in \mathbb{R}^{d_1} \setminus \{0\}} \mathrm{eff}_{T',T}(\vartheta_0, \vartheta_1) > 1,$$

bzw.

$$\mathrm{H}_2: \inf_{\vartheta_1 \in \mathbb{R}^{d_1} \setminus \{0\}} \mathrm{eff}_{T',T}(\vartheta_0, \vartheta_1) \geq 1, \ \mathrm{K}_2: \inf_{\vartheta_1 \in \mathbb{R}^{d_1} \setminus \{0\}} \mathrm{eff}_{T',T}(\vartheta_0, \vartheta_1) < 1.$$

bzw.

$$\mathrm{H}_3: \ \forall\, \vartheta_1 \in \mathbb{R}^{d_1} \setminus \{0\} : \mathrm{eff}_{T',T}(\vartheta_0, \vartheta_1) = 1, \ \mathrm{K}_3: \ \exists\, \vartheta_1 \in \mathbb{R}^{d_1} \setminus \{0\} : \mathrm{eff}_{T',T}(\vartheta_0, \vartheta_1) \neq 1.$$

Unter dem Hinweis auf die Ausführungen im Unterabschnitt 4.5 können die Testprobleme mit Hilfe der Matrix

$$\tilde{E}(\vartheta_0) = A_{\zeta_1'}(\vartheta_0) - A_{\zeta_1}(\vartheta_0)$$

charakterisiert werden. Es sei für $l \in \mathbb{N}$ eine Abbildung

$$p : \partial\Theta_0 \to \mathbb{R}^l$$

gegeben, und wir nehmen an, dass die Matrix $\tilde{E}(\vartheta_0)$ tatsächlich nur von $p(\vartheta_0)$, und nicht von ϑ_0 selber abhängt. Konkret nehmen wir die Existenz einer stetigen Abbildung

$$\tilde{H} : \mathbb{R}^l \to \mathcal{S}_{d_1}$$

an, wobei für $m \in \mathbb{N}$ $\mathcal{S}_m$ die Menge der reellen, symmetrischen $m \times m$-Matrizen sei, mit

$$\tilde{H}\big(p(\vartheta_0)\big) = \tilde{E}(\vartheta_0) \;\forall\; \vartheta_0 \in \partial\Theta_0.$$

Es sei $\vartheta_0 \in \partial\Theta_0$ fest gewählt. Wir setzen $p_0 := p(\vartheta_0)$. Es sei $\hat{p}_n$ ein Schätzer für p_0 , der ohne Einschränkung auf dem zugrunde liegenden statistischen Raum definiert sei, d.h. für jedes $n \in \mathbb{N}$ sei

$$\hat{p}_n : (\Omega, \mathfrak{A}) \to (\mathbb{R}^l, \mathfrak{B}^l).$$

Unter dem Hinweis auf Bemerkung 4.5.2 sollen unsere Testgrößen auf den Eigenwerten der symmetrischen Matrix $\tilde{H}(\hat{p}_n)$ basieren.

5.2.1 Bemerkung: Für $m \in \mathbb{N}$ sei $\mathcal{S}_m$ die Menge der reellen, symmetrischen $m \times m$-Matrizen, $\mathcal{S}_m^i,\; i \in \{1, \ldots, m\}$, die Menge der reellen, symmetrischen $m \times m$-Matrizen mit einfachem i-kleinstem Eigenwert. Dies bedeutet, dass die nach wachsender Größe geordneten Eigenwerte $\lambda_{0,1} \leq \cdots \leq \lambda_{0,m}$ aller $A_0 \in \mathcal{S}_m^i$ $\lambda_{0,1} \leq \cdots \leq \lambda_{0,i-1} < \lambda_{0,i} < \lambda_{0,i+1} \leq \cdots \leq \lambda_{0,m}$ erfüllen. Über den vech-Operator vech : $\mathcal{S}_m \to \mathbb{R}^{\frac{m(m+1)}{2}}$, siehe Beispiel 4.2.10, ist $\mathcal{S}_m$ identifizierbar mit $\mathbb{R}^{\frac{m(m+1)}{2}}$ und in diesem Sinne $\mathcal{S}_m^i$ eine offene Teilmenge von $\mathcal{S}_m$. Schließlich ist die Menge der Matrizen in $\mathcal{S}_m$ mit mehrfachem i-kleinstem Eigenwert in diesem Sinne eine abgeschlossene Teilmenge von $\mathcal{S}_m$, was aus der stetigen Abhängigkeit der Eigenwerte von den Einträgen der zugrunde liegenden Matrix, siehe Satz 1.2 in 5.1.1 in [39], folgt. Für $A_0 \in \mathcal{S}_m^i$ sei $\lambda_{0,i}$ der i-kleinste Eigenwert von A_0 mit dem zugehörigen normierten Eigenvektor $\beta_{0,i}$. Es gibt eine offene Umgebung U von A_0 und bezüglich der entsprechenden Borelschen σ-Algebren messbar wählbare Funktionen $\lambda_i : \mathcal{S}_m \to \mathbb{R},\; \beta_i : \mathcal{S}_m \to \mathbb{R}^m$, deren Einschränkungen auf U stetig differenzierbar sind und die Eigenschaft haben, dass

$$\begin{aligned}
&\lambda_i(A_0) = \lambda_{0,i},\; \beta_i(A_0) = \beta_{0,i},\\
&\forall\; A \in U: \;\; A\beta_i(A) = \lambda_i(A)\beta_i(A),\\
&\forall\; A \in \mathcal{S}_p: \;\; \beta_i(A)^t\beta_i(A) = \mathrm{I}_m
\end{aligned}$$

ist, siehe Theorem 1 in [25]. Es bezeichne $D \in \mathbb{R}^{m^2 \times \frac{m(m+1)}{2}}$ die eindeutig bestimmte Matrix mit der Eigenschaft

$$\forall_{A \in \mathcal{S}_m} : \; \mathrm{vec}(A) = D\,\mathrm{vech}(A),$$

siehe hierzu auch Unterabschnitt 3.8 in [26]. Dabei ist der Operator vech wie in Beispiel 4.2.10 definiert, und der Operator vec für $\ell \in \mathbb{N}$ gemäß

$$\begin{aligned}&\mathrm{vec} : \mathbb{R}^{\ell \times \ell} \to \mathbb{R}^{\ell^2},\\ &\mathrm{vec}(M) := (m_{1,1}, \ldots, m_{\ell,1}, \ldots, m_{\ell,1}, \ldots, m_{\ell,\ell})^t, \; M = (m_{i,j})_{1 \le i,j \le \ell},\end{aligned}$$

gegeben. D heißt Duplikationsmatrix und hat den Rang $\frac{m(m+1)}{2}$. Mit $G = (D^t D)^{-1} D^t$ ist

$$\forall_{A \in \mathcal{S}_m} : \; \mathrm{vech}(A) = G\,\mathrm{vec}(A).$$

Für $B \in \mathcal{S}_m^i$ ist mit $\beta_i = \beta_i(B)$ als Eigenvektor zum Eigenwert $\lambda_i = \lambda_i(B)$ der Gradient

$$\left.\frac{\partial \lambda_i(A)}{\mathrm{vech}(A)}\right|_{A=B} = (\beta_i^t \otimes \beta_i^t) D.$$

Dabei bezeichne $\otimes$ das Kroneckerprodukt für Matrizen. Siehe Theorem 1 in [25].

Die nach wachsender Größe geordneten Eigenwerte von $\tilde{H}(\hat{p}_n)$ seien durch $\tilde{\lambda}_{1,n} \le \cdots \le \tilde{\lambda}_{d_1,n}$ gegeben. Ist für $p \in \mathbb{R}^l$ $\tilde{\lambda}_i(p)$ der Eigenwert der symmetrischen Matrix $\tilde{H}(p)$, so gilt also

$$\tilde{\lambda}_{i,n} := \tilde{\lambda}_i(\hat{p}_n), \; i = 1, \ldots, d_1.$$

Um eine Arbeitsgrundlage zur Konstruktion von Tests für die genannten Testprobleme zu haben, definieren wir die folgenden Bedingungen an den Schätzer $\hat{p}_n$ und an die Matrix $\tilde{E}$:

(E1) Es existiere eine offene Umgebung $U \subset \mathbb{R}^l$ um p_0 und eine stetige Abbildung $V : U \to \mathbb{R}^{l \times l}$, sodass $V(p_0) \in \mathbb{R}^{l \times l}$ eine symmetrische positiv definite Matrix ist, und die Verteilungskonvergenz $\sqrt{n}(\hat{p}_n - p_0) \xrightarrow{\mathrm{v}} \mathrm{N}_l(0, V(p_0))$ bei zugrunde liegendem Parameter ϑ_0 mit $p_0 = p(\vartheta_0)$ für $n \to \infty$ gilt.

(E2) Es existiere eine offene Menge $U \subset \mathbb{R}^l$ um p_0 sodass die Abbildung $\mathrm{vech}(\tilde{H}) : \mathbb{R}^l \to \mathbb{R}^{d_1(d_1+1)/2}$ stetig differenzierbar auf U ist mit zugehöriger Jacobimatrix $\mathcal{J}_{\tilde{H}}(p_0) \in \mathbb{R}^{d_1(d_1+1)/2 \times l}$, die vollen Zeilenrang hat.

(E3) Es sei der kleinste Eigenwert $\tilde{\lambda}_1(p_0)$ sowie der größte Eigenwert $\tilde{\lambda}_{d_1}(p_0)$ der Matrix $\tilde{H}(p_0)$ einfach.

5.2.2 Bemerkung: In vielen Beispielen kann für $\hat{p}_n$ ein Maximum-Likelihood-Schätzer gewählt werden. In diesem Fall ist unter üblichen Voraussetzungen (E1) erfüllt, und die Matrix $V(p_0) \in \mathbb{R}^{l \times l}$ ist durch die inverse Matrix der zugehörigen Fisher-Informationsmatrix gegeben.

5.2.3 Bemerkung: Die üblichen Voraussetzungen, die im Rahmen der asymptotischen Verteilungstheorie von Likelihood-Quotienten-Tests gestellt werden, sind hinreichend für die Bedingung (E1) und die stetige Differenzierbarkeitsbedingung in (E2). Vergleiche hierzu mit den Bedingungen (B1) - (B5) aus Unterabschnitt 4.2 und den Ausführungen in Unterabschnitt 4.2.

Zur Behandlung des Testproblems H_1 gegen K_1 bzw. H_2 gegen K_2 betrachten wir unter der Voraussetzung der Gültigkeit von (E1), (E2) und (E3) für $\hat{p}_n \in U$, mit einer Umgebung U um p_0 wie in den Bedingungen (E1) und (E2) gewählt, die Testgröße

$$F_n^1 := \frac{\sqrt{n}\tilde{\lambda}_{d_1,n}}{\sqrt{M_{d_1,n}}},$$

bzw. die Testgröße

$$F_n^2 := -\frac{\sqrt{n}\tilde{\lambda}_{1,n}}{\sqrt{M_{1,n}}},$$

mit

$$M_{i,n} := w_i(\hat{p}_n)^t \mathcal{J}_{\tilde{H}}(\hat{p}_n) V(\hat{p}_n) \mathcal{J}_{\tilde{H}}(\hat{p}_n)^t w_i(\hat{p}_n),\ i = 1, \dots, d_1,$$

wobei für $i = 1, \dots, d_1$ der Vektor $w_i(p) \in \mathbb{R}^{d_1(d_1+1)/2}$, $p \in U$, durch die Komponenten

$$\big(w_i(p)\big)_j := \big(\beta_i(p)^t \otimes \beta_i(p)^t\big) De_j,\ j = 1, \dots, d_1(d_1+1)/2,$$

festgelegt ist. Dabei ist $\beta_i(p) \in \mathbb{R}^{d_1}$ ein normierter Eigenvektor zum Eigenwert $\tilde{\lambda}_i(p)$ der Matrix $\tilde{H}(p)$, $i = 1, \dots, d_1$, e_j der j-te Einheitsvektor in $\mathbb{R}^{d_1(d_1+1)/2}$, $j = 1, \dots, d_1(d_1+1)/2$, und D die Duplikationsmatrix, wie in Bemerkung 5.2.1 erläutert.

An dieser Stelle soll bemerkt werden, dass unter der Voraussetzung der Gültigkeit von (E1), (E2) und (E3)

$$\lim_{n\to\infty} P_{\vartheta_0}(M_{1,n} \leq 0) = \lim_{n\to\infty} P_{\vartheta_0}(M_{d_1,n} \leq 0) = 0$$

gilt, was sich wie folgt begründen lässt: Es sei

$$M_i := w_i(p_0)^t \mathcal{J}_{\tilde{H}}(p_0) V(p_0) \mathcal{J}_{\tilde{H}}(p_0)^t w_i(p_0),\ i = 1, \dots, d_1.$$

Wegen (E1) und (E2) ist die Matrix $\mathcal{J}_{\tilde{H}}(p_0)V(p_0)\mathcal{J}_{\tilde{H}}(p_0)^t$ positiv definit und es ist $\mathcal{J}_{\tilde{H}} V \mathcal{J}_{\tilde{H}}^t$ in einer Umgebung um p_0 stetig. Für $i = 1, \dots, d_1$ kann der Vektor $w_i(p_0)$ nicht verschwinden, da $\beta_i(p_0)$ nicht verschwindet. Daher gilt $M_1 > 0$ und $M_{d_1} > 0$. Wegen der Stetigkeit der Abbildung $\mathrm{vech}(\tilde{H})$ auf einer Umgebung U um p_0 gemäß (E2) und wegen (E3) sind β_1 und β_{d_1} auf einer Umgebung um p_0 stetig und damit auch w_1 und w_{d_1} auf einer Umgebung um p_0 stetig. Betrachte hierzu die oben stehende Argumentation sowie Theorem

1 in [25]. Da (E1) die stochastische Konvergenz von $\hat{p}_n$ gegen p_0 impliziert, folgt mit dem Satz von der stetigen Abbildung die stochastische Konvergenz von $M_{1,n}$ gegen M_1 sowie die stochastische Konvergenz von $M_{d_1,n}$ gegen M_{d_1} und damit

$$\lim_{n\to\infty} P_{\vartheta_0}(M_{1,n} \leq 0) = \lim_{n\to\infty} P_{\vartheta_0}(M_1 - M_{1,n} \geq M_1) \leq \lim_{n\to\infty} P_{\vartheta_0}(|M_1 - M_{1,n}| \geq M_1) = 0,$$

sowie analog

$$\lim_{n\to\infty} P_{\vartheta_0}(M_{d_1,n} \leq 0) = \lim_{n\to\infty} P_{\vartheta_0}(M_{d_1} - M_{d_1,n} \geq M_{d_1}) \leq \lim_{n\to\infty} P_{\vartheta_0}(|M_{d_1} - M_{d_1,n}| \geq M_{d_1}) = 0.$$

Um das Grenzverhalten der genannten Eigenwerte zu ermitteln, und damit das der Testgrößen zu erhalten, verallgemeinern wir zunächst einen Satz aus dem Buch von Anderson ([1] Theorem 13.5.1), der die asymptotische Verteilung von Eigenwerten von Wishart-verteilten Matrizen beschreibt, auf allgemein symmetrische und asymptotisch normalverteilte Matrizen. Hierbei lässt sich Andersons Argumentation größtenteils übernehmen.

5.2.4 Lemma: *Es sei für alle $m \in \mathbb{N}$ $Z_m \in \mathbb{R}^{q\times q}$ eine zufällige symmetrische Matrix mit den nach wachsender Größe geordneten Eigenwerten $l_m^1 \leq \cdots \leq l_m^q$. Ferner gelte $\sqrt{m}(\mathrm{vech}(Z_m) - \mathrm{vech}(Z)) \xrightarrow{\mathrm{v}} \mathrm{N}_{q(q+1)/2}(0, A)$ für $m \to \infty$, mit einer symmetrischen Matrix $Z \in \mathbb{R}^{q\times q}$ und einer symmetrischen positiv semidefiniten Matrix $A \in \mathbb{R}^{q(q+1)/2\times q(q+1)/2}$, wobei Z die nach wachsender Größe geordneten Eigenwerte $\lambda_1 \leq \cdots \leq \lambda_q$ besitze. Für ein gegebenes $i \in \{1, \ldots, q\}$ sei der Eigenwert λ_i einfach. Dann gilt*

$$\sqrt{m}(l_m^i - \lambda_i) \xrightarrow{\mathrm{v}} \mathrm{N}(0, w_i^t A w_i) \text{ für } m \to \infty,$$

wobei der Vektor $w_i \in \mathbb{R}^{q(q+1)/2}$ durch die Komponenten $(w_i)_j := \big(\beta_i^t \otimes \beta_i^t\big) D e_j,\ j = 1, \ldots, q(q+1)/2$, festgelegt ist. Dabei ist $\beta_i \in \mathbb{R}^{q\times q}$ ein normierter Eigenvektor zum Eigenwert λ_i der Matrix Z.

Beweis: Wir betrachten die Abbildungen

$$L_i : \{\mathrm{vech}(B); B \in \mathbb{R}^{q\times q} \text{ ist symmetrisch}\} \to \mathbb{R},\ i = 1, \ldots, q,$$

wobei $L_1(\mathrm{vech}(B)) \leq \cdots \leq L_q(\mathrm{vech}(B))$ die nach wachsender Größe angeordneten Eigenwerte der symmetrischen Matrix $B \in \mathbb{R}^{q\times q}$ seien. Mit Theorem 7 und Note 2 des Unterabschnitts 6.8 in [26] erhält man wegen der Einfachheit des Eigenwertes λ_i, dass w_i der Grandient von L_i an der Stelle $\mathrm{vech}(Z)$ ist. Mit dem Fehlerfortpflanzungsgesetz folgt die Behauptung. □

Nach diesen Vorbereitungen können wir nun den folgenden Satz formulieren, um das Grenzverhalten der Testgrößen zu beschreiben. Hierzu seien $\tilde{\lambda}_1 := \tilde{\lambda}_1(p_0), \ldots, \tilde{\lambda}_{d_1} := \tilde{\lambda}_{d_1}(p_0)$ die Eigenwerte der Matrix $\tilde{H}(p_0)$.

5.2.5 Satz: *Es gelte (E1) - (E3). Dann sind folgende Aussagen erfüllt:*

a) Falls $\tilde{\lambda}_{d_1} = 0$, so folgt $F_n^1 \xrightarrow{\mathrm{v}} \mathrm{N}(0,1)$ für $n \to \infty$, und falls $\tilde{\lambda}_1 = 0$, so folgt $F_n^2 \xrightarrow{\mathrm{v}} \mathrm{N}(0,1)$ für $n \to \infty$.

b) Falls $\tilde{\lambda}_{d_1} < 0$, so folgt für alle $c \in \mathbb{R}$ $\lim_{n\to\infty} P_{\vartheta_0}(F_n^1 \geq c) = 0$, und falls $\tilde{\lambda}_1 > 0$, so folgt für alle $c \in \mathbb{R}$ $\lim_{n\to\infty} P_{\vartheta_0}(F_n^2 \geq c) = 0$.

c) Falls $\tilde{\lambda}_{d_1} > 0$, so folgt für alle $c \in \mathbb{R}$ $\lim_{n\to\infty} P_{\vartheta_0}(F_n^1 \leq c) = 0$, und falls $\tilde{\lambda}_1 < 0$, so folgt für alle $c \in \mathbb{R}$ $\lim_{n\to\infty} P_{\vartheta_0}(F_n^2 \leq c) = 0$.

Beweis: Nur für F_n^1; für F_n^2 analog. Es gelte $\tilde{\lambda}_{d_1} = 0$. Aufgrund von (E1) und (E2) sowie des Fehlerfortpflanzungsgesetzes gilt

$$\sqrt{n}\Big(\text{vech}\left(\tilde{H}(\hat{p}_n)\right) - \text{vech}\left(\tilde{H}(p_0)\right)\Big) \xrightarrow{\text{v}} \mathrm{N}_{d_1(d_1+1)/2}\big(0, \mathcal{J}_{\tilde{H}}(p_0)V(p_0)\mathcal{J}_{\tilde{H}}(p_0)^t\big) \text{ für } \; n \to \infty.$$

Da die Matrizen $\tilde{H}(\hat{p}_n)$ und $\tilde{H}(p_0)$ symmetrisch sind, folgt mit

$$M_{d_1} = w_{d_1}(p_0)^t \mathcal{J}_{\tilde{H}}(p_0)V(p_0)\mathcal{J}_{\tilde{H}}(p_0)^t w_{d_1}(p_0),$$

sowie wegen (E3), mit Hilfe von Lemma 5.2.4

$$\sqrt{n}\tilde{\lambda}_{d_1,n} \xrightarrow{\text{v}} \mathrm{N}(0, M_{d_1}) \text{ für } \; n \to \infty.$$

Wie oben begründet hat man die stochastische Konvergenz von $M_{d_1,n}$ gegen M_{d_1}. Mit dem Lemma von Slutsky folgt die Verteilungskonvergenz der Testgröße gegen eine Standardnormalverteilung.

Nun gelte $\tilde{\lambda}_{d_1} < 0$. Da der größte Eigenwert eine stetige Funktion in den Einträgen der zugrunde liegenden Matrix ist, siehe Satz 1.2 in 5.1.1 in [39], und wegen der Stetigkeit der Abbildung vech$(\tilde{H})$ auf einer Umgebung U um p_0 gemäß (E2), folgt mit dem Satz von der stetigen Abbildung die stochastische Konvergenz von $\tilde{\lambda}_{d_1,n}$ gegen $\tilde{\lambda}_{d_1}$. Daraus folgt für beliebiges $c \in \mathbb{R}$ unter Berücksichtigung der stochastischen Konvergenz von $M_{d_1,n}$ gegen M_{d_1}

$$\begin{aligned}\lim_{n\to\infty} P_{\vartheta_0}(F_n^1 \geq c) &= \lim_{n\to\infty} P_{\vartheta_0}\Big(\tilde{\lambda}_{d_1,n} - \tilde{\lambda}_{d_1} \geq \frac{\sqrt{M_{d_1,n}}c}{\sqrt{n}} - \tilde{\lambda}_{d_1}\Big) \\ &\leq \lim_{n\to\infty} P_{\vartheta_0}\Big(|\tilde{\lambda}_{d_1,n} - \tilde{\lambda}_{d_1}| \geq \frac{\sqrt{M_{d_1,n}}c}{\sqrt{n}} - \tilde{\lambda}_{d_1}\Big) = 0.\end{aligned}$$

Im Fall $\tilde{\lambda}_{d_1} > 0$ kann analog argumentiert werden. □

Für ein gegebenes Testniveau $\alpha \in (0, 1)$ schlagen wir also für das Testproblem H_i gegen K_i, $i = 1, 2$, den Test

$$\text{„Verwirf } \mathrm{H}_i, \text{ falls } F_n^i \geq u_{1-\alpha}\text{“}, \; i = 1, 2,$$

vor, wobei $u_{1-\alpha}$ das $(1-\alpha)$-Quantil der Standardnormalverteilung sei. Dann halten wir folgendes fest:

5.2.6 Korollar: *Es gelte (E1) - (E3). Dann handelt es sich bei dem auf der Testgröße F_n^i basierenden vorgeschlagenen Test für das Testproblem* H_i *gegen* K_i*,* $i = 1, 2$*, um einen Test, der für jedes* $\alpha \in (0, 1)$ *das Testniveau* α *asymptotisch einhält, und der zudem konsistent ist.*

Beweis: Nur für $i = 1$, sonst analog. Es gelte die Hypothese, d.h. es gilt

$$\sup_{\vartheta_1 \in \mathbb{R}^{d_1} \setminus \{0\}} \operatorname{eff}_{T',T}(\vartheta_0, \vartheta_1) \leq 1.$$

Da in Lemma 4.5.2 in der vorliegenden Situation auch die Umkehrungen gelten, folgt $\tilde{\lambda}_{d_1} \leq 0$, und damit mit Satz 5.2.5 Teil a) und b), dass der Test das Testniveau asymptotisch einhält. Nun gelte die Alternative, also

$$\sup_{\vartheta_1 \in \mathbb{R}^{d_1} \setminus \{0\}} \operatorname{eff}_{T',T}(\vartheta_0, \vartheta_1) > 1.$$

Da in Lemma 4.5.2 in der vorliegenden Situation auch die Umkehrungen gelten, folgt $\tilde{\lambda}_{d_1} > 0$, und damit mit Satz 5.2.5 Teil c), dass der Test konsistent ist. □

Nun möchten wir einen konsistenten Test für das Testproblem H_3 gegen K_3 studieren. Hierzu seien die von ihrem Betrag her nach wachsender Größe geordneten Eigenwerte von $\tilde{H}(\hat{p}_n)$ durch $\tilde{\lambda}_{(1),n}, \ldots, \tilde{\lambda}_{(d_1),n}$ gegeben, d.h. es gilt $|\tilde{\lambda}_{(1),n}| \leq \cdots \leq |\tilde{\lambda}_{(d_1),n}|$. Analog seien die von ihrem Betrag her nach wachsender Größe geordneten Eigenwerte von $\tilde{H}(p_0)$ durch $\tilde{\lambda}_{(1)}, \ldots, \tilde{\lambda}_{(d_1)}$ gegeben, d.h. es gilt $|\tilde{\lambda}_{(1)}| \leq \cdots \leq |\tilde{\lambda}_{(d_1)}|$.

Wir definieren die Testgröße

$$F_n^3 := n\tilde{\lambda}^2_{(d_1),n}.$$

Zudem sei für $p \in U$, mit einer Umgebung U um p_0 wie in den Bedingungen (E1) und (E2) gewählt, $c_{p;1-\alpha}$ das $(1-\alpha)$-Quantil der Verteilung $\mathcal{L}(L_{d_1}|p)$, wobei $L_1 \leq \cdots \leq L_{d_1}$ die nach wachsender Größe geordneten Eigenwerte einer Matrix $A^2 \in \mathbb{R}^{d_1 \times d_1}$ seien, und $A \in \mathbb{R}^{d_1 \times d_1}$ eine symmetrische Matrix mit

$$\operatorname{vech}(A) \sim \mathrm{N}_{d_1(d_1+1)/2}\big(0, \mathcal{J}_{\tilde{H}}(p)V(p)\mathcal{J}_{\tilde{H}}(p)^t\big) \tag{5.2.1}$$

sei. Es bezeichne $\mathcal{L}(A|p)$ die Verteilung von A bei zugrunde liegendem $p \in U$ und $\mathcal{L}(A^2|p)$ die Verteilung von A^2 bei zugrunde liegendem $p \in U$. Die Matrix A^2 ist in jedem Fall positiv semidefinit und wegen (5.2.1) sogar mit Wahrscheinlichkeit 1 positiv definit. Schließlich ist die Determinante von A ein nicht identisch verschwindendes Polynom in Veränderlichen gegeben durch die Komponenten von $\operatorname{vech}(A)$. Wegen (5.2.1) und der positiven Definitheit von $\mathcal{J}_{\tilde{H}}(p_0)V(p_0)\mathcal{J}_{\tilde{H}}(p_0)^t$ folgt mit dem Lemma aus [30] $\det A \neq 0$ mit Wahrscheinlichkeit 1.

Nach Theorem 3.2.17 aus [28] besitzen die Eigenwerte einer positiv definiten $d_1 \times d_1$-Matrix mit Lebesgue-Dichte f die gemeinsame Lebesgue-Dichte

$$(l_1, \ldots, l_{d_1}) \longmapsto \frac{\pi^{\frac{{d_1}^2}{2}}}{\Gamma_{d_1}\big(\frac{d_1}{2}\big)} \prod_{1 \leq j < i \leq d_1} (l_i - l_j) \int_{O_{d_1}} f\big(H \operatorname{diag}(l_{d_1}, \ldots l_1) H'\big) \mathrm{d}H,$$
$$\text{für } l_{d_1} > \cdots > l_1 > 0, \text{ und } 0 \text{ sonst.}$$

Hierbei wird das innere Integral über die Menge $O_{d_1} := \{A \in \mathbb{R}^{d_1 \times d_1}; A^t A = \mathrm{I}_{d_1}\}$ bezüglich der Gleichverteilung auf O_{d_1} durchgeführt und Γ_m ist die verallgemeinerte Gammafunktion $\Gamma_m(t) := \pi^{m(m-1)/4} \prod_{i=1}^m \Gamma(t - (i-1)/2),\ m \in \mathbb{N},\ t > (m-1)/2$.

Ausintegrieren über $l_1, l_2, \ldots, l_{d_1-1}$ liefert daher für $\mathcal{L}(L_{d_1}|p)$ die zugehörige Lebesgue-Dichte

$$k_p(l_{d_1}) := \int \cdots \int \frac{\pi^{\frac{d_1^2}{2}}}{\Gamma_{d_1}\big(\frac{d_1}{2}\big)} \prod_{1 \leq j < i \leq d_1} (l_i - l_j) \int_{O_{d_1}} f_p\big(H \operatorname{diag}(l_{d_1}, \ldots, l_1) H'\big) \mathrm{d}H \mathrm{d}l_1 \ldots \mathrm{d}l_{d_1 - 1},$$
$$\text{für } l_{d_1} > 0, \text{ und } 0 \text{ sonst,} \tag{5.2.2}$$

wobei f_p die Dichte der oben genannten Matrix A^2 bei zugrunde liegendem $p \in U$ sei.

Für ein gegebenes Testniveau $\alpha \in (0,1)$ schlagen wir für das Testproblem H_3 gegen K_3 den Test

$$\text{„Verwirf } \mathrm{H}_3, \text{ falls } F_n^3 \geq c_{\hat{p}_n; 1-\alpha}\text{“}$$

vor. Um zu zeigen, dass dieser Test asymptotisch das Testniveau exakt einhält, und um die Konsistenz des Tests zu zeigen, werden wir nun folgenden Satz beweisen:

5.2.7 Satz: *Es gelte (E1) und (E2). Für die Eigenwerte der Matrix $\tilde{H}(p_0)$ gelte $\tilde{\lambda}_1 = \cdots = \tilde{\lambda}_{d_1} = 0$. Dann folgt, dass F_n^3 für $n \to \infty$ in Verteilung gegen $\mathcal{L}(L_{d_1}|p_0)$ konvergiert. Falls dagegen $\tilde{\lambda}_i \neq 0$ für ein $i \in \{1, \ldots, d_1\}$, so folgt für alle $c \in \mathbb{R}$*

$$\lim_{n \to \infty} P_{\vartheta_0}(F_n^3 \geq c) = 1.$$

Beweis: Es gelte $\tilde{\lambda}_1 = \cdots = \tilde{\lambda}_{d_1} = 0$. Da die Matrix $\tilde{H}(p_0)$ symmetrisch ist, impliziert dies $\tilde{H}(p_0) = 0$. Dann gilt natürlich auch $\tilde{H}(p_0)^2 = 0$. Aufgrund von (E1) und (E2) sowie des Fehlerfortpflanzungsgesetzes gilt analog zum Beweis von Satz 5.2.5

$$\sqrt{n}\Big(\operatorname{vech}\big(\tilde{H}(\hat{p}_n)\big) - \operatorname{vech}\big(\tilde{H}(p_0)\big)\Big) \xrightarrow{\mathrm{v}} \mathrm{N}_{d_1(d_1+1)/2}\big(0, \mathcal{J}_{\tilde{H}}(p_0) V(p_0) \mathcal{J}_{\tilde{H}}(p_0)^t\big) \text{ für } n \to \infty.$$

Wegen $\tilde{H}(p_0) = 0$ konvergiert also $\sqrt{n}\tilde{H}(\hat{p}_n)$ für $n \to \infty$ in Verteilung gegen $\mathcal{L}(A|p_0)$. Daher konvergiert $n\tilde{H}(\hat{p}_n)^2$ für $n \to \infty$ in Verteilung gegen $\mathcal{L}(A^2|p_0)$. Da $n\tilde{\lambda}^2_{(d_1),n}$ der größte

Eigenwert der Matrix $n\tilde{H}(\hat{p}_n)^2$ ist, und die extremalen Eigenwerte eine stetige Funktion der zugrunde liegenden Matrix sind (siehe z.B. Theorem 5.2 in [37]), folgt die erste Behauptung.

Nun gelte $\tilde{\lambda}_i \neq 0$ für ein $i \in \{1, \dots, d_1\}$. Dies impliziert $\tilde{H}(p_0) \neq 0$ und natürlich auch $\tilde{H}(p_0)^2 \neq 0$. Da der größte Eigenwert eine stetige Funktion der zugrunde liegenden Matrix ist, mit der stochastischen Konvergenz von $\hat{p}_n$ gegen p_0 und aufgrund der Stetigkeit der beteiligten Abbildungen, die in (E1) und (E2) gefordert ist, folgt mit dem Satz von der stetigen Abbildung die stochastische Konvergenz von $\tilde{\lambda}^2_{(d_1),n}$ gegen $\tilde{\lambda}^2_{(d_1)}$. Da $\tilde{\lambda}^2_{(d_1)}$ der größte Eigenwert der symmetrischen positiv semidefiniten Matrix $\tilde{H}(p_0)^2$ ist, und diese Matrix nicht verschwindet, gilt $\tilde{\lambda}_{(d_1)} > 0$. Somit folgt für beliebiges $c \in \mathbb{R}$

$$\lim_{n\to\infty} P_{\vartheta_0}(F_n^3 \geq c) = \lim_{n\to\infty} P_{\vartheta_0}\left(\tilde{\lambda}^2_{(d_1),n} - \tilde{\lambda}^2_{(d_1)} \geq \frac{c}{n} - \tilde{\lambda}^2_{(d_1)}\right) = 1,$$

und es ist alles gezeigt. □

5.2.8 Korollar: *Es gelte (E1) und (E2). Bei dem auf der Testgröße F_n^3 basierenden vorgeschlagenen Test für das Tesproblem H_3 gegen $\tilde{\mathrm{K}}_3$ handelt es sich um einen Test, der für jedes $\alpha \in (0,1)$ das Testniveau α asymptotisch exakt einhält, und der zudem konsistent ist.*

Beweis: Es gelte die Hypothese, d.h. es gilt

$$\forall\ \vartheta_1 \in \mathbb{R}^{d_1} \setminus \{0\} : \mathrm{eff}_{T',T}(\vartheta_0, \vartheta_1) = 1.$$

Da in Lemma 4.5.2 in der vorliegenden Situation auch die Umkehrungen gelten, folgt $\tilde{\lambda}_1 = \tilde{\lambda}_{d_1} = 0$. Mit Satz 5.2.7 folgt für $n \to \infty$ die Verteilungskonvergenz von F_n^3 gegen $\mathcal{L}(L_{d_1}|p_0)$ bei zugrunde liegendem p_0. Da L_{d_1} eine Lebesgue-Dichte und damit eine stetige Verteilungsfunktion besitzt, gilt

$$\lim_{n\to\infty} \sup_{x\in\mathbb{R}} |P_{\vartheta_0}(F_n^3 \geq x) - P_{\vartheta_0}(L_{d_1} \geq x)| = 0. \tag{5.2.3}$$

Es sei $(p_n)_{n\in\mathbb{N}}$ eine reelle Folge mit $p_n \in U$ für alle $n \in \mathbb{N}$ und $\lim_{n\to\infty} p_n = p_0$. Dann folgt aufgrund der Stetigkeit der beteiligten Abbildungen, die in (E1) und (E2) gefordert ist, für $n \to \infty$ mit dem Lemma von Scheffé die schwache Konvergenz der Verteilung

$$\mathrm{N}_{d_1(d_1+1)/2}\big(0, \mathcal{J}_{\tilde{H}}(p_n)V(p_n)\mathcal{J}_{\tilde{H}}(p_n)^t\big)$$

gegen die Verteilung

$$\mathrm{N}_{d_1(d_1+1)/2}\big(0, \mathcal{J}_{\tilde{H}}(p_0)V(p_0)\mathcal{J}_{\tilde{H}}(p_0)^t\big).$$

Dies impliziert die schwache Konvergenz von $\mathcal{L}(A^2|p_n)$ gegen $\mathcal{L}(A^2|p_0)$ für $n \to \infty$. Da die extremalen Eigenwerte eine stetige Funktion der zugrunde liegenden Matrix sind folgt mit dem Satz von der stetigen Abbildung für $n \to \infty$ die schwache Konvergenz von $\mathcal{L}(L_{d_1}|p_n)$ gegen $\mathcal{L}(L_{d_1}|p_0)$. Da $\mathcal{L}(L_{d_1}|p_0)$ eine streng monoton wachsende Verteilungsfunktion besitzt, was man anhand der Gestalt der zugehörigen Dichte in (5.2.2) erkennt, konvergieren die Quantile mit, d.h.

$$\lim_{n\to\infty} c_{p_n;1-\alpha} = c_{p_0;1-\alpha}.$$

Da (E1) auch die stochastische Konvergenz von $\hat{p}_n$ gegen p_0 impliziert, folgt gemäß dieser Argumentation die stochastische Konvergenz von $c_{\hat{p}_n;1-\alpha}$ gegen $c_{p_0;1-\alpha}$, und somit zusammen mit (5.2.3), dass der Test das Testniveau asymptotisch exakt einhält.

Nun gelte die Alternative, also

$$\exists\, \vartheta_1 \in \mathbb{R}^{d_1} \setminus \{0\} : \mathrm{eff}_{T',T}(\vartheta_0, \vartheta_1) \neq 1.$$

Dann folgt mit Lemma 4.5.2 $\tilde{\lambda}_i \neq 0$ für ein $i \in \{1, \ldots, d_1\}$. Mit Satz 5.2.7 und der stochastischen Konvergenz von $c_{\hat{p}_n;1-\alpha}$ gegen $c_{p_0;1-\alpha}$ folgt daraus die Konsistenz des Tests, womit alles gezeigt ist. □

5.2.9 Bemerkung: Die in (E3) genannte Bedingung, dass der kleinste und größte Eigenwert der Matrix $\tilde{H}(p_0)$ einfach ist, ist eine nichttriviale Bedingung, auf die aus technischen Gründen nicht verzichtet werden kann. Dagegen kann bei den Tests, die im folgenden Unterabschnitt 5.3 betrachtet werden, auf eine solche Bedingung verzichtet werden. Dafür sind die dort betrachteten Tests nur unter bestimmten aber sinnvollen Alternativen konsistent.

5.2.10 Bemerkung: In manchen Anwendungen, wie z.B. in den Abschnitten 8 und 9, ist es sinnvoll Tests basierend auf den Eigenwerten des empirischen Gegenstücks zur Matrix

$$\tilde{\tilde{E}}(\vartheta_0) = A_{\zeta_1'}(\vartheta_0)^{-1} - A_{\zeta_1}(\vartheta_0)^{-1}$$

anstelle der vorgeschlagenen Tests basierend auf den Eigenwerten des empirischen Gegenstücks zur Matrix

$$\tilde{E}(\vartheta_0) = A_{\zeta_1'}(\vartheta_0) - A_{\zeta_1}(\vartheta_0)$$

zu betrachten. In der Tat können die in diesem Unterabschnitt behandelten Testprobleme äquivalent gemäß

$$\mathrm{H}_1: \ \tilde{E}(\vartheta_0) \text{ ist negativ semidefinit}, \ \mathrm{K}_1: \ \tilde{E}(\vartheta_0) \text{ ist nicht negativ semidefinit},$$

bzw.

$$\mathrm{H}_2: \ \tilde{E}(\vartheta_0) \text{ ist positiv semidefinit}, \ \mathrm{K}_2: \ \tilde{E}(\vartheta_0) \text{ ist nicht positv semidefinit},$$

bzw.

$$\mathrm{H}_3: \ \tilde{E}(\vartheta_0) = 0, \ \mathrm{K}_3: \ \tilde{E}(\vartheta_0) \neq 0$$

formuliert werden. Aufgrund von Lemma A.1.3 im Anhang ist aber auch eine äquivalente Formulierung der Testprobleme gemäß

$$\mathrm{H}_1: \ \tilde{\tilde{E}}(\vartheta_0) \text{ ist positiv semidefinit}, \ \mathrm{K}_1: \ \tilde{\tilde{E}}(\vartheta_0) \text{ ist nicht positiv semidefinit},$$

bzw.

$$\mathrm{H}_2: \ \tilde{\tilde{E}}(\vartheta_0) \text{ ist negativ semidefinit}, \ \mathrm{K}_2: \ \tilde{\tilde{E}}(\vartheta_0) \text{ ist nicht negativ semidefinit},$$

bzw.

$$\mathrm{H}_3: \ \tilde{\tilde{E}}(\vartheta_0) = 0, \ \mathrm{K}_3: \ \tilde{\tilde{E}}(\vartheta_0) \neq 0$$

möglich. Die Konstruktion von Tests basierend auf den Eigenwerten des empirischen Gegenstücks zur Matrix $\tilde{\tilde{E}}(\vartheta_0)$ sowie das Studium des Grenzverhaltens der zugehörigen Testgrößen funktioniert analog zu den in diesem Unterabschnitt präsentierten Überlegungen.

5.2.11 Bemerkung: Die in diesem Unterabschnitt vorgeschlagenen Tests basieren auf einer Charakterisierung der Testprobleme mit Hilfe der Matrix

$$\tilde{E}(\vartheta_0) = A_{\zeta_1'}(\vartheta_0) - A_{\zeta_1}(\vartheta_0).$$

Dieses Vorgehen wurde motiviert durch das Resultat in Lemma 4.5.2. Motiviert durch das Resultat von Lemma 4.5.1 können mit Hilfe der gleichen Methoden alternativ auch Tests basierend auf einer Charakterisierung der Testprobleme mit Hilfe der Matrix

$$E(\vartheta_0) = A_{\zeta_1'}(\vartheta_0) A_{\zeta_1}(\vartheta_0)^{-1}$$

studiert werden, die das Testniveau asymptotisch einhalten, und zudem konsistent sind.

5.3 Konsistente Tests unter sinnvollen Alternativen

In diesem Unterabschnitt sollen alternative Tests zu denen in Unterabschnitt 5.2 vorgestellten Tests studiert werden, die unter bestimmten sinnvollen Alternativen ebenfalls konsistent sind. Annahmen und Bezeichnungen aus Unterabschnitt 5.2 werden übernommen.

Wie wir in Korollar 5.2.6 sahen, wird im Zusammenhang mit den in Unterabschnitt 5.2 vorgeschlagenen Tests aus technischen Gründen gefordert, dass die Matrix $\tilde{E}(\vartheta_0)$ bei Gültigkeit der Hypothese H_i, $i = 1, 2$, paarweise verschiedene Eigenwerte besitzt. Dagegen kann bei den in diesem Unterabschnitt studierten Tests auf solch eine Bedingung verzichtet werden. Im Gegenzug sind allerdings die hier betrachteten Tests nur unter bestimmten aber sinnvollen Alternativen konsistent.

Wir betrachten nun die Testprobleme mit eingeschränkten Alternativen, d.h.

$$\mathrm{H}_1: \sup_{\vartheta_1 \in \mathbb{R}^{d_1} \setminus \{0\}} \mathrm{eff}_{T',T}(\vartheta_0, \vartheta_1) \leq 1, \ \tilde{\mathrm{K}}_1: \inf_{\vartheta_1 \in \mathbb{R}^{d_1} \setminus \{0\}} \mathrm{eff}_{T',T}(\vartheta_0, \vartheta_1) > 1,$$

bzw.

$$\mathrm{H}_2: \inf_{\vartheta_1 \in \mathbb{R}^{d_1}\setminus\{0\}} \mathrm{eff}_{T',T}(\vartheta_0,\vartheta_1) \geq 1,\ \tilde{\mathrm{K}}_2: \sup_{\vartheta_1 \in \mathbb{R}^{d_1}\setminus\{0\}} \mathrm{eff}_{T',T}(\vartheta_0,\vartheta_1) < 1,$$

bzw.

$$\begin{aligned} &\mathrm{H}_3: \ \forall\, \vartheta_1 \in \mathbb{R}^{d_1} \setminus \{0\} : \mathrm{eff}_{T',T}(\vartheta_0,\vartheta_1) = 1, \\ &\tilde{\mathrm{K}}_3: \inf_{\vartheta_1 \in \mathbb{R}^{d_1}\setminus\{0\}} \mathrm{eff}_{T',T}(\vartheta_0,\vartheta_1) > 1 \text{ oder } \sup_{\vartheta_1 \in \mathbb{R}^{d_1}\setminus\{0\}} \mathrm{eff}_{T',T}(\vartheta_0,\vartheta_1) < 1. \end{aligned}$$

5.3.1 Bemerkung: Bei den betrachteten Alternativen handelt es sich um sinnvolle Alternativen im folgenden Sinn: Handelt es sich bei dem Test für das Testproblem H_1 gegen $\tilde{\mathrm{K}}_1$ um einen unter $\tilde{\mathrm{K}}_1$ konsistenten Test, so konvergiert die Wahrscheinlichkeit die Hypothese, dass die Effizienz zugehörig zum ersten Experiment stets besser ist als die Effizienz zugehörig zum zweiten Experiment, nicht zu verwerfen, im Fall, dass die Effizienz zugehörig zum zweiten Experiment stets besser ist als die Effizienz zugehörig zum ersten Experiment, gegen null.

Handelt es sich bei dem Test für das Testproblem H_2 gegen $\tilde{\mathrm{K}}_2$ um einen unter $\tilde{\mathrm{K}}_2$ konsistenten Test, so konvergiert die Wahrscheinlichkeit, die Hypothese, dass die Effizienz zugehörig zum zweiten Experiment stets besser ist als die Effizienz zugehörig zum ersten Experiment nicht zu verwerfen, im Fall, dass die Effizienz zugehörig zum ersten Experiment stets besser ist als die Effizienz zugehörig zum zweiten Experiment, gegen null.

Handelt es sich bei dem Test für das Testproblem H_3 gegen $\tilde{\mathrm{K}}_3$ um einen unter $\tilde{\mathrm{K}}_3$ konsistenten Test, so konvergiert die Wahrscheinlichkeit, die Hypothese, dass die Effizienz zugehörig zum zweiten Experiment stets gleich der Effizienz zugehörig zum ersten Experiment ist nicht zu verwerfen, im Fall, dass die Effizienz zugehörig zum ersten Experiment stets besser ist als die Effizienz zugehörig zum zweiten Experiment, oder dass die Effizienz zugehörig zum zweiten Experiment stets besser ist als die Effizienz zugehörig zum ersten Experiment, gegen null.

Wie in Unterabschnitt 5.2 sollen die hier betrachteten Tests auf den Eigenwerten des empirischen Gegenstückes zur Matrix

$$\tilde{E}(\vartheta_0) = A_{\zeta_1'}(\vartheta_0) - A_{\zeta_1}(\vartheta_0)$$

basieren.

Konkret sollen die hier betrachteten Testgrößen auf der Spur der Matrix basieren, und wir werden uns zu Nutze machen, dass

$$\mathrm{tr}\big(\tilde{E}(\vartheta_0)\big) = \tilde{\lambda}_1 + \cdots + \tilde{\lambda}_{d_1}$$

gilt. Schließlich gelten die Implikationen

$$\tilde{\lambda}_{d_1} \begin{Bmatrix} < \\ \leq \end{Bmatrix} 0 \;\Rightarrow\; \operatorname{tr}\big(\tilde{E}(\vartheta_0)\big) \begin{Bmatrix} < \\ \leq \end{Bmatrix} 0$$

und

$$\tilde{\lambda}_1 \begin{Bmatrix} > \\ \geq \end{Bmatrix} 0 \;\Rightarrow\; \operatorname{tr}\big(\tilde{E}(\vartheta_0)\big) \begin{Bmatrix} > \\ \geq \end{Bmatrix} 0.$$

Lassen wir die betrachteten Testgrößen auf der Spur der Matrix $\tilde{H}(\hat{p}_n)$ basieren, so vermeiden wir technische Komplikationen beim Studium des Grenzverhaltens der Testgröße, so wie sie in Bemerkung 5.2.9 beschrieben sind und müssen keine einschränkende Bedingungen an die Eigenwerte der Matrix $\tilde{H}(p_0)$ stellen, so wie es in Korollar 5.2.6 notwendig ist.

Zur Behandlung des Testproblems H_i gegen $\tilde{\mathrm{K}}_i$, $i = 1, 2, 3$, definieren wir unter der Voraussetzung der Gültigkeit von (E1) und (E2) für $\hat{p}_n \in U$, mit einer Umgebung U um p_0 wie in den Bedingungen (E1) und (E2) gewählt, die Testgrößen

$$\tilde{F}_n^1 := \frac{\sqrt{n}\operatorname{tr}\big(\tilde{H}(\hat{p}_n)\big)}{\sqrt{\tilde{M}_n}},\; \tilde{F}_n^2 := -\tilde{F}_n^1,\; \tilde{F}_n^3 := |\tilde{F}_n^1|,$$

mit

$$\tilde{M}_n := \tilde{M}_n(\hat{p}_n) := v^t \mathcal{J}_{\tilde{H}}(\hat{p}_n) V(\hat{p}_n) \mathcal{J}_{\tilde{H}}(\hat{p}_n)^t v,$$

wobei der Verktor $v \in \mathbb{R}^{d_1(d_1+1)/2}$ durch

$$v := (1, \underbrace{0, \ldots, 0}_{(d_1-1)\text{-mal}}, 1, \underbrace{0, \ldots, 0}_{(d_1-2)\text{-mal}}, 1, \ldots, 1, 0, 1)$$

gegeben ist. Es kann ähnlich wie in Unterabschnitt 5.2 argumentiert werden um die stochastische Konvergenz von $\tilde{M}_n$ gegen $\tilde{M} := \tilde{M}(p_0)$ zu zeigen und damit

$$\lim_{n\to\infty} P_{\vartheta_0}(\tilde{M}_n \leq 0) = 0$$

zu begründen.

Dann können wir bezüglich des Grenzverhaltens der Testgrößen folgendes Resultat formulieren:

5.3.2 Satz: *Es gelte (E1) und (E2). Dann gilt:*

a) Falls $\tilde{\lambda}_1 = \cdots = \tilde{\lambda}_{d_1} = 0$, so folgt $\tilde{F}_n^i \xrightarrow{\mathrm{v}} \mathrm{N}(0,1)$ für $n \to \infty$, $i = 1, 2$ und $\tilde{F}_n^3 \xrightarrow{\mathrm{v}} \mathrm{HN}(0,1)$ für $n \to \infty$, wobei $\mathrm{HN}(0,1)$ die Verteilung des Betrages einer standardnormalverteilten Zufallsvariable bezeichne.

b) Falls $\tilde{\lambda}_1, \ldots, \tilde{\lambda}_{d_1} \leq 0$, und $\tilde{\lambda}_i < 0$ für ein $i \in \{1, \ldots, d_1\}$, so folgt für alle $c \in \mathbb{R}$ $\lim_{n\to\infty} P_{\vartheta_0}(\tilde{F}_n^1 \geq c) = 0$, und für alle $c \in \mathbb{R}$ $\lim_{n\to\infty} P_{\vartheta_0}(\tilde{F}_n^i \leq c) = 0,\ i = 2, 3$.

c) Falls $\tilde{\lambda}_1, \ldots, \tilde{\lambda}_{d_1} \geq 0$, und $\tilde{\lambda}_i > 0$ für ein $i \in \{1, \ldots, d_1\}$, so folgt für alle $c \in \mathbb{R}$ $\lim_{n\to\infty} P_{\vartheta_0}(\tilde{F}_n^i \leq c) = 0,\ i = 1, 3$, und für alle $c \in \mathbb{R}$ $\lim_{n\to\infty} P_{\vartheta_0}(\tilde{F}_n^2 \geq c) = 0$.

Beweis: Nur für $i = 1$; sonst kann analog vorgegangen werden. Aufgrund von (E1) und (E2) sowie des Fehlerfortpflanzungsgesetzes gilt analog zum Beweis von Satz 5.2.5

$$\sqrt{n}\Big(\operatorname{vech}\big(\tilde{H}(\hat{p}_n)\big) - \operatorname{vech}\big(\tilde{H}(p_0)\big)\Big) \xrightarrow{\text{v}} \mathrm{N}_{d_1(d_1+1)/2}\big(0, \mathcal{J}_{\tilde{H}}(p_0)V(p_0)\mathcal{J}_{\tilde{H}}(p_0)^t\big) \text{ für } n \to \infty.$$

Fassen wir hier die Spur als Abbildung $\operatorname{tr} : \mathbb{R}^{d_1(d_1+1)/2} \to \mathbb{R}$ auf, so folgt durch erneute Anwendung des Fehlerfortpflanzungsgesetzes

$$\sqrt{n}\Big(\operatorname{tr}\big(\tilde{H}(\hat{p}_n)\big) - \operatorname{tr}\big(\tilde{H}(p_0)\big)\Big) \xrightarrow{\text{v}} \mathrm{N}_{d_1(d_1+1)/2}\big(0, v^t\mathcal{J}_{\tilde{H}}(p_0)V(p_0)\mathcal{J}_{\tilde{H}}(p_0)^t v\big) \text{ für } n \to \infty.$$

Weiter ist

$$\operatorname{tr}\big(\tilde{H}(p_0)\big) = \tilde{\lambda}_1 + \cdots + \tilde{\lambda}_{d_1}.$$

Nun gelte $\tilde{\lambda}_1 = \cdots = \tilde{\lambda}_{d_1} = 0$. Dann gilt

$$\operatorname{tr}\big(\tilde{H}(p_0)\big) = 0,$$

und mit

$$\tilde{M} = v^t\mathcal{J}_{\tilde{H}}(p_0)V(p_0)\mathcal{J}_{\tilde{H}}(p_0)^t v$$

folgt

$$\sqrt{n}\operatorname{tr}\big(\tilde{H}(\hat{p}_n)\big) \xrightarrow{\text{v}} \mathrm{N}(0, \tilde{M}) \text{ für } n \to \infty.$$

Hieraus folgt mit der Stetigkeit der beteiligten Abbildungen, die in (E1) und (E2) gefordert ist, und da (E1) auch die stochastische Konvergenz von $\hat{p}_n$ gegen p_0 impliziert, mit dem Satz von der stetigen Abbildung und dem Lemma von Slutsky die Verteilungskonvergenz der Testgröße gegen eine Standardnormalverteilung.

Falls $\tilde{\lambda}_1, \ldots, \tilde{\lambda}_{d_1} \leq 0$, und $\tilde{\lambda}_i < 0$ für ein $i \in \{1, \ldots, d_1\}$, so folgt

$$\operatorname{tr}\big(\tilde{H}(p_0)\big) < 0.$$

Da die Spur eine stetige Funktion der zugrunde liegenden Matrix ist, sowie mit der Stetigkeit der beteiligten Abbildungen, die in (E1) und (E2) gefordert ist, und da (E1) auch die stochastische Konvergenz von $\hat{p}_n$ gegen p_0 impliziert, folgt mit dem Satz von der stetigen

Abbildung die stochastische Konvergenz von $\mathrm{tr}\big(\tilde{H}(\hat{p}_n)\big)$ gegen $\mathrm{tr}\big(\tilde{H}(p_0)\big)$. Daraus folgt für beliebiges $c \in \mathbb{R}$ unter Berücksichtigung der stochastischen Konvergenz von $\tilde{M}_n$ gegen $\tilde{M}$

$$\begin{aligned}\lim_{n\to\infty} P_{\vartheta_0}(\tilde{F}_n^1 \geq c) &= \lim_{n\to\infty} P_{\vartheta_0}\Big(\mathrm{tr}\big(\tilde{H}(\hat{p}_n)\big) - \mathrm{tr}\big(\tilde{H}(p_0)\big) \geq \frac{\sqrt{\tilde{M}_n} c}{\sqrt{n}} - \mathrm{tr}\big(\tilde{H}(p_0)\big)\Big) \\ &\leq \lim_{n\to\infty} P_{\vartheta_0}\Big(\big|\mathrm{tr}\big(\tilde{H}(\hat{p}_n)\big) - \mathrm{tr}\big(\tilde{H}(p_0)\big)\big| \geq \frac{\sqrt{\tilde{M}_n} c}{\sqrt{n}} - \mathrm{tr}\big(\tilde{H}(p_0)\big)\Big) = 0.\end{aligned}$$

Falls $\tilde{\lambda}_1, \ldots, \tilde{\lambda}_{d_1} \geq 0$, und $\tilde{\lambda}_i > 0$ für ein $i \in \{1, \ldots, d_1\}$, so folgt

$$\mathrm{tr}\big(\tilde{H}(p_0)\big) > 0,$$

und es kann analog argumentiert werden. □

Für ein gegebenes Testniveau $\alpha \in (0, 1)$ lautet der zugehörige Test für das Testproblem H_i gegen $\tilde{\mathrm{K}}_i$, $i = 1, 2$,

$$„\text{Verwirf } \mathrm{H}_i, \text{ falls } \tilde{F}_n^i \geq u_{1-\alpha}“, \ i = 1, 2.$$

Für dass Testproblem H_3 gegen K_3 kann der Test

$$„\text{Verwirf } \mathrm{H}_3, \text{ falls } \tilde{F}_n^3 \geq u_{1-\alpha/2}“$$

formuliert werden.

5.3.3 Korollar: *Es gelte (E1) und (E2). Dann handelt es sich bei dem auf der Testgröße* $\tilde{F}_n^i$ *basierenden vorgeschlagenen Test für das Testproblem* H_i *gegen* $\tilde{\mathrm{K}}_i$, $i = 1, 2, 3$, *um einen Test, der für jedes* $\alpha \in (0, 1)$ *das Testniveau* α *asymptotisch einhält und der zudem konsistent ist. Der auf der Testgröße* $\tilde{F}_n^3$ *basierenden vorgeschlagenen Test hält zudem das Testniveau* α *asymptotisch exakt ein.*

Beweis: Nur für $i = 1$, sonst analog. Es gelte die Hypothese, also

$$\sup_{\vartheta_1 \in \mathbb{R}^{d_1} \setminus \{0\}} \mathrm{eff}_{T',T}(\vartheta_0, \vartheta_1) \leq 1.$$

Mit Lemma 4.5.1 folgt dann $\tilde{\lambda}_1 \leq \cdots \leq \tilde{\lambda}_{d_1} \leq 0$, sodass mit Satz 5.3.2 Teil a) und b) folgt, dass der Test das Testniveau asymptotisch einhält. Nun gelte die Alternative, also

$$\inf_{\vartheta_1 \in \mathbb{R}^{d_1} \setminus \{0\}} \mathrm{eff}_{T',T}(\vartheta_0, \vartheta_1) > 1.$$

Mit Lemma 4.5.1 folgt dann $0 < \tilde{\lambda}_1 \leq \cdots \leq \tilde{\lambda}_{d_1}$, sodass mit Satz 5.3.2 Teil c) folgt, dass der Test konsistent ist. □

5.3.4 Bemerkung: In manchen Anwendungen, wie z.B. in Unterabschnitt 8.12, ist es sinnvoll Tests basierend auf den Eigenwerten des empirischen Gegenstückes zur Matrix

$$\tilde{\tilde{E}}(\vartheta_0) := A_{\zeta_1'}(\vartheta_0)^{-1} - A_{\zeta_1}(\vartheta_0)^{-1}$$

anstelle der vorgeschlagenen Tests basierend auf den Eigenwerten des empirischen Gegenstückes zur Matrix

$$\tilde{E}(\vartheta_0) = A_{\zeta_1'}(\vartheta_0) - A_{\zeta_1}(\vartheta_0)$$

zu betrachten. In der Tat können die in diesem Unterabschnitt behandelten Testprobleme äquivalent gemäß

$$\mathrm{H}_1 : \ \tilde{E}(\vartheta_0) \text{ ist negativ semidefinit, } \tilde{\mathrm{K}}_1 : \ \tilde{E}(\vartheta_0) \text{ ist positiv definit,}$$

und

$$\mathrm{H}_2 : \ \tilde{E}(\vartheta_0) \text{ ist positiv semidefinit, } \tilde{\mathrm{K}}_2 : \ \tilde{E}(\vartheta_0) \text{ ist negativ definit}$$

sowie

$$\mathrm{H}_3 : \ \tilde{E}(\vartheta_0) = 0, \ \tilde{\mathrm{K}}_3 : \ \tilde{E}(\vartheta_0) \text{ ist positiv oder negativ definit}$$

formuliert werden. Aufgrund von Lemma A.1.3 im Anhang ist aber auch eine äquivalente Formulierung der Testprobleme gemäß

$$\mathrm{H}_1 : \ \tilde{\tilde{E}}(\vartheta_0) \text{ ist positiv semidefinit, } \tilde{\mathrm{K}}_1 : \ \tilde{\tilde{E}}(\vartheta_0) \text{ ist negativ definit,}$$

und

$$\mathrm{H}_2 : \ \tilde{\tilde{E}}(\vartheta_0) \text{ ist negativ semidefinit, } \tilde{\mathrm{K}}_2 : \ \tilde{\tilde{E}}(\vartheta_0) \text{ ist positiv definit}$$

sowie

$$\mathrm{H}_3 : \ \tilde{\tilde{E}}(\vartheta_0) = 0, \ \tilde{\mathrm{K}}_3 : \ \tilde{\tilde{E}}(\vartheta_0) \text{ ist positiv oder negativ definit}$$

möglich. Die Konstruktion von Tests basierend auf den Eigenwerten des empirischen Gegenstücks zur Matrix $\tilde{\tilde{E}}(\vartheta_0)$ sowie das Studium des Grenzverhaltens der zugehörigen Testgrößen funktioniert analog zu den in diesem Unterabschnitt präsentierten Überlegungen.

5.3.5 Bemerkung: Die in diesem Unterabschnitt vorgeschlagenen Tests basierend auf der Spur des empirischen Gegenstückes zur Matrix

$$\tilde{E}(\vartheta_0) = A_{\zeta_1'}(\vartheta_0) - A_{\zeta_1}(\vartheta_0).$$

Dieses Vorgehen wurde motiviert durch das Resultat in Lemma 4.5.2. Motiviert durch das Resultat von Lemma 4.5.1 können mit Hilfe der gleichen Methoden alternativ auch Tests basierend auf der Spur des empirischen Gegenstückes zur Matrix

$$E(\vartheta_0) = A_{\zeta_1'}(\vartheta_0) A_{\zeta_1}(\vartheta_0)^{-1}$$

studiert werden, die das Testniveau asymptotisch einhalten, und zudem konsistent sind. Mit Hilfe der gleichen Methoden können in diesem Fall auch Tests basierend auf der Determinante des empirischen Gegenstückes zur Matrix $E(\vartheta_0)$ studiert werden. Dabei kann man sich zu Nutze machen, dass

$$\det\big(E(\vartheta_0)\big) = \prod_{i=1}^{d_1} \lambda_i$$

gilt, wenn $\lambda_1 \leq \cdots \leq \lambda_{d_1}$ die nach wachsender Größe geordneten Eigenwerte der Matrix $E(\vartheta_0)$ sind. Schließlich gilt wegen $0 \leq \lambda_1 \leq \cdots \leq \lambda_{d_1}$, was wegen Lemma 4.5.1 gilt, auch

$$\lambda_{d_1} \begin{Bmatrix} < \\ \leq \end{Bmatrix} 1 \;\Rightarrow\; \det\big(E(\vartheta_0)\big) \begin{Bmatrix} < \\ \leq \end{Bmatrix} 1,$$

sowie

$$\lambda_1 \begin{Bmatrix} > \\ \geq \end{Bmatrix} 1 \;\Rightarrow\; \det\big(E(\vartheta_0)\big) \begin{Bmatrix} > \\ \geq \end{Bmatrix} 1.$$

Diese Tests basierend auf der Determinante des empirischen Gegenstückes zur Matrix $E(\vartheta_0)$ dienen dann auch als konsistente Tests für die Testprobleme

$$\mathrm{H}_4: \ \widetilde{\mathrm{eff}}_{T',T}(\vartheta_0) \leq 1, \ \mathrm{K}_4: \ \widetilde{\mathrm{eff}}_{T',T}(\vartheta_0) > 1$$

sowie

$$\mathrm{H}_5: \ \widetilde{\mathrm{eff}}_{T',T}(\vartheta_0) \geq 1, \ \mathrm{K}_5: \ \widetilde{\mathrm{eff}}_{T',T}(\vartheta_0) < 1$$

und

$$\mathrm{H}_6: \ \widetilde{\mathrm{eff}}_{T',T}(\vartheta_0) = 1, \ \mathrm{K}_6: \ \widetilde{\mathrm{eff}}_{T',T}(\vartheta_0) \neq 1.$$

6 Vergleich von verbundener und unabhängiger Stichprobenerhebung

6.1 Modell

Aufbauend auf den Ergebnissen aus Abschnitt 4 studieren wir in diesem Abschnitt die in Unterabschnitt 4.5 definierte Effizienz in der Situation, dass es sich bei den Experimenten um die verbundene und unabhängige Stichprobenerhebung handelt, so wie es zu Beginn dieser Arbeit beschrieben ist. Die hier erzielten Ergebnisse können in den Abschnitten 8 und 9 angewendet werden.

Es soll also die verbundene Stichprobenerhebung mit der unabhängigen Stichprobenerhebung verglichen werden. Orientiert am einführenden Beispiel in Unterabschnitt 1.1 betrachten wir dazu folgendes Modell: Für $m \geq 2$ sind im Fall der verbundenen Stichprobenerhebung durch $\binom{X_1}{Y_1}, \ldots, \binom{X_n}{Y_n}$ unabhängig und identisch verteilte Zufallsvektoren $\binom{X_i}{Y_i} : (\Omega, \mathfrak{A}) \to (\mathbb{R}^m, \mathfrak{B}^m)$, $i = 1, \ldots, n$, gegeben, sodass im Allgemeinen für $\vartheta \in \Theta$

$$P_\vartheta^{(X_1,Y_1)} \neq P_\vartheta^{X_1} \otimes P_\vartheta^{Y_1}$$

gilt. Es ist $\zeta_1 = w\binom{X_1}{Y_1}$, mit einer Abbildung $w : (\mathbb{R}^m, \mathfrak{B}^m) \to (\mathbb{R}^l, \mathfrak{B}^l)$, mit $l \in \mathbb{N}$. Es gilt also für den Bildraum $(R, \mathfrak{S})$ der Zufallsvariablen ζ_1 $R \subseteq \mathbb{R}^l$ und $\mathfrak{S} = \mathfrak{B}^l_R$. Falls m eine gerade Zahl ist und $X_i : (\Omega, \mathfrak{A}) \to (\mathbb{R}^{m/2}, \mathfrak{B}^{m/2})$, $Y_i : (\Omega, \mathfrak{A}) \to (\mathbb{R}^{m/2}, \mathfrak{B}^{m/2})$, $i = 1, \ldots, n$, betrachten wir häufig den Fall

$$w(x) = (\mathrm{I}_{m/2 \times m/2}, 0_{m/2 \times m/2})x - (0_{m/2 \times m/2}, \mathrm{I}_{m/2 \times m/2})x, \; x \in \overline{\mathbb{R}}^m, \tag{6.1.1}$$

d.h. $\zeta_1 = X_1 - Y_1$.

Im Fall der unabhängigen Stichprobenerhebung sind durch $\binom{X_1'}{Y_1'}, \ldots, \binom{X_n'}{Y_n'}$ unabhängig und identisch verteilte Zufallsvektoren $\binom{X_i'}{Y_i'} : (\Omega, \mathfrak{A}) \to (\mathbb{R}^m, \mathfrak{B}^m)$, $i = 1, \ldots, n$, gegeben, sodass für alle $\vartheta \in \Theta$

$$P_\vartheta^{(X_1',Y_1')} = P_\vartheta^{X_1'} \otimes P_\vartheta^{Y_1'} \tag{6.1.2}$$

gilt. Es ist $\zeta_1' = w'\binom{X_1'}{Y_1'}$, mit einer Abbildung $w' : (\mathbb{R}^m, \mathfrak{B}^m) \to (\mathbb{R}^{l'}, \mathfrak{B}^{l'})$, mit $l' \in \mathbb{N}$. Es gilt also für den Bildraum $(R', \mathfrak{S}')$ der Zufallsvariablen ζ_1' $R' \subseteq \mathbb{R}^{l'}$ und $\mathfrak{S}' = \mathfrak{B}^{l'}_{R'}$. Falls m eine gerade Zahl ist und $X_i' : (\Omega, \mathfrak{A}) \to (\mathbb{R}^{m/2}, \mathfrak{B}^{m/2})$, $Y_i' : (\Omega, \mathfrak{A}) \to (\mathbb{R}^{m/2}, \mathfrak{B}^{m/2})$, $i = 1, \ldots, n$, betrachten wir auch hier häufig den Fall, dass w' analog zu (6.1.1) gegeben ist, d.h. $\zeta_1' = X_1' - Y_1'$. Zudem gilt für alle $\vartheta \in \Theta$

$$P_\vartheta^{X_1} = P_\vartheta^{X_1'} \text{ und } P_\vartheta^{Y_1} = P_\vartheta^{Y_1'}. \tag{6.1.3}$$

Wir gehen hier von einer bestimmten Modellannahme aus, die ein Spezialfall der bereits in Unterabschnitt 4.1 getroffenen Modellannahme darstellt. Es handelt sich um die Situation aus Bemerkung 4.2.8 und Unterabschnitt 4.4. Wir gehen von der Existenz eines $\vartheta_0^1 \in \Theta^1$ mit $\{\vartheta_0^1\} \subsetneq \Theta^1$ und

$$\Theta_0 = \{(\vartheta_1, \vartheta_2, \vartheta_2') \in \Theta; \vartheta_1 = \vartheta_0^1\}$$

aus. Es ist wegen $\emptyset \neq \{\vartheta_0^1\} \subsetneq \Theta^1$ auch $\emptyset \neq \Theta_0 \subsetneq \Theta$. Wir setzen $\Theta_1^1 := \Theta^1 \setminus \{\vartheta_0^1\}$, und weiter

$$\tilde{\Theta}_0 := \{(\vartheta_1, \vartheta_2) \in \tilde{\Theta}; \vartheta_1 = \vartheta_0^1\},\ \tilde{\Theta}_1 := \tilde{\Theta} \setminus \tilde{\Theta}_0,$$

sowie

$$\tilde{\Theta}_0' := \{(\vartheta_1, \vartheta_2') \in \tilde{\Theta}'; \vartheta_1 = \vartheta_0^1\},\ \tilde{\Theta}_1' := \tilde{\Theta}' \setminus \tilde{\Theta}_0'.$$

Dann gilt

$$\Theta^1 = \{\vartheta_1 \in \mathbb{R}^{d_1}; \exists\ \vartheta_2 \in \mathbb{R}^{k-d_1} : (\vartheta_1, \vartheta_2) \in \tilde{\Theta}\} = \{\vartheta_1 \in \mathbb{R}^{d_1}; \exists\ \vartheta_2' \in \mathbb{R}^{k'-d_1} : (\vartheta_1, \vartheta_2') \in \tilde{\Theta}'\}$$

und

$$\Theta_1^1 = \{\vartheta_1 \in \mathbb{R}^{d_1}; \exists\ \vartheta_2 \in \mathbb{R}^{k-d_1} : (\vartheta_1, \vartheta_2) \in \tilde{\Theta}_1\} = \{\vartheta_1 \in \mathbb{R}^{d_1}; \exists\ \vartheta_2' \in \mathbb{R}^{k'-d_1} : (\vartheta_1, \vartheta_2') \in \tilde{\Theta}_1'\}.$$

Dies lässt erkennen, dass wegen $\emptyset \neq \{\vartheta_0^1\} \subsetneq \Theta^1$ auch $\emptyset \neq \tilde{\Theta}_0 \subsetneq \tilde{\Theta}$ sowie $\emptyset \neq \tilde{\Theta}_0' \subsetneq \tilde{\Theta}'$ gilt.

Nun kann mit $\vartheta = (\vartheta_1, \vartheta_2, \vartheta_2') \in \Theta$ das Testproblem (1.2.1),

$$\mathrm{H}:\ \vartheta \in \Theta_0,\ \mathrm{K}:\ \vartheta \in \Theta_1,$$

äquivalent gemäß

$$\mathrm{H}:\ \vartheta_1 = \vartheta_0^1,\ \mathrm{K}:\ \vartheta_1 \in \Theta_1^1,$$

bzw.

$$\mathrm{H}:\ (\vartheta_1, \vartheta_2) \in \tilde{\Theta}_0,\ \mathrm{K}:\ (\vartheta_1, \vartheta_2) \in \tilde{\Theta}_1,$$

bzw.

$$\mathrm{H}:\ (\vartheta_1, \vartheta_2') \in \tilde{\Theta}_0',\ \mathrm{K}:\ (\vartheta_1, \vartheta_2') \in \tilde{\Theta}_1',$$

formuliert werden.

Wir nehmen an, dass die Mengen $\tilde{\Theta}$ und $\tilde{\Theta}'$ offen und konvex sind. Hier sollten zur Effizienzuntersuchung sinnvollerweise Alternativen der Form

$$\begin{pmatrix} \vartheta_0^1 \\ \vartheta_2 \\ \vartheta_2' \end{pmatrix} + \begin{pmatrix} \vartheta_1 \\ 0 \\ 0 \end{pmatrix}$$

mit $\vartheta_1 \in \mathbb{R}^{d_1} \setminus \{0\}$, $(\vartheta_0^1, \vartheta_2, \vartheta_2') \in \Theta_0$, betrachtet werden, die $(\vartheta_0^1 + \vartheta_1, \vartheta_2, \vartheta_2') \in \Theta_1$ erfüllen.

Im Folgenden wollen wir Annahmen an die Zufallsvariablen ζ_1 und ζ_1' bei den zugrunde liegenden injektiven Parametrisierungen durch $\tilde{\Theta}$ und $\tilde{\Theta}'$ formulieren. Der Einfachheit halber formulieren wir diese Annahmen nur für ζ_1 und $\tilde{\Theta}$; sie sollen natürlich in analoger Weise auch für ζ_1' und $\tilde{\Theta}'$ gelten. Dabei werden die im Zusammenhang mit ζ_1 und $\tilde{\Theta}$ eingeführten Größen im Zusammenhang mit ζ_1' und $\tilde{\Theta}'$ mit einem Strich versehen.

Wir nehmen an, dass ein σ-endliches Maß μ auf $(R, \mathfrak{S})$ existiert, sodass für jedes $\vartheta \in \tilde{\Theta}$ $Q_\vartheta^{\zeta_1} \ll \mu$ gilt, und $f(\cdot;\vartheta)$ eine μ-Dichte von $Q_\vartheta^{\zeta_1}$ ist. Wir setzen generell voraus, dass $f(z;\vartheta) > 0$ für alle $\vartheta \in \tilde{\Theta}$ und alle $z \in R$. Es sei die Abbildung $\vartheta \mapsto f(z;\vartheta)$, $\vartheta \in \tilde{\Theta}$, für alle $z \in R$ stetig differenzierbar. Zudem nehmen wir an, dass die Fisher-Informationsmatrix aus $Q_\vartheta^{\zeta_1}$, $i_{\zeta_1}(\vartheta)$, für alle $\vartheta \in \tilde{\Theta}$ existiert, endliche Einträge besitzt und positiv definit ist. Mit der am Anfang von Unterabschnitt 4.2 eingeführten Notation und $\vartheta = (\vartheta_1, \ldots, \vartheta_k)^t \in \tilde{\Theta}$ ist dann

$$i_{\zeta_1}(\vartheta) = \left(\mathrm{E}_{\vartheta_0} \left(\frac{\partial}{\partial \vartheta_i} L(\zeta_1, \vartheta) \frac{\partial}{\partial \vartheta_j} L(\zeta_1, \vartheta) \right) \right)_{\substack{1 \le i \le k \\ 1 \le j \le k}}.$$

Dann sei

$$i_{\zeta_1}(\vartheta) = \begin{pmatrix} i_{\zeta_1}^{1,1}(\vartheta) & i_{\zeta_1}^{1,2}(\vartheta) \\ i_{\zeta_1}^{2,1}(\vartheta) & i_{\zeta_1}^{2,2}(\vartheta) \end{pmatrix},$$

mit $i_{\zeta_1}^{1,1}(\vartheta) \in \mathbb{R}^{d_1 \times d_1}$, $i_{\zeta_1}^{1,2}(\vartheta) \in \mathbb{R}^{d_1 \times (k-d_1)}$, $i_{\zeta_1}^{2,1}(\vartheta) \in \mathbb{R}^{(k-d_1) \times d_1}$, $i_{\zeta_1}^{2,2}(\vartheta) \in \mathbb{R}^{(k-d_1) \times (k-d_1)}$. D.h. $i_{\zeta_1}^{1,1}(\vartheta)$ ist die Informationsmatrix aus ζ_1, wenn für $\vartheta = (\vartheta_1, \vartheta_2) \in \tilde{\Theta}$ der Parameter $\vartheta_1 \in \Theta^1$ als einziger variabler Parameter und ϑ_2 als fest und gegeben betrachtet wird.

Ferner soll für alle $\vartheta \in \tilde{\Theta}$ $\mathrm{E}_\vartheta(\zeta_1)$ existieren und endliche Einträge besitzen. Falls m eine gerade Zahl ist und $X_i : (\Omega, \mathfrak{A}) \to (\mathbb{R}^{m/2}, \mathfrak{B}^{m/2})$, $Y_i : (\Omega, \mathfrak{A}) \to (\mathbb{R}^{m/2}, \mathfrak{B}^{m/2})$, $i = 1, \ldots, n$, verlangen wir zusätzlich, dass für alle $(\vartheta_1, \vartheta_2, \vartheta_2') \in \Theta$ $\mathrm{E}_{(\vartheta_1, \vartheta_2, \vartheta_2')}(X_1 - Y_1)$, $\mathrm{E}_{(\vartheta_1, \vartheta_2, \vartheta_2')}(X_1)$ und $\mathrm{E}_{(\vartheta_1, \vartheta_2, \vartheta_2')}(Y_1)$ existieren und endliche Einträge besitzen. Falls w gemäß (6.1.1) gegeben ist, d.h. $\zeta_1 = X_1 - Y_1$, so hängt

$$\mathrm{E}_{(\vartheta_1, \vartheta_2, \vartheta_2')}(X_1 - Y_1) = \mathrm{E}_{(\vartheta_1, \vartheta_2, \vartheta_2')}(X_1) - \mathrm{E}_{(\vartheta_1, \vartheta_2, \vartheta_2')}(Y_1), \; (\vartheta_1, \vartheta_2, \vartheta_2') \in \Theta,$$

nicht mehr von ϑ_2' ab. Gleiches gilt auch im Fall der Invertierbarkeit von w mit Inverser $w^{-1} : (\mathbb{R}^l, \mathfrak{B}^l) \to (\mathbb{R}^m, \mathfrak{B}^m)$. Berücksichtige dabei die Abbildung $p : (\mathbb{R}^l, \mathfrak{B}^l) \to (\mathbb{R}^{m/2}, \mathfrak{B}^{m/2})$ gegeben durch

$$p(x) := (\mathrm{I}_{m/2 \times m/2}, 0_{m/2 \times m/2}) w^{-1}(x) - (0_{m/2 \times m/2}, \mathrm{I}_{m/2 \times m/2}) w^{-1}(x), \tag{6.1.4}$$

d.h. $p(\zeta_1) = p(w(\begin{smallmatrix} X_1 \\ Y_1 \end{smallmatrix})) = X_1 - Y_1$. Daher kann in diesen Fällen auch $\mathrm{E}_\vartheta(X_1 - Y_1)$, $\vartheta \in \tilde{\Theta}$, geschrieben werden.

Es sei für alle $1 \leq i \leq k$

$$\frac{\partial}{\partial \vartheta_i} \int f(z;\vartheta) \mathrm{d}\mu(z) = \int \frac{\partial}{\partial \vartheta_i} f(z;\vartheta) \mathrm{d}\mu(z) = 0, \ \vartheta \in \tilde{\Theta},$$

sowie

$$\frac{\partial}{\partial \vartheta_i} \int z f(z;\vartheta) \mathrm{d}\mu(z) = \int z \frac{\partial}{\partial \vartheta_i} f(z;\vartheta) \mathrm{d}\mu(z), \ \vartheta \in \tilde{\Theta},$$

erfüllt. Falls m eine gerade Zahl ist und $X_i : (\Omega, \mathfrak{A}) \to (\mathbb{R}^{m/2}, \mathfrak{B}^{m/2})$, $Y_i : (\Omega, \mathfrak{A}) \to (\mathbb{R}^{m/2}, \mathfrak{B}^{m/2})$, $i = 1, \dots, n$, und w invertierbar ist, so gelte ebenfalls für alle $1 \leq i \leq k$

$$\frac{\partial}{\partial \vartheta_i} \int p(z) f(z;\vartheta) \mathrm{d}\mu(z) = \int p(z) \frac{\partial}{\partial \vartheta_i} f(z;\vartheta) \mathrm{d}\mu(z), \ \vartheta \in \tilde{\Theta}.$$

Es soll die übliche Schreibweise für die Definitheit von Matrizen benutzt werden, d.h. für zwei Matrizen $A, B \in \mathbb{R}^{h \times h}$, $h \in \mathbb{N}$, bedeute $A > B$, dass die Matrix $A - B$ positiv definit sei. Entsprechendes gelte für positive Semidefinitheit ($A \geq B$), für negative Definitheit ($A < B$) und für negative Semidefinitheit ($A \leq B$).

Weiter benutzen wir für die Kreuzkovarianzmatrix zweier $\mathbb{R}^h$-wertiger Zufallsvektoren U und W, $h \in \mathbb{N}$, die Schreibweise

$$\mathrm{Cov}(U, W) := \mathrm{E}\left((U - \mathrm{E}(U))(W - \mathrm{E}(W))^t \right)$$

und im Fall $U = W$ für die Kovarianzmatrix die Schreibweise

$$\mathrm{Cov}(U) := \mathrm{Cov}(U, U),$$

immer unter der Voraussetzung, dass diese existieren und endliche Einträge besitzen. Im Fall $h = 1$ ist natürlich $\mathrm{Cov}(U, W) = \mathrm{E}((U - \mathrm{E}(U))(W - \mathrm{E}(W)))$ die bekannte Kovarianz zwischen U und W, und $\mathrm{Cov}(U) = \mathrm{Var}(U)$ die Varianz von U. Ferner sei im Fall $h = 1$

$$\rho(U, W) := \frac{\mathrm{Cov}(U, W)}{\sqrt{\mathrm{Var}(U)\mathrm{Var}(W)}}$$

die Korrelation von U und W, falls $\mathrm{Var}(U) > 0$ und $\mathrm{Var}(W) > 0$.

Es existiere für alle $\vartheta \in \tilde{\Theta}$ $\mathrm{Cov}_\vartheta(\zeta_1)$ und besitze endliche Einträge. Falls m eine gerade Zahl ist und $X_i : (\Omega, \mathfrak{A}) \to (\mathbb{R}^{m/2}, \mathfrak{B}^{m/2})$, $Y_i : (\Omega, \mathfrak{A}) \to (\mathbb{R}^{m/2}, \mathfrak{B}^{m/2})$, $i = 1, \dots, n$, verlangen wir zusätzlich, dass für alle $(\vartheta_1, \vartheta_2, \vartheta_2') \in \Theta$ $\mathrm{Cov}_{(\vartheta_1, \vartheta_2, \vartheta_2')}(X_1)$, $\mathrm{Cov}_{(\vartheta_1, \vartheta_2, \vartheta_2')}(Y_1)$, $\mathrm{Cov}_{(\vartheta_1, \vartheta_2, \vartheta_2')}(X_1 - Y_1)$ und $\mathrm{Cov}_{(\vartheta_1, \vartheta_2, \vartheta_2')}(X_1, Y_1)$ existieren und endliche Einträge besitzen. Falls w gemäß (6.1.1) gegeben ist, d.h. $\zeta_1 = X_1 - Y_1$, so hängt

$$\begin{aligned} \mathrm{Cov}_{(\vartheta_1, \vartheta_2, \vartheta_2')}(X_1 - Y_1) = &\mathrm{Cov}_{(\vartheta_1, \vartheta_2, \vartheta_2')}(X_1) + \mathrm{Cov}_{(\vartheta_1, \vartheta_2, \vartheta_2')}(Y_1) \\ &- \mathrm{Cov}_{(\vartheta_1, \vartheta_2, \vartheta_2')}(X_1, Y_1) - \mathrm{Cov}_{(\vartheta_1, \vartheta_2, \vartheta_2')}(X_1, Y_1)^t, \ (\vartheta_1, \vartheta_2, \vartheta_2') \in \Theta, \end{aligned}$$

nicht mehr von ϑ_2' ab. Gleiches gilt auch im Fall der Invertierbarkeit von w. Berücksichtige dabei die Abbildung p gegeben durch (6.1.4), d.h. $p(\zeta_1) = p(w(\binom{X_1}{Y_1})) = X_1 - Y_1$. Daher kann in diesen Fällen auch $\mathrm{Cov}_\vartheta(X_1 - Y_1)$, $\vartheta \in \tilde{\Theta}$, geschrieben werden.

Für $h \in \mathbb{N}$ sei eine Abbildung[6] $g : R \to \mathbb{R}^h$ gegeben, $g = (g_1, \ldots, g_h)^t$, für die für alle $i = 1, \ldots, h$ und alle $\vartheta \in \tilde{\Theta}$ $\mathrm{E}_\vartheta(g(\zeta_1))$ existiert und endliche Einträge besitzt und für die die Abbildung $\vartheta \mapsto \mathrm{E}_\vartheta(g(\zeta_1))$ für alle $\vartheta \in \tilde{\Theta}$ differenzierbar ist. Dann definieren wir

$$\psi_{g(\zeta_1)}(\vartheta) := \Big(\frac{\partial}{\partial \vartheta_j} \mathrm{E}_\vartheta\left(g_i(\zeta_1)\right)\Big)_{\substack{1\le i\le h\\ 1\le j\le k}}, \ \vartheta \in \tilde{\Theta}. \tag{6.1.5}$$

Mit $(\vartheta_1, \vartheta_2) \in \tilde{\Theta}$, $\vartheta_1 = (\vartheta_{1,1}, \ldots, \vartheta_{1,d_1})^t \in \Theta^1$, definieren wir

$$\psi^1_{g(\zeta_1)}(\vartheta) := \Big(\frac{\partial}{\partial \vartheta_{1,j}} \mathrm{E}_\vartheta\left(g_i(\zeta_1)\right)\Big)_{\substack{1\le i\le h\\ 1\le j\le d_1}}, \ \vartheta \in \tilde{\Theta}.$$

D.h. unter den genannten Bedingungen gilt

$$\psi_{g(\zeta_1)}(\vartheta) = \big(\psi^1_{g(\zeta_1)}(\vartheta), \psi^2_{g(\zeta_1)}(\vartheta)\big),$$

mit einer Matrix $\psi^2_{g(\zeta_1)}(\vartheta) \in \mathbb{R}^{h\times(k-d_1)}$, $\vartheta \in \Theta$. Zudem setzen wir im Fall der Invertierbarkeit der Matrix $\psi^1_{g(\zeta_1)}(\vartheta)^t \psi^1_{g(\zeta_1)}(\vartheta)$ für alle $\vartheta \in \tilde{\Theta}$,

$$\varphi_{g(\zeta_1)}(\vartheta) := \big(\psi^1_{g(\zeta_1)}(\vartheta)^t \psi^1_{g(\zeta_1)}(\vartheta)\big)^{-1} \psi^1_{g(\zeta_1)}(\vartheta)^t, \ \vartheta \in \tilde{\Theta}.$$

Weiter setzen wir unter der Voraussetzung, dass die Matrix $\psi_{g(\zeta_1)}(\vartheta)^t \psi_{g(\zeta_1)}(\vartheta)$ für alle $\vartheta \in \tilde{\Theta}$ invertierbar ist,

$$\phi_{g(\zeta_1)}(\vartheta) := (\mathrm{I}_{d_1\times d_1}, 0_{d_1\times(k-d_1)})\big(\psi_{g(\zeta_1)}(\vartheta)^t \psi_{g(\zeta_1)}(\vartheta)\big)^{-1} \psi_{g(\zeta_1)}(\vartheta)^t, \ \vartheta \in \tilde{\Theta}.$$

Motiviert durch die Resultate in Abschnitt 4 und unter Beachtung der Bemerkung 4.2.8, betrachten wir mit $\vartheta^0 := (\vartheta_0^1, \vartheta_2, \vartheta_2')$, $\vartheta_0 := (\vartheta_0^1, \vartheta_2)$ und $\vartheta_0' := (\vartheta_0^1, \vartheta_2')$ Effizienzen der Form

$$\mathrm{eff}_{T',T}(\vartheta^0, \vartheta_1) = \frac{\vartheta_1^t A^{1,1}_{\zeta_1'}(\vartheta_0')\vartheta_1}{\vartheta_1^t A^{1,1}_{\zeta_1}(\vartheta_0)\vartheta_1},$$

oder

$$\tilde{\mathrm{eff}}_{T',T}(\vartheta^0) = \frac{\det A_{\zeta_1'}(\vartheta_0')}{\det A_{\zeta_1}(\vartheta_0)},$$

[6] In den folgenden Ausführungen kann g die Identität, d.h. $g(\zeta_1) = \zeta_1$, falls m eine gerade Zahl ist und $X_i : (\Omega, \mathfrak{A}) \to (\mathbb{R}^{m/2}, \mathfrak{B}^{m/2})$, $Y_i : (\Omega, \mathfrak{A}) \to (\mathbb{R}^{m/2}, \mathfrak{B}^{m/2})$, $i = 1, \ldots, n$, und w invertierbar ist, die Abbildung p definiert in (6.1.4), d.h. $g(\zeta_1) = p(\zeta_1) = p(w(\binom{X_1}{Y_1})) = X_1 - Y_1$, oder im Zusammenhang mit Exponentialfamilien eine entprechende suffiziente Statistik t sein, d.h. $g(\zeta_1) = t(\zeta_1)$.

mit[7] $\vartheta_1 \in \mathbb{R}^{d_1} \setminus \{0\}$, wobei

$$A^{1,1}_{\zeta_1}(\vartheta_0) = i^{1,1}_{\zeta_1}(\vartheta_0) - i^{1,2}_{\zeta_1}(\vartheta_0) i^{2,2}_{\zeta_1}(\vartheta_0)^{-1} i^{2,1}_{\zeta_1}(\vartheta_0) \text{ oder } A^{1,1}_{\zeta_1}(\vartheta_0) = i^{1,1}_{\zeta_1}(\vartheta_0),$$

sowie

$$A^{1,1}_{\zeta_1'}(\vartheta_0') = (i^{1,1}_{\zeta_1'}(\vartheta_0') - i^{1,2}_{\zeta_1'}(\vartheta_0') i^{2,2}_{\zeta_1'}(\vartheta_0')^{-1} i^{2,1}_{\zeta_1'}(\vartheta_0') \text{ oder } A^{1,1}_{\zeta_1'}(\vartheta_0) = i^{1,1}_{\zeta_1'}(\vartheta_0')$$

gilt. Dabei ist zu bemerken, dass die Matrix $A^{1,1}_{\zeta_1'}(\vartheta_0') \in \mathbb{R}^{d_1 \times d_1}$ nur von der Verteilung von ζ_1' abhängt, also nur von den Randverteilungen $Q^{X_1}_{\vartheta_0}$ und $Q^{Y_1}_{\vartheta_0}$ der verbundenen Stichprobe, und darüber hinaus nicht von weiteren Abhängigkeitsstrukturen zwischen X_1 und Y_1.

