Name : ______________________
Matr.-Nr.: ______________________
Hörsaal : ______________________

Formelsammlung zur Vorlesung Statistik II für Studierende der Soziologie Magisterstudierende im Sommersemester 2007

Prof. Dr. Thomas Augustin
Christiane Dargatz

Einleitung: Grundaufgaben der induktiven Statistik und Überblick über die Veranstaltung

a) Deskriptiv versus Induktiv

b) Typische Beispiele

c) Typische Fragestellungen der induktiven Statistik

d) Inferenzfehler

e) Zur Wahrscheinlichkeitsrechnung

1. Grundbegriffe der Wahrscheinlichkeitsrechnung und wichtige Verteilungsmodelle

1.0 Mengen und elementare Mengenoperationen

- Eine Menge ist eine Zusammenfassung verschiedener Objekte zu einem Ganzen.
- Rechenregeln für beliebige Mengen A, B und C

a) Kommutativgesetze	:	$A \cap B = B \cap A, \quad A \cup B = B \cup A.$
b) Assoziativgesetze	:	$(A \cap B) \cap C = A \cap (B \cap C).$
		$(A \cup B) \cup C = A \cup (B \cup C).$
c) Distributivgesetze	:	$(A \cup B) \cap C = (A \cap C) \cup (B \cap C).$
		$(A \cap B) \cup C = (A \cup C) \cap (B \cup C).$
d) De Morgansche Regeln	:	$\overline{(A \cup B)} = \overline{A} \cap \overline{B}.$
		$\overline{(A \cap B)} = \overline{A} \cup \overline{B}.$
e) Aus $A \subset B$ folgt	:	$\overline{B} \subset \overline{A}.$
f) Für die Differenzmenge $A \backslash B$ gilt	:	$A \backslash B = A \cap \overline{B}.$
g) Für die Potenzmenge $\mathcal{P}(A)$ gilt	:	$\lvert\mathcal{P}(A)\rvert = 2^{\lvert A\rvert}.$

1.1 Grundlegendes zum Begriff "Wahrscheinlichkeit"

a) Wahrscheinlichkeit in der Alltagssprache

b) Klassische Aspekte und Meilensteile

- Wahrscheinlichkeit im Glückspiel
- Wahr-schein-lichkeit
- Port-Royal-Logik
- Jacob Bernoulli (1654-1705)
- Wahrscheinlichkeiten als Ausdruck göttlicher Ordnung
- Laplacescher Wahrscheinlichkeitsbegriff (kombinator. Wahrscheinlichkeitsbegriff)

c) Aktuelle interpretatorische Hauptrichtungen

1) objektivistisch / frequentistische Richtungen (aleatorische Wahrscheinlichkeiten)
2) subjektivistische (epistemische) Richtungen
3) logischer Wahrscheinlichkeitsbegriff
4) mathematisch formaler Wahrscheinlichkeitsbegriff

1.2 Axiomensystem von Kolmogorov

- Festlegen eines *Ergebnisraums* (Grundraum, Stichprobenraum) Ω, der die Menge aller möglichen *Ergebnisse* ω enthält.
- Teilmengen A von Ω bezeichnet man als *Ereignisse*, einelementige als *Elementarereignisse*.

Def 1.1 *Axiome von Kolmogorov*
Eine Funktion P, die Ereignissen aus Ω (reelle) Zahlen zuordnet, heißt *Wahrscheinlichkeit*, wenn gilt

$$(K1) \quad P(A) \geq 0 \text{ für alle Ereignisse } A \ (\subset \Omega) \tag{1.2.1}$$
$$(K2) \quad P(\Omega) = 1 \tag{1.2.2}$$
$$(K3) \quad \text{Falls } A \cap B = \emptyset, \quad \text{dann ist } P(A \cup B) = P(A) + P(B). \tag{1.2.3}$$

Zentrale Rechenregeln

- $P(\overline{A}) = 1 - P(A), \qquad \text{insb. } P(\emptyset) = 0$ (1.2.4)
- $P(A \cup B) = P(A) + P(B) - P(A \cap B)$ für nicht notwendig disjunkte A, B. (1.2.5)
- Falls $A_1, A_2, \ldots, A_n$ paarweise disjunkt sind, also $A_i \cap A_j = \emptyset$ für $i \neq j$, dann gilt:
  $$P(\bigcup_{i=1}^{n} A_i) = P(A_1 \cup A_2 \cup \ldots \cup A_n) = P(A_1) + P(A_2) + \ldots + P(A_n) \tag{1.2.6}$$
- Ist $B_1, B_2, \ldots, B_n$ eine *vollständige Zerlegung* von Ω, d.h. gilt $\bigcup_{i=1}^{n} B_i = \Omega$ und $B_i \cap B_j = \emptyset$ für $i \neq j$, so gilt für jedes Ereignis A:

$$P(A) = \sum_{i=1}^{n} P(A \cap B_i) \qquad (1.2.7)$$

Bem 1.4 *Reine Zufallsauswahl,* $n = 1$
Gegeben sei eine endliche Grundgesamtheit $\widetilde{\Omega}$ und ein Merkmal $\widetilde{X}$ mit den Ausprägungen $a_1, \ldots, a_k$ und den relativen Häufigkeiten $f_1, \ldots, f_k$. Es werde eine reine Zufallsauswahl eines Elements ω durchgeführt und die diesbezügliche Ausprägung $\widetilde{X}(\omega)$ von $\widetilde{X}$ erhoben.
Diese Experiment lässt sich durch den Ergebnisraum $\Omega = \{a_1, a_2, \ldots, a_k\}$ und die Wahrscheinlichkeitsbewertung

$$P(\{a_j\}) = f_j, \quad j = 1, \ldots, k \qquad (1.2.8)$$

beschreiben.

1.3 Stochastische Unabhängigkeit, bedingte Wahrscheinlichkeiten und Koppelung

1.3.1 Stochastische Unabhängigkeit

Def 1.5 Zwei Ereignisse A und B heißen *(stochastisch) unabhängig beim Wahrscheinlichkeitsmaß* P, wenn gilt

$$P(A \cap B) = P(A) \cdot P(B) \qquad (1.3.1),$$

widrigenfalls heißen sie *stochastisch abhängig.*

1.3.2 Bedingte Wahrscheinlichkeiten

Def 1.8 Gegeben seien zwei Ereignisse A und B, wobei $P(B) > 0$ sei:
Dann heißt:

$$P(A|B) := \frac{P(A \cap B)}{P(B)} \qquad (1.3.2)$$

bedingte Wahrscheinlichkeit von A gegeben B oder *bedingte Wahrscheinlichkeit von A unter der Bedingung B.*

1.3.3 Koppelung von unabhängigen Experimenten, unabhängige Wiederholungen

Bem 1.10 *Unabhängige Koppelung*
Gegeben sei eine Reihe von Zufallsexperimenten, beschrieben durch die Ergebnisräume Ω_i und die Wahrscheinlichkeitsbewertungen $P_i, \quad i = 1, \ldots, n$. Sind die Experimente unabhängig, so setzt man für beliebige $A_i \subset \Omega_i, \quad i = 1, \ldots, n$:

$$P(A_1 \cap A_2 \cap \ldots \cap A_n) = P_1(A_1) \cdot P_2(A_2) \cdot \ldots \cdot P_n(A_n). \qquad (1.3.2a)$$

Dies beschreibt eine Wahrscheinlichkeit auf $\Omega := \Omega_1 \times \Omega_2 \times \ldots \times \Omega_n$.
Besteht keine Verwechselungsgefahr, so wird das Subskript $_i$ bei P_i weggelassen.

Bem und Bsp 1.11 *Reine Zufallsstichprobe vom Umfang n*
Betrachtet werde die Situation von Bem 1.4. Es werde eine *(geordnete) Zufallsstichprobe von Umfang n bezüglich des Merkmals* $\widetilde{X}$ entnommen, d.h. eine (geordnete) Zufallsauswahl mit

Zurücklegen von n Elementen $\omega_1, \ldots, \omega_n$ mit $\omega_i \in \widetilde{\Omega}, i = 1, \ldots, n$ entnommen und die diesbezüglichen Ausprägungen $\widetilde{X}(\omega_i)$ von $\widetilde{X}$ erhoben.
Sei für $j = 1, \ldots, k,\ i = 1, \ldots, n$, mit A_{ij} das Ereignis "$\widetilde{X}(\omega_i) = a_j$" bezeichnet, so gilt für beliebige $j_1, j_2, \ldots, j_n$:
$P(A_{1j_1} \cap A_{2j_2} \cap \ldots \cap A_{nj_n}) = P(A_{1j_1}) \cdot P(A_{2j_2}) \cdot \ldots \cdot P(A_{nj_n}) = f_{j_1} \cdot f_{j_2} \cdot \ldots \cdot f_{j_n}$ (1.3.3)

