§24 Weitere wichtige Verteilungen

▼

Wir werden uns in diesem Kapitel mit weiteren wichtigen Verteilungen befassen und diskutieren, bei welchen Fragestellungen diese Verteilungen auftreten. Speziell untersuchen wir dabei die logarithmische Normalverteilung, die im Rahmen der Risikoforschung oft verwendete Extremwertverteilung und die bei Fragen hinsichtlich der Lebensdauer von technischen Geräten oft zum Einsatz gelangende Weibullverteilung.

24.1 Die logarithmische Normalverteilung $\mathcal{LN}[\mu, \sigma]$

Wir beginnen mit einem für die logarithmische Normalverteilung $\mathcal{LN}[\mu, \sigma]$ zentralen Transformationssatz:

24.1.1 Satz: Ist die Zufallsvariable Z $\mathcal{N}[\mu, \sigma]$-verteilt, so ist die Zufallsvariable $Y = \text{Exp}[Z]$ $\mathcal{LN}[\mu, \sigma]$-verteilt. Ist die Zufallsvariable Y $\mathcal{LN}[\mu, \sigma]$-verteilt, so ist die Zufallsvariable $Z = \text{Log}[Y]$ $\mathcal{N}[\mu, \sigma]$-verteilt.

▼

Aus diesem Satz zusammen mit dem zentralen Grenzverteilungssatz folgt sofort die für Anwendungen wichtige

24.1.2 Bemerkung: Sind die positiven Zufallsvariablen $Y_1, Y_2, \ldots, Y_n$ vollständig unabhängig und identisch verteilt (wie beim zentralen Grenzverteilungssatz genügt bereits die schwächere Forderung, dass keine dieser Zufallsvariablen die anderen "dominieren" darf) und ist n groß, so ist ihr Produkt $Y_1\, Y_2 \ldots Y_n$ wegen

$$Y_1\, Y_2 \ldots Y_n = \text{Exp}[\text{Log}[Y_1\, Y_2 \ldots Y_n]] = \text{Exp}[\text{Log}[Y_1] + \text{Log}[Y_2] + \ldots + \text{Log}[Y_n]]$$

näherungsweise logarithmisch normalverteilt.

Zu dieser Bemerkung sind einige Ergänzungen angebracht:

■ Setzt sich eine Meßgröße aus vielen vollständig unabhängigen, positiven Einflussgrößen multiplikativ zusammen, so ist diese Meßgröße näherungsweise $\mathcal{LN}[\mu,\sigma]$-verteilt.

■ Der Durchmesser Z von Mahlgut (Sand, Schotter, Salz,…) und damit auch der Durchmesser von Fremdkörpereinschlüssen (in Glas, Stahl,…) ist in erster Näherung logarithmisch normalverteilt. Diesem experimentell nachzuweisenden Sachverhalt liegt die Modellvorstellung zu Grunde, dass sich der Mahlvorgang aus einer großen Anzahl von unabhängigen Brechvorgängen zusammensetzt, wobei beim i-ten Brechvorgang ein Korn mit dem zufälligen Durchmesser D in lauter Bruchstücke mit den zufälligen Durchmessern $Y_i\, D$ zerfällt. Besteht daher das Ausgangsmaterial aus Körnern mit dem zufälligen Durchmesser X so zerfallen diese beim Mahlen (bestehend aus n Brechvorgängen) in Körner mit dem zufälligen Durchmesser $Z = Y_n\, Y_{n-1} \ldots Y_1\, X$.

Wir fassen die bereits bekannten Eigenschaften der logarithmischen Normalverteilung $\mathcal{LN}[\mu, \sigma]$ zusammen:

24.1.3 Bemerkung: Die logarithmische Normalverteilung $\mathcal{LN}[\mu, \sigma]$ besitzt den Träger

$$\mathbb{T} = [0, \infty[$$

die Verteilungsdichte

$$\mathbb{f}[z] = \begin{cases} \frac{1}{\sqrt{2\pi}\; z\,\sigma}\, e^{-(\mathrm{Log}[z]-\mu)^2/(2\,\sigma^2)} & \text{für } z \ge 0 \\ 0 & \text{für } z < 0 \end{cases}$$

und die Verteilungsfunktion

$$\mathbb{F}[z] = \begin{cases} 0 & \text{für } z < 0 \\ \frac{1}{\sqrt{2\pi}\; z\,\sigma} \int_0^z e^{-(\mathrm{Log}[t]-\mu)^2/(2\,\sigma^2)}\, dt = \frac{1}{2}\,(1 + \mathrm{Erf}[\frac{\mathrm{Log}[z]-\mu}{\sqrt{2}\;\sigma}]) & \text{für } z \ge 0 \end{cases}$$