Wie wir sahen ergeben sich solche Effizienzen, wenn in beiden Experimenten ein optimales Verfahren in einem bestimmten Sinn gewählt wird, und gewisse Voraussetzungen erfüllt sind. Wie bereits in der Einleitung zu dieser Arbeit begründet, wird ein Anwender in der Regel in beiden Situationen ein optimales Verfahren wählen, wenn ein solches existiert. Daher ist es sinnvoll, solche Effizienzen weiter zu studieren.

6.2 Die Rolle der Kovarianz

In der Tat spielt in vielen Fällen die Kovarianz zwischen X_1 und Y_1 (oder im univariaten Fall äquivalent formuliert: die Korrelation zwischen X_1 und Y_1) eine wichtige Rolle, wenn es darum geht, einen Vergleich von verbundener und unabhängiger Stichprobenerhebung anhand der Effizienz der zugehörigen Tests durchzuführen.

Im univariaten Fall kann dieser Sachverhalt in vielen Fällen salopp so formuliert werden, dass die verbundene Stichprobenerhebung der unabhängigen vorzuziehen ist, wenn die Korrelation zwischen X_1 und Y_1 nur groß genug ist. Vergleiche hierzu mit Beispiel 1.1.1 zu Beginn dieser Arbeit.

Wie wir im Laufe dieses Abschnittes feststellen werden, trifft dies nicht alleine im Fall der multivariaten Normalverteilung zu, wo bekanntlich die Abhängigkeitsstruktur vollständig durch die Kovarianzmatrix beschrieben wird.

Mit Hilfe der Informationsungleichung, siehe Satz 2.133 in [41], erhält man Kriterien dafür, dass die verbundene Stichprobenerhebung anhand der betrachteten Effizienz der unabhängigen Stichprobenerhebung vorzuziehen ist.

6.2.1 Satz: *Es gelte $A^{1,1}_{\zeta_1}(\vartheta_0) = i^{1,1}_{\zeta_1}(\vartheta_0)$. Dann gelten die folgenden Implikationen:*

a) Es sei $\psi^1_{\zeta_1}(\vartheta_0)^t \psi^1_{\zeta_1}(\vartheta_0)$ invertierbar. Dann gilt die Implikation

$$A^{1,1}_{\zeta_1'}(\vartheta_0')^{-1} \geq \varphi_{\zeta_1}(\vartheta_0) \mathrm{Cov}_{\vartheta_0}(\zeta_1) \varphi_{\zeta_1}(\vartheta_0)^t \;\Rightarrow\; \sup_{\vartheta_1 \in \mathbb{R}^{d_1} \setminus \{0\}} \mathrm{eff}_{T',T}(\vartheta^0, \vartheta_1) \leq 1.$$

[7] Beachte dabei Bemerkung 4.5.3.

b) Es sei m eine gerade Zahl und $X_1 : (\Omega, \mathfrak{A}) \to (\mathbb{R}^{m/2}, \mathfrak{B}^{m/2})$, $Y_1 : (\Omega, \mathfrak{A}) \to (\mathbb{R}^{m/2}, \mathfrak{B}^{m/2})$. Weiter sei $\zeta_1 = w((\begin{smallmatrix} X_1 \\ Y_1 \end{smallmatrix}))$ und w invertierbar, oder $\zeta_1 = X_1 - Y_1$. Zudem sei $\psi^1_{X_1-Y_1}(\vartheta_0)^t \psi^1_{X_1-Y_1}(\vartheta_0)$ invertierbar. Dann gilt die Implikation

$$\begin{aligned} &A^{1,1}_{\zeta_1'}(\vartheta_0')^{-1} \geq \varphi_{X_1-Y_1}(\vartheta_0)\big(\mathrm{Cov}_{\vartheta^0}(X_1) + \mathrm{Cov}_{\vartheta^0}(Y_1) \\ &- \mathrm{Cov}_{\vartheta^0}(X_1, Y_1) - \mathrm{Cov}_{\vartheta^0}(X_1, Y_1)^t\big)\varphi_{X_1-Y_1}(\vartheta_0)^t \;\Rightarrow\; \sup_{\vartheta_1 \in \mathbb{R}^{d_1} \setminus \{0\}} \mathrm{eff}_{T',T}(\vartheta^0, \vartheta_1) \leq 1. \end{aligned}$$

c) Es sei $X_1 : (\Omega, \mathfrak{A}) \to (\mathbb{R}, \mathfrak{B})$, $Y_1 : (\Omega, \mathfrak{A}) \to (\mathbb{R}, \mathfrak{B})$. Weiter sei $\zeta_1 = w((\begin{smallmatrix} X_1 \\ Y_1 \end{smallmatrix}))$ und w invertierbar, oder $\zeta_1 = X_1 - Y_1$. Zudem sei $\psi^1_{X_1-Y_1}(\vartheta_0)^t \psi^1_{X_1-Y_1}(\vartheta_0)$ invertierbar. Dann gilt

$$\begin{aligned} \mathrm{eff}_{T',T}(\vartheta^0) = \tilde{\mathrm{eff}}_{T',T}(\vartheta^0) \leq &\frac{1}{\varphi_{X_1-Y_1}(\vartheta_0)\varphi_{X_1-Y_1}(\vartheta_0)^t} A^{1,1}_{\zeta_1'}(\vartheta_0')\big(\mathrm{Var}_{\vartheta^0}(X_1) + \mathrm{Var}_{\vartheta^0}(Y_1) \\ &- 2\sqrt{\mathrm{Var}_{\vartheta^0}(X_1)\mathrm{Var}_{\vartheta^0}(Y_1)}\rho_{\vartheta^0}(X_1, Y_1)\big). \end{aligned}$$

d) Es sei $\psi^1_{\zeta_1}(\vartheta_0)^t \psi^1_{\zeta_1}(\vartheta_0)$ invertierbar. Dann ist

$$\tilde{\mathrm{eff}}_{T',T}(\vartheta^0) \leq \det A^{1,1}_{\zeta_1'}(\vartheta_0) \det\big(\varphi_{\zeta_1}(\vartheta_0)\mathrm{Cov}_{\vartheta_0}(\zeta_1)\varphi_{\zeta_1}(\vartheta_0)^t\big).$$

e) Es sei m eine gerade Zahl und $X_1 : (\Omega, \mathfrak{A}) \to (\mathbb{R}^{m/2}, \mathfrak{B}^{m/2})$, $Y_1 : (\Omega, \mathfrak{A}) \to (\mathbb{R}^{m/2}, \mathfrak{B}^{m/2})$. Weiter sei $\zeta_1 = w((\begin{smallmatrix} X_1 \\ Y_1 \end{smallmatrix}))$ und w invertierbar, oder $\zeta_1 = X_1 - Y_1$. Zudem sei $\psi^1_{X_1-Y_1}(\vartheta_0)^t \psi^1_{X_1-Y_1}(\vartheta_0)$ invertierbar. Dann gilt

$$\begin{aligned} \tilde{\mathrm{eff}}_{T',T}(\vartheta^0) \leq \det A^{1,1}_{\zeta_1'}(\vartheta_0) \det\Big(&\varphi_{X_1-Y_1}(\vartheta_0)\big(\mathrm{Cov}_{\vartheta^0}(X_1) + \mathrm{Cov}_{\vartheta^0}(Y_1) \\ &- \mathrm{Cov}_{\vartheta^0}(X_1, Y_1) - \mathrm{Cov}_{\vartheta^0}(X_1, Y_1)^t\big)\varphi_{X_1-Y_1}(\vartheta_0)^t\Big). \end{aligned}$$

Beweis: Im Beweis von Teil a) liefert die Informationsungleichung

$$\mathrm{Cov}_{\vartheta_0}(\zeta_1) \geq \psi^1_{\zeta_1}(\vartheta_0) i^{1,1}_{\zeta_1}(\vartheta_0)^{-1} \psi^1_{\zeta_1}(\vartheta_0)^t,$$

äquivalent

$$\mathrm{Cov}_{\vartheta_0}(\zeta_1) - \psi^1_{\zeta_1}(\vartheta_0) i^{1,1}_{\zeta_1}(\vartheta_0)^{-1} \psi^1_{\zeta_1}(\vartheta_0)^t \geq 0.$$

Daraus folgt

$$\begin{aligned} &\big(\psi^1_{\zeta_1}(\vartheta_0)^t \psi^1_{\zeta_1}(\vartheta_0)\big)^{-1} \psi^1_{\zeta_1}(\vartheta_0)^t \big(\mathrm{Cov}_{\vartheta_0}(\zeta_1) \\ &- \psi^1_{\zeta_1}(\vartheta_0) i^{1,1}_{\zeta_1}(\vartheta_0)^{-1} \psi^1_{\zeta_1}(\vartheta_0)^t\big) \psi^1_{\zeta_1}(\vartheta_0) \big(\psi^1_{\zeta_1}(\vartheta_0)^t \psi^1_{\zeta_1}(\vartheta_0)\big)^{-1} \geq 0, \end{aligned}$$

was äquivalent ist zu

$$\varphi_{\zeta_1}(\vartheta_0)\mathrm{Cov}_{\vartheta_0}(\zeta_1)\varphi_{\zeta_1}(\vartheta_0)^t - i^{1,1}_{\zeta_1}(\vartheta_0)^{-1} \geq 0.$$

Wegen Lemma 4.5.2 ist es ausreichend,

$$A^{1,1}_{\zeta_1'}(\vartheta_0') - i^{1,1}_{\zeta_1}(\vartheta_0) \leq 0$$

zu zeigen, was wegen Lemma A.1.3 des Anhangs äquivalent ist zu

$$A^{1,1}_{\zeta_1'}(\vartheta_0')^{-1} - i^{1,1}_{\zeta_1}(\vartheta_0)^{-1} \geq 0.$$

Es ist

$$\begin{aligned}&A^{1,1}_{\zeta_1'}(\vartheta_0')^{-1} - i^{1,1}_{\zeta_1}(\vartheta_0)^{-1}\\ =&\underbrace{A^{1,1}_{\zeta_1'}(\vartheta_0')^{-1} - \varphi_{\zeta_1}(\vartheta_0)\mathrm{Cov}_{\vartheta_0}(\zeta_1)\varphi_{\zeta_1}(\vartheta_0)^t}_{\geq 0} + \underbrace{\varphi_{\zeta_1}(\vartheta_0)\mathrm{Cov}_{\vartheta_0}(\zeta_1)\varphi_{\zeta_1}(\vartheta_0)^t - i^{1,1}_{\zeta_1}(\vartheta_0)^{-1}}_{\geq 0} \geq 0.\end{aligned}$$

Somit folgt die Aussage in Teil a).

Für Teil b) folgt im Fall $\zeta_1 = X_1 - Y_1$ die Aussage aus Teil a). Im Fall $\zeta_1 = w(\left(\begin{smallmatrix} X_1 \\ Y_1 \end{smallmatrix}\right))$ und w invertierbar folgt mit Hilfe der Informationsungleichung unter Heranziehung der Abbildung p definiert in (6.1.4)

$$\mathrm{Cov}_{\vartheta_0}(p(\zeta_1)) \geq \psi^1_{p(\zeta_1)}(\vartheta_0) i^{1,1}_{\zeta_1}(\vartheta_0)^{-1} \psi^1_{p(\zeta_1)}(\vartheta_0)^t,$$

und wegen $p(\zeta_1) = p(w(\left(\begin{smallmatrix} X_1 \\ Y_1 \end{smallmatrix}\right))) = X_1 - Y_1$ folgt der Rest von Teil b) analog zu Teil a).

Teil c) folgt analog zu Teil b) mit der Informationsungleichung.

Für Teil d) folgt wie oben zunächst

$$\varphi_{\zeta_1}(\vartheta_0)\mathrm{Cov}_{\vartheta_0}(\zeta_1)\varphi_{\zeta_1}(\vartheta_0)^t - i^{1,1}_{\zeta_1}(\vartheta_0)^{-1} \geq 0.$$

Dies impliziert

$$\det i^{1,1}_{\zeta_1}(\vartheta_0) \det\left(\varphi_{\zeta_1}(\vartheta_0)\mathrm{Cov}_{\vartheta_0}(\zeta_1)\varphi_{\zeta_1}(\vartheta_0)^t\right) \geq 1,$$

und damit

$$\tilde{\mathrm{eff}}_{T',T}(\vartheta^0) = \frac{\det A^{1,1}_{\zeta_1'}(\vartheta_0')}{\det i^{1,1}_{\zeta_1}(\vartheta_0)} \leq \det A^{1,1}_{\zeta_1'} \det\left(\varphi_{\zeta_1}(\vartheta_0)\mathrm{Cov}_{\vartheta_0}(\zeta_1)\varphi_{\zeta_1}(\vartheta_0)^t\right).$$

Teil e) folgt analog zu Teil b) und d) mit der Informationsungleichung. □

6.2.2 Satz: *Es sei $A^{1,1}_{\zeta_1}(\vartheta_0) = i^{1,1}_{\zeta_1}(\vartheta_0) - i^{1,2}_{\zeta_1}(\vartheta_0) i^{2,2}_{\zeta_1}(\vartheta_0)^{-1} i^{2,1}_{\zeta_1}(\vartheta_0)$. Dann gelten die folgenden Implikationen:*

a) Es sei $\psi_{\zeta_1}(\vartheta_0)^t \psi_{\zeta_1}(\vartheta_0)$ invertierbar. Dann gilt die Implikation

$$A^{1,1}_{\zeta_1'}(\vartheta_0')^{-1} \geq \phi_{\zeta_1}(\vartheta_0)\mathrm{Cov}_{\vartheta_0}(\zeta_1)\phi_{\zeta_1}(\vartheta_0)^t \;\Rightarrow\; \sup_{\vartheta_1 \in \mathbb{R}^{d_1} \setminus \{0\}} \mathrm{eff}_{T',T}(\vartheta^0, \vartheta_1) \leq 1.$$

b) Es sei m eine gerade Zahl und $X_1 : (\Omega, \mathfrak{A}) \to (\mathbb{R}^{m/2}, \mathfrak{B}^{m/2})$, $Y_1 : (\Omega, \mathfrak{A}) \to (\mathbb{R}^{m/2}, \mathfrak{B}^{m/2})$. Weiter sei $\zeta_1 = w(\binom{X_1}{Y_1})$ und w invertierbar, oder $\zeta_1 = X_1 - Y_1$. Zudem sei $\psi_{X_1-Y_1}(\vartheta_0)^t \psi_{X_1-Y_1}(\vartheta_0)$ invertierbar. Dann gilt die Implikation

$$\begin{aligned}&A^{1,1}_{\zeta_1'}(\vartheta_0')^{-1} \geq \phi_{X_1-Y_1}(\vartheta_0)\big(\mathrm{Cov}_{\vartheta^0}(X_1) + \mathrm{Cov}_{\vartheta^0}(Y_1)\\ &- \mathrm{Cov}_{\vartheta^0}(X_1, Y_1) - \mathrm{Cov}_{\vartheta^0}(X_1, Y_1)^t\big)\phi_{X_1-Y_1}(\vartheta_0)^t \ \Rightarrow \sup_{\vartheta_1 \in \mathbb{R}^{d_1}\setminus\{0\}} \mathrm{eff}_{T',T}(\vartheta^0, \vartheta_1) \leq 1.\end{aligned}$$

c) Es sei $X_1 : (\Omega, \mathfrak{A}) \to (\mathbb{R}, \mathfrak{B})$, $Y_1 : (\Omega, \mathfrak{A}) \to (\mathbb{R}, \mathfrak{B})$. Weiter sei $\zeta_1 = w(\binom{X_1}{Y_1})$ und w invertierbar, oder $\zeta_1 = X_1 - Y_1$. Zudem sei $\psi_{X_1-Y_1}(\vartheta_0)^t \psi_{X_1-Y_1}(\vartheta_0)$ invertierbar. Dann gilt

$$\begin{aligned}&\mathrm{eff}_{T',T}(\vartheta^0) = \tilde{\mathrm{eff}}_{T',T}(\vartheta^0)\\ \leq &\frac{1}{\phi_{X_1-Y_1}(\vartheta_0)\phi_{X_1-Y_1}(\vartheta_0)^t} A^{1,1}_{\zeta_1'}(\vartheta_0')\big(\mathrm{Var}_{\vartheta^0}(X_1) + \mathrm{Var}_{\vartheta^0}(Y_1)\\ &- 2\sqrt{\mathrm{Var}_{\vartheta^0}(X_1)\mathrm{Var}_{\vartheta^0}(Y_1)}\rho_{\vartheta^0}(X_1, Y_1)\big).\end{aligned}$$

d) Es sei $\psi_{\zeta_1}(\vartheta_0)^t \psi_{\zeta_1}(\vartheta_0)$ invertierbar. Dann ist

$$\tilde{\mathrm{eff}}_{T',T}(\vartheta^0) \leq \det A^{1,1}_{\zeta_1'}(\vartheta_0) \det\big(\phi_{\zeta_1}(\vartheta_0)\mathrm{Cov}_{\vartheta_0}(\zeta_1)\phi_{\zeta_1}(\vartheta_0)^t\big).$$

e) Es sei m eine gerade Zahl und $X_1 : (\Omega, \mathfrak{A}) \to (\mathbb{R}^{m/2}, \mathfrak{B}^{m/2})$, $Y_1 : (\Omega, \mathfrak{A}) \to (\mathbb{R}^{m/2}, \mathfrak{B}^{m/2})$. Weiter sei $\zeta_1 = w(\binom{X_1}{Y_1})$ und w invertierbar, oder $\zeta_1 = X_1 - Y_1$. Zudem sei $\psi_{X_1-Y_1}(\vartheta_0)^t \psi_{X_1-Y_1}(\vartheta_0)$ invertierbar. Dann gilt

$$\begin{aligned}\tilde{\mathrm{eff}}_{T',T}(\vartheta^0) \leq \det A^{1,1}_{\zeta_1'}(\vartheta_0) \det\Big(&\phi_{X_1-Y_1}(\vartheta_0)\big(\mathrm{Cov}_{\vartheta^0}(X_1) + \mathrm{Cov}_{\vartheta^0}(Y_1)\\ &- \mathrm{Cov}_{\vartheta^0}(X_1, Y_1) - \mathrm{Cov}_{\vartheta^0}(X_1, Y_1)^t\big)\phi_{X_1-Y_1}(\vartheta_0)^t\Big).\end{aligned}$$

Beweis: Im Beweis von Teil a) liefert die Informationsungleichung

$$\mathrm{Cov}_{\vartheta_0}(\zeta_1) \geq \psi_{\zeta_1}(\vartheta_0) i_{\zeta_1}(\vartheta_0)^{-1} \psi_{\zeta_1}(\vartheta_0)^t,$$

äquivalent

$$\mathrm{Cov}_{\vartheta_0}(\zeta_1) - \psi_{\zeta_1}(\vartheta_0) i_{\zeta_1}(\vartheta_0)^{-1} \psi_{\zeta_1}(\vartheta_0)^t \geq 0.$$

Daraus folgt

$$\begin{aligned}&(\mathrm{I}_{d_1\times d_1}, 0_{d_1\times(k-d_1)})\big(\psi_{\zeta_1}(\vartheta_0)^t\psi_{\zeta_1}(\vartheta_0)\big)^{-1}\psi_{\zeta_1}(\vartheta_0)^t\big(\mathrm{Cov}_{\vartheta_0}(\zeta_1)\\ &- \psi_{\zeta_1}(\vartheta_0) i_{\zeta_1}(\vartheta_0)^{-1}\psi_{\zeta_1}(\vartheta_0)^t\big)\psi_{\zeta_1}(\vartheta_0)\big(\psi_{\zeta_1}(\vartheta_0)^t\psi_{\zeta_1}(\vartheta_0)\big)^{-1}(\mathrm{I}_{d_1\times d_1}, 0_{d_1\times(k-d_1)})^t \geq 0,\end{aligned}$$

was wegen

$$\big(i_{\zeta_1}(\vartheta_0)^{-1}\big)_{\substack{1\le i\le d_1\\1\le j\le d_1}} = \big(i^{1,1}_{\zeta_1}(\vartheta_0) - i^{1,2}_{\zeta_1}(\vartheta_0)i^{2,2}_{\zeta_1}(\vartheta_0)^{-1}i^{2,1}_{\zeta_1}(\vartheta_0)\big)^{-1} = A^{1,1}_{\zeta_1}(\vartheta_0)^{-1},$$

äquivalent ist zu

$$\phi_{\zeta_1}(\vartheta_0)\mathrm{Cov}_{\vartheta_0}(\zeta_1)\phi_{\zeta_1}(\vartheta_0)^t - A^{1,1}_{\zeta_1}(\vartheta_0)^{-1} \ge 0.$$

Wegen Lemma 4.5.1 ist es ausreichend,

$$A^{1,1}_{\zeta_1'}(\vartheta_0') - A^{1,1}_{\zeta_1}(\vartheta_0) \le 0$$

zu zeigen, was wegen Lemma A.1.3 des Anhangs äquivalent ist zu

$$A^{1,1}_{\zeta_1'}(\vartheta_0')^{-1} - A^{1,1}_{\zeta_1}(\vartheta_0)^{-1} \ge 0.$$

Es ist

$$\begin{aligned}
&A^{1,1}_{\zeta_1'}(\vartheta_0')^{-1} - A^{1,1}_{\zeta_1}(\vartheta_0)^{-1}\\
=\,&\underbrace{A^{1,1}_{\zeta_1'}(\vartheta_0')^{-1} - \phi_{\zeta_1}(\vartheta_0)\mathrm{Cov}_{\vartheta_0}(\zeta_1)\phi_{\zeta_1}(\vartheta_0)^t}_{\ge 0} + \underbrace{\phi_{\zeta_1}(\vartheta_0)\mathrm{Cov}_{\vartheta_0}(\zeta_1)\phi_{\zeta_1}(\vartheta_0)^t - A^{1,1}_{\zeta_1}(\vartheta_0)^{-1}}_{\ge 0} \ge 0.
\end{aligned}$$

Somit folgt die Aussage in Teil a).

Für Teil b) folgt im Fall $\zeta_1 = X_1 - Y_1$ die Aussage aus Teil a). Im Fall $\zeta_1 = w(\binom{X_1}{Y_1})$ und w invertierbar folgt mit Hilfe der Informationsungleichung unter Heranziehung der Abbildung p definiert in (6.1.4)

$$\mathrm{Cov}_{\vartheta_0}(p(\zeta_1)) \ge \psi_{p(\zeta_1)}(\vartheta_0)i_{\zeta_1}(\vartheta_0)^{-1}\psi_{p(\zeta_1)}(\vartheta_0)^t,$$

und wegen $p(\zeta_1) = p(w(\binom{X_1}{Y_1})) = X_1 - Y_1$ folgt der Rest von Teil b) analog zu Teil a).

Teil c) folgt analog zu Teil b) mit der Informationsungleichung.

Für Teil d) folgt wie oben zunächst

$$\phi_{\zeta_1}(\vartheta_0)\mathrm{Cov}_{\vartheta_0}(\zeta_1)\phi_{\zeta_1}(\vartheta_0)^t - A^{1,1}_{\zeta_1}(\vartheta_0)^{-1} \ge 0.$$

Hieraus folgt

$$\det A^{1,1}_{\zeta_1}(\vartheta_0)\det\big(\phi_{\zeta_1}(\vartheta_0)\mathrm{Cov}_{\vartheta_0}(\zeta_1)\phi_{\zeta_1}(\vartheta_0)^t\big) \ge 1,$$

und damit

$$\tilde{\mathrm{eff}}_{T',T}(\vartheta^0) = \frac{\det A^{1,1}_{\zeta_1'}(\vartheta_0')}{\det A^{1,1}_{\zeta_1}(\vartheta_0)} \le \det A^{1,1}_{\zeta_1'}\det\big(\phi_{\zeta_1}(\vartheta_0)\mathrm{Cov}_{\vartheta_0}(\zeta_1)\phi_{\zeta_1}(\vartheta_0)^t\big).$$

Teil e) folgt analog zu Teil b) und d) mit der Informationsungleichung. □

6.2.3 Bemerkung: Die in den Sätzen 6.2.1 und 6.2.2 erarbeiteten Kriterien hängen im Fall von $\zeta_1 = \binom{X_1}{Y_1}$ oder $\zeta_1 = X_1 - Y_1$ nur von den Randverteilungen $Q_{\vartheta_0}^{X_1}$ und $Q_{\vartheta_0}^{Y_1}$ der verbundenen Stichprobe und der Kovarianzmatrix von ζ_1 ab; weitere Abhängigkeitsstrukturen zwischen X_1 und Y_1 spielen dabei keine Rolle. Zudem sind die in den Sätzen 6.2.1 und 6.2.2 erarbeiteten Kriterien in dem Sinne scharf, dass in speziellen Situationen die Kriterien nicht nur hinreichend, sondern sogar notwendig sind. Siehe hierzu Bemerkung 6.3.6 und Bemerkung 6.3.15. Zudem sind die in den Sätzen 6.2.1 a), b) und c) und 6.2.2 a), b) und c) erarbeiteten Kriterien unter den genannten Voraussetzungen auch hinreichende Kriterien für $\widetilde{\mathrm{eff}}_{T',T}(\vartheta^0) \leq 1$. Beachte hierzu Lemma 4.5.4.

6.3 Exponentialfamilien

In diesem Unterabschnitt gelte die Bezeichnungen aus den Unterabschnitten 6.1 und 6.2. Zunächst wollen wir einen Fall betrachten, wie er zum Beispiel in Abschnitt 8 vorliegt. Es sei

$$A_{\zeta_1}^{1,1}(\vartheta_0) = i_{\zeta_1}^{1,1}(\vartheta_0).$$

Wir gehen weiter davon aus, dass eine injektive und stetig differenzierbare Abbildung $L : \tilde{\Theta} \to \mathbb{R}^k$ mit überall regulärer Funktionalmatrix sowie Abbildungen $L_1 : \mathbb{R}^k \to \mathbb{R}^{d_1}$ und $L_2 : \mathbb{R}^{k-d_1} \to \mathbb{R}^{k-d_1}$ existieren, sodass

$$L(\vartheta_1, \vartheta_2) = \begin{pmatrix} L_1(\vartheta_1, \vartheta_2) \\ L_2(\vartheta_2) \end{pmatrix} \text{ für alle } (\vartheta_1, \vartheta_2) \in \tilde{\Theta}$$

erfüllt ist, sowie

$$f\big(\cdot; (\vartheta_1, \vartheta_2)\big) = \exp\Big(L_1(\vartheta_1, \vartheta_2)^t T_{\vartheta_2}(\cdot) - C_{\vartheta_2}\big(L_1(\vartheta_1, \vartheta_2)\big)\Big) R_{\vartheta_2}(\cdot) \text{ für alle } (\vartheta_1, \vartheta_2) \in \tilde{\Theta} \tag{6.3.1}$$

ist, mit $C_{\vartheta_2} : \mathbb{R}^{d_1} \to \mathbb{R}$, $T_{\vartheta_2} : (\mathbb{R}^l, \mathfrak{B}^l) \to (\mathbb{R}^{d_1}, \mathfrak{B}^{d_1})$ und $R_{\vartheta_2} : (\mathbb{R}^l, \mathfrak{B}^l) \to (\mathbb{R}, \mathfrak{B})$ μ-f.ü. positiv.

Aus den Voraussetzungen folgt, dass die Menge $\Lambda := L(\tilde{\Theta})$ offen ist, siehe Satz 171.2 Teil a) in [13]. Wir definieren $S_{L(\vartheta)} := Q_{\vartheta}^{\zeta_1}$, $\vartheta \in \tilde{\Theta}$. Es ist die Familie von Verteilungen von ζ_1 durch die Elemente von Λ injektiv parametrisiert. Dies bedeutet, dass mit $\{S_\lambda; \lambda \in \Lambda\}$ eine Familie von Wahrscheinlichkeitsmaßen auf dem Messraum $(\mathbb{R}^l, \mathfrak{B}^l)$ mit $\{S_\lambda; \lambda \in \Lambda\} = \{Q_{\vartheta}^{\zeta_1}; \vartheta \in \tilde{\Theta}\}$ gegeben ist. Elemente $\lambda \in \Lambda$ sollen in der Folge gemäß $\lambda = (\lambda_1, \lambda_2) \in \mathbb{R}^{d_1} \times \mathbb{R}^{k-d_1}$ geschrieben werden. Allgemein ist zu bemerken, dass im Fall $d_1 = k$ die Komponente λ_2 verschwindet.

Wir halten im Folgenden ϑ_2 fest und setzen

$$\Theta_2^1 := \{\vartheta_1 \in \Theta^1; \exists \vartheta_2' : (\vartheta_1, \vartheta_2, \vartheta_2') \in \Theta\}.$$

und

$$\Lambda^1 := \big\{\lambda_1 \in \mathbb{R}^{d_1}; \exists \lambda \in \Lambda : \lambda = \big(\lambda_1, L_2(\vartheta_2)\big)\big\}$$

Diese Mengen sind offen. Wir definieren $\ell(\cdot) := L_1(\cdot, \vartheta_2)$, $t := T_{\vartheta_2}$, $c := C_{\vartheta_2}$ und $r := R_{\vartheta_2}$, also

$$f\big(\cdot; (\vartheta_1, \vartheta_2)\big) = \exp\big(\ell(\vartheta_1)^t t(\cdot) - c(\ell(\vartheta_1)\big) r(\cdot),\ \vartheta_1 \in \Theta_2^1.$$

Dann ist die Familie von Verteilungen von ζ_1 eine d_1-parametrische Exponentialfamilie im Parameter $\lambda_1 \in \Lambda^1$. Zudem definieren wir $q_{\vartheta_1}^{\zeta_1} := Q_{(\vartheta_1,\vartheta_2)}^{\zeta_1}$, $\vartheta_1 \in \Theta_2^1$, und $s_{\lambda_1} := S_{(\lambda_1, L_2(\vartheta_2))}$, $\lambda_1 \in \Lambda^1$.

Wir nehmen an, dass für alle $\gamma \in \mathbb{R}$ und alle $a \in \mathbb{R}^{d_1}$ $\mu(a^t t \neq \gamma, r > 0) > 0$ gilt. Zudem existiere kein $a \in \mathbb{R}^{d_1}$ und kein $b \in \mathbb{R}$, sodass $a^t \ell(\vartheta_1) = b$ für alle $\vartheta_1 \in \Theta_2^1$ gilt.

Die Abbildung c ist auf Λ^1 beliebig oft differenzierbar und die zugehörige Hessematrix von c ist auf Λ^1 positiv definit. Siehe Satz 1.164 a) und b) in [41]. Es sei mit $\lambda_1 = (\lambda_{1,1}, \dots, \lambda_{1,d_1})^t$ der Nabla-Ableitungsoperator ∇_{λ_1} durch

$$\nabla_{\lambda_1} := \Big(\frac{\partial}{\partial \lambda_{1,1}}, \dots, \frac{\partial}{\partial \lambda_{1,d_1}}\Big)^t$$

definiert. Wegen $\nabla_{\lambda_1} c(\lambda_1) = \mathrm{E}_{\lambda_1}(t(\zeta_1))$, $\lambda_1 \in \Lambda^1$, ist die Existenz und Endlichkeit von $\mathrm{E}_{\lambda_1}(t(\zeta_1))$, $\lambda_1 \in \Lambda^1$, sichergestellt. Es sei die Abbildung $e : \Lambda^1 \to E := \{\mathrm{E}_{\lambda_1}(t(\zeta_1)); \lambda_1 \in \Lambda^1\}$ gegeben durch

$$e(\lambda_1) := \mathrm{E}_{\lambda_1}\big(t(\zeta_1)\big),\ \lambda_1 \in \Lambda^1,$$

d.h. $\eta \in E$ ist ein Momentenparameter.

Wegen $\nabla_{\lambda_1} c(\lambda_1) = \mathrm{E}_{\lambda_1}(t(\zeta_1))$, $\lambda_1 \in \Lambda^1$, und da c auf Λ_1 beliebig oft differenzierbar ist, ist die Abbildung e auf Λ_1 stetig differenzierbar, wobei die zugehörige Jacobimatrix die Hessematrix von c ist, und daher auf Λ_1 positiv definit ist. Daher existiert nach dem Satz über die Umkehrabbildung, siehe Satz 171.1 in [13], die inverse Abbildung $e^{-1} : E \to \Lambda^1$ und diese ist ebenfalls auf E differenzierbar. Wir schreiben auch $\overline{\ell} := e^{-1}$.

Zudem sei die Abbildung $\overline{e} : \Theta_2^1 \to \tilde{E} := \{\mathrm{E}_{\vartheta_1}(t(\zeta_1)); \vartheta_1 \in \Theta_2^1\}$ durch

$$\overline{e}(\vartheta_1) := \mathrm{E}_{\vartheta_1}\big(t(\zeta_1)\big),\ \vartheta_1 \in \Theta_2^1,$$

gegeben. Es gilt $\tilde{E} = E$. Wegen $\overline{e} = e \circ \ell$ und der Differenzierbarkeit von e und ℓ in Λ^1 bzw. Θ_2^1 ist die Abbildung $\overline{e}$ auf Θ_2^1 differenzierbar.

Wir setzen $\eta_0 := \overline{e}(\vartheta_0^1)$ und damit $\lambda_0^1 := \overline{\ell}(\eta_0)$. Dann ist

$$\begin{aligned}\lambda_0^1 = \overline{\ell}(\eta_0) = e^{-1}(\eta_0) = e^{-1}\big(\overline{e}(\vartheta_0^1)\big) = e^{-1}\Big(\operatorname{E}_{\vartheta_0^1}\big(t(\zeta_1)\big)\Big)\\ = e^{-1}\Big(\operatorname{E}_{\ell(\vartheta_0^1)}\big(t(\zeta_1)\big)\Big) = e^{-1}\Big(e\big(\ell(\vartheta_0^1)\big)\Big) = \ell(\vartheta_0^1).\end{aligned}$$

In der Tat lassen sich im vorliegenden Fall durch den Übergang von der Parametrisierung durch Θ^1 in die Parametrisierung durch Λ^1 die zugehörigen Fisher-Informationsmatrizen leichter berechnen und so Ergebnisse beim Studium der Effizienz erzielen.

6.3.1 Satz: *Es sei $\vartheta_0 = (\vartheta_0^1, \vartheta_2)$. Zudem sei $A_{\zeta_1}^{1,1}(\vartheta_0) = i_{\zeta_1}^{1,1}(\vartheta_0)$. Es ist die Matrix*

$$\psi_{T_{\vartheta_2}(\zeta_1)}^1(\vartheta_0)^t \operatorname{Cov}_{\vartheta_0}(T_{\vartheta_2}(\zeta_1))^{-1}\psi_{T_{\vartheta_2}(\zeta_1)}^1(\vartheta_0)$$

positiv definit und

$$A_{\zeta_1}^{1,1}(\vartheta_0) = \psi_{T_{\vartheta_2}(\zeta_1)}^1(\vartheta_0)^t \operatorname{Cov}_{\vartheta_0}\big(T_{\vartheta_2}(\zeta_1)\big)^{-1}\psi_{T_{\vartheta_2}(\zeta_1)}^1(\vartheta_0).$$

Beweis: Es sei $\vartheta_1 = (\vartheta_{1,1}, \dots, \vartheta_{1,d_1})^t$ und $\eta = (\eta_1, \dots, \eta_{d_1})^t$. Wir definieren folgende zusätzliche Nabla-Ableitungsoperatoren:

$$\nabla_{\vartheta_1} := \Big(\frac{\partial}{\partial\vartheta_{1,1}}, \dots, \frac{\partial}{\partial\vartheta_{1,d_1}}\Big)^t,\ \nabla_\eta := \Big(\frac{\partial}{\partial\eta_1}, \dots, \frac{\partial}{\partial\eta_{d_1}}\Big)^t.$$

Zudem bezeichne $I_{\zeta_1}(\lambda_0^1)$ die Fisher-Informationsmatrix von ζ_1 bei Parametrisierung durch Λ_1 an der Stelle λ_0^1, und $\mathbb{I}_{\zeta_1}(\eta_0)$ die Fisher-Informationsmatrix von ζ_1 bei Parametrisierung durch E an der Stelle η_0. Es gilt

$$\nabla_{\lambda_1} c(\lambda_1)_{|_{\lambda_1=\lambda_0^1}} = \operatorname{E}_{\lambda_0^1}\big(T_{\vartheta_2}(\zeta_1)\big)$$

sowie

$$I_{\zeta_1}(\lambda_0^1) = \nabla_{\lambda_1}\nabla_{\lambda_1}^t c(\lambda_1)_{|_{\lambda_1=\lambda_0^1}} = \operatorname{Cov}_{\lambda_0^1}\big(T_{\vartheta_2}(\zeta_1)\big),$$

wobei die Existenz beider Ausdrücke sowie die positive Definitheit der Informationsmatrix $I_{\zeta_1}(\lambda_0^1)$ nach Annahme sichergestellt ist.

Es erfolgt die Transformation der Informationsmatrizen unter den verschiedenen Parametrisierungen durch

$$\mathbb{I}_{\zeta_1}(\eta_0) = \big(\nabla_\eta \overline{\ell}(\eta)^t\big)^t_{|_{\eta=\eta_0}} I_{\zeta_1}(\lambda_0^1)\nabla_\eta \overline{\ell}(\eta)^t_{|_{\eta=\eta_0}}$$

und mit $\vartheta_0 = (\vartheta_0^1, \vartheta_2)$

$$i_{\zeta_1}^{1,1}(\vartheta_0) = \big(\nabla_{\vartheta_1}\overline{e}(\vartheta_1)^t\big)^t_{|_{\vartheta_1=\vartheta_0^1}} \mathbb{I}_{\zeta_1}(\eta_0)\nabla_{\vartheta_1}\overline{e}(\vartheta_1)^t_{|_{\vartheta_1=\vartheta_0^1}}.$$

Dann ist

$$\nabla_{\lambda_1} e(\lambda_1)^t_{|\lambda_1=\lambda_0^1} = \nabla_{\lambda_1}\nabla^t_{\lambda_1} c(\lambda_1)_{|\lambda_1=\lambda_0^1} = I_{\zeta_1}(\lambda_0^1),$$

und deshalb wegen der Invertierbarkeit der Abbildung e

$$\nabla_\eta \overline{\ell}(\eta)^t_{|\eta=\eta_0} = \nabla_\eta e^{-1}(\eta)^t_{|\eta=\eta_0} = \big(\nabla_{\lambda_1} e(\lambda_1)^t_{|\lambda_1=\lambda_0^1}\big)^{-1} = I_{\zeta_1}(\lambda_0^1)^{-1} = \mathrm{Cov}_{\lambda_0^1}\big(T_{\vartheta_2}(\zeta_1)\big)^{-1}.$$

Damit folgt

$$\mathbb{I}_{\zeta_1}(\eta_0) = \big(\nabla_\eta \overline{\ell}(\eta)^t\big)^t_{|\eta=\eta_0} I_{\zeta_1}(\lambda_0^1)\nabla_\eta \overline{\ell}(\eta)^t_{|\eta=\eta_0} = \mathrm{Cov}_{\lambda_0^1}\big(T_{\vartheta_2}(\zeta_1)\big)^{-1},$$

und damit mit $\lambda_0^1 = \ell(\vartheta_0^1)$

$$\begin{aligned} i^{1,1}_{\zeta_1}(\vartheta_0) &= \big(\nabla_{\vartheta_1}\overline{e}(\vartheta_1)^t\big)^t_{|\vartheta_1=\vartheta_0^1} \mathbb{I}_{\zeta_1}(\eta_0)\nabla_{\vartheta_1}\overline{e}(\vartheta_1)^t_{|\vartheta_1=\vartheta_0^1} \\ &= \big(\nabla_{\vartheta_1}\overline{e}(\vartheta_1)^t\big)^t_{|\vartheta_1=\vartheta_0^1} \mathrm{Cov}_{\lambda_0^1}\big(T_{\vartheta_2}(\zeta_1)\big)^{-1}\nabla_{\vartheta_1}\overline{e}(\vartheta_1)^t_{|\vartheta_1=\vartheta_0^1} \\ &= \big(\nabla_{\vartheta_1}\overline{e}(\vartheta_1)^t\big)^t_{|\vartheta_1=\vartheta_0^1} \mathrm{Cov}_{\vartheta_0^1}\big(T_{\vartheta_2}(\zeta_1)\big)^{-1}\nabla_{\vartheta_1}\overline{e}(\vartheta_1)^t_{|\vartheta_1=\vartheta_0^1}. \end{aligned}$$

Wegen $\nabla_{\vartheta_1}\overline{e}(\vartheta_1)^t|_{\vartheta_1=\vartheta_0^1} = \psi^1_{T_{\vartheta_2}(\zeta_1)}(\vartheta_0)$ liefert dies die Behauptung. □

6.3.2 Korollar: *Es sei $\vartheta_0 = (\vartheta_0^1, \vartheta_2)$. Zudem sei $A^{1,1}_{\zeta_1}(\vartheta_0) = i^{1,1}_{\zeta_1}(\vartheta_0)$. Ist $d_1 = l$ und ist die Abbildung T_{ϑ_2} durch die Identität gegeben, so ist die Matrix*

$$\psi^1_{\zeta_1}(\vartheta_0)^t \mathrm{Cov}_{\vartheta_0}(\zeta_1)^{-1}\psi^1_{\zeta_1}(\vartheta_0)$$

positiv definit und man erhält

$$A^{1,1}_{\zeta_1}(\vartheta_0) = \psi^1_{\zeta_1}(\vartheta_0)^t \mathrm{Cov}_{\vartheta_0}(\zeta_1)^{-1}\psi^1_{\zeta_1}(\vartheta_0).$$

Falls m eine gerade Zahl ist, $X_1 : (\Omega, \mathfrak{A}) \to (\mathbb{R}^{m/2}, \mathfrak{B}^{m/2})$, $Y_1 : (\Omega, \mathfrak{A}) \to (\mathbb{R}^{m/2}, \mathfrak{B}^{m/2})$, $d_1 = m/2$, $\zeta_1 = X_1 - Y_1$ und die Abbildung T_{ϑ_2} durch die Identität gegeben ist, so ist die Matrix

$$\begin{aligned} &\psi^1_{\zeta_1}(\vartheta_0)^t \mathrm{Cov}_{\vartheta_0}(\zeta_1)^{-1}\psi^1_{\zeta_1}(\vartheta_0)\psi^1_{X_1-Y_1}(\vartheta_0)^t(\mathrm{Cov}_{\vartheta^0}(X_1) + \mathrm{Cov}_{\vartheta^0}(Y_1) - \mathrm{Cov}_{\vartheta^0}(X_1, Y_1) \\ &\quad - \mathrm{Cov}_{\vartheta^0}(X_1, Y_1)^t)^{-1}\psi^1_{X_1-Y_1}(\vartheta_0) \end{aligned}$$

positiv definit und man hat

$$\begin{aligned} &A^{1,1}_{\zeta_1}(\vartheta_0) \\ =&\psi^1_{X_1-Y_1}(\vartheta_0)^t\big(\mathrm{Cov}_{\vartheta^0}(X_1) + \mathrm{Cov}_{\vartheta^0}(Y_1) - \mathrm{Cov}_{\vartheta^0}(X_1, Y_1) - \mathrm{Cov}_{\vartheta^0}(X_1, Y_1)^t\big)^{-1}\psi^1_{X_1-Y_1}(\vartheta_0). \end{aligned}$$

Beweis: Falls die Abbildung T_{ϑ_2} durch die Identität gegeben ist, erhält man mit Satz 6.3.1

$$A^{1,1}_{\zeta_1}(\vartheta_0) = \psi^1_{\zeta_1}(\vartheta_0)^t \mathrm{Cov}_{\vartheta_0}(\zeta_1)^{-1}\psi^1_{\zeta_1}(\vartheta_0).$$

Falls m eine gerade Zahl ist, $X_1 : (\Omega, \mathfrak{A}) \to (\mathbb{R}^{m/2}, \mathfrak{B}^{m/2})$, $Y_1 : (\Omega, \mathfrak{A}) \to (\mathbb{R}^{m/2}, \mathfrak{B}^{m/2})$, $\zeta_1 = X_1 - Y_1$, $m/2 = d_1$ und die Abbildung T_{ϑ_2} durch die Identität gegeben ist, erhält man mit Satz 6.3.1

$$\begin{aligned}&A^{1,1}_{\zeta_1}(\vartheta_0)\\ =&\psi^1_{X_1-Y_1}(\vartheta_0)^t \big(\mathrm{Cov}_{\vartheta^0}(X_1) + \mathrm{Cov}_{\vartheta^0}(Y_1) - \mathrm{Cov}_{\vartheta^0}(X_1, Y_1) - \mathrm{Cov}_{\vartheta^0}(X_1, Y_1)^t\big)^{-1} \psi^1_{X_1-Y_1}(\vartheta_0).\end{aligned}$$

Damit ist alles gezeigt. $\square$

6.3.3 Beispiel: Es sei m eine gerade Zahl, $X_1 : (\Omega, \mathfrak{A}) \to (\mathbb{R}^{m/2}, \mathfrak{B}^{m/2})$ und $Y_1 : (\Omega, \mathfrak{A}) \to (\mathbb{R}^{m/2}, \mathfrak{B}^{m/2})$. Weiter liege das zu Beginn dieser Arbeit genannte Testproblem

$$\mathrm{H}:\ \mathcal{L}(X_1) = \mathcal{L}(Y_1),\ \mathrm{K}:\ \mathcal{L}(X_1) \neq \mathcal{L}(Y_1),$$

vor, und wir gehen davon aus, dass sich die Verteilungen von X_1 und Y_1 nur in ihrem Erwartungswert unterscheiden. Konkret sei $\zeta_1 = X_1 - Y_1$,

$$\vartheta_1 = \mathrm{E}_\vartheta(X_1 - Y_1),\ \vartheta = (\vartheta_1, \vartheta_2) \in \tilde{\Theta},$$

und das Testproblem lasse sich äquivalent schreiben als

$$\mathrm{H}:\ \vartheta_1 = 0,\ \mathrm{K}:\ \vartheta_1 \in \Theta^1 \setminus \{0\}.$$

In diesem Fall ist es oft sinnvoll, im verbundenen Stichprobenfall Tests basierend auf Beobachtungen der Form $X_1 - Y_1, \ldots, X_n - Y_n$ vorzuschlagen. Dann ist $\psi^1_{\zeta_1}(\vartheta_0)$ gegeben durch die Einheitsmatrix. Zudem sei $A^{1,1}_{\zeta_1}(\vartheta_0) = i^{1,1}_{\zeta_1}(\vartheta_0)$. Es sei die Familie von Verteilungen $\{Q^{X_1-Y_1}_\vartheta; \vartheta \in \Theta\}$ gemäß (6.3.1) für ϑ_2 mit $\vartheta = (\vartheta^1_0, \vartheta_2) \in \tilde{\Theta}_0$ eine Exponentialfamilie mit suffizienter Statistik T_{ϑ_2}. Zudem sei die Abbildung T_{ϑ_2} durch die Identität gegeben. Es ist

$$\mathrm{eff}_{T',T}(\vartheta^0, \vartheta_1) = \frac{\vartheta_1^t A^{1,1}_{\zeta_1'}(\vartheta_0')\vartheta_1}{\vartheta_1^t\big(2\mathrm{Cov}_{\vartheta^0}(X_1) - \mathrm{Cov}_{\vartheta^0}(X_1, Y_1) - \mathrm{Cov}_{\vartheta^0}(X_1, Y_1)^t\big)^{-1}\vartheta_1},$$

sowie

$$\widetilde{\mathrm{eff}}_{T',T}(\vartheta^0) = \det A^{1,1}_{\zeta_1'}(\vartheta_0') \det\big(2\mathrm{Cov}_{\vartheta^0}(X_1) - \mathrm{Cov}_{\vartheta^0}(X_1, Y_1) - \mathrm{Cov}_{\vartheta^0}(X_1, Y_1)^t\big),$$

speziell im eindimensionalen Fall

$$\mathrm{eff}_{T',T}(\vartheta^0) = \widetilde{\mathrm{eff}}_{T',T}(\vartheta^0) = 2A^{1,1}_{\zeta_1'}(\vartheta_0')\mathrm{Var}_{\vartheta^0}(X_1)\big(1 - \rho_{\vartheta^0}(X_1, Y_1)\big).$$

Konkret kann z.B. der Fall betrachtet werden, dass

$$\begin{pmatrix} X_1 \\ Y_1 \end{pmatrix} \sim \mathrm{N}_{2d_1}\left(\begin{pmatrix} a \\ b \end{pmatrix}, \begin{pmatrix} \Sigma & \Pi \\ \Pi^t & \Sigma \end{pmatrix}\right)$$

gilt, wobei $a, b \in \mathbb{R}^{d_1}$ und $\Sigma, \Pi \in \mathbb{R}^{d_1 \times d_1}$ unbekannt seien, Σ und die zugrunde liegende Kovarianzmatrix von (X_1, Y_1) positiv definit seien. Dann ist

$$X_1 - Y_1 \sim \mathrm{N}_{d_1}(\mu, \Gamma),$$

mit $\mu := a - b$ und $\Gamma := 2\Sigma - (\Pi + \Pi^t)$, und Γ ist positiv definit. Es ist also $\Theta^1 = \mathbb{R}^{d_1}$, $\vartheta_0^1 = 0 \in \mathbb{R}^{d_1}$, $\tilde{\Theta} = \{(a, \mathrm{vech}(A)); a \in \mathbb{R}^{d_1}, A \in \mathbb{R}^{d_1 \times d_1}$ symmetrisch positiv definit$\}$. Zudem ist $\vartheta_1 = \mu$ und $\vartheta_2 = \mathrm{vech}(\Gamma)$. Die Dichte von $X_1 - Y_1$ ist gegeben durch

$$\begin{aligned} f\big(x; (\vartheta_1, \vartheta_2)\big) &= \Big(\frac{1}{\sqrt{2\pi}}\Big)^{d_1} \frac{1}{\sqrt{\det \Gamma}} \exp\Big(-\frac{1}{2}(x-\mu)^t \Gamma^{-1}(x-\mu)\Big) \\ &= \Big(\frac{1}{\sqrt{2\pi}}\Big)^{d_1} \frac{1}{\sqrt{\det \Gamma}} \exp\Big(\mu^t \Gamma^{-1} x - \frac{1}{2}\mu^t \Gamma^{-1}\mu - \frac{1}{2} x^t \Gamma^{-1} x\Big), \; x \in \mathbb{R}^{d_1}. \end{aligned}$$

Wir definieren $L(\vartheta_1, \vartheta_2) := \binom{\mathrm{vech}^{-1}(\vartheta_2)^{-1}\vartheta_1}{\vartheta_2}$, $L_1(\vartheta_1, \vartheta_2) := (\mathrm{vech}^{-1}(\vartheta_2))^{-1}\vartheta_1$ und $L_2(\vartheta_2) := \vartheta_2$, $(\vartheta_1, \vartheta_2) \in \tilde{\Theta}$. Dann ist $\Lambda = \{\binom{a}{\mathrm{vech}(A)}; a \in \mathbb{R}^{d_1}, A \in \mathbb{R}^{d_1 \times d_1}$ symmetrisch positiv definit$\}$ und $\Lambda^1 = \mathbb{R}^{d_1}$. Die Abbildung $L : \tilde{\Theta} \to \Lambda$ ist stetig differenzierbar und injektiv mit nirgends verschwindender Funktionaldeterminante. Ferner ist die Abbildung C_{ϑ_2} gegeben durch

$$C_{\vartheta_2} : \Lambda^1 \to \mathbb{R}, \; c_{\vartheta_2}(\lambda_1) := -\frac{1}{2}\lambda_1^t \,\mathrm{vech}^{-1}(\vartheta_2)\lambda_1,$$

und die Abbildung R_{ϑ_2} gegeben durch

$$\begin{aligned} &R_{\vartheta_2} : (\mathbb{R}^{d_1}, \mathfrak{B}^{d_1}) \to (\mathbb{R}, \mathfrak{B}) \\ &R_{\vartheta_2}(x) := \Big(\frac{1}{\sqrt{2\pi}}\Big)^{d_1} \frac{1}{\sqrt{\det \mathrm{vech}^{-1}(\vartheta_2)}} \exp\Big(-\frac{1}{2} x^t \big(\mathrm{vech}^{-1}(\vartheta_2)\big)^{-1} x\Big). \end{aligned}$$

Vergleiche mit den Ausführungen in Abschnitt 8.

6.3.4 Bemerkung: Die aus Satz 6.3.1 hervorgehende Effizienz hängt im Fall von $\zeta_1 = \binom{X_1}{Y_1}$ oder $\zeta_1 = X_1 - Y_1$ nur von den Randverteilungen $Q_{\vartheta_0}^{X_1}$ und $Q_{\vartheta_0}^{Y_1}$ der verbundenen Stichprobe, vom Erwartungswert $\mathrm{E}_{\vartheta_0}(T_{\vartheta_2}(\zeta_1))$ und der Kovarianzmatrix $\mathrm{Cov}_{\vartheta_0}(T_{\vartheta_2}(\zeta_1))$ ab; weitere Abhängigkeitsstrukturen zwischen X_1 und Y_1 spielen dabei keine Rolle.

6.3.5 Korollar: *Es sei $\vartheta_0 = (\vartheta_0^1, \vartheta_2)$. Zudem sei $A_{\zeta_1}^{1,1}(\vartheta_0) = i_{\zeta_1}^{1,1}(\vartheta_0)$. Dann gilt:*

a) Es gilt die Äquivalenz

$$A_{\zeta_1'}^{1,1}(\vartheta_0') \leq \psi_{T_{\vartheta_2}(\zeta_1)}^1(\vartheta_0)^t \mathrm{Cov}_{\vartheta_0}\big(T_{\vartheta_2}(\zeta_1)\big)^{-1} \psi_{T_{\vartheta_2}(\zeta_1)}^1(\vartheta_0) \;\Leftrightarrow\; \sup_{\vartheta_1 \in \mathbb{R}^{d_1} \setminus \{0\}} \mathrm{eff}_{T',T}(\vartheta^0, \vartheta_1) \leq 1,$$

sowie die Äquivalenz

$$A_{\zeta_1'}^{1,1}(\vartheta_0') \geq \psi_{T_{\vartheta_2}(\zeta_1)}^1(\vartheta_0)^t \mathrm{Cov}_{\vartheta_0}\big(T_{\vartheta_2}(\zeta_1)\big)^{-1} \psi_{T_{\vartheta_2}(\zeta_1)}^1(\vartheta_0) \;\Leftrightarrow\; \inf_{\vartheta_1 \in \mathbb{R}^{d_1} \setminus \{0\}} \mathrm{eff}_{T',T}(\vartheta^0, \vartheta_1) \geq 1$$

b) Ist $d_1 = l$ und ist die Abbildung T_{ϑ_2} durch die Identität gegeben, gilt die Äquivalenz

$$A^{1,1}_{\zeta_1'}(\vartheta_0') \leq \psi^1_{\zeta_1}(\vartheta_0)^t \mathrm{Cov}_{\vartheta_0}(\zeta_1)^{-1} \psi^1_{\zeta_1}(\vartheta_0) \;\Leftrightarrow\; \sup_{\vartheta_1 \in \mathbb{R}^{d_1}\setminus\{0\}} \mathrm{eff}_{T',T}(\vartheta^0, \vartheta_1) \leq 1,$$

sowie die Äquivalenz

$$A^{1,1}_{\zeta_1'}(\vartheta_0') \geq \psi^1_{\zeta_1}(\vartheta_0)^t \mathrm{Cov}_{\vartheta_0}(\zeta_1)^{-1} \psi^1_{\zeta_1}(\vartheta_0) \;\Leftrightarrow\; \inf_{\vartheta_1 \in \mathbb{R}^{d_1}\setminus\{0\}} \mathrm{eff}_{T',T}(\vartheta^0, \vartheta_1) \geq 1.$$

c) Ist m eine gerade Zahl, $X_1 : (\Omega, \mathfrak{A}) \to (\mathbb{R}^{m/2}, \mathfrak{B}^{m/2})$, $Y_1 : (\Omega, \mathfrak{A}) \to (\mathbb{R}^{m/2}, \mathfrak{B}^{m/2})$, $\zeta_1 = X_1 - Y_1$, $d_1 = m/2$ und die Abbildung T_{ϑ_2} durch die Identität gegeben, so gilt die Äquivalenz

$$\begin{aligned} A^{1,1}_{\zeta_1'}(\vartheta_0') \leq & \psi^1_{X_1-Y_1}(\vartheta_0)^t \big(\mathrm{Cov}_{\vartheta^0}(X_1) + \mathrm{Cov}_{\vartheta^0}(Y_1) \\ & - \mathrm{Cov}_{\vartheta^0}(X_1, Y_1) - \mathrm{Cov}_{\vartheta^0}(X_1, Y_1)^t\big)^{-1} \psi^1_{X_1-Y_1}(\vartheta_0) \\ & \Leftrightarrow \sup_{\vartheta_1 \in \mathbb{R}^{d_1}\setminus\{0\}} \mathrm{eff}_{T',T}(\vartheta^0, \vartheta_1) \leq 1, \end{aligned}$$

sowie die Äquivalenz

$$\begin{aligned} A^{1,1}_{\zeta_1'}(\vartheta_0') \geq & \psi^1_{X_1-Y_1}(\vartheta_0)^t \big(\mathrm{Cov}_{\vartheta^0}(X_1) + \mathrm{Cov}_{\vartheta^0}(Y_1) \\ & - \mathrm{Cov}_{\vartheta^0}(X_1, Y_1) - \mathrm{Cov}_{\vartheta^0}(X_1, Y_1)^t\big)^{-1} \psi^1_{X_1-Y_1}(\vartheta_0) \\ & \Leftrightarrow \inf_{\vartheta_1 \in \mathbb{R}^{d_1}\setminus\{0\}} \mathrm{eff}_{T',T}(\vartheta^0, \vartheta_1) \geq 1. \end{aligned}$$

Beweis: Dies folgt unmittelbar aus Satz 6.3.1, Korollar 6.3.2 und Lemma 4.5.2. □

6.3.6 Bemerkung: Es sei $\vartheta_0 = (\vartheta_0^1, \vartheta_2)$. Weiter sei $d_1 = l$, $\psi^1_{T_{\vartheta_2}}(\vartheta_0)$ eine quadratische invertierbare Matrix und es sei $A^{1,1}_{\zeta_1}(\vartheta_0) = i^{1,1}_{\zeta_1}(\vartheta_0)$. Dann gilt die Äquivalenz

$$A^{1,1}_{\zeta_1'}(\vartheta_0') \leq \psi^1_{T_{\vartheta_2}(\zeta_1)}(\vartheta_0)^t \mathrm{Cov}_{\vartheta_0}\big(T_{\vartheta_2}(\zeta_1)\big)^{-1} \psi^1_{T_{\vartheta_2}(\zeta_1)}(\vartheta_0) \;\Leftrightarrow\; \sup_{\vartheta_1 \in \mathbb{R}^{d_1}\setminus\{0\}} \mathrm{eff}_{T',T}(\vartheta^0, \vartheta_1) \leq 1,$$

äquivalent

$$A^{1,1}_{\zeta_1'}(\vartheta_0')^{-1} \geq \psi^1_{T_{\vartheta_2}(\zeta_1)}(\vartheta_0)^{-1} \mathrm{Cov}_{\vartheta_0}\big(T_{\vartheta_2}(\zeta_1)\big) \big(\psi^1_{T_{\vartheta_2}(\zeta_1)}(\vartheta_0)^t\big)^{-1} \;\Leftrightarrow\; \sup_{\vartheta_1 \in \mathbb{R}^{d_1}\setminus\{0\}} \mathrm{eff}_{T',T}(\vartheta^0, \vartheta_1) \leq 1.$$

Ist darüber hinaus die Abbildung T_{ϑ_2} durch die Identität gegeben, gilt die Äquivalenz

$$A^{1,1}_{\zeta_1'}(\vartheta_0') \leq \psi^1_{\zeta_1}(\vartheta_0)^t \mathrm{Cov}_{\vartheta_0}(\zeta_1)^{-1} \psi^1_{\zeta_1}(\vartheta_0) \;\Leftrightarrow\; \sup_{\vartheta_1 \in \mathbb{R}^{d_1}\setminus\{0\}} \mathrm{eff}_{T',T}(\vartheta^0, \vartheta_1) \leq 1,$$

äquivalent

$$A^{1,1}_{\zeta_1'}(\vartheta_0')^{-1} \geq \psi^1_{\zeta_1}(\vartheta_0)^{-1} \mathrm{Cov}_{\vartheta_0}(\zeta_1) \big(\psi^1_{\zeta_1}(\vartheta_0)^t\big)^{-1} \;\Leftrightarrow\; \sup_{\vartheta_1 \in \mathbb{R}^{d_1}\setminus\{0\}} \mathrm{eff}_{T',T}(\vartheta^0, \vartheta_1) \leq 1.$$

Weiter gilt

$$\tilde{\text{eff}}_{T',T}(\vartheta^0) = \det A^{1,1}_{\zeta_1'}(\vartheta_0) \det\Big(\psi^1_{\zeta_1}(\vartheta_0)^{-1}\text{Cov}_{\vartheta_0}(\zeta_1)(\psi^1_{\zeta_1}(\vartheta_0)^t)^{-1}\Big).$$

Ist weiter m eine gerade Zahl, $X_1 : (\Omega, \mathfrak{A}) \to (\mathbb{R}^{m/2}, \mathfrak{B}^{m/2})$, $Y_1 : (\Omega, \mathfrak{A}) \to (\mathbb{R}^{m/2}, \mathfrak{B}^{m/2})$, $\zeta_1 = X_1 - Y_1$, $d_1 = m/2$ und die Abbildung T_{ϑ_2} durch die Identität gegeben, haben wir die Äquivalenz

$$\begin{aligned} A^{1,1}_{\zeta_1'}(\vartheta_0') \leq & \psi^1_{X_1-Y_1}(\vartheta_0)^t\big(\text{Cov}_{\vartheta^0}(X_1) + \text{Cov}_{\vartheta^0}(Y_1) \\ & - \text{Cov}_{\vartheta^0}(X_1, Y_1) - \text{Cov}_{\vartheta^0}(X_1, Y_1)^t\big)^{-1}\psi^1_{X_1-Y_1}(\vartheta_0) \\ & \Leftrightarrow \sup_{\vartheta_1 \in \mathbb{R}^{d_1}\setminus\{0\}} \text{eff}_{T',T}(\vartheta^0, \vartheta_1) \leq 1, \end{aligned}$$

äquivalent

$$\begin{aligned} A^{1,1}_{\zeta_1'}(\vartheta_0')^{-1} \geq & \psi^1_{X_1-Y_1}(\vartheta_0)^{-1}\big(\text{Cov}_{\vartheta^0}(X_1) + \text{Cov}_{\vartheta^0}(Y_1) \\ & - \text{Cov}_{\vartheta^0}(X_1, Y_1) - \text{Cov}_{\vartheta^0}(X_1, Y_1)^t\big)\big(\psi^1_{X_1-Y_1}(\vartheta_0)^t\big)^{-1} \\ & \Leftrightarrow \sup_{\vartheta_1 \in \mathbb{R}^{d_1}\setminus\{0\}} \text{eff}_{T',T}(\vartheta^0, \vartheta_1) \leq 1. \end{aligned}$$

Wenn $\psi^1_{\zeta_1}(\vartheta_0)$ eine quadratische invertierbare Matrix ist, gilt $\varphi_{\zeta_1}(\vartheta_0) = \psi^1_{\zeta_1}(\vartheta_0)^{-1}$. Dann sind die in Satz 6.2.1 entprechend genannten Kriterien nicht nur hinreichend, sondern sogar notwendig.

6.3.7 Bemerkung: Es sei $\vartheta_0 = (\vartheta_0^1, \vartheta_2)$ und $\vartheta_0' = (\vartheta_0^1, \vartheta_2')$. Es sei m eine gerade Zahl, $X_1 : (\Omega, \mathfrak{A}) \to (\mathbb{R}^{m/2}, \mathfrak{B}^{m/2})$, $Y_1 : (\Omega, \mathfrak{A}) \to (\mathbb{R}^{m/2}, \mathfrak{B}^{m/2})$, $d_1 = m/2$, $\zeta_1 = X_1 - Y_1$ und $\zeta_1' = X_1' - Y_1'$. Zudem sei die Effizienz gegeben durch

$$\text{eff}_{T',T}(\vartheta^0, \vartheta_1) = \frac{\vartheta_1^t i^{1,1}_{X_1'-Y_1'}(\vartheta_0')\vartheta_1}{\vartheta_1^t i^{1,1}_{X_1-Y_1}(\vartheta_0)\vartheta_1}.$$

Es sei die Familie von Verteilungen $\{Q^{X_1-Y_1}_\vartheta; \vartheta \in \Theta\}$ gemäß (6.3.1) eine Exponentialfamilie mit suffizienter Statistik T_{ϑ_2}.

Ist die Familie von Verteilungen $\{M^{X_1'-Y_1'}_{\vartheta'}; \vartheta' \in \tilde{\Theta}'\}$ gemäß (6.3.1) für ϑ_2' eine Exponentialfamilie mit suffizienter Statistik $T'_{\vartheta_2'}$, so kann unter analogen Voraussetzungen analog zum Beweis von Satz 6.3.1 gezeigt werden, dass in diesem Fall die Effizienz die Gestalt

$$\text{eff}_{T',T}(\vartheta^0, \vartheta_1) = \frac{\vartheta_1^t \psi^1_{t'_{\vartheta_2'}(X_1'-Y_1')}(\vartheta_0')^t \text{Cov}_{\vartheta_0'}\big(t'_{\vartheta_2'}(X_1' - Y_1')\big)^{-1}\psi^1_{t'_{\vartheta_2'}(X_1'-Y_1')}(\vartheta_0')\vartheta_1}{\vartheta_1^t \psi^1_{t_{\vartheta_2}(X_1-Y_1)}(\vartheta_0)^t \text{Cov}_{\vartheta_0}\big(t_{\vartheta_2}(X_1 - Y_1)\big)^{-1}\psi^1_{t_{\vartheta_2}(X_1-Y_1)}(\vartheta_0)\vartheta_1}$$

besitzt. Sind nun T_{ϑ_2} und $T'_{\vartheta_2'}$ durch die Identität gegeben, so gilt wegen (6.1.3)

$$\text{E}_{\vartheta_0}(X_1 - Y_1) = \text{E}_{\vartheta_0'}(X_1' - Y_1'),$$

d.h. $\mathrm{E}_{\vartheta_0}(X_1 - Y_1) = \mathrm{E}_{\vartheta_0'}(X_1' - Y_1')$ hängt nur noch von ϑ_0^1 ab. Daher gilt sogar $\psi^1_{X_1-Y_1}(\vartheta_0) = \psi^1_{X_1'-Y_1'}(\vartheta_0')$. Damit ist

$$\mathrm{eff}_{T',T}(\vartheta^0, \vartheta_1) = \frac{\vartheta_1^t \psi^1_{X_1-Y_1}(\vartheta_0)^t \mathrm{Cov}_{\vartheta_0'}(X_1' - Y_1')^{-1} \psi^1_{X_1-Y_1}(\vartheta_0)\vartheta_1}{\vartheta_1^t \psi^1_{X_1-Y_1}(\vartheta_0)^t \mathrm{Cov}_{\vartheta_0}(X_1 - Y_1)^{-1} \psi^1_{X_1-Y_1}(\vartheta_0)\vartheta_1}.$$

Es gilt die Implikation

$$\mathrm{Cov}_{\vartheta_0}(X_1 - Y_1) \le \mathrm{Cov}_{\vartheta_0'}(X_1' - Y_1') \;\Rightarrow\; \sup_{\vartheta_1 \in \mathbb{R}^{d_1} \setminus \{0\}} \mathrm{eff}_{T',T}(\vartheta^0, \vartheta_1) \le 1,$$

und die Implikation

$$\mathrm{Cov}_{\vartheta_0}(X_1 - Y_1) \ge \mathrm{Cov}_{\vartheta_0'}(X_1' - Y_1') \;\Rightarrow\; \inf_{\vartheta_1 \in \mathbb{R}^{d_1} \setminus \{0\}} \mathrm{eff}_{T',T}(\vartheta^0, \vartheta_1) \ge 1.$$

Weiter ist

$$\mathrm{Cov}_{\vartheta_0}(X_1 - Y_1) = \mathrm{Cov}_{\vartheta^0}(X_1) + \mathrm{Cov}_{\vartheta^0}(Y_1) - \mathrm{Cov}_{\vartheta^0}(X_1, Y_1) - \mathrm{Cov}_{\vartheta^0}(X_1, Y_1)^t,$$

und wegen (6.1.2) und (6.1.3)

$$\begin{aligned}\mathrm{Cov}_{\vartheta_0'}(X_1' - Y_1') &= \mathrm{Cov}_{\vartheta^0}(X_1') + \mathrm{Cov}_{\vartheta^0}(Y_1') - \mathrm{Cov}_{\vartheta^0}(X_1', Y_1') - \mathrm{Cov}_{\vartheta^0}(X_1', Y_1')^t \\ &= \mathrm{Cov}_{\vartheta^0}(X_1') + \mathrm{Cov}_{\vartheta^0}(Y_1') = \mathrm{Cov}_{\vartheta^0}(X_1) + \mathrm{Cov}_{\vartheta^0}(Y_1).\end{aligned}$$

Daher gelten die Implikationen

$$\mathrm{Cov}_{\vartheta^0}(X_1, Y_1) + \mathrm{Cov}_{\vartheta^0}(X_1, Y_1)^t \ge 0 \;\Rightarrow\; \sup_{\vartheta_1 \in \mathbb{R}^{d_1} \setminus \{0\}} \mathrm{eff}_{T',T}(\vartheta^0, \vartheta_1) \le 1$$

und

$$\mathrm{Cov}_{\vartheta^0}(X_1, Y_1) + \mathrm{Cov}_{\vartheta^0}(X_1, Y_1)^t \le 0 \;\Rightarrow\; \inf_{\vartheta_1 \in \mathbb{R}^{d_1} \setminus \{0\}} \mathrm{eff}_{T',T}(\vartheta^0, \vartheta_1) \ge 1.$$

Wegen Lemma 4.5.4 gelten dann auch die Implikationen

$$\mathrm{Cov}_{\vartheta^0}(X_1, Y_1) + \mathrm{Cov}_{\vartheta^0}(X_1, Y_1)^t \ge 0 \;\Rightarrow\; \widetilde{\mathrm{eff}}_{T',T}(\vartheta^0) \le 1$$

und

$$\mathrm{Cov}_{\vartheta^0}(X_1, Y_1) + \mathrm{Cov}_{\vartheta^0}(X_1, Y_1)^t \le 0 \;\Rightarrow\; \widetilde{\mathrm{eff}}_{T',T}(\vartheta^0) \ge 1.$$

6.3.8 Beispiel: Betrachten wir eine Situation aufbauend auf Beispiel 6.3.3: Es sei m eine gerade Zahl und $X_1 : (\Omega, \mathfrak{A}) \to (\mathbb{R}^{m/2}, \mathfrak{B}^{m/2})$, $Y_1 : (\Omega, \mathfrak{A}) \to (\mathbb{R}^{m/2}, \mathfrak{B}^{m/2})$. Weiter liege das zu Beginn dieser Arbeit genannte Testproblem

$$\mathrm{H}: \ \mathcal{L}(X_1) = \mathcal{L}(Y_1), \ \mathrm{K}: \ \mathcal{L}(X_1) \neq \mathcal{L}(Y_1),$$

vor, und wir gehen davon aus, dass sich die Verteilungen von X_1 und Y_1 nur in ihrem Erwartungswert unterscheiden. Es sei $\zeta_1 = X_1 - Y_1$, $\zeta_1' = X_1' - Y_1'$, sowie

$$\vartheta_1 = \mathrm{E}_\vartheta(X_1 - Y_1) = \mathrm{E}_{\vartheta'}(X_1' - Y_1'), \ \vartheta = (\vartheta_1, \vartheta_2) \in \tilde{\Theta}, \ \vartheta' = (\vartheta_1, \vartheta_2') \in \tilde{\Theta}',$$

und das Testproblem lasse sich äquivalent schreiben als

$$\mathrm{H}: \ \vartheta_1 = 0, \ \mathrm{K}: \ \vartheta_1 \in \Theta^1 \setminus \{0\}.$$

In diesem Fall ist es häufig sinnvoll, Tests basierend auf Beobachtungen der Form $X_1 - Y_1, \ldots, X_n - Y_n$ und $X_1' - Y_1', \ldots, X_n' - Y_n'$ vorzuschlagen. Dann ist $\psi^1_{\zeta_1}(\vartheta_0) = \psi^1_{\zeta_1'}(\vartheta_0')$ gegeben durch die Identität. Zudem sei die Effizienz durch

$$\mathrm{eff}_{T',T}(\vartheta^0, \vartheta_1) = \frac{\vartheta_1^t i^{1,1}_{\zeta_1'}(\vartheta_0')\vartheta_1}{\vartheta_1^t i^{1,1}_{\zeta_1}(\vartheta_0)\vartheta_1}$$

gegeben.

Es sei die Familie von Verteilungen $\{Q^{X_1 - Y_1}_\vartheta; \vartheta \in \tilde{\Theta}\}$ gemäß (6.3.1) für ϑ_2, mit $\vartheta_0 = (\vartheta_0^1, \vartheta_2)$, eine Exponentialfamilie mit suffizienter Statistik T_{ϑ_2}. Zudem sei die Abbildung T_{ϑ_2} durch die Identität gegeben. Weiter sei die Familie von Verteilungen $\{M^{X_1' - Y_1'}_{\vartheta'}; \vartheta' \in \tilde{\Theta}'\}$ gemäß (6.3.1) für ϑ_2', mit $\vartheta_0' = (\vartheta_0^1, \vartheta_2')$, eine Exponentialfamilie mit suffizienter Statistik $T'_{\vartheta_2'}$. Zudem sei die Abbildung $T'_{\vartheta_2'}$ durch die Identität gegeben. Dann gelten die Implikationen

$$\mathrm{Cov}_{\vartheta^0}(X_1, Y_1) + \mathrm{Cov}_{\vartheta^0}(X_1, Y_1)^t \geq 0 \ \Rightarrow \sup_{\vartheta_1 \in \mathbb{R}^{d_1} \setminus \{0\}} \mathrm{eff}_{T',T}(\vartheta^0, \vartheta_1) \leq 1$$

und

$$\mathrm{Cov}_{\vartheta^0}(X_1, Y_1) + \mathrm{Cov}_{\vartheta^0}(X_1, Y_1)^t \leq 0 \ \Rightarrow \inf_{\vartheta_1 \in \mathbb{R}^{d_1} \setminus \{0\}} \mathrm{eff}_{T',T}(\vartheta^0, \vartheta_1) \geq 1,$$

sowie die Implikationen

$$\mathrm{Cov}_{\vartheta^0}(X_1, Y_1) + \mathrm{Cov}_{\vartheta^0}(X_1, Y_1)^t \geq 0 \ \Rightarrow \ \tilde{\mathrm{eff}}_{T',T}(\vartheta^0) \leq 1$$

und

$$\mathrm{Cov}_{\vartheta^0}(X_1, Y_1) + \mathrm{Cov}_{\vartheta^0}(X_1, Y_1)^t \leq 0 \ \Rightarrow \ \tilde{\mathrm{eff}}_{T',T}(\vartheta^0) \geq 1.$$

Konkret kann z.B. der Fall betrachtet werden, dass

$$\begin{pmatrix} X_1 \\ Y_1 \end{pmatrix} \sim \mathrm{N}_{2d_1}\left(\begin{pmatrix} a \\ b \end{pmatrix}, \begin{pmatrix} \Sigma & \Pi \\ \Pi^t & \Sigma \end{pmatrix} \right)$$

gilt, wobei $a, b \in \mathbb{R}^{d_1}$ und $\Sigma, \Pi \in \mathbb{R}^{d_1 \times d_1}$ unbekannt seien, und Σ und die zugrunde liegende Kovarianzmatrix von $\left(\begin{smallmatrix} X_1 \\ Y_1 \end{smallmatrix}\right)$ positiv definit seien. Dann gilt wegen (6.1.2) und (6.1.3) auch

$$\begin{pmatrix} X_1' \\ Y_1' \end{pmatrix} \sim \mathrm{N}_{2d_1}\left(\begin{pmatrix} a \\ b \end{pmatrix}, \begin{pmatrix} \Sigma & 0 \\ 0 & \Sigma \end{pmatrix} \right).$$

Dass in diesem konkreten Beispiel alle oben genannten Voraussetzungen erfüllt sind, kann den Ausführungen aus Beispiel 6.3.3 entnommen werden.

6.3.9 Bemerkung: Die in Bemerkung 6.3.7 und Beispiel 6.3.8 erarbeiteten Kriterien sind unter den genannten Voraussetzungen auch hinreichende Kriterien für $\widetilde{\mathrm{eff}}_{T',T}(\vartheta^0) \leq 1$ bzw. $\widetilde{\mathrm{eff}}_{T',T}(\vartheta^0) \geq 1$. Beachte hierzu Lemma 4.5.4.

Nun wollen wir einen Fall betrachten, wie er zum Beispiel in Abschnitt 9 vorliegt. Wir gehen dazu davon aus, dass

$$A_{\zeta_1}^{1,1}(\vartheta_0) = i_{\zeta_1}^{1,1}(\vartheta_0) - i_{\zeta_1}^{1,2}(\vartheta_0) i_{\zeta_1}^{2,2}(\vartheta_0)^{-1} i_{\zeta_1}^{2,1}(\vartheta_0).$$

Weiter nehmen wir an, dass eine injektive und stetig differenzierbare Abbildung $\ell : \mathbb{R}^k \to \mathbb{R}^k$ mit überall regulärer Funktionalmatrix existiert, sodass

$$f(\cdot;\vartheta) = \exp\Big(\ell(\vartheta)^t t(\cdot) - c\big(\ell(\vartheta)\big)\Big) r(\cdot) \text{ für alle } \vartheta \in \tilde{\Theta} \tag{6.3.2}$$

ist, mit $c : \mathbb{R}^k \to \mathbb{R}$, $t : (\mathbb{R}^l, \mathfrak{B}^l) \to (\mathbb{R}^k, \mathfrak{B}^k)$ und $r : (\mathbb{R}^l, \mathfrak{B}^l) \to (\mathbb{R}, \mathfrak{B})$ μ-f.ü. positiv. Mit $\Lambda := \ell(\tilde{\Theta})$ ist die Familie von Verteilungen von ζ_1 also eine k-parametrische Exponentialfamilie im Parameter $\lambda \in \Lambda$.

Aus den Voraussetzungen folgt, dass die Menge Λ offen ist, siehe Satz 171.2 Teil a) in [13]. Wir definieren $S_{\ell(\vartheta)} := Q_\vartheta^{\zeta_1}$, $\vartheta \in \tilde{\Theta}$. Es ist die Familie von Verteilungen von ζ_1 durch die Elemente von Λ injektiv parametrisiert. Dies bedeutet, dass mit $\{S_\lambda; \lambda \in \Lambda\}$ eine Familie von Wahrscheinlichkeitsmaßen auf dem Messraum $(\mathbb{R}^l, \mathfrak{B}^l)$ mit $\{S_\lambda; \lambda \in \Lambda\} = \{Q_\vartheta^{\zeta_1}; \vartheta \in \tilde{\Theta}\}$ gegeben ist.

Wir nehmen an, dass für alle $\gamma \in \mathbb{R}$ und alle $a \in \mathbb{R}^k$ $\mu(a^t t \neq \gamma, r > 0) > 0$ gilt. Zudem existiere kein $a \in \mathbb{R}^k$ und kein $b \in \mathbb{R}$, sodass $a^t \ell(\vartheta) = b$ für alle $\vartheta \in \tilde{\Theta}$ gilt.

Die Abbildung c ist auf Λ beliebig oft differenzierbar und die zugehörige Hessematrix von

c ist auf Λ positiv definit. Siehe Satz 1.164 a) und b) in [41]. Es sei mit $\lambda_1 = (\lambda_1, \ldots, \lambda_k)^t$ der Nabla-Ableitungsoperator ∇_λ durch

$$\nabla_\lambda := \Big(\frac{\partial}{\partial \lambda_1}, \ldots, \frac{\partial}{\partial \lambda_k}\Big)^t$$

definiert. Wegen $\nabla_\lambda c(\lambda) = \mathrm{E}_\lambda(t(\zeta_1)),\ \lambda \in \Lambda$, ist die Existenz und Endlichkeit von $\mathrm{E}_\lambda(t(\zeta_1))$, $\lambda \in \Lambda$, sichergestellt. Es sei die Abbildung $e : \Lambda \to E := \{\mathrm{E}_\lambda(t(\zeta_1)); \lambda \in \Lambda\}$ gegeben durch

$$e(\lambda) := \mathrm{E}_\lambda\big(t(\zeta_1)\big),\ \lambda \in \Lambda,$$

d.h. $\eta \in E$ ist ein Momentenparameter.

Wegen $\nabla_\lambda c(\lambda) = \mathrm{E}_\lambda(t(\zeta_1)),\ \lambda \in \Lambda$, und da c auf Λ beliebig oft differenzierbar ist, ist die Abbildung e auf Λ stetig differenzierbar, wobei die zugehörige Jacobimatrix die Hessematrix von c ist, und daher auf Λ positiv definit ist. Daher existiert nach dem Satz über die Umkehrabbildung, siehe Satz 171.1 in [13], die inverse Abbildung $e^{-1} : E \to \Lambda$ und diese ist ebenfalls auf E differenzierbar. Wir schreiben auch $\overline{\ell} := e^{-1}$.

Zudem sei die Abbildung $\overline{e} : \tilde{\Theta} \to \tilde{E} := \{\mathrm{E}_\vartheta(t(\zeta_1)); \vartheta \in \tilde{\Theta}\}$ durch

$$\overline{e}(\vartheta) := \mathrm{E}_\vartheta\big(t(\zeta_1)\big),\ \vartheta \in \tilde{\Theta},$$

gegeben. Es gilt $\tilde{E} = E$. Wegen $\overline{e} = e \circ \ell$ und der Differenzierbarkeit von e und ℓ auf Λ bzw. $\tilde{\Theta}$ ist die Abbildung $\overline{e}$ auf $\tilde{\Theta}$ differenzierbar.

Es sei $\vartheta_0 \in \tilde{\Theta}_0$ fest gewählt. Wir setzen $\eta_0 := \overline{e}(\vartheta_0)$ und damit $\lambda_0 := \overline{\ell}(\eta_0)$. Dann ist

$$\begin{aligned}\lambda_0 = \overline{\ell}(\eta_0) = e^{-1}(\eta_0) = e^{-1}\big(\overline{e}(\vartheta_0)\big) &= e^{-1}\Big(\mathrm{E}_{\vartheta_0}\big(t(\zeta_1)\big)\Big)\\ = e^{-1}\Big(\mathrm{E}_{\ell(\vartheta_0)}\big(t(\zeta_1)\big)\Big) = e^{-1}\Big(e\big(\ell(\vartheta_0)\big)\Big) &= \ell(\vartheta_0).\end{aligned}$$

In der Tat lassen sich im vorliegenden Fall durch den Übergang von der Parametrisierung durch $\tilde{\Theta}$ in die Parametrisierung durch Λ die zugehörigen Fisher-Informationsmatrizen leichter berechnen und so Ergebnisse beim Studium der Effizienz erzielen.

6.3.10 Satz: *Es sei $A^{1,1}_{\zeta_1}(\vartheta_0) = i^{1,1}_{\zeta_1}(\vartheta_0) - i^{1,2}_{\zeta_1}(\vartheta_0) i^{2,2}_{\zeta_1}(\vartheta_0)^{-1} i^{2,1}_{\zeta_1}(\vartheta_0)$. Es ist die Matrix*

$$((\mathrm{I}_{d_1 \times d_1}, 0_{d_1 \times (k-d_1)})(\psi_{t(\zeta_1)}(\vartheta_0)^t \mathrm{Cov}_{\vartheta_0}(t(\zeta_1))^{-1} \psi_{t(\zeta_1)}(\vartheta_0))^{-1} (\mathrm{I}_{d_1 \times d_1}, 0_{d_1 \times (k-d_1)})^t)^{-1}$$

positiv definit und

$$A^{1,1}_{\zeta_1}(\vartheta_0) = \Big((\mathrm{I}_{d_1 \times d_1}, 0_{d_1 \times (k-d_1)})\big(\psi_{t(\zeta_1)}(\vartheta_0)^t \mathrm{Cov}_{\vartheta_0}\big(t(\zeta_1)\big)^{-1} \psi_{t(\zeta_1)}(\vartheta_0)\big)^{-1} (\mathrm{I}_{d_1 \times d_1}, 0_{d_1 \times (k-d_1)})^t\Big)^{-1}.$$

Beweis: Wir definieren folgende zusätzliche Nabla-Ableitungsoperatoren:

$$\nabla_\vartheta := \Big(\frac{\partial}{\partial \vartheta_1}, \ldots, \frac{\partial}{\partial \vartheta_d}\Big)^t,\ \nabla_\eta := \Big(\frac{\partial}{\partial \eta}, \ldots, \frac{\partial}{\partial \eta_d}\Big)^t.$$

Zudem bezeichne $I_{\zeta_1}(\lambda_0)$ die Fisher-Informationsmatrix von ζ_1 bei Parametrisierung durch Λ an der Stelle λ_0, und $\mathbb{I}_{\zeta_1}(\eta_0)$ die Fisher-Informationsmatrix von ζ_1 bei Parametrisierung durch E an der Stelle η_0. Es gilt

$$\nabla_\lambda c(\lambda)_{|\lambda=\lambda_0} = \mathrm{E}_{\lambda_0}\big(t(\zeta_1)\big)$$

sowie

$$I_{\zeta_1}(\lambda_0) = \nabla_\lambda \nabla_\lambda^t c(\lambda)_{|\lambda=\lambda_0} = \mathrm{Cov}_{\lambda_0}\big(t(\zeta_1)\big),$$

wobei die Existenz beider Ausdrücke sowie die positive Definitheit der Informationsmatrix $I_{\zeta_1}(\lambda_0)$ nach Annahme sichergestellt ist.

