

Einführung in die Stochastik - Mitschrieb

Vorlesung im Wintersemester 2011/2012

Sarah Lutteropp

25. Oktober 2011

Inhaltsverzeichnis

1	Deskriptive Statistik	3
1.1	Der Grundraum	3
1.2	Absolute und relative Häufigkeit	3
1.3	Histogramm	3
1.4	Lagemaße	3
1.5	Streuungsmaße	5
1.6	Empirischer Korrelationskoeffizient	5
2	Ereignisse und Zufallsvariablen	7
2.1	Definition	7
2.2	Beispiele	7
2.3	Bemerkung (Mengentheoretische Operationen)	7
2.4	Definition	8
2.5	Definition	8
2.6	Definition	9
2.7	Bemerkungen (Rechenregeln für Indikatorfunktionen)	9
2.8	Definition	9
3	Diskrete Wahrscheinlichkeitsräume	11
3.1	Motivation	11
3.2	Definition	11
3.3	Folgerung	12
3.4	Satz	13
3.5	Definition + Satz	13

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

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

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

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

1.3 Histogramm

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

1.4 Lagemaße

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

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

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

1.4.1 Arithmetisches Mittel

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

Fakt $\sum_{j=1}^n (x_j - t)^2 \xrightarrow{t} \text{Min}$

Lösung: $t = \bar{x}$

"Prinzip der kleinsten Quadrate"

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

1.4.2 Median, Quantile

$x_1, \dots, x_n \in \mathbb{R} \Rightarrow x_{(1)} \leq x_{(2)} \leq \dots \leq x_{(n)}$ 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, \dots, x_n .

Fakt $\sum_{j=1}^n |x_j - x_{1/2}| = \min_t \sum_{j=1}^n |x_j - t|$ Übungsaufgabe

Bemerkung Der Median ist "robust" gegenüber "Ausreißern". Ist etwa $x_1 = \dots = x_9 = 1$ 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, \dots, 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 $\sigma: \mathbb{R}^n \rightarrow \mathbb{R}$ mit

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

heißt **Streuungsmaß**.

1.5.1 Empirische Varianz

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

1.5.2 Empirische Standardabweichung

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

1.5.3 Spannweite

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

1.5.4 Quartilsabstand

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

1.6 Empirischer Korrelationskoeffizient

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

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

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

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

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

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

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

"lineare Regression"

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

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

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

Folgerung $|r_{xy}| \leq 1$

Es gilt $r_{xy} = \pm 1 \Leftrightarrow$ Punktwolke 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 **Elementarerereignisse**. 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)
 $\Omega = \{0, 1\}^m (A = \{\omega = (\omega_1, \dots, \omega_m) : \sum_{j=1}^m \omega_j \geq k\} \text{ Ereignis })$
- (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 $A, B, A_1, A_2, \dots \subset \Omega$.

$A \cap B = \{\omega \in \Omega : \omega \in A \text{ und } \omega \in B\} \hat{=} \text{"A und B treten ein"}$

$A \cup B \hat{=} \text{"A oder B treten ein"}$

$\bar{A} \equiv A^c := \Omega \setminus A = \{\omega \in \Omega : \omega \notin A\} \hat{=} \text{"A tritt nicht ein"}$

$A \setminus B = A \cap B^c \hat{=}$ "A tritt ein, aber nicht B"

$A \subset B \hat{=}$ "wenn A, dann B"

$\emptyset \hat{=}$ "unmögliches Ereignis"

$\Omega \hat{=}$ "sicheres Ereignis"

Abkürzung $AB = A \cap B$

2.4 Definition

Eine Abbildung $X: \Omega \rightarrow \mathbb{R}$ 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 $X^{-1}: \mathcal{P}(\mathbb{R}) \rightarrow \mathcal{P}(\Omega) = \{A: A \subset \Omega\}$ ist definiert durch

