Kapitel 5

Bedingte Wahrscheinlichkeiten und Unabhängigkeit

Mitunter erhält man über das Ergebnis eines zufälligen Versuches Vorinformationen. Dann entsteht die Frage, wie sich für den Betrachter, den man als "Insider" bezeichnen könnte, die Wahrscheinlichkeiten der mit dem Versuch verbundenen Ereignisse ändern.

5.1 Definition und Eigenschaften

Es sei $(\Omega, \mathfrak{A}, P)$ ein Wahrscheinlichkeitsraum. Wir verbinden mit $(\Omega, \mathfrak{A}, P)$ wieder die Vorstellung, es handle sich um einen zufälligen Versuch mit möglichen Versuchsergebnissen ω aus Ω, mit den zufälligen Ereignissen A aus $\mathfrak{A}$, die mit dem Versuch verbunden sind, und mit der Wahrscheinlichkeitsverteilung P auf $\mathfrak{A}$, die jedem Ereignis A aus $\mathfrak{A}$ eine Wahrscheinlichkeit $P(A)$ zuordnet.
Es seien A und B zwei Ereignisse, d.h. $A, B \in \mathfrak{A}$.

Angenommen, vor Bekanntwerden des Ausganges dieses zufälligen Versuches erhält man als Vorinformation die Nachricht, dass das Ereignis B eingetreten ist. Wie ändert sich dadurch die Wahrscheinlichkeit $P(A)$ für das Eintreten von A? Um davon eine Vorstellung zu bekommen, nehmen wir wieder das empirische Gesetz der großen Zahlen zu Hilfe.

Wir stellen uns vor, der Versuch wird n-mal unter gleichartigen Bedingungen und unabhängig voneinander ausgeführt. Dann gilt entsprechend diesem Erfahrungsgesetz

$$\frac{n(A)}{n} \approx P(A), \frac{n(B)}{n} \approx P(B).$$

Wir betrachten von dieser Versuchsreihe der Länge n jetzt nur diejenigen Versuche, bei denen B eingetreten ist. Davon gibt es $n(B)$ an der Zahl. Wir können diese Reihe als eine neue Reihe von Versuchen ansehen, bei denen immer B eintritt, das Eintreten von B erscheint damit als eine zusätzliche Versuchsbedingung.
Will man die Wahrscheinlichkeit von A im Rahmen dieser neuen Versuche (unter der Bedingung, dass B eintritt) berechnen, so wäre ein Kandidat die Anzahl aller Versuche, bei denen A und B eintreten, geteilt durch die Anzahl aller Versuche, bei denen B eintritt:

$$\frac{n(A \cap B)}{n(B)} = \frac{n(A \cap B)}{n} : \frac{n(B)}{n}.$$

Das führt uns auf folgende Definition:

Definition 5.1 *Es seien* $(\Omega, \mathfrak{A}, P)$ *ein Wahrscheinlichkeitsraum,* $A, B \in \mathfrak{A}$ *mit* $P(B) > 0$*. Dann heißt*

$$P(A|B) := \frac{P(A \cap B)}{P(B)} \tag{5.1}$$

die bedingte Wahrscheinlichkeit von A unter der Bedingung, dass das Ereignis B eintritt *(kurz: die bedingte Wahrscheinlichkeit von A unter der Bedingung B).*

Beispiel 5.2

a) Beim Werfen zweier Würfel betrachten wir die Ereignisse

A: = "Der zweite Würfel zeigt eine gerade Zahl" und

B : = " Die Augensumme ist acht". Dann gilt

$$P(A|B) = \frac{P(A \cap B)}{P(B)} = \frac{P(\{(6,2),(4,4),(2,6)\})}{P(\{(2,6),(3,5),(4,4),(5,3),(6,2)\})} = \frac{3}{5} \neq P(A) = \frac{1}{2}.$$

b) Bei 6800 rein zufällig aus der Bevölkerung ausgewählten Personen wurden Augen- und Haarfarbe registriert. Das Ergebnis ist in folgender Tabelle enthalten.

k (Augen) \ l (Haar)	1 (hellblond)	2 (dunkelblond)	3 (schwarz)	4 (rot)	Zusammen $n_{k.}$
1 (blau)	1768	807	189	47	2811
2 (grau oder grün)	946	1387	746	53	3132
3 (braun)	115	438	288	16	857
Zusammen $n_{.l}$	2829	2632	1223	116	6800

(aus: Fisz, M., Wahrscheinlichkeitsrechnung und Mathematische Statistik, Deutscher Verlag der Wissenschaften, Berlin 1965.)

Die Tabelle erlaubt die Berechnung bedingter Wahrscheinlichkeiten, die die Zusammenhänge zwischen beiden Merkmalen Haar- und Augenfarbe deutlicher herausstellen als die Originaltabelle. Zum Beispiel gilt:

A_3 := "Eine rein zufällig ausgewählte Person hat braune Augen"

B_3 := "Eine rein zufällig ausgewählte Person hat schwarze Haare"

$$P(A_3|B_3) = \frac{P(A_3 \cap B_3)}{P(B_3)} = \frac{\frac{288}{6800}}{\frac{1223}{6800}} = \frac{288}{1223} = 0,234,$$

dagegen gilt die "absolute" Wahrscheinlichkeit

$$P(A_3) = 857/6800 = 0,126.$$

Entsprechend berechne man z.B. $P(A_2|B_3), P(B_3|A_3)$ usw..

Es seien $A, B, A_k \in \mathfrak{A}, k \geq 1,$ mit $P(B) > 0$. Dann bestehen folgende Eigenschaften, die sich unmittelbar aus der Definition ergeben:

Aussage 5.3 *(Eigenschaften bedingter Wahrscheinlichkeiten)*

1. $0 \leq P(A|B) \leq 1$
2. $P(\Omega|B) = 1, \quad P(\emptyset|B) = 0$
3. $P\big(\bigcup_{k=1}^{\infty} A_k|B) = \sum_{k=1}^{\infty} P(A_k|B\big),$

   falls $A_k \cap A_l = \emptyset$ *für alle* k, l *mit* $k \neq l$.

Bemerkung 5.4 Aus 1. - 3. folgt, dass die Abbildung

$$A \to P(A|B), \quad A \in \mathfrak{A},$$

ebenfalls eine Wahrscheinlichkeitsverteilung auf $\mathfrak{A}$ ist. Der Wahrscheinlichkeitsraum $(\Omega, \mathfrak{A}, P(\cdot|B))$ ist das mathematische Modell für den ursprünglichen Versuch mit der zusätzlichen Bedingung, dass B eintritt.
Unmittelbar aus der Definition der bedingten Wahrscheinlichkeit ergibt sich weiterhin

Aussage 5.5 *(Multiplikationssatz)*

$$P(A \cap B) = P(A|B)P(B) = P(B|A)P(A), \tag{5.2}$$

und allgemeiner:

$$P(A_1 \cap \ldots \cap A_n) =$$

$$P(A_n|A_1 \cap \ldots \cap A_{n-1})P(A_{n-1}|A_1 \cap \ldots \cap A_{n-2}) \ldots P(A_2|A_1) \cdot P(A_1) \tag{5.3}$$

falls $P(A_1 \cap \ldots \cap A_{n-1}) > 0$.