Es erfolgt die Transformation der Informationsmatrizen unter den verschiedenen Parametrisierungen durch

$$\mathbb{I}_{\zeta_1}(\eta_0) = \big(\nabla_\eta \overline{\ell}(\eta)^t\big)^t_{|\eta=\eta_0} I_{\zeta_1}(\lambda_0) \nabla_\eta \overline{\ell}(\eta)^t_{|\eta=\eta_0}$$

und

$$i_{\zeta_1}(\vartheta_0) = \big(\nabla_\vartheta \overline{e}(\vartheta)^t\big)^t_{|\vartheta=\vartheta_0} \mathbb{I}_{\zeta_1}(\eta_0) \nabla_\vartheta \overline{e}(\vartheta_1)^t_{|\vartheta=\vartheta_0}.$$

Dann ist

$$\nabla_{\lambda_1} e(\lambda)^t_{|\lambda_1=\lambda_0} = \nabla_\lambda \nabla_\lambda^t c(\lambda)_{|\lambda=\lambda_0} = I_{\zeta_1}(\lambda_0),$$

und deshalb wegen der Invertierbarkeit der Abbildung e

$$\nabla_\eta \overline{\ell}(\eta)^t_{|\eta=\eta_0} = \nabla_\eta e^{-1}(\eta)^t_{|\eta=\eta_0} = \big(\nabla_\lambda e(\lambda)^t_{|\lambda=\lambda_0}\big)^{-1} = I_{\zeta_1}(\lambda_0)^{-1} = \mathrm{Cov}_{\lambda_0}\big(t(\zeta_1)\big)^{-1}.$$

Damit folgt

$$\mathbb{I}_{\zeta_1}(\eta_0) = \big(\nabla_\eta \overline{\ell}(\eta)^t\big)^t_{|\eta=\eta_0} I_{\zeta_1}(\lambda_0) \nabla_\eta \overline{\ell}(\eta)^t_{|\eta=\eta_0} = \mathrm{Cov}_{\lambda_0}\big(t(\zeta_1)\big)^{-1},$$

und damit mit $\lambda_0 = \ell(\vartheta_0)$ und $\nabla_\vartheta \overline{e}(\vartheta)^t|_{\vartheta=\vartheta_0} = \psi_{t(\zeta_1)}(\vartheta_0)$

$$\begin{aligned}
i_{\zeta_1}(\vartheta_0) &= \big(\nabla_\vartheta \overline{e}(\vartheta)^t\big)^t_{|\vartheta=\vartheta_0} \mathbb{I}_{\zeta_1}(\eta_0) \nabla_\vartheta \overline{e}(\vartheta)^t_{|\vartheta=\vartheta_0} \\
&= \big(\nabla_\vartheta \overline{e}(\vartheta)^t\big)^t_{|\vartheta=\vartheta_0} \mathrm{Cov}_{\lambda_0}\big(t(\zeta_1)\big)^{-1} \nabla_\vartheta \overline{e}(\vartheta)^t_{|\vartheta=\vartheta_0} \\
&= \big(\nabla_\vartheta \overline{e}(\vartheta)^t\big)^t_{|\vartheta=\vartheta_0} \mathrm{Cov}_{\vartheta_0}\big(t(\zeta_1)\big)^{-1} \nabla_\vartheta \overline{e}(\vartheta)^t_{|\vartheta_1=\vartheta_0} \\
&= \psi_{t(\zeta_1)}(\vartheta_0)^t_{|\vartheta=\vartheta_0} \mathrm{Cov}_{\vartheta_0}\big(t(\zeta_1)\big)^{-1} \psi_{t(\zeta_1)}(\vartheta_0).
\end{aligned}$$

Dies ist äquivalent zu

$$i_{\zeta_1}(\vartheta_0)^{-1} = \big(\psi_{t(\zeta_1)}(\vartheta_0)^t \mathrm{Cov}_{\vartheta_0}\big(t(\zeta_1)\big)^{-1} \psi_{t(\zeta_1)}(\vartheta_0)\big)^{-1}.$$

Wegen

$$\big(i_{\zeta_1}(\vartheta_0)^{-1}\big)_{\substack{1\le i\le d_1\\1\le j\le d_1}} = \big(i^{1,1}_{\zeta_1}(\vartheta_0) - i^{1,2}_{\zeta_1}(\vartheta_0) i^{2,2}_{\zeta_1}(\vartheta_0)^{-1} i^{2,1}_{\zeta_1}(\vartheta_0)\big)^{-1} = A^{1,1}_{\zeta_1}(\vartheta_0)^{-1},$$

folgt daraus

$$A^{1,1}_{\zeta_1}(\vartheta_0)^{-1} = (\mathrm{I}_{d_1\times d_1}, 0_{d_1\times(k-d_1)})\big(\psi_{t(\zeta_1)}(\vartheta_0)^t \mathrm{Cov}_{\vartheta_0}\big(t(\zeta_1)\big)^{-1}\psi_{t(\zeta_1)}(\vartheta_0)\big)^{-1}(\mathrm{I}_{d_1\times d_1}, 0_{d_1\times(k-d_1)})^t$$

und somit die Behauptung. □

6.3.11 Korollar: *Es sei* $A^{1,1}_{\zeta_1}(\vartheta_0) = i^{1,1}_{\zeta_1}(\vartheta_0) - i^{1,2}_{\zeta_1}(\vartheta_0) i^{2,2}_{\zeta_1}(\vartheta_0)^{-1} i^{2,1}_{\zeta_1}(\vartheta_0)$. *Ist* $k = l$ *und ist die Abbildung* t *durch die Identität gegeben, so ist die Matrix*

$$((\mathrm{I}_{d_1\times d_1}, 0_{d_1\times(k-d_1)})(\psi_{\zeta_1}(\vartheta_0)^t \mathrm{Cov}_{\vartheta_0}(\zeta_1)^{-1}\psi_{\zeta_1}(\vartheta_0))^{-1}(\mathrm{I}_{d_1\times d_1}, 0_{d_1\times(k-d_1)})^t)^{-1}$$

positiv definit, und man erhält

$$A^{1,1}_{\zeta_1}(\vartheta_0) = \Big((\mathrm{I}_{d_1\times d_1}, 0_{d_1\times(k-d_1)})\big(\psi_{\zeta_1}(\vartheta_0)^t \mathrm{Cov}_{\vartheta_0}(\zeta_1)^{-1}\psi_{\zeta_1}(\vartheta_0)\big)^{-1}(\mathrm{I}_{d_1\times d_1}, 0_{d_1\times(k-d_1)})^t\Big)^{-1}.$$

Es sei m *eine gerade Zahl,* $X_1 : (\Omega, \mathfrak{A}) \to (\mathbb{R}^{m/2}, \mathfrak{B}^{m/2})$, $Y_1 : (\Omega, \mathfrak{A}) \to (\mathbb{R}^{m/2}, \mathfrak{B}^{m/2})$, $k = m/2$, $\zeta_1 = X_1 - Y_1$ *und die Abbildung* t *durch die Identität gegeben. Dann ist die Matrix*

$$\begin{aligned}\Big(&(\mathrm{I}_{d_1\times d_1}, 0_{d_1\times(k-d_1)})\big(\psi_{X_1-Y_1}(\vartheta_0)^t(\mathrm{Cov}_{\vartheta^0}(X_1) + \mathrm{Cov}_{\vartheta^0}(Y_1) - \mathrm{Cov}_{\vartheta^0}(X_1, Y_1)\\ &- \mathrm{Cov}_{\vartheta^0}(X_1, Y_1)^t)^{-1}\psi_{X_1-Y_1}(\vartheta_0)\big)^{-1}(\mathrm{I}_{d_1\times d_1}, 0_{d_1\times(k-d_1)})^t\Big)^{-1}\end{aligned}$$

positiv definit und man hat

$$\begin{aligned}A^{1,1}_{\zeta_1}(\vartheta_0) =&\Big((\mathrm{I}_{d_1\times d_1}, 0_{d_1\times(k-d_1)})\big(\psi_{X_1-Y_1}(\vartheta_0)^t\big(\mathrm{Cov}_{\vartheta^0}(X_1) + \mathrm{Cov}_{\vartheta^0}(Y_1)\\ &- \mathrm{Cov}_{\vartheta^0}(X_1, Y_1) - \mathrm{Cov}_{\vartheta^0}(X_1, Y_1)^t\big)^{-1}\psi_{X_1-Y_1}(\vartheta_0)\big)^{-1}(\mathrm{I}_{d_1\times d_1}, 0_{d_1\times(k-d_1)})^t\Big)^{-1}.\end{aligned}$$

Beweis: Dies folgt mit Hilfe von Satz 6.3.10 analog zum Beweis von Korollar 6.3.2. □

6.3.12 Beispiel: Es sei m eine gerade Zahl und $X_1 : (\Omega, \mathfrak{A}) \to (\mathbb{R}^{m/2}, \mathfrak{B}^{m/2})$, $Y_1 : (\Omega, \mathfrak{A}) \to (\mathbb{R}^{m/2}, \mathfrak{B}^{m/2})$. Weiter sei $\zeta_1 = X_1 - Y_1$ und

$$\binom{\vartheta_1}{\vartheta_2} = \vartheta = \mathrm{E}_\vartheta(X_1 - Y_1),\ (\vartheta_1, \vartheta_2) \in \tilde{\Theta}.$$

Wir betrachten für ein gegebenes $\vartheta^1_0 \in \Theta^1$ das Testproblem

$$\mathrm{H}:\ \vartheta_1 = \vartheta^1_0,\ \mathrm{K}:\ \vartheta_1 \in \Theta^1 \setminus \{\vartheta^1_0\}.$$

Zudem sei $A^{1,1}_{\zeta_1}(\vartheta_0) = i^{1,1}_{\zeta_1}(\vartheta_0) - i^{1,2}_{\zeta_1}(\vartheta_0) i^{2,2}_{\zeta_1}(\vartheta_0)^{-1} i^{2,1}_{\zeta_1}(\vartheta_0)$. Es ist $\psi_{\zeta_1}(\vartheta_0)$ die Identität. Weiter sei die Familie von Verteilungen $\{Q^{X_1-Y_1}_\vartheta; \vartheta \in \Theta\}$ gemäß (6.3.2) eine Exponentialfamilie mit suffizienter Statistik t. Zudem sei die Abbildung t durch die Identität gegeben. Dann gilt

$$\begin{aligned}\mathrm{eff}_{T',T}(\vartheta^0,\vartheta_1) =&\vartheta_1^t A^{1,1}_{\zeta_1'}(\vartheta_0')\vartheta_1\Bigg(\vartheta_1^t\Big((\mathrm{I}_{d_1\times d_1},0_{d_1\times(k-d_1)})\big(\mathrm{Cov}_{\vartheta^0}(X_1)+\mathrm{Cov}_{\vartheta^0}(Y_1)\\&-\mathrm{Cov}_{\vartheta^0}(X_1,Y_1)-\mathrm{Cov}_{\vartheta^0}(X_1,Y_1)^t\big)^{-1}(\mathrm{I}_{d_1\times d_1},0_{d_1\times(k-d_1)})^t\Big)^{-1}\vartheta_1\Bigg)^{-1},\end{aligned}$$

sowie

$$\begin{aligned}\mathrm{eff}^{\sim}_{T',T}(\vartheta^0) =&\det A^{1,1}_{\zeta_1'}(\vartheta_0')\det\Big((\mathrm{I}_{d_1\times d_1},0_{d_1\times(k-d_1)})\big(\mathrm{Cov}_{\vartheta^0}(X_1)+\mathrm{Cov}_{\vartheta^0}(Y_1)\\&-\mathrm{Cov}_{\vartheta^0}(X_1,Y_1)-\mathrm{Cov}_{\vartheta^0}(X_1,Y_1)^t\big)^{-1}(\mathrm{I}_{d_1\times d_1},0_{d_1\times(k-d_1)})^t\Big).\end{aligned}$$

Konkret kann z.B. der Fall betrachtet werden, dass

$$\begin{pmatrix}X_1\\Y_1\end{pmatrix}\sim\mathrm{N}_{2k}\left(\begin{pmatrix}a\\b\end{pmatrix},\begin{pmatrix}\Sigma&\Pi\\\Pi^t&\Delta\end{pmatrix}\right)$$

gilt, wobei $a,b\in\mathbb{R}^k$ unbekannt, und $\Sigma,\Delta,\Pi\in\mathbb{R}^{k\times k}$ bekannt seien, und Σ, Δ und die zugrunde liegende Kovarianzmatrix von $\left(\begin{smallmatrix}X_1\\Y_1\end{smallmatrix}\right)$ positiv definit seien. Dann ist

$$X_1-Y_1\sim\mathrm{N}_d(\mu,\Gamma),$$

mit $\mu:=a-b$ und $\Gamma:=\Sigma+\Delta-(\Pi+\Pi^t)$, und Γ ist positiv definit. Es ist also $\Theta^1=\mathbb{R}^{d_1}$, $\tilde{\Theta}=\mathbb{R}^{d_1}\times\mathbb{R}^{k-d_1}$, $d_1\le k$. Zudem ist $\vartheta_1=\mu_1$, $\vartheta_2=\mu_2$ und $\vartheta=\mu$ mit $\mu=(\mu_1,\mu_2)$. Die Dichte von X_1-Y_1 ist gegeben durch

$$\begin{aligned}f\big(x;(\vartheta_1,\vartheta_2)\big)&=\Big(\frac{1}{\sqrt{2\pi}}\Big)^k\frac{1}{\sqrt{\det\Gamma}}\exp\Big(-\frac{1}{2}(x-\mu)^t\Gamma^{-1}(x-\mu)\Big)\\&=\Big(\frac{1}{\sqrt{2\pi}}\Big)^k\frac{1}{\sqrt{\det\Gamma}}\exp\Big(\mu^t\Gamma^{-1}x-\frac{1}{2}\mu^t\Gamma^{-1}\mu-\frac{1}{2}x^t\Gamma^{-1}x\Big),\ x\in\mathbb{R}^k.\end{aligned}$$

Wir definieren die Abbildung $\ell(\vartheta):=\Gamma^{-1}\vartheta$, $\vartheta\in\tilde{\Theta}$. Die Abbildung $\ell:\tilde{\Theta}\to\Lambda$ ist stetig differenzierbar und injektiv mit nirgends verschwindender Funktionaldeterminante. Ferner ist $\Lambda=\mathbb{R}^k$. Es ist die Abbildung c gegeben durch

$$c:\Lambda\to\mathbb{R},\ c(\lambda):=-\frac{1}{2}\lambda^t\Gamma\lambda,$$

und die Abbildung r gegeben durch

$$r:(\mathbb{R}^k,\mathfrak{B}^k)\to(\mathbb{R},\mathfrak{B}),\ r(x):=\Big(\frac{1}{\sqrt{2\pi}}\Big)^k\frac{1}{\sqrt{\det\Gamma}}\exp\Big(-\frac{1}{2}x^t\Gamma^{-1}x\Big).$$

6.3.13 Bemerkung: Die aus Satz 6.3.10 resultierende Effizienz hängt im Fall von $\zeta_1 = \binom{X_1}{Y_1}$ oder $\zeta_1 = X_1 - Y_1$ nur von den Randverteilungen $Q^{X_1}_{\vartheta_0}$ und $Q^{Y_1}_{\vartheta_0}$ der verbundenen Stichprobe, vom Erwartungswert $\mathrm{E}_{\vartheta_0}(t(\zeta_1))$ und der Kovarianzmatrix $\mathrm{Cov}_{\vartheta_0}(t(\zeta_1))$ ab; weitere Abhängigkeitsstrukturen zwischen X_1 und Y_1 spielen dabei keine Rolle.

6.3.14 Korollar: *Es sei $A^{1,1}_{\zeta_1}(\vartheta_0) = i^{1,1}_{\zeta_1}(\vartheta_0) - i^{1,2}_{\zeta_1}(\vartheta_0) i^{2,2}_{\zeta_1}(\vartheta_0)^{-1} i^{2,1}_{\zeta_1}(\vartheta_0)$. Dann gilt:*

a) Es gilt die Äquivalenz

$$A^{1,1}_{\zeta_1'}(\vartheta_0') \leq \Big((\mathrm{I}_{d_1 \times d_1}, 0_{d_1 \times (k-d_1)}) \big(\psi_{t(\zeta_1)}(\vartheta_0)^t \mathrm{Cov}_{\vartheta_0}\big(t(\zeta_1)\big)^{-1} \psi_{t(\zeta_1)}(\vartheta_0) \big)^{-1} (\mathrm{I}_{d_1 \times d_1}, 0_{d_1 \times (k-d_1)})^t \Big)^{-1} \Leftrightarrow \sup_{\vartheta_1 \in \mathbb{R}^{d_1} \setminus \{0\}} \mathrm{eff}_{T',T}(\vartheta^0, \vartheta_1) \leq 1,$$

sowie die Äquivalenz

$$A^{1,1}_{\zeta_1'}(\vartheta_0') \geq \Big((\mathrm{I}_{d_1 \times d_1}, 0_{d_1 \times (k-d_1)}) \big(\psi_{t(\zeta_1)}(\vartheta_0)^t \mathrm{Cov}_{\vartheta_0}\big(t(\zeta_1)\big)^{-1} \psi_{t(\zeta_1)}(\vartheta_0) \big)^{-1} (\mathrm{I}_{d_1 \times d_1}, 0_{d_1 \times (k-d_1)})^t \Big)^{-1} \Leftrightarrow \inf_{\vartheta_1 \in \mathbb{R}^{d_1} \setminus \{0\}} \mathrm{eff}_{T',T}(\vartheta^0, \vartheta_1) \geq 1.$$

b) Ist $k = l$ und ist die Abbildung t durch die Identität gegeben, gilt die Äquivalenz

$$A^{1,1}_{\zeta_1'}(\vartheta_0') \leq \Big((\mathrm{I}_{d_1 \times d_1}, 0_{d_1 \times (k-d_1)}) \big(\psi_{\zeta_1}(\vartheta_0)^t \mathrm{Cov}_{\vartheta_0}(\zeta_1)^{-1} \psi_{\zeta_1}(\vartheta_0) \big)^{-1} (\mathrm{I}_{d_1 \times d_1}, 0_{d_1 \times (k-d_1)})^t \Big)^{-1} \Leftrightarrow \sup_{\vartheta_1 \in \mathbb{R}^{d_1} \setminus \{0\}} \mathrm{eff}_{T',T}(\vartheta^0, \vartheta_1) \leq 1,$$

sowie die Äquivalenz

$$A^{1,1}_{\zeta_1'}(\vartheta_0') \geq \Big((\mathrm{I}_{d_1 \times d_1}, 0_{d_1 \times (k-d_1)}) \big(\psi_{\zeta_1}(\vartheta_0)^t \mathrm{Cov}_{\vartheta_0}(\zeta_1)^{-1} \psi_{\zeta_1}(\vartheta_0) \big)^{-1} (\mathrm{I}_{d_1 \times d_1}, 0_{d_1 \times (k-d_1)})^t \Big)^{-1} \Leftrightarrow \inf_{\vartheta_1 \in \mathbb{R}^{d_1} \setminus \{0\}} \mathrm{eff}_{T',T}(\vartheta^0, \vartheta_1) \geq 1.$$

c) Ist m eine gerade Zahl, $X_1 : (\Omega, \mathfrak{A}) \to (\mathbb{R}^{m/2}, \mathfrak{B}^{m/2})$, $Y_1 : (\Omega, \mathfrak{A}) \to (\mathbb{R}^{m/2}, \mathfrak{B}^{m/2})$, $\zeta_1 = X_1 - Y_1$, $k = m/2$ und die Abbildung t durch die Identität gegeben, so gilt die Äquivalenz

$$A^{1,1}_{\zeta_1'}(\vartheta_0') \leq \bigg((\mathrm{I}_{d_1 \times d_1}, 0_{d_1 \times (k-d_1)}) \Big(\psi_{X_1 - Y_1}(\vartheta_0)^t \big(\mathrm{Cov}_{\vartheta^0}(X_1) + \mathrm{Cov}_{\vartheta^0}(Y_1) - \mathrm{Cov}_{\vartheta^0}(X_1, Y_1) - \mathrm{Cov}_{\vartheta^0}(X_1, Y_1)^t \big)^{-1} \psi_{X_1 - Y_1}(\vartheta_0) \Big)^{-1} (\mathrm{I}_{d_1 \times d_1}, 0_{d_1 \times (k-d_1)})^t \bigg)^{-1}$$
$$\Leftrightarrow \sup_{\vartheta_1 \in \mathbb{R}^{d_1} \setminus \{0\}} \mathrm{eff}_{T',T}(\vartheta^0, \vartheta_1) \leq 1,$$

sowie die Äquivalenz

$$A^{1,1}_{\zeta_1'}(\vartheta_0') \geq \Big((\mathrm{I}_{d_1\times d_1}, 0_{d_1\times(k-d_1)}) \big(\psi_{X_1-Y_1}(\vartheta_0)^t \big(\mathrm{Cov}_{\vartheta_0}(X_1) + \mathrm{Cov}_{\vartheta_0}(Y_1) - \mathrm{Cov}_{\vartheta_0}(X_1,Y_1) - \mathrm{Cov}_{\vartheta_0}(X_1,Y_1)^t \big)^{-1} \psi_{X_1-Y_1}(\vartheta_0) \big)^{-1} (\mathrm{I}_{d_1\times d_1}, 0_{d_1\times(k-d_1)})^t \Big)^{-1}$$
$$\Leftrightarrow \inf_{\vartheta_1\in\mathbb{R}^{d_1}\setminus\{0\}} \mathrm{eff}_{T',T}(\vartheta^0,\vartheta_1) \geq 1.$$

Beweis: Dies folgt unmittelbar aus Satz 6.3.10, Korollar 6.3.11 und Lemma 4.5.2. □

6.3.15 Bemerkung: Es sei $k = l$ und $\psi_{t(\zeta_1)}(\vartheta_0)$ eine quadratische invertierbare Matrix. Zudem sei $A^{1,1}_{\zeta_1}(\vartheta_0) = i^{1,1}_{\zeta_1}(\vartheta_0) - i^{1,2}_{\zeta_1}(\vartheta_0) i^{2,2}_{\zeta_1}(\vartheta_0)^{-1} i^{2,1}_{\zeta_1}(\vartheta_0)$. Dann gilt die Äquivalenz

$$A^{1,1}_{\zeta_1'}(\vartheta_0') \leq \Big((\mathrm{I}_{d_1\times d_1}, 0_{d_1\times(k-d_1)}) \big(\psi_{t(\zeta_1)}(\vartheta_0)^t \mathrm{Cov}_{\vartheta_0}\big(t(\zeta_1)\big)^{-1} \psi_{t(\zeta_1)}(\vartheta_0) \big)^{-1} (\mathrm{I}_{d_1\times d_1}, 0_{d_1\times(k-d_1)})^t \Big)^{-1} \Leftrightarrow \sup_{\vartheta_1\in\mathbb{R}^{d_1}\setminus\{0\}} \mathrm{eff}_{T',T}(\vartheta^0,\vartheta_1) \leq 1,$$

äquivalent

$$A^{1,1}_{\zeta_1'}(\vartheta_0')^{-1} \geq (\mathrm{I}_{d_1\times d_1}, 0_{d_1\times(k-d_1)}) \psi_{t(\zeta_1)}(\vartheta_0)^{-1} \mathrm{Cov}_{\vartheta_0}\big(t(\zeta_1)\big) \big(\psi_{t(\zeta_1)}(\vartheta_0)^t\big)^{-1} (\mathrm{I}_{d_1\times d_1}, 0_{d_1\times(k-d_1)})^t \Leftrightarrow \sup_{\vartheta_1\in\mathbb{R}^{d_1}\setminus\{0\}} \mathrm{eff}_{T',T}(\vartheta^0,\vartheta_1) \leq 1.$$

Ist zusätzlich die Abbildung t durch die Identität gegeben, gilt die Äquivalenz

$$A^{1,1}_{\zeta_1'}(\vartheta_0') \leq \Big((\mathrm{I}_{d_1\times d_1}, 0_{d_1\times(k-d_1)}) \big(\psi_{\zeta_1}(\vartheta_0)^t \mathrm{Cov}_{\vartheta_0}(\zeta_1)^{-1} \psi_{\zeta_1}(\vartheta_0) \big)^{-1} (\mathrm{I}_{d_1\times d_1}, 0_{d_1\times(k-d_1)})^t \Big)^{-1}$$
$$\Leftrightarrow \sup_{\vartheta_1\in\mathbb{R}^{d_1}\setminus\{0\}} \mathrm{eff}_{T',T}(\vartheta^0,\vartheta_1) \leq 1,$$

äquivalent

$$A^{1,1}_{\zeta_1'}(\vartheta_0')^{-1} \geq (\mathrm{I}_{d_1\times d_1}, 0_{d_1\times(k-d_1)}) \psi_{\zeta_1}(\vartheta_0)^{-1} \mathrm{Cov}_{\vartheta_0}(\zeta_1) \big(\psi_{\zeta_1}(\vartheta_0)^t\big)^{-1} (\mathrm{I}_{d_1\times d_1}, 0_{d_1\times(k-d_1)})^t$$
$$\Leftrightarrow \sup_{\vartheta_1\in\mathbb{R}^{d_1}\setminus\{0\}} \mathrm{eff}_{T',T}(\vartheta^0,\vartheta_1) \leq 1.$$

Weiter ist

$$\tilde{\mathrm{eff}}_{T',T}(\vartheta^0) = \det A^{1,1}_{\zeta_1'}(\vartheta_0) \det\Big((\mathrm{I}_{d_1\times d_1}, 0_{d_1\times(k-d_1)}) \psi_{\zeta_1}(\vartheta_0)^{-1} \mathrm{Cov}_{\vartheta_0}(\zeta_1) \big(\psi_{\zeta_1}(\vartheta_0)^t\big)^{-1} (\mathrm{I}_{d_1\times d_1}, 0_{d_1\times(k-d_1)})^t \Big).$$

Ist weiter m eine gerade Zahl, $X_1 : (\Omega, \mathfrak{A}) \to (\mathbb{R}^{m/2}, \mathfrak{B}^{m/2})$, $Y_1 : (\Omega, \mathfrak{A}) \to (\mathbb{R}^{m/2}, \mathfrak{B}^{m/2})$, $\zeta_1 = X_1 - Y_1$ und $k = m/2$, haben wir die Äquivalenz

$$\begin{aligned} &A^{1,1}_{\zeta'_1}(\vartheta'_0) \leq \Bigg((\mathrm{I}_{d_1 \times d_1}, 0_{d_1 \times (k-d_1)}) \Big(\psi_{X_1 - Y_1}(\vartheta_0)^t \big(\mathrm{Cov}_{\vartheta^0}(X_1) + \mathrm{Cov}_{\vartheta^0}(Y_1) \\ &- \mathrm{Cov}_{\vartheta^0}(X_1, Y_1) - \mathrm{Cov}_{\vartheta^0}(X_1, Y_1)^t \big)^{-1} \psi_{X_1 - Y_1}(\vartheta_0) \Big)^{-1} (\mathrm{I}_{d_1 \times d_1}, 0_{d_1 \times (k-d_1)})^t \Bigg)^{-1} \\ &\Leftrightarrow \sup_{\vartheta_1 \in \mathbb{R}^{d_1} \setminus \{0\}} \mathrm{eff}_{T', T}(\vartheta^0, \vartheta_1) \leq 1, \end{aligned}$$

äquivalent

$$\begin{aligned} &A^{1,1}_{\zeta'_1}(\vartheta'_0)^{-1} \geq (\mathrm{I}_{d_1 \times d_1}, 0_{d_1 \times (k-d_1)}) \psi_{X_1 - Y_1}(\vartheta_0)^{-1} \big(\mathrm{Cov}_{\vartheta^0}(X_1) + \mathrm{Cov}_{\vartheta^0}(Y_1) \\ &- \mathrm{Cov}_{\vartheta^0}(X_1, Y_1) - \mathrm{Cov}_{\vartheta^0}(X_1, Y_1)^t \big) \big(\psi_{X_1 - Y_1}(\vartheta_0)^t \big)^{-1} (\mathrm{I}_{d_1 \times d_1}, 0_{d_1 \times (k-d_1)})^t \\ &\Leftrightarrow \sup_{\vartheta_1 \in \mathbb{R}^{d_1} \setminus \{0\}} \mathrm{eff}_{T', T}(\vartheta^0, \vartheta_1) \leq 1. \end{aligned}$$

ist $\psi_{\zeta_1}(\vartheta_0)$ eine quadratische invertierbare Matrix, gilt $\phi_{\zeta_1}(\vartheta_0) = (\mathrm{I}_{d_1 \times d_1}, 0_{d_1 \times (k-d_1)}) \psi_{\zeta_1}(\vartheta_0)^{-1}$. Dann sind die in Satz 6.2.2 entsprechend genannten Kriterien nicht nur hinreichend, sondern sogar notwendig.

6.3.16 Bemerkung: Es sei m eine gerade Zahl, $X_1 : (\Omega, \mathfrak{A}) \to (\mathbb{R}^{m/2}, \mathfrak{B}^{m/2})$ und $Y_1 : (\Omega, \mathfrak{A}) \to (\mathbb{R}^{m/2}, \mathfrak{B}^{m/2})$. Weiter sei $\zeta_1 = X_1 - Y_1$ und $\zeta'_1 = X'_1 - Y'_1$. Zudem sei die Effizienz gegeben durch

$$\mathrm{eff}_{T', T}(\vartheta^0, \vartheta_1) = \frac{\vartheta_1^t \big(i^{1,1}_{\zeta'_1}(\vartheta'_0) - i^{1,2}_{\zeta'_1}(\vartheta'_0) i^{2,2}_{\zeta'_1}(\vartheta'_0)^{-1} i^{2,1}_{\zeta'_1}(\vartheta'_0) \big) \vartheta_1}{\vartheta_1^t \big(i^{1,1}_{\zeta_1}(\vartheta_0) - i^{1,2}_{\zeta_1}(\vartheta_0) i^{2,2}_{\zeta_1}(\vartheta_0)^{-1} i^{2,1}_{\zeta_1}(\vartheta_0) \big) \vartheta_1}.$$

Es sei die Familie von Verteilungen $\{Q_\vartheta^{X_1 - Y_1}; \vartheta \in \tilde{\Theta}\}$ gemäß (6.3.2) eine Exponentialfamilie mit suffizienter Statistik t.

Ist die Familie von Verteilungen $\{M_{\vartheta'}^{X'_1 - Y'_1}; \vartheta' \in \tilde{\Theta}'\}$ gemäß (6.3.2) eine Exponentialfamilie mit suffizienter Statistik t', so kann unter analogen Voraussetzung analog zum Beweis von Satz 6.3.1 gezeigt werden, dass in diesem Fall die Effizienz die Gestalt

$$\begin{aligned} &\mathrm{eff}_{T', T}(\vartheta^0, \vartheta_1) \\ =&\vartheta_1^t \Big((\mathrm{I}_{d_1 \times d_1}, 0_{d_1 \times (k'-d_1)}) \big(\psi_{t'(X'_1 - Y'_1)}(\vartheta'_0)^t \mathrm{Cov}_{\vartheta'_0} \big(t'(X'_1 - Y'_1) \big)^{-1} \psi_{t'(X'_1 - Y'_1)}(\vartheta'_0) \big)^{-1} \\ &(\mathrm{I}_{d_1 \times d_1}, 0_{d_1 \times (k'-d_1)})^t \Big)^{-1} \vartheta_1 \\ &\Bigg(\vartheta_1^t \Big((\mathrm{I}_{d_1 \times d_1}, 0_{d_1 \times (k-d_1)}) \big(\psi_{t(X_1 - Y_1)}(\vartheta_0)^t \mathrm{Cov}_{\vartheta_0} \big(t(X_1 - Y_1) \big)^{-1} \psi_{t(X_1 - Y_1)}(\vartheta_0) \big)^{-1} \\ &(\mathrm{I}_{d_1 \times d_1}, 0_{d_1 \times (k-d_1)})^t \Big)^{-1} \vartheta_1 \Bigg)^{-1}. \end{aligned}$$

besitzt. Sind nun t' und t durch die Identität gegeben und gilt $k = k'$, so gilt wegen (6.1.3)

$$\mathrm{E}_{\vartheta_0}(X_1 - Y_1) = \mathrm{E}_{\vartheta_0'}(X_1' - Y_1'),$$

d.h. $\mathrm{E}_{\vartheta_0}(X_1 - Y_1) = \mathrm{E}_{\vartheta_0'}(X_1' - Y_1')$ hängt nur noch von ϑ_0^1 ab. Dann gilt sogar $\psi_{X_1-Y_1}(\vartheta_0) = \psi_{X_1'-Y_1'}(\vartheta_0')$. Unter diesen Voraussetzungen vereinfacht sich die Effizienz zu

$$\begin{aligned}
&\mathrm{eff}_{T',T}(\vartheta^0, \vartheta_1)\\
=&\vartheta_1^t\Big(\big(\mathrm{I}_{d_1\times d_1}, 0_{d_1\times(k-d_1)}\big)\big(\psi_{X_1-Y_1}(\vartheta_0)^t\mathrm{Cov}_{\vartheta_0'}(X_1' - Y_1')^{-1}\psi_{X_1-Y_1}(\vartheta_0)\big)^{-1}\\
&\big(\mathrm{I}_{d_1\times d_1}, 0_{d_1\times(k-d_1)}\big)^t\Big)^{-1}\vartheta_1\\
&\Big(\vartheta_1^t\Big(\big(\mathrm{I}_{d_1\times d_1}, 0_{d_1\times(k-d_1)}\big)\big(\psi_{X_1-Y_1}(\vartheta_0)^t\mathrm{Cov}_{\vartheta_0}(X_1 - Y_1)^{-1}\psi_{X_1-Y_1}(\vartheta_0)\big)^{-1}\\
&\big(\mathrm{I}_{d_1\times d_1}, 0_{d_1\times(k-d_1)}\big)^t\Big)^{-1}\vartheta_1\Big)^{-1}.
\end{aligned}$$

Wegen der positiven Definitheit der beteiligten Matrizen gilt

$$\begin{aligned}
&\Big(\big(\mathrm{I}_{d_1\times d_1}, 0_{d_1\times(k-d_1)}\big)\big(\psi_{X_1-Y_1}(\vartheta_0)^t\mathrm{Cov}_{\vartheta_0'}(X_1' - Y_1')^{-1}\psi_{X_1-Y_1}(\vartheta_0)\big)^{-1}\\
&\big(\mathrm{I}_{d_1\times d_1}, 0_{d_1\times(k-d_1)}\big)^t\Big)^{-1}\\
\begin{Bmatrix}\geq\\ \leq\end{Bmatrix}&\Big(\big(\mathrm{I}_{d_1\times d_1}, 0_{d_1\times(k-d_1)}\big)\big(\psi_{X_1-Y_1}(\vartheta_0)^t\mathrm{Cov}_{\vartheta_0}(X_1 - Y_1)^{-1}\psi_{X_1-Y_1}(\vartheta_0)\big)^{-1}\\
&\big(\mathrm{I}_{d_1\times d_1}, 0_{d_1\times(k-d_1)}\big)^t\Big)^{-1}\\
\Leftrightarrow\ &\big(\mathrm{I}_{d_1\times d_1}, 0_{d_1\times(k-d_1)}\big)\big(\psi_{X_1-Y_1}(\vartheta_0)^t\mathrm{Cov}_{\vartheta_0'}(X_1' - Y_1')^{-1}\psi_{X_1-Y_1}(\vartheta_0)\big)^{-1}\\
&\big(\mathrm{I}_{d_1\times d_1}, 0_{d_1\times(k-d_1)}\big)^t\\
\begin{Bmatrix}\leq\\ \geq\end{Bmatrix}&\big(\mathrm{I}_{d_1\times d_1}, 0_{d_1\times(k-d_1)}\big)\big(\psi_{X_1-Y_1}(\vartheta_0)^t\mathrm{Cov}_{\vartheta_0}(X_1 - Y_1)^{-1}\psi_{X_1-Y_1}(\vartheta_0)\big)^{-1}\\
&\big(\mathrm{I}_{d_1\times d_1}, 0_{d_1\times(k-d_1)}\big)^t.
\end{aligned}$$

Es ist

$$\begin{aligned}
&\big(\psi_{X_1-Y_1}(\vartheta_0)^t\mathrm{Cov}_{\vartheta_0'}(X_1' - Y_1')^{-1}\psi_{X_1-Y_1}(\vartheta_0)\big)^{-1}\\
\begin{Bmatrix}\leq\\ \geq\end{Bmatrix}&\big(\psi_{X_1-Y_1}(\vartheta_0)^t\mathrm{Cov}_{\vartheta_0}(X_1 - Y_1)^{-1}\psi_{X_1-Y_1}(\vartheta_0)\big)^{-1}\\
\Rightarrow\ &\big(\mathrm{I}_{d_1\times d_1}, 0_{d_1\times(k-d_1)}\big)\big(\psi_{X_1-Y_1}(\vartheta_0)^t\mathrm{Cov}_{\vartheta_0'}(X_1' - Y_1')^{-1}\psi_{X_1-Y_1}(\vartheta_0)\big)^{-1}\\
&\big(\mathrm{I}_{d_1\times d_1}, 0_{d_1\times(k-d_1)}\big)^t\\
\begin{Bmatrix}\leq\\ \geq\end{Bmatrix}&\big(\mathrm{I}_{d_1\times d_1}, 0_{d_1\times(k-d_1)}\big)\big(\psi_{X_1-Y_1}(\vartheta_0)^t\mathrm{Cov}_{\vartheta_0}(X_1 - Y_1)^{-1}\psi_{X_1-Y_1}(\vartheta_0)\big)^{-1}\\
&\big(\mathrm{I}_{d_1\times d_1}, 0_{d_1\times(k-d_1)}\big)^t.
\end{aligned}$$

Weiter ist

$$
\begin{aligned}
&\big(\psi_{X_1-Y_1}(\vartheta_0)^t \mathrm{Cov}_{\vartheta_0'}(X_1'-Y_1')^{-1}\psi_{X_1-Y_1}(\vartheta_0)\big)^{-1}\\
\begin{Bmatrix}\le\\ \ge\end{Bmatrix}&\big(\psi_{X_1-Y_1}(\vartheta_0)^t \mathrm{Cov}_{\vartheta_0}(X_1-Y_1)^{-1}\psi_{X_1-Y_1}(\vartheta_0)\big)^{-1}\\
\Leftrightarrow\ &\psi_{X_1-Y_1}(\vartheta_0)^t \mathrm{Cov}_{\vartheta_0'}(X_1'-Y_1')^{-1}\psi_{X_1-Y_1}(\vartheta_0)\\
\begin{Bmatrix}\ge\\ \le\end{Bmatrix}&\psi_{X_1-Y_1}(\vartheta_0)^t \mathrm{Cov}_{\vartheta_0}(X_1-Y_1)^{-1}\psi_{X_1-Y_1}(\vartheta_0).
\end{aligned}
$$

Zudem gilt

$$
\begin{aligned}
&\mathrm{Cov}_{\vartheta_0'}(X_1'-Y_1')^{-1} \begin{Bmatrix}\ge\\ \le\end{Bmatrix} \mathrm{Cov}_{\vartheta_0}(X_1-Y_1)^{-1}\\
\Rightarrow\ &\psi_{X_1-Y_1}(\vartheta_0)^t \mathrm{Cov}_{\vartheta_0'}(X_1'-Y_1')^{-1}\psi_{X_1-Y_1}(\vartheta_0)\\
\begin{Bmatrix}\ge\\ \le\end{Bmatrix}&\psi_{X_1-Y_1}(\vartheta_0)^t \mathrm{Cov}_{\vartheta_0}(X_1-Y_1)^{-1}\psi_{X_1-Y_1}(\vartheta_0).
\end{aligned}
$$

Dann gilt die Implikation

$$
\mathrm{Cov}_{\vartheta_0}(X_1-Y_1) \le \mathrm{Cov}_{\vartheta_0'}(X_1'-Y_1') \ \Rightarrow\ \sup_{\vartheta_1\in\mathbb{R}^{d_1}\setminus\{0\}} \mathrm{eff}_{T',T}(\vartheta^0,\vartheta_1) \le 1,
$$

und die Implikation

$$
\mathrm{Cov}_{\vartheta_0}(X_1-Y_1) \ge \mathrm{Cov}_{\vartheta_0'}(X_1'-Y_1') \ \Rightarrow\ \inf_{\vartheta_1\in\mathbb{R}^{d_1}\setminus\{0\}} \mathrm{eff}_{T',T}(\vartheta^0,\vartheta_1) \ge 1.
$$

Es ist

$$
\mathrm{Cov}_{\vartheta_0}(X_1-Y_1) = \mathrm{Cov}_{\vartheta^0}(X_1) + \mathrm{Cov}_{\vartheta^0}(Y_1) - \mathrm{Cov}_{\vartheta^0}(X_1,Y_1) - \mathrm{Cov}_{\vartheta^0}(X_1,Y_1)^t,
$$

und wegen (6.1.2) und (6.1.3)

$$
\begin{aligned}
\mathrm{Cov}_{\vartheta_0'}(X_1'-Y_1') &= \mathrm{Cov}_{\vartheta^0}(X_1') + \mathrm{Cov}_{\vartheta^0}(Y_1') - \mathrm{Cov}_{\vartheta^0}(X_1',Y_1') - \mathrm{Cov}_{\vartheta^0}(X_1',Y_1')^t\\
&= \mathrm{Cov}_{\vartheta^0}(X_1') + \mathrm{Cov}_{\vartheta^0}(Y_1') = \mathrm{Cov}_{\vartheta^0}(X_1) + \mathrm{Cov}_{\vartheta^0}(Y_1).
\end{aligned}
$$

Daher gelten die Implikationen

$$
\mathrm{Cov}_{\vartheta^0}(X_1,Y_1) + \mathrm{Cov}_{\vartheta^0}(X_1,Y_1)^t \ge 0 \ \Rightarrow\ \sup_{\vartheta_1\in\mathbb{R}^{d_1}\setminus\{0\}} \mathrm{eff}_{T',T}(\vartheta^0,\vartheta_1) \le 1
$$

und

$$
\mathrm{Cov}_{\vartheta^0}(X_1,Y_1) + \mathrm{Cov}_{\vartheta^0}(X_1,Y_1)^t \le 0 \ \Rightarrow\ \inf_{\vartheta_1\in\mathbb{R}^{d_1}\setminus\{0\}} \mathrm{eff}_{T',T}(\vartheta^0,\vartheta_1) \ge 1.
$$

Wegen Lemma 4.5.4 gelten dann auch die Implikationen

$$
\mathrm{Cov}_{\vartheta^0}(X_1,Y_1) + \mathrm{Cov}_{\vartheta^0}(X_1,Y_1)^t \ge 0 \ \Rightarrow\ \widetilde{\mathrm{eff}}_{T',T}(\vartheta^0) \le 1
$$

und

$$
\mathrm{Cov}_{\vartheta^0}(X_1,Y_1) + \mathrm{Cov}_{\vartheta^0}(X_1,Y_1)^t \le 0 \ \Rightarrow\ \widetilde{\mathrm{eff}}_{T',T}(\vartheta^0) \ge 1.
$$

6.3.17 Beispiel: Betrachten wir die Situation aus Beispiel 6.3.12: Es sei m eine gerade Zahl und $X_1 : (\Omega, \mathfrak{A}) \to (\mathbb{R}^{m/2}, \mathfrak{B}^{m/2})$, $Y_1 : (\Omega, \mathfrak{A}) \to (\mathbb{R}^{m/2}, \mathfrak{B}^{m/2})$. Weiter sei $\zeta_1 = X_1 - Y_1$, $\zeta_1' = X_1' - Y_1'$, $k = k'$, $\tilde{\Theta} = \tilde{\Theta}'$ und

$$\left(\begin{smallmatrix} \vartheta_1 \\ \vartheta_2 \end{smallmatrix}\right) = \vartheta = \mathrm{E}_\vartheta(X_1 - Y_1) = \left(\begin{smallmatrix} \vartheta_1 \\ \vartheta_2' \end{smallmatrix}\right) = \vartheta' = \mathrm{E}_{\vartheta'}(X_1' - Y_1'),\ (\vartheta_1, \vartheta_2) \in \tilde{\Theta}.$$

Wir betrachten für das gegebene $\vartheta_0^1 \in \Theta^1$ das Testproblem

$$\mathrm{H}:\ \vartheta_1 = \vartheta_0^1,\ \mathrm{K}:\ \vartheta_1 \in \Theta^1 \setminus \{\vartheta_0^1\}.$$

Die Effizienz sei gegeben durch

$$\mathrm{eff}_{T',T}(\vartheta^0, \vartheta_1) = \frac{\vartheta_1^t \big(i_{\zeta_1'}^{1,1}(\vartheta_0') - i_{\zeta_1'}^{1,2}(\vartheta_0') i_{\zeta_1'}^{2,2}(\vartheta_0')^{-1} i_{\zeta_1'}^{2,1}(\vartheta_0')\big)\vartheta_1}{\vartheta_1^t \big(i_{\zeta_1}^{1,1}(\vartheta_0) - i_{\zeta_1}^{1,2}(\vartheta_0) i_{\zeta_1}^{2,2}(\vartheta_0)^{-1} i_{\zeta_1}^{2,1}(\vartheta_0)\big)\vartheta_1}.$$

Dann ist $\psi_{\zeta_1}(\vartheta_0) = \psi_{\zeta_1'}(\vartheta_0')$ gegeben durch die Einheitsmatrix.

Es sei die Familie von Verteilungen $\{Q_\vartheta^{X_1 - Y_1}; \vartheta \in \tilde{\Theta}\}$ gemäß (6.3.2) eine Exponentialfamilie mit suffizienter Statistik t. Zudem sei die Abbildung t durch die Identität gegeben.

Weiter sei die Familie von Verteilungen $\{M_{\vartheta'}^{X_1' - Y_1'}; \vartheta' \in \tilde{\Theta}'\}$ gemäß (6.3.2) eine Exponentialfamilie mit suffizienter Statistik t'. Zudem sei die Abbildung t' durch die Identität gegeben. Dann gelten die Implikationen

$$\mathrm{Cov}_{\vartheta^0}(X_1, Y_1) + \mathrm{Cov}_{\vartheta^0}(X_1, Y_1)^t \geq 0 \ \Rightarrow \sup_{\vartheta_1 \in \mathbb{R}^{d_1} \setminus \{0\}} \mathrm{eff}_{T',T}(\vartheta^0, \vartheta_1) \leq 1$$

und

$$\mathrm{Cov}_{\vartheta^0}(X_1, Y_1) + \mathrm{Cov}_{\vartheta^0}(X_1, Y_1)^t \leq 0 \ \Rightarrow \inf_{\vartheta_1 \in \mathbb{R}^{d_1} \setminus \{0\}} \mathrm{eff}_{T',T}(\vartheta^0, \vartheta_1) \geq 1,$$

sowie die Implikationen

$$\mathrm{Cov}_{\vartheta^0}(X_1, Y_1) + \mathrm{Cov}_{\vartheta^0}(X_1, Y_1)^t \geq 0 \ \Rightarrow \ \tilde{\mathrm{eff}}_{T',T}(\vartheta^0) \leq 1$$

und

$$\mathrm{Cov}_{\vartheta^0}(X_1, Y_1) + \mathrm{Cov}_{\vartheta^0}(X_1, Y_1)^t \leq 0 \ \Rightarrow \ \tilde{\mathrm{eff}}_{T',T}(\vartheta^0) \geq 1.$$

Konkret kann z.B. der Fall betrachtet werden, dass

$$\begin{pmatrix} X_1 \\ Y_1 \end{pmatrix} \sim \mathrm{N}_{2k}\left(\begin{pmatrix} a \\ b \end{pmatrix}, \begin{pmatrix} \Sigma & \Pi \\ \Pi^t & \Delta \end{pmatrix} \right)$$

gilt, wobei $a, b \in \mathbb{R}^k$ unbekannt und $\Sigma, \Delta, \Pi \in \mathbb{R}^{k \times k}$ bekannt seien, und Σ, Δ und die zugrunde liegende Kovarianzmatrix von $\left(\begin{smallmatrix} X_1 \\ Y_1 \end{smallmatrix}\right)$ positiv definit seien. Dann gilt wegen (6.1.2) und (6.1.3) auch

$$\begin{pmatrix} X_1' \\ Y_1' \end{pmatrix} \sim \mathrm{N}_{2k}\left(\begin{pmatrix} a \\ b \end{pmatrix}, \begin{pmatrix} \Sigma & 0 \\ 0 & \Delta \end{pmatrix} \right).$$

Das in diesem konkreten Fall alle oben genannten Voraussetzungen erfüllt sind, kann den Ausführungen aus Beispiel 6.3.12 entnommen werden.

6.3.18 Bemerkung: Die in Bemerkung 6.3.16 und Beispiel 6.3.17 erarbeiteten Kriterien sind unter den genannten Voraussetzungen auch hinreichende Kriterien für $\widetilde{\mathrm{eff}}_{T',T}(\vartheta^0) \leq 1$ bzw. $\widetilde{\mathrm{eff}}_{T',T}(\vartheta^0) \geq 1$. Beachte hierzu Lemma 4.5.4.

7 Vergleiche bei Schätzverfahren

7.1 Einleitung

Wie wir am Anfang der Arbeit ausgeführt haben, kann der Vergleich zweier Experimente, z. B. von unabhängiger und verbundener Stichprobenerhebung, nicht nur anhand der Qualität von verschiedenen Testverfahren, sondern auch anhand der Qualität verschiedener Punktschätzverfahren oder der Qualität verschiedener Konfidenzbereichsschätzverfahren vorgenommen werden. Diese beiden Ansätze wollen wir in diesem Abschnitt genauer betrachten.

Die Ausführungen in den Abschnitten 5 und 6 beziehen sich auf Effizienzen statistischer Tests, wie sie unter bestimmten Voraussetzungen vorkommen, wenn bei beiden Experimenten ein optimaler Test verwendet wird; siehe Abschnitt 4. In der Tat erhält man unter bestimmten Voraussetzungen ähnliche Effizienzen für Schätzer, wenn man optimale Schätzer betrachtet. Genauso erhält man unter bestimmten Voraussetzungen die gleichen Effizienzen für Konfidenzbereichsschätzer, wenn man optimale Konfidenzbereichsschätzer betrachtet.

Insofern lassen sich die Resultate aus den Abschnitten 5 und 6 auch anwenden, wenn eine Fragestellung bezüglich eines Punkt- oder Konfidenzbereichsschätzverfahrens vorliegt und der Vergleich zweier Experimente anhand der Qualität verschiedener Punktschätz- bzw. Konfidenzbereichsschätzverfahren vorgenommen wird.

7.2 Punktschätzverfahren

Es sei für ein $d \in \mathbb{N}$ $\emptyset \neq \Theta \subset \mathbb{R}^d$. Auf dem zugrunde liegenden statistischen Raum $(\Omega, \mathfrak{A}, \mathfrak{P})$, mit $\mathfrak{P} = \{P_\vartheta; \vartheta \in \Theta\}$, sei eine Folge von Schätzern $(t_n)_{n\in\mathbb{N}}$ für den unbekannten Parameter $\vartheta \in \Theta$ mit

$$t_n : (\Omega, \mathfrak{A}) \to (\mathbb{R}^d, \mathfrak{B}^d)$$

gegeben. Wir bezeichnen mit $t := (t_n)_{n\in\mathbb{N}}$ die Folge von Schätzern für ϑ.

Bei der Definition der relativen Effizienz zweier Schätzer orientieren wir uns an van der Vaart ([38]). Wir betrachten zunächst nur den Fall $d = 1$. Wir nehmen an, dass die Folge von Schätzern für ϑ

$$\sqrt{n}(t_n - \vartheta) \xrightarrow{\mathrm{v}} \mathrm{N}\big(0, \sigma_t^2(\vartheta)\big) \text{ für } n \to \infty$$

erfülle, mit einem $\sigma_t^2(\vartheta) > 0$. Es sei eine Sequenz natürlicher Zahlen $(N_{t,i})_{i\in\mathbb{N}}$ gegeben mit $\lim_{i\to\infty} N_{t,i} = \infty$ und

$$\sqrt{i}(t_{N_{t,i}} - \vartheta) \xrightarrow{\mathrm{v}} \mathrm{N}(0, 1) \text{ für } i \to \infty.$$

Sind dann t_n und t'_n zwei Schätzer mit den genannten Eigenschaften, so heißt der Grenzwert

$$\tilde{\tilde{\mathrm{eff}}}_{t',t} := \tilde{\tilde{\mathrm{eff}}}_{t',t}(\vartheta) := \lim_{i\to\infty} N_{t',i}/N_{t,i}$$

relative Effizienz der Schätzer t und t' an der Stelle ϑ. Dieser Grenzwert existiert, wie man wie folgt sieht:

Es ist

$$\frac{\sqrt{i}}{\sqrt{N_{t,i}}} \underbrace{\sqrt{N_{t,i}}(t_{N_{t,i}} - \vartheta)}_{\xrightarrow{\text{v}} \mathrm{N}(0,\sigma_t^2(\vartheta))} = \sqrt{i}(t_{N_{t,i}} - \vartheta) \xrightarrow{\text{v}} \mathrm{N}(0,1) \text{ für } i \to \infty,$$

woraus $\lim_{i\to\infty} \sqrt{i}/\sqrt{N_{t,i}} = 1/\sigma_t(\vartheta)$, mit $\sigma_t(\vartheta) := \sqrt{\sigma_t^2(\vartheta)}$, folgt. Analog kann für die Folge von Schätzern t' argumentiert werden. Daher gilt

$$\lim_{i\to\infty} \frac{N_{t',i}}{N_{t,i}} = \lim_{i\to\infty} \left(\frac{\frac{\sqrt{i}}{\sqrt{N_{t,i}}}}{\frac{\sqrt{i}}{\sqrt{N_{t',i}}}} \right)^2 = \frac{\sigma_t^2(\vartheta)}{\sigma_{t'}^2(\vartheta)} = \tilde{\tilde{\mathrm{eff}}}_{t',t}(\vartheta).$$

Die relative Effizienz an der Stelle ϑ ist also durch das Verhältnis der Varianzen der Grenzverteilung gegeben, was sehr plausibel ist.

Wollen wir nun für $d > 1$ eine Effizienz in dieser Art und Weise definieren, so stellen wir fest, dass sich dieses Konzept der relativen Effizienz nicht automatisch auf den mehrdimensionale Fall übertragen lässt. Dies liegt daran, dass im Allgemeinen für $d > 1$ keine Sequenz natürlicher Zahlen $(N_{t,i})_{i\in\mathbb{N}}$ existieren wird, sodass $\lim_{i\to\infty} N_{t,i} = \infty$ und $\sqrt{i}(t_{N_{t,i}} - \vartheta) \xrightarrow{\text{v}} \mathrm{N}_d(0, \mathrm{I}_d)$ für $i \to \infty$ gilt.

Dennoch lässt sich dieses Konzept auf höherdimensionale Situationen verallgemeinern. Es sei $d \geq 1$. Wir nehmen an, dass

$$\sqrt{n}(t_n - \vartheta) \xrightarrow{\text{v}} \mathrm{N}_d\big(0, \Sigma_t(\vartheta)\big) \text{ für } n \to \infty$$

gilt, mit einer symmetrischen und positiv definiten Matrix $\Sigma_t(\vartheta) \in \mathbb{R}^{d\times d}$.

Nun besteht die Möglichkeit, die relative Effizienz zweier Folgen von Schätzern t und t' mit dieser Eigenschaft als das Verhältnis der Determinanten der zugehörigen Kovarianzmatrizen zu definieren, also

$$\tilde{\mathrm{eff}}_{t',t} := \tilde{\mathrm{eff}}_{t',t}(\vartheta) := \frac{\det \Sigma_t(\vartheta)}{\det \Sigma_{t'}(\vartheta)}$$

zu setzen; siehe z.B. Serfling ([36]). Dann sind für $d = 1$ die Definitionen von $\tilde{\mathrm{eff}}_{t',t}$ und $\tilde{\tilde{\mathrm{eff}}}_{t',t}$ äquivalent.

Wir wollen hier zusätzlich ein anderes Vorgehen studieren. Dazu sei für alle $\vartheta_1 \in \mathbb{R}^d \setminus \{0\}$ eine Sequenz natürlicher Zahlen $(N_{t,i}(\vartheta_1))_{i\in\mathbb{N}}$ gegeben mit

$$\sqrt{i}(\vartheta_1^t t_{N_{t,i}(\vartheta_1)} - \vartheta_1^t \vartheta) \xrightarrow{\mathrm{v}} \mathrm{N}(0,1) \text{ für } i \to \infty.$$

Sind dann t_n und t'_n zwei Schätzer für ϑ mit den genannten Eigenschaften, so heißt der Grenzwert

$$\mathrm{eff}_{t',t} := \mathrm{eff}_{t',t}(\vartheta, \vartheta_1) := \lim_{i\to\infty} N_{t',i}(\vartheta_1)/N_{t,i}(\vartheta_1)$$

relative Effizienz der Schätzer t und t'. Dieser Grenzwert existiert, wie man wie folgt sieht:

Es gilt

$$\frac{\sqrt{i}}{\sqrt{N_{t,i}(\vartheta_1)}} \underbrace{\sqrt{N_{t,i}(\vartheta_1)}(\vartheta_1^t t_{N_{t,i}(\vartheta_1)} - \vartheta_1^t \vartheta)}_{\xrightarrow{\mathrm{v}} \mathrm{N}(0, \vartheta_1^t \Sigma_t(\vartheta)\vartheta_1)} = \sqrt{i}(\vartheta_1^t t_{N_{t,i}(\vartheta_1)} - \vartheta_1^t \vartheta) \xrightarrow{\mathrm{v}} \mathrm{N}(0,1) \text{ für } i \to \infty,$$

woraus $\lim_{i\to\infty} \sqrt{i}/\sqrt{N_{t,i}(\vartheta_1)} = 1/\sqrt{\vartheta_1^t \Sigma_t(\vartheta)\vartheta_1}$ folgt. Analog kann für die Folge von Schätzern t' argumentiert werden. Daher erhält man

$$\mathrm{eff}_{t',t}(\vartheta, \vartheta_1) = \lim_{i\to\infty} \frac{N_{t',i}(\vartheta_1)}{N_{t,i}(\vartheta_1)} = \lim_{i\to\infty} \left(\frac{\frac{\sqrt{i}}{\sqrt{N_{t,i}(\vartheta_1)}}}{\frac{\sqrt{i}}{\sqrt{N_{t',i}(\vartheta_1)}}} \right)^2 = \frac{\vartheta_1^t \Sigma_t(\vartheta)\vartheta_1}{\vartheta_1^t \Sigma_{t'}(\vartheta)\vartheta_1}.$$

Dann sind für $d = 1$ die Definitionen von $\widetilde{\mathrm{eff}}_{t',t}$, $\tilde{\tilde{\mathrm{eff}}}_{t',t}$ und $\mathrm{eff}_{t',t}$ äquivalent. Für $d \geq 1$ gibt die folgende Bemerkung einen Zusammenhang zwischen den Effizienzbegriffen $\widetilde{\mathrm{eff}}_{t',t}$ und $\mathrm{eff}_{t',t}$ an.

7.2.1 Bemerkung: Es gelten folgende Implikationen:

a) $\sup_{\vartheta_1 \in \mathbb{R}^d\setminus\{0\}} \mathrm{eff}_{t',t}(\vartheta, \vartheta_1) < 1 \ \Rightarrow\ \widetilde{\mathrm{eff}}_{t',t}(\vartheta) < 1$.

b) $\sup_{\vartheta_1 \in \mathbb{R}^d\setminus\{0\}} \mathrm{eff}_{t',t}(\vartheta, \vartheta_1) \leq 1 \ \Rightarrow\ \widetilde{\mathrm{eff}}_{t',t}(\vartheta) \leq 1$.

c) $\mathrm{eff}_{t',t}(\vartheta, \vartheta_1) = 1 \ \forall\ \vartheta_1 \in \mathbb{R}^d \setminus \{0\} \ \Rightarrow\ \widetilde{\mathrm{eff}}_{t',t}(\vartheta) = 1$.

d) $\inf_{\vartheta_1 \in \mathbb{R}^d\setminus\{0\}} \mathrm{eff}_{t',t}(\vartheta, \vartheta_1) \geq 1 \ \Rightarrow\ \widetilde{\mathrm{eff}}_{t',t}(\vartheta) \geq 1$.

e) $\inf_{\vartheta_1 \in \mathbb{R}^d\setminus\{0\}} \mathrm{eff}_{t',t}(\vartheta, \vartheta_1) > 1 \ \Rightarrow\ \widetilde{\mathrm{eff}}_{t',t}(\vartheta) > 1$.

Dies folgt analog zum Beweis von Lemma 4.5.4.

Wie in Abschnitt 6 gehen wir nun von einer bestimmten Modellannahme aus, die ein

Spezialfall der bereits in Unterabschnitt 4.1 getroffenen Modellannahme darstellt. Es handelt sich um die Situation aus Bemerkung 4.2.8 und Unterabschnitt 4.4, d.h. es liegt das Modell aus Unterabschnitt 4.1 vor, wobei für ein gegebenes $\vartheta_0^1 \in \Theta^1$ $\{\vartheta_0^1\} = \Theta_0^1$ ist. Wir nehmen an, dass die Mengen $\tilde{\Theta}$ und $\tilde{\Theta}'$ offen und konvex sind.

Im Folgenden wollen wir Annahmen an die Zufallsvariablen ζ_1 und ζ_1' bei den zugrunde liegenden injektiven Parametrisierungen durch $\tilde{\Theta}$ und $\tilde{\Theta}'$ formulieren. Zudem werden wir zugehörige Schätzer t_n und t_n' basierend auf $\zeta_1, \ldots, \zeta_n$ und $\zeta_1', \ldots, \zeta_n'$ einführen und diese ebenfalls mit Bedingungen versehen. Der Einfachheit halber führen wir dies nur für ζ_1, $\tilde{\Theta}$ und t_n aus; in analoger Weise sollen alle Definitionen und Annahmen auch für ζ_1', $\tilde{\Theta}'$, und t_n' gelten. Dabei werden die im Zusammenhang mit ζ_1, $\tilde{\Theta}$ und t_n eingeführten Größen im Zusammenhang mit ζ_1', $\tilde{\Theta}'$, und t_n' mit einem Strich versehen.

Es sei $\vartheta \in \tilde{\Theta}$ und $t = (t_n)_{n\in\mathbb{N}}$ eine Folge von Schätzern für den unbekannten Parameter ϑ basierend auf der Stichprobe $\zeta_1, \ldots, \zeta_n$, d.h. es ist $h_n : (R^n, \mathfrak{S}^n) \to (\mathbb{R}^k, \mathfrak{B}^k)$ für jedes $n \in \mathbb{N}$ und $t_n = h_n(\zeta_1, \ldots, \zeta_n)$.

Wir nehmen an, dass ein σ-endliches Maß μ auf $(R, \mathfrak{S})$ existiert, sodass für jedes $\vartheta \in \tilde{\Theta}$ $Q_\vartheta^{\zeta_1} \ll \mu$ gilt und $f(\cdot;\vartheta)$ eine μ-Dichte von $Q_\vartheta^{\zeta_1}$ ist. Wir setzen generell voraus, dass $f(z;\vartheta) > 0$ für alle $\vartheta \in \tilde{\Theta}$ und alle $z \in R$. Es sei die Abbildung $\vartheta \mapsto f(z;\vartheta)$, $\vartheta \in \tilde{\Theta}$, für alle $z \in R$ stetig differenzierbar. Zudem nehmen wir an, dass für alle $\vartheta \in \tilde{\Theta}$ die Fisher-Informationsmatrix aus $Q_\vartheta^{\zeta_1}$, $i_{\zeta_1}(\vartheta)$, existiert, endliche Einträge besitzt und positiv definit ist. Mit der am Anfang von Unterabschnitt 4.2 eingeführten Notation und $\vartheta = (\vartheta_1, \ldots, \vartheta_k)^t \in \tilde{\Theta}$ ist dann

$$i_{\zeta_1}(\vartheta) = \left(\mathrm{E}_\vartheta \left(\frac{\partial}{\partial \vartheta_i} L(\zeta_1, \vartheta) \frac{\partial}{\partial \vartheta_j} L(\zeta_1, \vartheta) \right) \right)_{\substack{1\le i\le k\\ 1\le j\le k}}, \ \vartheta \in \tilde{\Theta},$$

Dann sei

$$i_{\zeta_1}(\vartheta) = \begin{pmatrix} i_{\zeta_1}^{1,1}(\vartheta) & i_{\zeta_1}^{1,2}(\vartheta) \\ i_{\zeta_1}^{2,1}(\vartheta) & i_{\zeta_1}^{2,2}(\vartheta) \end{pmatrix}, \ \vartheta \in \tilde{\Theta},$$

mit $i_{\zeta_1}^{1,1}(\vartheta) \in \mathbb{R}^{d_1\times d_1}$, $i_{\zeta_1}^{1,2}(\vartheta) \in \mathbb{R}^{d_1\times(k-d_1)}$, $i_{\zeta_1}^{2,1}(\vartheta) \in \mathbb{R}^{(k-d_1)\times d_1}$, $i_{\zeta_1}^{2,2}(\vartheta) \in \mathbb{R}^{(k-d_1)\times(k-d_1)}$. D.h. $i_{\zeta_1}^{1,1}(\vartheta)$ ist die Informationsmatrix aus ζ_1, wenn für $\vartheta = (\vartheta^1, \vartheta_2) \in \tilde{\Theta}$ der Parameter ϑ^1 als einziger variabler Parameter, und ϑ_2 als fest und gegeben betrachtet wird.

Zudem existiere für alle $n \in \mathbb{N}$ und alle $\vartheta \in \tilde{\Theta}$ $\mathrm{E}_\vartheta(t_n)$, besitze endliche Einträge und es gelte für alle $\vartheta = (\vartheta_1, \ldots, \vartheta_k)^t \in \tilde{\Theta}$

$$\frac{\partial}{\partial \vartheta_i} \int f(z;\vartheta) \mathrm{d}\mu(z) = \int \frac{\partial}{\partial \vartheta_i} f(z;\vartheta) \mathrm{d}\mu(z) = 0, \ i = 1, \ldots, k,$$

sowie für alle $\vartheta = (\vartheta_1, \dots, \vartheta_k)^t \in \tilde{\Theta}$ und alle $n \in \mathbb{N}$

$$\frac{\partial}{\partial \vartheta_i} \int h_n(z_1, \dots, z_n) f(z_1; \vartheta) \cdots f(z_n; \vartheta) \mathrm{d}\mu(z_1) \dots \mathrm{d}\mu(z_n)$$
$$= \int h_n(z_1, \dots, z_n) \frac{\partial}{\partial \vartheta_i} (f(z_1; \vartheta) \cdots f(z_n; \vartheta)) \mathrm{d}\mu(z_1) \dots \mathrm{d}\mu(z_n), \; i = 1, \dots, k.$$

Dann sei mit $t_n = (t_{n,1}, \dots, t_{n,k})^t$

$$\psi_{t_n}(\vartheta) := \Big(\frac{\partial}{\partial \vartheta_j} \mathrm{E}_{\vartheta_0}(t_{n,i}) \Big)_{\substack{1 \le i \le k \\ 1 \le j \le k}}, \; \vartheta \in \tilde{\Theta}.$$

Für alle $n \in \mathbb{N}$ existiere $\mathrm{Cov}_\vartheta(t_n)$ und besitze endliche Einträge. Dann liefert die Cramér-Rao-Ungleichung

$$\mathrm{Cov}_\vartheta(t_n) \ge \frac{1}{n} \psi_{t_n}(\vartheta) i_{\zeta_1}(\vartheta)^{-1} \psi_{t_n}(\vartheta)^t, \; \vartheta \in \tilde{\Theta}, \; n \in \mathbb{N},$$

wobei hier Ungleichungen immer im Sinn der Definitheit der Differenz der Matrizen zu lesen sind; siehe hierzu Unterabschnitt 6.1. Gilt speziell $\mathrm{E}_\vartheta(t_n) = \vartheta$, $\vartheta \in \tilde{\Theta}$, so folgt $\psi_{t_n}(\vartheta) = \mathrm{I}_{k \times k}, \; n \in \mathbb{N}$, und damit

$$\mathrm{Cov}_\vartheta(t_n) \ge \frac{1}{n} i_{\zeta_1}(\vartheta)^{-1}, \; n \in \mathbb{N}.$$

In diesem Sinn kann der Schätzer t_n aufgrund seiner Eigenschaft

$$\sqrt{n}(t_n - \vartheta) \xrightarrow{\mathrm{v}} \mathrm{N}_k\big(0, \Sigma_t(\vartheta)\big) \text{ für } n \to \infty$$

asymptotisch optimal bei $\vartheta \in \tilde{\Theta}$ genannt werden, wenn

$$\Sigma_t(\vartheta) = i_{\zeta_1}(\vartheta)^{-1} \tag{7.2.1}$$

gilt. Siehe hierzu z.B. [38]. Im Folgenden wollen wir ausschließlich solche Schätzer t_n betrachten.

Es sei für $l \in \mathbb{N}$ eine Abbildung $p : (\mathbb{R}^k, \mathfrak{B}^k) \to (\mathbb{R}^l, \mathfrak{B}^l)$ gegeben. Dann ist $p(t_n)$ ein Schätzer für $p(\vartheta)$. Ist die Abbildung p in $\vartheta \in \tilde{\Theta}$ differenzierbar mit zugehöriger Jacobimatrix $\mathcal{J}_p(\vartheta)$, so liefert das Fehlerfortpflanzungsgesetz

$$\sqrt{n}\big(p(t_n) - p(\vartheta)\big) \xrightarrow{\mathrm{v}} \mathrm{N}_l\big(0, \mathcal{J}_p(\vartheta) i_{\zeta_1}(\vartheta)^{-1} \mathcal{J}_p(\vartheta)^t\big) \text{ für } n \to \infty.$$

Betrachten wir nun speziell die Abbildung $p : (\mathbb{R}^k, \mathfrak{B}^k) \to (\mathbb{R}^{d_1}, \mathfrak{B}^{d_1})$ gegeben durch $p(x) := x_1$, $x = (x_1, x_2) \in \mathbb{R}^{d_1} \times \mathbb{R}^{k-d_1}$, so ist $\mathcal{J}_p(\vartheta) = (\mathrm{I}_{d_1 \times d_1}, 0_{d_1 \times (k-d_1)})$, $\vartheta \in \tilde{\Theta}$, und wir erhalten mit $\vartheta = \binom{\vartheta^1}{\vartheta_2} \in \tilde{\Theta}$,

$$\sqrt{n}\big(p(t_n) - \vartheta^1\big) \xrightarrow{\mathrm{v}} \mathrm{N}_{d_1}\big(0, (\mathrm{I}_{d_1 \times d_1}, 0_{d_1 \times (k-d_1)}) i_{\zeta_1}(\vartheta)^{-1} (\mathrm{I}_{d_1 \times d_1}, 0_{d_1 \times (k-d_1)})^t\big) \text{ für } n \to \infty.$$

Es ist

$$i^{1,1}_{\zeta_1}(\vartheta) - i^{1,2}_{\zeta_1}(\vartheta) i^{2,2}_{\zeta_1}(\vartheta)^{-1} i^{2,1}_{\zeta_1}(\vartheta)$$

das Schurkomplement von $i^{2,2}_{\zeta_1}(\vartheta_0)$ in $i_{\zeta_1}(\vartheta_0)$. Wegen

$$\big(i_{\zeta_1}(\vartheta)^{-1}\big)_{\substack{1\le i\le d_1\\1\le j\le d_1}} = \big(i^{1,1}_{\zeta_1}(\vartheta) - i^{1,2}_{\zeta_1}(\vartheta) i^{2,2}_{\zeta_1}(\vartheta)^{-1} i^{2,1}_{\zeta_1}(\vartheta)\big)^{-1}$$

gilt

$$\sqrt{n}\big(p(t_n) - \vartheta^1\big) \xrightarrow{\mathrm{v}} \mathrm{N}_{d_1}\Big(0, \big(i^{1,1}_{\zeta_1}(\vartheta) - i^{1,2}_{\zeta_1}(\vartheta) i^{2,2}_{\zeta_1}(\vartheta)^{-1} i^{2,1}_{\zeta_1}(\vartheta)\big)^{-1}\Big) \text{ für } n \to \infty.$$

Somit ist die relative Effizienz zweier solcher Schätzer t und t' für ϑ^1_0 mit $\vartheta^0 = (\vartheta^1_0, \vartheta_2, \vartheta'_2) \in \Theta_0$, $\vartheta_0 = \binom{\vartheta^1_0}{\vartheta_2} \in \tilde{\Theta}_0$, $\vartheta'_0 = \binom{\vartheta^1_0}{\vartheta'_2} \in \tilde{\Theta}'_0$, und $\vartheta_1 \in \mathbb{R}^{d_1} \setminus \{0\}$ gegeben durch

$$\mathrm{eff}_{t',t}(\vartheta^0, \vartheta_1) = \frac{\vartheta^t_1\big(i^{1,1}_{\zeta_1}(\vartheta_0) - i^{1,2}_{\zeta_1}(\vartheta_0) i^{2,2}_{\zeta_1}(\vartheta_0)^{-1} i^{2,1}_{\zeta_1}(\vartheta_0)\big)^{-1}\vartheta_1}{\vartheta^t_1\big(i^{1,1}_{\zeta'_1}(\vartheta'_0) - i^{1,2}_{\zeta'_1}(\vartheta'_0) i^{2,2}_{\zeta'_1}(\vartheta'_0)^{-1} i^{2,1}_{\zeta'_1}(\vartheta'_0)\big)^{-1}\vartheta_1}.$$

Nun sei $t = (t_n)_{n\in\mathbb{N}}$ eine Folge von Schätzern für den unbekannten Parameter $\vartheta^1 \in \Theta^1$ basierend auf der Stichprobe $\zeta_1, \ldots, \zeta_n$, d.h. es ist $h_n : (R^n, \mathfrak{S}^n) \to (\mathbb{R}^{d_1}, \mathfrak{B}^{d_1})$ für jedes $n \in \mathbb{N}$ und $t_n = h_n(\zeta_1, \ldots, \zeta_n)$. Diese Folge von Schätzern erfülle

$$\sqrt{n}(t_n - \vartheta^1) \xrightarrow{\mathrm{v}} \mathrm{N}_{d_1}\big(0, \Sigma_t(\vartheta)\big) \text{ für } n \to \infty,$$

mit einer symmetrischen und positiv definiten Matrix $\Sigma_t(\vartheta) \in \mathbb{R}^{d_1\times d_1}$, $\vartheta = \binom{\vartheta^1}{\vartheta_2} \in \tilde{\Theta}$.

Argumentiert man wie oben, so kann diese Folge von Schätzern als optimal bei $\vartheta \in \tilde{\Theta}$ bezeichnet werden, wenn

$$\Sigma_t(\vartheta) = i^{1,1}_{\zeta_1}(\vartheta)^{-1} \tag{7.2.2}$$

gilt. Schließlich ist $i^{1,1}_{\zeta_1}(\vartheta)$ die Informationsmatrix aus $Q^{\zeta_1}_{\vartheta}$, wenn für $\vartheta = \binom{\vartheta^1}{\vartheta_2} \in \tilde{\Theta}$ der Parameter ϑ^1 als einziger variabler Parameter, und ϑ_2 als fest gegeben betrachtet wird. Siehe hierzu z.B. [38].

Somit ist die relative Effizienz zweier solcher Schätzer t und t' für ϑ^1_0 mit $\vartheta^0 = (\vartheta^1_0, \vartheta_2, \vartheta'_2) \in \Theta_0$, $\vartheta_0 = \binom{\vartheta^1_0}{\vartheta_2} \in \tilde{\Theta}_0$, $\vartheta'_0 = \binom{\vartheta^1_0}{\vartheta'_2} \in \tilde{\Theta}'_0$ und $\vartheta_1 \in \mathbb{R}^{d_1} \setminus \{0\}$ gegeben durch

$$\mathrm{eff}_{t',t}(\vartheta^0, \vartheta_1) = \frac{\vartheta^t_1 i^{1,1}_{\zeta_1}(\vartheta_0)^{-1}\vartheta_1}{\vartheta^t_1 1^{1,1}_{\zeta'_1}(\vartheta'_0)^{-1}\vartheta_1}.$$

Legen wir also für die betrachteten Schätzer die Optimalität gegeben durch (7.2.1) oder (7.2.2) zugrunde, so ergeben sich mit $\vartheta^0 = (\vartheta_0^1, \vartheta_2, \vartheta_2') \in \Theta_0$, $\vartheta_0 = \binom{\vartheta_0^1}{\vartheta_2} \in \tilde{\Theta}_0$, $\vartheta_0' = \binom{\vartheta_0^1}{\vartheta_2'} \in \tilde{\Theta}_0'$ und $\vartheta_1 \in \mathbb{R}^{d_1} \setminus \{0\}$ Effizienzen der Form

$$\mathrm{eff}_{t',t}(\vartheta^0, \vartheta_1) = \frac{\vartheta_1^t A_{\zeta_1}^{1,1}(\vartheta_0)^{-1}\vartheta_1}{\vartheta_1^t A_{\zeta_1'}^{1,1}(\vartheta_0')^{-1}\vartheta_1},$$

oder aber gemäß [36] Effizienzen der Form

$$\tilde{\mathrm{eff}}_{t',t}(\vartheta^0) = \frac{\det A_{\zeta_1'}^{1,1}(\vartheta_0')}{\det A_{\zeta_1}^{1,1}(\vartheta_0)},$$

wobei

$$A_{\zeta_1}^{1,1}(\vartheta_0) = (i_{\zeta_1}^{1,1}(\vartheta_0) - i_{\zeta_1}^{1,2}(\vartheta_0) i_{\zeta_1}^{2,2}(\vartheta_0)^{-1} i_{\zeta_1}^{2,1}(\vartheta_0) \text{ oder } A_{\zeta_1}^{1,1}(\vartheta_0) = i_{\zeta_1}^{1,1}(\vartheta_0),$$

sowie

$$A_{\zeta_1'}^{1,1}(\vartheta_0') = (i_{\zeta_1'}^{1,1}(\vartheta_0') - i_{\zeta_1'}^{1,2}(\vartheta_0') i_{\zeta_1'}^{2,2}(\vartheta_0')^{-1} i_{\zeta_1'}^{2,1}(\vartheta_0') \text{ oder } A_{\zeta_1'}^{1,1}(\vartheta_0') = i_{\zeta_1'}^{1,1}(\vartheta_0')$$

gilt.

Um eine Verbindung zu den Effizienzen für statistische Tests herzustellen, erinnern wir daran, dass wir unter Beachtung von Bemerkung 4.2.8 mit Hilfe der Resultate aus Abschnitt 4 sahen, dass die Effizienz von Tests T und T' mit $\vartheta^0 = (\vartheta_0^1, \vartheta_2, \vartheta_2') \in \Theta_0$, $\vartheta_0 = \binom{\vartheta_0^1}{\vartheta_2} \in \tilde{\Theta}_0$, $\vartheta_0' = \binom{\vartheta_0^1}{\vartheta_2'} \in \tilde{\Theta}_0'$ und $\vartheta_1 \in \mathbb{R}^{d_1} \setminus \{0\}$ häufig die Gestalt

$$\mathrm{eff}_{T',T}(\vartheta^0, \vartheta_1) = \frac{\vartheta_1^t A_{\zeta_1'}^{1,1}(\vartheta_0')\vartheta_1}{\vartheta_1^t A_{\zeta_1}^{1,1}(\vartheta_0)\vartheta_1},$$

oder die Gestalt

$$\tilde{\mathrm{eff}}_{T',T}(\vartheta^0) = \frac{\det A_{\zeta_1'}^{1,1}(\vartheta_0')}{\det A_{\zeta_1}^{1,1}(\vartheta_0)}$$

besitzt. Davon wollen wir auch hier ausgehen. Es gilt also $\tilde{\mathrm{eff}}_{t',t}(\vartheta^0, \vartheta_1) = \tilde{\mathrm{eff}}_{T',T}(\vartheta^0, \vartheta_1)$. Betrachten wir die Efiizienzen

$$\mathrm{eff}_{t',t}(\vartheta^0, \vartheta_1) = \frac{\vartheta_1^t A_{\zeta_1}^{1,1}(\vartheta_0)^{-1}\vartheta_1}{\vartheta_1^t A_{\zeta_1'}^{1,1}(\vartheta_0')^{-1}\vartheta_1} \text{ und } \mathrm{eff}_{T',T}(\vartheta^0, \vartheta_1) = \frac{\vartheta_1^t A_{\zeta_1'}^{1,1}(\vartheta_0')\vartheta_1}{\vartheta_1^t A_{\zeta_1}^{1,1}(\vartheta_0)\vartheta_1},$$

so können wir folgendes Lemma festhalten.

7.2.2 Lemma: *Es gelten die folgenden Äquivalenzen*

a) $\sup_{\vartheta_1 \in \mathbb{R}^{d_1} \setminus \{0\}} \mathrm{eff}_{t',t}(\vartheta^0, \vartheta_1) < 1 \;\Leftrightarrow\; \sup_{\vartheta_1 \in \mathbb{R}^{d_1} \setminus \{0\}} \mathrm{eff}_{T',T}(\vartheta^0, \vartheta_1) < 1.$

b) $\sup_{\vartheta_1 \in \mathbb{R}^{d_1} \setminus \{0\}} \mathrm{eff}_{t',t}(\vartheta^0, \vartheta_1) \leq 1 \;\Leftrightarrow\; \sup_{\vartheta_1 \in \mathbb{R}^{d_1} \setminus \{0\}} \mathrm{eff}_{T',T}(\vartheta^0, \vartheta_1) \leq 1.$

c) $\mathrm{eff}_{t',t}(\vartheta^0, \vartheta_1) = 1 \;\forall\; \vartheta_1 \in \mathbb{R}^{d_1} \setminus \{0\} \;\Leftrightarrow\; \mathrm{eff}_{T',T}(\vartheta^0, \vartheta_1) = 1 \;\forall\; \vartheta_1 \in \mathbb{R}^{d_1} \setminus \{0\}.$

d) $\inf_{\vartheta_1 \in \mathbb{R}^{d_1} \setminus \{0\}} \mathrm{eff}_{t',t}(\vartheta^0, \vartheta_1) \geq 1 \;\Leftrightarrow\; \inf_{\vartheta_1 \in \mathbb{R}^{d_1} \setminus \{0\}} \mathrm{eff}_{T',T}(\vartheta^0, \vartheta_1) \geq 1.$

e) $\inf_{\vartheta_1 \in \mathbb{R}^{d_1} \setminus \{0\}} \mathrm{eff}_{t',t}(\vartheta^0, \vartheta_1) > 1 \;\Leftrightarrow\; \inf_{\vartheta_1 \in \mathbb{R}^{d_1} \setminus \{0\}} \mathrm{eff}_{T',T}(\vartheta^0, \vartheta_1) > 1.$

Beweis: Es hat die Matrix $A^{1,1}_{\zeta_1}(\vartheta_0)^{-1} A^{1,1}_{\zeta_1'}(\vartheta_0')$ die gleichen Eigenwerte wie die Matrix $A^{1,1}_{\zeta_1'}(\vartheta_0') A^{1,1}_{\zeta_1}(\vartheta_0)^{-1}$. Mit analogen Überlegungen wie im Beweis von Lemma 4.5.1 folgt damit die Behauptung. □

7.2.3 Bemerkung: Aufgrund dieser Ergebnisse können die Tests zum Vergleich von Experimenten aus Abschnitt 5 auch verwendet werden, wenn nicht entsprechende Effizienzen von statistischen Tests, sondern Effizienzen von Schätzern in der oben beschriebenen Form betrachtet werden. Schließlich gelten alle Resultate aus Abschnitt 5 auch in der in diesem Unterabschnitt betrachteten Situation.

Genauso treffen aufgrund dieser Resultate alle Ergebnisse zum Vergleich von unabhängiger und verbundener Stichprobenerhebung aus Abschnitt 6 zu, wenn nicht entsprechende Effizienzen von statistischen Tests, sondern Effizienzen von Schätzern in der oben beschriebenen Form betrachtet werden. Dabei bekommt man natürlich für die expliziten Formeln zur Berechnung der Effizienz bei Annahme einer mehrparametrischen Exponentialfamilie entsprechend modifizierte Ergebnisse.

7.2.4 Beispiel: Betrachten wir die Situation aus Beispiel 4.2.10, d.h. wir vergleichen die verbundene Stichprobenerhebung mit der unabhängigen Sichprobenerhebung.

Im unabhängigen Stichprobenfall seien mit $X_1', \ldots, X_n'$ und $Y_1', \ldots, Y_n'$ zwei voneinander unabhängige Stichproben von unabhängigen $\mathbb{R}^{d_1}$-wertigen Zufallsvektoren gegeben, wobei die $X_1', \ldots, X_n'$ jeweils die gleiche $\mathrm{N}_{d_1}(a, \Sigma)$-Verteilung besitzen, und die $Y_1', \ldots, Y_n'$ jeweils die gleiche $\mathrm{N}_{d_1}(b, \Lambda)$-Verteilung besitzen, mit $a, b \in \mathbb{R}^{d_1}$ unbekannt, und $\Sigma, \Lambda \in \mathbb{R}^{d_1 \times d_1}$ symmetrisch positiv definit und ebenfalls unbekannt.

Im verbundenen Stichprobenfall liege dagegen eine Stichprobe $\binom{X_1}{Y_1}, \ldots, \binom{X_n}{Y_n}$ von unabhängigen $\mathbb{R}^{2d_1}$-wertigen Zufallsvektoren vor, die jeweils die gleiche Verteilung, konkret

$$\begin{pmatrix} X_1 \\ Y_1 \end{pmatrix} \sim \mathrm{N}_{2d_1}\left(\begin{pmatrix} a \\ b \end{pmatrix}, \begin{pmatrix} \Sigma & \Pi \\ \Pi^t & \Lambda \end{pmatrix} \right)$$

besitzen, wobei $\Pi \in \mathbb{R}^{d_1 \times d_1}$ ebenfalls unbekannt sei, sodass die zugrunde liegende Kovarianzmatrix positiv definit sei.

Es sollen Schätzer für $\mu := a - b$ betrachtet werden. Es sei $\zeta_i := X_i - Y_i$, $\zeta_i' := X_i' - Y_i'$, $i = 1, \ldots, n$. Dann ist

$$t_n := \overline{X}_n - \overline{Y}_n = \overline{\zeta}_n$$

der erwartungstreue Maximum-Likelihood-Schätzer für μ im verbundenen Stichprobenfall, und

$$t_n' := \overline{X_n'} - \overline{Y_n'} = \overline{\zeta_n'}$$

der erwartungstreue Maximum-Likelihood-Schätzer für μ im unabhängigen Stichprobenfall. Es gilt

$$\zeta_1' \sim \mathrm{N}_{d_1}(\mu, \Gamma_1),$$

mit $\Gamma_1 := \Sigma + \Lambda$, und

$$\zeta_1 \sim \mathrm{N}_{d_1}(\mu, \Gamma_2),$$

mit $\Gamma_2 := \Sigma + \Lambda - (\Pi + \Pi^t)$. Es liegen die Parametermengen Θ, $\tilde{\Theta}$ und $\tilde{\Theta}'$ wie in Beispiel 4.2.10 vor. Es ist

$$\sqrt{n}(t_n - \mu) \xrightarrow{\mathrm{v}} \mathrm{N}_{d_1}\big(0, i_{\zeta_1}^{1,1}(\vartheta_0)\big) \text{ für } n \to \infty,$$

und weiter

$$\sqrt{n}(t_n' - \mu) \xrightarrow{\mathrm{v}} \mathrm{N}_{d_1}\big(0, i_{\zeta_1'}^{1,1}(\vartheta_0')\big) \text{ für } n \to \infty.$$

Zudem gilt

$$i_{\zeta_1}^{1,1}(\vartheta_0) = \Gamma_2^{-1} \text{ und } i_{\zeta_1'}^{1,1}(\vartheta_0') = \Gamma_1^{-1},$$

sodass sich für die Effizienz

$$\mathrm{eff}_{t',t}(\vartheta^0, \vartheta_1) = \frac{\vartheta_1^t i_{\zeta_1}^{1,1}(\vartheta_0)^{-1} \vartheta_1}{\vartheta_1^t i_{\zeta_1'}^{1,1}(\vartheta_0')^{-1} \vartheta_1} = \frac{\vartheta_1^t \Gamma_2 \vartheta_1}{\vartheta_1^t \Gamma_1 \vartheta_1}$$

bzw.

$$\tilde{\mathrm{eff}}_{t',t}(\vartheta^0, \vartheta_1) = \frac{\det i_{\zeta_1'}^{1,1}(\vartheta_0')}{\det i_{\zeta_1}^{1,1}(\vartheta_0)} = \frac{\det \Gamma_2}{\det \Gamma_1}$$

ergibt.

Nun betrachten wir im verbundenen Stichprobenfall den erwartungstreuen Schätzer

$$t_n := \begin{pmatrix} \overline{\zeta}_n \\ \mathrm{vech}(S_n) \end{pmatrix}$$

für $\binom{\mu}{\mathrm{vech}(\Gamma_2)}$, mit $S_n = \frac{1}{n-1}\sum_{i=1}^n(\zeta_i - \overline{\zeta}_n)(\zeta_i - \overline{\zeta}_n)^t$, und im unabhängigen Stichprobenfall den erwartungstreuen Schätzer

$$t'_n := \begin{pmatrix} \overline{\zeta'_n} \\ \mathrm{vech}(S'_n) \end{pmatrix}$$

für $\binom{\mu}{\mathrm{vech}(\Gamma_1)}$, mit $S'_n = \frac{1}{n-1}\sum_{i=1}^n(\zeta'_i - \overline{\zeta'_n})(\zeta'_i - \overline{\zeta'_n})^t$. Dann ist

$$\sqrt{n}(t_n - \mu) \xrightarrow{\mathrm{v}} \mathrm{N}_{d_1+d_1(d_1+1)/2}\big(0, i_{\zeta_1}(\vartheta_0)\big) \text{ für } n \to \infty,$$

und weiter

$$\sqrt{n}(t'_n - \mu) \xrightarrow{\mathrm{v}} \mathrm{N}_{d_1+d_1(d_1+1)/2}\big(0, i_{\zeta'_1}(\vartheta'_0)\big) \text{ für } n \to \infty.$$

Zudem gilt

$$i^{1,2}_{\zeta_1}(\vartheta_0) = i^{1,2}_{\zeta'_1}(\vartheta'_0) = 0,$$

sodass sich für die Schätzer $p(t_n) = (\mathrm{I}_{d_1\times d_1}, 0_{d_1\times d_1(d_1+1)/2})t_n$ und $p(t'_n) = (\mathrm{I}_{d_1\times d_1}, 0_{d_1\times d_1(d_1+1)/2})t'_n$ das Ergebnis

$$\mathrm{eff}_{t',t}(\vartheta^0,\vartheta_1) = \frac{\vartheta_1^t\big(i^{1,1}_{\zeta_1}(\vartheta_0) - i^{1,2}_{\zeta_1}(\vartheta_0)i^{2,2}_{\zeta_1}(\vartheta_0)^{-1}i^{2,1}_{\zeta_1}(\vartheta_0)\big)^{-1}\vartheta_1}{\vartheta_1^t\big(i^{1,1}_{\zeta'_1}(\vartheta'_0) - i^{1,2}_{\zeta'_1}(\vartheta'_0)i^{2,2}_{\zeta'_1}(\vartheta'_0)^{-1}i^{2,1}_{\zeta'_1}(\vartheta'_0)\big)^{-1}\vartheta_1} = \frac{\vartheta_1^t i^{1,1}_{\zeta_1}(\vartheta_0)^{-1}\vartheta_1}{\vartheta_1^t i^{1,1}_{\zeta'_1}(\vartheta'_0)^{-1}\vartheta_1} = \frac{\vartheta_1^t\Gamma_2\vartheta_1}{\vartheta_1^t\Gamma_1\vartheta_1}$$

bzw.

$$\tilde{\mathrm{eff}}_{t',t}(\vartheta^0,\vartheta_1) = \frac{\det\big(i^{1,1}_{\zeta'_1}(\vartheta'_0) - i^{1,2}_{\zeta'_1}(\vartheta'_0)i^{2,2}_{\zeta'_1}(\vartheta'_0)^{-1}i^{2,1}_{\zeta'_1}(\vartheta'_0)\big)}{\det\big(i^{1,1}_{\zeta_1}(\vartheta_0) - i^{1,2}_{\zeta_1}(\vartheta_0)i^{2,2}_{\zeta_1}(\vartheta_0)^{-1}i^{2,1}_{\zeta_1}(\vartheta_0)\big)} = \frac{\det i^{1,1}_{\zeta'_1}(\vartheta'_0)}{\det i^{1,1}_{\zeta_1}(\vartheta_0)} = \frac{\det\Gamma_2}{\det\Gamma_1}$$

bestätigt. Vergleiche mit dem Resultat in Beispiel 4.2.10.

7.3 Konfidenzbereichsschätzverfahren

Wie in Abschnitt 6 und Unterabschnitt 7.2 gehen wir hier von einer bestimmten Modellannahme aus, die ein Spezialfall der bereits in Unterabschnitt 4.1 getroffenen Modellannahme darstellt. Es handelt sich um die Situation aus Bemerkung 4.2.8 und Unterabschnitt 4.4, d.h. es liegt das Modell aus Unterabschnitt 4.1 vor, wobei für ein gegebenes $\vartheta^1_0 \in \Theta^1$ $\{\vartheta^1_0\} = \Theta^1_0$ ist.

Im Folgenden wollen wir Annahmen an die Zufallsvariablen ζ_1 und ζ'_1 bei den zugrunde liegenden injektiven Parametrisierungen durch $\tilde{\Theta}$ und $\tilde{\Theta}'$ formulieren. Zudem werden wir zugehörige Konfidenzbereichschätzer B_n und B'_n basierend auf $\zeta_1,\dots,\zeta_n$ und $\zeta'_1,\dots,\zeta'_n$ einführen und diese ebenfalls mit Bedingungen versehen. Der Einfachheit halber führen wir dies nur für ζ_1, $\tilde{\Theta}$ und B_n aus; in analoger Weise sollen alle Definitionen und Annahmen auch für ζ'_1, $\tilde{\Theta}'$ und B'_n gelten. Dabei werden die im Zusammenhang mit ζ_1, $\tilde{\Theta}$ und B_n

eingeführten Größen im Zusammenhang mit ζ_1', $\tilde{\Theta}'$, und B_n' mit einem Strich versehen.

Für ein $\alpha \in (0,1)$ sei durch $1-\alpha$ ein Konfidenzniveau gegeben, und es sei durch $(B_n)_{n\in\mathbb{N}}$, mit $B_n : (\Omega, \mathfrak{A}) \to \mathfrak{P}(\Theta^1)$ für jedes $n \in \mathbb{N}$ eine Folge von Konfidenzbereichen für den unbekannten Parameter $\vartheta_1 \in \Theta^1$ zum (asymptotisch eingehaltenen) Konfidenzniveau $1-\alpha$ gegeben.

