Kombinatorik

	Mit Wiederholung	**Ohne Wiederholung**
Variation (Auswahl *mit* Reihenfolge)	$\overline{V_n^k} = n^k$	$V_n^k = \binom{n}{k} \cdot k!$
Kombination (Auswahl *ohne* Reihenfolge)	$\overline{K_n^k} = \binom{n+k-1}{k}$	$K_n^k = \binom{n}{k}$
Permutation (Anordnung)	$P_n^{n_1,\dots,n_r} = \dfrac{n!}{n_1! \cdot n_2! \cdots n_r!}$	$P_n = n!$

Grundlagen der Wahrscheinlichkeitsrechnung

Laplace-Whk.

$$P(A) = \frac{|A|}{|\Omega|} = \frac{\#\text{Günstige Ereignisse}}{\#\text{Alle Ereignisse}}$$

Voraussetzung: Endlich viele Ereignisse, alle mit der selben Whk.

Siebformel

$$P(A \cup B) = P(A) + P(B) - P(A \cap B)$$

Bedingte Whk.

$$P(A|B) = \frac{P(A \cap B)}{P(B)}, \quad P(B) \neq 0$$

Totale Wahrscheinlichkeit

$$P(A) = P(A|B_1) \cdot P(B_1) + \cdots + P(A|B_n) \cdot P(B_n)$$

Satz von Bayes

$$P(A|B_i) \cdot P(B_i) = P(B_i|A) \cdot P(A)$$

Voraussetzung: $\bigcup_{i=1}^{n} B_i = \Omega, \quad B_i \cap B_j = \varnothing \quad \forall i \neq j$

Unabhängigkeit (Ereignisse)

$$A, B \text{ unabhängig} \quad \Longleftrightarrow \quad P(A \cap B) = P(A) \cdot P(B)$$

Zufallsverteilungen

	Diskret (zählbar)	**Stetig** (überabzählbar)
Verteilungsfunktion	$F_X(x) := \mathbb{P}(X \le x) = \sum_{x_i \le x} p_i$	$F_X(x) := \mathbb{P}(X \le x) = \int_{-\infty}^{x} f(t)\,\mathrm{d}t$ (1) $\mathbb{P}(X \le b) = F_X(b)$ (2) $\mathbb{P}(X > a) = 1 - F_X(a)$ (3) $\mathbb{P}(a < X \le b) = F_X(b) - F_X(a)$
Whk.- & Dichtefunktion	$p_i := \mathbb{P}(X = x_i) \in [0,1]$ $\sum_{i \in I} p_i = 1$	$f(x) = \frac{\mathrm{d}}{\mathrm{d}x} F_X(x) \ge 0$ $\int_{\mathbb{R}} f(x)\,\mathrm{d}x = 1$
Erwartungswert	$\mathbb{E}(X) := \sum_{i \in I} x_i \cdot p_i,$ falls $\mathbb{E}(\lvert X \rvert) = \sum_{i \in I} \lvert x_i \rvert \cdot p_i < \infty$	$\mathbb{E}(X) := \int_{\mathbb{R}} x \cdot f(x)\,\mathrm{d}x,$ falls $\mathbb{E}(\lvert X \rvert) = \int_{\mathbb{R}} \lvert x \rvert \cdot f(x)\,\mathrm{d}x < \infty$
Varianz	$\mathrm{Var}(X) := \mathbb{E}\Big((X - \mathbb{E}(X))^2\Big) \overset{\circledast}{=} \mathbb{E}(X^2) - [\mathbb{E}(X)]^2$ $\circledast$ falls $\mathbb{E}(X^2) = \sum_{i \in I} x_i^2 \cdot p_i < \infty$	$\mathrm{Var}(X) := \mathbb{E}\Big((X - \mathbb{E}(X))^2\Big) \overset{\circledast}{=} \mathbb{E}(X^2) - [\mathbb{E}(X)]^2$ $\circledast$ falls $\mathbb{E}(X^2) = \int_{\mathbb{R}} x^2 \cdot f(x)\,\mathrm{d}x < \infty$

Rechenregeln für Erwartungswert & Varianz

	Erwartungswert	**Varianz**
Addititivät	$\mathbb{E}(X+Y) = \mathbb{E}(X) + \mathbb{E}(Y)$	$\mathrm{Var}(X+Y) = \mathrm{Var}(X) + \mathrm{Var}(Y) + 2 \cdot \mathrm{Cov}(X,Y)$
Skalare Multiplikation	$\mathbb{E}(aX) = a\mathbb{E}(X)$	$\mathrm{Var}(aX) = a^2\mathrm{Var}(X)$
Konstante	$\mathbb{E}(a) = a$	$\mathrm{Var}(a) = 0$

Beispiel: Seien $X_1, \ldots, X_n$ i.i.d. mit $\mathbb{E}(X_i) = \mu$ und $\mathrm{Var}(X_i) = \sigma^2$.