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

Bemerkung

- $X^{-1}(A \cap B) = X^{-1}(A) \cap X^{-1}(B), A, B \subset \mathbb{R}$
- $X^{-1}(A \cup B) = X^{-1}(A) \cup X^{-1}(B)$
- $X^{-1}\left(\bigcup_{j=1}^{\infty} A_j\right) = \bigcup_{j=1}^{\infty} X^{-1}(A_j)$
- $X^{-1}\left(\bigcap_{j=1}^{\infty} A_j\right) = \bigcap_{j=1}^{\infty} X^{-1}(A_j)$

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

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

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 $a \in \mathbb{R}$

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

2.6 Definition

Sei $A \subset \Omega$. Die Funktion $1_A: \Omega \rightarrow \mathbb{R}$ 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)

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

2.8 Definition

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

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

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

Bemerkung

- $\{X = 0\} = \{\omega : X(\omega) = 0\} = A_1^c \cap \dots \cap A_n^c$
- $\{X = n\} = A_1 \cap \dots \cap A_n$
- $\{X = k\} = \text{"genau } k \text{ der Ereignisse } A_1, \dots, A_n \text{ treten ein"} =$

$$\bigcup_{T \subset \{1, \dots, n\}, |T|=k} \left(\bigcap_{j \in T} A_j \cap \bigcap_{j \notin T} A_j^c \right)$$
 $(T \subset \{1, \dots, n\}, |T| = \text{card } T = k)$

Kapitel 3

Diskrete Wahrscheinlichkeitsräume

3.1 Motivation

Zufallsexperiment mit Ausgängen in Ω

n -malige, ‘unabhängige’ Wiederholung

\Rightarrow Ergebnis $(a_1, \dots, a_n) \in \Omega^n$

$r_n(A) := \frac{1}{n} \sum_{j=1}^n 1_A(a_j)$, $A \subset \Omega$ 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:

$r_n(A) \underset{n \rightarrow \infty}{\rightsquigarrow} ?$

3.2 Definition

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

- (P1) $\mathbb{P}(A) \geq 0$, $A \subset \Omega$
- (P2) $\mathbb{P}(\Omega) = 1$
- (P3) $\mathbb{P}\left(\bigcup_{j=1}^{\infty} A_j\right) = \sum_{j=1}^{\infty} \mathbb{P}(A_j)$, $A_i \cap A_j = \emptyset$, $i \neq j$

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) $\mathbb{P}(\bigcup_{j=1}^n A_j) = \sum_{j=1}^n \mathbb{P}(A_j), A_i \cap A_j = \emptyset, i \neq j$
- 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) $\mathbb{P}(\bigcup_{j=1}^{\infty} A_j) \leq \sum_{j=1}^{\infty} \mathbb{P}(A_j)$ (Subadditivität)
- h) $A_n \subset A_{n+1}, n \in \mathbb{N} \Rightarrow \mathbb{P}(\bigcup_{n=1}^{\infty} A_n) = \lim_{n \rightarrow \infty} \mathbb{P}(A_n)$ (Stetigkeit von unten)
- i) $A_n \supset A_{n+1}, n \in \mathbb{N} \Rightarrow \mathbb{P}(\bigcap_{n=1}^{\infty} A_n) = \lim_{n \rightarrow \infty} \mathbb{P}(A_n)$ (Stetigkeit von oben)

Beweis • a): $A_j = \emptyset, j \in \mathbb{N}$ (P3) ¹ $\mathbb{P}(\emptyset) = 0$.

• b): $A_{n+1} = A_{n+2} = \dots = \emptyset$ in P3!

• c) + f): Für $A \subset \Omega$ gilt nach b) (für $n = 2$):

$$1 = \mathbb{P}(\Omega) = \mathbb{P}(A \cup A^c) \stackrel{(b)}{=} \mathbb{P}(A) + \mathbb{P}(A^c)$$

• d): Nach b) gilt $\mathbb{P}(A) = \mathbb{P}(A \setminus B) + \mathbb{P}(A \cap B)$, $\mathbb{P}(B) = \mathbb{P}(B \setminus A) + \mathbb{P}(A \cap B)$ und

$$\text{somit } \mathbb{P}(A) + \mathbb{P}(B) - \mathbb{P}(A \cap B) = \mathbb{P}(A \setminus B) + \mathbb{P}(B \setminus A) + \mathbb{P}(A \cap B) \stackrel{(b)}{=} \mathbb{P}(A \cup B)$$

• e): Wegen $B = A \cup (B \setminus A)$ folgt² $\mathbb{P}(B) = \mathbb{P}(A) + \mathbb{P}(B \setminus A) \geq \mathbb{P}(A)$

• g): $B_1 := A_1, B_2 := A_2 \setminus A_1, \dots, B_n := A_n \setminus (\bigcup_{j=1}^{n-1} A_j), n \geq 2$.

