Einführung in die Stochastik - Mitschrieb

Vorlesung im Wintersemester 2011/2012

Sarah Lutteropp

15. November 2011

Inhaltsverzeichnis

Vorwort

Dies ist ein Mitschrieb der Vorlesung “Einführung in die Stochastik” vom Wintersemester 2011/2012 am Karlsruher Institut für Technologie, die von Herrn Prof. Dr. Günther Last gehalten wird.

Kapitel 1

Deskriptive Statistik

1.1 Der Grundraum

$\emptyset \neq \Omega =$ Grundraum (Grundgesamtheit, Merkmalsraum, Stichprobenraum)
Annahme: Ω ist diskret(endlich oder abzählbar unendlich) (Häufig $\Omega \subseteq \mathbb{R}$)

1.2 Absolute und relative Häufigkeit

$x_1, \ldots, x_n \in \Omega$ ("Daten")
$h(\omega) = \text{card}\,\{j \in \{1, \ldots, n\} \colon x_j = \omega\}\,, \omega \in \Omega$, absolute Häufigkeit von ω

Bemerkung $\sum\limits_{\omega \in \Omega} h(\omega) = n$

Definition $\frac{1}{n} h(\omega) =$ relative Häufigkeit von ω
$h(A) = \text{card}\,\{j \in \{1, \ldots, n\} \colon x_j \in A\}\,, A \subset \Omega =$ absolute Häufigkeit von A,
$\frac{1}{n} h(A) =$ relative Häufigkeit von A

1.3 Histogramm

$x_1, \ldots, x_n \in \mathbb{R}, b_1 < b_2 < \ldots < b_s$ mit $b_1 \leq \min\limits_{1 \leq i \leq n} x_i, b_s > \max\limits_{1 \leq i \leq n} x_i$
TODO: BILD
$d_j(b_{j+1} - b_j) = h([b_j, b_{j+1})) = \text{card}\,\{i \in \{1, \ldots, n\} \colon b_j \leq x_i < b_{j+1}\}$

1.4 Lagemaße

Definition Ein **Lagemaß** ist eine Abbildung $l \colon \mathbb{R}^n \to \mathbb{R}$ mit

$$l(x_1 + a, \ldots, x_n + a) = l(x_1, \ldots, x_n) + a$$

"Verschiebungskovarianz". $x_1, \ldots, x_n, a \in \mathbb{R}$

1.4.1 Arithmetisches Mittel

$x_1, \ldots, x_n \in \mathbb{R}, \bar{x} := \frac{1}{n} \sum\limits_{j=1}^{n} x_j$ "Schwerpunkt der Daten"

Fakt $\sum\limits_{j=1}^{n} (x_i - t)^2 \xrightarrow{t} \text{Min}$
Lösung: $t = \bar{x}$
"Prinzip der kleinsten Quadrate"

Beweis $\frac{1}{n} \sum\limits_{j=1}^{n} (x_j - t)^2 = t^2 - 2\bar{x}t + \frac{1}{n} \sum\limits_{j=1}^{n} x_j^2 = (t - \bar{x})^2 + \frac{1}{n} \sum\limits_{j=1}^{n} x_j^2 - (\bar{x})^2$

1.4.2 Median, Quantile

$x_1, \ldots, x_n \in \mathbb{R} \Rightarrow x_{(1)} \leq x_{(2)} \leq \ldots \leq x_{(n)}$ geordnete Stichprobe

Definition

$$x_{1/2} := \begin{cases} x_{(\frac{n+1}{2})} & \text{, falls } n \text{ ungerade} \\ \frac{1}{2}(x_{(\frac{n}{2})} + x_{(\frac{n}{2}+1)}) & \text{, falls } n \text{ gerade} \end{cases}$$

heißt **Median** von $x_1, \ldots, x_n$.

Fakt $\sum\limits_{j=1}^{n} |x_j - x_{1/2}| = \min\limits_{t} \sum\limits_{j=1}^{n} |x_j - t|$ Übungsaufgabe

Bemerkung Der Median ist "robust" gegenüber "Ausreißern". Ist etwa $x_1 = \ldots = x_9 = 1$ und $x_{10} = 1000 (n = 10)$, so gilt $\bar{x} = 100,9, x_{1/2} = 1$

Definition Für 0 < p < 1 heißt

$$x_p := \begin{cases} x_{(\lfloor n \cdot p+1 \rfloor)} & \text{, falls } n \cdot p \notin \mathbb{N} \\ \frac{1}{2}(x_{(n \cdot p)} + x_{(n \cdot p+1)}) & \text{, falls } n \cdot p \in \mathbb{N} \end{cases}$$

p-Quantil von $x_1, \ldots, x_n$.

Interpretation Mindestens $p \cdot 100\%$ der Daten liegen links von x_p und mindestens $(1 - p) \cdot 100\%$ liegen rechts von x_p.
$x_{1/4}$ = unteres Quartil, $x_{3/4}$ = oberes Quartil

1.5 Streuungsmaße

Definition Eine Abbildung $\sigma\colon \mathbb{R}^n \to \mathbb{R}$ mit

$$\sigma(x_1 + a, \ldots, x_n + a) = \sigma(x_1, \ldots, x_n) \text{ (Translationsinvarianz)}$$

heißt **Streuungsmaß**.

1.5.1 Empirische Varianz

$s^2 := \frac{1}{n-1} \sum_{j=1}^{n} (x_j - \bar{x})^2$ = **empirische Varianz** von $x_1, \ldots, x_n$

1.5.2 Empirische Standardabweichung

$s := +\sqrt{s^2}$ = **empirische Standardabweichung** von $x_1, \ldots, x_n$

1.5.3 Spannweite

$x_{(n)} - x_{(1)}$ = **Spannweite** von $x_1, \ldots, x_n$

1.5.4 Quartilsabstand

$x_{(3/4)} - x_{(1/4)}$ = **Quartilsabstand** von $x_1, \ldots, x_n$

1.6 Empirischer Korrelationskoeffizient

$(x_1, y_1), \ldots, (x_n, y_n) \in \mathbb{R}^2$ TODO: BILD

Gesucht: Gerade $y = a + b \cdot x$ so, dass

$$(*) \sum_{j=1}^{n} (y_j - a - bx_j)^2 \xrightarrow{a,b} \text{Min}$$

Definition $\sigma_x^2 = \frac{1}{n} \sum_{j=1}^{n} (x_j - \bar{x})^2$ $\sigma_y^2 = \frac{1}{n} \sum_{j=1}^{n} (y_j - \bar{y})^2$

$\sigma_{xy} = \frac{1}{n} \sum_{j=1}^{n} (x_j - \bar{x})(y_j - \bar{y})$ **empirische Kovarianz** $\sigma_x^2 > 0, \sigma_y^2 > 0$.