1.3.4 Koppelung abhängiger Experimente, Satz von der totalen Wahrscheinlichkeit, Markovmodelle

Satz 1.14 *Satz von der totalen Wahrscheinlichkeit*
Gegeben sei eine vollständige Zerlegung $A_1, A_2 \ldots, A_k$. Dann gilt für jedes Ereignis B

$$P(B) = \sum_{j=1}^{k} P(B|A_j) \cdot P(A_j) \qquad (1.3.4)$$

Bem 1.15 *Koppelung abhängiger Experimente*
Gegeben seien n Experimente, beschrieben durch die Grundräume $\Omega_i = \{a_{i1}, \ldots, a_{ik_i}\}$ und die Wahrscheinlichkeiten $P_i, i = 1, \ldots, n$. Bezeichnet man für beliebiges $i = 1, \ldots, n$ und $j = 1, \ldots, k_i$, mit A_{ij} jeweils das zu a_{ij} gehörige Elementarereignis (also das Ereignis "a_{ij} tritt ein"), so gilt:

$$P(A_{1j_1} \cap A_{2j_2} \cap \ldots \cap A_{nj_n}) =$$
$$P_1(A_{1j_1}) \cdot P_2\left(A_{2j_2}|A_{1j_1}\right) \cdot P_3\left(A_{3j_3}|A_{1j_1} \cap A_{2j_2}\right) \cdot \ldots \cdot P_n\left(A_{nj_n}|A_{1j_1} \cap A_{2j_2} \cap \ldots \cap A_{n-1j_{n-1}}\right) \qquad (1.3.5)$$

Korollar 1.16 zu Satz 1.14
Sei $A_1, A_2, \ldots, A_k$ eine vollständige Zerlegung. Dann gilt für beliebige Ereignisse B und C mit $P(C) > 0$:

$$P(B|C) = \sum_{j=1}^{k} P(B|A_j \cap C) \cdot P(A_j|C) \qquad (1.3.6)$$

Def 1.17 *Markovmodelle*
Gilt in der Situation von Bem 1.15 $\Omega_1 = \Omega_2 = \ldots = \Omega_n = \{a_1, \ldots, a_k\}$ und sind alle bedingten Wahrscheinlichkeiten nur vom jeweils unmittelbar vorhergehenden Zeitpunkt abhängig, d.h. gilt

$$P(A_{i+1,j_{i+1}}|A_{1j_1} \cap A_{2j_2} \cap \ldots \cap A_{ij_i}) = P(A_{i+1,j_{i+1}}|A_{ij_i}) \qquad (1.3.7)$$

so spricht man von einem Markovmodell mit *den Zuständen* $a_1, \ldots, a_k$. Sind die sog. Übergangswahrscheinlichkeiten in (1.3.7) unabhängig von der Zeit, gilt also $P(A_{i+1,j}|A_{il}) \equiv p_{jl}$ für alle i, j, l, so heißt das Markovmodell *homogen*.

1.3.5 Das Theorem von Bayes

Falls $P(B) > 0$ gilt:

$$P(A|B) = P(B|A) \cdot \frac{P(A)}{P(B)} \qquad (1.3.8)$$

Satz 1.20 *Theorem von Bayes*
Sei $A_1, \dots A_k$ eine vollständige Zerlegung von Ω (wobei für mindestens ein $i, i = 1, \dots k$, $P(A_i) > 0$ und $P(B|A_i) > 0$ erfüllt sei.) Dann gilt

$$P(A_j|B) = \frac{P(B|A_j) \cdot P(A_j)}{\sum_{i=1}^{k} P(B|A_i) \cdot P(A_i)} \qquad (1.3.9)$$

1.4 Zufallsvariablen, Verteilungsfunktion, Dichte

1.4.1 Diskrete Zufallselemente und Zufallsvariablen

Def und Bem 1.24

a) Gegeben seien ein diskreter, d.h. höchstens abzählbarer, Ergebnisraum Ω und die Wahrscheinlichkeit P auf Ω.
Jede Abbildung

$$\begin{array}{rcl} X : \Omega & \mapsto & \Omega_X \\ \omega & \mapsto & X(\omega) \end{array}$$

heißt Zufalls*element*. Setzt man für jede *Realisation* $x \in \Omega_X$

$$P_X(\{x\}) := P(\{X = x\}) := P(\{\omega | X(\omega) = x\}) \qquad (1.4.1)$$

so erhält man eine Wahrscheinlichkeit auf Ω_X. (Oft wird auch $P(X = x)$ statt $P(\{X = x\})$ geschrieben.)
Es ist häufig üblich, bei P_X den Index wegzulassen, also $P(\{x\})$ statt $P_X(\{x\})$ zu schreiben.
P_X heißt *Wahrscheinlichkeitsverteilung* von X.

b) Ist $\Omega_X = \mathbb{R}$, so bezeichnet man das Zufallselement X als *Zufallsvariable*.

Bem und Bsp 1.25
$(X_1, X_2, \dots, X_n)$: *i.i.d. Stichprobe* oder *reine Zufallsstichprobe* des Merkmals $\widetilde{X}$.

gemeinsame Verteilung:
$P(\{X_1 = x_1\} \cap \{X_2 = x_2\} \cap \dots \cap \{X_n = x_n\}) =$
$P(\{X_1 = x_1\}) \cdot P(\{X_2 = x_2\}) \cdot \dots \cdot P(\{X_n = x_n\})$

1.4.2 Die Verteilungsfunktion

Bem 1.27 Die Wahrscheinlichkeitsverteilung einer Zufallsvariable X kann man durch die *Verteilungsfunktion*

$$F(x) := P(X \leq x) \qquad (1.4.3)$$

eineindeutig erklären.
Die Wahrscheinlichkeit anderer Ereignisse ergibt sich aus dem (geeignet verallgemeinerten) dritten Kolmogorovschen Axiom.
Es gilt dann

$$P(a < X \leq b) = F(b) - F(a) \qquad (1.4.4)$$

1.4.3 Stetige Zufallsvariablen

$$P_X(\{x\}) = 0 \text{ für jedes } x \in \mathbb{R} \qquad (1.4.5)$$

Die Wahrscheinlichkeitsverteilung ist durch die Verteilungsfunktion $F(x) = P(\{X \leq x\})$ eineindeutig festgelegt. Es gilt dann:

$$P(a < X \leq b) = P(a \leq X \leq b) = P(a < X < b) = P(a \leq X < b) = F(b) - F(a) \qquad (1.4.6)$$

Def 1.30 Gegeben sei eine stetige Zufallsvariable X mit differenzierbarer Verteilungsfunktion $F_X(x)$.
Dann heißt die Ableitung von $F(x)$ nach x, also

$$f(x) = \frac{dF(x)}{dx} \qquad (1.4.7)$$

Dichte der Zufallsvariablen X.

Satz 1.31 Es gilt dann in der Situation von Def 1.30

$$F(x) = \int_{-\infty}^{x} f(u)\ du \qquad (1.4.8)$$

und für beliebige reelle Zahlen a und b mit $a < b$ und damit

$$P(a \leq X \leq b) = P(a < X \leq b) = P(a < X < b) = P(a \leq X < b) = \int_{a}^{b} f(x)\,dx. \qquad (1.4.9)$$

1.4.4 Lebensdauern; Hazardrate und Survivorfunktion

Bem 1.34 Die Verteilung einer nichtnegativen stetigen Zufallsvariable X wird auch eineindeutig durch die sog. *Überlebensfunktion* (Survivorfunktion)

$$S(x) := P(X \geq x) = 1 - F(x) \qquad (1.4.10)$$

und durch die *Hazardrate*

$$\lambda(x) := \lim_{h \to 0} \frac{P(x \leq X \leq x + h | X \geq x)}{h} \qquad (1.4.11)$$

beschrieben.
Es gilt:

$$S(x) = \exp(-\int_{0}^{x} \lambda(u) du) \qquad (1.4.12)$$

also

$$F(x) = 1 - \exp(-\int_0^x \lambda(u)du) \qquad (1.4.13)$$

und

$$f(x) = \lambda(x) \cdot S(x) \qquad (1.4.14)$$

1.4.5 Unabhängigkeit von Zufallsvariablen

Def 1.35 Zwei Zufallsvariablen X und Y mit den Verteilungsfunktionen F_X und F_Y heißen *stochastisch unabhängig*, falls für alle x und y gilt

$$P(\{X \leq x\} \cap \{Y \leq y\}) = P(\{X \leq x\}) \cdot P(\{Y \leq y\}) = F_X(x) \cdot F_Y(y),$$

widrigenfalls stochastisch abhängig.

1.5 Erwartungswert und Varianz

1.5.1 Diskrete Zufallsvariablen

Def 1.34 Gegeben sei eine diskrete Zufallsvariable X mit der Wahrscheinlichkeitsverteilung P. Die Menge

$$\mathcal{X} := \{x \in \mathbb{R} | P(\{x\}) > 0\}$$

heißt *Träger von X*.

Def 1.35 Gegeben sei eine diskrete Zufallsvariable X mit Träger $\mathcal{X}$. Dann heißt

$$\mathbb{E}X := \sum_{x \in \mathcal{X}} x \cdot P(X = x) \qquad (1.5.1)$$

Erwartungswert von X,

$$VarX = \mathbb{V}(X) := \mathbb{E}((X - \mathbb{E}(X))^2) = \sum_{x \in \mathcal{X}} (x - \mathbb{E}(X))^2 \cdot P(X = x) \qquad (1.5.2)$$

Varianz von X

und $\sigma_X := \sqrt{VarX}$ *Standardabweichung* von X.

