TECHNISCHE UNIVERSITÄT DRESDEN

Fakultät Wirtschaftswissenschaften Lehrstuhl für Quantitative Verfahren, insbesondere Statistik

STATISTIK IM BACHELORSTUDIUM

9. Auflage, 2016

Prof. Dr. Stefan Huschens

Inhaltsverzeichnis

0 Vorbemerkungen **1**

1 Merkmale und Daten **3**

1.1 Häufigkeiten 3
1.2 Verteilungsfunktion und Quantile 6
1.3 Stetige Klassierung 7
1.4 Ergänzungen 9

2 Auswertung eindimensionaler Daten **11**

2.1 Mittelwerte 12
2.2 Streuungsmessung 14
2.3 Ergänzungen 15

3 Messzahlen und Indizes **17**

3.1 Messzahlen 17
3.2 Preis-, Mengen- und Wertindizes 21
3.3 Ergänzungen 23

4 Auswertung mehrdimensionaler Daten **25**

4.1 Korrelationskoeffizient 25
4.2 Rangkorrelation 27
4.3 Kontingenzkoeffizient 28
4.4 Deskriptive lineare Regression 30
4.5 Ergänzungen 33

5 Zeitreihenanalyse **35**

5.1 Komponentenmodelle 36
5.2 Bestimmung der glatten Komponente durch gleitende Durchschnitte 36
5.3 Bestimmung der Saisonkomponente 38
5.4 Ergänzungen 39

6 Zufallsexperimente und Wahrscheinlichkeit **41**

6.1 Zufallsexperimente 43
6.2 Wahrscheinlichkeit 45
6.3 Unabhängigkeit von Ereignissen 47
6.4 Ergänzungen 49

7 Zufallsvariable und Verteilung **51**

7.1 Zufallsvariable 51

7.2 Verteilungsfunktion 52
7.3 Ergänzungen 54

8 Diskrete Verteilungen 55
8.1 Wahrscheinlichkeitsfunktion 55
8.2 Bernoulli-Verteilung 56
8.3 Binomialverteilung 57
8.4 Poisson-Verteilung 59
8.5 Ergänzungen 60

9 Stetige Verteilungen 63
9.1 Dichtefunktion 63
9.2 Rechteckverteilung 64
9.3 Exponentialverteilung 65
9.4 Normalverteilung 67
9.5 Ergänzungen 69

10 Erwartungswert und Varianz 71
10.1 Erwartungswert 71
10.2 Varianz und Standardabweichung 74
10.3 Ergänzungen 77

11 Gemeinsame Verteilung von Zufallsvariablen 79
11.1 Gemeinsame Verteilung von zwei Zufallsvariablen 79
11.2 Unabhängigkeit von zwei Zufallsvariablen 81
11.3 Kovarianz und Korrelation 83
11.4 Verteilung mehrerer Zufallsvariablen 84
11.5 Summen unabhängiger Zufallsvariablen 85
11.6 Randverteilung 85
11.7 Ergänzungen 86

12 Parameterschätzung 89
12.1 Grundgesamtheit und Stichprobe 89
12.2 Stichprobenfunktionen (Statistiken) und Parameter 92
12.3 Erwartungstreue 95
12.4 Wahrscheinlichkeitsverteilung eines Schätzers 96
12.5 Ergänzungen 98

13 Konfidenzintervalle 101
13.1 Konfidenzniveau und Konfidenzintervall 101
13.2 Konfidenzintervall für μ einer Normalverteilung 103
13.2.1 σ^2 bekannt 103
13.2.2 σ^2 unbekannt 104
13.3 Konfidenzintervall für σ^2 einer Normalverteilung 105
13.4 Konfidenzintervall für π einer Bernoulli-Verteilung 106
13.5 Ergänzungen 106

14 Grundstruktur statistischer Tests 109

14.1 Schätz- und Testproblem . . . 109
14.2 Beispiel zum Hypothesentest . . . 110
14.3 Allgemeine Teststruktur . . . 111

15 Drei Tests für die Normalverteilung 113
15.1 Der Gauß-Test . . . 114
15.2 Der t-Test . . . 116
15.3 Test für eine Varianz . . . 117

16 Fehlerwahrscheinlichkeit, Signifikanzniveau und p-Wert 119
16.1 Fehler 1. und 2. Art . . . 119
16.2 Signifikanzniveau . . . 120
16.3 p-Wert . . . 120
16.4 Tests und Statistiksoftware . . . 122
16.5 Ergänzungen . . . 123

17 Tests für den Vergleich zweier Parameter 127
17.1 Zweistichproben-Gauß-Test . . . 128
17.2 Zweistichproben-t-Test . . . 128
17.3 Vergleich zweier Varianzen (F-Test) . . . 129
17.4 Der t-Differenzentest . . . 130
17.5 Ergänzungen . . . 132

18 Approximative Testverfahren 133
18.1 Approximativer Gauß-Test . . . 134
18.2 Approximativer Zwei-Stichproben-Gauß-Test . . . 134
18.3 Test für eine Wahrscheinlichkeit . . . 135
18.4 Vergleich zweier Wahrscheinlichkeiten . . . 136
18.5 Chiquadrat-Anpassungstest . . . 136
18.6 Chiquadrat-Unabhängigkeitstest . . . 138
18.7 Ergänzungen . . . 139

19 Grundlagen der asymptotischen Statistik 141
19.1 Konsistenz von Schätzern . . . 142
19.2 Schwaches Gesetz der großen Zahlen . . . 143
19.3 Zentraler Grenzwertsatz der Statistik . . . 145
19.4 Asymptotisch begründete Tests . . . 146
19.5 Ergänzungen . . . 147

20 Bedingte Wahrscheinlichkeit und bedingte Verteilung 151
20.1 Bedingte Wahrscheinlichkeiten . . . 151
20.2 Bedingte Verteilung . . . 153
20.3 Ergänzungen . . . 154

A Formelsammlung 157
A.1 Allgemeine mathematische Notation . . . 157
A.2 Beschreibende Statistik . . . 157
A.3 Wahrscheinlichkeitsrechnung . . . 161

A.4 Schließende Statistik 166
A.5 Tabellen 170
Tabelle 1: Verteilungsfunktion der Binomialverteilung 170
Tabelle 2: Verteilungsfunktion der Poisson-Verteilung 171
Tabelle 3: Verteilungsfunktion der Standard-Normalverteilung 172
Tabelle 4: Quantile der t-Verteilung 173
Tabelle 5: Quantile der χ^2-Verteilung 174
Tabelle 6: Quantile der F-Verteilung 175

B Altgriechisches Alphabet **177**

Kapitel 0

Vorbemerkungen

Zum Skript

Dieses Skript enthält die meisten der in der Vorlesung zum Modul „Statistik“ verwendeten Folien, ergänzende Bemerkungen und Beispiele und im Anhang eine Formelsammlung. Viele Kapitel in diesem Skript haben einen ergänzenden Abschnitt mit der Bezeichnung „Ergänzungen“, der über die verwendeten Folien hinausgeht und das vertiefende Selbststudium unterstützen soll. Er enthält auch Hinweise und Beispiele, welche die Verknüpfung zwischen dem in der Vorlesung behandelten Material und den Ausführungen in Lehrbüchern erleichtern sollen.

Hinweise zur Literatur

Die Ausführungen zur beschreibenden (deskriptiven) Statistik in den Kapiteln 1 bis 5 stützen sich wesentlich auf

- Mosler, K., Schmid, F.: Beschreibende Statistik und Wirtschaftsstatistik, 4. Aufl., Springer: Berlin, Heidelberg 2009.

Die Ausführungen zur Wahrscheinlichkeitsrechnung in den Kapiteln 6 bis 11 und zur schließenden (induktiven) Statistik in den Kapiteln 12 bis 20 stützen sich auf

- Mosler, K., Schmid, F.: Wahrscheinlichkeitsrechnung und schließende Statistik, 4. Aufl., Springer: Berlin, Heidelberg 2011.

Vorausgesetzte Grundlagen der Schulmathematik können im Selbststudium mit Hilfe von

- Cramer, E., Nešlehová, J.: Vorkurs Mathematik – Arbeitsbuch zum Studienbeginn in Bachelor-Studiengängen, 6. Aufl., Springer: Berlin, Heidelberg 2015

aufgefrischt werden. Ein nützliches Nachschlage- und Übersichtswerk und ein hilfreicher Begleiter für das gesamte Bachelor-, Master- und Diplomstudium ist

- Rinne, H.: Taschenbuch der Statistik, 4. Aufl., Verlag Harri Deutsch: Frankfurt am Main 2008.

0-1

Prof. Dr. Stefan Huschens
Lehrstuhl für Quantitative Verfahren, insbesondere Statistik
Technische Universität Dresden

STATISTIK im Bachelorstudium

Kapitel 1-5:

Beschreibende (deskriptive) Statistik

Kapitel 6-11:

Wahrscheinlichkeitsrechnung

Kapitel 12-20:

Schließende (induktive) Statistik

0-2

Bemerkung 0.1 (Seitenzählung und Numerierung)

- Die Seitenzählung der Folien – jeweils in der rechten obere Ecke – ist nach dem Muster **Kapitel-Seite** aufgebaut, z. B. ist $\boxed{\text{3-4}}$ die Seitenkennung für die vierte Folie im dritten Kapitel.
- In jedem Kapitel sind die Bemerkungen, Beispiele, Definitionen und Sätze nach dem Schema **Kapitel.Nummer** fortlaufend numeriert.

Bemerkung 0.2 (Literatur)
Die ersten fünf Kapitel stützen sich auf das Lehrbuch

> Mosler, K., Schmid, F.: **Beschreibende Statistik und Wirtschaftsstatistik**, 4. Aufl., Springer: Berlin, Heidelberg 2009,

die weiteren Kapitel stützen sich auf das Lehrbuch

> Mosler, K., Schmid, F.: **Wahrscheinlichkeitsrechnung und schließende Statistik**, 4. Aufl., Springer: Berlin, Heidelberg 2011.

0-3

Bemerkung 0.3 (Notation)

- Die Notation folgt weitgehend den Lehrbüchern von Mosler/Schmid.
- Anstelle des Dezimalkommas wird in diesem Skript ein Dezimalpunkt verwendet, also z. B.
$$0.2 = \frac{2}{10} .$$
- Mosler/Schmid verwenden in ihrem Band „Beschreibende Statistik und Wirtschaftsstatistik“ Dezimalkommas und in ihrem Band „Wahrscheinlichkeitsrechnung und schließende Statistik“ Dezimalpunkte.

Kapitel 1

Merkmale und Daten

1-1

1. Merkmale und Daten

1.1 Häufigkeiten

1.2 Verteilungsfunktion und Quantile

1.3 Stetige Klassierung

1-2

1.1 Häufigkeiten

Beispiel 1.1 (Einheiten und Grundgesamtheit)

- Die Studierenden im Hörsaal sind **statistische Einheiten** $e_1, e_2, \ldots, e_n$.
- Sie bilden die **Grundgesamtheit** oder statistische Masse

$$G = \{e_1, e_2, \ldots, e_n\}.$$

- $n = |G|$ ist der **Umfang der Masse** oder die **Anzahl der Einheiten**.
- An den Einheiten können statistische **Merkmale** mit alternativen **Merkmalswerten** (Merkmalsausprägungen, Merkmalsmodalitäten) beobachtet oder gemessen werden.
- Die Einheiten heißen dann auch **Merkmalsträger**.

1-3

Beispiel 1.2 (Merkmale)

- **Geschlecht** ist ein **qualitatives binäres** Merkmal mit den Merkmalswerten {männlich, weiblich} bzw. $\{1, 2\}$ mit der **Kodierung** männlich = 1, weiblich = 2
- **Herkunftsland (Nationalität)** ist ein **qualitatives kategoriales** Merkmal mit den Merkmalswerten {Deutschland, Türkei, Italien, Serbien und Montenegro, Griechenland, Polen, Kroatien, Sonstige} oder $\{1, 2, \ldots, 8\}$, kodiert z. B. durch Deutschland = 1, Türkei = 2, Italien = 3, ..., Sonstige = 8.
- **Körpergröße** ist ein **quantitatives** Merkmal und zwar ein im Prinzip **kontinuierliches (stetiges)** Merkmal, aber praktisch **diskret** gemessenes Merkmal, z. B. in cm: Merkmalswerte in $\{151, 153, ..., 205\}$

1-4

Bemerkung 1.3 (Daten und Merkmalswerte)

- X sei ein Merkmal, das in einer Grundgesamtheit $G = \{e_1, e_2, \ldots, e_n\}$ beobachtet wird und durch Zahlen kodiert ist.
- $x_i \in \mathbb{R}$ bezeichnet den an der Einheit $e_i \in G$ beobachteten Wert des Merkmals X. Durch
$$(x_1, x_2, \ldots, x_n)$$
ist die **Urliste** der n **Daten** oder **Beobachtungen** gegeben.
- Es gibt $1 \leq J \leq n$ **verschiedene Merkmalswerte** ξ_1, ξ_2, $\ldots$, ξ_J, die in den Daten auftreten. Es ist
  - $J = 1$, falls alle x_i gleich sind,
  - $J = n$, falls alle x_i voneinander verschieden sind, und
  - $J < n$, falls Merkmalswerte mehrfach auftreten.

1-5

Beispiel 1.4 (Merkmal Geschlecht)

- Das Merkmal Geschlecht sei kodiert durch die $J = 2$ verschiedenen Merkmalswerte $\xi_1 = 0$ für männlich und $\xi_2 = 1$ für weiblich. Von $n = 348$ Personen sind 204 männlich und 144 weiblich.
- Aus den Daten
$$(x_1, x_2, \ldots, x_n) = (0, 0, 1, 0, 1, 1, \ldots, 0, 1)$$
ergeben sich die **absoluten Häufigkeiten**
$$n_1 = 204, \quad n_2 = 144$$
und die **relativen Häufigkeiten**
$$f_1 = \frac{n_1}{n} = \frac{204}{348} = 0.586 = 58.6\%, \quad f_2 = \frac{n_2}{n} = \frac{144}{348} = 0.414 = 41.4\%\,.$$

1-6

Definition 1.5 (Absolute und relative Häufigkeiten)
Gegeben seien n Daten (Beobachtungen) $x_1, x_2, \ldots, x_n$ mit den J verschiedenen Merkmalswerten $\xi_1, \xi_2, \ldots, \xi_J$.

- Für $j = 1, \ldots, J$ ist

  $$n_j = |\{i \mid x_i = \xi_j\}| = \text{„Anzahl der Daten mit dem Merkmalswert } \xi_j\text{“}$$

  die **absolute Häufigkeit** des Merkmalswertes ξ_j.

- Für $j = 1, \ldots, J$ ist

  $$f_j \overset{\text{def}}{=} \frac{n_j}{n}, \quad j = 1, \ldots, J$$

  die **relative Häufigkeit** des Merkmalswertes ξ_j.

1-7

Bemerkung 1.6 (Eigenschaften)

- n_j gibt die **Anzahl** der Daten mit dem Merkmalswert ξ_j an.
- f_j gibt den **Anteil** der Daten mit dem Merkmalswert ξ_j an.
- Es gilt

  $$\sum_{j=1}^{J} n_j = n \quad \text{und} \quad \sum_{j=1}^{J} f_j = 1.$$

1-8

Beispiel 1.7 (Merkmal Herkunftsland)
Aus den Daten $(x_1, x_2, \ldots, x_{348}) = (1, 3, 1, 1, 1, 8, 1, 1, 3, 1, 1, 6, ..., 1)$ wird eine **Häufigkeitstabelle** mit absoluten und relativen Häufigkeiten erstellt.

Herkunftsland	$j = \xi_j$	n_j	f_j
Deutschland	1	320	91.95%
Türkei	2	7	2.01%
Italien	3	2	0.57%
Serbien und Montenegro	4	2	0.57%
Griechenland	5	1	0.29%
Polen	6	1	0.29%
Kroatien	7	1	0.29%
Sonstige	8	14	4.02%
	$\sum_{j=1}^{8}$	348	100%

Durch Rundungsfehler addiert sich die letzte Spalte zu 99.99%.

1-9

1.2 Verteilungsfunktion und Quantile

Definition 1.8 (Verteilungsfunktion)
Gegeben seien die n Daten $x_1, x_2, \ldots, x_n$. Dann heißt die Funktion

$$F : \mathbb{R} \to [0, 1],$$

$$F(x) \overset{\text{def}}{=} \frac{|\{i \mid x_i \leq x\}|}{n} = \frac{\text{„Anzahl der Daten mit } x_i \leq x\text{“}}{n}, \quad x \in \mathbb{R}$$

(empirische) Verteilungsfunktion (engl.: *empirical distribution function*).

1-10

Bemerkung 1.9 (Eigenschaften)

- Für J Merkmalswerte $\xi_1, \xi_2, \ldots, \xi_J$ mit den absoluten Häufigkeiten $n_1, n_2, \ldots, n_J$ und den relativen Häufigkeiten $f_1, f_2, \ldots, f_J$ gilt

$$F(x) = \frac{1}{n} \sum_{r:\xi_r \leq x} n_r = \sum_{r:\xi_r \leq x} f_r, \quad x \in \mathbb{R}\,.$$

- Die Verteilungsfunktion F ist eine **Treppenfunktion**, d. h. sie ist **stückweise konstant**, und hat die **Sprungstellen** ξ_j mit den **Sprunghöhen** f_j.
- Falls $\xi_1 < \xi_2 < \ldots < \xi_J$, gilt

$$0 < F(\xi_1) < F(\xi_2) < \ldots < F(\xi_J) = 1.$$

1-11

Definition 1.10 (p-Quantil)
Für $0 < p < 1$ heißt

$$\begin{aligned} \tilde{x}_p &\overset{\text{def}}{=} \min\{x \in \mathbb{R} \mid F(x) \geq p\} \\ &= \text{„kleinster Wert } x \in \mathbb{R} \text{ mit der Eigenschaft, dass } F(x) \geq p\text{“} \end{aligned}$$

das p**-Quantil** (engl.: p-*quantile*) oder das p**-Fraktil** der Daten.

1-12

Bemerkung 1.11 (Berechnung des p-Quantils)

- Für aufsteigend geordnete Beobachtungen $x_1 \leq x_2 \leq \ldots \leq x_n$ gilt

$$\tilde{x}_p = \begin{cases} x_{np} & \text{falls } np \text{ ganzzahlig} \\ x_{[np]+1} & \text{sonst} \end{cases},$$

  wobei $[np]$ den **ganzzahligen Teil** von np bezeichnet.

- Z. B. ergibt sich für $n = 3$ und $p = 1/2$ der ganzzahlige Teil $[np] = [3/2] = 1$ und der Index $[np] + 1 = 2$. Für $x_1 \leq x_2 \leq x_3$ gilt also

$$\tilde{x}_{1/2} = x_{[3/2]+1} = x_2.$$

1-13

Definition 1.12 (Spezielle Quantile und Boxplot)

- Das Quantil $\tilde{x}_{0.5}$ heißt auch **Median** (engl.: *median*).
- Die Quantile $\tilde{x}_{0.25}, \tilde{x}_{0.5}$ und $\tilde{x}_{0.75}$ heißen **Quartile**.
- Eine spezielle graphische Darstellung des Minimums der Beobachtungen, der Quartilswerte $\tilde{x}_{0.25}, \tilde{x}_{0.5}$ und $\tilde{x}_{0.75}$ sowie des Maximums der Beobachtungen heißt **Boxplot** oder **Schachteldiagramm**.

1-14

1.3 Stetige Klassierung

Beispiel 1.13 (Körpergröße in cm)

- **Geordnete Beobachtungen:** ($n = 348$)

$$151, 153, \ldots, 168, 168, 168, \ldots, 205$$

- **Klassierung** (Klassenbildung)

  $[140, 150[$ entspricht der Klasse „140 bis unter 150“, usw.

  $]140, 150]$ entspricht der Klasse „über 140 bis 150“, usw.

  Beide Systeme der Klassierung sind üblich.

- Die beiden Randklassen werden häufig durch links bzw. rechts abgeschlossene Intervalle definiert.

1-15

Bemerkung 1.14 (Bezeichnungen)

- $J \in \mathbb{N}$ bezeichnet die Anzahl der Klassen.
- K_j bezeichne die j-te Klasse für $j = 1, 2, \ldots, J$.
- n_j bezeichne die **Anzahl der Daten in der Klasse** K_j.
- n_j heißt auch **Klassenhäufigkeit** oder **Besetzungszahl** der j-ten Klasse.
- Mit
$$f_j = \frac{n_j}{n}$$
wird der **Anteil der Daten in der Klasse** K_j bezeichnet.
- Es gilt
$$\sum_{j=1}^{J} n_j = n \quad \text{und} \quad \sum_{j=1}^{J} f_j = 1\,.$$

1-16

Beispiel 1.15 (Klassierte Daten)

Häufigkeitstabelle für klassierte Daten

j	K_j	n_j	f_j in %
1	[150,160]	14	4.0
2	]160,170]	81	23.3
3	]170,180]	141	40.5
4	]180,190]	94	27.0
5	]190,200]	16	4.6
6	]200,210]	2	0.6
Summe		348	100.0

1-17

Bemerkung 1.16 (Histogramm)
Ein **Histogramm** ist eine graphische Darstellung für klassierte Daten nach dem Prinzip:

- Die **Rechteckflächen** über den Intervallen sind proportional zu den Klassenhäufigkeiten.
- Bei **gleichen** Klassenbreiten bedeutet dies, dass die Höhe proportional zur Klassenhäufigkeit ist.
- Bei **ungleichen** Klassenbreiten muss die Höhe geeignet berechnet werden.

Definition 1.17 (Empirische Dichte)
Die Intervallobergrenzen seien mit x_j^o und die Intervalluntergrenzen mit x_j^u bezeichnet. Dann heißt der Quotient

$$\hat{f}_j \stackrel{\text{def}}{=} \frac{f_j}{x_j^o - x_j^u}$$

für $j = 1, 2, \ldots, J$ **empirische Dichte** der Daten in der Klasse K_j.

1-18

Bemerkung 1.18 (Eigenschaften)

- Die empirische Dichte führt zu einem Histogramm, wenn die empirische Dichte als Höhe für das jeweilige Intervall gewählt wird.
- Es gilt für die Einzelflächen

$$\hat{f}_j(x_j^o - x_j^u) = f_j$$

und daher für die Gesamtfläche

$$\sum_{j=1}^{J} \hat{f}_j(x_j^o - x_j^u) = 1\,.$$

1-19

Beispiel 1.19 (Körpergröße klassiert)

- $n = 348$, $J = 3$, ungleiche Klassenbreite
- Klassenbreite: Intervallobergrenze - Intervalluntergrenze $= x_j^o - x_j^u$

Häufigkeitstabelle für klassierte Daten bei ungleicher Klassenbreite

j	K_j	$x_j^o - x_j^u$	n_j	$f_j = \frac{n_j}{n}$	$\hat{f}_j = \frac{f_j}{x_j^o - x_j^u}$
1	[130,160]	30	14	0.040	0.00133
2	]160,180]	20	222	0.638	0.03190
3	]180,220]	40	112	0.322	0.00805

1.4 Ergänzungen

Bemerkung 1.a (Alternative Bezeichnungen) Die empirische Verteilungsfunktion wird auch mit $\hat{F}$ oder F_n bezeichnet. Diese Bezeichnungen werden insbesondere im Zusammenhang mit der schließenden (induktiven) Statistik verwendet, wo es erforderlich ist, einerseits eine **theoretische Verteilungsfunktion** bzw. eine Verteilungsfunktion des Merkmals in der Grundgesamtheit und andererseits eine **empirische Verteilungsfunktion** bzw. eine Verteilungsfunktion des Merkmals in der Stichprobe zu unterscheiden.

Definition 1.b (Kumulierte relative Häufigkeiten) Gegeben seien J Merkmalswerte $\xi_1, \xi_2, \ldots, \xi_J$ mit den relativen Häufigkeiten f_j. Für $j = 1, \ldots, J$ sind

$$F(\xi_j) \stackrel{\text{def}}{=} \sum_{r:\xi_r \le \xi_j} f_r$$

die **kumulierten relativen Häufigkeiten**.

Bemerkung 1.c (Spezielle Quantile) Die Quantile $\tilde{x}_{0.2}$, $\tilde{x}_{0.4}$,..., $\tilde{x}_{0.8}$ heißen **Quintile**. Die Quantile $\tilde{x}_{0.1}$, $\tilde{x}_{0.2}$,..., $\tilde{x}_{0.9}$ heißen **Dezile**. Die Quantile $\tilde{x}_{0.01}, \tilde{x}_{0.02}, \ldots, \tilde{x}_{0.99}$ heißen **Centile oder Perzentile**.

Kapitel 2

Auswertung eindimensionaler Daten

2-1

2. Auswertung eindimensionaler Daten

2.1 Mittelwerte
2.2 Streuungsmessung

2-2

Bemerkung 2.1 (Maßzahlen)

- Ziel ist die Charakterisierung einer Häufigkeitsverteilung durch **einen Wert** bzw. **wenige Zahlenwerte** (Kennzahlen), sogenannte **statistische Maßzahlen**.
- Die wichtigsten Maßzahlen sind die in den folgenden beiden Abschnitten behandelten **Mittelwerte** und **Streuungsmaße**.

2-3

2.1 Mittelwerte

Definition 2.2 (Arithmetisches Mittel)
Gegeben seien die Daten $x_1, \ldots, x_n$. Dann heißt

$$\bar{x} \stackrel{\text{def}}{=} \frac{1}{n} \sum_{i=1}^{n} x_i$$

arithmetisches Mittel (engl.: *mean*).

Bemerkung 2.3 (Diskret klassierte Daten)
Treten die Merkmalsausprägungen $\xi_1, \ldots, \xi_J$ mit den absoluten Häufigkeiten n_j bzw. relativen Häufigkeiten f_j auf, so gilt

$$\bar{x} = \frac{1}{n} \sum_{j=1}^{J} n_j \xi_j = \sum_{j=1}^{J} f_j \xi_j .$$

2-4

Definition 2.4 (Gewichtetes Mittel)
Gegeben seien die Daten $x_1, \ldots, x_n$ und ein **Gewichtsvektor**

$$w = (w_1, w_2, \ldots, w_n)$$

mit den **Gewichten** $w_i \geq 0$ für $i = 1, \ldots, n$ und $\sum_{i=1}^{n} w_i = 1$, dann heißt

$$\bar{x}_w \stackrel{\text{def}}{=} \sum_{i=1}^{n} w_i x_i$$

gewichtetes Mittel (oder gewogenes Mittel) zum Gewichtsvektor w.

Bemerkung 2.5 (Arithmetisches Mittel als Spezialfall)
Das arithmetische Mittel ergibt sich als Spezialfall mit den Gewichten $w_i = \frac{1}{n}$ für $i = 1, \ldots, n$, da

$$\sum_{i=1}^{n} \frac{1}{n} x_i = \frac{1}{n} \sum_{i=1}^{n} x_i = \bar{x} .$$

2-5

Definition 2.6 (Modus)
Gegeben seien Beobachtungen $x_1, \ldots, x_n$. Eine Merkmalsausprägung ξ_j mit

$$n_j = \max_{i=1}^{J} n_i \quad \text{bzw.} \quad f_j = \max_{i=1}^{J} f_i$$

heißt **Modus** (engl.: *mode*) oder **Modalwert**.

Bemerkung 2.7

- Alternative Bezeichnungen für den Modus sind **Modalwert**, **häufigster Wert** und **dichtester Wert**.
- Wenn der Modus eindeutig ist und die sogenannte **Häufigkeitsfunktion** $\xi_j \mapsto n_j$ für $j = 1, \ldots, J$ keine weiteren lokalen Maxima besitzt, so spricht man von einer **unimodalen** oder **eingipfligen Häufigkeitsverteilung**.

2-6

Definition 2.8
Für positive Beobachtungen $x_1, \ldots, x_n$ heißt

$$\bar{x}_G \stackrel{\text{def}}{=} \sqrt[n]{\prod_{i=1}^{n} x_i}$$

geometrisches Mittel (engl.: *geometric mean*).

2-7

Beispiel 2.9 (Durchschnittliche Zuwachsrate)

- Tabelle mit Daten, Zuwachsraten und Zuwachsfaktoren:

t	x_t	$w_t = \frac{x_t - x_{t-1}}{x_{t-1}}$	$m_{t-1,t} = 1 + w_t = \frac{x_t}{x_{t-1}}$
0	100		
1	88.6	- 11.4%	0.886
2	97.2	+ 9.7%	1.097
3	104.1	+ 7.1%	1.071
4	111.9	+ 7.5%	1.075

- Die
$$w_t = \frac{x_t - x_{t-1}}{x_{t-1}} = \frac{x_t}{x_{t-1}} - 1$$
heißen **Zuwachsraten** oder (diskrete) **Wachstumsraten**.

- Welche durchschnittliche Zuwachsrate führt von $x_0 = 100$ zu $x_4 = 111.9$?

2-8

- Die Quotienten
$$m_{t-1,t} = \frac{x_t}{x_{t-1}}$$
heißen auch **Zuwachsfaktoren** oder **Wachstumsfaktoren**.

- Der **durchschnittliche Zuwachsfaktor** von der Zeit 0 bis zur Zeit 4 wird durch das geometrische Mittel der Zuwachsfaktoren definiert,
$$\bar{m}_G \stackrel{\text{def}}{=} \sqrt[4]{0.886 \cdot 1.097 \cdot 1.071 \cdot 1.075} = \sqrt[4]{1.119} = 1.0285\,.$$

- Die **durchschnittliche Zuwachsrate** von der Zeit 0 bis zur Zeit 4 ist
$$\bar{w} \stackrel{\text{def}}{=} \bar{m}_G - 1 = 2.85\%$$
und es gilt
$$x_4 = (1 + \bar{w})^4 x_0\,.$$

- Achtung: $\bar{w}$ bezeichnet hier nicht das arithmetische Mittel der w_t.

2-9

2.2 Streuungsmessung

Definition 2.10 (Varianz und Standardabweichung)

- Die Maßzahl

$$s^2 \stackrel{\text{def}}{=} \frac{1}{n}\sum_{i=1}^{n}(x_i - \bar{x})^2$$

heißt **Varianz** (engl.: *variance*) (auch mittlere quadratische Abweichung oder Streuung).

- Die Maßzahl

$$s \stackrel{\text{def}}{=} \sqrt{s^2}$$

heißt **Standardabweichung** (engl.: *standard deviation*).

2-10

Satz 2.11 (Verschiebungsdarstellung der Varianz)
Es gilt

$$s^2 = \frac{1}{n}\sum_{i=1}^{n} x_i^2 - \bar{x}^2.$$

Bemerkung 2.12 (Eigenschaften)

- Es ist $s^2 \geq 0$ und $s \geq 0$.
- Es ist $s = 0$ genau dann, wenn $x_1 = x_2 = \ldots = x_n$.

2-11

Bemerkung 2.13 (Herleitung der Verschiebungsdarstellung)

$$\begin{aligned}
\frac{1}{n}\sum_{i=1}^{n}(x_i - \bar{x})^2 &= \frac{1}{n}\sum_{i=1}^{n}(x_i^2 - 2x_i\bar{x} + \bar{x}^2) \\
&= \frac{1}{n}\sum_{i=1}^{n} x_i^2 - \frac{1}{n}\sum_{i=1}^{n} 2x_i\bar{x} + \frac{1}{n}\sum_{i=1}^{n}\bar{x}^2 \\
&= \frac{1}{n}\sum_{i=1}^{n} x_i^2 - 2\bar{x}\frac{1}{n}\sum_{i=1}^{n} x_i + \frac{1}{n}n\bar{x}^2 \\
&= \frac{1}{n}\sum_{i=1}^{n} x_i^2 - 2\bar{x}^2 + \bar{x}^2 \\
&= \frac{1}{n}\sum_{i=1}^{n} x_i^2 - \bar{x}^2
\end{aligned}$$

2.3 Ergänzungen

Beispiel 2.a (Mittlere Studienzeit) Wie misst man die "mittlere" Studienzeit? Die Antwort hängt von der inhaltlichen Fragestellung ab.

- Jeweils 50% der Studierenden haben eine längere bzw. kürzere Studienzeit. Eine Aussage über den **mittleren** Studierenden im Sinne einer Rangbildung erfolgt durch den **Median der Studienzeit**.
- Das Produkt $n\bar{x}$ ist die Gesamtstudienzeit. Das arithmetische Mittel $\bar{x}$ ist eher geeignet zu Aussagen über die durchschnittliche Auslastung und die durchschnittlichen Kosten.

Bemerkung 2.b Für das geometrische Mittel gilt

$$\ln(\bar{x}_G) = \frac{1}{n}\sum_{i=1}^{n} \ln(x_i)\,.$$

Bemerkung 2.c (Alternative Definitionen des Medians)

1. Manchmal wird der Median durch die folgende Definition festgelegt:
   Jeder Wert $x_{MED} \in \mathbb{R}$ mit
   $$\frac{|\{i|x_i \leq x_{MED}\}|}{n} \geq 1/2 \quad \text{und} \quad \frac{|\{i|x_i \geq x_{MED}\}|}{n} \geq 1/2$$
   heißt Median. Eventuell ist der Median nach dieser Definition nicht eindeutig. Der Quantilswert $\tilde{x}_{0,5}$ ist immer ein Median im Sinn dieser Definition.
2. Falls n ungerade ist, gilt $x_{MED} = x_{\frac{n+1}{2}}$.
3. Für $n = 4$ und $x_1 < x_2 < x_3 < x_4$ gilt $\tilde{x}_{0.5} = x_2$ und jeder Wert in der Menge
   $$\{x|x_2 \leq x \leq x_3\}$$
   ist ein Median.
4. Manchmal wird der Median für gerade n durch
   $$x_{MED} \stackrel{\text{def}}{=} \frac{x_{\frac{n}{2}} + x_{\frac{n}{2}+1}}{2}$$
   festgelegt. Dann ist unter Umständen $x_{MED} \neq \tilde{x}_{0.5}$. Der Median heißt dann auch **Zentralwert**.

Kapitel 3

Messzahlen und Indizes

3-1

3. Messzahlen und Indizes

3.1 Messzahlen
3.2 Preis-, Mengen- und Wertindizes

3-2

3.1 Messzahlen

Bemerkung 3.1 (Messzahl)

- Eine **Messzahl** bezieht sich auf ein Merkmal und zwei statistische Massen, die sich sachlich, räumlich oder zeitlich unterscheiden.
- Entsprechend unterscheidet man **Messzahlen des sachlichen, räumlichen** oder **zeitlichen Vergleichs**.

3-3

Beispiel 3.2 (Messzahlen)

- Das Geschlechterverhältnis am 1. 1. 2005, gemessen durch den Quotienten

$$\frac{\text{Männer in Deutschland am 1.1.2005}}{\text{Frauen in Deutschland am 1.1.2005}},$$

  ist eine **Messzahl des sachlichen Vergleichs**.

- Die Einwohnerrelation, gemessen durch den Quotienten

$$\frac{\text{Einwohner von Deutschland am 1.1.2005}}{\text{Einwohner von Frankreich am 1.1.2005}},$$

  ist eine **Messzahl des räumlichen Vergleichs**.

- Der Quotient

$$\frac{\text{Einwohner von Deutschland am 1.1.2005}}{\text{Einwohner von Deutschland am 1.1.2000}}$$

  ist eine **Messzahl des zeitlichen Vergleichs**.

3-4

Bemerkung 3.3 (Zeitreihe)

- Eine **Zeitreihe** ist eine zeitlich geordnete Folge $x_{t_0}, x_{t_1}, \ldots, x_T$ von Werten eines Merkmals X zu den Zeiten $t_0 < t_1 < \ldots < T$.

- Wenn die Differenzen jeweils aufeinanderfolgender Zeitindizes gleich sind,

$$t_1 - t_0 = t_2 - t_1 = \ldots,$$

  spricht man von **äquidistanten** Zeiten und schreibt einfach

$$x_t \quad \text{mit} \quad t = 0, 1, \ldots, T.$$

Beispiel 3.4 (Zeitreihen)

- Umsatz eines Unternehmens im Monat t.

- Arbeitslosenzahl in Deutschland am Ende des Monats t.

3-5

Definition 3.5 (Messzahlen)
Gegeben seien eine Zeitreihe x_t für $t = t_0, t_1, \ldots, T$, eine **Basiszeit** $s \in \{t_0, t_1, \ldots, T\}$ und eine **Berichtszeit** $t \in \{t_0, t_1, \ldots, T\}$, dann ist der Quotient

$$m_{s,t} \overset{\text{def}}{=} \frac{x_t}{x_s}$$

eine **Messzahl für die Berichtszeit** t **zur Basiszeit** s.

Bemerkung 3.6 (Zeitreihe von Messzahlen)

- Aus einer Zeitreihe

$$x_t \quad \text{für} \quad t = t_0, t_1, \ldots, T$$

  für das Merkmal X erhält man zu einer fixierten Basiszeit $s \in \{t_0, t_1, \ldots, T\}$ eine **Zeitreihe von Messzahlen**

$$m_{s,t} \quad \text{für} \quad t = t_0, t_1, \ldots, T.$$

- Häufig ist $s = t_0$.

3-6

Beispiel 3.7 (Messzahlen zum Basisjahr 2000)
Gegebene Daten:

$$x_{1999} = 360, \quad x_{2000} = 400, \quad x_{2001} = 420, \quad x_{2002} = 440$$

Zeitreihe der Messzahlen zum Basisjahr 2000:

$$\begin{aligned} m_{2000,1999} &= \tfrac{x_{1999}}{x_{2000}} = \tfrac{360}{400} = 0.90 \\ m_{2000,2000} &= \tfrac{x_{2000}}{x_{2000}} = \tfrac{400}{400} = 1.00 \\ m_{2000,2001} &= \tfrac{x_{2001}}{x_{2000}} = \tfrac{420}{400} = 1.05 \\ m_{2000,2002} &= \tfrac{x_{2002}}{x_{2000}} = \tfrac{440}{400} = 1.10 \end{aligned}$$

3-7

Bemerkung 3.8 (Messzahlen in Prozent)
Manchmal werden Messzahlen auch durch Multiplikation mit dem Faktor 100 so **normiert**, dass im Basisjahr der Wert 100 vorliegt, der dann als 100 Prozent interpretiert werden kann, z. B.

$$1.00 = 100\%, \quad 1.05 = 105\%, \quad 0.90 = 90\%\,.$$

3-8

Beispiel 3.9 (Umbasierung von Messzahlen)

	t	0	1	...	r	...
	x_t	150	165	...	200	...
Basiszeit alt: 0	$m_{0,t}$	1	1.10	...	$1.3\bar{3}$	...
Basiszeit neu: r	$m_{r,t}$	?	?	...	1	...

Man erhält z. B. $m_{r,1}$ direkt aus x_1 und x_r als

$$m_{r,1} = \frac{x_1}{x_r} = \frac{165}{200} = 0.825$$

oder durch Umbasierung von der Basiszeit 0 auf die neue Basiszeit r,

$$m_{r,1} = \frac{m_{0,1}}{m_{0,r}} = \frac{1.10}{1.3\bar{3}} = 0.825\,.$$

Analog erhält man

$$m_{r,0} = \frac{m_{0,0}}{m_{0,r}} = \frac{1}{1.3\bar{3}} = 0.75\,.$$

3-9

Bemerkung 3.10 (Umbasierung)

- Allgemein gilt für die Umbasierung von der Basiszeit s auf die Basiszeit r der Zusammenhang

$$m_{r,t} = \frac{x_t}{x_r} = \frac{\frac{x_t}{x_s}}{\frac{x_r}{x_s}} = \frac{m_{s,t}}{m_{s,r}}\,.$$

- Die Messzahlen $m_{r,t}$ können also aus den Messzahlen $m_{s,t}$ ohne Kenntnis der Zeitreihe x_t berechnet werden.
- Es gilt die Formel

$$m_{s,t} = m_{s,r} \cdot m_{r,t}$$

für die **Zirkularität von Messzahlen**.

3-10

Beispiel 3.11 (Verkettung von Messzahlenreihen)

- Gegeben seien zwei Folgen von Messzahlen mit verschiedenen Basiszeiten, z. B.

t	0	1	2	3	4
$m_{0,t}$	1.0	1.1	1.2		**?**
$m_{2,t}$	**?**		1.0	1.3	1.4

- Unter Verwendung der Zirkularität erhält man z. B. den fehlenden Wert

$$m_{0,4} = m_{0,2} \cdot m_{2,4} = 1.2 \cdot 1.4 = 1.68$$

und den fehlenden Wert

$$m_{2,0} = \frac{m_{0,0}}{m_{0,2}} = \frac{1}{1.2} = 0.8\bar{3}\,.$$

3-11

Bemerkung 3.12 (Verkettung)
Gegeben sind die Folgen von Messzahlen

$$m_{0,t} \quad \text{für} \quad t = 0, 1, \ldots, s$$

und

$$m_{s,t} \quad \text{für} \quad t = s, s+1, \ldots, T.$$

Gesucht sind durchgehende Folgen von Messzahlen zur Basiszeit 0 und zur Basiszeit s.

- Für die Basiszeit 0 erhält man die zusätzlichen Werte

$$m_{0,t} = m_{0,s} \cdot m_{s,t}, \quad t = s+1, \ldots, T\,.$$

- Für die Basiszeit s erhält man die zusätzlichen Werte

$$m_{s,t} = \frac{m_{0,t}}{m_{0,s}}\,, \quad t = 0, 1, \ldots, s-1\,.$$

3-12

3.2 Preis-, Mengen- und Wertindizes

Bemerkung 3.13 (Warenkorb)

- n **Güter**: $i = 1, 2, \ldots, n$ zu zwei Zeitpunkten t und s.
- $p_t(i)$: **Preis** (engl.: *price*) des Gutes i zum Zeitpunkt t
- $q_t(i)$: **Menge** (engl.: *quantity*) des Gutes i zum Zeitpunkt t
- $v_t(i) = p_t(i)q_t(i)$: **Wert** (engl.: *value*) des Gutes i zum Zeitpunkt t
- Der Wert des Warenkorbes zum Zeitpunkt t ist daher

$$\sum_{i=1}^{n} v_t(i) = \sum_{i=1}^{n} p_t(i)q_t(i) \,.$$

3-13

Definition 3.14 (Preis- und Mengenindizes)

	Preis*index*	**Mengen***index*
La*speyres*	$I^p_{La;0,t} \stackrel{\text{def}}{=} \frac{\sum_{i=1}^n p_t(i)q_0(i)}{\sum_{i=1}^n p_0(i)q_0(i)}$	$I^q_{La;0,t} \stackrel{\text{def}}{=} \frac{\sum_{i=1}^n p_0(i)q_t(i)}{\sum_{i=1}^n p_0(i)q_0(i)}$
Pa*asche*	$I^p_{Pa;0,t} \stackrel{\text{def}}{=} \frac{\sum_{i=1}^n p_t(i)q_t(i)}{\sum_{i=1}^n p_0(i)q_t(i)}$	$I^q_{Pa;0,t} \stackrel{\text{def}}{=} \frac{\sum_{i=1}^n p_t(i)q_t(i)}{\sum_{i=1}^n p_t(i)q_0(i)}$

3-14

Bemerkung 3.15

- Ein Preisindex soll Preisänderungen, ein Mengenindex soll Mengenänderungen messen.
- **Laspeyresindizes**:
  Die Mengen beim Preisindex bzw. die Preise beim Mengenindex werden aus der **Basis**periode genommen und konstant gehalten.
- **Paascheindizes**:
  Die Mengen beim Preisindex bzw. die Preise beim Mengenindex werden aus der **Berichts**periode genommen und konstant gehalten.

3-15

Beispiel 3.16 (Prinzip eines Aktienindex)

- Kurse $p_t(i)$ für $i = 1, 2, \ldots, n$. Beispielsweise enthält der DAX $n = 30$ verschiedene Aktien.
- $q_t(i)$: Stückzahlen (i = 1, 2, …, n)
- Indexkonstruktion

$$I_{0,t} = \frac{\sum_{i=1}^{n} p_t(i) q_0(i)}{\sum_{i=1}^{n} p_0(i) q_0(i)} \quad (= I^p_{La;0,t})$$

3-16

Definition 3.17 (Wertindex)

$$I^v_{0,t} = \frac{\sum\limits_{i=1}^{n} p_t(i) q_t(i)}{\sum\limits_{i=1}^{n} p_0(i) q_0(i)}$$

ist ein **Wertindex** mit Berichtszeit t zur Basiszeit 0.

Bemerkung 3.18 (Zusammenhang zwischen den Indizes)
Es gilt

$$I^v_{0,t} = I^p_{Pa;0,t} I^q_{La;0,t} = I^p_{La;0,t} I^q_{Pa;0,t} \,.$$

3-17

Bemerkung 3.19 (Wägungsschema eines Index)
Ein Preisindex nach Laspeyres kann als gewichtetes Mittel (gewogener Mittelwert) von Preismesszahlen dargestellt werden. Es gilt

$$\begin{aligned} I^p_{La;0,t} &= \frac{\sum_{i=1}^{n} p_t(i) q_0(i)}{\sum_{i=1}^{n} p_0(i) q_0(i)} = \sum_{i=1}^{n} \frac{q_0(i)}{\sum_{i=1}^{n} p_0(i) q_0(i)} \, p_t(i) \\ &= \sum_{i=1}^{n} \frac{p_0(i) q_0(i)}{\sum_{i=1}^{n} p_0(i) q_0(i)} \, \frac{p_t(i)}{p_0(i)} = \sum_{i=1}^{n} w_i m_{0,t}(i) \end{aligned}$$

mit den **Preismesszahlen**

$$m_{0,t}(i) \overset{\text{def}}{=} \frac{p_t(i)}{p_0(i)}$$

und den **Gewichten** (Umsatzanteile in der Basiszeit)

$$w_i \overset{\text{def}}{=} \frac{p_0(i) q_0(i)}{\sum_{i=1}^{n} p_0(i) q_0(i)}, \quad w_i \geq 0, \quad \sum_{i=1}^{n} w_i = 1.$$

3.3 Ergänzungen

Bemerkung 3.a (Verhältniszahlen) Allgemein ist eine **Verhältniszahl** ein **Quotient** zweier statistischer Größen. Speziell unterscheidet man Gliederungszahlen, Beziehungszahlen und Messzahlen.

1. Eine **Gliederungszahl** bezieht sich auf *ein* statistisches Merkmal und charakterisiert den Anteil einer Teilgesamtheit an einer Grundgesamtheit. Beispiele für Gliederungszahlen sind der Anteil der weiblichen Studenten an allen Studierenden und der Anteil des Einkommens der weiblichen Studenten am Gesamteinkommen aller Studierenden.

2. Eine **Beziehungszahl** bezieht sich auf zwei Merkmale und setzt sachlich verschiedenartige Größen in eine (sinnvolle) Beziehung. Beispiele für Beziehungszahlen sind die Bevölkerungsdichte, gemessen durch „Einwohner je qkm“, und PKW-Dichte, gemessen durch „Kfz pro 1000 Einwohner“.

3. Eine **Messzahl** bezieht sich auf ein Merkmal und zwei statistische Massen, die sich sachlich, räumlich oder zeitlich unterscheiden.

Kapitel 4

Auswertung mehrdimensionaler Daten

4-1

4. Auswertung mehrdimensionaler Daten

4.1 Korrelationskoeffizient
4.2 Rangkorrelation
4.3 Kontingenzkoeffizient
4.4 Deskriptive lineare Regression

4-2

4.1 Korrelationskoeffizient

Definition 4.1 (Streuungsdiagramm)
Eine zweidimensionale graphische Darstellung der Beobachtungen (x_i, y_i) für $i = 1, 2, \ldots, n$ von zwei Merkmalen X und Y an n Merkmalsträgern heißt **Streuungsdiagramm** (engl. *scatterplot*).

Definition 4.2 (Kovarianz)
Die Maßzahl

$$s_{XY} \overset{\text{def}}{=} \frac{1}{n} \sum_{i=1}^{n} (x_i - \bar{x})(y_i - \bar{y})$$

heißt **Kovarianz** (engl. *covariance*) von X und Y.

Satz 4.3 Für die Kovarianz gilt die **Verschiebungsdarstellung**

$$s_{XY} = \frac{1}{n} \sum_{i=1}^{n} x_i y_i - \bar{x}\bar{y}\,.$$

4-3

Definition 4.4 (Korrelationskoeffizient)
Gegeben seien n Beobachtungspaare (x_i, y_i) für $i = 1, \ldots, n$ für zwei Variablen X und Y mit den Standardabweichungen

$$s_X \stackrel{\text{def}}{=} \sqrt{\frac{1}{n}\sum_{i=1}^{n}(x_i - \bar{x})^2} \quad > \quad 0,$$

$$s_Y \stackrel{\text{def}}{=} \sqrt{\frac{1}{n}\sum_{i=1}^{n}(y_i - \bar{y})^2} \quad > \quad 0$$

und der Kovarianz s_{XY}. Dann heißt

$$r_{XY} \stackrel{\text{def}}{=} \frac{s_{XY}}{s_X s_Y}$$

Korrelationskoeffizient von X und Y.

4-4

Bemerkung 4.5 (Eigenschaften)

- **Symmetrie**
$$r_{XY} = r_{YX}$$
- **Normierung**
$$-1 \leq r_{XY} \leq +1$$
- **Unkorreliertheit** von X und Y
$$r_{XY} = 0 \quad \Longleftrightarrow \quad s_{XY} = 0$$
- **Exakter positiver affin-linearer Zusammenhang**
$$r_{XY} = +1 \quad \Longleftrightarrow \quad \text{für alle } i \text{ gilt} \quad y_i = a + bx_i, \quad a \in \mathbb{R}, \quad b > 0$$
- **Exakter negativer affin-linearer Zusammenhang**
$$r_{XY} = -1 \quad \Longleftrightarrow \quad \text{für alle } i \text{ gilt} \quad y_i = a + bx_i, \quad a \in \mathbb{R}, \quad b < 0$$

4-5

Bemerkung 4.6 (Skalenniveau)
Eine sinnvolle Interpretation des Korrelationskoeffizienten setzt **metrisch skalierte Daten** voraus.

Bemerkung 4.7 (Bezeichnungen)
Zur Abgrenzung von alternativen Konzepten spricht man auch von der **Produkt-Momenten-Korrelation** oder **dem Korrelationskoeffizienten** von **Bravais** und **Pearson**.

4-6

4.2 Rangkorrelation

Bemerkung 4.8 (Skalenniveau)
Das Konzept der **Rangkorrelation** setzt für zwei Merkmale mindestens **ordinal skalierte Daten** voraus.

Definition 4.9 (Bindungen)
In den Beobachtungen $x_1, x_2, \ldots, x_n$ liegen **keine Bindungen** (engl: *ties*) vor, falls alle x_i voneinander verschieden sind. Eine **Bindung** liegt vor, wenn zwei oder mehrere Beobachtungen den gleichen Wert haben.

Beispiel 4.10 (Bindungen)

$$x_1 < \boxed{x_2 = x_3 = x_4} < x_5 < \boxed{x_6 = x_7}$$

4-7

Definition 4.11 (Ränge)
Gegeben seien Beobachtungen $x_1, x_2, \ldots, x_n$ ohne Bindungen.
Falls x_i in der aufsteigend geordneten Folge der x-Werte an der r-ten Stelle steht, sind durch

$$R_X(x_i) \stackrel{\text{def}}{=} r, \quad , i = 1, \ldots, n$$

eindeutige **Rangzahlen** oder kurz **Ränge** definiert.

