Lemma 92

$$I := \int_{-\infty}^{\infty} e^{-x^2/2}\,\mathrm{d}\,x = \sqrt{2\pi}.$$

Beweis:

Wir berechnen zunächst I^2:

$$\begin{aligned} I^2 &= \left(\int_{-\infty}^{\infty} e^{-x^2/2}\,\mathrm{d}\,x\right)\left(\int_{-\infty}^{\infty} e^{-y^2/2}\,\mathrm{d}\,y\right) \\ &= \int_{-\infty}^{\infty}\int_{-\infty}^{\infty} e^{-(x^2+y^2)/2}\,\mathrm{d}\,x\,\mathrm{d}\,y\,. \end{aligned}$$

Wir gehen nun zu Polarkoordinaten über und setzen $x := r\cos\phi$ und $y := r\sin\phi$. Dann ist

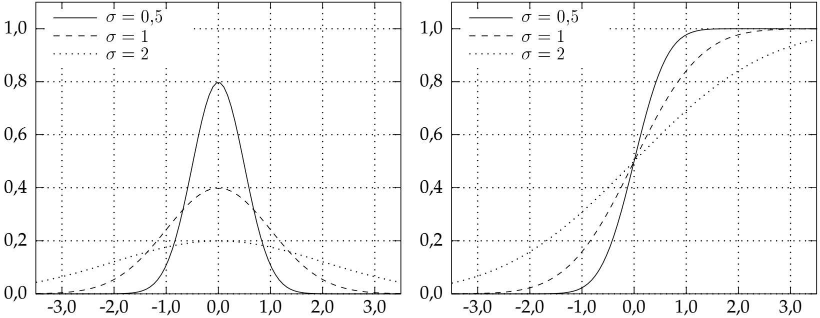
$$\begin{vmatrix} \frac{\partial x}{\partial r} & \frac{\partial y}{\partial r} \\ \frac{\partial x}{\partial \phi} & \frac{\partial y}{\partial \phi} \end{vmatrix} = \begin{vmatrix} \cos\phi & \sin\phi \\ -r\sin\phi & r\cos\phi \end{vmatrix} = r(\cos^2\phi + \sin^2\phi) = r$$

Beweis (Forts.):

und wir erhalten

$$
\begin{aligned}
I^2 &= \int_0^{2\pi} \int_0^{\infty} e^{-r^2/2} r \,\mathrm{d}\, r \,\mathrm{d}\, \phi \;=\; \int_0^{2\pi} \left[-e^{-r^2/2}\right]_0^{\infty} \mathrm{d}\, \phi \\
&= \int_0^{2\pi} 1 \,\mathrm{d}\, \phi \;=\; 2\pi.
\end{aligned}
$$

$\square$



Dichte und Verteilung von $\mathcal{N}(0, \sigma^2)$

Satz 93 (Lineare Transformation der Normalverteilung)

Sei X eine normalverteilte Zufallsvariable mit $X \sim \mathcal{N}(\mu, \sigma^2)$. Dann gilt für beliebiges $a \in \mathbb{R} \setminus \{0\}$ und $b \in \mathbb{R}$, dass $Y = aX + b$ normalverteilt ist mit $Y \sim \mathcal{N}(a\mu + b, a^2\sigma^2)$.

Beweis:

Wir betrachten zunächst den Fall „$a > 0$“:

$$\begin{aligned}\Pr[Y \leq y] = \Pr[aX + b \leq y] &= \Pr\left[X \leq \frac{y-b}{a}\right]\\ &= \frac{1}{\sqrt{2\pi}\sigma} \cdot \int_{-\infty}^{(y-b)/a} \exp\left(-\frac{(u-\mu)^2}{2\sigma^2}\right) \mathrm{d}\, u.\end{aligned}$$

Nach der Substitution $u = (v - b)/a$ und $\mathrm{d}\, u = (1/a) \cdot \mathrm{d}\, v$ erhalten wir

Beweis (Forts.):

$$\Pr[Y \le y] = \frac{1}{\sqrt{2\pi}a\sigma} \cdot \int_{-\infty}^{y} \exp\left(-\frac{(v - a\mu - b)^2}{2a^2\sigma^2}\right) \mathrm{d}\, v \,.$$

Also $Y \sim \mathcal{N}(a\mu + b, a^2\sigma^2)$. Für $a < 0$ verläuft der Beweis analog. $\square$

Sei also X eine beliebige $\mathcal{N}(\mu, \sigma^2)$-verteilte Zufallsvariable X und $Y := \frac{X-\mu}{\sigma}$.

