F - 2 Unendliche Wahrscheinlichkeitsräume

Definition F.45 (Diskreter Wahrscheinlichkeitsraum)

Seien Ω eine höchstens abzählbare Menge und $P : \mathcal{P}(\Omega) \to [0,1]$ eine Funktion. Dann heißt (Ω, P) ein **diskreter Wahrscheinlichkeitsraum**, wenn folgendes gilt:

$$P(\Omega) = 1. \tag{24}$$

Für jede Folge $A_1, A_2, \ldots$ paarweiser disjunkter Teilmengen von Ω ist

$$P\Big(\bigcup_{i=1}^{\infty} A_i\Big) = \sum_{i=1}^{\infty} P(A_i). \tag{25}$$

Eigenschaft (25) heißt σ-**Additivität**.

Vorsicht: bei der Summation ist die Summierbarkeit (absolute Konvergenz) i.a. nicht gewährleistet.

Beispiel F.46 (für einen unendlichen diskreten Wahrscheinlichkeitsraum)

(Poisson-Verteilung)
Eine bestimmte Masse einer radioaktiven Substanz zerfällt. Die Anzahl der Zerfälle $X_{[0,T]}$ im Zeitintervall $[0, T]$ ist eine Zufallsvariable. Dabei nehmen wir an, dass die Gesamtzahl der radioaktiven Teilchen sich im betrachteten Zeitraum nicht wesentlich ändert. Als mathematisches Modell nehmen wir die Verteilung

$$P_\lambda(X_{[0,T]} = k) \;=\; e^{-\lambda T}\frac{(\lambda T)^k}{k!} \qquad \text{für } k \in \{0, 1, 2, \ldots\}, \tag{26}$$

mit einem Parameter $\lambda > 0$, die in der folgenden Abbildung illustriert ist.

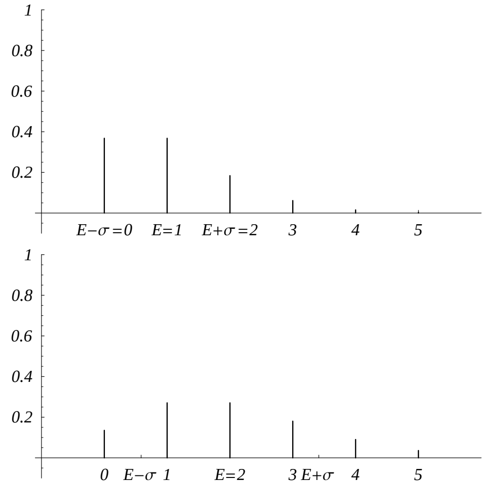

Abbildung: Stabdiagramme von Poisson-Verteilungen mit den Parametern $\lambda = 1$ und $T = 1$, bzw. $T = 2$

Es gilt für den Erwartungswert, das zweite Moment und die Varianz der Verteilung:

$$
\begin{aligned}
E(X_{[0,T]}) &= \sum_{k=0}^{\infty} k \cdot P_\lambda(X = k) = \sum_{k=0}^{\infty} k\, e^{-\lambda T}\frac{(\lambda T)^k}{k!} \\
&= \lambda T \cdot e^{-\lambda T} \sum_{k=1}^{\infty} \frac{(\lambda T)^{k-1}}{(k-1)!} = \lambda T \cdot e^{-\lambda T} \sum_{l=0}^{\infty} \frac{(\lambda T)^l}{l!} \\
&= \lambda T \cdot e^{-\lambda T} \cdot e^{\lambda T} = \lambda T, \\
E((X_{[0,T]})^2) &= \sum_{k=0}^{\infty} k^2 \cdot P_\lambda(X = k) = \ldots = (\lambda T)^2 + \lambda T
\end{aligned}
$$

(Übungsaufgabe 6, Serie 6)

$$\mathrm{Var}(X_{[0,T]}) \quad = \quad E((X_{[0,T]})^2) - (E(X_{[0,T]}))^2 = \lambda T.$$

Des weiteren gilt

$$\frac{dE(X_{[0,T]})}{dT} = \lambda,$$

d.h. λ ist die Zerfallsrate $= \frac{\text{mittlere Anzahl der Zerfälle}}{\text{Zeit}}$.

