

Korrekturblatt zum Lehrbuch S&S - 1. und 2. Auflage

aktueller Stand: 19.01.2024

Korrekturen der 2. Auflage: S. 2-6

Korrekturen der 1. Auflage: S. 7-37

Für Hinweise auf Fehler, grammatikalische Unzulänglichkeiten und sonstige Unstimmigkeiten sind wir immer sehr dankbar. Vergewissern Sie sich bitte zunächst an dieser Stelle, ob ein Fehler bereits „gemeldet“ wurde, bevor Sie uns einen Hinweis schicken. Hinweise können Sie uns mündlich oder schriftlich zukommen lassen. In letzterem Fall führen wir in der Regel die Namen der Hinweisgeber auf einer „Dankesliste“ auf, sofern diese nicht widersprechen. Unsere Mailadressen lauten: stocker@uni-mannheim.de oder steinke@uni-mannheim.de.

Dank geht an...

Avci, Umut
Bartelsen, Tim
Hartwig, Sophie
Hirzel, Simon
Holste, Lion
Lange, Konstantin
Nitschke, José Soler
Schäfer, Henning
Schichtel, Joschka
Scholz, Arthur
Strifler, Sina
Tschirley, Pascale
Veit, Jonas
Verch, Jonas
Wiedemer, Zacharias

Korrekturblatt zur 2. Auflage (2022)

S. 127

ab. Analog würde man bei den Männern 12 Raucher erwarten. Die Anzahl der Raucher weicht hier um 2.4 nach oben hin ab. Geschlecht und Raucherstatus sind auf jeden Fall abhängig.

9.6

S. 155

gung. Je näher die Korrelation bei 0 liegt, desto weniger ist eine lineare Richtungstendenz auszumachen.

Sofern $\tilde{s}_X^2 > 0$ und $\tilde{s}_Y^2 > 0$ sind, gilt $r_{XY} = 0$ gerade dann, wenn $\tilde{s}_{XY} = 0$ ist.

\tilde{s}_Y^2

S. 286

• Beispiel S1-a fortgesetzt • Die Dichtefunktion von Beispiel S1-a ist gegeben durch

$$f_X(x) = 1.2I_{(0,0.2)}(x) + 0.5I_{(0.2,1]}(x) + 0.15I_{(1,2]}(x).$$

ersetzen durch 0.5

S. 294

Somit implizieren die Werte

$$f_{X|Y}(0|0) = 0.25 \quad \text{und} \quad f_{X|Y}(1|0) = 0.25$$

0.75

S. 301

$$f_{X|Y}(x|y) = \frac{f(x, y)}{f_Y(y)} = \frac{(0.5x + 0.25y)I_{[0,1]}(x)I_{[0,2]}(y)}{(0.25y + 0.25)I_{[0,2]}(y)} = \frac{2x + y}{y + 1}I_{[0,1]}(x),$$

für $y \in [0, 2]$. Für $y \in [0, 1]$ ist die bedingte Dichte nicht definiert. Die bedingte Dichte

[0, 2]

S. 313

• Beispiel D3-a • Tabelle 7.1.6 zeigt ein Beispiel einer dreidimensionalen diskreten Verteilung. Die Realisationsmöglichkeiten von Y sind 0, 1 und 2, die von X und Z jeweils 0 und 1. Gemäß Lesart der Tabelle gilt dann beispielsweise

$$P(X = 0, Y = 0, Z = 0) = 0.04, \quad P(X = 0, Y = 0, Z = 1) = 0.07$$

oder $P(X = 1, Y = 2, Z = 1) = 0.04$.

0.00

S. 315

bestimmt. Beispielsweise gilt dann für $Z = 0$:

$$\begin{aligned} f_{XY|Z}(0, 0|0) &= \frac{0.04}{0.9} \approx 0.04, & f_{XY|Z}(0, 1|0) &= \frac{0.07}{0.9} \approx 0.08, \\ f_{XY|Z}(0, 2|0) &= \frac{0.26}{0.9} \approx 0.29, & f_{XY|Z}(0, 0|1) &= \frac{0.11}{0.9} \approx 0.12, \\ f_{XY|Z}(0, 1|0) &= \frac{0.10}{0.9} \approx 0.11, & f_{XY|Z}(0, 2|0) &= \frac{0.32}{0.9} \approx 0.36. \end{aligned}$$

links unten: 1, 1 | 0

rechts oben: 1, 0 | 0

rechts unten: 1, 2 | 0

S. 360

$f_{Y|X}(y|x) = I_{[0,1]}(y)$ für $x \in [0, 1]$ bzw. $f_{X|Y}(x|y) = I_{[0,1]}(x)$ für $y \in [0, 1]$.

