Elementare Stochastik

Universität Regensburg

(Dr. Werner Stich)

Julian Palme

(Stand: 9. Juli 2014)

Dies ist eine Mitschrift der Vorlesung 'Elementare Stochastik', gelesen von Herrn Dr. Werner Stich an der Universität Regensburg. Dieses Dokument ist KEIN offizielles Skript und wurde gesetzt in LaTeX von Julian Palme.

Inhaltsverzeichnis

Kapitel I: Beschreibende Statistik **1**

§ 1 Merkmale (Variablen) und Merkmalstypen **1**
1.1 Ein Beispiel . . . 1
1.2 Bezeichnungen . . . 1
1.3 Merkmalstyp, Messniveau . . . 2
1.4 Bezeichnung . . . 3
1.5 Notiz . . . 3
1.6 Bezeichnungen . . . 3
1.7 Häufigkeiten . . . 3

§ 2 Beschreibung eindimensionaler (univarianter) Datenreihen **5**
2.1 Datenreihe eines Experiments (Summe von 4 Würfeln) . . . 5
2.2 Statistische Maßzahlen . . . 5
2.3 Notiz . . . 6
2.4 arithmetisches Mittel (Def) . . . 6
2.5 Notiz . . . 7
2.6 Beispiel . . . 7
2.7 Beispiel . . . 7
2.8 Physikalische Interpretation des arithmetischen Mittels . . . 7
2.9 Notiz . . . 8
2.10 Linearitätseigenschaft des arithmetischen Mittels (Satz) . . . 8
2.11 Minimalitätseigenschaft des arithmetischen Mittels bzgl. der quadratischen Abweichung (Satz) . . . 9
2.12 Stichproben-Varianz, Standardabweichung, Streuungsmaßzahlen (Def) . . . 9
2.13 Notiz . . . 9
2.14 Beispiel . . . 10
2.15 Notiz . . . 10
2.16 Satz . . . 10
2.17 Beispiel . . . 11
2.18 empirische Verteilungsfunktion, kumulierte relative Häufigkeit (Def) . . . 11
2.19 Beispiel . . . 11
2.20 empirischer Median, Zentralwert, zentrales Wertepaar (Def) . . . 12
2.21 Median, Zentralwert, zentrales Wertepaar (Bsp) . . . 12
2.22 Notiz . . . 12
2.23 Grafik zum Median . . . 13
2.24 Minimalitätseigenschaft des Median bzgl. der Abweichung zum Betrag (Satz) 14
2.25 α-Quantil, r-tes Perzentil (Def) . . . 14
2.26 Notiz und Definition . . . 14
2.27 Skizze zur groben Orientierung . . . 15
2.28 Grafiken zu den Quantilen . . . 15
2.29 Vergleich arithmetisches Mittel – Median . . . 16
2.30 geometrisches und harmonisches Mittel (Def) . . . 17
2.31 Bemerkung . . . 18
2.32 Bemerkung . . . 18
2.33 Beispiel . . . 18
2.34 arithmetisches Mittel, geometrisches Mittel (Bsp) . . . 19
2.35 Modalwert (Def) . . . 20

2.36 gewichtetes arithmetisches Mittel für I- oder V-Skala (Def) . . . 20
2.37 Berechnungsvariante für die Stichprobenvarianz (Bem) . . . 20
2.38 Boxplots . . . 21

§ 3 Beschreibung zweidimensionaler Datenreihen 22
3.1 Beispiel . . . 22
3.2 Graphische Darstellung . . . 22
3.3 Beispiele . . . 23
3.4 Pearsonscher Korrelationskoeffizient (Def) . . . 24
3.5 Notiz . . . 24
3.6 Bemerkungen . . . 25
3.7 Verhalten von r bei linearen Transformationen der Messwerte (Bem) . . . 25
3.8 Skizze zu 3.7 . . . 25
3.9 Bemerkung zur Interpretation von r . . . 26
3.10 Lineare Regression . . . 26
3.11 Modellvorstellung zur linearen Regression . . . 26
3.12 Zur Bestimmung von a und b der Regressionsgerade . . . 27
3.13 Bemerkung (für $n \geq 2$) . . . 28
3.14 Linearisierung durch Transformation . . . 29
3.15 Anmerkung . . . 30
3.16 Kontingenzkoeffizient . . . 31
3.17 Beispiel . . . 31
3.18 Vorüberlegungen . . . 31
3.19 Kontingenzkoeffizient (Def) . . . 32
3.20 Bemerkung . . . 32

Kapitel II: Wahrscheinlichkeitsrechnung 33

§ 4 Der Begriff Wahrscheinlichkeit 33
4.1 Bezeichnungen . . . 33
4.2 Würfel (Bsp) . . . 33
4.3 Relative Häufigkeiten (Def) . . . 33
4.4 Bemerkungen . . . 34
4.5 Anmerkungen . . . 34
4.6 Wahrscheinlichkeit (Def) . . . 34
4.7 Geometrische Analogie für die Wahrscheinlichkeitsverteilung . . . 35
4.8 Ω endlich (Satz) . . . 35

§ 5 Gleichverteilung und Zufallsvariable 37
5.1 Gleichverteilung/Laplace-Verteilung (Def) . . . 37
5.2 Beispiel . . . 37
5.3 Zufallsvariable/Zufallsgröße, Ausprägungen (Def) . . . 37
5.4 Notiz . . . 38
5.5 Graphische Darstellung . . . 38
5.6 Bemerkung . . . 39
5.7 Beispiel . . . 39

§ 6 Urnenmodelle 42
6.1 Regeln mit Ergebnismengen . . . 42
6.2 Bemerkung . . . 42
6.3 Fakultät, Binomialkoeffizient (Def) . . . 43

6.4 Bemerkung . . . 43
6.5 Satz . . . 43
6.6 Hinweis . . . 44

§ 7 Baumdiagramme und Pfadregeln 45
7.1 Beispiel . . . 45
7.2 Pfadmultiplikationsregel . . . 45
7.3 Pfadadditionsregel . . . 45
7.4 Beispiel zu 7.1 . . . 45

§ 8 Bedingte Wahrscheinlichkeiten und unabhängige Ereignisse 46
8.1 Bedingte Wahrscheinlichkeit (Def) . . . 46
8.2 Beispiel . . . 46
8.3 Multiplikationssatz (Satz) . . . 46
8.4 Bemerkung . . . 47
8.5 stochastisch unabhängig (Def) . . . 47
8.6 Satz . . . 47
8.7 Bemerkung . . . 47
8.8 Disjunktion (Zerlegung) von Ω (Def) . . . 47
8.9 Satz . . . 47
8.10 Beispiel . . . 47
8.11 Bayessche Formel (Satz) . . . 48
8.12 Beispiel . . . 48

§ 9 Verteilungen 49
9.1 Ziehen ohne Zurücklegen (Def) . . . 49
9.2 Notiz . . . 49
9.3 Bemerkung . . . 49
9.4 Beispiel . . . 49
9.5 Grafiken zur Binomialverteilung . . . 51
9.6 hypergeometrisch verteilte Zufallsvariable (Def) . . . 51
9.7 Bemerkung . . . 51
9.8 Beispiel . . . 51
9.9 stetig verteilte Zufallsvariable (Def) . . . 51
9.10 Veranschaulichung Wahrscheinlichkeiten $\longleftrightarrow$ Verteilungsfunktion . . . 52
9.11 (standard-)normalverteilte Zufallsvariable (Def) . . . 53
9.12 Bemerkung . . . 54
9.13 Skizze zur Dichtefunktion . . . 54
9.14 Bemerkung . . . 54
9.15 Tabelle zu Werten von Φ . . . 55
9.16 Beispiel . . . 55
9.17 Approximation der Binomialverteilung durch die Normalverteilung (Bem) . . . 55
9.18 Beispiel . . . 55

§ 10 Erwartungswert, Varianz, Covarianz 56
10.1 Erwartungswert (Def) . . . 56
10.2 Definition . . . 56
10.3 Rechenregeln für Erwartungswerte (Bem) . . . 56
10.4 Notiz . . . 56
10.5 Varianz (Def) . . . 57
10.6 Bemerkung . . . 57
10.7 Standardabweichung (Def) . . . 57

10.8 Rechenregeln für die Varianz (Bem) . . . 57
10.9 Notiz . . . 57
10.10 Definition und Notiz . . . 58
10.11 Covarianz, Korrelationskoeffizient (Def) . . . 58
10.12 Notiz . . . 58
10.13 Rechenregeln im Zusammenhang mit der Covarianz (Satz) . . . 58
10.14 Beispiel . . . 59
10.15 Satz . . . 60

Kapitel III: Schließende Statistik **61**

§ 11 Signifikanztests **61**
11.1 Signifikanztest, Nullhypothese (Def) . . . 61
11.2 Notiz und Definition . . . 61
11.3 Beispiel . . . 61

Anhang A Grafiken zur Binomialverteilung **64**

Anhang B Normalverteilungstabelle **66**

Stichwortverzeichnis

Kapitel I:
Beschreibende Statistik

§ 1 Merkmale (Variablen) und Merkmalstypen

1.1 Ein Beispiel

Bei einer Befragung von Hörern einer Vorlesung werden folgende Daten ermittelt:

(1) Familienstand

(2) Studienrichtung

(3) Interesse am Vorlesungsgegenstand
(außerordentlich interessiert, sehr interessiert, interessiert, kaum interessiert, überhaupt nicht interessiert)

(4) Anzahl der Geschwister

(5) Anzahl der bereits studierten Hochschulsemester

(6) Körpergröße

(7) Gewicht

Zusammengehörigkeit der Merkmale (vom Typ her)

keine Ordnung vorgegeben	Ordnung	Zahlen diskret	Zahlen kontinuierlich
(1), (2)	(3)	(4), (5)	(6), (7)

1.2 Bezeichnungen

(Beobachtungs-)Merkmal: (gefragte) Eigenschaft, die beobachtet oder gemessen wird (z. B. Familienstand, Semesteranzahl, Körpergröße ...)

(Merkmals-)Ausprägung: Werte (Ergebnisse), die bei einem Merkmal auftreten (z. B. ledig, 8 Semester, 171 cm)

Im mathematischen Modell sind Ausprägungen die Werte, die eine Zufallsvariable annehmen kann.

1.3 Merkmalstyp, Messniveau

Merkmalstyp	**Messniveau**	**Beispiele**
qualitatives Merkmal	Nominalskala	(1), (2), Geschlecht, Farbe, Bundesland
Rangmerkmale	Ordinalskala	(3), Rangplätze bei Wettbewerben
quantitatives Merkmal (metrisches Merkmal)	Intervallskala	Nullpunkt NICHT absolut festgelegt, z. B. $(*)$
diskret / stetig	Verhältnisskala	Nullpunkt eindeutig festgelegt, z. B. $(**)$

$(*)$ Bei **Intervallskalen** sind Addition und Subtraktion der Messwerte zur Bestimmung eines Mittelwertes sinnvoll, NICHT jedoch die Multiplikation und Division der Werte untereinander.

Beispiel: Messreihe in Celsius und Fahrenheit

Zwischen Celsius und Fahrenheit gelten folgende Umrechnungsformeln (lineare Transformation):

$$t\,^\circ\mathrm{F} = 1.8 \cdot t\,^\circ\mathrm{C} + 32 \qquad ; \qquad t\,^\circ\mathrm{C} = (t\,^\circ\mathrm{F} - 32) : 1.8$$

Eine mögliche Messreihe wäre

MRC	10°	12°	20°	°C
MRF	50°	53.6°	68°	°F

arithmetisches Mittel zu MRC (Messreihe in Celsius):

$$14^\circ\mathrm{C} \longrightarrow 1.8 \cdot 14^\circ\mathrm{C} + 32 = 57.2^\circ\mathrm{F}$$

arithmetisches Mittel zu MRF (Messreihe in Fahrenheit):

$$57.2^\circ\mathrm{F}$$

Die beiden Werte stimmen überein.

Anders bei der Berechnung des geometrischen Mittels (Definition siehe später): Geometrisches Mittel zu MRC:

$$\sqrt[3]{10 \cdot 12 \cdot 20} \approx 13.388^\circ\mathrm{C} \longrightarrow \approx 56.09^\circ\mathrm{F}$$

Geometrisches Mittel zu MRF:

$$\sqrt[3]{50 \cdot 53.6 \cdot 68} \approx 56.695^\circ\mathrm{F}$$

$(**)$ Bei **Verhältnisskalen** sind auch Multiplikation und Division der Messwerte untereinander sinnvoll.

1.4 Bezeichnung

Merkmale mit nur zwei Ausprägungen (z. B. ja–nein, Erfolg–Misserfolg) heißen auch **dichotone Merkmale**.

1.5 Notiz

Quantitative Merkmale können auch in der Form von Nominal- oder Ordinalskalen erfasst werden.

Beispiel

Würmer werden in drei Gruppen eingeteilt

$$\begin{aligned}
&\text{Wert} < 7 \;\longmapsto\; 1 \text{ klein}\\
&7 \leq \text{Wert} \leq 10 \;\longmapsto\; 2 \text{ mittel}\\
&\text{Wert} > 10 \;\longmapsto\; 3 \text{ groß}
\end{aligned}$$

1.6 Bezeichnungen

M sei eine Menge von Individuen oder Objekten. Die Erhebung eines Merkmals kann man als Abbildung

$$X\colon M \longrightarrow A$$

auffassen, wobei A die Ausprägungen bezeichnet. Nummeriert man die Elemente von M von $1, \ldots, n$, so ist

$$X\colon \{1, \ldots, n\} \longrightarrow A\ ,\ i \longmapsto X(i) \in A$$

Statt $X(i)$ schreibt man auch x_i und nennt $x_1, \ldots, x_n$ Daten- oder Messreihe. Erfasst man mehrere Merkmale einer Menge M zusammen, so erhält man Abbildungen

$$X_1\colon \{1, \ldots, n\} \longrightarrow A_1\ ,\ X_2\colon \{1, \ldots, n\} \longrightarrow A_2 \quad \text{usw.}$$

Entsprechend erhält man Datenreihen

$$\begin{aligned}
&x_1^1, \ldots, x_n^1 \;\longrightarrow\; x_1, \ldots, x_n\\
&x_1^2, \ldots, x_n^2 \;\longrightarrow\; y_1, \ldots, y_n
\end{aligned}$$

1.7 Häufigkeiten

Sei $X\colon M \longrightarrow A$. Dann heißt

$$h(X = a) := \#\{X = a\} = \#\left\{m \in M \;\middle|\; X(m) = a\right\}$$

die **(absolute) Häufigkeit** der Ausprägung a und

$$r(X = a) := \frac{h(X = a)}{\#M}$$

die **relative Häufigkeit** von a.

Es gilt:

$$\sum_{a \in M} h(X = a) = \#M$$

$$\sum_{a \in M} r(X = a) = 1$$

§ 2 Beschreibung eindimensionaler (univarianter) Datenreihen

Sei $M = \{1, \ldots, n\}$ und $X\colon M \longrightarrow A = \mathbb{R}$ eine Abbildung. Die Abbildung X definiert eine eindimensionale Datenreihe (Messreihe)

$$x_1, \ldots, x_n$$

(Die Datenreihe kann von einer Gesamterhebung stammen, n ist die Anzahl der Mitglieder des betrachteten Kollektivs M), oder die Datenreihe kann eine Stichprobe darstellen (n ist dann der Umfang der Stichprobe $M \subset M'$).
Bemerkung: Grundgesamtheit: $\#M' \gg \#M$

Die x_i sind Werte (Merkmalsausprägungen) einer (Zufalls-)Variablen (eines Merkmals), z. B. Ergebnis eines Würfelwurfs, Länge eines Wurmes ...
Für Merkmale mit einer O-, I- und V-Skala kann man die Werte $x_1, \ldots, x_n$ der Größe nach anordnen. Man erhält dann die geordnete Datenreihe.

2.1 Datenreihe eines Experiments (Summe von 4 Würfeln)

$X\colon \{1, \ldots, 20\} \longrightarrow \{4, \ldots, 24\} \subset \mathbb{R}$.

i	1	2	3	4	5	6	7	8	9	10
x_i	16	12	23	13	13	15	13	9	16	11

i	11	12	13	14	15	16	17	18	19	20
x_i	12	12	14	7	7	11	12	14	13	7

Dazu gehörige geordnete Datenreihe:

i	1	2	3	4	5	6	7	8	9	10
x_i	7	7	7	9	11	11	12	12	12	12

i	11	12	13	14	15	16	17	18	19	20
x_i	13	13	13	13	14	14	15	16	16	23

Absolute und relative Häufigkeiten:

a	7	8	9	10	11	12	13	14	15	16	...	23	24	$\sum$
$h(X=a)$	3	0	1	0	2	4	4	2	1	2	...	1	0	20
$r(X=a)$	$\frac{3}{20}$	0	$\frac{1}{20}$	0	$\frac{1}{10}$	$\frac{1}{5}$	$\frac{1}{5}$	$\frac{1}{10}$	$\frac{1}{20}$	$\frac{1}{10}$	...	$\frac{1}{20}$	0	1

2.2 Statistische Maßzahlen

Statistische Maßzahlen dienen

- für eine knappe Beschreibung einer Datenreihe (sinnvoll bei Stichproben und Gesamterhebungen)
- bei Stichproben auch als Schätzwerte unbekannter Parameter der zugrunde liegenden Grundgesamtheit (der Population, des Experiments)

Es gibt Maßzahlen, die insbesondere die Lage und andere, die die Streuung einer Datenreihe beschreiben (gilt nur für O-, I- und V-Skala).

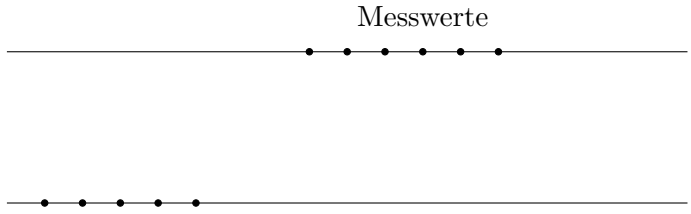


Abbildung 2.1: *unterschiedliche Lage (des Zentrums der Messwerte), Streuung der Messwerte etwa gleich*

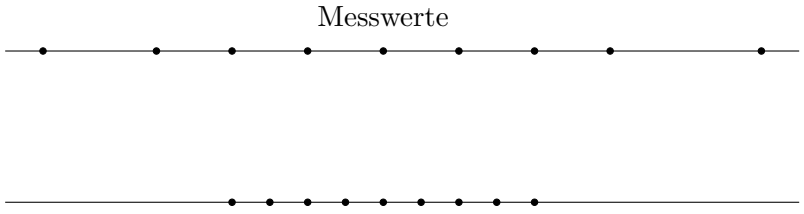


Abbildung 2.2: *Zentrum der Messwerte etwa gleich, Streuung der Messwerte verschieden*

2.3 Notiz

Für eine Datenreihe $x_1, \ldots, x_n$ vermitteln das Maximum, Minimum und die Spannweite (=Differenz zwischen größtem und kleinstem Wert) einen Eindruck über die Bandbreite der Messreihe.

Minimum und Maximum sind auch für eine grobe Überprüfung der Datenreihe nützlich (Eingabefehler, die zu unrealistischen oder falschen Werten führen, können schnell ermittelt werden).

2.4 Definition (arithmetisches Mittel)

Sei $M = \{1, \ldots, n\}$, $X \colon M \longrightarrow A$ eine Abbildung und $A \subset \mathbb{R}$ (z. B. für die Ausprägungen einer Intervall- oder Verhältnisskala). Zu X gehört die Datenreihe $x_1, \ldots, x_n$. Dann heißt

$$\bar{x} := \frac{1}{n}(x_1 + \ldots + x_n)$$

das **arithmetische Mittel**[1] der Datenreihe. Es ist die am häufigsten verwendete Maßzahl, um den mittleren Wert einer Datenreihe zu charakterisieren.

[1] *englisch* arithmetic mean

2.5 Notiz

Für $A = \{a_1, \ldots, a_n\}$ kann das arithmetische Mittel aus den absoluten oder relativen Häufigkeiten berechnet werden durch

(i) $\bar{x} = \frac{1}{n}\left(\sum_{j=1}^{k} a_j h(X = a_j)\right)$

(ii) $\bar{x} = \sum_{j=1}^{k} a_j r(X = a_j)$

2.6 Beispiel

Daten aus 2.1: $X = \{1, \ldots, 20\} \longrightarrow A = \{4, \ldots, 24\} \subset \mathbb{R}$.

$$\bar{x} = \frac{1}{20}(16 + 12 + \ldots + 7) = \frac{250}{20} = 12.5$$

zu (i): $\bar{x} = \frac{1}{20}\left(\sum_{j=4}^{24} j \cdot h(X = j)\right) = \frac{1}{20}(4 \cdot 0 + 5 \cdot 0 + 6 \cdot 0 + 7 \cdot 3 + 8 \cdot 0 + \ldots + 23 \cdot 1) = 12.5$

zu (ii): $\bar{x} = \sum_{j=4}^{24} j \cdot r(X = j) = 4 \cdot 0 + \ldots + 7 \cdot \frac{3}{20} + \ldots + 23 \cdot \frac{1}{20} = 12.5$

2.7 Beispiel

Versicherungsschäden in Tabellenform

a_i (in 1000 €)	$h(X = a_i)$
0	86
5	10
20	3
100	1

Arithmetisches Mittel

$$\bar{x} = \frac{1}{100}(0 \cdot 86 + 5 \cdot 10 + 20 \cdot 3 + 100 \cdot 1) = 2.1$$

$\bar{x}$ braucht man zur Berechnung der Prämie (+ Verwaltungskosten, Steuern usw.)