Verschiebungssatz: Es gilt:

$$VarX = \mathbb{E}(X^2) - (\mathbb{E}X)^2 \qquad (1.5.3)$$

1.5.2 Stetige Zufallsvariable

Def 1.38 Sei X eine stetige Zufallsvariable mit Dichte $f(x)$. Dann heißt

$$\mathbb{E}X := \int_{-\infty}^{\infty} x \cdot f(x)dx \qquad (1.5.4)$$

Erwartungswert von X,

$$VarX := \mathbb{V}(X) := \mathbb{E}((X - \mathbb{E}(X))^2 = \int_{-\infty}^{\infty} (x - \mathbb{E}(X))^2 \cdot f(x)\ dx \qquad (1.5.5)$$

Varianz von X

und $\sigma_X := \sqrt{VarX}$ *Standardabweichung* von X.

1.5.3 Allgemeine Rechenregeln für Erwartungswert und Varianz

Satz 1.39

$$\begin{array}{lllll}
\text{a)} & \mathbb{E}(aX + bY) & = & a \cdot \mathbb{E}(X) + b \cdot \mathbb{E}(Y) & (1.5.6) \\
 & \text{Insb. } \mathbb{E}(a) & = & a & \\
 & \text{und } \mathbb{E}(aX) & = & a \cdot \mathbb{E}(X) & \\
 & \text{und } \mathbb{E}(X + Y) & = & \mathbb{E}(X) + \mathbb{E}(Y) & \\
\text{b)} & Var(aX + b) & = & a^2 \cdot VarX & (1.5.7) \\
\text{c)} & \multicolumn{3}{l}{\text{sind } X \text{ und } Y \text{ unabhängig, so gilt}} & \\
 & \mathbb{E}(X \cdot Y) & = & \mathbb{E}(X) \cdot \mathbb{E}(Y) & (1.5.8) \\
 & Var(X + Y) & = & Var(X) + Var(Y) & (1.5.9)
\end{array}$$

Def und Bem 1.40 Die Zufallsvariable

$$Z := \frac{X - \mathbb{E}(X)}{\sqrt{Var(X)}} \qquad (1.5.10)$$

heißt *standardisierte Zufallsvariable*; es gilt $\mathbb{E}(Z) = 0$ und $Var(Z) = 1$.

1.6 Wichtige Verteilungsmodelle

1.6.1 Binomialverteilung

Def 1.41 Seien n und k natürliche Zahlen oder 0.

$$\begin{array}{ll}
\text{i)} & n! := n \cdot (n-1) \cdot (n-2) \cdot \ldots \cdot 2 \cdot 1 \\
 & 0! := 1 \qquad (1.6.1) \\
\text{ii)} & \binom{n}{k} := \dfrac{n!}{k!(n-k)!} \text{ heißt } \textit{Binomialkoeffizient.}
\end{array}$$

<u>Bem 1.42</u> *Das Grundmodell der Binomialverteilung*

- n unabhängige Wiederholungen eines Experiments, bei dem ein Ereignis A mit Wahrscheinlichkeit π eintritt.
- Betrachtet wird die Zufallsvariable X, die die Anzahl der Versuche zählt, bei denen A eintritt.
- häufig A: "Treffer", X: Anzahl der "Treffer".
- Es gilt *für alle* $x \in \{0, 1, \ldots, n\}$

$$P(X = x) = \binom{n}{x} \cdot \pi^x \cdot (1 - \pi)^{n-x} \tag{1.6.2}$$

(und $P(X = x) = 0$ für alle anderen x)

- X heißt *binomialverteilt*, das Wahrscheinlichkeitsmaß aus (1.6.2) heißt *Binomialverteilung. Abkürzung* $X \sim B(n, \pi)$
- Es gilt für binomialverteiltes X

  - $\mathbb{E}X = n \cdot \pi$ (1.6.3)
  - $VarX = n \cdot \pi(1 - \pi)$ (1.6.4)

<u>Bem 1.44</u> *Zwei weitere Eigenschaften der Binomialverteilung:*

- Symmetrieeigenschaft:
  Sei $X \sim B(n, \pi)$ und $Y = n - X$. Dann ist $Y \sim B(n, 1 - \pi)$.
- Summeneigenschaft:
  Seien $X \sim B(n, \pi)$ und $Y \sim B(m, \pi)$. Sind X und Y unabhängig, so ist
  $X + Y \sim B(n + m, \pi)$.

1.6.2 Normalverteilung

<u>Def 1.46</u> *Definition*
Eine stetige Zufallsvariable X heißt *normalverteilt* mit den Parametern μ und σ^2, in Zeichen $X \sim \mathcal{N}(\mu, \sigma^2)$, wenn für ihre Dichte gilt:

$$f(x) = \frac{1}{\sqrt{2\pi} \cdot \sigma} \exp\left(-\frac{1}{2\sigma^2}(x - \mu)^2\right), \; x \in \mathbb{R}, \tag{1.6.5}$$

und *standardnormalverteilt*, in Zeichen $X \sim \mathcal{N}(0; 1)$, falls $\mu = 0$ und $\sigma^2 = 1$.

Die *Dichte der Standardnormalverteilung* wird oft mit $\varphi(x)$ bezeichnet, also

$$\varphi(x) = \frac{1}{\sqrt{2\pi}} \exp\left(-\frac{1}{2}x^2\right), \qquad (1.6.6)$$

die zugehörige Verteilungsfunktion $\int_{-\infty}^{x} \varphi(u)\, du$ mit $\Phi(x)$.

Bem 1.48 *Grundlegendes zum Rechnen mit Normalverteilungen*

- $\Phi(-x) = 1 - \Phi(x)$ (1.6.7)
- Ist $X \sim \mathcal{N}(\mu, \sigma^2)$, so ist die zugehörige standardisierte Zufallsvariable $Z = \dfrac{X-\mu}{\sigma}$ standardnormalverteilt.
- $P(X \le a) = \Phi\left(\dfrac{a-\mu}{\sigma}\right)$ (1.6.8)

Prop 1.49 *Abgeschlossenheit gegenüber Linearkombinationen*:
Seien X_1 und X_2 unabhängig und $X_i \sim \mathcal{N}(\mu_i, \sigma_i^2),\ i = 1, 2$. Ferner seien b, a_1, a_2 feste reelle Zahlen. Dann gilt

$$Y_1 := a_1X_1 + b \sim \mathcal{N}(a_1\mu_1 + b; a_1^2\sigma_1^2) \qquad (1.6.9)$$

und

$$Y_2 := a_1X_1 + a_2X_2 \sim \mathcal{N}(a_1\mu_1 + a_2\mu_2; a_1^2\sigma_1^2 + a_2^2\sigma_2^2) \qquad (1.6.10)$$

1.7 Unabhängige und identische Wiederholungen, Grenzwertsätze und Approximationen

1.7.1 Das i.i.d. Modell

Betrachtet werden diskrete oder stetige Zufallsvariablen $X_1, \dots, X_n$, die unabhängig sind und dieselbe Verteilung besitzen ("$X_1, \dots, X_n$ i.i.d." (independently identically distributed)). Ferner existiere der Erwartungswert μ und die Varianz σ^2, die Verteilungsfunktion werde mit F bezeichnet.

1.7.2 Das Schwache Gesetz der großen Zahlen

Satz 1.51 *Theorem von Bernoulli*

Seien $X_1, \dots, X_n$, i.i.d. mit $X_i \in \{0, 1\}$ und $P(X_i = 1) = \pi$. Dann gilt für $H_n = \frac{1}{n}\sum_{i=1}^{n} X_i$ (relative Häufigkeit der "Einsen") und beliebig kleines $\epsilon > 0$

$$\lim_{n\to\infty} P(|H_n - \pi| \le \epsilon) = 1 \qquad (2.1.1)$$

Satz 1.52 *Schwaches Gesetz der großen Zahl*
Gegeben seien $X_1, \dots, X_n$, i.i.d. Zufallsvariablen mit (existierendem) Erwartungswert μ und

(existierender) Varianz σ^2. Dann gilt für $\overline{X}_n := \frac{1}{n}\sum_{i=1}^{n} X_i$ und beliebiges $\epsilon > 0$:

$$\lim_{n\to\infty} P(|\overline{X}_n - \mu| \leq \epsilon) = 1 \qquad (2.2.2)$$

Schreibweise: $\overline{X}_n \underset{P}{\longrightarrow} \mu$ ("Stochastische Konvergenz"; "$\overline{X}_n$ konvergiert in Wahrscheinlichkeit gegen μ.")