$$S_n := \sum_{i=1}^{n} X_i \implies \mathbb{E}(S_n) = n\mu, \quad \mathrm{Var}(S_n) = n\sigma^2$$

$$\overline{X}_n := \frac{1}{n}\sum_{i=1}^{n} X_i \implies \mathbb{E}(\overline{X}_n) = \mu, \quad \mathrm{Var}(\overline{X}_n) = \frac{\sigma^2}{n}$$

Verteilungsmodelle

Gleichverteilung
$X \sim \mathrm{U}\{x_1, \dots, x_n\}$

Jedes Ereignis $x_1, x_2, \dots, x_n$ ist gleich wahrscheinlich.

$$\mathbb{P}(X = k) = \frac{1}{n}, \quad \left(\text{für } x_k = k: \quad \mathbb{E}(X) = \frac{n+1}{2}, \quad \mathrm{Var}(X) = \frac{n^2-1}{12}\right)$$

Binomialverteilung
$X \sim \mathrm{B}(n, p)$

$X = \{\text{Anzahl Erfolge bei } n \text{ Versuchen } (\textbf{mit} \text{ Wiederholung})\}$

- **Mit** Wiederholung = Jeder Erfolg hat die selbe Whk. p
- Feste Anzahl an Versuchen

$$\mathbb{P}(X = k) = \binom{n}{k} p^k (1-p)^{n-k}, \quad \mathbb{E}(X) = np, \quad \mathrm{Var}(X) = np(1-p)$$

Hypergeometrische Verteilung
$X \sim \mathrm{Hyp}(N, M, n)$

$X = \{\text{Anzahl Erfolge bei } n \text{ Versuchen } (\textbf{ohne} \text{ Wiederholung})\}$

- **Ohne** Wiederholung
- Feste Anzahl an Versuchen

$$\mathbb{P}(X = k) = \frac{\binom{M}{k}\binom{N-M}{n-k}}{\binom{N}{n}}, \quad \mathbb{E}(X) = n\frac{M}{N}, \quad \mathrm{Var}(X) = n\frac{M(N-M)}{N^2}\frac{N-n}{N-1}$$

Poissonverteilung
$X \sim \mathrm{Poi}(\lambda)$

$X = \{\text{Anzahl Erfolge (innerhalb einer Zeit)}\}$

- „Verteilung der seltenen Ereignisse“
- Beliebig viele Erfolge möglich (innerhalb einer Zeit)

$$\mathbb{P}(X = k) = \frac{\lambda^k}{k!} \mathrm{e}^{-\lambda}, \quad \mathbb{E}(X) = \lambda, \quad \mathrm{Var}(X) = \lambda$$

Geometrische Verteilung
$X \sim \mathrm{G}(p)$

$X = \{\text{Anzahl Versuche } \textbf{bis} \text{ (einschließlich) zum ersten Erfolg}\}$

- Beim Eintreffen des ersten Erfolges wird das Experiment beendet
- Jeder Erfolg hat die selbe Wahrscheinlichkeit p
- Gedächtnislosigkeit: $\mathbb{P}(X \leq k + n \mid X > n) = \mathbb{P}(X \leq k)$

$$\mathbb{P}(X = k) = p(1-p)^{k-1}, \quad \mathbb{E}(X) = \frac{1}{p}, \quad \mathrm{Var}(X) = \frac{1-p}{p^2}$$

Gleichverteilung
$X \sim \mathrm{U}[a,b]$

Alle Teilintervalle von $[a,b]$ gleicher Länge sind gleich wahrscheinlich.

$$f_X(t) = \begin{cases} \frac{1}{b-a} & t \in [a,b] \\ 0 & \text{sonst} \end{cases}, \quad F_X(x) = \begin{cases} 0 & x < a \\ \frac{x-a}{b-a} & x \in [a,b] \\ 1 & x > b \end{cases}$$

$$\mathbb{E}(X) = \frac{b+a}{2}, \quad \mathrm{Var}(X) = \frac{(b-a)^2}{12}$$

Exponential-verteilung
$X \sim \mathrm{Exp}(\lambda)$

$X = \{\text{Dauer von Zeitintervallen}\}$, z.B. Wartezeit, Lebensdauer, …

- Gedächtnislosigkeit: $P(X \leq x+t \,|\, X > t) = P(X \leq x)$
- $X = \{\text{Wartezeit zw. 2 Erfolgen}\}$, $Y = \{\text{Anzahl Erfolge (innerhalb 1 Zeit-EH)}\}$
  $X \sim \mathrm{Exp}(\lambda) \Longleftrightarrow Y \sim \mathrm{Poi}(\lambda)$

$$f_X(t) = \begin{cases} \lambda \mathrm{e}^{-\lambda t} & t \geq 0 \\ 0 & \text{sonst} \end{cases}, \quad F_X(x) = \begin{cases} 1 - \mathrm{e}^{-\lambda x} & x \geq 0 \\ 0 & \text{sonst} \end{cases}$$

$$\mathbb{E}(X) = \frac{1}{\lambda}, \quad \mathrm{Var}(X) = \frac{1}{\lambda^2}$$