Lösung von (*): $b^* = \frac{\sigma_{xy}}{\sigma_{x^2}}, a^* = \bar{y} - b^* \cdot \bar{x}$

$\min_{a,b} \sum_{j=1}^{n} (y_j - a - bx_j)^2 \stackrel{!}{=} \min_{b} \sum_{j=1}^{n} (y_i - \bar{y} - b(x_j - \bar{x}))^2 = \ldots$

"lineare Regression"

Einsetzen von a^* und b^* in die Zielfunktion:

$$0 \leq \sum_{j=1}^{n} (y_j - a^* - b^* x_j)^2 = \ldots = n\sigma_y^2 (1 - (\frac{\sigma_{xy}}{\sigma_x \sigma_y})^2)$$

Definition $r_{xy} := \frac{\sigma_{xy}}{\sigma_x \sigma_y}$ heißt **empirischer Korrelationskoeffizient** (*Pearson*).

Folgerung $|r_{xy}| \leq 1$
Es gilt $r_{xy} = \pm 1 \Leftrightarrow$ Punktewolke liegt exakt auf der Geraden $y = a^* + b^* x$.
Dabei ist $b^* > 0$, falls $r_{xy} = 1$ und $b^* < 0$, falls $r_{xy} = -1$.
Dieser empirische Korrelationskoeffizient ist ein Maß für die (affin) lineare Abhängigkeit zwischen den x_j *und den* y_j.

Kapitel 2

Ereignisse und Zufallsvariablen

2.1 Definition

Gegeben sei eine Grundmenge Ω. Die Elemente von Ω heißen **Elementarereignisse**. Teilmengen von Ω heißen **Ereignisse**. (Idee: $\omega \in \Omega$ ist Ausgang eines zufälligen Versuchs.)

Interpretation Ein Ereignis $A \subset \Omega$ "tritt ein", wenn $\omega \in A$.

2.2 Beispiele

- (i) (Münzwurf)
  $\Omega = \{0, 1\}$(oder $\Omega = \{W, Z\}$)
- (ii) (m Münzwürfe)
  $\Omega = \{0, 1\}^m (A = \{\omega = (\omega_1, \dots, \omega_m) : \sum_{j=1}^{m} \omega_j \geq k\}$ Ereignis)
- (iii) Werfen von 2 Würfeln
  $\Omega = \{1, \dots, 6\}^2$
- (iv) Brownsche Bewegung
  (TODO: BILD) Bewegung eines Blütenpollens in einer Flüssigkeit
  $\Rightarrow$ Zukunftsmusik
  $\Omega = C([0, 1], \mathbb{R}^2)$

2.3 Bemerkung (Mengentheoretische Operationen)

Seien $A, B, A_1, A_2, \ldots \subset \Omega$.
$A \cap B = \{\omega \in \Omega : \omega \in A \text{ und } \omega \in B\} \mathrel{\hat{=}}$ "A und B treten ein"
$A \cup B \mathrel{\hat{=}}$ "A oder B treten ein"
$\bar{A} \equiv A^c := \Omega \backslash A = \{\omega \in \Omega : \omega \notin A\} \mathrel{\hat{=}}$ "A tritt nicht ein"

$A \backslash B = A \cap B^c \mathrel{\widehat{=}}$ "A tritt ein, aber nicht B"
$A \subset B \mathrel{\widehat{=}}$ "wenn A, dann B"
$\emptyset \mathrel{\widehat{=}}$ "unmögliches Ereignis"
$\Omega \mathrel{\widehat{=}}$ "sicheres Ereignis"

Abkürzung $AB = A \cap B$

2.4 Definition

Eine Abbildung $X \colon \Omega \to \mathbb{R}$ heißt (reelle) **Zufallsvariable**. Für $\omega \in \Omega$ heißt $X(\omega)$ **Realisierung** der Zufallsvariable zu ω.

Idee Mit $\omega \in \Omega$ bekommt auch $X(\omega)$ einen zufälligen Charakter.

Definition $X^{-1} \colon \mathcal{P}(\mathbb{R}) \to \mathcal{P}(\Omega) = \{A \colon A \in \Omega\}$ ist definiert durch

$$X^{-1}(A) = \{\omega \in \Omega \colon X(\omega) \in A\} \text{ ("Urbild von A unter X")}$$

Bemerkung

- $X^{-1}(A \cap B) = X^{-1}(A) \cap X^{-1}(B), A, B \subset \mathbb{R}$
- $X^{-1}(A \cup B) = X^{-1}(A) \cup X^{-1}(B)$
- $X^{-1}(\bigcup\limits_{j=1}^{\infty} A_j) = \bigcup\limits_{j=1}^{\infty} X^{-1}(A_j)$
- $X^{-1}(\bigcap\limits_{j=1}^{\infty} A_j) = \bigcap\limits_{j=1}^{\infty} X^{-1}(A_j)$

Vereinbarung Es sei X eine Zufallsvariable und $t \in \mathbb{R}$. Wir setzen

- $\{X = t\} := \{\omega \colon X(\omega) = t\} (= X^{-1}(t))$
- $\{X \geq t\} := \{\omega \colon X(\omega) \geq t\}$

2.5 Definition

Sind X, Y Zufallsvariablen, so definiert man

- $(X + Y)(\omega) = X(\omega) + Y(\omega)$
- $(X - Y)(\omega) = X(\omega) - Y(\omega)$
- $(X \cdot Y)(\omega) = X(\omega) \cdot Y(\omega)$

$\omega \in \Omega$, neue Zufallsvariablen $X+Y, X-Y, X \cdot Y$
analog für $a \in \mathbb{R}$

- $aX(\omega) = a \cdot (X(\omega))$
- $\min(X,Y) = (X \wedge Y)(\omega) := \min\{X(\omega), Y(\omega)\} \ldots$

2.6 Definition

Sei $A \subset \Omega$. Die Funktion $1_A \colon \Omega \to \mathbb{R}$ ist definiert durch

$$1_A(\omega) = \begin{cases} 1 & \text{, falls } \omega \in A \\ 0 & \text{, falls } \omega \notin A \end{cases}$$

und heißt **Indikatorfunktion** von A.

