Lehrstuhl IV
Stochastik & Analysis

Fachschaft Mathematik
Uni Dortmund

Fachschaft WiMa

Stochastik I

Wahrscheinlichkeitsrechnung

Skriptum nach einer Vorlesung von Hans-Peter Scheffler

Letzte Änderung: 28. November 2002

Gesetzt mit LaTeX und LyX

Vorwort:

Dieses Script wurde in Zusammenarbeit der Fachschaften Mathematik & WirtschaftsMathematik mit dem Lehrstuhl IV erarbeitet. Es basiert auf der Vorlesung Stochastik I, gelesen von H-Doz. Dr. H.-P. Scheffler in den Sommersemestern 1998, 2000, 2001 und den Zusatzübungen für Lehramt Sek. II (Sommersemester 2000) gehalten von Dipl. Math. Sonja Menges. Zu den jeweiligen Kapiteln sind die Aufgaben der Übungszettel (Sommersemester 2000) aufgeteilt worden. Die Lösungen der Aufgaben werden nicht ins Netz gestellt, um den zukünftigen Übungsbetrieb Stochastik I nicht überflüssig zu machen. Die Zusatzveranstaltung für das Lehramt - Sek. II bildet den Anhang A. Im Anhang B befinden sich Kopien der in den drei Semestern gestellten Klausuren, ihre Lösungen sind als Kopiervorlage in der Fachschaft erhältlich. In Verweisen werden die Nummern der Sätze, Definitionen, ... in runden Klammern gegeben, z.B. (1.10) oder (a).
Ich habe versucht, alles richtig wiederzugeben, es ist jedoch „wahrscheinlich“, daß ich Fehler gemacht habe. Deshalb wendet euch bitte mit Fehlermeldungen, Anregungen zuerst an mich:

stk@fsmath.mathematik.uni-dortmund.de

Die Verwendung des „ß” ist in voller Absicht geschehen, also kein Fehler. Fehlermeldungen bitte so detailliert wie möglich.

Bei den oben genannten Mitarbeitern des Lehrstuhls IV wollen wir uns im Namen der Fachschaft recht herzlich für ihre Unterstützung bedanken. Ferner gilt unser Dank Thorsten Camps für seine tatkräftige Mithilfe.

Der Setzer

Hans-Peter Scheffler

Sonja Menges

Inhaltsverzeichnis

0. Kapitel: Motivation

Aufgabe der Stochastik ist es mathematische Modelle zur Beschreibung „zufälliger Phänomene" zu entwickeln und damit Vorhersagen über deren Ausgang zu machen. Diese Phänomene sind außermathematisch, und ebenso ist die Frage, ob unsere Modelle diese Phänomene adäquat beschreiben nicht mit mathematischen Methoden zu entscheiden, sondern durch Experimente. Die mathematischen Modelle verwenden im wesentlichen die (naive) Mengenlehre und (darauf aufbauend) die Maßtheorie. Im Gegensatz zur Analysis und linearen Algebra sind also die mathematischen Objekte der Stochastik Mengen sowie Maße (Wahrscheinlichkeitsmaße), d.h. Abbildungen von Mengensystemen nach $[0,1]$ im Vergleich zu allgemen Maßen: $M \to \mathbb{R}_0^+$. Dies macht eine Schwierigkeit der Stochastik gegenüber der Analysis deutlich. Die andere Schwierigkeit ist das Modellbildungsproblem: Zu einem gegebenen „Phänomen" muß ein mathematisches Modell (Mengensystem, Wahrscheinlichkeitsmaß) konstruiert werden, daß dieses „Phänomen" beschreibt.

Analysis, Lineare Algebra, Numerik	Stochastik
Objekte, Abbildungen $f : \mathbb{R}^d \to \mathbb{R}^m$	Grundmenge Ω, $\mathcal{POT}(\Omega) = \{ A \mid A \subset \Omega\}$, $\mathfrak{A} \subset \mathcal{POT}(\Omega)$, Abbildung P: $\mathfrak{A} \mapsto [0,1]$

0.1 Beispiele

Phänomen	beobachtete Größe	mathematisches Modell	
Werfen einer Münze	Kopf/Zahl	$\Omega = \{K, Z\} \cong \{0.1\}$	*diskretes*
eines Würfels	Augenzahl $\{1,\dots,6\}$	$\Omega = \{1,2,3,4,5,6\}$	*Problem*
Wetter	Druck/ Temperatur/ Wind	$\Omega = \mathbb{R}^d$	*Kontinuierliches*
Börsenschwankung	Kurs	$\Omega = \mathbb{R}_+$	*Problem*
Ausfallraten von	Anzahl der defekten	$\Omega = \mathbb{Z}_+$	*endliche*
Produkten	Teile		*Anzahl*

Man nennt den Raum Ω den Stichprobenraum; er ist die Grundlage des mathematischen Modells.

0.2 Beispiel

Werfen mit zwei Würfeln $(rot, blau)$. Meßgrößen: (a,b) mit $1 \le a, b \le 6$. $\Omega = \{1,\dots,6\} \times \{1,\dots,6\}$. Interressiert nur die Augensumme $s = a+b$, so ist $\Omega = \{2,\dots,12\}$ ein möglicher Stichprobenraum.

Im allgemeinen interressiert man sich nicht nur für die Wahrscheinlichkeit von (diskreten) Elementen $i \in \Omega$, sondern für die Wahrscheinlichkeit von Teilmengen $E \subset \Omega$, den Ereignissen. Da mit $\omega \in \Omega$, $\{\omega\} \subset \Omega$ gilt, reicht es die Wahrscheinlichkeit für Ereignisse $E \subset \Omega$ zu definieren.

$$\begin{array}{lcl} \Omega \supset E & \mapsto & \mathbf{P}(E) \in [0,1] \\ \text{Ereignis} & & \text{Wahrscheinlichkeit} \\ & & \cong \text{Bewertung der Unsicherheit.} \end{array}$$

0.3 Beispiel

Werfen mit einem idealen Würfel $\Omega = \{1,\dots,6\}$. Plausibel ist, daß $\mathbf{P}(\{i\}) = \frac{1}{6}$, für $i \in \Omega$, gilt. Sei nun $E =$ „gerade Zahl" $= \{2,4,6\}$, es ist plausibel, daß $\mathbf{P}(E) = \frac{1}{2}$ gilt. Also

$$\frac{1}{2} = \mathbf{P}(E) = \mathbf{P}(\{2\} \cup \{4\} \cup \{6\}) = \mathbf{P}(\{2\}) + \mathbf{P}(\{4\}) + \mathbf{P}(\{6\}) = \frac{1}{6} + \frac{1}{6} + \frac{1}{6}$$

(Hierzu wird die Additivität gezeigt). $\mathbf{P}$ sollte also erfüllen: $E_1, E_2 \subset \Omega$ mit $E_1 \cap E_2 = \emptyset$ $\Rightarrow$ $\mathbf{P}(E_1 \cup E_2) = \mathbf{P}(E_1) + \mathbf{P}(E_2)$.

Aufgaben:

Aufgabe 0.1: Mengentheorie

Es seien $I \neq \emptyset$, $G \neq \emptyset$ Mengen und $A, B, C, M_i, N_i \in \mathcal{Pot}(G)$ für $i \in I$. Zeigen Sie:

(a) $A \cap B = A \setminus (A \setminus B)$

(b) $(A \cup B) \cap A^{\complement} = B \cap A^{\complement}$

(c) $(A \cup B)^{\complement} = A^{\complement} \cap B^{\complement}$

(d) $A \setminus \bigcup_{i \in I} M_i = \bigcap_{i \in I} (A \setminus M_i)$

(e) $A \setminus \bigcap_{i \in I} M_i = \bigcup_{i \in I} (A \setminus M_i)$

(f) $\bigcup_{i \in I} (M_i \cap N_i) \subset \left(\bigcup_{i \in I} M_i \right) \cap \left(\bigcup_{i \in I} N_i \right)$

(g) $\bigcap_{i \in I} (M_i \cup N_i) \supset \left(\bigcap_{i \in I} M_i \right) \cup \left(\bigcap_{i \in I} N_i \right)$

(h) Zeigen Sie anhand eines Gegenbeispiels, daß in (f) und (g) Gleichheit im allgemeinen nicht gilt.

$A \triangle B := (A \setminus B) \cup (B \setminus A)$ heißt die *symmetrische Differenz* von A und B. Man beweise die folgenden Eigenschaften:

(i) $A \triangle B = (A \cup B) \setminus (A \cap B)$

(j) $(A \triangle B) \cap C = (A \cap C) \triangle (B \cap C)$

(k) $A \triangle G = A^{\complement}$, $A \triangle A = \emptyset$

(l) $A = A \triangle B \Leftrightarrow B = \emptyset$

(m) $A \triangle B = B \setminus A \quad$ falls $A \subset B$

(n) $A \triangle B = A \cup B \quad$ falls $A \subset B^{\complement}$

1. Kapitel: Die σ-Algebra der Ereignisse und W.-Maße

In diesem Kapitel werden die zwei für die Stochastik grundlegenden Begriffe „Ereignis" und „Wahrscheinlichkeit" mathematisch definiert.

1.1 Definition (Stichprobenraum)

Als <u>Stichprobenraum</u> (oder <u>Merkmalmenge</u>) bezeichnen wir eine nichtleere Menge $\Omega \neq \emptyset$. Ω sollte möglichst adäquat die Ergebnisse des Experiments beschreiben. Die Wahl von Ω ist nicht eindeutig; es ist aber zu hoffen, daß die Ergebnisse des Modells nicht von der Wahl von Ω abhängen. Falls Ω endlich ist, stellt die Kombinatorik eine wichtige Methode dar, die seit dem 17. Jahrhundert in der Wahrscheinlichkeitsrechnung zentral ist. Ist Ω unendlich, so werden maßtheoretische Methoden wichtig, die *Kolmogoroff* (etwa 1930) in die Wahrscheinlichkeitstheorie eingeführt hat.

1.2 Beispiele

1.2.1 siehe 0.1

1.2.2 Skatspiel

Die Karten werden von 1 bis 32 durchnummeriert, etwa $1 \mathrel{\hat{=}}$ Kreuz Bube, $\ldots, 32 \mathrel{\hat{=}}$ Karo 7. Dann kann man

$$\Omega = \{(A, B, C) \mid A, B, C \subset \{1, \ldots 32\}\,;\ A \cap B = A \cap C = B \cap C = \emptyset;\ \#A = \#B = \#C = 10\}$$

setzen und so den Stichprobenraum aller möglichen Skatspiele definieren. Dabei bezeichne, für eine endliche Menge A, $\#A$ die Anzahl der Elemene von A.

1.3 Beispiele (wiederholtes/ zusammengesetztes Experiment)

1.3.1 Zusammengesetztes Experiment

Es wird zuerst eine Münze geworfen und dann gewürfelt:

$$\Omega = \{(W, 1)\,, (W, 2)\,, \ldots, (W, 6)\,, (Z, 1)\,, \ldots, (Z, 6)\} = \{W, Z\} \times \{1, \ldots, 6\}$$

(karthesisches Produkt).

1.3.2 Wiederholtes Experiment

Das n-fache Werfen einer Münze wird durch

$$\Omega = \{W, Z\}^n = \{\,(\omega_1, \ldots, \omega_n) \mid \omega_j \in \{W, Z\}\}$$

modelliert. Das n-Tupel $(\omega_1, \ldots, \omega_n)$ symbolisiert eine Folge von Experimenten, bei der der j-te Versuch den Wert ω_j ergeben hat.

1.3.3 Unendlich oft wiederholtes Experiment

Wird das Experiment unendlich oft wiederholt, so ist

$$\Omega = \{W, Z\}^{\mathbb{N}} = \{(\omega_1, \omega_2, \ldots) \mid \omega_j \in \{W, Z\}\}$$

ein geeigneter Stichprobenraum.

1.4 Konstruktion

Wir haben bisher den Stichprobenraum Ω als Modell zur Beschreibung der möglichen Ausgänge eines Experimentes eingeführt. Wir wollen aber auch kompliziertere Ereignisse wie „In 100 Versuchen wurde zwischen 40 und 60 mal Zahl geworfen" modellieren. Dies geschieht durch gewisse Teilmengen von Ω und logische Verknüpfungen, die durch Mengenoperatoren erzeugt werden:

$$\begin{aligned}
\text{sicheres Ereignis} &\;\widehat{=}\; \Omega \\
\text{unmögliches Ereignis} &\;\widehat{=}\; \emptyset \\
\text{Negation eines Ereignisses} &\;\widehat{=}\; A^{\complement} = \Omega \setminus A \\
\text{mindestens eines der beiden} &\;\widehat{=}\; A \cup B \\
\text{alle beide} &\;\widehat{=}\; A \cap B \\
\text{Das erste, aber nicht das zweite} &\;\widehat{=}\; A \setminus B \\
\text{mindestens eines aus einer Folge} &\;\widehat{=}\; \bigcup_{n=1}^{\infty} A_n \\
\text{alle aus einer Folge} &\;\widehat{=}\; \bigcap_{n=1}^{\infty} A_n
\end{aligned}$$

1.5 Beispiele

1.5.1 ungerade Würfelzahl

Der Würfel zeigt eine ungerade Zahl

$$\widehat{=}\, A_1 = \{1,3,5\} \subset \Omega_1 = \{1,2,3,4,5,6\}$$

1.5.2 Würfelsumme

Die Würfelsumme zweier Würfel ist ≥ 10

$$\widehat{=}\, A_2 = \{(6,4),(4,6),(5,5),(6,5),(5,6),(6,6)\} \subset \Omega_2 = \{1,\ldots,6\}^2$$

Ein geeignetes System von Mengen, das wir zur Beschreibung von Ereignissen verwenden, ist gegeben durch:

1.6 Definition (σ-Algebra)

Ein System $\mathfrak{A} \subset \mathcal{P}ot(\Omega)$ von Teilmengen von Ω heißt <u>σ-Algebra</u> auf Ω, falls

$$\begin{aligned}
(\sigma A_1) &: \quad \Omega \in \mathfrak{A} \\
(\sigma A_2) &: \quad A \in \mathfrak{A} \Rightarrow A^{\complement} \in \mathfrak{A} \\
(\sigma A_3) &: \quad \text{Ist } A_n \in \mathfrak{A}\, \forall n \geq 1 \Rightarrow \bigcup_{n=1}^{\infty} A_n \in \mathfrak{A}
\end{aligned}$$

Die $A \in \mathfrak{A}$ heißen <u>Ereignisse</u>, die $\{\omega\}$, $\omega \in \Omega$, Elementarereignisse (falls sie zu $\mathfrak{A}$ gehören).

1.7 Bemerkung

Die Tatsache, daß wir als System der „Ereignisse" i.a. <u>nicht</u> die komplette Potenzmenge $\mathcal{POT}(\Omega)$ verwenden, ist auf den ersten Blick befremdlich; sie ist aber praktikabel, unvermeidbar und sogar sinnvoll:

- Man kann zeigen, daß auf überabzählbaren Ω es prinzipiell unmöglich ist, sinnvoll Wahrscheinlichkeitsmaße auf ganz $\mathcal{POT}(\Omega)$ zu definieren, sondern nur auf einer kleineren σ-Algebra.

1.8 Satz

Es sei $\mathfrak{A}$ eine σ-Algebra auf $\Omega \neq \emptyset$ und $A, B, A_n \in \mathfrak{A}$ für alle $n \in \mathbb{N}$. Dann gilt

(a) $\emptyset \in \mathfrak{A}$, $A \cap B \in \mathfrak{A}$, $A \setminus B \in \mathfrak{A}$, $A \cup B \in \mathfrak{A}$

(b) $\bigcap\limits_{n \in \mathbb{N}} A_n \in \mathfrak{A}$

(c) **(i)** $\liminf\limits_{n \to \infty} A_n = \bigcup\limits_{n \geq 1} \bigcap\limits_{k \geq n} A_k \in \mathfrak{A}$

(ii) $\limsup\limits_{n \to \infty} A_n = \bigcap\limits_{n \geq 1} \bigcup\limits_{k \geq n} A_k \in \mathfrak{A}$

1.9 Bemerkung

$$\begin{aligned} \omega \in \liminf A_n \quad &\Leftrightarrow \quad \omega \in A_n \text{ für fast alle } n \text{ (d.h. bis auf endlich viele } n) \\ \omega \in \limsup A_n \quad &\Leftrightarrow \quad \omega \in A_n \text{ für } \infty\text{-viele } n \end{aligned}$$

1.8 Beweis

(a) $\emptyset = \Omega^{\complement} \in \mathfrak{A}$ wegen (σA_1) und (σA_2).

(b) $\bigcap\limits_{n \in \mathbb{N}} A_n = \Big(\bigcup\limits_{n \in \mathbb{N}} A_n^{\complement}\Big)^{\complement} \in \mathfrak{A}$ da $\bigcup\limits_{n \in \mathbb{N}} A_n^{\complement} \in \mathfrak{A}$ nach (σA_2) und (σA_3),

$A \cap B = \bigcap\limits_{n \in \mathbb{N}} A_n \in \mathfrak{A}$, falls $A_0 = A$, $A_1 = B$, $A_n = \Omega$ $(n \geq 2)$,

$A \setminus B = A \cap B^{\complement} \in \mathfrak{A}$,

$A \cup B = \Big(A^{\complement} \cap B^{\complement}\Big)^{\complement} \in \mathfrak{A}$.

(c) Setzt man für $n \in \mathbb{N}$, $B_n = \bigcup\limits_{k \geq n} A_k$, so ist das eine abzählbare Vereinigung von Mengen aus $\mathfrak{A}$ in $\mathfrak{A}$ für jedes $n \in \mathbb{N}$.

$$\Rightarrow \limsup A_n = \bigcap_{n \in \mathbb{N}} \bigcup_{k \geq n} A_k = \bigcap_{n \in \mathbb{N}} B_n \in \mathfrak{A}.$$

Ebenso für $\liminf A_n$.

$\square$

Falls Ω endlich oder abzählbar unendlich ist, ist alles viel einfacher:

1.10 Satz

Sei Ω abzählbar. Die einzige σ-Algebra, die alle Elementarereignisse $\{\omega\}$ für $\omega \in \Omega$ enthält, ist $\mathfrak{A} = \mathcal{POT}(\Omega)$.

Beweis: Es reicht $\mathcal{POT}(\Omega) \subset \mathfrak{A}$ zu zeigen:
Sei $A \in \mathcal{POT}(\Omega)$ beliebig. Falls A abzählbar unendlich, etwa $A = \{\omega_1, \omega_2, \ldots\}$ so ist

$$A = \bigcup_{n\in\mathbb{N}} \{\omega_n\} \in \mathfrak{A} \text{ nach } (\sigma A_3)\,.$$

Falls A $= \{\omega_1, \ldots, \omega_n\}$ endlich, ist $A = \bigcup_{j=1}^{n} \{\omega_j\} \in \mathfrak{A}$. □

σ-Algebren werden häufig nicht explizit (durch Angabe aller zugehörigen Ereignisse) sondern implizit durch Angabe von Grundereignissen und Verwendung der Axiome (1.6) angegeben. Z.B. wandelt (1.10) die implizite Forderung „alle Elementarereignisse gehören dazu” in die explizite Angabe $\mathfrak{A} = \mathcal{POT}(\Omega)$ um.

1.11 Lemma

Es sei $I \neq \emptyset$ eine Indexmenge und für $i \in I$ $\mathfrak{A}_i$ eine σ-Algebra auf Ω. Dann ist auch $\bigcap_{i\in I} \mathfrak{A}_i$ eine σ-Algebra.

Beweis:
$\Omega \in \mathfrak{A}_i$ (nach (σA_1)) für alle $i \in I \Rightarrow \Omega \in \bigcap_{i\in I} \mathfrak{A}_i$

(σA_2), (σA_3) ebenso. □

1.12 Satz und Definition

Sei $\mathcal{E} \in \mathcal{POT}(\Omega)$ ein Mengensystem. Dann ist

$$\mathfrak{A}(\mathcal{E}) := \bigcap_{\substack{\mathfrak{A}\ \sigma\text{-Algebra} \\ \mathcal{E}\subset\mathfrak{A}}} \mathfrak{A}$$

eine σ-Algebra, und zwar die kleinste, die $\mathcal{E}$ enthält. Man nennt $\mathfrak{A}(\mathcal{E})$ die von $\mathcal{E}$ erzeugte σ-Algebra.

Beweis:
Es ist $\mathcal{POT}(\Omega)$ eine σ-Algebra mit $\mathcal{E} \subset \mathcal{POT}(\Omega)$. Also enthält der Durchschitt mindestens $\mathcal{POT}(\Omega)$, ist also insbesondere nicht leer, d.h. wohldefiniert, und (1.11) zeigt, daß $\mathfrak{A}(\mathcal{E})$ eine σ-Algebra ist. Offenbar gilt $\mathcal{E} \subset \mathfrak{A}(\mathcal{E})$.

Ist $\mathfrak{A}'$ irgendeine σ-Algebra, die $\mathcal{E}$ enthält, so kommt $\mathfrak{A}'$ im Schnitt vor, also $\mathfrak{A}' \supset \bigcap_{\substack{\mathfrak{A}\ \sigma\text{-Algebra} \\ \mathcal{E}\subset\mathfrak{A}}} \mathfrak{A} = \mathfrak{A}(\mathcal{E})$.

$\mathfrak{A}(\mathcal{E})$ ist also die kleinste von $\mathcal{E}$ erzeugte σ-Algebra. □

1.13 Bemerkung

Satz (1.10) stellt also einen Zusammenhang zwischen impliziter und expliziter Definition einer σ-Algebra dar, denn er besagt:

$$\Omega \text{ abzählbar} \Rightarrow \underset{\substack{\uparrow \\ \text{implizit}}}{\mathfrak{A}(\text{Elementarereignisse})} = \underset{\substack{\uparrow \\ \text{explizit}}}{\mathcal{POT}(\Omega)}$$

Das Konstruktionsprinzip der erzeugten σ-Algebra wird nun erstmals wichtig für die aus zwei Teilexperimenten zusammengesetzten Experimente, die ja als Stichprobenraum nach (1.3) Produktmengen haben.

1.14 Beispiel (zwei Experimente)

Es werde das erste Experiment durch $(\Omega_1, \mathfrak{A}_1)$ beschrieben, sowie das zweite durch $(\Omega_2, \mathfrak{A}_2)$. Sei $\Omega = \Omega_1 \times \Omega_2$. Wir wollen mindestens die Teilmengen

$$\begin{aligned} A_1 \times \Omega_2 &\ \widehat{=}\ \text{beim ersten Experiment geschieht } A_1 \text{ und} \\ \Omega_1 \times A_2 &\ \widehat{=}\ \text{beim zweiten Experiment geschieht } A_2 \end{aligned}$$

in der σ-Algebra haben. Wir verwenden deshalb auf $\Omega = \Omega_1 \times \Omega_2$ die <u>Produkt-σ-Algebra</u>

$$\mathfrak{A}_1 \otimes \mathfrak{A}_2 := \mathfrak{A}\left(\{A_1 \times \Omega_2,\, \Omega_1 \times A_2 \mid A_1 \in \mathfrak{A}_1,\, A_2 \in \mathfrak{A}_2\}\right).$$

Insbesondere gehören dazu die Mengen

$$A_1 \times A_2 = (A_1 \times \Omega_2) \cap (\Omega_1 \times A_2) \ \text{ falls } A_j \in \mathfrak{A}_j.$$

1.14.1 Spezialfall

Sind Ω_1, Ω_2 abzählbar, so auch $\Omega_1 \times \Omega_2$. Falls $\mathfrak{A}_1 = \mathcal{POT}(\Omega_1)$, $\mathfrak{A}_2 = \mathcal{POT}(\Omega_2)$, so ist jedes Elementarereignis $\{(\omega_1, \omega_2)\} \in \Omega_1 \times \Omega_2$, da $\{(\omega_1, \omega_2)\} = \{\omega_1\} \times \{\omega_2\} \in \mathfrak{A}_1 \otimes \mathfrak{A}_2$ und somit nach (1.10) $\mathfrak{A}_1 \otimes \mathfrak{A}_2 = \mathcal{POT}(\Omega_1 \times \Omega_2)$.

1.14.2 Zweimaliger Würfelwurf

Der zweimalige Wurf eines Würfels werde durch $\Omega = \Omega_1 \times \Omega_2$ mit $\Omega_1 = \Omega_2 = \{1, \ldots, 6\}$ modelliert. Ist $\mathfrak{A}_j = \mathcal{POT}(\Omega_j)$ so ist die Produkt-σ-Algebra $\mathfrak{A}_1 \otimes \mathfrak{A}_2$ die Potenzmenge $\mathcal{POT}(\Omega)$ und folglich z.B.

$$\begin{aligned} \text{1. Wurf ungerade} &\ \widehat{=}\ \{1,3,5\} \times \{1, \ldots, 6\} \\ \text{Summe} \geq 10 &\ \widehat{=}\ \{(6,4), (4,6), (5,5), (6,5), (5,6), (6,6)\} \\ \text{Beide Würfe gleich} &\ \widehat{=}\ \{(1,1), (2,2), \ldots, (6,6)\} \end{aligned}$$

1.15 Beispiel

Ein Experiment, das durch $(\Omega, \mathfrak{A})$ beschrieben wird, soll mehrmals wiederholt werden. Die Zahl der Wiederholungen werde durch eine Menge I indiziert, also etwa

$$\left.\begin{array}{l} I = \{1, \ldots, n\} \\ I = \mathbb{N} \\ I = \mathbb{R}_+ \end{array}\right\} \text{ bei } \left\{\begin{array}{l} n \text{ Wiederholungen} \\ \infty\text{-vielen Wiederholungen} \\ \text{kontinuierlichem Experiment} \end{array}\right.$$

$I \,\widehat{=}$ Menge der Zeitpunkte.
Stichprobenraum:

$$\Omega^I = \{(\omega_i \mid i \in I) \mid \omega_i \in \Omega\} \qquad \omega_i \text{ beschreibt das Ereignis } \left\{\begin{array}{l} \text{des } i\text{-ten Experiments} \\ \text{der } i\text{-ten Beobachtung} \end{array}\right. .$$

Auf Ω^I soll nun eine geeignete σ-Algebra definiert werden. Dazu sollen zumindest Ereignisse gehören, die über ein bestimmtes Experiment $j \in I$ eine Aussage machen, also die Mengen

$$Z_j(A) = \left\{(\omega_i) \in \Omega^I \;\middle|\; \omega_j \in A,\, \omega_i \text{ beliebig für } i \neq j\right\}.$$

Die von dem System

$$\mathcal{E} = \{ Z_j (A) \mid j \in I,\ A \in \mathfrak{A} \}$$

erzeugte σ-Algebra heißt Produkt-σ-Algebra

$$\mathfrak{A}^{\otimes I} = \mathfrak{A} (\{ Z_j (A) \mid j \in I,\ A \in \mathfrak{A} \}) .$$

Falls $I = \{1, \ldots, n\}$ schreiben wir dafür auch $\mathfrak{A}^{\otimes n}$. In diesem Fall gehören die Mengen

$$A_1 \times \ldots \times A_n = \bigcap_{j=1}^{n} Z_j (A_j)$$

zu $\mathfrak{A}^{\otimes n}$, falls alle $A_j \in \mathfrak{A}_j$.

1.15.1 Bemerkung

Ist Ω höchstens abzählbar und $\mathfrak{A} = \mathcal{POT}(\Omega)$, so ist nach (1.10) auch $\mathfrak{A}^{\otimes n} = \mathcal{POT}(\Omega)$.

1.15.2 Bemerkung

Falls I unendlich ist und Ω mehr als ein Element hat, dann ist Ω^I überabzählbar und (1.15.1) gilt nicht.

1.15.3 Beispiel

Ein Würfel werde n-mal geworfen $\Omega = \{1, \ldots, 6\}^n$, dann gilt für die σ-Algebra:

$$(\mathcal{POT}\{1, \ldots, 6\})^{\otimes n} = \mathcal{POT}(\{1, \ldots, 6\}^n) .$$

Das Ereignis

$$\begin{aligned} \text{mindestens eine Sechs} \ &\widehat{=}\ \{ (\omega_1, \ldots, \omega_n) \mid \omega_j = 6 \text{ für mindestens ein } j \in \{1, \ldots, n\} \} \\ &=\ Z_1 \{6\} \cup Z_2 \{6\} \cup \ldots \cup Z_n \{6\} \end{aligned}$$

gehört also zur σ-Algebra.

1.16 Beispiel (Die Borel'sche σ-Algebra auf $\mathbb{R}$)

$\mathbb{R}$ ist überabzählbar und die Potenzmenge als Menge von Ereignissen ist zu groß. Man möchte aber zumindest die Intervalle dabei haben und auch die einelementigen Mengen.

1.16.1 Definition

Die Borel'sche σ-Algebra $\mathfrak{B}(\mathbb{R})$ auf $\mathbb{R}$ ist die kleinste σ-Algebra, die alle Intervalle $I =]a, b]$ mit $a < b$ enthält, also

$$\mathfrak{B}(\mathbb{R}) = \mathfrak{A} (\{]a, b] \mid a < b;\ a, b \in \mathbb{R} \}) .$$

Die Elemente von $\mathfrak{B}(\mathbb{R})$ heißen Borel-Mengen.

1.16.2 Eigenschaften

(1) Jedes Elementarereignis $\{a\} \in \mathfrak{B}(\mathbb{R})$,

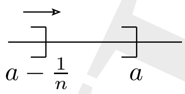
(2) $]a,b[\,,]-\infty,b[\,,]a,\infty[\in \mathfrak{B}(\mathbb{R})$ offene Intervalle,

(3) $[a,b]\,,]-\infty,b]\,, [a,\infty[\in \mathfrak{B}(\mathbb{R})$ abgeschlossene Intervalle

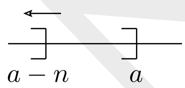
(4) $[a,b[\,,]a,b] \in \mathfrak{B}(\mathbb{R})$.

Beweis:
Es ist

$$\{a\} \quad = \quad \bigcap_{n\geq 1} \left]a-\frac{1}{n}, a\right] \in \mathfrak{B}(\mathbb{R})$$



$$\begin{aligned}]a,b[\quad &= \quad]a,b] \setminus \{b\} \in \mathfrak{B}(\mathbb{R}) \\ [a,b[\quad &= \quad]a,b[\cup \{a\} \in \mathfrak{B}(\mathbb{R}) \\ ]-\infty,b] \quad &= \quad \bigcup_{n\geq 1}]a-n,b] \in \mathfrak{B}(\mathbb{R}) \end{aligned}$$



und die übrigen gehen genauso.

1.16.3 Beispiel

Die Menge aller Zahlen in $[0,1]$ mit „7” in der Dezimalbruchentwicklung nach dem Komma ist eine Borel-Menge, da $A_1 = [0,7;0,8[$. Genauso ist die Menge der Zahlen mit „7” in der 2-ten Stelle hinter dem Komma, nämlich $A_2 = [0,07;0,08[\cup [0,17;0,18[\cup \ldots \cup [0,97;0,98[$ in $\mathfrak{B}(\mathbb{R})$ und analog $A_n \hat{=} \{7 \text{ in der } n\text{-ten Stelle}\} \in \mathfrak{B}(\mathbb{R})$. Nach (1.8) (vgl. (1.9)) gehört also auch $\limsup\limits_{n\to\infty} A_n \hat{=} \{7 \text{ kommt } \infty\text{-oft vor}\}$ zu $\mathfrak{B}(\mathbb{R})$.

1.16.4 Satz

Jede offene Teilmenge von $\mathbb{R}$ gehört zu $\mathfrak{B}(\mathbb{R})$.

Beweis:

$$U \subset \mathbb{R} \text{ offen } \Leftrightarrow \forall x \in U\, \exists \varepsilon > 0 :]x-\varepsilon, x+\varepsilon[\subset U.$$

Offensichtlich gehören die offenen Mengen $\emptyset, \mathbb{R}$ zu $\mathfrak{B}(\mathbb{R})$. Sei nun $U \neq \emptyset$ und $U \neq \mathbb{R}$. Es ist $U \cap \mathbb{Q} = \{q_0, q_1, \ldots\}$ abzählbar. Zu $q_n \in U \cap \mathbb{Q}$ wählen wir ein $\varepsilon_n > 0$, so daß $]q_n - \varepsilon_n, q_n + \varepsilon_n[\subset U$ und $]q_n - 2\varepsilon_n, q_n + 2\varepsilon_n[\not\subset U$
$\llbracket$ Geht, da $U \neq \mathbb{R}$: Wähle „maximales” $\varepsilon > 0$ mit $]x-\varepsilon, x+\varepsilon[\subset U$ für $x \in U$ $\rrbracket$
Dann gilt: $U = \bigcup\limits_{n\in\mathbb{N}}]q_n - \varepsilon_n, q_n + \varepsilon_n[$.
„$\supset$”: Gilt nach Definition von ε_n, q_n.
„$\subset$”: Sei $x \in U$ beliebig. Dann existiert ein $\varepsilon > 0$ mit $]x-\varepsilon, x+\varepsilon[\subset U$. Da $\mathbb{Q}$ dicht in $\mathbb{R}$ liegt, gibt es ein q_n mit $|q_n - x| < \dfrac{\varepsilon}{3}$.

Behauptung: $\frac{\varepsilon}{3} \leq \varepsilon_n$.

Beweis: Annahme $\frac{\varepsilon}{3} > \varepsilon_n$. Sei $z \in]q_n - 2\varepsilon_n, q_n + 2\varepsilon_n[$ beliebig

$$\begin{aligned} &\Rightarrow \quad |z - x| \leq |z - q_n| + |q_n - x| < 2\varepsilon_n + \frac{\varepsilon}{3} < 2\frac{\varepsilon}{3} + \frac{\varepsilon}{3} = \varepsilon \\ &\Rightarrow \quad]q_n - 2\varepsilon_n, q_n + 2\varepsilon_n[\subset]x - \varepsilon, x + \varepsilon[\subset U \quad \lightning \\ &\Rightarrow \quad x \in]q_n - \varepsilon_n, q_n + \varepsilon_n[\\ &\Rightarrow \quad x \in \bigcup_{n\in\mathbb{N}}]q_n - \varepsilon_n, q_n + \varepsilon_n[\in \mathfrak{B}(\mathbb{R}) \end{aligned}$$

da nach (1.16.2) $]q_n - \varepsilon_n, q_n + \varepsilon_n[\in \mathfrak{B}(\mathbb{R})$ für alle $n \geq 1$. □

1.16.5 Korollar

Jede abgeschlossene Teilmenge von $\mathbb{R}$ gehört zu $\mathfrak{B}(\mathbb{R})$.

Beweis:

Ist $A \subset \mathbb{R}$ abgeschlossen $\Rightarrow A^{\complement} \subset \mathbb{R}$ offen $\underset{(1.16.4)}{\Rightarrow} A^{\complement} \in \mathfrak{B}(\mathbb{R}) \underset{(\sigma A_2)}{\Rightarrow} A = \left(A^{\complement}\right)^{\complement} \in \mathfrak{B}(\mathbb{R})$. □

Die Borel'sche σ-Algebra $\mathfrak{B}(\mathbb{R})$ enthält also viele Mengen.

Nachdem wir bisher den Stichprobenraum Ω und die σ-Algebra der Ereignisse $\mathfrak{A}$ definiert haben, benötigen wir nur noch eine Abbildung $\mathbf{P} : \mathfrak{A} \to [0, 1]$ die uns sagt „wie wahrscheinlich" ein Ereignis $A \in \mathfrak{A}$ ist.

Dabei sollte $\mathbf{P}$ die folgenden „plausiblen" Eigenschaften besitzen:

(1) $\mathbf{P}(\Omega) = 1$ (Normierung)

(2) $\mathbf{P}(A \cup B) = \mathbf{P}(A) + \mathbf{P}(B)$, falls $A \cap B = \emptyset$ (endliche Additivität)

1.17 Definition

Es sei $\mathfrak{A}$ eine σ-Algebra auf Ω. Eine Abbildung $\mathbf{P} : \mathfrak{A} \to \mathbb{R}_+$ heißt Wahrscheinlichkeitsmaß auf Ω, falls

$$\begin{aligned} (WM_1) \quad &: \quad \mathbf{P}\left(\bigcup_{n\in\mathbb{N}} A_n\right) = \sum_{n\in\mathbb{N}} \mathbf{P}(A_n) \\ & \qquad \text{für jede disjunkte Famile } (A_n)_{n\in\mathbb{N}} \text{ von Ereignissen } A_n \in \mathfrak{A} \\ (WM_2) \quad &: \quad \mathbf{P}(\Omega) = 1 \end{aligned}$$

(Die Familie $(A_n)_n$ heißt disjunkt, falls $A_n \cap A_m = \emptyset$ für $n \neq m$.)

Das Tripel $(\Omega, \mathfrak{A}, \mathbf{P})$ wird als Wahrscheinlichkeitsraum bezeichnet.

Bemerkung: Aus (1.17) folgt bereits: $\mathbf{P}\left(A^{\complement}\right) = 1 - \mathbf{P}(A)$, da $A \cup A^{\complement} = \Omega$.

1.18 Beispiel (Laplace'scher Wahrscheinlichkeitsraum)

Ist Ω eine endliche Menge, so definiert

$$\mathbf{P} : \mathcal{POT}(\Omega) \to \mathbb{R}_+ \qquad \mathbf{P}(A) := \frac{\#A}{\#\Omega}$$

ein Wahrscheinlichkeitsmaß. Diese diskrete Gleichverteilung wird unter anderem dann verwendet, wenn aus Symmetriegründen alle Elementarereignisse als gleichwahrscheinlich angensehen werden. Wir bezeichnen die diskrete Gleichverteilung auf Ω mit $\mathfrak{L}_\Omega$, $(\Omega, \mathcal{POT}(\Omega), \mathfrak{L}_\Omega)$ heißt Laplace-Raum.

1.19 Beispiele

1.19.1 Fairer Münzwurf

Modell: $\Omega = \{W, Z\}$ mit $\mathfrak{A} = \mathcal{POT}(\Omega)$.
Gleichverteilung: $\mathbf{P}(A) = \frac{1}{2}$ für $A \subset \Omega$.

1.19.2 Verfälschte Münze

Wappen $\widehat{=}$ 1, Zahl $\widehat{=}$ 0.
Modell: $\Omega = \{0, 1\}$, $\mathfrak{A} = \mathcal{POT}(\Omega)$.
Ist die Münze nicht fair, so liegt ein Bernoulli-Experiment vor: Es sei $p \in [0, 1]$. Dann ist durch $\mathfrak{B}_{1,p} : \mathcal{POT}(\Omega) \to \mathbb{R}_+$; $\mathfrak{B}_{1,p}(\emptyset) = 0$, $\mathfrak{B}_{1,p}(\Omega) = 1$, $\mathfrak{B}_{1,p}(1) = p$, $\mathfrak{B}_{1,p}(0) = 1 - p$ die allgemeine Form eines Wahrscheinlichkeitsmaßes auf einer 2-elementigen Menge gegeben. Für $p = \frac{1}{2}$ erhalten wir wieder die Gleichverteilung aus (1.19.1).

1.19.3 Wiederholtes Werfen einer fairen Münze

Modell: $\Omega = \{0, 1\}^n$ mit $\mathfrak{A} = \mathcal{POT}(\Omega)$ und der Laplace-Verteilung $\mathbf{P} = \mathfrak{L}_\Omega$. Das Ereignis „alle Würfe gleich" $\widehat{=}$ $A = \{(0, \ldots, 0), (1, \ldots, 1)\}$ und es gilt:

$$\mathbf{P}(A) = \frac{2}{2^n} = 2^{1-n}$$

1.19.4 Binary search

Aus einer angeordneten Menge von $2^n - 1$ Elementen wird ein Schlüsselwort gesucht, indem man im ersten Schritt das 2^{n-1}-te Element mit dem zu suchenden vergleicht. Besteht Gleichheit, ist man fertig; ist das zu suchende Element kleiner, so macht man rekursiv mit den ersten $2^{n-1} - 1$ Elementen weiter; ist es größer, so setzt man das Intervallhalbierungsverfahren mit den $2^{n-1} - 1$ Elementen $2^{n-1} + 1, \ldots, 2^n - 1$ fort.
Modell: $\Omega = \{0, 1, \ldots, 2^n - 1\}$ wobei 0 $\widehat{=}$ Element nicht gefunden.
Die Teilmenge $\left\{2^{n-2}, 2^{n-1}, 3 \cdot 2^{n-2}\right\}$ ist dann zum Beispiel das Ereignis „man braucht höchstens 2 Suchschritte" und $A = \{0, 1, 3, \ldots, 2^n - 1\}$ ist „man braucht n Schritte" $\mathbf{P} = \mathfrak{L}_\Omega$. Es folgt $\mathbf{P}(A) = \frac{1}{2} + 2^{-n}$.

1.19.5 Punktmaß

Es sei $\mathfrak{A}$ eine σ-Algebra auf Ω und $x \in \Omega$. Dann ist durch

$$\mathcal{E}_x(A) = 1_A(x) = \begin{cases} 1 & x \in A \\ 0 & x \notin A \end{cases}$$

ein Wahrscheinlichkeitsmaß auf Ω definiert.
$\llbracket$ Es ist $\mathcal{E}_x(\Omega) = 1$, da $x \in \Omega$ und für eine disjunkte Familie $(A_n)_n \subset \mathfrak{A}$ gilt entweder

(i) x liegt in genau einer der $A_n \Rightarrow x \in \bigcup_{n \in \mathbb{N}} A_n \Rightarrow \mathcal{E}_x\left(\bigcup_{n \in \mathbb{N}} A_n\right) = 1$ und $\mathcal{E}_x(A_n) = 1$ für genau ein n.

(ii) x liegt in keinem der $A_n \Rightarrow x \notin \bigcup_{n \in \mathbb{N}} A_n$ $\rrbracket$

1.20 Satz

Es sei $(\Omega, \mathfrak{A}, \mathbf{P})$ ein Wahrscheinlichkeitsraum.

(a)
$$\mathbf{P}(\emptyset) = 0$$

(b) $\forall A, B \in \mathfrak{A}$ gilt
$$\mathbf{P}(A \cup B) = \mathbf{P}(A) + \mathbf{P}(B) - \mathbf{P}(A \cap B);$$
insbesondere gilt $\mathbf{P}(A \cup B) = \mathbf{P}(A) + \mathbf{P}(B)$, falls A und B disjunkt sind.

(c) Im Fall $A, B \in \mathfrak{A}$ und $A \subset B$ gilt
$$\mathbf{P}(B \setminus A) = \mathbf{P}(B) - \mathbf{P}(A).$$
Insbesondere $\mathbf{P}(A) \leq \mathbf{P}(B)$ und $\mathbf{P}\left(A^{\complement}\right) = 1 - \mathbf{P}(A)$.

(d) Es sei $(A_n)_n \subset \mathfrak{A}$ eine aufsteigende Folge, d.h. $A_n \subset A_{n+1}$ $\forall n \in \mathbb{N}$. Dann gilt:
$$\mathbf{P}\left(\bigcup_{n\in\mathbb{N}} A_n\right) = \lim_{n\to\infty} \mathbf{P}(A_n)$$
(Stetigkeit von unten)

(e) Es sei $(A_n)_n \subset \mathfrak{A}$ eine absteigende Folge, d.h. $A_n \supset A_{n+1}$ $\forall n \in \mathbb{N}$. Dann gilt:
$$\mathbf{P}\left(\bigcap_{n\in\mathbb{N}} A_n\right) = \lim_{n\to\infty} \mathbf{P}(A_n)$$

Beweis:

(a) Setzen wir $A_0 = \Omega$ und $A_n = \emptyset$ $\forall n \geq 1$
$$\underset{\substack{(WM_1)\\(WM_2)}}{\Rightarrow} 1 = \mathbf{P}(\Omega) = \mathbf{P}\left(\bigcup_{n\in\mathbb{N}_0} A_n\right) = \mathbf{P}(A_0) + \sum_{n\geq 1} \mathbf{P}(\emptyset) \geq 1 + \mathbf{P}(\emptyset)$$
$\Rightarrow\ 0 \geq \mathbf{P}(\emptyset) \geq 0$.

(b) 1. Fall: $A \cap B = \emptyset$. Setze $A_0 = A$, $A_1 = B$ und $A_n = \emptyset$ für $n \geq 2$
$$\overset{(a)}{\Rightarrow} \mathbf{P}(A \cup B) = \mathbf{P}\left(\bigcup_{n\in\mathbb{N}_0} A_n\right) = \mathbf{P}(A) + \mathbf{P}(B).$$
2. Fall: $A, B \in \mathfrak{A}$ beliebig $\Rightarrow B \setminus A$ und A disjunkt und $B \setminus A$ und $A \cap B$ disjunkt
$$\begin{aligned}
\Rightarrow \mathbf{P}(A \cup B) &= \mathbf{P}((B \setminus A) \cup A) = \mathbf{P}(B \setminus A) + \mathbf{P}(A) \\
&= \mathbf{P}(B \setminus A) + \mathbf{P}(A \cap B) + \mathbf{P}(A) - \mathbf{P}(A \cap B) \\
&= \mathbf{P}((B \setminus A) \cup (A \cap B)) + \mathbf{P}(A) - \mathbf{P}(A \cap B) \\
&= \mathbf{P}(B) + \mathbf{P}(A) - \mathbf{P}(A \cap B).
\end{aligned}$$

(c) Folgt sofort aus (b).

(d) Betrachte die disjunkte Folge $(B_n)_n$ mit $B_0 = A_0$; $B_{n+1} = A_{n+1} \setminus A_n$. Dann gilt:

$$\mathbf{P}(A_n) = \mathbf{P}\left(\bigcup_{m=0}^{n} B_m\right) = \sum_{m=0}^{n} \mathbf{P}(B_m)$$

$$\begin{aligned}\Rightarrow \mathbf{P}\left(\bigcup_{n\in\mathbb{N}} A_n\right) &= \mathbf{P}\left(\bigcup_{n\in\mathbb{N}} B_n\right) = \sum_{n\in\mathbb{N}} \mathbf{P}(B_n) \\ &= \lim_{n\to\infty} \sum_{m=0}^{n} \mathbf{P}(B_m) = \lim_{n\to\infty} \mathbf{P}(A_n).\end{aligned}$$

(e) Folgt aus (c) und (d).

1.21 Anwendung (n-faches Würfeln)

Modell: $\Omega = \{1,\ldots,6\}^n$, $\mathfrak{A} = \mathcal{POT}(\Omega)$, $\mathbf{P} = \mathfrak{L}_\Omega$

$$\begin{aligned} A &= \{(\omega_1,\ldots,\omega_n) \mid \omega_j = 6 \text{ für mindestens ein } j \leq n\} \\ &\hat{=} \text{ mindestens eine } 6\end{aligned}$$

Um $\mathbf{P}(A)$ zu berechen ist es einfacher $\mathbf{P}\left(A^{\complement}\right)$ zu berechnen und dann (1.20.c) zu verwenden. Es ist

$$A^{\complement} = \{(\omega_1,\ldots,\omega_n) \mid \omega_j \leq 5 \text{ für alle } j\}$$

Somit ist $\mathbf{P}(A) = 1 - \mathbf{P}\left(A^{\complement}\right) = 1 - \left(\frac{5}{6}\right)^n$.

Im Fall $n = 13$ gilt $\mathbf{P}(A) \approx 0,907$. Man müßte also 13 mal würfeln um mit Wahrscheinlichkeit von mindestens 90% eine (oder mehrere) Sechsen zu würfeln.

Es ist oft mühsam für alle Ereignisse $A \in \mathfrak{A}$ den Wert von $\mathbf{P}(A)$ anzugeben. Diskrete, d.h. abzählbare Wahrscheinlichkeitsräume können durch den Begriff der Zähldichte einfacher beschrieben werden.

1.22 Definition Zähldichte

Es sei Ω abzählbar. Eine Funktion $f : \Omega \to \mathbb{R}_+$ heißt Zähldichte, falls

$$\sum_{\omega\in\Omega} f(\omega) = 1.$$

1.23 Satz

Es sei Ω abzählbar und $\mathfrak{A} = \mathcal{POT}(\Omega)$. Zwischen Wahrscheinlichkeitsmaßen $\mathbf{P}$ und Zähldichten f auf Ω besteht folgender bijektiver Zusammenhang:

(a) Ist $\mathbf{P}$ ein Wahrscheinlichkeitsmaß auf Ω, so ist durch

$$f_{\mathbf{P}} : \Omega \to \mathbb{R}_+,\ f_{\mathbf{P}}(\omega) := \mathbf{P}(\{\omega\}),\ \omega \in \Omega$$

eine Zähldichte $f_{\mathbf{P}}$ definiert.

(b) Für jede Zähldichte f gibt es genau ein Wahrscheinlichkeitsmaß $\mathbf{P}$ mit $f_{\mathbf{P}} = f$, nämlich $\mathbf{P}(A) := \sum_{\omega \in A} f(\omega)$.

Beweis:

(a) Es ist $f_{\mathbf{P}}(\omega) = \mathbf{P}(\{\omega\}) \geq 0$ und wegen (WM_1) und (WM_2) gilt

$$\sum_{\omega \in \Omega} f_{\mathbf{P}}(\omega) = \sum_{\omega \in \Omega} \mathbf{P}(\{\omega\}) = \mathbf{P}\left(\bigcup_{\omega \in \Omega} \{\omega\} \right) = \mathbf{P}(\Omega) = 1.$$

(b) Es ist $\mathbf{P}(\Omega) = \sum_{\omega \in \Omega} f(\omega) = 1$. Ist $(A_n)_n$ eine disjunkte Familie so gilt wegen der Assoziativität der Summanden

$$\mathbf{P}\left(\bigcup_{n \in \mathbb{N}} A_n \right) = \sum_{\omega \in \bigcup_{n \in \mathbb{N}} A_n} f(\omega) = \sum_{n \in \mathbb{N}} \sum_{\omega \in A_n} f(\omega) = \sum_{n \in \mathbb{N}} \mathbf{P}(A_n)$$

$\Rightarrow$ $\mathbf{P}$ ist ein Wahrscheinlichkeitsmaß.
Ist $\mathbf{Q}$ ein weiteres Wahrscheinlichkeitsmaß mit $f_{\mathbf{Q}} = f$ so folgt für alle $A \subset \Omega$

$$\mathbf{Q}(A) = \mathbf{Q}\left(\bigcup_{\omega \in A} \{\omega\} \right) = \sum_{\omega \in A} \mathbf{Q}(\{\omega\}) = \sum_{\omega \in A} f_{\mathbf{Q}}(\omega) = \sum_{\omega \in A} f(\omega) = \mathbf{P}(A).$$

□

Es genügt also zur vollständigen Beschreibung eines diskreten Wahrscheinlichkeitsmaßes seine Zähldichte anzugeben.

1.24 Beispiel

1.24.1 Gleichverteilung

Die Gleichverteilung $\mathfrak{L}_\Omega$ auf einer endlichen Menge Ω hat die konstante Funktion $f(\omega) = \frac{1}{\#\Omega}$ als Zähldichte.

1.24.2 Binomialverteilung

Es seien $p \in [0,1]$ und $n \geq 1$, $\Omega = \{0, \ldots, n\}$, $\mathfrak{A} = \mathcal{POT}(\Omega)$. Dann ist durch

$$f(k) = \binom{n}{k} p^k (1-p)^{n-k}$$

(für $k \in \Omega$) eine Zähldichte definiert, denn $f(k) \geq 0$ und

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

Das durch f definierte Wahrscheinlichkeitsmaß $\mathcal{B}_{n,p}$ mit $\mathcal{B}_{n,p}\{k\} = \binom{n}{k} p^k (1-p)^{n-k}$ heißt Binomialverteilung mit Parameter n und p. Im Fall $n = 1$ erhalten wir die Bernoulli-Verteilung aus (1.19.2).

1.24.3 Poisson-Verteilung

Ein wichtiges Wahrscheinlichkeitsmaß auf $\mathbb{N}$ ist die Poisson-Verteilung Π_λ mit Parameter $\lambda \geq 0$. Sie ist durch ihre Zähldichte

$$f(n) := \Pi_\lambda(\{n\}) := \mathbf{e}^{-\lambda}\frac{\lambda^n}{n!}$$

($n \in \mathbb{N}$) definiert.

$\llbracket$ Es ist $f(n) \geq 0$ und $\displaystyle\sum_{n=0}^{\infty} \mathbf{e}^{-\lambda}\frac{\lambda^n}{n!} = \mathbf{e}^{-\lambda} \cdot \mathbf{e}^{\lambda} = \mathbf{e}^0 = 1$ $\rrbracket$

1.24.4 Geometrische Verteilung

Die geometrische Verteilung $\mathcal{G}_p$ ist für $p \in]0,1[$ das Wahrscheinlichkeitsmaß auf $\mathbb{N}$ mit Zähldichte $f(n) = p(1-p)^n$.

1.25 Beispiel (Kontinuierliche Gleichverteilung auf $[0,1]$)

Der naive Versuch eine Gleichverteilung auf $\Omega = [0,1]$ durch eine Zähldichte $f(\omega) = \varepsilon$ zu erhalten schlägt fehl:
Ist $\varepsilon > 0 \Rightarrow \mathbf{P}(\Omega) = \displaystyle\sum_{\omega\in\Omega} \varepsilon = \infty \cdot \varepsilon = \infty \neq 1$ und ist $\varepsilon = 0 \Rightarrow \mathbf{P}(\Omega) = \displaystyle\sum_{\omega\in\Omega} 0 = 0 \neq 1$.
Wir zeigen nun, daß unsere Probleme unter anderem daher rühren, daß $\mathfrak{A} = \mathcal{POT}(\Omega)$ als σ-Algebra zu groß ist. Wir formulieren den Begriff der Gleichverteilung durch die Translationsinvarianz, d.h. für alle $A \in \mathfrak{A}$, $x \in \mathbb{R}$, für die verschobene Menge $A + x = \{a + x \mid a \in A\} \subset [0,1]$ ist, muß gelten $\mathbf{P}(A + x) = \mathbf{P}(A)$, wobei $\mathbf{P}$ die „Gleichverteilung" auf $[0,1]$ sein soll.

1.25.1 Bemerkung

Es gibt kein auf ganz $\mathcal{POT}([0,1])$ definiertes translationsinvariantes Wahrscheinlichkeitsmaß $\mathbf{P}$.
Beweis: (Vitali) $\rightsquigarrow$ Stochastik II

Es ist also unvermeidlich, daß wir eine kleinere σ-Algebra auf $\Omega = [0,1]$ wählen. Wir verwenden die Borel'sche σ-Algebra

$$\mathcal{B}([0,1]) = \{A \in \mathcal{B}(\mathbb{R}) \mid A \subset [0,1]\}$$

(Spur σ-Algebra)

1.25.2 Faktum

Es gibt genau ein Wahrscheinlichkeitsmaß $\mathbf{P} =: \lambda^1_{[0,1]}$ auf $\mathcal{B}([0,1])$, das translationsinvariant ist. Es ist charakterisiert durch die Eigenschaft

$$\lambda^1_{[0,1]}(]a,b]) = b - a \quad \text{für } 0 \leq a < b \leq 1.$$

Es folgt

$$\begin{aligned}
\lambda^1_{[0,1]}(\{a\}) &\overset{*}{=} \lim_{n\to\infty} \lambda^1_{[0,1]}\left(\left]a - \frac{1}{n}, a\right]\right) \\
&= \lim_{n\to\infty} a - \left(a - \frac{1}{n}\right) \\
&= \lim_{n\to\infty} \frac{1}{n} = 0
\end{aligned}$$

(Bei $*$ geht die Stetigkeit von oben ein!)

$$\Rightarrow \quad \lambda^1_{[0,1]}\left([a,b]\right) = \lambda^1_{[0,1]}\left(]a,b[\right) = \lambda^1_{[0,1]}\left([a,b[\right) = b-a$$

Die Translationsinvarianz für Intervalle kann man direkt ablesen:

$$\lambda^1_{[0,1]}\left(]a+x,b+x]\right) = (b+x)-(a+x) = b-a = \lambda^1_{[0,1]}\left(]a,b]\right).$$

1.25.3 Beispiel

Wähle $\omega \in [0,1] = \Omega$ zufällig. Für $n \geq 1$ betrachten wir das Ereignis: "die k-te Nachkommastelle der Dezimalbruchentwicklung (eindeutig, falls Periode 9 ausgeschlossen!) ist 7".

$$A_n := \dot{\bigcup_{k=0}^{10^{n-1}-1}} \left[\frac{k}{10^{n-1}} + \frac{7}{10^n}, \frac{k}{10^{n-1}} + \frac{8}{10^n}\right[$$

($A \dot{\cup} B$ bedeutet: $A \cup B$ und $A \cap B = \emptyset$)
Es ist

$$\begin{aligned}\lambda^1_{[0,1]}\left(A_n\right) &= \sum_{k=0}^{10^{n-1}-1} \left[\left(\frac{k}{10^{n-1}} + \frac{8}{10^n}\right) - \left(\frac{k}{10^{n-1}} + \frac{7}{10^n}\right)\right] \\ &= 10^{n-1} \cdot \frac{1}{10^n} = \frac{1}{10}.\end{aligned}$$

Wir wollen nun wissen mit welcher Wahrscheinlichkeit irgendwo eine 7 vorkommt, also die Wahrscheinlichkeit von

$$\bigcup_{n=1}^{\infty} A_n = A.$$

Deren Komplement (nirgendwo eine 7) ist

$$B = A^{\complement} = \bigcap_{n=1}^{\infty} A_n^{\complement} = \bigcap_{n=1}^{\infty} B_n,$$

wobei

$$\begin{aligned} B_n &= \dot{\bigcup_{k_1,\ldots,k_n \in \{0,\ldots,9\}\setminus\{7\}}} \left[\frac{k_1}{10} + \ldots + \frac{k_n}{10}, \frac{k_1}{10} + \ldots + \frac{k_n+1}{10}\right[\\ &\mathrel{\widehat{=}} \text{die ersten } n\text{-Stellen sind} \neq 7 \end{aligned}$$

$$\Rightarrow \quad \lambda^1_{[0,1]}\left(B_n\right) = \lim_{n\to\infty}\left(\frac{9}{10}\right)^n = 0$$

$$\Rightarrow \quad \lambda^1_{[0,1]}\left(A\right) = 1$$

d.h. mit Wahrscheinlichkeit 1 kommt in jeder zufällig gezogenen Zahl $\omega \in [0,1]$ eine 7 vor.

Aufgaben:

Aufgabe 1.1:

Ritter de Méré glaubte mit 4 Würfeln mindestens eine Sechs zu werfen habe dieselbe Wahrscheinlichkeit, wie mit 2 Würfeln bei 24 Würfen mindestens eine Doppelsechs zu werfen. Stimmt dies?

(a) Geben Sie für beide Spiele je einen geeigneten Stichprobenraum Ω an.

(b) Formulieren Sie die Ereignisse „mindestens eine Sechs" und „mindestens eine Doppelsechs" als Teilmenge des entsprechenden Stichprobenraums.

(c) Berechnen Sie die Wahrscheinlichkeiten und entscheiden Sie ob der Ritter de Méré richtig lag.

Aufgabe 1.2:

Geben sie bei den Aufgaben eine geeignete σ-Algebra zur mathematischen Beschreibung des Experiments an.

(a) Ein Würfel wird geworfen und abhängig vom Ergebnis i des Wurfes aus einer von sechs Urnen $U_1, \ldots, U_6$ eine Kugel gezogen, wobei in der i-ten Urne i rote und $7 - i$ schwarze Kugeln liegen. Stellen sie das Ereignis „eine rote Kugel wurde gezogen" mathematisch dar.

(b) Beim Eiskunstlauf beurteilen die Richter die Leistungen von sechs Läufern. Jeder der Richter darf genau einmal jede Note $1, \ldots, 6$ vergeben. Stellen Sie das Ergebnis „ein Läufer erhält nur die Note 6 von allen Richtern" dar.

Aufgabe 1.3:

$\mathfrak{A} \subset \mathcal{POT}(\Omega)$ sei eine σ-Algebra, $(A_n)_n \subset \mathfrak{A}$ eine Folge von Ereignissen.

(a) Zeigen sie, daß

$$\begin{aligned} \omega \in \liminf_{n\to\infty} A_n \quad &\Leftrightarrow \quad \omega \in A_n \text{ für fast alle } n \\ \omega \in \limsup_{n\to\infty} A_n \quad &\Leftrightarrow \quad \omega \in A_n \text{ für unendlich viele } n \end{aligned}$$

(b) Zeigen Sie, daß die folgenden zusammengesetzten Ereignisse in $\mathfrak{A}$ liegen:

(i) Mindestens zwei der Ereignisse (A_n) treten ein.

Mit einem Run der Länge k bezeichnet man eine Folge von k hintereinander auftretenden Ereignissen aus (A_n).

(ii) Runs beliebiger Länge treten auf.

(iii) Für jedes $n \in \mathbb{N}$ gibt es nur endlich viele Runs der Länge n.

(*Hinweis*: Drücken Sie die Ereignisse mit Hilfe der Mengenoperatoren $\cap$, $\cup$ und $\setminus$, durch die Ereignisse (A_n) aus.)

Aufgabe 1.4:

Es sei $I \neq \emptyset$ eine Indexmenge und für $i \in I$ sei $\mathfrak{A}_i$ eine σ-Algebra auf Ω. Zeigen Sie, daß dann auch

$$\mathfrak{A} = \bigcap_{i \in I} \mathfrak{A}_i$$

eine σ-Algebra ist.

Bitte geben Sie bei den folgenden beiden Aufgaben einen geeigneten Wahrscheinlichkeitsraum zur mathemathischen Beschreibung des Experiments an.

Aufgabe 1.5:

Auf einem Parkplatz mit zwölf Plätzen stehen acht Autos, wobei die vier freien Plätze alle nebeneinander sind. Untersuchen Sie die Frage, ob diese Anordnung zufällig ist, indem Sie die Wahrscheinlichkeit dieses Ereignisses bei zufälliger Anordnung der acht Autos berechnen.

Aufgabe 1.6:

In einem Multiple-Choice-Test mit 20 Aufgaben sind pro Aufgabe 5 Antworten vorgesehen, wovon genau eine richtig ist. Ein risikofreudiger Kandidat kreuzt die Antworten zufällig an. Wie viele Möglichkeiten hat er:

(a) Den Bogen auszufüllen.

(b) So auszufüllen, daß alle Antworten richtig sind.

(c) So auszufüllen, daß alle Antworten falsch sind.

(d) Wie groß ist die Wahrscheinlichkeit, daß er alle Antworten richtig, bzw. falsch hat?

Aufgabe 1.7:

Es sei $p \in \,]0,1]$. Zeigen Sie, daß es genau ein Wahrscheinlichkeitsmaß $\mathcal{G}_p$ auf $(\Omega := \mathbb{N}_0, \mathcal{POT}(\mathbb{N}_0))$ gibt mit $\mathcal{G}_p(\{0,1,\ldots,n\}) = 1-(1-p)^{n+1}$ für alle $n \in \mathbb{N}_0$. Berechenen Sie die Wahrscheinlichkeit für das Auftreten einer geraden Zahl. (Hinweis: Berechnen Sie zuerst die Zähldichte von $\mathcal{G}_p$.)

Aufgabe 1.8:

Spieler A wirft sechs Würfel und gewinnt bei wenigstens einem Sechser. Spieler B wirft mit zwölf Würfeln und gewinnt bei wenigstens zwei Sechsen.

Welcher Spieler hat die größere Gewinnwahrscheinlichkeit?

Aufgabe 1.9:

$\mathfrak{A} \subset \mathcal{POT}(\Omega)$ sei eine σ-Algebra, $(\mu_i)_{i\in\mathbb{N}}$ eine Folge von Wahrscheinlichkeitsmaßen auf $(\Omega, \mathfrak{A})$ und $(p_i)_{i\in\mathbb{N}} \subset [0,1]$ mit $\sum\limits_{i=1}^{n} p_i = 1$.

Zeigen Sie, daß durch

$$\nu(A) = \sum_{i=1}^{\infty} p_i \mu_i(A) \quad \text{für } A \in \mathfrak{A}$$

ein Wahrscheinlichkeitsmaß auf $(\Omega, \mathfrak{A})$ definiert wird.

2. Kapitel: Der Laplace'sche W.-Raum und Kombinatorik

Wird als stochastisches Modell die Gleichverteilung $\mathfrak{L}_\Omega$ auf einer endlichen Menge Ω gewählt (geschieht immer dann, wenn die Elementarereignisse wegen ihrer Gleichartigkeit als gleichwahrscheinlich angesehen werden), ist die Bestimmung der Zahl der Elemente von $A \subset \Omega$ wesentlich, wegen $\mathfrak{L}_\Omega(A) = \frac{\#A}{\#\Omega}$. Diese Fragestellung ist Gegenstand der Kombinatorik. Dabei interessiert uns nur die Zahl der Möglichkeiten, wie aus einer gegebenen Menge mehrere Elemente ausgewählt werden können, etwa Ziehen aus einer Urne mit/ ohne Zurücklegen usw.

2.1 Definition (Permutation/ Kombination)

Es sei $A \neq \emptyset$ und $k \geq 1$.

(a) **(i)** Jedes k-Tupel $(a_1, \ldots, a_k)$ mit $a_i \in A$ heißt k-Permutation aus A mit Wiederholung, A^k ist die Menge der k-Permutationen aus A mit Wiederholung. Die Menge A^k beschreibt das Ziehen von Kugeln mit Zurücklegen unter Beachtung der Reihenfolge.

(ii) Eine k-Permutation aus A ohne Wiederholung ist ein k-Tupel $(a_1, \ldots, a_k)$ mit $a_i \neq a_j$ für alle $i \neq j$. Die Menge aller solcher Permutationen bezeichnen wir mit $\mathcal{P}_k^A$. Sie beschreibt das Ziehen ohne Zurücklegen unter Beachtung der Reihenfolge.

Kommt es auf die Reihenfolge nicht an, so spricht man von Kombinationen:

(b) **(i)** Als k-Kombination ohne Widerholung bezeichnen wir jede k-elementige Teilmenge $\{a_1, \ldots, a_k\} \subset A$; die Menge dieser Kombinationen nennen wir K_k^A.

(ii) k-Kombination mit Wiederholung bezeichnen wir als $\{a_1, \ldots, a_k\} \subset A$, wobei jetzt der Fall $a_j = a_i$ nicht ausgeschlossen ist. (Wir identifizieren $\{1,1,2\}$ mit $\{1,2,1\}$ und $\{2,1,1\}$ aber nicht mit $\{1,2\}$.) Die Menge aller k-Kombinationen mit Wiederholung aus A bezeichnen wir mit M_k^A.

2.2 Satz

Es sei $A \neq \emptyset$ und $\#A = n$ und $k \geq 1$. Dann gilt für

(1) k-Permutationen mit Wiederholung:

$$\#A^k = n^k$$

(2) k-Permutationen ohne Wiederholung:

$$\#\mathcal{P}_k^A = (n)_k = n \cdot (n-1) \cdot \ldots \cdot (n-k+1) = k! \binom{n}{k}$$

(3) k-Kombination ohne Wiederholung:

$$\#K_k^A = \binom{n}{k}$$

(4) k-Kombination mit Wiederholung:

$$\#M_k^A = \binom{n+k-1}{k}$$

Beweis:

(a) $\#A^k = (\#A)^k = n^k$

(b) Für a_1 gibt es n Möglichkeiten, für a_2 gibt es $(n-1)$ Möglichkeiten und für a_j gibt es, nachdem $a_1, \ldots, a_{j-1}$ festgelegt sind, $(n-j+1)$ Möglichkeiten. Insgesamt gibt es also $n \cdot (n-1) \cdot \ldots \cdot (n-k+1) = k!\binom{n}{k}$ Möglichkeiten.

(c) Kommt es auf die Reihenfolge nicht an, so fallen jeweils $k!$ Permutationen, die aus denselben Elementen bestehen, zu einer k-Kombination ohne Wiederholung zusammen. Deren Anzahl beträgt also $\frac{1}{k!} \cdot k!\binom{n}{k} = \binom{n}{k} = \frac{n!}{k!(n-k)!}$.

(d) Ohne Einschränkung (OE): $A = \{1, \ldots, n\}$. Jeder k-Kombination mit Wiederholung $\{a_1, \ldots, a_k\}$ (wobei wir die a_i so aufschreiben können, daß $a_1 \leq a_2 \leq \ldots \leq a_k$ ist) ordnen wir eine k-Kombination ohne Wiederholung

$$f(\{a_1, \ldots, a_k\}) = \{a_1 + 1, a_2 + 2, \ldots, a_k + k\}$$

aus Elementen von $A' = \{2, \ldots, n+k\}$ zu. f ist offensichtlich injektiv und auch surjektiv, denn das Urbild von $\{a'_1, \ldots, a'_n\} \in K_k^{A'}$ ist $\{a'_1 - 1, a'_2 - 2, \ldots, a'_k - k\} \in M_k^A$.

$$\Rightarrow \quad \#M_k^A = \#K_k^{A'} = \binom{\#A'}{k} = \binom{n+k-1}{k}$$

$\square$

2.3 Beispiele

2.3.1 Skat

Die Wahrscheinlichkeit, beim Skat alle 4 Buben auf die Hand zu bekommen ist $\frac{\binom{28}{6}}{\binom{32}{10}} \approx 0,0058$.

Denn:
<u>Modell</u>: $\Omega = K_{10}^{\{1,\ldots 32\}}$ (Laplace-Raum) hat $\binom{32}{10}$ Elemente. Das Ereignis

$$\text{„4 Buben"} \mathrel{\hat{=}} A := \{\omega \in \Omega \mid 4, 12, 20, 28 \in \omega\}$$

besteht aus den Teilmengen von $\{1, \ldots, 32\}$ die die 4 Buben $4, 12, 20, 28$ sowie noch 6 andere aus den restlichen 28 Karten beliebig ausgewählt enthält, also $\#A = \binom{28}{6}$.

2.3.2 Urne

Aus einer Urne, in der $N = r + s$ Kugeln und zwar r rote und s schwarze liegen, werden $n \leq N$ Kugeln (ohne Zurücklegen und ohne Reihenfolge) gezogen. Gesucht ist die Wahrscheinlichkeit, daß genau k davon rot sind.

$$\underbrace{1, \ldots, r}_{\text{rot}}, \underbrace{r+1, \ldots, r+s}_{\text{schwarz}} \qquad \Omega = K_n^{\{1,\ldots,r+s\}} \text{ mit } \mathbf{P} = \mathfrak{L}_\Omega$$

$$\begin{aligned} A_k &:= \{\omega \in \Omega \mid \#(\omega \cap \{1, \ldots, r\}) = k\} \\ &= \{\omega \subset \{1, \ldots, r+s\} \mid \#(\omega \cap \{1, \ldots, r\}) = k \text{ und } \#(\omega \cap \{r+1, \ldots, r+s\}) = n-k\} \end{aligned}$$

Der k-elementige Teil der roten Kugeln läßt sich auf genau $\binom{r}{k}$-Arten auswählen und der $(n-k)$-elementige Teil der schwarzen auf $\binom{s}{n-k}$.

$$\Rightarrow \quad \#A_k = \binom{r}{k} \cdot \binom{s}{n-k} \quad \Rightarrow \quad \mathbf{P}(A_k) = \frac{\binom{r}{k} \cdot \binom{s}{n-k}}{\binom{r+s}{n}},$$

denn $\#\Omega = \#K_n^{\{1,\ldots,r+s\}} = \binom{r+s}{n}$.

Wegen $\Omega = \dot{\bigcup}_{k=0}^{n} A_k \quad \Rightarrow \quad \sum_{k=0}^{n} \mathbf{P}(A_k) = 1$

$$\Rightarrow \quad \mathcal{H}_{n,N,r}\{k\} := \frac{\binom{r}{k} \cdot \binom{N-r}{n-k}}{\binom{N}{n}} \qquad 0 \leq k \leq n$$

definiert ein Wahrscheinlichkeitsmaß auf $\{0, \ldots, n\}$, die Hypergeometrische Verteilung.

2.3.3 Hypergeomertische Verteilung

In einer Lieferung von 10000 Schrauben seien 2% defekt. Nimmt man 100 Schrauben heraus, so ist die Wahrscheinlichkeit für genau k defekte gerade $\mathcal{H}_{100,10000,200}\{k\}$.

k	0	1	2	3	4	5
$\mathcal{H}_{100,10000,200}\{k\}$	$0,131$	$0,271$	$0,275$	$0,183$	$0,090$	$0,035$

2.4 Beispiele

Die Mengen A^k, $\mathcal{P}_k^A$, K_k^A, M_k^A können nicht nur zur Modellierung des Ziehens von k Kugeln aus einer Urne mit n Kugeln verwendet werden.

Andere Sichtweise: Verteilung von k Gegenständen auf n Zellen: Das Tupel $(a_1, \ldots, a_k)$ beschreibt dann das Einlegen eines Gegenstandes in Zelle a_1, eines in Zelle a_2 usw.

Belegungsmodell	Urnenmodell	mathematisches Modell
k Gegenstände auf n Zellen	k mal Ziehen aus einer Urne mit n Kugeln	Mengen $A = \{1, \ldots, n\}$
(1) unterscheidbare Objekte	**(1)** mit Berücksichtigung der Reihenfolge	**(1)** Permutationen
(a) mit Ausschlußprinzip	**(a)** ohne Zurücklegen	**(a)** $\Omega = \mathcal{P}_k^A$
(b) ohne Ausschlußprinzip	**(b)** mit Zurücklegen	**(b)** $\Omega = A^k$
(2) Nicht unterscheidbare Objekte	**(2)** Ohne Berücksichtigung der Reihenfolge	**(2)** Kombinationen
(a) mit Ausschlußsprinzip	**(a)** ohne Zurücklegen	**(a)** $\Omega = K_k^A$
(b) ohne Ausschlußprinzip	**(b)** mit Zurücklegen	**(b)** $\Omega = M_k^A$

2.4.1 Geburtstag

Gesucht ist die Wahrscheinlichkeit, daß in einer Gruppe von k Personen ein Geburtstag mehrfach vorkommt.

Modell: $n = 365$ Zellen und k unterscheidbare Objekte, Mehrfachbildungen sind möglich.

$$\Omega = \{1, \ldots, n\}^k \qquad \mathfrak{A} = \mathcal{POT}(\Omega) \qquad \mathbf{P} = \mathfrak{L}_\Omega$$

Sei

$$\begin{aligned} E &\triangleq \text{„ein Geburtstag kommt mehrfach vor"} \\ &= \{\omega = (\omega_1, \ldots, \omega_k) \in \Omega \mid \exists i, j \in \{1, \ldots, k\} : \omega_i = \omega_j\} \end{aligned}$$

$$\begin{aligned}
\Rightarrow \quad E^{\complement} &\;\widehat{=}\; \text{„alle Geburtstage verschieden"} \\
&= \{\omega = (\omega_1, \ldots, \omega_k) \in \Omega \mid \omega_i \neq \omega_j \; \forall i \neq j\} \\
&= \mathcal{P}_k^{\{1,\ldots,n\}}
\end{aligned}$$

$$\begin{aligned}
\Rightarrow \quad \mathbf{P}(E) = 1 - \mathbf{P}\left(E^{\complement}\right) &= 1 - \frac{\# \mathcal{P}_k^{\{1,\ldots,n\}}}{\# \Omega} \\
&= 1 - \frac{(n)_k}{n^k} = 1 - \frac{n \cdot (n-1) \cdot \ldots \cdot (n-k+1)}{n^k} \\
&= 1 - \left(1 - \frac{1}{n}\right) \cdot \left(1 - \frac{2}{n}\right) \cdot \ldots \cdot \left(1 - \frac{k-1}{n}\right)
\end{aligned}$$

Für $x \in [0,1]$ gilt $1 - x \leq e^{-x}$

$$\begin{aligned}
\Rightarrow \quad \mathbf{P}(E) &= 1 - \left(1 - \frac{1}{n}\right) \cdot \left(1 - \frac{2}{n}\right) \cdot \ldots \cdot \left(1 - \frac{k-1}{n}\right) \\
&\geq 1 - \exp\left(-\frac{1}{n} - \frac{2}{n} - \ldots - \frac{k-1}{n}\right) \\
&= 1 - \exp\left(-\frac{k(k-1)}{2n}\right)
\end{aligned}$$

Für $n = 365$ ist $\mathbf{P}(E) \geq \frac{1}{2}$ für $k \geq 23$, für $k = 70$ ist $\mathbf{P}(E) \geq 0{,}9987$.

Ein wichtiges Hilfsmittel, daß in der Kombinatorik häufig verwendet wird, aber für beliebige Wahrscheinlichkeitsmaße gilt, ist die Siebformel:

2.5 Satz (*Sylvester, Poincaré*)

Es sei $(\Omega, \mathfrak{A}, \mathbf{P})$ ein Wahrscheinlichkeitsraum und $A_1, \ldots, A_n \in \mathfrak{A}$. Die Siebformel liefert die Wahrscheinlichkeit, daß mindestens eines der $A_1, \ldots, A_n$ eintritt:

$$\mathbf{P}\left(\bigcup_{j=1}^{n} A_j\right) = \sum_{k=1}^{n} (-1)^{k-1} \sum_{\substack{J \subset \{1,\ldots,n\} \\ \# J = k}} \mathbf{P}\left(\bigcap_{j \in J} A_j\right).$$

Beweis: Entweder direkt aus der Definition des Wahrscheinlichkeitsmaßes oder später (folgt aus Satz (2.7.b), der später bewiesen wird).