2.8 Physikalische Interpretation des arithmetischen Mittels

Sei $A = \{a_1, \ldots, a_k\}$. Dann gilt:

$$\begin{aligned}&(a_1 - \bar{x}) \cdot h(X = a_1) + \ldots + (a_k - \bar{x}) \cdot h(X = a_k) = \\ &= \underbrace{a_1 \cdot h(X = a_1) + \ldots + a_k \cdot h(X = a_k)}_{=n \cdot \bar{x}} - \bar{x} \underbrace{\left(h(X = a_1) + \ldots + h(X = a_k)\right)}_{=n} = 0\end{aligned}$$

Interpretation:

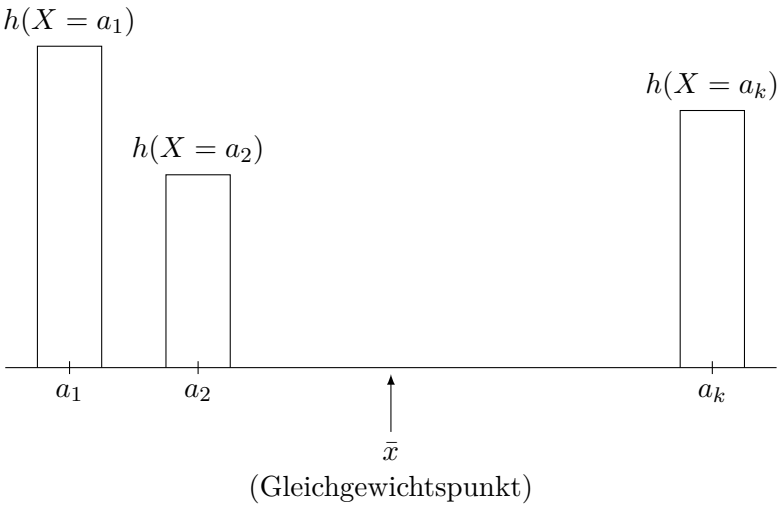


Abbildung 2.3: *Physikalische Interpretation des arithmetischen Mittels*

2.9 Notiz

Für $x_1, \ldots, x_n \in \mathbb{R}$ gilt:

$$\sum_{i=1}^{n}(x_i - \bar{x}) = 0$$

Beweis

$$\sum_{i=1}^{n}(x_i - \bar{x}) = \underbrace{\left(\sum_{i=1}^{n} x_i\right)}_{n\cdot\bar{x}} - n \cdot \bar{x}$$

□

2.10 Satz (Linearitätseigenschaft des arithmetischen Mittels)

Vor: Es seien $x_1, \ldots, x_n \in \mathbb{R}$ und $y_1, \ldots, y_n \in \mathbb{R}$ zwei Datenreihen, $a, b \in \mathbb{R}$.

Beh:

(a) $z_i = x_i + y_i \implies \bar{z} = \bar{x} + \bar{y}$

(b) $z_i = a \cdot x_i + b \implies \bar{z} = a \cdot \bar{x} + b$

Beweis

Übungsaufgabe

□

2.11 Satz (Minimalitätseigenschaft des arithmetischen Mittels bzgl. der quadratischen Abweichung)

Vor: Es sei $x_1, \ldots, x_n \in \mathbb{R}$. Dann ist:

Beh:

$$\sum_{i=1}^{n}(x_i - \bar{x})^2 \leq \sum_{i=1}^{n}(x_i - d)^2 \quad \forall\, d \in \mathbb{R}$$

Beweis

Sei $d \in \mathbb{R}$ vorgegeben.

$$\begin{aligned}
\sum_{i=1}^{n}(x_i - d)^2 &= \sum_{i=1}^{n}\left[(x_i - \bar{x}) + (\bar{x} - d)\right]^2 \\
&= \sum_{i=1}^{n}\left[(x_i - \bar{x})^2 + 2(x_i - \bar{x})(\bar{x} - d) + (\bar{x} - d)^2\right] \\
&= \sum_{i=1}^{n}(x_i - \bar{x})^2 + 2(\bar{x} - d) \cdot \underbrace{\sum_{i=1}^{n}(x_i - \bar{x})}_{\overset{2.9}{=}0} + \underbrace{n(\bar{x} - d)^2}_{\geq 0} \\
&\geq \sum_{i=1}^{n}(x_i - \bar{x})^2
\end{aligned}$$

□

2.12 Definition (Stichproben-Varianz, Standardabweichung, Streuungsmaßzahlen)

Es sei $x_1, \ldots, x_n \in \mathbb{R}$ eine Datenreihe einer I- oder V-Skala.

$$s^2 := \frac{1}{n-1} \cdot \sum_{i=1}^{n}(x_i - \bar{x})^2$$

heißt **Stichproben-Varianz** und

$$s := \sqrt{s^2}$$

heißt **Standardabweichung** der Datenreihe $x_1, \ldots, x_n$ bezüglich der Stichproben-Varianz. s^2 und s sind sogenannte **Streuungsmaßzahlen**.

2.13 Notiz

Liegen die Daten (Messwerte) in Form einer Häufigkeitstabelle vor, wobei $A = \{a_1, \ldots, a_k\}$, so gilt:

$$s^2 = \frac{1}{n-1} \cdot \sum_{j=1}^{k}(a_j - \bar{x})^2 \cdot h(X = a_j)$$

2.14 Beispiel

Datenreihe aus 2.1:
Stichproben-Varianz ($\bar{x} = 12.5$ nach 2.6)

$$s^2 = \frac{1}{19}(12 - 12.5)^2 + (17 - 12.5)^2 + \ldots + (7 - 12.5)^2) \approx 13.42$$

Dazu gehörige Standardabweichung:

$$s \approx \sqrt{13.42} \approx 3.66$$

2.15 Notiz

(a) s ist ein sogenannter **erwartungstreuer Schätzwert** für die Standardabweichung einer Grundgesamtheit. Bei Gesamterhebungen und in manchen Lehrbüchern wird

$$s'^2 := \frac{1}{n} \cdot \sum_{i=1}^{n} (x_i - \bar{x})^2$$

als Varianz und

$$s' := \sqrt{s'^2}$$

als Standardabweichung verwendet.

(b) Umrechung:

$$s'^2 = \frac{n-1}{n} \cdot s^2 \quad , \quad s' = \sqrt{s'^2} = \sqrt{\frac{n-1}{n}} \cdot s$$

Bei sehr großen n ist der Unterschied zwischen s und s' vernachlässigbar, z. B. ist $n = 1000 \implies \frac{n-1}{n} = 0.999$, $\sqrt{\frac{n-1}{n}} = 0.9995$.

2.16 Satz

Vor: Es sei $x_1, \ldots, x_n \in \mathbb{R}$ eine Datenreihe und $z_i = ax_i + b$ $(a, b \in \mathbb{R})$

Beh: $s_z^2 = a^2 s_x^2$ und $s_z = |a| \cdot s_x$

Anmerkung:
$a = 1$ bedeutet, dass die Datenreihe x_i um die Konstante b verschoben wird, d. h. das arithmetische Mittel ändert sich um den Wert b, Varianz und Standardabweichung ändern sich NICHT.

Beweis

Übungsaufgabe.

□

2.17 Beispiel

$x_1, \ldots, x_n$ Temperaturen in °C, $\bar{x}$ arithmetisches Mittel, s_x^2 Varianz, s_x Standardabweichung zur Datenreihe x_i. Dann ist ($t°\mathrm{F} = 1.8 \cdot t°\mathrm{C} + 32$)

$\bar{z} = 1.8\bar{x} + 32$ das arithmetische Mittel in °F

$s_z^2 = 1.8^2 \cdot s_x^2$ die Varianz in °F

$s_z = 1.8 \cdot s_x$ die Standardabweichung in °F

2.18 Definition (empirische Verteilungsfunktion, kumulierte relative Häufigkeit)

X sei ein Merkmal mit Ausprägung $A = \{a_1, \ldots, a_k\}$ mit $a_1 < a_2 < \ldots < a_k$
($X\colon M \longrightarrow A$, M Individuen, Objekte)
Die Funktion $F\colon \mathbb{R} \longrightarrow [0, 1]$ definiert durch

$$F(x) = r(X \leq x) := \sum_{a_j \leq x} r(X = a_j)$$

heißt **(empirische) Verteilungsfunktion** des Merkmals X.

$r(X \leq x)$ heißt **kumulierte relative Häufigkeit**.

2.19 Beispiel

Note	1	2	3	4	5	6	
absolute Häufigkeit	3	5	9	6	2	0	$n = 25$
relative Häufigkeit	0.12	0.20	0.36	0.24	0.08	0	$\sum = 1.00$

x	< 1	< 2	< 3	< 4	< 5	≥ 5
$F(x)$	0	0.12	0.32	0.68	0.92	1.0

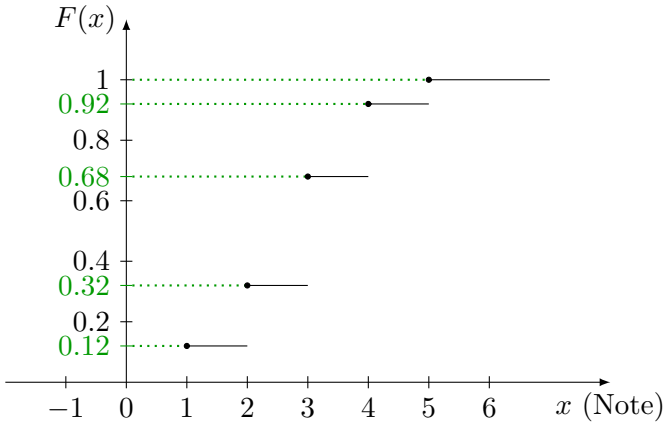


Abbildung 2.4: *(empirische) Verteilungsfunktion des Merkmals X*

2.20 Definition (empirischer Median, Zentralwert, zentrales Wertepaar)

$x_1, \ldots, x_n$ seien Werte einer O-, I- oder V-Skala.

(a) Der **empirische Median**[2] ist eine Lagemaßzahl

$$\tilde{x} := \begin{cases} x_{\left(\frac{n+1}{2}\right)} & \text{falls } n \text{ ungerade} \\ \frac{1}{2} \cdot \left(x_{\left(\frac{n}{2}\right)} + x_{\left(\frac{n+2}{2}\right)} \right) & \text{falls } n \text{ gerade} \end{cases}$$

In manchen Lehrbüchern wird der Median auch folgendermaßen definiert:

$$\tilde{x}' := \begin{cases} x_{\left(\frac{n+1}{2}\right)} & \text{falls } n \text{ ungerade} \\ x_{\left(\frac{n}{2}\right)} & \text{falls } n \text{ gerade} \end{cases}$$

(b) Für n ungerade heißt $x_{\left(\frac{n+1}{2}\right)}$ auch **Zentralwert** der Datenreihe und für n gerade heißt $\left(x_{\left(\frac{n}{s}\right)}, x_{\left(\frac{n+2}{2}\right)} \right)$ **zentrales Wertepaar** der Datenreihe.

2.21 Beispiel (Median, Zentralwert, zentrales Wertepaar)

(a) Messreihe: 1 2 $\overbrace{\underbrace{3}_{\tilde{x}=\tilde{x}'}}^{\text{Median, Zentralwert}}$ 4 5

Messreihe: 1 2 $\overbrace{\underbrace{3 \quad 4}_{\tilde{x}=3.5 \ ; \ \tilde{x}'=3}}^{\text{(3,4) zentrales Wertepaar}}$ 5 7

(b) $x_1, \ldots, x_{20}$ sei die Datenreihe aus 2.1:
$n = 20$, $\tilde{x} = \frac{1}{2} \left(x_{(10)} + x_{(11)} \right) = \frac{1}{2}(12 + 13) = 12.5 \quad , \quad (\tilde{x}') = 12$

2.22 Notiz

$X \colon \underbrace{\{1, \ldots, n\}}_{=M} \longrightarrow A.$

Eigenschaft des Median

$$\# \left\{ j \in M \;\middle|\; x_j < \tilde{x} \right\} \leq \frac{n}{2}$$
$$\# \left\{ j \in M \;\middle|\; x_j > \tilde{x} \right\} \leq \frac{n}{2}$$
$$\# \left\{ j \in M \;\middle|\; x_j \leq \tilde{x} \right\} \geq \frac{n}{2}$$

Mindestens die Hälfte der Messwerte haben einen Wert kleiner oder gleich dem Median.

$$\# \left\{ j \in M \;\middle|\; x_j < \tilde{x} \right\} \leq \frac{n}{2} \leq \# \left\{ j \in M \;\middle|\; x_j \leq \tilde{x} \right\}$$

[2] *englisch* median

2.23 Grafik zum Median

Modulprüfung!

Ablesen des Median aus dem Graphen der Verteilungsfunktion:

1. Fall

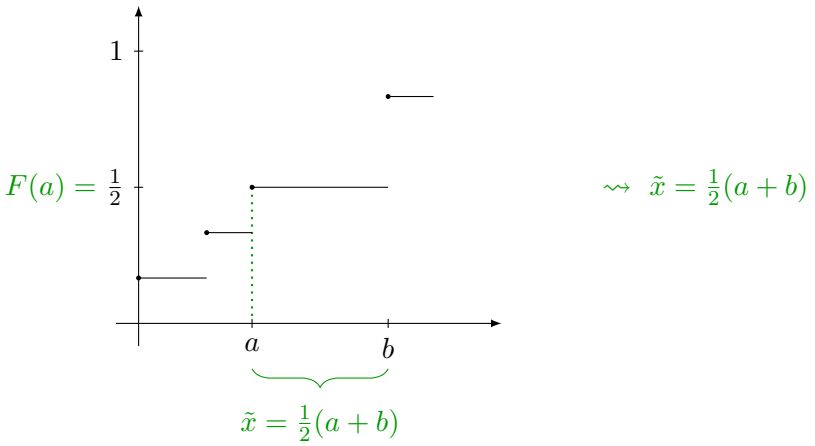


Abbildung 2.5: *z. B. n gerade,* $n = 10$

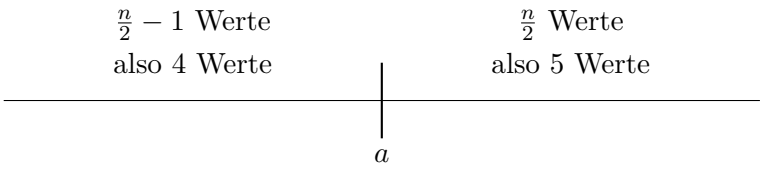


Abbildung 2.6: $F(a) = r(X \leq a) = \frac{5}{10} = \frac{1}{2} \quad , \quad \tilde{x} = \frac{1}{2}\left(x_{(\frac{n}{2})} + x_{(\frac{n+2}{2})}\right)$

2. Fall

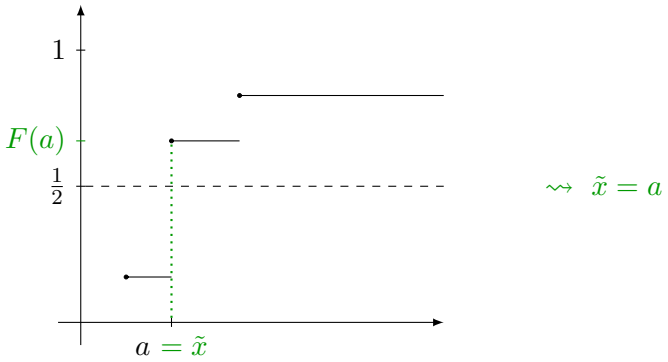


Abbildung 2.7: *z. B. n ungerade,* $n = 11$

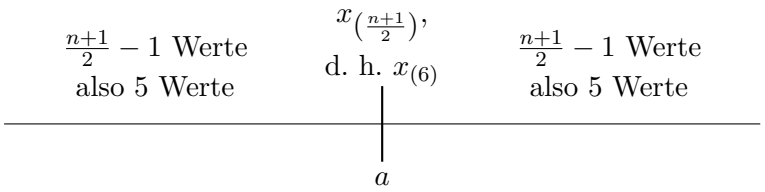


Abbildung 2.8: $F(a) = r(X \leq a) = \frac{6}{11} > \frac{1}{2}$, $\tilde{x} = x_{(\frac{n+1}{2})}$

Beispiel: Datenreihe aus 2.19 $\stackrel{\wedge}{=}$ 2. Fall, $\tilde{x} = 3$

2.24 Satz (Minimalitätseigenschaft des Median bzgl. der Abweichung zum Betrag)

Es seine $x_1, \ldots, x_n \in \mathbb{R}$. Dann ist

$$\sum_{i=1}^{n} |x_i - \tilde{x}| \leq \sum_{i=1}^{n} |x_i - d| \quad \forall\, d \in \mathbb{R}$$

Beweis

Teilweise für $n = 1, 2, 3$ findet man in „Fischer G.: Stochastik ...“
Ein vollständiger Beweis befindet sich in „Bosch K.: Einführung in die Statistik“

□

2.25 Definition (α-Quantil, r-tes Perzentil)

$x_1, \ldots, x_n$ seien Werte einer O-, I- oder V-Skala.

(a) Für $0 < \alpha < 1$ ist das **α-Quantil** definiert durch

$$x_\alpha := \begin{cases} \frac{1}{2}\left(x_{(n\alpha)} + x_{(n\alpha+1)}\right) & \text{falls } n\alpha \text{ ganzzahlig ist} \\ x_{([n\alpha+1])} & \text{sonst} \end{cases}$$

wobei $[n\alpha + 1]$ größte ganze Zahl ist, die kleiner oder gleich $n\alpha + 1$ ist, z. B. $[7.2] = 7$, $[7.8] = 7$, $[7] = 7$.

(b) Das $\frac{r}{100}$-Quantil (für $r = 1, \ldots, 99$) heißt auch ***r*-tes Perzentil**.

2.26 Notiz und Definition

Das 50-te Perzentil $\left(\alpha = \frac{1}{2}\right)$ ist der Median.
Das 25-te Perzentil $\left(\alpha = \frac{1}{4}\right)$ heißt auch **unteres Quartil**.
Das 75-te Perzentil $\left(\alpha = \frac{3}{4}\right)$ heißt auch **oberes Quartil**.

Der **Quartilabstand** ist ein Streuungsmaß zum Median und ist definiert durch

$$\text{Quartilabstand} := \text{oberes Quartil} - \text{unteres Quartil} = x_{\frac{3}{4}} - x_{\frac{1}{4}}$$

(nur sinnvoll für I- oder V-Skala).

2.27 Skizze zur groben Orientierung

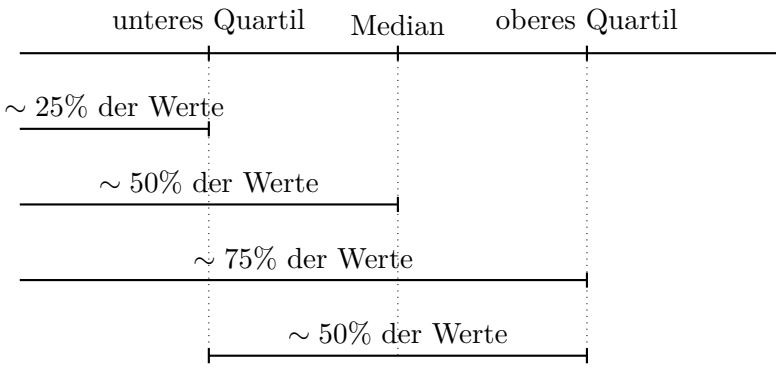


Abbildung 2.9

2.28 Grafiken zu den Quantilen

Für $0 < \alpha < 1$ gilt:

$$r(X < x_\alpha) \leq \alpha \leq r(X \leq x_\alpha) = F(x_\alpha)$$

Ablesen der Quantile aus einer Verteilungsfunktion zu einem Merkmal X.

1. Fall

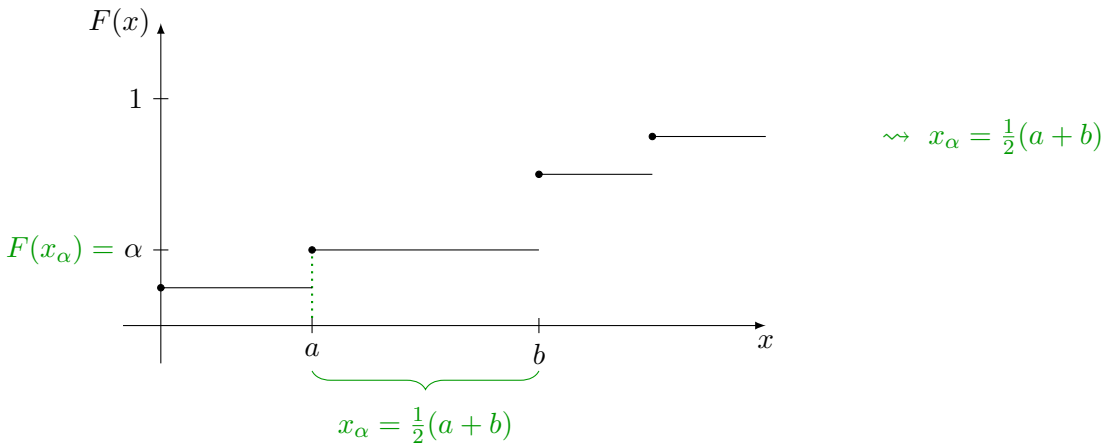


Abbildung 2.10

2. Fall

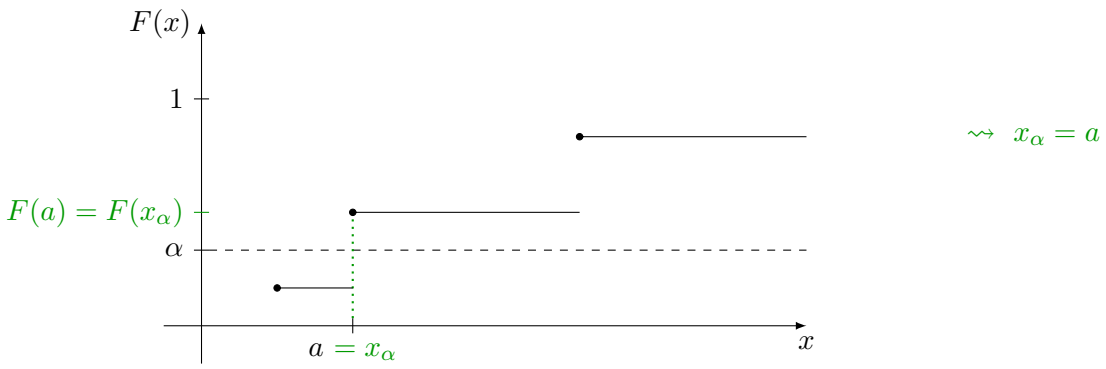


Abbildung 2.11

2.29 Vergleich arithmetisches Mittel – Median

(a) Median ist unempfindlicher als das arithmetische Mittel gegenüber sogenannten Ausreißerwerten.
Beispiel:

(1) geordnete Messreihe: 2 5 7 7 8 10 12

$$n = 7 \quad ; \quad \tilde{x} = x_{\left(\frac{7+1}{2}\right)} = x_{(4)} = 7 \quad ; \quad \bar{x} = \frac{51}{7} \approx 7.28$$

(2) geordnete Messreihe: 2 5 7 7 8 10 12 30

$$n = 8 \quad ; \quad \tilde{x} = \frac{x_{(4)} + x_{(5)}}{2} = \frac{7+8}{2} = 7.5 \quad ; \quad \bar{x} = \frac{81}{8} = 10.125$$

(b) Das arithmetische Mittel ist ein Durchschnittswert für die Datenreihe (bzw. Schätzwert für den Erwartungswert einer Zufallsgröße). Der Median teilt die Datenreihe in 2 etwa gleich große Bereiche.
Beispiel:
Bei 50 Glühlampen wird gemessen, nach welcher Zeitspanne sie ausfallen.
arithmetisches Mittel $\Longrightarrow$ durchschnittliche Lebensdauer (Kosten!)
Median $\Longrightarrow$ Zeitspanne, zu der noch 50 % der Lampen in Ordnung sind (Ausfallsicherheit!)