Dann ist nach Satz 93 Y $\mathcal{N}(0,1)$-verteilt. Y heißt auch normiert.

Ferner gilt

$$\begin{aligned}\Pr[a < X \leq b] &= \Pr\left[\frac{a-\mu}{\sigma} < Y \leq \frac{b-\mu}{\sigma}\right] \\ &= \Phi\left(\frac{b-\mu}{\sigma}\right) - \Phi\left(\frac{a-\mu}{\sigma}\right)\ .\end{aligned}$$

Satz 94

X sei $\mathcal{N}(0,1)$-verteilt. Dann gilt

$$\mathbb{E}[X] = 0 \textit{ und } \quad \mathrm{Var}[X] = 1.$$

Beweis:

$$\mathbb{E}[X] = \frac{1}{\sqrt{2\pi}} \int_{-\infty}^{\infty} x \cdot \exp\left(-\frac{x^2}{2}\right) \mathrm{d}\, x.$$

Da der Integrand punktsymmetrisch zu $(0,0)$ ist, folgt $\mathbb{E}[X] = 0$.

Beweis (Forts.):

Mittels Lemma 92 und durch partielle Integration erhalten wir

$$\begin{aligned}\sqrt{2\pi} &= \int_{-\infty}^{\infty} \exp\left(-\frac{x^2}{2}\right) \mathrm{d}\,x \\ &= \underbrace{x \exp\left(-\frac{x^2}{2}\right)\Big|_{-\infty}^{\infty}}_{=\,0} + \int_{-\infty}^{\infty} x^2 \cdot \exp\left(-\frac{x^2}{2}\right) \mathrm{d}\,x\end{aligned}$$

Daraus folgt, dass $\mathbb{E}[X^2] = 1$ ist und somit $\mathrm{Var}[X] = \mathbb{E}[X^2] - \mathbb{E}[X]^2 = 1$. □

Satz 95

X sei $\mathcal{N}(\mu, \sigma^2)$-verteilt. Dann gilt

$$\mathbb{E}[X] = \mu \text{ und } \mathrm{Var}[X] = \sigma^2\,.$$

Beweis:

$Y := \frac{X-\mu}{\sigma}$ ist standardnormalverteilt. Ferner gilt gemäß der Rechenregeln für Erwartungswert und Varianz

$$\mathbb{E}[X] = \mathbb{E}[\sigma Y + \mu] = \sigma \cdot \mathbb{E}[Y] + \mu = \mu$$

und

$$\mathrm{Var}[X] = \mathrm{Var}[\sigma Y + \mu] = \sigma^2 \cdot \mathrm{Var}[Y] = \sigma^2.$$

$\square$

2.3 Exponentialverteilung

Die Exponentialverteilung ist in gewisser Weise das kontinuierliche Analogon zur geometrischen Verteilung. Wie die geometrische Verteilung ist sie „gedächtnislos“. Sie spielt daher vor allem bei der Modellierung von Wartezeiten eine große Rolle.

Definition 96

Eine Zufallsvariable X heißt exponentialverteilt mit dem Parameter λ, $\lambda > 0$, wenn sie die Dichte

$$f(x) = \begin{cases} \lambda \cdot e^{-\lambda x} & \text{falls } x \geq 0, \\ 0 & \text{sonst} \end{cases}$$

besitzt.

Für die entsprechende Verteilungsfunktion gilt (für $x \geq 0$)

$$F(x) = \int_0^x \lambda \cdot e^{-\lambda t} \, \mathrm{d}\, t = \left[-e^{-\lambda t} \right]_0^x = 1 - e^{-\lambda x}.$$

Für $x < 0$ gilt selbstverständlich $F(x) = 0$.