Eine $\mathcal{LN}[\mu, \sigma]$-verteilte Zufallsvariable Z besitzt den Erwartungswert

$$\mathbb{E}[Z] = e^{\mu+\sigma^2/2}$$

und die Varianz

$$\mathbb{V}[Z] = e^{2\,\mu+\sigma^2}(e^{\sigma^2} - 1)$$

Es folgen eine Reihe von Beispielen, mit denen gezeigt wird, wie sich die logarithmische Normalverteilung bei der Behandlung von konkreten Problemen einsetzen lässt:

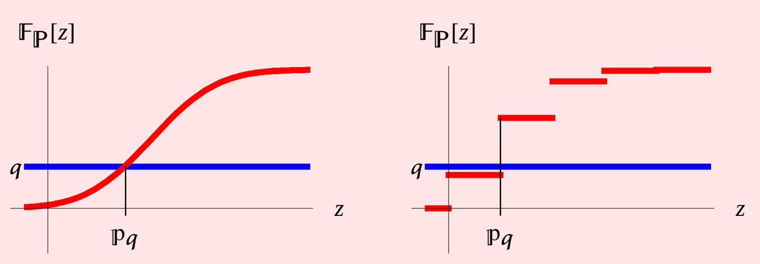
24.1.4 Beispiel: Die Zufallsvariablen $X_1, X_2, \ldots, X_n$ seien vollständig unabhängig und im Intervall $[0, 1]$ gleichverteilt. Man bestimme für großes n die Verteilung ihres Produkts $Z = X_1\, X_2 \ldots X_n$.

▼

Für das folgende Beispiel benötigen wir den Begriff des q-Quantils einer Verteilung $\mathbb{P}$, den wir im Rahmen der Statistik noch oft verwenden werden:

24.1.5 Definition: Unter dem q-Quantil $\mathbb{p}_q$ einer Verteilung $\mathbb{P}$ mit Verteilungsfunktion $\mathbb{F}_\mathbb{P}$ versteht man den eindeutig bestimmten Wert $\mathbb{p}_q = \mathrm{Min}[z \in \mathbb{R} \mid \mathbb{F}_\mathbb{P}[z] \ge q]$. Das 50%-Quantil bezeichnet man als Median.

Mit der folgenden Zeichnung soll der Begriff des q-Quantils $\mathbb{p}_q$ einer Verteilung $\mathbb{P}$ veranschaulicht werden: Links das Quantil einer stetigen Verteilung (es gilt stets $\mathbb{F}_\mathbb{P}[\mathbb{p}_q] = q$), rechts das Quantil einer diskreten Verteilung (es gilt stets $\mathbb{F}_\mathbb{P}[\mathbb{p}_q^{\,-}] \le q$ und $\mathbb{F}_\mathbb{P}[\mathbb{p}_q] > q$):



24.1.6 Beispiel: Die Beanspruchbarkeit R und die Beanspruchung S eines Bauteils seien unabhängig und mit den Parametern μ_R und σ_R bzw μ_S und σ_S logarithmisch normalverteilt. In welcher Beziehung stehen der zentrale Sicherheitsfaktor $\nu_0 = m_R / m_S$ und die Versagenswahrscheinlichkeit $p_f = \mathbb{P}[\{R < S\}]$ zueinander? (m_R bzw m_S bezeichnet dabei den Median - also das 50% Quantil - von R bzw S).

▼

24.1.7 Beispiel: Vom Durchmesser X eines zufällig ausgewählten Sandkorns einer bestimmten Sorte kennt man den Erwartungswert $\mathbb{E}[X] = 0.6$ mm und die Streuung $\mathbb{S}[X] = 0.1$ mm. Wieviel Prozent des Volumens der ganzen Sandmenge besteht aus Körnern mit einem Durchmesser, der kleiner als 0.6 mm ist?