2.7 Bemerkungen (Rechenregeln für Indikatorfunktionen)

- $1_\emptyset \equiv 0$
- $1_\Omega \equiv 1$
- $(1_A)^2 = 1_A$
- $1_{A^c} = 1 - 1_A$
- $1_{A \cap B} = 1_A \cdot 1_B$
- $1_{A \cup B} = 1_A + 1_B - 1_{A \cap B}$
- $A \subset B \Leftrightarrow 1_A \leq 1_B$
- $1_{A \triangle B} = |1_A - 1_B|$

2.8 Definition

Seien $A_1, \ldots, A_n \subset \Omega$. Die Zufallsvariable

$$X := \sum_{j=1}^{n} 1_{A_j}$$

heißt **Zählvariable** oder **Indikatorsumme**.

Bemerkung

- $\{X = 0\} = \{\omega\colon X(\omega) = 0\} = A_1^c \cap \ldots A_n^c$
- $\{X = n\} = A_1 \cap \ldots \cap A_n$
- $\{X = k\}$ = "genau k der Ereignisse $A_1, \ldots, A_n$ treten ein" $= \bigcup\limits_{T \subset \{1,\ldots,n\}, |T|=k} \big(\bigcap\limits_{j \in T} A_j \cap \bigcap\limits_{j \notin T} A_j^c\big)$
  $(T \subset \{1, \ldots, n\}, |T| = \text{card } T = k)$

Kapitel 3

Diskrete Wahrscheinlichkeitsräume

3.1 Motivation

Zufallsexperiment mit Ausgängen in Ω
n-malige, 'unabhängige' Wiederholung
$\Rightarrow$ Ergebnis $(a_1, \dots, a_n) \in \Omega^n$
$r_n(A) := \frac{1}{n} \sum\limits_{j=1}^{n} 1_A(a_j), A \subset \Omega$ relative Häufigkeit von A
$0 \leq r_n(A) \leq 1, r_n(\emptyset) = 0, r_n(\Omega) = 1$
$r_n(A \cup B) = r_n(A) + r_n(B), A \cap B = \emptyset$
empirisches Gesetz über Stabilisierung relativer Häufigkeiten:
$r_n(A) \underset{n \to \infty}{\rightsquigarrow} ?$

3.2 Definition

Ein Paar $(\Omega, \mathbb{P})$ bestehend aus einer diskreten Menge $\Omega \neq \emptyset$ und einer Funktion $\mathbb{P} \colon \mathcal{P} \to \mathbb{R}$ heißt **diskreter Wahrscheinlichkeitsraum**, falls:

- (P1) $\mathbb{P}(A) \geq 0, A \subset \Omega$
- (P2) $\mathbb{P}(\Omega) = 1$
- (P3) $\mathbb{P}(\bigcup\limits_{j=1}^{\infty} A_j) = \sum\limits_{j=1}^{\infty} \mathbb{P}(A_j), A_i \cap A_j = \emptyset, i \neq j$
  Diese Eigenschaft heißt σ-Additivität.

Man nennt $\mathbb{P}$ **Wahrscheinlichkeitsmaß (auf Ω)** (oder Wahrscheinlichkeitsverteilung) und $\mathbb{P}(A)$ heißt **Wahrscheinlichkeit von A**.

3.3 Folgerung

- a) $\mathbb{P}(\emptyset) = 0$
- b) $\mathbb{P}(\bigcup\limits_{j=1}^{n} A_j) = \sum\limits_{j=1}^{n} \mathbb{P}(A_j), A_i \cap A_j = \emptyset, i \neq j$
- c) $0 \leq \mathbb{P}(A) \leq 1, A \subset \Omega$
- d) $\mathbb{P}(A \cup B) = \mathbb{P}(A) + \mathbb{P}(B) - \mathbb{P}(A \cap B), A, B \subset \Omega$
- e) $A \subset B \Rightarrow \mathbb{P}(A) \leq \mathbb{P}(B)$ (Monotonie)
- f) $\mathbb{P}(A^c) = 1 - \mathbb{P}(A)$ (Komplementärwahrscheinlichkeit)
- g) $\mathbb{P}(\bigcup\limits_{j=1}^{\infty} A_j) \leq \sum\limits_{j=1}^{\infty} A_j$ (Subadditivität)
- h) $A_n \subset A_{n+1}, n \in \mathbb{N} \Rightarrow \mathbb{P}(\bigcup\limits_{n=1}^{\infty} A_n) = \lim\limits_{n\to\infty} \mathbb{P}(A_n)$ (Stetigkeit von unten)
- i) $A_n \supset A_{n+1}, n \in \mathbb{N} \Rightarrow \mathbb{P}(\bigcap\limits_{n=1}^{\infty} A_n) = \lim\limits_{n\to\infty} \mathbb{P}(A_n)$ (Stetigkeit von oben)

