Wahrscheinlichkeitsrechnung und Statistik HS2009 [1]

Stefan Heule
Contributors: Severin Heiniger, Andrea Helfenstein, Pascal Spörri

30. Januar 2010

Inhaltsverzeichnis

I Wahrscheinlichkeitstheorie 1

1 Grundlagen 1
1.1 Axiome von Kolmogorow 1
1.2 Grundbegriffe 1

2 Diskrete Zuvallsvariablen 1
2.1 Kennzahlen 1
2.2 Diskrete uniforme Verteilung (Laplace-Modell) 2
2.3 Bernoulli-Verteilung 2
2.4 Binomial-Verteilung 2
2.5 Poisson-Verteilung 2
2.6 Geometrische Verteilung 3
2.7 Unabhängigkeit von Ereignissen 3

3 Stetige Zuvallsvariablen 3
3.1 Kennzahlen 3
3.2 Stetige uniforme Verteilung $\mathcal{U}(a, b)$ 4
3.3 Exponentialverteilung $\mathrm{Exp}(\lambda)$ 4
3.4 Normalverteilung $\mathcal{N}(\mu, \sigma^2)$ 4
3.5 Poissonprozess 4
3.6 Simulation von Zufallsvariablen 4
3.7 Unabhängigkeit von Zufallsvariablen 4

4 Mehrere Zufallsvariablen 5
4.1 Die iid Annahme 5
4.2 Funktionen von Zufallsvariablen 5
4.3 Das Gesetz der grossen Zahlen (GGZ) 6
4.4 Der zentrale Grenzwertsatz 6
4.5 Chebyshev Ungleichung 6

5 Gemeinsame und bedingte Wahrscheinlichkeiten 6
5.1 Bedingte Wahrscheinlichkeit 6
5.2 Gemeinsame und bedingte diskrete Verteilungen 7
5.3 Gemeinsame und bedingte stetige Verteilungen 7
5.4 Kovarianz und Korrelation 7
5.5 Zweidimensionale Normalverteilung 8

II Statistik 9

6 Deskriptive Statistik 9
6.1 Kennzahlen 9
6.2 Graphische Darstellung 9

7 Statistischer Test 10
7.1 Allgemeines Vorgehen 10
7.2 Vertrauensintervall 11
7.3 z-Test 11
7.4 t-Test 12

9 Punktschätzungen 12
9.1 Binomial- und Poissonverteilung 12
9.2 Normalverteilung 12
9.3 Allgemeine Methoden 12

10 Vergleich zweier Stichproben 13
10.1 Gepaarte Vergleiche 13
10.2 Ungepaarte Vergleiche - Zwei Stichproben-Tests 13

III Appendix **15**

A Tabellen und Tafeln **15**
A.1 Integral- und Differentialrechnung . . . 15
A.2 Verschiedenes . . . 16
A.3 Regeln für Integrale . . . 16

Teil I
Wahrscheinlichkeitstheorie

1 Grundlagen

1.1 Axiome von Kolmogorow

1. $P(A) \geq 0$: Wahrscheinlichkeiten sind stets nicht-negativ.
2. $P(\Omega) = 1$: Das sichere Ereignis Ω hat Wahrscheinlichkeit 1.
3. $P(A \cup B) = P(A) + P(B)$ für alle sich gegenseitig ausschliessenden Ereignisse A und B (d.h. $A \cap B = \emptyset$).

1.1.1 Implikationen aus den Axiomen von Kolmogorow

1. $P(\emptyset) = 0$
2. $P(A^c) = 1 - P(A)$
3. $P(A \cup B) = P(A) + P(B) - P(A \cap B)$
4. $P(\bigcup_{i=1}^n A_i) = \sum_{i=1}^n P(A_i)$ falls $A_i \cap A_j = \emptyset \quad \forall i, j$

1.2 Grundbegriffe

$$\begin{aligned}
\text{Zufallsvariable } X: \quad & X : \Omega \to \mathbb{R} \text{ mit Ereignisraum } \Omega \\
\text{Verteilungsfunktion von } X: \quad & F_X : \mathbb{R} \to [0,1], \quad t \mapsto F_X(t) := P(X \leq t) = P(\{\omega \in \Omega | X(\omega) \leq t\})
\end{aligned}$$

2 Diskrete Zuvallsvariablen

$$\begin{aligned}
\text{Diskrete Zufallsvariable:} \quad & \text{Wertebereich } \mathcal{W}(X) \text{ ist abzählbar} \\
\text{Gewichtsfunktion:} \quad & p_X(x_k) := P[X = x_k] = P[\{\omega \in \Omega | X(\omega) = x_k\}] \text{ für } k = 1, 2, \ldots \\
\text{Verteilungsfunktion (Treppenfunktion):} \quad & F_X(t) = P[X \leq t] = \sum_{k \text{ mit } x_k \leq t} p_X(x_k)
\end{aligned}$$

2.1 Kennzahlen

$$\begin{aligned}
\text{Erwartungswert:} \quad & \mu = \mathrm{E}(X) = \sum_{x_i \in \mathcal{W}(X)} x_i \cdot P(X = x_i) = \sum_{x_i \in \mathcal{W}(X)} x_i \cdot p(x_i) \\
& \mathrm{E}(g(X)) = \sum_{x_i \in \mathcal{W}(X)} g(x_i) \cdot P(X = x_i) \\
& \mathrm{E}(aX + b) = a\,\mathrm{E}(X) + b
\end{aligned}$$

$$\begin{aligned}
\text{Varianz:} \quad & \mathrm{Var}(X) = \sum_{x_i \in \mathcal{W}(X)} (x_i - \mu)^2 \cdot P(X = x_i) \\
& \mathrm{Var}(X) = \mathrm{E}((X - \mu)^2) = \mathrm{E}(X^2) - \mathrm{E}(X)^2 \\
& \mathrm{Var}(aX + b) = a^2 \cdot \mathrm{Var}(X)
\end{aligned}$$

$$\begin{aligned}
\text{Standardabweichung:} \quad & \sigma = \sqrt{\mathrm{Var}(X)} \\
& \sigma_{aX+b} = |a| \cdot \sigma_X
\end{aligned}$$

2.2 Diskrete uniforme Verteilung (Laplace-Modell)

n gleich wahrscheinliche Ereignisse, wie z.B. Wurf einer Zahl auf einem Würfel.

Wahrscheinlichkeitsfunktion	$\begin{cases} \frac{1}{n} & \text{für } a \leq x \leq b \\ 0 & \text{sonst} \end{cases}$
Kumulative Funktion	$\frac{\lfloor x \rfloor - a + 1}{n} \quad \text{für } a \leq x \leq b$
Parameter	$a, b \in \mathbb{Z}, n = b - a + 1$
Support	$x \in \{a, a+1, \ldots, b-1, b\}$

Erwartungswert	$\frac{a+b}{2}$
Varianz	$\frac{n^2-1}{12}$
Median	$\frac{a+b}{2}$
Modus	jeder Wert in $\{a, a+1, \ldots, b-1, b\}$

2.3 Bernoulli-Verteilung

Ein einzelnes Ereignis tritt ein, oder nicht, mit fester Wahrscheinlichkeit p.

Wahrscheinlichkeitsfunktion	$\begin{cases} 1-p & \text{für } x = 0 \\ p & \text{für } x = 1 \end{cases}$
Kumulative Funktion	$1 - p \quad \text{für } 0 \leq x < 1$
Parameter	$p \in [0, 1]$
Support	$x \in \{0, 1\}$

Erwartungswert	p
Varianz	$p \cdot (1-p)$
Median	N/A
Modus	$0 \ (p \leq 0.5), 1 \ (p \geq 0.5)$

2.4 Binomial-Verteilung

Anzahl "Erfolge" bei n unabhängigen Experimenten mit zwei Ausgängen und festem Erfolgsparameter p.