Normal-verteilung
$X \sim \mathcal{N}(\mu, \sigma^2)$

$X = \{\text{Messfehler/Abweichungen/}\dots\}$

- Verteilung symmetrisch um den Maximalwert der Dichte $x_{\max} = \mu$
- Wendepunkte der Dichte bei $x_{\mathrm{w}} = \mu \pm \sigma$
- Standardisieren: $Z = \frac{X-\mu}{\sigma} \sim \mathcal{N}(0,1)$

$$f_X(t) = \frac{1}{\sqrt{2\pi}\sigma} \mathrm{e}^{-\frac{1}{2}\left(\frac{t-\mu}{\sigma}\right)^2}, \quad F_X(x) = \Phi\left(\frac{x-\mu}{\sigma}\right), \quad \mathbb{E}(X) = \mu, \quad \mathrm{Var}(X) = \sigma^2$$

Approximation von Verteilungen

Näherungen

$$\begin{tikzcd}
\mathbf{Hyp}(N,M,n) \arrow[r, "\frac{n}{N} < 0.05", "p = \frac{M}{N}"'] \arrow[dr, "n\frac{M}{N}\left(1-\frac{M}{N}\right) \geq 9"'] & \mathbf{B}(n,p) \arrow[r, "\substack{np \leq 10 \\ n/p \geq 1500}", "\lambda = np"'] \arrow[d, "np(1-p) \geq 9"] & \mathbf{Poi}(\lambda) \arrow[dl, "\lambda \geq 9"] \\
& \mathcal{N}\left(\mu = \mathbb{E}(X),\, \sigma^2 = \mathrm{Var}(X)\right) &
\end{tikzcd}$$

Univariate Statistik

Grundgesamtheit

- Menge der Elemente/Untersuchungseinheiten
- Identifikationskriterien = **Festgelegte** Eigenschaften *(örtlich, zeitlich, sachlich)*

„Menge der Teilnehmenden an einem Statistik Kurs im deutschsprachigen Raum im Jahr 2025."

Erhebungsmerkmale

- **Variable** Eigenschaften der Elemente/Untersuchungseinheiten
- Untersuchte Merkmalswerte = Ausprägungen

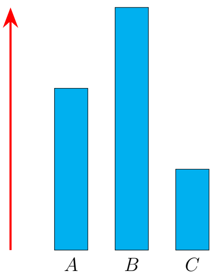
„Geschlecht: weiblich, männlich, …"
„Schulnoten: 1, 2, 3, 4, 5, 6"
„Einkommen: 0 €, 500 €, 1000 €, 2000 €, …"

Mächtigkeit

- diskret = Anzahl der Ausprägungen ist **zählbar**
- stetig = Anzahl der Ausprägungen ist **überabzählbar**
- quasistetig = diskret, aber aus praktischen Gründen als stetig behandelt
- diskretisiert = stetig, aber nur in diskreter Form erfassbar

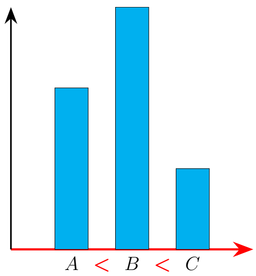
Skalenniveaus
von Merkmalen

Nominalskala



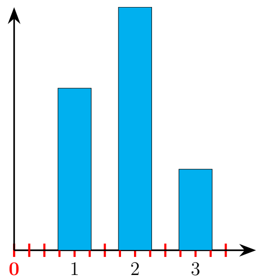
- Säulen-, Balken-, Kreisdiag.
- Abs. & Rel. Häufigkeiten
- Modus
- χ^2-basierte Kontingenzkoeffizienten

Ordinalskala



- Boxplot
- Emp. Verteilungsfkt.
- p-Quantile (Bsp. Median)
- Spannweite
- Quartilabstand
- Dispersionskoeffizient
- Konkordanz, Diskordanz
- γ, τ_a, τ_b, τ_c, r_S - Korr.-koeffs.