2.6 Beispiel

Auf einem Ball, an dem n Paare teilnehmen, werden alle Paare zufällig zusammengesetzt. Wie groß ist die Wahrscheinlichkeit, daß sich mindestens ein Paar wiederfindet?

Modell: $\Omega = \mathcal{P}_n^{\{1,\ldots,n\}}$, $\mathfrak{A} = \mathcal{POT}\{\Omega\}$, $\mathbf{P} = \mathfrak{L}_\Omega$.

Ein Elementarereignis $\omega = (\omega_1, \ldots, \omega_n)$ bedeutet dabei:

Dame 1 tanzt mit Partner von Dame ω_1

Dame 2 tanzt mit Partner von Dame ω_2

usw.

Uns interessiert die Wahrscheinlichkeit des Ereignisses

$$A = \{\omega \in \Omega \mid \omega_j = j \text{ für mindestens ein } j = 1, \ldots, n\}$$

Für $J \subset \{1,\dots,n\}$ sei $A_J := \{\omega \in \Omega \mid \omega_j = j \,\forall j \in J\} \cong \mathcal{P}_{n-l}^{\{1,\dots,n\}\setminus J}$ falls $\#J = l$.

$$\Rightarrow \quad \mathbf{P}\left(A_J\right) = \frac{\#A_J}{\#\Omega} = \frac{(n-l)!}{n!} \quad \text{für } \#J = l. \qquad (*)$$

Nach der Siebformel gilt dann:

$$\begin{aligned}
\mathbf{P}\left(A\right) = \mathbf{P}\left(\bigcup_{j=1}^{n} A_{\{j\}}\right) &= \sum_{l=1}^{n} (-1)^{l-1} \sum_{\substack{J\subset\{1,\dots,n\}\\ \#J=l}} \mathbf{P}\left(\bigcap_{j\in J} A_{\{j\}}\right)\\
&= \sum_{l=1}^{n} (-1)^{l-1} \sum_{\substack{J\subset\{1,\dots,n\}\\ \#J=l}} \mathbf{P}\left(A_J\right)\\
&\overset{(*)}{=} \sum_{l=1}^{n} (-1)^{l-1} \sum_{\substack{J\subset\{1,\dots,n\}\\ \#J=l}} \frac{(n-l)!}{n!}\\
&= \sum_{l=1}^{n} (-1)^{l-1} \frac{(n-l)!}{n!} \cdot \#K_l^{\{1,\dots,n\}}\\
&= \sum_{l=1}^{n} (-1)^{l-1} \binom{n}{l} \frac{(n-l)!}{n!} = \sum_{l=1}^{n} \frac{(-1)^{l-1}}{l!}\\
&= 1 - \frac{1}{\mathbf{e}} + \sum_{l=n+1}^{\infty} \frac{(-1)^l}{l!} \approx 0,632 \quad \text{schon für } n \text{ relativ klein.}
\end{aligned}$$

Die Siebformel ist ein Spezialfall der folgenden Aussage, die die Wahrscheinlichkeit, daß genau bzw. mindestens k der Ereignisse $A_1,\dots,A_n$ eintreten, angibt:

2.7 Satz (Ein-Ausschluß-Prinzip)

Es sei $(\Omega, \mathfrak{A}, \mathbf{P})$ ein Wahrscheinlichkeitsraum und $A_1,\dots,A_n \in \mathfrak{A}$. Für $1 \le k \le n$ sei

$$B_k := \bigcup_{\substack{J\subset\{1,\dots,n\}\\ \#J=k}} \bigcap_{j\in J} A_j \cap \bigcap_{j\notin J} A_j^{\complement}$$

(„genau k der Ereignisse treten ein")

$$C_k := \bigcup_{\substack{J\subset\{1,\dots,n\}\\ \#J=k}} \bigcap_{j\in J} A_j$$

(„mindestens k der Ereignisse treten ein").
Es bezeichne

$$S_k := \sum_{\substack{J\subset\{1,\dots,n\}\\ \#J=k}} \mathbf{P}\left(\bigcap_{j\in J} A_j\right)$$

für $1 \le k \le n$. Dann gilt:

(a)

$$\mathbf{P}\left(B_k\right) = \sum_{l=k}^{n} (-1)^{l-k} \binom{l}{k} \cdot S_l$$

(b)

$$\mathbf{P}(C_k) = \sum_{l=k}^{n} (-1)^{l-k} \binom{l-1}{k-1} \cdot S_l$$

Beweis:
Der Beweis von (a) kommt später (siehe (3.26)).
(b): Wir zeigen, wie (b) aus (a) folgt:
Beweis durch absteigende Induktion nach k:

$\boxed{k = n}$ Es ist $C_n = \bigcap_{j \in \{1,\dots,n\}} A_j$

$$\Rightarrow \quad \mathbf{P}(C_n) = \mathbf{P}\left(\bigcap_{j=1}^{n} A_j\right) = S_n.$$

$\boxed{k+1 \rightsquigarrow k \geq 1}$ Es ist $C_k = B_k \,\dot{\cup}\, C_{k+1}$

$$\Rightarrow \quad \mathbf{P}(C_k) = \mathbf{P}(B_k) + \mathbf{P}(C_{k+1})$$

Nach Induktionsvoraussetzung und Teil (a) folgt:

$$\begin{aligned}
\mathbf{P}(C_k) &= \mathbf{P}(B_k) + \mathbf{P}(C_{k+1}) \\
&= \sum_{l=k}^{n} (-1)^{l-k} \binom{l}{k} \cdot S_l + \sum_{l=k+1}^{n} (-1)^{l-k-1} \binom{l-1}{k} \cdot S_l \\
&= S_k + \sum_{l=k+1}^{n} (-1)^{l-k} \left(\binom{l}{k} - \binom{l-1}{k} \right) \cdot S_l \\
&= S_k + \sum_{l=k+1}^{n} (-1)^{l-k} \binom{l-1}{k-1} \cdot S_l \\
&= \sum_{l=k}^{n} (-1)^{l-k} \binom{l-1}{k-1} \cdot S_l,
\end{aligned}$$

da

$$\begin{aligned}
\binom{l}{k} - \binom{l-1}{k} &= \frac{l!}{k! \cdot (l-k)!} - \frac{(l-1)!}{k! \cdot (l-1-k)!} = \frac{l!}{k! \cdot (l-k)!} - \frac{(l-1)! \cdot (l-k)}{k! \cdot (l-k)!} \\
&= \frac{l! - (l-1)! \cdot (l-k)}{k! \cdot (l-k)!} = \frac{l \cdot (l-1)! - (l-1)! \cdot (l-k)}{k \cdot (k-1)! \cdot (l-k)!} \\
&= \frac{(l-(l-k)) \cdot (l-1)!}{k \cdot (k-1)! \cdot (l-k)!} = \frac{k \cdot (l-1)!}{k \cdot (k-1)! \cdot (l-k)!} \\
&= \frac{(l-1)!}{(k-1)! \cdot (l-k)!} = \binom{l-1}{k-1}
\end{aligned}$$

□

2.8 Beispiele (Ein-Ausschluß-Prinzip)

2.8.1 Skat

Wie groß ist die Wahrscheinlichkeit, beim Skat nach dem Austeilen 3 Buben zu bekommen?
Modell: $\Omega = K_{10}^{\{1,\dots,32\}}$, $\mathfrak{A} = \mathcal{POT}(\Omega)$, $\mathbf{P} = \mathfrak{L}_\Omega$, Buben: $1, 2, 3, 4$

Für $1 \leq j \leq 4$ sei

$$A_j = \{\omega \in \Omega \mid j \in \omega\} \mathrel{\hat{=}} \text{„man bekommt Bube } j\text{“}.$$

Gesucht ist dann die Wahrscheinlichkeit von

$$\begin{aligned} B_3 &\mathrel{\hat{=}} \text{„genau 3 Buben“}, \\ C_3 &\mathrel{\hat{=}} \text{„mindestens 3 Buben“}. \end{aligned}$$

Es ist

$$S_3 = \sum_{\substack{J \subset \{1,\ldots,4\} \\ \#J=3}} \mathbf{P}\left(\bigcap_{j\in J} A_j\right).$$

Das Ereignis 3 bestimmte Buben zu bekommen ist isomorph zu $K_7^{\{1,\ldots,29\}}$ (3 Buben fest, bleiben 7 Karten aus 29).

$\Rightarrow$ Für $J \subset \{1,\ldots,4\}$ und $\#J = 3$ gilt:

$$\mathbf{P}\left(\bigcap_{j\in J} A_j\right) = \frac{\#K_7^{\{1,\ldots,29\}}}{\#\Omega} = \frac{\binom{29}{7}}{\binom{32}{10}} = \frac{3}{124}$$

$$\begin{aligned} \Rightarrow \quad S_3 &= \frac{3}{124} \cdot \sum_{\substack{j \subset \{1,\ldots,4\} \\ \#J=3}} 1 = \frac{3}{124} \cdot \#K_3^{\{1,\ldots,4\}} = \frac{3}{124} \cdot \binom{4}{3} = \frac{12}{124} = \frac{3}{31} \\ S_4 &= \mathbf{P}\left(\bigcap_{j=1}^{4} A_j\right) \underset{(2.3.1)}{=} \frac{\binom{28}{6}}{\binom{32}{10}} = \frac{21}{3596} \end{aligned}$$

$$\begin{aligned} \Rightarrow \quad \mathbf{P}(B_3) &= \binom{3}{3} \cdot S_3 - \binom{4}{3} \cdot S_4 = \frac{66}{899} \approx 0,0734 \\ \mathbf{P}(C_3) &= \binom{2}{2} \cdot S_3 - \binom{3}{2} \cdot S_4 = \frac{285}{3596} \approx 0,0793 \end{aligned}$$

2.8.2 Sortieren

Beim alphabetischen Sortieren einer Liste mit n Einträgen ist bei bestimmten Algorithmen (z.B. quicksort) die Rechendauer besonders groß, wenn viele Elemente richtig sortiert sind.

Gesucht ist die Wahrscheinlichkeit, daß genau k Elemente am richtigen Platz sind.

Modell: $\Omega = \mathcal{P}_n^{\{1,\ldots n\}}$, $\mathfrak{A} = \mathcal{POT}(\Omega)$, $\mathbf{P} = \mathfrak{L}_\Omega$.

$$B_k = \{\omega \in \Omega \mid \omega_j = j \text{ für genau } k \text{ Zahlen } j \in \{1,\ldots,n\}\}.$$

Sei $A_j = \{\omega \in \Omega \mid \omega_j = j\}$ und für $J \subset \{1,\ldots,n\}$ sei $A_J = \{\omega \in \Omega \mid \omega_j = j \text{ für alle } j \in J\}$.

In (2.6) (siehe Stern $(*)$) haben wir gesehen, daß

$$S_l = \binom{n}{l} \cdot \mathbf{P}\left(A_{\{1,\ldots,l\}}\right) = \frac{1}{l!}.$$

$$\begin{aligned} (2.7.a) \quad \Rightarrow \quad \mathbf{P}(B_k) &= \sum_{l=k}^{n} (-1)^{l-k} \cdot \binom{l}{k} \cdot \frac{1}{l!} \\ &= \sum_{l=k}^{n} (-1)^{l-k} \cdot \frac{1}{k!\,(l-k)!} \\ &= \frac{1}{k!} \cdot \sum_{l=0}^{n-k} \frac{(-1)^l}{l!} \approx \frac{1}{k! \cdot \mathbf{e}}. \end{aligned}$$

Aufgaben:

Aufgabe 2.1:

(a) Zeigen Sie, daß die Menge

$$\left\{f : \{1, \ldots, r\} \to \{1, \ldots, n\} \mid \forall i \in \{1, \ldots, n\} : \#f^{-1}(\{i\}) = k_i\right\}$$

aus $\dfrac{r!}{k_1! \cdot \ldots \cdot k_n!}$ Elementen besteht, wobei $k_i \geq 0$, $\sum\limits_{j=1}^{n} k_j = r$ und $r \geq n$ gilt.

(b) Bestimmen Sie mittels (a) die Anzahl der surjektiven Abbildungen zwischen $\{1, \ldots, r\}$ und $\{1, \ldots, n\}$.

(c) Bestimmen Sie mit der Siebformel die Anzahl der surjektiven Abbildungen zwischen $\{1, \ldots, r\}$ und $\{1, \ldots, n\}$, indem Sie zunächst die Kardinalitäten der Menge der nicht surjektiven Abbildungen bestimmen.
(*Hinweis*: Die Siebformel für Kardinalitäten folgt aus Satz (2.5), indem man $\mathbf{P} = \mathfrak{L}_\Omega$ wählt und dann mit $\#\Omega$ multipliziert!)

Aufgabe 2.2:

(a) In jeder Packung des Waschmittels SOREIN befindet sich ein Coupon mit einem der sechs Buchstaben des Namens. Man erhält eine Packung SOREIN umsonst, wenn man sämtliche Buchstaben des Namens zusammen hat. Berechnen Sie die Wahrscheinlichkeit in zehn Packungen die nötigen Coupons zu finden, wenn alle 6^{10} Buchstabenkombinationen gleichwahrscheinlich sind.

(b) Auch die Firma OSOREIN startet eine ähnliche Kampangne; man erhält zwei Packungen umsonst, wenn man den Namen OSOREIN mit sieben Coupons bilden kann. Wie groß ist hier die Wahrscheinlichkeit, in zehn Packungen die nötigen Buchstaben zu finden?
(*Hinweis*: Aufgabe (2.1) ist nützlich.)

Aufgabe 2.3:

Man plaziere auf einem Schachbrett zufällig 8 Türme. Wie groß ist die Wahrscheinlichkeit dafür, daß keiner der Türme einen anderen Schlagen kann?
(*Hinweis*: Man zähle ab, auf wie viele Arten sich 8 Türme auf dem 8×8-Schachbrett so aufstellen lassen, daß in jeder Zeile und in jeder Spalte genau ein Turm steht.)

Aufgabe 2.4:

Eine Firma stellt drei verschiedene Artikel a, b und c her. Sie berichtet, daß von 1000 befragten Haushalten 70 mindestens a und b, 98 mindestens b und c, 119 mindestens a und c, 49 alle drei und 190 mindestens zwei der Artikel benutzen. Kann man diesen Angaben Glauben schenken?
(*Hinweis*: Untersuchen Sie mit der Siebformel die Kardinalitäten der Menge aller Haushalte, die mindestens zwei Artikel benutzen.)

Aufgabe 2.5:

Vom großen Mathematiker und Raucher S. Banach ist überliefert, daß er stets in beiden Hosentaschen eine Streichholzschachtel hatte. Zum Anzünden einer Ziagarette wählte er willkürlich eine der beiden Schachteln aus und entnahm ihr ein Streichholz.
Beide Streichholzschachteln enthalten zu Beginn je n Streichhölzer. Bestimmen Sie die Wahrscheinlichkeit dafür, daß nach wiederholtem Anzünden einer Zigarette in der einen Schachtel noch k Streichhölzer sind, wenn aus der anderen Schachel das letzte Streichholz entnommen wird.

3. Kapitel: Wahrscheinlichkeitsmaß mit Dichte auf $\mathbb{R}^d$

Im letzten Kapitel haben wir endliche bzw. abzählbar unendliche Wahrscheinlichkeitsräume betrachtet. Nach (1.23) sind die Wahrscheinlichkeitsmaße dann durch ihre Zähldichte eindeutig bestimmt, d.h. es gilt

$$\mathbf{P}(A) = \sum_{\omega \in A} \mathbf{P}(\{\omega\}).$$

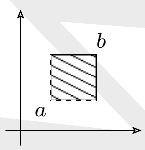
Häufig treten aber auch Meßgrößen auf, die reelle Zahlen oder Vektoren im $\mathbb{R}^d$ sind. Wie in (1.25) gezeigt, gilt im Falle der kontinuierlichen Gleichverteilung auf $[0,1]$: $\lambda^1\{\omega\} = 0$ für alle $\omega \in [0,1]$. Die von uns jetzt betrachteten Wahrscheinlichkeitsmaße auf $\mathbb{R}^d$ werden die Eigenschaft $\mathbf{P}(\{\omega\}) = 0\, \forall \omega \in \mathbb{R}^d$ besitzen. Gleichwohl werden auch diese Maße wieder mit Hilfe einer Funktion $f : \mathbb{R}^d \to \mathbb{R}_+$ definiert, ihrer Lebesgue-Dichte im Unterschied zur Zähldichte aus §1. Dazu müssen wir zuerst das Integral einer positiven meßbaren Funktion einführen, sowie eine kleinere σ-Algebra als die Potenzmenge. (1.25.1) gilt hier entsprechend!

3.1 Definition (Borel'sche σ-Algebra)

(a) Für $a = (a_1, \ldots, a_d)$, $b = (b_1, \ldots, b_d) \in \mathbb{R}^d$ sei

$$]a,b] = \{x = (x_1, \ldots, x_d) \mid a_j < x_j \leq b_j \text{ für alle } 1 \leq j \leq n\}$$

ein halboffener Quader.



(a) Die Borel'sche σ-Algebra auf $\mathbb{R}^d$ ist die kleinste σ-Algebra $\mathfrak{B}\left(\mathbb{R}^d\right)$, die alle halboffenen Quader $]a,b]$ mit $a, b \in \mathbb{R}^d$ enthält, d.h.

$$\mathfrak{B}\left(\mathbb{R}^d\right) = \mathfrak{A}\left(\{]a,b] \mid a,b \in \mathbb{R}^d\}\right).$$

Jedes Element $E \in \mathfrak{B}\left(\mathbb{R}^d\right)$ nennen wir (Borel-)meßbar oder Borelmenge.

(b) Für $\Omega \in \mathfrak{B}\left(\mathbb{R}^d\right)$ sei

$$\mathfrak{B}(\Omega) = \left\{ A \in \mathfrak{B}\left(\mathbb{R}^d\right) \,\middle|\, A \subset \Omega \right\}$$

die σ-Algebra der Borel'schen Mengen auf Ω.

3.2 Satz

$\mathfrak{B}\left(\mathbb{R}^d\right)$ enthält alle offenen und alle abgeschlossenen Teilmengen von $\mathbb{R}^d$.

Beweis:

Es sei $A \subset \mathbb{R}^d$ offen. Es sei

$$M = \left\{]a,b] \,\middle|\, a,b \in \mathbb{Q}^d;\ ]a,b] \subset A \right\}.$$

Dann ist M abzählbar und somit ist, wegen (σA_3),

$$A' := \bigcup_{Q \in M} Q \in \mathfrak{B}\left(\mathbb{R}^d\right).$$

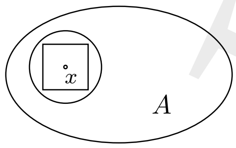
Wir zeigen nun: $A = A'$, wodurch die Behauptung bewiesen ist.
Trivialerweise gilt $A' \subset A$.
Umgekehrt („$\supset$"): Zu $x \in A \, \exists \varepsilon > 0$ mit

$$K_\varepsilon(x) = \left\{ y \in \mathbb{R}^d \,\middle|\, \|x-y\| < \varepsilon \right\}.$$

In $K_\varepsilon(x)$ gibt es ein $]a,b] \in M$ mit $x \in \,]a,b]$ und somit

$$x \in \bigcup_{Q \in M} Q = A',$$

d.h. $A \subset A'$, insgesamt also $A = A'$.



Ist $B \subset \mathbb{R}^d$ abgeschlossen, so ist $B^{\complement}$ offen und deshalb eine Borelmenge und wegen (σA_1)

$$\Rightarrow \quad B \in \mathfrak{B}\left(\mathbb{R}^d\right).$$

$\square$

3.3 Definition (Maß)

Es sei $\Omega \neq \emptyset$ und $\mathfrak{A}$ eine σ-Algebra auf Ω. Eine Abbildung

$$\mu : \mathfrak{A} \to \mathbb{R}_+ \cup \{\infty\}$$

mit $\mu(\emptyset) = 0$ und

$$\mu\left(\bigcup_{n\in\mathbb{N}} A_n\right) = \sum_{n\in\mathbb{N}} \mu(A_n)$$

für jede disjunkte Folge $(A_n) \in \mathfrak{A}$ heißt <u>Maß</u> auf Ω. (vgl. (1.17)!)

3.4 Konvention (Rechnen mit ∞)

Auf $\mathbb{R}_+ \cup \{\infty\}$ gilt :

$$\begin{aligned} a + \infty = \infty + a &= \infty && \text{für alle } a \in \mathbb{R}_+ \\ a \cdot \infty = \infty \cdot a &= \infty && \text{für alle } a > 0 \\ 0 \cdot \infty = \infty \cdot 0 &= 0. \end{aligned}$$

3.5 Bemerkung

Die für Wahrscheinlichkeitsmaße in Satz (1.20) bewiesenen Aussagen (a), (b), die Monotonieaussage $\mu(A) \leq \mu(B)$ für $A \subset B$ sowie (d) bleiben für allgemeine Maße gültig. Bei der Stetigkeit von oben (1.20.e) ist zusätzlich $\mu(A_0) < \infty$ zu fordern.
Das für uns wichtige Maß auf dem $\mathbb{R}^d$ ist das <u>Lebesgue-Maß λ^d</u>. Die Existenz wird in Stochastik II bewiesen.

3.6 Faktum

Es gibt auf $\left(\mathbb{R}^d, \mathfrak{B}\left(\mathbb{R}^d\right)\right)$ ein eindeutig bestimmtes Maß λ^d mit

$$\lambda^d(]a,b]) = \prod_{j=1}^{d} (b_j - a_j)$$

für alle Quader mit $a_j < b_j$ für $1 \leq j \leq d$. Dieses Maß wird <u>Lebesgue(-Borel)'sches Maß</u> genannt.

3.7 Beispiele

3.7.1 $\lambda^1, \lambda^2, \lambda^3$

Im Fall $d = 1$ mißt λ^1 die Gesamtlänge einer Teilmenge, λ^2 die Fläche, λ^3 das Volumen.

3.7.2 $\mathbf{P}(A) = \dfrac{\lambda^d(A)}{\lambda^d(\Omega)}$

Es sei $\Omega \in \mathfrak{B}\left(\mathbb{R}^d\right)$ mit $0 < \lambda^d(\Omega) < \infty$. Dann ist auf $(\Omega, \mathfrak{B}(\Omega))$ ein Wahrscheinlichkeitsmaß

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

durch

$$\mathbf{P}(A) := \frac{\lambda^d(A)}{\lambda^d(\Omega)}$$

definiert.

$\llbracket$ (WM_1) folgt aus der entsprechenden Eigenschaft von λ^d und (WM_2) gilt wegen

$$\mathbf{P}(\Omega) = \frac{\lambda^d(\Omega)}{\lambda^d(\Omega)} = 1.$$

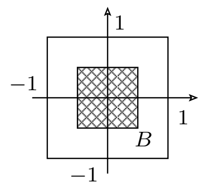
$\rrbracket$

$\mathbf{P}$ heißt kontinuierliche Gleichverteilung auf Ω. Den Fall $d = 1$, $\Omega = [0,1]$ hatten wir schon im Beispiel (1.25) betrachtet.

3.7.3 „Quadratische Zielscheibe"

Es sei $\Omega =]-1,1] \times]-1,1]$ eine quadratische Zielscheibe. Ist die Treff-Fläche durch $B :=]-a,a] \times]-a,a]$ mit $0 < a < 1$ gegeben, so ist bei gleichverteilten Treffern auf Ω die Wahrscheinlichkeit einen Treffer in B zu landen

$$\mathbf{P}(B) = \frac{\lambda^2(B)}{\lambda^2(\Omega)} = \frac{2a \cdot 2a}{2 \cdot 2} = a^2.$$



Um eine weitere große Klasse von Wahrscheinlichkeitsmaßen auf dem $\mathbb{R}^d$ zu definieren, müssen wir das Integral einer nichtnegativen integrierbaren Funktion definieren

Dazu benötigen wir:

3.8 Definition (Meßbarkeit)

Es sei $\mathfrak{A}$ eine σ-Algebra auf Ω. Eine Funktion $f : \Omega \to \overline{\mathbb{R}} := \mathbb{R} \cup \{-\infty, \infty\}$ heißt <u>meßbar</u>, wenn:

Für alle $\alpha \in \mathbb{R}$:

$$\{\omega \in \Omega \mid f(\omega) < \alpha\} \in \mathfrak{A}.$$

3.9 Beispiele

3.9.1 Stetige Funktionen

Im Fall $\Omega = \mathbb{R}^d$, $\mathfrak{A} = \mathfrak{B}\left(\mathbb{R}^d\right)$ ist jede stetige Funktion $f : \mathbb{R}^d \to \mathbb{R}$ meßbar.

$\llbracket$ $\{\omega \in \Omega \mid f(\omega) < \alpha\} = f^{-1}(]-\infty, \alpha[)$ offen, da $]-\infty, \alpha[\subset \mathbb{R}$ offen und nach Satz (3.2) somit in $\mathfrak{B}\left(\mathbb{R}^d\right)$. $\rrbracket$

3.9.2 Komponierte Funktionen

Die Menge aller meßbaren Funktionen $f : \Omega \to \mathbb{R}$ ist ein $\mathbb{R}$-Vektorraum. Sind $f, g : \Omega \to \mathbb{R}$ meßbar und $c \in \mathbb{R}$, so sind

$$c \cdot f : \omega \mapsto c \cdot f(\omega) \quad \text{und} \quad f + g : \omega \mapsto f(\omega) + g(\omega)$$

meßbar.

⟦ Klar für $c \cdot f$. Für $f + g$ folgt dies aus

$$\{\omega \in \Omega \mid f(\omega) + g(\omega) < \alpha\} = \bigcup_{\beta \in \mathbb{Q}} \{\omega \in \Omega \mid f(\omega) < \beta\} \cap \{\omega \in \Omega \mid g(\omega) < \alpha - \beta\} \in \mathfrak{A}$$

„$\supset$“ Klar!

„$\subset$“ Sei $f(\omega) + g(\omega) < \alpha$ und $\beta \in \mathbb{Q}$ mit $f(\omega) < \beta < \alpha - g(\omega)$ gewählt. Nach Voraussetzung gilt

$$f(\omega) < \alpha - g(\omega) \quad \Rightarrow \quad \omega \in \{\omega \in \Omega \mid f(\omega) < \beta\} \cap \{\omega \in \Omega \mid g(\omega) < \alpha - \beta\}$$

⟧

3.9.3 Indikatorfunktion

Die Indikatorfunktion $1_A : \Omega \to \mathbb{R}$ ist für $A \subset \Omega$ durch

$$1_A(\omega) = \begin{cases} 1 & \omega \in A \\ 0 & \omega \in A^{\complement} \end{cases}$$

definiert.

1_A ist genau dann meßbar, wenn $A \in \mathfrak{A}$.

⟦

$$\{\omega \in \Omega : 1_A(\omega) < \alpha\} = \left\{ \begin{array}{ll} \emptyset & \alpha \leq 0 \\ A^{\complement} & 0 < \alpha \leq 1 \\ \Omega & \alpha > 1 \end{array} \right\} \in \mathfrak{A}$$

$$\Leftrightarrow \quad A^{\complement} \in \mathfrak{A} \quad \Leftrightarrow \quad A \in \mathfrak{A}$$

⟧

3.9.4 Meßbare Elementarfunktionen

Jede Funktion der Gestalt

$$g = \sum_{j=1}^{n} \alpha_j \cdot 1_{A_j} \quad \text{mit } \alpha_j \in \mathbb{R}_+, A_j \in \mathfrak{A}$$

ist nach (3.9.2) und (3.9.3) meßbar. Funktionen dieser Gestalt heißen <u>meßbare Elementarfunktionen</u>. Die Menge aller solcher Funktionen wird mit $\mathcal{E}(\mathfrak{A})$ bezeichnet. Elementarfunktionen sind gerade die meßbaren Funktionen $g \geq 0$, die nur endlich viele Werte annehmen.

3.9.5 Folge meßbarer Elementarfunktionen

Es sei (f_n) eine Folge meßbarer Funktionen. Dann ist auch $\inf_{n \in \mathbb{N}} f_n$ und $\sup_{n \in \mathbb{N}} f_n$ meßbar.

⟦

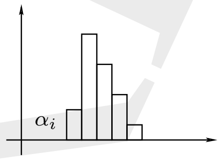
$$\left\{\omega \in \Omega \,\middle|\, \inf_{n \in \mathbb{N}} f_n(\omega) < \alpha\right\} = \bigcup_{n \in \mathbb{N}} \{\omega \in \Omega \mid f_n(\omega) < \alpha\} \in \mathfrak{A}$$

⟧

3.10 Definition (Integral von Elementarfunktionen)

Für $g \in \mathcal{E}(\mathfrak{A})$ und ein Maß μ auf $(\Omega, \mathfrak{A})$ definieren wir

$$\int g \,\mathbf{d}\mu := \int_\Omega g(\omega)\ \mathbf{d}\mu(\omega) := \sum_{x \in \mathbb{R}} x \cdot \mu\{\omega \in \Omega \mid g(\omega) = x\}.$$



Diese Summe hat nur endlich viele Summanden.

3.11 Beispiele

3.11.1

Für $A \in \mathfrak{A}$ gilt

$$\int 1_A \,\mathbf{d}\mu = 0 \cdot \mu\left(A^{\complement}\right) + 1 \cdot \mu(A) = \mu(A).$$

Es folgt leicht, daß die Elementarfunktion $g = \sum_{j=1}^{n} \alpha_j \cdot 1_{A_j}$ das Integral $\int g \,\mathbf{d}\mu = \sum_{j=1}^{n} \alpha_j \cdot \mu(A_j)$ hat.

⟦ Trivial für disjunkte A_j und dann durch Zerlegung in disjunkte Bestandteile. ⟧

3.11.2

$$\int c \cdot g \,\mathbf{d}\mu = c \cdot \int g \,\mathbf{d}\mu \quad \text{für } g \in \mathcal{E}(\mathfrak{A}),\ c \geq 0$$

3.11.3

$$\int g_1 + g_2 \,\mathbf{d}\mu = \int g_1 \,\mathbf{d}\mu + \int g_2 \,\mathbf{d}\mu \quad \text{für } g_1, g_2 \in \mathcal{E}(\mathfrak{A})$$

3.11.4

$$\int g_1 \,\mathbf{d}\mu \leq \int g_2 \,\mathbf{d}\mu \quad \text{falls } g_1, g_2 \in \mathcal{E}(\mathfrak{A}) \text{ mit } g_1(\omega) \leq g_2(\omega) \text{ für alle } \omega \in \Omega$$

⟦ Dies folgt trivialerweise aus der Definition und (3.11.1) ⟧

3.12 Definition (Integral meßbarer Funktionen)

Es sei $f : \Omega \to \mathbb{R}_+ \cup \{\infty\}$ meßbar. Das Integral von f bezüglich μ ist definiert durch

$$\int f \,\mathbf{d}\mu := \int_\Omega f(\omega)\ \mathbf{d}\mu(\omega) := \sup_{\substack{g \in \mathcal{E}(\mathfrak{A}) \\ g \leq f}} \int g \,\mathbf{d}\mu.$$

3.13 Bemerkung

Ist $f \in \mathcal{E}(\mathfrak{A})$, so ergibt sich durch die doppelte Definition von $\int f \,\mathbf{d}\mu$ natürlich kein Widerspruch: Wegen (3.11.4) gilt für $f \in \mathcal{E}(\mathfrak{A})$

$$\sup_{\substack{g \in \mathcal{E}(\mathfrak{A}) \\ g \leq f}} \int g \,\mathbf{d}\mu \leq \int f \,\mathbf{d}\mu$$

und „=“ gilt, da $g = f$ im Supremum vorkommt. Es folgt $\int f_1 \,\mathbf{d}\mu \leq \int f_2 \,\mathbf{d}\mu$ für f_1, f_2 meßbar mit $f_1 \leq f_2$.

Die Definition (3.12) ist für die praktische Berechnung des Integrals meist unbrauchbar. Um eine praktikablere Berechnungsmethode im Falle $\mu = \lambda^d$ angeben zu können, benötigen wir:

3.14 Satz (*Beppo-Levi*; monotone Konvergenz)

Es seien $f_n : \Omega \to \mathbb{R}_+ \cup \{\infty\}$ meßbare Funktionen mit $f_n \le f_{n+1}$ sowie $f = \sup\limits_{n\in\mathbb{N}} f_n$. Dann gilt:

$$\lim_{n\to\infty} \int f_n \,\mathbf{d}\mu = \int f \,\mathbf{d}\mu.$$

Beweis:

„$\le$“ Es sei $n \in \mathbb{N}$ und $g \in \mathcal{E}(\mathfrak{A})$ mit $g \le f_n$. Dann gilt $g \le f$ und nach (3.13)

$$\int g \,\mathbf{d}\mu \le \int f \,\mathbf{d}\mu \quad \Rightarrow \quad \int g \,\mathbf{d}\mu \le \int f_n \,\mathbf{d}\mu \le \int f \,\mathbf{d}\mu \quad \Rightarrow \quad \lim_{n\to\infty} \int f_n \,\mathbf{d}\mu \le \int f \,\mathbf{d}\mu.$$

„$\ge$“

<u>1. Fall</u>: $\displaystyle\int f \,\mathbf{d}\mu < \infty$:

Sei $\varepsilon > 0$. Wir zeigen, daß es ein $n_0 \in \mathbb{N}$ mit $\displaystyle\int f_{n_0} \,\mathbf{d}\mu > \int f \,\mathbf{d}\mu - \varepsilon$ gibt

$$\Rightarrow \quad \forall n \ge n_0 \quad \int f_n \,\mathbf{d}\mu \ge \int f_{n_0} \,\mathbf{d}\mu > \int f \,\mathbf{d}\mu - \varepsilon \quad \Rightarrow \quad „\ge“.$$

Nach Definition von $\displaystyle\int f \,\mathbf{d}\mu$ existiert ein $g \in \mathcal{E}(\mathfrak{A})$, $g \le f$ mit

$$\int g \,\mathbf{d}\mu \ge \int f \,\mathbf{d}\mu - \frac{\varepsilon}{2}.$$

Da $\displaystyle\int g \,\mathbf{d}\mu \le \int f \,\mathbf{d}\mu < \infty$, ist nach Definition (3.10) $\mu\{\omega : g(\omega) > 0\} < \infty$. Wähle $\delta > 0$ mit $\delta\mu\{\omega : g(\omega) > 0\} \le \dfrac{\varepsilon}{3}$ und setze $h := g \cdot 1_{\{\omega : g(\omega) > \delta\}}$

$$\begin{aligned}
\Rightarrow \quad g(\omega) - h(\omega) &= g(\omega) - g(\omega) \cdot 1_{\{\omega : g(\omega) > \delta\}}(\omega) \\
&= \left\{ \begin{array}{cl} 0 & g(\omega) = 0 \\ g(\omega) & 0 < g(\omega) \le \delta \\ 0 & g(\omega) > \delta \end{array} \right\} \le \delta \cdot 1_{\{\omega : g(\omega) > 0\}}
\end{aligned}$$

$$\Rightarrow \quad \int (g - h) \,\mathbf{d}\mu \le \delta \cdot \mu\{\omega : g(\omega) > 0\} < \frac{\varepsilon}{3}$$

$$\Rightarrow \quad \int h \,\mathbf{d}\mu > \int g \,\mathbf{d}\mu - \frac{\varepsilon}{3}.$$

Sei nun $A_n = \{\omega \in \Omega \mid f_n(\omega) < h(\omega)\}$. Wegen $f_n(\omega) \uparrow f(\omega) \ge h(\omega)$ falls $f(\omega) > 0$ gilt $A_n \downarrow \emptyset$ und wegen der Stetigkeit von oben und

$$\mu(A_0) \le \mu\{\omega : h(\omega) > 0\} \le \mu\{\omega : g(\omega) > 0\} < \infty$$

folgt $\lim\limits_{n\to\infty} \mu(A_n) = 0$. Es sei $M := \max\limits_{\omega\in\Omega} h(\omega)$. Dann gibt es ein n_0 mit $\mu(A_{n_0}) < \dfrac{\varepsilon}{3M}$. Daraus folgt

$$\begin{aligned}
\int f_{n_0} \,\mathbf{d}\mu \ge \int f_{n_0} 1_{\Omega\setminus A_{n_0}} \,\mathbf{d}\mu \quad &\ge \quad \int h \underbrace{\left(1 - 1_{A_{n_0}}\right)}_{1_{\Omega\setminus A_{n_0}}} \mathbf{d}\mu = \int h \,\mathbf{d}\mu - \int h \cdot 1_{A_{n_0}} \,\mathbf{d}\mu \\
&> \quad \int g \,\mathbf{d}\mu - \frac{\varepsilon}{3} - \int M \cdot 1_{A_{n_0}} \,\mathbf{d}\mu > \int f \,\mathbf{d}\mu - \frac{2\varepsilon}{3} - M \cdot \frac{\varepsilon}{3M} \\
&= \quad \int f \,\mathbf{d}\mu - \varepsilon.
\end{aligned}$$

2. Fall: $\int f \, \mathbf{d}\mu = \infty$: Hier sind 2 Fälle möglich:

2.1. Fall: $\exists g \in \mathcal{E}(\mathfrak{A})$ mit $g \leq f$ und $\int g \, \mathbf{d}\mu = \infty \Rightarrow \exists \varepsilon > 0,\ A \in \mathfrak{A}$ mit $\varepsilon \cdot 1_A \leq f$ und $\mu(A) = \infty$. Sei $A_n = A \cap \left\{\omega : f_n(\omega) \geq \frac{\varepsilon}{2}\right\}$. Da $f_n \uparrow f$ folgt mit der Stetigkeit von unten $\mu(A_n) \to \infty$ und damit $\int f_n \, \mathbf{d}\mu \geq \int \frac{\varepsilon}{2} \cdot 1_{A_n} \, \mathbf{d}\mu = \frac{\varepsilon}{2} \cdot \mu(A_n) \to \infty$.

2.2. Fall: Für alle $c > 0 \, \exists g \in \mathcal{E}(\mathfrak{A})$ mit $c \leq \int g \, \mathbf{d}\mu < \infty$ und dann folgt wie im endlichen Fall $\int f \, \mathbf{d}\mu$ mit $A_n = \left\{\omega : f_n(\omega) < \frac{g(\omega)}{2}\right\}$, daß $\lim_{n\to\infty} \int f_n \, \mathbf{d}\mu > \frac{c}{2}$ für jedes $c > 0$. □

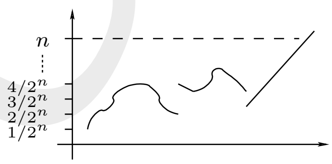
3.15 Lemma

Für jede meßbare Funktion $f \geq 0$ existiert eine Folge $f_n \in \mathcal{E}(\mathfrak{A})$ mit $f_n \leq f_{n+1} \to f$.

Beweis:

Setze

$$f_n(x) = \begin{cases} \frac{k}{2^n} & \text{falls } \frac{k}{2^n} \leq f(x) < \frac{k+1}{2^n} \leq n \\ n & \text{falls } f(x) \geq n. \end{cases}$$



□

3.16 Korollar (Eigenschaften des Integrals)

3.16.1

$$\int c \cdot f \, \mathbf{d}\mu = c \cdot \int f \, \mathbf{d}\mu \quad \text{für } c \geq 0 \text{ und } f \geq 0 \text{ meßbar.}$$

3.16.2

$$\int f_1 + f_2 \, \mathbf{d}\mu = \int f_1 \, \mathbf{d}\mu + \int f_2 \, \mathbf{d}\mu \quad \text{für } f_1, f_2 \geq 0 \text{ meßbar.}$$

Beweis:

Sind nämlich $g_n \uparrow f_1, h_n \uparrow f_2$ Folgen von Elementarfunktionen nach (3.15) so gilt $g_n + h_n \uparrow f_1 + f_2$ und mit *Beppo-Levi* (3.14)

$$\begin{aligned} \int f_1 + f_2 \, \mathbf{d}\mu &\underset{B.-L.}{=} \lim_{n\to\infty} \int g_n + h_n \, \mathbf{d}\mu = \lim_{n\to\infty} \int g_n \, \mathbf{d}\mu + \lim_{n\to\infty} \int h_n \, \mathbf{d}\mu \\ &\underset{B.-L.}{=} \int f_1 \, \mathbf{d}\mu + \int f_2 \, \mathbf{d}\mu \end{aligned}$$

□

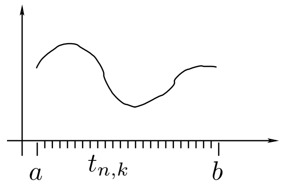
3.17 Satz

f sei eine stückweise stetige Funktion von $[a,b]$ nach $\mathbb{R}_+$. Dann gilt

$$\begin{aligned} \int 1_{[a,b]} \cdot f \, \mathbf{d}\lambda^1 &= \int 1_{]a,b]} \cdot f \, \mathbf{d}\lambda^1 = \int 1_{[a,b[} \cdot f \, \mathbf{d}\lambda^1 \\ &= \int 1_{]a,b[} \cdot f \, \mathbf{d}\lambda^1 = \int_a^b f(x) \, \mathbf{d}x, \end{aligned}$$

wobei letzteres als Riemann-Integral zu verstehen ist.

Beweis:
Die ersten 4 Integrale unterscheiden sich nur um $1_{\{a\}} \cdot f(a)$ (beziehungsweise $1_{\{b\}} \cdot f(b)$) und $\int 1_{\{a\}} \cdot f(a)\, \mathbf{d}\lambda^1 = f(a) \cdot \lambda^1 \{a\} = 0$.

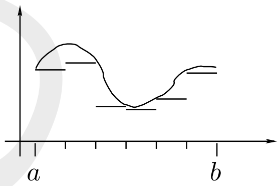


Es reicht ohne Einschränkung den Fall $f : [a,b] \to \mathbb{R}_+$ stetig zu betrachten (sonst Zerlegung in Teilintervalle). Für $n \geq 1$ sei $t_{n,k} = a + k \cdot \frac{b-a}{2^n}$ $(0 \leq k \leq 2^n)$ eine Zerlegung von $[a,b]$ und

$$\alpha_{n,k} = \min_{x \in [t_{n,k}, t_{n,k+1}]} f(x) \quad \text{für } 0 \leq k \leq 2^n - 1.$$

Dann ist $f_n = \sum_{k=0}^{2^n-1} \alpha_{n,k} \cdot 1_{]t_{n,k}, t_{n,k+1}]}$ eine Elementarfunktion mit

$$\int f_n \, \mathbf{d}\lambda^1 = \underbrace{\sum_{k=0}^{2^n-1} \alpha_{n,k} \cdot \frac{b-a}{2^n}}_{\text{Untersumme}} \xrightarrow{n\to\infty} \int_a^b f(x)\, \mathbf{d}x.$$



Andererseits gilt $f_n \leq f_{n+1}$ und da f stetig ist $f_n(x) \to f(x)$ für alle $x \in]a,b]$. Also ist nach *Beppo-Levi*

$$\int 1_{]a,b]} \cdot f \, \mathbf{d}\lambda^1 = \lim_{n\to\infty} \int f_n \, \mathbf{d}\lambda^1 = \int_a^b f(x)\, \mathbf{d}x.$$

3.18 Bemerkung

Ist $f : \mathbb{R} \to \mathbb{R}_+$ stückweise stetig, so gilt

$$\int f \, \mathbf{d}\lambda^1 = \int_{-\infty}^{\infty} f(x)\, \mathbf{d}x.$$

Wir sind nun in der Lage die angekündigten Maße mit Dichten zu definieren:

3.19 Satz

Es sei $f : \mathbb{R}^d \to \mathbb{R}_+ \cup \{\infty\}$ eine meßbare Funktion mit $\int f \, \mathbf{d}\lambda^d = 1$. Dann wird durch

$$\mathbf{P}(A) = \int 1_A \cdot f \, \mathbf{d}\lambda^d \quad A \in \mathfrak{B}\left(\mathbb{R}^d\right)$$

ein Wahrscheinlichkeitsmaß $\mathbf{P}$ auf $\mathbb{R}^d$ definiert.

Beweis:
Es ist

$$\mathbf{P}\left(\mathbb{R}^d\right) = \int 1_{\mathbb{R}^d} \cdot f \, \mathbf{d}\lambda^d = \int f \, \mathbf{d}\lambda^d = 1.$$

Die σ-Additivität folgt aus dem Satz von der monotonen Konvergenz: Ist $(A_n) \subset \mathfrak{B}\left(\mathbb{R}^d\right)$ eine disjunkte Folge und $A = \bigcup_{n\in\mathbb{N}} A_n$ so setzt man

$$f_m = 1_{\bigcup_{n=0}^m A_n} \cdot f = \sum_{n=0}^m 1_{A_n} \cdot f$$

$\Rightarrow \quad f_m \uparrow 1_A \cdot f.$

Dann folgt aus Satz (3.14)

$$\begin{aligned}\mathbf{P}\left(\bigcup_{n\in\mathbb{N}} A_n\right) = \mathbf{P}(A) &= \int 1_A \cdot f\,\mathbf{d}\lambda^d = \lim_{m\to\infty}\int f_m\,\mathbf{d}\lambda^d \\ &= \lim_{m\to\infty}\sum_{n=0}^{m}\int 1_{A_n}\cdot f\,\mathbf{d}\lambda^d = \lim_{m\to\infty}\sum_{n=0}^{m}\mathbf{P}(A_n) = \sum_{n=0}^{\infty}\mathbf{P}(A_n).\end{aligned}$$

$\square$

Die Funktion f nennen wir Lebesgue-Dichte von $\mathbf{P}$. Sämtliche Wahrscheinlichkeitsmaße die man so erhält sind diffus, d.h. für alle $\omega \in \mathbb{R}^d$ gilt $\mathbf{P}\{\omega\} = \int 1_{\{\omega\}}\cdot f\,\mathbf{d}\lambda^d = f(\omega)\cdot\lambda^d\{\omega\} = 0$.

3.20 Beispiele

(a) Es sei $\nu > 0$ und $f : \mathbb{R} \to \mathbb{R}_+$; $f(x) = \begin{cases} \nu\cdot\mathbf{e}^{-\nu x} & x \geq 0 \\ 0 & x < 0 \end{cases}$

$$\Rightarrow \quad \int f\,\mathbf{d}\lambda^d = \int_0^\infty \nu\cdot\mathbf{e}^{-\nu x}\,\mathbf{d}x = \left[-\mathbf{e}^{-\nu x}\right]_0^\infty = 1.$$

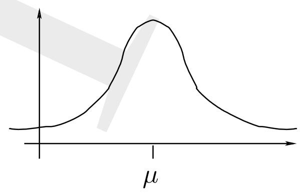
Das mit dieser Dichte definierte Wahrscheinlichkeitsmaß heiß Exponentialverteilung mit Parameter ν und wird mit $\mathcal{E}_\nu$ bezeichnet. Häufig wird $\mathcal{E}_\nu$ als Modell für die Lebensdauer eines Bauteils verwendet:
Für $T \geq 0$ ist $\mathcal{E}_\nu([T,\infty[) = \int_T^\infty \nu\cdot\mathbf{e}^{-\nu x}\,\mathbf{d}x = \mathbf{e}^{-\nu T}$ die Wahrscheinlichkeit, daß das Bauteil mindestens bis zum Zeitpunkt T funktioniert.

(b) Gauß'sche Normalverteilung mit Parametern $\mu \in \mathbb{R}$ und $\sigma^2 > 0$:
Dichte:

$$f(x) := \frac{1}{\sqrt{2\pi\sigma^2}}\mathbf{e}^{-\frac{(x-\mu)^2}{2\sigma^2}}$$

Bezeichnung:

$\mathcal{N}_{\mu,\sigma^2}$



(c) Leider läßt sich die Wahrscheinlichkeit von Intervallen nicht mehr so einfach bestimmen, da die Stammfunktion

$$\Phi(t) = \frac{1}{\sqrt{2\pi}}\int_{-\infty}^{t}\mathbf{e}^{-\frac{x^2}{2}}\,\mathbf{d}x$$

nicht in expliziter Form darstellbar ist. Es ist $\mathcal{N}_{0,1}(]-\infty,t]) = \Phi(t)$.
Eigenschaften:

(i) $\Phi(t) = 1 - \Phi(-t)$

(ii) $\mathcal{N}_{\mu,\sigma^2}([a,b]) = \Phi\left(\frac{b-\mu}{\sigma}\right) - \Phi\left(\frac{a-\mu}{\sigma}\right)$

Es reicht daher die Funktion Φ für Werte $t \geq 0$ zu tabellieren. Die Gauß'sche Normalverteilung ist eine der wichtigsten Verteilungen; sie tritt unter anderem bei der Modellierung von Meßfehlern auf.

(d) Cauchy-Verteilung
Es sei
$$f(x) = \frac{1}{\pi(1+x^2)}.$$
Dann gilt
$$\int f \,\mathbf{d}\lambda^1 = \int_{-\infty}^{\infty} \frac{\mathbf{d}x}{\pi(1+x^2)} = \left[\frac{1}{\pi}\arctan(x)\right]_{-\infty}^{\infty} = 1,$$
also ist f eine Lebesgue-Dichte.

Wahrscheinlichkeitsmaße auf $\mathbb{R}$ können auch durch den (älteren) Begriff der Verteilungsfunktion beschrieben werden:

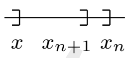
3.21 Definition ($F_{\mathbf{P}}(x)$)

Es sei $\mathbf{P}$ ein Wahrscheinlichkeitsmaß auf $(\mathbb{R}, \mathfrak{B}(\mathbb{R}))$. Dann heißt die Funktion
$$F_{\mathbf{P}}(x) := \mathbf{P}(]-\infty, x]),$$
$x \in \mathbb{R}$ die Verteilungsfunktion von $\mathbf{P}$.

3.22 Eigenschaften

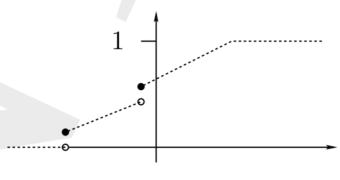
(a) $F_{\mathbf{P}}$ ist monoton wachsend und $\mathbf{P}(]a,b]) = F_{\mathbf{P}}(b) - F_{\mathbf{P}}(a)$ für $a < b$.

(b) $F_{\mathbf{P}}$ ist stetig von rechts
⟦
$$\begin{aligned} x_n \downarrow x \quad \Rightarrow \quad \lim_{n\to\infty} F_{\mathbf{P}}(x_n) &= \lim_{n\to\infty} \mathbf{P}(]-\infty, x_n]) = \mathbf{P}\left(\bigcap_{n\geq 1}]-\infty, x_n]\right) \\ &= \mathbf{P}(]-\infty, x]) = F_{\mathbf{P}}(x) \end{aligned}$$

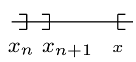


⟧

(c) $F_{\mathbf{P}}$ besitzt linksseitige Grenzwerte.



⟦
$$\begin{aligned} x_n \uparrow x \quad \Rightarrow \quad \lim_{n\to\infty} F_{\mathbf{P}}(x_n) &= \lim_{n\to\infty} \mathbf{P}(]-\infty, x_n]) \\ &= \mathbf{P}\left(\bigcup_{n\geq 1}]-\infty, x_n]\right) \\ &= \mathbf{P}(]-\infty, x[) \end{aligned}$$

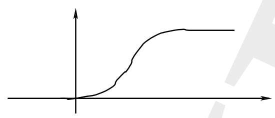


⟧

3.23 Beispiele

(a) Für $\nu > 0$ und $\mathcal{E}_\nu$ gilt

$$F_{\mathcal{E}_\nu} = \mathcal{E}_\nu\left(]-\infty, x]\right) = \int_{-\infty}^{x} 1_{[0,\infty[}(t) \cdot \nu \mathbf{e}^{-\nu t}\, \mathbf{d}t = \begin{cases} 0 & x \leq 0 \\ 1 - \mathbf{e}^{-\nu x} & x > 0 \end{cases}.$$



(b) Die Verteilungsfunktion von $\mathcal{N}_{\mu,\sigma^2}$ ist Φ; allgemein gilt für $\mathbf{P} := \mathcal{N}_{\mu,\sigma^2}$:

$$F_{\mathbf{P}}(x) = \Phi\left(\frac{x-\mu}{\sigma}\right).$$

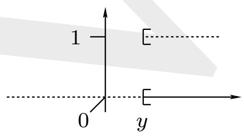
(c) Sei $y \in \mathbb{R}$. Ein Modell für das sichere Eintreten von y ist durch den Wahrscheinlichkeitsraum $(\mathbb{R}, \mathfrak{B}(\mathbb{R}), \mathcal{E}_y)$ gegeben, wobei

$$\mathcal{E}_y(A) = \begin{cases} 1 & y \in A \\ 0 & y \notin A \end{cases}.$$

Es gilt

$$F_{\mathcal{E}_y}(x) = \mathcal{E}_y\left(]-\infty, x]\right) = \begin{cases} 0 & x < y \\ 1 & x \geq y \end{cases}.$$

$\mathcal{E}_y$ heißt Dirac-Verteilung oder Punktmaß in y.



Um auch Maße auf $\mathbb{R}^d$ für $d \geq 2$ berechnen zu können benötigen wir ein maßtheoretisches Hilfsmittel:

3.24 Faktum (Satz von *Fubini*)

Für jede meßbare Funktion $f : \mathbb{R}^d \to \mathbb{R}_+$ gilt:

$$\int_{\mathbb{R}^d} f(x)\, \mathbf{d}\lambda^d = \int_{\mathbb{R}} \int_{\mathbb{R}^{d-1}} f(y)\, \mathbf{d}\lambda^{d-1}(y)\, \mathbf{d}\lambda^1(y).$$

3.25 Beispiele

(a) Zweidimensionale Normalverteilung mit Korrelation ρ.
Es sei $-1 < \rho < 1$ und $f : \mathbb{R}^2 \to \mathbb{R}_+$

$$f(x_1, x_2) = \frac{1}{2\pi\sqrt{1-\rho^2}} \exp\left(-\frac{1}{2(1-\rho^2)}\left(x_1^2 - 2\rho x_1 x_2 + x_2^2\right)\right).$$

Daß $\int f\, \mathbf{d}\lambda^2 = 1$ folgt aus dem Satz von Fubini:

$$\int f\, \mathbf{d}\lambda^2 = \int \int f(x_1, x_2)\, \mathbf{d}\lambda^1(x_1)\, \mathbf{d}\lambda^1(x_2).$$

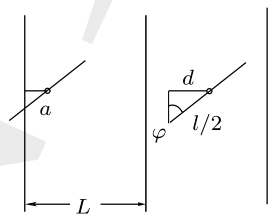
Nun ist

$$\begin{aligned}
& \int f(x_1, x_2)\, \mathbf{d}\lambda^1(x_1) \\
= \; & \frac{1}{2\pi\sqrt{1-\rho^2}} \int_{-\infty}^{\infty} \exp\left(-\frac{1}{2(1-\rho^2)}\left(x_1^2 - 2\rho x_1 x_2 + x_2^2\right)\right) \mathbf{d}x_1 \\
= \; & \frac{1}{2\pi\sqrt{1-\rho^2}} \int_{-\infty}^{\infty} \exp\left(-\frac{1}{2(1-\rho^2)}(x_1 - \rho x_2)^2 - \frac{1}{2(1-\rho^2)}\left(1-\rho^2\right) x_2^2\right) \mathbf{d}x_1 \\
= \; & \frac{1}{2\pi\sqrt{1-\rho^2}} \int_{-\infty}^{\infty} \exp\left(-\frac{1}{2(1-\rho^2)}(x_1 - \rho x_2)^2\right) \mathbf{d}x_1 \cdot \exp\left(\frac{-x_2^2}{2}\right) \\
= \; & \frac{1}{\sqrt{2\pi}} \cdot \frac{1}{\sqrt{2\pi(1-\rho^2)}} \int_{-\infty}^{\infty} \exp\left(-\frac{(x_1-\rho x_2)^2}{2(1-\rho^2)}\right) \mathbf{d}x_1 \cdot \exp\left(\frac{-x_2^2}{2}\right) \\
= \; & \frac{1}{\sqrt{2\pi}} \cdot \exp\left(\frac{-x_2^2}{2}\right) \cdot \mathcal{N}_{\rho x_2, 1-\rho^2}(\mathbb{R}) \\
= \; & \frac{1}{\sqrt{2\pi}} \cdot \exp\left(\frac{-x_2^2}{2}\right)
\end{aligned}$$

$$\begin{aligned}
\Rightarrow \quad \int f\, \mathbf{d}\lambda^2 \; &= \; \int\int f(x_1, x_2)\, \mathbf{d}\lambda^1(x_1)\, \mathbf{d}\lambda^1(x_2) \\
&= \; \int_{-\infty}^{\infty} \frac{1}{\sqrt{2\pi}} \exp\left(\frac{-x_2^2}{2}\right) \mathbf{d}x_2 \; = \; \mathcal{N}_{0,1}(\mathbb{R}) \; = \; 1.
\end{aligned}$$

Wir werden später ausrechnen, daß der Korrelationskoeffizient von $\mathbf{P}$ mit Dichte f tatsächlich ρ ist.

(b) Buffonsches Nadelproblem:
Eine Nadel der Länge l wird auf einen Boden geworfen, auf dem parallele Linien im Abstand $L \geq l$ gezeichnet sind. Wie groß ist die Wahrscheinlichkeit eine Linie zu treffen?
Wir beschreiben die Lage der Nadel durch den Winkel $\varphi \in [0, \pi[$ zwischen der Nadel und den Linien und den Abstand $a \in \left[0, \frac{L}{2}\right]$ zwischen dem Mittelpunkt der Nadel und der nächstgelegenen Linie.



Modell: $\Omega = \left[0, \frac{L}{2}\right] \times [0, \pi[\subset \mathbb{R}^2$ und $\mathbf{P}$ sei die Gleichverteilung auf Ω, d.h.

$$\mathbf{P}(A) = \frac{\lambda^2(A)}{\lambda^2(\Omega)} = \frac{2}{\pi \cdot L} \cdot \lambda^2(A).$$

Das Ereignis „Schnitt“ ist dann gerade

$$A := \left\{ (a, \varphi) \in \Omega \;\middle|\; a \leq \frac{l}{2} \cdot \sin(\varphi) \right\}$$

$$\begin{aligned}
\Rightarrow \quad \mathbf{P}(A) &= \frac{2}{\pi \cdot L} \cdot \lambda^2(A) = \frac{2}{\pi \cdot L} \cdot \int 1_A(x)\, \mathbf{d}\lambda^2(x) \\
&= \frac{2}{\pi \cdot L} \cdot \int \int 1_A(a,\varphi)\, \mathbf{d}\lambda^1(a)\, \mathbf{d}\lambda^1(\varphi) \\
&= \frac{2}{\pi \cdot L} \cdot \int_0^{\pi} \int_0^{\frac{l}{2}\sin(\varphi)} 1\, \mathbf{d}a\, \mathbf{d}\varphi = \frac{2}{\pi \cdot L} \cdot \int_0^{\pi} \frac{l}{2} \cdot \sin(\varphi)\, \mathbf{d}\varphi \\
&= \frac{2}{\pi \cdot L} \cdot \frac{l}{2} \cdot 2 = \frac{2 \cdot l}{\pi \cdot L}.
\end{aligned}$$

Dieses Ergebnis wurde um 1850 von dem Züricher Astronom *R. Wolf* benutzt um die Zahl π auf experimentellem Weg zu bestimmen. Er führte das Experiment 5000 mal aus und bestimmte die Näherung

$$\pi \approx 3,1596.$$

3.26 Beweis von Satz (2.7.a)

Das Ereignis

$$B_k = \bigcup_{\substack{J \subset \{1,\dots,n\} \\ \# J = k}} \bigcap_{j \in J} A_j \cap \bigcap_{j \notin J} A_j^{\complement} \quad \text{(disjunkte Vereinigung)}$$

hat die Indikatorfunktion

$$\begin{aligned}
1_{B_k} &= \sum_{\substack{K \subset \{1,\dots,n\} \\ \# K = k}} \prod_{j \in K} 1_{A_j} \cdot \prod_{j \notin K} \left(1 - 1_{A_j}\right) \\
&= \sum_{\substack{K \subset \{1,\dots,n\} \\ \# K = k}} \prod_{j \in K} 1_{A_j} \cdot \sum_{l=0}^{n-k} (-1)^l \sum_{\substack{L \subset \{1,\dots,n\} \setminus K \\ \# L = l}} \prod_{j \in L} 1_{A_j} \\
&= \sum_{l=0}^{n-k} (-1)^l \sum_{\substack{K \subset \{1,\dots,n\} \\ \# K = k}} \sum_{\substack{L \subset \{1,\dots,n\} \setminus K \\ \# L = l}} \prod_{j \in K \cup L} 1_{A_j}.
\end{aligned}$$

Es gibt $\binom{k+l}{k}$ Möglichkeiten, die Menge $J = K \cup L$ aus einer k- und einer l-elementigen Menge zusammenzusetzen. (Wähle aus $k + l$ Elementen k aus!)

$$\begin{aligned}
\Rightarrow \quad 1_{B_k} &= \sum_{l=0}^{n-k} (-1)^l \binom{k+l}{k} \sum_{\substack{J \subset \{1,\dots,n\} \\ \# J = k+l}} \prod_{j \in J} 1_{A_j} \\
&= \sum_{l=k}^{n} (-1)^{l-k} \binom{l}{k} \sum_{\substack{J \subset \{1,\dots,n\} \\ \# J = l}} 1_{\bigcap\limits_{j \in J} A_j}.
\end{aligned}$$

Integriert man diese Identität nach $\mathbf{P}$, so folgt aus den üblichen Eigenschaften (3.11) die Behauptung. $\square$

Aufgaben:

Aufgabe 3.1:

Es seien $\mathfrak{A}$ eine σ-Algebra auf Ω und $A_n \in \mathfrak{A}$ für $n \in \mathbb{N}$.
Beweisen Sie:

$$1_{\bigcup\limits_{n\in\mathbb{N}} A_n} = \sup_{n\in\mathbb{N}} 1_{A_n} \quad \text{und} \quad 1_{\bigcap\limits_{n\in\mathbb{N}} A_n} = \prod_{n\in\mathbb{N}} 1_{A_n} = \inf_{n\in\mathbb{N}} 1_{A_n}$$

$$1_{\limsup\limits_{n\in\mathbb{N}} A_n} = \limsup_{n\in\mathbb{N}} 1_{A_n} \quad \text{und} \quad 1_{\liminf\limits_{n\in\mathbb{N}} A_n} = \liminf_{n\in\mathbb{N}} 1_{A_n}$$

$$1_{\bigcup\limits_{n\in\mathbb{N}} A_n} = \sum_{n\in\mathbb{N}} 1_{A_n} \quad \text{falls die } A_n \text{ disjunkt sind}$$

$$1_{A_1 \setminus A_0} = 1_{A_1} - 1_{A_0} \quad \text{falls } A_0 \subset A_1$$

Aufgabe 3.2:

Es sei $\mathfrak{A}$ eine σ-Algebra auf Ω. Zeigen Sie:

(a) $f : \Omega \to \mathbb{R}$ meßbar $\Leftrightarrow$ $\{\omega \in \Omega : f(\omega) \leq \alpha\} \in \mathfrak{A}$ für alle $\alpha \in \mathbb{R}$.

(b) $f : \Omega \to \mathbb{R}$ meßbar $\Rightarrow$ f^2 meßbar.

(c) $f, g : \Omega \to \mathbb{R}$ meßbar $\Rightarrow$ $f \cdot g$ meßbar.
(*Hinweis*: Zeigen Sie zunächt, daß $(f+g)^2$ meßbar ist.)

Aufgabe 3.3:

Stellen Sie fest, ob es sich bei den angegebenen Funktionen $f : \mathbb{R} \to \mathbb{R}$ um Wahrscheinlichkeitsdichten handelt:

(a) $f(x) = \left(\frac{3}{2}x^2 + x\right) \cdot 1_{[-1,1]}(x)$

(b) $f(x) = \mathbf{e}^{-x^2+x} \cdot 1_{[0,1]}(x)$

(c) $f(x) = (1 - |x|) \cdot 1_{[-1,1]}(x)$

(d) $f_{\alpha,\beta}(x) = \alpha\beta x^{\beta-1}\mathbf{e}^{-\alpha x^\beta} \cdot 1_{]0,\infty[}(x) \qquad (\alpha, \beta > 0)$

(e) $f_\gamma(x) = \frac{1}{(\gamma-1)!}x^{\gamma-1}\mathbf{e}^{-x} \cdot 1_{]0,\infty[}(x) \qquad (\gamma \in \mathbb{N})$

Aufgabe 3.4:

Ein Detektor ist in der Lage, zwei Signale getrennt zu empfangen, wenn sie mindestens eine Zeit $t > 0$ auseinander liegen. Berechnen Sie die Wahrscheinlichkeit, daß zwei zufällig im Zeitintervall $[0, T]$ eintreffende Signale getrennt empfangen werden.

Aufgabe 3.5:

(a) Zeigen Sie, daß durch $\mu(A) = \#A$, $A \in \mathcal{POT}(\mathbb{N})$, ein Maß auf $(\mathbb{N}, \mathcal{POT}(\mathbb{N}))$ definiert wird.

(b) Zeigen Sie, daß jede Abbildung $f : \mathbb{N} \to \mathbb{R}_+$ meßbar (bezüglich $\mathcal{POT}(\mathbb{N})$) ist.

(c) Berechnen Sie $\int f \, \mathbf{d}\mu$.

Aufgabe 3.6:

(a) Zeigen Sie, daß für $\mu \in \mathbb{R}$ und $\sigma > 0$ durch

$$f(x) = \frac{1}{\sqrt{2\pi\sigma^2}} \exp\left(-\frac{(x-\mu)^2}{2\sigma^2} \right)$$

die Lebesgue-Dichte eines Wahrscheinlichkeitsmaßes definiert ist, das wir mit $\mathcal{N}_{\mu,\sigma^2}$ bezeichnen.
(*Hinweis*: Benutzen Sie, daß $\int\limits_{-\infty}^{\infty} \mathbf{e}^{-\frac{1}{2}u^2}\, \mathbf{d}u = \sqrt{2\pi}$ gilt.)

(b) Für $t \in \mathbb{R}$ sei $\Phi(t) = \mathcal{N}_{0,1}(]-\infty, t])$ die Verteilungsfunktion von $\mathcal{N}_{0,1}$. Zeigen Sie

(i) $\Phi(t) = 1 - \Phi(-t)$

(ii) $\mathcal{N}_{\mu,\sigma^2}([a,b]) = \Phi\left(\frac{b-\mu}{\sigma}\right) - \Phi\left(\frac{a-\mu}{\sigma}\right)$.

Aufgabe 3.7: Bertrand'sches Paradoxon

Es wird in einem Kreis mit Radius 1 zufällig eine Sehne gezogen. Wie groß ist die Wahrscheinlichkeit, daß die Sehne länger als die Seite des eingeschriebenen gleichseitigen Dreiecks, $\sqrt{3}$, ist? Dabei sollen folgende Modellbildungen des Begriffs einer *zufälligen Sehne* verwendet werden:

(a) Man wählt zufällig einen Punkt A in der Kreischreibe und zieht die Sehne mit Mittelpunkt A. (eine Sehne ist eindeutig durch ihren Mittelpunkt festgelegt.)

(b) Man wählt auf einem festen Durchmesser des Kreises einen zufälligen Punkt A und zeichnet die Sehne senkrecht auf diesem Durchmesser durch A.

(c) Ein Endpunkt B der Sehne wird auf der Kreislinie festgelegt. Im Mittelpunkt M des Kreises wird ein zweiter radius MC mit einem zufäligen, auf $]-\pi, \pi]$ gleichverteilten Winkel zu MB eingezeichnet und die Sehne BC gewählt.

Berechnen Sie in allen drei Modellen die Verteilungsfunktion der Länge der Sehne.

4. Kapitel: Zufallsvariable und Unabhängikeit

Bisher haben wir immer Wahrscheinlichkeiten von Ereignissen berechnent. Häufig interessiert man sich für Zahlen, die von einem Merkmal abhängen, wie z.B. die Augensumme beim Werfen mehrerer Würfel und nicht für die einzelnen Würfel insgesamt. Mathematisch wird dies durch den Begriff der Zufallsvariable beschrieben.

4.1 Definition (Zufallsvariable/ Zufallsvektor)

Es sei $(\Omega, \mathfrak{A}, \mathbf{P})$ ein Wahrscheinlichkeitsraum, Ω' eine weitere Menge und $\mathfrak{A}' \subset \mathcal{POT}(\Omega')$ eine σ-Algebra. Eine Abbildung

$$X : \Omega \to \Omega'$$

heißt Zufallsvariable (mit Werten in Ω')

$$:\Leftrightarrow \quad \forall A' \in \mathfrak{A}' : X^{-1}(A') = \{\omega \in \Omega \mid X(\omega) \in A'\} \in \mathfrak{A} = \{X \in A'\}.$$

Will man die zugehörigen σ-Algebren angeben sagt man X ist eine $\mathfrak{A} - \mathfrak{A}'$ Zufallsvariable (ZV) oder schreibt

$$X : (\Omega, \mathfrak{A}) \to (\Omega', \mathfrak{A}').$$

Im Fall $X : \Omega \to \mathbb{R}$, $\mathfrak{A}' = \mathfrak{B}(\mathbb{R})$ spricht man von reellen Zufallsvariablen und im Fall $X : \Omega \to \mathbb{R}^d$, $\mathfrak{A}' = \mathfrak{B}(\mathbb{R}^d)$ von einem Zufallsvektor.

4.2 Interpretation

Faßt man Ω als Menge von möglichen Zuständen eines Systems auf (die im allgemeinem nicht direkt beobachtet werden können), so kann man den Wert $X(\omega)$ als Beobachtung einer vom unbekannten zufälligen Zustand ω abhängenden Meßgröße ansehen.

4.3 Beispiele

(a) Würfeln mit 2 Würfeln und Beobachtung der Augensumme

$$\Omega = \{(\omega_1, \omega_2) \mid \omega_1, \omega_2 \in \{1, \dots, 6\}\} \qquad \mathbf{P} = \mathfrak{L}_\Omega \qquad \mathfrak{A} = \mathcal{POT}(\Omega)$$

$$X(\omega_1, \omega_2) = \omega_1 + \omega_2.$$

Man kann $\Omega' = \mathbb{N}$ oder $\{2, \dots, 12\}$ oder $\mathbb{R}$ wählen. Die Bedingung $\{X \in A'\} \in \mathfrak{A}$ ist trivialerweise erfüllt, da $\mathfrak{A} = \mathcal{POT}(\Omega)$.

(b) n-facher Münzwurf:
Modell:

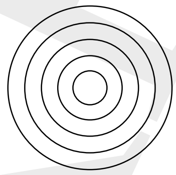
$$\Omega = \{0,1\}^n \quad (1 \mathrel{\hat{=}} \text{Zahl}, 0 \mathrel{\hat{=}} \text{Wappen}) \qquad \mathfrak{A} = \mathcal{POT}(\Omega) \qquad \mathbf{P} = \mathfrak{L}_\Omega$$

Für $j \in \{1, \dots, n\}$ sei $X_j : \Omega \to \{0,1\}$, $X_j(\omega_1, \dots, \omega_n) = \omega_j$ das Ergebnis des j-ten Versuchs und

$$Z(\omega_1, \dots, \omega_n) = \sum_{j=1}^{n} \omega_j$$

die Zahl der „1“. All dies sind Zufallsvariablen, da $\mathfrak{A} = \mathcal{POT}(\Omega)$.

(c) Treffer auf einer Zielscheibe:
Zielschreibe mit Radius 1 und n Ringen gleicher Breite konzentrisch aufgemalt. Wir interessieren uns für die Zufallsvarible X, die die Nummer des getroffenen Ringes angibt.



Modell:

$$\Omega = \left\{(x, y) \in \mathbb{R}^2 \,\middle|\, x^2 + y^2 < 1\right\}$$

$$R_j = \left\{(x, y) \in \mathbb{R}^2 \,\middle|\, \frac{j}{n} \leq \sqrt{x^2 + y^2} < \frac{j+1}{n}\right\}$$

für $0 \leq j \leq n - 1$.
Dann kann man $X(\omega) = j$ für $\omega \in R_j$ setzen, d.h.

$$X = \sum_{j=0}^{n-1} j \cdot 1_{R_j}.$$

Dies ist eine $\mathfrak{B}(\mathbb{R})-\mathcal{POT}(\{0, \ldots, n-1\})$ Zufallsvariable, denn für $A' \in \mathcal{POT}(\{0, \ldots, n-1\})$ ist

$$X^{-1}(A') = \bigcup_{j \in A'} R_j \in \mathfrak{B}(\Omega),$$

da

$$R_j = \left\{(x, y) \,\middle|\, \sqrt{x^2 + y^2} < \frac{j+1}{n}\right\} \setminus \left\{(x, y) \,\middle|\, \sqrt{x^2 + y^2} < \frac{j}{n}\right\} \in \mathfrak{B}(\Omega)$$

als Differenz offenen Mengen.

4.4 Eigenschaften

4.4.1 Lemma

Eine Abbildung $X : \Omega \to \mathbb{R}$ ist genau dann eine reelle Zufallsvariable, wenn

$$\{\omega \in \Omega : X(\omega) < \alpha\} \in \mathfrak{A}$$

für alle $\alpha \in \mathbb{R}$.

Beweis:

Ist X eine reelle Zufallsvariable, so folgt wegen $]-\infty, \alpha[\in \mathfrak{B}(\mathbb{R})$, daß $\{X < \alpha\} \in \mathfrak{A}$ für alle $\alpha \in \mathbb{R}$.
Sei umgekehrt

$$\{\omega \in \Omega : X(\omega) < \alpha\} = X^{-1}(]-\infty, \alpha[) \in \mathfrak{A}$$

für alle $x \in \mathbb{R}$.
Es sei

$$\mathcal{C} := \left\{C \in \mathfrak{B}(\mathbb{R}) \mid X^{-1}(C) \in \mathfrak{A}\right\}.$$

Dann ist $\mathcal{C}$ eine σ-Algebra, denn:
$\llbracket$

$$X^{-1}(\mathbb{R}) = \Omega \in \mathfrak{A} \quad \Rightarrow \quad \mathbb{R} \in \mathcal{C}$$

Für $C \in \mathcal{C}$ gilt wegen $X^{-1}\left(C^{\complement}\right) = \left(X^{-1}(C)\right)^{\complement} \in \mathfrak{A}$, da $X^{-1}(C) \in \mathfrak{A} \quad \Rightarrow \quad C^{\complement} \in \mathfrak{A}$ und schließlich:

$$\begin{aligned} \text{Ist } (C_n)_n \subset \mathcal{C} \quad &\Rightarrow \quad X^{-1}(C_n) \in \mathfrak{A} \quad \forall n \\ &\Rightarrow \quad \bigcup_{n \in \mathbb{N}} X^{-1}(C_n) = X^{-1}\left(\bigcup_{n \in \mathbb{N}} C_n\right) \in \mathfrak{A} \\ &\Rightarrow \quad \bigcup_{n \in \mathbb{R}} C_n \in \mathcal{C}. \end{aligned}$$

$\rrbracket$

Nach Voraussetzung enthält $\mathcal{C}$ die Intervalle $]-\infty, \alpha[$, $\alpha \in \mathbb{R}$. Da $\mathcal{C}$ eine σ-Algebra, folgt leicht, daß $\mathcal{C}$ alle Intervalle der Form $]a, b]$ mit $a, b \in \mathbb{R}$ enthält. Nun ist $\mathfrak{B}(\mathbb{R})$ die kleinste σ-Algebra, die $\{]a, b]\}$ enthält

$$\Rightarrow \quad \mathfrak{B}(\mathbb{R}) \subset \mathcal{C} \subset \mathfrak{B}(\mathbb{R})$$

$$\Rightarrow \quad \text{Für alle } A' \in \mathfrak{B}(\mathbb{R}) = \mathcal{C} \quad \Rightarrow \quad X^{-1}(A') \in \mathfrak{A},$$

d.h. X ist eine reelle Zufallsvariable. □

4.4.2

Sind X, Y reelle Zufallsvariable, so auch $X + Y$ (siehe (3.9.2)). Ebenso $X - Y$, $\min(X, Y)$, $\max(X, Y)$ usw.
Jede reelle Zahl $\alpha \in \mathbb{R}$ identifiziert man mit der deterministischen Zufallsvariable ($\hat{=}$ konstante Funktion) $\omega \mapsto \alpha$ von Ω nach $\mathbb{R}$.

4.4.3 Kompositionsregel

Ist $X : (\Omega, \mathfrak{A}) \to (\Omega', \mathfrak{A}')$ und $Y : (\Omega', \mathfrak{A}') \to (\Omega'', \mathfrak{A}'')$, so ist

$$Y \circ X : \begin{cases} (\Omega, \mathfrak{A}) \to (\Omega'', \mathfrak{A}'') \\ \omega \mapsto Y(X(\omega)) \end{cases}$$

⟦

$$\begin{aligned} A'' \in \mathfrak{A}'' \quad &\Rightarrow \quad A' = Y^{-1}(A'') \in \mathfrak{A}' \\ &\Rightarrow \quad X^{-1}(A') \in \mathfrak{A} \\ &\Rightarrow \quad \{\omega \in \Omega \mid Y \circ X(\omega) \in A''\} \in \mathfrak{A} \end{aligned}$$

⟧

4.4.4

Die Kompositionsregel (4.4.3) wird vor allem in der Situation angewandt, wo $Y := f : \mathbb{R} \to \mathbb{R}$ eine Funktion ist. Wir haben in (3.9.1) gesehen, daß jede stetige Funktion meßbar ist.

$$\Rightarrow \quad X : \Omega \to \mathbb{R} \text{ Zufallsvariable und } f : \mathbb{R} \to \mathbb{R} \text{ stetig}$$

$$\Rightarrow \quad f(X) : \Omega \to \mathbb{R} \text{ Zufallsvariable.}$$

Insbesondere sind X^m ($m \in \mathbb{N}$), $\mathbf{e}^X$, $\sin(X)$, $\frac{1}{X^2+1}$ usw. reelle Zufallsvariable.