Außerdem sind X und Y identisch verteilt. Wie bereits ermittelt, gilt

$$E(X) = E(Y) = 0.5 \quad \text{und} \quad \text{Var}(X) = \text{Var}(Y) = 1/12.$$

$f_{X|Y}(x|y)$ (ein X im Index zu viel)

S. 361

Der einfacheren Notation wegen setzen wir $U := E(Y|X)$ und erhalten schließlich

$$P(U = 1.4) = 0.6 \quad \text{und} \quad P(U = 1.6) = 0.4,$$

$P(U = 1.4) = 0.6$

S. 399

$$\begin{aligned} P(X \leq t + s | X > t) &= \frac{P(t < X \leq t + s)}{P(X > t)} = \frac{P(X \leq t + s) - P(X \leq t)}{1 - P(X \leq t)} \\ &= \frac{1 - e^{-\lambda(t+s)} - (1 - e^{-\lambda t})}{1 - (1 - e^{-\lambda t})} = \frac{-e^{-\lambda t} e^{-\lambda s} + e^{-\lambda t}}{e^{-\lambda t}} \\ &= \frac{-e^{-\lambda t}(1 - e^{-\lambda s})}{e^{-\lambda t}} = 1 - e^{-\lambda s} = F_X(s). \end{aligned}$$

plus (Minuszeichen ist falsch)

S. 466

schen X und Y). Für den Ausdruck (8.3.4) erhalten wir

$$E(XYZ|Z = 0) = 0 \cdot E(XY|Z = 0) = 0 \quad \text{bzw.}$$

$$E(XYZ|Z = 1) = 1 \cdot E(XY|Z = 0) = 1.$$

S. 470

• **Beispiel 8.3.4** • Sei $(X, Y)^T$ *zweidimensional* (bivariat) *normalverteilt*, d.h.

$$\begin{pmatrix} X \\ Y \end{pmatrix} \sim N_2\left(\begin{pmatrix} \mu_X \\ \mu_Y \end{pmatrix}, \begin{pmatrix} \sigma_X^2 & \sigma_{XY} \\ \sigma_{XY} & \sigma_Y^2 \end{pmatrix}\right).$$

Gemäß (8.3.8) und (8.3.9) sind dann X_1 und X_2 jeweils normalverteilt mit

$$X \sim N(\mu_X, \sigma_X^2) \quad \text{bzw.} \quad Y \sim N(\mu_Y, \sigma_Y^2).$$

Aus (8.3.6) lässt sich beispielsweise die Verteilung der gewichteten Summe $c_0 + c_1X + c_2Y$ ableiten. Mit entsprechender Wahl von

$$c_0 = c_0 \quad \text{und} \quad \mathbf{C} = (c_1 \quad c_2)$$

erhält man nach einigen Vereinfachungsschritten

$$c_0 + c_1X + c_2Y \sim N(c_0 + c_1\mu_X + c_2\mu_Y, c_1^2\sigma_X^2 + c_2^2\sigma_Y^2 + 2c_1c_2\sigma_{XY}). \quad (8.3.14)$$

X und Y

c_2^2

S. 576

$$H_0 : \pi = 0.5 \quad \text{vs.} \quad H_1 : \pi \neq 0.5.$$

Falls

$$S_{30} > 15 + \sqrt{7.5} \cdot z_{0.95}, \quad (11.1.4)$$

>

S. 605

Denn um mit dem Wert 19 gerade noch verwerfen zu können, müsste der kritische Wert gleich 20 sein. Analog erhält man für 15 korrekte Antworten als *p*-Wert

$$P(S_{30} < 16) = P(S_{30} \leq 15) \approx 0.571.$$

0.572

S. 675

hergeleitet werden kann. Entsprechend würde dann gelten:

$$\beta_1 = \sigma_{XY}/\sigma_X^2.$$

Für die Eigenschaft

$$\sigma_U^2 \approx \sigma_U^2(1 - \rho_{XY}^2)$$

σ_Y^2

S. 677

$$\begin{aligned}\tilde{S}_{XY} &= \frac{1}{n} \sum_{i=1}^n (X_i - \bar{X})(Y_i - \bar{Y}) \\ &= \frac{1}{n} \sum_{i=1}^n (X_i - \bar{X})(\beta_0 + \beta_1 X_i + U_i - \beta_0 - \beta_1 \bar{X} - \bar{U}) \\ &= \beta_1 \frac{1}{n} \sum_{i=1}^n (X_i - \bar{X})^2 + \frac{1}{n} \sum_{i=1}^n (X_i - \bar{X})(U_i - \bar{U}) \\ &= \beta_1 \tilde{S}_X^2 + \frac{1}{n} \sum_{i=1}^n (X_i - \bar{X}) U_i \quad (+) \quad \bar{U} \frac{1}{n} \sum_{i=1}^n (X_i - \bar{X}) \\ &= \beta_1 \tilde{S}_X^2 + \frac{1}{n} \sum_{i=1}^n (X_i - \bar{X}) U_i,\end{aligned}$$

Minuszeichen

Korrekturblatt zur 1. Auflage (2017)

wie folgt...

S. 30

Notation und Bezeichnungen bei Urlisten

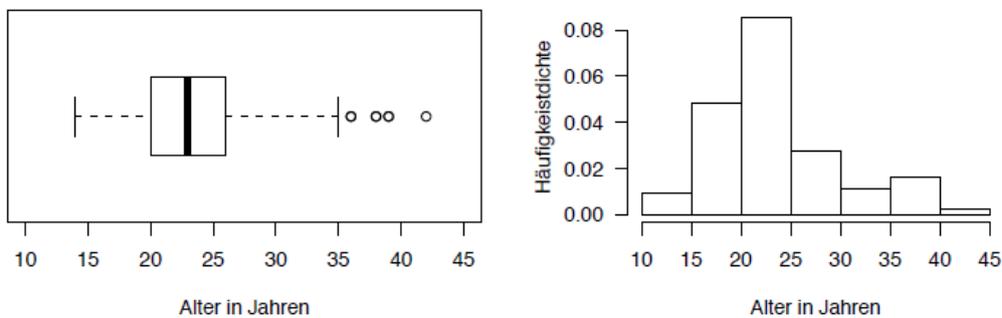
n_j	Absolute Häufigkeit der j -ten Ausprägung
f_j	Relative Häufigkeit der j -ten Ausprägung
x_i für $i = 1, 2, \dots, n$	i heißt Indexvariable, $\{1, 2, \dots, n\}$ heißt Indexmenge
n_1, n_2, \dots, n_k	Absolute Häufigkeitsverteilung
f_1, f_2, \dots, f_{n_k}	Relative Häufigkeitsverteilung

gehört hier nicht rein

f_k

S. 39

Abb. 3.2.8: Boxplot und Histogramm –
Alter von 87 Frauen bei der Geburt des ersten Kindes



Daten: R-Paket „UsingR“ (Datensatz „Baby“)

Es handelt sich um den Datensatz „firstchi“

S. 46

Tabelle 3.3.2: Häufigkeitsverteilung bei Klassenbildung mit kumulierten Anteilen

j	Klasse $(c_{j-1}, c_j]$	\tilde{n}_j	\tilde{f}_j	\tilde{f}_j/d_j	$\tilde{F}_n(c_j)$
1	$(0,10]$	10	0.067	0.0067	0.067
2	$(10,20]$	10	0.267	0.0267	0.333

d_j

S. 71

$$d^* \approx \frac{1}{n} \sum_{j=1}^k \tilde{n}_j |m_j - \bar{x}| = \sum_{j=1}^k \tilde{f}_j |m_i - \bar{x}|$$

für die mittlere absolute Abweichung vom arithmetischen Mittel und

$$d \approx \frac{1}{n} \sum_{j=1}^k \tilde{n}_j |m_j - \tilde{x}_{0.5}| = \sum_{j=1}^k \tilde{f}_j |m_i - \tilde{x}_{0.5}|$$

Summen laufen jeweils bis k, j statt i

S. 72

$$d^* \approx \frac{1}{n} \sum_{i=1}^n \tilde{n}_j |m_j - \bar{x}| = \sum_{i=1}^n \tilde{f}_j |m_i - \bar{x}| \quad \text{bzw.}$$

$$d \approx \frac{1}{n} \sum_{i=1}^n \tilde{n}_j |m_j - \tilde{x}_{0.5}| = \sum_{i=1}^n \tilde{f}_j |m_i - \tilde{x}_{0.5}|,$$

