



**TECHNISCHE
UNIVERSITÄT
DRESDEN**

Fakultät Mathematik Institut für Stochastik, Professur für Angewandte Stochastik

STOCHASTIK

Prof. Dr. Anita Behme

Sommersemester 2019

Autor : Eric Kunze
E-Mail : eric.kunze@mailbox.tu-dresden.de

Inhaltsverzeichnis

0	Einleitung	2
1	Grundbegriffe der Wahrscheinlichkeitstheorie	3
1.1	Wahrscheinlichkeitsräume	3
1.2	Zufallsvariablen	7
2	Erste Standardmodelle	10
2.1	Diskrete Gleichverteilungen	10
2.2	Urnenmodelle	10
2.2.1	Urnenmodell mit Zurücklegen: Multinomial-Verteilung	10
2.2.2	Urnenmodell ohne Zurücklegen: Hypergeometrische Verteilung	12
2.3	Poisson-Approximation und -Verteilung	13
3	Bedingte Wahrscheinlichkeiten und Unabhängigkeit	14
3.1	Bedingte Wahrscheinlichkeiten	14
3.2	(Un)abhängigkeit	18

Kapitel 0

EINLEITUNG

Literatur

- *Georgii* : Stochastik (5. Auflage)
- *Schilling* : Wahrscheinlichkeit
- *Bauer* : Wahrscheinlichkeitstheorie (5. Auflage)
- *Krengel* : Einführung in die W-Theorie und Statistik
- *Dehling & Haupt* : Einführung in die W-Theorie und Statistik

Was ist Stochastik ?

Altgriechisch „Stochastikos“ ($\sigma\tau\chi\alpha\tau\iota\kappa\zeta$) \leadsto „Befarfsinnig im Vermuten“

Fragestellungen stammen insbesondere aus dem Glücksspiel, heute vielmehr auch aus der Versicherungs- und Finanzmathematik - überall da, wo Zufall / Risiko / Chance auftaucht.

Was ist mathematische Stochastik ?

- Beschreibt zufällige Phänomene in einer exakten Sprache.
Bsp.: „Beim Würfeln erscheint jedes sechste Mal (im Schnitt) die Augenzahl 6“ \leadsto Gesetz der großen Zahlen
- lässt sich in zwei Teilgebiete unterteilen: Wahrscheinlichkeitstheorie & Statistik
Die W-Theorie beschreibt und untersucht konkret gegebene Zufallssituationen. Dagegen zieht die Statistik Schlussfolgerungen aus Beobachtungen. Dabei benötigt sie die Modelle der W-Theorie - umgekehrt benötigt auch die W-Theorie die Statistik zur Bestätigung der Modelle.
- In diesem Semester konzentrieren wir uns auf die Wahrscheinlichkeitstheorie.

Kapitel 1

GRUNDBEGRIFFE DER

WAHRSCHEINLICHKEITSTHEORIE

1.1 Wahrscheinlichkeitsräume

Ergebnisraum

Welche möglichen Ausgänge eines zufälligen Geschehens interessieren uns?

■ Beispiel

Würfeln: Augenzahl, aber nicht Lage, Fallhöhe, usw.

Definition 1.1 (Ergebnisraum)

Die Menge der relevanten Ergebnisse eines Zufallsgeschehens nennen wir **Ergebnisraum** und bezeichnen diesen mit Ω .

■ Beispiel

- Würfeln: $\Omega = \{1, 2, \dots, 6\}$
- Wartezeiten: $\Omega = \mathbb{R}_+ = [0, \infty)$ (also überabzählbar)

Ereignisse

Oft interessiert man sich gar nicht für das konkrete Ergebnis des Zufallsexperiments, sondern nur für das Eintreten gewisser Ereignisse.

■ Beispiel

Würfeln: Zahl ist > 3

Wartezeiten: Wartezeit ist ≤ 5 Minuten

Wir wollen also Teilmengen des Ergebnisraums betrachten, d.h. Elemente von $\mathcal{P}(\Omega)$ (Potenzmenge), denen eine Wahrscheinlichkeit zugeordnet werden kann d.h. welche *messbar* sind.

Definition 1.2 (Ereignisraum)

Sei $\Omega \neq \emptyset$ ein Ergebnisraum und \mathcal{F} eine σ -Algebra auf Ω , d.h. eine Familie von Teilmengen von Ω , sodass

- (1) $\Omega \in \mathcal{F}$
- (2) $A \in \mathcal{F} \Rightarrow A^c \in \mathcal{F}$
- (3) $A_1, A_2, \dots \in \mathcal{F} \Rightarrow \bigcup_{i \geq 1} A_i \in \mathcal{F}$

Dann heißt (Ω, \mathcal{F}) **Ereignisraum** oder messbarer Raum.

Wahrscheinlichkeit

Wir ordnen nun den Ereignissen Wahrscheinlichkeiten mittels einer Abbildung $\mathbb{P}: \mathcal{F} \rightarrow [0, 1]$ zu, sodass

(N) Normierung: $\mathbb{P}(\Omega) = 1$

(A) Additivität: Für paarweise disjunkte Ereignisse $A_1, A_2, \dots \in \mathcal{F}$ ist $\mathbb{P}\left(\bigcup_{i \geq 1} A_i\right) = \sum_{i \geq 1} \mathbb{P}(A_i)$.

(N), (A) und die Nichtnegativität von \mathbb{P} werden als Kolmogorov-Axiome bezeichnet (nach Kolmogorov: Grundbegriffe der Wahrscheinlichkeitstheorie, 1933).

Definition 1.3 (Wahrscheinlichkeit)

Sei (Ω, \mathcal{F}) ein Ereignisraum und $\mathbb{P}: \mathcal{F} \rightarrow [0, 1]$ eine Abbildung mit den Eigenschaften (N) und (A). Dann heißt \mathbb{P} **Wahrscheinlichkeitsmaß** oder auch **Wahrscheinlichkeitsverteilung**.

Aus der Definition folgen direkt die folgenden Eigenschaften:

Satz 1.4 (Rechenregeln für Wahrscheinlichkeitsmaße)

Sei \mathbb{P} ein W-Maß auf einem Ereignisraum (Ω, \mathcal{F}) und $A, B, A_1, A_2, \dots \in \mathcal{F}$. Dann gilt:

- (1) $\mathbb{P}(\emptyset) = 0$
- (2) Endliche Additivität: $\mathbb{P}(A \cup B) + \mathbb{P}(A \cap B) = \mathbb{P}(A) + \mathbb{P}(B)$ und $\mathbb{P}(A) + \mathbb{P}(A^c) = 1$
- (3) Monotonie: $A \subseteq B \Rightarrow \mathbb{P}(A) \leq \mathbb{P}(B)$
- (4) σ -Subadditivität: $\mathbb{P}\left(\bigcup_{i \geq 1} A_i\right) \leq \sum_{i \geq 1} \mathbb{P}(A_i)$
- (5) σ -Stetigkeit: Wenn $A_n \nearrow A$ (d.h. $A_1 \subseteq A_2 \subseteq \dots$ und $A = \bigcup_{i=1}^{\infty} A_i$) oder $A_n \searrow A$, so gilt $\mathbb{P}(A_n) \rightarrow \mathbb{P}(A)$ für $n \rightarrow \infty$

Beweis. siehe MINT oder Schillings Lehrbuch

□

Beispiel 1.5

Für einen beliebigen Ereignisraum (Ω, \mathcal{F}) und ein beliebiges Element $\xi \in \Omega$ definiert

$$\delta_{\xi}(A) := \begin{cases} 1 & \xi \in A \\ 0 & \text{sonst} \end{cases}$$

ein (degeneriertes) W-Maß auf (Ω, \mathcal{F}) , welches wir als **Dirac-Maß** oder Dirac-Verteilung bezeichnen.