1.7.3 Hauptsatz der Statistik

Satz 1.53 *Hauptsatz der Statistik*:
$X_1, \ldots, X_n$ i.i.d. mit Verteilungsfunktion F
Sei $F_n^{X_1,\ldots,X_n}(x)$ die empirische Verteilungsfunktion der ersten n Beobachtungen und

$$D_n := \sup_x |F_n^{X_1,\ldots,X_n}(x) - F(x)|, \qquad (2.2.4)$$

so gilt für jedes $c > 0$

$$\lim_{n\to\infty} P(D_n > c) = 0\ . \qquad (2.2.5)$$

1.7.4 Der zentrale Grenzwertsatz

Satz 1.54 *Zentraler Grenzwertsatz*
Seien $X_1, \ldots, X_n$ i.i.d. mit $\mathbb{E}(X_i) = \mu$ und $Var(X_i) = \sigma^2 > 0$ und

$$Z_n = \frac{1}{\sqrt{n}} \sum_{i=1}^{n} \left(\frac{X_i - \mu}{\sigma} \right) \qquad (2.3.1)$$

Dann gilt: Z_n ist *asymptotisch standardnormalverteilt*, in Zeichen: $Z_n \overset{asym}{\sim} \mathcal{N}(0;1)$, d.h. es gilt für jedes z

$$\lim_{n\to\infty} P(Z_n \leq z) = \Phi(z). \qquad (2.3.2)$$

Approximation der Binomialverteilung (mit Stetigkeitskorrektur)

$$P(X \leq x) \approx \Phi\left(\frac{x + 0.5 - n\pi}{\sqrt{n\pi(1-\pi)}} \right) \qquad (2.3.5)$$

und

$$P(X = x) \approx \Phi\left(\frac{x + 0.5 - n\pi}{\sqrt{n\pi(1-\pi)}} \right) - \Phi\left(\frac{x - 0.5 - n\pi}{\sqrt{n\pi(1-\pi)}} \right) \qquad (2.3.6)$$

1.8 Mehrdimensionale Zufallselemente

1.8.1 Grundbegriffe

Def. 1.58 Betrachtet werden zwei eindimensionale Zufallselemente X und Y (zu demselben Zufallsexperiment). Die Wahrscheinlichkeit

$$P(X = x_i, Y = y_j) := P(\{X = x_i\} \cap \{Y = y_j\})$$

in Abhängigkeit von x_i und y_j heißt *gemeinsame Verteilung* der mehrdimensionalen Zufallsvariable $\binom{X}{Y}$ und der Variablen X und Y.
Randwahrscheinlichkeiten:

$$p_{i\cdot} = P(X = x_i) = \sum_{j=1}^{m} P(X = x_i, Y = y_j) \qquad (3.1.2)$$

$$p_{\cdot j} = P(Y = y_j) = \sum_{i=1}^{k} P(X = x_i, Y = y_j) \qquad (3.1.3)$$

bedingte Verteilungen:

$$P(X = x_i | Y = y_j) = \frac{P(X = x_i, Y = y_j)}{P(Y = y_j)} \qquad (3.1.5)$$

$$P(Y = y_j | X = x_i) = \frac{P(X = x_i, Y = y_j)}{P(X = x_i)} \qquad (3.1.6)$$

stetiger Fall:
zweidimensionale Dichtefunktionen $f(x, y)$:

$$P(a \le X \le b, c \le Y \le d) = \int_a^b \left(\int_c^d f(x, y) dy \right) dx$$

1.8.2 Kovarianz und Korrelation

Def 1.59 *Kovarianz*
X, Y Zufallsvariable. Dann heißt

$$\sigma_{X,Y} := Cov(X, Y) := \mathbb{E}((X - \mathbb{E}(X))(Y - \mathbb{E}(Y))) \qquad (3.2.1)$$

Kovarianz von X und Y
Rechenregeln:

- $Cov(X, X) = Var(X)$ (3.2.2)
- $Cov(X, Y) = \mathbb{E}(XY) - \mathbb{E}(X) \cdot \mathbb{E}(Y)$ (3.2.3)
- $Cov(X, Y) = Cov(Y, X)$ (3.2.4)
- Mit $\tilde{X} = a_X X + b_X$ und $\tilde{Y} = a_Y Y + b_Y$ ist
  $Cov(\tilde{X}, \tilde{Y}) = a_X \cdot a_Y \cdot Cov(X, Y)$ (3.2.5)
- $Var(X + Y) = Var(X) + Var(Y) + 2 \cdot Cov(X, Y)$ (3.2.6)

Def. 1.60 Zwei Zufallsgrößen X, Y mit $Cov(X, Y) = 0$ heißen *unkorreliert.*

Def. 1.62 *Korrelationskoeffizient*
Gegeben seien zwei Zufallsgrößen X und Y. Dann heißt

$$\rho(X,Y) = \frac{Cov(X,Y)}{\sqrt{VarX}\sqrt{VarY}} \tag{3.2.7}$$

Korrelationskoeffizient von X und Y.

Eigenschaften des Korrelationskoeffizienten

i) Mit $\tilde{X} = a_X X + b_X$ und $\tilde{Y} = a_Y Y + b_Y$ ist $|\rho(\tilde{X},\tilde{Y})| = |\rho(X,Y)|$.

ii) $-1 \leq \rho(X,Y) \leq 1$

iii) $|\rho(X,Y)| = 1 \Longleftrightarrow Y = aX + b$

iv) Ist $VarX > 0$, $Var(Y) > 0$, so ist
$\rho(X,Y) = 0$ genau dann, wenn $Cov(X,Y) = 0$.

2. Induktive Statistik

2.1 Grundprinzipien der induktiven Statistik

2.2 Punktschätzung

2.2.1 Schätzfunktion, typische Beispiele

Def 2.1 Seien $X_1, \ldots, X_n$ i.i.d. Wiederholungen. Jede Funktion

$$T = g(X_1, \ldots, X_n)$$

heißt *Schätzer* oder *Schätzfunktion*.

Typische Beispiele für Schätzfunktionen

i) $\overline{X} = g(X_1, \ldots, X_n) = \frac{1}{n} \sum_{i=1}^{n} X_i$ (2.2.1)

arithmetisches Mittel der Stichprobe.

Falls $X_i = \begin{cases} 0 \\ 1 \end{cases} \Longrightarrow \overline{X}$ relative Häufigkeit in der Stichprobe

ii) $\tilde{S^2} = g(X_1, \ldots, X_n) = \frac{1}{n} \sum_{i=1}^{n} (X_i - \overline{X})^2$ (2.2.2)

Stichprobenvarianz

iii) $S^2 = g(X_1, \ldots, X_n) = \frac{1}{n-1} \sum_{i=1}^{n} (X_i - \overline{X})^2 = \frac{1}{n-1} \left(\sum_{i=1}^{n} X_i^2 - n \cdot (\overline{X})^2 \right)$ (2.2.3)

korrigierte Stichprobenvarianz

iv) $X_{(n)} = g(X_1, \ldots, X_n) = \max_{i=1,\ldots,n} X_i$ (2.2.4)

größter Stichprobenwert

v) $X_{(1)} = g(X_1, \ldots, X_n) = \min_{i=1,\ldots,n} X_i$ (2.2.5)

kleinster Stichprobenwert

2.2.2 Gütekriterien

Erwartungstreue, Bias

Def 2.5 Gegeben sei eine i.i.d. Stichprobe $X_1, \ldots, X_n$ und eine Schätzfunktion $T = g(X_1, \ldots, X_n)$ (mit existierendem Erwartungswert).

a) T heißt *erwartungstreu für den Parameter* ϑ, falls gilt

$$\mathbb{E}_\vartheta(T) = \vartheta \qquad (2.2.7)$$

für alle ϑ.

b) Die Größe

$$Bias_\vartheta(T) := \mathbb{E}_\vartheta(T) - \vartheta \qquad (2.2.8)$$

heißt *Bias* (oder Verzerrung) der Schätzfunktion.

2.2.3 Effizienz

Def 2.10 *Effizienz*

i) Gegeben seien zwei erwartungstreue Schätzfunktionen T_1 und T_2 für einen Parameter ϑ.
Ist $Var_\vartheta(T_1) \leq Var_\vartheta(T_2)$ für alle ϑ
und $Var_{\vartheta^*}(T_1) < Var_{\vartheta^*}(T_2)$ für ein ϑ^*,
so heißt T_1 *effizienter als* T_2.

ii) Eine für ϑ erwartungstreue Schätzfunktion T^* heißt *UMVU-Schätzfunktion* für ϑ (*u*niformly *m*inimum *v*ariance *u*nbiased), falls

$$Var_\vartheta(T^*) \leq Var_\vartheta(T)$$

für alle ϑ und für alle erwartungstreuen Schätzer T.