Summen laufen jeweils bis k, j statt i

S. 78

$$\bar{x} = \sum_{i=1}^k \tilde{f}_j \bar{x}_j = \frac{3}{12} \cdot 1.0 + \frac{4}{12} \cdot 2.0 + \frac{5}{12} \cdot 1.46 = 1.525.$$

Sofern die Urliste bekannt ist, kann das arithmetische Mittel natürlich auch direkt durch Mittelung der 12 Werte bestimmt werden. Die durchgezogene horizontale Linie markiert die Lage des Gesamtmittelwerts. Die gestrichelten Linien kennzeichnen die Lagen der Gruppenmittelwerte. Für die interne Varianz ergibt sich

$$\sum_{i=1}^k \tilde{f}_j \tilde{s}_j^2 = \frac{3}{12} \cdot 0.06 + \frac{4}{12} \cdot 0.05 + \frac{5}{12} \cdot 0.0464 = 0.051$$

j statt i

S. 80**Quantilsabstände**

Für metrische Beobachtungswerte ist der α -Quantilsabstand definiert als

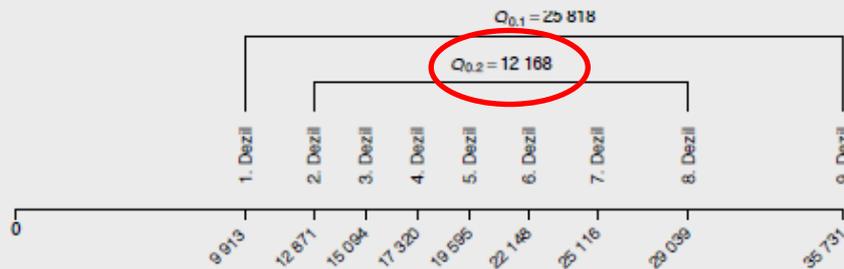
$$Q_\alpha = \tilde{x}_{1-\alpha} - \tilde{x}_\alpha \quad \text{für } \alpha \in (0, 0.5).$$

Er entspricht der Spannweite der mittleren $100 \times (1 - \alpha)\%$ aller Werte. Im Speziellen heißt $Q_{0.25}$ auch Interquartilsabstand.

$100 \times (1 - 2\alpha)\%$

S. 80

Abb. 4.5.2: Quantilsabstände des Nettoäquivalenzeinkommens in Euro (2012)



Daten: Statistisches Bundesamt [2014a]

16168

S. 81

dass sich 90% aller Einkommen unterhalb von 36 000 Euro bewegen. Die Spannweite der mittleren 80% aller Einkommen beträgt ca. 26 000 Euro und die der mittleren 60% nur noch etwa 12 000 Euro.

16000

S. 87

Quantilskoeffizient der Schiefe

Gegeben seien metrische Beobachtungswerte x_1, \dots, x_n . Dann ist durch

$$Q_\alpha = \frac{(\tilde{x}_{1-\alpha} - \tilde{x}_{0.5}) - (\tilde{x}_{0.5} - \tilde{x}_\alpha)}{\tilde{x}_{1-\alpha} - \tilde{x}_\alpha}, \text{ für } \alpha \in (0, 0.5),$$

der α -Quantilskoeffizient der Schiefe gegeben.

Speziell wird $QS_{0.25}$ als **Quantilskoeffizient der Schiefe** bezeichnet.

Für $QS_\alpha > 0$ sind die mittleren $(1 - 2\alpha) \times 100\%$ der Werte rechtsschief, für $QS_\alpha < 0$ linksschief und für $QS_\alpha = 0$ symmetrisch verteilt.