Wahrscheinlichkeitsfunktion	$\binom{n}{x} \cdot p^x \cdot (1-p)^{n-x}$
Kumulative Funktion	$\sum_{i=0}^{\lfloor x \rfloor} \binom{n}{i} \cdot p^i \cdot (1-p)^{n-i}$
Parameter	$n \in \mathbb{N}_0, p \in [0, 1]$
Support	$x \in \{0, \ldots, n\}$

Erwartungswert	np
Varianz	$np(1-p)$
Median	$\lfloor np \rfloor$
Modus	$\lfloor (n+1)p \rfloor$ oder $\lfloor (n+1)p \rfloor - 1$

2.5 Poisson-Verteilung

Verteilungsmodell für (diskrete) Zähldaten (nicht nach oben beschränkt). Damit können zum Beispiel die Anzahl der Ausfälle einer Komponente in einem Interval t modelliert werden.

Wahrscheinlichkeitsfunktion	$e^{-\lambda} \cdot \frac{\lambda^x}{x!}$
Kumulative Funktion	$e^{-\lambda} \sum_{i=0}^{x} \frac{\lambda^i}{i!}$
Parameter	$\lambda \in [0, \infty)$
Support	$x \in \mathbb{N}_0$

Erwartungswert	λ
Varianz	λ
Median	$\approx \lfloor \lambda + \frac{1}{3} - \frac{0.02}{\lambda} \rfloor$
Modus	$\lfloor \lambda \rfloor$ oder $\lambda - 1$ wenn $\lambda \in \mathbb{Z}$

2.5.1 Poissonapproximation einer Binomialverteilung

Sei $X \sim \text{Poisson}(\lambda)$ und $Y \sim \text{Binom}(n, p)$. Für $n \to \infty$ und $p \to 0$ so dass $n \cdot p = \lambda$ gilt

$$P(Y = k) \to P(X = k) \ \ (n \to \infty) \quad k = 1, 2, \ldots$$

Eine Faustregel besagt, dass diese Näherung brauchbar ist, wenn $n \geq 50$ und $p \leq 0.05$.

2.6 Geometrische Verteilung

Diskrete Wartezeiten (nicht nach oben beschränkt), Anzahl Wiederholungen bis ein Ereignis eintrifft.

Wahrscheinlichkeitsfunktion	$p \cdot (1-p)^{x-1}$
Kumulative Funktion	$1-(1-p)^x$
Parameter	$p \in [0,1]$
Support	$x \in \mathbb{N} = \{1, 2, \ldots\}$
Erwartungswert	$\frac{1}{p}$
Varianz	$\frac{1-p}{p^2}$
Median	$\lceil \frac{-\log 2}{\log(1-p)} \rceil$
Modus	1

2.7 Unabhängigkeit von Ereignissen

Zwei Ereignisse heissen stochastisch unabhängig, wenn $P(A \cap B) = P(A) \cdot P(B)$ gilt. Zwischen solchen Ereignissen besteht kein kausaler Zusammenhang.

3 Stetige Zuvallsvariablen

Eine Zufallsvariable X mit Verteilungsfunktion $F_X(t) = P(X \leq t)$ heisst stetig mit Dichtefunktion $f_X : \mathbb{R} \to [0, \infty)$ falls gilt

$$F_X(t) = \int_{-\infty}^{t} f_X(s) \mathrm{ds} \quad \text{für alle t} \in \mathbb{R}$$

Anschaulicher lässt sich die Dichtefunktion beschreiben durch folgenden Grenzwert, wonach $f_X(t) = F'_X(t)$ gilt.

$$f_X(t) = \lim_{\varepsilon \to 0} \frac{P(t < X \leq t + \varepsilon)}{\varepsilon}$$

Sie erfüllt die Eigenschaft $\int\limits_{-\infty}^{\infty} f_X(t) \mathrm{d}t = 1$.

3.1 Kennzahlen

$$\begin{aligned}
\text{Erwartungswert:} \quad & \mu = \mathrm{E}(X) = \int_{-\infty}^{\infty} x \cdot f_X(x) \, \mathrm{d}x \\
& \mathrm{E}(g(X)) = \int_{-\infty}^{\infty} g(x) \cdot f_X(x) \, \mathrm{d}x \\
\text{Varianz:} \quad & \mathrm{Var}(X) = \mathrm{E}\left(X^2\right) - \mathrm{E}\left(X\right)^2 = \int_{-\infty}^{\infty} (x - \mathrm{E}(X))^2 \cdot f_X(x) \, \mathrm{d}x \\
& \mathrm{Var}(aX + b) = a^2 \, \mathrm{Var}(X) \\
\text{Standardabweichung:} \quad & \sigma = \sqrt{\mathrm{Var}(X)} \\
& \sigma_{aX+b} = |a| \cdot \sigma_X
\end{aligned}$$

$$\begin{aligned}
\alpha\text{-Quantile:} \quad & \text{Sei} \quad 0 < \alpha < 1 \\
& P(X \leq q(a)) = \alpha \qquad \Longrightarrow \qquad F_X(q(a)) = \alpha \\
& \qquad\qquad\qquad\qquad\quad \Longrightarrow \qquad q(a) = F_X^{-1}(a) \\
\text{Median:} \quad & \tfrac{1}{2}\text{-Quantil} \\
\text{Lineare Transformation:} \quad & Y = aX + b \quad \Longrightarrow \quad f_Y(x) = \tfrac{1}{|a|} f_X\left(\tfrac{x-b}{|a|}\right)
\end{aligned}$$

3.2 Stetige uniforme Verteilung $\mathcal{U}(a,b)$

Rundungsfehler, Simulation (Zufallsvariablen), Formalisierung der völligen Ignoranz (alle Ereignisse sind gleich wahrscheinlich).

Dichtefunktion	$\begin{cases} \frac{1}{b-a} & \text{für } x \in [a,b] \\ 0 & \text{sonst} \end{cases}$	Erwartungswert	$\frac{1}{2}(a+b)$
Verteilungsfunktion	$\frac{x-a}{b-a}$ für $x \in [a,b]$	Varianz	$\frac{1}{12}(b-a)^2$
Parameter	$-\infty < a < b < \infty$	Median	$\frac{1}{2}(a+b)$
Support	$x \in [a,b]$	Modus	jeder Wert in $[a,b]$

3.3 Exponentialverteilung $\mathrm{Exp}(\lambda)$

Stetige Wartezeiten. Stetige Version der geometrische Verteilung.

Dichtefunktion	$\begin{cases} \lambda \cdot e^{-\lambda x} & \text{für } x \geq 0 \\ 0 & \text{sonst} \end{cases}$	Erwartungswert	$\frac{1}{\lambda}$
Verteilungsfunktion	$1 - e^{-\lambda x}$ für $x > 0$	Varianz	$\frac{1}{\lambda^2}$
Parameter	$\lambda > 0$	Median	$\frac{\ln 2}{\lambda}$
Support	$x \in [0,\infty)$	Modus	0

Die Exponentialverteilung ist "gedächnislos", im Sinne von $P(T > t+s \mid T > t) = P(T > s)$. Beweis:

$$P(T > t+s \mid T > t) = \frac{P(T > t+s, X > t)}{P(X > t)} = \frac{P(T > t+s)}{P(X > t)} = \frac{e^{-\lambda(t+s)}}{e^{-\lambda t}} = e^{-\lambda s} = P((X > s)$$

3.4 Normalverteilung $\mathcal{N}(\mu,\sigma^2)$

Streuung von Messwerten um ihren Mittelwert.

Dichtefunktion	$\frac{1}{\sqrt{2\pi\sigma^2}} \exp\left(-\frac{(x-\mu)^2}{2\sigma^2}\right)$	Erwartungswert	μ
Verteilungsfunktion	$\Phi\left(\frac{x-\mu}{\sigma}\right)$	Varianz	σ^2
Parameter	$\mu \in \mathbb{R}, \sigma^2 \geq 0$	Median	μ
Support	$x \in \mathbb{R}\ (\sigma^2 > 0), x = \mu\ (\sigma^2 = 0)$	Modus	μ

Spezialfall Standardnormalverteilung, mit $\mu = 0, \sigma^2 = 1$. Dichtefunktion: $\varphi(x)$. Kumulative Verteilungsfunktion $\Phi(x)$ mit $\Phi(-x) = 1 - \Phi(x)$ und $\Phi^{-1}(x) = -\Phi^{-1}(1-x)$.