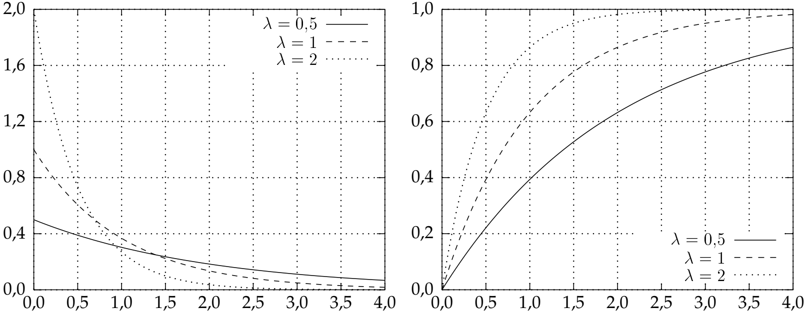
$$
\begin{aligned}
\mathbb{E}[X] &= \int_0^{\infty} t \cdot \lambda \cdot e^{-\lambda t} \,\mathrm{d}\, t \\
&= \left[t \cdot (-e^{-\lambda t})\right]_0^{\infty} + \int_0^{\infty} e^{-\lambda t} \,\mathrm{d}\, t \\
&= 0 + \left[-\frac{1}{\lambda} \cdot e^{-\lambda t}\right]_0^{\infty} = \frac{1}{\lambda}\,.
\end{aligned}
$$

Analog erhalten wir

$$\begin{aligned}\mathbb{E}[X^2] &= \int_0^\infty t^2 \cdot \lambda \cdot e^{-\lambda t} \,\mathrm{d}\, t \\ &= \left[t^2 \cdot (-e^{-\lambda t})\right]_0^\infty + \int_0^\infty 2t \cdot e^{-\lambda t} \,\mathrm{d}\, t \\ &= 0 + \frac{2}{\lambda} \cdot \mathbb{E}[X] = \frac{2}{\lambda^2}\end{aligned}$$

und somit

$$\mathrm{Var}[X] = \mathbb{E}[X^2] - \mathbb{E}[X]^2 = \frac{1}{\lambda^2}\,.$$



Dichte und Verteilung der Exponentialverteilung

2.3.1 Eigenschaften der Exponentialverteilung

Satz 97 (Skalierung exponentialverteilter Variablen)

Sei X eine exponentialverteilte Zufallsvariable mit dem Parameter λ. Für $a > 0$ ist die Zufallsvariable $Y := aX$ wieder exponentialverteilt mit dem Parameter λ/a.

Beweis:

$$\begin{aligned} F_Y(x) &= \Pr[Y \le x] = \Pr[aX \le x] \\ &= \Pr\left[X \le \frac{x}{a}\right] = F_X\left(\frac{x}{a}\right) \\ &= 1 - e^{-\frac{\lambda x}{a}} . \end{aligned}$$

$\square$

Gedächtnislosigkeit

Satz 98 (Gedächtnislosigkeit)

Eine (positive) kontinuierliche Zufallsvariable X mit Wertebereich $\mathbb{R}^+$ ist genau dann exponentialverteilt, wenn für alle $x, y > 0$ gilt, dass

$$\Pr[X > x + y \mid X > y] = \Pr[X > x]\,. \qquad (*)$$

Beweis:

Sei X exponentialverteilt mit Parameter λ. Dann gilt

$$\begin{aligned}
\Pr[X > x + y \mid X > y] &= \frac{\Pr[X > x + y, X > y]}{\Pr[X > y]} \\
&= \frac{\Pr[X > x + y]}{\Pr[X > y]} \\
&= \frac{e^{-\lambda(x+y)}}{e^{-\lambda y}} = e^{-\lambda x} = \Pr[X > x]\,.
\end{aligned}$$

Beweis (Forts.):

Sei umgekehrt X eine kontinuierliche Zufallsvariable, die die Gleichung $(*)$ erfüllt. Wir definieren $g(x) := \Pr[X > x]$. Für $x, y > 0$ gilt

$$\begin{aligned} g(x+y) &= \Pr[X > x+y] \\ &= \Pr[X > x+y \mid X > y] \cdot \Pr[X > y] \\ &= \Pr[X > x] \cdot \Pr[X > y] = g(x)g(y)\,. \end{aligned}$$

Daraus folgt durch wiederholte Anwendung

$$g(1) = g\Big(\underbrace{\frac{1}{n} + \cdots + \frac{1}{n}}_{n\text{-mal}}\Big) = \left(g\Big(\frac{1}{n}\Big)\right)^n \text{ für alle } n \in \mathbb{N}$$

und somit insbesondere auch $g(1/n) = (g(1))^{1/n}$.