QS_α

S. 87

Beispiel 4.7.1: Schiefe der Einkommensverteilung

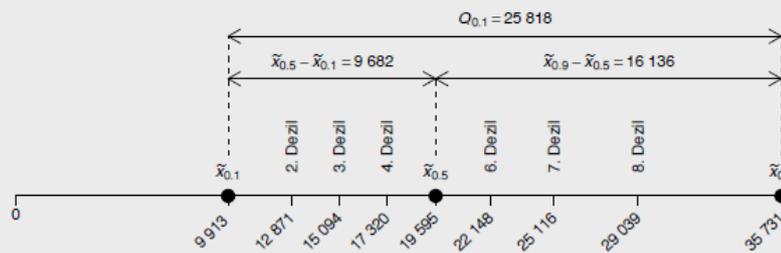
Gegeben sei die gleiche Situation wie in Beispiel 4.5.2 (Abschnitt 4.5.2). Die Einkommensverteilung für Deutschland weist eine deutliche Rechtsschiefe auf. Mit

$$\tilde{x}_{0.1} = 9913, \tilde{x}_{0.5} = 19595 \text{ und } \tilde{x}_{0.9} = 35731$$

ergibt sich als 0.1-Quantilkoeffizient der Schiefe

$$Q_{0.1} = \frac{(35731 - 19595) - (19595 - 9913)}{35731 - 9913} = \frac{16136 - 9682}{25818} \approx 0.25.$$

Abb. 4.7.2: 0.1-Quantilkoeffizient der Schiefe für die Einkommensverteilung



Daten: Statistisches Bundesamt [2014]

Analog erhält man

$$Q_{0.2} = \frac{(29039 - 19595) - (19595 - 12871)}{29039 - 12871} \approx 0.17 \text{ und}$$

QS_{0.1}

QS_{0.2}

S. 88

$$Q_{0.3} = \frac{(25116 - 19595) - (19595 - 15094)}{25116 - 15094} \approx 0.10.$$

Mit abnehmendem α -Wert nimmt die Rechtsschiefe zu. Dies ist darauf zurückzuführen, dass die Verteilung am oberen Rand „ausfranst“ und der Einkommensunterschied umso stärker ausfällt, je mehr man sich den beiden Rändern der Verteilung annähert.

QS_{0.3}

S. 94

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

+2

S. 110

$$f_{11}^{Y|X} = \frac{n_{11}}{n_{1\bullet}} = \frac{8}{40} = 0.20$$

4/40 = 0.10

S. 113

$$f_{1j}^{Y|X} = f_{1\bullet}, f_{2j}^{Y|X} = f_{2\bullet}, \dots, f_{kj}^{Y|X} = f_{k\bullet} \quad \text{für } j = 1, \dots, l.$$

umgekehrt, also: X|Y

S. 115

tatsächliche Häufigkeit um -2.4 nach unten hin „von der Unabhängigkeit“ ab. Analog würde man bei den Männern 12 Raucher erwarten. Die Anzahl der Raucher weicht hier um 2.4 nach oben hin ab. Geschlecht und Raucherstatus sind auf jeden Fall abhängig.

9.6

S. 118

$$\chi^2 = \frac{36(4 \cdot 12 - 12 \cdot 8)^2}{12 \cdot 24 \cdot 16 \cdot 20} = 0.9.$$

besser: 16×20×12×24 (Reihenfolge korrespondierend zu Formel)

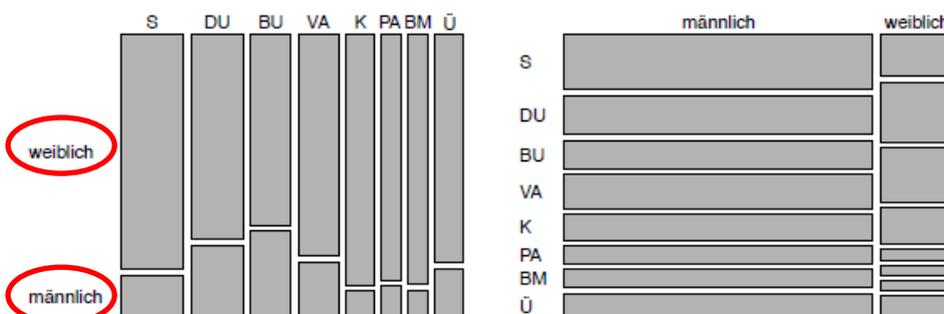
S. 126

Dadurch entsprechen die Flächen der Säulen den absoluten Beträgen der Abweichungen (5.1.9). Die Intention des Assoziationsplots besteht darin, die Unterschiede der bedingten Verteilungen herauszustellen.

sind die Flächen der Säulen proportional zu den ...