3.5 Poissonprozess

Der Poisson-Prozess ist ein Erneuerungsprozess, dessen Zuwächse unabhängig und poissonverteilt mit Parameter λ sind. λ wird auch Intensität oder Rate genannt, da pro Zeiteinheit genau λ Sprünge erwartet werden. Die Höhe jedes Sprunges ist eins, die Zeiten zwischen den Sprüngen sind exponentialverteilt mit Parameter λ. Der Poissonprozess ist also ein diskreter Prozess in stetiger Zeit.
Die Anzahl Ereignisse auf dem Intervall $(s, s+t]$ ist Poisson-verteilt mit Parameter $\lambda \cdot t$.

3.6 Simulation von Zufallsvariablen

Sei $U \sim Uniform([0,1])$ und $F : \mathbb{R} \to [0,1]$ bijektiv. Dann hat die Zufallsvariable $X = F^{-1}(U)$ die kumulative Verteilungsfunktion F. Beweis:

$$F_X(x) = P(X \leq x) = P(F^{-1}(U) \leq x) = P(U \leq F(x)) = F(x)$$

3.7 Unabhängigkeit von Zufallsvariablen

Zwei Zufallsvariablen X_1 und X_2 heissen unabhängig, falls für alle $A_1, A_2 \subseteq \mathbb{R}$ gilt

$$P(X_1 \in A_1, X_2 \in A_2) = P(X_1 \in A_1) \cdot P(X_2 \in A_2)$$

4 Mehrere Zufallsvariablen

Für beliebige Zufallsvariablen gilt

$$\begin{aligned}
\mathrm{E}\left(\sum_{i=1}^{n} X_i\right) &= \sum_{i=1}^{n} \mathrm{E}(X_i) \\
\mathrm{Var}(X) &= E(X^2) - E(X)^2 \\
\mathrm{Var}(X) &= E((X-\mu)^2) \\
\mathrm{Var}(aX+bY+c) &= a^2\,\mathrm{Var}(X) + b^2\,\mathrm{Var}(Y) + 2ab\,\mathrm{Cov}(X,Y)
\end{aligned}$$

Für unkorrelierte Zufallsvariablen X_i gilt

$$\begin{aligned}
\mathrm{Var}\left(\sum_{i=1}^{n} X_i\right) &= \sum_{i=1}^{n} \mathrm{Var}(X_i) \\
\mathrm{Var}(\overline{X}_n) &= \frac{\sigma^2}{n}
\end{aligned}$$

Sind X_i, X und Y unabhängig, so gilt

$$\begin{aligned}
\mathrm{E}(XY) &= \mathrm{E}(X)\,\mathrm{E}(Y) \\
\mathrm{E}\left(\prod X_i\right) &= \prod \mathrm{E}(X_i)
\end{aligned}$$

4.1 Die iid Annahme

Zufallsvariablen $X_1, X_2, \ldots, X_n$ sind *iid* (unabhängig und gleichverteilt) falls

- $X_1, X_2, \ldots, X_n$ sind unabhängig
- $X_i \sim F \quad \forall i = 1, \ldots, n$

4.2 Funktionen von Zufallsvariablen

Seien $X_1, \ldots, X_n \overset{iid}{\sim} F$. Dann ist $Y = g(X_1, \ldots, X_n)$ mit $g : \mathbb{R}^n \longrightarrow \mathbb{R}$ eine neue Zufallsvariable. Wichtige Vertreter sind die Summe von Zufallsvariablen und die relative Häufigkeit.

$$S_n = X_1 + \ldots + X_n \qquad \overline{X}_n = \frac{1}{n} S_n$$

Die Verteilungen von S_n und $\overline{X}_n$ sind im allgemeinen aber schwierig zu bestimmen.

4.2.1 Bekannte Verteilungen

$$\begin{aligned}
X_i &\overset{iid}{\sim} \mathrm{Poisson}(\lambda_i) &&\Rightarrow S_n \sim \mathrm{Poisson}(\textstyle\sum \lambda_i) \\
X_i &\overset{iid}{\sim} \mathrm{Bernoulli}(p) &&\Rightarrow S_n \sim \mathrm{Binomial}(n,p) \\
X_i &\overset{iid}{\sim} \mathcal{N}(\mu, \sigma^2) &&\Rightarrow S_n \sim \mathcal{N}(n\cdot\mu, n\cdot\sigma^2)
\end{aligned}$$

4.2.2 Erwartungswerte und Varianzen

$$\begin{aligned}
\mathrm{E}(S_n) &= n \cdot \mathrm{E}(X_1) & \mathrm{E}(\overline{X}_n) &= \mathrm{E}(X_1) \\
\mathrm{Var}(S_n) &= n \cdot \mathrm{Var}(X_1) & \mathrm{Var}(\overline{X}_n) &= \tfrac{1}{n}\,\mathrm{Var}(X_1) \\
\sigma_{S_n} &= \sqrt{n} \cdot \sigma_{X_1} & \sigma_{\overline{X}_n} &= \tfrac{1}{\sqrt{n}} \cdot \sigma_{X_1}
\end{aligned}$$

4.2.3 Minimum mehrerer Zufallsvariablen

Seien $X_1, \ldots, X_n$ unabhängig verteilt mit Verteilungsfunktion F_{X_i}, so gilt für das Minimum $Y = \min\{X_1, \ldots, X_n\}$

$$\begin{aligned}
F_Y(t) &= P(Y \le t) = 1 - P(Y > t) = 1 - P(\min\{X_1, \ldots, X_n\} > t) \\
&= 1 - P(X_1 > t, \ldots X_n > t) = 1 - \prod_{i=1}^{n} P(X_i > t) = 1 - \prod_{i=1}^{n} (1 - F_{X_i}(t))
\end{aligned}$$

Gilt weiter $F_{X_i} = 1 - e^{-\lambda t}$, so folgt $\quad F_Y(t) = 1 - e^{-n\lambda t}$.

4.3 Das Gesetz der grossen Zahlen (GGZ)

Seien $X_1, \ldots, X_n$ iid mit Erwartungswert $\mu = \mathrm{E}(X_1)$. Dann gilt: $\overline{X}_n \to \mu \quad (n \to \infty)$

4.3.1 Monte-Carlo-Integration

Seien $x_1, \ldots, x_n$ (simulierte) Realisierungen von $X_1, \ldots, X_n \sim \int f(x)\mathrm{d}x$ für eine Dichtefunktion f. Dann gilt

$$I = \int_{-\infty}^{\infty} g(x)f(x)\,\mathrm{d}x = \mathrm{E}(g(X)) \approx \frac{1}{n}\sum_{i=1}^{n} g(x_i)$$

4.4 Der zentrale Grenzwertsatz

Seien $X_1, \ldots, X_n$ iid mit Erwartungswert $\mu = \mathrm{E}(X_i)$ und Varianz $\mathrm{Var}(X_i) = \sigma^2$. Dann gilt für $n \to \infty$

$$\overline{X}_n \sim \mathcal{N}(\mu, \tfrac{1}{n}\sigma^2) \quad \text{und} \quad S_n \sim \mathcal{N}(n\mu, n\sigma^2)$$

bzw.

$$P(S_n \leq x) = P\left(\frac{S_n - n\mu}{\sigma\sqrt{n}} \leq \frac{x - n\mu}{\sigma\sqrt{n}}\right) \approx \Phi\left(\frac{x - n\mu}{\sigma\sqrt{n}}\right)$$

4.5 Chebyshev Ungleichung

Die Chebyshev-Ungleichung liefert eine Abschätzung von Wahrscheinlichkeiten, auch wenn die genaue Verteilungsfunktion nicht bekannt ist.
Seien $X_1, \ldots, X_n$ iid mit Erwartungswert $\mu = \mathrm{E}(X_1)$ und Varianz $\mathrm{Var}(X_1) = \sigma^2$. Dann gilt

$$P(\,|\overline{X}_n - \mu| > c) \quad \leq \quad \frac{\sigma^2}{n \cdot c^2}$$

5 Gemeinsame und bedingte Wahrscheinlichkeiten

Hier werden zweistufige Zufallsexperimente betrachtet. Die erste Stufe ist das Eintreten bzw. nicht Eintreten eines Ereignis B und die zweite Stufe das Eintreten bzw. nicht Eintreten eines Ereignis A.

