Eine Auswahl wichtiger Definitionen und Aussagen zur Vorlesung »Stochastik für Informatiker und Regelschullehrer«

Werner Linde

WS 2008/09

Inhaltsverzeichnis

1 Wahrscheinlichkeiten

1.1 Wahrscheinlichkeitsräume

1.1.1 Grundraum

Der **Grundraum** (meist mit Ω bezeichnet) ist eine Menge, die mindestens alle bei einem stochastischen Versuch oder Vorgang auftretenden Ergebnisse enthält. Die Teilmengen von Ω heißen **Ereignisse**, die einpunktigen Teilmengen nennt man **Elementarereignisse**.

1.1.2 Eintreten eines Ereignisses

Ein Ereignis $A \subseteq \Omega$ **tritt ein**, wenn das beim Versuch oder dem Vorgang beobachtete zufällige Ergebnis in der Menge A liegt.

1.1.3 σ-Algebra

Auf dem Grundraum Ω wird ein System $\mathcal{A} \subseteq \mathcal{P}(\Omega)$ von Ereignissen ausgezeichnet, denen man in sinnvoller Weise die Wahrscheinlichkeit ihres Eintretens zuordnen kann. Aus naheliegenden Gründen fordert man, dass $\mathcal{A}$ eine **σ-Algebra** bildet, d. h. $\mathcal{A}$ erfüllt folgende Eigenschaften:

(i) $\emptyset \in \mathcal{A}$.

(ii) Aus $A \in \mathcal{A}$ folgt $A^c \in \mathcal{A}$.

(iii) $A_1, A_2, \ldots \in \mathcal{A}$ impliziert $\bigcup_{j=1}^{\infty} A_j \in \mathcal{A}$.

Ist Ω höchstens abzählbar unendlich, so kann man als σ-Algebra stets die Potenzmenge $\mathcal{P}(\Omega)$ von Ω nehmen.

1.1.4 Wahrscheinlichkeitsmaß

Ein **Wahrscheinlichkeitsmaß** (oder eine **Wahrscheinlichkeitsverteilung**) $\mathbb{P}$ ist eine Abbildung von $\mathcal{A}$ nach [0,1], die jedem Ereignis $A \in \mathcal{A}$ die Wahrscheinlichkeit seines Eintretens zuordnet und folgende Eigenschaften besitzt:

(i) Es gilt $\mathbb{P}(\emptyset) = 0$ und $\mathbb{P}(\Omega) = 1$.

(ii) $\mathbb{P}$ ist σ-additiv, d. h. für disjunkte $A_j \in \mathcal{A}$ folgt

$$\mathbb{P}\Big(\bigcup_{j=1}^{\infty} A_j\Big) = \sum_{j=1}^{\infty} \mathbb{P}(A_j) \; .$$

1.1.5 Wahrscheinlichkeitsraum

Das Tripel $(\Omega, \mathcal{A}, \mathbb{P})$ heißt **Wahrscheinlichkeitsraum**. Zufällige Experimente werden durch geeignete Wahrscheinlichkeitsräume beschrieben.

1.1.6 Eigenschaften von Wahrscheinlichkeitsmaßen

(i) Jedes Wahrscheinlichkeitsmaß ist auch **endlich additiv**, d. h. sind $A_1, \ldots, A_n$ aus $\mathcal{A}$ disjunkt, so folgt

$$\mathbb{P}\Big(\bigcup_{j=1}^{n} A_j\Big) = \sum_{j=1}^{n} \mathbb{P}(A_j) \ .$$

(ii) Wahrscheinlichkeitsmaße sind **monoton**, d. h. gilt für $A, B \in \mathcal{A}$ die Inklusion $A \subseteq B$, so impliziert dies $\mathbb{P}(A) \leq \mathbb{P}(B)$.

(iii) Für $A, B \in \mathcal{A}$ mit $A \subseteq B$ folgt $\mathbb{P}(B \setminus A) = \mathbb{P}(B) - \mathbb{P}(A)$. Insbesondere ergibt sich hieraus $\mathbb{P}(A^c) = \mathbb{P}(\Omega \setminus A) = 1 - \mathbb{P}(A)$ für $A \in \mathcal{A}$.

(iv) Wahrscheinlichkeitsmaße sind **stetig von oben**, d. h. gilt für $A_j \in \mathcal{A}$ die Aussage $A_1 \supseteq A_2 \supseteq \cdots$, so folgt

$$\mathbb{P}\Big(\bigcap_{j=1}^{\infty} A_j\Big) = \lim_{j\to\infty} \mathbb{P}(A_j) \ .$$

(v) Wahrscheinlichkeitsmaße sind auch **stetig von unten**, d. h. gilt für $A_j \in \mathcal{A}$ die Aussage $A_1 \subseteq A_2 \subseteq \cdots$, so folgt

$$\mathbb{P}\Big(\bigcup_{j=1}^{\infty} A_j\Big) = \lim_{j\to\infty} \mathbb{P}(A_j) \ .$$

1.2 Typen von Wahrscheinlichkeitsmaßen

1.2.1 Wahrscheinlichkeitsmaße auf höchstens abzählbar unendlichen Grundräumen

Bei einem Experiment seien höchstens abzählbar unendlich viele Versuchsergebnisse möglich. Dann kann man entweder $\Omega = \{\omega_1, \ldots, \omega_N\}$ oder aber $\Omega = \{\omega_1, \omega_2, \ldots\}$ wählen. Als σ-Algebra nimmt man in diesen Fällen stets die Potenzmenge $\mathcal{P}(\Omega)$. Setzt man

$$p_i := \mathbb{P}(\{\omega_i\}) \,, \quad 1 \leq i \leq N \quad \text{bzw.} \quad i = 1{,}2, \ldots , \tag{1}$$

dann erhält man Zahlen mit den Eigenschaften

(i) $p_i \geq 0$ und

(ii) $\sum_{i=1}^{N} p_i = 1$ bzw. $\sum_{i=1}^{\infty} p_i = 1$.

Für eine Menge $A \subseteq \Omega$ folgt dann

$$\mathbb{P}(A) := \sum_{\{i\colon \omega_i \in A\}} p_i \,. \tag{2}$$

Umgekehrt, gibt man eine Folge $(p_i)_{i\geq 1}$ reeller Zahlen mit Punkt (i) und Punkt (ii) vor, so wird durch Gleichung 2 ein Wahrscheinlichkeitsmaß $\mathbb{P}$ auf $\mathcal{P}(\Omega)$ definiert. Für endliche oder abzählbar unendliche Grundräume Ω hat man also folgende Äquivalenz:

$$\{\mathbb{P}\colon \mathbb{P} \text{ Wahrscheinlichkeitsmaß auf } \mathcal{P}(\Omega)\} \Longleftrightarrow \{(p_i)_{i\geq 1}\colon (p_i)_{i\geq 1} \text{ erfüllen } ((\text{i})) \text{ und } ((\text{ii}))\}$$

Die Zuordnung erfolgt über Gleichung 1 bzw. Gleichung 2.

1.2.2 Diskrete Wahrscheinlichkeitsmaße

Sei nunmehr Ω ein beliebiger Grundraum (nicht notwendig endlich oder abzählbar unendlich). Ein Wahrscheinlichkeitsmaß $\mathbb{P}$ auf $(\Omega, \mathcal{P}(\Omega))$ heißt **diskret**, wenn es eine höchstens abzählbar unendliche Teilmenge $D \subseteq \Omega$ mit $\mathbb{P}(D) = 1$ gibt. Mit $D = \{\omega_1, \omega_2, \dots\}$ gilt dann für $A \subseteq \Omega$ wie zuvor

$$\mathbb{P}(A) := \sum_{\{i\colon \omega_i \in A\}} p_i \,,$$

wobei $p_i := \mathbb{P}(\{\omega_i\})$. Auf höchstens abzählbar unendlichen Grundräumen ist somit **jedes** Wahrscheinlichkeitsmaß diskret.