Der Beweis von (5.3) erfolgt mittels vollständiger Induktion unter Verwendung von (5.2).

Aussage 5.6 *("Satz von der totalen Wahrscheinlichkeit")*
Es sei $(Z_i, i \in I)$ mit $I \subseteq N_0$ eine Zerlegung von Ω in Ereignisse Z_i aus $\mathfrak{A}$, die alle eine positive Wahrscheinlichkeit besitzen:

$$Z_i \cap Z_j = \emptyset \textit{ falls } i \neq j; j \in I, \bigcup_{i \in I} Z_i = \Omega, P(Z_i) > 0, \quad i \in I.$$

(Bei jeder Versuchsdurchführung tritt also eines und nur eines der Ereignisse Z_i ein.)

Dann gilt für jedes Ereignis $A \in \mathfrak{A}$:

$$P(A) = \sum_{i \in I} P(A|Z_i)P(Z_i). \tag{5.4}$$

Beweis:

$$P(B) = P(B \cap \Omega) = P(B \cap (\bigcup_{i \in I} Z_i)) =$$

$$P(\bigcup_{i \in I}(B \cap Z_i)) = \sum_{i \in I} P(B \cap Z_i) = \sum_{i \in I} P(B|Z_i)P(Z_i).$$

$\square$

Folgerungen 5.7 Wenn $Z_1 = A, Z_2 = \bar{A}, 0 < P(A) < 1$ gilt, so ist

$$P(B) = P(B|A)P(A) + P(B|\bar{A})P(\bar{A}). \tag{5.5}$$

Beispiel 5.8 Aus einer Urne mit R roten und S schwarzen Kugeln ($R + S = M$) werden auf gut Glück nacheinander und ohne Zurücklegen zwei Kugeln gezogen. Wir definieren:

$A_1 :=$ "Die erste Kugel ist rot", $A_2 :=$ "Die zweite Kugel ist rot".

Wie groß sind $P(A_1)$ und $P(A_2)$?

Die Wahrscheinlichkeit $P(A_1)$ ist einfach zu bestimmen, da es sich beim ersten Ziehen um ein Laplace-Experiment handelt:

$$P(A_1) = \frac{R}{M}.$$

Wie groß die Wahrscheinlichkeit $P(A_2)$ ist, ist nicht unmittelbar ersichtlich. Zu ihrer Berechnung greifen wir auf (5.6) zurück:

$$P(A_2) = P(A_2|A_1)P(A_1) + P(A_2|\bar{A}_1)P(\bar{A}_1)$$

Wir wissen bereits, dass $P(A_1) = \frac{R}{M}$ und $P(\bar{A}_1) = 1 - P(A_1) = \frac{S}{M}$ gilt.

Die bedingten Wahrscheinlichkeiten $P(A_2|A_1)$ und $P(A_2|\bar{A}_1)$ sind ebenfalls einfach zu bestimmen:

$$P(A_2|A_1) = \frac{R-1}{M-1} \quad , \quad P(A_2|\bar{A}_1) = \frac{R}{M-1}$$

Folglich ergibt sich

$$P(A_2) = \frac{R-1}{M-1} \cdot \frac{R}{M} + \frac{R}{M-1} \cdot \frac{S}{M} = \frac{R}{M}.$$

Diese Wahrscheinlichkeit ist genauso groß wie $P(A_1)$. Man beachte, dass man bei der Berechnung von $P(A_2)$ nicht annimmt, dass bekannt ist, ob A_1 eingetreten ist oder nicht.

Es seien $(Z_i, i \in I)$ eine Zerlegung von Ω wie in Aussage 5.6 und $B \in \mathfrak{A}$ mit $P(B) > 0$.

Aussage 5.9 *("Bayes'sche Formel"):*
Es gilt für jedes $j \in I$:

$$P(Z_j|B) = \frac{P(B|Z_j)P(Z_j)}{\sum\limits_{i \in I} P(B|Z_i)P(Z_i)}.$$

Beweis: Nach Definition 5.1 gilt

$$P(Z_j|B) = \frac{P(Z_j \cap B)}{P(B)} = \frac{P(B|Z_j)P(Z_j)}{P(B)}.$$

Auf den Nenner wird nunmehr der Satz der totalen Wahrscheinlichkeit angewandt, Aussage 5.6. □

Beispiel 5.10 Reihenuntersuchung auf Tuberkulose (Tbc)

$A :=$ "Der Test auf Tbc ist beim untersuchten Patienten positiv"

$B :=$ "Der untersuchte Patient leidet unter Tbc"

Aus langen Testserien weiß man

$P(B) = \frac{1}{109}, \quad P(A|B) = 0,96, \quad P(\bar{A}|\bar{B}) = 0,92$ und somit insbesondere

$P(A|\bar{B}) = 0,08.$

Mit $Z_1 = B, Z_2 = \bar{B}$ ergibt sich

$$P(B|A) = \frac{P(A|B)P(B)}{P(A|B)P(B) + P(A|\bar{B})P(\bar{B})} = \frac{0,96 \cdot \frac{1}{109}}{0,96 \cdot \frac{1}{109} + 0,08 \cdot \frac{108}{109}} = 0,1.$$

Das bedeutet für die gegebenen Erfahrungswerte, dass bei positivem Testausgang die Wahrscheinlichkeit, dass der untersuchte Patient tatsächlich an Tbc leidet, nur 0,1 beträgt.

Bedingte Wahrscheinlichkeiten im Laplace-Modell

Es sei $(\Omega, \mathscr{P}(\mathfrak{A}), P)$ ein Laplace-Modell mit $\#\Omega = N$.
Weiterhin seien $A, B \subseteq \Omega$ mit $N(B) = \#B > 0$.

Angenommen, wir wissen bereits, dass das Ereignis B bei diesem zufälligen Versuch eintritt. Dann gibt es für uns nur noch $N(B)$ mögliche Versuchsausgänge ω. Alle diese Ausgänge sehen wir wieder als gleich wahrscheinlich an. Bei $N(A \cap B) = \#(A \cap B)$ von ihnen tritt auch A ein.