▼

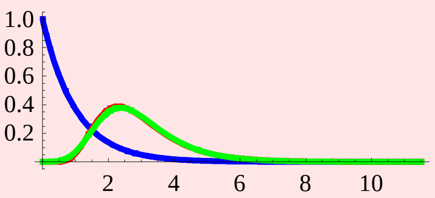
24.2 Die Extremwertverteilung $\mathcal{E}xtrem[\mu, \beta]$

Vor allem in der Risikoforschung ist man an der Verteilung des Maximums einer großen Anzahl von Zufallsvariablen interessiert. Besonderes Interesse kommt in diesem Zusammenhang dem folgenden Satz zu, den wir zwar ohne Beweis anführen, aber zumindest graphisch veranschaulichen wollen:

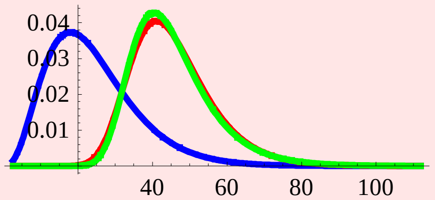
24.2.1 Satz: Sind die stetigen Zufallsvariablen $Z_1, Z_2, \ldots, Z_n$ vollständig unabhängig und identisch verteilt (mit nach oben exponentiell abnehmender Dichte), so ist für große n ihr Maximum $\text{Max}[Z_1, Z_2, \ldots, Z_n]$ näherungsweise $\mathcal{E}\text{xtrem}[\mu, \beta]$-verteilt, wobei die Parameter μ und β geeignet zu wählen sind.

Mit dem Befehl Extremwerte[dist, *n*] können wir die Aussage dieses Satzes graphisch veranschaulichen: Die blaue Kurve gibt dabei die Verteilungsdichte der Verteilung *dist* an; die rote Kurve entspricht der exakten Verteilungsdichte der Zufallsvariablen $Y = \text{Max}[Z_1, Z_2, \ldots, Z_n]$ für den Fall, dass die n Zufallsvariablen $Z_1, Z_2, \ldots, Z_n$ vollständig unabhängig und gemäß der Verteilung *dist* verteilt sind; bei der grünen Kurve handelt es sich um die Verteilungsdichte einer Extremwertverteilung, deren Parameter μ und β so gewählt sind, dass der Erwartungswert und die Varianz dieser Extremwertverteilung mit dem Erwartungswert und der Varianz der Zufallsvariablen Y übereinstimmen. Wie man durch Experimentieren zeigen kann, stimmen für stetige Verteilungen *dist*, deren Verteilungsdichten nach oben hin exponentiell abnehmen und große n die rote und die grüne Kurve weitgehend überein:

```
Extremwerte[ExponentialDistribution[1], 10]
```

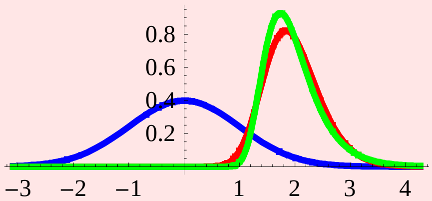



```
Extremwerte[GammaDistribution[4, 6], 10]
```

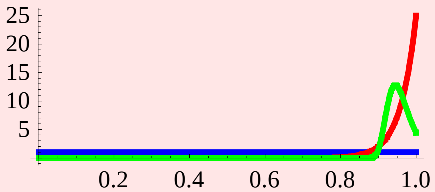


Nimmt die Verteilungsdichte der zur Diskussion stehenden Verteilung nach oben hin aber nicht exponentiell ab (die Verteilungsdichte der Normalverteilung ist für große Werte von z proportional zu $\text{Exp}[-z^2/2]$, die Verteilungsdichten der Gleichverteilung sowie der Betaverteilung sind für große Werte von z gleich 0), so gilt die Aussage von Satz 24.2.1 nicht - die roten und grünen Kurven unterscheiden sich mehr oder weniger deutlich:

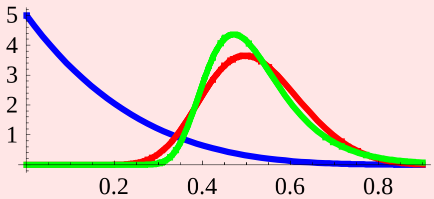
Extremwerte[NormalDistribution[0, 1], 25]



Extremwerte[UniformDistribution[{0, 1}], 25]



Extremwerte[BetaDistribution[1, 5], 25]



Wir fassen die wichtigsten Eigenschaften der Extremwertverteilung $\mathcal{E}xtrem[\mu, \beta]$ wieder zusammen:

24.2.2 Bemerkung: Die Extremwertverteilung $\mathcal{E}xtrem[\mu, \beta]$ besitzt den Träger

$$\mathbb{T} = \mathbb{R}$$

die Verteilungsdichte

$$\mathbb{f}[z] = \frac{1}{\beta}\, e^{(\mu-z)/\beta - e^{(\mu-z)/\beta}}$$

und die Verteilungsfunktion

$$\mathbb{F}[z] = \frac{1}{\beta} \int_0^z e^{(\mu-t)/\beta - e^{(\mu-t)/\beta}}\, dt = e^{-e^{(\mu-t)/\beta}}$$

Eine $\mathcal{E}xtrem[\mu, \beta]$-verteilte Zufallsvariable Z besitzt den Erwartungswert

$$\mathbb{E}[Z] = \gamma\, \beta + \mu$$

wobei $\gamma = 0.577216$ die sogenannte Euler'sche Konstante bezeichnet, und die Varianz

$$\mathbb{V}[Z] = \pi^2\, \beta^2 / 6$$

Außerdem gilt der folgende Transformationssatz: Ist die Zufallsvariable Z extremwertverteilt mit den Parametern μ und β, so ist für alle $\xi > 0$ und $\eta \in \mathbb{R}$ die Zufallsvariable $Y = \xi\, Z + \eta$ extremwertverteilt mit den Parametern $\xi\, \mu + \eta$ und $\beta\, \xi$. Daraus folgt speziell: Ist die Zufallsvariable Z extremwertverteilt mit den Parametern μ und β, so ist die Zufallsvariable $Y = (Z - \mu) / \beta$ extremwertverteilt mit den Parametern 0 und 1 (man spricht in diesem Zusammenhang von der Standardisierung der Extremwertverteilung).

▼

Mit einigen Beispielen zeigen wir wieder, wie sich die Extremwertverteilung bei der Behandlung von konkreten Problemen einsetzen lässt (man vergleiche dazu auch Beispiel 21.3.8):

24.2.3 Beispiel: Aufgrund langjähriger Beobachtungen wurde für einen Fluss bei einer Streuung von 1.50 m ein mittlerer Hochwasserstand von 5.50 m ermittelt. Im Rahmen der Planung eines Flussregulierungsprojekts soll nun ein Damm errichtet werden, dessen Krone nur in maximal 2 % aller Hochwasserfälle überschritten werden darf. Wie hoch soll der Damm werden?

▼

24.2.4 Beispiel: Aus den Aufzeichnungen der Zentralanstalt für Meteorologie geht hervor, dass die Mindesttemperatur Y im Monat Mai in Wien bei einer Streuung von 5.3 Grad einen Mittelwert von 6.2 Grad aufweist. Mit welcher Wahrscheinlichkeit muss man in Wien im Mai mit Frost rechnen?

▼

24.2.5 Beispiel: In einer seismisch aktiven Region soll ein neues Bauwerk errichtet werden. Die Zeitpunkte, zu denen in dieser Region relevante Erdbeben (das sind Erdbeben einer Stärke größer 4 auf der Richterskala) auftreten, bilden erfahrungsgemäß einen Poissonprozess mit Intensität $\lambda = 0.2$ (wobei als Zeiteinheit ein Jahr dient). Von der bei einem derartigen Erdbeben auftretenden maximalen Horizontalbeschleunigung Z kennt man auf Grund von Messdaten aus der Vergangenheit den Erwartungswert $\mathbb{E}[Z] = 0.2\,g$ und die Streuung $\mathbb{S}[Z] = 0.3\,g$. Das Bauwerk soll so bemessen werden, dass es Horizontalbeschleunigungen bis zu einem Wert von $b = 0.5\,g$ unbeschadet übersteht. Wie groß ist die Wahrscheinlichkeit dafür, dass dieses Bauwerk 25 Jahre unbeschädigt bleibt? Wie groß müsste b sein, damit die Wahrscheinlichkeit dafür, dass dieses Bauwerk 25 Jahre unbeschädigt bleibt, größer als 90 % ist?