(c) Verhalten von Median und arithmetischem Mittel bei verschiedenen Verteilungsfunktionen: „**symmetrische Verteilung**“, „**linkssteile Verteilung**“ und „**rechtssteile Verteilung**“

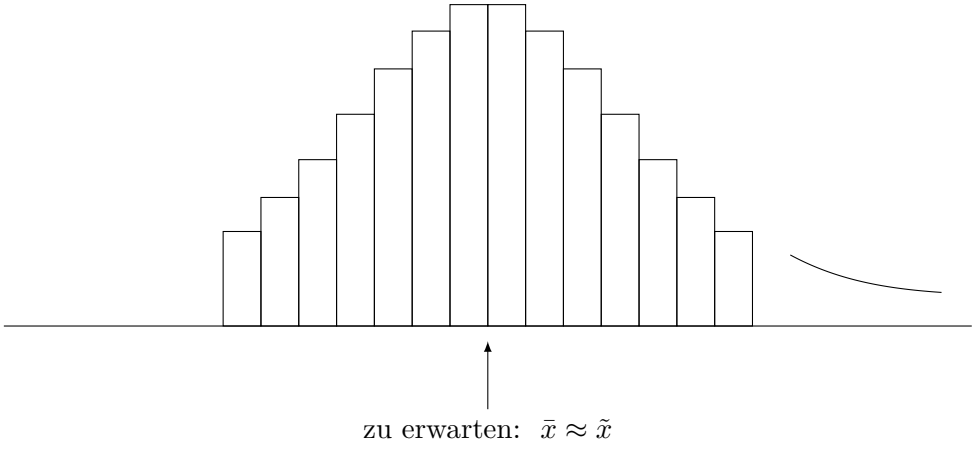


Abbildung 2.12: *symmetrische Verteilung*

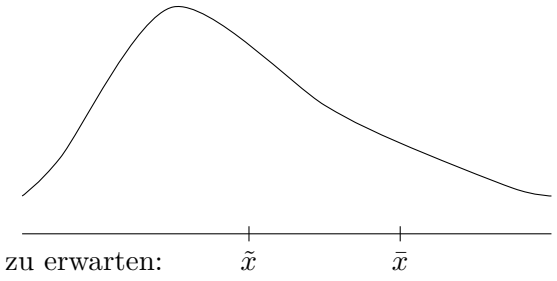


Abbildung 2.13: *linkssteile Verteilung*

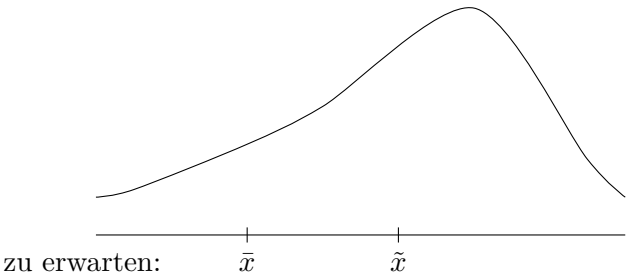


Abbildung 2.14: *rechtssteile Verteilung*

2.30 Definition (geometrisches und harmonisches Mittel)

Es seien $x_1, \ldots, x_n \in \mathbb{R}$, $x_i \geq 0$ für alle $i \in \{1, \ldots, n\}$, Messwerte einer Verhältnisskala. Dann heißt

(a) $\bar{x}_G := \sqrt[n]{x_1 \cdot \ldots \cdot x_n}$ das **geometrische Mittel**

(b) $\bar{x}_H := n \cdot \frac{1}{\frac{1}{x_1} + \ldots + \frac{1}{x_n}}$ das **harmonische Mittel**.

Hgew steht für Gewichtung

(c) Variante des harmonischen Mittels:

$$\bar{x}_{Hgew} = \sum_{j=1}^{n} g_j \cdot \frac{1}{\frac{g_1}{x_1} + \ldots + \frac{g_n}{x_n}}$$

wobei g_j Zusatzinformationen zu x_j darstellen.

2.31 Bemerkung

Es seien $x_1, \ldots, x_n \in \mathbb{R}$, $x_i > 0$ für alle $i \in \{1, \ldots, n\}$.

(a) $y_i = \ln(x_i) \implies \bar{x}_G = \exp(\bar{y})$ (logarithmische Koordinatentransformation)

(b) (i) $\bar{x}_G \leq \bar{x}$

(ii) $\bar{x}_G = \bar{x} \iff x_1 = \ldots = x_n$

Beweis

Übungsaufgabe.

□

2.32 Bemerkung

Das geometrische Mittel wird vor allem bei Wachstumsprozessen (z. B. Zinseszins-Berechnungen) verwendet.
Das harmonische Mittel ist z. B. bei der Berechnung von Durchschnittsgeschwindigkeiten sinnvoll.

2.33 Beispiel

Ein Fahrzeug fährt die ersten 100 km mit einer Geschwindigkeit von 150 $\frac{\text{km}}{\text{h}}$, weitere 100 km mit 50 $\frac{\text{km}}{\text{h}}$.
Wie groß ist die Durchschnittsgeschwindigkeit?

$$\bar{x}_H = \left(\underbrace{100}_{g_1} + \underbrace{100}_{g_2} \right) \cdot \frac{1}{\frac{100}{150} + \frac{100}{50}} = \frac{\overbrace{2}^{=n}}{\underbrace{\frac{1}{150}}_{=x_1} + \underbrace{\frac{1}{50}}_{=x_2}} = \frac{600}{8} = 75 \left(\frac{\text{km}}{\text{h}} \right)$$

$\bar{x} = (150 + 50) : 2 = 100$ ist falsche, denn

$$\frac{100}{150} \mathrel{\hat{=}} \frac{2}{3}\,\text{h}\,, \quad \frac{100}{50} \mathrel{\hat{=}} 2\,\text{h} \quad \rightsquigarrow \quad 100 \cdot \left(\frac{2}{3} + 2 \right) \approx 267\,(\text{km})$$

Aber $75 \cdot \left(\frac{2}{3} + 2 \right) = 75 \cdot \frac{8}{3} = 200\,(\text{km})$.

2.34 Beispiel (arithmetisches Mittel, geometrisches Mittel)

Bevölkerungsentwicklung einer aufstrebenden Gemeinde von 2007 bis 2011

Jahr	Bevölkerung am 31.12.	Zuwachs gegenüber Vorjahr
2007	$l_0 = 20000$	
2008	32000	60 %
2009	41600	30 %
2010	44387	6.7 %
2011	$l_4 = 48826$	10 %

Durchschnittlicher Zuwachs mit arithmetischem Mittel

$$\bar{x} = \frac{1}{4}(60 + 30 + 6.7 + 10) = 26.675$$

Variante:

$$\bar{x} = \frac{1}{4}(1.6 + 1.3 + 1.067 + 1.1) - 1 = 0.26675$$

Jahr	Bevölkerung	durchschnittlicher Zuwachs
2007	$C_0 = 20000$	
2008	25335	26.675 %
2009	32093	26.675 %
2010	40654	26.675 %
2011	$C_4 = 51498$	26.675 %

$$C_4 = 1.26675^4 \cdot C_0 \approx 51498$$

Durchschnittlicher Zuwachs mit geometrischem Mittel

$$\bar{x}_G = \sqrt[4]{1.6 \cdot 1.3 \cdot 1.067 \cdot 1.1} - 1 \approx 0.2499859 \approx 0.25$$

Jahr	Bevölkerung	durchschnittlicher Zuwachs
2007	$D_0 = 20000$	
2008	25000	25 %
2009	31250	25 %
2010	39063	25 %
2011	$D_4 = 48829$	25 %

$D_4 = 1.25^4 \cdot D_0 \approx 48828$

$D_4 = 1.2499859^4 \cdot D_0 \approx 48826$

2.35 Definition (Modalwert)

Vor allem bei Nominalskalen wird der **Modalwert** $\bar{x}_D$ als Maßzahl verwendet. $\bar{x}_D$ bezeichnet die Ausprägung, die am häufigsten vorkommt, d. h.

$$h(\bar{x}_D) = \max \left\{ h(a) \;\middle|\; a \in A \right\}$$

Gibt es mehrere Ausprägungen mit der größten Häufigkeit, so gibt es mehrere Modalwerte.

Beispiel

Augenfarbe:

Farbe	Codierung	$h(.)$
braun	1	10
grün	2	4
grau	3	2

$$\bar{x}_D = 1 \text{ (braun)}$$

2.36 Definition (gewichtetes arithmetisches Mittel für I- oder V-Skala)

$x_1, \ldots, x_n \in \mathbb{R}$ sei eine Datenreihe mit positiven Gewichten $g_i > 0$, $i = 1, \ldots, n$. Dann heißt

$$\bar{x}_{gew} = \frac{1}{\sum\limits_{i=1}^{n} g_i} \cdot \sum_{i=1}^{n} g_i x_i$$

das **gewichtete arithmetische Mittel**.
Für $g_i = 1$ für alle $i \in \{1, \ldots, n\}$ ist $\bar{x}_{gew} = \bar{x}$.

Beispiel

Zeit	30 Minuten	90 Minuten	15 Minuten
Durchschnittsgeschwindigkeit in $\frac{\text{km}}{\text{h}}$	50	90	60

$$\bar{x}_{gew} = \frac{1}{\underbrace{\frac{1}{2} + \frac{3}{2} + \frac{1}{4}}_{g_i \text{ in Stunden}}} \cdot \left(\frac{1}{2} \cdot 50 + \frac{3}{2} \cdot 90 + \frac{1}{4} \cdot 60 \right) = \frac{4}{9} \cdot 175 \approx 77.8 \left(\frac{\text{km}}{\text{h}} \right)$$

2.37 Bemerkung (Berechnungsvariante für die Stichprobenvarianz)

Vor: $x_1, \ldots, x_n \in \mathbb{R}$ sei eine Datenreihe.

Beh: $$s^2 = \underbrace{\frac{1}{n-1} \cdot \left(\sum_{i=1}^{n} x_i^2 - n\bar{x}^2 \right)}_{(*)} = \underbrace{\frac{1}{n-1} \left(\sum_{i=1}^{n} x_i^2 - \frac{1}{n} \left(\sum_{i=1}^{n} x_i \right)^2 \right)}_{(**)}$$

Vorüberlegung: (i) $\sum_{i=1}^{n} \bar{x}^2 = n \cdot \bar{x}^2$ (ii) $\frac{1}{n} \cdot \sum_{i=1}^{n} x_i = \bar{x} \overset{n \neq 0}{\iff} \sum_{i=1}^{n} x_i = n\bar{x}$

Beweis

$$
\begin{aligned}
s^2 &= \frac{1}{n-1} \cdot \sum_{i=1}^{n} (x_i - \bar{x})^2 = \frac{1}{n-1} \left(\sum_{i=1}^{n} (x_i^2 - 2x_i\bar{x} + \bar{x}^2) \right) \\
&= \frac{1}{n-1} \left(\sum_{i=1}^{n} x_i^2 - 2\bar{x} \cdot \underbrace{\sum_{i=1}^{n} x_i}_{n\bar{x}} + \underbrace{\sum_{i=1}^{n} \bar{x}^2}_{n\bar{x}^2} \right) = \frac{1}{n-1} \left(\sum_{i=1}^{n} x_i^2 - n\bar{x}^2 \right) \implies (*) \\
&= \frac{1}{n-1} \left(\sum_{i=1}^{n} x_i^2 - n \cdot \left(\frac{1}{n} \cdot \sum_{i=1}^{n} x_i \right)^2 \right) = \frac{1}{n-1} \left(\sum_{i=1}^{n} x_i^2 - \frac{1}{n} \cdot \left(\sum_{i=1}^{n} x_i \right)^2 \right) \implies (**)
\end{aligned}
$$

□

2.38 Boxplots

Geeignet für O-, I- und V-Skalen. Grafische Darstellung von Minimum, unteres Quartil, Median, oberes Quartil, Maximum.

Beispiel

Gegeben sei die geordnete Messreihe (Datenreihe) $x_{(1)}$ bis $x_{(11)}$:

$$2 \ , \ 3 \ , \ 3 \ , \ 4 \ , \ 5 \ , \ 5 \ , \ 6 \ , \ 7 \ , \ 8 \ , \ 11 \ , \ 14$$

Minimum: 2, Maximum: 14
unteres Quartil: $x_{\frac{1}{4}} = x_{\left(\left[11 \cdot \frac{1}{4} + 1\right]\right)} = x_{(3)} = 3$
Median: $\tilde{x} = x_{\left(\frac{11+1}{2}\right)} = x_{(6)} = 5$
oberes Quartil: $x_{\frac{3}{4}} = x_{\left(\left[11 \cdot \frac{3}{4} + 1\right]\right)} = x_{(9)} = 8$
Boxplot:

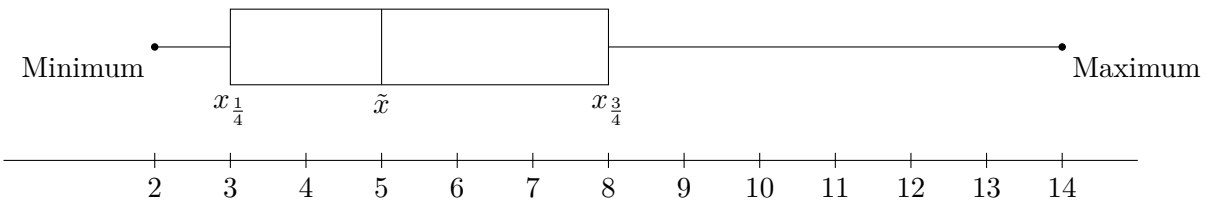


Abbildung 2.15: *Boxplot*

§ 3 Beschreibung zweidimensionaler Datenreihen

Problemstellung

Existiert ein Zusammenhang zwischen zwei Merkmalen?
Gegeben ist also eine zweidimensionale Datenreihe (Stichprobe)

$$(x_1, y_1), (x_2, y_2), \ldots, (x_n, y_n)$$

Jedes Wertepaar (x_i, y_i) mit $i = 1, \ldots, n$ gehört in der Regel zu einer Untersuchungseinheit (Person, Objekt, ...).

3.1 Beispiel

Körpergröße x (Merkmal 1), Gewicht y (Merkmal 2)

Person	x	y	(x_i, y_i)
1	175	70	$\rightsquigarrow \ (x_1, y_1) = (175, 70)$
2	180	72	$\rightsquigarrow \ (x_2, x_2) = (180, 72)$
⋮	⋮	⋮	⋮
20	169	65	$\rightsquigarrow \ (x_{20}, y_{20}) = (169, 65)$

Dazugehörige eindimensionale Datenreihen:
$(x_1, \ldots, x_{20})$ Datenreihe zu X und $(y_1, \ldots, y_{20})$ Datenreihe zu Y.

3.2 Graphische Darstellung

Für die graphische Darstellung das Zusammenhangs zweier Merkmale werden **Punktwolken** (Streudiagramme[3]) verwendet. Dazu werden die Wertepaare (x_i, y_i) in ein Koordinatensystem eingetragen. Die Form der Punktwolke gibt einen Hinweis auf die Stärke des Zusammenhangs.

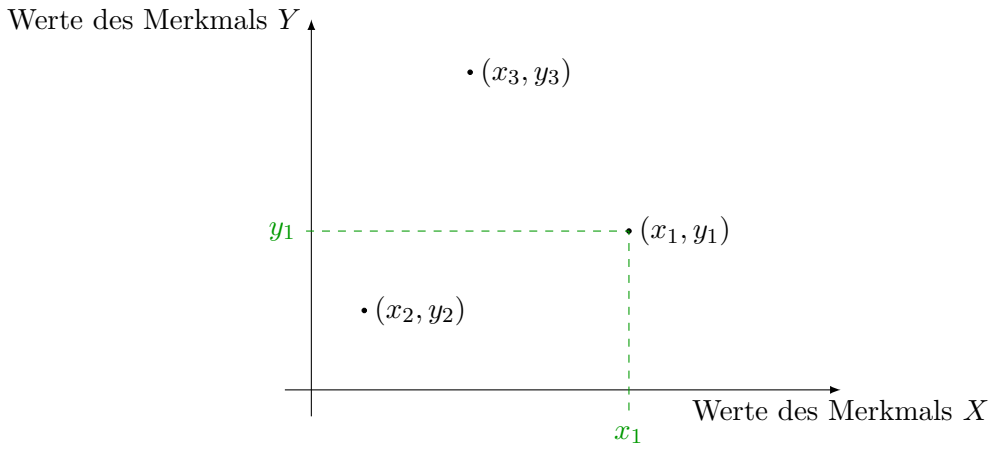


Abbildung 3.1: *Punktwolke (Streudiagramm)*

[3] *englisch* scatterplots

3.3 Beispiele

(a)

i	1	2	3	3	5	6	7
x	10	16	7	19	15	3	11
y	11	5	14	2	6	18	10

$\rightsquigarrow$ Datenreihe $(x_1, y_1), \ldots, (x_7, y_7)$

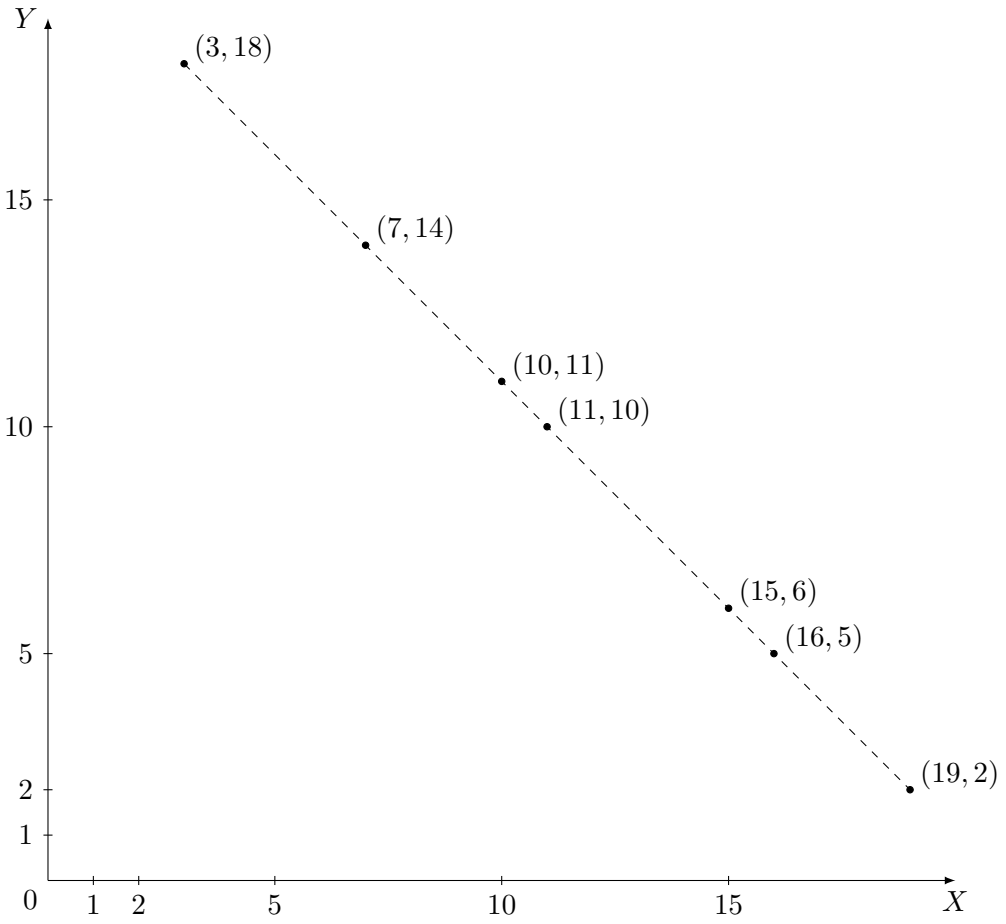


Abbildung 3.2: *Starker (sogar funktionaler) Zusammenhang;* $y = 21 - x$.