Also gilt

$$P(A|B) = \frac{N(A \cap B)}{N(B)} = \frac{N(A \cap B)/N}{N(B)/N} = \frac{P(A \cap B)}{P(B)}$$

Die Definition der bedingten Wahrscheinlichkeit $P(B|A)$ ist also nicht nur aus der Sicht des empirischen Gesetzes der großen Zahlen, sondern auch aus der Sicht der Laplace-Modelle vernünftig.

Beispiel 5.11 Im Nachbarraum werden zwei reguläre Würfel geworfen. Wir erhalten die Nachricht, dass die Augensumme kleiner als sieben ist. Wie groß ist unter dieser Bedingung die Wahrscheinlichkeit dafür, dass mindestens einer der Würfel eine Eins zeigt? $\left(= \frac{3}{5} \right)$

Bedingte Wahrscheinlichkeiten in mehrstufigen Versuchen

Wir haben in Abschnitt 4.7 die Wahrscheinlichkeiten $P(A)$ für Ereignisse A berechnet, die mit mehrstufigen zufälligen Versuchen zusammenhängen. Mit den dortigen Bezeichnungen gilt:

$$P(A) = \sum_{\omega \in A} P(\{\omega\}) \text{ und} \tag{5.6}$$

$$P(\{\omega\}) = p^{(1)}(\omega_1) p^{(2)}_{\omega_1}(\omega_2) \cdots p^{(n)}_{\omega_1 \cdots \omega_{n-1}}(\omega_n) \tag{5.7}$$

für $\omega = (\omega_1, \omega_2, \cdots, \omega_n)$.

Dabei bildeten die Größen $p^{(1)}(\omega_1)$ die Anfangsverteilung und Größen $p^{(k)}_{\omega_1 \cdots \omega_{t-1}}(\omega_k)$ die vorgegebene "Übergangswahrscheinlichkeiten", aus denen $P(\{\omega\})$ berechnet wird. Auf der Grundlage der so konstruierten Wahrscheinlichkeitsverteilung Pbestimmen wir nunmehr gewisse bedingte Wahrscheinlichkeiten.

Wir beginnen mit einem zweistufigen Versuch mit höchstens abzählbar vielen Ausgängen. Dann haben wir

$$P(A) = \sum_{\omega \in A} P(\{\omega\}) \text{ und}$$

$$P(\{\omega\}) = p^{(1)}(\omega_1) p^{(2)}_{\omega_1}(\omega_2), \quad \omega = (\omega_1, \omega_2),$$

wobei die Anfangsverteilung $p^{(1)}(\omega_1), \omega_1 \in \Omega_1$, und die Übergangswahrscheinlichkeiten $p^{(2)}_{\omega_1}(\omega_2), \omega_1 \in \Omega_1, \omega_2 \in \Omega_2$, gegeben seien.

Setzt man

$A := \{\omega \in \Omega |$ In der zweiten Versuchsstufe erscheint der Ausgang $\omega_2\}$ und

$B := \{\omega \in \Omega |$ In der ersten Versuchsstufe erscheint der Ausgang $\omega_1\}$,

so gilt

$$P(A|B) = \frac{P(A \cap B)}{P(B)} = \frac{\sum\limits_{\omega \in A \cap B} P(\{\omega\})}{\sum\limits_{\omega \in B} P(\{\omega\})} =$$

$$\frac{p^{(1)}(\omega_1) p^{(2)}_{\omega_1}(\omega_2)}{\sum\limits_{\omega_2' \in \Omega_2} p^{(1)}(\omega_1) p^{(2)}_{\omega_1}(\omega_2')} = p^{(2)}_{\omega_1}(\omega_2). \tag{5.8}$$

Das heißt, die Wahrscheinlichkeit $p^{(2)}_{\omega_1}(\omega_2)$ ist die bedingte Wahrscheinlichkeit $P(A|B)$ bezüglich des Maßes P, was durchaus der Anschauung und Konstruktion entspricht.

Wir wollen im Folgenden zeigen, dass in jedem n-stufigen Versuch auch die Größen $p^{(k)}_{\omega_1\cdots\omega_{k-1}}(\omega_k)$ als bedingte Wahrscheinlichkeiten bezüglich der Verteilung P auffassen lassen.

Definiert man

$A := \{\omega \in \Omega|$ In der k-ten Versuchsstufe erscheint der Versuchsausgang $\omega_k\}$

und

$B := \{\omega \in \Omega|$ In den ersten $k-1$ Versuchen erscheinen jeweils $\omega_1, \omega_2, \cdots, \omega_{k-1}\}$, so erhalten wir

Aussage 5.12

$$P(A|B) = \frac{P(A\cap B)}{P(B)} = \frac{\sum\limits_{\omega\in A\cap B} P(\{\omega\})}{\sum\limits_{\omega\in B} P(\{\omega\})} =$$

$$= p^{(k)}_{\omega_1\cdots\omega_{k-1}}(\omega_k).$$

Beweis: Wegen Formel (5.9) haben wir

$$P(A|B) = \frac{P(A\cap B)}{P(B)} = \frac{\sum\limits_{\omega\in A\cap B} P(\{\omega\})}{\sum\limits_{\omega\in B} P(\{\omega\})} =$$

$$\frac{\sum\limits_{\omega'_{k+1},\cdots,\omega'_n} p^{(1)}(\omega_1)\cdots p^{(k)}_{\omega_1}\cdots_{\omega_{k-1}}(\omega_k) p^{(k+1)}_{\omega_1\cdots\omega_k}(\omega'_{k+1})\cdots p^{(n)}_{\omega_1\cdots\omega_{n-1}}(\omega'_n)}{\sum\limits_{\omega'_k,\omega'_{k+1},\cdots,\omega'_n} p^{(1)}(\omega_1)\cdots p^{(k)}_{\omega_1\cdots\omega_{k-1}}(\omega'_k) p^{(k+1)}_{\omega_1\cdots\omega'_k}(\omega'_{k+1})\cdots p^{(n)}_{\omega_1\cdots\omega'_{n-1}}(\omega'_n)}$$

$$= p^{(k)}_{\omega_1\cdots\omega_{k-1}}(\omega_k).$$

□

Das bedeutet, die "Bausteine" $p^{(k)}_{\omega_1\cdots\omega_{k-1}}(\omega_k)$, aus denen sich P zusammen mit der Anfangsverteilung $p^{(1)}(\cdot)$ ergibt, erweisen sich aus der Sicht von P gerade als die bedingte Wahrscheinlichkeit dafür, dass der k-te Versuch mit ω_k endet, unter der Bedingung, dass in den vorangegangenen Versuchen die Resultate $\omega_1, \omega_2, \cdots, \omega_{k-1}$ auftraten.

5.2 Unabhängigkeit

Wenn die Vorinformation, dass das Ereignisse B eintreten wird, die Wahrscheinlichkeit des Eintretens des Ereignisses A nicht verändert, so sagt man, A und B seien voneinander unabhängige Ereignisse.