Beispiel 4.12 (Rangbestimmung)
Gegeben sind die $n = 3$ Beobachtungspaare $(x_1, y_1) = (15, 300)$, $(x_2, y_2) = (11, 200)$ und $(x_3, y_3) = (14, 100)$.

i	x_i	$R_X(x_i)$	y_i	$R_Y(y_i)$
1	15	3	300	3
2	11	1	200	2
3	14	2	100	1

4-8

Definition 4.13 (Rangkorrelation)
Gegeben seien $n \geq 2$ Beobachtungen ohne Bindungen (x_i, y_i) mit den Rangzahlen $(R_X(x_i), R_Y(y_i))$. Mit $\bar{R}_X$ und $\bar{R}_Y$ seien die arithmetischen Mittelwerte der Rangzahlen bezeichnet. Dann heißt der aus den Rangzahlen berechnete Korrelationskoeffizient

$$r_{XY}^R \stackrel{\text{def}}{=} \frac{\frac{1}{n}\sum\limits_{i=1}^{n}(R_X(x_i) - \bar{R}_X)(R_Y(y_i) - \bar{R}_Y)}{\sqrt{\frac{1}{n}\sum\limits_{i=1}^{n}(R_X(x_i) - \bar{R}_X)^2}\sqrt{\frac{1}{n}\sum\limits_{i=1}^{n}(R_Y(y_i) - \bar{R}_Y)^2}}$$

Rangkorrelationskoeffizient von Spearman.

4-9

Bemerkung 4.14 (Berechnung)

- Es gilt

$$\bar{R}_X = \bar{R}_Y = \frac{n+1}{2}\,.$$

- Umformungen von r^R_{XY} führen zu

$$r^R_{XY} = 1 - \frac{6\sum\limits_{i=1}^{n}(R_X(x_i) - R_Y(y_i))^2}{n(n^2-1)}\,.$$

4-10

Bemerkung 4.15 (Eigenschaften)

- **Symmetrie**

$$r^R_{XY} = r^R_{YX}$$

- **Normierung** (Wertebereich)

$$-1 \le r^R_{XY} \le 1$$

- **Vollständig gleichgerichteter Zusammenhang**

$$r^R_{XY} = +1 \quad\Longleftrightarrow\quad R_X(x_i) = R_Y(y_i) \quad \text{für alle } i$$

- **Vollständig gegenläufiger Zusammenhang**

$$r^R_{XY} = -1 \quad\Longleftrightarrow\quad R_X(x_i) = n - R_Y(y_i) + 1 \quad \text{für alle } i$$

- r^R_{XY} ist ein **Maß des monotonen Zusammenhangs**.

4-11

4.3 Kontingenzkoeffizient

Gegeben seien n Beobachtungen $(x_1, y_1), \ldots, (x_n, y_n)$ von zwei **kategorialen Merkmalen** X und Y mit den möglichen **Merkmalskombinationen**

$$(\xi_j, \eta_k) \qquad j = 1, \ldots, J, \quad k = 1, \ldots, K.$$

Definition 4.16 (Gemeinsame Häufigkeiten)
Durch

$$n_{jk} \stackrel{\text{def}}{=} |\{i | (x_i, y_i) = (\xi_j, \eta_k)\}|, \quad j = 1, \ldots, J, \quad k = 1, \ldots, K$$

sind die **gemeinsamen absoluten Häufigkeiten** und durch

$$f_{jk} \stackrel{\text{def}}{=} \frac{n_{jk}}{n}, \quad j = 1, \ldots, J, \quad k = 1, \ldots, K$$

die **gemeinsamen relativen Häufigkeiten** definiert.

4-12

Definition 4.17 (Kontingenztabelle)
Eine zweidimensionale Tabelle der gemeinsamen absoluten Häufigkeiten n_{jk} oder der gemeinsamen relativen Häufigkeiten f_{jk} heißt **Kontingenztabelle**.

Definition 4.18 (Randhäufigkeiten)
Die Häufigkeiten

$$n_{j\cdot} \stackrel{\text{def}}{=} \sum_{k=1}^{K} n_{jk}, \quad n_{\cdot k} \stackrel{\text{def}}{=} \sum_{j=1}^{J} n_{jk}$$

heißen **absolute Randhäufigkeiten**. Durch

$$f_{j\cdot} \stackrel{\text{def}}{=} \sum_{k=1}^{K} f_{jk}, \quad f_{\cdot k} \stackrel{\text{def}}{=} \sum_{j=1}^{J} f_{jk}$$

sind die **relativen Randhäufigkeiten** gegeben.

4-13

Aufbau einer Kontingenztabelle

		Merkmal Y						
		η_1	η_2	$\dots$	η_k	$\dots$	η_K	$\sum_{k=1}^{K}$
	ξ_1	n_{11}	n_{12}	$\dots$	n_{1k}	$\dots$	n_{1K}	$n_{1\cdot}$
	ξ_2	n_{21}	n_{22}	$\dots$	n_{2k}	$\dots$	n_{2K}	$n_{2\cdot}$
	$\dots$	$\dots$	$\dots$	$\dots$	$\dots$	$\dots$	$\dots$	$\dots$
Merkmal X	ξ_j	n_{j1}	n_{j2}	$\dots$	n_{jk}	$\dots$	n_{jK}	$n_{j\cdot}$
	$\dots$	$\dots$	$\dots$	$\dots$	$\dots$	$\dots$	$\dots$	$\dots$
	ξ_J	n_{J1}	n_{J2}	$\dots$	n_{Jk}	$\dots$	n_{JK}	$n_{J\cdot}$
	$\sum_{j=1}^{J}$	$n_{\cdot 1}$	$n_{\cdot 2}$	$\dots$	$n_{\cdot k}$	$\dots$	$n_{\cdot K}$	n

4-14

Bemerkung 4.19
Es gilt

$$n = \sum_{j=1}^{J}\sum_{k=1}^{K} n_{jk} = \sum_{j=1}^{J} n_{j\cdot} = \sum_{k=1}^{K} n_{\cdot k}.$$

Definition 4.20 (Chiquadratmaßzahl)
Die Maßzahl

$$\chi^2 \stackrel{\text{def}}{=} \sum_{j=1}^{J}\sum_{k=1}^{K} \frac{(n_{jk} - \tilde{n}_{jk})^2}{\tilde{n}_{jk}} \quad \text{mit} \quad \tilde{n}_{jk} \stackrel{\text{def}}{=} \frac{n_{j\cdot}}{n} \cdot \frac{n_{\cdot k}}{n} \cdot n = \frac{n_{j\cdot} n_{\cdot k}}{n}$$

heißt **Chiquadratmaßzahl**.

4-15

Definition 4.21 (Deskriptive Unabhängigkeit)
Die zwei Variablen X und Y heißen **deskriptiv unabhängig**, wenn für alle $j = 1, \ldots, J$ und $k = 1, \ldots, K$ die Beziehung

$$n_{jk} = \frac{n_{j\cdot} n_{\cdot k}}{n}$$

gilt.

Bemerkung 4.22
Bei deskriptiver Unabhängigkeit gilt

$$\frac{n_{jk}}{n} = \frac{n_{j\cdot}}{n} \cdot \frac{n_{\cdot k}}{n},$$

$$n_{jk} = \tilde{n}_{jk}$$

und

$$\chi^2 = 0\,.$$

4-16

Definition 4.23
Es sei $\min\{J, K\} \geq 2$. Dann heißt

$$C = C_{XY} = \sqrt{\frac{\chi^2}{n + \chi^2} \cdot \frac{\min\{J, K\}}{\min\{J, K\} - 1}}$$

Kontingenzkoeffizient.

Bemerkung 4.24

- Es gilt
$$0 \leq C \leq 1\,.$$
- Bei deskriptiver Unabhängigkeit gilt $C = 0$.

4-17

4.4 Deskriptive lineare Regression

Bemerkung 4.25 (Problemstellung)

- Es sind Beobachtungen (x_i, y_i) für $i = 1, \ldots, n$ gegeben.
- Gesucht ist ein Zusammenhang
$$y = f(x)\,.$$
- Dabei heißen y **Regressand**, x **Regressor** und f **Regressionsfunktion**.

Bemerkung 4.26 (Alternative Bezeichnungen)
Anstelle von Regressand und Regressor spricht man auch von

- **abhängiger** und **unabhängiger** Variable,
- **erklärter** und **erklärender** Variable,
- **endogener** und **exogener** Variable.

4-18

Bemerkung 4.27

- Wenn von einer linearen Regressionsfunktion die Rede ist, ist im Allgemeinen eine Funktion der Form

$$y = a + bx$$

gemeint.

- Im Fall $a = 0$ spricht man von einer homogenen, anderenfalls von einer inhomogenen Regression.
- Die Funktion $y = a+bx$ nennt man auch **affin-linear** zur Unterscheidung von der (im engeren Sinn) linearen Funktion $y = bx$.

4-19

Bemerkung 4.28 (Lineare Regressionsgerade)
a ist der **Achsenabschnitt** (das Absolutglied, die Inhomogenität) der Geraden $y = a + bx$.

$$b = \frac{dy}{dx}$$

ist die **Steigung** der Geraden. a und b, häufig aber nur b, heißen **Regressionskoeffizienten**. Eine Gerade

$$y = a + bx,$$

die sich den Beobachtungen

$$(x_i, y_i) \qquad (i = 1, 2, \ldots, n)$$

gut anpasst, eine sogenannte **Ausgleichsgerade**, erhält man mit der **Methode der kleinsten Quadrate**.

4-20

Bemerkung 4.29 (Methode der kleinsten Quadrate)

- Man bestimmt a und b so, dass die Quadratsumme

$$Q(\alpha, \beta) \overset{\text{def}}{=} \sum_{i=1}^{n} (y_i - (\alpha + \beta x_i))^2$$

minimal ist, d.h.

$$Q(a, b) = \min_{\alpha,\beta} Q(\alpha, \beta).$$

- Falls nicht alle x_i identisch sind, ist die Lösung der Optimierungsaufgabe eindeutig und durch

$$b = \frac{s_{XY}}{s_X^2} \tag{4.1}$$

und

$$a = \bar{y} - b\bar{x} \tag{4.2}$$

gegeben.

4-21

Bemerkung 4.30

- Es gilt
$$b = r_{XY} \frac{s_Y}{s_X} \, .$$
- Wegen (4.2) gilt
$$\bar{y} = a + b\bar{x}. \tag{4.3}$$
Die nach der Methode der kleinsten Quadrate bestimmte Regressionsgerade geht also durch den Punkt $(\bar{x}, \bar{y})$.

Definition 4.31 (Erklärte y-Werte)
Durch
$$\hat{y}_i \stackrel{\text{def}}{=} a + bx_i \quad (i = 1, 2, \ldots, n) \tag{4.4}$$
sind die **durch die Regression erklärten** y-Werte definiert.

4-22

Bemerkung 4.32

- Aus (4.4) folgt
$$\bar{\hat{y}} = \frac{1}{n} \sum_{i=1}^{n} \hat{y}_i = a + b\bar{x}, \tag{4.5}$$
und
$$s_{\hat{Y}}^2 = \frac{1}{n} \sum_{i=1}^{n} (\hat{y}_i - \bar{\hat{y}})^2 = b^2 s_X^2 \, .$$
- Durch Vergleich von (4.3) mit (4.5) erhält man
$$\bar{\hat{y}} = \bar{y} \tag{4.6}$$
und damit
$$s_{\hat{Y}}^2 = \frac{1}{n} \sum_{i=1}^{n} (\hat{y}_i - \bar{y})^2 \, .$$

4-23

Definition 4.33 (Residuen)
Durch
$$u_i \stackrel{\text{def}}{=} y_i - \hat{y}_i \quad (i = 1, 2, \ldots, n) \tag{4.7}$$
sind die **Residuen** (Singular: Residuum) der Regression definiert.

Bemerkung 4.34

- Aus (4.7) erhält man
$$\bar{u} = \frac{1}{n} \sum_{i=1}^{n} u_i = \bar{y} - \bar{\hat{y}}$$
- Wegen $\bar{\hat{y}} = \bar{y}$, vgl. (4.6), gilt
$$\bar{u} = 0$$
und damit
$$s_U^2 = \frac{1}{n} \sum_{i=1}^{n} u_i^2 \, .$$

4-24

Satz 4.35 (Varianzzerlegungssatz)
Es gilt

$$s_Y^2 = s_{\hat{Y}}^2 + s_U^2.$$

Definition 4.36
Im Zusammenhang mit der Varianzzerlegung heißen s_Y^2 **Gesamtvarianz**, $s_{\hat{Y}}^2$ **erklärte Varianz** und s_U^2 **Residualvarianz**.

Definition 4.37 (Bestimmtheitsmaß)
Die Maßzahl

$$R^2 \stackrel{\text{def}}{=} \frac{s_{\hat{Y}}^2}{s_Y^2} = 1 - \frac{s_U^2}{s_Y^2}$$

heißt **Bestimmtheitsmaß** oder **Determinationskoeffizient**.

4-25

Bemerkung 4.38
Es gilt

$$R^2 = r_{XY}^2 \in [0, 1]$$

und damit

$$r_{XY} = \sqrt{R^2} \quad \text{oder} \quad r_{XY} = -\sqrt{R^2}\,.$$

Bemerkung 4.39 (Extremfälle)

- $s_Y^2 = 0$, d. h. $y_1 = y_2 = \ldots = y_n$. Eine Regression ist nicht sinnvoll und R^2 ist nicht definiert.
- $s_Y^2 > 0$ und alle Punkte (x_i, y_i) liegen auf der Regressionsgeraden:

$$\hat{y}_i = y_i \quad (i = 1, 2, \ldots, n) \quad \Rightarrow \quad s_U^2 = 0, \quad R^2 = 1$$

- $s_Y^2 > 0$ und alle $\hat{y}_i$ sind gleich; kein erklärter Varianzanteil:

$$\hat{y}_i = \bar{y}, \quad b = 0, \quad a = \bar{y} \quad \Rightarrow \quad s_{\hat{Y}}^2 = 0, \quad R^2 = 0$$

4.5 Ergänzungen

Bemerkung 4.a (Herleitung der Verschiebungsdarstellung)

$$\begin{aligned}
s_{XY} &= \frac{1}{n}\sum_{i=1}^{n}(x_i - \bar{x})(y_i - \bar{y}) \\
&= \frac{1}{n}\sum_{i=1}^{n}(x_i y_i - x_i\bar{y} - \bar{x}y_i + \bar{x}\bar{y}) \\
&= \frac{1}{n}\sum_{i=1}^{n} x_i y_i - \frac{1}{n}\sum_{i=1}^{n} x_i\bar{y} - \frac{1}{n}\sum_{i=1}^{n} y_i\bar{x} + \frac{1}{n}\sum_{i=1}^{n}\bar{x}\bar{y} \\
&= \frac{1}{n}\sum_{i=1}^{n} x_i y_i - \bar{x}\bar{y} - \bar{x}\bar{y} + \bar{x}\bar{y} \\
&= \frac{1}{n}\sum_{i=1}^{n} x_i y_i - \bar{x}\bar{y}
\end{aligned}$$

Bemerkung 4.b (Berechnung des Korrelationskoeffizienten) Zur Berechnung des Korrelationskoeffizienten kann eine der folgenden äquivalenten Darstellungen verwendet werden:

$$
\begin{aligned}
r_{XY} &= \frac{\frac{1}{n}\sum_{i=1}^{n}(x_i-\bar{x})(y_i-\bar{y})}{\sqrt{\frac{1}{n}\sum_{i=1}^{n}(x_i-\bar{x})^2}\sqrt{\frac{1}{n}\sum_{i=1}^{n}(y_i-\bar{y})^2}} \\
&= \frac{\sum_{i=1}^{n}(x_i-\bar{x})(y_i-\bar{y})}{\sqrt{\sum_{i=1}^{n}(x_i-\bar{x})^2}\sqrt{\sum_{i=1}^{n}(y_i-\bar{y})^2}} \\
&= \frac{\sum_{i=1}^{n}x_iy_i-n\bar{x}\bar{y}}{\sqrt{\sum_{i=1}^{n}x_i^2-n\bar{x}^2}\sqrt{\sum_{i=1}^{n}y_i^2-n\bar{y}^2}}\,.
\end{aligned}
$$

Bemerkung 4.c (Berechnung des Regressionskoeffizienten) Es gilt

$$s_{XY} = \frac{1}{n}\sum_{i=1}^{n}(x_i-\bar{x})(y_i-\bar{y}) \quad \text{und} \quad s_X^2 = \frac{1}{n}\sum_{i=1}^{n}(x_i-\bar{x})^2,$$

so dass b aus (4.1) – nach Kürzen von $1/n$ – auch als

$$b = \frac{\sum_{i=1}^{n}(x_i-\bar{x})(y_i-\bar{y})}{\sum_{i=1}^{n}(x_i-\bar{x})^2}$$

geschrieben werden kann.

Bemerkung 4.d (Zur Rangkorrelation)

1. Für zwei streng monoton zunehmende Funktionen $g_1(x)$ und $g_2(y)$ haben die transformierten Daten
$$x_i^* \stackrel{\text{def}}{=} g_1(x_i), \quad y_i^* \stackrel{\text{def}}{=} g_2(y_i)$$
dieselben Rangzahlen
$$R_X(x_i^*) = R_X(x_i), \quad R_Y(y_i^*) = R_Y(y_i).$$
2. Daher berechnet sich aus den transformierten Daten (x_i^*, y_i^*) für $i = 1, \ldots, n$ derselbe Rangkorrelationskoeffizient wie aus den untransformierten Daten (x_i, y_i) für $i = 1, \ldots, n$.
3. Diese **Invarianz des Rangkorrelationskoeffizienten gegenüber streng monoton zunehmenden Transformationen** kennzeichnet die Eignung dieser Maßzahl für ordinale Merkmale.
4. Wenn die Rangzahlen nicht nach aufsteigender, sondern **absteigender Ordnung** vergeben werden, dann ist
$$\tilde{R}_X(x_i) = n + 1 - R_X(x_i)$$
die Rangzahl bei absteigender Ordnung, wobei $R_X(x_i)$ die Rangzahl bei aufsteigender Ordnung bezeichnet, und es gilt
$$(\tilde{R}_X(x_i) - \tilde{R}_Y(y_i))^2 = (R_X(x_i) - R_Y(y_i))^2,$$
so dass sich **derselbe Rangkorrelationskoeffizient** ergibt.
5. Für Daten mit **Bindungen** sind in der Literatur Modifikationen von r_{XY}^R vorgeschlagen worden.

Bemerkung 4.e (Zur Kontingenz)

1. **Kontingent** (engl.: *contingent*) bedeutet zufallsbedingt, zufällig; möglich, aber nicht notwendig. **Kontingenz** (engl.: *contingency*) bedeutet Zufälligkeit, Abhängigkeit vom Zufall, Möglichkeit.
2. Der Koeffizient C wird manchmal auch als **korrigierter Kontingenzkoeffizient** bezeichnet. Der **unkorrigierte Kontingenzkoeffizient** ist dann
$$\sqrt{\frac{\chi^2}{n+\chi^2}} \leq \sqrt{\frac{\min\{J,K\}-1}{\min\{J,K\}}} < 1\,.$$

Kapitel 5

Zeitreihenanalyse

5-1

5. Zeitreihenanalyse

5.1 Komponentenmodelle
5.2 Bestimmung der glatten Komponente durch gleitende Durchschnitte
5.3 Bestimmung der Saisonkomponente

5-2

Bemerkung 5.1

- **Finanzmarktzeitreihen**, z. B. Aktienkurse, Wechselkurse, Zinssätze, sind in der Regel irregulär, ohne saisonale Muster und werden in der Regel durch **stochastische Modelle** dargestellt.
- Ökonomische Umsatzreihen, Reihen von Arbeitslosenzahlen etc. sind saisonalen Schwankungen unterworfen und werden häufig durch deskriptive **Komponentenmodelle** dargestellt.

5-3

5.1 Komponentenmodelle

Bemerkung 5.2 (Komponenten einer Zeitreihe)

	Trendkomponente	langfristig, “glatt“, relativ gleichförmig
	Konjunkturkomponente oder zyklische Komponente	mittelfristig, zyklisch, Zykluslänge: 3 bis 7 Jahre
g_i	**glatte** Komponente	Trend + Konjunktur
s_i	**Saison**komponente	zyklisch, Jahreszeiteneinfluss, Wetter, Ferienzeiten, Festtage, Erntezyklen, ...
u_i	**Restkomponente** oder irreguläre Komponente	irreguläre Einflüsse, Meßfehler, Zufallsfehler

5-4

Bemerkung 5.3 (Komponentenmodelle)

- Additives Komponentenmodell:

$$y_i = g_i + s_i + u_i, \quad i = 1, \ldots, n$$

- Multiplikatives Komponentenmodell:

$$y_i = g_i \cdot s_i \cdot u_i, \quad i = 1, \ldots, n$$

5-5

5.2 Bestimmung der glatten Komponente durch gleitende Durchschnitte

Bemerkung 5.4 (Annahmen)

- Es liegen Beobachtungen y_i zu den Zeitpunkten t_i für $i = 1, \ldots, n$ vor.
- Es liegt ein additives Komponentenmodell vor.
- Es wird vorausgesetzt, dass die Zeitpunkte t_i **äquidistant** sind.

5-6

Definition 5.5 (Gleitender Durchschnitt der Ordnung λ)

- Gleitender Durchschnitt der Ordnung $\lambda = 3$:

$$\tilde{g}_i \stackrel{\text{def}}{=} \frac{y_{i-1} + y_i + y_{i+1}}{3}$$

- Gleitender Durchschnitt **ungerader Ordnung** $\lambda = 2l+1$ für $l = 1, 2, \ldots$:

$$\begin{aligned} \tilde{g}_i &\stackrel{\text{def}}{=} \frac{y_{i-l} + \ldots + y_{i-2} + y_{i-1} + y_i + y_{i+1} + y_{i+2} + \ldots + y_{i+l}}{2l+1} \\ &= \frac{1}{2l+1} \sum_{h=-l}^{l} y_{i+h} \end{aligned}$$

5-7

- Gleitender Durchschnitt der Ordnung $\lambda = 2$:

$$\tilde{g}_i \stackrel{\text{def}}{=} \frac{\frac{1}{2}y_{i-1} + y_i + \frac{1}{2}y_{i+1}}{2} = \frac{1}{4}y_{i-1} + \frac{1}{2}y_i + \frac{1}{4}y_{i+1}$$

- Gleitender Durchschnitt **gerader** Ordnung $\lambda = 2l$ für $l = 1, 2, \ldots$:

$$\tilde{g}_i \stackrel{\text{def}}{=} \frac{\frac{1}{2}y_{i-l} + \sum\limits_{h=-l+1}^{l-1} y_{i+h} + \frac{1}{2}y_{i+l}}{2l}$$

Bemerkung 5.6

- Im Fall einer ungeraden Ordnung ist $\tilde{g}_i$ ein arithmetisches Mittel.
- Im Fall einer geraden Ordnung ist $\tilde{g}_i$ ein gewichtetes Mittel.

5-8

Bemerkung 5.7 (Schätzung der glatten Komponente)

- Im Fall von **Monatsdaten** (12 Beobachtungen pro Jahr) wird die glatte Komponente durch einen gleitenden Durchschnitt der Ordnung $\lambda = 12$ geschätzt.
- Im Fall von **Vierteljahresdaten** (4 Beobachtungen pro Jahr) wird die glatte Komponente durch einen gleitenden Durchschnitt der Ordnung $\lambda = 4$ geschätzt.
- Durch die Glättung tritt eine **Reihenverkürzung** ein. Bei einer Durchschnittsbildung der Ordnung 12 (bzw. 4) gehen jeweils zu Anfang der Beobachtungen und zu Ende der Beobachtungen 6 (bzw. 2) Werte verloren.

5-9

5.3 Bestimmung der Saisonkomponente

Bemerkung 5.8 (Modellannahmen)

- **Äquidistante** Zeitpunkte $t_1 < t_2 < \ldots < t_n$.
- Annahme eines **additiven Komponentenmodells** der Form

$$y_i = g_i + s_i + u_i, \quad i = 1, \ldots, n\,.$$

mit **konstanter Saisonfigur**. Es gibt eine kleinste Zahl K, die sogenannte **Periodenlänge**, so dass für die Saisonkomponente

$$s_{k+j\cdot K} = s_k, \quad k = 1, \ldots, K, \quad j = 0, 1, 2, \ldots$$

gilt. Außerdem gilt die Normierungseigenschaft

$$\sum_{k=1}^{K} s_k = 0\,.$$

5-10

Bemerkung 5.9

- Es gibt nur K sich periodisch wiederholende Werte für die Saisonkomponente, so dass nur $s_1, s_2, \ldots, s_K$ bestimmt werden müssen.
- Bei Monatsdaten ist $K = 12$, bei Vierteljahresdaten ist $K = 4$.

5-11

Bemerkung 5.10 (Phasendurchschnittsverfahren)

1. Bestimmung der glatten Komponente $\tilde{g}_i$ durch gleitende Durchschnitte der Länge $\lambda = K$ und Berechnung der **trendbereinigten Reihe**

$$d_i = y_i - \tilde{g}_i\,.$$

2. Für jeden Wert der Saisonphase k bestimmt man den **Rohwert**

$$\bar{d}_k = \frac{1}{J_k} \sum_j d_{k+j\cdot K} \quad k = 1, \ldots, K\,,$$

wobei über die J_k verfügbaren Werte $d_{k+j\cdot K}$ summiert wird.

3. Berechnung der **normierten** Schätzwerte $\hat{s}_k$ für s_k,

$$\hat{s}_k = \bar{d}_k - \frac{1}{K} \sum_{i=1}^{K} \bar{d}_i, \quad k = 1, \ldots, K\,.$$

5-12

Bemerkung 5.11

- Für die Rohwerte $\bar{d}_k$ gilt im allgemeinen nicht

$$\sum_{k=1}^{K} \bar{d}_k = 0\,.$$

- Die normierten Schätzwerte $\hat{s}_k$ erfüllen die Bedingung

$$\sum_{k=1}^{K} \hat{s}_k = 0\,.$$

- Hier wird im Unterschied zu Mosler/Schmid die Bezeichnung J_k verwendet, da J_k mit k variiert. J_k hängt von der Anzahl n der Ausgangswerte und der Anzahl der durch die gleitende Durchschnittsbildung fehlenden Werte zu Beginn und Ende der Zeitreihe ab.

5-13

Bemerkung 5.12 (Saisonbereinigung)
Aus der Ausgangsreihe

$$y_1, y_2, \ldots, y_n$$

und den Schätzwerten

$$\hat{s}_1, \hat{s}_2, \ldots, \hat{s}_K$$

für eine konstante Saisonfigur erhält man die **saisonbereinigten Zeitreihenwerte**

$$y_i^s = y_i - \hat{s}_k$$

für

$$i = k + j \cdot K, \quad k = 1, \ldots, K \quad \text{und} \quad j = 0, 1, \ldots\,.$$

5.4 Ergänzungen

Bemerkung 5.a (Trendkoeffizienten)
Die glatte Komponente kann alternativ durch einen **linearen Trend** bestimmt werden, dabei ist

$$g_i = a + bt_i, \quad i = 1, \ldots, n$$

unterstellt. Die **Trendkoeffizienten** a und b können nach der Methode der kleinsten Quadrate bestimmt werden. Es handelt sich um eine lineare Regression, wobei die Zeiten t_i die Werte des Regressors sind. Es ergibt sich dann

$$b = \frac{\sum\limits_{i=1}^{n} t_i y_i - n\bar{t}\bar{y}}{\sum\limits_{i=1}^{n} t_i^2 - n\bar{t}^2}$$

und

$$a = \bar{y} - b\bar{t}$$

mit

$$\bar{y} = \frac{1}{n} \sum_{i=1}^{n} y_i \quad \text{und} \quad \bar{t} = \frac{1}{n} \sum_{i=1}^{n} t_i\,.$$

Kapitel 6

Zufallsexperimente und Wahrscheinlichkeit

6-1

6. Zufallsexperimente und Wahrscheinlichkeit

6.1 Zufallsexperimente
6.2 Wahrscheinlichkeit
6.3 Unabhängigkeit von Ereignissen

6-2

Bemerkung 6.1 (Mengentheoretische Grundbegriffe)

- x ist ein **Element** (*element*) der **Menge** (*set*) A:

$$x \in A$$

- A ist eine **Teilmenge** (*subset*) von B:

$$A \subset B \overset{\text{def}}{\iff} (x \in A \implies x \in B)$$

- **Leere Menge** (*empty set, null set*):

$$\emptyset \quad (\text{auch } \{\})$$

- Die Menge aller Teilmengen einer Menge A heißt **Potenzmenge** von A (*power set of A*):

$$\mathcal{P}(A) \overset{\text{def}}{=} \{B \mid B \subset A\}$$

Eine alternative Bezeichnung für die Potenzmenge ist 2^A.

6-3

- **Durchschnitt** (*intersection*) der Mengen A und B:

$$A \cap B \stackrel{\text{def}}{=} \{x \,|\, x \in A, x \in B\} = \{x \,|\, x \in A \textbf{ und } x \in B\}$$

- **Durchschnitt** der Mengen $A_1, A_2, \ldots, A_n$:

$$\bigcap_{i=1}^{n} A_i \stackrel{\text{def}}{=} A_1 \cap A_2 \cap \cdots \cap A_n = \{x \,|\, x \in A_i \text{ für } i = 1, \ldots, n\}$$

- **Durchschnitt** der Mengenfolge $A_1, A_2, \ldots$:

$$\bigcap_{i=1}^{\infty} A_i \stackrel{\text{def}}{=} \{x \,|\, x \in A_i \text{ für } i = 1, 2, \ldots\}$$

6-4

- **Vereinigung** (*union*) der Mengen A und B:

$$A \cup B \stackrel{\text{def}}{=} \{x \,|\, x \in A \textbf{ oder } x \in B\}$$

- **Vereinigung** der Mengen $A_1, A_2, \ldots, A_n$:

$$\bigcup_{i=1}^{n} A_i \stackrel{\text{def}}{=} A_1 \cup A_2 \cup \cdots \cup A_n$$

- **Vereinigung** der Mengenfolge $A_1, A_2, \ldots$:

$$\bigcup_{i=1}^{\infty} A_i \stackrel{\text{def}}{=} \{x \,|\, x \in A_i \text{ für mindestens ein } i \in \{1, 2, \ldots\}\}$$

6-5

- **Differenzmenge** (*set difference*):

$$A \setminus B \stackrel{\text{def}}{=} \{x \,|\, x \in A, x \notin B\}$$

- **Komplementmenge** (*complement*) von A bzgl. Ω mit $A \subset \Omega$:

$$\bar{A} \stackrel{\text{def}}{=} \Omega \setminus A$$

- Die Mengen A und B sind **disjunkt** (*disjoint, mutually exclusive*) genau dann, wenn
$$A \cap B = \emptyset \,.$$

- **Kartesisches Produkt** (*Cartesian product*) von zwei Mengen A und B:

$$A \times B \stackrel{\text{def}}{=} \{(x, y) | x \in A, y \in B\}$$

6-6

Bemerkung 6.2 (Spezielle Mengen)

- **Menge der reellen Zahlen** (*set of real numbers*): $\mathbb{R}$
- **Menge der natürlichen Zahlen** (*set of natural numbers*):

$$\mathbb{N} \overset{\text{def}}{=} \{1, 2, \ldots\}$$

- **Intervalle** (*intervals*)

$$]a, b[\overset{\text{def}}{=} \{x \in \mathbb{R} \mid a < x < b\}$$

$$[a, b] \overset{\text{def}}{=} \{x \in \mathbb{R} \mid a \leq x \leq b\}$$

$$]-\infty, b\,] \overset{\text{def}}{=} \{x \in \mathbb{R} \mid x \leq b\}$$

6-7

6.1 Zufallsexperimente

Bemerkung 6.3 (Zufallsexperiment)
Beispiele für Zufallsexperimente (*random experiments*) oder zufällige Versuche:

- Lostrommel (Lotto), Glücksrad, Roulette,...
- Würfelwurf, Münzwurf, Werfen einer Reißzwecke
- Messungen bei geplanten Experimenten in den Naturwissenschaften (zufälliger Messfehler)
- Zählungen und Erhebungen (Erhebungsfehler, Antwortfehler,...)
- Sozialwissenschaftliche Befragung (Stichprobenfehler, Antwortvariabilität,...)

6-8

Bemerkung 6.4 (Eigenschaften eines Zufallsexperiments)

1. Alle möglichen Ergebnisse ω des Zufallsexperimentes sind *a priori* bekannt und können durch die Ergebnismenge Ω beschrieben werden.
2. Der Experimentausgang ist unbekannt, aber den Ereignissen sind Wahrscheinlichkeiten zugeordnet.
3. Das Experiment ist (zumindest gedanklich) wiederholbar.

Bemerkung 6.5 (Ergebnismenge)

- Die **Ergebnismenge** (oder Grundgesamtheit) Ω ist die Menge aller möglichen **Ergebnisse** (*outcomes*) $\omega_1, \omega_2, \omega_3, \ldots$ eines Zufallsexperimentes.
- Im Folgenden wird die Ergebnismenge Ω immer als nichtleere Menge vorausgesetzt.

6-9

Definition 6.6 (Ereignisse)
Gegeben sei eine Ergebnismenge Ω.

- Teilmengen von Ω heißen **Ereignisse** (*events*), $A, B, \ldots \subset \Omega$.
- Ω heißt **sicheres Ereignis** (*sure event*).
- Die leere Menge $\emptyset$ heißt **unmögliches Ereignis**.
- Eine einelementige Teilmenge (*singleton*) von Ω, $\{\omega\} \subset \Omega$ mit $\omega \in \Omega$, heißt **Elementarereignis**.

Bemerkung 6.7

- Es gilt $\emptyset \subset \Omega$ und $\Omega \subset \Omega$.
- Zu unterscheiden sind die **Ergebnisse** $\omega \in \Omega$ und die **Elementarereignisse** $\{\omega\} \subset \Omega$.
  Z. B. gilt $1 \in \mathbb{N}$, $\{1\} \subset \mathbb{N}$ und $\{1\} \in \mathcal{P}(\mathbb{N})$.

6-10

Definition 6.8 (Ereignissystem)
Gegeben sei eine Ergebnismenge Ω. Eine nichtleere Teilmenge $\mathcal{A}$ von $\mathcal{P}(\Omega)$, der Potenzmenge von Ω, heißt **Ereignissystem** (System von Ereignissen).

Bemerkung 6.9

- $\mathcal{A}$ ist ebenfalls eine Menge.
- Die Elemente von $\mathcal{A}$ sind Teilmengen von Ω.

6-11

Beispiel 6.10 (Würfelwurf)
Für $i = 1, \ldots, 6$ bezeichne ω_i das **Ergebnis**: "i Augen liegen oben".

- Ergebnismenge: $\Omega = \{\omega_1, \omega_2, \ldots, \omega_6\}$
- Die Potenzmenge enthält $2^6 = 64$ Elemente, d. h. Teilmengen von Ω:

$$\mathcal{P}(\Omega) = \{A \mid A \subset \Omega\} = \left\{ \begin{array}{l} \emptyset, \{\omega_1\}, \{\omega_2\}, \ldots, \{\omega_6\}, \\ \{\omega_1, \omega_2\}, \{\omega_1, \omega_3\}, \ldots, \{\omega_5, \omega_6\}, \\ \{\omega_1, \omega_2, \omega_3\}, ..., \{\omega_4, \omega_5, \omega_6\}, \\ \cdots, \\ \{\omega_1, \omega_2, \ldots, \omega_6\} \end{array} \right\}$$

- $\{\omega_6\}$ ist das Ereignis „Die Augenzahl sechs wurde gewürfelt".
- $\{\omega_2, \omega_4, \omega_6\}$ ist das Ereignis „Eine gerade Augenzahl wurde gewürfelt".

6-12

Beispiel 6.11 (Fairer Würfel)
(fair bedeutet hier: gleichwahrscheinliche Elementarereignisse)

$$P(\{\omega_i\}) = \frac{1}{6}, \qquad i = 1, \ldots, 6$$

$$P(\{\omega_2, \omega_4, \omega_6\}) = P(\{\omega_2\}) + P(\{\omega_4\}) + P(\{\omega_6\}) = \frac{1}{6} + \frac{1}{6} + \frac{1}{6} = \frac{1}{2}$$

„Definition“ der Wahrscheinlichkeit nach Laplace:

$$P(\{\omega_2, \omega_4, \omega_6\}) = \frac{\text{Anzahl der günstigen Fälle}}{\text{Anzahl der möglichen Fälle}} = \frac{|\{\omega_2, \omega_4, \omega_6\}|}{|\Omega|} = \frac{3}{6}$$

6-13

6.2 Wahrscheinlichkeit

Bemerkung 6.12 (Axiomensystem von Kolmogoroff)

- Axiom 1 (**Nichtnegativität**): $P(A) \geq 0$ gilt für jedes Ereignis A.
- Axiom 2 (**Normierung**):

$$P(\Omega) = 1$$

- Axiom 3 (**Additivität**):

$$A \cap B = \emptyset \quad \Longrightarrow \quad P(A \cup B) = P(A) + P(B)$$

- Axiom 3' (σ-Additivität, Totaladditivität):

$$P\left(\bigcup_{i=1}^{\infty} A_i\right) = \sum_{i=1}^{\infty} P(A_i), \quad \text{falls} \quad A_i \cap A_j = \emptyset \quad \text{für alle} \quad i \neq j$$

6-14

Bemerkung 6.13

- Im Folgenden werden stets die Axiome 1, 2 und 3' für eine **Wahrscheinlichkeit** (*probability*) P unterstellt.
- Das Axiom 3 folgt aus dem Axiom 3', siehe unten Satz 6.17.
- Formal ist P eine reellwertige Funktion, deren Definitionsbereich ein **Ereignissystem** $\mathcal{A} \subset \mathcal{P}(\Omega)$ ist. Damit sinnvoll Wahrscheinlichkeiten berechnet werden können, muss $\mathcal{A}$ genügend viele Ereignisse enthalten, andererseits ist $\mathcal{P}(\Omega)$ häufig zu groß.

Definition 6.14 (Komplementäres Ereignis)
Zu jedem Ereignis $A \in \mathcal{A}$ heißt

$$\bar{A} \overset{\text{def}}{=} \Omega \setminus A$$

das zu A **komplementäre** Ereignis oder Komplementärereignis.

6-15

Bemerkung 6.15
Alle Aussagen der Wahrscheinlichkeitsrechnung (oder Wahrscheinlichkeitsalgebra) ergeben sich als logische, beweisbare Folgerungen aus den Axiomen.

6-16

Satz 6.16 (Wahrscheinlichkeit des unmöglichen Ereignisses)
Das unmögliche Ereignis hat die Wahrscheinlichkeit Null, d. h. es gilt

$$P(\emptyset) = 0.$$

Satz 6.17 (Axiom 3)
Für zwei Ereignisse A und B gilt

$$A \cap B = \emptyset \quad \Longrightarrow \quad P(A \cup B) = P(A) + P(B).$$

Satz 6.18
Für jedes Ereignis A gilt

$$P(\bar{A}) = 1 - P(A)$$

und

$$P(A) \leq 1.$$

6-17

Satz 6.19
Für beliebige Ereignisse A und B gilt

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

Beispiel 6.20 (Würfelwurf)
Für

$$A = \{\omega_2, \omega_4, \omega_6\}, \quad B = \{\omega_1, \omega_2, \omega_3\}$$

gilt

$$A \cup B = \{\omega_1, \omega_2, \omega_3, \omega_4, \omega_6\}$$

und

$$P(A) = P(B) = \frac{1}{2}, \quad P(A \cap B) = P(\{\omega_2\}) = \frac{1}{6},$$

$$P(A \cup B) = \frac{1}{2} + \frac{1}{2} - \frac{1}{6} = \frac{5}{6}.$$

6-18

Definition 6.21 (Paarweise disjunkt)
Die Ereignisse $A_1, A_2, \ldots, A_n$ heißen **paarweise disjunkt**, falls

$$A_i \cap A_j = \emptyset \qquad \text{für } i, j = 1, \ldots, n \text{ mit } i \neq j.$$

Definition 6.22 (Vollständige Zerlegung, Partition)
Die paarweise disjunkten Ereignisse $A_1, A_2, \ldots, A_n$ mit $A_i \neq \emptyset$ für $i = 1, \ldots, n$ bilden eine **vollständige Zerlegung** oder **Partition** von Ω, falls

$$\bigcup_{i=1}^{n} A_i = \Omega$$

erfüllt ist.

6-19

Satz 6.23
Bilden die Ereignisse $A_1, A_2, \ldots, A_n$ eine vollständige Zerlegung von Ω, dann gilt

$$\sum_{i=1}^{n} P(A_i) = P(\Omega) = 1$$

und für jedes Ereignis B gilt

$$P(B) = \sum_{i=1}^{n} P(A_i \cap B).$$

6-20

6.3 Unabhängigkeit von Ereignissen

Definition 6.24 (Unabhängigkeit von zwei Ereignissen)
Zwei Ereignisse A und B heißen **(stochastisch) unabhängig** (*stochastically independent*) genau dann, wenn

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

6-21

Beispiel 6.25 (Fortsetzung Würfelbeispiel)

- Die Ereignisse $A_1 = \{\omega_1, \omega_3, \omega_5\}$ und $A_2 = \{\omega_2, \omega_4, \omega_6\}$ sind **nicht unabhängig**, da
$$P(A_1)\,P(A_2) = \frac{1}{2}\,\frac{1}{2} = \frac{1}{4}$$
und
$$P(A_1 \cap A_2) = P(\emptyset) = 0\,.$$
- Die Ereignisse $A_1 = \{\omega_1, \omega_3, \omega_5\}$ und $C = \{\omega_1, \omega_2, \omega_3, \omega_4\}$ sind **unabhängig**, da
$$P(C)\,P(A_1) = \frac{4}{6}\,\frac{1}{2} = \frac{1}{3}$$
und
$$P(C \cap A_1) = P(\{\omega_1, \omega_3\}) = \frac{1}{3}\,.$$

6-22

Definition 6.26
Die n Ereignisse $A_1, \ldots, A_n$ heißen **paarweise unabhängig** (*pairwise independent*), falls

$$P(A_i \cap A_j) = P(A_i) \cdot P(A_j) \qquad \text{für alle } i, j = 1, \ldots, n \text{ mit } i \neq j.$$

Bemerkung 6.27
Aus der paarweisen Unabhängigkeit von mehr als zwei Ereignissen folgt nicht die im Folgenden definierte vollständige Unabhängigkeit.

6-23

Definition 6.28
Die n Ereignisse $A_1, \ldots, A_n$ heißen **vollständig unabhängig** (*totally independent, mutually independent*) oder kurz **unabhängig** , falls für jede Auswahl von m Indizes $i_1, \ldots, i_m \in \{1, \ldots, n\}$, $2 \leq m \leq n$,

$$P(A_{i_1} \cap A_{i_2} \cap \ldots \cap A_{i_m}) = P(A_{i_1}) \cdot \ldots \cdot P(A_{i_m})$$

gilt.

Bemerkung 6.29
Drei Ereignisse A, B und C sind unabhängig, falls sie paarweise unabhängig sind, d.h. wenn

$$P(A \cap B) = P(A)P(B), P(A \cap C) = P(A)P(C), P(B \cap C) = P(B)P(C)$$

gilt, und falls **zusätzlich** zur paarweisen Unabhängigkeit

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

gilt.

6-24

Beispiel 6.30 (Wein und stochastische Unabhängigkeit)
Im Durchschnitt ist jede zehnte Weinflasche durch Korkgeschmack verdorben: $p = 0.10 = 10\%$. Es wird angenommen, dass die Ereignisse stochastisch unabhängig sind.
Wie groß ist die Wahrscheinlichkeit p_2, dass zwei hintereinander geöffnete Weinflaschen verdorben sind?

$$p_2 = p \cdot p = p^2 = 0.01$$

Wie groß ist die Wahrscheinlichkeit p_3, dass drei hintereinander geöffnete Weinflaschen verdorben sind?

$$p_3 = p \cdot p \cdot p = p^3 = 0.001$$

Wie groß ist die Wahrscheinlichkeit q_{10}, dass zehn hintereinander geöffnete Weinflaschen **nicht** verdorben sind?

$$q_{10} = (1-p)^{10} = 0.9^{10} = 0.3487 = 34.87\%$$

6.4 Ergänzungen

Bemerkung 6.a (Zur Selbstkontrolle) Im Kapitel 6 eingeführte Begriffe und Konzepte: Zufallsexperiment, Ergebnis, Ergebnismenge, Ereignis, Elementarereignis, Ereignissystem, unmögliches Ereignis, sicheres Ereignis, fairer Würfel, gleichwahrscheinlich, Laplace-Wahrscheinlichkeit, günstige und mögliche Fälle, Axiomensystem von Kolmogoroff, Nichtnegativität, Normierung, σ-Additivität, paarweise disjunkt, komplementäres Ereignis, Zerlegung, Unabhängigkeit von Ereignissen.

Beweis von Satz 6.16: Setze $A_1 = \Omega$ und $A_2 = A_3 = \ldots = \emptyset$. Mit Axiom 3 folgt

$$P(\Omega) = P\left(\bigcup_{i=1}^{n} A_i\right) = P(\Omega) + \sum_{i=2}^{\infty} P(\emptyset).$$

Wegen $P(\Omega) = 1$ (Axiom 2) kann die Gleichung

$$1 = 1 + \sum_{i=2}^{\infty} P(\emptyset)$$

nur mit $P(\emptyset) = 0$ erfüllt werden.

Beweis von Satz 6.17: Setze $A_1 = A$, $A_2 = B$ und $A_3 = A_4 = \ldots = \emptyset$. Dann gilt

$$P(A \cup B) = P\left(\bigcup_{i=1}^{\infty} A_i\right) = P(A) + P(B) + \sum_{i=3}^{\infty} P(A_i) = P(A) + P(B).$$

Beweis von Satz 6.18: Die erste Aussage folgt aus

$$P(A) + P(\bar{A}) = P(\Omega) = 1.$$

Die zweite Aussage folgt aus

$$P(A) = 1 - P(\bar{A})$$

und der Nichtnegativität von $P(\bar{A})$ (Axiom 1).

Beispiel 6.b (Zur paarweisen Unabhängigkeit) Es sei $\Omega = \{\omega_1, \omega_2, \omega_3, \omega_4\}$ mit $P(\{\omega_i\}) = \frac{1}{4}$ für $i = 1, \ldots, 4$. Für die Ereignisse $A = \{\omega_1, \omega_2\}$, $B = \{\omega_2, \omega_3\}$, $C = \{\omega_1, \omega_3\}$ verifiziert man leicht

$$P(A) = P(B) = P(C) = \frac{1}{2},$$

$$P(A \cap B) = P(B \cap C) = P(A \cap C) = \frac{1}{4}$$

$$P(A \cap B \cap C) = 0.$$

- Die drei Ereignisse A, B und C sind paarweise unabhängig.
- Die drei Ereignisse A, B und C sind nicht unabhängig.

Beispiel 6.c (Zur paarweisen Unabhängigkeit) Es sei $\Omega = \{\omega_1, \omega_2, \ldots, \omega_8\}$ mit $P(\{\omega_i\}) = \frac{1}{8}$ für $i = 1, \ldots, 8$. Für die Ereignisse $A = \{\omega_1, \omega_2, \omega_3, \omega_4\}$, $B = \{\omega_1, \omega_2, \omega_5, \omega_6\}$, $C = \{\omega_3, \omega_4, \omega_5, \omega_6\}$, $D = \{\omega_2, \omega_3, \omega_6, \omega_7\}$ verifiziert man leicht

$$P(A) = P(B) = P(C) = P(D) = \frac{1}{2},$$

$$P(A \cap B) = P(A \cap C) = P(A \cap D) = P(B \cap C) = P(B \cap D) = \frac{1}{4}$$

$$P(A \cap B \cap C) = 0 \quad \text{und} \quad P(A \cap B \cap D) = \frac{1}{8}.$$

- Die drei Ereignisse A, B und C sind paarweise unabhängig, aber nicht unabhängig.
- Die drei Ereignisse A, B und D sind unabhängig.
- Die vier Ereignisse A, B, C und D sind nicht unabhängig.

Definition 6.d (Ereignis-σ-Algebra) Enthält ein Ereignissystem $\mathcal{A} \subset \mathcal{P}(\Omega)$

1. das Ereignis Ω,
2. zu jedem Ereignis $A \in \mathcal{A}$ auch das komplementäre Ereignis $\bar{A}$, und
3. zu je abzählbar unendlich vielen Ereignissen $A_1, A_2, \ldots$ auch deren Vereinigung und Durchschnitt,

so nennt man $\mathcal{A}$ eine **Ereignis-σ-Algebra** auf Ω.

Kapitel 7

Zufallsvariable und Verteilung

7-1

7. Zufallsvariable und Verteilung

7.1 Zufallsvariable
7.2 Verteilungsfunktion

7-2

7.1 Zufallsvariable

Bemerkung 7.1
Gegeben sei ein Zufallsexperiment mit der Ergebnismenge Ω. In der Regel will man die Ergebnisse durch reelle Zahlen repräsentieren bzw. die $\omega \in \Omega$ sind Merkmalsträger und man interessiert sich für die Merkmalswerte $X(\omega)$. Formal interessieren reellwertige Abbildungen (Funktionen)

$$X : \Omega \to \mathbb{R}, \quad \omega \mapsto X(\omega),$$

für welche die Wahrscheinlichkeiten

$$P(X \leq x) \stackrel{\text{def}}{=} P(\{\omega \in \Omega \mid X(\omega) \leq x\})$$

für alle $x \in \mathbb{R}$ bestimmbar sind.

7-3

Beispiel 7.2 (Würfel)

$$\begin{array}{ccccccccc} & \Omega & = & \{\omega_1, & \omega_2, & \omega_3, & \omega_4, & \omega_5, & \omega_6\} \\ X: & & & \downarrow & \downarrow & & & & \downarrow \\ & & & 1 & 2 & & & & 6 \end{array}$$

$$P(X=2) = P(\{\omega \in \Omega \mid X(\omega) = 2\}) = P(\{\omega_2\}) = \frac{1}{6}$$

$$P(X \le 2) = P(\{\omega \in \Omega \mid X(\omega) \le 2\}) = P(\{\omega_1, \omega_2\}) = \frac{2}{6}$$

Beispiel 7.3 (Körpergröße)

$$\begin{array}{ccccccccc} & \Omega & = & \{\omega_1, & \omega_2, & \omega_3, & \cdots, & \omega_9, & \omega_{10}\} \\ X: & & & \downarrow & \downarrow & \downarrow & & \downarrow & \downarrow \\ & & & 175 & 190 & 175 & \neq 175 & 175 & 180 \end{array}$$

$$P(X=175) = P(\{\omega \in \Omega \mid X(\omega) = 175\}) = P(\{\omega_1, \omega_3, \omega_9\}) = \frac{3}{10}$$

7-4

Definition 7.4 (Zufallsvariable)
Gegeben sei eine Ergebnismenge Ω mit einem Ereignissystem $\mathcal{A}$ und einer für die Ereignisse in $\mathcal{A}$ definierten Wahrscheinlichkeit P.
Eine Funktion $X : \Omega \to \mathbb{R}$ heißt **Zufallsvariable** (*random variable*), falls sich alle Wahrscheinlichkeiten der Form

$$P(X \le x) = P(\{\omega \in \Omega \mid X(\omega) \le x\}), \quad x \in \mathbb{R},$$

angeben lassen.

