Wahrscheinlichkeitsrechnung und Statistik

Manuel Kauers
nach einer Vorlesung von
Prof. Dr. M. Stieglitz

21. Oktober 2000

Inhaltsverzeichnis

1 **Bedingte Wahrscheinlichkeit** 1

2 **Zähldichten und Riemann-Dichten** 2

3 **Verteilungsfunktionen** 3

4 **Kopplung** 4

5 **Zufallsvariablen** 4

6 **Kenngrößen** 7

7 **Grundlagen der Statistik** 10

8 **Punktschätzung** 11

9 **Intervallschätzung** 11

10 **Maximum-Likelihood-Schätzer** 12

11 **Testen von Hypothesen** 13

12 **χ^2-Anpassungstest** 14

13 **Test auf Unabhängigkeit** 15

1 Bedingte Wahrscheinlichkeit

Def. 1 (Bedingte Wahrscheinlichkeit zweier Ereignisse) Seien $A, B \subseteq \Omega$ Ereignisse. Dann heißt

$$P(A \mid B) := P_B(A) := \frac{P(A \cap B)}{P(B)}$$

die *bedingte Wahrscheinlichkeit* von A unter B.

Satz 1 (Totale Wahrscheinlichkeit) Seien B_i n Ereignisse mit $\Omega = \bigcup_{i \in \{1 \ldots n\}} B_i$. Dann gilt:

$$P(A) = \sum_{i \in \{1 \ldots n\}} P(A \cup B_i) = \sum_{i \in \{1 \ldots n\}} P(B_i) \cdot P(A \mid B_i)$$

Bem. B ist eine disjunkte Zerlegung von Ω.

Satz 2 (Bedingte Wahrscheinlichkeit bei mehr als zwei Ereignissen) Es seien $A, B, C \subseteq \Omega$ Ereignisse. Dann gilt:

$$P(A \cap B \cap C) = P(A) \cdot P(B \mid A) \cdot P(C \mid (A \cap B))$$

Analog für mehr als drei Ereingisse.

Satz 3 (Formel von Bayes, Umkehrung von Bäumen) Seien B_i n Ereignisse mit $\Omega = \bigcup_{i \in \{1 \ldots n\}} B_i$. Dann gilt:

$$P(B_k \mid A) = \frac{P(B_k)P(A|B_k)}{\sum_{i=1}^{n} P(B_i)P(A \mid B_k)}$$

Def. 2 (stochastische Unabhängigkeit bei zwei Ereignissen) Zwei Ereignisse A, B heißen *stochastisch unabhängig*, falls gilt:

$$P(A \cap B) = P(A) \cdot P(B)$$

Bem.

- A, B stochastisch unabhängig und $P(A) \neq 0 \Longrightarrow P(B \mid A) = P(B)$
- A, B stochastisch unabhängig $\Longrightarrow A^C, B$ stochastisch unabhängig

Def. 3 (Stochastische Unabhängigkeit bei n Ereignissen) $A_1 \ldots A_n$ heißen *stochastisch unabhängig* $:\Longleftrightarrow$

$$\begin{array}{l} \forall \quad A_{i_1} \ldots A_{i_k} \text{Teilmengen von } A_1 \ldots A_n : \\ P(A_{i_1} \cap A_{i_2} \cap \cdots \cap A_{i_k}) = P(A_{i_1})P(A_{i_2}) \cdots P(A_{i_k}) \end{array}$$

Bem. Aus paarweiser stochastischer Unabhängigkeit folgt i.A. nicht die stochastische Unabhängigkeit für mehr als zwei Ereignisse.

2 Zähldichten und Riemann-Dichten

Satz 4 Ist P ein W-Maß in Ω und $f : \Omega \to \mathbb{R}$, $f(\omega) = P(\{\omega\})$, dann gilt:

1. $\forall \quad \omega \in \Omega : \qquad f(\omega) \geq 0$
2. $\sum_{\omega \in \Omega} f(\omega) = 1$

3. $P(A) = \sum\limits_{\omega \in A} f(\omega)$

Bem. Jede solche Abbildung $f : \Omega \to \mathbb{R}$ definiert ein W-Maß mit $P(\{\omega\}) = f(\omega)$

Def. 4 (Zähldichte) Eine Funktion $f : \Omega \to \mathbb{R}$ mit den Eigenschaften aus Satz 4 heißt *Zähldichte*.

Def. 5 (RIEMANN-Dichte) Sei $f : \mathbb{R} \to \mathbb{R}$ mit $f(x) \geq 0$ und $\int\limits_{-\infty}^{\infty} f(x)dx = 1$. Dann heißt f RIEMANN-*Dichte* und f definiert ein eindeutiges W-Maß durch

$$P([a,b]) = \int\limits_a^b f(x)dx \qquad P(\{a\}) = 0 \qquad (a, b \in \mathbb{R})$$

3 Verteilungsfunktionen

Def. 6 (Verteilungsfunktion) Ist P ein W-Maß, dann heißt $F : \mathbb{R} \to \mathbb{R}$ mit $F(x) := P((-\infty, x])$ $(x \in \mathbb{R})$ die *Verteilungsfunktion* von P.

Bem.