Beispiel für eine Verteilung ohne endlichen Erwartungswert siehe Übungsaufgabe 7, Serie 6.

Kontinuierliche Wahrscheinlichkeitsräume

hier: Ω Intervall, z.B. $[0,1]$, $[0,\infty[$, $]-\infty,\infty[$.

Definition F.47

(Wahrscheinlichkeitsmaße mit einer Dichtefunktion)
Sei $\Omega = [a,b]$ ein Intervall mit $a < b$. 1. Eine **Wahrscheinlichkeitsdichte** auf Ω ist eine integrierbare Funktion $f : \Omega \to \mathbb{R}$ mit

1. Nicht-Negativität:

$$f \geq 0, \text{d.h. } f(\omega) \geq 0 \text{ für alle } \omega \in \Omega.$$

2. Normiertheit:

$$\int_a^b f(\omega) d\omega = 1.$$

Die Definition im Falle von (halb-) offenen Intervallen Ω ist analog.

2. Das zur Dichte f gehörende **Wahrscheinlichkeitsmaß** P ist auf Intervallen durch

$$P([a_0, b_0]) = \int_{a_0}^{b_0} f(\omega)\, d\omega \tag{27}$$

definiert, wie in der folgenden Abbildung illustriert.

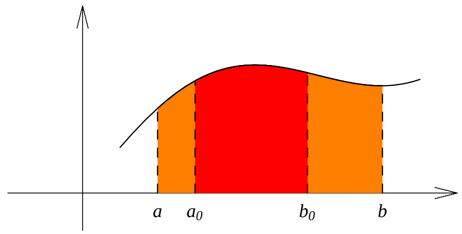

Abbildung: Wahrscheinlichkeitsdichte: Die Fläche über dem Intervall $[a_0, b_0]$ ist gleich der Wahrscheinlichkeit dieses Intervalls

3. Die Integralfunktion F von f, definiert durch

$$F(x) = \int_a^x f(\omega)\, d\omega,$$

heißt **Verteilungsfunktion** von P.

4. Eine **reelle Zufallsvariable** ist eine Funktion

$$X : \Omega \to \mathbb{R}.$$

Ihr **Erwartungswert** ist

$$E(X) := \int_a^b X(\omega) f(\omega)\, d\omega, \tag{28}$$

falls das Integral in (28) existiert, und ihre **Varianz** ist

$$\mathrm{Var}(X) := \int_a^b (X(\omega) - E(X))^2 f(\omega)\, d\omega, \tag{29}$$

sofern die Integrale in (28) und (29) existieren.

Bemerkung: Erwartungswert und Varianz einer Wahrscheinlichkeitsverteilung auf $\mathbb{R}$

Wir bezeichnen mit

$$\mu = \int_a^b x \cdot f(x)\, dx \tag{30}$$

den **Erwartungswert der Verteilung** und mit

$$\sigma^2 = \int_a^b (x - \mu)^2 f(x)\, dx \tag{31}$$

ihre **Varianz**, sofern diese Integrale existieren.

(Formaler Bezug durch die Zufallsvariable $X(x) = x$.)

Beispiel F.48

(Gleichverteilung auf einem beschränkten Intervall)
Die **Gleichverteilung** auf $[a, b]$ ist durch die Dichtefunktion

$$f : [a, b] \to \mathbb{R}, \qquad x \mapsto \frac{1}{b - a},$$

gegeben.

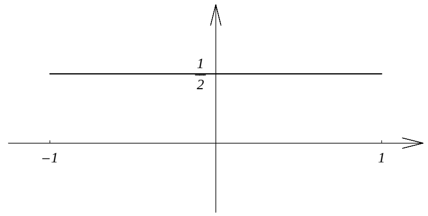

Abbildung: Gleichverteilung auf dem Intervall $[-1, 1]$

Es gelten

$$f(x) = \frac{1}{b - a} > 0$$

und

$$\int_a^b f(x)\, dx = 1,$$

d.h. f ist also tatsächlich eine Wahrscheinlichkeitsdichte.