Kardinalskala



- Histogramm
- Artihm./Harm./Geom. Mittel
- Varianz/Standardabweichung
- Variationskoeffizient
- Konzentrationsrate
- Konzentrationskurve
- Herfindahl-Index
- Lorenzkurve
- Gini-Koeffizient
- Kovarianz
- r_B - Korr.-koeff.
- Regressionskoeffizienten

UNKLASSIERTE DATEN
(geordnete Gruppen $j = 1, \dots, k$)

KLASSIERTE DATEN
(innerhalb gleichverteilt)

Empirische Verteilungsfunktion $F(x) =$

Unklassierte Daten:

$$F(x) = \begin{cases} 0 & x < x_1 \\ \sum_{i=1}^{j} f_i & x_j \le x < x_{j+1} \\ 1 & x \ge x_k \end{cases}$$

Klassierte Daten:

$$F(x) = \begin{cases} 0 & x < x_1 \\ \sum_{i=1}^{j-1} f_i + \dfrac{x - x_j^{\mathrm{u}}}{x_j^{\mathrm{o}} - x_j^{\mathrm{u}}} f_j & x_j^{\mathrm{u}} \le x < x_j^{\mathrm{o}} \\ 1 & x \ge x_k \end{cases}$$

	UNKLASSIERTE DATEN (geordnete Gruppen $j = 1, \dots, k$)	**KLASSIERTE DATEN** (innerhalb gleichverteilt)

Lagemaße

	Unklassierte Daten	Klassierte Daten
Modus $x_{\text{mod}} =$	Häufigste Ausprägung	$x^{\text{u}}_{\text{M}} + \frac{f^*_{\text{M}} - f^*_{\text{M}-1}}{2f^*_{\text{M}} - f^*_{\text{M}-1} - f^*_{\text{M}+1}} \cdot (x^{\text{o}}_{\text{M}} - x^{\text{u}}_{\text{M}})$ **Hinweis:** $f_i^* = \frac{f_i}{x_i^{\text{o}} - x_i^{\text{u}}}$ *„Dichte der Klasse i“*
p-Quantil $x_p =$	$\begin{cases} x_{(\lceil np \rceil)} & np \notin \mathbb{Z} \\ \dfrac{x_{(np)} + x_{(np+1)}}{2} & np \in \mathbb{Z} \end{cases}$	$x^{\text{u}}_{\text{Q}} + \frac{p - F_{\text{Q}-1}}{f_{\text{Q}}} \cdot (x^{\text{o}}_{\text{Q}} - x^{\text{u}}_{\text{Q}})$
Median	$x_{\text{med}} =$ Das mittlere **50%-Quantil** $x_{0{,}5}$	
Arithmetisches Mittel $\bar{x} =$	$\frac{1}{n}\sum_{i=1}^{k} n_i x_i = \sum_{i=1}^{k} f_i x_i$	$\frac{1}{n}\sum_{i=1}^{k} n_i \bar{x}_i = \sum_{i=1}^{k} f_i \bar{x}_i$
Harmonisches Mittel $\bar{x}_{\text{H}} =$	$\frac{n}{\sum_{i=1}^{k} \frac{n_i}{x_i}} = \frac{1}{\sum_{i=1}^{k} \frac{f_i}{x_i}}$	$\frac{n}{\sum_{i=1}^{k} \frac{n_i}{\bar{x}_i}} = \frac{1}{\sum_{i=1}^{k} \frac{f_i}{\bar{x}_i}}$
Geometrisches Mittel $\bar{x}_{\text{G}} =$	$\sqrt[n]{\prod_{i=1}^{k} x_i^{n_i}} = \prod_{i=1}^{k} x_i^{f_i}$	$\sqrt[n]{\prod_{i=1}^{k} \bar{x}_i^{n_i}} = \prod_{i=1}^{k} \bar{x}_i^{f_i}$

Streuungsmaße

Spannweite

$$R = x_{\max} - x_{\min}$$

Quartilabstand

$$Q = x_{0{,}75} - x_{0{,}25}$$

Dispersionskoeffizient

$$q = \frac{Q}{x_{\text{med}}}, \quad x_{\text{med}} > 0$$

Empirische Varianz $\tilde{s}^2 =$ (deskr. Statistik)

Unklassierte Daten:

$$\frac{1}{n}\sum_{i=1}^{k} n_i\,(x_i - \overline{x})^2 = \overline{x^2} - \overline{x}^2$$

Klassierte Daten:

$$\tilde{s}^2_{\text{innerhalb}} + \tilde{s}^2_{\text{zwischen}} = \frac{1}{n}\sum_{i=1}^{k} n_i \tilde{s}_i^2 + \frac{1}{n}\sum_{i=1}^{k} n_i (\overline{x}_i - \overline{x})^2$$

Anmerkung: In der induktiven Statistik wird $s^2 = \frac{n}{n-1}\tilde{s}^2$ verwendet.

Variationskoeffizient

$$v = \frac{\tilde{s}}{\bar{x}}, \quad \bar{x} > 0$$

Abhängigkeitsmaße

Empirische Kovarianz
(deskr. Statistik)

$$\tilde{s}_{xy} = \frac{1}{n}\sum_{i=1}^{n}(x_i - \bar{x})(y_i - \bar{y}) = \overline{xy} - \bar{x}\cdot\bar{y}$$

Anmerkung: In der induktiven Statistik wird $s_{xy} = \frac{n}{n-1}\tilde{s}_{xy}$ verwendet.