Beweis • a): $A_j = \emptyset, j \in \mathbb{N}$ (P3) [1] $\mathbb{P}(\emptyset) = 0$.
• b): $A_{n+1} = A_{n+2} = \ldots = \emptyset$ in P3!
• c) + f): Für $A \subset \Omega$ gilt nach b) (für n = 2):
$1 = \mathbb{P}(\Omega) = \mathbb{P}(A \cup A^c) \overset{(b)}{=} \mathbb{P}(A) + \mathbb{P}(A^c)$
• d): Nach b) gilt $\mathbb{P}(A) = \mathbb{P}(A \backslash B) + \mathbb{P}(A \cap B), \mathbb{P}(B) = \mathbb{P}(B \backslash A) + \mathbb{P}(A \cap B)$ und somit $\mathbb{P}(A) + \mathbb{P}(B) - \mathbb{P}(A \cap B) = \mathbb{P}(A \backslash B) + \mathbb{P}(B \backslash A) + \mathbb{P}(A \cap B) \overset{(b)}{=} \mathbb{P}(A \cup B)$
• e): Wegen $B = A \cup (B \backslash A)$ folgt[2] $\mathbb{P}(B) = \mathbb{P}(A) + \mathbb{P}(B \backslash A) \geq \mathbb{P}(A)$
• g): $B_1 := A_1, B_2 := A_2 \backslash A_1, \ldots, B_n := A_n \backslash (\bigcup\limits_{j=1}^{n-1} A_j), n \geq 2$.
Dann gilt $B_n \subset A_n$ und $\bigcup\limits_{j=1}^{n} B_j = \bigcup\limits_{j=1}^{n} A_j$ sowie $B_i \cap B_j = \emptyset, i \neq j$.
Es folgt aus (P3):
$\mathbb{P}(\bigcup\limits_{j=1}^{\infty} A_j) \overset{!}{=} \mathbb{P}(\bigcup\limits_{j=1}^{n} B_j) \overset{(P3)}{=} \sum\limits_{j=1}^{n} \mathbb{P}(B_j) \overset{e)}{\leq} \sum\limits_{j=1}^{n} \mathbb{P}(A_j)$ (∞ ist zugelassen)
• h) + i): Übungsaufgabe

[1] $\mathbb{P}(\emptyset) = \mathbb{P}(\emptyset \cup \emptyset) = \mathbb{P}(\emptyset) + \mathbb{P}(\emptyset) = 2 \cdot \mathbb{P}(\emptyset)$
[2] (aus der Additivität)

3.4 Satz

Seien $A_1, \dots, A_n \subset \Omega$. Setze

$$S_k := \sum_{1 \le i_1 < \ldots < i_k \le n} \mathbb{P}(A_{i_1} \cap \ldots \cap A_{i_k})$$

Dann gilt

- a) $\mathbb{P}(\bigcup\limits_{j=1}^{n} A_j) = \sum\limits_{k=1}^{n} (-1)^{k-1} S_k$ 'Siebformel'
- b) $\mathbb{P}(\bigcup\limits_{j=1}^{n} A_j) \le \sum\limits_{k=1}^{2s+1} (-1)^{k-1} S_k, s = 0, \ldots, \lfloor \frac{n-1}{2} \rfloor$
  $\mathbb{P}(\bigcup\limits_{j=1}^{n} A_j) \ge \sum\limits_{k=1}^{2s} (-1)^{k-1} S_k, s = 1, \ldots, \lfloor \frac{n}{2} \rfloor$

Beweisidee für Siebformel vollständige Induktion nach n:
n=2: $\mathbb{P}(A_1 \cup A_2) \overset{(d)}{=} \mathbb{P}(A_1) + \mathbb{P}(A_2) - \mathbb{P}(A_1 \cap A_2) = S_1 - S_2$
n=3: $\mathbb{P}(\underbrace{A_1 \cup A_2} \cup A_3) \overset{(d)}{=} \mathbb{P}(A_1 \cup A_2) + \mathbb{P}(A_3) - \mathbb{P}((A_1 \cup A_2) \cap A_3)$ [3]
$\overset{(d)}{=} \mathbb{P}(A_1) + \mathbb{P}(A_2) - \mathbb{P}(A_1 \cap A_2) + \mathbb{P}(A_3) - \mathbb{P}(A_1 \cap A_3) - \mathbb{P}(A_2 \cap A_3) + \mathbb{P}(A_1 \cap A_2 \cap A_3) = S_1 - S_2 + S_3$

3.5 Definition + Satz

a) Sei $(\Omega, \mathbb{P})$ diskreter Wahrscheinlichkeitsraum. Dann heißt $p\colon \Omega \to \mathbb{R}$ definiert durch $p(\omega) := \mathbb{P}(\{\omega\})$ **Wahrscheinlichkeitsfunktion** (von $\mathbb{P}$).
Es gilt $\mathbb{P}(A) = \sum\limits_{\omega \in A} p(\omega), A \subset \Omega$.
b) Sind Ω diskret und $p\colon \Omega \to \mathbb{R}$ eine Abbildung mit $p(\omega) \ge 0$ und $\sum\limits_{\omega \in \Omega} p(\omega) = 1$, so erhält man vermöge $\mathbb{P}(A) := \sum\limits_{\omega \in A} p(\omega)$ einen diskreten Wahrscheinlichkeitsraum.

Beweis • a) σ-Additivität $(A = \bigcup\limits_{\omega \in A} \{\omega\})$
• b) σ-Additivität: Großer Umordnungssatz! (Analysis)

3.6 Definition

$|\Omega| =: n < \infty$. Definiere $\mathbb{P}(A) = \frac{|A|}{n}$. Dann heißt $(\Omega, \mathbb{P})$ (ein diskreter Wahrscheinlichkeitsraum!) **Laplace-Raum**. Man nennt $\mathbb{P}$ **Gleichverteilung** auf Ω.
('homogene Münze', 'Würfeln', ...)

[3] $(A_1 \cup A_2) \cap A_3 = (A_1 \cap A_3) \cup (A_2 \cap A_3)$

3.7 Definition

Sei $\Omega \neq \emptyset$ beliebig! $(\Omega, \mathbb{P})$ diskreter Wahrscheinlichkeitsraum $\Leftrightarrow$ $\exists$ abzählbare Menge $\Omega_0 \subset \Omega$, $\exists p \colon \Omega \to [0, \infty)$ mit $p(\omega = 0)$ für alle $\omega \notin \Omega_0$, und $\sum\limits_{\omega \in \Omega_0} p(\omega) = 1$, und $\mathbb{P}(A) = \sum\limits_{\omega \in A \cap \Omega_0} p(\omega)$, $A \subset \Omega$.

