

Die Weibull-Verteilung und das Problem garantierter Mindestlebensdauern

Harald Rosemann*

Bei Lebensdauerversuchen wird oft die Weibullverteilung mit drittem Parameter angesetzt. Sie verspricht eine anfänglich absolute Fehlerfreiheit bis $t = L_0$; diese Zusage kennzeichnet deterministisches Verhalten. Erst im Anschluß an diese erste Phase erlaubt sie zufälliges Versagen; das ist stochastisches Geschehen. Dieses Modell mit zwei aufeinander folgenden Abschnitten von unvereinbar gegensätzlicher Natur paßt nicht zur Vorstellung eines kontinuierlich fortschreitenden Abnutzungs, Ermüdungs oder Verschleißprozesses, solange keine äußeren Einflüsse auf ihn einwirken. An Stelle dieses in sich widersprüchlichen Modells werden Varianten zur Weibullverteilung von durchgängig stochastischer Natur vorgeschlagen und untersucht, die mit verminderter Versagenswahrscheinlichkeit starten, bevor sie zu normalem Weibull-Verhalten übergehen.

1 Einleitung

Werkstoffe können verschleifen, sie können ermüden. Als Folge versagen Bauteile, zu nicht vorhersehbaren Zeiten, den individuellen, statistisch verteilten Lebensdauern. Sie entsprechen recht oft der Weibullverteilung als derjenigen, die höchste Zielwerte liefert bei Parameterschätzungen mittels Optimierungsverfahren wie der Methode der 'Maximum-Likelihood'. Die ursprüngliche Weibullverteilung ist durch zwei Parameter festgelegt.

* Prof.Dr.-Ing.habil. Harald Rosemann
am Institut für Maschinenkonstruktion und Tribologie,
Leibniz Universität Hannover
rosemann@imkt.uni-hannover.de

Man hat versucht, bei anfänglich oft schwach ausgeprägtem Ausfallverhalten, eine modifizierte Variante mit drei Parametern einzusetzen. Diese Variante ist in ständigem Gebrauch, wie bereits aus einigen der ersten Ergebnisse einer Suchanfrage im Internet nach dem Begriff 'Weibullverteilung' hervorgeht. Der zusätzliche dritte, auch Schwellenparameter genannt, steht für eine gewisse Anfangs- oder Einlaufzeit, in der eine - angeblich - absolute und totale Fehlerfreiheit garantiert sein soll. Ist diese Zusage begründet, kann man ihr vertrauen?

2 Das Problem

2.1 Die Weibullverteilung mit zwei Parametern

In vielen Lebensdauerversuchen zeigt sich die ursprüngliche Weibullverteilung mit zwei Parametern als gut geeignet, die gemessenen Werte zu repräsentieren. $F(t)$ steht allgemein für Verteilungsfunktion, $W(t)$ speziell für die Abhängigkeit nach Weibull:

$$F(t) = W(t) = \begin{cases} 1 - e^{-(t/T)^\beta}, & t \geq 0, \quad \beta > 0, \quad T > 0 \\ 0, & t < 0 \end{cases} \quad (1)$$

Wesentliches Kennzeichen ist, daß in der Exponentialfunktion die Zeit t ihrerseits zur Potenz β erhoben ist. T wird charakteristische Zeit genannt; ganz gleich, welchen Wert β hat, stets ist: $W(T) = 1 - 1/e \approx 0.632$.

Bei $t = 0$ beginnt die Verteilungsfunktion $W(t)$ mit t von 0 aus monoton anzusteigen, um sich für große t dem Niveau 1 anzunähern. Von ihren Werten liest man die Wahrscheinlichkeit ab, mit der bis zur Zeit t ein Ausfall oder Versagen zu erwarten ist. Mit $W(t) = 0$ zeigt sie für Zeiten $t < 0$, daß die Wirkung nicht vor der Ursache eintreten kann, daß also ein Ausfall erst nach Beginn der schädigenden Belastung zu erwarten ist; das schließt ein vorheriges Versagen grundsätzlich aus, die Wahrscheinlichkeit für negative Lebenszeiten ist Null.