Wollen wir die Qualität zweier solcher Folgen von Konfidenzbereichen vergleichen, so kann dafür die in Abschnitt 3 eingeführte Volumen-Effizienz benutzt werden. Das in Unterabschnitt 3 vorgestellte Konzept der Volumen-Effizienz kann in offensichtlicher Weise auf die vorliegende Situation der zusammengesetzten Hypothese erweitert werden. Siehe hierzu die Erklärung am Anfang des Unterabschnitts 4.4.

Alternativ können aber auch relative Effizienzen von statistischen Tests, so wie sie in Abschnitt 2 beschrieben sind, verwendet werden, was die folgenden Überlegungen zeigen.

Es existiere für jedes $\alpha \in (0,1)$ eine Abbildung $b_n(\cdot,\alpha) : R^n \to \mathfrak{P}(\Theta^1)$, sodass für alle $\vartheta = (\vartheta_1, \vartheta_2) \in \tilde{\Theta}$

$$\{(z_1,\ldots,z_n) \in R^n; \vartheta_1 \in b_n((z_1,\ldots,z_n),\alpha)\} \in \mathfrak{S}^n \text{ für jedes } \alpha \in (0,1)$$

gilt und $B_n = b_n(\zeta_1,\ldots,\zeta_n,\alpha)$ ist.

Betrachten wir für $\vartheta_0^1 \in \Theta^1$ das Testproblem

$$\mathrm{H}: \ \vartheta_1 = \vartheta_0^1 \ \ \mathrm{K}: \ \vartheta_1 \in \Theta^1 \setminus \{\vartheta_0^1\}.$$

Dann ist für jedes $n \in \mathbb{N}$ der Test, festgelegt durch die Vorschrift

„Verwirf H, falls $\vartheta_0^1 \notin B_n$“,

ein Test für das Testproblem H gegen K zum (asymptotisch eingehaltenen) Testniveau α. Wir nennen die so entstandene Folge von Tests $T = (T_n)_{n\in\mathbb{N}}$, und schreiben $B_{T,n} := B_n$.

Gehen wir zusätzlich davon aus, dass für alle $\vartheta_0^1 \in \Theta^1$ eine Folge von Testgrößen $(T_n(\vartheta_0^1))_{n\in\mathbb{N}}$ und eine Folge von kritischen Werten $(c_n(\vartheta_0^1))_{n\in\mathbb{N}}$ existiere mit

$$Q_\vartheta\big(T_n(\vartheta_0^1) > c_n(\vartheta_0^1)\big) = Q_\vartheta(\vartheta_0^1 \notin B_{T,n}) \text{ für alle } \vartheta \in \tilde{\Theta}.$$

Dies kann stets durch die Setzung $T_n(\vartheta_0^1) = \mathrm{I}(\vartheta_0^1 \notin B_{T,n})$ und $c_n(\vartheta_0^1) = 0$ erreicht werden.

Der oben genannte Test lässt sich nun äquivalent gemäß

„Verwirf H, falls $T_n(\vartheta_0^1) > c_n(\vartheta_0^1)$“

formulieren. Es können daher zum Vergleich solcher Tests relative Effizienzen herangezogen werden, so wie sie in Abschnitt 2 beschrieben sind.

Diese Überlegungen zeigen, dass beim Vergleich von Experimenten anhand der Qualität von Konfidenzbereichsschätzverfahren auf relative Effizienzen für statistische Tests zurückgegriffen werden kann.

Gehen wir davon aus, dass zwei Folgen von Tests T und T' entsprechend dieser Überlegungen gegeben sind. Natürlich kann es sich bei diesen Tests um optimale Tests im Bahadur- oder Hodges-Lehmann Sinn, oder um Likelihood-Quotienten-Tests für das beschriebene Testproblem handeln. Somit motivieren die Ergebnisse der Unterabschnitte 4.2 und 4.3 mit $\vartheta^0 = (\vartheta_0^1, \vartheta_2, \vartheta_2') \in \Theta$, $\vartheta_0 = (\vartheta_0^1, \vartheta_2) \in \tilde{\Theta}$, $\vartheta_0' = (\vartheta_0^1, \vartheta_2') \in \tilde{\Theta}'$ und $\vartheta_1 \in \mathbb{R}^{d_1} \setminus \{0\}$ die Betrachtung von Effizienzen der Form

$$\mathrm{eff}_{T',T}(\vartheta^0, \vartheta_1) = \frac{\vartheta_1^t A_{\zeta_1'}^{1,1}(\vartheta_0')\vartheta_1}{\vartheta_1^t A_{\zeta_1}^{1,1}(\vartheta_0)\vartheta_1},$$

die im vorliegenden Fall als Effizienzen von Konfidenzbereichsschätzverfahren betrachtet werden können. Wir können folgende Bemerkung festhalten.

7.3.1 Bemerkung: Es können alle Tests zum Vergleich von Experimenten aus Abschnitt 5 auch verwendet werden, wenn nicht entsprechende Effizienzen von statistischen Tests, sondern Effizienzen von Konfidenzbereichsschätzverfahren in der oben beschriebenen Form betrachtet werden. Schließlich gelten alle Resultate aus Abschnitt 5 auch in der in diesem Unterabschnitt betrachteten Situation.

Genauso treffen alle Ergebnisse zum Vergleich von unabhängiger und verbundener Stichprobenerhebung aus Abschnitt 6 zu, wenn nicht entsprechende Effizienzen von statistischen Tests, sondern Effizienzen von Konfidenzbereichsschätzverfahren in der oben beschriebenen Form betrachtet werden.

8 Vergleiche bei multivariater Normalverteilung

8.1 Grundlegendes Modell

In diesem Abschnitt werden die verbundene und unabhängige Stichprobenerhebung unter der Annahme miteinander verglichen, dass die Beobachtungen multivariaten Normalverteilungen folgen.

Wir werden sehen, dass die Annahme der multivariaten Normalverteilung auf eine Situation führt, die als Beispiel für die in den Abschnitten 3, 4, 5 und 6 erarbeiteten theoretischen Resultate dienen kann. Allerdings können unter dieser Annahme Ergebnisse erzielt werden, die über die in den genannten Abschnitten erarbeiteten Resultate hinaus gehen.

Wir erinnern an dieser Stelle an das zu Beginn von Abschnitt 6 eingeführte Modell, das stets beim Vergleich von unabhängiger und verbundener Stichprobenerhebung zugrunde liegen soll und das sich am einführenden Beispiel in Unterabschnitt 1.1 orientiert. In diesem Abschnitt betrachteten wir konkret folgende Situation: Es sei $d \in \mathbb{N}$ und $n > d$. Auf dem zugrunde liegenden Wahrscheinlichkeitsraum seien im verbundenen Stichprobenfall durch $\binom{X_1}{Y_1}, \ldots, \binom{X_n}{Y_n}$ unabhängige $\mathbb{R}^{2d}$-wertige Zufallsvektoren gegeben, wobei $X_1, \ldots, X_n, Y_1, \ldots, Y_n$, jeweils $\mathbb{R}^d$-wertige Zufallsvektoren seien. Die $\binom{X_1}{Y_1}, \ldots, \binom{X_n}{Y_n}$ seien jeweils identisch multivariat normalverteilt gemäß[8].

$$\begin{pmatrix} X_1 \\ Y_1 \end{pmatrix} \sim \mathrm{N}_{2d}\left(\begin{pmatrix} a \\ b \end{pmatrix}, \begin{pmatrix} \Sigma & \Sigma\Sigma_1 \\ \Sigma\Sigma_2 & \Sigma \end{pmatrix}\right). \tag{8.1.1}$$

Dabei seien $a, b \in \mathbb{R}^d$ und $\Sigma, \Sigma_1, \Sigma_2 \in \mathbb{R}^{d \times d}$ unbekannt und mit $\mathcal{P}_l := \{A \in \mathcal{S}_l; A \text{ ist positiv definit}\}$ und $\mathcal{S}_l := \{A \in \mathbb{R}^{l \times l}; A \text{ ist symmetrisch}\}$, $l \in \mathbb{N}$, gelte

$$\begin{pmatrix} \Sigma & \Sigma\Sigma_1 \\ \Sigma\Sigma_2 & \Sigma \end{pmatrix} \in \mathcal{P}_{2d}. \tag{8.1.2}$$

Im unabhängigen Stichprobenfall seien mit $X_1', \ldots, X_n'$ und $Y_1', \ldots, Y_n'$ zwei voneinander unabhängige Stichproben auf dem zugrunde liegenden Wahrscheinlichkeitsraum gegeben. Wegen (6.1.3) soll $\mathcal{L}(X_1) = \mathcal{L}(X_1')$ und $\mathcal{L}(Y_1) = \mathcal{L}(Y_1')$ gelten und deshalb aufgrund von (8.1.1)

$$X_1' \sim \mathrm{N}_d(a, \Sigma), \; Y_1' \sim \mathrm{N}_d(b, \Sigma).$$

Zudem folgt aus (8.1.2) auch $\Sigma \in \mathcal{P}_d$. Das zu Beginn dieser Arbeit vorgestellte Testproblem (1.1.3),

$$\mathrm{H}: \; \mathcal{L}(X_1) = \mathcal{L}(Y_1), \; \mathrm{K}: \; \mathcal{L}(X_1) \neq \mathcal{L}(Y_1),$$

[8] Gilt $\binom{X_1}{Y_1} \sim \mathrm{N}_{2d}(\binom{a}{b}, \begin{pmatrix} \Sigma & \Pi \\ \Pi^t & \Sigma \end{pmatrix})$, mit einer entsprechenden Matrix $\Pi \in \mathbb{R}^{d \times d}$, so bedeutet dies $\Sigma_1 = \Sigma^{-1}\Pi$ und $\Sigma_2 = \Sigma^{-1}\Pi^t$, falls Σ invetierbar ist.

kann hier äquivalent gemäß

$$\mathrm{H}:\ a = b,\ \mathrm{K}:\ a \neq b$$

formuliert werden.

8.2 Verbundener Stichprobenfall

Es gelte also im verbundenen Stichprobenfall

$$\begin{pmatrix} X_1 \\ Y_1 \end{pmatrix} \sim \mathrm{N}_{2d}\left(\begin{pmatrix} a \\ b \end{pmatrix}, \begin{pmatrix} \Sigma & \Sigma\Sigma_1 \\ \Sigma\Sigma_2 & \Sigma \end{pmatrix}\right).$$

Setzen wir $\zeta_i := X_i - Y_i$, $1 \leq i \leq n$, so sind $\zeta_1, \ldots, \zeta_n$ unabhängig und identisch verteilt mit

$$\zeta_1 \sim \mathrm{N}_d(\mu, \Gamma),$$

mit $\mu := a - b$ und $\Gamma := \Sigma(2\,\mathrm{I}_d - (\Sigma_1 + \Sigma_2))$, wobei $\Gamma \in \mathcal{P}_d$ gilt. Wir können das Testproblem auch so formulieren:

$$\mathrm{H}:\ \mu = 0,\ \mathrm{K}:\ \mu \neq 0.$$

Als Testgröße für dieses Testproblem wählen wir die auf $\zeta_1, \ldots, \zeta_n$ basierende Hotellingsche T^2-Testgröße, genauer

$$T_n = \frac{n(n-d)}{d(n-1)} \overline{\zeta}_n^t S_n^{-1} \overline{\zeta}_n,$$

mit $\overline{\zeta}_n := \frac{1}{n}\sum_{i=1}^n \zeta_i$ und $S_n := \frac{1}{n-1}\sum_{i=1}^n (\zeta_i - \overline{\zeta}_n)(\zeta_i - \overline{\zeta}_n)^t$, wobei S_n f.s. invertierbar ist, da $n > d$ gilt. Die Hotellingsche T^2-Testgröße und ihre Eigenschaften zind z.B. bei Muirhead ([28]) beschrieben. Es gilt

$$T_n \stackrel{v}{=} \frac{W_1/d}{W_2/(n-d)} \sim \mathrm{F}_{d,n-d}(\delta),$$

mit $\delta = n\mu'\Gamma^{-1}\mu$ und unabhängigen Zufallsvariablen $W_1 \sim \chi_d^2(\delta)$ und $W_2 \sim \chi_{n-d}^2$, siehe hierzu Theorem 1.3.6 und Theorem 6.3.1 aus [28]. Zu einem vorgegebenen Testniveau $\alpha \in (0,1)$ formulieren wir den Hotellingschen T^2-Test

$$\text{„Verwirf H, falls } T_n \geq \mathrm{F}_{d,n-d}^{\alpha}\text{“}.$$

Dabei ist $\mathrm{F}_{d,n-d}^{\alpha}$ das $(1-\alpha)$-Quantil der $\mathrm{F}_{d,n-d}$-Verteilung. Die Testgröße T_n ist äquivalent zur Likelihood-Quotienten-Testgröße für das Testproblem H gegen K, wenn man als Stichprobe $\zeta_1, \ldots, \zeta_n$ zugrunde legt. Siehe Theorem 6.3.2 aus [28]. Es sei

$$\Phi_d := \{\phi_A; \phi_A : \mathbb{R}^d \times \mathcal{D}_d \to \mathbb{R}^d \times \mathcal{D}_d, (x, D) \mapsto (Ax, ADA^t), A \in \mathrm{GL}(d, \mathbb{R})\},$$

mit $\mathrm{GL}(d,\mathbb{R}) = \{A \in \mathbb{R}^{d\times d}; A \text{ ist invertierbar}\}$, der Raum der Transformationen auf $\mathbb{R}^d \times \mathcal{D}_d$, welcher von der linearen Gruppe $\mathrm{GL}(d,\mathbb{R})$ induziert wird. Man sieht sofort, dass die Testgröße T_n invariant unter Transformationen aus Φ_d ist. Darüber hinaus ist der oben formulierte Hotellingsche T^2-Test unter den Tests, welche invariant unter Transformationen aus Φ_d sind, ein gleichmäßig bester Test zum Testniveau $\alpha \in (0,1)$ für das Testproblem H gegen K, wenn man als Stichprobe $\zeta_1,\dots,\zeta_n$ zugrunde legt. Siehe Theorem 6.3.4 aus [28].

8.3 Hodges-Lehmann-Index im verbundenen Stichprobenfall

In der Tat ist der Hotellingsche T^2-Test ein optimaler Test im Hodges-Lehmann Sinn für das Testproblem H gegen K, wenn man als Stichprobe $\zeta_1,\dots,\zeta_n$ zugrunde legt. Dies hat Kourouklis ([22]) gezeigt. Zur Bestimmung der Kullback-Leiber-Information bemerken wir, dass dieser Test auf unabhängigen Zufallsvariablen $\zeta_1,\dots,\zeta_n$ basiert, die jeweils $\mathrm{N}_d(\mu,\Gamma)$-verteilt sind.

Ganz allgemein berechnet sich die Kullback-Leibler-Information basierend auf der Verteilung einer Zufallsvariablen V mit der Verteilung $\mathrm{N}_p(\nu,\Lambda)$, $p \in \mathbb{N}$, für $\vartheta_0 = (\nu_0,\Lambda_0), \vartheta = (\nu,\Lambda)$, mit $\vartheta_0,\vartheta \in \mathbb{R}^p \times \mathcal{P}_p$, gemäß

$$\begin{aligned}
&K(\vartheta_0,\vartheta) \qquad\qquad (8.3.1)\\
=&\int \log \frac{\mathrm{d}P^V_{\vartheta_0}}{\mathrm{d}P^V_{\vartheta}}(x)\mathrm{d}P^V_{\vartheta_0}(x)\\
=&\log\left(\sqrt{\frac{\det\Lambda}{\det\Lambda_0}}\right) - \frac{1}{2}\int \big((x-\nu_0)^t\Lambda_0^{-1}(x-\nu_0) - (x-\nu)^t\Lambda^{-1}(x-\nu)\big)\mathrm{d}N_p(\nu_0,\Lambda_0)(x)\\
=&-\frac{1}{2}\log\det(\Lambda_0\Lambda^{-1}) + \frac{1}{2}(\nu-\nu_0)^t\Lambda^{-1}(\nu-\nu_0) + \frac{1}{2}\operatorname{tr}(\Lambda_0\Lambda^{-1}) - \frac{p}{2}.
\end{aligned}$$

Dieser Ausdruck muss in unserem Fall für $p = d$ bei gegebenem $\vartheta = (\mu,\Gamma)$ über alle $\vartheta_0 \in \tilde{\Theta}_0 = \{0\} \times \mathcal{P}_p$ minimiert werden.

Hierzu seien $l_1,\dots,l_p > 0$ die Eigenwerte der Matrix $\Lambda_0\Lambda^{-1}$. Dann folgt mit der Ungleichung $\log x \le x - 1,\ x > 0$,

$$\begin{aligned}
K(\vartheta_0,\vartheta) &= -\frac{1}{2}\log\det(\Lambda_0\Lambda^{-1}) + \frac{1}{2}(\nu-\nu_0)^t\Lambda^{-1}(\nu-\nu_0) + \frac{1}{2}\operatorname{tr}(\Lambda_0\Lambda^{-1}) - \frac{p}{2}\\
&= -\frac{1}{2}\sum_{i=1}^{p}(\log l_i - l_i + 1) + \frac{1}{2}(\nu-\nu_0)^t\Lambda^{-1}(\nu-\nu_0) \ge \frac{1}{2}(\nu-\nu_0)^t\Lambda^{-1}(\nu-\nu_0),
\end{aligned}$$

wobei die Gleichheit gilt, wenn $l_1 = \dots = l_p = 1$, äquivalent $\Lambda_0 = \Lambda$ gilt.

Damit ergibt sich in unserem Fall für $\mu \neq 0$

$$\liminf_{n\to\infty} \frac{1}{n}\log\beta_n(\alpha,\mu) \ge -K(\tilde{\Theta}_0,\vartheta) = -\frac{1}{2}\mu^t\Gamma^{-1}\mu.$$

Für eine Zufallsvariable $V \sim \chi^2_m(\nu)$, mit $m \in \mathbb{N}$ und $\nu \geq 0$, ist die momenterzeugende Funktion gegeben durch

$$\mathrm{E}\left(\exp(tV)\right) = \exp\Big(\frac{t\nu}{1-2t}\Big)(1-2t)^{-\frac{m}{2}},\ t < \frac{1}{2}.$$

Damit und mit der Markov-Ungleichung ergibt sich mit $\tau < 0$ für genügend großes n

$$\begin{aligned}\beta_n(\alpha,\mu) = P_\mu(T_n < c_n) &= P_\mu\Big(\frac{W_1/d}{W_2/(n-d)} \leq c_n\Big) = P_\mu\Big(\tau W_1 - \frac{\tau d}{n-d}c_n W_2 \geq 0\Big)\\ &\leq \mathrm{E}_\mu\left(\exp(\tau W_1)\right)\mathrm{E}_\mu\left(\exp\Big(-\frac{\tau d}{n-d}c_n W_2\Big)\right)\\ &= \exp\Big(\frac{\tau\delta}{1-2\tau}\Big)(1-2\tau)^{-\frac{d}{2}}\Big(1+2\frac{\tau d}{n-d}c_n\Big)^{-\frac{n-d}{2}}.\end{aligned}$$

Daraus folgt

$$\limsup_{n\to\infty}\frac{1}{n}\log\beta_n(\alpha,\mu) \leq \frac{\tau}{1-2\tau}\mu^t\Gamma^{-1}\mu.$$

Lässt man jetzt $\tau \to -\infty$ laufen, erhält man

$$\limsup_{n\to\infty}\frac{1}{n}\log\beta_n(\alpha,\mu) \leq -\frac{1}{2}\mu^t\Gamma^{-1}\mu = -K(\tilde{\Theta}_0,\vartheta),$$

und damit die Hodges-Lehmann-Optimalität für das Testproblem H gegen K, wenn man als Stichprobe $\zeta_1,\ldots,\zeta_n$ zugrunde legt. Vergleiche hierzu mit Beispiel 3.2 aus [22]. Für den Hodges-Lehmann-Index der Testgröße für das Testproblem H gegen K gilt daher mit $\mu = a - b$

$$d_T(\mu) = \frac{1}{2}\mu^t\Gamma^{-1}\mu.$$

8.4 Bahadur-Steigung im verbundenen Stichprobenfall

Der Hotellingsche T^2-Test ist ebenfalls optimal im Bahadur Sinn für das Testproblem H gegen K, wenn man als Stichprobe $\zeta_1,\ldots,\zeta_n$ zugrunde legt. Um dies zu zeigen gibt es verschiedene Herangehensweisen. Wir wählen hier eine andere als Koziol in der Arbeit [24], indem wir Lemma 2.3.2 verwenden.

8.4.1 Satz: *Der Hotellingsche* T^2*-Test ist optimal im Bahadur Sinn für das Testproblem H gegen K, wenn man als Stichprobe* $\zeta_1,\ldots,\zeta_n$ *zugrunde legt. Mit* $\mu = a - b$ *ist die Bahadur-Steigung*

$$c_T(\mu) = \frac{1}{2}\log(1+\mu^t\Gamma^{-1}\mu).$$

Beweis: Aufgrund des starken Gesetzes der großen Zahlen ist

$$\lim_{n\to\infty} \frac{d(n-1)}{n(n-d)} T_n = \mu^t \Gamma^{-1} \mu \text{ f.s.}$$

erfüllt. Weiter erhalten wir wieder mit Hilfe der Markov-Ungleichung für $t \geq 0$

$$P_0\Big(\frac{d(n-1)}{n(n-d)} T_n \geq t\Big) = P_0\Big(\frac{d(n-1)}{n(n-d)} \frac{W_1/d}{W_2/(n-d)} \geq t\Big) = P_0\Big(\frac{n-1}{2n} W_1 - \frac{t}{2} W_2 \geq 0\Big)$$
$$\leq \mathrm{E}_0\left(\exp\left(\frac{n-1}{2n} W_1\right)\right) \mathrm{E}_0\left(\exp\left(-\frac{t}{2} W_2\right)\right) = n^{\frac{d}{2}} (1+t)^{-\frac{n-d}{2}}.$$

Daraus folgt

$$\limsup_{n\to\infty} \frac{1}{n} \log P_0\Big(\frac{d(n-1)}{n(n-d)} T_n \geq t\Big) \leq -\frac{1}{2} \log(1+t),\ t \geq 0.$$

Wenden wir uns nun der Kullback-Leibler-Information zu. Mit Hilfe der in Unterabschnitt 8.3 gefundenen Darstellung (8.3.1) ergibt sich durch zyklische Vertauschung innerhalb des Spuroperators für die Kullback-Leibler-Information basierend auf der Verteilung einer Zufallsvariablen V mit der Verteilung $\mathrm{N}_p(\nu, \Lambda)$, $p \in \mathbb{N}$, für $\vartheta_0 = (\nu_0, \Lambda_0), \vartheta = (\nu, \Lambda)$, mit $\vartheta_0, \vartheta \in \mathbb{R}^p \times \mathcal{P}_d$,

$$\begin{aligned}
&K(\vartheta, \vartheta_0)\\
&= -\frac{1}{2} \log\det(\Lambda\Lambda_0^{-1}) + \frac{1}{2}(\nu-\nu_0)^t \Lambda_0^{-1}(\nu-\nu_0) + \frac{1}{2}\operatorname{tr}(\Lambda\Lambda_0^{-1}) - \frac{p}{2}\\
&= -\frac{1}{2} \log\det(\Lambda\Lambda_0^{-1}) + \frac{1}{2}\operatorname{tr}\Big(\big(\Lambda + (\nu-\nu_0)(\nu-\nu_0)^t\big)\Lambda_0^{-1}\Big) - \frac{p}{2}\\
&= -\frac{1}{2} \log\det\Big(\big(\Lambda + (\nu-\nu_0)(\nu-\nu_0)^t\big)\Lambda_0^{-1}\Big) + \frac{1}{2}\operatorname{tr}\Big(\big(\Lambda + (\nu-\nu_0)(\nu-\nu_0)^t\big)\Lambda_0^{-1}\Big) - \frac{p}{2}\\
&\quad + \frac{1}{2} \log\det\Big(\big(\Lambda + (\nu-\nu_0)(\nu-\nu_0)^t\big)\Lambda_0^{-1}\Big) - \frac{1}{2}\log\det(\Lambda\Lambda_0^{-1})\\
&= -\frac{1}{2} \log\det\Big(\big(\Lambda + (\nu-\nu_0)(\nu-\nu_0)^t\big)\Lambda_0^{-1}\Big) + \frac{1}{2}\operatorname{tr}\Big(\big(\Lambda + (\nu-\nu_0)(\nu-\nu_0)^t\big)\Lambda_0^{-1}\Big) - \frac{p}{2}\\
&\quad - \frac{1}{2} \log\det\Big(\Lambda\big(\Lambda + (\nu-\nu_0)(\nu-\nu_0)^t\big)^{-1}\Big).
\end{aligned}$$

Dieser Ausdruck muss nun für $p = d$ bei gegebenem $\vartheta = (\mu, \Gamma)$ über alle $\vartheta_0 \in \tilde{\Theta}_0 = \{0\} \times \mathcal{P}_d$ minimiert werden. Mit den gleichen Überlegungen wie in Unterabschnitt 8.3 erhalten wir zunächst

$$K(\vartheta, \tilde{\Theta}_0) = -\frac{1}{2} \log\det\big(\Gamma(\Gamma + \mu\mu^t)^{-1}\big),$$

und daraus mit Hilfe von Sylvesters Determinantentheorem

$$K(\vartheta, \tilde{\Theta}_0) = -\frac{1}{2} \log\det\left(\Gamma(\Gamma + \mu\mu^t)^{-1}\right) = \frac{1}{2}\log\det(\mathrm{I}_d + \Gamma^{-1}\mu\mu^t) = \frac{1}{2}\log(1 + \mu^t\Gamma^{-1}\mu).$$

Siehe hierzu auch Beispiel 3.2 aus [24]. Mit Hilfe von Lemma 2.3.2 folgt die Optimalität des Hotellingschen T^2-Tests im Bahadur Sinn für das Testproblen H gegen K, wenn man als Stichprobe $\zeta_1, \ldots, \zeta_n$ zugrunde legt. □

8.5 Unabhängiger Stichprobenfall

Es liegen im unabhängigen Stichprobenfall mit $X_1', \ldots, X_n'$ und $Y_1', \ldots, Y_n'$ zwei voneinenander unabhängige Stichproben mit

$$X_1' \sim \mathrm{N}_d(a, \Sigma),\; Y_1' \sim \mathrm{N}_d(b, \Sigma)$$

vor. Als Testgröße für das Testproblem H gegen K wählen wir die Hotellingsche Zweistichproben-T^2-Testgröße

$$T_n' = \frac{n(2n-1-d)}{4d(n-1)}(\overline{X_n'} - \overline{Y_n'})^t (S_n^{x,y})^{-1}(\overline{X_n'} - \overline{Y_n'}),$$

mit $\overline{X_n'}$ bzw. $\overline{Y_n'}$ als empirischer Erwartungswertvektor der ersten bzw. zweiten Stichprobe und $S_n^{x,y} := \frac{1}{2}(S_n^x + S_n^y)$, mit S_n^x bzw. S_n^y als empirische Kovarianzmatrix der ersten bzw. zweiten Stichprobe. Dann gilt

$$T_n' \stackrel{v}{=} \frac{W_1'/d}{W_2'/(2n-1-d)} \sim \mathrm{F}_{d,2n-1-d}\left(\delta'\right),$$

mit $\delta' = \frac{n}{2}\mu^t\Sigma^{-1}\mu$ und unabhängigen $W_1' \sim \chi_d^2(\delta')$ und $W_2' \sim \chi^2_{2n-1-d}$, wobei $\mu = a - b$ sei. Siehe hierzu Theorem 1.3.6 und Theorem 6.3.1 aus [28]. Zu einem vorgegebenen Testniveau $\alpha \in (0,1)$ formulieren wir den Hotellingschen Zweistichproben-T^2-Test

$$„\text{Verwirf H, falls } T_n' \geq \mathrm{F}^{\alpha}_{d,2n-1-d}\text{“}.$$

In der Tat handelt es sich bei der Testgröße T_n' um eine zur Likelihood-Quotienten-Testgröße für das Testproblem H gegen K äquivalente Testgröße. Dies folgt aus den Ausführungen zu Beginn des Abschnitts 10.7 in [28]. Allerdings lässt sich dies auch direkt nachrechnen, und zwar wie folgt.

Es sind die Maximum-Likelihood-Schätzer für a, für b und für Σ gegeben durch $\overline{X_n'}$, $\overline{Y_n'}$ und $S_n^{x,y}$, und die Maximum-Likelihood-Schätzer unter H für a und Σ sind gegeben durch

$$\hat{a}_n := \frac{1}{2n}\sum_{i=1}^n (X_i' + Y_i'),\; \hat{\Sigma}_n := \frac{1}{2n}\sum_{i=1}^n \left((X_i' - \hat{a}_n)(X_i' - \hat{a}_n)^t + (Y_i' - \hat{a}_n)(Y_i' - \hat{a}_n)^t\right).$$

Dies zeigt man folgendermaßen. Es sei $\zeta_i' := \binom{X_i'}{Y_i'}$, $i = 1, \ldots, n$. Die Dichte der $\zeta_1', \ldots \zeta_n'$ lautet

$$\begin{aligned}
&L(a,b,\Sigma)\\
:=&\left(\frac{1}{2\pi}\right)^{nd}\left(\det\begin{pmatrix}\Sigma & 0\\ 0 & \Sigma\end{pmatrix}\right)^{-\frac{n}{2}}\exp\left(-\frac{1}{2}\mathrm{tr}\begin{pmatrix}\Sigma & 0\\ 0 & \Sigma\end{pmatrix}^{-1}\sum_{i=1}^{n}\begin{pmatrix}X_i'-a\\ Y_i'-b\end{pmatrix}\begin{pmatrix}X_i'-a\\ Y_i'-b\end{pmatrix}^t\right).
\end{aligned}$$

Allgemein ergibt sich daraus

$$\begin{aligned}
&L(a,b,\Sigma)\\
=&\left(\frac{1}{2\pi}\right)^{nd}(\det\Sigma)^{-n}\exp\left(-\frac{1}{2}\mathrm{tr}\begin{pmatrix}\Sigma & 0\\ 0 & \Sigma\end{pmatrix}^{-1}\sum_{i=1}^{n}\begin{pmatrix}X_i'-\overline{X_n'}\\ Y_i'-\overline{Y_n'}\end{pmatrix}\begin{pmatrix}X_i'-\overline{X_n'}\\ Y_i'-\overline{Y_n'}\end{pmatrix}^t\right)\\
&\exp\left(-\frac{1}{2}\mathrm{tr}\begin{pmatrix}\Sigma & 0\\ 0 & \Sigma\end{pmatrix}^{-1}n\begin{pmatrix}\overline{X_n'}-a\\ \overline{Y_n'}-b\end{pmatrix}\begin{pmatrix}\overline{X_n'}-a\\ \overline{Y_n'}-b\end{pmatrix}^t\right)\\
=&\left(\frac{1}{2\pi}\right)^{nd}(\det\Sigma)^{-n}\exp\left(-\frac{1}{2}\mathrm{tr}\Sigma^{-1}\sum_{i=1}^{n}\big((X_i'-\overline{X_n'})(X_i'-\overline{X_n'})^t+(Y_i'-\overline{Y_n'})(Y_i'-\overline{Y_n'})^t\big)\right)\\
&\exp\left(-\frac{1}{2}\mathrm{tr}\begin{pmatrix}\Sigma & 0\\ 0 & \Sigma\end{pmatrix}^{-1}n\begin{pmatrix}\overline{X_n'}-a\\ \overline{Y_n'}-b\end{pmatrix}\begin{pmatrix}\overline{X_n'}-a\\ \overline{Y_n'}-b\end{pmatrix}^t\right)\\
=&\left(\frac{1}{2\pi}\right)^{nd}(\det\Sigma)^{-n}\exp\left(-\frac{1}{2}\mathrm{tr}\Sigma^{-1}\sum_{i=1}^{n}\big((X_i'-\overline{X_n'})(X_i'-\overline{X_n'})^t+(Y_i'-\overline{Y_n'})(Y_i'-\overline{Y_n'})^t\big)\right)\\
&\exp\left(-\frac{1}{2}n\begin{pmatrix}\overline{X_n'}-a\\ \overline{Y_n'}-b\end{pmatrix}^t\begin{pmatrix}\Sigma & 0\\ 0 & \Sigma\end{pmatrix}^{-1}\begin{pmatrix}\overline{X_n'}-a\\ \overline{Y_n'}-b\end{pmatrix}\right).
\end{aligned}$$

Daher ergeben sich als Maximum-Likelihood-Schätzer die genannten Ausdrücke; vergleiche mit der Argumentation im Beweis von Theorem 3.1.5 in [28]. Unter H ergibt sich dagegen aus der Dichte

$$\begin{aligned}
&L(a,a,\Sigma)\\
=&\left(\frac{1}{2\pi}\right)^{nd}(\det\Sigma)^{-n}\exp\left(-\frac{1}{2}\mathrm{tr}\Sigma^{-1}\sum_{i=1}^{n}\big((X_i'-a)(X_i'-a)^t+(Y_i'-a)(Y_i'-a)^t\big)\right)\\
=&\left(\frac{1}{2\pi}\right)^{nd}(\det\Sigma)^{-n}\exp\left(-\frac{1}{2}\mathrm{tr}\Sigma^{-1}\sum_{i=1}^{n}\big((X_i'-\hat{a}_n)(X_i'-\hat{a}_n)^t+(Y_i'-\hat{a}_n)(Y_i'-\hat{a}_n)^t\big)\right)\\
&\exp\left(-\frac{1}{2}\mathrm{tr}\Sigma^{-1}\Big(\sum_{i=1}^{n}\big((X_i'-a)(X_i'-a)^t+(Y_i'-a)(Y_i'-a)^t\big)\right.\\
&\left.-\sum_{i=1}^{n}\big((X_i'-\hat{a}_n)(X_i'-\hat{a}_n)^t+(Y_i'-\hat{a}_n)(Y_i'-\hat{a}_n)^t\big)\Big)\right).
\end{aligned}$$

Es ist

$$
\begin{aligned}
&\sum_{i=1}^n \big((X_i'-a)(X_i'-a)^t+(Y_i'-a)(Y_i'-a)^t\big)\\
&\quad-\sum_{i=1}^n \big((X_i'-\hat a_n)(X_i'-\hat a_n)^t+(Y_i'-\hat a_n)(Y_i'-\hat a_n)^t\big)\\
&=\sum_{i=1}^n X_i'(X_i')^t-n\overline{X_n'}a^t-na\overline{X_n'}^t+2naa^t+\sum_{i=1}^n Y_i'(Y_i')^t-n\overline{Y_n'}a^t-na\overline{Y_n'}^t\\
&\quad-(\sum_{i=1}^n X_i'(X_i')^t-n\overline{X_n'}\hat a_n^t-n\hat a_n\overline{X_n'}^t+2n\hat a_n\hat a_n^t+\sum_{i=1}^n Y_i'(Y_i')^t-n\overline{Y_n'}\hat a_n^t-n\hat a_n\overline{Y_n'}^t)\\
&=-n\overline{X_n'}(a-\hat a_n)^t-n(a-\hat a_n)\overline{X_n'}^t+2n(aa^t-\hat a_n\hat a_n^t)-n\overline{Y_n'}(a-\hat a_n)^t-n(a-\hat a_n)\overline{Y_n'}^t\\
&=-n(\overline{X_n'}+\overline{Y_n'})(a-\hat a_n)^t-n(a-\hat a_n)(\overline{X_n'}+\overline{Y_n'})^t+2n(aa^t-\hat a_n\hat a_n^t)\\
&=-2n\hat a_n(a-\hat a_n)^t-2n(a-\hat a_n)\hat a_n^t+2n(aa^t-\hat a_n\hat a_n^t)=2n(\hat a_n-a)(\hat a_n-a)^t,
\end{aligned}
$$

und somit ergeben sich mit der gleichen Argumentation wie im allgemeinen Fall die genannten Ausdrücke als Maximum-Likelihood-Schätzer. Ferner ist

$$
\begin{aligned}
\hat\Sigma_n=&\frac{1}{2n}\sum_{i=1}^n \big((X_i'-\hat a_n)(X_i'-\hat a_n)^t+(Y_i'-\hat a_n)(Y_i'-\hat a_n)^t\big)\big)\\
=&\frac{1}{2n}\sum_{i=1}^n\Big(\big(X_i'-\overline{X_n'}+(\overline{X_n'}-\overline{Y_n'})/2\big)\big(X_i'-\overline{X_n'}+(\overline{X_n'}-\overline{Y_n'})/2\big)^t\\
&+\big(Y_i'-\overline{Y_n'}-(\overline{X_n'}-\overline{Y_n'})/2\big)\big(Y_i'-\overline{Y_n'}+(\overline{X_n'}-\overline{Y_n'})/2\big)^t\Big)\\
&=S_n^{x,y}+\frac14(\overline{X_n'}-\overline{Y_n'})(\overline{X_n'}-\overline{Y_n'})^t,
\end{aligned}
$$

und somit ergibt sich als Likelihood-Quotienten-Testgröße

$$
\begin{aligned}
\frac{L(\hat a_n,\hat a_n,\hat\Sigma_n)}{L(\overline{X}_n,\overline{Y}_n,S_n^{x,y})}&=\left(\frac{\det S_n^{x,y}}{\det\hat\Sigma_n}\right)^n=\left(\det\left(\mathrm{I}_d+\frac14(S_n^{x,y})^{-1}(\overline{X_n'}-\overline{Y_n'})(\overline{X_n'}-\overline{Y_n'})^t\right)\right)^{-n}\\
&=\left(1+\frac14(\overline{X_n'}-\overline{Y_n'})^t(S_n^{x,y})^{-1}(\overline{X_n'}-\overline{Y_n'})\right)^{-n}=\left(1+\frac{n(2n-1-d)}{d(n-1)}T_n'\right)^{-n},
\end{aligned}
$$

also eine streng monoton fallende Funktion der Testgröße T_n'.

Es sei

$$
\begin{aligned}
\Psi_d:=\{\psi_{(A,c)};\psi_{(A,c)}:\mathbb{R}^d\times\mathbb{R}^d\times\mathcal{D}_d\to\mathbb{R}^d\times\mathbb{R}^d\times\mathcal{D}_d,\\
(x,y,D)\mapsto(Ax+c,Ay+c,ADA^t),(A,c)\in\mathrm{QL}(d,\mathbb{R})\},
\end{aligned}
$$

mit $\mathrm{QL}(d,\mathbb{R})=\{(A,c);A\in\mathrm{GL}(d,\mathbb{R}),c\in\mathbb{R}^d\}$, der Raum der Transformationen auf $\mathbb{R}^d\times\mathbb{R}^d\times\mathcal{D}_d$, der durch die affine Gruppe[9] $\mathrm{QL}(d,\mathbb{R})$ induziert wird. Man erkennt sofort,

[9]Die zugehörige Gruppenoperation $\circ$ ist über $(A_1,c_1)\circ(A_2,c_2):=(A_1A_2,A_1c_2+c_1)$ definiert.

dass die Testgröße T_n' invariant unter Transformationen aus Ψ_d ist. Darüber hinaus kann gezeigt werden, dass der oben formulierte Hotellingsche Zweistichproben-T^2-Test unter den Tests, welche invariant unter Transformationen aus Ψ_d sind, ein gleichmäßig bester Test zum Testniveau $\alpha \in (0,1)$ für das Testproblem H gegen K ist. Siehe hierzu Theorem 6.3.5 aus [28].

8.6 Hodges-Lehmann-Index im unabhängigen Stichprobenfall

8.6.1 Satz: *Der Hotellingsche Zweistichproben-*T^2*-Test ist ein optimaler Test im Hodges-Lehmann Sinn für das Testproblem H gegen K, und es ist mit* $\mu = a - b$

$$d_{T'}(\mu) = \frac{1}{4}\mu^t \Sigma^{-1} \mu.$$

Beweis: Man bemerke, dass mit $\zeta_1', \dots \zeta_n'$ Beobachtungen von unabhängig und identisch verteilten Zufallsvariablen vorliegen, mit

$$\zeta_1' = \begin{pmatrix} X_1' \\ Y_1' \end{pmatrix} \sim \mathrm{N}_{2d}\left(\begin{pmatrix} a \\ b \end{pmatrix}, \begin{pmatrix} \Sigma & 0 \\ 0 & \Sigma \end{pmatrix} \right).$$

Es ist für $a, b, c \in \mathbb{R}^d$ und $a \neq b$

$$\begin{aligned}
&\frac{1}{2}\begin{pmatrix} c-a \\ c-b \end{pmatrix}^t \begin{pmatrix} \Sigma^{-1} & 0 \\ 0 & \Sigma^{-1} \end{pmatrix} \begin{pmatrix} c-a \\ c-b \end{pmatrix} \\
=&\frac{1}{2}\big((c-a)^t\Sigma^{-1}(c-a) + (c-b)^t\Sigma^{-1}(c-b)\big) \\
=&\big(c - \frac{1}{2}(a+b)\big)^t\Sigma^{-1}\big(c - \frac{1}{2}(a+b)\big) + \frac{1}{4}(a-b)^t\Sigma^{-1}(a-b).
\end{aligned}$$

Da Σ positiv definit ist, und damit auch Σ^{-1} positiv definit ist, wird dieser Ausdruck in c genau dann minimal, wenn $c = (a+b)/2$ gewählt wird. Somit erhalten wir gemäß der in Unterabschnitt 8.3 für die Annahme der multivariaten Normalverteilung berechneten Formel für die Kullback-Leibler-Information mit analogen Überlegungen wie im Unterabschnitt 8.3

$$K\big(\tilde{\Theta}_0', \vartheta\big) = \frac{1}{4}(b-a)^t\Sigma^{-1}(b-a) = \frac{1}{4}\mu^t\Sigma^{-1}\mu,$$

mit $\tilde{\Theta}_0' = \{(c,c) \in \mathbb{R}^{2d}; a \in \mathbb{R}^d\} \times \mathcal{S}_d$ und $\vartheta = (a, b, \Sigma)$. Andererseits ergibt sich analog zur Rechnung für die Hotellingsche T^2-Testgröße aus Unterabschnitt 8.3

$$\limsup_{n\to\infty} \frac{1}{n}\log\beta_n(\alpha,\mu) \leq -\frac{1}{4}\mu^t\Sigma^{-1}\mu = -K(\tilde{\Theta}_0', \vartheta),$$

und damit, dass der Hotellingsche Zweistichproben-T^2-Test ein optimaler Test im Hodges-Lehmann Sinn für das Testproblem H gegen K ist, mit zugehögigem Hodges-Lehmann-Index

$$d_{T'}(\mu) = \frac{1}{4}\mu^t\Sigma^{-1}\mu,$$

woraus die Behauptung folgt. □

8.7 Bahadur-Steigung im unabhängigen Stichprobenfall

Der Hotellingsche Zweistichproben-T^2-Test ist auch ein optimaler Test im Bahadur Sinn für das Testproblem H gegen K. Dies zeigen wir wieder mit Lemma 2.3.2.

8.7.1 Satz: *Der Hotellingsche Zweistichproben-T^2-Test ist ein optimaler Test im Bahadur Sinn für das Testproblem H gegen K und es gilt mit $\mu = a - b$*

$$c_{T'}(\mu) = \frac{1}{2}\log(1 + \frac{1}{2}\mu^t\Sigma^{-1}\mu).$$

Beweis: Es ist

$$\lim_{n\to\infty} \frac{4d(n-1)}{n(2n-1-d)} T'_n = \mu^t\Sigma^{-1}\mu \text{ f.s.}$$

erfüllt. Mit der Markov-Ungleichung gilt für $t \geq 0$

$$\begin{aligned}
&P_0\left(\frac{4d(n-1)}{n(2n-1-d)} T'_n \geq t\right) \\
=&P_0\left(\frac{4d(n-1)}{n(2n-1-d)} \frac{W'_1/d}{W'_2/(2n-1-d)} \geq t\right) = P_0\left(\frac{n-1}{2n}W'_1 - \frac{t}{8}W'_2 \geq 0\right) \\
\leq& \mathrm{E}_0\left(\exp\left(\frac{n-1}{2n}W'_1\right)\right)\mathrm{E}_0\left(\exp\left(-\frac{t}{8}W'_2\right)\right) = n^{\frac{d}{2}}\left(1+\frac{t}{4}\right)^{-\frac{2n-1-d}{2}}.
\end{aligned}$$

Daraus folgt mit $\exp(-(1+t)) \leq \exp(-(1+t)/2)$ für alle $t \geq -1$

$$\limsup_{n\to\infty} \frac{1}{n}\log P_0\left(\frac{4d(n-1)}{n(2n-1-d)} T'_n \geq t\right) \leq -\log(1+\frac{t}{4}) \leq -\frac{1}{2}\log(1+\frac{t}{2}),\ t \geq 0.$$

Außerdem folgt analog zu den Überlegungen in den Beweisen von den Sätzen 8.4.1 und 8.6.1

$$K(\vartheta, \tilde{\Theta}'_0) = \frac{1}{2}\log(1 + \frac{1}{2}\mu^t\Sigma^{-1}\mu),$$

sodass aus Lemma 2.3.2 die Optimalität des Hotellingsche Zweistichproben-T^2-Tests im Bahadur Sinn für das Testproblem H gegen K folgt. □

8.8 Bahadur- und Hodges-Lehmann-Effizienz der Tests

Wir können nun die Ergebnisse aus den Unterabschnitten 8.3, 8.4, 8.6 und 8.7 zusammenfassen.

8.8.1 Satz: *Die asymptotische relative Effizienz im Hodges-Lehmann Sinn der Tests T und T' lautet mit $\mu = a - b$*

$$e^{\mathrm{HL}}_{T',T} = e^{\mathrm{HL}}_{T',T}(\mu) = \frac{\frac{1}{2}\mu^t\Sigma^{-1}\mu}{\mu^t\Gamma^{-1}\mu},$$

und die asymptotische relative Effizienz im Bahadur Sinn der Tests T und T' ist mit $\mu = a - b$ gegeben durch

$$e^{\mathrm{B}}_{T',T} = e^{\mathrm{B}}_{T',T}(\mu) = \frac{\log(1 + \frac{1}{2}\mu^t\Sigma^{-1}\mu)}{\log(1 + \mu^t\Gamma^{-1}\mu)}.$$

Beweis: Dies folgt aus den Ausführungen in Unterabschnitt 8.3 und aus den Sätzen 8.4.1, 8.6.1 und 8.7.1. □

8.8.2 Bemerkung: Für eine reelle Folge $(\nu_i)_{i\in\mathbb{N}}$ mit $\nu_i \neq 0$ für alle $i \in \mathbb{N}$ und $\lim_{i\to\infty} \nu_i = 0$ und für ein $\eta \in \mathbb{R}^d \setminus \{0\}$ betrachte die Alternativen $\mu_i := \nu_i\eta$, $i \in \mathbb{N}$. Dann ergibt sich für die lokale Hodges-Lehmann-Effizienz mit $\mu = a - b$

$$e^{\mathrm{LHL}}_{T',T}(\eta) = \frac{\frac{1}{2}\eta^t\Sigma^{-1}\eta}{\eta^t\Gamma^{-1}\eta},$$

und mit Hilfe der Regel von L'Hospital für die lokale Bahadur-Effizienz

$$\begin{aligned} e^{\mathrm{LB}}_{T',T}(\eta) &= \lim_{i\to\infty} e^{\mathrm{B}}_{T',T}(\mu_i) = \lim_{i\to\infty} \frac{\log(1 + \nu_i^2\frac{1}{2}\eta^t\Sigma^{-1}\eta)}{\log(1 + \nu_i^2\eta^t\Gamma^{-1}\eta)} \\ &= \lim_{i\to\infty} \frac{1 + \nu_i^2\eta^t\Gamma^{-1}\eta}{1 + \nu_i^2\frac{1}{2}\eta^t\Sigma^{-1}\eta} \frac{\eta^t\Sigma^{-1}\eta}{2\eta^t\Gamma^{-1}\eta} = \frac{\frac{1}{2}\eta^t\Sigma^{-1}\eta}{\eta^t\Gamma^{-1}\eta}. \end{aligned}$$

8.9 Pitman-Effizienz der Tests

Da es sich bei den betrachteten Tests jeweils um den Likelihood-Quotienten-Test handelt, und die in Unterabschnitt 4.2 genannten Bedingungen hier erfüllt sind, kann die Pitma-Effizienz der Tests T und T' mit Hilfe von Korollar 4.2.5 angegeben werden. Allerdings kann in dieser konkreten Situation die Pitman-Effizienz der Tests auch anders ermittelt werden, indem die in diesem Fall bekannten finiten Verteilungen der Testgrößen unter Alternativen, also die vorliegenden nichtzentralen F-Verteilungen, ausgenutzt werden. Dabei kann die Pitman-Effizienz für eine größere Klasse von Alternativen erhalten werden, als bei der Anwendung von Korollar 4.2.5.

8.9.1 Satz: *Betrachten wir Folgen von Alternativen $(\mu_i)_{i\in\mathbb{N}}$ definiert durch $\mu_i := \nu_i\eta$, $i \in \mathbb{N}$, mit einer reellen Folge $(\nu_i)_{i\in\mathbb{N}}$ mit $\nu_i \neq 0$ für alle $i \in \mathbb{N}$ und $\lim_{i\to\infty} \nu_i = 0$, sowie $\eta \in \mathbb{R}^d \setminus \{0\}$ beliebig. Die asymptotische relative Effizienz im Pitman Sinn der Tests T und T' lautet im Fall ihrer Existenz*

$$e^{\mathrm{P}}_{T',T} = e^{\mathrm{P}}_{T',T}(\eta) = \frac{\frac{1}{2}\eta^t\Sigma^{-1}\eta}{\eta^t\Gamma^{-1}\eta}.$$

Beweis: Um die asymptotische relative Effizienz im Pitman Sinn der Testgrößen T und T' zu bestimmen, halten wir zunächst fest, dass für $b \in [0, \infty)$

$$X_{n,b} := \frac{\sum_{i=1}^{d}(U_i + \sqrt{b/d})^2}{\sum_{i=1}^{n-d} V_i^2/(n-d)} \longrightarrow \sum_{i=1}^{d}(U_i + \sqrt{b/d})^2 \sim \chi_d^2(b) \text{ f.s. für } n \to \infty$$

und

$$T_n d \overset{v}{=} \frac{\sum_{i=1}^{d}(U_i + \sqrt{n\mu^t\Gamma^{-1}\mu/d})^2}{\sum_{i=1}^{n-d} V_i^2/(n-d)} = X_{n,n\mu'\Gamma^{-1}\mu},$$

gilt, wenn $U_1, \ldots, U_d, V_1, \ldots$ unabhängige und standardnormalverteilte Zufallsvariablen sind. Dies impliziert, dass die Verteilungsfunktion von $X_{n,b}$ für $n \to \infty$ gleichmäßig gegen die auf $[0, \infty)$ stetige und streng monoton wachsende Verteilungsfunktion der $\chi_d^2(b)$-Verteilung, welche wir mit G_b bezeichnen wollen, konvergiert. Ebenso gilt für $b \in [0, \infty)$

$$X'_{n,b} := \frac{\sum_{i=1}^{d}(U_i + \sqrt{b/d})^2}{\sum_{i=1}^{2n-1-d} V_i^2/(2n-1-d)} \longrightarrow \sum_{i=1}^{d}(U_i + \sqrt{b/d})^2 \sim \chi_d^2(b) \text{ f.s. für } n \to \infty,$$

und

$$T'_n d \overset{v}{=} \frac{\sum_{i=1}^{d}\left(U_i + \sqrt{n\mu^t\Sigma^{-1}\mu/(2d)}\right)^2}{\sum_{i=1}^{2n-1-d} V_i^2/(2n-1-d)} = X'_{n,\frac{n}{2}\mu'\Sigma^{-1}\mu},$$

was ebenfalls impliziert, dass die Verteilungsfunktion von $X'_{n,b}$ für $n \to \infty$ gleichmäßig gegen G_b, die stetige und auf $[0, \infty)$ streng monoton wachsende Verteilungsfunktion der $\chi_d^2(b)$-Verteilung, konvergiert. Damit folgt aus

$$P_0(T_n d \geq \mathrm{F}_{d,n-d}^{\alpha}\, d) = P_0(T'_n d \geq \mathrm{F}_{d,2n-1-d}^{\alpha}\, d) = \alpha \text{ für alle } n \in \mathbb{N}$$

sofort

$$\lim_{n\to\infty} F_{d,n-d}^{\alpha} d = \lim_{n\to\infty} F_{d,2n-1-d}^{\alpha} d = G_0^{-1}(1-\alpha).$$

Es sei ein $\beta \in (0, 1-\alpha)$ und die Folge $(\mu_i)_{i\in\mathbb{N}}$ entsprechend der Voraussetzung gegeben. Wie in Unterabschnitt 2.4 definiert, gilt im Fall der Existenz der Pitman-Effizienz

$$\lim_{i\to\infty} P_{\mu_i}(T_{N_{T,i}} < \mathrm{F}_{d,N_{T,i}-d}^{\alpha}) = \lim_{i\to\infty} P_{\mu_i}(T'_{N_{T',i}} < \mathrm{F}_{d,2N_{T',i}-1-d}^{\alpha}) = \beta.$$

Daraus folgt aufgrund der vorherigen Überlegungen die Existenz eines $c \in (0, \infty)$ mit

$$\lim_{i\to\infty} N_{T,i}\mu_i^t\Gamma^{-1}\mu_i = \lim_{i\to\infty} N_{T',i}\mu_i^t\Sigma^{-1}\mu_i/2 = c,$$

wobei c durch die Eigenschaft

$$\beta = G_c(G_0^{-1}(1-\alpha))$$

eindeutig bestimmt ist. Dies sieht man wie folgt: Für $\nu_1, \nu_2 \in \mathbb{N}$ und $b \geq 0$ sei $Z \sim F_{\nu_1,\nu_2}(b)$. Für jedes $x > 0$ ist die Abbildung

$$b \longmapsto P_b(Z \leq x),\ b \geq 0,$$

stetig und streng monoton fallend, siehe Satz 2.36 in [41]. Darüber hinaus ist

$$\lim_{b\to\infty} P_b(Z \leq x) = 0.$$

Daher ist für jedes $x > 0$ die Abbildung

$$b \longmapsto G_b(x),\ b \geq 0,$$

stetig und streng monoton fallend mit

$$\lim_{b\to\infty} G_b(x) = 0.$$

Wegen $\beta \in (0, 1-\alpha)$ existiert ein eindeutig bestimmtes $c \in (0,\infty)$ mit

$$\lim_{n\to\infty} P\left(\frac{\sum_{i=1}^d (U_i + \sqrt{c/d})^2}{\sum_{i=1}^{n-d} V_i^2/(n-d)} < d\,\mathrm{F}_{d,n-d}^{\alpha}\right) = G_c\big(G_0^{-1}(1-\alpha)\big) = \beta.$$

Angenommen, $(N_{T,i}\mu_i^t\Gamma^{-1}\mu_i)_{i\in\mathbb{N}}$ konvergiert nicht gegen dieses c, d.h.

$$\exists\ \varepsilon > 0\ \forall\ i_\varepsilon \in \mathbb{N}\ \exists\ i \geq i_\varepsilon :\ |N_{T,i}\mu_i^t\Gamma^{-1}\mu_i - c| > \varepsilon.$$

Dann existiert ein $\varepsilon > 0$ und eine Teilfolge $(N_{T,i_k}\mu_{i_k}^t\Gamma^{-1}\mu_{i_k})_{k\in\mathbb{N}}$ mit

$$N_{T,i_k}\mu_{i_k}^t\Gamma^{-1}\mu_{i_k} > c+\varepsilon\ \forall\ k \in \mathbb{N} \text{ oder } N_{T,i_k}\mu_{i_k}^t\Gamma^{-1}\mu_{i_k} < c-\varepsilon\ \forall\ k \in \mathbb{N}.$$

Ohne Einschränkung gelte die erste Eigenschaft und es sei die genannte Teilfolge die Folge selbst. Dann erhält man

$$\begin{aligned}
\underbrace{P_{\mu_i}(T_{N_{T,i}} < \mathrm{F}_{d,N_{T,i}-d}^{\alpha})}_{\to\beta} &= P_{\mu_i}(T_{N_{T,i}}d < d\,\mathrm{F}_{d,N_{T,i}-d}^{\alpha}) \\
&= P\left(\frac{\sum_{i=1}^d (U_i + \sqrt{N_{T,i}\mu_i^t\Gamma^{-1}\mu_i/d})^2}{\sum_{i=1}^{N_{T,i}-d} V_i^2/(N_{T,i}-d)} < d\,\mathrm{F}_{d,N_{T,i}-d}^{\alpha}\right) \\
&< P\left(\frac{\sum_{i=1}^d (U_i + \sqrt{(c+\varepsilon)/d})^2}{\sum_{i=1}^{N_{T,i}-d} V_i^2/(N_{T,i}-d)} < d\,\mathrm{F}_{d,N_{T,i}-d}^{\alpha}\right) \\
&\longrightarrow G_{c+\varepsilon}\big(G_0^{-1}(1-\alpha)\big) < G_c\big(G_0^{-1}(1-\alpha)\big) = \beta \text{ für } i \to \infty,
\end{aligned}$$

also einen Widerspruch. Analog kann $\lim_{i\to\infty} N_{T',i}\mu_i^t\Sigma^{-1}\mu_i/2 = c$ gezeigt werden. Also gilt

$$\lim_{i\to\infty} \frac{N_{T',i}}{N_{T,i}} \frac{\mu_i^t\Sigma^{-1}\mu_i/2}{\mu_i^t\Gamma^{-1}\mu_i} = 1,$$

woraus schließlich für Folgen von Alternativen in der beschriebenen Form

$$e^{\mathrm{P}}_{T',T} = e^{\mathrm{P}}_{T',T}(\eta) = \frac{\frac{1}{2}\eta^t\Sigma^{-1}\eta}{\eta^t\Gamma^{-1}\eta}$$

als asymptotische relative Effizienz im Pitman Sinn der Testgrößen T und T' folgt. $\square$

8.9.2 Bemerkung: Es lässt sich unter Beachtung von Bemerkung 2.4.1 mit den Überlegungen präsentiert in 3.8.3 in [32] zeigen, dass die Pitman-Effizienz für $\nu_i = 1/\sqrt{i}$, $i \in \mathbb{N}$, existiert.

8.10 Volumen-Effizienz der Tests

Das in Unterabschnitt 3 vorgestellte Konzept der Volumen-Effizienz kann in offensichtlicher Weise auf die vorliegende Situation der zusammengesetzten Hypothese erweitert werden. Siehe hierzu die Erklärung am Anfang des Unterabschnitts 4.4.

Im verbundenen Stichprobenfall gilt bei zugrunde liegendem $\mu \in \mathbb{R}^d$

$$\frac{n(n-d)}{d(n-1)}(\overline{\zeta}_n - \mu)^t S_n^{-1}(\overline{\zeta}_n - \mu) \sim \mathrm{F}_{d,n-d}.$$

Daher ist für $\alpha \in (0,1)$

$$\Big\{\mu \in \mathbb{R}^d; \frac{n(n-d)}{d(n-1)}(\overline{\zeta}_n - \mu)^t S_n^{-1}(\overline{\zeta}_n - \mu) \le \mathrm{F}^{\alpha}_{d,n-d}\Big\}.$$

ein Konfidenzbereich, genauer ein Konfidenzellipsoid, für den unbekannten Parameter $\mu \in \mathbb{R}^d$ zum Konfidenzniveau $1-\alpha$. Mit Hilfe dieses Konfidenzellipsoids lässt sich der Hotellingsche T^2-Test formulieren.

Im unabhängigen Stichprobenfall gilt bei zugrunde liegendem $\mu \in \mathbb{R}^d$ analog

$$\frac{n(2n-1-d)}{4d(n-1)}(\overline{X'_n} - \overline{Y'_n} - \mu)^t (S_n^{x,y})^{-1}(\overline{X'_n} - \overline{Y'_n} - \mu) \sim \mathrm{F}_{d,2n-1-d},$$

sodass sich als zugehöriger Konfidenzellipsoid für den unbekannten Parameter $\mu \in \mathbb{R}^d$ zum Konfidenzniveau $1-\alpha$

$$\Big\{\mu \in \mathbb{R}^d; \frac{n(2n-1-d)}{4d(n-1)}(\overline{X'_n} - \overline{Y'_n} - \mu)^t (S_n^{x,y})^{-1}(\overline{X'_n} - \overline{Y'_n} - \mu) \le \mathrm{F}^{\alpha}_{d,2n-1-d}\Big\}$$

ergibt. Mit Hilfe dieses Konfidenzellipsoiden lässt sich der Hotellingsche Zweistichproben-T^2-Test formulieren.

8.10.1 Satz: *Die Volumen-Effizienz der Testgrößen T_n und T'_n hängt nicht mehr vom unbekannten Parameter $\mu \in \mathbb{R}^d$ ab und es ist*

$$e^{\mathrm{vol}}_{T',T,n} = \frac{(2n-1-d)^d}{2^d(n-d)^d} \frac{(\mathrm{F}^\alpha_{d,n-d})^d}{(\mathrm{F}^\alpha_{d,2n-1-d})^d} \frac{\det(\Gamma)}{\det(\Sigma)} \left(\frac{\Gamma(\frac{n}{2})\Gamma(\frac{2n-d-1}{2})}{\Gamma(\frac{n-d}{2})\Gamma(\frac{2n-1}{2})} \right)^2.$$

Beweis: Im verbundenen Stichprobenfall besitzt das Konfidenzellipsoid das Volumen

$$\mathrm{Vol}_{T,n} = \frac{\pi^{d/2}}{\Gamma(d/2+1)} \sqrt{\frac{d^d(n-1)^d}{n^d(n-d^d)} (\mathrm{F}^\alpha_{d,n-d})^d \det(S_n)},$$

welches nicht mehr vom unbekannten Parameter $\mu \in \mathbb{R}^d$ abhängt. Zur Berechnung des Volumens eines Ellipsoiden siehe Beispiel 1 in Abschnitt 9.1 in [21]. Bekanntlich besitzt S_n als empirische Kovarianzmatrix basierend auf entsprechenden normalverteilten Zufallsvariablen die Wishart-Verteilung $\mathrm{W}_d(n-1, 1/(n-1)\Gamma)$, siehe hierzu [28]. Nach Theorem 3.2.15 aus [28] gilt dann

$$(n-1)^d \frac{\det(S_n)}{\det(\Gamma)} \overset{\mathrm{v}}{=} \prod_{i=1}^{d} C_{n-i},$$

mit unabhängigen Zufallsvariablen $C_{n-d}, \ldots, C_{n-1}$, wobei C_j Chi-Quadrat verteilt sei mit j Freiheitsgraden, $j = n-d, \ldots, n-1$. Also ist

$$\mathrm{E}(\mathrm{Vol}_{T,n}) = \frac{\pi^{d/2}}{\Gamma(d/2+1)} \sqrt{\frac{d^d}{n^d(n-d)^d} (\mathrm{F}^\alpha_{d,n-d})^d \det(\Gamma)} \prod_{i=1}^{d} \mathrm{E}\left(\sqrt{C_{n-i}}\right)$$

der Erwartungswert des Volumens des Konfidenzellipsoiden. Besitzt C_k eine Chi-Quadrat-Verteilung mit k Freiheitsgeraden, $k > 1$, so ist $\sqrt{C_k}$ Chi-verteilt mit k Freiheitsgeraden, d.h. $\sqrt{C_k}$ besitzt die Dichte

$$x \longmapsto \frac{2^{1-\frac{k}{2}} x^{k-1} e^{-\frac{x^2}{2}}}{\Gamma(\frac{k}{2})}, \text{ für } x > 0, \text{ und } 0 \text{ sonst},$$

und damit ist

$$\mathrm{E}\left(\sqrt{C_k}\right) = \sqrt{2} \frac{\Gamma(\frac{k+1}{2})}{\Gamma(\frac{k}{2})}.$$

Schließlich erhält man für das erwartete Volumen die Darstellung

$$\mathrm{E}(\mathrm{Vol}_{T,n}) = \frac{\pi^{d/2}}{\Gamma(d/2+1)} \sqrt{\frac{d^d}{n^d(n-d)^d} (\mathrm{F}^\alpha_{d,n-d})^d 2^d \det(\Gamma)} \frac{\Gamma(\frac{n}{2})}{\Gamma(\frac{n-d}{2})}.$$

Im unabhängigen Stichprobenfall besitzt das Konfidenzellipsoid das Volumen

$$\mathrm{Vol}_{T',n} = \frac{\pi^{d/2}}{\Gamma(d/2+1)}\sqrt{\frac{4^d d^d (n-1)^d}{n^d(2n-1-d)^d}(\mathrm{F}^{\alpha}_{d,2n-1-d})^d \det(S_n^{x,y})}.$$

Dieses Volumen hängt ebenfalls nicht vom unbekannten Parameter $\mu \in \mathbb{R}^d$ ab. Da S_n^x und S_n^y jeweils die Wishart-Verteilung $\mathrm{W}_d(n-1, 1/(n-1)\Sigma)$ besitzen und unabhängig sind, besitzt nach Theorem 3.2.4 aus [28] $2S_n^{x,y}$ die Wishart-Verteilung $\mathrm{W}_d(2(n-1), 1/(n-1)\Sigma)$. Ebenfalls nach Theorem 3.2.15 aus [28] gilt dann

$$(n-1)^d \frac{\det(2S_n^{x,y})}{\det(\Sigma)} \stackrel{\mathrm{v}}{=} \prod_{i=1}^{d} C_{2n-1-i},$$

mit unabhängigen Zufallsvariablen $C_{2n-1-d}, \ldots, C_{2n-2}$, wobei C_j Chi-Quadrat verteilt sei mit j Freiheitsgraden, $j = 2n-1-d, \ldots, 2n-2$. Also gilt

$$\begin{aligned}\mathrm{E}(\mathrm{Vol}_{T',n}) &= \frac{\pi^{d/2}}{\Gamma(d/2+1)}\sqrt{\frac{2^d d^d}{n^d(2n-1-d)^d}(\mathrm{F}^{\alpha}_{d,2n-1-d})^d \det(\Sigma)} \prod_{i=1}^{d} \mathrm{E}\left(\sqrt{C_{2n-1-i}}\right) \\ &= \frac{\pi^{d/2}}{\Gamma(d/2+1)}\sqrt{\frac{4^d d^d}{n^d(2n-1-d)^d}(\mathrm{F}^{\alpha}_{d,2n-1-d})^d \det(\Sigma)} \frac{\Gamma(\frac{2n-1}{2})}{\Gamma(\frac{2n-d-1}{2})}.\end{aligned}$$

So ergibt sich für die Volumen-Effizienz der Testgrößen T_n und T'_n

$$\begin{aligned}e^{\mathrm{vol}}_{T',T,n} &= \left(\frac{\frac{\pi^{d/2}}{\Gamma(d/2+1)}\sqrt{\frac{2^d d^d}{n^d(n-d)^d}(\mathrm{F}^{\alpha}_{d,n-d})^d \det(\Gamma)}\frac{\Gamma(\frac{n}{2})}{\Gamma(\frac{n-d}{2})}}{\frac{\pi^{d/2}}{\Gamma(d/2+1)}\sqrt{\frac{4^d d^d}{n^d(2n-1-d)^d}(\mathrm{F}^{\alpha}_{d,2n-1-d})^d \det(\Sigma)}\frac{\Gamma(\frac{2n-1}{2})}{\Gamma(\frac{2n-d-1}{2})}}\right)^2 \\ &= \frac{(2n-1-d)^d}{2^d(n-d)^d}\frac{(\mathrm{F}^{\alpha}_{d,n-d})^d}{(\mathrm{F}^{\alpha}_{d,2n-1-d})^d}\frac{\det(\Gamma)}{\det(\Sigma)}\left(\frac{\Gamma(\frac{n}{2})\Gamma(\frac{2n-d-1}{2})}{\Gamma(\frac{n-d}{2})\Gamma(\frac{2n-1}{2})}\right)^2,\end{aligned}$$

und so die Behauptung. □

8.10.2 Korollar: *Die asymptotische Volumen-Effizienz der Testgrößen T_n und T'_n hängt nicht mehr vom Konfidenzniveau $1-\alpha$ ab und es ist*

$$e^{\mathrm{Lvol}}_{T',T} = \frac{\det(\Gamma)}{\det(2\Sigma)} = \lim_{n\to\infty} e^{\mathrm{vol}}_{T',T,n}.$$

Beweis: Mit den Überlegungen und Bezeichnungen aus dem Beweis von Satz 8.10.1 gilt mit dem starken Gesetz der großen Zahlen

$$\begin{aligned}\sqrt{\frac{n^{2d}}{\prod_{i=1}^{d}(n-i)}\frac{1}{(\mathrm{F}^{\alpha}_{d,n-d})^{d}}\mathrm{Vol}_{T,n}} &\overset{\mathrm{v}}{=} \frac{\pi^{d/2}}{\Gamma(d/2+1)}\sqrt{\frac{d^d n^{2d}}{n^d(n-d)^d}\det(\Gamma)}\prod_{i=1}^{d}\sqrt{\frac{C_{n-i}}{n-i}} \\ &\overset{\mathrm{v}}{=} \frac{\pi^{d/2}}{\Gamma(d/2+1)}\sqrt{\frac{d^d n^{2d}}{n^d(n-d)^d}\det(\Gamma)}\prod_{i=1}^{d}\sqrt{\frac{\sum_{j=1}^{n-i}U_{i,j}^2}{n-i}} \\ &\longrightarrow \frac{\pi^{d/2}}{\Gamma(d/2+1)}\sqrt{d^d\det(\Gamma)}\ P\text{-f.s. für } n\to\infty,\end{aligned}$$

mit unabhängigen Folgen von unabhängigen standardnormalverteilten Zufallsvariablen $(U_{i,j})_{j\in\mathbb{N}}$, $i=1,\ldots,d$. Ebenfalls mit den Überlegungen und Bezeichnungen aus dem Beweis von Satz 8.10.1 und mit $\lim_{n\to\infty}\mathrm{F}^{\alpha}_{d,2n-1-d}/\mathrm{F}^{\alpha}_{d,n-d}=1$, was bereits im Beweis von Satz 8.9.1 begründet wurde, gilt ebenso mit dem starken Gesetz der großen Zahlen

$$\begin{aligned}&\sqrt{\frac{n^{2d}}{\prod_{i=1}^{d}(n-i)}\frac{1}{(\mathrm{F}^{\alpha}_{d,n-d})^{d}}\mathrm{Vol}_{T',n}} \\ &\overset{\mathrm{v}}{=}\sqrt{\frac{n^{2d}}{\prod_{i=1}^{d}(n-i)}\frac{1}{(\mathrm{F}^{\alpha}_{d,n-d})^{d}}}\frac{\pi^{d/2}}{\Gamma(d/2+1)}\sqrt{\frac{2^d d^d}{n^d(2n-1-d)^d}(\mathrm{F}^{\alpha}_{d,2n-1-d})^d\det(\Sigma)}\prod_{i=1}^{d}\sqrt{\frac{C_{2n-1-i}}{2n-1-i}} \\ &\overset{\mathrm{v}}{=}\sqrt{\frac{\prod_{i=1}^{d}(2n-1-i)}{\prod_{i=1}^{d}(n-i)}}\frac{\pi^{d/2}}{\Gamma(d/2+1)}\sqrt{\frac{2^d d^d n^{2d}}{n^d(2n-1-d)^d}\frac{(\mathrm{F}^{\alpha}_{d,2n-1-d})^d}{(\mathrm{F}^{\alpha}_{d,n-d})^d}\det(\Sigma)}\prod_{i=1}^{d}\sqrt{\frac{\sum_{j=1}^{2n-1-i}U_{i,j}^2}{2n-1-i}} \\ &\longrightarrow \sqrt{2^d}\frac{\pi^{d/2}}{\Gamma(d/2+1)}\sqrt{d^d\det(\Sigma)}\ P\text{-f.s. für } n\to\infty,\end{aligned}$$

sodass

$$e^{\mathrm{Lvol}}_{T',T}=\frac{\det(\Gamma)}{2^d\det(\Sigma)}=\frac{\det(\Gamma)}{\det(2\Sigma)}$$

ist.

Weiter gilt für $y\in\mathbb{R}$

$$\frac{\Gamma(x)}{\Gamma(x+y)}\sim\frac{1}{x^y}\ \text{für } x\to\infty.$$

Damit folgt

$$\lim_{n\to\infty}\left(\frac{\Gamma(\frac{n}{2})\Gamma(\frac{2n-d-1}{2})}{\Gamma(\frac{n-d}{2})\Gamma(\frac{2n-1}{2})}\right)^2=\lim_{n\to\infty}\left(\frac{\Gamma(\frac{n}{2})}{\Gamma(\frac{n}{2}-\frac{d}{2})}\frac{\Gamma(\frac{2n-1}{2}-\frac{d}{2})}{\Gamma(\frac{2n-1}{2})}\right)^2=\frac{1}{2^d}.$$

Zudem gilt

$$\lim_{n\to\infty} \frac{(2n-1-d)^d}{2^d(n-d)^d} \frac{(\mathrm{F}^{\alpha}_{d,n-d})^d}{(\mathrm{F}^{\alpha}_{d,2n-1-d})^d} = 1,$$

sodass auch

$$\lim_{n\to\infty} e^{\mathrm{vol}}_{T',T,n} = \frac{\det(\Gamma)}{\det(2\Sigma)}$$

ist. □

8.11 Vergleich der Testverfahren

Anhand der oben berechneten Hodges-Lehmann-, Bahadur- und Pitman-Effizienz können wir nun die beiden Verfahren, unabhängige und verbundene Stichprobenerhebung, vergleichen. Wie in Lemma 4.5.1 können dazu die Matrix

$$E := (2\Sigma)^{-1}\Gamma = \mathrm{I}_d - \frac{1}{2}(\Sigma_1 + \Sigma_2)$$

und ihre nach wachsender Größe geordneten Eigenwerte $\lambda_1 \leq \cdots \leq \lambda_d$, herangezogen werden.

8.11.1 Korollar: *Es ist*

$$\inf_{\mu\in\mathbb{R}^d\setminus\{0\}} e^{\mathrm{B}}_{T',T}(\mu) \leq 1 \leq \sup_{\mu\in\mathbb{R}^d\setminus\{0\}} e^{\mathrm{B}}_{T',T}(\mu).$$

Für die Pitman-Effizienz betrachten wir Folgen von Alternativen $(\mu_i)_{i\in\mathbb{N}}$ definiert durch $\mu_i := \nu_i\eta$, $i \in \mathbb{N}$, wobei die reelle Folge $(\nu_i)_{i\in\mathbb{N}}$ mit $\nu_i \neq 0$ für $i \in \mathbb{N}$ und $\lim_{i\to\infty} \nu_i = 0$, sowie $\eta \in \mathbb{R}^d \setminus \{0\}$ beliebig gewählt werden können. Dann gilt:

a) $\sup_{\mu\in\mathbb{R}^d\setminus\{0\}} e^{\mathrm{HL}}_{T',T}(\mu) < 1 \Leftrightarrow \sup_{\eta\in\mathbb{R}^d\setminus\{0\}} e^{\mathrm{P}}_{T',T}(\eta) < 1 \Leftrightarrow \lambda_d < 1.$

b) $\sup_{\mu\in\mathbb{R}^d\setminus\{0\}} e^{\mathrm{HL}}_{T',T}(\mu) \leq 1 \Leftrightarrow \sup_{\mu\in\mathbb{R}^d\setminus\{0\}} e^{\mathrm{B}}_{T',T}(\mu) = 1 \Leftrightarrow \sup_{\eta\in\mathbb{R}^d\setminus\{0\}} e^{\mathrm{P}}_{T',T}(\eta) \leq 1 \Leftrightarrow \lambda_d \leq 1.$

c) $\inf_{\mu\in\mathbb{R}^d\setminus\{0\}} e^{\mathrm{HL}}_{T',T}(\mu) \geq 1 \Leftrightarrow \inf_{\mu\in\mathbb{R}^d\setminus\{0\}} e^{\mathrm{B}}_{T',T}(\mu) = 1 \Leftrightarrow \inf_{\eta\in\mathbb{R}^d\setminus\{0\}} e^{\mathrm{P}}_{T',T}(\eta) \geq 1 \Leftrightarrow \lambda_1 \geq 1.$

d) $\inf_{\mu\in\mathbb{R}^d\setminus\{0\}} e^{\mathrm{HL}}_{T',T}(\mu) > 1 \Leftrightarrow \inf_{\eta\in\mathbb{R}^d\setminus\{0\}} e^{\mathrm{P}}_{T',T}(\eta) > 1 \Leftrightarrow \lambda_1 > 1.$

Beweis: Wir benötigen zum Beweis der Aussagen die Gestalt der Effizienzen gegeben durch die Sätze 8.8.1 und 8.9.1. Es sei eine Folge $(\mu_i)_{i\in\mathbb{N}}$ gemäß $\mu_i := a_i\theta$, gegeben, $i \in \mathbb{N}$, mit einer reellen Folge $(a_i)_{i\in\mathbb{N}}$ mit $a_i \neq 0$ und $\lim_{i\to\infty} a_i = \infty$, und $\theta \in \mathbb{R}^d \setminus \{0\}$. Mit Hilfe der Regel von L'Hospital und der positiven Definitheit der beteiligten Matrizen erhält man

$$\begin{aligned}\lim_{i\to\infty} e^{\mathrm{B}}_{T',T}(\mu_i) = \lim_{i\to\infty} \frac{c_{T'}(\mu_i)}{c_T(\mu_i)} = \lim_{i\to\infty} \frac{\log(1 + a_i^2\frac{1}{2}\theta'\Sigma^{-1}\theta)}{\log(1 + a_i^2\theta'\Gamma^{-1}\theta)} = \lim_{n\to\infty} \frac{1 + a_i^2\theta'\Gamma^{-1}\theta}{1 + a_i^2\frac{1}{2}\theta'\Sigma^{-1}\theta} \frac{\theta'\Sigma^{-1}\theta}{2\theta'\Gamma^{-1}\theta} \\ = \lim_{i\to\infty} \frac{2\theta'\Gamma^{-1}\theta}{\theta'\Sigma^{-1}\theta} \frac{\theta'\Sigma^{-1}\theta}{2\theta'\Gamma^{-1}\theta} = 1,\end{aligned}$$

womit

$$\inf_{\mu\in\mathbb{R}^d\setminus\{0\}} e^{\mathrm{B}}_{T',T}(\mu) \le 1 \le \sup_{\mu\in\mathbb{R}^d\setminus\{0\}} e^{\mathrm{B}}_{T',T}(\mu)$$

folgt. Da weiter die Äquivalenz

$$e^{\mathrm{B}}_{T',T} = \frac{c_{T'}}{c_T} = \frac{\frac{1}{2}\log(1+\frac{1}{2}\mu'\Sigma^{-1}\mu)}{\frac{1}{2}\log(1+\mu'\Gamma^{-1}\mu)} \begin{Bmatrix} \le \\ \ge \end{Bmatrix} 1 \iff \frac{\frac{1}{2}\mu'\Sigma^{-1}\mu}{\mu'\Gamma^{-1}\mu} = e^{\mathrm{HL}}_{T',T} \begin{Bmatrix} \le \\ \ge \end{Bmatrix} 1$$

gilt, folgt die Behauptung mit Lemma 4.5.1. □

8.11.2 Beispiel: Es sei $\Sigma_1 + \Sigma_2$ eine Diagonalmatrix mit Diagonalelementen σ_i, $i = 1,\dots,d$. Dann besitzt E die Eigenwerte

$$\lambda_i = 1 - \frac{1}{2}\sigma_i,\ i = 1,\dots,d.$$

In diesem Fall gilt:

a) $\sup_{\mu\in\mathbb{R}^d\setminus\{0\}} e^{\mathrm{HL}}_{T',T}(\mu) < 1 \Leftrightarrow \sup_{\eta\in\mathbb{R}^d\setminus\{0\}} e^{\mathrm{P}}_{T',T}(\eta) < 1 \Leftrightarrow \sigma_i > 0$ für $i = 1,\dots,d$.

b) $\sup_{\mu\in\mathbb{R}^d\setminus\{0\}} e^{\mathrm{HL}}_{T',T}(\mu) \le 1 \Leftrightarrow \sup_{\mu\in\mathbb{R}^d\setminus\{0\}} e^{\mathrm{B}}_{T',T}(\mu) = 1 \Leftrightarrow \sup_{\eta\in\mathbb{R}^d\setminus\{0\}} e^{\mathrm{P}}_{T',T}(\eta) \le 1 \Leftrightarrow \sigma_i \ge 0$ für $i = 1,\dots,d$.

c) $\inf_{\mu\in\mathbb{R}^d\setminus\{0\}} e^{\mathrm{HL}}_{T',T}(\mu) \ge 1 \Leftrightarrow \inf_{\mu\in\mathbb{R}^d\setminus\{0\}} e^{\mathrm{B}}_{T',T}(\mu) = 1 \Leftrightarrow \inf_{\eta\in\mathbb{R}^d\setminus\{0\}} e^{\mathrm{P}}_{T',T}(\eta) \ge 1 \Leftrightarrow \sigma_i \le 0$ für $i = 1,\dots,d$.

d) $\inf_{\mu\in\mathbb{R}^d\setminus\{0\}} e^{\mathrm{HL}}_{T',T}(\mu) > 1 \Leftrightarrow \inf_{\eta\in\mathbb{R}^d\setminus\{0\}} e^{\mathrm{P}}_{T',T}(\eta) > 1 \Leftrightarrow \sigma_i < 0$ für $i = 1,\dots,d$.

8.11.3 Beispiel: Sei $\Sigma_1 + \Sigma_2$ symmetrisch. Für das Spektrum von E gilt

$$\sigma(E) = \Big\{\lambda + 1; \lambda \in \sigma\Big(-\frac{1}{2}(\Sigma_1 + \Sigma_2)\Big)\Big\}.$$

Bei quadratischen symmetrischen Matrizen lässt sich bekanntlich die positive Definitheit bzw. die negative Definitheit bzw. die Semidefinitheit durch die Vorzeichen der Eigenwerte charakterisieren. Daher gilt:

a) $\sup_{\mu\in\mathbb{R}^d\setminus\{0\}} e^{\mathrm{HL}}_{T',T}(\mu) < 1 \Leftrightarrow \sup_{\eta\in\mathbb{R}^d\setminus\{0\}} e^{\mathrm{P}}_{T',T}(\eta) < 1 \Leftrightarrow \Sigma_1 + \Sigma_2$ ist positiv definit.

b) $\sup_{\mu\in\mathbb{R}^d\setminus\{0\}} e^{\mathrm{HL}}_{T',T}(\mu) \le 1 \Leftrightarrow \sup_{\mu\in\mathbb{R}^d\setminus\{0\}} e^{\mathrm{B}}_{T',T}(\mu) = 1 \Leftrightarrow \sup_{\eta\in\mathbb{R}^d\setminus\{0\}} e^{\mathrm{P}}_{T',T}(\eta) \le 1 \Leftrightarrow \Sigma_1 + \Sigma_2$ ist positiv semidefinit.

c) $\inf_{\mu\in\mathbb{R}^d\setminus\{0\}} e^{\mathrm{HL}}_{T',T}(\mu) \ge 1 \Leftrightarrow \inf_{\mu\in\mathbb{R}^d\setminus\{0\}} e^{\mathrm{B}}_{T',T}(\mu) = 1 \Leftrightarrow \inf_{\eta\in\mathbb{R}^d\setminus\{0\}} e^{\mathrm{P}}_{T',T}(\eta) \ge 1 \Leftrightarrow \Sigma_1 + \Sigma_2$ ist negativ semidefinit.

d) $\inf_{\mu\in\mathbb{R}^d\setminus\{0\}} e^{\mathrm{HL}}_{T',T}(\mu) > 1 \Leftrightarrow \inf_{\eta\in\mathbb{R}^d\setminus\{0\}} e^{\mathrm{P}}_{T',T}(\eta) > 1 \Leftrightarrow \Sigma_1 + \Sigma_2$ ist negativ definit.

Betrachten wir nun die Volumen-Effizienz der Tests. Dann können wir folgendes notieren:

8.11.4 Korollar: *Es gilt*

$$e^{\mathrm{vol}}_{T',T,n} \begin{Bmatrix} < \\ = \\ > \end{Bmatrix} 1 \iff \det\big(\mathrm{I}_d - \frac{1}{2}(\Sigma_1 + \Sigma_2)\big) \begin{Bmatrix} < \\ = \\ > \end{Bmatrix} k_n,$$

mit

$$k_n := \frac{(n-d)^d}{(2n-1-d)^d} \frac{(\mathrm{F}^{\alpha}_{d,2n-1-d})^d}{(\mathrm{F}^{\alpha}_{d,n-d})^d} \frac{\Gamma(\frac{n-d}{2})^2 \Gamma(\frac{2n-1}{2})^2}{\Gamma(\frac{n}{2})^2 \Gamma(\frac{2n-d-1}{2})^2},$$

und es gilt

$$e^{\mathrm{Lvol}}_{T',T} \begin{Bmatrix} < \\ = \\ > \end{Bmatrix} 1 \iff \det\big(\mathrm{I}_d - \frac{1}{2}(\Sigma_1 + \Sigma_2)\big) \begin{Bmatrix} < \\ = \\ > \end{Bmatrix} 1.$$

Beweis: Mit Hilfe von Satz 8.10.1 erhalten wir für die Volumen-Effizienz der Testgrößen T_n und T'_n

$$\begin{aligned} e^{\mathrm{vol}}_{T',T,n} & \frac{(2n-1-d)^d}{2^d(n-d)^d} \frac{(\mathrm{F}^{\alpha}_{d,n-d})^d}{(\mathrm{F}^{\alpha}_{d,2n-1-d})^d} \frac{\det\Big(\Sigma\big(2\,\mathrm{I}_d - (\Sigma_1 + \Sigma_2)\big)\Big)}{\det(\Sigma)} \left(\frac{\Gamma(\frac{n}{2})\Gamma(\frac{2n-d-1}{2})}{\Gamma(\frac{n-d}{2})\Gamma(\frac{2n-1}{2})}\right)^2 \\ &= \frac{(2n-1-d)^d}{(n-d)^d} \frac{(\mathrm{F}^{\alpha}_{d,n-d})^d}{(\mathrm{F}^{\alpha}_{d,2n-1-d})^d} \det\big(\mathrm{I}_d - \frac{1}{2}(\Sigma_1 + \Sigma_2)\big) \left(\frac{\Gamma(\frac{n}{2})\Gamma(\frac{2n-d-1}{2})}{\Gamma(\frac{n-d}{2})\Gamma(\frac{2n-1}{2})}\right)^2, \end{aligned}$$

woraus die erste Behauptung folgt. Mit Hilfe von Korollar 8.10.2 erhalten wir für die asymptotische Volumen-Effizienz der Testgrößen T_n und T'_n

$$e^{\mathrm{Lvol}}_{T',T} = \frac{\det\Big(\Sigma\big(2\,\mathrm{I}_d - (\Sigma_1 + \Sigma_2)\big)\Big)}{\det(2\Sigma)} = \det\big(\mathrm{I}_d - \frac{1}{2}(\Sigma_1 + \Sigma_2)\big),$$

woraus die zweite Behauptung folgt. □

8.11.5 Bemerkung: Auch die asymptotische Volumen-Effizienz hängt nur von der Matrix $\mathrm{I}_d - \frac{1}{2}(\Sigma_1 + \Sigma_2)$ ab, die auch beim Studium der Hodges-Lehmann-, Bahadur- und Pitman-Effizienz die entscheidende Rolle spielt.