(b)

i	1	2	3	3	5	6	7
x	6	2	5	1	3	6	2
y	2	4	5	2	4	2	5

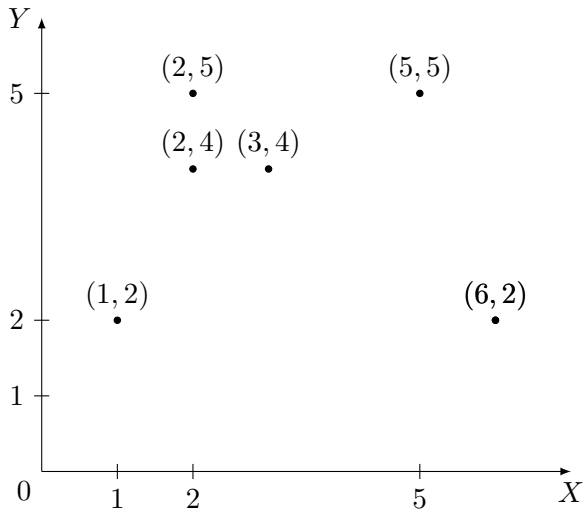


Abbildung 3.3: *KEIN Zusammenhang ersichtlich (z. B. Werfen von 2 fairen Würfeln)*

3.4 Definition (Pearsonscher Korrelationskoeffizient)

Für metrische Merkmale wird als statistische Maßzahl für die Abhängigkeit (Korrelation) zweier Merkmale häufig der **Pearsonsche[4] Korrelationskoeffizient** verwendet. Er ist definiert durch

$$r = r_{x,y} := \frac{\overbrace{s_{x,y}}^{\substack{\text{enthält Informationen} \\ \text{über evtl. Zusammenhang}}}}{\underbrace{s_x \cdot s_y}_{\text{Normierungsfaktoren}}}$$

wobei

$$s_{x,y} := \frac{1}{n-1} \cdot \sum_{i=1}^{n} (x_i - \bar{x})(y_i - \bar{y})$$

empirische Kovarianz,

$$s_x := \sqrt{\frac{1}{n-1} \cdot \sum_{i=1}^{n} (x_i - \bar{x})^2}$$

Standardabweichung der Datenreihe $(x_1, \ldots, x_n)$,

$$s_y := \sqrt{\frac{1}{n-1} \cdot \sum_{i=1}^{n} (y_i - \bar{y})^2}$$

Standardabweichung der Datenreihe $(y_1, \ldots, y_n)$.

3.5 Notiz

Für die Berechnung von $s_{x,y}$ kann auch folgende Formel verwendet werden:

$$s_{x,y} = \frac{1}{n-1} \cdot \left(\left(\sum_{i=1}^{n} x_i y_i \right) - (n\bar{x} \cdot \bar{y}) \right)$$

3.6 Bemerkungen

(a) r ist ein Maß für die lineare Abhängigkeit zwischen zwei Merkmalen.

(b) Es gilt:

(i) $-1 \leq r \leq 1$ (Beweis: Übungsaufgabe)

(ii) $r = 1 \iff y_i = ax_i + b$ mit $a, b \in \mathbb{R}$, $a > 0$

(iii) $r = -1 \iff y_i = ax_i + b$ mit $a, b \in \mathbb{R}$, $a < 0$

Sind $(x_1, \ldots, x_n)$, $(y_1, \ldots, y_n)$ Stichproben von zwei unabhängigen Merkmalen X und Y, so ist zu erwarten, dass r „nahe“ bei 0 liegt.

(c) Ist $r = 0$ oder „nahe“ bei 0, so bedeutet dies NICHT, dass die beiden Merkmale unabhängig sind (es besteht evtl. ein nichtlinearer Zusammenhang).

3.7 Bemerkung (Verhalten von r bei linearen Transformationen der Messwerte)

$a_1x + b_1$ bedeute $x_i \xrightarrow{\text{transformiert in}} a_1x_i + b_1$ für $i = 1, \ldots, n$
$a_2y + b_2$ bedeute $y_i \xrightarrow{\text{transformiert in}} a_2y_i + b_2$ für $i = 1, \ldots, n$

$$\implies r_{(a_1x+b_1),(a_2y+b_2)} = \left.\begin{cases} r_{x,y} & \text{falls } a_1, a_2 > 0 \\ r_{x,y} & \text{falls } a_1, a_2 < 0 \\ -r_{x,y} & \text{falls } a_1 > 0,\ a_2 < 0 \\ -r_{x,y} & \text{falls } a_1 < 0,\ a_2 > 0 \end{cases}\right\} = \frac{a_1 \cdot a_2}{|a_1 \cdot a_2|} \cdot r_{x,y}$$

Beweis

Übungsaufgabe.

□

3.8 Skizze zu 3.7

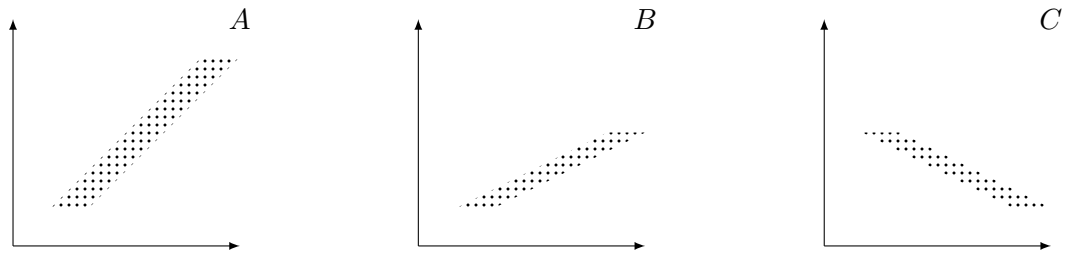


Abbildung 3.4: $r_A \approx r_B$; $r_C \approx -r_A$

Der Grad der Steigung der Punktwolke wirkt sich NICHT auf r aus.

3.9 Bemerkung zur Interpretation von r

Bei der statistischen Analyse des Zusammenhangs zweier Merkmale geht es NICHT um den Nachweis von Kausalitäten (Ursache $\longrightarrow$ Wirkung). Dieser Nachweis ist sachlogischen Argumenten vorbehalten.

r ist nur ein Maß dafür, wie stark die Werte der beiden Datenreihen gemeinsam variieren. Ein hoher Korrelationskoeffizient ist NICHT automatisch ein Hinweis auf einen direkten Zusammenhang zwischen den Merkmalen.

3.10 Lineare Regression

r ist ein Maßstab für die Stärke eines linearen Zusammenhangs. Ziel der Regressionsrechnung ist es, die Abhängigkeit des einen Merkmals von einem anderen Merkmal durch eine Funktionsgleichung zu beschreiben.

Man nennt dabei X **unabhängige Variable** und Y **abhängige Variable**.

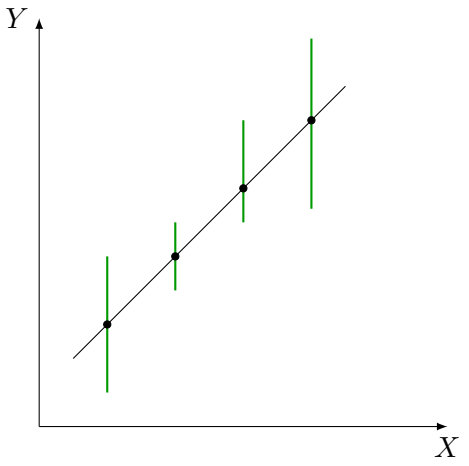
Welches Merkmal as unabhängig und welches als abhängig gelten soll, ergibt sich meistens aus der Problemstellung, z. B. Alter – Blutdruck: Alter unabhängig, Blutdruck abhängig.

Bei der **linearen Regression** soll die Form der Abhängigkeit durch eine **lineare Funktion** (eine sogenannte **Regressionsgerade** oder **Ausgleichsgerade**) dargestellt werden, d. h. in der Form

$$Y = a \cdot X + b \quad \text{mit } a, b \in \mathbb{R}\ ,\ a \neq 0$$

3.11 Modellvorstellung zur linearen Regression

(a) Jedem Wert x der Variablen X wird ein „wahrer Wert“ y der Variablen Y zugeordnet. Bei Messungen wird der zu einem x-Wert gewonnene y-Wert mehr oder minder stark vom wahren y-Wert abweichen (wegen Zufallseinflüssen).



„wahre Werte“ liegen auf der Gerade $Y = aX + b$

Streubereich der Werte für Y bei einem Experiment

Abbildung 3.5

(b) Bei einem Experiment erhält man eine zweidimensionale Datenreihe $(x_1, y_1), \ldots, (x_n, y_n)$

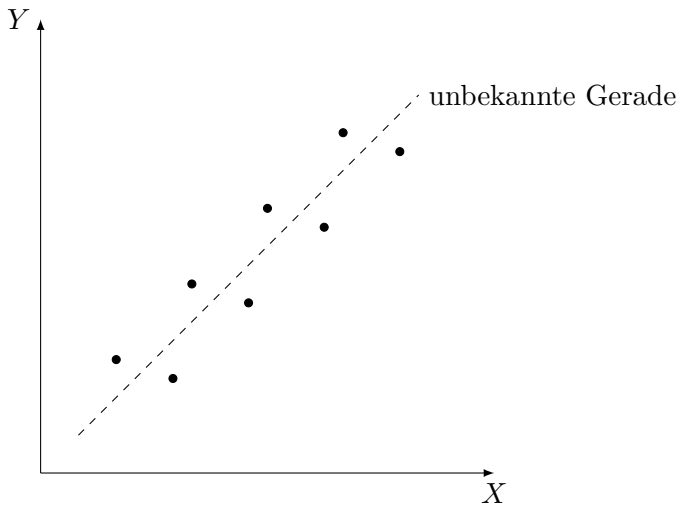


Abbildung 3.6

Wie erhält man aus der zweidimensionalen Datenreihe eine Gerade (Regressionsgerade), die den unbekannten linearen Zusammenhang $Y = aX + b$ „gut“ beschreibt?

Idee: Die Parameter a und b sollen so gewählt werden, dass die Gerade $Y = aX + b$ die Punktwolke möglichst „gut approximiert“.

3.12 Zur Bestimmung von a und b der Regressionsgerade

... verwendet man i. A. die **Methode der kleinsten Quadrate**[5]:

Die **Summe der Quadrate der Abstände** aller Punkte von der gesuchten Geraden soll minimal sein. Die Abstände werden dabei durch Parallelen zu Y-Achse bestimmt.

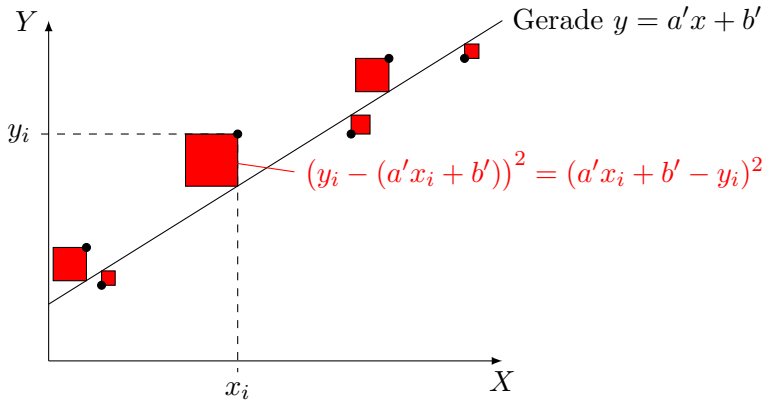


Abbildung 3.7

[5] *englisch* least square problem

Problem:
Gesucht sind $a, b \in \mathbb{R}$ mit der Eigenschaft

$$\sum_{i=1}^{n}(ax_i + b - y_i)^2 = \min_{(a',b')\in\mathbb{R}^2} \sum_{i=1}^{n}(a'x_i + b' - y_i)^2$$

Lösung:

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

wobei a **Regressionskoeffizient**[6] (Steigung der Geraden) und

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

die **Regressionskonstante**[7] (Schnittpunkt mit y-Achse) darstellen. $\bar{x}$ und $\bar{y}$ sind die arithmetischen Mittel zu $(x_1, \ldots, x_n)$ bzw. $(y_1, \ldots, y_n)$.

3.13 Bemerkung (für $n \geq 2$)

(1) Für den Regressionskoeffizienten gilt:

$$a = \frac{(n-1)s_{x,y}}{(n-1)s_x^2} = \frac{s_{x,y}}{s_x^2}$$

(2) Es gilt: Der Punkt $(\bar{x}, \bar{y})$ liegt stets auf der Regressionsgeraden.
Genauer gilt: Für $\hat{y}_i = ax_i + b$ ($i = 1, \ldots, n$; (a, b) aus der Regressionsgeraden) gilt

$$\bar{\hat{y}} = \bar{y} = a\bar{x} + b$$

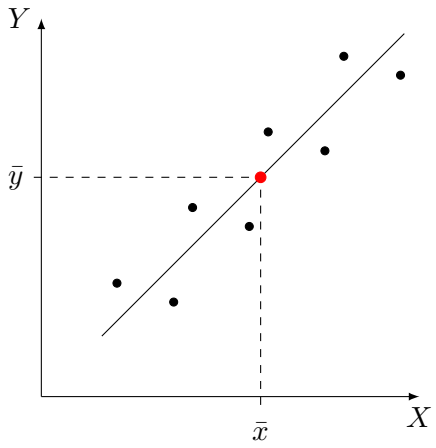


Abbildung 3.8

[6] *englisch* slope
[7] *englisch* intercept

Beispiele für die Bestimmung der Regressionsgeraden werden in der Zentralübung behandelt.

3.14 Linearisierung durch Transformation

Modulprüfung!

Durch Transformation der Messwerte lassen sich einige häufig vorkommende nichtlineare Zusammenhänge mit Hilfe der linearen Regression behandeln.

(a) **Exponentialfunktion** (einfach logarithmische Transformation):
Die Regressionsfunktion habe die Form

$$y = b \cdot a^x \quad \text{mit } b, a > 0\ ,\ x \in \mathbb{R}$$

Durch Logarithmieren auf beiden Seiten erhält man wegen

$$\underbrace{\ln(y)}_{=:y'} = \ln(b \cdot a^x) = \underbrace{\ln(b)}_{:=b'} + \underbrace{\ln(a)}_{=:a'} \cdot x$$
$$y' = b' + a' \cdot x$$

nach der der Transformation einen linearen Zusammenhang zwischen x und y'. Mit linearer Regressionsrechnung lassen sich a' und b' bestimmen und daraus erhält man

$$a = \exp(a') \quad , \quad b = \exp(b')$$

x	1	1.5	2	2.5	3	3.5	4
y	1	1.4	2	2.8	4	5.7	8
$\ln(y)$	0	0.3	0.7	1	1.4	1.7	2.1

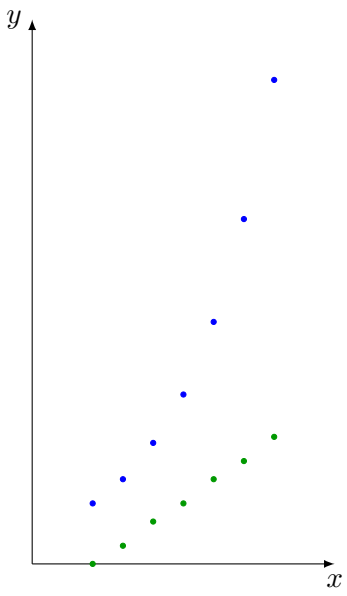


Abbildung 3.9

Durchführung:
Die zweidimensionale Datenreihe $(x_1, y_1), \ldots, (x_n, y_n)$ soll „möglichst gut“ durch eine Funktion der Form

$$y = b \cdot a^x$$

beschrieben werden.

1. Schritt: Erzeuge transformierte Datenreihe (Vor. $y_i > 0$):
$(x_1, \ln(y_1)), \ldots, (x_n, \ln(y_n))$ (einfache logarithmische Transformation)
2. Schritt: Berechne a' und b' zur transformierten Datenreihe mit linearer Regression
3. Schritt: Berechne $a = \exp(a')$ und $b = \exp(b')$

(b) **Potenzfunktion** (doppelt logarithmische Transformation)
Die Regressionsfunktion habe die Form

$$y = b \cdot x^a \quad \text{mit } b > 0\ ,\ x > 0\ ,\ a \in \mathbb{R}$$

Durch Logarithmieren auf beiden Seiten erhält man wegen

$$\underbrace{\ln(y)}_{=:y'} = \ln(b \cdot x^a) = \underbrace{\ln(b)}_{=:b'} + a \cdot \underbrace{\ln(x)}_{=:x'}$$
$$y' = b' + a \cdot x'$$

einen linearen Zusammenhang zwischen x' und y'.
Mit linearer Regression lassen sich a und b' bestimmen und $b = \exp(b')$ berechnen.

Durchführung:
Die zweidimensionale Datenreihe $(x_1, y_1), \ldots, (x_n, y_n)$ soll „möglichst gut“ durch eine Funktion der Form

$$y = b \cdot x^a$$

beschrieben werden.

1. Schritt: Erzeuge transformierte Datenreihe (Vor. $x_i, y_i > 0$):
$\big(\ln(x_1), \ln(y_1)\big), \ldots, \big(\ln(x_n), \ln(y_n)\big)$ (doppelt logarithmische Transformation)
2. Schritt: Berechne a und b' zur transformierten Datenreihe mit linearer Regression
3. Schritt: Berechne $b = \exp(b')$

3.15 Anmerkung

Für Ordinalskalen wird oft der „Spearmansche Rangkorrelationskoeffizient“ verwendet.

3.16 Kontingenzkoeffizient

Bei der Analyse des Zusammenhangs zwischen nominalskalierten Merkmalen wird statt „Korrelation“ meisten der Begriff **Kontingenz** verwendet.

Zur Analyse des Zusammenhangs werden die Daten in einer sogenannten **Kontingenztafel** (Feldertafel) zusammengefasst.

3.17 Beispiel

Zusammenhang zwischen Augen- und Haarfarbe, $n = 100$ Personen:

y Haar / x Auge	blond	rot	braun	schwarz	Z_i Zeilensumme	Randwahrscheinlichkeit, relative Häufigkeit
grün	5	1	19	5	30	0.3
blau	24	1	3	0	28	0.28
braun	7	3	17	15	42	0.42
S_j Spaltensumme	36	5	39	20	100	
Randwahrscheinlichkeit, relative Häufigkeit	0.36	0.05	0.39	0.2		

3.18 Vorüberlegungen

(a) Gäbe es zwischen Merkmal X (Augenfarbe) und Y (Haarfarbe) KEINEN Zusammenhang, d. h. X und Y sind „unabhängig im Sinne der Wahrscheinlichkeitstheorie“, so wäre

$$\text{Wkt}(X = \text{grün und } Y = \text{blond}) = \text{Wkt}(X = \text{grün}) \cdot \text{Wkt}(Y = \text{blond}) \overset{3.17}{=} 0.3 \cdot 0.36 = 0.108$$

Die erwartete Häufigkeit für das Ereignis „X = grün und Y = blond“ ist dann, unter der Voraussetzung, dass KEIN Zusammenhang vorliegt,

$$0.108 \cdot 100 = 10.8$$

Die erwartete Häufigkeit (unter der Voraussetzung, dass KEIN Zusammenhang vorliegt) sei mit E_{ij} bezeichnet. Für E_{ij} gilt:

$$E_{ij} = \underbrace{\frac{Z_i}{n} \cdot \frac{S_j}{n}}_{\substack{\text{Produkt rel.} \\ \text{Häufigkeiten}}} \cdot n = \frac{Z_i \cdot S_j}{n}$$

(b) Im Beispiel 3.17 ist $i = 1, 2, 3$, $j = 1, 2, 3, 4$, $n = 100$. Besteht ein starker Zusammenhang zwischen den beiden Merkmalen, so wird der beobachtete Wert B_{ij} (im Beispiel: $B_{11} = 5$, $B_{12} = 1 \ldots$) stark vom erwarteten Wert E_{ij} abweichen.

3.19 Definition (Kontingenzkoeffizient)

(a) Als Hilfsgröße für den Zusammenhang bei nominalskalierten Merkmalen (aber auch bei Klassenbildung metrischer Merkmale) verwendet man

$$\chi^2 := \sum_{i=1}^{c_1} \sum_{j=1}^{c_2} \frac{(B_{ij} - E_{ij})^2}{E_{ij}}$$

(Chi-Quadrat), wobei c_1 die Anzahl der Zeilen und c_2 die Anzahl der Spalten der Kontingenztafel bezeichnen.

(b) Die Normierung von χ^2 führt zu einer Maßzahl für den Zusammenhang:

$$C := \sqrt{\frac{\chi^2}{\chi^2 + n}}$$

heißt **Kontingenzkoeffizient**.

3.20 Bemerkung

(1) C ist kleiner als 1 (natürlich ≥ 0).

(2) Der maximale Wert von C, bezeichnet mit $C_{\max}$, hängt von der Form der Kontingenztafel ab. Es gilt:

$$C_{\max} = \sqrt{\frac{m-1}{m}}$$

mit $m := \min(c_1, c_2)$.

Zum Beispiel ist $C_{\max}(3,3) = \sqrt{\frac{2}{3}} \approx 0.816$, $C_{\max}(4,4) \approx 0.866$.

(3) Wegen (2) verwendet man statt C auch den sogenannten korrigierten Kontingenzkoeffizienten C_k.

$$C_k := \frac{C}{C_{\max}} = \sqrt{\frac{\chi^2 \cdot m}{(\chi^2 + n) \cdot (m-1)}}$$

wobei n die Anzahl der Fälle und $m = \min(c_1, c_2)$ bezeichnen.

Kapitel II: Wahrscheinlichkeitsrechnung

§ 4 Der Begriff Wahrscheinlichkeit

Was ist Wahrscheinlichkeit im mathematischen Sinn?

4.1 Bezeichnungen

Bei einem Zufallsexperiment seien $\omega_1, \ldots, \omega_n$ die möglichen **Ergebnisse**. Die Menge

$$\Omega := \{\omega_1, \ldots, \omega_n\}$$

heißt **Ergebnismenge**. Ziel ist es, für jedes $\omega \in \Omega$ eine reelle Zahl

$$P(\omega) \quad \text{mit } 0 \leq P(\omega) \leq 1$$

zu erklären, als Wahrscheinlichkeit dafür, dass das Ergebnis ω eintritt.

Eine Teilmenge $A \subset \Omega$ ($A = \Omega$ ist dabei auch möglich) wird als **Ereignis** bezeichnet. Gesucht ist wieder eine Zahl

$$P(A) \in \mathbb{R} \quad \text{mit } 0 \leq P(A) \leq 1$$

die Wahrscheinlichkeit dafür, dass $\omega \in A$, d. h. das Ereignis A, eintritt.
$A = \{\omega\}$ heißt **Elementarereignis**.