Beispiel 1.6

Wir betrachten das Zufallsexperiment "Würfeln mit einem fairen, 6-seitigen Würfel" mit der Ergebnismenge $\Omega = \{1, \dots, 6\}$ und Ereignisraum $\mathcal{F} = \mathcal{P}(\Omega)$. Setzen wir aus Symmetriegründen

$$\mathbb{P}(A) = \frac{\#A}{6}$$

mit $\#A = |A|$ = Kardinalität. Dies definiert ein W-Maß.

Beispiel 1.7 (Wartezeiten an der Bushaltestelle)

Ergebnisraum $\Omega = \mathbb{R}_+$ und Ereignisraum Borel'sche σ -Algebra $\mathcal{F} = \mathcal{B}(\mathbb{R}_+)$. Ein mögliches W-Maß können wir durch

$$\mathbb{P}(A) := \int_A \lambda e^{-\lambda x} \, dx$$

für einen Parameter $\lambda > 0$ festlegen. (offensichtlich gelten $\mathbb{P}(\Omega) = 1$ und die σ -Additivität aufgrund der σ -Additivität des Integrals). Wir bezeichnen dieses Maß als **Exponentialverteilung**. (Warum gerade dieses Maß für Wartezeiten gut geeignet ist, sehen wir später.)

Satz 1.8 (Konstruktion von WMaßen mit Dichten)

Sei (Ω, \mathcal{F}) ein Ereignisraum.

- Ω abzählbar, $\mathcal{F} = \mathcal{P}(\Omega)$:

Sei $\rho = (\rho(\omega))_{\omega \in \Omega}$ eine Folge in $[0, 1]$ in $\sum_{\omega \in \Omega} \rho(\omega) = 1$, dann definiert

$$\mathbb{P}(A) = \sum_{\omega \in A} \rho(\omega), \quad A \in \mathcal{F}$$

ein (diskretes) WMaß \mathbb{P} auf (Ω, \mathcal{F}) . ρ wird als **Zähldichte** bezeichnet. Umgekehrt definiert jedes WMaß \mathbb{P} auf (Ω, \mathcal{F}) mittels $\rho(\omega) = \mathbb{P}(\{\omega\})$, $\omega \in \Omega$ eine Folge ρ mit den obigen Eigenschaften.

- $\Omega \subseteq \mathbb{R}^n$, $\mathcal{F} = \mathcal{B}(\Omega)$:

Sei $\rho: \Omega \rightarrow [0, \infty)$ eine Funktion, sodass

- (1) $\int_{\Omega} \rho(x) \, dx = 1$
- (2) $\{x \in \Omega: \rho(x) \leq c\} \in \mathcal{B}(\Omega)$ für alle $c > 0$

dann definiert ρ ein WMaß \mathbb{P} auf (Ω, \mathcal{F}) durch

$$\mathbb{P}(A) = \int_A \rho(x) \, dx = \int_A \rho \, d\lambda, \quad A \in \mathcal{B}(\Omega)$$

Das Integral interpretieren wir stets als Lebesgue-Integral bzgl. Lebesgue-Maß λ . ρ bezeichnet wir als **Dichte**, **Dichtefunktion** oder **Wahrscheinlichkeitsdichte** von \mathbb{P} und nennen ein solches \mathbb{P} (absolut) **stetig** (bzgl. dem Lebesgue-Maß).

Beweis. Der diskrete Fall ist klar. Im stetigen Fall folgt die Behauptung aus den bekannten Eigenschaften des Lebesgue-Integrals (\nearrow Schilling MINT, Lemma 8.9) \square

► **Bemerkung.** ■ Die eindeutige Beziehung zwischen Dichte und Wahrscheinlichkeitsmaß überträgt sich nicht auf den stetigen Fall.

▷ Nicht jedes Wahrscheinlichkeitsmaß auf $(\Omega, \mathcal{B}(\Omega))$, $\Omega \subset \mathbb{R}^n$ besitzt eine Dichte.

▷ Zwei Dichtefunktionen definieren dasselbe Wahrscheinlichkeitsmaß, wenn sie sich nur auf einer Menge von Lebesgue-Maß 0 unterscheiden.

- Jede auf $\Omega \subset \mathbb{R}^n$ definierte Dichtefunktion ρ lässt sich auf ganz \mathbb{R}^n fortsetzen durch $\rho(x) = 0, x \notin \Omega$. Das erzeugte WMaß auf $(\mathbb{R}^n, \mathcal{B}(\mathbb{R}^n))$ lässt mit den WMaß auf $(\Omega, \mathcal{B}(\Omega))$ identifizieren.
- Mittels Dirac-Maß δ_x können auch jedes diskrete WMaß auf $\Omega \subset \mathbb{R}^n$ als WMaß auf $\mathbb{R}^n, \mathcal{B}(\mathbb{R}^n)$ interpretieren:

$$\mathbb{P}(A) = \sum_{\omega \in A} \rho(\omega) = \int_A d\left(\sum_{\omega \in \Omega} \rho(\omega) \delta_{\omega}\right) \quad A \in \mathcal{B}(\mathbb{R}^n)$$

- stetige und diskrete WMaße lassen sich kombinieren z.B. definiert

$$\mathbb{P}(A) = \frac{1}{2} \delta_0 + \frac{1}{2} \int_A \mathbb{1}_{[0,1]}(x) \, dx, \quad A \in \mathcal{B}(\mathbb{R})$$

ein WMaß auf $(\mathbb{R}, \mathcal{B}(\mathbb{R}))$.

Abschließend erinnern wir uns an:

Satz 1.9 (Eindeutigkeitssatz für Wahrscheinlichkeitsmaße)

Sei (Ω, \mathcal{F}) Ereignisraum und \mathbb{P} ein WMaß auf (Ω, \mathcal{F}) . Sei $\mathcal{F} = \sigma(\mathcal{G})$ für ein \cap -stabiles Erzeugendensystem $\mathcal{G} \subset \mathcal{P}(\Omega)$. Dann ist \mathbb{P} bereits durch seine Einschränkung $\mathbb{P}|_{\mathcal{G}}$ eindeutig bestimmt.

Beweis. ↗ Schilling MINT, Satz 4.5. □

Insbesondere definiert z.B.

$$\mathbb{P}([0, a)) = \int_0^a \lambda e^{-\lambda x} dx = 1 - e^{-\lambda a}, a > 0$$

bereits die Exponentialverteilung aus Beispiel 1.7.

Definition 1.10 (Gleichverteilung)

Ist Ω endlich, so heißt das WMaß mit konstanter Zähldichte

$$\rho(\omega) = \frac{1}{|\Omega|}$$

die **(diskrete) Gleichverteilung** auf Ω und wird mit $U(\Omega)$ notiert ($U = \text{uniform}$).

Ist $\Omega \subset \mathbb{R}^n$ eine Borelmenge mit Lebesgue-Maß $0 < \lambda^n(\Omega) < \infty$ so heißt das WMaß auf $(\Omega, \mathcal{B}(\Omega))$ mit konstanter Dichtefunktion

$$\rho(x) = \frac{1}{\lambda^n(\Omega)}$$

die **(stetige) Gleichverteilung** auf Ω . Sie wird ebenso mit $U(\Omega)$ notiert.

Wahrscheinlichkeitsräume

Definition 1.11 (Wahrscheinlichkeitsraum)

Ein Tripel $(\Omega, \mathcal{F}, \mathbb{P})$ mit Ω, \mathcal{F} Ereignisraum und \mathbb{P} WMaß auf (Ω, \mathcal{F}) , nennen wir **Wahrscheinlichkeitsraum**.