Wiederholung $(\Omega, \mathbb{P})$ Wahrscheinlichkeitsraum
$p \colon \Omega \longrightarrow [0,1]$, $\sum\limits_{\omega \in \Omega} p(\omega) = 1$
$\mathbb{P}(A) := \sum\limits_{\omega \in A} p(\omega), A \subset \Omega$
$p(\omega) := \mathbb{P}(\{\omega\})$

Ω allgemeine Menge, $\Omega_0 \subset \Omega$ diskret
$p \colon \Omega \longrightarrow [0,1]$, $\sum\limits_{\omega \in \Omega_0} p(\omega) = 1, p(\omega) = 0, \omega \notin \Omega_0$
$\mathbb{P}(A) := \sum\limits_{\omega \in A} p(\omega) := \sum\limits_{\omega \in A \cap \Omega_0} p(\omega)$
$\Omega_0 =$ Träger von $\mathbb{P}$

3.8 Definition

$(\Omega, \mathbb{P})$ diskreter Wahrscheinlichkeitsraum mit Träger Ω_0. Es sei $X \colon \Omega \to \mathbb{R}$ eine Zufallsvariable. Dann heißt die Funktion $\mathbb{P}^X : \mathbb{P}(\mathbb{R}) \to \mathbb{R}$ definiert durch $\mathbb{P}^X(B) := \mathbb{P}\left(X^{-1}(B)\right), B \subset \mathbb{R}$ **Verteilung um** X.

3.9 Satz

In der Situation von Definition 3.8 ist $(\mathbb{R}, \mathbb{P}^X)$ ein diskreter Wahrscheinlichkeitsraum mit Träger $B_0 := X(\Omega_0) = \{X(\omega) \colon \omega \in \Omega_0\}$

Beweis. Für $B \subset \mathbb{R}$.

$$\begin{aligned} \mathbb{P}^X(B) &= \mathbb{P}(\{\omega \colon X(\omega) \in B\}) \\ &\overset{!}{=} \mathbb{P}(\{\omega \colon X(\omega) \in B \cap B_0\}) \end{aligned}$$

Definiert man für $t \in \mathbb{R}$

$$p_t = \mathbb{P}(\{\omega \colon X(\omega) = t\}) = \mathbb{P}(X = t)$$

so ergibt sich aus der σ-Additivität von $\mathbb{P}$

$$\mathbb{P}^X(B) = \sum_{t \in B \cap B_0} \mathbb{P}(\{\omega \colon X(\omega) = t\}) = \sum_{t \in B \cap B_0} p_t$$

$\square$

Kapitel 4

Kombinatorik

$|A| = card(A) =$ Anzahl der Elemente einer endlichen Menge A

4.1 Grundregeln

$A_1, \ldots, A_k$ endliche Menge

(i) $A_i \cap A_j = \emptyset, i \neq j \Rightarrow \left|\bigcup\limits_{j=1}^{n} A_j\right| = \sum\limits_{j=1}^{n} A_j$

(ii) $|A_1 \times \ldots \times A_n| = \prod\limits_{j=1}^{k} |A_j|$

4.2 Satz

Es sollen k-Tupel $(a_1, \ldots, a_k)$ durch sukzessives Festlegen von $a_1, a_2, \ldots, a_k$ nach folgenden Regeln gebildet werden:

- es gibt j_1 Möglichkeiten für die Wahl von a_1
- es gibt (dann) j_2 Möglichkeiten für die Wahl von a_2
- $\ldots$
- es gibt (dann) j_k Möglichkeiten für die Wahl von a_k

Dann gibt es genau $j_1 \cdot \ldots \cdot j_k$ solcher Tupel.

4.3 Beispiel (Urnenmodelle)

Betrachte Urne mit n durchnummerierten Kugeln.Es werden k Kugeln nach folgenden Regeln gezogen: $(M := \{1, \ldots, n\})$

Zurücklegen (Wiederholung) \ Beachtung der Reihenfolge	ja	nein
ja	k-Permutationen aus M mit Wiederholung, Per_k^n	k-Kombinationen aus M mit Wiederholung, Kom_k^n
nein	k-Permutationen aus M ohne Wiederholung, $Per_{k,\neq}^n$	k-Kombinationen aus M ohne Wiederholung, $Kom_{k,\neq}^n$

4.4 Definition

$M = \{1, \ldots, n\} (n \in \mathbb{N})$

- $Per_k^n := M^k$
- $Per_{k,\neq}^n := \{(a_1, \ldots, a_k) \in M^k \colon a_i \neq a_j \text{ für } i \neq j\}$
- $Kom_k^n := \{(a_1, \ldots, a_k =\in M^k \colon a_1 \leq a_2 \leq \ldots \leq a_k\}$
- $(Kom_{k,\neq}^n := \{(a_1, \ldots, a_k) \in M^k \colon a_1 < a_2 < \ldots < a_k\}$

4.5 Satz

- (i) $|Per_k^n| = n^k$
- (ii) $|Per_{k,\neq}^n| := n^{\underline{k}} = n \cdot (n-1) \cdot \ldots \cdot (n-k+1)$
- (iii) $|Kom_k^n| = \binom{n+k-1}{k}$
- (iv) $|Kom_{k,\neq}^n| = \binom{n}{k}$

Beweis. (i): 4.1.(ii)
(ii) Satz 4.2
(iv) Betrachte Äquivalenzrelation

$$(a_1, \ldots, a_k) \sim (b_1, \ldots, b_k) \Leftrightarrow \{a_1, \ldots, a_k\} = \{b_1, \ldots, b_k\}$$

auf $Per_{k,\neq}^n$. Jede Äquivalenzklasse hat $k!$ Elemente! Es folgt

$$|Kom_{k,\neq}^n| \cdot k! = |Per_{k,\neq}^n| = n^{\underline{k}}$$

(iii) Die Abbildung $g \colon Kom_k^n \to Kom_{k,\neq}^{n+k-1}$ definiert durch

$$(a_1, \ldots, a_k) \mapsto (a_1, a_2 + 1, a_3 + 2, \ldots, a_k + k - 1)$$

ist eine Bijektion! (Umkehrabbildung!) Es folgt(!)

$$|Kom_k^n| = |Kom_{k,\neq}^{n+k-1}| = \binom{n+k-1}{k}$$

□

4.6 Beispiel (Geburtstagsproblem)

Wie groß ist die Wahrscheinlichkeit, dass unter k rein zufällig ausgewählten Personen mindestens zwei am selben Tag Geburtstag haben?

