Zusammenfassung
Stochastik I + II

Stephan Kuschel
Vorlesung von Dr. Nagel

Stochastik I: WS 2007/08 Stochastik II: SS 2008
zuletzt aktualisiert: 7. Juli 2009

Da diese Zusammenfassung den Menschen, die sie lesen helfen soll bitte ich darum, Fehler und andere Verbesserungsideen an mich weiterzuleiten: Vorname.Nachname@uni-jena.de (entsprechend Deckblatt ersetzen)

Inhaltsverzeichnis

Teil I
Stochastik I

Inhaltsverzeichnis

1 Wahrscheinlichkeitsraum

1.1 Wahrscheinlichkeitsraum

Wahrscheinlichkeitsraum $[\Omega, \mathfrak{A}, P]$, Ereignis $A \in \mathfrak{A}$, dann gilt:

- $\Omega \in \mathfrak{A}$
- $A \in \mathfrak{A} \Rightarrow A^{\complement} \in \mathfrak{A} \qquad \forall A \subseteq \Omega$ $\qquad \mathfrak{A} \subseteq p(\Omega)$ ist σ-Algebra.
- $\forall A_i \in \mathfrak{A} \Rightarrow \bigcup_{i=1}^{\infty} A_i \in \mathfrak{A}$

Axiomensystem von Kolmogorov: $P : \mathfrak{A} \to [0,1]$

- $P(\Omega) = 1$
- $\boxed{\begin{aligned} A_i \cap A_j &= \emptyset \\ P\left(\bigcup_{i=1}^{\infty} A_i\right) &= \sum_{i=1}^{\infty} P(A_i) \end{aligned}}$ (σ-Additivität von P)

Folgerungen:

- $P(\emptyset) = 0$ $P(\Omega) = 1$
- $P(A^{\complement}) = 1 - P(A)$
- $P(A \cup B) = P(A) + P(B) - P(A \cap B)$
- $A \subseteq B \Rightarrow P(A) \leq P(B)$
  $\Rightarrow P(B \setminus A) = P(B) - P(A)$
- $A_1 \subseteq A_2 \subseteq ... \Rightarrow P\left(\bigcup_i A_i\right) = \lim_{i\to\infty} P(A_i)$
- $A_1 \supseteq A_2 \supseteq ... \Rightarrow P\left(\bigcap_i A_i\right) = \lim_{i\to\infty} P(A_i)$

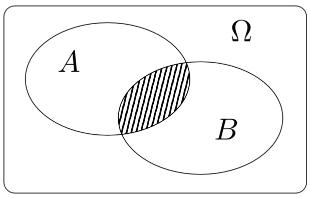
1.2 Beschreibungsmöglichkeiten für Wahrscheinlichkeitsmaße

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

1.3 Spezielle Wahrscheinlichkeitsräume

siehe Verteilungen

1.4 Bedingte Wahrscheinlichkeiten



$$\boxed{P(A|B) = \frac{P(A \cap B)}{P(B)}}$$

- A_1, A_2 disjunkt $\Leftrightarrow P(A_1 \cup A_2|B) = P(A_1|B) + P(A_2|B)$
- $P(A \cap B) = P(A|B) \cdot P(B)$
  $= P(B|A) \cdot P(A)$

$$\Rightarrow P(B|A) = P(A|B) \cdot \frac{P(B)}{P(A)}$$

- Entnahme ohne Zurücklegen:

$$P(A_1 \cap ... \cap A_n) = P(A_1) \cdot P(A_2|A_1) \cdot P(A_3|A_2 \cap A_1) \cdot ...$$

1.5 stochastische Unabhängigkeit

A, B stochastisch unabhängig $\Leftrightarrow P(A \cap B) = P(A) \cdot P(B)$

A_i stochastisch vollständig unabhängig $\Leftrightarrow P\left(\bigcap_i A_i\right) = \prod_i P(A_i)$

- paarweise stochastische Unabhängigkeit ist etwas anderes!
- (A, B) unabhängig $\Rightarrow (A, B^{\complement}), (A^{\complement}, B), (A^{\complement}, B^{\complement})$ unabhängig

2 Zufällige Variablen, Zufallsgrößen, zufällige Vektoren

2.1 Zufällige Variablen

$g : \Omega \to \Omega'$
$g^{-1} : p(\Omega') \to p(\Omega)$
$g^{-1}(A') : \{\omega \in \Omega : g(\omega) \in A'\}$ mit $A' \subseteq \Omega'$

2.2 Zufallsgrößen

$X : \Omega \to \mathbb{R}$
$P_X(B) = P(X^{-1}(B)) \qquad B \in \mathbb{R}$

Verteilungsfunktion: $F_X(x) = P(X \leq x) \qquad x \in \mathbb{R}$

- $P(a < X < b) = F_X(b) - F_X(a)$
- $P(X = a) = F_X(a) - F_X(a - 0)$
- $F_X(-\infty) = 0$
- $F_X(\infty) = 1$

2.3 Unabhängigkeit von Zufallsgrößen

X, Y sollen unabh. heißen, wenn all Paare von Ereignissen, die Mithilfe von X, Y formuliert werden können unabhängig sind.

