Grundlegendes zum Rechnen mit Normalverteilungen:

- Es gilt:

$$\Phi(-x) = 1 - \Phi(x)$$

  (folgt aus der Symmetrie der Dichte).

- Tabelliert sind die Werte der Verteilungsfunktion $\Phi(z) = P(Z \leq z)$ für $z \geq 0$.

  Ablesebeispiel: $\Phi(1.75) = 0.9599$

- Funktionswerte für negative Argumente: $\Phi(-z) = 1 - \Phi(z)$
- Die z-Quantile ergeben sich über die Umkehrfunktion.

  Beispielsweise ist $z_{0.9599} = 1.75$ und $z_{0.9750} = 1.96$.

Rechnen mit der Normalverteilung

- Zentrale Idee: X zu standardnormalverteilter Zufallsvariable umformen, d.h. standardisieren.
- Dabei muss die rechte Seite analog mit transformiert werden:

$$\begin{aligned} X \leq a \quad &\Leftrightarrow \quad X - \mu \leq a - \mu \\ &\Leftrightarrow \quad \frac{X-\mu}{\sigma} \leq \frac{a-\mu}{\sigma} \end{aligned}$$

das heißt

$$P(X \leq a) = P(\frac{X-\mu}{\sigma} \leq \frac{a-\mu}{\sigma}).$$

Wegen

$$\frac{X-\mu}{\sigma} \sim \mathrm{N}(0;1)$$

gilt dann

$$P(\frac{X-\mu}{\sigma} \leq \frac{a-\mu}{\sigma}) = \Phi(\frac{a-\mu}{\sigma}),$$

so dass sich der folgende Zusammenhang ergibt:

$$P(X \leq a) = \Phi(\frac{a-\mu}{\sigma}).$$

Beispiel: IQ

Der IQ st so konstruiert, dass er in der Bevölkerung normalverteilt ist mit Erwartungswert 100 und Standardabweichung 15.

- $P(IQ \leq 110) = ?$
- $P(IQ \leq 70) =$
- Wie groß ist q, damit gilt $P(IQ \geq q) = 0.01$?

Abgeschlossenheit gegenüber Linearkombinationen

Seien X_1 und X_2 unabhängig und $X_i \sim N(\mu_i, \sigma_i^2)$, $i = 1, 2$. Ferner seien b, a_1, a_2 feste reelle Zahlen. Dann gilt

$$\begin{aligned} Y_1 &:= a_1 X_1 + b \sim N(a_1\mu_1 + b; a_1^2\sigma_1^2) \\ Y_2 &:= a_1 X_1 + a_2 X_2 \sim N(a_1\mu_1 + a_2\mu_2; a_1^2\sigma_1^2 + a_2^2\sigma_2^2). \end{aligned}$$

Das Ergebnis lässt sich auf mehrere Summanden verallgemeinern.

Lebensdauerverteilungen

Zur Modellierung von Lebensdauern (im weiteren Sinne) ist die Normalverteilung selten geeignet, da

- Nur Werte > 0 möglich
- Symmetrie selten sinnvoll

Typische Verteilungen sind die Exponentialverteilung, die Gammaverteilung, und die Weibullverteilung

Lebensdauern und Ereignisanalyse

Moderner Zweig vieler empirischer Untersuchungen: Lebensdaueranalyse bzw. allgemeiner Ereignisanalyse. Im Folgenden nur eine kurze Einführung, weiterführende Texte sind

- Rohwer und Pötter (2001): *Grundzüge der sozialwissenschaftlichen Statistik*, Teil III. Juventa, Soziologische Grundlagentexte.
- Blossfeld, Hamerle, Mayer (1986): *Ereignisanalyse: Statistische Theorie und Anwendungen in den Wirtschafts- und Sozialwissenschaften*. Campus.
- Diekmann und Mitter (1984): *Methoden zur Analyse von Zeitverläufen*. Teubner.
- Blossfeld und Rohwer (1995): *Techniques of Event History Modelling*. Erlbaur.

Lebensdauerverteilungen(2)

Betrachtet wird die Zufallsgröße „Zeit bis zu einem Ereignis“, z.B. Tod, Rückkehr aus Arbeitslosigkeit, Konkurs. Um den zeitlichen Aspekt (time) zu betonen, wird die interessierende Zufallsvariable häufig mit T statt mit X bezeichnet.
Bedingt durch die spezielle Anwendung, werden in der Lebensdaueranalyse häufig nicht die Dichte oder die Verteilungsfunktion betrachtet, sondern alternative Charakterisierungen einer Wahrscheinlichkeitsverteilung.