5.1 Bedingte Wahrscheinlichkeit

Die bedingte Wahrscheinlichkeit von Ereignis A gegeben B ist

$$P(A \mid B) = \frac{P(A \cap B)}{P(B)} \quad \text{falls } P(B) \neq 0 \qquad\qquad P(A \cap B) = P(A \mid B) \cdot P(B)$$

5.1.1 Satz der totalen Wahrscheinlichkeit und Satz von Bayes

Seien $B_1, \ldots, B_k$ die Ereignisse der 1. Stufe mit

$$B_1 \cup \ldots \cup B_k = \Omega \qquad\qquad B_i \cap B_j = \emptyset \quad (i \neq j)$$

dann gilt für beliebige Ereignisse A der 2. Stufe

$$P(A) = \sum_{j=1}^{k} P(A \mid B_j) \cdot P(B_j) \qquad \overset{k=2}{\Longrightarrow} \qquad P(A) = P(A \mid B) \cdot P(B) + P(A \mid B^{\complement}) \cdot P(B^{\complement})$$

und

$$P(B_i \mid A) = \frac{P(A \mid B_i) \cdot P(B_i)}{\sum\limits_{j=1}^{k} P(A \mid B_j) \cdot P(B_j)} = \frac{P(A \mid B_i) \cdot P(B_i)}{P(A)}$$

5.1.2 Rechenregeln

Komplement	$P(A^\complement) = 1 - P(A)$	Multiplikation	$P(A \cap B) = P(A \mid B) \cdot P(B)$
	$P(A^\complement \mid S) = 1 - P(A \mid S)$	A, B unabhängig	$P(A \cap B) = P(A) \cdot P(B)$
Addition	$P(A \cup B) = P(A) + P(B) - P(A \cap B)$		$P(A \cap B \mid S) = P(A \mid S) \cdot P(B \mid S)$

5.2 Gemeinsame und bedingte diskrete Verteilungen

Die Gemeinsame Verteilung von X und Y ist vollständig beschrieben durch $P(X = x, Y = y)$ für alle (x, y).

$$\text{Randverteilung von } X: \quad P(X = x) = \sum_{\text{alle } y} P(X = x, Y = y)$$

$$\text{Bedingte Wahrscheinlichkeit:} \quad P(X = x \mid Y = y) = \frac{P(X = x, Y = y)}{P(Y = y)} \quad \text{falls } P(Y = y) \neq 0$$

5.3 Gemeinsame und bedingte stetige Verteilungen

Gemeinsame stetige Verteilungen können wie im diskreten Fall durch die kumulative Verteilungsfunktion beschrieben werden: $F_{XY}(x, y) = P(X \leq x, Y \leq y)$.

$$\text{Gemeinsame Dichte:} \quad f_{XY}(x, y) = \frac{\partial^2}{\partial x \partial y} F_{XY}(x, y)$$

$$\text{Randdichte von } X: \quad f_X(x) = \int_{-\infty}^{\infty} f_{XY}(x, y) \, \mathrm{d}y$$

$$\text{Bedingte Dichte:} \quad f_{Y|X}(y \mid X = x) = \frac{f_{XY}(x, y)}{f_X(x)}$$

$$P(Y \in A \mid X = x) = \int_A f_{Y|X}(y \mid X = x) \, \mathrm{d}y$$

5.3.1 Unabhängige Zufallsvariablen

Die Zufallsvariablen X und Y heissen unabhängig, falls für alle x, y gilt: $f_{XY}(x, y) = f_X(x) \cdot f_Y(y)$

5.3.2 Erwartungswert bei mehreren Zufallsvariablen

Seien X und Y zwei stetige Zufallsvariablen und $g(x, y) : \mathbb{R}^2 \to \mathbb{R}$ eine Funktion. Dann

$$\mathrm{E}(g(X, Y)) = \iint g(x, y) \cdot f_{XY}(x, y) \, \mathrm{d}x \, \mathrm{d}y$$

5.4 Kovarianz und Korrelation

$$\text{Kovarianz:} \quad \mathrm{Cov}(X, Y) = \mathrm{E}\left((X - \mathrm{E}(X))\,(Y - E(Y))\right)$$

$$\mathrm{Cov}(X, X) = Var(X)$$

$$\text{Korrelation:} \quad \mathrm{Corr}(X, Y) = \rho_{XY} = \frac{\mathrm{Cov}(X, Y)}{\sqrt{\mathrm{Var}(X)} \cdot \sqrt{\mathrm{Var}(Y)}} = \frac{\mathrm{Cov}(X, Y)}{\sigma_X \cdot \sigma_Y}$$

$\rho_{XY} \in [-1, 1]$ als dimensionsloses Mass des linearen Zusammenhanges

$\rho_{XY} = \pm 1 \quad \Longleftrightarrow \quad Y = a \pm bX \quad \text{mit } b > 0, a \in \mathbb{R}$

Rechenregeln:

$\mathrm{Cov}(X, Y) = \mathrm{E}(X \cdot Y) - \mathrm{E}(X) \cdot \mathrm{E}(Y)$

$\mathrm{Cov}(\cdot, \cdot)$ ist bilinear: $\mathrm{Cov}(a + bX, c + dY) = b \cdot d \cdot \mathrm{Cov}(X, Y)$

$\mathrm{Cov}(X + Y, Z) = \mathrm{Cov}(X, Z) + \mathrm{Cov}(Y, Z)$

$\mathrm{Var}(X + Y) = \mathrm{Var}(X) + \mathrm{Var}(Y) + 2 \cdot \mathrm{Cov}(X, Y)$

Cov ist symmetrisch

Sind die Zufallsvariablen X und Y unabhängig, so gilt $\mathrm{Cov}(X, Y) = \mathrm{Corr}(X, Y) = 0$. Die Umkehrung gilt jedoch im Allgemeinen nicht.

5.4.1 Lineare Prognose

Die beste lineare Prognose $\widehat{Y}$ von Y durch X im Sinn der Minimierung des Progonosefehlers $\mathrm{E}(Y - \widehat{Y})^2)$ ist gegeben als

$$\widehat{Y} = \mu_Y + \frac{\mathrm{Cov}(X,Y)}{\mathrm{Var}(X)} \cdot (X - \mu_X) \qquad \mathrm{E}((Y - \widehat{Y})^2) = (1 - \rho_{XY}^2) \cdot \mathrm{Var}(Y) = \mathrm{Var}(Y) - \frac{\mathrm{Cov}(X,Y)^2}{\mathrm{Var}(X)}$$

5.5 Zweidimensionale Normalverteilung

Die wichtigste zweidimensionale Verteilung ist die Normalverteilung mit Erwartungswert $\boldsymbol{\mu} = \begin{pmatrix} \mu_1 \\ \mu_2 \end{pmatrix}$ und Kovarianzmatrix $\boldsymbol{\Sigma} \in \mathbb{R}^{2,2}$ mit $\Sigma_{11} = \mathrm{Var}(X)$, $\Sigma_{22} = \mathrm{Var}(Y)$ und $\Sigma_{12} = \Sigma_{21} = \mathrm{Cov}(X,Y)$. Sie hat die Dichte

$$f_{XY}(x,y) = \frac{1}{2\pi\sqrt{\det \boldsymbol{\Sigma}}} \cdot \exp\left(-\tfrac{1}{2} \begin{pmatrix} x - \mu_X & y - \mu_Y \end{pmatrix} \cdot \boldsymbol{\Sigma} \cdot \begin{pmatrix} x - \mu_X \\ y - \mu_Y \end{pmatrix} \right)$$