- $P(X \in B_1, Y \in B_2) = P(X \in B_1) \cdot P(Y \in B_2) \iff X, Y$ unabh.
- $P(X_1 \in B_1, X_2 \in B_2, \dots X_n \in B_n) = P(X_1 \in B_1) \cdot \ldots \cdot P(X_n \in B_n)$
  $\implies X_1, \dots, X_n$ vollständig unabhängig
  $\implies X_1, \dots, X_n$ i.i.d., falls $X_1, \dots, X_n$ Zufallsgrößen über demselben W.-Raum

2.3.1 diskrete Zufallsgrößen

X, Y unabhängig $\iff P(X = x, Y = y) = P(X = x) \cdot P(Y = y)$

$X \sim$ geometrisch verteilt: („Gedächtnislosigkeit“)
$P(X = k + l | X \geq k) = P(X = l)$

2.3.2 stetige Zufallsgrößen

$X = (X_1, \dots, X_n)$

- Randverteilungsfunktion: $F_{X_i}(x) = P(X_1 \in \mathbb{R}, \dots, x_i \leq x, \dots, X_n \in \mathbb{R})$
- gemeinsame Verteilungsfunktion: $F_X(x_1, \dots, x_n) = P(X_1 \leq x_1, \dots, X_n \leq x_n)$

$\boxed{X_1, \dots, X_n \text{ unabhängig} \iff F_X(x_1, \dots, x_n) = F_{X_1}(x_1) \cdot \ldots \cdot F_{X_n}(x_n)}$

Dichtefunktion $F_X(x_1, \dots, x_n) = \displaystyle\int_{-\infty}^{x_n} \dots \int_{-\infty}^{x_1} f_X(t_1, \dots, t_n) dt_1 \dots dt_n$

mit $\displaystyle\int_{\mathbb{R}^n} f_X(t) dt = 1 \iff$ f heißt Dichtefunktion

Randdichte $f_{x_i} = \displaystyle\int_{\mathbb{R} \setminus \text{span}(x_i)} f_X(t)$

3 Verteilungsgesetze von transformierten Zufallsgrößen

3.1 Transformationen von 1dim. Zufallsgrößen

$F_{g(x)}(x) = P(g(X) \leq x) = P(X \in g^{-1}((-\infty, x]))$
oder Darstellung $F_{g(x)}(x) = \int_{-\infty}^{x} k(t)dt \Rightarrow g(x)$ hat VD k

3.2 Summe zweier Zufallsgrößen

Seien X_1, X_2 unabhängig
$P(X_1 + X_2 = s) = \sum_{x_1} P(X_1 = x_1, X_2 = s - x_2) = \sum_{x_1} P(X_1 = x_1) \cdot P(X_2 = s - x_1)$
$\Rightarrow X_1 \sim \Pi_{\lambda_1}, \quad X_2 \sim \Pi_{\lambda_2}, \quad X_1 + X_2 \sim \Pi_{\lambda_1+\lambda_2}$
$f_{X_1+X_2} = \int_{-\infty}^{\infty} f_{X_1}(t) f_{X_2}(s-t)dt$
$\Rightarrow X_1 \sim N_{\mu_1,\sigma_1^2}, \quad X_2 \sim N_{\mu_2,\sigma_2^2}, \quad X_1 + X_2 \sim N_{\mu_1+\mu_2,\sigma_1^2+\sigma_2^2}$
$f_{X_1-X_2} = \int_{-\infty}^{\infty} f_{X_1}(t) f_{X_2}(t-s)dt$

3.3 Produkt & Quotient zweier Zufallsgrößen

$f_{X_1 \cdot X_2}(s) = \int_{-\infty}^{\infty} \frac{1}{|t|} f_{X_1}\left(\frac{s}{t}\right) f_{X_2}(t)dt$
$f_{\frac{X_1}{X_2}}(s) = \int_{-\infty}^{\infty} |t| f_{X_1}(st) f_{X_2}(t)dt$

3.4 Injektive diffbare Transformationen von zufälligen Vektoren

$$f_Y(u) = \frac{f_X(T^{-1}(u))}{|\det T'(T^{-1}(u))|}$$

4 Erwartungswert, Varianz, Kovarianz

- Erwartungswert $\hat{=}$ Mittelwert
- Varianz $\hat{=}$ mittlere quadratische Abweichung