1.2.3 Wahrscheinlichkeitsdichten

Eine stückweise stetige Funktion $p\colon \mathbb{R} \mapsto \mathbb{R}$ heißt **Wahrscheinlichkeitsdichte**, wenn

(i) $p(x) \geq 0$ für $x \in \mathbb{R}$ und

(ii) $\int_{-\infty}^{\infty} p(x)\,\mathrm{d}x = 1$

gelten.

1.2.4 Borel-σ-Algebra

Mit $\mathcal{B}(\mathbb{R})$ bezeichnet man die kleinste σ-Algebra von Mengen aus $\mathbb{R}$, die die halboffenen Intervalle enthält. Man nennt $\mathcal{B}(\mathbb{R})$ die σ-Algebra der **Borelmengen**. Elemente von $\mathcal{B}(\mathbb{R})$ sind z. B. alle offenen oder abgeschlossenen Mengen, deren abzählbaren Vereinigungen und Durchschnitte usw.

1.2.5 Stetige Wahrscheinlichkeitsmaße

Gegeben sei eine Wahrscheinlichkeitsdichte p. Dann existiert ein eindeutig bestimmtes Wahrscheinlichkeitsmaß $\mathbb{P}\colon \mathcal{B}(\mathbb{R}) \mapsto [0{,}1]$ mit

$$\mathbb{P}([\alpha,\beta]) = \mathbb{P}((\alpha,\beta]) = \int\limits_{\alpha}^{\beta} p(x)\,\mathrm{d}x$$

für alle reelle Zahlen $\alpha < \beta$. Das so erzeugte Wahrscheinlichkeitsmaß $\mathbb{P}$ heißt **stetig** und p nennt man die **Dichte** von $\mathbb{P}$. Stetige Wahrscheinlichkeitsmaße beschreiben Vorgänge, bei denen überabzählbar viele reelle Zahlen als Ergebnis auftreten können (z. B. Lebenszeiten, Messwerte etc.).

1.3 Die wichtigsten diskreten Wahrscheinlichkeitsverteilungen

1.3.1 Einpunktverteilung

Gegeben sei ein $\omega_0 \in \Omega$, fest aber beliebig. Dann wird durch

$$\delta_{\omega_0}(A) := \begin{cases} 1 & : \omega_0 \in A \\ 0 & : \omega_0 \notin A \end{cases}$$

die **Einpunktverteilung** in ω_0 (oder das **Diracsche $\boldsymbol{\delta}$-Maß** in ω_0) definiert. Der Wahrscheinlichkeitsraum $(\Omega, \mathcal{P}(\Omega), \delta_{\omega_0})$ beschreibt Vorgänge, bei denen mit Wahrscheinlichkeit 1 genau ω_0 eintritt (deterministische Vorgänge).

1.3.2 Gleichverteilung auf N Punkten

Gegeben seien N Punkte $\omega_1, \ldots, \omega_N \in \Omega$. Das Maß $\mathbb{P}$ auf $\mathcal{P}(\Omega)$ mit

$$\mathbb{P} := \frac{1}{N} \sum_{i=1}^{N} \delta_{\omega_i}$$

heißt **Gleichverteilung** auf $\{\omega_1, \ldots, \omega_N\}$. Für ein Ereignis A gilt dann

$$\mathbb{P}(A) = \frac{\operatorname{card}\{i \le N \colon \omega_i \in A\}}{N} = \frac{\text{Anzahl der günstigen Fälle für A}}{\text{Anzahl der möglichen Fälle}} \,.$$

1.3.3 Binomialverteilung

Sei $\Omega = \{0, \ldots, n\}$ und sei $p \in [0,1]$ vorgegeben. Dann wird durch

$$B_{n,p}(\{k\}) := \binom{n}{k} p^k (1-p)^{n-k}, \quad k = 0, \ldots, n \ ,$$

ein Wahrscheinlichkeitsmaß $B_{n,p}$ auf $\mathcal{P}(\Omega)$ definiert. Man nennt $B_{n,p}$ **Binomialverteilung** mit den Parametern n und p. Die Zahl $B_{n,p}(\{k\})$ gibt die Wahrscheinlichkeit an, dass man bei n unabhängigen Versuchen genau k-mal Erfolg hat. Dabei ist die Erfolgswahrscheinlichkeit in jedem einzelnen Versuch p, die für Misserfolg $1-p$.

1.3.4 Hypergeometrische Verteilung

Gegeben seien Zahlen

$M, N, n \in \mathbb{N}_0$ mit $M, n \leq N$. Dann wird durch

$$H_{N,M,n}(\{m\}) := \frac{\binom{M}{m}\binom{N-M}{n-m}}{\binom{N}{n}}$$

ein Wahrscheinlichkeitsmaß $H_{N,M,n}$ auf $\mathcal{P}(\{0, \ldots, n\})$ definiert. Man nennt $H_{N,M,n}$ **hypergeometrische Verteilung** mit den Parametern N, M und n. Sind in einer Lieferung von N Geräten M Stück defekt, so beschreibt $H_{N,M,n}(\{m\})$ die Wahrscheinlichkeit, dass man in einer zufällig entnommenen Stichprobe vom Umfang n genau m defekte Geräte beobachtet.

1.3.5 Poissonverteilung

Es sei $\Omega = \mathbb{N}_0 = \{0,1,2, \ldots\}$. Für eine Zahl $\lambda > 0$ definiert man die **Poissonverteilung** mit Parameter λ durch

$$P_\lambda(\{k\}) := \frac{\lambda^k}{k!} \mathrm{e}^{-\lambda} \quad \text{für} \quad k \in \mathbb{N}_0 \ .$$

Die Bedeutung der Poissonverteilung ergibt sich aus folgendem Satz:

Satz 1.1 *Gegeben sei eine Zahl $\lambda > 0$. Für $n \in \mathbb{N}$ setze man $p_n := \lambda/n$. Dann folgt für alle $k \in \mathbb{N}_0$ stets*

$$\lim_{n \to \infty} B_{n,p_n}(\{k\}) = P_\lambda(\{k\}) \ .$$

Inhaltlich bedeutet dies: Führt man sehr viele unabhängige Versuche durch (n Stück), bei denen jeweils nur mit sehr kleiner Wahrscheinlichkeit p Erfolg eintreten kann, so ist die Anzahl der insgesamt beobachteten Erfolge approximativ gemäß P_λ verteilt, wobei $\lambda = n \cdot p$.