Anstelle der Bezugsgröße T nutzt man gern die L_{10} -Lebensdauer und formt die Beziehung um zu:

$$F(t) = W(t) = \begin{cases} 1 - e^{\ln(0.9) \left[\frac{t}{L_{10}} \right]^\beta}, & t \geq 0, \quad \beta > 0, \quad L_{10} > 0, \\ 0, & t < 0. \end{cases} \quad (2)$$

Wieder unabhängig von dem Wert, den β annimmt, ist stets $W(L_{10}) = 0.1$ und L_{10} gibt die Zeit an, bis zu der 10% der Ausfälle zu erwarten sind.

2.2 Die Weibullverteilung mit drei Parametern

Bei einigen Anwendungen stellt man fest, daß zu Beginn nicht so viele Ausfälle zu verzeichnen sind, wie man es nach der ursprünglichen Weibullverteilung erwarten sollte. Die Abweichung wird damit erklärt, daß Alterungsprozesse wie Verschleiß, Abnutzung oder Ermüdung eine gewisse Zeit benötigen, ehe Schäden auftreten, die zum Ausfall führen. Deswegen führten Snare [1] und später Bergling [2] einen dritten Parameter L_0 ein, um den sie die Verteilungsfunktion $W(t)$ nach rechts verschoben und gemäß

$$F(t) = W(t) = \begin{cases} 1 - e^{\ln(0.9) \left[\frac{t-L_0}{L_{10}-L_0} \right]^\beta}, & \begin{cases} t \geq L_0, \\ \beta > 0, \quad L_{10} > 0, \\ 0 \leq L_0 < L_{10} \end{cases} \\ 0, & t < L_0 \end{cases} \quad (3)$$

die Kurve dieser geschätzten Verteilung 'besser' durch die Punktwolke der frühen Ausfälle zu leiten. Das sieht in den Diagrammen auch durchaus gut aus. Und ebenso wie derjenige Parametersatz überlegen ist, der in der Optimierung eines Schätzverfahrens wie der 'Maximum-Likelihood' die höheren Zielwerte liefert, so mag auch die Weibullverteilung mit drei Parametern dabei besser abschneiden als die mit zweien und somit als eher geeignet erscheinen, die gemessenen Werte zu repräsentieren; das ist die eine Seite.

2.3 Der Konflikt

Aus anderer Sicht hingegen steht hinter der verschobenen Verteilung mit dem Schwellenparameter L_0 ein Modell, das sich aus zwei Phasen zusammensetzt. Die erste ist für $t < L_0$ von rein deterministischer Natur, die zweite für größere t ist hingegen stochastisch.

Die beiden einander wesensfremden Bereiche stoßen an der Grenze L_0 aufeinander, einer von vornherein festen Grenze, die nicht vom Zufall abhängt.

Im ersten Teil herrscht absolute Fehlerfreiheit. Das Modell sichert dort eine durch nichts zu erschütternde Garantie zu, daß vor $t = L_0$ keinerlei Versagen auftreten kann und wird. Ist das realistisch? Kann man einem bestimmten Ereignis, hier dem sehr frühen Ausfall, das Etikett 'unmöglich' anheften, nur weil man es in zahlreichen Versuchen noch nicht beobachtet hat? Ausschließlich aus Messungen läßt sich eine derart grundlegende Feststellung nicht ableiten, seien die Daten auch noch so zahlreich. Selbst wenn eine Schätzung \hat{L}_0 aus einer Stichprobe existiert und rein formal nach Park [3] berechnet werden kann, das allein ist für sich genommen kein Beweis für die Existenz einer fehlerfreien Zeit L_0 in der Grundgesamtheit.

