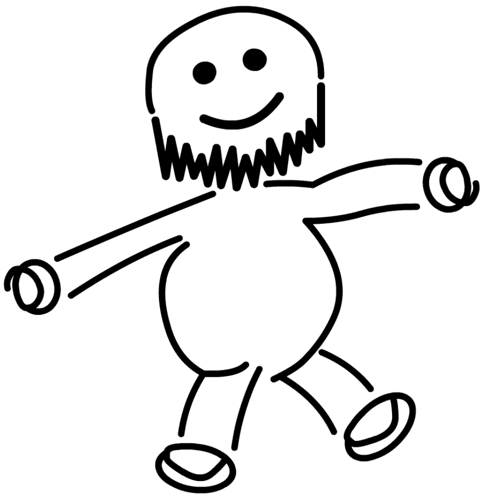
Stochastik



Daniel Scholz im Winter 2005 / 2006

Überarbeitete Version vom 3. Februar 2006.

Inhaltsverzeichnis

Bezeichnungen

Es sollen zunächst einige Bezeichnungen und Konventionen festgelegt werden, die im Folgenden verwendet werden.

Die Potenzmenge einer Menge A wird mit $\mathcal{P}(A)$ beschreiben, die Mächtigkeit von A mit $|A|$, also

$$|A| \;=\; \text{Anzahl der Elemente aus } A.$$

Für die Differenz zweier Mengen A und B gilt

$$A \setminus B \;=\; \{x \in A \mid x \notin B\}.$$

Die Vereinigung ist $A \cup B$ und der Durchschnitt $A \cap B$. Weiter beschreibt A^C oder $\overline{A}$ das Komplement $\Omega \setminus A$ und

$$A \vartriangle B \;=\; A \setminus B \,\cup\, B \setminus A.$$

ist die symmetrische Differenz.

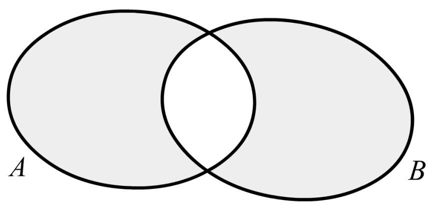


Abbildung 1: Veranschaulichung der symmetrischen Differenz.

Sei $f : M \to N$ eine beliebige Abbildung. Für ein $n \in N$ schreiben wir dann

$$(f = n) \;:=\; \{f = n\} \;:=\; \{m \in M \mid f(m) = n\}.$$

Für $f : \mathbb{N} \to \mathbb{N}$ mit $f(x) = x^2$ gilt also zum Beispiel

$$(f = 4) \;=\; \{f = 4\} \;=\; \{x \in \mathbb{N} \mid f(x) = 4\} \;=\; \{-2, 2\}.$$

1 Diskrete Wahrscheinlichkeitsräume

1.1 Laplace Verteilung

Grundlegende Definitionen

Es sei stets

$$\Omega \;=\; \{\omega_1, \omega_2, \ldots, \omega_n\}$$

eine beliebige endliche Menge, der ***Grundraum***. Die ω_i heißen ***Elementarereignisse*** und sind alle gleich wahrscheinlich. Teilmengen von Ω heißen ***Ereignisse***. Der Grundraum Ω selber heißt das ***sichere Ereignis*** und $\emptyset$ ist das ***unmögliche Ereignis***.

Seien nun A_i beliebige Ereignisse, also $A_i \subset \Omega$. Dann gilt die ***DeMorgansche Regel***

$$\left(\bigcup_{i\in I} A_i\right)^C \;=\; \bigcap_{i\in I} A_i^C \qquad \text{und} \qquad \left(\bigcap_{i\in I} A_i\right)^C \;=\; \bigcup_{i\in I} A_i^C,$$

dabei ist I eine endliche Indexmenge.

Beispiel 1.1.1

Beim Würfeln mit einem fairen Würfel gilt

$$\Omega \;=\; \{1, 2, 3, 4, 5, 6\}.$$

Ein Ereignis A ist also zum Beispiel gerade Augenzahl, also $A = \{2, 4, 6\}$.

Definition 1.1.2

Sei Ω ein Grundraum und sei $\mathcal{P}(\Omega)$ die Potenzmenge von Ω. Die Abbildung

$$\begin{aligned} P : \mathcal{P}(\Omega) &\;\to\; [0,1] \\ A &\;\mapsto\; |A|/|\Omega| \end{aligned}$$

heißt die ***Laplace Verteilung*** oder ***Gleichverteilung*** auf Ω. Es gilt also $P(\{\omega_i\}) = 1/n$ für alle Elementarereignisse ω_i.

(Ω, P) heißt ***Wahrscheinlichkeitsraum***.

Definition 1.1.3

Sei Ω ein Grundraum und sei $\mathcal{P}(\Omega)$ die Potenzmenge von Ω. Eine Abbildung

$$P : \mathcal{P}(\Omega) \to [0,1]$$

heißt ein ***Wahrscheinlichkeitsmaß***, wenn P die folgenden Eigenschaften besitzt:

(1) $P(\Omega) = 1$.

(2) $P(A) \geq 0$.

(3) Aus $A, B \subset \mathcal{P}(\Omega)$ und $A \cap B = \emptyset$ folgt $P(A \cup B) = P(A) + P(B)$.

Die Laplace Verteilung ist also ein mögliches Wahrscheinlichkeitsmaß. Um damit nun wirklich Wahrscheinlichkeiten ausrechnen zu können, wir Kombinatorik benötigt, dazu dienen die folgenden Definitionen.

Der Binomialkoeffizient

Der ***Binomialkoeffizient*** für zwei ganze Zahlen $0 \leq k \leq n$ wird definiert durch

$$\binom{n}{k} \;:=\; \frac{n!}{k! \cdot (n-k)!}.$$

Der Name stammt von der ***Binomischen Formel***

$$(x+y)^n \;=\; \sum_{k=0}^{n} \binom{n}{k} x^k y^{n-k}.$$

Diese Formel ergibt sich, indem man $x_1 = x_2 = .. = x_n = x$ und $y_1 = y_2 = .. = y_n = y$ setzt, man erhält dann

$$(x+y)^n \;=\; (x_1 + y_1)(x_2 + y_2) \ldots (x_n + y_n).$$

Wendet man nun Kombinatorik auf die Anzahl der x und der y beim Ausmultiplizieren an, so erhält man gerade die beschriebene Formel.

Als Folgerungen ergeben sich nun:

(1) Setzt man $x = y = 1$, so ergibt sich sofort

$$\sum_{k=0}^{n} \binom{n}{k} \;=\; 2^n.$$

(2) Setzt man $x = -1$ und $y = 1$, so ergibt sich sofort

$$\sum_{k=0}^{n} (-1)^k \binom{n}{k} \;=\; 0.$$

(3) Leitet man die Binomische Formel ab und setzt dann wieder $x = y = 1$, so ergibt sich

$$\sum_{k=0}^{n} k \binom{n}{k} \;=\; n2^{n-1}.$$

Definition 1.1.4

Sei $\Omega = \{\omega_1, .., \omega_n\}$ ein Grundraum. Dann gilt:

(1) Die Menge der (k, n) ***Permutationen*** mit Wiederholung ist

$$\Omega^k \;:=\; \{(x_1, .., x_k) \mid x_i \in \Omega\}.$$

(2) Die Menge der (k, n) ***Permutationen*** ohne Wiederholung ist

$$\mathrm{Perm}_k^n(\Omega) \;:=\; \{(x_1, .., x_k) \mid x_i \in \Omega,\; x_i \neq x_n \;\forall\; i \neq j\}.$$

(3) Die Menge der (k, n) ***Kombinationen*** mit Wiederholung ist

$$\mathrm{Kom}_k^n(\Omega, \mathrm{m.W.}) \;:=\; \{(\omega_{i_1}, .., \omega_{i_k}) \mid i_1 \leq .. \leq i_k\}.$$

(4) Die Menge der (k, n) ***Kombinationen*** ohne Wiederholung ist

$$\mathrm{Kom}_k^n(\Omega) \;:=\; \{(\omega_{i_1}, .., \omega_{i_k}) \mid i_1 < .. < i_k\} \;=\; \{A \subset \Omega \mid |A| = k\}.$$

Satz 1.1.5

Sei $\Omega = \{\omega_1, .., \omega_n\}$ ein Grundraum. Dann gilt:

$$\begin{aligned}
|\Omega^k| \;&=\; n^k, \\
|\mathrm{Perm}_k^n(\Omega)| \;&=\; (n)_k \;:=\; n \cdot (n-1) \cdot \ldots \cdot (n-k+1), \\
|\mathrm{Kom}_k^n(\Omega, \mathrm{m.W.})| \;&=\; \binom{n+k-1}{k}, \\
|\mathrm{Kom}_k^n(\Omega)| \;&=\; \binom{n}{k}.
\end{aligned}$$

Beweis

Es gilt

$$\Omega^k \;=\; \Omega \times \Omega \times \ldots \times \Omega,$$

damit ergibt sich sofort $|\Omega^k| = n \cdot n \cdot .. \cdot n = n^k$.

Bei den (k, n) Permutationen ohne Wiederholung gibt es für die erste Komponente n Wahlmöglichkeiten, für die zweite $n - 1$, für die dritte $n - 2$

und so weiter. Für die k te Komplonente erhalten wir schließlich $n-k+1$ Möglichkeiten, also gilt

$$|\mathrm{Perm}_k^n(\Omega)| \;=\; n\cdot(n-1)\cdot\ldots\cdot(n-k+1) \;=\; (n)_k.$$

Jede (k,n) Kombination ohne Wiederholung liefert genau $k!$ verschiedene (k,n) Permutationen. Damit ergibt sich

$$\begin{aligned}|\mathrm{Kom}_k^n(\Omega)| \;&=\; \frac{|\mathrm{Perm}_k^n(\Omega)|}{k!} \;=\; \frac{n(n-1)..(n-k+1)}{k!}\cdot\frac{(n-k)!}{(n-k)!}\\ &=\; \frac{n!}{k!\cdot(n-k)!} \;=\; \binom{n}{k}.\end{aligned}$$

Es soll nun von den (k,n) Kombinationen ohne Wiederholung auf die (k,n) Kombinationen mit Wiederholung geschlossen werden. Sei dazu

$$\Omega^* \;:=\; \{\omega_1,\ldots,\omega_n,\,\omega_{n+1},\ldots,\omega_{n+k-1}\},$$

es wurden also $n-1$ künstliche Variablen eingeführt. Damit gibt es eine bijektive Abbildung

$$\begin{aligned}\mathrm{Kom}_k^n(\Omega,\mathrm{m.W.}) \;&\to\; \mathrm{Kom}_k^n(\Omega^*)\\ (\omega_{i_1},..,\omega_{i_k}) \;&\mapsto\; (\omega_{i_1},\omega_{i_1+1},..,\omega_{i_k+k-1})\end{aligned}$$

und es ergibt sich damit wieder

$$|\mathrm{Kom}_k^n(\Omega,\mathrm{m.W.})| \;=\; \binom{n+k-1}{k},$$

was die Behauptung zeigt. □

Bemerkung 1.1.6

Die Menge der (k,n) Kombinationen mit Wiederholung ist isomorph zu einer ***mehrfachen Besetzung***, also

$$\mathrm{Kom}_k^n(\Omega,\mathrm{m.W.}) \;\cong\; D_k^n \;:=\; \left\{(x_1,..,x_n)\;\middle|\; x_i\in\{0,1,..,k\},\ \sum_{i=1}^n x_i = k\right\}.$$

Die Menge der (k,n) Kombinationen ohne Wiederholung ist isomorph zu einer ***einfachen Besetzung***, also

$$\mathrm{Kom}_k^n(\Omega) \;\cong\; C_k^n \;:=\; \left\{(x_1,..,x_n)\;\middle|\; x_i\in\{0,1\},\ \sum_{i=1}^n x_i = k\right\}.$$

C_k^n und D_k^n lassen sich als Besetzungszahlen interpretieren, wenn man k Kugeln auf n Zellen verteilen möchte. Es ergibt sich also sofort wieder

$$|C_k^n| \;=\; \binom{n}{k} \qquad \text{und} \qquad |D_k^n| \;=\; \binom{n+k-1}{k}.$$

Das Zellen- und das Urnenmodell

Die beiden elementaren Anwendungen dieser Kombinatorischen Überlegungen sind das Zellen- und das Urnenmodell:

Das Zellenmodell

Es sollen k Kugeln auf n numerierte Zellen verteilt werden.

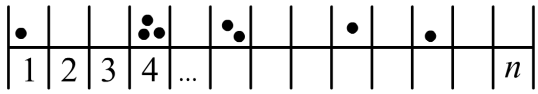


Abbildung 1.1: Das Zellenmodell.

(1) Es gibt n^k Möglichkeiten, um k unterscheidbare Kugeln mit Mehrfachbesetzung auf n Zellen zu verteilen.

(2) Es gibt $(n)_k$ Möglichkeiten, um k unterscheidbare Kugeln ohne Mehrfachbesetzung auf n Zellen zu verteilen.

(3) Es gibt $\binom{n+k-1}{k}$ Möglichkeiten, um k ununterscheidbare Kugeln mit Mehrfachbesetzung auf n Zellen zu verteilen.

(4) Es gibt $\binom{n}{k}$ Möglichkeiten, um k ununterscheidbare Kugeln ohne Mehrfachbesetzung auf n Zellen zu verteilen.

Das Urnenmodell

Aus einer Urne mit n numerierten Kugeln sollen nun k Kugeln gezogen werden.

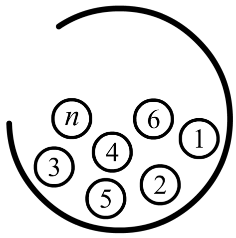


Abbildung 1.2: Das Urnenmodell.

(1) Es gibt n^k Möglichkeiten, um k Kugeln in Reihenfolge mit zurücklegen zu ziehen.

(2) Es gibt $(n)_k$ Möglichkeiten, um k Kugeln in Reihenfolge ohne zurücklegen zu ziehen.

(3) Es gibt $\binom{n+k-1}{k}$ Möglichkeiten, um k Kugeln one Reihenfolge mit zurücklegen zu ziehen.

(4) Es gibt $\binom{n}{k}$ Möglichkeiten, um k Kugeln ohne Reihenfolge ohne zurücklegen zu ziehen.

Beispiel 6 aus 49

Um die Anzahl der Möglichkeiten beim Spiel 6 aus 49 zu berechnen, wenden wir das Urnenmodell mit einer Urne mit 49 numerierten Kugeln an, aus welcher 6 Kugeln ohne Reihenfolge ohne zurücklegen gezogen werden. Demnach gibt es genau

$$\binom{49}{6} = 13\,983\,816$$

Möglichkeiten.

Beispiel Sehtest

Bei einem Sehtest werden 3 Karten vorgelegt, die jeweils eine der Ziffern aus $\Omega = \{0, 1, .., 9\}$ im Rot-Grün-Kontrast zeigen. Es sollen nun berechnet werden, wie groß die Wahrscheinlichkeit ist, dass ein farbenblinder Proband mindestens zwei richtige Antworten nennt. Dazu kann die Laplace Verteilung genutzt werden, da der Proband beliebig raten wird.

(1) Die Karten werden nacheinander gezogen und danach wieder zurückgelegt. Dies entspricht dem Urnenmodell, indem 3 Kugeln in Reihenfolge mit zurücklegen gezogen werden. Wir haben

$$\Omega_1 = \{0, 1, \ldots, 9\}^3, \qquad \text{also} \qquad |\Omega_1| = 1000.$$

Es seien (a_1, a_2, a_3) die gezeigten Karten. Insgeamt wählt der Proband nun zufällig ein Elemente aus Ω_1 aus. Es sei nun A_1 die Menge der Möglichkeiten, bei der mindestens zwei richtige Antworten gegeben wurden. Es gilt

$$\begin{aligned} A_1 = {} & \{(a_1, a_2, a_3)\} \cup \{(a_1, a_2, \omega_3) \mid \omega_3 \neq a_3\} \cup \\ & \{(a_1, \omega_2, a_1) \mid \omega_2 \neq a_2\} \cup \{(\omega_1, a_2, a_1) \mid \omega_1 \neq a_1\}. \end{aligned}$$

Es folgt $|A_1| = 1+9+9+9 = 28$, somit wurden mit einer Wahrscheinlichkeit von

$$P(A_1) = \frac{|A_1|}{|\Omega_1|} = \frac{28}{1000} = 2.8\,\%$$

mindestens zwei richtige Antworten erraten.

(2) Die Karten werden wieder nacheinander gezogen, aber danach nicht zurückgelegt. Damit haben wir

$$\Omega_2 \;=\; \mathrm{Perm}_3^{10}(\Omega), \qquad \text{also} \qquad |\Omega_2| \;=\; 10 \cdot 9 \cdot 8 \;=\; 720.$$

Die Menge A_2 der Möglichkeiten, bei der mindestens zwei richtige Antworten gegeben wurden, ist nun

$$\begin{aligned} A_2 \;=\;& \{(a_1, a_2, a_3)\} \cup \{(a_1, a_2, \omega_3) \mid \omega_3 \neq a_i,\; i = 1, .., 3\} \cup \\ & \{(a_1, \omega_2, a_1) \mid \omega_2 \neq a_i,\; i = 1, .., 3\} \cup \\ & \{(\omega_1, a_2, a_1) \mid \omega_1 \neq a_i,\; i = 1, .., 3\}. \end{aligned}$$

Damit folgt $|A_2| = 1 + 7 + 7 + 7 = 22$, also

$$P(A_2) \;=\; \frac{|A_2|}{|\Omega_2|} \;=\; \frac{22}{720} \;\approx\; 3.06\,\%.$$

(3) Die drei Karten werden nun gleichzeitig gezeigt, der Proband nennt also nur 3 Ziffern ohne Zuordnung. Damit ergibt sich

$$\Omega_3 \;=\; \mathrm{Kom}_3^{10}(\Omega), \qquad \text{also} \qquad |\Omega_3| \;=\; \binom{10}{3} \;=\; 120.$$

Die Menge A_3 der Möglichkeiten, bei der mindestens zwei richtige Antworten gegeben wurden, ist nun

$$A_3 \;=\; \{(a_1, a_2, a_3)\} \cup \{(a_1, a_2, \omega_3) \mid \omega_3 \neq a_i,\; i = 1, .., 3\},$$

also $|A_3| = 1 + 7$. Wir erhalten somit eine Wahrscheinlichkeit von

$$P(A_3) \;=\; \frac{|A_3|}{|\Omega_3|} \;=\; \frac{8}{120} \;=\; \frac{1}{15} \;\approx\; 6.66\,\%.$$

Beispiel Hashing

Beim Hashing sollen k Daten zufällig auf n Felder verteilt werden, dabei gilt $k \leq n$ und es sind Kollisionen möglich, das heißt es können mehrere Daten in ein Feld abgelegt werden. Es sei nun $A_{k,n}$ die Menge aller Kollisionen und

$$\Omega \;=\; \{(x_1, .., x_k) \mid 1 \leq x_i \leq n,\; i = 1, .., k\}.$$

Betachten wir die Menge ohne Kollisionen, so erhalten wir

$$A_{k,n}^C \;=\; \mathrm{Perm}_k^n(\Omega).$$

Die Wahrscheinlichkeit für keine Kollision ist also

$$P(A_{k,n}^C) \;=\; \frac{|\mathrm{Perm}_k^n(\Omega)|}{|\Omega|} \;=\; \frac{(n)_k}{n^k}.$$

Draus ergibt sich nun die Wahrscheinlichkeit für eine Kollision:

$$P(A_{k,n}) \;=\; 1 - P(A_{k,n}^C) \;=\; 1 - \frac{(n)_k}{n^k}.$$

Es soll nun noch untersucht werden, wie k zu wählen ist, um bei einer vorgegebenen Wahrscheinlichkeit $p \in]0,1[$ keine Kollision zu haben. Dazu wird die folgende Approximation durchgeführt:

$$\begin{aligned} P(A_{k,n}^C) &= \prod_{i=0}^{k-1} \frac{n-i}{n} = \prod_{i=0}^{k-1} 1 - \frac{i}{n} = \exp\left(\sum_{i=0}^{k-1} \log\left(1 - \frac{i}{n}\right)\right) \\ &\leq \exp\left(-\sum_{i=0}^{k-1} \frac{i}{n}\right) = \exp\left(\frac{-k(k-1)}{2n}\right). \end{aligned}$$

Die Ungleichung kommt daher zustande, dass $\log(1-x) \leq -x$ für jedes $x \in]0,1[$ gilt. Wir haben nun

$$p \;=\; P(A_{k,n}^C) \;\approx\; \exp\left(-\frac{k^2}{2n}\right).$$

Wählt man also

$$k \;\approx\; \sqrt{|\log(p)| \cdot 2n}\,,$$

so kommt es mit einer Wahrscheinlichkeit von p zu keiner Kollision.

Beispiel Qualitätskontrolle

Bei einer Lieferung von n Waren werden zur Kontrolle zufällig eine Probe von k herausgenommen. In der Lieferung sind $s \leq n$ defekte Waren enthalten.

Es soll nun untersucht werden, wie groß die Wahrscheinichkeit dafür ist, dass in der Probe r defekte Waren vorgefunden werden.

Sei nun

$$\Omega \;=\; \{1, \ldots, s, s+1, \ldots, n\},$$

dabei bezeichnet $D = \{1, .., s\}$ die s defekten und $I = \{s+1, .., n\}$ die $n-s$ intakten Waren. Sei nun weiter

$$\Omega_1 \;=\; \mathrm{Kom}_k^n(\Omega).$$

Die Menge A_r der Möglichen Kombinationen für r defekte Waren ergibt sich nun aus

$$\begin{aligned} A_r &= \{A \in \Omega_1 \mid |A \cap D| = r\} \\ &= \{A \in \Omega_1 \mid |A \cap D| = r,\ |A \cap I| = k - r\} \end{aligned}$$

Damit ergibt sich die folgende Wahrscheinlichkeit für r defekte Waren in den k Probewaren:

$$P(A_r) \;=\; \frac{\binom{s}{r}\cdot\binom{n-s}{k-r}}{\binom{n}{k}}.$$

Qualitätskontrolle als 6 aus 49

Auch mit der Qualitätskontrolle lässt sich 6 aus 49 Problem lösen:

Wir haben $n = 49$ Kugeln. Von denen werden $k = 6$ gezogen und es gibt auch genau $s = 6$ defekte, also in diesem Falle richtige, Kugeln. Um die Wahrscheinlichkeit für die 6 richtigen zu erhalten, müssen also auch die $r = 6$ richtigen Kugeln in der Ziehung enthalten sein.

Damit ergibt sich

$$P(A_r) \;=\; \frac{\binom{6}{6}\cdot\binom{43}{0}}{\binom{49}{6}} \;=\; \frac{1}{\binom{49}{6}}.$$

Der folgende Satz soll nun die Berechnung der Wahrscheinlichkeiten von Vereinigungen von Mengen vereinfachen.

Satz 1.1.7 (Siebformel)

Sei P eine Laplace Verteilung auf $\Omega = \{\omega_1, .., \omega_r\}$ und seien $A_1, .., A_n \subset \Omega$.

Dann gilt

$$\begin{aligned} P\left(\bigcup_{i=1}^{n} A_i\right) &= \sum_{k=1}^{n}(-1)^{k+1}\cdot \sum_{\substack{I\subset\{1,..,n\}\\ |I|=k}} P\left(\bigcap_{i\in I} A_i\right) \\ &= \sum_{i=1}^{n} P(A_i) \;-\; \sum_{i<j}^{n} P(A_i\cap A_j) \;+ \\ &\qquad \sum_{i<j<k}^{n} P(A_i\cap A_j\cap A_k) \;-\; \ldots \;+\; (-1)^{n-1}P(A_1\cap ..\cap A_n). \end{aligned}$$

Beweis

Es reicht zu zeigen, dass die Mächtigkeit der beiden Mengen gleich ist, dazu betrachten wir alle $\omega \in \Omega$.

Gilt $\omega \notin \bigcup_{i=1}^{n} A_i$, so erhalten wir auf der linken wie auf der rechten Seite der Gleichung einen Beitrag von 0. Gilt $\omega \in \bigcup_{i=1}^{n} A_i$, so ergibt dies auf der linken Seite einen Beitrag von 1. Es bleibt also zu zeigen, dass wir dann auch auf der rechten Seite einen Beitrag von 1 erhalten.

Seien dazu $A_{i_1}, .., A_{i_k}$ genau diejenigen Mengen, in denen ω enthalten ist.

Damit ergibt sich

$$\begin{aligned}
& \sum_{i=1}^{n} |P(A_i)| \;-\; \sum_{\substack{i,j\\ i\neq j}}^{n} |P(A_i\cap A_j)| \;+\; \ldots \;+\; (-1)^{n-1}|P(A_1\cap .. \cap A_n)| \\
= \;& \binom{k}{1} - \binom{k}{2} + \binom{k}{3} - \ldots + (-1)^{k-1}\binom{k}{k} \\
= \;& -\sum_{i=1}^{k}(-1)^i\binom{k}{i} \;=\; 1 - \sum_{i=0}^{k}(-1)^i\binom{k}{i} \;=\; 1,
\end{aligned}$$

dabei wurde die Folgerung **(2)** der Binomischen Formel von Seite 5 verwendet. □

Beispiel Rencontre Problem

Sei

$$\Omega \;=\; \{\pi \mid \pi : \{1,..,n\} \to \{1,..,n\}\} \;=\; \mathrm{Perm}_n^n(\{1,..,n\}).$$

(Ω, P) sei nun ein Laplace Experiment der zufälligen Anordnung. Es soll untersucht werden, wie groß die Wahrscheinlichkeit einer fixpunktfreien Permutatin ist, also

$$A \;=\; \{\pi \in \Omega \mid \pi(i) \neq i \text{ für } i = 1,..,n\}.$$

Es ist einfacher mit dem Komplement zu rechnen. Sei dazu

$$B_i \;=\; \{\pi \in \Omega \mid \pi(i) = i\}, \qquad \text{also} \qquad A^C \;=\; \bigcup_{i=1}^{n} B_i.$$

Es gilt $|\Omega| = n!$ und $|B_i| = (n-1)!$, also

$$P(B_i) \;=\; \frac{1}{n}.$$

Für $B_i \cap B_j$ gilt $|B_i \cap B_j| = (n-2)!$, also

$$P(B_i \cap B_j) \;=\; \frac{1}{n(n-1)}$$

und so weiter. Damit kann nun die Siebformel angewendet werden:

$$\begin{aligned}
P(A) \;&=\; 1 - P(A^C) \;=\; 1 - P\left(\bigcup_{i=1}^{n} B_i\right) \\
&=\; 1 - \sum_{i=1}^{n} P(B_i) - \sum_{\substack{i,j\\ i\neq j}}^{n} P(B_i \cap B_j) + \ldots + (-1)^{n-1} P(A_1 \cap .. \cap A_n) \\
&=\; 1 - n\cdot\frac{1}{n} + \binom{n}{2}\frac{1}{n(n-1)} - \ldots + (-1)^n\binom{n}{n}\frac{1}{n!} \\
&=\; \frac{1}{2!} - \ldots + (-1)^n\frac{1}{n!} \;=\; \sum_{i=0}^{n}(-1)^i\frac{1}{i!} \;\stackrel{n\to\infty}{\longrightarrow}\; e^{-1}.
\end{aligned}$$

Die folgende Abschätzung dient dazu, wie genau e^{-1} die Wahrscheinlichkeit $P(A)$ approximiert:

$$\left|P(A) - \frac{1}{e}\right| \;=\; \left|\sum_{i=n+1}^{\infty} (-1)^i \frac{1}{i!}\right| \;\leq\; \frac{1}{(n+1)!},$$

für $n = 4$ erhält man also eine Abweichung von maximal $0.84\,\%$.

Mit dieser Methode lässt sich die Eulersche Zahl e also auch recht gut experimentell bestimmen.

1.2 Diskrete Wahrscheinlichkeitsräume

Der Begriff des Wahrscheinlichkeismaßes soll nun auch auf abzählbare Mengen Ω verallgemeinert werden, dazu zunächst die folgende grundlegende Definition.

Definition 1.2.1

Sei Ω ein nicht leerer Grundraum.

Eine Abbildung $f : \Omega \to [0,1]$ heißt ***Zähldichte*** oder ***Zählfunktion*** auf Ω, wenn es eine abzählbare Menge $\Omega_0 \subset \Omega$ gibt, für die gilt:

(1) Es ist $\{f \neq 0\} := \{\omega \in \Omega \mid f(\omega) \neq 0\} \subset \Omega_0$.

(2) Es ist $\sum_{\omega\in\Omega_0} f(\omega) = 1$.

Eine Abbildung $P : \mathcal{P}(\Omega) \to [0,1]$ heißt ***diskretes Wahrscheinlichkeitsmaß*** auf Ω, wenn es eine Zähldichte f auf Ω gibt, so dass für alle $A \subset \Omega$ gilt:

$$P(A) \;=\; \sum_{\omega\in A} f(\omega) \;=\; \sum_{\omega\in A\cap\Omega_0} f(\omega).$$

(Ω, P) heißt dann ***diskreter Wahrscheinlichkeitsraum***.

Bemerkung

Zu einer Zähldichte gibt es immer genau einen diskreten Wahrscheinlichkeitsraum und umgekehrt.

Bernoulli Verteilung

Sei $\Omega = \{0,1\}$ und sei $\vartheta \in [0,1]$.

Durch $f(0) = 1 - \vartheta$ und $f(1) = \vartheta$ wird dann eine Zähldichte f auf Ω definiert. Das zugehörige diskrete Wahrscheinlichkeitsmaß ist dann die ***Bernoulli Verteilung*** mit dem Parameter ϑ.

Schreibweise: $\mathcal{L}(1, \vartheta)$.

Binomialverteilung

Sei $\Omega = \{0, 1, .., n\}$ und sei $\vartheta \in [0, 1]$.

Dann wird durch

$$f(i) \;=\; \binom{n}{i}\vartheta^i(1-\vartheta)^{n-i} \qquad \text{für} \qquad i \in \Omega$$

eine Zähldichte f definiert, denn nach der Binomischen Formel gilt

$$\sum_{i=0}^{n} f(i) \;=\; \sum_{i=0}^{n}\binom{n}{i}\vartheta^i(1-\vartheta)^{n-i} \;=\; (\vartheta + 1 - \vartheta)^n \;=\; 1.$$

Das zugehörige diskrete Wahrscheinlichkeitsmaß ist die ***Binomialverteilung*** mit den Parametern n und ϑ.

Schreibweise: $\mathcal{L}(n, \vartheta)$.

$\mathcal{L}(n, \vartheta)$ liefert also gerade die Anzahl der Erfolge bei n unabhängigen Experimenten mit der Erfolgswahrscheinlichkeit von ϑ für jedes Experiment.

Münzwurf

Bei einem vierfachen Münzwurf ist also $\Omega = \{0, 1, 2, 3, 4\}$ und $\vartheta = \frac{1}{2}$. Damit ergibt sich

$$\begin{aligned} f(0) = f(4) \;&=\; 1 \cdot \left(\frac{1}{2}\right)^4 \;=\; 0.0625\,, \\ f(1) = f(3) \;&=\; 4 \cdot \left(\frac{1}{2}\right)^4 \;=\; 0.25\,, \\ f(2) \;&=\; 6 \cdot \left(\frac{1}{2}\right)^4 \;=\; 0.375\,. \end{aligned}$$

Hypergeometrische Verteilung

Sei $P = H(n, s, k)$ ein diskretes Wahrscheinlichkeitsmaß wie im Beispiel der Qualitätskontrolle auf Seite 11. Dann wird durch

$$f(r) \;=\; \frac{\binom{s}{r}\cdot\binom{n-s}{k-r}}{\binom{n}{k}}$$

eine Zähldichte definiert, dabei gilt

$$\max\{0, k+s-n\} \;\le\; r \;\le\; \min\{k, s\}.$$

Poisson Verteilung

Sei $\Omega = \mathbb{N} \cup \{0\}$ und sei $\lambda > 0$.

Dann wird durch

$$f(i) \;=\; \frac{\lambda^i}{i!} e^{-\lambda} \qquad \text{für} \qquad i \in \Omega$$

eine Zähldichte f definiert, denn es gilt

$$\sum_{i=0}^{\infty} f(i) \;=\; \sum_{i=0}^{\infty} \frac{\lambda^i}{i!} e^{-\lambda} \;=\; e^{\lambda} e^{-\lambda} \;=\; 1.$$

Das zugehörige diskrete Wahrscheinlichkeitsmaß ist die ***Poisson Verteilung*** $\mathcal{P}(\lambda)$.

Seltene Ereignis Versuch

Beim seltene Ereignis Versuch von Rutherford und Geiger werden die Anzahl der α Teilchen in gewissen Zeitabständen beim radioaktiven Zerfall gemessen.

Ein Zeitabschnitt sei 7.5 Sekunden lang und n_i sei die Anzahl der Zeitabschnitte mit i zerfallenen α Teilchen. Experimentell wurde mit $n = 2608$ gemessen:

i	0	1	2	3	4	5	6	7	8	9	10
n_i	57	203	383	525	532	408	273	139	45	27	16

Tabelle 1.1: Anzahl der zerfallenen α Teilchen.

Dabei gilt nun

$$f_\lambda(i) \;=\; \frac{n_i}{n} \qquad \text{mit} \qquad \lambda \;=\; \frac{\sum_{i=0}^{10} i n_i}{n} \;\approx\; 3.87\,.$$

Der Wert λ gibt also gerade die durchschnittliche Anzahl der pro Zeitintervall zerfallenen α Teilchen an.

Satz 1.2.2 über seltene Ereignisse

Sei $(\vartheta_n)_{n\in\mathbb{N}}$ eine Folge von reellen Zahlen mit $\vartheta \in]0,1[$ für alle $n \in \mathbb{N}$ und mit

$$\lim_{n\to\infty} n\vartheta_n \;=\; \lambda \;>\; 0.$$

Weiter seien f_n die Zähldichten der Binomialverteilungen $P_n = \mathcal{L}(n, \vartheta_n)$, die auf $\Omega = \mathbb{N} \cup \{0\}$ definiert sind, und es sei f die Zähldichte der Poisson Verteilung $P = \mathcal{P}(\lambda)$, auch auf $\Omega = \mathbb{N} \cup \{0\}$ definiert.

Dann gilt

$$\lim_{n\to\infty} f_n(i) \;=\; f(i) \qquad \text{für alle } \; i \in \Omega = \mathbb{N} \cup \{0\}.$$

Bemerkung

Dieser Satz besagt also: Für n unabhängige Experimente mit einer Erfolgswahrscheinlichkeit von $\vartheta = \frac{\lambda}{n}$ ist die Anzahl der Erfolge gleich $\mathcal{P}(\lambda)$.

Mit Satz 1.2.2 wurde nun punktweise Konvergenz gezeigt, es gilt jedoch auch gleichmäßige Konvergenz:

Satz 1.2.3

Sei $(\vartheta_n)_{n\in\mathbb{N}}$ eine Folge von reellen Zahlen mit $\vartheta \in]0,1[$ für alle $n \in \mathbb{N}$ und mit

$$\lim_{n\to\infty} n\vartheta_n \;=\; \lambda \;>\; 0.$$

Weiter seien $P_n = \mathcal{L}(n, \vartheta_n)$ die Binomialverteilungen und $P = \mathcal{P}(\lambda)$ die Poisson Verteilung.

Dann gilt

$$\sup_{A\subset\mathbb{N}\cup\{0\}} |P_n(A) - P(A)| \;\overset{n\to\infty}{\longrightarrow}\; 0.$$

Dieser Satz folgt aus 1.2.2 und aus dem folgenden Lemma:

Lemma von Scheffé

Seien f_n und f Zähldichten auf Ω und Ω_0 ein abzählbarer Träger.

Wenn $f_n(\omega) \to f(\omega)$ für $n \to \infty$ für alle $\omega \in \Omega_0$, dann gilt auch

$$\sum_{\omega\in\Omega_0} |f_n(\omega) - f(\omega)| \;\overset{n\to\infty}{\longrightarrow}\; 0.$$

Satz 1.2.4

Sei P ein diskretes Wahrscheinlichkeitsmaß auf Ω. Dann gilt:

(1) $P(A^C) = 1 - P(A)$ für alle $A \subset \Omega$.

(2) Für $A \subset B$ folgt $P(A) \leq P(B)$. Diese Eigenschaft heißt ***Isotonie***.

(3) P ist σ-***additiv***, das heißt für paarweise disjunkte $(A_i)_{i\in\mathbb{N}} \subset \Omega$ gilt

$$P\left(\bigcup_{i=1}^{\infty} A_i\right) \;=\; \sum_{i=1}^{\infty} P(A_i).$$

(4) Es gilt $P(\Omega) = 1$.

Beweis

Es ist $P(\Omega) = \sum_{\omega\in\Omega} f(\omega) = 1$, somit wurde **(4)** gezeigt. Weiter ist

$$1 \;=\; P(\Omega) \;=\; \sum_{\omega\in\Omega} f(\omega) = \sum_{\omega\in A} f(\omega) + \sum_{\omega\in A^C} f(\omega) \;=\; P(A) + P(A^C)$$

und **(1)** ist damit klar. Sei nun $A \subset B$ und $B = A \cup (B \setminus A)$. Damit folgt

$$P(B) \;=\; \sum_{\omega\in B} f(\omega) \;=\; \sum_{\omega\in A} f(\omega) + \sum_{\omega\in B\setminus A} f(\omega) \;=\; P(A) + P(B\setminus A) \;\geq\; P(A),$$

also gilt auch **(2)**. Nach dem Umordnungssatz folgt weiter

$$P\left(\bigcup_{i=1}^{\infty} A_i\right) \;=\; \sum_{\omega\in\bigcup_{i=1}^{\infty} A_i} f(\omega) \;=\; \sum_{i=1}^{\infty}\left(\sum_{\omega\in A_i} f(\omega)\right) \;=\; \sum_{i=1}^{\infty} P(A_i)$$

und auch **(3)** wurde gezeigt. □

Definition 1.2.5

Sei Ω ein Grundraum, sei $\mathcal{P}(\Omega)$ die Potenzmenge von Ω und sei

$$P : \mathcal{P}(\Omega) \to [0, 1].$$

P heißt genau dann ***Wahrscheinlichkeitsinhalt*** auf $\mathcal{P}(\Omega)$, wenn P additiv ist, wenn also für disjunkte Mengen $A \subset \Omega$ gerade $P(A \cup B) = P(A) + P(B)$ gilt.

P heißt genau dann ***Wahrscheinlichkeitsmaß*** auf $\mathcal{P}(\Omega)$, wenn P σ-additiv ist, wenn also für paarweise disjunkte $(A_i)_{i\in\mathbb{N}} \subset \Omega$

$$P\left(\bigcup_{i=1}^{\infty} A_i\right) \;=\; \sum_{i=1}^{\infty} P(A_i)$$

gilt.

Bemerkung

Damit gilt nun:

Ist P ein diskretes Wahrscheinlichkeitsmaß auf Ω, dann ist P genau dann ein Wahrscheinlichkeitsmaß auf $\mathcal{P}(\Omega)$, wenn P σ-additiv ist.

Ist P ein Wahrscheinlichkeitsmaß auf $\mathcal{P}(\Omega)$, dann ist P genau dann ein diskretes Wahrscheinlichkeitsmaß auf Ω, wenn P ***diskret*** ist, wenn es also eine abzählbare Menge $\Omega_0 \subset \Omega$ gibt mit $P(\Omega_0) = 1$.

Weiter soll nun untersucht werden, was Wahrscheinlichkeitsmaße von Wahrscheinlichkeitsinhalte unterscheidet.

Satz 1.2.6

Sei $P : \mathcal{P}(\Omega) \to [0,1]$ ein Wahrscheinlichkeitsmaß auf $\mathcal{P}(\Omega)$. Dann gilt

(1) Für $A \subset B$ folgt $P(A) \leq P(B)$.

(2) Für $(A_i)_{1 \leq i \leq n} \in \mathcal{P}(\Omega)$ folgt

$$P\left(\bigcup_{i=1}^{n} A_i\right) \;\leq\; \sum_{i=1}^{n} P(A_i).$$

(3) Für paarweise disjunkte $(A_i)_{1 \leq i \leq \infty} \in \mathcal{P}(\Omega)$ folgt

$$P\left(\bigcup_{i=1}^{\infty} A_i\right) \;\geq\; \sum_{i=1}^{\infty} P(A_i).$$

P heißt damit *σ-**superadditiv***.

Satz 1.2.7

Sei $P : \mathcal{P}(\Omega) \to [0,1]$ ein Wahrscheinlichkeitsinhalt auf $\mathcal{P}(\Omega)$. Dann sind folgende Aussagen äquivalent:

(1) P ist σ-additiv, also ein Wahrscheinlichkeitsmaß auf $\mathcal{P}(\Omega)$.

(2) P ist stetig von oben, das heißt es gilt für $A_n \downarrow A$ gerade $P(A_n) \downarrow P(A)$.

(3) P ist stetig von unten, das heißt es gilt für $A_n \uparrow A$ gerade $P(A_n) \uparrow P(A)$.

(4) P ist stetig in $\emptyset$, das heißt es gilt für $A_n \downarrow \emptyset$ gerade $P(A_n) \downarrow 0$.