Mean Square Error:

$$MSE_\vartheta(T) := \mathbb{E}(T-\vartheta)^2 \tag{2.2.9}$$
$$MSE_\vartheta(T) = Var_\vartheta(T) + (Bias_\vartheta(T))^2 \tag{2.2.10}$$

2.2.4 Konstruktionsprinzipien guter Schätzer

Def 2.11

- Gegeben sei die Realisation $x_1, \ldots, x_n$ einer *i.i.d.* Stichprobe.
  Die Funktion in ϑ
  $$Lik(\vartheta || x_1, \ldots, x_n) = \begin{cases} \prod_{i=1}^{n} P_\vartheta(X_i = x_i) & X_i \text{ diskret} \\ & \text{falls} \\ \prod_{i=1}^{n} f_\vartheta(x_i) & X_i \text{ stetig} \end{cases} \tag{2.2.11}$$
  heißt *Likelihood* des Parameters ϑ bei der Beobachtung $x_1, \ldots, x_n$.
- Derjenige Wert $\widehat{\vartheta} = \widehat{\vartheta}(x_1, \ldots, x_n)$, der $Lik(\vartheta || x_1, \ldots, x_n)$ maximiert, heißt *Maximum-Likelihood-Schätzwert*; die zugehörige Schätzfunktion $T(X_1, \ldots, X_n)$ *Maximum-Likelihood-Schätzer.*

2.3. Intervallschätzung / Konfidenzintervalle

2.3.1 Motivation und Hinführung

2.3.2 Definition von Konfidenzintervallen

Def 2.13
Gegeben sei eine i.i.d. Stichprobe $X_1, \ldots, X_n$ zur Schätzung eines Parameters ϑ und eine Zahl $\gamma \in (0; 1)$. Ein zufälliges Intervall $\mathcal{C}(X_1, \ldots, X_n)$ heißt *Konfidenzintervall* zum *Sicherheitsgrad* γ (Konfidenzniveau γ), falls für jedes ϑ

$$P_\vartheta(\vartheta \in \mathcal{C}(X_1, \ldots, X_n)) \geq \gamma. \qquad (2.3.1).$$

2.3.3 Typische Beispiele für die Konstruktion von Konfidenzintervallen

Bsp 2.15 Konfidenzintervall für den Mittelwert eines normalverteilten Merkmals bei *bekannter Varianz*

$$\left[\overline{X} \pm \frac{z_{\frac{1+\gamma}{2}} \cdot \sigma}{\sqrt{n}}\right] \tag{2.3.2}$$

Def 2.17 Gegeben sei eine i.i.d. Stichprobe $X_1, \ldots, X_n$ mit $X_i \sim \mathcal{N}(\mu, \sigma^2)$. Dann heißt die Verteilung von

$$Z = \frac{\overline{X} - \mu}{S} \cdot \sqrt{n}$$

t-Verteilung (oder Student-Verteilung) mit $\nu = n - 1$ *Freiheitsgraden.*
In Zeichen: $Z \sim t(\nu)$

Bsp 2.16 Konfidenzintervall für den Mittelwert eines normalverteilten Merkmals bei *unbekannter Varianz*

$$\left[\overline{X} \pm \frac{t_{\frac{1+\gamma}{2}}(n-1) \cdot S}{\sqrt{n}}\right] \tag{2.3.3}$$

Bsp 2.18b *Konfidenzintervall für den Anteil π, basierend auf $\hat{\pi} = \bar{X}$*

$$\left[\overline{X} \pm z_{\frac{1+\gamma}{2}} \cdot \sqrt{\frac{\overline{X}(1-\overline{X})}{n}}\right] \tag{2.3.4}$$

2.4. Grundprinzipien statistischer Hypothesentests

2.4.1 Motivation und Hinführung

2.4.2 Die prinzipielle Vorgehensweise bei einem parametrischen statistischen Test

2.4.3 Typische Tests I: Ein-Stichproben-Gauss-Test, t-Test und Test auf einen Anteilswert

Einstichproben-Testprobleme

Formulierung der Hypothesen:

- Zweiseitiges Testproblem:

$$(a) \qquad H_0 : \theta = \theta_0 \quad \text{vs.} \quad H_1 : \theta \neq \theta_0$$

- Einseitige Testprobleme:

$$(b) \qquad H_0 : \theta \geq \theta_0 \quad \text{vs.} \quad H_1 : \theta < \theta_0$$

$$(c) \qquad H_0 : \theta \leq \theta_0 \quad \text{vs.} \quad H_1 : \theta > \theta_0$$

Verteilung	θ	Teststatistik	Ablehnbereiche
$X \sim \mathcal{N}(\mu, \sigma^2)$, σ^2 bekannt	μ	$T = \frac{\bar{X} - \mu_0}{\sigma}\sqrt{n}$	(a) $\lvert T\rvert > z_{1-\frac{\alpha}{2}}$ (b) $T < -z_{1-\alpha}$ (c) $T > z_{1-\alpha}$
$X \sim \mathcal{N}(\mu, \sigma^2)$, σ^2 unbekannt	μ	$T = \frac{\bar{X} - \mu_0}{S}\sqrt{n}$	(a) $\lvert T\rvert > t_{1-\frac{\alpha}{2}}(n-1)$ (b) $T < -t_{1-\alpha}(n-1)$ (c) $T > t_{1-\alpha}(n-1)$
X beliebig verteilt, $n \geq 30$, $\mathrm{E}(X) = \mu$, $\mathrm{Var}(X) = \sigma^2$ bekannt	μ	$T = \frac{\bar{X} - \mu_0}{\sigma}\sqrt{n}$	(a) $\lvert T\rvert > z_{1-\frac{\alpha}{2}}$ (b) $T < -z_{1-\alpha}$ (c) $T > z_{1-\alpha}$
X beliebig verteilt, $n \geq 30$, $\mathrm{E}(X) = \mu$, $\mathrm{Var}(X) = \sigma^2$ unbekannt	μ	$T = \frac{\bar{X} - \mu_0}{S}\sqrt{n}$	(a) $\lvert T\rvert > z_{1-\frac{\alpha}{2}}$ (b) $T < -z_{1-\alpha}$ (c) $T > z_{1-\alpha}$
X dichotom $\Leftrightarrow \sum_{i=1}^{n} X_i \sim B(n, \pi)$	π	$T = \frac{\bar{X} - \pi_0}{\sqrt{\pi_0(1-\pi_0)}}\sqrt{n}$	(a) $\lvert T\rvert > z_{1-\frac{\alpha}{2}}$ (b) $T < -z_{1-\alpha}$ (c) $T > z_{1-\alpha}$

2.4.4 Lagevergleiche aus unabhängigen Stichproben

Zweistichproben-Mittelwertsvergleiche

Bezeichnungen:
$X_1, \dots, X_n$ unabhängig und identisch verteilt mit $\mathbb{E}(X_i) = \mu_X$
$Y_1, \dots, Y_m$ unabhängig und identisch verteilt $\mathbb{E}(Y_i) = \mu_Y$
$X_1, \dots, X_n, Y_1, \dots, Y_m$ unabhängig

Formulierung der Hypothesen:

- Zweiseitiges Testproblem:

$$(a) \qquad H_0 : \mu_X = \mu_Y \quad \text{vs.} \quad H_1 : \mu_X \neq \mu_Y$$

- Einseitige Testprobleme:

$$(b) \qquad H_0 : \mu_X \geq \mu_Y \quad \text{vs.} \quad H_1 : \mu_X < \mu_Y$$

$$(c) \qquad H_0 : \mu_X \leq \mu_Y \quad \text{vs.} \quad H_1 : \mu_X > \mu_Y$$

Verteilung	Teststatistik	Ablehnbereiche
$X \sim \mathcal{N}(\mu_X, \sigma_X^2)$, $Y \sim \mathcal{N}(\mu_Y, \sigma_Y^2)$ σ_X^2, σ_Y^2 bekannt	$T = \frac{\bar{X} - \bar{Y}}{\sqrt{\frac{\sigma_X^2}{n} + \frac{\sigma_Y^2}{m}}}$	(a) $\lvert T\rvert > z_{1-\frac{\alpha}{2}}$ (b) $T < -z_{1-\alpha}$ (c) $T > z_{1-\alpha}$
$X \sim \mathcal{N}(\mu_X, \sigma_X^2)$, $Y \sim \mathcal{N}(\mu_Y, \sigma_Y^2)$ $\sigma_X^2 = \sigma_Y^2$ unbekannt	$T = \frac{\bar{X} - \bar{Y}}{\sqrt{\left(\frac{1}{n} + \frac{1}{m}\right) \frac{(n-1)S_X^2 + (m-1)S_Y^2}{n+m-2}}}$	(a) $\lvert T\rvert > t_{1-\frac{\alpha}{2}}(n+m-2)$ (b) $T < -t_{1-\alpha}(n+m-2)$ (c) $T > t_{1-\alpha}(n+m-2)$
X, Y beliebig verteilt und $n, m > 30$, oder $X \sim \mathcal{N}(\mu_X, \sigma_X^2)$, $Y \sim \mathcal{N}(\mu_Y, \sigma_Y^2)$, jeweils σ_X^2, σ_Y^2 unbekannt	$T = \frac{\bar{X} - \bar{Y}}{\sqrt{\frac{S_X^2}{n} + \frac{S_Y^2}{m}}}$	(a) $\lvert T\rvert > z_{1-\frac{\alpha}{2}}$ (b) $T < -z_{1-\alpha}$ (c) $T > z_{1-\alpha}$

2.4.5 Gauss-Test und t-Test für verbundene Stichproben

Verbundene Stichproben:
Annahmen: $(X_1, Y_1), \dots, (X_n, Y_n)$ i.i.d., $X_i \sim \mathcal{N}(\mu_X, \sigma_X^2)$, $Y_i \sim \mathcal{N}(\mu_Y, \sigma_Y^2)$, Hypothesen wie oben.
Zurückführen auf Einstichprobenproblem durch Betrachten von $D_i = X_i - Y_i$ mit
$D_i \sim \mathcal{N}(\mu_D, \sigma_D^2)$, wobei $\mu_D = \mu_X - \mu_Y$ und $\sigma_D^2 = \sigma_X^2 + \sigma_Y^2 - 2 \cdot \sigma_{X,Y}$, sowie der Hypothesen:

a) $H_0 : \mu_D = 0$ vs. $H_1 : \mu_D \neq 0$

b) $H_0 : \mu_D \geq 0$ vs. $H_1 : \mu_D < 0$

c) $H_0 : \mu_D \leq 0$ vs. $H_1 : \mu_D > 0$

Testgröße: $T = \frac{\bar{D}}{S_D} \cdot \sqrt{n}$ mit $\bar{D} = \frac{1}{n} \cdot \sum_{i=1}^{n} D_i$ und $S_D^2 = \frac{1}{n-1} \sum_{i=1}^{n} (D_i - \bar{D})^2$.