Antwort Betrachte $\Sigma = Per_k^n$ mit $n = 365$, und der Laplace-Verteilung. Es sei $A := \{(a_1, \dots, a_k) \in \Omega\colon$ es gibt $i, j \in \{1, \dots, k\}$ mit $i \neq j, a_i = a_j\}$. Es gilt

$$\begin{aligned}
\mathbb{P}(A) &= 1 - \mathbb{P}(A^c) \\
&= 1 - \mathbb{P}(Per^n_{k,\neq}) \\
&\overset{!}{=} 1 - \frac{|Per^n_{k,\neq}|}{card \Omega} \\
&= 1 - \frac{n \cdot (n-1) \cdot \ldots \cdot (n-k+1)}{n^k} \\
&= 1 - \frac{n}{n} \cdot \frac{(n-1)}{n} \cdot \ldots \cdot \frac{n-k+1}{n} \\
&= 1 - (1 - \frac{1}{n}) \cdot \ldots \cdot (1 - \frac{k-1}{n})
\end{aligned}$$

$\underline{k = 23:}$ $\mathbb{P}(A) \approx 0,507 > \frac{1}{2}$

$n = \binom{49}{6}, k = 4004, \mathbb{P}(A) = 0,5001 > \frac{1}{2}$

4.7 Beispiel

n Personen bringen (zu einer Feier) je ein Geschenk mit. Geschenke werden "rein zufällig" verteilt. Mit welcher Wahrscheinlichkeit bekommt mindestens eine Person ihr eigenes Geschenk?

$\rightsquigarrow$ Siebformel!

4.8 Beispiel (Besetzungsmodelle)

k Teilchen sollen auf n nummerierte Fächer verteilt werden. Analogie zu Urnenmodell: Nummer der Kugel $\widehat{=}$ Nummer des Fachs, Nummer der Ziehung $\widehat{=}$ Nummer des Teilchens

Mehrfachbesetzungen / Unterscheidbare Teilchen	ja	nein
ja	Per_k^n Maxwell-Boltzmann	Kom_k^n Bose-Einstein-Statistik
nein	$Per^n_{k,\neq}$ Fermi-Dirak-Statistik	$Kom^n_{k,\neq}$

Statistische Physik

Kapitel 5

Der Erwartungswert

$p(\omega) = \mathbb{P}(\{\omega\})$, $(\Omega, \mathbb{P})$ diskreter Wahrscheinlichkeitsraum

5.1 Definition

- Der Erwartungswert einer Zufallsvariablen $X\colon \Omega \to \mathbb{R}$ existiert (genauer: X ist integrierbar bezüglich $\mathbb{P}$), falls

$$\sum_{\omega \in \Omega} |X(\omega)| p(\omega) < \infty \tag{5.1}$$

  In diesem Fall heißt

$$\mathbb{E}X = \mathbb{E}[X] := \sum_{\omega \in \Omega} X(\omega) p(\omega)$$

  (Physik: $< X >= \mathbb{E}[X]$) **Erwartungswert von** X.

- Ist $X \geq 0$ eine Zufallsvariable, so heißt

$$\mathbb{E}X := \sum_{\omega \in \Omega} X(\omega) p(\omega) \in [0, \infty]$$

  ebenfalls Erwartungswert von X.

5.2 Satz

Sei $L^1 \equiv L^1(\mathbb{P}) := \{X\colon X \text{ erfüllt } 5.1\}$. Dann ist L^1 ein reeller Vektorraum. Genauer:

- (i) $\mathbb{E}[X + Y] = \mathbb{E}X + \mathbb{E}Y, X, Y \in L^1$
- (ii) $\mathbb{E}[aX] = a\mathbb{E}X, X \in L^1, a \in \mathbb{R}$
- (iii) $\mathbb{E}1_A = \mathbb{P}(A), A \subset \Omega$

- (iv) $X \leq Y \Rightarrow \mathbb{E}X \leq \mathbb{E}Y$
- (v) $|\mathbb{E}X| \leq \mathbb{E}|X|$

Beweis. (i) $|(X+Y)(\omega)| \leq |X(\omega)| + |Y(\omega)|$.
Also $X + Y \in L^1(\mathbb{P})$ und

$$\sum_{\omega\in\Omega}(X(\omega)+Y(\omega))p(\omega) \stackrel{!}{=} \sum_{\omega\in\Omega} X(\omega)p(\omega) + \sum_{\omega\in\Omega} Y(\omega)p(\omega)$$

(ii) analog
(iii) $\mathbb{E}1_A = \sum\limits_{\omega\in\Omega} p(\omega) = \mathbb{P}(A)$
(iv) + (v) Übungsaufgabe □

5.3 Folgerung

Seien $A_1, \ldots, A_n \subset \Omega$ und $X := \sum\limits_{j=1}^{n} 1_{A_j}$. Dann gilt $\mathbb{E}X = \sum\limits_{j=1}^{n} \mathbb{P}(A_j)$.
(Gilt auch für ∞ viele Ereignisse.)

5.4 Satz (Transformationsformel)

Seien $X\colon \Omega \Rightarrow \mathbb{R}$ und $g\colon \mathbb{R} \to \mathbb{R}$. Definiere $g(X)\colon \Omega \to \mathbb{R}$ durch

$$g(X)(\omega) = g(X(\omega)).$$

Dann ist $g(X) \in L^1(\mathbb{P})$ genau dann, wenn

$$\sum_{x\colon \mathbb{P}(X=x)>0} |g(x)|\mathbb{P}(X=x) < \infty$$

[1]

In diesem Fall gilt

$$\mathbb{E}g(x) = \sum_{x\colon \mathbb{P}(X=x)>0} g(x)\mathbb{P}(X=x)$$

Beweis. Es gilt

$$\sum_{\omega\in\Omega} |g(X(\omega))|p(\omega) = \sum_{x\in\mathbb{R}\colon \mathbb{P}(X=x)>0} |g(x)| \sum_{\omega\in\Omega\colon X(\omega)=x} p(\omega)$$

$$\Omega = \bigcup_{x\in\mathbb{R}\colon X(\omega)=x, \mathbb{P}(X=x)>0} \{\omega\colon X(\omega)=x\} \cup \Omega', \mathbb{P}(\Omega') = 0$$