Für $\boldsymbol{X} = \begin{pmatrix} X_1 \\ X_2 \end{pmatrix} \sim \mathcal{N}_2(\boldsymbol{\mu}, \boldsymbol{\Sigma})$ gilt: $\qquad X_1, X_2$ unabhängig $\quad \Longleftrightarrow \quad \rho_{X_1 X_2} = 0$

Teil II

Statistik

6 Deskriptive Statistik

6.1 Kennzahlen

n bezeichnet die Stichprobengrösse, und $x_1, x_2, \ldots x_n$ sind die gemessenen Werte. Sind diese aufsteigend sortiert, so werden sie folgendermassen bezeichnet:

$$x_{(1)} \leq x_{(2)} \leq x_{(3)} \leq \ldots \leq x_{(n)}$$

Arithmetisches Mittel	$\overline{x} = \frac{1}{n}\sum_{i=1}^{n} x_i$
empirische Varianz	$s^2 = \frac{1}{n-1}\sum_{i=1}^{n}(x_i - \overline{x})^2$
empirische Standardabweichung	$s = \sqrt{s^2}$
empirisches α-Quantil	$x_{(k)}$ mit minimalem $k \in \mathbb{N}$, sodass $k > \alpha \cdot n$
	ist $\alpha \cdot n$ eine Ganzzahl, so nimmt man $\frac{1}{2}(x_{(\alpha n)} + x_{(\alpha n+1)})$
	$\alpha \cdot 100\%$ sind approximativ $\leq$ dem α-Quantil
	$(1-\alpha) \cdot 100\%$ sind approximativ $>$ dem α-Quantil
unteres Quartil	0.25-Quantil
empirischer Median	0.5-Quantil
oberes Quartil	0.75-Quantil
Quartilsdifferenz	oberes Quartil − unteres Quartil
Rechtsschiefe (positive Schiefe)	Werte, die kleiner sind als der Mittelwert, sind häufiger zu beobachten,
	so dass sich der Gipfel links vom Mittelwert befindet.
	Der rechte Teil des Graphs ist somit flacher als der linke.
Linksschiefe (negative Schiefe)	Definition symmetrisch.

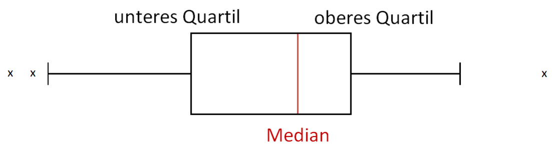
6.2 Graphische Darstellung

6.2.1 Histogramm

Für das Histogramm werden Klassen $(c_{k-1}, c_k]$ gebildet und berechnet die absoluten Häufigkeiten h_k in diesem Intervall. Nun werden Balken mit einer Höhe proportional zu $h_k/(c_k - c_{k-1})$ gezeichnet.

6.2.2 Boxplot

Beim Boxplot gibt es eine Rechteck, welches die mittleren 50% der Daten enthält. Zusätzlich wird mit einer Trennlinie der Median eingezeichnet. Nach oben und unten gibt es zusätzlich Linien bis zum kleinsten bzw. grössten "normalen" Wert. Dazu definiert man willkürlich, dass ein Wert normal ist, falls er höchstens 1.5 mal die Quartilsdifferenz von einem der Quartile entfernt ist. Zuletzt werden Ausreisser noch durch Sterne markiert.



Im Boxplot erkennt man viel besser die Lage und Streuung der Daten, nicht aber die "Form" der Verteilung. Ausserdem lassen sich sehr gut mehrere Datensätze vergleichen.

6.2.3 Empirische, kumulative Verteilungsfunktion

Die empirische Verteilungsfuktion ist eine Treppenfunktion, die an der Stelle $x_{(i)}$ von $(i-1)/n$ nach i/n springt.

$$\widehat{F}(x) = \frac{1}{n} \cdot |\{i \mid x_i \leq x\}|$$

6.2.4 QQ-Plot

QQ-Plots ("Quantil-Quantil-Plot") dienen dazu die Abweichung der Daten von einer gewählten Modell-Verteilung F grafisch zu überprüfen. Dazu betrachtet man die theoretischen und empirischen α_k-Quantile mit $\alpha_k = (k - 0.5)/n$.
Die empirischen α_k-Quantile entsprechen Dabei gerade den geordneten Daten $x_{(k)}$. Man trägt also die Punkte $F^{-1}(\alpha_k)$ (die theoretischen Quantile der kumulativen Verteilungsfunktion F) gegen $x_{(k)}$ (die empirischen Quantile) ab. Falls nun die wahre Verteilung $\tilde{F}$ mit der gewählten Verteilung F übereinstimmt, so liefert der QQ-Plot approximativ eine Gerade durch den Nullpunkt mit Steigung 1.

$$\begin{array}{ll} \text{x-Achse:} & \text{theoretische Quantile von } F\text{, also } F^{-1}(\alpha_k) \\ \text{y-Achse:} & \text{empirische Quantile, also } x_{(k)} \end{array}$$

6.2.5 Normal-Plot

Normal-Plots sind ein Spezialfall des QQ-Plots, wobei als Modell-Verteilungsfunktion eine Standardnormalverteilung verwendet wird, also $F \sim \mathcal{N}(0,1)$. Zusätzlich zu den normalen Eigenschaften eines QQ-Plots gilt hier weiter folgende nützliche Eigenschaft: Sind die Daten Normalverteilt (nicht unbedingt standardnormal), so zeigt der Normal-Plot stets eine Gerade. Diese hat jedoch im Allgemeinen Steigung σ und den y-Achsenabschnitt μ.

7 Statistischer Test

7.1 Allgemeines Vorgehen

Beobachtet wird $X = x$, Modell: $X \sim F_\theta$ mit unbekanntem Parameter θ.

1. Modell: statistisches Modell für die Daten (Verteilung)
2. Nullhypothese: H_0: $\theta = \theta_0$
3. Alternativhypothese:
$$H_A := \begin{cases} \theta \neq \theta_0 & \text{2-seitig} \\ \theta > \theta_0 & \text{1-seitig von oben} \\ \theta < \theta_0 & \text{1-seitig von unten} \end{cases}$$
4. Bestimme die Teststatistik T (z.B. $T = S_n$ oder so, dass T eine normierte Variante von X darstellt)
5. Verteilung der Teststatistik unter der Nullhypothese
6. Wahl des Signifikanzniveaus α

a)
   7. Konstruiere Verwerfungsbereich K so, dass $P_{H_0}(X \in K) \leq \alpha$
   8. beobachteter Wert bezüglich der Teststatistik
   9. Entscheidung: verwerfe H_0 falls $x \in K$, sonst belasse H_0

b)
   7. Berechne den P-Wert. Der P-Wert ist die Wahrscheinlichkeit, dass unter der Nullhypothese H_0 ein zufälliger Versuch mindestens so "extrem" ausfällt, wie der beobachtete Wert t:
   $$p_{\text{links}} = P_{H_0}(T \leq t) \qquad p_{\text{rechts}} = P_{H_0}(T \geq t) \qquad p_{\text{beidseitig}} = 2 \cdot \min\{p_{\text{links}}, p_{\text{rechts}}\}$$
   Anders ausgedrückt ist der P-Wert das kleinste Signifikanzniveau, bei dem bei gegebener Beobachtung t H_0 gerade noch verworfen werden kann.
   8. Entscheidung: verwerfe H_0 falls P-Wert $\leq \alpha$, sonst belasse H_0

7.1.1 Fehler 1. Art

Wahrscheinlichkeit, dass der Test H_0 verwirft, obschon H_0 stimmt.

$$P(\text{Fehler 1. Art}) = P_{H_0}(X \in K) \leq \alpha$$

7.1.2 Fehler 2. Art

Wahrscheinlichkeit, dass der Test H_0 belässt, obschon H_A stimmt.