Beweis (Forts.):

Da X nur positive Werte annimmt, muss es ein $n \in \mathbb{N}$ geben mit $g(1/n) > 0$. Wegen $0 < g(1) \leq 1$ muss es daher auch ein $\lambda \geq 0$ geben mit $g(1) = e^{-\lambda}$.

Nun gilt für beliebige $p, q \in \mathbb{N}$

$$g(p/q) = g(1/q)^p = g(1)^{p/q},$$

und somit $g(r) = e^{-\lambda r}$ für alle $r \in \mathbb{Q}^+$.

Aufgrund der Stetigkeit folgt daraus

$$g(x) = e^{-\lambda x}\,.$$

$\square$

Beispiel 99

Über das Cäsium-Isotop $^{134}_{55}\text{Cs}$ ist bekannt, dass es eine mittlere Lebensdauer von ungefähr 3,03 Jahren oder $1{,}55 \cdot 10^6$ Minuten besitzt. Die Zufallsvariable X messe die Lebenszeit eines bestimmten $^{134}_{55}\text{Cs}$-Atoms. X ist exponentialverteilt mit dem Parameter

$$\lambda = \frac{1}{\mathbb{E}[X]} = \frac{1}{1{,}55 \cdot 10^6} \approx 0{,}645 \cdot 10^{-6} \left[\frac{1}{\text{min}}\right]$$

Da λ den Kehrwert einer Zeit als Einheit besitzt, spricht man von der Zerfallsrate. Auch bei anderen Anwendungen ist es üblich, λ als Rate einzuführen.

2.3.2 Exponentialverteilung als Grenzwert der geometrischen Verteilung

Erinnerung: Die Poisson-Verteilung lässt sich als Grenzwert der Binomialverteilung darstellen.

Wir betrachten eine Folge geometrisch verteilter Zufallsvariablen X_n mit Parameter $p_n = \lambda/n$. Für ein beliebiges $k \in \mathbb{N}$ ist die Wahrscheinlichkeit, dass $X_n \le k \cdot n$, gleich

$$\begin{aligned}
\Pr[X_n \le kn] &= \sum_{i=1}^{kn} (1-p_n)^{i-1} \cdot p_n \;=\; p_n \cdot \sum_{i=0}^{kn-1} (1-p_n)^i \\
&= p_n \cdot \frac{1-(1-p_n)^{kn}}{p_n} \;=\; 1 - \left(1 - \frac{\lambda}{n}\right)^{kn} .
\end{aligned}$$

Wegen $\lim_{n\to\infty}(1-\frac{\lambda}{n})^n = e^{-\lambda}$ gilt daher für die Zufallsvariablen $Y_n := \frac{1}{n}X_n$, dass

$$\begin{aligned}\lim_{n\to\infty} \Pr[Y_n \le t] &= \lim_{n\to\infty} \Pr[X_n \le t\cdot n]\\ &= \lim_{n\to\infty}\left[1-\left(1-\frac{\lambda}{n}\right)^{tn}\right]\\ &= 1-e^{-\lambda t}\,.\end{aligned}$$

Die Folge Y_n der (skalierten) geometrisch verteilten Zufallsvariablen geht also für $n\to\infty$ in eine exponentialverteilte Zufallsvariable mit Parameter λ über.