8.11.6 Korollar: *Für die Pitman-Effizienz betrachten wir Folgen von Alternativen $(\mu_i)_{i\in\mathbb{N}}$ definiert durch $\mu_i := \nu_i \eta$, $i \in \mathbb{N}$, wobei die reelle Folge $(\nu_i)_{i\in\mathbb{N}}$, mit $\nu_i \neq 0$ für alle $i \in \mathbb{N}$ und $\lim_{i\to\infty} \nu_i = 0$ sowie $\eta \in \mathbb{R}^d \setminus \{0\}$ beliebig gewählt werden können. Dann gilt:*

a) $\sup_{\mu\in\mathbb{R}^d\setminus\{0\}} e^{\mathrm{HL}}_{T',T}(\mu) < 1 \Leftrightarrow \sup_{\eta\in\mathbb{R}^d\setminus\{0\}} e^{\mathrm{P}}_{T',T}(\eta) < 1 \Leftrightarrow \lambda_d < 1 \Rightarrow e^{\mathrm{Lvol}}_{T',T} < 1.$

b) $\sup_{\mu\in\mathbb{R}^d\setminus\{0\}} e^{\mathrm{HL}}_{T',T}(\mu) \leq 1 \Leftrightarrow \sup_{\mu\in\mathbb{R}^d\setminus\{0\}} e^{\mathrm{B}}_{T',T}(\mu) = 1 \Leftrightarrow \sup_{\eta\in\mathbb{R}^d\setminus\{0\}} e^{\mathrm{P}}_{T',T}(\eta) \leq 1 \Leftrightarrow \lambda_d \leq 1 \Rightarrow e^{\mathrm{Lvol}}_{T',T} \leq 1.$

c) $\inf_{\mu\in\mathbb{R}^d\setminus\{0\}} e^{\mathrm{HL}}_{T',T}(\mu) \geq 1 \Leftrightarrow \inf_{\mu\in\mathbb{R}^d\setminus\{0\}} e^{\mathrm{B}}_{T',T}(\mu) = 1 \Leftrightarrow \inf_{\eta\in\mathbb{R}^d\setminus\{0\}} e^{\mathrm{P}}_{T',T}(\eta) \geq 1 \Leftrightarrow \lambda_1 \geq 1 \Rightarrow e^{\mathrm{Lvol}}_{T',T} \geq 1.$

d) $\inf_{\mu\in\mathbb{R}^d\setminus\{0\}} e^{\mathrm{HL}}_{T',T}(\mu) > 1 \Leftrightarrow \inf_{\eta\in\mathbb{R}^d\setminus\{0\}} e^{\mathrm{P}}_{T',T}(\eta) > 1 \Leftrightarrow \lambda_1 > 1 \Rightarrow e^{\mathrm{Lvol}}_{T',T} > 1.$

Beweis: Dies folgt mit Korollar 8.11.1, Korollar 8.11.4 und Lemma 4.5.4. □

8.11.7 Beispiel: Im Fall $d = 1$ lässt sich die Hodges-Lehmann-, Pitman-, und in manchen Fällen auch die Bahadur-Effizienz durch die asymptotische Volumen-Effizienz charakterisieren. Es sei $\sigma := \Sigma_1 + \Sigma_2$. Dann gilt konkret:

a) $e^{\mathrm{HL}}_{T',T} < 1 \Leftrightarrow e^{\mathrm{P}}_{T',T} < 1 \Leftrightarrow e^{\mathrm{Lvol}}_{T',T} < 1 \Leftrightarrow \sigma > 0.$

b) $e^{\mathrm{HL}}_{T',T} \leq 1 \Leftrightarrow \sup_{\mu\in\mathbb{R}_*} e^{\mathrm{B}}_{T',T} = 1 \Leftrightarrow e^{\mathrm{P}}_{T',T} \leq 1 \Leftrightarrow e^{\mathrm{Lvol}}_{T',T} \leq 1 \Leftrightarrow \sigma \geq 0.$

b) $e^{\mathrm{HL}}_{T',T} \geq 1 \Leftrightarrow \inf_{\mu\in\mathbb{R}_*} e^{\mathrm{B}}_{T',T} = 1 \Leftrightarrow e^{\mathrm{P}}_{T',T} \geq 1 \Leftrightarrow e^{\mathrm{Lvol}}_{T',T} \geq 1 \Leftrightarrow \sigma \leq 0.$

d) $e^{\mathrm{HL}}_{T',T} > 1 \Leftrightarrow e^{\mathrm{P}}_{T',T} > 1 \Leftrightarrow e^{\mathrm{Lvol}}_{T',T} > 1 \Leftrightarrow \sigma < 0.$

8.11.8 Bemerkung: Um hier auf alternativem Weg Ergebnisse für die Pitman-Effizienz, die lokale Bahadur- und Hodges Lehmann-Effizienz oder die Volumen-Effizienz der Tests mit Hilfe der Resultate aus den Abschnitten 4 und 6 zu erhalten, kann folgende Parametrisierung benutzt werden: Wir wählen als gemeinsame Parametermenge der Verteilungen von ζ_1 und ζ_1'

$$\begin{aligned}\Theta =& \big\{ \big(\mu, \mathrm{vech}(\Gamma), \left(\begin{smallmatrix} b \\ \mathrm{vech}(\Sigma)\end{smallmatrix}\right)\big) \in \mathbb{R}^d \times \mathbb{R}^{d(d+1)/2} \times \mathbb{R}^{d+d(d+1)/2}; \\ & \Sigma, \Gamma \in \mathcal{S}_d, \exists\ \Pi \in \mathbb{R}^{d\times d} : \Gamma = 2\Sigma - (\Pi + \Pi^t), \left(\begin{smallmatrix} \Sigma & \Pi \\ \Pi^t & \Sigma \end{smallmatrix}\right) \in \mathcal{P}_{2d} \big\}.\end{aligned}$$

Dann ist $\Theta^1 = \mathbb{R}^d$, $\vartheta_0^1 = 0 \in \mathbb{R}^d$, und

$$\Theta_0 = \big\{ \big(\mu, \mathrm{vech}(\Gamma), \left(\begin{smallmatrix} b \\ \mathrm{vech}(\Sigma)\end{smallmatrix}\right)\big) \in \Theta; \mu = 0 \big\}.$$

Die injektive Partametrisierung der Verteilung von ζ_1 bzw. ζ_1' ist dann gegeben durch

$$\tilde{\Theta} = \big\{ \big(\mu, \mathrm{vech}(\Gamma)\big) \in \mathbb{R}^d \times \mathbb{R}^{d(d+1)/2}; \Gamma \in \mathcal{P}_d \big\}$$

bzw.

$$\tilde{\Theta}' = \big\{ \big(\mu, \left(\begin{smallmatrix} b \\ \mathrm{vech}(\Sigma)\end{smallmatrix}\right)\big) \in \mathbb{R}^d \times \mathbb{R}^{d+d(d+1)/2}; \Sigma \in \mathcal{P}_d \big\},$$

also

$$\tilde{\Theta}_0 = \big\{ \big(\mu, \mathrm{vech}(\Gamma)\big) \in \mathbb{R}^d \times \mathbb{R}^{d(d+1)/2}; \mu = 0, \Gamma \in \mathcal{P}_d \big\}$$

bzw.

$$\tilde{\Theta}_0' = \big\{ \big(\mu, \left(\begin{smallmatrix} b \\ \mathrm{vech}(\Sigma)\end{smallmatrix}\right)\big) \in \mathbb{R}^d \times \mathbb{R}^{d+d(d+1)/2}; \mu = 0, \Sigma \in \mathcal{P}_d \big\}.$$

8.12 Tests zur relativen Wirksamkeit

Um Tests zum Vergleich der Verfahren der verbundenen und unabhängigen Stichprobenerhebung anzugeben, kann hier auf die in Abschnitt 5 entwickelten Tests zurückgegriffen werden. Zunächst wollen wir konsistente Tests für die Testprobleme

$$\begin{aligned}
&\mathrm{H}_1: \sup_{\mu\in\mathbb{R}^d\setminus\{0\}} e^{\mathrm{HL}}_{T',T}(\mu) \le 1,\ \mathrm{K}_1: \sup_{\mu\in\mathbb{R}^d\setminus\{0\}} e^{\mathrm{HL}}_{T',T}(\mu) > 1\\
\Leftrightarrow\ &\mathrm{H}_1: \sup_{\mu\in\mathbb{R}^d\setminus\{0\}} e^{\mathrm{B}}_{T',T}(\mu) = 1,\ \mathrm{K}_1: \sup_{\mu\in\mathbb{R}^d\setminus\{0\}} e^{\mathrm{B}}_{T',T}(\mu) > 1\\
\Leftrightarrow\ &\mathrm{H}_1: \sup_{\eta\in\mathbb{R}^d\setminus\{0\}} e^{\mathrm{P}}_{T',T}(\eta) \le 1,\ \mathrm{K}_1: \sup_{\eta\in\mathbb{R}^d\setminus\{0\}} e^{\mathrm{P}}_{T',T}(\eta) > 1,
\end{aligned}$$

sowie

$$\begin{aligned}
&\mathrm{H}_2: \inf_{\mu\in\mathbb{R}^d\setminus\{0\}} e^{\mathrm{HL}}_{T',T}(\mu) \ge 1,\ \mathrm{K}_2: \inf_{\mu\in\mathbb{R}^d\setminus\{0\}} e^{\mathrm{HL}}_{T',T}(\mu) < 1\\
\Leftrightarrow\ &\mathrm{H}_2: \inf_{\mu\in\mathbb{R}^d\setminus\{0\}} e^{\mathrm{B}}_{T',T}(\mu) = 1,\ \mathrm{K}_2: \inf_{\mu\in\mathbb{R}^d\setminus\{0\}} e^{\mathrm{B}}_{T',T}(\mu) < 1\\
\Leftrightarrow\ &\mathrm{H}_2: \inf_{\eta\in\mathbb{R}^d\setminus\{0\}} e^{\mathrm{P}}_{T',T}(\eta) \ge 1,\ \mathrm{K}_2: \inf_{\eta\in\mathbb{R}^d\setminus\{0\}} e^{\mathrm{P}}_{T',T}(\eta) < 1,
\end{aligned}$$

und

$$\begin{aligned}
&\mathrm{H}_3:\ \forall\,\mu\in\mathbb{R}^d\setminus\{0\}: e^{\mathrm{HL}}_{T',T}(\mu) = 1,\ \mathrm{K}_3:\ \exists\,\mu\in\mathbb{R}^d\setminus\{0\}: e^{\mathrm{HL}}_{T',T}(\mu) \ne 1\\
\Leftrightarrow\ &\mathrm{H}_3:\ \forall\,\mu\in\mathbb{R}^d\setminus\{0\}: e^{\mathrm{B}}_{T',T}(\mu) = 1,\ \mathrm{K}_3:\ \exists\,\mu\in\mathbb{R}^d\setminus\{0\}: e^{\mathrm{B}}_{T',T}(\mu) \ne 1\\
\Leftrightarrow\ &\mathrm{H}_3:\ \forall\,\eta\in\mathbb{R}^d\setminus\{0\}: e^{\mathrm{P}}_{T',T}(\eta) = 1,\ \mathrm{K}_3:\ \exists\,\eta\in\mathbb{R}^d\setminus\{0\}: e^{\mathrm{P}}_{T',T}(\eta) \ne 1
\end{aligned}$$

aufstellen. Dafür werden wir die in Unterabschnitt 5.2 entwickelte Theorie zur Anwendung bringen.

Wir partitionieren die empirische Kovarianzmatrix basierend auf n Beobachtungen der verbundenen Stichprobe $\tilde{S}_n$ mit Hilfe der zufälligen $d \times d$-Matrizen $\tilde{S}_n^{1,1}, \tilde{S}_n^{1,2}, \tilde{S}_n^{2,1}$ und $\tilde{S}_n^{2,2}$ gemäß

$$\tilde{S}_n = \begin{pmatrix} \tilde{S}_n^{1,1} & \tilde{S}_n^{1,2} \\ \tilde{S}_n^{2,1} & \tilde{S}_n^{2,2} \end{pmatrix}.$$

Wir erinnern daran, dass die zugrunde liegende unbekannte Kovarianzmatrix durch

$$\tilde{S} := \begin{pmatrix} \Sigma & \Sigma\Sigma_1 \\ \Sigma\Sigma_2 & \Sigma \end{pmatrix} \in \mathcal{P}_{2d}$$

gegeben ist. Entsprechend der in Unterabschnitt 5.2 eingeführten Bezeichnungen ist hier

$$\tilde{E}(\vartheta_0) = (2\Sigma)^{-1} - \Gamma^{-1},$$

und das zu dieser Matrix entsprechende empirische Gegenstück ist gegeben durch

$$(\tilde{S}_n^{1,1} + \tilde{S}_n^{2,2})^{-1} - \big(\tilde{S}_n^{1,1} + \tilde{S}_n^{2,2} - (\tilde{S}_n^{2,1} + \tilde{S}_n^{1,2})\big)^{-1}.$$

Um Tests für die genannten Testprobleme zu formulieren, gehen wir aber unter Hinweis auf Bemerkung 5.2.10 stattdessen zur Matrix

$$\tilde{\tilde{E}}(\vartheta_0) = 2\Sigma - \Gamma = \Sigma\Sigma_1 + \Sigma\Sigma_2 = \Sigma\Sigma_2 + (\Sigma\Sigma_2)^t,$$

bzw. zu ihrem empirischen Gegenstück

$$(\tilde{S}_n^{1,1} + \tilde{S}_n^{2,2}) - \big(\tilde{S}_n^{1,1} + \tilde{S}_n^{2,2} - (\tilde{S}_n^{2,1} + \tilde{S}_n^{1,2})\big) = \tilde{S}_n^{2,1} + (\tilde{S}_n^{2,1})^t$$

über, und lassen unsere Tests auf den nach wachsender Größe geordneten Eigenwerten $\tilde{\lambda}_{1,n} \leq \cdots \leq \tilde{\lambda}_{d,n}$ dieser Matrix basieren.

Wir definieren $\hat{p}_n := \operatorname{vec}(\tilde{S}_n^{2,1})$ und $p_0 := \operatorname{vec}(\Sigma\Sigma_2)$. Dann hängt $\tilde{\tilde{E}}(\vartheta_0)$ ausschließlich von p_0 ab. Es ist also $\tilde{\tilde{E}}(\vartheta_0)$ eine Matrix

$$\tilde{H}(p_0) = \Sigma\Sigma_2 + (\Sigma\Sigma_2)^t,$$

mit der in Unterabschnitt 5.2 eingeführten Abbildung $\tilde{H}$, und

$$\tilde{H}(\hat{p}_n) = \tilde{S}_n^{2,1} + (\tilde{S}_n^{2,1})^t.$$

Zur Behandlung des Testproblems H_1 gegen K_1 verwenden wir die Testgröße

$$F_n^1 = -\frac{\sqrt{n}\tilde{\lambda}_{d,n}}{\sqrt{M_{d,n}}},$$

und zur Behandlung des Testproblems H_2 gegen K_2 die Testgröße

$$F_n^2 = \frac{\sqrt{n}\tilde{\lambda}_{1,n}}{\sqrt{M_{1,n}}},$$

mit

$$M_{i,n} = w_i(\hat{p}_n)^t \mathcal{J}_{\tilde{H}}(\hat{p}_n) V(\hat{p}_n) \mathcal{J}_{\tilde{H}}(\hat{p}_n)^t w_i(\hat{p}_n), \; i = 1, \ldots, d,$$

wobei für $i = 1, \ldots, d$ die Größen $w_i(\hat{p}_n), \mathcal{J}_{\tilde{H}}(\hat{p}_n)$ und $V(\hat{p}_n)$ die empirischen Gegenstücke zu $w_i(p_0), \mathcal{J}_{\tilde{H}}(p_0)$ und $V(p_0)$ sind, die wie folgt definiert werden:

Für $i = 1, \ldots, d$ ist der Vektor $w_i(p_0) \in \mathbb{R}^{d(d+1)/2}$ durch die Komponenten

$$\big(w_i(p_0)\big)_j := \big(\beta_i(p_0)^t \otimes \beta_i(p_0)^t\big) D e_j, \; j = 1, \ldots, d(d+1)/2,$$

festgelegt. Dabei ist $\beta_i(p_0) \in \mathbb{R}^d$ der normierte Eigenvektor zum Eigenwert $\tilde{\lambda}_i(p_0)$ der Matrix $\tilde{H}(p_0)$, $i = 1, \ldots, d$, und D die in Unterabschnitt 5.2 eingeführte Duplikationsmatrix. Weiter ist

$$V(p_0) := B \operatorname{Iss}(\tilde{S}) B^t,$$

wobei die Matrix $\operatorname{Iss}(\tilde{S}) \in \mathcal{S}_{2d(2d+1)/2}$ die Isserlis-Matrix von $\tilde{S}$ ist, die durch

$$\sqrt{n}\big(\operatorname{vech}(\tilde{S}_n) - \operatorname{vech}(\tilde{S})\big) \xrightarrow{\text{v}} N_{2d(2d+1)/2}\big(0, \operatorname{Iss}(\tilde{S})\big) \text{ für } n \to \infty$$

festgelegt ist. Diese Isserlis-Matrix lässt sich wie folgt genauer beschreiben. Es sei durch M eine zufällige $2d \times 2d$-Matrix mit

$$\sqrt{n}\big(\operatorname{vech}(\tilde{S}_n) - \operatorname{vech}(\tilde{S})\big) \xrightarrow{\text{v}} \operatorname{vech}(M) \text{ für } n \to \infty$$

gegeben, wobei

$$\operatorname{vech}(M) \sim N_{2d(2d+1)/2}\big(0, \operatorname{Iss}(\tilde{S})\big)$$

gelte, mit einer Matrix $\operatorname{Iss}(\tilde{S}) \in \mathcal{S}_{2d(2d+1)/2}$. Siehe Theorem 3.4.3 in [1]. Dann gilt

$$\operatorname{Iss}(\tilde{S}) = \operatorname{Cov}\big(\operatorname{vech}(M)\big)$$

und mit $M = (m_{i,j})_{1 \le i,j \le 2d}$ und $\tilde{S} = (s_{i,j})_{1 \le i,j \le 2d}$

$$\operatorname{Cov}(m_{i,j}, m_{k,l}) = s_{i,k}s_{j,l} + s_{i,l}s_{j,k}, \ 1 \le i \le j \le 2d.$$

Siehe hierzu Theorem 3.4.4 in [1]. Weiter ist $B \in \mathbb{R}^{d^2 \times 2d(2d+1)/2}$ die Matrix, die durch die Eigenschaft

$$B\operatorname{vech}(K) = \operatorname{vec}(K^{2,1}),$$

mit

$$K = \begin{pmatrix} K^{1,1} & K^{1,2} \\ K^{2,1} & k^{2,2} \end{pmatrix} \in \mathcal{S}_{2d}$$

und $K^{1,1}$, $K^{1,2}$, $K^{2,1}$ und $K^{2,2}$ $d \times d$-Matrizen, gegeben ist. Darüber hinaus ist $\mathcal{J}_{\tilde{H}}(p_0) \in \mathbb{R}^{d(d+1)/2 \times d^2}$ die Jacobi-Matrix der stetig differenzierbaren Abbildung $\operatorname{vech}(\tilde{H}) : \mathbb{R}^{d^2} \to \mathbb{R}^{d(d+1)/2}$.

Es hängt $B \operatorname{Iss}(\tilde{S}) B^t$ nur von p_0 ab und zwar stetig. Weiter ist

$$\det \operatorname{Iss}(\tilde{S}) = 2^{2d} (\det \tilde{S})^{2d+1},$$

siehe [34]. Daher ist wegen der positiven Definitheit von $\tilde{S}$ auch die Isserlis-Matrix $\operatorname{Iss}(\tilde{S})$ positv definit. Da die Matrix B vollen Zeilenrang besitzt, ist auch die Matrix $V(p_0) =$

$B\,\mathrm{Iss}(\tilde{S})B^t$ positiv definit. Diese Matrix hängt außerdem stetig von p_0 ab. Zudem besitzt die Jacobi-Matrix $\mathcal{J}_{\tilde{H}}(p_0)$ vollen Zeilenrang. Wie oben erläutert gilt

$$\sqrt{n}\big(\mathrm{vech}(\tilde{S}_n) - \mathrm{vech}(\tilde{S})\big) \xrightarrow{\mathrm{v}} N_{2d(2d+1)/2}\big(0, \mathrm{Iss}(\tilde{S})\big) \text{ für } n \to \infty.$$

Daher gilt auch

$$\sqrt{n}\big(B\mathrm{vech}(\tilde{S}_n) - B\mathrm{vech}(\tilde{S})\big) \xrightarrow{\mathrm{v}} N_{2d(2d+1)/2}\big(0, B\mathrm{Iss}(\tilde{S})B^t\big) \text{ für } n \to \infty,$$

äquivalent

$$\sqrt{n}\big(\mathrm{vec}(\tilde{S}_n^{2,1}) - \mathrm{vec}(\Sigma\Sigma_2)\big) \xrightarrow{\mathrm{v}} N_{2d(2d+1)/2}\big(0, V(p_0)\big) \text{ für } n \to \infty,$$

äquivalent

$$\sqrt{n}(\hat{p}_n - p_0) \xrightarrow{\mathrm{v}} N_{2d(2d+1)/2}\big(0, V(p_0)\big) \text{ für } n \to \infty.$$

Somit sind (E1) und (E2) aus Unterabschnitt 5.2 erfüllt. Wie in Unterabschnitt 5.2 argumentiert wurde, gilt im Fall der Einfachheit des kleinsten und größten Eigenwerts der Matrix $\tilde{H}(p_0)$

$$\lim_{n\to\infty} P_{\vartheta_0}(M_{1,n} \le 0) = \lim_{n\to\infty} P_{\vartheta_0}(M_{d_1,n} \le 0) = 0,$$

wenn ϑ_0 der zu p_0 korrespondierende Parameter ist.

Für ein gegebenes Testniveau $\alpha \in (0,1)$ schlagen wir für das Testproblem H_i gegen K_i, $i = 1, 2$, den Test

$$\text{„Verwirf } \mathrm{H}_i, \text{ falls } F_n^i \ge u_{1-\alpha}\text{“}, \; i = 1, 2,$$

vor, und können folgendes festhalten.

8.12.1 Korollar: *Es sei der kleinste und größte Eigenwert der Matrix $\tilde{H}(p_0)$ einfach. Dann handelt es sich bei dem auf der Testgröße F_n^i basierenden vorgeschlagenen Test für das Testproblem H_i gegen K_i, $i = 1, 2$, um einen Test, der für jedes $\alpha \in (0,1)$ das Testniveau α asymptotisch einhält, und der zudem konsistent ist.*

Beweis: Wie oben begründet sind (E1) und (E2) aus Unterabschnitt 5.2 erfüllt. Da nach Annahme auch die Bedingung (E3) aus aus Unterabschnitt 5.2 erfüllt ist, folgt mit Korollar 5.2.6 die Behauptung. □

Zur Behandlung des Testproblems H_3 gegen K_3 betrachten wir die Testgröße

$$F_n^3 = n\tilde{\lambda}_{(d),n}^2.$$

Dabei seien $\tilde{\lambda}_{(1),n}, \ldots, \tilde{\lambda}_{(d),n}$ die von ihrem Betrag her nach wachsender Größe geordneten Eigenwerte von $\tilde{H}(\hat{p}_n)$, d.h. es gilt $|\tilde{\lambda}_{(1),n}| \le \cdots \le |\tilde{\lambda}_{(d),n}|$.

Für ein gegebenes Testniveau $\alpha \in (0,1)$ schlagen wir für das Testproblem H_3 gegen K_3 den Test

$$„\text{Verwirf } \mathrm{H}_3, \text{ falls } F_n^3 \geq c_{\hat{p}_n;1-\alpha}“$$

vor. Hierbei ist $c_{p_0;1-\alpha}$ das $(1-\alpha)$-Quantil der Verteilung mit der Lebesgue-Dichte

$$k_{p_0}(l_d) = \int \dots \int \frac{\pi^{\frac{d^2}{2}}}{\Gamma_d(\frac{d}{2})} \prod_{1 \leq j < i \leq d} (l_i - l_j) \int_{O_d} f_{p_0}\big(H \operatorname{diag}(l_d, \dots, l_1) H'\big) \mathrm{d}H \mathrm{d}l_1 \dots \mathrm{d}l_{d-1},$$

für $l_d > 0$, und 0 sonst.

Hierbei wird das innere Integral über die Menge $O_d = \{A \in \mathbb{R}^{d \times d}; A^t A = \mathrm{I}_d\}$ bezüglich der Gleichverteilung auf O_d durchgeführt, Γ_m sei die verallgemeinerte Gammafunktion $\Gamma_m(t) := \pi^{m(m-1)/4} \prod_{i=1}^m \Gamma(t-(i-1)/2)$, $m \in \mathbb{N}$, $t > (m-1)/2$, und f_{p_0} sei die Lebesgue-Dichte einer Matrix $A^2 \in \mathbb{R}^{d \times d}$, wobei $A \in \mathbb{R}^{d \times d}$ eine symmetrische Matrix mit $\operatorname{vech}(A) \sim \mathrm{N}_{d(d+1)/2}(0, \mathcal{J}_{\tilde{H}}(p_0) V(p_0) \mathcal{J}_{\tilde{H}}(p_0)^t)$ sei; siehe Unterabschnitt 5.2. Dann halten wir folgendes Resultat fest.

8.12.2 Korollar: *Bei dem auf der Testgröße F_n^3 basierenden vorgeschlagenen Test für das Testproblem H_3 gegen K_3 handelt es sich um einen Test, der für jedes $\alpha \in (0,1)$ das Testniveau α asymptotisch exakt einhält, und der zudem konsistent ist.*

Beweis: Die Aussage folgt aus Korollar 5.2.8. □

Nun wollen wir konsitente Tests unter sinnvollen Alternativen gemäß den Ausführungen aus Unterabschnitt 5.3 angeben. Betrachten wir also das Testproblem

$$\begin{aligned} &\mathrm{H}_1 : \sup_{\mu \in \mathbb{R}^d \setminus \{0\}} e^{\mathrm{HL}}_{T',T}(\mu) \leq 1, \ \tilde{\mathrm{K}}_1 : \inf_{\mu \in \mathbb{R}^d \setminus \{0\}} e^{\mathrm{HL}}_{T',T}(\mu) > 1 \\ \Leftrightarrow\ &\mathrm{H}_1 : \sup_{\eta \in \mathbb{R}^d \setminus \{0\}} e^{\mathrm{P}}_{T',T}(\eta) \leq 1, \ \tilde{\mathrm{K}}_1 : \inf_{\eta \in \mathbb{R}^d \setminus \{0\}} e^{\mathrm{P}}_{T',T}(\eta) > 1. \end{aligned}$$

Nach Korollar 8.11.1 lässt sich die Hypothese auch in der Form

$$\mathrm{H}_1 : \sup_{\mu \in \mathbb{R}^d \setminus \{0\}} e^{\mathrm{B}}_{T',T}(\mu) = 1$$

äquivalent formulieren. Zudem betrachten wir das Testproblem

$$\begin{aligned} &\mathrm{H}_2 : \inf_{\mu \in \mathbb{R}^d \setminus \{0\}} e^{\mathrm{HL}}_{T',T}(\mu) \geq 1, \ \tilde{\mathrm{K}}_2 : \sup_{\mu \in \mathbb{R}^d \setminus \{0\}} e^{\mathrm{HL}}_{T',T}(\mu) < 1 \\ \Leftrightarrow\ &\mathrm{H}_2 : \inf_{\eta \in \mathbb{R}^d \setminus \{0\}} e^{\mathrm{P}}_{T',T}(\eta) \geq 1, \ \tilde{\mathrm{K}}_2 : \sup_{\eta \in \mathbb{R}^d \setminus \{0\}} e^{\mathrm{P}}_{T',T}(\eta) < 1. \end{aligned}$$

Auch hier ist die Hypothese äquivalent zu

$$\mathrm{H}_2: \inf_{\mu \in \mathbb{R}^d \setminus \{0\}} e^{\mathrm{B}}_{T',T}(\mu) = 1.$$

Schließlich betrachten wir noch das Testproblem

$$\mathrm{H}_3: \ \forall\, \mu \in \mathbb{R}^d \setminus \{0\} : e^{\mathrm{HL}}_{T',T}(\mu) = 1,\ \tilde{\mathrm{K}}_3: \sup_{\mu \in \mathbb{R}^d \setminus \{0\}} e^{\mathrm{HL}}_{T',T}(\mu) < 1 \text{ oder } \inf_{\mu \in \mathbb{R}^d \setminus \{0\}} e^{\mathrm{HL}}_{T',T}(\mu) > 1$$

$$\Leftrightarrow\ \mathrm{H}_3: \ \forall\, \eta \in \mathbb{R}^d \setminus \{0\} :\ e^{\mathrm{P}}_{T',T}(\mu) = 1,\ \tilde{\mathrm{K}}_3: \sup_{\eta \in \mathbb{R}^d \setminus \{0\}} e^{\mathrm{P}}_{T',T}(\eta) < 1 \text{ oder } \inf_{\eta \in \mathbb{R}^d \setminus \{0\}} e^{\mathrm{P}}_{T',T}(\eta) > 1.$$

Hier ist die Hypothese ebenfalls äquivalent zu

$$\mathrm{H}_3: \forall\, \mu \in \mathbb{R}^d \setminus \{0\} :\ e^{\mathrm{B}}_{T',T}(\mu) = 1.$$

Diese Testprobleme sollen unter Hinweis auf Bemerkung 5.3.4 im Wesentlichen erneut mit Hilfe der Matrix $\tilde{H}(\hat{p}_n)$ behandelt werden. Zur Behandlung des Testproblems H_i gegen $\tilde{\mathrm{K}}_i$, $i = 1, 2, 3$, betrachten wir die Testgrößen

$$\tilde{F}^1_n = \frac{\sqrt{n}\mathrm{tr}\big(\tilde{H}(\hat{p}_n)\big)}{\sqrt{\tilde{M}_n}},\ \tilde{F}^2_n = -\tilde{F}^1_n,\ \tilde{F}^3_n = |\tilde{F}^1_n|,$$

mit

$$\tilde{M}_n = \tilde{M}_n(\hat{p}_n) = v^t \mathcal{J}_{\tilde{H}}(\hat{p}_n) V(\hat{p}_n) \mathcal{J}_{\tilde{H}}(\hat{p}_n)^t v,$$

wobei der Verktor $v \in \mathbb{R}^{d(d+1)/2}$ durch

$$v = (1, \underbrace{0, \dots, 0}_{(d-1)\text{-mal}}, 1, \underbrace{0, \dots, 0}_{(d-2)\text{-mal}}, 1, \dots, 1, 0, 1)$$

gegeben ist. Wie oben gezeigt sind hier die Bedingungen (E1) und (E2) aus Unterabschnitt 5.2 erfüllt. Daher gilt

$$\lim_{n \to \infty} P_{\vartheta_0}(\tilde{M}_n \leq 0) = 0,$$

was in Unterabschnitt 5.3 begründet wurde.

Für ein gegebenes Testniveau $\alpha \in (0, 1)$ lautet dann der zugehörige Test für das Testproblem H_i gegen $\tilde{\mathrm{K}}_i$, $i = 1, 2$,

$$\text{„Verwirf } \mathrm{H}_i, \text{ falls } \tilde{F}^i_n \geq u_{1-\alpha}\text{“},\ i = 1, 2.$$

Für das Testproblem H_3 gegen K_3 kann der Test

$$\text{„Verwirf } \mathrm{H}_3, \text{ falls } \tilde{F}^3_n \geq u_{1-\alpha/2}\text{“}$$

formuliert werden. Dann können wir folgendes festhalten.

8.12.3 Korollar: *Bei dem auf der Testgröße $\tilde{F}_n^i$ basierenden vorgeschlagenen Test für das Testproblem* H_i *gegen* $\tilde{\mathrm{K}}_i$, $i = 1, 2$, *handelt es sich um einen Test, der für jedes* $\alpha \in (0,1)$ *das Testniveau* α *asymptotisch einhält, und der zudem konsistent ist. Bei dem auf der Testgröße* $\tilde{F}_n^3$ *basierenden vorgeschlagenen Test für das Testproblem* H_3 *gegen* $\tilde{\mathrm{K}}_3$ *handelt es sich darüber hinaus um einen Test, der für jedes* $\alpha \in (0,1)$ *das Testniveau* α *asymptotisch exakt einhält und der zudem konsistent ist.*

Beweis: Die Aussage folgt aus Korollar 5.3.3. □

Im Fall $d = 1$ bietet sich hier ein Vorgehen gemäß Bemerkung 5.2.11 bzw. Bemerkung 5.3.5 an. Hierzu definieren wir

$$D_n^1 := \sqrt{n}\left(\frac{\tilde{S}_n^{1,1} + \tilde{S}_n^{2,2} - (\tilde{S}_n^{2,1} + \tilde{S}_n^{1,2})}{\tilde{S}_n^{1,1} + \tilde{S}_n^{2,2}} - 1\right), \ D_n^2 := -D_n^1, \ D_n^3 := |D_n^1|.$$

8.12.4 Lemma: *Es gelte* $\Gamma = 2\Sigma$. *Dann ist die Testgröße* D_n^i, $i = 1, 2, 3$, *verteilungsfrei.*

Beweis: Es ist zu beachten, dass für $d = 1$ im Fall $\Gamma = 2\Sigma$ sofort $\Sigma_1 = \Sigma_2 = 0$ folgt, und deshalb wegen

$$\begin{pmatrix} X_1 \\ \vdots \\ X_n \\ Y_1 \\ \vdots \\ Y_n \end{pmatrix} \stackrel{\mathrm{v}}{=} \begin{pmatrix} \sqrt{\Sigma}U_1 + a \\ \vdots \\ \sqrt{\Sigma}U_n + a \\ \sqrt{\Sigma}V_1 + b \\ \vdots \\ \sqrt{\Sigma}V_n + b \end{pmatrix}$$

mit $\binom{U_1}{V_1}, \ldots, \binom{U_n}{V_n}$ unabhängig und je $\mathrm{N}_2(0, \mathrm{I}_2)$-verteilt, auch

$$\begin{aligned} D_n^1 &= \sqrt{n}\left(\frac{\frac{1}{n}\sum_{i=1}^n \left((X_i - \overline{X}_n)^2 + (Y_i - \overline{Y}_n)^2 - 2(X_i - \overline{X}_n)(Y_i - \overline{Y}_n)\right)}{\frac{1}{n}\sum_{i=1}^n \left((X_i - \overline{X}_n)^2 + (Y_i - \overline{Y}_n)^2\right)} - 1\right) \\ &\stackrel{\mathrm{v}}{=} \sqrt{n}\left(\frac{\frac{1}{n}\sum_{i=1}^n \left((U_i - \overline{U}_n)^2 + (V_i - \overline{V}_n)^2 - 2(U_i - \overline{U}_n)(V_i - \overline{V}_n)\right)}{\frac{1}{n}\sum_{i=1}^n \left((U_i - \overline{U}_n)^2 + (V_i - \overline{V}_n)^2\right)} - 1\right) \end{aligned}$$

gilt, d.h. D_n^i, $i = 1, 2, 3$, ist im Fall $\Gamma = 2\Sigma$ verteilungsfrei. □

Für ein gegebenes Testniveau $\alpha \in (0,1)$ schlagen wir für das Testproblem H_i gegen K_i, $i = 1, 2, 3$, folgenden Test vor:

$$\text{„Verwirf } \mathrm{H}_i, \text{ falls } D_n^i \geq t_{n;1-\alpha}^i\text{“}, \ i = 1, 2, 3,$$

wobei $t_{n;1-\alpha}^i$ das $(1-\alpha)$-Quantil[10] der Verteilung von D_n^i, $i = 1, 2, 3$, im Fall $(X_1, Y_1) \sim \mathrm{N}_2(0, \mathrm{I}_2)$ sei. Dann kann folgendes festgehalten werden:

[10] Diese Quantile können mit Hilfe eines Monte-Carlo-Verfahrens beliebig genau approximiert werden.

8.12.5 Korollar: *Bei dem auf der Testgröße D_n^i basierenden vorgeschlagenen Test für das Testproblem H_i gegen K_i, $i = 1, 2$, handelt es sich um einen Test, der für jedes $\alpha \in (0, 1)$ das Testniveau α asymptotisch einhält, und der zudem konsistent ist. Bei dem auf der Testgröße D_n^3 basierenden vorgeschlagenen Test für das Tesproblem H_3 gegen H_3 handelt es sich darüber hinaus um einen Test, der für jedes $\alpha \in (0, 1)$ das Testniveau α exakt einhält und der zudem konsistent ist.*

Beweis: Dies folgt mit Lemma 8.12.4 und mit den in Unterabschnitt 5.2 angestellten Überlegungen. $\square$

8.12.6 Bemerkung: Wie in Bemerkung 5.3.5 erläutert können analog konsistente Tests für die Testprobleme

$$\mathrm{H}_4: \ e_{T',T}^{\mathrm{Lvol}} \leq 1, \ \mathrm{K}_4: \ e_{T',T}^{\mathrm{Lvol}} > 1$$

sowie

$$\mathrm{H}_5: \ e_{T',T}^{\mathrm{Lvol}} \geq 1, \ \mathrm{K}_5: \ e_{T',T}^{\mathrm{Lvol}} < 1$$

und

$$\mathrm{H}_6: \ e_{T',T}^{\mathrm{Lvol}} = 1, \ \mathrm{K}_6: \ e_{T',T}^{\mathrm{Lvol}} \neq 1$$

basierend auf der Determinante der Matrix

$$(\tilde{S}_n^{1,1} + \tilde{S}_n^{2,2})^{-1}\big(\tilde{S}_n^{1,1} + \tilde{S}_n^{2,2} - (\tilde{S}_n^{2,1} + \tilde{S}_n^{1,2})\big)$$

konstruiert werden.

9 Vergleiche bei Kontingenztafeln (Multinomialverteilungen)

9.1 Das zugrunde liegende Modell

In diesem Abschnitt kommen die Ergebnisse aus den Abschnitten 4, 5 und 6 zum Einsatz. Es seien $(X_1, Y_1), \ldots, (X_n, Y_n)$ unabhängige und identisch verteilte zweidimensionale Zufallsvektoren mit dem Wertebereich $\{1, \ldots, d\}^2$, wobei $d \in \mathbb{N}$ gegeben sei. Zudem sei

$$p_{i,j} := P\big((X_1, Y_1) = (i, j)\big),\ (i, j) \in \{1, \ldots, d\}^2,$$

mit $(p_{1,1}, \ldots, p_{d,d}) \in (0, 1)^{d^2}$, $\sum_{i,j=1}^{d} p_{i,j} = 1$ und

$$p := (p_{1,1}, \ldots, p_{d,1}, p_{1,2}, \ldots, p_{d,2}, \ldots, p_{1,d-1}, \ldots, p_{d,d-1}, p_{1,d}, \ldots, p_{d-1,d})^t$$

unbekannt.

Wir definieren

$$\Delta_s := \{(q_1, \ldots, q_{s-1})^t; (q_1, \ldots, q_{s-1}) \in (0, 1)^{s-1}, \sum_{i=1}^{s-1} q_i < 1\},\ s \in \mathbb{N},$$

und

$$\Delta_s^0 := \{(q_1, \ldots, q_{s-1})^t; (q_1, \ldots, q_{s-1}) \in [0, 1]^{s-1}, \sum_{i=1}^{s-1} q_i \leq 1\},\ s \in \mathbb{N}.$$

Für $(q_1, \ldots, q_{s-1})^t \in \Delta_s^0$ setzen wir zudem $q_s := 1 - \sum_{i=1}^{s-1} q_i$. Elemente

$$q = (q_{1,1}, \ldots, q_{s,1}, q_{1,2}, \ldots, q_{s,2}, \ldots, q_{1,s-1}, \ldots, q_{s,s-1}, q_{1,s}, \ldots, q_{s-1,s})^t \in \Delta_{s^2}^0$$

schreiben wir auch kürzer in der Form $q = (q_{1,1}, \ldots, q_{s-1,s})^t$.

Es seien $m, s \in \mathbb{N}$ und $q := (q_1, \ldots, q_{s-1})^t \in \Delta_s^0$. Für einen $\mathbb{R}^{s-1}$-wertigen Zufallsvektor $M := (M_1, \ldots, M_{s-1})^t$ schreiben wir im Folgenden $M \sim \mathfrak{M}(m; q_1, \ldots, q_s)$, falls für den $\mathbb{R}^s$-wertigen Zufallsvektor $(M_1, \ldots, M_{s-1}, m - \sum_{i=1}^{s-1} M_i)^t$ die Verteilungseigenschaft $(M_1, \ldots, M_{s-1}, m - \sum_{i=1}^{s-1} M_i)^t \sim \mathfrak{M}(m; q_1, \ldots, q_s)$ gilt. Zudem setzen wir dann $M_s := m - \sum_{i=1}^{s-1} M_i$. Zufallsvektoren

$$\begin{aligned} M &= (M_{1,1}, \ldots, M_{s,1}, M_{1,2}, \ldots, M_{s,2}, \ldots, M_{1,s-1}, \ldots, M_{s,s-1}, M_{1,s}, \ldots, M_{s-1,s})^t \\ &\sim \mathfrak{M}(m; q_{1,1}, \ldots, q_{s,s}) \end{aligned}$$

schreiben wir auch kürzer in der Form $M = (M_{1,1}, \ldots, M_{s-1,s})^t$.

Für $i = 1, \ldots, n$ kann (X_i, Y_i) mit dem $\mathbb{R}^{d^2-1}$-wertigen Zufallsvektor N_i mit

$$N_i \sim \mathfrak{M}(1; p_{1,1}, \ldots, p_{d,d}),\ i = 1, \ldots, n,$$

identifiziert werden, wobei

$$N_i := \Big(\mathrm{I}\big((X_i, Y_i) = (1,1)\big), \dots, \mathrm{I}\big((X_i, Y_i) = (d-1, d)\big) \Big)^t, \; i = 1, \dots, n.$$

Im Folgenden werden die Zufallsvektoren $N_1, \dots, N_n$ anstelle der Zufallsvektoren $(X_1, Y_1), \dots, (X_n, Y_n)$ betrachtet.

Mit

$$N_{i,j} := \sum_{k=1}^{n} \mathrm{I}\big((X_k, Y_k) = (i,j)\big), \; (i,j) \in \{1, \dots, d\}^2,$$

ist

$$N := (N_{1,1}, \dots, N_{d-1,d})^t \sim \mathfrak{M}(n; p_{1,1}, \dots, p_{d,d}).$$

Nach Konstruktion ist dann $N = \sum_{i=1}^{n} N_i$. Das starke Gesetz der großen Zahlen liefert

$$\lim_{n \to \infty} N/n = (p_{1,1}, \dots, p_{d-1,d})^t \; P_p\text{-f.s.}$$

Mit der Definition

$$p_{i\cdot} := \sum_{k=1}^{d} p_{i,k}, \; p_{\cdot i} := \sum_{k=1}^{d} p_{k,i}, \; i \in \{1, \dots, d\},$$

und mit

$$N_{i\cdot} := \sum_{k=1}^{n} \mathrm{I}(X_k = i), \; N_{\cdot i} := \sum_{k=1}^{n} \mathrm{I}(Y_k = i), \; i \in \{1, \dots, d\},$$

ist

$$N_{*\cdot} := (N_{1\cdot}, \dots, N_{d-1\cdot})^t \sim \mathfrak{M}(n; p_{1\cdot}, \dots, p_{d\cdot}), \; N_{\cdot *} := (N_{\cdot 1}, \dots, N_{\cdot d-1})^t \sim \mathfrak{M}(n; p_{\cdot 1}, \dots, p_{\cdot d}).$$

Im unabhängigen Stichprobenfall liegen mit $X_1', \dots, X_n'$ und $Y_1', \dots, Y_n'$ zwei voneinander unabhängige Stichproben vor, wobei

$$\mathcal{L}(X_1) = \mathcal{L}(X_1') \text{ und } \mathcal{L}(Y_1) = \mathcal{L}(Y_1') \tag{9.1.1}$$

ist. Für $i = 1, \dots, n$ kann (X_i', Y_i') wegen (9.1.1) mit einem $\mathbb{R}^{2d-2}$-wertigen Zufallsvektor N_i' mit

$$N_i' \sim \mathfrak{M}(1; p_{1\cdot}, \dots, p_{d\cdot}) \otimes \mathfrak{M}(1; p_{\cdot 1}, \dots, p_{\cdot d}), \; i = 1, \dots, n,$$

identifiziert werden, wobei

$$N_i' := \big(\mathrm{I}(X_i' = 1), \dots, \mathrm{I}(X_i' = d-1), \mathrm{I}(Y_i' = 1), \dots, \mathrm{I}(Y_i' = d-1) \big)^t, \; i = 1, \dots, n.$$

Im Folgenden werden die Zufallsvektoren $N'_1, \ldots, N'_n$ anstelle der Zufallsvektoren $(X'_1, Y'_1), \ldots, (X'_n, Y'_n)$ betrachtet.

Mit

$$N'_{i\cdot} := \sum_{k=1}^{n} \mathrm{I}(X'_k = i),\ N'_{\cdot i} := \sum_{k=1}^{n} \mathrm{I}(Y'_k = i),\ i \in \{1, \ldots, d\},$$

gilt wegen (9.1.1)

$$N'_{*\cdot} := (N'_{1\cdot}, \ldots, N'_{d-1\cdot})^t \sim \mathfrak{M}(n; p_{1\cdot}, \ldots, p_{d\cdot}),\ N'_{\cdot *} := (N'_{\cdot 1}, \ldots, N'_{\cdot d-1})^t \sim \mathfrak{M}(n; p_{\cdot 1}, \ldots, p_{\cdot d}),$$

wobei $N'_{*\cdot}$ und $N'_{\cdot *}$ unabhängig sind. Nach Konstruktion ist dann $\binom{N'_{*\cdot}}{N'_{\cdot *}} = \sum_{i=1}^n N'_i$. Das Gesetz der großen Zahlen liefert auch hier

$$\lim_{n\to\infty} N'_{*\cdot}/n = (p_{1\cdot}, \ldots, p_{d-1\cdot})^t,\ \lim_{n\to\infty} N'_{\cdot *}/n = (p_{\cdot 1}, \ldots, p_{\cdot d-1})^t\ P_p\text{-f.s.}.$$

Das zu Beginn dieser Arbeit vorgestellte Testproblem

$$\mathrm{H}:\ \mathcal{L}(X_1) = \mathcal{L}(Y_1),\ \mathrm{K}:\ \mathcal{L}(X_1) \neq \mathcal{L}(Y_1),$$

kann hier äquivalent gemäß

$$\mathrm{H}:\ (p_{1\cdot}, \ldots, p_{d-1\cdot}) = (p_{\cdot 1}, \ldots, p_{\cdot d-1}),\ \mathrm{K}:\ (p_{1\cdot}, \ldots, p_{d-1\cdot}) \neq (p_{\cdot 1}, \ldots, p_{\cdot d-1})$$

formuliert werden.

9.2 Allgemeine Betrachtungen

Wir betrachten zunächst ein allgemeines Einstichprobenproblem bei Multinomialverteilung. Dazu seien $m, s \in \mathbb{N}$ und für $q := (q_1, \ldots, q_{s-1}) \in \Delta_s^0$ sei $M := (M_1, \ldots, M_{s-1}) \sim \mathfrak{M}(m; q_1, \ldots, q_s)$.

Für ein $l = (l_1, \ldots, l_{s-1})^t \in \{(i_1, \ldots, i_{s-1})^t; (i_1, \ldots, i_{s-1}) \in \{0, \ldots, m\}^{s-1}, \sum_{j=1}^{s-1} i_j \leq m\}$ setzen wir $l_s := 1 - \sum_{i=1}^{s-1} l_i$.

Es sei $\emptyset \neq \Pi_0 \subsetneq \Pi = \Delta_s$ und wir betrachten das allgemeine Testproblem

$$\mathrm{H}_0:\ q \in \Pi_0,\ \mathrm{K}_0:\ q \in \Pi_1 := \Pi \setminus \Pi_0.$$

Eine Abbildung $\hat{q} : \Delta_s^0 \to \overline{\Pi_0}$ heiße Maximum-Schätzfunktion für $q \in \Pi$ unter H_0, falls

$$P_{\hat{q}(l/m)}(M = l) = \sup_{q \in \Pi_0} P_q(M = l),$$
$$l \in \Big\{(i_1, \ldots, i_{s-1})^t; (i_1, \ldots, i_{s-1}) \in \{0, \ldots, m\}^{s-1}, \sum_{j=1}^{s-1} i_j \leq m\Big\},$$

erfüllt ist. Wir schreiben $\hat{q} = (\hat{q}_1, \ldots, \hat{q}_{s-1})^t$. Ist $\hat{q}$ Maximum-Schätzfunktion für $q \in \Pi$ unter H_0 und gilt $\hat{q} : (\Delta_s^0, \mathfrak{B}^{s-1}_{\Delta_s^0}) \to (\overline{\Pi_0}, \mathfrak{B}^{s-1}_{\overline{\Pi_0}})$, so ist $\hat{q}(M/m)$ Maximum-Likelihood-Schätzer[11] für $q \in \Pi$ unter H_0 bei zugrunde liegendem M.

Es besitzt die Kullback-Leibler-Information aus der Verteilung von M für $m = 1$ die Gestalt

$$K(q, \tilde{q}) = \sum_{i=1}^{s} q_i \log \frac{q_i}{\tilde{q}_i} = \sum_{i=1}^{s} q_i \log q_i - \sum_{i=1}^{s} q_i \log \tilde{q}_i, \;\; q, \tilde{q} \in \Delta_s^0,$$

wobei an dieser Stelle $0/0 := 1$, $\log 0 := -\infty$, $0 \log 0 := 0$ und $a/0 := \infty$ für $a \in (0, \infty)$ gesetzt werden soll. Daher gilt

$$\underset{\tilde{q} \in \Pi_0}{\operatorname{argmin}} K(q, \tilde{q}) = \underset{\tilde{q} \in \Pi_0}{\operatorname{argmax}} \sum_{i=1}^{s} q_i \log \tilde{q}_i, \;\; q \in \Delta_s^0.$$

Weiter ist

$$\frac{1}{m} \log P_{\tilde{q}}(M = l) = \frac{1}{m} \log \left(\frac{m!}{l_1! \ldots l_s!} \tilde{q}_1^{l_1} \cdots \tilde{q}_s^{l_s} \right) = \frac{1}{m} \log \frac{m!}{l_1! \ldots l_s!} + \sum_{i=1}^{s} \frac{l_i}{m} \log \tilde{q}_i,$$
$$l \in \left\{ (i_1, \ldots, i_{s-1})^t ; (i_1, \ldots, i_{s-1}) \in \{0, \ldots, m\}^{s-1}, \sum_{j=1}^{s-1} i_j \leq m \right\}, \;\; \tilde{q} \in \Delta_s^0,$$

und daher

$$\underset{\tilde{q} \in \Pi_0}{\operatorname{argmax}} P_{\tilde{q}}(M = l) = \underset{\tilde{q} \in \Pi_0}{\operatorname{argmax}} \frac{1}{m} \log P_{\tilde{q}}(M = l) = \underset{\tilde{q} \in \Pi_0}{\operatorname{argmax}} \sum_{i=1}^{s} \frac{l_i}{m} \log \tilde{q}_i,$$
$$l \in \left\{ (i_1, \ldots, i_{s-1})^t ; (i_1, \ldots, i_{s-1}) \in \{0, \ldots, m\}^{s-1}, \sum_{j=1}^{s-1} i_j \leq m \right\}.$$

Deshalb ist eine Abbildung $\hat{q} : \Delta_s^0 \to \overline{\Pi_0}$ Maximum-Schätzfunktion, wenn

$$K(q, \Pi_0) = K\big(q, \hat{q}(q)\big) \text{ für alle } q \in \Delta_s^0 \tag{9.2.1}$$

gilt. Eine Abbildung $\hat{q} : \Delta_s^0 \to \overline{\Pi_0}$ die (9.2.1) erfüllt existiert stets, siehe hierzu Lemma 4.3 (a) in der Arbeit [15] von Hoeffding. Daher existiert auch stets eine Maximum-Schätzfunktion für $q \in \Pi$ unter H_0. Gilt $\hat{q} : (\Delta_s^0, \mathfrak{B}^{s-1}_{\Delta_s^0}) \to (\overline{\Pi_0}, \mathfrak{B}^{s-1}_{\overline{\Pi_0}})$, so existiert also auch ein Maximum-Likelihood-Schätzer für $q \in \Pi$ unter H_0.

Nun betrachten wir ein allgemeines Zweistichprobenproblem. Es sei also mit $O :=$

[11] Dieser Maximum-Likelihood-Schätzer für $q \in \Pi$ unter H_0 nimmt also nicht zwingend Werte in Π_0 an.

$(O_1,\ldots,O_{s-1})^t \sim \mathfrak{M}(m;r_1,\ldots,r_s)$, $r := (r_1,\ldots,r_{s-1})^t \in \Delta_s^0$, ein weiterer multinomialverteilter Zufallsvektor gegeben, wobei M und O unabhängig seien.

Für $\emptyset \neq \Pi_0' \subsetneq \Pi' = \Delta_s^2$ betrachten wir das Testproblem

$$\mathrm{H}_0':\ (q,r)\in\Pi_0',\ \mathrm{K}_0':\ (q,r)\in\Pi_1' := \Pi'\setminus\Pi_0'.$$

Eine Abbildung $(\hat{q},\hat{r}) : (\Delta_s^0)^2 \to \overline{\Pi_0'}$ heiße eine Maximum-Schätzfunktion für $(q,r)\in\Pi'$ unter H_0', falls

$$P_{(\hat{q},\hat{r})(l/m,j/m)}(M=l,O=j) = \sup_{(q,r)\in\Pi_0'} P_{q,r}(M=l,O=j),$$
$$l,j \in \Big\{(i_1,\ldots,i_{s-1})^t;(i_1,\ldots,i_{s-1})\in\{0,\ldots,m\}^{s-1},\sum_{j=1}^{s-1} i_j \le m\Big\},$$

erfüllt ist. Wir schreiben $(\hat{q},\hat{r}) := ((\hat{q}_1,\ldots,\hat{q}_{s-1})^t,(\hat{r}_1,\ldots,\hat{r}_{s-1})^t)$. Ist $(\hat{q},\hat{r})$ Maximum-Schätzfunktion für $(q,r)\in\Pi'$ unter H_0', und gilt $(\hat{q},\hat{r}) : ((\Delta_s^0)^2,\mathfrak{B}^{2(s-1)}_{(\Delta_s^0)^2}) \to (\overline{\Pi_0'},\mathfrak{B}^{2(s-1)}_{\overline{\Pi_0'}})$, so ist $(\hat{q},\hat{r})(M/m,O/m)$ Maximum-Likelihood-Schätzer[12] für $(q,r)\in\Pi'$ unter H_0' bei zugrunde liegendem (M,O).

Es besitzt die Kullback-Leibler-Information aus der Verteilung von (M,O) für $m=1$ die Gestalt

$$K\big((q,r),(\tilde{q},\tilde{r})\big) = \sum_{i,j=1}^{s} q_i r_j \log\frac{q_i r_j}{\tilde{q}_i\tilde{r}_j} = \sum_{i=1}^{s} q_i\log\frac{q_i}{\tilde{q}_i} + \sum_{j=1}^{s} r_j\log\frac{r_j}{\tilde{r}_j}$$
$$= \sum_{i=1}^{s} q_i\log q_i - \sum_{i=1}^{s} q_i\log\tilde{q}_i + \sum_{j=1}^{s} r_j\log r_j - \sum_{j=1}^{s} r_j\log\tilde{r}_j,\ (q,r),(\tilde{q},\tilde{r})\in(\Delta_s^0)^2.$$

Daher gilt

$$\underset{(\tilde{q},\tilde{r})\in\Pi_0'}{\operatorname{argmin}} K\big((q,r),(\tilde{q},\tilde{r})\big) = \underset{(\tilde{q},\tilde{r})\in\Pi_0'}{\operatorname{argmax}}\Big(\sum_{i=1}^{s} q_i\log\tilde{q}_i + \sum_{j=1}^{s} r_j\log\tilde{r}_j\Big),\ (q,r)\in(\Delta_s^0)^2.$$

Weiter ist

$$\frac{1}{m}\log P_{(\tilde{q},\tilde{r})}(M=l,O=j) = \frac{1}{m}\log\Big(\frac{m!}{l_1!\ldots l_s!}\tilde{q}_1^{l_1}\cdots\tilde{q}_s^{l_s}\frac{m!}{j_1!\ldots j_s!}\tilde{r}_1^{j_1}\cdots\tilde{r}_s^{j_s}\Big)$$
$$=\frac{1}{m}\log\frac{m!}{l_1!\ldots l_s!} + \sum_{i=1}^{s}\frac{l_i}{m}\log\tilde{q}_i + \frac{1}{m}\log\frac{m!}{j_1!\ldots j_s!} + \sum_{i=1}^{s}\frac{j_i}{m}\log\tilde{r}_i,$$
$$l,j \in \Big\{(i_1,\ldots,i_{s-1})^t;(i_1,\ldots,i_{s-1})\in\{0,\ldots,m\}^{s-1},\sum_{j=1}^{s-1} i_j \le m\Big\},\ (\tilde{q},\tilde{r})\in(\Delta_s^0)^2,$$

[12]Dieser Maximum-Likelihood-Schätzer für $(q,r)\in\Pi'$ unter H_0' nimmt also nicht zwingend Werte in Π_0' an.

und daher

$$\underset{(\tilde q,\tilde r)\in\Pi_0'}{\operatorname{argmax}} P_{(\tilde q,\tilde r)}(M=l,O=j)=\underset{(\tilde q,\tilde r)\in\Pi_0'}{\operatorname{argmax}}\frac{1}{m}\log P_{(\tilde q,\tilde r)}(M=l,O=j)$$

$$=\underset{(\tilde q,\tilde r)\in\Pi_0'}{\operatorname{argmax}}\Big(\sum_{i=1}^{s}\frac{l_i}{m}\log\tilde q_i+\sum_{i=1}^{s}\frac{j_i}{m}\log\tilde r_i\Big),\ l,j\in\Big\{(i_1,\dots,i_{s-1})\in\{0,\dots,m\}^{s-1};\sum_{j=1}^{s-1}i_j\le m\Big\}.$$

Deshalb ist eine Abbildung $(\hat q,\hat r):(\Delta_s^0)^2\to\overline{\Pi_0'}$ Maximum-Schätzfunktion, wenn

$$K\big((q,r),\Pi_0\big)=K\big((q,r),(\hat q,\hat r)(q,r)\big)\text{ für alle }(q,r)\in(\Delta_s^0)^2 \tag{9.2.2}$$

gilt. Existiert eine Abbildung $(\hat q,\hat r):(\Delta_s^0)^2\to\overline{\Pi_0'}$ die (9.2.2) erfüllt, so existiert auch stets eine Maximum-Schätzfunktion für $(q,r)\in\Pi'$ unter H_0'. Gilt $\hat q:((\Delta_s^0)^2,\mathfrak{B}^{2(s-1)}_{(\Delta_s^0)^2})\to(\overline{\Pi_0'},\mathfrak{B}^{2(s-1)}_{\overline{\Pi_0'}})$, so existiert in diesem Fall also auch ein Maximum-Likelihood-Schätzer für $(q,r)\in\Pi'$ unter H_0'.

Wir werden nun mit Hilfe von Lemma 4.3 (a) in der Arbeit [15] von Hoeffding im nächsten Lemma beweisen, dass eine Abbildung $(\hat q,\hat r):(\Delta_s^0)^2\to\overline{\Pi_0'}$, die (9.2.2) erfüllt, stets existiert. Dazu sei $S:=(S_{1,1},\dots,S_{s-1,s})^t\sim\mathfrak{M}(m;v_{1,1},\dots,v_{s,s})$, mit $v:=(v_{1,1},\dots,v_{s-1,s})^t\in\Delta_{s^2}^0$. Ferner sei

$$w.:=(w_{1\cdot},\dots,w_{s-1\cdot},w_{\cdot1},\dots,w_{\cdot s-1})^t,\ w\in\Delta_{s^2}^0,$$

und es sei

$$\begin{aligned}\tilde\Pi_0:=&\big\{(v_{1,1},\dots,v_{s-1,s})^t\in\Delta_{s^2}^0;v.\in\Pi_0'\big\}\\&\cap\big\{(v_{1,1},\dots,v_{s-1,s})^t\in\Delta_{s^2}^0;v_{i,j}=v_{i\cdot}v_{\cdot j},\ (i,j)\in\{1,\dots,s\}^2\big\}.\end{aligned}$$

9.2.1 Lemma: *Es existiert eine Abbildung $\hat v:\Delta_{s^2}^0\to\overline{\tilde\Pi_0}$, $\hat v:=(\hat v_{1,1},\dots,\hat v_{s-1,s})^t$, für die*

$$K(w,\tilde\Pi_0)=K\big(w,\hat v(w)\big)\ \textit{für alle}\ w\in\Delta_{s^2}^0 \tag{9.2.3}$$

gilt, wobei hier $K(w,\tilde w)$, $w,\tilde w\in\Delta_{s^2}^0$, die Kullback-Leibler-Information aus S sei. Für diese Abbildung gilt auch

$$\sum_{i,j=1}^{s}w_{i,j}\log\hat v_{i,j}(w)=\sum_{i=1}^{s}w_{i\cdot}\log\hat v_{i\cdot}(w)+\sum_{j=1}^{s}w_{\cdot j}\log\hat v_{\cdot j}(w),\ w\in\Delta_{s^2}^0,$$

und es hängen $\hat v_{i\cdot}(w)$ und $\hat v_{\cdot j}(w)$ nur von $w.$ ab, $w\in\Delta_{s^2}^0$. Weiter existiert eine Abbildung $(\hat q,\hat r):=((\hat q_1,\dots,\hat q_{s-1})^t,(\hat r_1,\dots,\hat r_{s-1})^t)$, die (9.2.2) erfüllt und die festgelegt ist durch

$$\hat q_i(w.)=\hat v_{i\cdot}(w),\ \hat r_j(w.)=\hat v_{\cdot j}(w),\ w\in\Delta_{s^2}^0,\ (i,j)\in\{1,\dots,s\}^2.$$

Beweis: Nach Lemma 4.3 (a) in [15] existiert eine Abbildung $\hat{v} : \Delta^0_{s^2} \to \overline{\tilde{\Pi}_0}$ die (9.2.3) erfüllt. Für diese gilt

$$\begin{aligned}\sum_{i,j=1}^{s} w_{i,j} \log \hat{v}_{i,j}(w) &= \sup_{v \in \tilde{\Pi}_0} \sum_{i,j=1}^{s} w_{i,j} \log v_{i,j} = \sup_{v \in \tilde{\Pi}_0} \sum_{i,j=1}^{s} w_{i,j} \log(v_{i\cdot} v_{\cdot j}) \\ &= \sup_{v \in \tilde{\Pi}_0} \Bigg(\sum_{i=1}^{s} w_{i\cdot} \log v_{i\cdot} + \sum_{j=1}^{s} w_{\cdot j} \log v_{\cdot j} \Bigg),\end{aligned}$$

und wegen

$$\begin{aligned}\overline{\tilde{\Pi}_0} \subset & \overline{\big\{ (v_{1,1}, \ldots, v_{s-1,s})^t \in \Delta^0_{s^2}; v_{\cdot} \in \Pi'_0 \big\}} \\ & \cap \big\{ (v_{1,1}, \ldots, v_{s-1,s})^t \in \Delta^0_{s^2}; v_{i,j} = v_{i\cdot} v_{\cdot j},\ (i,j) \in \{1, \ldots, s\}^2 \big\}\end{aligned}$$

auch $\hat{v}_{i,j}(w) = \hat{v}_{i\cdot}(w) \hat{v}_{\cdot j}(w)$, $(i,j) \in \{1, \ldots, s\}$, $w \in \Delta^0_{s^2}$, und damit

$$\begin{aligned}\sum_{i,j=1}^{s} w_{i,j} \log \hat{v}_{i,j}(w) &= \sum_{i,j=1}^{s} w_{i,j} \log \big(\hat{v}_{i\cdot}(w) \hat{v}_{\cdot j}(w) \big) \\ &= \sum_{i=1}^{s} w_{i\cdot} \log \hat{v}_{i\cdot}(w) + \sum_{j=1}^{s} w_{\cdot j} \log \hat{v}_{\cdot j}(w),\ w \in \Delta^0_{s^2},\end{aligned}$$

also auch

$$\sum_{i=1}^{s} w_{i\cdot} \log \hat{v}_{i\cdot}(w) + \sum_{j=1}^{s} w_{\cdot j} \log \hat{v}_{\cdot j}(w) = \sup_{v \in \tilde{\Pi}_0} \Bigg(\sum_{i=1}^{s} w_{i\cdot} \log v_{i\cdot} + \sum_{j=1}^{s} w_{\cdot j} \log v_{\cdot j} \Bigg),\ w \in \Delta^0_{s^2}.$$

Damit hängen $\hat{v}_{i\cdot}(w)$ und $\hat{v}_{\cdot j}(w)$ nur von $w_{\cdot}$ ab, $w \in \Delta^0_{s^2}$. Somit lässt sich eine Abblidung $(\hat{q}, \hat{r}) : (\Delta^0_s)^2 \to \overline{\Pi'_0}$ durch die Eigenschaft

$$\hat{q}_i(w_{\cdot}) = \hat{v}_{i\cdot}(w),\ \hat{r}_j(w_{\cdot}) = \hat{v}_{\cdot j}(w),\ w \in \Delta^0_{s^2},\ (i,j) \in \{1, \ldots, s\}^2,$$

wohldefinieren, die

$$\sum_{i=1}^{s} v_i \log \hat{q}_i(v,w) + \sum_{j=1}^{s} w_j \log \hat{r}_j(v,w) = \sup_{(q,r) \in \Pi'_0} \Bigg(\sum_{i=1}^{s} v_i \log q_i + \sum_{j=1}^{s} w_j \log r_j \Bigg),$$
$$(v,w) \in (\Delta^0_s)^2,$$

leistet, und somit auch Gleichung (9.2.2) erfüllt. □

9.3 Verbundener Stichprobenfall

Wir müssen hier Abhängigkeiten zwischen X_1 und Y_1 berücksichtigen. Zur Behandlung des Testproblems der marginalen Homogenität, d.h.

$$\mathrm{H}: \ (p_{1\cdot}, \ldots, p_{d-1\cdot}) = (p_{\cdot 1}, \ldots, p_{\cdot d-1}), \ \mathrm{K}: \ (p_{1\cdot}, \ldots, p_{d-1\cdot}) \neq (p_{\cdot 1}, \ldots, p_{\cdot d-1}),$$

schlagen wir als Testgröße die Likelihood-Quotienten-Testgröße

$$T_{1,n} := 2 \sum_{i,j=1}^{d} N_{i,j} \log \frac{N_{i,j}}{n\hat{p}(N/n)_{i,j}}, \tag{9.3.1}$$

oder die Waldsche Testgröße, hier χ^2-Testgröße,

$$T_{2,n} := \sum_{i,j=1}^{d} \frac{\big(N_{i,j} - n\hat{p}(N/n)_{i,j}\big)^2}{n\hat{p}(N/n)_{i,j}} \tag{9.3.2}$$

vor, wobei $\hat{p} = (\hat{p}_{1,1}, \ldots, \hat{p}_{d-1,d})^t$ eine Maximum-Schätzfunktion für $p = (p_{1,1}, \ldots, p_{d-1,d})^t$ unter H ist. Diese existiert; siehe Unterabschnitt 9.2.

Es ist anzumerken, dass sowohl die Likelihood-Quotienten-Testgröße als auch die Waldsche Testgröße nur auf dem Zufallsvektor $N = \sum_{i=1}^{n} N_i$ basiert.

Betrachte für $q \in \Delta_{d^2}^0$ das konvexe Optimierungsproblem der Minimierung der Abbildung

$$p \longmapsto -\sum_{i,j=1}^{d} q_{i,j} \log p_{i,j}$$

unter den Nebenbedingungen

$$p_{i\cdot} = p_{\cdot i}, \ p_{i,j} \geq 0, \ (i,j) \in \{1, \ldots, d\}^2, \ \sum_{i,j=1}^{d} p_{i,j} = 1.$$

Eine Lösung dieses Optimierungsproblems existiert und liefert die genannte Maximum-Schätzfunktion $\hat{p}$ für den zugrunde liegenden Parameter p unter H; siehe Unterabschnitt 9.2.

Betrachten wir dagegen das Optimierungsproblem für $q \in \Delta_{d^2}$ der Minimierung der Abbildung

$$p \longmapsto -\sum_{i,j=1}^{d} q_{i,j} \log p_{i,j} \tag{9.3.3}$$

unter den Nebenbedingungen

$$p_{i\cdot} = p_{\cdot i},\ p_{i,j} \geq 0,\ (i,j) \in \{1,\dots,d\}^2,\ \sum_{i,j=1}^{d} p_{i,j} = 1, \tag{9.3.4}$$

so muss wegen $q_{i,j} > 0$ für alle $(i,j) \in \{1,\dots,d\}^2$ für die Lösung $\hat{p}$ auch $\hat{p}_{i,j} > 0$ für alle $(i,j) \in \{1,\dots,d\}^2$ gelten, da die zu minimierende Funktion ansonsten den Wert $+\infty$ annehmen würde. Also muss für $q \in \Delta_{d^2}$ für die Lösung des Optimierungsprobelms $\hat{p}$ auch $\hat{p} \in \Delta_{d^2}$ gelten. Da also eine Lösung dieser Minimierungsaufgabe existiert und diese Lösung in Δ_{d^2} liegen muss, ist die Lösung auch ein lokales Minimum von (9.3.3) unter den Nebenbedingungen (9.3.4) und daher nach Theorem 4.3.6 aus [6] notwendiger Weise Komponente der Lösung des Gleichungssystems

$$\begin{aligned}
0 &= -\frac{q_{i,j}}{p_{i,j}} + \lambda_i - \lambda_j + \mu,\\
0 &= p_{i\cdot} - p_{\cdot i},\\
0 &= \sum_{i,j=1}^{d} p_{i,j} - 1,\ (i,j) \in \{1,\dots,d\}^2,\ (p_{1,1},\dots,p_{d,d})^t \in \mathbb{R}^{d^2},\ \lambda \in \mathbb{R}^d,\ \mu \in \mathbb{R}.
\end{aligned} \tag{9.3.5}$$

Nach Theorem 4.3.7 aus [6] ist das Lösen von (9.3.5) auch ein hinreichendes Kriterium für das globale Minimum der Funktion (9.3.3) unter den Nebenbedingungen (9.3.4). Die gesuchte Lösung von (9.3.5) lässt sich für $d > 2$ nicht explizit angeben und kann nur iterativ numerisch bestimmt werden. Wir wollen diese Lösung wieder mit $\hat{p}$ bezeichnen.

Es ist für jedes $p \in \Delta_{d^2}$ die Abbildung $q \mapsto K(q,p)$ $(\mathfrak{B}^{d^2-1}_{\Delta_{d^2}}, \mathfrak{B})$-messbar und für jedes $p \in \Delta_{d^2}$ die Abbildung $q \mapsto K(p,q)$ stetig auf Δ_{d^2}. Zudem ist die Abbildung $q \mapsto K(q,\hat{p}(q))$ $(\mathfrak{B}^{d^2-1}_{\Delta_{d^2}}, \mathfrak{B})$-messbar. Somit folgt aus Hilfssatz 6.7 in [27], dass die Lösung $\hat{p}$ von (9.3.5) als Funktion von $q \in \Delta_{d^2}$ ohne Einschränkung als $(\mathfrak{B}^{d^2-1}_{\Delta_{d^2}}, \mathfrak{B})$-messbar angenommen werden kann.

Somit folgt aus der allgemeinen Theorie für Likelihood-Quotienten-Tests und Waldsche Tests, dass bei Gültigkeit von H

$$T_{k,n} \xrightarrow{\mathrm{v}} \chi^2_{d-1} \text{ für } n \to \infty,\ k = 1,2,$$

gilt. Siehe hierfür z.B. [27]. Für das Testproblem H gegen K schlagen wir daher den Likelihood-Quotienten-Test bzw. den Waldschen Test

$$\text{„verwirf } \mathrm{H}_1, \text{ falls } T_{k,n} > \chi^2_{d-1;1-\alpha}\text{“},\ k = 1,2,$$

vor, wobei $\chi^2_{d-1;1-\alpha}$ das $(1-\alpha)$-Quantil der Chi-Quadrat-Verteilung mit $d-1$ Freiheitsgraden sei. Wie wir weiter unten sehen werden, ist der Likelihood-Quotienten-Test optimal im Bahadur- und Hodges-Lehmann Sinn.

9.4 Unabhängiger Stichprobenfall

Gehen wir davon aus, dass mit $X_1', \dots, X_n'$ und $Y_1', \dots, Y_n'$ zwei voneinander unabhängige Stichproben vorliegen. Zur Behandlung des Testproblems der Homogenität, d.h.

$$\mathrm{H}: \ (p_{1\cdot}, \dots, p_{d-1\cdot}) = (p_{\cdot 1}, \dots, p_{\cdot d-1}), \ \mathrm{K}: \ (p_{1\cdot}, \dots, p_{d-1\cdot}) \neq (p_{\cdot 1}, \dots, p_{\cdot d-1}),$$

schlagen wir als Testgröße die Likelihood-Quotienten-Testgröße

$$T_{1,n}' := 2 \sum_{i=1}^{d} N_{i\cdot}' \log \frac{N_{i\cdot}'}{n\hat{p}_{i\cdot}'(N_{*\cdot}'/n, N_{\cdot *}'/n)} + 2 \sum_{i=1}^{d} N_{\cdot i}' \log \frac{N_{\cdot i}'}{n\hat{p}_{\cdot i}'(N_{*\cdot}'/n, N_{\cdot *}'/n)}, \tag{9.4.1}$$

oder die Waldsche Testgröße, hier χ^2-Testgröße,

$$T_{2,n}' := \sum_{i=1}^{d} \frac{\big(N_{i\cdot}' - n\hat{p}_{i\cdot}'(N_{*\cdot}'/n, N_{\cdot *}'/n)\big)^2}{n\hat{p}_{i\cdot}'(N_{*\cdot}'/n, N_{\cdot *}'/n)} + \sum_{i=1}^{d} \frac{\big(N_{\cdot i}' - n\hat{p}_{\cdot i}'(N_{*\cdot}'/n, N_{\cdot *}'/n)\big)^2}{n\hat{p}_{\cdot i}'(N_{*\cdot}'/n, N_{\cdot *}'/n)} \tag{9.4.2}$$

vor, wobei $(\hat{p}_{*\cdot}', \hat{p}_{\cdot *}') := ((\hat{p}_{1\cdot}', \dots, \hat{p}_{d-1\cdot}')^t, (\hat{p}_{\cdot 1}', \dots, \hat{p}_{\cdot d-1}')^t)$ eine Maximum-Schätzfunktion für $(p_{*\cdot}, p_{\cdot *}) := ((p_{1\cdot}, \dots, p_{d-1\cdot})^t, (p_{\cdot 1}, \dots, p_{\cdot d-1})^t)$ unter H ist. Diese existiert; siehe Unterabschnitt 9.2.

Es ist anzumerken, dass sowohl die Likelihood-Quotienten-Testgröße als auch die Waldsche Testgröße nur auf dem Zufallsvektor $(N_{*\cdot}', N_{\cdot *}')^t = \sum_{i=1}^n N_i'$ basiert.

Betrachte für $(q_{*\cdot}, q_{\cdot *}) := ((q_{1\cdot}, \dots, q_{d-1\cdot})^t, (q_{\cdot 1}, \dots, q_{\cdot d-1})^t) \in (\Delta_d^0)^2$ das konvexe Optimierungsproblem der Minimierung der Abbildung

$$(p_{*\cdot}, p_{\cdot *}) \longmapsto -\sum_{i=1}^{d} q_{i\cdot} \log p_{i\cdot} - \sum_{i=1}^{d} q_{\cdot i} \log p_{\cdot i}$$

unter den Nebenbedingungen

$$p_{i\cdot} = p_{\cdot i}, \ p_{i\cdot} \geq 0, \ p_{\cdot i} \geq 0, \ i \in \{1, \dots, d\}, \ \sum_{i=1}^{d} p_{i\cdot} = 1, \ \sum_{i=1}^{d} p_{\cdot i} = 1.$$

Die Lösung dieses Optimierungsproblems existiert und liefert die genannte Maximums-Schätzfunktion $(\hat{p}_{*\cdot}', \hat{p}_{\cdot *}')$ für den zugrunde liegenden Parameter $(p_{*\cdot}, p_{\cdot *})$ unter H; siehe Unterabschnitt 9.2.

Betrachten wir dagegen das Optimierungsproblem für $(q_{*\cdot}, q_{\cdot *}) \in \Delta_d^2$ der Minimierung der Abbildung

$$(p_{*\cdot}, p_{\cdot *}) \longmapsto -\sum_{i=1}^{d} q_{i\cdot} \log p_{i\cdot} - \sum_{i=1}^{d} q_{\cdot i} \log p_{\cdot i} \tag{9.4.3}$$

unter den Nebenbedingungen

$$p_{i\cdot} = p_{\cdot i},\ p_{i\cdot} \geq 0,\ p_{\cdot i} \geq 0,\ i \in \{1, \dots, d\},\ \sum_{i=1}^{d} p_{i\cdot} = 1,\ \sum_{i=1}^{d} p_{\cdot i} = 1, \tag{9.4.4}$$

so muss wegen $q_{i\cdot} > 0$ und $q_{\cdot i} > 0$ für alle $i \in \{1, \dots, d\}$ für die Lösung $(\hat{p}'_{*\cdot}, \hat{p}'_{\cdot *})$ auch $\hat{p}'_{i\cdot} > 0$ und $\hat{p}'_{\cdot i} > 0$ für alle $i \in \{1, \dots, d\}$ gelten, da die zu minimierende Funktion ansonsten den Wert $+\infty$ annehmen würde. Also muss für $(q_{*\cdot}, q_{\cdot *}) \in \Delta_d^2$ für die Lösung des Optimierungsprobelms $(\hat{p}'_{*\cdot}, \hat{p}'_{\cdot *})$ auch $(\hat{p}'_{*\cdot}, \hat{p}'_{\cdot *}) \in \Delta_d^2$ gelten. Da also eine Lösung dieser Minimierungsaufgabe existiert und diese Lösung in Δ_d^2 liegen muss, ist die Lösung dieser Minimierungsaufgabe auch ein lokales Minimum von (9.4.3) unter den Nebenbedingungen (9.4.4) und daher nach Theorem 4.3.6 aus [6] notwendiger Weise Komponente der Lösung des Gleichungssystems

$$\begin{aligned}
0 &= -\frac{q_{i\cdot}}{p_{i\cdot}} + \lambda_i + \mu_1, \\
0 &= -\frac{q_{\cdot i}}{p_{\cdot i}} - \lambda_i + \mu_2, \\
0 &= p_{i\cdot} - p_{\cdot i}, \\
0 &= \sum_{i=1}^{d} p_{i\cdot} - 1 \\
0 &= \sum_{i=1}^{d} p_{\cdot i} - 1,\ i \in \{1, \dots, d\},\ ((p_{1\cdot}, \dots, p_{d\cdot})^t, (p_{\cdot 1}, \dots, p_{\cdot d})^t) \in \mathbb{R}^d \times \mathbb{R}^d,\ \lambda \in \mathbb{R}^d,\ \mu \in \mathbb{R}^2.
\end{aligned} \tag{9.4.5}$$

Nach Theorem 4.3.7 aus [6] ist das Lösen von (9.4.5) auch ein hinreichendes Kriterium für das globale Minimum der Funktion (9.4.3) unter den Nebenbedingungen (9.4.4). Wir wollen die Lösung von (9.4.5) wieder mit $(\hat{p}'_{*\cdot}, \hat{p}'_{\cdot *})$ bezeichnen. Als Lösung von (9.4.5) ergibt sich

$$(\hat{p}'_{*\cdot}, \hat{p}'_{\cdot *})(q_{*\cdot}, q_{\cdot *}) = \Big(\frac{q_{*\cdot} + q_{\cdot *}}{2}, \frac{q_{*\cdot} + q_{\cdot *}}{2}\Big).$$

Damit besitzen die Testgrößen die Gestalt

$$T'_{1,n} = 2\sum_{i=1}^{d} N'_{i\cdot} \log \frac{N'_{i\cdot}}{\frac{N'_{i\cdot}+N'_{\cdot i}}{2}} + 2\sum_{i=1}^{d} N'_{\cdot i} \log \frac{N'_{\cdot i}}{\frac{N'_{i\cdot}+N'_{\cdot i}}{2}} \tag{9.4.6}$$

und

$$T'_{2,n} = \sum_{i=1}^{d} \frac{\big(N'_{i\cdot} - \frac{N'_{i\cdot}+N'_{\cdot i}}{2}\big)^2}{\frac{N'_{i\cdot}+N'_{\cdot i}}{2}} + \sum_{i=1}^{d} \frac{\big(N'_{\cdot i} - \frac{N'_{i\cdot}+N'_{\cdot i}}{2}\big)^2}{\frac{N'_{i\cdot}+N'_{\cdot i}}{2}}. \tag{9.4.7}$$

Die Lösung von (9.4.5) ist als Funktion von $(q_{*\cdot}, q_{\cdot *}) \in \Delta_d^2$ stetig und damit als eine solche Abbildung auch messbar.

Somit folgt aus der allgemeinen Theorie für Likelihood-Quotienten-Tests und Waldsche Tests, dass bei Gültigkeit von H

$$T'_{k,n} \xrightarrow{\text{v}} \chi^2_{d-1} \text{ für } n \to \infty,\ k = 1, 2,$$

gilt. Siehe hierfür z.B. [27]. Für das Testproblem H gegen K schlagen wir daher den Likelihood-Quotienten-Test bzw. den Waldschen Test

$$\text{„verwirf H}_1,\ \text{falls } T'_{k,n} > \chi^2_{d-1;1-\alpha}\text{“},\ k = 1, 2,$$

vor, wobei $\chi^2_{d-1;1-\alpha}$ das $(1-\alpha)$-Quantil der Chi-Quadrat-Verteilung mit $d-1$ Freiheitsgraden sei. Wie wir weiter unten sehen werden, ist der Likelihood-Quotienten-Test optimal im Bahadur- und Hodges-Lehmann Sinn.

9.5 Bahadur-Effizienz der Tests

Die auf den Testgrößen (9.3.1) und(9.4.6) basierenden vorgeschlagenen Likelihood-Qutienten-Tests sind jeweils optimal im Bahadur Sinn. Dies folgt aus Theorem 2 der Arbeit [3] von Bahadur; beachte dabei Bemerkung 3. (b) und (e).

Somit ergibt sich für die Bahadur-Steigung im unabhängigen Stichprobenfall bei zugrunde liegender Alternative $(p_{*\cdot}, p_{\cdot *}) \in \Delta_d^2$

$$\sum_{i=1}^{d} p_{i\cdot} \log \frac{p_{i\cdot}}{\frac{p_{i\cdot}+p_{\cdot i}}{2}} + \sum_{i=1}^{d} p_{\cdot i} \log \frac{p_{\cdot i}}{\frac{p_{i\cdot}+p_{\cdot i}}{2}},$$

und für die Bahadur-Steigung im verbundenen Stichprobenfall bei zugrunde liegender Alternative $p \in \Delta_{d^2}$

$$\sum_{i,j=1}^{d^2} p_{i,j} \log \frac{p_{i,j}}{\hat{p}(p)_{i,j}},$$

wobei $\hat{p} = (\hat{p}_{1,1}, \ldots, \hat{p}_{d-1,d})^t$ Lösung von (9.3.5) ist. Somit kann auch die Bahadur-Effizienz der Likelihood-Quotienten-Tests basierend auf den Testgrößen (9.3.1) und (9.4.6) für $d > 2$ nicht explizit angegeben werden.

Allerdings können hier die in Unterabschnitt 4.3 hergeleiteten allgemeinen Resultate zur lokalen Bahadur-Effizienz von optimalen Tests im Bahadur-Sinn angewendet werden. Hierzu müssen zunächst passende Parametrisierungen gemäß den Ausführungen am Anfang des Unterabschnitts 4.3 gewählt werden.

Wir definieren

$$z_i := p_{i\cdot} - p_{\cdot i},\ i = 1, \ldots, d-1,$$
$$s_i := p_{i\cdot} + p_{\cdot i},\ i = 1, \ldots, d-1.$$

Dann lässt sich das Testproblem äquivalent gemäß

$$\mathrm{H}:\ (z_1, \ldots, z_{d-1}) = 0,\ \mathrm{K}:\ (z_1, \ldots, z_{d-1}) \neq 0$$

formulieren. Es sei

$$\Theta = \Bigg\{ \big((z_1, \ldots, z_{d-1})^t, (s_1, \ldots, s_{d-1}, p_{1,1}, \ldots, p_{d-1,d-1})^t, (s_1, \ldots, s_{d-1})^t\big)$$
$$\in \mathbb{R}^{d-1} \times \mathbb{R}^{d-1+(d-1)^2} \times \mathbb{R}^{d-1}; p_{i,j} > 0, \frac{1}{2}(z_i + s_i) - \sum_{k=1}^{d-1} p_{i,k} > 0,$$
$$\frac{1}{2}(-z_j + s_j) - \sum_{k=1}^{d-1} p_{k,j} > 0, (i,j) \in \{1, \ldots, d-1\}^2, \sum_{i=1}^{d-1} s_i - \sum_{i,j=1}^{d-1} p_{i,j} < 1 \Bigg\} \subset \mathbb{R}^{d^2+d-2}.$$

Hierbei schreiben wir stets abkürzend $(p_{1,1}, \ldots, p_{d-1,d-1})^t$ für

$$(p_{1,1}, \ldots, p_{d-1,1}, p_{1,2}, \ldots, p_{d-1,2}, \ldots, p_{1,d-1}, \ldots, p_{d-1,d-1})^t$$

In der Notation aus Unterabschnitt 4.3 bedeutet dies

$$(\vartheta_1, \vartheta_2, \vartheta_2') = \big((z_1, \ldots, z_{d-1})^t, (s_1, \ldots, s_{d-1}, p_{1,1}, \ldots, p_{d-1,d-1})^t, (s_1, \ldots, s_{d-1})^t\big) \in \Theta.$$

Somit ist

$$\Theta^1 = \big\{(z_1, \ldots, z_{d-1})^t \in \mathbb{R}^{d-1}; \exists\ \big((s_1, \ldots, s_{d-1}, p_{1,1}, \ldots, p_{d-1,d-1})^t, (s_1, \ldots, s_{d-1})^t\big)$$
$$\in \mathbb{R}^{d-1+(d-1)^2} \times \mathbb{R}^{d-1} : \big((z_1, \ldots, z_{d-1})^t, (s_1, \ldots, s_{d-1}, p_{1,1}, \ldots, p_{d-1,d-1})^t, (s_1, \ldots, s_{d-1})^t\big) \in \Theta\big\}$$

und $\Theta_0^1 = \{0\} \subsetneq \Theta^1$, also

$$\Theta_0 = \Bigg\{ \big((0, \ldots, 0)^t, (s_1, \ldots, s_{d-1}, p_{1,1}, \ldots, p_{d-1,d-1})^t, (s_1, \ldots, s_{d-1})^t\big)$$
$$\in \mathbb{R}^{d-1} \times \mathbb{R}^{d-1+(d-1)^2} \times \mathbb{R}^{d-1}; p_{i,j} > 0, \frac{1}{2}s_i - \sum_{k=1}^{d-1} p_{i,k} > 0,$$
$$\frac{1}{2}s_j - \sum_{k=1}^{d-1} p_{k,j} > 0, (i,j) \in \{1, \ldots, d-1\}^2, \sum_{i=1}^{d-1} s_i - \sum_{i,j=1}^{d-1} p_{i,j} < 1 \Bigg\}.$$

Es liegt somit die Situation aus Bemerkung 4.2.8 vor. Weiter ist

$$\begin{aligned}\tilde{\Theta} =\Bigg\{ & \big((z_1,\dots,z_{d-1})^t,(s_1,\dots,s_{d-1},p_{1,1},\dots,p_{d-1,d-1})^t\big) \in \mathbb{R}^{d-1}\times\mathbb{R}^{d-1+(d-1)^2};\\ & p_{i,j}>0,\frac{1}{2}(z_i+s_i)-\sum_{k=1}^{d-1}p_{i,k}>0,\frac{1}{2}(-z_j+s_j)-\sum_{k=1}^{d-1}p_{k,j}>0,(i,j)\in\{1,\dots,d-1\}^2,\\ & \sum_{i=1}^{d-1}s_i-\sum_{i,j=1}^{d-1}p_{i,j}<1\Bigg\}\subset\mathbb{R}^{d^2-1}.\end{aligned}$$

und

$$\begin{aligned}\tilde{\Theta}' =\Bigg\{ & \big((z_1,\dots,z_{d-1})^t,(s_1,\dots,s_{d-1})^t\big)\in\mathbb{R}^{d-1}\times\mathbb{R}^{d-1};z_i+s_i>0,\\ & -z_i+s_i>0,i\in\{1,\dots,d-1\},\frac{1}{2}\sum_{i=1}^{d-1}(z_i+s_i)<1,\frac{1}{2}\sum_{i=1}^{d-1}(-z_i+s_i)<1\Bigg\}\subset\mathbb{R}^{2(d-1)}.\end{aligned}$$

In der Notation aus Unterabschnitt 4.3 bedeutet dies

$$(\vartheta_1,\vartheta_2)=\big((z_1,\dots,z_{d-1})^t,(s_1,\dots,s_{d-1},p_{1,1},\dots,p_{d-1,d-1})^t\big)\in\tilde{\Theta}$$

und

$$(\vartheta_1,\vartheta_2')=\big((z_1,\dots,z_{d-1})^t,(s_1,\dots,s_{d-1})^t\big)\in\tilde{\Theta}'.$$

Zudem ist

$$\begin{aligned}\tilde{\Theta}_0 =\Bigg\{ & \big((0,\dots,0)^t,(s_1,\dots,s_{d-1},p_{1,1},\dots,p_{d-1,d-1})^t\big)\in\mathbb{R}^{d-1}\times\mathbb{R}^{d-1+(d-1)^2};\\ & p_{i,j}>0,\frac{1}{2}s_i-\sum_{k=1}^{d-1}p_{i,k}>0,\frac{1}{2}s_j-\sum_{k=1}^{d-1}p_{k,j}>0,(i,j)\in\{1,\dots,d-1\}^2,\\ & \sum_{i=1}^{d-1}s_i-\sum_{i,j=1}^{d-1}p_{i,j}<1\Bigg\}.\end{aligned}$$

und

$$\tilde{\Theta}_0' =\Bigg\{\big((0,\dots,0)^t,(s_1,\dots,s_{d-1})^t\big)\in\mathbb{R}^{d-1}\times\mathbb{R}^{d-1};s_i>0,i\in\{1,\dots,d-1\},\frac{1}{2}\sum_{i=1}^{d-1}s_i<1\Bigg\}.$$

Drücken wir $p = (p_{1,1}, \ldots, p_{d-1,d})^t$ in der neuen Parametrisierung aus, so ergibt sich

$$\begin{aligned}
p_{i,j} &= p_{i,j},\ (i,j) \in \{1, \ldots, d-1\}^2, \\
p_{i,d} &= \frac{1}{2}(z_i + s_i) - \sum_{k=1}^{d-1} p_{i,k},\ i \in \{1, \ldots, d-1\} \\
p_{d,j} &= \frac{1}{2}(-z_j + s_j) - \sum_{k=1}^{d-1} p_{k,j},\ j \in \{1, \ldots, d-1\} \\
p_{d,d} &= 1 + \sum_{i,j=1}^{d-1} p_{i,j} - \sum_{i=1}^{d-1} s_i. \qquad (9.5.1)
\end{aligned}$$

Insbesondere ist

$$z_i \in \left(2\sum_{k=1}^{d-1} p_{i,k} - s_i, 2\sum_{k=1}^{d-1} p_{k,i} + s_i\right)\ i \in \{1, \ldots, d-1\}. \qquad (9.5.2)$$

Weiter ist

$$\begin{aligned}
p_{i\cdot} &= \frac{1}{2}(z_i + s_i),\ i \in \{1, \ldots, d-1\}, \\
p_{\cdot i} &= \frac{1}{2}(-z_i + s_i),\ i \in \{1, \ldots, d-1\}. \qquad (9.5.3)
\end{aligned}$$

Allgemein ist zu bemerken, dass sowohl die Likelihood-Quotienten-Testgröße als auch die Waldsche Testgröße nicht von der Wahl der Parametrisierung abhängt.

Es bezeichne $\phi : \tilde{\Theta} \to \Delta_{d^2}$ die Bijektion, die durch (9.5.1) gegeben ist. Analog bezeichne $\phi' : \tilde{\Theta}' \to \Delta_d^2$ die Bijektion, die durch (9.5.3) gegeben ist.

Setzen wir

$$\Lambda := \{(\log p_{1,1} - \log p_{d,d}, \ldots, \log p_{d-1,d} - \log p_{d,d})^t; p \in \Delta_{d^2}\}$$

und betrachten die Abbildung

$$\varphi : \Delta_{d^2} \to \Lambda,\ \varphi(p) := (\log p_{1,1} - \log p_{d,d}, \ldots, \log p_{d-1,d} - \log p_{d,d})^t,$$

so können wir festhalten, dass mit $\varphi \circ \phi : \tilde{\Theta} \to \Lambda$ eine bijektive und stetig differenzierbare Abbildung mit überall regulärer Funktionalmatrix existiert, die Gleichung (6.3.2) erfüllt. Daher ist die Familie von Verteilungen von N_1 in der Parametrisierung durch Elemente von Λ eine (d^2-1)-parametrische Exponentialfamilie mit suffizienter Statistik gegeben durch die Identität bezüglich des Zählmaßes auf $\mathbb{N}_0^{d^2-1}$.