(5) P ist σ-subadditiv, das heißt

$$P\left(\bigcup_{i=1}^{\infty} A_i\right) \;\leq\; \sum_{i=1}^{\infty} P(A_i).$$

Ist P zusätzlich noch diskret, dann gelten auch noch die folgenden zusätzlichen Äquivalenzen:

(6) Es gilt $\sum_{\omega \in \Omega} P(\{\omega\}) = 1$.

(7) Es gilt $\sum_{\omega \in A} P(\{\omega\}) = P(A)$ für alle $A \subset \Omega$, P ist also ein diskretes Wahrscheinlichkeitsmaß auf Ω.

Beispiel Dirac Maß

Sei $\omega \in \Omega$ fest und sei

$$f(\omega') \;=\; \begin{cases} 1 & \text{wenn} \quad \omega' = \omega \\ 0 & \text{wenn} \quad \omega' \neq \omega \end{cases} .$$

Dann ist f offensichtliche eine Zähldichte und das zugehörige Wahrscheinlichkeitsmaß heißt ***Dirac Maß***:

$$\delta_\omega(A) \;=\; \sum_{\omega' \in A} f(\omega') \;=\; 1_A(\omega)$$

mit der ***Indikator***– oder ***charakteristischen Funktion***

$$1_A(\omega) \;=\; \begin{cases} 1 & \text{wenn} \quad \omega \in A \\ 0 & \text{wenn} \quad \omega \notin A \end{cases} .$$

Satz 1.2.8 (Renyi Prinzip)

Sei P ein diskretes Wahrscheinlichkeitsmaß mit der Zähldichte f.

Dann gilt

$$P(A) \;=\; \sum_{\omega \in \Omega} f(\omega)\delta_\omega(A).$$

P ist dabei σ-additiv, also konvex für unendlich viele Summanden.

Seien weiter $A_i \subset \Omega$ und $a_i \in \mathbb{R}$ für $1 \leq i \leq n$ sowie $c \in \mathbb{R}$ so gewählt, dass

$$\sum_{i=1}^{n} a_i A_{A_i}(\omega) \;\geq\; c \qquad \text{für alle} \qquad \omega \in \Omega.$$

Dann gilt für alle diskrete Wahrscheinlichkeitsmaße P

$$\sum_{i=1}^{n} a_i P(A_i) \;\geq\; c.$$

Der zweite Teil dieser Satzes gilt auch für Gleichheit, also für $= c$.

Beweis

Der erste Teil ergibt sich nach

$$P(A) \;=\; \sum_{\omega \in A} f(\omega) \;=\; \sum_{\omega \in \Omega} f(\omega) 1_A(\omega) \;=\; \sum_{\omega \in \Omega} f(\omega)\delta_\omega(A),$$

der zweite Teil folt nach den Voraussetzungen aus

$$\begin{aligned} \sum_{i=1}^{n} a_i P(A_i) \;&=\; \sum_{i=1}^{n} a_i \sum_{\omega \in \Omega} f(\omega)\delta_\omega(A_i) \;=\; \sum_{i=1}^{n} f(\omega) \sum_{\omega \in \Omega} a_i \delta_\omega(A_i) \\ &\leq\; \sum_{\omega \in \Omega} f(\omega) \cdot c \;=\; 0, \end{aligned}$$

womit der Satz gezeigt ist. $\square$

Satz 1.2.9 (Siebformel für diskrete Wahrscheinlichkeitsmaße)

Sei P ein diskretes Wahrscheinlichkeitsmaß auf Ω und seien $A_1, .., A_n \subset \Omega$. Dann gilt ganz analog zur Siebformel für Laplace Verteilungen

$$\begin{aligned} P\left(\bigcup_{i=1}^{n} A_i\right) &= \sum_{k=1}^{n} (-1)^{k+1} \cdot \sum_{\substack{I \subset \{1,..,n\} \\ |I|=k}} P\left(\bigcap_{i \in I} A_i\right) \\ &= \sum_{i=1}^{n} P(A_i) - \sum_{i<j}^{n} P(A_i \cap A_j) + \\ &\quad \sum_{i<j<k}^{n} P(A_i \cap A_j \cap A_k) - \ldots + (-1)^{n-1} P(A_1 \cap .. \cap A_n). \end{aligned}$$

Beispiel asymptotische Dichte

Sei $\Omega = \mathbb{N}$ und sei

$$\gamma := \left\{ S \subset \mathbb{N} \;\middle|\; \text{der Grenzwert } \lim_{n \to \infty} \frac{|S \cap \{1, .., n\}|}{n} \text{ existiert} \right\}.$$

Nach dieser Definition gilt für γ:

(1) Mit $A \in \gamma$ folgt $A^C \in \gamma$.

(2) Für disjunkten $A, B \in \gamma$ folgt $A \cup B \in \gamma$.

(3) Mit $A \in \gamma$ und $k \in \mathbb{N}$ folgt $A + k = \{a + k \mid a \in A\} \in \gamma$.

Sei weiter $P(S)$ gerade der Grenzwert, also

$$P(S) = \lim_{n \to \infty} \frac{|S \cap \{1, .., n\}|}{n}.$$

Dann liegen die folgenden Mengen in γ, der Grenzwert existiert also:

(1) Für $S \subset \mathbb{N}$ mit $|S| < \infty$ gilt $P(S) = 0$.

(2) Für $S \subset \mathbb{N}$ mit $|S^C| < \infty$ gilt $P(S) = 1$.

(3) Für $A_k = \{n \in \mathbb{N} \mid k \text{ teilt } n\}$ gilt $P(A_k) = 1/k$.

(4) Für $A_0 = \{p \in \mathbb{N} \mid p \text{ ist Primzahl}\}$ folgt nach dem Primzahlsatz von Gauß $P(A_0) = 0$.

(5) Für n paarweise unterschiedliche Primzahlen p_i gilt nach **(3)** für

$$A_{\prod_{i=1}^{k} A_{p_i}} = \bigcap_{i=1}^{k} A_{p_i} \qquad \text{gerade} \qquad P\left(\bigcap_{i=1}^{k} A_{p_i}\right) = \prod_{i=1}^{n} P(A_{p_i}).$$

(6) Für $C_k = \{m \in \mathbb{N} \mid m \text{ und } k \text{ sind teilerfremd}\}$ mit $k = \prod_{i=1}^n p_i^{s_i}$ folgt

$$P(C_k) \;=\; \prod_{i=1}^{n} \left(1 - \frac{1}{p_i}\right).$$

Definition 1.2.10

Seien (Ω_1, P_1) und (Ω_2, P_2) diskrete Wahrscheinlichkeitsräume mit den Zähldichten f_1 und f_2. Sei nun $\Omega = \Omega_1 \times \Omega_2$ und $f(\omega_1, \omega_2) = f_1(\omega_1) f_2(\omega_2)$ für $(\omega_1, \omega_2) \in \Omega$.

Dann ist f eine Zähldichte auf Ω und das zughörige diskrete Wahrscheinlichkeits $P = P_1 \otimes P_2$ heißt ***Produktmaß*** von P_1 und P_2.

Bemerkungen

(1) Sei $\Omega_i' \subset \Omega_i$ ein abzählbarer Träger von P_i für $i = 1, 2$. Dann ist auch $\Omega_1' \times \Omega_2' \subset \Omega$ ein abzählbarer Träger von $P_1 \otimes P_2$.

(2) Analog kann die Situation auch für mehr als zwei Produkte durchgeführt werden.

(3) Derartige Produktexperimente benötigt man zur Modellierung von Versuchswiederholungen.

Proposition 1.2.11

Für zwei diskrete Wahrscheinlichkeitsräume (Ω_1, P_1) und (Ω_2, P_2) ist $P = P_1 \otimes P_2$ das eindeutig bestimmt Produktmaß auf $\Omega = \Omega_1 \times \Omega_2$ mit

$$P(A_1 \times A_2) \;=\; P_1(A_1) \cdot P_2(A_2)$$

für alle $A_1 \subset \Omega_1$ und $A_2 \subset \Omega_2$.

1.3 Bedingte Wahrscheinlichkeiten

Definition 1.3.1

Sei (Ω, P) ein disktreter Wahrscheinlichkeitsraum und sei $B \subset \Omega$.

Für $P(B) > 0$ heißt dann

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

für alle $A \subset \Omega$ die ***bedingte Wahrscheinlichkeit*** von A unter B. Falls $P(B) = 0$, dann gelte $P(A|B) := P(A)$.

Bemerkung

Nach dieser Definition bedeutet $P(A|B)$ die Wahrscheinlichkeit dafür, dass A eintritt, unter der Bedingung, dass auch die Vorinformation B eintritt.

Beispiel Familie

Wir betrachten eine Familie mit zwei Kindern, mindestens ein Kind ist ein Mädchen. Es soll nun geklärt werden, mit welcher Wahrscheinlichkeit auch das zweite Kind ein Mädchen ist.

Es stehe J für Junge und M für Mädchen. Weiter sei

$$\Omega \;=\; \{J, M\}^2 \;=\; \{(J,J),\,(J,M),\,(M,J),\,(M,M)\}$$

und wir betrachten die Laplace Verteilung P. Wir suchen nun die Wahrscheinlichkeit von $A = \{(M,M)\}$ under der Vorinformation B, welche durch die folgende Menge beschrieben werden kann:

$$B \;=\; \{(J,M),\,(M,J),\,(M,M)\}.$$

Damit erhalten wir

$$P(A|B) \;=\; \frac{P(A \cap B)}{P(B)} \;=\; \frac{\frac{1}{4}}{\frac{3}{4}} \;=\; \frac{1}{3}.$$

Proposition 1.3.2

Sei wieder (Ω, P) ein disktreter Wahrscheinlichkeitsraum und sei $B \subset \Omega$ mit $P(B) > 0$.

Dann ist

$$P(\,\cdot\,, B) : \mathcal{P}(\Omega) \to [0,1]$$

ein Wahrscheinlichkeitsmaß und heißt ***bedingte Wahrscheinlichkeitsverteilung*** von P unter B.

Beweis

Zunächsteinmal gilt

$$P(\Omega, B) \;=\; \frac{P(\Omega \cap B)}{P(B)} \;=\; \frac{P(B)}{P(B)} \;=\; 1.$$

Weiter muss noch gezeigt werden, dass $P(\cdot, B)$ σ-additiv ist. Seien dazu $(A_i)_{i\in\mathbb{N}} \subset \Omega$ paarweise disjunkte Teilmengen. Dann folgt aus der σ Additivität von P gerade

$$P\left(\bigcup_{i=1}^{\infty} A_i \middle| B\right) \;=\; \frac{P\left(\left(\bigcup_{i=1}^{\infty} A_i\right) \cap B\right)}{P(B)} \;=\; \frac{P\left(\bigcup_{i=1}^{\infty} A_i \cap B\right)}{P(B)}$$

$$= \frac{\sum_{i=1}^{\infty} P(A_i \cap B)}{P(B)} = \sum_{i=1}^{\infty} P(A_i|B),$$

also ist auch $P(\,\cdot\,, B)$ σ-additiv und damit ein Wahrscheinlichkeitsmaß. $\square$

Satz 1.3.3

Sei P eine bedingte Wahrscheinlichkeitsverteilung und seien $A, B, A_i, B_i \subset \Omega$ für $i \in \mathbb{N}$ und dabei B_i paarweise disjunkt. Dann:

(1) Es gilt die ***Bayessche Formel***

$$P(A|B) = P(B|A) \cdot \frac{P(A)}{P(B)}.$$

(2) Es gilt die ***Multiplikationsformel***

$$P\left(\bigcap_{i=1}^{n} A_i\right) = P(A_1) \cdot P(A_2|A_1) \cdot P(A_3|A_1 \cap A_2) \cdot \ldots \cdot P(A_n|A_1 \cap .. \cap A_{n-1}).$$

(3) Für ein $A \subset \bigcup_{i=1}^{\infty} B_i$ gilt die ***totale Wahrscheinlichkeit***

$$P(A) = \sum_{i=1}^{\infty} P(A|B_i)P(B_i) \quad \text{und} \quad P(B_k|A) = \frac{P(A|B_k)P(B_k)}{\sum_{i=1}^{\infty} P(A|B_i)P(B_i)}.$$

Beweis

Zunächst zu Teil **(1)**. Für $P(A) = 0$ gilt

$$P(A|B) = \frac{P(A \cap B)}{P(B)} \leq \frac{P(A)}{P(B)} = 0,$$

also wie behauptet $P(A|B) = 0$. Für den zweiten Spezialfall $P(B) = 0$ folgt nach Definition

$$P(A|B) \cdot P(B) = P(B|A) \cdot P(A).$$

Betrachten wir nun die Fälle mit $P(A) > 0$ und $P(B) > 0$. Dafür gilt

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

Nun zu Teil **(2)**. Es gilt gerade

$$\begin{aligned}
& P(A_1) \cdot \prod_{i=2}^{n} P(A_i|A_1 \cap .. \cap A_{i-1}) \\
= \; & P(A_1) \cdot \frac{P(A_1 \cap A_2)}{P(A_1)} \cdot \frac{P(A_1 \cap A_2 \cap A_3)}{P(A_1 \cap A_2)} \cdot \ldots \cdot \frac{P(A_1 \cap .. \cap A_n)}{P(A_1 \cap .. \cap A_{n-1})} \\
= \; & P(A_1 \cap .. \cap A_n),
\end{aligned}$$

womit auch die Multiplikationsformel gezeigt wurde. Bei der totalen Wahrscheinlichkeit gilt aufgrund der σ Additivität von P

$$\begin{aligned} P(A) &= P\left(A \cap \bigcup_{i=1}^{\infty} B_i\right) = P\left(\bigcup_{i=1}^{\infty} A \cap B_i\right) \\ &= \sum_{i=1}^{\infty} P(A \cap B_i) = \sum_{i=1}^{\infty} \frac{P(A \cap B_i)}{P(B_i)} P(B_i) = \sum_{i=1}^{\infty} P(A|B_i)P(B_i). \end{aligned}$$

Kombiniert man nun dieses Ergebnis mit Teil **(1)**, so erhält man auch die zweite Aussage zur totalen Wahrscheinlichkeit. □

Beispiel Gefangene

Es gibt drei Gefangene A, B und C, von denen einer freigelassen wird. Der Wärter darf den einzelnen Gefangenen nicht sagen, ob er freigelassen wird. A erfährt aber vom Wärter, dass C nicht freigelassen wird. Es soll nun geklärt werden, ob die Wahrscheinlichkeit dafür, dass A freigelassen wird, weiterhin 1/3 ist, da die Entscheidung bereits feststand, oder ob die Wahrscheinlichkeit nun 1/2 ist, da entweder A oder B freigelassen wird.

Wir schreiben F_A, F_B und F_C dafür, dass der jeweilige Gefangene freigelassen wird, und G_B und G_C dafür, dass B bzw. C weiterhin gefangen bleibt. Die vorliegende Situation zeigt nun die folgende Abbildung:

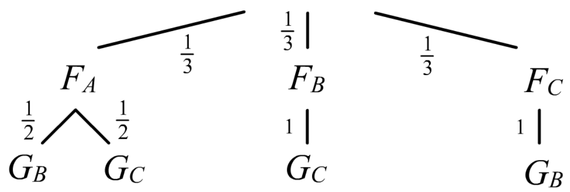


Abbildung 1.3: Beispiel Gefangene.

Nach der totalen Wahrscheinlichkeit folgt nun

$$\begin{aligned} P(F_A|G_C) &= \frac{P(G_C|F_A)P(F_A)}{P(G_C|F_A)P(F_A) + P(G_C|F_B)P(F_B) + P(G_C|F_C)P(F_C)} \\ &= \frac{\frac{1}{2}\cdot\frac{1}{3}}{\frac{1}{2}\cdot\frac{1}{3} + \frac{1}{3} + 0} = \frac{1}{3}. \end{aligned}$$

Beispiel Diagnose einer Krankeit

In einer Vorsorgeuntersuchung wird auf eine Erkrankung untersucht. Der Anteil der Kranken Testpersonen ist $\frac{1}{145}$. Der Test ist mit einer Wahrscheinlichkeit von 94 % negativ (N), wenn keine Erkrankung vorliegt (K^C). Anderherum ist der Test mit einer Wahrscheinlichkeit von 96 % positiv (N^C), wenn eine Erkrankung vorliegt (K).

Es soll nun geklärt werden, wie hoch die Wahrscheinlichkeit dafür ist, dass die Testperson wirklich krank ist, wenn der Test positiv ist.

Als Modell wählen wir

$$\Omega \;=\; \{K, K^C\} \times \{N, N^C\},$$

es sei P ein Wahrscheinlichkeitsmaß auf Ω und es sind

$$N \;=\; \{(K,N),\,(K^C,N)\} \qquad \text{und} \qquad K \;=\; \{(K,N),\,(K,N^C)\}.$$

Durch die Vorinformationen haben wir $P(K) = \frac{1}{145}$, $P(N, K^C) = 0.94$ und $P(N^C, K) = 0.96$. Wir suchen nun $P(K|N^C)$. Nach der Bayesschen und der Multiplikationsformel folgt

$$\begin{aligned} P(K|N^C) \;&=\; \frac{P(N^C|K)P(K)}{P(N^C)} \;=\; \frac{P(N^C|K)P(K)}{P(N^C|K)P(K) + P(N^C|K^C)P(K^C)} \\ &=\; \frac{0.96 \cdot \frac{1}{145}}{0.96 \cdot \frac{1}{145} + \left(1 - \frac{1}{145}\right) + (1 - 0.94)} \;\approx\; 0.1 \;=\; 10\,\%. \end{aligned}$$

Die Testperson ist also bei einem postivien Testergebnis nur zu 10 % wirklich erkrankt.

Beispiel Autorenidentifikation

Autor 1 und Autor 2 kommen beide als Autoren für ein Werk in Frage. Die Experten sind unentschlossen, 50 % sind für 1 und 50 % sind für 2. Aus Studien ist bekannt, dass die Häufigkeit des Wortes *also* in Texten einer Länge m mit $P(a_i, m)$ verteilt ist, wobei $a_1 = 0.6 \cdot 10^{-3}$ und $a_2 = 1.4 \cdot 10^{-3}$. Im vorliegenden Text der Länge $m = 2500$ ist die Häufigkeit des Wortes *also* gleich 5.

Als Modell für den Test wählen wir

$$\Omega \;=\; \mathbb{N} \cup \{0\} \times \{1,2\},$$

dabei ist $\mathbb{N} \cup \{0\}$ die Häufigkeit des Wortes *also* und $\{1,2\}$ steht für den jeweiligen Autor. Weiter seien

$$\pi_1 : \Omega \to \mathbb{N} \cup \{0\} \qquad \text{und} \qquad \pi_2 : \Omega \to \{1,2\}$$

die Projektionsfunktionen, also $\pi_1(n,i) = n$ und $\pi_2(n,i) = i$. Auf Ω wählen wir die Poission Verteilung und erhalten damit nach unseren Vorinformationen

$$P(\pi_2 = 1) \;=\; \frac{1}{2} \qquad \text{und} \qquad P(\pi_2 = 2) \;=\; \frac{1}{2}$$

und mit $k \in \mathbb{N} \cup \{0\}$

$$\begin{aligned} P(\{\pi_1 = k\}|\{\pi_2 = 1\}) &= \frac{(ma_1)^k}{k!}\mathrm{e}^{-ma_1} \quad \text{und} \\ P(\{\pi_1 = k\}|\{\pi_2 = 2\}) &= \frac{(ma_2)^k}{k!}\mathrm{e}^{-ma_2}. \end{aligned}$$

Nach der totalen Wahrscheinlichkeit folgt nun

$$\begin{aligned} P(\{\pi_1 = k\}) &= \frac{1}{2}P(\{\pi_1 = k\}|\{\pi_2 = 1\}) + \frac{1}{2}P(\{\pi_1 = k\}|\{\pi_2 = 2\}) \\ &= \frac{1}{2}\left(\frac{(ma_1)^k}{k!}\mathrm{e}^{-ma_1} + \frac{(ma_2)^k}{k!}\mathrm{e}^{-ma_2}\right). \end{aligned}$$

Damit folgt nun nach der Bayesschen Formel

$$\begin{aligned} P(\{\pi_2 = 1\}|\{\pi_1 = k\}) &= P(\{\pi_1 = k\}|\{\pi_2 = 1\}) \cdot \frac{\{\pi_2 = 1\}}{\{\pi_1 = k\}} \\ &= \frac{a_1^k \mathrm{e}^{-ma_1}}{a_1^k \mathrm{e}^{-ma_1} + a_2^k \mathrm{e}^{-ma_2}}. \end{aligned}$$

Mit unseren Werten ergibt sich also

$$\begin{aligned} P(\{\pi_2 = 1\}|\{\pi_1 = 5\}) &\approx 10\,\%, \\ P(\{\pi_2 = 2\}|\{\pi_1 = 5\}) &\approx 90\,\%, \end{aligned}$$

also sollte Autor 2 als Autor gewählt werden.

Beispiel Gambler's Ruin

Ein Spieler hat ein Startkapitel von k Talern. Es wird eine Münze geworfen, erscheint Wappen, so erhält er ein Taler, erscheint Zahl, so muss er ein Taler abgeben.

Hat der Spieler irgendwann gar keine Taler mehr, so bedeutet das der Ruin. Es soll nun die Wahrscheinlichkeit dafür berechnet werden, dass er $N > k$ Taler erreicht, ohne dass er zuvor in den Ruin fällt.

Als Modell wählen wir

$$\Omega = \bigcup_{k=1}^{N} (\Omega_0^k \cup \Omega_1^k),$$

dabei ist

$$\begin{aligned} \Omega_0^k = \{(\omega_1, .., \omega_n) \mid n \in \mathbb{N},\ \omega_i \in \{-1, 1\},\ 0 < k + \sum_{i=1}^{r} \omega_i < N \ \forall\ n, \\ r \leq n - 1,\ k + \sum_{i=1}^{n} \omega_i \in \{0, N\}\}, \end{aligned}$$

$$\Omega_1^k = \{(\omega_1, .., \omega_n) \mid n \in \mathbb{N},\ \omega_i \in \{-1, 1\},\ 0 < k + \sum_{i=1}^{n} \omega_i < N\ \forall\, n\},$$

Ω_0^k ist also die Menge aller Spiele, die beendet werden und Ω_1^k ist die Menge aller Spiele, die unendlich lange laufen.

Offenbar sind Ω_0^k und Ω_1^k disjunkt. (Ω, P) beschreibt nun das gegebene Spiel, dabei hat P folgende Eigenschaften:

(1) $P(\Omega_1^k) = 0$, die ergibt sich aus den Irrfahrten und Markowketten [siehe unten].

(2) Sei A_k die Menge dafür, dass beim Startkapitel k die 0 vor dem N erreicht wird. Dann gilt also

$$P(A_N) = 0 \qquad \text{und} \qquad P(A_0) = 1.$$

(3) B sei die Menge aller Möglichkeiten, dass beim ersten Wurf das Wappen erscheint, also

$$P(B) = \frac{1}{2}.$$

Damit erhalten wir auch

$$P(A_k|B) = P(A_{k+1}) \qquad \text{und} \qquad P(A_k|B^C) = P(A_{k-1}).$$

Wir definieren nun die Ruinwahrscheinlichkeit $f(k) = P(A_k)$ und erhalten damit

$$\begin{aligned} f(k) &= P(A_k) = P(A_k|B)P(B) + P(A_k|B^C)P(B) \\ &= \frac{1}{2}P(A_{k+1}) + \frac{1}{2}P(A_{k-1}) = \frac{1}{2}\left(f(k+1) + f(k-1)\right), \end{aligned}$$

also insbesondere für alle k

$$f(k+1) - f(k) = f(k) - f(k-1).$$

Nach den Randbedingungen aus **(2)** folgt

$$f(N) = 0 \qquad \text{und} \qquad f(0) = 1.$$

Damit erhalten wir

$$\begin{aligned} f(k+1) - f(k) &= f(k) - f(k-1) = \ldots = f(1) - f(0) \\ &= f(1) - 1 =: b, \\ f(k) &= b + f(k-1) = 2b + f(k-1) = \ldots = kb + 1. \end{aligned}$$

Nutzen wir nun wieder $f(N) = 0$, so muss $b = -1/N$ folgen. Die Ruinwahrscheinlichkeit ist also

$$f(k) = 1 - \frac{k}{N}.$$

1.4 Stochastische Unabhängigkeit

Definition 1.4.1

Sei (Ω, P) ein diskreter Wahrscheinlichkeitsraum und seien $A, B, A_i \subset \Omega$, dabei ist $i \in I$ und I eine beliebige Indexmenge.

(1) A und B heißen ***stochastisch unabhängig*** oder einfach nur ***unabhängig***, wenn $P(A|B) = P(A)$ gilt, wenn also

$$P(A \cap B) \;=\; P(A) \cdot P(B)$$

ist.

(2) $\{A\}_{i\in I}$ heißen ***stochastisch unabhängig*** oder einfach nur ***unabhängig***, wenn für alle endlichen $J \subset I$ gilt:

$$P\left(\bigcap_{i\in J} A_j\right) \;=\; \prod_{j\in J} P(A_j).$$

(3) $\{A\}_{i\in I}$ heißen ***paarweise stochastisch unabhängig*** oder einfach nur ***paarweise unabhängig***, wenn für alle $i, j \in I$ mit $i \neq j$ gilt, dass A_i und A_j stochastisch unabhängig sind.

Bemerkung

Aus der stochastischen Unabhängigkeit folgt die paarweise stochastische Unabhängigkeit, die Umkehrung ist im Allgemeinen falsch.

Beispiel Münzwurf

Es wird zweimal mit einer fairen Münze geworfen. Dazu definieren wir die folgenden Mengen:

$$\begin{aligned} A_1 &= \{\text{erster Wurf ist Kopf}\}, \\ A_2 &= \{\text{zweiter Wurf ist Kopf}\}, \\ A_3 &= \{\text{beide Würfe gleich}\}. \end{aligned}$$

Dann gilt

$$P(A_1 \cap A_2 \cap A_3) \;=\; \frac{1}{4} \qquad \text{und} \qquad P(A_1) \cdot P(A_2) \cdot P(A_3) \;=\; \frac{1}{8},$$

also sind A_1, A_2 und A_3 nicht stochastisch unabhängig. Es gilt aber

$$\begin{aligned} P(A_1 \cap A_2) &= \frac{1}{4} \quad \text{und} \quad P(A_1) \cdot P(A_2) = \frac{1}{4}, \\ P(A_2 \cap A_3) &= \frac{1}{4} \quad \text{und} \quad P(A_2) \cdot P(A_3) = \frac{1}{4}, \\ P(A_1 \cap A_3) &= \frac{1}{4} \quad \text{und} \quad P(A_1) \cdot P(A_3) = \frac{1}{4}, \end{aligned}$$

also sind A_1, A_2 und A_3 paarweise stochastisch unabhängig.

Beispiel Familie

Wir betrachten wieder eine Familie mit zwei Kindern. Dabei steht J für Junge und M für Mädchen, wir haben also

$$\Omega \;=\; \{J,M\}^2 \;=\; \{(J,J),\,(J,M),\,(M,J),\,(M,M)\}.$$

Es sei P die Laplace Verteilung und für $i=1,2$ definieren wir die Projektionsfunktionen

$$\pi_i : \Omega \to \{M,J\}.$$

Weiter definieren wir dazu die folgenden Mengen:

$$\begin{aligned} A_1 &= \{\text{erstes Kind ist ein Junge}\} \;=\; \{\pi_1 = J\}, \\ A_2 &= \{\text{beide Kinder unterschiedliches Geschlecht}\} \;=\; \{\pi_1 \neq \pi_2\}, \\ A_3 &= \{\text{erstes Kind ist ein Mädchen}\} \;=\; \{\pi_1 = M\}. \end{aligned}$$

Dann haben wir

$$\begin{aligned} P(A_1 \cap A_2) \;=\; P(\{(J,M)\}) \;=\; \frac{1}{4} \quad &\text{und} \quad P(A_1)\cdot P(A_2) \;=\; \frac{1}{4}, \\ P(A_2 \cap A_3) \;=\; P(\{(M,J)\}) \;=\; \frac{1}{4} \quad &\text{und} \quad P(A_2)\cdot P(A_3) \;=\; \frac{1}{4}, \\ P(A_1 \cap A_3) \;=\; P(\emptyset) \;=\; 0 \quad &\text{und} \quad P(A_1)\cdot P(A_3) \;=\; \frac{1}{4}, \end{aligned}$$

also sind A_1 und A_2 sowie A_2 und A_3 jeweils stochastisch unabhängig, A_1 und A_3 jedoch nicht. Die stochastische Unabhängigkeit ist also nicht transitiv.

Beispiel Projektionen

Wir betrachten die Menge $\Omega = \{0,1\}^3$ und für $i = 1,2,3$ seien

$$\pi_i : \Omega \to \{0,1\}$$

die Projektionsfunktionen. Dazu definieren wir die Mengen

$$\begin{aligned} A_1 &= \{\pi_1 + \pi_2 \;=\; 1\} \;=\; \{(\omega_1,\omega_2,\omega_2) \in \Omega \mid \omega_1 + \omega_2 = 1\}, \\ A_2 &= \{\pi_2 + \pi_3 \;=\; 1\} \;=\; \{(\omega_1,\omega_2,\omega_2) \in \Omega \mid \omega_2 + \omega_3 = 1\}, \\ A_3 &= \{\pi_1 + \pi_3 \;=\; 1\} \;=\; \{(\omega_1,\omega_2,\omega_2) \in \Omega \mid \omega_1 + \omega_3 = 1\}. \end{aligned}$$

Damit gilt

$$\begin{aligned} P(A_1 \cap A_2) \;=\; \frac{1}{4} \quad &\text{und} \quad P(A_1)\cdot P(A_2) \;=\; \frac{1}{4}, \\ P(A_2 \cap A_3) \;=\; \frac{1}{4} \quad &\text{und} \quad P(A_2)\cdot P(A_3) \;=\; \frac{1}{4}, \\ P(A_1 \cap A_3) \;=\; \frac{1}{4} \quad &\text{und} \quad P(A_1)\cdot P(A_3) \;=\; \frac{1}{4}, \end{aligned}$$

also sind A_1, A_2 und A_3 paarweise stochastisch unabhängig aber insgesamt nicht stochastisch unabhängig, da $A_1 \cap A_2 \cap A_3 = \emptyset$ gilt.

Beispiel Netzwerk

Wir berachten ein Netzwerk mit drei Knoten A, B und C und je zwei Verbindunge zwischen A und B und zwischen B und C [siehe Abbildung 1.4].

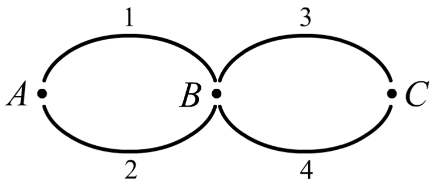


Abbildung 1.4: Veranschaulichung des Netzwerkes.

Für $i = 1, .., 4$ definieren wir

$$x_i = \begin{cases} 0, & \text{wenn die } i \text{ te Verbindung defekt} \\ 1, & \text{wenn die } i \text{ te Verbindung intakt} \end{cases} .$$

Wir haben die Vorinformation, dass die Defekte in den Verbindungen stochastisch unabhängig sind und dass die Defekte mit einer Wahrscheinlichkeit von ϑ auftreten, also $P(x_i = 0) = \vartheta$ für $i = 1, .., 4$.

Es soll nun untersucht werden, wie groß die Wahrscheinlichkeit für eine intakte Verbindung von A nach C ist.

Damit es eine intakte Verbindung von A nach C geben kann, muss es eine intakte Verbindung von A nach B und von B nach C geben. Wir erhalten also

$$\begin{aligned} & P((\{x_1 = 1\} \cup \{x_2 = 1\}) \cap (\{x_3 = 1\} \cup \{x_4 = 1\})) \\ = \ & (1 - P(x_1 = x_2 = 0)) \cdot (1 - P(x_3 = x_4 = 0)) \\ = \ & (1 - P(x_1 = 0)P(x_2 = 0)) \cdot (1 - P(x_3 = 0)P(x_4 = 0)) \ = \ (1 - \vartheta^2)^2 . \end{aligned}$$

Derartige Probleme werden in der Zuverlässigkeitstheorie behandelt.

Beispiel Hardy-Weinberg-Gesetz

Wie betrachten die Vererbung in diploiden Organismen, am Beispiel der Blütenfarbe haben wir zum Beispiel

Rot	Rosa	Weiß	Phänotypen,
RR	RW	WW	Genotypen.

Wie untersuchen nun die Genotypen der 0. Generation bzw. der Elterngeneration AA, Aa und aa und nehmen an, dass sie mit den konstanten Wahrscheinlichkeiten

$$P_0(AA) = u, \qquad P_0(Aa) = 2v \qquad \text{und} \qquad P_0(aa) = w$$

auftreten, dabei gilt natürlich $w+2v+w = 1$. Bei der Vererbung in diploiden Organismen gibt es eine unabhängige Rekombination der Gene.

Wir wollen die erste Kindergeneration P_1 untersuchen, also die Verteilungen in der ersten Generation.

Zunächst berechnen wir $P_1(AA)$, dazu die folgende Tabelle 1.2:

P_0-Konstellationen	$W(\text{konst})$	$W(AA\|\text{konst})$
$AA\ AA$	u^2	1
$AA\ Aa$	$2uv$	1/2
$Aa\ AA$	$2uv$	1/2
$Aa\ Aa$	$4v^2$	1/4
$aa\ AA$	uw	0
$AA\ aa$	uw	0
$Aa\ aa$	$2vw$	0
$aa\ Aa$	$2vw$	0
$aa\ aa$	v^2	0

Tabelle 1.2: Wahrscheinlichkeiten der ersten Kindergeneration.

Nach der totalen Wahrscheinlichkeit erhalten wir nun

$$\begin{aligned} P_1(AA) &= \sum W(AA|\text{konst}) \cdot W(\text{konst}) \\ &= u^2 + 2uv\frac{1}{2} + 2uv\frac{1}{2} + 4uv\frac{1}{4} \;=\; (u+v)^2. \end{aligned}$$

Aus Symmetriegründen erhalten wir

$$\begin{aligned} P_1(aa) &= (v+w)^2, \\ P_1(aA) &= 1 - P_1(AA) - P_1(aa) \;=\; 2(u+v)(v+w). \end{aligned}$$

Für die erste Kindergeneration habne wir also folgende Wahrscheinlichkeiten berechnet:

$$\begin{array}{ccc} AA & Aa & aa \\ u_1 := (u+v)^2 & 2v_1 := 2(u+v)(v+w) & w_1 := (v+w)^2 \end{array}$$

Damit erhalten wie analog auch für die zweite Kindergeneration

$$\begin{aligned} P_2(AA) &= (u_1+v_1)^2 \;=:\; u_2, \\ P_2(Aa) &= 2(u_1+v_1)(v_1+w_1) \;=:\; 2v_2, \\ P_2(aa) &= (v_1+w_1)^2 \;=:\; w_2. \end{aligned}$$

Setzen wir nun u, v und w ein, so erhalten wir

$$u_2 \;=\; ((u{+}v){+}(u{+}v)(v{+}w))^2 \;=\; (u{+}v)^2(\underbrace{(u+v)+(v+w)}_{=1}) \;=\; (u{+}v)^2 \;=\; u_1.$$

Wiederum aus Symmetriegründen erhalten wir

$$w_2 \;=\; w_1 \qquad \text{und} \qquad v_2 \;=\; v_1.$$

Die Genotypenverteilung ist ab der ersten Kindergeneration also stationär und damit wurde das ***Hardy-Weinberg-Gesetz*** mathematisch bewiesen.

Natürlich wurden dabei einige Annahmen vorgenommen, es wurden zum Beispiel keine Mutationen, keine Migration und keine Selektion berücksichtigt.

Über die stochastische Unabhängigkeit wollen wir nun noch die geometrische Verteilung definieren:

Geometrische Verteilung

Es sei (Ω, P') ein Wahrscheinlichkeitsraum und $(A_i)_{i\in\mathbb{N}} \subset \mathcal{P}(\Omega)$ seien stochastisch unabhängig mit $P'(A_i) = \vartheta$ für alle $i \in \mathbb{N}$.

Dann definieren wir $(\mathbb{N}, P)$ mit

$$f(n) \;=\; P(\{n\}) \;=\; P'(\underbrace{A_1 \cap .. \cap A_{n-1}}_{\text{kein Erfolg}} \cap \underbrace{A_n^C}_{\text{Erfolg}}) \;=\; (1-\vartheta)^{n-1} \cdot \vartheta.$$

Dadurch wird eine Zähldichte definiert, denn es gilt nach der geometrischen Reihe

$$\sum_{n=1}^{\infty} f(n) \;=\; \vartheta \cdot \sum_{n=1}^{\infty} (1-\vartheta)^{n-1} \;=\; \vartheta \cdot \sum_{n=0}^{\infty} (1-\vartheta)^n \;=\; \frac{\vartheta}{1-(1-\vartheta)} \;=\; 1.$$

Das zugehörige Wahrscheinlichkeitsmaß P heißt ***geometrische Verteilung***, welche die Wartezeit auf den ersten Erfolg beschreibt.

Satz 1.4.2

Seien $A, B, A_i \subset \Omega$ und $i \in I$, dabei ist $I \subset \mathbb{N}$ eine beliebige Indexmenge.

(1) Sind A und B stochastisch unabhängig, so aus A und B^C.

(2) Sind $\{A_i\}_{i\in I}$ stochastisch unabhängig, so ist für alle endlichen und disjunkten Mengen $J, K \subset I$ auch

$$\{A_j^C, A_k \mid j \in J,\, k \in K\}$$

stochastisch unabhängig.

Beweis

Zu Teil **(1)** haben wir

$$\begin{aligned} P(A \cap B^C) &= P(A \setminus (A \cap B)) \;=\; P(A) - P(A \cap B) \\ &= P(A) - P(A)(B) \;=\; P(A)(1 - P(B)) \;=\; P(A)P(B^C) \end{aligned}$$

und damit wurde die Behauptung gezeigt. Teil **(2)** ergibt sich durch Induktion über $|J|$ und aus Teil **(1)**. □

Korollar 1.4.3

Seien $A_i \subset \Omega$ mit $i \in \mathbb{N}$.

(1) $\{A_i\}_{1 \leq i \leq n}$ ist genau dann stochastisch unabhängig, wenn für alle $B_i \in \{A_i, A_i^C\}$

$$P\left(\bigcap_{i=1}^{n} B_i\right) = \prod_{i=1}^{n} P(B_i)$$

gilt.

(2) $\{A_i\}_{i \in \mathbb{N}}$ ist genau dann stochastisch unabhängig, wenn für alle $B_i \in \{A_i, A_i^C\}$

$$P\left(\bigcap_{i=1}^{\infty} B_i\right) = \prod_{i=1}^{\infty} P(B_i)$$

gilt.

Beispiel Bernoulliexperiment

Sei $\Omega = \{0,1\}^n$, sei $x = (x_1, .., x_n) \in \Omega$ und sei π_i die i te Projektion. Wir definieren Erfolgt mit 1 und Misserfolg mit 0, dabei sei $\vartheta \in [0,1]$ die Erfolgswahrscheinlichkeit.

Es soll nun ein Modell von n unabhängingen Experimenten formulert werden. Dazu definieren wir

$$A_i = \{x \in \Omega \mid x_i = 1\} = \{\pi_i = 1\}.$$

Das Problem ist nun die Konstruktion eines Wahrscheinlichkeitsmaßes P auf Ω mit den folgenden Eigenschaften:

(1) Es gilt $P(A_i) = \vartheta$ für alle $1 \leq i \leq n$.

(2) $A_1, .., A_n$ sind stochastisch unabhängig.

Betrachten wir die Zähldichte

$$f(x) = P(\{x\}) = \vartheta^{\sum_{i=1}^{n} x_i} \cdot \left(1 - \vartheta^{n - \sum_{i=1}^{n} x_i}\right) \qquad \text{für alle } \; x \in \Omega.$$

Damit haben wir gerade das gesuchte Wahrscheinlichkeitsmaßes P gefunden.

1.5 Zufallsvariablen

Definition 1.5.1

Sei (Ω, P) eine diskreter Wahrscheinlichkeitsraum.

(1) Jede Abbildung $X : \Omega \to \Omega'$ heißt ***Zufallsvariable***. Ist $\Omega' = \mathbb{R}$, so heißt X ***reelle Zufallsvariable***. Ist $\Omega' = \mathbb{R}^n$, so heißt X ***Zufallsvektor***.

(2) Wir definieren $X^{-1} : \mathcal{P}(\Omega') \to \mathcal{P}(\Omega)$, dabei

$$X^{-1}(A') \;=\; \{\omega \in \Omega \mid X(\omega) \in A'\} \;=:\; \{X \in A^{-1}\}.$$

X^{-1} heißt ***Urbildfunktion***.

Lemma 1.5.2 (Eigenschaften der Urbildfunktion)

Sei $f : \Omega \to \Omega'$ und seien $A', B', A'_i, B'_i \subset \Omega$ mit $i \in I$ und I ist eine beliebige Indexmenge. Dann gilt:

(1) $f^{-1}\left(\bigcap_{i\in I} A'_i\right) \;=\; \bigcap_{i\in I} f^{-1}(A'_i)$.

(2) $f^{-1}\left(\bigcup_{i\in I} A'_i\right) \;=\; \bigcup_{i\in I} f^{-1}(A'_i)$.

(3) Sind A'_i paarweise disjunkt, dann sind auch $f^{-1}(A'_i)$ paarweise disjunkt.