4.2 Beispiel (Würfel)

$\Omega = \{1, \ldots, c\}$; $A = \{1, 3, 5\} \subset \Omega$ ist ein Ereignis. $B = \{1\}$ ist ein Elementarereignis.
Für einen „fairen“ Würfel ist

$$P(\{i\}) = \frac{1}{6}$$

für $i = 1, \ldots, 6$ sinnvoll.

4.3 Definition (Relative Häufigkeiten)

Ein Zufallsexperiment werde k-mal durchgeführt. Seien $\omega \in \Omega$ und $A \subset \Omega$.

(a) Häufigkeiten des Eintretens von ω bzw. A

$H_k(\omega) =$ Anzahl der Male, in denen ω eintritt
$H_k(A) =$ Anzahl der Male, in denen $\omega \in A$ eintritt

(b) Relative Häufigkeiten des Eintretens von ω bzw. A

$$R_k(\omega) := \frac{1}{k} \cdot H_k(\omega)$$
$$R_k(A) := \frac{1}{k} \cdot H_k(A) = \sum_{\omega \in A} R_k(\omega)$$

4.4 Bemerkungen

Mit den Bezeichnungen aus 4.3 gelten:

(a) $0 \leq H_k(\omega) \leq k \ , \ 0 \leq H_k(A) \leq k$

(b) $0 \leq R_k(\omega) \leq 1 \ , \ \sum\limits_{\omega \in \Omega} R_k(\omega) = 1$

(c) $0 \leq R_k(A) \leq 1$ für $A \subset \Omega$

(d) $R_k(\Omega) = 1$

(e) $R_k(A \cup B) = R_k(A) + R_k(B)$ für $A, B \subset \Omega$, $\underline{A \cap B = \emptyset}$

4.5 Anmerkungen

Intuitiv hat man die Vorstellung, dass die Zahl $R_k(A)$ bei immer größer werdenden k gegen eine Zahl $P(A)$ „konvergiert“, die gleich der „Wahrscheinlichkeit“ dafür ist, dass das Ergebnis ω in A liegt.

Diese Idee ist zu ungenau.

4.6 Definition (Wahrscheinlichkeit)

(Axiomatische Definition der Wahrscheinlichkeit, Kolmogoroff 1933)

Ω sei eine (endliche) Menge und $\mathcal{A} \subset \mathcal{P}(\Omega)$, $\mathcal{P}(\Omega)$ ist die Potenzmenge von Ω, d. h. die Menge aller Teilmengen von Ω. Für $\mathcal{A}$ gelte: $\emptyset, \Omega \in \mathcal{A}$; $A, B \in \mathcal{A} \implies A \cup B \in \mathcal{A}$ und $\Omega \setminus A \in \mathcal{A}$.

Eine Wahrscheinlichkeitsverteilung ist eine Abbildung

$$P \colon \mathcal{A} \longrightarrow \mathbb{R} \ , \ A \longmapsto P(A)$$

mit folgenden Eigenschaften:

(1) $0 \leq P(A) \leq 1$ für alle $A \in \mathcal{A}$

(2) $P(\Omega) = 1$

(3) $P(A \cup B) = P(A) + P(B)$ falls $A \cap B = \emptyset \ , \ A, B \in \mathcal{A}$

Für nicht endliche Ω sind die Eigenschaften an $\mathcal{A}$ und Bedingung (3) zu modifizieren.

4.7 Geometrische Analogie für die Wahrscheinlichkeitsverteilung

Sei Ω das Quadrat mit Kantenlänge eins. Für gewisse Teilmengen $A \in \mathcal{A} \subset \mathcal{P}(\Omega)$ kann man den Flächeninhalt berechnen.
Definiert man $P(A) :=$ Fläche von A, so gelten die Axiome aus 4.6:

(a) $0 \leq P(A) \leq 1$

(b) $P(\Omega) = 1$

(c) $P(A \cup B) = P(A) + P(B)$ falls $A \cap B = \emptyset$

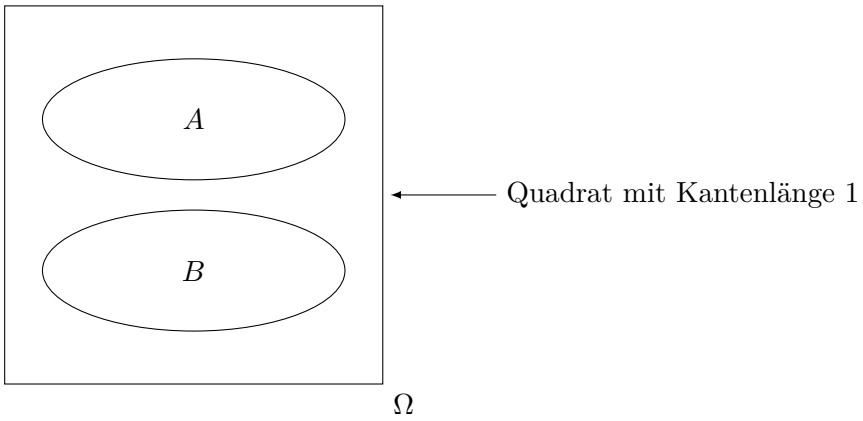


Abbildung 4.1: *Skizze zu (c)*

4.8 Satz (Ω endlich)

Sei Ω eine nicht leere Menge und $\mathcal{A} \subset \mathcal{P}(\Omega)$. $A_1, \ldots, A_r$ seien aus $\mathcal{A}$ und alle Operationen mit den A_i (Vereinigung, Durchschnitt usw.) seien ebenfalls aus $\mathcal{A}$. Erfüllt die Abbildung

$$P\colon \mathcal{A} \longrightarrow \mathbb{R}\ ,\ A \longmapsto P(A)$$

die Eigenschaften (1) bis (3) aus 4.6, so gilt:

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

(b) $P(A_1 \cup \ldots \cup A_r) = P(A_1) + \ldots + P(A_r)$, falls $A_i \cap A_j = \emptyset$ für $i \neq j$

(c) $P(\bar{A}) = 1 - P(A)$, wobei $\bar{A} = \Omega \setminus A$

(d) $A_1 \subset A_2 \implies P(A_1) \leq P(A_2)$

(e) $P(A_1 \cup A_2) = P(A_1) + P(A_2) - P(A_1 \cap A_2)$

(f) $P(A_1 \cup \ldots \cup A_r) \leq P(A_1) + \ldots + P(A_r)$

Beweis

(a) Folgt aus Axiomen (2) und (3) mit $A = \Omega$, $B = \emptyset$.

$$1 \overset{(2)}{=} P(\Omega) = P(\Omega \cup \emptyset) \overset{(3)}{=} P(\Omega) + P(\emptyset) \overset{(2)}{=} 1 + P(\emptyset)$$
$$\Longrightarrow \; P(\emptyset) = 0$$

(b) Wiederholte Anwendung von Axiom (3)

$$\begin{aligned} P(A_1 \cup \ldots \cup A_r) &= P\big(A_1 \cup (A_2 \cup \ldots \cup A_r)\big) \\ &\overset{(3)}{=} P(A_1) + P(A_2 \cup \ldots \cup A_r) \\ &\overset{(3)}{=} P(A_1) + P(A_2) + P(A_3 \cup \ldots \cup A_r) \\ &\overset{(3)}{=} \ldots \\ &\overset{(3)}{=} P(A_1) + \ldots + P(A_r) \end{aligned}$$

(c)

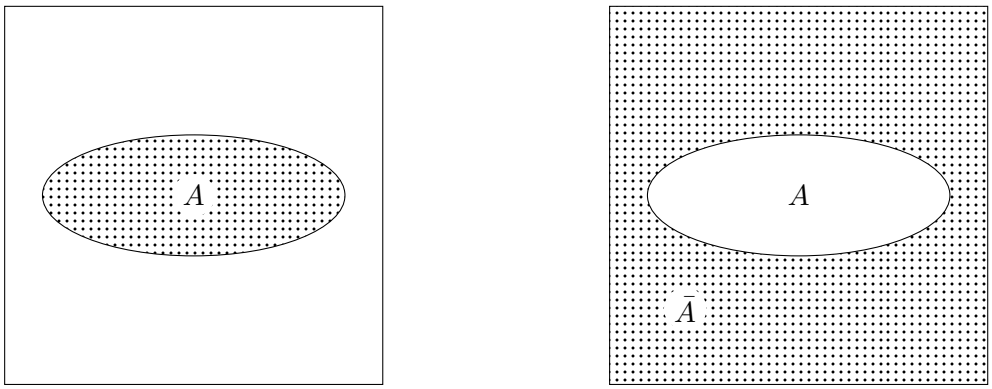


Abbildung 4.2

Folgt aus Axiomen (2) und (3).

$$1 \overset{(2)}{=} P(\Omega) = P(A \cup \bar{A}) \overset{(3)}{=} P(A) + P(\bar{A})$$
$$\Longrightarrow \; P(\bar{A}) = 1 - P(A)$$

(d) und (e) Übungsaufgabe.

(f) Wiederholte Anwendung von (e)

□

§ 5 Gleichverteilung und Zufallsvariable

5.1 Definition (Gleichverteilung/Laplace-Verteilung)

Es sei $\Omega = \{\omega_1, \ldots, \omega_n\}$. Durch

$$P(\omega) = \frac{1}{n} \quad \forall\, \omega \in \Omega$$

wird eine **Gleichverteilung** (oder **Laplace-Verteilung**) auf Ω definiert.

Ist $A \subset \Omega$, so gilt für die Wahrscheinlichkeit von A

$$P(A) = \frac{\#A}{\#\Omega}$$

Dabei bedeutet das Symbol # „Anzahl der Elemente von“.

5.2 Beispiel

Beim Glücksrad bedeutet Gleichverteilung, dass alle n Felder gleich groß sind. Bei $n = 37$ hat man ein Roulett mit $\Omega = \{0, 1, \ldots, 36\}$.

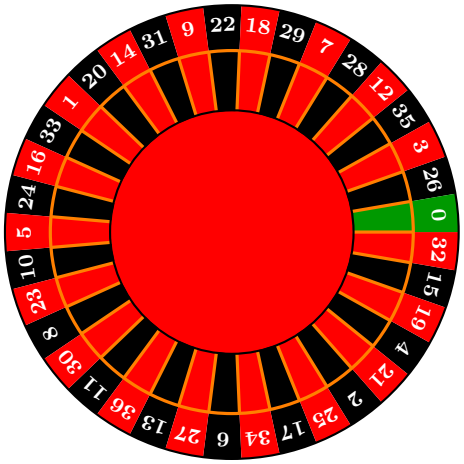


Abbildung 5.1: *Roulett*

5.3 Definition (Zufallsvariable/Zufallsgröße, Ausprägungen)

(a) Ein wichtiger Grundbegriff ist die **Zufallsvariable** (oder **Zufallsgröße**). Das ist eine Abbildung

$$X : \Omega \longrightarrow \mathbb{R}$$

Man kann sich dabei vorstellen, dass jedes Ergebnis ω eine gewisse Zahl $X(\omega)$ (etwa Gewinn oder Verlust) auslöst. Wie bei den „Merkmalen“ nennt man die Werte $a = X(\omega) \in \mathbb{R}$ **Ausprägungen** von X

(b) Für $a \in \mathbb{R}$ sei

$$\{X = a\} := \Big\{ \omega \in \Omega \;\Big|\; X(\omega) = a \Big\} \subset \Omega$$

und

$$P(X = a) := P(\{X = a\}) \in [0, 1]$$

die Wahrscheinlichkeit dafür, dass X den Wert a annimmt.

5.4 Notiz

Setzt man voraus, dass Ω endlich ist, dann hat auch X nur endlich viele Ausprägungen $a_1, \ldots, a_k$ und die von $a \in \mathbb{R}$ abhängige Funktion $P(X = a)$ hat nur an endlich vielen Stellen positive Werte.

5.5 Graphische Darstellung

(a) Darstellung von $P(X = a)$ durch ein Stabdiagramm.

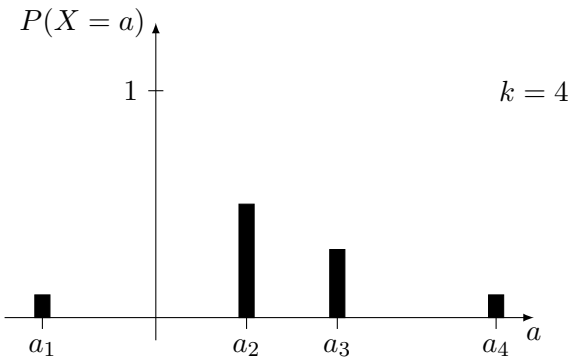


Abbildung 5.2

Dabei ist

$$\sum_{a \in \mathbb{R}} P(X = a) = 1$$

(Ω endlich, dann ist es eine endliche Summe)

(b) Darstellung durch eine Verteilungsfunktion

$$F_X \colon \mathbb{R} \longrightarrow [0, 1] \quad \text{mit } F_X(x) := P(X \leq x)$$

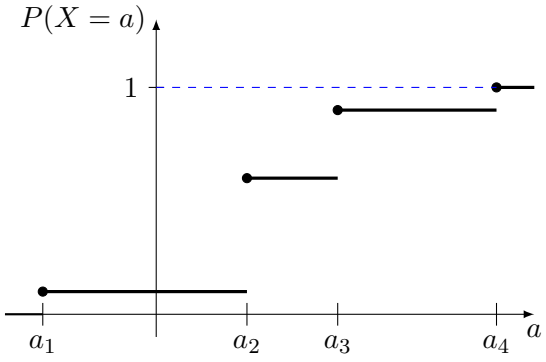


Abbildung 5.3

5.6 Bemerkung

Ist eine endliche Ergebnismenge Ω mit einer Wahrscheinlichkeitsverteilung P und eine Zufallsvariable

$$X : \Omega \longrightarrow \mathbb{R} \ , \ \omega \longmapsto X(\omega)$$

gegeben, so kann man den Wert $X(\omega)$ als ein neues von ω abhängiges Ergebnis betrachten; damit erhält man eine neuer Ergebnismenge

$$\Omega' := X(\Omega) \subset \mathbb{R}$$

Setzt man für X einen Wert $a \in \Omega'$

$$P'(a) := P(X = a)$$

so wird dadurch eine Wahrscheinlichkeitsverteilung P' auf Ω' erklärt.
Ein Vorteil von Ω' gegenüber Ω kann sein, dass $\Omega' \subset \mathbb{R}$. Die Verteilungsfunktion F_X von X ist durch P' bestimmt.

5.7 Beispiel

Wir würfeln mit einem Würfel zweimal hintereinander und bestimmen die Augensumme. Das Ergebnis liegt zwischen 2 und 12 und $\Omega' = \{2, 3, \ldots, 12\}$.

Gesucht ist eine geeignete Wahrscheinlichkeitsverteilung P' auf Ω', wenn der Würfel fair ist. Bei fairen Würfeln hat man Gleichverteilung der Wahrscheinlichkeit auf $\{1, \ldots, 6\}$. Das Ergebnis beim zweimaligen Wurf liegt in

$$\Omega = \{1, \ldots, 6\}^2 = \begin{Bmatrix} (1,1) & \ldots & (1,6) \\ \vdots & \ddots & \vdots \\ (6,1) & \ldots & (6,6) \end{Bmatrix}$$

Dabei ist es entscheidend zu bemerken, dass die Paare entsprechend der Reihenfolge der Würfe

geordnet sind, also etwa $(1,6) \neq (6,1)$. Auf diesem Ω ist Gleichverteilung angemessen (wie in Beispiel 5.2), also ist

$$P(\omega) = \frac{1}{36} \quad \forall\, \omega \in \Omega$$

Man erhält die Zufallsvariable

$$X\colon \Omega \longrightarrow \mathbb{R}\ ,\ \omega = (\omega_1, \omega_2) \longmapsto \omega_1 + \omega_2\ ,\ X(\Omega) = \Omega' \in \mathbb{R}$$

die mit der Augensumme bestimmt wird. Für $a \in \Omega'$ ist

$$P'(a) = P(X = a)$$

Dazu muss in Ω abgezählt werden, bei wieviel Paaren die Augensumme a entsteht. Das Ergebnis ist:

a	2	3	4	5	6	7	8	9	10	11	12
$36 \cdot P'(X = a)$	1	2	3	4	5	6	5	4	3	2	1

So gilt z. B.

$$P'(2) = \frac{1}{36}\ ,\ P'(5) = \frac{4}{36} = \frac{1}{9}\ ,\ P'(7) = \frac{6}{36} = \frac{1}{6}\ \ldots$$

P ist durch eine Gleichverteilung definiert.
P' und Ω' stellen KEINE Gleichverteilung dar.

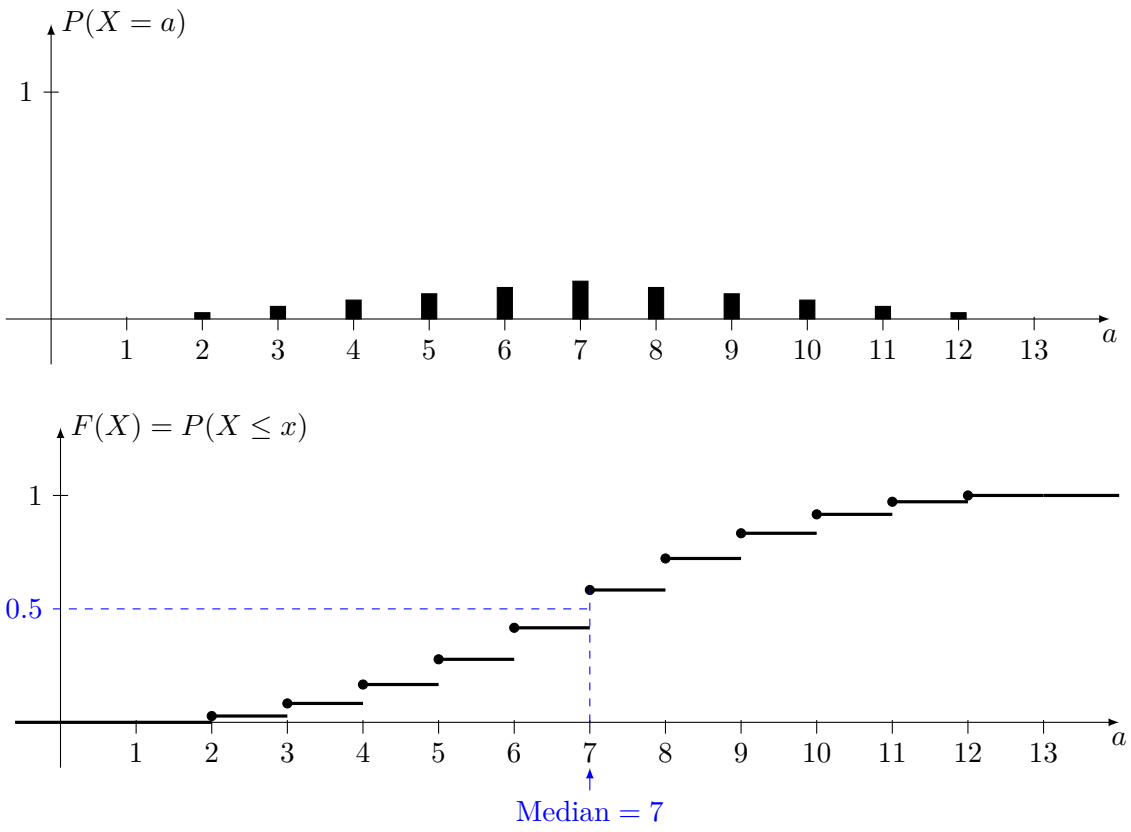


Abbildung 5.4: *Stabdiagramm und Verteilungsfunktion*

$$\bar{x} = \frac{1}{36}\left(\sum_{a=2}^{12} a \cdot \#(X = a)\right) = \frac{1}{36}(2 \cdot 1 + 3 \cdot 2 + 4 \cdot 3 + \ldots + 12 \cdot 1) = \frac{252}{36} = 7$$

(symmetrische Verteilungsform)

§ 6 Urnenmodelle

Viele Experimente kann man abstrakt beschreiben als Ziehung von Kugeln aus einer Urne.
Die Urne enthält n gleichartige Kugeln, nummeriert von eins bis n. Daraus werden k Kugeln gezogen.

6.1 Regeln mit Ergebnismengen

Sei $M = \{1, \ldots, n\}$. k Kugeln werden gezogen.

(1) Geordnet, mit Zurücklegen

$$\Omega_1 = \left\{ (a_1, \ldots, a_k) \in M^k \;\middle|\; 1 \leq a_i \leq n \right\}$$

(2) Geordnet, ohne Zurücklegen

$$\Omega_2 = \left\{ (a_1, \ldots, a_k) \in M^k \;\middle|\; a_i \neq a_j \text{ für } i \neq j \right\}$$

(3) Ungeordnet, mit Zurücklegen

$$\Omega_3 = \left\{ (a_1, \ldots, a_k) \in M^k \;\middle|\; 1 \leq a_1 \leq \ldots \leq a_k \leq n \right\}$$

(4) Ungeordnet, ohne Zurücklegen ($k \leq n$)

$$\Omega_4 = \left\{ (a_1, \ldots, a_k) \in M^k \;\middle|\; 1 \leq a_1 < a_2 < \ldots < a_k \leq n \right\}$$

Geordnet bedeute dabei, dass im Ergebnis die Reihenfolge der Züge berücksichtigt wird.
Ungeordnet bedeute, dass die Reihenfolge der Ziehungen unwesentlich ist; das Ergebnis wird nach der Größe der Kugelnummern sortiert.