Setzen wir

$$\Lambda' := \{(\log p_{1\cdot} - \log p_{d\cdot}, \ldots, \log p_{d-1\cdot} - \log p_{d\cdot},\\ \log p_{\cdot 1} - \log p_{\cdot d}, \ldots, \log p_{\cdot d-1} - \log p_{\cdot d})^t; (p_{*\cdot}, p_{\cdot *}) \in \Delta_d^2\}$$

und betrachten die Abbildung

$$\varphi' : \Delta_d^2 \to \Lambda',\ \varphi(p_{*\cdot}, p_{\cdot *}) := (\log p_{1\cdot} - \log p_{d\cdot}, \ldots, \log p_{d-1\cdot} - \log p_{d\cdot},\\ \log p_{\cdot 1} - \log p_{\cdot d}, \ldots, \log p_{\cdot d-1} - \log p_{\cdot d})^t,$$

so können wir festhalten, dass mit $\varphi' \circ \phi' : \tilde{\Theta}' \to \Lambda'$ eine bijektive und stetig differenzierbare Abbildung mit überall regulärer Funktionalmatrix existiert, die Gleichung (6.3.2) erfüllt. Daher ist die Familie von Verteilungen von N_1' in der Parametrisierung durch Elemente von Λ' eine $2(d-1)$-parametrische Exponentialfamilie mit suffizienter Statistik gegeben durch die Identität bezüglich des Zählmaßes auf $\mathbb{N}_0^{2(d-1)}$.

Für den nächsten Satz definieren wir für $\vartheta = (\vartheta_1, \ldots, \vartheta_{d^2-1})^t \in \tilde{\Theta}$ den Gradienten

$$\nabla_{d^2-1} := \Big(\frac{\partial}{\partial \vartheta_1}, \ldots, \frac{\partial}{\partial \vartheta_{d^2-1}}\Big)^t$$

und für $\vartheta' = (\vartheta_1', \ldots, \vartheta_{2(d-1)}')^t \in \tilde{\Theta}'$ den Gradienten

$$\nabla_{2(d-1)} := \Big(\frac{\partial}{\partial \vartheta_1'}, \ldots, \frac{\partial}{\partial \vartheta_{2(d-1)}'}\Big)^t,$$

Damit definieren wir Matrizen

$$\psi_{N_1}(\vartheta_0) := \big(\nabla_{d^2-1}\, \mathrm{E}_\vartheta(N_1)^t_{|\vartheta=\vartheta_0}\big)^t,\ \vartheta_0 \in \tilde{\Theta}_0,$$

und

$$\psi_{N_1'}(\vartheta_0') := \big(\nabla_{2(d-1)}\, \mathrm{E}_\vartheta(N_1')^t_{|\vartheta'=\vartheta_0'}\big)^t,\ \vartheta_0' \in \tilde{\Theta}_0'.$$

Hierbei handelt es sich um Matrizen der Form (6.1.5). Ist für $\vartheta \in \tilde{\Theta}$ $p = \phi(\vartheta)$, so gilt $\mathrm{E}_\vartheta(N_1) = p$. Ist für $\vartheta' \in \tilde{\Theta}'$ $\binom{p_{*\cdot}}{p_{\cdot *}} = \phi'(\vartheta')$, so gilt $\mathrm{E}_{\vartheta'}(N_1') = \binom{p_{*\cdot}}{p_{\cdot *}}$. Daher hängt die Matrix $\psi_{N_1}(\vartheta_0)$ bzw. $\psi_{N_1'}(\vartheta_0')$ nicht von $\vartheta_0 \in \tilde{\Theta}_0$ bzw. $\vartheta_0' \in \tilde{\Theta}_0'$ ab. Dies wird durch Betrachtung von (9.5.1) bzw. (9.5.3) klar. Wir schreiben also $\psi_{N_1}(\vartheta_0) = \psi_{N_1},\ \vartheta_0 \in \tilde{\Theta}_0$, bzw. $\psi_{N_1'}(\vartheta_0') = \psi_{N_1'},\ \vartheta_0' \in \tilde{\Theta}_0'$.

Die Matrizen ψ_{N_1} und $\psi_{N_1'}$ sind regulär, was man wie folgt begründen kann. Es handelt sich bei der Matrix ψ_{N_1} um die Jacobimatrix der Abbildung ϕ. Die Abbildung ϕ kann als die Einschränkung einer linearen Abbildung mit Definitionsbereich $\mathbb{R}^{d^2-1}$ auf $\tilde{\Theta}$ aufgefasst werden. Wäre diese Jacobimatrix singulär, so wäre $\phi(\tilde{\Theta}) = \Delta_{d^2}$ Teilmenge eines echten Unterraums des $\mathbb{R}^{d^2-1}$, da die Jacobimatrix konstant ist. Dies kann nicht sein. Daher muss die Matrix ψ_{N_1} regulär sein. Genauso kann argumentiert werden, um die Regularität der Matrix $\psi_{N_1'}$ zu begründen.

Es ist $i_{N_1}(\vartheta_0) = \psi_{N_1}^t \mathrm{Cov}_{\vartheta_0}(N_1)^{-1}\psi_{N_1}$, $\vartheta_0 \in \tilde{\Theta}_0$ und $i_{N_1'}(\vartheta_0') = \psi_{N_1'}^t \mathrm{Cov}_{\vartheta_0'}(N_1')^{-1}\psi_{N_1'}$, $\vartheta_0' \in \tilde{\Theta}_0'$, siehe den Beweis von Satz 6.3.10. Da ψ_{N_1} und $\psi_{N_1'}$ regulär sind, sind $i_{N_1}(\vartheta_0)$, $\vartheta_0 \in \tilde{\Theta}_0$, und $i_{N_1'}(\vartheta_0')$, $\vartheta_0' \in \tilde{\Theta}_0'$, positiv definit.

9.5.1 Korollar: *Es sei eine reellwertige Folge $(\nu_i)_{i\in\mathbb{N}}$ mit $\nu_i \neq 0$ für alle $i \in \mathbb{N}$ und $\lim_{i\to\infty} \nu_i = 0$ gegeben. Für $i \in \mathbb{N}$ betrachten wir Alternativen der Form*

$$\vartheta^i := \begin{pmatrix} 0 \\ \vartheta_2 \\ \vartheta_2' \end{pmatrix} + \nu_i \begin{pmatrix} \vartheta_1 \\ 0 \\ 0 \end{pmatrix}$$

mit $\vartheta_1 \in \mathbb{R}^{d-1} \setminus \{0\}$, $(0, \vartheta_2, \vartheta_2') \in \Theta_0$, die $(\nu_i\vartheta_1, \vartheta_2, \vartheta_2') \in \Theta_1$ erfüllen. Dann gilt mit $\vartheta_0 := \left(\begin{smallmatrix} 0 \\ \vartheta_2 \end{smallmatrix}\right)$, $\vartheta_0' := \left(\begin{smallmatrix} 0 \\ \vartheta_2' \end{smallmatrix}\right)$ und $\vartheta^0 := (0, \vartheta_2, \vartheta_2')$ für die Folgen von Tests $T_1 = (T_{1,n})_{n\in\mathbb{N}}$ und $T_1' = (T_{1,n}')_{n\in\mathbb{N}}$

$$e^{\mathrm{LB}}_{T_1',T_1}\big(\vartheta^0, \vartheta_1\big) = \frac{\vartheta_1^t A^{1,1}_{N_1'}(\vartheta_0')\vartheta_1}{\vartheta_1^t A^{1,1}_{N_1}(\vartheta_0)\vartheta_1},$$

mit

$$A^{1,1}_{N_1}(\vartheta_0) := \Big((\mathrm{I}_{(d-1)\times(d-1)}, 0_{(d-1)\times(d^2-d)})\psi_{N_1}^{-1}\mathrm{Cov}_{\vartheta_0}(N_1)\big(\psi_{N_1}^t\big)^{-1} (\mathrm{I}_{(d-1)\times(d-1)}, 0_{(d-1)\times(d^2-d)})^t \Big)^{-1},$$

und

$$A^{1,1}_{N_1'}(\vartheta_0') := \Big((\mathrm{I}_{(d-1)\times(d-1)}, 0_{(d-1)\times(d-1)})\psi_{N_1'}^{-1}\mathrm{Cov}_{\vartheta_0'}(N_1')\big(\psi_{N_1'}^t\big)^{-1} (\mathrm{I}_{(d-1)\times(d-1)}, 0_{(d-1)\times(d-1)})^t \Big)^{-1}.$$

Beweis: Im verbundenen Stichprobenfall sind in der vorliegenden Situation die Bedingungen (C1)-(C4) aus Unterabschnitt 4.3 erfüllt. Zum Nachweis der Bedingung (C4) ist festzustellen, dass die zu Beginn des Unterabschnitts 4.3 eingeführte Menge Δ hier gegeben ist durch

$$\Delta = \Big\{ (s_1, \ldots, s_{d-1}, p_{1,1}, \ldots, p_{d-1,d-1})^t \in \mathbb{R}^{d-1+(d-1)^2}; p_{i,j} > 0, \frac{1}{2}s_i - \sum_{k=1}^{d-1} p_{i,k} > 0, \frac{1}{2}s_j - \sum_{k=1}^{d-1} p_{k,j} > 0, (i,j) \in \{1, \ldots, d-1\}^2, \sum_{i=1}^{d-1} s_i - \sum_{i,j=1}^{d-1} p_{i,j} < 1 \Big\}.$$

Dann ist die zu Beginn des Unterabschnitts 4.3 eingeführte Abbildung g gegeben durch $g : \Delta \to \tilde{\Theta}$,

$$g\big((s_1, \ldots, s_{d-1}, p_{1,1}, \ldots, p_{d-1,d-1})^t\big) = \begin{pmatrix} (0, \ldots, 0)^t \\ (s_1, \ldots, s_{d-1}, p_{1,1}, \ldots, p_{d-1,d-1})^t \end{pmatrix}.$$

Es sei $\vartheta_0 \in \tilde{\Theta}_0 = \partial\tilde{\Theta}_0$. Gemäß der in Unterabschnitt 9.3 präsentierten Überlegungen liegt die Minimalstelle $\hat{p}$ der Abbildung (9.3.3) unter den Nebenbedingungen (9.3.4) in Δ_{d^2}. D.h. für die Minimalstelle $\hat{\eta} \in \overline{\Delta}$ der Abbildung

$$\eta \longmapsto K\big(\vartheta, g(\eta)\big),\ \eta \in \overline{\Delta}, \tag{9.5.4}$$

mit $\vartheta \in \tilde{\Theta}$, muss $\phi(g(\hat{\eta})) \in \Delta_{d^2}$ gelten. Dies bedeutet allerdings, dass auch $\hat{\eta} \in \Delta$ gelten muss. Damit muss die Minimalstelle $\hat{\eta}$ der Abbildung (9.5.4)

$$\int \nabla_h^t \Big(\log \frac{f(z;\vartheta)}{f\big(z; g(\eta)\big)} f(z;\vartheta) \Big)_{|_{\eta=\hat{\eta}}} \mathrm{d}\mu(z) = 0$$

erfüllen. Dabei gelten die in Unterabschnitt 4.3 eingeführten Bezeichnungen, d.h. f ist die Dichte von N_1 bezüglich des Zählmaßes μ auf $\mathbb{N}_0^{d^2-1}$. Da hier alle Bedingungen von Lemma 4.3.1 erfüllt sind, existiert nach Lemma 4.3.1 eine offene Umgebung U_{ϑ_0} um ϑ_0 und eine eindeutige und auf U_{ϑ_0} stetige Abbildung $\hat{\eta} : U_{\vartheta_0} \to \Delta$ mit

$$\int \nabla_h^t \Big(\log \frac{f(z;\vartheta)}{f\big(z; g(\eta)\big)} f(z;\vartheta) \Big)_{|_{\eta=\hat{\eta}(\vartheta)}} \mathrm{d}\mu(z) = 0 \text{ für alle } \vartheta \in U_{\vartheta_0}.$$

Somit ist wegen der Eindeutigkeit dieser Abbildung auch Bedingung (C4) aus Unterabschnitt 4.3 im verbundenen Stichprobenfall erfüllt.

Im unabhängigen Stichprobenfall sind in der vorliegenden Situation die Bedingungen (C1)-(C4) aus Unterabschnitt 4.3 erfüllt.

Also folgt insgesamt aufgrund der Optimalität der beteiligten Tests im Bahadur Sinn mit Korollar 4.3.5 für die lokale Bahadur-Effizienz der Tests

$$e^{\mathrm{LB}}_{T_1',T_1}\big(\vartheta^0, \vartheta_1\big) = \frac{\vartheta_1^t \big(i^{1,1}_{N_1'}(\vartheta_0') - i^{1,2}_{N_1'}(\vartheta_0') i^{2,2}_{N_1'}(\vartheta_0')^{-1} i^{2,1}_{N_1'}(\vartheta_0')\big)\vartheta_1}{\vartheta_1^t \big(i^{1,1}_{N_1}(\vartheta_0) - i^{1,2}_{N_1}(\vartheta_0) i^{2,2}_{N_1}(\vartheta_0)^{-1} i^{2,1}_{N_1}(\vartheta_0)\big)\vartheta_1},$$

wobei hier Bemerkung 4.2.8 berücksichtigt wurde. Dabei ist

$$i_{N_1}(\vartheta_0) = \begin{pmatrix} i^{1,1}_{N_1}(\vartheta_0) & i^{1,2}_{N_1}(\vartheta_0) \\ i^{2,1}_{N_1}(\vartheta_0) & i^{2,2}_{N_1}(\vartheta_0) \end{pmatrix}$$

die Informationsmatrix aus der Verteilung von N_1, mit $i^{1,1}_{N_1}(\vartheta_0) \in \mathbb{R}^{(d-1)\times(d-1)}$, $i^{1,2}_{N_1}(\vartheta_0) \in \mathbb{R}^{(d-1)\times(d^2-d)}$, $i^{2,1}_{N_1}(\vartheta_0) \in \mathbb{R}^{(d^2-d)\times(d-1)}$, $i^{2,2}_{N_1}(\vartheta_0) \in \mathbb{R}^{(d^2-d)\times(d^2-d)}$, und

$$i_{N_1'}(\vartheta_0') = \begin{pmatrix} i^{1,1}_{N_1'}(\vartheta_0') & i^{1,2}_{N_1'}(\vartheta_0') \\ i^{2,1}_{N_1'}(\vartheta_0') & i^{2,2}_{N_1'}(\vartheta_0') \end{pmatrix}$$

die Informationsmatrix aus der Verteilung von N_1', mit $i_{N_1'}^{1,1}(\vartheta_0') \in \mathbb{R}^{(d-1)\times(d-1)}$, $i_{N_1'}^{1,2}(\vartheta_0') \in \mathbb{R}^{(d-1)\times(d-1)}$, $i_{N_1'}^{2,1}(\vartheta_0') \in \mathbb{R}^{(d-1)\times(d-1)}$, $i_{N_1'}^{2,2}(\vartheta_0') \in \mathbb{R}^{(d-1)\times(d-1)}$. Mit Satz 6.3.10 ergibt sich

$$A_{N_1}^{1,1}(\vartheta_0) = i_{N_1}^{1,1}(\vartheta_0) - i_{N_1}^{1,2}(\vartheta_0) i_{N_1}^{2,2}(\vartheta_0)^{-1} i_{N_1}^{2,1}(\vartheta_0),$$

sowie

$$A_{N_1'}^{1,1}(\vartheta_0') = i_{N_1'}^{1,1}(\vartheta_0') - i_{N_1'}^{1,2}(\vartheta_0') i_{N_1'}^{2,2}(\vartheta_0')^{-1} i_{N_1'}^{2,1}(\vartheta_0').$$

Damit ist alles gezeigt. □

9.5.2 Bemerkung: Es ist

$$\mathrm{Cov}_p(N_1) = \mathrm{diag}(p^t) - pp^t,\ p \in \Delta_{d^2},$$

sowie

$$\mathrm{Cov}_p(N_1)^{-1} = \mathrm{diag}\left(\frac{1}{p_{1,1}}, \dots, \frac{1}{p_{d-1,d}}\right) - \frac{1}{p_{d,d}} \mathbb{1}_{(d^2-1)\times(d^2-1)},\ p \in \Delta_{d^2},$$

mit $\mathbb{1}_{(d^2-1)\times(d^2-1)}$ als $(d^2-1)\times(d^2-1)$-Matrix, bei der alle Einträge 1 sind. Weiter ist

$$\mathrm{Cov}_{(p_{*\cdot}, p_{\cdot *})}(N_1') = \begin{pmatrix} \mathrm{diag}(p_{*\cdot}^t) - p_{*\cdot} p_{*\cdot}^t & 0_{(d^2-1)\times(d^2-1)} \\ 0_{(d^2-1)\times(d^2-1)} & \mathrm{diag}(p_{\cdot *}^t) - p_{\cdot *} p_{\cdot *}^t \end{pmatrix},\ (p_{*\cdot}, p_{\cdot *}) \in \Delta_d^2,$$

sowie

$$\begin{aligned} &\mathrm{Cov}_{(p_{*\cdot}, p_{\cdot *})}(N_1')^{-1} \\ = &\begin{pmatrix} \mathrm{diag}\left(\frac{1}{p_{1\cdot}}, \dots, \frac{1}{p_{d-1\cdot}}\right) - \frac{1}{p_{d\cdot}} \mathbb{1}_{(d^2-1)\times(d^2-1)} & 0_{(d^2-1)\times(d^2-1)} \\ 0_{(d^2-1)\times(d^2-1)} & \mathrm{diag}\left(\frac{1}{p_{\cdot 1}}, \dots, \frac{1}{p_{\cdot d-1}}\right) - \frac{1}{p_{\cdot d}} \mathbb{1}_{(d^2-1)\times(d^2-1)} \end{pmatrix}, \\ &(p_{*\cdot}, p_{\cdot *}) \in \Delta_d^2. \end{aligned}$$

9.5.3 Beispiel: Wir betrachten den Spezialfall $d = 2$. Hier ist

$$A_{N_1}^{1,1}(\vartheta_0) = \frac{1}{(1, 0_{1\times 2}) \psi_{N_1}^{-1} \mathrm{Cov}_{\vartheta_0}(N_1) (\psi_{N_1'}^t)^{-1} (1, 0_{1\times 2})^t}$$

sowie

$$A_{N_1'}^{1,1}(\vartheta_0') = \frac{1}{(1,0) \psi_{N_1'}^{-1} \mathrm{Cov}_{\vartheta_0'}(N_1') (\psi_{N_1'}^t)^{-1} (1,0)^t}.$$

Insgesamt ergibt sich

$$e_{T_1', T_1}^{\mathrm{LB}}(\vartheta^0) = \frac{(1, 0_{1\times 2}) \psi_{N_1}^{-1} \mathrm{Cov}_{\vartheta_0}(N_1) (\psi_{N_1'}^t)^{-1} (1, 0_{1\times 2})^t}{(1,0) \psi_{N_1'}^{-1} \mathrm{Cov}_{\vartheta_0'}(N_1') (\psi_{N_1'}^t)^{-1} (1,0)^t},\ \vartheta^0 \in \Theta_0.$$

Konkret ist

$$\psi_{N_1} = \begin{pmatrix} 0 & 0 & 1 \\ -1/2 & 1/2 & -1 \\ 1/2 & 1/2 & -1 \end{pmatrix}, \ \psi_{N_1}^{-1} = \begin{pmatrix} 0 & -1 & 1 \\ 2 & 1 & 1 \\ 1 & 0 & 0 \end{pmatrix},$$

und mit $p_{1,2} = p_{2,1}$ unter H

$$\begin{aligned} \mathrm{Cov}_{\vartheta_0}(N_1) &= \begin{pmatrix} p_{1,1} & 0 & 0 \\ 0 & p_{2,1} & 0 \\ 0 & 0 & p_{1,2} \end{pmatrix} - \begin{pmatrix} p_{1,1}^2 & p_{1,1}p_{2,1} & p_{1,1}p_{1,2} \\ p_{1,1}p_{2,1} & p_{2,1}^2 & p_{1,2}p_{2,1} \\ p_{1,1}p_{1,2} & p_{1,2}p_{2,1} & p_{1,2}^2 \end{pmatrix} \\ &= \begin{pmatrix} p_{1,1} & 0 & 0 \\ 0 & p_{1,2} & 0 \\ 0 & 0 & p_{1,2} \end{pmatrix} - \begin{pmatrix} p_{1,1}^2 & p_{1,1}p_{1,2} & p_{1,1}p_{1,2} \\ p_{1,1}p_{1,2} & p_{1,2}^2 & p_{1,2}^2 \\ p_{1,1}p_{1,2} & p_{1,2}^2 & p_{1,2}^2 \end{pmatrix}, \ \vartheta_0 \in \tilde{\Theta}_0, \end{aligned}$$

wobei $p = \phi(\vartheta)$. Weiter ist

$$\psi_{N_1'} = \begin{pmatrix} 1/2 & 1/2 \\ -1/2 & 1/2 \end{pmatrix}, \ \psi_{N_1'}^{-1} = \begin{pmatrix} 1 & -1 \\ 1 & 1 \end{pmatrix},$$

und mit $p_{1\cdot} = p_{\cdot 1}$ unter H

$$\mathrm{Cov}_{\vartheta_0'}(N_1') = \begin{pmatrix} p_{1\cdot} - p_{1\cdot}^2 & 0 \\ 0 & p_{\cdot 1} - p_{\cdot 1}^2 \end{pmatrix} = \begin{pmatrix} p_{1\cdot}p_{2\cdot} & 0 \\ 0 & p_{1\cdot}p_{2\cdot} \end{pmatrix}, \ \vartheta_0' \in \tilde{\Theta}_0'.$$

Somit ergibt sich

$$e_{T_1',T_1}^{\mathrm{LB}}(\vartheta^0) = \frac{2p_{1,2}}{2p_{1\cdot}p_{2\cdot}} = \frac{p_{1,2}}{p_{1\cdot}p_{2\cdot}} = \frac{p_{1,2} + p_{2,1}}{p_{1\cdot}p_{2\cdot} + p_{\cdot 1}p_{\cdot 2}}, \ \vartheta^0 \in \Theta_0.$$

Wegen

$$\begin{aligned} p_{1,2} - p_{1\cdot}p_{2\cdot} &= \mathrm{Cov}_{\vartheta_0}\big(\mathrm{I}(X_1 = 1), \mathrm{I}(Y_1 = 2)\big) \\ &= -\mathrm{Cov}_{\vartheta_0}\big(\mathrm{I}(X_1 = 1), \mathrm{I}(Y_1 = 1)\big) = -\mathrm{Cov}_{\vartheta_0}(X_1, Y_1), \ \vartheta_0 \in \tilde{\Theta}_0, \end{aligned}$$

ist

$$e_{T_1',T_1}^{\mathrm{LB}}(\vartheta^0) \begin{Bmatrix} < \\ = \\ > \end{Bmatrix} 1 \ \Leftrightarrow \ \mathrm{Cov}_{\vartheta_0}(X_1, Y_1) \begin{Bmatrix} > \\ = \\ < \end{Bmatrix} 0, \ \vartheta^0 \in \Theta_0.$$

9.6 Hodges-Lehmann-Effizienz der Tests

Die auf den Testgrößen (9.3.1) und (9.4.6) basierenden vorgeschlagenen Likelihood-Qutienten-Tests sind jeweils optimal im Hodges-Lehmann Sinn. Dies folgt aus Theorem 3 in der Arbeit [35]. Um den zu dem jeweiligen Test zugehörigen Hodges-Lehmann-Index zu bestimmen, muss zunächst die jeweilige Kullback-Leibler-Information entsprechend minimiert werden.

Betrachte dazu im unabhängigen Stichprobenfall für $(q_{*\cdot}, q_{\cdot *}) := ((q_{1\cdot}, \dots, q_{d-1\cdot})^t, (q_{\cdot 1}, \dots, q_{\cdot d-1})^t) \in \Delta_d^2$ das konvexe Optimierungsproblem der Minimierung der Abbildung

$$(p_{*\cdot}, p_{\cdot *}) \longmapsto \sum_{i=1}^{d} p_{i\cdot} \log p_{i\cdot} + \sum_{i=1}^{d} p_{\cdot i} \log p_{\cdot i} - \sum_{i=1}^{d} p_{i\cdot} \log q_{i\cdot} - \sum_{i=1}^{d} p_{\cdot i} \log q_{\cdot i} \tag{9.6.1}$$

unter den Nebenbedingungen

$$p_{i\cdot} = p_{\cdot i},\ p_{i\cdot} \geq 0,\ p_{\cdot i} \geq 0,\ i \in \{1, \dots, d\},\ \sum_{i=1}^{d} p_{i\cdot} = 1,\ \sum_{i=1}^{d} p_{\cdot i} = 1. \tag{9.6.2}$$

Betrachten wir das Gleichungssystem

$$\begin{aligned}
0 &= \log p_{i\cdot} - \log q_{i\cdot} + \lambda_i + \mu_1 + 1, \\
0 &= \log p_{\cdot i} - \log q_{\cdot i} - \lambda_i + \mu_2 + 1, \\
0 &= p_{i\cdot} - p_{\cdot i}, \\
0 &= \sum_{i=1}^{d} p_{i\cdot} - 1 \\
0 &= \sum_{i=1}^{d} p_{\cdot i} - 1,\ i \in \{1, \dots, d\},\ ((p_{1\cdot}, \dots, p_{d\cdot})^t, (p_{\cdot 1}, \dots, p_{\cdot d})^t) \in \mathbb{R}^d \times \mathbb{R}^d,\ \lambda \in \mathbb{R}^d,\ \mu \in \mathbb{R}^2.
\end{aligned} \tag{9.6.3}$$

Nach Theorem 4.3.7 aus [6] ist das Lösen von (9.4.5) ein hinreichendes Kriterium für das globale Minimum der Funktion (9.6.1) unter den Nebenbedingungen (9.6.2). Die Lösung $(\hat{\tilde{p}}'_{*\cdot}, \hat{\tilde{p}}'_{\cdot *})$ von (9.6.3) ergibt sich aus

$$\hat{\tilde{p}}'_{i\cdot}(q_{*\cdot}, q_{\cdot *}) = \hat{\tilde{p}}'_{\cdot i}(q_{*\cdot}, q_{\cdot *}) = \frac{\sqrt{q_{i\cdot} q_{\cdot i}}}{\sum_{j=1}^{d} \sqrt{q_{j\cdot} q_{\cdot j}}},\ i \in \{1, \dots, d\}.$$

Diese Lösung ist als Funktion von $(q_{*\cdot}, q_{\cdot *}) \in \Delta_d^2$ stetig. Da die Funktion (9.6.1) stetig in $(\Delta_d^0)^2$ ist, handelt es sich hierbei auch um die globale Minimalstelle der Funktion (9.6.1) unter den Nebenbedingungen (9.6.2), die also insbesondere in Δ_d^2 liegt, also $(\hat{\tilde{p}}'_{*\cdot}, \hat{\tilde{p}}'_{\cdot *}) \in \Delta_d^2$ gilt.

Betrachte im verbundenen Stichprobenfall für $q \in \Delta_{d^2}$ das konvexe Optimierungsproblem der Minimierung der Abbildung

$$p \longmapsto \sum_{i,j=1}^{d} p_{i,j} \log p_{i,j} - \sum_{i,j=1}^{d} p_{i,j} \log q_{i,j} \tag{9.6.4}$$

unter den Nebenbedingungen

$$p_{i\cdot} = p_{\cdot i},\ p_{i,j} \geq 0,\ (i,j) \in \{1,\ldots,d\}^2,\ \sum_{i,j=1}^{d} p_{i,j} = 1. \tag{9.6.5}$$

Betrachten wir das Gleichungssystem

$$\begin{aligned} 0 &= \log p_{i,j} - \log q_{i,j} + \lambda_i - \lambda_j + \mu + 1, \\ 0 &= p_{i\cdot} - p_{\cdot i}, \\ 0 &= \sum_{i,j=1}^{d} p_{i,j} - 1,\ (i,j) \in \{1,\ldots,d\}^2,\ (p_{1,1},\ldots,p_{d,d})^t \in \mathbb{R}^{d^2},\ \lambda \in \mathbb{R}^d,\ \mu \in \mathbb{R}. \end{aligned} \tag{9.6.6}$$

Nach Theorem 4.3.7 aus [6] ist das Lösen von (9.6.6) ein hinreichendes Kriterium für das globale Minimum der Funktion (9.6.4) unter den Nebenbedingungen (9.6.5). Nach Lemma 4.8 aus der Arbeit [15] von Hoeffding existiert eine eindeutige globale Minimalstelle $\hat{\hat{p}}$ der Funktion (9.6.4) unter den Nebenbedingungen (9.6.5). Diese Minimalstelle liegt in Δ_{d^2}, also $\hat{\hat{p}} \in \Delta_{d^2}$; betrachte dazu den Beweis von Lemma 4.8 aus [15]. Die globale Minimalstelle $\hat{\hat{p}}$ ist auch eine lokale Minimalstelle und daher nach Theorem 4.3.6 aus [6] Lösung des Gleichungssystems (9.6.6). Die Lösung $\hat{\hat{p}}$ von (9.6.6) lässt sich für $d > 2$ nicht explizit angeben.

Somit ergibt sich für den Hodges-Lehmann-Index im unabhängigen Stichprobenfall bei zugrunde liegender Alternative $(p_{*\cdot}, p_{\cdot *}) \in \Delta_d^2$

$$\sum_{i=1}^{d} \frac{\sqrt{p_{i\cdot}p_{\cdot i}}}{\sum_{j=1}^{d}\sqrt{p_{j\cdot}p_{\cdot j}}} \log \frac{\frac{\sqrt{p_{i\cdot}p_{\cdot i}}}{\sum_{j=1}^{d}\sqrt{p_{j\cdot}p_{\cdot j}}}}{p_{i\cdot}} + \sum_{i=1}^{d} \frac{\sqrt{p_{i\cdot}p_{\cdot i}}}{\sum_{j=1}^{d}\sqrt{p_{j\cdot}p_{\cdot j}}} \log \frac{\frac{\sqrt{p_{i\cdot}p_{\cdot i}}}{\sum_{j=1}^{d}\sqrt{p_{j\cdot}p_{\cdot j}}}}{p_{\cdot i}},$$

und für den Hodges-Lehmann-Index im verbundenen Stichprobenfall bei zugrunde liegender Alternative $p \in \Delta_{d^2}$

$$\sum_{i,j=1}^{d^2} \hat{\hat{p}}(p)_{i,j} \log \frac{\hat{\hat{p}}(p)_{i,j}}{p_{i,j}},$$

wobei $\hat{\hat{p}} = (\hat{\hat{p}}_{1,1},\ldots,\hat{\hat{p}}_{d-1,d})^t$ Lösung von (9.6.6) ist. Somit kann auch die Hodges-Lehmann-Effizienz der Likelihood-Quotienten-Tests basierend auf den Testgrößen (9.3.1) und (9.4.6) für $d > 2$ nicht explizit angegeben werden. Allerdings können hier die in Unterabschnitt 4.3 hergeleiteten allgemeinen Resultate zur lokalen Hodges-Lehmann-Effizienz von optimalen Tests im Hodges-Lehmann Sinn angewendet werden. Mit den Notationen aus Unterabschnitt 9.5 können wir dann folgendes Resultat zur lokalen Hodges-Lehmann-Effizienz formulieren.

9.6.1 Korollar: *Es sei eine reellwertige Folge* $(\nu_i)_{i\in\mathbb{N}}$ *mit* $\nu_i \neq 0$ *für alle* $i \in \mathbb{N}$ *und* $\lim_{i\to\infty} \nu_i = 0$ *gegeben. Für* $i \in \mathbb{N}$ *betrachten wir Alternativen der Form*

$$\vartheta^i := \begin{pmatrix} 0 \\ \vartheta_2 \\ \vartheta_2' \end{pmatrix} + \nu_i \begin{pmatrix} \vartheta_1 \\ 0 \\ 0 \end{pmatrix}$$

mit $\vartheta_1 \in \mathbb{R}^{d-1} \setminus \{0\}$, $(0, \vartheta_2, \vartheta_2') \in \Theta_0$, *die* $(\nu_i\vartheta_1, \vartheta_2, \vartheta_2') \in \Theta_1$ *erfüllen. Dann gilt mit* $\vartheta_0 := \left(\begin{smallmatrix} 0 \\ \vartheta_2 \end{smallmatrix}\right)$, $\vartheta_0' := \left(\begin{smallmatrix} 0 \\ \vartheta_2' \end{smallmatrix}\right)$ *und* $\vartheta^0 := (0, \vartheta_2, \vartheta_2')$ *für die Folgen von Tests* $T_1 = (T_{1,n})_{n\in\mathbb{N}}$ *und* $T_1' = (T_{1,n}')_{n\in\mathbb{N}}$

$$e^{\mathrm{LHL}}_{T_1',T_1}\left(\vartheta^0, \vartheta_1\right) = \frac{\vartheta_1^t A^{1,1}_{N_1'}(\vartheta_0')\vartheta_1}{\vartheta_1^t A^{1,1}_{N_1}(\vartheta_0)\vartheta_1}.$$

Beweis: Im verbundenen Stichprobenfall sind in der vorliegenden Situation die Bedingungen (C1),(C2),(C3),(C5),(C6) und (C7) aus Unterabschnitt 4.3 erfüllt. Benutze zum Beweis von (C7) Lemma 4.3.6 und argumentiere analog zum Beweis von Korollar 9.5.1.

Im unabhängigen Stichprobenfall sind in der vorliegenden Situation die Bedingungen (C1),(C2),(C3),(C5),(C6) und (C7) aus Unterabschnitt 4.3 erfüllt.

Somit ergibt sich ingesamt analog zum Beweis von Korollar 9.5.1 durch Anwendung von Korollar 4.3.10 die Behauptung. □

9.6.2 Beispiel: Für das Beispiel $d = 2$ betrachte die Ausführungen in Beispiel 9.5.3. Es gilt

$$e^{\mathrm{LHL}}_{T_1',T_1}\left(\vartheta^0\right) = \frac{p_{1,2} + p_{2,1}}{p_{1\cdot}p_{2\cdot} + p_{\cdot 1}p_{\cdot 2}}, \; \vartheta^0 \in \Theta_0,$$

sowie

$$e^{\mathrm{LHL}}_{T_1',T_1}\left(\vartheta^0\right) \left\{\begin{matrix} < \\ = \\ > \end{matrix}\right\} 1 \;\Leftrightarrow\; \mathrm{Cov}_{\vartheta_0}(X_1, Y_1) \left\{\begin{matrix} > \\ = \\ < \end{matrix}\right\} 0, \; \vartheta^0 \in \Theta_0.$$

9.7 Pitman-Effizienz der Tests

Anhand der bereits angestellten Überlegungen lässt sich ein Resultat zur Pitman-Effizienz der betrachteten Tests begründen. Mit den Notationen aus Unterabschnitt 9.5 können wir dann folgendes Resultat zur Pitman-Effizienz formulieren.

9.7.1 Korollar: *Es sei eine reellwertige Folge* $(\nu_i)_{i\in\mathbb{N}}$ *mit* $\nu_i \neq 0$ *für alle* $i \in \mathbb{N}$ *und* $\lim_{i\to\infty} \nu_i = 0$ *gegeben. Für* $i \in \mathbb{N}$ *betrachten wir Alternativen der Form*

$$\vartheta^i := \begin{pmatrix} 0 \\ \vartheta_2 \\ \vartheta_2' \end{pmatrix} + \nu_i \begin{pmatrix} \vartheta_1 \\ 0 \\ 0 \end{pmatrix}$$

mit $\vartheta_1 \in \mathbb{R}^{d-1} \setminus \{0\}$, $(0, \vartheta_2, \vartheta_2') \in \Theta_0$, *die* $(\nu_i \vartheta_1, \vartheta_2, \vartheta_2') \in \Theta_1$ *erfüllen. Zudem seien die Abbildungen*

$$i \mapsto N_{T,i},\ i \mapsto N_{T',i},\ i \in \mathbb{N},$$

streng monoton wachsend. Dann gilt mit $\vartheta_0 := \binom{0}{\vartheta_2}$, $\vartheta_0' := \binom{0}{\vartheta_2'}$ *und* $\vartheta^0 := (0, \vartheta_2, \vartheta_2')$ *für die Folgen von Tests* $T_i = (T_{i,n})_{n\in\mathbb{N}}$ *und* $T_j' = (T_{j,n}')_{n\in\mathbb{N}}$, $i, j = 1, 2$, *im Fall der Existenz der Pitman-Effizienz*

$$e^{\mathrm{P}}_{T_j',T_i}\big(\vartheta^0, \vartheta_1\big) = \frac{\vartheta_1^t A^{1,1}_{N_1'}(\vartheta_0')\vartheta_1}{\vartheta_1^t A^{1,1}_{N_1}(\vartheta_0)\vartheta_1},\ i, j = 1, 2.$$

Beweis: Im verbundenen Stichprobenfall sind in der vorliegenden Situation die Bedingungen (B1) - (B4) aus Unterabschnitt 4.2 erfüllt.

Im verbundenen Stichprobenfall ist auch die Bedingung (B5) erfüllt, wie man wie folgt sieht: Für alle p_0 aus der Hypothesenmenge existiert eine offene Umgebung $U_{p_0} \subset \Delta_{d^2}$ um p_0 sodass $\hat{p}$ auf U_{p_0} stetig ist. Dies sieht man wie folgt. Es kann die Dichte von N_1 bezüglich des Zählmaßes μ auf $\mathbb{N}_0^{d^2-1}$ bei zugrunde liegendem Parameter $p \in \Delta_{d^2}$, die mit $\tilde{f}(\cdot, p)$ bezeichnet werden soll, zu einer Dichte von N_1 bezüglich des Zählmaßes μ auf $\mathbb{N}_0^{d^2-1}$ bei zugrunde liegendem Parameter $\phi^{-1}(p) \in \tilde{\Theta}$, die mit $f(\cdot, \phi^{-1}(p))$ bezeichnet werden soll, umgeschrieben werden, d.h. $\tilde{f}(\cdot, p) = f(\cdot, \phi^{-1}(p))$. Es bezeichne $\tilde{K}(\vartheta, \tilde{\vartheta})$ die Kullback-Leibler Information aus der Verteilung von N_1 bei zugrunde liegenden Parametern $\vartheta, \tilde{\vartheta} \in \tilde{\Theta}$, d.h. wir definieren

$$\tilde{K}(\vartheta, \tilde{\vartheta}) := \int \log \frac{f(z, \vartheta)}{f(z, \tilde{\vartheta})} f(z, \vartheta) \mathrm{d}\mu(z),\ \vartheta, \tilde{\vartheta} \in \tilde{\Theta}.$$

Wegen $\tilde{f}(\cdot, p) = f(\cdot, \phi^{-1}(p))$, $p \in \Delta_{d^2}$, gilt $K(p, \tilde{p}) := \tilde{K}(\phi^{-1}(p), \phi^{-1}(\tilde{p}))$. Es sei $U_{\phi^{-1}(p_0)}$ eine offene Umgebung um $\phi^{-1}(p_0)$ und $\hat{\eta} : U_{\phi^{-1}(p_0)} \to \Delta$ die eindeutige und in $U_{\phi^{-1}(p_0)}$ stetige Abbildung $\hat{\eta} : U_{\phi^{-1}(p_0)} \to \Delta$, die für $\vartheta \in U_{\phi^{-1}(p_0)}$ die Minimalstelle der Abbildung

$$\eta \longmapsto K\big(\vartheta, g(\eta)\big),\ \eta \in \overline{\Delta},$$

liefert, und deren Existenz im Beweis von Korollar 9.5.1 begründet wurde. Mit $U_{p_0} := \phi(U_{\phi^{-1}(p_0)})$ folgt

$$\tilde{K}\bigg(\phi^{-1}(p), g\Big(\hat{\eta}\big(\phi^{-1}(p)\big)\Big)\bigg) \le \tilde{K}\Big(\phi^{-1}(p), \phi^{-1}\big(\hat{p}(p)\big)\Big),\ p \in U_{p_0}.$$

Zudem gilt mit den Überlegungen aus Unterabschnitt 9.2

$$\begin{aligned}\tilde{K}\bigg(\phi^{-1}(p), g\Big(\hat{\eta}\big(\phi^{-1}(p)\big)\Big)\bigg) &= K\Bigg(p, \phi\bigg(g\Big(\hat{\eta}\big(\phi^{-1}(p)\big)\Big)\bigg)\Bigg) \ge K\big(p, \hat{p}(p)\big)\\ &= \tilde{K}\Big(\phi^{-1}(p), \phi^{-1}\big(\hat{p}(p)\big)\Big),\ p \in U_{p_0}.\end{aligned}$$

Also ist

$$\tilde{K}\Big(\phi^{-1}(p), g\Big(\hat{\eta}\big(\phi^{-1}(p)\big)\Big)\Big) = \tilde{K}\Big(\phi^{-1}(p), \phi^{-1}\big(\hat{p}(p)\big)\Big),\ p \in U_{p_0}.$$

Hieraus folgt wegen der Eindeutigkeit von $\hat{\eta}$ sowie der Injektivität von g auf Δ

$$g\Big(\hat{\eta}\big(\phi^{-1}(p)\big)\Big) = \phi^{-1}\big(\hat{p}(p)\big),\ p \in U_{p_0},$$

äquivalent

$$\hat{p}(p) = \phi\bigg(g\Big(\hat{\eta}\big(\phi^{-1}(p)\big)\Big)\bigg),\ p \in U_{p_0}.$$

Wegen der Stetigkeit von $\phi \circ g \circ \hat{\eta} \circ \phi^{-1}$ auf U_{p_0} folgt die Stetigkeit von $\hat{p}$ auf U_{p_0}.

Nun sei $(p_n)_{n\in\mathbb{N}}$ eine Folge mit $p_n \in \Delta_{d^2}$ für alle $n \in \mathbb{N}$ und $\lim_{n\to\infty} p_n = p_0$. Mit einem schwachen Gesetz der großen Zahlen für Dreiecksschemata, siehe Lemma A.1.2 im Anhang, folgt für alle $\varepsilon > 0$ $\lim_{n\to\infty} P_{p_n}(|N/n - p_0| \geq \varepsilon) = 0$. Daraus folgt (B5).

Im unabhängigen Stichprobenfall sind in der vorliegenden Situation die Bedingungen (B1)-(B5) aus Unterabschnitt 4.2 erfüllt.

Somit ergibt sich ingesamt analog zum Beweis von Korollar 9.5.1 durch Anwendung von Korollar 4.2.5 die Behauptung. □

9.7.2 Bemerkung: Es lässt sich unter Beachtung von Bemerkung 2.4.1 mit den Überlegungen präsentiert in 3.8.3 in [32] zeigen, dass die Pitman-Effizienz für $\nu_i = 1/\sqrt{i}$, $i \in \mathbb{N}$, existiert.

9.7.3 Beispiel: Für das Beispiel $d = 2$ betrachte die Ausführungen in Beispiel 9.5.3. Es gilt

$$e^{\mathrm{P}}_{T_j',T_i,n}(\vartheta^0) = \frac{p_{1,2} + p_{2,1}}{p_{1\cdot}p_{2\cdot} + p_{\cdot 1}p_{\cdot 2}},\quad \vartheta^0 \in \Theta_0,\ i,j = 1,2,$$

sowie

$$e^{\mathrm{P}}_{T_1',T_1}(\vartheta^0) \begin{Bmatrix} < \\ = \\ > \end{Bmatrix} 1 \ \Leftrightarrow\ \mathrm{Cov}_{\vartheta_0}(X_1, Y_1) \begin{Bmatrix} > \\ = \\ < \end{Bmatrix} 0,\ \vartheta^0 \in \Theta_0,\ i,j = 1,2.$$

9.8 Volumen-Effizienz der Tests

Das in Unterabschnitt 3 vorgestellte Konzept der Volumen-Effizienz kann in offensichtlicher Weise auf die vorliegende Situation der zusammengesetzten Hypothese erweitert werden. Siehe hierzu die Erklärung am Anfang des Unterabschnitts 4.4.

Anhand der bereits angestellten Überlegungen lässt sich ein Resultat zur Volumen-Effizienz der genannten Tests begründen. Mit den Notationen aus Unterabschnitt 9.5 können wir folgendes Resultat zur Volumen-Effizienz formulieren.

9.8.1 Satz: *Es sei* $\tilde{\vartheta} = (\vartheta_1, \vartheta_2, \vartheta_2')^t \in \Theta$, $\vartheta = (\vartheta_1, \vartheta_2)^t \in \tilde{\Theta}$ *und* $\vartheta' = (\vartheta_1, \vartheta_2')^t \in \tilde{\Theta}'$. *Dann hat man für die Folgen von Tests* $T_i = (T_{i,n})_{n\in\mathbb{N}}$ *und* $T_j' = (T_{j,n}')_{n\in\mathbb{N}}$, $i, j = 1, 2$,

$$\lim_{n\to\infty} e^{\mathrm{vol}}_{T_j', T_i, n}(\tilde{\vartheta}) = \frac{\det A^{1,1}_{N_1'}(\vartheta')}{\det A^{1,1}_{N_1}(\vartheta)}, \ \ i, j = 1, 2.$$

Beweis: Im verbundenen Stichprobenfall sind in der vorliegenden Situation die Bedingungen (D1) - (D4) aus Unterabschnitt 4.4 erfüllt.

Im verbundenen Stichprobenfall ist auch die Bedingung (D5) erfüllt, wie man wie folgt sieht: Es sei $\tilde{p} \in \Delta_{d^2}$ und $\vartheta_1^* := (z_1^*, \ldots, z_{d-1}^*)^t \in \Theta^1$. Analog zu den Ausführungen in Unterabschnitt 9.3 kann begründet werden, dass für $N/n \in \Delta_{d^2}$ $\hat{p}(N/n, \vartheta_1^*)$ Maximum-Likelihood-Schätzer für $\tilde{p}$ unter der Hypothese $\tilde{\vartheta}_1 = \vartheta_1^*$ ist, wobei $\tilde{\vartheta}_1 \in \Theta^1$ die zu $\phi^{-1}(\tilde{p})$ zugehörige Parameterkomponente sei, und $\hat{p} : \Delta_{d^2} \times \Theta^1 \to \Delta_{d^2}$ Lösung von

$$\begin{aligned}
0 &= -\frac{q_{i,j}}{p_{i,j}} + \lambda_i - \lambda_j + \mu, \\
0 &= p_{k\cdot} - p_{\cdot k} - z_k^*, \\
0 &= \sum_{i,j=1}^{d} p_{i,j} - 1, \ (i,j) \in \{1, \ldots, d\}^2, \\
& k \in \{1, \ldots, d-1\}, \ (p_{1,1}, \ldots, p_{d,d})^t \in \mathbb{R}^{d^2}, \ \lambda \in \mathbb{R}^d, \ \mu \in \mathbb{R},
\end{aligned}$$

sei. Es kann analog zu den Beweisen von Korollar 9.5.1 und Korollar 9.7.1 argumentiert werden, um die Gültigkeit von (D5) zu begründen.

Im unabhängigen Stichprobenfall sind in der vorliegenden Situation die Bedingungen (D1)-(D5) aus Unterabschnitt 4.4 erfüllt.

Um Korollar 4.4.7 hier anwenden zu können, werden wir die Bedingungen aus Lemma 4.4.5 und Lemma 4.4.6 nachweisen. Dabei beschränken wir uns auf die schwierigere Situation im verbundenen Stichprobenfall. Im unabhängigen Stichprobenfall können dann analoge Argumente angebracht werden.

Befassen wir uns zunächst mit den Bedingungen in Lemma 4.4.5. Es sei dazu $\alpha \in (0,1)$ und $p \in \Delta_{d^2}$ mit $\vartheta = \phi^{-1}(p)$. Es sei $k, l \in \{1, \ldots, d\}$ mit der Eigenschaft

$$p_{k,l} = \min_{1\le i,j\le n} p_{i,j}.$$

Es ist $p_{k,l} > 0$. Wir definieren

$$\varepsilon := \frac{p_{k,l}}{2} + \frac{\chi^2_{d_1;1-\alpha}}{8} - \sqrt{\left(\frac{p_{k,l}}{2} + \frac{\chi^2_{d_1;1-\alpha}}{8}\right)^2 - \frac{p_{k,l}^2}{4}}.$$

Dann ist $\varepsilon \in (0, p_{k,l}/2)$. Ferner setzen wir

$$\tilde{K}_1 := \Big\{ (\tilde{p}_{1,1}, \ldots, \tilde{p}_{d-1,d})^t \in \Delta_{d^2}; \tilde{p}_{i,j} \geq \varepsilon, i, j \in \{1, \ldots, d\} \Big\}.$$

Daraus folgt $\tilde{K}_1 \neq \emptyset$. Zudem wird $\tilde{K}_2 := \Delta_{d^2} \setminus \tilde{K}_1$, $K_1 := \phi^{-1}(\tilde{K}_1)$ und $K_2 := \phi^{-1}(\tilde{K}_2)$ gesetzt. Beachte nun die Notationen aus Unterabschnitt 4.4. Es ist $K_1, K_2 \in \mathfrak{B}^{d^2-1}_{\tilde{\Theta}}$, $K_1 \cap K_2 = \emptyset$ sowie $K_1 \cup K_2 = \tilde{\Theta}$. Weiter ist K_1 abgeschlossen und für alle $\tilde{\tilde{\vartheta}} \in K_1$ ist $i_{N_1}(\tilde{\tilde{\vartheta}}) = \psi^t_{N_1} \mathrm{Cov}_{\tilde{\tilde{\vartheta}}}(N_1)^{-1} \psi_{N_1}$ positiv definit. Daher existiert eine reelle Zahl $\lambda_{\min} > 0$ mit

$$\inf_{\tilde{\tilde{\vartheta}} \in K_1} \lambda_{\tilde{\tilde{\vartheta}}} > \lambda_{\min},$$

mit $\lambda_{\tilde{\tilde{\vartheta}}}$ als kleinster Eigenwert der Matrix $i_{N_1}(\tilde{\tilde{\vartheta}})$, $\tilde{\tilde{\vartheta}} \in K_1$. Es sei $\vartheta_1^* \in \Theta^1$. Dann ist

$$\begin{aligned}
&Q_\vartheta\big(T_n(\vartheta_1 + \vartheta_1^*/\sqrt{n}) < \chi^2_{d_1;1-\alpha}, (\vartheta_1 + \vartheta_1^*/\sqrt{n}) \in \Theta^1, \hat{\tilde{\vartheta}}_n^{\vartheta_1 + \vartheta_1^*/\sqrt{n}} \in K_2\big) \\
=&Q_\vartheta\Bigg(T_n(\vartheta_1 + \vartheta_1^*/\sqrt{n}) < \chi^2_{d_1;1-\alpha}, (\vartheta_1 + \vartheta_1^*/\sqrt{n}) \in \Theta^1, \hat{\tilde{\vartheta}}_n^{\vartheta_1 + \vartheta_1^*/\sqrt{n}} \in K_2, \\
&\forall\, i, j \in \{1, \ldots, d\} : \frac{N_{i,j}}{n} \geq p_{k,l} - \varepsilon\Bigg) \\
&+ Q_\vartheta\Bigg(T_n(\vartheta_1 + \vartheta_1^*/\sqrt{n}) < \chi^2_{d_1;1-\alpha}, (\vartheta_1 + \vartheta_1^*/\sqrt{n}) \in \Theta^1, \hat{\tilde{\vartheta}}_n^{\vartheta_1 + \vartheta_1^*/\sqrt{n}} \in K_2, \\
&\exists\, i, j \in \{1, \ldots, d\} : \frac{N_{i,j}}{n} < p_{k,l} - \varepsilon\Bigg).
\end{aligned}$$

Für jedes $\tilde{p} \in \tilde{K}_2$ existieren $r, s \in \{1, \ldots, d\}$ mit $\tilde{p}_{r,s} < \varepsilon$. Daher existieren $r, s \in \{1, \ldots, d\}$ sodass

$$\begin{aligned}
&Q_\vartheta\Bigg(T_n(\vartheta_1 + \vartheta_1^*/\sqrt{n}) < \chi^2_{d_1;1-\alpha}, (\vartheta_1 + \vartheta_1^*/\sqrt{n}) \in \Theta^1, \hat{\tilde{\vartheta}}_n^{\vartheta_1 + \vartheta_1^*/\sqrt{n}} \in K_2, \\
&\forall\, i, j \in \{1, \ldots, d\} : \frac{N_{i,j}}{n} \geq p_{k,l} - \varepsilon\Bigg) \\
=&Q_\vartheta\Bigg(\sum_{i,j=1}^{d} \frac{\big(N_{i,j} - n\hat{p}(N/n, \vartheta_1 + \vartheta_1^*/\sqrt{n})_{i,j}\big)^2}{n\hat{p}(N/n, \vartheta_1 + \vartheta_1^*/\sqrt{n})_{i,j}} < \chi^2_{d_1;1-\alpha}, \\
&(\vartheta_1 + \vartheta_1^*/\sqrt{n}) \in \Theta^1, \hat{p}(N/n, \vartheta_1 + \vartheta_1^*/\sqrt{n}) \in \tilde{K}_2, \forall\, i, j \in \{1, \ldots, d\} : \frac{N_{i,j}}{n} \geq p_{k,l} - \varepsilon\Bigg) \\
\leq&Q_\vartheta\Bigg(\frac{\big(N_{r,s}/n - \hat{p}(N/n, \vartheta_1 + \vartheta_1^*/\sqrt{n})_{r,s}\big)^2}{\hat{p}(N/n, \vartheta_1 + \vartheta_1^*/\sqrt{n})_{r,s}} < \chi^2_{d_1;1-\alpha}, \\
&(\vartheta_1 + \vartheta_1^*/\sqrt{n}) \in \Theta^1, \hat{p}(N/n, \vartheta_1 + \vartheta_1^*/\sqrt{n}) \in \tilde{K}_2, \frac{N_{r,s}}{n} \geq p_{k,l} - \varepsilon\Bigg)
\end{aligned}$$

$$
\begin{aligned}
=&Q_\vartheta\Bigg(\frac{\big(\overbrace{N_{r,s}/n-p_{k,l}+\varepsilon}^{\geq 0}+\overbrace{p_{k,l}-2\varepsilon}^{>0}+\overbrace{\varepsilon-\hat{p}(N/n,\vartheta_1+\vartheta_1^*/\sqrt{n})_{r,s}}^{>0}\big)^2}{\hat{p}(N/n,\vartheta_1+\vartheta_1^*/\sqrt{n})_{r,s}}<\chi^2_{d_1;1-\alpha},\\
&(\vartheta_1+\vartheta_1^*/\sqrt{n})\in\Theta^1,\hat{p}(N/n,\vartheta_1+\vartheta_1^*/\sqrt{n})\in\tilde{K}_2,\frac{N_{r,s}}{n}\geq p_{k,l}-\varepsilon\Bigg)\\
\leq&\,\mathrm{I}\left(\frac{(p_{k,l}-2\varepsilon)^2}{\varepsilon}<\chi^2_{d_1;1-\alpha}\right)=\mathrm{I}\left(\chi^2_{d_1;1-\alpha}<\chi^2_{d_1;1-\alpha}\right)=0.
\end{aligned}
$$

Weiter ist

$$
\begin{aligned}
&Q_\vartheta\Bigg(T_n(\vartheta_1+\vartheta_1^*/\sqrt{n})<\chi^2_{d_1;1-\alpha},(\vartheta_1+\vartheta_1^*/\sqrt{n})\in\Theta^1,\hat{\hat{\vartheta}}_n^{\vartheta_1+\vartheta_1^*/\sqrt{n}}\in K_2,\\
&\exists\, i,j\in\{1,\dots,d\}:\frac{N_{i,j}}{n}<p_{k,l}-\varepsilon\Bigg)\\
\leq&\,\mathrm{I}\left((\vartheta_1+\vartheta_1^*/\sqrt{n})\in\Theta^1\right)\sum_{i,j=1}^d Q_\vartheta\left(\frac{N_{i,j}}{n}<p_{k,l}-\varepsilon\right),
\end{aligned}
$$

und somit auch

$$
\begin{aligned}
&\int Q_\vartheta\big(T_n(\vartheta_1+\vartheta_1^*/\sqrt{n})<\chi^2_{d_1;1-\alpha},(\vartheta_1+\vartheta_1^*/\sqrt{n})\in\Theta^1,\hat{\hat{\vartheta}}_n^{\vartheta_1+\vartheta_1^*/\sqrt{n}}\in K_2\big)\mathrm{d}\vartheta_1^*\\
\leq&2^{d-1}n^{(d-1)/2}\sum_{i,j=1}^d Q_\vartheta\left(\frac{N_{i,j}}{n}<p_{k,l}-\varepsilon\right).
\end{aligned}
$$

Wir setzen $A_{i,j}:=\{(\tilde{p}_{1,1},\dots,\tilde{p}_{d-1,d})^t\in\Delta_{d^2};\tilde{p}_{i,j}<p_{k,l}-\varepsilon\}$, $i,j\in\{1,\dots,d\}$, und $A_n:=\{(i_{1,1}/n,\dots,i_{d-1,d}/n)^t;i_{1,1},\dots,i_{d-1,d}\in\mathbb{N}_0,i_{1,1}+\dots+i_{d-1,d}\leq n\}$. Dann ist mit Sanovs Theorem, siehe Example 5.4 in [4],

$$
Q_\vartheta\left(\frac{N_{i,j}}{n}<p_{k,l}-\varepsilon\right)\leq(n+1)^{d^2}\exp\big(-n\inf_{\tilde{p}\in A_{i,j}\cap A_n}K(\tilde{p},p)\big),\ i,j\in\{1,\dots,d\}.
$$

Weiter ist $A_{i,j}$ offen, $i,j\in\{1,\dots,d\}$, und Lemma 5.2 in [4] liefert $\lim_{n\to\infty}\inf_{\tilde{p}\in A_{i,j}\cap A_n} K(\tilde{p},p)=\inf_{\tilde{p}\in A_{i,j}}K(\tilde{p},p)>0$, $i,j\in\{1,\dots,d\}$. Damit folgt insgesamt

$$
\begin{aligned}
&Q_\vartheta\Bigg(T_n(\vartheta_1+\vartheta_1^*/\sqrt{n})<\chi^2_{d_1;1-\alpha},(\vartheta_1+\vartheta_1^*/\sqrt{n})\in\Theta^1,\hat{\hat{\vartheta}}_n^{\vartheta_1+\vartheta_1^*/\sqrt{n}}\in K_2,\\
&\exists\, i,j\in\{1,\dots,d\}:\frac{N_{i,j}}{n}<p_{k,l}-\varepsilon\Bigg)\\
\leq&(n+1)^{d^2}\sum_{i,j=1}^d\exp\big(-n\inf_{\tilde{p}\in A_{i,j}\cap A_n}K(\tilde{p},p)\big)\longrightarrow 0 \text{ für } n\to\infty
\end{aligned}
$$

sowie

$$\int Q_\vartheta\big(T_n(\vartheta_1+\vartheta_1^*/\sqrt{n})<\chi^2_{d_1;1-\alpha},(\vartheta_1+\vartheta_1^*/\sqrt{n})\in\Theta^1,\hat{\tilde{\vartheta}}_n^{\vartheta_1+\vartheta_1^*/\sqrt{n}}\in K_2\big)\mathrm{d}\vartheta_1^*$$
$$\leq 2^{d-1}n^{(d-1)/2}(n+1)^{d^2}\sum_{i,j=1}^d \exp\big(-n\inf_{\tilde{p}\in A_{i,j}\cap A_n}K(\tilde{p},p)\big)\longrightarrow 0 \text{ für } n\to\infty.$$

Somit sind die Voraussetzungen von Lemma 4.4.5 erfüllt.

Nun werden wir die Bedingungen aus Lemma 4.4.6 im verbundenen Stichprobenfall nachweisen. Dazu seien α, p, $p_{k,l}$, ε, $\tilde{K}_1$, $\tilde{K}_2$, K_1 und K_2 wie oben gewählt. Dann ist $K_1,K_2\in\mathfrak{B}_{\tilde{\Theta}}^{d^2-1}$, $K_1\cap K_2=\emptyset$, $K_1\cup K_2=\tilde{\Theta}$ und K_1 ist konvex und abgeschlossen. Beachte nun die Notationen aus Unterabschnitt 4.4. Weiter sei

$$\begin{aligned}W_{1,n}:=&\Big\{(x_1,\dots,x_n)\in\times_{i=1}^n\{0,1\}^{d^2-1};\forall\ i,j\in\{1,\dots,d\}:\frac{x_{i,j}}{n}\geq p_{k,l}-\varepsilon,\\&x:=x_1+\dots+x_n,x=(x_{1,1},\dots,x_{d-1,d})^t\Big\},\end{aligned}$$

sowie $W_{2,n}:=\times_{i=1}^n\{0,1\}^{d^2-1}\setminus W_{1,n}$. Dann ist $W_{1,n}\cap W_{2,n}=\emptyset$ und $W_{1,n}\cup W_{2,n}=\times_{i=1}^n\{0,1\}^{d^2-1}$. Es sei $\tilde{f}(\cdot;\tilde{p})$ die Dichte von N_1 bezüglich des Zählmaßes μ auf $\mathbb{N}_0^{d^2-1}$ bei zugrunde liegendem Parameter $\tilde{p}\in\Delta_{d^2}$. Dann ist

$$\tilde{f}(N_1;\tilde{p})=\prod_{i,j=1}^d\tilde{p}_{i,j}^{(N_1)_{i,j}}.$$

Damit ergibt sich mit $L(N_m,\tilde{p})=\log\tilde{f}(N_m;\tilde{p})$, $m\in\mathbb{N}$,

$$\begin{aligned}&\left(-\frac{1}{n}\sum_{m=1}^n\frac{\partial^2}{\partial\tilde{p}_{i,j}\partial\tilde{p}_{k,l}}L(N_m,\tilde{p})\right)_{\substack{1\leq i,j\leq d,(i,j)\neq(d,d)\\1\leq k,l\leq d,(k,l)\neq(d,d)}}=\frac{1}{n}\sum_{m=1}^n\operatorname{diag}\left(\frac{(N_m)_{1,1}}{\tilde{p}_{1,1}^2},\dots,\frac{(N_m)_{d-1,d}}{\tilde{p}_{d-1,d}^2}\right)\\&=\operatorname{diag}\left(\frac{N_{1,1}/n}{\tilde{p}_{1,1}^2},\dots,\frac{N_{d-1,d}/n}{\tilde{p}_{d-1,d}^2}\right),\ n\in\mathbb{N}.\end{aligned}$$

Wir setzen

$$\tilde{M}(q,\tilde{p}):=\operatorname{diag}\left(\frac{q_{1,1}}{\tilde{p}_{1,1}^2},\dots,\frac{q_{d-1,d}}{\tilde{p}_{d-1,d}^2}\right),\ (q,\tilde{p})\in\Delta_{d^2}^0\times\Delta_{d^2}.$$

Wegen der Kettenregel ist

$$\left(-\frac{1}{n}\sum_{j=1}^n\frac{\partial^2}{\partial\tilde{\tilde{\vartheta}}_i\partial\tilde{\tilde{\vartheta}}_l}L(N_j,\tilde{\tilde{\vartheta}})_{|_{\tilde{\tilde{\vartheta}}=\phi^{-1}(\tilde{p})}}\right)_{\substack{1\leq i\leq d^2-1\\1\leq l\leq d^2-1}}=\psi_{N_1}^t\tilde{M}(N/n,\tilde{p})\psi_{N_1},$$

mit der in Unterabschnitt 9.5 eingeführten regulären Matrix ψ_{N_1}. Dann definieren wir

$$M(q,\tilde{\tilde{\vartheta}}):=\psi_{N_1}^t\tilde{M}\big(q,\phi(\tilde{\tilde{\vartheta}})\big)\psi_{N_1},\ (q,\tilde{\tilde{\vartheta}})\in\Delta_{d^2}^0\times\tilde{\Theta}.$$

Es ist $\tilde{M}(q,\tilde{p})$ positiv definit für alle $q,\tilde{p} \in \Delta_{d^2}$. Da die Matrix ψ_{N_1} regulär ist, ist für $\tilde{\tilde{\vartheta}} \in K_1$ die Matrix $M(N/n,\tilde{\tilde{\vartheta}})$ positiv definit auf $\{(N_1,\ldots,N_n) \in W_{1,n}\}$ für alle $n \in \mathbb{N}$, und K_1 ist abgeschlossen. Daher existiert eine reelle Zahl $\lambda_{\min} > 0$ mit

$$\inf_{\tilde{\tilde{\vartheta}} \in K_1} \lambda_{n,\tilde{\tilde{\vartheta}}} > \lambda_{\min} \text{ auf } \{(N_1,\ldots,N_n) \in W_{1,n}\} \text{ für alle } n \in \mathbb{N},$$

mit $\lambda_{n,\tilde{\tilde{\vartheta}}}$ als kleinster Eigenwert der Matrix $M(N/n,\tilde{\tilde{\vartheta}})$, $\tilde{\tilde{\vartheta}} \in K_1$. Weiter ist

$$\begin{aligned}
&Q_\vartheta\big(T_n(\vartheta_1+\vartheta_1^*/\sqrt{n}) < \chi^2_{d_1;1-\alpha}, (\vartheta_1+\vartheta_1^*/\sqrt{n}) \in \Theta^1, (N_1,\ldots,N_n) \in W_{1,n},\\
&\hat{\vartheta}_n \in K_2 \text{ oder } \hat{\hat{\vartheta}}_n^{\vartheta_1+\vartheta_1^*/\sqrt{n}} \in K_2\big)\\
\leq &\mathrm{I}\left((\vartheta_1+\vartheta_1^*/\sqrt{n}) \in \Theta^1\right) Q_\vartheta\big((N_1,\ldots,N_n) \in W_{1,n}, \hat{\vartheta}_n \in K_2\big)\\
&+ Q_\vartheta\big(T_n(\vartheta_1+\vartheta_1^*/\sqrt{n}) < \chi^2_{d_1;1-\alpha}, (\vartheta_1+\vartheta_1^*/\sqrt{n}) \in \Theta^1, (N_1,\ldots,N_n) \in W_{1,n},\\
&\hat{\hat{\vartheta}}_n^{\vartheta_1+\vartheta_1^*/\sqrt{n}} \in K_2\big).
\end{aligned}$$

Es ist analog zu den obigen Überlegungen

$$\begin{aligned}
&Q_\vartheta\big(T_n(\vartheta_1+\vartheta_1^*/\sqrt{n}) < \chi^2_{d_1;1-\alpha}, (\vartheta_1+\vartheta_1^*/\sqrt{n}) \in \Theta^1, (N_1,\ldots,N_n) \in W_{1,n}, \hat{\hat{\vartheta}}_n^{\vartheta_1+\vartheta_1^*/\sqrt{n}} \in K_2\big)\\
=&Q_\vartheta\Bigg(T_n(\vartheta_1+\vartheta_1^*/\sqrt{n}) < \chi^2_{d_1;1-\alpha}, (\vartheta_1+\vartheta_1^*/\sqrt{n}) \in \Theta^1, \hat{\hat{\vartheta}}_n^{\vartheta_1+\vartheta_1^*/\sqrt{n}} \in K_2,\\
&\forall\, i,j \in \{1,\ldots,d\} : \frac{N_{i,j}}{n} \geq p_{k,l} - \varepsilon\Bigg) = 0.
\end{aligned}$$

Damit ist wegen $\varepsilon \in (0, p_{k,l}/2)$ auch

$$\begin{aligned}
&\int Q_\vartheta\big(T_n(\vartheta_1+\vartheta_1^*/\sqrt{n}) < \chi^2_{d_1;1-\alpha}, (\vartheta_1+\vartheta_1^*/\sqrt{n}) \in \Theta^1, (N_1,\ldots,N_n) \in W_{1,n},\\
&\hat{\vartheta}_n \in K_2 \text{ oder } \hat{\hat{\vartheta}}_n^{\vartheta_1+\vartheta_1^*/\sqrt{n}} \in K_2\big) \mathrm{d}\vartheta_1^*\\
\leq &2^{d-1} n^{(d-1)/2} Q_\vartheta\Big(\forall\, i,j \in \{1,\ldots,d\} : \frac{N_{i,j}}{n} \geq p_{k,l} - \varepsilon, \exists\, i,j \in \{1,\ldots,d\} : \frac{N_{i,j}}{n} < \varepsilon\Big) = 0.
\end{aligned}$$

Weiter ist

$$\begin{aligned}
&Q_\vartheta\big(T_n(\vartheta_1+\vartheta_1^*/\sqrt{n}) < \chi^2_{d_1;1-\alpha}, (\vartheta_1+\vartheta_1^*/\sqrt{n}) \in \Theta^1, (N_1,\ldots,N_n) \in W_{2,n}\big)\\
\leq &\mathrm{I}\left((\vartheta_1+\vartheta_1^*/\sqrt{n}) \in \Theta^1\right) Q_\vartheta\Big(\exists\, i,j \in \{1,\ldots,d\} : \frac{N_{i,j}}{n} < p_{k,l} - \varepsilon\Big),
\end{aligned}$$

und damit analog zu den obigen Überlegungen und mit $A_{i,j}$, $i,j \in \{1,\ldots,d\}$, und A_n wie oben definiert

$$\begin{aligned}
&Q_\vartheta\big(T_n(\vartheta_1+\vartheta_1^*/\sqrt{n}) < \chi^2_{d_1;1-\alpha}, (\vartheta_1+\vartheta_1^*/\sqrt{n}) \in \Theta^1, (N_1,\ldots,N_n) \in W_{2,n}\big)\\
\leq &(n+1)^{d^2} \sum_{i,j=1}^{d} \exp\big(-n \inf_{\tilde{p} \in A_{i,j} \cap A_n} K(\tilde{p},p)\big) \longrightarrow 0 \text{ für } n \to \infty,
\end{aligned}$$

sowie

$$\int Q_\vartheta\big(T_n(\vartheta_1+\vartheta_1^*/\sqrt{n}) < \chi^2_{d_1;1-\alpha}, (\vartheta_1+\vartheta_1^*/\sqrt{n}) \in \Theta^1, (N_1,\dots,N_n)\in W_{2,n}\big)\mathrm{d}\vartheta_1^*$$

$$\leq 2^{d-1} n^{(d-1)/2}(n+1)^{d^2}\sum_{i,j=1}^{d}\exp\big(-n\inf_{\tilde p\in A_{i,j}\cap A_n} K(\tilde p, p)\big) \longrightarrow 0 \text{ für } n\to\infty.$$

Schließlich gilt hier auch

$$\sum_{j=1}^{n}\nabla_k L(N_j,\tilde{\tilde{\vartheta}})|_{\tilde{\tilde{\vartheta}}=\hat\vartheta_n} = 0 \text{ auf } \{(N_1,\dots,N_n)\in W_{1,n}\} \text{ für alle } n\in\mathbb{N},$$

sodass die Bedingungen aus Lemma 4.4.6 im verbundenen Stichprobenfall erfüllt sind. Somit ergibt sich insgesamt analog zum Beweis von Korollar 9.5.1 durch Anwendung von Korollar 4.4.7 die Behauptung. □

9.8.2 Beispiel: Für das Beispiel $d=2$ betrachte die Ausführungen in Beispiel 9.5.3. Es gilt

$$\lim_{n\to\infty} e^{\mathrm{vol}}_{T_j',T_i,n}(\tilde\vartheta) = \frac{p_{1,2}+p_{2,1}}{p_{1\cdot}p_{2\cdot}+p_{\cdot 1}p_{\cdot 2}},\ \tilde\vartheta\in\Theta,\ i,j=1,2,$$

sowie

$$e^{\mathrm{vol}}_{T_1',T_1}(\tilde\vartheta)\begin{Bmatrix}<\\=\\>\end{Bmatrix}1 \ \Leftrightarrow\ \mathrm{Cov}_{\tilde\vartheta}(X_1,Y_1)\begin{Bmatrix}>\\=\\<\end{Bmatrix}0,\ \tilde\vartheta\in\Theta,\ i,j=1,2.$$

9.9 Tests zur relativen Wirksamkeit

Um Tests zum Vergleich der Verfahren der verbundenen und unabhängigen Stichprobenerhebung anzugeben, kann hier auf die in Abschnitt 5 entwickelten Tests zurückgegriffen werden. Bezugnehmend auf die zu Beginn von Unterabschnitt 9.5 eingeführte Parametrisierung sei $\vartheta_0 := \left(\begin{smallmatrix}0\\ \vartheta_2\end{smallmatrix}\right)$, $\vartheta_0' := \left(\begin{smallmatrix}0\\ \vartheta_2'\end{smallmatrix}\right)$ und $\vartheta^0 := (0,\vartheta_2,\vartheta_2')$. Zunächst wollen wir konsistente Tests für die Testprobleme

$$\begin{aligned}
&\mathrm{H}_1: \sup_{\vartheta_1\in\mathbb{R}^{d_1}\setminus\{0\}} \mathrm{e}^{\mathrm{LB}}_{T_1',T_1}(\vartheta^0,\vartheta_1)\leq 1,\ \mathrm{K}_1: \sup_{\vartheta_1\in\mathbb{R}^{d_1}\setminus\{0\}} \mathrm{e}^{\mathrm{LB}}_{T_1',T_1}(\vartheta^0,\vartheta_1)>1\\
\Leftrightarrow\ &\mathrm{H}_1: \sup_{\vartheta_1\in\mathbb{R}^{d_1}\setminus\{0\}} \mathrm{e}^{\mathrm{LHL}}_{T_1',T_1}(\vartheta^0,\vartheta_1)\leq 1,\ \mathrm{K}_1: \sup_{\vartheta_1\in\mathbb{R}^{d_1}\setminus\{0\}} \mathrm{e}^{\mathrm{LHL}}_{T_1',T_1}(\vartheta^0,\vartheta_1)>1\\
\Leftrightarrow\ &\mathrm{H}_1: \sup_{\vartheta_1\in\mathbb{R}^{d_1}\setminus\{0\}} \mathrm{e}^{\mathrm{P}}_{T_j',T_i}(\vartheta^0,\vartheta_1)\leq 1,\ \mathrm{K}_1: \sup_{\vartheta_1\in\mathbb{R}^{d_1}\setminus\{0\}} \mathrm{e}^{\mathrm{P}}_{T_j',T_i}(\vartheta^0,\vartheta_1)>1,\ i,j=1,2,
\end{aligned}$$

sowie

$$\begin{aligned}
&\mathrm{H}_2: \inf_{\vartheta_1\in\mathbb{R}^{d_1}\setminus\{0\}} \mathrm{e}^{\mathrm{LB}}_{T_1',T_1}(\vartheta^0,\vartheta_1)\geq 1,\ \mathrm{K}_2: \inf_{\vartheta_1\in\mathbb{R}^{d_1}\setminus\{0\}} \mathrm{e}^{\mathrm{LB}}_{T_1',T_1}(\vartheta^0,\vartheta_1)<1\\
\Leftrightarrow\ &\mathrm{H}_2: \inf_{\vartheta_1\in\mathbb{R}^{d_1}\setminus\{0\}} \mathrm{e}^{\mathrm{LHL}}_{T_1',T_1}(\vartheta^0,\vartheta_1)\geq 1,\ \mathrm{K}_2: \inf_{\vartheta_1\in\mathbb{R}^{d_1}\setminus\{0\}} \mathrm{e}^{\mathrm{LHL}}_{T_1',T_1}(\vartheta^0,\vartheta_1)<1\\
\Leftrightarrow\ &\mathrm{H}_2: \inf_{\vartheta_1\in\mathbb{R}^{d_1}\setminus\{0\}} \mathrm{e}^{\mathrm{P}}_{T_j',T_i}(\vartheta^0,\vartheta_1)\geq 1,\ \mathrm{K}_2: \inf_{\vartheta_1\in\mathbb{R}^{d_1}\setminus\{0\}} \mathrm{e}^{\mathrm{P}}_{T_j',T_i}(\vartheta^0,\vartheta_1)<1,\ i,j=1,2,
\end{aligned}$$

und

$$\begin{aligned}
&\mathrm{H}_3:\ \forall\,\vartheta_1\in\mathbb{R}^{d_1}\setminus\{0\}: e^{\mathrm{LB}}_{T_1',T_1}(\vartheta^0,\vartheta_1)=1,\ \mathrm{K}_3:\ \exists\,\vartheta_1\in\mathbb{R}^{d_1}\setminus\{0\}: e^{\mathrm{LB}}_{T_1',T_1}(\vartheta^0,\vartheta_1)\neq 1\\
\Leftrightarrow\ &\mathrm{H}_3:\ \forall\,\vartheta_1\in\mathbb{R}^{d_1}\setminus\{0\}: e^{\mathrm{LHL}}_{T_1',T_1}(\vartheta^0,\vartheta_1)=1,\ \mathrm{K}_3:\ \exists\,\vartheta_1\in\mathbb{R}^{d_1}\setminus\{0\}: e^{\mathrm{LHL}}_{T_1',T_1}(\vartheta^0,\vartheta_1)\neq 1\\
\Leftrightarrow\ &\mathrm{H}_3:\ \forall\,\vartheta_1\in\mathbb{R}^{d_1}\setminus\{0\}: e^{\mathrm{P}}_{T_j',T_i}(\vartheta^0,\vartheta_1)=1,\ \mathrm{K}_3:\ \exists\,\vartheta_1\in\mathbb{R}^{d_1}\setminus\{0\}: e^{\mathrm{P}}_{T_j',T_i}(\vartheta^0,\vartheta_1)\neq 1,\\
&i,j=1,2,
\end{aligned}$$

aufstellen. Beachte dabei Bemerkung 4.5.3. Dafür werden wir die in Unterabschnitt 5.2 entwickelte Theorie zur Anwendung bringen.

Unter Beachtung von Bemerkung 9.5.2 führen wir basierend auf n Beobachtungen der verbundenen Stichprobe mit $\hat{p}_n := N/n$ durch

$$S_n := \operatorname{diag}(\hat{p}_n^t) - \hat{p}_n\hat{p}_n^t$$

bzw. mit $\hat{p}_{*\cdot,n} := N_{*\cdot}/n$ und $\hat{p}_{\cdot*,n} := N_{\cdot*}/n$

$$S_n' := \begin{pmatrix} \operatorname{diag}(\hat{p}_{*\cdot,n}^t) - \hat{p}_{*\cdot,n}\hat{p}_{*\cdot,n}^t & 0_{(d-1)\times(d-1)} \\ 0_{(d-1)\times(d-1)} & \operatorname{diag}(\hat{p}_{\cdot*,n}^t) - \hat{p}_{\cdot*,n}\hat{p}_{\cdot*,n}^t \end{pmatrix}$$

Schätzer für $\operatorname{Cov}_{\vartheta_0}(N_1)$ und $\operatorname{Cov}_{\vartheta_0}(N_1')$ ein. Diese Schätzer basieren ausschließlich auf $\hat{p}_n$.

Entsprechend der in Unterabschnitt 5.2 eingeführten Bezeichnungen ist hier

$$\begin{aligned}
\tilde{E}(p_0) =&\Big((\mathrm{I}_{(d-1)\times(d-1)}, 0_{(d-1)\times(d-1)})\psi_{N_1'}^{-1}\operatorname{Cov}_{p_0}(N_1')\big(\psi_{N_1'}^t\big)^{-1}(\mathrm{I}_{(d-1)\times(d-1)}, 0_{(d-1)\times(d-1)})^t\Big)^{-1}\\
&-\Big((\mathrm{I}_{(d-1)\times(d-1)}, 0_{(d-1)\times(d^2-d)})\psi_{N_1}^{-1}\operatorname{Cov}_{p_0}(N_1)\big(\psi_{N_1}^t\big)^{-1}(\mathrm{I}_{(d-1)\times(d-1)}, 0_{(d-1)\times(d^2-d)})^t\Big)^{-1},
\end{aligned}$$

und wir betrachten zunächst das zu dieser Matrix entsprechende empirische Gegenstück,

$$\begin{aligned}
&\Big((\mathrm{I}_{(d-1)\times(d-1)}, 0_{(d-1)\times(d-1)})\psi_{N_1'}^{-1}S_n'\big(\psi_{N_1'}^t\big)^{-1}(\mathrm{I}_{(d-1)\times(d-1)}, 0_{(d-1)\times(d-1)})^t\Big)^{-1}\\
&-\Big((\mathrm{I}_{(d-1)\times(d-1)}, 0_{(d-1)\times(d^2-d)})\psi_{N_1}^{-1}S_n\big(\psi_{N_1}^t\big)^{-1}(\mathrm{I}_{(d-1)\times(d-1)}, 0_{(d-1)\times(d^2-d)})^t\Big)^{-1},
\end{aligned}$$

um Tests für die genannten Testprobleme zu formulieren. Schließlich gehen wir aber unter Hinweis auf Bemerkung 5.2.10 stattdessen zur Matrix

$$\begin{aligned}
\tilde{\tilde{E}}(p_0) =&(\mathrm{I}_{(d-1)\times(d-1)}, 0_{(d-1)\times(d-1)})\psi_{N_1'}^{-1}\operatorname{Cov}_{p_0}(N_1')\big(\psi_{N_1'}^t\big)^{-1}(\mathrm{I}_{(d-1)\times(d-1)}, 0_{(d-1)\times(d-1)})^t\\
&-(\mathrm{I}_{(d-1)\times(d-1)}, 0_{(d-1)\times(d^2-d)})\psi_{N_1}^{-1}\operatorname{Cov}_{p_0}(N_1)\big(\psi_{N_1}^t\big)^{-1}(\mathrm{I}_{(d-1)\times(d-1)}, 0_{(d-1)\times(d^2-d)})^t,
\end{aligned}$$

bzw. zu ihrem empirischen Gegenstück

$$\begin{aligned}
&(\mathrm{I}_{(d-1)\times(d-1)}, 0_{(d-1)\times(d-1)})\psi_{N_1'}^{-1}S_n'\big(\psi_{N_1'}^t\big)^{-1}(\mathrm{I}_{(d-1)\times(d-1)}, 0_{(d-1)\times(d-1)})^t\\
&-(\mathrm{I}_{(d-1)\times(d-1)}, 0_{(d-1)\times(d^2-d)})\psi_{N_1}^{-1}S_n\big(\psi_{N_1}^t\big)^{-1}(\mathrm{I}_{(d-1)\times(d-1)}, 0_{(d-1)\times(d^2-d)})^t
\end{aligned}$$

über, und lassen unsere Tests auf den nach wachsender Größe geordneten Eigenwerten $\tilde{\lambda}_{1,n} \leq \cdots \leq \tilde{\lambda}_{d-1,n}$ dieser Matrix basieren.

Es ist

$$\begin{aligned}
&\tilde{\tilde{E}}(p_0) = \tilde{H}(p_0) \\
=&(\mathrm{I}_{(d-1)\times(d-1)}, 0_{(d-1)\times(d-1)})\psi_{N_1'}^{-1}\mathrm{Cov}_{p_0}(N_1')\big(\psi_{N_1'}^t\big)^{-1}(\mathrm{I}_{(d-1)\times(d-1)}, 0_{(d-1)\times(d-1)})^t \\
&- (\mathrm{I}_{(d-1)\times(d-1)}, 0_{(d-1)\times(d^2-d)})\psi_{N_1}^{-1}\mathrm{Cov}_{p_0}(N_1)\big(\psi_{N_1}^t\big)^{-1}(\mathrm{I}_{(d-1)\times(d-1)}, 0_{(d-1)\times(d^2-d)})^t,
\end{aligned}$$

mit der in Unterabschnitt 5.2 eingeführten Abbildung $\tilde{H}$, und

$$\begin{aligned}
\tilde{H}(\hat{p}_n) =&(\mathrm{I}_{(d-1)\times(d-1)}, 0_{(d-1)\times(d-1))})\psi_{N_1'}^{-1}S_n'\big(\psi_{N_1'}^t\big)^{-1}(\mathrm{I}_{(d-1)\times(d-1)}, 0_{(d-1)\times(d-1))})^t \\
&- (\mathrm{I}_{(d-1)\times(d-1)}, 0_{(d-1)\times(d^2-d)})\psi_{N_1}^{-1}S_n\big(\psi_{N_1}^t\big)^{-1}(\mathrm{I}_{(d-1)\times(d-1)}, 0_{(d-1)\times(d^2-d)})^t.
\end{aligned}$$

Zur Behandlung des Testproblems H_1 gegen K_1 verwenden wir die Testgröße

$$F_n^1 = -\frac{\sqrt{n}\tilde{\lambda}_{d-1,n}}{\sqrt{M_{d,n}}},$$

und zur Behandlung des Testproblems H_2 gegen K_2 die Testgröße

$$F_n^2 = \frac{\sqrt{n}\tilde{\lambda}_{1,n}}{\sqrt{M_{1,n}}},$$

mit

$$M_{i,n} = w_i(\hat{p}_n)^t \mathcal{J}_{\tilde{H}}(\hat{p}_n) V(\hat{p}_n) \mathcal{J}_{\tilde{H}}(\hat{p}_n)^t w_i(\hat{p}_n),\ i = 1, \ldots, d-1,$$

wobei für $i = 1, \ldots, d-1$ die Größen $w_i(\hat{p}_n), \mathcal{J}_{\tilde{H}}(\hat{p}_n)$ und $V(\hat{p}_n)$ die empirischen Gegenstücke zu $w_i(p_0), \mathcal{J}_{\tilde{H}}(p_0)$ und $V(p_0)$ sind, die wie folgt definiert werden: Für $i = 1, \ldots, d-1$ ist der Vektor $w_i(p_0) \in \mathbb{R}^{(d-1)d/2}$ durch die Komponenten

$$\big(w_i(p_0)\big)_j := \big(\beta_i(p_0)^t \otimes \beta_i(p_0)^t\big) D e_j,\ j = 1, \ldots, (d-1)d/2,$$

festgelegt. Dabei ist $\beta_i(p_0) \in \mathbb{R}^{d-1}$ der normierte Eigenvektor zum Eigenwert $\tilde{\lambda}_i(p_0)$ der Matrix $\tilde{H}(p_0)$, $i = 1, \ldots, d-1$, und D die in Unterabschnitt 5.2 eingeführte Duplikationsmatrix. Weiter ist

$$V(p_0) := \mathrm{Cov}_{p_0}(N_1),$$

und $\mathcal{J}_{\tilde{H}}(p_0) \in \mathbb{R}^{(d-1)d/2\times(d^2-1)}$ ist die Jacobi-Matrix der Abbildung $\mathrm{vech}(\tilde{H}) : \mathbb{R}^{d^2-1} \to \mathbb{R}^{(d-1)d/2}$. Die Abbildung $\mathrm{vech}(\tilde{H})$ ist offensichtlich stetig differenzierbar in einer Umgebung um p_0.

Es hängt $V(p_0)$ stetig von p_0 ab, und wegen $p_0 \in \Delta_{d^2}$ ist $V(p_0)$ positiv definit. Wir nehmen an, dass die Jacobi-Matrix $\mathcal{J}_{\tilde{H}}(p_0)$ vollen Zeilenrang besitzt. Es gilt bei zugrunde liegendem Parameter $p_0 \in \Delta_{d^2}$ mit dem mehrdimensionalen zentralen Grenzwertsatz

$$\sqrt{n}(N/n - p_0) \xrightarrow{\mathrm{v}} N_{d^2-1}\big(0, V(p_0)\big) \text{ für } n \to \infty.$$

Somit sind die Bedingungen (E1) und (E2) aus Unterabschnitt 5.2 erfüllt. Wie in Unterabschnitt 5.2 argumentiert wurde, gilt im Fall der Einfachheit des kleinsten und größten Eigenwerts der Matrix $\tilde{H}(p_0)$

$$\lim_{n\to\infty} P_{p_0}(M_{1,n} \leq 0) = \lim_{n\to\infty} P_{p_0}(M_{d_1,n} \leq 0) = 0.$$

Für ein gegebenes Testniveau $\alpha \in (0,1)$ schlagen wir für das Testproblem H_i gegen K_i, $i = 1, 2$, den Test

$$\text{„Verwirf } \mathrm{H}_i, \text{ falls } F_n^i \geq u_{1-\alpha}\text{“}, \ i = 1, 2,$$

vor und können folgendes festhalten.

9.9.1 Korollar: *Es sei der kleinste und größte Eigenwert der Matrix $\tilde{H}(p_0)$ einfach. Dann handelt es sich bei dem auf der Testgröße F_n^i basierenden vorgeschlagenen Test für das Testproblem H_i gegen K_i, $i = 1, 2$, um einen Test, der für jedes $\alpha \in (0,1)$ das Testniveau α asymptotisch einhält und der zudem konsistent ist.*

Beweis: Wie oben erklärt sind die Bedingungen (E1) und (E2) aus Unterabschnitt 5.2 erfüllt. Da nach Annahme auch die Bedingung (E3) aus Korollar 5.2.6 erfüllt ist, folgt mit dem Korollar die Behauptung. □

Zur Behandlung des Testproblems H_3 gegen K_3 betrachten wir die Testgröße

$$F_n^3 = n\tilde{\lambda}^2_{(d-1),n}.$$

Dabei seien $\tilde{\lambda}_{(1),n}, \ldots, \tilde{\lambda}_{(d-1),n}$ die von ihrem Betrag her nach wachsender Größe geordneten Eigenwerte von $\tilde{H}(\hat{p}_n)$, d.h. es gilt $|\tilde{\lambda}_{(1),n}| \leq \cdots \leq |\tilde{\lambda}_{(d-1),n}|$.

Für ein gegebenes Testniveau $\alpha \in (0,1)$ schlagen wir für das Testproblem H_3 gegen K_3 den Test

$$\text{„Verwirf } \mathrm{H}_3, \text{ falls } F_n^3 \geq c_{\hat{p}_n;1-\alpha}\text{“}$$

vor. Hierbei ist $c_{p_0;1-\alpha}$ das $(1-\alpha)$-Quantil der Verteilung mit der Lebesgue-Dichte

$$k_{p_0}(l_{d-1}) = \int \cdots \int \frac{\pi^{\frac{(d-1)^2}{2}}}{\Gamma_{d-1}(\frac{d-1}{2})} \prod_{1\leq j<i\leq d-1} (l_i - l_j)$$

$$\int_{O_{d-1}} f_{p_0}\big(H \operatorname{diag}(l_{d-1}, \ldots, l_1)H'\big)\mathrm{d}H\mathrm{d}l_1 \ldots \mathrm{d}l_{d-2}, \text{ für } l_{d-1} > 0 \text{ und } 0 \text{ sonst.}$$

Dabei wird das innere Integral über die Menge $O_{d-1} = \{A \in \mathbb{R}^{(d-1)\times(d-1)}; A^t A = \mathrm{I}_{d-1}\}$ bezüglich der Gleichverteilung auf O_{d-1} durchgeführt, Γ_m sei die verallgemeinerte Gammafunktion $\Gamma_m(t) := \pi^{m(m-1)/4} \prod_{i=1}^m \Gamma(t-(i-1)/2),\ m \in \mathbb{N},\ t > (m-1)/2$, und f_{p_0} sei die Lebesgue-Dichte einer Matrix $A^2 \in \mathbb{R}^{(d-1)\times(d-1)}$, wobei $A \in \mathbb{R}^{(d-1)\times(d-1)}$ eine symmetrische Matrix mit $\mathrm{vech}(A) \sim \mathrm{N}_{(d-1)d/2}(0, \mathcal{J}_{\tilde{H}}(p_0)V(p_0)\mathcal{J}_{\tilde{H}}(p_0)^t)$ sei; siehe Unterabschnitt 5.2. Dann halten wir folgendes Resultat fest.