Den Begriff der Unabhängigkeit kann man erweitern auf Zufallsgrößen und auf Ereignissysteme wie σ-Algebren. Er steht im Mittelpunkt der klassischen Wahrscheinlichkeitstheorie und Mathematischen Statistik und liegt klassischen Gesetzen der großen Zahlen und zentralen Grenzwertsätzen zugrunde, die wir später kennen lernen werden.

5.2.1 Unabhängigkeit von Ereignissen

Es seien $(\Omega, \mathfrak{A}, P)$ ein Wahrscheinlichkeitsraum und $A, B \in \mathfrak{A}$ mit $P(B) \in (0,1)$.

Im Allgemeinen gilt $P(A|B) \neq P(A)$, d.h., die Kenntnis, dass B eingetreten ist, beeinflusst die Wahrscheinlichkeit des Eintretens von A.

In manchen Fällen gilt allerdings $P(A|B) = P(A)$. Das bedeutet, das Wissen, dass B eintritt, verändert die Wahrscheinlichkeit von A nicht. In diesem Fall gilt übrigens auch

$$P(A|\bar{B}) = \frac{P(A \cap \bar{B})}{P(\bar{B})} = \frac{P(A) - P(A \cap B)}{P(\bar{B})}$$

$$= \frac{P(A) - P(A|B)P(B)}{P(\bar{B})} = P(A). \tag{5.9}$$

Definition 5.13 *Zwei Ereignisse $A, B \in \mathfrak{A}$ heißen voneinander* stochastisch unabhängig *(bezüglich der Wahrscheinlichkeitsverteilung P), falls $P(A \cap B) = P(A)P(B)$ gilt. (Im Fall $P(B) > 0$ bedeutet das $P(A|B) = P(A)$, und für $P(A) > 0$ ergibt sich auch $P(B|A) = P(B)$.)*

Die Unabhängigkeit ist eine symmetrische Eigenschaft bezüglich A und B. Statt stochastisch unabhängig sagt man auch einfach unabhängig.

Aussage 5.14 *A, B unabhängig $\Longleftrightarrow A, \bar{B}$ unabhängig.*

Beweis: $P(A \cap B) = P(A)P(B) \Longleftrightarrow P(A \cap \bar{B}) =$

$$P(A) - P(A \cap B) = P(A)(1 - P(B)) = P(A)P(\bar{B}).$$

Folgerungen 5.15 A, B unabhängig $\Longleftrightarrow A, \bar{B}$ unabhängig $\Longleftrightarrow \bar{A}, B$ unabhängig $\Longleftrightarrow \bar{A}\bar{B}$ unabhängig.

Definition 5.16 *$(\Omega, \mathfrak{A}, P)$ sei ein Wahrscheinlichkeitsraum, I eine beliebige Indexmenge und $A_i \in \mathfrak{A}, i \in I$.*

a) Die $A_i, i \in I$, heißen voneinander paarweise stochastisch unabhängig *bezüglich der Wahrscheinlichkeitsverteilung P, wenn gilt:*

$$P(A_i \cap A_j) = P(A_i) \cdot P(A_j)$$

für beliebige Indizes $i, j \in I$ mit $i \neq j$.

b) Die A_i heißen voneinander (in ihrer Gesamtheit) stochastisch unabhängig *unter der Wahrscheinlichkeitsverteilung P, wenn gilt:*

$$P(A_i \cap \ldots \cap A_{i_m}) = P(A_{i_1}) \cdot \ldots \cdot P(A_{i_m})$$

für je endlich viele verschiedene Indizes $i_1, \ldots, i_m \in I$.

Statt in ihrer Gesamtheit unabhängiger Ereignisse sprechen wir auch einfach von *unabhängigen Ereignissen.*
Aus dieser Definition ergeben sich unmittelbar folgende Eigenschaften.

Folgerungen 5.17

1) Ist $(A_i, i \in I)$ eine Familie von unabhängigen Ereignissen, so sind auch die Ereignisse jeder Teilfamilie $(A_i, i \in I'), I' \subseteq I$, unabhängig.

2) Mit $(A_i, i \in I)$ ist auch $(C_i, i \in I)$, wobei $C_i = A_i$ oder $C_i = \bar{A}_i, i \in I$, gilt, eine Familie unabhängiger Ereignisse.

   Der Beweis von 2) erfolgt analog zum Beweis der Aussage (5.14) und ist nur schreibtechnisch komplizierter.

Aus der Unabhängigkeit der Ereignisse $(A_i, i \in I)$ folgt ihre paarweise Unabhängigkeit. Die Umkehrung gilt nicht, siehe Übungen.

Aussage 5.18 *Ist $(A_i, i \in I)$ mit $I \subseteq N_0$ eine Folge voneinander unabhängiger Ereignisse, so gilt*

1)

$$P\ (\textit{Alle } A_i, i \in I \textit{ treten ein}) = P(\bigcap_{i \in I} A_i) = \prod_{i \in I} P(A_i) \tag{5.10}$$

2)

$$\begin{aligned} P\ (\textit{Mindestens eines der } A_i, i \in I,\ \textit{tritt ein}) &= P(\bigcup_{i \in I} A_i) \\ &= 1 - \prod_{i \in I} (1 - P(A_i)) \end{aligned} \tag{5.11}$$

Beweis: Es sei $I_n = I \cap [0, n], n \geq 1$.

1) Aus der Unabhängigkeit der Ereignisse $A_i, i \in I$, folgt $P(\cap_{i\in I_n} A_i) = \prod_{i\in I_n} P(A_i)$ für jedes $n \geq 1$. Daraus folgt (...) durch Grenzübergang $n \to \infty$ unter Verwendung der Stetigkeit von P bezüglich monotoner Ereignisfolgen.