1.3.6 Geometrische Verteilung

Bei einem einzelnen Versuch trete Erfolg wieder mit Wahrscheinlichkeit p und Misserfolg mit Wahrscheinlichkeit $1-p$ auf. Man führt nun so lange unabhängige Versuche durch, bis man erstmals Erfolg beobachtet. Die Wahrscheinlichkeit, dass dies im $(k+1)$-ten Versuch mit $k \in \mathbb{N}_0$ geschieht, wird durch die **geometrische Verteilung** mit Parameter $p \in (0{,}1]$ beschrieben:

$$\mathbb{P}(\{k\}) := p \cdot (1-p)^k\,, \quad k \in \mathbb{N}_0\,.$$

1.4 Die wichtigsten stetigen Wahrscheinlichkeitsverteilungen

1.4.1 Gleichverteilung auf einem Intervall

Es sei $[a,b]$ ein endliches Intervall. Durch

$$p(x) := \begin{cases} \frac{1}{b-a} & : x \in [a,b] \\ 0 & : x \notin [a,b] \end{cases}$$

wird eine Wahrscheinlichkeitsdichte auf $\mathbb{R}$ definiert. Damit berechnet sich die Wahrscheinlichkeit eines Intervalls $[\alpha,\beta]$ durch

$$\mathbb{P}([\alpha,\beta]) = \int_\alpha^\beta p(x)\,\mathrm{d}x = \frac{\text{Länge von }([\alpha,\beta] \cap [a,b])}{b-a}\,. \tag{3}$$

Insbesondere ergibt sich im Fall $[\alpha,\beta] \subseteq [a,b]$ die Formel

$$\mathbb{P}([\alpha,\beta]) = \frac{\beta-\alpha}{b-a}\,,$$

d. h. die Wahrscheinlichkeit des Eintretens von $[\alpha,\beta] \subseteq [a,b]$ hängt nur von seiner Länge, nicht aber von seiner speziellen Lage innerhalb $[a,b]$, ab. Das durch Gleichung 3 erzeugte Wahrscheinlichkeitsmaß heißt **Gleichverteilung** auf dem Intervall $[a,b]$.

1.4.2 Exponentialverteilung

Gegeben sei eine Zahl $\lambda > 0$. Man definiert die **Exponentialverteilung** E_λ mit Parameter $\lambda > 0$ durch ihre Dichte

$$p(x) := \begin{cases} \lambda\,\mathrm{e}^{-\lambda x} & : x > 0 \\ 0 & : x \leq 0 \end{cases}\,.$$

Für ein Intervall $[\alpha,\beta] \subseteq [0,\infty)$ berechnet sich damit die Wahrscheinlichkeit seines Eintretens durch

$$E_\lambda([\alpha,\beta]) = \lambda \int_\alpha^\beta \mathrm{e}^{-\lambda x}\,\mathrm{d}x = \mathrm{e}^{-\lambda\alpha} - \mathrm{e}^{-\lambda\beta}\,.$$

1.4.3 Normalverteilung

Gegeben seien Zahlen $\mu \in \mathbb{R}$ und $\sigma > 0$. Die Funktion

$$p_{\mu,\sigma^2}(x) := \frac{1}{\sqrt{2\pi}\sigma}\, \mathrm{e}^{-(x-\mu)^2/2\sigma^2}\ , \quad x \in \mathbb{R}\ ,$$

erzeugt ein Wahrscheinlichkeitsmaß $\mathcal{N}(\mu, \sigma^2)$, das man **Normalverteilung** mit Mittelwert μ und Varianz σ^2 nennt. Es gilt dann

$$\mathcal{N}(\mu, \sigma^2)([\alpha, \beta]) = \frac{1}{\sqrt{2\pi}\sigma} \int_\alpha^\beta \mathrm{e}^{-(x-\mu)^2/2\sigma^2}\, \mathrm{d}x\ .$$

Im Fall $\mu = 0$ und $\sigma = 1$ erhält man die **Standardnormalverteilung** $\mathcal{N}(0,1)$. Wahrscheinlichkeiten des Eintretens von Intervallen berechnen sich in diesem Fall durch

$$\mathcal{N}(0,1)([\alpha, \beta]) = \frac{1}{\sqrt{2\pi}} \int_\alpha^\beta \mathrm{e}^{-x^2/2}\, \mathrm{d}x\ .$$

1.4.4 Gleichverteilung auf einer Menge im $\mathbb{R}^n$

Es sei $E \subseteq \mathbb{R}^n$ eine beschränkte und abgeschlossene Teilmenge, deren n-dimensionales Volumen $\mathrm{vol}_n(E)$ man berechnen kann. Man definiert die **Gleichverteilung auf E** durch den Ansatz

$$\mathbb{P}(A) = \frac{\mathrm{vol}_n(A \cap E)}{\mathrm{vol}_n(E)}\ .$$

Insbesondere ergibt sich für $A \subseteq E$ die Aussage

$$\mathbb{P}(A) = \frac{\mathrm{vol}_n(A)}{\mathrm{vol}_n(E)}\ ,$$

d. h., wie im eindimensionalen Fall hängt die Wahrscheinlichkeit des Eintretens einer Menge $A \subseteq E$ nur von deren Volumen ab, nicht aber von deren Lage innerhalb E noch von ihrer Gestalt.

1.5 Verteilungsfunktion

1.5.1 Definition

Für ein Wahrscheinlichkeitsmaß $\mathbb{P}$ auf $(\mathbb{R}, \mathcal{B}(\mathbb{R}))$ wird die **Verteilungsfunktion** $F\colon \mathbb{R} \mapsto \mathbb{R}$ durch

$$F(t) := \mathbb{P}((-\infty, t])\ , \quad t \in \mathbb{R}\ , \tag{4}$$

definiert.

Hinweis: Ist $\mathbb{P}$ ein diskretes Wahrscheinlichkeitsmaß auf $(\Omega, \mathcal{P}(\Omega))$ mit $\Omega \subseteq \mathbb{R}$, so modifiziert sich die Definition zu

$$F(t) := \mathbb{P}((-\infty, t] \cap \Omega)\,, \quad t \in \mathbb{R}\,.$$

1.5.2 Eigenschaften der Verteilungsfunktion

Satz 1.2 *Die Verteilungsfunktion F eines Wahrscheinlichkeitsmaßes besitzt folgende Eigenschaften:*

(i) $\lim_{t\to-\infty} F(t) = 0$ *und* $\lim_{t\to\infty} F(t) = 1$,

(ii) die Funktion F ist nichtfallend und

(iii) die Funktion F ist rechtsseitig stetig.

1.5.3 Weitere Eigenschaften von Verteilungsfunktionen

(a) Für jedes halboffene Intervall $(\alpha, \beta]$ gilt

$$\mathbb{P}((\alpha, \beta]) = F(\alpha) - F(\beta)\,.$$

(b) Die Funktion F besitzt in einem Punkt $t_0 \in \mathbb{R}$ genau dann einen Sprung der Höhe $h > 0$ (man hat $F(t_0) - F(t_0 - 0) = h$), wenn $\mathbb{P}(\{t_0\}) = h$ gilt. Insbesondere hat die Verteilungsfunktion eines diskreten Maßes Sprünge in den Punkten, wo die Masse des Maßes konzentriert ist. Dazwischen ist sie konstant.

(c) Ist F Verteilungsfunktion eines stetigen Wahrscheinlichkeitsmaßes $\mathbb{P}$ mit Dichte p, so berechnet sich F aus

$$F(t) = \int\limits_{-\infty}^{t} p(x)\,\mathrm{d}x\,, \quad t \in \mathbb{R}\,.$$