4.1 Erwartungswert

$\mathbb{E}X = \int_{-\infty}^{\infty} x f_X(x)dx, \qquad \mathbb{E}g(X) = \int_{-\infty}^{\infty} g(x) f_X(x)dx$
Existenz: $\int_{-\infty}^{\infty} |x| f_X(x)dx < \infty$

- $\mathbb{E}(aX_1 + b) = a\mathbb{E}X + b$
- $\mathbb{E}(X_1 + X_2) = \mathbb{E}X_1 + \mathbb{E}X_2$ (gilt immer)
- $\mathbb{E}(X_1 \cdot X_2) = \mathbb{E}X_1 \cdot \mathbb{E}X_2$ (Zgr. unabhängig!)

4.2 Varianz

$\text{var}X = \mathbb{E}(X - \mathbb{E}X)^2 = \mathbb{E}X^2 - (\mathbb{E}X)^2$

- $\text{var}(aX + b) = a^2\text{var}X$
- $\text{var}(X_1 \pm X_2) = \text{var}X_1 + \text{var}X_2, \quad X_1, X_2$ unabhängig

4.3 Kovarianz

- $\text{cov}(X_1, X_2) = \mathbb{E}(X_1 \cdot X_2) - (\mathbb{E}X_1) \cdot (\mathbb{E}X_2)$
- $\text{var}(X_1 + X_2) = \text{var}X_1 + \text{var}X_2 + 2\text{cov}(X_1 X_2)$
- $\text{cov}(X_1, X_2) = 0 \Leftrightarrow X_1, X_2$ unkorreliert
- $\text{cov}(X, aX + b) = a\text{var}X$
- $\text{cov}(X, X) = \text{var}X$
- $X_i \sim N_{\mu_i \sigma_i^2}: \quad \text{cov}(X_1, X_2) = \rho\sigma_1\sigma_2$
  $\rho_{X_1,X_2} = \dfrac{\text{cov}(X_1, X_2)}{\sqrt{\text{var}X_1 \cdot \text{var}X_2}}$

4.4 Die Kovarianzmatrix

$X = (X_1, \ldots, X_n)$ zufälliger Vektor
$\Sigma_X = \mathbb{E}(X - \mathbb{E}X)^T(X - \mathbb{E}X) = (\text{cov}(X_i, X_j))_{ij}$
n-dim Normalverteilung:
$N_{\mu,\Sigma}, n = 2: \quad \Sigma = \begin{pmatrix} \sigma_1^2 & \rho\sigma_1\sigma_2 \\ \rho\sigma_1\sigma_2 & \sigma_2^2 \end{pmatrix}$

5 Ungleichungen & Grenzwertsätze

5.1 Markov-Ungleichung

$$P(|X| \geq c) \leq \frac{\mathbb{E}(g(|X|))}{g(c)} \qquad \text{g monoton, nicht fallend}$$

5.2 Tschebyscheff Ungleichung

$$P(|X - \mathbb{E}X| \geq c) \leq \frac{\text{var}X}{c^2}$$

- $X \sim N_{\mu,\sigma^2}: \quad P(|X - \mu| \geq k\sigma) \leq \frac{\sigma^2}{k^2\sigma^2} = \frac{1}{k^2}$

5.3 Gesetz der großen Zahlen

$(X_i)_i$ i.i.d.

$$\lim_{n\to\infty} P\left(\frac{1}{n}\sum_{i=1}^{n} |X_i - \mathbb{E}X_1| > \epsilon\right) = 0$$

5.4 Der zentrale Grenzwertsatz

$(X_i)_i$ i.i.d. mit $\mathbb{E}X_i^2 < \infty, \quad \mathrm{var}X_i = \sigma^2 > 0, \quad \mathbb{E}X_i = m$

$$\lim_{n\to\infty} P\left(\frac{\sum_{i=1}^n X_i - n\mu}{\sqrt{n}\sigma} \leq x\right) = \Phi(x)$$

mit $\Phi(x)$ VF von $N_{0,1}$

- Summen von i.i.d.Zgr. sind asymptotisch normalverteilt.
- $P(X_i = 1) = p = 1 - P(X_i = 0)$
  $\lim\limits_{n\to\infty} P\left(\frac{\sum_{i=1}^n X_i - np}{\sqrt{np(1-p)}} \leq x\right) = \Phi(x)$
  mit Korrekturformel: $P(\sum_i X_i \leq k) = P(\sum_i X_i < k) = \Phi\left(\frac{k+\frac{1}{2}-np}{\sqrt{np(1-p)}}\right)$

Approx der Binomialverteilung:

- Poissonverteilung: $\lambda = np$, n groß, p klein
- Normalverteilung: $\mu = np, \ \sigma^2 = np(1-p)$

Teil II
Stochastik II: mathematische Statistik

Inhaltsverzeichnis

1 Stichproben und der statistische Raum

Stichproben

- math. Stichprobe: $X_1, \ldots, X_n$
- konkrete Stichprobe: $x_1, \ldots, x_n$

Der statistische Raum: $[\mathbb{R}^n, \mathcal{R}_n, \{P_\theta^{\otimes n}, \theta \in \Theta\}]$
$\Theta \subseteq \mathbb{R}^l, \quad l \geq 1, \quad P_\theta$ ist Wahrscheinlichkeitsmaß auf $[\mathbb{R}, \mathcal{R}]$ $\forall \theta$

1.1 Liste wichtiger Statistiken

1. Stichprobenmittel / empirischer Erwartungswert

$$T(x_1, \ldots, x_n) = \frac{1}{n} \sum_{i=1}^n x_i = \bar{x}$$

2. r-tes Stichprobenmoment / empirisches r-tes Moment

$$T(x_1, \ldots, x_n) = \frac{1}{n} \sum_{i=1}^n x_i^r$$

3. Stichprobenstreuung / empirische Varianz

$$T(x_1, \ldots, x_n) = \frac{1}{n} \sum_{i=1}^n (x_i - \bar{x})^2 = \bar{\sigma}^2$$

4. korrigierte Stichprobenstreuung / korrigierte empirische Varianz

$$T(x_1, \ldots, x_n) = \frac{1}{n-1} \sum_{i=1}^n (x_i - \bar{x})^2 = \hat{\sigma}^2 = \frac{n}{n-1} \bar{\sigma}^2$$

5. konkrete geordnete Stichprobe / Variationsreihe

$$T(x_1, \ldots, x_n) = (x_1^*, \ldots, x_n^*) \qquad \text{(sortieren)}$$
$$\text{mit } x_1^* \leq \ldots \leq x_n^*$$

6. i-te geordnete Statistik / i-te Rangstatistik

$$T(x_1, \ldots, x_n) = x_i^* \qquad \text{(nach Sortieren, i-tes Element)}$$

7. Spannweite einer Stichprobe

$$T(x_1, \ldots, x_n) = x_n^* - x_1^*$$

8. Stichprobenmedian / empirischer Zentralwert

$$T(x_1, \ldots, x_n) = \tilde{x}_{\frac{1}{2}} = \begin{cases} x^*_{\frac{n+1}{2}} & \text{falls } n \text{ ungerade} \\ \frac{1}{2}\left(x^*_{\frac{n}{2}} + x^*_{\frac{n}{2}+1}\right) & \text{falls } n \text{ gerade} \end{cases}$$

9. Stichproben-α-Quantil / empirisches α-Quantil, $\alpha \in (0,1)$

$$T(x_1,\dots,x_n) = \tilde{x}_\alpha = \begin{cases} x^*_{\lfloor n\alpha \rfloor + 1} & \text{falls } n\alpha \text{ nicht ganzzahlig} \\ \frac{1}{2}\left(x^*_{n\alpha} + x^*_{n\alpha+1}\right) & \text{falls } n\alpha \text{ ganzzahlig} \end{cases}$$

10. empirische Verteilungsfunktion

$$\begin{aligned} T(x_1,\dots,x_n) = \hat{F}(s) &= \frac{1}{n}\sum_{i=1}^n 1_{[x_i,\alpha)}(s) \\ &= \frac{1}{n}\left|\{i \in 1,\dots,n : x_i \le s\}\right| \end{aligned}$$

11. Histogramm oder rel. Häufigkeiten zu einer Vorgegebenen Klasseneinteilung $\Delta_1,\dots,\Delta_k$

$$T(x_1,\dots,x_n) = \left(\frac{1}{n}\sum_{i=1}^n 1_{\Delta_1}(x_i),\dots,\frac{1}{n}\sum_{i=1}^n 1_{\Delta_k}(x_i)\right)$$

- $\Delta_1,\dots\Delta_k$ paarweise Disjunkt, äquidistant
- Faustregel von STURGES: $k = 1 + 3.32\log_{10} n$

12. Box-Plot

$$T(x_1,\dots,x_n) = (\tilde{x}_{0.1}, \tilde{x}_{0.25}, \tilde{x}_{0.5}, \tilde{x}_{0.75}, \tilde{x}_{0.9})$$

2 Punktschätzungen

2.1 Punktschätzungen für Erwartungswert

$\hat{\mu}(x_1,\dots,x_n) = \bar{x}$
$\mathrm{var}_\theta \hat{\mu}(X_1,\dots,X_n) = \frac{1}{n}\mathrm{var}_\theta X_1$ falls $X_1,\dots,X_n$ i.i.d.