Bemerkung 7.5
Dazu muss

$$\{\omega \in \Omega \mid X(\omega) \le x\} \in \mathcal{A} \quad \text{für alle} \quad x \in \mathbb{R}$$

gelten.

7-5

7.2 Verteilungsfunktion

Definition 7.6 (Verteilungsfunktion von X)
Für jede Zufallsvariable X ist durch die Funktion $F_X : \mathbb{R} \to [0,1]$,

$$F_X(x) = P(X \le x) \quad \text{für alle} \quad x \in \mathbb{R},$$

die **Verteilungsfunktion** (*distribution function, d. f.*; *cumulative distribution function, c. d. f.*) von X definiert.

Bemerkung 7.7
Wenn nur eine Zufallsvariable X in Frage kommt, wird auch kurz F anstelle F_X geschrieben.

7-6

Bemerkung 7.8 (Eigenschaften)
Jede Verteilungsfunktion F einer Zufallsvariablen X hat die Eigenschaften

- $\lim_{x \to -\infty} F(x) = 0$
- $\lim_{x \to \infty} F(x) = 1$
- F ist monoton wachsend (monoton nicht abnehmend, *monotone non-decreasing*),

$$a < b \Rightarrow F(a) \leq F(b) \quad \text{für alle} \quad a, b \in \mathbb{R}$$

- F ist rechtsseitig stetig (*right-continuous, continuous on the right*),

$$\lim_{\substack{\varepsilon \to 0 \\ \varepsilon > 0}} F(x + \varepsilon) = F(x) \quad \text{für alle} \quad x \in \mathbb{R}\,.$$

7-7

Beispiel 7.9 (Treppenfunktion)
X sei die zufällige Augenzahl beim fairen Würfelwurf. Die Verteilungsfunktion $F(x) = P(X \leq x)$ von X ist wegen

$$P(X \leq x) = \sum_{i \leq x} P(X = i) \quad \text{und} \quad P(X = i) = \frac{1}{6} \quad \text{für } i = 1, \ldots, 6$$

eine **Treppenfunktion** (*step function*) mit Sprüngen (*jumps*) der Höhe 1/6 an den Stellen $1, 2, ..., 6$.

$$F(x) = \begin{cases} 0 & & x < 1 \\ 1/6 & & 1 \leq x < 2 \\ 2/6 & & 2 \leq x < 3 \\ 3/6 & \text{falls} & 3 \leq x < 4 \\ 4/6 & & 4 \leq x < 5 \\ 5/6 & & 5 \leq x < 6 \\ 1 & & 6 \leq x \end{cases}$$

7-8

Bemerkung 7.10
Wenn die Verteilungsfunktion einer Zufallsvariablen X stetig ist, dann gilt

$$P(X = x) = 0 \quad \text{für alle} \quad x \in \mathbb{R}\,.$$

Beispiel 7.11
Gegeben sei die stetige Verteilungsfunktion

$$F_X(x) = \begin{cases} 0 & & x < 0 \\ x & \text{falls} & 0 \leq x < 1 \\ 1 & & 1 \leq x \end{cases}$$

einer Zufallsvariablen X. Dann gilt

$$P(0 \leq X \leq 1) = 1$$

und

$$P(a \leq X \leq b) = b - a \quad \text{für alle} \quad 0 \leq a \leq b \leq 1\,.$$

7-9

Bemerkung 7.12
Eine gegebene Verteilungsfunktion legt unmittelbar die Wahrscheinlichkeiten

$$P(X \leq x) = P(X \in]-\infty, x]) = F_X(x)$$

für alle $x \in \mathbb{R}$ fest. Aus diesen erhält man die Wahrscheinlichkeiten komplizierterer Ereignisse. Für $a < b$ erhält man z. B.

$$P(a < X \leq b) = F_X(b) - F_X(a) .$$

Es gilt

$$]-\infty, b] =]-\infty, a] \cup]a, b] \quad \text{und} \quad]-\infty, a] \cap]a, b] = \emptyset .$$

Aus der Additivität für disjunkte Ereignisse folgt

$$P(X \leq b) = P(X \leq a) + P(a < X \leq b)$$

bzw.

$$F_X(b) = F_X(a) + P(a < X \leq b) .$$

7-10

Definition 7.13 (p-Quantil)

- Zu vorgegebenem $0 < p < 1$ heißt die Zahl

$$x_p = \min\{x \mid F_X(x) \geq p\}$$

  p-Quantil (*p-quantile*) oder p-Fraktil von X.
- Das Quantil $x_{0.5}$ heißt **Median** (*median*).
- $x_{0.25}$, $x_{0.5}$, $x_{0.75}$ sind die **Quartile** (*quartiles*).

7.3 Ergänzungen

Bemerkung 7.a (Zur Selbstkontrolle) Im Kapitel 7 eingeführte Begriffe und Konzepte: Zufallsvariable, Verteilungsfunktion, Wahrscheinlichkeitsverteilung, p-Quantil einer Zufallsvariablen, Median, Quartile.

Bemerkung 7.b Im Rahmen der deskriptiven Statistik (Kapitel 1) wurde das Symbol $\tilde{x}_p$ für das p-Quantil der Daten eingeführt, das aus der empirischen Verteilungsfunktion der Daten berechnet wird. Hier wird mit x_p das p-Quantil einer Zufallsvariablen X bezeichnet. Dieses ist eine Maßzahl der Wahrscheinlichkeitsverteilung von X und wird aus der Verteilungsfunktion F_X abgeleitet. Diese beiden Konzepte finden im Rahmen der schließenden Statistik (Kapitel 12 bis 20) zueinander. Wenn nämlich $\tilde{x}_p$ aus einer Stichprobe vom Umfang n berechnet wurde, dann ist $\tilde{x}_p$ ein Schätzwert für das p-Quantil x_p der Grundgesamtheit.

Bemerkung 7.c Die Quantile $x_{0.2}$, $x_{0.4}$, $x_{0.6}$, $x_{0.8}$ heißen **Quintile** (*quintiles*), die Quantile $x_{0.1}$, $x_{0.2}$, $x_{0.3}$, $x_{0.4}$, $x_{0.5}$, $x_{0.6}$, $x_{0.7}$,$x_{0.8}$, $x_{0.9}$ heißen **Dezile** und die Quantile $x_{0.01}$, $x_{0.02}$, $\ldots$, $x_{0.99}$ heißen **Centile** oder **Perzentile**.

Kapitel 8

Diskrete Verteilungen

8-1

8. Diskrete Verteilungen

8.1 Wahrscheinlichkeitsfunktion
8.2 Bernoulli-Verteilung
8.3 Binomialverteilung
8.4 Poisson-Verteilung

8-2

8.1 Wahrscheinlichkeitsfunktion

Definition 8.1 (Diskrete Zufallsvariable, Träger)

- Eine Zufallsvariable X und ihre Verteilung heißen **diskret** (*discrete*), falls es eine endliche oder abzählbar unendliche Menge $T_X \subset \mathbb{R}$ mit

$$P(X = x) > 0 \quad \text{für alle} \quad x \in T_X$$

und

$$\sum_{x \in T_X} P(X = x) = 1$$

gibt.

- Die Menge T_X heißt dann **Träger** (*support*) der Wahrscheinlichkeitsverteilung von X.

8-3

Beispiel 8.2
X bezeichne die Augenzahl beim Würfelwurf. X ist eine diskrete Zufallsvariable mit dem Träger $T_X = \{1, 2, \ldots, 6\}$.

Definition 8.3 (Wahrscheinlichkeitsfunktion)
Für eine diskrete Zufallsvariable X heißt die Funktion

$$f_X : \mathbb{R} \to [0,1], \quad x \mapsto f_X(x) = P(X = x)$$

Wahrscheinlichkeitsfunktion (*probability function, probability mass function*) von X.

Beispiel 8.4
X sei die Augenzahl beim Würfelwurf. X hat die Wahrscheinlichkeitsfunktion

$$f_X(x) = \begin{cases} \frac{1}{6} & \text{falls} \quad x \in \{1,2,3,4,5,6\} = T_X \\ 0 & \phantom{\text{falls}} \quad x \in \mathbb{R} \setminus T_X \,. \end{cases}$$

8-4

8.2 Bernoulli-Verteilung

Definition 8.5 (Bernoulli-Verteilung)
Es sei $0 < \pi < 1$.

- Eine Zufallsvariable X mit dem Träger $T_X = \{0, 1\}$ und der Wahrscheinlichkeitsfunktion

  $$f_X(x) = P(X = x) = \begin{cases} \pi & \text{für} \quad x = 1 \\ 1 - \pi & \phantom{\text{für}} \quad x = 0 \end{cases}$$

  heißt **Bernoulli-verteilt** mit dem **Bernoulli-Parameter** π.
- Schreibweise: $X \sim B(1, \pi)$.

8-5

Bemerkung 8.6 (Bernoulli-Experiment)

- Ein **Bernoulli-Experiment** ist ein Zufallsexperiment mit zwei möglichen Ergebnissen.
- Ein Ergebnis entspricht $X = 1$ und wird als „Erfolg“ bezeichnet, das andere Ergebnis entspricht $X = 0$ und wird als „Misserfolg“ bezeichnet.
- $''X = 1''$ zeigt das Auftreten eines interessierenden Ereignisses an.
- Der Bernoulli-Parameter π heißt daher auch **Erfolgsparameter**.

8-6

Bemerkung 8.7
Für irgendeine Zufallsvariable Y und $A \subset \mathbb{R}$ gelte

$$\pi \stackrel{\text{def}}{=} P(Y \in A) \in \,]0,1[\ ,$$

d. h. das Ereignis $Y \in A$ tritt mit der Wahrscheinlichkeit π ein. Für die durch

$$X = \begin{cases} 1 & \text{falls} \quad Y \in A \\ 0 & \phantom{\text{falls}} \quad Y \notin A \end{cases}$$

gebildete Zufallsvariable X gilt dann

$$X \sim B(1, \pi).$$

In diesem Zusammenhang bezeichnet man X auch als **Indikatorvariable**.

8-7

8.3 Binomialverteilung

Definition 8.8 (Fakultät und Binomialkoeffizient)

- Fakultät

$$0! \stackrel{\text{def}}{=} 1$$

$$n! \stackrel{\text{def}}{=} 1 \cdot 2 \cdots (n-1) \cdot n = \prod_{i=1}^{n} i \qquad \text{für} \quad n \in \mathbb{N}$$

- Binomialkoeffizient

$$\binom{n}{k} \stackrel{\text{def}}{=} \frac{n!}{k!(n-k)!} \qquad \text{für} \quad n, k \in \mathbb{N} \cup \{0\}, \quad k \leq n$$

8-8

Beispiel 8.9 (Binomialverteilung)
$n = 10$ Schüsse sind (unabhängig) auf ein Ziel gerichtet. Jeder Schuss trifft mit einer Wahrscheinlichkeit von $\pi = 0.9$.

- Die zufällige Trefferzahl X ist dann **binomialverteilt** mit den Parametern $n = 10$ und $\pi = 0.9$:

$$X \sim B(n, \pi).$$

- Die Wahrscheinlichkeit von z. B. vier Treffern ist

$$P(X = 4) = \binom{10}{4} 0.9^4\, 0.1^6.$$

8-9

Definition 8.10 (Binomialverteilung)
Es sei $0 < \pi < 1$ und $n \in \mathbb{N}$.

- Eine Zufallsvariable X mit dem Träger $T_X = \{0, 1, \ldots, n\}$ und der Wahrscheinlichkeitsfunktion

$$f_X(x) = P(X = x) = \binom{n}{x} \pi^x (1-\pi)^{n-x} \qquad \text{für} \quad x \in T_X$$

  heißt **binomialverteilt** mit den Parametern n und π.

- Schreibweise: $X \sim B(n, \pi)$.

Bemerkung 8.11 (Eigenschaften)

- Es gilt

$$\sum_{x=0}^{n} f_X(x) = 1\,.$$

8-10

Bemerkung 8.12 (Zur Interpretation)
X ist die zufällige Anzahl der Erfolge bei n unabhängigen Versuchen (unabhängigen Bernoulli-Experimenten) mit der Erfolgswahrscheinlichkeit π.

Beispiel 8.13 (Wein und Binomialverteilung)
Im Durchschnitt ist jede zehnte Weinflasche durch Korkgeschmack verdorben. Die Ereignisse sind als stochastisch unabhängig vorausgesetzt.
Wie groß sind dann die Wahrscheinlichkeiten, dass von zehn gekauften Weinflaschen keine und genau eine verdorben sind?
X bezeichne die Anzahl der verdorbenen Weinflaschen. Dann ist X binomialverteilt mit den Parametern $n = 10$ und $\pi = 0.10$ und es gilt

$$P(X = 0) = \binom{10}{0} \cdot 0.1^0 \cdot 0.9^{10} = 0.9^{10} = 0.3487,$$

$$P(X = 1) = \binom{10}{1} \cdot 0.1 \cdot 0.9^9 = 0.3874.$$

8-11

Beispiel 8.14 (Hochwasser)

- $\pi = 0.05$ bezeichnet die Wahrscheinlichkeit eines Hochwasser-Jahres.
- Annahme: Unabhängigkeit der Hochwasserereignisse verschiedener Jahre.
- X bezeichne die Anzahl der Hochwasser-Jahre in n Jahren, dann gilt

$$X \sim B(n,\, 0.05)\,.$$

- Für $n = 20$ erhält man

$$P(X = x) = \binom{20}{x} 0.05^x 0.95^{20-x} \qquad \text{für} \quad x \in \{0, 1, \ldots, 20\}$$

  und speziell

$$P(X = 0) = 0.358, \quad P(X \le 1) = 0.736,$$

$$P(X \le 2) = 0.925, \quad P(X \le 3) = 0.984\,.$$

8-12

8.4 Poisson-Verteilung

Beispiel 8.15 (Kreditkartengesellschaft)
Betrachtet wird die Notrufzentrale einer Kreditkartengesellschaft.

- X bezeichne die zufällige Zahl von Meldungen pro Stunde.
- Durchschnittlich gehen drei Meldungen pro Stunde ein.
- X folgt einer Poisson-Verteilung mit dem Parameter $\mu = 3$.
- Dann gilt z. B.
$$P(X = 2) = \frac{3^2}{2!} e^{-3} \approx 0.22\,.$$

8-13

Definition 8.16 (Poisson-Verteilung)
Es sei $\mu > 0$.

- Eine Zufallsvariable X mit dem Träger
$$T_X \stackrel{\text{def}}{=} \{0, 1, 2, \ldots\}$$
und der Wahrscheinlichkeitsfunktion
$$f_X(x) = P(X = x) = \frac{\mu^x}{x!} e^{-\mu} \qquad \text{für} \quad x \in T_X$$
heißt **Poisson-verteilt** mit dem Parameter μ.
- Schreibweise: $X \sim Poi(\mu)$.

8-14

Bemerkung 8.17 (Anwendung)

- Das Zählen von Ereignissen, die zufällig und voneinander unabhängig innerhalb eines Zeitintervalls auftreten, führt zur Poisson-Verteilung.
- Beispiele:
  - Telefonanrufe pro Stunde in einer Vermittlungsstelle
  - Emittierte α-Teilchen pro Minute beim radioaktiven Zerfall
- μ heißt in diesem Zusammenhang auch **Intensitätsrate**. Diese gibt die durchschnittliche Ereignishäufigkeit in einem Intervall der Länge 1 an.

8-15

Bemerkung 8.18 (Approximation)

- $X \sim B(n, \pi)$ mit n sehr groß und π sehr klein, dann ist X näherungsweise Poisson-verteilt mit $\mu = n\pi$.
- Als Faustregel für die Zulässigkeit der Approximation gilt, dass die drei Bedingungen
$$\pi \leq 0.1, \quad n \geq 50, \quad n\pi \leq 9$$
erfüllt sind.
- Wahrscheinlichkeiten einer Poisson-Verteilung lassen sich im allgemeinen leichter berechnen als Wahrscheinlichkeiten einer Binomialverteilung.
- Beispiel: Anzahl der falsch gedruckten Buchstaben pro Seite.

8.5 Ergänzungen

Bemerkung 8.a (Zur Selbstkontrolle) Im Kapitel 8 eingeführte Begriffe und Konzepte: Diskrete Zufallsvariable, Träger, Wahrscheinlichkeitsfunktion, Bernoulli-Verteilung, -parameter, -experiment, Erfolgsparameter, Binomialverteilung, Poisson-Verteilung.

Beispiel 8.b (Lotto 1) Beim Lotto 6 aus 49 gibt es

$$N \stackrel{\text{def}}{=} \binom{49}{6} = 13\,983\,816$$

verschiedene Tippreihen. Die Wahrscheinlichkeit, mit einer abgegebenen Tippreihe einen „Sechser“ (sechs richtige Zahlen) zu erreichen, ist daher

$$\frac{1}{N} = \frac{1}{13\,983\,816} = 0.00000715\%\,.$$

Es genügt, alle N verschiedenen Tippreihen abzugeben, um mit Sicherheit einen „Sechser“ (sechs Richtige) zu erhalten. Neben einem „Sechser“ ist die Anzahl der zusätzlich erzielten „Fünfer“

$$\binom{6}{5}\binom{43}{1} = 6 \cdot 43 = 258\,,$$

der „Vierer“

$$\binom{6}{4}\binom{43}{2} = 15 \cdot 903 = 13\,545$$

und der „Dreier“

$$\binom{6}{3}\binom{43}{3} = 20 \cdot 12341 = 246\,820\,.$$

Um den Jackpot zu knacken, muss der „Sechser“ mit der richtigen Superzahl ($\in \{0, 1, \ldots, 9\}$) kombiniert sein. Mit $10 \cdot N$ verschiedenen Tippreihen kann man also den Jackpot mit Sicherheit knacken.

Beispiel 8.c (Lotto 2)
Beim Lotto 6 aus 49 gibt es $N \stackrel{\text{def}}{=} 13\,983\,816$ Tippreihen und 10 Superzahlen. Wie groß ist die Wahrscheinlichkeit, dass der Jackpot geknackt wird („Sechser“ und richtige Superzahl), wenn n Tippreihen zufällig und voneinander unabhängig ausgewählt und jeweils zufällig mit einer Superzahl kombiniert werden?
Die Trefferzahl ist binomialverteilt mit den Parametern n und der Erfolgswahrscheinlichkeit

$$\pi \stackrel{\text{def}}{=} \frac{1}{10 \cdot N} = 0.715 \cdot 10^{-8}.$$

Die Poisson-Approximation führt zur Zufallsvariablen X mit dem Parameter $\mu = n\pi$ und den Wahrscheinlichkeiten

$$p_k \stackrel{\text{def}}{=} P(X = k) = \frac{\mu^k}{k!} e^{-\mu}, \quad k = 0, 1, 2, \ldots .$$

Beispielsweise erhält man für $n = 120\,000\,000$ den Parameter $\mu = 0.8581$ und die Wahrscheinlichkeiten

$$p_0 = e^{-\mu} = 42\%, \quad p_1 = \mu e^{-\mu} = 36\%.$$

Werden also $n = 120\,000\,000$ (120 Millionen) zufällige Tippreihen abgegeben, so ist $p_0 = 42\%$ die Wahrscheinlichkeit, dass niemand den Jackpot knackt, und $p_1 = 36\%$ ist die Wahrscheinlichkeit, dass genau eine Tippreihe den Jackpot knackt. Außerdem gilt $p_2 = 16\%$, $p_3 = 4\%$, $p_4 = 1\%$, $p_5 = 0.2\%$.

Beispiel 8.d (Lotto 3) Beim Lotto 6 aus 49 gibt es $N \stackrel{\text{def}}{=} 13\,983\,816$ verschiedene Tippreihen.
Wie groß ist die Wahrscheinlichkeit, dass eine spezielle Tippreihe, z. B. die Reihe {1,2,3,4,5,6}, auftaucht, wenn N Tippreihen zufällig und voneinander unabhängig ausgewählt werden?
Die Trefferwahrscheinlichkeit bei einer zufälligen Auswahl einer Tippreihe ist $\pi \stackrel{\text{def}}{=} \frac{1}{N} = 0.715 \cdot 10^{-7}$. Die Wahrscheinlichkeit für genau einen Treffer bei N Versuchen ist durch die Binomialwahrscheinlichkeit

$$\binom{N}{1} \pi (1-\pi)^{N-1} = (1-\pi)^{N-1} = 0.36787945$$

gegeben. Die Poisson-Approximation führt zu einer Poisson-verteilten Zufallsvariable X mit dem Poisson-Parameter $\lambda = 1$ und der Wahrscheinlichkeit für einen Erfolg von

$$P(X = 1) = \frac{1^1}{1!} e^{-1} = e^{-1} = 0.36787944.$$

Beispiel 8.e (Lotto 4) Beim Lotto 6 aus 49 gibt es $N \stackrel{\text{def}}{=} 13\,983\,816$ verschiedene Tippreihen.
Wie groß ist die Wahrscheinlichkeit, dass jede Tippreihe genau einmal gewählt wird, wenn N Tippreihen zufällig ausgewählt werden?
Die Wahrscheinlichkeit, bei der zweiten Wahl nicht die zuerst gewählte Tippreihe zu wählen, ist $\frac{N-1}{N}$. Die Wahrscheinlichkeit, bei der dritten Wahl nicht die beiden Reihen der ersten und zweiten Wahl zu wählen, ist $\frac{N-2}{N}$. Die Wahrscheinlichkeit, bei der letzten der N Ziehungen keine der bereits gezogenen $N-1$ verschiedenen Reihen zu wählen, ist $\frac{1}{N}$. Die gesuchte Wahrscheinlichkeit ist daher

$$p_N \stackrel{\text{def}}{=} \prod_{i=1}^{N-1} \frac{N-i}{N} = \frac{(N-1)!}{N^{N-1}} .$$

Zur Berechnung von p_N verwendet man vorteilhaft

$$\ln(p_N) = \ln\left(\frac{(N-1)!}{N^{N-1}}\right) = \ln((N-1)!) - (N-1)\ln(N) .$$

Die Funktion $\ln(n!)$ für $n \in \mathbb{N}$ steht entweder in Software zur Verfügung oder $n!$ kann mit Hilfe der Gammafunktion berechnet werden. Die Wahrscheinlichkeit p_N kann z. B. mit der Software GAUSS durch

```
exp(lnfact(N-1) - (N-1)*ln(N))
```

im Prinzip berechnet werden. Man erhält z. B. $p_{100} = 0.93 \cdot 10^{-42}$ und $p_{700} = 0.65 \cdot 10^{-303}$.

Bereits p_{1000} führt aber zu einer so winzigen Zahl, dass diese im Rahmen der Rechengenauigkeit des Programms nicht mehr von Null unterschieden werden kann.

Beispiel 8.f Beim wiederholten Werfen einer Münze bezeichne Y die Anzahl der Würfe bis erstmalig „Zahl" oben liegt. Bei stochastischer Unabhängigkeit der Ergebnisse ist Y eine diskrete Zufallsvariable mit dem Träger $\mathbb{N}$ und der Wahrscheinlichkeitsfunktion

$$f_Y(x) = P(Y = x) = \begin{cases} \frac{1}{2^x} & \text{falls } x \in \mathbb{N} \\ 0 & \text{falls } x \notin \mathbb{N}. \end{cases}$$

Dabei gilt

$$\sum_{x \in \mathbb{N}} f_Y(x) = \sum_{i=1}^{\infty} \frac{1}{2^i} = \frac{1}{2} + \frac{1}{4} + \frac{1}{8} + \frac{1}{16} + \ldots = 1$$

Kapitel 9

Stetige Verteilungen

9-1

9. Stetige Verteilungen

9.1 Dichtefunktion
9.2 Rechteckverteilung
9.3 Exponentialverteilung
9.4 Normalverteilung

9-2

9.1 Dichtefunktion

Definition 9.1 (Stetige Zufallsvariable, Dichtefunktion)

- Eine Zufallsvariable X und ihre Verteilung heißen **stetig** (*continuous*), falls es eine Funktion $f_X : \mathbb{R} \to [0, \infty[$ gibt, so dass

$$F_X(x) = P(X \leq x) = \int_{-\infty}^{x} f_X(t)dt \quad \text{für alle} \quad x \in \mathbb{R}$$

gilt.

- Die Funktion $f_X(x)$ heißt dann **Dichtefunktion** (*probability density function, p. d. f.*) oder auch kurz **Dichte** von X.

Bemerkung 9.2
Wenn F_X an einer Stelle x differenzierbar ist, dann gilt $F'_X(x) = f_X(x)$.

9-3

Bemerkung 9.3
Für eine stetige Zufallsvariable X mit der Dichtefunktion f_X gilt

$$f_X(x) \geq 0 \quad \text{für alle} \quad x \in \mathbb{R},$$

$$\int\limits_{-\infty}^{\infty} f_X(x)dx = 1,$$

$$P(a \leq X \leq b) = \int\limits_{a}^{b} f_X(x)dx$$

und

$$P(X = x) = 0 \quad \text{für alle} \quad x \in \mathbb{R}.$$

9-4

Beispiel 9.4
X habe die Dichtefunktion

$$f_X(x) = \begin{cases} 6(x - x^2) & \text{für } x \in [0,1] \\ 0 & \text{für } x \notin [0,1]. \end{cases}$$

Dann gilt

$$P\left(\frac{1}{6} \leq X \leq \frac{2}{3}\right) = \int\limits_{1/6}^{2/3} 6(x - x^2)dx = 6\left[\frac{x^2}{2} - \frac{x^3}{3}\right]_{1/6}^{2/3} = \frac{2}{3}\,.$$

9-5

9.2 Rechteckverteilung

Definition 9.5 (Rechteckverteilung)
Es sei $\alpha < \beta$.

- Eine Zufallsvariable X mit der Dichtefunktion

  $$f_X(x) \overset{\text{def}}{=} \begin{cases} \frac{1}{\beta - \alpha} & \text{für } x \in [\alpha, \beta] \\ 0 & \text{für } x \notin [\alpha, \beta] \end{cases}$$

  heißt **rechteckverteilt** im Intervall $[\alpha, \beta]$.

- Schreibweise: $X \sim R(\alpha, \beta)$.

9-6

Bemerkung 9.6

- Der Name Rechteckverteilung bezieht sich auf die Rechteckform der Dichtefunktion.
- Die Rechteckverteilung heißt auch **Gleichverteilung** oder **uniforme Verteilung** (*uniform distribution*) wegen der gleichförmigen Verteilung der Wahrscheinlichkeit.

9-7

Bemerkung 9.7
Für $X \sim R(\alpha, \beta)$ und $\alpha \leq c \leq d \leq \beta$ gilt

$$P(X \in [c, d]) = \frac{d - c}{\beta - \alpha}.$$

Bemerkung 9.8

- Die Verteilungsfunktion von $X \sim R(0, 1)$ ist

$$F_X(x) = \begin{cases} 0 & & x < 0 \\ x & \text{falls} & 0 \leq x < 1 \\ 1 & & 1 \leq x. \end{cases}$$

- Die Verteilungsfunktion von $X \sim R(\alpha, \beta)$ ist

$$F_X(x) = \begin{cases} 0 & & x < \alpha \\ \frac{x-\alpha}{\beta-\alpha} & \text{falls} & \alpha \leq x < \beta \\ 1 & & \beta \leq x. \end{cases}$$

9-8

9.3 Exponentialverteilung

Beispiel 9.9 (Exponentialverteilung)

- Betrachtet wird ein außergewöhnlicher Kurssturz (*Crash*), definiert durch einen DAX-Tagesverlust von über 8%. Ein solcher Kurssturz erfolgte zuletzt am 11. 9. 2001 und zum vorletzten mal am 28. 10. 1997.
- Modellierung der Zeit zwischen zwei Crashs
- Intensitätsrate: Crash/Jahr mit $\lambda = \frac{1}{10}$
  Mittlere Zwischenereigniszeit: $\frac{1}{\lambda} = 10$
- Annahme: konstante Intensitätsrate; Prozess ohne Gedächtnis
- Die zufällige Zwischenereigniszeit X folgt einer Exponentialverteilung mit dem Parameter $\lambda = 1/10$ und es gilt z. B.

$$P(X \leq 5) = \int_{-\infty}^{5} f_X(x)dx = 1 - e^{-\frac{5}{10}} = 39.3\%.$$

9-9

Definition 9.10 (Exponentialverteilung)
Es sei $\lambda > 0$.

- Eine Zufallsvariable X mit der Dichtefunktion

$$f_X(x) = \begin{cases} \lambda e^{-\lambda x} & \text{für } x \geq 0 \\ 0 & \text{für } x < 0 \end{cases}$$

  heißt **exponentialverteilt** mit dem Parameter λ.

- Schreibweise: $X \sim Exp(\lambda)$.

9-10

Bemerkung 9.11 (Eigenschaften)

- Die Verteilung der Zwischenereigniszeiten (Wartezeiten, Abfertigungszeiten,...) bei Poisson-verteilten Ereignishäufigkeiten pro Zeitintervall ist eine Exponentialverteilung.

- Dabei ist $1/\lambda$ die mittlere Zwischenereigniszeit und λ ist die mittlere Ereignishäufigkeit pro Zeitintervall.

- Die Verteilungsfunktion von $X \sim Exp(\lambda)$ ist

$$F_X(x) = \begin{cases} 1 - e^{-\lambda x} & \text{für } x \geq 0 \\ 0 & \text{für } x < 0. \end{cases}$$

- Für $X \sim Exp(\lambda)$ gilt $Y = \lambda X \sim Exp(1)$ mit der Dichtefunktion

$$f_Y(y) = e^{-y}, \quad y \geq 0.$$

9-11

Beispiel 9.12 (Exponentialverteilung)

- Betrachtet wird eine Notrufzentrale (siehe oben) mit durchschnittlich drei Meldungen pro Stunde. Die zufällige Anzahl der Meldungen pro Stunde sei Poisson-verteilt.

- X bezeichne jetzt die zufällige Zeit zwischen zwei Meldungen.

- X ist exponentialverteilt mit dem Parameter $\lambda = 3$; $X \sim Exp(\lambda)$.

- Die Wahrscheinlichkeit, dass zwischen zwei Meldungen mehr als eine Stunde vergeht, ist

$$\begin{aligned} P(X > 1) &= 1 - P(X \leq 1) \\ &= 1 - \int_0^1 3e^{-3x} dx \\ &= e^{-3} \\ &= 0.05\ . \end{aligned}$$

9-12

9.4 Normalverteilung

Definition 9.13 (Normalverteilung oder Gauß-Verteilung)
Es sei $\mu \in \mathbb{R}$ und $\sigma > 0$.

- Eine Zufallsvariable X mit der Dichtefunktion

$$f_X(x) \stackrel{\text{def}}{=} \frac{1}{\sqrt{2\pi}\sigma} e^{-\frac{(x-\mu)^2}{2\sigma^2}}$$

  heißt **normalverteilt** mit den Parametern μ und σ^2.

- Schreibweise: $X \sim N(\mu, \sigma^2)$.

9-13

Bemerkung 9.14 (Eigenschaften)
Die Dichtefunktion $f_X(x)$ von $X \sim N(\mu, \sigma^2)$ hat die folgenden Eigenschaften:

- $f_X(x) > 0 \quad$ für alle $\quad x \in \mathbb{R}$.
- $f_X(x)$ ist glockenförmig mit einem Maximum an der Stelle $x = \mu$.
- $f_X(x)$ ist symmetrisch zu μ.
- $f_X(x)$ hat Wendepunkte bei $\mu - \sigma$ und $\mu + \sigma$.
- Es gilt

$$\lim_{x\to\infty} f_X(x) = \lim_{x\to-\infty} f_X(x) = 0.$$

9-14

Definition 9.15 (Standard-Normalverteilung)
Die spezielle Normalverteilung $N(\mu, \sigma^2)$ mit den Parametern $\mu = 0$ und $\sigma^2 = 1$ heißt **Standard-Normalverteilung**.

Bemerkung 9.16 (Dichte- und Verteilungsfunktion)

- Eine standard-normalverteilte Zufallsvariable $U \sim N(0,1)$ hat die **Dichtefunktion**

$$f_U(u) = \varphi(u) \stackrel{\text{def}}{=} \frac{1}{\sqrt{2\pi}} e^{-\frac{u^2}{2}}$$

  und die **Verteilungsfunktion**

$$F_U(u) = \Phi(u) \stackrel{\text{def}}{=} \int_{-\infty}^{u} \varphi(t)dt.$$

- Die Werte von $\Phi(u)$ sind vertafelt (in Tabellen angegeben).
- Es gilt

$$\Phi(-u) = 1 - \Phi(u).$$

9-15

Bemerkung 9.17
Merkmalsbeobachtungen, die vielen unabhängigen Zufallseinflüssen ausgesetzt sind, sind häufig näherungsweise normalverteilt.

Beispiel 9.18

Messwert	=	wahrer Wert	+	Zufallswert (Messfehler)
X	=	μ	+	Z

Der Messfehler Z hat den Mittelwert 0 und den Streuungsparameter σ^2.

$$X \sim N(\mu, \sigma^2) \quad \text{bzw.} \quad Z \sim N(0, \sigma^2)$$

9-16

Bemerkung 9.19 (Ein-, Zwei-, ..., Sechs-Sigma-Bereiche)

- Für $k > 0$ heißt das Intervall $[\mu - k\sigma, \mu + k\sigma]$ k-**Sigma-Bereich.**
- Es gilt
$$P(X \in [\mu - k\sigma,\, \mu + k\sigma]) = P(\mu - k\sigma \leq X \leq \mu + k\sigma)\,.$$
- In der Regel ist k eine positive ganze Zahl. Speziell gilt für $k = 1, 2, \ldots, 6$ und beliebige μ und σ^2:
$$P(\mu - \sigma \leq X \leq \mu + \sigma) = 0.6826895$$
$$P(\mu - 2\sigma \leq X \leq \mu + 2\sigma) = 0.9544997$$
$$P(\mu - 3\sigma \leq X \leq \mu + 3\sigma) = 0.9973002$$
$$P(\mu - 4\sigma \leq X \leq \mu + 4\sigma) = 0.9999366$$
$$P(\mu - 5\sigma \leq X \leq \mu + 5\sigma) = 0.9999994$$
$$P(\mu - 6\sigma \leq X \leq \mu + 6\sigma) = 0.999999998$$

9-17

Bemerkung 9.20

- Wahrscheinlichkeiten für $X \sim N(\mu, \sigma^2)$ lassen sich auf Wahrscheinlichkeiten für $U \sim N(0, 1)$ zurückführen.
- Für $X \sim N(\mu, \sigma^2)$ gilt
$$\begin{aligned} F_X(x) &= P(X \leq x) \\ &= P\left(\frac{X - \mu}{\sigma} \leq \frac{x - \mu}{\sigma}\right) \\ &= P\left(U \leq \frac{x - \mu}{\sigma}\right) \\ &= \Phi\left(\frac{x - \mu}{\sigma}\right). \end{aligned}$$

9-18

Bemerkung 9.21
Für k-Sigma-Bereiche gilt

$$\begin{aligned} P(\mu - k\sigma \le X \le \mu + k\sigma) &= P\left(-k \le \frac{X-\mu}{\sigma} \le k\right) \\ &= P(-k \le U \le k) \\ &= P(U \le k) - P(U < -k) \\ &= P(U \le k) - P(U \le -k) \\ &= \Phi(k) - \Phi(-k) \\ &= \Phi(k) - (1 - \Phi(k)) \\ &= 2\Phi(k) - 1\,. \end{aligned}$$

9.5 Ergänzungen

Bemerkung 9.a (Zur Selbstkontrolle) Im Kapitel 9 eingeführte Begriffe und Konzepte: stetige Zufallsvariable, Dichtefunktion, Rechteckverteilung (uniforme Verteilung, Gleichverteilung), Exponentialverteilung, Normal- oder Gauß-Verteilung, Standard-Normalverteilung, Messfehlermodell, k-Sigma-Bereich.

Bemerkung 9.b (Eigenschaften der Normalverteilung)

1. Es gilt
$$\Phi(-u) = 1 - \Phi(u)$$
wegen
$$\Phi(-u) = P(U \le -u) = P(U \ge u) = P(U > u) = 1 - \Phi(u).$$

2. Für $X \sim N(\mu, \sigma^2)$ gilt
$$\begin{aligned} P(a \le X \le b) &= F_X(b) - F_X(a) \\ &= \Phi\left(\frac{b-\mu}{\sigma}\right) - \Phi\left(\frac{a-\mu}{\sigma}\right). \end{aligned}$$

3. Für $X \sim N(\mu, \sigma^2)$ gilt
$$f_X(x) = \frac{1}{\sigma}\varphi\left(\frac{x-\mu}{\sigma}\right),$$
da
$$f_X(x) = F_X'(x) = \frac{d}{dx}\Phi\left(\frac{x-\mu}{\sigma}\right) = \frac{1}{\sigma}\varphi\left(\frac{x-\mu}{\sigma}\right).$$

Beispiel 9.c (Renditeverteilung) Für die Tagesrendite X gelte $X \sim N(\mu, \sigma^2)$ mit den Parametern $\mu = 0.0004$ und $\sigma = 0.0123$.
1. Wie groß ist die Wahrscheinlichkeit einer negativen Tagesrendite?

$$\begin{aligned} P(X < 0) &= P(X \le 0) = P(X - \mu \le -\mu) = P\left(\frac{X-\mu}{\sigma} \le \frac{-\mu}{\sigma}\right) = \Phi\left(-\frac{4}{123}\right) \\ &= 1 - \Phi\left(\frac{4}{123}\right) = 1 - 0.51 = 0.49 \end{aligned}$$

2. Wie groß ist $P(X \le \mu - 4\sigma)$?

$$P(X \le \mu - 4\sigma) = P\left(\frac{X-\mu}{\sigma} \le -4\right) = \Phi(-4) = 1 - \Phi(4) = 1 - 0.999968 = 0.000032$$

3. Wie groß sind die k-Sigma-Bereiche für $k = 1, \ldots, 4$?

k	$[\mu - k\sigma, \mu + k\sigma]$	$P(\mu - k\sigma \leq X \leq \mu + k\sigma)$
1	$[-0.0119, 0.0127]$	68.27%
2	$[-0.0242, 0.0250]$	95.45%
3	$[-0.0365, 0.0373]$	99.73%
4	$[-0.0488, 0.0496]$	99.994%

Bemerkung 9.d (Abweichende Notation bei der Normalverteilung) Manchmal werden Normalverteilungen über die beiden Parameter μ und σ parametrisiert. Während hier, und überwiegend in der Literatur, $N(1, 4)$ eine Normalverteilung mit den Parametern $\mu = 1$ und $\sigma^2 = 4$ bezeichnet, ist dann mit $N(1, 4)$ eine Normalverteilung mit $\mu = 1$, $\sigma = 4$ und somit $\sigma^2 = 16$ gemeint.

Kapitel 10

Erwartungswert und Varianz

10-1

10. Erwartungswert und Varianz

10.1 Erwartungswert
10.2 Varianz und Standardabweichung

10-2

10.1 Erwartungswert

Definition 10.1 (Erwartungswert im endlichen Fall)
X sei diskret mit der Wahrscheinlichkeitsfunktion $f_X(x) = P(X = x)$ und dem endlichen Träger T_X. Dann heißt die Zahl

$$\mathbb{E}[X] \overset{\text{def}}{=} \sum_{x \in T_X} xP(X = x) = \sum_{x \in T_X} x f_X(x)$$

der **Erwartungswert** von X (*expectation of* X, *mean of* X).

Beispiel 10.2
X bezeichne die Augenzahl beim fairen Würfel. Dann gilt

$$\mathbb{E}[X] = \sum_{i=1}^{6} iP(X = i) = \frac{1}{6}\sum_{i=1}^{6} i = \frac{1}{6}21 = \frac{7}{2}.$$

10-3

Definition 10.3 (Erwartungswert im diskreten Fall)
X sei diskret mit der Wahrscheinlichkeitsfunktion $f(x) = P(X = x)$ und dem abzählbar unendlichen Träger $T_X = \{x_1, x_2, \ldots\}$. Falls die Reihe

$$\sum_{i=1}^{\infty} |x_i| f(x_i)$$

konvergiert, heißt

$$\mathbb{E}[X] \stackrel{\text{def}}{=} \sum_{i=1}^{\infty} x_i f(x_i)$$

Erwartungswert von X.

Bemerkung 10.4

- Aus der Konvergenz von $\sum_{i=1}^{\infty} |x_i| f(x_i)$ folgt die Konvergenz von $\sum_{i=1}^{\infty} x_i f(x_i)$ und damit die Existenz des Erwartungswertes.
- Falls die Reihe nicht konvergiert, sagt man: „Der Erwartungswert existiert nicht“.

10-4

Beispiel 10.5
Es sei $X \sim B(1, \pi)$. Dann ist der Träger $T_X = \{0, 1\}$ und es gilt

$$\mathbb{E}[X] = 0P(X = 0) + 1P(X = 1) = 0(1 - \pi) + 1\pi = \pi.$$

Definition 10.6 (Erwartungswert im stetigen Fall)
X sei stetig mit Dichtefunktion f.

$$\mathbb{E}[X] \stackrel{\text{def}}{=} \int_{-\infty}^{\infty} x f(x)\, dx$$

heißt **Erwartungswert** von X, falls das Integral als reelle Zahl existiert.

Beispiel 10.7
Für $X \sim R(\alpha, \beta)$ gilt

$$\mathbb{E}[X] = \int_{\alpha}^{\beta} x \frac{1}{\beta - \alpha}\, dx = \cdots = \frac{\alpha + \beta}{2}.$$

10-5

Beispiel 10.8
Für $X \sim N(\mu, \sigma^2)$ gilt

$$\mathbb{E}[X] = \mu.$$

Bemerkung 10.9

- Häufig verwendet man auch für Zufallsvariablen mit anderen Verteilungen die Bezeichnung μ_X oder noch kürzer μ als Bezeichnung von $\mathbb{E}[X]$. Bei der Verwendung von μ besteht Verwechselungsgefahr mit dem Parameter μ der Normalverteilung.
- Üblich ist auch die abkürzende Schreibweise $\mathbb{E}X$ für $\mathbb{E}[X]$.

10-6

Bemerkung 10.10 X sei eine Zufallsvariable. Dann ist

$$Y = g(X)$$

für viele Funktionen $g : \mathbb{R} \to \mathbb{R}$ ebenfalls eine Zufallsvariable,
z. B. für

$$g(x) = x^2, x^3, a + bx, \ldots .$$

Die Berechnung von $\mathbb{E}[Y]$ kann dabei mit der Verteilung von Y erfolgen, aber auch mit der Verteilung von X.

Satz 10.11
X sei diskret mit der Wahrscheinlichkeitsfunktion f_X und dem Träger T_X. Falls $\mathbb{E}[g(X)]$ existiert, gilt

$$\mathbb{E}[g(X)] = \sum_{x \in T_X} g(x) f_X(x).$$

10-7

Beispiel 10.12
X bezeichne die Augenzahl beim fairen Würfel. Dann ist $T_X = \{1, 2, 3, 4, 5, 6\}$ und $f_X(x) = P(X = x) = \frac{1}{6}$ für $x \in T_X$.

- X^2 ist eine Zufallsvariable mit dem Erwartungswert

$$\mathbb{E}\left[X^2\right] = \sum_{x \in T_X} x^2 f_X(x) = \frac{1}{6}\left(1^2 + 2^2 + 3^2 + 4^2 + 5^2 + 6^2\right) = \frac{91}{6}.$$

- $1/X$ ist eine Zufallsvariable mit dem Erwartungswert

$$\mathbb{E}\left[\frac{1}{X}\right] = \sum_{x \in T_X} \frac{1}{x} f_X(x) = \frac{1}{6}\left(1 + \frac{1}{2} + \frac{1}{3} + \frac{1}{4} + \frac{1}{5} + \frac{1}{6}\right) = \frac{49}{120}.$$

- Es ist

$$\mathbb{E}\left[X^2\right] \neq (\mathbb{E}[X])^2 = \left(\frac{7}{2}\right)^2 = \frac{49}{4} \quad \text{und} \quad \mathbb{E}\left[\frac{1}{X}\right] \neq \frac{1}{\mathbb{E}[X]} = \frac{2}{7}.$$

10-8

Satz 10.13
X sei stetig mit der Dichtefunktion f_X. Falls $\mathbb{E}[g(X)]$ existiert, gilt

$$\mathbb{E}[g(X)] = \int_{-\infty}^{\infty} g(x) f_X(x)\, dx.$$

Beispiel 10.14
Es sei $X \sim R(0, 1)$.

$$\mathbb{E}[X^2] = \int_{-\infty}^{\infty} x^2 f_X(x)\, dx = \int_0^1 x^2\, dx = \frac{1}{3}.$$

10-9

Bemerkung 10.15 (Regeln für den Erwartungswert)

- Erwartungswert einer linearen Transformation:
$$\mathbb{E}[a + bX] = a + b\mathbb{E}[X]$$
- Im Allgemeinen ist
$$\mathbb{E}[g(X)] \neq g(\mathbb{E}[X])\,.$$
- Erwartungswert einer Summe und einer Linearkombination von Zufallsvariablen:
$$\mathbb{E}[X + Y] = \mathbb{E}[X] + \mathbb{E}[Y]$$
$$\mathbb{E}[aX + bY] = a\mathbb{E}[X] + b\mathbb{E}[Y]$$

10-10

Beispiel 10.16
$X_1, \ldots, X_n$ seien Zufallsvariablen mit demselben Erwartungswert $\mu = \mathbb{E}[X_i]$. Für
$$\bar{X} \overset{\text{def}}{=} \frac{1}{n}\sum_{i=1}^{n} X_i$$
gilt $\mathbb{E}[\bar{X}] = \mu$. Dies zeigen die Umformungen
$$\mathbb{E}[\bar{X}] = \mathbb{E}\left[\frac{1}{n}\sum_{i=1}^{n} X_i\right] = \frac{1}{n}\mathbb{E}\left[\sum_{i=1}^{n} X_i\right] = \frac{1}{n}\sum_{i=1}^{n}\mathbb{E}[X_i] = \frac{1}{n}n\mu = \mu.$$

10-11

10.2 Varianz und Standardabweichung

Bemerkung 10.17
Die Maßzahlen Varianz und Standardabweichung messen die Streuung einer Zufallsvariablen um den Erwartungswert.

Definition 10.18 (Varianz)
$\mathbb{E}[X^2]$ existiere. Dann heißt die reelle Zahl
$$\mathbb{V}[X] \overset{\text{def}}{=} \mathbb{E}[(X - \mathbb{E}[X])^2]$$
Varianz von X.

Bemerkung 10.19
Falls $\mathbb{E}[X^2]$ existiert, dann existieren auch $\mathbb{E}[X]$ und $\mathbb{V}[X]$.

10-12

Satz 10.20 (Verschiebungsdarstellung der Varianz)
Falls $\mathbb{E}[X^2]$ *existiert, gilt*

$$\mathbb{V}[X] = \mathbb{E}[X^2] - (\mathbb{E}[X])^2.$$

Beispiel 10.21
X bezeichne die zufällige Augenzahl beim fairen Würfelwurf. Es wurden bereits $\mathbb{E}[X] = \frac{7}{2}$ und $\mathbb{E}[X^2] = \frac{91}{6}$ berechnet. Die Varianz der Zufallsvariablen X ist

$$\mathbb{V}[X] = \mathbb{E}[X^2] - (\mathbb{E}[X])^2 = \frac{91}{6} - \frac{49}{4} = \frac{35}{12}.$$

10-13

Beispiel 10.22

- Für $X \sim R(\alpha, \beta)$ gilt
$$\mathbb{V}[X] = \frac{(\beta - \alpha)^2}{12}.$$
- Es sei $X \sim N(\mu, \sigma^2)$. Dann gilt
$$\mathbb{V}[X] = \sigma^2.$$

10-14

Bemerkung 10.23 (Eigenschaften der Varianz)

- $\mathbb{V}[X] \geq 0$
- $\mathbb{V}[X] = 0 \quad \Longleftrightarrow \quad P(X = \mathbb{E}[X]) = 1$
- $\mathbb{V}[X]$ ist ein Streuungsmaß.
- $\mathbb{V}[a + X] = \mathbb{V}[X]$
- $\mathbb{V}[bX] = b^2\mathbb{V}[X]$
- $\mathbb{V}[a + bX] = b^2\mathbb{V}[X]$

Bemerkung 10.24
Häufig wird auch für Zufallsvariablen mit anderen Verteilungen die Bezeichnung

$$\sigma_X^2 \stackrel{\text{def}}{=} \mathbb{V}[X]$$

gewählt. Wenn verkürzend σ^2 für σ_X^2 geschrieben wird, besteht eine Verwechselungsmöglichkeit mit dem Parameter σ^2 der Normalverteilung.

10-15

Definition 10.25 (Standardabweichung)
Falls $\mathbb{V}[X]$ existiert, heißt

$$\sigma[X] \stackrel{\text{def}}{=} \sqrt{\mathbb{V}[X]}$$

Standardabweichung von X.

Bemerkung 10.26 (Eigenschaften der Standardabweichung)

- $\sigma[X] \geq 0$
- $\sigma[a + X] = \sigma[X]$
- $\sigma[bX] = |b|\sigma[X]$
- $\sigma[a + bX] = |b|\sigma[X]$

10-16

Definition 10.27 (Standardisierung)
X sei eine Zufallsvariable mit $0 < \sigma[X] < \infty$. Dann heißt

$$\tilde{X} \stackrel{\text{def}}{=} \frac{X - \mathbb{E}[X]}{\sigma[X]}$$

standardisierte Zufallsvariable zu X, d. h. $\mathbb{E}[\tilde{X}] = 0$ und $\sigma[\tilde{X}] = 1$.

Satz 10.28
Z sei eine standardisierte Zufallsvariable. Dann gilt

$$\mathbb{E}[Z] = 0 \quad \textit{und} \quad \sigma[Z] = \mathbb{V}[Z] = 1.$$

Beispiel 10.29
Es sei $X \sim N(\mu, \sigma^2)$. Dann gilt

$$\tilde{X} = \frac{X - \mu}{\sigma} \sim N(0, 1).$$

10-17

Bemerkung 10.30 (Spezielle Verteilungen)

$X \sim$	$\mathbb{E}[X] =$	$\mathbb{V}[X] =$	$\sigma[X] =$
$N(0,1)$	0	1	1
$N(\mu,\sigma^2)$	μ	σ^2	σ
$B(1,\pi)$	π	$\pi(1-\pi)$	$\sqrt{\pi(1-\pi)}$
$B(n,\pi)$	$n\pi$	$n\pi(1-\pi)$	$\sqrt{n\pi(1-\pi)}$
$Poi(\mu)$	μ	μ	$\sqrt{\mu}$
$Exp(1)$	1	1	1
$Exp(\lambda)$	$\frac{1}{\lambda}$	$\frac{1}{\lambda^2}$	$\frac{1}{\lambda}$
$R(0,1)$	$\frac{1}{2}$	$\frac{1}{12}$	$\frac{1}{\sqrt{12}}$
$R(\alpha,\beta)$	$\frac{\alpha+\beta}{2}$	$\frac{(\beta-\alpha)^2}{12}$	$\frac{\beta-\alpha}{\sqrt{12}}$

10.3 Ergänzungen

Bemerkung 10.a (Zur Selbstkontrolle) In diesem Kapitel eingeführte Begriffe und Konzepte: Erwartungswert, Regeln für das Rechnen mit Erwartungswerten, Varianz, Verschiebungsdarstellung der Varianz, Standardabweichung, Standardisierung.