Sei X eine Zufallsvariable, deren Verteilung die Dichte f hat, also $X = x$.
Der Erwartungswert ist

$$E(X) \quad = \quad \int_a^b \frac{1}{b - a} \cdot x\, dx = \frac{1}{b - a} \cdot \frac{1}{2}(b^2 - a^2) = \frac{b + a}{2},$$

also gleich dem Mittelpunkt des Intervalls $[a, b]$.

Zur Berechnung der Varianz benutzen wir

$$\mathrm{Var}(X) \quad = \quad E\big((X - E(X))^2\big) = E(X^2) - \big(E(X)\big)^2.$$

Wir müssen also noch das **zweite Moment** $E(X^2)$ von X berechnen.

$$E(X^2) \quad = \quad \int_a^b \frac{1}{b-a} x^2 \, dx = \frac{1}{b-a} \cdot \frac{1}{3}(b^3 - a^3) = \frac{1}{3}(b^2 + ab + a^2).$$

Damit erhalten wir

$$\mathrm{Var}(X) \quad = \quad \frac{1}{3}(b^2 + ab + a^2) - \frac{1}{4}(b^2 + 2ab + a^2) = \frac{1}{12}(b-a)^2.$$

Die Varianz hängt also nur von der Intervalllänge ab. Physikalisch kann man den Erwartungswert von X als *Schwerpunkt* bei homogener Massenverteilung interpretieren, und die Varianz ist proportional zum **Trägheitsmoment**, also proportional zum **mittleren quadratischen Abstand** zum Schwerpunkt.

Beispiel F.49 (Exponentialverteilungen auf $[\mathbf{0}, \infty)$)

Die **Exponentialverteilung** mit Parameter $\lambda > 0$ ist gegeben durch die Dichte

$$f_\lambda : [0, \infty) \to \mathbb{R}, \qquad t \mapsto \lambda e^{-\lambda t}.$$

Sie tritt z.B. beim durch den Poisson-Prozeß modellierten radioaktiven Zerfall auf (s. Beispiel F.46) Die **Wartezeit** bis zum ersten Zerfall ist eine Zufallsvariable, deren Verteilung die Dichte f_λ hat.

(siehe auch Übungsaufgabe 8, Serie 6)

Beispiel F.50 (Normalverteilungen)

Die **Normalverteilung** $\mathcal{N}(\mu, \sigma^2)$ mit Erwartungswert μ und Varianz σ^2 hat die Dichte

$$f_{\mu,\sigma^2}(x) = \frac{1}{\sigma\sqrt{2\pi}} e^{\left(\frac{-(x-\mu)^2}{2\sigma^2}\right)}. \tag{32}$$

Die Normalverteilung $\mathcal{N}(0,1)$ mit Erwartungswert 0 und Varianz 1 heißt **Standard-Normalverteilung**.

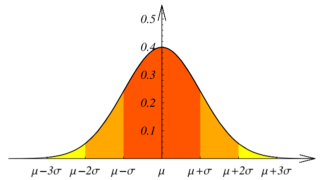

Abbildung: Die Standard-Normalverteilung mit ihrem σ-, 2σ- und 3σ-Intervall

Durch die Normalverteilung werden viele gestreute Größen, wie z.B. Körperlängen von Personen in einer Bevölkerung beschrieben, allerdings nur in einem hinreichend kleinen Intervall um die Durchschnittsgröße herum, denn natürlich gibt es keinen Menschen mit negativer Größe oder von 3m Länge. Solche Verteilungen haben mit den Normalverteilungen die typische Glockenform gemeinsam. Mathematisch wird der Zustand zwischen der Normalverteilung und mehrfach wiederholten Experimenten (z.B. mehrfacher Münzwurf) durch den **zentralen Grenzwertsatz** (Satz F.53) hergestellt.

$f_{\mu,\sigma^2}(x)$ ist eine Wahrscheinlichkeitsdichte, d.h. $f_{\mu,\sigma^2}(x) \geq 0 \; \forall x$ und Normiertheit ist erfüllt:

Das uneigentliche Integral $0 < \int\limits_{-\infty}^{\infty} e^{-x^2} dx < \infty$ existiert (Majorante).