---

[1] $\mathbb{P}(X=x) = \mathbb{P}(\{\omega \in \Omega\colon X(\omega) = x\})$

$$= \sum |g(X(\omega))|p(\omega) = \sum_{\omega \notin \Omega'} |g(X(\omega))|p(\omega)$$

$$= \sum_{x \in \mathbb{R}:\, \mathbb{P}(X=x)>0} \; \sum_{\omega \in \{\omega \in \Omega:\, X(\omega)=x\}} |g(X(\omega))|p(\omega)$$

$$= \sum_{x\cdots} |g(x)|\mathbb{P}(X = x)$$

Ist das endlich, so gilt die Rechnung auch ohne Betragsstriche!
Insbesondere gilt

$$\mathbb{E}X = \sum_{x \in \mathbb{R}} x\mathbb{P}(X = x)\,(g(x) \equiv x)$$

$\square$

5.5 Beispiele

- Würfelwurf, X=Augenzahl, $\mathbb{P}(X = j) = \frac{1}{6}$.
  Also
  $$\mathbb{E}X = \sum_{j=1}^{6} j \cdot \mathbb{P}(X = j) = \frac{6 \cdot 7}{2} \cdot \frac{1}{6} = \frac{7}{2} = 3,5$$
- Zweifacher Würfelwurf, X=Maximum der Augenzahlen ($\Omega = \{1, \ldots, 6\}^2, \mathbb{P} =$ Gleichverteilung)
  $$\mathbb{P}(X = 1) = \frac{1}{36}$$
  $$\mathbb{P}(X = 2) = p((1,2)) + p((2,1)) + p((2,2)) = \frac{3}{36}$$
  Allgemein: $\mathbb{P}(X = j) = \frac{2j-1}{36}, j = 1, \ldots, 6$
  Es folgt
  $$\mathbb{E}X = \sum_{j=1}^{6} j \cdot \frac{2j-1}{36} \overset{?}{\approx} 4,47$$

Kapitel 6

Die hypergeometrische Verteilung und die Binomialverteilung

Urne mit Kugeln $\underbrace{1, 2, \ldots, r}_{rot}, \underbrace{r+1, \ldots, r+s}_{schwarz}$

$r, s \in \mathbb{N}_0, r+s > 0$.

6.1 Definition

- n mal Ziehen ohne Zurücklegen
- $a_j :=$ Nummer der j-ten gezogenen Kugel
- $\Omega = Per^{r+s}_{n,\neq}$
- $\mathbb{P} =$ Gleichverteilung ("unabhängiges", "rein zufälliges" Ziehen)
- $A_j := \{(a_1, \ldots, a_n) \in \Omega \colon a_j \leq r\} \hat{=} \{$j-te gezogene Kugel ist rot$\}$
- $X := \sum\limits_{j=1}^{n} 1_{A_j} =$ Anzahl der gezogenen roten Kugeln

$\mathbb{P}^X$ (die Verteilung von X) heißt **hypergeometrische Verteilung** mit Parametern r, s, n, kurz:

$$X \sim Hyp(n, r, s), n \leq r+s$$

$$\mathbb{P}^X = Hyp(n, r, s)$$

6.2 Satz

Es gilt

- (i) $\mathbb{E}X = n \cdot \frac{r}{r+s}$
- (ii) $\mathbb{P}(X = k) = \frac{\binom{r}{s}\binom{s}{n-k}}{\binom{r+s}{n}}, k = 0, \dots, r \wedge n$

Beweis. (i) Es gilt (Symmetrieargument!) $|A_j| = r \cdot (r+s-1)^{\underline{n-1}}$
$|\Omega| = (r+s)^{\underline{n}} \Rightarrow \mathbb{P}(A_j) = \frac{|A_j|}{|\Omega|} = \frac{r}{r+s}$
Aus 5.3 folgt $\mathbb{E}X = n \cdot \frac{r}{r+s}$
(ii) $|\{X = k\}| \stackrel{!}{=} \binom{n}{k} r^{\underline{k}} s^{\underline{n-k}}$
$\Rightarrow \mathbb{P}(X = k) = \frac{\binom{n}{k} r^{\underline{k}} s^{\underline{n-k}}}{(r+s)^{\underline{n}}} = \frac{\binom{r}{k}\binom{s}{n-k}}{\binom{r+s}{n}}$ □

6.3 Motivation

X Zufallsvariable, $\sum\limits_{k=1}^{r} \mathbb{P}(X = x_n) = 1$
$X_1, X_2, \dots, X_n$ "unabhängige" Wiederholungen von X (= Ergebnis eines zufälligen Versuchs)
$\bar{X} := \frac{1}{n}(X_1 + \dots + X_n)$ Zufallsvariable!
Mit $h_j := card\{i \in \{1, \dots, n\} \colon X_i = x_j\}$ gilt $\bar{X} \stackrel{!}{=} \frac{1}{n}(h_1 x_1 + h_2 x_2 + \dots + h_n x_n)$
empirisches Gesetz über Stabilität relativer Häufigkeiten
$\underset{n\to\infty}{\to} \mathbb{P}(X = x_1)x_1 + \dots + \mathbb{P}(X = x_r)x_r \stackrel{!}{=} \mathbb{E}X$
$X \sim Hyp(n, r, s) = \mathbb{P}^X, n \leq r + s$
$\mathbb{P}(X = k) = \frac{\binom{r}{k}\binom{s}{n-k}}{\binom{r+s}{n}}, k = 0, \dots, n$
Wegen $\binom{m}{l} := 0$ für $m < l$ gilt: $\mathbb{P}(X = k) = 0$ für $k < r$ und für $n - k > s(k < n - s)$

6.4 Definition

Binomialverteilung:

- n maliges Ziehen aus einer Urne mit $r + s$ Kugeln mit Zurücklegen
- $\Omega = Per_n^{r+s} = \{(a_1, \dots, a_n) \colon 1 \leq a_i \leq r + s, i = 1, \dots, n\}$
- $\mathbb{P}$ Gleichverteilung

$X := \sum\limits_{j=1}^{n} 1_{A_j}, A_j := \{(a_1, \dots, a_n) \in \Omega \colon a_j \leq r\}$
$\mathbb{P}^X$ heißt Binomialverteilung mit Parametern n und $p := \frac{r}{r+s}$. Man schreibt auch $Bin(n, p) := \mathbb{P}^X$.

6.5 Satz

Es gilt

1. $\mathbb{E}X = np$
2. $\mathbb{P}(X = k) = \binom{n}{k} p^k (1-p)^{n-k}, 0 \leq k \leq n$

Beweis.