(4) $f^{-1}(\emptyset) \;=\; \emptyset$ und $f^{-1}(\Omega) \;=\; \Omega$.

(5) Aus $A' \subset B'$ folgt $f^{-1}(A') \subset f^{-1}(B')$.

(6) $f^{-1}(A' \setminus B') \;=\; f^{-1}(A') \setminus f^{-1}(B')$.

(7) $f^{-1}((B')^C) \;=\; (f^{-1}(B'))^C$.

(8) Sei $g : \Omega' \to \Omega''$ und $B'' \subset \Omega''$, dann $(g \cdot f)^{-1}(B'') \;=\; f^{-1}(g^{-1}(B''))$.

Proposition 1.5.3

Sei (Ω, P) ein diskreter Wahrscheinlichkeitsraum und $X : \Omega \to \Omega'$ eine Zufallsvariable. Dann ist

$$\begin{aligned} P^X : \mathcal{P}(\Omega') &\;\to\; [0,1] \\ A' &\;\mapsto\; P(X^{-1}(A')) \end{aligned}$$

ein Wahrscheinlichkeitsmaß auf Ω'.

P^X heißt die ***Verteilung*** von X bezüglich P und (Ω', P^X) ist ein Wahrscheinlichkeitsraum.

Beispiel Würfel

Es werden zwei faire Würfel geworfen, wir benutzen also

$$\Omega \;=\; \{1,\dots,6\}^2 \;=\; \{(i,j) \mid 1 \le i,j \le 6\}$$

und die Laplace Verteilung P auf Ω. Weiter definieren wir die Zufallsvariable

$$X : \Omega \to \{1,\dots,12\} \qquad \text{mit} \qquad X(i,j) \;=\; i+j,$$

also die Augensumme. Damit können wir nun die Verteilung von X bestimmen:

$$\begin{aligned}
P^X(\{1\}) &= P(X^{-1}(\{1\})) = P(\emptyset) = 0,\\
P^X(\{2\}) &= P(X^{-1}(\{2\})) = P(\{(1,1)\}) = 1/36,\\
P^X(\{3\}) &= P(X^{-1}(\{3\})) = P(\{(1,2),(2,1)\}) = 2/36,\\
P^X(\{4\}) &= P(X^{-1}(\{4\})) = P(\{(1,3),(3,1),(2,2)\}) = 3/36,\\
&\;\vdots\\
P^X(\{10\}) &= P(X^{-1}(\{10\})) = P(\{(4,6),(6,4),(5,5)\}) = 3/36,\\
P^X(\{11\}) &= P(X^{-1}(\{11\})) = P(\{(5,6),(6,5)\}) = 2/36,\\
P^X(\{12\}) &= P(X^{-1}(\{12\})) = P(\{(6,6)\}) = 1/36.
\end{aligned}$$

Beispiel Binomialverteilung

Sei $\Omega = \{0,1\}^n$ und sei

$$P(\{x\}) \;=\; \vartheta^{\sum_{i=1}^n x_i} \cdot \left(1 - \vartheta^{n-\sum_{i=1}^n x_i}\right)$$

für alle $x = (x_1,..,x_n) \in \Omega$, dabei $\vartheta \in [0,1]$. Wir betrachten also wieder n unabhängige Experimente mit der Erfolgswahrscheinlichkeit ϑ pro Experiment.

S_n sei eine Zufallsvariable auf (Ω, P), dabei

$$S_n : \Omega \to \mathbb{N} \cup \{0\} \qquad \text{mit} \qquad S_n(x) \;=\; \sum_{i=1}^n x_i.$$

S_n entspricht also der Anzahl der Erfolge bei n Experimenten. Für $k > n$ gilt offenbar $P^{S_n}(\{k\}) = 0$. Für $k \le n$ haben wir

$$P^{S_n}(\{k\}) \;=\; P(S_n^{-1}(\{k\})) \;=\; \sum_{\substack{x\in\Omega\\ \sum_{i=1}^n x_i = k}} \vartheta^k (1-\vartheta)^{n-k} \;=\; \binom{n}{k} \vartheta^k (1-\vartheta)^{n-k}.$$

Dies ist die Zähldichte der Binomialverteilung $\mathcal{L}(n,\vartheta)$. Die Binomialverteilung beschreibt also die Anzahl der Erfolge bei n unabhängigen Experimenten.

Definition 1.5.4

Sei (Ω, P) ein diskreter Wahrscheinlichkeitsraum und seien $X_i : \Omega \to \Omega_i$ Zufallsvariablen.

$(X_i)_{i\in I}$ heißen genau dann ***stochastisch unabhängig***, wenn für alle $A_i \subset \Omega_i$ gerade $(X_i \in A_i)_{i\in I}$ stochastisch unabhängig ist.

Für die Zähldichte einer Verteiliung P^{X_i} schreiben wir nun f_{X_i}.

Proposition 1.5.5

Sei (Ω, P) ein diskreter Wahrscheinlichkeitsraum, seien $X_i : \Omega \to \Omega_i$ Zufallsvariablen mit $1 \leq i \leq n$ und sei $X = (X_1, .., X_n)$.

Dann sind $X_1, .., X_n$ genau dann stochastisch unabhängig, wenn

$$P^X = \bigotimes_{i=1}^{n} P^{X_i}$$

gilt.

Bemerkung

Sei $\Omega = \Omega_1 \times \Omega_2$ und sei P ein diskretes Wahrscheinlichkeitsmaß auf Ω mit

$$P^{\pi_i} = P_i \qquad \text{für} \qquad i = 1, 2.$$

Dann sind π_1 und π_2 genau dann stochastisch unabhängig, wenn $P = P_1 \otimes P_2$ gilt.

P^{π_i} heißen ***Randverteilungen*** von P.

Proposition 1.5.6

Seien $X_i : \Omega \to \Omega_i$ für $1 \leq i \leq n$ stochastisch unabhängige Zufallsvariablen auf einem diskreten Wahrscheinlichkeitsraum (Ω, P) und seien $f_i : \Omega_i \to \Omega_i'$ die zugehörigen Zähldichten.

Dann sind auch $(f_i \circ X_i)_{1\leq i\leq n}$ stochastisch unabhängig.

Bemerkung

Insbesondere sind also für alle $1 \leq m < n$ auch

$$y_1 = f_1(X_1, .., X_m) \qquad \text{und} \qquad y_2 = f_2(X_{m+1}, .., X_n)$$

stochastisch unabhängig.

Definition 1.5.7

Sei (Ω, P) ein Wahrscheinlichkeitsraum und seien

$$X : \Omega \to \Omega_1 \qquad \text{und} \qquad Y : \Omega \to \Omega_2$$

zwei Zufallsvariablen. Dann heißt für $y \in \Omega_2$

$$P_{X|y} : \mathcal{P}(\Omega_1) \to [0,1] \qquad \text{mit} \qquad P_{X|y}(A) \;=\; P(\{X \in A\} \mid \{Y = y\})$$

die ***bedingte Verteilung*** von X und $Y = y$.

$P_{X|y}$ ist ein Wahrscheinlichkeitsmaß auf Ω_1.

Bemerkung

Falls $P(\{Y = y\}) = 0$, dann ist $P_{X|y}(A) = P(\{X \in A\})$.

Wir müssen daher nicht zwischen P_X und P^X unterscheiden.

Satz 1.5.8

Sei (Ω, P) ein Wahrscheinlichkeitsraum und seien

$$X : \Omega \to \Omega_1 \qquad \text{und} \qquad Y : \Omega \to \Omega_2$$

zwei Zufallsvariablen. Dann gilt:

(1) Für $x \in \Omega_1$ und für $f_y(y) > 0$ hat $P_{X|y}$ die bedingte Zähldichte

$$f_{X|y}(x) \;=\; \frac{f(x,y) \cdot (x,y)}{f_y(y)}.$$

Gilt $f_y(y) = 0$, so ist $f_{X|y}(x) = f_X(x)$.

(2) Für alle $(x,y) \in \Omega_1 \times \Omega_2$ ist

$$f_{(X,Y)}(x,y) \;=\; f_{X|y}(x) f_y(y).$$

(3) X und Y genau dann stochastisch unabhängig, wenn für alle $y \in \Omega_2$

$$f_{X|y}(x) \;=\; f_X(x)$$

gilt.

Satz 1.5.9 (Faltungsformel)

Seien $X : \Omega \to \Omega_X$ und $Y : \Omega \to \Omega_Y$ stochastisch unabhängige und diskrete Zufallsvariablen mit den Zähldichten f_X und f_Y.

Dann hat die Summe $Z = X + Y$ die Zähldichte

$$f_Z(k) \;=\; f_{X+Y}(k) \;=\; \sum_{x \in \Omega_X} f_X(x) f_Y(k-x) \;=\; \sum_{y \in \Omega_Y} f_X(k-y) f_Y(y).$$

Beispiel Binomialverteilung

Seien X und Y stochastisch unabhängige binomialverteilte Zufallsvariablen mit $\mathcal{L}(n,\vartheta)$ bzw. $\mathcal{L}(m,\vartheta)$. Wir haben also die Zähldichten

$$\begin{aligned} f_X(k) &= \binom{n}{k}\vartheta^k(1-\vartheta)^{n-k}, \\ f_Y(k) &= \binom{m}{k}\vartheta^k(1-\vartheta)^{m-k}. \end{aligned}$$

Damit erhalten wir nach der Faltungsformel

$$\begin{aligned} f_{X+Y}(k) &= \sum_{i=0}^{k}\binom{n}{i}\vartheta^i(1-\vartheta)^{n-i}\binom{m}{k-i}\vartheta^{k-i}(1-\vartheta)^{m-(k-i)} \\ &= \vartheta^k(1-\vartheta)^{n+m-k}\sum_{i=0}^{k}\binom{n}{i}\binom{m}{k-i} \\ &= \binom{n+m}{k}\vartheta^k(1-\vartheta)^{n+m-k}, \end{aligned}$$

dabei wurde die Identität

$$\sum_{i=0}^{k}\binom{n}{i}\binom{m}{k-i} = \binom{n+m}{k}$$

verwendet. Die Zufallsvariable $Z = X+Y$ ist also $\mathcal{L}(n+m,\vartheta)$ binomialverteilt.

Beispiel Poissonverteilung

Seien X und Y stochastisch unabhängige Zufallsvariablen, die Poisson verteilt sind mit Parameter λ bzw. μ. Wir haben damit die Zähldichten

$$\begin{aligned} f_X(k) &= \frac{\lambda^k}{k!}\mathrm{e}^{-\lambda}, \\ f_Y(k) &= \frac{\mu^k}{k!}\mathrm{e}^{-\mu}. \end{aligned}$$

Die Faltungsformel und die Binomische Formel liefern dann

$$\begin{aligned} f_{X+Y}(k) &= \sum_{i=0}^{k}\frac{\lambda^i}{i!}\mathrm{e}^{-\lambda}\frac{\mu^{k-i}}{(k-i)!}\mathrm{e}^{-\mu} \\ &= \mathrm{e}^{-(\lambda+\mu)}\frac{1}{k!}\sum_{i=0}^{k}\frac{k!}{i!(k-i)!}\lambda^i\mu^{k-i} = \frac{(\lambda+\mu)^k}{k!}\mathrm{e}^{-(\lambda+\mu)}. \end{aligned}$$

Die Zufallsvariable $Z = X+Y$ ist also Poisson verteilt mit Parameter $\lambda+\mu$.

1.6 Reelle Zufallsvariablen

Wie betrachten nun stets einen diskreten Wahrscheinlichkeitsraum (Ω, P) mit einer reellen Zufallsvariablen $X : \Omega \to \mathbb{R}$.

Definition 1.6.1

Sei also (Ω, P) ein diskreter Wahrscheinlichkeitsraum mit einer reellen Zufallsvariablen $X : \Omega \to \mathbb{R}$.

Die Funktion

$$F_X : \mathbb{R} \to \mathbb{R} \qquad \text{mit} \qquad F_X(x) \;=\; \sum_{\substack{x' \in \Omega \\ x' \leq x}} P^X(\{x'\}) \;=\; P(X \leq x)$$

heißt ***Verteilungsfunktion*** von X bezüglich P^X.

Proposition 1.6.2

Für eine Verteilungsfunktion F_X gilt:

(1) F_X ist isoton, das heißt aus $x \leq y$ folgt $F_X(x) \leq F_X(y)$.

(2) F_X ist rechtsseitig stetig.

(3) Es ist $\lim_{x \to \infty} F_X(x) = 1$ und $\lim_{x \to -\infty} F_X(x) = 0$.

Lemma 1.6.3

Eine Verteilung P^X ist eindeutig durch die zugehörige Verteilungsfunktion F_X bestimmt.

Beispiel Münzwurf

Bei dem einfachen Münzwurf mit einer fairen Münze definieren wir eine Zufallsvariable X mit

$$X(\text{Kopf}) \;=\; -1 \qquad \text{und} \qquad X(\text{Zahl}) \;=\; 1.$$

Wir haben also

$$P(X = -1) \;=\; P(X = 1) \;=\; \frac{1}{2}.$$

Die Verteilungsfunktion zu diesem Beispiel ist in Abbildung 1.5 dargestellt.

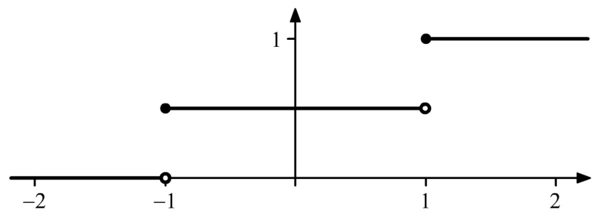


Abbildung 1.5: Verteilungsfunktion des einfachen Münzwurfes.

Beispiel Würfel

Die Zufallsvariable X gebe die Augensumme beim einmaligen Werfen eines fairen Würfels an. Dann haben wir

$$P(\{X = i\}) \;=\; \frac{1}{6} \qquad \text{für} \;\; 1 \leq i \leq 6$$

und 0 sonst. Die Verteilungsfunktion zu diesem Beispiel ist in Abbildung 1.6 dargestellt.

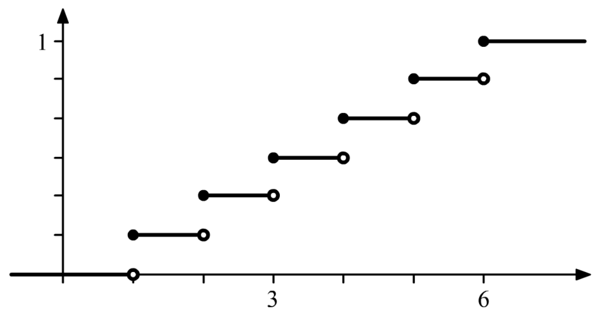


Abbildung 1.6: Verteilungsfunktion des einfachen Würfelwurfes.

Poisson Verteilung

Sei P^X Poisson Verteilt mit dem Parameter $\lambda > 0$, also $P^X = \mathcal{P}(\lambda)$. Dann gilt

$$F_X(r) \;=\; \sum_{k=0}^{r} \mathrm{e}^{-\lambda} \frac{\lambda^k}{k!} \;=\; \frac{1}{r!} \int_{\lambda}^{\infty} x^r \mathrm{e}^{-x} \,\mathrm{d}x.$$

Der Integralausdruck beschreibt die unvollständige Γ-Funktion. Diese Verteilungsfunktion wurde in Abbildung 1.7 angedeutet.

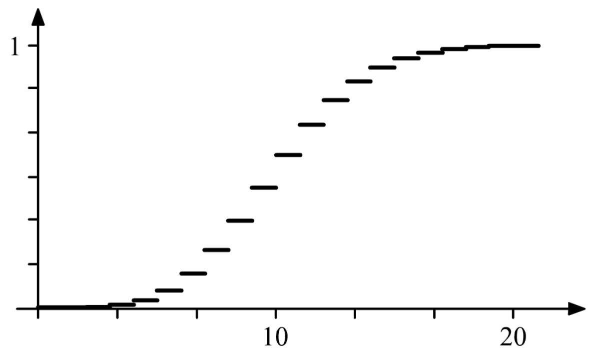


Abbildung 1.7: Verteilungsfunktion der Poisson Verteilung für $\lambda = 10$.

Binomialverteilung

Sei $P^X = \mathfrak{L}(n, \vartheta)$ für ein $\vartheta \in [0,1]$ die Binomialverteilung. Dann gilt

$$F_X(r) \;=\; \sum_{k=0}^{r} \binom{n}{r} \vartheta(1-\vartheta)^{n-k} \;=\; \binom{n}{r}(n-r) \int_{\vartheta}^{1} x^r (1-x)^{n-r-1}\, \mathrm{d}x.$$

Der Integralausdruck beschreibt die unvollständige β-Funktion.

1.7 Erwartungswert

Um den Erwartungswert einzuführen, müssen zunächst einige Notationen eingeführt werden:

Definition 1.7.1

Sei $(a_i)_{i\in I} \in \mathbb{R}$, dabei ist I eine beliebige Indexmenge.

Für $a_i \geq 0$ definieren wir

$$\sum_{i\in I} a_i \;=\; \left\{ \sum_{i\in K} a_i \;\middle|\; K \subset I,\ |K| < \infty \right\}.$$

Für $a_i \in \mathbb{R}$ definieren wir, dass $\sum_{i\in I} a_i$ genau dann exisitert, wenn

$$\sum_{i\in I} a_i^+ \;<\; \infty \qquad \text{oder wenn} \qquad \sum_{i\in I} a_i^- \;<\; \infty$$

gilt, dabei ist

$$a_i^+ \;:= \begin{cases} a_i & \text{für } a_i \geq 0 \\ 0 & \text{für } a_i < 0 \end{cases} \qquad \text{und} \qquad a_i^- \;:= \begin{cases} a_i & \text{für } a_i \leq 0 \\ 0 & \text{für } a_i > 0 \end{cases}.$$

Diese Definition erlaubt es uns stets den Umordnungssatz anzuwenden.

Definition 1.7.2

Sei (Ω, P) ein diskreter Wahrscheinlichkeitsraum und sei $X : \Omega \to \mathbb{R}$ eine reelle Zufallsvariable. Dann heißt

$$EX = \sum_{\omega \in \Omega} X(\omega) P(\{\omega\})$$

genau dann der ***Erwartungswert*** von X, wenn die Summe existiert, also endlich ist. Existiert der Erwartungswert, so heißt X auch ***integrierbar***.

Bemerkung

Es ist wichtig, dass die Summe existert, denn sonst würde der Umordnungssatz nicht gelten und die Definition des Erwartungswertes wäre nicht eindeutig. Der Erwartungswert einer Messreihe wäre dann von der Reihenfolge der Messungen abhängig.

Ist $X \geq 0$, so ist der Erwartungswert aber eindeutig bestimmt.

Proposition 1.7.3

Sei X eine reelle Zufallsvariable und der Erwartungswert EX existiere.

Dann gilt

$$EX = \sum_{x \in \mathbb{R}} x f_X(x) = \sum_{x \in \mathbb{R}} x P^x(\{x\}).$$

Beweis

Ohne Einschränkung können wir annehmen, dass $X \geq 0$ gilt. Damit erhalten wir

$$\begin{aligned} EX &= \sum_{\omega \in \Omega} X(\omega) P(\{\omega\}) = \sum_{x \in \mathbb{R}} \sum_{\omega \in \{X = x\}} X(\omega) P(\{\omega\}) \\ &= \sum_{x \in \mathbb{R}} x P(\{X = x\}) = \sum_{x \in \mathbb{R}} x P^X(\{x\}) = \sum_{x \in \mathbb{R}} x f_X(x), \end{aligned}$$

was die Behauptung zeigt. □

Beispiel Würfelwurf

Für einen fairen Würfel führen wir die reelle Zufallsvariable

$$X : \{1, \ldots, 6\} \to \{1, \ldots, 6\} \qquad \text{mit} \qquad X(i) = i$$

ein und betrachten P^X als Laplace Verteilung auf $\{1, .., 6\}$. Dann erhalten wir

$$EX = \sum_{i=1}^{6} i \frac{1}{6} = 3.5 \,.$$

Beispiel Poisson Verteilung

Sei $P^N = \mathcal{P}(\lambda)$ die Poisson Verteilung mit den Parameter $\lambda > 0$, also

$$P(\{N = n\}) \;=\; \mathrm{e}^{-\lambda}\frac{\lambda^n}{n!}.$$

Dann erhalten wir den Erwartungswert

$$\begin{aligned} EN &= \sum_{n=0}^{\infty} n \cdot \mathrm{e}^{-\lambda}\frac{\lambda^n}{n!} \;=\; \lambda \mathrm{e}^{-\lambda}\sum_{n=1}^{\infty}\frac{\lambda^{n-1}}{(n-1)!} \\ &= \lambda \mathrm{e}^{-\lambda}\sum_{n=0}^{\infty}\frac{\lambda^n}{n!} \;=\; \lambda \mathrm{e}^{-\lambda}\mathrm{e}^{\lambda} \;=\; \lambda. \end{aligned}$$

Beispiel Binomialverteilung

Sei $P^X = \mathcal{L}(n, \vartheta)$ für ein $\vartheta \in [0,1]$ die Binomialverteilung. Dann gilt nach der Binomischen Formel

$$\begin{aligned} EX &= \sum_{k=0}^{n} k\binom{n}{k}\vartheta^k(1-\vartheta)^{n-k} \\ &= n\vartheta \cdot \sum_{k=1}^{n}\frac{(n-1)!}{(n-k)!(k-1)!}\vartheta^{k-1}(1-\vartheta)^{n-k} \\ &= n\vartheta \cdot \sum_{k=0}^{n-1}\binom{n-1}{k}\vartheta^k(1-\vartheta)^{n-1-k} \;=\; n\vartheta. \end{aligned}$$

Beispiel geometrische Verteilung

Seien $(X_n)_{n\in\mathbb{N}}$ stochastisch unabhängig und $\mathcal{L}(1, \vartheta)$ verteilte Zufallsvariablen für ein $\vartheta \in]0,1[$, also

$$P(X_n = 1) \;=\; \vartheta \;=\; 1 - P(X_n = 0)$$

für alle $n \in \mathbb{N}$. Wir definieren wie üblich

$$W \;:=\; \inf\{n \in \mathbb{N} \mid X_n \;=\; 1\} \qquad \text{und} \qquad \inf\{\emptyset\} \;=\; \infty.$$

Dann haben wir

$$f_W(n) \;=\; P(\{W = n\}) \;=\; (1-\vartheta)^{n-1}\vartheta,$$

denn es gilt

$$\{W = n\} \;=\; \{X_1 = 0\} \cap \{X_2 = 0\} \cap \ldots \cap \{X_{n-1} = 0\} \cap \{X_n = 1\}.$$

Der Erwartungswert ergibt sich aus der folgenden Rechnung:

$$\begin{aligned} EW &= \sum_{n=1}^{\infty} n f_W(n) \;=\; \sum_{n=1}^{\infty} n(1-\vartheta)^{n-1}\vartheta \\ &= \vartheta \left(\sum_{n=1}^{\infty} n x^{n-1} \right)\Big|_{x=1-\vartheta} \;=\; \vartheta \left(\frac{\mathrm{d}}{\mathrm{d}x} \sum_{n=0}^{\infty} x^n \right)\Big|_{x=1-\vartheta} \\ &= \vartheta \left[\frac{\mathrm{d}}{\mathrm{d}x} \left(\frac{1}{1-x} \right) \right]_{x=1-\vartheta} \;=\; \vartheta \frac{1}{(1-x)^2}\Big|_{x=1-\vartheta} \;=\; \frac{1}{\vartheta}. \end{aligned}$$

Proposition 1.7.4 (Transformationsformel)

Sei (Ω, P) ein diskreter Wahrscheinlichkeitsraum, sei $X : \Omega \to \Omega'$ eine Zufallsvariable und sei $g : \Omega' \to \mathbb{R}$, so dass $g \circ X$ integrierbar ist.

Dann gilt

$$E(g \circ X) \;=\; \sum_{x \in \Omega'} g(x) f_X(x).$$

Proposition 1.7.5

Seien X und Y reelle Zufallsvariablen auf einem diskreter Wahrscheinlichkeitsraum (Ω, P).

Dann gilt:

(1) Ist $X \geq 0$, so folgt $EX \geq 0$.

(2) Es gilt ***Linearität***: Für $a, b \in \mathbb{R}$ und integrierbare X, Y ist auch $aX + bY$ integrierbar und es gilt

$$E(aX + bY) \;=\; aEX + bEY.$$

(3) Ist $Z : \Omega \to \mathbb{R}$ mit $Z(\omega) = 1$ für alle $\omega \in \Omega$, dann gilt $EZ = 1$.

(4) Es gilt der ***Multiplikationssatz***: Sind X, Y stochastisch unabhängig und integrierbar, so ist auch XY integrierbar und es gilt

$$E(XY) \;=\; EX \cdot EY.$$

Punkt **(4)** gilt im Allgemeinen wirklich nur dann, wenn X und Y stochastisch unabhängig sind.

Beispiel 1.7.6

Wir betrachten noch einmal ein einfaches Beispiel zu reellen Zufallsvariablen:

Sei $\Omega = \{\omega_1, \omega_2, \omega_3\}$ mit

$$P(\omega_i) \;=\; \frac{1}{3} \qquad \text{für} \;\; i \;=\; 1, 2, 3\,.$$

Weiter definieren wir die reellen Zufallsvariablen $X, Y : \Omega \to \{1, 2, 3\}$ durch

$$\begin{array}{lll} X(\omega_1) \;=\; 1, & X(\omega_2) \;=\; 2, & X(\omega_3) \;=\; 3, \\ Y(\omega_1) \;=\; 2, & Y(\omega_2) \;=\; 3, & Y(\omega_3) \;=\; 1 \end{array}$$

und suchen nun die Verteilungen P^X und P^Y. Wir erhalten aber sofort

$$\begin{aligned} P^X(\{1\}) &= P(X=1) = P(\omega_1) = 1/3, \\ P^X(\{2\}) &= P(X=2) = P(\omega_2) = 1/3, \\ P^X(\{3\}) &= P(X=3) = P(\omega_3) = 1/3, \end{aligned}$$

$$\begin{aligned} P^Y(\{1\}) &= P(Y=1) = P(\omega_3) = 1/3, \\ P^Y(\{2\}) &= P(Y=2) = P(\omega_1) = 1/3, \\ P^Y(\{3\}) &= P(Y=3) = P(\omega_2) = 1/3, \end{aligned}$$

die Zufallsvariablen X und Y haben also dieselben Verteilungen. Auch die Erwartungswerte sind gleich:

$$\begin{aligned} EX &= \sum_{\omega \in \Omega} X(\omega) P(\omega) \;=\; 1 \cdot \frac{1}{3} + 2 \cdot \frac{1}{3} + 3 \cdot \frac{1}{3} \;=\; 2, \\ EY &= \sum_{\omega \in \Omega} Y(\omega) P(\omega) \;=\; 2 \cdot \frac{1}{3} + 3 \cdot \frac{1}{3} + 1 \cdot \frac{1}{3} \;=\; 2, \end{aligned}$$

im Allgemeinen haben Zufallsvariablen mit gleichen Erwartungswerten aber natürlich unterschiedliche Verteilungen. Nun soll noch die Verteilung von $X + Y$ untersucht werden. Wir erhalten

$$X + Y : \Omega \to \{3, 4, 5\},$$

dies ist aus Abbildung 1.8 abzulesen.
Es gilt dabei

$$\begin{aligned} P^{X+Y}(\{3\}) &= P(\omega_1) = 1/3, \\ P^{X+Y}(\{4\}) &= P(\omega_3) = 1/3, \\ P^{X+Y}(\{5\}) &= P(\omega_2) = 1/3. \end{aligned}$$

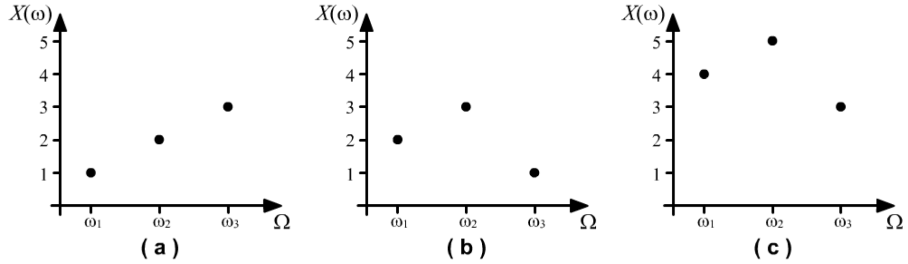


Abbildung 1.8: Verteilungen von X, Y und von $X+Y$.

Um zu prüfen, ob X und $X+Y$ stochastisch unabhängig sind, muss deren gemeinsame Verteilung untersucht werden. Es lässt sich einfach ein Gegenbeispiel finden:

$$\begin{aligned} P^{(X+Y,X)}(\{3\}\times\{1\}) &= P((X+Y)^{-1}(\{3\}),\, X^{-1}(\{1\})) \;=\; P(\omega_1) \\ &= \frac{1}{3} \neq \frac{1}{9} = \frac{1}{3}\cdot\frac{1}{3} \;=\; P(\omega_1)\cdot P(\omega_1) \\ &= P^{X+Y}(\{3\}) + P^{X}(\{1\}). \end{aligned}$$

Tabelle 1.3 veranschaulicht noch einmal die Randverteilungen von $X+Y$ und von X.

	3	4	5	$\sum$
1	1/3	0	0	1/3
2	0	0	1/3	1/3
3	0	1/3	0	1/3
$\sum$	1/3	1/3	1/3	

Tabelle 1.3: Tabelle zur Randverteilung von $X+Y$ und von X.

1.8 Varianz und Standardabweichung

Definition 1.8.1

Seien X und Y reelle Zufallsvariablen auf (Ω, P) und X^2 sowie Y^2 seien integrierbar.

Dann ist

(1) $\mathrm{Var}(X) = E((X-EX)^2)$ die ***Varianz*** von X,

(2) $\rho_X = \sqrt{\mathrm{Var}(X)}$ die ***Standardabweichung*** oder ***Streuung*** von X,

(3) $\mathrm{Cov}(X,Y) = E((X-EX)(Y-EY))$ die ***Kovarianz*** von X und Y,

(4) $\rho_{XY} = \frac{\mathrm{Cov}(X,Y)}{\rho_X \cdot \rho_Y}$ die ***Korrelation*** von X und Y,

(5) X und Y heißen ***unkorreliert***, wenn $\mathrm{Cov}(X,Y) = 0$ gilt.

Bemerkungen

Es gilt $|X| \leq 1 + X^2$, wenn also X^2 integrierbar ist, so auch $|X|$ und dann auch $(X - EX)^2$.

Der Erwartungswert EX kann als Schwerpunkt einer Verteilung interpretiert werden. Die Varianz $\mathrm{Var}(X)$ beschreibt dann das Trägheitsmoment bei einer Drehung um diesen Schwerpunkt.

Proposition 1.8.2

Seien X, Y, Z sowie $X_1, .., X_n$ quadratintegrierbare reelle Zufallsvariablen und seinen $a, b, c, d \in \mathbb{R}$.

Dann gilt:

(1) $\mathrm{Var}(X) \;=\; E(X^2) - (EX)^2.$

(2) $\mathrm{Var}(aX + b) \;=\; a^2 \cdot \mathrm{Var}(X).$

(3) $\mathrm{Cov}(X,Y) \;=\; E(XY) - (EX)(EY).$

(4) $\mathrm{Cov}(aX + b, cY + d) \;=\; ac \cdot \mathrm{Cov}(X,Y).$

(5) Formel von ***Bienaymé***

$$\mathrm{Var}\left(\sum_{i=1}^{n} X_i\right) \;=\; \sum_{i=1}^{n} \mathrm{Var}(X_i) + \sum_{i \neq j} \mathrm{Cov}(X_i, X_j).$$

(6) Sind X und Y stochastisch unabhängig, dann sind X und Y unkorreliert.

(7) Sind $X_1, .., X_n$ stochastisch unabhängig, dann gilt

$$\mathrm{Var}\left(\sum_{i=1}^{n} X_i\right) \;=\; \sum_{i=1}^{n} \mathrm{Var}(X_i).$$

Bemerkung

Damit ergeben sich noch weitere einfache Rechenregeln: Sind X und Y stochastisch unabhängig, so gilt zum Beispiel auch

$$\mathrm{Cov}(X,Y) \;=\; E(XY) - (EX)(EY) \;=\; (EX)(EY) - (EX)(EY) \;=\; 0.$$

Satz 1.8.3 (Cauchy-Schwarz-Ungleichung)

Seien X und Y quadrierbare reelle Zufallsvariablen. Dann gilt

$$(EXY)^2 \;\leq\; EX^2 \cdot EY^2$$

mit Gleichheit genau dann, wenn es $a, b \in \mathbb{R}$ mit $P(aX + bY = 0) = 1$ gibt, dabei dürfen a und b nicht beide gleichzeitig 0 sein.

Folgerung 1.8.4

Wenden wir Satz 1.8.3 auf die Zufallsvariablen $X - EX$ und $Y - EY$ an, so erhalten wir

$$|\rho_{XY}| \;=\; \frac{\operatorname{Cov}(X,Y)}{\rho_X \cdot \rho_Y} \;=\; \frac{|E((X-EX)(Y-EY))|}{\sqrt{E(X-EX)^2}\sqrt{E(Y-EY)^2}} \;\leq\; 1.$$

Lineare Regression

Seien X und Y reelle Zufallsvariablen mit $\operatorname{Var}(X), \operatorname{Var}(Y) > 0$ und seien $a, b \in \mathbb{R}$. Dann betrachten wir das Problem, bei welchem wir

$$E(X - (a + bY))^2$$

minimieren wollen. Durch Differenzieren erhalten wir die Lösungen

$$b^* \;=\; \rho_{XY}\frac{\rho_X}{\rho_Y} \qquad \text{und} \qquad a^* \;=\; EX - b^*EY \;=\; EX - \rho_{XY}\frac{\rho_X}{\rho_Y}EY.$$

Dieses Ergebnis wollen wir nun interpretieren: Betrachten wir zweidimensionale Zufallsvariablen (X_1, Y_1) bis (X_n, Y_n), die unabhängig und identisch verteilt sind, dann wird durch das Minimierungsproblem

$$E(X - (a + bY))^2$$

die Gerade bestimmt, bei welcher der quadratische Abstand der Geraden zu den X Komponenten minimiert wird. Die Gerade

$$a^* + b^*Y$$

heißt ***Regressionsgerade***.

Regressionsgerade für unterschiedliche Korrelationen

In Abbildung 1.10 **(a)** erhalten wir ein kleines $\rho_{XY} > 0$, das heißt die Abstände der Punkte zur Geraden sind groß und $\rho_{XY} > 0$ entspricht einer positiven Steigung.

Bei **(b)** ist keine mögliche Gerade erkennbar.

In Beispiel **(c)** erhalten wir ein $\rho_{XY} \approx -1$. Die Werde lassen sich gut durch eine Gerade approximieren und wir haben eine negative Steigung.

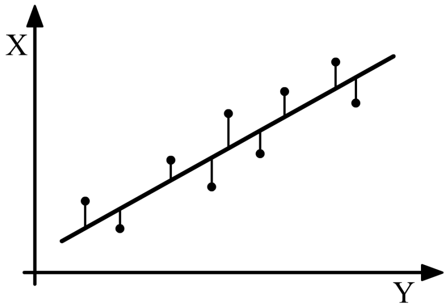


Abbildung 1.9: Beispiel einer Regressionsgeraden.

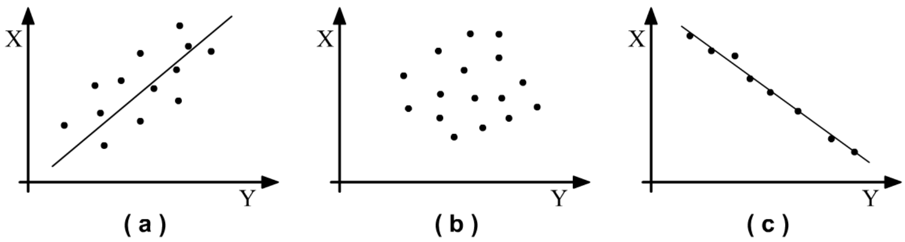


Abbildung 1.10: Regressionsgeraden zu unterschiedlichen Korrelationen.

Beispiel Binomialverteilung

Sei $P^X = \mathfrak{L}(n, \vartheta)$ die Binomialverteilung. Dann wissen wir bereits $EX = n\vartheta$ und es gilt

$$P^X \;=\; P^{S_n} \qquad \text{mit} \qquad S_n \;=\; \sum_{i=1}^{n} X_i,$$

dabei sind X_i unabhängige Bernoulliverteilte Zufallsvariablen. Nach Satz 1.8.2 erhalten wir damit

$$\begin{aligned} \operatorname{Var}(X) &= \operatorname{Var}(S_n) \;=\; \sum_{i=1}^{n} \operatorname{Var}(X_i) \;=\; \sum_{i=1}^{n} (EX_i^2 - (EX_i)^2) \\ &= n(\vartheta - \vartheta^2) \;=\; n\vartheta(1-\vartheta). \end{aligned}$$

Beispiel 1.8.5

Sei Z eine reelle Zufallsvariable auf (Ω, P) mit $P(Z > 0) = 1$. Weitern seien Z und $1/N$ integrierbar.

Setzen wir $X = \sqrt{Z}$ und $Y = 1/\sqrt{Z}$, dann erhalten wir nach der Ungleichung von Cauchy-Schwarz

$$1 \;=\; (EXY)^2 \;\leq\; EX^2 \cdot EY^2 \;=\; EZ \cdot E\frac{1}{Z},$$

also gilt

$$E\frac{1}{Z} \;\geq\; \frac{1}{EZ}$$

mit Gleichheit, wenn es ein $a > 0$ mit $P(Z = a) = 1$ gibt.

Satz 1.8.6 (Chebychev Ungleichung)

Sei X eine reelle Zufallsvariable auf (Ω, P) mit $EX^2 < \infty$.

Dann gilt für alle $\varepsilon > 0$

$$P(|X - EX| \geq \varepsilon) \;\leq\; \frac{\operatorname{Var}(X)}{\varepsilon^2}.$$

Beispiel 1.8.7

Sei $X : \Omega \to \Omega' = \{(x_1, .., x_n) \mid x_i \neq x_j \text{ für } i \neq j|\}$ und sei P^X Laplace verteilt auf Ω'. Weiter definieren wir

$$\mu \;=\; EX \;=\; \frac{1}{n}\sum_{i=1}^{n} x_i \qquad \text{und} \qquad \sigma^2 \;=\; \operatorname{Var}(X) \;=\; \frac{1}{n}\sum_{i=1}^{n}(x_i - \mu)^2.$$

Dann gilt

$$P^X(\{x \in \Omega' \mid |X - EX| \geq \varepsilon\}) \;=\; \frac{1}{n}|\{x \in \Omega' \mid |X - EX| \geq \varepsilon\}| \;\leq\; \frac{\sigma^2}{\varepsilon^2}.$$

Betrachten wir $\varepsilon = \sigma \cdot \delta$, dann folgt

$$|\{x \in \Omega' \mid |X - EX| \geq \delta\}| \;\leq\; \frac{n}{\delta^2}.$$

Die Wahl von $\delta = 2$ zeigt also, dass 75 % aller Elemente aus Ω' im Intervall

$$[\mu - 2\sigma,\ \mu + 2\sigma]$$

liegen.

Proposition 1.8.8

Sei $\mathbb{R}^+ = \{x \in \mathbb{R} \mid x \geq 0\}$, sei $f : \mathbb{R}^+ \to \mathbb{R}^+$ monoton wachsende und sei X eine reelle Zufallsvariable auf (Ω, P), so dass $f \circ |X|$ integrierbar ist.

Dann gilt für alle $\varepsilon > 0$

$$P(|X| \geq \varepsilon) \;\leq\; \frac{E(f(|X|))}{f(\varepsilon)}.$$

Folgerung 1.8.9 (Markov Ungleichung)

Sei X eine reelle Zufallsvariable auf (Ω, P).

Dann gilt für alle $\varepsilon > 0$

$$P(|X| \geq \varepsilon) \;\leq\; \frac{1}{\varepsilon} E(|X|).$$

Beispiel 1.8.10

Sei X eine Poisson verteilte Zufallsvariable mit Parameter λ. Dann gilt

$$E(|X|) \;=\; EX \;=\; \lambda$$

und für $\varepsilon = 2\lambda$ erhalten wir

$$P(X \geq 2\lambda) \;=\; P(|X| \geq 2\lambda) \;\leq\; \frac{1}{2\lambda} EX \;=\; \frac{1}{2}.$$

Definition 1.8.11

Seien $X_n, X : \Omega \to \mathbb{R}$ für $n \in \mathbb{N}$ reelle Zufallsvariablen auf Ω.

X_n konvergiert genau dann ***stochastisch*** gegen X, wenn für alle $\varepsilon > 0$

$$\lim_{n \to \infty} P(|X_n - X| \geq \varepsilon) \;=\; 0$$

gilt. Schreibweise:

$$X_n \;\xrightarrow{P}\; X.$$

Satz 1.8.12

Seien $X_n : \Omega \to \mathbb{R}$ reelle quadratintegrierbare Zufallsvariablen mit $\mu_n = EX_n$ und mit $\sigma_n^2 = \mathrm{Var}(X_n)$. Weiter seien die Zufallsvariablen X_n paarweise unkorreliert.