2.4.6 χ^2-Test(s)

χ2-Unabhängigkeitstest:

- Annahmen: $(X_1, Y_1), \dots, (X_n, Y_n)$ unabhängige Wiederholungen von (X, Y)
- Hypothesen:

$$\begin{array}{ll} H_0: & X \text{ und } Y \text{ sind stochastisch unabhängig} \\ H_1: & X \text{ und } Y \text{ sind nicht stochastisch unabhängig} \end{array}$$

$$\begin{array}{cc|ccc|c} & & & Y & & \\ & & 1 & \dots & m & \\ \hline & 1 & h_{11} & \dots & h_{1m} & h_{1\cdot} \\ X & \vdots & \vdots & & \vdots & \vdots \\ & \text{k} & h_{k1} & \dots & h_{km} & h_{k\cdot} \\ \hline & & h_{\cdot 1} & \dots & h_{\cdot m} & n \end{array}$$

- Teststatistik:

$$T = \sum_{i=1}^{k} \sum_{j=1}^{m} \frac{(h_{ij} - \tilde{h}_{ij})^2}{\tilde{h}_{ij}} \quad \text{mit} \quad \tilde{h}_{ij} = \frac{h_{i\cdot} h_{\cdot j}}{n}$$

  Alternative Form:

$$T = \sum_{i=1}^{k} \sum_{j=1}^{m} n \cdot \frac{\left(\frac{h_{ij}}{n} - \frac{h_{i\cdot} h_{\cdot j}}{n^2} \right)^2}{\frac{h_{i\cdot} h_{\cdot j}}{n^2}}$$

- Ablehnungsbereich:

$$T > \chi^2_{1-\alpha}((k-1) \cdot (m-1))$$

2.4.7 Zur praktischen Anwendung statistischer Tests - Typische Fallen

2.5. Inferenz bei der linearen Regression

2.5.1 Erinnerung an das Grundmodell

2.5.2 Lineare Einfachregression

Regressionsansatz:

$$Y_i = \beta_0 + \beta_1 X_i + \epsilon_i \ , \quad i = 1, \ldots, n. \qquad (2.5.1)$$

Normalverteilungsannahme:

$$\epsilon_i \sim \mathcal{N}(0, \sigma^2) \quad \Leftrightarrow \quad Y_i \sim \mathcal{N}(\beta_0 + \beta_1 X_i \ , \sigma^2), \quad i = 1, \ldots, n. \qquad (2.5.2)$$

Kleinste-Quadrate-Schätzer:

$$\hat{\beta}_0 = \bar{Y} - \hat{\beta}_1 \bar{X} \qquad (2.5.4), \qquad \hat{\beta}_1 = \frac{\sum\limits_{i=1}^{n} (X_i - \bar{X})(Y_i - \bar{Y})}{\sum\limits_{i=1}^{n} (X_i - \bar{X})^2} \qquad (2.5.3)$$

Residuen:

$$\hat{\epsilon}_i = Y_i - \hat{\beta}_0 - \hat{\beta}_1 X_i \ , \quad i = 1, \ldots, n. \qquad (2.5.6)$$

Schätzer für die Varianz $\boldsymbol{\sigma^2}$:

$$\hat{\sigma}^2 = \frac{1}{n-2} \sum_{i=1}^{n} \hat{\epsilon}_i^2 \qquad (2.5.5)$$

Verteilung der geschätzten Regressionskoeffizienten ($\sum$ heißt $\sum_{i=1}^{n}$):

$$\hat{\beta}_0 \sim \mathcal{N}(\beta_0, \sigma^2_{\hat{\beta}_0}) \quad \text{mit} \quad \widehat{\mathrm{Var}}(\hat{\beta}_0) = \hat{\sigma}^2_{\hat{\beta}_0} = \hat{\sigma}^2 \frac{\sum X_i^2}{n \sum (X_i - \bar{X})^2} = \hat{\sigma}^2 \frac{\sum X_i^2}{n(\sum X_i^2 - n\bar{X}^2)}$$

$$\hat{\beta}_1 \sim \mathcal{N}(\beta_1, \sigma^2_{\hat{\beta}_1}) \quad \text{mit} \quad \widehat{\mathrm{Var}}(\hat{\beta}_1) = \hat{\sigma}^2_{\hat{\beta}_1} = \frac{\hat{\sigma}^2}{\sum (X_i - \bar{X})^2} = \frac{\hat{\sigma}^2}{\sum X_i^2 - n\bar{X}^2}$$

Verteilung der standardisierten Schätzfunktionen:

$$\frac{\hat{\beta}_0 - \beta_0}{\hat{\sigma}_{\hat{\beta}_0}} \sim t(n-2) \quad \text{mit} \quad \hat{\sigma}_{\hat{\beta}_0} = \hat{\sigma} \frac{\sqrt{\sum X_i^2}}{\sqrt{n \sum (X_i - \bar{X})^2}} = \hat{\sigma} \frac{\sqrt{\sum X_i^2}}{\sqrt{n(\sum X_i^2 - n\bar{X}^2)}} \qquad (2.5.7)$$

$$\frac{\hat{\beta}_1 - \beta_1}{\hat{\sigma}_{\hat{\beta}_1}} \sim t(n-2) \quad \text{mit} \quad \hat{\sigma}_{\hat{\beta}_1} = \frac{\hat{\sigma}}{\sqrt{\sum (X_i - \bar{X})^2}} = \frac{\hat{\sigma}}{\sqrt{\sum X_i^2 - n\bar{X}^2}} \qquad (2.5.8)$$

(γ)-Konfidenzintervalle für $\boldsymbol{\beta_0}$ und $\boldsymbol{\beta_1}$:

$$\text{für } \beta_0: \quad \left[\hat{\beta}_0 - \hat{\sigma}_{\hat{\beta}_0} t_{\frac{1+\gamma}{2}}(n-2),\ \hat{\beta}_0 + \hat{\sigma}_{\hat{\beta}_0} t_{\frac{1+\gamma}{2}}(n-2)\right]$$

$$\text{für } \beta_1: \quad \left[\hat{\beta}_1 - \hat{\sigma}_{\hat{\beta}_1} t_{\frac{1+\gamma}{2}}(n-2),\ \hat{\beta}_1 + \hat{\sigma}_{\hat{\beta}_1} t_{\frac{1+\gamma}{2}}(n-2)\right]$$

Teststatistiken:

$$T_{\beta_0^*} = \frac{\hat{\beta}_0 - \beta_0^*}{\hat{\sigma}_{\hat{\beta}_0}} \quad \text{und} \quad T_{\beta_1^*} = \frac{\hat{\beta}_1 - \beta_1^*}{\hat{\sigma}_{\hat{\beta}_1}}$$

Hypothesen und Ablehnbereiche:

Hypothesen			Ablehnbereich
$H_0 : \beta_0 = \beta_0^*$	vs.	$H_1 : \beta_0 \neq \beta_0^*$	$\lvert T_{\beta_0^*}\rvert > t_{1-\frac{\alpha}{2}}(n-2)$
$H_0 : \beta_1 = \beta_1^*$	vs.	$H_1 : \beta_1 \neq \beta_1^*$	$\lvert T_{\beta_1^*}\rvert > t_{1-\frac{\alpha}{2}}(n-2)$
$H_0 : \beta_0 \geq \beta_0^*$	vs.	$H_1 : \beta_0 < \beta_0^*$	$T_{\beta_0^*} < -t_{1-\alpha}(n-2)$
$H_0 : \beta_1 \geq \beta_1^*$	vs.	$H_1 : \beta_1 < \beta_1^*$	$T_{\beta_1^*} < -t_{1-\alpha}(n-2)$
$H_0 : \beta_0 \leq \beta_0^*$	vs.	$H_1 : \beta_0 > \beta_0^*$	$T_{\beta_0^*} > t_{1-\alpha}(n-2)$
$H_0 : \beta_1 \leq \beta_1^*$	vs.	$H_1 : \beta_1 > \beta_1^*$	$T_{\beta_1^*} > t_{1-\alpha}(n-2)$