Korrelation & Abhängigkeit

$$\begin{aligned} &X, Y \text{ unabhängig} && \Longrightarrow && X, Y \text{ unkorreliert } (s_{xy} = 0) \\ &X, Y \text{ korreliert } (s_{xy} \neq 0) && \Longrightarrow && X, Y \text{ abhängig} \\ & && \cancel{\Longleftarrow} && \end{aligned}$$

Korrelationskoeffizient
Bravais-Pearson r_{B}

$$r_{\mathrm{B}} = \frac{\tilde{s}_{xy}}{\tilde{s}_x\cdot\tilde{s}_y} = \frac{s_{xy}}{s_x\cdot s_y} = \frac{\overline{xy} - \bar{x}\cdot\bar{y}}{\sqrt{\overline{x^2} - \bar{x}^2}\cdot\sqrt{\overline{y^2} - \bar{y}^2}}$$

- Berücksichtigt die *Abstände* zwischen den Beobachtungen
- $r_{\mathrm{B}} \in [-1, 1]$ von „negativ linear abhänig" bis „positive linear abhängig"
- $r_{\mathrm{B}} = 0 \iff s_{xy} = 0 \iff X, Y$ unkorreliert (= linear unabhängig)

Korrelationskoeffizient
Spearman r_{S}

$$r_{\mathrm{S}} = \frac{\tilde{s}_{\mathrm{rg}(x)\mathrm{rg}(y)}}{\tilde{s}_{\mathrm{rg}(x)}\cdot\tilde{s}_{\mathrm{rg}(y)}} = \frac{n^3 - n - \frac{1}{2}(T_x^* + T_y^*) - 6\cdot\sum_{i=1}^{n} d_i^2}{\sqrt{n^3 - n - T_x^*}\cdot\sqrt{n^3 - n - T_y^*}} \overset{\circledast}{=} 1 - \frac{6\cdot\sum_{i=1}^{n} d_i^2}{n(n^2-1)}$$

$\circledast$ falls Ränge eindeutig $\iff T_x^* = T_y^* = 0$ (keine Bindungen)

- Gleiche Ränge? → erstelle k_x, k_y Rangklassen, bilde Durchschnittsränge und $T_\bullet^* = \sum_{j=1}^{k_\bullet}(t_j^3 - t_j)$ mit t_j = Anzahl Beobachtungen gleichen Rangs in Klasse j
- $d_i = \mathrm{rg}(x_i) - \mathrm{rg}(y_i)$ sind die Rangdifferenzen
- $r_{\mathrm{S}} \in [-1, 1]$ von „negativer" bis „positiver" monotoner Zusammenhang

Quadratische Kontingenz

$$\chi^2 = \sum_{i=1}^{k}\sum_{j=1}^{\ell}\frac{\left(n_{ij} - \frac{n_{i\bullet}\cdot n_{\bullet j}}{n}\right)^2}{\frac{n_{i\bullet}\cdot n_{\bullet j}}{n}} = n\cdot\left(\sum_{i=1}^{k}\sum_{j=1}^{\ell}\frac{n_{ij}^2}{n_{i\bullet}\cdot n_{\bullet j}} - 1\right) = n\cdot\left(\sum_{i=1}^{k}\sum_{j=1}^{\ell}\frac{f_{ij}^2}{f_{i\bullet}\cdot f_{\bullet j}} - 1\right)$$

- Es werden die quadratischen Abstände zwischen *beobachteten* und *unter Unabhängigkeit erwarteten* Häufigkeiten ins Verhältnis zu den *unter Unabhängigkeit erwarteten* Häufigkeiten gesetzt und addiert
- $\chi^2 \in [0, n(\min(k,\ell) - 1)]$ von „Unabhängigkeit" bis „exakter Zusammenhang"
- Normiertes Kontingenzmaß (Cramérs) $V = \sqrt{\frac{\chi^2}{n\cdot(\min(k,\ell)-1)}}$
- Normierter (korrigierter) Kontingenzkoeffizient $C_{\mathrm{korr.}} = \sqrt{\frac{\chi^2}{\chi^2+n}}\cdot\sqrt{\frac{\min(k,\ell)}{\min(k,\ell)-1}}$

Regressionsanalyse

Grundidee

$$\sum_{i=1}^{n}(y_i - \hat{y}_i)^2 \to \min$$

- Gesucht ist eine Funktion $\hat{y} := \hat{y}(\boldsymbol{x}, \hat{\boldsymbol{\beta}})$, bei der die Summe der quadratischen (vertikalen) Abweichungen von der vorliegenden Stichprobe minimal ist
- Merkmal X ist gegeben oder beeinflussbar *(unabhängige Variable)*
- Merkmal Y ist beobachtete Reaktion auf X *(abhängige Variable)*