9.9.2 Korollar: *Bei dem auf der Testgröße F_n^3 basierenden vorgeschlagenen Test für das Testproblem* H_3 *gegen* K_3 *handelt es sich um einen Test, der für jedes $\alpha \in (0,1)$ das Testniveau α asymptotisch exakt einhält, und der zudem konsistent ist.*

Beweis: Die Aussage folgt aus Korollar 5.2.8. □

Nun wollen wir konsistente Tests unter sinnvollen Alternativen gemäß den Ausführungen aus Unterabschnitt 5.3 angeben. Betrachten wir also das Testproblem

$$\begin{aligned}
&\mathrm{H}_1: \sup_{\vartheta_1\in\mathbb{R}^{d_1}\setminus\{0\}} \mathrm{e}^{\mathrm{LB}}_{T_1',T_1}(\vartheta^0,\vartheta_1) \le 1,\ \tilde{\mathrm{K}}_1: \inf_{\vartheta_1\in\mathbb{R}^{d_1}\setminus\{0\}} \mathrm{e}^{\mathrm{LB}}_{T_1',T_1}(\vartheta^0,\vartheta_1) > 1\\
\Leftrightarrow\ &\mathrm{H}_1: \sup_{\vartheta_1\in\mathbb{R}^{d_1}\setminus\{0\}} \mathrm{e}^{\mathrm{LHL}}_{T_1',T_1}(\vartheta^0,\vartheta_1) \le 1,\ \tilde{\mathrm{K}}_1: \inf_{\vartheta_1\in\mathbb{R}^{d_1}\setminus\{0\}} \mathrm{e}^{\mathrm{LHL}}_{T_1',T_1}(\vartheta^0,\vartheta_1) > 1\\
\Leftrightarrow\ &\mathrm{H}_1: \sup_{\vartheta_1\in\mathbb{R}^{d_1}\setminus\{0\}} \mathrm{e}^{\mathrm{P}}_{T_j',T_i}(\vartheta^0,\vartheta_1) \le 1,\ \tilde{\mathrm{K}}_1: \inf_{\vartheta_1\in\mathbb{R}^{d_1}\setminus\{0\}} \mathrm{e}^{\mathrm{P}}_{T_j',T_i}(\vartheta^0,\vartheta_1) > 1,\ i,j=1,2.
\end{aligned}$$

Zudem betrachten wir das Testproblem

$$\begin{aligned}
&\mathrm{H}_2: \inf_{\vartheta_1\in\mathbb{R}^{d_1}\setminus\{0\}} \mathrm{e}^{\mathrm{LB}}_{T_1',T_1}(\vartheta^0,\vartheta_1) \ge 1,\ \tilde{\mathrm{K}}_2: \sup_{\vartheta_1\in\mathbb{R}^{d_1}\setminus\{0\}} \mathrm{e}^{\mathrm{LB}}_{T_1',T_1}(\vartheta^0,\vartheta_1) < 1\\
\Leftrightarrow\ &\mathrm{H}_2: \inf_{\vartheta_1\in\mathbb{R}^{d_1}\setminus\{0\}} \mathrm{e}^{\mathrm{LHL}}_{T_1',T_1}(\vartheta^0,\vartheta_1) \ge 1,\ \tilde{\mathrm{K}}_2: \sup_{\vartheta_1\in\mathbb{R}^{d_1}\setminus\{0\}} \mathrm{e}^{\mathrm{LHL}}_{T_1',T_1}(\vartheta^0,\vartheta_1) < 1\\
\Leftrightarrow\ &\mathrm{H}_2: \inf_{\vartheta_1\in\mathbb{R}^{d_1}\setminus\{0\}} \mathrm{e}^{\mathrm{P}}_{T_j',T_i}(\vartheta^0,\vartheta_1) \ge 1,\ \tilde{\mathrm{K}}_2: \sup_{\vartheta_1\in\mathbb{R}^{d_1}\setminus\{0\}} \mathrm{e}^{\mathrm{P}}_{T_j',T_i}(\vartheta^0,\vartheta_1) < 1,\ i,j=1,2.
\end{aligned}$$

Schließlich betrachten wir noch das Testproblem

$$\begin{aligned}
&\mathrm{H}_3:\ \forall\, \vartheta_1 \in \mathbb{R}^{d_1}\setminus\{0\}: \mathrm{e}^{\mathrm{LB}}_{T_1',T_1}(\vartheta^0,\vartheta_1) = 1,\\
&\tilde{\mathrm{K}}_3: \sup_{\vartheta_1\in\mathbb{R}^{d_1}\setminus\{0\}} \mathrm{e}^{\mathrm{LB}}_{T_1',T_1}(\vartheta^0,\vartheta_1) < 1 \text{ oder } \inf_{\vartheta_1\in\mathbb{R}^{d_1}\setminus\{0\}} \mathrm{e}^{\mathrm{LB}}_{T_1',T_1}(\vartheta^0,\vartheta_1) > 1\\
\Leftrightarrow\ &\mathrm{H}_3:\ \forall\, \vartheta_1 \in \mathbb{R}^{d_1}\setminus\{0\}:\ \mathrm{e}^{\mathrm{LHL}}_{T_1',T_1}(\vartheta^0,\vartheta_1) = 1,\\
&\tilde{\mathrm{K}}_3: \sup_{\vartheta_1\in\mathbb{R}^{d_1}\setminus\{0\}} \mathrm{e}^{\mathrm{LHL}}_{T_1',T_1}(\vartheta^0,\vartheta_1) < 1 \text{ oder } \inf_{\vartheta_1\in\mathbb{R}^{d_1}\setminus\{0\}} \mathrm{e}^{\mathrm{LHL}}_{T_1',T_1}(\vartheta^0,\vartheta_1) > 1\\
\Leftrightarrow\ &\mathrm{H}_3:\ \forall\, \vartheta_1 \in \mathbb{R}^{d_1}\setminus\{0\}:\ \mathrm{e}^{\mathrm{P}}_{T_1',T_1}(\vartheta^0,\vartheta_1) = 1,\\
&\tilde{\mathrm{K}}_3: \sup_{\vartheta_1\in\mathbb{R}^{d_1}\setminus\{0\}} \mathrm{e}^{\mathrm{P}}_{T_j',T_i}(\vartheta^0,\vartheta_1) < 1 \text{ oder } \inf_{\vartheta_1\in\mathbb{R}^{d_1}\setminus\{0\}} \mathrm{e}^{\mathrm{P}}_{T_j',T_i}(\vartheta^0,\vartheta_1) > 1,\ i,j=1,2.
\end{aligned}$$

Beachte Bemerkung 4.5.3. Diese Testprobleme sollen unter Hinweis auf Bemerkung 5.3.4 erneut mit Hilfe der Matrix $\tilde{H}(\hat{p}_n)$ behandelt werden. Zur Behandlung des Testproblems H_i gegen $\tilde{\mathrm{K}}_i$, $i = 1, 2, 3$, betrachten wir die Testgrößen

$$\tilde{F}_n^1 = \frac{\sqrt{n}\operatorname{tr}\big(\tilde{H}(\hat{p}_n)\big)}{\sqrt{\tilde{M}_n}}, \ \tilde{F}_n^2 = -\tilde{F}_n^1, \ \tilde{F}_n^3 = |\tilde{F}_n^1|,$$

mit

$$\tilde{M}_n = \tilde{M}_n(\hat{p}_n) = v^t \mathcal{J}_{\tilde{H}}(\hat{p}_n) V(\hat{p}_n) \mathcal{J}_{\tilde{H}}(\hat{p}_n)^t v,$$

wobei der Verktor $v \in \mathbb{R}^{(d-1)d/2}$ durch

$$v = (1, \underbrace{0, \ldots, 0}_{(d-2)\text{-mal}}, 1, \underbrace{0, \ldots, 0}_{(d-3)\text{-mal}}, 1, \ldots, 1, 0, 1)$$

gegeben ist. Wie oben gezeigt sind hier die Bedingungen (E1) und (E2) aus Unterabschnitt 5.2 erfüllt. Daher gilt

$$\lim_{n\to\infty} P_{p_0}(\tilde{M}_n \leq 0) = 0,$$

was in Unterabschnitt 5.3 begründet wurde.

Für ein gegebenes Testniveau $\alpha \in (0, 1)$ lautet dann der zugehörige Test für das Testproblem H_i gegen $\tilde{\mathrm{K}}_i$, $i = 1, 2$,

$$„\text{Verwirf } \mathrm{H}_i, \text{ falls } \tilde{F}_n^i \geq u_{1-\alpha}“, \ i = 1, 2.$$

Für dass Testproblem H_3 gegen K_3 kann der Test

$$„\text{Verwirf } \mathrm{H}_3, \text{ falls } \tilde{F}_n^3 \geq u_{1-\alpha/2}“$$

formuliert werden. Dann können wir folgendes festhalten.

9.9.3 Korollar: *Bei dem auf der Testgröße $\tilde{F}_n^i$ basierenden vorgeschlagenen Test für das Testproblem H_i gegen $\tilde{\mathrm{K}}_i$, $i = 1, 2$, handelt es sich um einen Test, der für jedes $\alpha \in (0, 1)$ das Testniveau α asymptotisch einhält und der zudem konsistent ist. Bei dem auf der Testgröße $\tilde{F}_n^3$ basierenden vorgeschlagenen Test für das Testproblem H_3 gegen $\tilde{\mathrm{K}}_3$ handelt es sich darüber hinaus um einen Test, der für jedes $\alpha \in (0, 1)$ das Testniveau α asymptotisch exakt einhält, und der zudem konsistent ist.*

Beweis: Die Aussage folgt aus Korollar 5.3.3. □

9.9.4 Bemerkung: Wie in Bemerkung 5.3.5 erläutert können analog konsistente Tests für die Testprobleme

$$\mathrm{H}_4: \; e^{\mathrm{Lvol}}_{T'_j,T_i}(\vartheta^0) \leq 1, \; \mathrm{K}_4: \; e^{\mathrm{Lvol}}_{T'_j,T_i}(\vartheta^0) > 1, \; i,j = 1,2,$$

sowie

$$\mathrm{H}_5: \; e^{\mathrm{Lvol}}_{T'_j,T_i}(\vartheta^0) \geq 1, \; \mathrm{K}_5: \; e^{\mathrm{Lvol}}_{T'_j,T_i}(\vartheta^0) < 1, \; i,j = 1,2,$$

und

$$\mathrm{H}_6: \; e^{\mathrm{Lvol}}_{T'_j,T_i}(\vartheta^0) = 1, \; \mathrm{K}_6: \; e^{\mathrm{Lvol}}_{T'_j,T_i}(\vartheta^0) \neq 1, \; i,j = 1,2,$$

basierend auf der Determinante der Matrix

$$\begin{aligned}&\Big((\mathrm{I}_{(d-1)\times(d-1)}, 0_{(d-1)\times(d-1)})\psi^{-1}_{N'_1} S'_n \big(\psi^t_{N'_1}\big)^{-1} (\mathrm{I}_{(d-1)\times(d-1)}, 0_{(d-1)\times(d-1)})^t\Big)^{-1} \\ &(\mathrm{I}_{(d-1)\times(d-1)}, 0_{(d-1)\times(d^2-d)})\psi^{-1}_{N_1} S_n \big(\psi^t_{N_1}\big)^{-1} (\mathrm{I}_{(d-1)\times(d-1)}, 0_{(d-1)\times(d^2-d)})^t\end{aligned}$$

konstruiert werden.

9.9.5 Beispiel: Für $d = 2$ ergibt sich mit Hilfe der Zwischenergebnisse aus Beispiel 9.5.3 für p aus der Hypothesenmenge

$$\tilde{H}(p) = \mathrm{tr}\tilde{H}(p) = \det \tilde{H}(p) = \tilde{\lambda}_{1,n} = \tilde{\lambda}_{(1),n} = (p_{1,1} + p_{1,2})\big(1 - (p_{1,1} + p_{1,2})\big) - p_{1,2}.$$

Daraus folgt

$$\mathcal{J}_{\tilde{H}}(p) = \big(1 - 2(p_{1,1} + p_{1,2}), 0, -2(p_{1,1} + p_{1,2})\big)$$

sowie mit $w_1(p) = 1$

$$\begin{aligned}&w_1(p)^t \mathcal{J}_{\tilde{H}}(p) V(p) \mathcal{J}_{\tilde{H}}(p)^t w_1(p) \\ =&\big(1 - 2(p_{1,1} + p_{1,2})\big)^2 (p_{1,1} - p^2_{1,1}) + 4(p_{1,1} + p_{1,2})^2 (p_{1,2} - p^2_{1,2}) \\ &+ 4\big(1 - 2(p_{1,1} + p_{1,2})\big)(p_{1,1} + p_{1,2}) p_{1,1} p_{1,2}.\end{aligned}$$

Durch Einsetzen des Schätzers $\hat{p}_n = N/n$ ergeben sich die betrachteten Testgrößen und die vorgeschlagenen Tests.

10 Vergleiche bei nichtparametrischen Lagealternativen

10.1 Modell

Das hier betrachtete Modell unterscheidet sich zu den in den Abschnitten 8 und 9 betrachteten Modellen insofern, dass die Resultate aus den Abschnitten 4, 5 und 6 hier nicht angewendet werden können. Dies liegt daran, dass für die hier untersuchten Tests im Allgemeinen nicht die in den Abschnitten 4, 5 und 6 vorausgesetzten Optimalitätseigenschaften nachgewiesen werden können.

Gehen wir davon aus, dass im verbundenen Stichprobenfall durch $\binom{X_1}{Y_1}, \ldots, \binom{X_n}{Y_n}$ unabhängige und identisch verteilte 2-dimensionale Zufallsvektoren vorliegen. Es gelte

$$X_1 \sim F(\cdot),\ Y_1 \sim F(\cdot - \vartheta),$$

mit einem unbekannten $\vartheta \in \mathbb{R}$ und einer evtl. als unbekannt aber als stetig vorausgesetzten Verteilungsfunktion F. Ferner gelte

$$X_1 - Y_1 \sim G(\cdot + \vartheta),$$

mit einer unbekannten aber ebenfalls als stetig vorausgesetzten Verteilungsfunktion G, die zu einer um Null symmetrischen Verteilung gehöre.

Im unabhängigen Stichprobenfall liegen mit $X_1', \ldots, X_n'$ und $Y_1', \ldots, Y_n'$ zwei voneinander unabhängige Stichproben vor, wobei

$$\mathcal{L}(X_1) = \mathcal{L}(X_1') \text{ und } \mathcal{L}(Y_1) = \mathcal{L}(Y_1'),$$

und damit

$$X_1' \sim F(\cdot),\ Y_1' \sim F(\cdot - \vartheta),$$

gelte. Dann gilt auch

$$X_1' - Y_1' \sim \tilde{F}(\cdot + \vartheta),$$

mit einer stetigen Verteilungsfunktion $\tilde{F}$, die zu einer um Null symmetrischen Verteilung gehört, und die von F abhängt.

Das zu Beginn dieser Arbeit vorgestellte Testproblem (1.1.3),

$$\mathrm{H}:\ \mathcal{L}(X_1) = \mathcal{L}(Y_1),\ \mathrm{K}:\ \mathcal{L}(X_1) \neq \mathcal{L}(Y_1),$$

kann hier äquivalent gemäß

$$\mathrm{H}:\ \vartheta = 0,\ \mathrm{K}:\ \vartheta \neq 0$$

formuliert werden.

10.1.1 Beispiel: Es besitze $\binom{X_1}{Y_1}$ die Dichte $\binom{x}{y} \mapsto v(x, y-\vartheta)$ bezüglich des Lebesgue-Borelschen Maßes, $x, y \in \mathbb{R}$, mit einer zweidimensionalen Dichte v bezüglich des Lebesgue-Borelschen Maßes, die $v(x,y) = v(y,x) \ \forall \ x, y \in \mathbb{R}$ erfülle. Dann besitzt Y_1 die Dichte $y \mapsto \int v(x, y-\vartheta)\mathrm{d}x =: f(y;\vartheta)$, $y \in \mathbb{R}$, und X_1 besitzt die Dichte

$$x \longmapsto \int v(x,y)\mathrm{d}y = \int v(y,x)\mathrm{d}y = f(x;0), x \in \mathbb{R}.$$

Setzen wir $h(z) := \int v(z+y, y)\mathrm{d}y$, $z \in \mathbb{R}$, so erhalten wir mit der Substitution $y \mapsto y+\vartheta$, dass $X_1 - Y_1$ die Dichte

$$z \longmapsto \int v(z+y, y-\vartheta)\mathrm{d}y = \int v(z+\vartheta+y, y)\mathrm{d}y = h(z+\vartheta), \ z \in \mathbb{R},$$

besitzt. Ferner erhalten wir durch die Substitution $y \mapsto y+z$

$$h(-z) = \int v(-z+y, y)\mathrm{d}y = \int v(y, z+y)\mathrm{d}y = \int v(z+y, y)\mathrm{d}y = h(z), \ z \in \mathbb{R},$$

sodass $\binom{X_1}{Y_1}$ die genannten Bedingungen erfüllt.

10.1.2 Beispiel: Es besitze $\binom{X_1}{Y_1}$ die Dichte $\binom{x}{y} \mapsto v(x, y-\vartheta)$ bezüglich des Lebesgue-Borelschen Maßes, $x, y \in \mathbb{R}$, mit einer zweidimensionalen Dichte v bezüglich des Lebesgue-Borelschen Maßes, die $v(x,y) = v(-x,-y) \ \forall \ x, y \in \mathbb{R}$ und $\int v(x,y)\mathrm{d}y = \int v(y,x)\mathrm{d}y \ \forall \ x \in \mathbb{R}$ erfülle. Dann besitzt Y_1 die Dichte $y \mapsto \int v(x, y-\vartheta)\mathrm{d}x =: f(y;\vartheta)$, $y \in \mathbb{R}$, und X_1 besitzt die Dichte

$$x \longmapsto \int v(x,y)\mathrm{d}y = \int v(y,x)\mathrm{d}y = f(x;0), x \in \mathbb{R}.$$

Setzen wir $h(z) := \int v(z+y, y)\mathrm{d}y$, $z \in \mathbb{R}$, so erhalten wir mit der Substitution $y \mapsto y+\vartheta$, dass $X_1 - Y_1$ die Dichte

$$z \longmapsto \int v(z+y, y-\vartheta)\mathrm{d}y = \int v(z+\vartheta+y, y)\mathrm{d}y = h(z+\vartheta), \ z \in \mathbb{R},$$

besitzt. Ferner erhalten wir durch die Substitution $y \mapsto -y$

$$h(-z) = \int v(-z+y, y)\mathrm{d}y = \int v(-z-y, -y)\mathrm{d}y = \int v(z+y, y)\mathrm{d}y = h(z), \ z \in \mathbb{R},$$

sodass $\binom{X_1}{Y_1}$ die genannten Bedingungen erfüllt.

10.2 Verbundener Stichprobenfall

Im verbundenen Stichprobenfall sind Abhängigkeiten zwischen X_1 und Y_1 zugelassen. Klassische Zweistichprobentests für die Hypothese der Homogenität sind deshalb hier nicht anwendbar. Allerdings kann zur Verifizierung der Hypothese ein Symmetrietest basierend auf den Differenzen $\zeta_i := X_i - Y_i$, $1 \leq i \leq n$, herangezogen werden. Als Testgröße für das Testproblem H gegen K wählen wir die Wilcoxon-Vorzeichen-Rang Teststatistik.

$$T_n := \sum_{i=1}^{n} R_i^+ \, \mathrm{I}(\zeta_i > 0),$$

wobei R_i^+ den Rang von $|\zeta_i|$ in den Beobachtungen $|\zeta_1|, \ldots, |\zeta_n|$, $1, \leq i \leq n$, bezeichne. Für ein gegebenes Testniveau $\alpha \in (0,1)$ schlagen wir als Test den zweiseitigen Wilcoxon-Vorzeichen-Rang Test

$$\text{„Verwirf H, falls } T_n < \frac{1}{2}n(n+1) - c_n \text{ oder } T_n > c_n\text{“}$$

vor, wobei c_n ein $(1-\alpha/2)$-Quantil der Verteilung von T_n bei Gültigkeit der Hypothese bezeichne. Es ist für jedes $n \in \mathbb{N}$ $c_n \in \mathbb{N}_0$ und $0 \leq c_n \leq n(n+1)/2$. Man beachte, dass die Verteilung von T_n bei Gültigkeit der Hypothese der Symmetrie verteilungsfrei und um den Erwartungswert $\mathrm{E}_0(T_n) = n(n+1)/4$ symmetrisch ist. Für die Testgröße T_n ist die alternativen Darstellung

$$T_n = \sum_{1 \leq i \leq j \leq n} \mathrm{I}(\zeta_i + \zeta_j > 0)$$

bekannt. Beim vorgeschlagenen zweiseitigen Wilcoxon-Vorzeichen-Rang Test handelt es sich hier um einen konsistenten Test zur Alternative K, da $P_\vartheta(\zeta_1 + \zeta_2 > 0) \neq 1/2$ für $\vartheta \neq 0$ gilt. Siehe 5.7 in [9].

10.3 Bahadur-Steigung im verbundenen Stichprobenfall

Die Bahadur-Steigung lässt sich mit Hilfe eines Resultats, welches im Buch von Nikitin ([29]) genannt wird, lokal beschreiben. Hierzu sei

$$W_n := \frac{T_n - n(n+1)/4}{\sqrt{n(n+1)(2n+1)/24}}.$$

10.3.1 Satz: *Für die Bahadur-Steigung gilt im verbundenen Stichprobenfall*

$$c_T(\vartheta) = \frac{3}{2}\big(P_0(\zeta_1 + \zeta_2 > 2\vartheta) - 1/2\big)^2\big(1 + o(1)\big) \text{ für } \vartheta \to 0.$$

Beweis: Es ist der auf der Testgröße T_n basierende Test äquivalent zum Test

$$\text{„Verwirf H, falls } W_n < -\frac{c_n - n(n+1)/4}{\sqrt{n(n+1)(2n+1)/24}} \text{ oder } W_n > \frac{c_n - n(n+1)/4}{\sqrt{n(n+1)(2n+1)/24}}\text{“},$$

bzw. äquivalent zum Test

$$\text{„Verwirf H, falls } |W_n| > \frac{c_n - n(n+1)/4}{\sqrt{n(n+1)(2n+1)/24}}\text{“}.$$

Anhand der alternativen Darstellung der Testgröße T_n erkennt man ferner, dass sich T_n schreiben lässt als

$$T_n := \binom{n}{2} U_n + \sum_{i=1}^{n} \mathrm{I}(\zeta_i > 0),$$

mit $U_n := \binom{n}{2}^{-1} \sum_{1\le i<j\le n} \mathrm{I}(\zeta_i + \zeta_j > 0)$. Da U_n eine U-Statistik ist, gilt im Fall $\vartheta \neq 0$ mit 5.4 Theorem A aus [36]

$$\frac{W_n}{\sqrt{n}} = \frac{U_n - 1/2}{\sqrt{1/3}} + o_P(1) \xrightarrow{P_\vartheta} \sqrt{3}\big(P_\vartheta(\zeta_1 + \zeta_2 > 0) - 1/2\big) \text{ für } n \to \infty.$$

Es sei $J(u) := \sqrt{12}u,\ u \in [0,1]$, und $J_0 := \int_0^1 J(u)\mathrm{d}u/2 = \sqrt{3}/2$. Weiter definieren wir

$$a_n(u) := \sqrt{12} \sum_{j=1}^{n} \frac{j}{n} \mathrm{I}\left(\frac{j-1}{n} < u \le \frac{j}{n}\right),\ u \in [0,1].$$

Dann gilt

$$S_n := \frac{\sqrt{12}}{n^2} \sum_{i=1}^{n} R_i^+ \,\mathrm{I}(\zeta_i > 0) - J_0 = \frac{1}{n} \sum_{i=1}^{n} a_n\Big(\frac{R_i^+}{n+1}\Big) \mathrm{I}(\zeta_i > 0) - J_0$$

und auch $\lim_{n\to\infty} a_n = J$ in $L^2([0,1], \mathfrak{B}_{[0,1]}, \lambda_{[0,1]})$ sowie $\int_0^1 J(u)^2 \mathrm{d}u/4 = 1$. Somit gilt mit Theorem 4.2.3 aus [29]

$$\lim_{n\to\infty} \frac{1}{n} \log G_n(t) = -h(t) = -\frac{1}{2} t^2 \big(1 + o(1)\big) \text{ für } t \downarrow 0,$$

mit $G_n(t) := P_0(S_n \ge t),\ t \in \mathbb{R}$, und einer Funktion $h : \mathbb{R} \to \mathbb{R}$, die in $t > 0$ genügend klein stetig ist. Nun sei

$$g_n(t) := \frac{\sqrt{12}}{n^2} t - J_0,\ h_n(t) := \frac{t - n(n+1)/4}{\sqrt{n(n+1)(2n+1)/24}},\ t \in \mathbb{R}.$$

Dann gilt

$$\begin{aligned} B_n &:= g_n\big(h_n^{-1}(|W_n|)\big) = \sqrt{12} \frac{\sqrt{n^2(n+1)(2n+1)/24}}{n^2} \frac{|W_n|}{\sqrt{n}} + \sqrt{12} \frac{n(n+1)/4}{n^2} - J_0 \\ &\xrightarrow{P_\vartheta} \sqrt{3} |P_\vartheta(\zeta_1 + \zeta_2 > 0) - 1/2| \text{ für } n \to \infty, \end{aligned}$$

und deshalb mit Lemma 2.3.3

$$\lim_{n\to\infty} \frac{1}{n} \log G_n(B_n) = -\frac{3}{2} \big(P_\vartheta(\zeta_1 + \zeta_2 > 0) - 1/2\big)^2 \big(1 + o(1)\big) \text{ in } P_\vartheta\text{-Wahrscheinlichkeit},$$

wobei $o(1)$ einen Term bezeichne, der für $\vartheta \to 0$ gegen $0 \in \mathbb{R}$ konvergiere. Da nach Konstruktion $S_n = g_n(T_n)$ und $W_n = h_n(T_n)$ gilt, und damit $S_n = g_n(h_n^{-1}(W_n))$, folgt insgesamt mit $H_n(t) := P_0(|W_n| \geq t),\ t \in \mathbb{R}$, unter Ausnutzung der Symmetrie der Verteilung von W_n bei Gültigkeit der Hypothese

$$\begin{aligned}
&-\frac{3}{2}\big(P_\vartheta(\zeta_1 + \zeta_2 > 0) - 1/2\big)^2\big(1 + o(1)\big) = \lim_{n\to\infty} \frac{1}{n} \log G_n(B_n) \\
&= \lim_{n\to\infty} \frac{1}{n} \log\left(\frac{1}{2} H_n\Big(h_n\big(g_n^{-1}(B_n)\big)\Big)\right) = \lim_{n\to\infty} \frac{1}{n} \log H_n(|W_n|) \text{ in } P_\vartheta\text{-Wahrscheinlichkeit,}
\end{aligned}$$

wobei $o(1)$ einen Term bezeichne, der für $\vartheta \to 0$ gegen $0 \in \mathbb{R}$ konvergiere. Damit folgt die Behauptung. $\square$

10.4 Unabhängiger Stichprobenfall

Im unabhängigen Stichprobenfall liegen mit $X_1', \ldots, X_n'$ und $Y_1', \ldots, Y_n'$ zwei voneinander unabhängige Stichproben vor. Wir wählen die Mann-Whitney Form der Wilcoxon-Mann-Whitney Testgröße, konkret

$$T_n' := \sum_{i=1}^{n} \sum_{j=1}^{n} \mathrm{I}(X_i' > Y_j'),$$

und als Test für die Hypothese H gegen die Alternative K schlagen wir für $\alpha \in (0,1)$ den zweiseitigen Wilcoxon-Mann-Whitney Test

$$\text{„Verwirf H, falls } T_n' < n^2 - c_n' \text{ oder } T_n' > c_n'\text{“}$$

vor, wobei c_n' ein $(1 - \alpha/2)$-Quantil der Verteilung von T_n' bei Gültigkeit der Hypothese bezeichne. Es ist für jedes $n \in \mathbb{N}$ $c_n' \in \mathbb{N}_0$ und $0 \leq c_n' \leq n^2$. Man beachte auch hier, dass die Verteilung von T_n' bei Gültigkeit der Hypothese verteilungsfrei und um den Erwartungswert $\mathrm{E}_0(T_n') = n^2/2$ symmetrisch ist. Der vorgeschlagene zweiseitige Wilcoxon-Mann-Whitney Test ist hier ein konsistenter Test zur Alternative K, da $P_\vartheta(X_1' > Y_1') \neq 1/2$ für $\vartheta \neq 0$ gilt. Siehe 6.6 in [9]. Für die Testgröße ist auch die alternative Darstellung

$$T_n' := \sum_{i=1}^{n} R_i - \frac{n(n+1)}{2}$$

bekannt. Hierbei stehen die $R_1, \ldots, R_n$ für die Ränge der $X_1', \ldots, X_n'$ in der zusammengelegten Stichprobe $X_1', \ldots, X_n', Y_1', \ldots, Y_n'$.

Es bezeichne im Folgenden $\zeta_i' := X_i' - Y_i'$, $1 \leq i \leq n$. Wegen $\zeta_1' \overset{\mathrm{v}}{=} X_1 - Y_2$ wollen wir manchmal auch $\zeta_1' {=} X_1 - Y_2$ schreiben, wenn es nur auf die Verteilung von ζ_1' ankommt.

10.5 Bahadur-Steigung im unabhängigen Stichprobenfall

Auch hier lässt sich die Bahadur-Steigung mit Hilfe eines Resultats, welches im Buch von Nikitin ([29]) genannt wird, lokal beschreiben. Hierfür sei

$$W_n' := \frac{T_n' - n^2/2}{\sqrt{n^2(2n+1)/12}}.$$

10.5.1 Satz: *Für die Bahadur-Steigung gilt im unabhängigen Stichprobenfall*

$$c_{T'}(\vartheta) = 3\big(P_0(X_1' - Y_1' > \vartheta) - 1/2\big)^2\big(1 + o(1)\big) \textit{ für } \vartheta \to 0.$$

Beweis: Es ist der auf der Testgröße T_n basierende Test äquivalent zm Test

$$\text{„Verwirf H, falls } W_n' < -\frac{c_n' - n^2/2}{\sqrt{n^2(2n+1)/12}} \text{ oder } W_n' > \frac{c_n' - n^2/2}{\sqrt{n^2(2n+1)/12}}\text{“},$$

bzw. äquivalent zum Test

$$\text{„Verwirf H, falls } |W_n'| > \frac{c_n' - n^2/2}{\sqrt{n^2(2n+1)/12}}\text{“}.$$

T_n' lässt sich schreiben als

$$T_n' := \binom{n}{2} U_n' + \sum_{i=1}^{n} \mathrm{I}(X_i' > Y_i'),$$

mit $U_n' := \binom{n}{2}^{-1} \sum_{1 \le i < j \le n} (\mathrm{I}(X_i' > Y_j') + \mathrm{I}(X_j' > Y_i'))$. Da U_n' eine U-Statsitik ist, gilt im Fall $\vartheta \neq 0$ mit 5.4 Theorem A aus [36]

$$\frac{W_n'}{\sqrt{n}} = \frac{U_n' - 1}{\sqrt{2/3}} + o_P(1) \xrightarrow{P_\vartheta} \sqrt{6}\big(P_\vartheta(X_1' > Y_1') - 1/2\big) \text{ für } n \to \infty.$$

Es sei $J'(u) := \sqrt{12}(u - 1/2),\ u \in [0,1]$, und $N := 2n$. Weiter definieren wir

$$a_N'(u) := \sqrt{12} \sum_{j=1}^{N} \frac{j}{N} \mathrm{I}\left(\frac{j-1}{N} < u \le \frac{j}{N}\right) - \sqrt{3},\ u \in [0,1].$$

Dann gilt

$$S_N' := \frac{\sqrt{12}}{N^2} \sum_{i=1}^{n} R_i - \frac{\sqrt{3}}{2} = \frac{1}{N} \sum_{i=1}^{n} a_N'\Big(\frac{R_i}{N+1}\Big),$$

und auch $\lim_{N\to\infty} a'_N = J'$ in $L^2([0,1],\mathfrak{B}_{[0,1]},\lambda_{[0,1]})$ sowie $\int_0^1 J'(u)\mathrm{d}u = 0$ und $\int_0^1 J'(u)^2\mathrm{d}u = 1$. Somit gilt mit Theorem 3.1.2 aus [29] (siehe auch die Theoreme 1.6.11 und 1.6.12 in [29])

$$\lim_{N\to\infty} \frac{1}{N} \log G'_N(t) = -h'(t) = -2t^2\big(1+o(1)\big) \text{ für } t \downarrow 0,$$

mit $G'_N(t) := P_0(S'_N \geq t),\ t \in \mathbb{R}$, und einer Funktion $h' : \mathbb{R} \to \mathbb{R}$, die in $t > 0$ genügend klein stetig ist. Nun sei

$$g'_n(t) := \frac{\sqrt{3}}{2n^2}\Big(\frac{n(n+1)}{2} + t\Big) - \frac{\sqrt{3}}{2},\ h'_n(t) := \frac{t - n^2/2}{\sqrt{n^2(2n+1)/12}},\ t \in \mathbb{R}.$$

Dann gilt

$$\begin{aligned} B'_n := g'_n\big({h'}_n^{-1}(|W'_n|)\big) &= \frac{\sqrt{12}(2n^2+n)}{8n^2} + \frac{\sqrt{12}\sqrt{n^3(2n+1)/12}}{4n^2}\frac{|W'_n|}{\sqrt{n}} - \frac{\sqrt{3}}{2} \\ &\xrightarrow{P_\vartheta} \frac{\sqrt{3}}{2}|P_\vartheta(X'_1 > Y'_1) - 1/2| \text{ für } n \to \infty, \end{aligned}$$

und deshalb mit Lemma 2.3.3

$$\lim_{N\to\infty} \frac{1}{N} \log G'_N(B'_n) = -\frac{3}{2}\big(P_\vartheta(X'_1 > Y'_1) - 1/2\big)^2\big(1+o(1)\big) \text{ in } P_\vartheta\text{-Wahrscheinlichkeit},$$

wobei $o(1)$ einen Term bezeichne, der für $\vartheta \to 0$ gegen $0 \in \mathbb{R}$ konvergiere. Da nach Konstruktion $S'_N = g'_n(T'_n)$ und $W'_n = h'_n(T'_n)$ gilt, und damit $S'_N = g'_n({h'}_n^{-1}(W'_n))$, folgt insgesamt mit $H'_n(t) := P_0(|W'_n| \geq t),\ t \in \mathbb{R}$, unter Ausnutzung der Symmetrie der Verteilung von W'_n bei Gültigkeit der Hypothese

$$\begin{aligned} &-3\big(P_\vartheta(X'_1 > Y'_1) - 1/2\big)^2\big(1+o(1)\big) = \lim_{n\to\infty} \frac{1}{n} \log G'_N(B'_n) \\ &= \lim_{n\to\infty} \frac{1}{n} \log\left(\frac{1}{2}H'_n\Big(h'_n\big({g'}_n^{-1}(B'_n)\big)\Big)\right) = \lim_{n\to\infty} \frac{1}{n} \log H'_n(|W'_n|) \text{ in } P_\vartheta\text{-Wahrscheinlichkeit}, \end{aligned}$$

wobei $o(1)$ einen Term bezeichne, der für $\vartheta \to 0$ gegen $0 \in \mathbb{R}$ konvergiere. Damit folgt die Behauptung. □

10.6 Volumen-Effizienz der Tests

Es sei $\zeta'_i = X'_i - Y'_i$, $i = 1, \ldots, n$. Im unabhängigen Stichprobenfall besitzt $\sum_{i=1}^n \sum_{j=1}^n \mathrm{I}(X'_i > Y'_j - \vartheta)$ die gleiche Verteilung wie T'_n bei Gültigkeit der Hypothese. Damit gilt

$$P_\vartheta\bigg(n^2 - c'_n \leq \sum_{i=1}^n \sum_{j=1}^n \mathrm{I}(X'_i > Y'_j - \vartheta) \leq c'_n\bigg) \geq 1 - \alpha.$$

Somit ist

$$\left\{\vartheta \in \mathbb{R}; n^2 - c_n' \leq \sum_{i=1}^n \sum_{j=1}^n \mathrm{I}(X_i' > Y_j' - \vartheta) \leq c_n'\right\}$$

ein Konfidenzbereich zum Konfidenzniveau $1-\alpha$ für den unbekannten Parameter $\vartheta \in \mathbb{R}$, mit dessen Hilfe sich der Wilcoxon-Mann-Whitney Test durchführen lässt. Es seien $V_{(1)} \leq \cdots \leq V_{(n^2)}$ die nach wachsender Größe geordneten Beobachtungen $-(X_i' - Y_j')$, $1 \leq i, j \leq n$, deren Erwartungswert existiere und endlich sei. Wegen

$$n^2 - c_n' \leq \sum_{i=1}^n \sum_{j=1}^n \mathrm{I}(X_i' > Y_j' - \vartheta) \leq c_n' \iff V_{(n^2 - c_n')} < \vartheta \leq V_{(c_n'+1)}$$

ist $(V_{(n^2-c_n')}, V_{(c_n'+1)}]$ ein Konfidenzintervall zum Konfidenzniveau $1-\alpha$ für ϑ.

10.6.1 Bemerkung: Der Erwartungswert der zugehörigen Länge des Intervalls hängt nicht mehr vom zugrunde liegenden unbekannten Parameter $\vartheta \in \mathbb{R}$ ab. Schließlich ist

$$\begin{aligned} P_\vartheta(V_{(c_n'+1)} - \vartheta \geq t) &= P_\vartheta\left(\sum_{i=1}^n \sum_{j=1}^n \mathrm{I}(X_i' > Y_j' - \vartheta - t) \leq c_n'\right) \\ &= P_0\left(\sum_{i=1}^n \sum_{j=1}^n \mathrm{I}(X_i' > Y_j' - t) \leq c_n'\right), \ t \in \mathbb{R}, \end{aligned}$$

und genauso

$$\begin{aligned} P_\vartheta(V_{(n^2-c_n')} - \vartheta < t) &= P_\vartheta\left(\sum_{i=1}^n \sum_{j=1}^n \mathrm{I}(X_i' > Y_j' - \vartheta - t) \geq n^2 - c_n'\right) \\ &= P_0\left(\sum_{i=1}^n \sum_{j=1}^n \mathrm{I}(X_i' > Y_j' - t) \geq n^2 - c_n'\right), \ t \in \mathbb{R}, \end{aligned}$$

und somit wegen $\mathrm{E}_\vartheta(V_{(c_n'+1)} - V_{(n^2-c_n')}) = \mathrm{E}_\vartheta(V_{(c_n'+1)} - \vartheta) - \mathrm{E}_\vartheta(V_{(n^2-c_n')} - \vartheta)$ keine Abhängigkeit mehr von $\vartheta \in \mathbb{R}$ vorhanden.

Im verbundenen Stichprobenfall besitzt $\sum_{1 \leq i \leq j \leq n} \mathrm{I}(\zeta_i + \zeta_j + 2\vartheta > 0)$ die gleiche Verteilung wie T_n bei Gültigkeit der Hypothese. Also gilt

$$P_\vartheta\left(\frac{1}{2}n(n+1) - c_n \leq \sum_{1 \leq i \leq j \leq n} \mathrm{I}(\zeta_i + \zeta_j + 2\vartheta > 0) \leq c_n\right) \geq 1 - \alpha.$$

Damit ist

$$\left\{\frac{1}{2}n(n+1) - c_n \leq \sum_{1 \leq i \leq j \leq n} \mathrm{I}(\zeta_i + \zeta_j + 2\vartheta > 0) \leq c_n\right\}$$

ein Konfidenzbereich zum Konfidenzniveau $1-\alpha$ für den unbekannten Parameter $\vartheta \in \mathbb{R}$, mit dessen Hilfe sich der Wilcoxon-Vorzeichen-Rang Test durchführen lässt. Es seien $W_{(1)} \leq \cdots \leq W_{(n(n+1)/2)}$ die nach wachsender Größe geordneten Beobachtungen $-(\zeta_i + \zeta_j)/2$, $1 \leq i \leq j \leq n$, deren Erwartungswert existiere und endlich sei. Wegen

$$\frac{1}{2}n(n+1) - c_n \leq \sum_{1 \leq i \leq j \leq n} \mathrm{I}(\zeta_i + \zeta_j + 2\vartheta > 0) \leq c_n \;\Leftrightarrow\; W_{(n(n+1)/2-c_n)} < \vartheta \leq W_{(c_n+1)}$$

ist $(W_{(n(n+1)/2-c'_n)}, W_{(c'_n+1)}]$ ein Konfidenzintervall zum Konfidenzniveau $1-\alpha$ für ϑ.

10.6.2 Bemerkung: Analog zum unabhängigem Stichprobenfall erkennt man auch hier, dass der Erwartungswert der zugehörigen Länge des Intervalls nicht mehr vom unbekannten Parameter $\vartheta \in \mathbb{R}$ abhängt. Somit hängt auch die Volumen-Effizienz der Testfolge T relativ zur Testfolge T',

$$e^{\mathrm{vol}}_{T',T,n} = \left(\frac{\mathrm{E}(W_{(c'_n+1)} - W_{(n(n+1)/2-c'_n)})}{\mathrm{E}(V_{(c_n+1)} - V_{(n^2-c_n)})} \right)^2,$$

nicht mehr vom unbekannten Parameter $\vartheta \in \mathbb{R}$ ab.

Um die asymptotische Volumen-Effizienz der Testfolge T relativ zur Testfolge T' zu bestimmen, benötigen wir ein paar Vorbereitungen. Wir definieren

$$\begin{aligned}\sigma'_n(t)^2 :=& n^3\, \mathrm{E}_0 \Bigg(\bigg(\mathrm{E}_0 \Big(\mathrm{I}\big(t/\sqrt{n} > -(x_1 - Y'_2)\big) + \mathrm{I}\big(t/\sqrt{n} > -(X'_2 - y_1)\big)\Big) \\ &- \mathrm{E}_0 \Big(\mathrm{I}\big(t/\sqrt{n} > -(X'_1 - Y'_2)\big) + \mathrm{I}\big(t/\sqrt{n} > -(X'_2 - Y'_1)\big)\Big)\bigg)^2_{\big|(x_1,y_1)=(X'_1,Y'_1)} \Bigg), \; t \in \mathbb{R}.\end{aligned}$$

10.6.3 Lemma: *Im unabhängigen Stichprobenfall sei die Funktion $\tilde{F}: t \mapsto P_0(\zeta'_1 \leq t)$ auf einer offenen Umgebung U mit $0 \in U$ zweimal stetig differenzierbar mit $\tilde{f}(t) := \frac{\mathrm{d}}{\mathrm{d}t}\tilde{F}(t)$, $t \in U$, als Ableitung. Dann gilt*

$$\lim_{n\to\infty} \frac{\sigma'_n(t)^2}{n^3} = \frac{1}{6}, \; t \in \mathbb{R},$$

sowie

$$\lim_{n\to\infty} \frac{c'_n - 2\binom{n}{2}P_0\big(t/\sqrt{n} > -(X'_1 - Y'_1)\big)}{\sigma'_n(t)} = u_{1-\frac{\alpha}{2}} - \sqrt{6}\tilde{f}(0)t, \; t \in \mathbb{R},$$

und auch

$$\lim_{n\to\infty} \frac{n^2 - c'_n - 2\binom{n}{2}P_0\big(t/\sqrt{n} > -(X'_1 - Y'_1)\big)}{\sigma'_n(t)} = -u_{1-\frac{\alpha}{2}} - \sqrt{6}\tilde{f}(0)t, \; t \in \mathbb{R}.$$

Beweis: Es ist

$$1-\frac{\alpha}{2}=\lim_{n\to\infty}P_0(T'_n\leq c'_n)=\lim_{n\to\infty}P_0\left(W'_n\leq\frac{c'_n-n^2/2}{\sqrt{n^2(2n+1)/12}}\right),$$

und da wie bereits im Beweis von Satz 10.5.1 erwähnt $W'_n\xrightarrow{\text{v}}\mathrm{N}(0,1)$ für $n\to\infty$ gilt, und daher insbesondere die Quantilfunktion der Grenzverteilung streng monoton wachsend ist, folgt

$$\lim_{n\to\infty}\frac{c'_n-n^2/2}{\sqrt{n^2(2n+1)/12}}=u_{1-\frac{\alpha}{2}}.$$

Ferner folgt aufgrund der Stetigkeit von F unter $\vartheta=0$

$$\begin{aligned}&\lim_{n\to\infty}F(X'_1+t/\sqrt{n})=F(X'_1)\sim\mathfrak{R}(0,1),\\&\lim_{n\to\infty}F(Y'_1-t/\sqrt{n})=F(Y'_1)\sim\mathfrak{R}(0,1),\ t\in\mathbb{R},\end{aligned}$$

und deshalb und aufgrund der Stetigkeit von $\tilde{F}$ in $0\in\mathbb{R}$ folgt mit dem Satz von der majorisierten Konvergenz

$$\begin{aligned}\lim_{n\to\infty}\frac{\sigma'_n(t)^2}{n^3}=&\lim_{n\to\infty}\mathrm{E}_0\Big(F(X'_1+t/\sqrt{n})^2+2F(X'_1+t/\sqrt{n})\big(1-F(Y'_1-t/\sqrt{n})\big)\\&+\big(1-F(Y'_1-t/\sqrt{n})\big)^2-4F(X'_1+t/\sqrt{n})\big(1-\tilde{F}(-t/\sqrt{n})\big)\\&-4\big(1-F(Y'_1-t/\sqrt{n})\big)\big(1-\tilde{F}(t/\sqrt{n})\big)+4\big(1-\tilde{F}(-t/\sqrt{n})\big)^2\Big)\\=&\mathrm{E}_0\Big(F(X'_1)^2+2F(X'_1)\big(1-F(Y'_1)\big)+\big(1-F(Y'_1)\big)^2\\&-4F(X'_1)\big(1-\tilde{F}(0)\big)-4\big(1-F(Y'_1)\big)\big(1-\tilde{F}(0)\big)+4\big(1-\tilde{F}(0)\big)^2\Big)\\=&\mathrm{E}_0\big(F(X'_1)^2+F(Y'_1)^2-2F(X'_1)F(Y'_1)\big)\\=&\mathrm{E}_0\big(F(X'_1)^2\big)+\mathrm{E}_0\big(F(Y'_1)^2\big)-2\,\mathrm{E}_0\big(F(X'_1)\big)\,\mathrm{E}_0\big(F(Y'_1)\big)=\frac{1}{6},\ t\in\mathbb{R}.\end{aligned}$$

Insgesamt folgt also

$$\lim_{n\to\infty}\frac{c'_n-n^2/2}{\sqrt{n^2(2n+1)/12}}\frac{\sqrt{n^2(2n+1)/12}}{\sigma'_n(t)}=u_{1-\frac{\alpha}{2}},\ t\in\mathbb{R},$$

und mit einer Taylorentwicklung von $\tilde{F}$ in $0\in\mathbb{R}$

$$\begin{aligned}&\frac{n^2/2-2\binom{n}{2}P_0\big(t/\sqrt{n}>-(X'_1-Y'_1)\big)}{\sigma'_n(t)}\sim\frac{\sqrt{n}\Big(1/2-\big(1-\tilde{F}(-t/\sqrt{n})\big)\Big)}{\sqrt{1/6}}\\=&\frac{\sqrt{n}\Big(1/2-\big(1-\tilde{F}(0)+\tilde{f}(0)t/\sqrt{n}+o(1/\sqrt{n})\big)\Big)}{\sqrt{1/6}}\longrightarrow-\sqrt{6}\tilde{f}(0)t\text{ für }n\to\infty,\ t\in\mathbb{R}.\end{aligned}$$

Also erhält man

$$\begin{aligned}
&\lim_{n\to\infty} \frac{c_n' - 2\binom{n}{2}P_0\big(t/\sqrt{n} > -(X_1' - Y_1')\big)}{\sigma_n'(t)} \\
=&\lim_{n\to\infty} \Big(\frac{c_n' - n^2/2}{\sqrt{n^2(2n+1)/12}} \frac{\sqrt{n^2(2n+1)/12}}{\sigma_n'(t)} + \frac{n^2/2 - 2\binom{n}{2}P_0\big(t/\sqrt{n} > -(X_1' - Y_1')\big)}{\sigma_n'(t)} \Big) \\
=&u_{1-\frac{\alpha}{2}} - \sqrt{6}\tilde{f}(0)t, \ t \in \mathbb{R},
\end{aligned}$$

und auch

$$\begin{aligned}
&\lim_{n\to\infty} \frac{n^2 - c_n' - 2\binom{n}{2}P_0\big(t/\sqrt{n} > -(X_1' - Y_1')\big)}{\sigma_n'(t)} \\
=&\lim_{n\to\infty} \Big(- \frac{c_n' - n^2/2}{\sqrt{n^2(2n+1)/12}} \frac{\sqrt{n^2(2n+1)/12}}{\sigma_n'(t)} + \frac{n^2/2 - 2\binom{n}{2}P_0\big(t/\sqrt{n} > -(X_1' - Y_1')\big)}{\sigma_n'(t)} \Big) \\
=& - u_{1-\frac{\alpha}{2}} - \sqrt{6}\tilde{f}(0)t, \ t \in \mathbb{R},
\end{aligned}$$

und so die Behauptung. □

Wir definieren

$$\begin{aligned}
&\sigma_n(t)^2 \\
:=&n^3 \operatorname{E}_0 \left(\left(\operatorname{E}_0 \left(\operatorname{I}\left(2t/\sqrt{n} > -(z_1 + \zeta_2)\right)\right) - \operatorname{E}_0 \left(\operatorname{I}\left(2t/\sqrt{n} > -(\zeta_1 + \zeta_2)\right)\right)\right)^2_{|z_1=\zeta_1} \right), \ t \in \mathbb{R}.
\end{aligned}$$

10.6.4 Lemma: *Im verbundenen Stichprobenfall sei die Funktion $\tilde{G} : t \mapsto P_0(\zeta_1 + \zeta_2 \le t)$ auf einer offenen Umgebung U' mit $0 \in U'$ zweimal stetig differenzierbar mit $\tilde{g}(t) := \frac{\mathrm{d}}{\mathrm{d}t}\tilde{G}(t)$, $t \in U'$, als Ableitung. Dann gilt*

$$\lim_{n\to\infty} \frac{\sigma_n(t)^2}{n^3} = \frac{1}{12}, \ t \in \mathbb{R},$$

sowie

$$\lim_{n\to\infty} \frac{c_n - \binom{n}{2}P_0\big(2t/\sqrt{n} > -(\zeta_1 + \zeta_2)\big)}{\sigma_n(t)} = u_{1-\frac{\alpha}{2}} - \sqrt{12}\tilde{g}(0)t, \ t \in \mathbb{R},$$

und auch

$$\lim_{n\to\infty} \frac{n(n+1)/2 - c_n - \binom{n}{2}P_0\big(t/\sqrt{n} > -(\zeta_1 + \zeta_2)\big)}{\sigma_n(t)} = -u_{1-\frac{\alpha}{2}} - \sqrt{12}\tilde{g}(0)t, \ t \in \mathbb{R}.$$

Beweis: Es ist

$$1 - \frac{\alpha}{2} = \lim_{n\to\infty} P_0(T_n \le c_n) = \lim_{n\to\infty} P_0\bigg(W_n \le \frac{c_n - n(n+1)/4}{\sqrt{n(n+1)(2n+1)/24}}\bigg),$$

und da wie bereits im Beweis von Satz 10.3.1 erwähnt $W_n \xrightarrow{\mathrm{v}} \mathrm{N}(0,1)$ für $n \to \infty$ gilt, und daher insbesondere die Quantilfunktion der Grenzverteilung stetig ist, folgt

$$\lim_{n\to\infty} \frac{c_n - n(n+1)/4}{\sqrt{n(n+1)(2n+1)/24}} = u_{1-\frac{\alpha}{2}}.$$

Ferner folgt aufgrund der Stetigkeit von G unter $\vartheta = 0$

$$\lim_{n\to\infty} G(\zeta_1 + 2t/\sqrt{n}) = G(\zeta_1) \sim \mathfrak{R}(0,1),\ t \in \mathbb{R},$$

und deshalb und aufgrund der Stetigkeit von $\tilde{G}$ in $0 \in \mathbb{R}$ folgt mit dem Satz von der majorisierten Konvergenz

$$\begin{aligned}
\lim_{n\to\infty} \frac{\sigma_n(t)^2}{n^3} &= \lim_{n\to\infty} \mathrm{E}_0 \Big(\big(1 - G(-\zeta_1 - 2t/\sqrt{n})\big)^2 \\
&\quad - 2\big(1 - G(-\zeta_1 - 2t/\sqrt{n})\big)\big(1 - \tilde{G}(-2t/\sqrt{n})\big) + \big(1 - \tilde{G}(-2t/\sqrt{n})\big)^2 \Big) \\
&= \lim_{n\to\infty} \mathrm{E}_0 \left(G(\zeta_1 + 2t/\sqrt{n})^2 - 2G(\zeta_1 + 2t/\sqrt{n})\tilde{G}(2t/\sqrt{n}) + \tilde{G}(2t/\sqrt{n})^2 \right) \\
&= \mathrm{E}_0 \left(G(\zeta_1)^2 - 2G(\zeta_1)\tilde{G}(0) + \tilde{G}(0)^2 \right) = \mathrm{E}_0 \left(G(\zeta_1)^2 - G(\zeta_1) + \frac{1}{4} \right) \\
&= \mathrm{E}_0 \left(G(\zeta_1)^2 \right) - \mathrm{E}_0 \left(G(\zeta_1) \right) + \frac{1}{4} = \frac{1}{12},\ t \in \mathbb{R}.
\end{aligned}$$

Insgesamt folgt also

$$\lim_{n\to\infty} \frac{c_n - n(n+1)/4}{\sqrt{n(n+1)(2n+1)/24}} \frac{\sqrt{n(n+1)(2n+1)/24}}{\sigma_n(t)} = u_{1-\frac{\alpha}{2}},\ t \in \mathbb{R},$$

und mit einer Taylorentwicklung von $\tilde{G}$ in $0 \in \mathbb{R}$

$$\begin{aligned}
&\frac{n(n+1)/4 - \binom{n}{2} P_0\big(2t/\sqrt{n} > -(\zeta_1 + \zeta_2)\big)}{\sigma_n(t)} \sim \frac{1}{2} \frac{\sqrt{n}\Big(1/2 - \big(1 - \tilde{G}(-2t/\sqrt{n})\big)\Big)}{\sqrt{1/12}} \\
=& \frac{1}{2} \frac{\sqrt{n}\Big(1/2 - \big(1 - \tilde{G}(0) + \tilde{g}(0)2t/\sqrt{n} + o(1/\sqrt{n})\big)\Big)}{\sqrt{1/12}} \longrightarrow -\sqrt{12}\tilde{g}(0)t \text{ für } n \to \infty,\ t \in \mathbb{R}.
\end{aligned}$$

Also erhält man

$$\begin{aligned}
&\lim_{n\to\infty} \frac{c_n - \binom{n}{2} P_0\big(2t/\sqrt{n} > -(\zeta_1 + \zeta_2)\big)}{\sigma_n(t)} \\
&= \lim_{n\to\infty} \Big(\frac{c_n - n(n+1)/4}{\sqrt{n(n+1)(2n+1)/24}} \frac{\sqrt{n(n+1)(2n+1)/24}}{\sigma_n(t)} \\
&\quad + \frac{n(n+1)/4 - \binom{n}{2} P_0\big(2t/\sqrt{n} > -(\zeta_1 + \zeta_2)\big)}{\sigma_n(t)} \Big) = u_{1-\frac{\alpha}{2}} - \sqrt{12}\tilde{g}(0)t,\ t \in \mathbb{R},
\end{aligned}$$

und auch

$$
\begin{aligned}
&\lim_{n\to\infty} \frac{n(n+1)/2 - c_n - \binom{n}{2} P_0\big(2t/\sqrt{n} > -(\zeta_1+\zeta_2)\big)}{\sigma_n(t)} \\
&= \lim_{n\to\infty} \Big(- \frac{c_n - n(n+1)/4}{\sqrt{n(n+1)(2n+1)/24}} \frac{\sqrt{n(n+1)(2n+1)/24}}{\sigma_n(t)} \\
&\quad + \frac{n(n+1)/4 - \binom{n}{2} P_0\big(2t/\sqrt{n} > -(\zeta_1+\zeta_2)\big)}{\sigma_n(t)} \Big) = -u_{1-\frac{\alpha}{2}} - \sqrt{12}\tilde{g}(0)t, \ t \in \mathbb{R},
\end{aligned}
$$

und so die Behauptung. □

10.6.5 Satz: *Es seien die Funktionen $\tilde{F}$ und $\tilde{G}$ auf offenenen Umgebungen U mit $0 \in U$ und U' mit $0 \in U'$ zweimal stetig differenzierbar mit $\tilde{f}(t) = \frac{\mathrm{d}}{\mathrm{d}t}\tilde{F}(t)$, $t \in U$, und $\tilde{g}(t) := \frac{\mathrm{d}}{\mathrm{d}t}\tilde{G}(t)$, $t \in U'$, als Ableitungen. Ferner gelte $\tilde{f}(0) \neq 0$ und $\tilde{g}(0) \neq 0$. Dann gilt*

$$
\begin{aligned}
\sqrt{n}(V_{(c'_n+1)} - \vartheta) &\xrightarrow{\mathrm{v}} \mathrm{N}\Big(\frac{u_{1-\frac{\alpha}{2}}}{\sqrt{6}\tilde{f}(0)}, \frac{1}{6\tilde{f}(0)^2} \Big) \text{ für } n \to \infty, \\
\sqrt{n}(V_{(n^2-c'_n)} - \vartheta) &\xrightarrow{\mathrm{v}} \mathrm{N}\Big(-\frac{u_{1-\frac{\alpha}{2}}}{\sqrt{6}\tilde{f}(0)}, \frac{1}{6\tilde{f}(0)^2} \Big) \text{ für } n \to \infty, \\
\sqrt{n}(W_{(c_n+1)} - \vartheta) &\xrightarrow{\mathrm{v}} \mathrm{N}\Big(\frac{u_{1-\frac{\alpha}{2}}}{\sqrt{12}\tilde{g}(0)}, \frac{1}{12\tilde{g}(0)^2} \Big) \text{ für } n \to \infty, \\
\sqrt{n}(W_{(n(n+1)/2-c_n)} - \vartheta) &\xrightarrow{\mathrm{v}} \mathrm{N}\Big(-\frac{u_{1-\frac{\alpha}{2}}}{\sqrt{12}\tilde{g}(0)}, \frac{1}{12\tilde{g}(0)^2} \Big) \text{ für } n \to \infty.
\end{aligned}
$$

Beweis: Mit Lemma 10.6.3 gilt $\liminf_{n\to\infty} \sigma'_n(t)/n^{3/2} > 0$, $t \in \mathbb{R}$. Damit folgt mit einem zentralen Grenzwertsatz für U-Statistiken mit einem von n abhängigen Kern (siehe Korollar A.1.5 im Anhang), sowie mit Lemma 10.6.3, für $\vartheta = 0$

$$
\begin{aligned}
&\frac{\sum_{i=1}^n \sum_{j=1}^n \mathrm{I}\left(t/\sqrt{n} > -(X'_i - Y'_j)\right) - 2\binom{n}{2} P_0\big(t/\sqrt{n} > -(X'_1 - Y'_1)\big)}{\sigma'_n(t)} \\
=& \frac{1}{\sigma'_n(t)} \Bigg(\sum_{1\le i<j\le n} \Big(\mathrm{I}\left(t/\sqrt{n} > -(X'_i - Y'_j)\right) + \mathrm{I}\left(t/\sqrt{n} > -(X'_j - Y'_i)\right) \Big) \\
&- 2\binom{n}{2} P_0\big(t/\sqrt{n} > -(X'_1 - Y'_1)\big) \Bigg) + \frac{1}{\sigma'_n(t)} \sum_{i=1}^n \mathrm{I}\left(t/\sqrt{n} > -(X'_i - Y'_i)\right) \\
&\xrightarrow{\mathrm{v}} \mathrm{N}(0,1) \text{ für } n \to \infty, \ t \in \mathbb{R}.
\end{aligned}
$$

Da die Verteilungsfunktion dieser Zufallsvariablen aufgrund der Stetigkeit der Verteilungsfunktion der Normalverteilung gleichmäßig konvergiert, folgt mit Lemma 10.6.3

$$\lim_{n\to\infty} P_\vartheta\big(\sqrt{n}(V_{(c'_n+1)}-\vartheta)\geq t\big) = \lim_{n\to\infty} P_\vartheta\Bigg(\sum_{i=1}^n\sum_{j=1}^n \mathrm{I}(X'_i > Y'_j-\vartheta-t/\sqrt{n})\leq c'_n\Bigg)$$

$$=\lim_{n\to\infty} P_0\Bigg(\sum_{i=1}^n\sum_{j=1}^n \mathrm{I}\left(t/\sqrt{n} > -(X'_i-Y'_j)\right)\leq c'_n\Bigg)$$

$$=\lim_{n\to\infty} P_0\Bigg(\frac{\sum_{i=1}^n\sum_{j=1}^n \mathrm{I}\left(t/\sqrt{n} > -(X'_i-Y'_j)\right)-2\binom{n}{2}P_0\big(t/\sqrt{n}>-(X'_1-Y'_1)\big)}{\sigma'_n(t)}\leq$$

$$\frac{c'_n-2\binom{n}{2}P_0\big(t/\sqrt{n}>-(X'_1-Y'_1)\big)}{\sigma'_n(t)}\Bigg) = P\big(Z\leq u_{1-\frac{\alpha}{2}}-\sqrt{6}\tilde{f}(0)t\big)$$

$$=P\Bigg(\frac{-Z+u_{1-\frac{\alpha}{2}}}{\sqrt{6}\tilde{f}(0)}\geq t\Bigg),\ t\in\mathbb{R},$$

mit $Z\sim \mathrm{N}(0,1)$, und somit

$$\sqrt{n}(V_{(c'_n+1)}-\vartheta)\xrightarrow{\text{v}}\mathrm{N}\Bigg(\frac{u_{1-\frac{\alpha}{2}}}{\sqrt{6}\tilde{f}(0)},\frac{1}{6\tilde{f}(0)^2}\Bigg)\text{ für } n\to\infty.$$

Ebenso folgt mit Lemma 10.6.3

$$\lim_{n\to\infty} P_\vartheta\big(\sqrt{n}(V_{(n^2-c'_n)}-\vartheta)< t\big) = \lim_{n\to\infty} P_\vartheta\Bigg(\sum_{i=1}^n\sum_{j=1}^n \mathrm{I}(X'_i > Y'_j-\vartheta-t/\sqrt{n})\geq n^2-c'_n\Bigg)$$

$$=\lim_{n\to\infty} P_0\Bigg(\sum_{i=1}^n\sum_{j=1}^n \mathrm{I}\left(t/\sqrt{n} > -(X'_i-Y'_j)\right)\geq n^2-c'_n\Bigg)$$

$$=\lim_{n\to\infty} P_0\Bigg(\frac{\sum_{i=1}^n\sum_{j=1}^n \mathrm{I}\left(t/\sqrt{n} > -(X'_i-Y'_j)\right)-2\binom{n}{2}P_0\big(t/\sqrt{n}>-(X'_1-Y'_1)\big)}{\sigma'_n(t)}\geq$$

$$\frac{n^2-c'_n-2\binom{n}{2}P_0\big(t/\sqrt{n}>-(X'_1-Y'_1)\big)}{\sigma'_n(t)}\Bigg) = P\big(Z\geq -u_{1-\frac{\alpha}{2}}-\sqrt{6}\tilde{f}(0)t\big)$$

$$=P\Bigg(\frac{-Z-u_{1-\frac{\alpha}{2}}}{\sqrt{6}\tilde{f}(0)}< t\Bigg),\ t\in\mathbb{R},$$

und damit

$$\sqrt{n}(V_{(n^2-c'_n)}-\vartheta)\xrightarrow{\text{v}}\mathrm{N}\Bigg(-\frac{u_{1-\frac{\alpha}{2}}}{\sqrt{6}\tilde{f}(0)},\frac{1}{6\tilde{f}(0)^2}\Bigg)\text{ für } n\to\infty.$$

Analog folgt im verbundenen Stichprobenfall mit Lemma 10.6.4

$$\sqrt{n}(W_{(c_n+1)}-\vartheta)\xrightarrow{\text{v}}\mathrm{N}\Bigg(\frac{u_{1-\frac{\alpha}{2}}}{\sqrt{12}\tilde{g}(0)},\frac{1}{12\tilde{g}(0)^2}\Bigg)\text{ für } n\to\infty,$$

$$\sqrt{n}(W_{((n(n+1)/2-c_n)}-\vartheta)\xrightarrow{\text{v}}\mathrm{N}\Bigg(-\frac{u_{1-\frac{\alpha}{2}}}{\sqrt{12}\tilde{g}(0)},\frac{1}{12\tilde{g}(0)^2}\Bigg)\text{ für } n\to\infty,$$

womit alles gezeigt ist. $\square$

Nun definieren wir für feste aber beliebige $a, b \in \mathbb{R}$

$$\begin{aligned}\sigma_n'(s,t)^2 :=& n^3 \,\mathrm{E}_0 \Bigg(\Bigg(\mathrm{E}_0 \Big(a\,\mathrm{I}\left(s/\sqrt{n} > -(x_1 - Y_2')\right) + a\,\mathrm{I}\left(s/\sqrt{n} > -(X_2' - y_1)\right) \\ &+ b\,\mathrm{I}\left(t/\sqrt{n} > -(x_1 - Y_2')\right) + b\,\mathrm{I}\left(t/\sqrt{n} > -(X_2' - y_1)\right)\Big) \\ &- \mathrm{E}_0 \Big(a\,\mathrm{I}\left(s/\sqrt{n} > -(X_1' - Y_2')\right) + a\,\mathrm{I}\left(s/\sqrt{n} > -(X_2' - Y_1')\right) \\ &+ b\,\mathrm{I}\left(t/\sqrt{n} > -(X_1' - Y_2')\right) + b\,\mathrm{I}\left(t/\sqrt{n} > -(X_2' - Y_1')\right)\Big)\Bigg)^2_{|(x_1,y_1)=(X_1',Y_1')}\Bigg), \\ & s,t \in \mathbb{R}.\end{aligned}$$

10.6.6 Lemma: *Es gilt*

$$\lim_{n\to\infty} \frac{\sigma_n'(s,t)^2}{n^3} = \frac{(a+b)^2}{6}, \ \ s,t \in \mathbb{R}.$$

Beweis: Aufgrund der Stetigkeit von F gilt für $\vartheta = 0$

$$\begin{aligned}&\lim_{n\to\infty} F(X_1' + t/\sqrt{n}) = F(X_1') \sim \mathfrak{R}(0,1), \\ &\lim_{n\to\infty} F(Y_1' - t/\sqrt{n}) = F(Y_1') \sim \mathfrak{R}(0,1), \ \ t \in \mathbb{R},\end{aligned}$$

und deshalb und aufgrund der Stetigkeit von $\tilde{F}$ in $0 \in \mathbb{R}$ folgt mit den Überlegungen aus dem Beweis von Lemma 10.6.3 und dem Satz von der majorisierten Konvergenz

$$\begin{aligned}&\lim_{n\to\infty} \frac{\sigma_n'(s,t)^2}{n^3} \\ =& \lim_{n\to\infty} \mathrm{E}_0 \Bigg(\Bigg(a\Big(F(X_1' + s/\sqrt{n}) + 1 - F(Y_1' - s/\sqrt{n}) - 2\big(1 - \tilde{F}(-s/\sqrt{n})\big)\Big) \\ &+ b\Big(F(X_1' + t/\sqrt{n}) + 1 - F(Y_1' - t/\sqrt{n}) - 2\big(1 - \tilde{F}(-t/\sqrt{n})\big)\Big)\Bigg)^2\Bigg) \\ =& \frac{a^2+b^2}{6} + 2ab \lim_{n\to\infty} \mathrm{E}_0 \Big(F(X_1' + s/\sqrt{n}) F(X_1' + t/\sqrt{n}) \\ &+ F(X_1' + s/\sqrt{n})\big(1 - F(Y_1' - t/\sqrt{n})\big) - 2F(X_1' + s/\sqrt{n})\big(1 - \tilde{F}(-t/\sqrt{n})\big) \\ &+ \big(1 - F(Y_1' - s/\sqrt{n})\big) F(X_1' + t/\sqrt{n}) + \big(1 - F(Y_1' - s/\sqrt{n})\big)\big(1 - F(Y_1' - t/\sqrt{n})\big) \\ &- 2\big(1 - F(Y_1' - s/\sqrt{n})\big)\big(1 - \tilde{F}(-t/\sqrt{n})\big) - 2\big(1 - \tilde{F}(-s/\sqrt{n})\big) F(X_1' + t/\sqrt{n}) \\ &- 2\big(1 - \tilde{F}(-s/\sqrt{n})\big)\big(1 - F(Y_1' - t/\sqrt{n})\big) + 4\big(1 - \tilde{F}(-s/\sqrt{n})\big)\big(1 - \tilde{F}(-t/\sqrt{n})\big)\Big)\end{aligned}$$

$$=\frac{a^2+b^2}{6}+2ab\Big(\mathrm{E}_0\left(F(X_1')^2\right)+\mathrm{E}_0\left(F(Y_1')^2\right)-2\,\mathrm{E}_0\left(F(X_1')\right)\mathrm{E}_0\left(F(Y_1')\right)\Big)$$

$$=\frac{a^2+b^2}{6}+\frac{2ab}{6}=\frac{(a+b)^2}{6},\ \ s,t\in\mathbb{R},$$

also die Behauptung. □

Es sei für feste aber beliebige $a,b\in\mathbb{R}$

$$\sigma_n(s,t)^2:=n^3\,\mathrm{E}_0\Bigg(\bigg(\mathrm{E}_0\Big(a\,\mathrm{I}\left(2t/\sqrt{n}>-(z_1+\zeta_2)\right)+b\,\mathrm{I}\left(2t/\sqrt{n}>-(z_1+\zeta_2)\right)\Big)$$
$$-\mathrm{E}_0\Big(a\,\mathrm{I}\left(2t/\sqrt{n}>-(\zeta_1+\zeta_2)\right)+b\,\mathrm{I}\left(2t/\sqrt{n}>-(\zeta_1+\zeta_2)\right)\Big)\bigg)^2_{|z_1=\zeta_1}\Bigg),\ \ s,t\in\mathbb{R}.$$

10.6.7 Lemma: *Es gilt*

$$\lim_{n\to\infty}\frac{\sigma_n(s,t)^2}{n^3}=\frac{(a+b)^2}{12},\ \ t\in\mathbb{R}.$$

Beweis: Betrachte die stetige Funktion $G:t\mapsto P_0(\zeta_1\le t)$. Es gilt aufgrund der Stetigkeit von G für $\vartheta=0$

$$\lim_{n\to\infty}G(\zeta_1+2t/\sqrt{n})=G(\zeta_1)\sim\mathfrak{R}(0,1),\ \ t\in\mathbb{R},$$

und deshalb und aufgrund der Stetigkeit von $\tilde{G}$ in $0\in\mathbb{R}$ folgt mit den Überlegungen aus dem Beweis von Lemma 10.6.4 und dem Satz von der majorisierten Konvergenz

$$\lim_{n\to\infty}\frac{\sigma_n(s,t)^2}{n^3}$$
$$=\lim_{n\to\infty}\mathrm{E}_0\Bigg(\bigg(a\Big(1-G(-\zeta_1-2s/\sqrt{n})-\big(1-\tilde{G}(-2s/\sqrt{n})\Big)$$
$$+b\Big(1-G(-\zeta_1-2s/\sqrt{n})-\big(1-\tilde{G}(-2t/\sqrt{n})\Big)\bigg)^2\Bigg)$$
$$=\frac{a^2+b^2}{12}+2ab\lim_{n\to\infty}\mathrm{E}_0\left(G(\zeta_1+2s/\sqrt{n})G(\zeta_1+2t/\sqrt{n})-G(\zeta_1+2s/\sqrt{n})\tilde{G}(2t/\sqrt{n})\right.$$
$$\left.-\tilde{G}(2s/\sqrt{n})G(\zeta_1+2t/\sqrt{n})+\tilde{G}(2s/\sqrt{n})\tilde{G}(2t/\sqrt{n})\right)$$
$$=\frac{a^2+b^2}{12}+2ab\Big(\mathrm{E}_0\left(G(\zeta_1)^2\right)-\mathrm{E}_0\left(G(\zeta_1)\right)+\frac{1}{4}\Big)=\frac{a^2+b^2}{12}+\frac{2ab}{12}=\frac{(a+b)^2}{12},\ \ s,t\in\mathbb{R},$$

und so die Behauptung. □

10.6.8 Satz: *Es seien die Funktionen $\tilde{F}$ und $\tilde{G}$ auf offenen Umgebungen U mit $0 \in U$ und U' mit $0 \in U'$ zweimal stetig differenzierbar mit $\tilde{f}(t) = \frac{\mathrm{d}}{\mathrm{d}t}\tilde{F}(t)$, $t \in U$, und $\tilde{g}(t) := \frac{\mathrm{d}}{\mathrm{d}t}\tilde{G}(t)$, $t \in U'$, als Ableitungen. Ferner gelte $\tilde{f}(0) \neq 0$ und $\tilde{g}(0) \neq 0$. Dann gilt für die asymptotische Volumen-Effizienz der Testfolge T relativ zur Testfolge T'*

$$e^{\mathrm{Lvol}}_{T',T} = \frac{1}{2}\left(\frac{\tilde{f}(0)}{\tilde{g}(0)}\right)^2.$$

Beweis: Mit Lemma 10.6.6 gilt $\liminf_{n\to\infty} \sigma'_n(s,t)/n^{3/2} > 0$, $s,t \in \mathbb{R}$, $a,b \in \mathbb{R}$, $a \neq 0$ oder $b \neq 0$. Damit folgt mit einem zentralen Grenzwertsatz für U-Statistiken mit einem von n abhängigen Kern (siehe Korollar A.1.5 im Anhang), dass für alle $a,b \in \mathbb{R}$, mit $a \neq 0$ oder $b \neq 0$, für $\vartheta = 0$

$$\begin{aligned}&\Bigg(\sum_{i=1}^{n}\sum_{j=1}^{n}\Big(a\,\mathrm{I}\big(s/\sqrt{n} > -(X'_i - Y'_j)\big) + b\,\mathrm{I}\big(t/\sqrt{n} > -(X'_i - Y'_j)\big)\Big)\\ &- \left(2a\binom{n}{2}P_0\big(s/\sqrt{n} > -(X'_1 - Y'_1)\big) + 2b\binom{n}{2}P_0\big(t/\sqrt{n} > -(X'_1 - Y'_1)\big)\right)\Bigg)\Bigg/\sigma'_n(s,t)\\ &\xrightarrow{\mathrm{v}} \mathrm{N}(0,1) \text{ für } n \to \infty,\ s,t \in \mathbb{R},\end{aligned}$$

gilt. Mit Lemma 10.6.3 und Lemma 10.6.6 erhält man weiter, dass für alle $a,b \in \mathbb{R}$, mit $a \neq 0$ oder $b \neq 0$,

$$\begin{aligned}&\lim_{n\to\infty}\left(a\frac{c'_n - 2\binom{n}{2}P_0\big(s/\sqrt{n} > -(X'_1 - Y'_1)\big)}{\sigma'_n(s,t)} + b\frac{n^2 - c'_n - 2\binom{n}{2}P_0\big(t/\sqrt{n} > -(X'_1 - Y'_1)\big)}{\sigma'_n(s,t)}\right)\\ =&\lim_{n\to\infty}\Bigg(a\frac{\sqrt{n^3}}{\sigma'_n(s,t)}\frac{\sigma'_n(s)}{\sqrt{n^3}}\frac{c'_n - 2\binom{n}{2}P_0\big(s/\sqrt{n} > -(X'_1 - Y'_1)\big)}{\sigma'_n(s)}\\ &+ b\frac{\sqrt{n^3}}{\sigma'_n(s,t)}\frac{\sigma'_n(t)}{\sqrt{n^3}}\frac{n^2 - c'_n - 2\binom{n}{2}P_0\big(t/\sqrt{n} > -(X'_1 - Y'_1)\big)}{\sigma'_n(t)}\Bigg)\\ =&\frac{1}{|a+b|}\big((a-b)u_{1-\frac{\alpha}{2}} - \sqrt{6}\tilde{f}(0)(as+bt)\big),\ s,t \in \mathbb{R},\end{aligned}$$

gilt. Es ist mit Lemma 10.6.6 $\lim_{n\to\infty}\sigma'_n(s,t)/\sqrt{n^3} = |a+b|/\sqrt{6}$, sodass mit dem Lemma von Slutsky für alle $a,b \in \mathbb{R}$, mit $a \neq 0$ oder $b \neq 0$, für $\vartheta = 0$

$$\begin{aligned}&\Bigg(\sum_{i=1}^{n}\sum_{j=1}^{n}\Big(a\,\mathrm{I}\big(s/\sqrt{n} > -(X'_i - Y'_j)\big) + b\,\mathrm{I}\big(t/\sqrt{n} > -(X'_i - Y'_j)\big)\Big)\\ &- \left(2a\binom{n}{2}P_0\big(s/\sqrt{n} > -(X'_1 - Y'_1)\big) + 2b\binom{n}{2}P_0\big(t/\sqrt{n} > -(X'_1 - Y'_1)\big)\right)\Bigg)\Bigg/\sqrt{n^3}\\ &- \left(a\frac{c'_n - 2\binom{n}{2}P_0\big(s/\sqrt{n} > -(X'_1 - Y'_1)\big)}{\sqrt{n^3}} + b\frac{n^2 - c'_n - 2\binom{n}{2}P_0\big(t/\sqrt{n} > -(X'_1 - Y'_1)\big)}{\sqrt{n^3}}\right)\\ &\xrightarrow{\mathrm{v}} \mathrm{N}\left(-\frac{1}{\sqrt{6}}\big((a-b)u_{1-\frac{\alpha}{2}} - \sqrt{6}\tilde{f}(0)(as+bt)\big), \frac{(a+b)^2}{6}\right) \text{ für } n \to \infty,\ s,t \in \mathbb{R},\end{aligned}$$

gilt. Da $a, b \in \mathbb{R}$, mit $a \neq 0$ oder $b \neq 0$, beliebig waren, folgt mit dem Satz von Cramér und Wold, dass der Zufallsvektor für $\vartheta = 0$

$$\Bigg(\frac{\sum_{i=1}^n \sum_{j=1}^n \mathrm{I}\left(s/\sqrt{n} > -(X_i' - Y_j')\right) - 2\binom{n}{2} P_0\big(s/\sqrt{n} > -(X_1' - Y_1')\big)}{\sqrt{n^3}} \\
- \frac{c_n' - 2\binom{n}{2} P_0\big(s/\sqrt{n} > -(X_1' - Y_1')\big)}{\sqrt{n^3}}, \\
\frac{\sum_{i=1}^n \sum_{j=1}^n \mathrm{I}\left(t/\sqrt{n} > -(X_i' - Y_j')\right) - 2\binom{n}{2} P_0\big(t/\sqrt{n} > -(X_1' - Y_1')\big)}{\sqrt{n^3}} \\
- \frac{n^2 - c_n' - 2\binom{n}{2} P_0\big(t/\sqrt{n} > -(X_1' - Y_1')\big)}{\sqrt{n^3}}\Bigg), \quad s, t \in \mathbb{R},$$

für $n \to \infty$ in Verteilung gegen eine zweidimensionale Normalverteilung konvergiert.

Analog folgt mit Lemma 10.6.4 und Lemma 10.6.7, dass für alle $a, b \in \mathbb{R}$, mit $a \neq 0$ oder $b \neq 0$, für $\vartheta = 0$

$$\Bigg(\sum_{1 \leq i \leq j \leq n}^{n} \Big(a\,\mathrm{I}\left(2s/\sqrt{n} > -(\zeta_i + \zeta_j)\right) + b\,\mathrm{I}\left(2t/\sqrt{n} > -(\zeta_i + \zeta_j)\right)\Big) \\
- \left(2a\binom{n}{2} P_0\big(2s/\sqrt{n} > -(\zeta_1 + \zeta_2)\big) + 2b\binom{n}{2} P_0\big(2t/\sqrt{n} > -(\zeta_1 + \zeta_2)\big)\right)\Bigg) \Big/ \sigma_n(s,t) \\
\xrightarrow{\mathrm{v}} \mathrm{N}(0,1) \text{ für } n \to \infty, \ s, t \in \mathbb{R},$$

sowie

$$\Bigg(\sum_{1 \leq i \leq j \leq n}^{n} \Big(a\,\mathrm{I}\left(2s/\sqrt{n} > -(\zeta_i + \zeta_j)\right) + b\,\mathrm{I}\left(2t/\sqrt{n} > -(\zeta_i + \zeta_j)\right)\Big) \\
- \left(2a\binom{n}{2} P_0\big(2s/\sqrt{n} > -(\zeta_1 + \zeta_2)\big) + 2b\binom{n}{2} P_0\big(2t/\sqrt{n} > -(\zeta_1 + \zeta_2)\big)\right)\Bigg) \Big/ \sqrt{n^3} \\
- \left(a\frac{c_n - 2\binom{n}{2} P_0\big(2s/\sqrt{n} > -(\zeta_1 + \zeta_2)\big)}{\sqrt{n^3}} + b\frac{n(n+1)/2 - c_n - 2\binom{n}{2} P_0\big(2t/\sqrt{n} > -(\zeta_1 + \zeta_2)\big)}{\sqrt{n^3}}\right) \\
\xrightarrow{\mathrm{v}} \mathrm{N}\left(-\frac{1}{\sqrt{12}}\big((a-b)u_{1-\frac{\alpha}{2}} - \sqrt{12}\tilde{g}(0)(as+bt)\big), \frac{(a+b)^2}{12}\right) \text{ für } n \to \infty, \ s, t \in \mathbb{R},$$

gilt, und auch, dass der Zufallsvektor für $\vartheta = 0$

$$\left(\frac{\sum_{1\leq i\leq j\leq n}^{n} \mathrm{I}\left(2s/\sqrt{n} > -(\zeta_i+\zeta_j)\right) - 2\binom{n}{2} P_0\big(2s/\sqrt{n} > -(\zeta_1+\zeta_2)\big)}{\sqrt{n^3}} \right.$$
$$- \frac{c_n - 2\binom{n}{2} P_0\big(2s/\sqrt{n} > -(\zeta_1+\zeta_2)\big)}{\sqrt{n^3}},$$
$$\frac{\sum_{1\leq i\leq j\leq n}^{n} \mathrm{I}\left(2t/\sqrt{n} > -(\zeta_i+\zeta_j)\right) - 2\binom{n}{2} P_0\big(2/\sqrt{n} > -(\zeta_1+\zeta_2)\big)}{\sqrt{n^3}}$$
$$\left. - \frac{n(n+1)/2 - c_n - 2\binom{n}{2} P_0\big(2t/\sqrt{n} > -(\zeta_1+\zeta_2)\big)}{\sqrt{n^3}} \right), \ s,t \in \mathbb{R},$$

für $n \to \infty$ in Verteilung gegen eine zweidimensionale Normalverteilung konvergiert. Schließlich folgt

$$\lim_{n\to\infty} P_\vartheta\big(\sqrt{n}(V_{(c'_n+1)} - \vartheta) \geq s, \sqrt{n}(V_{(n^2-c'_n)} - \vartheta) < t\big)$$
$$= \lim_{n\to\infty} P_\vartheta\left(\sum_{i=1}^{n}\sum_{j=1}^{n} \mathrm{I}(X'_i > Y'_j - \vartheta - s/\sqrt{n}) \leq c'_n, \sum_{i=1}^{n}\sum_{j=1}^{n} \mathrm{I}(X'_i > Y'_j - \vartheta - t/\sqrt{n}) \geq n^2 - c'_n \right)$$
$$= \lim_{n\to\infty} P_0\left(\frac{\sum_{i=1}^{n}\sum_{j=1}^{n} \mathrm{I}\left(s/\sqrt{n} > -(X'_i - Y'_j)\right) - 2\binom{n}{2} P_0\big(s/\sqrt{n} > -(X'_1 - Y'_1)\big)}{\sqrt{n^3}} - \right.$$
$$- \frac{c'_n - 2\binom{n}{2} P_0\big(s/\sqrt{n} > -(X'_1 - Y'_1)\big)}{\sqrt{n^3}} \leq 0,$$
$$\frac{\sum_{i=1}^{n}\sum_{j=1}^{n} \mathrm{I}\left(t/\sqrt{n} > -(X'_i - Y'_j)\right) - 2\binom{n}{2} P_0\big(t/\sqrt{n} > -(X'_1 - Y'_1)\big)}{\sqrt{n^3}}$$
$$\left. - \frac{n^2 - c'_n - 2\binom{n}{2} P_0\big(t/\sqrt{n} > -(X'_1 - Y'_1)\big)}{\sqrt{n^3}} \geq 0 \right) \in (0,1), \ s,t \in \mathbb{R},$$

was bedeutet, dass ein zweidimensionaler Zufallsvektor (A_1, A_2) existiert, sodass

$$\big(\sqrt{n}(V_{(c'_n+1)} - \vartheta), \sqrt{n}(V_{(n^2-c'_n)} - \vartheta)\big) \xrightarrow{\text{v}} (A_1, A_2) \text{ für } n \to \infty$$

gilt. Wieder mit dem Satz von Cramér und Wold folgt daraus, dass auch

$$\sqrt{n}(V_{(c'_n+1)} - V_{(n^2-c'_n)}) = \sqrt{n}(V_{(c'_n+1)} - \vartheta) - \sqrt{n}(V_{(n^2-c'_n)} - \vartheta) \xrightarrow{\text{v}} A_1 - A_2$$
$$\text{für } n \to \infty$$

gilt. Mit Satz 10.6.5 folgt

$$\mathrm{E}(A_1 - A_2) = \mathrm{E}(A_1) - \mathrm{E}(A_2) = \frac{2u_{1-\frac{\alpha}{2}}}{\sqrt{6}\tilde{f}(0)}.$$

Analog folgt, dass ein zweidimensionaler Zufallsvektor (B_1, B_2) existiert, sodass

$$\left(\sqrt{n}(W_{(c_n+1)} - \vartheta), \sqrt{n}(W_{(n(n+1)/2-c_n)} - \vartheta)\right) \xrightarrow{\mathrm{v}} (B_1, B_2) \text{ für } n \to \infty,$$

und auch

$$\sqrt{n}(W_{(c_n+1)} - W_{(n(n+1)/2-c_n)}) \xrightarrow{\mathrm{v}} B_1 - B_2 \text{ für } n \to \infty,$$

sowie

$$\mathrm{E}(B_1 - B_2) = \frac{2u_{1-\frac{\alpha}{2}}}{\sqrt{12}\tilde{g}(0)}$$

gilt. Damit ist alles gezeigt. □

10.7 Pitman-Effizienz und lokale Bahadur-Effizienz der Tests

Mit Hilfe von analogen Überlegungen zu denen in den vorangegangenen Unterabschnitten, konkret aufgrund der asymptotischen Normalität von U-Statistiken, lässt sich mit dem bekannten Resultat zur Pitman-Effizienz erwähnt am Endes des Unterabschnittes 2.4 die Pitman-Effizienz der Testfolge T relativ zur Testfolge T' angeben. Vergleiche mit der Arbeit von Hollander ([16]).

Es existiere eine Dichte ℓ bezüglich des Lebesgue-Borelschen Maßes von (X_1, Y_1) bei Gültigkeit der Hypothese. Dann sei f eine Dichte bezüglich des Lebesgue-Borelschen Maßes von X_1, d.h. $f(x) = \frac{\mathrm{d}}{\mathrm{d}x}F(x)$, $x \in \mathbb{R}$, und g eine Dichte bezüglich des Lebesgue-Borelschen Maßes von ζ_1 bei Gültigkeit der Hypothese, d.h. $g(x) = \frac{\mathrm{d}}{\mathrm{d}x}G(x)$, $x \in \mathbb{R}$. Ferner sei $\tilde{f}$ eine Dichte bezüglich des Lebesgue-Borelschen Maßes von ζ_1' bei Gültigkeit der Hypothese, d.h. $\tilde{f}(x) = \frac{\mathrm{d}}{\mathrm{d}x}\tilde{F}(x)$, $x \in \mathbb{R}$, und $\tilde{g}$ eine Dichte bezüglich des Lebesgue-Borelschen Maßes von $\zeta_1 + \zeta_2$ bei Gültigkeit der Hypothese, d.h. $\tilde{g}(x) = \frac{\mathrm{d}}{\mathrm{d}x}\tilde{G}(x)$, $x \in \mathbb{R}$. Es gelte $\tilde{f}(0) \neq 0$ und $\tilde{g}(0) \neq 0$.