6.2 Bemerkung

Es gelten:

(a)
$$\begin{array}{ccc} \Omega_1 & \supset & \Omega_2 \\ \bigcup & & \bigcup \\ \Omega_3 & \supset & \Omega_4 \end{array}$$

(b) $\Omega_4 = \Omega_2 \cap \Omega_3$

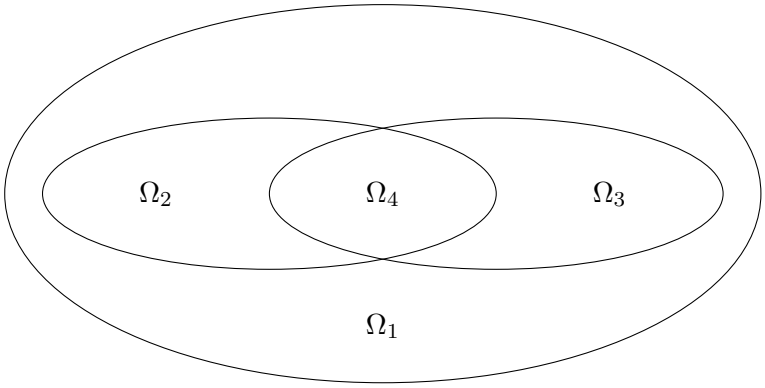


Abbildung 6.1: *Venndiagramm zu 6.2*

6.3 Definition (Fakultät, Binomialkoeffizient)

$n! := n(n-1) \cdot \ldots \cdot 2 \cdot 1$ (n-Fakultät)
$\binom{n}{k} := \frac{n!}{(n-k)!k!}$ (n über k), heißt Binomialkoeffizient; dabei ist $0! = 1$ und für $k = 0$ oder $k > n$ wird $\binom{n}{k} := 0$ gesetzt.

6.4 Bemerkung

Es gilt:

(a) $(a+b)^n = \sum_{k=0}^{n} \binom{n}{k} a^k b^{n-k}$ (Binomische Formel)

(b) $\binom{n}{k} = \binom{n}{n-k}$

(c) Die Binomialkoeffizienten sind ganzzahlig.

(d) $\sum_{k=0}^{n} \binom{n}{k} = 2^n$

Beweis

Siehe Analysis-Vorlesung.

□

6.5 Satz

Es gilt ($\#M$ bezeichne die Anzahl der Elemente von M):

(a) $\#\Omega_1 = n^k$

(b) $\#\Omega_2 = \frac{n!}{(n-k)!}$

(c) $\#\Omega_3 = \binom{n+k-1}{k}$

(d) $\#\Omega_4 = \binom{n}{k}$

Beweis

[Fischer][8], Seite 83ff.

□

6.6 Hinweis

Eine Veranschaulichung von Ω_i $(i = 1, \ldots, 4)$ für $n = 6$ und $k = 2$ durch Quadrate findet man bei [Fischer], Seite 76.

[8] Fischer, Gerd: Stochastik einmal anders. Vieweg Verlag 2005.

§ 7 Baumdiagramme und Pfadregeln

Baumdiagramme dienen der Berechnung von Wahrscheinlichkeiten bei mehrstufigen Zufallsexperimenten.

7.1 Beispiel

Zweimaliger Wurf einer Münze.

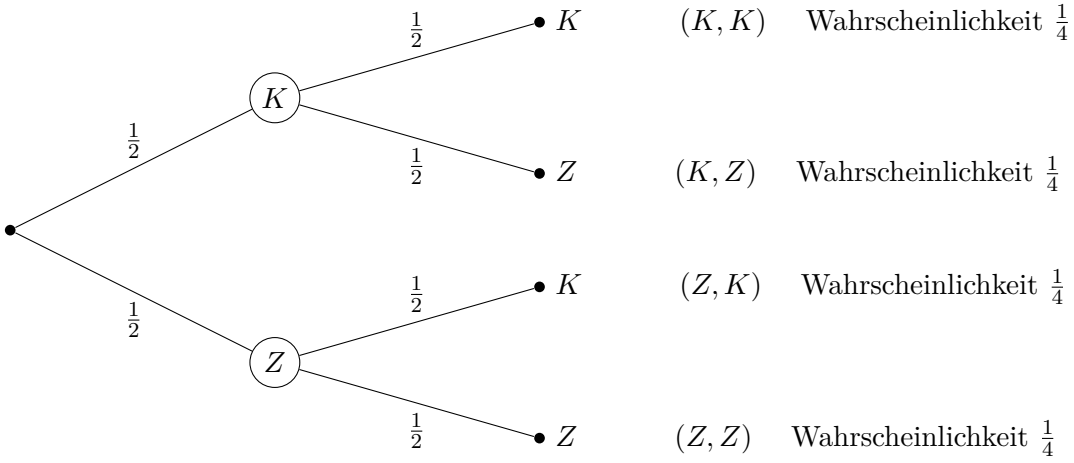


Abbildung 7.1: *Baumdiagramm*

7.2 Pfadmultiplikationsregel

Die Wahrscheinlichkeit, dass ein Ergebnis – z. B. (Z, K) – eintritt, berechnet man, indem man die Wahrscheinlichkeiten entlang des Pfades miteinander multipliziert.

7.3 Pfadadditionsregel

Die Wahrscheinlichkeit, dass ein Ereignis – z. B. $\{(Z, K); (K, Z)\}$ – eintritt, berechnet man, indem man die Pfad-Wahrscheinlichkeiten alles Punkte, die zu dem Ereignis gehören, addiert.

7.4 Beispiel zu 7.1

Wahrscheinlichkeit $\big((Z, Z)\big) = \frac{1}{2} \cdot \frac{1}{2} = \frac{1}{4}$
Wahrscheinlichkeit $\big(\{(Z, K); (K, Z)\}\big) = \frac{1}{4} + \frac{1}{4} = \frac{1}{2}$

Weitere Beispiele in der Zentralübung.

§ 8 Bedingte Wahrscheinlichkeiten und unabhängige Ereignisse

Ω sei eine (endliche) Menge, $\mathcal{A} \subset \mathcal{P}(\Omega)$, $A, B \ldots \in \mathcal{A}$. $P\colon \mathcal{A} \longrightarrow \mathbb{R}$ eine Wahrscheinlichkeitsverteilung.

8.1 Definition (Bedingte Wahrscheinlichkeit)

$A, B \in \mathcal{A}$, $P(B) > 0$. Dann heißt

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

die **bedingte Wahrscheinlichkeit** von A unter der Bedingung B.

Bemerkung: $P(B|B) = 1$

8.2 Beispiel

Nachwahl einer Person in den Betriebsrat. Kandidaten:

	angestellt	Arbeiter	Zeilensumme
weiblich	3	2	5
männlich	3	5	8
Spaltensumme	6	7	13

Ereignisse:
A „die gewählte Person ist angestellt“
M „die gewählte Person ist männlich“

Wahl soll einem Laplace-Experiment genügen. Dann gilt:

$$P(A) = \frac{6}{13} \quad , \quad P(M) = \frac{8}{13} \quad , \quad P(A \cap M) = \frac{3}{13}$$

Nach der Wahl und vor der Veröffentlichung des Wahlergebnisses ist bekannt geworden, dass ein Mann gewählt wurde. Es gilt:

$$P(A|M) = \frac{3}{8} = \frac{\frac{3}{13}}{\frac{8}{13}} = \frac{P(A \cap M)}{P(M)} \neq P(A)$$

8.3 Satz (Multiplikationssatz)

$A, B \in \mathcal{A}$, $P(B) > 0$. Dann gilt:

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

8.4 Bemerkung

Sei $B \in \mathcal{A}$ und $P(B) > 0$. Durch $P_B(A) := P(A|B)$ wird eine Wahrscheinlichkeitsverteilung definiert.

$$P_B \colon \left\{ C \cap B \;\middle|\; C \in \mathcal{A} \right\} \longrightarrow \mathbb{R}$$

erfüllt die Axiome aus 4.6.

8.5 Definition (stochastisch unabhängig)

Sei $B \in \mathcal{A}$ und $0 < P(B) < 1$. Dann heißt $A \in \mathcal{A}$ (stochastisch) **unabhängig** von B
$:\Longleftrightarrow \; P(A|B) = P(A|\bar{B})$

8.6 Satz

$A, B \in \mathcal{A}$, $0 < P(B) < 1$. Dann gilt:
A (stochastisch) unabhängig von B
$\Longleftrightarrow \; P(A|B) = P(A)$
$\Longleftrightarrow \; P(A \cap B) = P(A) \cdot P(B)$

am häufigsten verwendet

8.7 Bemerkung

$A, B \in \mathcal{A}$ und $0 < P(A) < 1$, $0 < P(B) < 1$.
A unabhängig von $B \iff B$ unabhängig von A.

Bemerkung: $P(A|B) \overset{\text{i.A.}}{\neq} P(B|A)$

8.8 Definition (Disjunktion (Zerlegung) von Ω)

$A_1, \ldots, A_n \in \mathcal{A}$ bilden eine Disjunktion (Zerlegung) von Ω
$:\Longleftrightarrow \; A_1 \cup \ldots \cup A_n = \Omega$ und $A_i \cap A_j = \emptyset$ für $i \neq j$.

Bemerkung: A, B heißen disjunkt $:\Longleftrightarrow \; A \cap B = \emptyset$

8.9 Satz

$A_1, \ldots, A_n$ Disjunktion von Ω, $P(A_i) > 0 \quad \forall\, i$ $\left(\text{damit ist } \sum_{i=1}^{n} P(A_i) = 1\right)$.
Dann gilt für $B \in \mathcal{A}$:

$$P(B) = \sum_{i=1}^{n} P(B|A_i) \cdot P(A_i)$$

8.10 Beispiel

Bekannt seien

$$P(A) = \frac{3}{5}\,, \quad \left[P(\bar{A}) = \frac{2}{5}\right] \,, \quad P(B|A) = \frac{5}{9}\,, \quad P(B|\bar{A}) = \frac{6}{9}$$

Berechne $P(B)$.
Es gilt: $A \cup \bar{A} = \Omega$ und $A \cap \bar{A} = \emptyset$. Mit Satz 8.9 erhalten wir

$$P(B) = P(B|A) \cdot P(A) + P(B|\bar{A}) \cdot P(\bar{A}) = \frac{5}{9} \cdot \frac{3}{5} + \frac{6}{9} \cdot \frac{2}{5} = \frac{3}{5}$$

8.11 Satz (Bayessche Formel)

$A_1, \ldots, A_n$ Disjunktionen von Ω, $P(A_i) > 0 \quad \forall\, i$ und $B \in \mathcal{A}$ mit $P(B) > 0$. Dann gilt:

$$P(A_i|B) \overset{8.3,8.1}{=} \frac{P(B|A_i) \cdot P(A_i)}{P(B)} \overset{8.9}{=} \frac{P(B|A_i) \cdot P(A_i)}{\sum\limits_{i=1}^{n} P(B|A_i) \cdot P(A_i)}$$

8.12 Beispiel

Bekannt seien

$$P(A_1) = 0.2\,,\ P(A_2) = 0.3\,,\ P(A_3) = 0.5\,,\ P(B|A_1) = 0.02\,,\ P(B|A_2) = 0.04\,,\ P(B|A_3) = 0.07$$

mit A_1, A_2, A_3 disjunkt. Ohne Einschränkung kann angenommen werden, dass A_1, A_2, A_3 eine Disjunktion von Ω ist.
Dann ist

$$P(B) = \sum_{i=1}^{3} P(B|A_i) \cdot P(A_i) = 0.051$$

$$P(A_1|B) = \frac{0.02 \cdot 0.2}{0.051} = \frac{4}{51} \qquad \left[\neq 0.02 = P(B|A_1) \right]$$

$$P(A_2|B) = \frac{12}{51} \qquad \left[\neq P(B|A_2) \right]$$

$$P(A_3|B) = \frac{35}{51} \qquad \left[\neq P(B|A_3) \right]$$

§ 9 Verteilungen

(Ordnung in die vielfältigen Wahrscheinlichkeitsmodelle bringen.)

9.1 Definition (Ziehen ohne Zurücklegen)

Eine Zufallsvariable X mit Werten in $\{0, 1, \ldots, n\}$ heißt **binomialverteilt** mit den Parametern $n \in \mathbb{N}$ und $p \in [0, 1]$, wenn gilt

$$(*) \quad b_{n;p}(k) = B(n; p; k) := P(X = k) = \binom{n}{k} p^k (1-p)^{n-k}$$

für $k \in \{0, 1, \ldots, n\}$

9.2 Notiz

Eine Herleitung der Formel $(*)$ aus 9.1 über Urnenmodelle findet man bei [Fischer], Abschnitt 2.4.

9.3 Bemerkung

X sei binomialverteilt mit den Parametern n und p. Dann gilt:

(a) $P(l \leq X \leq m) = \sum_{k=l}^{m} P(X = k) = \sum_{k=l}^{m} b_{n;p}(k)$

(b) $P(0 \leq X \leq n) = \sum_{k=0}^{n} \binom{n}{k} p^k (1-p)^{n-k} \overset{\text{BLS}}{=} \big(p + (1-p)\big)^n = 1$

(c) $b_{n;0.5}(k) = \frac{1}{2^n} \binom{n}{k}$

(d) $P(X \geq k) = 1 - P(X < k) = 1 - \sum_{k=0}^{k-1} b_{n;p}(l)$

Bemerkung: $\sum_{k=0}^{k-1} b_{n;p}(l)$ findet man oft in Tabellen

9.4 Beispiel

Ein fairer Würfel wird n-mal geworfen. Zu berechnen sind die Wahrscheinlichkeiten folgender Ereignisse:

(a) Genau k-mal die Augenzahl 6.

(b) Mindestens k-mal die Augenzahl 1.

(c) Genau k-mal eine gerade Augenzahl.

(d) Mindestens k-mal eine gerade Augenzahl.

Lösung

(a) Die Zufallsvariable X zähle die Anzahl „Eine 6 wird gewürfelt“.
X ist dann binomialverteilt mit den Parametern n und $p = \frac{1}{6}$.

$$P(X = k) = n_{n;\frac{1}{6}}(k) = \binom{n}{k} \left(\frac{1}{6}\right)^k \left(\frac{5}{6}\right)^{n-k} = \binom{n}{k} \cdot \frac{5^{n-k}}{6^n}$$

z. B. $n = 10$, $k = 5$:

$$P(X = 5) = b_{10;\frac{1}{6}}(5) = \binom{10}{5} \cdot \frac{5^5}{6^{10}} \approx 0.013$$

(b) Die Zufallsvariable X zähle die Anzahl „Eine 1 wird gewürfelt“.
X ist dann binomialverteilt mit den Parametern n und $p = \frac{1}{6}$.

$$P(X \geq k) = 1 - P(X < k) = 1 - \sum_{l=0}^{k-1} b_{n;p}(l)$$

z. B. $n = 10$, $k = 5$:

$$P(X \geq 5) = 1 - \underbrace{\sum_{l=0}^{4} b_{10;\frac{1}{6}}(l)}_{\approx 0.98454 \text{ (Tabelle)}} \approx 0.01546$$

(c) Die Zufallsvariable X zähle „Eine gerade Zahl wird gewürfelt“. X ist dann binomialverteilt mit den Parametern n und $p = \frac{1}{2}$.

$$P(X = k) = b_{n;\frac{1}{2}}(k) = \frac{1}{2^n} \cdot \binom{n}{k}$$

z. B. $n = 10$, $k = 3$:

$$P(X = 3) = \underbrace{b_{n;\frac{1}{2}}(3)}_{\approx 0.11719 \text{ (Tabelle)}} = \underbrace{\frac{1}{2^{10}} \cdot \binom{10}{3}}_{= \frac{1}{2^{10}} \cdot \frac{10!}{7! \cdot 3!} = \frac{1}{2^{10}} \cdot \frac{10 \cdot 9 \cdot 8}{3 \cdot 2 \cdot 1}} \approx 0.1171825$$

(d) X, n, p wie bei (c).

$$P(X \geq k) = 1 - P(X < k) = 1 - \sum_{l=0}^{k-1} n_{n;\frac{1}{2}}(l)$$

z. B. $n = 10$, $k = 3$:

$$P(X \geq 3) = 1 - \underbrace{\sum_{l=0}^{2} b_{n;\frac{1}{2}}(l)}_{\text{(Tabelle)}} \approx 0.9453$$

9.5 Grafiken zur Binomialverteilung

(siehe Anhang A)

Wir betrachten Diagramme zur Binomialverteilung $B(n;p)$.
Stabdiagramm: Je kleiner p, desto linkssteiler die Verteilung. Je größer p, desto rechtssteiler die Verteilung.
Für die Modulprüfung sollte man in der Lage sein, Stabdiagramm und Verteilungsfunktion zu zeichnen und p zu einem Histogramm (bzgl. links- und rechtssteil) zuzuordnen.

9.6 Definition (hypergeometrisch verteilte Zufallsvariable)

(Ziehen ohne Zurücklegen. Vgl. [Fischer] für dazugehöriges Urnenmodell)
Seien $n, r, N \in \mathbb{N}$ mit $n, r \leq N$.
Eine Zufallsvariable X heißt **hypergeometrisch verteilt** mit Parametern n, r, N, wenn gilt

$$(*) \quad P(X = k) = \frac{\binom{r}{k}\binom{N-r}{n-k}}{\binom{N}{n}}$$

wobei $k \geq \max(0, n - N + r)$ und $k \leq \min(n, r)$.
Notation: $H(n;N;r)$-verteilt

9.7 Bemerkung

$$\lim_{N\to\infty} \frac{\binom{r}{k}\binom{N-r}{n-k}}{\binom{N}{n}} = \binom{n}{k} p^k (1-p)^{n-k}$$

unter der Voraussetzung $p = \frac{r}{N}$ und $\lim\limits_{N\to\infty} \frac{r(N)}{N} = p$.

9.8 Beispiel

Eine Lieferung von 100 Dioden enthalte genau vier fehlerhafte. Aus der Lieferung werden ohne Zurücklegen fünf Dioden entnommen. Die Zufallsvariable X beschreibe die Anzahl der fehlerhaften Dioden unter den fünf entnommenen Dioden.
$n = 5, r = 4, N = 100$. Es gilt für $k = 1$:

$$P(X = 1) = \frac{\binom{4}{1}\binom{96}{4}}{\binom{100}{5}} \approx 0.1765$$

9.9 Definition (stetig verteilte Zufallsvariable)

Eine Zufallsvariable X heißt **stetig verteilt** mit der Dichte f
$:\Longleftrightarrow$ f ist integrierbar und für die Verteilungsfunktion $F : \mathbb{R} \longrightarrow [0, 1]$ gilt:

$$F(x) = \int_{-\infty}^{x} f(t)\,\mathrm{d}t \quad \forall\, x \in \mathbb{R}$$

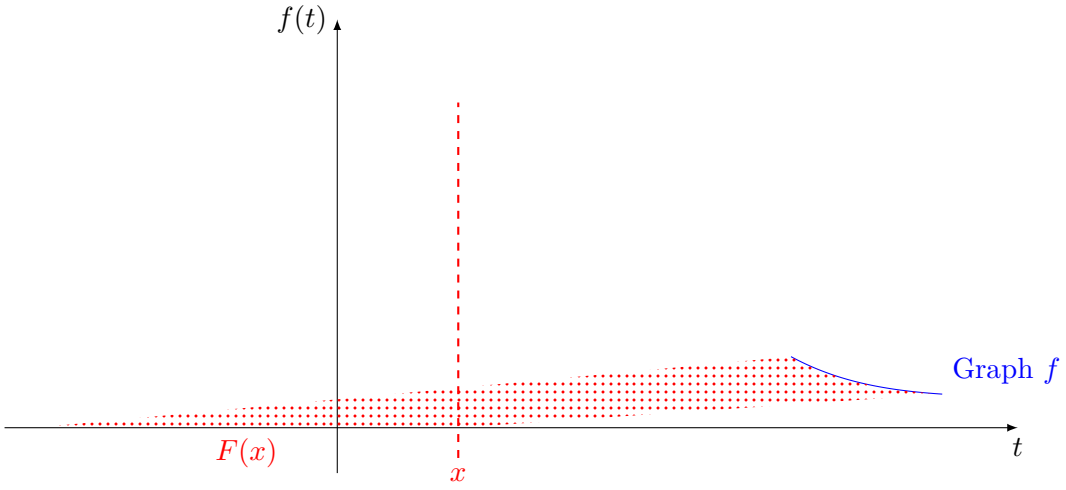


Abbildung 9.1

9.10 Veranschaulichung Wahrscheinlichkeiten $\longleftrightarrow$ Verteilungsfunktion

Sei X stetig verteilte Zufallsgröße mit Dichte f. Dann gelten folgende Veranschaulichungen:

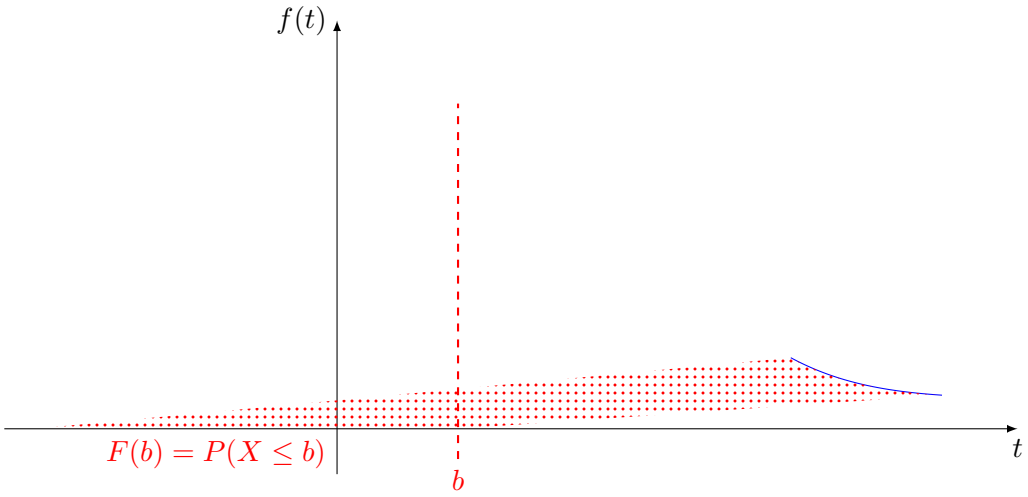


Abbildung 9.2: $F(b) = P(X \leq b)$

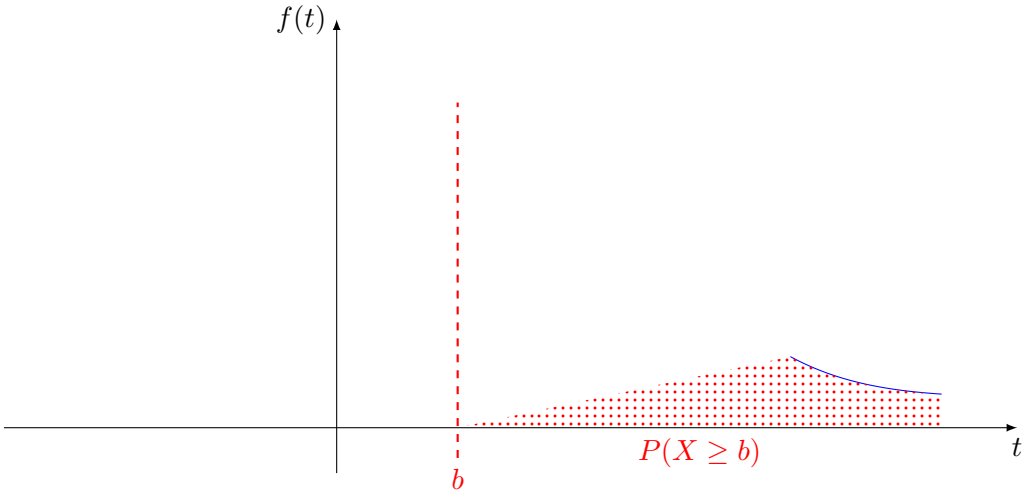


Abbildung 9.3: $P(X \geq b) = 1 - P(X < b) \overset{(*)}{=} 1 - P(X \leq b) = 1 - F(b)$
$(*)$ *X stetig verteilt*