Bemerkung 10.b (Funktion einer Zufallsvariablen) X sei eine Zufallsvariable. Dann ist

$$Y = g(X)$$

für viele Funktionen $g : \mathbb{R} \to \mathbb{R}$ ebenfalls eine Zufallsvariable. Wegen

$$X : \Omega \to \mathbb{R}, \quad g : \mathbb{R} \to \mathbb{R}$$

kann Y als Zufallsvariable auf Ω interpretiert werden:

$$Y : \Omega \to \mathbb{R}, \quad Y(\omega) = g(X(\omega)).$$

Die Verteilungsfunktion von $Y = g(X)$ ergibt sich eindeutig aus der Verteilungsfunktion von X:

$$P(Y \leq y) = P(g(X) \leq y) = P(\{\omega | g(X(\omega)) \leq y\}).$$

Bemerkung 10.c (Erwartungswert einer linearen Funktion) Mit der die affin-linearen Funktion $\ell(x) = a + bx$ lässt sich

$$\mathbb{E}[a + bX] = a + b\mathbb{E}[X]$$

in der Form

$$\mathbb{E}[\ell(X)] = \ell(\mathbb{E}[X])$$

schreiben. Für beliebige Funktionen g gilt aber im Allgemeinen

$$\mathbb{E}[g(X)] \neq g(\mathbb{E}[X])\,.$$

Kapitel 11

Gemeinsame Verteilung von Zufallsvariablen

11-1

11. Gemeinsame Verteilung von Zufallsvariablen

11.1 Gemeinsame Verteilung von zwei Zufallsvariablen
11.2 Unabhängigkeit von zwei Zufallsvariablen
11.3 Kovarianz und Korrelation
11.4 Verteilung mehrerer Zufallsvariablen
11.5 Summen unabhängiger Zufallsvariablen
11.6 Randverteilung

11-2

11.1 Gemeinsame Verteilung von zwei Zufallsvariablen

Bemerkung 11.1
Werden bei einem Zufallsexperiment mindestens zwei Merkmale gleichzeitig beobachtet, so führt dies zur **gemeinsamen Wahrscheinlichkeitsverteilung** (oder kurz **gemeinsamen Verteilung**) mehrerer Zufallsvariablen.

Beispiel 11.2
Bei der gleichzeitigen Messung der Körpergröße X (in cm) und des Körpergewichts Y (in kg) einer zufällig ausgewählten Person ist (X, Y) eine zweidimensionale Zufallsvariable mit möglichen Werten

$$(x, y) \in \,]0, \infty[\,^2 \subset \mathbb{R}^2$$

und es interessieren z. B. Wahrscheinlichkeiten der Art

$$P(X \geq 185, Y \geq 90)\,.$$

11-3

Bemerkung 11.3

- Zunächst wird nur die gemeinsame Verteilung von zwei Zufallsvariablen betrachtet, die mit X und Y bezeichnet werden, später wird die gemeinsame Verteilung von n Zufallsvariablen betrachtet, die dann mit $X_1, \ldots, X_n$ bezeichnet werden.
- Es wird davon ausgegangen, dass die betrachteten Zufallsvariablen X und Y so durch ein gemeinsames Zufallsexperiment bestimmt sind, dass die Wahrscheinlichkeiten

  $$P(X \leq x, Y \leq y) \quad \text{für alle} \quad x, y \in \mathbb{R}$$

  bestimmt werden können.
- Man sagt dann auch, dass eine gemeinsame Wahrscheinlichkeitsverteilung oder kurz eine **gemeinsame Verteilung** der Zufallsvariablen X und Y existiert.

11-4

Definition 11.4 (Gemeinsame Verteilungsfunktion)
X und Y seien Zufallsvariablen mit einer gemeinsamen Verteilung. Die Funktion $F_{XY} : \mathbb{R}^2 \to [0, 1]$,

$$F_{XY}(x, y) = P(X \leq x, Y \leq y),$$

heißt **gemeinsame Verteilungsfunktion** der Zufallsvariablen X und Y.

11-5

Definition 11.5 (Gemeinsame diskrete Verteilung)
Die gemeinsame Verteilung von zwei Zufallsvariablen X und Y heißt **diskret**, falls es endlich oder abzählbar unendlich viele Werte $x_1, x_2, \ldots$ und $y_1, y_2, \ldots$ gibt, so dass

$$\sum_j \sum_k P(X = x_j, Y = y_k) = 1$$

gilt.

Definition 11.6 (Gemeinsame Wahrscheinlichkeitsfunktion)
Die gemeinsame Verteilung der Zufallsvariablen X und Y sei diskret. Die Funktion $f_{XY} : \mathbb{R}^2 \to [0, 1]$,

$$f_{XY}(x, y) = P(X = x, Y = y),$$

heißt **gemeinsame Wahrscheinlichkeitsfunktion** der Zufallsvariablen X und Y.

11-6

Definition 11.7 (Gemeinsame stetige Verteilung)

- Die gemeinsame Verteilung von zwei Zufallsvariablen X und Y heißt **stetig**, falls es eine Funktion $f_{XY} : \mathbb{R}^2 \to [0, \infty[$ gibt, so dass

$$F_{XY}(x, y) = \int_{-\infty}^{y} \int_{-\infty}^{x} f_{XY}(z, v) dz dv$$

  für alle $x, y \in \mathbb{R}$ gilt.

- Die Funktion f_{XY} heißt dann **gemeinsame Dichtefunktion** der Zufallsvariablen X und Y.

11-7

11.2 Unabhängigkeit von zwei Zufallsvariablen

Definition 11.8 (Stochastische Unabhängigkeit)
Die Zufallsvariablen X und Y mit der gemeinsamen Verteilungsfunktion F_{XY} und den Verteilungsfunktionen F_X und F_Y heißen **stochastisch unabhängig**, wenn

$$F_{XY}(x, y) = F_X(x) F_Y(y) \quad \text{für alle } x, y \in \mathbb{R}.$$

Bemerkung 11.9
Häufig sagt man kurz **unabhängig** anstatt stochastisch unabhängig, wenn keine Verwechselungsgefahr mit anderen Unabhängigkeitsbegriffen, z. B. der linearen Unabhängigkeit, besteht.

11-8

Satz 11.10
Die Zufallsvariablen X und Y mit gemeinsamer Verteilung sind genau dann stochastisch unabhängig, wenn

$$P(X \in A, Y \in B) = P(X \in A) \cdot P(Y \in B)$$

für alle $A \subset \mathbb{R}$ und $B \subset \mathbb{R}$ gilt, für welche die Wahrscheinlichkeiten $P(X \in A)$ und $P(Y \in B)$ definiert sind.

11-9

Beispiel 11.11
Die Zufallsvariable W bezeichne die zufällige Augenzahl beim fairen Würfelwurf. Die beiden Zufallsvariablen

$$X = \begin{cases} 0 & \text{falls } W \in \{1,3,5\} \\ 1 & \text{falls } W \in \{2,4,6\}, \end{cases}$$

und

$$Y = \begin{cases} 0 & \text{falls } W \in \{1,2,3\} \\ 1 & \text{falls } W \in \{4,5,6\} \end{cases}$$

sind **nicht** stochastisch unabhängig, da

$$\begin{aligned} P(X \leq 0, Y \leq 0) &= P(W \in \{1,3,5\} \cap \{1,2,3\}) = P(W \in \{1,3\}) = \frac{2}{6} = \frac{1}{3} \\ &\neq P(X \leq 0)P(Y \leq 0) = \frac{1}{2} \cdot \frac{1}{2} = \frac{1}{4}\,. \end{aligned}$$

11-10

Für X und

$$Z = \begin{cases} 0 & \text{falls } W \in \{3,4,5,6\} \\ 1 & \text{falls } W \in \{1,2\} \end{cases}$$

gilt

$$P(X \leq 0, Z \leq 0) = P(W \in \{1,3,5\} \cap \{3,4,5,6\}) = P(W \in \{3,5\}) = \frac{1}{3}$$

und

$$P(X \leq 0)P(Z \leq 0) = \frac{1}{2} \cdot \frac{2}{3} = \frac{1}{3}\,,$$

d. h. die **Ereignisse** „$X \leq 0$“ und „$Z \leq 0$“ sind stochastisch unabhängig.
Analog zeigt man

$$P(X \leq x, Z \leq z) = P(X \leq x) \cdot P(Z \leq z) \quad \text{für alle} \quad x, z \in \mathbb{R}\,.$$

Also sind die **Zufallsvariablen** X und Z stochastisch unabhängig.

11-11

Satz 11.12 (Kriterium für Unabhängigkeit)

- *Die Zufallsvariablen X und Y seien gemeinsam stetig verteilt mit der gemeinsamen Dichtefunktion f_{XY} und den Dichtefunktionen f_X und f_Y. Die Zufallsvariablen X und Y sind genau dann stochastisch unabhängig, wenn*

  $$f_{XY}(x,y) = f_X(x) \cdot f_Y(y) \quad \text{für alle } x, y \in \mathbb{R}$$

  gilt.

- *Die Zufallsvariablen X und Y seien gemeinsam diskret verteilt. X und Y sind genau dann stochastisch unabhängig, wenn*

  $$P(X = x, Y = y) = P(X = x) \cdot P(Y = y) \quad \text{für alle } x, y \in \mathbb{R}$$

  gilt.

11-12

11.3 Kovarianz und Korrelation

Bemerkung 11.13

- Wenn die Zufallsvariablen X und Y eine gemeinsame Wahrscheinlichkeitsverteilung besitzen, so ist für eine vorgegebene Funktion $g : \mathbb{R}^2 \to \mathbb{R}$ in der Regel auch $g(X, Y)$ eine Zufallsvariable, für die sich der Erwartungswert $\mathbb{E}[g(X, Y)]$ berechnen lässt.
- Im Fall einer gemeinsamen diskreten Verteilung von zwei Zufallsvariablen X und Y mit zweidimensionaler Wahrscheinlichkeitsfunktion f_{XY} gilt

  $$\mathbb{E}[g(X,Y)] = \sum_x \sum_y g(x,y)P(X = x, Y = y) = \sum_x \sum_y g(x,y) f_{XY}(x,y)\,.$$

  Für den Spezialfall $g(X, Y) = XY$ gilt

  $$\mathbb{E}[XY] = \sum_x \sum_y xyP(X = x, Y = y) = \sum_x \sum_y xy f_{XY}(x,y)\,.$$

11-13

Definition 11.14 (Kovarianz)
Falls $\mathbb{E}[X], \mathbb{E}[Y]$ und $\mathbb{E}[XY]$ existieren, heißt

$$\sigma_{XY} \stackrel{\text{def}}{=} \mathbb{C}\text{ov}[X,Y] \stackrel{\text{def}}{=} \mathbb{E}\left[\big(X - \mathbb{E}[X]\big)\big(Y - \mathbb{E}[Y]\big)\right]$$

Kovarianz von X und Y.

Bemerkung 11.15 (Eigenschaften der Kovarianz)

- $\mathbb{C}\text{ov}[X,Y] = \mathbb{E}[XY] - \mathbb{E}[X]\mathbb{E}[Y]$
- $\mathbb{C}\text{ov}[X,X] = \mathbb{V}[X]$
- $\mathbb{C}\text{ov}[Y,X] = \mathbb{C}\text{ov}[X,Y]$
- $\mathbb{V}[X + Y] = \mathbb{V}[X] + \mathbb{V}[Y] + 2\mathbb{C}\text{ov}[X,Y]$
- $\mathbb{V}[aX \pm bY] = a^2\mathbb{V}[X] \pm 2ab\mathbb{C}\text{ov}[X,Y] + b^2\mathbb{V}[Y]$

11-14

- Wenn X und Y stochastisch unabhängig sind, dann gilt $\mathbb{C}\text{ov}[X,Y] = 0$.
- Aus $\mathbb{C}\text{ov}[X,Y] = 0$ folgt nicht, dass X und Y stochastisch unabhängig sind.

Definition 11.16 (Korrelationskoeffizient)
Falls X und Y eine gemeinsame Wahrscheinlichkeitsverteilung besitzen und $0 < \mathbb{V}[X] < \infty$ sowie $0 < \mathbb{V}[Y] < \infty$ gelten, heißt

$$\rho_{XY} \stackrel{\text{def}}{=} \mathbb{C}\text{orr}[X,Y] \stackrel{\text{def}}{=} \frac{\mathbb{C}\text{ov}[X,Y]}{\sqrt{\mathbb{V}[X]\mathbb{V}[Y]}} = \frac{\mathbb{C}\text{ov}[X,Y]}{\sigma[X]\sigma[Y]}$$

Korrelationskoeffizient von X und Y.

11-15

Bemerkung 11.17 (Eigenschaften)

- $-1 \leq \rho_{XY} \leq 1$
- $\rho_{YX} = \rho_{XY}$
- Es sei $a \in \mathbb{R}$ und $b > 0$, dann gelten

$$Y = a + bX \quad \Longrightarrow \quad \rho_{XY} = 1,$$

$$Y = a - bX \quad \Longrightarrow \quad \rho_{XY} = -1.$$

Definition 11.18 (Unkorreliert, vollkommen korreliert)

- X und Y heißen **unkorreliert**, falls $\rho_{XY} = 0$ gilt.
- X und Y heißen **vollkommen korreliert**, falls $\rho_{XY} \in \{-1, 1\}$ gilt.

11-16

11.4 Verteilung mehrerer Zufallsvariablen

Bemerkung 11.19

- Für n Zufallsvariablen $X_1, \ldots, X_n$ seien die Wahrscheinlichkeiten

$$P(X_1 \leq x_1, \ldots, X_n \leq x_n) \quad \text{für alle} \quad x_1, \ldots, x_n \in \mathbb{R}$$

  durch ein gemeinsames Zufallsexperiment bestimmt. Man sagt, dass eine gemeinsame Wahrscheinlichkeitsverteilung oder kurz eine **gemeinsame Verteilung** von $X_1, \ldots, X_n$ existiert.

- Man spricht dann auch von einem **n-dimensionalen Zufallsvektor** $(X_1, \ldots, X_n)$ mit den Komponenten $X_1, \ldots, X_n$.

11-17

Definition 11.20 (Stochastische Unabhängigkeit)
Die Zufallsvariablen $X_1, \ldots, X_n$ mit gemeinsamer Verteilung heißen **stochastisch unabhängig**, wenn

$$P(X_1 \leq x_1, \ldots, X_n \leq x_n) = \prod_{i=1}^{n} P(X_i \leq x_i) \quad \text{für alle } x_1, \ldots, x_n \in \mathbb{R}.$$

Bemerkung 11.21

- Aus der stochastischen Unabhängigkeit folgt die paarweise Unabhängigkeit, d. h. dass jeweils zwei verschiedene Zufallsvariablen X_i und X_j mit $i \neq j$ stochastisch unabhängig sind.
- Aus der paarweisen stochastischen Unabhängigkeit folgt aber nicht die stochastische Unabhängigkeit im Sinn von Definition 11.20.

11-18

11.5 Summen unabhängiger Zufallsvariablen

Satz 11.22 (Binomialverteilung)

1. *Für stochastisch unabhängige Zufallsvariablen $X_i \sim B(1, \pi)$, $i = 1, \ldots, n$, gilt*
$$\sum_{i=1}^{n} X_i \sim B(n, \pi).$$
2. *Die Zufallsvariablen $X_1, ..., X_k$ seien stochastisch unabhängig mit $X_i \sim B(n_i, \pi)$ für $i = 1, \ldots, k$, dann gilt*
$$\sum_{i=1}^{k} X_i \sim B\left(\sum_{i=1}^{k} n_i, \pi\right).$$

11-19

Satz 11.23 (Summen unabhängiger Zufallsvariablen)
Die Zufallsvariablen $X_1, ..., X_n$ seien stochastisch unabhängig.

- *Aus $X_i \sim Poi(\mu_i)$, $i = 1, \ldots, n$, folgt*
$$\sum_{i=1}^{n} X_i \sim Poi\left(\sum_{i=1}^{n} \mu_i\right).$$
- *Aus $X_i \sim N(\mu_i, \sigma_i^2)$, $i = 1, \ldots, n$, folgt*
$$\sum_{i=1}^{n} X_i \sim N\left(\sum_{i=1}^{n} \mu_i, \sum_{i=1}^{n} \sigma_i^2\right).$$

11-20

11.6 Randverteilung

Bemerkung 11.24
Mit der gemeinsamen Verteilung von zwei Zufallsvariablen X und Y liegen auch die Verteilungen von X und Y fest, denn es gilt

$$P(X \leq x) = P(X \leq x, Y \in \mathbb{R}), \quad x \in \mathbb{R}$$

und
$$P(Y \leq y) = P(X \in \mathbb{R}, Y \leq y), \quad y \in \mathbb{R}\,.$$

11-21

Definition 11.25 (Randwahrscheinlichkeiten)
Ist die gemeinsame Verteilung von X und Y diskret, wobei es endlich oder abzählbar unendlich viele Werte $x_1, x_2, \ldots$ und $y_1, y_2, \ldots$ mit

$$\sum_j \sum_k P(X = x_j, Y = y_k) = 1$$

gibt, dann sind

$$P(X = x_j) = \sum_k P(X = x_j, Y = y_k), \quad j = 1, 2, \ldots,$$

und

$$P(Y = y_k) = \sum_j P(X = x_j, Y = y_k), \quad k = 1, 2, \ldots,$$

die **Randwahrscheinlichkeiten** von X bzw. Y.

11.7 Ergänzungen

Bemerkung 11.a (Zur Selbstkontrolle) In diesem Kapitel eingeführte Begriffe und Konzepte: gemeinsame Verteilung, gemeinsame Verteilungs-, Wahrscheinlichkeits- und Dichtefunktion, stochastische Unabhängigkeit von Zufallsvariablen, Kovarianz, Korrelationskoeffizient, unkorreliert, vollkommen korreliert, Randverteilung, Randwahrscheinlichkeiten.

Bemerkung 11.b Im Fall einer zweidimensionalen Dichtefunktion f_{XY} gilt

$$\mathbb{E}[g(X,Y)] = \int_{-\infty}^{+\infty} \int_{-\infty}^{+\infty} g(x,y) f_{XY}(x,y) dx dy \,.$$

Für den Spezialfall $g(X,Y) = XY$ gilt

$$\mathbb{E}[XY] = \int_{-\infty}^{+\infty} \int_{-\infty}^{+\infty} xy f_{XY}(x,y) dx dy.$$

Definition 11.c (Gemeinsame Verteilungsfunktion) $X_1, \ldots, X_n$ seien Zufallsvariablen mit einer gemeinsamen Verteilung. Die Funktion $F_{X_1, X_2, \ldots, X_n} : \mathbb{R}^n \to [0,1]$,

$$F_{X_1, X_2, \ldots, X_n}(x_1, \ldots, x_n) = P(X_1 \leq x_1, \ldots, X_n \leq x_n),$$

heißt **gemeinsame Verteilungsfunktion** der Zufallsvariablen $X_1, \ldots, X_n$ oder n-dimensionale Verteilungsfunktion des Zufallsvektors $(X_1, \ldots, X_n)$.

Satz 11.d (Stochastische Unabhängigkeit) Die Zufallsvariablen $X_1, \ldots, X_n$ mit den Verteilungsfunktionen $F_{X_1}, \ldots, F_{X_n}$ und der gemeinsamen Verteilungsfunktion $F_{X_1, \ldots, X_n}$ sind genau dann stochastisch unabhängig, wenn

$$F_{X_1, \ldots, X_n}(x_1, \ldots, x_n) = \prod_{i=1}^{n} F_{X_i}(x_i) \quad \text{für alle } x_1, \ldots, x_n \in \mathbb{R}.$$

Definition 11.e (Randverteilungsfunktion) F_{XY} sei die gemeinsame Verteilungsfunktion der Zufallsvariablen X und Y. Die Funktion $F_X : \mathbb{R} \to [0,1]$,

$$F_X(x) = \lim_{y \to \infty} F_{XY}(x,y),$$

heißt **Randverteilungsfunktion von** X. Die Funktion $F_Y : \mathbb{R} \to [0,1]$,

$$F_Y(y) = \lim_{x \to \infty} F_{XY}(x,y),$$

heißt **Randverteilungsfunktion von** Y.

Bemerkung 11.f F_X ist die Verteilungsfunktion von X und F_Y ist die Verteilungsfunktion von Y.

Definition 11.g (Randwahrscheinlichkeitsfunktion) Ist die gemeinsame Verteilung von X und Y diskret mit den Randwahrscheinlichkeiten

$$P(X = x_j), \quad j = 1, 2, \ldots,$$

und

$$P(Y = y_k), \quad k = 1, 2, \ldots,$$

dann sind

$$f_X(x) = P(X = x), \quad x \in \mathbb{R}$$

und

$$f_Y(y) = P(Y = y), \quad y \in \mathbb{R}$$

die **Randwahrscheinlichkeitsfunktionen** von X bzw. Y.

Bemerkung 11.h f_X ist die Wahrscheinlichkeitsfunktion von X. und f_Y ist die Wahrscheinlichkeitsfunktion von Y.

Definition 11.i (Randdichtefunktion) Ist die gemeinsame Verteilung von X und Y stetig mit der Dichtefunktion f_{XY}, dann heißt die Funktion $f_X : \mathbb{R} \to [0, \infty[$,

$$f_X(x) = \int_{-\infty}^{\infty} f_{XY}(x, y)\, dy$$

Randdichtefunktion von X und die Funktion $f_Y : \mathbb{R} \to [0, \infty[$,

$$f_Y(y) = \int_{-\infty}^{\infty} f_{XY}(x, y)\, dx$$

heißt **Randdichtefunktion** von Y.

Bemerkung 11.j f_X ist die Dichtefunktion von X und f_Y ist die Dichtefunktion von Y.

Kapitel 12

Parameterschätzung

12-1

12. Parameterschätzung

12.1 Grundgesamtheit und Stichprobe
12.2 Stichprobenfunktionen (Statistiken) und Parameter
12.3 Erwartungstreue
12.4 Wahrscheinlichkeitsverteilung eines Schätzers

12-2

12.1 Grundgesamtheit und Stichprobe

Bemerkung 12.1 (Stichprobe, Zufallsstichprobe)

- Eine **Stichprobenziehung** (oder Stichprobenentnahme) (engl.: *sampling*) ist die **zufällige** Entnahme von Einheiten aus einer **Grundgesamtheit** (engl.: *population*).
- Das Ergebnis der Stichprobenziehung ist eine **Stichprobe** (engl.: *sample*).
- Zufällig bedeutet nicht willkürlich, sondern realisiert durch ein Zufallsexperiment mit *uneingeschränkter* oder *reiner* **Zufallsauswahl**, wobei jede Einheit der Grundgesamtheit die gleiche Wahrscheinlichkeit hat, gewählt zu werden.
- Das Wort Stichprobe entstammt der Warenprüfung.

12-3

Bemerkung 12.2 (Stichprobe oder Totalerhebung?)
Gründe für eine Stichprobenerhebung anstatt einer Totalerhebung sind

- Zeitersparnis,
- Kostenersparnis,
- die Unmöglichkeit der Totalerhebung z. B. bei der Warenprüfung (Lebensdauer von Glühbirnen, Geschmacksüberprüfung, ...) oder
- das Vorliegen einer unendlichen, fiktiven Grundgesamtheit.

12-4

Bemerkung 12.3 (Ziehungsschema)
Es gibt zwei grundsätzliche Ziehungsschemata zur Durchführung einer Zufallsstichprobe:

a) Ziehen **ohne** Zurücklegen (engl.: *sampling without replacement*)

b) Ziehen **mit** Zurücklegen (engl.: *sampling with replacement*)

12-5

Bemerkung 12.4 (Stichprobenumfang)

- Im Folgenden gehen wir von einer **Stichprobe vom Umfang** n (engl.: *sample size*) aus einer Grundgesamtheit mit N Elementen aus.
- Es ist $n \in \mathbb{N}$ und $N \in \mathbb{N}$, wobei

  $$\mathbb{N} \stackrel{\text{def}}{=} \{1, 2, 3, \ldots\}$$

  die **Menge der natürlichen Zahlen** bezeichnet.
- Beim Ziehen **ohne** Zurücklegen gilt immer $n \leq N$.
- Beim Ziehen **mit** Zurücklegen ist theoretisch auch $n > N$ möglich.

12-6

Bemerkung 12.5 (Ziehungsschema und Unabhängigkeit)
Bei n aufeinanderfolgenden Ziehungen werde ein statistisches Merkmal beobachtet. Dann ergeben sich n Zufallsvariablen $X_1, \ldots, X_n$.

- Beim Ziehen **mit** Zurücklegen sind die Zufallsvariablen $X_1, \ldots, X_n$ stochastisch unabhängig.
- Beim Ziehen **ohne** Zurücklegen sind die Zufallsvariablen $X_1, \ldots, X_n$ nicht unabhängig und in der Regel paarweise negativ korreliert.

12-7

Bemerkung 12.6 (Die i.i.d.-Annahme)

- Im Folgenden wird in der Regel von stochastisch unabhängigen Zufallsvariablen ausgegangen. Bei den später behandelten Schätz- und Testverfahren ist eine immer wieder verwendete Standardannahme:

  Die Zufallsvariablen X_i $(i = 1, 2, \ldots, n)$ sind **stochastisch unabhängig und identisch verteilt**.

- Dafür ist die folgende Bezeichnung üblich:
  i.i.d. (*independent identically distributed*), $\overset{i.i.d.}{\sim}$
- Z. B. schreibt man

  $$X_i \overset{i.i.d.}{\sim} N(\mu, \sigma^2), \quad i = 1, \ldots, n$$

  für Zufallsvariablen, die stochastisch unabhängig verteilt mit derselben Normalverteilung $N(\mu, \sigma^2)$ sind.

12-8

Definition 12.7 (Zufallsstichprobe)

- Die Zufallsvariablen $X_1, X_2, \ldots, X_n$ heißen **Zufallsstichprobe aus** X (oder aus der Verteilung von X), wenn alle X_i wie X verteilt sind.
- Die Zahl n heißt **Stichprobenumfang** (*sample size*) oder Stichprobenlänge.
- Die Werte $x_1, x_2, \ldots, x_n$ der Zufallsvariablen $X_1, X_2, \ldots, X_n$ heißen **Realisierung der Stichprobe**, **Wert der Stichprobe** oder **Stichprobenwerte**.
- Eine Zufallsstichprobe $X_1, X_2, \ldots, X_n$ aus X heißt **einfache Zufallsstichprobe** aus X, falls die Zufallsvariablen $X_1, X_2, \ldots, X_n$ stochastisch unabhängig sind.

12-9

Bemerkung 12.8 (Einfache Zufallsstichprobe)

- Bei einer einfachen Zufallsstichprobe liegen i.i.d.-Variablen vor.
- Eine einfache Zufallsstichprobe entsteht z. B. beim Ziehen mit Zurücklegen.
- Im Folgenden wird stets eine einfache Zufallsstichprobe zugrundegelegt und diese wird kurz als **Stichprobe** bezeichnet

Bemerkung 12.9 (Stichprobenvariable)
Die i-te Zufallsvariable X_i heißt auch die i-te **Stichprobenvariable**.

Bemerkung 12.10
Ausgehend von einer Zufallsstichprobe $X_1, X_2, \ldots, X_n$ werden Funktionen der Stichprobenvariablen verwendet, um von der Stichprobe auf Eigenschaften der Grundgesamtheit zu schließen.

12-10

12.2 Stichprobenfunktionen (Statistiken) und Parameter

Beispiel 12.11 (Stichprobenfunktion)

- (X_1, X_2) bezeichne das zufällige Ergebnis von zwei unabhängigen Münzwürfen.
- Für die einzelnen Zufallsvariablen X_1 und X_2 gilt

$$P(X_i = 0) = P(X_i = 1) = \frac{1}{2}, \quad i = 1, 2\,.$$

- Die Zufallsvariablen X_1 und X_2 sind Bernoulli-verteilt mit dem Bernoulli-Parameter $\pi = 1/2$, d. h.

$$X_i \overset{i.i.d.}{\sim} B(1, 1/2), \quad i = 1, 2\,.$$

  Es gilt

$$\mathbb{E}[X_i] = \frac{1}{2}, \quad \mathbb{V}[X_i] = \frac{1}{4}\,.$$

12-11

- Für die gemeinsame Verteilung der Zufallsvariablen X_1 und X_2 gilt

$$P(X_1 = 0, X_2 = 0) = P(X_1 = 0)P(X_2 = 0) = \frac{1}{4}$$

  und allgemein gilt

$$P\left((X_1, X_2) = (x_1, x_2)\right) = P\left(X_1 = x_1, X_2 = x_2\right) = \frac{1}{4}$$

  für

$$(x_1, x_2) \in \{(0,0), (0,1), (1,0), (1,1)\} \subseteq \mathbb{R}^2.$$

- Die Summenvariable

$$Y = g(X_1, X_2) = X_1 + X_2$$

  ist ebenfalls eine Zufallsvariable und ein einfaches Beispiel einer sogenannten **Stichprobenfunktion** $g(X_1, X_2)$.

12-12

- Die Stichprobenfunktion Y hat die Wahrscheinlichkeitsverteilung

$$P(Y=0) = P(X_1=0, X_2=0) = \frac{1}{4},$$

$$P(Y=1) = P(X_1=0, X_2=1) + P(X_1=1, X_2=0) = \frac{1}{4} + \frac{1}{4} = \frac{1}{2},$$

und

$$P(Y=2) = P(X_1=1, X_2=1) = \frac{1}{4}\,.$$

- Die Zufallsvariable Y ist binomialverteilt mit den Parametern $n=2$ und dem Bernoulli-Parameter $\pi = 1/2$,

$$Y \sim B(2, 1/2).$$

Es gilt

$$\mathbb{E}[Y] = 1, \quad \mathbb{V}[Y] = \frac{1}{2}.$$

12-13

Definition 12.12 (Stichprobenfunktion, Statistik)
$X_1, \ldots, X_n$ sei eine Zufallsstichprobe, dann heißt eine Zufallsvariable $g(X_1, \ldots, X_n)$ **Stichprobenfunktion** *oder* **Statistik***.*

Bemerkung 12.13 (Schätzfunktionen, Prüfgrößen)
Weiter unten werden spezielle Stichprobenfunktionen, die zur Parameterschätzung verwendet werden, als **Schätzfunktionen** und solche, die zur Durchführung statistischer Tests verwendet werden, als **Testfunktionen** oder **Prüfgrößen** bezeichnet.

12-14

Bemerkung 12.14 (Beispiele von Stichprobenfunktionen)

- **Merkmalssumme**

$$\sum_{i=1}^{n} X_i$$

- **Stichprobenmittel**

$$\bar{X} \stackrel{\text{def}}{=} \frac{1}{n} \sum_{i=1}^{n} X_i$$

- **Stichprobenvarianz** oder **mittlere quadratische Abweichung** vom Stichprobenmittel $\bar{X}$

$$S^2 \stackrel{\text{def}}{=} \frac{1}{n} \sum_{i=1}^{n} (X_i - \bar{X})^2$$

- **Stichprobenstandardabweichung**

$$S \stackrel{\text{def}}{=} \sqrt{S^2}$$

12-15

Bemerkung 12.15

- Der Ausgangspunkt ist die **unbekannte Verteilung** eines Merkmales X in einer Grundgesamtheit.
- Es interessieren **Parameter** (Kennzahlen) θ_1, θ_2, ... dieser Verteilung.
- Mit einer Zufallsstichprobe $X_1, \ldots, X_n$ aus X soll auf die Parameter geschlossen werden.

12-16

Beispiel 12.16 (Endliche Grundgesamtheit)

- $\xi_1, \xi_2, \ldots, \xi_N$ sind die Werte eines Merkmals X in einer endlichen Grundgesamtheit vom Umfang N.
- Der **Mittelwert**

$$\theta_1 \stackrel{\text{def}}{=} \mu \stackrel{\text{def}}{=} \frac{1}{N}\sum_{i=1}^{N} \xi_i$$

und die **Varianz** des Merkmals X in der Grundgesamtheit

$$\theta_2 \stackrel{\text{def}}{=} \sigma^2 \stackrel{\text{def}}{=} \frac{1}{N}\sum_{i=1}^{N} (\xi_i - \mu)^2$$

sind die unbekannten **Parameter** der Grundgesamtheit.

- Mit einer Zufallsstichprobe $X_1, \ldots, X_n$ aus X soll auf die Parameter geschlossen werden.

12-17

Beispiel 12.17 (Verteilung als Grundgesamtheit)

- Bekannt ist das Verteilungsgesetz eines Merkmals X in einer unendlichen Grundgesamtheit, $X \sim N(\mu, \sigma^2)$.
- Unbekannt sind die Parameter $\theta_1 = \mu$ und $\theta_2 = \sigma^2$.
- Mit einer Zufallsstichprobe $X_1, \ldots, X_n$ aus X soll auf die Parameter geschlossen werden.

12-18

Bemerkung 12.18 (Parameterschätzung)

- θ sei ein **unbekannter Parameter** der Grundgesamtheit.
- Dieser soll aus einer **Stichprobe** vom Umfang n geschätzt werden.
- Ein **Schätzwert** (engl.: *estimate*)

$$\hat{\theta}(x_1, \ldots, x_n)$$

  für θ ist eine Realisation eines geeigneten **Schätzers** (engl.: *estimator*)

$$\hat{\theta} = \hat{\theta}(X_1, \ldots, X_n).$$

- Ein Schätzer heißt auch **Schätzfunktion**.

12-19

Beispiel 12.19 Es sei $X \sim N(\mu, \sigma^2)$, $n > 1$.

Parameter θ	Schätzwert	Schätzer
μ	$\bar{x} = \frac{1}{n}\sum_{i=1}^{n} x_i$	$\bar{X} = \frac{1}{n}\sum_{i=1}^{n} X_i$
σ^2	$s^2 = \frac{1}{n}\sum_{i=1}^{n}(x_i - \bar{x})^2$	$S^2 = \frac{1}{n}\sum_{i=1}^{n}(X_i - \bar{X})^2$
σ	$s = \sqrt{s^2}$	$S = \sqrt{S^2}$

12-20

12.3 Erwartungstreue

Bemerkung 12.20

- Ein Schätzer $\hat{\theta} = \hat{\theta}(X_1, \ldots, X_n)$, z. B.

$$\bar{X} = \frac{1}{n}\sum_{i=1}^{n} X_i,$$

  ist eine Zufallsvariable.
- Eine Theorie geeigneter Schätzer basiert daher auf Eigenschaften der Zufallsvariablen $\hat{\theta}$ (z. B. $\bar{X}$) bzw. deren Wahrscheinlichkeitsverteilung, nicht aber auf Eigenschaften einer einzelnen Realisation $\hat{\theta}(x_1, \ldots, x_n)$ (z. B. $\bar{x}$).
- Eine wünschenswerte Eigenschaft ist die Erwartungstreue.

12-21

Definition 12.21 (Erwartungstreue, Unverzerrtheit)
Ein Schätzer $\hat{\theta}$ heißt **erwartungstreu** oder **unverzerrt** (engl. *unbiased*) für den Parameter θ, wenn

$$\mathbb{E}[\hat{\theta}(X_1, \ldots, X_n)] = \theta$$

gilt.

Satz 12.22 (Erwartungstreue Schätzung von μ)
Die X_i $(i = 1, 2, \ldots, n)$ seien unabhängig und identisch verteilt und es existiere $\mu \stackrel{\text{def}}{=} \mathbb{E}[X_i]$. Dann ist $\bar{X}$ ein erwartungstreuer Schätzer für μ, d. h.

$$\mathbb{E}[\bar{X}] = \mu.$$

12-22

Satz 12.23 (Erwartungstreue Schätzung von σ^2)
Die X_i $(i = 1, 2, \ldots, n)$ seien unabhängig und identisch verteilt und es existiere $\sigma^2 \stackrel{\text{def}}{=} \mathbb{V}[X_i]$.

1. *Es gilt*
$$\mathbb{E}\left[S^2\right] = \frac{n-1}{n}\sigma^2,$$
d. h. die Stichprobenvarianz S^2 ist **kein** *erwartungstreuer Schätzer für σ^2.*

2. *Die korrigierte Stichprobenvarianz*
$$S^{*2} \stackrel{\text{def}}{=} \frac{1}{n-1}\sum_{i=1}^{n}(X_i - \bar{X})^2$$
ist ein erwartungstreuer Schätzer für σ^2,
$$\mathbb{E}[S^{*2}] = \sigma^2.$$

12-23

12.4 Wahrscheinlichkeitsverteilung eines Schätzers

Bemerkung 12.24

- Häufig lassen sich für einen Schätzer nicht nur der Erwartungswert und die Varianz bestimmen, sondern es kann die Wahrscheinlichkeitsverteilung des Schätzers explizit angegeben werden.
- Zwei Beispiele sind in den nächsten Sätzen angegeben:
  - Die Verteilung des Stichprobenmittels $\bar{X}$ bei normalverteilten und bei Bernoulli-verteilten Beobachtungen.
  - Die Verteilung der Stichprobenvarianz S^2 bei normalverteilten Beobachtungen.

12-24

Satz 12.25 (Verteilung des Stichprobenmittelwertes)

1. *Wenn die* X_i *unabhängig und identisch verteilt mit* $\mathbb{E}[X_i] = \mu$ *und* $\mathbb{V}[X_i] = \sigma^2$ *sind, dann gilt*

$$\mathbb{E}[\bar{X}] = \mu \quad \textit{und} \quad \mathbb{V}[\bar{X}] = \frac{\sigma^2}{n}\,.$$

2. *Für* $X_i \overset{i.i.d.}{\sim} N(\mu, \sigma^2)$ *gilt*

$$\bar{X} \sim N\left(\mu, \frac{\sigma^2}{n}\right) \qquad \textit{und} \qquad \frac{\bar{X} - \mu}{\sigma}\sqrt{n} \sim N(0,1).$$

3. *Für* $X_i \overset{i.i.d.}{\sim} B(1, \pi)$ *gilt*

$$\sum_{i=1}^{n} X_i = n\bar{X} \sim B(n, \pi).$$

12-25

Bemerkung 12.26
Interessiert man sich für die Verteilung der Schätzer S^2 und S^{*2} für die Varianz σ^2, so gelangt man für normalverteilte Beobachtungen zur sogenannten χ^2-Verteilung (Chiquadrat-Verteilung).

Definition 12.27 (χ^2-Verteilung)
Es sei $\nu \in \mathbb{N}$ und $U_1, \ldots, U_\nu \overset{i.i.d.}{\sim} N(0,1)$.

1. Die Verteilung der Zufallsvariablen

$$Q = \sum_{i=1}^{\nu} U_i^2$$

heißt χ^2**-Verteilung mit dem Parameter** ν. Schreibweise: $Q \sim \chi^2_\nu$.

2. Der Parameter ν heißt auch **Anzahl der Freiheitsgrade**.

3. Das p-Quantil einer χ^2-Verteilung mit dem Parameter ν wird mit $\chi^2_{\nu,p}$ bezeichnet.

12-26

Bemerkung 12.28 (Eigenschaften von $Q \sim \chi^2_\nu$)

- Q ist eine stetige Zufallsvariable mit $P(Q > 0) = 1$.
- Es gilt

$$\mathbb{E}[Q] = \nu \quad \text{und} \quad \mathbb{V}[Q] = 2\nu\,.$$

12-27

Satz 12.29 (Verteilung der Stichprobenvarianz S^2)
Für $X_i \overset{i.i.d.}{\sim} N(\mu, \sigma^2)$ gilt

$$\frac{nS^2}{\sigma^2} \sim \chi^2_{n-1}.$$

Bemerkung 12.30 (Verteilung von S^{*2})
Wegen

$$nS^2 = \sum_{i=1}^{n}(X_i - \bar{X})^2 = (n-1)S^{*2}$$

gilt auch

$$\frac{(n-1)S^{*2}}{\sigma^2} \sim \chi^2_{n-1}.$$

12.5 Ergänzungen

Bemerkung 12.a (Standardfehler) Die Standardabweichung des Stichprobenmittels $\bar{X}$,

$$\sigma[\bar{X}] = \sqrt{\mathbb{V}[\bar{X}]} = \frac{\sigma}{\sqrt{n}}\,,$$

heißt **Standardfehler** (engl. *standard error*) von $\bar{X}$. Häufig heißt auch der Schätzwert

$$\hat{\sigma}[\bar{X}] = \frac{\hat{\sigma}}{\sqrt{n}}$$

für $\sigma[\bar{X}]$ mit $\hat{\sigma} = s$ oder mit $\hat{\sigma} = s^*$ **Standardfehler der Mittelwertschätzung**. Allgemeiner wird die Standardabweichung oder die geschätzte Standardabweichung eines erwartungstreuen Schätzers $\hat{\theta}$ für einen Parameter θ als Standardfehler des Schätzers bezeichnet.

Beispiel 12.b (Erwartungstreue Schätzung von μ) Für X_i i.i.d. mit $\mathbb{E}[X_i] = \mu$ gilt

$$\mathbb{E}\left[\sum_{i=1}^{n} c_i X_i\right] = \sum_{i=1}^{n} c_i \mathbb{E}[X_i] = \sum_{i=1}^{n} c_i \mu = \mu \sum_{i=1}^{n} c_i$$

für $c_1, \ldots, c_n \in \mathbb{R}$. Die Stichprobenfunktion $\sum_{i=1}^{n} c_i X_i$ ist also genau dann ein erwartungstreuer Schätzer für den Parameter μ, falls $\sum_{i=1}^{n} c_i = 1$. Zum Beispiel ist $\bar{X}$ ein erwartungstreuer Schätzer für μ.

Bemerkung 12.c (Konstruktion von Schätzern) Es gibt verschiedene Methoden zur Konstruktion von Parameterschätzern. Die wichtigsten Verfahren sind

- die **Methode der kleinsten Quadrate**
- die **Maximum-Likelihood-Methode**,
- die **Momentenmethode** und
- **Bayesianische Schätzverfahren**.

Beispiel 12.d (Ziehungsschemata)

1. Ziehen *ohne* Zurücklegen (ZoZ) bedeutet beispielsweise die Entnahme von $n = 2$ Kugeln ohne Zurücklegen aus einer Urne mit $N = 3$ Kugeln, die als Kugeln A, B und C identifizierbar sind. Das Ergebnis der Stichprobenziehung ist ein **geordnetes Paar** (erste Ziehung, zweite Ziehung). Es gibt **sechs** mögliche gleichwahrscheinliche Stichprobenergebnisse:

$$(A, B), (A, C), (B, A), (B, C), (C, A), (C, B)\,.$$

Die Anzahl der möglichen Stichprobenergebnisse ist $N(N-1) = 3 * 2 = 6$. Allgemein ist die Anzahl der möglichen Stichproben beim Ziehen ohne Zurücklegen $N(N-1) \cdot \ldots \cdot (N-n+1)$.

2. Ziehen *mit* Zurücklegen (ZmZ) bedeutet z. B. die Entnahme von $n = 2$ Kugeln mit Zurücklegen aus einer Urne mit $N = 3$ Kugeln, die als Kugeln A, B und C identifizierbar sind. Ein Stichprobenergebnis ist ein **geordnetes Paar** (erste Ziehung, zweite Ziehung). Es gibt **neun** mögliche gleichwahrscheinliche Stichproben:
$$(A, A), (A, B), (A, C), (B, A), (B, B), (B, C), (C, A), (C, B), (C, C)$$
Die Anzahl der möglichen Stichproben ist $N^2 = 3^2 = 9$. Allgemein ist die Anzahl der möglichen Stichproben beim Ziehen mit Zurücklegen N^n.

3. Bei der Entnahme von $n = 2$ Kugeln ohne Zurücklegen und **ohne Berücksichtigung der Reihenfolge** aus einer Urne mit $N = 3$ Kugeln, die als Kugeln A, B und C identifizierbar sind. Ein Stichprobenergebnis ist eine **Menge** { erste Kugel, zweite Kugel }. Es gibt **drei** mögliche, gleichwahrscheinliche Stichprobenergebnisse
$$\{A, B\}, \{A, C\}, \{B, C\}\,.$$
Die Anzahl der möglichen Stichproben ist $\binom{3}{2} = 3$. Allgemein ist die Anzahl der möglichen Stichproben beim Ziehen ohne Zurücklegen ohne Berücksichtigung der Reihenfolge durch den Binomialkoeffizienten $\binom{N}{n}$ gegeben.

4. Das interessierende **Merkmal in der Grundgesamtheit** sei das Gewicht der Kugeln, z. B.
$$x_A = 100, \quad x_B = 100, \quad x_C = 130\,.$$
Das zufällige Ziehen **einer** Kugel und die Ermittlung des Gewichtes wird beschrieben durch eine Zufallsvariable X mit den beiden möglichen **Realisationen** 100 und 130 und mit der **diskreten Wahrscheinlichkeitsverteilung**
$$P(X = 100) = 2/3, \quad P(X = 130) = 1/3\,.$$
Die Merkmalsverteilung in der Grundgesamtheit mit den relativen Häufigkeiten 2/3 und 1/3 wird dann zur Wahrscheinlichkeitsverteilung bei zufälliger Entnahme einer Einheit. Der **Mittelwert** des Merkmals Gewicht in der Grundgesamtheit ist
$$\bar{x} = \frac{100 + 100 + 130}{3} = 110\,.$$
Der **Erwartungswert** der Zufallsvariablen X, die das zufällige Ergebnis beim Ziehen einer Kugel beschreibt, ist
$$\mathbb{E}[X] = 100 \cdot P(X = 100) + 130 \cdot P(X = 130) = 100 \cdot \frac{2}{3} + 130 \cdot \frac{1}{3} = 110.$$
Bei der Entnahme mehrerer Einheiten bezeichne X_1 bezeichne das zufällige Ergebnis der ersten Ziehung, X_2 das zufällige Ergebnis der zweiten Ziehung usw. Die Realisationen von (X_1, X_2) sind dann Paare (x_1, x_2) mit $x_i \in \{100, 130\}$ für $i = 1, 2$. Der Mittelwert einer Stichprobe vom Umfang $n = 2$,
$$\bar{X} \stackrel{\text{def}}{=} \frac{X_1 + X_2}{2}\,,$$
ist ein einfaches Beispiel einer **Stichprobenfunktion**. $\bar{X}$ ist eine Zufallsvariable mit den möglichen Realisationen 100, 115 und 130 beim Ziehen mit Zurücklegen und den möglichen Realisationen 100 und 115 beim Ziehen ohne Zurücklegen, während der Mittelwert der Grundgesamtheit den Wert 110 hat.

Bemerkung 12.e (Eigenschaften von $Q \sim \chi^2_\nu$) Die standardisierte Variable
$$\frac{Q - \nu}{\sqrt{2\nu}}$$
ist asymptotisch standardnormalverteilt, d. h.
$$\lim_{\nu \to \infty} P\left(\frac{Q - \nu}{\sqrt{2\nu}} \leq x\right) = \Phi(x), \quad x \in \mathbb{R},$$
wobei Φ die Verteilungsfunktion der Standardnormalverteilung ist. Eine Approximation für große ν ist daher
$$\chi^2_\nu \approx N(\nu, 2\nu)\,.$$

Kapitel 13

Konfidenzintervalle

13-1

13. Konfidenzintervalle

13.1 Konfidenzniveau und Konfidenzintervall
13.2 Konfidenzintervall für μ einer Normalverteilung
13.2.1 σ^2 bekannt
13.2.2 σ^2 unbekannt
13.3 Konfidenzintervall für σ^2 einer Normalverteilung
13.4 Konfidenzintervall für π einer Bernoulli-Verteilung

13-2

13.1 Konfidenzniveau und Konfidenzintervall

Bemerkung 13.1 (Punkt- und Intervallschätzung)

- Grundidee der **Punktschätzung:**

$$(x_1, \dots, x_n) \mapsto \hat{\theta}(x_1, \dots, x_n) \in \mathbb{R}$$

mit dem Ziel, dass der Schätzwert $\hat{\theta}(x_1, \dots, x_n)$ nahe bei dem Parameter θ liegt.

- Grundidee der **Intervallschätzung:**

$$(x_1, \dots, x_n) \mapsto [\hat{\theta}_u(x_1, \dots, x_n), \hat{\theta}_o(x_1, \dots, x_n)] \subset \mathbb{R}$$

mit dem Ziel, dass das Intervall den Parameter θ überdeckt,

$$\theta \in [\hat{\theta}_u(x_1, \dots, x_n), \hat{\theta}_o(x_1, \dots, x_n)]\,.$$

13-3

Definition 13.2 (Konfidenzintervall, Konfidenzniveau)
$\hat{\theta}_u = \hat{\theta}_u(X_1, \ldots, X_n)$ und $\hat{\theta}_o = \hat{\theta}_o(X_1, \ldots, X_n)$ seien Stichprobenfunktionen mit

$$\hat{\theta}_u \leq \hat{\theta}_o \,.$$

Außerdem gelte

$$P(\hat{\theta}_u \leq \theta \leq \hat{\theta}_o) = 1 - \alpha$$

für ein vorgegebenes α mit $0 < \alpha < 1$.

- Das zufällige Intervall $[\hat{\theta}_u, \hat{\theta}_o]$ heißt **Konfidenzintervall** (engl.: *confidence interval*) für θ zum **Konfidenzniveau** (engl.: *confidence level*) $1 - \alpha$.
- Eine Realisation $[\hat{\theta}_u(x_1, \ldots, x_n), \hat{\theta}_o(x_1, \ldots, x_n)]$ des zufälligen Intervalls $[\hat{\theta}_u, \hat{\theta}_o]$ heißt **Wert des Konfidenzintervalls**, Ergebnis der Intervallschätzung, Schätzintervall oder **konkretes Konfidenzintervall**.

13-4

Bemerkung 13.3

- $(1 - \alpha)$ heißt auch **Überdeckungswahrscheinlichkeit** (engl.: *coverage probability*), da
$$P([\hat{\theta}_u, \hat{\theta}_o] \ni \theta) = 1 - \alpha \,.$$
- α heißt auch **Irrtumswahrscheinlichkeit** (engl.: *error probability*). Es gilt
$$P(\theta \notin [\hat{\theta}_u, \hat{\theta}_o]) = \alpha \,.$$
- Für jeden Wert $[\hat{\theta}_u(x_1, \ldots, x_n), \hat{\theta}_o(x_1, \ldots, x_n)]$ des zufälligen Konfidenzintervalles ist die Aussage
$$\theta \in [\hat{\theta}_u(x_1, \ldots, x_n), \hat{\theta}_o(x_1, \ldots, x_n)]$$
entweder falsch oder richtig. Es kann daher **keine** Wahrscheinlichkeitsaussage für das konkrete Konfidenzintervall gemacht werden.