Dabei mag man aus numerischer Sicht kaum einen Unterschied empfinden zwischen 'genau Null' und sehr, sehr klein, sagen wir einem Milliardstel oder noch weniger. Qualitativ hingegen ist das 'unmögliche Ereignis' grundverschieden von einem mit geringer Eintrittswahrscheinlichkeit; das erste beruht auf abstrakter Definition, das andere ist eine Sache der realen Welt; im ersten Fall darf man völlig unbesorgt sein, im anderen können vorsorgende Maßnahmen erforderlich werden.

Überdies bedürfte es einer von außen kommenden Einwirkung, um den Übergang in die zweite Phase zu vollziehen, so wie bei der ursprünglichen Weibullverteilung (1) das Einschalten der Belastung bei $t = 0$ den Schädigungsprozeß erst in Gang setzt. Wo aber ist der Schalter, der die Versagensfreiheit bei $t = L_0$ beendet? Was läßt wie und warum den gleichförmig und sprunghaft ablaufenden Prozeß des Verschleißens ganz unvermittelt und übergangslos umschlagen von der einen Verhaltensart in die andere?

Man steht hier vor einer unbefriedigenden Situation, es gibt einen Konflikt. Einerseits hat man die beste Verteilung (unter den getesteten), andererseits trifft sie von ihrer Aussage und inneren Bedeutung her nicht zu auf den stetig und kontinuierlich verlaufenden Vorgang, von dem die Meßwerte herrühren. Man kann sich behelfen und pragmatisch sagen: das ist eine Näherung. Dennoch muß man sich wappnen, einen von außen kommenden Anspruch abzuwehren, man solle für einen Schaden einstehen, weil man ja selbst die frühe Fehlerfreiheit zugesichert habe. Das ist ein Dilemma, aus dem nur eines heraushilft: eine hinsichtlich des Schätzverfahrens noch besser geeignete Verteilung, die sich problemfrei deuten läßt.

3 Neuer Ansatz

3.1 Hyperbel anstelle von Geraden

In Frage steht also, ob sich eine vermittelnde Zwischenlösung finden läßt, die den Weibullcharakter bewahrt und verzögertes Ausfallverhalten erfaßt, ohne Mißdeutungen zuzulassen. Zunächst ist es sinnvoll, zur Vereinfachung der Gleichungen auf $(L_{10} - L_0)$ normierte Größen mit $t' = t/(L_{10} - L_0)$ and $L'_0 = L_0/(L_{10} - L_0)$. zu definieren. Dann schreiben wir die Verschiedenheit der beiden Verteilungen mit $g_2(t') = t'$ und $g_3(t') = t' - L'_0$ als Funktionen von t' ; der Index zählt die Parameter. Die Funktionen $g(t')$ sind jeweils die Basis, die nach Weibull in der Verteilungsfunktion zur Potenz β erhoben wird.

Diese zwei Funktionen zeigt Bild 1 in Abhängigkeit von t' und L'_0 als zwei parallele Linien, g_3 rechts, punktiert ausgeführt, um $L'_0 = 0.05$ verschoben gegen g_2 links, gestrichelt

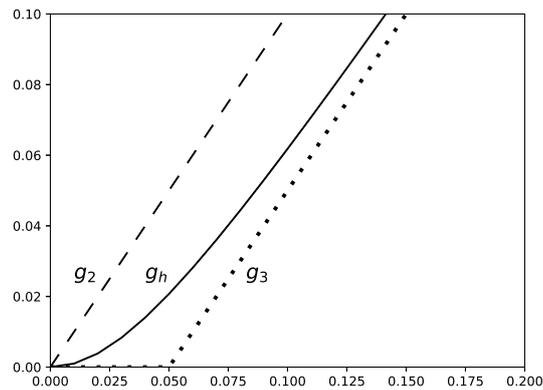


Abb. 1: $g(t')$ über t' , $L'_0 = 0.05$

gezeichnet. In den Zwischenraum läßt sich eine neue Kurve legen. Sie sollte bei $t' = 0$ vom Niveau 0 aus ansteigen und sich dann im weiteren Verlauf g_3 annähern. Es liegt nahe, mit

$$g_h(t') = -L'_0 + \sqrt{t'^2 + L_0'^2}, \quad t' \geq 0, \quad L'_0 \geq 0 \quad (4)$$

den Ast einer Hyperbel zu wählen, wobei wir mit L'_0 dieselbe Bezeichnung für einen entsprechenden Parameter einsetzen. Diesen Verlauf zeigt die durchgehende Linie. In der Nähe von $t' = 0$ verhält sich $g_h(t')$ wie $t'^2/2L'_0$, beginnt also mit horizontaler Tangente.¹