1.2 Zufallsvariablen

Zufallsvariablen dienen dazu von einem gegebenen Ereignisraum (Ω, \mathcal{F}) zu einem Modellausschnitt Ω', \mathcal{F}' überzugehen. Es handelt sich also um Abbildungen $X: \Omega \rightarrow \Omega'$. Damit wir auch jedem Ereignis in \mathcal{F}' eine Wahrscheinlichkeit zuordnen können, benötigen wir

$$A' \in \mathcal{F}' \Rightarrow X^{-1}A' \in \mathcal{F}$$

d.h. X sollte **messbar** sein.

Definition 1.12 (Zufallsvariable)

Seien (Ω, \mathcal{F}) und (Ω', \mathcal{F}') Ereignisräume. Dann heißt jede messbare Abbildung

$$X: \Omega \rightarrow \Omega'$$

Zufallsvariable (von (Ω, \mathcal{F}) nach (Ω', \mathcal{F}') / auf (Ω', \mathcal{F}') oder **Zufallselement**.

Beispiel 1.13

- (1) Ist Ω abzählbar und $\mathcal{F} = \mathcal{P}(\Omega)$, so ist jede Abbildung $X: \Omega \rightarrow \Omega'$ messbar und damit eine Zufallsvariable.
- (2) Ist $\Omega \subset \mathbb{R}^n$ und $\mathcal{F} = \mathcal{B}(\Omega)$, so ist jede stetige Funktion $X: \Omega \rightarrow \mathbb{R}$ messbar und damit eine Zufallsvariable.

Satz 1.14

Sei $(\Omega, \mathcal{F}, \mathbb{P})$ ein Wahrscheinlichkeitsraum und X eine Zufallsvariable von (Ω, \mathcal{F}) nach (Ω', \mathcal{F}') . Dann definiert

$$\mathbb{P}'(A') := \mathbb{P}(X^{-1}(A')) = \mathbb{P}(\{X \in A'\}), \quad A' \in \mathcal{F}'$$

ein Wahrscheinlichkeitsmaß auf (Ω', \mathcal{F}') , welches wir als **Wahrscheinlichkeitsverteilung von X unter \mathbb{P}** bezeichnen.

Beweis. Aufgrund der Messbarkeit von X ist die Definition sinnvoll. Zudem gelten

$$\mathbb{P}'(\Omega') = \mathbb{P}(X^{-1}(\Omega')) = \mathbb{P}(\Omega) = 1$$

und für $A'_1, A'_2, \dots \in \mathcal{F}'$ paarweise disjunkt.

$$\begin{aligned} \mathbb{P}'\left(\bigcup_{i \geq 1} A'_i\right) &= \mathbb{P}\left(X^{-1}\left(\bigcup_{i \geq 1} A'_i\right)\right) = \mathbb{P}\left(\bigcup_{i \geq 1} X^{-1}(A'_i)\right) \\ &= \sum_{i \geq 1} \mathbb{P}(X^{-1}A'_i) \quad \text{da auch } X^{-1}A'_1, X^{-1}A'_2, \dots \text{ paarweise disjunkt sind} \\ &= \sum_{i \geq 1} \mathbb{P}'(A_i) \end{aligned}$$

Also ist \mathbb{P}' ein Wahrscheinlichkeitsmaß. □

► **Bemerkung.** ■ Aus Gründen der Lesbarkeit schreiben wir in der Folge

$$\mathbb{P}(X \in A) = \mathbb{P}(\{\omega: X(\omega) \in A\})$$

■ Ist $X = \text{id}$, so fallen die Begriffe Wahrscheinlichkeitsmaß und Wahrscheinlichkeitsverteilung zusammen.

- In der (weiterführenden) Literatur zur Wahrscheinlichkeitstheorie wird oft auf die Angabe eines zugrundeliegenden Wahrscheinlichkeitsraumes verzichtet und stattdessen eine “Zufallsvariable mit Verteilung \mathbb{P} auf Ω ” eingeführt. Gemeint ist (fast) immer X als Identität auf $(\Omega, \mathcal{F}, \mathbb{P})$ mit $\mathcal{F} = \mathcal{P}(\Omega)$ oder $\mathcal{F} = \mathcal{B}(\Omega)$.
- Für die Verteilung von X unter \mathbb{P} schreibe \mathbb{P}_X und $X \sim \mathbb{P}_X$ für die Tatsache, dass X gemäß \mathbb{P}_X verteilt ist.

Definition 1.15

Zwei Zufallsvariablen sind **identisch verteilt**, wenn sie dieselbe Verteilung haben.

Von besonderen Interesse sind für uns die Zufallsvariablen, die nach $(\mathbb{R}, \mathcal{B}(\mathbb{R}))$ abbilden, sogenannte **reelle Zufallsvariablen**.

Da die halboffenen Intervalle $\mathcal{B}(\mathbb{R})$ erzeugen, ist die Verteilung einer reellen Zufallsvariable durch die Werte $(-\infty, c]$, $c \in \mathbb{R}$ eindeutig festgelegt.

Definition 1.16 (Verteilungsfunktion)

Sei $(\mathbb{R}, \mathcal{B}(\mathbb{R}), \mathbb{P})$ ein Wahrscheinlichkeitsraum, so heißt

$$F: \mathbb{R} \rightarrow [0, 1] \text{ mit } x \mapsto \mathbb{P}((-\infty, x))$$

(kumulative) **Verteilungsfunktion von \mathbb{P}** . Ist X eine reelle Zufallsvariable auf beliebigem WRaum $(\Omega, \mathcal{F}, \mathbb{P})$, so heißt

$$F: \mathbb{R} \rightarrow [0, 1] \text{ mit } x \mapsto \mathbb{P}(X \leq x) = \mathbb{P}(X \in (-\infty, x])$$

die (kumulative) Verteilungsfunktion von X .

Beispiel 1.17

$(\mathbb{R}, \mathcal{B}(\mathbb{R}), \mathbb{P})$ mit \mathbb{P} Exponentialverteilung mit Parameter $\lambda > 0$

$$\mathbb{P}(A) = \int_{A \cap [0, \infty)} \lambda e^{-\lambda x} dx \quad A \in \mathcal{B}(\mathbb{R})$$

Dann ist

$$F(x) = \mathbb{P}((-\infty, x)) = \begin{cases} 0 & x \leq 0 \\ \int_0^x \lambda e^{-\lambda y} dy = 1 - e^{-\lambda x} & x > 0 \end{cases}$$

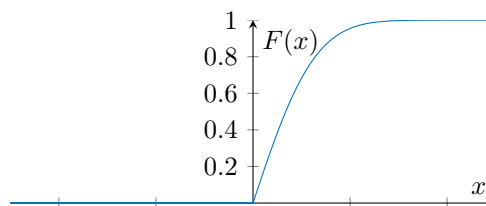


Abbildung 1.1: Verteilungsfunktion der Exponentialverteilung

Beispiel 1.18

Das Würfeln mit einem fairen, sechseitigen Würfel kann mittels einer reellen Zufallsvariablen

$$X: \{1, 2, \dots, 6\} \rightarrow \mathbb{R} \text{ mit } x \mapsto x$$

modelliert werden.

Es folgt als Verteilungsfunktion

$$F(x) = \mathbb{P}(X \leq x) = \mathbb{P}(X^{-1}(-\infty, x]) = \mathbb{P}((-\infty, x]) = \frac{1}{6} \sum_{i=1}^6 \mathbb{1}_{\{i \leq x\}}$$

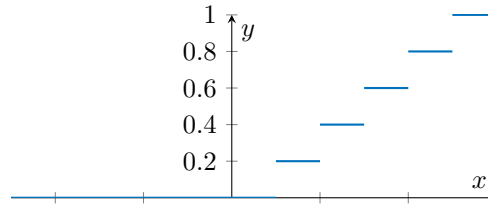


Abbildung 1.2: Verteilungsfunktion des Würfel-experiments

Diese Erkenntnisse lassen sich auch verallgemeinern:

Satz 1.19

Ist \mathbb{P} ein Wahrscheinlichkeitsmaß auf $(\mathbb{R}, \mathcal{B}(\mathbb{R}))$ und F die zugehörige Verteilungsfunktion, so gelten

- (1) F ist monoton wachsend
- (2) F ist rechtsseitig stetig
- (3) $\lim_{x \rightarrow -\infty} F(x) = 0$ und $\lim_{x \rightarrow \infty} F(x) = 1$

Umgekehrt existiert zu jeder Funktion $F: \mathbb{R} \rightarrow [0, 1]$ mit Eigenschaften (1) bis (3) eine reelle Zufallsvariable auf $((0, 1), \mathcal{B}((0, 1)), \mathbb{U}((0, 1)))$ mit Verteilungsfunktion F .

