Zufällig oder systematisch?

Die Bestimmung von Mischverteilungen im Weibull-Netz

Curt Ronniger, München

Das Ausfallverhalten von Bauteilen wird in der Regel mit Hilfe der Weibull-Verteilung untersucht. Die Grenze von der reinen Verteilung zu in der Praxis häufig auftretenden Mischverteilungen ist allerdings unklar. Ein Automobilhersteller entwickelte ein Verfahren, mit dem sich zufällige und systematisch bedingte Streuungen unterscheiden lassen.

Um bessere Informationen über das Ausfallverhalten von Bauteilen zu erhalten, wird aus den Ausfällen üblicherweise eine Weibull-Verteilung bestimmt [1, 2]. Aus den Parametern dieser Verteilung, z.B. aus der Steigung der Ausgleichsgeraden, kann auf bestimmte Eigenschaften, wie etwa Verschleiß, geschlossen werden. In der Realität findet man aber nicht nur reine, sondern sehr oft Mischverteilungen. Sie sind ein Hinweis dafür, dass mehrere Ausfallmechanismen zur Wirkung kommen. Unklar ist allerdings bei einem nicht stetigen Verlauf die Grenze, ab welcher man eine Mischverteilung annehmen kann. Und die Frage ist offen, bis zu welcher Abweichung von der Ausgleichsgeraden eine zufällige Streuung vorliegt oder eine systematisch bedingte.

Bei der BMW AG in München wurde ein neues Verfahren entwickelt, mit dem sich diejenigen Verteilungen, die für bestimmte Fahrstreckenabschnitte charakteristisch sind, bestimmen lassen. Die Beurteilung einer Mischverteilung erfolgte im Unternehmen bislang subjektiv und führte bei Gewährleistungsanalysen, insbesondere mit dem Zulieferer, zu unterschiedlichen Interpretationen. Es erwies sich deshalb als notwendig, die Auswertungen auf fundierte Grundlagen zu stellen und somit objektiv bewerten zu können.

Für eine objektive Bewertung

Für eine Mischverteilung gibt es unterschiedliche Ursachen. Sind mehrere Ausfallmechanismen vorhanden, so kann es vorkommen, dass der Verlauf im Weibull-Netz nicht geradlinig ist, sondern eine Krümmung aufweist. Die Bestimmung einer Mischverteilung ist bei einer relativ geringen Probenanzahl schwierig, da die einzelnen Punkte immer um die Ausgleichsgerade herum streuen, und nicht immer handelt es sich bei ungeradlinigen Verläufen um eine Mischverteilung. Bei Automobilen kann es je nach Datenaufbereitung und -eingang auch vorkommen, dass die Daten bei höheren Laufstrecken ausgehen, weil noch nicht alle Kunden die entsprechende Kilometerleistung erreicht haben. Der Verlauf der Kurve ist dann rechtsabknickend. Zur Bestimmung dieses Sachverhaltes wendet man ein Prognoseverfahren an [2–4]. Im folgenden wird davon ausgegangen, dass die Daten korrekt aufbereitet wurden.

Grundsätzlich sollte man sich zunächst die Ausfallpunkte im Weibull-Netz anschauen [5]. Ist der Verlauf der Punkte unter Ausschluss der eben genannten Nebeneinflüsse immer noch rechts gekrümmt, so kann der Grund hierfür auch eine so genannte ausfallfreie Zeit t_o sein. Rein visuell spricht hierfür schon ein weich abgerundeter Rechtsknick. Das Vorhandensein einer ausfallfreien Zeit hat nichts mit einer Mischverteilung zu tun. Es muss deshalb geprüft werden, ob $t_o > 0$ der Grund für einen gekrümmten Verlauf ist, bevor man sich einem Test auf Mischverteilung zuwendet.