¹ Will man im Gegensatz zu dem verzögert auftretenden Versagen besonders häufige Frühausfälle erfassen, bietet sich mit $g_h(t') = \sqrt{t'^2 + 2t'L'_0}$ ein anderer Hyperbelast an, der bei $t' = 0$ wie die Quadratwurzel zunächst stark ansteigt.

3.2 Vergleich der Verteilungsfunktionen

Die drei Versionen für $g(t')$ führen mit $W(g(t'))$ zu drei Verteilungsfunktionen nach Weibull, wobei das jeweilige $g(t')$ das ursprüngliche t' ersetzt; Kurzformen wie W_2 stehen für $W(g_2(t'))$. Die beiden folgenden Bilder 2 und 3 geben den Verlauf wieder, links in

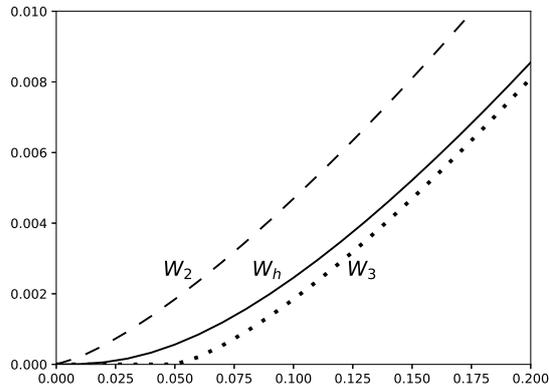


Abb. 2: $W(g(t'))$ über t' ,
Maßstab linear, $L'_0 = 0.05$.

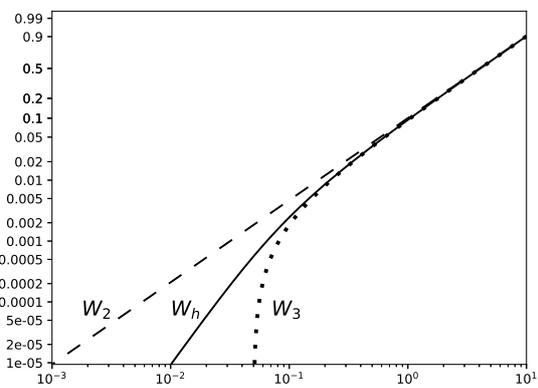


Abb. 3: $W(g(t'))$ über t'
im Weibullnetz, $L'_0 = 0.05$.

linear geteilten Achsen, rechts im Weibullnetz, das die ursprüngliche Weibullverteilung als gerade Linie zeigt. Für die Berechnungen wurde $\beta = 1.35$ gesetzt.

Das erwünschte sinnvolle Verhalten tritt deutlich hervor. Links in Bild 2 startet W_h bei Null erst langsamer als das ursprüngliche W_2 , um sich im weiteren Verlauf dem W_3 mehr und mehr anzuschmiegen. Im Weibullnetz rechts im Bild 3 verläuft W_h zu Beginn zwischen den beiden anderen Kurven, steiler von unten kommend als W_2 aber nicht so abrupt bei einem festen Wert startend wie W_3 bei $t' = L'_0$.