Regressions-gerade

$$\boxed{\hat{y} = \hat{\beta}_0 + \hat{\beta}_1 x} \quad \Longrightarrow \quad \hat{\beta}_0 = \bar{y} - \hat{\beta}_1 \bar{x}, \qquad \hat{\beta}_1 = \frac{\overline{xy} - \bar{x} \cdot \bar{y}}{\overline{x^2} - (\bar{x})^2} \left(= r_{\mathrm{B}} \cdot \frac{\tilde{s}_y}{\tilde{s}_x} = r_{\mathrm{B}} \cdot \frac{s_y}{s_x}\right)$$

- $\hat{\beta}_0$: Schnittpunkt der Regressionsgerade mit y-Achse *(Regressionskonstante)*
- $\hat{\beta}_1$: Anstieg der Regressionsgerade; Durchschnittliche Änderung von Y, wenn sich X um eine Einheit erhöht *(Regressionskoeffizient)*
- $\hat{\beta}_1 = 0 \;\Leftrightarrow\; r_{\mathrm{B}} = 0 \;\Leftrightarrow\;$ X und Y linear unabhängig
- Der Punkt $(\bar{x}, \bar{y})$ liegt auf der Regressionsgerade *(Gravitationszentrum)*
- $\sum\limits_{i=1}^{n} \hat{\varepsilon}_i = \sum\limits_{i=1}^{n} (y_i - \hat{y}_i) = 0$, bzw. $\bar{\hat{y}} = \bar{y}$ *(Fehlerausgleich)*

Schätztheorie

Punktschätzung

Momenten-schätzer

Setze k theoretische und empirische Momente gleich und stelle nach $\hat{\vartheta}_1, \ldots, \hat{\vartheta}_k$ um. z.B. $\mathbb{E}(X^i) \stackrel{!}{=} \overline{X^i}$ für alle $i = 1, \ldots, k$.

Maximum-Likelihood-Schätzer

$$\mathscr{L}(\boldsymbol{x}, \vartheta) = \prod_{i=1}^{n} p_\vartheta(x_i) \to \max \qquad \text{bzw.} \qquad \mathscr{L}(\boldsymbol{x}, \vartheta) = \prod_{i=1}^{n} f_\vartheta(x_i) \to \max$$

Der *Maximum-Likelihood-Schätzer* $\hat{\vartheta}_{\mathrm{ML}}$ ist der Wert, mit dem das Eintreten der vorliegenden Stichprobe $\boldsymbol{x}$ am wahrscheinlichsten ist.

Kleinste-Quadrate-Schätzer

$$Q(\boldsymbol{x}, \vartheta) = \sum_{i=1}^{n} \left(x_i - \mathbb{E}(X_i)\right)^2 \to \min$$

Der *Kleinste-Quadrate-Schätzer* $\hat{\vartheta}_{\mathrm{KQ}}$ ist der Wert, bei dem die Summe der quadratischen Abweichungen von der vorliegenden Stichprobe $\boldsymbol{x}$ minimal ist.

Erwartungs-treue

$$\mathbb{E}(\hat{\vartheta}) = \vartheta$$

Eine Schätzfunktion $\hat{\vartheta}$ heißt *erwartungstreu*, wenn ihr Erwartungswert gleich dem zu schätzenden Parameter ϑ ist.
Anmerkung: Schätzer heißt *asymptotisch erwartungstreu*, falls $\lim\limits_{n\to\infty} \mathbb{E}(\hat{\vartheta}) = \vartheta$.

Effizienz

Eine Schätzfunktion $\hat{\vartheta}$ heißt *effizient*, falls sie *erwartungstreu* ist und die kleinste Varianz unter allen erwartungstreuen Schätzern hat.

Konsistenz

$$P(|\hat{\vartheta} - \vartheta| < \varepsilon) \xrightarrow[n\to\infty]{} 1 \quad (\forall \varepsilon > 0)$$

Eine Schätzfunktion $\hat{\vartheta}$, die *(asymptotisch) erwartungstreu* ist und für die gilt $\mathrm{Var}(\hat{\vartheta}) \xrightarrow[n\to\infty]{} 0$, ist *konsistent*.

Intervallschätzung

Konfidenzintervalle $\mathbb{P}\Big(\vartheta \in C_\vartheta(X)\Big) \geq 1-\alpha$

- Ein $(1-\alpha)$-Konfidenzintervall $C_\vartheta(X)$, basierend auf einem Punktschätzer $\hat{\vartheta}$, enthält den echten (unbekannten) Parameter ϑ in (mindestens) $1-\alpha$ [$\cdot 100\%$] aller möglichen Stichproben.
- Ein Konfidenzintervall wird umso „besser“, je …
  … kleiner/kürzer/genauer das $C_\theta(X)$ ist.
  … kleiner die Irrtumswahrscheinlichkeit α ist.
  … größer der Stichprobenumfang n ist.
- $n = \text{const.}$: Länge $\downarrow \iff \alpha \uparrow$