$$\beta(p) = P(\text{Fehler 2. Art}) = P_{H_A}(X \notin K)$$

7.1.3 Macht/Power

Die Macht eines Tests ist definiert als

$$1 - \beta(p) = P(H_0 \text{ wird verworfen, und } H_A \text{ korrekterweise angenommen})$$

7.2 Vertrauensintervall

Das Vertrauens- oder Konfidenzintervall I ist die Menge jener Parameterwerte, welche mit der Beobachtung $X = x$ verträglich sind.

$$I = \{\theta_0 \mid \text{Test mit Alternativhypothese } H_A : \theta \star \theta_0 \text{ und Signifikanzniveau } \alpha \text{ belässt } H_0\}, \qquad \text{wobei } \star \in \{<, >, \neq\}$$

Es gilt:

$$P_\theta(\theta \in I) \geq 1 - \alpha$$

Normalapproximation (für 2-seitiges Problem):

$$\begin{aligned}
X_i \sim \text{Poisson}(\lambda) : \quad & I = \widehat{\lambda} \pm \Phi^{-1}\left(1 - \frac{\alpha}{2}\right)\sqrt{\tfrac{1}{n}\widehat{\lambda}} \qquad \text{mit } \widehat{\lambda} = \overline{X}_n \\
X \sim \text{Binom}(n,p) : \quad & I = \widehat{p} \pm \Phi^{-1}\left(1 - \frac{\alpha}{2}\right)\sqrt{\widehat{p}\,(1-\widehat{p}) \cdot \tfrac{1}{n}} \qquad \text{mit } \widehat{p} = \frac{x}{n}
\end{aligned}$$

7.3 z-Test

$$\begin{aligned}
\text{Annahme:} \quad & \sigma_X^2 \text{ bekannt} \\
\text{Teststatistik:} \quad & Z = \sqrt{n}\,\frac{\overline{X}_n - \mu_0}{\sigma_X} && Z \sim N(0,1) \\
\text{Verwerfungsbereich:} \quad & |Z| > \Phi^{-1}\left(1 - \frac{\alpha}{2}\right) && \text{für } H_A : \mu \neq \mu_0 \\
& Z < -\Phi^{-1}(1-\alpha) && \text{für } H_A : \mu < \mu_0 \\
& Z > \Phi^{-1}(1-\alpha) && \text{für } H_A : \mu > \mu_0 \\
\text{Vertrauensintervall:} \quad & I = \left[\overline{x} - \frac{\sigma}{\sqrt{n}} \cdot \Phi^{-1}\left(1 - \frac{\alpha}{2}\right), \overline{x} + \frac{\sigma}{\sqrt{n}} \cdot \Phi^{-1}\left(1 - \frac{\alpha}{2}\right)\right] && \text{für 2-seitiges Problem} \\
\text{Fehler 2. Art:} \quad & P_\mu(Z \notin K) && \text{für } H_A : \mu < \mu_0 \\
\text{(wenn } \mu \text{ der wahre Parameter ist)} \quad & = P_\mu\left(\sqrt{n}\,\frac{\overline{X}_n - \mu + \mu_0 - \mu_0}{\sigma_X} > -\Phi^{-1}(1-\alpha)\right) \\
& = P_\mu\left(\sqrt{n}\,\frac{\overline{X}_n - \mu_0}{\sigma_X} > -\Phi^{-1}(1-\alpha) + (\mu - \mu_0)\frac{\sqrt{n}}{\sigma_X}\right) \\
& = 1 - \Phi\left(-\Phi^{-1}(1-\alpha) + (\mu - \mu_0)\frac{\sqrt{n}}{\sigma_X}\right)
\end{aligned}$$

7.4 t-Test

Schätzung für σ^2: $\widehat{\sigma}^2 = \frac{1}{n-1}\sum_{i=1}^{n}(X_i - \overline{X}_n)^2$

Teststatistik: $T = \sqrt{n}\,\frac{\overline{X}_n - \mu_0}{\widehat{\sigma}}$ Student-t verteilt mit $n-1$ Freiheitsgraden

Verwerfungsbereich: $|T| > (t_{n-1})^{-1}\left(1 - \frac{\alpha}{2}\right)$ für $H_A : \mu \neq \mu_0$

$T < -(t_{n-1})^{-1}(1-\alpha)$ für $H_A : \mu < \mu_0$

$T > (t_{n-1})^{-1}(1-\alpha)$ für $H_A : \mu > \mu_0$

Vertrauensintervall: $I = \left[\overline{x}_n - \frac{\widehat{\sigma}}{\sqrt{n}} \cdot t_{n-1,1-\frac{\alpha}{2}}, \overline{x}_n + \frac{\widehat{\sigma}}{\sqrt{n}} \cdot t_{n-1,1-\frac{\alpha}{2}}\right]$ für 2-seitiges Problem

$I = \left[\overline{x}_n - \frac{\widehat{\sigma}}{\sqrt{n}} \cdot t_{n-1,1-\alpha}, \infty\right]$ für $H_A : \mu < \mu_0$

P-Wert: $2 \cdot (1 - F_{t_{n-1}}(|z|))$ für 2-seitiges Problem

$1 - F_{t_{n-1}}(z)$ für 1-seitiges Problem

9 Punktschätzungen

9.1 Binomial- und Poissonverteilung

$$X \sim \text{Binomial}(n,p) \qquad \widehat{p} = \frac{X}{n}$$
$$X \sim \text{Poisson}(\lambda) \qquad \widehat{\lambda} = X$$

9.2 Normalverteilung

Schätzer	Schätzwert	
$\widehat{\mu} = \frac{1}{n}\sum_{i=1}^{n} x_i = \overline{x}_n$	$\widehat{\mu} = \frac{1}{n}\sum_{i=1}^{n} X_i = \overline{X}_n$	mit $\mathrm{E}(\widehat{\mu}) = \mu$
$\widehat{\sigma}^2 = \frac{1}{n-1}\sum_{i=1}^{n}(x_i - \widehat{\mu})^2$	$\widehat{\sigma}^2 = \frac{1}{n-1}\sum_{i=1}^{n}(X_i - \widehat{\mu})^2$	mit $\mathrm{E}(\widehat{\sigma}^2) = \sigma^2$

9.3 Allgemeine Methoden

Beobachtungen $x_1, x_2, \ldots, x_n$ werden aufgefasst als Realisierungen von $X_1, X_2, \ldots, X_n \overset{iid}{\sim} F_\theta$, wobei F eine bekannte Verteilungsfunktion mit unbekanntem Parameter θ der Dimension r ist.

9.3.1 Momentenmethode

Der unbekannte Parameter θ kann ausgedrückt durch Momente $\mu_k = \mathrm{E}[X^k], \quad 1 \le k \le p$ werden mit

$$\theta_j = g_j(\mu_1, \mu_2, \ldots, \mu_p), \quad 1 \le j \le r \quad (= \dim(\theta))$$

Schätze Momente μ_k mit

$$\widehat{\mu}_k = \frac{1}{n}\sum_{i=1}^{n} X_i^k$$

und berechne den geschätzten Parameter mit $\widehat{\theta}_j = g_j(\widehat{\mu}_1, \widehat{\mu}_2, \ldots, \widehat{\mu}_p)$.
Beispiele

1. $X_i \sim \text{Poisson}(\lambda), \quad \theta = \lambda, \quad r = 1$
   $\mathrm{E}(X_i) = \lambda \quad \Longrightarrow \quad \mu_1 = \lambda$
   $\theta = g(\mu_1) = \mu_1 \quad \Longrightarrow \quad \widehat{\theta} = g(\widehat{\mu}_1) = \widehat{\mu}_1 = \frac{1}{n}\sum_{i=1}^{n} X_i$
2. $X_i \sim \mathcal{N}(\mu, \sigma^2), \quad \theta = (\mu, \sigma^2), \quad r = 2$
   $\mu = \mathrm{E}(X_i) = \mu_1$
   $\sigma^2 = \mathrm{Var}(X_i) = \mathrm{E}(X_i^2) - (\mathrm{E}(X_i))^2 = \mu_2 - \mu_1^2$