- $P((a,b]) = F(b) - F(a)$
- Ist f RIEMANN-Dichte zu P, so gilt für die Verteilungsfunktion

$$F(x) = \int\limits_{-\infty}^{x} f(t)dt$$

Def. 7 (Gemischte Verteilung) Hat das W-Maß P über $\mathbb{R}$ die Darstellung

$$P(A) = \alpha_d P_d(A) + \alpha_s P_s(A)$$

für ein diskretes W-Maß P_d, ein stetiges W-Maß P_s, einem $\alpha_d \in [0,1]$ und $\alpha_s = 1 - \alpha_d$, dann heißt P *gemischte Verteilung*.

Satz 5 Sei F eine Verteilungsfunktion. Dann gilt:

1. F ist monoton wachsend (isoton).
2. $\lim\limits_{x \to -\infty} F(x) = 0$, $\lim\limits_{x \to \infty} F(x) = 1$
3. $P(\{x_0\}) = F(x_0) - \lim\limits_{x \to x_0 -} F(x)$

4 Kopplung

Def. 8 (Übergangszähldichten) Hat ein mehrstufiger Versuch die Zähldichte

$$f : (\omega_1, \ldots, \omega_n) \mapsto f(\omega_1, \ldots, \omega_n)$$

mit

$$f(\omega_1, \ldots, \omega_n) = f_1(\omega_1) f_2^1(\omega_1, \omega 2) \cdots f_n^{n-1}(\omega_1, \ldots, \omega_n),$$

dann heißten die abhängigen Zähldichten $f_i^{i-1}(\omega_1, \ldots, \omega_{i-1}; \omega_i)$ *Übergangszähldichten* von $\Omega_1 \times \Omega_2 \times \cdots \times \Omega_{i-1}$ nach Ω_i. Die Parameter $\omega_1, \ldots, \omega_{i-1}$ heißen *Vorgeschichte zur Stufe i*.

Bem.

- Die Zähldichte f heißt *Kopplung* von $f_1, f_1^2, \ldots, f_n^{n-1}$ und wird kurz geschieben als

  $$f = f_1 \otimes f_1^2 \otimes \cdots \otimes f_n^{n-1}$$

  Analog heißen die von $\omega_1, \ldots, \omega_{i-1}$ abhängigen W-Maße P_i^{i-1} *Übergangs-W-Maße* und man schreibt

  $$P = P_1 \otimes P_1^2 \otimes \cdots \otimes P_n^{n-1}$$

  für das zu f gehörende W-Maß.

- Ist die Vorgeschichte unrelevant, so sind die f_i^{i-1} ebenfalls Z- bzw. R-Dichten und

  $$f(\omega_1, \ldots, \omega_n) = f_1(\omega_1) f_2(\omega_2) \cdots f_n(\omega_n) = \prod_{i=1}^{n} f_i(\omega_i)$$

  heißt *Produktdichte*.

  z.B. Unabhängige Wiederholdung desselben Versuchs.

Satz 6 (Wahrscheinlichkeit im n-stufigen Versuch) In einem n-stufigen Versuch mit den W-Maßen $P_1, P_2, \ldots, P_n$ gilt für ein Ereignis $A = A_1 \times A_2 \times \cdots \times A_n$:

$$P(A) = P(A_1 \times A_2 \times \cdots \times A_n) = P_1(A_1) P_2(A_2) \cdots P_n(A_n)$$

Bem. Dies gilt nur bei unabhängigen Versuche.

Bsp. Ziehen mit Zurücklegen

5 Zufallsvariablen

Def. 9 Sei $X : \Omega \to \Omega'$ eine Abbildung, $A' \subset \Omega'$. Dann sei

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

Analog seien die Ereignisse $(X = a)$, $(b < X < c)$, $(X > d)$, ... definiert.

Def. 10 (Urbildfunktion) Sei $X : \Omega \to \Omega'$ eine Abbildung, $A' \subset \Omega'$. Dann heißt $A := (X \in A')$ das *Urbild* von A'. Die Zuordnung $X^{-1} : A' \mapsto X^{-1}(A') := (X \in A')$ heißt *Urbildfunktion.*

Bem.

- X^{-1} ist nicht die Umkehrfunktion.
- X^{-1} existiert immer.

Def. 11 (Zufallsvariable) Sei $(\Omega, \mathcal{A}, P)$ ein W-Modell und Ω' eine Bildmenge mit Ereignissystem $\mathcal{A}'$. Eine Abbildung $X : \Omega \to \Omega'$ heißt *Zufallsvariable* $:\Longleftrightarrow$

$$\forall \quad A' \in \mathcal{A}' : \qquad X^{-1}(A') \in \mathcal{A}$$

Eine solche Abbildung X heißt *meßbar.*

Bem. Im Rahmen der Vorlesung erfüllen alle Abbildungen $X : \Omega \to \Omega'$ diese Eigenschaft und sind somit Zufallsvariablen (insbesondere, da $\mathcal{A} = \mathcal{P}(\Omega)$ festgelegt wurde)

Def. 12 (Bildmaß P^X, Verteilung der Zufallsvariablen X) Sei $X : \Omega \to \Omega'$ eine Zufallsvariable, $A' \subset \Omega$. Dann heißt das W-Maß P^X mit

$$A' \mapsto P^X(A') := P(X^{-1}(A')) = P((X \in A'))$$

Bildmaß von P (unter X) oder *Verteilung von X (bezüglich P)* $(\Omega', \mathcal{A}', P^X)$ heißt *Bildmodell* zu $(\Omega, \mathcal{A}, P)$.

Bem.