▼

Abschließend befassen wir uns noch mit dem Zusammenhang zwischen der Exponentialverteilung $\mathcal{E}[1]$ und der Extremwertverteilung $\mathcal{E}$xtrem[0, 1]:

24.2.6 Satz: Ist die Zufallsvariable Z $\mathcal{E}xtrem[0,1]$-verteilt, so ist die Zufallsvariable $Y = e^{-Z}$ $\mathcal{E}[1]$-verteilt. Ist die Zufallsvariable Z $\mathcal{E}[1]$-verteilt, so ist die Zufallsvariable $Y = \mathrm{Log}[Z]$ $\mathcal{E}xtrem[0,1]$-verteilt.

▼

24.3 Die Weibullverteilung $\mathcal{W}[\alpha, \beta]$

Wir fassen die wichtigsten Eigenschaften der Weibullverteilung $\mathcal{W}[\alpha, \beta]$ wieder zusammen:

24.3.1 Bemerkung: Die Weibullverteilung $\mathcal{W}[\alpha, \beta]$ besitzt den Träger

$$\mathbb{T} = [0, \infty[$$

die Verteilungsdichte

$$\mathbb{f}[z] = \begin{cases} \alpha\, \beta^{-\alpha}\, z^{\alpha-1}\, e^{-(z/\beta)^\alpha} & \text{für } z \geq 0 \\ 0 & \text{sonst} \end{cases}$$

und die Verteilungsfunktion

$$\mathbb{F}[z] = \begin{cases} 1 - e^{-(z/\beta)^\alpha} & \text{für } z \geq 0 \\ 0 & \text{sonst} \end{cases}$$

Eine $\mathcal{W}[\alpha, \beta]$-verteilte Zufallsvariable Z besitzt den Erwartungswert

$$\mathbb{E}[Z] = \beta\, \Gamma[1 + 1/\alpha]$$

und die Varianz

$$\mathbb{V}[Z] = \beta^2\, (\Gamma[1 + 2/\alpha] - \Gamma[1 + 1/\alpha]^2)$$

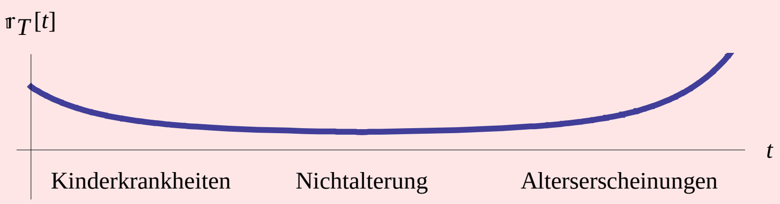
Wenn es um die Lebensdauer von technischen Geräten geht, wird sehr oft die Weibullverteilung verwendet. Der Grund dafür ist die Tatsache, dass sich mit einer Weibullverteilung das Altern eines Geräts gut beschreiben lässt.

24.3.2 Definition: Beschreibt die positive und stetige Zufallsvariable T die Lebensdauer eines Geräts, so nennt man die Abbildung $\mathbb{r}_T :]0, \infty[\to \mathbb{R}$ mit

$$\mathbb{r}_T[t] = \lim_{h\to 0} \frac{\mathbb{P}[\{T \in [t, t+h]\} \mid \{T > t\}]}{h} = \lim_{h\to 0} \frac{\mathbb{P}[\{T \in [t, t+h]\}]}{h\,\mathbb{P}[\{T > t\}]} = \frac{\mathbb{f}_T[t]}{1 - \mathbb{F}_T[t]}$$

die Risikofunktion von T. Für beliebige $t > 0$ und kleine $h > 0$ beschreibt die Größe $h\,\mathbb{r}_T[t]$ damit die Wahrscheinlichkeit dafür, dass dieses Gerät nach Erreichen des Alters t innerhalb des Intervalls $[t, t+h]$ ausfällt. Der Wert $\mathbb{r}_T[t]$ gibt somit die Ausfallsrate des Geräts zum Zeitpunkt t an.