Frühe Ausfälle sind so nach dem Hyperbelansatz (4) weniger wahrscheinlich als bei der ursprünglichen Weibullverteilung W_2 aber eben auch nicht gänzlich unmöglich vor $t' = L'_0$ wie bei W_3 . Zu größeren Werten t' hin verschmelzen die Kurven, im Weibullnetz deutlich zu sehen. Bild 2 zeigt mit nicht verzerrten Koordinaten nur den Ausschnitt bei kleinen t' ; erweitert auf bis zu $t' = 10$ wie im Weibullnetz würde man die drei Kurven vor allem bei größeren Zeiten t' kaum unterscheiden können.

Das gesetzte Ziel, das Kernanliegen ist erreicht, ein brauchbarer Ersatz gefunden. Er ist von durchgehend stochastischer Natur ohne deterministischen Anteil; mit anfänglich klei-

nen Wahrscheinlichkeiten kann er ein verzögertes Ausfallverhalten erfassen. Es werden keine zweifelhafte Garantien mehr benötigt, keine dubiosen Umschaltmechanismen.

4 Erweiterung des Hyperbelansatzes

4.1 Weiterer Ersatz der Geraden

Ist das Potential des ersten Ansatzes nun erschöpft oder läßt er sich noch weiter verfolgen, läßt er sich aufweiten? Der charakteristische Verlauf des Hyperbelanstes sollte erhalten bleiben; wie kann man ihn variieren? Indem wir die Quadratwurzel und die zweite Potenz verallgemeinern, erhalten wir mit dem neuen Parameter c , bei $g_c(t')$ als Index eingesetzt, den modifizierten Ansatz:²

$$g_c(t') = -L'_0 + \left[(t')^c + (L'_0)^c \right]^{1/c}, \quad t' \geq 0, \quad L'_0 \geq 0, \quad c \geq 1 \quad (5)$$

Die Kurve $g_c(t')$ wächst monoton mit t' , wie zuvor bereits die Hyperbel nach Gleichung (4); ersetzt man in der Weibullverteilung das Argument t' durch $g_c(t')$, bleibt der Charakter einer Verteilung ebenso erhalten.

Bild 4 zeigt zwischen den beiden bekannten Ausgangsgeraden, gestrichelt und punktiert, eine Schar von vier durchgehenden Kurven. Die Liste führt die zugehörigen Ansätze für

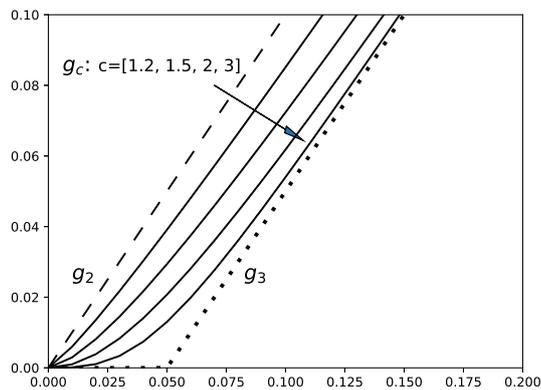


Abb. 4: $g(t')$ über t' , $L'_0 = 0.05$

²Werte im Bereich $0 < c < 1$ führen zu vermehrten Frühausfällen

c auf, der Pfeil zeigt in Richtung wachsender Werte. Bei $t' = 0$ starten die Kurven mehr oder weniger platt mit t' wie t'/cL_0^{c-1} , können sich also mit höheren c recht eng und über eine längere Dauer an die Zeitachse anschmiegen.

Die neue Formel füllt nicht nur die Fläche zwischen den beiden ursprünglichen Geraden, sie hat darüber hinaus die schöne Eigenschaft, daß sie mit $c = 1$ die ursprüngliche Weibullverteilung enthält und daß die andere Grenzfall für $c \rightarrow \infty$ ist.