2) Mit $A_i, i \in I$, sind auch die Ereignisse $\bar{A}_i, i \in I$, voneinander unabhängig. Folglich ist wegen der eben bewiesenen Eigenschaft a)

$$P(\bigcup_{i\in I} A_i) = 1 - P(\bigcap_{i\in I} \bar{A}_i) = 1 - \prod_{i\in I} P(\bar{A}_i), \text{ woraus sich (...) ergibt.} \qquad \square$$

Beispiel 5.19 Die Wahrscheinlichkeit dafür, dass beim rein zufälligen Austeilen eines Skatspieles zwei Buben im Skat liegen, beträgt $p_1 = 0,0121$. Niemand rechnet also bei einem einmaligen Spiel damit, dass dieses Ereignis tatsächlich eintritt. Die Wahrscheinlichkeit dafür, dass dieses Ereignis in einer Serie von n voneinander unabhängigen Spielen mindestens einmal stattfindet, ist dagegen $p_n = 1 - (1 - p_1)^n$. Für $n = 20$ ergibt das $p_{20} = 0,216$, und für $n = 50$ erhalten wir $p_{50} = 0,456$. Ab welchem n kann man darauf wetten, dass in einer Serie von n Skatspielen mindestens einmal zwei Buben im Skat liegen? (Vgl. Abschnitt 3.4a)

Aussage 5.20 *(2. Lemma von Borel-Cantelli)*
Es sei $(\Omega, \mathfrak{A}, P)$ ein Wahrscheinlichkeitsraum, und $(A_n, n \geq 1)$ sei eine Folge voneinander unabhängiger Ereignisse aus $\mathfrak{A}$. Dann folgt aus $\sum_{n\geq 1} P(A_n) = \infty$, dass $P(\limsup_{n\to\infty} A_n) = 1$ gilt.

Beweis:

$$1 - P(\limsup A_n) = P(\overline{\bigcap_n \bigcup_{m\geq n} A_m}) =$$

$$P(\bigcup_{n=1}^{\infty} \bigcap_{m\geq n} \bar{A}_m) = \lim_{n\to\infty} P(\bigcap_{m\geq n} \bar{A}_m) =$$

$$\lim_{n\to\infty} \lim_{k\to\infty} P(\bigcap_{m=n}^{k} \bar{A}_m) = \lim_{n\to\infty} \lim_{k\to\infty} \prod_{m=n}^{k} (1 - P(A_m)).$$

Unter Ausnutzung von $1 - x \leq \mathrm{e}^{-x}, x \in R_1$, folgt

$$1 - P(\limsup A_n) \leq \lim_{n\to\infty} \lim_{k\to\infty} \exp[-\sum_{m=n}^{k} P(A_m)] = 0.$$

□

5.2.2 Unabhängigkeit von σ-Algebren

Wir erweitern in diesem Punkt den Unabhängigkeitsbegriff auf allgemeine Strukturen, da wir ihn später in dieser Form benötigen werden.

Definition 5.21 *Es seien* $(\Omega, \mathfrak{A}, P)$ *ein Wahrscheinlichkeitsraum,* I *eine Indexmenge und* γ_i *für jedes* i *ein Mengensystem mit* $\gamma_i \subseteq \mathfrak{A}$. *Die* γ_i *heißen* voneinander unabhängig, *falls für jede endliche Auswahl* $\mathfrak{J} \subseteq I$ *und jede Auswahl* $A_j \in \gamma_j,\ j \in \mathfrak{J}$, *gilt:*

$$P(\bigcap_{j\in\mathfrak{J}} A_j) = \prod_{j\in\mathfrak{J}} P(A_j). \tag{5.12}$$

Beispiel 5.22

$$\Omega = [0,1), \mathfrak{A} = \mathfrak{B}_{[0,1)}, P = \lambda_{[0,1)}, \gamma_1 = \{[0,\frac{1}{2}), [\frac{1}{2},1)\},$$

$$\gamma_2 = \{[\frac{1}{4},\frac{3}{4}), ([0,\frac{1}{4}) \cup [\frac{3}{4},1))\}$$

γ_1 und γ_2 sind bez. P unabhängige Ereignissysteme.

Die folgende Aussage ist nützlich zum Nachweis, dass σ-Algebren voneinander unabhängig sind, da die sie erzeugenden Semialgebren im Allgemeinen von wesentlich einfacherer Gestalt sind.

Aussage 5.23 *Sind die Mengensysteme* $\gamma_i, i \in I$, *aus der vorangegangenen Definition voneinander unabhängige Semiringe, so sind auch die von ihnen erzeugten* σ*-Algebren* $\mathfrak{A}_i = \sigma(\gamma_i),\ i \in I$, *voneinander unabhängig.*

Beweis: Es sei $\mathfrak{J}$ eine endliche Teilmenge von I und für jedes $j \in \mathfrak{J}$ sei A_j ein Element von γ_j. Nach Voraussetzung gilt

$$P(\bigcap_{j\in\mathfrak{J}} A_j) = \prod_{j\in\mathfrak{J}} P(A_j). \tag{5.13}$$

Fixiert man ein $j_0 \in \mathfrak{J}$ und alle A_j mit $j \in \mathfrak{J} \setminus \{j_0\}$, so stellen beide Seiten von (8) σ-additive Mengenfunktionen auf der Semialgebra γ_{j_0} dar, die sich eindeutig zu Wahrscheinlichkeitsmaßen auf $\mathfrak{A}_{j_0}$ fortsetzen lassen. Also gilt (...) bei festen $A_j, (j \in \mathfrak{J} \setminus \{j_0\})$, für alle A_{j_0} aus $\mathfrak{A}_{j_0}$. Diese Überlegung wird sukzessiv fortgesetzt, und so ergibt sich die Behauptung. □

Auch die folgende Aussage wird im Weiteren nützlich sein.

Aussage 5.24 *Es sei $\mathfrak{A}_i, i \in I$, eine Familie voneinander unabhängiger Teil-σ-Algebren von $\mathfrak{A}$. Weiterhin seien $\mathfrak{J}_1$ und $\mathfrak{J}_2$ zwei disjunkte Teilmengen von I. Dann sind auch $\mathfrak{A}_{\mathfrak{J}_1} := \sigma\Big(\bigcup_{i\in\mathfrak{J}_1} \mathfrak{A}_i\Big)$ und $\mathfrak{A}_{\mathfrak{J}_2} := \sigma\Big(\bigcup_{i\in\mathfrak{J}_2} \mathfrak{A}_i\Big)$ voneinander unabhängig.*

Beweis: $\gamma_1 = \bigcup_{i\in\mathfrak{J}_1} \mathfrak{A}_i$ und $\gamma_2 = \bigcup_{i\in\mathfrak{J}_2} \mathfrak{A}_i$ sind nach Voraussetzung zwei unabhängige Ereignissysteme, die überdies Algebren sind. Daraus folgt auf Grund der Aussage 5.23, dass auch $\mathfrak{A}_{\mathfrak{J}_1}$ und $\mathfrak{A}_{\mathfrak{J}_2}$ unabhängig sind. □