Beweis. Ist F Verteilungsfunktion, so folgt mit 1.4

$$x \leq y \Rightarrow F(x) = \mathbb{P}((-\infty, x]) \stackrel{1.4.3}{\leq} \mathbb{P}((-\infty, y]) = F(y)$$

und

$$\lim_{m \searrow c} F(x) = \lim_{m \searrow c} \mathbb{P}((-\infty, x]) \stackrel{\sigma\text{-Stetigkeit}}{=} \mathbb{P}((-\infty, c]) = F(c)$$

sowie

$$\lim_{x \rightarrow -\infty} F(x) \stackrel{1.4.5}{=} \mathbb{P}(\emptyset) \stackrel{1.4.1}{=} 0 \quad \text{und} \quad \lim_{x \rightarrow \infty} F(x) \stackrel{1.4.5}{=} \mathbb{P}(\mathbb{R}) = 1$$

Umgekehrt wähle

$$X(u) := \inf \{x \in \mathbb{R}: F(x) \geq u\}, \quad u \in (0, 1)$$

Dann ist X eine “linkseitige Inverse” von F (auch **Quantilfunktion** oder **verallgemeinerte Inverse**). Wegen (3) gilt $-\infty < X(u) < \infty$ und zudem

$$\{X \leq x\} = (0, F(x)) \cap (0, 1) \in \mathcal{B}((0, 1))$$

Da diese halboffene Mengen ein Erzeugendensystem von $\mathcal{B}(\mathbb{R})$ bilden, folgt bereits die Messbarkeit von X , also ist X eine Zufallsvariable. Insbesondere hat die Menge $\{X \leq x\}$ gerade LEBESGUE-Maß $F(x)$ und damit hat X die Verteilungsfunktion F . \square

Korollar 1.20

Ist \mathbb{P} Wahrscheinlichkeitsmaß auf $(\mathbb{R}, \mathcal{B}(\mathbb{R}))$ und F die zugehörige Verteilungsfunktion. Dann besitzt \mathbb{P} genau eine Dichtefunktion ρ , wenn F stetig differenzierbar ist, denn dann gilt

$$F(x) = \int_0^x \rho(x) \, dx \quad \text{bzw.} \quad \rho(x) = F'(x)$$

Beweis. Folgt aus Satz 1.8, der Definition 1.16 der Verteilungsfunktion und dem Eindeutigkeitssatz 1.9. \square

Kapitel 2

ERSTE STANDARDMODELLE

— Diskrete Verteilungen —

2.1 Diskrete Gleichverteilungen

Erinnerung (Definition 1.10)

Ist Ω endlich, so heißt Wahrscheinlichkeitsmaß mit Zähldichte $\rho(\omega) = \frac{1}{\#\Omega}$ für $\omega \in \Omega$ **(diskrete) Gleichverteilung** auf $\Omega \rightsquigarrow \mathcal{U}(\Omega)$

Es gilt das für jedes $A \in \mathcal{P}(\Omega)$

$$\mathbb{P}(A) = \frac{\#A}{\#\Omega}$$

Anwendungsbeispiele sind faires Würfeln, fairer Münzwurf, Zahlenlotto, ...

2.2 Urnenmodelle

Ein “Urnenmodell” ist eine abstrakte Darstellung von Zufallsexperimenten, bei denen zufällig Stichproben aus einer gegebenen Menge “gezogen” werden. Eine Urne ist ein Behälter in welchem sich farbige/nummerierte Kugeln befinden, die ansonsten ununterscheidbar sind. Aus der Urne ziehe man blind/zufällig eine oder mehrere Kugeln und notiere ihre Farbe/Zahl.

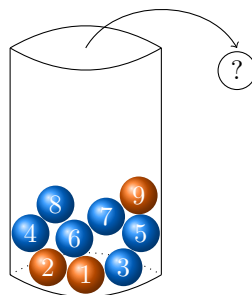


Abbildung 2.1: Urnenmodell mit nummerierten, farbigen Kugeln

2.2.1 Urnenmodell mit Zurücklegen: Multinomial-Verteilung

Gegeben sei eine Urne mit N Kugeln, verschiedenfarbig mit Farben aus E , wobei $|E| \geq 2$

Man ziehe n Stichproben/Kugeln, wobei nach jedem Zug die Kugel wieder zurückgelegt wird. Uns interessiert die Farbe in jedem Zug, setze also

$$\Omega = E^n \text{ und } \mathcal{F} = \mathcal{P}(\Omega)$$

Zur Bestimmung eines geeigneten Wahrscheinlichkeitsmaßes nummerieren wir die Kugeln mit $1, \dots, N$, so dass alle Kugeln der Farbe $a \in E$ eine Nummer aus $F_a \subset \{1, \dots, N\}$ tragen. Würden wir die Nummern notieren, so wäre

$$\bar{\Omega} = \{1, \dots, N\}^n \text{ und } \bar{\mathcal{F}} = \mathcal{P}(\bar{\Omega})$$

und wir könnten die Gleichverteilung $\bar{\mathbb{P}} = U(\bar{\Omega})$ als WMaß für einen einzelnen Zug verwenden. Für den Übergang zu Ω konstruieren wir Zufallsvariablen. Die Farbe im i -ten Zug wird beschrieben durch

$$X_i: \begin{cases} \bar{\Omega} & \rightarrow & E \\ \bar{\omega} = (\bar{\omega}_1, \dots, \bar{\omega}_n) & \mapsto & a \text{ falls } \bar{\omega}_i \in F_a \end{cases}$$

Der Zufallsvektor

$$X = (X_1, \dots, X_n): \bar{\Omega} \rightarrow \Omega$$

beschreibt dann die Abfolge der Farben. Für jedes $\omega \in \Omega$ gilt dann

$$\{X = \omega\} = F_{\omega_1} \times \dots \times F_{\omega_n} = \bigtimes_{i=1}^n F_{\omega_i}$$

und damit

$$\mathbb{P}(\{\omega\}) = \bar{\mathbb{P}}(X^{-1}(\{\omega\})) = \mathbb{P}(X = \omega) = \frac{|F_{\omega_1}| \cdots |F_{\omega_n}|}{|\bar{\Omega}|} = \prod_{i=1}^n \frac{|F_{\omega_i}|}{N} =: \prod_{i=1}^n \rho(\omega_i)$$

Zähldichten, die sich als Produkt von Zähldichten schreiben lassen, werden auch als **Produktdichten** bezeichnet (\nearrow §3 Unabhängigkeit).