13-5

Bemerkung 13.4
Häufig verwendete α-Werte bei Konfidenzintervallen:

$$0.10 = 10\%, \quad 0.05 = 5\%, \quad 0.01 = 1\%, \quad 0.001 = 0.1\%$$

13-6

13.2 Konfidenzintervall für μ einer Normalverteilung

13.2.1 σ^2 bekannt

Satz 13.5 (Konfidenzintervall für μ)
Für $X_i \overset{i.i.d.}{\sim} N(\mu, \sigma^2)$ ist

$$\left[\bar{X} - u_{1-\frac{\alpha}{2}} \frac{\sigma}{\sqrt{n}}, \bar{X} + u_{1-\frac{\alpha}{2}} \frac{\sigma}{\sqrt{n}}\right]$$

ein Konfidenzintervall für μ zum Niveau $1-\alpha$.

Bemerkung 13.6
Mit einer gegebenen Realisation $\bar{x}$ von $\bar{X}$ ist

$$\left[\bar{x} - u_{1-\frac{\alpha}{2}} \frac{\sigma}{\sqrt{n}}, \bar{x} + u_{1-\frac{\alpha}{2}} \frac{\sigma}{\sqrt{n}}\right]$$

der **Wert des Konfidenzintervalls** für μ zum Niveau $1-\alpha$.

13-7

Bemerkung 13.7 (p-Quantil)

- u_p bezeichne das p-Quantil (p-Fraktil) für $U \sim N(0,1)$, d. h.

$$P(U \leq u_p) = p, \quad 0 < p < 1\,.$$

- Die Quantile (Fraktile) u_p von $U \sim N(0,1)$ sind in Tabellen zu finden.
- Es gilt

$$u_{\alpha/2} = -u_{1-\alpha/2},$$

  da die Dichtefunktion von U symmetrisch zu Null ist.

13-8

Bemerkung 13.8 (Zur Interpretation)

- Das zufällige Intervall

$$\left[\bar{X} - u_{1-\frac{\alpha}{2}} \frac{\sigma}{\sqrt{n}}, \bar{X} + u_{1-\frac{\alpha}{2}} \frac{\sigma}{\sqrt{n}}\right]$$

  ist ein Konfidenzintervall zum Konfidenzniveau $1-\alpha$ für den Parameter μ. Es enthält (überdeckt) mit der Wahrscheinlichkeit $1-\alpha$ den unbekannten Parameter μ.

- Das Intervall

$$\left[\bar{x} - u_{1-\frac{\alpha}{2}} \frac{\sigma}{\sqrt{n}}, \bar{x} + u_{1-\frac{\alpha}{2}} \frac{\sigma}{\sqrt{n}}\right]$$

  ist ein Wert des Konfidenzintervalls, der aus dem Wert $(x_1, \ldots, x_n)$ der (zufälligen) Stichprobe $(X_1, \ldots, X_n)$ berechnet ist.

13-9

Bemerkung 13.9 (Theoretischer Hintergrund)
Wegen $X_i \stackrel{i.i.d.}{\sim} N(\mu, \sigma^2)$ ist die sogenannte **Gauß-Statistik**

$$\frac{\bar{X}-\mu}{\sigma}\sqrt{n}$$

standardnormalverteilt, vgl. Satz 12.25. Daher gilt

$$\begin{aligned}
1-\alpha &= P\left(u_{\alpha/2} \le \frac{\bar{X}-\mu}{\sigma}\sqrt{n} \le u_{1-\alpha/2}\right) \\
&= P\left(-u_{1-\alpha/2} \le \frac{\bar{X}-\mu}{\sigma}\sqrt{n} \le u_{1-\alpha/2}\right) \\
&= P\left(-\bar{X} - u_{1-\alpha/2}\frac{\sigma}{\sqrt{n}} \le -\mu \le -\bar{X} + u_{1-\alpha/2}\frac{\sigma}{\sqrt{n}}\right) \\
&= P\left(\bar{X} - u_{1-\alpha/2}\frac{\sigma}{\sqrt{n}} \le \mu \le \bar{X} + u_{1-\alpha/2}\frac{\sigma}{\sqrt{n}}\right) \\
&= P\left(\mu \in \left[\bar{X} - u_{1-\alpha/2}\frac{\sigma}{\sqrt{n}},\ \bar{X} + u_{1-\alpha/2}\frac{\sigma}{\sqrt{n}}\right]\right).
\end{aligned}$$

13.2.2 σ^2 unbekannt

13-10

Bemerkung 13.10

1. Das in Satz 13.5 angegebene Konfidenzintervall für μ kann nicht verwendet werden, wenn σ unbekannt ist.
2. Der unbekannte Parameter σ wird durch S geschätzt und anstelle des Quantils $u_{1-\alpha/2}$ einer Standardnormalverteilung wird ein Quantil aus einer sogenannten t-Verteilung verwendet.

13-11

Satz 13.11 (Konfidenzintervall für μ)
Für $X_i \stackrel{i.i.d.}{\sim} N(\mu, \sigma^2)$ ist

$$\left[\bar{X} - t_{n-1,1-\frac{\alpha}{2}}\frac{S}{\sqrt{n-1}}, \bar{X} + t_{n-1,1-\frac{\alpha}{2}}\frac{S}{\sqrt{n-1}}\right]$$

ein Konfidenzintervall für μ zum Niveau $1-\alpha$.

Bemerkung 13.12

- $t_{n-1,1-\frac{\alpha}{2}}$ bezeichnet das $1-\frac{\alpha}{2}$-Quantil einer t-Verteilung mit dem Parameter $n-1$.
- Es gilt
$$\frac{S}{\sqrt{n-1}} = \frac{S^*}{\sqrt{n}}.$$

13-12

Definition 13.13 (t-Verteilung)
Die Zufallsvariablen $U \sim N(0,1)$ und $Q \sim \chi^2_\nu$ für $\nu \in \mathbb{N}$ seien stochastisch unabhängig.

1. Dann heißt die Verteilung von
$$W \stackrel{\text{def}}{=} \frac{U}{\sqrt{\frac{Q}{\nu}}}$$
t-Verteilung mit Parameter ν. Der Parameter ν heißt auch **Anzahl der Freiheitsgrade**.
2. Schreibweise: $W \sim t_\nu$.
3. $t_{\nu,p}$ bezeichnet das p-Quantil von t_ν.

Bemerkung 13.14
Bei Anwendungen ist häufig $\nu = n-1$, wobei n die Anzahl der Beobachtungen ist, oder allgemeiner $\nu = n - l$, wobei n die Anzahl der Beobachtungen und l die Anzahl geschätzter Parameter ist.

13-13

Bemerkung 13.15 (Eigenschaften von $W \sim t_\nu$)

- $\mathbb{E}[W]$ existiert nicht für $\nu = 1$. $\mathbb{E}[W] = 0$ für $\nu \geq 2$
- $\mathbb{V}[W]$ existiert nicht für $\nu = 1, 2$. $\mathbb{V}[W] = \dfrac{\nu}{\nu - 2}$ für $\nu \geq 3$
- Die Dichtefunktion von W ist symmetrisch zu Null, daher gilt
$$t_{\nu,1-p} = -t_{\nu,p}\,.$$
- W ist asymptotisch standardnormalverteilt, d. h.
$$\lim_{\nu\to\infty} P(W \leq x) = \Phi(x), \quad x \in \mathbb{R},$$
wobei Φ die Verteilungsfunktion der Standardnormalverteilung ist.
- Approximation für große ν (z. B. für $\nu \geq 30$):
$$t_{\nu,p} \approx u_p\,.$$

13-14

13.3 Konfidenzintervall für σ^2 einer Normalverteilung

Satz 13.16 (Konfidenzintervall für σ^2)
Unter der Voraussetzung $X_i \stackrel{i.i.d.}{\sim} N(\mu, \sigma^2)$ ist
$$\left[\frac{nS^2}{\chi^2_{n-1,1-\frac{\alpha}{2}}}, \frac{nS^2}{\chi^2_{n-1,\frac{\alpha}{2}}}\right]$$
ein Konfidenzintervall für σ^2 zum Niveau $1-\alpha$.

Bemerkung 13.17
Ein Wert des Konfidenzintervalls für σ^2 ist
$$\left[\frac{ns^2}{\chi^2_{n-1,1-\frac{\alpha}{2}}}, \frac{ns^2}{\chi^2_{n-1,\frac{\alpha}{2}}}\right].$$

13-15

Bemerkung 13.18 (Symmetrie)

- Der Wert des Konfidenzintervalls für σ^2 ist **asymmetrisch** in dem Sinn, dass s^2 nicht in der Intervallmitte liegt.
- Symmetrisch ist die Aufteilung von α. Man spricht auch von **zentralen** Konfidenzintervallen.

Satz 13.19 (Konfidenzintervall für σ)
Unter der Voraussetzung $X_i \overset{i.i.d.}{\sim} N(\mu, \sigma^2)$ *ist*

$$\left[\sqrt{\frac{nS^2}{\chi^2_{n-1,1-\frac{\alpha}{2}}}}, \sqrt{\frac{nS^2}{\chi^2_{n-1,\frac{\alpha}{2}}}}\right]$$

ein Konfidenzintervall für σ *zum Niveau* $1-\alpha$.

13-16

13.4 Konfidenzintervall für π einer Bernoulli-Verteilung

Bemerkung 13.20 (Konfidenzintervall für π)

- Die X_i seien unabhängig und identisch Bernoulli-verteilt, $X_i \overset{i.i.d.}{\sim} B(1, \pi)$ mit $0 < \pi < 1$. Dann gilt $\mu \overset{\text{def}}{=} \mathbb{E}[X_i] = \pi$ und $\sigma^2 \overset{\text{def}}{=} \mathbb{V}[X_i] = \pi(1-\pi)$.
- Basierend auf dem Punktschätzer $\hat{\pi} = \hat{\mu} = \bar{X}$ ist

$$\left[\hat{\pi} - u_{1-\frac{\alpha}{2}}\sqrt{\frac{\hat{\pi}(1-\hat{\pi})}{n}}, \hat{\pi} + u_{1-\frac{\alpha}{2}}\sqrt{\frac{\hat{\pi}(1-\hat{\pi})}{n}}\right].$$

  ein **approximatives Konfidenzintervall** für π mit

$$P\left(\pi \in \left[\hat{\pi} - u_{1-\frac{\alpha}{2}}\sqrt{\frac{\hat{\pi}(1-\hat{\pi})}{n}}, \hat{\pi} + u_{1-\frac{\alpha}{2}}\sqrt{\frac{\hat{\pi}(1-\hat{\pi})}{n}}\right]\right) \approx 1-\alpha\,.$$

- Faustregel für die Anwendung: $n\hat{\pi}(1-\hat{\pi}) > 9$.

13.5 Ergänzungen

Bemerkung 13.a (Theoretischer Hintergrund) Die Konstruktion des Konfidenzintervalls aus Satz 13.11 beruht auf der stochastischen Unabhängigkeit der Variablen $\bar{X}$ und S^2 und der Verteilung der sogenannten t-Statistik.

Satz 13.b *Unter der Voraussetzung* $X_i \overset{i.i.d.}{\sim} N(\mu, \sigma^2)$ *sind die beiden Zufallsvariablen*

$$\frac{\bar{X}-\mu}{\sigma}\sqrt{n} \sim N(0,1) \quad \textit{und} \quad \frac{nS^2}{\sigma^2} \sim \chi^2_{n-1}$$

stochastisch unabhängig und es gilt

$$T \overset{\text{def}}{=} \frac{\bar{X}-\mu}{S}\sqrt{n-1} \sim t_{n-1}.$$

Bemerkung 13.c T ist die sogenannte **t-Statistik**. Es gilt

$$\frac{\frac{\bar{X}-\mu}{\sigma}\sqrt{n}}{\sqrt{\frac{\frac{nS^2}{\sigma^2}}{n-1}}} = \frac{\bar{X}-\mu}{S}\sqrt{n-1} = \frac{\bar{X}-\mu}{S^*}\sqrt{n} = T\,.$$

Bemerkung 13.d (Zu Satz 13.16)
Aus Satz 12.29 folgt:

$$\begin{aligned} 1-\alpha &= P\left(\chi^2_{n-1,\frac{\alpha}{2}} \le \frac{nS^2}{\sigma^2} \le \chi^2_{n-1,1-\frac{\alpha}{2}}\right) = P\left(\chi^2_{n-1,\frac{\alpha}{2}}\sigma^2 \le nS^2 \le \chi^2_{n-1,1-\frac{\alpha}{2}}\sigma^2\right) \\ &= P\left(\frac{nS^2}{\chi^2_{n-1,1-\frac{\alpha}{2}}} \le \sigma^2 \le \frac{nS^2}{\chi^2_{n-1,\frac{\alpha}{2}}}\right) = P\left(\sigma^2 \in \left[\frac{nS^2}{\chi^2_{n-1,1-\frac{\alpha}{2}}}, \frac{nS^2}{\chi^2_{n-1,\frac{\alpha}{2}}}\right]\right) \end{aligned}$$

Beispiel 13.e (DAX-Rendite)

1. DAX-Kurs am 30.12.1987: 1000, am 29.12.1995: 2254, am 30.12.1998: 5002.
2. Das stochastische Standardmodell für die Kurse K_t und die Tagesrenditen $X_t = \ln(K_t/K_{t-1})$ ist
$$K_t = K_{t-1}e^{X_t}, \quad X_t \text{ i.i.d. } N(\mu, \sigma^2)\,.$$
3. Die durchschnittliche Tagesrendite $\bar{x}$ aus den drei Jahren 1996 bis 1998 ergibt sich bei 250 Handelstagen pro Jahr mit dem Ansatz
$$5002 = 2254 * \exp\left(\sum_{t=1}^{750} x_t\right) = 2254e^{750\bar{x}}$$
als
$$\bar{x} = \frac{\ln\left(\frac{5002}{2254}\right)}{750} = 0.0011 = 0.11\%\,.$$
4. Mit der Annahme für die Tagesvolatilität (Standardabweichung der X_t)
$$\sigma = 1.5\% = 0.015$$
ergibt sich ein Wert des Konfidenzintervalls für die erwartete Tagesrendite $\mu = \mathbb{E}[X_i]$ zu vorgegebenem $\alpha = 0.05$ als
$$\begin{aligned} \left[0.0011 - 1.96\frac{0.015}{\sqrt{750}}, 0.0011 + 1.96\frac{0.015}{\sqrt{750}}\right] &= [-0.00003, 0.00217] \\ &= [-0.003\%, 0.217\%]\,. \end{aligned}$$
Dies entspricht dem Konfidenzintervall für die erwartete Jahresrendite (Faktor 250)
$$[-0.0075, 0.5425] = [-0.75\%, 54.25\%]\,.$$
5. Bei Berücksichtigung von weiteren 7 Jahren (DAX-Kurs am 30. 12. 2005: 5408) führt der Ansatz
$$5002e^{1750\bar{x}} = 5408$$
zu einer **durchschnittlichen Tagesrendite** von
$$\bar{x} = \frac{\ln\left(\frac{5408}{5002}\right)}{1750} = 0.0000446 = 0.00446\%$$
und durch Multiplikation mit 250 zu einer **durchschnittlichen Jahresrendite** von 1.11% für die Jahre 1999 bis 2005.
6. Die im Beispiel unterstellte **Tagesvolatilität** von 1.5% entspricht einer **Jahresvolatilität** (250 Handelstage) von etwa 24%. Zwischen der Jahresvolatilität $\sigma_{(250)}$ und der **Tagesvolatilität** σ besteht der Zusammenhang
$$\sigma = \frac{\sigma_{(250)}}{\sqrt{250}},$$
da
$$\sigma^2_{(250)} = \mathbb{V}\left[\sum_{t=1}^{250} X_t\right] = \sum_{t=1}^{250} \mathbb{V}[X_t] = 250\sigma^2.$$
Bei einer Jahresvolatilität (250 Handelstage) von
$$\sigma_{(250)} = 24\% = 0.24$$
ergibt sich eine Tagesvolatilität von
$$\sigma = \frac{\sigma_{(250)}}{\sqrt{250}} = 0.01518 = 1.518\%.$$

Bemerkung 13.f (Approximatives Konfidenzintervall für π) Für Bernoulli-verteilte Zufallsvariablen gilt

$$S^2 = \frac{1}{n}\sum_{i=1}^{n}(X_i - \bar{X})^2 = \bar{X}(1-\bar{X}) = \hat{\pi}(1-\hat{\pi}),$$

so dass sich das approximative Konfidenzintervall für π auch als

$$\left[\hat{\pi} - u_{1-\frac{\alpha}{2}}\frac{S}{\sqrt{n}}, \hat{\pi} + u_{1-\frac{\alpha}{2}}\frac{S}{\sqrt{n}}\right]$$

schreiben lässt.

Bemerkung 13.g (Approximative Konfidenzintervalle für μ) Die X_i seien unabhängig und identisch verteilt (i.i.d.) mit $\mu \overset{\text{def}}{=} \mathbb{E}[X_i]$ und $0 < \sigma^2 \overset{\text{def}}{=} \mathbb{V}[X_i] < \infty$.

- Wenn σ^2 **bekannt** ist, dann ist

  $$\left[\bar{X} - u_{1-\frac{\alpha}{2}}\frac{\sigma}{\sqrt{n}}, \bar{X} + u_{1-\frac{\alpha}{2}}\frac{\sigma}{\sqrt{n}}\right]$$

  ein **approximatives Konfidenzintervall** für μ zum Niveau $1-\alpha$. Es gilt

  $$P\left(\bar{X} - u_{1-\frac{\alpha}{2}}\frac{\sigma}{\sqrt{n}} \leq \mu \leq \bar{X} + u_{1-\frac{\alpha}{2}}\frac{\sigma}{\sqrt{n}}\right) \approx 1-\alpha$$

  und die Approximation ist umso besser, je größer n ist. Eine Faustregel für die Verwendung ist $n \geq 40$.

- Wenn σ^2 **unbekannt** ist, dann ist

  $$\left[\bar{X} - u_{1-\frac{\alpha}{2}}\frac{S}{\sqrt{n}}, \bar{X} + u_{1-\frac{\alpha}{2}}\frac{S}{\sqrt{n}}\right]$$

  mit

  $$S \overset{\text{def}}{=} \sqrt{\frac{1}{n}\sum_{i=1}^{n}(X_i - \bar{X})^2}$$

  ein **approximatives Konfidenzintervall** für μ zum Niveau α. Der Stichprobenumfang n muss hinreichend groß sein, Faustregel: $n \geq 40$.

Kapitel 14

Grundstruktur statistischer Tests

14-1

14. Grundstruktur statistischer Tests

14.1 Schätz- und Testproblem
14.2 Beispiel zum Hypothesentest
14.3 Allgemeine Teststruktur

14-2

14.1 Schätz- und Testproblem

Bemerkung 14.1 (Schätzproblem)

- Der unbekannte Parameter θ der Grundgesamtheit soll geschätzt werden.
- Ein **Schätzwert** $\hat{\theta}(x_1, \ldots, x_n)$ für θ ist die Realisation einer geeigneten Statistik (**= Schätzfunktion**) $\hat{\theta} = \hat{\theta}(X_1, \ldots, X_n)$.
- Beispiel: $\hat{\mu} = \bar{X}$ ist eine Schätzfunktion für den Erwartungswert μ und $\bar{x}$ ist ein Schätzwert für μ.

14-3

Bemerkung 14.2 (Testproblem)

- Eine **Hypothese** über die Grundgesamtheit (z. B. $\mu > 0$) soll überprüft werden.
- Zu einer geeigneten Statistik (**Prüfgröße** oder **Testfunktion**)

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

erhält man eine konkrete Realisation

$$t = T(x_1, \ldots, x_n),$$

mit deren Hilfe man über die Hypothese entscheidet.
- Beispiel: Realisationen von $\bar{X}$, die "weit" von 1000 entfernt sind, sprechen gegen die Hypothese $\mu = 1000$.
- Aber "weit" hängt vom Stichprobenumfang n und von der Varianz σ^2 ab!

14-4

14.2 Beispiel zum Hypothesentest

Beispiel 14.3 (Abfüllanlage)

- Betrachtet wird eine Abfüllanlage mit dem Sollgewicht $\mu_0 = 1000[g]$ und **bekannter** Standardabweichung der Abfüllmenge $\sigma = 3[g]$.
- Beobachtet wird die **Stichprobe**

$$x_1 = 1002, x_2 = 1007, x_3 = 997, x_4 = 1014$$

vom Umfang $n = 4$ mit dem Mittelwert

$$\bar{x} = \frac{1}{4} \sum_{i=1}^{4} x_i = 1005.$$

- Die interessierende **Fragestellung** ist:
Wird die Hypothese $\mu = \mu_0$ (hier $\mu = 1000$) durch die Beobachtung von $\bar{x} = 1005$ erschüttert? Weicht also $\bar{x}$ statistisch signifikant von μ_0 ab?

14-5

- Das **stochastische Modell** ist die stochastische Unabhängigkeit und Normalverteilung der Abfüllungen:

$$X_i \overset{i.i.d.}{\sim} N(\mu, \sigma^2)$$

- Für den **zufälligen Mittelwert** gilt daher

$$\bar{X} \sim N\left(\mu, \frac{\sigma^2}{n}\right).$$

- Die **Prüfgröße** ist

$$T \overset{\text{def}}{=} \frac{\bar{X} - \mu_0}{\sigma}\sqrt{n}\,.$$

14-6

- Falls $\mu = \mu_0$ (Hypothese richtig), gilt:

$$T = \frac{\bar{X} - \mu_0}{\sigma}\sqrt{n} \sim N(0,1).$$

- Falls $\mu > \mu_0$, gilt: $\bar{X}$ liegt in der Nähe von μ, **große** Werte von T sind wahrscheinlicher als im Fall $\mu = \mu_0$.

- Falls $\mu < \mu_0$, gilt: $\bar{X}$ liegt in der Nähe von μ, **kleine** Werte von T sind wahrscheinlicher als im Fall $\mu = \mu_0$.

- Für vorgegebenes kleines $\alpha > 0$ gilt

$$P\left(T \in]-\infty, -u_{1-\frac{\alpha}{2}}[\cup]u_{1-\frac{\alpha}{2}}, \infty[\right) = P(|T| > u_{1-\frac{\alpha}{2}}) = \alpha,$$

falls $\mu = \mu_0$.

14-7

- Beispiele für Quantile der Normalverteilung:

α	$u_{\frac{\alpha}{2}} = -u_{1-\frac{\alpha}{2}}$	$u_{1-\frac{\alpha}{2}}$
0.05	-1.96	1.96
0.01	-2.58	2.58

- $\bar{x} = 1005, \mu_0 = 1000, \sigma = 3, n = 4$ ergibt

$$t = \frac{1005 - 1000}{3} 2 = \frac{10}{3}.$$

- Da $t > 2.58$, wird die Hypothese $\mu = 1000$ auf dem Signifikanzniveau $\alpha = 0.01$ verworfen.

- t ist ein realisierter, beobachteter Wert der Zufallsvariablen T.

14-8

14.3 Allgemeine Teststruktur

Bemerkung 14.4 (Komponenten eines Tests)

- **Nullhypothese** (engl. *null hypothesis*) H_0,
  z. B. $H_0 : \mu = \mu_0$, wobei $\mu_0 \in \mathbb{R}$ vorgegeben.

- **Gegenhypothese** H_1, z. B. $H_1 : \mu \neq \mu_0$.

- Vorgegebenes **Signifikanzniveau** (engl.: *level of significance, significance level*) α

- **Prüfgröße** T

- **Kritischer Bereich** (engl.: *critical region*) K

- **Testentscheidung**:
  Es wird ein realisierter Wert t der Zufallsvariablen T beobachtet.
  H_0 wird zugunsten von H_1 abgelehnt, falls $t \in K$.
  H_0 kann nicht zugunsten von H_1 abgelehnt werden, falls $t \notin K$.

14-9

Bemerkung 14.5 (Alternative Bezeichnungen)

- Die Gegenhypothese heißt auch **Alternativhypothese**.
- Die Prüfgröße wird auch als **Teststatistik** (engl.: *test statistic*), Prüfvariable, Testfunktion, Testvariable oder Testgröße bezeichnet.
- Der kritische Bereich heißt auch **Ablehnbereich** (engl.: *region of rejection*) oder **Verwerfungsbereich**.

Bemerkung 14.6

- Die Prüfgröße $T = T(X_1, \ldots, X_n)$ ist eine **Zufallsvariable**, die von der Stichprobe abhängt.
- Der berechnete Wert $t(x_1, \ldots, x_n)$ der Prüfgröße ist eine **Realisation** der Zufallsvariablen T mit $t \in K$ oder $t \notin K$.

14-10

Bemerkung 14.7 (Signifikanz und Relevanz)

- Die statistische Signifikanz darf nicht verwechselt werden mit inhaltlicher Relevanz für eine bestimmte Fragestellung.
- Sehr kleine Abweichungen von der Nullhypothese können zwar statistisch signifikant sein, aber eventuell irrelevant für eine Anwendung.

Kapitel 15

Drei Tests für die Normalverteilung

15-1

15. Drei Tests für die Normalverteilung

15.1 Der Gauß-Test
15.2 Der t-Test
15.3 Test für eine Varianz

15-2

Bemerkung 15.1

- Für alle Tests in diesem Kapitel gilt die **Normalverteilungsvoraussetzung**:
$$X_i \overset{i.i.d.}{\sim} N(\mu, \sigma^2).$$
- Der Gauß- und der t-Test sind Tests über den Parameter μ.
- Der **Gauß-Test** setzt voraus, dass σ^2 bekannt ist.
- Beim **t-Test** ist σ^2 unbekannt und wird aus den Beobachtungen geschätzt.
- Der **Varianz-Test** ist ein Test über den Parameter σ^2 der Normalverteilung.

15-3

15.1 Der Gauß-Test

Bemerkung 15.2 (Schematische Darstellung)

- **Zweck:** Test über den Parameter μ einer Normalverteilung bei bekannter Varianz $\sigma^2 > 0$.
- **Voraussetzung:** $X_i \overset{i.i.d.}{\sim} N(\mu, \sigma^2)$
- **Gegeben:** Signifikanzniveau $\alpha \in (0,1)$, $\quad \sigma^2 > 0$, $\quad \mu_0 \in \mathbb{R}$,
- **Null- und Gegenhypothese:** $H_0 : \mu = \mu_0$, $\quad H_1 : \mu \neq \mu_0$
- **Prüfgröße:**

$$T = \frac{\bar{X} - \mu_0}{\sigma}\sqrt{n}$$

- **Kritischer Bereich:**

$$K = \left]-\infty, -u_{1-\frac{\alpha}{2}}\right[\, \cup \, \left]u_{1-\frac{\alpha}{2}}, \infty\right[$$

15-4

Bemerkung 15.3 (Kritischer Wert)

- In der Regel besteht der kritische Bereich aus einem Intervall oder der Vereinigung von zwei Intervallen. Die Intervallgrenzen heißen auch **kritische Werte**.
- Die kritischen Werte $u_{\frac{\alpha}{2}} = -u_{1-\frac{\alpha}{2}}$ und $u_{1-\frac{\alpha}{2}}$ für vorgegebenes α sind Quantile (Fraktile) der Testverteilung. $u_{\frac{\alpha}{2}} = -u_{1-\frac{\alpha}{2}}$ ergibt sich aus der Symmetrie der Dichtefunktion der Standardnormalverteilung.

Bemerkung 15.4 (Testverteilung)

- Die **Testverteilung** ist die Verteilung der Prüfgröße T, falls $\mu = \mu_0$, d. h. falls H_0 richtig ist. Dann gilt beim Gauß-Test

$$T \sim N(0,1)\,.$$

- Sehr große und sehr kleine Werte von T sind unter H_1 wahrscheinlicher als unter H_0.

15-5

Bemerkung 15.5 (Ein- und zweiseitiger Gauß-Test)

H_0	H_1	K
$\mu = \mu_0$	$\mu \neq \mu_0$	$\left]-\infty, -u_{1-\frac{\alpha}{2}}\right[\, \cup \, \left]u_{1-\frac{\alpha}{2}}, \infty\right[$
$\mu = \mu_0$ oder $\mu \geq \mu_0$	$\mu < \mu_0$	$\left]-\infty, -u_{1-\alpha}\right[$
$\mu = \mu_0$ oder $\mu \leq \mu_0$	$\mu > \mu_0$	$\left]u_{1-\alpha}, \infty\right[$

15-6

Bemerkung 15.6 (Zusammengesetzte Nullhypothesen)

- $H_0 : \mu = \mu_0$ ist ein Beispiel einer **einfachen Nullhypothese** (engl.: *simple null hypothesis*). Es gilt

$$P(T \in K) = \alpha,$$

falls H_0 richtig ist.

- Bei einer **zusammengesetzten Nullhypothese** (engl.: *composite null hypothesis*) $\mu \geq \mu_0$ (bzw. $\mu \leq \mu_0$) gilt

$$P(T \in K) \leq \alpha,$$

falls H_0 richtig ist.

15-7

Beispiel 15.7 (DAX-Jahresrendite I)
Die durchschnittliche Jahresrendite des DAX in einem Zeitraum von 25 Jahren sei $\bar{x} = 10\%$ bei einer Volatilität (p. a.) von $\sigma = 25\%$. Ist $\bar{x}$ signifikant ($\alpha = 5\%$) von Null verschieden?

- Standardmodell für Renditen: $X_t \overset{i.i.d.}{\sim} N(\mu, \sigma^2)$
- Hypothesen: $H_0 : \mu = 0, \quad H_1 : \mu \neq 0$
- Der berechnete Wert der Prüfgröße ist

$$t = \frac{\bar{x}}{\sigma}\sqrt{n} = \frac{0.10}{0.25}\sqrt{25} = 2\,.$$

- Der kritische Wert ist $u_{1-\alpha/2} = 1.96$ und der kritische Bereich ist

$$K = \,]-\infty, -1.96[\, \cup \,]1.96, \infty[\;\;.$$

- Testentscheidung: H_0 wird verworfen, da $t \in K$.

15-8

Beispiel 15.8 (DAX-Jahresrendite II)
Die durchschnittliche Rendite (p. a.) des DAX in einem Zeitraum von 25 Jahren sei $\bar{x} = 10\%$ bei einer Volatilität (p. a.) von $\sigma = 25\%$. Ist $\bar{x}$ signifikant ($\alpha = 5\%$) größer als $\mu_0 = 4\%$?

- Modell für die Renditen: $X_t \overset{i.i.d.}{\sim} N(\mu, \sigma^2)$
- Hypothesen: $H_0 : \mu \leq \mu_0, \quad H_1 : \mu > \mu_0$
- Der berechnete Wert der Prüfgröße ist

$$t = \frac{\bar{x} - \mu_0}{\sigma}\sqrt{n} = \frac{0.06}{0.25}\sqrt{25} = 1.2\,.$$

- Der kritische Wert ist $u_{1-\alpha} = 1.645$.
- Kritischer Bereich

$$K = \,]1.645, \infty[$$

- Testentscheidung: H_0 wird **nicht** verworfen, da $t \notin K$.

15-9

15.2 Der t-Test

Bemerkung 15.9 (Schematische Darstellung)

- **Zweck:** Test über den Parameter μ einer Normalverteilung bei unbekannter Varianz σ^2.
- **Voraussetzung:** $X_i \overset{i.i.d.}{\sim} N(\mu, \sigma^2)$
- **Gegeben:** Signifikanzniveau $\alpha \in (0,1)$, Schätzer S^* oder S für σ, $\mu_0 \in \mathbb{R}$
- **Nullhypothese und Gegenhypothese:** $H_0 : \mu = \mu_0, \quad H_1 : \mu \neq \mu_0$
- **Prüfgröße:**
$$T = \frac{\bar{X} - \mu_0}{S^*}\sqrt{n} = \frac{\bar{X} - \mu_0}{S}\sqrt{n-1}$$
- **Kritischer Bereich:**
$$K = \left]-\infty, -t_{n-1,1-\frac{\alpha}{2}}\right[\cup \left]t_{n-1,1-\frac{\alpha}{2}}, \infty\right[$$

15-10

Bemerkung 15.10 (Ein- und zweiseitige t-Tests)

H_0	H_1	K
$\mu = \mu_0$	$\mu \neq \mu_0$	$\left]-\infty, -t_{n-1,1-\frac{\alpha}{2}}\right[\cup \left]t_{n-1,1-\frac{\alpha}{2}}, \infty\right[$
$\mu = \mu_0$ oder $\mu \geq \mu_0$	$\mu < \mu_0$	$\left]-\infty, -t_{n-1,1-\alpha}\right[$
$\mu = \mu_0$ oder $\mu \leq \mu_0$	$\mu > \mu_0$	$\left]t_{n-1,1-\alpha}, \infty\right[$

Bemerkung 15.11 (Approximation der t-Verteilung)
Approximation für große n:
$t_{n-1,p} \approx u_p$, falls n groß (z. B. für $n \geq 30$)

15-11

Bemerkung 15.12 (Theoretischer Hintergrund)
Die sogenannte **t-Statistik**
$$\frac{\bar{X} - \mu_0}{S^*}\sqrt{n}$$
hat eine t-Verteilung mit $n-1$ Freiheitsgraden, falls $\mu = \mu_0$.

Bemerkung 15.13 (Normalverteilungsvoraussetzung)

- Eine **zwingende** Voraussetzung für die Anwendung des **t-Tests** ist die Normalverteilung der Grundgesamtheit.
- Wenn die Normalverteilungsvoraussetzung verletzt ist, kommt für große n eine approximative Variante des **Gauß-Tests** zum Einsatz.

15-12

15.3 Test für eine Varianz

Bemerkung 15.14 (Schematische Darstellung)

- **Zweck:** Test über den Parameter σ^2 einer Normalverteilung.
- **Voraussetzung:** $X_i \overset{i.i.d.}{\sim} N(\mu, \sigma^2)$
- **Gegeben:** Signifikanzniveau α, $\sigma_0^2 > 0$
- **Nullhypothese, Gegenhypothese:** $H_0 : \sigma^2 = \sigma_0^2, \quad H_1 : \sigma^2 \neq \sigma_0^2$
- **Prüfgröße:**

$$T = \frac{(n-1)S^{*2}}{\sigma_0^2} = \frac{nS^2}{\sigma_0^2}$$

- **Kritischer Bereich:**

$$K = \left[0, \chi^2_{n-1,\frac{\alpha}{2}}\right[\,\cup\, \left]\chi^2_{n-1,1-\frac{\alpha}{2}}, \infty\right[$$

15-13

Bemerkung 15.15 (Ein- und zweiseitige Varianz-Tests)

H_0	H_1	K
$\sigma^2 = \sigma_0^2$	$\sigma^2 \neq \sigma_0^2$	$\left[0, \chi^2_{n-1,\frac{\alpha}{2}}\right[\,\cup\, \left]\chi^2_{n-1,1-\frac{\alpha}{2}}, \infty\right[$
$\sigma^2 = \sigma_0^2$ oder $\sigma^2 \geq \sigma_0^2$	$\sigma^2 < \sigma_0^2$	$\left[0, \chi^2_{n-1,\alpha}\right[$
$\sigma^2 = \sigma_0^2$ oder $\sigma^2 \leq \sigma_0^2$	$\sigma^2 > \sigma_0^2$	$\left]\chi^2_{n-1,1-\alpha}, \infty\right[$

Bemerkung 15.16 (Theoretischer Hintergrund)

- Wenn $\sigma^2 = \sigma_0^2$ richtig ist, gilt $T \sim \chi^2_{n-1}$, vgl. Satz 12.29.
- Falls $\sigma^2 < \sigma_0^2$ liegt die Verteilung von T näher bei Null.
- Falls $\sigma^2 > \sigma_0^2$ liegt die Verteilung von T näher bei ∞.

Kapitel 16

Fehlerwahrscheinlichkeit, Signifikanzniveau und p-Wert

16-1

16. Fehlerwahrscheinlichkeit, Signifikanzniveau und p-Wert

16.1 Fehler 1. und 2. Art
16.2 Signifikanzniveau
16.3 p-Wert
16.4 Tests und Statistiksoftware

16-2

16.1 Fehler 1. und 2. Art

Definition 16.1 (Fehler 1. und 2. Art)

- Der **Fehler 1. Art** (engl.: *error of first kind, error of first type, type I error*) wird gemacht, falls H_0 abgelehnt wird, obwohl H_0 richtig ist.
- Der **Fehler 2. Art** (engl.: *error of second kind, error of second type, type II error*) wird gemacht, falls H_0 nicht abgelehnt wird, obwohl H_0 falsch ist.
- Die Wahrscheinlichkeit, den Fehler 1. Art zu begehen, heißt **Fehlerwahrscheinlichkeit 1. Art**.
- Die Wahrscheinlichkeit, den Fehler 2. Art zu begehen, heißt **Fehlerwahrscheinlichkeit 2. Art**.

16-3

16.2 Signifikanzniveau

Bemerkung 16.2 (α und Fehlerwahrscheinlichkeit 1. Art)

- Das vorgegebene Signifikanzniveau α **beschränkt** die Fehlerwahrscheinlichkeit 1. Art nach oben, es gilt also

  $$\text{„Fehlerwahrscheinlichkeit 1. Art"} \leq \alpha.$$

- Beim Gauß-Test mit der **einfachen** Nullhypothese $H_0 : \mu = \mu_0$ gilt

  $$P(T \in K || \mu_0) = \alpha,$$

  es gilt also „Fehlerwahrscheinlichkeit 1. Art“ $= \alpha$.

- Beim Gauß-Test mit **zusammengesetzter** Nullhypothese $H_0 : \mu \leq \mu_0$ gilt

  $$P(T \in K || \mu) \leq \alpha \quad \text{für} \quad \mu \leq \mu_0,$$

  es gilt also „Fehlerwahrscheinlichkeit 1. Art“ $\leq \alpha$.

16-4

Bemerkung 16.3 (Teststrategie)

- Parametertests werden in der Regel so durchgeführt, dass eine Ablehnung von H_0 angestrebt wird, um dadurch eine Bestätigung von H_1 zu erhalten.
- Falls es zu einer Ablehnung von H_0 kommt, geschieht diese Entscheidung mit einer durch das vorgegebene Signifikanzniveau α **kontrollierten Fehlerwahrscheinlichkeit 1. Art**.
- Im Fall einer Nichtablehnung von H_0 ist dagegen die **Fehlerwahrscheinlichkeit 2. Art in der Regel unbekannt**. Die Nichtablehnung der Nullhypothese H_0 ist dann nur eine schwache Bestätigung von H_0.

16-5

16.3 p-Wert

Bemerkung 16.4 (p-Wert)

- Eine Testentscheidung kann häufig auch dadurch getroffen werden, dass das vorgegebene Signifikanzniveau α mit dem sogenannten **p-Wert** (engl.: *p-value*) verglichen wird.
- Der p-Wert hängt von den Hypothesen und dem beobachteten Wert t der Prüfgröße T ab.

16-6

Bemerkung 16.5 (Gauß-Test und p-Wert)

- Beim Gauß-Test mit den Hypothesen $H_0 : \mu = \mu_0$ und $H_1 : \mu > \mu_0$ und dem berechneten Wert t ist der p-Wert die Wahrscheinlichkeit

$$p = P(T > t||\mu_0)\,. \tag{16.1}$$

  Dies ist die Wahrscheinlichkeit für alle Werte der Teststatistik T, die stärker als der beobachtete Wert t zugunsten von H_1 sprechen.

- Beim Gauß-Test mit den Hypothesen $H_0 : \mu = \mu_0$ und $H_1 : \mu < \mu_0$ ist der p-Wert die Wahrscheinlichkeit

$$p = P(T < t||\mu_0)\,. \tag{16.2}$$

16-7

- Beim Gauß-Test mit den Hypothesen $H_0 : \mu = \mu_0$ und $H_1 : \mu \neq \mu_0$ und dem beobachteten Wert t ist der p-Wert die Wahrscheinlichkeit

$$p = P(|T| > |t|\,||\mu_0), \tag{16.3}$$

  das ist die Wahrscheinlichkeit für alle Werte der Teststatistik, die stärker gegen H_0 sprechen als der beobachtete Wert t.

16-8

Bemerkung 16.6 (p-Wert und Signifikanzniveau)

- **Allgemein gilt:** Der p-Wert ist bei gegebenem Wert t der Teststatistik das kleinste Signifikanzniveau, bei dem H_0 zugunsten von H_1 abgelehnt wird.

- Wird der p-Wert in Abhängigkeit von der Gegenhypothese mit (16.1), (16.2) oder (16.3) berechnet, so ergibt sich folgende einfache **Entscheidungsregel**:

  Bei vorgegebenem Signifikanzniveau α wird H_0 zugunsten von H_1 abgelehnt, falls $p \leq \alpha$.

- Im Unterschied zu α hängt der p-Wert von den Daten ab, er wird manchmal auch als **empirisches Signifikanzniveau** bezeichnet.

16-9

16.4 Tests und Statistiksoftware

Bemerkung 16.7 (p-Wert und Statistiksoftware)

- Statistiksoftware wie z. B. SPSS, SAS usw. sieht in der Regel weder die Vorgabe eines Signifikanzniveaus vor, noch die Angabe, welche Gegenhypothese getestet werden soll.
- Da der p-Wert von der Gegenhypothese abhängt – vergleiche die drei unterschiedlichen Wahrscheinlichkeiten (16.1), (16.2) und (16.3) – kann der p-Wert nicht unmittelbar berechnet werden.
- Statistiksoftware gibt vielmehr eine Wahrscheinlichkeit aus, mit der sich der p-Wert bestimmen lässt und die in Spezialfällen mit dem p-Wert übereinstimmt.

16-10

Bemerkung 16.8 (Gauß-Test und Statistiksoftware)

- Statistiksoftware (z. B. SPSS) berechnet beim Gauß-Test in der Regel die folgende Wahrscheinlichkeit $\tilde{p}$ ohne expliziten Bezug auf eine Gegenhypothese:

$$\tilde{p} = \begin{cases} P(T > t||\mu_0) & & t > 0 \\ P(T > t||\mu_0) = P(T < t||\mu_0) = \frac{1}{2} & \text{für} & t = 0 \\ P(T < t||\mu_0) & & t < 0 \end{cases} . \tag{16.4}$$

  Es gilt immer $\tilde{p} \leq 1/2$.
- Die Wahrscheinlichkeit $\tilde{p}$ wird z. B. in deutschsprachigen Ausgaben von SPSS als **Signifikanz** bezeichnet, wird manchmal aber auch etwas irreführend als p-Wert bezeichnet.

16-11

Bemerkung 16.9 (Beziehung zwischen p-Wert und $\tilde{p}$)

1. Für die beiden Tests $H_0 : \mu = \mu_0$ versus $H_1 : \mu > \mu_0$ und $H_0 : \mu \leq \mu_0$ versus $H_1 : \mu > \mu_0$ gilt

$$p = \begin{cases} \tilde{p} & & t > 0 \\ \tilde{p} = 1/2 & \text{für} & t = 0 \\ 1 - \tilde{p} & & t < 0 \end{cases} .$$

2. Für die beiden Tests $H_0 : \mu = \mu_0$ versus $H_1 : \mu < \mu_0$ und $H_0 : \mu \geq \mu_0$ versus $H_1 : \mu < \mu_0$ gilt

$$p = \begin{cases} 1 - \tilde{p} & & t > 0 \\ \tilde{p} = 1/2 & \text{für} & t = 0 \\ \tilde{p} & & t < 0 \end{cases} .$$

3. Für den Test $H_0 : \mu = \mu_0$ versus $H_1 : \mu \neq \mu_0$ gilt

$$p = 2\tilde{p}.$$

16-12

Bemerkung 16.10 (Gauß-Test und Excel)

- Die Funktion GTEST in Excel 2003 ermöglicht die Durchführung von Gauß-Tests. Neben der Angabe der Beobachtungen ist die Eingabe von μ_0 und σ möglich. Berechnet wird die Wahrscheinlichkeit

$$p^* = P(T > t||\mu_0), \tag{16.5}$$

für die alle Werte zwischen 0 und 1 möglich sind.

- Diese Wahrscheinlichkeit ermöglicht die Berechnung des p-Wertes in Abhängigkeit von den zu testenden Hypothesen.

16-13

Bemerkung 16.11 (Beziehung zwischen p-Wert und p^*)

1. Für die beiden Tests $H_0 : \mu = \mu_0$ versus $H_1 : \mu > \mu_0$ und $H_0 : \mu \leq \mu_0$ versus $H_1 : \mu > \mu_0$ gilt
$$p = p^* .$$
2. Für die beiden Tests $H_0 : \mu = \mu_0$ versus $H_1 : \mu < \mu_0$ und $H_0 : \mu \geq \mu_0$ versus $H_1 : \mu < \mu_0$ gilt
$$p = 1 - p^* .$$
3. Für den beidseitigen Test $H_0 : \mu = \mu_0$ versus $H_1 : \mu \neq \mu_0$ gilt
$$p = 2 \min\{p^*, 1 - p^*\} .$$

16.5 Ergänzungen

Beispiel 16.a (Einseitiger Gauß-Test)

1. $X_i \overset{i.i.d.}{\sim} N(\theta, \sigma^2)$, σ^2 bekannt, $\theta_0 \in \mathbb{R}$ ist gegeben (z. B. $\theta_0 = 0$):
$$H_0 : \theta \leq \theta_0, \quad H_1 : \theta > \theta_0, \quad T = \frac{\bar{X} - \theta_0}{\sigma}\sqrt{n}, \quad K = (u_{1-\alpha}, \infty).$$
2. $T \in K$ d. h. $T > u_{1-\alpha}$ führt zur Ablehnung von H_0.
3. Die Verteilung von T hängt von θ ab. Es gilt $T \sim N(0, 1)$, falls $\theta = \theta_0$.
4. Die **Fehlerwahrscheinlichkeit 1. Art**, d. h. die Wahrscheinlichkeit, die Nullhypothese zu verwerfen, falls sie richtig ist, ist bei diesem Test
$$P(T \in K||\theta) \quad \text{für} \quad \theta \leq \theta_0.$$
Diese ist nicht konstant, sondern hängt von θ ab. Dabei gilt
$$P(T \in K||\theta_0) = \alpha, \tag{16.6}$$

$$P(T \in K||\theta) < \alpha, \quad \theta < \theta_0 \tag{16.7}$$

und

$$\lim_{\theta \to -\infty} P(T \in K||\theta) = 0\,. \tag{16.8}$$

5. Die **Fehlerwahrscheinlichkeit 2. Art**, d. h. die Wahrscheinlichkeit, die Nullhypothese beizubehalten, obwohl sie falsch ist, ist bei diesem Test

$$P(T \notin K||\theta) \quad \text{für} \quad \theta > \theta_0\,.$$

Falls H_0 falsch ist, ist

$$P(T \in K||\theta) \quad \text{für} \quad \theta > \theta_0$$

die Wahrscheinlichkeit mit der die Nullhypothese abgelehnt wird.

6. Im Beispiel gilt

$$P(T \in K||\theta) > \alpha \quad \text{für} \quad \theta > \theta_0 \tag{16.9}$$

und

$$\lim_{\theta \to \infty} P(T \in K||\theta) = 1\,. \tag{16.10}$$

7. Wünschenswert ist, dass die Fehlerwahrscheinlichkeit 2. Art möglichst klein ist und damit $P(T \in K||\theta)$ für $\theta > \theta_0$ möglichst groß ist.

Bemerkung 16.b (Zur Rolle von Null- und Gegenhypothese)

1. In der **Regel** versucht man, mit einem Test die Nullhypothese zu verwerfen. Das gewünschte Testergebnis ist also die Verwerfung der Nullhypothese, um die Gegenhypothese zu stützen. Falls die Nullhypothese verworfen wird, geschieht dies mit bekannter Fehlerwahrscheinlichkeit, denn das vorgegebene Signifikanzniveau α beschränkt die Wahrscheinlichkeit für den Fehler 1. Art. Falls die Nullhypothese dagegen nicht verworfen werden kann, ist in der Regel die Wahrscheinlichkeit für den Fehler 2. Art unbekannt, daher ist die Nichtverwerfung der Nullhypothese nur eine schwache Stützung der Nullhypothese.

2. **Ausnahmen** sind Tests, bei denen die Wahrscheinlichkeit für den Fehler 2. Art bei der Testdurchführung (z. B. durch geeignete Wahl des Stichprobenumfangs) kontrolliert wird und sogenannte Anpassungstests (engl.: *goodness-of-fit tests*).

Bemerkung 16.c (Fehlerquellen am Beispiel des Gauß-Tests)

1. Die Verwendung von Statistiksoftware verleitet zur folgenden **fehlerhaften Testdurchführung**, bei der eine einseitige Gegenhypothese in Abhängigkeit von den Daten bzw. vom realisierten Wert t formuliert wird. Es ist angenommen, dass $\tilde{p}$ aus (16.4) gegeben ist.

   (a) Es wird z. B. $H_0 : \mu = \mu_0$ gegen $H_1 : \mu > \mu_0$ dann getestet, wenn $t > 0$, und $H_0 : \mu = \mu_0$ gegen $H_1 : \mu < \mu_0$ dann getestet, wenn $t < 0$.

   (b) Dieses Vorgehen wird realisiert, indem man H_0 verwirft, falls $\tilde{p} < \alpha$, und dabei entweder $H_1 : \mu > \mu_0$ formuliert, falls $t > 0$, oder $H_1 : \mu < \mu_0$ formuliert, falls $t < 0$.

   (c) Als Folge dieses Vorgehens, ist die Fehlerwahrscheinlichkeit 1. Art nicht, wie eigentlich gewünscht, α, sondern 2α. Die Wahrscheinlichkeit, eine richtige Nullhypothese $H_0 : \mu = \mu_0$ zu verwerfen, ist also verdoppelt.

   Damit das Signifikanzniveau α eingehalten wird, muss die **einseitige Gegenhypothese unabhängig von den Daten formuliert** werden.

2. Bei der Verwendung von Statistiksoftware besteht beim zweiseitigen Gauß-Test die Fehlerquelle, dass ein vorgegebenes Signifikanzniveau α (z. B. $\alpha = 5\%$) mit dem von der Software ausgegebenen Wert $\tilde{p}$ anstatt mit $2\tilde{p}$ verglichen wird. Als Folge dieses Vorgehens, ist die Fehlerwahrscheinlichkeit 1. Art nicht, wie eigentlich gewünscht, α, sondern 2α. Die Wahrscheinlichkeit, eine richtige Nullhypothese $H_0 : \mu = \mu_0$ zu verwerfen, ist also verdoppelt.

3. Die Möglichkeit, einen zweiseitigen Test zum Signifikanzniveau α durchzuführen, indem man $\tilde{p}$ verdoppelt, beruht allerdings auf der Symmetrie der Normalverteilung und lässt sich nicht auf Tests mit asymmetrischer Verteilung der Teststatistik übertragen.

4. Ein Gauß-Test unter Verwendung von $\tilde{p}$ aus (16.4) wird folgendermaßen durchgeführt. Vorgegeben ist ein kleines α $(0 < \alpha < 1/2)$, z. B. $\alpha = 5\%$.

(a) Für die Tests $H_0 : \mu = \mu_0$ oder $H_0 : \mu \leq \mu_0$ versus $H_1 : \mu > \mu_0$ wird H_0 abgelehnt, falls $\tilde{p} < \alpha$ **und** $t > 0$.

(b) Für die Tests $H_0 : \mu = \mu_0$ oder $H_0 : \mu \geq \mu_0$ versus $H_1 : \mu < \mu_0$ wird H_0 abgelehnt, falls $\tilde{p} < \alpha$ **und** $t < 0$.