4.5 Beispiel (Unendlicher Münzwurf)

Modell:

$$\Omega = \{0,1\}^{\mathbb{N}} \qquad \mathfrak{A} = (\mathcal{POT}(\{0,1\}))^{\otimes \mathbb{N}},$$

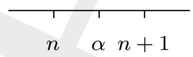
wobei diese σ-Algebra nach (1.15) die kleinste σ-Algebra auf Ω ist, zu der alle Ereignisse $Z_j(A) = \{(\omega_0, \omega_1, \ldots) \mid \omega_j \in A\}$ mit $j \in \mathbb{N}$, $A \in \mathcal{POT}(\{0,1\})$ („der j-te Wurf hat Eigenschaft A“) gehören.
Sei für $j \in \mathbb{N}$:

$$X_j : \Omega \to \{0,1\} \qquad X_j(\omega_0, \omega_1, \ldots) = \omega_j$$

(„Ergebnis des j-ten Wurfs“).
Wegen $X_j^{-1}(A) = Z_j(A) \in \mathfrak{A}$ für alle $A \in \mathcal{POT}(\{0,1\})$ ist $X_j : (\Omega, \mathfrak{A}) \to (\{0,1\}, \mathcal{POT}(\{0,1\}))$ eine Zufallsvariable. Wir interessieren uns für die Zeit, in der erstmals „1“ auftritt:

$$T(\omega) = \inf\{n \in \mathbb{N} \mid X_n(\omega) = 1\}.$$

$T : \Omega \to \mathbb{N} \cup \{\infty\}$, wobei $\inf\{\emptyset\} = +\infty$ gesetzt wird. Daß T eine Zufallsvariable ist sieht man mit Lemma (4.4.1). Für $\alpha > 0$ wähle $n \in \mathbb{N}$ mit $n < \alpha \leq n+1$



$$\begin{aligned} \Rightarrow \quad \{\omega \in \Omega \mid T(\omega < \alpha)\} &= \{\omega \in \Omega \mid T(\omega) \leq n\} \\ &= \{\omega \in \Omega \mid X_j(\omega) = 1 \text{ für mindestens ein } j \leq n\} \\ &= \bigcup_{j=0}^{n} \{\omega \in \Omega \mid X_j(\omega = 1)\} = \bigcup_{j=0}^{n} Z_j(\{j\}) \in \mathfrak{A}. \end{aligned}$$

4.6 Bezeichnung

Ist $X : (\Omega, \mathfrak{A}) \to (\Omega', \mathfrak{A}')$ eine Zufallsvariable und $A' \in \mathfrak{A}'$, so bezeichne:

$$\{X \in A'\} := \{\omega \in \Omega \mid X(\omega) \in A'\}$$

das Ereignis „der Wert von X liegt in A'“. Nach (4.1) ist $\{X \in A'\} \in \mathfrak{A}$, also tatsächlich ein Ereignis. Falls $\{X\} \in \mathfrak{A}'$, dann ist auch

$$\{X = x\} := \{\omega \in \Omega \mid X(\omega) = x\} \in \mathfrak{A}$$

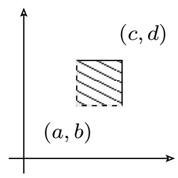
und analog ist $\{X \neq x\}$ zu definieren. Sei nun $\Omega' = \mathbb{R}$ und $\mathfrak{A}' = \mathfrak{B}(\mathbb{R})$ und $X, Y : \Omega \to \mathbb{R}$ reelle Zufallsvariable

$$\Rightarrow \quad (X, Y) : \omega \mapsto (X(\omega), Y(\omega)) \in \mathbb{R}^2$$

ist ein <u>Zufallsvektor</u>.

⟦ Für $a < c$ und $b < d$ gilt

$$\{(X, Y) \in](a, b), (c, d)]\} = \{X \in]a, c]\} \cap \{Y \in]b, d]\} \in \mathfrak{A}$$



und da $\mathfrak{B}(\mathbb{R}^2)$ nach Definition die kleinste σ-Algebra ist, die alle Rechtecke der Form $](a, b), (c, d)]$ enthält, folgt wie im Beweis von (4.4.4), daß $\{(X, Y) \in B\} \in \mathfrak{A}$ für alle $B \in \mathfrak{B}(\mathbb{R}^2)$. ⟧

$\Rightarrow \quad \{X = Y\} = \{\omega \in \Omega \mid X(\omega) = Y(\omega)\} \in \mathfrak{A}$, da $D = \{(x, x) \in \mathbb{R}^2 \mid x \in \mathbb{R}\}$ abgeschlossen, also $D \in \mathfrak{B}(\mathbb{R}^2)$ und $\{X = Y\} = \{(x, y) \in D\}$. Ebenso $\{X \neq Y\}$ und $\{X < Y\}, \ldots$

Die Werte die eine Zufallsvariable $X : \Omega \to \Omega'$ annimmt, treten im allgemeinen mit unterschiedlichen Wahrscheinlichkeiten auf, die wir durch eine Wahrscheinlichkeit auf $(\Omega', \mathfrak{A}')$ beschreiben können, der Verteilung von X. Dazu benötigen wir:

4.7 Satz

Es sei $X : (\Omega, \mathfrak{A}) \to (\Omega', \mathfrak{A}')$ eine Zufallsvariable auf einem Wahrscheinlichkeitsraum $(\Omega, \mathfrak{A}, \mathbf{P})$. Dann wird durch

$$\mathbf{P}_X : \mathfrak{A}' \to \mathbb{R}_+ \qquad \mathbf{P}_X(A') = \mathbf{P}\{X \in A'\} = \mathbf{P}\{\omega \in \Omega \mid X(\omega) \in A'\}$$

ein Wahrscheinlichkeitsmaß auf $(\Omega', \mathfrak{A}')$ definiert.

Beweis:

Trivialerweise gilt $\mathbf{P}_X(A') \geq 0$ und $\mathbf{P}_X(\Omega') = \mathbf{P}\{\omega \in \Omega \mid X(\omega) \in \Omega'\} = \mathbf{P}(\Omega) = 1$. Sei (A'_n) eine disjunkte Folge in $\mathfrak{A}'$

$$\Rightarrow \quad A_n := X^{-1}(A'_n) = \{\omega \in \Omega \mid X(\omega) \in A'_n\} \in \mathfrak{A}$$

ist eine disjunkte Folge in $\mathfrak{A}$.

$$\Rightarrow \quad \mathbf{P}_X\left(\bigcup_{k\in\mathbb{N}_0} A'_k\right) = \mathbf{P}\left\{X \in \bigcup_{k\in\mathbb{N}_0} A'_k\right\} = \mathbf{P}\left(\bigcup_{k\in\mathbb{N}_0}\{X\in A'_k\}\right)$$

$$= \mathbf{P}\left(\bigcup_{k\in\mathbb{N}_0} A_k\right) = \sum_{k=0}^{\infty}\mathbf{P}(A_k) = \sum_{k=0}^{\infty}\mathbf{P}_X(A'_k)$$

$\square$

4.8 Definition (Verteilung X unter $\mathbf{P}$)

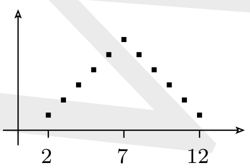
Das in Satz (4.7) eingeführte Wahrscheinlichkeitsmaß $\mathbf{P}_X$ auf $(\Omega', \mathfrak{A}')$ heißt <u>Verteilung von X unter $\mathbf{P}$</u>.

4.9 Beispiele

(a) Um die Verteilung der Augensumme in (4.3.a) zu beschreiben genügt es (siehe (1.23)), ihre Zähldichte anzugeben. Man rechnet leicht nach

$$\mathbf{P}_X\{k\} = \mathbf{P}\{X=k\} = \mathbf{P}\{(\omega_1,\omega_2)\in\Omega \mid \omega_1+\omega_2 = k\} = \begin{cases} \dfrac{k-1}{36} & \text{für } k \le 7 \\ \dfrac{13-k}{36} & \text{für } k \ge 7 \end{cases}$$

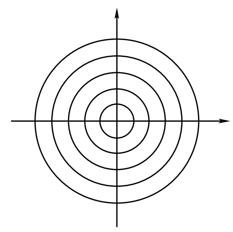
$\mathbf{P}_X$ heißt <u>diskrete Dreiecksverteilung</u> auf $(\{2,\dots,12\}, \mathcal{POT}(\{2,\dots,12\}))$.



(b) Die Zähldichte der Verteilung der Nummer des getroffenen Ringes in (4.3.c) ist im Fall von auf der ganzen Scheibe gleichverteilten Treffern:

$$\mathbf{P}_X\{j\} = \mathbf{P}\{X=j\} = \mathbf{P}\{R_j\} = \frac{\lambda^2(R_j)}{\lambda^2(\Omega)}$$

$$= \frac{1}{\pi}\left(\pi\left(\frac{j+1}{n}\right)^2 - \pi\left(\frac{j}{n}\right)^2\right) = \frac{2j+1}{n^2}$$



für $j = 0,\dots,n-1$.

(c) (vgl. (4.4.3)) Ist $X : (\Omega,\mathfrak{A}) \to (\Omega',\mathfrak{A}')$ und $Y : (\Omega',\mathfrak{A}') \to (\Omega'',\mathfrak{A}'')$ und $\mathbf{P}$ ein Wahrscheinlichkeitsmaß auf $(\Omega,\mathfrak{A})$, so ist die Verteilung von $Y\circ X$ das Wahrscheinlichkeitsmaß $(\mathbf{P}_X)_Y$ auf Ω''.

(d) $X : \Omega \to]0,1]$ mit $\mathbf{P}_X = \lambda^1\big|_{]0,1]}$ und $Y := -\ln(X) \quad \Rightarrow \quad \mathbf{P}_Y = \mathcal{E}_1$ (Exponentialverteilung).
$\llbracket$ Sei $A \in \mathfrak{A}(\mathbb{R}_+)$ und $B := \{\mathbf{e}^{-a} \mid a \in A\}$

$$\Rightarrow \quad \mathbf{P}_Y(A) = \mathbf{P}\{-\ln\circ X \in A\} = \mathbf{P}\{X\in B\} = \mathbf{P}_X(B) = \lambda^1\big|_{]0,1]}(B)$$

$$= \int 1_B(x)\,\mathbf{d}\lambda^1(x) = \int_{-\infty}^{\infty} 1_B(x)\,\mathbf{d}x$$

$$\overset{x=\mathbf{e}^{-u}}{=} \int_0^{\infty} 1_B\left(\mathbf{e}^{-u}\right)\cdot\mathbf{e}^{-u}\,\mathbf{d}u = \int_0^{\infty} 1_A(u)\cdot\mathbf{e}^{-u}\,\mathbf{d}u = \mathcal{E}_1(A)$$

Andere Möglichkeit: Über Verteilungsfunktionen:

$$\begin{aligned} F_{\mathbf{P}_Y}(y) &= \mathbf{P}\{-\ln(X) \leq y\} = \mathbf{P}\left\{X \geq \mathbf{e}^{-y}\right\} = \lambda^1\left(\left[\mathbf{e}^{-y}, 1\right]\right) \\ &= \begin{cases} 1 - \mathbf{e}^{-y} & \text{für } y \geq 0 \\ 0 & \text{für } y < 0 \end{cases} \end{aligned}$$

und dies ist (vgl. (3.23.a) $= F_{\mathcal{E}_1}(y)$ $\rrbracket$

(e) Es sei $X : \Omega \to \mathbb{R}$ eine Zufallsvariable mit $\mathbf{P}_X = \mathcal{N}_{0,1} \quad \Rightarrow \quad X^2$ besitzt eine $\mathcal{X}_1^2$-Verteilung mit einem Freiheitsgrad, d.h. $\mathbf{P}_{X^2}$ hat die Lebesgue -Dichte:

$$x \mapsto \frac{1}{\sqrt{2\pi}} \cdot \frac{1}{\sqrt{x}} \cdot \mathbf{e}^{-x/2} \cdot 1_{]0,\infty[}(x).$$

$\llbracket$ Für $A \in \mathfrak{B}(\mathbb{R})$ sei $B := \left\{a \in \mathbb{R} \mid a^2 \in A\right\} \in \mathfrak{B}(\mathbb{R})$,

$$\begin{aligned} \Rightarrow \quad \mathbf{P}_{X^2}(A) &= \mathbf{P}\left\{X^2 \in A\right\} = \mathbf{P}\{X \in B\} = \frac{1}{\sqrt{2\pi}} \int 1_B(x) \cdot \mathbf{e}^{-x^2/2}\, \mathbf{d}x \\ &= \frac{2}{\sqrt{2\pi}} \int 1_B(x) \cdot 1_{]0,\infty[}(x) \cdot \mathbf{e}^{-x^2/2}\, \mathbf{d}\lambda^1(x) \\ &\overset{y=x^2}{=} \frac{2}{\sqrt{2\pi}} \int 1_B(\sqrt{y}) \cdot 1_{]0,\infty[}(y) \cdot \mathbf{e}^{-y/2} \cdot \frac{1}{2\sqrt{y}}\, \mathbf{d}\lambda^1(y) \\ &= \frac{1}{\sqrt{2\pi}} \int 1_A(y) \cdot 1_{]0,\infty[}(y) \cdot \frac{1}{\sqrt{y}} \cdot \mathbf{e}^{-y/2}\, \mathbf{d}\lambda^1(y) \\ &= \mathcal{X}_1^2(A) \end{aligned}$$

$\rrbracket$

Eine sehr wichtige Annahme für die Modellierung ist die Unabhängigkeit zwischen Teilen des Experiments, z.B. zwischen den einzelnen Würfen einer Münze. Mathematisch wird dies durch folgende Definition modelliert:

4.10 Definition (Stochastische Unabhängigkeit)

Es sei $(\Omega, \mathfrak{A}, \mathbf{P})$ ein Wahrscheinlichkeitsraum, $I \neq \emptyset$ eine Indexmenge und $A_j \in \mathfrak{A}$ für alle $j \in I$. Die Ereignisse $(A_j)_{j \in I}$ heißen <u>(stochastisch) unabhängig</u>, falls

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

für alle endlichen $J \subset I$.

4.11 Beispiele

(a) Einmal Würfeln mit fairem Würfel:
<u>Modell</u>:

$$\Omega = \{1, \ldots, 6\} \qquad \mathfrak{A} = \mathcal{POT}(\Omega) \qquad \mathbf{P} = \mathfrak{L}_\Omega$$

$$\begin{aligned} A_1 &\hat{=} \text{„gerade Zahl“} = \{2, 4, 6\} \\ A_2 &\hat{=} \text{„mindestens 5“} = \{5, 6\} \end{aligned}$$

Dann sind A_1 und A_2 unabhängig, da

$$\mathbf{P}(A_1 \cap A_2) = \mathbf{P}\{6\} = \frac{1}{6} \quad \text{und} \quad \mathbf{P}(A_1) = \frac{1}{2}, \mathbf{P}(A_2) = \frac{1}{3} \quad \Rightarrow \quad \mathbf{P}(A_1) \cdot \mathbf{P}(A_2) = \frac{1}{6}.$$

(b) Wenn I drei oder mehr Elemente hat, genügt es nicht die „paarweise Unabhängigkeit" $\mathbf{P}(A_j \cap A_k) = \mathbf{P}(A_j) \cdot \mathbf{P}(A_k)$ für alle $j, k \in I$ zu zeigen um die Unabhängigkeit von $(A_j)_{j \in I}$ nachzuweisen:

$$\Omega = \{0,1,2,3\} \qquad \mathbf{P} = \mathfrak{L}_\Omega \qquad A_1 = \{0,1\}, \, A_2 = \{0,2\}, \, A_3 = \{1,2\}$$

$$\Rightarrow \quad \mathbf{P}(A_1 \cap A_2) = \mathbf{P}(A_1 \cap A_3) = \mathbf{P}(A_2 \cap A_3) = \frac{1}{4}$$

$$\text{und} \quad \mathbf{P}(A_i) = \frac{1}{2} \quad \Rightarrow \quad \mathbf{P}(A_i) \cdot \mathbf{P}(A_j) = \frac{1}{2} \cdot \frac{1}{2} = \frac{1}{4} \quad i, j \in \{1,2,3\}$$

d.h. A_1, A_2, A_3 sind paarweise unabhängig, aber

$$\mathbf{P}(A_1 \cap A_2 \cap A_3) = \mathbf{P}(\emptyset) = 0 \neq \mathbf{P}(A_1) \cdot \mathbf{P}(A_2) \cdot \mathbf{P}(A_3) = \frac{1}{2} \cdot \frac{1}{2} \cdot \frac{1}{2} = \frac{1}{8}.$$

Auch für Zufallsvariablen läßt sich der Begriff der Unabhängigkeit definieren und zwar durch Zurückführung auf Ereignisse:

4.12 Definition (Unabhängigkeit von Zufallsvariablen)

Es seien für $j \in I$ $X_j : (\Omega, \mathfrak{A}) \to (\Omega_j, \mathfrak{A}_j)$ Zufallsvariablen auf einem Wahrscheinlichkeitsraum $(\Omega, \mathfrak{A}, \mathbf{P})$. $(X_j)_{j \in I}$ heißen <u>unabhängig</u>

$$:\Leftrightarrow \quad \forall A_j \in \mathfrak{A}_j \text{ sind die Ereignisse } (\{X_j \in A_j\})_{j \in I} \text{ unabhängig.}$$

4.13 Bemerkung

(a) $(X_j)_{j \in I}$ unabhängig

$$\Leftrightarrow \quad \forall J \subset I \text{ endlich } \forall A_j \in \mathfrak{A}_j \ (j \in J) : \mathbf{P}\{X_j \in A_j \text{ für alle } j \in J\} = \prod_{j \in J} \mathbf{P}\{X_j \in A_j\}.$$

(b) Ist I eine endliche Menge, so genügt es

$$\mathbf{P}\{X_j \in A_j \text{ für alle } j \in I\} = \prod_{j \in I} \mathbf{P}\{X_j \in A_j\}$$

für alle $A_j \in \mathfrak{A}_j$ zu zeigen, da man $A_j = \Omega_j$ für $j \notin J$ wählen kann.

Im Falle von abzählbaren Räumen Ω_j, $\mathfrak{A}_j = \mathcal{POT}(\Omega_j)$ kann man dies einfacher mit der Zähldichte nachprufen:

4.14 Satz

Es seien für $j \in I$ Ω_j abzählbar und $\mathfrak{A}_j = \mathcal{POT}(\Omega_j)$. Die Zufallsvariable $X_j : \Omega \to \Omega_j$ sind genau dann unabhängig, falls für alle $J \subset I$ endlich

$$\mathbf{P}\{X_j = \omega_j \text{ für alle } j \in J\} = \prod_{j \in J} \mathbf{P}\{X_j = \omega_j\}$$

für alle $\omega_j \in \Omega_j$ $(j \in J)$ gilt. Ist I endlich, reicht es dies für $J = I$ zu zeigen.

Beweis:

„$\Rightarrow$": Klar, da man $A_j = \{\omega_j\} \in \mathfrak{A}_j$ wählen kann.

„ $\Leftarrow$ “ : Seien $A_j \in \mathfrak{A}_j$ für $j \in J \subset I$. Dann gilt:

$$\begin{aligned}
\mathbf{P}\{X_j \in A_j \text{ für alle } j \in J\} &= \sum_{j \in J} \mathbf{P}\{X_j \in A_j \text{ für } j \in J\} \\
&= \sum_{\substack{\omega_j \in A_j \\ j \in J}} \mathbf{P}\{X_j = \omega_j \text{ für } j \in J\} \\
&= \sum_{\substack{\omega_j \in A_j \\ j \in J}} \prod_{j \in J} \mathbf{P}\{X_j = \omega_j\} = \prod_{j \in J} \sum_{\substack{\omega_j \in A_j \\ j \in J}} \mathbf{P}\{X_j = \omega_j\} \\
&= \prod_{j \in J} \mathbf{P}\{X_j \in A_j\}
\end{aligned}$$

nach dem Distributivgesetz. $\square$

4.15 Beispiele

(a) (vgl. Bsp. (4.3.b)) (n-facher Münzwurf)

$$\Omega = \{0,1\}^n \qquad \mathfrak{A} = \mathcal{POT}(\Omega) \qquad \mathbf{P} = \mathfrak{L}_\Omega$$

Die Zufallsvariable $X_j : \Omega \to \{0,1\}$, $X_j(\omega_1, \ldots, \omega_n) = \omega_j$ (Ergebnis des j-ten Wurfs) sind unabhängig, denn:

$$\mathbf{P}\{X_j = x_j\} = 2^{-n} \cdot \#\{(\omega_1, \ldots, \omega_n) \in \Omega \mid \omega_j = x_j\} = 2^{-n} \cdot 2^{n-1} = \frac{1}{2}$$

und folglich

$$\mathbf{P}\{X_j = x_j \text{ für } 1 \leq j \leq n\} = 2^{-n} = \prod_{j=1}^{n} \mathbf{P}\{X_j = x_j\}.$$

(b) Allgemein gilt: Sind $\Omega_1, \ldots, \Omega_n$ endliche Mengen, $\Omega = \Omega_1 \times \ldots \times \Omega_n$, $X_j : \Omega \to \Omega_j$ die Projektion $X_j(\omega_1, \ldots, \omega_n) := \omega_j$ und $\mathbf{P} = \mathfrak{L}_\Omega$, $\mathfrak{A} = \mathcal{POT}(\Omega)$, so sind die Zufallsvariablen $X_1, \ldots, X_n$ unabhängig und haben die Verteilung $\mathbf{P}_{X_j} = \mathfrak{L}_{\Omega_j}$.

$[\![$

$$\begin{aligned}
\mathbf{P}\{X_j = x_j\} &= \frac{\#\{(\omega_1, \ldots, \omega_n) \mid \omega_j = x_j\}}{\#\Omega} \\
&= \frac{\#\Omega_1 \cdot \ldots \cdot \#\Omega_{j-1} \cdot 1 \cdot \#\Omega_{j+1} \cdot \ldots \cdot \#\Omega_n}{\#\Omega} \\
&= \frac{1}{\#\Omega_j} \quad \Rightarrow \quad \mathbf{P}_{X_j} = \mathfrak{L}_{\Omega_l}
\end{aligned}$$

und weiterhin:

$$\mathbf{P}\{X_j = x_j \text{ für } 1 \leq j \leq n\} = \frac{1}{\#\Omega} = \prod_{j=1}^{n} \frac{1}{\#\Omega_j} = \prod_{j=1}^{n} \mathbf{P}\{X_j = x_j\}$$

$]\!]$

(c) Sind $(X_j)_{j \in I}$ unabhängige Zufallsvariable, $X_j : \Omega \to \Omega_j$ und sind $Y_j : (\Omega_j, \mathfrak{A}_j) \to (\Omega'_j, \mathfrak{A}'_j)$ weitere Zufallsvarible, so sind auch die Zufallsvariablen $(Y_j \circ X_j)_{j \in I}$ unabhängig.

$[\![$

$$\{Y_j \circ X_j \in A'_j\} = \{X_j \in \underbrace{\{Y_j \in A_j\}}_{\in \mathfrak{A}_j}\}$$

$]\!]$

(d) Es sei $\Omega = [0,1[$, $\mathfrak{A} = \mathfrak{B}(\Omega)$ und für $n \geq 1$, $\omega \in [0,1[$ sei $X_n(\omega)$ die n-te Stelle in der Binärdarstellung von ω (eindeutig durch Ausschluß von 1!)
Dann sind die $(X_n)_{n\in\mathbb{N}}$ unabhängig und für $\mathbf{P} = \lambda^1\big|_\Omega$ gilt $\mathbf{P}_{X_n} = \mathfrak{L}_{\{0,1\}}$.
⟦ Es sei für $1 \leq j \leq n$ $x_j \in \{0,1\}$. Dann ist

$$\begin{aligned} & \{\omega \in [0,1[\mid X_j(\omega) = x_j \text{ für } 1 \leq j \leq n\} \\ = \; & \{\omega \in [0,1[\mid \text{die ersten } n\text{-Stellen nach dem Komma von } \omega \text{ sind } x_1, \ldots, x_n\} \\ = \; & \left[\sum_{j=1}^n x_j 2^{-j}, \sum_{j=1}^n x_j 2^{-j} + 2^{-n}\right[\\ & \qquad\qquad \Rightarrow \quad \mathbf{P}\{X_j = x_j \text{ für } 1 \leq j \leq n\} = 2^{-n} \end{aligned}$$

Sei nun $J \subset \mathbb{N}$ endlich und $n = \max(J)$.

$$\begin{aligned} \Rightarrow \quad \mathbf{P}\{X_j = x_j \text{ für } j \in J\} &= \sum_{\substack{x_j \in \{0,1\} \\ j \in \{1,\ldots,n\}\setminus J}} \mathbf{P}\{X_j = x_j \text{ für } 1 \leq j \leq n\} \\ &= 2^{n-\#J} \cdot 2^{-n} = 2^{-\#J} \\ &= \prod_{j\in J} 2^{-1} = \prod_{j\in J} \mathbf{P}\{X_j = x_j\} \end{aligned}$$

⟧

Zur Konstruktion eines Wahrscheinlichkeitsraumes $(\Omega, \mathfrak{A}, \mathbf{P})$ und unabhängiger Zufallsvariable $X_j : \Omega \to \Omega_j$ mit vorgegebener Verteilung $\mathbf{P}_{X_j} = \mathbf{P}_j$ beschränken wir uns hier auf endlich viele Zufallsvariable mit diskreter Verteilung beziehungsweise einer Verteilung mit Dichte.

4.16 Satz

Sind $(\Omega_j, \mathcal{POT}(\Omega_j), \mathbf{P}_j)$ abzählbare Wahrscheinlichkeitsräume $(1 \leq j \leq n)$, so existieren ein abzählbarer Wahrscheinlichkeitsraum $(\Omega, \mathcal{POT}(\Omega), \mathbf{P})$ und Zufallsvariable $X_j : \Omega \to \Omega_j$, die unabhängig sind und die Verteilung $\mathbf{P}_{X_j} = \mathbf{P}_j$ haben.

Beweis:
Sei $\Omega := \Omega_1 \times \ldots \times \Omega_n$, $X_j : \Omega \to \Omega_j$ die Projektion $X_j(\omega_1, \ldots, \omega_n) := \omega_j$. Die Funktion $(\omega_1, \ldots, \omega_n) \mapsto \mathbf{P}_1(\{\omega_1\}) \cdot \ldots \cdot \mathbf{P}_n(\{\omega_n\})$ ist eine Zähldichte, denn

$$\sum_{(\omega_1,\ldots\omega_n)\in\Omega} \mathbf{P}_1(\{\omega_1\}) \cdot \ldots \cdot \mathbf{P}_n(\{\omega_n\}) = \prod_{1\leq j\leq n} \sum_{\omega_j\in\Omega_j} \mathbf{P}_j(\{\omega_j\}) = 1^n = 1$$

Sei $\mathbf{P}$ das zugehörige Wahrscheinlichkeitsmaß auf Ω nach Satz (1.23).

$$\begin{aligned} \Rightarrow \quad \mathbf{P}\{X_j = \omega_j\} &= \mathbf{P}\{(\omega_1, \ldots, \omega_n) \in \Omega \mid x_j = \omega_j\} \\ &= \left(\underbrace{\sum_{x_1\in\Omega_1} \cdots \sum_{x_n\in\Omega_n}}_{\Omega_j \text{ fehlt}} \prod_{\substack{k=1 \\ k\neq j}}^n \mathbf{P}_k\{x_k\}\right) \cdot \mathbf{P}_j\{\omega_j\} \\ &= \left(\prod_{\substack{k=1 \\ k\neq j}}^n \sum_{x_k\in\Omega_k} \mathbf{P}_k\{x_k\}\right) \cdot \mathbf{P}_j\{\omega_j\} \\ &= 1^{n-1} \cdot \mathbf{P}_j\{\omega_j\} = \mathbf{P}_j\{\omega_j\} \end{aligned}$$

d.h. $\mathbf{P}_{X_j} = \mathbf{P}_j$. Weiter gilt für alle $(\omega_1, \ldots, \omega_n) \in \Omega$

$$\mathbf{P}\{X_j = \omega_j \text{ für } 1 \leq j \leq n\} = \mathbf{P}\{(\omega_1, \ldots, \omega_n)\} = \prod_{j=1}^n \mathbf{P}_j\{x_j\} = \prod_{j=1}^n \mathbf{P}\{X_j = \omega_j\}$$

und damit folgt aus Satz (4.14) die Unabhängigkeit. □

4.17 Satz

Es seien Wahrscheinlichkeitsmaße $\mathbf{P}_j$ auf $\Omega_j = \mathbb{R}$ mit Lebesgue-Dichten $f_j : \mathbb{R} \to \mathbb{R}_+$ vorgegeben $(1 \leq j \leq n)$. Dann ist durch die Lebesgue-Dichte

$$f : (x_1, \ldots, x_n) \mapsto f_1(x_1) \cdot \ldots \cdot f_n(x_n)$$

auf $\Omega := \mathbb{R}^n$ ein Wahrscheinllichkeitsmaß $\mathbf{P}$ auf $(\mathbb{R}^n, \mathfrak{B}(\mathbb{R}^n))$ definiert, für das die Zufallsvariablen $X_j : \Omega \to \mathbb{R}$ mit $X_j(\omega_1, \ldots, \omega_n) := \omega_j$ unabhängig sind und $\mathbf{P}_{X_j} = \mathbf{P}_j$ gilt.

Beweis:

Seien $A_j \in \mathfrak{B}(\mathbb{R})$ für $1 \leq j \leq n$.

$$\begin{aligned}
\Rightarrow \quad & \int 1_{A_1 \times \ldots \times A_n}(x) \cdot f_1(x_1) \cdot \ldots \cdot f_n(x_n)\, \mathbf{d}\lambda^n(x) \\
= \quad & \int \cdots \int 1_{A_1}(x_1) \times \ldots \times 1_{A_n}(x_n) \cdot f_1(x_1) \cdot \ldots \cdot f_n(x_n)\, \mathbf{d}\lambda^1(x_n) \cdot \ldots \cdot \mathbf{d}\lambda^1(x_1) \\
= \quad & \int 1_{A_1}(x_1) \cdot f_1(x_1)\, \mathbf{d}\lambda^1(x_1) \cdot \ldots \cdot \int 1_{A_n}(x_n) \cdot f_n(x_n)\, \mathbf{d}\lambda^1(x_n) \\
= \quad & \mathbf{P}_1(A_1) \cdot \ldots \cdot \mathbf{P}_n(A_n) \quad \text{nach dem Satz von Fubini.}
\end{aligned}$$

Setzen wir $A_j = \mathbb{R}$ für $1 \leq j \leq n$, so folgt $\int f\, \mathbf{d}\lambda^n = 1$, also ist f die Dichte eines Wahrscheinlichkeitsmaßes $\mathbf{P}$ auf $\mathbb{R}^n$. Nach obiger Identität folgt

$$\begin{aligned}
\mathbf{P}\{X_j \in A_j \text{ für } 1 \leq j \leq n\} &= \mathbf{P}(\{(x_1, \ldots, x_n) \in \mathbb{R}^n \mid x_j \in A_j\}) \\
&= \mathbf{P}(A_1 \times \ldots \times A_n) \\
&= \int 1_{A_1 \times \ldots \times A_n}(x) \cdot f(x)\, \mathbf{d}\lambda^n(x) \\
&= \mathbf{P}_1(A_1) \cdot \ldots \cdot \mathbf{P}_n(A_n) \\
&= \prod_{j=1}^n \mathbf{P}_j(A_j).
\end{aligned}$$

Sei nun $k \leq n$ fest, $A \in \mathfrak{B}(\mathbb{R})$, $A_j = \mathbb{R}$ für $j \neq k$, $A_k = A$

$$\begin{aligned}
\Rightarrow \quad \mathbf{P}\{X_k \in A\} &= \mathbf{P}_{X_k}(A) \\
&= \mathbf{P}\{X_j \in A_j \text{ für } 1 \leq j \leq n\} \\
&= \mathbf{P}_k(A)
\end{aligned}$$

$$\Rightarrow \quad \mathbf{P}_{X_k} = \mathbf{P}_k$$

Für $A_j \in \mathfrak{B}(\mathbb{R})$ beliebig folgt dann

$$\mathbf{P}\{X_j \in A_j \text{ für } 1 \leq j \leq n\} = \prod_{j=1}^n \mathbf{P}_j(A_j) = \prod_{j=1}^n \mathbf{P}\{X_j = A_j\}$$

d.h. die $(X_j)_{j=1,\ldots,n}$ sind unabhängig □

4.18 Definition (Produktmaß)

Das in Satz (4.16) und (4.17) konstruierte Wahrscheinlichkeitsmaß auf $\Omega_1 \times \ldots \times \Omega_n$ heißt Produktmaß und wird mit $\mathbf{P}_1 \otimes \ldots \otimes \mathbf{P}_n$ bezeichnet.

4.19 Anwendungen

(a) Sind $M_1, \ldots, M_n \in \mathfrak{B}(\mathbb{R})$ mit $0 < \lambda^1(M_j)$, so hat die Gleichverteilung auf $\Omega = M_1 \times \ldots \times M_n$ die Eigenschaft, daß die Projektionen $X_j : \Omega \to M_j$ unabhängig sind und die Verteilung von X_j die Gleichverteilung auf M_j ist.

(b) Ein Experiment habe zwei mögliche Ausgänge $1 \mathrel{\hat{=}}$ Erfolg, $0 \mathrel{\hat{=}}$ Mißerfolg die mit Wahrscheinlichkeit p beziehungsweise $1-p$ $(0 \leq p \leq 1)$ auftreten. Es werde n-mal unabhängig wiederholt.
<u>Modell</u>:

$$\Omega = \{0,1\}^n \qquad \mathfrak{A} = \mathcal{POT}(\Omega) \qquad \mathbf{P} = \mathcal{B}_{1,p} \otimes \ldots \otimes \mathcal{B}_{1,p}$$

wobei $\mathcal{B}_{1,p}$ die Bernoulli-Verteilung auf $\{0,1\}$ ist. Die Zähldichte von $\mathbf{P}$ ist gegeben durch

$$\Omega \ni (\omega_1, \ldots, \omega_n) \mapsto \prod_{j=1}^{n} \mathcal{B}_{1,p}\{\omega_j\} = p^k \cdot (1-p)^{n-k}$$

mit $k := \#\{j \leq n \mid \omega_j = 1\}$, die Anzahl der Einsen in $(\omega_1, \ldots, \omega_n)$.
Sei $X : \Omega \to \mathbb{N}$, $X(\omega_1, \ldots, \omega_n) := \#\{j \leq n \mid \omega_j = 1\}$ die Zufallsvariable, die die Anzahl der Erfolge angibt. Da man k Erfolge auf $\binom{n}{k}$ Arten unter n Experimenten erhalten kann, ist

$$\mathbf{P}\{X = k\} = \binom{n}{k} \cdot p^k \cdot (1-p)^{n-k}.$$

Das durch diese Zähldichte definierte Wahrscheinlichkeitsmaß auf $\mathbb{N}$ heißt <u>Binomialverteilung</u> mit Parametern n, p und wird mit $\mathcal{B}_{n,p}$ bezeichnet (vgl. (1.24.2)).
Ist $X_j : \Omega \to \{0,1\}$ die Projektion, so gilt nach Satz (4.16), daß $\mathbf{P}_{X_j} = \mathcal{B}_{1,p}$ und es folgt, daß $X = \sum\limits_{j=1}^{n} X_j$ die Summe von n unabhängigen $\mathcal{B}_{1,p}$-verteilten Zufallsvariable ist.

(c) Ähnlich wie man in (b) die Binomialverteilung als Verteilung der Summe von n unabhängigen $\mathcal{B}_{1,p}$-verteilten Zufallsvariablen erhält, kann man die <u>negative Binomialverteilung</u> aus der geometrischen Verteilung gewinnen (siehe (1.24.4)). Es seien $X_1, \ldots, X_n$ n unabhängige Zufallsvariable mit Verteilung $\mathcal{G}_p$, $p \in]0,1]$ (d.h. $\mathbf{P}\{X_j = k\} = p \cdot (1-p)^k$)

$$\begin{aligned} \Rightarrow \quad \mathbf{P}\{X_j = k_j \text{ für } 1 \leq j \leq n\} &= \prod_{j=1}^{n} \mathcal{G}_p\{k_j\} \\ &= p \cdot (1-p)^{k_1} \cdot \ldots \cdot p \cdot (1-p)^{k_n} \\ &= p^n \cdot (1-p)^k \end{aligned}$$

wobei $k = k_1 + \ldots + k_n$. Für die Verteilung der Zufallsvariable $X = X_1 + \ldots + X_n$ gilt also:

$$\begin{aligned} \mathbf{P}_X\{k\} &= \mathbf{P}\{X_1 + \ldots + X_n\} \\ &= \sum_{\substack{(k_1, \ldots, k_n) \in \mathbb{N}^n \\ k_1 + \ldots + k_n = k}} \mathbf{P}\{X_j = k_j \text{ für } 1 \leq j \leq n\} \\ &= \sum_{\substack{(k_1, \ldots, k_n) \in \mathbb{N}^n \\ k_1 + \ldots + k_n = k}} p^n \cdot (1-p)^k \\ &= f_n(k) \cdot p^n \cdot (1-p)^k \end{aligned}$$

wobei $f_n(k)$ die Anzahl der Summanden, d.h. die Anzahl der $(k_1, \ldots, k_n) \in \mathbb{N}^n$ mit $k_1 + \ldots + k_n = k$ ist. Da $\mathbf{P}_X$ ein Wahrscheinlichkeitsmaß ist

$$\Rightarrow \quad 1 = \sum_{k=0}^{\infty} \mathbf{P}_X\{k\} = \sum_{k=0}^{\infty} f_n(k) \cdot p^n \cdot (1-p)^k = 1 \quad \text{für alle } 0 < p \leq 1$$

Nach der Taylor'schen Formel.

$$\Rightarrow \quad \sum_{k=0}^{\infty} f_n(k) \cdot (1-p)^k = p^{-n} = (1-(1-p))^{-n} = \sum_{k=0}^{\infty} \binom{-n}{k} \cdot (1-p)^k$$

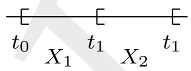
Koeffizientenvergleich liefert

$$f_n(k) = \binom{-n}{k} = \binom{n+k-1}{k}$$

$$\Rightarrow \quad \mathbf{P}_X\{k\} = \mathbf{P}\{X = k\} = \binom{n+k-1}{k} \cdot p^n \cdot (1-p)^k$$

für $k \geq 0$. Die Verteilung von X bezeichnen wir mit $\mathcal{NB}_{n,p}$, die negative Binomialverteilung mit Parametern n und p. Sie tritt auf als Wartezeit bis zum n-ten Erfolg in einer Folge von unabhängigen Experimenten mit Erfolgswahrscheinlichkeit p.

(d) Bei der Beobachtung einer radioaktiven Substanz werden im Zeitintervall $[t_0, t_1[$ insgesamt X_1 Zerfälle beobachtet und im Zeitintervall $[t_1, t_2[$ X_2.



Experimentell ergibt sich, daß X_1 und X_2 unabhängig sind und die Poisson-Verteilung Π_{λ_1} und Π_{λ_2} mit $\lambda_1, \lambda_2 > 0$ haben, d.h.

$$\mathbf{P}\{X_j = k\} = \Pi_{\lambda_j}\{k\} = \mathbf{e}^{-\lambda_j} \cdot \frac{\lambda_j^k}{k!}.$$

Unter diesem Modell können wir die Verteilung der Zahl $X_1 + X_2$ der Zerfälle in $[t_0, t_2[$ bestimmen:

$$\begin{aligned}
\mathbf{P}_{X_1+X_2}\{n\} &= \mathbf{P}\{X_1 + X_2 = n\} \\
&= \sum_{k=0}^{n} \mathbf{P}\{X_1 = k,\ X_2 = n-k\} \\
&\overset{(*)}{=} \sum_{k=0}^{n} \mathbf{P}\{X_1 = k\} \cdot \mathbf{P}\{X_2 = n-k\} \\
&= \sum_{k=0}^{n} \mathbf{e}^{-\lambda_1} \cdot \frac{\lambda_1^k}{k!} \cdot \mathbf{e}^{-\lambda_2} \cdot \frac{\lambda_2^{n-k}}{(n-k)!} \\
&= \frac{1}{n!} \cdot \mathbf{e}^{-(\lambda_1+\lambda_2)} \cdot \sum_{k=0}^{n} \frac{n!}{k! \cdot (n-k)!} \cdot \lambda_1^k \cdot \lambda_2^{n-k} \\
&= \mathbf{e}^{-(\lambda_1+\lambda_2)} \cdot \frac{(\lambda_1+\lambda_2)^n}{n!} \\
&= \Pi_{\lambda_1+\lambda_2}\{n\}
\end{aligned}$$

Bei $(*)$ geht die Unaghängigkeit von X_1 und X_2 ein.
Also hat die Zufallsvariable $X_1 + X_2$ die Verteilung $\Pi_{\lambda_1+\lambda_2}$.

Aufgaben:

Aufgabe 4.1:

Es sei $\mathfrak{A}$ eine σ-Algebra auf Ω. Zeigen Sie, daß eine Abbildung $X : \Omega \to \mathbb{R}^d$ mit $X(\omega) = (X_1(\omega), \ldots, X_d(\omega))$ genau dann ein Zufallsvektor ist, wenn X_j für ein $1 \leq j \leq d$ eine reelle Zufallsvariable ist.

Aufgabe 4.2:

(a) Zeigen Sie, daß beim zweifachen Wurf eines Würfels je zwei der Ereignisse

$$\begin{array}{lcl} A & \hat{=} & \text{erster Wurf ergibt wenigstens 4 Augen} \\ B & \hat{=} & \text{zweiter Wurf ergibt wenigstens 4 Augen} \\ C & \hat{=} & \text{Augensumme gerade} \end{array}$$

unabhängig sind, daß aber A, B und C nicht voneinander unabhängig sind.

(b) Ein Zufallsvektor (X_1, X_2) sei gleichverteilt auf $\left\{(x, y) \in \mathbb{R}^2 : x^2 + y^2 < 1\right\}$. Untersuchen Sie, ob X_1 und X_2 voneinander unabhängig sind.

Aufgabe 4.3:

Es seien $X, Y : \Omega \to \mathbb{N}$ unabhängige Zufallsvariable, die beide die geometrische Verteilung $\mathcal{G}_p$ mit $p \in]0, 1]$ haben. Zeigen Sie, daß dann die Zufallsvariablen $\min(X, Y)$ und $X - Y$ unabhängig sind.

Aufgabe 4.4:

Es seien X, Y unabhängige Zufallsvariable, deren Verteilungen die Verteilungsfunktionen $F := F_{\mathbf{P}_X}$ und $G := G_{\mathbf{P}_Y}$ haben. Berechnen Sie die Verteilungsfunktionen der Zufallsvariablen $\max(X, Y)$ und $\min(X, Y)$.

Was ergibt sich für exponentialverteilte X, Y?

Aufgabe 4.5:

Es sei $\mathbf{P}$ die Gleichverteilung auf $\Omega = \{0, 1\}^n$ und für $1 \leq j \leq n$ $A_j := \{(\omega_1, \ldots, \omega_n) \in \Omega : \omega_j = 1\}$ sowie $B := \left\{(\omega_1, \ldots, \omega_n) \in \Omega : \sum\limits_{j=1}^{n} \omega_j = 1 \mod 2\right\}$. Untersuchen Sie, welche der folgenden Systeme unabhängig sind:

$$(A_1, \ldots, A_n, B) \quad \text{bzw.} \quad (A_1, \ldots, A_n) \quad \text{bzw.} \quad (A_2, \ldots, A_n, B)\,.$$

Aufgabe 4.6:

Ein Stab der Länge 1 wird in zwei Stücke gebrochen. Das längere der beiden Stücke wird ein zweites Mal gebrochen. Es sei U der auf $]0, 1[$ gleichverteilte Teilpunkt beim ersten Brechen und V die Länge des größeren Stücks. Weiter seien Y und Z die Längen der Teilstücke nach dem zweiten Brechen, also $1 - V$, Y, Z die Längen der drei Teilstücke.

(a) Zeigen Sie, daß V gleichverteilt auf $\left[\frac{1}{2}, 1\right[$ ist.

(b) Zeigen Sie, daß $W := \frac{Y}{V}$ und V unabhängig sind und daß W gleichverteilt in $]0, 1[$ ist.

(c) Bestimmen Sie die Wahrscheinlichkeit, daß das beim ersten Brechen entstandene kürzere Teilstück das kürzere der drei Stücke ist.

Aufgabe 4.7:

$X_1, \ldots, X_n$ seien unabhängige und identisch verteilte Zufallsvariable, die monoton wachsend angeordnet seien zu einem Zufallsvektor $\big(X_{[1]}, \ldots, X_{[n]}\big)$.

(a) Bestimmen Sie die Verteilungsfunktion von $X_{[k]}$ für ein $k = 1, \ldots, n$.

(b) Bestimmen Sie speziell die Verteilungsfunktion von $X_{[1]}$ für den Fall, daß die X_i exponentialverteilt mit Parameter $\lambda > 0$ sind.

5. Kapitel: Erwartungswert und Varianz von Zufallsvariablen

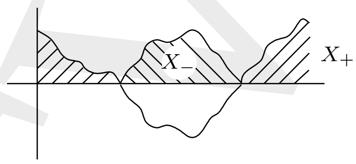
Interessiert man sich schon vor der Durchführung eines Experiments für das Ergebnis einer Zufallsvariablen X, so kann man zwar dies nicht exakt vorhersagen, aber immerhin einen „mittleren Wert“ ihrer Realisation. Dazu dient unter anderem der Begriff des Erwartungswerts. (Anwendungen: Fairer Einsatz beim Glücksspiel usw. ...) Vorher müssen wir noch den Integralbegriff aus Definition (3.12) für meßbare nichtnegative Funktionen auf Funktionen mit beliebigen Vorzeichen erweitern:
Erinnerung: Es sei $X : (\Omega, \mathfrak{A}) \to (\mathbb{R}, \mathfrak{B}(\mathbb{R}))$ eine Zufallsvariable und $\mathbf{P}$ ein Wahrscheinlichkeitsmaß auf $(\Omega, \mathfrak{A})$.
1. Schritt: Ist $X \geq 0$ eine Zufallsvariable, die nur endlich viele Werte annimmt, d.h. $X = \sum\limits_{j=1}^{n} x_j \cdot 1_{A_j}$, wobei $x_j > 0$ und $A_j \in \mathfrak{A}$, so ist

$$\int X \, \mathbf{dP} := \sum_{j=1}^{n} x_j \cdot \mathbf{P}(A_j)$$

2. Schritt: Ist $X : \Omega \to \mathbb{R}_+$ beliebige Zufallsvariable, so ist

$$\int X \, \mathbf{dP} := \sup_{\substack{Y \leq X \\ Y \in \mathcal{E}(\mathfrak{A})}} \int Y \, \mathbf{dP} \quad \in \mathbb{R}_+ \cup \{\infty\}$$

3. Schritt: Es sei $X : \Omega \to \mathbb{R}$ Zufallsvariable. Wir zerlegen X in $X_+ := \max\{X, 0\}$ und $X_- := \max\{-X, 0\}$. Dann sind X_+ und X_- nichtnegative Zufallsvariable und nach dem zweiten Schritt ist $\int X_+ \, \mathbf{dP}$ und $\int X_- \, \mathbf{dP}$ definiert, weiter gilt $X = X_+ - X_-$.



Wir definieren:

5.1 Definition (Integrierbar, Erwartungswert)

Eine reelle Zufallsvariable $X : \Omega \to \mathbb{R}$ heißt integrierbar, falls $\int X_+ \, \mathbf{dP} < \infty$ und $\int X_- \, \mathbf{dP} < \infty$. Wir nennen für solche X

$$\int X \, \mathbf{dP} := \int X_+ \, \mathbf{dP} - \int X_- \, \mathbf{dP}$$

das Integral von X bezüglich $\mathbf{P}$. Wir bezeichnen $\mathbf{E}(X) = \mathbf{E_P}(X) := \int X \, \mathbf{dP}$ als den Erwartungswert von X (unter $\mathbf{P}$). Ist $X \geq 0$ so setzen wir $\mathbf{E}(X) = \int X \, \mathbf{dP}$ auch in dem Fall, daß das Integral $+\infty$ ist. Im Fall $X \geq 0$ ist $X_+ = X$ und $X_- = 0$, also steht Definition (5.1) nicht im Widerspruch zu Definition (3.12),.

5.2 Beispiele

(a) Es sei X das Ergebnis beim Würfeln, d.h. $\Omega = \{1, \ldots, 6\}$, $\mathfrak{A} = \mathcal{POT}(\Omega)$ und $\mathbf{P} = \mathfrak{L}_\Omega$ und $X : \Omega \to \mathbb{R}$, $X(\omega) = \omega$. Dann ist nach Definition (3.10) das Integral der endlichwertigen Funktion

$$X = \sum_{j=1}^{6} j \cdot 1_{\{j\}} \text{ gleich } \mathbf{E_P}(X) = \sum_{j=1}^{6} j \cdot \mathbf{P}\{j\} = (1+2+3+4+5+6) \cdot \frac{1}{6} = 3,5.$$

(b) Allgemeiner sei X gleichverteilt auf einer endlichen Menge $I \subset \mathbb{R}$, d.h. $\mathbf{P}\{X = i\} = \frac{1}{\#I}$ $\forall i \in I$. Dann ist

$$\mathbf{E}_{\mathbf{P}}(X) = \sum_{i \in I} i \cdot \mathbf{P}\{X = i\} = \frac{1}{\#I} \sum_{i \in I} i$$

der <u>Mittelwert</u> aller Elemente von I.

(c) Es sei X $\mathcal{B}_{1,p}$-verteilt mit $p \in [0,1]$, d.h. $\mathbf{P}\{X = 0\} = 1 - p$, $\mathbf{P}\{X = 1\} = p$. Dann gilt

$$\mathbf{E}_{\mathbf{P}}(X) = 0 \cdot \mathbf{P}\{X = 0\} + 1 \cdot \mathbf{P}\{X = 1\} = p.$$

(d) Es sei $A \in \mathfrak{A}$. Dann gilt

$$\mathbf{E}_{\mathbf{P}}(1_A) = 1 \cdot \mathbf{P}\{1_A = 1\} + 0 \cdot \mathbf{P}\{1_A = 0\} = \mathbf{P}(A).$$

(e) Sei $\alpha \in \mathbb{R}$ beliebig, sowie $X = \alpha = \alpha \cdot 1_\Omega$ die deterministische Zufallsvariable mit Wert α. Dann gilt

$$\mathbf{E}(X) = \alpha \cdot 1 = \alpha.$$

5.3 Bemerkungen

Im folgenden seien stets $X, Y : \Omega \to \mathbb{R}$ reelle Zufallsvariable.

(a) X ist genau dann integrierbar (bezüglich $\mathbf{P}$), wenn $\mathbf{E}(|X|) < \infty$.
$\llbracket$ Es ist $\int X_+ \, \mathbf{dP} + \int X_- \, \mathbf{dP} = \int |X| \, \mathbf{dP}$ $\rrbracket$

(b) Sind X und Y integrierbar, so ist auch $X + Y$ integrierbar und es gilt $\mathbf{E}(X+Y) = \mathbf{E}(X) + \mathbf{E}(Y)$ (Additivität).
$\llbracket$ Sind $X, Y \geq 0$, so wurde dies in (3.16) gezeigt. Sind $0 \leq Y \leq X \Rightarrow 0 \leq X - Y$ und damit

$$\mathbf{E}(X) = \mathbf{E}((X-Y)+Y) = \mathbf{E}(X-Y) + \mathbf{E}(Y) \quad \Leftrightarrow \quad \mathbf{E}(X-Y) = \mathbf{E}(X) - \mathbf{E}(Y).$$

Nun gilt (!)

$$(X+Y)_+ \leq X_+ + Y_+ \quad \text{und} \quad (X+Y)_- \leq X_- + Y_-$$
$$\text{sowie} \quad X_+ + Y_+ - (X+Y)_+ = X_- + Y_- - (X+Y)_-$$

$$\begin{aligned}
\Rightarrow \quad \mathbf{E}(X+Y) &= \mathbf{E}\big((X+Y)_+\big) - \mathbf{E}\big((X+Y)_-\big) \\
&= \mathbf{E}(X_+ + Y_+) - \mathbf{E}\big(X_+ + Y_+ - (X+Y)_+\big) \\
&\qquad - \mathbf{E}(X_- + Y_-) + \mathbf{E}\big(X_- + Y_- - (X+Y)_-\big) \\
&= \mathbf{E}(X_+) + \mathbf{E}(Y_+) - \mathbf{E}(X_-) - \mathbf{E}(Y_-) \\
&= \mathbf{E}(X) + \mathbf{E}(Y)
\end{aligned}$$

$\rrbracket$

(c) Ist X integrierbar und $c \in \mathbb{R}$, so ist auch $c \cdot X$ integrierbar und $\mathbf{E}(c \cdot X) = c \cdot \mathbf{E}(X)$.
$\llbracket$ Hier nur $c \leq 0$ (anderer Fall analog!):

$$(c \cdot X)_+ = |c| \cdot X_- \quad \text{und} \quad (c \cdot X)_- = |c| \cdot X_+$$

$$\begin{aligned}
\Rightarrow \quad \mathbf{E}(c \cdot X) &= \mathbf{E}\big((c \cdot X)_+\big) - \mathbf{E}\big((c \cdot X)_-\big) = \mathbf{E}(|c| \cdot X_-) - \mathbf{E}(|c| \cdot X_+) \\
&= |c| \cdot (\mathbf{E}(X_-) - \mathbf{E}(X_+)) = |c| \cdot (-\mathbf{E}(X)) = (-|c|) \cdot \mathbf{E}(X) \\
&= c \cdot \mathbf{E}(X)
\end{aligned}$$

$\rrbracket$

(d) Sind X, Y integrierbar, oder ≥ 0 und gilt $X \leq Y$, so folgt $\mathbf{E}(X) \leq \mathbf{E}(Y)$, denn

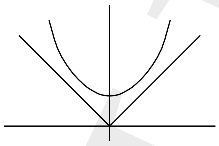
$$\begin{aligned} Y - X \geq 0 \quad &\Rightarrow \quad \mathbf{E}(Y - X) \geq 0 \\ &\Rightarrow \quad \mathbf{E}(Y) - \mathbf{E}(X) \geq 0 \\ &\Rightarrow \quad \mathbf{E}(X) \leq \mathbf{E}(Y) \end{aligned}$$

$\rrbracket$

(e) Ist X^2 integrierbar $\Rightarrow$ X integrierbar

$\llbracket$ Es gilt

$$|X| \leq \frac{1}{2} \cdot \left(1 + X^2\right)$$



$$\underset{\text{(d)}}{\Rightarrow} \quad \mathbf{E}(|X|) \leq \frac{1}{2} \cdot \mathbf{E}\left(1 + X^2\right) = \frac{1}{2} + \frac{1}{2} \cdot \mathbf{E}\left(X^2\right) < \infty$$

und die Behauptung folgt aus (5.3.a) $\rrbracket$

Im Fall von abzählbaren Wahrscheinlichkeitsräumen oder -verteilungen mit Dichte gibt es einfachere Formeln als Definition (5.1) zur Berechnung von $\mathbf{E}(X)$. Dazu benötigen wir

5.4 Satz (Transformationsformel)

Es sei $(\Omega, \mathfrak{A}, \mathbf{P})$ ein Wahrscheinlichkeitsraum, $\mathfrak{A}'$ eine σ-Algebra auf Ω', $X : (\Omega, \mathfrak{A}) \to (\Omega', \mathfrak{A}')$ eine Zufallsvariable und $g : \Omega' \to \mathbb{R}$ eine reelle Zufallsvariable.
Dann ist $g(X) := g \circ X$ genau dann integrierbar bezüglich $\mathbf{P}$, wenn g integrierbar bezüglich $\mathbf{P}_X$ ist und es gilt

$$\begin{tikzcd} \Omega \arrow[r, "X"'] \arrow[rr, bend left, "g \circ X"] & \Omega' \arrow[r, "g"'] & \mathbb{R} \end{tikzcd}$$

$$\mathbf{P} \rightsquigarrow \mathbf{P}_X$$

$$\mathbf{E_P}(g \circ X) = \int g \, \mathbf{dP}_X = \int g(X) \, \mathbf{dP}(X)$$

Beweis:

(I) g nimmt nur endlich viele Werte an, also $g = \sum_{j=1}^{n} \alpha_j \cdot 1_{A_j}$ mit $A_j \in \mathfrak{A}'$

$$\begin{aligned} \Rightarrow \quad \mathbf{E_P}(g \circ X) \quad &= \quad \sum_{j=1}^{n} \int \alpha_j \cdot 1_{\{X \in A_j\}} \, \mathbf{dP} = \sum_{j=1}^{n} \alpha_j \cdot \mathbf{P}\{X \in A_j\} \\ &= \quad \sum_{j=1}^{n} \alpha_j \cdot \mathbf{P}_X(A_j) = \int g \, \mathbf{dP}_X \end{aligned}$$

(II) g $\geq$ 0 meßbar. Nach Lemma (3.15) existiert eine Folge $(g_n) \subset \mathcal{E}(\mathfrak{A}')$ mit $g_n \uparrow g$. $\Rightarrow g_n \circ X \uparrow g \circ X$ und nach Beppo-Levi folgt:

$$\begin{aligned} \mathbf{E_P}(g \circ X) \quad &= \quad \int g \circ X \, \mathbf{dP} = \lim_{n \to \infty} \int g_n \circ X \, \mathbf{dP} \\ &= \quad \lim_{n \to \infty} \int g_n \, \mathbf{dP}_X = \int g \, \mathbf{dP}_X . \end{aligned}$$

Wegen (5.4.a) folgt daraus auch die behauptete Äquivalenz.

(III) g beliebige Zufallsvariable: Wegen $(g \circ X)_+ = g_+ \circ X$ und $(g \circ X)_- = g_- \circ X$ folgt

$$\begin{aligned}\mathbf{E_P}(g \circ X) &= \mathbf{E_P}(g_+ \circ X) - \mathbf{E_p}(g_- \circ X) = \int g_+ \, \mathbf{dP}_X - \int g_- \, \mathbf{dP}_X \\ &= \int g \, \mathbf{dP}_X.\end{aligned}$$

□

5.5 Folgerungen

5.5.1 Korollar

Es sei X eine reelle Zufallsvariable auf einem Wahrscheinlichkeitsraum $(\Omega, \mathfrak{A}, \mathbf{P})$. Dann hängt der Erwartungswert $\mathbf{E_P}(X)$ nur von der Verteilung $\mathbf{P}_X$ von X ab und nicht von der speziellen Wahl des Wahrscheinlichkeitsraumes $(\Omega, \mathfrak{A}, \mathbf{P})$.

Beweis:

Ist $X' : \Omega' \to \mathbb{R}$ eine weitere Zufallsvariable auf $(\Omega', \mathfrak{A}', \mathbf{P}')$ mit $\mathbf{P}'_{X'} = \mathbf{P}_X$ so folgt aus Satz (5.4) für $g = \mathrm{id}_{\mathbb{R}}$:

$$\mathbf{E_P}(X) = \mathbf{E_P}(g \circ X) = \int \mathrm{id}_{\mathbb{R}} \, \mathbf{dP}_X = \int \mathrm{id}_{\mathbb{R}} \, \mathbf{dP}'_{X'} = \mathbf{E_{P'}}(X')$$

□

5.5.2 Korollar

Ist X eine Zufallsvariable, deren Werte sämtlich in einer abzählbaren Teilmenge $A \subset \mathbb{R}$ liegen, so ist X integrierbar genau dann, wenn

$$\sum_{x \in A} |x| \cdot \mathbf{P}\{X = x\}$$

endlich ist und es gilt

$$\mathbf{E_P}(X) = \sum_{x \in A} x \cdot \mathbf{P}\{X = x\}.$$

Beweis:

(I) Hat $X \geq 0$ nur endlich viele Werte, so folgt dies aus der Definition des Integrals.

(II) Wenn $X \geq 0$ die abzählbar vielen Werte $x_0, x_1, \ldots$ annimmt, also $X = \sum_{y=0}^{\infty} x_j \cdot 1_{A_j}$ mit disjunkten Mengen $A_j \in \mathfrak{A} \Rightarrow X_n = \sum_{j=0}^{n} x_j \cdot 1_{A_j} \uparrow X$. Mit Beppo-Levi folgt

$$\begin{aligned}\mathbf{E}(X) &= \lim_{n \to \infty} \mathbf{E}(X_n) = \lim_{n \to \infty} \sum_{j=1}^{n} x_j \cdot \mathbf{P}(A_j) \\ &= \sum_{j=0}^{\infty} x_j \cdot \mathbf{P}\{X = x_j\} = \sum_{x \in A} x \cdot \mathbf{P}\{X = x\}\end{aligned}$$

da $A = \{x_0, x_1, \ldots\}$. Da $x_j \geq 0 \; \forall j$ folgt auch die im Satz behauptete Äquivalenz.

(III) Ist X eine beliebige Zufalsvariable mit abzählbar vielen Werten, so folgt die Behauptung durch Zerlegung: $X = X_+ - X_-$, $A = A_+ \cup A_-$ mit $A_+ = \{x \in A \mid x \geq 0\} \cup \{0\}$ und $A_- = \{x \in A \mid x \leq 0\} \cup \{0\}$.

$$\begin{aligned} \Rightarrow \quad \mathbf{E}(X_+) &= \sum_{x \in A_+} x \cdot \mathbf{P}\{X_+ = x\} = \sum_{x \in A_+} x \cdot \mathbf{P}\{X = x\} \\ \mathbf{E}(X_-) &= \sum_{x \in A_-} |x| \cdot \mathbf{P}\{X_- = x\} = \sum_{x \in A_-} |x| \cdot \mathbf{P}\{X = x\} \end{aligned}$$

$$\begin{aligned} \Rightarrow \quad \mathbf{E}(X) &= \mathbf{E}(X_+) - \mathbf{E}(X_-) \\ &= \sum_{x \in A_+} x \cdot \mathbf{P}\{X = x\} - \sum_{x \in A_-} |x| \cdot \mathbf{P}\{X = x\} \\ &= \sum_{x \in A} x \cdot \mathbf{P}\{X = x\} \end{aligned}$$

□

5.5.3 Korollar

Es seien X eine reelle Zuvallsvariable, deren Verteilung $\mathbf{P}_X$ eine Lebesgue-Dichte f besitzt und $g : \mathbb{R} \to \mathbb{R}$ meßbar. Dann ist $g(X) = g \circ X$ integrierbar genau dann, wenn

$$\int |g(x)| \cdot f(x) \, \mathbf{d}\lambda^1(x) < \infty$$

und es gilt

$$\mathbf{E}_{\mathbf{P}}(g(X)) = \int g(x) \cdot f(x) \, \mathbf{d}\lambda^1(x).$$

Beweis:
Es sei Q das Wahrscheinlichkeitsmaß auf $\mathbb{R}$ mit Dichte f, d.h. $Q(A) = \int 1_A \cdot f \, \mathbf{d}\lambda^1$ für $A \in \mathfrak{A}(\mathbb{R})$.

(I) g nimmt nur endlich viele Werte ≥ 0 an, d.h. $g = \sum_{j=1}^{n} x_j \cdot 1_{A_j}$ mit $x_j \geq 0$ und $A_j \in \mathfrak{B}(\mathbb{R})$ paarweise disjunkt.

$$\begin{aligned} \Rightarrow \quad \int g \, \mathbf{d}Q &= \sum_{j=1}^{n} x_j \cdot Q(A_j) = \sum_{j=1}^{n} x_j \cdot \int 1_{A_j} \cdot f \, \mathbf{d}\lambda^1(x) \\ &= \int \sum_{j=1}^{n} x_j \cdot 1_{A_j}(x) \cdot f(x) \, \mathbf{d}\lambda^1(x) = \int g(x) \cdot f(x) \, \mathbf{d}\lambda^1(x). \end{aligned}$$

(II) $g \geq 0$: Nach Lemma (3.15) existiert eine Folge $(g_n) \subset \mathcal{E}(\mathfrak{B}(\mathbb{R}))$ mit $g_n \uparrow g$. Nach Beppo-Levi gilt

$$\begin{aligned} \int g \, \mathbf{d}Q &= \lim_{n \to \infty} \int g_n \, \mathbf{d}Q = \lim_{n \to \infty} \int g_n \cdot f \, \mathbf{d}\lambda^1 \\ &\underset{\text{B.-L.}}{=} \int g \cdot f \, \mathbf{d}\lambda^1 \end{aligned}$$

da $g_n(x) \cdot f(x) \uparrow g(x) \cdot f(x)$ mit $n \to \infty$, da $f \geq 0$.

(III) Ist $g = g_+ - g_-$ beliebig

$$\begin{aligned}\Rightarrow \quad \int g \, \mathbf{d}Q &= \int g_+ \, \mathbf{d}Q - \int g_- \, \mathbf{d}Q \\ &= \int g_+ \cdot f \, \mathbf{d}\lambda^1 - \int g_- \cdot f \, \mathbf{d}\lambda^1 = \int g \cdot f \, \mathbf{d}\lambda^1.\end{aligned}$$

Wir wenden nun dies auf $Q = \mathbf{P}_X$ an. Nach Satz (5.4) folgt

$$\mathbf{E_P}(g \circ X) = \int g \, \mathbf{dP}_X = \int g \, \mathbf{d}Q = \int g(x) \cdot f(x) \, \mathbf{d}\lambda^1(x)$$

□

5.5.4 Folgerung

Falls X integrierbar ist und die Verteilung von X die Dichte f hat, so gilt

$$\mathbf{E_P}(X) = \int x \cdot f(x) \, \mathbf{d}\lambda^1(x).$$

5.6 Beispiele

(a) Es sei $X : \Omega \to \mathbb{R}$ eine Zufallsvariable mit geometrischer Verteilung $\mathcal{G}_p$ (wobei $p \in \,]0,1]$). Dann gilt

$$\mathbf{E}(X) = \frac{1}{p} - 1.$$

⟦ In (4.19.c) hatten wir die negative Binomialverteilung $\mathcal{NB}_{2,p}$ auf $\mathbb{N}$ mit der Zähldichte

$$\mathcal{NB}_{2,p}\{k\} = \binom{k+1}{k} \cdot p^2 \cdot (1-p)^k \quad \text{für } k \in \mathbb{N}$$

eingeführt.

$$\Rightarrow \quad \sum_{k=0}^{\infty} (k+1) \cdot p^2 \cdot (1-p)^k = 1,$$

da $\binom{k+1}{k} = k+1$.

$$\begin{aligned}\Rightarrow \quad \mathbf{E}(X) &= \sum_{k=0}^{\infty} k \cdot \mathbf{P}\{X = k\} = \sum_{k=0}^{\infty} k \cdot p \cdot (1-p)^k \\ &= \frac{1}{p} \cdot \sum_{k=0}^{\infty} (k+1) \cdot p^2 \cdot (1-p)^k - \sum_{k=0}^{\infty} p \cdot (1-p)^k \\ &= \frac{1}{p} - p \cdot \frac{1}{1-(1-p)} = \frac{1}{p} - 1\end{aligned}$$

⟧

(b) Es sei X eine reelle Zufallsvariable mit Normalverteilung $\mathcal{N}_{\mu,\sigma^2}$.

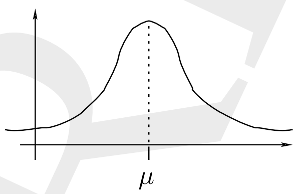
(1) Sei $\mu = 0$, $\sigma^2 = 1$.

$$\begin{aligned}\Rightarrow \quad \mathbf{E}(|X|) &= \frac{1}{\sqrt{2\pi}} \int_{-\infty}^{\infty} |x| \cdot \mathbf{e}^{-x^2/2} \, \mathbf{d}x = \frac{2}{\sqrt{2\pi}} \int_0^{\infty} x \cdot \mathbf{e}^{-x^2/2} \, \mathbf{d}x \\ &= \left[\frac{2}{\sqrt{2\pi}} \cdot \left(-\mathbf{e}^{-x^2/2} \right) \right]_0^{\infty} = \frac{2}{\sqrt{2\pi}} < \infty,\end{aligned}$$

also besitzt X einen Erwartungswert. Es folgt $\mathbf{E}(X) = 0$.

(2) $\mu \in \mathbb{R}$, $\sigma^2 > 0$ beliebig: Setze $X' = \dfrac{X-\mu}{\sigma}$. Nach (3.20.b) hat X' eine $\mathcal{N}_{0,1}$-Verteilung.

$$\begin{aligned} \Rightarrow \quad 0 = \mathbf{E}\left(X'\right) = \mathbf{E}\left(\frac{1}{\sigma}\left(X-\mu\right)\right) = \frac{1}{\sigma}\mathbf{E}\left(X\right) - \frac{\mu}{\sigma} \\ \Leftrightarrow \quad \mathbf{E}\left(X\right) = \mu. \end{aligned}$$



(c) Es sei nun X eine reelle Zufallsvariable mit der Cauchy-Verteilung γ_α; die Lebegue-Dichte ist dann

$$x \mapsto \frac{\alpha}{\pi\left(\alpha^2 + x^2\right)}$$

auf $\mathbb{R}$. Trotz ihrer Symmetrie um 0 ist der Erwartungswert aber nicht $= 0$. Wegen

$$\begin{aligned} \mathbf{E}\left(|X|\right) &= \int_{-\infty}^{\infty} \frac{\alpha \cdot |x|}{\pi\left(\alpha^2+x^2\right)}\,\mathbf{d}x = \frac{\alpha}{\pi} \cdot \int_0^\infty \frac{2 \cdot x}{\alpha^2+x^2}\,\mathbf{d}x \\ &= \left[\frac{\alpha}{\pi} \cdot \ln\left(\alpha^2+x^2\right)\right]_0^\infty = \infty \end{aligned}$$

existiert der Erwartungswert nicht.

(d) Es sei X eine Zufallsvariable, die $\mathcal{B}_{n,p}$-verteilt ist mit $n \geq 1$ und $p \in [0,1]$.
Für $n = 1$ gilt nach (5.2.c) $\mathbf{E}\left(X\right) = p$.
Für $n \geq 1$ gilt $\mathbf{E}\left(X\right) = n \cdot p$.

$\llbracket$ Sind $X_1, \ldots, X_n$ unabhängige $\mathcal{B}_{1,p}$-verteilte Zufallsvariable, so ist nach (4.19.b) $\overline{X} = \sum\limits_{j=1}^n X_j$ $\mathcal{B}_{n,p}$-verteilt. Aus Korollar (5.5.1) und (5.3.b) folgt dann

$$\mathbf{E}\left(X\right) = \mathbf{E}\left(\overline{X}\right) = \sum_{j=1}^n \mathbf{E}\left(X_j\right) = \sum_{j=1}^n p = n \cdot p.$$

$\rrbracket$

Während der Erwartungswert $\mathbf{E}\left(X\right)$ eine Aussage über den mittleren Wert, den eine Zufallsvariable X annimmt, macht, ist darin noch keine Information enthalten, wie weit die Zufallsvariable von diesem mittleren Wert im Mittel abweicht. Dazu braucht man:

5.7 Definition (Varianz)

Es sei $X : \Omega \to \mathbb{R}$ eine Zufallsvariable auf einem Wahrscheinlichkeitsraum $(\Omega, \mathfrak{A}, \mathbf{P})$, für die X^2 integrierbar ist. (Man nennt X quadratisch integrierbar.) Die Zahl

$$\mathbf{V_P}\left(X\right) = \mathbf{E_P}\left(\left(X - \mathbf{E_P}\left(X\right)\right)^2\right)$$

nennt man <u>Varianz</u> von X. Die <u>Streuung</u> von X ist

$$\sigma\left(X\right) = \sqrt{\mathbf{V_P}\left(X\right)};$$

diese Größe hat die gleiche Dimension wie X.

5.8 Bemerkungen

(a) Wegen $(X - \mathbf{E_P}(X))^2 = X^2 - 2\mathbf{E_P}(X) \cdot X + (\mathbf{E_P}(X))^2$ und da nach (5.3.e) auch X integrierbar ist, ist $\mathbf{V}_P(\mathbf{X})$ wohldefiniert und ≥ 0.

(b) Es ist

$$\mathbf{V_P}(X) = \mathbf{E_P}\left(X^2\right) - (\mathbf{E_P}(X))^2 .$$

⟦

$$\begin{aligned}
\mathbf{V_P}(X) &= \mathbf{E_P}\left((X - \mathbf{E_P}(X))^2\right) = \mathbf{E_P}\left(X^2 - 2\mathbf{E_P}(X) \cdot X + (\mathbf{E_P}(X))^2\right) \\
&= \mathbf{E_P}\left(X^2\right) - 2\mathbf{E_P}(X) \cdot \mathbf{E_P}(X) + (\mathbf{E_P}(X))^2 \\
&= \mathbf{E_P}\left(X^2\right) - (\mathbf{E_P}(X))^2
\end{aligned}$$

⟧

(c) Wie schon der Erwartungswert hängt $\mathbf{V_P}$ nur von der Verteilung $\mathbf{P}_X$ von X ab. Wir schreiben oft kurz $\mathbf{V}(X)$ statt $\mathbf{V_P}(X)$.

5.9 Beispiele

(a) (Vergleiche (5.2.a)) Für die Augenzahl X eines Würfels gilt

$$\mathbf{E}\left(X^2\right) = \frac{1}{6} \cdot \sum_{j=1}^{6} j^2 = \frac{91}{6}$$

und damit

$$\mathbf{V}(X) = \mathbf{E}\left(X^2\right) - (\mathbf{E}(X))^2 = \frac{91}{6} - \frac{49}{4} = \frac{35}{12}.$$

Die Streuung beträgt

$$\sigma(X) = \sqrt{\mathbf{V}(X)} = \sqrt{\frac{35}{12}} \approx 1,708.$$

(b) Es sei X $\mathcal{B}_{1,p}$-verteilt.

$$\Rightarrow \quad \mathbf{V}(X) = \mathbf{E}\left(X^2\right) - (\mathbf{E}(X))^2 = \mathbf{E}(X) - (\mathbf{E}(X))^2 = p - p^2 = p \cdot (1-p)$$

(c) Ist $X = \alpha \in \mathbb{R}$ eine deterministische Zufallsvariable, so gilt $\mathbf{E}\left(X^2\right) = \alpha^2 \Rightarrow \mathbf{V}(X) = 0$. Dies ist plausibel. Die Umkehrung wird in den Übungsaufgaben gezeigt.

(d) Ist X eine beschränkte Zufallsvariable, d.h. $|X| \leq B$, $B \in \mathbb{R}_+$, so existieren Erwartungswert und Varianz von X.

(e) Für jede quadratisch integrierbare Zufallsvariable X gilt

$$\mathbf{V}(c \cdot X) = c^2 \cdot \mathbf{V}(X)$$

für alle $c \in \mathbb{R}$.

(f) Ist X eine $\mathcal{N}_{\mu,\sigma^2}$-verteilte Zufallsvariable, so gilt

$$\mathbf{V}(X) = \sigma^2.$$

⟦ Betrachte $X' = \dfrac{X - \mu}{\sigma}$. Dann ist X' $\mathcal{N}_{0,1}$-verteilt. Also ist

$$\mathbf{E}\left(X'^2\right) = \int x^2 \, \mathbf{dP}_{X'}(x) = \frac{1}{\sqrt{2\pi}} \int_{-\infty}^{\infty} x^2 \cdot \mathbf{e}^{-x^2/2} \, \mathbf{d}x.$$

Nun ist

$$\begin{aligned}\int_{-\infty}^{\infty} x^2 \cdot \mathbf{e}^{-x^2/2}\,\mathbf{d}x &= \int_{-\infty}^{\infty} \underbrace{(-x)}_{u} \cdot \underbrace{(-x)\cdot \mathbf{e}^{-x^2/2}}_{v'}\,\mathbf{d}x \\ &= \Big[\underbrace{(-x)}_{u} \cdot \underbrace{\mathbf{e}^{-x^2/2}}_{v}\Big]_{-\infty}^{\infty} - \int_{-\infty}^{\infty} \underbrace{(-1)}_{u'} \cdot \underbrace{\mathbf{e}^{-x^2/2}}_{v} = \sqrt{2\pi}\end{aligned}$$

$$\begin{aligned}\Rightarrow \quad \mathbf{V}(X') &= \mathbf{E}\left(X'^2\right) = 1 \quad \text{und somit} \\ \mathbf{V}(X) &= \mathbf{E}\left((X-\mu)^2\right) = \mathbf{E}\left(\sigma^2 \cdot X'^2\right) = \sigma^2 \cdot \mathbf{E}\left(X'^2\right) = \sigma^2.\end{aligned}$$

⫫

5.10 Satz (*Cauchy-Schwarz*'sche Ungleichung)

Es seien X, Y quadratisch integrierbare Zufallsvariable. Dann besitzt die Zufallsvariable XY einen Erwartungswert und es gilt

$$|\mathbf{E}(XY)| \leq \sqrt{|\mathbf{E}(X^2) \cdot \mathbf{E}(Y^2)|}$$

Beweis:
Es sei $\lambda > 0$ beliebig. Wegen

$$X^2 - 2\lambda XY + \lambda^2 Y^2 = (X - \lambda Y)^2 \geq 0$$

folgt

$$XY \leq \frac{1}{2\lambda} X^2 + \frac{\lambda}{2} Y^2$$

und (da diese Ungleichung auch für $-Y$ anstelle von Y gilt), folgt

$$|XY| \leq \frac{1}{2\lambda} X^2 + \frac{\lambda}{2} Y^2.$$

Da die rechte Seite einen Erwartungswert besitzt, gilt dies auch für die linke Seite.

$$\Rightarrow \quad |\mathbf{E}(XY)| \leq \mathbf{E}(|XY|) \leq \frac{1}{2\lambda}\mathbf{E}\left(X^2\right) + \frac{\lambda}{2}\mathbf{E}\left(Y^2\right).$$

<u>1. Fall</u>: $\mathbf{E}\left(Y^2\right) = 0$. Mit $\lambda \to 0$ folgt dann $\mathbf{E}(XY) = 0$ und damit die Behauptung.