- P^X ist W-Maß $\Longleftrightarrow X^{-1}(\Omega') = \Omega \wedge X^{-1}(\sum\limits_{i=1}^{\infty} A'_i) = \sum\limits_{i=1}^{\infty} X^{-1}(A'_i)$
- Schreibweisen:
  $P(X \in A) := P((X \in A))$
  $P(X \in A \mid X \in B) := P((X \in A) \cup (X \in B))$
  $P(X \in A, X \in B) := P((X \in A) \cap (X \in B))$

Satz 7 (Zähldichte f^X, Verteilungsfunktion F^X) Sei $X : \Omega \to \Omega'$ Zufallsvariable. Dann gilt:

1. Ist Ω' abzählbar, so ist $f^X : \Omega \to \mathbb{R}$ mit

   $$f^X(\omega') := P(X = \omega') \qquad (\omega' \in \Omega)$$

   die Zähldichte von P^X

2. Ist $\Omega' = \mathbb{R}$, so ist $F^X : \mathbb{R} \to \mathbb{R}$ mit

   $$F^X(t) = P(X \leq t) \qquad (t \in \mathbb{R})$$

   die Verteilungsfunktion von P^X.

Def. 13 (Projektion, Randverteilung) Sei $\Omega = \Omega_1 \times \Omega_2 \times \cdots \times \Omega_n$, dann heißt die Zufallsvariable

$$X_i : \Omega \to \Omega_i; \qquad \omega \mapsto \omega_i \qquad i \in \{1, \ldots, n\}$$

die i-te Projektion. Die Verteilung P^{X_i} heißt *i-te Randverteilung.*

Satz 8 (Berechnung von P^{X_i})

1. allgemein

$$P^{X_i}(A_i) = P(X_i \in A_i) = P(\Omega_1 \times \Omega_2 \times \cdots \times \Omega_n)$$

2. Ist Ω_i abzählbar und f Zähldichte von P, dann ist f^{X_i} mit

$$f^{X_i}(\omega_i) = \sum_{\omega_1 \in \Omega_1} \cdots \sum_{\omega_{i-1} \in \Omega_{i-1}} \sum_{\omega_{i+1} \in \Omega_{i+1}} \cdots \sum_{\omega_n \in \Omega n} f(\omega_1, \ldots, \omega_n)$$

eine Zähldichte von P^{X_i} und heißt *i-te Randdichte.*

3. Ist $\Omega_i = \mathbb{R}$, f R-Dichte zu P, dann ist f^{X_i} mit

$$f^{X_i}(x_i) = \int_{-\infty}^{\infty} \cdots \int_{-\infty}^{\infty} f(x_1, x_2, \ldots, x_n) dx_1 dx_2 \ldots dx_{i-1} dx_{i+1} \ldots dx_n$$

eine R-Dichte von P^{X_i}

Bem. Bei unabhänginger Kopplung vereinfacht sich die Berechnung zu

$$\begin{aligned} P^{X_i}(A_i) &= P(X_i \in A_i) \\ &= P_1(\Omega_1) \cdots P_{i-1}(\Omega_{i-1}) \cdot P_{i+1}(\Omega i + 1) \cdots P_n(\Omega_n) \end{aligned}$$

Def. 14 (Gemeinsame Verteilung) Sei $\Omega = \Omega_1 \times \Omega_2 \times \cdots \times \Omega_n$ und $Y_i : \Omega \to \Omega_i$ ($i \in \{1, \ldots, n\}$) eine Zufallsvariable. Dann heißt die Zufallsvariable

$$\begin{aligned} &Y : \Omega \to \Omega' := \Omega_1 \times \Omega_2 \times \ldots \Omega_n; \\ &Y := (Y_1, Y_2, \ldots, Y_n) \end{aligned}$$

Zufallsvektor. Die Verteilung $P^Y = P^{(Y_1, \ldots, Y_n)}$ heißt dann *gemeinsame Verteilung von $Y_1, \ldots, Y_n$.*

Bem. Die Berechung einer gemeinsamen Verteilung ist i.A. sehr aufwendig.

Satz 9 Ist $P^{(Y_1, \ldots, Y_n)}$ eine gemeinsame Verteilung, so ist P^{Y_i} die i-te Randverteilung von Y_i.

Def. 15 (Unabhängigkeit von Zufallsvariablen) Seien $X_i : \Omega \to \Omega_i$, ($i \in \{1, \ldots, n\}$) Zufallsvariablen mit

$$\begin{aligned} &\forall \quad A_i \in \mathcal{A}_i, 1 \le i \le n : \\ &P^{(X_1, \ldots, X_n)}(A_1 \times A_2 \times \cdots \times A_n) = P^{X_1}(A_1) \cdots P^{X_n}(A_n) \end{aligned}$$

Dann heißen die X_i *stochastisch unabhängig.*

Satz 10 Sind $X_1, \ldots, X_n$ Zufallsvariablen mit $X_i : \Omega \to \Omega_i$ und Z- bzw. R-Dichten $f_1, \ldots, f_n$. Dann gilt:

$$\begin{array}{ll} & X_1, \ldots, X_n \text{ sind stochastisch unabhängig} \\ \Longleftrightarrow & \exists\, f : \Omega \to \mathbb{R} : f \text{ Produktdichte zu } P^{(X_1,\ldots,X_n)} \end{array}$$

Satz 11 Sind $X_1, \ldots, X_n$ stochastisch unabhängig. Dann sind auch stochastisch unabhängig:

1. Permutationen von $X_1, \ldots, X_n$
2. Teilmengen von $X_1, \ldots, X_n$
3. Vektoren disjunkter Teilmengen, z.B.
$$Z_1 := (X_1, \ldots, X_k), Z_2 := (X_{k+1}, \ldots, X_n)$$
4. Funktionen von $X_1, \ldots, X_n$, z.B. $Z_1 := \sqrt{X_1^2 - X_n}, Z_2 := X_2 + \frac{1}{X_n}$
5. Jede Konstante Zufallsvariable ist von allen Zufallsvariablen unabhängig
6. Sind $Y_1, \ldots, Y_{n-1}$ stochastisch unabhängig und $(Y_1, \ldots, Y_{n-1}), Y_n$ stochastisch unabhängig, dann sind auch $Y_1, \ldots, Y_n$ stochastisch unabhängig

Satz 12 Beliebige Ereignisse $A_1, \ldots, A_n \subseteq \Omega$ sind stochastisch unabhängig $\Longleftrightarrow$ Die Zufallsvariablen $X_i := \mathbf{1}_{A_i} \quad (i \in \{1, \ldots, n\})$ sind stochastisch unabhängig

Satz 13 (Dichte einer Summe von Zufallsfunktionen) Seien $X, Y : \Omega \to \Omega'$ Zufallsvariablen, $f^{(X,Y)}$ die gemeinsame Dichte von X und Y. Sei $Z := X + Y$. Dann gilt:

1. $\Omega' = \mathbb{Z} \quad \Longrightarrow \quad f^Z(z) = \sum\limits_{x \in \mathbb{Z}} f^{(X,Y)}(x, z - x)$
2. $\Omega' = \mathbb{R} \quad \Longrightarrow \quad f^Z(z) = \int\limits_{-\infty}^{\infty} f^{(X,Y)}(x, z - x)dx$

Bem. Sind X, Y stochastisch unabhängig, dann ist $f^Z(z) = \sum\limits_{x \in \Omega'} f^X(x) f^Y(y)$.

Def. 16 (Die Faltung) Seien $X, Y : \Omega \to \Omega'$ Zufallsvariablen mit den Verteilungen P^X, P^Y. Dann heißt

$$P^X * P^Y := P^{X+Y}$$

Faltung von X und Y.

6 Kenngrößen

Def. 17 (Modalwert, Median, Quantil) Sei $X : \Omega \to \Omega'$ Zufallsvariable, P^X W-Maß zu X, f^X Dichte zu X, F^X Verteilungsfunktion von X.

1. Hat f^X an einer Stelle x_0 einen Maximalwert, so heißt x_0 *Modalwert* zu P^X.

2. Eine Stelle x_0 heißt *Median*, falls gilt:

$$F^X(m-) \leq \frac{1}{2} \leq F^X(m)$$

3. $u_\alpha \in (0,1)$ heißt *α-Quantil* von P^X, falls gilt:

$$F^X(u_\alpha -) \leq \alpha \leq F^X(u_\alpha +)$$

Def. 18 (Erwartungswert bei diskreten Verteilungen) $X : \Omega \to \Omega' \subset \mathbb{R}$ sei eine diskret verteilte Zufallsvariable. Weiter sei $X \geq 0$ oder Ω' endlich. Dann heißt

$$E(X) := EX := \sum_{k \in \Omega'} kP(X = k) = \sum_{k \in \Omega'} kf^X(k)$$

Erwartungswert von X.

Bem.

- Der Erwartungswert von X entspricht dem Mittelwert von P^X
- Er ist unabhängig von Ω
- $\exists \quad X_1, X_2; X_1 \neq X_2 : P^{X_1} = P^{X_2}$

Satz 14 Sei $X : \Omega \to \Omega' \subset \mathbb{R}$ eine diskret verteilte Zufallsvariable.

1. Sind Ω, Ω' abzählbar, so gilt:

$$EX := \sum_{k \in \Omega'} kP(X = k) = \sum_{\omega \in \Omega} X(\omega) f^X(\omega)$$

2. Ist $g : \Omega' \to \Omega''$ eine Abbildung und sind Ω', Ω'' abzählbar, so gilt:

$$Eg(X) := \sum_{m \in \Omega''} mP(g(X) = m) = \sum_{k \in \Omega'} g(k)P(X = k)$$

3. Ist $Y : \Omega \to \Omega'_2$ eine diskrete Zufallsvariable und $h : \Omega' \times \Omega'_2 \to \Omega''$ eine Abbildung und sind $\Omega', \Omega'_2, \Omega''$ abzählbar, dann gilt

$$\begin{aligned} Eh(X,Y) &:= \sum_{m \in \Omega''} mP(h(X,Y) = m) \\ &= \sum_{k \in \Omega'} \sum_{l \in \Omega'_2} h(k,l) \cdot P(X = k, Y = l) \end{aligned}$$

Def. 19 (Erwartungswert bei stetigen Verteilungen) Sei $X : \Omega \to \mathbb{R}$ eine Zufallsvariable, f^X R-Dichte von X. Dann heißt

$$EX := \int_{-\infty}^{\infty} xf^X(x)dx$$

Erwartungswert von X, falls $\int_0^\infty xf^X(x)dx < \infty$ oder $\int_{-\infty}^0 |x|f^X(x)dx < \infty$.