5.2.3 Unabhängigkeit in mehrstufigen Experimenten

Es sei $(\Omega, \mathfrak{A}, P)$ ein n-stufiges zufälliges Experiment mit den Einzelexperimenten $(\Omega_k, \mathfrak{A}_k, P_k),\ k = 1, \ldots, n$. Das heißt, es seien
$\Omega = \prod_1^{n} \Omega_k$, $\quad \mathfrak{A} = \prod_1^{n\,\otimes} \mathfrak{A}_k$ und P eine Wahrscheinlichkeitsverteilung auf $\mathfrak{A}$.
Mögliche Ausgänge für das n-stufige Experiment sind somit die Folgen $\omega = (\omega_1, \omega_2, \ldots, \omega_n)$ von Ausgängen ω_k der Einzelexperimente. Der Zusammenhang zwischen P und den P_k ist gegeben durch

$$P_k(A_k) = P(\Omega_1 \times \ldots \times \Omega_{k-1} \times A_k \times \Omega_{k+1} \times \ldots \times \Omega_n) \tag{5.14}$$

für alle $A_k \in \mathfrak{A}_k$ und $k = 1, \ldots, n$. Wir identifizieren jedes Ereignis $A_k \in \mathfrak{A}_k$ mit dem ihm zugeordneten Ereignis A'_k im Gesamtexperiment $(\Omega, \mathfrak{A}, P)$:

$$A'_k := \Omega_1 \times \ldots \times \Omega_{k-1} \times A_k \times \Omega_{k+1} \times \ldots \times \Omega_n, \ k = 1, \ldots, n,$$

und setzen $\mathfrak{A}'_k := \{A'_k \in \mathfrak{A} : A_k \in \mathfrak{A}_k\}$,
Dann bedeutet (...)

$$P_k(A_k) = P(A'_k), \quad A_k \in \mathfrak{A}_k, \ k = 1, \ldots, n.$$

Außerdem gilt mit dieser Terminologie

$$P(A'_1 \cap A'_2 \cap \ldots \cap A'_n) = P(A_1 \times A_2 \times \ldots \times A_n). \tag{5.15}$$

Definition 5.25 *Wir sagen, die Einzelexperimente $(\Omega_k, \mathfrak{A}_k, P_k)$ seien im Rahmen des Gesamtexperimentes $(\Omega, \mathfrak{A}, P)$ voneinander unabhängig, falls die σ-Algebren $\mathfrak{A}'_k$ bezüglich P voneinander unabhängig sind, d.h., falls gilt:*

$$P(A'_1 \ldots \cap A'_n) = \prod_1^n P(A'_k), \quad A'_k \in \mathfrak{A}'_k, \ k = 1, \ldots, n. \tag{5.16}$$

Aussage 5.26 *Die Einzelexperimente $(\Omega_k, \mathfrak{A}_k, P_k)$ sind im Rahmen des Gesamtexperimentes $(\Omega, \mathfrak{A}, P)$ genau dann unabhängig, wenn P das Produktmaß der $P_k, \ k = 1, \ldots, n$, ist, d.h., falls für jede Auswahl $A_k \in \mathfrak{A}_k, k = 1, 2, \ldots, n$ gilt:*

$$P(A_1 \times \ldots \times A_n) = \prod_1^n P_k(A_k). \tag{5.17}$$

Beweis: (...) ist mit (...) identisch, siehe (...) und (...).

Beide Seiten von (...) bilden σ-additive Mengenfunktionen auf der Semialgebra $\gamma = \{A_1 \times A_2 \times \ldots \times A_n | A_k \in \mathfrak{A}, \ k = 1, \ldots, n\}$ und sind dort gleich. Wegen der eindeutigen Fortsetzbarkeit stimmen sie auch auf $\mathfrak{A} = \sigma(\gamma)$ überein.
$\square$

Folgerungen 5.27 Sind alle $\Omega_k,\ k = 1, \ldots, n$, höchstens abzählbar, $\mathfrak{A}_k = \mathfrak{P}(\Omega_k)$, und ist P gegeben durch die Einzelwahrscheinlichkeiten

$$P(\{\omega\}) = p^{(1)}(\omega_1)p^{(2)}_{\omega_1}(\omega_2) \cdot \ldots \cdot p^{(n)}_{\omega_1\omega_2\ldots\omega_{n-1}}(\omega_n) \text{ mit } \omega = (\omega_1 \ldots, \omega_n),$$

so sind die Einzelexperimente genau dann voneinander unabhängig (bez. P), wenn die Übergangswahrscheinlichkeiten

$$p^{(k)}_{\omega_1\ldots\omega_{k-1}}(\omega_k) =: p^{(k)}(\omega_k),\ k = 2, \ldots, n, \tag{5.18}$$

nicht von $\omega_1, \ldots, \omega_{k-1}$ abhängen.

In diesem Fall gilt

$$P_k(\{\omega_k\}) = p^{(k)}(\omega_k) \text{ und}$$

$$P(\{\omega\}) = \prod_{k=1}^{n} p^{(k)}(\omega_k), \qquad \omega = (\omega_1, \ldots, \omega_n).$$

Dieser Sachverhalt entspricht dem anschaulichen Unabhängigkeitsbegriff bei der Konstruktion des mehrstufigen Experimentes aus den Größen $p^{(k)}_{\omega_1\cdots\omega_{k-1}}$.

5.2.4 Unabhängigkeit von Zufallsgrößen

Es seien X und Y zwei Zufallsgrößen auf einem Wahrscheinlichkeitsraum mit Werten in $(E, \mathfrak{E})$ bzw. $(F, \mathfrak{F})$.

Definition 5.28 *Die Zufallsgrößen X und Y heißen voneinander unabhängig, falls für jede Wahl von Mengen $C \in \mathfrak{E}$ und $D \in \mathfrak{F}$ die Ereignisse $\{X \in C\}$ und $\{Y \in D\}$ voneinander unabhängig sind, d.h., falls gilt:*

$$P(X \in C, Y \in D) = P(X \in C)P(Y \in D),\ C \in \mathfrak{E}, D \in \mathfrak{F}. \tag{5.19}$$

Aus dieser Definition ist wegen $\{X \in C, Y \in D\} = X^{-1}(C) \cap Y^{-1}(D)$ offenbar, dass die Unabhängigkeit von X und Y äquivalent damit ist, dass die σ-Algebren $\mathfrak{A}^X = X^{-1}(\mathfrak{E}) = \big\{\{X \in C\} | C \in \mathfrak{E}\big\}$ und $\mathfrak{A}^Y = Y^{-1}(\mathfrak{F}) = \big\{\{Y \in D\} | D \in \mathfrak{F}\big\}$ voneinander unabhängig sind.

Die Eigenschaft (...) lässt sich wegen $\{X \in C, Y \in D\} = \{(X,Y) \in C \times D\}$ auch schreiben als

$$P^{(X,Y)}(C \times D) = P^X(C)P^Y(D), \quad C \in \mathfrak{E}, D \in \mathfrak{F}.$$

Die Unabhängigkeit der beiden Zufallsgrößen X und Y bedeutet also, dass ihre gemeinsame Wahrscheinlichkeitsverteilung $P^{(X,Y)}$ gleich dem Produktmaß ihrer beiden Randverteilungen P^X und P^Y ist. Die Unabhängigkeit der zwei Zufallsgrößen X und Y ist somit eine Eigenschaft ihrer gemeinsamen Verteilungen $P^{(X,Y)}$. Im Fall der Unabhängigkeit von X und Y ist die gemeinsame Verteilung $P^{(X,Y)}$ durch ihre Randverteilungen P^X und P^Y also eindeutig bestimmt.