(c) Für den Test $H_0 : \mu = \mu_0$ versus $H_1 : \mu \neq \mu_0$ wird H_0 abgelehnt, falls $2\tilde{p} < \alpha$.

5. Für die Wahrscheinlichkeit p^* aus (16.5) gilt

$$p^* = P(T > t||\mu_0) = 1 - P(T \leq t||\mu_0) = 1 - \Phi(t),$$

da T standardnormalverteilt ist, falls $\mu = \mu_0$.

Bemerkung 16.d (Eine Fehlinterpretation) Ein Irrtum: große Stichprobenumfänge seien „nicht sinnvoll", da man „dann jede Nullhypothese verwerfen könne".

Dieser Irrtum beruht auf einer Fehlinterpretation des folgenden Sachverhaltes: Bei einem Gauß-Test wird die Nullhypothese $H_0 : \mu \leq 0$, falls sie richtig ist, mit zunehmendem Stichprobenumfang bei immer kleineren Werten von $\bar{x} > 0$ abgelehnt.

Dies bedeutet aber **nicht**, dass eine richtige Nullhypothese mit größerer Wahrscheinlichkeit abgelehnt wird, wenn der Stichprobenumfang zunimmt, denn der Gauß-Test beruht *nicht* auf $\bar{X}$, sondern auf der Statistik

$$T = \frac{\bar{X}}{\sigma}\sqrt{n}.$$

Beim einseitigen Gauß-Test für $H_0 : \mu \leq 0$ vs. $H_1 : \mu > 0$ können drei Fälle unterschieden werden:

1. Im Fall $\mu = 0$ ist die Fehlerwahrscheinlichkeit 1. Art für **jeden** Stichprobenumfang

$$P(T > u_{1-\alpha}||0) = 1 - \Phi\left(u_{1-\alpha}\right) = \alpha.$$

2. Im Fall $\mu < 0$ ist die Fehlerwahrscheinlichkeit 1. Art

$$P(T > u_{1-\alpha}||\mu) = 1 - \Phi\left(u_{1-\alpha} - \frac{\mu}{\sigma}\sqrt{n}\right) < \alpha\,.$$

Diese **sinkt** streng monoton mit wachsendem Stichprobenumfang.

3. Im Fall $\mu > 0$ ist die Fehlerwahrscheinlichkeit 2. Art

$$P(T \leq u_{1-\alpha}||\mu) = \Phi\left(u_{1-\alpha} - \frac{\mu}{\sigma}\sqrt{n}\right) < 1 - \alpha\,.$$

Diese **sinkt** ebenfalls streng monoton mit wachsendem Stichprobenumfang.

Bemerkung 16.e (Vertrauenswahrscheinlichkeit)

1. Bei gegebenem Signifikanzniveau α wird $1 - \alpha$ manchmal als Vertrauenswahrscheinlichkeit bezeichnet[1]. Diese Bezeichnung wird damit gerechtfertigt, dass bei einer einfachen Nullhypothese $1 - \alpha$ in der Regel die Wahrscheinlichkeit ist, mit der eine richtige Nullhypothese beibehalten wird. Die Bezeichnung ist aber deswegen eher unglücklich, weil das Ziel eines Tests in der Regel nicht die Beibehaltung einer richtigen Nullhypothese mit hoher Vertrauenswahrscheinlichkeit ist, sondern die Verwerfung einer falschen Nullhypothese mit kleiner Fehlerwahrscheinlichkeit 1. Art ist. Vgl. dazu Bemerkung 16.b.
2. Manchmal wird auch das Konfidenzniveau bei der Konstruktion von Konfidenzintervallen als Vertrauenswahrscheinlichkeit bezeichnet.[2]

Bemerkung 16.f (Irrtumswahrscheinlichkeit) Das vorgegebene Signifikanzniveau α eines Tests wird manchmal als Irrtumswahrscheinlichkeit bezeichnet[3].

1. Im Allgemeinen ist die Irrtumswahrscheinlichkeit eines Tests über $H_0 : \theta \in \Theta_0$ mit der Prüfgröße T und dem kritischen Bereich K die maximale (oder allgemeiner supremale) Fehlerwahrscheinlichkeit 1. Art, also z. B.

$$\max_{\theta\in\Theta_0} P(T \in K||\theta) \quad \text{bzw.} \quad \sup_{\theta\in\Theta_0} P(T \in K||\theta). \tag{16.11}$$

[1] Z. B. S. 70 in: Backhaus, K., Erichson, P., Plinke, W., Weiber, R.: Multivariate Analysemethoden. Eine anwendungsorientierte Einführung. 11. Aufl. Springer: Berlin, Heidelberg, New York 2006.

[2] A. a. O., S. 77.

[3] A. a. O., S. 71.

2. Bei einem Test mit Signifikanzniveau α kann die Irrtumswahrscheinlichkeit im Sinn von (16.11) kleiner als das vorgegebene Signifikanzniveau α sein.

3. Wenn die Irrtumswahrscheinlichkeit im Sinn von (16.11) mit dem vorgegebenen Signifikanzniveau α übereinstimmt, liegt ein sogenannter Test mit dem Umfang α vor.

4. Bei vielen Testverfahren, so auch auch bei allen im vorangegangenen Kapitel und im folgenden Kapitel behandelten Testverfahren, wird der kritische Bereich so konstruiert, dass die Irrtumswahrscheinlichkeit im Sinn von (16.11) gleich dem vorgegebenen Signifikanzniveau ist.

Kapitel 17

Tests für den Vergleich zweier Parameter

17-1

17. Tests für den Vergleich zweier Parameter

17.1 Zweistichproben-Gauß-Test
17.2 Zweistichproben-t-Test
17.3 Vergleich zweier Varianzen (F-Test)
17.4 Der t-Differenzentest

17-2

Bemerkung 17.1

- Bei den folgenden drei Tests werden Parameter von **zwei normalverteilten Grundgesamtheiten** verglichen.
- Vorausgesetzt sind jeweils zwei **unabhängige** normalverteilte Stichproben
$$X_1, X_2, \ldots, X_n \overset{i.i.d.}{\sim} N(\mu_X, \sigma_X^2),$$
$$Y_1, Y_2, \ldots, Y_m \overset{i.i.d.}{\sim} N(\mu_Y, \sigma_Y^2).$$
- Unabhängige Stichproben bedeutet, dass die Komponenten des Vektors
$$(X_1, X_2, \ldots, X_n, Y_1, Y_2, \ldots, Y_m)$$
insgesamt stochastisch unabhängig sind.
- Im vierten Test geht es um zwei Merkmale **derselben Grundgesamtheit**.

17-3

17.1 Zweistichproben-Gauß-Test

Bemerkung 17.2 (Schematische Darstellung)

- **Zweck:** Vergleich der Erwartungswerte von zwei normalverteilten Grundgesamtheiten mit **bekannten** Varianzen σ_X^2 und σ_Y^2.
- **Voraussetzungen:** Zwei *unabhängige* normalverteilte Stichproben
$$X_1, \dots, X_n \overset{i.i.d.}{\sim} N(\mu_X, \sigma_X^2), \quad Y_1, \dots, Y_m \overset{i.i.d.}{\sim} N(\mu_Y, \sigma_Y^2).$$
- **Hypothesen:**
$$H_0 : \mu_X = \mu_Y, \quad H_1 : \mu_X \neq \mu_Y$$
- **Prüfgröße und kritischer Bereich:**
$$T = \frac{\bar{X} - \bar{Y}}{\sqrt{\frac{\sigma_X^2}{n} + \frac{\sigma_Y^2}{m}}}, \quad K = \,]-\infty, u_{\frac{\alpha}{2}}[\,\cup\,]u_{1-\frac{\alpha}{2}}, \infty[$$

17-4

Bemerkung 17.3 (Theoretischer Hintergrund)

- Es gilt
$$\mathbb{E}[\bar{X} - \bar{Y}] = \mu_X - \mu_Y = 0,$$
falls H_0 richtig ist, und es gilt
$$\mathbb{V}[\bar{X} - \bar{Y}] = \mathbb{V}[\bar{X}] + \mathbb{V}[\bar{Y}] = \frac{\sigma_X^2}{n} + \frac{\sigma_Y^2}{m}.$$
- Unter den Voraussetzungen ist T standardnormalverteilt, falls H_0 richtig ist.

17-5

17.2 Zweistichproben-t-Test

Bemerkung 17.4 (Schematische Darstellung)

- **Zweck:** Vergleich der Erwartungswerte von zwei Normalverteilungen mit unbekannten, aber identischen Varianzen $\sigma_X^2 = \sigma_Y^2$.
- **Voraussetzung:** Zwei unabhängige normalverteilte Stichproben
$$X_1, \dots, X_n \overset{i.i.d.}{\sim} N(\mu_X, \sigma_X^2), \quad Y_1, \dots, Y_m \overset{i.i.d.}{\sim} N(\mu_Y, \sigma_Y^2), \quad \sigma_X^2 = \sigma_Y^2.$$
- **Hypothesen:** $H_0 : \mu_X = \mu_Y, \quad H_1 : \mu_X \neq \mu_Y$
- **Prüfgröße:**
$$T = \sqrt{\frac{nm(n+m-2)}{n+m}} \frac{\bar{X} - \bar{Y}}{\sqrt{(n-1)S_X^{*2} + (m-1)S_Y^{*2}}}$$
- Der **kritische Bereich** ist $K = \,]-\infty, t_{n+m-2,\frac{\alpha}{2}}[\,\cup\,]t_{n+m-2,1-\frac{\alpha}{2}}, \infty[$.

17-6

17.3 Vergleich zweier Varianzen (F-Test)

Bemerkung 17.5 (Schematische Darstellung)

- **Zweck:** Vergleich der Varianzen von zwei Normalverteilungen
- **Voraussetzung:** Zwei unabhängige normalverteilte Stichproben

$$X_1, \ldots, X_n \overset{i.i.d.}{\sim} N(\mu_X, \sigma_X^2), \quad Y_1, \ldots, Y_m \overset{i.i.d.}{\sim} N(\mu_Y, \sigma_Y^2).$$

- **Hypothesen:** $H_0 : \sigma_X^2 = \sigma_Y^2, \quad H_1 : \sigma_X^2 \neq \sigma_Y^2$
- **Prüfgröße:**

$$T = \frac{S_X^{*2}}{S_Y^{*2}}$$

- **Kritischer Bereich:**

$$K = \left[0, F_{n-1,m-1,\frac{\alpha}{2}}\right[\; \cup \; \left]F_{n-1,m-1,1-\frac{\alpha}{2}}, \infty\right[$$

17-7

Bemerkung 17.6

- $F_{n,m,p}$ bezeichnet das p-Quantil der nach *R. A. Fisher* benannten F-Verteilung mit n und m Freiheitsgraden.
- Die p-Quantile der F-Verteilung sind vertafelt.
- Die Eigenschaft

$$F_{n,m,1-p} = \frac{1}{F_{m,n,p}}$$

  ermöglicht es, z. B. Quantile $F_{n,m,0.05}$ aus einer Tabelle mit den Quantilen $F_{n,m,0.95}$ zu bestimmen, indem die Rolle von n und m vertauscht wird und der Kehrwert gebildet wird.

17-8

Beispiel 17.7 (F-Verteilung)

- Gegeben: $n = 11, \quad m = 6, \quad \alpha = 0.1, \quad 1 - \frac{\alpha}{2} = 0.95$
- Gesucht: $K = [0, F_{10,5,0.05}[\; \cup \;]F_{10,5,0.95}, \infty[$
- Tabellenwerte:

$$F_{10,5,0.95} = 4.735$$

$$F_{10,5,0.05} = \frac{1}{F_{5,10,0.95}} = \frac{1}{3.326} = 0.301$$

- Kritischer Bereich:

$$K = [0, 0.301[\; \cup \;]4.735, \infty[$$

17-9

Definition 17.8 (F-Verteilung)
Die Zufallsvariablen $A \sim \chi^2_n$ und $B \sim \chi^2_m$ seien stochastisch unabhängig.

1. Dann heißt die Verteilung von
$$Z = \frac{\frac{A}{n}}{\frac{B}{m}}$$
F-Verteilung mit n und m Freiheitsgraden.
2. Schreibweise: $Z \sim F_{n,m}$.
3. Man bezeichnet n als die Anzahl der **Zählerfreiheitsgrade** und m als die Anzahl der **Nennerfreiheitsgrade**.

17-10

17.4 Der t-Differenzentest

- **Zweck:** Test über die Differenz der Mittelwerte von zwei Merkmalen X und Y **derselben** Grundgesamtheit.
- **Voraussetzungen:**
  1. $\mathbb{E}[X_i] = \mu_X$ und $\mathbb{E}[Y_i] = \mu_Y$ für $i = 1, \ldots, n$.
  2. Die Differenzen sind normalverteilt,
  $$D_i \overset{\text{def}}{=} X_i - Y_i \overset{i.i.d.}{\sim} N(\mu_D, \sigma^2_D).$$
- Hypothesen über
$$\mu_D \overset{\text{def}}{=} \mathbb{E}[D_i] = \mathbb{E}[X_i] - \mathbb{E}[Y_i] = \mu_X - \mu_Y$$
können mit den Differenzen D_i getestet werden.
- **Hypothesen:**
$$H_0 : \mu_D = \mu_0, \quad H_1 : \mu_D \neq \mu_0$$

17-11

- **Prüfgröße:**
$$T = \frac{\bar{D} - \mu_0}{S^*_D}\sqrt{n} = \frac{\bar{D} - \mu_0}{S_D}\sqrt{n-1}$$
mit
$$\bar{D} = \frac{1}{n}\sum_{i=1}^{n} D_i$$
und
$$S^*_D = \sqrt{\frac{1}{n-1}\sum_{i=1}^{n}(D_i - \bar{D})^2}\,, \quad S_D = \sqrt{\frac{1}{n}\sum_{i=1}^{n}(D_i - \bar{D})^2}\,.$$
- **Kritischer Bereich:**
$$K = \left]-\infty, t_{n-1,\frac{\alpha}{2}}\right[\;\cup\; \left]t_{n-1,1-\frac{\alpha}{2}}, \infty\right[$$

17-12

Bemerkung 17.9
Eine häufige Anwendung ist ein Test mit $\mu_0 = 0$, d. h. es wird getestet, ob die Mittelwerte gleich sind.

$$H_0 : \mu_X = \mu_Y , \quad H_1 : \mu_X \neq \mu_Y$$

Bemerkung 17.10
Nicht X_i und Y_i sind als unabhängig vorausgesetzt, sondern (X_i, Y_i) und (X_j, Y_j) sind für $i \neq j$ unabhängig und identisch verteilt.
Man spricht auch von *verbundenen* Stichproben: $(X_i, Y_i), i = 1, \ldots, n$ sind unabhängige und identisch verteilte Beobachtungen aus einer zweidimensionalen Verteilung.

17-13

Beispiel 17.11 (Verbundene und unverbundene Stichproben)

- *verbundene* Stichproben:

  Zwei Merkmale, n Personen: $(x_1, y_1), (x_2, y_2), \ldots, (x_n, y_n)$

  (x_j, y_j) gehört zu j-ter Person wie z.B. (Statistiknote, Mathematiknote).

- *unabhängige* Stichproben:

  n_1 Personen aus Grundgesamtheit G_1: $(x_1, x_2, ..., x_{n_1})$
  n_2 Personen aus Grundgesamtheit G_2: $(y_1, y_2, ..., y_{n_2})$

  Evtl., aber nicht zwingend gilt $n_1 = n_2 = n$.

  x_i sind die Statistiknoten von n_1 männlichen Studierenden,
  y_j sind die Statistiknoten von n_2 weiblichen Studierenden.

17-14

Bemerkung 17.12

- Für die Varianz $\sigma_D^2 \stackrel{\text{def}}{=} \mathbb{V}[D_i]$ gilt

$$\begin{aligned} \sigma_D^2 = \mathbb{V}[X_i - Y_i] &= \mathbb{V}[X_i] - 2\mathbb{C}\mathrm{ov}[X_i, Y_i] + \mathbb{V}[Y_i] \\ &= \sigma_X^2 - 2\mathbb{C}\mathrm{ov}[X_i, Y_i] + \sigma_Y^2 \end{aligned}$$

- Es muss nicht $\sigma_X^2 = \sigma_Y^2$ gelten.

- Nicht hinreichend für die Voraussetzung $D_i \stackrel{i.i.d.}{\sim} N(\mu_D, \sigma_D^2)$ ist, dass $X_i \stackrel{i.i.d.}{\sim} N(\mu_X, \sigma_X^2)$ und $Y_i \stackrel{i.i.d.}{\sim} N(\mu_Y, \sigma_Y^2)$ gilt.

- In der Praxis wird der t-Differenzentest häufig missbräuchlich, nämlich ohne Rücksicht auf die Voraussetzung, angewendet.

17.5 Ergänzungen

Bemerkung 17.a

Eine hinreichende Voraussetzung für die Anwendung des t-Differenzentests ist, dass die (X_i, Y_i) zweidimensional (bivariat) normalverteilt sind. Wenn diese Voraussetzung erfüllt ist, sind auch die Differenzen normalverteilt.

Kapitel 18

Approximative Testverfahren

18-1

18. Approximative Testverfahren

18.1 Approximativer Gauß-Test
18.2 Approximativer Zwei-Stichproben-Gauß-Test
18.3 Test für eine Wahrscheinlichkeit
18.4 Vergleich zweier Wahrscheinlichkeiten
18.5 Chiquadrat-Anpassungstest
18.6 Chiquadrat-Unabhängigkeitstest

18-2

Bemerkung 18.1

- Alle Verfahren in diesem Kapitel sind approximativ.
- Der **Approximationsfehler** wird mit zunehmendem Stichprobenumfang kleiner.
- Asymptotisch, für $n \to \infty$, verschwindet der Approximationsfehler.
- Dem Nachteil des Approximationsfehlers steht der Vorteil gegenüber, dass keine normalverteilte Grundgesamtheit erforderlich ist.

18-3

18.1 Approximativer Gauß-Test

Bemerkung 18.2 (σ^2 bekannt)

- **Zweck**: Test über den Erwartungswert einer Verteilung.
- **Voraussetzungen**: Die X_i seien unabhängig und identisch verteilt (i.i.d.) mit $\mu \stackrel{\text{def}}{=} \mathbb{E}[X_i]$ und $0 < \sigma^2 \stackrel{\text{def}}{=} \mathbb{V}[X_i] < \infty$.
- Der auf der **Prüfgröße**
$$T = \frac{\bar{X} - \mu_0}{\sigma}\sqrt{n}$$
beruhende Gauß-Test für die Nullhypothese $H_0 : \mu = \mu_0$ ist approximativ gültig.
- Faustregel für die Anwendung: $n \geq 40$.
- Der Approximationsfehler besteht darin, dass das vorgegebene Signifikanzniveau α nur approximativ eingehalten wird.

18-4

Bemerkung 18.3 (σ^2 unbekannt)

- Wenn die Varianz σ^2 unbekannt, aber endlich ist, kann diese in der Prüfgröße durch den Schätzer S^2 ersetzt werden.
- Der mit der modifizierten Prüfgröße
$$T = \frac{\bar{X} - \mu_0}{S}\sqrt{n}$$
durchgeführte Gauß-Test für die Nullhypothese $H_0 : \mu = \mu_0$ ist approximativ gültig.
- Faustregel für die Anwendung: $n \geq 40$.

18-5

18.2 Approximativer Zwei-Stichproben-Gauß-Test

Bemerkung 18.4

- **Zweck**: Vergleich der Erwartungswerte von zwei Grundgesamtheiten mit bekannten Varianzen σ_X^2 und σ_Y^2.
- **Voraussetzungen**: Zwei *unabhängige* Stichproben
$$X_1, X_2, \ldots, X_n \text{ i.i.d.} \quad \mathbb{E}[X_i] = \mu_X, \quad \mathbb{V}[X_i] = \sigma_X^2,$$
$$Y_1, Y_2, \ldots, Y_m \text{ i.i.d.} \quad \mathbb{E}[Y_i] = \mu_Y, \quad \mathbb{V}[Y_i] = \sigma_Y^2.$$

18-6

Bemerkung 18.5 (Bekannte Varianzen)

- Der Zwei-Stichproben-Gauß-Test mit der Prüfgröße

$$T = \frac{\bar{X} - \bar{Y}}{\sqrt{\frac{\sigma_X^2}{n} + \frac{\sigma_Y^2}{m}}}$$

  für die Nullhypothese $H_0 : \mu_X = \mu_Y$ ist approximativ gültig.

- Faustregel für die Anwendung: $n \geq 40$, $m \geq 40$.

18-7

Bemerkung 18.6 (Unbekannte Varianzen)

- Wenn die Varianzen σ_X^2 und σ_Y^2 unbekannt, aber endlich sind, können diese in der Prüfgröße durch die Schätzer S_X^2 und S_Y^2 ersetzt werden.

- Der mit der Prüfgröße

$$\frac{\bar{X} - \bar{Y}}{\sqrt{\frac{S_X^2}{n} + \frac{S_Y^2}{m}}}$$

  durchgeführte Zwei-Stichproben-Gauß-Test für die Nullhypothese $H_0 : \mu_X = \mu_Y$ ist approximativ gültig.

- Faustregel für die Anwendung: $n \geq 40$, $m \geq 40$.

18-8

18.3 Test für eine Wahrscheinlichkeit

Bemerkung 18.7

- **Zweck:** Test über eine Wahrscheinlichkeit π.

- **Voraussetzung:** Bernoulli-verteilte Beobachtungen $X_i \overset{i.i.d.}{\sim} B(1, \pi)$.

- **Nullhypothese, Gegenhypothese:**

$$H_0 : \pi = \pi_0, \quad H_1 : \pi \neq \pi_0$$

- **Prüfgröße, kritischer Bereich:**

$$T = \frac{\hat{\pi} - \pi_0}{\sqrt{\pi_0(1 - \pi_0)}} \sqrt{n}, \quad K = \left]-\infty, u_{\frac{\alpha}{2}}\right[\cup \left]u_{1-\frac{\alpha}{2}}, \infty\right[$$

  $\hat{\pi} = \bar{X}$ ist der zufällige Anteil (die relative Häufigkeit) in der Stichprobe.

- **Faustregel für die Anwendung:** $n\pi_0(1 - \pi_0) > 9$

18-9

18.4 Vergleich zweier Wahrscheinlichkeiten

Bemerkung 18.8

- **Zweck:** Vergleich von zwei Wahrscheinlichkeiten π_X und π_Y
- **Voraussetzung**: Zwei unabhängige Stichproben:
  $X_1, \ldots, X_n \overset{i.i.d.}{\sim} B(1, \pi_X)$ und $Y_1, \ldots, Y_m \overset{i.i.d.}{\sim} B(1, \pi_Y)$
- **Nullhypothese, Gegenhypothese:**

$$H_0 : \pi_X = \pi_Y, \quad H_1 : \pi_X \neq \pi_Y$$

- **Prüfgröße, kritischer Bereich:**

$$T = \frac{\hat{\pi}_X - \hat{\pi}_Y}{\sqrt{\frac{\hat{\pi}_X(1-\hat{\pi}_X)}{n} + \frac{\hat{\pi}_Y(1-\hat{\pi}_Y)}{m}}}, \quad K = \left]-\infty, u_{\frac{\alpha}{2}}\right[\cup \left]u_{1-\frac{\alpha}{2}}, \infty\right[$$

  $\hat{\pi}_X = \bar{X}$ und $\hat{\pi}_Y = \bar{Y}$ sind die Anteile in den beiden Stichproben.

18-10

- **Faustregel für die Anwendung:**

$$n\hat{\pi}_X(1 - \hat{\pi}_X) > 9 \quad \text{und} \quad m\hat{\pi}_Y(1 - \hat{\pi}_Y) > 9$$

18-11

18.5 Chiquadrat-Anpassungstest

Bemerkung 18.9

- **Zweck**: Überprüfung, ob die Beobachtungen mit einer theoretischen Verteilung verträglich sind, die durch die vorgegebene Verteilungsfunktion F_0 festgelegt ist.
- **Voraussetzung**: Die X_i sind stochastisch unabhängig und identisch verteilt (i.i.d.) mit der Verteilungsfunktion F.
- **Hypothesen:**

$$H_0 : F = F_0, \quad H_1 : F \neq F_0$$

- **Testdurchführung**
  Gegeben ist eine vollständige **Zerlegung** (Klassierung) $A_1, \ldots, A_J$ mit den zugehörigen Wahrscheinlichkeiten

$$\pi_j = P(X_i \in A_j || F_0), \qquad \sum_{j=1}^{J} \pi_j = 1$$

  und den **zufälligen absoluten Häufigkeiten**

$$N_j = \text{Anzahl der } X_i \in A_j, \qquad \sum_{j=1}^{J} N_j = n.$$

- **Prüfgröße:**

$$T = \sum_{j=1}^{J} \frac{(N_j - n\pi_j)^2}{n\pi_j} = \left(\sum_{j=1}^{J} \frac{N_j^2}{n\pi_j} \right) - n$$

- **Kritischer Bereich:**

$$K = \left]\chi^2_{J-1,1-\alpha}, \infty\right[$$

Bemerkung 18.10

- Da der Test auf einer Approximation beruht, die asymptotisch, d. h. für $n \to \infty$, begründet ist, darf der Stichprobenumfang nicht zu klein sein. Eine in der Literatur häufig genannte **Faustregel** verlangt

$$n\pi_j \geq 5 \quad \forall\, j = 1, \ldots, J.$$

- Große Werte der Prüfgröße führen zur Verwerfung von H_0, da

$$\frac{N_j}{n} \approx \pi_j, \quad N_j \approx n\pi_j,$$

  falls H_0 richtig und n hinreichend groß ist.

18-15

18.6 Chiquadrat-Unabhängigkeitstest

- **Zweck**
  Überprüfung der stochastischen Unabhängigkeit zweier Merkmale X und Y.
- **Voraussetzungen**
  1. Unabhängige zweidimensionale Beobachtungen (X_i, Y_i) für $i = 1, \dots, n$.
  2. Stichprobenumfang n hinreichend groß, da der Test auf einer Approximation beruht, die umso besser ist, je größer n ist.
- **Hypothesen**
  $H_0 : X$ und Y sind unabhängig, $H_1 : X$ und Y sind nicht unabhängig.

18-16

- **Klassierung beider Merkmale als Voraussetzung für die Anwendung**
  - Disjunkte **Intervalle** $A_1, \dots, A_J$ für die x-Werte und $B_1, \dots, B_K$ für die y-Werte.
  - Die **beobachteten Häufigkeiten** in $A_j \times B_k$ sind

$$N_{jk} \stackrel{\text{def}}{=} \text{Anzahl der } (X_i, Y_i) \text{ mit } X_i \in A_j \text{ und } Y_i \in B_k\,.$$

  - Die **Randhäufigkeiten** sind

$$N_{j.} \stackrel{\text{def}}{=} \sum_{k=1}^{K} N_{jk}, \quad N_{.k} \stackrel{\text{def}}{=} \sum_{j=1}^{J} N_{jk}\,.$$

  - Die **theoretischen Häufigkeiten** sind:

$$\tilde{N}_{jk} \stackrel{\text{def}}{=} n \frac{N_{j.}}{n} \frac{N_{.k}}{n} = \frac{N_{j.} N_{.k}}{n}$$

18-17

Bemerkung 18.11
Die theoretischen Häufigkeiten sind erwartete Häufigkeiten, falls H_0 richtig ist, und falls die relativen Häufigkeiten $N_{j.}/n$ und $N_{.k}/n$ als Wahrscheinlichkeiten interpretiert werden.

- **Prüfgröße**

$$T = \sum_{j=1}^{J} \sum_{k=1}^{K} \frac{(N_{jk} - \tilde{N}_{jk})^2}{\tilde{N}_{jk}} = n \left(\sum_{j=1}^{J} \sum_{k=1}^{K} \frac{N_{jk}^2}{N_{j.} N_{.k}} - 1 \right)$$

- **Kritischer Bereich**

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

18-18

Bemerkung 18.12 (Erforderlicher Stichprobenumfang)

- Der Chiquadrat-Unabhängigkeitstest ist ein approximatives Verfahren, das auf der asymptotischen Verteilung von T beruht. Daher sollte der Stichprobenumfang hinreichend groß sein.
- In Abhängigkeit von der Klassierung erhält man beobachtete Werte $n_{j.}$ und $n_{.k}$ für die zufälligen Häufigkeiten $N_{j.}$ und $N_{.k}$.
- Faustregel für die Anwendung:

$$\tilde{n}_{jk} \stackrel{\text{def}}{=} \frac{n_{j.}n_{.k}}{n} \geq 5 \quad \text{für alle} \quad j = 1, \ldots, J, \quad k = 1, \ldots, K\,.$$

Bemerkung 18.13
Der Chiquadrat-Unabhängigkeitstest heißt manchmal auch **Kontingenztest**.

18.7 Ergänzungen

Beispiel 18.a (Beispiel zum Chiquadrat-Anpassungstest)

1. **Daten**: Monatshöchststand des Elbepegels in Dresden (Quelle: Sächsische Zeitung) gerundet auf Meter. $n = 57$ Beobachtungen von August 2002 bis April 2007: 9, 3, 4, 5, 5, 7, 4, 4, 3, 2, 1, 1, 1, 1, 2, 1, 2, 3, 4, 4, 3, 2, 3, 1, 1, 2, 1, 3, 2, 3, 5, 6, 3, 2, 2, 3, 3, 2, 2, 1, 2, 2, 3, 7, 7, 4, 5, 4, 4, 2, 1, 2, 2, 3, 4, 4, 2

2. **Fragestellung:** Sind diese Beobachtungen mit einer Poisson-Verteilung gut verträglich? Aus langjähriger Beobachtung sei der Parameter $\mu = 3$ als bekannt vorausgesetzt. (Hinweis: das arithmetische Mittel der 57 Beobachtungen ist 3.05).

3. **Nullhypothese**: Die Beobachtungen $x_1, \ldots, x_n$ sind Realisationen einer Poisson-Verteilung mit dem Parameter $\mu = 3$.

4. Für eine Poisson-verteilte Zufallsvariable $X \sim Poi(\mu)$ gilt

$$p_k \stackrel{\text{def}}{=} P(X = k) = \frac{\mu^k}{k!} e^{-\mu}, \quad k = 0, 1, 2, \ldots$$

n_k bezeichne die Anzahl der Beobachtungen mit $x_i = k$. np_k ist die erwartete Häufigkeit für den Merkmalswert k bei n Versuchen.

Merkmalswert k	0	1	2	3	4	5
Beobachtete Häufigkeit n_k	0	10	16	12	10	4
Wahrscheinlichkeit p_k	0.0498	0.149	0.224	0.224	0.168	0.101
Erwartete Häufigkeit np_k	2.84	8.51	12.77	12.77	9.58	5.75

Merkmalswert k	6	7	8	9	≥ 10	≥ 0
Beobachtete Häufigkeit n_k	1	3	0	1	0	57
Wahrscheinlichkeit p_k	0.050	0.0216	0.00810	0.00270	0.0011	1
Erwartete Häufigkeit np_k	2.87	1.23	0.46	0.15	0.06	57

Die Wahrscheinlichkeit für die letzte offene Klasse mit $k \geq 10$ ergibt sich als

$$P(X \geq 10) = 1 - P(X \leq 9) = 1 - \sum_{k=0}^{9} p_k.$$

5. Eine **Klassierung** der Daten erfolgt mit dem Ziel: „Die erwartete Häufigkeit ist in jeder Klasse ≥ 5."

Merkmalswerte	≤ 1	2	3	4	≥ 5
Klasse j	1	2	3	4	5
Beobachtete Häufigkeit n_j	10	16	12	10	9
Wahrscheinlichkeit π_j	0.199	0.224	0.224	0.168	0.185
Erwartete Häufigkeit $n\pi_j$	11.35	12.74	12.77	9.58	10.53

   Die Wahrscheinlichkeiten werden bei der Klassierung addiert:

$$\pi_1 = p_0 + p_1, \quad \pi_5 = \sum_{i \geq 5} p_i$$

6. Der Wert der Testgröße (Chiquadrat-Anpassungstest) ist

$$t = \sum_{j=1}^{J} \frac{(n_j - n\pi_j)^2}{n\pi_j} = \frac{(10 - 11.35)^2}{11.35} + \cdots + \frac{(9 - 10.53)^2}{10.53} = 1.265 \,.$$

7. Hier gilt $J = 5$ (Anzahl der Klassen). Die Nullhypothese „H_0: Die Daten kommen aus einer Poisson-Verteilung mit dem Parameter $\mu = 3$" kann abgelehnt werden, falls $t > \chi^2_{J-1,1-\alpha}$.

8. Es gilt z. B. $\chi^2_{4,0.90} = 7.78$, $\chi^2_{4,0.95} = 9.49$ und $\chi^2_{4,0.99} = 13.28$. Mit einer Irrtumswahrscheinlichkeit von $\alpha = 0.10$ kann H_0 nicht verworfen werden.

9. Selbst wenn man eine erheblich größere Irrtumswahrscheinlichkeit einräumt, eine richtige Nullhypothese fehlerhaft zu verwerfen, wird H_0 nicht verworfen. Dies verdeutlicht der p-Wert.

10. Zum gegebenen Wert t der Prüfgröße T ist durch die Wahrscheinlichkeit p mit

$$\chi^2_{J-1,1-p} = t$$

    der sogenannte **p-Wert** gegeben. Der p-Wert ist das kleinste Signifikanzniveau, bei dem die Nullhypothese abgelehnt werden kann. Im Beispiel ist $p = 0.867$ der p-Wert zu $t = 1.265$. Nur für $p \leq \alpha$ kann die Nullhypothese abgelehnt werden.

11. Statistische Programme wie z. B. SPSS, SAS usw. geben üblicherweise den p-Wert an oder eine Wahrscheinlichkeit, aus der sich der p-Wert berechnen lässt, z. B. $\tilde{p}$ mit $\chi^2_{J-1,\tilde{p}} = t$.

12. Falls H_0 richtig ist, dann ist p die Wahrscheinlichkeit, mit der die Prüfgröße T einen größeren Wert als den beobachteten Wert t hat,

$$p = P(T \geq t || H_0) \,.$$

Definition 18.b (Indikatorfunktion einer Menge) Für eine beliebige Menge $A \subset \mathbb{R}$ heißt die Funktion $\mathbf{1}_A : \mathbb{R} \to \{0, 1\}$,

$$\mathbb{1}_A(x) \stackrel{\text{def}}{=} \begin{cases} 1, & \text{falls} \quad x \in A \\ 0, & \text{falls} \quad x \notin A \end{cases}$$

Indikatorfunktion der Menge A.

Bemerkung 18.c (Zählen und Indikatorfunktion)

- $x_1, \ldots, x_n$ seien reelle Zahlen und $A \subset \mathbb{R}$ sei eine Menge. Dann ist

$$n_A = \sum_{i=1}^{n} \mathbb{1}_A(x_i)$$

  die Anzahl der Zahlen $x_1, \ldots, x_n$, die in der Menge A liegen.

- Es gilt $n_A \in \{0, 1, \ldots, n\}$.

- $X_1, X_2, \ldots, X_n$ seien zufällige Beobachtungen. Dann gibt die Zufallsvariable

$$N_A = \sum_{i=1}^{n} \mathbb{1}_A(X_i)$$

  die **zufällige Anzahl von Beobachtungen in** A an.

Kapitel 19

Grundlagen der asymptotischen Statistik

19-1

19. Grundlagen der asymptotischen Statistik

19.1 Konsistenz von Schätzern
19.2 Schwaches Gesetz der großen Zahlen
19.3 Zentraler Grenzwertsatz der Statistik
19.4 Asymptotisch begründete Tests

19-2

Bemerkung 19.1

- In diesem Kapitel wird die theoretische Begründung für die zuvor angegebenen approximativ gültigen Verfahren geliefert.
- Dabei werden asymptotische Betrachtungen für einen über alle Grenzen wachsenden Stichprobenumfang ($n \to \infty$) angestellt.

19-3

19.1 Konsistenz von Schätzern

Bemerkung 19.2
Die Wahrscheinlichkeitsverteilung des Schätzers

$$\bar{X}_n \stackrel{\text{def}}{=} \frac{1}{n}\sum_{i=1}^{n} X_i$$

für den Parameter μ konzentriert sich für $n \to \infty$ zunehmend um μ. Genauer ist dies eine Eigenschaft der Folge von Schätzern

$$\bar{X}_1 = X_1,\ \bar{X}_2 = \frac{X_1 + X_2}{2}, \bar{X}_3 = \frac{X_1 + X_2 + X_3}{3}, \ldots, \bar{X}_n = \frac{1}{n}\sum_{i=1}^{n} X_i, \ldots$$

Bemerkung 19.3
Die Konsistenz ist eine **asymptotische Eigenschaft** eines Schätzers $\hat{\theta}_n = \hat{\theta}_n(X_1, \ldots, X_n)$ für einen Parameter θ.

19-4

Bemerkung 19.4
Da die Schätzer $\hat{\theta}_n$ Zufallsvariablen sind, muss zunächst definiert werden, was unter der Konvergenz einer Folge von Zufallsvariablen (Schätzern) gegen eine Konstante (Parameter) verstanden werden soll.

Definition 19.5 (Konvergenz nach Wahrscheinlichkeit)
Eine Folge von Zufallsvariablen $Y_1, Y_2, \ldots$ **konvergiert nach Wahrscheinlichkeit** gegen eine Zahl $x \in \mathbb{R}$ genau dann, wenn

$$\lim_{n\to\infty} P(|Y_n - x| \geq \varepsilon) = 0 \quad \forall \varepsilon > 0. \tag{19.1}$$

Man schreibt

$$\operatorname*{p\,lim}_{n\to\infty} Y_n = x\,.$$

Bemerkung 19.6
Die Konvergenzaussage aus Gleichung (19.1) ist äquivalent zu

$$\lim_{n\to\infty} P(|Y_n - x| < \varepsilon) = 1 \quad \forall \varepsilon > 0.$$

19-5

Bemerkung 19.7
Die **Konvergenz nach Wahrscheinlichkeit** wird auch **Konvergenz in Wahrscheinlichkeit** oder **stochastische Konvergenz** genannt.

Bemerkung 19.8
Es gibt auch andere Konvergenzarten für Zufallsvariablen, z. B. die Verteilungskonvergenz oder die fast sichere Konvergenz.

19-6

Definition 19.9 (Konsistenz)
Eine Folge von Schätzern

$$\hat{\theta}_1 = \hat{\theta}_1(X_1), \hat{\theta}_2 = \hat{\theta}_2(X_1, X_2), \dots, \hat{\theta}_n = \hat{\theta}_n(X_1, \dots, X_n), \dots$$

heißt **konsistent** für θ, falls $\hat{\theta}_n$ nach Wahrscheinlichkeit gegen θ konvergiert.

Beispiel 19.10 (Konsistente Schätzer)

1. Für stochastisch unabhängige Zufallsvariablen X_i mit $\mathbb{E}[X_i] = \mu$ gilt: $\bar{X}_n$ ist ein konsistenter Schätzer für μ.
2. Für stochastisch unabhängige Zufallsvariablen X_i mit $\mathbb{V}[X_i] = \sigma^2$ gilt:
   (a) S^2 ist ein konsistenter Schätzer für σ^2.
   (b) S^{*2} ist ein konsistenter Schätzer für σ^2.
   (c) S ist ein konsistenter Schätzer für σ.
   (d) S^* ist ein konsistenter Schätzer für σ.

19-7

Bemerkung 19.11
Eine hinreichende Bedingung für die Konsistenz ist

$$\lim_{n\to\infty} \mathbb{E}[\hat{\theta}_n] = \theta, \quad \lim_{n\to\infty} \mathbb{V}[\hat{\theta}_n] = 0.$$

Beispiel 19.12
Für stochastisch unabhängige Zufallsvariablen X_i mit $\mathbb{E}[X_i] = \mu$ und $\mathbb{V}[X_i] = \sigma^2$ gilt

$$\mathbb{E}[\bar{X}_n] = \mu \quad \text{und} \quad \lim_{n\to\infty} \mathbb{V}[\bar{X}_n] = \lim_{n\to\infty} \frac{\sigma^2}{n} = 0,$$

so dass $\bar{X}_n$ ein konsistenter Schätzer für μ ist. Dies ist ein Spezialfall des sogenannten **Schwachen Gesetzes der großen Zahlen**.

Definition 19.13 (Asymptotische Erwartungstreue)
Die Eigenschaft

$$\lim_{n\to\infty} \mathbb{E}[\hat{\theta}_n] = \theta,$$

nennt man asymptotische Unverzerrtheit oder asymptotische Erwartungstreue des Schätzers $\hat{\theta}_n$ für θ.

19-8

19.2 Schwaches Gesetz der großen Zahlen

Beispiel 19.14 (Idealer Würfel)
Stochastisches Modell des wiederholten Würfelwurfes: Die X_i sind stochastisch unabhängig und identisch verteilt (i. i. d.) mit

$$P(X_i = k) = \frac{1}{6}, \quad k = 1, ..., 6$$

und

$$\mu = \mathbb{E}[X_i] = 3.5\,.$$

In welchem Sinn gilt

$$\bar{x}_n \overset{\text{def}}{=} \frac{1}{n}\sum_{i=1}^{n} x_i \to 3.5 \quad \text{für} \quad n \to \infty$$

für eine gegebene Folge von Beobachtungen $x_1, x_2, x_3, ..., x_n, ...$ mit $x_i \in \{1, 2, 3, 4, 5, 6\}$?

19-9

Satz 19.15
Die Zufallsvariablen $X_1, X_2, \ldots, X_n, \ldots$ *seien identisch verteilt und stochastisch unabhängig mit* $\mu \overset{\text{def}}{=} \mathbb{E}[X_i]$ *und endlicher Varianz* $\sigma^2 \overset{\text{def}}{=} \mathbb{V}[X_i]$*. Für die Mittelwerte*

$$\bar{X}_1 \overset{\text{def}}{=} X_1, \quad \bar{X}_2 \overset{\text{def}}{=} \frac{1}{2}\sum_{i=1}^{2} X_i, \quad \ldots, \quad \bar{X}_n \overset{\text{def}}{=} \frac{1}{n}\sum_{i=1}^{n} X_i, \ldots$$

gilt dann

$$\mathbb{E}[\bar{X}_n] = \mu, \quad \mathbb{V}[\bar{X}_n] = \frac{\sigma^2}{n}$$

und

$$\lim_{n\to\infty} \mathbb{V}[\bar{X}_n] = 0.$$

19-10

Bemerkung 19.16
Für wachsenden Stichprobenumfang konzentriert sich also die Verteilung von $\bar{X}_n$ um μ. Dies gilt allgemeiner, nämlich auch ohne die Voraussetzung einer endlichen Varianz.

Satz 19.17 (Schwaches Gesetz der großen Zahlen)
Die X_i *(*$i = 1, 2, \ldots$*) seien i. i. d. mit Erwartungswert* $\mu \overset{\text{def}}{=} \mathbb{E}[X_i]$*. Dann konvergiert* $\bar{X}_n$ *nach Wahrscheinlichkeit gegen* μ*.*

Bemerkung 19.18
Es gibt auch Wahrscheinlichkeitsverteilungen mit unendlicher Varianz, die z. B. im Bereich der Finanzmarktstochastik eine gewisse Rolle spielen.

Satz 19.19 (Spezialfall von Bernoulli)
Es sei $X_i \overset{i.i.d.}{\sim} B(1, \pi)$*. Dann konvergiert* $\bar{X}_n$ *nach Wahrscheinlichkeit gegen* π*.*

19-11

Bemerkung 19.20
Für $X_i \overset{i.i.d.}{\sim} B(1, \pi)$ gilt:

- $\pi = P(X_i = 1) = \mathbb{E}[X_i]$ ist die Erfolgswahrscheinlichkeit.
- $\sum_{i=1}^{n} X_i \sim B(n, \pi)$ ist die Anzahl (= absolute Häufigkeit) der Erfolge.
- $\bar{X}_n$ ist die relative Häufigkeit der Erfolge.
- Die relative Häufigkeit konvergiert nach Wahrscheinlichkeit gegen die Wahrscheinlichkeit π.

19-12

19.3 Zentraler Grenzwertsatz der Statistik

Bemerkung 19.21 (Verteilung von $\bar{X}_n$)
Für $X_i \overset{i.i.d.}{\sim} N(\mu, \sigma^2)$ gilt

$$\bar{X}_n \sim N\left(\mu, \frac{\sigma^2}{n}\right), \qquad \frac{\bar{X}_n - \mu}{\sigma}\sqrt{n} \sim N(0,1).$$

Satz 19.22 (Zentraler Grenzwertsatz der Statistik)
Die X_i $(i = 1, 2, \ldots)$ seien i. i. d. mit Varianz $0 < \sigma^2 \overset{\text{def}}{=} \mathbb{V}[X_i] < \infty$ und Erwartungswert $\mu \overset{\text{def}}{=} \mathbb{E}[X_i]$. Für die standardisierten Zufallsvariablen

$$Y_n \overset{\text{def}}{=} \frac{\bar{X}_n - \mathbb{E}[\bar{X}_n]}{\sqrt{\mathbb{V}[\bar{X}_n]}} = \frac{\bar{X}_n - \mu}{\sigma}\sqrt{n} \tag{19.2}$$

gilt dann

$$\lim_{n\to\infty} P(Y_n \le x) = \Phi(x) \quad \textit{für alle} \quad x \in \mathbb{R}. \tag{19.3}$$

19-13

Bemerkung 19.23

- Φ bezeichnet die Verteilungsfunktion der Standardnormalverteilung $N(0,1)$.
- $P(Y_n \le x)$ ist der Wert der Verteilungsfunktion von Y_n an der Stelle x.
- Wenn (19.3) erfüllt ist, sagt man auch: Y_n **besitzt (für $n \to \infty$) die Standardnormalverteilung als Grenzverteilung**.
- Eine bessere Bezeichnung wäre „Grenzverteilungssatz“.
- Wenn (19.2) und (19.3) erfüllt sind, sagt man auch: **Die Folge $\bar{X}_n$ ist asymptotisch normalverteilt**.
- Für "große" n wird Y_n als **näherungsweise** $N(0,1)$-verteilt und $\bar{X}_n$ als **näherungsweise** $N\left(\mu, \frac{\sigma^2}{n}\right)$-verteilt unterstellt.

19-14

Beispiel 19.24 (Würfel)
Stochastisches Modell: Für $i = 1, \ldots, n$ sind die X_i i.i.d. mit

$$P(X_i = k) = \frac{1}{6}, \quad k = 1, ..., 6.$$

Es gilt

$$\mu = \mathbb{E}[X_i] = \frac{1}{6}(1+2+3+4+5+6) = \frac{7}{2},$$
$$\mathbb{E}[X_i^2] = \frac{1}{6}(1+2^2+3^2+4^2+5^2+6^2) = \frac{91}{6}$$

und

$$\sigma^2 = \mathbb{V}[X_i] = \frac{91}{6} - \left(\frac{7}{2}\right)^2 = \frac{35}{12}\,.$$

19-15

Beispiel 19.25 (Fortsetzung)
Allgemein:

$$\mathbb{V}[\bar{X}_n] = \frac{\sigma^2}{n}, \quad \sigma[\bar{X}_n] = \sqrt{\mathbb{V}[\bar{X}_n]}$$

Für $n = 600$:

$$\mathbb{V}[\bar{X}_{600}] = \frac{1}{600} \cdot \frac{35}{12} = \frac{35}{7200} = 0.00486, \quad \sigma[\bar{X}_{600}] = 0.0697$$

$\bar{X}_{600}$ besitzt approximativ die Verteilung $N(\mu_{\bar{X}_{600}}, \sigma^2_{\bar{X}_{600}})$ mit

$$\mu_{\bar{X}_{600}} = 3.5, \quad \sigma^2_{\bar{X}_{600}} = 0.00486.$$

$$\mu_{\bar{X}_{600}} - 2\sigma_{\bar{X}_{600}} = 3.36, \quad \mu_{\bar{X}_{600}} + 2\sigma_{\bar{X}_{600}} = 3.64$$

$$\mu_{\bar{X}_{600}} - 3\sigma_{\bar{X}_{600}} = 3.29, \quad \mu_{\bar{X}_{600}} + 3\sigma_{\bar{X}_{600}} = 3.71$$

19-16

19.4 Asymptotisch begründete Tests

Bemerkung 19.26 (Approximative Gauß-Tests) Es wird die theoretische Begründung für die in den Bemerkungen 18.2 und 18.3 vorgestellten approximativen Gauß-Tests nachgeliefert.

- Nach dem zentralen Grenzwertsatz der Statistik (vgl. Satz 19.22) hat die Prüfgröße

$$T = \frac{\bar{X}_n - \mu_0}{\sigma} \sqrt{n} \tag{19.4}$$

  aus Bemerkung 18.2 unter $H_0 : \mu = \mu_0$ für $n \to \infty$ die Standardnormalverteilung als Grenzverteilung.

19-17

- Falls H_0 richtig ist, gilt

$$\begin{aligned} P(T \in K) &= 1 - P\left(u_{\frac{\alpha}{2}} \le T \le u_{1-\frac{\alpha}{2}}\right) \\ &= 1 - \left(P\left(T \le u_{1-\frac{\alpha}{2}}\right) - P\left(T < u_{\frac{\alpha}{2}}\right)\right) \\ &= 1 - P\left(T \le u_{1-\frac{\alpha}{2}}\right) + P\left(T < u_{\frac{\alpha}{2}}\right) \end{aligned}$$

  und daher

$$\begin{aligned} \lim_{n\to\infty} P(T \in K) &= \lim_{n\to\infty} \left[1 - P\left(T \le u_{1-\frac{\alpha}{2}}\right) + P\left(T < u_{\frac{\alpha}{2}}\right)\right] \\ &= 1 - \lim_{n\to\infty} P\left(T \le u_{1-\frac{\alpha}{2}}\right) + \lim_{n\to\infty} P\left(T < u_{\frac{\alpha}{2}}\right) \\ &= 1 - \Phi\left(u_{1-\frac{\alpha}{2}}\right) + \Phi\left(u_{\frac{\alpha}{2}}\right) \\ &= 1 - \left(1 - \frac{\alpha}{2}\right) + \frac{\alpha}{2} \\ &= \alpha\,. \end{aligned}$$

19-18

- Die modifizierte Prüfgröße

$$T = \frac{\bar{X} - \mu_0}{S}\sqrt{n} \tag{19.5}$$

aus Bemerkung 18.3, bei der σ durch den konsistenten Schätzer S ersetzt ist, wird durch

$$T = \frac{\bar{X} - \mu_0}{\sigma}\sqrt{n} \cdot \frac{\sigma}{S}$$

auf die Prüfgröße aus (19.4) und den Faktor σ/S zurückgeführt.

- Zusammen mit einem bekannten Hilfssatz der Statistik, dem **Lemma von Slutzky**, ergibt sich, dass die Prüfgröße aus (19.5) dieselbe Grenzverteilung wie die Prüfgröße aus (19.4) besitzt, so dass sich ebenfalls

$$\lim_{n\to\infty} P(T \in K) = \alpha$$

ergibt.

19.5 Ergänzungen

Bemerkung 19.a (Zur Selbstkontrolle) Im Kapitel 19 eingeführte Begriffe und Konzepte: Konsistenz, schwaches Gesetz der großen Zahlen, asymptotische Erwartungstreue, zentraler Grenzwertsatz der Statistik.