3. Mehrere kontinuierliche Zufallsvariablen

3.1 Mehrdimensionale Dichten

Beobachtung

Zu zwei kontinuierlichen Zufallsvariablen X, Y wird der zugrunde liegende gemeinsame Wahrscheinlichkeitsraum über $\mathbb{R}^2$ durch eine integrierbare (gemeinsame) Dichtefunktion $f_{X,Y} : \mathbb{R}^2 \to \mathbb{R}_0^+$ mit

$$\int_{-\infty}^{\infty} \int_{-\infty}^{\infty} f_{X,Y}(x,y) \,\mathrm{d}\, x \,\mathrm{d}\, y = 1$$

beschrieben. Für ein Ereignis $A \subseteq \mathbb{R}^2$ (das aus abzählbar vielen geschlossenen oder offenen Bereichen gebildet sein muss) gilt

$$\Pr[A] = \int_A f_{X,Y}(x,y) \,\mathrm{d}\, x \,\mathrm{d}\, y.$$

Unter einem Bereich B verstehen wir dabei Mengen der Art

$$B = \{(x, y) \in \mathbb{R}^2 \mid a \leq x \leq b, c \leq y \leq d\} \quad \text{mit } a, b, c, d \in \mathbb{R}.$$

Dabei können die einzelnen Intervallgrenzen auch „offen“ bzw. $\pm\infty$ sein.

Analog zum eindimensionalen Fall ordnen wir der Dichte $f_{X,Y}$ eine (gemeinsame) Verteilung $F_{X,Y} : \mathbb{R}^2 \to [0,1]$ zu:

$$F_{X,Y}(x,y) = \Pr[X \leq x, Y \leq y] = \int_{-\infty}^{y} \int_{-\infty}^{x} f_{X,Y}(u,v) \,\mathrm{d}\,u \,\mathrm{d}\,v.$$

3.2 Randverteilungen und Unabhängigkeit

Definition 100

Sei $f_{X,Y}$ die gemeinsame Dichte der Zufallsvariablen X und Y. Die Randverteilung der Variablen X ist gegeben durch

$$F_X(x) = \Pr[X \leq x] = \int_{-\infty}^{x} \left[\int_{-\infty}^{\infty} f_{X,Y}(u,v) \, \mathrm{d}\, v \right] \mathrm{d}\, u.$$

Analog nennen wir

$$f_X(x) = \int_{-\infty}^{\infty} f_{X,Y}(x,v) \, \mathrm{d}\, v$$

die Randdichte von X. Entsprechende Definitionen gelten symmetrisch für Y.

Definition 101

Zwei kontinuierliche Zufallsvariablen X und Y heißen unabhängig, wenn

$$\Pr[X \leq x, Y \leq y] = \Pr[X \leq x] \cdot \Pr[Y \leq y]$$

für alle $x, y \in \mathbb{R}$ gilt.

Dies ist gleichbedeutend mit

$$F_{X,Y}(x,y) = F_X(x) \cdot F_Y(y)\,.$$

Differentiation ergibt

$$f_{X,Y}(x,y) = f_X(x) \cdot f_Y(y)\,.$$

Für mehrere Zufallsvariablen $X_1, \ldots, X_n$ gilt analog: $X_1, \ldots, X_n$ sind genau dann unabhängig, wenn

$$F_{X_1,\ldots,X_n}(x_1, \ldots, x_n) = F_{X_1}(x_1) \cdot \ldots \cdot F_{X_n}(x_n)$$

bzw.

$$f_{X_1,\ldots,X_n}(x_1, \ldots, x_n) = f_{X_1}(x_1) \cdot \ldots \cdot f_{X_n}(x_n)$$

für alle $x_1, \ldots, x_n \in \mathbb{R}$.

3.3 Warteprobleme mit der Exponentialverteilung

Warten auf mehrere Ereignisse

Satz 102

Die Zufallsvariablen $X_1, \ldots, X_n$ *seien unabhängig und exponentialverteilt mit den Parametern* $\lambda_1, \ldots, \lambda_n$. *Dann ist auch* $X := \min\{X_1, \ldots, X_n\}$ *exponentialverteilt mit dem Parameter* $\lambda_1 + \ldots + \lambda_n$.