<u>2. Fall</u>: $\mathbf{E}\left(Y^2\right) > 0$. Setze $\lambda = \sqrt{\frac{\mathbf{E}(X^2)}{\mathbf{E}(Y^2)}}$

$$\begin{aligned}\Rightarrow \quad |\mathbf{E}(XY)| &\leq \frac{1}{2\sqrt{\mathbf{E}(X^2)/\mathbf{E}(Y^2)}}\mathbf{E}\left(X^2\right) + \frac{\sqrt{\mathbf{E}(X^2)/\mathbf{E}(Y^2)}}{2}\mathbf{E}\left(Y^2\right) \\ &= \sqrt{\mathbf{E}(X^2) \cdot \mathbf{E}(Y^2)}\end{aligned}$$

□

Satz (5.10) ermöglicht die folgende Definition:

5.11 Definition (Kovarianz, unkorreliert)

Es seien X, Y quadratisch integrierbare Zufallsvariable. Dann heißt

$$\operatorname{Kov}(X, Y) = \mathbf{E}(XY) - \mathbf{E}(X) \cdot \mathbf{E}(Y)$$

die <u>Kovarianz</u> von X und Y. Die Kovarianz ist ein Maß für die affin-lineare Abhängigkeit zwischen X und Y. Falls $\operatorname{Kov}(X, Y) = 0$, so nennt man X und Y <u>unkorreliert</u>.

5.12 Satz

(a) Es seien X, Y quadratisch integrierbare Zufallsvariable.

$$\Rightarrow \quad |\mathrm{Kov}\,(X,Y)| \leq \sqrt{\mathbf{V}\,(X) \cdot \mathbf{V}\,(Y)}$$

(b) Sind $X_1, \ldots, X_n$ quadratisch integrierbare Zufallsvariable, dann existiert $\mathbf{E}\left(\left(\sum\limits_{j=1}^{n} X_j\right)^2\right)$ und es gilt

$$\mathbf{V}\left(\sum_{j=1}^{n} X_j\right) = \sum_{j=1}^{n} \mathbf{V}\,(X_j) + 2 \cdot \sum_{1 \leq j < k}^{n} \mathrm{Kov}\,(X_j, X_k)$$

(Formel von *Bienaimée*)

(c) Sind zusätzlich die Zufallsvariablen unabhängig, so sind die (X_j) unkorreliert und

$$\mathbf{V}\left(\sum_{j=1}^{n} X_j\right) = \sum_{j=1}^{n} \mathbf{V}\,(X_j)\,.$$

Beweis:

(a) Es gilt $\mathrm{Kov}\,(X,Y) = \mathbf{E}\,((X - \mathbf{E}\,(X))\,(Y - \mathbf{E}\,(Y)))$. Aus Satz (5.10) folgt dann

$$\begin{aligned} |\mathrm{Kov}\,(X,Y)| &= |\mathbf{E}\,((X - \mathbf{E}\,(X))\,(Y - \mathbf{E}\,(Y)))| \\ &\leq \sqrt{\mathbf{E}\left((X - \mathbf{E}\,(X))^2\right) \mathbf{E}\left((Y - \mathbf{E}\,(Y))^2\right)} \\ &= \sqrt{\mathbf{V}\,(X) \cdot \mathbf{V}\,(Y)} \end{aligned}$$

(b) Die erste Behauptung folgt aus

$$\left(\sum_{j=1}^{n} X_j\right)^2 = \sum_{j,k=1}^{n} X_j \cdot X_k \text{ und Satz (5.10)}$$

$$\begin{aligned} \Rightarrow \quad \mathbf{V}\left(\sum_{j=1}^{n} X_j\right) &= \mathbf{E}\left(\left(\sum_{j=1}^{n} X_j\right)^2\right) - \left(\mathbf{E}\left(\sum_{j=1}^{n} X_j\right)\right)^2 \\ &= \sum_{j=1}^{n} \mathbf{E}\left(X_j^2\right) + 2 \cdot \sum_{j<k} \mathbf{E}\,(X_j X_k) \\ &\qquad - \sum_{j=1}^{n} (\mathbf{E}\,(X_j))^2 - 2 \cdot \sum_{j<k} \mathbf{E}\,(X_j)\,\mathbf{E}\,(X_k) \\ &= \sum_{j=1}^{n} \mathbf{V}\,(X_j) + 2 \cdot \sum_{j<k} \mathrm{Kov}\,(X_j, X_k)\,. \end{aligned}$$

Für den Beweis von (c) brauchen wir:

5.13 Lemma

Sind X, Y integrierbare und unabhängige Zufallsvariable, so besitzt $X \cdot Y$ einen Erwartungswert und es gilt

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

Beweis:
Wir zeigen (scheinbar) etwas mehr als behauptet, nämlich

$$\mathbf{E}((f \circ X) \cdot (g \circ Y)) = \mathbf{E}(f \circ X) \cdot \mathbf{E}(g \circ Y)$$

für alle meßbaren $f, g : \mathbb{R} \to \mathbb{R}_+$, falls entweder $f, g \geq 0$ oder $f \circ X$ und $g \circ Y$ integrierbar sind.

(I) f, g nehmen nur endlich viele Werte ≥ 0 an:

$$f(\mathbb{R}) = \{x_1, \ldots, x_m\} \qquad g(\mathbb{R}) = \{y_1, \ldots, y_n\}.$$

$$\begin{aligned} \Rightarrow \quad \mathbf{E}((f \circ X) \cdot (g \circ Y)) &= \sum_{j=1}^{m} \sum_{k=1}^{n} x_j \cdot y_k \cdot \mathbf{P}\{f \circ X = x_j,\, g \circ Y = y_k\} \\ &= \sum_{j=1}^{m} \sum_{k=1}^{n} x_j \cdot y_k \cdot \mathbf{P}\{f \circ X = x_j\} \cdot \mathbf{P}\{g \circ Y = y_k\} \\ &= \sum_{j=1}^{m} x_j \cdot \mathbf{P}\{f \circ X = x_j\} \cdot \sum_{k=1}^{n} y_k \cdot \mathbf{P}\{g \circ Y = y_k\} \\ &= \mathbf{E}(f \circ X) \cdot \mathbf{E}(g \circ Y). \end{aligned}$$

(II) $f, g \geq 0$. Wähle $f_n, g_n \in \mathcal{E}(\mathfrak{B}(\mathbb{R}))$ mit $f_n \uparrow f$, $g_n \uparrow g$ nach Lemma (3.15)

$$\Rightarrow \quad (f_n \circ X) \cdot (g_n \circ Y) \uparrow (f \circ X) \cdot (g \circ Y)$$

und nach Beppo-Levi folgt

$$\begin{aligned} \mathbf{E}((f \circ X) \cdot (g \circ Y)) &= \lim_{n \to \infty} \mathbf{E}((f_n \circ X) \cdot (g_n \circ Y)) \\ &= \lim_{n \to \infty} \mathbf{E}(f_n \circ X) \cdot \mathbf{E}(g_n \circ Y) \\ &= \mathbf{E}(f \circ X) \cdot \mathbf{E}(g \circ Y). \end{aligned}$$

(III) Schreibe $f = f_+ - f_-$ und $g = g_+ - g_-$ und benutze (II).

Mit $f(x) = g(x) = |x| \geq 0$ folgt $\mathbf{E}(|X \cdot Y|) = \mathbf{E}(|X|) \cdot \mathbf{E}(|Y|) < \infty$ also ist $X \cdot Y$ integrierbar und mit $f(x) = g(x) = x$ folgt die Behauptung. □

Beweis von (5.12.**c**)

Nach Lemma (5.13) ist

$$\operatorname{Kov}(X_j, Y_k) = \mathbf{E}(X_j \cdot Y_k) - \mathbf{E}(X_j) \cdot \mathbf{E}(Y_k) = 0 \quad \text{für } j \neq k,$$

damit sind X_j, Y_k unkorreliert und aus (b) folgt dann die Behauptung. □

5.14 Definition (Korrelation)

Es seien X, Y quadratisch integrierbare Zufallsvariablen. Dann heißt die Zahl

$$\operatorname{Kor}(X, Y) := \begin{cases} \frac{\operatorname{Kov}(X,Y)}{\sqrt{\mathbf{V}(X) \cdot \mathbf{V}(Y)}} & \text{falls } \mathbf{V}(X) > 0,\, \mathbf{V}(Y) > 0 \\ 0 & \text{sonst} \end{cases}$$

die <u>Korrelation</u> von X und Y.

5.15 Bemerkungen

(a) Wegen (5.12.a) gilt $-1 \leq \operatorname{Kor}(X,Y) \leq 1$.

(b) Wegen (5.12.c) gilt für unkorrelierte und damit erst recht unabhängige Zufallsvariable

$$\operatorname{Kor}(X,Y) = 0.$$

5.16 Beispiel

Es sei X $\mathcal{B}_{n,p}$-verteilt mit $n \geq 1$, $p \in]0,1[$. Dann gilt

$$\mathbf{V}(X) = n \cdot p \cdot (1-p).$$

⟦ Sind $X_1, \ldots, X_n$ unabhängige $\mathcal{B}_{1,p}$-verteilte Zufallsvariable, so hat nach (4.19.b) die Zufallsvariable $\overline{X} = \sum\limits_{j=1}^{n} X_j$ eine $\mathcal{B}_{n,p}$-Verteilung.

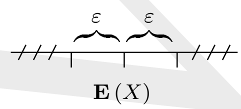
$$\Rightarrow \quad \mathbf{V}\left(\overline{X}\right) = \mathbf{V}(X) = \mathbf{V}\left(\sum_{j=1}^{n} X_j\right) = \sum_{j=1}^{n} \mathbf{V}(X_j) = n \cdot p \cdot (1-p),$$

da nach (5.9.b) $\mathbf{V}(X_j) = p \cdot (1-p)$ gilt. ⟧

5.17 Satz (*Tschebyscheff*'sche Ungleichung)

Es sei X eine quadratisch integrierbare Zufallsvariable. Dann gilt für alle $\varepsilon > 0$:

$$\mathbf{P}\{|X - \mathbf{E}(X)| \geq \varepsilon\} \leq \frac{\mathbf{V}(X)}{\varepsilon^2}$$



Beweis:

$$\begin{aligned}
\mathbf{V}(X) &= \mathbf{E}\left((X - \mathbf{E}(X))^2\right) \\
&\geq \mathbf{E}\left(1_{\{|X-\mathbf{E}(X)|\geq\varepsilon\}} \cdot (X - \mathbf{E}(X))^2\right) \\
&\geq \varepsilon^2 \cdot \mathbf{E}\left(1_{\{|X-\mathbf{E}(X)|\geq\varepsilon\}}\right) \\
&= \varepsilon^2 \cdot \mathbf{P}\{|X - \mathbf{E}(X)| \geq \varepsilon\}.
\end{aligned}$$

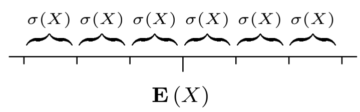
□

5.18 Bemerkung

Die obige Ungleichung präsentiert die Verwendung der Streuung $\sigma(X)$ als Maß für die Breite der Verteilung von X: die Wahrscheinlichkeit für eine Abweichung von $\mathbf{E}(X)$ um mindestens $c \cdot \sigma(X)$ beträgt höchstens $\frac{1}{c^2}$.

⟦ Setze $\varepsilon = c \cdot \sigma(X)$

$$\Rightarrow \quad \mathbf{P}\{|X - \mathbf{E}(X)| \geq c \cdot \sigma(X)\} \leq \frac{\mathbf{V}(X)}{c^2 \cdot (\sigma(X))^2} = \frac{1}{c^2}.$$



⟧

Einer der wichtigsten Sätze der Wahrscheinlichkeitsrechnung ist das von *Jakob Bernoulli* etwa 1688 entdeckte Gesetz der großen Zahl. Es macht eine Aussage über den Mittelwert $\overline{X_n} := \frac{1}{n}\sum_{i=1}^{n} X_i$ der ersten n identisch verteilten Messungen und den Erwartungswert $\mathbf{E}(X_i)$.

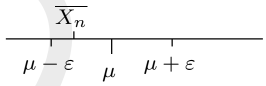
5.19 Satz (Schwaches Gesetz der großen Zahlen)

Es sei $X_1, X_2, \dots$ eine Folge von quadratisch integrierbaren und unkorrelierten Zufallsvariablen mit gleichem Erwartungswert $\mu = \mathbf{E}(X_n)$ und beschränkter Varianz $\mathbf{V}(X_n) \le M$ für alle $n \ge 1$. Wir setzen $\overline{X_n} := \frac{1}{n}\sum_{i=1}^{n} X_i$ für $n \ge 1$. Dann gilt für alle $\varepsilon > 0$

$$\mathbf{P}\left\{\left|\overline{X_n} - \mu\right| \ge \varepsilon\right\} \xrightarrow{n\to\infty} 0.$$

Beweis:
Es gilt

$$\mathbf{E}\left(\overline{X_n}\right) = \frac{1}{n}\sum_{i=1}^{n}\mathbf{E}(X_i) = \frac{1}{n}\cdot n\mu = \mu$$



und

$$\begin{aligned}\mathbf{V}\left(\overline{X_n}\right) &= \mathbf{V}\left(\frac{1}{n}\sum_{i=1}^{n} X_i\right) \underset{(5.9.e)}{=} \frac{1}{n^2}\cdot\mathbf{V}\left(\sum_{i=1}^{n} X_i\right) \\ &\underset{(5.12.c)}{=} \frac{1}{n^2}\cdot\sum_{i=1}^{n}\mathbf{V}(X_i) \le \frac{1}{n^2}\cdot n\cdot M = \frac{M}{n}.\end{aligned}$$

Mit der Tschebyscheff'schen Ungleichung (5.17) folgt dann

$$\begin{aligned}\mathbf{P}\left\{\left|\overline{X_n} - \mu\right| \ge \varepsilon\right\} &= \mathbf{P}\left\{\left|\overline{X_n} - \mathbf{E}\left(\overline{X_n}\right)\right| \ge \varepsilon\right\} \\ &\le \frac{1}{\varepsilon^2}\cdot\mathbf{V}\left(\overline{X_n}\right) \le \frac{M}{\varepsilon^2}\cdot\frac{1}{n} \xrightarrow{n\to\infty} 0\end{aligned}$$

□

Die in Satz (5.19) bewiesene Annäherung des Stichprobenmittels $\overline{X_n}$ an den gemessenen Erwartungswert der Mengen X_j bezeichnet man als stochastische Konvergenz:

5.20 Definition (Stochastische Konvergenz)

Es sei $(\Omega, \mathfrak{A}, \mathbf{P})$ ein Wahrscheinlichkeitsraum und $Y_n : \Omega \to \mathbb{R}$ Zufallsvariable. Weiter sei $Y : \Omega \to \mathbb{R}$ Zufallsvariable. Man sagt, daß Y_n gegen Y stochastisch konvergiert, falls für alle $\varepsilon > 0$

$$\mathbf{P}\left\{|Y_n - Y| \ge \varepsilon\right\} \xrightarrow{n\to\infty} 0.$$

Wir schreiben in diesem Fall $Y_n \xrightarrow{\mathbf{P}} Y$ für $n \to \infty$.

5.21 Bemerkungen

(a) Satz (5.19) besagt, daß $\overline{X_n} \xrightarrow{\mathbf{P}} \mu$ für $n \to \infty$.

(b) In seiner ursprünglichen Form behandelt das schwache Gesetz der großen Zahlen von *J. Bernoulli* den folgenden Fall:
Seien $A_1, A_2, \dots \in \mathfrak{A}$ unabhängige Ereignisse mit gleicher Wahrscheinlichkeit $p = \mathbf{P}(A_j)$ und K_n die Anzahl der Erfolge unter den ersten n Ereignissen, d.h.

$$K_n(\omega) = \#\{j \le n \mid \omega \in A_j\} = \sum_{j=1}^{n} 1_{A_j}(\omega)$$

für $\omega \in \Omega$.

Dann konvergiert die relative Häufigkeit $\frac{K_n}{n}$ stochastisch gegen die Erfolgswahrscheinlichkeit

$$p = \mathbf{P}(A_j) \qquad \text{d.h. } \frac{K_n}{n} \xrightarrow{\mathbf{P}} p \quad \text{für } n \to \infty.$$

⟦ Sei $X_j = 1_{A_j}$. Dann sind die X_j unabhängig und haben die Verteilung $\mathbf{P}_{X_j} = \mathcal{B}_{1,p}$
Nach (5.12.c) sind die X_j unkorreliert und es gilt

$$\mathbf{E}(X_j) = \mathbf{E}\left(1_{A_j}\right) = \mathbf{P}(A_j) = p$$

und nach (5.9.b) gilt

$$\mathbf{V}(X_j) = p \cdot (1-p) =: M.$$

Satz (5.19) liefert dann die Behauptung. ⟧

5.22 Anwendung (Numerische Integration/ Monte-Carlo-Methode)

Es sei $f : [0,1] \to \mathbb{R}$ meßbar (z.B. stetig). Gesucht ist eine Näherung für

$$\int_0^1 f(x)\, \mathbf{d}\lambda^1(x).$$

Ist f „glatt“, d.h. hinreichend oft differenzierbar, so ist

$$\int_0^1 f(x)\, \mathbf{d}\lambda^1(x) = \int_0^1 f(x)\, \mathbf{d}x$$

und für dieses Riemann-Integral stellt die Numerik Quadraturformeln zu näherungsweisen Berechnung des Integrals bereit. Ist f nur stetig oder besitzt f sogar viele Unstetigkeitsstellen, funktionieren die klassischen Quadraturformeln schlecht.
Das schwache Gesetz der großen Zahlen liefert eine andere Methode, die Monte-Carlo-Methode:
Es seien $X_j : \Omega \to [0,1]$ unkorrelierte Zufallsvariable mit Verteilung $\mathbf{P}_{X_j} = \lambda^1\big|_{[0,1]}$. Dann gilt, falls $\int_0^1 f^2(x)\, \mathbf{d}\lambda^1(x) < \infty$:

$$\frac{1}{n} \cdot \sum_{j=1}^n f(X_j) \xrightarrow{\mathbf{P}} \int_0^1 f(x)\, \mathbf{d}\lambda^1(x) \quad \text{für } n \to \infty.$$

⟦ Die Zufallsvariablen $Y_j = f(X_j)$ sind wegen der Transformationsformel (5.4)

$$\mathbf{E}\left(Y_j^2\right) = \int f^2(x)\, \mathbf{d}\mathbf{P}_{X_j}(x) = \int_0^1 f^2(x)\, \mathbf{d}\lambda^1(x) < \infty$$

quadratisch integrierbar und unkorreliert. Es ist

$$\mathbf{E}(Y_j) = \mathbf{E}(f(X_j)) = \int_0^1 f(x)\, \mathbf{d}\mathbf{P}_{X_j}(x) = \int_0^1 f(x)\, \mathbf{d}\lambda^1(x) = \mu.$$

Weiter ist

$$\mathbf{V}(Y_j) = \mathbf{E}\left(Y_j^2\right) - (\mathbf{E}(Y_j))^2 = \int_0^1 f^2(x)\, \mathbf{d}\lambda^1(x) - \left(\int_0^1 f(x)\, \mathbf{d}\lambda^1(x)\right)^2 =: M < \infty.$$

Nach Satz (5.19) gilt dann

$$\frac{1}{n} \cdot \sum_{j=1}^n Y_j = \frac{1}{n} \cdot \sum_{j=1}^n f(X_j) \xrightarrow{\mathbf{P}} \int_0^1 f(x)\, \mathbf{d}\lambda^1(x) = \mu.$$

Für den Fehler ergibt sich aus dem Beweis von Satz (5.19) die Fehlerabschätzung:

$$\mathbf{P}\left\{\left|\frac{1}{n} \cdot \sum_{j=1}^n f(X_j) - \int_0^1 f(x)\, \mathbf{d}\lambda^1(x)\right| \geq \varepsilon\right\} \leq \frac{1}{n} \cdot \frac{1}{\varepsilon^2} \cdot \left[\int_0^1 f^2(x)\, \mathbf{d}\lambda^1(x) - \left(\int_0^1 f(x)\, \mathbf{d}\lambda^1(x)\right)^2\right]$$

Aufgaben:

Aufgabe 5.1:

(a) Es sei X Poisson-verteilt mit Parameter $\lambda > 0$. Berechnen Sie $\mathbf{E}(X^m)$ für $m = 1, 2, 3$ und geben Sie eine Rekursionsformel für $m \geq 4$ an.

(b) Es sei X gleichverteilt im Intervall $[a, b]$ $(a < b)$. Berechnen Sie $\mathbf{E}(X)$ und $\mathbf{E}\left(X^2\right)$.

(c) Ein Betrieb stellt Fußbälle her. Produktionsbedingt schwankt der Radius R der Fußbälle, R sei gleichverteilt auf $[r - \varepsilon, r + \varepsilon]$. Berechnen Sie den mittleren Radius, die mittlere Oberfläche und das mittlere Volumen der Fußbälle.

Aufgabe 5.2:

Die Zufallsvariable X_i bezeichne das Ereignis des i-ten Ziehens beim n-fachen Ziehen ohne Zurücklegen aus einer Urne, die M schwarze und $N - M$ gelbe Kugeln enthalte; $N, M \in \mathbb{N}$. Zeigen Sie:

(a)

$$\mathbf{P}\{X_1 = i_1, \ldots, X_n = i_n\} = \frac{M \cdot \ldots \cdot (M - r + 1) \cdot (N - M) \cdot \ldots \cdot (N - M - s + 1)}{N \cdot (N - 1) \cdot \ldots \cdot (N - n + 1)}$$

wobei $i_j \in \{0, 1\} \mathrel{\hat{=}} \{\text{gelb}, \text{schwarz}\}$ für $1 \leq j \leq n$ und $r = \sum i_j$, $s = \sum (1 - i_j)$.

(b) Alle X_i sind identisch verteilt.

(c) Bestimmen Sie den Erwartungswert der hypergeometrischen Verteilung $\mathcal{H}_{n,N,M}$ und vergleichen Sie ihn mit dem Erwartungswert, den man beim Ziehen mit Zurücklegen aus einer Urne gleichen Typs erhält.

Aufgabe 5.3:

Wir betrachten ein aus mehreren Komponenten bestehendes Aggregat. Die Lebensdauer der einzelnen Komponenten (ab ihrem Einschalten) seien unabhängige Zufallsvariable mit der gleichen Exponentialverteilung $\mathcal{E}_\lambda$. Berechen Sie in den folgenden Situationen jeweils den Erwartungswert der Lebensdauer des Aggregats.

(a) Es gibt zwei Komponenten, wobei die zweite Komponente erst dann eingeschaltet wird, wenn die erste ausfällt (Aggregat mit „kalter Reserve“).

(b) Aggregat mit zwei Komponenten, die gleichzeitig eingeschaltet sind und parallel arbeiten („heiße Reserve“).

(c) Zur Erhöhung der Sicherheit bei (b) wird eine dritte Komponente aktiviert, sobald eine der ersten beiden ausfällt. Diese arbeitet dann parallel zur verbliebenen. (*Hinweis*: Zeigen Sie, daß für unabhängige $\mathcal{E}_\lambda$-verteilte X, Y die Zufallsvariable $\max\{X, Y\} - \min\{X, Y\}$ auch die Verteilung $\mathcal{E}_\lambda$ hat.)

Aufgabe 5.4:

X, Y seien quadratisch integrierbare Zufallsvariable. Zeigen Sie:

(a) Aus $X \leq Y$ und $\mathbf{E}(X) = \mathbf{E}(Y)$ folgt $\mathbf{P}\{X = Y\} = 1$.

(b) Varianz ist genau dann $\mathbf{V}(X) = 0$, wenn ein $c \in \mathbb{R}$ existiert mit $\mathbf{P}\{X = c\} = 1$.

(c) Ist $\operatorname{Kor}(X, Y) = \pm 1$, so existieren Konstanten $a, b \in \mathbb{R}$ mit $a \neq 0$, so daß

$$\mathbf{P}\{y = aX + b\} = 1.$$

Aufgabe 5.5:

$(X_n)_{n\in\mathbb{N}}$ sei eine Folge quadratisch integrierbarer Zufallsvariblen mit $\mathbf{E}(X_n) = \mu$, $\mathbf{V}(X_n) \leq c$ für alle $n \in \mathbb{N}$ mit $\mu \in \mathbb{R}$ und $c \geq 0$, sowie $\operatorname{Kov}(X_i, X_j) \to 0$ für $|i - j| \to \infty$. Zeigen Sie, daß für die Folge (X_n) das schwache Gesetz der großen Zahlen gilt, d.h.

$$\mathbf{P}\left\{\left|\frac{1}{n}\sum_{i=1}^{n} X_i - \mu\right| > \varepsilon\right\} \to \infty$$

für alle $\varepsilon > 0$.

Aufgabe 5.6:

Bestimmen Sie für die näherungsweise Berechnung des Integrals

$$\int_{-1}^{1} \sin(\pi x)\, \mathbf{d}x$$

mit der Monte-Carlo Methode ein n, so daß mit einer Wahrscheinlichkeit von wenigstens 95% der Fehler kleiner als $0,005$ ist. (*Hinweis*: Schätzen Sie den Fehler mit der Tschebyscheff-Ungleichung ab.)

Aufgabe 5.7:

Ein Schwarzhändler kauft bei Karten Krause die Karten für n EUR, die er samstags für $n + m$ EUR weiterverkauft. Manchmal bleibt er auf Karten sitzen, an anderen Tagen hätte er viele Karten mehr verkaufen können.

Wie viele Karten sollte er einkaufen, um seinen mittleren Gewinn zu maximieren, wenn die zufällige Anzahl seiner Kunden geometrisch verteilt mit Parameter $p \in \,]0, 1[$ ist?

Aufgabe 5.8:

$X : \Omega \to \mathbb{R}$ sei gleichverteilt in $[-1, 1]$. Zeigen Sie:

(a) $(Y_k := \sin(k\pi X))_{k\geq 1}$ sind paarweise unkorreliert.

(b) $\frac{1}{N}\sum\limits_{k=1}^{N} \sin(k\pi X) \to 0$ stochastisch.

6. Kapitel: Approximationen der Binomialverteilung

Die für viele Anwendungen wichtige Binomialverteilung $\mathcal{B}_{n,p}$, definiert durch

$$\begin{aligned}\mathcal{B}_{n,p}\{k\} &= \binom{n}{k} \cdot p^k \cdot (1-p)^{n-k} \\ &= \frac{n!}{k! \cdot (n-k)!} \cdot p^k \cdot (1-p)^{n-k} \quad \text{für } 0 \leq k \leq n\end{aligned}$$

ist für große n und k schwer zu berechnen. Wir stellen im Folgenden zwei Möglichkeiten vor, um $\mathcal{B}_{n,p}$ näherungsweise zu berechnen.

I) Poisson-Approximation

Es seien $X_1, \ldots, X_n$ unabhängige Zufallsvariable mit Verteilungen $\mathcal{B}_{1,p_j}$ $(p_j \in [0,1])$ und $S := \sum_{j=1}^{n} X_j$; Π_λ Poissonverteilung mit Parameter $\lambda > 0$, d.h. $\Pi_\lambda\{k\} = \mathbf{e}^{-\lambda} \cdot \frac{\lambda^k}{k!}$.

Interpretation:

$$\begin{aligned} X_j &\mathrel{\hat{=}} \text{Fälligwerden einer Versicherungspolice} \\ & \quad\ \text{(passiert mit Wahrscheinlichkeit } p_j) \\ S &\mathrel{\hat{=}} \text{Anzahl der eingetretenen Versicherungsfälle}\end{aligned}$$

6.1 Satz

Es seien X_j für $1 \leq j \leq n$ unabhängige $\mathcal{B}_{1,p_j}$-verteilte Zufallsvariable und

$$S = \sum_{j=1}^{n} X_j \qquad \lambda = \sum_{j=1}^{n} p_j.$$

Dann gilt:

$$\sum_{k=0}^{\infty} \left|\mathbf{P}\{S=k\} - \Pi_\lambda\{k\}\right| \leq 2 \cdot \sum_{j=1}^{n} p_j^2.$$

Beweis:

Für alle j sei $\Omega_j = \{-1, 0\} \cup \mathbb{N}$, $\mathfrak{A}_j = \mathcal{POT}(\Omega_j)$ und das Wahrscheinlichkeitsmaß $\mathbf{P}_j$ durch

$$\mathbf{P}_j\{\omega\} := \begin{cases} \mathbf{e}^{-p_j} - (1-p_j) & \omega = -1 \\ 1 - p_j & \omega = 0 \\ \mathbf{e}^{-p_j} \cdot \frac{p_j^\omega}{\omega!} & \omega \geq 1 \end{cases}$$

definiert. Auf dem Produktraum $\Omega = \Omega_1 \times \ldots \times \Omega_n$, $\mathfrak{A} = \mathcal{POT}(\Omega)$ verwenden wir das Produktmaß $\mathbf{P} = \mathbf{P}_1 \otimes \ldots \otimes \mathbf{P}_n$, das durch

$$\mathbf{P}\{(\omega_1, \ldots, \omega_n)\} := \mathbf{P}_1\{\omega_1\} \cdot \ldots \cdot \mathbf{P}_n\{\omega_n\}$$

definiert ist. Weiter sei $\overline{X_j} : \Omega \to \{0, 1\}$ durch

$$\overline{X_j}(\omega) = \begin{cases} 0 & \omega_j = 0 \\ 1 & \omega_j \neq 0 \end{cases}$$

definiert. Nach (4.14) und (4.15.a) sind die $\overline{X_j}$ $(1 \leq j \leq n)$ unabhängig und da

$$\begin{aligned} &\mathbf{P}\left\{\overline{X_j} = 0\right\} = \mathbf{P}_j\{0\} = 1 - p_j \\ \Rightarrow \quad &\mathbf{P}_{\overline{X_j}} = \mathcal{B}_{1,p_j} = \mathbf{P}_{X_j}.\end{aligned}$$

Sei $\overline{S} := \sum\limits_{j=1}^{n} \overline{X_j}$

$$\Rightarrow \quad \mathbf{P}_S = \mathbf{P}_{\overline{S}}.$$

Weiter seien $Y_j : \Omega \to \mathbb{N}_0$,

$$Y_j(\omega) = \begin{cases} 0 & \omega_j \leq 0 \\ \omega_j & \omega_j \geq 1 \end{cases}$$

$\Rightarrow Y_j$ $(1 \leq j \leq n)$ sind unabhängig und

$$\mathbf{P}\{Y_j = k\} = \mathbf{P}_j\{k\} = \mathbf{e}^{-p_j} \cdot \frac{p_j^k}{k!}$$

für $k \geq 1$ und

$$\mathbf{P}\{Y_j = 0\} = \mathbf{P}_j\{-1, 0\} = \mathbf{e}^{-p_j} - (1-p_j) + (1-p_j) = \mathbf{e}^{-p_j}$$

$$\Rightarrow \quad \mathbf{P}_{Y_j} = \Pi_{p_j}$$

Sei $T = \sum\limits_{j=1}^{n} Y_j$

$$\underset{\substack{(4.19.b)\\ \&\ \text{Induktion}}}{\Rightarrow} \quad \mathbf{P}_T = \Pi_\lambda \quad \text{mit } \lambda = \sum_{j=1}^{n} p_j$$

$$\begin{aligned}
\Rightarrow \quad & \sum_{k=0}^{\infty} \left|\mathbf{P}\{S = k\} - \Pi_\lambda\{k\}\right| \\
= \quad & \sum_{k=0}^{\infty} \left|\mathbf{P}\{\overline{S} = k\} - \mathbf{P}\{T = k\}\right| \\
= \quad & \sum_{k=0}^{\infty} \left|\mathbf{P}\{S = k = T\} + \mathbf{P}\{S = k \neq T\} - \mathbf{P}\{T = k = S\} - \mathbf{P}\{T = k \neq S\}\right| \\
\leq \quad & \sum_{k=0}^{\infty} \mathbf{P}\{S = k \neq T\} + \sum_{k=0}^{\infty} \mathbf{P}\{T = k \neq S\} \\
= \quad & 2 \cdot \mathbf{P}\{S \neq T\} \ = \ 2 \cdot \mathbf{P}\left\{\sum_{j=1}^{n} \overline{X_j} \neq \sum_{j=1}^{n} Y_j\right\} \\
\overset{X_j, Y_j \geq 0}{=} \quad & 2 \cdot \mathbf{P}\{X_j \neq Y_j \ \text{für mindestens ein} \ 1 \leq j \leq n\} \\
\leq \quad & 2 \cdot \sum_{j=1}^{n} \mathbf{P}\{X_j \neq Y_j\} \ = \ \clubsuit
\end{aligned}$$

Es ist

$$\begin{aligned}
\mathbf{P}\{X_j = Y_j\} &= \mathbf{P}\{\omega \in \Omega : \omega_j = 0 \ \text{oder} \ 1\} \\
&= \mathbf{P}_j\{0\} + \mathbf{P}_j\{1\} \ = \ 1 - p_j + p_j \cdot \mathbf{e}^{-p_j}
\end{aligned}$$

$$\begin{aligned}
\Rightarrow \quad \mathbf{P}\{X_j \neq Y_j\} &= 1 - \mathbf{P}\{X_j = Y_j\} \\
&= p_j \cdot \left(1 - \mathbf{e}^{-p_j}\right) \ \leq \ p_j^2 \quad \text{da } 1 - \mathbf{e}^{-x} \leq x \ \text{für } x \in [0,1]
\end{aligned}$$

$$\Rightarrow \quad \clubsuit \ \leq \ 2 \cdot \sum_{j=1}^{n} p_j^2$$

□

6.2 Korollar

Ist $(p(n))_{n\geq 1}$ eine Folge in $[0,1]$ mit $n \cdot p(n) \xrightarrow{n\to\infty} \lambda > 0$, so gilt für alle $k \in \mathbb{N}_0$:

$$\mathcal{B}_{n,p(n)}\{k\} \to \Pi_\lambda\{k\}$$

Beweis:
In (6.1) setzen wir $p_j = p(n)$ für $1 \leq j \leq n$

$$\Rightarrow \quad \mathbf{P}\{S = k\} = \mathcal{B}_{n,p(n)}\{k\}$$

$$\begin{aligned}\Rightarrow \quad \left|\mathcal{B}_{n,p(n)}\{k\} - \Pi_\lambda\{k\}\right| &\leq \left|\mathcal{B}_{n,p(n)}\{k\} - \Pi_{n\cdot p(n)}\{k\}\right| + \left|\Pi_{n\cdot p(n)}\{k\} - \Pi_\lambda\{k\}\right| \\ &\leq 2\cdot\sum_{j=1}^n p_j^2 + \underbrace{\left|\Pi_{p(n)}\{k\} - \Pi_\lambda\{k\}\right|}_{\xrightarrow{n\to\infty} 0}\end{aligned}$$

Wegen

$$2\cdot\sum_{j=1}^n p_j^2 = 2\cdot n\cdot p(n)^2 = 2\cdot\underbrace{n\cdot p(n)}_{\xrightarrow{n\to\infty}\lambda}\cdot\frac{n\cdot p(n)}{n} \xrightarrow{n\to\infty} 0$$

folgt die Behauptung. □

6.3 Beispiel

Im Hörsaal sitzen etwa 100 Studenten. Die Wahrscheinlichkeit, daß heute eine bestimmte Person Geburtstag hat ist $\frac{1}{365}$.

$$\Rightarrow \quad S \mathrel{\hat{=}} \{\text{„Anzahl der Personen im Hörsaal die heute Geburtstag haben.“}\}$$

ist $\mathcal{B}_{100,\frac{1}{365}}$-verteilt. Sei $\lambda = \frac{100}{365} \approx 0,2739$.

$$(6.2) \quad \Rightarrow \quad \mathcal{B}_{100,\frac{1}{365}}\{k\} \approx \Pi_{\frac{100}{365}}\{k\}.$$

Für $k = 0, 1, 2$ ergeben sich

$$\Pi_\lambda\{0\} \approx 0,76035 \qquad \Pi_\lambda\{1\} \approx 0,2083 \qquad \Pi_\lambda\{2\} \approx 0,02853.$$

II) Zentraler Grenzwertsatz

Die Poisson-Approximation funktioniert nur, wenn $n\cdot p$ „relativ“ klein ist, also p „sehr“ klein. Für „größere“ Werte von p bietet sich eine Version des zentralen Grenzwertsatzes an. Dazu:

6.4 Faktum (*Stirling*'sche Formel)

Für $n \geq 1$ gilt

$$\sqrt{2\pi n}\cdot\left(\frac{n}{\mathbf{e}}\right)^n \leq n! \leq \sqrt{2\pi n}\cdot\left(\frac{n}{\mathbf{e}}\right)^n\cdot\mathbf{e}^{\frac{1}{12n}}.$$

Beweis:
Siehe Heuser, Analysis I, S. 501 – 502 oder Kaballo, Analysis I, 36.13.

6.5 Definition

Es seien $t_n, u_n > 0$. Falls $\lim\limits_{n\to\infty} \frac{t_n}{u_n} = 1$ so sagen wir, daß <u>t_n und u_n asymptotisch gleich</u> sind. Schreibweise:

$$t_n \sim u_n.$$

6.6 Bemerkung

(a) $\sim$ ist eine Äquivalenzrelation auf der Menge aller positiven Folgen.

(b)

$$(6.4) \quad \Rightarrow \quad n! \sim \sqrt{2\pi n} \cdot \left(\frac{n}{\mathbf{e}}\right)^n$$

(c) Es gilt

$$\mathcal{B}_{2n,\frac{1}{2}}\{n\} \sim \frac{1}{\sqrt{\pi n}}.$$

⟦

$$\begin{aligned}
\mathcal{B}_{2n,\frac{1}{2}}\{n\} &= \binom{2n}{n} \cdot \left(\frac{1}{2}\right)^n \cdot \left(1 - \frac{1}{2}\right)^{2n-n} \\
&= \frac{(2n)!}{(n!)^2} \cdot 2^{-2n} \\
&\sim \frac{\sqrt{2\pi 2n} \cdot \left(\frac{2n}{\mathbf{e}}\right)^{2n}}{\left(\sqrt{2\pi n} \cdot \left(\frac{n}{\mathbf{e}}\right)^n\right)^2} \cdot 2^{-2n} \\
&= \frac{\sqrt{\pi n}}{\pi n} = \frac{1}{\sqrt{\pi n}}
\end{aligned}$$

⟧

6.7 Satz (Lokaler zentraler Grenzwertsatz)

Es seien $0 < p < 1$ und (a_n), (b_n) Folgen mit

$$\frac{a_n - np}{n^{\frac{2}{3}}} \xrightarrow{n\to\infty} 0 \quad \text{und} \quad \frac{b_n - np}{n^{\frac{2}{3}}} \xrightarrow{n\to\infty} 0.$$

Dann gilt

$$\mathcal{B}_{n,p}\{k\} \sim \frac{1}{\sqrt{2\pi np(1-p)}} \cdot \exp\left(-\frac{(k-np)^2}{2np(1-p)}\right)$$

für $n \to \infty$, gleichmäßig für alle $a_n \leq k \leq b_n$.
Beweis:
Sei

$$Q_n\{k\} = \frac{\mathcal{B}_{n,p}\{k\}}{\exp\left(-\frac{(k-np)^2}{2np(1-p)}\right)} \cdot \sqrt{2\pi np(1-p)}.$$

Dann ist die Aussage äquivalent zu:

$$\max_{a_n \leq k \leq b_n} |Q_n\{k\} - 1| \xrightarrow{n\to\infty} 0.$$

Sei nun $a_n \leq k_n \leq b_n$ so gewählt, daß $Q_n\{k_n\} = \max\limits_{a_n \leq k \leq b_n} Q_n\{k\}$. Dann ist zu zeigen, daß

$$Q_n\{k_n\} \xrightarrow{n\to\infty} 1.$$

Es gilt

$$\frac{k_n - np}{n^{\frac{2}{3}}} \to 0 \quad \text{für } n \to \infty.$$

Sei $q = 1 - p$, $x_n = \dfrac{k_n - np}{\sqrt{npq}}$

$$\Rightarrow \quad k_n = np + \sqrt{npq} \cdot x_n$$

und es gilt $\dfrac{x_n}{n^{\frac{1}{6}}} = \dfrac{k_n - np}{n^{\frac{2}{3}}} \cdot \dfrac{1}{n^{-\frac{1}{2}} \cdot \sqrt{npq}} = \underbrace{\dfrac{k_n - np}{n^{\frac{2}{3}}}}_{n\to\infty} \cdot \dfrac{1}{\sqrt{pq}} \xrightarrow{n\to\infty} 0.$

$$\Rightarrow \quad \frac{k_n}{n} = p + \frac{\sqrt{pq} \cdot x_n}{n^{\frac{1}{2}}} \xrightarrow{n\to\infty} p,$$

also $\dfrac{k_n}{n} \sim p$.

Mit der Stirling'schen Formel (6.4) folgt:

$$\begin{aligned}
\mathcal{B}_{n,p}\{k_n\} &= \frac{n!}{k_n! \cdot (n-k_n)!} \cdot p^{k_n} \cdot q^{n-k_n} \\
&\sim \frac{\sqrt{2\pi n} \cdot \left(\frac{n}{\mathbf{e}}\right)^n \cdot p^{k_n} \cdot q^{n-k_n}}{\sqrt{2\pi k_n} \cdot \left(\frac{k_n}{\mathbf{e}}\right)^{k_n} \cdot \sqrt{2\pi(n-k_n)} \cdot \left(\frac{n-k_n}{\mathbf{e}}\right)^{n-k_n}} \\
&= \sqrt{\frac{n}{2\pi k_n (n-k_n)}} \cdot \left(\frac{np}{k_n}\right)^{k_n} \cdot \left(\frac{nq}{n-k_n}\right)^{n-k_n}.
\end{aligned}$$

Nun gilt

$$\begin{aligned}
\sqrt{\frac{n}{2\pi k_n (n-k_n)}} &= \frac{1}{\sqrt{2\pi n \frac{k_n}{n}\left(1 - \frac{k_n}{n}\right)}} \\
&\sim \frac{1}{\sqrt{2\pi np(1-p)}}.
\end{aligned}$$

Sei nun $g :]0,1[\to \mathbb{R};\ g(t) = t \cdot \ln\left(\frac{t}{p}\right) + (1-t) \cdot \ln\left(\frac{1-t}{q}\right)$. Dann gilt:

$$\begin{aligned}
\exp\left(-n \cdot g\left(\frac{k_n}{n}\right)\right) &= \exp\left(-n \cdot \frac{k_n}{n} \cdot \ln\left(\frac{k_n}{np}\right) - n \cdot \left(1 - \frac{k_n}{n}\right) \cdot \ln\left(\frac{1 - \frac{k_n}{n}}{q}\right)\right) \\
&= \exp\left(-k_n \cdot \ln\left(\frac{k_n}{np}\right) - n \cdot \left(1 - \frac{k_n}{n}\right) \cdot \ln\left(\frac{n-k_n}{nq}\right)\right) \\
&= \left(\frac{np}{k_n}\right)^{k_n} \cdot \left(\frac{nq}{n-k_n}\right)^{n-k_n}.
\end{aligned}$$

Nun ist

$$\begin{aligned}
g(p) &= 0 \\
g'(t) &= \ln\left(\frac{t}{p}\right) + t \cdot \frac{1}{\frac{t}{p}} \cdot \frac{1}{p} - \ln\left(\frac{1-t}{q}\right) - (1-t) \cdot \frac{1}{\frac{1-t}{q}} \cdot \frac{1}{q} \\
&= \ln\left(\frac{t}{p}\right) - \ln\left(\frac{1-t}{q}\right) \\
\Rightarrow \quad g'(p) &= 0 \\
g''(t) &= \frac{1}{t} + \frac{1}{1-t} \\
\Rightarrow \quad g''(p) &= \frac{1}{p} + \frac{1}{1-p} = \frac{1}{p \cdot q}
\end{aligned}$$

$$\begin{aligned}
\text{Taylor} \quad \Rightarrow \quad g(t) &= g(p) + g'(p) \cdot (t-p) + \frac{1}{2} \cdot g''(p) \cdot (t-p)^2 + r(t-p) \\
&= \frac{(t-p)^2}{2pq} + r(t-p)
\end{aligned}$$

wobei $|r(t-p)| \leq c \cdot |t-p|^3$ für eine Konstante $c > 0$.

$$\Rightarrow \quad -n \cdot g\left(\frac{k_n}{n}\right) = -n \cdot \frac{\left(\frac{k_n}{n} - p\right)^2}{2pq} - n \cdot r\left(\frac{k_n}{n} - p\right).$$

Wegen

$$\frac{k_n}{n} = p + \sqrt{\frac{pq}{n}} \cdot x_n$$

ist

$$-n \frac{\left(\frac{k_n}{n} - p\right)^2}{2pq} = -\frac{x_n^2}{2}$$

und

$$\begin{aligned}
\left| n \cdot r\left(\frac{k_n}{n} - p\right) \right| &\leq c \cdot n \cdot \left| \sqrt{\frac{pq}{n}} \cdot x_n \right|^3 \\
&= c \cdot n \cdot \frac{\sqrt{pq}^3}{n^{\frac{3}{2}}} |x_n|^3 \\
&= \overline{c} \cdot \left(\frac{|x_n|}{n^{\frac{1}{6}}}\right)^3 \xrightarrow{n \to \infty} 0 \qquad \left(\overline{c} := c \cdot \sqrt{pq}^3\right)
\end{aligned}$$

$$\begin{aligned}
\Rightarrow \quad \left(\frac{np}{k_n}\right)^{k_n} \cdot \left(\frac{nq}{n-k_n}\right)^{n-k_n} &= \exp\left(-n \cdot g\left(\frac{k_n}{n}\right)\right) \\
&= \exp\left(-\frac{x_n^2}{2}\right) \cdot \exp\left(-n \cdot r\left(\frac{k_n}{n} - p\right)\right) \\
&\sim \exp\left(-\frac{x_n^2}{2}\right).
\end{aligned}$$

Zusammen mit der Asymptotik des ersten Faktors folgt nun die Behauptung. □

Aus Satz (6.7) folgt, daß der maximale Wert der Zähldichte von $\mathcal{B}_{n,p}$ in der Nähe von $n \cdot p$ liegt. Im folgenden sei (vgl. (3.20.b))

$$\Phi(x) = \mathcal{N}_{0,1}(]-\infty, x]) = \frac{1}{\sqrt{2\pi}} \cdot \int_{-\infty}^{x} \mathbf{e}^{-\frac{t^2}{2}} \, \mathbf{d}t$$

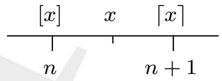
die Verteilungsfunktion der Standardnormalverteilung.

6.8 Satz (Zentraler Grenzwertsatz von *deMoivre* und *Laplace*)

Es sei $p \in \,]0,1[$ fest und für $n \geq 1$ sei S_n eine Zufallsvariable mit Verteilung $\mathcal{B}_{n,p}$. Dann gilt für alle $a < b$:

$$\lim_{n\to\infty} \mathbf{P}\left\{a \leq \frac{S_n - np}{\sqrt{np\,(1-p)}} \leq b\right\} = \Phi\,(b) - \Phi\,(a)\,.$$

Beweis:



Es seien $a_n = \lceil np + a\sqrt{npq}\rceil$ (aufrunden!) und $b_n = \lfloor np + b \cdot \sqrt{npq}\rfloor$ (Gaußklammer!) $\Rightarrow \dfrac{a_n - np}{n^{\frac{2}{3}}} \xrightarrow{n\to\infty} 0$, ebenso für b_n. Diese erfüllen die Voraussetzungen von Satz (6.7).

Dann gilt

$$\begin{aligned}
\mathbf{P}\left\{a \leq \frac{S_n - np}{\sqrt{npq}} \leq b\right\} &= \mathbf{P}\left\{a_n \leq S_n \leq b_n\right\} \\
&= \mathcal{B}_{n,p}\left\{a_n, a_n + 1, \ldots, b_n\right\} = \sum_{k=a_n}^{b_n} \mathcal{B}_{n,p}\left\{k\right\}.
\end{aligned}$$

Setze $x_{n,k} = \dfrac{k - np}{\sqrt{npq}}$.

$\underset{(6.7)}{\Rightarrow}$ Es gibt eine Nullfolge ε_n mit

$$\begin{aligned}
(1-\varepsilon_n) \cdot \frac{\exp\left(-\frac{x_{n,k}^2}{2}\right)}{\sqrt{2\pi npq}} &\leq \mathcal{B}_{n,p}\left\{k\right\} \\
&\leq (1+\varepsilon_n) \cdot \frac{\exp\left(-x_{n,k}^2\right)}{\sqrt{2\pi npq}}
\end{aligned}$$

für alle $a_n \leq k \leq b_n$.

Setze

$$\varphi\,(x) = \frac{1}{\sqrt{2\pi}} \cdot \exp\left(-\frac{x^2}{2}\right) \quad \text{und} \quad I_n = \sum_{k=a_n}^{b_n} \frac{1}{\sqrt{npq}} \cdot \varphi\,(x_{n,k})$$

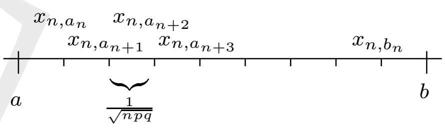
$$\Rightarrow \quad (1-\varepsilon_n) \cdot I_n \;\leq\; \mathbf{P}\left\{a \leq \frac{S_n - np}{\sqrt{npq}} \leq b\right\} \;\leq\; (1+\varepsilon_n) \cdot I_n.$$

Bleibt zu zeigen:

$$I_n \xrightarrow{n\to\infty} \Phi\,(b) - \Phi\,(a)\,.$$

Für n fest ist $x_{n,k}$ $(a_n \leq k \leq b_n)$ eine äquidistante Unterteilung des Intervalls $[a,b]$ mit Gitterabstand $\dfrac{1}{\sqrt{npq}}$,

$$\text{Analysis I} \quad \Rightarrow \quad I_n \to \int_a^b \varphi\,(t)\;\mathbf{d}t = \Phi\,(b) - \Phi\,(a)\,.$$



□

6.9 Beispiel (Wahlvorhersage)

Es soll die Mindestgröße n einer Stichprobe für die Vorhersage des Stimmanteils p einer Partei bestimmt werden, so daß die Wahrscheinlichkeit für einen Fehler der Vorhersage von höchstens 1% mindestens 95% beträgt.

Jeder Wähler aus der Stichprobe entscheidet sich unabhängig mit Wahrscheinlichkeit p für die Partei.

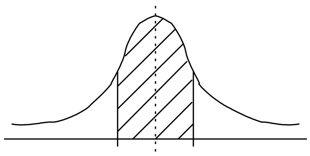
$\Rightarrow$ Verteilung der Stimmenzahl S_n der Partei ist $\mathcal{B}_{n,p}$ (bei unbekannten p).

Nach (5.20.2.b) ist $\frac{S_n}{n}$ eine „gute“ Schätzung von p. Die gesuchte Wahrscheinlichkeit beträgt

$$\begin{aligned}\mathbf{P}\left\{-0,01 \leq \frac{S_n}{n} - p \leq 0,01\right\} &= \mathbf{P}\left\{\frac{-0,01\sqrt{n}}{\sqrt{pq}} \leq \frac{S_n - np}{\sqrt{npq}} \leq \frac{0,01\sqrt{n}}{\sqrt{pq}}\right\} \\ &\approx \Phi\left(\frac{0,01\sqrt{n}}{\sqrt{pq}}\right) - \Phi\left(\frac{-0,01\sqrt{n}}{\sqrt{pq}}\right)\end{aligned}$$

Nun gilt (wegen der Monotonie, nach Tabelle)

$$\Phi(x) - \Phi(-x) \geq 0,95 \quad \Leftrightarrow \quad x \geq \Phi^{-1}(0,975) \approx 1,96$$



$$\Rightarrow \quad \frac{0,01 \cdot \sqrt{n}}{\sqrt{pq}} \geq 1,96 \quad \Leftrightarrow \quad n \geq 196^2 \cdot pq$$

Der maximale Wert von pq ist $\frac{1}{4}$ $\left(p = q = \frac{1}{2}\right)$

$$\Rightarrow \quad n \geq 9604$$

6.10 Korollar

Es gilt für $S_n^* = \frac{S_n - np}{\sqrt{np(1-p)}}$ und $b \in \mathbb{R}$, daß

$$\begin{aligned}\mathbf{P}\{S_n^* \leq b\} &\rightarrow \Phi(b) \quad \text{und} \\ \mathbf{P}\{S_n^* \geq b\} &\rightarrow 1 - \Phi(b).\end{aligned}$$

Beweis:

Für jedes $a < b$ gilt

$$\begin{aligned}\liminf_{n\to\infty} \mathbf{P}\{S_n^* \leq b\} &\geq \liminf_{n\to\infty} \mathbf{P}\{a \leq S_n^* \leq b\} \\ &= \Phi(b) - \Phi(a).\end{aligned}$$

Da Φ eine Verteilungsfunktion ist, gilt $\lim\limits_{a\to-\infty} \Phi(a) = 0$.

$$\Rightarrow \quad \liminf_{n\to\infty} \mathbf{P}\{S_n^* \leq b\} \geq \Phi(b).$$

Für $c > b$ gilt

$$\begin{aligned}\limsup_{n\to\infty} \mathbf{P}\{S_n^* \leq b\} &= 1 - \liminf_{n\to\infty} \mathbf{P}\{S_n^* > b\} \\ &\leq 1 - \liminf_{n\to\infty} \mathbf{P}\{b < S_n^* \leq c\} \\ &= 1 - \Phi(c) + \Phi(b).\end{aligned}$$

Da $\lim\limits_{c\to\infty} \Phi(c) = 1$ folgt

$$\limsup_{n\to\infty} \mathbf{P}\{S_n^* \leq b\} \leq \Phi(b)$$

und daraus folgt die Behauptung. □

6.11 Beispiel

Wahl mit zwei Parteien ABC und XYZ. Wahlprogramme so ähnlich, daß jeder der n Wähler mit Wahrscheinlichkeit $\frac{1}{2}$ für eine der beiden Parteien stimmt.

Gesucht:

Minimale Anzahl m der Wähler, die wir bestechen müssen, so daß wir mit mindestens 90% Sicherheit die Mehrheit der Stimmen erhalten.

Da sich die nicht bestochenen Wähler mit Wahrscheinlichkeit $\frac{1}{2}$ für unsere Partei entscheiden und wir schon m Stimmen sicher haben, genügen uns $\frac{n}{2} - m$ zusätzliche Stimmen.

Die Wahrscheinlichkeit, daß wir mindestens so viele erhalten ist

$$\mathcal{B}_{n-m,\frac{1}{2}}\left\{\frac{n}{2} - m, \ldots, n-m\right\} \overset{!}{\geq} 0,9.$$

(6.10) $\Rightarrow$ S_{n-m} sei $\mathcal{B}_{n-m,\frac{1}{2}}$-verteilt:

$$\begin{aligned}
\mathbf{P}\left\{S_{n-m} \geq \frac{n}{2} - m\right\} &= \mathbf{P}\left\{\frac{S_{n-m} - \frac{n-m}{2}}{\sqrt{(n-m)\cdot\frac{1}{2}\cdot\frac{1}{2}}} \geq \frac{-m}{\sqrt{n-m}}\right\} \\
&\approx 1 - \Phi\left(\frac{-m}{\sqrt{n-m}}\right) \overset{!}{\geq} 0,9.
\end{aligned}$$

$$\Leftrightarrow \quad \Phi\left(\frac{-m}{\sqrt{n-m}}\right) \leq 0,1.$$

Tabelle:

$$\Phi(-1,28) \approx 0,1.$$

Also

$$\frac{-m}{\sqrt{n-m}} \leq -1,28 \quad \Leftrightarrow \quad m \geq 1,28 \cdot \sqrt{n-m} \approx 1,28 \cdot \sqrt{n},$$

da m sehr viel kleiner als n ist.

$$n = 5 \cdot 10^7 \quad \Rightarrow \quad m \geq 9051,$$

das heißt es genügt $0,018\%$ der Wähler zu bestechen.

Aufgaben:

Aufgabe 6.1:

Seien $n \geq 1$ und $p \in \left]0,1\right[$. Finden Sie eine ganze Zahl k_0 derart, daß die Zähldichte der Binomialverteilung $\mathcal{B}_{n,p}$ auf $\{0,\ldots,k_0\}$ monoton steigend und auf $\{k_0, k_0+1,\ldots,n\}$ monoton fallend ist. Für welche n, p ist k_0 eindeutig bestimmt?

Aufgabe 6.2:

Es seien $p_n \in \left]0,1\right]$ reelle Zahlen mit $\lim_{n\to\infty} n\,(1-p_n) = \lambda$. Zeigen Sie die folgende Approximation der negativen Binomialverteilung durch die Poisson-Verteilung:

$$\lim_{n\to\infty} \mathcal{NB}_{n,p_n}\{k\} = \Pi_\lambda\{k\}$$

für alle $k \in \mathbb{N}$.

Hinweis: Aus $nx_n \to \lambda$ folgt $(1+x_n)^n \to \mathbf{e}^\lambda$.

Aufgabe 6.3:

Bei einem Spiel werden insgesamt $2n$ Runden gespielt, wobei jeder der beiden Spieler bei jeder Runde (unabhängig von dem anderen) mit Wahrscheinlichkeit $\frac{1}{2}$ gewinnt. Z_{2n} sei die Zahl der Runden, die der bessere Spieler mehr gewonnen hat als der schlechtere. Zeigen Sie, daß für jede Folge (j_n) in $2\mathbb{N}$, $j_n \to \infty$, mit $j_n \leq C\sqrt{n}$ gilt

$$\sqrt{\pi n}\cdot \mathbf{P}\{Z_{2n} = j_n\} \sim 2\cdot \mathbf{e}^{\frac{-j_n^2}{4n}}.$$

Aufgabe 6.4:

In einem Gefäß befinden sich 10^{22} Gasmoleküle, die sich jeweils mit Wahrscheinlichkeit $\frac{1}{2}$ in der rechten beziehungsweise linken Seite des Behälters aufhalten. Bestimmen Sie mit dem Satz von *deMoivre* und *Laplace* approximativ die Wahrscheinlichkeit dafür, daß sich die Zahlen der Moleküle in den beiden Seiten um mehr als $10^{-9}\%$ der Gesamtzahl unterscheiden.

Aufgabe 6.5:

Es seien (a_n), (b_n) Folgen mit $a_n > 0$, $b_n > 0$ für alle $n \in \mathbb{N}$.

$$a_n \sim b_n :\Leftrightarrow \lim_{n\to\infty} \frac{a_n}{b_n} = 1$$

(a_n und b_n sind asymptotisch gleich). Zeigen Sie

(a) $\sim$ ist eine Äquivalenzrelation;

(b) $a_n \sim b_n,\ c_n \sim d_n \Rightarrow a_n c_n \sim b_n d_n$;

(c) $a_n \sim b_n \Leftrightarrow \dfrac{1}{a_n} \sim \dfrac{1}{b_n}$.

7. Kapitel: Bedingte Wahrscheinlichkeiten

Wenn wir zusätzliche Informationen über ein Experiment haben, ändert sich die Wahrscheinlichkeit mit der wir ein bestimmtes Ereignis einschätzen. Ein extremes Beispiel ist, wenn wir bereits wissen, daß ein Münzwurf Wappen ergeben hat, dann werden wir auf Grund dieser Information $\mathbf{P}\{W\} = 1$ annehmen.

Beispiel:

Hat jemand beim Skat 3 Buben auf der Hand, so wird man die Wahrscheinlichkeit, daß im Stock noch ein Bube liegt, nicht mehr so hoch einschätzen, wie vor dem Aufheben der Karten ($\frac{7}{31}$ vor dem Aufheben, $\frac{1}{11}$ danach).

Sei $(\Omega, \mathfrak{A}, \mathbf{P})$ ein Wahrscheinlichkeitsraum:

$$A \in \mathfrak{A} \qquad \overline{Q}(A) := \mathbf{P}(A \cap B)$$

$\overline{Q}$ ist kein Wahrscheinlichkeitsmaß, da $\overline{Q}(\Omega) = \mathbf{P}(B) < 1$ möglich. Besser ist:

$$\tilde{Q}(A) := \frac{\mathbf{P}(A \cap B)}{\mathbf{P}(B)} \quad \Rightarrow \quad \tilde{Q}(B) = \frac{\mathbf{P}(B)}{\mathbf{P}(B)} = 1.$$

Dieses Phänomen wird durch den Begriff der bedingten Wahrscheinlichkeit beschrieben:

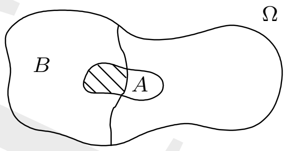
7.1 Definition (Bedingte Wahrscheinlichkeit)

Es seien $(\Omega, \mathfrak{A}, \mathbf{P})$ ein Wahrscheinlichkeitsraum und $A, B \in \mathfrak{A}$ mit $\mathbf{P}(B) > 0$. Dann heißt die Zahl

$$\mathbf{P}(A \mid B) := \frac{\mathbf{P}(A \cap B)}{\mathbf{P}(B)}$$

die bedingte Wahrscheinlichkeit von A unter B.

Sind $X : \Omega \to \Omega_0$ und $Y : \Omega \to \Omega_1$ Zufallsvariable, sowie $A = \{X \in A'\}$, $B = \{Y \in B'\}$ so schreiben wir $\mathbf{P}(A \mid B) = \mathbf{P}(X \in A' \mid Y \in B')$.



7.2 Beispiele

(a) Aus einem Skatspiel werden nacheinander zwei Karten gezogen (ohne Zurücklegen)

$$\begin{array}{rcl} A & \hat{=} & \text{„die zweite Karte ist ein As“} \\ B & \hat{=} & \text{„die erste Karte ist ein As“} \\ B^{\complement} & \hat{=} & \text{„die erste Karte ist kein As“} \end{array}$$

$$\Rightarrow \quad \mathbf{P}(A \mid B) = \frac{3}{31} \approx 0{,}097 \qquad \mathbf{P}\left(A \mid B^{\complement}\right) = \frac{4}{31} \approx 0{,}129.$$

Modell:

$$\Omega = \mathcal{P}_2^{\{1,\ldots,32\}} \qquad \mathbf{P} = \mathcal{L}_\Omega \qquad \mathfrak{A} = \mathcal{POT}(\Omega)$$

Asse: 1, 9, 17, 25.

$$\begin{array}{rcl} A & = & \{(\omega_1, \omega_2) \in \Omega \mid \omega_2 \in \{1, 9, 17, 25\}\} \\ B & = & \{(\omega_1, \omega_2) \in \Omega \mid \omega_1 \in \{1, 9, 17, 25\}\} \\ B^{\complement} & = & \{(\omega_1, \omega_2) \in \Omega \mid \omega_2 \notin \{1, 9, 17, 25\}\} \end{array}$$

Es ist

$$\#\Omega = (32)_2 = 32 \cdot 31 = 992 \qquad \#B = 4 \cdot 31 = 124 \qquad \#B^{\complement} = 28 \cdot 31 = 868$$

$$\#(A \cap B) = 4 \cdot 3 = 12 \qquad \#\left(A \cap B^{\complement}\right) = 28 \cdot 4 = 112$$

$\Rightarrow$ Behauptung. □

(b) Ein Experiment mit Erfolgswahrscheinlichkeit $p \in]0,1[$ wird n-mal unabhängig wiederholt.

$$\begin{array}{lcl} S & \hat{=} & \text{Anzahl der Erfolge} \\ T & \hat{=} & \text{Nummer des Experiments, bei dem zum ersten Mal Erfolg eintritt} \end{array}$$

Falls kein Erfolg eintritt, wird $T = n+1$ gesetzt.
Dann gilt (die plausible Tatsache)

$$\mathbf{P}\{T = k \mid S = 1\} = \frac{1}{n}$$

für $1 \leq k \leq n$.
Es gilt

$$\begin{aligned} \mathbf{P}_S = \mathcal{B}_{n,p} \quad \Rightarrow \quad \mathbf{P}\{S = 1\} &= \mathcal{B}_{n,p}\{1\} = \binom{n}{1} \cdot p^1 \cdot (1-p)^{n-1} \\ &= n \cdot p \cdot (1-p)^{n-1} \\ \mathbf{P}\{T = k,\, S = 1\} &= p \cdot (1-p)^{n-1} \end{aligned}$$

$$\Rightarrow \quad \mathbf{P}\{T = k \mid S = 1\} = \frac{\mathbf{P}\{T = k,\, S = 1\}}{\mathbf{P}\{S = 1\}} = \frac{p \cdot (1-p)^{n-1}}{n \cdot p \cdot (1-p)^{n-1}} = \frac{1}{n}$$

□

7.3 Eigenschaften

Es sei $B \in \mathfrak{A}$ mit $\mathbf{P}(B) > 0$.

(a) Durch $\mathfrak{A} \ni C \mapsto \mathbf{P}(C \mid B)$ wird ein Wahrscheinlichkeitsmaß auf $(\Omega, \mathfrak{A})$ definiert mit $\mathbf{P}(B \mid B) = 1$ und $\mathbf{P}\left(B^{\complement} \mid B\right) = 0$.
⟦ Sei

$$Q(C) = \mathbf{P}(C \mid B) \quad \Rightarrow \quad Q(\Omega) = 1$$

d.h. (WM_2) aus (1.17) ist erfüllt.
Sind (C_n) disjunkte Mengen aus $\mathfrak{A} \Rightarrow C_n \cap B$ disjunkt.

$$\begin{aligned} Q\left(\bigcup_{n=1}^{\infty} C_n\right) &= \frac{1}{\mathbf{P}(B)} \cdot \mathbf{P}\left(B \cap \bigcup_{n=1}^{\infty} C_n\right) = \frac{1}{\mathbf{P}(B)} \cdot \mathbf{P}\left(\bigcup_{n=1}^{\infty} (B \cap C_n)\right) \\ &= \frac{1}{\mathbf{P}(B)} \cdot \sum_{n=1}^{\infty} \mathbf{P}(C_n \cap B) = \sum_{n=1}^{\infty} \frac{\mathbf{P}(C_n \cap B)}{\mathbf{P}(B)} \\ &= \sum_{n=1}^{\infty} \mathbf{P}(C_n \mid B) = \sum_{n=1}^{\infty} Q(C_n) \end{aligned}$$

⟧

(b) A und B sind genau dann unabhängig, falls $\mathbf{P}(A \mid B) = \mathbf{P}(A)$.

$\llbracket$

$$\begin{aligned} A \text{ und } B \text{ unabhängig} \quad &\Leftrightarrow \quad \mathbf{P}(A \cap B) = \mathbf{P}(A) \cdot \mathbf{P}(B) \\ &\Leftrightarrow \quad \mathbf{P}(A) = \frac{\mathbf{P}(A \cap B)}{\mathbf{P}(B)} = \mathbf{P}(A \mid B) \end{aligned}$$

$\rrbracket$

(c) Umkehrformel: Falls $\mathbf{P}(A) > 0$ so gilt

$$\mathbf{P}(B \mid A) = \frac{\mathbf{P}(A \mid B) \cdot \mathbf{P}(B)}{\mathbf{P}(A)}$$

$\llbracket$

$$\mathbf{P}(B \mid A) = \frac{\mathbf{P}(B \cap A)}{\mathbf{P}(A)} = \frac{\mathbf{P}(B \cap A)}{\mathbf{P}(B)} \cdot \frac{\mathbf{P}(B)}{\mathbf{P}(A)} = \mathbf{P}(A \mid B) \cdot \frac{\mathbf{P}(B)}{\mathbf{P}(A)}$$

$\rrbracket$

(d) Sind $A_1, \ldots, A_n \in \mathfrak{A}$ Ereignisse, so gilt:

$$\mathbf{P}\left(\bigcap_{k=1}^{n} A_k\right) = \mathbf{P}(A_1) \cdot \mathbf{P}(A_2 \mid A_1) \cdot \mathbf{P}(A_3 \mid A_1 \cap A_2) \cdot \ldots \cdot \mathbf{P}(A_n \mid A_1 \cap \ldots \cap A_{n-1}).$$

Dies bleibt gültig, auch in dem Fall, daß eine der auftretenden Wahrscheinlichkeiten nicht definiert ist, wenn man $0 \cdot \text{undefiniert} = 0$ setzt.

$\llbracket$ 1. Fall: $\mathbf{P}(A_1 \cap \ldots \cap A_k) > 0$ für alle $k = 1, \ldots, n-1$

$$\begin{aligned} \Rightarrow \quad & \mathbf{P}(A_1) \cdot \mathbf{P}(A_2 \mid A_1) \cdot \ldots \cdot \mathbf{P}(A_n \mid A_1 \cap \ldots \cap A_{n-1}) \\ = \quad & \mathbf{P}(A_1) \cdot \frac{\mathbf{P}(A_1 \cap A_2)}{\mathbf{P}(A_1)} \cdot \frac{\mathbf{P}(A_1 \cap A_2 \cap A_3)}{\mathbf{P}(A_1 \cap A_2)} \cdot \ldots \cdot \frac{\mathbf{P}(A_1 \cap \ldots \cap A_n)}{\mathbf{P}(A_1 \cap \ldots \cap A_{n-1})} \\ = \quad & \mathbf{P}(A_1 \cap \ldots \cap A_n) \end{aligned}$$

2. Fall: $\mathbf{P}(A_1 \cap \ldots \cap A_k) = 0$ für ein $k = 1, \ldots, n$

$$\Rightarrow \quad \mathbf{P}(A_1 \cap \ldots \cap A_n) \leq \mathbf{P}(A_1 \cap \ldots \cap A_k) = 0 = \text{linke Seite}$$

und die rechte Seite ist „$= 0$“ nach Konvention. $\rrbracket$

7.4 Satz (Formel von der totalen Wahrscheinlichkeit/ Formel von *Bayes*)

Es seien $B_1, B_2, \ldots$ disjunkte Ereignisse und $A \in \mathfrak{A}$ mit $A \subset \bigcup_{n=1}^{\infty} B_n$. Dann gilt

$$\mathbf{P}(A) = \sum_{n=1}^{\infty} \mathbf{P}(B_n) \cdot \mathbf{P}(A \mid B_n)$$

und (im Fall $\mathbf{P}(A) > 0$)

$$\mathbf{P}(B_k \mid A) = \frac{\mathbf{P}(B_k) \cdot \mathbf{P}(A \mid B_k)}{\sum\limits_{n=1}^{\infty} \mathbf{P}(B_n) \cdot \mathbf{P}(A \mid B_n)}.$$

Beweis:
Es ist

$$\mathbf{P}(A) = \mathbf{P}\left(A \cap \bigcup_{n=1}^{\infty} B_n\right) = \mathbf{P}\left(\bigcup_{n=1}^{\infty} (A \cap B_n)\right) = \sum_{n=1}^{\infty} \mathbf{P}(A \cap B_n) = \sum_{n=1}^{\infty} \mathbf{P}(B_n) \cdot \mathbf{P}(A \mid B_n),$$

diese Formel ist auch im Fall $\mathbf{P}(B_n) = 0$ gültig, wenn man $0 \cdot \text{undefiniert} = 0$ setzt.

Ist $\mathbf{P}(A) > 0$ so gilt nach der Umkehrformel

$$\begin{aligned}\mathbf{P}(B_k \mid A) &= \frac{\mathbf{P}(B_k) \cdot \mathbf{P}(A \mid B_k)}{\mathbf{P}(A)} \\ &= \frac{\mathbf{P}(B_k) \cdot \mathbf{P}(A \mid B_k)}{\sum\limits_{n=1}^{\infty} \mathbf{P}(B_n) \cdot \mathbf{P}(A \mid B_n)}.\end{aligned}$$

□

7.5 Beispiele

7.5.1 Signalübertragung

Über einen Datenkanal wird eine Folge von Bits übertragen. Durch Störung wird eine gesendete 0 mit Wahrscheinlichkeit ε_0 als 1 empfangen und eine gesendete 1 mit Wahrscheinlichkeit ε_1 als 0 empfangen. Das Verhältnis der Anzahl der gesendeten Einsen zu der der Nullen beträgt ρ ($\rho \in [0,1]$). Wie groß ist die Wahrscheinlichkeit, daß das übertragene Bit korrekt ist, wenn wir 0 bezeiehungsweise 1 empfangen haben?

$$A_i \,\hat{=}\, „i \text{ wurde gesendet“} \qquad B_i \,\hat{=}\, „i \text{ wurde empfangen“}$$

Dann wissen wir

$$\mathbf{P}(B_1 \mid A_0) = \varepsilon_0 \qquad \mathbf{P}(B_0 \mid A_1) = \varepsilon_1 \qquad \frac{\mathbf{P}(A_1)}{\mathbf{P}(A_0)} = \rho.$$

Die Formel von Bayes liefert

$$\begin{aligned}\mathbf{P}(A_0 \mid B_0) &= \frac{\mathbf{P}(A_0) \cdot \mathbf{P}(B_0 \mid A_0)}{\mathbf{P}(A_0) \cdot \mathbf{P}(B_0 \mid A_0) + \mathbf{P}(A_1) \cdot \mathbf{P}(B_0 \mid A_1)} \\ &= \left(1 + \frac{\mathbf{P}(A_1)}{\mathbf{P}(A_0)} \cdot \frac{\mathbf{P}(B_0 \mid A_1)}{\mathbf{P}(B_0 \mid A_0)}\right)^{-1} \\ &\overset{(*)}{=} \left(1 + \frac{\rho \cdot \varepsilon_1}{1 - \varepsilon_0}\right)^{-1}\end{aligned}$$

Bei $(*)$ geht

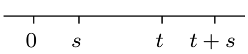
$$\mathbf{P}(B_0 \mid A_0) = 1 - \mathbf{P}\left(B_0^{\complement} \mid A_0\right) = 1 - \mathbf{P}(B_1 \mid A_0) = 1 - \varepsilon_0$$

ein.
Analog:

$$\mathbf{P}(A_1 \mid B_1) = \left(1 + \frac{\varepsilon_0}{\rho \cdot (1 - \varepsilon_1)}\right)^{-1}$$

7.5.2 Lebensdauer eines Bauteils

Ein besonders einfaches Modell für die (zufällige) Zeit T bis zum Ausfallen eines Bauteils besteht in der Annahme, daß das Teil, wenn es eine bestimmte Zeit überstanden hat, danach mit der gleichen Wahrscheinlichkeit eine weitere Zeitperiode s durchhält, wie direkt nach dem Einbau.



Es sei $T : \Omega \to \mathbb{N}$ die Zufallsvariable, die die Lebenszeit beschreibt. Dann muß

$$(*) \quad \mathbf{P}\{T \geq t + s \mid T \geq t\} = \mathbf{P}\{T \geq s\} \quad \forall s, t \in \mathbb{N}$$

gelten. Man nennt dann die Verteilung von T gedächtnislos.