Prognose:

$$\hat{Y}_0 = \hat{\alpha} + \hat{\beta} X_0$$

Konfidenzintervall für $\boldsymbol{Y_0}$:

$$\left[\hat{Y}_0 - t_{\frac{1+\gamma}{2}}(n-2)\hat{\sigma}\sqrt{1 + \frac{1}{n} + \frac{(X_0 - \bar{X})^2}{\sum X_i^2 - n\bar{X}^2}}\ , \hat{Y}_0 + t_{\frac{1+\gamma}{2}}(n-2)\hat{\sigma}\sqrt{1 + \frac{1}{n} + \frac{(X_0 - \bar{X})^2}{\sum X_i^2 - n\bar{X}^2}}\ \right]$$

2.5.3 Multiple lineare Regression

Regressionsansatz:

$$Y_i = \beta_0 + \beta_1 X_{i1} + \beta_2 X_{i2} + \ldots + \beta_p X_{ip} + \epsilon_i\ , \quad i = 1, \ldots, n.$$

Normalverteilungsannahme:

$$\epsilon_i \sim \mathcal{N}(0, \sigma^2) \quad \Leftrightarrow \quad Y_i \sim \mathcal{N}(\beta_0 + \beta_1 X_{i1} + \ldots + \beta_p X_{ip}\ , \sigma^2), \quad i = 1, \ldots, n.$$

Gefittete Werte:

$$\hat{Y}_i = \hat{\beta}_0 + \hat{\beta}_1 X_{i1} + \hat{\beta}_2 X_{i2} + \ldots + \hat{\beta}_p X_{ip}$$

Residuen:

$$\hat{\epsilon}_i = Y_i - \hat{Y}_i \ , \quad i = 1, \ldots, n.$$

Schätzer für die Varianz $\boldsymbol{\sigma^2}$:

$$\hat{\sigma}^2 = \frac{1}{n-p-1} \sum_{i=1}^{n} \hat{\epsilon}_i^2 = \frac{1}{n-p-1} \sum_{i=1}^{n} (Y_i - \hat{Y}_i)^2$$

Verteilung der standardisierten Schätzfunktionen:

$$\frac{\hat{\beta}_j - \beta_j}{\hat{\sigma}_j} \sim t(n-p-1) \ , \quad j = 0, \ldots, p$$

($\boldsymbol{\gamma}$)-Konfidenzintervalle für $\boldsymbol{\beta_j}$:

$$\left[\hat{\beta}_j - \hat{\sigma}_j t_{\frac{1+\gamma}{2}}(n-p-1),\ \hat{\beta}_j + \hat{\sigma}_j t_{\frac{1+\gamma}{2}}(n-p-1)\right]$$

Teststatistiken:

$$T_{\beta_j^*} = \frac{\hat{\beta}_j - \beta_j^*}{\hat{\sigma}_j} \ , \quad j = 0, \ldots, p$$

Hypothesen und Ablehnbereiche:

Hypothesen	Ablehnbereich
$H_0 : \beta_j = \beta_j^*$ vs. $H_1 : \beta_j \neq \beta_j^*$	$\lvert T_{\beta_j^*} \rvert > t_{1-\frac{\alpha}{2}}(n-p-1)$
$H_0 : \beta_j \geq \beta_j^*$ vs. $H_1 : \beta_j < \beta_j^*$	$T_{\beta_j^*} < -t_{1-\alpha}(n-p-1)$
$H_0 : \beta_j \leq \beta_j^*$ vs. $H_1 : \beta_j > \beta_j^*$	$T_{\beta_j^*} > t_{1-\alpha}(n-p-1)$

SPSS-Output einer multiplen Regression:

Coefficients[a]

Model		Unstandardized Coefficients B	Std. Error	t	Sig.
1	(Constant)	$\hat{\beta}_0$	$\hat{\sigma}_0$	$T_{\beta_0^*}$	p-Wert
	X_1	$\hat{\beta}_1$	$\hat{\sigma}_1$	$T_{\beta_1^*}$	"
	X_2	$\hat{\beta}_2$	$\hat{\sigma}_2$	$T_{\beta_2^*}$	"
	$\vdots$	$\vdots$	$\vdots$	$\vdots$	"
	X_p	$\hat{\beta}_p$	$\hat{\sigma}_p$	$T_{\beta_p^*}$	"

a Dependent Variable: Y

Testprobleme:

- t ist der Wert der t-Statistik für

$$H_0 : \beta_j = 0 \quad \text{vs.} \quad H_1 : \beta_j \neq 0 \ ,$$

d.h. $T_{\beta_j^*} = \dfrac{\hat{\beta}_j}{\hat{\sigma}_j}, \quad j = 0, \dots, p$

2.5.4 Varianzanalyse

A. Verteilungstabellen

A1. Standardnormalverteilung

Tabelliert sind die Werte der Verteilungsfunktion $\Phi(z) = P(Z \leq z)$ für $z \geq 0$.
Ablesebeispiel: $\Phi(1.75) = 0.9599$
Funktionswerte für negative Argumente: $\Phi(-z) = 1 - \Phi(z)$
Die z-Quantile ergeben sich über die Umkehrfunktion.
Beispielsweise ist $z_{0.9599} = 1.75$ und $z_{0.9750} = 1.96$.

	0.00	0.01	0.02	0.03	0.04	0.05	0.06	0.07	0.08	0.09
0.0	0.5000	0.5040	0.5080	0.5120	0.5160	0.5199	0.5239	0.5279	0.5319	0.5359
0.1	0.5398	0.5438	0.5478	0.5517	0.5557	0.5596	0.5636	0.5675	0.5714	0.5753
0.2	0.5793	0.5832	0.5871	0.5910	0.5948	0.5987	0.6026	0.6064	0.6103	0.6141
0.3	0.6179	0.6217	0.6255	0.6293	0.6331	0.6368	0.6406	0.6443	0.6480	0.6517
0.4	0.6554	0.6591	0.6628	0.6664	0.6700	0.6736	0.6772	0.6808	0.6844	0.6879
0.5	0.6915	0.6950	0.6985	0.7019	0.7054	0.7088	0.7123	0.7157	0.7190	0.7224
0.6	0.7257	0.7291	0.7324	0.7357	0.7389	0.7422	0.7454	0.7486	0.7517	0.7549
0.7	0.7580	0.7611	0.7642	0.7673	0.7704	0.7734	0.7764	0.7794	0.7823	0.7852
0.8	0.7881	0.7910	0.7939	0.7967	0.7995	0.8023	0.8051	0.8078	0.8106	0.8133
0.9	0.8159	0.8186	0.8212	0.8238	0.8264	0.8289	0.8315	0.8340	0.8365	0.8389
1.0	0.8413	0.8438	0.8461	0.8485	0.8508	0.8531	0.8554	0.8577	0.8599	0.8621
1.1	0.8643	0.8665	0.8686	0.8708	0.8729	0.8749	0.8770	0.8790	0.8810	0.8830
1.2	0.8849	0.8869	0.8888	0.8907	0.8925	0.8944	0.8962	0.8980	0.8997	0.9015
1.3	0.9032	0.9049	0.9066	0.9082	0.9099	0.9115	0.9131	0.9147	0.9162	0.9177
1.4	0.9192	0.9207	0.9222	0.9236	0.9251	0.9265	0.9279	0.9292	0.9306	0.9319
1.5	0.9332	0.9345	0.9357	0.9370	0.9382	0.9394	0.9406	0.9418	0.9429	0.9441
1.6	0.9452	0.9463	0.9474	0.9484	0.9495	0.9505	0.9515	0.9525	0.9535	0.9545
1.7	0.9554	0.9564	0.9573	0.9582	0.9591	0.9599	0.9608	0.9616	0.9625	0.9633
1.8	0.9641	0.9649	0.9656	0.9664	0.9671	0.9678	0.9686	0.9693	0.9699	0.9706
1.9	0.9713	0.9719	0.9726	0.9732	0.9738	0.9744	0.9750	0.9756	0.9761	0.9767
2.0	0.9772	0.9778	0.9783	0.9788	0.9793	0.9798	0.9803	0.9808	0.9812	0.9817
2.1	0.9821	0.9826	0.9830	0.9834	0.9838	0.9842	0.9846	0.9850	0.9854	0.9857
2.2	0.9861	0.9864	0.9868	0.9871	0.9875	0.9878	0.9881	0.9884	0.9887	0.9890
2.3	0.9893	0.9896	0.9898	0.9901	0.9904	0.9906	0.9909	0.9911	0.9913	0.9916
2.4	0.9918	0.9920	0.9922	0.9925	0.9927	0.9929	0.9931	0.9932	0.9934	0.9936
2.5	0.9938	0.9940	0.9941	0.9943	0.9945	0.9946	0.9948	0.9949	0.9951	0.9952
2.6	0.9953	0.9955	0.9956	0.9957	0.9959	0.9960	0.9961	0.9962	0.9963	0.9964
2.7	0.9965	0.9966	0.9967	0.9968	0.9969	0.9970	0.9971	0.9972	0.9973	0.9974
2.8	0.9974	0.9975	0.9976	0.9977	0.9977	0.9978	0.9979	0.9979	0.9980	0.9981
2.9	0.9981	0.9982	0.9982	0.9983	0.9984	0.9984	0.9985	0.9985	0.9986	0.9986
3.0	0.9987	0.9987	0.9987	0.9988	0.9988	0.9989	0.9989	0.9989	0.9990	0.9990
3.1	0.9990	0.9991	0.9991	0.9991	0.9992	0.9992	0.9992	0.9992	0.9993	0.9993
3.2	0.9993	0.9993	0.9994	0.9994	0.9994	0.9994	0.9994	0.9995	0.9995	0.9995
3.3	0.9995	0.9995	0.9995	0.9996	0.9996	0.9996	0.9996	0.9996	0.9996	0.9997
3.4	0.9997	0.9997	0.9997	0.9997	0.9997	0.9997	0.9997	0.9997	0.9997	0.9998
3.5	0.9998	0.9998	0.9998	0.9998	0.9998	0.9998	0.9998	0.9998	0.9998	0.9998