1. $|A_j| = r \cdot (r+s)^{n-1}$
   $|\Omega| = (r+s)^n \rightsquigarrow \mathbb{P}(A_j) = \frac{|A_j|}{|\Omega|} = \frac{r}{r+s} = p$
   Folgerung 5.3 $\rightsquigarrow \mathbb{E}X = np$.
2. $card\{X = k\} = \binom{n}{k} r^k s^{n-k}$
   $\rightsquigarrow \mathbb{P}(X = k) = \frac{\binom{n}{k} r^k s^{n-k}}{(r+s)^k (r+s)^{n-k}}$

□

Bemerkung $Bin(n, p)$ ist für jedes $p \in [0, 1]$ definiert.

Kapitel 7

Mehrstufige Experimente

7.1 Beispiel

Urne mit einer roten und drei schwarzen Kugeln
1. Experiment Kugel ziehen, Farbe notieren, Kugel und eine weitere Kugel derselben Farbe zurücklegen
2. Experiment Erneut Kugel ziehen
Modell: $\Omega := \{0,1\} \times \{0,1\}, \quad (0 \widehat{=} s, 1 \widehat{=} r)$

Konstruktion von $\mathbb{P}$ $\quad p(\omega) := \mathbb{P}(\{\omega\})$

$$\left.\begin{aligned} p(1,1) &:= \tfrac{1}{4} \cdot \tfrac{2}{5} = \tfrac{2}{20} = \tfrac{1}{10} \\ p(1,0) &:= \tfrac{1}{4} \cdot \tfrac{3}{5} = \tfrac{3}{20} \\ p(0,1) &:= \tfrac{3}{4} \cdot \tfrac{1}{5} = \tfrac{3}{20} \\ p(0,0) &:= \tfrac{3}{4} \cdot \tfrac{4}{5} = \tfrac{12}{20} \end{aligned}\right\} \text{1. Pfadregel}$$

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

Betrachte $B := \{(1,1),(0,1)\}$. Dann gilt

$$\mathbb{P}(B) = p(1,1) + p(0,1) = \text{(2. Pfadregel)}$$

$$= \frac{2}{20} + \frac{3}{20} = \frac{1}{4} \overset{!}{=} \mathbb{P}(\text{erste Kugel ist rot})$$

(TODO: Bild(Baumdiagramm))

7.2 Definition

Mehrstufige Experimente $\quad \Omega = \Omega_1 \times \ldots \times \Omega_n$ ($\Omega_j \widehat{=}$ Grundraum für j-tes Teilexperiment)
$\omega = (a_1, \ldots, a_n) \in \Omega$
Problem: Definiere $p(\omega) = \mathbb{P}(\{\omega\})$

1. Startverteilung $p_1\colon \Omega_1 \to [0,1] \quad \sum\limits_{\omega\in\Omega_1} p_1(\omega) = 1$
2. Übergangswahrscheinlichkeiten $p_2(a_2|a_1) \geq 0 \quad \sum\limits_{a_2\in\Omega_2} p_2(a_2|a_1) \overset{!}{=} 1$

   ($p_2(a_2|a_1)\hat{=}$ Wahrscheinlichkeit, dass 2. Versuch das Ergebnis a_2 liefert unter der Bedingung, dass 1. Versuch Ergebnis a_1 geliefert hat.)
   $p_3(a_3|a_1,a_2) \geq 0 \quad \sum\limits_{a_3\in\Omega_3} p_3(a_3|a_1,a_2) = 1$
   ...
   $p_n(a_n|a_1,\ldots,a_{n-1}) \geq 0 \quad \sum\limits_{a_n\in\Omega_n} p_n(a_n|a_1,\ldots,a_{n-1}) = 1$

Setze für $\omega = (a_1,\ldots,a_n) \in \Omega$

$p(\omega) := p_1(a_1)\cdot p_2(a_2|a_1)\cdot p_3(a_3|a_1,a_2)\cdot\ldots\cdot p_n(a_n|a_1,\ldots,a_{n-1})$ 1. Pfadregel

Schließlich sei

$\mathbb{P}(A) := \sum\limits_{\omega\in A} p(\omega), \quad A \subset \Omega$ Produkt von Übergangswahrscheinlichkeiten

7.3 Satz

$(\Omega,\mathbb{P})$ ist diskreter Wahrscheinlichkeitsraum.

Beweis. zu zeigen: $\sum\limits_{\omega\in\Omega} p(\omega) = 1$
Induktion (oder direkt)! Zum Beispiel gilt für $n = 2$

$$\sum_{\omega\in\Omega} p(\omega) = \sum_{(a_1,a_2)\in\Omega_1\times\Omega_2} p_1(a_1)p_2(a_2|a_1) = \sum_{a_1\in\Omega_1}\sum_{a_2\in\Omega_2} p_1(a_1)p_2(a_2|a_1)$$

$$\sum_{a_1\in\Omega_1} p_1(a_1)\cdot 1 = 1.$$

□

7.4 Beispiel

Unabhängige Experimente $(\Omega_j,\mathbb{P}_j), j = 1,\ldots,n$, diskrete Wahrscheinlichkeitsräume, $p_i(a_i) = \mathbb{P}_i(\{a_i\})$
Idee: "Unabhängiges" Durchführen der zugehörigen Experimente

$$\Omega := \Omega_1 \times \ldots \times \Omega_n, p(\omega) := p_1(a_1)\cdot\ldots\cdot p_n(a_n), \omega = (a_1,\ldots,a_n) \in \Omega$$

$\big(p_2(a_2|a_1) = p_2(a_2),\ldots,p_n(a_n|a_1,\ldots,a_{n-1}) = p_n(a_n)\big)$

$$\mathbb{P}(A) = \sum_{\omega\in A} p(\omega)$$

Man nennt $\mathbb{P}$ das **Produkt** von $\mathbb{P}_1, \ldots, \mathbb{P}_n$ und schreibt

$$\mathbb{P} := \bigotimes_{i=1}^{n} \mathbb{P}_i.$$

z.B. kann $\Omega = \Omega_1 \times \Omega_2, \quad \Omega_1 = \Omega_2 = \{1, \ldots, 6\}$
$p_1(a_1) = p_2(a_2) = \frac{1}{6}$
Dann ist

$$p(a_1, a_2) = \frac{1}{6} \cdot \frac{1}{6} = \frac{1}{36}$$

und $\mathbb{P}$ ist die Laplace-Verteilung auf Ω.