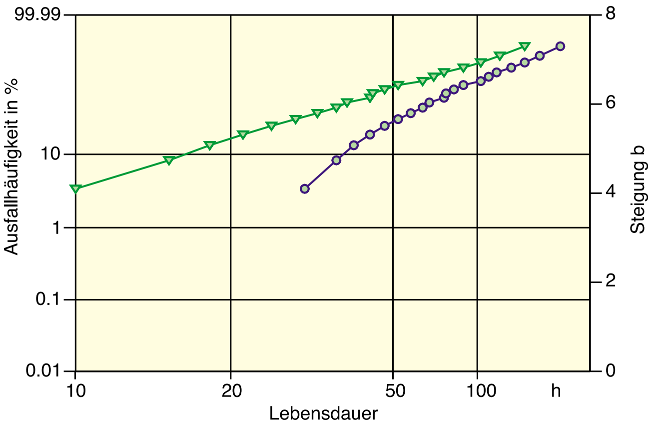

Bild 1. Darstellung der Ausfälle mit und ohne Berücksichtigung der ausfallfreien Zeit

Ausfallfreie Zeit

In Verbindung mit der ausfallfreien Zeit spricht man von der 3-parametrigen Weibull-Verteilung (Summenhäufigkeit):

$$H = 1 - e^{-\left(\frac{t-t_o}{T-t_o}\right)^b}$$

Zur Bestimmung von t_o existieren verschiedene Methoden [1]. Eine eindeutige Formel hierfür gibt es jedoch nicht. Grundsätzlich gilt, dass die ausfallfreie Zeit t_o zwischen 0 und dem Wert des ersten ausgefallenen Teiles liegen muss. In der Regel liegt ihr Wert auch sehr nahe vor dem des ersten Ausfalls. Folgendes Verfahren bietet sich an: Man lässt t_o in kleinen Schritten das Intervall zwischen $t > 0$ und dem ersten Ausfall t_{min} durchlaufen und berechnet bei jedem Schritt den Korrelationskoeffizienten der Ausgleichsgeraden. Je größer der Wert des Korrelationskoeffizienten ist, desto ge-