Behauptung: Es gilt $\mathbf{P}_T = \mathcal{G}_p$ für ein $p \in \,]0,1[$.
Beweis: Sei $F(t) := \mathbf{P}\{T \geq t\}$. Dann folgt aus $(*)$

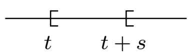
$$\begin{aligned} F(s) &= \mathbf{P}\{T \geq s+t \mid T \geq t\} \\ &= \frac{\mathbf{P}\{T \geq s+t, T \geq t\}}{\mathbf{P}\{T \geq t\}} \\ &= \frac{\mathbf{P}\{T \geq s+t\}}{\mathbf{P}\{T \geq t\}} = \frac{F(s+t)}{F(t)} \end{aligned}$$

für alle $s, t \in \mathbb{N}$ mit $F(t) > 0$.

$$\Rightarrow \quad F(s+t) = F(s) \cdot F(t)$$

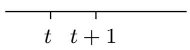
für alle $s, t \in \mathbb{N}$.
Sei $p = 1 - F(1)$

$$\begin{aligned} \Rightarrow \quad F(t) &= F(1 + (t-1)) \\ &= F(1) \cdot F(t-1) \\ &\;\vdots \\ &= F(1)^t = (1-p)^t \quad \forall t \in \mathbb{N}. \end{aligned}$$



Wegen $\lim\limits_{t\to\infty} F(t) = \lim\limits_{t\to\infty} \mathbf{P}\{T \geq t\} = 0 \Rightarrow p \in \,]0,1]$

$$\begin{aligned} \Rightarrow \quad \mathbf{P}_T\{t\} &= \mathbf{P}\{T = t\} = \mathbf{P}\{T \geq t\} - \mathbf{P}\{T \geq t+1\} \\ &= F(t) - F(t+1) = (1-p)^t - (1-p)^{t+1} \\ &= (1-p)^t \cdot (1 - (1-p)) = p \cdot (1-p)^t = \mathcal{G}_p\{t\} \end{aligned}$$



Sei umgekehrt $\mathbf{P}_T = \mathcal{G}_p$

$$\begin{aligned} \Rightarrow \quad \mathbf{P}\{T \geq t\} &= \sum_{k=t}^{\infty} p \cdot (1-p)^k = p \cdot (1-p)^t \cdot \sum_{k=0}^{\infty} (1-p)^k \\ &= (1-p)^t \cdot p \cdot \frac{1}{1 - (1-p)} = (1-p)^t \end{aligned}$$

$$\begin{aligned} \Rightarrow \quad \mathbf{P}\{T \geq t+s \mid T \geq t\} &= \frac{\mathbf{P}\{T \geq t+s\}}{\mathbf{P}\{T \geq t\}} \\ &= \frac{(1-p)^{t+s}}{(1-p)^t} = (1-p)^s = \mathbf{P}\{T \geq s\} \end{aligned}$$

□

Aufgaben:

Aufgabe 7.1:

Bei einer Reihenuntersuchung auf HIV wird ein Test verwendet, der mit einer Wahrscheinlichkeit von 98% positiv ausfällt, wenn die untersuchte Person infiziert ist, während er für Nichtinfizierte mit einer Wahrscheinlichkeit von 99% negativ ausfällt. Berechnen Sie mit Hilfe der Formel von Bayes, in Abhängigkeit vom Anteil p der infizierten Personen der Gesamtbevölkerung, die Wahrscheinlichkeit dafür, daß eine untersuchte Person

(a) trotz negativem Testergebnis erkrankt ist beziehungsweise

(b) trotz positivem Ergebnis gesund ist.

Skizzieren Sie diese Wahrscheinlichkeiten als Funktion von p in den Intervallen $[0, 0.1]$ beziehungsweise $[0.9, 1]$.

Aufgabe 7.2:

Eine Maschine A stellt Werkstücke mit einer Ausschußquote von 20% her, eine Maschine B arbeitet doppelt so schnell wie die Maschine A bei einer halbierten Ausschußquote von 10%. Die beiden Maschinen packen die Werkstücke zu Zehnerpackungen. Man wählt zufällig eine Packung aus, diese enthält drei Ausschußstücke. Mit welcher Wahrscheinlichkeit wurde ihr Inhalt von der Maschine B produziert?

Aufgabe 7.3:

In einem zweistufigen Experiment wird zuerst die zufällige Anzahl N der durchzuführenden Versuche bestimmt und danach N mal unabhängig ein Versuch wiederholt, der jeweils mit Wahrscheinlichkeit $p \in [0, 1]$ gelingt. Bestimmen Sie die Verteilung der Gesamtzahl der Erfolge X, falls N Poisson-verteilt mit Parameter $\lambda > 0$ ist und berechnen Sie $\mathbf{E}(X)$.

Aufgabe 7.4:

In einem Kraftwerk sind zum Betrieb der Kühlanlage zwei voneinander unabhängige Pumpen installiert. Zum Betrieb jeder Pumpe dient ein eigener Generator. Falls einer davon ausfällt, kann er noch durch das Notstromaggregat ersetzt werden. Das gesamte Kühlsystem ist so lange funktionsfähig, wie mindestens eine Pumpe funktioniert. An jeder Komponente des Systems könnte während der Zeit t unabhängig voneinander Störungen auftreten, und zwar an den Pumpen mit Wahrscheinlichkeit p_1, an den Generatoren mit Wahrscheinlichkeit p_2 und am Notstromaggregat mit Wahrscheinlichkeit p_3.

(a) Bestimmen Sie die Wahrscheinlichkeit, daß die Kühlanlage während der Zeit t funktioniert.

(b) Bestimmen Sie die Ausfallwahrscheinlichkeit des obigen Systems mit $p_1 = p_2 = 0,001$ und $p_3 = 0,005$. Vergleichen Sie damit die Ausfallwahrscheinlichkeit eines einfachen Kühlsystems, das aus einer Pumpe und einem Generator besteht.

(c) Nun hat das Kühlsystem aber doch versagt. Bestimmen Sie die Wahrscheinlichkeit, daß nur die beiden Pumpen defekt sind.

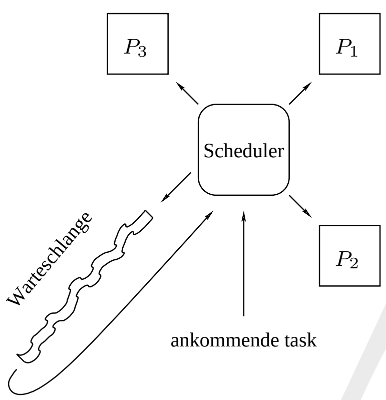
8. Kapitel: Markoff-Ketten

In diesem Kapitel untersuchen wir zufällige zeitliche Prozesse, deren zukünftige Entwicklung zwar von der Gegenwart, nicht aber von der gesamten Vorgeschichte abhängt. Dabei ist die Zeit durch die diskrete Menge $\mathbb{N}$ beschrieben und die möglichen Zustände des Prozesses liegen in einem abzählbaren Zustandsraum I.

8.1 Beispiel

Es sei $(D_1, D_2, \ldots)$ eine Folge identisch verteilter unabhängiger Zufallsvariabler mit Werten in $\mathbb{Z}$ und $X_m = \sum_{j=1}^{m} D_j$, $X_0 := 0$. Wir bezeichen einen solchen stochastischen Prozeß $(X_0, X_1, X_2, \ldots)$ als Irrfahrt auf $I = \mathbb{Z}$. Im Spezialfall $\mathbf{P}\{D_1 = +1\} = \mathbf{P}\{D_1 = -1\} = \frac{1}{2}$ sprechen wir von der einfachen symmetrischen Irrfahrt. Es ist wegen $X_{m+1} = X_m + D_{m+1}$ klar, daß der Zustand zum Zeitpunkt $m+1$ nur vom Zustand zum Zeitpunkt m abhängt, aber nicht von der weiteren Vergangenheit, denn D_{m+1} ist nach Vorraussetzung von X_m unabhängig.

8.2 Beispiel (Rechner mit $c \geq 1$ Prozessoren)



Der Scheduler teilt so lange noch Prozessoren frei sind jede ankommende task einem freien Prozessor zu, der diese dann vollständig abarbeitet. Sind keine Prozessoren frei, wird die ankommende task in eine Warteschlange mit der Kapazität $m \geq 0$ gesetzt; ist auch diese voll, wird die task abgewiesen. Zur Beschreibung des Zustands des Systems verwenden wir die Zahl $i \in I := \{0, 1, \ldots, c+m\}$ der tasks die gerade bearbeitet werden oder warten.
D.h. im Fall $i < c$ werden neuankommende tasks sofort bearbeitet; im Fall $c \leq i < c+m$ in die Warteschlange gesetzt und im Fall $i = c+m$ abgewiesen. Auch hier ist plausibel, daß der Zustand X_{m+1} des Systems zum Zeitpunkt $m+1$ nur von X_m abhängt, aber nicht von der weiteren Vergangenheit.

8.3 Definition

Es sei $(X_0, X_1, \ldots)$ eine Folge von Zufallsvariablen mit Werten in einer endlichen oder abzählbar unendlichen Menge I. Dann heißt $(X_0, X_1, \ldots)$ eine (zeithomogene) Markoff-Kette mit Übergangsmatrix (oder Übergangskern) $K = (K_{i,j})_{i,j \in I} \in M^{I \times I}(\mathbb{R})$ für alle $m \in \mathbb{N}$:

$$\mathbf{P}\{X_{m+1} = j \mid X_0 = i_0, \ldots, X_m = i_m\} = K_{i_m, j}$$

für alle $j \in I$ und alle $i_0, \ldots, i_m \in I$, bei denen die bedingte Wahrscheinlichkeit definiert ist.
Bemerkung:
Die Markoff-Eigenschaft besteht also darin, daß der Zustand zum Zeitpunkt $m+1$ von der Vergangenheit nur über den unmittelbar vorhergehenden Zustand X_m abhängt. Die zeitliche Homogenität drückt sich dadurch aus, daß der Übergangskern K nicht von m abhängt.

8.4 Beispiel

Bei der Irrfahrt aus (7.1) gilt für alle $i_0, \ldots, i_m, j \in \mathbb{Z}$ (falls wir $X_0 := 0$ setzen):

$$\begin{aligned} \mathbf{P}\{X_{m+1} = j, X_m = i_m, \ldots, X_0 = i_0\} &= \mathbf{P}\{X_0 = i_0, D_1 = i_1 - i_0, \ldots, D_{m+1} = j - i_m\} \\ &= \delta_{i_0,0} \cdot \prod_{k=1}^{m} \mathbf{P}\{D_k = i_k - i_{k-1}\} \cdot \mathbf{P}\{D_{m+1} = j - i_m\}. \end{aligned}$$

Analog:

$$\mathbf{P}\{X_m = i_m, \ldots, X_0 = i_0\} = \delta_{i_0,0} \cdot \prod_{k=1}^{m} \mathbf{P}\{D_k = i_k - i_{k-1}\}.$$

Daraus folgt, falls $\mathbf{P}\{X_m = i_m, \ldots, X_0 = i_0\} > 0$

$$\begin{aligned}\mathbf{P}\{X_{m+1} = j \mid X_m = i_m, \ldots, X_0 = i_0\} &= \mathbf{P}\{D_{m+1} = j - i_m\} \\ &= \mathbf{P}\{D_1 = j - i_m\} =: K_{i_m,j}.\end{aligned}$$

Also ist die Irrfahrt eine Markoff-Kette.

8.5 Satz

Es sei $(X_0, X_1, \ldots)$ eine Markoff-Kette mit Übergangsmatrix K.

(a) Es gilt $\mathbf{P}\{X_m = j \mid X_m = i\} = K_{i,j}$ für alle $m \in \mathbb{N}$ und $i, j \in I$ mit $\mathbf{P}\{X_m = i\} > 0$.

(b) Für alle $i \in I$, so daß $\mathbf{P}\{X_m = i\} > 0$ für mindestens ein $m \in \mathbb{N}$, gilt $K_{i,j} \geq 0 \, \forall j \in I$ und $\sum_{j \in I} K_{i,j} = 1$.

(c) Es sei $\sigma = (\sigma_i : i \in I)$ der durch $\sigma_i := \mathbf{P}\{X_0 = i\}$ definierte Zeilenvektor, der die <u>Startverteilung</u> beschreibt. Dann gilt für alle $m \in \mathbb{N}$:

$$\mathbf{P}\{X_m = i\} = (\sigma K^m)_i$$

(wobei $K^m = K \cdot \ldots \cdot K$ die m-te Potenz von K ist).

Beweis:

(a) Mit der Formel für die totale Wahrscheinlichkeit (Satz (7.4)) und der Definition (8.3) gilt:

$$\begin{aligned}\mathbf{P}\{X_{m+1} = j, X_m = i\} &= \mathbf{P}\left(\dot{\bigcup_{i_0,\ldots,i_{m-1} \in I}} \{X_{m+1} = j, X_m = i, \ldots, X_0 = i_0\} \right) \\ &= \sum_{i_0,\ldots,i_{m-1} \in I} \mathbf{P}\{X_{m+1} = j, X_m = i, \ldots, X_0 = i_0\} \\ &= \sum_{i_0,\ldots,i_{m-1} \in I} K_{i,j} \cdot \mathbf{P}\{X_m = i, \ldots, X_0 = i_0\} \\ &= K_{i,j} \cdot \mathbf{P}\left(\dot{\bigcup_{i_0,\ldots,i_{m-1} \in I}} \{X_m = i, \ldots, X_0 = i_0\} \right) \\ &= K_{i,j} \cdot \mathbf{P}\{X_m = i\}\end{aligned}$$

$$\Rightarrow \quad K_{i,j} = \frac{\mathbf{P}\{X_{m+1} = j, X_m = i\}}{\mathbf{P}\{X_m = i\}} = \mathbf{P}\{X_m = j \mid X_m = i\}.$$

(b) Nach (7.3.a) ist $M \mapsto \mathbf{P}\{X_{m+1} \in M \mid X_m = i\}$ ein Wahrscheinlichkeitsmaß auf I, falls $\mathbf{P}\{X_m = i\} > 0$.

$$\underset{\text{(a)}}{\Rightarrow} \quad K_{i,j} = \mathbf{P}\{X_{m+1} = j \mid X_m = i\} \geq 0 \quad \text{und}$$

$$\begin{aligned}\sum_{j \in I} K_{i,j} &= \sum_{j \in I} \mathbf{P}\{X_{m+1} = j \mid X_m = i\} \\ &= \mathbf{P}\{X_{m+1} \in I \mid X_m = i\} = 1.\end{aligned}$$

(c) Induktion nach m:

$\boxed{m = 0}$

$$K^0 = I \quad \Rightarrow \quad \mathbf{P}\{X_0 = i\} = \left(\sigma K^0\right)_i = (\sigma I)_i = \sigma_i.$$

$\boxed{m \rightsquigarrow m+1}$ Mit der Formel für die totale Wahrscheinlichkeit (7.4) und (a) gilt:

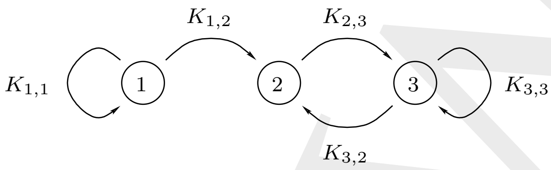
$$\begin{aligned} \mathbf{P}\{X_{m+1} = j\} &= \sum_{i \in I} \mathbf{P}\{X_m = i\} \cdot \mathbf{P}\{X_{m+1} = j \mid X_m = i\} \\ &= \sum_{i \in I} (\sigma K^m)_i \cdot K_{i,j} \\ &= \left(\sigma K^{m+1}\right)_j . \end{aligned}$$

□

8.6 Bemerkungen

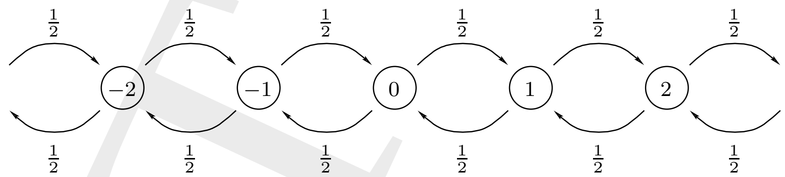
(a) Die Übergangsmatrix $K = (K_{i,j})$ läßt sich durch einen <u>bewerteten Graphen</u> beschreiben, wobei die Knoten gerade die Zustände $i \in I$ sind und die gerichtete Kante von i nach j die Bewertung $K_{i,j}$ bekommt. Kanten mit $K_{i,j} = 0$ werden dabei weggelassen.

<u>Beispiel 1</u>: $I = \{1, 2, 3\}$



$$\Leftrightarrow \quad K = \begin{pmatrix} K_{1,1} & K_{1,2} & 0 \\ 0 & 0 & K_{2,3} \\ 0 & K_{3,2} & K_{3,3} \end{pmatrix}$$

<u>Beispiel 2</u>: symmetrische Irrfahrt



$$\Leftrightarrow \begin{pmatrix} 0 & \frac{1}{2} & 0 & \cdots & \cdots & 0 \\ \frac{1}{2} & 0 & \frac{1}{2} & & & \vdots \\ 0 & \frac{1}{2} & \ddots & \ddots & & \vdots \\ \vdots & & \ddots & \ddots & \ddots & 0 \\ \vdots & & & \ddots & \ddots & \frac{1}{2} \\ 0 & \cdots & \cdots & 0 & \frac{1}{2} & 0 \end{pmatrix} \quad \text{mit } K_{i,j} = \begin{cases} \frac{1}{2} & \text{falls } |i-j| = 1 \\ 0 & \text{sonst.} \end{cases}$$

(b) (8.5.c) erlaubt es uns die Verteilung von X_m durch den Vektor $\sigma \cdot K^m$ ihrer Zähldichte auszudrücken. Insbesondere beschreibt σ die Startverteilung des Prozesses, d.h. die Verteilung von X_0.

(c) Die Eigenschaft der Übergangsmatrix (8.5.b) drückt man dadurch aus, daß man K <u>zeilenstochastisch</u> nennt. Dabei kann man auch ohne die Gültigkeit der Voraussetzung $\mathbf{P}\{X_m = i\} > 0$ für ein $m \in \mathbb{N}$ o.B.d.A. K als zeilenstochastisch annehmen, da im Fall $\mathbf{P}\{X_m = i\} = 0\, \forall m \in \mathbb{N}$ der Zustand $i \in I$ nie getroffen wird und wir $I \setminus \{i\}$ als Zustandsraum wählen können.

(d) Sind $K, L \in M^{I \times I}(\mathbb{R})$ zeilenstochastisch, so ist auch $K \cdot L$ zeilenstochastisch.
$[\![$ Für $i, k \in I$ gilt

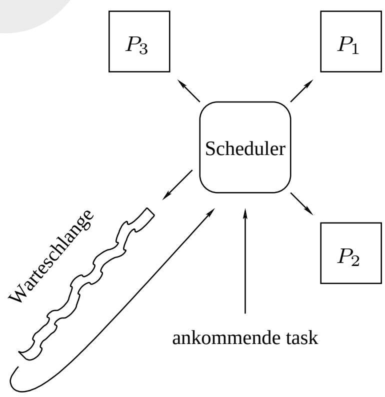
$$(K \cdot L)_{i,k} = \sum_{j \in I} K_{i,j} \cdot L_{j,k} \geq 0$$

$$\begin{aligned}
\Rightarrow \quad \sum_{k \in I} (K \cdot L)_{i,k} &= \sum_{k \in I} \sum_{j \in I} K_{i,j} \cdot L_{j,k} \\
&= \sum_{j \in I} K_{i,j} \cdot \sum_{k \in I} L_{j,k} = \sum_{j \in I} K_{i,j} \cdot 1 = 1.
\end{aligned}$$

$]\!]$

8.7 Beispiel (vgl. (8.2))

Rechner mit $c \geq 1$ Prozessoren und Warteschlange der Länge $m \geq 0$.
<u>Modellbildung</u>:



- Wir nehmen an, daß pro Zeiteinheit die Wahrscheinlichkeit für die Beendigung einer bestimmten task $\mu > 0$ ist.
- Mit Wahrscheinlichkeit $\lambda \geq 0$ trifft eine neue task ein. Diese Wahrscheinlichkeiten seien unabhängig von der Vorgeschichte.
- Weiter nehmen wir an, daß die Zeiteinheit so klein ist, daß das gleichzeitige Eintreffen beziehungsweise Fertigwerden vernachlässigt werden kann; insbesondere soll $\lambda + c \cdot \mu < 1$ gelten.
- Es bezeichne X_m den Zustand des Systems $I = \{0, 1, \ldots, c + m\}$.

Die Übergangsmatrix $(K_{i,j}) = K$ können wir nun mit (8.5.a) berechenen:
Für $0 \leq i \leq c$ gilt:

$$K_{i,j} = \begin{cases} i \cdot \mu & j = i - 1 \\ \lambda & j = i + 1 \\ 1 - \lambda - i \cdot \mu & j = i \\ 0 & \text{sonst.} \end{cases}$$

Für $c < i < c + m$ gilt:

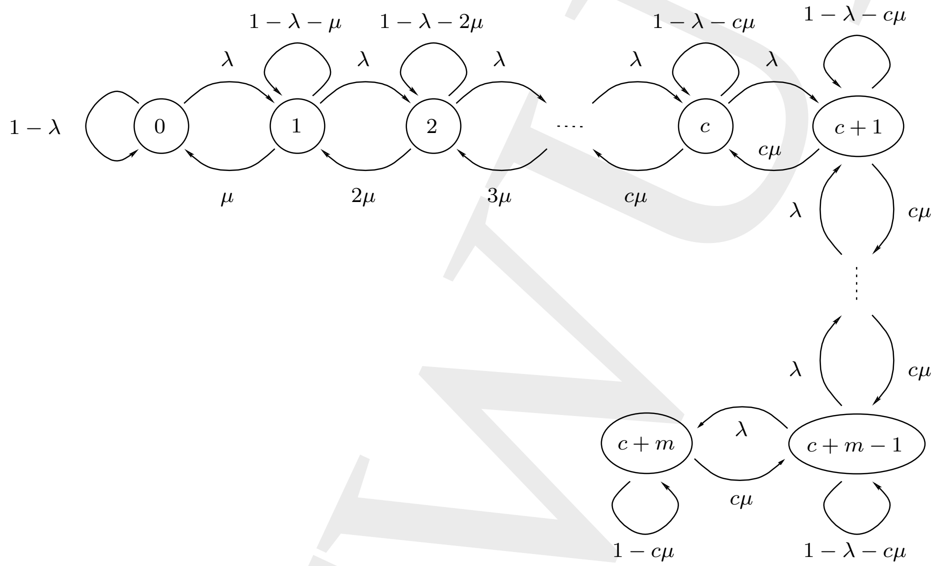
$$K_{i,j} = \begin{cases} c \cdot \mu & j = i - 1 \\ \lambda & j = i + 1 \\ 1 - \lambda - c \cdot \mu & j = i \\ 0 & \text{sonst.} \end{cases}$$

Und für $i = c + m$ gilt:

$$K_{c+m,j} = \begin{cases} c \cdot \mu & j = c + m - 1 \\ 1 - c \cdot \mu & j = c + m \\ 0 & \text{sonst.} \end{cases}$$

$\Rightarrow \quad K := \left(\mathbf{P}\left\{X_{m+1} = j\right\} \mid X_m = i\right)_{i,j}$

$$= \begin{pmatrix} 1-\lambda & \lambda & 0 & \cdots & \cdots & 0 \\ \mu & 1-\lambda-\mu & \lambda & 0 & & \vdots \\ 0 & 2\mu & 1-\lambda-2\mu & \lambda & 0 & \vdots \\ \vdots & 0 & \ddots & \ddots & \ddots & 0 \\ \vdots & & 0 & c\mu & 1-\lambda-c\lambda & \lambda \\ 0 & \cdots & \cdots & 0 & c\mu & 1-c\mu \end{pmatrix}$$



Frage: Asymptotisches Verhalten von $\mathbf{P}_{X_n}$ also von $\sigma \cdot K^n$

$$\lim_{n\to\infty} \sigma \cdot K^n = ???$$

8.8 Satz (*Markoff* - 1907)

Es sei $(X_n)_{n\in\mathbb{N}}$ eine Markoff-Kette mit endlichem Zustandsraum I und Übergangsmatrix K. Es existiere ein $N \geq 1$, $j_0 \in I$, so daß $K^N_{i,j_0} > 0$ für alle $i \in I$. Dann gibt es eine Zähldichte $\pi \in \mathbb{R}^I_+$ mit $\mathbf{P}\left\{X_n = i\right\} \to \pi_i$ für alle $i \in I$, und π ist unabhängig von der Startverteilung.

Beweis:

Da I endlich

$$\Rightarrow \quad \varepsilon := \min_{i\in I} K^N_{i,j_0} > 0 \text{ ist wohldefiniert.}$$

Wir betrachten den vollständigen metrischen Raum

$$X := \left\{ \pi \in \mathbb{R}^I_+ \,\middle|\, \sum_{i\in I} \pi_i = 1 \right\} \quad \text{mit } \mathrm{d}(\pi,\rho) := \sum_{i\in I} \left|\pi_i - \rho_i\right|.$$

Dann ist $\varphi : X \to X$, $\varphi(\sigma) = \sigma \cdot K$ eine Kontraktion, d.h.

$$\begin{aligned}
\mathrm{d}(\varphi(\sigma), \varphi(\pi)) &= \mathbf{d}(\sigma K, \pi K) = \sum_{j \in J} \left| \sum_{i \in I} \sigma_j K_{ij} - \pi_j K_{ij} \right| \\
&\leq \sum_{j \in J} \sum_{i \in I} |\sigma_j K_{ij} - \pi_j K_{ij}| = \sum_{j \in J} |\sigma_j - \pi_j| \cdot \underbrace{\sum_{i \in I} K_{ij}}_{=1} \\
&= \sum_{j \in J} |\sigma_j - \pi_j| = \mathrm{d}(\sigma, \pi)
\end{aligned}$$

Wegen $\sum\limits_{i \in I} (\pi_i - \rho_i) = 1 - 1 = 0$ gilt für $\varphi^N(\sigma) = \sigma \cdot K^N$, daß

$$\begin{aligned}
\mathrm{d}\left(\varphi^N(\pi), \varphi^N(\rho)\right) &= \sum_{j \in I} \left| \left(\pi \cdot K^N\right)_j - \left(\rho \cdot K^N\right)_j \right| \\
&= \sum_{j \in I} \left| \left((\pi - \rho) \cdot K^N\right)_j \right| \\
&= \sum_{j \in I} \left| \sum_{i \in I} (\pi_i - \rho_i) K^N_{i,j} \right| \\
&= \sum_{j \in I \setminus \{j_0\}} \left| \sum_{i \in I} (\pi_i - \rho_i) K^N_{i,j} \right| + \left| \sum_{i \in I} (\pi_i - \rho_i) \cdot K^N_{i,j_o} \right| \\
&\leq \sum_{j \in I \setminus \{j_0\}} \sum_{i \in I} |\pi_i - \rho_i| \cdot K^N_{i,j} + \left| \sum_{i \in I} (\pi_i - \rho_i) \cdot K^N_{i,j_o} \right| \\
&= \sum_{i \in I} \sum_{j \in I \setminus \{j_0\}} |\pi_i - \rho_i| \cdot K^N_{i,j} + \left| \sum_{i \in I} (\pi_i - \rho_i) \cdot \left(\left(K^N_{i,j_o} - \varepsilon\right) + \varepsilon\right) \right| \\
&\leq \sum_{i \in I} |\pi_i - \rho_i| \cdot \left(\underbrace{\sum_{j \in I} K^N_{i,j}}_{=1} - \varepsilon \right) + \left| \underbrace{\sum_{i \in I} (\pi_i - \rho_i)}_{=0} \cdot \varepsilon \right| \\
&= (1 - \varepsilon) \cdot \mathrm{d}(\pi, \rho).
\end{aligned}$$

Nach dem Banach'schen Fixpunktsatz konvergiert also $\varphi^{kN}(\sigma) = \sigma \cdot K^{kN}$ unabhängig vom Startwert σ, für $k \to \infty$, gegen den eindeutig bestimmten Fixpunkt π von φ^N. Wegen $\varphi^N(\varphi(\pi)) = \varphi\left(\varphi^N(\pi)\right) = \varphi(\pi)$ und $\varphi^N(\pi) = \pi$ und der Eindeutigkeit folgt $\varphi(\pi) = \pi$. Da φ Kontraktion

$$\Rightarrow \quad \mathrm{d}(\varphi(\sigma), \pi) = \mathrm{d}(\varphi(\sigma), \varphi(\pi)) \leq \mathrm{d}(\sigma, \pi)$$

und somit für $n \geq kN$: ($n = kN + r$ mit $0 \leq r < N$)

$$\mathrm{d}(\varphi^n(\sigma), \pi) \leq \mathrm{d}\left(\varphi^{kN}(\sigma), \pi\right) \xrightarrow{n \to \infty} 0$$

d.h. $\sigma \cdot K^n \to \pi$. □

8.9 Korollar

Unter den Voraussetzungen von Satz (8.8) gilt $K^n \to \overline{K}$, wobei

$$\overline{K} = \begin{pmatrix} \leftarrow \pi \rightarrow \\ \vdots \\ \leftarrow \pi \rightarrow \end{pmatrix}.$$

Beweis:
Für $l \in I$ sei $\left(\sigma^{(l)}\right)_{i \in I} = (\delta_{l,i})_{i \in I} = \big(0, \ldots, 0, \underset{l.\text{te Stelle}}{\underset{\uparrow}{1}}, 0, \ldots, 0\big)$. Dann ist $\sigma^{(l)}$ eine Startverteilung und nach (8.8) gilt $\sigma^{(l)} : K^n \xrightarrow{n\to\infty} \pi$, wobei π unabhängig von $l \in I$ ist. Für $i, l \in I$ gilt dann:

$$\left(\sigma^{(l)} \cdot K^n\right)_i = (K^n)_{l,i} \xrightarrow{n\to\infty} \pi_i$$

$$\Rightarrow \quad K^n \to \overline{K} \quad \text{wobei } \overline{K}_{l,i} = \pi_i \quad \text{d.h. } \overline{K} = \begin{pmatrix} \leftarrow \pi \rightarrow \\ \vdots \\ \leftarrow \pi \rightarrow \end{pmatrix}.$$

□

8.10 Satz

Es sei $K \in M^{I \times I}(\mathbb{R})$ eine zeilenstochastische Matrix und $\sigma \in \mathbb{R}_+^I$ eine Startverteilung, für die $\pi = \lim\limits_{n\to\infty} \sigma \cdot K^n$ existiert. Dann ist π eine Zähldichte (d.h. $\pi_i \geq 0$ und $\sum\limits_{i \in I} \pi_i = 1$) und es gilt $\pi \cdot K = \pi$.
Beweis:
Hier nur für den Fall, daß I endlich ist. Nach (8.5.c) ist $\sigma \cdot K^n$ eine Zähldichte für alle $n \geq 1$ $\Rightarrow$ π Zähldichte. Es ist

$$\pi \cdot K = \left(\lim_{n\to\infty} \sigma \cdot K^n\right) \cdot K = \lim_{n\to\infty} \sigma \cdot K^{n+1} = \pi.$$

□

8.11 Definition

Es sei $K \in M^{I \times I}(\mathbb{R})$ zeilenstochastisch. Dann heißt jede Zähldichte $\pi \in \mathbb{R}_+^I$ mit $\pi \cdot K = \pi$ ein <u>stationärer Vektor</u> (zu K/ zur Markoff-Kette $(X_0, X_1, \ldots)$ mit Übergangsmatrix K).

8.12 Bemerkung

Unter den Voraussetzungen von Satz (8.8) gilt: Die Grenzverteilung π ist der eindeutig bestimmte stationäre Vektor zu K. D.h. um $\lim\limits_{n\to\infty} \mathbf{P}\{X_n = i\} = \pi_i$ zu berechnen löst man das lineare Gleichungsystem

$$\begin{aligned} \pi \cdot K &= \pi \quad \Leftrightarrow \quad \pi \cdot (K - I) = 0 \quad \Leftrightarrow \quad (K - I)^t \cdot \pi = 0 \\ \sum_{i \in I} \pi_i &= 1 \quad \text{mit } \pi_i \geq 1. \end{aligned}$$

8.13 Korollar

Es sei π ein stationärer Vektor einer Markoff-Kette $(X_0, X_1, \ldots)$ mit Übergangsmatrix K und Startverteilung $\mathbf{P}\{X_0 = i\} = \pi_i \, \forall i \in I$. Dann gilt $\mathbf{P}\{X_n = i\} = \pi_i$ für alle $n \in \mathbb{N}$, $i \in I$.
Beweis:
π stationärer Vektor

$$\Rightarrow \quad \pi \cdot K = \pi \quad \Rightarrow \quad \pi \cdot K^2 = \pi \cdot K = \pi \quad \overset{\text{induktiv}}{\Rightarrow} \quad \pi \cdot K^n = \pi$$

$$\Rightarrow \quad \mathbf{P}\{X_n = i\} = (\pi \cdot K^n)_i = \pi_i \quad \forall i \in I, \, n \in \mathbb{N}.$$

□

8.14 Beispiel (vgl. (8.2) und (8.7))

Wir wenden Satz (8.8) auf die Übergangsmatrix K in (8.7) an. Da K eine 3-Band-Matrix ist, folgt $K^N_{i,j_0} > 0$ für alle $i \in I$ und jedes $j_0 \in I$ für $N = c + m - 1$. Nach Satz (8.8) existiert also eine eindeutig stationäre Verteilung π, die nach (8.10) das Gleichungssystem $\pi \cdot K = \pi$ für $\pi = (\pi_0, \dots, \pi_{c+m})$ erfüllt

$$K = \begin{pmatrix} \ddots & & 0 \\ & \ddots & \\ 0 & & \ddots \end{pmatrix} \qquad K^2 = \begin{pmatrix} \ddots & & 0 \\ & \ddots & \\ 0 & & \ddots \end{pmatrix}$$

$$\Rightarrow \quad \begin{cases} \mu\pi_1 &= \lambda\pi_0 \\ (k+1)\cdot\mu\pi_{k+1} &= (\lambda + k\mu)\cdot\pi_k - \lambda\pi_{k-1} \qquad \text{für } 1 \le k \le c-1 \\ c\mu\pi_{k+1} &= (\lambda + k\mu)\cdot\pi_k - \lambda\pi_{k-1} \qquad \text{für } c \le k \le c+m-1 \\ c\mu\pi_{c+m} &= \lambda\pi_{c+m-1}. \end{cases}$$

Sei $\rho = \dfrac{\lambda}{\mu}$. Dividiert man durch $\mu > 0$, so folgt

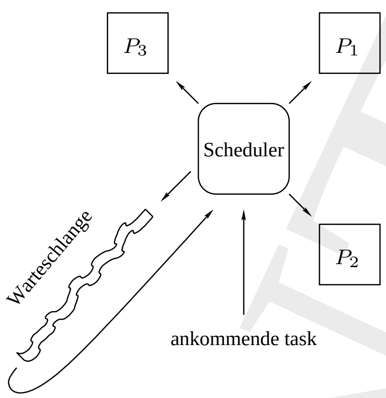
$$\pi_k = \frac{\rho^k}{k!}\cdot\pi_0 \quad \text{für } k \le c \quad \text{und} \quad \pi_k = \frac{\rho^k}{c!\cdot c^k}\cdot\pi_0 \quad \text{für } c+1 \le k \le c+m.$$

Da $\sum\limits_{i=0}^{c+m} \pi_i = 1$ folgt

$$\pi_0 = \left(\sum_{k=0}^{c-1} \frac{\rho^k}{k!} + \frac{\rho^c}{c!} \cdot \frac{1 - \left(\frac{\rho}{c}\right)^{m+1}}{1 - \frac{\rho}{c}} \right)^{-1}.$$

<u>Interpretation</u>:

Die Zahl $\rho = \dfrac{\lambda}{\mu}$ gibt an, wieviel mal mehr tasks ankommen, als ein Prozessor bearbeiten kann.



(i) Im Überlauffall $\rho \ge c$ ist $\pi_k \le \pi_{k+1}$ für alle $k = 0, \dots, c+m-1$. D.h. der Zustand $c+m$ hat die größte Wahrscheinlichkeit, das System lehnt, wegen Überlastung, also in der Regel die tasks ab. Vergrößern der Warteschlange nützt hier nichts.

(ii) Ist $\rho < c$, so ist der Zustand mit der höchsten Wahrscheinlichkeit π_k für $k = [\rho] < c$ erreicht, also im Bereich wo noch Prozessoren frei sind. Im Bereich der Warteschlangenplätze $\{c+1, \dots, c+m\}$ nehmen die Wahrscheinlichkeiten exponentiell ab, und zwar umso schneller, je kleiner $\dfrac{\rho}{c}$ ist. Man kann also die Warteschlange umso kleiner dimensionieren, je kleiner der Quotient ist. Die Wahl von m wird dann dadurch bestimmt, wie unwahrscheinlich das Abweisen einer task sein soll.

Aufgaben:

Aufgabe 8.1:

Am Dortmunder Weihnachtsmarkt reihen sich auf dem Westenhellweg insgesamt n Glühweinstände. Eine Person startet beim ersten Stand dieser Reihe und bleibt in jeder Zeiteinheit mit gleicher Wahrscheinlichkeit beim aktuellen Stand oder geht zu einem der (ein oder zwei) Nachbarstände. Mit welcher Wahrscheinlichkeit befindet sie sich nach langer Zeit am j-ten Stand?

(Die Möglichkeit der Einlieferung in eine Ausnüchterungszelle dürfen Sie außer acht lassen!)

Aufgabe 8.2:

Ein etwas zerstreuter Professor besitzt zwei Schirme. Jeden Morgen geht er zur Uni, und, wenn es regnet, nimmt er seien Schirm mit - falls er daheim gerade einen Schirm hat. Abends geht er wieder nach Hause und nimmt, wenn es regnet und er einen Schirm in der Uni hat, einen Schirm mit. Wenn es nicht regnet, trägt er keinen Schirm mit sich herum. Es werde angenommen, daß es bei jedem Gang unabhängig von der Vorgeschichte, mit Wahrscheinlichkeit $p \in]0, 1[$ regnet. Berechnen Sie die Wahrscheinlichkeit, daß der Professor bei einem Gang naß wird.

Hinweis: Verwenden Sie als Zustandsraum $I = \{0, 1, 2\}$ die Zahl der Schirme, die beim Weggehen am jeweiligen Ort vorhanden sind.

9. Kapitel: Schätzung statistischer Parameter

Bisher hatten wir stets ein stochastisches Modell vorgefunden und daraus Folgerungen über Wahrscheinlichkeiten zufälliger Ereignisse gezogen. In der Praxis hat man jedoch meist kein Modell vorgegeben, sondern muß versuchen, dieses aufgrund der experimentell gefundenen Daten zu (re-) konstruieren; dies ist Aufgabe der Statistik.

Zum Beispiel, ist man sich oft beim Werfen einer Münze nicht sicher, ob diese mit Wahrscheinlichkeit $\frac{1}{2}$ Wappen zeigt, oder verfälscht ist. Um dies zu untersuchen, wiederholen wir das Experiment n mal und erhalten k mal „Wappen“. Es ist klar, daß wir den Stichprobenraum als $\Omega = \{0, 1, \ldots, n\}$ mit $\mathfrak{A} = \mathcal{POT}(\Omega)$ beschreiben können, aber wir wissen nur, daß das „richtige Modell“ durch die Binomialverteilung $\mathcal{B}_{n,p}$ beschrieben wäre, kennen aber $p \in [0, 1]$ nicht. Aufgrund vom Ereignis k unseres Versuchs wollen wir eine Aussage über das „richtige“ p machen.

Ein weiteres Beispiel tritt bei einer Messung auf, die n mal wiederholt wird, und die Meßwerte $x_1, \ldots, x_n$ ergab. Nehmen wir an, daß sich jede zufällige Messung $X_j = \mu + F_j$ zusammensetzt aus dem richtigen Wert μ und einer zufälligen Störung F_j, die unabhängig und $\mathcal{N}_{0,\sigma^2}$-verteilt sei, so ist das Experiment beschrieben durch $\Omega = \mathbb{R}^n$, $\mathfrak{A} = \mathfrak{B}(\mathbb{R}^n)$ und einer Familie von möglichen Wahrscheinlichkeitsmaßen $\mathbf{P}_{\mu,\sigma^2}$, deren Dichte nach (3.20.b) und (3.24):

$$(x_1, \ldots, x_n) \mapsto \prod_{j=1}^{n} \frac{1}{\sqrt{2\pi\sigma^2}} \cdot \mathbf{e}^{-\frac{(x_j-\mu)^1}{2\sigma^2}} = \frac{1}{(2\pi)^{\frac{n}{2}} \cdot \sigma^n} \cdot \exp\left(\frac{-1}{2\sigma^2} \sum_{j=1}^{n} (x_j - \mu)^2\right)$$

ist, wobei aber μ und σ nicht bekannt sind.

9.1 Definition

Ein statistisches Experiment ist ein Tripel $(\Omega, \mathfrak{A}, \mathcal{P})$, wobei $\mathcal{P}$ eine Familie von Wahrscheinlichkeitsmaßen auf $(\Omega, \mathfrak{A})$ ist. Häufig liegt eine Parametrisierung $\mathcal{P} = (\mathbf{P}_\vartheta : \vartheta \in \Theta)$ vor mit Parameterraum Θ und Wahrscheinlichkeitsmaßen $\mathbf{P}_\vartheta$ auf $(\Omega, \mathfrak{A})$ für jedes $\vartheta \in \Theta$. Wir sagen dann, daß ein parametrisches Experiment vorliegt.

Aufgabe der Statistik ist es nun, aufgrund der (zufälligen) Stichprobe $\omega \in \Omega$ eine Aussage über den richtigen Wert von ϑ zu machen beziehungsweise das richtige $\mathbf{P} \in \mathcal{P}$.

Dafür gibt es 2 besonders wichtige Ansätze:

(1) Für ϑ soll eine Schätzung $S(\omega)$ aufgrund von ω gemacht werden. Da die Stichprobe ω und damit auch $S(\omega)$ zufällig sind, wird diese Schätzung meist vom richtigen Wert abweichen und wir müssen versuchen, S so zu konstruieren, daß der mittlere Fehler möglichst klein wird.

(2) Wir denken uns die möglichen Wahrscheinlichkeitsmaße in zwei Teilmengen zerlegt, $\mathcal{P} = \mathcal{P}_0 \uplus \mathcal{P}_1$ und wollen nur entscheiden, welche der beiden vorliegt (z.B. beim Würfeln $\mathcal{P}_0 = \left\{\frac{1}{2}\right\} \mathrel{\widehat{=}}$ Münze fair und $\mathcal{P}_1 = \left[0, \frac{1}{2}\right[\cup \left]\frac{1}{2}, 1\right] \mathrel{\widehat{=}}$ Münze verfälscht). Aufgrund von Stichprobe ω treffen wir nun die Entscheidung $t(\omega) \in \{0, 1\}$ mit $0 \mathrel{\widehat{=}} \mathcal{P}_0$ wird beibehalten oder $1 \mathrel{\widehat{=}} \mathcal{P}_0$ wird verworfen (d.h. wir folgern, daß das richtige $\mathbf{P}$ zu $\mathcal{P}_1$ gehört). $t : \Omega \to \{0, 1\}$ heißt Test und wir wollen durch geschickte Wahl von t erreichen, daß die Fehler 1. Art: $\mathbf{P}\{t = 1\}$ für $\mathbf{P} \in \mathcal{P}_0$ (Nullhypothese ist richtig, wird aber verworfen) und Fehler 2. Art: $\mathbf{P}\{t = 0\}$ für $\mathbf{P} \in \mathcal{P}_1$ (Nullhypothese $\mathcal{P}_0$ ist falsch, wird aber bestätigt) beide möglichst klein sind.

9.2 Beispiele

9.2.1 Münzwurf

Hier ist $\Omega = \{0, 1, \ldots, n\}$, $\mathfrak{A} = \mathcal{POT}(\Omega)$ und $\mathcal{P} = \{\mathcal{B}_{n,p} : p \in [0, 1]\}$ und es liegt ein parametrisches Experiment vor.

9.2.2 Messung

$\Omega = \mathbb{R}^n$, $\mathfrak{A} = \mathfrak{B}(\mathbb{R})^n$ und $\mathcal{P} = \left\{\mathbf{P}_{\mu,\sigma^2} : \mu \in \mathbb{R},\, \sigma^2 > 0\right\}$ ist wieder ein parametrisches Experiment, was aber als Modellannahme zu hinterfragen ist, weil ja eine $\mathcal{N}_{0,\sigma^2}$-Verteilung für die Fehler schon eine „starke" Annahme ist. Hier könnte man auch ein nicht parametrisches Modell verwenden, was in der modernen Statistik getan wird, aber oft komplizierte Verfahren nötig macht.

9.3 Definition

Es sei $(\Omega, \mathfrak{A}, \{\mathcal{P}_\vartheta : \vartheta \in \Theta\})$ ein statistisches Experiment und $g : \Theta \to \mathbb{R}$ eine Funktion. Ein „Schätzer" für den unbekannten Parameter $g(\vartheta)$ ist eine Funktion S von Ω nach $\mathbb{R}$. Wird eine Stichprobe ω erzielt, so wird der Wert $S(\omega)$ als Schätzung für die unbekannte Zahl $g(\vartheta)$ angegeben. Besitzt S^2 einen Erwartungswert bezüglich $\mathcal{P}_\vartheta$ für alle $\vartheta \in \Theta$, so heißt $\mathbf{E}_\vartheta\left((S - g(\vartheta))^2\right)$ der mittlere quadratische Fehler von S unter ϑ.
Wir wollen im folgenden versuchen diesen Fehler für alle $\vartheta \in \Theta$ möglichst klein zu machen.

9.4 Beispiele (Maximum-Likelihood-Schätzer)

Eine sehr weit anwendbare Methode, für ein gegebenes Schätzproblem einen Schätzer zu konstruieren, wird durch die Maximum-Likelihood-Schätzer gegeben:

9.4.1

Sei zuerst Ω abzählbar. Ist eine Stichprobe $\omega \in \Omega$ erzielt worden, so schätzt man den unbekannten Wert ϑ durch dasjenige $\hat{\vartheta}(\omega) \in \Theta$, für das die Wahrscheinlichkeit, daß gerade die aufgefundene Stichprobe vorliegt, maximal ist. Im Fall eines diskreten Experiments suchen wir also nach dem Wert $\hat{\vartheta}(\omega)$, der die Likelihood-Funktion $\vartheta \mapsto \mathbf{P}_\vartheta\{\omega\}$ maximiert. $\hat{\vartheta}$ heißt also Maximum-Likelihood-Schätzer für ϑ, falls $\mathbf{P}_{\hat{\vartheta}(\omega)}\{\omega\} = \max\limits_{\vartheta \in \Theta}\{\mathbf{P}_\vartheta\{\omega\}\}$ für alle $\omega \in \Omega$ gilt.

Als Maximum-Likelihood-Schätzung für $g(\vartheta)$ wählen wir dann $S(\omega) := g\left(\hat{\vartheta}(\omega)\right)$. Es ist klar, daß im allgemeinen das obige Maximum nicht angenommen werden muß und dann auch kein Maximum-Likelihood-Schätzer existiert. Ebenso ist selbst im Fall der Existenz $\hat{\vartheta}(\omega)$ nicht notwendigerweise eindeutig bestimmt.

9.4.2

Es soll die Wahrscheinlichkeit des Ereignisses „Wappen" aufgrund der n-fachen Wiederholung des Münzwurfs geschätzt werden. Das statistische Modell ist hier also

$$\left(\Omega := \{0,1\}^n, \mathfrak{A} = \mathcal{POT}(\Omega), \{\mathbf{P}_\vartheta : \vartheta \in \Theta := [0,1]\}\right),$$

wobei $\mathbf{P}_\vartheta\{(\omega_1, \ldots, \omega_n)\} = \vartheta^k \cdot (1-\vartheta)^{n-k}$ und $k := \sum\limits_{j=1}^{n} \omega_j$ die Zahl der Erfolge ist; geschätzt werden soll der Parameter $g(\vartheta) := \vartheta$. Um die Maximum-Likelihood-Schätzung $\hat{\vartheta}(\omega)$ für die Stichprobe $\omega = (\omega_1, \ldots, \omega_n)$ zu berechnen, setzen wir nun die Ableitung der Funktion

$$\vartheta \mapsto \mathbf{P}_\vartheta\{(\omega_1, \ldots, \omega_n)\} = \binom{n}{k} \cdot \vartheta^k \cdot (1-\vartheta)^{n-k}$$

Null, d.h.

$$\binom{n}{k} \cdot \left(k \cdot \vartheta^{k-1} \cdot (1-\vartheta)^{n-k} - (n-k) \cdot \vartheta^k \cdot (1-\vartheta)^{n-k-1}\right) = 0.$$

Dies ist genau dann der Fall, wenn $k \cdot (1-\vartheta) - (n-k) \cdot \vartheta = 0$, d.h. $\hat{\vartheta}(\omega) = \frac{k}{n}$, und wir erhalten als Maximum-Likelihood-Schätzung $\overline{X}_n(\Omega) := \frac{1}{n} \cdot \sum\limits_{j=1}^{n} \omega_j$ die relative Häufigkeit des Ereignisses.

9.4.3

In einer Lieferung von insgesamt t Bauteilen soll der unbekannte Anteil der defekten Exemplare geschätzt werden, indem man (um nicht die gesamte Lieferung untersuchen zu müssen) nur eine Stichprobe der Größe $n \leq t$ prüft. Wir berechnen zuerst die Wahrscheinlichkeit dafür, daß genau k defekte Teile in der Stichprobe gefunden werden, wenn insgesamt ϑ defekte und $t - \vartheta$ korrekte Exemplare geliefert wurden, als

$$\mathcal{H}_{n,t,\vartheta}\left\{k\right\} = \frac{\binom{\vartheta}{k} \cdot \binom{t-\vartheta}{n-k}}{\binom{t}{n}}.$$

Dieses Wahrscheinlichkeitsmaß $\mathcal{H}_{n,t,\vartheta}$ auf $\{0, 1, \ldots, n\}$ ist die in (2.3.2) eingeführte hypergeometrische Verteilung.
Es liegt somit das statistische Experiment $\left(\Omega := \{0, 1, \ldots, n\}, \mathcal{POT}(\Omega), \mathcal{H}_{n,t,\vartheta} : 0 \leq \vartheta \leq t\right)$ vor, und der Wert des Parameters $g(\vartheta) := \frac{\vartheta}{t}$ ist zu schätzen. Um für den vorgefundenen Wert k der Stichprobe das Maximum der Funktion $\vartheta \mapsto \mathcal{H}_{n,t,\vartheta}\{k\}$ zu bestimmen, untersuchen wir, für welche ϑ diese Funktion wächst, d.h.

$$\binom{\vartheta - 1}{k} \cdot \binom{t - \vartheta + 1}{n - k} \leq \binom{\vartheta}{k} \cdot \binom{t - \vartheta}{n - k}$$

gilt. Dies ist genau dann der Fall, wenn

$$\frac{(\vartheta - 1)! \cdot (t - \vartheta + 1)!}{(\vartheta - 1 - k)! \cdot (k + t - n - \vartheta + 1)!} \leq \frac{\vartheta! \cdot (t - \vartheta)!}{(\vartheta - k)! \cdot (k + t - n - \vartheta)!},$$

also wenn

$$(\vartheta - k) \cdot (t - \vartheta + 1) \leq \vartheta \cdot (k + t - n - \vartheta + 1).$$

Eine leichte Rechnung zeigt, daß diese Bedingung zu $\vartheta \leq \frac{k\cdot(t+1)}{n}$ äquivalent ist, und da ϑ eine ganze Zahl ist zu $\vartheta \leq \left[\frac{k\cdot(t+1)}{n}\right]$. Das Maximum wird daher in $\hat{\vartheta}(k) := \left[\frac{k\cdot(t+1)}{n}\right]$ angenommen, und die Maximum-Likelihood-Schätzung für $g(\vartheta) := \frac{\vartheta}{t}$ beträgt:

$$S(k) = \frac{\left[\frac{k\cdot(t+1)}{n}\right]}{t}.$$

9.4.4

Im Fall von Verteilungen $\mathbf{P}_\vartheta$ mit stetigen Lebesgue-Dichten f_ϑ auf $\mathbb{R}^n$ (wo die Wahrscheinlichkeiten von Elementarereignissen $\{\omega\}$ 0 sind) bestimmen wir $\hat{\vartheta}(\omega)$ als Maximumstelle der Likelihood-Funktion $\vartheta \mapsto f_\vartheta(\omega)$, wenn $\omega \in \mathbb{R}^n$ gemessen wurde.

9.4.5

Bei einer n mal wiederholten Messung einer unbekannten Größe $\mu \in \mathbb{R}$ treten normalverteilte Meßfehler mit Streuung $\sigma > 0$ auf; wir verwenden also das statistische Experiment $\left(\mathbb{R}^n, \mathfrak{B}(\mathbb{R}^n), \{\mathbf{P}_\vartheta : \vartheta = (\mu, \sigma) \in \mathbb{R} \times \mathbb{R}\}\right)$, wobei das Maß $\mathbf{P}_{\mu,\sigma}$ die Lebesgue-Dichte

$$x \mapsto f_{\mu,\sigma}(x) := \frac{1}{\sqrt[n]{2\pi\sigma^2}} \cdot \exp\left(-\frac{1}{2\sigma^2} \cdot \sum_{j=1}^{n} (x_j - \mu)^2\right)$$

auf $\mathbb{R}^n$ hat.
Um die Maximum-Likelihood-Schätzung im Fall des Meßergebnisses $x = (x_1, \ldots, x_n)$ zu bestimmen, müssen wir also das Paar (μ, σ) berechnen, für das $f_{\mu,\sigma}(x)$ am größten ist. Dazu setzen

wir die partiellen Ableitungen $\frac{\partial}{\partial\mu} f_{\mu,\sigma}(x)$ und $\frac{\partial}{\partial\sigma} f_{\mu,\sigma}(x)$ Null; dies führt für den ersten Fall zur Gleichung

$$\frac{1}{\sigma^2} \cdot \sum_{j=1}^{n} (x_j - \mu) \cdot f_{\mu,\sigma}(x) = 0,$$

deren einzige Lösung

$$\hat{\mu} := \overline{x} := \frac{1}{n} \cdot \sum_{j=1}^{n} x_j$$

ist. Nun setzen wir die andere partielle Ableitung Null, was nach kurzer Rechnung auf die Gleichung

$$\left(\sum_{j=1}^{n} (x_j - \mu)^2 - n\sigma^2 \right) \cdot \frac{1}{\sigma^3} \cdot f_{\mu,\sigma}(x) = 0$$

führt. Das einzig mögliche Extremum liegt also bei

$$(\hat{\mu}, \hat{\sigma}) = \left(\overline{x}, \sqrt{\frac{1}{n} \cdot \sum_{k=1}^{n} (x_j - \overline{x})^2} \right),$$

und daß tatsächlich ein Maximum vorliegt folgt daraus, daß $f_{\mu,\sigma}(x)$ für μ beziehungsweise $\sigma \to 0$ oder ∞ gegen 0 geht.
Die Maximum-Likelihood-Schätzung des Parameters $g(\mu, \sigma) = \mu$ ist somit der Mittelwert

$$S(x) = g(\hat{\mu}(x), \hat{\sigma}(x)) = \overline{x},$$

während die Maximum-Likelihood-Schätzung für die Varianz $g_1(\mu, \sigma) := \sigma^2$ durch den Schätzer

$$S_1(X) = g_1(\hat{\mu}(x), \hat{\sigma}(x)) = \frac{1}{n} \cdot \sum_{j=1}^{n} (x_j - \overline{x})^2$$

ist.

Um Schätzer auszuschließen, die im Mittel zu niedrige oder zu hohe Schätzungen ergeben, wird der Begriff des unverfälschten Schätzers eingeführt:

9.5 Definition

Ein Schätzer S für g heißt unverfälscht oder erwartungstreu, falls S für jedes $\vartheta \in \Theta$ bezüglich $\mathbf{P}_\vartheta$ einen Erwartungswert besitzt (wir bezeichnen diesen, um das verwendete Maß klarzustellen, dann mit $\mathbf{E}_\vartheta(S) := \mathbf{E}_{\mathbf{P}_\vartheta}(S)$) und $\mathbf{E}_\vartheta(S) = g(\vartheta)$ für alle $\vartheta \in \Theta$ gilt.

9.6 Beispiele

9.6.1

In (9.4.2) sei $\overline{X}_n$ der Maximum-Likelihood-Schätzer und X_j das Ergebnis des j-ten Experiments, $X_j(\omega_1, \ldots, \omega_n) := \omega_j$ für $1 \leq j \leq n$. Diese Schätzer sind beide erwartungstreu, denn

$$\mathbf{E}_\vartheta(X_j) = \mathbf{P}_\vartheta\left\{(\omega_1, \ldots, \omega_n) \in \{0,1\}^n : \omega_j = 1\right\} = \vartheta$$

und

$$\mathbf{E}_\vartheta\left(\overline{X}_n\right) = \frac{1}{n} \cdot \sum_{j=1}^{n} \mathbf{E}_\vartheta(X_j) = \vartheta.$$

Da X_j den größten Teil der Daten nicht berücksichtigt, ist jetzt schon plausibel, daß es kein allzu guter Schätzer ist; wir werden später sehen, daß $\overline{X}_n$ der beste dieser Schätzer ist.

9.6.2

Wir betrachten das statistische Experiment $(\Omega = \mathbb{N}, \mathcal{POT}(\Omega), \{\mathcal{G}_\vartheta : \vartheta \in]0,1]\})$; $X : \Omega \to \mathbb{R}$ sei durch $X(\omega) := \omega$ definiert. Dann ist nach (5.6.1)

$$\mathbf{E}_\vartheta(X) = \frac{1}{\vartheta} - 1,$$

also X ein erwartungstreuer Schätzer für $g(\vartheta) := \frac{1}{\vartheta}$. Ein erwartungstreuer Schätzer für ϑ selbst ist nun nicht etwa $\frac{1}{X+1}$, denn $\mathbf{E}_\vartheta\left(\frac{1}{X+1}\right) = -\vartheta \cdot \ln\left(\frac{\vartheta}{1-\vartheta}\right)$, sondern der Schätzer $S := 1_{\{0\}}$ wegen $\mathbf{E}_\vartheta(S) = \mathcal{G}_\vartheta(0) = \vartheta$.

9.6.3

Es sei $(X_1, \ldots, X_n)$ eine Folge unabhängiger identisch verteilter Zufallsvariablen mit gemeinsamer Verteilung $\mathbf{P}_\vartheta$ (bei unbekanntem $\vartheta \in \Theta$) derart, daß $\mu(\vartheta) := \mathbf{E}_\vartheta(X_1)$ für alle $\vartheta \in \Theta$ existiert. Dann ist das Stichprobenmittel

$$\overline{X}_n := \frac{1}{n} \cdot \sum_{j=1}^{n} X_j$$

ein erwartungtreuer Schätzer für μ. Der in (9.4.5) gefundene Maximum-Likelihood-Schätzer für den Erwartungswert einer normalverteilten Meßgröße ist also erwartungstreu.

9.6.4

Setzen wir in (9.6.3) noch zusätzlich voraus, daß $\sigma^2(\vartheta) := \mathbf{V}_\vartheta(X_1)$ für alle $\vartheta \in \Theta$ existiert. Dann ist

$$S := \frac{1}{n-1} \cdot \sum_{j=1}^{n} \left(X_j - \overline{X}_n\right)^2$$

ein erwartungtreuer Schätzer für σ^2.
Dies folgt aus

$$\begin{aligned}
\mathbf{E}_\vartheta\left(\left(X_k - \overline{X}_n\right)^2\right) &= \mathbf{V}_\vartheta\left(X_k - \overline{X}_n\right) \\
&= \mathbf{V}_\vartheta\left(\frac{n-1}{n} \cdot X_k - \frac{1}{n} \cdot \sum_{j \neq k} X_j\right) \\
&= \frac{(n-1)^2}{n^2} \cdot \mathbf{V}_\vartheta(X_k) + \frac{1}{n^2} \cdot \sum_{j \neq k} \mathbf{V}_\vartheta(X_j) \\
&= \frac{(n-1)^2}{n^2} \cdot \mathbf{V}_\vartheta(X_1) + \frac{n-1}{n^2} \cdot \mathbf{V}_\vartheta(X_1) \\
&= \frac{n-1}{n} \cdot \sigma^2(\vartheta)
\end{aligned}$$

für alle $k = 1, \ldots, n$ und damit

$$\mathbf{E}_\vartheta(S) = \frac{1}{n-1}\cdot$$

Es stellt sich also heraus, daß der in (9.4.5) erhaltene Maximum-Likelihood-Schätzer für die Varianz einer normalverteilten Meßgröße nicht erwartungstreu ist.

Wir suchen nun unter den erwartungstreuen Schätzern S (von denen es mehrere geben kann, vergleiche (9.6.a)) denjenigen, dessen mittlerer quadratischer Fehler $\mathbf{E}_\vartheta\left((S-g(\vartheta))^2\right)$ am kleinsten ist.

9.7 Definition

Es sei S ein Schätzer für den Parameter g derart, daß $\mathbf{V}_\vartheta(S)$ für alle $\vartheta \in \Theta$ existiert. Dann heißt S erwartungstreuer Minimalschätzer, falls S erwartungstreu ist und für jeden weiteren erwartungstreuen Schätzer T

$$\mathbf{E}_\vartheta\left((S-g(\vartheta))^2\right) \leq \mathbf{E}_\vartheta\left((T-g(\vartheta))^2\right)$$

für alle $\vartheta \in \Theta$ gilt. Wegen der Erwartungstreue von S und T ist dies genau dann der Fall, wenn $\mathbf{V}_\vartheta(S) \leq \mathbf{V}_\vartheta(T)$, weshalb man auch vom Minimalvarianzschätzer spricht.

wir zeigen nun, daß die relative Häufigkeit im Beispiel (9.4.1) einen erwartungstreuen Minimalschätzer darstellt.

9.8 Satz

Es sei $\Omega := \{0,1\}^n$ und $\mathbf{P}_\vartheta := \mathcal{B}_{1,\vartheta} \otimes \ldots \otimes \mathcal{B}_{1,\vartheta}$ (mit n Faktoren) für $\vartheta \in [0,1]$, sowie $X_j(\omega_1, \ldots, \omega_n) := \omega_j$. Dann ist

$$\overline{X}_n := \frac{1}{n} \cdot \sum_{j=1}^{n} X_j$$

der erwartungstreue Minimalschätzer für $g(\vartheta) := \vartheta$.

Beweis:

Wegen (9.6.3) ist $\overline{X}_n$ erwartungstreu. Sei nun $S := \sum_{j=1}^{n} X_j$, T ein weiterer erwartungstreuer Schätzer und $t_k := \sum_{\omega \in \Omega : S(\omega)=k} T(\omega)$ für $0 \leq k \leq n$. Da T erwartungstreu ist und $\mathbf{P}_\vartheta\{\omega\} = \vartheta^k \cdot (1-\vartheta)^{n-k}$ für alle ω mit $S(\omega) = k$, gilt

$$\vartheta = \mathbf{E}_\vartheta(T) = \sum_{k=0}^{n} \vartheta^k \cdot (1-\vartheta)^{n-k} \cdot \sum_{S(\omega)=k} T(\omega) = \sum_{k=0}^{n} \vartheta^k \cdot (1-\vartheta)^{n-k} \cdot t_k.$$

Nun ist auch $\overline{X}_n$ erwartungstreu, und deshalb folgt

$$\sum_{k=0}^{n} \frac{k}{n} \cdot \binom{n}{k} \cdot \left(\frac{\vartheta}{1-\vartheta}\right)^k = \frac{\mathbf{E}_\vartheta\left(\overline{X}_n\right)}{(1-\vartheta)^n} = \frac{\mathbf{E}_\vartheta(T)}{(1-\vartheta)^n} = \sum_{k=0}^{n} \left(\frac{\vartheta}{1-\vartheta}\right)^k \cdot t_k$$

für alle $\vartheta \in [0,1[$, d.h.

$$\sum_{k=0}^{n} \frac{k}{n} \cdot \binom{n}{k} \cdot x^k = \sum_{k=0}^{n} t_k \cdot x^k$$

für alle $x := \frac{\vartheta}{1-\vartheta} > 0$. Wegen der Eindeutigkeit der Taylor-Entwicklung gilt $t_k = \frac{k}{n} \cdot \binom{n}{k}$ für $0 \leq k \leq n$.

Da die Summe der Quadrate von $\binom{n}{k}$ Zahlen bei vorgegebener Summe der Zahlen minimal ist, wenn alle Zahlen gleich sind (Cauchy-Schwarz'sche-Ungleichung), folgt leicht, daß

$$\sum_{S(\omega)=k} (T(\omega))^2 \geq \binom{n}{k}^{-1} \cdot \left(\sum_{S(\omega)=k} T(\omega)\right)^2 = \binom{n}{k}^{-1} \cdot t_k^2$$

gilt, und somit

$$\begin{aligned}
\mathbf{E}_\vartheta\left(T^2\right) &= \sum_{k=0}^{n} \vartheta^k \cdot (1-\vartheta)^{n-k} \cdot \sum_{S(\omega)=k} (T(\omega))^2 \\
&\geq \sum_{k=0}^{n} \vartheta^k \cdot (1-\vartheta)^{n-k} \cdot \binom{n}{k}^{-1} \cdot t_k^2 \\
&= \sum_{k=0}^{n} \vartheta^k \cdot (1-\vartheta)^{n-k} \cdot \binom{n}{k} \cdot \frac{k^2}{n^2} \\
&= \mathbf{E}_\vartheta\left(\overline{X}_n^2\right).
\end{aligned}$$

Nun gilt im Falle von erwartungstreuen Schätzern für den mittleren quadratischen Fehler

$$\mathbf{E}_\vartheta\left((T-g(\vartheta))^2\right) = \mathbf{E}_\vartheta\left((T-\mathbf{E}_\vartheta(T))^2\right) = \mathbf{V}_\vartheta(T) = \mathbf{E}_\vartheta\left(T^2\right) - (g(\vartheta))^2$$

$$\geq \mathbf{E}_\vartheta\left(\overline{X}_n^2\right) - (g(\vartheta))^2 = \mathbf{E}_\vartheta\left(\left(\overline{X}_n - g(\vartheta)\right)^2\right)$$

$\square$

Anhang A: Sonderveranstaltung - Lehramt Sek. II

1. Das Ziegenspiel

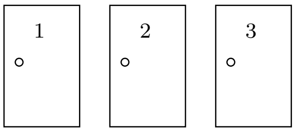
Bei einer Fernseh-Show darf der Gewinner seinen Preis durch ein Glücksspiel auswählen: Er steht vor 3 geschlossenen Türen und weiß, daß hinter einer von ihnen ein Auto steht.

Nachdem der Spieler sich für eine der drei Türen entschieden hat, macht der Quizmaster, der weiß wo das Auto steht, eine der beiden restlichen Türen auf, hinter der eine Ziege steht.

Er erlaubt dem Spieler jedoch, sich jetzt nochmals zu entscheiden und die dritte Tür zu wählen. Soll der Spieler das tun?

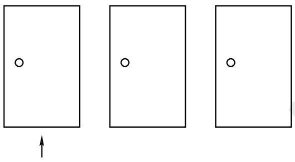
1.1 Berechnung der Wahrscheinlichkeiten

Schritt 0:



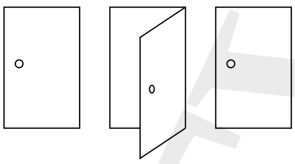
zwei Ziegen und ein Auto

Schritt 1:



Spieler wählt eine Tür

Schritt 2:



Der Quizmaster öffnet eine der beiden übrigen Türen. hinter der eine Ziege steht.

Schritt 3: Spieler kann (muß aber nicht) wechseln.

Der Spieler kann zwei Strategien verfolgen:

(1) Spieler wechselt die Tür

(2) Spieler behält die zuerst gewählte Tür

Frage: Welche der Strategien sollte er verfolgen?

In der zweiten Strategie entscheidet sich der Spieler zufällig für eine der drei Türen. Er beachtet nicht, was der Quizmaster danach macht.

Spieler verliert

$\Leftrightarrow$ Der Tip war falsch.

$\Leftrightarrow$ Hinter der gewählten Tür steht eine Ziege. Die Wahrscheinlichkeit dafür ist $\frac{2}{3}$.

Also: Der Spieler gewinnt mit Wahrscheinlichkeit $\frac{1}{3}$.

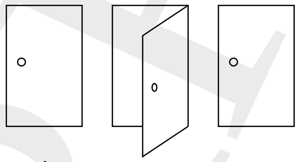
In der ersten Strategie verliert der Spieler beim Wechseln.

$\Leftrightarrow$ Der Tip im ersten Schritt war richtig

$\Leftrightarrow$ Hinter der getippten Tür steht das Auto, also mit Wahrscheinlichkeit $\frac{1}{3}$.

Also: Der Spieler gewinnt mit Wahrscheinlichkeit $\frac{2}{3}$ bei der ersten Strategie.

Wenn man die Vorgeschichte (Schritt 0 - Schritt 2) außer acht läßt, also nur weiß, daß hinter einer der 2 geschlossenen Türen ein Auto steht und hinter der anderen eine Ziege, so ist die Wahrscheinlichkeit zu gewinnen (wenn man sich zum Beispiel durch Werfen einer Münze entscheidet) glleich $\frac{1}{2}$.



Diese ist kleiner als die Gewinn-Wahrscheinlichkeit bei der ersten Strategie.

Die Frage ist also, ob man nicht auch aus der Vorgeschichte Informationen gewinnen kann, die auf die Tür hindeuten, hinter der das Auto steht.

Tatsächlich ist dies so. Der Quizmaster ist nicht frei von der Wahl der Tür, die er öffnet.

Der Spieler wähle im ersten Schritt Tür 1.

- A_1: Steht das Auto hinter Tür 1, so kann er jede der beiden Türen $2 \vee 3$ öffnen, er öffnet in diesem Fall Tür 2 mit Wahrscheinlichkeit $\frac{1}{3} \cdot \frac{1}{2}$.
- A_2: Steht das Auto hinter Tür 2, so wird der Quizmaster diese mit Wahrscheinlichkeit $\frac{1}{3} \cdot 0 = 0$ öffnen.
- A_3: Steht das Auto hinter Tür 3, so wird Tür 2 mit Wahrscheinlichkeit $\frac{1}{3} \cdot 1 = \frac{1}{3}$ geöffnet.

Das Ereignis

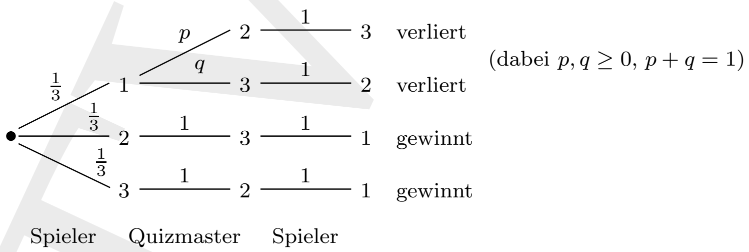
$$A = \{\text{Quizmaster öffnet Tür } 2\} = A_1 \cup A_3$$

setzt sich aus den Ereignissen A_1 und A_3 zusammen, wobei A_3 wahrscheinlicher ist als A_1.

Deshalb ist es „vernünftig“, beim Eintreten von A darauf zu setzen, daß das wahrscheinlichere Ereignis $A_3 \subset A = A_1 \cup A_3$ eingetreten ist.

Andere Lösungsmethode:

Verwendung eines Baumdiagramms $\Omega = \{1, 2, 3\}$, ohne Einschränkung: $1 \mathrel{\hat{=}}$ Auto.



$$\mathbf{P}\{\text{Spieler gewinnt}\} = \frac{1}{3} \cdot 1 \cdot 1 + \frac{1}{3} \cdot 1 \cdot 1 = \frac{1}{3} + \frac{1}{3} = \frac{2}{3}$$

$$\mathbf{P}\{\text{Spieler verliert}\} = \frac{1}{3} \cdot p \cdot 1 + \frac{1}{3} \cdot q \cdot 1 = \frac{1}{3}(p+q) = \frac{1}{3} \cdot 1 = \frac{1}{3}.$$

2. Deskriptive (=beschreibende) Statistik

2.1 Einige Grundbegriffe

Eine Beobachtungsmenge ist in der beschreibenden Statistik eine Menge, deren Elemente auf ein bestimmtes Merkmal hin betrachtet werden; die Elemente dieser Beobachtungsmenge werden Beobachtungseinheiten oder Merkmalsträger genannt.

Für statistische Zwecke sind nur solche Merkmale von Interesse, die verschiedene Realisierungen - genannt Ausprägungen - besitzten.

Bei einer statistischen Erhebung wird von jeder Einheit einer Beobachtungsmenge festgehalten, welche der vorgegebenen Ausprägungen sie besitzt. Eine statistische Erhebung ist vom Umfang n, wenn bei dieser Erhebung n Beobachtungseinheiten erfaßt werden.

2.1.1 Einteilung der Merkmale

- Quantitative Merkmale: Merkmale, deren Ausprägungen sich durch Messen oder Zählen bestimmen lassen.
- Qualitative Merkmale: Merkmale, deren Ausprägungen sich weder durch Messen noch durch Auszählen bestimmen lassen.
- Anordbare Merkmale: Qualitative Merkmale, deren Ausprägungen sich anordnen lassen.

Eine Urliste (Beobachtungsreihe) bekommt man, wenn man bei einer Erhebung die Ausprägungen der einzelnen Beobachtungseinheiten in der Reihenfolge aufschreibt, wie man sie bei der Erhebung gewinnt. Die in einer Urliste stehenden Ausprägungen bilden die bei einer Erhebung gewonnenen Daten. Bezeichnung: $x_1, \ldots, x_n$.

2.2 Auswerten von Erhebungen durch Häufigkeiten

2.2.1 Definition (absolute/ relative Häufigkeit)

Die absolute Häufigkeit einer Ausprägung a_1 in einer Erhebung ist die Anzahl n_i, mit der a_i in dieser Erhebung vorkommt.

Die ralative Häufigkeit von a_i in einer Erhebung vom Umfang n ist

$$h_i := \frac{\text{absolute Häufigkeit}}{\text{Umfang}} = \frac{n_i}{n}.$$

Merkmal x habe s Ausprägungen, dann ist

$$\sum_{i=1}^{s} n_i = n \qquad \sum_{i=1}^{s} h_i = 1.$$

2.2.2 Graphische Darstellung

- Stabdiagramme
- Histogramme
- Säulendiagramme
- Kreisdiagramme

2.2.3 Klassifizierung der Ausprägungen

Die Ausprägungen eines quantitativen Merkmals liegen im Intervall $[b, c]$. $[b, c]$ wird in Intervallklassen der Form $[b_0, b_1], [b_1, b_2], \ldots, [b_{k-1}, b_k]$ mit $b = b_0$ und $b_k = c$ zerlegt.

$$\begin{aligned} n_i^* &= \text{Anzahl der in } k_i \text{ liegenden Beobachtungswerte} \\ h_i^* &= \frac{n_i^*}{n} \text{ realtive Häufigkeit von } k_i \text{ in einer Erhebung vom Umfang } n. \end{aligned}$$

2.2.4 Beschreibung von Häufigkeiten durch Parameter

- Lageparameter: Parameter, um die sich die Beobachtungswerte der Erhebung gruppieren.
- Streuungsparameter: Parameter, die die Streuung der Beobachtungswerte um einen Lageparameter beschreiben.

2.2.5 Definition

Der Durchschnitt $\overline{x}$ der Urliste $x_1, \ldots, x_n$ eines quantitativen Mermals ist gegeben durch

$$\overline{x} = \frac{1}{n} \cdot (x_1 + \ldots + x_n).$$

Es gilt:

$$\sum_{i=1}^{n} (x_i - \overline{x}) = 0.$$

2.2.6 Beispiel

Bei einen Test in zwei Klassen ergeben sich folgende Noten:
Klasse a: $2, 5, 4, 2, 4, 4, 3, 1, 4, 6, 3, 4, 3, 4, 5$;
Klasse b: $3, 5, 3, 4, 4, 3, 6, 2, 1, 5, 6, 4, 3, 2, 4, 5, 3, 6, 2, 1$.
Für die Durchschnitte ergeben sich also folgende Werte

$$\begin{aligned} \overline{x} &= 3,6 \quad \text{Klasse } a \\ \overline{y} &= 3,6 \quad \text{Klasse } b \end{aligned}$$

und für die angeordneten Urlisten folgt:
Klasse a: $1, 2, 2, 3, 3, 3, 4, 4, 4, 4, 4, 4, 5, 5, 6$;
Klasse b: $1, 1, 2, 2, 2, 3, 3, 3, 3, 3, 4, 4, 4, 4, 5, 5, 5, 6, 6, 6$.

2.2.7 Definition (Median)

Es sei $x_1, \ldots, x_n$ eine Urliste eines qualitativen oder angeordneten Merkmals. $x_{(1)}, \ldots, x_{(n)}$ sei die angeordnete Urliste. Der Median $\tilde{x}$ der Urliste ist jeder Wert, für den gilt: Höchstens 50% aller Beobachtungswerte liegen vor dem Median und höchstens 50% der Beobachtungswerte kommen nach dem Median.

2.2.8 Satz (ohne Beweis)

Für den Median $\tilde{x}$ gilt:

- $$\tilde{x} = x_{\left(\frac{n+1}{2}\right)} \quad \text{falls } n \text{ ungerade.}$$

- Bei geradem n kann für $\tilde{x}$ jeder Wert zwischen $x_{\left(\frac{n}{2}\right)}$ und $x_{\left(\frac{x}{2}+1\right)}$ einschließlich dieser beiden Werte genommen werden.
$$\tilde{x} = \frac{1}{2} \cdot \left(x_{\left(\frac{n}{2}\right)} + x_{\left(\frac{n}{2}+1\right)}\right)$$

2.2.9 Definition (p-Quantil)

Der p-Quantil einer Urliste, eines qualitativen oder anordbaren Merkmals ist jeder Wert, für den gilt:

Höchstens 100% der Beobachtungswerte liegen vor dem Quantil und höchstens $(100 - 100 \cdot p)\,\%$ der Beobachtungswerte kommen nach dem Quantil ($0 < p < 1$).

Bezeichnung:

$$\tilde{x}_p = \begin{cases} \frac{1}{2} \cdot \left(x_{(n \cdot p)} + x_{(n \cdot p + 1)}\right) & \text{falls } n \cdot p \in \mathbb{N} \\ x_{([n \cdot p]+1)} & \text{sonst.} \end{cases}$$

Bemerkung:

Der Merdian ist der $\frac{1}{2}$-Quantil.

