Einführung in die Stochastik - Mitschrieb

Vorlesung im Wintersemester 2011/2012

Sarah Lutteropp

20. Oktober 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")
$h(\omega) = \operatorname{card}\{j \in \{1, \ldots, n\} \colon x_j = \omega\}, \omega \in \Omega$, absolute Häufigkeit von ω

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

Definition $\frac{1}{n} h(\omega) =$ relative Häufigkeit von ω
$h(A) = \operatorname{card}\{j \in \{1, \ldots, n\} \colon x_j \in A\}, A \subset \Omega =$ 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
$d_j(b_{j+1} - b_j) = h([b_j, b_{j+1})) = \operatorname{card}\{i \in \{1, \ldots, n\} \colon b_j \leq x_i < b_{j+1}\}$

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

$s^2 := \frac{1}{n-1} \sum\limits_{j=1}^{n} (x_j - \bar{x})^2$ = **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 $\sigma_x^2 = \frac{1}{n} \sum\limits_{j=1}^{n} (x_j - \bar{x})^2$ $\sigma_y^2 = \frac{1}{n} \sum\limits_{j=1}^{n} (y_j - \bar{y})^2$

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

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

$\min\limits_{a,b} \sum\limits_{j=1}^{n} (y_j - a - bx_j)^2 \overset{!}{=} \min\limits_{b} \sum\limits_{j=1}^{n} (y_i - \bar{y} - b(x_j - \bar{x}))^2 = \ldots$

"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)
  $\Omega = \{0,1\}^m (A = \{\omega = (\omega_1, \ldots, \omega_m) : \sum_{j=1}^{m} \omega_j \geq k\}$ Ereignis)
- (iii) Werfen von 2 Würfeln
  $\Omega = \{1, \ldots, 6\}^2$
- (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 $A, B, A_1, A_2, \ldots \subset \Omega$.
$A \cap B = \{\omega \in \Omega : \omega \in A$ 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"

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

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

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

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