Def. 20 (Erwartungswert bei gemischten Veteilungen) Sei $X : \Omega \to \mathbb{R}$ eine Zufallsvariable mit $P^X = \alpha_d P_d^X + \alpha_s P_s^X$ als Verteilung. Dabei seien P_d^X eine diskrete Verteilung, P_s^X eine stetige Verteilung, $\alpha_d \in [0,1]$ beliebig und $\alpha_s = 1 - \alpha_d$. Existieren die Erwartungswerte $E_d X$ von P_d^X und $E_s X$ von P_s^X, dann sei

$$EX := \alpha_d E_d X + \alpha_s E_s X$$

der *Erwartungswert* von X.

Satz 15 Seien $X, Y : \Omega \to \mathbb{R}$ Zufallsvariablen; $a, b \in \mathbb{R}$. Dann gilt:

1. $X < Y, EX, EY$ existiern $\implies EX < EY$
2. $E(a + bX) = a + bEX$
3. $E(X + Y) = EX + EY$, falls $EX, EY \neq \pm\infty$
4. Sind X, Y stochastisch unabhängig und ihre Erwartungswerte existieren und sind endlich, so gilt $E(XY) = EXEY$
5. $E|X| = EX^+ + EX^-$. Dabei sei $a^+ := \max\{0, a\}$ und $a^- := \max\{0, -a\}$
6. EX existiert $\implies |EX| \leq E|X|$

Def. 21 (Varianz) Sei $X : \Omega \to \Omega' \subset \mathbb{R}$ eine Zufallsvariable, $EX < \infty$. Dann heißen

$$\operatorname{Var} X := E(X - EX)^2 = EX^2 - (EX)^2$$

die *Varianz* von X und $\operatorname{Str} X := \sqrt{\operatorname{Var} X}$ die *Streuung* von X.

Satz 16 (Eigenschaften von Varianz und Streuung) Sei $X : \Omega \to \Omega' \subset \mathbb{R}$ eine Zufallsvariable mit $VarX, \operatorname{Str} X$ existieren; $a, b \in \mathbb{R}$. Dann gilt:

1. $\operatorname{Var}(a + b \cdot X) = b^2 \cdot \operatorname{Var} X$
   $\operatorname{Str}(a + b \cdot X) = |b| \operatorname{Str} X$
2. $\operatorname{Str} X = 0 \iff \operatorname{Var} X = 0 \iff P(X = EX) = 1$
3. Sei $Y : \Omega \to \Omega'' \subset \mathbb{R}$ eine weitere Zufallsvariable, X, Y stochastisch unabhängig, so gilt: $\operatorname{Var}(X + Y) = \operatorname{Var} X + \operatorname{Var} Y$

Def. 22 (Kovarianz, Korrelationskoeffizient) Seien $X : \Omega \to \Omega' \subset \mathbb{R}$ und $Y : \Omega \to \Omega'' \subset \mathbb{R}$ Zufallsvariablen mit $EX^2 < \infty$, $EY^2 < \infty$. Dann heißt

$$\operatorname{Kov}(X, Y) := E(XY) - EXEY = E[(X - EX)(Y - EY)]$$

die *Kovarianz* von X und Y.
Ist $\operatorname{Str} X \neq 0$ und $\operatorname{Str} Y \neq 0$, so heißt

$$\operatorname{korr}(X, y) := \frac{\operatorname{Kov}(X, Y)}{\operatorname{Str} X \operatorname{Str} Y}$$

der *Korrelationskoeffizient* von X und Y. Ist $\operatorname{Str} X = 0$ oder $\operatorname{Str} Y = 0$, so sei $\operatorname{korr}(X, Y) := \operatorname{Kov}(X, Y) = 0$

Bem.

- $\operatorname{Kov}(X, Y)$ und $\operatorname{korr}(X, Y)$ sind Maße für die Abhängigkeit von X und Y.
- Es gilt $-1 \leq \operatorname{korr}(X, Y) \leq 1$. Der Korrelationskoeffizient wird daher auch als „normierte Kovarianz" bezeichnet.
- Die Voraussetzungen $EX^2 < \infty, EY^2 < \infty$ gewährleisten die Existenz und Endlichkeit von $E(XY), EX$ und EY.

Satz 17 Seien $X, Y : \Omega \to \mathbb{R}$ Zufallsvariablen mit $EX^2 < \infty, EY^2 < \infty$. Dann gilt:

1. $\operatorname{Kov}(X, X) = \operatorname{Var} X$
2. $\operatorname{Kov}(X, Y) = \operatorname{Kov}(Y, X)$
3. $\operatorname{Var}(X + Y) = \operatorname{Var} X + \operatorname{Var} Y + 2 \operatorname{Kov}(X, Y)$
4. X, Y stochastisch unabhängig $\implies$ $\operatorname{Kov}(X, Y) = 0$

Bem. Die Umkehrung von 4. gilt nicht.

Satz 18 (Regressionsgerade) Seien $X, Y : \Omega \to \mathbb{R}$ Zufallsvariablen mit endlichen EX^2, EY^2. Weiter seien $\operatorname{Var} X, \operatorname{Var} y \neq 0$. Die mittlere quadratische Abweichung zwischen Y und $a + bX$ ist minimal, wenn gilt:

$$a = EY - bEX, \qquad b = \operatorname{korr}(X, Y) \frac{\operatorname{Str} X}{\operatorname{Str} Y}$$

7 Grundlagen der Statistik

Bisher war $(\Omega, \mathcal{A}, P)$ bzw. P^X das zugrundegelegte Modell und P, P^X waren bekannt. In der Statistik ist nun die Verteilung P unbekannt. Stattdessen sind Kenngrößen und/oder „Familien" von Verteilungen vorgegeben, die die „wahre" Verteilung P enthalten.