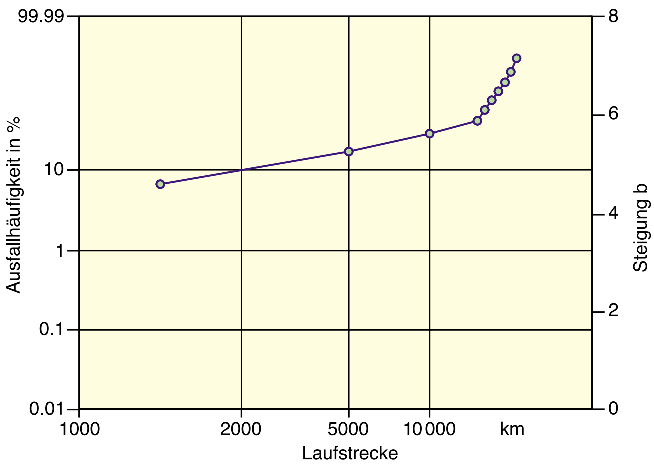

Bild 2. Typischer Verlauf einer Mischverteilung

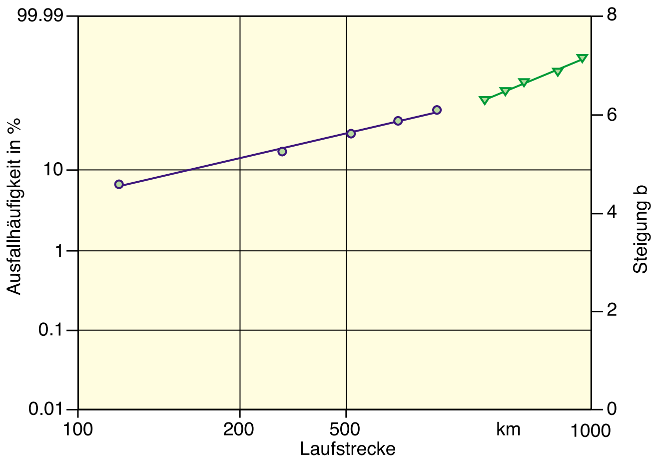

Bild 3. Beispiel für die getrennten Abschnitte einer Mischverteilung

nauer liegen die Punkte im Weibull-Netz auf einer Geraden. t_0 ist dann der Wert, bei dem der Korrelationskoeffizient am höchsten ist und sich somit die Ausgleichsgerade am besten approximieren lässt. Grafisch bedeutet dies, dass die Punkte im Weibull-Netz um den Betrag t_0 nach links verschoben aufgetragen werden (Bild 1). Die Punkte ergeben dann eine bessere Geradlinigkeit. Man kann natürlich statistisch den Korrelationskoeffizienten der Ausgleichsgeraden mit t_0 mit den einschlägigen Methoden dahingehend testen, ob die ausfallfreie Zeit signifikant vorhanden ist oder nicht (F-Test zur Prüfung der Linearität oder t-Test für den Vergleich der Regressionen mit und ohne t_0) [6].

Ist an Stelle des vorher beschriebenen weichen abgerundeten Verlaufes ein markanter Rechtsknick an nur einer Stelle mit sonst unterschiedlichen geraden Verläufen zu sehen, so ist eine Mischverteilung sehr wahrscheinlich. Das Gleiche gilt insbesondere bei einem links verlaufenden Knick, bei dem keine ausfallfreie Zeit t_0 in Frage kommt (abgesehen von vorgeschädigten Teilen mit negativem t_0).

Mischverteilung oder nicht?

Zur Bestimmung, ob an einem vorhandenen Knick auf eine Mischverteilung geschlossen werden darf, wurde folgendes Verfahren entwickelt: Man ermittelt jeweils zwei Ausgleichsgeraden aus den Punkten im Anfangsbereich und aus denen im hinteren Bereich. Begonnen wird mit den ersten 3 Punkten. Hat eine Auswertung 10 Ausfälle, so hat der hintere Abschnitt die letzten 7 Punkte (Bild 2). Im nächsten Schritt fasst man die ersten 4 Punkte und die letzten 6 zusammen usw. Dies geht soweit, bis der zweite Abschnitt nur noch 3 Punkte hat. Eine Ausgleichsgerade im Weibull-Netz wird gebildet aus den linearisierten Werten nach folgenden Gleichungen:

$$X = \ln(t)$$

$$Y = \ln\left(\ln\left(\frac{1}{1-H}\right)\right)$$

Die Ausfallhäufigkeiten H werden für einen Stichprobenumfang $n >= 50$ bestimmt durch:

$$H = \frac{i}{n+1}$$

wobei i die Ordnungszahl der aufsteigend sortierten Ausfälle ist. Für $n<50$ gilt:

$$H = \frac{i-0.3}{n+0.4}$$

Durch die Ermittlung der Koeffizienten für die lineare Regression $Y = b \cdot X + a$ ergibt sich die später benötigte Steigung b, die mit dem Formparameter b der Weibull-Gleichung identisch ist.

Weiter vergleicht man die Korrelationskoeffizienten der jeweiligen Abschnitte und wählt einen möglichen Trennungspunkt aus, an dem die Korrelationskoeffizienten beider Ausgleichsgeraden im Mittel am besten sind. Zu beachten ist für die Bestimmung der Ausfallhäufigkeiten nach den genannten Gleichungen, dass für die Teilabschnitte für n stets der gesamte Umfang einzusetzen ist, sonst ergibt sich kein stetiger Verlauf am Trennpunkt. Möchte man sich den Rechenaufwand sparen, so kann vereinfacht der Trennpunkt auch grafisch im Weibull-Netz geschätzt werden.

Vergleich statistisch

Nun muss nur noch ein Test herangezogen werden, um festzustellen, ob die beiden Abschnitte den Verlauf der Ausfälle besser repräsentieren als eine Ausgleichsgerade über alle Punkte zusammen. Wie bei fast allen statistischen Tests stellt sich die Frage, ob die Abweichungen zufällig oder systematisch sind. Zur Behandlung dieser Fragestellung gibt es verschiedene Möglichkeiten. Eine ist der Vergleich von zwei Weibull-Verteilungen [1]. Mit obigen Gleichungen wird eine Aussagewahrscheinlichkeit ermittelt, wann zwei Verteilungen signifikant unterschiedlich sind. Bezogen auf den vorliegenden Fall sind die beiden Teilgeraden, die durch den ermittelten Trennpunkt ermittelt wurden, zu vergleichen. Für die Aussagewahrscheinlichkeit AW gilt:

$$AW = 1 - \frac{1}{e^{e^{y'}}}$$

$$y' = -0{,}3507 + 1{,}4752\,y - 0{,}1954\,y^2$$

$$y = \sqrt{1-q}\,(t_{qB}^2 - t_{qA}^2)\,\frac{\ln\left(\frac{1}{1-q}\right)}{2\sqrt{q\,t_{qA}t_{qB}\,z}}$$

$$z = \left(\frac{t_{qA}/b_A}{\sqrt{n_A}} + \frac{t_{qB}/b_B}{\sqrt{n_B}}\right)\left(\frac{t_{qB}/b_A}{\sqrt{n_A}} + \frac{t_{qA}/b_B}{\sqrt{n_B}}\right)$$

Dabei sind q der betrachtete Summen-%-Ausfallbereich, $t_{q,A}$ die Lebensdauer des ersten Teilabschnitts bei q-%-Ausfällen und $t_{q,B}$ die des zweiten, b_A der Steigungsparameter der Weibullverteilung des ersten Teilabschnitts und b_B der des zweiten sowie n_A der Stichprobenumfang des ersten Teilabschnitts und n_B der des zweiten.

Als Grenzwert dafür, wann eine Mischverteilung vorliegt, ist eine Aussagewahrscheinlichkeit von mehr als 50 % anzusetzen, denn mit dem Wert für AW wird ausgedrückt, wie groß die Wahrscheinlichkeit ist, dass die zu vergleichenden Verteilungen unterschiedlich sind.

Steigungen vergleichen

Ein weiteres neues Verfahren, das sich bereits im Einsatz bewährt hat, ist der Vergleich der beiden Steigungen mit dem Vertrauensbereich der Steigung der Gesamtgeraden. Diese ergibt sich aus:

$$b \pm b\left(1 - u_{1-\alpha/2}\,\frac{0.78}{\sqrt{n}}\right)$$

mit dem Schwellenwert $u_{1\text{-alpha}/2}$ (Tabelle 1). Die Schwellenwerte gelten für n>50. Nicht zu verwechseln ist dieser Vertrauensbereich mit dem oft im Weibull-Netz als tailliert dargestellten Bereich um die Ausgleichsgerade herum.

Liegt nun die geringere Steigung der beiden Teilabschnitte unterhalb des unteren Vertrauensbereiches und/oder die größere über dem oberen Vertrauensbereich, so ist mit einer Mischverteilung zu rechnen. Eine entsprechende Analyse und Interpretation der Steigungen ist dann für die getrennten Abschnitte durchzuführen.

Bei der Bestimmung für die Steigung b ist in den jeweiligen Abschnitten darauf zu achten, dass diese nicht aus einer getrennten Weibull-Darstellung herausgelesen werden, sondern – wie oben beschrieben – aus der linearen Korrelation berechnet werden müssen. Werden die beiden Abschnitte alleine für sich im Weibull-Netz eingetragen, ergeben sich zu große Steigungen und der Verlauf am Trennpunkt ist versetzt und nicht stetig.

Geringer Aufwand

Obwohl die beschriebenen Verfahren streng genommen erst ab Stichprobenumfängen von n ≥ 50 gültig sind, wird zur Vereinfachung ein Beispiel mit 10 Ausfällen dargestellt. Folgende Laufstrecken bis zum Ausfall wurden festgestellt: 1200, 2500, 3400, 4200, 5000, 6200, 6800, 7400, 8600, 9600 km. Nach dem eben beschriebenen Verfahren ergibt sich die Trennung der beiden Abschnitte zwischen den Punkten 5 und 6 (Bild 3):

Die Steigung des ersten Abschnittes wird zu 1,5, die des zweiten zu 2,7 berechnet. Der Vertrauensbereich der Steigung erstreckt sich damit für den gesamten Bereich bei α = 90 % über $1{,}03 < b_{ges} < 2{,}43$. Somit überschreitet die Steigung des Abschnittes 2 den Vertrauensbereich nach oben und es ist die Hypothese einer Mischverteilung zu bestätigen (Tabelle 1). Das Ergebnis ist allerdings knapp. Wäre der letzte Ausfall nicht bei 9600 km sondern bei 10 200 km erfolgt, so wäre keine Überschreitung des Vertrauensbereichs mehr gegeben.