9.3.2 Maximum-likelihood Schätzer

Die Likelihood-Funktion $L(\theta)$ bestimmt die Wahrscheinlichkeit der Daten unter θ. θ wird dann so geschätzt, dass die Daten maximale Wahrscheinlichkeit haben.

$$L(\theta) = L(\theta, X_1, X_2, \ldots, X_n) = \begin{cases} p_\theta(X_1) \cdot p_\theta(X_2) \cdot \ldots \cdot p_\theta(X_n) & \text{falls } X_i \text{ diskret} \\ f_\theta(X_1) \cdot f_\theta(X_2) \cdot \ldots \cdot f_\theta(X_n) & \text{falls } X_i \text{ stetig} \end{cases}$$

Der Maximum-likelihood Schätzer ist nun

$$\widehat{\theta}_{MLE} = \underset{\theta}{\operatorname{argmax}}\, L(\theta) = \underset{\theta}{\operatorname{argmin}} -l(\theta) \qquad \text{wobei} \quad l(\theta) = \begin{cases} \sum_{i=1}^{n} \log p_\theta(X_i) & \text{falls } X_i \text{ diskret} \\ \sum_{i=1}^{n} \log f_\theta(X_i) & \text{falls } X_i \text{ stetig} \end{cases}$$

10 Vergleich zweier Stichproben

Seien zwei Stichproben

$$\begin{array}{ll} x_1, x_2, \ldots, x_n & \text{unter Versuchsbedingung 1} \\ y_1, y_2, \ldots, y_m & \text{unter Versuchsbedingung 2} \end{array}$$

aufgefasst als Realisierungen von Zufallsvariabeln $X_1, \ldots, X_n, Y_1, \ldots, Y_m$.

10.1 Gepaarte Vergleiche

Beide Versuche werden an derselben Veruchseinheit durchgeführt (Blockbildung). Es gilt $n = m$. Betrachte die Differenzen $u_i = x_i - y_i$ als Realisierungen von U_i mit $n = 1, 2, \ldots, n$.

10.1.1 Vorzeichen-Test

$$\begin{array}{rll} H_0: & \text{Median der Verteilung von } U_i = 0 & \\ H_A: & \text{Median der Verteilung von } U_i \lessgtr 0 \text{ oder } \neq 0 & \\ \text{Teststatistik:} & V = \#\{U_i > 0\} & \text{(Anzahl positiver Vorzeichen)} \\ \text{Verteilung von } V: & V \sim \text{Binomial}(n, P(U_i > 0)) \overset{\text{unter } H_0}{=} \text{Binomial}(n, \frac{1}{2}) & \end{array}$$

10.1.2 Wilcoxon-Test (Rangnummern-Test)

Seien $U_1, U_2, \ldots, U_n \overset{iid}{\sim} F_\mu$ für eine symmetrische Verteilung F_μ mit $\mu = E(U_i) = F_\mu^{-1}(\frac{1}{2})$ (Median).

$$\begin{array}{rl} H_0: & \mu = 0 \\ H_A: & \mu \lessgtr 0 \\ \text{Teststatistik:} & \text{Rang}(|U_i|) = R_i \qquad R_i = k \text{ heisst, dass } U_i \text{ der } k\text{-te Wert unter } |U_1|, |U_2|, \ldots, |U_n| \text{ ist.} \\ & \text{Gibt es mehrere } U_i \text{ mit demselben Betrag, so erhalten alle den durchschnittlichen Rang} \\ & V_i = \mathbb{1}_{[U_i > 0]} \\ & W = \sum_{i=1}^{n} R_i \cdot V_i \\ \text{Verteilung von } W \text{ unter } H_0: & \text{Wilcoxon-Verteilung} \end{array}$$

10.2 Ungepaarte Vergleiche - Zwei Stichproben-Tests

Die Versuchseinheiten werden zufällig einer der beiden Versuchsbedingungen zugeordnet (Randomisierung). Im Allgemeinen gilt $n \neq m$.

10.2.1 2-Stichproben t-Test

Seien $X_1, X_2, \ldots, X_n \sim N(\mu_x, \sigma_x^2)$ und $Y_1, Y_2, \ldots, Y_n \sim N(\mu_y, \sigma_y^2)$ unabhängig.

Annahme: $\sigma^2 = \sigma_x^2 = \sigma_y^2$

Nullhypothese: $H_0 : \mu_x = \mu_y$

Alternativhypothese: $H_A : \mu_x \lesseqgtr \mu_y \quad$ oder $\quad \mu_x \neq \mu_y$

Teststatistik: $T = \dfrac{\overline{X}_n - \overline{Y}_m}{S_{pool}\sqrt{\frac{1}{n} + \frac{1}{m}}}$

$$S_{pool}^2 = \frac{1}{n+m-2}\left(\sum_{i=1}^{n}(X_i - \overline{X}_n)^2 + \sum_{j=1}^{m}(Y_j - \overline{Y}_m)^2\right)$$

$$S_{pool}^2 = \frac{1}{2}\left(\widehat{\sigma}_x^2 + \widehat{\sigma}_y^2\right), \quad \text{falls } n = m$$

Verteilung von T unter H_0: $T \sim t_{n+m-2}$ (Student-t-Verteilung mit $n+m-2$ Freiheitsgraden)

Teil III
Appendix

A Tabellen und Tafeln

A.1 Integral- und Differentialrechnung

$f(x)$	$F(x)$
x^n	$\frac{1}{n+1}x^{n+1}$
$(ax+b)^n$	$\frac{1}{a\cdot(n+1)}(ax+b)^{n+1}$
$\sqrt{x}$	$\frac{2}{3}x^{\frac{3}{2}}$
$\sqrt[n]{x}$	$\frac{n}{n+1}x^{\frac{1}{n}+1}$
$\frac{1}{ax+b}$	$\frac{1}{a}\ln\lvert ax+b\rvert$
$\frac{ax+b}{cx+d}$	$\frac{ax}{c}-\frac{ad-bc}{c^2}\ln\lvert cx+d\rvert$
$\frac{1}{x^2-a^2}$	$\frac{1}{2a}\ln\left\lvert\frac{x-a}{x+a}\right\rvert$
$\sqrt{a^2+x^2}$	$\frac{x}{2}f(x)+\frac{a^2}{2}\ln(x+f(x))$
$\sqrt{a^2-x^2}$	$\frac{x}{2}\sqrt{a^2-x^2}-\frac{a^2}{2}\arcsin\frac{x}{\lvert a\rvert}$
$\sqrt{x^2-a^2}$	$\frac{x}{2}f(x)-\frac{a^2}{2}\ln\left(x+f(x)\right)$
$\frac{1}{\sqrt{x^2+a^2}}$	$\ln(x+\sqrt{x^2+a^2})$
$\frac{1}{\sqrt{x^2-a^2}}$	$\ln(x+\sqrt{x^2-a^2})$
$\frac{1}{\sqrt{a^2-x^2}}$	$\arcsin(\frac{x}{\lvert a\rvert})$
$\frac{1}{\sqrt{1-x^2}}$	$\arcsin(x)$
$\frac{-1}{\sqrt{1-x^2}}$	$\arccos(x)$
$\frac{1}{x^2+a^2}$	$\frac{1}{a}\arctan(\frac{x}{a})$
$\frac{-1}{1+x^2}$	$\operatorname{arccot}(x)$
$\frac{1}{\sqrt{x^2+1}}$	$\operatorname{arsinh}(x)$
$\frac{-1}{\sqrt{x^2-1}}$	$\operatorname{arcosh}(x)$
$\frac{1}{1-x^2}$	$\operatorname{artanh}(x)$
$\sin(ax+b)$	$-\frac{1}{a}\cos(ax+b)$
$\cos(ax+b)$	$\frac{1}{a}\sin(ax+b)$
$\tan(x)$	$-\ln\lvert\cos(x)\rvert$
$\cot(x)$	$\ln\lvert\sin(x)\rvert$