4.2 Vergleich der Verteilungsfunktionen

Das Erscheinungsbild der zugehörigen Verteilungsfunktionen, links in gleichmäßig geteilten Koordinaten, rechts im Weibullnetz, fällt nun aus, wie man es vermuten konnte; zwischen den beiden ursprünglichen Verläufen stehen mit unterschiedlichem c beliebig

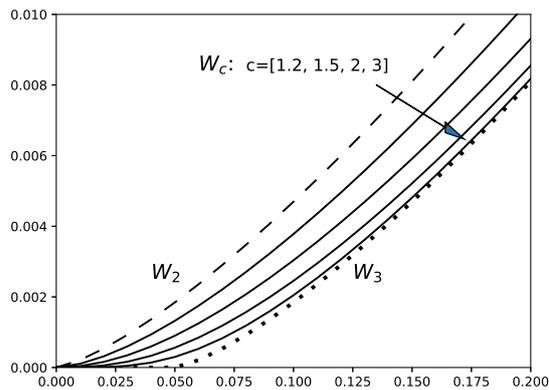


Abb. 5: $W(g(t'))$ über t' ,
Maßstab linear, $L'_0 = 0.05$.

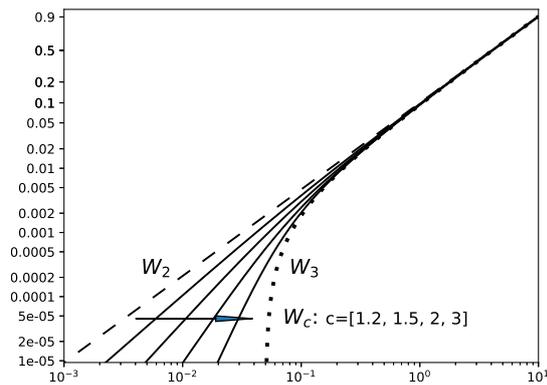


Abb. 6: $W(g(t'))$ über t'
im Weibullnetz, $L'_0 = 0.05$.

viele Zwischenlösungen bereit. Im Bild 6, rechts, laufen die Kurven von unten zunächst fast geradlinig aufsteigend mit einer mehr oder weniger starken Krümmung in die Weibullgerade W_2 ein. Weil die Reihenentwicklung von $g_c(t')$ für kleine Zeiten t' mit der Ordnung t'^c startet, beträgt die Anfangssteigung der W_c im Weibullnetz $c\beta$.

4.3 Besondere Eigenschaften

Als Beispiel wiederholt Bild 7 die Darstellung des ersten Hyperbelansatzes (4). Zusätzlich verdeutlicht eine Reihe von Kringeln den nahezu geradlinigen Anfangsverlauf mit auf

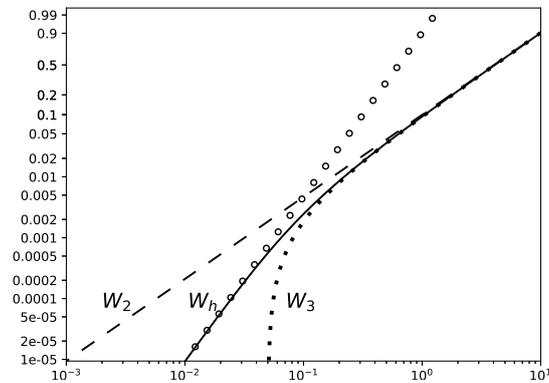


Abb. 7: $W(g(t'))$ und Asymptote über t' ,
im Weibullnetz, $L'_0 = 0.05$

2β verdoppelter Steigung und setzt ihn nach oben hin fort. Demnach bilden zwei Geraden eine konservative Näherung. Das erinnert an die alte Regel zur Auslegung von Wälzlagern, nach der für Lebensdauern unter L_{10} der Wert für β auf 1.5 zu erhöhen ist.³

4.4 Kurzer Blick auf das Schätzen der Parameter

Bei der ursprünglichen Weibullverteilung mit zwei Parametern ermittelt man aus den gemessenen Lebensdauern die Schätzwerte $\hat{\beta}_2$ und \hat{L}_{10-2} . Jede der Messungen hat Einfluß auf jede dieser beiden Werte. Allenfalls wirken große und kleine Ausfallzeiten ein wenig mehr auf das Ergebnis für die Steigung $\hat{\beta}_2$ im Weibullnetz ein als die mittleren, die ihrerseits etwas mehr Gewicht ausüben bei \hat{L}_{10-2} .