Insbesondere gilt für alle $t \in \mathbb{R}$, in denen p stetig ist, die Gleichung

$$F'(t) = \left(\frac{\mathrm{d}F}{\mathrm{d}t}\right)(t) = p(t)\,.$$

1.6 Bedingte Verteilungen

1.6.1 Definition

Es sei $(\Omega, \mathcal{A}, \mathbb{P})$ ein Wahrscheinlichkeitsraum. Dann wird für $B \in \mathcal{A}$ mit $\mathbb{P}(B) > 0$ die **bedingte Wahrscheinlichkeit** $\mathbb{P}(\,\cdot\,|B)$ (oder die Wahrscheinlichkeit von A unter der Bedingung B) durch

$$\mathbb{P}(A|B) := \frac{\mathbb{P}(A \cap B)}{\mathbb{P}(B)} \quad \text{für} \;\; A \in \mathcal{A} \tag{5}$$

definiert. Sie gibt die Wahrscheinlichkeit dafür an, dass A eintritt, unter der Bedingung, dass B bereits eingetreten ist. Häufig verwendet man Gleichung 5 auch in der Form

$$\mathbb{P}(A \cap B) = \mathbb{P}(B)\,\mathbb{P}(A|B)\;.$$

1.6.2 Eigenschaften

Satz 1.3 *Die Abbildung*

$$A \mapsto \mathbb{P}(A|B)$$

von $\mathcal{A}$ nach $[0{,}1]$ ist ein Wahrscheinlichkeitsmaß mit den zusätzlichen Eigenschaften

$$\mathbb{P}(B|B) = 1 \quad \textit{und} \quad \mathbb{P}(B^c|B) = 0\;.$$

1.6.3 Formel über die totale Wahrscheinlichkeit

Satz 1.4 *Gegeben seien disjunkte Mengen $B_1, \ldots, B_n$ in $\mathcal{A}$ mit $\mathbb{P}(B_j) > 0$. Dann gilt für $A \in \mathcal{A}$ mit $A \subseteq \bigcup_{j=1}^n B_j$ die Aussage*

$$\mathbb{P}(A) = \sum_{j=1}^{n} \mathbb{P}(B_j) \cdot \mathbb{P}(A|B_j)\;.$$

Bemerkung: Insbesondere gilt der Satz im Fall $\bigcup_{j=1}^n B_j = \Omega$ für alle $A \in \mathcal{A}$.

1.6.4 Formel von Bayes

Zur Berechnung von a posteriori Wahrscheinlichkeiten ist die Formel von Bayes wichtig. Sie besagt das folgende:

Satz 1.5 *Unter den Voraussetzungen aus Satz 1.4 an $B_1, \ldots, B_n$ und A folgt für $\mathbb{P}(A) > 0$ die Identität*

$$\mathbb{P}(B_k|A) = \frac{\mathbb{P}(B_k) \cdot \mathbb{P}(A|B_k)}{\sum_{j=1}^n \mathbb{P}(B_j) \cdot \mathbb{P}(A|B_j)} \quad \textit{für} \;\; k = 1, \ldots, n\;. \tag{6}$$

Bemerkung: Den Nenner in Gleichung 6 kann man (falls bekannt) durch $\mathbb{P}(A)$ ersetzen.

1.7 Unabhängigkeit von Ereignissen

1.7.1 Unabhängigkeit von zwei Ereignissen

Gegeben seien zwei Ereignisse A, B aus einem Wahrscheinlichkeitsraum $(\Omega, \mathcal{A}, \mathbb{P})$. Dann heißen A und B (stochastisch) **unabhängig**, wenn

$$\mathbb{P}(A \cap B) = \mathbb{P}(A)\,\mathbb{P}(B)$$

gilt.

1.7.2 Eigenschaften

Die $\emptyset$ und Ω sind von jeder Menge $A \in \mathcal{A}$ unahängig. Sind A und B unahängig, dann gilt dies auch für die Paare A und B^c bzw. A^c und B^c.

1.7.3 Unabhängigkeit von n Ereignissen

Die Ereignisse $A_1, \dots, A_n$ aus $\mathcal{A}$ heißen (stochastisch) **unabhängig**, wenn für alle Teilmengen $I \subseteq \{1, \dots, n\}$ stets

$$\mathbb{P}\Big(\bigcap_{i \in I} A_i\Big) = \prod_{i \in I} \mathbb{P}(A_i) \tag{7}$$

gilt. Man kann dies auch wie folgt formulieren: Für alle $m \geq 2$ und alle $1 \leq i_1 < \cdots < i_m \leq n$ hat man

$$\mathbb{P}(A_{i_1} \cap \cdots \cap A_{i_m}) = \mathbb{P}(A_{i_1}) \cdots \mathbb{P}(A_{i_m}) \;. \tag{8}$$

Die Ereignisse $A_1, \dots, A_n$ aus $\mathcal{A}$ heißen **paarweise unabhängig**, wenn jeweils zwei Ereignisse aus $A_1, \dots, A_n$ unabhängig sind, d. h. Gleichung 7 muss nur für $\operatorname{card}(I) = 2$ bzw. Gleichung 8 nur für $m = 2$ erfüllt sein.

1.7.4 Eigenschaften

Unabhängige Mengen $A_1, \dots, A_n$ sind auch paarweise unabhängig. Die Umkehrung ist i. a. falsch. Ebenso falsch ist, dass aus

$$\mathbb{P}(A_1 \cap \cdots \cap A_n) = \mathbb{P}(A_1) \cdots \mathbb{P}(A_n)$$

stets die Unabhängigkeit der A_j folgt.

Sind $A_1, \dots, A_n$ unabhängig, so gilt dies auch für $(A_j)_{j \in J}$ mit $J \subseteq \{1, \dots, n\}$.

2 Zufallsvariable

2.1 Definition und Verteilungsgesetz

2.1.1 Das vollständige Urbild

Für eine Abbildung $X\colon \Omega \mapsto \mathbb{R}$ und eine Teilmenge $B \subseteq \mathbb{R}$ wird das **vollständige Urbild** von B unter X durch

$$X^{-1}(B) := \{\omega \in \Omega\colon X(\omega) \in B\}$$

definiert. Verkürzend schreibt man auch $X^{-1}(B) = \{X \in B\}$.

2.1.2 Zufällige Größen

Sei Ω eine Menge, die mit einer σ-Algebra $\mathcal{A}$ versehen ist. Eine Abbildung $X\colon \Omega \mapsto \mathbb{R}$ heißt **zufällige Größe** oder **reellwertige Zufallsvariable** oder **zufällige reelle Zahl**, wenn für jedes $t \in \mathbb{R}$ die Menge $\{\omega \in \Omega\colon X(\omega) \leq t\}$ zur σ-Algebra $\mathcal{A}$ gehört.
Bemerkung: In diesem Fall gilt dann auch $X^{-1}(B) \in \mathcal{A}$ für jede Borelmenge $B \subseteq \mathbb{R}$.

2.1.3 Verteilungsgesetz einer zufälligen Größe

Sei $(\Omega, \mathcal{A}, \mathbb{P})$ ein Wahrscheinlichkeitsraum. Für eine zufällige Größe $X\colon \Omega \mapsto \mathbb{R}$ ist die Abbildung $\mathbb{P}_X\colon \mathcal{B}(\mathbb{R}) \mapsto [0,1]$ mit

$$\mathbb{P}_X(B) = \mathbb{P}\big(X^{-1}(B)\big) = \mathbb{P}\{\omega \in \Omega\colon X(\omega) \in B\} = \mathbb{P}(\{X \in B\}) = \mathbb{P}(X \in B)$$

sinnvoll definiert.

Satz 2.1 *Die Abbildung* $\mathbb{P}_X$ *ist ein Wahrscheinlichkeitsmaß auf* $(\mathbb{R}, \mathcal{B}(\mathbb{R}))$*.*

Man nennt $\mathbb{P}_X$ das **Verteilungsgesetz** von X (bzgl. $\mathbb{P}$).