Zum Beispiel: siehe (2.2.6)

Klasse a:

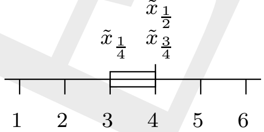
$$\begin{aligned} \tilde{x}_{\frac{1}{2}} &= x_{\left(\left[15 \cdot \frac{1}{2}\right]+1\right)} = x_{(7+1)} = x_{(8)} = 4 \\ \tilde{x}_{\frac{1}{4}} &= x_{\left(\left[15 \cdot \frac{1}{4}\right]+1\right)} = x_{(3+1)} = x_{(4)} = 3 \\ \tilde{x}_{\frac{3}{4}} &= x_{\left(\left[15 \cdot \frac{3}{4}\right]+1\right)} = x_{(11+1)} = x_{(12)} = 4 \end{aligned}$$

Klasse b:

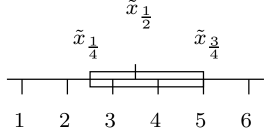
$$\begin{aligned} \tilde{x}_{\frac{1}{4}} &= \frac{1}{2} \cdot \left(x_{\left(20 \cdot \frac{1}{4}\right)} + x_{\left(20 \cdot \frac{1}{4}+1\right)}\right) = \frac{1}{2} \cdot \left(x_{(5)} + x_{(6)}\right) = \frac{1}{2} \cdot (2+3) = 2,5 \\ \tilde{x}_{\frac{1}{4}} &= \frac{1}{2} \cdot \left(x_{\left(20 \cdot \frac{1}{2}\right)} + x_{\left(20 \cdot \frac{1}{2}+1\right)}\right) = \frac{1}{2} \cdot \left(x_{(10)} + x_{(11)}\right) = \frac{1}{2} \cdot (3+4) = 3,5 \\ \tilde{x}_{\frac{3}{4}} &= \frac{1}{2} \cdot \left(x_{\left(20 \cdot \frac{3}{4}\right)} + x_{\left(20 \cdot \frac{3}{4}+1\right)}\right) = \frac{1}{2} \cdot \left(x_{(15)} + x_{(16)}\right) = \frac{1}{2} \cdot (5+5) = 5 \end{aligned}$$

Zur Veranschaulichung:

Klasse a:



Klasse b:



2.2.10 Definition

Sei $x_1, \ldots, x_n$ eine Urliste eines quantitativen Merkmals mit Durchschnitt $\overline{x}$. Dann ist

$$s^2 := \frac{1}{n} \cdot \sum_{i=1}^{n} (x_i - \overline{x})^2$$

die Varianz (mittlere quadratische Abweichung),

$$s := \sqrt{s^2}$$

die Standardabweichung und

$$d := \frac{1}{n} \cdot \sum_{i=1}^{n} |x_i - \overline{x}|$$

die mittlere absolute Abweichung.
Es gilt:

$$s^2 = \frac{1}{n} \cdot \sum_{i=1}^{n} x_i^2 - \left(\frac{1}{n} \cdot \sum_{i=1}^{n} x_i \right)^2 .$$

2.3 Beispiel

Ein Würfel wird 600-mal geworfen und die jeweils geworfene Augenzahl notiert.

2.3.1 Urliste:

2	6	1	6	5	1	6	2	2	5	6	1	2	3	3	2	6	5	4	6
6	3	3	6	4	5	4	1	6	2	6	2	2	3	2	1	4	2	3	6
6	5	2	3	2	3	6	1	2	6	1	3	4	6	5	3	6	6	5	4
6	5	3	1	4	1	1	1	3	3	5	4	5	6	5	3	2	2	3	1
6	6	2	4	2	3	6	4	3	5	2	2	6	3	5	3	2	5	5	3
4	5	4	3	3	5	4	6	2	1	1	4	3	4	2	3	4	5	5	4
6	2	6	1	3	5	3	2	4	6	3	2	4	1	1	1	5	6	1	5
2	2	2	4	1	2	1	6	5	6	4	4	4	2	5	6	5	4	3	1
6	4	6	1	1	3	2	6	4	2	1	4	6	6	2	5	5	3	2	2
6	3	6	3	2	3	6	5	5	1	5	2	6	6	6	4	6	6	5	4

1	1	1	5	3	6	4	4	5	2	6	1	5	3	3	6	2	6	5	1
4	3	1	6	1	5	1	6	5	4	5	5	2	1	2	4	5	6	3	2
3	3	2	6	5	6	5	5	4	3	6	3	3	3	6	4	4	5	6	3
2	1	3	6	3	6	4	2	3	5	5	3	6	2	6	4	1	2	3	2
3	3	4	5	5	3	3	1	2	5	4	5	1	4	5	5	6	4	2	1
3	5	3	3	2	5	2	2	6	5	5	1	6	2	2	2	5	6	6	2
1	6	2	2	6	6	3	6	2	6	3	2	6	6	1	2	4	4	3	2
3	3	2	1	4	4	2	5	3	1	2	2	5	3	1	1	3	3	1	3
4	1	1	6	1	3	2	6	5	4	5	5	6	5	6	5	5	5	6	4
3	1	3	2	3	2	4	1	3	4	1	5	2	4	1	5	1	1	1	6

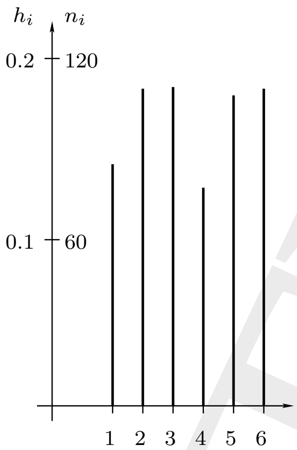
4	6	5	2	4	3	6	5	6	1	1	3	1	3	1	6	3	2	2	5
5	3	5	2	5	4	5	1	4	1	5	2	2	3	4	2	4	2	2	6
3	6	2	2	3	3	6	5	2	5	3	3	1	3	3	2	5	4	3	5
3	5	2	2	4	6	3	5	3	4	3	5	6	1	1	2	5	3	3	5
6	1	1	2	5	2	3	6	4	1	2	4	1	6	5	5	6	3	3	4
3	6	4	3	5	2	1	4	6	5	6	3	2	6	4	1	5	4	2	4
6	3	6	1	5	1	3	4	1	2	2	2	6	5	6	2	1	6	4	6
3	6	2	4	5	5	1	5	3	5	3	5	1	6	1	5	3	2	5	2
4	6	2	3	6	2	2	6	1	6	3	5	5	3	3	6	1	4	2	2
2	3	5	1	5	2	5	1	3	4	4	5	1	5	4	2	2	6	2	3

2.3.2 Häufigkeiten

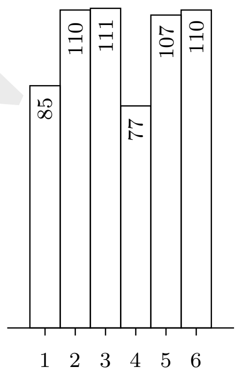
Augenzahl a	1	2	3	4	5	6
absolute Häufigkeit n_i	85	110	111	77	107	110
relative Häufigkeit h_i	$0,142$	$0,183$	$0,185$	$0,128$	$0,178$	$0,183$

2.3.3 Diagramme

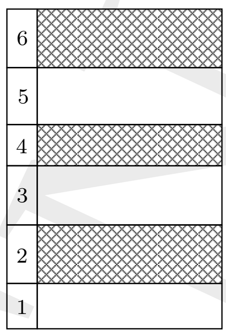
Stabdiagramm:



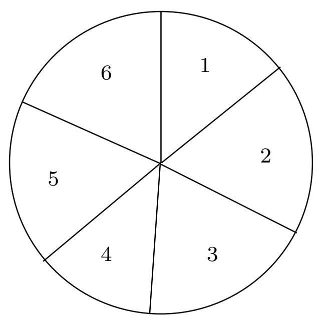
Histogramm:



Säulendiagramm:



Kreisdiagramm:



3. Binomialverteilung

3.1 Definition aus der Vorlesung

$\Omega = \{0, 1, \ldots, n\}$ $(n \geq 1)$, $\mathfrak{A} = \mathcal{POT}\{\Omega\}$, $p \in [0, 1]$. Dann ist durch

$$f : \{0, 1, \ldots, n\} \to \mathbb{R}_+ \qquad f(k) := \binom{n}{k} \cdot p^k \cdot (1-p)^{n-k}$$

eine Zähldichte definiert.

Das zugehörige Wahrscheinlichkeitsmaß auf $(\Omega, \mathfrak{A})$ heißt Binomialverteilung $\mathcal{B}_{n,p}$.

Speziell für $n = 1$: $\mathcal{B}_{1,p}$ Bernoulli-Verteilung. Ein Zufallsexperiment heißt Bernoulli-Experiment, wenn es nur zwei Ergebnisse hat: Treffer ($\widehat{=} 1$) und Niete ($\widehat{=} 0$). p sei die Wahrscheinlichkeit für Treffer, $q = 1 - p$ die Wahrscheinlichkeit für Niete.

3.1.1 Beispiel

Eine Münze wird 4-mal geworfen. Jeder Wurf ist ein Bernoulli-Experiment mit der Ergebnismenge $\{0, 1\}$, wobei $0 \,\widehat{=}\,$ Zahl, $1 \,\widehat{=}\,$ Wappen.

3.1.2 Bezeichnung

Ein Zufallsexperiment, das aus n unabhängigen Durchführungen desselben Bernoulli-Experiments besteht, heißt Bernoulli-Kette der Länge n.

Ein Bernoulli-Experiment mit der Wahrscheinlichkeit p für Treffer werde n-mal durchgeführt. Die Durchführungen seien unabhängig, dann beträgt die Wahrscheinlichkeit für genau k Treffer $(k = 0, 1, \ldots, n)$ in dieser Bernoulli-Kette

$$\mathcal{B}_{n,p}(\{k\}) = \binom{n}{k} \cdot p^k \cdot (1-p)^{n-k}.$$

3.1.3 Beispiel

Eine verfälschte Münze wird 4-mal geworfen.

$$p = \mathbf{P}(\text{Wappen} \,\widehat{=}\, 1) = \frac{5}{8} \qquad q = 1 - p = \mathbf{P}(\text{Zahl} \,\widehat{=}\, 0) = \frac{3}{8}$$

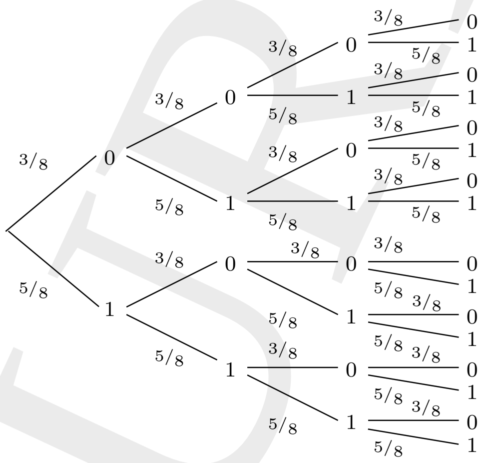
$$\mathbf{P}\{\text{Genau zweimal Wappen}\} = \binom{4}{2} \cdot \left(\frac{5}{8}\right)^2 \cdot \left(\frac{3}{8}\right)^2 \approx 0,33.$$

3.1.4 Anschaulich

Günstige Pfade:

$$(0,0,1,1)$$
$$(0,1,0,1)$$
$$(1,0,0,1)$$
$$(0,1,1,0)$$
$$(1,0,1,0)$$
$$(1,1,0,0)$$

$$\# = 6 = \binom{4}{2}$$



Jeder Pfad hat die Wahrscheinlichkeit $\left(\frac{5}{8}\right)^2 \cdot \left(\frac{3}{8}\right)^2$, der Binomialkoeffizient gibt die Anzahl der günstigen Pfade an. Auf $\binom{4}{2}$ Möglichkeiten (allgemein $\binom{n}{k}$) kann man die 2 (allgemein k) Treffer im Tupel der Länge 4 (allgemein n) unterbringen.

Zur Vereinfachung der Schreibweise führt man eine Zufallsvariable X ein.

X nimmt die Werte $0, 1, \ldots, n$ an und gibt die Anzahl der Treffer in der Bernoulli-Kette an.

$$\begin{aligned}
\mathbf{P}\{\text{Genau } k \text{ Treffer}\} &= \mathbf{P}\{X = k\} = \mathcal{B}_{n,p}(\{k\}) = \binom{n}{k} \cdot p^k \cdot (1-p)^{n-k} \\
\mathbf{P}\{\text{Mindestens } k \text{ Treffer}\} &= \mathbf{P}\{X \geq k\} = \mathbf{P}\{X = k\} + \mathbf{P}\{X = k+1\} + \ldots + \mathbf{P}\{X = n\} \\
&\hat{=} \mathcal{B}_{n,p}(\{k, k+1, \ldots, n\})
\end{aligned}$$

3.1.5 Beispiel

Eine Maschine produziert Bleistifte; 8% der Produktion ist Ausschuß. Wie groß ist die Wahrscheinlichkeit, daß von 10 zufällig herausgegriffenen Bleistiften mehr als 5 Ausschuß sind?

Bernoulli-Kette der Länge $n = 10$.

Trefferwahscheinlichkeit $p = 0,08$ (Wahrscheinlichkeit für Ausschuß).

X gibt die Anzahl der Treffer an (die Anzahl der Ausschußstücke)

$$\begin{aligned}
\mathbf{P}\{\text{Mehr als 5 Ausschuß}\} &= \mathbf{P}\{X > 5\} = \mathbf{P}\{X \geq 6\} = \mathbf{P}\{X = 6\} + \mathbf{P}\{X = 7\} \\
&\quad + \mathbf{P}\{X = 8\} + \mathbf{P}\{X = 9\} + \mathbf{P}\{X = 10\} \\
&= \binom{10}{6} \cdot 0,08^6 \cdot 0,92^4 + \binom{10}{7} \cdot 0,08^7 \cdot 0,92^3 \\
&\quad + \binom{10}{8} \cdot 0,08^8 \cdot 0,92^2 + \binom{10}{9} \cdot 0,08^9 \cdot 0,92^1 \\
&\quad + \binom{10}{10} \cdot 0,08^{10} \cdot 0,92^0 \\
&= 210 \cdot 0,08^6 \cdot 0,92^4 + 120 \cdot 0,08^7 \cdot 0,92^3 \\
&\quad + 45 \cdot 0,08^8 \cdot 0,92^2 + 20 \cdot 0,08^9 \cdot 0,92 + 0,08^{10} \\
&\approx 4,15 \cdot 10^{-5}
\end{aligned}$$

Problem: Die Berechnung von $\mathcal{B}_{n,p}(\{k\})$ wird mit zunemendem n immer mühsamer. Daher liegt die Binomialverteilung für ausgewählte Werte von n, p und k tabelliert vor.

3.2 Tabelle der Wahrscheinlichkeitsfunktion $\mathcal{B}_{n,p}$

Bei der Erstellung der Tabelle wird der Rechenaufwand dadurch reduziert, daß für die Funktion $\mathcal{B}_{n,p}$ eine Symmetriebeziehung gilt:

3.2.1 Symmetriebeziehung

$$\mathcal{B}_{n,p}(\{k\}) = \mathcal{B}_{n,1-p}(\{n-k\})$$

Es genügt also. die Funktionswerte für $p \leq 0,5$ zu berechnen. Die Ermittlung der Werte für $p \geq 0,5$ erfolgt dann über die Symmetriebeziehung oder, um sich Rechenarbeit zu erparen, über den rechten Eingang (= rechte untere Ecke, die Tabelle wird mit der rechten und unteren Skala gelesen).

Bemerkung:

Die Wahrscheinlichkeiten sind auf 4 Stellen gerundet.

Deshalb kann es vorkommen, daß sich eine Summe $\neq 1$ ergibt. Nicht aufgeführte Werte sind (auf 4 Stellen gerundet) $0,0000$.

3.2.2 Tabellen

							p							
n	k	$0,02$	$0,03$	$0,04$	$0,05$	$0,10$	$\frac{1}{6}$	$0,20$	$0,30$	$\frac{1}{3}$	$0,40$	$0,50$	k	n
10	0	0,8171	0,7374	0,6648	0,5987	0,3487	0,1615	0,1074	0,0282	0,0176	0,0060	0,0010	10	10
	1	0,1667	0,2281	0,2770	0,3151	0,3874	0,3230	0,2684	0,1211	0,0867	0,0403	0,0098	9	
	2	0,0153	0,0317	0,0519	0,0746	0,1937	0,2907	0,3020	0,2335	0,1951	0,1209	0,0439	8	
	3	0,0008	0,0026	0,0058	0,0105	0,0574	0,1550	0,2013	0,2668	0,2601	0,2150	0,1172	7	
	4		0,0001	0,0004	0,0010	0,0112	0,0543	0,0881	0,2001	0,2276	0,2508	0,2051	6	
	5				0,0001	0,0015	0,0130	0,0264	0,1029	0,1366	0,2007	0,2461	5	
	6					0,0001	0,0022	0,0055	0,0368	0,0569	0,1115	0,2051	4	
	7						0,0002	0,0008	0,0090	0,0163	0,0425	0,1172	3	
	8							0,0001	0,0014	0,0030	0,0106	0,0439	2	
	9								0,0001	0,0003	0,0016	0,0098	1	
	10	Nicht aufgeführte Werte sind (auf 4 Dezimalstellen) 0,0000									0,0001	0,0010	0	
n	k	$0,98$	$0,97$	$0,96$	$0,95$	$0,90$	$\frac{5}{6}$	$0,80$	$0,70$	$\frac{2}{3}$	$0,60$	$0,50$	k	n

							p							
n	k	$0,02$	$0,03$	$0,04$	$0,05$	$0,10$	$\frac{1}{6}$	$0,20$	$0,30$	$\frac{1}{3}$	$0,40$	$0,50$	k	n
15	0	0,7386	0,6333	0,5421	0,4633	0,2059	0,0649	0,0352	0,0047	0,0023	0,0005	0,0000	15	15
	1	0,2261	0,2938	0,3388	0,3658	0,3432	0,1947	0,1319	0,0305	0,0171	0,0047	0,0005	14	
	2	0,0323	0,0636	0,0988	0,1348	0,2669	0,2726	0,2309	0,0916	0,0599	0,219	0,0032	13	
	3	0,0029	0,0085	0,0178	0,0307	0,1285	0,2363	0,2501	0,1700	0,1299	0,0634	0,0139	12	
	4	0,0002	0,0008	0,0022	0,0049	0,0428	0,1418	0,1876	0,2186	0,1948	0,1268	0,0417	11	
	5		0,0001	0,0002	0,0006	0,0105	0,0624	0,1032	0,2061	0,2143	0,1859	0,0916	10	
	6					0,0019	0,0208	0,0430	0,1472	0,1786	0,2066	0,1527	9	
	7					0,0003	0,0053	0,0138	0,0811	0,1148	0,1771	0,1964	8	
	8						0,0011	0,0035	0,0348	0,0574	0,1181	0,1964	7	
	9						0,0002	0,0007	0,0116	0,0223	0,0612	0,1527	6	
	10							0,0001	0,0030	0,0067	0,0245	0,0916	5	
	11								0,0006	0,0015	0,0074	0,0139	4	
	12								0,0001	0,0003	0,0016	0,0139	3	
	13										0,0003	0,0032	2	
	14											0,0005	1	
	15	Nicht aufgeführte Werte sind (auf 4 Dezimalstellen) 0,0000											0	
n	k	$0,98$	$0,97$	$0,96$	$0,95$	$0,90$	$\frac{5}{6}$	$0,80$	$0,70$	$\frac{2}{3}$	$0,60$	$0,50$	k	n

		p												
n	k	0,02	0,03	0,04	0,05	0,10	$\frac{1}{6}$	0,20	0,30	$\frac{1}{3}$	0,40	0,50	k	n
	0	0,6676	0,5438	0,4420	0,3585	0,1216	0,0261	0,0115	0,0008	0,0003	0,0000	0,0000	20	
	1	0,2725	0,3364	0,3683	0,3774	0,2702	0,1043	0,0576	0,0068	0,0030	0,0005	0,0000	19	
	2	0,0528	0,0988	0,1458	0,1887	0,2852	0,1982	0,1369	0,0278	0,0143	0,0031	0,0002	18	
	3	0,0065	0,0183	0,0364	0,0596	0,1901	0,2379	0,2054	0,0716	0,0429	0,0123	0,0011	17	
	4	0,0006	0,0024	0,0065	0,0133	0,0898	0,2022	0,2182	0,1304	0,0911	0,0350	0,0046	16	
	5		0,0002	0,0009	0,0022	0,0319	0,1294	0,1746	0,0789	0,1457	0,0746	0,0148	15	
	6			0,0001	0,0003	0,0089	0,0647	0,1091	0,1916	0,1821	0,1659	0,0739	14	
	7					0,0020	0,0259	0,0545	0,1643	0,1821	0,1659	0,0739	13	
	8					0,0004	0,0084	0,0222	0,1144	0,1480	0,1797	0,1201	12	
	9					0,0001	0,0022	0,0074	0,0654	0,0987	0,1597	0,1602	11	
20	10						0,0005	0,0020	0,0308	0,0543	0,1171	0,1762	10	20
	11						0,001	0,0005	0,0120	0,0247	0,0710	0,1602	9	
	12							0,0001	0,0039	0,0092	0,0355	0,1201	8	
	13								0,0010	0,0028	0,0146	0,0739	7	
	14								0,0002	0,0007	0,0049	0,0370	6	
	15									0,0001	0,0013	0,0148	5	
	16										0,0003	0,0046	4	
	17											0,0011	3	
	18											0,0002	2	
	19												1	
	20	Nicht aufgeführte Werte sind (auf 4 Dezimalstellen) 0,0000											0	
n	k	0,98	0,97	0,96	0,95	0,90	$\frac{5}{6}$	0,80	0,70	$\frac{2}{3}$	0,60	0,50	k	n

3.2.3 Beispiele

$$n = 10 \quad p = 0.3$$
$$\mathcal{B}_{10,0.3}(4) = 0,2001$$

$$n = 20 \quad p = 0.9$$
$$\mathcal{B}_{20,0.9}(12) = \mathcal{B}_{20,0.1}(8) = 0,0004$$

3.3 Tabelle der Summenfunktion $F_{n,p}$

$$\begin{aligned} F_{n,p}(k) &= \mathbf{P}\{X \leq k\} = \mathcal{B}_{n,p}(0) + \mathcal{B}_{n,p}(1) + \ldots + \mathcal{B}_{n,p}(k) \\ \mathbf{P}\{X > k\} &= 1 - \mathbf{P}\{X \leq k\} = 1 - F_{n,p}(k) \\ \mathbf{P}\{X \geq k\} &= \mathbf{P}\{X > k-1\} = 1 - F_{n,p}(k-1) \end{aligned}$$

3.3.1 Symmetriebeziehung

$$\begin{aligned} F_{n,p}(k) &= \mathbf{P}\{X \leq k\} = \mathbf{P}\{Y \geq n-k\} \\ &= 1 - \mathbf{P}\{Y < n-k\} = 1 - \mathbf{P}\{Y \leq n-k-1\} \\ &= 1 - F_{n,1-p}(n-k-1) \end{aligned}$$

X: Anzahl der Treffer, Y: Anzahl der Nieten

Für $p \geq 0,5$ erhält man die Werte $F_{n,p}(k)$, indem man den (beim rechten Eingang) abgelesenen Tabellenwert von 1 subtrahiert. Nicht aufgeführte Werte sind (auf 4 Dezimalstellen) $1,0000$.

3.3.2 Tabellen

		p												
n	k	0,02	0,03	0,04	0,05	0,10	$\frac{1}{6}$	0,20	0,30	$\frac{1}{3}$	0,40	0,50	k	n
	0	0,6676	0,5438	0,4420	0,3585	0,1216	0,0261	0,0115	0,0008	0,0003	0,0000	0,0000	19	
	1	0,9401	0,8802	0,8103	0,7358	0,3917	0,1304	0,0692	0,0076	0,0033	0,0005	0,0000	18	
	2	0,9929	0,9790	0,9561	0,9245	0,6769	0,3287	0,2061	0,0355	0,0176	0,0036	0,0002	17	
	3	0,9994	0,9973	0,9926	0,9841	0,8670	0,5665	0,4114	0,1071	0,0604	0,0160	0,0013	16	
	4		0,9997	0,9990	0,9974	0,9568	0,7687	0,6296	0,2375	0,1515	0,0510	0,0059	15	
	5			0,9999	0,9997	0,9887	0,8982	0,8042	0,4164	0,2972	0,1256	0,0207	14	
	6					0,9976	0,9629	0,9133	0,6080	0,4793	0,2500	0,0577	13	
	7					0,9996	0,9887	0,9679	0,7723	0,6615	0,4159	0,1316	12	
	8					0,9999	0,9972	0,9900	0,8867	0,8095	0,5956	0,2517	11	
20	9						0,9994	0,9974	0,9520	0,9081	0,7553	0,4119	10	20
	10						0,9999	0,9994	0,9829	0,9624	0,8725	0,5881	9	
	11							0,9999	0,9949	0,9870	0,9435	0,7483	8	
	12								0,9987	0,9963	0,9790	0,8684	7	
	13								0,9997	0,9991	0,9935	0,9423	6	
	14									0,9998	0,9984	0,9793	5	
	15										0,9997	0,9941	4	
	16											0,9987	3	
	17											0,9998	2	
	18	Nicht aufgeführte Werte sind (auf 4 Dezimalstellen) 1,0000											1	
n	k	0,98	0,97	0,96	0,95	0,90	$\frac{5}{6}$	0,80	0,70	$\frac{2}{3}$	0,60	0,50	k	n

		p												
n	k	0,02	0,03	0,04	0,05	0,10	$\frac{1}{6}$	0,20	0,30	$\frac{1}{3}$	0,40	0,50	k	n
	0	0,3642	0,2181	0,1299	0,0769	0,0052	0,0001	0,0000	0,0000	0,0000	0,0000	0,0000	49	
	1	0,7358	0,5553	0,4005	0,2794	0,0338	0,0012	0,0002	0,0000	0,0000	0,0000	0,0000	48	
	2	0,9216	0,8108	0,6767	0,5405	0,1117	0,0066	0,0013	0,0000	0,0000	0,0000	0,0000	47	
	3	0,9822	0,9372	0,8609	0,7604	0,2503	0,0238	0,0057	0,0000	0,0000	0,0000	0,0000	46	
	4	0,9968	0,9832	0,9510	0,8964	0,4312	0,0643	0,0185	0,0002	0,0000	0,0000	0,0000	55	
	5	0,9995	0,9963	0,9856	0,9622	0,6161	0,1388	0,0480	0,0007	0,0001	0,0000	0,0000	44	
	6	0,9999	0,9993	0,9964	0,9882	0,7702	0,2506	0,1034	0,0025	0,0005	0,0000	0,0000	43	
	7		0,9999	0,9992	0,9968	0,8779	0,3911	0,1904	0,0073	0,0017	0,0000	0,0000	42	
	8			0,9999	0,9992	0,9421	0,5421	0,3073	0,0183	0,0050	0,0002	0,0000	41	
	9				0,9998	0,9755	0,6830	0,4437	0,0402	0,0127	0,0008	0,0000	40	
	10					0,9906	0,7986	0,5836	0,0789	0,0284	0,0022	0,0000	39	
	11					0,9968	0,8827	0,7107	0,1390	0,0570	0,0057	0,0000	38	
	12					0,9990	0,9373	0,8139	0,2229	0,1035	0,0133	0,0002	37	
	13					0,9997	0,9693	0,8894	0,3279	0,1715	0,0280	0,0005	36	
	14					0,9999	0,9862	0,9393	0,4468	0,2612	0,0540	0,0013	35	
	15						0,9943	0,9692	0,5692	0,3690	0,0955	0,0033	34	
	16						0,9978	0,9856	0,6839	0,4868	0,1561	0,0077	33	
	17						0,9992	0,9937	0,7822	0,6046	0,2369	0,0164	32	
50	18						0,9998	0,9975	0,8594	0,7126	0,3356	0,0325	31	50
	19						0,9999	0,9991	0,9152	0,8036	0,4465	0,0595	30	
	20							0,9997	0,9522	0,8741	0,5610	0,1013	29	
	21							0,9999	0,9749	0,9244	0,6701	0,1611	28	
	22								0,9877	0,9576	0,7660	0,2399	27	
	23								0,9944	0,9778	0,8438	0,3359	26	
	24								0,9976	0,9892	0,9022	0,4439	25	
	25								0,9991	0,9951	0,9427	0,5561	24	
	26								0,9997	0,9979	0,9686	0,6641	23	
	27								0,9999	0,9992	0,9840	0,7601	22	
	28									0,9997	0,9924	0,8389	21	
	29									0,9999	0,9960	0,8987	20	
	30										0,9986	0,9405	19	
	31										0,9995	0,9675	18	
	32										0,9998	0,9836	17	
	33										0,9999	0,9923	16	
	34											0,9967	15	
	35											0,9987	14	
	36											0,9995	13	
	37											0,9998	12	
	38	Nicht aufgeführte Werte sind (auf 4 Dezimalstellen) 1,0000											11	
n	k	0,98	0,97	0,96	0,95	0,90	$\frac{5}{6}$	0,80	0,70	$\frac{2}{3}$	0,60	0,50	k	n

4. Hypergeometrische Verteilung/ Urnenmodelle

4.1 Hypergeometrische Verteilung

(Vergleiche Vorlesung: Beispiel (2.3.2)) In einer Urne befinden sich N Kugeln, von denen M schwarz und $N-M$ weiß sind. Es werden n Kugeln ohne Zurücklegen gezogen. Beschreibt die Zufallsvariable X die Anzahl der schwarzen unter den n gezogenen Kugeln, so gilt:

$$\mathbf{P}\{X=k\} = \frac{\binom{M}{k}\cdot\binom{N-M}{M-k}}{\binom{N}{M}} \quad \text{für } k=0,1,\ldots,n \ (M\le N,\ n\le N).$$

Vorlesung: $\mathcal{H}_{n,N,M}(\{k\})$ mit

$$\mathcal{H}_{n,N,M}(\{k\}) = \frac{\binom{M}{k}\cdot\binom{N-M}{M-k}}{\binom{N}{M}}$$

heißt hypergeometrische Verteilung auf $(\Omega=\{0,1,\ldots,n\}, \mathfrak{A}=\mathcal{POT}(\Omega))$.

4.1.1 Aufgabe

Von den 10 Blitzbirnen einer Schachtel sind 2 schon benutzt worden. Mit welcher Wahrscheinlichkeit sind unter 5 zufällig entnommenen Blitzbirnen mindestens 4 unbenutzte, wenn

(a) mit Zurücklegen

(b) ohne Zurücklegen (*Hilfe*: geeignetes Urnenmodell)

gezogen wird?

Zu (a):
Beschreibt X die Anzahl der unbenutzten Blitzbirnen unter den 5 entnommenen, so ist X $\mathcal{B}_{5;0,8}$-verteilt.

$$\begin{aligned}
\mathbf{P}\{X\ge 4\} &= \mathbf{P}\{X=4\}+\mathbf{P}\{X=5\}\\
&= \mathcal{B}_{5;0,8}(4)+\mathcal{B}_{5;0,8}(5)\\
&= \binom{5}{4}\cdot 0,8^4\cdot 0,2^1+\binom{5}{5}\cdot 0,8^5\cdot 0,2^0 \approx 0,73728
\end{aligned}$$

Zu (b):

$$\begin{aligned}
\mathbf{P}\{\text{Mindestens 4 unbenutze}\} &= \mathbf{P}\{\text{Genau 4 unbenutze}\}+\mathbf{P}\{\text{Genau 5 unbenutze}\}\\
&= \frac{\binom{8}{4}\cdot\binom{2}{1}}{\binom{10}{5}}+\frac{\binom{8}{5}\cdot\binom{2}{0}}{\binom{10}{5}} = \frac{70\cdot 2+56\cdot 1}{252} = \frac{7}{9} = 0,\overline{7}
\end{aligned}$$

4.2 Urnenmodelle (I)

Urne mit $n=r+s$ Kugeln; r rote, s schwarze. Es werden n $(\le N)$ Kugeln gezogen.

4.2.1 Ziehen ohne Zurücklegen

Gesucht: Die Wahrscheinlichkeit, daß genau k gezogene Kugeln rot sind.
1. Modell: (ohne Reihenfolge)

$$\Omega = K_n^{\{1,\ldots N\}} = \{A\subseteq\{1,\ldots,N\} : \#A=n\} \qquad \mathfrak{A}=\mathcal{POT}(\Omega) \qquad \mathbf{P}=\mathfrak{L}_\Omega$$

$$\begin{aligned} A_k &= \{\omega \in \Omega \mid \#(\omega \cap \{1,\ldots,r\}) = k\} \\ &= \{\omega \in \Omega \mid \#(\omega \cap \{1,\ldots,r\}) = k, \#(\omega \cap \{r+1,\ldots,N\}) = n-k\} \\ &\cong K_k^{\{1,\ldots r\}} \times K_{n-k}^{\{r+1,\ldots N\}} \end{aligned}$$

$$\Rightarrow \quad A_k = \binom{r}{k} \cdot \binom{N-r}{n-k} \qquad \#\Omega = \binom{N}{n}$$

Die k roten Kugeln können wir aus $\binom{r}{k}$ Möglichkeiten ziehen, die $n-k$ schwarzen Kugeln aus $\binom{s}{n-k} = \binom{N-r}{n-k}$ Möglichkeiten.

$$\Rightarrow \quad \mathbf{P}(A_k) = \frac{\#A_k}{\#\Omega} = \frac{\binom{r}{k} \cdot \binom{N-r}{n-k}}{\binom{N}{n}}.$$

<u>2. Modell</u>: (mit Reihenfolge)

$$\Omega = \mathcal{P}_n^{\{1,\ldots,N\}} = \{\{\omega_1,\ldots,\omega_n\} : \omega_i \in \{1,\ldots,N\}, \omega_i \neq \omega_j \, \forall i \neq j\}$$

$$\mathfrak{A} = \mathcal{POT}(\Omega) \qquad \mathbf{P} = \mathfrak{L}_\Omega$$

$$A_k = \{\omega \in \Omega \mid \#\{i : \omega_i \in \{1,\ldots,r\}\} = k\}$$

Die roten Kugel können wir aus $\binom{n}{k}$ Möglichkeiten ziehen, die k roten aus $(r)_k$ Möglichkeiten und die $n-k$ schwarzen aus $(N-r)_{n-k}$ Möglichkeiten.

$$\Rightarrow \quad \#A_k = \binom{n}{k} \cdot (r)_k \cdot (N-r)_{n-k} \qquad \#\Omega = (N)_n = \binom{N}{n} \cdot n!$$

$$\begin{aligned} \Rightarrow \quad \mathbf{P}(A_k) &= \frac{\binom{n}{k} \cdot (r)_k \cdot (N-r)_{n-k}}{\binom{N}{n} \cdot n!} \\ &= \frac{\binom{n}{k} \cdot k! \cdot \binom{r}{k} \cdot (n-k)! \cdot \binom{N-r}{n-k}}{\binom{N}{n} \cdot n!} \\ &= \frac{\frac{n!}{k!\cdot(n-k)!} \cdot k! \cdot \frac{r!}{k!\cdot(r-k)!} \cdot (n-k)! \cdot \frac{(N-r)!}{(n-k)!\cdot((N-r)-(n-k))!}}{\frac{N!}{n!\cdot(N-n)!} \cdot n!} \\ &= \frac{n! \cdot r! \cdot (N-r)! \cdot (N-n)!}{N! \cdot k! \cdot (n-k)! \cdot (r-k)! \cdot ((N-r)-(n-k))!} \end{aligned}$$

Natürlich ist dieses Ergebnis gleich dem Ergebnis im ersten Modell. Es gilt nämlich:

$$\begin{aligned} \frac{\binom{r}{k} \cdot \binom{N-r}{n-k}}{\binom{N}{n}} &= \frac{\frac{r!}{k!\cdot(r-k)!} \cdot \frac{(N-r)!}{(n-k)!\cdot((N-r)-(n-k))!}}{\frac{N!}{n!\cdot(N-n)!}} \\ &= \frac{r! \cdot (N-r)! \cdot n! \cdot (N-n)!}{k! \cdot (r-k)! \cdot (n-k)! \cdot ((N-r)-(n-k))! \cdot N!}. \end{aligned}$$

Dies ist auch nicht verwunderlich, da die „tatsächliche“ Wahrscheinlichkeit nicht davon abhängt, ob wir in unserem Modell die Kugeln durchnummerieren und die Reihenfolge des Auftretens notieren oder nicht. Welches mathematische Modell man jeweils wählt, ist Geschmackssache und ändert nichts am Ergebnis!

4.2.2 Ziehen mit Zurücklegen

Gesucht: Die Wahrscheinlichkeit, daß genau k Kugeln rot sind.
Modell: (mit Reihenfolge)

$$\Omega = \{1, \ldots, r, r+1, \ldots, N\}^n \qquad \mathfrak{A} = \mathcal{POT}(\Omega) \qquad \mathbf{P} = \mathfrak{L}_\Omega$$

$$A_k = \{\omega \in \Omega \mid \# \{i : \omega_i \in \{1, \ldots, r\}\} = k\}$$

In den Kugeln sind die k roten auf $\binom{n}{k}$ Möglichkeiten angeordnet, für die k roten Kugeln gibt es r^k Möglichkeiten gezogen zu werden und für die $(n-k)$ schwarzen $(N-r)^{n-k}$ Möglichkeiten.

$$\Rightarrow \quad \# A_k = \binom{n}{k} \cdot r^k \cdot (N-r)^{n-k} \qquad \#\Omega = N^n$$

$$\begin{aligned} \Rightarrow \quad \mathbf{P}(A_k) &= \frac{\binom{n}{k} \cdot r^k \cdot (N-r)^{n-k}}{N^n} \\ &= \binom{n}{k} \cdot \left(\frac{r}{N}\right)^k \cdot \left(\frac{N-r}{N}\right)^{n-k} \\ &= \binom{n}{k} \cdot \left(\frac{r}{N}\right)^k \cdot \left(1 - \frac{r}{N}\right)^{n-k} \\ &= \mathcal{B}_{n, \frac{r}{N}}\{k\} \quad \text{Binomialverteilung.} \end{aligned}$$

4.3 Urnenmodelle (II)

Urne mit r roten und $s = N - r$ schwarzen Kugeln. Es wird n-mal gezogen und es sei

$$B_k := \text{„beim } k\text{-ten Zug rot“.}$$

4.3.1 ohne Zurücklegen

$$\Omega = \mathcal{P}_n^{\{1, \ldots N\}} = \{(\omega_1, \ldots, \omega_n) : \omega_i \in \{1, \ldots, N\}, \, \omega_i \neq \omega_j \; \forall i \neq j\}$$

$$\mathfrak{A} = \mathcal{POT}(\Omega) \qquad \mathbf{P} = \mathfrak{L}_\Omega \qquad \#\Omega = (N)_n$$

$$\begin{aligned} B_k &= \{\omega \in \Omega \mid \omega_k \in \{1, \ldots, r\}\} \\ &\cong \mathcal{P}_{n-1}^{\{1, \ldots, N-1\}} \times \{1, \ldots, r\} \\ \Rightarrow \quad \# B_k &= (N-1)_{n-1} \cdot r \end{aligned}$$

$$\begin{aligned} \Rightarrow \quad \mathbf{P}(B_k) &= \frac{(N-1)_{n-1} \cdot r}{(N)_n} = \frac{\binom{N-1}{n-1} \cdot (n-1)! \cdot r}{\binom{N}{n} \cdot n!} \\ &= \frac{(N-1)! \cdot (n-1)! \cdot n! \cdot (N-n)! \cdot r}{(n-1)! \cdot ((N-1) - (n-1))! \cdot N! \cdot n!} \\ &= \frac{(N-1)! \cdot (N-n)! \cdot r}{(N-n)! \cdot N!} = \frac{r}{N} = \frac{r}{r+s}, \end{aligned}$$

unabhängig von k, n $(k \leq n)$.

4.3.2 mit Zurücklegen

$$\Omega = \{1, \ldots, N\}^n \qquad \mathfrak{A} = \mathcal{POT}(\Omega) \qquad \mathbf{P} = \mathfrak{L}_\Omega$$

$$\begin{aligned} B_k &= \{\omega \in \Omega \mid \omega_k \in \{1, \ldots, r\}\} \\ &\cong \{1, \ldots, N\}^n \times \{1, \ldots, r\} \end{aligned}$$

$$\Rightarrow \quad \mathbf{P}(B_k) = \frac{N^{n-1} \cdot r}{N^n} = \frac{r}{N} = \frac{r}{r+s},$$

unabhängig von k, n $(k \leq n)$ und gleich dem Ergebnis in $(A.4.3.1)$!

5. Kombinatorische Probleme

5.1 Aufgaben

5.1.1 Aufgabe

Ein Kurs besteht aus 7 Mädchen und 13 Jungen. Zur Vorbereitung der Studienfahrt wird ein Dreierausschuß ausgelost. Mit welcher Wahrscheinlichkeit

(a) wird die Kurssprecherin in den Ausschuß gelost;

(b) besteht der Ausschuß nur aus Jungen;

(c) enthält der Ausschuß höchstens einen Jungen;

(d) enthält der Ausschuß einen Jungen und ein Mädchen, wenn vorher festgelegt wurde, daß die Kurssprecherin auf jeden Fall dem Ausschuß angehören muß?

Zu (a):

$$\Omega = \big\{A \subseteq \{\underbrace{1,\ldots,7}_{\text{Mädchen}},\underbrace{8,\ldots,20}_{\text{Jungen}}\} \mid \#A = 3\big\} \qquad \mathfrak{A} = \mathcal{POT}(\Omega) \qquad \mathbf{P} = \mathfrak{L}_\Omega$$

Es gibt

$$\#\Omega = \binom{20}{3} = 1140$$

mögliche Ausschußzusammensetzungen.
Ohne Einschränkung sei die Nr. 1 die Kurssprecherin:

$$\begin{aligned} B_1 &= \{\text{Kurssprecherin wird in den Ausschuß gelost}\} \\ &= \{\omega \in \Omega \mid 1 \in \omega\}, \end{aligned}$$

also folgt

$$\#B_1 = \binom{19}{2} = 171.$$

$$\Rightarrow \quad \mathbf{P}(B_1) = \frac{\binom{19}{2}}{\binom{20}{3}} = \frac{171}{1140} = \frac{3}{20} = 0,15$$

Zu (b):
Ω wie zuvor, $\mathbf{P} = \mathfrak{L}_\Omega$.

$$\begin{aligned} B_2 &= \{\text{Ausschuß besteht nur aus Jungen}\} \\ &= \{\omega \in \Omega \mid \#(\omega \cap \{8,\ldots,20\}) = 3\} \\ &= \{\omega \in \Omega \mid \omega \subset \{8,\ldots,20\}\} \\ \#B_2 &= \binom{13}{3} = 286 \end{aligned}$$

$$\Rightarrow \quad \mathbf{P}(B_2) = \frac{\binom{13}{3}}{\binom{20}{3}} = \frac{286}{1140} \approx 0,25$$

Zu (c):
Ω, $\mathbf{P}$ wie oben.

$$\begin{aligned}
B_3 &= \{\text{Ausschuß enthält höchstens einen Jungen}\} \\
&= \{\omega \in \Omega \mid \#(\omega \cap \{8,\dots,20\}) \leq 1\} \\
&= \{\omega \in \Omega \mid \#(\omega \cap \{8,\dots,20\}) = 1\} \mathbin{\dot\cup} \{\omega \in \Omega \mid \#(\omega \cap \{8,\dots,20\}) = 0\} \\
&= \{\omega \in \Omega \mid \#(\omega \cap \{8,\dots,20\}) = 1 \text{ und } \#(\omega \cap \{1,\dots,7\}) = 2\} \\
&\qquad \mathbin{\dot\cup} \{\omega \in \Omega \mid \#(\omega \cap \{8,\dots,20\}) = 0 \text{ und } \#(\omega \cap \{1,\dots,7\}) = 3\} \\
&= \{\omega \in \Omega \mid \#(\omega \cap \{8,\dots,20\}) = 1 \text{ und } \#(\omega \cap \{1,\dots,7\}) = 2\} \\
&\qquad \mathbin{\dot\cup} \{\omega \in \Omega \mid \omega \subset \{1,\dots,7\}\}
\end{aligned}$$

$$\Rightarrow \quad \#B_3 = \binom{13}{1} \cdot \binom{7}{2} + \binom{7}{3} = 13 \cdot 21 + 35 = 308$$

$$\Rightarrow \quad \mathbf{P}(B_3) = \frac{308}{1140} \approx 0,27$$

Zu (d):

$$\Omega = \{A \subset \{2,\dots,20\} \mid \#A = 2\} \qquad \mathfrak{A} = \mathcal{POT}(\Omega) \qquad \mathbf{P} = \mathfrak{L}_\Omega \qquad \#\Omega = \binom{19}{2} = 171$$

$$\begin{aligned}
B_4 &= \{\omega \in \Omega \mid \#(\omega \cap \{2,\dots,7\}) = 1 \text{ und } \#(\omega \cap \{8,\dots,20\}) = 1\} \\
\#B_4 &= \binom{6}{1} \cdot \binom{13}{1} = 6 \cdot 13 = 78
\end{aligned}$$

$$\Rightarrow \quad \mathbf{P}(B_4) = \frac{78}{171} \approx 0,46$$

5.1.2 Aufgabe

Wie viele unterschiedliche Anordnungen der Buchstaben des Wortes $MISSISSIPPI$ gibt es?

— — — — — — — — — — —

11 Stellen sind mit Buchstaben zu besetzen.
Die 4 „S“ könnenn auf $\binom{11}{4}$ Arten untergebracht werden, die 4 „I“ dann auf $\binom{7}{4}$ Arten, die 2 „P“ auf $\binom{3}{2}$ und das „M“ liegt dann fest $\binom{1}{1}$.
$\Rightarrow$ Insgesamt

$$\binom{11}{4} \cdot \binom{7}{4} \cdot \binom{3}{2} \cdot 1 = 330 \cdot 35 \cdot 3 = 34690$$

unterschiedliche Anordnungen.

5.1.3 Aufgabe

In einem Aufzug, der noch 6 Stockwerke fährt, sind 4 Personen, die voneinander unabhängig aussteigen.
Wie groß ist die Wahrscheinlichkeit, daß

(a) alle in verschiedenen Stockwerken,

(b) genau zwei in einem Stockwerk,

(c) alle 4 im gleichen Stockwerk,

(d) mindestens 3 im gleichen Stockwerk

aussteigen?

$$\Omega = \{(\omega_1, \omega_2, \omega_3, \omega_4) : \omega_i \in \{1, \ldots, 6\}\,,\, i = 1, \ldots, 4\} = \{1, \ldots, 6\}^4 \qquad \mathfrak{A} = \mathcal{POT}(\Omega)$$
$$\mathbf{P} = \mathfrak{L}_\Omega \quad \text{Gleichverteilung}$$

ω_i gibt das Stockwerk an, in dem Person i austeigt.

$$\#\Omega = 6^4 = 1296$$

Zu (a):

$$\begin{aligned}
B_1 &= \{\text{„alle steigen in verschieden Stockwerken aus“}\} \\
&= \{(\omega_1, \omega_2, \omega_3, \omega_4) \in \Omega \mid \omega_i \neq \omega_j \;\forall i \neq j\} \\
\#B_1 &= 6 \cdot 5 \cdot 4 \cdot 3 = 360
\end{aligned}$$

$$\Rightarrow \quad \mathbf{P}(B_1) = \frac{\#B_1}{\#\Omega} = \frac{360}{1296} = \frac{5}{18} = 0,2\overline{7} \approx 0,28$$

Zu (b):

$$\begin{aligned}
B_2 &= \{\text{„genau zwei steigen in einem Stockwerk aus“}\} \\
&= \{(\omega_1, \omega_2, \omega_3, \omega_4) \in \Omega \mid \#\{\omega_1, \omega_2, \omega_3, \omega_4\} = 3\} \\
\#B_2 &= \binom{4}{2} \cdot 6 \cdot 5 \cdot 4 \;=\; 6 \cdot 6 \cdot 5 \cdot 4 \;=\; 720
\end{aligned}$$

Die zwei Personen sind beliebig aus den 4 gezogen, d.h. $\binom{4}{2}$, sie haben 6 Ausstigsmöglichkeiten, für die dritte Person bleiben dann noch 5 Ausstigsmöglichkeiten und für die vierte noch 4.

$$\Rightarrow \quad \mathbf{P}(B_2) = \frac{\#B_2}{\#\Omega} = \frac{720}{1296} = 0,\overline{5} \approx 0,56$$

Zu (c):

$$\begin{aligned}
B_3 &= \{\text{„Alle 4 steigen im gleichen Stockwerk aus“}\} \\
&= \{(\omega_1, \omega_2, \omega_3, \omega_4) \in \Omega \mid \omega_1 = \omega_2 = \omega_3 = \omega_4\} \\
\#B_3 &= 6
\end{aligned}$$

$$\Rightarrow \quad \mathbf{P}(B_3) = \frac{\#B_3}{\#\Omega} = \frac{6}{1296} \approx 4,6 \cdot 10^{-3}$$

Zu (d):

$$\begin{aligned}
B_4 &= \{\text{„Mindestens 3 steigen im gleichen Stockwerk aus“}\} \\
&= \{\text{„Alle 4 steigen im gleichen Stockwerk aus“}\} \\
&\qquad \uplus \{\text{„Genau 3 steigen im gleichen Stockwerk aus“}\} \\
&= \{(\omega_1, \omega_2, \omega_3, \omega_4) \in \Omega \mid \omega_1 = \omega_2 = \omega_3 = \omega_4\} \\
&\qquad \uplus \{(\omega_1, \omega_2, \omega_3, \omega_4) \in \Omega \mid \#\{\omega_1, \omega_2, \omega_3, \omega_4\} = 2\} \\
\#B_4 &= 6 + 6 \cdot \binom{4}{3} \cdot 5 \;=\; 6 + 6 \cdot 4 \cdot 5 \;=\; 126
\end{aligned}$$

$$\Rightarrow \quad \mathbf{P}(B_4) = \frac{\#B_4}{\#\Omega} = \frac{126}{1296} \approx 0,097$$

5.1.4 Aufgabe

(1) Wie viele „Worte“ (auch sinnlose) aus 5 Buchstaben lassen sich aus den 3 Konsonanten b, d, n und den 2 Vokalen a, e bilden?

(2) Wie viele verschiedene „Worte“ aus 5 Buchstaben lassen sich aus 3 verschiedenen Konsonanten (der 21 Konsonanten) und 2 verschiedenen Vokalen (der 5 Vokale) bilden?

Zu (a):

$$\begin{aligned} \Omega &= \{(\omega_1, \omega_2, \omega_3, \omega_4, \omega_5) \mid \omega_i \in \{b, d, n, a, e\}, \, \omega_i \neq \omega_j \; \forall i \neq j\} \\ \#\Omega &= 5 \cdot 4 \cdot 3 \cdot 2 \cdot 1 = 5! = 120 \end{aligned}$$

Zu (b):
Auswählen der 3 verschiedenen Konsonanten: $\binom{21}{3}$ Möglichkeiten.
Auswählen der 2 verschiedenen Vokale: $\binom{5}{2}$ Möglichkeiten.
Mit diesen 5 gewählten Buchstaben kann man 5! (auch sinnlose) Worte bilden.

$$\Rightarrow \quad \binom{21}{3} \cdot \binom{5}{2} \cdot 5! = 1330 \cdot 10 \cdot 120 = 1.596.000 \approx 1,6 \text{ Millionen}$$

verschiedene „Worte“ können gebildet werden.

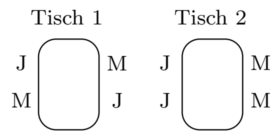
5.1.5 Aufgabe

Die Volleyballmanschaften der Mädchen und der Jungen waren so erfolgreich, daß sie zum Sportlerball der Stadt eingeladen werden. 6 Mädchen und 10 Jungen folgen der Einladung und finden ihre Namen auf Tischkarten an 4 Vierertischen. Ihre 16 Tischkarten wurden aus einem Korb gezogen, erst 4 für den 1. Tisch, dann 4 für den 2. Tisch usw.
Mit welcher Wahrscheinlichkeit sitzen

(a) gleich viel Mädchen wie Jungen an den ersten beiden Tischen,

(b) nur Mädchen am ersten und nur Jungen am zweiten Tisch?

Zu (a):
An jedem der beiden Tische sitzen 2 Jungen und 2 Mädchen.



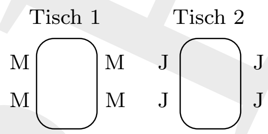
$$\{\underbrace{1, \ldots, 6}_{\text{Mädchen}}, \underbrace{7, \ldots, 16}_{\text{Jungen}}\}$$

$$\begin{aligned} \Omega &= \{(T_1, T_2) : T_i \subset \{1, \ldots, 16\}, \, \#T_i = 4, \, T_1 \cap T_2 = \emptyset\} \\ &\quad \text{z.B.} \;\; (\{4, 7, 8, 12\}, \{2, 9, 10, 16\}) \\ \#\Omega &= \binom{16}{4} \cdot \binom{12}{4} = 1820 \cdot 499 = 900900 \\ A &= \{(T_1, T_2) \in \Omega \mid \# (T_{1,2} \cap \{1, \ldots, 6\}) = 2, \, \# (T_{1,2} \cap \{7, \ldots, 16\}) = 2\} \\ \#A &= \binom{6}{2} \cdot \binom{10}{2} \cdot \binom{4}{2} \cdot \binom{8}{2} = 15 \cdot 45 \cdot 6 \cdot 28 = 113400 \end{aligned}$$

$$\Rightarrow \quad \mathbf{P}(A) = \frac{\#A}{\#\Omega} = \frac{113400}{900900} \approx 0,13$$

Zu (b):

Nur Mädchen sitzen am ersten Tisch und nur Jungen sitzen am zweiten.
Ω wie oben.



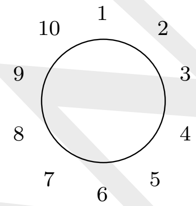
$$\begin{aligned} B &= \{(T_1, T_2) \in \Omega \mid \#(T_1 \cap \{1, \ldots, 6\}) = 4, \#(T_2 \cap \{7, \ldots, 16\}) = 4\} \\ \#B &= \binom{6}{4} \cdot \binom{10}{4} = 19 \cdot 210 = 3150 \end{aligned}$$

$$\Rightarrow \quad \mathbf{P}(B) = \frac{\#B}{\#\Omega} = \frac{3150}{900900} \approx 3,5 \cdot 10^{-3}$$

5.1.6 Aufgabe

Auf wie viele Arten können 5 Männer und 5 Frauen an einem runden Tisch sitzen, wenn die Tischnachbarn verschiedenen Geschlechts sein sollen?

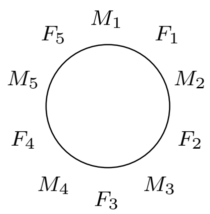
(1) Es interessiert, wer auf welchem Stuhl am Tisch sitzt. Die Stühle seien von 1 bis 10 durchnummeriert.



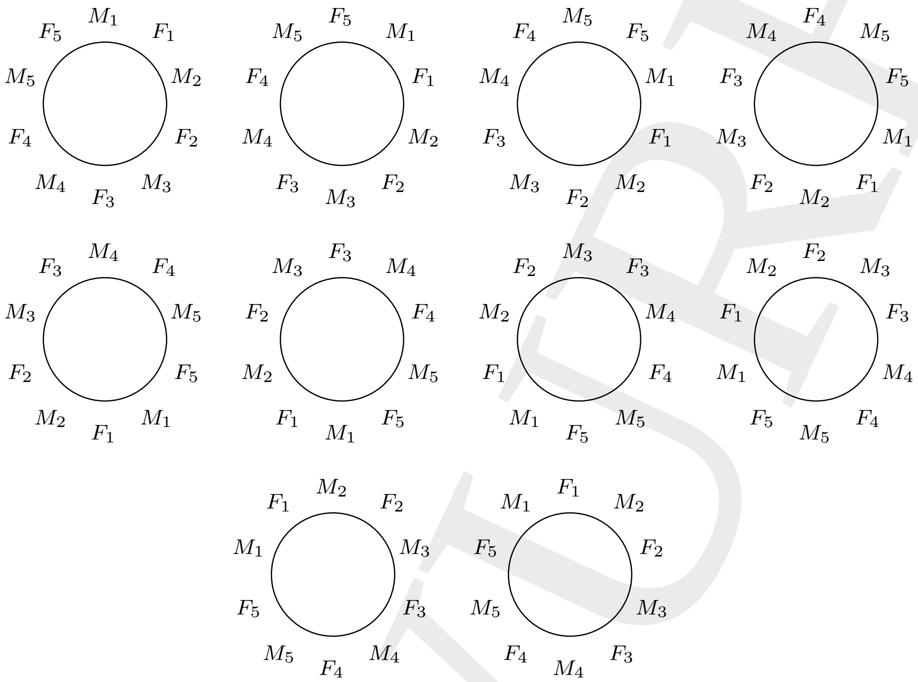
- Falls auf Stuhl 1 eine Frau sitzt, so liegen die Stühle der 4 übrigen Frauen fest: Stuhl 3, Stuhl 5, Stuhl 7, Stuhl 9. Es gibt 5! Möglichkeiten, die Frauen auf die Stühle $1, 3, 5, 7, 9$ zu verteilen.
  Es gibt ebenfalls 5! Möglichkeiten die 5 Männer auf die Stühle $2, 4, 6, 8, 10$ zu verteilen.
- Falls auf Stuhl 1 ein Mann sitzt, so erhält man mit den selben Überlegungen $5! \cdot 5!$ Möglichkeiten, die Frauen und Männer auf die Stühle zu verteilen.

$\Rightarrow$ Insgesamt $2 \cdot 5! \cdot 5! = 2880$ Möglichkeiten.

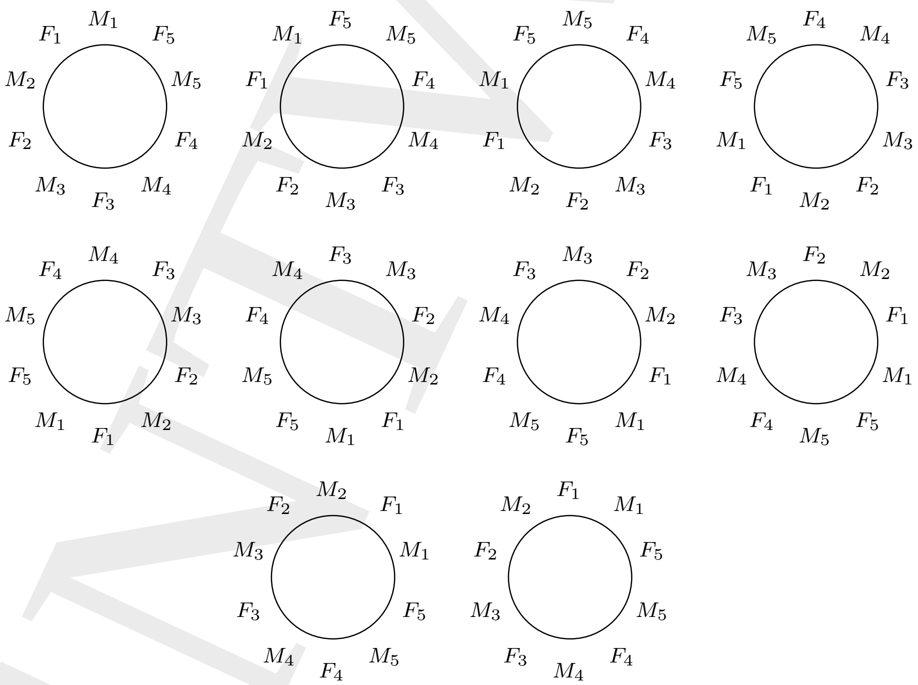
(2) Es interessiert lediglich die Anordnung der 10 Persohnen, d.h. wer neben wem sitzt. Zwischen rechtem und linkem Nachbarn wird nicht unterschieden.
Man unterscheidet zunächst die 10 Sitzplätze, wie in Teil (1).
Betrachte eine feste Annordnung:



<u>Frage</u>: Welche Anordnungen aus Teil (1) (d.h. mit Unterscheidung der Sitzplätze) führen auf diese Anordnung, bei der nun lediglich die Nachbarn interessieren?
Durch „Weiterrücken“ jeder Person um einen Platz:



Durch „Spiegelung" an einem (festen) Durchmesser des Tisches (dadurch werden rechter und linker Nachbar vertauscht).



Spiegelungen an anderen Durchmessern liefern keine neuen Anordnungen.
Also führen 20 verschiedene Anordnungen mit Unterscheidung der Sitzplätze auf dieselbe Anordnung, bei der lediglich die Nachbarn interessieren.

$$\Rightarrow \quad \frac{2 \cdot 5! \cdot 5!}{20} = 1440 \text{ Anordnungen.}$$

6. Gauß'sche Normalverteilung

6.1 Wahrscheinlichkeitsmaße mit Dichten

Es sei $f : \mathbb{R}^d \to \mathbb{R}_+ \cup \{\infty\}$ eine meßbare Funktion mit

$$\int f \, \mathbf{d}\lambda^d = 1,$$

dann wird durch

$$\mathbf{P}(A) = \int 1_A \cdot f \, \mathbf{d}\lambda^d \quad A \in \mathfrak{B}\left(\mathbb{R}^d\right)$$

ein Wahrscheinlichkeitsmaß auf $\mathbb{R}^d$ definiert. f heißt Lebesgue-Dichte von $\mathbf{P}$.

6.2 Normalverteilung

Gauß'sche Normalverteilung mit Parametern $\mu \in \mathbb{R}$ und $\sigma^2 > 0$; Wahrscheinlichkeitsmaß mit Lebesgue-Dichte

$$f(x) = \frac{1}{\sqrt{2\pi\sigma^2}} \cdot \mathbf{e}^{-\frac{(x-\mu)^2}{2\sigma^2}}$$

auf $(\mathbb{R}, \mathfrak{B}(\mathbb{R}))$ $\mathcal{N}_{\mu,\sigma^2}$.
Für $\mu = 0$, $\sigma^2 = 1$ erhalten wir die Standard-Normalverteilung $\mathcal{N}_{0,1}$, ein Wahrscheinlichkeitsmaß auf $\mathbb{R}$.

$$\begin{aligned}
\mathcal{N}_{0,1}([a;b]) &= \frac{1}{\sqrt{2\pi}} \cdot \int_a^b \mathbf{e}^{-\frac{x^2}{2}} \, \mathbf{d}x \\
\mathcal{N}_{0,1}(]-\infty;t]) &= \frac{1}{\sqrt{2\pi}} \cdot \int_{-\infty}^t \mathbf{e}^{-\frac{x^2}{2}} \, \mathbf{d}t =: \Phi(t)
\end{aligned}$$

Problem: Die Stammfunktion ist nicht in expliziter Form darstellbar, deshalb liegt die Funktion Φ tabelliert vor.

6.2.1 Eigenschaften

(i) $\Phi(t) = 1 - \Phi(-t)$,

(ii) $\mathcal{N}_{\mu,\sigma^2}([a;b]) = \Phi\left(\frac{b-\mu}{\sigma}\right) - \Phi\left(\frac{a-\mu}{\sigma}\right)$

Es reicht daher aus, die Funktion Φ für Werte $t \geq 0$ zu tabellieren.

$$\mathcal{N}_{\mu,\sigma^2}(]-\infty;t]) = \Phi\left(\frac{t-\mu}{\sigma}\right)$$

Abschätzungen:

$$\begin{aligned}
\mathcal{N}_{\mu,\sigma^2}([\mu-\sigma;\mu+\sigma]) &= \Phi\left(\frac{\mu+\sigma-\mu}{\sigma}\right) - \Phi\left(\frac{\mu-\sigma-\mu}{\sigma}\right) = \Phi(1) - \Phi(-1) \\
&= \Phi(1) - (1 - \Phi(1)) = 2 \cdot \Phi(1) - 1 = 2 \cdot 0,8413 - 1 = 0,6826
\end{aligned}$$

$$\begin{aligned}
\mathcal{N}_{\mu,\sigma^2}([\mu-2\sigma;\mu+2\sigma]) &= \Phi\left(\frac{\mu+2\sigma-\mu}{\sigma}\right) - \Phi\left(\frac{\mu-2\sigma-\mu}{\sigma}\right) = \Phi(2) - \Phi(-2) \\
&= \Phi(2) - (1 - \Phi(2)) = 2 \cdot \Phi(2) - 1 = 2 \cdot 0,9772 - 1 = 0,9544
\end{aligned}$$

$$\begin{aligned}\mathcal{N}_{\mu,\sigma^2}\left([\mu-3\sigma;\mu+3\sigma]\right) &= \Phi\left(\frac{\mu+3\sigma-\mu}{\sigma}\right)-\Phi\left(\frac{\mu-3\sigma-\mu}{\sigma}\right) = \Phi(3)-\Phi(-3)\\ &= \Phi(3)-(1-\Phi(3)) = 2\cdot\Phi(3)-1 = 2\cdot 0,9987-1 = 0,9974\end{aligned}$$

Bei der Normalverteilung liegen die beobachteten Werte (wenn deren Anzahl hinreichend groß ist) für

rund 68% im Intervall $[\mu-\sigma;\mu+\sigma]$,

rund 95% im Intervall $[\mu-2\sigma;\mu+2\sigma]$,

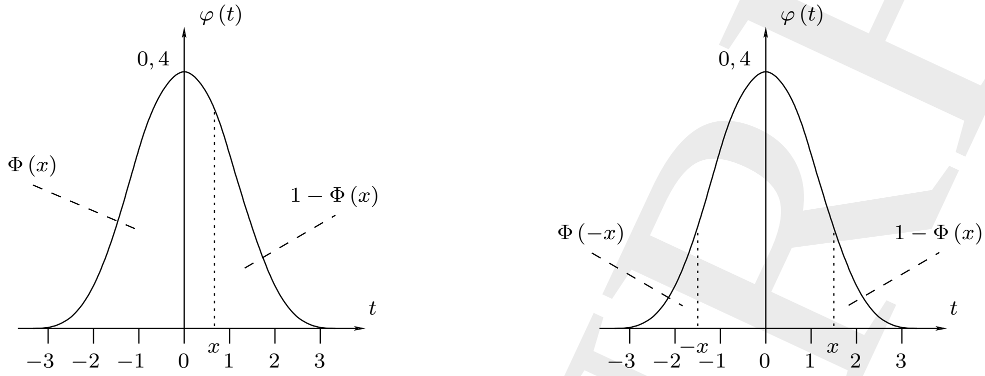
rund 99% im Intervall $[\mu-3\sigma;\mu+3\sigma]$.

Es ist also praktisch sicher, daß nur Werte aus dem sogenannten 3σ-Intervall auftreten. Diesen Sachverhalt bezeichnet man als 3σ-Regel.

6.2.2 Tabelle

Gauß'sche Summenfunktion Φ

x	0	1	2	3	4	5	6	7	8	9
0,0	0,5000	0,5040	0,5080	0,5120	0,5160	0,5199	0,5239	0,5279	0,5319	0,5359
0,1	0,5398	0,5438	0,5478	0,5517	0,5557	0,5596	0,5636	0,5675	0,5714	0,5753
0,2	0,5793	0,5832	0,5871	0,5910	0,5948	0,5987	0,6026	0,6064	0,6103	0,6141
0,3	0,6179	0,6217	0,6255	0,6293	0,6331	0,6368	0,6406	0,6443	0,6480	0,6517
0,4	0,6554	0,6591	0,6628	0,6664	0,6700	0,6736	0,6772	0,6808	0,6844	0,6879
0,5	0,6915	0,6950	0,6985	0,7019	0,7054	0,7088	0,7123	0,7157	0,7190	0,7224
0,6	0,7257	0,7291	0,7324	0,7357	0,7389	0,7422	0,7454	0,7486	0,7517	0,7549
0,7	0,7580	0,7611	0,7642	0,7673	0,7703	0,7734	0,7764	0,7794	0,7823	0,7852
0,8	0,7881	0,7910	0,7939	0,7967	0,7995	0,8023	0,8051	0,8078	0,8106	0,8133
0,9	0,8159	0,8186	0,8212	0,8238	0,8264	0,8289	0,8315	0,8340	0,8365	0,8389
1,0	0,8413	0,8438	0,8461	0,8485	0,8508	0,8531	0,8554	0,8577	0,8599	0,8621
1,1	0,8643	0,8665	0,8686	0,8708	0,8729	0,8749	0,8770	0,8790	0,8810	0,8830
1,2	0,8849	0,8869	0,8888	0,8907	0,8925	0,8944	0,8962	0,8980	0,8997	0,9015
1,3	0,9032	0,9049	0,9066	0,9082	0,9099	0,9115	0,9131	0,9147	0,9164	0,9177
1,4	0,9192	0,9207	0,9222	0,9236	0,9251	0,9265	0,9279	0,9292	0,9306	0,9319
1,5	0,9332	0,9345	0,9357	0,9370	0,9382	0,9394	0,9406	0,9418	0,9429	0,9441
1,6	0,9452	0,9463	0,9474	0,9484	0,9495	0,9505	0,9515	0,9525	0,9535	0,9545
1,7	0,9554	0,9564	0,9573	0,9582	0,9591	0,9599	0,9608	0,9616	0,9625	0,9633
1,8	0,9641	0,9649	0,9656	0,9664	0,9671	0,9678	0,9686	0,9693	0,9699	0,9706
1,9	0,9713	0,9719	0,9726	0,9732	0,9738	0,9744	0,9750	0,9756	0,9761	0,9767
2,0	0,9772	0,9778	0,9783	0,9788	0,9793	0,9798	0,9803	0,9808	0,9812	0,9817
2,1	0,9821	0,9826	0,9830	0,9834	0,9838	0,9842	0,9846	0,9850	0,9854	0,9857
2,2	0,9861	0,9864	0,9868	0,9871	0,9875	0,9878	0,9881	0,9884	0,9887	0,9890
2,3	0,9893	0,9896	0,9898	0,9901	0,9904	0,9906	0,9909	0,9911	0,9913	0,9916
2,4	0,9918	0,9920	0,9922	0,9925	0,9927	0,9929	0,9931	0,9932	0,9934	0,9936
2,5	0,9938	0,9940	0,9941	0,9943	0,9945	0,9946	0,9948	0,9949	0,9951	0,9952
2,6	0,9953	0,9955	0,9956	0,9957	0,9959	0,9960	0,9961	0,9962	0,9963	0,9964
2,7	0,9965	0,9966	0,9967	0,9968	0,9969	0,9970	0,9971	0,9972	0,9973	0,9974
2,8	0,9974	0,9975	0,9976	0,9977	0,9977	0,9978	0,9979	0,9979	0,9980	0,9981
2,9	0,9981	0,9982	0,9982	0,9983	0,9984	0,9984	0,9985	0,9985	0,9986	0,9986
3,0	0,9987	0,9987	0,9987	0,9988	0,9988	0,9989	0,9989	0,9989	0,9990	0,9990
3,1	0,9990	0,9991	0,9991	0,9991	0,9992	0,9992	0,9992	0,9992	0,9993	0,9993
3,2	0,9993	0,9993	0,9994	0,9994	0,9994	0,9994	0,9994	0,9995	0,9995	0,9995
3,3	0,9995	0,9995	0,9996	0,9996	0,9996	0,9996	0,9996	0,9996	0,9996	0,9997
3,4	0,9997	0,9997	0,9997	0,9997	0,9997	0,9997	0,9997	0,9997	0,9997	0,9998



6.2.3 Einige besondere Werte

x	1,2816	1,6449	1,9600	2,3263	2,5785	3,0902	3,2905
$\Phi(x)$	0,9	0,95	0,975	0,99	0,995	0,999	0,9995

6.2.4 Beispiele

$\mu = 3$, $\sigma^2 = 4$

$$\mathcal{N}_{\mu,\sigma^2}(]-\infty;4]) = \Phi\left(\frac{4-3}{2}\right) = \Phi(0,5) = 0,6915$$

$$\begin{aligned}
\mathcal{N}_{\mu,\sigma^2}([2,5;4]) &= \Phi\left(\frac{4-3}{2}\right) - \Phi\left(\frac{2,5-3}{4}\right) = \Phi(0,5) - \Phi(-0,25) \\
&= \Phi(0,5) - (1-\Phi(0,25)) \\
&= 0,6915 - 1 + 0,5987 = 0,2902
\end{aligned}$$

6.2.5 Aufgabe

Die Körpergröße (in *cm*) von Kindern eines Jahrgangs sei annähernd normalverteilt mit $\mu = 90$ und $\sigma = 8$.

(a) Wieviel Prozent dieser Kinder sind höchstens 87 *cm* groß?

(b) Wieviel Prozent dieser Kinder sind mindestens 86 *cm* und höchstens 95 *cm* groß?

Zu (a):

$$\begin{aligned}
\mathcal{N}_{90,64}(]-\infty,87]) &= \Phi\left(\frac{87-90}{8}\right) = \Phi\left(-\frac{3}{8}\right) = 1-\Phi\left(\frac{3}{8}\right) \\
&= 1-\Phi(0,375) \approx 1-\Phi(0,38) = 1-0,6480 \\
&= 0,3520
\end{aligned}$$

Zu (b):

$$\begin{aligned}
\mathcal{N}_{90,64}([86,95]) &= \Phi\left(\frac{95-90}{8}\right) - \Phi\left(\frac{86-90}{8}\right) = \Phi\left(\frac{5}{8}\right) - \Phi\left(-\frac{1}{2}\right) \\
&= \Phi(0,625) - (1-\Phi(0,5)) \approx \Phi(0,63) - (1-\Phi(0,5)) \\
&= 0,7357 - 1 + 0,6915 = 0,4272
\end{aligned}$$

7. Testprobleme

7.0.1 Beispiel

Um zu untersuchen, ob ein neu erschienenes Präparat den Blutdruck senkt, wurde folgendes Experiment gemacht:

13 Versuchspersonen wurde vor und 2 Stunden nach Einnahme des Präparats der Blutduck gemessen.

Ergebnisse:

Nummer	1	2	3	4	5	6	7	8	9	10	11	12	13
vor	129	122	124	105	110	112	109	115	102	95	106	123	99
nach	125	120	121	109	110	102	92	95	97	90	111	115	85
Vorzeichen	+	+	+	−	0	+	+	+	+	+	−	+	+

7.0.2 Mathematische Behandlung

Wir streichen in der weiteren Betrachtung das 5. Meßergebnis, da dort nichts passiert ist.

7.1 Urnenmodell

In einer Urne liegen weiße (für plus) und schwarze (für minus) Kugeln. Das Mischungsverhältnis ist unbekannt. Es seien aber sehr viele Kugeln in der Urne. Wir ziehen aus dieser Urne (mit Zurücklegen) eine „kleine“ Anzahl von Kugeln.

Über das Mischungsverhältnis der Kugeln können wir nur eine Vermutung oder Hypothese aufstellen.

Wir vermuten, daß das Präparat den Blutdruck nicht senkt. Dann sollten in der Urne beide Sorten der Kugeln gleich oft vorkommen. Also

$$\mathbf{P}(\text{„weiß“}) = \mathbf{P}(\text{„schwarz“}) = \frac{1}{2}.$$

Das obige Experiment entspricht 12-maligem Ziehen mit Zurücklegen. Es wurden 10 weiße und 2 schwarze Kugeln gezogen.

Ein Verfahren, die aufgestellte Hypothese über eine Wahrscheinlichkeitsverteilung aufgrund eines Stichprobenergebnisses zu überprüfen heißt Test. Hier haben wir, da die Hypothese anhand von Vorzeichen überprüft wird, einen Vorzeichentest:

7.2 Vorzeichentest

Nehmen wir an, daß unsere Hypothese wahr ist, so geschieht dies mit Wahrscheinlichkeit

$$\begin{aligned}\mathbf{P}(10 \text{ weiße Kugeln}) &= \binom{12}{10} \cdot \frac{1}{2^{12}} \\ &= \frac{12!}{10! \cdot 2!} \cdot \frac{1}{2^{12}} = 0,01611.\end{aligned}$$

Wir betrachten die noch extremeren Fälle

$$\begin{aligned}\mathbf{P}(11 \text{ weiße Kugeln}) &= \binom{12}{11} \cdot \frac{1}{2^{12}} = 0,00293 \\ \mathbf{P}(12 \text{ weiße Kugeln}) &= \binom{12}{12} \cdot \frac{1}{2^{12}} = 0,0002\end{aligned}$$

$$\Rightarrow \quad \mathbf{P}(\text{Anzahl weißer Kugeln} \geq 10) = 0,01928.$$

Falls also unsere Hypothese gilt, so folgt, daß das Ereignis „Anzahl weißer Kugeln $\geq$ 10“ nur mit Wahrscheinlichkeit $\approx 0,02$ eintritt. Diese Wahrscheinlichkeit deutet man als Indiz gegen die Richtigkeit der Hypothese und verwirft sie somit.

Der Bereich der Anzahl der Pluszeichen, bei dem die Hypothese verworfen wird, heißt <u>kritischer Bereich</u> oder <u>Ablehnungsbereich</u> K. Entsprechend heißt der Bereich, indem die Hypothese nicht verworfen wird, <u>Annahmebereich</u> $\overline{K} = K^{\complement}$.