Etwas anders sieht es aus bei den vier Parametern des erweiterten Ansatzes. Den neuen Größen L_{0-c} und c kommt von sich aus Einfluß und Wirkmöglichkeit nur im Anfangsbereich zu; demzufolge hängen ihre Schätzungen \hat{L}_{0-c} und \hat{c} auch nur von den Zeiten der ersten frühen Versagensfälle ab. Damit verbunden ist eine verminderte Abhängigkeit

³ Diese Modifikation ist nach [4] in den Auslegungsbeiwert a_1 eingerechnet.

der beiden anderen Schätzungen $\hat{\beta}_c$ und \hat{L}_{10_c} von diesen anfänglichen Meßwerten. Man benötigt also eine hinreichend große Zahl an Frühausfällen um die neuen Parameter einigermaßen zutreffend und verlässlich schätzen zu können. Soweit bislang die um den Schwellenwert L_0 auf drei Parameter erweiterte Weibullverteilung als brauchbar erschien, mögen in diesen Fällen wohl auch hinreichend viele Werte für den neuen Ansatz zur Verfügung stehen. Überdies läßt sich denken, daß es typische Werte für spezielle Anwendungen gibt, so wie üblicherweise bei Wälzlagern mit vorwiegender Punktberührung β zu 1.11 gesetzt wird und bei Mischformen aus Punkt- und Linienberührung zu 1.35.

5 Fazit

Soweit bei kontinuierlich verlaufenden Verschleiß- und Ermüdungsprozessen keine äußeren Einflüsse vorliegen, ist die Weibullverteilung mit drei Parametern kein gültiges Modell für die Verteilung der Lebensdauern; sie kann allenfalls als pragmatische Näherung angesehen werden. Aus ihrem Ansatz läßt sich mithin auch keine garantierte Lebensdauer L_0 ableiten. In der hier vorgestellten Variante wird die Zeit t ersetzt durch einen Verlauf nach Art einer Hyperbel. Damit erfaßt dieser neue Ansatz verzögertes Ausfallverhalten in einem voll stochastischen Modell und vermeidet Deutungsschwierigkeiten vor allem in Hinsicht auf Garantiezusagen.

Dank

Meinen sehr geschätzten Kollegen M.Sc. Josephine Kelley und Prof.Dr.-Ing. Gerhard Poll am IMKT danke ich für die kritische Durchsicht sowie für wertvolle Anregungen und regen Zuspruch bei der Abfassung dieses Manuskriptes.

Literatur

- [1] Snare, B., Neuere Erkenntnisse über die Zuverlässigkeit von Wälzlagern; *Die Kugellager-Zeitschrift*, Nr 162 (1969) S. 3-7.

Remark: Figure 2 refers to an L_{10} of 15 Million rotations; in the text, however, it states: "Die Lager liefen ... bei ... einer Belastung, die nach dem Katalog einer L_{10} -Lebensdauer von 10 Millionen Umdrehungen entspricht." (The bearings ran for a load that, according to the catalogue, corresponds to an L_{10} service lifetime of 10 Million revolutions.)

- [2] Bergling, G., Betriebszuverlässigkeit von Wälzlagern; *Die Kugellager-Zeitschrift*, Jahrgang 51, Heft Nr.188, (1976), S.1-10.

Remark: Figure 3 (agrees with Figure 2 in [1]) shows an L_{10} of 15 Million revolutions; in the legend, a different value of 10 Million revolutions is stated.

- [3] Park, C., A Note on the Existence of the Location Parameter Estimate of the Three-Parameter Weibull Model Using the Weibull Plot; *Mathematical problems in engineering* (2018), S. 1-6.

- [4] ISO 281:2007, Rolling bearings - Dynamic load ratings and rating life. (2007, and previous versions)