A2. Students *t*-Verteilung

Tabelliert sind die Quantile für n Freiheitsgrade.
Für das Quantil $t_{1-\alpha}(n)$ gilt $F(t_{1-\alpha}(n)) = 1 - \alpha$.
Links vom Quantil $t_{1-\alpha}(n)$ liegt die Wahrscheinlichkeitsmasse $1 - \alpha$.
Ablesebeispiel: $t_{0.99}(20) = 2.528$
Die Quantile für $0 < 1 - \alpha < 0.5$ erhält man aus $t_\alpha(n) = -t_{1-\alpha}(n)$

Approximation für $n > 30$:

$$t_\alpha(n) \approx z_\alpha \quad (z_\alpha \text{ ist das } (\alpha)\text{-Quantil der Standardnormalverteilung})$$

n	0.6	0.8	0.9	0.95	0.975	0.99	0.995	0.999	0.9995
1	0.3249	1.3764	3.0777	6.3138	12.706	31.821	63.657	318.31	636.62
2	0.2887	1.0607	1.8856	2.9200	4.3027	6.9646	9.9248	22.327	31.599
3	0.2767	0.9785	1.6377	2.3534	3.1824	4.5407	5.8409	10.215	12.924
4	0.2707	0.9410	1.5332	2.1318	2.7764	3.7469	4.6041	7.1732	8.6103
5	0.2672	0.9195	1.4759	2.0150	2.5706	3.3649	4.0321	5.8934	6.8688
6	0.2648	0.9057	1.4398	1.9432	2.4469	3.1427	3.7074	5.2076	5.9588
7	0.2632	0.8960	1.4149	1.8946	2.3646	2.9980	3.4995	4.7853	5.4079
8	0.2619	0.8889	1.3968	1.8595	2.3060	2.8965	3.3554	4.5008	5.0413
9	0.2610	0.8834	1.3830	1.8331	2.2622	2.8214	3.2498	4.2968	4.7809
10	0.2602	0.8791	1.3722	1.8125	2.2281	2.7638	3.1693	4.1437	4.5869
11	0.2596	0.8755	1.3634	1.7959	2.2010	2.7181	3.1058	4.0247	4.4370
12	0.2590	0.8726	1.3562	1.7823	2.1788	2.6810	3.0545	3.9296	4.3178
13	0.2586	0.8702	1.3502	1.7709	2.1604	2.6503	3.0123	3.8520	4.2208
14	0.2582	0.8681	1.3450	1.7613	2.1448	2.6245	2.9768	3.7874	4.1405
15	0.2579	0.8662	1.3406	1.7531	2.1314	2.6025	2.9467	3.7328	4.0728
16	0.2576	0.8647	1.3368	1.7459	2.1199	2.5835	2.9208	3.6862	4.0150
17	0.2573	0.8633	1.3334	1.7396	2.1098	2.5669	2.8982	3.6458	3.9651
18	0.2571	0.8620	1.3304	1.7341	2.1009	2.5524	2.8784	3.6105	3.9216
19	0.2569	0.8610	1.3277	1.7291	2.0930	2.5395	2.8609	3.5794	3.8834
20	0.2567	0.8600	1.3253	1.7247	2.0860	2.5280	2.8453	3.5518	3.8495
21	0.2566	0.8591	1.3232	1.7207	2.0796	2.5176	2.8314	3.5272	3.8193
22	0.2564	0.8583	1.3212	1.7171	2.0739	2.5083	2.8188	3.5050	3.7921
23	0.2563	0.8575	1.3195	1.7139	2.0687	2.4999	2.8073	3.4850	3.7676
24	0.2562	0.8569	1.3178	1.7109	2.0639	2.4922	2.7969	3.4668	3.7454
25	0.2561	0.8562	1.3163	1.7081	2.0595	2.4851	2.7874	3.4502	3.7251
26	0.2560	0.8557	1.3150	1.7056	2.0555	2.4786	2.7787	3.4350	3.7066
27	0.2559	0.8551	1.3137	1.7033	2.0518	2.4727	2.7707	3.4210	3.6896
28	0.2558	0.8546	1.3125	1.7011	2.0484	2.4671	2.7633	3.4082	3.6739
29	0.2557	0.8542	1.3114	1.6991	2.0452	2.4620	2.7564	3.3962	3.6594
30	0.2556	0.8538	1.3104	1.6973	2.0423	2.4573	2.7500	3.3852	3.6460
∞	0.2533	0.8416	1.2816	1.6449	1.9600	2.3263	2.5758	3.0903	3.2906

A3. $\chi 2$-Verteilung

Tabelliert sind die Quantile für n Freiheitsgrade.
Für das Quantil $\chi^2_{1-\alpha,n}$ gilt $F(\chi^2_{1-\alpha,n}) = 1 - \alpha$.
Links vom Quantil $\chi^2_{1-\alpha,n}$ liegt die Wahrscheinlichkeitsmasse $1 - \alpha$.
Ablesebeispiel: $\chi^2_{0.95,10} = 18.307$

Approximation für $n > 30$:

$$\chi^2_\alpha(n) \approx \frac{1}{2}(z_\alpha + \sqrt{2n-1})^2 \quad (z_\alpha \text{ ist das } 1-\alpha\text{-Quantil der Standardnormalverteilung})$$

n	0.01	0.025	0.05	0.1	0.5	0.9	0.95	0.975	0.99
1	0.0002	0.0010	0.0039	0.0158	0.4549	2.7055	3.8415	5.0239	6.6349
2	0.0201	0.0506	0.1026	0.2107	1.3863	4.6052	5.9915	7.3778	9.2103
3	0.1148	0.2158	0.3518	0.5844	2.3660	6.2514	7.8147	9.3484	11.345
4	0.2971	0.4844	0.7107	1.0636	3.3567	7.7794	9.4877	11.143	13.277
5	0.5543	0.8312	1.1455	1.6103	4.3515	9.2364	11.070	12.833	15.086
6	0.8721	1.2373	1.6354	2.2041	5.3481	10.645	12.592	14.449	16.812
7	1.2390	1.6899	2.1674	2.8331	6.3458	12.017	14.067	16.013	18.475
8	1.6465	2.1797	2.7326	3.4895	7.3441	13.362	15.507	17.535	20.090
9	2.0879	2.7004	3.3251	4.1682	8.3428	14.684	16.919	19.023	21.666
10	2.5582	3.2470	3.9403	4.8652	9.3418	15.987	18.307	20.483	23.209
11	3.0535	3.8157	4.5748	5.5778	10.341	17.275	19.675	21.920	24.725
12	3.5706	4.4038	5.2260	6.3038	11.340	18.549	21.026	23.337	26.217
13	4.1069	5.0088	5.8919	7.0415	12.340	19.812	22.362	24.736	27.688
14	4.6604	5.6287	6.5706	7.7895	13.339	21.064	23.685	26.119	29.141
15	5.2293	6.2621	7.2609	8.5468	14.339	22.307	24.996	27.488	30.578
16	5.8122	6.9077	7.9616	9.3122	15.338	23.542	26.296	28.845	32.000
17	6.4078	7.5642	8.6718	10.085	16.338	24.769	27.587	30.191	33.409
18	7.0149	8.2307	9.3905	10.865	17.338	25.989	28.869	31.526	34.805
19	7.6327	8.9065	10.117	11.651	18.338	27.204	30.144	32.852	36.191
20	8.2604	9.5908	10.851	12.443	19.337	28.412	31.410	34.170	37.566
21	8.8972	10.283	11.591	13.240	20.337	29.615	32.671	35.479	38.932
22	9.5425	10.982	12.338	14.041	21.337	30.813	33.924	36.781	40.289
23	10.196	11.689	13.091	14.848	22.337	32.007	35.172	38.076	41.638
24	10.856	12.401	13.848	15.659	23.337	33.196	36.415	39.364	42.980
25	11.524	13.120	14.611	16.473	24.337	34.382	37.652	40.646	44.314
26	12.198	13.844	15.379	17.292	25.336	35.563	38.885	41.923	45.642
27	12.879	14.573	16.151	18.114	26.336	36.741	40.113	43.195	46.963
28	13.565	15.308	16.928	18.939	27.336	37.916	41.337	44.461	48.278
29	14.256	16.047	17.708	19.768	28.336	39.087	42.557	45.722	49.588
30	14.953	16.791	18.493	20.599	29.336	40.256	43.773	46.979	50.892