Die folgende Aussage erlaubt es, die Unabhängigkeit zweier reellwertiger Zufallsgrößen anhand ihrer gemeinsamen Verteilungsfunktion zu prüfen.

Aussage 5.29 *Zwei reellwertige Zufallsgrößen X und Y über einem Wahrscheinlichkeitsraum $(\Omega, \mathfrak{A}, P)$ sind genau dann unabhängig, wenn für ihre gemeinsame Verteilungsfunktion $F_{(X,Y)}$ und ihre Randverteilungsfunktionen F_X und F_Y gilt:*

$$F_{(X,Y)}(x,y) = F_X(x)F_Y(y), \qquad x, y \in R_1. \tag{5.20}$$

Beweis: Sind X und Y unabhängig, so gilt

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

Gilt dagegen (...), so ist mit $F = F_{X,Y}$

$$\begin{gathered} P(X \in (a_1, b_1], Y \in (a_2, b_2]) = \\ F(b_1, b_2) - F(a_1, b_2) - F(b_1, a_2) + F(a_1, a_2) = \\ F_X(b_1)F_Y(b_2) - F_X(a_1)F_Y(b_2) - F_X(b_1)F_Y(a_2) + F_X(a_1)F_Y(a_2) = \\ (F_X(b_1) - F_X(a_1))(F_X(b_2) - F_Y(a_2)) = \\ P(X \in (a_1, b_1]) \cdot P(Y \in (a_2, b_2]). \end{gathered} \tag{5.21}$$

Es sei γ der Semiring aller halboffenen Intervalle $(a,b] : \gamma = \{(a,b] | -\infty < a < b < \infty\}$. Bekanntlich gilt $\sigma(\gamma) = \mathfrak{B}_1$. Die Gleichung (...) besagt, dass die

Semiringe $X^{-1}(\gamma)$ und $Y^{-1}(\gamma)$ unabhängig sind. Somit sind auch $\sigma(X^{-1}(\gamma)) = X^{-1}(\sigma(\gamma)) = X^{-1}(\mathfrak{B}_1)$ und $\sigma(Y^{-1}(\gamma)) = Y^{-1}(\sigma(\gamma)) = X^{-1}(\mathfrak{B}_1)$ unabhängig, was aber mit der Unabhängigkeit von X und Y äquivalent ist. $\square$

Die Definitionen und Aussagen dieses Abschnittes lassen sich analog zu Abschnitt 5.2. auf beliebige Familien $(X_i, i \in I)$ von Zufallsgrößen über einem Wahrscheinlichkeitsraum $(\Omega, \mathfrak{A}, P)$ verallgemeinern. Wir beschränken uns auf folgende

Definition 5.30 *Es seien I irgendeine Indexmenge und $X_i, i \in I$, Zufallsgrößen über einem Wahrscheinlichkeitsraum $(\Omega, \mathfrak{A}, P)$ mit Werten in $(E_i, \mathfrak{E}_i), i \in I$. Die Zufallsgrößen $X_i, i \in I$, heißen voneinander unabhängig, falls für jede endliche Teilmenge $\mathfrak{J}$ von I und jede Auswahl $B_j \in \mathfrak{E}_j, j \in \mathfrak{J}$, gilt:*

$$P(\bigcap_{j\in\mathfrak{J}}\{X_j \in B_j\}) = \prod_{j\in J} P(X_j \in B_j).$$

Wir betrachten noch die zwei Sonderfälle, dass (X, Y) diskret verteilt ist bzw. eine gemeinsame Dichte besitzt.

Diskret verteilte zufällige Vektoren

Die gemeinsame Verteilung der Zufallsgrößen X, Y sei diskret und gegeben durch $((x_i, y_j), p_{ij}), i \in I, i \in \mathfrak{J}$.

a) X und Y sind genau dann unabhängig, wenn

$$p_{ij} = p_{i\cdot}p_{\cdot j}, \ i \in I, j \in \mathfrak{J} \tag{5.22}$$

gilt, wobei nach Definition $p_{i\cdot} = \sum_j p_{ij}, i \in I$, und $p_{\cdot j} = \sum_i p_{ij}, j \in \mathfrak{J}$, die Einzelwahrscheinlichkeiten von X bzw. Y sind.

Beweis: (5.25) folgt aus (5.22) mit $C = \{x_i\}, D = \{y_j\}$. Umgekehrt ergibt sich (5.22) aus (5.25) mittels

$$P(X \in C, Y \in D) = \sum_{\substack{i:x_i \in C,\\ j:y_j \in D}} p_{ij} = \sum_{\substack{i:x_i \in C,\\ j:y_j \in D}} p_{i\cdot} p{\cdot}j = \sum_{i:x_i \in C} p_{i\cdot} \cdot \sum_{j:y_j \in D} p_{\cdot j} =$$

$$P(X \in C)P(Y \in D)$$

□

b) Sind X und Y reellwertig und unabhängig, existieren die Erwartungswerte EX, EY und sind sie beide endlich, so existiert auch der Erwartungswert $E(XY)$ und ist endlich.

Überdies gilt in diesem Fall:

$$E(XY) = EX\,EY. \tag{5.23}$$

Beweis: Nach Voraussetzung ist $\sum_i |x_i| p_{i\cdot} < \infty$ und $\sum_j |y_i| p_{\cdot j} < \infty$. Folglich gilt auch

$$\sum_{i,j} |x_i y_j| p_{ij} = \sum_{i,j} |x_i y_j| p_{i\cdot} p_{\cdot j} = \sum_i |x_i| p_{i\cdot} \cdot \sum_j |y_j| p_{\dot{j}} < \infty.$$

Außerdem gilt für diese absolut konvergente Reihe

$$\sum_{i,j} x_i y_i p_{ij} = \sum_i x_i \sum_j y_i p_{ij} = \sum_i x_i p_{i\cdot} \sum_j y_i p_{\cdot j}.$$

□

Die Formel (5.26) impliziert, dass X und Y unter den genannten Voraussetzungen unkorreliert sind:

$$\mathrm{Kov(X,Y) = E(XY) - EXEY = 0.}$$

Sind überdies die Streuungen D^2X und D^2Y endlich, so folgt aus der Unabhängigkeit von X und Y die Gleichung

$$D^2(X+Y) = D^2X + D^2Y.$$

Beweis:

$$\begin{aligned} D^2(X+Y) = E(X+Y-EX-EY)^2 &= D^2X + D^2Y + 2\,\mathrm{Cov(X,Y)} \\ &= D^2X + D^2Y. \end{aligned}$$

□

c) Die Einzelwahrscheinlichkeiten p_{ij} werden wie folgt in einer Tabelle aufgeschrieben.