Gilt dann

$$\lim_{n \to \infty} \frac{1}{n^2} \sum_{i=1}^{n} \sigma_i^2 \;=\; 0,$$

so folgt

$$\frac{1}{n} \sum_{i=1}^{n} (X_i - \mu_i) \;\xrightarrow{P}\; 0.$$

Satz 1.8.13 (Schwaches Gesetz großer Zahlen)

Sei $(X_n)_{n\in\mathbb{N}}$ eine Folge von stochastisch unabhängigen Zufallsvariablen mit gleichen Erwartungswerten $\mu = EX$ und es gelte für alle $n \in \mathbb{N}$ gerade $EX_n^2 < \infty$.

Dann gilt für $n \to \infty$

$$\frac{1}{n}\sum_{i=1}^{n} X_i \quad \xrightarrow{P} \quad \mu.$$

Beispiel 1.8.14

Seien $(X_n)_{n\in\mathbb{N}}$ stochastisch unabhängig und identisch Bernoilli verteilt, also

$$P^{X_i} \;=\; \mathfrak{L}(1,\vartheta).$$

Dann folgt sofort für $n \to \infty$

$$\frac{1}{n}\sum_{i=1}^{n} X_i \quad \xrightarrow{P} \quad \vartheta.$$

Beispiel Bernsteinpolynome

Sei $I \subset \mathbb{R}$ ein abgeschlossenes Intervall, sei $f : I \to \mathbb{R}$ stetig, sei $\delta > 0$ und sei dazu

$$w(f,\delta) \;=\; \sup_{\substack{x,y\in I\\ |x-y|<\delta}} |f(x) - f(y)|$$

das ***Stetigkeitsmodul*** von f. Weiter definieren wir

$$w(f,\infty) \;=\; \sup_{x,y\in I} |f(x) - f(y)|.$$

Aus der Analysis ist bekannt, dass f genau dann gleichmäßig stetig auf I ist, wenn $w(f,\delta) \to 0$ für $\delta \to 0$ gilt. Sei $P^{S_n(t)} = \mathfrak{L}(n,t)$ mit $t \in [0,t]$. Dann heißen die Polynome

$$B_n(f,t) \;=\; Ef(S_n(t)/n) \;=\; \sum_{n=0}^{n} f\left(\frac{k}{n}\right) \cdot \binom{n}{k} t^k (1-t)^{n-k}$$

Bernsteinpolynome vom Grad n. Diese Polynome approximieren gerade die Funktion f, was auch der nächste Satz zeigt:

Satz 1.8.15

Sei $I \subset \mathbb{R}$ ein abgeschlossenes Interval und sei $f : I \to \mathbb{R}$ stetig. Dann gibt es eine Folge $(P_n(t))_{n\in\mathbb{N}}$ von Polynomen mit

$$\sup_{t\in I} |f(t) - P_n(t)| \quad \to \quad 0 \qquad \text{für } \; n \to \infty.$$

Jede stetige Funktion lässt sich also beliebig nahe durch ein Polynom approximieren.

1.9 Ganzzahlige Zufallsvariablen

Wie betrachten nun stets einen diskreten Wahrscheinlichkeitsraum (Ω, P) mit einer ganzzahligen Zufallsvariablen $X : \Omega \to \mathbb{N} \cup \{0\}$.

Definition 1.9.1

Sei f_X die Zähldichte einer ganzzahligen Zufallsvariablen X. Dann definieren wir die ***erzeugenden Funktion*** von X durch

$$M_X(s) \;=\; Es^X \;=\; \sum_{n=0}^{\infty} f_X(n)s^n.$$

Für weitere Untersuchungen von erzeugenden Funktionen benötigen wir einige Grundlagen zu Potenzreihen:

Bemerkung 1.9.2 (Potenzreihen)

Seien $(a_n)_{n\geq 0}$ eine reelle Folge. Dann ist

$$A(s) \;=\; \sum_{n=0}^{\infty} a_n s^n$$

eine Potenzreihen. Diese Reihe konvergiert für alle $s \in]-R, R[$, dabei ist R der Konvergenzradius dieser Potenzreihe. Es gilt:

(1) $A(s)$ ist auf $]-R, R[$ stetig und konvergiert absolut sowie gleichmäßig.

(2) Für $s \in]-R, R[$ gilt

$$\frac{\mathrm{d}}{\mathrm{d}s} A(s) \;=\; \sum_{n=1}^{\infty} n a_n s^{n-1}.$$

(3) ***Abelscher Konvergenzsatz***: Aus $a_n \geq 0$ für alle $n \geq 0$ folgt

$$\lim_{s\to R} A(s) \;=\; A(R) \;\leq\; \infty.$$

(4) Gilt für abzählbar viele $s \in]-\varepsilon, \varepsilon[$ für alle $\varepsilon > 0$ gerade

$$\sum_{n=0}^{\infty} a_n s^n = \sum_{n=0}^{\infty} b_n s^n,$$

dann folgt $a_n = b_n$ für alle $n \geq 0$.

(5) Seien $a = (a_n)_{n \geq 0}$ und $b = (b_n)_{n \geq 0}$ zwei reelle Folgen. Dann wird die ***Faltung*** gegeben durch

$$c = (c_n)_{n \geq 0} \qquad \text{mit} \qquad c_n = a_0 b_n + a_1 b_{n-1} + \ldots + a_n b_0.$$

Damit gilt gerade

$$\left(\sum_{n=0}^{\infty} c_n s^n \right) = \left(\sum_{n=0}^{\infty} a_n s^n \right) \cdot \left(\sum_{n=0}^{\infty} b_n s^n \right)$$

für alle $|s| < \min\{R_a, R_b\}$, wobei R_a und R_b die Konvergenzradien der entsprechenden Potenzreihen sind.

Aus diesen Überlegungen erhalten wir nun die folgenden Aussagen:

Bemerkung 1.9.3

(1) Es gilt

$$\sum_{k=0}^{n} \binom{n}{k}^2 = \binom{2n}{n}.$$

(2) Eine erzeugenden Funktion $M_X(s)$ hat immer einen Konvergenzradius $R \geq 1$, da

$$M_X(s) = \sum_{n=0}^{\infty} f_X(n) s^n \qquad \text{und} \qquad \sum_{n=0}^{\infty} f_X(n) = 1.$$

Defintion 1.9.4

Sei Q ein Wahrscheinlichkeitsmaß auf $\mathbb{N} \cup \{0\}$. Dann ist

$$M_Q(s) = \sum_{n=0}^{\infty} f_Q(n) s^n$$

die ***erzeugenden Funktion*** von Q.

Proposition 1.9.5

(1) Jede erzeugenden Funktion $M_X(s)$ ist stetig, monoton wachsend und konvex auf $[0, R[$. Weiter gilt

$$M_X(0) \;=\; f_X(0) \qquad \text{sowie} \qquad M_X(1) \;=\; 1.$$

(2) Es gilt

$$EX \;=\; \lim_{s\to 1} \frac{\mathrm{d}}{\mathrm{d}s} M_X(s) \;=\; M_X'(1) \;\leq\; \infty.$$

(3) Für das k te faktorielle Moment gilt

$$E(X(X-1)..(X-k+1)) \;=\; \lim_{s\to 1} \frac{\mathrm{d}^k}{\mathrm{d}s^k} M_X(s) \;=\; M_X^{(k)}(1) \;\leq\; \infty.$$

(4) Für die Varianz gilt

$$\mathrm{Var}(X) \;=\; M_X''(1) + M_X'(1) - (M_X'(1))^2.$$

Beispiel Binomialverteilung

Sei $P^X = \mathfrak{L}(n, \vartheta)$ Binomial verteilt mit Parameter ϑ. Dann gilt nach der Binomischen Formel

$$\begin{aligned} M_X(s) &= \sum_{k=0}^{n} \binom{n}{k} \vartheta^k (1-\vartheta)^{n-1} \cdot s^k \;=\; \sum_{k=0}^{n} \binom{n}{k} (\vartheta s)^k (1-\vartheta)^{n-1} \\ &= (1 - \vartheta + \vartheta s)^n \end{aligned}$$

für alle $s \in \mathbb{R}$. Damit folgt

$$\begin{aligned} M_X'(s) &= n\vartheta(1-\vartheta+\vartheta s)^{n-1}, \\ M_X''(s) &= n(n-1)\vartheta^2(1-\vartheta+\vartheta s)^{n-2} \end{aligned}$$

und wir erhalten sofort die bekannten Ergebnisse

$$\begin{aligned} EX &= M_X'(1) \;=\; n\vartheta, \\ \mathrm{Var}(X) &= n(n-1)\vartheta^2 + n\vartheta - n^2\vartheta^2 \;=\; n\vartheta(1-\vartheta). \end{aligned}$$

Proposition 1.9.6

Seien $N, X, Y, X_i : \Omega \to \mathbb{N} \cup \{0\}$ für $i \in \mathbb{N}$ ganzzahlige Zufallsvariablen auf (Ω, P).

(1) Sind X und Y stochastisch unabhängig, dann gilt

$$M_{X+Y}(s) \;=\; M_X(s) \cdot M_Y(s).$$

(2) Sind $\{X_i\}$ stochastisch unabhängig und identisch verteilt sowie auch N und $\{X_i\}$ stochastisch unabhängig, dann gilt

$$M_S(t) \;=\; M_N(M_{X_1}(t)),$$

dabei ist $S = \sum_{n=1}^{N} X_i$ die Summe der Zufallsvariablen.

Beispiel Binomialverteilung

Seien $X_1, .., X_n$ stochastsich unabhängige Zufallsvariablen, die identisch Bernoulli verteilt sind, also $P^{X_k} = \mathcal{L}(1, \vartheta)$. Sei weiter

$$S_n \;=\; \sum_{k=1}^{n} X_k.$$

Wir haben für $k = 1, .., n$

$$M_{X_k}(t) \;=\; Et^{X_k} \;=\; 1 - \vartheta + \vartheta t.$$

Damit folgt nach 1.9.6 wieder das bekannte Ergebnis

$$M_{S_n}(t) \;=\; \prod_{k=1}^{n} M_{X_k}(t) \;=\; (1 - \vartheta + \vartheta t)^n$$

Dieses Beispiel zeigt auch, dass die Darstellung eindeutig ist.

Beispiel Poisson Verteilung

Sei X Poisson verteilt mit Parameter λ, also $P^X = \mathcal{P}(\lambda)$. Dann gilt

$$M_X(t) \;=\; Et^X \;=\; \sum_{k=0}^{\infty} \frac{\lambda^k}{k!} \mathrm{e}^{-\lambda} t^k \;=\; \mathrm{e}^{\lambda t - \lambda}.$$

Seien nun $X_1, .., X_n$ stochastsich unabhängige Zufallsvariablen, die Poisson verteilt sind, also $P^{X_k} = \mathcal{P}(\lambda_k)$. Sei weiter

$$S_n \;=\; \sum_{k=1}^{n} X_k.$$

Dann gilt wieder nach 1.9.6

$$M_{S_n}(t) \;=\; \prod_{k=1}^{n} M_{X_k}(t) \;=\; \prod_{k=1}^{n} \mathrm{e}^{\lambda_k (t-1)} \;=\; \mathrm{e}^{\left(\sum_{k=1}^{n} \lambda_k\right) \cdot (t-1)},$$

also folgt

$$P^{S_n} \;=\; \mathcal{P}\left(\sum_{k=1}^{n} \lambda_k\right).$$

Beispiel radioaktiver Zerfall

Eine radioaktive Probe sendet pro Minute N neongrüne α Teilchen aus. Die Wahrscheinlichkeit, dass eine neongrünes α Teilchen vom Messgerät erfasst wird, sei ϑ. Gesucht ist die Verteilung der registrierten α Teilchen.

N sei dazu Poisson verteilt mit Parameter λ, also $P^N = \mathcal{P}(\lambda)$. Die Zufallsvariablen X_k entsprechen dem k ten Teilchen, das gesendet wird. Diese Zufallsvarialben sind also Bernoulli verteilt mit Parameter ϑ: $P^{X_k} = \mathcal{L}(1, \vartheta)$. Weiter sei

$$S_n = \sum_{k=1}^{n} X_k$$

die Anzahl der registrierten Teilcen. Dann erhalten wir nach 1.9.6

$$\begin{aligned} M_S(t) &= M_N(M_{X_1}(t)) = M_N(1 - \vartheta + \vartheta t) \\ &= \mathrm{e}^{\lambda(1-\vartheta+\vartheta t-1)} = \mathrm{e}^{\lambda\vartheta(1-t)}, \end{aligned}$$

somit ist $P^S = \mathcal{P}(\lambda\vartheta)$.

Satz 1.9.7 (Waldsche Identität)

Seien $\{X_i\}$ stochastisch unabhängig und identisch verteilt sowie auch $N \geq 1$ und $\{X_i\}$ stochastisch unabhängig.

Dann gilt für $S_n = \sum_{k=1}^{N} X_k$

$$ES = EN \cdot EX_1.$$

Satz 1.9.8 (Stetigkeitssatz)

Seien Q_n für $n \in \mathbb{N} \cup \{0\}$ Wahrscheinlichkeitsmaße auf $n \in \mathbb{N} \cup \{0\}$. Dann sind die folgenden Aussagen äquivalent:

(1) $\lim_{n\to\infty} Q_n(\{k\}) = Q_0(\{k\})$ für alle $k \in \mathbb{N} \cup \{0\}$.

(2) $\lim_{n\to\infty} M_{Q_n}(t) = M_{Q_0}(\{t\})$ für alle $t \in]0, 1[$.

Auch mit diesem Satz erhalten wir noch einmal die Poissonapproximation:

Korollar 1.9.9 (Poissonapproximation)

Seien $Q_n = \mathcal{L}(n, \vartheta_n)$ mit $\vartheta_n \in [0, 1[$ und es gelte $n\vartheta \to \lambda > 0$ für $n \to \infty$.

Dann gilt

$$\lim_{n\to\infty} Q_n(\{k\}) = Q_0(\{k\}) = \frac{\lambda^k}{k!}\mathrm{e}^{-\lambda} \qquad \text{für } k \in \mathbb{N} \cup \{0\},$$

somit ist $Q_0 = \mathcal{P}(\lambda)$.

Verteilungsprozesse

Wie betrachten die Entwicklung in Generationen.

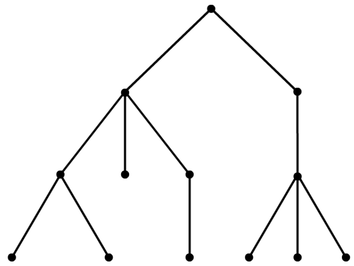


Abbildung 1.11: Beispiel eines Verteilungsprozesses.

Dazu definieren wir die Zufallsvariablen Z_i, die die Anzahl der Mitglieder in der i ten Generation beschreibt. Es wird angenommen, dass jedes Mitglied unabhängig voneinander dieselbe Reproduktionsverteilung mit der Zähldichte f und der erzeugenden Funktion G besitzt. Wir starten mit $Z_0 = 1$.

Es werden also $(X_n^k)_{n\in\mathbb{N}}$ unabhängige gleichverteilte Zufallsvariablen auf (Ω, P) mit Zähldichte f und Werten in $\mathbb{N} \cup \{0\}$ gestartet. Dazu definieren wir

$$Z_{k+1} \;=\; \sum_{n=1}^{Z_k} X_n^k$$

für die k te Generation. Für das Beispiel aus Abbildung 1.11 haben wir also

$$\begin{array}{llll}
Z_0 = 1 & X_1^0 = 2 & & \\
Z_1 = 2 & X_1^1 = 3, & X_2^1 = 1 & \\
Z_2 = 4 & X_1^2 = 2, & X_2^2 = 2, & X_3^2 = 1, \quad X_4^2 = 3 \\
\vdots & & &
\end{array}$$

Die Folge (Z_k) heißt ***Verteilungsprozess*** oder ***Gallon-Watson Prozess***. Wir definieren die erzeugenden Funktion G_k von Z_k also

$$G_k(s) \;=\; M_{Z_k}(s)$$

und erhalten $G_1 = G$. Damit lässt sich zeigen, dass für den Erwartungswert $\mu = EZ_1$ und die Varianz $\sigma^2 = \mathrm{Var}(Z_1)$ gerade

$$EZ_k \;=\; \mu^k \qquad \text{und} \qquad \mathrm{Var}(Z_k) \;=\; \begin{cases} k\sigma^2 & \text{für} \quad \mu = 1 \\ \frac{\sigma^2(\mu^k-1)}{\mu-1}\mu^{k-1} & \text{für} \quad \mu \neq 1 \end{cases}$$

gilt. Betrachten wir den Fall, dass die X_n^k geometrisch verteilt sind, dann gilt

$$f_{X_n^k}(k) \;=\; (1-\vartheta)\vartheta^k \qquad \text{für} \;\; k \;\geq\; 0$$

mit der erzeugenden Funktion

$$\begin{aligned} G(s) &= \sum_{k=0}^{\infty}(1-\vartheta)\vartheta^k s^k = (1-\vartheta)\frac{1}{1-\vartheta s}, \\ G'(s) &= (1-\vartheta)\frac{\vartheta}{(1-\vartheta s)^2}. \end{aligned}$$

Wir erhalten also den Erwartungswert

$$\mu = G'(1) = \frac{\vartheta}{1-\vartheta}.$$

Induktiv erhalten wir

$$G_k(s) = \begin{cases} \frac{k-(k-1)s}{\mu+1-ks} & \text{für } \vartheta = \frac{1}{2} \\ \frac{(1-\vartheta)\left(\vartheta^k-(1-\vartheta)^k-\vartheta s(\vartheta^{k-1}-(1-\vartheta)^{k-1})\right)}{\vartheta^{k+1}-(1-\vartheta)^{k+1}-\vartheta s(\vartheta^k-(1-\vartheta)^k)} & \text{für } \vartheta \neq \frac{1}{2} \end{cases}.$$

Damit kann nun die Aussterbewahrscheinlichkeit zum Zeitpunkt k berechnet werden, denn es gilt

$$P(Z_k = 0) = G_k(0) = \begin{cases} \frac{k}{k+1} & \text{für } \vartheta = \frac{1}{2} \\ \frac{(1-\vartheta)(\vartheta^k-(1-\vartheta)^k)}{\vartheta^{k+1}-(1-\vartheta)^{k+1}} & \text{für } \vartheta \neq \frac{1}{2} \end{cases}.$$

Nach dem Stetigkeitssatz erhalten wir nun die gesuchte Aussterbewahrscheinlichkeit η:

$$\begin{aligned} \eta &= \lim_{k\to\infty} P(Z_k = 0) = P\left(\lim_{k\to\infty}(Z_k = 0)\right) \\ &= \lim_{k\to\infty} \begin{cases} \frac{k}{k+1} & \text{für } \vartheta = \frac{1}{2} \\ \frac{(1-\vartheta)(\vartheta^k-(1-\vartheta)^k)}{\vartheta^{k+1}-(1-\vartheta)^{k+1}} & \text{für } \vartheta \neq \frac{1}{2} \end{cases} = \begin{cases} 1 & \text{für } \vartheta \leq \frac{1}{2} \\ \frac{1-\vartheta}{\vartheta} & \text{für } \vartheta > \frac{1}{2} \end{cases}. \end{aligned}$$

Die Aussterbewahrscheinlichkeit ist gleich 1, wenn $\mu \leq 1$, wenn es im Mittel also höchstens einen Nachkommen gibt. Für $\vartheta = 3/4$ erhalten wir die Aussterbewahrscheinlichkeit $\eta = 1/3$.

Allgemein gilt der folgende Satz:

Satz 1.9.10

Sei $(Z_n)_{n\in\mathbb{N}}$ ein Verzweigungsprozess mit $P(Z_1 = 1) < 1$ und sei

$$\eta = \lim_{n\to\infty} P(Z_n = 0)$$

die Aussterbewahrscheinlichkeit. Weiter sei G die erzeugenden Funktion wie zuvor.

Dann gilt:

(1) Ist $\mu = EZ_1 \leq 1$, dann folgt $\eta = 1$.

(2) Ist $\mu = EZ_1 > 1$, dann ist η die eindeutige Lösung der Fixpunktgleichung $G(s) = s$ für $0 \leq s < 1$.

1.10 Aufgaben

Aufgabe 1.10.1

Berechne die Wahrscheinlichkeit dafür, dass beim dreifachen Würfeln eines fairen Würfels ...

(1) ... das Maximum gleich 4 ist.

(2) ... das Minimum kleiner oder gleich 4 ist.

Lösung Teil 1

Sei $\Omega = \{1, .., 6\}^3$, also $|\Omega| = 216$, und sei P die Laplace Verteilung auf Ω. Weiter sei A die Menge aller Möglichen Würfe, bei denen das Maximum gleich 4 ist. Für diese Menge gilt

$$A \;=\; \frac{\{1,2,3,4\}^3}{\{1,2,3\}^3}, \qquad \text{also} \qquad |A| \;=\; 4^3 - 3^3 \;=\; 37.$$

Wir erhalten also die gesuchte Wahrscheinlichkeit von

$$P(A) \;=\; \frac{|A|}{|\Omega|} \;=\; \frac{37}{216} \;\approx\; 17.2\,\%.$$

Lösung Teil 2

Es sei Ω wie in Aufgabenteil 1 und nun sei B die Menge der möglichen Würfe, bei denen das Minimum kleiner oder gleich 4 ist. Damit gilt $B^C = \{5, 6\}^3$ mit $|B^C| = 2^3 = 8$, somit erhalten wir die gesuchte Wahrscheinlichkeit

$$P(B) \;=\; 1 - P(B^C) \;=\; 1 - \frac{8}{216} \;=\; \frac{26}{27} \;\approx\; 96{,}3\,\%.$$

Aufgabe 1.10.2

Beim Skatspiel erhält jeder der drei Spieler 10 Karten, während die restlichen 2 Karten in den Skat gelegt werden.

(1) Berechne die Anzahl der Arten, auf welche die 32 Karten verteilt werden können.

(2) Berechne die Wahrscheinlichkeit dafür, dass der Kreuzbube im Skat liegt, wenn keine weiteren Vorinformationen vorliegen.

(3) Berechne die Wahrscheinlichkeit dafür, dass genau ein Bube im Skat liegt, wenn keine weiteren Vorinformationen vorliegen.

(4) Berechne die Wahrscheinlichkeit dafür, dass zwei Buben im Skat liegt, wenn keine weiteren Vorinformationen vorliegen.

Lösung Teil 1

Die 32 Karten können auf

$$\binom{32}{10}\binom{22}{10}\binom{12}{10} \approx 2.75 \cdot 10^{15}$$

Arten verteilt werden, denn der erste Spieler zieht 10 aus 32, der zweite Spieler 10 aus 22 und der dritte Spieler 10 aus 12. Die 2 Karten im Skat sind danach eindeutig festgelegt.

Lösung Teil 2

Es sei $M = \{1, .., 32\}$ und

$$\Omega = \{W \subset M \mid |W| = 2\}, \qquad \text{also} \qquad |\Omega| = \binom{32}{2} = 496.$$

Damit ergibt sich die Wahrscheinlichkeit P_1 dafür, dass der Kreuzbube im Skat liegt, nach

$$P_1 = \frac{\binom{1}{1}\binom{31}{1}}{\binom{32}{2}} = \frac{1}{16} = 6.25\,\%.$$

Lösung Teil 3

Es Ω wie im Aufgabenteil zuvor. Die Wahrscheinlichkeit P_2 dafür, dass genau ein Bube im Skat liegt, ist nun

$$P_1 = \frac{\binom{4}{1}\binom{28}{1}}{\binom{32}{2}} = \frac{7}{31} \approx 22.6\,\%.$$

Lösung Teil 4

Es Ω wie in den Aufgabenteilen zuvor. Die Wahrscheinlichkeit P_3 dafür, dass nun zwei Buben im Skat liegt, ist

$$P_1 = \frac{\binom{4}{2}\binom{28}{0}}{\binom{32}{2}} = \frac{3}{248} \approx 2.4\,\%.$$

Aufgabe 1.10.3

In den USA gibt es 52 Bundesstaaten, von denen jeder 2 Senatoren stellt. Aus diesen werden zufällig 52 ausgewählt.

(1) Berechne die Wahrscheinlichkeit dafür, dass ein fester Staat vertreten ist.

(2) Berechne die Wahrscheinlichkeit dafür, dass jeder Staat vertreten ist.

Lösung Teil 1

Wir betrachten die Laplace Verteilung P auf der Menge $\Omega = \mathrm{Kom}_{52}^{104}$. Sei A die Menge der Möglichkeiten, dass ein fester Staat vertreten ist. Damit ergibt sich nun die gesuchte Wahrscheinlichkeit nach

$$P(A) \;=\; \frac{|A|}{|\Omega|} \;=\; \frac{\binom{102}{51}\binom{2}{1} + \binom{102}{50}\binom{2}{2}}{\binom{104}{52}} \;=\; \frac{155}{206} \;\approx\; 75.2\,\%.$$

Lösung Teil 2

Wir betrachten wieder die Laplace Verteilung P auf der Menge $\Omega = \mathrm{Kom}_{52}^{104}$ und nun sei B die Menge der Möglichkeiten, dass jeder Staat einen Senator stellt, dies sind zwei für jeden Staat, also $|B| = 2^{52}$. Wir erhalten also die gesuchte Wahrscheinlichkeit

$$P(B) \;=\; \frac{|B|}{|\Omega|} \;=\; \frac{2^{52}}{\binom{104}{52}} \;=\; 2.8 \cdot 10^{-15}.$$

Aufgabe 1.10.4

In einer Urne sind N Kugeln enthalten, die von 1 bis N durchnumeriert sind. Aus dieser Urne werden n Kugeln mit Rücklegen gezogen. Berechne die Wahrscheinlichkeit dafür, dass die dabei gezogenen Kugelnummern strikt wachsen.

Lösung

Es sei $\Omega = \{1, .., N\}$. Die Menge der möglichen Fälle ist damit einfach Ω^n mit $|\Omega^n| = N^n$. Sei weiter A die Menge dafür, dass die gezogenen Kugelnummern strikt wachsen. Für diese Menge gilt nun

$$|A| \;=\; \frac{\mathrm{Perm}_N^n(\Omega)}{n!} \;=\; \frac{(n)_N}{n!} \;=\; \binom{N}{n}.$$

Damit erhalten wir die gesuchte Wahrscheinlichkeit

$$P(A) \;=\; \frac{|A|}{|\Omega|} \;=\; \frac{\binom{N}{n}}{N^n}.$$

Aufgabe 1.10.5

Eine Urne enthalte 30 Kugeln, 5 weiße, 10 blaue und 15 rote. Drei Kugeln werden zufällig ohne zurücklegen gezogen.

(1) Berechne die Wahrscheinlichkeit $P(A)$, dass alle drei Kugeln weiß sind.

(2) Berechne die Wahrscheinlichkeit $P(B)$, dass mindestens eine Kugel weiß ist.

(3) Berechne die Wahrscheinlichkeit $P(C)$, dass eine Kugel weiß, eine blau und eine rot ist.

Lösung Teil 1

Es sei P die Laplace Verteilung. Damit erhalten wir

$$P(A) \;=\; \frac{\binom{5}{3}\cdot\binom{25}{0}}{\binom{30}{3}} \;=\; \frac{1}{406} \;\approx\; 0.25\,\%.$$

Lösung Teil 2

Wir erhalten

$$P(B) \;=\; \frac{\binom{5}{1}\cdot\binom{25}{2}}{\binom{30}{3}} + \frac{\binom{5}{2}\cdot\binom{25}{1}}{\binom{30}{3}} + \frac{\binom{5}{3}\cdot\binom{25}{0}}{\binom{30}{3}} \;=\; \frac{22}{203} \;\approx\; 43.35\,\%.$$

Lösung Teil 3

Wir erhalten

$$P(C) \;=\; \frac{\binom{5}{1}\cdot\binom{10}{1}\cdot\binom{15}{1}}{\binom{30}{3}} \;=\; \frac{75}{406} \;\approx\; 18.47\,\%.$$

Aufgabe 1.10.6

Die Ziffern $2, 3, 4, 5, 6, 7$ und 8 werden zufällig in einer Reihenfolge aufgeschrieben. Bestimme die Wahrscheinlichkeiten dafür, dass die so gebildete siebenstellige Zahl durch 2, durch 3 oder durch 4 teilbar ist.

Lösung

Sei P die Laplace Verteilung und sei A das Ereignis, dass die gebildete siebenstellige Zahl durch 2 ist. Dies ist genau dann der Fall, wenn die Zahl mit $2, 4, 6$ oder mit 8 endet. Wir erhalten also

$$P(A) \;=\; \frac{4}{7}.$$

Sei B das Ereignis, dass die gebildete siebenstellige Zahl durch 3 ist. Eine Zahl ist genau dann durch 3 teilbar, wenn ihre Quersumme durch 3 teilbar ist. Die Quersumme unserer siebenstelligen Zahl ist immer 35 und somit nicht durch 3 teilbar, also

$$P(B) \;=\; 0.$$

Sei nun C das Ereignis, dass die siebenstellige Zahl durch 4 ist. Dazu muss die aus den letzten beiden Ziffern der siebenstellige Zahl untersucht werden. Für die letzten beiden Ziffern haben wir genau $6 \cdot 7 = 42$ Möglichkeiten. Davon sind die folgenden Möglichkeiten durch 4 teilbar:

$$N = \{24, 28, 32, 36, 46, 52, 56, 64, 68, 72, 76, 84\}.$$

Damit ergibt sich die gesuchte Wahrscheinlichkeit von

$$P(C) = \frac{|N|}{42} = \frac{14}{42} = \frac{2}{7}.$$

Aufgabe 1.10.7

Die Ecken eines Würfels sind gleichmäßig schräg abgeschliffen, so dass der Würfel auch seinen Seite nun Ecken liegen kann. Die Wahrscheinlichkeit jeder Ecke sei $1/4$ mal so groß wie die einer Seite. Berechne die Wahrscheinlichkeit dafür, dass eine 6 gewürfelst wird.

Lösung

Der Würfel hat 8 Ecken und 6 Seiten. Sei P die Laplace Verteilung und A das Ereignis, dass der Würfel auf der 3 liegenbleibt. Gewichten wir nun die Seiten mit 4 und die Ecken mit 1, so erhalten wir

$$P(A) = \frac{4}{8 \cdot 1 + 6 \cdot 4} = \frac{1}{8}.$$

Aufgabe 1.10.8

Unter 32 Karten befinden sich 4 Asse. Alle Karten werden nebeneinander hingelegt. Bestimme die Wahrscheinlichkeit dafür, dass die neunte Karte das zweite Ass ist.

Lösung

Sei P die Laplace Verteilung und A das Ereignis, dass unter den ersten 8 aufgedeckten Karten genau ein Ass ist. Dann erhalten wir nach der hypergeometrischen Verteilung gerade

$$P(A) = \frac{\binom{4}{1} \cdot \binom{28}{7}}{\binom{32}{8}} = \frac{4 \cdot 1184040}{10518300} \approx 0.4503\,.$$

Sei B nun das Ereignis dafür, dass die neunte Karte genau das zweite Ass ist. Dann erhalten wir

$$P(B) = P(A) \cdot \frac{3}{24} \approx 0.0563 = 5.63\,\%,$$

da wir noch genau 3 Asse in 24 Karten übrig haben.

Aufgabe 1.10.9

Es sei $\mathbb{R}[x_1, .., x_n]$ der Polynomring mit $n \geq 2$ Variablen. Ein ***Monom*** vom Grad m ist ein Polynom der Form

$$x_1^{i_1} \cdot \ldots \cdot x_n^{i_n}$$

mit $\sum_{j=1}^{n} i_j = m$ und $i_j \in \mathbb{N} \cup \{0\}$ für $j = 1, .., n$.

Berechne die Anzahl der Monome vom Grad m in $\mathbb{R}[x_1, .., x_n]$.

Lösung

Wir untersuchen also alle Polynome der Form

$$x_1^{i_1} \cdot \ldots \cdot x_n^{i_n}$$

mit $\sum_{j=1}^{n} i_j = m$. Betrachten wir das Zellenmodell mit n Zellen, auf die m ununterscheidbare Kugeln mit Mehrfachbesetztung verteilt werden. Dieses Modell beschreibt genau die Situation der Monome vom Grad m mit n Variablen. Somit gibt es nach dem Zellenmodell genau

$$\binom{n+m-1}{m}$$

derartige Monome.

Aufgabe 1.10.10

Drei Kartenspiele mit jeweils n verschiedenen Karten werden getrennt gemischt und in drei Reihen hingelegt:

$$\begin{array}{cccccc l} \square & \square & \square & \square & \ldots & \square & \text{1. Spiel} \\ \square & \square & \square & \square & \ldots & \square & \text{2. Spiel} \\ \square & \square & \square & \square & \ldots & \square & \text{3. Spiel} \end{array}$$

Berechne die Wahrscheinlichkeit q_n, dass wenigstens eine der Spalten aus gleichen Karten zusammengesetzt ist.

Lösung

Zunächst sei

$$\Omega \;=\; \{\pi \mid \pi : \{1, .., n\} \to \{1, .., n\},\ \pi \text{ bijektiv}\}$$

die Menge aller n Permutationen. Da wir drei Kartenspiele mit jeweils n Karten haben, betrachten wir die Menge Ω^3 mit $|\Omega^3| = (n!)^3$.

Die gesuchte Wahrscheinlichkeit q_n ist nun $P(A)$ mit

$$A \;=\; \{(\pi_1, \pi_2, \pi_3) \in \Omega^3 \mid \pi_1(i) = \pi_2(i) = \pi_3(i) \text{ für mindestens ein } i\},$$

dabei ist P die Laplace Verteilung. Betrachten wir weiter für $i = 1, .., n$ die Mengen

$$B_i \;=\; \{(\pi_1, \pi_2, \pi_3) \in \Omega^3 \mid \pi_1(i) = \pi_2(i) = \pi_3(i)\},$$

dann gilt $A = \bigcup_{i=1}^n B_i$. Um

$$q_n \;=\; P(A) \;=\; P\left(\bigcup_{i=1}^n B_i\right)$$

zu berechnen, kann also die Siebformel angewendet werden. Es gilt für alle $1 \leq i \leq n$, $i < j \leq n$ und $j < k \leq n$ nun

$$\begin{aligned} P(B_i) &= \frac{1}{n^2}, \\ P(B_i \cap B_j) &= \frac{1}{n^2(n-1)^2}, \\ P(B_i \cap B_i \cap B_k) &= \frac{1}{n^2(n-1)^2(n-2)^2} \end{aligned}$$

und so weiter. Nach der Siebformel erhalten wir also

$$\begin{aligned} q_n \;=\; P(A) \;=\; P\left(\bigcup_{i=1}^n B_i\right) & \\ = \frac{1}{n^2} - \binom{n}{2}\frac{1}{n^2(n-1)^2} + \binom{n}{3}\frac{1}{n^2(n-1)^2(n-2)^2} - \ldots + (-1)^{n-1}\frac{1}{(n!)^2} & \\ = \binom{n}{1}\frac{((n-1)!)^2}{(n!)^2} - \binom{n}{2}\frac{((n-2)!)^2}{(n!)^2} + \ldots + (-1)^{n-1}\binom{n}{n}\frac{((n-n)!)^2}{(n!)^2} & \\ = \sum_{k=1}^n (-1)^{k-1}\binom{n}{k}\frac{((n-k)!)^2}{(n!)^2} \;=\; \sum_{k=1}^n (-1)^{k-1}\frac{(n-k)!}{n! \cdot k!}. & \end{aligned}$$

Für kleine n erhalten wir damit die Wahrscheinlichkeiten, die in der folgenden Tabelle 1.4 gezeigt sind:

n	1	2	3	4	5	6	7	8
q_n	1	$\frac{1}{4}$	$\frac{5}{18}$	$\frac{41}{192}$	$\frac{71}{400}$	$\frac{3137}{20736}$	$\frac{557563}{4233600}$	$\frac{9022171}{77414400}$
in %	100	25	27.8	21.3	17.8	15.1	13.2	11.7

Tabelle 1.4: Wahrscheinlichkeiten q_k für kleine n.

Weiter gilt

$$P(A) \;=\; P\left(\bigcup_{i=1}^n B_i\right) \;\leq\; \sum_{i=1}^n P(B_i) \;=\; \sum_{i=1}^n \frac{1}{n^2} \;=\; \frac{1}{n}.$$

Für $n \to \infty$ ist die Wahrscheinlichkeit dafür, dass eine Spalte aus gleichen Karten zusammengesetzt ist, gleich 0.

Aufgabe 1.10.11

Es werden zwei Würfel geworfen, die die Ziffern $1, .., 6$ tragen. Es bezeichne r_i für $i = 2, .., 12$ die Wahrscheinlichkeit dafür, dass die Augensumme der beiden Würfel gleich i ist.

Zeige, dass es keine zwei verfälschten Würfel gibt, so dass $r_2 = .. = r_{12}$ gilt.

Lösung

Angenommen es gäbe zwei solche verfälschten Würfel, dann würde

$$r_2 \;=\; \ldots \;=\; r_{12} \;>\; 0$$

gelten. Sei $P_i(k)$ die Wahrscheinlichkeit dafür, dass mit dem i ten Würfel die Zahl k gewürfelt wird. Dann gilt $r_2 = P_1(1) \cdot P_2(1)$. Somit sind $P_1(1)$ und $P_2(1)$ größer 0. Weiter gilt $r_{12} = P_1(6) \cdot P_2(6)$, also sind auch $P_1(6)$ und $P_2(6) > 0$. Es gilt nun

$$\begin{aligned} r_7 \;&=\; P_1(1) \cdot P_2(6) + P_1(6) \cdot P_2(1) + P_1(2) \cdot P_2(5) + P_1(5) \cdot P_2(2) \\ &\quad +P_1(3) \cdot P_2(4) + P_1(4) \cdot P_2(3). \end{aligned}$$

Daraus erhalten wir

$$r_7 \;\geq\; P_1(1) \cdot P_2(6) + P_1(6) \cdot P_2(1) \;\leq\; r_2 \;=\; r_6$$

und es folgt

$$\frac{P_1(1) \cdot P_2(6)}{P_1(6)} + P_2(1) \;\leq\; P_2(6).$$

Somit gilt aber auch $P_2(1) < P_2(6)$, da der linke Summand > 0 ist. Analog erhalten wir auch $P_2(1) > P_2(6)$, was zum Widerspruch führt. Es gibt also keine zwei erfälschten Würfel mit $r_2 = .. = r_{12}$.

Aufgabe 1.10.12

Es werden drei faire Würfel geworfen. Berechne die bedingte Wahrscheinlichkeit dafür, dass unter der Vorinformation, dass keine zwei Würfel die gleiche Augenzahl haben, eine 6 gewürfelt wird.

Lösung

Es seien B die Möglichkeiten der Vorinformation, dass also keine zwei Würfel die gleiche Augenzahl haben. Es gilt damit $|B| = 6 \cdot 5 \cdot 4 = 120$.

Weiter seien A die Möglichkeiten, dass eine 6 gewürfelt wird. Die Anzahl der Möglichkeiten, dass eine 6 gewürfelt wird, aber keine zwei Würfel die gleiche Augenzahl haben, ist $5 \cdot 4 + 5 \cdot 4 + 5 \cdot 4 = 60$, also gilt $|A \cap B| = 60$.

Damit erhalten wir die gesuchte bedingte Wahrscheinlichkeit

$$P(A|B) = \frac{P(A\cap B)}{P(B)} = \frac{\frac{60}{6^3}}{\frac{120}{6^3}} = \frac{\frac{60}{216}}{\frac{120}{216}} = \frac{1}{2} = 50\,\%.$$

Aufgabe 1.10.13

Zwei Schützen schießen abwechselnd auf ein Ziel, bis einer trifft. Die Trefferwahrscheinlichkeit pro Schuß ist bei Schütze 1 gleich p_1 und bein Schütze 2 gleich p_2. Schütze 1 beginnt das Duell.

Berechne die Wahrscheinlichkeiten dafür, dass Schütze 1 bzw. Schütze 2 gewinnt. Gib weiter Bedingungen für p_1 und p_2 dafür an, dass beide Schützen die gleiche Siegeswahrscheinlichkeit haben.

Lösung

Zunächst nehmen wir an, dass $0 < p_1, p_2 < 1$ gilt.

Die folgende Tabelle 1.5 zeigt die Wahrscheinlichkeiten dafür, dass Schütze 1 bzw. Schütze 2 im jeweiligen Versuch das Duell gewinnt:

	Schütze 1	Schütze 2
1. Versuch	p_1	$(1-p_1)p_2$
2. Versuch	$(1-p_1)(1-p_2)p_1$	$(1-p_1)^2(1-p_2)p_2$
3. Versuch	$(1-p_1)^2(1-p_2)^2p_1$	$(1-p_1)^3(1-p_2)^2p_2$
4. Versuch	$(1-p_1)^3(1-p_2)^3p_1$	$(1-p_1)^4(1-p_2)^3p_2$

Tabelle 1.5: Erfolgswahrscheinlichkeiten der Schützen.