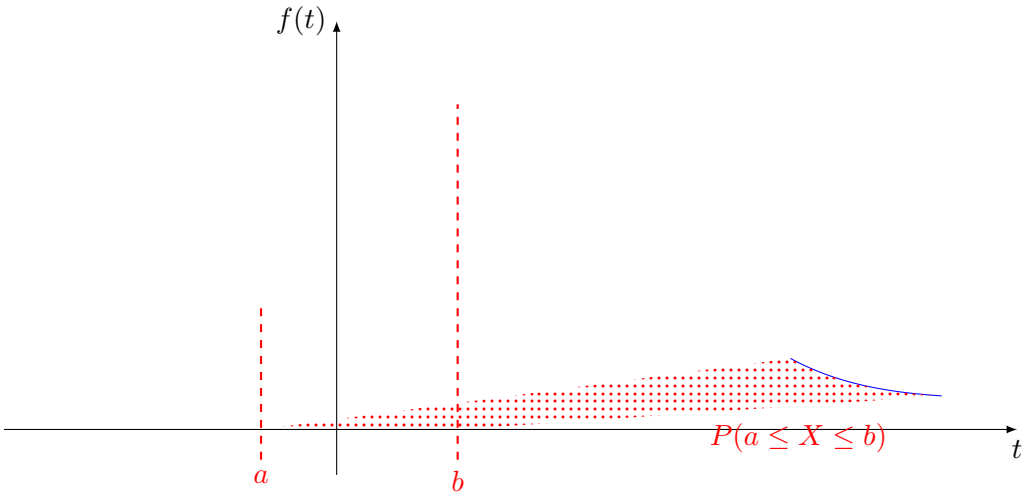


Abbildung 9.4: $P(a \leq X \leq b) = F(b) - F(a) = P(X \leq b) - P(X \leq a) = P(X \leq b) - P(X < a)$

9.11 Definition ((standard-)normalverteilte Zufallsvariable)

(a) Es seien $\mu, t \in \mathbb{R}$ und $\sigma > 0$. Eine Zufallsvariable X heißt **normalverteilt** mit Parametern μ und σ (kurz: $N(\mu, \sigma)$-verteilt $(*)$), falls X stetig verteilt ist mit Dichte f definiert durch

$$f(t) = \frac{1}{\sigma \cdot \sqrt{2\pi}} \cdot \exp\left(-\frac{1}{2}\left(\frac{t-\mu}{\sigma}\right)^2\right) \quad \forall\, t \in \mathbb{R}$$

Formeln bzgl. Verteilungen für Modulprüfung wichtig!

μ ist der Erwartungswert von X und σ die Standardabweichung von X.
$(*)$ ▽ In manchen Büchern wird die Abkürzung $N(\mu, \sigma^2)$ benutzt.

(b) Eine $N(0,1)$-verteilte Zufallsvariable heißt **standard-normalverteilt**.

9.12 Bemerkung

(a) Ist X eine $N(0,1)$-verteilte Zufallsvariable, so gilt für die Verteilungsfunktion Φ von X ($\Phi\colon \mathbb{R} \longrightarrow [0,1]$):

$$\Phi(x) = \frac{1}{\sqrt{2\pi}} \cdot \int_{-\infty}^{x} \exp\left(-\frac{1}{2}t^2\right) \mathrm{d}t$$

(b) Es gilt:

$$\varphi(x) = \frac{1}{\sqrt{2\pi}} \cdot \exp\left(-\frac{1}{2}x^2\right) \quad \forall\, x \in \mathbb{R}$$

ist die **Dichtefunktion** zur $N(0,1)$-verteilten Zufallsvariable.

(c) Es gilt:

sehr wichtige Formeln

$$\varphi(-x) = \varphi(x) \quad \forall\, x \in \mathbb{R}$$

und

$$\Phi(-x) = 1 - \Phi(x) \quad \forall\, x \in \mathbb{R}$$

9.13 Skizze zur Dichtefunktion

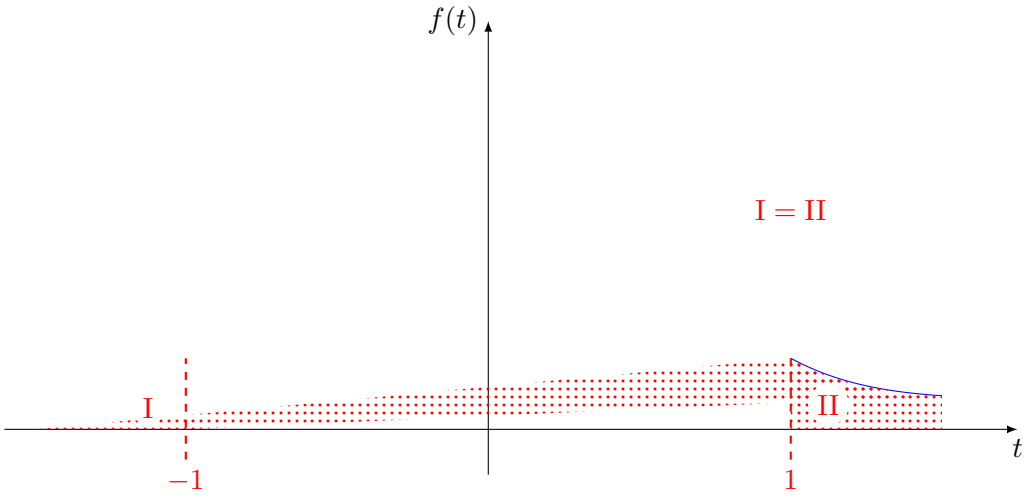


Abbildung 9.5: *Skizze zur Dichtefunktion*

9.14 Bemerkung

(a) Ist X eine $N(\mu,\sigma)$-verteilte Zufallsvariable, so ist

$$U := \frac{X - \mu}{\sigma}$$

eine $N(0,1)$-verteilte Zufallsvariable.

(b) Ist X eine $N(\mu,\sigma)$-verteilte Zufallsvariable, so gilt:

$$P(X < b) \overset{(*)}{=} P(X \leq b) = \Phi\left(\frac{b-\mu}{\sigma}\right)$$

(∗) X stetig verteilt

Für die Bestimmung von Wahrscheinlichkeiten zu normalverteilten Zufallsvariablen genügt die Kenntnis der Verteilungsfunktion Φ der $N(0,1)$-Verteilung für Werte $x \geq 0$.

9.15 Tabelle zu Werten von Φ

x	0.00	0.25	0.5	0.75	1	1.25	1.5	2	3
$\approx \Phi(x)$	0.5	0.5987	0.6915	0.7734	0.8413	0.8944	0.9332	0.9772	0.999

9.16 Beispiel

(a) X sei $N(\mu,\sigma)$-verteilt. Dann gilt:

$$\begin{aligned} P(\mu-\sigma<X<\mu+\sigma) &= P(X<\mu+\sigma)-P(X\leq\mu-\sigma) \\ &= \Phi\left(\frac{\mu+\sigma-\mu}{\sigma}\right)-\Phi\left(\frac{\mu-\sigma-\mu}{\sigma}\right) \\ &= \Phi(1)-\Phi(-1) \\ &= \Phi(1)-\big(1-\Phi(1)\big) = 2\Phi(1)-1 \approx 1.6826-1 = 0.6826 \end{aligned}$$

(b) X sei $N(1,2)$-verteilt.

$$P(X\geq 2) = 1-\underbrace{P(X<2)}_{=P(X\leq 2)} = 1-\Phi\left(\frac{2-1}{2}\right) = 1-\Phi(0.5) \approx 1-0.6915 = 0.3085$$

9.17 Bemerkung (Approximation der Binomialverteilung durch die Normalverteilung)

Für $p \in]0,1[$ und „große n“ (mindestens $n\cdot p\cdot(1-p)>9$) ist

$$B(n;p;k) \approx \frac{1}{\sigma}\cdot\varphi\left(\frac{k-\mu}{\sigma}\right)$$

wobei $\mu = n\cdot p$, $\sigma = \sqrt{np(1-p)}$.

9.18 Beispiel

Berechne $B(100;0.3;40)$ mit Hilfe der Normalverteilung.
$n=100$, $p=0.3$, $k=40$. Somit ist

$$\mu = n\cdot p = 30$$
$$\sigma = \sqrt{np(1-p)} = \sqrt{100\cdot 0.3\cdot 0.7} = \sqrt{21}$$

$$\begin{aligned} B(100;0.3;40) &= \frac{1}{\sigma}\cdot\varphi\left(\frac{k-\mu}{\sigma}\right) = \frac{1}{\sqrt{21}}\cdot\varphi\left(\frac{40-30}{\sqrt{21}}\right) \\ &\approx \frac{1}{\sqrt{21}}\cdot\varphi(2.18) \overset{\text{Tabelle}}{\approx} \frac{1}{\sqrt{21}}\cdot 0.037 \approx 0.008 \end{aligned}$$

§ 10 Erwartungswert, Varianz, Covarianz

10.1 Definition (Erwartungswert)

Unter dem **Erwartungswert** einer Zufallsvariablen

$$X\colon \Omega \longrightarrow \mathbb{R}$$

versteht man die Zahl

$$E(X) := \sum_{\omega\in\Omega} X(\omega)\cdot P(\omega) = \sum_{k=1}^{r} a_k \cdot P(X = a_k)$$

wenn $a_1, \ldots, a_r \in \mathbb{R}$ die Ausprägungen, d. h. die paarweise verschiedenen Werte von X bezeichnen.

Für stetige Zufallsvariablen mit Dichtefunktion f gilt:

$$E(X) = \int_{-\infty}^{\infty} t\cdot f(t)\,\mathrm{d}t$$

10.2 Definition

Sind zwei Zufallsvariablen $X, Y\colon \Omega \longrightarrow \mathbb{R}$ und $\lambda \in \mathbb{R}$ gegeben, so kann man daraus neue Zufallsvariablen erklären durch

$$\begin{aligned}(X+Y)(\omega) &:= X(\omega) + Y(\Omega)\\ (\lambda X)(\omega) &:= \lambda\cdot X(\omega)\\ (X\cdot Y)(\omega) &:= X(\omega)\cdot Y(\omega)\end{aligned}$$

10.3 Bemerkung (Rechenregeln für Erwartungswerte)

Seien $X, Y\colon \Omega \longrightarrow \mathbb{R}$ Zufallsvariablen und $\lambda \in \mathbb{R}$. Dann gilt:

$$\begin{aligned}E(X+Y) &= E(X) + E(Y)\\ E(\lambda X) &= \lambda E(X)\end{aligned}$$

10.4 Notiz

Im Allgemeinen ist $E(X\cdot Y) \overset{\text{i.A.}}{\neq} E(X)\cdot E(Y)$ $(*)$.
Gleichheit gilt in $(*)$, wenn die Zufallsvariablen „(stochastisch) unabhängig“ sind (siehe später).

10.5 Definition (Varianz)

Sei $X\colon \Omega \longrightarrow \mathbb{R}$ eine Zufallsvariable und $\mu := E(X)$.
Dann ist die **Varianz** von X definiert durch

$$\mathrm{Var}(X) = \begin{cases} \sum\limits_{\omega\in\Omega} \big(X(\omega)-\mu\big)^2 \cdot P(\omega)\ ,\ P(\omega)\neq 0 & \text{für endlich viele } \omega\in\Omega \\ \int\limits_{-\infty}^{\infty} (t-\mu)^2 f(t)\,\mathrm{d}t & \text{für stetige Zufallsvariablen mit Dichte } f \end{cases}$$

10.6 Bemerkung

Für Ω endlich oder $X(\Omega)$ endlich kann man gleiche Werte von X zusammenfassen und erhält

$$\mathrm{Var}(X) = \sum_{k=1}^{r} (a_k-\mu)^2 \cdot P(X=a_k)$$

wobei $a_1,\ldots,a_r$ die verschiedenen Werte von X sind.

10.7 Definition (Standardabweichung)

Die Zahl

$$r_x := \sqrt{\mathrm{Var}(X)}$$

heißt **Standardabweichung** von X.

10.8 Bemerkung (Rechenregeln für die Varianz)

Es sei X eine Zufallsvariable, $\mu = E(X)$ und $\lambda\in\mathbb{R}$. Dann gilt:

(a) $\mathrm{Var}(X) = E\big((X-\mu)^2\big) = E(X^2) - \mu^2$

(b) $\mathrm{Var}(\lambda X) = \lambda^2 \cdot \mathrm{Var}(X)$

(c) $\mathrm{Var}(X+\lambda) = \mathrm{Var}(X)$

10.9 Notiz

Im Allgemeinen ist $\mathrm{Var}(X+Y) \overset{\text{i.A.}}{\neq} \mathrm{Var}(X) + \mathrm{Var}(Y)$.

Beispiel

$Y = X$, $\mathrm{Var}(X+X) = \mathrm{Var}(2X) = 4\,\mathrm{Var}(X) \overset{\text{i.A.}}{\neq} \mathrm{Var}(X) + \mathrm{Var}(X)$

10.10 Definition und Notiz

X sei eine Zufallsvariable, $\mu = E(X)$ und σ_X die Standardabweichung von X. Die Zufallsvariable

$$X^{\varphi} := \frac{X - \mu}{\sigma_X}$$

heißt die **Standardisierung** von X. Allgemein heißt X standardisiert, wenn $E(X) = 0$ und $\mathrm{Var}(X) = 1$.

10.11 Definition (Covarianz, Korrelationskoeffizient)

Sei Ω endlich, $X, Y \colon \Omega \in \mathbb{R}$ seien Zufallsvariablen, $\mu := E(X)$, $\nu := E(Y)$.
Dann heißt

$$\sigma_{x,y} = \mathrm{Cov}(X,Y) = \sum_{\omega \in \Omega} \big(X(\omega) - \mu\big)\big(Y(\omega) - \nu\big) = E\big((X-\mu)(Y-\nu)\big)$$

die **Covarianz** von X und Y und

$$\varrho = \varrho_{x,y} := \frac{\mathrm{Cov}(X,Y)}{\sqrt{\mathrm{Var}(X) \cdot \mathrm{Var}(Y)}} = \frac{\mathrm{Cov}(X,Y)}{\sigma_x \cdot \sigma_y}$$

heißt **Korrelationskoeffizient** von X und Y.

10.12 Notiz

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

10.13 Satz (Rechenregeln im Zusammenhang mit der Covarianz)

X und Y seien Zufallsvariablen. Dann gilt:

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

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

(c) Sind X und Y (stochastisch) unabhängig, d. h.

$$P(X = a, Y = b) = P(X = a) \cdot P(Y = b)$$

so ist

$$E(X \cdot Y) = E(X) \cdot E(Y) \quad \text{und} \quad \mathrm{Var}(X+Y) = \mathrm{Var}(X) + \mathrm{Var}(Y)$$

(d) $X_1, \ldots, X_n$ seien unabhängige Zufallsvariablen, d. h.

$$P(X_1 = a_1, \ldots, X_n = a_n) = P(X_1 = a_1) \cdot \ldots \cdot P(X_n = a_n)$$

so gilt

$$\mathrm{Var}(X_1 + \ldots + X_n) = \mathrm{Var}(X_1) + \ldots + \mathrm{Var}(X_n)$$

10.14 Beispiel

X und Y seien Zufallsvariablen mit gemeinsamer Verteilung

X \ Y	1	2	3	4
1	0.05	0.1	0	0
2	0.1	0.2	0.05	0
3	0	0.1	0.1	0.2
4	0	0	0.05	0.05

$P(X = 1, Y = 1) = 0.05$
$P(X = 1, Y = 2) = 0.1$
$\dots$

Berechnen Sie die Covarianz $\sigma_{x,y}$ und den Korrelationskoeffizienten ϱ.
Es gilt

x_i	1	2	3	4
$P(X = x_i)$	0.15	0.35	0.4	0.1

Y_i	1	2	3	4
$P(X = x_i)$	0.15	0.4	0.2	0.25

$E(X) = 1 \cdot 0.15 + 2 \cdot 0.35 + 3 \cdot 0.4 + 4 \cdot 0.1 = 2.45$
$E(Y) = 1 \cdot 0.15 + 2 \cdot 0.4 + 3 \cdot 0.2 + 4 \cdot 0.25 = 2.55$

X \ Y	1	2	3	4	$P(X = x_i)$
1	0.05	0.1	0	0	0.15
2	0.1	0.2	0.05	0	0.35
3	0	0.1	0.1	0.2	0.4
4	0	0	0.05	0.05	0.1
$P(Y = y_i)$	0.15	0.4	0.2	0.25	1

$E(X^2) = 1 \cdot 0.15 + 4 \cdot 0.35 + 9 \cdot 0.4 + 16 \cdot 0.1 = 6.75$

$E(Y^2) = 1 \cdot 0.15 + 4 \cdot 0.4 + 9 \cdot 0.2 + 16 \cdot 0.25 = 7.55$

$E(X \cdot Y) = \sum_{i,k=1,\ldots,4} x_i y_k \cdot P(X = x_i, Y = y_k) = \ldots = 6.85$

$\sigma_x^2 = \mathrm{Var}(X) = \sum_{i=1}^{4} \big(x_i - E(X)\big)^2 \cdot P(X = x_i)$

oder $\mathrm{Var}(X) = E(X^2) - \big(E(X)\big)^2 = 6.75 - 2.45^2 = 0.7475$

$\sigma_y^2 = \mathrm{Var}(Y) = E(Y^2) - \big(E(Y)\big)^2 = 7.55 - 2.55^2 = 1.0475$

$\sigma_{x,y} = \mathrm{Cov}(X, Y) = E(X \cdot Y) - E(X) \cdot E(Y) = 6.85 - 2.45 \cdot 2.55 = 0.6025$

$\varrho = \dfrac{\mathrm{Cov}(X, Y)}{\sigma_x \cdot \sigma_y} = \dfrac{0.6025}{\sqrt{0.7475} \cdot \sqrt{1.0475}} \approx 0.6809$

10.15 Satz

Vor: $X_1, \ldots, X_n$ seien paarweise unabhängige Zufallsvariablen mit $E(X_i) = \mu$, $\mathrm{Var}(X_i) = \sigma^2$ für $i = 1, \ldots, n$. Weiter sei $\bar{X} = \frac{1}{n} \sum_{i=1}^{n} X_i$, $S^2 = \frac{1}{n-1} \sum_{i=1}^{n} (X_i - \bar{X})^2$.

Beh: $E(S^2) = \sigma^2$, d. h. S^2 ist erwartungstreu.

Kapitel III: Schließende Statistik

Als exemplarisches Beispiel dient ein Signifikanztest.

§ 11 Signifikanztests

11.1 Definition (Signifikanztest, Nullhypothese)

Ein Test, bei dem entschieden werden soll, ob eine Hypothese H_0 abgelehnt wird oder nicht, heißt **Signifikanztest**. Die zu testende Hypothese H_0 heißt **Nullhypothese**.

11.2 Notiz und Definition

Bei einem Signifikanztest wählt man eine Zahl $\alpha \in \mathbb{R}$ mit $0 < \alpha < 1$. Der Test ist dann so zu konstruieren, dass die Wahrscheinlichkeit(H_0 wird abgelehnt, obwohl H_0 richtig ist)$\leq \alpha$ ist.
α heißt **Signifkanzniveau**.
Übliche Werte für α sind 0.05 und 0.01.

11.3 Beispiel

(i) Jemand bezweifelt, dass beim Werfen einer bestimmten Münze die beiden Ereignisse W:„Wappen liegt oben“ und Z:„Zahl liegt oben“ gleich wahrscheinlich sind.

(ii) $P(W)$ und $P(Z)$ sind unbekannt.
Nullhypothese H_0: $P(W) = P(Z) = \frac{1}{2}$.
Als Signifikanzniveau α wählen wir $\alpha = 0.05$.

(iii) Für eine Entscheidung soll die Münze 200 mal („unabhängig“) geworfen werden (unabhängige Wiederholung eines Zufallsexperiments). Gezählt wird, wie oft das Wappen oben liegt. X_{200} sei die dadurch erhaltene Zahl.

(iv) Für welche Werte von X_{200} soll die Nullhypothese H_0 abgelehnt werden?
Vorgabe: $\alpha = 0.05$
Ist H_0 richtig, so ist X_{200} eine binomialverteilte Zufallsvariable, d. h.

$$P_{H_0}(X_{200} = k) = \binom{200}{k} \left(\frac{1}{2}\right)^{200}$$

und der Wert X_{200} wird in der Nähe von 100 liegen.

(v) Gesucht ist deshalb ein $c \in \mathbb{N}$, $c \in [0, 200]$, mit der Eigenschaft

$$P_{H_0}(|X_{200} - 100| > c) = P_{H_0}(X < 100 - c) + P_{H_0}(X > 100 + c) \leq \alpha$$

Bei diesem Beispiel handelt es sich um einen zweiseitigen Signifikanztest.