Wird das obige Experiment sehr oft durchgeführt und lehnt man die Hypothese immer bei 10 und mehr Pluszeichen ab, so begeht man in höchstens $\alpha = 0,02$ (2%) der Fälle einen Fehler, denn man verwirft dann irrtümlich eine wahre Hypothese. Dieser Fehler, die Hypothese zu verwerfen, obwohl sie stimmt, heißt <u>Irrtumswahrscheinlichkeit</u> α des Tests.

Nach Festlegung der Irrtumswahrscheinlichkeit α kann man den Ablehnungsbereich K ermitteln.

7.2.1 Ermittlung des Ablehnbereichs K

Bei n Paaren von Meßwerten (vorher/ nachher) ergibt sich ein n-Tupel von Vorzeichen. Unter der Annahme der Hypothese, daß „+“ und „−“ gleichwahrscheinlich sind, ergibt sich für die Anzahl der „+“-Zeichen

$$\mathbf{P}\left(k \text{ „}+\text{“-Zeichen}\right) = \binom{n}{k} \cdot \frac{1}{2^n} \quad \text{für } k = 0, \ldots n.$$

Um die Irrtumswahrscheinlichkeit α maximal auszunutzen, bestimmt man die kleinste Zahl $g \in \{0, \ldots, n\}$, für die gilt

$$\mathbf{P}\left(\text{Anzahl der „}+\text{“-Zeichen} \geq g\right) = \left(\binom{n}{g} \cdot \binom{n}{g+1} \cdot \ldots \cdot \binom{n}{n}\right) \cdot \frac{1}{2^n} \leq \alpha$$

$$\underbrace{0, 1, \ldots, g-1}_{\overline{K}}, \underbrace{g, \ldots, n}_{K}$$

Anschließend prüft man nach, ob die Anzahl der „+“-Zeichen im Experiment zu K gehört ($\rightarrow$ Ablehnung der Hypothese) oder nicht, d.h. zu $\overline{K}$ gehört ($\rightarrow$ Beibehaltung der Hypothese). Gehört die Anzahl der „+“-Zeichen zu K, so ist der Fehler den man macht, wenn man in diesem Fall die Hypothese verwirft kleiner gleich α.

7.3 (rechtsseitiger) Vorzeichentest

Um mit Hilfe eines <u>rechtsseitigen Vorzeichentests</u> die Hypothese, daß Plus- und Minuszeichen gleich wahrscheinlich sind, zu testen geht man folgendermaßen vor:

(1) Vorgeben einer Irrtumswahrscheinlichkeit α ($\alpha = 0,05$; $\alpha = 0,02$; $\alpha = 0,01$).

(2) Bestimmung des Ablehnungsbereichs $K = \{g, \ldots, n\}$, d.h.

$$\mathbf{P}\left\{\text{Anzahl der „}+\text{“-Zeichen} \geq g\right\} \leq \alpha$$

(kleinstes g mit dieser Eigenschaft).

(3) Man verwirft die Hypothese, falls die Anzahl der „+“-Zeichen zu K gehört. Fällt der Wert in $K^{\complement}$ so kann die Hypothese nicht abgelehnt werden.

7.3.1 Beispiele

(a) Um bei Ratten den Einfluß von Eisen auf den Hämoglobingehalt im Blut zu untersuchen, wurde zunächst durch Fütterung von Kuhmilch der Hämoglobingehalt gesenkt. Anschließend erhielden die Ratten 2 Wochen lang täglich einen Zusatz von $0,5\,mg$ reinem Eisen. Vor Beginn und nach Beendigung der Zugabe von Eisen wurde der Hämoglobingehalt ($\frac{g}{100\,cm^3}$ Blut) gemessen.

Ratte Nummer	1	2	3	4	5	6	7	8	9	10	11	12
vor Beginn	3,4	3,0	3,0	3,4	3,7	4,0	2,9	2,9	3,1	2,8	2,8	2,4
nach Zugabe	4,4	2,2	3,4	3,3	3,5	4,1	5,2	4,5	3,5	4,0	4,1	2,7

Kann man aufgrund dieser Meßwerte mit einer Irrtumswahrscheinlichkeit von 5% behaupten, daß Eisenzugabe den Hämoglobingealt bei Ratten erhöht?
Antwort:

1	2	3	4	5	6	7	8	9	10	11	12
+	−	+	−	−	+	+	+	+	+	+	+

- Die Hypothese lautet: Eisenzugabe erhöht den Hämoglobingehalt bei Ratten nicht; „+“ und „−“-Zeichen treten mit der gleichen Wahrscheinlichkeit auf.
- Stichprobenumfang $n = 12$; Irrtumswahrscheinlichkeit $\alpha = 0,005$.
- X: Anzahl der „+“-Zeichen bei 12 Paaren von Meßwerten.
- Ermittlung des Ablehnbereichs K:
  Gesucht ist bei wahrer Hypothese die kleinste Zahl g mit

$$\begin{aligned}
\mathbf{P}\{X \geq g\} &\leq \alpha = 0,05 \\
\mathbf{P}\{X \geq 10\} &= \left(\binom{12}{10} + \binom{12}{11} + \binom{12}{12}\right) \cdot \frac{1}{2^{12}} \\
&= (66 + 12 + 1) \cdot \frac{1}{2^{12}} \approx 0,0193 \leq 0,05 \\
\mathbf{P}\{X \geq 9\} &= \frac{299}{4096} \approx 0,073 \geq 0,05 \\
&\Rightarrow \quad g = 10
\end{aligned}$$

  Ablehnungsbereich: $K = \{10, 11, 12\}$.
- Aus dem obigen Paaren von Meßwerten ergeben 9 „+”-Zeichen. Da $9 \notin K$ kann man die Hypothese nicht ablehnen.

(b) Um zu untersuchen, ob bei Zwillingen der Erstgeborene ein höheres Geburtsgewicht hat als der Zweitgeborene, wurde bei Zwillingen das Geburtsgewicht (in g) gemessen.

Zwillingspaar Nr.	1	2	3	4	5	6	7	8	9	10
Erstgeborener	3440	4500	3820	3540	3650	3690	3750	3800	3650	3200
Zweitgeborener	3700	4080	3200	3700	3550	3350	3500	3700	3420	3250

Kann man aufgrund dieser Daten sagen, daß bei Zwillingen die Erstgeborenen ein höheres Geburtsgewicht haben als die Zweitgeborenen? (Irrtumswahrscheinlichkeit 5%)
Antwort:

1	2	3	4	5	6	7	8	9	10
−	+	+	−	+	+	+	+	+	−

- Die Hypothese lautet: Die Erstgeborenen haben kein höheres Geburtsgewicht als die Zweitgeborenen; „+“- und „−“-Zeichen treten mit gleicher Wahrscheinlichkeit auf.
- Stichprobenumfang $n = 10$; Irrtumswahrscheinlichleit $\alpha = 0,05$.
- X: Anzahl der“+“-Zeichen bei 10 Paaren von Meßwerten.

- Ermittlung des Ablehnunugsbereichs K:
  Gesucht ist bei wahrer Hypothese die kleinste ganze Zahl g mit

$$\begin{aligned}
\mathbf{P}\{X \geq g\} &\leq 0,05 \\
\mathbf{P}\{X \geq 8\} &= \left(\binom{10}{8} + \binom{10}{9} + \binom{10}{10}\right) \cdot \frac{1}{2^{10}} \\
&= (45 + 10 + 1) \cdot \frac{1}{2^{10}} = \frac{56}{1024} \approx 0,0546 \geq 0,05 \\
\mathbf{P}\{X \geq 9\} &= (10 + 1) \cdot \frac{1}{2^{10}} = \frac{11}{1024} \approx 0,0107 \leq 0,05 \\
&\Rightarrow \quad g = 9
\end{aligned}$$

  Ablehnungsbereich: $K = \{9, 10\}$

- Aus den obigen Paaren von Meßwerten ergeben sich 7 „+”-Zeichen. Da $7 \notin K$, kann die Hypothese nicht abgelehnt werden.

(c) Da man bei Tupajas (Halbaffen) festgestellt hatte, daß die linke und die rechte Nebenniere unterschiedlich gebaut sind, sollte untersucht werden, ob bei diesen Tieren das Gewicht (in mg) der linken Nebenniere größer ist als das der rechten.

Tier Nr.	1	2	3	4	5	6	7	8	9	10	11	12
linke Nebenniere	4,1	6,6	7,9	10,8	2,6	7,2	10,0	11,5	3,3	8,2	5,8	6,7
rechte Nebenniere	5,9	5,7	7,6	9,5	2,4	4,9	7,6	9,0	2,6	7,7	5,0	6,3

Kann man aufgrund dieser Messungen behaupten, daß bei Tupajas die linke Nebeniere ein größeres Gewicht hat als die rechte? (Irrtumswahrscheinlichkeit 5%)
Antwort:

1	2	3	4	5	6	7	8	9	10	11	12
−	+	+	+	+	+	+	+	+	+	+	+

- Hypothese: Die linke Nebenniere hat kein größeres Gewicht als die rechte; „+“- und „−“-Zeichen treten mit gleicher Wahrscheinlichkeit auf.
- Stichprobenumfang $n = 12$; Irrtumswahrscheinlichleit $\alpha = 0,05$.
- X: Anzahl der“+“-Zeichen bei 12 Paaren von Meßwerten.
- Ermittlung des Ablehnungsbereichs K:
  Gesucht ist bei wahrer Hypothese die kleinste ganze Zahl g mit

$$\mathbf{P}\{X \geq g\} \leq 0,05$$

  Aus Aufgabe (A.7.2.7.a) folgt: $g = 10$, Ablehnungsbereich $K = \{10, 11, 12\}$
- Aus den Paaren von Meßwerten ergeben sich 11 „+”-Zeichen. Da $11 \in K$, wird die Hypothese abgelehnt.
  Mit einer Irrtumswahrscheinlichkeit von 5% kann man behaupten, daß die linke Nebenniere bei Tapajas ein höheres Gewicht hat als die rechte.

7.4 Zweiseitiger Vorzeichentest

Ein einseitiger Vorzeichentest liegt nahe, wenn man vermutet, daß eine der beiden „Behandlungen“ eine bessere Wirkung zeigt als die andere. Ist dies nicht der Fall, und möchte man nur überprüfen, ob die beiden „Behandlungen“ unterschiedliche Wirkungen zeigen, führt man einen zweiseitigen Vorzeichentest durch.

Die Hypothese, daß beide „Behandlungen“ doch gleiche Wirkungen zeigen, wird dann verworfen, wenn die Anzahl X der „+“-Zeichen „große“ oder „kleine“ Werte annimmt.

Der Ablehnbereich kann wie folgt beschrieben werden:

$$K = \{0, \ldots, g_l\} \cup \{g_r, \ldots, n\}\,.$$

Die vorgegebene Irrtumswahrscheinlichkeit α wird nun in zweimal $\frac{\alpha}{2}$ aufgeteilt.

Zu bestimmen sind:

g_l möglichst groß mit $\mathbf{P}\{X \le g_l\} \le \frac{\alpha}{2}$;

g_r möglichst groß mit $\mathbf{P}\{X \ge g_r\} \le \frac{\alpha}{2}$.

7.4.1 Beispiele

(a) Um festzustellen, ob sich der Triglyceridgehalt von Blutplasma während der Lagerung ändert, wurde von 14 Plasmaproben der Triglyceridgehalt (in $\frac{mg}{100ml}$) sofort nach der Entnahme und 8 Monate später nach der Lagerung in gefrorenem Zustand gemessen.

Probe Nr.	1	2	3	4	5	6	7	8	9	10	11	12	13	14
1. Messung	74	80	177	136	169	89	125	130	129	83	42	91	105	42
2. Messung	65	86	185	132	182	96	124	127	115	81	48	81	117	53

Liefern die beiden Messungen unterschiedliche Werte? (Irrtumswahrscheinlichkeit 1%)

Antwort:

1	2	3	4	5	6	7	8	9	10	11	12	13	14
+	−	−	+	−	−	+	+	+	+	−	+	−	−

- Hypothese: Blutplasma ändert während der Lagerung seinen Triglyceridgehalt nicht.
- Stichprobenumfang $n = 14$; Irrtumswahrscheinlichkeit $\alpha = 0,01$.
- X: Anzahl der "+"-Zeichen („−"-Zeichen) bei 14 Paaren von Meßwerten.
- Ermittlung des Ablehnbereichs K
  Gesucht ist bei wahrer Hypothese die kleinste ganze Zahl g_r mit

$$\begin{aligned}
\mathbf{P}\{X \ge g_r\} &\le \frac{\alpha}{2} = 0,005 \\
\mathbf{P}\{X \ge 12\} &= \left(\binom{14}{12} + \binom{14}{13} + \binom{14}{14}\right) \cdot \frac{1}{2^{14}} = (91 + 14 + 1) \cdot \frac{1}{2^{14}} \\
&= 106 \cdot \frac{1}{2^{14}} \approx 0,00647 \ge 0,005 \\
\mathbf{P}\{X \ge 13\} &= 15 \cdot \frac{1}{2^{14}} \approx 0,0009155 \le 0,005 \\
&\Rightarrow \quad g_r = 13
\end{aligned}$$

Gesucht ist bei wahrer Hypothese die kleinste ganze Zahl g_l mit

$$\begin{aligned}
\mathbf{P}\{X \le g_l\} &\le \frac{\alpha}{2} = 0,005 \\
\mathbf{P}\{X \le 1\} &= \left(\binom{14}{1} + \binom{14}{0}\right) \cdot \frac{1}{2^{14}} = (14 + 1) \cdot \frac{1}{2^{14}} \\
&= 15 \cdot \frac{1}{2^{14}} \approx 0,000916 \le 0,005 \\
\mathbf{P}\{X \le 2\} &= 106 \cdot \frac{1}{2^{14}} \approx 0,00647 \ge 0,005 \\
&\Rightarrow \quad g_l = 1
\end{aligned}$$

Ablehnungsbereich $K = \{0, 1\} \cup \{13, 14\}$

- Die Paare von Meßwerten ergeben 7 „+“- und 7 „−“-Zeichen, wenn man die Werte der 2. Messung von denen der 1. Messung abzieht.
  Da $7 \notin K$, muß die Hypothese beibehalten werden.

(b) Man vermutet, daß die Erträge zweier neuer Züchtungen von Tomatensorten nicht gleich sind. Zum Vergleich wurde von jeder Sorte je eine Pflanze in Töpfe gepflanzt und unter gleichen Bedingungen aufgezogen. Die Tabelle enthält die Ertäge (in g) jeder Pflanze.

Pflanzenpaar Nr.	1	2	3	4	5	6	7	8	9	10
Sorte I	1450	1230	1885	985	1675	1610	1050	1340	1250	880
Sorte II	1120	1100	1970	820	1230	1220	875	1245	1325	760

Liefern beide Sorten unterschiedliche Erträge? (Irrtumswahrscheinlichkeit 5%)
Antwort:

1	2	3	4	5	6	7	8	9	10
+	+	−	+	+	+	+	+	−	+

- Hypothese: Beide Sorten liefern gleiche Erträge.
- Stichprobenumfang $n = 10$; Irrtumswahrscheinlichkeit $\alpha = 0,05$.
- X: Anzahl der "+"-Zeichen („−"-Zeichen) bei 10 Paaren von Meßwerten.
- Ermittlung des Ablehnbereichs K
  Gesucht ist bei wahrer Hypothese die kleinste ganze Zahl g mit

$$
\begin{aligned}
\mathbf{P}\{X \geq g\} &\leq 0,025 \\
\mathbf{P}\{X \geq 8\} &= \left(\binom{10}{8} + \binom{10}{9} + \binom{10}{10}\right) \cdot \frac{1}{2^{10}} = (45 + 10 + 1) \cdot \frac{1}{2^{10}} \\
&= 56 \cdot \frac{1}{2^{10}} \approx 0,0547 \geq 0,025 \\
\mathbf{P}\{X \geq 9\} &= \left(\binom{10}{9} + \binom{10}{10}\right) \cdot \frac{1}{2^{10}} = (10 + 1) \cdot \frac{1}{2^{10}} \\
&= 11 \cdot \frac{1}{2^{10}} \approx 0,0107 \leq 0,025 \\
&\Rightarrow \quad g_r = 9
\end{aligned}
$$

  Aus der Symmetrie der Binomialverteilung folgt

$$g_l = 10 - 9 = 1$$

  Ablehnbereich $K = \{0, 1\} \cup \{9, 10\}$
- Aus den Meßwerten ergeben sich 8 „+“- und 2 „−“-Zeichen, wenn man die Erträge der 2. Sorte von denen der 1. Sorte abzieht.
  Da $8 \notin K$, kann die Hypothese nicht abgelehnt werden.

(c) Um die Beliebtheit zweier neuer Brotsorten I und II zu untersuchen, wurden 15 zufällig ausgewählte Testpersonen in einem „Blindversuch“ (die Testpersonen wissen nicht, welche Sorte sie verzehren) aufgefordert, für jede Sorte eine der Geschmacksnoten 1 (sehr schmackhaft) bis 4 (nicht schmackhaft) zu vergeben.

Probe Nr.	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15
Sorte I	2	4	1	3	2	2	2	3	2	2	2	2	3	1	2
Sorte II	1	2	3	2	3	1	1	4	3	1	1	2	2	2	1

Prüfe durch einen zweiseitigen Test, ob beide Brotsorten gleich schmackhaft sind oder nicht. (Irrtumswahrscheinlichkeit 5%)
Antwort: Dem interessierten Leser überlassen.

7.5 Einfache Nullhypothese/ zweiseitiger Signifikanztest

In einer Urne mit sehr vielen Kugeln sind angeblich 30% aller Kugeln weiß. Aufgrund der Ergebnisse von 100 Ziehungen mit Zurücklegen aus der Urne soll entschieden werden, ob man der Angabe trauen kann oder nicht.

Über die Wahrscheinlichkeit, daß bei einer einmaligen Ziehung aus dieser Urne eine weiße Kugel erscheint, wird also die Hypothese $p = 0,3$ aufgestellt. Diese zu überprüfende Hypothese heißt Nullhypothese H_0. Schreibweise:

$$H_0 : p = 0,3$$

Die Hypothese $p \neq 0,3$ wird Gegenhypothese H_1 genannt; Schreibweise:

$$H_1 : p \neq 0,3$$

Hypothesen der Form $p = p_0$, die also durch genau einen Wert festgelegt sind, nennt man einfache Hypothese im Unterschied zu Hypothesen, zum Beispiel der Form $p \neq p_0$, die als zusammengesetzt bezeichnet werden.

H_0 ist im vorliegenden Fall also eine einfache Hypothese und H_1 eine zusammengesetzte Hypothese. Die Überprüfung von H_0 gegen H_1 wird anhand einer Stichprobe vorgenommen:

Stichprobenumfang:

$$n = 100$$

$$X : \text{Anzahl der weißen Kugeln in der Stichprobe}$$

Bei wahrer Nullhypothese ist X $\mathcal{B}_{100;0,3}$-verteilt. Sehr „kleine“ oder sehr „große“ Werte von X deutet man als Indiz gegen die Richtigkeit von H_0. (Man wird also dann H_0 ablehnen, oder wie man auch sagt, verwerfen.) Die Menge der Werte von X, bei deren Eintreten H_0 verworfen wird, heißt Ablehnungsbereich K. Der Annahmebereich $\overline{K} = K^{\complement}$ wird durch die übrigen Werte von X gebildet, bei deren Eintreten man H_0 nicht ablehnt. Nun kann aber X auch Werte aus dem Ablehnungsbereich K annehmen, obwohl H_0 zutrifft. Die Wahrscheinlichkeit dafür, die Nullhypothese zu verwerfen, obwohl sie zutrifft, heißt Irrtumswahrscheinlichkeit α.

Wird die Nullhypothese abgelehnt, spricht man von einem signifikanten Unterschied zwischen der in der Nullhypothese angenommenen und der tatsächlich vorliegenden Verteilung. Bei Ablehnung von H_0 ist damit keineswegs „bewiesen“, daß H_0 falsch ist.

Aber es wird mit einer Irrtumswahrscheinlichkeit α festgelegt, daß der beobachtete Wert von X mit der Nullhypothese unverträglich ist. Fällt der Wert von X dagegen nicht in den Ablehnungsbereich, so ist für den Entscheidungspozeß wenig gewonnen, denn in diesem Fall kann man nicht sagen, daß H_0 sich als richtig erwiesen hat. Man kann nur feststellen, daß der Wert nicht im Widerspruch zur Nullhypothese steht. Man sagt deshalb auch „nur“ die Nullhypothese kann nicht abgelehnt werden.

Ein Verfahren, das auf die Widerlegung der Nullhypothese abzielt heißt Signifikanztest

7.5.1 Schema eines zweiseitigen Signifikanztests

(1) Wie lauten die Nullhypothese H_0 und die Gegenhypothese H_1?

(2) Wie groß sind der Stichprobenumfang n und die Irrtumswahrscheinlichkeit α?

(3) Festlegung von X; wie ist X, wenn H_0 zutrifft, verteilt?

(4) Wie lautet der Ablehnungsbereich?

(5) Wie wird aufgrund der Stichprobe entschieden?

7.5.2 Beispiel

Der Oberbürgermeister einer Stadt erhielt bei der letzten Wahl 60% der Stimmen. Bei einer Befragung vor der nächsten Wahl bevorzugten von 100 zufällig ausgewählten Personen 48 den bisherigen Oberbürgermeister. Kann man hieraus mit einer Irrtumswahrscheinlichkeit von 5% schließen, daß sich sein Stimmanteil seit der letzten Wahl verändert hat?

(1)

$$H_0 : p = 0,6 \qquad H_1 : p \neq 0,6$$

(2) Stichprobenumfang $n = 100$; Irrtumswahrscheinlichkeit $\alpha = 0,05$.

(3) X: Anzahl der Personen (von 100), die den bisherigen Oberbürgermeister bevorzugen. X ist bei wahrer Nullhypothese $\mathcal{B}_{100;0,6}$-verteilt.

(4) <u>Gesucht</u>: Größte ganze Zahl g_l, so daß $\mathbf{P}\{X \leq g_l\} \leq \frac{\alpha}{2} = 0,025$ und die kleinste ganze Zahl g_r, so daß $\mathbf{P}\{X \geq g_r\} \leq \frac{\alpha}{2} = 0,025$.
Aus einer Tabelle für die Summenfunktion $F_{n,p}$ der Binomialverteilung (hier $F_{100;0,6}$) folgt dann:

$$\begin{aligned} \mathbf{P}\{X \leq 50\} &= 0,0271 \geq 0,025 \\ \mathbf{P}\{X \leq 49\} &= 0,0168 \leq 0,025 \end{aligned}$$

$$\Rightarrow \quad g_l = 49$$

und

$$\begin{aligned} \mathbf{P}\{X \geq g_r\} &= 1 - \mathbf{P}\{X < g_r\} = 1 - \mathbf{P}\{X \leq g_r - 1\} = 1 - F_{n,p}(g_r - 1) \\ \mathbf{P}\{X \geq 70\} &= 1 - F_{100;0,6}(69) = 0,0248 \leq 0,025 \\ \mathbf{P}\{X \geq 69\} &= 0,0398 \geq 0,025 \end{aligned}$$

$$\Rightarrow \quad g_r = 70$$

Ablehnungsbereich:

$$K = \{0, 1, \ldots, 49\} \cup \{70, \ldots, 100\}$$

(5) Da $48 \in K$ wird H_0 abgelehnt. Man entscheidet sich also dafür, daß der Stimmanteil des bisherigen Oberbürgermeisters zum Zeitpunkt der Umfrage nicht 60% betrug.

7.5.3 Fehler beim Signifikanztest

Bei einem Signifikanztest wird die Entscheidung darüber, ob die Nullhypothese abgelehnt wird oder nicht, aufgrund des Ergebnisses eines Zufallsexperimentes getroffen. Folglich können bei dieser Entscheidung Fehler begangen werden.

Einige Fehler kennen wir schon: Den Fehler, welchen man begeht, wenn nach Festlegung der Nullhypothese X einen Wert aus dem Ablehnungsbereich annimmt und hierdurch die Nullhypothese abgelehnt wird, owohl sie in Wirklichkeit richtig ist. $\rightsquigarrow$ Fehler 1. Art.

Einen ganz anderen Fehler begeht man, wenn X einen Wert aus dem Annahmehmebereich annimmt und hierdurch die Nullhypothese nicht abgelehnt werden kann, obwohl sie in Wirklichkeit falsch ist. $\rightsquigarrow$ Fehler 2. Art.

		Zustand der Wirklichkeit	
		H_0 wahr	H_0 falsch
Entscheidung: Die Hypothese H_0 wird	abgelehnt	Fehler 1. Art	richtige Entscheidung
	beibehalten	richtige Entscheidung	Fehler 2. Art

Fehler 1. Art: Eine richtige Hypothese wird abgelehnt.

Fehler 2. Art: Eine falsche Hypothese wird nicht abgelehnt.

Beachte: Da bei einem Test zwei verschiedene Fehler möglich sind, kann man nicht sagen, die Irrtumswahrscheinlichkeit sei die Wahrscheinlichkeit, sich bei einem Test zu irren. Sie ist vielmehr die Wahrscheinlichkeit, eine zutreffende Nullhypothese abzulehnen.

Bei einem Signifikanztest wird für die Wahrscheinlichkeit, einen Fehler 1. Art zu begehen, eine Irrtumswahrscheinlichkeit α angegeben.

Die Wahrscheinlichkeit für einen Fehler 2. Art wird mit β bezeichnet. Die Größe von β kann bei zusammengesetzter Gegenhypothese nicht berechnet werden, denn es ist im allgemeinen unbekannt, welcher Wert von p in Wirklichkeit zutrifft.
Sie kann ebenfalls für willkürlich angenommene Werte der Gegenhypothese ermittelt werden.
Nur in dem Sonderfall, daß neben H_0 auch H_1 eine einfache Hypothese ist, läßt sich das Risiko 2. Art berechnen; ein Test dieser Art heißt Alternativtest.

7.6 Zusammengesetzte Nullhypothese, einseitiger Signifikanztest

Bei den bisherigen Problemen war die Nullhypothese einfach und die Gegenhypothese zusammengesetzt. Im folgenden betrachten wir solche Fälle, bei denen auch die Nullhypothese zusammengesetzt ist.

Beispiel:
Der Lieferant eines Massenartikels behauptet gegenüber einem Abnehmer, daß der Ausschußanteil höchstens 5% beträgt. Um dies zu überprüfen, kommen sie überein, eine Stichprobe mit Zurücklegen zu ziehen und aufgrund des Ergebnisses der Stichprobe über die Annahme beziehungsweise Ablehnung der Lieferung zu entscheiden.

$$\begin{array}{ll} H_0 : p \leq 0,05 & \text{Nullhypothese} \\ H_1 : p > 0,05 & \text{Gegenhypothese} \end{array} \quad \text{rechtsseitiger Signifikanztest}$$

Allgemein ist also die zusammengesetzte Nullhypothese $H_0 : p \leq p_0$ gegen die ebenfalls zusammengesetzte Gegenhypothese $H_1 : p > p_0$ zu testen.
Da man H_0 nur bei einer großen Anzahl von Ausschußstücken in der Stichprobe ablehnen wird, nennt man diesen Test einen rechtsseitigen Signifikanztest.

$$\begin{array}{ll} H_0 : p \geq p_0 & \text{Nullhypothese} \\ H_1 : p < p_0 & \text{Gegenhypothese} \end{array} \quad \text{linksseitiger Signifikanztest}$$

Beide Arten von Tests heißen einseitige Tests.

Bei einem einseitigen Test müssen die entsprechenden 5 Schritte wie beim zweiseitigen Test durchgeführt werden. Dabei muß allerdings der Begriff Irrtumswahrscheinlichkeit etwas verallgemeinert werden.
Bei einem einseitigen Test ist die Verteilung nicht mehr eindeutig. Die Wahrscheinlichkeit einen Fehler 1. Art bei gleichbleibendem Ablehnungsbereich ist aber sowohl bei $H_0 : p \leq p_0$ als auch bei $H_0 : p \geq p_0$ für $p = p_0$ am größten.
Daher versteht man (bei einem einseitigen Signifikanztest) unter der Irrtumswahrscheinlichkeit diese maximale Wahrscheinlichkeit.

Da wie bei einem zweiseitigen auch bei einem einseitigen Signifikanztest die Gegenhypothese H_1 zusammengesetzt ist, kann hier im allgemeinen das Risiko 2. Art nicht berechnet werden.

7.6.1 Beispiele

(a) In einer Fabrik werden Tüten mit Mehl abgefüllt. Aus bisherigen Überprüfungen ist bekannt, daß höchstens 2% aller Tüten weniger als $1kg$ Mehl enthalten.
Bei einer erneuten Überprüfung findet man unter 150 Paketen 5 mit einem Gewicht unter $1kg$.

- Kann man hieraus mit einer Irrtumswahrscheinlichkeit von 5% schließen, daß sich der Anteil der Tüten mit weniger als $1kg$ Gewicht erhöht hat?
- Mit welcher Wahrscheinlichkeit ist man weiter der Meinung, höchstens 2% der Tüten enthalten weniger als $1kg$ Mehl, obwohl sich dieser Anteil in Wirklichkeit auf 6% erhöht hat?

Zur ersten Frage:

(1)

$$H_0 : p \leq 0,02 \qquad H_1 : p > 0,02$$

(2) Stichprobenumfang: $n = 150$; Irrtumswahrscheinlichkeit $\alpha = 0,05$.

(3) X : Anzahl der Tüten unter den 150, die weniger als $1kg$ Mehl enthalten. X ist bei wahrer Nullhypothese im Extremfall $\mathcal{B}_{150;0,02}$-verteilt.

(4) Gesucht ist die kleinste ganze Zahl g mit $\mathbf{P}\{X \geq g\} \leq 0,05$.
Wir gehen über das Gegenereignis:

$$\mathbf{P}\{X \leq g-1\} \geq 0,95.$$

$$\begin{array}{rrcl}
g=4: & \mathbf{P}\{X \leq 3\} & = & \binom{150}{3} \cdot (0,02)^3 \cdot (0,98)^{147} + \binom{150}{2} \cdot (0,02)^2 \cdot (0,98)^{148} \\
 & & & +\binom{150}{1} \cdot (0,02)^1 \cdot (0,98)^{149} + \binom{150}{0} \cdot (0,02)^0 \cdot (0,98)^{150} \\
 & & = & 0,2263138361697 + 0,2247846886281 \\
 & & & +0,147844962990 + 0,04829602124349 \\
 & & = & 0,6472395089413 \\
g=5: & \mathbf{P}\{X \leq 4\} & = & \mathbf{P}\{X \leq 3\} + \binom{150}{4} \cdot (0,02)^4 \cdot (0,98)^{146} \\
 & & = & 0,6472395089413 + 0,1697353771273 \\
 & & = & 0,8169748860686 \\
g=6: & \mathbf{P}\{X \leq 5\} & = & \mathbf{P}\{X \leq 4\} + \binom{150}{5} \cdot (0,02)^5 \cdot (0,98)^{145} \\
 & & = & 0,8169748860686 + 0,1011484288187 \\
 & & = & 0,9181233148873 \\
g=7: & \mathbf{P}\{X \leq 6\} & = & \mathbf{P}\{X \leq 5\} + \binom{150}{6} \cdot (0,02)^6 \cdot (0,98)^{144} \\
 & & = & 0,9181233148873 + 0,04988612985957 \\
 & & = & 0,968009444769
\end{array}$$

$$\Rightarrow \quad g = 7$$

Ablehnungsbereich: $K = \{7, 8, \ldots, 150\}$

(5) Da $5 \notin K$ kann die Hypothese nicht abgelehnt werde. Man geht also mit der Irrtumswahrscheinlichkeit $0,05$ weiter davon aus, daß höchstens 2% aller Tüten weniger als $1kg$ Mehl enthalten

Zur zweiten Frage:
Gesucht ist die Wahrscheinlichkeit dafür, daß ein Wert aus dem Annahmebereich $K = \{0, \ldots, 6\}$ angenommen wird, obwohl in Wirklichkeit $p = 0,06$ gilt:

$$\begin{array}{rcl}
\beta & = & \binom{150}{0} \cdot (0,06)^0 \cdot (0,94)^{150} + \binom{150}{1} \cdot (0,06)^1 \cdot (0,94)^{149} \\
 & & +\binom{150}{2} \cdot (0,06)^2 \cdot (0,94)^{148} + \binom{150}{3} \cdot (0,06)^3 \cdot (0,94)^{147} \\
 & & +\binom{150}{4} \cdot (0,06)^4 \cdot (0,94)^{146} + \binom{150}{5} \cdot (0,06)^5 \cdot (0,94)^{145} \\
 & & +\binom{150}{6} \cdot (0,06)^6 \cdot (0,94)^{144} \\
 & \approx & 0,1984
\end{array}$$

(b) Der Vertreter einer Kaffeemarke behauptet, daß mindestens 70% aller Kunden, die Kaffee kaufen, die von ihm vertriebene Marke wählen. Bei einer Überprüfung wählen von 100 Kaffeekäufern nur 59 die Marke des Vertreters.

- Läßt sich hieraus mit der Irrtumswahrscheinlichkeit von 5% ein Widerspruch gegen die Behauptung des Vertreters herleiten?
- Mit welcher Wahrscheinlichkeit wird die Behauptung des Vertreter fälschlicherweise angenommen, wenn in Wirklichkeit nur 50% aller Kaffekäufer die von ihm vertriebene Marke kaufen?

<u>Zur ersten Frage</u>:

(1)

$$H_0 : p \geq 0,7 \qquad H_1 : p < 0,7$$

(2) Stichprobenumfang $n = 100$; Irrtumswahrscheinlichkeit $\alpha = 0,05$.

(3) X : Anzahl der Kaffeekäufer unter den 100, die die Marke des Vertreters wählen. X ist bei wahrer Nullhypothese im Extremfall $\mathcal{B}_{100;0,7}$-verteilt

(4) Gesucht ist die größte ganze Zahl g mit $\mathbf{P}\{X \leq g\} \leq 0,05$

$$\begin{array}{lrcl} g = 62: & \mathbf{P}\{X \leq 62\} & = & 1 - 0,9470 = 0,053 \geq 0,05 \\ g = 61: & \mathbf{P}\{X \leq 61\} & = & 1 - 0,9660 = 0,034 \leq 0,05 \end{array}$$

$$\Rightarrow \quad g = 61$$

Ablehnungsbereich: $K = \{0, 1, \ldots, 61\}$

(5) Da $59 \in K$, wird die Hypothese abgelehnt. Man kann also mit einer Irrtumswahrscheinlichkeit $\alpha = 0,05$ sagen, daß die Behauptung des Vertreters unzutreffend ist.

<u>Zur zweiten Frage</u>:
Gesucht ist die Wahrscheinlichkeit, daß einWert aus $\overline{K} = \{62, \ldots, 100\}$ angenommen wird, wobei in Wirklichkeit $p = 0,5$ ist. Y sei $\mathcal{B}_{100;0,5}$-verteilt

$$\beta = \mathbf{P}\{Y \geq 62\} = 1 - \mathbf{P}\{Y \leq 61\} = 1 - 0,9895 = 0,0105.$$

8. Approximation der Binomialverteilung

8.1 Näherungsformel von *Moivre-Laplace*

Für eine $\mathcal{B}_{n;p}$-verteilte Zufallsvariable gilt bei großen Werten von n:

$$\mathbf{P}\{k_1 \leq X \leq k_2\} \approx \Phi(x_2) - \Phi(x_1),$$

mit

$$x_1 = \frac{k_1 - n \cdot p}{\sqrt{n \cdot p \cdot (1-p)}} \quad \text{und} \quad x_2 = \frac{k_2 - n \cdot p}{\sqrt{n \cdot p \cdot (1-p)}}.$$

Die Näherungsformel liefert brauchbare Werte, wenn die Faustregel $n \cdot p \cdot (1-p) > 9$ erfüllt ist.

Für $k_1 = 0$ liegt $\Phi\left(\frac{0 - n \cdot p}{\sqrt{n \cdot p \cdot (1-p)}}\right)$ bei großen n unterhalb der Genauigkeitsgrenze. Damit ergibt sich für $\mathcal{B}_{n;p}$-verteilte Zufallsvariable X bei großen n:

$$\mathbf{P}\{X \leq k\} \approx \Phi\left(\frac{k - n \cdot p}{\sqrt{n \cdot p \cdot (1-p)}}\right).$$

Demnach gilt: Ist für eine Binomialverteilung die Bedingung $n \cdot p \cdot (1-p) > 9$ erfüllt, kann man bei einem Signifikanztest den Ablehnungsbereich näherungsweise mit Hilfe der zugehörigen Normalverteilung ermitteln.

8.2 Ungleichung von *Tschebyscheff*

Zur Erinnerung:

Tschebyscheff-Ungleichung (5.17): Es sei X eine quadratisch integrierbare Zufallsvariable. Dann gilt für alle $\varepsilon > 0$:

$$\mathbf{P}\{|X - \mathbf{E}(X)| \geq \varepsilon\} \leq \frac{\mathbf{V}(X)}{\varepsilon^2}.$$

Die Abschätzung gilt für alle Verteilungen, sie kann also im Einzelfall sehr grob sein.
Wähle in der Ungleichung für ε Vielfache von $\sigma = \sigma(X) = \sqrt{\mathbf{V}(X)}$: $\varepsilon = k \cdot \sigma$

$$\mathbf{P}\{|X - \mu| \geq k \cdot \sigma\} \leq \frac{\sigma^2}{k^2 \cdot \sigma^2} = \frac{1}{k^2} \qquad (\mu = \mathbf{E}(X))$$

Die Wahrscheinlichkeit, daß X einen Wert annimmt, der von μ um mindestens das k-fache der Standardabweichung σ abweicht, ist demnach höchstens $\frac{1}{k^2}$.

Spezialfälle $k = 1, 2, 3$:

$$\begin{aligned}
\mathbf{P}\{|X - \mu| \geq \sigma\} &\leq 1 \\
\mathbf{P}\{|X - \mu| \geq 2\sigma\} &\leq \frac{1}{4} = 0,25 \\
\mathbf{P}\{|X - \mu| \geq 3\sigma\} &\leq \frac{1}{9} \approx 0,111
\end{aligned}$$

Bei langen Serien von Durchführungen des Zufallsexperimentes kann man davon ausgehen, daß nicht mehr als etwa 11% der auftretenden Werte außerhalb des 3σ-Intervalls liegen werden. (Prinzipiell kann es jedoch im Einzelfall auch anders sein.)

Wird ein Zufallsexperiment sehr oft wiederholt, so liegen von den Werten, welche die Zufallsvariable X dabei annimmt

im 2σ-Intervall $]\mu - 2\sigma; \mu + 2\sigma[$ mindestens etwa 75%
im 3σ-Intervall $]\mu - 3\sigma; \mu + 3\sigma[$ mindestens etwa 89%

Sei X $\mathcal{B}_{n;p}$-verteilt, X zählt die Anzahl der Treffer in einer Bernoulli-Kette der Länge n:

$$X = X_1 + \ldots + X_n, \text{ wobei } X_i \, \mathcal{B}_{1;p}\text{-verteilt ist,}$$

daß heißt $\mathbf{P}\{X_i = 1\} = p$, $\mathbf{P}\{X_i = 0\} = q = 1 - p$. Die (X_i) sind unabhängig.

$$\overline{X} = \frac{1}{n}\sum_{i=1}^{n} X_i = \frac{1}{n} \cdot X \text{ beschreibt die relativen Häufigkeiten für Treffer.}$$

$$\begin{aligned}
\mathbf{E}\left(\overline{X}\right) &= \mathbf{E}\left(\frac{1}{n}\sum_{i=1}^{n} X_i\right) = \frac{1}{n}\sum_{i=1}^{n} \underbrace{\mathbf{E}\left(X_i\right)}_{=p} = \frac{1}{n} \cdot n \cdot p = p \\
\mathbf{V}\left(\overline{X}\right) &= \mathbf{V}\left(\frac{1}{n}\sum_{i=1}^{n} X_i\right) = \frac{1}{n^2}\sum_{i=1}^{n} \underbrace{\mathbf{V}\left(X_i\right)}_{=p(1-p)} = \frac{1}{n^2} \cdot n \cdot p\,(1-p) = \frac{p\,(1-p)}{n}
\end{aligned}$$

$$\begin{aligned}
\mathbf{P}\left\{\left|\overline{X} - p\right| \geq \varepsilon\right\} &\leq \frac{1}{\varepsilon^2} \cdot \frac{1}{n} \cdot p\,(1-p) \\
&\overset{(*)}{\leq} \frac{1}{\varepsilon^2} \cdot \frac{1}{n} \cdot \frac{1}{2} \cdot \frac{1}{2} = \frac{1}{4n\varepsilon^2}.
\end{aligned}$$

Zu $(*)$: Das Produkt $p \cdot q$ nimmt seinen Maximalwert bei $p, q = \frac{1}{2}$ an, da:

$$f(p) = p \cdot q = p \cdot (1-p) = p - p^2$$

$$f'(p) = 1 - 2p \quad \Rightarrow \quad f'(p) = 0 \quad \Leftrightarrow \quad p = \frac{1}{2}$$

$$f''(p) = -2 < 0 \quad \Rightarrow \quad \text{Maximum.}$$

Ist X eine $\mathcal{B}_{n;p}$-verteilte Zufallsvariable, dann gilt für $\overline{X} = \frac{1}{n} \cdot X$:

$$\mathbf{P}\left\{\left|\overline{X} - p\right| < \varepsilon\right\} \geq 1 - \frac{1}{4n\varepsilon^2}.$$

⟦

$$\mathbf{P}\left\{\left|\overline{X} - p\right| \geq \varepsilon\right\} \leq \frac{1}{4n\varepsilon^2} \quad \Rightarrow \quad \mathbf{P}\left\{\left|\overline{X} - p\right| < \varepsilon\right\} = 1 - \mathbf{P}\left\{\left|\overline{X} - p\right| \geq \varepsilon\right\} \geq 1 - \frac{1}{4n\varepsilon^2}$$

⟧

9. Aufgaben zu Erwartungswert und Varianz

9.1 Aufgabe 1: Roulette

Das Roulette-Spiel ist ein Zufallsexperiment mit der Ereignismenge

$$\Omega = \{0, 1, 2, \ldots, 35, 36\} \quad \text{und} \quad \mathbf{P} = \mathcal{L}_\Omega .$$

Es sei

$$A = \{1, 2, \ldots, 11, 12\},$$

also:

$$\mathbf{P}(A) = \frac{12}{37}.$$

(a) Das Zufallsexperiment wird 1-mal durchgeführt. Spieler I setzt 10 EUR darauf, daß A eintritt. Ist dies der Fall, so bekommt er den 3-fachen Einsatz ausgezahlt. Tritt A nicht ein, so erfolgt keine Auszahlung.
Berechne den Erwartungswert des Gewinns für Spieler I.

(b) Spieler II setzt 10 EUR auf A. Tritt A ein, so erhält er den 3-fachen Einsatz ausbezahlt und hört mit dem Spielen auf. Tritt A nicht ein, so verdoppelt er den vorhergegangenen Einsatz und spielt entsprechend weiter, so lange, bis zum ersten Mal A eintritt (Verdopplungsstrategie). Er hat jedoch nur 70 EUR zur Verfügung.
Berechne den Erwartungswert seines Gewinns.
Wie ändert sich der Erwartungswert, wenn dem Spieler 150 EUR statt 70 EUR zur Verfügung stehen?

Zu (a):

X: Gewinn für Spieler I, bei einer Durchführung.

Wahrscheinlichkeitsverteilung:

x_i	-10	$+20$
$\mathbf{P}\{X = x_i\}$	$\frac{25}{37}$	$\frac{12}{37}$

$$\mathbf{E}(X) = -10 \cdot \frac{25}{37} + 20 \cdot \frac{12}{37} = -\frac{10}{37} \approx -0,27$$

Der Erwartungswert für den Gewinn von Spieler I beträgt also ungefähr $-0,27$ EUR.

Zu (b):

Zur ersten Frage:

Z: Gewinn von Spieler II

A in der ersten Durchführung:

$$\mathbf{P}\{Z = 20\} = \mathbf{P}(A) = \frac{12}{37}$$

A in der zweiten Durchführung $(A^{\complement}, A)$:

$$\mathbf{P}\{Z = 30\} = \mathbf{P}\left\{A^{\complement}\right\} \cdot \mathbf{P}\{A\} = \frac{25}{37} \cdot \frac{12}{37}$$

$$\left.\begin{array}{l} \text{10 EUR Verlust im ersten Spiel} \\ \text{20 EUR im zweiten Spiel gesetzt} \\ \quad \rightsquigarrow \text{ 60 EUR werden ausgezahlt} \end{array}\right\} \rightsquigarrow \text{ 30 EUR Gewinn}$$

A in der dritten Durchführung $(A^{\complement}, A^{\complement}, A)$:

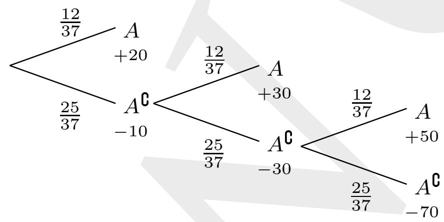
$$\mathbf{P}\{Z=50\} = \mathbf{P}\left\{A^{\complement}\right\} \cdot \mathbf{P}\left\{A^{\complement}\right\} \cdot \mathbf{P}\{A\} = \left(\frac{25}{37}\right)^2 \cdot \frac{12}{37}$$

$$\left.\begin{array}{l} \text{10 EUR Verlust im ersten Spiel} \\ \text{20 EUR Verlust im zweiten Spiel} \\ \text{40 EUR im dritten Spiel gesetzt} \\ \quad\rightsquigarrow \text{ 120 EUR werden ausgezahlt} \end{array}\right\} \rightsquigarrow \text{ 50 EUR Gewinn}$$

A in keiner Durchführung $(A^{\complement}, A^{\complement}, A^{\complement})$:

$$\mathbf{P}\{Z=-70\} = \mathbf{P}\left\{A^{\complement}\right\} \cdot \mathbf{P}\left\{A^{\complement}\right\} \cdot \mathbf{P}\left\{A^{\complement}\right\} = \left(\frac{25}{37}\right)^3$$

$$\left.\begin{array}{l} \text{10 EUR Verlust im ersten Spiel} \\ \text{20 EUR Verlust im zweiten Spiel} \\ \text{40 EUR Verlust im dritten Spiel} \end{array}\right\} \rightsquigarrow \text{ 70 EUR Verlust } \rightsquigarrow \left\{\begin{array}{l} \text{Kein weiteres Spiel} \\ \text{mehr möglich.} \end{array}\right.$$



$$\mathbf{E}(Z) = -70 \cdot \left(\frac{25}{37}\right)^3 + 50 \cdot \left(\frac{25}{37}\right)^2 \cdot \frac{12}{37} + 30 \cdot \frac{25}{37} \cdot \frac{12}{37} + 20 \cdot \frac{12}{37} \approx -1,129.$$

<u>Zur zweiten Frage</u>:

$$Z'\text{: Gewinn von Spieler } II$$

$$\begin{aligned} \mathbf{P}\{Z'=20\} &= \frac{12}{37} \\ \mathbf{P}\{Z'=30\} &= \frac{25}{37} \cdot \frac{12}{37} \\ \mathbf{P}\{Z'=50\} &= \left(\frac{25}{37}\right)^2 \cdot \frac{12}{37} \end{aligned}$$

A in der vierten Durchführung $(A^{\complement}, A^{\complement}, A^{\complement}, A)$:

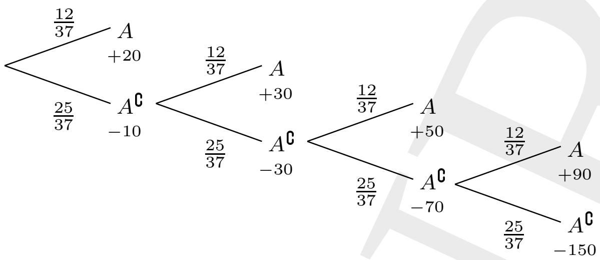
$$\mathbf{P}\{Z'=90\} = \mathbf{P}\left\{A^{\complement}\right\} \cdot \mathbf{P}\left\{A^{\complement}\right\} \cdot \mathbf{P}\left\{A^{\complement}\right\} \cdot \mathbf{P}\{A\} = \left(\frac{25}{37}\right)^3 \cdot \frac{12}{37}$$

$$\left.\begin{array}{l} \text{10 EUR Verlust im ersten Spiel} \\ \text{20 EUR Verlust im zweiten Spiel} \\ \text{40 EUR Verlust im zweiten Spiel} \\ \text{80 EUR im vierten Spiel gesetzt} \\ \quad\rightsquigarrow \text{ 240 EUR werden ausgezahlt} \end{array}\right\} \rightsquigarrow \text{ 90 EUR Gewinn}$$

A in keiner Durchführung $(A^{\complement}, A^{\complement}, A^{\complement}, A^{\complement})$:

$$\mathbf{P}\{Z'=0\} = \mathbf{P}\left\{A^{\complement}\right\} \cdot \mathbf{P}\left\{A^{\complement}\right\} \cdot \mathbf{P}\left\{A^{\complement}\right\} \cdot \mathbf{P}\left\{A^{\complement}\right\} = \left(\frac{25}{37}\right)^4$$

$$\left.\begin{array}{l}\text{10 EUR Verlust im ersten Spiel}\\ \text{20 EUR Verlust im zweiten Spiel}\\ \text{40 EUR Verlust im dritten Spiel}\\ \text{80 EUR Verlust im vierten Spiel}\end{array}\right\} \rightsquigarrow \text{150 EUR Verlust} \rightsquigarrow \left\{\begin{array}{l}\text{Kein weiteres Spiel}\\ \text{mehr möglich.}\end{array}\right.$$



$$\begin{aligned}\mathbf{E}(Z') &= -150\cdot\left(\frac{25}{37}\right)^4 + 90\cdot\left(\frac{25}{37}\right)^3\cdot\frac{12}{37} + 50\cdot\left(\frac{25}{37}\right)^2\cdot\frac{12}{37} + 30\cdot\frac{25}{37}\cdot\frac{12}{37} + 20\cdot\frac{12}{37}\\ &\approx -1{,}796.\end{aligned}$$

9.2 Aufgabe 2: Würfel

Zwei ideale Würfel W_1 und W_2 haben folgende Netze:

(a) Die Zufallsvariablen X_1 und X_2 beschreiben die Augenzahlen beim einmaligen Werfen mit dem W_1- beziehungsweise W_2-Würfel.
Bestimme den Erwartungswert und die Varianz von X_1 und X_2.
Die Zufallsvariable X beschreibe die Augensumme beim gleichzeitigen Wurf mit einem W_1- und einem W_2-Würfel.
Bestimme die Wahrscheinlichkeitsverteilung, den Erwartungswert und die Varianz von X.

(b) Die Zufallsvariable Y beschreibe die Augensumme beim gleichzeitigen Wurf mit zwei W_1- und zwei W_2-Würfeln.
Ermittle die Wahrscheinlichkeitsverteilung von Y.

(c) Die Zufallsvariable Z beschreibe die Augensumme beim gleichzeitigen Wurf mit 15 W_1- und 30 W_2-Würfeln.
Bestimme ein möglichst kleines, um $\mathbf{E}(Z)$ symmetrisches Intervall I, so daß Z mit einer Wahrscheinlichkeit von mindestens 95% einen Wert aus I annimmt. (Approximiere mit der Normalverteilung.)

(d) Mit Hilfe eines Computerprogramms soll ein W_1-Würfel simuliert werden.
Zur Überprüfung der Nullhypothese

$$H_0 : \text{„Der elektronische Würfel hat den Augenzahlerwartungswert } \frac{13}{6}\text{“}$$

macht man eine Stichprobe im Umfang von 6000 elektronischen „Würfen“ mit folgenden Ergebnis:

Augenzahl	1	2	3
Häufigkeit	2060	1008	2932

Kann danach H_0 mit einer Irrtumswahrscheinlichkeit von 5% verworfen werden? (Zweiseitiger Test)

Zu (a):

X_1: Augenzahl beim einmaligen Werfen mit einem W_1-Würfel

X_2: Augenzahl beim einmaligen Werfen mit einem W_2-Würfel

x_i	1	2	3
$\mathbf{P}\{X_1 = x_i\}$	$\frac{1}{3}$	$\frac{1}{6}$	$\frac{1}{2}$
$\mathbf{P}\{X_2 = x_i\}$	$\frac{1}{2}$	$\frac{1}{6}$	$\frac{1}{3}$

$$\begin{aligned}
\mathbf{E}(X_1) &= 1 \cdot \mathbf{P}\{X_1 = 1\} + 2 \cdot \mathbf{P}\{X_1 = 2\} + 3 \cdot \mathbf{P}\{X_1 = 3\} \\
&= 1 \cdot \frac{1}{3} + 2 \cdot \frac{1}{6} + 3 \cdot \frac{1}{2} \\
&= \frac{2+2+9}{6} = \frac{13}{6} = 2\frac{1}{6} \\
\mathbf{E}(X_2) &= 1 \cdot \mathbf{P}\{X_2 = 1\} + 2 \cdot \mathbf{P}\{X_2 = 2\} + 3 \cdot \mathbf{P}\{X_2 = 3\} \\
&= 1 \cdot \frac{1}{2} + 2 \cdot \frac{1}{6} + 3 \cdot \frac{1}{3} \\
&= \frac{3+2+6}{6} = \frac{11}{6} = 1\frac{5}{6}
\end{aligned}$$

Zur Erinnerung (Formel für die Berechnung der Varianz):

$$\begin{aligned}
\mathbf{V}(X) &= \mathbf{E}\left((X - \mathbf{E}(X))^2\right) \\
&= \mathbf{E}\left(X^2 - 2\mathbf{E}(X)X + (\mathbf{E}(X))^2\right) \\
&= \mathbf{E}\left(X^2\right) - \mathbf{E}(2\mathbf{E}(X)X) + \mathbf{E}\left((\mathbf{E}(X))^2\right) \\
&= \mathbf{E}\left(X^2\right) - 2\mathbf{E}(X)\mathbf{E}(X) + (\mathbf{E}(X))^2 \\
&= \mathbf{E}\left(X^2\right) - (\mathbf{E}(X))^2
\end{aligned}$$

Also:

$$\begin{aligned}
\mathbf{V}(X_1) &= \left(1 - \frac{13}{6}\right)^2 \cdot \mathbf{P}\{X_1 = 1\} + \left(2 - \frac{13}{6}\right)^2 \cdot \mathbf{P}\{X_1 = 2\} + \left(3 - \frac{13}{6}\right)^2 \cdot \mathbf{P}\{X_1 = 3\} \\
&= \left(-\frac{7}{6}\right)^2 \cdot \frac{1}{3} + \left(-\frac{1}{6}\right)^2 \cdot \frac{1}{6} + \left(\frac{5}{6}\right)^2 \cdot \frac{1}{2} \\
&= \frac{49}{36} \cdot \frac{1}{3} + \frac{1}{36} \cdot \frac{1}{6} + \frac{25}{36} \cdot \frac{1}{2} \\
&= \frac{1}{36} \cdot \left(\frac{98+1+75}{6}\right) = \frac{1}{36} \cdot \frac{174}{6} = \frac{29}{36}.
\end{aligned}$$

Oder:

$$\begin{aligned}
\mathbf{V}(X_1) &= \mathbf{E}\left(X_1^2\right) - (\mathbf{E}(X))^2 \\
&= 1 \cdot \frac{1}{3} + 4 \cdot \frac{1}{6} + 9 \cdot \frac{1}{2} - \left(\frac{13}{6}\right)^2 \\
&= \frac{2+4+27}{6} - \frac{169}{36} = \frac{198-169}{36} = \frac{29}{36}.
\end{aligned}$$

Und:

$$\begin{aligned}
\mathbf{V}(X_2) &= \left(1-\frac{11}{6}\right)^2\cdot\frac{1}{2}+\left(2-\frac{11}{6}\right)^2\cdot\frac{1}{6}+\left(3-\frac{11}{6}\right)^2\cdot\frac{1}{3}\\
&= \frac{25}{36}\cdot\frac{1}{2}+\frac{1}{36}\cdot\frac{1}{6}+\frac{49}{36}\cdot\frac{1}{3}\\
&= \frac{1}{36}\cdot\left(\frac{75+1+98}{6}\right)=\frac{1}{36}\cdot\frac{174}{6}=\frac{29}{36}.
\end{aligned}$$

X: Augensumme beim gleichzeitigen Wurf mit einem W_1- und einem W_2-Würfel

$$X := X_1 + X_2$$

Wahrscheinlichkeitsverteilung von X:

x_i	2	3	4	5	6
$\mathbf{P}\{X=x_i\}$	$\frac{1}{6}$	$\frac{5}{36}$	$\frac{7}{18}$	$\frac{5}{36}$	$\frac{1}{6}$

Denn:

$$\begin{aligned}
\mathbf{P}\{X=2\} &= \mathbf{P}\{X_1=1,X_2=1\}\overset{(*)}{=}\mathbf{P}\{X_1=1\}\cdot\mathbf{P}\{X_2=1\}\\
&= \frac{1}{3}\cdot\frac{1}{2}=\frac{1}{6}\\
\mathbf{P}\{X=3\} &= \mathbf{P}\{X_1=1,X_2=2\}+\mathbf{P}\{X_1=2,X_2=1\}\\
&\overset{(*)}{=} \mathbf{P}\{X_1=1\}\cdot\mathbf{P}\{X_2=2\}+\mathbf{P}\{X_1=2\}\cdot\mathbf{P}\{X_2=1\}\\
&= \frac{1}{3}\cdot\frac{1}{6}+\frac{1}{6}\cdot\frac{1}{2}=\frac{5}{36}\\
\mathbf{P}\{X=4\} &= \mathbf{P}\{X_1=1,X_2=3\}+\mathbf{P}\{X_1=2,X_2=2\}+\mathbf{P}\{X_1=3,X_2=1\}\\
&\overset{(*)}{=} \mathbf{P}\{X_1=1\}\cdot\mathbf{P}\{X_2=3\}+\mathbf{P}\{X_1=2\}\cdot\mathbf{P}\{X_2=2\}\\
&\qquad +\mathbf{P}\{X_1=3\}\cdot\mathbf{P}\{X_2=1\}\\
&= \frac{1}{3}\cdot\frac{1}{3}+\frac{1}{2}\cdot\frac{1}{2}+\frac{1}{6}\cdot\frac{1}{6}=\frac{7}{18}\\
\mathbf{P}\{X=5\} &= \mathbf{P}\{X_1=2,X_2=3\}+\mathbf{P}\{X_1=3,X_2=2\}\\
&\overset{(*)}{=} \mathbf{P}\{X_1=2\}\cdot\mathbf{P}\{X_2=3\}+\mathbf{P}\{X_1=3\}\cdot\mathbf{P}\{X_2=2\}\\
&= \frac{1}{6}\cdot\frac{1}{3}+\frac{1}{2}\cdot\frac{1}{6}=\frac{5}{36}\\
\mathbf{P}\{X=6\} &= \mathbf{P}\{X_1=3,X_2=3\}\overset{(*)}{=}\mathbf{P}\{X_1=3\}\cdot\mathbf{P}\{X_2=3\}\\
&= \frac{1}{3}\cdot\frac{1}{2}=\frac{1}{6}
\end{aligned}$$

Für den Erwartungswert von X ergibt sich:

$$\mathbf{E}(X)=\mathbf{E}(X_1+X_2)\overset{(*)}{=}\mathbf{E}(X_1)+\mathbf{E}(X_2)=\frac{13}{6}+\frac{11}{6}=\frac{24}{6}=4,$$

oder

$$\mathbf{E}(X)=2\cdot\frac{1}{6}+3\cdot\frac{5}{36}+4\cdot\frac{7}{18}+5\cdot\frac{5}{36}+6\cdot\frac{1}{6}=4$$

Für die Varianz folgt:

$$\begin{aligned}
\mathbf{V}(X) = \mathbf{V}(X_1 + X_2) &= \mathbf{E}\left((X_1 + X_2)^2\right) - (\mathbf{E}(X_1 + X_2))^2 \\
&\overset{(\#)}{=} \mathbf{E}\left(X_1^2\right) + 2 \cdot \mathbf{E}(X_1 \cdot X_2) + \mathbf{E}\left(X_2^2\right) - (\mathbf{E}(X_1) + \mathbf{E}(X_2))^2 \\
&\overset{(*)}{=} \mathbf{E}\left(X_1^2\right) + 2 \cdot \mathbf{E}(X_1) \cdot \mathbf{E}(X_2) + \mathbf{E}\left(X_2^2\right) \\
&\qquad - (\mathbf{E}(X_1))^2 - 2 \cdot \mathbf{E}(X_1) \cdot \mathbf{E}(X_2) - (\mathbf{E}(X_2))^2 \\
&= \mathbf{E}\left(X_1^2\right) - (\mathbf{E}(X_1))^2 + \mathbf{E}\left(X_2^2\right) - (\mathbf{E}(X_2))^2 \\
& \mathbf{V}(X_1) + \mathbf{V}(X_2) = \frac{29}{36} + \frac{29}{36} = \frac{29}{18}
\end{aligned}$$

Zu $(*)$: X_1 und X_2 sind unabhängig. Zu $(\#)$: Additivität des Erwartungswertes.

Zu (b):

Y: Augensumme beim gleichzeitigen Wurf mit zwei W_1- und zwei W_2-Würfeln

X_{11}: Augenzahl des ersten W_1-Würfels
X_{12}: Augenzahl des zweiten W_1-Würfels
X_{21}: Augenzahl des ersten W_2-Würfels
X_{22}: Augenzahl des zweiten W_2-Würfels

$$X' := X_{11} + X_{12} \qquad X'' := X_{12} + X_{22}$$

x''_j \ x'_i	2	3	4	5	6	$\mathbf{P}\left\{X'' = x''_j\right\}$
2	$\frac{36}{1296}$	$\frac{30}{1296}$	$\frac{84}{1296}$	$\frac{30}{1296}$	$\frac{36}{1296}$	$\frac{6}{36}$
3	$\frac{30}{1296}$	$\frac{25}{1296}$	$\frac{70}{1296}$	$\frac{25}{1296}$	$\frac{30}{1296}$	$\frac{5}{36}$
4	$\frac{84}{1296}$	$\frac{70}{1296}$	$\frac{196}{1296}$	$\frac{70}{1296}$	$\frac{84}{1296}$	$\frac{14}{36}$
5	$\frac{30}{1296}$	$\frac{25}{1296}$	$\frac{70}{1296}$	$\frac{25}{1296}$	$\frac{30}{1296}$	$\frac{5}{36}$
6	$\frac{36}{1296}$	$\frac{30}{1296}$	$\frac{84}{1296}$	$\frac{30}{1296}$	$\frac{36}{1296}$	$\frac{6}{36}$
$\mathbf{P}\left\{X' = x'_i\right\}$	$\frac{6}{36}$	$\frac{5}{36}$	$\frac{14}{36}$	$\frac{5}{36}$	$\frac{6}{36}$	

$$\mathbf{P}\left\{X' = x'_i, X'' = x''_j\right\} = \mathbf{P}\left\{X' = x'_i\right\} \cdot \mathbf{P}\left\{X'' = x''_j\right\}$$

Wahrscheinlichkeitsverteilung von Y $(Y = X' + X'')$:

y_i	4	5	6	7	8	9	10	11	12
$\mathbf{P}\{Y = y_i\}$	$\frac{36}{1296}$	$\frac{60}{1296}$	$\frac{193}{1296}$	$\frac{200}{1296}$	$\frac{318}{1296}$	$\frac{200}{1296}$	$\frac{193}{1296}$	$\frac{60}{1296}$	$\frac{36}{1296}$

Zu (c):

$$\begin{aligned}
\mathbf{E}(Z) &\overset{(\clubsuit)}{=} \mathbf{E}(Z_1) + \mathbf{E}(Z_2) \overset{(\clubsuit)}{=} 15 \cdot \mathbf{E}(W_1) + 30 \cdot \mathbf{E}(W_2) \\
&= 15 \cdot \frac{13}{6} + 30 \cdot \frac{11}{6} = \frac{5 \cdot 13}{2} + 5 \cdot 11 = \frac{65}{2} + 55 = 87\frac{1}{2}
\end{aligned}$$

$$\begin{aligned}\mathbf{V}(Z) &\overset{(\clubsuit)}{=} 15\cdot\mathbf{V}(W_1)+30\cdot\mathbf{V}(W_2)\\ &= 15\cdot\frac{29}{36}+30\cdot\frac{29}{36} = \frac{45\cdot 29}{36} = \frac{5\cdot 29}{4} = 36,25\end{aligned}$$

Zu ($\clubsuit$): Additivität des Erwartungswertes, W_1, W_2 sind unabhängig.

<u>Gesucht</u>: Kleinste Zahl k (mit zwei Nachkommastellen), so daß $\mathbf{P}\{x\in I\} \geq 0,95$, mit $I=[\mathbf{E}(Z)-k;\mathbf{E}(Z)+k]$.

<u>Approximation mit der Normalverteilung</u>:

$$\mu=\mathbf{E}=87,5 \qquad \sigma^2=\mathbf{V}=36,25$$

$$\begin{aligned}0,95 \leq \mathcal{N}_{87,5;36,25}\left([87,5-k;87,5+k]\right) &= \Phi\left(\frac{k}{\sqrt{36,25}}\right)-\Phi\left(\frac{-k}{\sqrt{36,25}}\right)\\ &= \Phi\left(\frac{k}{\sqrt{36,25}}\right)-\left(1-\Phi\left(\frac{k}{\sqrt{36,25}}\right)\right)\\ &= 2\cdot\Phi\left(\frac{k}{\sqrt{36,25}}\right)-1\end{aligned}$$

$$\begin{aligned}\Rightarrow\quad 0,95 \leq 2\cdot\Phi\left(\frac{k}{\sqrt{36,25}}\right)-1 \quad&\Rightarrow\quad 1,95 \leq 2\cdot\Phi\left(\frac{k}{\sqrt{36,25}}\right)\\ &\Rightarrow\quad 0,975 \leq \Phi\left(\frac{k}{\sqrt{36,25}}\right)\\ &\overset{(A.6.2.2)}{\Rightarrow}\quad 1,96 \leq \frac{k}{\sqrt{36,25}}\\ &\Rightarrow\quad k \geq 1,96\cdot\sqrt{36,25} \approx 11,80\end{aligned}$$

$$\Rightarrow\quad I = [87,5-k;87,5+k] = [75,7;99,3]$$

Zu (d):

$$H_0:\mathbf{E}_0 = \frac{13}{6} \approx 2,1667 \qquad \alpha = 0,05$$

$$\begin{aligned}\mathbf{E}_1 &= \frac{1\cdot 2060+2\cdot 1008+3\cdot 2932}{6000}\\ &= \frac{12872}{6000} \approx 2,1253\end{aligned}$$

<u>Gesucht sind</u>:

g_l, so daß $\mathbf{P}\{x\leq\mathbf{E}_0-g_l\}\leq\frac{\alpha}{2}=0,025$ und g_r, so daß $\mathbf{P}\{x\geq\mathbf{E}_0+g_r\}\leq\frac{\alpha}{2}=0,025$

Wir approximieren mit der Normalverteilung:

$$\begin{aligned}0,025 \geq \mathbf{P}\{x\leq\mathbf{E}_0-g_l\} &= \mathcal{N}_{\frac{13}{6};\frac{29}{36}}\left(]-\infty;\mathbf{E}_0-g_l]\right)\\ &= \Phi\left(\frac{\frac{13}{6}-\left(\frac{13}{6}-g_l\right)}{\sqrt{\frac{29}{36}}}\right) = \Phi\left(\frac{g_l\cdot 6}{\sqrt{26}}\right)\end{aligned}$$

Nach Tabelle $(A.6.2.2)$ folgt (in diesem Bereich ist die Funktion annähernd linear): $0,025=\Phi(0,51)$.

$$\begin{aligned}\Rightarrow\quad 0,51 \geq \frac{g_l\cdot 6}{\sqrt{29}} \quad&\Rightarrow\quad g_l \leq \frac{0,51\cdot\sqrt{29}}{6}\\ &\Rightarrow\quad g_l \lesssim 0,46\end{aligned}$$

Da $\mathbf{E}_1 \approx 2,1253$ im Intervall $[\mathbf{E}_0 - g_l; \mathbf{E}_0] = [1,7067; 2,1253]$ liegt, sind wir hier eigentlich schon fertig: Die Hypothese kann nicht verworfen werden.

Der Vollständigkeit halber berechnen wir jedoch noch g_r

$$0,025 \geq \mathbf{P}\{x \geq \mathbf{E}_0 + g_r\} = \mathcal{N}_{\frac{13}{6};\frac{29}{36}}\left([\mathbf{E}_0 + g_r; \infty[\right)$$

$$\begin{aligned}
\Rightarrow \quad 0,025 \geq 1 - \mathbf{P}\{x \leq \mathbf{E}_0 + g_r\} &= \mathcal{N}_{\frac{13}{6};\frac{29}{36}}\left(]-\infty; \mathbf{E}_0 + g_r]\right) \\
&= \Phi\left(\frac{\frac{13}{6} - \left(\frac{13}{6} + g_r\right)}{\sqrt{\frac{29}{36}}}\right) = \Phi\left(\frac{-g_r \cdot 6}{\sqrt{26}}\right) \\
&= 1 - \Phi\left(\frac{g_r \cdot 6}{\sqrt{26}}\right)
\end{aligned}$$

$$\Rightarrow \quad 0,975 \geq \Phi\left(\frac{g_r \cdot 6}{\sqrt{26}}\right) \overset{(A.6.2.2)}{\Rightarrow} 1,96 \geq \frac{g_r \cdot 6}{\sqrt{26}}$$

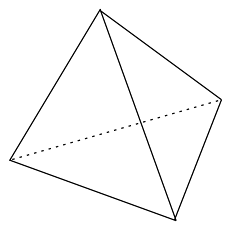
$$\Rightarrow \quad g_r \leq \frac{1,96 \cdot \sqrt{29}}{6} \quad \Rightarrow \quad g_r \lesssim 1,79$$

Annahmebereich:

$$[\mathbf{E}_0 - g_l; \mathbf{E}_0 + g_r] \approx [1,7067; 3,9153].$$

9.3 Aufgabe 3: Tetraeder-Würfel

Die vier Seitenflächen eines Tetraeders tragen die Zahlen 1, 2, 3 und 4. Die Augenzahl eines Wurfes ist die Zahl auf der Standfläche. Alle vier Seiten fallen mit der gleichen Wahrscheinlichkeit.



(a) Der Tetraeder wird zweimal geworfen. Die Zufallsvariable X beschreibe die Augensumme in den beiden Würfen. Ermittle die Wahrscheinlichkeitsverteilung von X. Berechne den Erwartungswert und die Standardabweichung von X.

(b) Zwei Spieler A und B verwenden den Tetraeder zu folgendem Spiel:
A zahlt zunächst 1 EUR als Einsatz an B. Dann wirft er den Tetraeder zweimal. Fällt in keinem der Würfe die „1“, so erhält A von B die Augensumme in EUR. Fällt mindestens einmal die „1“, so bekomt B nochmals 5 EUR von A. Zeige, daß das Spiel nicht fair ist. Wie müßte der Einsatz abgeändert werden, damit das Spiel fair ist?

(c) Der Tetraeder wird 4 mal geworfen. Berechne die Wahrscheinlichkeiten der Ereignisse:

- Jedesmal die „1“;
- Zweimal die „1“;
- Alle Augenzahlen verschieden;
- Mindestens drei gleiche Augenzahlen.

(d) Der Tetraeder wird so lange geworfen, bis die Augensumme mindestens 4 beträgt. Die Zufallsvariable Y beschreibe die Anzahl der hierzu benötigten Würfe. Ermittle die Wahrscheinlichkeitsverteilung von Y. Wie viele Würfe sind durchschnittlich erforderlich, bis die Augensumme mindestens 4 beträgt?

(e) Schätze mit Hilfe der Tschebyscheffschen Ungleichung ab, wie oft man den Tetraeder mindestens werfen muß, damit sich die relative Häufigkeit für die „1“ von der Wahrscheinlichkeit für „1“ mit einer Wahrscheinlichkeit von mindestens 90% um weniger als $0,01$ unterscheidet. Welche Abschätzung für die Anzahl der Würfe ergibt sich, wenn man die Binomialverteilung durch die Normalverteilung approximiert?

(f) Um zu prüfen, ob ein vorgelegter Tetraeder als ideal gelten kann, wurde er 2000 mal geworfen. Man erhielt das folgende Ergebnis:

Augenzahl	1	2	3	4
Absolute Häufigkeit	460	507	518	515

Kann man daraus mit einer Irrtumswahrscheinlichkeit von 5% schließen, daß der Tetraeder nicht ideal ist?

Zu (a):

X: Augensumme in den beiden Würfen
X': Augenzahl im ersten Wurf
X'': Augenzahl im zweiten Wurf

$x''_j \backslash x'_i$	1	2	3	4	$\mathbf{P}\left\{X'' = x''_j\right\}$
1	$\frac{1}{16}$	$\frac{1}{16}$	$\frac{1}{16}$	$\frac{1}{16}$	$\frac{1}{4}$
2	$\frac{1}{16}$	$\frac{1}{16}$	$\frac{1}{16}$	$\frac{1}{16}$	$\frac{1}{4}$
3	$\frac{1}{16}$	$\frac{1}{16}$	$\frac{1}{16}$	$\frac{1}{16}$	$\frac{1}{4}$
4	$\frac{1}{16}$	$\frac{1}{16}$	$\frac{1}{16}$	$\frac{1}{16}$	$\frac{1}{4}$
$\mathbf{P}\left\{X' = x'_i\right\}$	$\frac{1}{4}$	$\frac{1}{4}$	$\frac{1}{4}$	$\frac{1}{4}$	

Wahrscheinlichkeitsverteilung von X:

x_i	2	3	4	5	6	7	8
$\mathbf{P}\left\{X = x_i\right\}$	$\frac{1}{16}$	$\frac{1}{8}$	$\frac{3}{16}$	$\frac{1}{4}$	$\frac{3}{16}$	$\frac{1}{8}$	$\frac{1}{16}$

$$
\begin{aligned}
\mathbf{E}(X) &= 2\cdot\frac{1}{16} + 3\cdot\frac{1}{8} + 4\cdot\frac{3}{16} + 5\cdot\frac{1}{4} + 6\cdot\frac{3}{16} + 7\cdot\frac{1}{8} + 8\cdot\frac{1}{16} \;=\; 5 \\
\mathbf{V}(X) &= (2-5)^2\cdot\frac{1}{16} + (3-5)^2\cdot\frac{1}{8} + (4-5)^2\cdot\frac{3}{16} \\
&\qquad + (6-5)^2\cdot\frac{3}{16} + (7-5)^2\cdot\frac{1}{8} + (8-5)^2\cdot\frac{1}{16} \;=\; 2\frac{1}{2} \\
\sigma(X) &= \sqrt{\frac{5}{2}} \approx 1,58
\end{aligned}
$$

Zu (b):

$$\mathbf{P}\{\text{In keinem der Würfe eine „1“}\} = \frac{9}{16}$$

$$\mathbf{P}\{\text{Mindestens eine „1“}\} = 1 - \frac{9}{16} = \frac{7}{16}$$

Sei Z der Gewinn von Spieler A:

z_i	$+3$	$+4$	$+5$	$+6$	$+7$	-6
$\mathbf{P}\left\{Z = z_i\right\}$	$\frac{1}{16}$	$\frac{2}{16}$	$\frac{3}{16}$	$\frac{2}{16}$	$\frac{1}{16}$	$\frac{7}{16}$

$$\Rightarrow \quad \mathbf{E}(Z) = 3\cdot\frac{1}{16} + 4\cdot\frac{2}{16} + 5\cdot\frac{3}{16} + 6\cdot\frac{2}{16} + 7\cdot\frac{1}{16} - 6\cdot\frac{7}{16} = \frac{3}{16} > 0$$

Also ist das Spiel nicht fair!

Sei a der Einsatz:

$$\begin{aligned}
& (4-a)\cdot\frac{1}{16} + (5-a)\cdot\frac{2}{16} + (6-a)\cdot\frac{3}{16} + (7-a)\cdot\frac{2}{16} + (8-a)\cdot\frac{1}{16} \\
& \quad - (5+a)\cdot\frac{7}{16} \overset{!}{=} 0 \\
\Leftrightarrow \quad & \frac{4+10+18+14+8-35}{16} + a\cdot\frac{-1-2-3-2-1-7}{16} = 0 \\
\Leftrightarrow \quad & \frac{19}{16} + a\cdot(-1) = 0 \\
\Leftrightarrow \quad & a = \frac{19}{16} = 1,1875 \approx 1,19.
\end{aligned}$$

Der Einsatz müßte also $1,19$ EUR betragen, damit das Spiel fair ist. (Da A pro Spiel bei einem Einsatz von 1 EUR durchschnittlich $\frac{3}{16}$ EUR gewinnt, ist das Spiel bei einem Einsatz von $\left(1+\frac{3}{16}\right)$ EUR fair, das heißt $\mathbf{E}$(„Gewinn“) $= 0$.)