Sehr oft interessiert uns bei einem Urnenexperiment nicht die Reihenfolge der gezogenen Farben, sondern nur die Anzahl der Kugeln in Farbe $a \in E$ nach n Zügen. Dies entspricht

$$\hat{\Omega} = \left\{ k = (k_a)_{a \in E} \in \mathbb{N}_0^{|E|} : \sum_{a \in E} k_a = n \right\} \text{ und } \hat{\mathcal{F}} = \mathcal{P}(\hat{\Omega})$$

Den Übergang $\Omega \rightarrow \hat{\Omega}$ beschreiben wir durch die Zufallsvariablen

$$Y_a(\omega): \begin{cases} \Omega & \rightarrow & \mathbb{N}_0 \\ \omega = (\omega_1, \dots, \omega_n) & \mapsto & \sum_{a \in E} \mathbb{1}_{\{a\}}(\omega_i) \end{cases}$$

und

$$Y = (Y_a)_{a \in E}: \Omega \rightarrow \hat{\Omega}$$

Wir erhalten

$$\begin{aligned} \mathbb{P}(Y = k) &= \mathbb{P}(Y_a = k_a : a \in E) \\ &= \sum_{\omega \in \Omega: Y(\omega) = k} \prod_{i=1}^n \rho(\omega_i) = \sum_{\omega \in \Omega: Y(\omega) = k} \prod_{a \in E} \rho(a)^{k_a} = \binom{n}{(k_a)_{a \in E}} \prod_{a \in E} \rho(a)^{k_a} \end{aligned}$$

wobei

$$\binom{n}{(k_1, \dots, k_l)} := \begin{cases} \frac{n!}{k_1! \cdots k_l!} & \text{falls } \sum_{i=1}^l k_i = n \\ 0 & \text{sonst} \end{cases}$$

der Multinomialkoeffizient ist, welcher die Anzahl der Möglichkeiten beschreibt n Objekte in l Gruppen aufzuteilen, sodass die Gruppe i gerade k_i Objekte enthält.

Definition 2.1

Sei $l \geq 2$, $p = (p_1, \dots, p_l)$ eine Zähl-dichte und $n \in \mathbb{N}$, dann heißt die Verteilung auf

$$\left\{ k = (k_i)_{i=1, \dots, l} \in \mathbb{N}_0^l : \sum_{i=1}^l k_i = n \right\}$$

mit Zähl-dichte

$$m((k_1, \dots, k_l)) = \binom{n}{k_1, \dots, k_l} \prod_{i=1}^l p_i^{k_i}$$

Multinomialverteilung mit Parametern n und p . Wir schreiben auch $\text{Multi}(n, p)$.

Beispiel 2.2

Eine Urne enthalte nur schwarze (“1”) und weiße (“0”) Kugeln ($E = \{0, 1\}$) und es sei $\rho(1) = p$ gerade die Proportion der schwarzen Kugeln (Wahrscheinlichkeit bei einem Zug schwarz zu ziehen). Dann ist die Wahrscheinlichkeit in n Zügen k -mal schwarz zu ziehen

$$\binom{n}{k} \prod_{i=0,1} \rho(i)^{k_i} = \binom{n}{k} \cdot p^k \cdot (1-p)^{n-k}$$

Ein solches (wiederholtes) Experiment mit nur zwei möglichen Ergebnissen und feste Wahrscheinlichkeit $p \in [0, 1]$ nennen wir auch (wiederholtes) **BERNOULLI-Experiment**.

Definition 2.3

Sei $p \in [0, 1]$ und $n \in \mathbb{N}$, dann heißt die Verteilung mit Zähl-dichte

$$\rho(k) = \binom{n}{k} p^k (1-p)^{n-k} \quad k \in \{0, 1, \dots, n\}$$

Binomialverteilung auf $\{0, \dots, n\}$ mit Parameter p (Erfolgswahrscheinlichkeit). Wir schreiben auch $\text{Bin}(n, p)$. Im Fall $n = 1$ nennen wir die Verteilung mit Zähl-dichte

$$\rho(0) = 1 - p \quad \rho(1) = p$$

auch **Bernoulliverteilung** mit Parameter p und schreiben $\text{Bernoulli}(p)$.

2.2.2 Urnenmodell ohne Zurücklegen: Hypergeometrische Verteilung

Gegeben sei ein Urne mit N Kugeln verschiedener Farben aus E mit $|E| \geq 2$. Es werden $n \leq N$ Stichproben entnommen, wobei die gezogenen Kugeln nicht in die Urne zurückgelegt werden.

Beispiel 2.4

Eine Urne enthalte S schwarze (“1”) und W weiße (“0”) Kugeln, d.h. $E = \{0, 1\}$ und $S + W = N$. Dann ist die Wahrscheinlichkeit in n Zügen ohne Zurücklegen gerade s schwarze und w weiße Kugeln zu ziehen

$$\rho(s) = \frac{\binom{W}{w} \cdot \binom{S}{s}}{\binom{N}{n}} \quad 0 \leq s \leq S, 0 \leq w \leq W, s + w = n, S + W = N$$

Definition 2.5

Seien $N \in \mathbb{N}, W \leq N, n \leq N$. Dann heißt die Verteilung auf $\{0, \dots, n\}$ mit Zähldichte

$$\rho(w) = \frac{\binom{W}{w} \cdot \binom{N-W}{n-w}}{\binom{N}{n}} \quad w = \max\{0, n - N + W\}, \dots, \min\{W, n\}$$

Hypergeometrische Verteilung mit Parametern N, W und n . Wir schreiben auch $\text{Hyper}(N, W, n)$.

2.3 Poisson-Approximation und -Verteilung

$\text{Bin}(n, p)$ ist zwar explizit und elementar definiert, jedoch für große n mühsam auszuwerten. Für seltene Ereignisse (n groß, p klein) kann man daher eher den folgenden Satz anwenden.

Satz 2.6 (Poisson-Approximation)

Sei $\lambda > 0$ und $(p_n)_{n \in \mathbb{N}}$ eine Folge in $[0, 1]$ mit

$$n \cdot p_n \rightarrow \lambda, \quad n \rightarrow \infty$$

Dann gilt für alle $k \in \mathbb{N}_0$ für die Zähldichte der $\text{Bin}(n, p_n)$ -Verteilung

$$\lim_{n \rightarrow \infty} \binom{n}{k} p_n^k (1 - p_n)^{n-k} = e^{-\lambda} \frac{\lambda^k}{k!}$$

Beweis. Sei $k \in \mathbb{N}_0$ fix, dann

$$\binom{n}{k} = \frac{n!}{k!(n-k)!} = \frac{n^k}{k!} \frac{n(n-1) \cdots (n-k+1)}{n^k} = \frac{n^k}{k!} \cdot 1 \cdot \left(1 - \frac{1}{n} \cdots \frac{k-1}{n}\right) \xrightarrow{n \rightarrow \infty} \frac{n^k}{k!}$$

wobei $a(l) \xrightarrow{n \rightarrow \infty} b(l) \Leftrightarrow \frac{a(l)}{b(l)} \xrightarrow{n \rightarrow \infty} 1$. Damit gilt

$$\begin{aligned} \binom{n}{k} p_n^k (1 - p_n)^{n-k} &\xrightarrow{n \rightarrow \infty} \frac{n^k}{k!} p_n^k (1 - p_n)^{n-k} \\ &\xrightarrow{n \rightarrow \infty} \frac{\lambda^k}{k!} (1 - p_n)^n = \frac{\lambda^k}{k!} \left(1 - \frac{n \cdot p_n}{n}\right)^n \\ &\xrightarrow{n \rightarrow \infty} \frac{\lambda^k}{k!} e^{-\lambda} \end{aligned}$$

Der erhaltene Grenzwert liefert die Zähldichte auf \mathbb{N}_0 , denn

$$\sum_{k=0}^{\infty} \frac{\lambda^k}{k!} e^{-\lambda} = e^{-\lambda} \sum_{k=0}^{\infty} \frac{\lambda^k}{k!} = e^{-\lambda} e^{\lambda} = 1$$

Definition 2.7

Sei $\lambda > 0$. Dann heißt das auf $(\mathbb{N}_0, \mathbb{P}(\mathbb{N}_0))$ definierte Wahrscheinlichkeitsmaß mit

$$\mathbb{P}(\{k\}) = \frac{\lambda^k}{k!} e^{-\lambda} \quad k \in \mathbb{N}_0,$$

Poissonverteilung mit Parameter λ . Wir schreiben auch $\text{Poisson}(\lambda)$.