Wendet man das erste beschriebene Verfahren über den Vergleich der Weibull-Geraden an, so ergibt sich bei einer Ausfallhäufigkeit von 10 % eine Aussagewahrscheinlichkeit von 50,4 %. Bei dem Grenzwert von 50 % zeigt sich auch hier eine knappe Bestätigung der Hypothese für die Mischverteilung. Setzt man wiederum für den letzten Wert 10 200 km ein, so ergibt sich eine Aussagewahrscheinlichkeit von 49,9 %, die Hypothese wäre gerade abzulehnen.

In einem konkreten Beispiel aus dem Motorenbereich wurde für die Wasserpumpe eines Sechszylinders festgestellt, dass im Anfangsbereich mit einer Steigung von nur wenig über 1 so genannte Zufallsausfälle vorlagen. Der hintere Abschnitt hatte im Gegensatz dazu b>2. Erst ab einer bestimmten Kilometerleistung trat also ein typisches Verschleißverhalten auf. Nur hierfür war die weitere Betrachtung mit der entsprechenden Steigung zielführend. Die Zufallsausfälle durch Vorschädigungen waren getrennt analysiert worden. Eine Fehlinterpretation wurde vermieden, und das Bauteil konnte gezielt verbessert werden.

α	90 %	95 %	99 %	99,9 %
u	1,645	1,960	2,516	3,291

Tabelle 1. Vertrauensbereich

Der Aufwand zur Durchführung ist gering und der Test einfach anzuwenden. Auf der anderen Seite erhält man die Aussage, ob tatsächlich eine Mischverteilung vorliegt oder nicht, und die Berücksichtigung einer möglichen Mischverteilung ergibt eine gezieltere Bestimmung der Steigung b.

Content in Short

Random or systematic evaluation? The determination of mixed distributions in the Weibull network. As a rule, component failure rates are investigated with the aid of the Weibull distribution theory. The borderline between pure distribution and the mixed distributions which frequently occur in practice, however, is unclear. A car manufacturer has developed a process by which a distinction can be made between random and systematically conditioned degrees of scatter.

Literatur

1 Verband der Deutschen Automobilindustrie (VDA) (Hrsg.): Qualitätskontrolle in der Automobilindustrie. Zuverlässigkeitssicherung bei Automobilherstellern und Lieferanten. Verfahren und Beispiele, Band 3. VDA, Frankfurt/M. 1984

2 Ronniger, C.: Zuverlässigkeitsanalyse mit Weibull in Entwicklung und Serie. ATZ (1999) 11, S. 942

3 Eckel, G.: Bestimmung des Anfangsverlaufs der Zuverlässigkeitsfunktion von Automobilen. QZ 22 (1977) 9, S. 206–208

4 Weibull – Diagramme, Analysen und Statistik nach Industrie-Standard. http://www.weibull.de

5 Deutsche Gesellschaft für Qualität (DGQ) (Hrsg.): Das Lebensdauernetz. Beuth Verlag, Berlin 1975

6 Sachs, L.: Angewandte Statistik. Springer, Heidelberg 1983

Der Autor dieses Beitrags

Dipl.-Ing. (FH) Curt Ronniger, geb. 1959, studierte Maschinenbau/Fahrzeugtechnik an der Fachhochschule in Coburg. Seit 1985 ist er im Motorenversuch bei der BMW AG, München, tätig. Seit 1990 ist er als SE-Team-Leiter für die Aggregateantriebe der Benzinmotoren zuständig. In dieser Zeit setzte er sich intensiv mit statistischen Methoden auseinander und spezialisierte sich auf die Weibull-Analyse.