$C_\mu(x)$ **Normalverteilung oder $n \geq 30$ (ZGWS)**

σ bekannt

$$\left(\bar{x} - \frac{\sigma}{\sqrt{n}} \cdot z_{1-\frac{\alpha}{2}} \; , \; \bar{x} + \frac{\sigma}{\sqrt{n}} \cdot z_{1-\frac{\alpha}{2}}\right)$$

σ unbekannt

$$\left(\bar{x} - \frac{s}{\sqrt{n}} \cdot t_{n-1,1-\frac{\alpha}{2}} \; , \; \bar{x} + \frac{s}{\sqrt{n}} \cdot t_{n-1,1-\frac{\alpha}{2}}\right), \quad \text{mit } s^2 = \frac{n}{n-1}\left(\overline{x^2} - \bar{x}^2\right)$$

$C_{\sigma^2}(x)$ **Normalverteilung**

μ bekannt

$$\left(\frac{n \cdot s_*^2}{\chi^2_{n,1-\frac{\alpha}{2}}}, \frac{n \cdot s_*^2}{\chi^2_{n,\frac{\alpha}{2}}}\right), \quad \text{mit } s_*^2 = \frac{1}{n}\sum_{i=1}^{n}(x_i - \mu)^2$$

μ unbekannt

$$\left(\frac{(n-1) \cdot s^2}{\chi^2_{(n-1),1-\frac{\alpha}{2}}}, \frac{(n-1) \cdot s^2}{\chi^2_{n-1,\frac{\alpha}{2}}}\right), \quad \text{mit } s^2 = \frac{1}{n-1}\sum_{i=1}^{n}(x_i - \bar{x})^2 = \frac{n}{n-1}\left(\overline{x^2} - \bar{x}^2\right)$$

$C_p(x)$ **Binomialverteilung und $n\bar{x}(1-\bar{x}) > 9$ (ZGWS)**

$$\left(\bar{x} - \sqrt{\frac{\bar{x}(1-\bar{x})}{n}} \cdot z_{1-\frac{\alpha}{2}} \; , \; \bar{x} + \sqrt{\frac{\bar{x}(1-\bar{x})}{n}} \cdot z_{1-\frac{\alpha}{2}}\right)$$

Testtheorie

Hypothesen

H_0 : Was bisher angenommen wurde. „Das Alte“
H_1 : Der Verdacht (Gegenteil von H_0). „Das Neue“

Fehlerarten

Test \ Realität	H_0 wahr	H_1 wahr
H_0 behalten	✓	Fehler 2. Art
H_1 wählen	Fehler 1. Art	✓

- Fehler 1. Art ist schlimmer als Fehler 2. Art
- Um ihn zu kontrollieren: P(**„Fehler 1. Art“**) $\leq \alpha$ festlegen
- Test wählen, der P(**„Fehler 2. Art“**) minimiert

Testschema

(1) Signifikanzniveau α festlegen
(2) Hypothesen bestimmen („=“ immer bei H_0)
(3) Testgröße berechnen
(4) Ablehnungsbereich für H_0 aufstellen („Kritischer Bereich“)
(5) Entscheidung

Einfacher Gaußtest (μ)
σ bekannt

H_0	H_1	H_0 ablehnen, falls:
$\mu = \mu_0$	$\mu \neq \mu_0$	$T \in \left(-\infty\ ,\ -z_{1-\frac{\alpha}{2}}\right) \cup \left(z_{1-\frac{\alpha}{2}}\ ,\ \infty\right)$
$\mu \geq \mu_0$	$\mu < \mu_0$	$T \in (-\infty\ ,\ -z_{1-\alpha})$
$\mu \leq \mu_0$	$\mu > \mu_0$	$T \in (z_{1-\alpha}\ ,\ \infty)$

$$T = \frac{\bar{x} - \mu_0}{\sigma}\sqrt{n}, \quad T \sim \mathcal{N}(0,1) \text{ unter } H_0.$$

Einfacher t-Test (μ)
σ unbekannt

H_0	H_1	H_0 ablehnen, falls:
$\mu = \mu_0$	$\mu \neq \mu_0$	$T \in \left(-\infty\ ,\ -t_{n-1;\,1-\frac{\alpha}{2}}\right) \cup \left(t_{n-1;\,1-\frac{\alpha}{2}}\ ,\ \infty\right)$
$\mu \geq \mu_0$	$\mu < \mu_0$	$T \in (-\infty\ ,\ -t_{n-1;\,1-\alpha})$
$\mu \leq \mu_0$	$\mu > \mu_0$	$T \in (t_{n-1;\,1-\alpha}\ ,\ \infty)$

$$T = \frac{\bar{x} - \mu_0}{s}\sqrt{n}, \quad T \sim t_{n-1} \text{ unter } H_0.$$