$i\backslash^j$	1	2	3	$\ldots\ j\ \ldots$	
1	p_{11}	p_{12}	$\cdots$	$\cdots p_{1j} \cdots$	$p_{1\cdot}$
2	p_{21}	p_{22}		$\cdot$	$p_{2\cdot}$
3	$\cdot$			$\cdot$	
$\cdot$	$\cdot$			$\cdot$	
$\cdot$	$\cdot$			$\cdot$	
$\cdot$	$\cdot$			$\cdot$	
i	p_{i1}	$\cdots$		$\cdots p_{ij} \cdots$	$p_{i\cdot}$
$\cdot$	$\cdot$			$\cdot$	
	$p_{\cdot 1}$	$p_{\cdot 2}$		$p_{\cdot j}$	1

Teilt man die i-te Zeile durch den Wert $p_{i.}$ sofern $p_{i\cdot} > 0$ gilt, so erhält man eine Wahrscheinlichkeitsverteilung

$$\left(\frac{p_{ij}}{p_{i.}}, j \geq 1\right),$$

für die gilt

$$\frac{p_{ij}}{p_{i.}} = \frac{P(X = x_i, Y = y_j)}{P(X = x_i)} = P(Y = y_i \,|\, X = x_i).$$

Analog liefert die j-te Spalte, geteilt durch $p_{.j}$

$$\left(\frac{p_{ij}}{p_{.j}}, i \geq 1\right)$$

$$\text{mit } \frac{p_{ij}}{p_{.j}} = P(X = x_i | Y = y_j).$$

Bei fest gewählten y_j bilden die $P(X = x_i | Y = y_i), i \in \mathscr{J}$, die Einzelwahrscheinlichkeiten einer diskreten Verteilung, die man als *bedingte Verteilung von X unter der Bedingung* $\{Y = y_j\}$ bezeichnet. Analog definierten die $P(Y = y_j | X = x_i)$ die bedingte Verteilung von Y unter der Bedingung $\{X = x_i\}$.

Zufällige Vektoren mit gemeinsamer Dichte

a) Besitzt (X, Y) eine Dichte $f_{(X,Y)}$ und sind f_X und f_Y die entsprechenden Randverteilungsdichten, so sind X und Y genau dann unabhängig, wenn

$$f_{X,Y}(x, y) = f_X(x) f_{Y(y)} \quad \lambda \otimes \lambda \; - \; \text{f.ü.} \tag{5.24}$$

gilt.

Beweis: Aus (....) folgt (....) durch Differentation nach x und nach y. Aus (...) ergibt sich (...) nach Definition der Dichten. □

Der Beweis folgt unmittelbar aus (...), wir kommen später im allgemeineren Rahmen darauf zurück.

b) Zweidimensionale Normalverteilung: Es sei (X, Y) ein $N(\mu, \sum)$-verteilter zufälliger Vektor mit den Parametern $\mu_1, \mu_2, \sigma_1^2, \sigma_2^2, \rho$. Genau dann sind X und Y unabhängig, wenn $\rho = 0$ gilt.

Index

σ-Algebra, 15
σ-Stetigkeit von Wahrscheinlichkeitsverteilungen, 27

Algebra, 15
Anfangsverteilung, 111
Axiomensystem der Wahrscheinlichkeitstheorie, 24

Bayes'sche Formel, 122
bedingte Wahrscheinlichkeit, 118
  im Laplace-Modell, 123
  im mehrstufigen Versuch, 124
Binomialverteilung, 77
  Erwartungswert, 85
  erzeugende Funktion, 106
  Varianz, 90
Bonferroni-Ungleichungen, 30
Borel-Cantelli
  1. Lemma von, 29
  2. Lemma von, 130

Ein- und Ausschlussformel, 29
Einpunktverteilung, 76
  Erwartungswert, 85
  erzeugende Funktion, 106
  Varianz, 90
Ereignis
  fast sicheres, 49
  fast unmögliches, 49
  zufälliges, 11
Erwartungswert
  diskret, 84, 86
erzeugende Funktion, 102
Exponentialverteilung
  Verteilungsfunktion, 65

geometrische Verteilung, 77
  Erwartungswert, 85
  erzeugende Funktion, 106
  Varianz, 90
gleichmäßige Verteilung
  diskret, 33
gleichmäßige Verteilung, diskret, 76
  Erwartungswert, 85
  erzeugende Funktion, 106
  Varianz, 90

hypergeometrische Verteilung, 81–83
  Erwartungswert, 85
  erzeugende Funktion, 106
  Varianz, 90

Korrelationskoeffizient, 98
Kovarianz, 98

Laplace
  Experiment, 32

Münzenwurf, 16, 34
Median, 58
Moment
  diskret, 87, 88

diskret, zentriert, 87, 88
Multiplikationssatz für Wahrscheinlichkeiten, 120

negative Binomialverteilung, 78
Erwartungswert, 85
erzeugende Funktion, 106
Varianz, 90

Pfadregel
erste, 110
zweite, 110
Poissonverteilung, 77
Erwartungswert, 85
erzeugende Funktion, 106
Varianz, 90
Polya'sches Urnenschema, 111

Quantil, 58
unteres, oberes, 58

Randverteilung
diskret, 93
Regressionsgerade, 101

Streuung, *siehe* Varianz

totale Wahrscheinlichkeit, Satz von, 121

Uebergangsverteilung, 111
Unabhängigkeit
in mehrstufigen Experimenten, 133
von σ-Algebren, 131
von Ereignissen, 127, 128
von Ereignissen, paarweise, 128
von Mengensystemen, 131
von Zufallsgrößen, 134, 136
Ungleichung
von Cauchy-Schwarz, 97
von Tschebychev
diskret, 90
unkorreliert, 99
Urnenmodelle, 43–45

Varianz
diskret, 89–90
Verteilung
diskrete, 75
gemeinsame von U und V, diskret, 92
Wahrscheinlichkeits-, P^X, 52
Verteilungsdichte, 63
Verteilungsfunktion
der Zufallsgröße X, 55
diskret, 79

Wahrscheinlichkeitsmaß, 24
Wahrscheinlichkeitsraum, 25

zufälliger Vektor, 53
diskret, zweidimensional, 91
Funktionen diskreter, 94
zufälliger Versuch, 9
mehrstufig, 106–115
Zufallsgröße, 51
diskret, 78
reellwertige, 53