2.1.4 Typen von zufälligen Größen

Eine zufällige Größe X heißt **diskret**, wenn $\mathbb{P}_X$ ein diskretes Wahrscheinlichkeitsmaß ist. Damit hat $\mathbb{P}_X$ die Gestalt

$$\mathbb{P}_X(B) = \sum_{\{i\colon x_i \in B\}} p_i$$

mit geeigneten $x_i \in \mathbb{R}$ und $p_i \geq 0$. Die x_i sind die möglichen Werte von X, d. h. es gilt $\mathbb{P}(X \in \{x_1, x_2, \dots\}) = 1$, und

$$p_i = \mathbb{P}\{\omega \in \Omega \colon X(\omega) = x_i\} \, .$$

Eine zufällige Größe X heißt **stetig**, falls $\mathbb{P}_X$ ein stetiges Wahrscheinlichkeitsmaß ist. Das gilt genau dann, wenn mit einer Wahrscheinlichkeitsdichte p für alle $\alpha < \beta$ die Gleichung

$$\mathbb{P}_X([\alpha, \beta]) = \mathbb{P}\{\omega \in \Omega \colon \alpha \leq X(\omega) \leq \beta\} = \int_\alpha^\beta p(x) \, \mathrm{d}x$$

erfüllt ist. Die Funktion p nennt man auch **Verteilungsdichte** (oder einfach **Dichte**) von X.

2.1.5 Speziell verteilte diskrete zufällige Größen

Eine zufällige Größe X heißt **gleichverteilt** auf einer endlichen Menge oder **binomialverteilt** oder **Poissonverteilt** etc., wenn $\mathbb{P}_X$ von diesem Typ ist. In allen diesen Fällen ist X diskret. Zum Beispiel ist X gemäß $B_{n,p}$ verteilt (man schreibt auch $X \sim B_{n,p}$), falls für $0 \leq k \leq n$ stets

$$\mathbb{P}_X(\{k\}) = \mathbb{P}\{\omega \in \Omega \colon X(\omega) = k\} = \binom{n}{k} p^k (1-p)^{n-k}$$

gilt. Analog ist X gemäß P_λ verteilt, sofern für $k \in \mathbb{N}_0$

$$\mathbb{P}_X(\{k\}) = \mathbb{P}\{\omega \in \Omega \colon X(\omega) = k\} = \frac{\lambda^k}{k!} \mathrm{e}^{-\lambda} \, .$$

2.1.6 Speziell verteilte stetige zufällige Größen

Eine zufällige Größe X heißt **gleichverteilt** auf einem Intervall, oder **exponentialverteilt** oder **normalverteilt** etc., wenn $\mathbb{P}_X$ von diesem Typ ist. Alle diese zufälligen Größen sind stetig. Zum Beispiel ist X gleichverteilt auf $[a, b]$, falls für alle $\alpha < \beta$ stets

$$\mathbb{P}_X([\alpha, \beta]) = \mathbb{P}\{\omega \in \Omega \colon \alpha \leq X(\omega) \leq \beta\} = \frac{\text{Länge von}\,([\alpha, \beta] \cap [a, b])}{b - a}$$

gilt. Oder X ist $\mathcal{N}(\mu, \sigma^2)$-verteilt (man schreibt $X \sim \mathcal{N}(\mu, \sigma^2)$), sofern

$$\mathbb{P}_X([\alpha, \beta]) = \mathbb{P}\{\omega \in \Omega \colon \alpha \leq X(\omega) \leq \beta\} = \frac{1}{\sqrt{2\pi}\sigma} \int_\alpha^\beta \mathrm{e}^{-(x-\mu)^2/2\sigma^2} \, \mathrm{d}x \, .$$

2.1.7 Identisch verteilte zufällige Größen

Zwei zufällige Größen X und Y sind **identisch verteilt**, wenn $\mathbb{P}_X = \mathbb{P}_Y$ gilt, d. h. für alle $B \in \mathcal{B}(\mathbb{R})$ hat man

$$\mathbb{P}\{\omega \in \Omega\colon X(\omega) \in B\} = \mathbb{P}\{\omega \in \Omega\colon Y(\omega) \in B\} .$$

Man schreibt dann $X \overset{d}{=} Y$.

2.1.8 Verteilungsfunktion einer zufälligen Größe

Die **Verteilungsfunktion** F_X einer zufälligen Größe ist die Verteilungsfunktion ihres Verteilungsgesetzes, d. h., es gilt

$$F_X(t) = \mathbb{P}_X\big((-\infty, t]\big) = \mathbb{P}\{\omega \in \Omega\colon X(\omega) \le t\} , \quad t \in \mathbb{R} .$$

Für zwei zufällige Größen X und Y gilt genau dann $X \overset{d}{=} Y$, wenn man $F_X = F_Y$ hat.

Die Funktion F_X besitzt die Eigenschaften aus Satz 1.2.

2.2 Zufällige Vektoren und Unabhängigkeit zufälliger Größen

2.2.1 Zufällige Vektoren

Sei Ω eine Menge mit einer σ-Algebra $\mathcal{A}$. Eine Abbildung $\vec{X}\colon \Omega \mapsto \mathbb{R}^n$ heißt (n-dimensionaler) **zufälliger Vektor**, wenn seine Koordinatenabbildungen $X_j\colon \Omega \mapsto \mathbb{R}$ alle zufällige Größen sind. Dabei sind wie üblich die X_j durch

$$\vec{X}(\omega) = (X_1(\omega), \ldots, X_n(\omega)) , \quad \omega \in \Omega ,$$

definiert.

2.2.2 Gemeinsames Verteilungsgesetz

Sei $(\Omega, \mathcal{A}, \mathbb{P})$ ein Wahrscheinlichkeitsraum. Dann definiert man wie im eindimensionalen Fall das Verteilungsgesetz $\mathbb{P}_{\vec{X}}$ von $\vec{X}$ durch

$$\mathbb{P}_{\vec{X}}(B) := \mathbb{P}\big(\vec{X}^{-1}(B)\big) = \mathbb{P}\{\omega \in \Omega\colon (X_1(\omega), \ldots, X_n(\omega)) \in B\} .$$

Im Spezialfall $B = B_1 \times \cdots \times B_n$ für Borelmengen $B_j \subseteq \mathbb{R}$ folgt

$$\mathbb{P}_{\vec{X}}(B) = \mathbb{P}(X_1 \in B_1, \ldots, X_n \in B_n) .$$

Deshalb nennt man $\mathbb{P}_{\vec{X}}$ auch **gemeinsames Verteilungsgesetz** der zufälligen Größen $X_1, \ldots, X_n$.

2.2.3 Randverteilungen

Für einen zufälligen Vektor $\vec{X}$ nennt man die Verteilungsgesetze $\mathbb{P}_{X_j}$, $1 \leq j \leq n$, die **Randverteilungen** von $\vec{X}$. Hierbei sind wie zuvor die zufälligen Größen X_j die zugehörigen Koordinatenabbildungen.

Satz 2.2 *Die Randverteilungen berechnen sich aus der gemeinsamen Verteilung durch*

$$\mathbb{P}_{X_j}(B) = \mathbb{P}_{\vec{X}}(\mathbb{R} \times \cdots \times \underset{\uparrow\ j\text{-te Stelle}}{B} \times \cdots \times \mathbb{R}) \,, \quad B \in \mathcal{B}(\mathbb{R}) \,.$$

Damit bestimmt die gemeinsame Verteilung die zugehörigen Randverteilungen.

Bemerkung: Die Umkehrung der obigen Aussage ist i. a. falsch, d. h. es existieren zufällige Vektoren $\vec{X} = (X_1, \ldots, X_n)$ und $\vec{Y} = (Y_1, \ldots, Y_n)$ mit $\mathbb{P}_{X_j} = \mathbb{P}_{Y_j}$, $1 \leq j \leq n$, aber mit $\mathbb{P}_{\vec{X}} \neq \mathbb{P}_{\vec{Y}}$.