2.2 Punktschätzung für die Varianz

$$\hat{\sigma}^2(x_1,\dots,x_n) = \frac{1}{n-1}\sum_{i=1}^n (x_i - \bar{x})^2$$

2.3 Gütekriterien

- T erwartungstreue Punktschätzung für γ
  $\Leftrightarrow \mathbb{E}_\theta T(X_1,\dots,X_n) = \gamma(\theta) \quad \forall \theta \in \Theta$
- T_1 effizienter als T_2
  $\Leftrightarrow \mathrm{var}_\theta T_1(X_1,\dots,X_n) < \mathrm{var}_\theta T_2(X_1,\dots,X_n) < \infty$
- T^* bester erwartungstreuer Schätzer (BUE - best unbiased estimator)
  $\Leftrightarrow \mathrm{var}_\theta T^*(X_1,\dots,X_n) \le \mathrm{var}_\theta T_i(X_1,\dots,X_n)\ \forall \theta \in \Theta$ und T_i erwartungstreu $\forall i$

Mögliche andere Kriterien

- $\mathbb{E}_\theta(T(X_1,\dots,X_n) - \theta)^2$
- Asymptot. Verhalten für $n \to \infty$

Anmerkung: $\hat{\mu} = \bar{x}$ ist bester erwartungstreuer Schätzer, falls Grundgesamtheit normalverteilt, poissonverteilt, binomialverteilt, aber nicht bei Gleichverteilung auf (a,b), a,b unbekannt.

2.4 Die Maximum-Likelihood-Methode

$\theta \in \Theta$ suchen für das diese Stichprobe den Maximalwert der Wahrscheinlichkeitsdichte liefert.

- Likelihood-Funktion:
  $L: \ \theta \times \mathbb{R}^n \to [0, \infty)$
  $$\begin{aligned} L(\theta, x_1, \ldots, x_n) &= \prod_{i=1}^{n} f_\theta(x_i) && \text{(stetig)} \\ &= \prod_{i=1}^{n} P_\theta(X_i = x_i) && \text{(diskret)} \end{aligned}$$
- Maximum-Likelihood Schätzwert $\hat{\theta}^*$
  $L(\hat{\theta}^*, x_1, \ldots, x_n) \geq L(\theta, x_1, \ldots, x_n) \quad \forall \theta \in \Theta$
- Zur Berechnung häufig $\ln L$ betrachten.

3 Verteilungen

$X_i \sim N_{0,1} \qquad Z_i \sim N_{\mu,\sigma^2}$

3.1 χ^2-Verteilung

$$f_Y(x) = \begin{cases} 0 & x < 0 \\ \frac{1}{2^{\frac{r}{2}}\Gamma\left(\frac{r}{2}\right)} x^{\frac{r}{2}-1} e^{-\frac{x^2}{2}} & x \geq 0 \end{cases}$$

- $\sum_{i=1}^{r} X_i^2 \sim \chi_r^2$
- $\frac{1}{\sigma^2} \sum_{i=1} n(Z_i - \mu)^2 \sim \chi_n^2$
- $\frac{1}{\sigma^2} \sum_{i=1} n(Z_i - \bar{Z})^2 \sim \chi_{n-1}^2$
- $Y_1 \sim \chi_{r_1}^2 \qquad Y_2 \sim \chi_{r_2}^2 \qquad Y_1 + Y_2 \sim \chi_{r_1+r_2}^2$
- $\chi_2^2 = \varepsilon_{\frac{1}{2}}$

3.2 t-Verteilung

$$f_Y(x) = \begin{cases} \frac{\Gamma\left(\frac{r+1}{2}\right)}{\Gamma\left(\frac{r}{2}\right)\sqrt{\pi r}} \left(1 + \frac{x^2}{r}\right)^{-\frac{r+1}{2}} & x \in \mathbb{R} \end{cases}$$

- $f_Y(-x) = f_Y(x)$
- $r = 1$: Cauchy-Verteilung
- $r \to \infty$: Normalverteilung mit $\mu = 0, \sigma^2 = 1$
- $\frac{X_0}{\sqrt{\frac{1}{r}\sum_{i=1}^{r} X_i^2}} \sim t_r$
- $\sqrt{n}\frac{\bar{Z} - \mu}{\sqrt{\frac{1}{n-1}\sum_{i=1}^{n}(X_i - \bar{X})^2}} \sim t_{n-1}$