Zu der Funktion e^{-x^2} gibt es keine *elementare* Stammfunktion.
Man kann aber berechnen: (Transformation in Polarkoordinaten)

$$\int\limits_{-\infty}^{\infty} e^{-x^2} dx = \sqrt{\pi}$$

Wir erhalten die Normiertheit der Dichtefunktion:

$$\int\limits_{-\infty}^{\infty} \frac{1}{\sigma\sqrt{2\pi}} e^{\left(\frac{-(x-\mu)^2}{2\sigma^2}\right)} dx = 1$$

Erwartungswert und Varianz einer $\mathcal{N}(\mu, \sigma^2)$-verteilten Zufallsvariablen X_{μ,σ^2}:

$$E(X_{\mu,\sigma^2}) = \int\limits_{-\infty}^{\infty} x \cdot f_{\mu,\sigma^2}(x) \, dx = \mu$$

$$\mathrm{Var}(X_{\mu,\sigma^2}) = E(X_{0,\sigma^2}^2) - E(X_{0,\sigma^2})^2 = \sigma^2 - 0 = \sigma^2$$

(invariant bezüglich Verschiebung)

Verteilungsfunktion der Standard-Normalverteilung

Definition F.51

Die **Verteilungsfunktion** (s. Definition F.47) **der Standard-Normalverteilung** ist

$$\begin{aligned} \Phi : \mathbb{R} &\rightarrow \mathbb{R}, \\ \Phi(z) &= \int_{-\infty}^{z} f_{0,1}(x) \, dx. \end{aligned}$$

Graphen der Dichte $f_{0,1}$ und von Φ siehe Abbildung.

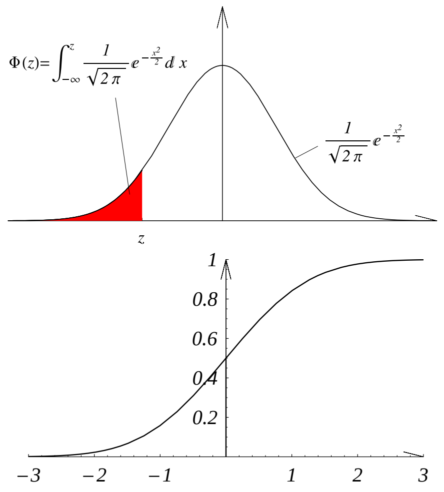

Abbildung: Die Standard-Normalverteilung und ihre Verteilungsfunktion

Bemerkung zur Verteilungsfunktion der Standard - Normalverteilung

- Es gibt keine Darstellung von Φ durch *elementare* Funktionen.
- Werte von Φ lassen sich aber beliebig genau numerisch berechnen, und für diskrete Werte von z liegen die Funktionswerte tabellarisch vor (z.B. Bronstein, Taschenbuch der Mathematik).
- Dadurch kann man schnell Integrale der Form

$$\int_a^b f_{0,1}(x)\,dx = \Phi(b) - \Phi(a)$$

auswerten.
- Wegen

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

enthalten solche Tabellen z.B. nur die Werte für nicht-negative z.
- Für symmetrische Intervalle $[-z, z]$ (mit $z > 0$) gilt:

$$\int_{-z}^{z} f_{0,1}(x)\,dx = \Phi(z) - \Phi(-z) = \Phi(z) - (1 - \Phi(z)) = 2\Phi(z) - 1.$$

Einige spezielle Werte von Φ:

$$\begin{aligned}
\Phi(0) &= 0.5, \\
\Phi(1) &\approx 0.8413 \quad \Rightarrow \quad \int_{-1}^{1} f_{0,1}(y)\,dy \approx 0.6826, \\
\Phi(2) &\approx 0.9772 \quad \Rightarrow \quad \int_{-2}^{2} f_{0,1}(y)\,dy \approx 0.9544, \\
\Phi(3) &\approx 0.9986 \quad \Rightarrow \quad \int_{-3}^{3} f_{0,1}(y)\,dy \approx 0.9972.
\end{aligned}$$

Aus der zweiten Zeile folgt z.B., dass bei irgendeiner Normalverteilung dem Intervall $[\mu - \sigma, \mu + \sigma]$ mit Radius σ (Streuung) um den Erwartungswert μ herum eine Wahrscheinlichkeit von etwa 68% zugeordnet wird. Bei einem Experiment mit vielen voneinander unabhängigen $\mathcal{N}(\mu, \sigma^2)$-verteilten Messungen liegen ungefähr 68% der Meßwerte in diesem Intervall.