Def. 23 (Statistisches Modell) Sei Ω ein Merkmalraum (Stichprobenumfang), $\mathcal{A}$ eine σ-Algebra über Ω und W eine Menge von Verteilungen über $(\Omega, \mathcal{A})$. Dann heißt $(\Omega, \mathcal{A}, W)$ *statistisches Modell.*

Bem. Im Rahmen der Vorlesung gelte stets: $\Omega := \mathbb{R}^n, \mathcal{A} := \mathcal{B}^n$ und W bestehe aus der Familie von Produktmaßen, d.h.:

$$W := \{\bigotimes_{1}^{n} Q : Q \in \mathcal{V}\}$$

mit $\mathcal{V} :=$ {W-Maße über $\mathcal{B}$ mit endlichem Erwartungswert und endlicher Varianz}.
Des weiteren seien die Stichproben $X_1, \ldots, X_n$ unabhängig und identisch verteilt (u.i.v.)

8 Punktschätzung

Def. 24 (Schätzfunktion) Sei $(X_1, \ldots, X_n)$ eine Stichprobe und g eine Kenngröße von $P^{X_i} \in \mathcal{V}$. Eine Funktion $h = h(X_1, \ldots, X_n)$ heißt *Schätzfunktion*, wenn ihre Funktionswerte als Schätzungen für g betrachtet werden.

Bem. Der Erwartungswert von h sollte mit g übereinstimmen und die Varianz sollte minimal sein.

Def. 25 (Erwartungstreuer Schätzer) Eine Schätzfuktion $h = h(X_1, \ldots, X_n)$ für $g = g(P^{X_i})$ heißt *erwartungstreuer Schätzer*, wenn gilt:

$$\forall \quad P^{X_i} \in \mathcal{V}: \qquad Eh(X_1, \ldots, X_n) = g(P^{X_i})$$

Bsp. Die Schätzfunktion h mit

$$h(X_1, \ldots, X_n) = \frac{1}{n} \sum_{i=1}^{n} X_i$$

(arithmetisches Mittel) ist für alle möglichen $\mathcal{V}$ ein erwartungstreuer Schätzer für die Kenngröße $g = EX_1$

Def. 26 (Minimum-Varianz-Schätzer) Eine Schätzfunktion $h = h(X_1, \ldots, X_n)$ für eine Kenngröße $g = g(P^{X_i})$ heißt *Minimum-Varianz-Schätzer*, wenn h erwartungstreu ist und die Varianz von h für alle Verteilungen P^{X_i} minimal ist.

Satz 19 (Schätzung von $\operatorname{Var} X_1$ bei unbekanntem EX_1) Für $\operatorname{Var} X_1$ ist $\overline{V}_n$ mit

$$\overline{V}_n := \frac{1}{n-1} \left[\sum_{i=1}^{n} X_i^2 - \frac{1}{n} \left(\sum_{i=1}^{n} X_i \right)^2 \right]$$

ein erwartungstreuer Schätzer.

Bem.

- Man beachte, daß im Nenner $(n-1)$ statt – wie erwartet – n steht.
- Bei bekanntem EX_1 setze man

  $$V_n := \frac{1}{n} \sum_{i=1}^{n} (X_i - EX_i)^2$$

  als Schätzfunktion. V_n ist auch erwartungstreu.

9 Intervallschätzung

Def. 27 (Konfidenzintervall) Sei $(X_1, \ldots, X_n)$ eine Stichprobe mit Verteilungsfamilie $\mathcal{V}$. Ein Intervall

$$I = [\underline{g}(X_1, \ldots, X_n), \overline{g}(X_1, \ldots, X_n)]$$

heißt *Intervallschätzung* oder *Konfidenzintervall* zum Niveau $1-\alpha$ für die Kenngröße $g = g(P^{X_i})$ mit $P^{X_i} \in \mathcal{V}$, wenn gilt:

$$\forall \quad P^{X_i} \in \mathcal{V} : P(g \in I) \geq 1 - \alpha$$

Die Funktionen $\underline{g}, \overline{g} : \mathbb{R}^n \to \mathbb{R}$ heißen *Intervallschätzer*, $(1 - \alpha)$ heißt *Sicherheitsniveau*.

Satz 20 (Intervallschätzung für die Normalverteilung) Ist $\mathcal{V} = \{\mathcal{N}(a, \sigma_0^2), a \in \mathbb{R}, \sigma_0 \text{ fest}\}$, so ergibt sich zum Niveau $(1-\alpha)$ für $EX = a$ das Konfidenzintervall

$$[\overline{X}_n - u_{1-\frac{\alpha}{2}} \frac{\sigma_0}{\sqrt{n}}, \overline{X}_n + u_{1-\frac{\alpha}{2}} \frac{\sigma_0}{\sqrt{n}}]$$

wobei u_x das x-Quantil der $\mathcal{N}(0,1)$-Verteilung sei.