(vi) Berechnung der Wahrscheinlichkeiten:
Es gilt: $n \cdot p(1-p) = 200 \cdot \frac{1}{2} \cdot \frac{1}{2} = 50 \overset{9.17}{>} 9$
d. h. zur Berechnung für die binomialverteilte Zufallsvariable X_{200} kann die Normalverteilung verwendet werden.
Für X_{200} gilt:

$$\mu = n \cdot p = 100$$
$$\sigma = \sqrt{n \cdot p \cdot (1-p)} = \sqrt{50}$$

Es gilt:

$$P_{H_0}(|X_{200} - 100| > c) \leq \alpha \iff 1 - P_{H_0}(|X_{200} - 100| > c) = P_{H_0}(|X_{200} - 100| \leq c) \geq 1 - \alpha$$

$$\begin{aligned} P_{H_0}(|X_{200} - 100| \leq c) &= P_{H_0}(100 - c \leq X_{200} \leq 100 + c) \\ &\overset{(\sun)}{\approx} \Phi\left(\frac{100 + c - \overbrace{100}^{=\mu}}{\underbrace{\sqrt{50}}_{=\sigma}}\right) - \Phi\left(\frac{100 - c - 100}{\sqrt{50}}\right) \\ &= \Phi\left(\frac{c}{\sqrt{50}}\right) \underbrace{- \Phi\left(\frac{-c}{\sqrt{50}}\right)}_{= \Phi\left(\frac{c}{\sqrt{50}}\right) - 1} = 2\Phi\left(\frac{c}{\sqrt{50}}\right) - 1 \end{aligned}$$

($\sun$) Standardisierung; Approximation durch Normalverteilung
Also gilt:

$$\begin{aligned} & P_{H_0}(|X_{200} - 100| > c) \leq \alpha \\ \iff & 2\Phi\left(\frac{c}{\sqrt{50}}\right) - 1 \geq 1 - \alpha \\ \iff & \Phi\left(\frac{c}{\sqrt{50}}\right) \geq 1 - \frac{\alpha}{2} \quad (*) \end{aligned}$$

Gesucht ist das kleineste $c \in [0, 200]$, $c \in \mathbb{N}$, mit der Eigenschaft

$$\Phi\left(\frac{c}{\sqrt{50}}\right) \geq 1 - \frac{\alpha}{2}$$

(vii) Bestimmung von c
Gesucht ist ein c mit

$$(*) \quad \Phi\left(\frac{c}{\sqrt{50}}\right) \approx 1 - \frac{\alpha}{2} = 0.975$$

($\geq$, kleinstes c mit dieser Eigenschaft)

Dazu betrachten wir die Tabelle der $N(0,1)$-Verteilung

$$\Phi(X) = 0.975 \text{ für } X \approx 1.96$$

Mit $(*)$ folgt

$$\frac{c}{\sqrt{50}} \approx 1.96 \quad \text{d. h. } c \approx 13.8$$

Wähle $c = 14$.

(viii) Entscheidungsregel

H_0 wird abgelehnt für

$$X_{200} < 100 - 14 = 86$$

oder

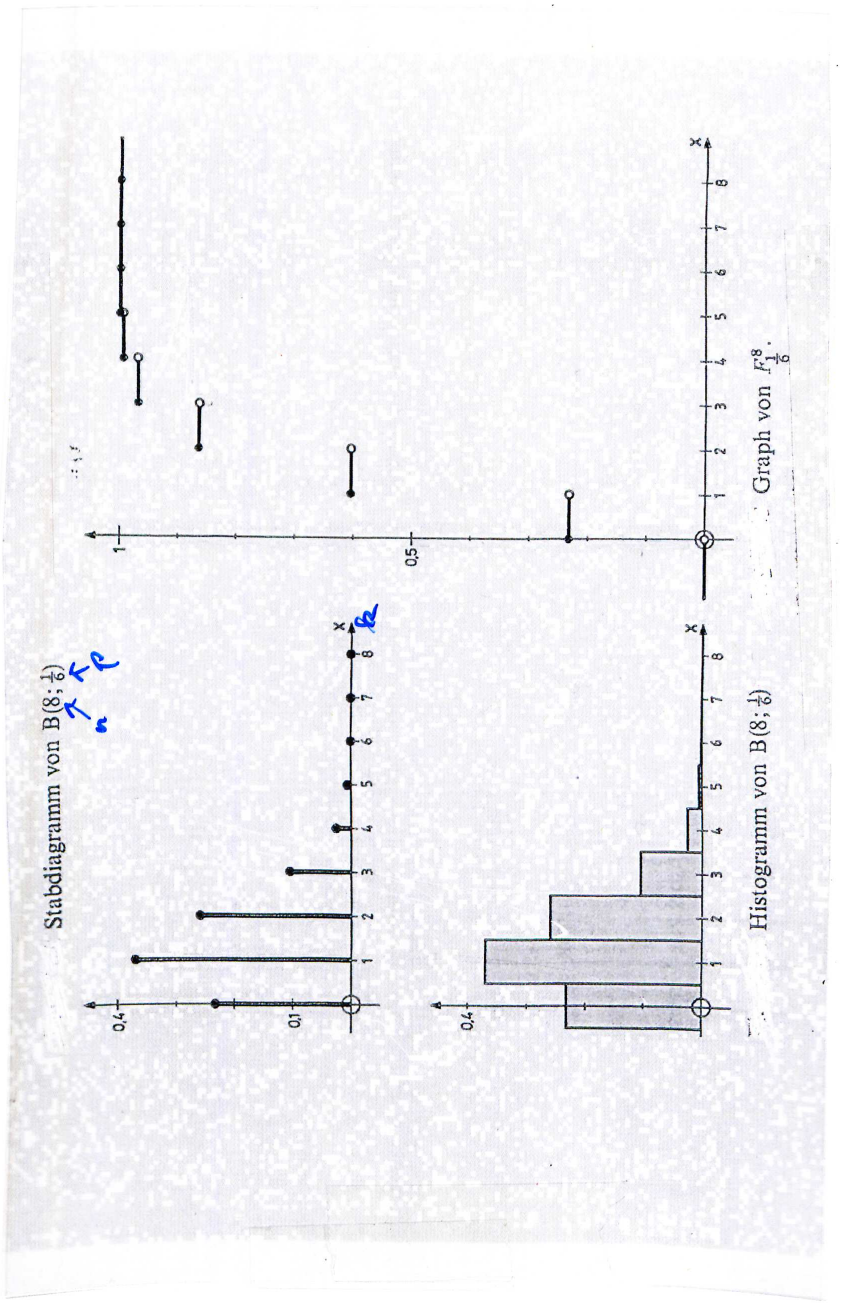
$$X_{200} > 100 + 14 = 114$$

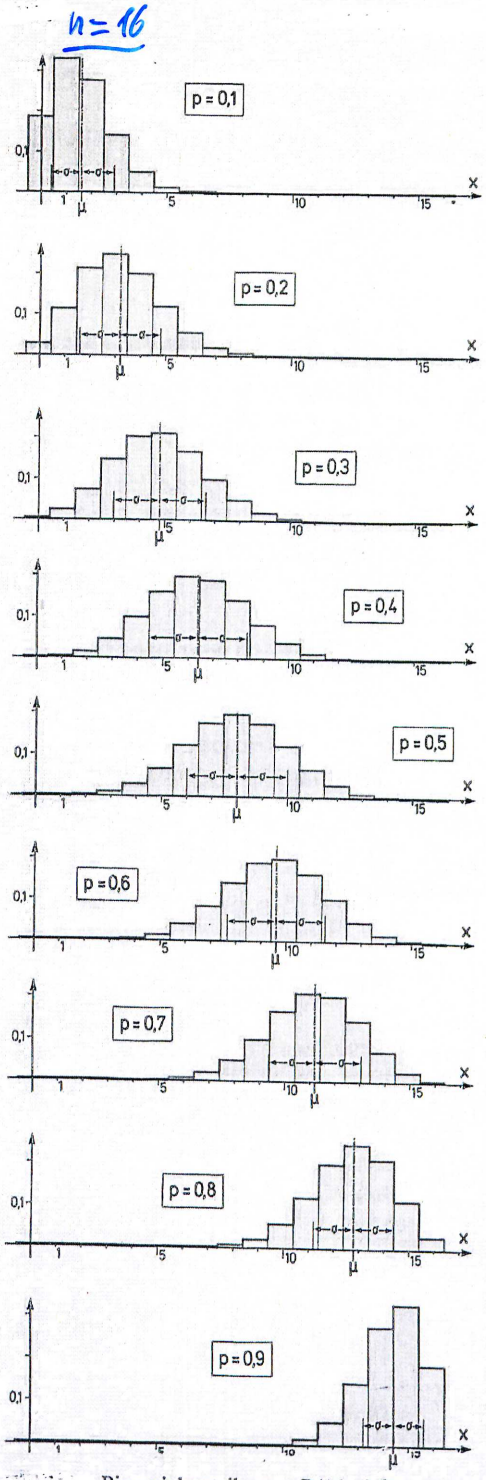
Für $X_{200} \in [86, 114]$ wird H_0 NICHT abgelehnt. Dies heißt NICHT, dass für $X_{200} \in [86, 114]$ H_0 beseitigt ist.

Anhang A Grafiken zur Binomialverteilung

Eesko (LR) S 916

(S. 5) Grafiken zur Binomialverteilung (viele GR(PS))





	$p = 0{,}1$		$p = 0{,}2$		$p = 0{,}3$	
x	B(x)	$F(x)$	B(x)	$F(x)$	B(x)	$F(x)$
0	.185	.185	.028	.028	.003	.003
1	.329	.515	.113	.141	.023	.026
2	.275	.789	.211	.352	.073	.099
3	.142	.932	.246	.598	.146	.246
4	.051	.983	.200	.798	.204	.450
5	.014	.997	.120	.918	.210	.660
6	.003	.999	.055	.973	.165	.825
7	.000	1.000	.020	.993	.101	.926
8			.006	.999	.049	.974
9			.001	1.000	.019	.993
10					.006	.998
11					.001	1.000

	$p = 0{,}4$		$p = 0{,}5$		$p = 0{,}6$	
x	B(x)	$F(x)$	B(x)	$F(x)$	B(x)	$F(x)$
1	.003	.003				
2	.015	.018	.002	.002		
3	.047	.065	.009	.011	.001	.001
4	.101	.167	.028	.038	.004	.005
5	.162	.329	.067	.105	.014	.019
6	.198	.527	.122	.227	.039	.058
7	.189	.716	.175	.402	.084	.142
8	.142	.858	.196	.598	.142	.284
9	.084	.942	.175	.773	.189	.473
10	.039	.981	.122	.895	.198	.671
11	.014	.995	.067	.962	.162	.833
12	.004	.999	.028	.989	.101	.935
13	.001	1.000	.009	.998	.047	.982
14			.002	1.000	.015	.997
15					.003	1.000

	$p = 0{,}7$		$p = 0{,}8$		$p = 0{,}9$	
x	B(x)	$F(x)$	B(x)	$F(x)$	B(x)	$F(x)$
5	.001	.002				
6	.006	.007				
7	.019	.026	.001	.001		
8	.049	.074	.006	.007		
9	.101	.175	.020	.027	.000	.001
10	.165	.340	.055	.082	.003	.003
11	.210	.550	.120	.202	.014	.017
12	.204	.754	.200	.402	.051	.068
13	.146	.901	.246	.648	.142	.211
14	.073	.974	.211	.859	.275	.485
15	.023	.997	.113	.972	.329	.815
16	.003	1.000	.028	1.000	.185	1.000

Binomialverteilungen B(16; p) für verschiedene Parameterwerte p

Anhang B Normalverteilungstabelle

Normalverteilung

x	$\varphi(x)$	$\phi(x)$	x	$\varphi(x)$	$\phi(x)$	x	$\varphi(x)$	$\phi(x)$	x	$\varphi(x)$	$\phi(x)$
0,00	0,39894	0,50000	0,60	0,33322	0,72575	1,20	0,19419	0,88493	1,80	0,07895	0,96407
0,01	0,39892	0,50399	0,61	0,33121	0,72907	1,21	0,19186	0,88686	1,81	0,07754	0,96485
0,02	0,39886	0,50798	0,62	0,32918	0,73237	1,22	0,18954	0,88877	1,82	0,07614	0,96562
0,03	0,39876	0,51197	0,63	0,32713	0,73565	1,23	0,18724	0,89065	1,83	0,07477	0,96638
0,04	0,39862	0,51595	0,64	0,32506	0,73891	1,24	0,18494	0,89251	1,84	0,07341	0,96712
0,05	0,39844	0,51994	0,65	0,32297	0,74215	1,25	0,18265	0,89435	1,85	0,07206	0,96784
0,06	0,39822	0,52392	0,66	0,32086	0,74537	1,26	0,18037	0,89617	1,86	0,07074	0,96856
0,07	0,39797	0,52790	0,67	0,31874	0,74857	1,27	0,17810	0,89796	1,87	0,06943	0,96926
0,08	0,39767	0,53188	0,68	0,31659	0,75175	1,28	0,17585	0,89973	1,88	0,06814	0,96995
0,09	0,39733	0,53586	0,69	0,31443	0,75490	1,29	0,17360	0,90147	1,89	0,06687	0,97062
0,10	0,39695	0,53983	0,70	0,31225	0,75804	1,30	0,17137	0,90320	1,90	0,06562	0,97128
0,11	0,39654	0,54380	0,71	0,31006	0,76115	1,31	0,16915	0,90490	1,91	0,06438	0,97193
0,12	0,39608	0,54776	0,72	0,30785	0,76424	1,32	0,16694	0,90658	1,92	0,06316	0,97257
0,13	0,39559	0,55172	0,73	0,30563	0,76730	1,33	0,16474	0,90824	1,93	0,06195	0,97320
0,14	0,39505	0,55567	0,74	0,30339	0,77035	1,34	0,16256	0,90988	1,94	0,06077	0,97381
0,15	0,39448	0,55962	0,75	0,30114	0,77337	1,35	0,16038	0,91149	1,95	0,05959	0,97441
0,16	0,39387	0,56356	0,76	0,29887	0,77637	1,36	0,15822	0,91309	1,96	0,05844	0,97500
0,17	0,39322	0,56749	0,77	0,29659	0,77935	1,37	0,15608	0,91466	1,97	0,05730	0,97558
0,18	0,39253	0,57142	0,78	0,29431	0,78230	1,38	0,15395	0,91621	1,98	0,05618	0,97615
0,19	0,39181	0,57535	0,79	0,29200	0,78524	1,39	0,15183	0,91774	1,99	0,05508	0,97670
0,20	0,39104	0,57926	0,80	0,28969	0,78814	1,40	0,14973	0,91924	2,00	0,05399	0,97725
0,21	0,39024	0,58317	0,81	0,28737	0,79103	1,41	0,14764	0,92073	2,01	0,05292	0,97778
0,22	0,38940	0,58706	0,82	0,28504	0,79389	1,42	0,14556	0,92220	2,02	0,05186	0,97831
0,23	0,38853	0,59095	0,83	0,28269	0,79673	1,43	0,14350	0,92364	2,03	0,05082	0,97882
0,24	0,38762	0,59483	0,84	0,28034	0,79955	1,44	0,14146	0,92507	2,04	0,04980	0,97932
0,25	0,38667	0,59871	0,85	0,27798	0,80234	1,45	0,13943	0,92647	2,05	0,04879	0,97982
0,26	0,38568	0,60257	0,86	0,27562	0,80511	1,46	0,13742	0,92785	2,06	0,04780	0,98030
0,27	0,38466	0,60642	0,87	0,27324	0,80785	1,47	0,13542	0,92922	2,07	0,04682	0,98077
0,28	0,38361	0,61026	0,88	0,27086	0,81057	1,48	0,13344	0,93056	2,08	0,04586	0,98124
0,29	0,38251	0,61409	0,89	0,26848	0,81327	1,49	0,13147	0,93189	2,09	0,04491	0,98169
0,30	0,38139	0,61791	0,90	0,26609	0,81594	1,50	0,12952	0,93319	2,10	0,04398	0,98214
0,31	0,38023	0,62172	0,91	0,26369	0,81859	1,51	0,12758	0,93448	2,11	0,04307	0,98257
0,32	0,37903	0,62552	0,92	0,26129	0,82121	1,52	0,12566	0,93574	2,12	0,04217	0,98300
0,33	0,37780	0,62930	0,93	0,25888	0,82381	1,53	0,12376	0,93699	2,13	0,04128	0,98341
0,34	0,37654	0,63307	0,94	0,25647	0,82639	1,54	0,12188	0,93822	2,14	0,04041	0,98382
0,35	0,37524	0,63683	0,95	0,25406	0,82894	1,55	0,12001	0,93943	2,15	0,03955	0,98422
0,36	0,37391	0,64058	0,96	0,25164	0,83147	1,56	0,11816	0,94062	2,16	0,03871	0,98461
0,37	0,37255	0,64431	0,97	0,24923	0,83398	1,57	0,11632	0,94179	2,17	0,03788	0,98500
0,38	0,37115	0,64803	0,98	0,24681	0,83646	1,58	0,11450	0,94295	2,18	0,03706	0,98537
0,39	0,36973	0,65173	0,99	0,24439	0,83891	1,59	0,11270	0,94408	2,19	0,03626	0,98574
0,40	0,36827	0,65542	1,00	0,24197	0,84134	1,60	0,11092	0,94520	2,20	0,03547	0,98610
0,41	0,36678	0,65910	1,01	0,23955	0,84375	1,61	0,10915	0,94630	2,21	0,03470	0,98645
0,42	0,36526	0,66276	1,02	0,23713	0,84614	1,62	0,10741	0,94738	2,22	0,03394	0,98679
0,43	0,36371	0,66640	1,03	0,23471	0,84849	1,63	0,10567	0,94845	2,23	0,03319	0,98713
0,44	0,36213	0,67003	1,04	0,23230	0,85083	1,64	0,10396	0,94950	2,24	0,03246	0,98745
0,45	0,36053	0,67364	1,05	0,22988	0,85314	1,65	0,10226	0,95053	2,25	0,03174	0,98778
0,46	0,35889	0,67724	1,06	0,22747	0,85543	1,66	0,10059	0,95154	2,26	0,03103	0,98809
0,47	0,35723	0,68082	1,07	0,22506	0,85769	1,67	0,09893	0,95254	2,27	0,03034	0,98840
0,48	0,35553	0,68439	1,08	0,22265	0,85993	1,68	0,09728	0,95352	2,28	0,02965	0,98870
0,49	0,35381	0,68793	1,09	0,22025	0,86214	1,69	0,09566	0,95449	2,29	0,02898	0,98899
0,50	0,35207	0,69146	1,10	0,21785	0,86433	1,70	0,09405	0,95543	2,30	0,02833	0,98928
0,51	0,35029	0,69497	1,11	0,21546	0,86650	1,71	0,09246	0,95637	2,31	0,02768	0,98956
0,52	0,34849	0,69847	1,12	0,21307	0,86864	1,72	0,09089	0,95728	2,32	0,02705	0,98983
0,53	0,34667	0,70194	1,13	0,21069	0,87076	1,73	0,08933	0,95818	2,33	0,02643	0,99010
0,54	0,34482	0,70540	1,14	0,20831	0,87286	1,74	0,08780	0,95907	2,34	0,02582	0,99036
0,55	0,34294	0,70884	1,15	0,20594	0,87493	1,75	0,08628	0,95994	2,35	0,02522	0,99061
0,56	0,34105	0,71226	1,16	0,20357	0,87698	1,76	0,08478	0,96080	2,36	0,02463	0,99086
0,57	0,33912	0,71566	1,17	0,20121	0,87900	1,77	0,08329	0,96164	2,37	0,02406	0,99111
0,58	0,33718	0,71904	1,18	0,19886	0,88100	1,78	0,08183	0,96246	2,38	0,02349	0,99134
0,59	0,33521	0,72240	1,19	0,19652	0,88298	1,79	0,08038	0,96327	2,39	0,02294	0,99158

Stichwortverzeichnis

A

α-Quantil 14
arithmetisches Mittel 6
Ausgleichsgerade 26
Ausprägung 1
Ausprägungen 37

B

Binomialverteilung 49

C

Covarianz 58

D

Dichtefunktion 54

E

Elementarereignis 33
Ereignis 33
Ergebnismenge 33
Ergebnisse 33
Erwartungswert 56
Exponentialfunktion 29

G

Gleichverteilung 37

H

Häufigkeit
- absolute 3
- kumulierte relative 11
- relative 4

I

Intervallskalen 2

K

Kontingenz 31
Kontingenzkoeffizient 32
Kontingenztafel 31
Korrelationskoeffizient 58
- Pearsonscher 24

Kovarianz
- empirische 24

L

Laplace-Verteilung 37

M

Median
- empirischer 12

Merkmal 1
Merkmale
- dichotone 3

Methode der kleinsten Quadrate 27
Mittel
- geometrisches 17
- gewichtetes arithmetisches 20
- harmonisches 17

Modalwert 20

N

Nullhypothese 61

P

Potenzfunktion 30
Punktwolken 22

Q

Quartil
- oberes 14
- unteres 14

Quartilabstand 14

R

Perzentil
- r-tes 14

Regression
lineare 26
Regressionsgerade 26
Regressionskoeffizient 28
Regressionskonstante 28

S

Schätzwert
erwartungstreuer 10
Signifikanztest 61
Signifkanzniveau 61
Standardabweichung 9, 57
Standardisierung 58
Stichproben-Varianz 9
Streuungsmaßzahlen 9

U

unabhängig
stochastisch 47

V

Variable
abhängige 26
unabhängige 26
Varianz 57
Verhältnisskalen 2
Verteilung
linkssteile 17
rechtssteile 17
symmetrische 17
Verteilungsfunktion
empirische 11

W

Wahrscheinlichkeit
bedingte 46
Wertepaar
zentrales 12

Z

Zentralwert 12
Zufallsgröße 37
Zufallsvariable 37
hypergeometrisch verteilt 51
normalverteilt 53
standard-normalverteilt 53
stetig verteilt 51