Sei nun S_{1_n} das Ereignis, dass Schütze 1 in der n ten Runde gewinnt und analog S_{2_n} jenes, das Schütze 2 in der n ten Runde gewinnt. Man erkennt anhand von Tabelle 1.5, dass gilt:

$$\begin{aligned}
P(S_{1_n}) &= p_1(1-p_1)^{n-1}(1-p_1)^{n-1},\\
P(S_{2_n}) &= p_2(1-p_1)^{n}(1-p_2)^{n-1}\\
&= p_2(1-p_1)(1-p_1)^{n-1}(1-p_2)^{n-1}
\end{aligned}$$

Weiter sei S_1 das Ereignis, dass Schütze 1 dass Spiel gewinnt, und S_2 jenes, das Schütze 2 gewinnt. Um nun die zugehörigen Wahrscheinlichkeiten zu berechnen, müssen alle möglichen Spielausgänge aufsummiert werden. Somit gilt

$$P(S_1) = \sum_{k=0}^{\infty} p_1\cdot(1-p_1)^k(1-p_2)^k = p_1\cdot\sum_{k=0}^{\infty}((1-p_1)(1-p_2))^k,$$

$$P(S_2) \quad = \quad p_2(1-p_1) \cdot \sum_{k=0}^{\infty} ((1-p_1)(1-p_2))^k.$$

Mit der Annahme $0 < p_1, p_2 < 1$ gilt stets

$$(1-p_1)(1-p_2) \ < \ 1\,,$$

also sind $P(S_1)$ und $P(S_2)$ nach der geometrischen Reihe konvergent und es folgt

$$P(S_1) \quad = \quad \frac{p_1}{1-((1-p_1)(1-p_2))} \quad = \quad \frac{p_1}{p_1+p_2-p_1p_2},$$
$$P(S_2) \quad = \quad \frac{p_2(1-p_1)}{1-((1-p_1)(1-p_2))} \quad = \quad \frac{p_2-p_1p_2}{p_1+p_2-p_1p_2}.$$

Wenn beide Schützen die gleiche Siegeswahrscheinlichkeit haben sollen, muss $P(S_1) = P(S_2)$ gelten und somit $p_1 = p_2(1-p_1)$, also

$$p_2 \ = \ \frac{p_1}{1-p_1}\,.$$

Für $p_1 \geq \frac{1}{2}$ folgt damit $p_2 \geq 1$, was nicht erlaubt ist.

Für alle $0 < p_1 < \frac{1}{2}$ und

$$p_2 \ = \ \frac{p_1}{1-p_1}$$

erhalten wir aber wie gefordert $P(S_1) = P(S_2) = \frac{1}{2}$.

Ist also zum Beispiel $p_1 = \frac{1}{4}$, so muss $p_2 = \frac{1}{3}$ gelten, damit beide Schützen die gleiche Siegeswahrscheinlichkeit haben.

Aufgabe 1.10.14

Eine Familie vermisst ihren Hund. Es gibt drei Möglichkeiten:

A = Der Hund ist heimgelaufen und wartet vor der Haustür,

B = Der Hund hat einen Knochen auf dem Spielplatz gefunden,

C = Der Hund streunt am Waldrand herum.

Aus den Gewohnheiten des Hundes sind die folgenden Wahrscheinlichkeiten für diese Möglichkeiten bekannt:

$$P(A) \ = \ \frac{1}{4}, \qquad P(B) \ = \ \frac{1}{2} \qquad \text{und} \qquad P(C) \ = \ \frac{1}{4}.$$

Ein Kind sucht auf dem Spielplatz und ein anderes Kind am Waldrand nach dem Hund. Wenn der Hund auf dem Spielplatz ist, so wird er zu 90 % gefunden. Wenn der Hund am Waldrand ist, so wird er zu 50 % gefunden.

(1) Berechne die Wahrscheinlichkeit dafür, dass eines der Kinder den Hund findet.

(2) Berechne die bedingte Wahrscheinlichkeit dafür, dass der Hund vor der Haustür wartet, wenn keines der Kinder den Hund angetroffen hat.

Lösung Teil 1

Sei D das Ereignis, dass eines der Kinder den Hund findet. Die Wahrscheinlichkeit $P(D)$ erhalten wir nach der totalen Wahrscheinlichkeit:

$$\begin{aligned} P(D) &= P(D|A) \cdot P(A) + P(D|B) \cdot P(B) + P(D|C) \cdot P(C) \\ &= 0 + \frac{9}{10} \cdot \frac{1}{2} + \frac{1}{2} \cdot \frac{1}{2} \;=\; \frac{23}{40} \;=\; 57.5\,\%. \end{aligned}$$

Lösung Teil 2

Nun nuten wir das Ergebnis aus Aufgabenteil 1 und wendne die Bayessche Formel an:

$$P(A|D^C) \;=\; P(D^C|A) \cdot \frac{P(A)}{P(D^C)} \;=\; 1 \cdot \frac{\frac{1}{4}}{1 - \frac{23}{40}} \;=\; \frac{10}{17} \;\approx\; 58.82\,\%.$$

Aufgabe 1.10.15

A_1, A_2 und A_3 seien paarweise stochastisch unabhängige Ereignisse mit $P(A_i) = p$ für $i = 1, 2, 3$ und mit $P(A_1 \cap A_2 \cap A_3) = 0$. Finde den Wert für p, für den $P(A_1 \cup A_2 \cup A_3)$ maximal wird.

Lösung

Nach der Siebformel und der stochastischen Unabhängigkeit erhalten wir

$$\begin{aligned} P(A_1 \cup A_2 \cup A_3) &= P(A_1) + P(A_2) + P(A_3) - P(A_1 \cap A_2) \\ &\quad - P(A_2 \cap A_3) - P(A_1 \cap A_3) + P(A_1 \cap A_2 \cap A_3) \\ &= p + p + p - p^2 - p^2 - p^2 \\ &= 3p - 3p^2. \end{aligned}$$

Gesucht ist nun das Maximum der Funktion

$$f : [0,1] \to \mathbb{R} \qquad \text{mit} \qquad f(p) \;=\; 3p - 3p^2.$$

Bei dieser nach unten geöffneten Parabell erhalten wir durch $f'(p) = 3 - 6p$ unser gesuchtes Maximum sofort bei $p = 1/2$.

Aufgabe 1.10.16

In einer Urne befinden sich 3 schwarze, 5 blaue und 2 rote Kugeln. Berechne die Wahrscheinlichkeit dafür, dass beim Ziehen von 2 Kugeln ohne Zurücklegen beide die gleiche Farbe haben.

Lösung

Sei P die Laplace Verteilung und sei A das Ereigniss, dass beim Ziehen von zwei Kugelen beide Kugeln die gleiche Farbe haben. Dann erhalten wir nach der totalen Wahrscheinlichkeit gerade

$$P(A) \;=\; \frac{3}{10}\cdot\frac{2}{9}+\frac{1}{2}\cdot\frac{4}{9}+\frac{1}{5}\cdot\frac{1}{9} \;=\; \frac{6}{90}+\frac{20}{90}+\frac{2}{90} \;=\; \frac{14}{45}.$$

Aufgabe 1.10.17

Es seien A, B und C drei Ereignisse, zu denen wir die folgendne Wahrscheinlichkeiten kennen:

$$P(A) \;=\; \frac{1}{4}, \qquad P(B^C) \;=\; \frac{2}{3}, \qquad P(C) \;=\; \frac{1}{2}, \qquad P(A^C\cap B) \;=\; \frac{1}{4},$$

$$P(B^C\cup C^C) \;=\; \frac{5}{6} \qquad \text{und} \qquad P(A\cap C) \;=\; 0.$$

Berechne $P(A\cup B\cup C)$.

Lösung

Bevor die Siebformel angewendet werden kann, müssen noch einige Wahrscheinlichkeiten berechnet werden. Es gilt $P(B) = 1 - P(B^C) = \frac{1}{3}$. Weiter folgt aus der DeMorganschen Regel

$$P(B^C\cup C^C) \;=\; P((B\cap C)^C) \;=\; 1-P(B\cap C) \;=\; \frac{5}{6},$$

also $P(B\cap C) = \frac{1}{6}$. Natürlich gilt $P(A\cap B\cap C) = 0$, da bereits $P(A\cap C) = 0$ ist. Es ergibt sich noch

$$P(A^C\cap B) \;=\; \frac{1}{4} \;=\; \frac{3}{4}\cdot\frac{1}{3} \;=\; P(A^C)\cdot P(B),$$

also sind A^C und B stochastisch unabhängig und damit auch A und B. Somit erhalten wir

$$P(A\cap B) \;=\; P(A)\cdot P(B) \;=\; \frac{1}{4}\cdot\frac{1}{3} \;=\; \frac{1}{12}.$$

Nun kann die Siebformel angewendet werden:

$$\begin{aligned} P(A\cup B\cup C) \;&=\; P(A)+P(B)+P(C)-P(A\cap B)-P(A\cap C)-\\ &\;\; P(B\cap C)+P(A\cap B\cap C)\\ &=\; \frac{1}{4}+\frac{1}{3}+\frac{1}{2}-\frac{1}{12}-0-\frac{1}{6}+0 \;=\; \frac{5}{6}. \end{aligned}$$

Aufgabe 1.10.18

Die Wahrscheinlichkeiten, dass eine Familie k Kinder hat, sei wie folgt verteilt:

Anzahl Kinder	0	1	2	3	4	5	≥ 6
Wahrscheinlichkeit	0.3	0.2	0.2	0.15	0.1	0.05	0.0

Jungen und Mädchen seien gleichverteilt.

Es wird zufällig eine Familie mit einem Jungen ausgewählt. Berechne die Wahrscheinlichkeit, dass dieser Junge eine Schwester hat.

Lösung

Es sei A die Menge dafür, dass ein Mädchen in einer festen Familie lebt und es sei B die Menge, dass ein Jungen in einer festen Familie lebt. P sei die Laplace Verteilung. Dann gilt nach der totalen Wahrscheinlichkeit

$$\begin{aligned} P(A) \;=\; P(B) &= \frac{1}{2}\cdot 0.2 + \frac{3}{4}\cdot 0.2 + \frac{7}{8}\cdot 0.15 + \frac{15}{16}\cdot 0.1 + \frac{31}{32}\cdot 0.05 \\ &= 0.523438\,. \end{aligned}$$

Nun berechnen wir die Wahrscheinlichkeit dafür, dass ein Junge und ein Mädchen in einer festen Familie leben, also

$$P(A\cap B) \;=\; \frac{2}{4}\cdot 0.2 + \frac{6}{8}\cdot 0.15 + \frac{14}{16}\cdot 0.1 + \frac{30}{32}\cdot 0.05 \;=\; 0.346875\,.$$

Die Lösung des Problems liefert damit die folgende bedingte Wahrscheinlichkeit

$$P(A|B) \;=\; \frac{P(A\cap B)}{P(B)} \;=\; \frac{0.346875}{0.523438} \;\approx\; 0.662687\,.$$

Aufgabe 1.10.19

Eine natürliche Zahl I werde mit der Wahrscheinlichkeit $P(\{I = n\}) = 1/2^n$ ausgewählt. Nimmt I nun den Wert n an, so wird eine Münze einmal geworfen. Mit einer Wahrscheinlichkeit von e^{-n} fällt diese so, dass Zahl erscheint.

Gib einen Wahrscheinlichkeitsraum an und berechen die Wahrscheinlichkeit dafür, dass Zahl erscheint.

Lösung

Es wird also zunächst eine natürliche Zahl n ermittelt und danach eine Münze geworfen. Unser Grundraum Ω ergibt sich somit aus

$$\Omega \;=\; \mathbb{N}\times\{K, Z\},$$

dabei steht K für Kopf und Z für Zahl. Wir definieren dazu wie gefordert die Funktion

$$f(\omega) \;=\; \begin{cases} \frac{1}{2^n}\left(1-e^{-n}\right) & \text{wenn} \quad x = K \\ \frac{1}{2^n} e^{-n} & \text{wenn} \quad x = Z \end{cases} \qquad \text{für alle} \;\; \omega \;=\; (n,x) \;\in\; \Omega.$$

Nun müssen wir zeigen, dass f eine Zähldichte ist, dass also $\sum_{\omega\in\Omega} f(\omega) = 1$ gilt. Nach der geometrischen Reihe folgt aber sofort

$$\begin{aligned} \sum_{\omega\in\Omega} f(\omega) \;&=\; \sum_{n\in\mathbb{N}} \frac{1}{2^n}\left(1-e^{-n}\right) + \frac{1}{2^n}e^{-n} \;=\; \sum_{n=1}^{\infty} \frac{1}{2^n} - \frac{1}{2^n}e^{-n} + \frac{1}{2^n}e^{-n} \\ &=\; \sum_{n=1}^{\infty} \frac{1}{2^n} \;=\; \sum_{n=0}^{\infty} \frac{1}{2^n} - 1 \;=\; \frac{1}{1-\frac{1}{2}} - 1 \;=\; 2-1 \;=\; 1. \end{aligned}$$

Für alle $A \subset \Omega$ mit $P(A) = \sum_{\omega\in A} f(\omega)$ ist demnach (Ω, P) der gesuchte Wahrscheinlichkeitsraum.

Die Wahrscheinlichkeit dafür, dass Zahl erscheint, ergibt sich nun nach

$$\begin{aligned} \sum_{n\in\mathbb{N}} \frac{1}{2^n}e^{-n} \;&=\; \sum_{n=1}^{\infty} \left(\frac{1}{2e}\right)^n \;=\; \sum_{n=0}^{\infty} \left(\frac{1}{2e}\right)^n - 1 \;=\; \frac{1}{1-\frac{1}{2e}} - 1 \\ &=\; \frac{2e}{2e-1} - \frac{2e-1}{2e-1} \;=\; \frac{1}{2e-1} \;\approx\; 22.5\,\%. \end{aligned}$$

Aufgabe 1.10.20

Sei $\Omega = \mathbb{N} \cup \{0\}$ und sei

$$f(n) \;=\; 10^{-(n \text{ div } 9)+1},$$

dabei ist div die ganzzahlige Division, wir haben also

$$f(0) = \ldots = f(8) = 0.1, \qquad f(9) = \ldots = f(17) = 0.01, \;\; \ldots .$$

Zeige:

(1) f ist eine Zähldichte auf Ω.

(2) Zusammen mit der durch die Zähldichte f gegebenen Verteilung P ist (Ω, P) ein Wahrscheinlichkeitsraum.

(3) Die Abbilung $P : \mathcal{P}(\Omega) \to [0,1]$ ist surjektive.

Lösung Teil 1

Ω ist abzählbar, wir haben also nur noch $\sum_{n\in\Omega} f(n) = 1$ zu zeigen. Es gilt aber gerade

$$\begin{aligned}\sum_{n\in\Omega} f(n) &= \sum_{n=0}^{\infty} f(n) = 9\cdot\sum_{n=1}^{\infty} 10^{-n} = 9\cdot\sum_{n=0}^{\infty}\left(\frac{1}{10}\right)^n - 9 \\ &= 9\cdot\frac{1}{1-\frac{1}{10}} - 9 = 10-9 = 1\end{aligned}$$

und somit folgt die Behauptung.

Lösung Teil 2

Wir definieren die Verteilung P durch

$$P(A) = \sum_{n\in A} f(n) \qquad \text{für alle} \quad A\subset\Omega.$$

Wir haben damit nun die σ Additivität zu zeigen.

Seien $\{A_i\}_{i\in\mathbb{N}}$ paarweise disjunkte Teilmengen von Ω. Dann gilt aber sofort

$$P\left(\bigcup_{i=1}^{\infty} A_i\right) = \sum_{i=1}^{\infty} P(A_i),$$

da für alle $k\in\mathbb{N}$ aufgrund der Disjunktheit 10^{-k} maximal 9 mal in der Summe aufsummiert wird, es gibt also keinen Zehnerüberlauf. Dies zeigt die σ Additivität und somit ist (Ω, P) ein Wahrscheinlichkeitsraum.

Lösung Teil 3

Es ist klar, dass $P(\emptyset) = 0$ gilt. Aus der Analysis ist bekannt, dass es zu jeder reellen Zahl $0 < x \leq 1$ eine Folge $(a_k)_{k\in\mathbb{N}}$ von Ziffern gibt, so dass

$$x = \sum_{k=1}^{\infty} a_k 10^{-k}$$

gilt. Dabei ist $0 \leq a_k \leq 9$ für alle $k\in\mathbb{N}$.

Wir müssen nun jedes a_k durch eine Menge von natürlichen Zahlen identifizieren, so dass all diese Mengen disjunkt sind.

Sei dazu

$$A_k = \{9(k-1)+r \mid r = 0,..,a_k-1\}.$$

Für a_k gilt also $A_k = \emptyset$. Die Mengen A_k sind für $k\in\mathbb{N}$ disjunkt und wir erhalten für alle $0 < x \leq 1$

$$x = f\left(\bigcup_{k=1}^{\infty} A_k\right) = \sum_{k=1}^{\infty} a_k 10^{-k},$$

womit die Surjektivität von $P : \mathcal{P}(\Omega) \to [0,1]$ gezeigt wurde.

Aufgabe 1.10.21

Der Mensch hat 6 mögliche Genotypen der Blutgruppe, nämlich

$$AA, \quad BB, \quad 00, \quad AB, \quad A0, \quad B0,$$

die in der Elterngeneration mit den Wahrscheinlichkeiten

$$p_1, \quad p_2, \quad p_3, \quad p_4, \quad p_5, \quad p_6$$

vorkommen. Berechne die Verteilung der Genotypen unter der ersten Kindergeneration bei zufälligen Paarungen.

Lösung

Um die erste Kindergeneration P_1 zu untersuchen, berechnen wir die einzelnen Wahrscheinlichkeiten

$$P_1(AA), \quad P_1(BB), \quad P_1(00), \quad P_1(AB), \quad P_1(A0), \quad P_1(B0),$$

analog zum Beispiel 1.4.1 auf Seite 31. Die Wahrscheinlichkeit $P_1(AA)$ erhalten wir aus Tabelle 1.6.

P_0-Konstellationen	W(konst)	$W(AA\|\text{konst})$
$AA\ AA$	p_1^2	1
$AA\ AB$	p_1p_2	1/2
$AB\ AA$	p_1p_2	1/2
$AA\ A0$	p_1p_5	1/2
$A0\ AA$	p_1p_5	1/2
$AB\ AB$	p_4^2	1/4
$AB\ A0$	p_4p_5	1/4
$A0\ AB$	p_4p_5	1/4
$A0\ A0$	p_5^2	1/4

Tabelle 1.6: Wahrscheinlichkeiten der ersten Kindergeneration.

Aus keiner weitern P_0-Konstellationen kann ein Kind mit der Blutgruppe AA erzeugt werden. Wir erhalten also

$$P_1(AA) \;=\; p_1^2 + p_1p_4 + p_1p_5 + \frac{1}{4}p_4^2 + \frac{1}{5}p_4^2 + \frac{1}{2}p_4p_5\,.$$

Analog können wir auch die anderen gesuchten Wahrscheinlichkeiten berechnen und erhalten

$$P_1(BB) \;=\; p_2^2 + p_2p_4 + p_2p_6 + \frac{1}{4}p_4^2 + \frac{1}{4}p_6^2 + \frac{1}{2}p_4p_6\,,$$

$$
\begin{aligned}
P_1(00) &= p_3^2 + p_3p_5 + p_3p_6 + \frac{1}{4}p_5^2 + \frac{1}{4}p_6^2 + \frac{1}{2}p_5p_6\,, \\
P_1(AB) &= 2p_1p_2 + p_1p_4 + p_1p_6 + p_2p_4 + p_2p_5 + \frac{1}{2}p_4p_5 + \frac{1}{2}p_4p_6 \\
&\quad + \frac{1}{2}p_5p_6 + \frac{1}{2}p_4^2\,, \\
P_1(A0) &= 2p_1p_3 + p_1p_5 + p_1p_6 + p_3p_4 + p_3p_5 + \frac{1}{2}p_4p_5 + \frac{1}{2}p_4p_6 \\
&\quad + \frac{1}{2}p_5p_6 + \frac{1}{2}p_5^2\,, \\
P_1(B0) &= 2p_2p_3 + p_2p_5 + p_2p_6 + p_3p_4 + p_3p_6 + \frac{1}{2}p_4p_5 + \frac{1}{2}p_4p_6 \\
&\quad + \frac{1}{2}p_5p_6 + \frac{1}{2}p_6^2\,.
\end{aligned}
$$

Aufgabe 1.10.22

In einer Glühbirnenfabrik stellen 3 Maschinen X, Y und Z von der Gesamtproduktion 20 %, 30 % und 50 % her. Dabei sind im Mittel 2 % von der Maschine X, 4 % von Y und 7 % von Z hergestellten Birnen defekt. Aus der Gesamtproduktion wird eine Glühbirne entnommen, die sich als fehlerhaft herausstellt. Berechne die Wahrscheinlichkeiten dafür, dass die defekte Birne von Maschine X, Y bzw. Z produziert wurde.

Lösung

Sei A das Ereignis, dass eine aus der Gesamtproduktion zufällig gezogene Glühbirne von Maschine X produziert wurde. Seien B und C analog definiert. Sei weiter G das Ereignis, dass eine aus der Gesamtproduktion zufällig gezogene Glühbirne defekt ist. Gesucht sind also $P(A|G)$, $P(B|G)$ und $P(C|G)$.

Es gilt nach Aufgabenstellung

$$P(A) = 0.2, \qquad P(B) = 0.3 \qquad \text{und} \qquad P(C) = 0.5\,.$$

Aus der Aufgabenstellung folgt ebenfalls, dass die Wahrscheinlichkeiten dafür, dass eine von Maschine X, Y oder Z produzierte Glühbirne defekt ist,

$$P(G|A) = 0.02, \qquad P(G|B) = 0.04 \qquad \text{und} \qquad P(G|C) = 0.07$$

sind. Die Wahrscheinlichkeit dafür, dass eine aus der Gesamtproduktion zufällig entnommene Glühbirne defekt ist, ergibt sich nach der totalen Wahrscheinlichkeit zu

$$P(G) = 0.2 \cdot 0.02 + 0.3 \cdot 0.04 + 0.5 \cdot 0.07 = 0.051\,.$$

Schließlich ergibt sich nach der Bayesschen Formel

$$\begin{aligned}
P(A|G) &= P(G|A)\frac{P(A)}{P(G)} = \frac{0.02 \cdot 0.2}{0.051} \approx 7.84\,\%,\\
P(B|G) &= P(G|B)\frac{P(B)}{P(G)} = \frac{0.04 \cdot 0.3}{0.051} \approx 23.53\,\%,\\
P(C|G) &= P(G|C)\frac{P(C)}{P(G)} = \frac{0.07 \cdot 0.5}{0.051} \approx 68.63\,\%.
\end{aligned}$$

Aufgabe 1.10.23

Es seien X_N für $N \in \mathbb{N}$ Zufallsvariablen mit einer hypergeometrischen Verteilung mit Parametern den Parametern $(N, [pN], n)$, dabei ist $[x]$ die größte natürliche Zahl $\leq x$. Zeige, dass

$$\lim_{N\to\infty} P(X_N = k) = \binom{n}{k} p^k (1-p)^{n-k}$$

gilt, dass die hypergeometrische Verteilung also gegen die Binomialverteilung konvergiert.

Lösung

Es gilt

$$\begin{aligned}
P(X_N = k) &= \frac{\binom{[pN]}{k}\binom{N-[pN]}{n-k}}{\binom{N}{n}}\\
&= \frac{[pN]!}{k!([pN]-k)!}\,\frac{(N-[pN])!}{(n-k)!(N-[pN]-n+k)!}\,\frac{n!(N-n)!}{N!}\\
&= \frac{n!}{k!(n-k)!}\,\frac{[pN]!(N-[pN])!(N-n)!}{N!([pN]-k)!(N-[pN]-n+k)!}\\
&= \binom{n}{k}\frac{[pN]\ldots([pN]-k+1)\cdot(N-[pN])\ldots(N-[pN]-n+k+1)}{N(N-1)\ldots(N-n+1)}.
\end{aligned}$$

Betrachten wir den Grenzwert $\lim_{N\to\infty}$, so folgt wegen $\lim_{N\to\infty}[pN]/N = p$

$$\begin{aligned}
&\lim_{N\to\infty} P(X_N = k)\\
&= \binom{n}{k}\frac{N^k(p\ldots(p-\frac{k+1}{N}))N^{n-k}((1-p)\ldots(1-p-\frac{n+k+1}{N}))}{N^n(1(1-\frac{1}{N})\ldots(1-\frac{n+1}{N}))}\\
&= \binom{n}{k}\frac{p\ldots(p-\frac{k+1}{N})(1-p)\ldots(1-p-\frac{n+k+1}{N})}{1(1-\frac{1}{N})\ldots(1-\frac{n+1}{N})}.
\end{aligned}$$

Da für alle $x \in \mathbb{R}$ gerade $\lim_{N\to\infty}\frac{x}{N} = 0$ gilt, folgt wie behauptet

$$\lim_{N\to\infty} P(X_N = k) = \binom{n}{k}\frac{p^k(1-p)^{n-k}}{1} = \binom{n}{k}p^k(1-p)^{n-k}.$$

Aufgabe 1.10.24

Die k te von $n \geq 3$ Urnen enthalte k schwarze und $n-k$ weiße Kugeln, wobei n ungerade sei. Nun wird eine Urne zufällig ausgewählt und daraus eine Kugel gezogen. Berechne unter der Annahme, dass die gezogene Kugel scharz ist, die Wahrscheinlichkeit p_n dafür, dass die gewählte Urne danach noch mindestens soviele schwarze wie weiße Kugeln enthält.

Lösung

Das Beispiel $n = 7$ verdeutlicht Abbildung 1.12.

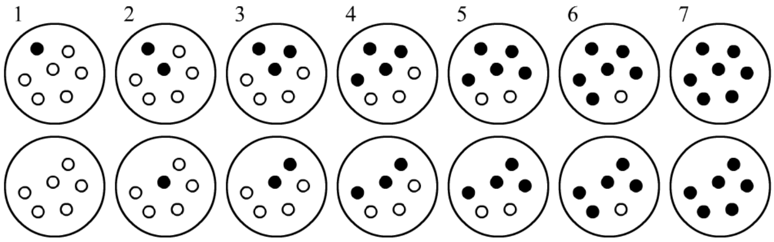


Abbildung 1.12: Urnen vor und nach dem Ziehen einer schwarzen Kugel.

Mit B bezeichnen wir die Möglichkeiten, dass eine schwarze Kugel gezogen wird, und mit A bezeichnen wir die Möglichkeiten, dass nach der Ziehung noch mehr schwarze als weiße Kugeln in der Urne enthalten sind. Sei P die Laplace Verteilung. Dann haben wir nach der Aufgabenstellung

$$P(B) \;=\; \sum_{k=1}^{n} \frac{1}{n} \cdot \frac{k}{n} \qquad \text{und} \qquad P(A) \;=\; \sum_{k=\frac{n+1}{2}}^{n} \frac{k}{n}.$$

Damit kann nun die gesuchte bedingte Wahrscheinlichkeit $p_n = P(A|B)$ berechnet werden:

$$\begin{aligned} p_n \;&=\; P(A|B) \;=\; \frac{P(A\cap B)}{P(B)} \;=\; \frac{\sum_{k=\frac{n+1}{2}}^{n} \frac{1}{n}\cdot\frac{k}{n}}{\sum_{k=1}^{n} \frac{1}{n}\cdot\frac{k}{n}} \;=\; \frac{\sum_{k=\frac{n+1}{2}}^{n} k}{\sum_{k=1}^{n} k} \\ &=\; \frac{\frac{n(n+1)}{2} - \frac{\frac{n-1}{2}(\frac{n-1}{2}+1)}{2}}{\frac{n(n+1)}{2}} \;=\; \frac{n^2+n-\frac{1}{4}n^2+\frac{1}{2}n-\frac{1}{4}-\frac{1}{2}n+\frac{1}{2}}{n^2+n} \\ &=\; \frac{\frac{3}{4}n^2+n+\frac{1}{4}}{n^2+n} \;=\; \frac{3}{4}+\frac{1}{4n}. \end{aligned}$$

Betrachten wir den Grenzwert für $n \to \infty$, so erhalten wir

$$\lim_{n\to\infty} p_n \;=\; \frac{3}{4}.$$

Aufgabe 1.10.25

Es wird solange mit einem fairen Würfel gewürfelt, bis im n ten Wurf eine 6 kommt. Berechen die Wahrscheinlichkeit dafür, dass $n = 2$ gilt unter der Annahme, dass n gerade ist.

Lösung

Sei Ω die Menge aller Würfelfolgen, die mit einer 6 enden und bei denen zuvor keine 6 vorkam. Sei weiter A die Menge dieser Folgen, die aus nur zwei Würfen besteht, bei denen der zweite Wurf gleich eine 6 war. Sei weiter B die Menge der Folgen, die aus einer geraden Anzahl von Würfen besteht.

Ist P die Laplace Verteilung auf Ω, dann erhalten wir

$$P(A) \;=\; \frac{5}{6}\cdot\frac{1}{6} \;=\; \frac{5}{36}.$$

Um $P(B)$ zu berechnen, müssen wir die folgende Reihe betrachten, da die erste 6 theoretisch erst im Unendlichen gewürfelt werden könnte:

$$\begin{aligned}
P(B) &= \frac{1}{6}\cdot\left(\frac{5}{6}\right)^1 + \frac{1}{6}\cdot\left(\frac{5}{6}\right)^3 + \frac{1}{6}\cdot\left(\frac{5}{6}\right)^5 + \frac{1}{6}\cdot\left(\frac{5}{6}\right)^7 + \ldots \\
&= \frac{1}{6}\cdot\sum_{k=1}^{\infty}\left(\frac{5}{6}\right)^{2k-1} \;=\; \frac{1}{6}\cdot\sum_{k=0}^{\infty}\left(\frac{5}{6}\right)^{2k+1} \;=\; \frac{1}{6}\cdot\frac{5}{6}\cdot\sum_{k=0}^{\infty}\left(\frac{5}{6}\right)^{2k} \\
&= \frac{1}{6}\cdot\frac{5}{6}\cdot\frac{1}{1-\left(\frac{5}{6}\right)^2} \;=\; \frac{5}{36}\cdot\frac{36}{11} \;=\; \frac{5}{11},
\end{aligned}$$

dabei wurde die geometrische Reihe verwendet.

Gesucht ist die bedingte Wahrscheinlichkeit $P(A|B)$. Es ist sofort einsichtig, dass $P(B|A) = 1$ gilt, denn eine Folge aus 2 Würfen ist auch eine Folge mit einer geraden Anzahl von Würfen. Nach Bayesschen Formel erhalten wir also

$$P(A|B) \;=\; P(B|A)\cdot\frac{P(A)}{P(B)} \;=\; \frac{\frac{5}{36}}{\frac{5}{11}} \;=\; \frac{11}{36} \;\approx\; 30.6\,\%.$$

Aufgabe 1.10.26

Die Wahrscheinlichkeit in einer Auster eine Perle zu finden ist $\vartheta = 0.001$. Gib eine Wahrscheinlichkeitsmaß zur Berechnung dafür an, unter $n = 5000$ Austern mindestens zwei Perlen zu finden, und berechne diese Wahrscheinlichkeit.

Lösung

Eine Perle in einer Auster zu finden ist ein seltendes Ereignis. Wir approximieren das Problem daher durch die Poisson Verteilung mit dem Parameter

$$\lambda \;=\; n \cdot \vartheta \;=\; 5000 \cdot 0.001 \;=\; 5.$$

Damit erhalten wir die Zähldichte

$$f(n) \;=\; \frac{\lambda^n}{n!}\mathrm{e}^{-\lambda} \;=\; \frac{5^n}{n!}\mathrm{e}^{-5} \qquad \text{für} \;\; n \in \mathbb{N} \cup \{0\}.$$

Die Wahrscheinlichkeit P_2, dass mindestens zwei Perlen gefunden werden, ist nun

$$P_2 \;=\; 1 - f(0) - f(1) \;=\; 1 - \mathrm{e}^{-5} - 5\mathrm{e}^{-5} \;=\; 1 - \frac{6}{\mathrm{e}^5} \;\approx\; 95.96\,\%.$$

Aufgabe 1.10.27

Drei unterscheidbare faire Würfel werden geworfen und es bezeichne w_i für $i = 1, 2, 3$ das Ergebnis des jeweiligen Würfels. Bestimme die Verteilung der Zufallsvariablen $X_1 = \min\{w_1, w_2, w_3\}$ und $X_2 = \max\{w_1, w_2, w_3\}$.

Lösung

Wir wollen also die Zufallsvariablen

$$X_k : \{1, \ldots, 6\}^3 \to \{1, \ldots, 6\} \qquad \text{für} \;\; k = 1, 2$$

mit

$$\begin{aligned} X_1(w_1, w_2, w_3) &= \min\{w_1, w_2, w_3\} \qquad \text{und} \\ X_2(w_1, w_2, w_3) &= \max\{w_1, w_2, w_3\} \end{aligned}$$

untersuchen. Wir erhalten dafür sofort

$$\begin{aligned}
P^{X_1}(\{1\}) &= P(X_1 = 1) = \frac{6^3 - 5^3}{6^3} = \frac{91}{216}, \\
P^{X_1}(\{2\}) &= P(X_1 = 2) = \frac{5^3 - 4^3}{6^3} = \frac{61}{216}, \\
P^{X_1}(\{3\}) &= P(X_1 = 3) = \frac{4^3 - 3^3}{6^3} = \frac{37}{216}, \\
P^{X_1}(\{4\}) &= P(X_1 = 4) = \frac{3^3 - 2^3}{6^3} = \frac{19}{216}, \\
P^{X_1}(\{5\}) &= P(X_1 = 5) = \frac{2^3 - 1^3}{6^3} = \frac{7}{216}, \\
P^{X_1}(\{6\}) &= P(X_1 = 6) = \frac{1^3}{6^3} = \frac{1}{216}
\end{aligned}$$

und $P^{X_1}(\{k\}) = 0$ für $k \notin \{1,..,6\}$. Aus Symmetriegründen erhalten wir analog

$$\begin{aligned}
P^{X_2}(\{1\}) &= P(X_2 = 1) &= \frac{1^3}{6^3} &= \frac{1}{216}, \\
P^{X_2}(\{2\}) &= P(X_2 = 2) &= \frac{2^3 - 1^3}{6^3} &= \frac{7}{216}, \\
P^{X_2}(\{3\}) &= P(X_2 = 3) &= \frac{3^3 - 2^3}{6^3} &= \frac{19}{216}, \\
P^{X_2}(\{4\}) &= P(X_2 = 4) &= \frac{4^3 - 3^3}{6^3} &= \frac{37}{216}, \\
P^{X_2}(\{5\}) &= P(X_2 = 5) &= \frac{5^3 - 4^3}{6^3} &= \frac{61}{216}, \\
P^{X_2}(\{6\}) &= P(X_2 = 6) &= \frac{6^3 - 5^3}{6^3} &= \frac{91}{216}.
\end{aligned}$$

Aufgabe 1.10.28

In Göttingen gibt es N Haushalte. Wählt man aus diesen N Haushalten zufällig einen aus, so ist die Wahrscheinlichkeit gleich p_k, dass in diesem Haushalt k Kinder wohnen mit $0 \le k \le m$ und m sei dabei die maximale Anzahl an Kindern in einem Haushalt. Nun wird zufällig ein Kind aus der Gesamtheit aller Kinder in Göttingen ausgewählt und es sei X die Anzahl der Kinder, die in dem betreffenden Haushalt lebt. Bestimme die Wahrscheinlichkeitsverteilung von X.

Lösung

Es gilt

$$X : \Omega \to \{1, \ldots, m\}.$$

Gesucht ist $P(X = k)$. Das Auswählen eines Kindes und somit eines Haushaltes ist Laplace verteilt. Sei A_k die Anzahl von Kindern in Göttingen, die in einem Haushalt mit k Kindern wohnen. Sei H_k die Anzahl der Haushalte in Göttingen, in denen k Kinder wohnen. Sei weiterhin M die Anzahl von Kindern, die in Göttingen wohnen. Dann gilt

$$P(X = k) = \frac{A_k}{M} = \frac{H_k \cdot k}{M}.$$

Es gilt

$$H_k = p_k \cdot N \qquad \text{und} \qquad M = \sum_{k=1}^{m} N p_k k\,.$$

Somit folgt

$$P(X = k) = \frac{p_k \cdot N \cdot k}{\sum_{k=1}^{m} N p_k k} = \frac{p_k k}{\sum_{k=1}^{m} p_k k}\,.$$

Summieren wir die Einzelwahrscheinlichkeiten auf, so erhalten wir wie immer

$$\begin{aligned}\sum_{k=0}^{m} P(X=k) &= \frac{p_0 \cdot 0}{\sum_{k=1}^{m} p_k k} + \frac{p_1 \cdot 1}{\sum_{k=1}^{m} p_k k} + \ldots + \frac{p_m \cdot m}{\sum_{k=1}^{m} p_k k} \\ &= 0 + \frac{\sum_{k=1}^{m} p_k k}{\sum_{k=1}^{m} p_k k} = 1\,.\end{aligned}$$

Aufgabe 1.10.29

Sei $\Omega = \{\omega_1, \omega_2, \omega_3\}$ und sei

$$P(\omega_1) = P(\omega_2) = P(\omega_3) = \frac{1}{3}.$$

Weiter seien die Zufallsvariablen X, Y und Z gegeben durch

$$\begin{array}{lll} X(\omega_1) = 1, & X(\omega_2) = 2, & X(\omega_3) = 3, \\ Y(\omega_1) = 2, & Y(\omega_2) = 3, & Y(\omega_3) = 1, \\ Z(\omega_1) = 3, & Z(\omega_2) = 1, & Z(\omega_3) = 2. \end{array}$$

Bestimme die Verteilungen von $X+Y$, von $Y+Z$ und von $\sqrt{(X^2+Y^2)Z}$.

Lösung

Zunächst zeigen wir, dass X, Y und Z die gleichen Wahrscheinlichkeitsverteilungen besitzen, denn es gilt für

$$X, Y, Z : \{1,2,3\} \to \{1,2,3\}$$

gerade

$$\begin{aligned} P^X(\{1\}) &= P(X=1) = P(\omega_1) = 1/3, \\ P^X(\{2\}) &= P(X=2) = P(\omega_2) = 1/3, \\ P^X(\{3\}) &= P(X=3) = P(\omega_3) = 1/3, \\[1ex] P^Y(\{1\}) &= P(Y=1) = P(\omega_3) = 1/3, \\ P^Y(\{2\}) &= P(Y=2) = P(\omega_1) = 1/3, \\ P^Y(\{3\}) &= P(Y=3) = P(\omega_2) = 1/3, \\[1ex] P^Z(\{1\}) &= P(Y=1) = P(\omega_2) = 1/3, \\ P^Z(\{2\}) &= P(Y=2) = P(\omega_3) = 1/3, \\ P^Z(\{3\}) &= P(Y=3) = P(\omega_1) = 1/3. \end{aligned}$$

Weiter erhalten wir $X+Y : \{1,2,3\} \to \{3,4,5\}$

$$\begin{aligned} P^{X+Y}(\{3\}) &= P(X+Y=3) = P(\omega_1) = 1/3, \\ P^{X+Y}(\{4\}) &= P(X+Y=4) = P(\omega_3) = 1/3, \\ P^{X+Y}(\{5\}) &= P(X+Y=5) = P(\omega_2) = 1/3 \end{aligned}$$

und ebenso $Y+Z : \{1,2,3\} \to \{3,4,5\}$

$$\begin{aligned} P^{Y+Z}(\{3\}) &= P(Y+Z=3) = P(\omega_3) = 1/3, \\ P^{Y+Z}(\{4\}) &= P(Y+Z=4) = P(\omega_2) = 1/3, \\ P^{Y+Z}(\{5\}) &= P(Y+Z=5) = P(\omega_1) = 1/3. \end{aligned}$$

Für $\sqrt{(X^2+Y^2)Z} : \{1,2,3\} \to \{\sqrt{13}, \sqrt{15}, \sqrt{20}\}$ ergibt sich

$$\begin{aligned} P^{\sqrt{(X^2+Y^2)Z}}(\{\sqrt{13}\}) &= P\left(\sqrt{(X^2+Y^2)Z} = \sqrt{13}\right) = P(\omega_2) = 1/3, \\ P^{\sqrt{(X^2+Y^2)Z}}(\{\sqrt{15}\}) &= P\left(\sqrt{(X^2+Y^2)Z} = \sqrt{15}\right) = P(\omega_1) = 1/3, \\ P^{\sqrt{(X^2+Y^2)Z}}(\{\sqrt{20}\}) &= P\left(\sqrt{(X^2+Y^2)Z} = \sqrt{20}\right) = P(\omega_3) = 1/3. \end{aligned}$$

Aufgabe 1.10.30

Seien $X_1, .., X_n$ unabhängige gleichverteilte Zufallsvariablen mit Werten in $\{1, .., N\}$. Berechne

$$E\left(\max\{X_1, .., X_n\}\right) + E\left(\min\{X_1, .., X_n\}\right).$$

Lösung

Es gilt

$$X_i : \Omega \to \Omega'_i \subset \{1, .., N\}$$

mit

$$P(X_i = k) = \frac{1}{|\Omega'_i|}$$

für $i = 1, .., n$ und $k \in \Omega'_i$. Wegen der Unabhängigkeit, der Gleichverteilung und der Symmetrie erhalten wir

$$\begin{aligned} & E\left(\max\{X_1, .., X_n\}\right) + E\left(\min\{X_1, .., X_n\}\right) \\ = & \sum_{k=1}^{N} k \cdot P(\max = k) + \sum_{k=1}^{N} k \cdot P(\min = k) \\ = & \sum_{k=1}^{N} k \cdot P(\max = k) + \sum_{k=1}^{N} k \cdot P(\max = N-k+1) \\ = & \sum_{k=1}^{N} k \cdot P(\max = k) + \sum_{k=1}^{N} P(\max = k(N-k+1)) \\ = & \sum_{k=1}^{N} k \cdot P(\max = k) + \sum_{k=1}^{N} (N-k+1) \cdot P(\max = k) \end{aligned}$$

$$
\begin{aligned}
&= \sum_{k=1}^{N} (k + N - k + 1) \cdot P(\max = k) \\
&= (N+1) \cdot \sum_{k=1}^{N} P(\max = k) \;=\; (N+1) \cdot 1 \;=\; N + 1\,.
\end{aligned}
$$

Aufgabe 1.10.31

Aus einer Urne mit N Kugeln wird mit Zurücklegen gezogen. Berechen die erwartete Anzahl der Ziehungen, die benötigt wird, bis jede Kugel mindestens einmal gezogen wurde.