$f(x)$	$F(x)$
$\frac{1}{\sin(x)}$	$\ln\left\lvert\tan(\frac{x}{2})\right\rvert$
$\frac{1}{\cos(x)}$	$\ln\left\lvert\tan(\frac{x}{2}+\frac{\pi}{4})\right\rvert$
$\sin^2(x)$	$\frac{1}{2}(x-\sin(x)\cos(x))$
$\cos^2(x)$	$\frac{1}{2}(x+\sin(x)\cos(x))$
$\tan^2(x)$	$\tan(x)-x$
$\cot^2(x)$	$-\cot(x)-x$
$\arcsin(x)$	$x\arcsin(x)+\sqrt{1-x^2}$
$\arccos(x)$	$x\arccos(x)-\sqrt{1-x^2}$
$\arctan(x)$	$x\arctan(x)-\frac{1}{2}\ln(1+x^2)$
$\sin^n(x)$	$s_n=-\frac{1}{n}\sin^{n-1}x\cos x+\frac{n-1}{n}s_{n-2}$
	$s_0=x \quad s_1=-\cos(x)$
$\cos^n(x)$	$c_n=\frac{1}{n}\sin x\cos^{n-1}x+\frac{n-1}{n}c_n$
	$c_0=x \quad c_1=\sin(x)$
$\sinh(x)$	$\cosh(x)$ und umgekehrt
$\tanh(x)$	$\ln(\cosh(x))$
$\frac{f'(x)}{f(x)}$	$\ln\lvert f(x)\rvert$
$\ln\lvert x\rvert$	$x\cdot(\ln\lvert x\rvert-1)$
$\frac{1}{x}(\ln x)^n$	$\frac{1}{n+1}(\ln x)^{n+1} \quad n\neq-1$
$\frac{1}{x}\ln x^n$	$\frac{1}{2n}(\ln x^n)^2 \quad n\neq0$
$\frac{1}{x\ln x}$	$\ln\lvert\ln x\rvert \quad x>0,x\neq1$
a^{bx+k}	$\frac{1}{b\ln a}a^{bx+k}$
$x\cdot e^{cx}$	$\frac{cx-1}{c^2}\cdot e^{cx}$
$x^n\ln x$	$\frac{x^{n+1}}{n+1}\left(\ln x-\frac{1}{n+1}\right) \quad n\neq-1$
$e^{cx}\sin(ax+b)$	$\frac{e^{cx}(c\sin(ax+b)-a\cos(ax+b))}{a^2+c^2}$
$e^{cx}\cos(ax+b)$	$\frac{e^{cx}(c\cos(ax+b)+a\sin(ax+b))}{a^2+c^2}$

$f(x)$	$f'(x)$
x^n	nx^{n-1}
$\sqrt{x}$	$\frac{1}{2\sqrt{x}}$
$\sqrt[n]{x}$	$\frac{1}{n}x^{\frac{1}{n}-1}$
$\frac{1}{f(x)}$	$\frac{-f'(x)}{(f(x))^2}$
$\tan(x)$	$\tan^2(x)+1=\frac{1}{\cos^2(x)}$
$\cot(x)$	$-(1+\cot^2(x))=\frac{-1}{\sin^2(x)}$

$f(x)$	$f'(x)$
$\ln\lvert x\rvert$	$\frac{1}{x}$
$\log_a\lvert x\rvert$	$\frac{1}{x\ln a}=\log_a(e)\frac{1}{x}$
a^{cx}	$a^{cx}\cdot c\ln a$
x^x	$x^x\cdot(1+\ln x) \quad x>0$
$(x^x)^x$	$(x^x)^x(x+2x\ln(x)) \quad x>0$
$x^{(x^x)}$	$x^{(x^x)}(x^{x-1}+\ln x\cdot x^x(1+\ln x)))$

A.2 Verschiedenes

A.2.1 Binomialkoeffizient

$$\binom{n}{k-1} + \binom{n}{k} = \binom{n+1}{k} \qquad \sum_{k=0}^{n} \binom{n}{k} = 2^n$$
$$(a+b)^n = \sum_{k=0}^{n} \binom{n}{k} a^{n-k} b^k \qquad \binom{n}{k} = \binom{n}{n-k}$$

A.2.2 Kombinatorik

Ziehen von k Elementen aus einer Menge mit n Elementen

	geordnet	ungeordnet
mit zurücklegen	n^k	$\binom{n+k-1}{k}$
ohne zurücklegen	$\frac{n!}{(n-k)!}$	$\binom{n}{k} = \frac{n!}{k!(n-k)!}$

A.2.3 Mitternachtsformel

$$ax + bx + c = 0 \qquad \Longrightarrow \qquad x_{1,2} = \frac{-b \pm \sqrt{b^2 - 4ac}}{2a}$$

A.2.4 Ausklammern

$$x^n - y^n = (x-y)(x^{n-1} + x^{n-2}y + x^{n-3}y^2 + \ldots + xy^{n-2} + y^{n-1})$$
$$x^n - 1 = (x-1)(x^{n-1} + x^{n-2} + x^{n-3} + \ldots + x + 1)$$

A.3 Regeln für Integrale

A.3.1 Partielle Integration

Seien $f, g \in C^0(]a,b[)$ mit Stammfunktionen $F, G \in C^1(]a,b[)$, dann gilt

$$\int f'(x) \cdot g(x) \, \mathrm{d}x = f(x) \cdot g(x) - \int f(x) \cdot g'(x) \, \mathrm{d}x$$

A.3.2 Substitutionsregel

Seien $f, g \in C^1(]a,b[)$. Dann gilt

$$\int_{x_0}^{x_1} f'(g(x)) \, g'(x) \, \mathrm{d}x = (f \circ g)\Big|_{x=x_0}^{x_1} = f(g(x_1)) - f(g(x_0)) = \int_{g(x_0)}^{g(x_1)} f'(z) \, \mathrm{d}z$$

A.3.3 Standardsubstitutionen

Integral	Substitution	Differential	Bemerkungen
$\int f(g(x), g'(x))dx$	$t = g(x)$	$dx = \frac{dt}{g'(x)}$	Lsg: $\frac{1}{2}[f(x)^2] + C$
$\int f((ax+b))dx$	$t = ax + b$	$dx = \frac{dt}{a}$	Lsg: $\frac{1}{a}\int f(u)du)$
$\int f(x, \sqrt{ax+b})dx$	$x = \frac{t^2-b}{a}$	$dx = \frac{2tdt}{a}$	$t \geq 0$
$\int f(x, \sqrt{ax^2+bx+c})dx$	$x = \alpha t + \beta$	$dx = \alpha dt$	wähle α und β so, dass gilt $ax^2 + bx + c = \gamma \cdot (\pm t^2 \pm 1)$
$\int f(x, \sqrt{a^2-x^2})dx$	$x = a \cdot \sin t$	$dx = a \cdot \cos t dt$	$-\frac{\pi}{2} \leq t \leq \frac{\pi}{2}$
$\int f(x, \sqrt{a^2+x^2})dx$	$x = a \cdot \sinh t$	$dx = a \cdot \cosh t dt$	$t \in \mathbb{R}$
$\int f(x, \sqrt{x^2-a^2})dx$	$x = a \cdot \cosh t$	$dx = a \cdot \sinh t dt$	$t \geq 0$
$\int f(e^x, \sinh x, \cosh x)dx$	$e^x = t$	$dx = \frac{dt}{t}$	$t > 0$, und dabei gilt: $\sinh x = \frac{t^2-1}{2t}$ $\cosh x = \frac{t^2+1}{2t}$
$\int f(\sin x, \cos x)dx$	$\tan \frac{x}{2} = t$	$dx = \frac{2dt}{1+t^2}$	$-\frac{\Pi}{2} < t < \frac{\Pi}{2}$, und dabei gilt: $\sin x = \frac{2t}{1+t^2}$, $\cos x = \frac{1-t^2}{1+t^2}$