3.3 F-Verteilung

$$f_Y(x) = \begin{cases} 0 & x < 0 \\ (\frac{s}{r})^{\frac{s}{2}}\, x^{\frac{s}{2}-1}\, (1+\frac{s}{r}x)^{-\frac{s+r}{2}} & x \geq 0 \end{cases}$$

- $\dfrac{\frac{1}{s}\sum_{i=r+1}^{r+s} X_i^2}{\frac{1}{r}\sum_{i=1}^{r} X_i^2} \sim F_{s,r}$

4 Konfidenzintervalle

$(1-\alpha)$-Konfidenzintervall $\quad P_\theta^{\otimes n}(C(X_1,\dots,X_n) \ni \gamma(\theta)) \geq 1-\alpha \quad \forall \theta \in \Theta$
Häufig $(1-\alpha) \in \{0.9; 0.95; 0.99\}$

4.1 Konfidenzintervall für Erwartungswert bei bekannter Varianz

$$C(x_1,\dots,x_n) = \left[\bar{x} - \frac{\sigma}{\sqrt{n}}\Phi^{-1}(1-\frac{\alpha}{2}),\ \bar{x} + \frac{\sigma}{\sqrt{n}}\Phi^{-1}(1-\frac{\alpha}{2})\right]$$

ist $(1-\alpha)$-KI

4.2 Konfidenzintervall für Erwartungswert bei unbekannter Varianz

$$C(x_1,\dots,x_n) = \left[\bar{x} - \frac{\sqrt{\hat{\sigma}^2}}{\sqrt{n}}t_{n-1,1-\frac{\alpha}{2}},\ \bar{x} + \frac{\sqrt{\hat{\sigma}^2}}{\sqrt{n}}t_{n-1,1-\frac{\alpha}{2}}\right]$$

ist $(1-\alpha)$-KI

4.3 Konfidenzintervall für Varianz bei unbekanntem Erwartungswert

$$C(x_1,\dots,x_n) = \left[\frac{(n-1)\hat{\sigma}^2}{\chi^2_{n-1,1-\frac{\alpha}{2}}},\ \frac{(n-1)\hat{\sigma}^2}{\chi^2_{n-1,\frac{\alpha}{2}}}\right]$$

weil $\dfrac{(n-1)\hat{\sigma}^2}{\sigma^2} \sim \chi^2_{n-1}$

5 Tests

5.1 Grundbegriffe

Θ partitionieren in Θ_0, Θ_1
$H_0:\ \theta \in \Theta_0$ Nullhypothese
$H_1:\ \theta \in \Theta_1$ Alternativhypothese

	H_0 wahr	H_1 wahr
H_0 wählen	✓	Fehler 2. Art
H_0 ablehnen	Fehler 1. Art	✓

2.Art: irrtümliche Annahme von H_1
1.Art: irrtümliche Ablehnung von H_0

- Test: $D : \mathbb{R}^n \to \{H_0, H_1\}$
- Testgröße: $T : \mathbb{R}^n \to \mathbb{R}$

- Kritischer Bereich $K \subseteq \mathbb{R}$ sodass
  $T(x_1, \dots, x_n) \in K \Rightarrow H_1$ wählen
  $T(x_1, \dots, x_n) \notin K \Rightarrow H_0$ wählen
- Gütefunktion des Tests D
  $\beta_D : \theta \to [0, 1] \quad$ mit
  $\beta(\theta) = P_\theta^{\otimes n}(D(X_1, \dots, X_n) = H_1)$
- Signifikanzniveau zum Niveau α
  $\beta_D(\theta) \leq \alpha \quad \forall \theta \in \Theta$
  $\beta_D(\theta) \geq \sup\limits_{\theta' \in \Theta_0} \beta_D(\theta') \quad \forall \theta \in \Theta_1 \qquad$ (Unverfälschtheit)

5.2 Tests für Normalverteilte Grundgesamtheit

5.2.1 Gausstest: Prüfung des Erwartungswertes bei bekannter Varianz

(a) $H_0 : \mu = \mu_0 \qquad H_1 : \mu \neq \mu_0$

$$T(X_1, \dots, X_n) = \sqrt{n}\frac{\bar{X} - \mu_0}{\sigma_0} \underset{\substack{\uparrow \\ \text{falls } \mu=\mu_0 \text{ (!)}}}{\sim} N_{0,1}$$

$$K = \left(-\infty, \Phi^{-1}\left(\frac{\alpha}{2}\right)\right) \cup \left(\Phi^{-1}\left(1 - \frac{\alpha}{2}\right), \infty\right)$$