10 Maximum-Likelihood-Schätzer

Gesucht wird eine Schätzfunktion für den Parameter einer Verteilung

Def. 28 (Likelihood-Funktion) Die Abbildung L_ϑ mit

$$L_\vartheta(X_1, \ldots, X_n) := \prod_{i=1}^{n} f_\vartheta(x_i)$$

heißt *Likelihood-Funktion*. Dabei ist $X_1, \ldots, X_n$ die Stichprobe, f_ϑ die (mit ϑ parametrisierte) Dichte der Verteilung der X_i.

Vorgehensweise

Bestimme die Maximalstelle der Likelihood-Funktion in Abhängigkeit der Stichprobe $X_1, \ldots, X_n$. Dieser Term ist eine Schätzfunktion für ϑ.

Trick

Durch Logarithmieren der Likelihodd-Funktion ändern sich die Extremstellen nicht, die Terme werden aber einfacher.

Bsp. Sei $X \sim F_\vartheta$ eine Zufallsvariable mit

$$F_\vartheta(x) := \left(1 - (1 + x^2)^\vartheta\right) \cdot \mathbf{1}_{(x>0)}(x)$$

Bestimmen Sie den Maximum-Likelihood-Schätzer für X.

$$f_\vartheta(x) = F'_\vartheta(x) = \frac{2x\vartheta}{(1+x^2)^{\vartheta+1}} \cdot \mathbf{1}_{(x>0)}(x)$$

$$\Longrightarrow \quad L_\vartheta(x_1,\ldots,x_n) = \prod_{i=1}^n f_\vartheta(x_i) = \prod_{i=1}^n \frac{2x_i\vartheta}{(1+x_i^{\vartheta+1})}$$

$$\Longrightarrow \quad \ln L_\vartheta(x_1,\ldots,x_n) = n\ln 2 + n\ln\vartheta + \sum_{i=1}^n \ln x_i - \sum_{i=1}^n \ln(1+x_i^2)$$

$$\Longrightarrow \quad \frac{\partial \ln L_\vartheta(x_1,\ldots,x_n)}{\partial x} = \frac{n}{\vartheta} - \sum_{i=1}^n \ln(1+x_i^2)$$

$$\Longrightarrow \quad \vartheta = \frac{n}{\sum\limits_{i=1}^n \ln(1+x_i^2)} \quad \text{ist M-L-Schätzer.}$$

11 Testen von Hypothesen

Ziel ist es, eine Entscheidung zu treffen, z.B. ob ein Medikament auf den Markt gebracht werden sollte, wenn k von n Personen durch die Verwendung genesen sind und man vom Medikament eine „Genesungswahrscheinlichkeit“ von p fordert.

Vorgehensweise

1. Formuliere das Gegenteil des Wunsches als Hyphothese H. Ziel wird es sein, H abzulehnen.
2. Wähle eine Stichprobe $T = (X_1,\ldots,X_n)$ und die Irrtumswahrscheinlichkeit α.
3. Bestimme die zu testende Zufallsvariable X und ihre Verteilung.
4. Bestimme den Ablehnungsbereich A zu H, also den Bereich, in dem H abgeleht wird, falls $T \in A$.
5. Entscheide durch Test:
   $T \in A \quad \Longrightarrow \quad H$ ablehen
   $T \notin A \quad \Longrightarrow \quad H$ nicht ablehnen

Bsp. Es wurde ein Medikament entwickelt, das mit einer Wahrscheinlichkeit p wirkt. Es soll vermarktet werden, wenn $p > 0.5$ ist. Getestet werden soll dies zum Niveau 90%.

1. Hypothese $H : p \in [0; \frac{1}{2})$
2. Stichprobe (aus klinischem Test): 14 von 20 Patienten wurden gesund.
   $\alpha = 1 - 0.9 = 0.1$
3. X : „Anzahl der genesenen Patienten“, $X \sim B(n,p)$
4. $A = \{k,\ldots,20\}$ (mit zunächst freiem k)

5. Bedingung: $\forall \quad p \in [0; \frac{1}{2}) : P_p(T \in A) \leq \alpha$
   hier ist

$$P_p(T \in A) = P_p(X \geq k) = \sum_{j=k}^{n} \binom{n}{j} p^j (1-p)^{n-j}$$

   Mit Ablesung aus der Tabelle folgt

$$P(T \in A) \leq \alpha \quad \Longleftrightarrow \quad k \geq 15$$

   $\Longrightarrow \quad T \notin A \quad \Longrightarrow \quad H$ wird abgelehnt
   $\Longrightarrow$ Medikament wird nicht vermarktet.

Fehler

Fehler 1. Art Die Hypothese wird abgelehnt, obwohl sie wahr ist.

Fehler 2. Art Die Hypothese wird nicht abgelehnt, obwohl sie falsch ist.

Die Hypothese sollte so gewählt sein, daß der Fehler erster Art der gravierendere ist. Dessen Größe kann man mit geeignetem α regulieren. Der Fehler zweiter Art ergibt sich daraus zwangsläufig.

12 χ^2-Anpassungstest

Von einer Zufallsvariable ist die Verteilungsklasse nicht bekannt. Es soll getestet werden, ob die Zufallsvariable nach einer bestimmten Verteilung verteilt ist.

Vorgehensweise

Die Zufallsvariable sei X.

1. Wähle eine Verteilung V und ein Testniveau $1 - \alpha$ aus.
2. Annahme: X ist nicht nach V verteilt.
3. $H_i := \#\{$ Versuche mit Ergebnis i$\}$ (absolute Häufigkeiten)
4. Es sei

$$Z_k := \sum_{i=1}^{n} \frac{(H_i - np_i)^2}{np_i}$$

5. Die Hypothese wird abgelehnt, falls Z_k im Ablehnungsbereich

$$A := (\chi^2_{k-1,1-\alpha}, \infty)$$

   liegt.