Die Poisson-Verteilung ist ein natürliches Modell für die Anzahl von zufälligen, seltenen Ereignissen (z.B. Tore im Fußballspiel, Schadensfälle einer Versicherung)

Kapitel 3

BEDINGTE WAHRSCHEINLICHKEITEN UND UNABHÄNGIGKEIT

3.1 Bedingte Wahrscheinlichkeiten

Beispiel 3.1

Das Würfeln mit zwei fairen, sechsseitigen Würfeln können wir mit

$$\Omega = \{(i, j) : i, j \in \{1, \dots, 6\}\} \quad \text{und} \quad \mathbb{P} = \mathbb{U}(\Omega)$$

modellieren. Da $|\Omega| = 36$ gilt also $\mathbb{P}(\{\omega\}) = 1/36$ für alle $\omega \in \Omega$. Betrachten wir das Ereignis

$$A = \{(i, j) \in \Omega : i + j = 8\}$$

dann folgt $\mathbb{P}(A) = \frac{5}{36}$. Werden die beiden Würfe nacheinander ausgeführt, so kann nach dem ersten Wurf eine Neubewertung der Wahrscheinlichkeit von A erfolgen. Ist beispielsweise

$$B = \{(i, j) \in \Omega, i = 4\}$$

eingetreten, so kann die Summe 8 nur durch eine weitere 4 realisiert werden, also mit Wahrscheinlichkeit

$$\frac{1}{6} = \frac{|A \cap B|}{|B|}.$$

Das Eintreten von B führt also dazu, dass das Wahrscheinlichkeitsmaß \mathbb{P} durch ein neues Wahrscheinlichkeitsmaß \mathbb{P}_B ersetzt werden muss. Hierbei sollte gelten:

$$\text{Renormierung:} \quad \mathbb{P}_B = 1 \tag{R}$$

$$\text{Proportionalität:} \quad \text{Für alle } A \subseteq \mathcal{F} \text{ mit } A \subseteq B \text{ gilt } \mathbb{P}_B(A) = c_B \cdot \mathbb{P}(A) \text{ mit einer Konstante } c_B \tag{P}$$

Lemma 3.2

Sei $(\Omega, \mathcal{F}, \mathbb{P})$ Wahrscheinlichkeitsraum und $B \in \mathcal{F}$ mit $\mathbb{P}(B) > 0$. Dann gibt es genau ein Wahrscheinlichkeitsmaß \mathbb{P}_B auf (Ω, \mathcal{F}) mit den Eigenschaften (R) und (P). Dieses ist gegeben durch

$$\mathbb{P}_B(A) = \frac{\mathbb{P}(A \cap B)}{\mathbb{P}(B)} \quad \text{für alle } A \in \mathcal{F}$$

Beweis. Offenbar erfüllt \mathbb{P}_B wie definiert (R) und (P). Umgekehrt erfülle \mathbb{P}_B die Eigenschaften (R) und (P). Dann folgt für $A \in \mathcal{F}$

$$\mathbb{P}_B(A) = \mathbb{P}_B(A \cap B) + \underbrace{\mathbb{P}_B(A \setminus B)}_{=0, \text{ wegen (R)}} \stackrel{(P)}{=} c_B \cdot \mathbb{P}(A \cap B).$$

Für $A = B$ folgt zudem aus (R)

$$1 = \mathbb{P}_B(B) = c_B \cdot \mathbb{P}(B)$$

also $c_B = \mathbb{P}(B)^{-1}$. □

Definition 3.3

Sei $(\Omega, \mathcal{F}, \mathbb{P})$ Wahrscheinlichkeitsraum und $B \in \mathcal{F}$ mit $\mathbb{P}(B) > 0$. Dann heißt

$$\mathbb{P}(A \mid B) := \frac{\mathbb{P}(A \cap B)}{\mathbb{P}(B)} \text{ mit } A \in \mathcal{F}$$

die **bedingte Wahrscheinlichkeit** von A gegeben B . Falls $\mathbb{P}(B) = 0$, setze $\mathbb{P}(A \mid B) = 0$ für alle $A \in \mathcal{F}$.

Beispiel 3.4

In der Situation Beispiel 3.1 gilt $A \cap B = \{(4, 4)\}$ und damit

$$\mathbb{P}(A \mid B) = \frac{\mathbb{P}(A \cap B)}{\mathbb{P}(B)} = \frac{\frac{1}{36}}{\frac{1}{6}} = \frac{1}{6}$$

Aus Definition 3.3 ergibt sich das folgende Lemma.

Lemma 3.5 (Multiplikationsformel)

Sei $(\Omega, \mathcal{F}, \mathbb{P})$ ein Wahrscheinlichkeitsraum und $A_1, \dots, A_n \in \mathcal{F}$. Dann gilt

$$\mathbb{P}(A_1 \cap \dots \cap A_n) = \mathbb{P}(A_1) \mathbb{P}(A_2 \mid A_1) \dots \mathbb{P}(A_n \mid A_1 \cap \dots \cap A_{n-1})$$

Beweis. Ist $\mathbb{P}(A_1 \cap \dots \cap A_n) = 0$, so gilt auch $\mathbb{P}(A_n \mid A_1 \cap \dots \cap A_{n-1}) = 0$. Andernfalls sind alle Faktoren der rechten Seite ungleich Null und

$$\begin{aligned} & \mathbb{P}(A_1) \mathbb{P}(A_2 \mid A_1) \dots \mathbb{P}(A_n \mid A_1 \cap \dots \cap A_{n-1}) \\ &= \mathbb{P}(A_1) \cdot \frac{\mathbb{P}(A_1 \cap A_2)}{\mathbb{P}(A_1)} \dots \frac{\mathbb{P}(A_1 \cap \dots \cap A_n)}{\mathbb{P}(A_1 \cap \dots \cap A_{n-1})} \\ &= \mathbb{P}(A_1 \cap \dots \cap A_n) \end{aligned}$$

Stehen die A_i in Lemma 3.5 in einer (zeitlichen) Abfolge, so liefert Formel einen Hinweis wie Wahrscheinlichkeitsmaße für **Stufenexperimente** konstruiert werden können. Ein Stufenexperiment aus n nacheinander ausgeführten Teilexperimenten lässt sich als **Baumdiagramm** darstellen.

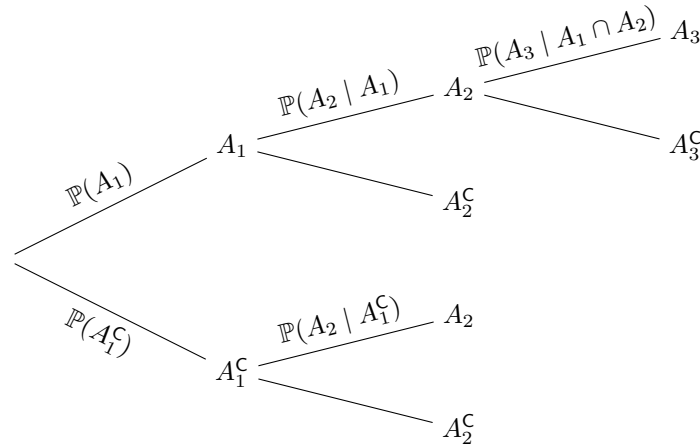


Abbildung 3.1: Darstellung eines Stufenexperiments

Satz 3.6 (Konstruktion des Wahrscheinlichkeitsmaßes eines Stufenexperiments)