Lösung

Wir betrachten die Zufallsvariablen

$$X \;=\; X_1 + \ldots + X_N,$$

wobei X_k für $k = 2, .., N$ die Anzahl der Ziehungen zwischen dem ersten Erscheinen der $(k-1)$ ten und der k ten Kugel beschreibt, also die Wartezeit auf den ersten Erfolg. Dies heißt, dass P^{X_k} geometrisch verteilt ist mit dem Parameter

$$\vartheta_k \;=\; \frac{N - k + 1}{N},$$

wir haben also

$$\vartheta_1 \;=\; 1, \qquad \vartheta_2 = \frac{N-1}{N}, \qquad \text{bis} \qquad \vartheta_N \;=\; \frac{1}{N}.$$

Weiter ist bekannt, dass der Erwartungswerte EX einer geometrischen Verteilung X mit Parameter ϑ gerade $EX = 1/\vartheta$ ist. Dadurch erhalten wir die gesuchte Anzahl der Ziehungen $Z(N)$:

$$
\begin{aligned}
Z(N) &= E\left(\sum_{i=1}^{N} X_i\right) \;=\; \sum_{i=1}^{N} EX_i \;=\; \sum_{i=1}^{N} \frac{N}{N-k+1} \\
&= N \cdot \sum_{i=1}^{N} \frac{1}{N-k+1} \;=\; N \cdot \left(\frac{1}{N} + \frac{1}{N-1} + \ldots + \frac{1}{2} + 1\right) \\
&\approx N \cdot (\log N + \gamma)\,,
\end{aligned}
$$

dabei ist $\gamma \approx 0.57721$ die Eulersche Konstante.

Für $N = 13$ erwarten wir also nach $Z(13) = 41.34$ Ziehung, dass jede Kugel einmal gezogen wurde.

Aufgabe 1.10.32

Seien X und Y reelle Zufallsvariablen mit $EX = EY = 0$, mit $\mathrm{Var}(X) = \mathrm{Var}(Y) = 1$ und mit $\mathrm{Cov}(X,Y) = p$. Zeige, dass dann

$$E(\max(X^2, Y^2)) \;\leq\; 1 + \sqrt{1 - p^2}$$

gilt. Dabei kann die Identität $\max(a,b) = (a+b)/2 + |a-b|/2$ für alle $a, b \in \mathbb{R}$ verwendet werden.

Lösung

Es gilt nach Definition und nach Voraussetzung

$$\mathrm{Cov}(X,Y) \;=\; E(XY) - (EX)(EY) \;=\; E(XY) \;=\; p$$

und

$$\mathrm{Var}(X) \;=\; E(X^2) - (EX)^2 \;=\; E(X^2) \;=\; 1$$

sowie analog $E(Y^2) = 1$. Damit erhalten wir

$$\begin{aligned}
& E(\max(X^2, Y^2)) \;=\; \frac{1}{2} \cdot (E(X^2) + E(Y^2) + E(|X^2 - Y^2|)) \\
=\; & \frac{1}{2} \cdot (2 + E(|X^2 - Y^2|)) \;=\; 1 + \frac{1}{2} \cdot E(|(X+Y)(X-Y)|)) \\
\leq\; & 1 + \frac{1}{2} \cdot \sqrt{E((X+Y)^2)} \cdot \sqrt{E((X-Y)^2)} \\
=\; & 1 + \frac{1}{2} \cdot \sqrt{E(X^2) + 2E(XY) + E(Y^2)} \cdot \sqrt{E(X^2) - 2E(XY) + E(Y^2)} \\
=\; & 1 + \frac{1}{2} \cdot \sqrt{(2 + 2p)} \cdot \sqrt{(2 - 2p)} \;=\; 1 + \frac{1}{2}\sqrt{4 - 4p^2} \\
=\; & 1 + \sqrt{1 - p^2}.
\end{aligned}$$

Dabei wurde die Cauchy-Schwarz Ungleichung

$$|E(XY)| \;\leq\; \sqrt{E(X^2)} \cdot \sqrt{E(Y^2)}$$

verwendet.

Aufgabe 1.10.33

Ein betrunkener Junge kommt im Dunkeln nach Hause. Die Haustür ist verschlossen und er hat genau n Schlüssel in seiner Tasche, von denen genau einer passt. Er wählt zufällig einen aus und falls dieser nicht passt,

(1) ... legt er ihn beiseite.

(2) ... steckt er ihn zurück in die Tasche.

Er probiert jeweils solange, bis er das Haus betreten kann. Die Zufallsvariable X beschreibe die Anzahl der zum Öffnen der Tür benötigten Versuche. Gib jeweils ein geeignetes Modell an und berechne EX sowie $\mathrm{Var}(X)$.

Lösung Teil 1

Ein geeignetes Modell für das Problem ist die Annahme, dass X Laplace verteilt ist, also

$$P^X(\{k\}) \;=\; P(X=k) \;=\; \frac{1}{N} \qquad \text{für} \;\; k=1,..,N.$$

Damit berechnen wir den Erwartungswert

$$EX \;=\; \sum_{k=1}^{N} k \cdot \frac{1}{N} \;=\; \frac{1}{N}\sum_{k=1}^{N} k \;=\; \frac{1}{N}\cdot\frac{N(N+1)}{2} \;=\; \frac{N+1}{2}$$

und erhalten damit auch die Varianz

$$\begin{aligned}
\operatorname{Var}(X) &= E(X-E(X))^2 \;=\; EX^2-(EX)^2 \\
&= \sum_{k=1}^{N} k^2\cdot\frac{1}{N} - \frac{(N+1)^2}{4} \;=\; \frac{1}{N}\cdot\sum_{k=1}^{N} k^2 - \frac{(N+1)^2}{4} \\
&= \frac{1}{N}\cdot\frac{N(N+1)(2N+1)}{6} - \frac{(N+1)^2}{4} \\
&= \frac{(N+1)(2N+1)}{6} - \frac{N^2+2N+1}{4} \;=\; \frac{n^2-1}{12}.
\end{aligned}$$

Lösung Teil 2

Ein geeignetes Modell für dieses Problem ist die Annahme, dass X mit dem Paramter $p=1/N$ geometrisch verteilt ist, also

$$P^X(\{k\}) \;=\; P(X=k) \;=\; (1-p)^{k-1}p \;=\; \frac{(N-1)^{k-1}}{N^k} \qquad \text{für} \;\; k\in\mathbb{N}.$$

Es sei $q=1-p=(N-1)/N$. Damit erhalten wir den Erwartungswert

$$\begin{aligned}
E(X) &= \sum_{n=1}^{\infty} n(1-p)^{n-1}p \\
&= p\left(\sum_{n=1}^{\infty} nq^{n-1}\right) \;=\; p\left(\sum_{n=0}^{\infty}(n+1)q^n\right) \\
&= p\left(\frac{\mathrm{d}}{\mathrm{d}q}\sum_{n=0}^{\infty} q^n\right) \;=\; p\left(\frac{\mathrm{d}}{\mathrm{d}q}\left(\frac{1}{1-q}\right)\right) \\
&= p\frac{1}{(1-q)^2} \;=\; \frac{p}{(1-(1-p))^2} \;=\; \frac{1}{p} \\
&= \frac{1}{\frac{1}{N}} \;=\; N,
\end{aligned}$$

dabei haben wir die geometrische Reihe verwendet. Um nun die Varianz zu erhalten, berechnen wir zunächst $E(X^2)$. Es gilt

$$E(X^2) = E(X^2) - E(X) + E(X) = E(X(X-1)) + EX.$$

Weiterhin gilt

$$\begin{aligned} E(X(X-1)) &= \sum_{n=1}^{\infty} n(n-1)q^{n-1}p = p\sum_{n=1}^{\infty}(n+1)nq^n \\ &= pq\frac{\mathrm{d}^2}{\mathrm{d}q^2}\sum_{n=1}^{\infty} q^n = pq\frac{\mathrm{d}^2}{\mathrm{d}q^2}\frac{1}{(1-q)} \\ &= \frac{2pq}{(1-q)^3} = \frac{2pq}{p^3} = \frac{2q}{p^2}. \end{aligned}$$

Somit folgt

$$\begin{aligned} \mathrm{Var}(X) &= E(X^2) - (EX)^2 = E(X(X-1)) + EX - (EX)^2 \\ &= \frac{2q}{p} + \frac{1}{p} - \frac{1}{p^2} = \frac{2-2p+p-1}{p^2} \\ &= \frac{1-p}{p^2} = \frac{q}{p^2}. \end{aligned}$$

Mit $p = 1/N$ und $q = 1 - 1/N = (N-1)/N$ folgt schließlich

$$\mathrm{Var}(X) = \frac{q}{p^2} = \frac{N-1}{N}\frac{N^2}{1} = N^2 - N.$$

Aufgabe 1.10.34

Seien X und Y stochastisch unabhängige Zufallsvariablen.

(1) X und Y seien geometrisch verteilt mit Parameter p. Bestimme die Verteilung von $X + Y$.

(2) X und Y seien negativ binomial verteilt mit Parametern r und p bzw. mit s und p, es gilt also für $k \geq 0$

$$\begin{aligned} f_X(k) &= \binom{r+k-1}{k} p^r (1-p)^k, \\ f_Y(k) &= \binom{s+k-1}{k} p^s (1-p)^k. \end{aligned}$$

Bestimme die Verteilung von $X + Y$.

Dabei kann die Faltungsformel genutzt werden.

Lösung Teil 1

Wir haben die Zähldichten

$$f_X(k) \;=\; f_Y(k) \;=\; (1-p)^{k-1}p \qquad \text{für} \;\; k \in \mathbb{N}$$

und erhalten damit nach der Faltungsformel die Verteilung $X+Y$ durch die Zähldichte

$$\begin{aligned} f_{X+Y}(z) \;&=\; \sum_{k=1}^{z} f_X(k) f_Y(z-k) \;=\; \sum_{k=1}^{z} (1-p)^{k-1} p (1-p)^{z-k-1} p \\ &=\; p^2 \cdot \sum_{k=1}^{z} (1-p)^{k-1+z-k-1} \;=\; p^2 \cdot \sum_{k=1}^{z} (1-p)^{z-2} \\ &=\; (1-p)^{z-2} p^2 \cdot \sum_{k=1}^{z} 1 \;=\; (1-p)^{z-2} p^2 z. \end{aligned}$$

Lösung Teil 2

Durch die Zähldichten aus der Aufgabenstellung und durch die Faltungsformel erhalten wir die gesuchte Verteilung durch die Zähldichte

$$\begin{aligned} f_{X+Y}(z) \;&=\; \sum_{k=0}^{z} f_X(k) f_Y(z-k) \\ &=\; \sum_{k=0}^{z} \binom{r+k-1}{k} p^r (1-p)^k \binom{s+z-k-1}{z-k} p^s (1-p)^{z-k} \\ &=\; p^{r+s}(1-p)^z \sum_{k=0}^{z} \binom{r+k-1}{k} \binom{s+z-k-1}{z-k} \\ &=\; p^{r+s}(1-p)^z \binom{r+k-1+s+z-k-1}{z} \\ &=\; \binom{(r+s)+z-2}{z} p^{r+s}(1-p)^z. \end{aligned}$$

Dabei haben wir die Identität

$$\sum_{k=0}^{z} \binom{m}{k} \binom{n}{z-k} \;=\; \binom{m+n}{z}$$

für nicht negative ganze Zahlen m, n und z verwendet.

Aufgabe 1.10.35

Sei X eine diskrete Zufallsvariable mit symmetrischer Verteilung, es gilt also

$$P(X=x) \;=\; P(X=-x)$$

für alle $x \in \mathbb{R}$. Zeige, dass X und X^2 unkorreliert sind und nur dann stochastisch unabhängig sind, wenn X in maximal zwei Werten eine positive Wahrscheinlichkeit annimmt.

Lösung

Sei $X : \Omega \to \Omega' \subset \mathbb{R}$. Aufgrund der Symmetrie von X erhalten wir dann die Erwartungswerte

$$\begin{aligned}
EX &= \sum_{x\in\Omega'} x \cdot P(X=x) \;=\; \sum_{\substack{x\in\Omega'\\ x>0}} x \cdot P(X=x) + \sum_{\substack{x\in\Omega'\\ x>0}} -x \cdot P(X=x) \\
&= \sum_{\substack{x\in\Omega'\\ x>0}} x \cdot P(X=x) - \sum_{\substack{x\in\Omega'\\ x>0}} x \cdot P(X=x) \;=\; 0, \\
E(X^3) &= \sum_{x\in\Omega'} x^3 \cdot P(X=x) \\
&= \sum_{\substack{x\in\Omega'\\ x>0}} x^3 \cdot P(X=x) + \sum_{\substack{x\in\Omega'\\ x>0}} (-x)^3 \cdot P(X=-x) \\
&= \sum_{\substack{x\in\Omega'\\ x>0}} x^3 \cdot P(X=x) - \sum_{\substack{x\in\Omega'\\ x>0}} x^3 \cdot P(X=-x) \;=\; 0.
\end{aligned}$$

Nach den Rechenregeln für die Kovarianz ergibt sich damit

$$\begin{aligned}
\mathrm{Cov}(X, X^2) &= E(XX^2) - (EX) \cdot (E(X^2)) \\
&= E(X^3) - (EX) \cdot (E(X^2)) \;=\; 0 - 0 \cdot (E(X^2)) \;=\; 0,
\end{aligned}$$

also sind X und X^2 unkorreliert.

Angenommen X nimmt nur einen Wert an, dann muss dieser Wert wegen der Symmetrie gleich 0 sein. Dann haben wir

$$P(X=0) \cdot P(X^2=0) \;=\; 1 \cdot 1 \;=\; 1 \;=\; P(X=0, X^2=0),$$

also sind in diesem Falle X und X^2 stochastisch unabhängig.

Angenommen X nimmt genau zwei Werte an. Dann nimmt X die Werte $x > 0$ und $-x$ an und es gilt

$$P(X=x) \;=\; P(X=-x) \;=\; \frac{1}{2}.$$

Tabelle 1.7 zeigt die gemeinsame Verteilung von X und X^2 sowie deren Randverteilungen.
Es ist sofort zu erkennen, dass X und X^2 stochastisch unabhängig sind.

X^2/X	x	$-x$	$\sum$
x^2	$1/2$	$1/2$	1
$\sum$	$1/2$	$1/2$	

Tabelle 1.7: Gemeinsame Verteilung, wenn X zwei Werte annimmt.

Nun nehme X mehr als zwei Werte an, darunter etwa x und $-x$ mit $x \neq 0$. Dann gilt

$$P(X = x) \ = \ P(X = -x) \ = \ p \qquad \text{und} \qquad P(X^2 = x) \ = \ 2p$$

für ein $0 < p < 1/2$, da X ja mehr als zwei Werte annimmt. Damit haben wir

$$P(X = x) \cdot P(X^2 = x) \ = \ 2p^2 \ \neq \ p \ = \ P(X = x, X^2 = x),$$

also sind X und X^2 nur dann stochastisch unabhängig, wenn X weniger als drei Werte annimmt.

Aufgabe 1.10.36

Sei (X_k) eien Folge von stochastisch unabhängigen und identisch Bernoulli verteilten Zufallsvariablen mit Parameter ϑ. Zeige, dass es dann ein $c > 0$ gibt, so dass für alle $\varepsilon > 0$

$$P\left(|(X_1X_{n+1} + X_2X_{n+2} + \ldots + X_nX_{n+n})/n - c| \ \geq \ \varepsilon\right) \quad \longrightarrow \quad 0$$

für $n \to \infty$ gilt.

Lösung

Wir definieren die Zufallsvariable

$$Z \ = \ \frac{1}{n} \cdot (X_1X_{n+1} + X_2X_{n+2} + \ldots + X_nX_{n+n}) \ = \ \frac{1}{n} \cdot \left(\sum_{k=1}^{n} X_kX_{n+k}\right)$$

und berechnen dessen Varianz:

$$\begin{aligned} EZ \ &= \ E\left(\frac{1}{n} \cdot \left(\sum_{k=1}^{n} X_kX_{n+k}\right)\right) \ = \ \frac{1}{n} \cdot \sum_{k=1}^{n} E(X_kX_{n+k}) \\ &= \ \frac{1}{n} \cdot \sum_{k=1}^{n} E(X_k)E(X_{n+k}) \ = \ \frac{1}{n} \cdot \sum_{k=1}^{n} \vartheta^2 \ = \ \frac{1}{n} \cdot n\vartheta^2 \ = \ \vartheta^2. \end{aligned}$$

Dabei wurde ausgenutzt, dass der Erwartungswert linear ist, dass die Folge (X_k) stochastisch unabhängig und identisch Bernoulli verteilt ist.

Mit Z haben wir eine reelle Zufallsvariable auf (Ω, P) mit $EX^2 < \infty$. Setzen wir nun $c = \vartheta^2$, so erhalten wir nach der Tschebycheff Ungleichung für alle $\varepsilon > 0$

$$\begin{aligned} & P(|Z - EZ| \geq \varepsilon) \\ = \; & P\left(|(X_1 X_{n+1} + X_2 X_{n+2} + \ldots + X_n X_{n+n})/n - c| \;\geq\; \varepsilon\right) \\ \leq \; & \frac{\mathrm{Var}(Z)}{\varepsilon^2}. \end{aligned}$$

Wir haben also noch zu zeigen, dass für $n \to \infty$ gerade $\mathrm{Var}(Z) \to 0$ gilt. Dies folgt sofort aus der folgenden Rechnung:

$$\begin{aligned} \mathrm{Var}(Z) &= \mathrm{Var}\left(\sum_{k=1}^{n} \frac{X_k\, X_{n+k}}{n}\right) \;=\; \sum_{k=1}^{n} \mathrm{Var}\left(\frac{X_k\, X_{n+k}}{n}\right) \\ &= \sum_{k=1}^{n} \mathrm{Var}\left(E\left(\frac{X_k^2 X_{n+k}^2}{n^2}\right) - E\left(\frac{X_k\, X_{n+k}}{n}\right)^2\right) \\ &= \sum_{k=1}^{n} \mathrm{Var}\left(\frac{1}{n^2}\left(E\left(X_k^2 X_{n+k}^2\right) - E\left(X_k\, X_{n+k}\right)^2\right)\right) \\ &= \frac{1}{n^2}\sum_{k=1}^{n}(E(X_k^2)E(X_{n+k}^2) - \vartheta^4) \;=\; \frac{1}{n}\sum_{k=1}^{n}(\vartheta^2 \cdot \vartheta^2 - \vartheta^4) \\ &= \frac{1}{n^2}\sum_{k=1}^{n} 0 \;=\; 0. \end{aligned}$$

Dabei haben wir mehrfach ausgenutzt, dass der Erwartungswert linear ist und dass die Folge (X_k) stochastisch unabhängig sowie identisch Bernoulli verteilt ist.

Aufgabe 1.10.37

X und Y seien zwei Zufallsvariablen mit der folgenden gemeinsamen Verteilung:

X/Y	-2	-1	1	2
-1	$1/16$	$2/16$	$2/16$	$1/16$
0	$1/16$	$1/16$	$1/16$	$1/16$
1	$1/16$	$2/16$	$2/16$	$1/16$

Berechne die Randverteilungen von X und von Y und bestimmt EX, EY, $\mathrm{Var}(X)$, $\mathrm{Var}(Y)$ sowie $\mathrm{Cov}(X, Y)$. Entscheide, ob X und Y stochastisch unabhängig sind.

X/Y	-2	-1	1	2	$P(X=l)$
-1	1/16	2/16	2/16	1/16	6/16
0	1/16	1/16	1/16	1/16	4/16
1	1/16	2/16	2/16	1/16	6/16
$P(Y=k)$	3/16	5/16	5/16	3/16	1

Tabelle 1.8: Randverteilungen von X und Y.

Lösung

Zunächst ergeben sich die Randverteilungen durch Tabelle 1.8.
Für die Erwartungswerte ergibt sich damit

$$\begin{aligned} EX &= -1 \cdot \frac{6}{16} + 0 \cdot \frac{4}{16} + 1 \cdot \frac{6}{16} = 0, \\ EY &= -2 \cdot \frac{3}{16} - 1 \cdot \frac{5}{16} + 1 \cdot \frac{5}{16} + 2 \cdot \frac{3}{16} = 0. \end{aligned}$$

Für die Varianz erhalten wir

$$\begin{aligned} \mathrm{Var}(X) &= E(X^2) - (EX)^2 = (-1)^2 \cdot \frac{6}{16} + 1^2 \cdot \frac{6}{16} - 0 = \frac{12}{16} = \frac{3}{4}, \\ \mathrm{Var}(Y) &= \left(\frac{4 \cdot 3}{16} + \frac{5}{16} + \frac{5}{16} + \frac{4 \cdot 3}{16}\right) - 0 = \frac{34}{16} = \frac{17}{8}. \end{aligned}$$

Für die Kovarianz folgt

$$\begin{aligned} \mathrm{Cov}(X,Y) &= E(XY) - EX \cdot EY = E(XY) \\ &= \left(-1 \cdot -2 \cdot \frac{1}{16}\right) + \left(-1 \cdot -1 \cdot \frac{1}{8}\right) + \left(-1 \cdot \frac{1}{8}\right) + \left(-2 \cdot \frac{1}{16}\right) \\ &\quad + \left(-2 \cdot \frac{1}{16}\right) + \left(-1 \cdot \frac{1}{8}\right) + \left(1 \cdot \frac{1}{8}\right) + \left(2 \cdot \frac{1}{16}\right) \\ &= \frac{2}{16} + \frac{1}{8} - \frac{1}{8} - \frac{2}{16} - \frac{2}{16} - \frac{1}{8} + \frac{1}{8} + \frac{2}{16} = 0. \end{aligned}$$

Wenn X und Y stochastisch unabhängig wären, dann müsste

$$P(X=k, Y=l) = P(X=k) \cdot P(Y=l)$$

für alle $k \in \{-2,-1,1,2\}$ und $l \in \{-1,0,1\}$ gelten. Es gilt jedoch zum Beispiel

$$P(X=1, Y=1) = \frac{1}{8} \neq \frac{5}{16} \cdot \frac{6}{16} = P(X=1) \cdot P(Y=1),$$

also sind X und Y nicht unabhängig.

Aufgabe 1.10.38

Zeige, dass

$$g(t) \;=\; \frac{1}{(2-t)^2}$$

die erzeugenden Funktion einer Zufallsvariablen X ist und berechne $f_X(n) = P(X = n)$ für $n \geq 0$.

Lösung

Für $|t| < 2$ gilt nach der geometrischen Reihe

$$\begin{aligned}
g(t) &= \frac{1}{(2-t)^2} = \left(\frac{1}{2}\cdot\frac{1}{1-\frac{t}{2}}\right)^2 = \left(\frac{1}{2}\sum_{n=0}^{\infty}\left(\frac{t}{2}\right)^n\right)^2 \\
&= \left(\sum_{n=0}^{\infty}\frac{1}{2^{n+1}}t^n\right)^2 = \sum_{n=0}^{\infty}\left(\sum_{k=0}^{n}\frac{1}{2^{k+1}}\cdot\frac{1}{2^{n+1-k}}\right)t^n \\
&= \sum_{n=0}^{\infty}\left(\sum_{k=0}^{n}\frac{1}{2^{n+2}}\right)t^n = \sum_{n=0}^{\infty}\frac{n+1}{2^{n+2}}t^n,
\end{aligned}$$

dabei wurde die Faltung für Potenzreihen verwendet. Damit $g(t)$ die erzeugenden Funktion einer Zufallsvariablen X ist, muss

$$\sum_{n=0}^{\infty}\frac{n+1}{2^{n+2}} \;=\; 1$$

gelten. Für $t = 1$ erhalten wir

$$\sum_{n=0}^{\infty}\frac{n+1}{2^{n+2}} \;=\; \left(\sum_{n=0}^{\infty}\frac{1}{2^{n+1}}\right)\cdot\left(\sum_{n=0}^{\infty}\frac{1}{2^{n+1}}\right)$$

und wieder nach der geometrischen Reihe gilt

$$\sum_{n=0}^{\infty}\frac{1}{2^{n+1}} \;=\; \frac{1}{2}\cdot\sum_{n=0}^{\infty}\left(\frac{1}{2}\right)^n \;=\; \frac{1}{2}\cdot\frac{1}{1-\frac{1}{2}} \;=\; 1,$$

also gilt auch

$$\sum_{n=0}^{\infty}\frac{n+1}{2^{n+2}} \;=\; 1\cdot 1 \;=\; 1.$$

Somit ist

$$f_X(n) \;=\; P(X = n) \;=\; \frac{n+1}{2^{n+2}}$$

eine Zähldichte und es gilt

$$g(t) \;=\; \sum_{n=0}^{\infty} f_X(n)t^n \;=\; \sum_{n=0}^{\infty}\frac{n+1}{2^{n+2}}t^n$$

mit dem Konvergenzradius $R = 2$.

Aufgabe 1.10.39

Für ein $0 < b < 1$ sei

$$P(\{k\}) \;=\; a\frac{b^k}{k} \qquad \text{für} \quad k \in \mathbb{N}.$$

Zeige, dass dadurch für ein geeignetes $a \in \mathbb{R}$ eine Verteilung auf $\mathbb{N}$ definiert wird und berechne die erzeugenden Funktion dieser Verteilung.

Lösung

Damit wir eine Verteilung auf $\mathbb{N}$ haben, muss

$$\sum_{k=1}^{\infty} P(\{k\}) \;=\; \sum_{k=1}^{\infty} a\frac{b^k}{k} \;=\; 1$$

gelten. Nach der Potenzreihe von $\log(1-x)$ für alle $|x| < 1$ folgt

$$\sum_{k=1}^{\infty} \frac{b^k}{k} \;=\; -\log(1-b),$$

also erhalten wir gerade für

$$a \;=\; -\frac{1}{\log(1-b)}$$

eine Verteilung. Die erzeugenden Funktion ist

$$\begin{aligned} g(t) \;&=\; \sum_{k=1}^{\infty} a\frac{b^k}{k}t^k \;=\; a\sum_{k=1}^{\infty}\frac{(tb)^k}{k} \\ &=\; -a\log(1-tb) \;=\; \frac{\log(1-tb)}{\log(1-b)} \end{aligned}$$

Aufgabe 1.10.40

Es seien X und Y zwei Zufallsvariablen, die Poisson verteilt sind mit Parameter λ bzw. μ, dabei gelte $\mu > \lambda > 0$. Weiter sei $Z = Y - X$ und X und Z seien stochastisch unabhängig.

Zeige, dass dann Z Poisson verteilt mit Parameter $\mu - \lambda$ ist.

Lösung

Zunächst haben wir die Zähldichten

$$f_X(k) \;=\; \frac{\lambda^k}{k!}\mathrm{e}^{-\lambda} \qquad \text{und} \qquad f_Y(k) \;=\; \frac{\mu^k}{k!}\mathrm{e}^{-\mu}$$

für $k \geq 0$. Weiter gilt $X + Z = X + Y - X = Y$, also folgt sofort

$$f_{X+Z}(k) \;=\; f_Y(k) \;=\; \frac{\mu^k}{k!}\mathrm{e}^{-\mu}.$$

Angenommen, Z ist wirklich Poisson verteilt mit Parameter $\mu - \lambda$, dann muss auch nach der Faltungsformel

$$f_{X+Z}(k) \;=\; f_Y(k)$$

folgen, da X und Z stochastisch unabhängig sind. Die angenommene Zähldichte ist

$$f_Z(k) \;=\; \frac{(\mu-\lambda)^k}{k!}\mathrm{e}^{-\mu+\lambda}.$$

Damit berechnen wir nach der Faltungsformel und der Binomischen Formel

$$\begin{aligned} f_{X+Z}(k) &= \sum_{i=0}^{k} \frac{\lambda^i}{i!}\mathrm{e}^{-\lambda}\frac{(\mu-\lambda)^{k-i}}{(k-i)!}\mathrm{e}^{-\mu+\lambda} \\ &= \mathrm{e}^{-\lambda-\mu+\lambda} \cdot \frac{1}{k!}\sum_{i=0}^{k}\frac{k!}{i!(k-i)!}\lambda^i(\mu-\lambda)^{k-i} \\ &= \mathrm{e}^{-\mu} \cdot \frac{(\lambda+\mu-\lambda)^k}{k!} \;=\; \frac{\mu^k}{k!}\mathrm{e}^{-\mu} \;=\; f_Y(k), \end{aligned}$$

demnach ist Z wirklich Poisson verteilt mit dem Parameter $\mu - \lambda > 0$.

Aufgabe 1.10.41

Zeige, dass

$$g(t) \;=\; a(\mathrm{e}^t - \mathrm{e}^{-t})$$

für ein geeignetes $a \in \mathbb{R}$ die erzeugenden Funktion einer Zufallsvariablen X ist und berechne EX sowie $\mathrm{Var}(X)$.

Lösung

Wir haben zunächst eine Zähldichte $f : \mathbb{N} \cup \{0\} \to \mathbb{R}$ zu finden, für die

$$g(t) \;=\; \sum_{n=0}^{\infty} f(n)t^n \;=\; a(\mathrm{e}^t - \mathrm{e}^{-t})$$

folgt. Für den Sinushyperbolicus gilt

$$\sinh(t) \;=\; \frac{1}{2}(\mathrm{e}^t - \mathrm{e}^{-t}) \;=\; \sum_{n=0}^{\infty}\frac{1}{(2n+1)!}\,t^{2n+1}.$$

Betrachten wir dazu die Reihe

$$\sum_{n=0}^{\infty} \frac{1}{(2n+1)!},$$

so ist nach dem Quotientenkriterium sofort klar, dass diese Reihe konvergent ist. Der Grenzwert sei $b > 0$. Wir setzen nun $a = 1/2b$ und führen die Funktion

$$f(n) = \begin{cases} 0 & \text{für } n \text{ gerade} \\ \frac{2a}{n!} & \text{für } n \text{ ungerade} \end{cases}$$

für $n \geq 0$ ein. Damit gilt

$$\sum_{n=0}^{\infty} f(n) = \sum_{\substack{n=0 \\ n \text{ ungerade}}}^{\infty} \frac{2a}{n!} = 2a \cdot \sum_{n=0}^{\infty} \frac{1}{(2n+1)!} = 2a \cdot b = 1,$$

also ist f eine Zähldichte auf $\mathbb{N} \cup \{0\}$. Zusammen erhalten wir wie zu zeigen war

$$\sum_{n=0}^{\infty} f(n)t^n = \sum_{n=0}^{\infty} \frac{2a}{(2n+1)!} t^{2n+1} = 2a \sinh(t) = a(\mathrm{e}^t - \mathrm{e}^{-t}).$$

Weiter gilt

$$g'(t) = a(\mathrm{e}^t + \mathrm{e}^{-t}) \qquad \text{und} \qquad g''(t) = a(\mathrm{e}^t - \mathrm{e}^{-t}),$$

damit erhalten wir sofort den Erwartungswert und die Varianz:

$$\begin{aligned} EX &= g'(1) = a\left(\mathrm{e} + \frac{1}{\mathrm{e}}\right), \\ \mathrm{Var}(X) &= g''(1) + g'(1) - (g'(1))^2 \\ &= a\left(\mathrm{e} - \frac{1}{\mathrm{e}} + \mathrm{e} + \frac{1}{\mathrm{e}}\right) - a^2\left(\mathrm{e} + \frac{1}{\mathrm{e}}\right)^2 = 2a\mathrm{e} - a^2\left(\mathrm{e} + \frac{1}{\mathrm{e}}\right)^2. \end{aligned}$$

Durch

$$\begin{aligned} b = \sinh(1) &= \sum_{n=0}^{\infty} \frac{1}{(2n+1)!} 1^{2n+1} \\ &= \sum_{n=0}^{\infty} \frac{1}{(2n+1)!} = \frac{1}{2}(\mathrm{e}^1 - \mathrm{e}^{-1}) \approx 1.175 \end{aligned}$$

erhalten wir $a = 1/2b \approx 0.425$ und damit die ungefähren Werte

$$EX = 1.313 \qquad \text{und} \qquad \mathrm{Var}(X) = 0.589\,.$$

2 Stetige Wahrscheinlichkeitsräume

Bisher haben wir zu einem Grundraum Ω stets Wahrscheinlichkeitsmaße

$$P : \mathcal{P}(\Omega) \to [0, 1]$$

betrachtet. Dabei ist es aber möglich, dass die Potenzmenge $\mathcal{P}$ zu mächtig ist. Wir werden daher ein Mengensystem $\mathcal{K} \subset \mathcal{P}(\Omega)$ einführen, welches wir Algebra nennen. Ein Wahrscheinlichkeitsmaß auf Ω werden wir dann durch geeignete Abbildungen

$$P : \mathcal{K} \to [0, 1]$$

definieren. Viele der Ideen aus den diskreten Wahrscheinlichkeitsräumen werden wir aber geeignet übertragen können. Es kann also hilfreich sein, Definitionen und Sätze aus diesem Kapitel jeweils mit dem vorherigen zu vergleichen.

2.1 Allgemeine Wahrscheinlichkeitsräume

Definition 2.1.1

Sei $\Omega \neq \emptyset$ ein Grundraum. Ein Mengensystem $\mathcal{K} \subset \mathcal{P}(\Omega)$ heißt ***σ-Algebra*** in Ω, wenn gilt:

(1) Es ist $\Omega \in \mathcal{K}$.

(2) Mit $A \in \mathcal{K}$ ist auch $A^C \in \mathcal{K}$.

(3) Mit $(A_i)_{i \in \mathbb{N}} \in \mathcal{K}$ ist auch $\bigcup\limits_{i=1}^{\infty} A_i \in \mathcal{K}$.

Das Tupel $(\Omega, \mathcal{K})$ heißt dann ***Meßraum***.

Bemerkung

Für eine σ-Algebra $\mathcal{K}$ ergeben sich nach Definition sofort die folgenden Aussagen:

(1) Es ist $\emptyset \in \mathcal{K}$.

(2) Nach der DeMorganschen Regel ist mit $(A_i)_{i\in\mathbb{N}} \in \mathcal{K}$ auch $\bigcap\limits_{i=1}^{\infty} A_i \in \mathcal{K}$.

(3) Für endlich viele $A_1, ..A_n \in \mathcal{K}$ gilt auch

$$A_1 \cup \ldots \cup A_n \in \mathcal{K} \qquad \text{und} \qquad A_1 \cap \ldots \cap A_n \in \mathcal{K}$$

Proposition 2.1.2

Für eine endliche Indexmenge I seien $\mathcal{K}_i \subset \mathcal{P}(\Omega)$ beliebige σ-Algebren.

Dann ist auch

$$\bigcap_{i\in I} \mathcal{K}_i$$

eine σ-Algebra in Ω.

Die drei Punkte der Definition lassen sich leicht zeigen.

Definition und Satz 2.1.3

Sei $\mathcal{E} \in \mathcal{P}(\Omega)$. Dann gibt es [bezüglich Inklusion] eine kleinste σ-Algebra $\mathcal{K}(\mathcal{E})$ in Ω, die $\mathcal{E}$ enthält. Es ist also

$$\mathcal{K}(\mathcal{E}) \;=\; \bigcap_{\substack{\mathcal{K}\text{ ist }\sigma\text{-Algebra}\\ \mathcal{E}\subset\Omega}} \mathcal{K}.$$

$\mathcal{K}(\mathcal{E})$ heißt die von $\mathcal{E}$ ***erzeugte σ-Algebra*** bzw. $\mathcal{E}$ heißt der ***Erzeuger*** von $\mathcal{K}(\mathcal{E})$.

Beispiel Borelsche Mengen

Sei $\Omega = \mathbb{R}^k$ und seien $a = (a_1, .., a_k)$ und $b = (b_1, .., b_k)$ aus $\mathbb{R}^k$ mit $a_i \le b_i$ für $i = 1, .., k$. Weiter sei

$$]a, b] \;=\; \prod_{i=1}^{k}]a_i, b_i] \;=\; \{x = (x_1, .., x_k) \in \mathbb{R}^k \mid a_i < x_i \le b_i,\ i = 1, .., k\}$$

ein halboffenes Intervall in $\mathbb{R}^k$. Dazu sei nun

$$\mathcal{E} \;=\; \{]a, b] \mid a_i \le b_i,\ i = 1, .., k\}.$$

Dann heißt $\mathcal{B}^k := \mathcal{K}(\mathcal{E})$ die ***Borelsche σ-Algebra***.

Es lässt sich zeigen, dass nicht nur halboffene sondern auch offene sowie geschlossene Intervalle in $\mathcal{B}^k$ liegen. Trotzdem gibt es Teilmengen des $\mathbb{R}^k$, die nicht in $\mathcal{B}^k$ enthalten sind. Derartige Menge können nicht explizit angegeben werden, deren Existenz kann jedoch durch das Auswahlaxiom gezeigt werden.

Die Mächtigkeit der Menge $\mathcal{B}^k$ ist gleich der Mächtigkeit von $\mathbb{R}$ und damit gleiner gleich der Mächtigkeit von $\mathbb{R}^k$.

Definition 2.1.4

Das Tripel $(\Omega, \mathcal{K}, P)$ heißt ***Wahrscheinlichkeitsraum***, wenn gilt:

(1) $\mathcal{K}$ ist eine σ-Algebra in Ω.

(2) $P : \mathcal{K} \to [0,1]$ ist eine Wahrscheinlichkeitsmaß auf $\mathcal{K}$.

Das heißt, dass P gerade σ-Additiv ist und dass $P(\mathcal{K}) = 1$ gilt.

Definition 2.1.5

Eine Funktion $F : \mathbb{R} \to [0,1]$ heißt ***Verteilungsfunktion*** auf $\mathbb{R}$, wenn gilt:

(1) F ist monoton wachsend.

(2) F ist rechtsseitig stetig.

(3) F ist normiert, es gilt also $\lim\limits_{x\to\infty} F(x) = 1$ und $\lim\limits_{x\to-\infty} F(x) = 0$.

Satz 2.1.6 (Korrespondenzsatz)

Ist P ein Wahrscheinlichkeitsmaß auf $(\mathbb{R}, \mathcal{B}^1)$, dann ist

$$F_P(x) \; = \; P(]-\infty, x]) \qquad \text{für } \; x \in \mathbb{R}$$

eine Verteilungsfunktion auf $\mathbb{R}$.

Ist umgekehrt F eine Verteilungsfunktion auf $\mathbb{R}$, dann gibt es genau ein Wahrscheinlichkeitsmaß P auf $(\mathbb{R}, \mathcal{B}^1)$ mit $F = F_P$.

Folgerung 2.1.7

Ist F eine Verteilungsfunktion und P das zugehörige Wahrscheinlichkeitsmaß, dann gilt für halboffene Intervalle $]a,b] \subset \mathbb{R}$

$$P(]a,b]) \; = \; F(b) - F(a).$$

Für paarweise disjunkte halboffene Intervalle $]a_i, b_i] \subset \mathbb{R}$ gilt

$$P\left(\bigcup_{i=1}^{n}]a_i, b_i]\right) \; = \; \sum_{i=1}^{n} (F(b_i) - F(a_i)).$$

Betrachten wir eine offene Menge $]a,b[\subset \mathbb{R}$, so gilt

$$\begin{aligned} P(]a,b[) \; &= \; \lim_{n\to\infty} P\left(\left]a,\, b - \frac{1}{n}\right]\right) \\ &= \; \lim_{n\to\infty} \left(F\left(b - \frac{1}{n}\right) - F(a)\right) \; = \; F(b-) - F(a). \end{aligned}$$

Dabei ist nun wesentlich, dass F rechtsseitig stetig ist, denn daher kann $b- \neq b$ gelten.

Definition 2.1.8

Eine Funktion $f : \mathbb{R} \to [0, \infty[$ heißt ***Riemanndichte*** auf $\mathbb{R}$, wenn gilt:

(1) f ist Riemann integrierbar auf $\mathbb{R}$.

(2) Es gilt
$$\int_{-\infty}^{\infty} f(x)\,\mathrm{d}x \;=\; 1.$$

Proposition 2.1.9

Sei f eine Riemanndichte auf $\mathbb{R}$. Dann gilt

(1) Die Funktion
$$F(x) \;=\; \int_{-\infty}^{x} f(y)\,\mathrm{d}y$$
ist eine stetige Verteilungsfunktion.

(2) Ist C_f die Menge der Stetigkeitsstellen von f und ist $x \in C_f$, dann gilt nach dem Hauptsatz der Analysis $F'(x) = f(x)$.

(3) Zu f gibt es nach dem Korrespondenzsatz genau ein Wahrscheinlichkeitsmaß $P = P_F$ mit der Verteilungsfunktion F.

f heißt dann einfach die ***Dichte*** von P_F.

(4) Für alle $x \in \mathbb{R}$ gilt $P_F(\{x\}) = 0$.