(b) $H_0 : \mu \geq \mu_0 \qquad H_1 : \mu < \mu_0$
Testgröße wie bei Punkt a

$$K = \left(-\infty, -\Phi^{-1}(1 - \alpha)\right)$$

5.2.2 t-Test: Prüfung des Erwartungswertes bei unbekannter Varianz

(a) $H_0 : \mu = \mu_0 \qquad H_1 : \mu \neq \mu_0$

$$T(X_1, \dots, X_n) = \frac{\sqrt{n}}{\sqrt{\hat{\sigma}^2}}(\bar{X} - \mu_0) \underset{\substack{\uparrow \\ \text{falls } \mu=\mu_0 \text{ (!)}}}{\sim} t_{n-1}$$

$$K = (-\infty, -t_{n-1,1-\frac{\alpha}{2}}) \cup (t_{n-1,1-\frac{\alpha}{2}}, \infty)$$

K ist ein wenig kleiner als beim entsprechenden Gauss-Test

- α kann auch als Überschreitungswahrscheinlichkeit gelesen werden, sodass $P_Ü(X_1, \dots, X_n) < \alpha$

(b) $H_0 : \mu \geq \mu_0 \qquad H_1 : \mu < \mu_0$
Testgröße wie bei Punkt a

$$K = (-\infty, -t_{n-1,1-\alpha})$$

5.2.3 χ^2-Test: Prüfung der Varianz bei unbekanntem Erwartungswert und normalverteilter Grundgesamtheit

(a) $H_0 : \sigma^2 = \sigma_0^2 \qquad H_1 : \sigma^2 \neq \sigma_0^2$

$$\begin{aligned} T(X_1, \ldots, X_n) &= \frac{\hat{\sigma}^2(n-1)}{\sigma_0^2} \\ &= \frac{1}{\sigma_0^2} \sum_{i=1}^{n} (X_i - \bar{X})^2 \underset{\text{falls } \sigma^2 = \sigma_0^2 \text{ (!)}}{\overset{\sim}{\uparrow}} \chi^2_{n-1} \\ K &= [0, \chi^2_{n-1,\frac{\alpha}{2}}) \cup (\chi^2_{n-1,1-\frac{\alpha}{2}}, \infty) \end{aligned}$$

(b) $H_0 : \sigma^2 \geq \sigma_0^2 \qquad H_1 : \sigma < \sigma_0^2$
Testgröße wie bei Punkt a

$$T(X_1, \ldots, X_n) \notin K \iff \frac{\alpha}{2} \leq p_{\ddot{u}} \leq 1 - \frac{\alpha}{2}$$
$$K = (0, \chi^2_{n-1,\alpha})$$

5.3 Zwei Testprobleme für disjunkte Verteilungen

5.3.1 Likelihood-Quotienten Methode

Allgemeines Prinzip zur Konstruktion von Tests bzw von Testgrößen.
Betrachten $[\mathbb{R}^n, \mathcal{R}_n\{P_{\theta_0}, P_{\theta_1}\}]$ d.h. $\theta = \{P_{\theta_0}, P_{\theta_1}\}$
$H_0 : \theta = \theta_0 \qquad H_1 : \theta = \theta_1$

$$\frac{L(\theta_1, x_1, \ldots, x_n)}{L(\theta_0, x_1, \ldots, x_n)} > c \qquad \text{dann } H_0 \text{ ablehnen}$$

c so wählen, dass Wahrscheinlichkeit für Fehler 1.Art $\beta(\theta_0) \leq \alpha_0$
Man kann zeigen, dass dieser Test, falls $\beta(\theta_0) = \alpha$ bester α-Signifikanztest ist, in dem Sinne, dass $\beta(\theta_1) \geq \beta_D(\theta_1) \quad \forall \alpha$-Signifikanztests D.

5.3.2 Testen von Hypothesen über Parameter der hypergeometrischen Verteilung

N - Gesamtzahl der Produkte in einer Lieferung
θ - Anzahl der fehlerhaften Produkte (unbekannt), $\theta \in \{0, 1, 2, \ldots, N\} = \Theta$
m - Anzahl geprüfter Produkte
Zufallsgröße: κ - Anzahl der fehlerhaften Produkte unter den geprüften Produkten. κ besitzt eine hypergeometrische Verteilung mit Parametern $N, \theta, m = \mathscr{H}_{N,\theta,m}$

$$P_\theta(\kappa = k) = \frac{\binom{\theta}{k}\binom{N-\theta}{m-k}}{\binom{N}{m}}$$

$H_0 : \theta \leq \theta_0 \qquad H_1 : \theta \geq \theta_1$
Kritischen Bereich aus Quantilen wählen.

- $N >> m \Rightarrow \mathscr{H}_{N,\theta,m} \approx B_{m,\frac{\theta}{N}}$
- Diese Binomialverteilung gegebenenfalls mithilfe des zentralen Grenzwertsatzes durch Normalverteilung oder durch Poissonverteilung mit $\lambda = \frac{m \cdot \theta}{N}$ nähern.