Die Risikofunktion $\mathbb{r}_T[t]$ der Lebensdauer T eines realen technischen Geräts hat üblicherweise die Form einer sogenannten Badewannenkurve. Der linke Bereich, in dem $\mathbb{r}_T[t]$ abnimmt, beschreibt die Kinderkrankheiten. Es handelt sich dabei üblicherweise um jenen Zeitbereich, in dem Garantie geleistet wird. Der mittlere Bereich, in dem $\mathbb{r}_T[t]$ konstant ist, beschreibt jenen Bereich, in dem sich das Gerät weitgehend wie ein nichtalterndes Gerät verhält. Der rechte Bereich, in dem $\mathbb{r}_T[t]$ zunimmt, beschreibt die Alterserscheinungen des Geräts.



24.3.3 Beispiel: Man plotte die Risikofunktion $\mathbb{r}_T[t]$ der Weibullverteilung $\mathcal{W}[\alpha, \beta]$ und der Extremwertverteilung $\mathcal{E}xtrem[\mu, \beta]$. Die Parameter dieser Verteilungen sollen dabei in dynamischer Weise abgeändert werden können.

▼

24.3.4 Beispiel: Von der Lebensdauer T eines Bauteils ist die Risikofunktion

$$\mathbb{r}_T[t] = \begin{cases} 0 & \text{für } t \le 0 \\ \lambda\, t^{\mu} & \text{für } t > 0 \end{cases}$$

bekannt. Man bestimme die Verteilung der Lebensdauer T dieses Bauteils.

▼

24.3.5 Beispiel: Ein Gerät besteht in der Serienschaltung von n gleichartigen Bauteilen mit vollständig unabhängigen $\mathcal{W}[\alpha, \beta]$-verteilten Lebensdauern $T_1, T_2, \ldots, T_n$. Gesucht ist der Erwartungswert und die Varianz der Lebensdauer T dieses Geräts.

▼

24.3.6 Beispiel: Die Zufallsvariable Z sei $\mathcal{W}[\alpha, \beta]$-verteilt. Man bestimme die Verteilungsfunktion $\mathbb{F}_Y$, die Verteilungsdichte $\mathbb{f}_Y$, den Erwartungswert $\mathbb{E}[Y]$ und die Varianz $\mathbb{V}[Y]$ der Zufallsvariablen $Y = 1/Z$. (Die Verteilung dieser Zufallsvariablen nennt man Frechetverteilung mit den Parametern α und β und bezeichnet sie mit $\mathcal{F}$rechet$[\alpha, \beta]$. Sie wird vielfach als Alternative zur Extremwertverteilung verwendet.)

▼

24.3.7 Beispiel: Wir befassen uns nochmals mit Beispiel 24.2.3, setzen nun aber voraus, dass der Hochwasserstand Z dieses Flusses einer Frechetverteilung genügt.

▼

Abschließend befassen wir uns noch mit dem Zusammenhang zwischen der Weibullverteilung und der Extremwertverteilung bzw der Weibullverteilung und der Exponentialverteilung:

24.3.8 Satz: Es gelten die folgenden Transformationen:
a) Ist die Zufallsvariable Z $\mathcal{W}[\alpha, \beta]$-verteilt, so genügt $Y = -\text{Log}[Z]$ einer $\mathcal{E}$xtrem$[-\text{Log}[\beta], 1/\alpha]$-Verteilung;
b) Ist die Zufallsvariable Z $\mathcal{E}$xtrem$[\mu, \beta]$-verteilt, so genügt $Y = e^{-Z}$ einer $\mathcal{W}[1/\beta, e^{-\mu}]$-Verteilung;
c) Ist die Zufallsvariable Z $\mathcal{W}[\alpha, \beta]$-verteilt, so genügt $Y = Z^{\alpha}$ einer $\mathcal{E}[\beta^{-\alpha}]$-Verteilung;
d) Ist die Zufallsvariable Z $\mathcal{E}[\lambda]$-verteilt, so genügt $Y = \sqrt[\alpha]{Z}$ einer $\mathcal{W}[\alpha, \lambda^{-1/\alpha}]$-Verteilung.

▼