Beweis:

Der allgemeine Fall folgt mittels Induktion aus dem für $n = 2$. Für die Verteilungsfunktion F_X gilt:

$$\begin{aligned} 1 - F_X(t) &= \Pr[X > t] = \Pr[\min\{X_1, X_2\} > t] \\ &= \Pr[X_1 > t, X_2 > t] \\ &= \Pr[X_1 > t] \cdot \Pr[X_2 > t] \\ &= e^{-\lambda_1 t} \cdot e^{-\lambda_2 t} = e^{-(\lambda_1 + \lambda_2)t}. \end{aligned}$$

$\square$

Anschaulich besagt Satz 102, dass sich die Raten addieren, wenn man auf das erste Eintreten eines Ereignisses aus mehreren unabhängigen Ereignissen wartet. Wenn beispielsweise ein Atom die Zerfallsrate λ besitzt, so erhalten wir bei n Atomen die Zerfallsrate $n\lambda$ (wie uns auch die Intuition sagt).

Poisson-Prozess

Wir hatten bei der Diskussion der geometrischen und der Poisson-Verteilung festgestellt:

Wenn der zeitliche Abstand der Treffer geometrisch verteilt ist, so ist ihre Anzahl in einer festen Zeitspanne binomialverteilt.

Im Grenzwert $n \to \infty$, wobei wir die Trefferwahrscheinlichkeit mit $p_n = \lambda/n$ ansetzen, konvergiert die geometrische Verteilung gegen die Exponentialverteilung und die Binomialverteilung gegen die Poisson-Verteilung. Im Grenzwert $n \to \infty$ erwarten wir deshalb die folgende Aussage:

Wenn man Ereignisse zählt, deren zeitlicher Abstand exponentialverteilt ist, so ist die Anzahl dieser Ereignisse in einer festen Zeitspanne Poisson-verteilt.

Seien $T_1, T_2 \ldots$ unabhängige exponentialverteilte Zufallsvariablen mit Parameter λ. Die Zufallsvariable T_i modelliert die Zeit, die zwischen Treffer $i-1$ und i vergeht.

Für den Zeitpunkt $t > 0$ definieren wir

$$X(t) := \max\{n \in \mathbb{N} \mid T_1 + \ldots + T_n \leq t\}.$$

$X(t)$ gibt also an, wie viele Treffer sich bis zur Zeit t (von Zeit Null ab) ereignet haben. Es gilt:

Fakt 103

Seien $T_1, T_2, \ldots$ *unabhängige Zufallsvariablen und sei* $X(t)$ *für* $t > 0$ *wie oben definiert. Dann gilt:* $X(t)$ *ist genau dann Poisson-verteilt mit Parameter* $t\lambda$*, wenn es sich bei* $T_1, T_2, \ldots$ *um exponentialverteilte Zufallsvariablen mit Parameter* λ *handelt.*

Zum Zufallsexperiment, das durch $T_1, T_2, \ldots$ definiert ist, erhalten wir für jeden Wert $t > 0$ eine Zufallsvariable $X(t)$. Hierbei können wir t als Zeit interpretieren und $X(t)$ als Verhalten des Experiments zur Zeit t. Eine solche Familie $(X(t))_{t>0}$ von Zufallsvariablen nennt man allgemein einen stochastischen Prozess. Der hier betrachtete Prozess, bei dem $T_1, T_2, \ldots$ unabhängige, exponentialverteilte Zufallsvariablen sind, heißt Poisson-Prozess und stellt ein fundamentales und zugleich praktisch sehr bedeutsames Beispiel für einen stochastischen Prozess dar.

Beispiel 104

Wir betrachten eine Menge von Jobs, die auf einem Prozessor sequentiell abgearbeitet werden. Die Laufzeiten der Jobs seien unabhängig und exponentialverteilt mit Parameter $\lambda = 1/30[1/s]$. Jeder Job benötigt also im Mittel $30s$.

Gemäß Fakt 103 ist die Anzahl von Jobs, die in einer Minute vollständig ausgeführt werden, Poisson-verteilt mit Parameter $t\lambda = 60 \cdot (1/30) = 2$.

Die Wahrscheinlichkeit, dass in einer Minute höchstens ein Job abgearbeitet wird, beträgt in diesem Fall ($t\lambda = 2$)

$$e^{-t\lambda} + t\lambda e^{-t\lambda} \approx 0{,}406\ .$$