2.2.4 Randverteilungen diskreter Vektoren

Wir betrachten hier nur den Fall $n = 2$. Ein zufälliger 2-dimensionaler Vektor hat die Gestalt (X, Y) mit vorgegebenen zufälligen Größen X und Y. Weiterhin seien X und Y diskret und die Folgen $(x_i)_{i\geq 1}$ bzw. $(y_j)_{j\geq 1}$ von reellen Zahlen bezeichnen die möglichen Werte von X bzw. Y. Dann nimmt der Vektor (X, Y) die Werte $(x_i, y_j)_{i,j\geq 1}$ an und für das Verteilungsgesetz von $\mathbb{P}_{(X,Y)}$, d. h. die gemeinsame Verteilung von X und Y, gilt

$$\mathbb{P}_{(X,Y)}(B) = \sum_{\{(i,j)\colon\ (x_i,y_j)\in B\}} p_{ij} \,, \quad B \in \mathcal{P}(\mathbb{R}^2) \,,$$

wobei

$$p_{ij} = \mathbb{P}_{(X,Y)}(\{(x_i, y_j)\}) = \mathbb{P}(X = x_i, Y = y_j) \,.$$

Für die Randverteilungen ergibt sich dann

$$\mathbb{P}_X(B) = \sum_{\{i\colon\ x_i\in B\}} q_i \qquad \text{und} \qquad \mathbb{P}_Y(B) = \sum_{\{j\colon\ y_j\in B\}} r_j \,, \quad B \in \mathcal{P}(\mathbb{R}) \,,$$

mit

$$q_i = \sum_{j=1}^{\infty} p_{ij} \qquad \text{und} \qquad r_j = \sum_{i=1}^{\infty} p_{ij} \quad .$$

2.2.5 Randverteilungen stetiger Vektoren

Zur besseren Übersichtlichkeit betrachten wir auch hier nur den Fall $n = 2$. Der 2-dimensionale Vektor (X, Y) sei wie oben definiert. Diesmal nehmen wir aber an, dass $\mathbb{P}_{(X,Y)}$ eine Dichte hat, es also eine Funktion $p\colon \mathbb{R}^2 \mapsto \mathbb{R}$ gibt, so dass für alle $\alpha < \beta$ und $\gamma < \delta$ stets

$$\mathbb{P}_{(X,Y)}\big([\alpha, \beta] \times [\gamma, \delta]\big) = \mathbb{P}\{\omega \in \Omega\colon \alpha \leq X(\omega) \leq \beta,\, \gamma \leq Y(\omega) \leq \delta\} = \int\limits_{\alpha}^{\beta} \int\limits_{\gamma}^{\delta} p(x, y)\,\mathrm{d}y\,\mathrm{d}x$$

gilt. Dann haben X bzw. Y Verteilungsdichten q und r mit

$$q(x) := \int\limits_{-\infty}^{\infty} p(x, y)\,\mathrm{d}y \qquad \text{und} \qquad r(y) := \int\limits_{-\infty}^{\infty} p(x, y)\,\mathrm{d}x\ .$$

2.2.6 Unabhängigkeit von zufälligen Größen

Gegeben seien n zufällige Größen $X_1, \ldots, X_n$ auf $(\Omega, \mathcal{A}, \mathbb{P})$. Gilt für beliebige Borelmengen $B_1, \ldots, B_n \in \mathcal{B}(\mathbb{R})$ stets

$$\mathbb{P}(X_1 \in B_1, \ldots, X_n \in B_n) = \mathbb{P}(X_1 \in B_1) \cdots \mathbb{P}(X_n \in B_n)\ , \tag{9}$$

so heißen $X_1, \ldots, X_n$ **unabhängig**.

Bemerkung 1: Die Unabhängigkeit der X_j ist äquivalent zu folgender Aussage: Für beliebige Borelmengen $B_j \in \mathcal{B}(\mathbb{R})$ sind die Ereignisse $\big(X_j^{-1}(B_j)\big)_{j=1}^{n}$ unabhängig. Das folgt aus der Tatsache, dass man in Gleichung 9 für gewisse vorgegebene B_j auch die reellen Zahlen $\mathbb{R}$ einsetzen kann.

Bemerkung 2: Es reicht aus, wenn Gleichung 9 mit Intervallen B_j der Form $(-\infty, t_j]$ für alle $t_j \in \mathbb{R}$ gilt. Die zufälligen Größen $X_1, \ldots, X_n$ sind also dann und nur dann unabhängig, wenn für alle $t_j \in \mathbb{R}$ stets

$$\mathbb{P}(X_1 \leq t_1, \ldots, X_n \leq t_n) = \mathbb{P}(X_1 \leq t_1) \cdots \mathbb{P}(X_n \leq t_n)$$

folgt.

Bemerkung 3: Aufgrund von Gleichung 9 ist die gemeinsame Verteilung von $X_1, \ldots, X_n$ im Fall der Unabhängigkeit eindeutig durch ihre Randverteilungen P_{X_j}, $1 \leq j \leq n$, bestimmt.

2.2.7 Spezialfälle

Besitzen X und Y die Eigenschaften aus Abschnitt 2.2.4, so sind X und Y dann und nur dann unabhängig, wenn

$$p_{ij} = q_i \cdot r_j \,, \quad 1 \le i,j < \infty \,.$$

Im stetigen Fall (Abschnitt 2.2.5) sind X und Y genau dann unabhängig, wenn

$$p(x,y) = q(x) \cdot r(y) \,, \quad x,y \in \mathbb{R} \,.$$

2.3 Rechnen mit zufälligen Größen

2.3.1 Transformationen

Eine Abbildung $f\colon \mathbb{R} \mapsto \mathbb{R}$ heißt **messbar**, wenn für jedes $t \in \mathbb{R}$ die Menge $\{x \in \mathbb{R}\colon f(x) \le t\}$ eine Borelmenge ist. Stetige Funktionen, Grenzwerte stetiger Funktionen oder auch monotone Funktionen besitzen diese Eigenschaft.

Satz 2.3 *Sei X eine zufällige Größe und sei $f\colon \mathbb{R} \mapsto \mathbb{R}$ messbar. Dann ist $Y := f(X)$ ebenfalls eine zufällige Größe.*

Allgemeine Aufgabe: Man bestimme $\mathbb{P}_Y$ mit Hilfe von $\mathbb{P}_X$ und f. Folgendes Beispiel illustriere die Situation: Sei U gleichverteilt auf [0,1], so ist mit $f(s) := 1 - s$ auch $Y := f(U) = 1 - U$ gleichverteilt auf [0,1].

2.3.2 Simulation stetiger zufälliger Größen

Sei X eine stetige zufällige Größe mit Verteilungsfunktion F_X. Wir nehmen an, dass mit zwei Zahlen $-\infty \le a < b \le \infty$ die Verteilungsfunktion $F_X(a) = 0$, $F_X(b) = 1$ erfülle und auf (a,b) streng wachsend sei. Dann existiert die inverse Funktion von F_X, die mit F_X^{-1} bezeichnet wird, und es gilt $F_X^{-1}\colon (0,1) \mapsto (a,b)$.

Satz 2.4 *Sei U eine auf [0,1] gleichverteilte zufällige Größe. Unter den obigen Voraussetzungen gilt dann für $Y := F_X^{-1}(U)$ die Aussage $X \stackrel{d}{=} Y$.*

Anwendung: Sind $u_1, \ldots, u_n$ unabhängig erzeugte reelle Zahlen, die gemäß der Gleichverteilung aus [0,1] gewählt wurden, so sind die Zahlen $x_j := F_X^{-1}(u_j)$ ebenfalls unabhängig und gemäß $\mathbb{P}_X$ verteilt.

2.3.3 Lineare Transformationen

Für reelle Zahlen $a \neq 0$ und $b \in \mathbb{R}$ betrachte man die lineare Transformation

$$Y := a\,X + b$$

einer zufälligen Größe X.