Bemerkung 19.b (Zum Zwei-Stichproben-Gauß-Test) Falls $H_0 : \mu_X = \mu_Y$ richtig ist, ist die Prüfgröße

$$T = \frac{\bar{X} - \bar{Y}}{\sqrt{\frac{\sigma_X^2}{n} + \frac{\sigma_Y^2}{m}}}$$

aus Bemerkung 18.5 standardisiert mit $\mathbb{E}[T] = 0$ und $\mathbb{V}[T] = 1$ und besitzt für $n \to \infty$ die Standardnormalverteilung als Grenzverteilung. Diese Grenzverteilung bleibt erhalten, wenn die Parameter σ_X^2 und σ_Y^2 unbekannt sind und wie in Bemerkung 18.6 durch die konsistenten Schätzer S_X^2 und S_Y^2 ersetzt werden.

Bemerkung 19.c (Zum Test für eine Wahrscheinlichkeit) Die X_i sind i. i. d. mit

$$\mathbb{E}[X_i] = \mu = \pi, \quad \mathbb{V}[X_i] = \sigma^2 = \pi(1-\pi).$$

Die Prüfgröße

$$\frac{\bar{X} - \pi_0}{\sqrt{\pi_0(1-\pi_0)}}\sqrt{n}$$

in Bemerkung 18.7 besitzt für $n \to \infty$ die Standardnormalverteilung als Grenzverteilung, falls $H_0 : \pi = \pi_0$.

Bemerkung 19.d (Test zum Vergleich zweier Wahrscheinlichkeiten) Unter den Voraussetzungen von Bemerkung 18.8 ist

$$\frac{\hat{\pi}_X - \hat{\pi}_Y}{\sqrt{\frac{\pi_X(1-\pi_X)}{n} + \frac{\pi_Y(1-\pi_Y)}{m}}},$$

falls $H_0 : \pi_X = \pi_Y$ richtig ist, asymptotisch standardnormalverteilt. Dieses asymptotische Verhalten bleibt erhalten, wenn

$$\frac{\pi_X(1-\pi_X)}{n} + \frac{\pi_Y(1-\pi_Y)}{m}$$

durch den konsistenten Schätzer

$$\frac{\hat{\pi}_X(1-\hat{\pi}_X)}{n} + \frac{\hat{\pi}_Y(1-\hat{\pi}_Y)}{m}$$

ersetzt wird.

Bemerkung 19.e (Zum χ^2-Anpassungstest) Die Prüfgröße T beim χ^2-Anpassungstest aus Bemerkung 18.9 besitzt, falls H_0 richtig ist, für $n \to \infty$ eine χ^2-Verteilung mit $J-1$ Freiheitsgraden als Grenzverteilung. Das Signifikanzniveau α beschränkt damit (asymptotisch) den Fehler 1. Art bezüglich der Nullhypothese

$$H_0 : P(X \in A_j) = \pi_j \quad \text{für} \quad j = 1, \dots, J\,.$$

Wenn diese Nullhypothese abgelehnt wird, wird erst recht die Nullhypothese

$$H_0 : F = F_0$$

abgelehnt, die $P(X \in A_j) = \pi_j$ für $j = 1, \dots, J$ impliziert.

Bemerkung 19.f Die beiden folgenden Sätze liefern die theoretische Begründung für die in den Bemerkung 13.g angegebenen approximativen Konfidenzintervalle.

Satz 19.g (σ^2 bekannt) *Die X_i seien i.i.d. mit $0 < \sigma^2 \overset{\text{def}}{=} \mathbb{V}[X_i] < \infty$. Dann ist*

$$I_n \overset{\text{def}}{=} \left[\bar{X} - u_{1-\frac{\alpha}{2}} \frac{\sigma}{\sqrt{n}}, \bar{X} + u_{1-\frac{\alpha}{2}} \frac{\sigma}{\sqrt{n}}\right]$$

ein Konfidenzintervall für μ mit asymptotischem Konfidenzniveau $(1-\alpha)$, d. h. es gilt

$$\lim_{n\to\infty} P(\mu \in I_n) = 1 - \alpha.$$

Beweis Der zentrale Grenzwertsatz der Statistik (Satz 19.22) besagt, dass die standardisierten Zufallsvariablen

$$Y_n = \frac{\bar{X}_n - \mu}{\sigma}\sqrt{n}$$

die Standardnormalverteilung als Grenzverteilung besitzen. Daher gilt

$$\begin{aligned}
\lim_{n\to\infty} P(u_{\frac{\alpha}{2}} \le Y_n \le u_{1-\frac{\alpha}{2}}) &= \lim_{n\to\infty} (P(Y_n \le u_{1-\frac{\alpha}{2}}) - P(Y_n < u_{\frac{\alpha}{2}}) \\
&= \lim_{n\to\infty} P(Y_n \le u_{1-\frac{\alpha}{2}}) - \lim_{n\to\infty} P(Y_n < u_{\frac{\alpha}{2}}) \\
&= \Phi(u_{1-\frac{\alpha}{2}}) - \Phi(u_{\frac{\alpha}{2}}) \\
&= 1 - \frac{\alpha}{2} - \frac{\alpha}{2} \\
&= 1 - \alpha.
\end{aligned}$$

Weiter gilt

$$\begin{aligned}
P(u_{\frac{\alpha}{2}} \le Y_n \le u_{1-\frac{\alpha}{2}}) &= P\left(u_{\frac{\alpha}{2}} \le \frac{\bar{X}_n - \mu}{\sigma}\sqrt{n} \le u_{1-\frac{\alpha}{2}}\right) \\
&= P\left(u_{\frac{\alpha}{2}} \frac{\sigma}{\sqrt{n}} \le \bar{X}_n - \mu \le u_{1-\frac{\alpha}{2}} \frac{\sigma}{\sqrt{n}}\right) \\
&= P\left(u_{\frac{\alpha}{2}} \frac{\sigma}{\sqrt{n}} - \bar{X}_n \le -\mu \le u_{1-\frac{\alpha}{2}} \frac{\sigma}{\sqrt{n}} - \bar{X}_n\right) \\
&= P\left(\bar{X}_n - u_{\frac{\alpha}{2}} \frac{\sigma}{\sqrt{n}} \ge \mu \ge \bar{X}_n - u_{1-\frac{\alpha}{2}} \frac{\sigma}{\sqrt{n}}\right) \\
&= P\left(\bar{X}_n - u_{1-\frac{\alpha}{2}} \frac{\sigma}{\sqrt{n}} \le \mu \le \bar{X}_n - u_{\frac{\alpha}{2}} \frac{\sigma}{\sqrt{n}}\right) \\
&= P\left(\bar{X}_n - u_{1-\frac{\alpha}{2}} \frac{\sigma}{\sqrt{n}} \le \mu \le \bar{X}_n + u_{1-\frac{\alpha}{2}} \frac{\sigma}{\sqrt{n}}\right) \\
&= P\left(\mu \in I_n\right)
\end{aligned}$$

mit $-u_{\frac{\alpha}{2}} = u_{1-\frac{\alpha}{2}}$.

Satz 19.h (σ^2 unbekannt) *Die X_i seien i.i.d. mit $0 < \sigma^2 \overset{\text{def}}{=} \mathbb{V}[X_i] < \infty$. Dann ist*

$$I_n \overset{\text{def}}{=} \left[\bar{X} - u_{1-\frac{\alpha}{2}} \frac{S}{\sqrt{n}}, \bar{X} + u_{1-\frac{\alpha}{2}} \frac{S}{\sqrt{n}}\right] \quad \textit{mit} \quad S \overset{\text{def}}{=} \sqrt{\frac{1}{n}\sum_{i=1}^{n}(X_i - \bar{X})^2},$$

ein Konfidenzintervall für μ mit asymptotischem Konfidenzniveau $(1-\alpha)$, d. h. es gilt

$$\lim_{n\to\infty} P(\mu \in I_n) = 1 - \alpha.$$

Beweis Beim Beweis von Satz 19.h muss zusätzlich zu den Überlegungen aus dem vorangegangenen Beweis die Konvergenz von S gegen σ für $n \to \infty$ (Gesetz der großen Zahlen) berücksichtigt werden. Zusammen mit dem Lemma von Slutzky ergibt sich, dass die Zufallsvariablen

$$\frac{\bar{X}_n - \mu}{S}\sqrt{n} = \frac{\sigma}{S}\frac{\bar{X}_n - \mu}{\sigma}\sqrt{n} = \frac{\sigma}{S}Y_n$$

dieselbe Grenzverteilung wie Y_n, nämlich die Standardnormalverteilung, besitzen. Der Satz 19.h ergibt sich dann mit analogen Umformungen wie im Beweis von Satz 19.g.

Bemerkung 19.i (Bernoulli-Verteilung) Das bereits in Bemerkung 13.20 angegebene approximative Konfidenzintervall

$$\left[\hat{\pi} - u_{1-\frac{\alpha}{2}}\sqrt{\frac{\hat{\pi}(1-\hat{\pi})}{n}}, \hat{\pi} + u_{1-\frac{\alpha}{2}}\sqrt{\frac{\hat{\pi}(1-\hat{\pi})}{n}}\right]$$

für den Erfolgsparameter π einer Bernoulli-Verteilung ergibt sich als Spezialfall des Konfidenzintervalls aus Satz 19.h, wenn man berücksichtigt, dass sich

$$S^2 = \frac{1}{n}\sum_{i=1}^{n}(X_i - \bar{X})^2 = \bar{X}(1-\bar{X}) = \hat{\pi}(1-\hat{\pi})$$

aus $X_i \overset{i.i.d.}{\sim} B(1, \pi)$ ergibt.

Kapitel 20

Bedingte Wahrscheinlichkeit und bedingte Verteilung

20-1

20. Bedingte Wahrscheinlichkeit und bedingte Verteilung

20.1 Bedingte Wahrscheinlichkeiten
20.2 Bedingte Verteilung

20-2

20.1 Bedingte Wahrscheinlichkeiten

Definition 20.1 (Bedingte Wahrscheinlichkeit)
Es sei B ein Ereignis mit $P(B) > 0$. Dann heißt

$$P(A|B) \stackrel{\text{def}}{=} \frac{P(A \cap B)}{P(B)}$$

die **bedingte Wahrscheinlichkeit** (*conditional probability*) **von** A **unter der Bedingung** B (oder **unter der Hypothese** B).

Beispiel 20.2 (Würfelwurf)
Für die Ereignisse $A = \{\omega_1, \omega_3, \omega_5\}$ („ungerade Augenzahl") und $B = \{\omega_1, \omega_2, \omega_3\}$ („Augenzahl ≤ 3") gilt

$$P(A|B) = \frac{P(A \cap B)}{P(B)} = \frac{P(\{\omega_1, \omega_3\})}{\frac{1}{2}} = \frac{\frac{1}{3}}{\frac{1}{2}} = \frac{2}{3}.$$

20-3

Satz 20.3
A und B seien unabhängig. Es sei $P(B) > 0$. Dann gilt

$$P(A|B) = P(A).$$

Bemerkung 20.4
Wegen der Unabhängigkeit gilt $P(A \cap B) = P(A)P(B)$. Somit gilt

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

20-4

Satz 20.5
Es gelte $P(B) > 0$ und $P(A|B) = P(A)$. Dann sind A und B unabhängig.

Bemerkung 20.6
Aus

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

folgt $P(A \cap B) = P(A)P(B)$ und somit die Unabhängigkeit der Ereignisse A und B.

20-5

Satz 20.7 (Multiplikationssatz für zwei Ereignisse)
Es sei B ein Ereignis mit $P(B) > 0$. Dann gilt

$$P(A \cap B) = P(A|B) \cdot P(B)\,.$$

Bemerkung 20.8
Die Aussage von Satz 20.7 ergibt sich unmittelbar durch Umformen der Definition der bedingten Wahrscheinlichkeit.

Satz 20.9 (Multiplikationssatz für drei Ereignisse)
Es seien B und C Ereignisse mit $P(B \cap C) > 0$. Dann gilt

$$P(A \cap B \cap C) = P(A|B \cap C) \cdot P(B|C) \cdot P(C)\,.$$

Bemerkung 20.10
Die zweimalige Anwendung von Satz 20.7 ergibt

$$P(A \cap B \cap C) = P(A|B \cap C) \cdot P(B \cap C) = P(A|B \cap C) \cdot P(B|C) \cdot P(C)\,.$$

Dabei ist $P(C) > 0$ wegen $P(B \cap C) > 0$.

20-6

Satz 20.11 (Satz von der totalen Wahrscheinlichkeit)
Bilden die Ereignisse $A_1, \dots, A_n$ eine vollständige Zerlegung von Ω mit $P(A_i) > 0$ für $i = 1, \dots, n$, dann gilt

$$P(B) = \sum_{i=1}^{n} P(B|A_i)P(A_i)$$

für jedes Ereignis B.

Beispiel 20.12 (Würfel)
Für die drei Ereignisse $A_1 = \{\omega_1, \omega_3, \omega_5\}$ („ungerade Augenzahl“),
$A_2 = \{\omega_2, \omega_4, \omega_6\}$ („gerade Augenzahl“) und
$B = \{\omega_1, \omega_2, \omega_3\}$ („Augenzahl ≤ 3“) gilt

$$P(B) = P(B|A_1)\,P(A_1) + P(B|A_2)\,P(A_2)$$

bzw.

$$\frac{1}{2} = \frac{2}{3}\,\frac{1}{2} + \frac{1}{3}\,\frac{1}{2}\,.$$

20-7

Bemerkung 20.13
Häufig interessiert die Berechnung der Wahrscheinlichkeiten $P(A_i|B)$ aus gegebenen Wahrscheinlichkeiten $P(B|A_i)$ und $P(A_i)$.

Satz 20.14 (Bayessches Theorem, Formel von Bayes)
Bilden die Ereignisse $A_1, \dots, A_n$ eine vollständige Zerlegung von Ω mit $P(A_i) > 0$ für $i = 1, \dots, n$ und ist B ein Ereignis mit $P(B) > 0$, dann gilt

$$P(A_i|B) = \frac{P(B|A_i)P(A_i)}{\sum\limits_{j=1}^{n} P(B|A_j)P(A_j)}, \quad i = 1, \dots, n\,.$$

Beispiel 20.15 (Würfel: Bayessches Theorem)
Aus $P(A_1) = P(A_2) = 1/2$, $P(B|A_1) = 2/3$ und $P(B|A_2) = 1/3$ folgt

$$P(A_2|B) = \frac{P(B|A_2)\,P(A_2)}{P(B|A_1)\,P(A_1) + P(B|A_2)\,P(A_2)} = \frac{1}{3}\,.$$

20-8

20.2 Bedingte Verteilung

Definition 20.16
f_{XY} bezeichne die Dichtefunktion oder Wahrscheinlichkeitsfunktion von (X,Y). f_X und f_Y seien die X und Y gehörigen Wahrscheinlichkeits- bzw. Dichtefunktionen.

- Falls f_Y an einer Stelle $y \in \mathbb{R}$ positiv ist, heißt die Funktion $f_{X|Y=y} : \mathbb{R} \to [0, \infty[$,

  $$f_{X|Y=y}(x) = \frac{f_{XY}(x,y)}{f_Y(y)}$$

  bedingte Dichtefunktion bzw. **bedingte Wahrscheinlichkeitsfunktion** von X unter der Bedingung $Y = y$.

- Analog ist $f_{Y|X=x}(y)$ für x mit $f_X(x) > 0$ definiert.

20.3 Ergänzungen

Bemerkung 20.a (Zur Selbstkontrolle) Im Kapitel 20 eingeführte Begriffe und Konzepte: bedingte Wahrscheinlichkeit, Satz von der totalen Wahrscheinlichkeit, Bayessches Theorem (Formel von Bayes), bedingte Dichtefunktion, bedingte Wahrscheinlichkeitsfunktion.

Beispiel 20.b (Gemeinsame und bedingte Verteilung)

1. Merkmal X: Körpergröße (klassiert)
   x_1 bis unter 165 cm
   x_2 über 165 cm

2. Merkmal Y: Körpergewicht (klassiert)
   y_1 bis unter 50 kg
   y_2 über 50 bis unter 75 kg
   y_3 über 75 kg

3. Die Anteile der **Merkmalskombinationen** (x_j, y_k) in der Grundgesamtheit werden zu Wahrscheinlichkeiten, wenn eine Person zufällig ausgewählt wird und an dieser die beiden Merkmale X und Y gemessen werden.

4. Tabelle der gemeinsamen Wahrscheinlichkeiten und der Randwahrscheinlichkeiten:

$P(X = x_j, Y = y_k)$	y_1	y_2	y_3	$\sum_k$
x_1	0.25	0.15	0.10	0.50
x_2	0.05	0.25	0.20	0.50
$\sum_j$	0.30	0.40	0.30	1.00

   Beispiel für Randwahrscheinlichkeit:

$$P(X = x_2) = \sum_{k=1}^{3} P(X = x_2, Y = y_k) = 0.05 + 0.25 + 0.20 = 0.50$$

   Beispiel für bedingte Wahrscheinlichkeit:

$$P(Y = y_2 | X = x_1) = \frac{P(X = x_1, Y = y_2)}{P(X = x_1)} = \frac{0.15}{0.50} = 0.30$$

5. **Randwahrscheinlichkeitsfunktion** von X:

$$f_X(x_1) = P(X = x_1) = 0.5,$$

$$f_X(x_2) = P(X = x_2) = 0.5,$$

$$f_X(x) = 0, \quad \text{falls} \quad x \notin \{x_1, x_2\}$$

   Analog gibt es eine Randwahrscheinlichkeitsfunktion von Y.

6. **Bedingte Wahrscheinlichkeitsfunktion** von Y bedingt auf $X = x_1$:

$$f_{Y|X=x_1}(y_1) = P(Y = y_1 | X = x_1) = 0.50,$$

$$f_{Y|X=x_1}(y_2) = P(Y = y_2 | X = x_1) = 0.30,$$

$$f_{Y|X=x_1}(y_3) = P(Y = y_3 | X = x_1) = 0.20,$$

$$f_{Y|X=x_1}(y) = P(Y = y | X = x_1) = 0, \quad \text{falls} \quad y \notin \{y_1, y_2, y_3\}$$

   Analog gibt es eine bedingte Wahrscheinlichkeitsfunktion von Y bedingt auf $X = x_2$ und bedingte Wahrscheinlichkeitsfunktionen von X bedingt auf $Y = y$ für $y \in \{y_1, y_2, y_3\}$.

Beispiel 20.c (DNA-Test als Massenscreening) Ausgangssituation: Vom unbekannten Täter liegt eine DNA-Spur vor. Für Person P wird eine DNA-Analyse auf Übereinstimmung durchgeführt.

- Ereignis A: Die DNA-Spur stammt von P. Ereignis $\bar{A}$: Die DNA-Spur stammt **nicht** von P.

- Ereignis B: „Bei der DNA-Analyse wird Übereinstimmung der DNA von P mit der DNA-Spur festgestellt“ (sogenannter positiver Befund).
- Die Wahrscheinlichkeit für einen richtigen positiven Befund ist

$$P(B|A) = p, \quad \text{z. B. } p = 99.9999\% \, .$$

- Die Wahrscheinlichkeit für einen falschen positiven Befund ist

$$P(B|\bar{A}) = \beta, \quad \text{z. B. } \beta = 10^{-9} = \frac{1}{1000000000} \, .$$

- Es gilt nicht notwendig $p + \beta = 1$.
- Wie groß ist die Wahrscheinlichkeit $P(A|B)$?

$$\begin{aligned} P(A|B) &= \frac{P(B|A)P(A)}{P(B|A)P(A) + P(B|\bar{A})P(\bar{A})} \\ &= \frac{pP(A)}{pP(A) + \beta(1 - P(A))} \end{aligned}$$

- Mit der sogenannten A-priori-Wahrscheinlichkeit

$$P(A) = \frac{1}{n} = \frac{1}{100000}$$

bei $n = 100000 = 10^5$ Verdächtigten ergibt sich

$$P(A|B) = \frac{p\frac{1}{n}}{p\frac{1}{n} + \beta(1 - \frac{1}{n})} = \frac{p}{p + (n-1)\beta} = \frac{p}{p + 99999\beta} \, .$$

Die Größenordnung von β im Vergleich zu n ist entscheidend für den Rückschluss von einem positiven Befund (B) auf die Täterzuordnung, d. h. für die Größe der Wahrscheinlichkeit $P(A|B)$:

- Für $\beta = 0$ ergibt sich $P(A|B) = 1$.
- Für $\beta = 10^{-9}$ und $n = 10^5$ ergibt sich $n\beta = 10^{-4}$ und

$$P(A|B) = \frac{p}{p + 10^{-4} - 10^{-9}} \approx 1.$$

- Für $\beta = 10^{-5}$ und $n = 10^5$ ergibt sich $n\beta = 1$ und

$$P(A|B) = \frac{p}{p + 1 - 10^{-5}} \approx \frac{1}{2} \, .$$

- Für $\beta = 10^{-3}$ und $n = 10^5$ ergibt sich $n\beta = 100$ und

$$P(A|B) \approx \frac{p}{p + 100 - 10^{-3}} \approx \frac{1}{101} \, .$$

- Für weiterführende Literatur zur Thematik und zur Größenordnung von p und β siehe

`http://www.decisions.ch/dissertation.html`,

insbesondere S. 169-177, und die dort angegebenen Quellen.

Anhang A

Formelsammlung

A.1 Allgemeine mathematische Notation

Symbol	Bedeutung	Bemerkung, Beispiel
$\mathbb{R}$	Menge der reellen Zahlen	
$\mathbb{N}$	Menge der natürlichen Zahlen	$\mathbb{N} = \{1, 2, \ldots\}$
$\stackrel{\text{def}}{=}$	Definitorisches Gleichheitszeichen	$\mathbb{N} \stackrel{\text{def}}{=} \{1, 2, \ldots\}$
$[x]$	Ganzzahliger Teil einer Zahl x	$\left[\frac{15}{4}\right] = [3.75] = 3$
$\|A\|$	Anzahl der Elemente einer Menge A	$\|\{4, 5, 7\}\| = 3$
$\sum_{i=1}^{n} x_i$	$x_1 + x_2 + \ldots + x_n$	$\sum_{i=1}^{0} x_i = 0$
$\prod_{i=1}^{n} x_i$	$x_1 \cdot x_2 \cdot \ldots \cdot x_n$	$\prod_{i=1}^{0} x_i = 1$

A.2 Beschreibende Statistik

Merkmale, Daten, Auswertung eindimensionaler Daten

Für ein statistisches Merkmal X mit J verschiedenen Merkmalswerten $\xi_1, \xi_2, \ldots, \xi_J$ liegt eine Urliste von n Daten (Beobachtungen) $x_1, x_2, \ldots, x_n$ vor.

Symbol, Definition	Bedeutung	Eigenschaften, Bemerkung
$n_j = \|\{i \mid x_i = \xi_j\}\|$	Absolute Häufigkeiten	$\sum_{j=1}^{J} n_j = n$
$f_j = \frac{n_j}{n}$	Relative Häufigkeiten	$\sum_{j=1}^{J} f_j = 1$
$F(x) = \frac{\|\{i \mid x_i \leq x\}\|}{n}, \quad x \in \mathbb{R}$	(Empirische) Verteilungsfunktion	$F(x) = \sum_{j:\xi_j \leq x} f_j$
$\tilde{x}_p = \min\{x \in \mathbb{R} \mid F(x) \geq p\}$, $0 < p < 1$	p-Quantil	Für $x_1 \leq x_2 \leq \ldots \leq x_n$: $\tilde{x}_p = \begin{cases} x_{np} & np \text{ ganzzahlig} \\ x_{[np]+1} & \text{sonst} \end{cases}$
$\tilde{x}_{0.5}$	Median	0.5-Quantil

$\bar{x} = \frac{1}{n}\sum_{i=1}^{n} x_i$	Arithmetisches Mittel	$\bar{x} = \frac{1}{n}\sum_{j=1}^{J} \xi_j n_j = \sum_{j=1}^{J} \xi_j f_j$
$\bar{x}_G = \sqrt[n]{\prod_{i=1}^{n} x_i}$	Geometrisches Mittel	$x_i > 0$ für $i = 1, \ldots, n$
$\bar{x}_w = \sum_{i=1}^{n} w_i x_i$	Gewichtetes Mittel	$w_i \geq 0$ für $i = 1, \ldots, n$; $\sum_{i=1}^{n} w_i = 1$
$s^2 = \frac{1}{n}\sum_{i=1}^{n}(x_i - \bar{x})^2 = \frac{1}{n}\left(\sum_{i=1}^{n} x_i^2\right) - \bar{x}^2$	Varianz	$s^2 = \frac{1}{n}\sum_{j=1}^{J}(\xi_j - \bar{x})^2 n_j = \sum_{j=1}^{J}(\xi_j - \bar{x})^2 f_j$
$s = \sqrt{s^2}$	Standardabweichung	

Verhältniszahlen, Meßzahlen und Indexzahlen

Symbol, Definition	Bedeutung, Eigenschaften
Meßzahlen des zeitlichen Vergleichs	
$m_{s,t} = \frac{x_t}{x_s}$	Meßzahl für die Berichtszeit t zur Basiszeit s
$m_{r,t} = \frac{m_{s,t}}{m_{s,r}}$	Umbasierung von Basiszeit s zu Basiszeit r
$m_{s,t} = m_{s,r} \cdot m_{r,t}$	Zirkularität von Meßzahlen
$m_{0,t} = m_{0,s} \cdot m_{s,t}, \quad t = s+1, \ldots, T$ $m_{s,t} = \frac{m_{0,t}}{m_{0,s}}, \quad t = 0, 1, \ldots, s-1$	Verkettung von $m_{0,t}$ für $t = 0, 1, \ldots, s$ und $m_{s,t}$ für $t = s, s+1, \ldots, T$
Indexzahlen	
t, 0	Berichtszeit (-periode), Basiszeit (-periode)
$p_t(i), q_t(i)$	Preis und Menge des Gutes i zur Zeit t
$I^p_{La;0,t} = \frac{\sum_{i=1}^{n} p_t(i)q_0(i)}{\sum_{i=1}^{n} p_0(i)q_0(i)} = \sum_{i=1}^{n} \frac{p_t(i)}{p_0(i)} \frac{p_0(i)q_0(i)}{\sum_{j=1}^{n} p_0(j)q_0(j)}$	Preisindex vom Typ Laspeyres
$I^p_{Pa;0,t} = \frac{\sum_{i=1}^{n} p_t(i)q_t(i)}{\sum_{i=1}^{n} p_0(i)q_t(i)}$	Preisindex vom Typ Paasche
$I^q_{La;0,t} = \frac{\sum_{i=1}^{n} p_0(i)q_t(i)}{\sum_{i=1}^{n} p_0(i)q_0(i)}$	Mengenindex vom Typ Laspeyres
$I^q_{Pa;0,t} = \frac{\sum_{i=1}^{n} p_t(i)q_t(i)}{\sum_{i=1}^{n} p_t(i)q_0(i)}$	Mengenindex vom Typ Paasche
$I^v_{0,t} = \frac{\sum_{i=1}^{n} p_t(i)q_t(i)}{\sum_{i=1}^{n} p_0(i)q_0(i)} = I^p_{Pa;0,t} I^q_{La;0,t} = I^p_{La;0,t} I^q_{Pa;0,t}$	Wertindex

Auswertung mehrdimensionaler Daten

Das Merkmal X hat die J verschiedenen Merkmalswerte $\xi_1, \ldots, \xi_J$. Das Merkmal Y hat die K verschiedenen Merkmalswerte $\eta_1, \ldots, \eta_K$. Gegeben sind n Beobachtungspaare (x_i, y_i) für $i = 1, \ldots, n$.

Symbol, Definition	Bedeutung, Eigenschaften
$\bar{x} = \frac{1}{n}\sum_{i=1}^{n} x_i, \quad \bar{y} = \frac{1}{n}\sum_{i=1}^{n} y_i$	Mittelwerte
$s_X^2 = \frac{1}{n}\sum_{i=1}^{n}(x_i - \bar{x})^2, \quad s_Y^2 = \frac{1}{n}\sum_{i=1}^{n}(y_i - \bar{y})^2$	Varianzen
$s_X = \sqrt{s_X^2}, \quad s_Y = \sqrt{s_Y^2}$	Standardabweichungen
Häufigkeiten für $j = 1, \ldots, J, k = 1, \ldots, K$	
$n_{jk} = \lvert\{i \mid (x_i, y_i) = (\xi_j, \eta_k)\}\rvert$	Gemeinsame absolute Häufigkeiten
$n_{j\cdot} = \sum_{k=1}^{K} n_{jk}, \quad n_{\cdot k} = \sum_{j=1}^{J} n_{jk}$	Absolute Randhäufigkeiten
$f_{jk} = \frac{n_{jk}}{n}$	Gemeinsame relative Häufigkeiten
$f_{j\cdot} = \sum_{k=1}^{K} f_{jk}, \quad f_{\cdot k} = \sum_{j=1}^{J} f_{jk}$	Relative Randhäufigkeiten
Zusammenhangsmaße	
$s_{XY} = \frac{1}{n}\sum_{i=1}^{n}(x_i - \bar{x})(y_i - \bar{y}) = \frac{1}{n}\sum_{i=1}^{n} x_i y_i - \bar{x}\bar{y}$	Kovarianz
$r_{XY} = \dfrac{\sum_{i=1}^{n}(x_i - \bar{x})(y_i - \bar{y})}{\sqrt{\sum_{i=1}^{n}(x_i - \bar{x})^2 \sum_{i=1}^{n}(y_i - \bar{y})^2}} = \dfrac{s_{XY}}{s_X s_Y}$	Korrelationskoeffizient (nach Bravais-Pearson)
$r_{XY}^R = 1 - \dfrac{6\sum_{i=1}^{n}\left(R_X(x_i) - R_Y(y_i)\right)^2}{n(n^2-1)}$	Rangkorrelationskoeffizient (nach Spearman)
$n_{jk} = \tilde{n}_{jk} = \dfrac{n_{j\cdot} n_{\cdot k}}{n}, \quad j = 1, \ldots, J, \quad k = 1, \ldots, K$	Deskriptive Unabhängigkeit
$\chi^2 = \sum_{j=1}^{J}\sum_{k=1}^{K} \dfrac{(n_{jk} - \tilde{n}_{jk})^2}{\tilde{n}_{jk}}$	Chiquadratmaßzahl
$C = C_{XY} = \sqrt{\dfrac{\chi^2}{\chi^2 + n} \cdot \dfrac{\min\{J, K\}}{\min\{J, K\} - 1}}$	Kontingenzkoeffizient
Deskriptive lineare Regression	
$b = \dfrac{s_{XY}}{s_X^2}, \quad a = \bar{y} - b\bar{x}$	Regressionskoeffizienten
$\hat{y}_i = a + b x_i$	Lineare Regressionsfunktion
$u_i = y_i - \hat{y}_i$	Regressionsresiduen
$s_{\hat{Y}}^2 = \frac{1}{n}\sum_{i=1}^{n}(\hat{y}_i - \bar{\hat{y}})^2 = b^2 s_X^2$	Erklärte Varianz
$s_U^2 = \frac{1}{n}\sum_{i=1}^{n} u_i^2$	Residualvarianz
$s_Y^2 = s_{\hat{Y}}^2 + s_U^2$	Varianzzerlegungssatz
$R^2 = \dfrac{s_{\hat{Y}}^2}{s_Y^2} = 1 - \dfrac{s_U^2}{s_Y^2} = r_{XY}^2$	Bestimmtheitsmaß oder Determinationskoeffizient

Zeitreihenanalyse

Symbol, Definition	Bedeutung, Eigenschaften
Gleitende Durchschnitte für äquidistante Zeitpunkte	
$\tilde{g}_i = \frac{y_{i-l}+\ldots+y_{i-1}+y_i+y_{i+1}+\ldots y_{i+l}}{2l+1} = \frac{1}{2l+1}\sum\limits_{h=-l}^{l} y_{i+h}$	Gleitende Durchschnittsbildung ungerader Ordnung $\lambda = 2l+1$ mit $l \in \mathbb{N}$
$\tilde{g}_i = \frac{1}{2l}\left(\frac{1}{2}y_{i-l} + \sum\limits_{h=-(l-1)}^{l-1} y_{i+h} + \frac{1}{2}y_{i+l}\right)$	Gleitende Durchschnittsbildung gerader Ordnung $\lambda = 2l$ mit $l \in \mathbb{N}$
Saisonbereinigung (Phasendurchschnittsverfahren)	
$y_i = g_i + s_i + u_i,\ i = 1,\ldots,n$	Additives Komponentenmodell
$s_{k+j\cdot K} = s_k,\ k = 1,\ldots,K,\ j = 0,1,\ldots,\ \sum\limits_{k=1}^{K} s_k = 0$	Konstante Saisonfigur
K	Periodenlänge der Saisonfigur (z. B. $K = 12$ bei Monatsdaten, $K = 4$ bei Quartalsdaten)
$\tilde{g}_i,\ i = l+1, l+2,\ldots,n-l$	Gleitende Durchschnittswerte der Ordnung K mit $l = K/2$ für gerades K und $l = (K-1)/2$ für ungerades K
$d_i = y_i - \tilde{g}_i, \quad i = l+1, l+2,\ldots,n-l$	Trendbereinigte Reihe
$\bar{d}_k = \frac{1}{J_k}\sum\limits_{j} d_{k+j\cdot K}, \quad k = 1,\ldots,K, \quad j = 0,1,\ldots$	Rohwert der Saisonzahl für die Saisonphase k, wobei J_k die Anzahl der zur Berechnung verfügbaren Werte $d_{k+j\cdot K}$ bezeichnet.[1]
$\hat{s}_k = \bar{d}_k - \frac{1}{K}\sum\limits_{i=1}^{K}\bar{d}_i, \quad k = 1,\ldots,K$	Normierte Saisonzahl für Saisonphase k
$y_i^s = y_i - \hat{s}_k$ für $i = k + j\cdot K$ mit $k = 1,\ldots,K$ und $j = 0,1,\ldots$	Saisonbereinigte Zeitreihe

[1] Die in Mosler/Schmid (2009) verwendete Notation J^* anstatt von J_k verdeutlicht nicht, dass J^* mit k variiert.

A.3 Wahrscheinlichkeitsrechnung

Ereignisse und Wahrscheinlichkeiten

Symbol, Definition	Bedeutung
Ω	Ergebnismenge, sicheres Ereignis
$\omega \in \Omega$	Ergebnis
$A \subset \Omega$	Ereignis
$P(A)$	Wahrscheinlichkeit des Ereignisses A
$\mathcal{P}(\Omega)$	Potenzmenge von Omega
$\mathcal{A} \subset \mathcal{P}(\Omega)$	Ereignissystem
$\emptyset$	leere Menge, unmögliches Ereignis
$\bar{A} = \Omega \setminus A$	Komplementärereignis
$P(A) \overset{\text{def}}{=} \frac{\lvert A\rvert}{\lvert\Omega\rvert} = \frac{\text{Anzahl der für } A \text{ günstigen Fälle}}{\text{Anzahl der möglichen Fälle}}$	Klassische Wahrscheinlichkeit (nach Laplace)
$P(A) \geq 0$	Axiom 1 (Nichtnegativität)
$P(\Omega) = 1$	Axiom 2 (Normierung)
$P(A \cup B) = P(A) + P(B), \quad \text{falls} \quad A \cap B = \emptyset$	Axiom 3 (Additivität)
$A_1, A_2, \ldots$ paarweise disjunkt $\Longrightarrow P\left(\bigcup_{i=1}^{\infty} A_i\right) = \sum_{i=1}^{\infty} P(A_i)$	Axiom 3' (σ-Additivität)
$P(\emptyset) = 0$	Wahrscheinlichkeit des unmöglichen Ereignisses
$P(\bar{A}) = 1 - P(A)$	Wahrscheinlichkeit des Komplementärereignisses
$P(A \setminus B) = P(A) - P(A \cap B)$	
$P(A \cup B) = P(A) + P(B) - P(A \cap B)$	
$P(B\|A) \overset{\text{def}}{=} \frac{P(A \cap B)}{P(A)} \quad \text{mit} \quad P(A) > 0$	Bedingte Wahrscheinlichkeit
$P(A \cap B) = P(A)P(B)$	A und B sind stochastisch unabhängig
$P(A_i \cap A_j) = P(A_i)P(A_j), \quad i,j = 1,\ldots,n, \quad i \neq j$	$A_1, A_2, \ldots, A_n$ sind paarweise unabhängig
$P(A_{i_1} \cap A_{i_2} \cap \ldots \cap A_{i_m}) = P(A_{i_1}) \cdot P(A_{i_2}) \cdot \ldots \cdot P(A_{i_m})$ für jede Auswahl von m Indizes $i_1, \ldots, i_m \in \{1, \ldots, n\}$ mit $2 \leq m \leq n$	$A_1, A_2, \ldots, A_n$ sind vollständig unabhängig
$A_i \cap A_j = \emptyset, \quad i,j = 1,\ldots,n, \quad i \neq j$	$A_1, A_2, \ldots, A_n$ sind paarweise disjunkt
$A_1, A_2, \ldots, A_n$ sind paarweise disjunkt und $\bigcup_{i=1}^{n} A_i = \Omega$	$A_1, \ldots, A_n$ bilden eine vollständige Zerlegung von Ω

- Für $P(A) > 0$ gilt

$$A \text{ und } B \text{ unabhängig} \iff P(B|A) = P(B)\,.$$

- Wenn die Ereignisse $A_1, \ldots, A_n$ eine vollständige Zerlegung von Ω bilden, dann gilt

$$\sum_{i=1}^{n} P(A_i) = P(\Omega) = 1$$

und für jedes Ereignis B gilt

$$P(B) = \sum_{i=1}^{n} P(A_i \cap B)\,.$$

- **Formel (Satz) von der totalen Wahrscheinlichkeit**
  Wenn die $A_1, A_2, \ldots, A_n$ eine vollständige Zerlegung von Ω mit $P(A_i) > 0$ für $i = 1, \ldots, n$ bilden, dann gilt
$$P(B) = \sum_{i=1}^{n} P(B|A_i)P(A_i)$$
  für jedes Ereignis B.

- **Formel (Theorem) von Bayes**
  Wenn die $A_1, A_2, \ldots, A_n$ eine vollständige Zerlegung von Ω mit $P(A_i) > 0$ für $i = 1, \ldots, n$ bilden, dann gilt
$$P(A_i|B) = \frac{P(B|A_i)P(A_i)}{\sum_{j=1}^{n} P(B|A_j)P(A_j)}, \quad i = 1, \ldots, n$$
  für jedes Ereignis B mit $P(B) > 0$.

Zufallsvariablen und Verteilungen

Symbol, Definition	Bedeutung
X	Zufallsvariable
$\mathbb{E}[X]$	Erwartungswert von X
$\mathbb{V}[X] = \mathbb{E}[(X - \mathbb{E}[X])^2]$	Varianz von X
$\sigma[X] = \sqrt{\mathbb{V}[X]}$	Standardabweichung von X
$\mathbb{C}\mathrm{ov}[X, Y] = \mathbb{E}[(X - \mathbb{E}[X])(Y - \mathbb{E}[Y])]$	Kovarianz von X und Y
$\rho_{XY} = \mathbb{C}\mathrm{orr}[X, Y] = \dfrac{\mathbb{C}\mathrm{ov}[X, Y]}{\sigma[X]\sigma[Y]}$	Korrelation von X und Y
$F_X(x) = P(X \leq x), \quad x \in \mathbb{R}$	Verteilungsfunktion von X
$x_p = \min\{x \mid F_X(x) \geq p\}, \quad 0 < p < 1$	p-Quantil von X
$x_{0.5}$	Median
$x_{0.25}$, $x_{0.5}$, $x_{0.75}$	Quartile

- Für die **lineare Transformation** $Z = a + bX$ gilt
$$\mathbb{E}[Z] = a + bE[X], \quad \mathbb{V}[Z] = b^2\mathbb{V}[X], \quad \sigma[Z] = |b|\sigma[X]\,.$$

- Für die **Linearkombination** $Z = aX \pm bY$ gilt
$$\mathbb{E}[Z] = a\mathbb{E}[X] \pm b\mathbb{E}[Y], \quad \mathbb{V}[Z] = a^2\mathbb{V}[X] \pm 2ab\mathbb{C}\mathrm{ov}[X, Y] + b^2\mathbb{V}[Y]\,.$$

- Es gelten die **Verschiebungssätze**
$$\mathbb{V}[X] = \mathbb{E}[X^2] - (\mathbb{E}[X])^2, \quad \mathbb{C}\mathrm{ov}[X, Y] = \mathbb{E}[XY] - \mathbb{E}[X]\mathbb{E}[Y]\,.$$

- Für eine Zufallsvariable X mit $\sigma[X] > 0$ ist
$$\tilde{X} = \frac{X - \mathbb{E}[X]}{\sigma[X]}$$
  die **standardisierte Zufallsvariable**. Für $\tilde{X}$ gilt $\mathbb{E}[\tilde{X}] = 0$ und $\mathbb{V}[\tilde{X}] = 1$.

Diskrete Zufallsvariablen

Symbol, Definition	Bedeutung
$f(x) = P(X = x), \quad x \in \mathbb{R}$	Wahrscheinlichkeitsfunktion
$F(x) = P(X \leq x) = \sum_{t \leq x} f(t), \quad x \in \mathbb{R}$	Verteilungsfunktion
$T_X \stackrel{\text{def}}{=} \{x \in \mathbb{R} \mid f(x) > 0\}$	Träger
$\mathbb{E}[X] = \sum_{x \in T_X} x f(x)$	Erwartungswert von X
$\mathbb{E}[g(X)] = \sum_{x \in T_X} g(x) f(x)$	Erwartungswert von $g(X)$ mit $g : \mathbb{R} \to \mathbb{R}$
$\mathbb{V}[X] = \sum_{x \in T_X} (x - \mathbb{E}[X])^2 f(x)$	Varianz von X

Spezielle diskrete Verteilungen

Verteilung, Symbol	Wahrscheinlichkeitsfunktion	$\mathbb{E}[X]$	$\mathbb{V}[X]$
Binomialverteilung, $X \sim B(n, \pi)$, $0 < \pi < 1$	$\binom{n}{x} \pi^x (1-\pi)^{n-x}$, $x \in \{0, 1, 2, \ldots, n\}$	$n\pi$	$n\pi(1-\pi)$
Bernoulliverteilung, $X \sim B(1, \pi)$, $0 < \pi < 1$	$\pi^x (1-\pi)^{1-x}, x \in \{0, 1\}$	π	$\pi(1-\pi)$
Poisson-Verteilung, $X \sim Poi(\mu)$, $\mu > 0$	$\frac{\mu^x}{x!} e^{-\mu}, \quad x \in \{0, 1, 2, \ldots\}$	μ	μ

Stetige Zufallsvariablen und Verteilungen

Symbol, Definition	Bedeutung
$f(x) \geq 0, \quad x \in \mathbb{R}, \int\limits_{-\infty}^{\infty} f(x)dx = 1$	Dichtefunktion
$F(x) = P(X \leq x) = \int\limits_{-\infty}^{x} f(t)dt, \quad x \in \mathbb{R}$ Es ist $f(x) = F'(x)$ an allen Stellen, an denen $F(x)$ differenzierbar ist.	Verteilungsfunktion
$\mathbb{E}[X] = \int\limits_{-\infty}^{+\infty} xf(x)dx$	Erwartungswert von X
$\mathbb{E}[g(X)] = \int\limits_{-\infty}^{+\infty} g(x)f(x)dx$	Erwartungswert von $g(X)$ mit $g : \mathbb{R} \to \mathbb{R}$
$\mathbb{V}[X] = \int\limits_{-\infty}^{+\infty} (x - \mathbb{E}[X])^2 f(x)dx$	Varianz von X

Spezielle stetige Verteilungen

Verteilung, Symbol	Dichtefunktion	$\mathbb{E}[X]$	$\mathbb{V}[X]$
Standard-Normalverteilung $X \sim N(0,1)$	$\varphi(x) = \frac{1}{\sqrt{2\pi}} e^{-\frac{x^2}{2}}$, $x \in \mathbb{R}$	0	1
Normalverteilung $X \sim N(\mu, \sigma^2)$, $\mu \in \mathbb{R}$, $\sigma^2 > 0$	$\frac{1}{\sqrt{2\pi}\sigma} e^{-\frac{(x-\mu)^2}{2\sigma^2}}$, $x \in \mathbb{R}$	μ	σ^2
Exponentialverteilung $X \sim Exp(\lambda)$, $\lambda > 0$	$\lambda e^{-\lambda x}$, $x \geq 0$	$\frac{1}{\lambda}$	$\frac{1}{\lambda^2}$
Rechteckverteilung $X \sim R(\alpha, \beta)$, $\alpha \in \mathbb{R}$, $\beta \in \mathbb{R}$, $\alpha < \beta$	$\frac{1}{\beta - \alpha}, \alpha \leq x \leq \beta$	$\frac{\alpha + \beta}{2}$	$\frac{(\beta - \alpha)^2}{12}$

Faustregeln für Approximationen

Approximation von

- $B(n, \pi)$ durch $N(\mu, \sigma^2)$ mit $\mu = n\pi$ und $\sigma^2 = n\pi(1 - \pi)$, falls $n\pi(1 - \pi) > 9$,
- $B(n, \pi)$ durch $Poi(\mu)$ mit $\mu = n\pi$, falls $\pi \leq 0.1$, $n \geq 50$ und $n\pi \leq 9$,
- $Poi(\mu)$ durch $N(\mu, \sigma^2)$ mit $\sigma^2 = \mu$, falls $\mu > 9$.

Symbol, Definition	Bedeutung
Gemeinsame Verteilung von zwei Zufallsvariablen	
$F_{XY}:\mathbb{R}^2\to[0,1]$, $F_{XY}(x,y)=P(X\le x,Y\le y)$	Gemeinsame Verteilungsfunktion von X und Y
$f_{XY}:\mathbb{R}^2\to[0,1]$, $f_{XY}(x,y)=P(X=x,Y=y)$	Gemeinsame Wahrscheinlichkeitsfunktion (diskrete Verteilung) von X und Y
$f_{XY}:\mathbb{R}^2\to\mathbb{R}$, $F_{XY}(x,y)=\int\limits_{-\infty}^{y}\int\limits_{-\infty}^{x} f_{XY}(z,v)dzdv$	Gemeinsame Dichtefunktion (stetige Verteilung) von X und Y
$\mathbb{E}[XY]=\sum\limits_j\sum\limits_k x_jy_kP(X=x_j,Y=y_k)$	Erwartungswert von XY für gemeinsame diskrete Verteilung
$\mathbb{E}[g(X,Y)]=\sum\limits_j\sum\limits_k g(x_j,y_k)P(X=x_j,Y=y_k)$	Erwartungswert von $g(X,Y)$ für gemeinsame diskrete Verteilung und $g:\mathbb{R}^2\to\mathbb{R}$
$\mathbb{E}[XY]=\int\limits_{-\infty}^{+\infty}\int\limits_{-\infty}^{+\infty} xyf_{XY}(x,y)dydx$	Erwartungswert von XY für gemeinsame stetige Verteilung
$\mathbb{E}[g(X,Y)]=\int\limits_{-\infty}^{+\infty}\int\limits_{-\infty}^{+\infty} g(x,y)f_{XY}(x,y)dydx$	Erwartungswert von $g(X,Y)$ für gemeinsame stetige Verteilung und $g:\mathbb{R}^2\to\mathbb{R}$
$F_X(x)=\lim\limits_{y\to\infty}F_{XY}(x,y),\quad x\in\mathbb{R}$	Randverteilungsfunktion von X
$P(X=x)=\sum_k P(X=x,Y=y_k),\quad x\in\mathbb{R}$	Randwahrscheinlichkeitsfunktion von X für gemeinsame diskrete Verteilung
$f_X(x)=\int\limits_{-\infty}^{+\infty} f_{XY}(x,y)dy,\quad x\in\mathbb{R}$	Randdichtefunktion von X für gemeinsame stetige Verteilung
$P(X=x\|Y=y_k)=\dfrac{P(X=x,Y=y_k)}{P(Y=y_k)},\quad x\in\mathbb{R}$	Bedingte Wahrscheinlichkeitsfunktion von X unter der Bedingung $Y=y_k$, falls $P(Y=y_k)>0$
$f_{X\|Y=y}(x)=\dfrac{f_{XY}(x,y)}{f_Y(y)},\quad x\in\mathbb{R}$	Bedingte Dichtefunktion von X unter der Bedingung $Y=y$, wobei $f_Y(y)>0$
$F_{XY}(x,y)=F_X(x)F_Y(y),\quad x,y\in\mathbb{R}$	Unabhängigkeit von X und Y (allgemein)
$f_{XY}(x,y)=f_X(x)f_Y(y),\quad x,y\in\mathbb{R}$	Unabhängigkeit von X und Y (diskrete oder stetige Verteilung)
Gemeinsame Verteilung von n Zufallsvariablen	
$F_{X_1,\ldots,X_n}:\mathbb{R}^n\to[0,1]$, $F_{X_1,\ldots,X_n}(x_1,\ldots,x_n)=P(X_1\le x_1,\ldots,X_n\le x_n)$	Gemeinsame Verteilungsfunktion
$f_{X_1,\ldots,X_n}:\mathbb{R}^n\to[0,1]$, $f_{X_1,\ldots,X_n}(x_1,\ldots,x_n)=P(X_1=x_1,\ldots,X_n=x_n)$	Gemeinsame Wahrscheinlichkeitsfunktion (diskrete Verteilung)
$f_{X_1,\ldots,X_n}:\mathbb{R}^n\to\mathbb{R}$, $F_{X_1,\ldots,X_n}(x_1,\ldots,x_n)$ $=\int\limits_{-\infty}^{x_1}\ldots\int\limits_{-\infty}^{x_n} f_{X_1,\ldots,X_n}(t_1,\ldots,t_n)dt_n\ldots dt_1$	Gemeinsame Dichtefunktion (stetige Verteilung)
$F_{X_1}:\mathbb{R}\to\infty$, $F_{X_1}(x_1)=\lim\limits_{x_2\to\infty}\cdots\lim\limits_{x_n\to\infty}F_{X_1,\ldots,X_n}(x_1,x_2,\ldots,x_n)$	Randverteilungsfunktion von X_1
$F_{X_1,\ldots,X_n}(x_1,\ldots,x_n)=\prod\limits_{i=1}^{n}F_{X_i}(x_i)$	Unabhängigkeit (allgemein)
$f_{X_1,\ldots,X_n}(x_1,\ldots,x_n)=\prod\limits_{i=1}^{n}f_{X_i}(x_i)$	Unabhängigkeit (diskrete oder stetige Verteilung)

A.4 Schließende Statistik

Symbol	Bedeutung
$\varphi(x)$	Dichtefunktion der Standard-Normalverteilung
$\Phi(x) = \int_{-\infty}^{x} \varphi(u)du$	Verteilungsfunktion der Standard-Normalverteilung
u_p	p-Quantil (p-Fraktil) der Standard-Normalverteilung
$t_{\nu,p}$	p-Quantil einer t-Verteilung mit ν Freiheitsgraden
$\chi^2_{\nu,p}$	p-Quantil einer χ^2-Verteilung mit ν Freiheitsgraden
$F_{n,m,p}$	p-Quantil einer F-Verteilung mit n und m Freiheitsgraden

Punktschätzung	
θ	Zu schätzender Parameter
Θ	Zugelassener Parameterraum, $\theta \in \Theta$
$\hat{\theta}(x_1, \ldots, x_n)$	Schätzwert für θ
$\hat{\theta} = \hat{\theta}(X_1, \ldots, X_n)$	Schätzfunktion, Schätzer für θ
$\mathbb{E}[\hat{\theta}] = \theta$ für alle $\theta \in \Theta$	Erwartungstreuer oder unverzerrter Schätzer für θ
$S^2 \stackrel{\text{def}}{=} \frac{1}{n}\sum_{i=1}^{n}(X_i - \bar{X})^2$	Schätzer für die Varianz σ^2
$S^{*2} \stackrel{\text{def}}{=} \frac{1}{n-1}\sum_{i=1}^{n}(X_i - \bar{X})^2$	korrigierter, erwartungstreuer Schätzer für die Varianz σ^2
$S \stackrel{\text{def}}{=} \sqrt{S^2}, \quad S^* \stackrel{\text{def}}{=} \sqrt{S^{*2}}$	Schätzer für die Standardabweichung σ

Intervallschätzung	
$\bar{X} \mp u_{1-\frac{\alpha}{2}} \frac{\sigma}{\sqrt{n}}$	Konfidenzintervall für μ, normalverteilte Grundgesamtheit, σ^2 bekannt
$\bar{X} \mp t_{n-1,1-\frac{\alpha}{2}} \frac{S}{\sqrt{n-1}} = \bar{X} \mp t_{n-1,1-\frac{\alpha}{2}} \frac{S^*}{\sqrt{n}}$	Konfidenzintervall für μ, normalverteilte Grundgesamtheit, σ^2 unbekannt
$\left[\frac{nS^2}{\chi^2_{n-1,1-\frac{\alpha}{2}}}, \frac{nS^2}{\chi^2_{n-1,\frac{\alpha}{2}}}\right] = \left[\frac{(n-1)S^{*2}}{\chi^2_{n-1,1-\frac{\alpha}{2}}}, \frac{(n-1)S^{*2}}{\chi^2_{n-1,\frac{\alpha}{2}}}\right]$	Konfidenzintervall für σ^2, normalverteilte Grundgesamtheit, μ unbekannt
$\hat{\pi} \mp u_{1-\frac{\alpha}{2}} \sqrt{\frac{\hat{\pi}(1-\hat{\pi})}{n}}$	Approximatives Konfidenzintervall für eine Wahrscheinlichkeit π
$\bar{X} \mp u_{1-\frac{\alpha}{2}} \frac{\sigma}{\sqrt{n}}$	Approximatives Konfidenzintervall für μ, beliebige Grundgesamtheit, σ^2 bekannt
$\bar{X} \mp u_{1-\frac{\alpha}{2}} \frac{S}{\sqrt{n}}$	Approximatives Konfidenzintervall für μ, beliebige Grundgesamtheit, σ^2 unbekannt

Testverfahren

- Die **Nullhypothese** $H_0 : \theta \in \Theta_0$ wird zugunsten der **Gegenhypothese** $H_1 : \theta \in \Theta_1$ abgelehnt, falls die **Prüfgröße** (Testgröße, Testfunktion) T im **kritischen Bereich** (**Verwerfungsbereich**, **Ablehnbereich**) K liegt. Es gilt $\Theta_0 \cap \Theta_1 = \emptyset$.
- Ein Test zum vorgegebenen **Signifikanzniveau** α erfüllt $P(T \in K||\theta) \leq \alpha$ für alle $\theta \in \Theta_0$. Im Fall einer einfachen Nullhypothese $H_0 : \theta = \theta_0$ gilt in der Regel $P(T \in K||\theta_0) = \alpha$. Im Fall einer zusammengesetzten Nullhypothese gilt in der Regel $\max\limits_{\theta \in \Theta_0} P(T \in K||\theta) = \alpha$ oder $\sup\limits_{\theta \in \Theta_0} P(T \in K||\theta) = \alpha$.