Bsp. Ein Würfel soll auf Gleichverteilung getestet werden.

1. $V := L(\{1, 2, \ldots, 6\}); \quad \alpha - 1 := 0.9$
2. Annahme: $V \neq L(\{1, 2, \ldots, 6\})$

3. absolute Häufigkeit nach $n = 200$ Würfen:

i	1	2	3	4	5	6
H_i	32	35	41	38	28	26

4. Mit $p_i = \frac{1}{6} \quad (i \in \{1, \ldots, 6\}), n = 200$ und H_i aus obiger Tabelle ergibt sich

$$Z_k = \sum_{i=1}^{k} \frac{(H_i - np_i)^2}{np_i} = 4.03329$$

5. $A = (\chi^2_{5,0.9}, \infty) = (9.36, \infty)$.
   $Z_k \notin A \quad \Longrightarrow \quad$ Es könnte eine Gleichverteilung sein.

13 Test auf Unabhängigkeit

Aus der Angabe von relativen Häufigkeiten für Zufallsvarialben X und Y soll mit einer vorgegebenen Wahrscheinlichkeit untersucht werden, ob X und Y sochastisch unabhängig sind.

Satz 21 (Test auf Unabhängigkeit) $X : \Omega \to \{1, 2, \ldots, r\}; Y : \Omega \to \{1, 2, \ldots, s\}$ seien Zufallsvariablen. Mit einer Wahrscheinlichkeit von $1 - \alpha$ kann man X und Y als stochastisch unabhängig annehmen, wenn mit

$$H_{jk} := \sum_{i=1}^{n} \mathbf{1}_{\{Y_i = j; Z_i = k\}} \qquad H_j^Y := \sum_{i=1}^{n} \mathbf{1}_{\{Y_i = j\}}$$

$$H_k^Z := \sum_{i=1}^{n} \mathbf{1}_{\{Z_i = k\}} \qquad p_{jk} := \frac{H_j^Y \cdot H_k^Z}{n^2}$$

gilt:

$$Z := \sum_{j=1}^{r} \sum_{k=1}^{s} \frac{(H_{jk} - np_{jk})^2}{np_{jk}} > \chi^2_{(r-1)(s-1),1-\alpha}$$

Bsp. Die Verteilung zweier Zufallsvariablen Y, Z sei gegeben durch

	Z_1	Z_2	Z_3	$H_j^Y = \sum$
Y_1	11	13	5	29
Y_2	16	17	11	44
Y_3	9	34	21	64
$H_k^Z = \sum$	36	64	37	$137 = n$

Prüfen sie Y und Z zu einem Niveau von 0.95 auf Unabhängigkeit.

$$((np_{jk})) = \begin{pmatrix} 7.6 & 13.5 & 7.8 \\ 11.5 & 20.6 & 11.9 \\ 16.8 & 29.9 & 17.3 \end{pmatrix}$$

$\Longrightarrow \quad Z = 9.924$
$\chi^2_{(3-1)(3-1);0.95} = \chi^2_{4;0.95} = 0.488 < Z$
$\Longrightarrow \quad$ Hypothese kann nicht abgelehnt werden.

Index

B

Bayes, Formel von . . . 2
Bildmaß . . . 5
Bildmodell . . . 5

C

Chi-Quadrat-Anpassungstest 14–15

E

Erwartungstreue . . . 11
Erwartungswert . . . 8, 9

F

Faltung . . . 7
Fehler
  erster Art . . . 14
  zweiter Art . . . 14

H

Hypothesen . . . 13–14

I

Intervallschätzer . . . 12
Intervallschätzung . . . 11

K

Kenngrößen . . . 7–10
Konfidenzintervall . . . 11
Kopplung . . . 4
Korrelationskoeffizient . . . 9
Kovarianz . . . 9
  normierte . . . 10

L

Likelihood-Funktion . . . 12

M

Median . . . 7
Meßbarkeit . . . 5
Modalwert . . . 7
Modell, statistisches . . . 10

P

Projektion . . . 6
Punktschätzung . . . 11

Q

Quantil . . . 7

R

Randverteilung . . . 6
Regressionsgerade . . . 10
Riemanndichte . . . 2–3

S

Schätzer . . . 11
  erwartungstreuer . . . 11
  Maximum-Likelihood- . . . 12–13
  Minimum-Varianz- . . . 11
Schätzfunktion . . . 11
Sicherheitsniveau . . . 12
Statistik . . . 10–15
statistisches Modell . . . 10
Stichprobe
  unabhängig identisch verteilte 10
Streuung . . . 9

U

Übergangszähldichten . . . 4
Unabhängigkeit
  stochastische . . . 2
  Test auf . . . 15
  Zufallsvariablen . . . 6
Urbildfunktion . . . 5

V

Varianz . . . 9
Versuch, n-stufiger . . . 4
Verteilung
  gemeinsame . . . 6
  gemischte . . . 3
Verteilungsfunktion . . . 3
Vorgeschichte . . . 4

W

Wahrscheinlichkeit
  bedingte . . . 1–2
  totale . . . 2

Z

Zähldichte . . . 2–3
Zufallsvariable . . . 4–7
  Unabhängigkeit . . . 6
  Verteilung . . . 5