Dann gilt $B_n \subset A_n$ und $\bigcup_{j=1}^n B_j = \bigcup_{j=1}^n A_j$ sowie $B_i \cap B_j = \emptyset, i \neq j$.

Es folgt aus (P3):

$$\mathbb{P}(\bigcup_{j=1}^{\infty} A_j) \stackrel{!}{=} \mathbb{P}(\bigcup_{j=1}^{\infty} B_j) \stackrel{(P3)}{=} \sum_{j=1}^{\infty} \mathbb{P}(B_j) \stackrel{e)}{\leq} \sum_{j=1}^{\infty} \mathbb{P}(A_j) \quad (\infty \text{ ist zugelassen})$$

• 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

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

Dann gilt

- a) $\mathbb{P}(\bigcup_{j=1}^n A_j) = \sum_{k=1}^n (-1)^{k-1} S_k$ ‘Siebformel’
- b) $\mathbb{P}(\bigcup_{j=1}^n A_j) \leq \sum_{k=1}^{2s+1} (-1)^{k-1} S_k, s = 0, \dots, \lfloor \frac{n-1}{2} \rfloor$
 $\mathbb{P}(\bigcup_{j=1}^n A_j) \geq \sum_{k=1}^{2s} (-1)^{k-1} S_k, s = 1, \dots, \lfloor \frac{n}{2} \rfloor$

Beweisidee für Siebformel vollständige Induktion nach n :

$$\underline{n=2}: \mathbb{P}(A_1 \cup A_2) \stackrel{(d)}{=} \mathbb{P}(A_1) + \mathbb{P}(A_2) - \mathbb{P}(A_1 \cap A_2) = S_1 - S_2$$

$$\underline{n=3}: \mathbb{P}(\underbrace{A_1 \cup A_2}_{\cup A_3}) \stackrel{(d)}{=} \mathbb{P}(A_1 \cup A_2) + \mathbb{P}(A_3) - \mathbb{P}((A_1 \cup A_2) \cap A_3) \quad ^3$$

$$\stackrel{(d)}{=} \mathbb{P}(A_1) + \mathbb{P}(A_2) - \mathbb{P}(A_1 \cap A_2) + \mathbb{P}(A_3) - \mathbb{P}(A_1 \cap A_3) - \mathbb{P}(A_2 \cap A_3) + \mathbb{P}(A_1 \cap A_2 \cap A_3) = S_1 - S_2 + S_3$$

3.5 Definition + Satz

a) Sei (Ω, \mathbb{P}) diskreter Wahrscheinlichkeitsraum. Dann heißt $p: \Omega \rightarrow \mathbb{R}$ definiert durch $p(\omega) := \mathbb{P}(\{\omega\})$ **Wahrscheinlichkeitsfunktion** (von \mathbb{P}).

Es gilt $\mathbb{P}(A) = \sum_{\omega \in A} p(\omega), A \subset \Omega$.

b) Sind Ω diskret und $p: \Omega \rightarrow \mathbb{R}$ eine Abbildung mit $p(\omega) \geq 0$ und $\sum_{\omega \in \Omega} p(\omega) = 1$, so erhält man vermöge $\mathbb{P}(A) := \sum_{\omega \in A} p(\omega)$ einen diskreten Wahrscheinlichkeitsraum.

Beweis • a) σ -Additivität ($A = \bigcup_{\omega \in A} \{\omega\}$)

• b) σ -Additivität: Großer Umordnungssatz! (Analysis)

3.6 Definition

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

(‘homogene Münze’, ‘Würfeln’, ...)

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

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