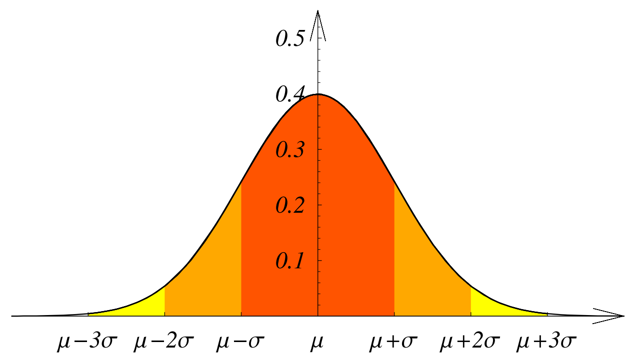

Abbildung: Die Standard-Normalverteilung mit ihrem σ-, 2σ- und 3σ-Intervall

Definition F.52 (α-Quantile der $\mathcal{N}(\mu, \sigma^2)$-Verteilung)

Sei $\alpha \in]0,1[$. Das α-**Quantil** der Standard-Normalverteilung ist die Zahl $z \in \mathbb{R}$ mit

$$\alpha = \int_{-\infty}^{z} f_{0,1}(x)\,dx = \Phi(z),$$

also

$$z = \Phi^{-1}(\alpha).$$

Bemerkung: Quantile für allgemeine Verteilungen, Median

Man kann α-Quantile allgemein für (diskrete oder kontinuierliche) reelle Verteilungen definieren.
Das $\frac{1}{2}$-Quantil heißt **Median** der Verteilung. Im Falle einer kontinuierlichen Verteilung auf einem Intervall $[a, b]$ mit überall positiver Dichte f ist der Median m die durch die Bedingung $P([a, m]) = \frac{1}{2}$ eindeutig festgelegte Zahl. Der Median ist im allgemeinen vom Erwartungswert verschieden.

Wahrscheinlichkeit: Sei X $\mathcal{N}(\mu, \sigma^2)$-verteilt.

$$\begin{aligned} P(X \in [a;b]) &= \int_a^b f_{\mu,\sigma^2}(x)dx \\ &= \int_a^b \frac{1}{\sigma} f_{0,1}\left(\frac{x-\mu}{\sigma}\right) dx \\ &= \int_{\frac{a-\mu}{\sigma}}^{\frac{b-\mu}{\sigma}} f_{0,1}(z)dz \end{aligned}$$

Verteilungsfunktion:

$$\Phi(z) = \int_{-\infty}^{z} f_{0,1}(z)dz$$

$$P(X \in [a;b]) = \Phi\left(\frac{b-\mu}{\sigma}\right) - \Phi\left(\frac{a-\mu}{\sigma}\right)$$

(Anwendung in Übungsaufgabe 5, Serie 6)

Transformation einer beliebigen Normalverteilung in die Standard-Normalverteilung

- Normalverteilung $\mathcal{N}(\mu, \sigma^2)$ (Erwartungswert μ, Varianz: σ^2)

$$f_{\mu,\sigma^2}(x) = \frac{1}{\sigma\sqrt{2\pi}} e^{\left(\frac{-(x-\mu)^2}{2\sigma^2}\right)}$$