Beispiel Gleichverteilung

Wir betrachten die Funktion
$$f(x) \;=\; \begin{cases} \frac{1}{b-a} & \text{für} \quad a \le x \le b \\ 0 & \text{sonst} \end{cases} .$$

Dann gilt $f \ge 0$, f ist Riemann integrierbar und
$$\int_{-\infty}^{\infty} f(x)\,\mathrm{d}x \;=\; \frac{1}{b-a}\int_a^b 1\,\mathrm{d}x \;=\; 1,$$

also ist f eine Dichte auf $\mathbb{R}$. Die Verteilungsfunktion ist
$$F(x) \;=\; \int_{-\infty}^{x} f(y)\,\mathrm{d}y \;=\; \frac{1}{b-a}\int_{-\infty}^{x} 1_{[a,b]}(y)\,\mathrm{d}y \;=\; \begin{cases} 0 & \text{für} \quad x \le a \\ \frac{x-a}{b-a} & \text{für} \quad a < x \le b \\ 1 & \text{für} \quad x > b \end{cases} .$$

Das Wahrscheinlichkeitsmaß $R(a,b) = P_F$ heißt ***Gleichverteilung*** auf dem Intervall $[a,b]$.

Beispiel Exponentialverteilung

Sei $\alpha > 0$ eine reelle Zahl und sei

$$f(x) \;=\; \alpha \mathrm{e}^{-\alpha x} \qquad \text{für } \; x \geq 0$$

und $f(x) = 0$ für $x < 0$. Dann gilt wieder $f \geq 0$, f ist Riemann integrierbar und durch

$$\int_{-\infty}^{\infty} f(x)\,\mathrm{d}x \;=\; \alpha \int_0^{\infty} \mathrm{e}^{-\alpha x}\,\mathrm{d}x \;=\; 1$$

ist f eine Dichte auf $\mathbb{R}$. Die Verteilungsfunktion ist

$$F(x) \;=\; \begin{cases} 0 & \text{für } \; x < 0 \\ 1 - \mathrm{e}^{-\alpha x} & \text{für } \; x \geq 0 \end{cases},$$

denn es gilt

$$\alpha \int_0^{x} \mathrm{e}^{-\alpha y}\,\mathrm{d}y \;=\; 1 - \mathrm{e}^{-\alpha x}.$$

Das Wahrscheinlichkeitsmaß $E(\alpha) = P_F$ heißt ***Exponentialverteilung*** mit Parameter α. Ähnlich wie die Poissonverteilung in diskreten Wahrscheinlichkeitsräumen wird auch die Exponentialverteilung zur Modellierung von Lebensdauern radioaktiver Teilchen oder Bauteile verwendet.

Beispiel Normalverteilung

Sei

$$\varphi(x) \;=\; \frac{1}{\sqrt{2\pi}} \mathrm{e}^{-\frac{1}{2}x^2}.$$

Dann ist φ eine Dichte auf $\mathbb{R}$. Dies lässt sich zeigen, indem man $\varphi(x)\varphi(y)$ betrachtet und mittels Transformationsformel über $\mathbb{R}^2$ integriert.

Das zu dieser Dichte gehörige Wahrscheinlichkeitsmaß $N(0,1) = P_F$ heißt ***Standard Normalverteilung***.

Diese Verteilung kann jedoch durch zwei Parameter modifiziert werden: Ist $\mu \in \mathbb{R}$ und $\sigma > 0$ reell, dann gilt

$$\varphi_{\mu,\sigma^2}(x) \;=\; \frac{1}{\sigma}\varphi\left(\frac{x-\mu}{\sigma}\right) \;=\; \frac{1}{\sqrt{2\pi\sigma^2}} \mathrm{e}^{-\frac{1}{2\sigma^2}(x-\mu)^2}.$$

φ_{μ,σ^2} ist ebenfalls eine Dichte und beschreibt eine Verschiebung des Maxima der Standard Normalverteilung. $N(\mu, \sigma^2) = P_F$ heißt ***Normalverteilung*** mit Mittelwert μ und Varianz σ^2.

Bis hierher wurden nur eindimensionale Dichten betrachtet. Dies soll nun für $(\mathbb{R}^k, \mathcal{B}^k)$ verallgemeinert werden.

Definition 2.1.10

Eine Funktion $f : \mathbb{R}^k \to [0, \infty[$ heißt ***Riemanndichte*** auf $\mathbb{R}^k$, wenn gilt:

(1) f ist Riemann integrierbar auf $\mathbb{R}^k$.

(2) Es gilt

$$\int_{\mathbb{R}^k} f(x)\,\mathrm{d}x \;=\; 1.$$

Proposition 2.1.11

Sei f eine Riemanndichte auf $\mathbb{R}^k$. Dann gibt es genau ein Wahrscheinlichkeitsmaß P auf $\mathcal{B}^k$ mit

$$P(]a,b]) \;=\; \int_{a_1}^{b_1} .. \int_{a_k}^{b_k} f(x)\,\mathrm{d}x_k..\mathrm{d}x_1$$

für alle $a = (a_1, .., a_k), b = (b_1, .., b_k) \in \mathbb{R}^k$ mit $a_i < b_i$ für $i = 1, .., k$.

f heißt dann die ***Dichte*** von P.

Sei umgekehrt P ein Wahrscheinlichkeitsmaß auf $\mathbb{R}^k$. Dann heißt

$$F_P(x) \;=\; P(]-\infty, x])$$

die eindeutig bestimmte ***Verteilungsfunktion*** von P.

Beispiel Gleichverteilung

Sei $G \subset \mathbb{R}^k$ offen und der Rand ∂G von G sei eine Nullmenge. Wir definieren

$$f(x) \;=\; \begin{cases} \frac{1}{\lambda(G)} & \text{für } \; x \in G \\ 0 & \qquad \text{sonst} \end{cases},$$

wobei $\lambda(G) = \int_{\mathbb{R}}^k 1_G(x)\,\mathrm{d}x$ das Volumen von G beschreibt. Damit ist f eine Dichte auf $\mathbb{R}^k$ und das zugehörige Wahrscheinlichkeitsmaß heißt ***Gleichverteilung*** auf G.

Proposition 2.1.12

Seien P_i Wahrscheinlichkeitsmaße auf $(\mathbb{R}, \mathcal{B}^k)$ mit Riemanndichten f_i für $i = 1, .., k$.

Dann ist

$$f(x) \;=\; \prod_{i=1}^{k} f_k(x_i) \qquad \text{für } \; x = (x_1, .., x_k) \in \mathbb{R}^k$$

eine Riemanndichte auf $\mathbb{R}^k$. Das zugehörige Wahrscheinlichkeitsmaß

$$P \;=\; \bigoplus_{i=1}^{k} P_i$$

heißt ***Produktmaß*** der P_1. Es gilt

$$P(]a,b]) \;=\; \prod_{i=1}^{k} P_i(]a_i,b_i]) \;=\; \prod_{i=1}^{k}\left(\int_{a_i}^{b_i} f_i(x)\,\mathrm{d}x\right)$$

für alle $a = (a_1,..,a_k), b = (b_1,..,b_k) \in \mathbb{R}^k$ mit $a_i < b_i$ für $i = 1,..,k$.

Definition 2.1.13

Seien $(\Omega_i, \mathcal{K}_i)$ für $i = 1, 2$ zwei Meßräume und sei $g : \Omega_1 \to \Omega_2$.

Die Abbildung $g(\mathcal{K}_1, \mathcal{K}_2)$ heißt genau dann ***messbar***, wenn $g^{-1}(A_2) \in \mathcal{K}_1$ für alle $A_2 \in \mathcal{K}_2$.

Schreibweise: $g : (\Omega_1, \mathcal{K}_1) \to (\Omega_2, \mathcal{K}_2)$.

Satz 2.1.14

Sei $\mathcal{K}_1$ eine σ-Algebra auf Ω_1 und sei

$$g : (\Omega_1, \mathcal{K}_1) \to (\Omega_2, \mathcal{K}_2)$$

eine Abbildung. Dann ist

$$\mathcal{K}_g \;=\; \{A_2 \subset \Omega_2 \mid g^{-1}(A_2) \in \mathcal{K}_1\}$$

eine σ-Algebra auf Ω_2.

Dies ist die größte σ-Algebra, so dass g messbar ist.

2.2 Transformation und Erwartungswert

Definition 2.2.1

Sei $g : (\Omega_1, \mathcal{K}_1) \to (\Omega_2, \mathcal{K}_2)$ und sei P ein Wahrscheinlichkeitsmaß auf $(\Omega_1, \mathcal{K}_1)$.

Dann heißt

$$P^g \qquad \text{mit} \qquad P^g(A_2) \;=\; P(g^{-1}(A_2))$$

für alle $A_2 \in \mathcal{K}_2$ die ***Verteilung*** von g bezüglich P.

P^g ist ein Wahrscheinlichkeitsmaß auf $(\Omega_2, \mathcal{K}_2)$.

Beispiel 2.2.2

Das Wahrscheinlichkeitsmaß P habe die Dichte f auf $(\mathbb{R}^2, \mathcal{B}^2)$. Wir definieren die erste Projektion

$$\pi_1 : \mathbb{R}^2 \to \mathbb{R} \qquad \text{mit} \qquad \pi_1(x,y) \;=\; x.$$

Da π_1 stetig und somit auch messbar ist, gilt

$$\begin{aligned} P^{\pi_1}(B) &= P(\pi^{-1}(B)) \;=\; P(B \times \mathbb{R}) \\ &= \int_{B\times\mathbb{R}} f(x,y)\,\mathrm{d}y\mathrm{d}x \;=\; \int_B \left(\int_{\mathbb{R}} f(x,y)\,\mathrm{d}y \right) \mathrm{d}x. \end{aligned}$$

P^{π_1} hat also die Dichte

$$h(x) \;=\; \int_{\mathbb{R}} f(x,y)\,\mathrm{d}y$$

und heißt ***Randverteilung*** von P.

Satz 2.2.3 (Transformationsformel)

Sei $G \subset \mathbb{R}$ ein Intervall und sei $g : G \to \mathbb{R}$ eine stetig differenzierbare und injektive Transformation. Sei weiter P ein Wahrscheinlichkeitsmaß auf $(\mathbb{R}, \mathcal{B}^1)$ mit Dichte f und Träger G, das heißt es gilt $f(x) = 0$ für $x \in \mathbb{R} \setminus G$.

Dann hat P^g die Dichte

$$h(x) \;=\; f(g^{-1}(x)) \cdot \left| \frac{\mathrm{d}}{\mathrm{d}x} g^{-1}(x) \right|$$

für $x \in g(G)$ und $h(x) = 0$ sonst.

Beispiel 2.2.4

Sei $P = R(0,1)$ gleichverteilt, wir betrachten also die Dichte

$$f(x) \;=\; \begin{cases} 1 & \text{für} \quad 0 < x \leq 1 \\ 0 & \qquad \text{sonst} \end{cases}.$$

Damit hat f den Träger $G =]0,1]$ und für alle $x \in G$ gilt $f(x) = 1$. Es sei weiter

$$g : G \to \mathbb{R} \qquad \text{mit} \qquad g(x) \;=\; \frac{1}{x}.$$

Damit ist g stetig differenzierbar und injektiv auf G. Für P^g erhalten wir somit die Dichte

$$\begin{aligned} h(x) &= f(g^{-1}(x)) \cdot \left| \frac{\mathrm{d}}{\mathrm{d}x} g^{-1}(x) \right| \\ &= 1 \cdot \left| \frac{\mathrm{d}}{\mathrm{d}x} \frac{1}{x} \right| \;=\; \left| -\frac{1}{x^2} \right| \;=\; \frac{1}{x^2} \end{aligned}$$

für $x > 1$ und $h(x) = 0$ sonst.

Satz 2.2.5 (Faltungsformel)

Sei P ein Wahrscheinlichkeitsmaß auf $(\mathbb{R}^2, \mathcal{B}^2)$ mit Dichte f. Weiter sei $g : \mathbb{R}^2 \to \mathbb{R}$ mit $g(x, y) = x + y$.

Dann hat P^g die Dichte

$$h(x) \;=\; \int_{\mathbb{R}} f(x-y, y)\,\mathrm{d}x.$$

Definition 2.2.6

Sei $(\Omega, \mathcal{K}, P)$ ein Wahrscheinlichkeitsraum und sei $X : (\Omega, \mathcal{K}) \to (\mathbb{R}, \mathcal{B}^1)$ eine stetige Zufallsvariable mit Dichte f.

Dann heißt

$$EX \;=\; \int x f(x)\,\mathrm{d}x$$

genau dann der ***Erwartungswert*** von X, wenn

$$\int |x| f(x)\,\mathrm{d}x \;<\; \infty$$

gilt. Damit erhalten wir auch die übliche Definition von Varianz und Kovarianz:

$$\mathrm{Var}(X) \;=\; E((X-EX)^2) \quad \text{und} \quad \mathrm{Cov}(X,Y) \;=\; E((X-EX)(Y-EY)).$$

Beispiel Gleichverteilung

Sei X auf $[a, b]$ gleichverteilt, also $P^X = R(a, b)$ $f(x) = 1/(b-a)$ für $x \in [a, b]$ und $f(x) = 0$ sonst. Dann gilt

$$EX \;=\; \int_{\mathbb{R}} x f(x)\,\mathrm{d}x \;=\; \frac{1}{b-a}\int_a^b x\,\mathrm{d}x \;=\; \frac{1}{b-a}\cdot\frac{b^2-a^2}{2} \;=\; \frac{a+b}{2}.$$

Beispiel Normalverteilung

Sei X normalverteilt, also $P^X = N(\mu, \sigma^2)$. Wir betrachten also die Dichte

$$\varphi_{\mu,\sigma^2}(x) \;=\; \frac{1}{\sqrt{2\pi\sigma^2}}\mathrm{e}^{-\frac{1}{2\sigma^2}(x-\mu)^2}.$$

Wir nutzen nun aus, dass für die Dichte der Standard Normalverteilung

$$\int_{-\infty}^{\infty} \frac{1}{\sqrt{2\pi}}\mathrm{e}^{-\frac{1}{2}y^2}\,\mathrm{d}y \;=\; 1$$

gilt und erhalten durch Substitution mit $(x-\mu)/2$ den Erwartungswert

$$\begin{aligned} EX &= \int_{-\infty}^{\infty} x \frac{1}{\sqrt{2\pi\sigma^2}} \mathrm{e}^{-\frac{1}{2\sigma^2}(x-\mu)^2} \,\mathrm{d}x = \int_{-\infty}^{\infty} (\sigma y + \mu) \frac{1}{\sqrt{2\pi}} \mathrm{e}^{-\frac{y^2}{2}} \,\mathrm{d}y \\ &= \frac{\sigma}{\sqrt{2\pi}} \int_{-\infty}^{\infty} y\mathrm{e}^{-\frac{y^2}{2}} \,\mathrm{d}y + \mu \int_{-\infty}^{\infty} \frac{1}{\sqrt{2\pi}} \mathrm{e}^{-\frac{y^2}{2}} \,\mathrm{d}y = \mu. \end{aligned}$$

Für die Varianz erhalten wir etwas komplizierter aber analog

$$\begin{aligned} \mathrm{Var}(X) &= E(X-\mu)^2 = \int_{-\infty}^{\infty} (x-\mu)^2 \frac{1}{\sqrt{2\pi\sigma^2}} \mathrm{e}^{-\frac{1}{2\sigma^2}(x-\mu)^2} \,\mathrm{d}x \\ &= \int_{-\infty}^{\infty} \sigma^2 y^2 \frac{1}{\sqrt{2\pi}} \mathrm{e}^{-\frac{y^2}{2}} \,\mathrm{d}x = \frac{\sigma^2}{\sqrt{2\pi}} \int_{-\infty}^{\infty} y \left(y\mathrm{e}^{-\frac{y^2}{2}} \right) \mathrm{d}x = \sigma^2. \end{aligned}$$

Definition 2.2.7

Zwei stetige Zufallsvariablen $X, Y : (\Omega, \mathcal{K}) \to (\mathbb{R}, \mathcal{B}^1)$ heißen ***stochastisch unabhängig***, wenn für alle $A = \{x \in B_1\}$ und $B = \{x \in B_2\}$ für $B_1, B_2 \in \mathcal{B}^1$

$$P(A \cap B) = P(A) \cdot P(B)$$

gilt.

Folgerungen 2.2.8

Wie im diskreten Falle gelten auch hier die folgenden Rechenregeln für zwei stetige Zufallsvariablen X und Y:

(1) Der Erwartungswert ist Linear, für $a, b \in \mathbb{R}$ gilt also

$$E(aX + bY) = aEX + bEY.$$

(2) Sind x und Y unabhängige stetige Zufallsvariablen, deren Erwartungswerte existieren, so gilt der ***Multiplikationssatz***

$$EXY = EX \cdot EY.$$

(3) Ist f die Dichte von f und $g : (\mathbb{R}, \mathcal{B}^1) \to (\mathbb{R}, \mathcal{B}^1)$, so dass $E(g \circ X)$ existiere. Dann gilt die ***Transformationsformel***

$$E(g \circ X) = \int g(x) f(x) \,\mathrm{d}x.$$

2.3 Das starke Gesetz großer Zahlen

Zunächst möchten wir festhalten, dass die Tschebycheff Ungleichung auch für reelle stetige Zufallsvariablen $X : (\Omega, \mathcal{K}) \to (\mathbb{R}, \mathcal{B}^1)$. Somit gilt auch das schwache Gesetz großer Zahlen: Für eine Folge $(X_n)_{n\in\mathbb{N}}$ von stochastisch unabhängigen Zufallsvariablen mit gleichen Erwartungswerten $EX = \mu$ und mit $EX_n^2 < \infty$ gilt

$$\frac{1}{n}\sum_{i=1}^{n} X_i \quad \xrightarrow{P} \quad \mu$$

für $n \to \infty$. Dabei betrachten wir die stochastische Konvergenz. Für das starke Gesetz großer Zahlen benötigen wir eine weitere Konvergenzaussage:

Definition 2.3.1

Seien $Y_n, Y : (\Omega, \mathcal{K}) \to (\mathbb{R}, \mathcal{B}^1)$ für $n \in \mathbb{N}$ Zufallsvariablen auf $(\Omega, \mathcal{K}, P)$.

Y_n konvergiert genau dann ***fast sicher*** gegen Y, wenn

$$P\left(\left\{\omega \in \Omega \;\middle|\; \lim_{n\to\infty} X_n(\omega) = X(\omega)\right\}\right) \;=\; 1$$

gilt. Schreibweise:

$$Y_n \to Y \;[P].$$

Proposition 2.3.2

Aus der fast sicheren Konvergenz folgt die stochastische Konvergenz.

Die Umkehrung ist im Allgemeinen falsch, es lassen sich recht einfach Gegenbeispiele konstruieren.

Satz 2.3.3 (Lemma von Borel-Cantelli)

Sei $(\Omega, \mathcal{K}, P)$ ein Wahrscheinlichkeitsraum, seien $A_n \in \mathcal{K}$ für $n \in \mathbb{N}$ und sei

$$\begin{aligned} A^* &= \limsup_{n\to\infty} A_n = \bigcap_{n=1}^{\infty}\bigcup_{m=n}^{\infty} A_m \\ &= \{\omega \in \Omega \mid \omega \in A_m \text{ für unendlich viele } m\}. \end{aligned}$$

(1) Gilt $\sum\limits_{n=1}^{\infty} P(A_n) < \infty$, so folgt $P(A^*) = 0$.

(2) Sind die A_n stochastisch unabhängig und gilt $\sum\limits_{n=1}^{\infty} P(A_n) = \infty$, so folgt $P(A^*) = 1$.

Beispiel 2.3.4

Wir betrachten unendlich viele Urnen, die für alle $n \in \mathbb{N}$ durchnummeriert sind. Die $n$te Urne enthält eine weiße und $n-1$ schwarke Kugeln. Es wird unabhängig aus jeder Urne eine Kugel gezogen.

Wir wollen untersuchen, wieviele weiße Kugeln gezogen werden. Dazu definieren wir mit A_n das Ereignis, dass die aus der $n$ten Urne gezogene Kugel weiß ist. Damit gilt

$$P(A_n) \;=\; \frac{1}{n} \qquad \text{für } \; n \in \mathbb{N}$$

und wir erhalten die harmonische Reihe

$$\sum_{n=1}^{\infty} P(A_n) \;=\; \sum_{n=1}^{\infty} \frac{1}{n} \;=\; \infty.$$

Nach dem Lemma von Borel-Cantelli folgt also

$$P(\limsup_{n\to\infty} A_n) \;=\; 1,$$

das heißt, es werden fast sicher unendlich viele Kugeln gezogen.

Beispiel Irrfahrten

Wir betrachten den Wahrscheinlichkeitsraum $(\Omega, \mathcal{K}, P)$ und defininieren für $n \in \mathbb{N}$ die Zufallsvariablen

$$X_n : (\mathcal{K}, \Omega) \to (\{-1,1\},\, \mathcal{P}(\{-1,1\})).$$

Die X_n seien stochastisch unabhängig und es gelte

$$P(X_n = 1) \;=\; \vartheta \qquad \text{sowie} \qquad P(X_n = -1) \;=\; 1 - \vartheta.$$

Weiter definieren wir

$$s_n \;=\; \sum_{k=1}^{n} X_n$$

und nennen die Folge $s = (s_n)$ eine ***Irrfahrt***.

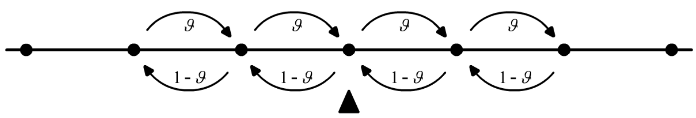


Abbildung 2.1: Irrfahrt mit Parameter ϑ.

Wir interessieren und für das Ereignis

$$B_n \;=\; \{S_n = 0\},$$

das heißt für die Wahrscheinlichkeit, dass (s_n) im n ten Schritt wieder am Startpunkt ist. Es gilt $B_n \in \mathcal{K}$ für alle $n \in \mathbb{N}$ und das Lemma von Borel-Cantelli liefert

$$P(\limsup_{n\to\infty} B_n) \;=\; 0,$$

wenn $\vartheta \neq 1/2$ und

$$P(\limsup_{n\to\infty} B_n) \;=\; 1$$

für $\vartheta = 1/2$.

Satz 2.3.5 (Starkes Gesetz großer Zahlen)

Version 1

Es sei $(X_n)_{n\in\mathbb{N}}$ eine Folge von reellen unkorrelierten Zufallsvariablen auf $(\Omega, \mathcal{K}, P)$, für die

$$\mathrm{Var}(X_k) \;\leq\; M \;<\; \infty$$

gilt. Dann folgt für $n \to \infty$

$$\frac{1}{n}\sum_{k=1}^{n}(X_k - EX_k) \quad\longrightarrow\quad 0\ [P].$$

Version 2

Es sei $(X_n)_{n\in\mathbb{N}}$ eine Folge von reellen, unabhängigen und identisch verteilten Zufallsvariablen auf $(\Omega, \mathcal{K}, P)$ mit $E|X_1| < \infty$

Dann gilt

$$\frac{1}{n}\sum_{k=1}^{n} X_k \quad\longrightarrow\quad EX_1\ [P].$$

Da die X_k identisch verteilt sind, gilt $E|X_k| = E|X_1| < \infty$ für alle $k \in \mathbb{N}$.

Beispiel Schätzer

Gegeben sind Daten $(x_1, .., x_n)$ un ein Modell, wie diese Daten abhängig von einem umbekannten Parameter erzeugt sind.

Ein ***Schätzer*** bezüglich dieser Daten ist eine Funktion $g(x_1, .., x_n)$ mit den folgenden drei Eigenschaften:

(1) Verschwindender quadratischer Fehler: Es gilt

$$E(g(x_1, .., x_2))^2 \quad\overset{n\to\infty}{\longrightarrow}\quad 0.$$

(2) Konsistenz: Je mehr Daten verwendet werden, desto besser wird das Ergebnis.

$$\lim_{n\to\infty} g(x_1, .., x_n) \;=\; 0\ [P].$$

(3) Erwartungstreue: Es wird kein systematischer Fehler gemacht, im Mittel erhalten wir das gewünschte Ergebnis.

$$E(g(x_1, .., x_n) \;=\; a.$$

Seien (X_k) reelle, unabhängige und identisch verteilte Zufallsvariablen mit $EX_k = \mu$ und $\mathrm{Var}(X_k) = \sigma^2$.

Stichprobenmittel

Das Stichprobenmittel

$$\overline{X_n} \;=\; \frac{1}{n}\sum_{k=1}^{n} X_k$$

ist ein Schätzer für μ, denn es gilt:

(1) Verschwindender quadratischer Fehler: Nach der Formel von Bienaymé folgt

$$E(\overline{X_n} - \mu)^2 \;=\; \mathrm{Var}(\overline{X_n}) \;=\; \frac{1}{n}\mathrm{Var}(X_1) \;=\; \frac{\sigma^2}{n} \;\stackrel{n\to\infty}{\longrightarrow}\; 0.$$

(2) Konsistenz: Nach dem starken Gesetz großer Zahlen konvergiert $\overline{X_n}$ fast sicher gegen μ.

(3) Erwartungstreue: Es gilt

$$E\overline{X_n} \;=\; \frac{1}{n}\sum_{k=1}^{n} EX_k \;=\; \frac{1}{n}\sum_{k=1}^{n}\mu \;=\; \mu.$$

Stichprobenvarianz

Die Stichprobenvarianz

$$S_n^2 \;=\; \frac{1}{n-1}\sum_{k=1}^{n}(X_k - \overline{X_n})^2$$

ist ein Schätzer σ^2, alle drei Axiome lassen sich etwas komplizierter als beim Stichprobenmittel nachweisen.

Beispiel Monte Carlo Simulation

Sei $g : [0,1] \to \mathbb{R}$ mit

$$\int_0^1 |g(x)|\,\mathrm{d}x \;<\; \infty.$$

Gesucht wird eine Annäherung von

$$\int_0^1 g(x)\,\mathrm{d}x.$$

Sei dazu $(X_n)_{n\in\mathbb{N}}$ eine Folge von unabhängigen gleichverteilten Zufallsvariablen auf $[0,1]$. Dann gilt

$$E|g(X_1)| \;=\; \int_0^1 |g(x)|\,\mathrm{d}x \;<\; \infty$$

und nach dem starken Gesetzt der großen Zahlen folgt

$$\frac{1}{n}\sum_{k=1}^{n} g(X_k) \quad\longrightarrow\quad Eg(X_1) \;=\; \int_0^1 g(x)\,\mathrm{d}x$$

fast sicher.

2.4 Zentraler Grenzwertsatz

Sei $(X_n)_{n\in\mathbb{N}}$ eine Folge von unabhängigen und identisch verteilten Zufallsvariablen mit

$$EX_1 \;=\; 0 \qquad \text{und} \qquad \mathrm{Var}(X_1) \;=\; EX_1^2 \;=\; \sigma^2 \;<\; \infty.$$

Das Ziel ist nun die Approximation der Verteilung von

$$S_n \;=\; \sum_{k=1}^{n} X_k$$

durch eine geeignete Normierung. Bei dem starken Gesetz großer Zahlen hatten wir die Normierung $1/n$ verwendet, wodurch die Folge der Zufallsvariablen unter den geforderten Bedingungen fast sicher gegen 0 konvergiert ist. Bei dem zentralen Grenzwertsatz werden wir eine etwas schwächere Normierung wählen, wodurch die Verteilung von S_n gegen die Normalverteilung konvergieren wird.

Dazu zunächst einige Definitionen:

Definition 2.4.1

Seien F und G zwei beliebige Verteilungsfunktionen und seien X und Y zwei beliebige Zufallsvariablen mit den Verteilungsfunktionen F_X und F_Y.

Der ***Abstand*** der Verteilungsfunktionen F und G wird gegeben durch

$$d(F,G) \;:=\; \sup_{x\in\mathbb{R}} |F(x) - G(x)| \;\leq\; 1 \;<\; \infty.$$

Der ***Abstand*** der Zufallsvariablen X und Y wird gegeben durch

$$d(X,Y) \;:=\; d(F_X, F_Y).$$

Satz 2.4.2

Seien X und Y zwei Zufallsvariablen mit $EY^1 \le \eta$ und sei

$$\varphi(x) \;=\; \frac{1}{\sqrt{2\pi}} \mathrm{e}^{-\frac{1}{2}x^2}$$

die Standard Normalverteilung. Dann gilt

$$d(X+Y, \varphi) \;\le\; d(X, \varphi) + 2 \cdot \eta^{1/3}.$$

Beispiel 2.4.3

Seien $(X_n)_{n\in\mathbb{N}}$ und $(Y_n)_{n\in\mathbb{N}}$ zwei Folgen von Zufallsvariablen, für die

$$d(X_n, \varphi) \;\to\; 0 \qquad \text{und} \qquad E(X_n - Y_n)^2 \;\to\; 0$$

für $n \to \infty$ gilt. Dann folgt auch

$$d(Y_n, \varphi) \;\to\; 0,$$

denn es gilt

$$\begin{aligned} d(Y_n, \varphi) \;&=\; d(X_n + (Y_n - X_n), \varphi) \\ &\le\; d(X_n, \varphi) + 2 \cdot (E(X_n - Y_n)^2)^{1/3} \;\to\; 0. \end{aligned}$$

Satz 2.4.4 (Zentraler Grenzwertsatz)

Version 1

Sei $(X_n)_{n\in\mathbb{N}}$ eine Folge von unabhängigen und identisch verteilten Zufallsvariablen auf $(\Omega, \mathcal{K}, P)$ mit

$$EX_1 \;=\; 0 \qquad \text{und} \qquad \mathrm{Var}(X_1) \;=\; EX_1^2 \;=\; 1.$$

X_1 nehme nur endlich viele Werte an, also $\#(X_1(\Omega)) < \infty$. Sei weiter

$$\varphi(x) \;=\; \frac{1}{\sqrt{2\pi}} \mathrm{e}^{-\frac{1}{2}x^2}$$

die Standard Normalverteilung und

$$S_n \;:=\; \sum_{k=1}^{n} X_k.$$

Dann gilt

$$d\left(\frac{S_n}{\sqrt{n}}, \varphi\right) \;\overset{n\to\infty}{\longrightarrow}\; 0.$$

Version 2

Sei $(X_n)_{n\in\mathbb{N}}$ eine Folge von unabhängigen und identisch verteilten Zufallsvariablen auf $(\Omega, \mathcal{K}, P)$ mit Erwartungswert η und Varianz $\sigma^2 > 0$. Weiter sei

$$\varphi(x) = \frac{1}{\sqrt{2\pi}} \mathrm{e}^{-\frac{1}{2}x^2}$$

die Standard Normalverteilung und es sei

$$S_n := \sum_{k=1}^{n} X_k \qquad \text{mit} \qquad S_n^* = \frac{S_n - n\eta}{\sigma\sqrt{n}}.$$

Dann gilt für alle $a \leq b$

$$\lim_{n\to\infty} P(a \leq S_n^* \leq b) = \int_a^b \varphi(x)\,\mathrm{d}x.$$

Beispiel 2.4.5

Eine faire Münze wird 200 mal geworfen. Es soll die Wahrscheinlichkeit dafür berechnet werden, dass die Anzahl der Kopfwürfe zwischen 95 und 105 liegt.

Dazu sei $n = 200$ es seien $X_1, .., X_n$ Bernoulli verteilte unabhängigen Zufallsvariablen mit Parameter $\vartheta = 1/2$. Wir haben also

$$\eta = EX_k = \vartheta = \frac{1}{2} \qquad \text{und} \qquad \sigma^2 = \mathrm{Var}(X_k) = \vartheta(1-\vartheta) = \frac{1}{4}.$$

Weiter sei

$$S_n = \sum_{k=1}^{n} X_n$$

die Anzahl der Kopfwürfe bei n Versuchen. Nach dem zentralen Grenzwertsatz erhalten wir

$$\begin{aligned} P(95 \leq S_n \leq 105) &= P\left(\frac{95 - n\eta}{\sigma\sqrt{n}} \leq S_n^* \leq \frac{105 - n\eta}{\sigma\sqrt{n}}\right) \\ &= P\left(-0.71 \leq S_n^* \leq 0.71\right) \\ &\approx \int_{-0.71}^{0.71} \varphi(x)\,\mathrm{d}x = 0.52\,. \end{aligned}$$

Zu ungefähr 52 % erhalten wir bei 200 unabhängigen Münzwürfen zwischen 95 und 105 mal Kopf.

Beispiel 2.4.6

Es seien $X_1, .., X_n$ die Augenzahlen bei unabhängigen Würfeln eines fairen Würfels. Wir haben also den Erwartungswert $\eta = 3.50$ und die Varianz $\sigma^2 = 35/12 \approx 2.92$ für alle X_k.

Wir wollen die Wahrscheinlichkeit untersuchen, dass bei $n = 100$ Würfen die Augensumme zwischen 340 und 360 liegt. Nach dem zentralen Grenzwertsatz erhalten wir

$$\begin{aligned} P(340 \leq S_n \leq 360) &= P\left(\frac{340 - n\eta}{\sigma\sqrt{n}} \leq S_n^* \leq \frac{360 - n\eta}{\sigma\sqrt{n}}\right) \\ &= P\left(-0.59 \leq S_n^* \leq 0.58\right) \\ &\approx \int_{-0.59}^{0.59} \varphi(x)\,\mathrm{d}x = 0.44\,. \end{aligned}$$

Zu ungefähr 44 % liegt die Augensumme bei $n = 100$ Würfen also zwischen 340 und 360.

Für $n = 1000$ erhalten wir bereits eine Wahrscheinlichkeit von ungefähr 94 %, dass die Augensumme zwischen 3400 und 3600 liegt.

Beispiel 2.4.7

In diesem Beispiel soll umgekehrt die Anzahl der Versuche ermittelt werden, so dass das Ergebnis mit vorgegebenen Wahrscheinlichkeit innerhalb eines gegebenen Intervalls liegt:

Ein fairer Würfel wird n mal gewürfelt und es soll die Anzahl der Würfe ermittelt werden, so dass die mittlere Augenzahl zu 99 % zwischen 3.4 und 3.6 liegt.

Dazu seien $X_1, .., X_n$ Zufallsvariablen, die die Augenzahlen vom n ten Wurf beschreiben. Wir haben damit

$$\mu = EX_k = 3.5 \qquad \text{und} \qquad \sigma^2 = \mathrm{Var}(X_k) = \frac{35}{12} \approx 2.917\,.$$

Weiter sei $S_n = \sum_{k=1}^n X_k$ die Augensumme und S_n/n die mittlere Augenzahl. Durch

$$\int_{-2.57}^{2.57} \varphi(x)\,\mathrm{d}x \approx 0.99$$

[dieser Wert lässt sich nachschlagen oder mittels PC berechnen], erhalten wir nach dem zentralen Grenzwertsatz

$$\begin{aligned} 0.99 &\approx \int_{-2.57}^{2.57} \varphi(x)\,\mathrm{d}x \\ &\approx P\left(-2.57 \leq S_n^* \leq 2.57\right) \end{aligned}$$

$$
\begin{aligned}
&\approx\ P\left(-2.57 \le \frac{S_n - n\eta}{\sigma\sqrt{n}} \le 2.57\right)\\
&\approx\ P\left(-2.57 \le \frac{S_n - 3.5\,n}{\sqrt{2.917\,n}} \le 2.57\right)\\
&\approx\ P\left(-2.57 \cdot \sqrt{2.917n} + 3.5\,n \le S_n \le 2.57 \cdot \sqrt{2.917n} + 3.5\,n\right)\\
&\approx\ P\left(\frac{-2.57 \cdot \sqrt{2.917}}{\sqrt{n}} + 3.5 \le \frac{S_n}{n} \le \frac{2.57 \cdot \sqrt{2.917}}{\sqrt{n}} + 3.5\right).
\end{aligned}
$$

Damit ergibt sich das gesuchte n aus

$$
\begin{aligned}
n &= \min\left\{k \in \mathbb{N} \;\middle|\; \frac{-2.57 \cdot \sqrt{2.917}}{\sqrt{k}} + 3.5 \ \ge\ 3.4\right\}\\
&= \min\left\{k \in \mathbb{N} \;\middle|\; \frac{-2.57 \cdot \sqrt{2.917}}{\sqrt{k}} \ \ge\ -0.1\right\}\\
&= \min\left\{k \in \mathbb{N} \;\middle|\; \frac{2.57 \cdot \sqrt{2.917}}{\sqrt{k}} \ \le\ -0.1\right\},
\end{aligned}
$$

denn für dieses n gilt auch

$$
n = \min\left\{k \in \mathbb{N} \;\middle|\; \frac{2.57 \cdot \sqrt{2.917}}{\sqrt{k}} + 3.5 \ \le\ 3.6\right\}
$$

Wir erhalten $n = 1927$ und müssen also 1927 mal würfeln, damit die mittlere Augenzahl zu 99 % zwischen 3.4 und 3.6 liegt.

2.5 Aufgaben

Aufgabe 2.5.1

Es seien X und Y unabhängige auf $[0, 1]$ gleichverteilte Zufallsvariablen. Berechne die bedingte Wahrscheinlichkeit

$$P(Y \ge 1/2 \,|\, Y \ge X^2).$$

Lösung

Für die bedingte Wahrscheinlichkeit gilt

$$P(Y \ge 1/2 \,|\, Y \ge X^2) \ = \ \frac{P(Y \ge 1/2\,,\, Y \ge X^2)}{P(Y \ge X^2)}.$$

Die nun auftretenen Wahrscheinlichkeiten erhalten wir durch Berechnung der Flächen aus Abbildung 2.2.

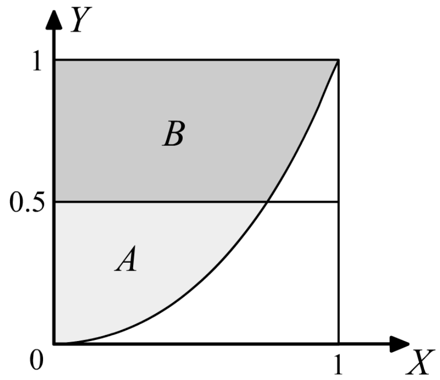


Abbildung 2.2: Gleichverteilte Zufallsvariablen X und Y.

Es gilt

$$\begin{aligned} A + B &= 1 - \int_0^1 x^2 \,\mathrm{d}x = 1 - \frac{1}{3} = \frac{2}{3} \approx 0.6666, \\ A &= \frac{1}{2} \cdot \frac{1}{\sqrt{2}} - \int_0^{\sqrt{2}} x^2 \,\mathrm{d}x = \frac{1}{3\sqrt{2}} \approx 0.2357, \\ B &= A + B - A = \frac{2}{3} - \frac{1}{3\sqrt{2}} \approx 0.4310\,. \end{aligned}$$

Damit folgt

$$\begin{aligned} P(Y \geq 1/2 \,|\, Y \geq X^2) &= \frac{P(Y \geq 1/2\,,\, Y \geq X^2)}{P(Y \geq X^2)} = \frac{B}{A+B} \\ &= \left(\frac{2}{3} - \frac{\sqrt{2}}{6}\right)\frac{3}{2} = 1 - \frac{1}{2\sqrt{2}} \approx 0.647\,. \end{aligned}$$

Aufgabe 2.5.2

Zeige, dass jede Verteilungsfunktion F an höchstens abzählbar vielen Stellung unstetig sein kann.

Lösung

Sei M die Menge aller Unstetigkeitsstellen von F. Nach Definition von Verteilungsfunktionen kann M keine Intervalle enthalten, denn sonst wäre F in diesem Intervall nicht rechtsseitig stetig. Die Menge M enthält also nur einzelne Punkte, von denen wir zeigen müssen, dass sie abzählbar sind.

Da M keine Intervalle enthält, gibt es zu jedem $a \in M$ ein geeignet kleines $t_a > 0$, so dass die Intervalle

$$[a - t_a, a + t_a] \subset \mathbb{R} \qquad \text{für alle } \ a \in M$$

paarweise disjunkt sind. Bilden wir nun die Abbildung

$$\varphi : M \to \mathbb{Q},$$

die jedem $a \in M$ eine beliebige rationale Zahl aus $[a - t_a, a + t_a]$ zuordnet. Da die $[a - t_a, a + t_a]$ paarweise disjunkt sind, ist die Abbildung φ injektiv. Da weiter $\mathbb{Q}$ abzählbar ist, muss auch M abzählbar sein. Somit ist jede Verteilungsfunktion F an höchstens abzählbar vielen Stellung unstetig.

Aufgabe 2.5.3

Es sei X eine Standard Normalverteilte Zufallsvariable. Berechne die Dichten von $\exp(X)$ und X^2.

Lösung Teil 1

Die Dichte der Standard Normalverteilung ist

$$\varphi(x) = \frac{1}{\sqrt{2\pi}} \exp\left(-\frac{1}{2}x^2\right).$$

Betrachten wir zunächst

$$g(x) = \exp(x) \qquad \text{mit} \qquad g^{-1}(x) = \log(x).$$

Es gilt $g(\mathbb{R}) =]0, \infty[$ und g ist streng monoton steigen, also bijektiv. Somit folgt nach der Transformationsformel für die Dichte von $\exp(X)$

$$\begin{aligned} f(y) &= \varphi(g^{-1}(y)) \cdot \left|\frac{\mathrm{d}}{\mathrm{d}x} g^{-1}(y)\right| \\ &= \frac{1}{\sqrt{2\pi}} \exp\left(-\frac{1}{2}(\log(y))^2\right) \cdot \left|\frac{1}{y}\right| \\ &= \frac{1}{\sqrt{2\pi} y} \exp\left(-\frac{1}{2}(\log(y))^2\right) \end{aligned}$$

für $y > 0$ und $f(y) = 0$ für $y \leq 0$.