Gegeben seien n Ergebnisräume $\Omega_i = \{\omega_i(1), \dots, \omega_i(k)\}, k \in \mathbb{N} \cup \{\infty\}$ und es sei $\Omega = \Omega_1 \times \dots \times \Omega_n$ der zugehörige Produktraum. Weiter seien \mathcal{F}_i σ -Algebren auf Ω_i und $\mathcal{F} = \bigotimes_{i=1}^n \mathcal{F}_i$ die Produkt- σ -Algebra auf Ω . Setze $\omega = (\omega_1, \dots, \omega_n)$ und

$$[\omega_1, \dots, \omega_m] := \{\omega_1\} \times \dots \times \{\omega_m\} \times \Omega_{m+1} \times \dots \times \Omega_n \quad (m \leq n)$$

und

$$\mathbb{P}(\{\omega_m\} \mid [\omega_1, \dots, \omega_{m-1}])$$

für die Wahrscheinlichkeit in der m -ten Stufe des Experiments ω_m zu beobachten, falls in den vorausgehenden Stufen $\omega_1, \dots, \omega_{m-1}$ beobachtet wurden. Dann definiert

$$\mathbb{P}(\{\omega\}) := \mathbb{P}(\{\omega_1\}) \prod_{m=2}^n \mathbb{P}(\{\omega_m\} \mid [\omega_1, \dots, \omega_{m-1}])$$

ein Wahrscheinlichkeitsmaß auf $(\Omega, \mathcal{F}, \mathbb{P})$.

Beweis. Nachrechnen! □

Beispiel 3.7 (POLYA-Urne)

Gegeben sei eine Urne mit s schwarzen und w weißen Kugeln. Bei jedem Zug wird die gezogene Kugel zusammen mit $c \in \mathbb{N}_0 \cup \{-1\}$ weiteren Kugeln derselben Farbe zurückgelegt.

- $c = 0$: Urnenmodell mit Zurücklegen
- $c = -1$: Urnenmodell ohne Zurücklegen

Beide haben wir schon in Kapitel 2.2 gesehen. Sei deshalb $c \in \mathbb{N}$ (Bsp. Modell für zwei konkurrierende Populationen). Ziehen wir n -mal, so haben wir ein n -Stufenexperiment mit

$$\Omega = \{0, 1\}^n \text{ mit } 0 = \text{“weiß”}, 1 = \text{“schwarz”} \quad (\Omega_i = \{0, 1\})$$

Zudem gelten im ersten Schritt

$$\mathbb{P}(\{0\}) = \frac{w}{s+w} \text{ und } \mathbb{P}(\{1\}) = \frac{s}{s+w}$$

sowie

$$\mathbb{P}(\{\omega_m\} \mid [\omega_1, \dots, \omega_{m-1}]) = \begin{cases} \frac{w+c(m-1-\sum_{i=1}^{m-1} \omega_i)}{s+w+c(m-1)} & \omega_m = 0 \\ \frac{s+c \sum_{i=1}^{m-1} \omega_i}{s+w+c(m-1)} & \omega_m = 1 \end{cases}$$

Mit Satz 3.6 folgt als Wahrscheinlichkeitsmaß auf $(\Omega, \mathcal{P}(\Omega))$

$$\begin{aligned} \mathbb{P}(\{(\omega_1, \dots, \omega_n)\}) &= \mathbb{P}(\{\omega_1\}) \prod_{m=2}^n \mathbb{P}(\{\omega_m\} \mid [\omega_1, \dots, \omega_{m-1}]) \\ &= \frac{\prod_{i=0}^{l-1} (s+c \cdot i) \prod_{j=0}^{n-l-1} (w+c \cdot j)}{\prod_{i=0}^n (s+w+c \cdot i)} \end{aligned}$$

mit $l = \sum_{i=1}^n \omega_i$. Definiere wir nun die Zufallsvariable

$$S_n : \Omega \rightarrow \mathbb{N}_0 \text{ mit } (\omega_1, \dots, \omega_n) \mapsto \sum_{i=1}^n \omega_i$$

welche die Anzahl der gezogenen schwarzen Kugeln modelliert, so folgt

$$\mathbb{P}(S_n = l) = \binom{n}{l} \frac{\prod_{i=0}^{l-1} (s + c \cdot i) \prod_{j=0}^{n-l-1} (w + c \cdot j)}{\prod_{i=0}^{n-1} (s + w + c \cdot i)}$$

Mittels $a := s/c$ und $b := w/c$ folgt

$$\mathbb{P}(S_n = l) = \binom{n}{l} \frac{\prod_{i=0}^{l-1} (-a - i) \prod_{j=0}^{n-l-1} (-b - j)}{\prod_{i=0}^{n-1} (-a - b - i)} = \frac{\binom{-a}{l} \binom{-b}{n-l}}{\binom{-a-b}{n}} \text{ mit } l \in \{0, \dots, n\}$$

Dies ist die **Polya-Verteilung** auf $\{0, \dots, n\}$ ($n \in \mathbb{N}$) mit Parametern $a, b > 0$.

Beispiel 3.8

Ein Student beantwortet eine Multiple-Choice-Frage mit vier Antwortmöglichkeiten, eine davon ist richtig. Er kennt die richtige Antwort mit Wahrscheinlichkeit $2/3$. Wenn er die richtige Antwort kennt, so wählt er diese aus. Andernfalls wählt er zufällig (gleichverteilt) eine Antwort. Betrachten wir

$$W = \{\text{richtige Antwort gewusst}\} \text{ und } R = \{\text{Richtige Antwort gewählt}\}$$

Dann gilt

$$\mathbb{P}(W) = \frac{2}{3} \quad \mathbb{P}(R | W) = 1 \quad \mathbb{P}(R | W^C) = \frac{1}{4}$$

Angenommen der Student gibt die richtige Antwort. Mit welcher Wahrscheinlichkeit hat er diese gewusst?
 $\rightarrow \mathbb{P}(W | R) = ?$

Satz 3.9

Sei $(\Omega, \mathcal{F}, \mathbb{P})$ ein Wahrscheinlichkeitsraum und $\Omega = \bigcup_{i \in I} B_i$ eine höchstens abzählbare Zerlegung in paarweise disjunkte Ereignisse $B_i \in \mathcal{F}$.

(1) *Satz von der totalen Wahrscheinlichkeit:* Für alle $A \in \mathcal{F}$ gilt

$$\mathbb{P}(A) = \sum_{i \in I} \mathbb{P}(A | B_i) \cdot \mathbb{P}(B_i) \quad (3.1)$$

(2) *Satz von Bayes:* Für alle $A \in \mathcal{F}$ mit $\mathbb{P}(A) > 0$ und alle $k \in I$ gilt

$$\mathbb{P}(B_k | A) = \frac{\mathbb{P}(A | B_k) \cdot \mathbb{P}(B_k)}{\sum_{i \in I} \mathbb{P}(A | B_i) \cdot \mathbb{P}(B_i)} \quad (3.2)$$

Beweis. Es gilt:

$$\sum_{i \in I} \mathbb{P}(A | B_i) \cdot \mathbb{P}(B_i) = \sum_{i \in I} \frac{\mathbb{P}(A \cap B_i)}{\mathbb{P}(B_i)} \cdot \mathbb{P}(B_i) = \sum_{i \in I} \mathbb{P}(A \cap B_i) \stackrel{\sigma\text{-Add.}}{=} \mathbb{P}(A)$$

und

$$\mathbb{P}(B_k | A) = \frac{\mathbb{P}(A \cap B_k)}{\mathbb{P}(A)} = \frac{\mathbb{P}(A | B_k) \cdot \mathbb{P}(B_k)}{\mathbb{P}(A)}$$

also folgt (b) aus (a). □

Beispiel 3.10

In der Situation von Beispiel 3.8 folgt mit (3.1)

$$\begin{aligned}\mathbb{P}(R) &= \mathbb{P}(R \mid W) \cdot \mathbb{P}(W) + \mathbb{P}(R \mid W^C) \cdot \mathbb{P}(W^C) \\ &= 1 \cdot \frac{2}{3} + \frac{1}{4} \cdot \frac{1}{3} = \frac{3}{4}\end{aligned}$$

und mit (3.2)

$$\mathbb{P}(W \mid R) = \frac{\mathbb{P}(R \mid W) \cdot \mathbb{P}(W)}{\mathbb{P}(R)} = \frac{1 \cdot \frac{2}{3}}{\frac{3}{4}} = \frac{8}{9}$$

für die gesuchte Wahrscheinlichkeit.