5.4 Anpassungstests (Kolmogorov-Smirnov)

$H_0 : F = F_0 \qquad H_1 : F \neq F_0$

$$T(x_1, \dots, x_n) = \sqrt{n} \sup_{t \in \mathbb{R}} \left| \hat{F}(t) - F_0(t) \right|$$

5.5 Zwei-Stichproben Test

$H_0 : \mu_1 \leq \mu_2 \qquad H_1 : \mu_1 > \mu_2$

$$T(X_1, \dots, X_n, Y_1, \dots, Y_n) = \sqrt{\frac{n \cdot m}{n+m}} \frac{\bar{X} - \bar{Y}}{\sqrt{\hat{\hat{\sigma}}^2}} \underset{\substack{\text{falls } \mu_1 = \mu_2 \\ \text{und } \mathrm{var} X_1 = \mathrm{var} X_2 = \sigma^2}}{\overset{\sim}{\uparrow}} t_{n+m-2}$$

$$\hat{\hat{\sigma}}^2 = \frac{1}{n+m-2} \left((n-1)\hat{\sigma}_X^2 + (m-1)\hat{\sigma}_Y^2 \right)$$

$$K = (t_{m+n-2, 1-\alpha}, \infty)$$

- Vergleich der Varianzen: F-Test
- $\sigma_1^2 \neq \sigma_2^2$: Welch Test

6 Stat. Methoden für 2-dim Stichproben (Multivariatstatistik)

6.1 Test auf Unabhängigkeit von Beobachtungsparametern

6.1.1 Randverteilungen

$F_{(X,Y)}(x, y) = F_X(x) \cdot F_Y(y) \quad \Longleftrightarrow \quad$ Unabhängigkeit von X und Y
mit $F_{(X,Y)}(x, y) = P(X \leq x, Y \leq y)$

6.1.2 Korrelationskoeffizient

$$\rho_{X,Y} = \frac{\mathrm{cov}(X, Y)}{\sqrt{\mathrm{var}X \cdot \mathrm{var}Y}} \quad \frac{\text{nur falls } X, Y \text{ normalverteilt} \Rightarrow}{\Leftarrow \text{gilt immer}} \quad X, Y \text{ unabhängig}$$

6.1.3 χ^2 Unabhängigkeitstest

Es seien $a_1, \dots, a_r \in \mathbb{R}$ und $b_1, \dots, b_s \in \mathbb{R}$

$$p_{ij} = P(X_1 = a_i, Y_1 = b_j) \qquad \text{mit } \sum_{i,j} p_{ij} = 1$$

Bezeichnung:

$$p_{i\cdot} = \sum_{j=1}^{s} p_{ij} = P(X_1 = a_i) \qquad p_{\cdot j} = \sum_{i=1}^{r} p_{ij} = P(Y_1 = b_j)$$

$H_0 : p_{ij} = p_{i\cdot} \cdot p_{\cdot j} \quad \forall_{i,j}$
$H_1 : p_{ij} \neq p_{i\cdot} \cdot p_{\cdot j} \quad$ für wenigstens ein Paar (i, j)

Zufallsgröße: H_{ij} - Anzahl $\{l : X_l = a_i, Y_l = b_j\} \mathrel{\hat{=}}$ Absolute Häufigkeit des Auftretens ddes Paares (a_i, b_j) in der Stichprobe.

$$H_{i\cdot} = \sum_{j=1}^{s} H_{ij} \qquad\qquad H_{\cdot j} = \sum_{i=1}^{r} H_{ij}$$

Testgröße:

$$T = n \cdot \sum_{j=1}^{s} \sum_{i=1}^{r} \frac{\left(H_{ij} - \frac{H_{i\cdot} \cdot H_{\cdot j}}{n}\right)^2}{H_{i\cdot} \cdot H_{\cdot j}} \underset{\substack{\text{asymptot. } n\to\infty \\ H_0 \text{ wahr}}}{\underset{\uparrow}{\sim}} \chi^2_{(r-1)(s-1)}$$

Kritischer Bereich:

$$K = \left[\chi^2_{(r-1)(s-1),1-\alpha}, \infty\right)$$

- Falls X_1 und Y_1 nicht diskret mithilfe von Klasseneinteilung diskretisieren
- ACHTUNG: Falls H_0 abgelehnt wird, dann Annahme dass notwendige Bedingung für die Unabhängigkeit verletzt ist!
  Falls H_0 angenommen wird, dann also keine Aussage über die Unabhängigkeitshypothese möglich

6.2 Regressionsanalyse

MKQ is BLUE
Gauss Markov Theorem
Zufällige Prozesse