Einfacher σ-Test
μ unbekannt

H_0	H_1	H_0 ablehnen, falls:
$\sigma = \sigma_0$	$\sigma \neq \sigma_0$	$T \in \left[0\ ,\ \chi^2_{n-1;\,\frac{\alpha}{2}}\right) \cup \left[\chi^2_{n-1;\,1-\frac{\alpha}{2}}\ ,\ \infty\right)$
$\sigma \geq \sigma_0$	$\sigma < \sigma_0$	$T \in \left[0\ ,\ \chi^2_{n-1;\,\alpha}\right)$
$\sigma \leq \sigma_0$	$\sigma > \sigma_0$	$T \in \left[\chi^2_{n-1;\,1-\alpha}\ ,\ \infty\right)$

$$T = (n-1)\frac{s^2}{\sigma_0^2}, \quad T \sim \chi^2_{n-1} \text{ unter } H_0.$$

Anteilswertetest (p)

$np_0(1-p_0) > 9$

H_0	H_1	H_0 **ablehnen, falls:**
$p = p_0$	$p \neq p_0$	$T \in \left(-\infty\ ,\ -z_{1-\frac{\alpha}{2}}\right) \cup \left(z_{1-\frac{\alpha}{2}}\ ,\ \infty\right)$
$p \geq p_0$	$p < p_0$	$T \in (-\infty\ ,\ -z_{1-\alpha})$
$p \leq p_0$	$p > p_0$	$T \in (z_{1-\alpha}\ ,\ \infty)$

$$T = \frac{\bar{x} - p_0}{\sqrt{p_0(1-p_0)}}\sqrt{n}, \quad T \sim \mathcal{N}(0,1) \text{ unter } H_0.$$

Binomialtest (p)

$np_0(1-p_0) \leq 9$

H_0	H_1	H_0 **ablehnen, falls:**
$p = p_0$	$p \neq p_0$	$T \in \{0, \ldots, \boldsymbol{c_o}\} \cup \{\boldsymbol{c_u} + 1, \ldots, n\}$
$p \geq p_0$	$p < p_0$	$T \in \{0, \ldots, \overline{\boldsymbol{c}}\}$
$p \leq p_0$	$p > p_0$	$T \in \{\underline{\boldsymbol{c}} + 1, \ldots, n\}$

$$T = \text{Anzahl Erfolge bei } n \text{ Versuchen}, \quad T \sim \mathrm{B}(n, p_0) \text{ unter } H_0.$$

$\boldsymbol{c_o}$ **:**	$F_T(\boldsymbol{c_o}) \leq \frac{\alpha}{2} < F_T(\boldsymbol{c_o} + 1)$	Die größte Zahl, bei der die Verteilungsfkt. der $B\,(n, p_0)$ höchstens $\frac{\alpha}{2}$ ist.
$\boldsymbol{c_u}$ **:**	$F_T(\boldsymbol{c_u} - 1) < 1 - \frac{\alpha}{2} \leq F_T(\boldsymbol{c_u})$	Die kleinste Zahl, bei der die Verteilungsfkt. der $B\,(n, p_0)$ mindestens $1 - \frac{\alpha}{2}$ ist.
$\overline{\boldsymbol{c}}$ **:**	$F_T(\overline{\boldsymbol{c}}) \leq \alpha < F_T(\overline{\boldsymbol{c}} + 1)$	Die größte Zahl, bei der die Verteilungsfkt. der $B\,(n, p_0)$ höchstens α ist.
$\underline{\boldsymbol{c}}$ **:**	$F_T(\underline{\boldsymbol{c}} - 1) < 1 - \alpha \leq F_T(\underline{\boldsymbol{c}})$	Die kleinste Zahl, bei der die Verteilungsfkt. der $B\,(n, p_0)$ mindestens $1 - \alpha$ ist.

$\boldsymbol{\chi^2}$**-Unabhängigkeitstest**

H_0	H_1	H_0 **ablehnen, falls:**
X, Y unabh.	X, Y abh.	$T \in \left(\chi^2_{(k-1)(l-1);\,1-\alpha}\ ,\ \infty\right)$

$$T = n \cdot \left(\left(\sum_{i=1}^{k}\sum_{j=1}^{l} \frac{n_{ij}^2}{n_{i\bullet}n_{\bullet j}}\right) - 1\right), \quad T \sim \chi^2_{(k-1)(l-1)} \text{ unter } H_0.$$

$\boldsymbol{\chi^2}$**-Anpassungstest**

H_0	H_1	H_0 **ablehnen, falls:**
$X \sim V$	$X \nsim V$	$T \in \left(\chi^2_{k-l-1;\,1-\alpha}\ ,\ \infty\right)$

$$T = \left(\sum_{i=1}^{k} \frac{n_i^2}{np_i}\right) - n \quad (np_i \geq 5), \quad T \sim \chi^2_{k-l-1} \text{ unter } H_0.$$

V = zu prüfende Verteilung

k = Anzahl der Klassen

l = Anzahl unbekannter Parameter