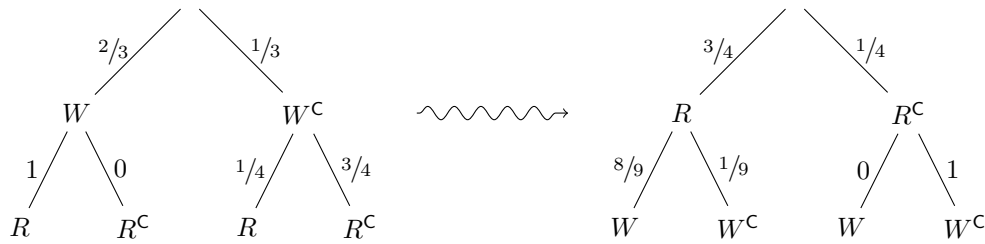


Abbildung 3.2: Bedingte Wahrscheinlichkeit im Baumdiagramm

3.2 (Un)abhängigkeit

In vielen Fällen besagt die Intuition über verschiedene Zufallsexperimente / Ereignisse, dass diese sich *nicht* gegenseitig beeinflussen. Für solche $A, B \in \mathcal{F}$ mit $\mathbb{P}(A) > 0$ und $\mathbb{P}(B) > 0$ sollte gelten

$$\mathbb{P}(A \mid B) = \mathbb{P}(A) \quad \text{und} \quad \mathbb{P}(B \mid A) = \mathbb{P}(B).$$

Definition 3.11

Sei $(\Omega, \mathcal{F}, \mathbb{P})$ Wahrscheinlichkeitsraum. Zwei Ereignisse $A, B \in \mathcal{F}$ heißt **(stochastisch) unabhängig** bezüglich \mathbb{P} , falls

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

Wir schreiben auch $A \perp\!\!\!\perp B$.

Beispiel 3.12

Wir betrachten wieder das Würfeln mit 2 fairen, sechsseitigen Würfeln.

$$\Omega = \{(i, j) \mid i, j \in \{1, \dots, n\}\} \quad \mathcal{F} = \mathcal{P}(\Omega) \quad \mathbb{P} = \mathbb{U}(\Omega)$$

Betrachte

$$A := \{(i, j) \in \Omega : i \text{ gerade}\}$$

$$B := \{(i, j) \in \Omega : j \leq 2\}$$

In diesem Fall, erwarten wir intuitiv die Unabhängigkeit von A und B . In der Tat ist $\mathbb{P}(A) = 1/2$, $\mathbb{P}(B) = 1/3$ und $\mathbb{P}(A \cap B) = 1/6$, womit $\mathbb{P}(A \cap B) = \mathbb{P}(A) \cdot \mathbb{P}(B)$ erfüllt ist. Betrachte nun

$$C := \{(i, j) \in \Omega : i + j = 7\}$$

$$D := \{(i, j) \in \Omega : i = 6\}$$

Dann gilt $\mathbb{P}(C) = 1/6$ und $\mathbb{P}(D) = 1/6$. Wegen $C \cap D = \{(6, 1)\}$ folgt

$$\mathbb{P}(C \cap D) = \frac{1}{36} = \frac{1}{6} \cdot \frac{1}{6} = \mathbb{P}(C) \cdot \mathbb{P}(D)$$

C und D sind also *stochastisch* unabhängig, obwohl eine kausale Abhängigkeit vorliegt.

Definition 3.13

Sei $(\Omega, \mathcal{F}, \mathbb{P})$ ein Wahrscheinlichkeitsraum und $I \neq \emptyset$ eine endliche Indexmenge. Dann heißt die Familie $(A_i)_{i \in I}$ von Ereignissen in \mathcal{F} **unabhängig** bezüglich \mathbb{P} , falls für alle $\emptyset \neq J \subseteq I$ gilt

$$\mathbb{P}\left(\bigcap_{i \in J} A_i\right) = \prod_{i \in J} \mathbb{P}(A_i)$$

Offensichtlich impliziert die Unabhängigkeit einer Familie die paarweise Unabhängigkeit je zweier Familienmitglieder nach Definition 3.11. Umgekehrt gilt dies nicht.

Beispiel 3.14 (Abhängigkeit trotz paarweiser Unabhängigkeit)

Wir betrachten ein zweifaches Bernoulliexperiment mit Erfolgswahrscheinlichkeit $1/2$, d.h.

$$\Omega = \{0, 1\}^2 \quad \mathcal{F} = \mathcal{P}(\Omega) \quad \mathbb{P} = \mathbb{U}(\Omega)$$

sowie

$$\begin{aligned} A &= \{1\} \times \{0, 1\} && \text{(Münzwurf: erster Wurf ist Zahl)} \\ B &= \{0, 1\} \times \{1\} && \text{(Münzwurf: zweiter Wurf ist Zahl)} \\ C &= \{(0, 0), (1, 1)\} && \text{(beide Würfe haben selbes Ergebnis)} \end{aligned}$$

Dann gelten $\mathbb{P}(A) = 1/2 = \mathbb{P}(B) = \mathbb{P}(C)$ und

$$\begin{aligned} \mathbb{P}(A \cap B) &= \mathbb{P}(\{(1, 1)\}) = 1/4 = \mathbb{P}(A) \cdot \mathbb{P}(B) \\ \mathbb{P}(A \cap C) &= \mathbb{P}(\{(1, 1)\}) = 1/4 = \mathbb{P}(A) \cdot \mathbb{P}(C) \\ \mathbb{P}(B \cap C) &= \mathbb{P}(\{(1, 1)\}) = 1/4 = \mathbb{P}(B) \cdot \mathbb{P}(C) \end{aligned}$$

Daraus folgt also paarweise Unabhängigkeit. Jedoch ist

$$\mathbb{P}(A \cap B \cap C) = \mathbb{P}(\{(1, 1)\}) = \frac{1}{4} \neq \mathbb{P}(A) \cdot \mathbb{P}(B) \cdot \mathbb{P}(C)$$

und A, B, C sind *nicht* stochastisch unabhängig.

Definition 3.15 (Unabhängige σ -Algebren)

Seien $(\Omega, \mathcal{F}, \mathbb{P})$ ein Wahrscheinlichkeitsraum, $I \neq \emptyset$ eine Indexmenge und (E_i, \mathcal{E}_i) Messräume.

(1) Die Familie $\mathcal{F}_i \subset \mathcal{F} (i \in I)$, heißen **unabhängig**, wenn für die $\emptyset \neq J \subseteq I$ mit $|J| < \infty$ gilt

$$\mathbb{P}\left(\bigcap_{i \in J} A_i\right) = \prod_{i \in J} \mathbb{P}(A_i) \quad \text{für beliebige } A_i \in \mathcal{F}_i (i \in J)$$

(2) Die Zufallsvariable $X_i: (\Omega, \mathcal{F}) \rightarrow (E_i, \mathcal{E}_i) (i \in I)$, heißen **unabhängig**, wenn die σ -Algebren

$$\sigma(X_i) = X_i^{-1}(\mathcal{E}_i) = \{\{X_i \in F\} : F \in \mathcal{E}_i\} \quad (i \in I)$$

unabhängig sind.