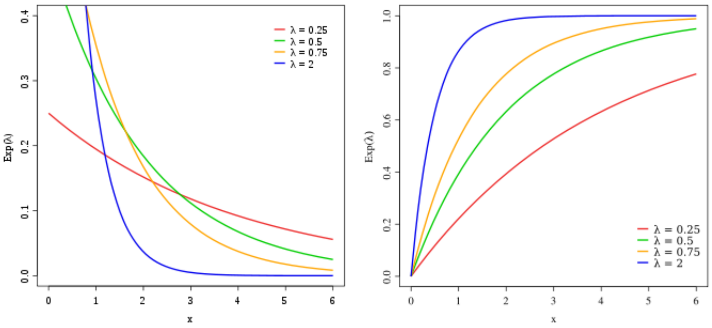
Die Exponentialverteilung

In Zeichen $\mathbf{X} \sim \mathbf{Ex}(\lambda)$ Der Parameter λ charakterisiert die Verteilung.
Der Erwartungswert (Lebenserwartung) ist $\frac{1}{\lambda}$.

1. **Modell:** X ist die Lebensdauer eines Objekts, das nicht altert.
2. **Dichte, Verteilungsfunktion und Momente**

$$
\begin{aligned}
f_X(x) &= \lambda e^{-\lambda x} \\
F_X(x) &= 1 - e^{-\lambda x} \\
E(X) &= \frac{1}{\lambda} \\
V(X) &= \frac{1}{\lambda^2}
\end{aligned}
$$

Dichte und Verteilungsfunktion



Satz

Die Verteilung einer nicht negativen, stetigen Zufallsvariable X wird eineindeutig durch die *Überlebensfunktion* (Survivorfunktion)

$$S(x) := P(X \geq x) = 1 - F(x)$$

und durch die *Hazardrate*

$$\lambda(x) := \lim_{h \to 0} \frac{P(x \leq X \leq x + h | X \geq x)}{h}$$

beschrieben.

Zur Interpretation der Hazardrate

- Teil 1: bedingte Wahrscheinlichkeit mit Argument $\{x \leq X \leq x + h\}$ (Tod zwischen den Zeitpunkten x und $x + h$)
- Teil 2: bedingendes Ereignis $\{X \geq x\}$: Überleben bis mindestens zum Zeitpunkt x
- Teil 3: Intensität relativ zur Größe des betrachteten Intervalls $[x, x + h]$ mit Breite h.
- Teil 4: Grenzwert h gegen 0 betrachten, d.h. h sehr klein machen.
- Insgesamt: grobe, anschauliche Deutung:

  Risiko, im nächsten Moment zu „sterben“, wenn man bis zum Zeitpunkt x „überlebt“ hat.
- Beachte: $\lambda(\cdot)$ ist keine Wahrscheinlichkeit, kann Werte zwischen 0 und unendlich annehmen.
- Sehr anschauliches Instrument zur Beschreibung von Lebensdauerverteilungen.

Zusammenhänge zwischen S, F und Hazard

Es gelten folgende Zusammenhänge

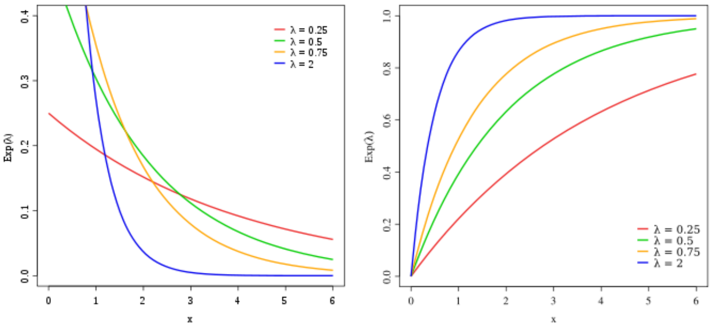
$$
\begin{aligned}
\lambda(x) &= \frac{f(x)}{S(x)} \\
S(x) &= \exp\left(-\int_0^x \lambda(u)du\right) \\
F(x) &= 1 - \exp\left(-\int_0^x \lambda(u)du\right) \\
f(x) &= \lambda(x) \cdot S(x)
\end{aligned}
$$

Warten auf den Bus

X ist die Zeit bis der Bus kommt (10 Minuten Takt) Für $x \in (0; 10)$ gilt:

$$\begin{aligned} f(x) &= 0.1 \\ F(x) &= 0.1x \\ S(x) &= 1 - 0.1x = 0.1 * (10 - x) \\ \lambda(x) &= \frac{f(x)}{S(x)} = \frac{1}{10 - x} \end{aligned}$$

Hazardrate und Survivorfunktion



Weibullverteilung

$\mathbf{X} \sim \mathbf{Wb}(\mathbf{c}, \alpha)$

1. **Modell:** Verteilung für Bruchfestigkeit von Materialien. Die Verteilung ist auch durch ihre Hazardrate charakterisiert und wird daher auch als Lebensdauerverteilung benutzt.

2. **Dichte, Verteilungsfunktion und Momente**

$$\begin{aligned} f(x) &= cx^{c-1}/\alpha^c \cdot \exp\left(-\left(\frac{x}{\alpha}\right)^c\right) \\ F(x) &= 1 - \exp\left(-\left(\frac{x}{\alpha}\right)^c\right) \end{aligned}$$

3. Hazardrate

$$\lambda(x) = \frac{f(x)}{1 - F(x)} = \frac{c}{\alpha}\left(\frac{x}{\alpha}\right)^{c-1}$$

4. Für c$=$ 1 erhält man die Exponentialverteilung

Beispiele Weibullverteilungen