Lösung Teil 2

Weiter betrachten wir

$$g(x) = x^2 \qquad \text{mit} \qquad g^{-1}(x) = \pm\sqrt{x}.$$

Es gilt $g(\mathbb{R}) = [0, \infty[$. Bei der Anwendung der Transformationformel zur Berechnung der Dichte von X^2 haben wir zu beachten, dass g nur auf den Intervallen $[0, \infty[$ und $]-\infty, 0]$ bijektiv ist. Als Umkehrfunktion von $g(x)$

erhalten wir $-\sqrt{x}$ bzw. $+\sqrt{x}$. Wir müssen die Transformation auf beide Umkehrfunktionen anwenden und erhalten die Dichte

$$\begin{aligned} f(y) &= \varphi(g^{-1}(y)) \cdot \left| \frac{\mathrm{d}}{\mathrm{d}x} g^{-1}(y) \right| \\ &= \frac{1}{\sqrt{2\pi}} \exp\left(-\frac{1}{2}y\right) \cdot \left| \frac{1}{2\sqrt{y}} \right| \;+\; \frac{1}{\sqrt{2\pi}} \exp\left(-\frac{1}{2}y\right) \cdot \left| -\frac{1}{2\sqrt{y}} \right| \\ &= \frac{1}{\sqrt{2\pi y}} \exp\left(-\frac{1}{2}y\right) \end{aligned}$$

für $y \geq 0$ und $f(y) = 0$ für $y < 0$.

Aufgabe 2.5.4

In der Ebene tangiert die Gerade L den Kreis K vom Durchmesser 1 im Punkt P. Auf der Kreislinie gegenüber von P liegt der Punkt Q. Eine weitere Gerade verbindet Q mit einem beweglichen Punkt R auf der Geraden L.

Der Winkel $\varphi = (P, Q, R)$ sei eine $-\pi/2, \pi/2$ gleichverteilte Zufallsvariable X. Die Länge Z der Verbindungsstrecke $\overline{PR}$ sei ebenfalls eine Zufallsvariable. Y beschreibe diese Zufallsvariable Z, falls der Winkel im Uhrzeigersinn durchlaufen wird.

Berechne die Dichte von Y.

Lösung

Zunächst verdeutlichen wir uns das Problem anhand von Abbildung 2.3.

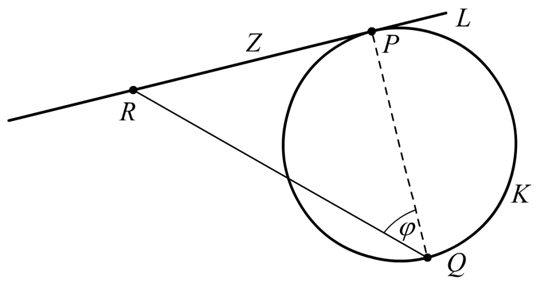


Abbildung 2.3: Bestimmung der Zufallsvariablen der Strecke $\overline{PR}$.

Die gleichverteilte Zufallsvariable X wird gegeben durch die Dichte

$$f_X(x) = \frac{1}{\pi}$$

für $-\pi/2 \leq x \leq \pi/2$ und $f_X(x) = 0$ sonst. Da der Durchmesser des Kreises gerade 1 ist, erhalten wir die Länge Z der Verbindungsstrecke $\overline{PR}$ durch

$$Z = \tan \varphi.$$

Damit haben wir einen direkten Zusammenhang zwischen den Zufallsvariablen X und Y gefunden: Um von der Dichte f_X auf die Dichte f_Y zu schließen, können wir die Transformationsformel mit der Transformation

$$g(x) \;=\; \tan x$$

anwenden. Weiter gilt

$$g^{-1}(x) \;=\; \arctan x \qquad \text{und} \qquad \frac{\mathrm{d}}{\mathrm{d}x} g^{-1}(x) \;=\; \frac{1}{1+x^2}.$$

Damit erhalten wir die gesuchte Dichte

$$f_Y(y) \;=\; f_X(g^{-1}(y)) \cdot \left| \frac{\mathrm{d}}{\mathrm{d}x} g^{-1}(y) \right| \;=\; \frac{1}{\pi} \cdot \left| \frac{1}{1+x^2} \right| \;=\; \frac{1}{\pi(1+x^2)}$$

für alle $y \in \mathbb{R}$, da $g^{-1}(y) \in [-\pi/2, \pi/2]$ für alle $y \in \mathbb{R}$.

Aufgabe 2.5.5

Seien X und Y unabhängige identisch verteilte Zufallsvariablen mit der Verteilungsfunktion F.

(1) Berechne die Verteilungsfunktionen von $V = \max\{X, Y\}$ und von $U = \min\{X, Y\}$.

(2) X und Y seien unabhängig und gleichverteilt auf $[0, 1]$. Berechne EV, EU sowie $\mathrm{Cov}(V, U)$.

Lösung Teil 1

Es sei F_V die Verteilungsfunktion von V. Da X und Y unabhängig sind, erhalten wir

$$\begin{aligned} F_V(t) \;&=\; P(t \le \max\{X, Y\}) \;=\; P(X \le t, Y \le t) \\ &=\; P(X \le t) \cdot P(Y \le t) \;=\; F(t) \cdot F(t) \;=\; F^2(t). \end{aligned}$$

Für die Verteilungsfunktion F_U von U betrachten wir die Gegenwahrscheinlichkeit und erhalten

$$\begin{aligned} F_U(t) \;&=\; P(t \le \min\{X, Y\}) \;=\; 1 - P(X > t, Y > t) \\ &=\; 1 - (P(X > t) \cdot P(Y > t)) \;=\; 1 - ((1 - F(t)) \cdot (1 - F(t))) \\ &=\; 1 - (1 - 2F(t) + F^2(t)) \;=\; 2F(t) - F^2(t). \end{aligned}$$

Lösung Teil 2

Für $t \in [0,1]$ ist die Verteilungsfunktion F der Gleichverteilung auf $[0,1]$ gerade $F(t) = t$. Somit erhalten wir für $t \in [0,1]$

$$F_V(t) \; = \; t^2 \qquad \text{und} \qquad F_U(t) \; = \; 2t - t^2$$

und dadruch die Dichten

$$f_V(t) \; = \; 2t \qquad \text{und} \qquad f_U(t) \; = \; 2 - 2t.$$

Damit lassen sich die Erwartungswerte berechnen:

$$\begin{aligned} EV &= \int_0^1 t f_V(t)\,\mathrm{d}t \; = \; 2\int_0^1 t^2\,\mathrm{d}t \; = \; \frac{2}{3}, \\ EU &= \int_0^1 t f_U(t)\,\mathrm{d}t \; = \; 2\int_0^1 t - t^2\,\mathrm{d}t \; = \; 1 - \frac{2}{3} \; = \; \frac{1}{3}. \end{aligned}$$

Die Kovarianz berechnet sich nun aus

$$\begin{aligned} \mathrm{Cov}(V,U) &= E(VU) - EUEV \; = \; E(\max\{X,Y\} \cdot \min\{X,Y\}) - \frac{2}{9} \\ &= E(XY) - \frac{2}{9} \; = \; EX \cdot EY - \frac{2}{9} \; = \; \frac{1}{2} \cdot \frac{1}{2} - \frac{2}{9} \; = \; \frac{1}{36}. \end{aligned}$$

Aufgabe 2.5.6

Sei (X_n) eine Folge von unabhängigen, identisch verteilten Zufallsvariablen mit Verteilungsfunktion F. Die relative Häufigkeit der $X_n(\omega)$ mit Werte $\leq t$ sei definiert durch

$$F_n(t,\omega) \; = \; \frac{1}{n}\sum_{k=1}^{n} 1_{]-\infty,t]}(X_k(\omega)).$$

Zeige, dass für alle $t \in \mathbb{R}$ die Folge $F_n(t,\omega)$ fast sicher gegen $F(t)$ konvergiert.

Lösung

Wir wollen das Problem auf das starke Gesetz großer Zahlen zurückführen.

Sei dazu $Y_k = 1_{]-\infty,t]}(X_k(\omega))$ für ein beliebiges $t \in \mathbb{R}$. Damit sind die Y_k unabhängig und identisch Bernoulli verteilt. Es gilt

$$P(Y_k = 1) \; = \; P(X_k \leq t) \; = \; F(t),$$

also haben die Bernoulli verteilten Zufallsvariablen Y_k den Parameter $F(t)$ und damit gilt auch $EY_k = F(t)$ für alle $k \in \mathbb{N}$. Wir können nun das starke Gesetz großer Zahlen anwenden und erhalten, dass

$$F_n(t,\omega) \; = \; \frac{1}{n}\sum_{k=1}^{n} Y_k \; = \; \frac{1}{n}\sum_{k=1}^{n} 1_{]-\infty,t]}(X_k(\omega))$$

für $n \to \infty$ fast sicher gegen $EY_k = F(t)$ konvergiert.

Aufgabe 2.5.7

Das Intervall $[0,1]$ wird in n disjunkte Teilintervalle der Längen $p_1,..,p_n$ unterteilt. Die Entropie dieser Partition definieren wir durch

$$h \;=\; -\sum_{k=0}^{n} p_i \log p_i.$$

Es sei (X_n) eine Folge von unabhängigen auf $[0,1]$ gleichverteilten Zufallsvariablen und es sei $Z_m(k)$ die Anzahl der $X_1,..,X_m$, die im $k$ten Teilintervall liegen. Weiter sei

$$R_m \;=\; \prod_{k=1}^{n} p_i^{Z_m(k)}.$$

Zeige, dass dann

$$\frac{1}{m}\log(R_m)$$

für $m \to \infty$ fast sicher gegen $-h$ konvergiert.

Lösung

Sei $Y_j = \log p_i$, falls X_j im $i$ten Teilintervall liegt. Damit sind die Y_j unabhängig und identisch verteilte Zufallsvariablen und es gilt

$$EY_j \;=\; \sum_{k=1}^{n} p_k \log p_k \;=\; -\,h.$$

Weiter haben wir

$$\log(R_m) \;=\; \log\left(\prod_{k=1}^{n} p_i^{Z_m(k)}\right) \;=\; \sum_{k=1}^{n} Z_m(k)\log(p_k) \;=\; Y_1 + \ldots + Y_m.$$

Nach dem starken Gesetz großer Zahlen konvergiert somit

$$\frac{1}{m}\log(R_m) \;=\; \frac{1}{m}\sum_{k=1}^{m} Y_k$$

für $m \to \infty$ fast sicher gegen $EY_j = -h$.

Aufgabe 2.5.8

Mit einem fairen Würfel wird 600 mal gewürfelt. Bestimme näherungsweise mit dem zentralen Grenzwertsatz die Wahrscheinlichkeit dafür, dass zwischen 90 und 110 mal eine 6 gewürfelt wird.

Lösung

Sei $n = 600$ und sei X_k das Ergebnis des k ten Wurfes für $k = 1, .., n$. Weiter sei $Y_k = 1$, wenn $X_k = 6$, und $Y_k = 0$ sonst. Dann sind die Zufallsvariablen $Y_1, .., Y_n$ unabhängig und idendtisch verteilt. Weiter sei

$$S_n \;=\; \sum_{k=1}^{n} Y_k$$

die Anzahl der Würfe, in denen eine 6 gewürfelt wird.

Gesucht ist nun die Wahrscheinlichkeit von $P(90 \le S_n \le 110)$. Es gilt

$$\mu \;=\; EY_k \;=\; \frac{1}{6} \qquad \text{und} \qquad \sigma^2 \;=\; \operatorname{Var}(Y_k) \;=\; \frac{1}{6} - \frac{1}{6^2} \;=\; \frac{5}{36} \;>\; 0$$

für $k = 0, .., n$, da die Zufallsvariablen Y_k Bernoulli verteilt sind mit Parameter 1/6. Mit

$$S_n^* \;=\; \frac{S_n - n\eta}{\sigma\sqrt{n}}$$

kann nun der zentrale Grenzwertsatz angewendet werden:

$$\begin{aligned} P(90 \le S_n \le 110) \;&=\; P\left(\frac{90 - n\eta}{\sigma\sqrt{n}} \;\le\; S_n^* \;\le\; \frac{110 - n\eta}{\sigma\sqrt{n}}\right) \\ &=\; P\left(-1.095 \;\le\; S_n^* \;\le\; 1.095\right) \\ &\approx\; \int_{-1.095}^{1.095} \varphi(x)\,\mathrm{d}x \;=\; 0.726\,. \end{aligned}$$

Es wird als zu ungefähr 72.6% zwischen 90 und 110 mal eine 6 gewürfelt.

Aufgabe 2.5.9

Berechne mit dem zentralen Grenzwertsatz die Wahrscheinlichkeit dafür, wie oft eine faire Münze ungefähr geworfen werden muss, damit mit einer Wahrscheinlichkeit von 95% bei 49% bis 51% der Würfe Kopf erscheint.

Lösung

Seien X_k unabhängige Bernoulli verteilte Zufallsvariablen mit Parameter $\vartheta = 1/2$, es gilt also

$$\eta \;=\; EX_k \;=\; \frac{1}{2} \qquad \text{und} \qquad \sigma^2 \;=\; \operatorname{Var}(X_k) \;=\; \frac{1}{2} - \frac{1}{4} \;=\; \frac{1}{4} \;>\; 0.$$

Sei weiter

$$S_n \;=\; \sum_{k=1}^{n} X_k$$

die Anzahl der Erfolge bei n Experimenten. Wir suchen nun ein möglichst kleines $n \in \mathbb{N}$ mit

$$P\left(n\frac{49}{100} \leq S_n \leq n\frac{51}{100}\right) \approx 0.95\,.$$

Es gilt

$$\int_{-1.96}^{1.96} \varphi(x)\,\mathrm{d}x \approx 0.95$$

und somit erhalten wir nach dem zentralen Grenzwertsatz

$$\begin{aligned}
0.95 &\approx \int_{-1.96}^{1.96} \varphi(x)\,\mathrm{d}x \\
&\approx P\left(-1.96 \leq S_n^* \leq 1.96\right) \\
&\approx P\left(-1.96 \leq \frac{S_n - n\eta}{\sigma\sqrt{n}} \leq 1.96\right) \\
&\approx P\left(-1.96 \leq \frac{S_n - 0.5\,n}{\sqrt{0.25\,n}} \leq 1.96\right) \\
&\approx P\left(-1.96 \cdot \sqrt{0.25n} + 0.5\,n \leq S_n \leq 1.96 \cdot \sqrt{0.25n} + 0.5\,n\right) \\
&\approx P\left(\frac{-1.96 \cdot \sqrt{0.25}}{\sqrt{n}} + 0.5 \leq \frac{S_n}{n} \leq \frac{1.96 \cdot \sqrt{0.25}}{\sqrt{n}} + 0.5\right).
\end{aligned}$$

Damit ergibt sich das gesuchte n aus

$$\begin{aligned}
n &= \min\left\{k \in \mathbb{N} \;\middle|\; \frac{-1.96 \cdot \sqrt{0.25}}{\sqrt{k}} + 0.5 \geq 0.49\right\} \\
&= \min\left\{k \in \mathbb{N} \;\middle|\; \frac{-1.96 \cdot \sqrt{0.25}}{\sqrt{k}} \geq -0.01\right\} \\
&= \min\left\{k \in \mathbb{N} \;\middle|\; \frac{1.96 \cdot \sqrt{0.25}}{\sqrt{k}} \leq 0.01\right\},
\end{aligned}$$

denn für dieses n gilt auch

$$n = \min\left\{k \in \mathbb{N} \;\middle|\; \frac{1.96 \cdot \sqrt{0.25}}{\sqrt{k}} + 0.5 \leq 0.51\right\}$$

Wir erhalten $n = 9604$ und müssen also 9604 mal eine Münze werfen, um mit einer Wahrscheinlichkeit von 95% bei 49% bis 51% der Würfe Kopf zu erhalten.

3 Markov Ketten

Abschließend werden wir in diesem Kapitel kurz auf Markov Ketten eingehen.

Dazu betrachten wir einen Wahrscheinlichkeitsraum $(\Omega, \mathcal{K}, P)$ und eine abzählbare Indexmenge I betrachten, so dass $J = \mathcal{P}(I)$ eine σ-Algebra ist. Weiter seien

$$X_n : (\Omega, \mathcal{K}) \to (I, J)$$

messbare Zufallsvariablen. Die Folge $X = (X_n)$ soll nun den zeitlichen Verlauf zu diskreten Zeitpunkten beschreiben.

Wir verwenden dazu nur Folgen (X_n) mit speziellen Eigenschaften, die die Analyse vereinfachen, aber trotzdem für viele Modelle in den Anwendungen geeignet sind.

3.1 Die Markov Eigenschaft

Definition 3.1.1

X heißt eine ***Markov Kette*** mit ***Zustandsraum*** I, wenn für alle $n \in \mathbb{N}$ und für alle $i_0, .., i_n \in I$ mit

$$P(X_0 = i_0, \ldots, X_n = i_n) \; > \; 0$$

die ***Markov Eigenschaft***

$$P(X_{n+1} = i_{n+1} \,|\, X_0 = i_0, \ldots, X_n = i_n) \;\; = \;\; P(X_{n+1} = i_{n+1} \,|\, X_n = i_n)$$

gilt.

Bemerkung

Die X_n lassen sich als Zustände zu den Zeitpunkten n interpretieren. Bei einer Markov Kette hängt der Zustand zum Zeitpunkt $n + 1$ als nur vom Zustand zum Zeitpunkt n ab und nicht von den Zuständen in der Vergangenheit.

Im Folgenden werden wir stets die Bedingung

$$P(X_0 = i_0, \ldots, X_n = i_n) \; > \; 0$$

fordern, auch wenn wir es nicht immer explizit angeben werden.

Lemma 3.1.2

Sei $(\Omega, \mathcal{K}, P)$ ein Wahrscheinlichkeitsraum, seien $C_k \in \mathcal{K}$ paarweise verschieden und sei

$$C \; = \; \bigcup_{k=0}^{\infty} C_k.$$

Weiter seien $A, B \in \mathcal{K}$ mit

$$P(A \,|\, B \cap C_k) \; = \; p$$

für alle $k \in \mathbb{N} \cup \{0\}$. Dann ist auch

$$P(A \,|\, B \cap C) \; = \; p.$$

Proposition 3.1.3

X ist genau dann eine Markov Kette, wenn für alle $n \in \mathbb{N}$ und alle $i_0, .., i_n \in I$ die Wahrscheinlichkeit

$$P(X_{n+1} = i_{n+1} \,|\, X_0 = i_0, \ldots, X_n = i_n)$$

unabhängig von $i_0, .., i_{n-1}$ ist.

3.2 Beisiele für Markov Ketten

Beispiel Irrfahrten

Sei $(Y_n)_{k \geq 0}$ eine Folge von unabhängigen Zufallsvariablen mit Werten in $I = \mathbb{Z}$, sei

$$X_n \; = \; \sum_{k=0}^{n} Y_n$$

und sei $X = (X_n)_{n \geq 0}$ mit $X_0 = Y_0$.

Dann ist X eine Markov Kette, dann da die Y_n unabhängig sind, gilt

$$\begin{aligned} & P(X_{n+1} = i_{n+1} \,|\, X_0 = i_0, \ldots, X_n = i_n) \\ = \; & \frac{P(X_0 = i_0, \ldots, X_n = i_n, X_{n+1} = i_{n+1})}{P(X_0 = i_0, \ldots, X_n = i_n)} \end{aligned}$$

$$
\begin{aligned}
&= \frac{P(Y_0 = i_0, Y_1 = i_1 - i_0, \ldots, Y_{n+1} = i_{n+1} - i_n)}{P(Y_0 = i_0, Y_1 = i_1 - i_0, \ldots, Y_n = i_n - i_{n-1})} \\
&= \frac{P(Y_0 = i_0) \cdot \prod_{k=1}^{n+1} P(Y_k = i_k - i_{k+1})}{P(Y_0 = i_0) \cdot \prod_{k=1}^{n} P(Y_k = i_k - i_{k+1})} \\
&= P(Y_{n+1} = i_{n+1} - i_n).
\end{aligned}
$$

Demnach sind wir nicht von den Zuständen $i_0, .., i_{n-1}$ abhängig und somit ist X eine Markov Kette.

Anwendung als Modell für Aktienkurse

Wir nehmen an, dass sich der Preis zu den Zeitpunkten $0, 1, 2, .., T$ stets verändert. Mit einer Wahrscheinlichkeit von p springt der Preis um eine Einheit nach oben und mit einer Wahrscheinlichkeit von $1-p$ um eine Einheit nach unten.

Dazu seien die X_k Bernoulli verteilt mit Parameter p. Es sei

$$
S_n = \sum_{k=1}^{n} X_k
$$

der Preis zum Zeitpunkt n. Dann haben wir im Beispiel gesehen, dass $S = (S_n)$ eine Markov Kette ist, der zukünftige Preis hängt also nur von dem gegenwärtigen Preis ab, nicht von Preisen aus der Vergangenheit.

Dieses Modell wird ***Binomial-Baum-Methode*** genannt und wird in der Anwendung zur numerischen Berechnung für komplizierte Modelle verwendet.

Beispiel Warteschlange

In diesem Beispiel soll die Warteschlange an einem Skilift simuliert werden.

Es seien $n = 0, 1, 2, ..$ Zeitpunkte, an denen der Lift mit je einer Person abfährt. Mit der Folge (Y_n) von unabhängigen Zufallsvariablen beschreiben wir die Anzahl der neu ankommenden Fahrgäste im Intervall $[n, n+1[$. Mit der Zufallsvariablen X_n beschreiben wir die Länge der Warteschlange unmittelbar vor der Abfahrt zur Zeit n. Es gilt also

$$
X_n = \max\{0,\ X_{n-1} - 1\} + Y_n,
$$

wobei wir $X_0 = i_0$ annehmen. Da die Folge (Y_n) aus unabhängigen Zufallsvariablen besteht, ist sind damit auch $Y_{n+1}, X_n, .., X_0$ unabhängig. Wir unterscheiden nun zwei Fälle:

Zunächst gelte $i_n \geq 1$. Dann haben wir

$$\begin{aligned} & P(X_{n+1} = i_{n+1}, X_n = i_n, \ldots, X_0 = i_0) \\ = \;& P(Y_{n+1} = i_{n+1} - i_n + 1, X_n = i_n, \ldots, X_0 = i_0) \\ = \;& P(Y_{n+1} = i_{n+1} - i_n + 1) \cdot P(X_n = i_n, \ldots, X_0 = i_0) \end{aligned}$$

und durch Division folgt

$$P(X_{n+1} = i_{n+1} \,|\, X_n = i_n, \ldots, X_0 = i_0) \;=\; P(Y_{n+1} = i_{n+1} - i_n + 1).$$

Damit ist die Wahrscheinlichkeit unabhängig von $i_0, .., i_{n-1}$, also ist $X = (X_n)$ eine Markov Kette.

Der zweite Fall $i_n = 0$ lässt sich sehr ähnlich zeigen.

Satz 3.2.1

Sei X eine Markov Kette. Dann gilt:

(1) Für alle $i_0, .., i_n \in I$ folgt

$$\begin{aligned} & P(X_0 = i_0, \ldots, X_n = i_n) \\ = \;& P(X_0 = i_0) \cdot P(X_1 = i_1 \,|\, X_0 = i_0) \cdot \ldots \cdot P(X_n = i_n \,|\, X_{n-1} = i_{n-1}). \end{aligned}$$

(2) Für $0 < n < N$ und $i_n \in I$ sowie $V \subset I^n$ und $F \subset I^{N-n}$ folgt

$$\begin{aligned} & P((x_{n+1}, .., X_n) \in F \,|\, X_n = i_n,\, (X_0, .., X_{n+1}) \in V) \\ = \;& P((x_{n+1}, .., X_n) \in F \,|\, X_n = i_n). \end{aligned}$$

(3) Für $k < m < n$ und $h, j \in I$ folgt die ***Chapman Kolmogrov Gleichung***

$$P(X_n = j \,|\, X_k = h) \;=\; \sum_{i \in I} P(X_n = j \,|\, X_n = i) \cdot P(X_m = i \,|\, X_k = h).$$

Definition 3.2.2

Eine Markov Kette heißt ***homogen*** oder Kette mit ***stationären Übergangswahrscheinlichkeiten***, wenn für alle $i, j \in I$

$$P(X_{n+1} = j \,|\, X_n = i) \;=:\; p_{ij}$$

unabhängig von n ist.

Die so erzeugte Matrix $P = (p_{ij})$ heißt eine ***stochastische Matrix***, denn es gilt

$$p_{ij} \;\geq\; 0 \qquad \text{und} \qquad \sum_{j \in I} p_{ij} \;=\; 1 \qquad \text{für} \quad i, j \in I.$$

Bespiel 3.2.3

Sei (Y_n) eine Folge von unabhängigen Bernoulli verteilten Zufallsvariablen mit Parameter 1/2. Sei weiter

$$X_n \;=\; 2Y_n + Y_{n+1}.$$

Dazu soll nun die Übergangsmatrix für X berechnet werden.

Es gilt $Y_k \in \{0,1\}$, somit erhalten wir den Zustandraum

$$X_k \;\in\; \{0,1,2,3\}.$$

Weiter folgt aus $X_n \in \{0,2\}$ gerade $Y_{n+1} = 0$ und aus $X_n \in \{1,3\}$ natürlich $Y_{n+1} = 1$. Damit erhalten wir aus

$$p_{ij} \;=\; P(X_{n+1} = j \,|\, X_n = i) \qquad \text{für} \quad i,j = 0,..,3$$

die gesuchte 4×4 Übergangsmatrix:

$$P \;=\; \begin{pmatrix} 1/2 & 1/2 & 0 & 0 \\ 0 & 0 & 1/2 & 1/2 \\ 1/2 & 1/2 & 0 & 0 \\ 0 & 0 & 1/2 & 1/2 \end{pmatrix}.$$

4 Anhang

4.1 Übersicht der diskreten Verteilungen

Verteilung 4.1.1 (Laplace Verteilung)

Sei $n \in \mathbb{N}$. Dann wird die Laplace Verteilung auf $\Omega = \{1, .., n\}$ durch die Zähldichte

$$f(i) \;=\; \frac{1}{n} \qquad \text{für} \qquad i \in \Omega$$

gegeben.

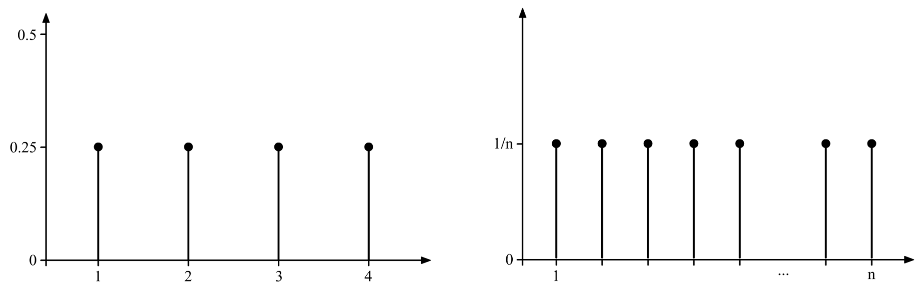


Abbildung 4.1: Laplace Verteilung mit $n = 4$ [links] und $n \in \mathbb{N}$ [rechts].

Für eine Laplace verteilte Zufallsvariable X gilt

$$EX \;=\; \frac{n+1}{2} \qquad \text{und} \qquad \operatorname{Var}(X) \;=\; \frac{n^2-1}{12}.$$

Verteilung 4.1.2 (Bernoulli Verteilung)

Sei $\vartheta \in [0, 1]$. Dann wird die Bernoulli Verteilung auf $\Omega = \{0, 1\}$ durch die Zähldichte

$$f(0) \;=\; 1 - \vartheta \qquad \text{und} \qquad f(1) \;=\; \vartheta$$

gegeben. Schreibweise: $\mathfrak{L}(1, \vartheta)$.

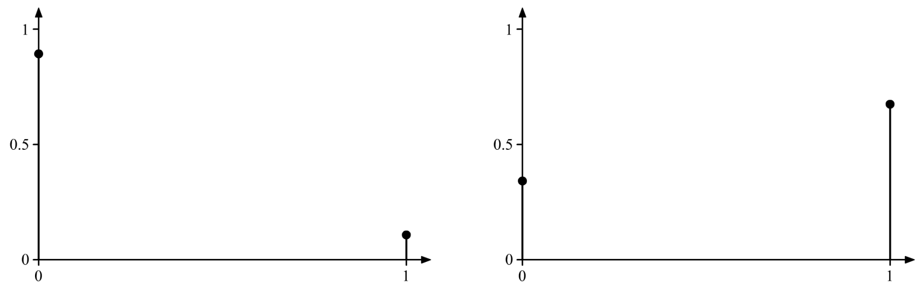


Abbildung 4.2: Bernoulli Verteilung mit $\vartheta = \frac{1}{10}$ [links] und $\vartheta = \frac{2}{3}$ [rechts].

Für eine Bernoulli verteilte Zufallsvariable X gilt

$$EX \ = \ \vartheta \qquad \text{und} \qquad \mathrm{Var}(X) \ = \ \vartheta - \vartheta^2.$$

Verteilung 4.1.3 (Binomialverteilung)

Sei $\vartheta \in [0,1]$. Dann wird die Binomialverteilung auf $\Omega = \{0, 1, .., n\}$ durch die Zähldichte

$$f(i) \ = \ \binom{n}{i} \vartheta^i (1-\vartheta)^{n-i} \qquad \text{für} \qquad i \in \Omega$$

gegeben. Schreibweise: $\mathfrak{L}(n, \vartheta)$.

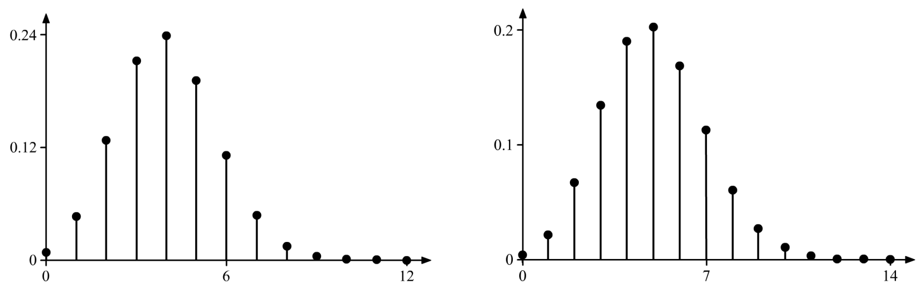


Abbildung 4.3: Binomialverteilung mit $n = 12$ und $\vartheta = \frac{1}{3}$ [links] sowie mit $n = 20$ und $\vartheta = \frac{1}{4}$ [rechts].

Für eine binomialverteilte Zufallsvariable X gilt

$$EX \ = \ n\vartheta \qquad \text{und} \qquad \mathrm{Var}(X) \ = \ n(\vartheta - \vartheta^2).$$

Verteilung 4.1.4 (Poisson Verteilung)

Sei $\lambda > 0$. Dann wird die Poisson Verteilung auf $\Omega = \mathbb{N} \cup \{0\}$ durch die Zähldichte

$$f(n) \;=\; \frac{\lambda^n}{n!}e^{-\lambda} \qquad \text{für} \qquad n \in \Omega$$

gegeben. Schreibweise: $\mathcal{P}(\lambda)$.

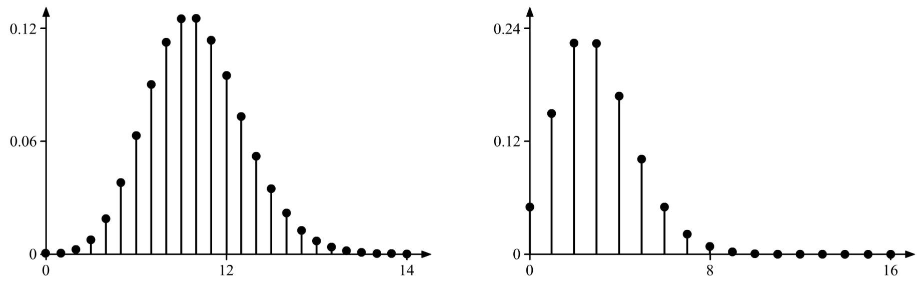


Abbildung 4.4: Poisson Verteilung mit $\lambda = 10$ [links] und $\lambda = 3$ [rechts].

Für eine Poisson verteilte Zufallsvariable X gilt

$$EX \;=\; \lambda \qquad \text{und} \qquad \mathrm{Var}(X) \;=\; \lambda.$$

Verteilung 4.1.5 (Hypergeometrische Verteilung)

Für geeignete Parameter n, s und k wird durch

$$f(r) \;=\; \frac{\binom{s}{r} \cdot \binom{n-s}{k-r}}{\binom{n}{k}} \qquad \text{für} \qquad \max\{0, k+s-n\} \;\leq\; r \;\leq\; \min\{k, s\}$$

eine Zähldichte gegeben. Schreibweise: $H(n, s, k)$.

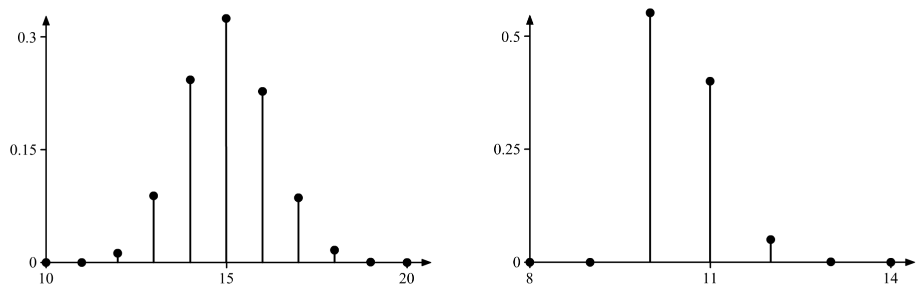


Abbildung 4.5: Hypergeometrische Verteilung mit $(n, s, k) = (32, 24, 20)$ [links] sowie mit $(n, s, k) = (16, 12, 14)$ [rechts].

Für eine hypergeometrisch verteilte Zufallsvariable X gilt

$$EX \;=\; k\frac{s}{n} \qquad \text{und} \qquad \mathrm{Var}(X) \;=\; k\frac{s}{n}\left(\frac{n-s}{n}\right)\frac{n-k}{n-1}.$$

Verteilung 4.1.6 (Geometrische Verteilung)

Sei $\vartheta \in [0,1]$. Dann wird die geometrische Verteilung auf $\Omega = \mathbb{N}$ durch die Zähldichte

$$f(n) \;=\; (1-\vartheta)^{n-1}\cdot\vartheta \qquad \text{für} \qquad n \in \Omega$$

gegeben. Schreibweise: $\mathcal{G}(\vartheta)$.

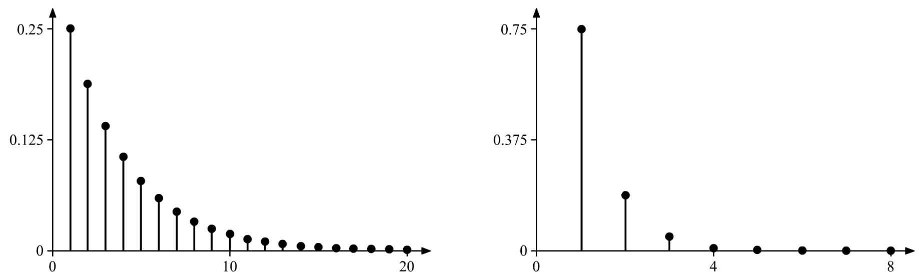


Abbildung 4.6: Geometrische Verteilung mit $\vartheta = \frac{1}{4}$ [links] und $\vartheta = \frac{3}{4}$ [rechts].

Für eine geometrisch verteilte Zufallsvariable X gilt

$$EX \;=\; \frac{1}{\vartheta} \qquad \text{und} \qquad \mathrm{Var}(X) \;=\; \frac{1-\vartheta}{\vartheta^2}.$$

4.2 Übersicht der stetigen Verteilungen

Verteilung 4.2.1 (Gleichverteilung)

Das Wahrscheinlichkeitsmaß $P = R(a,b)$ zu der Dichte

$$f(x) \;=\; \begin{cases} \frac{1}{b-a} & \text{für} \quad a \le x \le b \\ 0 & \text{sonst} \end{cases}$$

heißt Gleichverteilung auf dem Intervall $[a,b]$.

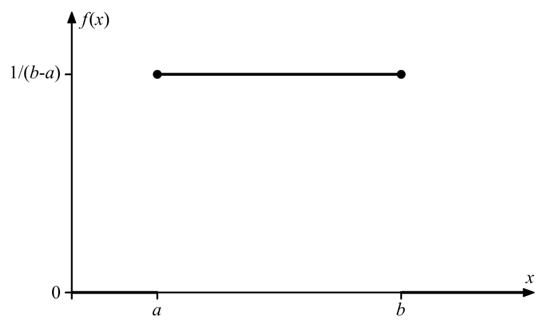


Abbildung 4.7: Gleichverteilung auf $[a,b]$.

Für eine gleichverteilte Zufallsvariable X gilt

$$EX \;=\; \frac{a+b}{2} \qquad \text{und} \qquad \mathrm{Var}(X) \;=\; \frac{(b-a)^2}{12}.$$

Verteilung 4.2.2 (Exponentialverteilung)

Sei $\alpha > 0$ eine reelle Zahl. Dann heißt das Wahrscheinlichkeitsmaß $P = E(\alpha)$ zu der Dichte

$$f(x) \;=\; \alpha \mathrm{e}^{-\alpha x} \qquad \text{für} \;\; x \geq 0$$

und $f(x) = 0$ für $x < 0$ Exponentialverteilung mit Parameter α.

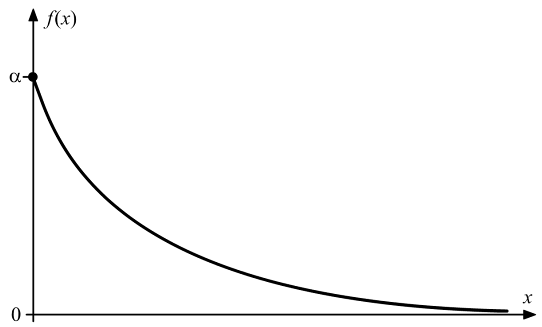


Abbildung 4.8: Exponentialverteilung mit Parameter α.

Für eine exponentiell verteilte Zufallsvariable X gilt

$$EX \;=\; \frac{1}{\alpha} \qquad \text{und} \qquad \mathrm{Var}(X) \;=\; \frac{1}{\alpha^2}.$$

Verteilung 4.2.3 (Normalverteilung)

Sei $\mu \in \mathbb{R}$ und $\sigma > 0$. Dann heißt das Wahrscheinlichkeitsmaß $P = N(\mu, \sigma^2)$ zu der Dichte

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

Normalverteilung mit Mittelwert μ und Varianz σ^2. Für $\mu = 0$ und $\sigma = 1$ ist

$$\varphi(x) \;=\; \frac{1}{\sqrt{2\pi}}\mathrm{e}^{-\frac{1}{2}x^2}$$

die Standard Normalverteilung $N = (0, 1)$.

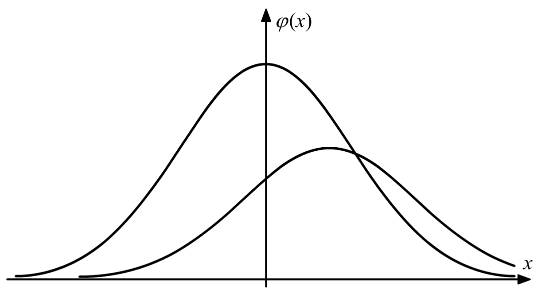


Abbildung 4.9: Normalverteilung mit Parameter μ und σ^2.

Für eine normalverteilte Zufallsvariable X gilt

$$EX \; = \; \mu \qquad \text{und} \qquad \mathrm{Var}(X) \; = \; \sigma^2.$$

Literaturverzeichnis

[1] Dehling, H. ; Haupt, B.: *Einführung in die Wahrscheinlichkeitstheorie und Statistik*. 2. Auflage. Springer Verlag Berlin Heidelberg, 2004. – ISBN 3–540–20380–X

[2] Koch, S.: *Übungen zur Vorlesung Einführung in die Stochastik*. – Übungszettel zur Vorlesung im Wintersemester 2003/2004 von Prof. S. Koch, Universität Göttingen.

[3] Scholz, D.: *Einführung in die Stochastik*. – Vorlesungsmitschrift im Wintersemester 2005/2006 zur Vorlesung von Prof. J. Woerner, Universität Göttingen.

[4] Scholz, D.: *Analytische Geometrie und lineare Algebra I und II*. Cuvillier Verlag Göttingen, 2004. – ISBN 3–86537–270–8

[5] Scholz, D.: *Differential- und Integralrechung I und II*. Cuvillier Verlag Göttingen, 2004. – ISBN 3–86537–271–6

[6] Woerner, J. ; Holzmann, H.: *Übungen zur Stochastik 1*. – Übungszettel zur Vorlesung im Wintersemester 2005/2006 von Prof. J. Woerner, Universität Göttingen.

Stichwortverzeichnis

G

H

I

K

L

M

N

P

Q

R