Gauß-Test Normalverteilte Grundgesamtheit, σ^2 bekannt.

$$T = \frac{\bar{X} - \mu_0}{\sigma}\sqrt{n}$$

H_0	H_1	H_0 wird abgelehnt, falls
$\mu = \mu_0$	$\mu \neq \mu_0$	$\lvert T\rvert > u_{1-\frac{\alpha}{2}}$
$\mu = \mu_0$ oder $\mu \geq \mu_0$	$\mu < \mu_0$	$T < -u_{1-\alpha}$
$\mu = \mu_0$ oder $\mu \leq \mu_0$	$\mu > \mu_0$	$T > u_{1-\alpha}$

- Approximativer Gauß-Test bei nicht normalverteilter Grundgesamtheit; Faustregel: $n \geq 40$.
- Approximativer Gauß-Test bei nicht normalverteilter Grundgesamtheit und geschätzter Varianz; σ wird durch S ersetzt; Faustregel: $n \geq 40$.

***t*-Test** Normalverteilte Grundgesamtheit, σ^2 unbekannt.

$$T = \frac{\bar{X} - \mu_0}{S^*}\sqrt{n} = \frac{\bar{X} - \mu_0}{S}\sqrt{n-1},$$

H_0	H_1	H_0 wird abgelehnt, falls
$\mu = \mu_0$	$\mu \neq \mu_0$	$\lvert T\rvert > t_{n-1,1-\frac{\alpha}{2}}$
$\mu = \mu_0$ oder $\mu \geq \mu_0$	$\mu < \mu_0$	$T < -t_{n-1,1-\alpha}$
$\mu = \mu_0$ oder $\mu \leq \mu_0$	$\mu > \mu_0$	$T > t_{n-1,1-\alpha}$

Zweistichproben-Gauß-Test Zwei normalverteilte Grundgesamtheiten, σ_X^2 und σ_Y^2 bekannt, unabhängige Stichproben.

$$T = \frac{\bar{X} - \bar{Y}}{\sqrt{\frac{\sigma_X^2}{n} + \frac{\sigma_Y^2}{m}}}$$

H_0	H_1	H_0 wird abgelehnt, falls
$\mu_X = \mu_Y$	$\mu_X \neq \mu_Y$	$\lvert T\rvert > u_{1-\frac{\alpha}{2}}$
$\mu_X = \mu_Y$ oder $\mu_X \geq \mu_Y$	$\mu_X < \mu_Y$	$T < -u_{1-\alpha}$
$\mu_X = \mu_Y$ oder $\mu_X \leq \mu_Y$	$\mu_X > \mu_Y$	$T > u_{1-\alpha}$

- Approximativer Zweistichproben-Gauß-Test bei nicht normalverteilter Grundgesamtheit; Faustregel: $m \geq 40$ und $n \geq 40$.
- Approximativer Zweistichproben-Gauß-Test bei nicht normalverteilter Grundgesamtheit und geschätzten Varianzen; σ_X^2 wird durch S_X^2 und σ_Y^2 wird durch S_Y^2 ersetzt; Faustregel: $m \geq 40$ und $n \geq 40$.

Zweistichproben-*t*-Test Zwei normalverteilte Grundgesamtheiten, σ_X^2 und σ_Y^2 unbekannt, $\sigma_X^2 = \sigma_Y^2$, unabhängige Stichproben.

$$T = \sqrt{\frac{nm(n+m-2)}{n+m}}\,\frac{\bar{X} - \bar{Y}}{\sqrt{(n-1)S_X^{*2} + (m-1)S_Y^{*2}}}$$

H_0	H_1	H_0 wird abgelehnt, falls
$\mu_X = \mu_Y$	$\mu_X \neq \mu_Y$	$\lvert T\rvert > t_{n+m-2,1-\frac{\alpha}{2}}$
$\mu_X = \mu_Y$ oder $\mu_X \geq \mu_Y$	$\mu_X < \mu_Y$	$T < -t_{n+m-2,1-\alpha}$
$\mu_X = \mu_Y$ oder $\mu_X \leq \mu_Y$	$\mu_X > \mu_Y$	$T > t_{n+m-2,1-\alpha}$

Differenzentest für verbundene Stichproben (*t*-Differenzentest) Unabhängige und normalverteilte Differenzen $D_i = X_i - Y_i$ für $i = 1, \ldots, n$, $\mu_D = \mathbb{E}[X_i] - \mathbb{E}[Y_i]$, $\bar{D} = \frac{1}{n}\sum\limits_{i=1}^{n} D_i$, $S_D = \sqrt{\frac{1}{n}\sum_{i=1}^{n}(D_i - \bar{D})^2}$.

$$T = \frac{\bar{D} - \mu_0}{S_D}\sqrt{n-1}$$

H_0	H_1	H_0 wird abgelehnt, falls
$\mu_D = \mu_0$	$\mu_D \neq \mu_0$	$\lvert T\rvert > t_{n-1,1-\frac{\alpha}{2}}$
$\mu_D = \mu_0$ oder $\mu_D \geq \mu_0$	$\mu_D < \mu_0$	$T < -t_{n-1,1-\alpha}$
$\mu_D = \mu_0$ oder $\mu_D \leq \mu_0$	$\mu_D > \mu_0$	$T > t_{n-1,1-\alpha}$

Test für eine Varianz σ^2 Normalverteilte Grundgesamtheit.

$$T = \frac{(n-1)S^{*2}}{\sigma_0^2} = \frac{nS^2}{\sigma_0^2}$$

H_0	H_1	H_0 wird abgelehnt, falls
$\sigma^2 = \sigma_0^2$	$\sigma^2 \neq \sigma_0^2$	$T < \chi^2_{n-1,\frac{\alpha}{2}}$ oder $T > \chi^2_{n-1,1-\frac{\alpha}{2}}$
$\sigma^2 = \sigma_0^2$ oder $\sigma^2 \geq \sigma_0^2$	$\sigma^2 < \sigma_0^2$	$T < \chi^2_{n-1,\alpha}$
$\sigma^2 = \sigma_0^2$ oder $\sigma^2 \leq \sigma_0^2$	$\sigma^2 > \sigma_0^2$	$T > \chi^2_{n-1,1-\alpha}$

Vergleich zweier Varianzen σ_X^2 und σ_Y^2 Zwei normalverteilte Grundgesamtheiten, unabhängige Stichproben vom Umfang n bzw. m.

$$T = \frac{S_X^{*2}}{S_Y^{*2}}$$

H_0	H_1	H_0 wird abgelehnt, falls
$\sigma_X^2 = \sigma_Y^2$	$\sigma_X^2 \neq \sigma_Y^2$	$T < F_{n-1,m-1,\frac{\alpha}{2}}$ oder $T > F_{n-1,m-1,1-\frac{\alpha}{2}}$
$\sigma_X^2 = \sigma_Y^2$ oder $\sigma_X^2 \geq \sigma_Y^2$	$\sigma_X^2 < \sigma_Y^2$	$T < F_{n-1,m-1,\alpha}$
$\sigma_X^2 = \sigma_Y^2$ oder $\sigma_X^2 \leq \sigma_Y^2$	$\sigma_X^2 > \sigma_Y^2$	$T > F_{n-1,m-1,1-\alpha}$

Test für eine Wahrscheinlichkeit π Approximativer Test, Faustregel: $n\pi_0(1-\pi_0) > 9$. $\hat{\pi} = \bar{X}$ ist der Anteilswert in der Stichprobe $X_1, \ldots, X_n \overset{i.i.d.}{\sim} B(1, \pi)$.

$$T = \frac{\hat{\pi} - \pi_0}{\sqrt{\pi_0(1-\pi_0)}}\sqrt{n}$$

H_0	H_1	H_0 wird abgelehnt, falls
$\pi = \pi_0$	$\pi \neq \pi_0$	$\lvert T \rvert > u_{1-\frac{\alpha}{2}}$
$\pi = \pi_0$ oder $\pi \geq \pi_0$	$\pi < \pi_0$	$T < -u_{1-\alpha}$
$\pi = \pi_0$ oder $\pi \leq \pi_0$	$\pi > \pi_0$	$T > u_{1-\alpha}$

Vergleich zweier Wahrscheinlichkeiten π_X und π_Y $\hat{\pi}_X = \bar{X}$ und $\hat{\pi}_Y = \bar{Y}$ sind die Anteilswerte der unabhängigen Stichproben $X_1, \ldots, X_n \overset{i.i.d.}{\sim} B(1, \pi_X)$ und $Y_1, \ldots, Y_m \overset{i.i.d.}{\sim} B(1, \pi_Y)$. Faustregel für die Anwendung $n\hat{\pi}_X(1-\hat{\pi}_X) > 9$ und $m\hat{\pi}_Y(1-\hat{\pi}_Y) > 9$.

$$T = \frac{\hat{\pi}_X - \hat{\pi}_Y}{\sqrt{\frac{\hat{\pi}_X(1-\hat{\pi}_X)}{n} + \frac{\hat{\pi}_Y(1-\hat{\pi}_Y)}{m}}}$$

H_0	H_1	H_0 wird abgelehnt, falls
$\pi_X = \pi_Y$	$\pi \neq \pi_0$	$\lvert T \rvert > u_{1-\frac{\alpha}{2}}$
$\pi_X = \pi_Y$ oder $\pi_X \geq \pi_Y$	$\pi_X < \pi_Y$	$T < -u_{1-\alpha}$
$\pi_X = \pi_Y$ oder $\pi_X \leq \pi_Y$	$\pi_X > \pi_Y$	$T > u_{1-\alpha}$

χ^2-Anpassungstest $H_0: F = F_0$, $H_1: F \neq F_0$; F_0 ist eine hypothetische Verteilung.

$$T = \sum_{j=1}^{J} \frac{(N_j - n\pi_j)^2}{n\pi_j} = \left(\sum_{j=1}^{J} \frac{N_j^2}{n\pi_j} \right) - n$$

mit

N_j :	absolute Häufigkeit
π_j :	Wahrscheinlichkeit für das j-te Intervall, falls H_0 richtig ist
n :	Stichprobenumfang
J :	Anzahl der disjunkten Intervalle

$\boxed{H_0 \text{ wird abgelehnt, falls } T > \chi^2_{J-1,1-\alpha}.}$

Faustregel für Anwendung: $n\pi_j \geq 5$ für $j = 1, \ldots, J$.

Werden r unbekannte Parameter aus der Verteilung der Häufigkeiten $N_1, \ldots, N_J$ mit Hilfe der Maximum-Likelihood-Methodik geschätzt, so ist der kritische Wert $\chi^2_{J-1-r,1-\alpha}$ anstatt $\chi^2_{J-1,1-\alpha}$

χ^2-Unabhängigkeitstest $H_0 : X$ und Y sind unabhängig; $H_1 : X$ und Y sind abhängig

$$T = \sum_{j=1}^{J}\sum_{k=1}^{K} \frac{(N_{jk} - \tilde{N}_{jk})^2}{\tilde{N}_{jk}} = n\left(\sum_{j=1}^{J}\sum_{k=1}^{K} \frac{N_{jk}^2}{N_{j.}N_{.k}} - 1\right)$$

mit

J			Anzahl der Intervalle für die Zufallsvariable X
K			Anzahl der Intervalle für die Zufallsvariable Y
N_{jk}			absolute Häufigkeit
n	$=$	$\sum\limits_{j=1}^{J}\sum\limits_{k=1}^{K} N_{jk}$	Stichprobenumfang
$\tilde{N}_{jk}$	$\overset{\text{def}}{=}$	$\frac{N_{j.}N_{.k}}{n}$	Schätzer für die erwartete Häufigkeit bei Unabhängigkeit
$N_{j.}$	$\overset{\text{def}}{=}$	$\sum\limits_{k=1}^{K} N_{jk}$	Randhäufigkeit
$N_{.k}$	$\overset{\text{def}}{=}$	$\sum\limits_{j=1}^{J} N_{jk}$	Randhäufigkeit

$\boxed{H_0 \text{ wird abgelehnt, falls } T > \chi^2_{(J-1)(K-1),1-\alpha}.}$

Faustregel für Anwendung: $\tilde{n}_{jk} \geq 5$ für $j = 1, \ldots, J$ und $k = 1, \ldots, K$.

p-Wert und Testdurchführung am Beispiel des Gaußtests

Eine Realisation t von T wird beobachtet. Der p-Wert wird mit Hilfe von t berechnet. H_0 wird abgelehnt, falls $p \leq \alpha$.

H_0	H_1	p-Wert berechnet aus t
$\mu = \mu_0$	$\mu \neq \mu_0$	$p = P(\lvert T\rvert > \lvert t\rvert \,\Vert \mu_0) = 2P(T > \lvert t\rvert \,\Vert \mu_0)$
$\mu = \mu_0$ oder $\mu \geq \mu_0$	$\mu < \mu_0$	$p = P(T < t \,\Vert \mu_0)$
$\mu = \mu_0$ oder $\mu \leq \mu_0$	$\mu > \mu_0$	$p = P(T > t \,\Vert \mu_0)$

Asymptotische Statistik

- Eine Folge von Zufallsvariablen $X_1, X_2, \ldots$ **konvergiert nach Wahrscheinlichkeit** gegen c, notiert als $\operatorname*{p\,lim}_{n\to\infty} X_n = c$, falls
$$\lim_{n\to\infty} P\left(|X_n - c| \geq \varepsilon\right) = 0 \quad \text{für alle } \varepsilon > 0.$$
- Das (schwache) **Gesetz der großen Zahlen** besagt, dass für X_i i.i.d. mit $\mathbb{E}[X_i] = \mu$ die Folge der Mittelwerte
$$\bar{X}_n \overset{\text{def}}{=} \frac{1}{n}\sum_{i=1}^{n} X_i, \quad n = 1, 2, \ldots$$
nach Wahrscheinlichkeit gegen μ konvergiert, d. h. $\operatorname*{p\,lim}_{n\to\infty} \bar{X}_n = \mu$.
- Der **zentrale Grenzwertsatz der Statistik** besagt, dass für X_i i.i.d. mit $\mathbb{E}[X_i] = \mu$ und $\mathbb{V}[X_i] = \sigma^2 < \infty$ die Verteilung der standardisierten Summen bzw. Mittelwerte
$$Y_n \overset{\text{def}}{=} \frac{\sum\limits_{i=1}^{n} X_i - n\mu}{\sqrt{n\sigma^2}} = \frac{\bar{X}_n - \mu}{\sigma}\sqrt{n}$$
gegen die Standard-Normalverteilung konvergiert,
$$\lim_{n\to\infty} P\left(Y_n \leq x\right) = \Phi(x), \quad x \in \mathbb{R}.$$
- Eine Folge $\hat{\theta}_1(X_1)$, $\hat{\theta}_2(X_1, X_2)$, $\ldots$, $\hat{\theta}_n(X_1, \ldots, X_n)$,… von Schätzern heißt **konsistent** für θ, falls die Folge $\hat{\theta}_n$ nach Wahrscheinlichkeit gegen θ konvergiert.

A.5 Tabellen

Tabelle 1: Verteilungsfunktion der Binomialverteilung

Tabelliert sind die Werte $P(X \leq x)$ für $X \sim B(n, \pi)$. Es gilt $P(X \leq n) = 1$.

		π							
n	x	0.01	0.05	0.10	0.15	0.20	0.30	0.40	0.50
1	0	0.990	0.950	0.900	0.850	0.800	0.700	0.600	0.500
2	0	0.980	0.903	0.810	0.723	0.640	0.490	0.360	0.250
	1	1.000	0.998	0.990	0.978	0.960	0.910	0.840	0.750
3	0	0.970	0.857	0.729	0.614	0.512	0.343	0.216	0.125
	1	1.000	0.993	0.972	0.939	0.896	0.784	0.648	0.500
	2	1.000	1.000	0.999	0.997	0.992	0.973	0.936	0.875
4	0	0.961	0.815	0.656	0.522	0.410	0.240	0.130	0.063
	1	0.999	0.986	0.948	0.890	0.819	0.652	0.475	0.313
	2	1.000	1.000	0.996	0.988	0.973	0.916	0.821	0.688
	3	1.000	1.000	1.000	0.999	0.998	0.992	0.974	0.938
5	0	0.951	0.774	0.590	0.444	0.328	0.168	0.078	0.031
	1	0.999	0.977	0.919	0.835	0.737	0.528	0.337	0.188
	2	1.000	0.999	0.991	0.973	0.942	0.837	0.683	0.500
	3	1.000	1.000	1.000	0.998	0.993	0.969	0.913	0.813
	4	1.000	1.000	1.000	1.000	1.000	0.998	0.990	0.969
6	0	0.941	0.735	0.531	0.377	0.262	0.118	0.047	0.016
	1	0.999	0.967	0.886	0.776	0.655	0.420	0.233	0.109
	2	1.000	0.998	0.984	0.953	0.901	0.744	0.544	0.344
	3	1.000	1.000	0.999	0.994	0.983	0.930	0.821	0.656
	4	1.000	1.000	1.000	1.000	0.998	0.989	0.959	0.891
	5	1.000	1.000	1.000	1.000	1.000	0.999	0.996	0.984
10	0	0.904	0.599	0.349	0.197	0.107	0.028	0.006	0.001
	1	0.996	0.914	0.736	0.544	0.376	0.149	0.046	0.011
	2	1.000	0.988	0.930	0.820	0.678	0.383	0.167	0.055
	3	1.000	0.999	0.987	0.950	0.879	0.650	0.382	0.172
	4	1.000	1.000	0.998	0.990	0.967	0.850	0.633	0.377
	5	1.000	1.000	1.000	0.999	0.994	0.953	0.834	0.623
	6	1.000	1.000	1.000	1.000	0.999	0.989	0.945	0.828
	7	1.000	1.000	1.000	1.000	1.000	0.998	0.988	0.945
	8	1.000	1.000	1.000	1.000	1.000	1.000	0.998	0.989
	9	1.000	1.000	1.000	1.000	1.000	1.000	1.000	0.999
20	0	0.818	0.358	0.122	0.039	0.012	0.001	0.000	0.000
	1	0.983	0.736	0.392	0.176	0.069	0.008	0.001	0.000
	2	0.999	0.925	0.677	0.405	0.206	0.035	0.004	0.000
	3	1.000	0.984	0.867	0.648	0.411	0.107	0.016	0.001
	4	1.000	0.997	0.957	0.830	0.630	0.238	0.051	0.006
	5	1.000	1.000	0.989	0.933	0.804	0.416	0.126	0.021
	6	1.000	1.000	0.998	0.978	0.913	0.608	0.250	0.058
	7	1.000	1.000	1.000	0.994	0.968	0.772	0.416	0.132
	8	1.000	1.000	1.000	0.999	0.990	0.887	0.596	0.252
	9	1.000	1.000	1.000	1.000	0.997	0.952	0.755	0.412
	10	1.000	1.000	1.000	1.000	0.999	0.983	0.872	0.588
	11	1.000	1.000	1.000	1.000	1.000	0.995	0.943	0.748
	12	1.000	1.000	1.000	1.000	1.000	0.999	0.979	0.868
	13	1.000	1.000	1.000	1.000	1.000	1.000	0.994	0.942
	14	1.000	1.000	1.000	1.000	1.000	1.000	0.998	0.979
	15	1.000	1.000	1.000	1.000	1.000	1.000	1.000	0.994
	16	1.000	1.000	1.000	1.000	1.000	1.000	1.000	0.999
	17	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000
	18	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000
	19	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000

Tabelle 2: Verteilungsfunktion der Poisson-Verteilung

Tabelliert sind die Werte $P(X \leq x)$ für $X \sim Poi(\mu)$. Hinweis: Häufig wird der Parameter der Poisson-Verteilung auch mit λ bezeichnet.

	μ											
x	0.001	0.005	0.010	0.015	0.020	0.025	0.030	0.035	0.040	0.045	0.050	0.055
0	0.999	0.995	0.990	0.985	0.980	0.975	0.970	0.966	0.961	0.956	0.951	0.946
1	1.000	1.000	1.000	1.000	1.000	1.000	1.000	0.999	0.999	0.999	0.999	0.999
2	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000
3	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000
4	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000

	μ											
x	0.060	0.065	0.070	0.075	0.080	0.085	0.090	0.095	0.100	0.200	0.300	0.400
0	0.942	0.937	0.932	0.928	0.923	0.919	0.914	0.909	0.905	0.819	0.741	0.670
1	0.998	0.998	0.998	0.997	0.997	0.997	0.996	0.996	0.995	0.982	0.963	0.938
2	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	0.999	0.996	0.992
3	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	0.999
4	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000
5	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000

	μ										
x	0.5	0.6	0.7	0.8	0.9	1.0	1.2	1.4	1.6	1.8	2.0
0	0.607	0.549	0.497	0.449	0.407	0.368	0.301	0.247	0.202	0.165	0.135
1	0.910	0.878	0.844	0.809	0.772	0.736	0.663	0.592	0.525	0.463	0.406
2	0.986	0.977	0.966	0.953	0.937	0.920	0.879	0.833	0.783	0.731	0.677
3	0.998	0.997	0.994	0.991	0.987	0.981	0.966	0.946	0.921	0.891	0.857
4	1.000	1.000	0.999	0.999	0.998	0.996	0.992	0.986	0.976	0.964	0.947
5	1.000	1.000	1.000	1.000	1.000	0.999	0.998	0.997	0.994	0.990	0.983
6	1.000	1.000	1.000	1.000	1.000	1.000	1.000	0.999	0.999	0.997	0.995
7	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	0.999	0.999
8	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000

	μ										
x	2.5	3.0	3.5	4.0	4.5	5.0	6.0	7.0	8.0	9.0	10.0
0	0.082	0.050	0.030	0.018	0.011	0.007	0.002	0.001	0.000	0.000	0.000
1	0.287	0.199	0.136	0.092	0.061	0.040	0.017	0.007	0.003	0.001	0.000
2	0.544	0.423	0.321	0.238	0.174	0.125	0.062	0.030	0.014	0.006	0.003
3	0.758	0.647	0.537	0.433	0.342	0.265	0.151	0.082	0.042	0.021	0.010
4	0.891	0.815	0.725	0.629	0.532	0.440	0.285	0.173	0.100	0.055	0.029
5	0.958	0.916	0.858	0.785	0.703	0.616	0.446	0.301	0.191	0.116	0.067
6	0.986	0.966	0.935	0.889	0.831	0.762	0.606	0.450	0.313	0.207	0.130
7	0.996	0.988	0.973	0.949	0.913	0.867	0.744	0.599	0.453	0.324	0.220
8	0.999	0.996	0.990	0.979	0.960	0.932	0.847	0.729	0.593	0.456	0.333
9	1.000	0.999	0.997	0.992	0.983	0.968	0.916	0.830	0.717	0.587	0.458
10	1.000	1.000	0.999	0.997	0.993	0.986	0.957	0.901	0.816	0.706	0.583
11	1.000	1.000	1.000	0.999	0.998	0.995	0.980	0.947	0.888	0.803	0.697
12	1.000	1.000	1.000	1.000	0.999	0.998	0.991	0.973	0.936	0.876	0.792
13	1.000	1.000	1.000	1.000	1.000	0.999	0.996	0.987	0.966	0.926	0.864
14	1.000	1.000	1.000	1.000	1.000	1.000	0.999	0.994	0.983	0.959	0.917
15	1.000	1.000	1.000	1.000	1.000	1.000	0.999	0.998	0.992	0.978	0.951
16	1.000	1.000	1.000	1.000	1.000	1.000	1.000	0.999	0.996	0.989	0.973
17	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	0.998	0.995	0.986
18	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	0.999	0.998	0.993
19	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	0.999	0.997
20	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	0.998
21	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	0.999
22	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000

Tabelle 3: Verteilungsfunktion der Standard-Normalverteilung

Tabelliert sind die Werte $\Phi(u) = P(U \leq u)$ für $U \sim N(0,1)$. Der Wert von u ergibt sich aus der Summe der jeweiligen Werte in der ersten Zeile und ersten Spalte. Beispiel: $0.1 + 0.02 = 0.12$ und $\Phi(0.12) = 0.5478$. Es gilt $\Phi(-u) = 1 - \Phi(u)$. Für die Quantile $u_p = \Phi^{-1}(p)$ gilt $u_{1-p} = -u_p$.

u	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

Einige häufig verwendete Quantile der Standardnormalverteilung:

p	0.950	0.975	0.990	0.995
u_p	1.645	1.960	2.326	2.576

Tabelle 4: Quantile der t-Verteilung

Tabelliert sind p-Quantile $t_{\nu,p}$ einer t-Verteilung mit ν Freiheitsgraden. Für $\nu \to \infty$ ergeben sich die Quantile einer Standard-Normalverteilung. Es gilt $t_{\nu,1-p} = -t_{\nu,p}$.

	p						
ν	0.995	0.990	0.975	0.950	0.925	0.900	0.750
1	63.66	31.82	12.71	6.314	4.165	3.078	1.000
2	9.925	6.965	4.303	2.920	2.282	1.886	0.816
3	5.841	4.541	3.182	2.353	1.924	1.638	0.765
4	4.604	3.747	2.776	2.132	1.778	1.533	0.741
5	4.032	3.365	2.571	2.015	1.699	1.476	0.727
6	3.707	3.143	2.447	1.943	1.650	1.440	0.718
7	3.499	2.998	2.365	1.895	1.617	1.415	0.711
8	3.355	2.896	2.306	1.860	1.592	1.397	0.706
9	3.250	2.821	2.262	1.833	1.574	1.383	0.703
10	3.169	2.764	2.228	1.812	1.559	1.372	0.700
11	3.106	2.718	2.201	1.796	1.548	1.363	0.697
12	3.055	2.681	2.179	1.782	1.538	1.356	0.695
13	3.012	2.650	2.160	1.771	1.530	1.350	0.694
14	2.977	2.624	2.145	1.761	1.523	1.345	0.692
15	2.947	2.602	2.131	1.753	1.517	1.341	0.691
16	2.921	2.583	2.120	1.746	1.512	1.337	0.690
17	2.898	2.567	2.110	1.740	1.508	1.333	0.689
18	2.878	2.552	2.101	1.734	1.504	1.330	0.688
19	2.861	2.539	2.093	1.729	1.500	1.328	0.688
20	2.845	2.528	2.086	1.725	1.497	1.325	0.687
21	2.831	2.518	2.080	1.721	1.494	1.323	0.686
22	2.819	2.508	2.074	1.717	1.492	1.321	0.686
23	2.807	2.500	2.069	1.714	1.489	1.319	0.685
24	2.797	2.492	2.064	1.711	1.487	1.318	0.685
25	2.787	2.485	2.060	1.708	1.485	1.316	0.684
26	2.779	2.479	2.056	1.706	1.483	1.315	0.684
27	2.771	2.473	2.052	1.703	1.482	1.314	0.684
28	2.763	2.467	2.048	1.701	1.480	1.313	0.683
29	2.756	2.462	2.045	1.699	1.479	1.311	0.683
30	2.750	2.457	2.042	1.697	1.477	1.310	0.683
31	2.744	2.453	2.040	1.696	1.476	1.309	0.682
32	2.738	2.449	2.037	1.694	1.475	1.309	0.682
33	2.733	2.445	2.035	1.692	1.474	1.308	0.682
34	2.728	2.441	2.032	1.691	1.473	1.307	0.682
35	2.724	2.438	2.030	1.690	1.472	1.306	0.682
36	2.719	2.434	2.028	1.688	1.471	1.306	0.681
37	2.715	2.431	2.026	1.687	1.470	1.305	0.681
38	2.711	2.429	2.024	1.686	1.469	1.304	0.681
39	2.708	2.426	2.023	1.685	1.468	1.304	0.681
40	2.704	2.423	2.021	1.684	1.468	1.303	0.681
50	2.678	2.403	2.009	1.676	1.462	1.299	0.679
60	2.660	2.390	2.000	1.671	1.458	1.296	0.679
70	2.648	2.381	1.994	1.667	1.456	1.294	0.678
80	2.639	2.374	1.990	1.664	1.453	1.292	0.678
90	2.632	2.368	1.987	1.662	1.452	1.291	0.677
100	2.626	2.364	1.984	1.660	1.451	1.290	0.677
150	2.609	2.351	1.976	1.655	1.447	1.287	0.676
200	2.601	2.345	1.972	1.653	1.445	1.286	0.676
300	2.592	2.339	1.968	1.650	1.443	1.284	0.675
400	2.588	2.336	1.966	1.649	1.442	1.284	0.675
600	2.584	2.333	1.964	1.647	1.441	1.283	0.675
800	2.582	2.331	1.963	1.647	1.441	1.283	0.675
1000	2.581	2.330	1.962	1.646	1.441	1.282	0.675
∞	2.576	2.326	1.960	1.645	1.440	1.282	0.674

Tabelle 5: Quantile der χ^2-Verteilung

Tabelliert sind p-Quantile $\chi^2_{\nu,p}$ einer χ^2-Verteilung mit ν Freiheitsgraden.

	p												
ν	0.995	0.990	0.975	0.950	0.900	0.750	0.500	0.250	0.100	0.050	0.025	0.010	0.005
1	7.879	6.635	5.024	3.841	2.706	1.323	0.455	0.102	0.016	0.004	0.001	0.000	0.000
2	10.60	9.210	7.378	5.991	4.605	2.773	1.386	0.575	0.211	0.103	0.051	0.020	0.010
3	12.84	11.34	9.348	7.815	6.251	4.108	2.366	1.213	0.584	0.352	0.216	0.115	0.072
4	14.86	13.28	11.14	9.488	7.779	5.385	3.357	1.923	1.064	0.711	0.484	0.297	0.207
5	16.75	15.09	12.83	11.07	9.236	6.626	4.351	2.675	1.610	1.145	0.831	0.554	0.412
6	18.55	16.81	14.45	12.59	10.64	7.841	5.348	3.455	2.204	1.635	1.237	0.872	0.676
7	20.28	18.48	16.01	14.07	12.02	9.037	6.346	4.255	2.833	2.167	1.690	1.239	0.989
8	21.95	20.09	17.53	15.51	13.36	10.22	7.344	5.071	3.490	2.733	2.180	1.647	1.344
9	23.59	21.67	19.02	16.92	14.68	11.39	8.343	5.899	4.168	3.325	2.700	2.088	1.735
10	25.19	23.21	20.48	18.31	15.99	12.55	9.342	6.737	4.865	3.940	3.247	2.558	2.156
11	26.76	24.73	21.92	19.68	17.28	13.70	10.34	7.584	5.578	4.575	3.816	3.053	2.603
12	28.30	26.22	23.34	21.03	18.55	14.85	11.34	8.438	6.304	5.226	4.404	3.571	3.074
13	29.82	27.69	24.74	22.36	19.81	15.98	12.34	9.299	7.041	5.892	5.009	4.107	3.565
14	31.32	29.14	26.12	23.68	21.06	17.12	13.34	10.17	7.790	6.571	5.629	4.660	4.075
15	32.80	30.58	27.49	25.00	22.31	18.25	14.34	11.04	8.547	7.261	6.262	5.229	4.601
16	34.27	32.00	28.85	26.30	23.54	19.37	15.34	11.91	9.312	7.962	6.908	5.812	5.142
17	35.72	33.41	30.19	27.59	24.77	20.49	16.34	12.79	10.09	8.672	7.564	6.408	5.697
18	37.16	34.81	31.53	28.87	25.99	21.60	17.34	13.68	10.86	9.390	8.231	7.015	6.265
19	38.58	36.19	32.85	30.14	27.20	22.72	18.34	14.56	11.65	10.12	8.907	7.633	6.844
20	40.00	37.57	34.17	31.41	28.41	23.83	19.34	15.45	12.44	10.85	9.591	8.260	7.434
21	41.40	38.93	35.48	32.67	29.62	24.93	20.34	16.34	13.24	11.59	10.28	8.897	8.034
22	42.80	40.29	36.78	33.92	30.81	26.04	21.34	17.24	14.04	12.34	10.98	9.542	8.643
23	44.18	41.64	38.08	35.17	32.01	27.14	22.34	18.14	14.85	13.09	11.69	10.20	9.260
24	45.56	42.98	39.36	36.42	33.20	28.24	23.34	19.04	15.66	13.85	12.40	10.86	9.886
25	46.93	44.31	40.65	37.65	34.38	29.34	24.34	19.94	16.47	14.61	13.12	11.52	10.52
26	48.29	45.64	41.92	38.89	35.56	30.43	25.34	20.84	17.29	15.38	13.84	12.20	11.16
27	49.65	46.96	43.19	40.11	36.74	31.53	26.34	21.75	18.11	16.15	14.57	12.88	11.81
28	50.99	48.28	44.46	41.34	37.92	32.62	27.34	22.66	18.94	16.93	15.31	13.56	12.46
29	52.34	49.59	45.72	42.56	39.09	33.71	28.34	23.57	19.77	17.71	16.05	14.26	13.12
30	53.67	50.89	46.98	43.77	40.26	34.80	29.34	24.48	20.60	18.49	16.79	14.95	13.79
40	66.77	63.69	59.34	55.76	51.81	45.62	39.34	33.66	29.05	26.51	24.43	22.16	20.71
50	79.49	76.15	71.42	67.50	63.17	56.33	49.33	42.94	37.69	34.76	32.36	29.71	27.99
60	91.95	88.38	83.30	79.08	74.40	66.98	59.33	52.29	46.46	43.19	40.48	37.48	35.53
70	104.2	100.4	95.02	90.53	85.53	77.58	69.33	61.70	55.33	51.74	48.76	45.44	43.28
80	116.3	112.3	106.6	101.9	96.58	88.13	79.33	71.14	64.28	60.39	57.15	53.54	51.17
90	128.3	124.1	118.1	113.1	107.6	98.65	89.33	80.62	73.29	69.13	65.65	61.75	59.20
100	140.2	135.8	129.6	124.3	118.5	109.1	99.33	90.13	82.36	77.93	74.22	70.06	67.33
150	198.4	193.2	185.8	179.6	172.6	161.3	149.3	138.0	128.3	122.7	118.0	112.7	109.1
200	255.3	249.4	241.1	234.0	226.0	213.1	199.3	186.2	174.8	168.3	162.7	156.4	152.2
250	311.3	304.9	295.7	287.9	279.1	264.7	249.3	234.6	221.8	214.4	208.1	200.9	196.2
300	366.8	359.9	349.9	341.4	331.8	316.1	299.3	283.1	269.1	260.9	253.9	246.0	240.7
400	476.6	468.7	457.3	447.6	436.6	418.7	399.3	380.6	364.2	354.6	346.5	337.2	330.9
600	693.0	683.5	669.8	658.1	644.8	623.0	599.3	576.3	556.1	544.2	534.0	522.4	514.5
800	906.8	896.0	880.3	866.9	851.7	826.6	799.3	772.7	749.2	735.4	723.5	709.9	700.7
1000	1119	1107	1090	1075	1058	1030	999.3	969.5	943.1	927.6	914.3	898.9	888.6

Tabelle 6: Quantile der F-Verteilung

Tabelliert sind p-Quantile $F_{n,m,p}$ einer $F_{n,m}$-Verteilung mit n Zähler- und m Nennerfreiheitsgraden. Es gilt $F_{n,m,1-p} = \dfrac{1}{F_{m,n,p}}$.

m	p	n										
		1	2	3	4	5	10	15	25	50	100	∞
1	0.990	4052	4999	5404	5624	5764	6056	6157	6240	6302	6334	6366
	0.975	647.8	799.5	864.2	899.6	921.8	968.6	984.9	998.1	1008	1013	1018
	0.950	161.4	199.5	215.7	224.6	230.2	241.9	245.9	249.3	251.8	253.0	254.3
	0.900	39.86	49.50	53.59	55.83	57.24	60.19	61.22	62.05	62.69	63.01	63.33
2	0.990	98.50	99.00	99.16	99.25	99.30	99.40	99.43	99.46	99.48	99.49	99.50
	0.975	38.51	39.00	39.17	39.25	39.30	39.40	39.43	39.46	39.48	39.49	39.50
	0.950	18.51	19.00	19.16	19.25	19.30	19.40	19.43	19.46	19.48	19.49	19.50
	0.900	8.526	9.000	9.162	9.243	9.293	9.392	9.425	9.451	9.471	9.481	9.491
3	0.990	34.12	30.82	29.46	28.71	28.24	27.23	26.87	26.58	26.35	26.24	26.13
	0.975	17.44	16.04	15.44	15.10	14.88	14.42	14.25	14.12	14.01	13.96	13.90
	0.950	10.13	9.552	9.277	9.117	9.013	8.785	8.703	8.634	8.581	8.554	8.526
	0.900	5.538	5.462	5.391	5.343	5.309	5.230	5.200	5.175	5.155	5.144	5.134
4	0.990	21.20	18.00	16.69	15.98	15.52	14.55	14.20	13.91	13.69	13.58	13.46
	0.975	12.22	10.65	9.979	9.604	9.364	8.844	8.657	8.501	8.381	8.319	8.257
	0.950	7.709	6.944	6.591	6.388	6.256	5.964	5.858	5.769	5.699	5.664	5.628
	0.900	4.545	4.325	4.191	4.107	4.051	3.920	3.870	3.828	3.795	3.778	3.761
5	0.990	16.26	13.27	12.06	11.39	10.97	10.05	9.722	9.449	9.238	9.130	9.020
	0.975	10.01	8.434	7.764	7.388	7.146	6.619	6.428	6.268	6.144	6.080	6.015
	0.950	6.608	5.786	5.409	5.192	5.050	4.735	4.619	4.521	4.444	4.405	4.365
	0.900	4.060	3.780	3.619	3.520	3.453	3.297	3.238	3.187	3.147	3.126	3.105
10	0.990	10.04	7.559	6.552	5.994	5.636	4.849	4.558	4.311	4.115	4.014	3.909
	0.975	6.937	5.456	4.826	4.468	4.236	3.717	3.522	3.355	3.221	3.152	3.080
	0.950	4.965	4.103	3.708	3.478	3.326	2.978	2.845	2.730	2.637	2.588	2.538
	0.900	3.285	2.924	2.728	2.605	2.522	2.323	2.244	2.174	2.117	2.087	2.055
15	0.990	8.683	6.359	5.417	4.893	4.556	3.805	3.522	3.278	3.081	2.977	2.868
	0.975	6.200	4.765	4.153	3.804	3.576	3.060	2.862	2.689	2.549	2.474	2.395
	0.950	4.543	3.682	3.287	3.056	2.901	2.544	2.403	2.280	2.178	2.123	2.066
	0.900	3.073	2.695	2.490	2.361	2.273	2.059	1.972	1.894	1.828	1.793	1.755
25	0.990	7.770	5.568	4.675	4.177	3.855	3.129	2.850	2.604	2.400	2.289	2.169
	0.975	5.686	4.291	3.694	3.353	3.129	2.613	2.411	2.230	2.079	1.996	1.906
	0.950	4.242	3.385	2.991	2.759	2.603	2.236	2.089	1.955	1.842	1.779	1.711
	0.900	2.918	2.528	2.317	2.184	2.092	1.866	1.771	1.683	1.607	1.565	1.518
50	0.990	7.171	5.057	4.199	3.720	3.408	2.698	2.419	2.167	1.949	1.825	1.683
	0.975	5.340	3.975	3.390	3.054	2.833	2.317	2.109	1.919	1.752	1.656	1.545
	0.950	4.034	3.183	2.790	2.557	2.400	2.026	1.871	1.727	1.599	1.525	1.438
	0.900	2.809	2.412	2.197	2.061	1.966	1.729	1.627	1.529	1.441	1.388	1.327
100	0.990	6.895	4.824	3.984	3.513	3.206	2.503	2.223	1.965	1.735	1.598	1.427
	0.975	5.179	3.828	3.250	2.917	2.696	2.179	1.968	1.770	1.592	1.483	1.347
	0.950	3.936	3.087	2.696	2.463	2.305	1.927	1.768	1.616	1.477	1.392	1.283
	0.900	2.756	2.356	2.139	2.002	1.906	1.663	1.557	1.453	1.355	1.293	1.214
∞	0.990	6.635	4.605	3.782	3.319	3.017	2.321	2.039	1.773	1.523	1.358	1.000
	0.975	5.024	3.689	3.116	2.786	2.566	2.048	1.833	1.626	1.428	1.296	1.000
	0.950	3.841	2.996	2.605	2.372	2.214	1.831	1.666	1.506	1.350	1.243	1.000
	0.900	2.706	2.303	2.084	1.945	1.847	1.599	1.487	1.375	1.263	1.185	1.000

Anhang B

Altgriechisches Alphabet

In der Statistik und Mathematik werden häufig Buchstaben des altgriechischen Alphabets verwendet.

Altgriechisches Alphabet			
Deutscher Name	englischer Name	Kleinbuchstabe	Großbuchstabe
Alpha	alpha	α	A
Beta	beta	β	B
Gamma	gamma	γ	Γ
Delta	delta	δ, ∂	Δ
Epsilon	epsilon	ϵ, ε	E
Zeta	zeta	ζ	Z
Eta	eta	η	H
Theta	theta	θ, ϑ	Θ
Iota	iota	ι	I
Kappa	kappa	κ	K
Lambda	lambda	λ	Λ
My	mu	μ	M
Ny	nu	ν	N
Xi	xi	ξ	Ξ
Omikron	omicron	o	O
Pi	pi	π	Π
Rho	rho	ρ, ϱ	P
Sigma	sigma	σ, ς	Σ
Tau	tau	τ	T
Ypsilon	upsilon	υ	Υ, Y
Phi	phi	ϕ, φ	Φ
Chi	chi	χ	X
Psi	psi	ψ	Ψ
Omega	omega	ω	Ω

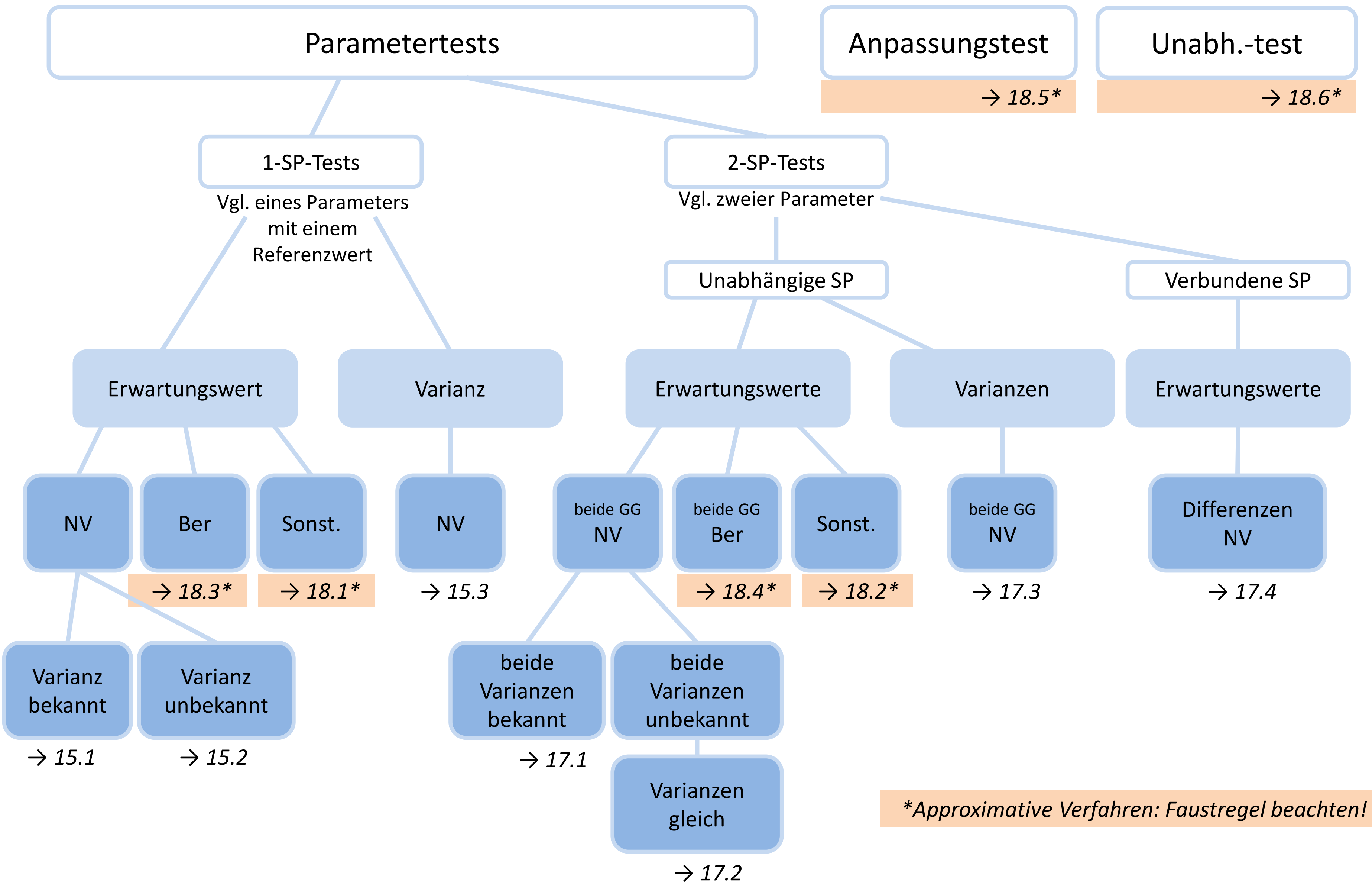
- Die Schreibvariante ∂ des griechischen Kleinbuchstaben Delta wird häufig als Symbol für partielle Differentiale verwendet und dann manchmal auch „del" gesprochen.
  Beispiel: Für die Funktion $f(x,y) = x^2y + 2y^2$ sind
  $$\frac{\partial f(x,y)}{\partial x} = 2xy \quad \text{und} \quad \frac{\partial f(x,y)}{\partial y} = x^2 + 4y$$
  die partiellen Ableitungen nach x und y.
- Die Schreibvariante $\in$ des griechischen Kleinbuchstaben Epsilon wird für das Elementzeichen bei der Mengennotation verwendet.
  Beispiel: Für die Menge $A = \{1,2,3\}$ gilt $2 \in A$ und $0 \notin A$.
- Der griechische Kleinbuchstabe Theta, θ, wird in der induktiven Statistik häufig verwendet, um den unbekannten Parameter einer Wahrscheinlichkeitsverteilung zu bezeichnen.
- Die Schreibvariante ς für den griechischen Kleinbuchstaben Sigma, die im Altgriechischen nur am Wortende verwendet wird, ist in der Statistik und Mathematik eher ungebräuchlich.

- Die Schreibvariante $\sum$ des griechischen Großbuchstaben Sigma wird für das Summenzeichen verwendet. Beispiel:
$$\sum_{i=1}^{3} i^2 = 1^2 + 2^2 + 3^2 = 1 + 4 + 9 = 14.$$

- Der griechische Kleinbuchstabe Pi, π, wird häufig zur Bezeichnung der Kreiskonstante verwendet, die das Verhältnis vom Umfang U zum Durchmesser D eines Kreises angibt, es gilt
$$D\pi = U.$$
In der Statistik wird π aber auch zur Bezeichnung von Wahrscheinlichkeiten und als Symbol für den Parameter einer Poisson-Verteilung verwendet.

- Die Schreibvariante $\prod$ des griechischen Großbuchstaben Pi wird für das Produktzeichen verwendet. Beispiel:
$$\prod_{i=1}^{3} i^2 = 1^2 \cdot 2^2 \cdot 3^2 = 1 \cdot 4 \cdot 9 = 36.$$

Testverfahren



D. Tillich, C. Wobar: 2015