Satz 2.5 *Im Fall $a > 0$ folgt*

$$F_Y(t) = F_X\left(\frac{t-b}{a}\right) .$$

Ist $a < 0$, so ergibt sich

$$F_Y(t) = 1 - \mathbb{P}\left(X < \frac{t-b}{a}\right) ,$$

also

$$F_Y(t) = 1 - F_X\left(\frac{t-b}{a}\right)$$

im Fall stetiger X.

Folgerung: Besitzt X die Verteilungsdichte p, so hat $Y = a\,X + b$ eine Dichte q, die sich aus p durch

$$q(t) = \frac{1}{|a|}\, p\left(\frac{t-b}{a}\right) , \quad t \in \mathbb{R} ,$$

ergibt.

2.3.4 Addition zufälliger Größen

Für zwei zufällige Größen X und Y wird ihre Summe $X + Y$ durch

$$(X + Y)(\omega) := X(\omega) + Y(\omega) , \qquad \omega \in \Omega ,$$

definiert.

Satz 2.6 *Sind X und Y zufällige Größen, so gilt dies auch für $X + Y$.*

Das Verteilungsgesetz der Summe $X + Y$ kann man für unabhängige zufällige Größen in einigen Fällen in einfacher Form angeben.

Satz 2.7 *Es seien X und Y unabhängige zufällige Größen.*

1. *Nehmen X und Y Werte in den ganzen Zahlen $\mathbb{Z}$ an, so folgt*

$$\mathbb{P}(X+Y=k) = \sum_{i=-\infty}^{\infty} \mathbb{P}(X=i)\,\mathbb{P}(Y=k-i)\;, \quad k \in \mathbb{Z}\;.$$

2. *Besitzen X und Y Werte in $\mathbb{N}_0$, so ergibt sich*

$$\mathbb{P}(X+Y=k) = \sum_{i=0}^{k} \mathbb{P}(X=i)\,\mathbb{P}(Y=k-i)\;, \quad k \in \mathbb{N}_0\;.$$

Im Fall stetiger zufälliger Größen gilt folgender Satz:

Satz 2.8 *Seien X und Y unabhängig mit Verteilungsdichten p und q. Dann besitzt $X+Y$ die Verteilungsdichte r mit*

$$r(x) = \int_{-\infty}^{\infty} p(x-y)\,q(y)\,\mathrm{d}y = \int_{-\infty}^{\infty} p(y)\,q(x-y)\,\mathrm{d}y\;.$$

Man nennt r die **Faltung** von p und q und schreibt $r = p * q$.

2.3.5 Addition speziell verteilter zufälliger Größen

Satz 2.9 *Im folgenden seien X und Y stets als unabhängig vorausgesetzt. Dann gilt:*

(a) *Aus $X \sim B_{n,p}$ und $Y \sim B_{m,p}$ folgt $X+Y \sim B_{n+m,p}$.*

(b) *Aus $X \sim P_\lambda$ und $Y \sim P_\mu$ erhält man $X+Y \sim P_{\lambda+\mu}$.*

(c) *Aus $X \sim \mathcal{N}(\mu_1, \sigma_1^2)$ und $Y \sim \mathcal{N}(\mu_2, \sigma_2^2)$ folgt $X+Y \sim \mathcal{N}(\mu_1+\mu_2, \sigma_1^2+\sigma_2^2)$.*

2.4 Erwartungswert

2.4.1 Erwartungswert diskreter zufälliger Größen

Eine zufällige Größe X nehme Werte $x_1, x_2, \ldots$ aus $[0,\infty)$ an. Dann definiert man den **Erwartungswert** von X durch

$$\mathbb{E}X := \sum_{i=1}^{\infty} x_i\,\mathbb{P}(X=x_i)\;.$$

Es gilt dann $0 \le \mathbb{E}X \le \infty$.

Sind nunmehr die Werte von X beliebige reelle Zahlen (nicht notwendig ≥ 0), so sagt man, dass X einen **Erwartungswert besitzt**, wenn

$$\sum_{i=1}^{\infty} |x_i| \, \mathbb{P}(X = x_i) < \infty \, .$$

In diesem Fall ist der Erwartungswert von X mit

$$\mathbb{E}X := \sum_{i=1}^{\infty} x_i \, \mathbb{P}(X = x_i)$$

eine wohldefinierte reelle Zahl.

2.4.2 Erwartungswert stetiger zufälliger Größen

Sei p die Verteilungsdichte einer zufälligen Größe X. Dann **besitzt *X* einen Erwartungswert**, wenn

$$\int_{-\infty}^{\infty} |x| \, p(x) \, \mathrm{d}x < \infty \, ,$$

und man definiert den **Erwartungswert** von X durch

$$\mathbb{E}X := \int_{-\infty}^{\infty} x \, p(x) \, \mathrm{d}x \, .$$

2.4.3 Beispiele zur Berechnung von Erwartungswerten

Verteilung	**Erwartungswert**
X gleichverteilt auf $x_1, \ldots, x_N$	$\mathbb{E}X = \frac{1}{N} \sum_{i=1}^{N} x_i$
$X \sim B_{n,p}$	$\mathbb{E}X = np$
$X \sim P_\lambda$	$\mathbb{E}X = \lambda$
X geometrisch verteilt mit Parameter p	$\mathbb{E}X = \frac{1-p}{p}$
X gleichverteilt auf $[a, b]$	$\mathbb{E}X = \frac{a+b}{2}$
$X \sim E_\lambda$	$\mathbb{E}X = \frac{1}{\lambda}$
$X \sim \mathcal{N}(\mu, \sigma^2)$	$\mathbb{E}X = \mu$

2.4.4 Eigenschaften des Erwartungswertes

Satz 2.10 *Der Erwartungswert einer zufälligen Größe hat folgende Eigenschaften:*

1. *Der Erwartungswert ist linear, d. h. für alle $a, b \in \mathbb{R}$ und zufällige Größen X und Y gilt*

$$\mathbb{E}(aX + bY) = a\,\mathbb{E}X + b\,\mathbb{E}Y\ .$$

2. *Sei X diskret mit möglichen Werten $x_1, x_2, \ldots$ aus $\mathbb{R}$. Dann existiert für eine Funktion $f\colon \mathbb{R} \mapsto \mathbb{R}$ der Erwartungswert $\mathbb{E}f(X)$ genau dann, wenn*

$$\sum_{i=1}^{\infty} |f(x_i)|\,\mathbb{P}(X = x_i) < \infty\ ,$$

und es gilt

$$\mathbb{E}f(X) = \sum_{i=1}^{\infty} f(x_i)\,\mathbb{P}(X = x_i)\ .$$

3. *Ist X stetig mit Verteilungsdichte p, so existiert für eine messbare Abbildung $f\colon \mathbb{R} \mapsto \mathbb{R}$ genau dann der Erwartungswert von $f(X)$, wenn*

$$\int_{-\infty}^{\infty} |f(x)|\,p(x)\,\mathrm{d}x < \infty\ ,$$

und man hat

$$\mathbb{E}f(X) = \int_{-\infty}^{\infty} f(x)\,p(x)\,\mathrm{d}x\ .$$

4. *Sind X und Y unabhängige zufällige Größen deren Erwartungswert existiert, so existiert auch der Erwartungswert von $X \cdot Y$, und es gilt*

$$\mathbb{E}(X \cdot Y) = \mathbb{E}X \cdot \mathbb{E}Y\ .$$

2.5 Varianz und Kovarianz

2.5.1 Momente

Sei $n \in \mathbb{N}$. Eine zufällige Größe X besitzt ein ***n*-tes Moment**, wenn $\mathbb{E}|X|^n < \infty$. Im diskreten Fall bedeutet dies

$$\sum_{i=1}^{\infty} |x_i|^n \mathbb{P}(X = x_i) < \infty$$

und im stetigen

$$\int_{-\infty}^{\infty} |x|^n \, p(x) \, \mathrm{d}x < \infty \, .$$

Insbesondere hat X ein erstes Moment, genau dann, wenn $\mathbb{E}X$ existiert.