Zu (c): Viermaliges Werfen eines Tetraeders:

$$\Omega = \{1,2,3,4\}^4 \qquad \mathbf{P} = \mathcal{L}_\Omega \ \text{Gleichverteilung} \qquad \#\Omega = 4^4$$

$$\begin{aligned}
\mathbf{P}(A) &= \frac{1}{4^4} = \frac{1}{256} \approx 0,0039 \\
\mathbf{P}(B) &= \mathbf{P}\{\text{Genau zweimal die „1“}\} = \frac{\binom{4}{2}\cdot 3\cdot 3}{4^4} = \frac{54}{256} \approx 0,2109 \\
\mathbf{P}(C) &= \frac{4\cdot 3\cdot 2\cdot 1}{4^4} = \frac{24}{256} = 0,09375 \\
\mathbf{P}(D) &= \frac{4\cdot\binom{4}{3}\cdot 3 + 4}{4^4} = \frac{52}{256} \approx 0,2031.
\end{aligned}$$

Zu (d):

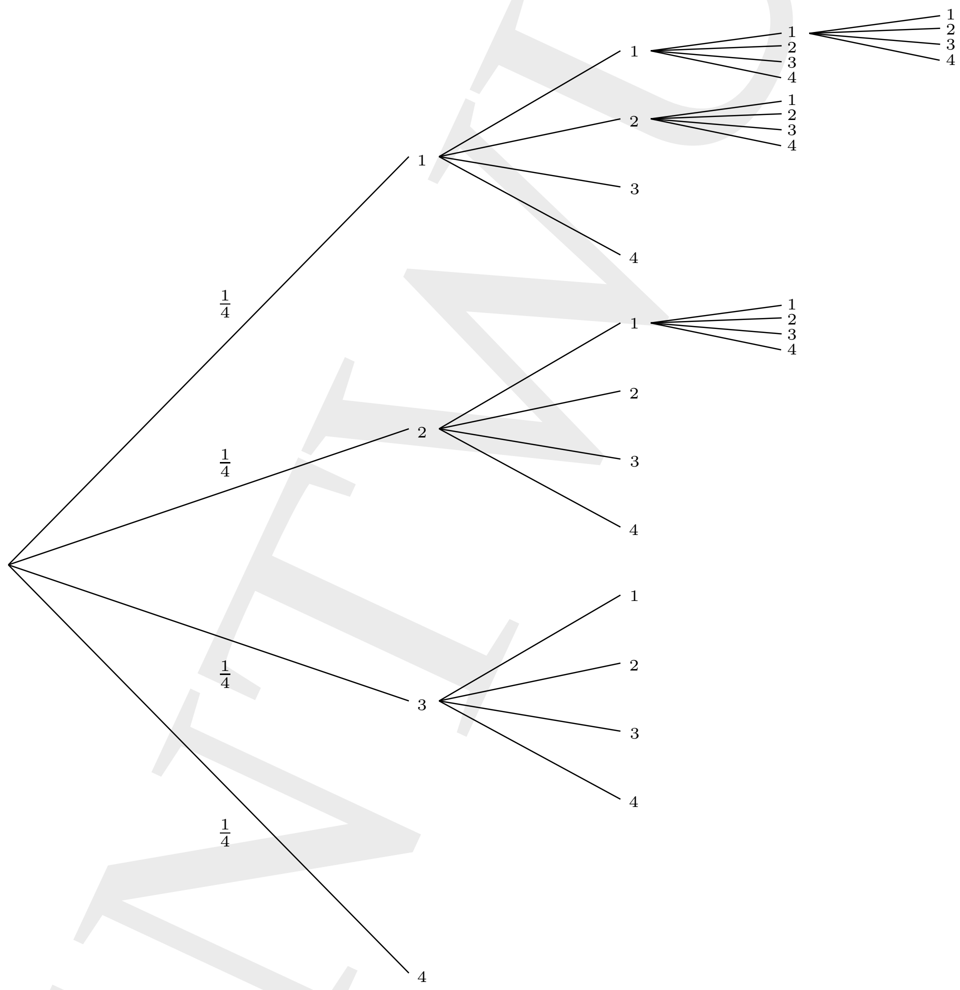
Y: Anzahl der benötigten Würfe, bis die Augensumme mindestens 4 beträgt.

<u>Wahrscheinlichkeitsverteilung von Y</u>:

y_i	1	2	3	4
$\mathbf{P}\{Y = y_i\}$	$\frac{1}{4}$	$\frac{9}{16}$	$\frac{11}{64}$	$\frac{1}{64}$

X_i: Ergebnis des i-ten Wurfs.

$$
\begin{aligned}
\mathbf{P}\{Y=1\} &= \mathbf{P}\{X_1=4\} = \frac{1}{4} \\
\mathbf{P}\{Y=2\} &= \mathbf{P}\{X_1=3\} + \mathbf{P}\{X_1=2; X_2=2\} + \mathbf{P}\{X_1=2; X_2=3\} \\
&\qquad + \mathbf{P}\{X_1=2; X_2=4\} + \mathbf{P}\{X_1=1; X_2=3\} + \mathbf{P}\{X_1=1; X_2=4\} \\
&= \frac{1}{4} + \frac{1}{4}\cdot\frac{3}{4} + \frac{1}{4}\cdot\frac{1}{2} = \frac{9}{16} \\
\mathbf{P}\{Y=3\} &= \mathbf{P}\{X_1=1; X_2=2\} + \mathbf{P}\{X_1=2; X_2=1\} \\
&\qquad + \mathbf{P}\{X_1=1; X_2=1; X_3=2\} + \mathbf{P}\{X_1=1; X_2=1; X_3=3\} \\
&\qquad + \mathbf{P}\{X_1=1; X_2=1; X_3=4\} \\
&= \frac{1}{4}\cdot\frac{1}{4}\cdot 1 + \frac{1}{4}\cdot\frac{1}{4}\cdot 1 + \frac{1}{4}\cdot\frac{1}{4}\cdot\frac{3}{4} = \frac{11}{64} \\
\mathbf{P}\{Y=4\} &= \mathbf{P}\{X_1=1; X_2=1; X_3=1\} = \frac{1}{4}\cdot\frac{1}{4}\cdot\frac{1}{4} = \frac{1}{64}
\end{aligned}
$$



$$
\Rightarrow \quad \mathbf{E}(Y) = 1\cdot\frac{1}{4} + 2\cdot\frac{9}{16} + 3\cdot\frac{11}{64} + 4\cdot\frac{1}{64} = 1\frac{61}{64}
$$

Also muß man durchschnittlich etwas weniger als zweimal würfeln.

Zu (e):

Für $\varepsilon = 0,01$ soll die Wahrscheinlichkeit für einen Abstand von weniger als ε zwischen $\overline{X}$ und $\mathbf{E}\left(\overline{X}\right)$ mindestens 90% betragen. Aus $(A.8.2)$ folgt:

$$\mathbf{P}\left\{\left|\overline{X} - \mathbf{E}\left(\overline{X}\right)\right| < \varepsilon\right\} = 1 - \mathbf{P}\left\{\left|\overline{X} - \mathbf{E}\left(\overline{X}\right)\right| \geq \varepsilon\right\} \geq 1 - \frac{1}{4 \cdot n \cdot \varepsilon^2} \overset{!}{\geq} 0,9.$$

$$\Rightarrow \quad n \geq \frac{1}{4 \cdot \varepsilon^2 \cdot 0,1} = \frac{10}{4 \cdot \varepsilon^2} = 25000.$$

Wenn man mit der Normalverteilung approximiert, folgt:

$$\mathbf{P}\left(\left|\frac{X}{n} - \frac{\mu}{n}\right| \leq 0,01\right) = \mathbf{P}\left(|X - \mu| \leq 0,01 \cdot n\right) \quad \Rightarrow \quad 2 \cdot \Phi\left(\frac{0,01 \cdot n}{\sigma}\right) - 1 \overset{!}{\geq} 0,9$$

und daraus:

$$n \geq 5071.$$

Zu (f):

X_i: Anzahl der Würfe, welche die Augenzahl i ergeben haben.

X_i ist $\mathcal{B}_{2000;\frac{1}{4}}$-verteilt, mit

$$\mu = n \cdot p = 2000 \cdot \frac{1}{4} = 500 \quad \text{und} \quad \sigma^2 = n \cdot p \cdot q = 2000 \cdot \frac{1}{4} \cdot \frac{3}{4} = 375 \quad \Rightarrow \quad \sigma \approx 19,4.$$

Es muß für jedes i getestet werden (Abschätzung mit der Tschebyscheff-Ungleichung):

$$\mathbf{P}\left(|X - \mu| \geq \varepsilon\right) \leq \frac{\sigma^2}{\varepsilon^2} \leq 0,025 \quad \Rightarrow \quad \varepsilon^2 \geq \frac{\sigma^2}{0,025} \quad \Rightarrow \quad \varepsilon \geq \frac{\sigma}{0,5} = 2 \cdot \sqrt{375} \approx 38,73.$$

Für den Annahmebereich folgt:

$$[\mu - \varepsilon; \mu + \varepsilon] \approx [461; 539].$$

Also müßte die Hypothese „Der Würfel ist ideal“ verworfen werden, da das Versuchsergebnis für $i = 1$ nicht im Annahmebereich liegt.

10. Bildverteilungen & Bedingte Wahrscheinlichkeiten

10.1 Bildverteilungen

Sei

$$X : (\Omega; \mathfrak{A}) \to (\Omega'; \mathfrak{A}')$$

eine Zufallsvariable, $\mathbf{P}$ ein Wahrscheinlichkeitsmaß auf $(\Omega, \mathfrak{A})$.
$\mathbf{P}_X$ definiert durch

$$\mathbf{P}_X(A') = \mathbf{P}\{X \in A'\} = \mathbf{P}\{\omega \in \Omega \mid X(\omega) \in A'\} \quad \forall A' \in \mathfrak{A}'$$

ist ein Wahrscheinlichkeitsmaß auf $(\Omega', \mathfrak{A}')$, die Verteilung von $\mathbf{P}$ unter X.
Experiment: Würfeln mit zwei idealen Würfeln.
Modell:

$$\Omega = \{1, \ldots, 6\}^2 = \{(\omega_1, \omega_2) : \omega_1, \omega_2 \in \{1, \ldots, 6\}\}$$
$$\mathfrak{A} = \mathcal{POT}(\Omega) \qquad \mathbf{P} = \mathcal{L}_\Omega \quad \text{(Gleichverteilung)}$$

Gesucht: Die Wahrscheinlichkeit, daß die Augensumme gleich k ist.
Methode: Bildverteilung:
Mögliche Werte für k:

$$k \in \{2, \ldots, 12\}$$

Es seien

$$S : \Omega \to \Omega' = \{2, \ldots, 12\} \qquad \mathfrak{A}' = \mathcal{POT}(\Omega)' \qquad S(\omega_1, \omega_2) = \omega_1 + \omega_2$$

S ist trivialerweise eine Zufallsvariable, da $\mathfrak{A} = \mathcal{POT}(\Omega)$.

$$\mathbf{P}' = \mathbf{P}_S \quad \underline{\text{Bildverteilung}}.$$

Sei $k \in \{2, \ldots, 12\}$:

$$\begin{aligned}
\mathbf{P}'(\{k\}) &= \mathbf{P}_S(\{k\}) = \mathbf{P}\{S \in \{k\}\} = \mathbf{P}\{S = k\} \\
&= \mathbf{P}\{(\omega_1, \omega_2) \in \Omega \mid S(\omega_1, \omega_2) = k\} = \mathbf{P}\{(\omega_1, \omega_2) \in \Omega \mid \omega_1 + \omega_2 = k\}
\end{aligned}$$

$$\begin{aligned}
\mathbf{P}'(\{2\}) &= \mathbf{P}\{(\omega_1, \omega_2) \in \Omega \mid \omega_1 + \omega_2 = 2\} \\
&= \mathbf{P}\{(1,1)\} = \frac{1}{36} \\
\mathbf{P}'(\{3\}) &= \mathbf{P}\{(\omega_1, \omega_2) \in \Omega \mid \omega_1 + \omega_2 = 3\} \\
&= \mathbf{P}\{(1,2), (2,1)\} = \frac{2}{36} \\
\mathbf{P}'(\{4\}) &= \mathbf{P}\{(\omega_1, \omega_2) \in \Omega \mid \omega_1 + \omega_2 = 4\} \\
&= \mathbf{P}\{(1,3), (2,2), (3,1)\} = \frac{3}{36} \\
\mathbf{P}'(\{5\}) &= \mathbf{P}\{(\omega_1, \omega_2) \in \Omega \mid \omega_1 + \omega_2 = 5\} \\
&= \mathbf{P}\{(1,4), (2,3), (3,2), (4,1)\} = \frac{4}{36} \\
&\vdots
\end{aligned}$$

Wahrscheinlichkeitsverteilung von S:

k	2	3	4	5	6	7	8	9	10	11	12
$\mathbf{P}'\{k\} = \mathbf{P}\{S = k\}$	$\frac{1}{36}$	$\frac{1}{18}$	$\frac{1}{12}$	$\frac{1}{9}$	$\frac{5}{36}$	$\frac{1}{6}$	$\frac{5}{36}$	$\frac{1}{9}$	$\frac{1}{12}$	$\frac{1}{18}$	$\frac{1}{36}$

Beachte: $\mathbf{P}'$ ist keine Laplace-Verteilung auf Ω'!

M sei das Minimum der gewürfelten Zahlen:

$$M : \Omega \to \Omega'' = \{1, \ldots, 6\} \qquad M(\omega_1, \omega_2) = \min\{\omega_1, \omega_2\},$$

$$\mathbf{P}'' = \mathbf{P}_M \quad \underline{\text{Bildverteilung}}.$$

Angabe der Zähldichte von $\mathbf{P}_M$: Sei $l \in \{1, \ldots, 6\}$:

$$\begin{aligned}
\mathbf{P}_M(\{l\}) &= \mathbf{P}\{(\omega_1, \omega_2) \in \Omega \mid \min\{\omega_1, \omega_2\} = l\} \\
\mathbf{P}_M(\{1\}) &= \mathbf{P}\{(1,1), (1,2), (1,3), (1,4), (1,5), (1,6), \\
&\qquad (2,1), (3,1), (4,1), (5,1), (6,1)\} = \frac{11}{36} \\
\mathbf{P}_M(\{2\}) &= \mathbf{P}\{(2,2), (2,3), (2,4), (2,5), (2,6), (3,2), (4,2), (5,2), (6,2)\} = \frac{9}{36} \\
\mathbf{P}_M(\{3\}) &= \mathbf{P}\{(3,3), (3,4), (3,5), (3,6), (4,3), (5,3), (6,3)\} = \frac{7}{36} \\
\mathbf{P}_M(\{4\}) &= \mathbf{P}\{(4,4), (4,5), (4,6), (5,4), (6,4)\} = \frac{5}{36} \\
\mathbf{P}_M(\{5\}) &= \mathbf{P}\{(5,5), (5,6), (6,5)\} = \frac{3}{36} \\
\mathbf{P}_M(\{6\}) &= \mathbf{P}\{(6,6)\} = \frac{1}{36}
\end{aligned}$$

$$(S, M) : \Omega \to \Omega' \times \Omega'' = \{2, \ldots, 12\} \times \{1, \ldots, 6\}$$

Verteilung von (S, M):

$$\begin{aligned}
\mathbf{P}_{(S,M)}(\{2,1\}) &= \mathbf{P}\{(S, M) = (2,1)\} \\
&= \mathbf{P}\{(\omega_1, \omega_2) \in \Omega \mid \omega_1 + \omega_2 = 2, \min\{\omega_1, \omega_2\} = 1\} \\
&= \mathbf{P}\{(1,1)\} = \frac{1}{36} \\
\mathbf{P}_{(S,M)}(\{3,1\}) &= \mathbf{P}\{(S, M) = (3,1)\} \\
&= \mathbf{P}\{(\omega_1, \omega_2) \in \Omega \mid \omega_1 + \omega_2 = 3, \min\{\omega_1, \omega_2\} = 1\} \\
&= \mathbf{P}\{(2,1), (1,2)\} = \frac{2}{36} \\
\mathbf{P}_{(S,M)}(\{4,1\}) &= \mathbf{P}\{(S, M) = (4,1)\} \\
&= \mathbf{P}\{(\omega_1, \omega_2) \in \Omega \mid \omega_1 + \omega_2 = 4, \min\{\omega_1, \omega_2\} = 1\} \\
&= \mathbf{P}\{(3,1), (1,3)\} = \frac{2}{36} \\
&\ \vdots \\
\mathbf{P}_{(S,M)}(\{6,2\}) &= \mathbf{P}\{(S, M) = (6,2)\} \\
&= \mathbf{P}\{(\omega_1, \omega_2) \in \Omega \mid \omega_1 + \omega_2 = 6, \min\{\omega_1, \omega_2\} = 2\} \\
&= \mathbf{P}\{(4,2), (2,4)\} = \frac{2}{36} \\
&\ \vdots
\end{aligned}$$

Gemeinsame Verteilung von S und M ($\mathbf{P}\{S=k, M=l\}$):

$k \backslash l$	1	2	3	4	5	6	$\mathbf{P}\{S=k\}$
2	$\frac{1}{36}$	0	0	0	0	0	$\frac{1}{36}$
3	$\frac{2}{16}$	0	0	0	0	0	$\frac{2}{36}$
4	$\frac{2}{36}$	$\frac{1}{36}$	0	0	0	0	$\frac{3}{36}$
5	$\frac{2}{36}$	$\frac{2}{36}$	0	0	0	0	$\frac{4}{36}$
6	$\frac{2}{36}$	$\frac{2}{36}$	$\frac{1}{36}$	0	0	0	$\frac{5}{36}$
7	$\frac{2}{16}$	$\frac{2}{36}$	$\frac{2}{36}$	$\frac{2}{36}$	0	0	$\frac{6}{36}$
8	0	$\frac{2}{36}$	$\frac{2}{36}$	$\frac{1}{36}$	0	0	$\frac{5}{36}$
9	0	0	$\frac{2}{36}$	$\frac{2}{36}$	0	0	$\frac{4}{36}$
10	0	0	0	$\frac{2}{36}$	$\frac{1}{36}$	0	$\frac{3}{36}$
11	0	0	0	0	$\frac{2}{36}$	0	$\frac{2}{36}$
12	0	0	0	0	0	$\frac{1}{36}$	$\frac{1}{36}$
$\mathbf{P}\{M=l\}$	$\frac{11}{36}$	$\frac{9}{36}$	$\frac{7}{36}$	$\frac{5}{36}$	$\frac{3}{36}$	$\frac{1}{36}$	

Das Ereignis $S=6$ ist zum Beispiel:

$$\begin{aligned}
\{S=6\} &= \{(\omega_1,\omega_2) \mid \omega_1+\omega_2=6\} \\
&= \{(\omega_1,\omega_2) \mid \omega_1+\omega_2=6,\ \min\{\omega_1,\omega_2\}=1\} \\
&\qquad \uplus \{(\omega_1,\omega_2) \mid \omega_1+\omega_2=6,\ \min\{\omega_1,\omega_2\}=2\} \\
&\qquad \uplus \{(\omega_1,\omega_2) \mid \omega_1+\omega_2=6,\ \min\{\omega_1,\omega_2\}=3\} \\
&\qquad \uplus \{(\omega_1,\omega_2) \mid \omega_1+\omega_2=6,\ \min\{\omega_1,\omega_2\}=4\} \\
&\qquad \uplus \{(\omega_1,\omega_2) \mid \omega_1+\omega_2=6,\ \min\{\omega_1,\omega_2\}=5\}
\end{aligned}$$

Beachte: S und M sind nicht unabhängig

Zum Beispiel:

$$\mathbf{P}\{S=2, M=1\} = \frac{1}{36} \neq \frac{1}{36}\cdot\frac{11}{36} = \mathbf{P}\{S=2\}\cdot\mathbf{P}\{M=1\}$$

10.2 Bedingte Wahrscheinlichkeiten

$$(\Omega, \mathfrak{A}, \mathbf{P}) \quad \text{Wahrscheinlichkeitsraum} \qquad A, B \in \mathfrak{A} \text{ mit } \mathbf{P}(B) > 0$$

$$\mathbf{P}(A \mid B) = \frac{\mathbf{P}(A \cap B)}{\mathbf{P}(B)}$$

Bedingte Wahrscheinlichkeit von A unter B.

10.2.1 Beispiel

Die Fernseh- und Lesegewohnheiten einer Bevölkerung werden statistisch untersucht. Aufgrund dieser Erhebung kann man davon ausgehen, daß 95% dieser Bevölkerung fernsehen; 89% der Personen, die fernsehen, lesen auch Zeitung und 97% der Personen, die nicht fernsehen, lesen aber Zeitung.
Mit welcher Wahrscheinlichkeit sieht eine Person dieser Bevölkerung

(a) fern und liest Zeitung;

(b) nicht fern, liest aber Zeitung?

$$A = \{\text{„Person sieht fern“}\} \qquad B = \{\text{„Person liest Zeitung“}\}$$

Gegeben:

$$\mathbf{P}(A) = 0,95 \qquad \mathbf{P}(B \mid A) = 0,89 \qquad \mathbf{P}\left(B \,\middle|\, A^{\complement}\right) = 0,97$$

Zu (a): Gesucht: $\mathbf{P}(A \cap B)$

$$\mathbf{P}(A \cap B) = \mathbf{P}(B \cap A) = \mathbf{P}(B \mid A) \cdot \mathbf{P}(A) = 0,89 \cdot 0,95 = 0,8455.$$

Zu (b): Gesucht: $\mathbf{P}\left(A^{\complement} \cap B\right)$

$$\mathbf{P}\left(A^{\complement} \cap B\right) = \mathbf{P}\left(B \cap A^{\complement}\right) = \mathbf{P}\left(B \,\middle|\, A^{\complement}\right) \cdot \mathbf{P}\left(A^{\complement}\right) = 0,05 \cdot 0,97 = 0,0485$$

Zusatz: Mit welcher Wahrscheinlichkeit liest eine Person dieser Bevölkerung Zeitung?

$$\begin{aligned}\mathbf{P}(B) &= \mathbf{P}(A) \cdot \mathbf{P}(B \mid A) + \mathbf{P}\left(A^{\complement}\right) \cdot \mathbf{P}\left(B \,\middle|\, A^{\complement}\right) \\ &= 0,95 \cdot 0,89 + 0,05 \cdot 0,97 \;=\; 0,8455 + 0,0485 \;=\; 0,894\end{aligned}$$

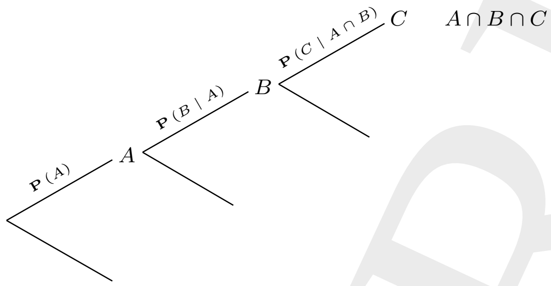
10.2.2 Formel von der totalen Wahrscheinlichkeit/ Formel von *Bayes*

Es seien $B_1, B_2, \ldots$ disjunkte Ereignisse mit $A \subseteq \bigcup_{n=1}^{\infty} B_n$. Dann gilt:

$$\mathbf{P}(A) = \sum_{n=1}^{\infty} \mathbf{P}(B_n) \cdot \mathbf{P}(A \mid B_n).$$

Diese Formel ist auch im Fall $\mathbf{P}(B_n) = 0$ gültig, wenn man $0 \cdot \text{undefiniert} = 0$ setzt.

$$\begin{aligned}\mathbf{P}(A \cap B) &= \mathbf{P}(A) \cdot \mathbf{P}(B \mid A) \\ \mathbf{P}(A \cap B \cap C) &= \mathbf{P}((A \cap B) \cap C) \\ &= \mathbf{P}(A \cap B) \cdot \mathbf{P}(C \mid A \cap B) \\ &= \mathbf{P}(A) \cdot \mathbf{P}(B \mid A) \cdot \mathbf{P}(C \mid A \cap B)\end{aligned}$$



10.2.3 Beispiel

Urne mit 5 weißen und 6 roten Kugeln. 3-maliges Ziehen ohne Zurücklegen.
Gesucht: Wahrscheinlichkeit, daß alle drei Kugeln weiß sind.
1. Lösung:

$$X\text{: Anzahl der weißen Kugeln}$$

X ist hypergeometrisch verteilt:

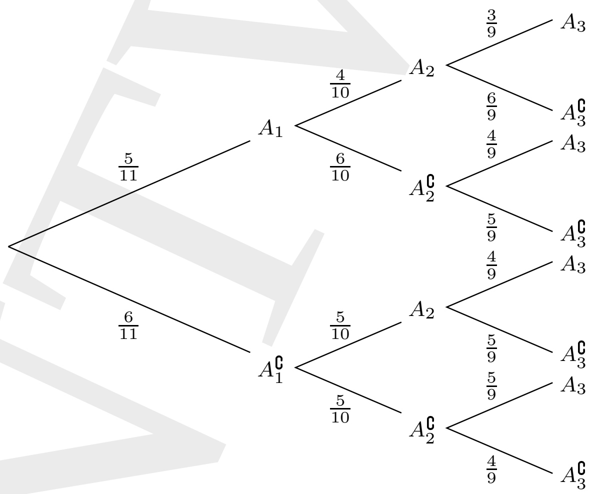
$$\mathbf{P}\{X = 3\} = \frac{\binom{5}{3} \cdot \binom{6}{0}}{\binom{11}{3}} = \frac{10}{165} = \frac{2}{33}.$$

2. Lösung:

$$A_i = \{\text{Im } i\text{-ten Zug eine weiße Kugel}\}$$

Gesucht:

$$\mathbf{P}(A_1 \cap A_2 \cap A_3).$$



$$\begin{aligned} \mathbf{P}(A_1 \cap A_2 \cap A_3) &= \mathbf{P}(A_1) \cdot \mathbf{P}(A_2 \mid A_1) \cdot \mathbf{P}(A_3 \mid A_1 \cap A_2) \\ &= \frac{5}{11} \cdot \frac{4}{10} \cdot \frac{3}{9} = \frac{2}{33} \end{aligned}$$

10.2.4 Beispiel

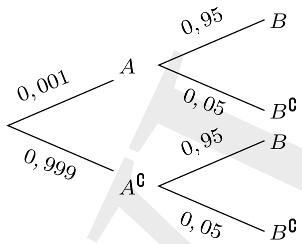
In einer Bevölkerung sind durchschnittlich $0,1\%$ aller Personen Tbc-krank. Ein medizinischer Test zur Tbc-Erkennung zeigt in 95% aller Fälle eine vorligende Erkrankung an, bei Gesunden zeigt der Test in 4% aller Fälle aber irrtümlich eine Erkrankung an. Aus der Bevölkerung wird eine Person zufällig ausgewählt. Mit welcher Wahrscheinlichkeit reagiert sie bei dem Test positiv?

$$A = \{\text{Person ist Tbc-krank}\} \qquad B = \{\text{Test ist positiv}\}$$

<u>Gegeben</u>:

$$\mathbf{P}(A) = 0,001 \qquad \mathbf{P}(B \mid A) = 0,95 \qquad \mathbf{P}\left(B \,\middle|\, A^{\complement}\right) = 0,04$$

<u>Baummodell</u>:



<u>Gesucht</u>: $\mathbf{P}(B)$

$$\begin{aligned}\mathbf{P}(B) &= \mathbf{P}(A)\cdot\mathbf{P}(B \mid A) + \mathbf{P}\left(A^{\complement}\right)\cdot\mathbf{P}\left(B \,\middle|\, A^{\complement}\right)\\ &= 0,001\cdot 0,95 + 0,999\cdot 0,04 \approx 0,04\end{aligned}$$

Das heißt in rund 4% aller Fälle ist der Test positiv. Der Prozentsatz der tatsächlichen Tbc-Kranken ist mit $0,1\%$ viel geringer. Die unterschiedlichen Werte beruhen auf den Fehlerquellen des Tests.

Eine Person reagiere auf den Test positiv. Mit welcher Wahrscheinlichkeit hat sie tatsächlich Tbc?

<u>Gesucht</u>: $\mathbf{P}(A \mid B)$

<u>Formel von *Bayes*</u>:

Im Fall $\mathbf{P}(B) > 0$ gilt:

$$\mathbf{P}(A_k \mid B) = \frac{\mathbf{P}(A_k)\cdot\mathbf{P}(B \mid A_k)}{\sum\limits_{n=1}^{\infty}\mathbf{P}(A_n)\cdot\mathbf{P}(B \mid A_n)}.$$

Damit ergibt sich:

$$\begin{aligned}\mathbf{P}(A \mid B) &= \frac{\mathbf{P}(A)\cdot\mathbf{P}(B \mid A)}{\underbrace{\mathbf{P}(A)\cdot\mathbf{P}(B \mid A) + \mathbf{P}\left(A^{\complement}\right)\cdot\mathbf{P}\left(B \,\middle|\, A^{\complement}\right)}_{\mathbf{P}(B)}}\\ &= \frac{0,001\cdot 0,95}{0,001\cdot 0,95 + 0,999\cdot 0,04} \approx 0,023.\end{aligned}$$

Durch die Fehlerquellen des Tests bedingt, haben überaschenderweise nur rund 2% der Personen, die beim Test positiv reagieren, auch tatsächlich Tbc.

10.2.5 Beispiel: Ziegenspiel

Vergleiche ersten Abschnitt auf Seite A-1: Bei einer Fernsehshow darf ein Kandidat zwischen drei Türen wählen, hinter einer steht ein Auto, hinter den beiden anderen Ziegen. Nach seiner Wahl öffnet der Quizmaster (der weiß, wo der Gewinn steht) eine Tür, hinter der eine Ziege steht.

Soll der Kandidat wechseln?

$$\begin{array}{rcl} A & \widehat{=} & \{\text{Auto hinter der zuerst gewählten Tür}\} \\ A' & \widehat{=} & \{\text{Ziege hinter der zuerst gewählten Tür}\} \end{array} \quad \text{mit} \quad \mathbf{P}(A) = \frac{1}{3} \quad \mathbf{P}(A') = \frac{2}{3}$$

$$B \;\widehat{=}\; \{\text{Auto hinter der nicht gewählten Tür}\}$$

$$\mathbf{P}(B \mid A) = 0 \qquad \mathbf{P}(B \mid A') = 1$$

$$\Rightarrow \quad \mathbf{P}(B) = \mathbf{P}(A) \cdot \mathbf{P}(B \mid A) + \mathbf{P}(A') \cdot \mathbf{P}(B \mid A') = \frac{1}{3} \cdot 0 + \frac{2}{3} \cdot 1 = \frac{2}{3}$$

Also sollte er wechseln, da das Auto mit einer Wahrscheinlichkeit von $\frac{2}{3}$ hinter der nicht gewählten Tür steht.

Anhang B: Klausuren

H-Doz. Dr. H.P. Scheffler | Sommersemester 1998

Klausur zur Vorlesung Stochastik I

08.07.1998

Name	
Vorname	
Geburtsdatum	
Matrikelnummer	

Studiengang		Leistungsnachweis		Qual. Stud. Nachweis	

Für Leistungsnachweise: Bearbeiten Sie alle 4 Aufgaben. Die Bearbeitungszeit beträgt 3 Stunden.

Für qualifizierte Studiennachweise: Wählen Sie 3 Aufgaben aus. Die Bearbeitungszeit beträgt 2 Stunden und 15 Minuten.
Bitte kreuzen Sie die bearbeiteten Aufgaben an. Die Bearbeitung anderer Aufgaben(teile) wird nicht bewertet.

Nummer	1	2	3	4	$\sum$
bearbeitet					
Punkte					

Jede Aufgabe soll auf einem gesonderten Blatt bearbeitet werden. Versehen Sie jedes Blatt mit Ihrem Namen und der Nummer der Aufgabe.

Achten Sie in der Klausur auf sorgfältige und exakte Formulierungen. Geben Sie, falls gefordert, Definitionen und Sätze stets exakt (mit allen Voraussetzungen) an. Es ist nicht ausreichend, Definitionen durch Schlagworte, durch Plausibilitätsbetrachtungen oder durch verbale Be- und Umschreibung des Sachverhalts zu ersetzen

Bei der Bearbeitung der einzelnen Aufgaben genügt es, entsprechende Formeln abzuleiten, exakte numerische Berechnungen, etwa mit Hilfe eines Taschenrechners sind nicht erforderlich.

Hilfsmittel (Bücher, Skripten) sind nicht zugelassen!

Aufgabe 1: $x\mathbf{P} + x\mathbf{P} + x\mathbf{P} + x\mathbf{P} = x\mathbf{P}$

Es wird mit drei (idealen) Würfeln geworfen. Bei jedem Wurf werden die Augenzahlen (i, j, k), $1 \leq i, j, k \leq 6$ notiert.

S_n respektive M_n bezeichnen die *Summe* respektive das *Minimum* der Augen beim n-ten Wurf mit den drei Würfeln.

(a) Bestimmen Sie die Wahrscheinlichkeit bei einem Wurf

- **(i)** die Augensumme $S_n = 16$, bzw.
- **(ii)** das Minimum $M_n = 5$ zu erhalten.
  (Kontrolle: $\mathbf{P}\{S_n = 16\} = \frac{1}{36}$, $\mathbf{P}\{M_n = 5\} = \frac{7}{6^3}$)

(b) **(i)** Geben Sie die Wahrscheinlichkeit $\mathbf{P}(\{M_n = 5\} \cap \{S_n = 16\})$ an.

- **(ii)** Sind (bei festem n) M_n und S_n unabhängig?

(c) Wie groß ist die Wahrscheinlichkeit unter n Würfen mit drei Würfeln mindestens einmal die Augensumme $S_i = 16$ zu erziehlen? $(1 \leq i \leq n)$

(d) Geben Sie die Definition der Unabhängigkeit

- **(i)** von Ereignissen,
- **(ii)** von Zufallsvariablen an.

Aufgabe 2: $x\mathbf{P} + x\mathbf{P} + x\mathbf{P} + x\mathbf{P} = x\mathbf{P}$

Für ein $x_0 > 0$ fest und $\alpha > 0$ sei

$$f(x) = c \cdot x^{-\alpha-1} 1_{]x_0,\infty[}(x)$$

für ein $c > 0$.

(a) Bestimmen Sie $c > 0$ so, daß f die Lebesgue-Dichte eines Wahrscheinlichkeitsmaßes μ auf $\mathbb{R}$ ist.

(b) Berechnen Sie die Verteilungsfunktion von μ.

(c) Es sei X eine Zufallsvariable mit Verteilung μ.

- **(i)** Bestimmen Sie die $\alpha > 0$ für die der Erwartungswert von X existiert und berechnen Sie diesen.
- **(ii)** Bestimmen Sie die $\alpha > 0$ für die die Varianz von X existiert und berechnen Sie diese.

(d) Definieren Sie

- **(i)** den Erwartungswert
- **(ii)** die Varianz

einer reellwertigen Zufallsvariablen X.

Aufgabe 3: $x\mathbf{P} + x\mathbf{P} + x\mathbf{P} + x\mathbf{P} = x\mathbf{P}$

In einem Friseur-Laden mit 9 Friseuren dauert ein Haarschnitt 10 Minuten. Herr Snyder betritt den Laden und sieht, daß alle 9 Friseure beschäftigt sind.

Die Zeitpunkte $T_1, \ldots, T_9$ des Fertigwerdens der Friseure $1, \ldots, 9$ seien in $[0, 10]$ gleichverteilt und unabhängig.

Es bezeichne T die Wartezeit von Herrn Snyder bis er bedient werden kann.

(a) **(i)** Geben Sie eine Formel für T an.

(ii) Zeigen Sie, daß die Verteilungsfunktion F_T von T die Gestalt

$$F_T(t) = \begin{cases} 0 & \text{für } t < 0 \\ 1 - \left(1 - \dfrac{t}{10}\right)^9 & \text{für } 0 \leq t \leq 10 \\ 1 & \text{für } t > 10 \end{cases}$$

hat.

(b) Berechnen die Dichte der Verteilung von T.

(c) Berechnen Sie die *mittlere Wartezeit* $\mathbf{E}(T)$ von Herrn Snyder.

(d) Definieren Sie die Begriffe

(i) Verteilungsfunktion einer Zufallsvariable;

(ii) Lebesgue-Dichte einer Verteilung.

Aufgabe 4: $x\mathbf{P} + x\mathbf{P} + x\mathbf{P} + x\mathbf{P} = x\mathbf{P}$

Es sei $X : \Omega \to \mathbb{R}_+$ eine Zufallsvariable. Für festes $s > 0$ heißt die Funktion

$$F_s(t) = \mathbf{P}\{X \leq t \,|X > s\} \quad \text{für } s < t$$

die Restlebensdauer von X.

(a) Zeigen Sie, daß F_s eine Verteilungsfunktion eines geeigneten Wahrscheinlichkeitsmaßes ist und geben Sie dieses an.

(b) Es sei nun X exponentialverteilt mit Parameter $\lambda > 0$ (d.h. die Verteilung von X hat die Dichte $x \mapsto \lambda \mathbf{e}^{-\lambda x} 1_{]0,\infty[}(x)$.) Berechnen Sie

- **(i)** $F_s(t)$
- **(ii)** $\mathbf{P}\left\{\frac{1}{\lambda} < X < \frac{2}{\lambda} \middle| X > \frac{1}{\lambda}\right\}$

(c) Es sei nun X eine reelle Zufallsvariable auf $\mathbb{R}_+$, deren Verteilungsfunktion F stetig auf $\mathbb{R}_+$ ist und $1 - F(s) > 0$. Zeigen Sie, daß dann äquivalent sind:

- **(i)** Für alle $0 < s < t$ gilt $F_s(t+s) = F(t)$
- **(ii)** F ist die Verteilungsfunktion einer exponential verteilten Zufallsvariable.

(Hinweis: Betrachten Sie $G(t) = 1 - F(t)$ und verwenden Sie, daß jede stetige Lösung der Funktionalgleichung $G(t) \cdot G(s) = G(t+s)$ für $s, t > 0$ von der Form $G(t) = \mathbf{e}^{bt}$ für ein geeignetes $b \in \mathbb{R}$ ist.)

(d)
- **(i)** Definieren Sie die bedingte Wahrscheinlichkeit von A unter B.
- **(ii)** Formulieren Sie den Satz von der totalen Wahrscheinlichkeit.

H-Doz. Dr. H.P. Scheffler Sommersemester 1998

Nachklausur zur Vorlesung Stochastik I

05.10.1998

Name	
Vorname	
Geburtsdatum	
Matrikelnummer	

Studiengang		Leistungsnachweis		Qual. Stud. Nachweis	

Für Leistungsnachweise: Bearbeiten Sie alle 6 Aufgaben. Die Bearbeitungszeit beträgt 3 Stunden. Zum Erwerb eines Leistungsnachweises müssen Sie mindestens 22 Punkte erreichen.

Für qualifizierte Studiennachweise: Wählen Sie von den Aufgaben 1 und 2 eine Aufgabe aus und von den Aufgaben 3 bis 6 wählen Sie 3 Aufgaben aus. Die Bearbeitungszeit beträgt 2 Stunden und 15 Minuten. Zum Erwerb eines qualifizierten Studiennachweises müssen Sie mindestens 15 Punkte erreichen.
Bitte kreuzen Sie die bearbeiteten Aufgaben an. Die Bearbeitung anderer Aufgaben(teile) wird nicht bewertet.

Nummer	1	2	3	4	5	6	$\sum$
bearbeitet							
Punkte							

Jede Aufgabe soll auf einem gesonderten Blatt bearbeitet werden. Versehen Sie jedes Blatt mit Ihrem Namen und der Nummer der Aufgabe.

Achten Sie in der Klausur auf sorgfältige und exakte Formulierungen. Geben Sie, falls gefordert, Definitionen und Sätze stets exakt (mit allen Voraussetzungen) an. Es ist nicht ausreichend, Definitionen durch Schlagworte, durch Plausibilitätsbetrachtungen oder durch verbale Be- und Umschreibung des Sachverhalts zu ersetzen

Bei der Bearbeitung der einzelnen Aufgaben genügt es, entsprechende Formeln abzuleiten, exakte numerische Berechnungen, etwa mit Hilfe eines Taschenrechners sind nicht erforderlich.

Hilfsmittel (Bücher, Skripten) sind nicht zugelassen!

Aufgabe 1: $x\mathbf{P} + x\mathbf{P} + x\mathbf{P} = 6\mathbf{P}$

Die 4 Seiten eines tetraederförmigen "Würfels" sind jeweils mit den Zahlen 1, 2, 3, 4 beschriftet und treten beim Würfeln mit gleicher Wahrscheinlichkeit auf. Der Tetraeder wird zweimal geworfen und die untenliegenden Zahlen X_1 und X_2 notiert.

(a) Bestimmen Sie die Zähldichte der Verteilung $\mathbf{P}_{X_1+X_2}$ von $X_1 + X_2$.

(b) Berechnen Sie den Erwartungswert von $X_1 + X_2$.

(c) Definieren Sie

- **(i)** die Zähldichte einer Verteilung auf einer abzählbaren Menge.
- **(ii)** den Erwartungswert einer diskreten Zufallsvariablen.

Aufgabe 2: $x\mathbf{P} + x\mathbf{P} + x\mathbf{P} + x\mathbf{P} = 6\mathbf{P}$

Ablenkeinheiten für Fernsehröhren werden einer sorgfältigen Endkontrolle unterzogen. Der automatisierte Kontrollvorgang weist folgende statistische Parameter auf: Die Wahrscheinlichkeit dafür, daß eine fehlerfreie Einheit auch als fehlerfrei erkannt wird, ist $0,98$; die Wahrscheinlichkeit dafür, daß eine defekte Einheit auch als defekt erkannt wird ist $0,95$. Die Wahrscheinlichkeit, daß eine Einheit defekt ist, beträgt $0,08$.
Berechnen Sie die Wahrscheinlichkeiten folgender Ereignisse:

(a) Eine durch die Kontrolle als fehlerfrei deklarierte Ablenkeinheit ist tatsächlich fehlerfrei.

(b) Eine durch die Kontrolle als defekt deklarierte Ablenkeinheit ist tatsächlich defekt.

(c) Eine durch die Kontrolle als fehlerfrei deklarierte Ablenkeinheit ist in Wirklichkeit defekt.

(d) **(i)** Definieren Sie die bedingte Wahrscheinlichkeit.

(ii) Geben Sie die Bayes'sche Formel an.

Aufgabe 3: $x\mathbf{P} + x\mathbf{P} + x\mathbf{P} + x\mathbf{P} = 10\mathbf{P}$

Bei einem Multiple-Choice-Test mit 10 Aufgaben sind pro Aufgabe 4 Antworten möglich, von denen genau eine richtig ist. Ein Kandidat, der schlecht vorbreitet ist, kreuzt die Antworten zufällig an.

(a) Bestimmen Sie die Wahrscheinlichkeit

- **(i)** alle Antworten richtig
- **(ii)** alle Antworten falsch

zu beantworten.

(b) Berechnen Sie die Wahrscheinlichkeit, daß mindestens 8 Aufgaben richtig sind. (Stellen sie das Ergebnis als Bruch dar.)

(c) Wie groß ist die Wahrscheinlichkeit, daß bei 100 schlecht vorbereiteten Teilnehmern am Test mindestens eine komplett richtige Lösung dabei ist?

(d) Definieren Sie

- **(i)** die Gleichverteilung auf einer endlichen Menge.
- **(ii)** die Binomialverteilung mit Parametern n, p. (Es reicht den zugehörigen Wahrscheinlichkeitsraum und die Zähldichte anzugeben.)

Aufgabe 4: $x\mathbf{P} + x\mathbf{P} + x\mathbf{P} + x\mathbf{P} = 10\mathbf{P}$

Auf den Strecken zwischen den Punkten $(1, 0)$ und $(1, 1)$ sowie zwischen $(2, 0)$ und $(2, 1)$ in der Ebene wird jeweils ein Punkt zufällig ausgewählt.

(a) Geben Sie ein geeignetes mathematisches Modell (unter der Annahme, daß die beiden Punkte unabhängig voneinander ausgewählt werden) an.

(b) Es bezeichne X_1 bzw. X_2 den zufällig ausgewählten Punkt als Element von $[0, 1]$. Zeigen Sie, daß X_1 und X_2 unabhängig sind.

(c) Berechnen Sie die Wahrscheinlichkeit dafür, daß der Schnitt der Verbindungsgerade dieser zwei Punkte mit der x-Achse rechts vom Nullpunkt liegt.

(d) Definieren Sie

(a) die kontinuierliche Gleichverteilung.

(b) die Unabhängigkeit einer Familie von Zufallsvariablen.

Aufgabe 5: $x\mathbf{P} + x\mathbf{P} + x\mathbf{P} + x\mathbf{P} = 10\mathbf{P}$

$X_i : \Omega \to \mathbb{R}_+$ seien Zufallsvariablen für $1 \leq i \leq n$, die nach einer Exponentialverteilung mit Parameter $\alpha > 0$ verteilt sind. Die $(X_i)_{i=1,\ldots,n}$ seien unabhängig. Sei $M_n = \min\limits_{1 \leq i \leq n} X_i$.

(a) Bestimmen Sie die Verteilungsfunktion F_{M_n} sowie die Dichte der Verteilung von M_n.

(b) Verifizieren Sie:

(i) $\mathbf{E}(X_i) = \dfrac{1}{\alpha}$

(ii) $\mathbf{E}\left(X_i^2\right) = \dfrac{2}{\alpha^2}$

und berechnen Sie Erwartungswert und Varianz von M_n.

(c) Es sei $\varepsilon > 0$. Schätzen Sie $\mathbf{P}\left\{\left|n \cdot M_n - \dfrac{1}{\alpha}\right| \geq \varepsilon\right\}$ mit der Tschebyscheff-Ungleichung ab.

(d) **(i)** Definieren Sie die Verteilungsfunktion einer Zufallsvariable.

(ii) Formulieren Sie die Tschebyscheff-Ungleichung.

Aufgabe 6: $x\mathbf{P} + x\mathbf{P} + x\mathbf{P} + x\mathbf{P} = 10\mathbf{P}$

Es sei $F(x) = \exp(-\exp(-x))$, $x \in \mathbb{R}$.

(a) Zeigen Sie, daß

- **(i)** F eine Verteilungsfunktion ist.
- **(ii)** die zugehörige Verteilung eine Dichte besitzt und geben Sie diese Dichte an.

(b) Es sei X eine Zufallsvariable mit Verteilungsfunktion $F_X = F$. Sei $Y = \mathbf{e}^X$. Geben Sie die Verteilungsfunktion von Y an und geben Sie - falls sie existiert - die Dichte an.

(c) X_1, X_2 seien unabhängige Zufallsvariablen mit Verteilungsfunktion F. Für $b \in \mathbb{R}$ sei $Y = \max(X_1 - b, X_2 - b)$. Geben Sie die Verteilungsfunktion von Y an. Wie muß $b \in \mathbb{R}$ gewählt werden, damit $F_Y = F$ gilt.

(d) Definieren Sie

- **(i)** den Begriff der Dichte einer Verteilung.
- **(ii)** den Begriff einer reellwertigen Zufallsvariablen.

H-Doz. Dr. H.P. Scheffler Sommersemester 2000

Klausur zur Vorlesung Stochastik I

11.07.2000

Name	

Matrikel-Nr.	

Wählen Sie aus den folgenden sechs Aufgaben fünf Aufgaben aus. Die maximal erreichbare Punktzahl finden Sie neben jeder Aufgabe. Tragen Sie die Nummer der gewählten Aufgaben in das folgende Kästchen ein:

Gewählte Aufgabe					

Die Bearbeitung anderer Aufgaben(teile) wird nicht bewertet.

Für einen qualifizierten Studiennachweis benötigen Sie mindestens $18 \pm \varepsilon$ Punkte. Für einen Leistungsnachweis sind mindestens $24 \pm \varepsilon$ Punkte notwendig. Sie haben zum Bearbeiten der Aufgaben 3 Stunden Zeit.

Bitte verwenden Sie für jede Aufgabe ein gesondertes Blatt, welches Sie mit Ihrem Namen und Matrikelnummer versehen.

Achten Sie in der Klausur auf sorgfältige und exakte Formulierungen. Geben Sie, falls gefordert, Definitionen oder Sätze stets exakt (mit allen Voraussetzungen) an.

Bei der Bearbeitung der einzelnen Aufgaben genügt es, entsprechende Formeln abzuleiten, exakte numerische Berechnungen, etwa mit Hilfe eines Taschenrechners sind nicht erforderlich.

Hilfsmittel (Bücher, Skripten) sind nicht zugelassen!

Aufgabe 1: 5P + 2P + 1P = 8P

Auf einer Party mit n Mathematikern wird eine Tombola veranstaltet. Zu gewinnen sind (der Gastgeber ist großzügig) $k \geq n$ Geschenke (in Form von Mathematikbüchern).

(a) Mit welcher Wahrscheinlichkeit bekommt jeder der Mathematiker mindestens ein Geschenk? (Es reicht eine Formel abzuleiten.)

(b) Geben Sie die Siebformel an.

(c) Definieren Sie die Gleichverteilung auf einer endlichen Menge.

Aufgabe 2: $3\mathrm{P} + 3\mathrm{P} + 2\mathrm{P} = 8\mathrm{P}$

(a) Ein gut durchgemischter Stapel mit je 18 roten und schwarzen Karten wird in zwei gleichgroße Stapel geteilt. Wie groß ist die Wahrscheinlichkeit, daß sich in den kleineren Stapeln wiederum je gleichviele rote und schwarze Karten befinden? (Geben Sie auch ein geeignetes mathematisches Modell an.)

(b) Um die *mediale* Begabung herauszufinden, stellt eine okkultistische Gesellschaft einer Versammlung von 500 Menschen die Aufgabe, das Ergebnis eines Versuches zu erraten. Hinter einem Wandschirm wird eine Münze 10-mal geworfen. Bei jedem Wurf ist die Wahrscheinlichkeit für Kopf oder Zahl jeweils $0,5$. Das Versuchsergebnis soll von den Zuschauern geraten werden. Als medial begabt gilt, wer höchstens einen Fehler in der Vorhersage macht. Wie groß ist die Wahrscheinlichkeit, daß sich in der Versammlung mindestens ein *medial begabter* Zuschauer befindet?

(c) Definieren Sie und geben Sie die Mächtigkeiten an von

- **(i)** $k-$Permutationen mit Wiederholung einer $n-$elementigen Menge.
- **(ii)** $k-$Permutationen ohne Wiederholung einer $n-$elementigen Menge.

Aufgabe 3: 3P + 3P + 2P = 8P

Wie groß ist die Wahrscheinlichkeit, daß die quadratische Gleichung $x^2 + px + q = 0$ genau zwei reelle Lösungen besitzt? Dabei sei der Zufallsvektor $X = (p, q)$

(a) in $[0, 1]^2$ gleichverteilt;

(b) in $[-1, 1]^2$ gleichverteilt.

(c) Definieren Sie:

- **(i)** die Gleichverteilung auf einer Teilmenge des $\mathbb{R}^d$,
- **(ii)** eine Verteilung mit Lebesgue-Dichte auf $\mathbb{R}^d$.

Aufgabe 4: 2P + 4P + 2P = 8P

(a) Ein Student, der an einem *wahr - falsch* Test teilnimmt, verfährt bei der Beantwortung der Fragen folgendermaßen: Sofern er die Antwort weiß, kreuzt er diese an, andernfalls wirft er eine Münze um sich für wahr oder falsch zu entscheiden. Die Wahrscheinlichkeit, daß dem Studenten die richtige Antwort bekannt ist, sei $\frac{3}{5}$. Wie groß ist dann die Wahrscheinlichkeit, daß er eine richtig markierte Antwort wußte?

(b) Bei der Übertragung der Zeichen *Punkt* und *Strich* in einem Fernmeldesystem werden durch Störungen im Mittel 5% der gesendeten Punkte als Striche und 3% der gesendeten Striche als Punkte empfangen. Das Verhältnis von gesendeten Punkten zu gesendeten Strichen sei $\frac{3}{4}$.

(i) Wie groß ist die Wahrscheinlichkeit, daß ein gesendetes Zeichen richtig empfangen wird?

(ii) Es wird *Punkt* empfangen. Wie groß ist die Wahrscheinlichkeit, daß auch *Punkt* gesendet wurde?

(c) Formulieren Sie den Satz von der totalen Wahrscheinlichkeit und die Bayes'sche Formel.

Aufgabe 5: 2P + 3P + 2P + 2P + 3P = 12P

Die Lebensdauer X eines Systems sei gleichverteilt mit der Verteilungsfunktion

$$F : x \mapsto \begin{cases} 1 - \mathbf{e}^{-\alpha x^{\beta}} & \text{falls } x > 0 \\ 0 & \text{falls } x \leq 0 \end{cases}$$

Dabei sind $\alpha, \beta > 0$ fest.

(a) Zeigen Sie, daß die Verteilung von X eine Lebesgue-Dichte besitzt und bestimmen Sie diese.

(b) Bestimmen Sie für $\beta = 1$ und $\beta = 2$ den Erwartungswert $\mathbf{E}(X)$.

(c) Bestimmen Sie die Verteilungsfunktion der Zufallsvariable $Y = X^{\beta}$.

(d) Bestimmen Sie die Verteilungsfunktion der Restlebensdauer, das heißt

$$F_h(x) := \mathbf{P}\{X \leq x + h \,|X > h\}$$

für ein festes $h > 0$.

(e) Sei $a > 0$. Weiter seien X_1, X_2 unabhängige Zufallsvariablen mit Verteilungsfunktion F. Geben Sie die Verteilungsfunktion von $Z = \min(aX_1, aX_2)$ an. Für welche a ist $F_Z = F$?

Aufgabe 6: $5\mathbf{P} + 5\mathbf{P} + 2\mathbf{P} = 12\mathbf{P}$

Eine Fabrik stellt zylindrische Dosen mit Durchmesser $2r_0$ und Höhe h_0 her. Produktionsbedingt schwanken die wahren Werte (r, h) um den Sollwert (r_0, h_0). Also ist $X = (r, h)$ als $\mathbb{R}^2$-wertige Zufallsvariable aufzufassen. Ihre Verteilung besitze die Dichte $f(r, h)$.

(a) Geben Sie eine Formel für den Erwartungswert und die Varianz der Oberfläche Fl und des Volumens Vol an. (Oberfläche $= 2\pi rh + 2\pi r^2$, Volumen $= \pi r^2 h$.)

(b) Berechnen Sie $\mathbf{E}(Fl)$ und $\mathbf{E}(Vol)$ explizit für eine Dichte der Gestalt

$$f(r, h) = \frac{1}{4\varepsilon\delta} \cdot 1_{[r_0-\varepsilon, r_0+\varepsilon]}(r) \cdot 1_{[h_0-\delta, h_0+\delta]}(h)$$

wobei $\varepsilon, \delta > 0$ sind.

(c) Geben Sie die Transformationsformel für $\mathbf{E}(g(X))$ an für den Spezialfall, daß die Verteilung von X eine Lebesgue-Dichte f besitzt.

H-Doz. Dr. H.P. Scheffler | Sommersemester 2000

Nachklausur zur Vorlesung Stochastik I

Name			Matrikel-Nr.	

Wählen Sie aus den folgenden sechs Aufgaben fünf Aufgaben aus. Die maximal erreichbare Punktzahl finden Sie neben jeder Aufgabe. Tragen Sie die Nummer der gewählten Aufgaben in das folgende Kästchen ein:

Gewählte Aufgabe					

Die Bearbeitung anderer Aufgaben(teile) wird nicht bewertet.

Für einen qualifizierten Studiennachweis benötigen Sie mindestens $16 \pm \varepsilon$ Punkte. Für einen Leistungsnachweis sind mindestens $22 \pm \varepsilon$ Punkte notwendig. Sie haben zum Bearbeiten der Aufgaben 3 Stunden Zeit.

Bitte verwenden Sie für jede Aufgabe ein gesondertes Blatt, welches Sie mit Ihrem Namen und Matrikelnummer versehen.

Achten Sie in der Klausur auf sorgfältige und exakte Formulierungen. Geben Sie, falls gefordert, Definitionen oder Sätze stets exakt (mit allen Voraussetzungen) an.

Bei der Bearbeitung der einzelnen Aufgaben genügt es, entsprechende Formeln abzuleiten, exakte numerische Berechnungen, etwa mit Hilfe eines Taschenrechners sind nicht erforderlich.

Hilfsmittel (Bücher, Skripten) sind nicht zugelassen!

Aufgabe 1: $2\mathbf{P} + 4\mathbf{P} + 2\mathbf{P} = 8\mathbf{P}$

Bei der Glücksspirale der Olympialotterie 1971 wurden die 7-ziffrigen Gewinnzahlen auf folgende Weise ermittelt: Aus einer Trommel, welche je 7 Kugeln mit den Ziffern $0, \ldots, 9$ enthielt, wurden nach Durchmischen 7 Kugeln entnommen und deren Ziffern in der Reihenfolge des Ziehens zu einer Zahl angeordnet.

(a) Geben Sie ein geeignetes mathemathisches Modell an.

(b) Zeigen Sie, daß die Gewinnwahrscheinlichkeiten der einzelnen (gleich teuren!) Lose verschieden sind. (Berechnen Sie die Wahrscheinlichkeit, daß $0, 0, 0, 0, 0, 0, 0$ sowie daß $1, 2, 3, 4, 5, 6, 7$ gezogen wurde und vergleichen Sie diese.)

(c) Definieren Sie und geben Sie die Mächtigkeiten an von

- **(i)** $k-$Permutationen mit Wiederholung einer $n-$elementigen Menge.
- **(ii)** $k-$Permutationen ohne Wiederholung einer $n-$elementigen Menge.

Aufgabe 2: 3P + 3P + 2P = 8P

Ein Elektronikfachmarkt wird von drei Lieferanten L_i $(i = 1, 2, 3)$ mit Disketten beliefert. Die Lieferanten liefern 20%, 30% bzw. 50% des Bedarfs. Erfahrungsgemäß sind unter 1000 gelieferten Disketten des Lieferanten L_i 4, 3 bzw. 20 Stück defekt.

(a) Wie groß ist die Wahrscheinlichkeit, daß ein Kunde in dem Fachmarkt eine defekte Diskette kauft?

(b) Ein Kunde beschwert sich, daß er eine defekte Diskette gekauft hat. Wie groß ist die Wahrscheinlichkeit daß diese Diskette aus der Lieferung des Lieferanten L_3 stammt?

(c) Geben Sie

(i) den Satz von der totalen Wahrscheinlichkeit und

(ii) die Bayes'sche Formel an.

Aufgabe 3: $1\mathbf{P} + 2\mathbf{P} + 3\mathbf{P} + 2\mathbf{P} = 8\mathbf{P}$

Die logistische Verteilung ist durch die Verteilungsfunktion

$$F : x \mapsto \frac{1}{1 + \mathbf{e}^{-x}}$$

definiert.

(a) Zeigen Sie, daß F eine Verteilungsfunktion ist.

(b) Zeigen Sie, daß diese Verteilung eine Dichte f besitzt und bestimmen Sie diese.

(c) Zeigen Sie: Ist X eine reelle Zufallsvariable mit Verteilungsfunktion F, so gilt für alle $h > 0$, daß

$$\mathbf{P}\{X \geq h\} = \mathbf{P}\{X \leq -h\}$$

(Man nennt diese Verteilung symmetrisch um 0.)

(d) Definieren Sie:

(i) Varianz einer Zufallsvariablen;

(ii) Lebesgue-Dichte einer Verteilung.

Aufgabe 4: 1P + 3P + 4P + 2P = 10P

Die Länge X eines Telefongesprächs sei eine $\mathbb{R}_+$-wertige Zufallsvariable, deren Verteilung die Lebesgue-Dichte

$$f(x) = C \cdot \mathbf{e}^{-2x} \cdot 1_{[0,\infty)}(x)$$

besitzt.

(a) Bestimmen Sie die Konstante C. (Kontrolle: $C = 2$.)

(b) Geben Sie die

- **(i)** Verteilungsfunktion von X an;
- **(ii)** *mittlere Länge* $\mathbf{E}(X)$ eines Telefonats an.

(c) Die Kosten eines Telefonats der (deterministischen) Länge x berechnen sich nach der Formel

$$K(x) = 1_{(0,3]}(x) + (1 + (x - 3)) \cdot 1_{(3,\infty)}(x)$$

(Sockelbetrag + linearer Anstieg)
Berechnen Sie die *mittleren Kosten* $\mathbf{E}(K(X))$ eines Telefonats.

(d) Formulieren Sie die Transformationsformel für eine Zufallsvariable deren Verteilung die Lebesgue-Dichte f besitzt.

Aufgabe 5: 2P + 4P + 2P + 2P = 10P

Es sei X eine auf $[-1, 1]$ gleichverteilte Zufallsvariable.

(a) Zeigen Sie, daß X und $Y = X^2$ unkorreliert sind.

(b) Zeigen Sie, daß die Zufallsvariablen $Y_k := \sin(\pi k X)$, $k \geq 1$ paarweise unkorreliert sind.

(c) Zeigen Sie, daß

$$\frac{1}{n}\sum_{k=1}^{n} \sin(\pi k X) \to 0 \quad \text{stochastisch für } n \to \infty$$

gilt.

(d) Formulieren Sie das schwache Gesetz der großen Zahlen.

Aufgabe 6: 3P + 1P + 2P + 2P = 8P

In einem Friseur-Laden mit 9 Friseuren dauert ein Haarschnitt 10 Minuten. Herr Snyder betritt den Laden und sieht, daß alle 9 Friseure beschäftigt sind.

Die Zeitpunkte $T_1, \ldots, T_9$ des Fertigwerdens der Friseure $1, \ldots, 9$ seien in $[0, 10]$ gleichverteilt und unabhängig.

Es bezeichne T die Wartezeit von Herrn Snyder bis er bedient werden kann.

(a) **(i)** Geben Sie eine Formel für T an.

(ii) Zeigen Sie, daß die Verteilungsfunktion F_T von T die Gestalt

$$F_T(t) = \begin{cases} 0 & \text{für } t < 0 \\ 1 - \left(1 - \dfrac{t}{10}\right)^9 & \text{für } 0 \leq t \leq 10 \\ 1 & \text{für } t > 10 \end{cases}$$

hat.

(b) Berechnen die Dichte der Verteilung von T.

(c) Berechnen Sie die *mittlere Wartezeit* $\mathbf{E}(T)$ von Herrn Snyder.

(d) Definieren Sie die Begriffe

(i) Verteilungsfunktion einer Zufallsvariable;

(ii) Unabhängigkeit einer Folge von Zufallsvariablen.

H-Doz. Dr. H.P. Scheffler Sommersemester 2001

Klausur zur Vorlesung Stochastik I

11.07.2001

Name	

Matrikel-Nr.	

Wählen Sie aus den folgenden fünf Aufgaben vier Aufgaben aus. Die maximal erreichbare Punktzahl finden Sie neben jeder Aufgabe. Tragen Sie die Nummer der gewählten Aufgaben in das folgende Kästchen ein:

Gewählte Aufgabe				

Die Bearbeitung anderer Aufgaben(teile) wird nicht bewertet.

Für einen qualifizierten Studiennachweis benötigen Sie mindestens $18 \pm \varepsilon$ Punkte. Für einen Leistungsnachweis sind mindestens $24 \pm \varepsilon$ Punkte notwendig. Sie haben zum Bearbeiten der Aufgaben 3 Stunden Zeit.

Bitte verwenden Sie für jede Aufgabe ein gesondertes Blatt, welches Sie mit Ihrem Namen und Matrikelnummer versehen.

Achten Sie in der Klausur auf sorgfältige und exakte Formulierungen. Geben Sie, falls gefordert, Definitionen oder Sätze stets exakt (mit allen Voraussetzungen) an.

Bei der Bearbeitung der einzelnen Aufgaben genügt es, entsprechende Formeln abzuleiten, exakte numerische Berechnungen, etwa mit Hilfe eines Taschenrechners sind nicht erforderlich.

Hilfsmittel (Bücher, Skripten) sind nicht zugelassen!

$\varepsilon = +2$

Aufgabe 1: $4\mathbf{P} + 6\mathbf{P} + 2\mathbf{P} = 12\mathbf{P}$

(a) Von 100 befragten Studenten hören 30 die Vorlesung A, 25 die Vorlesung B, und 12 die Vorlesung C. 11 Studenten hören die Vorlesungen A und B, 6 Studenten die Vorlesungen A und C, 5 Studenten die Vorlesung B und C und 3 Studenten hören alle drei Vorlesungen. Wie viele der 100 Befragten hören keine der drei Vorlesungen?

(b) Jede Packung einer bestimmten Cornflakesart enthalte eines von insgesamt 11 verschiedenen Sammelfotos des BVB. Es werden $k \geq 11$ Packungen gekauft. Wie groß ist die Wahrscheinlichkeit, daß mindestens 1 Sammelfoto an der vollständigen Serie fehlt?

(c) Geben Sie die Siebformel an.

Aufgabe 2: 3P + 7P + 2P = 12P

(a) Einem Kasten, der 16 weiße und 16 schwarze Schachfiguren enthält, werden nacheinander 3 Figuren zufällig und ohne Zurücklegen entnommen. Wie groß ist die Wahrscheinlichkeit, hierbei 3 schwarze Figuren zu bekommen?

(b) In einem Aufzug, der noch 6 Stockwerke fährt, sind 4 Personen, die voneinander unabhängig aussteigen. Wie groß ist die Wahrscheinlichkeit, daß

- **(i)** alle in verschiedenen Stockwerken,
- **(ii)** zwei in einem und die anderen beiden in jeweils verschiedenen Stockwerken,
- **(iii)** alle 4 im gleichen Stockwerk,
- **(iv)** mindestens 3 im gleichen Stockwerk aussteigen?

(c) Definieren Sie und geben Sie die Mächtigkeiten an von

- **(i)** $k-$Permutationen mit Wiederholung einer $n-$elementigen Menge.
- **(ii)** $k-$Permutationen ohne Wiederholung einer $n-$elementigen Menge.

Aufgabe 3: $4\mathbf{P} + 6\mathbf{P} + 2\mathbf{P} = 12\mathbf{P}$

(a) Von den Übungsaufgaben zu einer bestimmten Vorlesung werden 20% vom Professor, 50% vom Assistenten und 30% von einer Hilfskraft gestellt. Von den Aufgaben des Professors können Sie erfahrungsgemäß 80%, von denjenigen des Assistenten 60% und von denjenigen der Hilfskraft 50% lösen. Wieder einmal können Sie eine Übungsaufgabe nicht lösen. Mit welcher Wahrscheinlichkeit stammt die Aufgabe vom Professor, vom Assistenten bzw. von der Hilfskraft?

(b) In der Stadt Omega wird ein neuer Impfstoff gegen Stochastichitis erprobt.
Mit
$I \,\hat{=}\,$ {Menge der Einwohner von Omega, die vor Ausbruch der Epidemie geimpft wurden}
$S \,\hat{=}\,$ {Menge der Einwohner von Omega, die an Stochastichitis erkrankt sind}
$M \,\hat{=}\,$ {Menge der männlichen Einwohner von Omega}
gilt:

$\#\left(S \cap I \cap M\right) = 1000$
$\#\left(S \cap I^{\mathsf{C}} \cap M\right) = 14000$
$\#\left(S^{\mathsf{C}} \cap I \cap M\right) = 1000$
$\#\left(S^{\mathsf{C}} \cap I^{\mathsf{C}} \cap M\right) = 10000$
$\#\left(S \cap I \cap M^{\mathsf{C}}\right) = 4000$
$\#\left(S \cap I^{\mathsf{C}} \cap M^{\mathsf{C}}\right) = 10000$
$\#\left(S^{\mathsf{C}} \cap I \cap M^{\mathsf{C}}\right) = 1000$
$\#\left(S^{\mathsf{C}} \cap I^{\mathsf{C}} \cap M^{\mathsf{C}}\right) = 2000$

Unter der Annahme, daß jede Person mit gleicher Wahrscheinlichkeit an Stochastichitis erkranke, vergleiche man:

(i) $\mathbf{P}\left(S \middle| I \cap M\right)$ mit $\mathbf{P}\left(S \middle| I^{\mathsf{C}} \cap M\right)$

(ii) $\mathbf{P}\left(S \middle| I \cap M^{\mathsf{C}}\right)$ mit $\mathbf{P}\left(S \middle| I^{\mathsf{C}} \cap M^{\mathsf{C}}\right)$

(iii) $\mathbf{P}\left(S \middle| I\right)$ mit $\mathbf{P}\left(S \middle| I^{\mathsf{C}}\right)$

(c) Geben Sie die Formel von Bayes an.

Aufgabe 4: $4\mathbf{P} + 6\mathbf{P} + 2\mathbf{P} = 12\mathbf{P}$

(a) Der zweidimensionale Zufallsvektor (X, Y) sei gleichverteilt auf dem Einheitskreis $\Omega := \left\{ (x, y) \in \mathbb{R}^2 \middle| \, x^2 + y^2 \leq 1 \right\}$. Es sei R der Abstand eines zufällig gezogenen Punktes (x, y) zum Ursprung. Berechnen Sie:

- **(i)** Die Verteilungsfunktion von R;
- **(ii)** Die Dichte der Verteilung von R;
- **(iii)** Den Erwartungswert und die Varianz von R.

(b) Aus einer Urne mit n Kugeln, die von 1 bis n durchnummeriert sind, werden $k \geq 1$ Kugeln zufällig und mit Zurücklegen gezogen. Sei X die größte gezogene Zahl unter den k gezogenen.

- **(i)** Man gebe ein geeignetes mathematisches Modell an.
- **(ii)** Man berechne $\mathbf{P}\{X \leq m\}$ für $1 \leq m \leq n$.
- **(iii)** Man bestimme die Verteilung von X.
- **(iv)** Berechnen Sie $\mathbf{E}(X)$. (Es reicht eine Formel anzugeben!)

(c) Man definiere die Begriffe Verteilungsfunktion und Dichte der Verteilung einer Zufallsvariablen.

Aufgabe 5: 4P + 6P + 2P = 12P

(a) Die Verteilung der reellen Zufallsvariablen X habe die Dichte f_X. Zeigen Sie, daß die Verteilung von $Y := |X|$ die Dichte

$$f_Y(x) = \begin{cases} f_X(x) + f_X(-x) & \text{falls } x > 0 \\ 0 & \text{falls } x \leq 0 \end{cases}$$

besitzt.

(b) Es sei X auf $[-1, 1]$ gleichverteilt. Zeigen Sie:

(i) X und $Y := X^2$ sind unkorreliert;

(ii) Die Zufallsvariablen $Y_k := \sin(\pi k X)$, $k \geq 1$ sind paarweise unkorreliert.

(iii) $\frac{1}{n}\sum_{k=1}^{n} \sin(\pi k X) \to 0 \quad$ für $n \to \infty$ stochastisch.

(c) Geben Sie die Transformationsformel für $\mathbf{E}(g(X))$ an.

H-Doz. Dr. H.P. Scheffler Sommersemester 2001

Nachklausur zur Vorlesung Stochastik I

Name, Vorname:

Geburtsdatum:

Matrikelnummer:

Bearbeiten Sie bitte alle fünf Aufgaben. Die maximal erreichbare Punktezahl finden Sie neben jeder Aufgabe.

Für einen qualifizierten Studiennachweis benötigen Sie mindestens $18 \pm \varepsilon$ Punkte. Für einen Leistungsnachweis sind mindestens $24 \pm \varepsilon$ Punkte notwendig. Sie haben zum Bearbeiten der Aufgaben 3 Stunden Zeit.

Bitte verwenden Sie für jede Aufgabe ein gesondertes Blatt, welches Sie mit Ihrem Namen und Matrikelnummer versehen.

Achten Sie in der Klausur auf sorgfältige und exakte Formulierungen. Geben Sie, falls gefordert, Definitionen oder Sätze stets exakt (mit allen Voraussetzungen) an.

Bei der Bearbeitung der einzelnen Aufgaben genügt es, entsprechende Formeln abzuleiten, exakte numerische Berechnungen, etwa mit Hilfe eines Taschenrechners sind nicht erforderlich.

Hilfsmittel (Bücher, Skripten) sind nicht zugelassen!

Aufgabe 1: $8\mathbf{P} + 2\mathbf{P} = 10\mathbf{P}$

(a) An einer Theatergarderobe werden nach Ende der Vorstellung zunächst die Mäntel und anschließend (unabhängig davon) die Hüte zufällig zurückgegeben. Wie groß ist die Wahrscheinlichkeit, daß keiner der Besucher beide Kleidungsstücke (sowohl den eigenen Mantel, als auch den eigenen Hut) zurückerhält?

(b) Geben Sie die Siebformel an.

Aufgabe 2: 2P + 2P + 2P + 2P = 8P

Die Blutspender der UniDo werden nach einem neuen Testverfahren auf Antikörper gegen Viren vom Typ B untersucht. Spenden, bei denen der Test positiv ist, werden ausgesondert. Aufgrund klinischer Untersuchungen ist bekannt, daß der Test mit Wahrscheinlichkeit $1-\alpha$ positiv ausfällt, wenn das Spenderblut Antikörper enthält, und mit Wahrscheinlichkeit $1-\beta$ negativ ausfällt, wenn keine Antikörper vorliegen. Aufgrund früherer Untersuchungen weiß man, daß bei den Spendern der UniDo das Spenderblut mit Wahrscheinlichkeit γ Antikörper enthält. (Dabei sind $0 < \alpha, \beta, \gamma < 1$.)

(a) Wie groß ist die Wahrscheinlichkeit, daß eine Blutprobe mit Antikörpern beim Test nicht erkannt wird?

(b) Wie groß ist die Wahrscheinlichkeit, daß eine Blutprobe ausgesondert wird?

(c) Wie groß ist die Wahrscheinlichkeit, daß eine Spende, die ausgesondert wurde, antikörperfrei ist?

(d) Geben Sie

- **(i)** den Satz von der totalen Wahrscheinlichkeit und
- **(ii)** die Bayes'sche Formel an.

Aufgabe 3: $2\mathbf{P} + 1\mathbf{P} + 3\mathbf{P} + 4\mathbf{P} + 2\mathbf{P} = 12\mathbf{P}$

Sei $a > 0$ und $k \in \mathbb{N}$. Die Pareto-Verteilung $\mathbf{P}_{a;k}$ ist definiert durch die Verteilungsfunktion

$$F(x) = \begin{cases} 0 & x \leq a \\ 1 - a^k x^{-k} & x > a \end{cases}$$

(a) Zeigen Sie, daß F eine Verteilungsfunktion ist.

(b) Zeigen Sie, daß $\mathbf{P}_{a;k}$ eine Dichte besitzt und bestimmen Sie diese.

(c) Es sei X eine Zufallsvariable mit Verteilung $\mathbf{P}_{a;k}$. Für welche $j \in \mathbb{N}$ ist $\mathbf{E}\left(|X|^j\right) < \infty$?

(d) Es seien X, Y unabhängige $\mathbf{P}_{a;k}$-verteilte Zufallsvariable. Geben Sie die Verteilungsfunktionen und Dichten von $\min(X, Y)$ und $\max(X, Y)$ an.

(e) Definieren Sie die Begriffe:

- **(i)** Lebesgue-Dichte einer Verteilung;
- **(ii)** Verteilungsfunktion einer reellen Zufallsvariable.

Aufgabe 4: 4P + 2P + 2P + 2P = 10P

(a) Die Bauteile B_i eines Systems besitzen die Lebensdauer X_i, $1 \leq i \leq n$. Die X_i seien unabhängig und besitzen eine Exponentialverteilung mit Parameter $\lambda_i \geq 0$ für $1 \leq i \leq n$. Bestimmen Sie die Verteilung der Gersamtlebensdauer Z des Systems, falls

- **(i)** die B_i in Reihe geschaltet sind;
- **(ii)** die B_i parallel geschaltet sind.

(b) Es sei X eine Zufallsvariable mit Exponentialverteilung mit Parameter $\lambda > 0$. Berechnen Sie $\mathbf{E}(X)$ und $\mathbf{V}(X)$.

(c) Beweisen Sie: Sind $X_1, \ldots, X_n$ unabhängige reelle Zufallsvariablen, so gilt für die Verteilungsfunktion

$$F_{\max\{X_i:1\leq i\leq n\}}(x) = \prod_{i=1}^{n} F_{X_i}(x).$$

(d) **(i)** Definieren Sie die Varianz einer reellen Zufallsvariablen.

(ii) Geben Sie die Tschebyscheff-Ungleichung an.

Aufgabe 5: $5\mathbf{P} + 5\mathbf{P} + 2\mathbf{P} = 12\mathbf{P}$

(a) Es sei $(X_0, X_1, \ldots)$ eine Markovkette mit Übergangsmatrix

$$K = \begin{pmatrix} 0 & \frac{1}{2} & \frac{1}{2} \\ \frac{1}{2} & 0 & \frac{1}{2} \\ 0 & 0 & 1 \end{pmatrix}$$

(i) Geben Sie den Übergangsgraphen an.

(ii) Was kann man über das Konvergenzverhalten von K^n beziehungsweise $\mathbf{P}\{X_n = i\}$ sagen? Bestimmen Sie gegebenenfalls die Grenzwerte.

(b) Es sei $(X_0, X_1, \ldots)$ eine Markovkette mit Übergangsmatrix

$$K = \begin{pmatrix} 0 & 1 & 0 \\ 1 & 0 & 0 \\ 0 & 0 & 1 \end{pmatrix}$$

(i) Berechnen Sie K^n für alle $n \geq 1$.

(ii) Berechnen Sie die stationären Vektoren von K, das heißt die Zähldichten π mit $\pi \cdot K = \pi$.

(c) Formulieren Sie den Satz von Markov.

Index