- Standard-Normalverteilung $\mathcal{N}(0,1)$ (Erwartungswert 0, Varianz: 1)

$$f_{0,1}(x) = \frac{1}{\sqrt{2\pi}} e^{\left(\frac{-x^2}{2}\right)}$$

Umrechnung:

$$f_{\mu,\sigma^2}(x) = \frac{1}{\sigma\sqrt{2\pi}} e^{\left(\frac{-(x-\mu)^2}{2\sigma^2}\right)} = \frac{1}{\sigma}\frac{1}{\sqrt{2\pi}} e^{\left(-\frac{1}{2}\left(\frac{x-\mu}{\sigma}\right)^2\right)} = \frac{1}{\sigma} f_{0,1}\left(\frac{x-\mu}{\sigma}\right)$$

Der zentrale Grenzwertsatz, den wir hier in einer speziellen Version formulieren, erklärt die herausragende Bedeutung von Normalverteilungen für die Wahrscheinlichkeitstheorie und Statistik.

Satz F.53 (Zentraler Grenzwertsatz)

Sei $X_1, X_2, \ldots$ eine Folge von auf demselben Wahrscheinlichkeitsraum (Ω, P) definierten, paarweise unabhängigen reellen Zufallsvariablen, die alle dieselbe Verteilung haben mit

$$E(X_i) = \mu, \quad Var(X_i) = \sigma^2 > 0.$$

Sei $X^{(n)} = X_1 + \ldots X_n$, und sei $Z^{(n)} = \frac{X^{(n)} - n\mu}{\sigma\sqrt{n}}$. (Somit hat $Z^{(n)}$ den Erwartungswert 0 und die Varianz 1.)

Dann gilt für jedes Intervall $[a_0, b_0] \subset \mathbb{R}$:

$$\lim_{n\to\infty} P(Z^{(n)} \in [a_0, b_0]) = \int_{a_0}^{b_0} f_{0,1}(x)\,dx.$$

wobei $f_{0,1}$ die Dichte der Standard-Normalverteilung ist. Äquivalent dazu können wir schreiben:

$$\lim_{n\to\infty} P\left(\frac{X^{(n)} - n\mu}{\sigma\sqrt{n}} \in [a_0, b_0]\right) = \int_{a_0}^{b_0} f_{0,1}(x)\,dx.$$

Beispiel F.54 (Binomialverteilung für große n)

Die Binomialverteilung mit gegebenem Erfolgsparameter p wird für große n ungefähr gleich einer $\mathcal{N}(np, np(1-p))$ Normalverteilung:

$$P(k) = \binom{n}{k} p^k (1-p)^{n-k} \approx \frac{1}{\sqrt{2\pi}\sigma} e^{-\frac{(k-\mu)^2}{2\sigma^2}} \text{ mit } \mu = np \text{ und } \sigma^2 = np(1-p).$$

Dieser Sachverhalt, der für $p = 0.3$ und $n = 100$ in der folgenden Abbildung illustriert ist, folgt direkt aus dem zentralen Grenzwertsatz, denn die binomialverteilte Zufallsvariable K kann als Summe vieler unabhängiger Zufallsvariablen X_i aufgefasst werden, die jeweils nur die Werte 0 oder 1 (jeweils mit Wahrscheinlichkeit $(1-p)$ bzw. p) annehmen, und die den Erwartungswert p und die Varianz $p(1-p)$ haben.

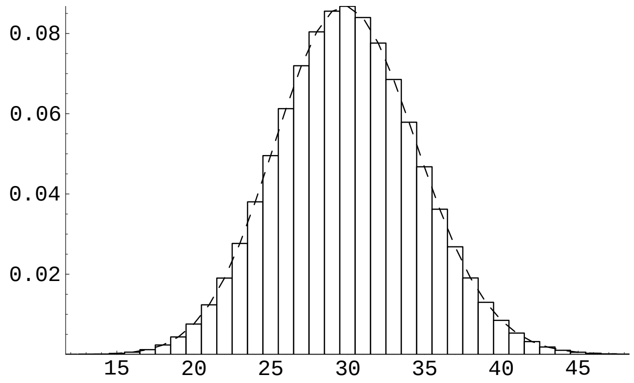

Abbildung: Histogramm der Binomialverteilung für $n = 100$ und $p = 0.3$, verglichen mit der $\mathcal{N}(np, np(1-p))$ Verteilung.