Satz 2.11 *Sei $1 \leq m \leq n$. Hat eine zufällige Größe X ein n-tes Moment, so besitzt sie auch ein m-tes Moment. Insbesondere hat jede zufällige Größe mit zweitem Moment einen Erwartungswert.*

2.5.2 Varianz

Es sei X eine zufällige Größe mit zweitem Moment. Sei $a := \mathbb{E}X$. Dann definiert man die **Varianz** (oder **Streuung**) von X durch

$$\mathbb{V}X := \mathbb{E}(X-a)^2 \, .$$

Die Varianz gibt den mittleren quadratischen Abstand einer zufälligen Größe X von ihrem Erwartungswert an. Sie ist ein Maß dafür, wie sehr die Werte von X um $\mathbb{E}X$ schwanken.

2.5.3 Eigenschaften der Varianz

Satz 2.12 *Im folgenden seien X und Y zufällige Größen mit zweiten Momenten. Dann gelten die folgenden Aussagen:*

1. *Mit $a := \mathbb{E}X$ berechnet sich die Varianz für diskrete zufällige Größen in der Form*

$$\mathbb{V}X = \sum_{i=1}^{\infty} (x_i - a)^2 \, \mathbb{P}(X = x_i) \, ,$$

   und im stetigen Fall hat man

$$\mathbb{V}X = \int_{-\infty}^{\infty} (x-a)^2 \, p(x) \, \mathrm{d}x \, .$$

2. *Es besteht die Identität*

$$\mathbb{V}X = \mathbb{E}X^2 - (\mathbb{E}X)^2 \, .$$

3. *Für eine konstante zufällige Größe X folgt $\mathbb{V}X = 0$.*

4. *Für $\alpha \in \mathbb{R}$ erhält man*

$$\mathbb{V}(\alpha\, X) = \alpha^2\, \mathbb{V}X\ .$$

5. *Sind X und Y unabhängig, dann gilt*

$$\mathbb{V}(X+Y) = \mathbb{V}X + \mathbb{V}Y\ .$$

2.5.4 Beispiele zur Berechnung von Varianzen

Verteilung	Varianz
X gleichverteilt auf $x_1, \ldots, x_N$	$\mathbb{V}X = \frac{1}{N}\sum_{i=1}^{N}(x_i - \mathbb{E}X)^2$
$X \sim B_{n,p}$	$\mathbb{V}X = n\,p\,(1-p)$
$X \sim P_\lambda$	$\mathbb{V}X = \lambda$
X geometrisch verteilt mit Parameter p	$\mathbb{V}X = \frac{1-p}{p^2}$
X gleichverteilt auf $[a,b]$	$\mathbb{V}X = \frac{(b-a)^2}{12}$
$X \sim E_\lambda$	$\mathbb{V}X = \frac{1}{\lambda^2}$
$X \sim \mathcal{N}(\mu, \sigma^2)$	$\mathbb{V}X = \sigma^2$

2.5.5 Kovarianz

Gegeben seien zwei zufällige Größen X und Y mit zweiten Momenten. Seien $a := \mathbb{E}X$ und $b := \mathbb{E}Y$. Dann wird die **Kovarianz** von X und Y durch

$$\mathrm{cov}(X,Y) := \mathbb{E}(X-a)(Y-b)$$

definiert.

Eigenschaften:

1. Sind X und Y diskret mit möglichen Werten $x_1, x_2, \ldots$ bzw. $y_1, y_2, \ldots$ aus $\mathbb{R}$, so berechnet sich die Kovarianz aus der Formel

$$\mathrm{cov}(X,Y) = \sum_{i,j=1}^{\infty}(x_i - a)(y_j - b)\, p_{ij}$$

wobei

$$p_{ij} = \mathbb{P}(X = x_i, Y = y_j)\ .$$

2. Hat die Verteilung des zufälligen Vektors (X, Y) eine Dichte $p \colon \mathbb{R}^2 \mapsto \mathbb{R}$, so ergibt sich die Kovarianz von X und Y aus

$$\operatorname{cov}(X, Y) = \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} (x - a)(y - b)\, p(x, y)\, \mathrm{d}x\, \mathrm{d}y \; .$$

3. Sind X und Y unabhängig, so impliziert dies $\operatorname{cov}(X, Y) = 0$, d. h. X und Y sind **unkorreliert**. Man beachte, dass aus der Unkorreliertheit i. a. nicht die Unabhängigkeit folgt.

4. Man hat

$$|\operatorname{cov}(X, Y)| \leq (\mathbb{V}X)^{1/2} (\mathbb{V}Y)^{1/2} \; . \tag{10}$$

2.5.6 Korrelationskoeffizient

Für zwei zufällige Größen X und Y mit zweiten Momenten definiert man ihren **Korrelationskoeffizienten** durch

$$\rho(X, Y) := \frac{\operatorname{cov}(X, Y)}{(\mathbb{V}X)^{1/2} (\mathbb{V}Y)^{1/2}} \; .$$

Aus Gleichung 10 folgt

$$-1 \leq \rho(X, Y) \leq 1 \; .$$

Für unkorrelierte zufällige Größen gilt $\rho(X, Y) = 0$. Der Korrelationskoeffizient ist ein Maß für den Grad der Abhängigkeit von X und Y. Je näher $\rho(X, Y)$ an 1 oder -1 liegt, desto größer ist die Abhängigkeit zwischen X und Y. Im stärksten Fall der Abhängigkeit von X und Y, nämlich $Y = X$ bzw. $Y = -X$, hat man $\rho(X, X) = 1$ bzw. $\rho(X, -X) = -1$.

Nutzungsbedingungen

Dieses Dokument wurde für die auf der Titelseite genannte Vorlesung erstellt und wird jetzt im Rahmen des Projekts »Vorlesungsskripte der Fakultät für Mathematik und Informatik« weiter betreut. Das Dokument wurde nach bestem Wissen und Gewissen angefertigt. Dennoch garantiert weder der auf der Titelseite genannte Dozent, die Personen, die an dem Dokument mitgewirkt haben, noch die Mitglieder des Projekts für dessen Fehlerfreiheit. Für etwaige Fehler und dessen Folgen wird von keiner der genannten Personen eine Haftung übernommen. Es steht jeder Person frei, dieses Dokument zu lesen, zu verändern oder auf anderen Medien verfügbar zu machen, solange ein Verweis auf die Internetadresse des Projekts `http://uni-skripte.lug-jena.de/` *enthalten ist.*

Diese Ausgabe trägt die Versionsnummer 1889 und ist vom 15. November 2008. Eine neue Ausgabe könnte auf der Webseite des Projekts verfügbar sein.

Jeder ist dazu aufgerufen, Verbesserungen, Erweiterungen und Fehlerkorrekturen für das Skript einzureichen bzw. zu melden oder diese selbst einzupflegen – einfach eine E-Mail an die Mailingliste `<uni-skripte@lug-jena.de>` *senden. Weitere Informationen sind unter der oben genannten Internetadresse verfügbar.*

Hiermit möchten wir allen Personen, die an diesem Skript mitgewirkt haben, vielmals danken:

- *Prof. Werner Linde (2008/09)*

Index

W

Z