10.7.1 Korollar: *Es gilt für die Pitman- und lokale Bahadur-Effizienz der Testfolge T relativ zur Testfolge T'*

$$e^{\mathrm{P}}_{T',T} = e^{\mathrm{LB}}_{T',T} = \frac{1}{2}\left(\frac{\tilde{f}(0)}{\tilde{g}(0)}\right)^2.$$

Sind zusätzlich die Funktionen $\tilde{F}$ und $\tilde{G}$ auf offenen Umgebungen U mit $0 \in U$ und U' mit $0 \in U'$ zweimal stetig differenzierbar, so ist

$$e^{\mathrm{P}}_{T',T} = e^{\mathrm{LB}}_{T',T} = e^{\mathrm{Lvol}}_{T',T}.$$

Beweis: Die Gestalt der Pitman-Effizienz folgt mit Hilfe der asymtotischen Normalität von U-Statistiken und mit dem bekannten Resultat zur Pitman-Effizienz erwähnt am Ende des

Unterabschnittes 2.4. Mit Hilfe des Satzes von L'Hospital und der Sätze 10.3.1 und 10.5.1 erhält man

$$e^{\mathrm{LB}}_{T',T} = \lim_{n\to\infty} 2\left(\frac{P_0(\zeta_1' > \vartheta) - 1/2}{P_0(\zeta_1 + \zeta_2 > 2\vartheta) - 1/2}\right)^2 = \frac{1}{2}\left(\frac{\tilde{f}(0)}{\tilde{g}(0)}\right)^2.$$

Mit Satz 10.6.8 folgt die letzte genannte Identität. □

10.7.2 Beispiel: Im Fall, dass $\binom{X_1}{Y_1}$ zweidimensional normalverteilt ist mit Erwartungswertvektor $(a, a+\vartheta)$, mit einem $a \in \mathbb{R}$, und Kovarianzmatrix

$$\Sigma := \sigma \begin{pmatrix} 1 & \rho \\ \rho & 1 \end{pmatrix},$$

mit $\sigma > 0$ und $\rho \in (-1, 1)$, liegen die Situationen aus den Beispielen 10.1.1 und 10.1.2 vor. Weiter ergibt sich $\tilde{f}(0) = \int f(t)^2 \mathrm{d}t = 1/\sqrt{4\pi\sigma}$ sowie $\tilde{g}(0) = \int g(t)^2 \mathrm{d}t = 1/\sqrt{8\pi\sigma(1-\rho)}$ und damit

$$e^{\mathrm{P}}_{T',T} = e^{\mathrm{LB}}_{T',T} = e^{\mathrm{Lvol}}_{T',T} = 1 - \rho.$$

Also ist die Effizienz kleiner als 1 (gleich 1 bzw. größer als 1), falls ρ größer als 0 (gleich 0 bzw. kleiner als 0) ist.

10.7.3 Beispiel: Die verbundene Stichprobe folge einer Gumbel-Morgenstern-Verteilung, d.h. für ein $\lambda \in [-1, 1]$ lässt sich die gemeinsame Verteilungsfunktion von X_1 und Y_1 schreiben als

$$M_{\lambda,\vartheta}(x, y) := P_{\vartheta,\lambda}(X_1 \le x, Y_1 \le y) = F(x)F(y-\vartheta)\Big(1 - \lambda\big(1 - F(x)\big)\big(1 - F(y-\vartheta)\big)\Big),$$
$$x, y \in \mathbb{R}.$$

Die zugehörige Dichte besitzt die Gestalt

$$m_{\lambda,\vartheta}(x, y) := \Big(1 - \lambda\big(1 - 2F(x)\big)\big(1 - 2F(y-\vartheta)\big)\Big) f(x) f(y-\vartheta), \ x, y \in \mathbb{R}.$$

Es liegt also die Situation aus Beispiel 10.1.1 vor. Im Fall, dass $f(x) = f(-x) \ \forall \, x \in \mathbb{R}$ gilt, liegt außerdem der Sachverhalt aus Beispiel 10.1.2 vor. Setzen wir

$$a := \frac{\sqrt{2}\int \Big(\int \big(1 - 2F(t)\big)\big(1 - 2F(y)\big) f(t+y) f(y) \mathrm{d}y\Big)^2 \mathrm{d}t}{\int f(t)^2 \mathrm{d}t},$$

$$b := -\frac{2\sqrt{2}\int \Big(\int f(t+y) f(y)\mathrm{d}y\Big)\Big(\int \big(1 - 2F(t+y)\big)\big(1 - 2F(y)\big) f(t+y) f(y) \mathrm{d}y\Big) \mathrm{d}t}{\int f(t)^2 \mathrm{d}t},$$

$$c := \frac{\sqrt{2}\int \Big(\int f(t+y) f(y) \mathrm{d}y\Big)^2 \mathrm{d}t}{\int f(t)^2 \mathrm{d}t},$$

so ergibt sich

$$\sqrt{e^{\mathrm{P}}_{T',T}} = \sqrt{e^{\mathrm{LB}}_{T',T}} = \sqrt{e^{\mathrm{Lvol}}_{T',T}} = \frac{1}{\sqrt{2}}\frac{\tilde{f}(0)}{\tilde{g}(0)} = \frac{1}{\sqrt{2}}\frac{\int f(t)^2\mathrm{d}t}{\int\big(\int m_{\lambda,0}(t+y,y)\mathrm{d}y\big)^2\mathrm{d}t} = \frac{1}{a\lambda^2+b\lambda+c}.$$

Gilt also $b^2-4a(c-1)<0$, so ist die Effizienz kleiner als 1 für alle $\lambda\in[-1,1]$. Gilt $b^2-4a(c-1)=0$, so ist die Effizienz gleich 1 für $\lambda=-b/(2a)$, falls $\lambda\in[-1,1]$, und kleiner als 1 für alle übrigen $\lambda\in[-1,1]$. Gilt dagegen $b^2-4a(c-1)>0$, so ist die Effizienz gleich 1 für $\lambda=-b/(2a)\pm\sqrt{(b^2-4a(c-1))/(4a^2)}$, falls $\lambda\in[-1,1]$, größer als 1 für $-b/(2a)-\sqrt{(b^2-4a(c-1))/(4a^2)}<\lambda<-b/(2a)+\sqrt{(b^2-4a(c-1))/(4a^2)}$, falls $\lambda\in[-1,1]$, und kleiner als 1 für alle übrigen $\lambda\in[-1,1]$. Hoeffdings Formel zur Berechnung der Kovarianz, siehe [14], liefert hier

$$\begin{aligned}\mathrm{Cov}_\lambda(X_1,Y_1) &= \int\int\big(M_{\lambda,\vartheta}(x,y)-F(x)F(y-\vartheta)\big)\mathrm{d}x\mathrm{d}y\\ &= -\lambda\int\int F(x)\big(1-F(x)\big)F(y-\vartheta)\big(1-F(y-\vartheta)\big)\mathrm{d}x\mathrm{d}y\\ &= -\lambda\Big(\int F(x)\big(1-F(x)\big)\mathrm{d}x\Big)^2 = -\lambda\,\mathrm{E}(|X_1-X_2|)^2,\end{aligned}$$

äquivalent

$$\lambda = -\frac{\mathrm{Cov}_\lambda(X_1,Y_1)}{\Big(\int F(x)\big(1-F(x)\big)\mathrm{d}x\Big)^2} = -\frac{\mathrm{Cov}_\lambda(X_1,Y_1)}{\mathrm{E}(|X_1-X_2|)^2}.$$

Somit lassen sich die genannten Bedingungen für die Effizienz gänzlich auf die Kovarianz zwischen X_1 und Y_1 sowie auf die Verteilungsfunktion F zurückführen.

10.8 Tests zur relativen Wirksamkeit

In diesem Unterabschnitt sollen Tests basierend auf Beobachtungen im verbundenen Stichprobenfall bei Gültigkeit der Hypothese[13] H zur empirischen Untersuchung der in Unterabschnitt 10.7 berechneten Pitman-, lokalen approximativen Bahadur- und asymptotischen Volumen-Effizienz vorgeschlagen und untersucht werden.

Wir betrachten die Testprobleme

$$\begin{aligned} &\mathrm{H}_1:\ e^{\mathrm{P}}_{T,T'}\le 1,\ \mathrm{K}_1:\ e^{\mathrm{P}}_{T,T'}>1,\\ \Leftrightarrow\ &\mathrm{H}_1:\ e^{\mathrm{LB}}_{T,T'}\le 1,\ \mathrm{K}_1:\ e^{\mathrm{LB}}_{T,T'}>1,\\ \Leftrightarrow\ &\mathrm{H}_1:\ e^{\mathrm{Lvol}}_{T,T'}\le 1,\ \mathrm{K}_1:\ e^{\mathrm{Lvol}}_{T,T'}>1,\end{aligned}$$

[13] Beachte Bemerkung 10.8.8.

bzw.

$$\begin{aligned}
&\mathrm{H}_2: \ e^{\mathrm{P}}_{T,T'} \geq 1, \ \mathrm{K}_2: \ e^{\mathrm{P}}_{T,T'} < 1,\\
\Leftrightarrow\ &\mathrm{H}_2: \ e^{\mathrm{LB}}_{T,T'} \geq 1, \ \mathrm{K}_2: \ e^{\mathrm{LB}}_{T,T'} < 1,\\
\Leftrightarrow\ &\mathrm{H}_2: \ e^{\mathrm{Lvol}}_{T,T'} \geq 1, \ \mathrm{K}_2: \ e^{\mathrm{Lvol}}_{T,T'} < 1,
\end{aligned}$$

bzw.

$$\begin{aligned}
&\mathrm{H}_3: \ e^{\mathrm{P}}_{T,T'} = 1, \ \mathrm{K}_3: \ e^{\mathrm{P}}_{T,T'} \neq 1,\\
\Leftrightarrow\ &\mathrm{H}_3: \ e^{\mathrm{LB}}_{T,T'} = 1, \ \mathrm{K}_3: \ e^{\mathrm{LB}}_{T,T'} \neq 1,\\
\Leftrightarrow\ &\mathrm{H}_3: \ e^{\mathrm{Lvol}}_{T,T'} = 1, \ \mathrm{K}_3: \ e^{\mathrm{Lvol}}_{T,T'} \neq 1.
\end{aligned}$$

Es is in diesem Unterabschnitt die Idee, Testgrößen basierend auf Kerndichteschätzern für die entsprechenden Dichten zu konstruieren. Diese Kerndichteschätzer sollen auf Beobachtungen $\binom{X_1}{Y_1}, \dots, \binom{X_n}{Y_n}$ der verbundenen Stichprobe bei Gültigkeit der Hypothese H basieren.

Es sei $k : \mathbb{R} \to \mathbb{R}_{\geq 0}$ eine um $0 \in \mathbb{R}$ symmetrische und beschränkte Kernfunktion mit $\int k(x)\mathrm{d}x = 1$, und $(h_n)_{n\in\mathbb{N}}$ eine Folge von Bandbreiten mit $h_n \in (0, \infty)$ und $\lim_{n\to\infty} h_n = 0$. Basierend auf der verbundenen Stichprobe bei Gültigkeit der Hypothese, d.h. es gilt in diesem Zusammenhang $\vartheta = 0$, definieren wir die Kerndichteschätzer

$$\begin{aligned}
\tilde{f}_n(t) :=& \frac{1}{2(n^2-n)h_n} \sum_{1\leq i\neq j\leq n} \Big(k\Big(\frac{t-(X_i-Y_j)}{h_n}\Big) + k\Big(\frac{t-(Y_i-X_j)}{h_n}\Big)\Big),\\
\tilde{g}_n(t) :=& \frac{1}{2(n^2-n)h_n} \sum_{1\leq i\neq j\leq n} \Big(k\Big(\frac{t-(\zeta_i+\zeta_j)}{h_n}\Big) + k\Big(\frac{t-(\zeta_i-\zeta_j)}{h_n}\Big)\Big), \ t \in \mathbb{R}.
\end{aligned}$$

Für $t = 0$ sind aufgrund der Symmetrie um $0 \in \mathbb{R}$ der Kernfunktion k diese Schätzer invariant unter Transformation der verbundenen Stichprobe bei $\vartheta = 0$ der Form

$$\Big(\binom{X_1}{Y_1}, \dots, \binom{X_n}{Y_n}\Big) \longmapsto \Big(\binom{X_{i_1}}{Y_{i_1}}, \dots, \binom{X_{i_n}}{Y_{i_n}}\Big), \tag{10.8.1}$$

für alle Permutationen $\{i_1, \dots, i_n\}$ von $\{1, \dots, n\}$, und sie sind invariant invariant unter der Transformation

$$\Big(\binom{X_1}{Y_1}, \dots, \binom{X_n}{Y_n}\Big) \longmapsto \Big(\binom{Y_1}{X_1}, \dots, \binom{Y_n}{X_n}\Big). \tag{10.8.2}$$

Ferner definieren wir eine Abbildung $L_n : \mathbb{R}^4 \to \mathbb{R}$ durch

$$\begin{aligned}
L_n(w_1, w_2) :=& -\frac{\sqrt{2}}{2h_n}\Big(k\Big(-\frac{x_1-y_1+x_2-y_2}{h_n}\Big) + k\Big(-\frac{x_1-y_1-(x_2-y_2)}{h_n}\Big)\Big)\\
&+ \frac{1}{2h_n}\Big(k\Big(-\frac{x_1-y_2}{h_n}\Big) + k\Big(-\frac{y_1-x_2}{h_n}\Big)\Big),\\
&w_1 = (x_1, y_1), w_2 = (x_2, y_2) \in \mathbb{R}^2.
\end{aligned}$$

Dann gilt $L_n((x_1, y_1), (x_2, y_2)) = L_n((y_1, x_1), (y_2, x_2))$.

Basierend auf der verbundenen Stichprobe bei Gültigkeit der Hypothese setzen wir $W_i := (X_i, Y_i)$, $i = 1, \ldots, n$, und definieren damit

$$\begin{aligned} M_n :=& 2\Bigg(\frac{1}{n(n-1)(n-2)} \sum_{\substack{1 \leq i_1, i_2, i_3 \leq n \\ i_l \neq i_m, l \neq m}} L_n(W_{i_1}, W_{i_2}) L_n(W_{i_1}, W_{i_3}) \\ & - \left(\frac{1}{n^2 - n} \sum_{1 \leq i \neq j \leq n} L_n(W_i, W_j) \right)^2 \Bigg)^{1/2}, \end{aligned}$$

falls

$$\frac{1}{n(n-1)(n-2)} \sum_{\substack{1 \leq i_1, i_2, i_3 \leq n \\ i_l \neq i_m, l \neq m}} L_n(W_{i_1}, W_{i_2}) L_n(W_{i_1}, W_{i_3}) - \left(\frac{1}{n^2 - n} \sum_{1 \leq i \neq j \leq n} L_n(W_i, W_j) \right)^2 > 0,$$

und $M_n := 0$ sonst. Nun betrachten wir die Testgrößen

$$E_n^1 := \frac{\sqrt{n}\big(\tilde{f}_n(0) - \sqrt{2}\tilde{g}_n(0)\big)}{M_n}, \; E_n^2 := -E_n^1, \; E_n^3 := |E_n^1|,$$

wobei an dieser Stelle die Konvention

$$\frac{a}{0} := \begin{cases} +\infty & , a > 0 \\ 0 & , a = 0 \\ -\infty & , a < 0 \end{cases}$$

gelte. Diese Testgrößen sind nach Konstruktion invariant unter den oben genannten Transformationen (10.8.1) und (10.8.2) der Stichprobe.

Um das asymptotische Verhalten der Testgrößen zu studieren, stellen wir zunächst Lemmata bereit. Zuvor definieren wir

$$\tilde{L}_n(w_1, w_2, w_3) := \frac{L_n(w_1, w_2)L_n(w_1, w_3) + L_n(w_2, w_1)L_n(w_2, w_3) + L_n(w_3, w_2)L_n(w_3, w_1)}{3},$$

$w_1, w_2, w_3 \in \mathbb{R}^2$.

10.8.1 Lemma: *Die Dichten $f, g, \tilde{f}, \tilde{g}$ seien beschränkt. Ferner gelte $\lim_{n\to\infty} nh_n = \infty$. Dann gilt:*

a) $\sup_{w\in\mathbb{R}^2} \mathrm{E}(L_n^2(w, W_2)) = o(n)$ *für $n \to \infty$,* $\sup_{n\in\mathbb{N}} \sup_{w\in\mathbb{R}^2} \mathrm{E}(|L_n(w, W_2)|) < \infty$ *sowie* $\sup_{n\in\mathbb{N}} \mathrm{E}(|L_n(W_1, W_2)|) < \infty$. *Weiter ist* $\mathrm{E}(L_n^2(W_1, W_2)) = o(n)$ *und* $\sup_{w_1,w_2\in\mathbb{R}^2} |L_n(w_1, w_2)| = o(n)$ *für $n \to \infty$.*

b) Mit $h_{n,i}(w_1,\dots,w_i) := \mathrm{E}(\tilde{L}_n(w_1,\dots,w_i,W_{i+1},\dots,W_3))$, $w_1,\dots,w_i \in \mathbb{R}^2$, *gilt* $\mathrm{E}(h_{n,i}^2(W_1,\dots,W_i)) = o(n^i)$ *für* $n\to\infty$, $i=1,2,3$.

Beweis: Da für $m\in\mathbb{N}$ die Ungleichung

$$(x_1+\dots+x_m)^2 \le mx_1^2+\dots+mx_m^2,\ x_i\in\mathbb{R},\ i=1,\dots,m, \tag{10.8.3}$$

gilt, welche aus der Cauchy-Schwarzschen Ungleichung folgt, ist

$$\begin{aligned}
&\mathrm{E}\left(L_n^2(w_1,W_2)\right)\\
\le&\,\mathrm{E}\left(\frac{4}{h_n^2}k^2\Big(-\frac{x_1-y_1+\zeta_2}{h_n}\Big)\right)+\mathrm{E}\left(\frac{1}{h_n^2}k^2\Big(-\frac{x_1-Y_2}{h_n}\Big)\right)+\mathrm{E}\left(\frac{1}{h_n^2}k^2\Big(-\frac{y_1-X_2}{h_n}\Big)\right)\\
=&\int\frac{4}{h_n^2}k^2\Big(-\frac{x_1-y_1+u}{h_n}\Big)g(u)\mathrm{d}u+\int\frac{1}{h_n^2}k^2\Big(-\frac{x_1-u}{h_n}\Big)f(u)\mathrm{d}u\\
&+\int\frac{1}{h_n^2}k^2\Big(-\frac{y_1-u}{h_n}\Big)f(u)\mathrm{d}u\\
=&\frac{4}{h_n}\int k^2(u)g\big(h_nu-(x_1-y_1)\big)\mathrm{d}u+\frac{1}{h_n}\int k^2(u)f(h_nu-x_1)\mathrm{d}u\\
&+\frac{1}{h_n}\int k^2(u)f(h_nu-y_1)\mathrm{d}u,
\end{aligned}$$

woraus wegen der Beschränktheit von f und g sowie wegen $\lim_{n\to\infty} nh_n = \infty$ $\sup_{w\in\mathbb{R}^2}\mathrm{E}(L_n^2(w,W_2)) = o(n)$ für $n\to\infty$ folgt.

Zudem ist

$$\begin{aligned}
&\mathrm{E}\left(|L_n(w,W_2)|\right)\\
\le&\sqrt{2}\,\mathrm{E}\left(\frac{1}{h_n}k\Big(-\frac{x-y+\zeta_2}{h_n}\Big)\right)\\
&+\mathrm{E}\left(\frac{1}{2h_n}k\Big(-\frac{x-Y_2}{h_n}\Big)\right)+\mathrm{E}\left(\frac{1}{2h_n}k\Big(-\frac{y-X_2}{h_n}\Big)\right)\\
=&\sqrt{2}\int\frac{1}{h_n}k\Big(-\frac{x-y+u}{h_n}\Big)g(u)\mathrm{d}u+\int\frac{1}{2h_n}k\Big(-\frac{x-u}{h_n}\Big)f(u)\mathrm{d}u\\
&+\int\frac{1}{2h_n}k\Big(-\frac{y-u}{h_n}\Big)f(u)\mathrm{d}u\\
=&\sqrt{2}\int k(u)g(h_nu+x-y)\mathrm{d}u+\frac{1}{2}\int k(u)f(h_nu+x)\mathrm{d}u+\frac{1}{2}\int k(u)f(h_nu+y)\mathrm{d}u,\\
&w=(x,y)\in\mathbb{R}^2,
\end{aligned}$$

woraus wegen der Beschränktheit von f und g $\sup_{n\in\mathbb{N}}\sup_{w\in\mathbb{R}^2}\mathrm{E}(|L_n(w,W_2)|)<\infty$ folgt.

Ähnlich ist

$$\begin{aligned}
&\mathrm{E}\left(|L_n(W_1,W_2)|\right)\\
\leq&\sqrt{2}\,\mathrm{E}\left(\frac{1}{h_n}k\Big(-\frac{X_1-Y_1+\zeta_2}{h_n}\Big)\right)\\
&+\mathrm{E}\left(\frac{1}{2h_n}k\Big(-\frac{X_1-Y_2}{h_n}\Big)\right)+\mathrm{E}\left(\frac{1}{2h_n}k\Big(-\frac{Y_1-X_2}{h_n}\Big)\right)\\
=&\sqrt{2}\,\mathrm{E}\left(\int\frac{1}{h_n}k\Big(-\frac{X_1-Y_1+u}{h_n}\Big)g(u)\mathrm{d}u\right)+\mathrm{E}\left(\int\frac{1}{2h_n}k\Big(-\frac{X_1-u}{h_n}\Big)f(u)\mathrm{d}u\right)\\
&+\mathrm{E}\left(\int\frac{1}{2h_n}k\Big(-\frac{Y_1-u}{h_n}\Big)f(u)\mathrm{d}u\right)\\
=&\sqrt{2}\,\mathrm{E}\left(\int k(u)g(h_nu+X_1-Y_1)\mathrm{d}u\right)+\mathrm{E}\left(\frac{1}{2}\int k(u)f(h_nu+X_1)\mathrm{d}u\right)\\
&+\mathrm{E}\left(\frac{1}{2}\int k(u)f(h_nu+Y_1)\mathrm{d}u\right),
\end{aligned}$$

woraus wegen der Beschränktheit von f und g $\sup_{n\in\mathbb{N}}\mathrm{E}(|L_n(W_1,W_2)|)<\infty$ folgt.

Ferner erhält man mit der Ungleichung (10.8.3)

$$\begin{aligned}
L_n^2(w_1,w_2)\leq&\frac{2}{h_n^2}\left(k^2\Big(-\frac{x_1-y_1+x_2-y_2}{h_n}\Big)+k^2\Big(-\frac{x_1-y_1-(x_2-y_2)}{h_n}\Big)\right)\\
&+\frac{1}{h_n^2}\left(k^2\Big(-\frac{x_1-y_2}{h_n}\Big)+k^2\Big(-\frac{y_1-x_2}{h_n}\Big)\right),\ w_1=(x_1,y_1),w_2=(x_2,y_2)\in\mathbb{R}^2,
\end{aligned}$$

und daher

$$\begin{aligned}
&\mathrm{E}\left(L_n^2(W_1,W_2)\right)\\
\leq&\,\mathrm{E}\left(\frac{4}{h_n^2}k^2\Big(-\frac{\zeta_1+\zeta_2}{h_n}\Big)\right)+\mathrm{E}\left(\frac{2}{h_n^2}k^2\Big(-\frac{X_1-Y_2}{h_n}\Big)\right)\\
=&\int\int\frac{4}{h_n^2}k^2\Big(-\frac{u+v}{h_n}\Big)g(u)\mathrm{d}ug(v)\mathrm{d}v+\int\int\frac{2}{h_n^2}k^2\Big(-\frac{u-v}{h_n}\Big)f(u)f(v)\mathrm{d}u\mathrm{d}v\\
=&\frac{4}{h_n}\int\int k^2(u)g(h_nu-v)g(v)\mathrm{d}u\mathrm{d}v+\frac{2}{h_n}\int\int k^2(u)f(h_nu-v)f(v)\mathrm{d}u\mathrm{d}v,
\end{aligned}$$

woraus wegen der Beschränktheit von f und g sowie wegen $\lim_{n\to\infty}nh_n=\infty$ $\mathrm{E}(L_n^2(W_1,W_2))=o(n)$ für $n\to\infty$ folgt.

$\sup_{w_1,w_2\in\mathbb{R}^2}|L_n(w_1,w_2)|=o(n)$ für $n\to\infty$ folgt wegen $\lim_{n\to\infty}nh_n=\infty$ und der Beschränktheit der Kernfunktion k sofort aus der Definition von L_n.

Für Teil b) ist

$$h_{n,1}(w_1)=\frac{1}{3}\,\mathrm{E}\left(L_n(w_1,W_2)\right)^2+\frac{2}{3}\,\mathrm{E}\left(L_n(W_2,w_1)L_n(W_2,W_3)\right),\ w_1\in\mathbb{R}^2,$$

und daher

$$
\begin{aligned}
\mathrm{E}\left(h_{n,1}^2(W_1)\right) =& \frac{1}{9}\,\mathrm{E}\Big(\,\mathrm{E}\left(L_n(w_1,W_2)\right)^4_{|w_1=W_1}\Big)\\
&+\frac{4}{9}\,\mathrm{E}\Big(\,\mathrm{E}\left(L_n(w_1,W_2)\right)^2\,\mathrm{E}\left(L_n(W_2,w_1)L_n(W_2,W_3)\right)_{|w_1=W_1}\Big)\\
&+\frac{4}{9}\,\mathrm{E}\Big(\,\mathrm{E}\left(L_n(W_2,w_1)L_n(W_2,W_3)\right)^2_{|w_1=W_1}\Big).
\end{aligned}
$$

Nach Teil a) gilt $\sup_{n\in\mathbb{N}}\sup_{w\in\mathbb{R}^2}\mathrm{E}(|L_n(w,W_1)|)<c\in(0,\infty)$, sodass

$$
\begin{aligned}
\mathrm{E}\left(h_{n,1}^2(W_1)\right) \leq& \frac{1}{9}c^4+\frac{4}{9}c^2\,\mathrm{E}\left(|L_n(W_2,W_1)L_n(W_2,W_3)|\right)\\
&+\frac{4}{9}\,\mathrm{E}\Big(\,\mathrm{E}\left(L_n(W_2,w_1)L_n(W_2,W_3)\right)^2_{|w_1=W_1}\Big)
\end{aligned}
$$

gilt. Es ist

$$
\begin{aligned}
&L_n(w_1,w_2)L_n(w_3,w_4)\\
=&\frac{1}{4h_n^2}\Big(-\sqrt{2}k\Big(-\frac{x_1-y_1+x_2-y_2}{h_n}\Big)-\sqrt{2}k\Big(-\frac{x_1-y_1-(x_2-y_2)}{h_n}\Big)\\
&+k\Big(-\frac{x_1-y_2}{h_n}\Big)+k\Big(-\frac{y_1-x_2}{h_n}\Big)\Big)\\
&\Big(-\sqrt{2}k\Big(-\frac{x_3-y_3+x_4-y_4}{h_n}\Big)-\sqrt{2}k\Big(-\frac{x_3-y_3-(x_4-y_4)}{h_n}\Big)\\
&+k\Big(-\frac{x_3-y_4}{h_n}\Big)+k\Big(-\frac{y_3-x_4}{h_n}\Big)\Big),\ w_i=(x_i,y_i)\in\mathbb{R}^2,\ i=1,\dots,4. \qquad (10.8.4)
\end{aligned}
$$

Durch Betrachtung der 16 Produkte in (10.8.4) kann analog zu oben aufgrund der Beschränktheit der Dichten f und g und mit $\lim_{n\to\infty}nh_n=\infty$ $\mathrm{E}(|L_n(W_2,W_1)L_n(W_2,W_3)|)=o(n)$ gezeigt werden. So ist z.B.

$$
\begin{aligned}
&\mathrm{E}\left(\frac{\sqrt{2}}{4h_n^2}k\Big(-\frac{X_2-Y_2+X_1-Y_1}{h_n}\Big)k\Big(-\frac{X_2-Y_3}{h_n}\Big)\right)\\
=&\mathrm{E}\left(-\frac{\sqrt{2}}{4h_n^2}\int k\Big(-\frac{X_2-Y_2+X_1-Y_1}{h_n}\Big)k\Big(\frac{X_2-u}{h_n}\Big)f(u)\mathrm{d}u\right)\\
=&\mathrm{E}\left(\frac{\sqrt{2}}{4h_n}\int k\Big(-\frac{X_2-Y_2+X_1-Y_1}{h_n}\Big)k(u)f(uh_n+X_2)\mathrm{d}u\right).
\end{aligned}
$$

Ebenso kann durch Betrachtung der 16 Produkte in (10.8.4) analog zu oben aufgrund der Beschränktheit der Dichten f und g und mit $\lim_{n\to\infty}nh_n=\infty$ sowie mit Ungleichung

(10.8.3) $\mathrm{E}(\mathrm{E}(L_n(W_2,w_1)L_n(W_2,W_3))|^2_{w_1=W_1}) = o(n)$ für $n\to\infty$ gezeigt werden. Es ist z.B.

$$
\begin{aligned}
&\mathrm{E}\left(\mathrm{E}\left(-\frac{\sqrt{2}}{4h_n^2}k\Big(-\frac{X_2-Y_2+x_1-y_1}{h_n}\Big)k\Big(-\frac{X_2-Y_3}{h_n}\Big)\right)^2_{|w_1=W_1}\right)\\
=&\,\mathrm{E}\left(\mathrm{E}\left(-\frac{\sqrt{2}}{4h_n^2}\int k\Big(-\frac{X_2-Y_2+x_1-y_1}{h_n}\Big)k\Big(-\frac{X_2-u}{h_n}\Big)f(u)\mathrm{d}u\right)^2_{|w_1=W_1}\right)\\
=&\,\mathrm{E}\left(\mathrm{E}\left(-\frac{\sqrt{2}}{4h_n}\int k\Big(-\frac{X_2-Y_2+x_1-y_1}{h_n}\Big)k(u)f(uh_n+X_2)\mathrm{d}u\right)^2_{|w_1=W_1}\right)\\
=&\int\frac{1}{8h_n^2}\,\mathrm{E}\left(\int k\Big(-\frac{X_2-Y_2+v}{h_n}\Big)k(u)f(uh_n+X_2)\mathrm{d}u\right)^2 g(v)\mathrm{d}v\\
=&\int\frac{1}{8h_n^2}\int\int\int k\Big(-\frac{x_2-y_2+v}{h_n}\Big)k(u)f(uh_n+x_2)\mathrm{d}u\ell(x_1,y_1)\mathrm{d}x_1\mathrm{d}y_1\\
&\int\int\int k\Big(-\frac{\tilde{x}_2-\tilde{y}_2+v}{h_n}\Big)k(u)f(uh_n+\tilde{x}_2)\mathrm{d}u\ell(\tilde{x}_2,\tilde{y}_2)\mathrm{d}\tilde{x}_2\mathrm{d}\tilde{y}_2g(v)\mathrm{d}v\\
=&\int\frac{1}{8h_n}\int\int\int k(v)k(u)f(uh_n+x_2)\mathrm{d}u\ell(x_1,y_1)\mathrm{d}x_1\mathrm{d}y_1\\
&\int\int\int k\Big(-\frac{\tilde{x}_2-\tilde{y}_2+vh_n-(x_2-y_2)}{h_n}\Big)k(u)f(uh_n+\tilde{x}_2)\\
&\mathrm{d}u\ell(\tilde{x}_2,\tilde{y}_2)\mathrm{d}\tilde{x}_1\mathrm{d}\tilde{y}_1g\big(vh_n-(x_2-y_2)\big)\mathrm{d}v.
\end{aligned}
$$

Damit folgt $\mathrm{E}(h^2_{n,1}(W_1)) = o(n)$ für $n\to\infty$.

Wegen

$$
\begin{aligned}
h_{n,2}(w_1,w_2) =&\frac{1}{3}L_n(w_1,w_2)\,\mathrm{E}\big(L_n(w_1,W_3)\big)+\frac{1}{3}L_n(w_1,w_2)\,\mathrm{E}\big(L_n(w_2,W_3)\big)\\
&+\frac{1}{3}\,\mathrm{E}\big(L_n(W_3,w_2)L_n(W_3,w_1)\big),\ w_1,w_2\in\mathbb{R}^2,
\end{aligned}
$$

ist mit der Ungleichung (10.8.3)

$$
\begin{aligned}
\mathrm{E}\big(h^2_{n,2}(W_1,W_2)\big)\le&\frac{1}{3}\,\mathrm{E}\Big(L_n^2(w_1,w_2)\,\mathrm{E}\big(L_n(w_1,W_3)\big)^2_{|(w_1,w_2)=(W_1,W_2)}\Big)\\
&+\frac{1}{3}\,\mathrm{E}\Big(L_n^2(w_1,w_2)\,\mathrm{E}\big(L_n(w_2,W_3)\big)^2_{|(w_1,w_2)=(W_1,W_2)}\Big)\\
&+\frac{1}{3}\,\mathrm{E}\Big(\,\mathrm{E}\big(L_n(W_3,w_2)L_n(W_3,w_1)\big)^2_{|(w_1,w_2)=(W_1,W_2)}\Big)
\end{aligned}
$$

und demnach

$$
\mathrm{E}\big(h^2_{n,2}(W_1,W_2)\big)\le\frac{2}{3}c^2\,\mathrm{E}\big(L_n^2(W_1,W_2)\big)+\frac{1}{3}\,\mathrm{E}\Big(\,\mathrm{E}\big(L_n(W_3,w_2)L_n(W_3,w_1)\big)^2_{|(w_1,w_2)=(W_1,W_2)}\Big).
$$

Mit Teil a) ist $\mathrm{E}(L_n^2(W_1, W_2)) = o(n)$ für $n \to \infty$. Zudem kann analog zu oben aufgrund der Beschränktheit der Dichten f und g und mit $\lim_{n\to\infty} nh_n = \infty$ $\mathrm{E}(\mathrm{E}(L_n(W_3, w_2)L_n(W_3, w_1))^2|_{(w_1,w_2)=(W_1,W_2)}) = o(n^2)$ für $n \to \infty$ gezeigt werden. Es folgt $\mathrm{E}(h_{n,2}^2(W_1, W_2)) = o(n^2)$ für $n \to \infty$.

Schließlich ist

$$\begin{aligned} h_{n,3}(w_1, w_2, w_3) =& \frac{1}{3} L_n(w_1, w_2)L_n(w_1, w_3) + \frac{1}{3} L_n(w_2, w_1)L_n(w_2, w_3) \\ &+ \frac{1}{3} L_n(w_3, w_2)L_n(w_3, w_1),\ w_1, w_2, w_3 \in \mathbb{R}^2, \end{aligned}$$

also mit der Ungleichung (10.8.3)

$$\mathrm{E}\left(h_{n,3}^2(W_1, W_2, W_3)\right) \leq \mathrm{E}\left(L_n^2(W_1, W_2)L_n^2(W_1, W_3)\right),\ w_1, w_2, w_3 \in \mathbb{R}^2.$$

Es kann analog zu oben $\mathrm{E}(L_n^2(W_1, W_2)L_n^2(W_1, W_3)) = o(n^3)$ für $n \to \infty$ gezeigt werden, sodass $\mathrm{E}(h_{n,3}^2(W_1, W_2, W_3)) = o(n^3)$ für $n \to \infty$ folgt, und damit b).

□

10.8.2 Lemma: *Die Dichten $f, g, \tilde{f}, \tilde{g}$ seien beschränkt. Es sei die Kernfunktion k, die Folge von Bandbreiten $(h_n)_{n\in\mathbb{N}}$ sowie die Dichten $\tilde{f}$ und $\tilde{g}$ dergestalt, dass $\lim_{n\to\infty} nh_n = \infty$ sowie*[14]

$$\lim_{n\to\infty} \sqrt{n}\,\mathrm{E}\left(\tilde{f}(h_n K) - \tilde{f}(0)\right) = \lim_{n\to\infty} \sqrt{n}\,\mathrm{E}\left(\tilde{g}(h_n K) - \tilde{g}(0)\right) = 0$$

gilt, wobei K eine Zufallsvariable mit Dichte k sei. Dann gilt:

a) Falls $\sqrt{2}\tilde{g}(0) = \tilde{f}(0)$, so folgt $\lim_{n\to\infty} \sqrt{n}\,\mathrm{E}(L_n(W_1, W_2)) = 0$.

b) Falls $\sqrt{2}\tilde{g}(0) < \tilde{f}(0)$, existiert eine Konstante $c > 0$ mit $\lim_{n\to\infty} \mathrm{E}(L_n(W_1, W_2)) = c$.

c) Falls $\sqrt{2}\tilde{g}(0) > \tilde{f}(0)$, existiert eine Konstante $d > 0$ mit $\lim_{n\to\infty} \mathrm{E}(L_n(W_1, W_2)) = -d$.

Beweis: Es gelte zunächst $\sqrt{2}\tilde{g}(0) = \tilde{f}(0)$. Dann folgt

$$\begin{aligned} &\lim_{n\to\infty} \sqrt{n}\,\mathrm{E}\left(L_n(W_1, W_2)\right) \\ &= \lim_{n\to\infty} \sqrt{n}\Bigg(- \sqrt{2}\,\mathrm{E}\left(\frac{1}{h_n} k\Big(- \frac{X_1 - Y_1 + X_2 - Y_2}{h_n}\Big)\right) + \sqrt{2}\tilde{g}(0) - \tilde{f}(0) \\ &\quad + \mathrm{E}\left(\frac{1}{h_n} k\Big(- \frac{X_1 - Y_2}{h_n}\Big)\right)\Bigg) \\ &= \lim_{n\to\infty} \sqrt{n}\Bigg(- \sqrt{2}\int \frac{1}{h_n} k\Big(- \frac{z}{h_n}\Big)\tilde{g}(z)\mathrm{d}z + \sqrt{2}\tilde{g}(0) - \tilde{f}(0) + \int \frac{1}{h_n} k\Big(- \frac{z}{h_n}\Big)\tilde{f}(z)\mathrm{d}z\Bigg) \end{aligned}$$

[14] Zum Nachweis dieser Bedingung in Anwendungsfällen beachte Bemerkung 10.8.6.

$$= \lim_{n\to\infty} \sqrt{n}\Bigg(-\sqrt{2}\int k(z)\tilde{g}(h_n z)\mathrm{d}z + \sqrt{2}\tilde{g}(0) - \tilde{f}(0) + \int k(z)\tilde{f}(h_n z)\mathrm{d}z\Bigg)$$
$$= -\lim_{n\to\infty} \sqrt{2}\sqrt{n}\,\mathrm{E}\big(\tilde{g}(h_n K) - \tilde{g}(0)\big) + \lim_{n\to\infty} \sqrt{n}\,\mathrm{E}\big(\tilde{f}(h_n K) - \tilde{f}(0)\big) = 0.$$

Nun gelte $\sqrt{2}\tilde{g}(0) < \tilde{f}(0)$. Dann existiert eine Konstante $c > 0$ mit $\sqrt{2}\tilde{g}(0) - \tilde{f}(0) + c = 0$. Es folgt analog zu den Überlegungen am Anfang des Beweises

$$\lim_{n\to\infty} \mathrm{E}\big(L_n(W_1, W_2)\big) = -\lim_{n\to\infty} \sqrt{2}\,\mathrm{E}\big(\tilde{g}(h_n K) - \tilde{g}(0)\big) + \lim_{n\to\infty} \mathrm{E}\big(\tilde{f}(h_n K) - \tilde{f}(0)\big) + c = c.$$

Nun gelte $\sqrt{2}\tilde{g}(0) > \tilde{f}(0)$. Dann existiert eine Konstante $d > 0$ mit $\sqrt{2}\tilde{g}(0) - \tilde{f}(0) - d = 0$. Es folgt analog zu den Überlegungen am Anfang des Beweises

$$\lim_{n\to\infty} \mathrm{E}\big(L_n(W_1, W_2)\big) = -\lim_{n\to\infty} \sqrt{2}\,\mathrm{E}\big(\tilde{g}(h_n K) - \tilde{g}(0)\big) - \lim_{n\to\infty} \mathrm{E}\big(\tilde{f}(h_n K) - \tilde{f}(0)\big) - d = -d,$$

womit alles gezeigt ist. □

Es sei

$$M_n' := 2\sqrt{\mathrm{E}\Bigg(\bigg(\mathrm{E}\big(L_n(w, W_2)\big)_{|w=W_1} - \mathrm{E}\big(L_n(W_1, W_2)\big)\bigg)^2\Bigg)}.$$

10.8.3 Lemma: *Die Dichten $f, g, \tilde{f}, \tilde{g}$ seien beschränkt. Es sei die Kernfunktion k, die Folge von Bandbreiten $(h_n)_{n\in\mathbb{N}}$ sowie die Dichten f und g dergestalt, dass $\lim_{n\to\infty} nh_n = \infty$ sowie*

$$\lim_{n\to\infty} \sqrt{n}\,\mathrm{E}\big(\tilde{f}(h_n K) - \tilde{f}(0)\big) = \lim_{n\to\infty} \sqrt{n}\,\mathrm{E}\big(\tilde{g}(h_n K) - \tilde{g}(0)\big) = 0$$

gilt, wobei K eine Zufallsvariable mit Dichte k sei. Ferner gelte[15]

$$\lim_{n\to\infty} \mathrm{E}\big(f(h_n K + x) - f(x)\big) = \lim_{n\to\infty} \mathrm{E}\big(g(h_n K + z) - g(z)\big) = 0$$
$$\text{für alle } x \in \mathbb{R} \text{ mit } f(x) \neq 0 \text{ und für alle } z \in \mathbb{R} \text{ mit } g(z) \neq 0.$$

Zudem sei

$$-\sqrt{2}g(\zeta_1) + \frac{1}{2}f(X_1) + \frac{1}{2}f(Y_1) \neq 0 \;\; P\text{-f.s.}$$

erfüllt. Dann gilt

$$\lim_{n\to\infty} M_n/\sqrt{n} = 0 \;\; \text{in Wahrscheinlichkeit.}$$

[15]Zum Nachweis dieser Bedingung in Anwendungsfällen beachte Bemerkung 10.8.7.

Gilt zusätzlich $\sqrt{2}\tilde{g}(0) = \tilde{f}(0)$, *so gilt*

$$\liminf_{n\to\infty} M'_n > 0 \text{ sowie } \lim_{n\to\infty} M_n/M'_n = 1 \text{ in Wahrscheinlichkeit.}$$

Beweis: Wegen

$$\mathrm{E}\left(\mathrm{E}\left(L_n(w, W_2)\right)^2_{|w=W_1}\right) = \mathrm{E}\left(\tilde{L}_n(W_1, W_2, W_3)\right)$$

ist

$$M'_n = 2\sqrt{\mathrm{E}\left(\tilde{L}_n(W_1, W_2, W_3)\right) - \mathrm{E}\left(L_n(W_1, W_2)\right)^2}.$$

Aus Lemma 10.8.1 a) und b) folgt damit

$$\lim_{n\to\infty} \frac{M'_n}{\sqrt{n}} = 0.$$

Nun soll noch

$$\lim_{n\to\infty} (M_n - M'_n) = 0 \text{ in Wahrscheinlichkeit}$$

gezeigt werden, womit dann insgesamt

$$\lim_{n\to\infty} \frac{M_n}{\sqrt{n}} = \lim_{n\to\infty} \left(\frac{M_n}{\sqrt{n}} - \frac{M'_n}{\sqrt{n}} + \frac{M'_n}{\sqrt{n}}\right) = \lim_{n\to\infty} \frac{M_n - M'_n}{\sqrt{n}} + \lim_{n\to\infty} \frac{M'_n}{\sqrt{n}} = 0$$
$$\text{in Wahrscheinlichkeit}$$

folgt. Mit

$$\begin{aligned}
&\frac{1}{n(n-1)(n-2)} \sum_{\substack{1\le i_1, i_2, i_3 \le n \\ i_l \ne i_m, l \ne m}} L_n(W_{i_1}, W_{i_2}) L_n(W_{i_1}, W_{i_3}) \\
=&\frac{1}{n(n-1)(n-2)} \sum_{\substack{1\le i_1, i_2, i_3 \le n \\ i_l \ne i_m, l \ne m}} \tilde{L}_n(W_{i_1}, W_{i_2}, W_{i_3})
\end{aligned}$$

ist

$$\begin{aligned}
\frac{1}{4}(M_n^2 - {M'_n}^2) =& \frac{1}{n(n-1)(n-2)} \sum_{\substack{1\le i_1, i_2, i_3 \le n \\ i_l \ne i_m, l \ne m}} \tilde{L}_n(W_{i_1}, W_{i_2}, W_{i_3}) - \mathrm{E}\left(\tilde{L}_n(W_1, W_2, W_3)\right) \\
&- \left(\frac{1}{n^2 - n} \sum_{1 \le i \ne j \le n} L_n(W_i, W_j)\right)^2 + \mathrm{E}\left(L_n(W_1, W_2)\right)^2.
\end{aligned}$$

Aufgrund von Lemma 10.8.1 Teil b) greift das schwache Gesetz der großen Zahlen für U-Statistiken, Satz A.1.6 aus dem Anhang, und es ist

$$\lim_{n\to\infty}\left(\frac{1}{n(n-1)(n-2)}\sum_{\substack{1\le i_1,i_2,i_3\le n\\ i_l\neq i_m, l\neq m}}\tilde{L}_n(W_{i_1},W_{i_2},W_{i_3})-\mathrm{E}\left(\tilde{L}_n(W_1,W_2,W_3)\right)\right)=0$$

in Wahrscheinlichkeit für $n\to\infty$.

Ebenso greift aufgrund von Lemma 10.8.1 Teil a) das schwache Gesetz der großen Zahlen für U-Statistiken, Satz A.1.6 aus dem Anhang, und mit Hilfe des dritten binomischen Lehrsatzes und Lemma 10.8.1 Teil a) ist

$$\left(\frac{1}{n^2-n}\sum_{1\le i\neq j\le n}L_n(W_i,W_j)\right)^2-\mathrm{E}\left(L_n(W_1,W_2)\right)^2\longrightarrow 0$$

in Wahrscheinlichkeit für $n\to\infty$.

Damit folgt insgesamt $\lim_{n\to\infty}(M_n-M_n')=0$ in Wahrscheinlichkeit.

Nun gelte zusätzlich $\sqrt{2}\tilde{g}(0)=\tilde{f}(0)$. Es ist

$$\begin{aligned}
&\mathrm{E}\left(L_n(w,W_2)\right)_{|w=W_1}\\
=&-\sqrt{2}\,\mathrm{E}\left(\frac{1}{h_n}k\Big(-\frac{z+\zeta_2}{h_n}\Big)\right)_{|z=\zeta_1}\\
&+\mathrm{E}\left(\frac{1}{2h_n}k\Big(-\frac{x-Y_2}{h_n}\Big)\right)_{|x=X_1}+\mathrm{E}\left(\frac{1}{2h_n}k\Big(-\frac{y-X_2}{h_n}\Big)\right)_{|y=Y_1}\\
=&-\sqrt{2}\int\frac{1}{h_n}k\Big(-\frac{\zeta_1+u}{h_n}\Big)g(u)\mathrm{d}u+\int\frac{1}{2h_n}k\Big(-\frac{X_1-u}{h_n}\Big)f(u)\mathrm{d}u\\
&+\int\frac{1}{2h_n}k\Big(-\frac{Y_1-u}{h_n}\Big)f(u)\mathrm{d}u\\
=&-\sqrt{2}\int k(u)g(h_nu+\zeta_1)\mathrm{d}u+\frac{1}{2}\int k(u)f(h_nu+X_1)\mathrm{d}u+\frac{1}{2}\int k(u)f(h_nu+Y_1)\mathrm{d}u\\
=&-\sqrt{2}\,\mathrm{E}\left(g(h_nK+z)\right)_{|z=\zeta_1}+\frac{1}{2}\,\mathrm{E}\left(f(h_nK+x)\right)_{|x=X_1}+\frac{1}{2}\,\mathrm{E}\left(f(h_nK+y)\right)_{|y=Y_1}\\
=&-\sqrt{2}\,\mathrm{E}\left(g(h_nK+z)\right)_{|z=\zeta_1}+\sqrt{2}g(\zeta_1)-\sqrt{2}g(\zeta_1)\\
&+\frac{1}{2}\,\mathrm{E}\left(f(h_nK+x)\right)_{|x=X_1}-\frac{1}{2}f(X_1)+\frac{1}{2}f(X_1)\\
&+\frac{1}{2}\,\mathrm{E}\left(f(h_nK+y)\right)_{|y=Y_1}-\frac{1}{2}f(Y_1)+\frac{1}{2}f(Y_1)
\end{aligned}$$

und daher mit den Voraussetzungen

$$\lim_{n\to\infty}\mathrm{E}\left(L_n(w,W_2)\right)_{|w=W_1}=-\sqrt{2}g(\zeta_1)+\frac{1}{2}f(X_1)+\frac{1}{2}f(Y_1)\neq 0\ P\text{-f.s.}$$

erfüllt. Damit folgt mit dem Lemma von Fatou

$$\begin{aligned}\liminf_{n\to\infty} \mathrm{E}\Big(\mathrm{E}\left(L_n(w,W_2)\right)^2_{|w=W_1}\Big) &\geq \mathrm{E}\Big(\liminf_{n\to\infty} \mathrm{E}\left(L_n(w,W_2)\right)^2_{|w=W_1}\Big) \\ &= \mathrm{E}\Big(\big(-\sqrt{2}g(\zeta_1)+\frac{1}{2}f(X_1)+\frac{1}{2}f(Y_1)\big)^2\Big) > 0.\end{aligned}$$

Aufgrund der genannten Voraussetzungen gilt mit Lemma 10.8.2 $\lim_{n\to\infty} \mathrm{E}(L_n(W_1,W_2)) = 0$, also insgesamt $\liminf_{n\to\infty} {M'_n}^2 > 0$, und damit auch

$$\liminf_{n\to\infty} M'_n > 0.$$

Mit $\lim_{n\to\infty}(M_n - M'_n) = 0$ in Wahrscheinlichkeit, was bereits gezeit wurde, folgt insgesamt

$$\lim_{n\to\infty}\left(\frac{M_n}{M'_n}-1\right) = \lim_{n\to\infty}\left(\frac{M_n}{M'_n}-\frac{M'_n}{M'_n}\right) = \lim_{n\to\infty}\left(\frac{M_n-M'_n}{M'_n}\right) = 0 \text{ in Wahrscheinlichkeit,}$$

womit alles gezeigt ist. □

10.8.4 Satz: *Die Dichten $f, g, \tilde{f}, \tilde{g}$ seien beschränkt. Es sei die Kernfunktion k, die Folge von Bandbreiten $(h_n)_{n\in\mathbb{N}}$ sowie die Dichten f und g dergestalt, dass $\lim_{n\to\infty} nh_n = \infty$ und*

$$\lim_{n\to\infty}\sqrt{n}\,\mathrm{E}\left(\tilde{f}(h_nK)-\tilde{f}(0)\right) = \lim_{n\to\infty}\sqrt{n}\,\mathrm{E}\left(\tilde{g}(h_nK)-\tilde{g}(0)\right) = 0$$

gilt, wobei K eine Zufallsvariable mit Dichte k sei. Ferner gelte

$$\lim_{n\to\infty}\mathrm{E}\left(f(h_nK+x)-f(x)\right) = \lim_{n\to\infty}\mathrm{E}\left(g(h_nK+z)-g(z)\right) = 0$$
$$\textit{für alle } x\in\mathbb{R} \textit{ mit } f(x)\neq 0 \textit{ und für alle } z\in\mathbb{R} \textit{ mit } g(z)\neq 0.$$

Zudem sei

$$-\sqrt{2}g(\zeta_1)+\frac{1}{2}f(X_1)+\frac{1}{2}f(Y_1)\neq 0 \;\; P\text{-}f.s.$$

erfüllt, und es gelte $\sqrt{2}\tilde{g}(0)=\tilde{f}(0)$. Dann folgt

$$E_n^i \xrightarrow{\mathrm{v}} \mathrm{N}(0,1),\;\; i=1,2, \;\; \textit{und } E_n^3 \xrightarrow{\mathrm{v}} \mathrm{HN}(0,1) \textit{ für } n\to\infty,$$

wobei $\mathrm{HN}(0,1)$ die Verteilung des Betrages einer standardnormalverteilten Zufallsvariable bezeichne.

Beweis: Nur für $i = 1$; der Rest folgt dann sofort. Bei dem Ausdruck

$$\sqrt{n}\bigg(\frac{1}{n^2-n}\sum_{1\leq i\neq j\leq n} L_n(W_i,W_j) - \mathrm{E}\left(L_n(W_1,W_2)\right)\bigg)$$

handelt es sich um eine U-Statsitik der Kernlänge 2 mit einem von n abhängigen Kern. Es folgt aufgrund von Lemma 10.8.1 Teil a) mit einem zentralen Grenzwertsatzes für U-Statistiken mit einem von n abhängigen Kern (siehe Korollar A.1.5 im Anhang)

$$\frac{\sqrt{n}\Big(\frac{1}{n^2-n}\sum_{1\le i\ne j\le n} L_n(W_i,W_j) - \mathrm{E}\left(L_n(W_1,W_2)\right)\Big)}{M_n'} \xrightarrow{\mathrm{v}} \mathrm{N}(0,1) \text{ für } n\to\infty.$$

Mit Lemma 10.8.2, Lemma 10.8.3 und dem Lemma von Slutsky folgt daraus

$$\begin{aligned}
E_n^1 &= \frac{\sqrt{n}\frac{1}{n^2-n}\sum_{1\le i\ne j\le n} L_n(W_i,W_j)}{M_n} \\
&= \frac{\sqrt{n}\frac{1}{n^2-n}\sum_{1\le i\ne j\le n} L_n(W_i,W_j)}{M_n'}\frac{M_n'}{M_n} \\
&= \frac{\sqrt{n}\Big(\frac{1}{n^2-n}\sum_{1\le i\ne j\le n} L_n(W_i,W_j) - \mathrm{E}\left(L_n(W_1,W_2)\right)\Big)}{M_n'}\frac{M_n'}{M_n} + \frac{\sqrt{n}\,\mathrm{E}\left(L_n(W_1,W_2)\right)}{M_n'} \\
&\xrightarrow{\mathrm{v}} \mathrm{N}(0,1) \text{ für } n\to\infty,
\end{aligned}$$

also die Behauptung. □

Für ein gegebenes Testniveau $\alpha \in (0,1)$ schlagen wir für das Testproblem H_i gegen K_i, $i=1,2$, den Test

$$\text{„verwirf } \mathrm{H}_i, \text{ falls } E_n^i \ge u_{1-\alpha}\text{“}, \ i=1,2,$$

vor, wobei $u_{1-\alpha}$ das $(1-\alpha)$-Quantil der Standardnormalverteilung sei. Für das Testproblem H_3 gegen K_3 schlagen wir dagegen den Test

$$\text{„verwirf } \mathrm{H}_3, \text{ falls } E_n^3 \ge u_{1-\alpha/2}\text{“}$$

vor. Zur Begründung dieses Vorschlags können wir mit Hilfe der bisher erzielten Resultate folgendes Korollar formulieren:

10.8.5 Korollar: *Die Dichten $f, g, \tilde{f}, \tilde{g}$ seien beschränkt. Es sei die Kernfunktion k, die Folge von Bandbreiten $(h_n)_{n\in\mathbb{N}}$ sowie die Dichten f und g dergestalt, dass* $\lim_{n\to\infty} nh_n = \infty$ *und*

$$\lim_{n\to\infty}\sqrt{n}\,\mathrm{E}\left(\tilde{f}(h_nK) - \tilde{f}(0)\right) = \lim_{n\to\infty}\sqrt{n}\,\mathrm{E}\left(\tilde{g}(h_nK) - \tilde{g}(0)\right) = 0$$

gilt, wobei K eine Zufallsvariable mit Dichte k sei. Ferner gelte

$$\lim_{n\to\infty}\mathrm{E}\left(f(h_nK+x) - f(x)\right) = \lim_{n\to\infty}\mathrm{E}\left(g(h_nK+z) - g(z)\right) = 0$$
für alle $x\in\mathbb{R}$ mit $f(x)\ne 0$ und für alle $z\in\mathbb{R}$ mit $g(z)\ne 0$.

Zudem sei

$$-\sqrt{2}g(\zeta_1) + \frac{1}{2}f(X_1) + \frac{1}{2}f(Y_1) \neq 0 \ P\text{-f.s.}$$

erfüllt. Dann handelt es sich bei dem auf der Testgröße E_n^i basierenden vorgeschlagenen Test für das Tesproblem H_i gegen K_i, $i = 1, 2$, um einen Test, der für jedes $\alpha \in (0, 1)$ das Testniveau α asymptotisch einhält und zudem konsistent ist. Bei dem auf der Testgröße E_n^3 basierenden vorgeschlagenen Test für das Tesproblem H_3 gegen K_3 handelt es sich darüber hinaus um einen Test, der für jedes $\alpha \in (0, 1)$ das Testniveau α asymptotisch exakt einhält und zudem konsistent ist.

Beweis: Nur für $i = 1$; sonst analog. Ohne Einschränkung nehmen wir hier stets $M_n > 0$ an. Es gelte zunächst $e_{T,T'}^{\mathrm{P}} = 1$. Mit Satz 10.8.4 folgt, dass der Test in diesem Fall das Testniveau asymptotisch exakt einhält. Nun gelte $e_{T,T'}^{\mathrm{P}} < 1$. Nach Lemma 10.8.2 existiert eine Konstante $d > 0$ mit

$$\lim_{n\to\infty} \mathrm{E}\left(L_n(W_1, W_2)\right) = -d,$$

und somit folgt aufgrund von Lemma 10.8.1 Teil b) mit einem schwachen Gesetz der großen Zahlen für U-Statistiken mit einem von n abhängigen Kern, Satz A.1.6 aus dem Anhang, sowie mit Lemma 10.8.3, für $\alpha \in (0, 1)$ beliebig

$$\begin{aligned}
&P(E_n^1 \geq u_{1-\alpha}) \\
=&P\left(\frac{1}{n^2 - n} \sum_{1 \leq i \neq j \leq n} L_n(W_i, W_j) - \mathrm{E}\left(L_n(W_1, W_2)\right) \geq \frac{M_n}{\sqrt{n}} u_{1-\alpha} - \mathrm{E}\left(L_n(W_1, W_2)\right)\right) \\
\leq&P\left(\left|\frac{1}{n^2 - n} \sum_{1 \leq i \neq j \leq n} L_n(W_i, W_j) - \mathrm{E}\left(L_n(W_1, W_2)\right)\right| \geq \frac{M_n}{\sqrt{n}} u_{1-\alpha} - \mathrm{E}\left(L_n(W_1, W_2)\right)\right) \\
&\longrightarrow 0 < \alpha \text{ für } n \to \infty.
\end{aligned}$$

Somit hält der Test insgesamt das Testniveau asymptotisch ein. Nun gelte $e_{T,T'}^{\mathrm{P}} > 1$. Dann existiert nach Lemma 10.8.2 eine Konstante $c > 0$ mit

$$\lim_{n\to\infty} \mathrm{E}\left(L_n(W_1, W_2)\right) = c,$$

und somit folgt wie im Beweis von Lemma 10.8.3 aufgrung von Lemma 10.8.1 Teil b) mit einem schwachen Gesetz der großen Zahlen für U-Statistiken mit einem von n abhängigen Kern, sowie mit Lemma 10.8.3, für $\alpha \in (0, 1)$ beliebig

$$\begin{aligned}
&P(E_n^1 < u_{1-\alpha}) = P(-E_n^1 > -u_{1-\alpha}) \\
=&P\left(-\frac{1}{n^2 - n} \sum_{1 \leq i \neq j \leq n} L_n(W_i, W_j) + \mathrm{E}\left(L_n(W_1, W_2)\right) > -\frac{M_n}{\sqrt{n}} u_{1-\alpha} + \mathrm{E}\left(L_n(W_1, W_2)\right)\right) \\
\leq&P\left(\left|\frac{1}{n^2 - n} \sum_{1 \leq i \neq j \leq n} L_n(W_i, W_j) - \mathrm{E}\left(L_n(W_1, W_2)\right)\right| > -\frac{M_n}{\sqrt{n}} u_{1-\alpha} + \mathrm{E}\left(L_n(W_1, W_2)\right)\right) \\
&\longrightarrow 0 \text{ für } n \to \infty,
\end{aligned}$$

d.h. der Test ist konsistent. □

10.8.6 Bemerkung: Um die Bedingung

$$\lim_{n\to\infty} \sqrt{n}\,\mathrm{E}\left(\tilde{f}(h_n K) - \tilde{f}(0)\right) = \lim_{n\to\infty} \sqrt{n}\,\mathrm{E}\left(\tilde{g}(h_n K) - \tilde{g}(0)\right) = 0$$

aus Lemma 10.8.2, Lemma 10.8.3, Satz 10.8.4 und Korollar 10.8.5 nachzuweisen, kann folgende Überlegung von Nutzen sein: Wir nehmen an, dass die Dichten $\tilde{f}$ und $\tilde{g}$ in einer offenen Umgebung U um $0 \in \mathbb{R}$ Hölder-stetig seien, d.h.

$$\forall\, x, y \in U : |\tilde{f}(x) - \tilde{f}(y)| \leq C|x-y|^{\alpha},\ \ |\tilde{g}(x) - \tilde{g}(y)| \leq C'|x-y|^{\alpha'},$$

mit Konstanten $C, C' > 0$ und $\alpha, \alpha' \in (0,1]$. Ferner sei die Folge von Bandbreiten durch $(1/n^{\beta})_{n\in\mathbb{N}}$, mit $1/2 - \min\{\alpha,\alpha'\}\beta < 0$, gegeben, und es sei die Kernfunktion k so gewählt, dass $\mathrm{E}(|K|^{\max\{\alpha,\alpha'\}}) < \infty$ gelte, wobei wie gesagt die Zufallsvariable K die Dichte k habe. Dann folgt

$$\begin{aligned}\left|\sqrt{n}\,\mathrm{E}\left(\tilde{f}(h_n K) - \tilde{f}(0)\right)\right| &\leq \sqrt{n}\,\mathrm{E}\left(|\tilde{f}(h_n K) - \tilde{f}(0)|\right) \leq C\sqrt{n}\,\mathrm{E}(|h_n K|^{\alpha}) \\ &= Cn^{1/2-\alpha\beta}\,\mathrm{E}(|K|^{\alpha}) \longrightarrow 0 \text{ für } n \to \infty.\end{aligned}$$

Analog erhält man

$$\lim_{n\to\infty} \left|\sqrt{n}\,\mathrm{E}\left(\tilde{g}(h_n K) - \tilde{g}(0)\right)\right| = 0,$$

sodass die Bedingung erfüllt ist. Im Fall der Lipschitz-Stetigkeit von $\tilde{f}$ und $\tilde{g}$ in einer offenen Umgebung U um $0 \in \mathbb{R}$, d.h. $\alpha = \alpha' = 1$, sollte $\beta > 1/2$ gelten.

10.8.7 Bemerkung: Die Bedingung

$$\begin{gathered}\lim_{n\to\infty} \mathrm{E}\left(f(h_n K + x) - f(x)\right) = \lim_{n\to\infty} \mathrm{E}\left(g(h_n K + z) - g(z)\right) = 0 \\ \text{für alle } x \in \mathbb{R} \text{ mit } f(x) \neq 0 \text{ und für alle } z \in \mathbb{R} \text{ mit } g(z) \neq 0\end{gathered}$$

aus Lemma 10.8.3, Satz 10.8.4 und Korollar 10.8.5 kann mit Hilfe entsprechender Hölder-Bedingungen an f und g und entsprechenden Bedingungen an die Folge von Bandbreiten und die Kerndichte analog zu Bemerkung 10.8.6 nachgewiesen werden. Allerdings genügt es zum Nachweis dieser Bedingung schon, dass die Kerndichte $\lim_{|x|\to\infty} |x|k(x) = 0$ erfüllt, und dass f und g stetig sind. Dann folgt

$$\lim_{n\to\infty} \mathrm{E}\left(f(h_n K + x)\right) = \lim_{n\to\infty} \int k(y)f(h_n y + x)\mathrm{d}y = f(x) \int k(y)\mathrm{d}y = f(x),\ \ x \in \mathbb{R},$$

und analog

$$\lim_{n\to\infty} \mathrm{E}\left(g(h_n K + z)\right) = g(z),\ \ z \in \mathbb{R};$$

siehe hierzu Satz A.1.7 im Anhang.

10.8.8 Bemerkung: Da für alle $\vartheta \in \mathbb{R}$

$$\frac{1}{2}\Big(\frac{\tilde{f}(0)}{\tilde{g}(0)}\Big)^2 = \frac{1}{2}\Big(\frac{\int f(t)^2 \mathrm{d}t}{\int g(t)^2 \mathrm{d}t}\Big)^2 = \frac{1}{2}\Big(\frac{\int f(t)^2 \mathrm{d}t}{\int g(t+\vartheta)^2 \mathrm{d}t}\Big)^2 = \frac{1}{2}\Big(\frac{\int f(t)^2 \mathrm{d}t}{\int \gamma_\vartheta(t)^2 \mathrm{d}t}\Big)^2$$

gilt, wobei $\gamma_\vartheta(\cdot) := g(\cdot + \vartheta)$ die Dichte bezüglich des Lebesgue-Borelschen Maßes von ζ_1 bei zugrunde liegendem Parameter $\vartheta \in \mathbb{R}$ bezeichne, können in analoger Weise Tests für die genannten Testprobleme basierend auf Beobachtungen im verbundenen Stichprobenfall mit $\vartheta \in \mathbb{R}$ konstruiert und studiert werden. Dazu bieten sich Tests basierend auf den Kerndichteschätzern

$$f_n(t) := \frac{1}{nh_n}\sum_{i=1}^n k\Big(\frac{t-X_i}{h_n}\Big),\ \gamma_n(t) := \frac{1}{nh_n}\sum_{i=1}^n k\Big(\frac{t-\zeta_i}{h_n}\Big),\ t \in \mathbb{R},$$

an. Entsprechende asymptotischen Eigenschaften von Testgrößen sowie die Eigenschaften der zugehörigen Tests (Niveaueinhaltung und Konsistenz) basierend auf diesen Kerndichteschätzern lassen sich unter den gleichen Bedingungen nachweisen.

A Anhang

A.1 Hilfsaussagen

Das folgende Lemma ist als Pratts Version des Satzes von der majorisierten Konvergenz bekannt; das z.B. im Buch von Gänssler und Stute ([8]) nachgeschlagen werden kann.

A.1.1 Lemma: *Auf einem Maßraum $(\Omega, \mathfrak{A}, \mu)$ seien Folgen von Funktionen $f_n : (\Omega, \mathfrak{A}) \to (\mathbb{R}, \mathfrak{B})$, $g_n : (\Omega, \mathfrak{A}) \to (\mathbb{R}, \mathfrak{B})$, $n \in \mathbb{N}$, gegeben. Für diese Folgen von Funktionen gelte $\lim_{n\to\infty} f_n = f$ μ-f.ü. und $\lim_{n\to\infty} g_n = g$ μ-f.ü. mit Funktionen $f : (\Omega, \mathfrak{A}) \to (\mathbb{R}, \mathfrak{B})$, $g : (\Omega, \mathfrak{A}) \to (\mathbb{R}, \mathfrak{B})$. Zudem gelte $|f_n| \le g_n$ μ-f.ü. für alle $n \in \mathbb{N}$, $\int g_n \mathrm{d}\mu < \infty$ für alle $n \in \mathbb{N}$, $\int g \mathrm{d}\mu < \infty$ und $\lim_{n\to\infty} \int g_n \mathrm{d}\mu = \int g \mathrm{d}\mu$. Dann gilt*

$$\lim_{n\to\infty} \int f_n \mathrm{d}\mu = \int f \mathrm{d}\mu.$$

Beweis: Siehe Lemma 1.11.16 in [8]. □

Das nachfolgende Lemma kann als ein schwaches Gesetz der großen Zahlen für Dreiecksschemata bezeichnet werden.

A.1.2 Lemma: *Es sei $n \in \mathbb{N}$ und $X_{n,1}, ..., X_{n,r_n}$ ein Dreiecksschema von unabhängigen Zufallsvariablen, je mit derselben Verteilung F_n mit $F_n \rightsquigarrow F$. Ferner gelte für alle $n \in \mathbb{N}$ $\int |x| \,\mathrm{d}F_n(x) < \infty$, $\int |x| \,\mathrm{d}F(x) < \infty$ und es gelte $\int |x| \,\mathrm{d}F_n(x) \to \int |x| \,\mathrm{d}F(x)$ und $r_n \to \infty$ für $n \to \infty$. Dann gilt mit $\mu := \int x \,\mathrm{d}F(x)$*

$$\frac{1}{r_n}\sum_{k=1}^{r_n} X_{n,k} \xrightarrow{P} \mu \text{ für } n \to \infty.$$

Beweis: Sei φ_n die Fourier-Transformierte von F_n, $\mu_n := \int x \,\mathrm{d}F_n(x)$ und $\tilde{\varphi}_n$ die Fourier-Transformierte von $\frac{1}{r_n}\sum_{k=1}^{r_n} X_{n,k}$. Für $z \in [0,\infty)$ ist

$$\begin{aligned}\tilde{\varphi}_n(z) = \left(\varphi_n\Big(\frac{z}{r_n}\Big)\right)^{r_n} &= \mathrm{E}\left(1 + i\frac{z}{r_n}X_{n,1} + e^{i\frac{z}{r_n}X_{n,1}} - 1 - i\frac{z}{r_n}X_{n,1}\right)^{r_n} \\ &= \left(1 + i\frac{z}{r_n}\mu_n + \frac{1}{r_n}\int_0^z \left(\varphi_n'\Big(\frac{t}{r_n}\Big) - \varphi_n'(0)\right)\mathrm{d}t\right)^{r_n}.\end{aligned}$$

O.B.d.A. gelte $X_{n,1} \to X$ f.s. mit einem $X \sim F$ (ansonsten wähle Zufallsvariablen $\tilde{X}_{n,1} \sim F_n, \tilde{X} \sim F$ auf einem geeigneten Wahrscheinlichkeitsraum mit der gewünschten Eigenschaft). Es gilt für alle $z \in [0,\infty)$ mit Pratts Version des Satzes der majorisierten Konvergenz (Lemma A.1.1)

$$\begin{aligned}\left|\int_0^z \left(\varphi_n'\Big(\frac{t}{r_n}\Big) - \varphi_n'(0)\right)\mathrm{d}t\right| &= \left|\int_0^z \mathrm{E}\left(iX_{n,1}e^{i\frac{t}{r_n}X_{n1}} - iX_{n,1}\right)\mathrm{d}t\right| \\ &\le \int_0^z \mathrm{E}(|X_{n,1}||e^{i\frac{t}{r_n}X_{n,1}} - 1|)\,\mathrm{d}t \longrightarrow 0 \text{ für } n \to \infty.\end{aligned}$$

Dann folgt mit $\mu_n \to \mu$ für $n \to \infty$

$$\tilde{\varphi}_n(z) \longrightarrow e^{i\mu z} \text{ für } n \to \infty.$$

Analog lässt sich für $z \in (-\infty, 0)$ $\tilde{\varphi}_n(z) \longrightarrow e^{i\mu z}$ für $n \to \infty$ zeigen. Hieraus folgt

$$\frac{1}{r_n}\sum_{k=1}^{r_n} X_{n,k} \xrightarrow{P} \mu \text{ für } n \to \infty,$$

also die Behauptung. □

Die folgende Aussage ist nützlich im Zusammenhang mit der Definitheit der Differenz $A - B$ zweier quadratischer symmetrischer positiv definiter Matrizen A und B. Sie kann in [11] nachgeschlagen werden. Dabei bedeute die Schreibweise $A - B \ge 0$, dass die Matrix $A - B$ positiv semidefinit ist, und die Schreibweise $A - B > 0$ bedeute entsprechend, dass die Matrix $A - B$ positiv definit ist.

A.1.3 Lemma: *Es seien $A, B \in \mathbb{R}^{h\times h}$, $h \in \mathbb{N}$, zwei symmetrische positv definite Matrizen. Dann gilt die Äquivalenz*

$$A - B \begin{Bmatrix} \ge \\ > \end{Bmatrix} 0 \;\Leftrightarrow\; B^{-1} - A^{-1} \begin{Bmatrix} \ge \\ > \end{Bmatrix} 0.$$

Beweis: Siehe Theorem 8.8 und die anschließende Bemerkung in [11]. □

Es sei $d \in \mathbb{N}$. Auf einem Wahrscheinlichkeitsraum $(\Omega, \mathfrak{A}, P)$ seien für $n \geq 2$ durch $W_1, \ldots, W_n$ unabhängige und identisch verteilte $\mathbb{R}^d$-wertige Zufallsvektoren gegeben. Zudem sei $L_n : (\mathbb{R}^{2d}, \mathfrak{B}^{2d}) \to (\mathbb{R}, \mathfrak{B})$ eine von n abhängige Kernfunktion der Kernlänge 2, die $L_n(x, y) = L_n(y, x)$ für alle $x, y \in \mathbb{R}^d$ erfüllt. Dann ist

$$\frac{1}{n^2 - n} \sum_{1 \leq i \neq j \leq n} L_n(W_i, W_j)$$

die zugehörige U-Statistik der Kernlänge 2 mit einem von n abhängigen Kern, und

$$\sqrt{n}\Big(\frac{1}{n^2 - n} \sum_{1 \leq i \neq j \leq n} L_n(W_i, W_j) - \mathrm{E}\left(L_n(W_1, W_2) \right) \Big).$$

Es gelte $\mathrm{E}(|L_n(W_1, W_2)|) < \infty$ und $\mathrm{E}(\mathrm{E}(L_n(w, W_2))^2|_{w=W_1}) < \infty$ für alle $n \in \mathbb{N}$. Dann gilt der folgende zentrale Grenzwertsatz für U-Statistiken mit einem von n abhängigen Kern aus der Arbeit von Jammalamadaka und Janson (Theorem 2.1 in [17]).

A.1.4 Satz: *Es sei*

$$\begin{aligned} \sigma_n^2 :=& \frac{1}{2} n^2 \Big(\mathrm{E}\left(L_n(W_1, W_2)^2 \right) - \mathrm{E}\left(L_n(W_1, W_2) \right)^2 \Big) \\ &+ n^3 \, \mathrm{E}\bigg(\Big(\mathrm{E}\left(L_n(w, W_2) \right)_{|w=W_1} - \mathrm{E}\left(L_n(W_1, W_2) \right) \Big)^2 \bigg). \end{aligned}$$

Ferner sei $\sigma_n := \sqrt{\sigma_n^2} > 0$ *und* L_n *beschränkt für alle* $n \in \mathbb{N}$. *Zudem gelte* $\sup_{x,y \in \mathbb{R}^d} |L_n(x, y)| = o(\sigma_n)$ *und* $\sup_{x \in \mathbb{R}^d} \mathrm{E}(|L_n(x, W_1)|) = o(\sigma_n / n)$ *für* $n \to \infty$. *Dann gilt*

$$\frac{\frac{1}{2} \sum_{1 \leq i \neq j \leq n} L_n(W_i, W_j) - \binom{n}{2} \mathrm{E}\left(L_n(W_1, W_2) \right)}{\sigma_n} \xrightarrow{\mathrm{v}} \mathrm{N}(0, 1) \text{ für } n \to \infty.$$

Beweis: Dies ist Theorem 2.1 in [17]. □

Wir definieren

$$M_n := 2 \sqrt{\mathrm{E}\bigg(\Big(\mathrm{E}\left(L_n(w, W_2) \right)_{|w=W_1} - \mathrm{E}\left(L_n(W_1, W_2) \right) \Big)^2 \bigg)}.$$

Dann können wir folgendes Korollar angeben:

A.1.5 Korollar: *Es sei* $M_n > 0$ *und* $\liminf_{n \to \infty} M_n > 0$. *Ferner sei* L_n *beschränkt für alle* $n \in \mathbb{N}$, $\sup_{x,y \in \mathbb{R}^d} |L_n(x, y)| = o(n^{3/2})$, $\sup_{x \in \mathbb{R}^d} \mathrm{E}(|L_n(x, W_1)|) = o(n^{1/2})$, $\mathrm{E}(L_n(W_1, W_2)^2) = o(n)$ *für* $n \to \infty$. *Dann gilt*

$$\frac{\sqrt{n}\Big(\frac{1}{n^2 - n} \sum_{1 \leq i \neq j \leq n} L_n(W_i, W_j) - \mathrm{E}\left(L_n(W_1, W_2) \right) \Big)}{M_n} \xrightarrow{\mathrm{v}} \mathrm{N}(0, 1) \text{ für } n \to \infty.$$

Beweis: Es ist für alle $n \in \mathbb{N}$ $\sigma_n \geq n^{3/2} M_n/2 > 0$. Weiter ist wegen $\sup_{x,y \in \mathbb{R}^d} |L_n(x,y)| = o(n^{3/2})$ für $n \to \infty$ und $\liminf_{n\to\infty} M_n > 0$

$$\frac{\sup_{x,y\in\mathbb{R}^d} |L_n(x,y)|}{\sigma_n} \leq \frac{\sup_{x,y\in\mathbb{R}^d} |L_n(x,y)|}{n^{3/2} M_n/2} \longrightarrow 0 \text{ für } n \to \infty.$$

Analog ist wegen $\sup_{x\in\mathbb{R}^d} \mathrm{E}(|L_n(x,W_1)|) = o(n^{1/2})$ für $n \to \infty$ und $\liminf_{n\to\infty} M_n > 0$

$$\frac{\sup_{x\in\mathbb{R}^d} \mathrm{E}(|L_n(x,W_1)|)}{\sigma_n/n} \leq \frac{\sup_{x\in\mathbb{R}^d} \mathrm{E}(|L_n(x,W_1)|)}{n^{1/2} M_n/2} \longrightarrow 0 \text{ für } n \to \infty.$$

Ferner ist mit der Jensenschen Ungleichung $\mathrm{E}(L_n(W_1,W_2))^2/n \leq \mathrm{E}(L_n(W_1,W_2)^2)/n \to 0$ für $n \to \infty$. Insgesamt folgt mit $\mathrm{E}(L_n(W_1,W_2)^2) = o(n)$, $\mathrm{E}(L_n(W_1,W_2)) = o(n^{1/2})$ für $n \to \infty$ und $\liminf_{n\to\infty} M_n > 0$

$$\begin{aligned} \lim_{n\to\infty} \frac{\sigma_n}{n^{3/2} M_n/2} = \lim_{n\to\infty} \Bigg(\frac{1}{2} \Bigg(& \frac{\mathrm{E}\left(L_n(W_1,W_2)^2\right)}{n} - \frac{\mathrm{E}\left(L_n(W_1,W_2)\right)^2}{n} \Bigg) \\ & + \mathrm{E}\left(\left(\mathrm{E}\left(L_n(w,W_2)\right)_{|w=W_1} - \mathrm{E}\left(L_n(W_1,W_2)\right) \right)^2 \right) \Bigg)^{1/2} \\ & \mathrm{E}\left(\left(\mathrm{E}\left(L_n(w,W_2)\right)_{|w=W_1} - \mathrm{E}\left(L_n(W_1,W_2)\right) \right)^2 \right)^{-1/2} = 1. \end{aligned}$$

Damit folgt mit Satz A.1.4 die Behauptung. □

Nun wird ein schwaches Gesetz der großen Zahlen für U-Statistiken mit einem von n abhängigen Kern formuliert. Hierzu sei $d, k \in \mathbb{N}$ und auf einem Wahrscheinlichkeitsraum $(\Omega, \mathfrak{A}, P)$ für $n \geq k$ durch $W_1, \ldots, W_n$ unabhängige und identisch verteilte $\mathbb{R}^d$-wertige Zufallsvektoren gegeben. Zudem sei $L_n : (\mathbb{R}^{kd}, \mathfrak{B}^{kd}) \to (\mathbb{R}, \mathfrak{B})$ eine von n abhängige Kernfunktion der Kernlänge k, sodass $L_n(W_{i_1}, \ldots, W_{i_k}) = L_n(W_1, \ldots, W_k)$ für alle Permutationen $(i_1, \ldots, i_k)$ von $(1, \ldots, k)$. Dann ist

$$\frac{1}{n(n-1)\cdots(n-k+1)} \sum_{\substack{1 \leq i_1,\ldots,i_k \leq n \\ i_l \neq i_m, l \neq m}} L_n(W_{i_1}, \ldots, W_{i_k})$$

die zugehörige U-Statistik der Kernlänge k mit einem von n abhängigen Kern. Es sei

$$h_{n,i}(x_1, \ldots, x_i) := \mathrm{E}\left(L_n(x_1, \ldots, x_i, X_{i+1}, \ldots, X_k)\right),\ x_1, \ldots, x_i \in \mathbb{R}^d,\ i \leq k,$$

und es gelte $\mathrm{E}(L_n^2(X_1, \ldots, X_k)) < \infty$.

A.1.6 Satz: *Es sei* $\mathrm{E}(h_{n,i}^2(X_1, \ldots, X_i)) = o(n^i)$, $i \leq k$, *für* $n \to \infty$. *Dann gilt*

$$\lim_{n\to\infty} \left| \frac{1}{n(n-1)\cdots(n-k+1)} \sum_{\substack{1\leq i_1,\ldots,i_k\leq n \\ i_l\neq i_m, l\neq m}} L_n(W_{i_1}, \ldots, W_{i_k}) - \mathrm{E}\left(L_n(W_1, \ldots, W_k)\right) \right| = 0$$

in Wahrscheinlichkeit.

Beweis: Es ist

$$\begin{aligned}
&\mathrm{Var}\left(\frac{1}{n(n-1)\cdots(n-k+1)} \sum_{\substack{1\leq i_1,\ldots,i_k\leq n \\ i_l\neq i_m, l\neq m}} L_n(W_{i_1}, \ldots, W_{i_k}) \right) \\
=&\frac{1}{\binom{n}{k}} \sum_{i=1}^{k} \binom{k}{i} \binom{n-k}{m-i} \mathrm{Var}\big(h_{n,i}(W_1, \ldots, W_i)\big),
\end{aligned}$$

siehe 5.2.1 Lemma A in [36]. Damit folgt mit der Tschebyscheff-Ungleichung für alle $\varepsilon > 0$

$$\begin{aligned}
&P\left(\left| \frac{1}{n(n-1)\cdots(n-k+1)} \sum_{\substack{1\leq i_1,\ldots,i_k\leq n \\ i_l\neq i_m, l\neq m}} L_n(W_{i_1}, \ldots, W_{i_k}) - \mathrm{E}\left(L_n(W_1, \ldots, W_k)\right) \right| \geq \varepsilon \right) \\
\leq&\frac{1}{\varepsilon^2} \frac{1}{\binom{n}{k}} \sum_{i=1}^{k} \binom{k}{i} \binom{n-k}{k-i} \mathrm{Var}\big(h_{n,i}(W_1, \ldots, W_i)\big) \\
\leq&\frac{1}{\varepsilon^2} \frac{1}{\binom{n}{k}} \sum_{i=1}^{k} \binom{k}{i} \binom{n-k}{k-i} \mathrm{E}\left(h_{n,i}^2(W_1, \ldots, W_i)\right) \\
\sim&\frac{1}{\varepsilon^2} \sum_{i=1}^{k} \binom{k}{i} \frac{k!}{(k-i)!} \frac{\mathrm{E}\left(h_{n,i}^2(W_1, \ldots, W_i)\right)}{n^i} \longrightarrow 0 \text{ für } n \to \infty,
\end{aligned}$$

woraus die Behauptung folgt. □

Nun sei $k : (\mathbb{R}, \mathfrak{B}) \to (\mathbb{R}, \mathfrak{B})$ eine Kernfunktion mit $\int |k(y)|\mathrm{d}y < \infty$ und $\lim_{|y|\to\infty} |yk(y)| = 0$. Ferner sei $g : (\mathbb{R}, \mathfrak{B}) \to (\mathbb{R}, \mathfrak{B})$ stetig mit $\int |g(y)|\mathrm{d}y < \infty$, und es sei $(h_n)_{n\in\mathbb{N}}$, $h_n > 0$ für alle $n \in \mathbb{N}$, eine Folge von Bandbreiten mit $\lim_{n\to\infty} h_n = 0$. Weiter sei für alle $n \in \mathbb{N}$ $\int |k(y)g(x - h_n y)|\mathrm{d}y < \infty,\ x \in \mathbb{R}$, und

$$g_n(x) := \int k(y)g(x - h_n y)\mathrm{d}y,\ x \in \mathbb{R}.$$

Dann haben wir das folgende Resultat aus dem Werk von Rao ([33]):

A.1.7 Satz: *Es gilt*

$$\lim_{n\to\infty} g_n(x) = g(x) \int k(y)\mathrm{d}y\ , x \in \mathbb{R}.$$

Beweis: Siehe Theorem 2.11 in [33]. □

Literatur

[1] Anderson, T. W. (2003). *An Introduction to Multivariate Statistical Analysis.* Wiley.

[2] Arcones, M. A. (2005). *Bahadur Efficiency of the likelihood ratio test.* Mathematical Methods of Statistics 14, 163-179.

[3] Bahadur, R. R. (1967). *An Optimal Property of the Likelihood Ratio Statistic.* Proc. Fifth Berkeley Symp. on Math. Statist. and Prob., Vol. 1, 13-26.

[4] Bahadur, R. R. (1971). *Some Limit Theorems in Statistics.* Regional Series in Applied Mathematics. SIAM, Philadelphia.

[5] Baringhaus, L. (1987). *Asymptotic Optimality of Multivariate Linear Hypothesis Tests.* Journal of Multivariate Analysis, Vol. 23, 303-311.

[6] Bazaraa, M. S., Shetty, C. M. (1970). *Nonlinear Programming Theory and Algorithms.* Wiley.

[7] Frishmuth, M. (1991). *Zweistichproben-Test oder Test bei verbundenen Stichproben?* Diplomarbeit, Leibniz Universität Hannover.

[8] Gänssler, P., Stute, W. (1977). *Wahrscheinlichkeitstheorie.* Springer.

[9] Gibbons, J. D., Chakraborti, S. (2011). *Nonparametric Statistical Inference.* Taylor and Francis.

[10] Groeneboom, P., Oosterhoff, J. (1977). *Bahadur Efficiency and Probabilities of Large Deviations.* Stat. Neerlandica 31, 1-24.

[11] Gruber, M. H. J. (2013). *Matrix Algebra for Linear Models.* Wiley.

[12] Hannan, E. J. (1956). *The Asymptotic Powers of Certain Tests Based on Multiple Correlations.* Journal of the Royal Statistical Society, Series B, Vol. 18, No. 2, 227-233.

[13] Heuser, H. (2008). *Lehrbuch der Analysis Teil 2.* 2. Auflage. Vieweg+Teubner.

[14] Hoeffding, W. (1940). *Maszstabinvariante Korrelationstheorie.* Schriften des Mathematischen Institiuts und des Instituts für Angewandte Mathematik der Universität Berlin, 5, 179-233.

[15] Hoeffding, W. (1965). *Asymptotically Optimal Tests for Multinomial Distributions.* Ann. Math. Stat. Vol. 36, No. 2, 369-401.

[16] Hollander, M. (1960). *Asymptotic Efficiency of Two Nonparametric Competitors of Wilcoxon's Two Sample Test.* Journal of the American Stat. Association. Vol. 62, No. 319, 939-949.

[17] Jammalamadaka, S. R., Janson, S. (1986). *Limit Theorems for a Triangular Scheme of U-Statistics with Applications to Inter-Point Distances.* The Annals of Prob. Vol. 14, No. 4, 1347-1358.

[18] Kallenberg, W. C. M. (1985). *On Moderate and Lare Deviations in Multinomial Distributions.* The Ann. of Stat. Vol. 13, No. 4, 1554-1580.

[19] Kallenberg, W. C. M., Kourouklis, S. (1992). *Hodges-Lehmann-Optimality of Tests.* Stat. & Prob. Letters Vol. 14, 31-38.

[20] Knüsel, L. (1977). *Vergleich zweier Testverfahren zur Symmetrieprüfung in 4-Felder-Tafeln.* Metrika Vol. 24, 175-186.

[21] Königsberger, K. (2004). *Analysis 2. 5., korrigierte Auflage.* Springer.

[22] Kourouklis, S. (1988). *Hodges-Lehmann-Efficacies of Certain Tests in Multivariate Analysis and Regression Analysis.* The Canadian Journal of Stat. Vol. 16, No. 1, 87-95.

[23] Kourouklis, S. (1989). *On the Relation between Hodges-Lehmann Efficiency and Pitman Efficiency.* The Canadian Journal of Stat. Vol. 17, No. 3, 311-318.

[24] Koziol, J. A. (1978). *Exact Slopes of Certain Multivariate Tests of Hypotheses.* The Ann. of Stat. Vol. 6, No. 4, 546-558.

[25] Magnus, J. R. (1985). *On Differentiating Eigenvalues and Eigenvectors.* Econometric Theorie 1, 179-191.

[26] Magnus, J. R., Neudecker, H. (1988). *Matrix Differential Calculus with Applications in Statistics and Econometrics.* Wiley.

[27] Müller-Funk, U., Witting, H. (1995). *Mathematische Statistik II - Asymptotische Statistik: Parametrische Modelle und nicht-parametrische Funktionale.* Teubner.

[28] Muirhead, R. J. (1982). *Aspects of Multivariate Statistical Theory.* Wiley.

[29] Nikitin, Y. (1995). *Asymptotic Efficiency of Nonparametric Tests.* Cambridge University Press.

[30] Okamoto, (1973). *Distinctness of the Eigenvalues of a Quadratic form in a Multivariate Sample.* The Ann. of Stat. Vol. 1, No. 4, 763-765.

[31] Pollak, M., Cohen, J. (1981). *A Comparison of the Independent-Samples t-Test and the Paired-Samples t-Test when the Observations are Nonnegatively Correlated Pairs.* J. Stat. Planning and Inference 5, 133-146.

[32] Puri, M. L., Sen, P. K. (1971). *Nonmarametric Methods in Multivariate Analysis.* Wiley.

[33] Rao, B. L. S. Prakasa (1983). *Nonparametric Functional Estimation.* Academic Press.

[34] Roverato, A., Whittaker, J. (1998). *The Isserlis Matrix and its Application to Non-Decomposable Graphical Gaussian Models.* Biometrika 85, 3, 711-725.

[35] Rublík, F. (1994). *On Hodges-Lehmann Optimality of LR Tests.* Kybernetika, Vol. 30, No. 2, 199-210.

[36] Serfling, R. S. (1980). *Approximation Theorems of Mathematical Statistics*. Wiley.

[37] Serre, D. (2010). *Matrices: Theory and Applications*. Springer.

[38] Van der Vaart, A. W. (1998). *Asymptotic Statistics*. Cambridge University Press.

[39] Werner, J. (1992). *Numerische Mathematik 2. Eigenwertaufgaben, lineare Optimierungsaufgaben, unrestringierte Optimierungsaufgaben*. Bd. 2. Vierweg.

[40] Wieand, H. S. (1976). *A Condition under which the Pitman and Bahadur Approaches to Efficiency Coincide*. The Ann. of Stat. Vol. 4, No. 5, 1003-1011.

[41] Witting, H. (1985). *Mathematische Statistik I - Parametrische Verfahren bei festem Stichprobenumfang*. Teubner.