S. 127

Abb. 5.1.7: Mosaikplots für gemeinsame Verteilungen



Daten: Statistisches Bundesamt [2014c]

vertauscht

S. 140

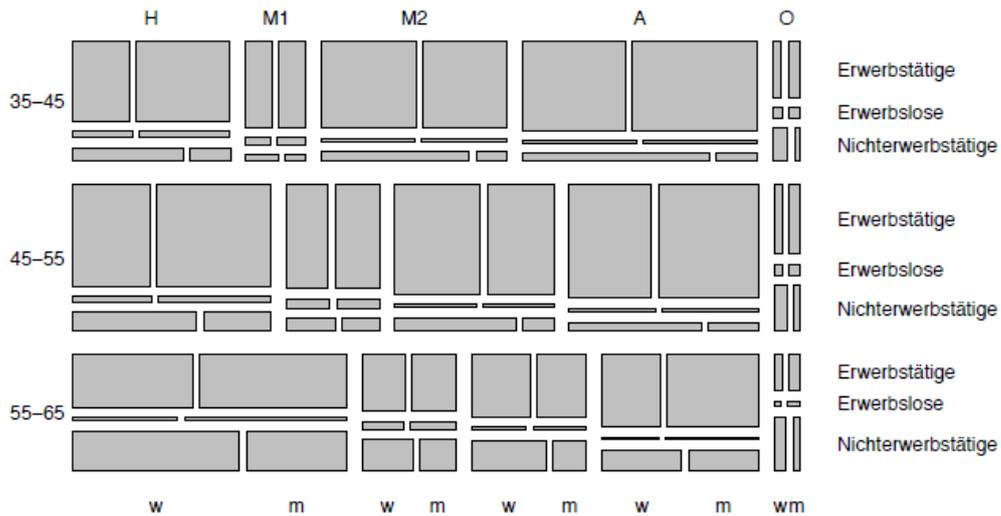
die Korrelation bei 0 liegt, desto weniger ist eine lineare Richtungstendenz auszumachen.

Sofern $\tilde{s}_X^2 > 0$ und $\tilde{s}_Y^2 > 0$ sind, gilt $r_{XY} = 0$ gerade dann, wenn $\tilde{s}_{XY} = 0$ ist.

\tilde{s}_Y^2

S. 175

Abb. 5.3.6: Mosaikplot: Alter, Schulbildung, Erwerbsstatus und Geschlecht

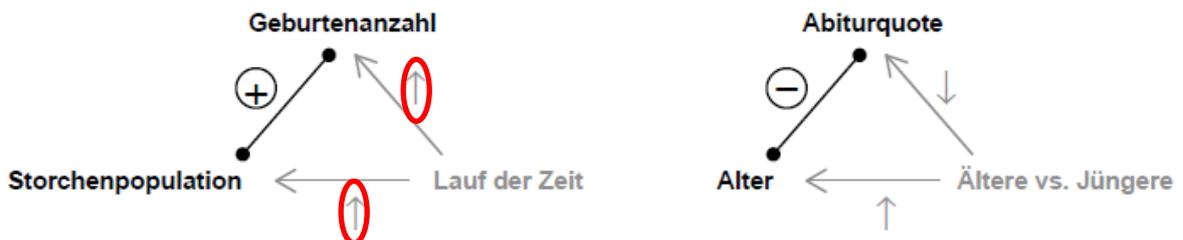


Quelle: Statistisches Bundesamt [2013b, S. 79–80]

Daten

S. 192

Abb. 5.3.17: Schematische Beispiele von Scheinabhängigkeiten



Pfeile müssen nach unten zeigen

S. 196

Tabelle 5.3.4: Erwerbstätigkeit, Alter und Schulabschluss erwerbsfähiger Personen

Erwerbstätigenquote nach Alter

35-45	84.8%
45-55	84.0%
55-65	61.2%

Erwerbstätigenquote nach Schulabschluss

ohne Abitur	73.7%
mit Abitur	86.3%

Anteil von Abiturienten nach Alter

35-45	84.8%
45-55	84.0%
55-65	61.2%

Erwerbstätigenquote nach Alter und Schulabschluss

	35-45	45-55	55-65
ohne Abitur	82.4%	81.5%	57.0%
mit Abitur	89.4%	90.4%	75.1%

84.9%

84.1%

61.3%

81.6%

35.4%

28.6%

23.5%

Wie man erkennen kann, nimmt die Erwerbstätigkeit erst ab einem Alter von ca. 55 Jahren schlagartig ab. Der totale Effekt beim Übergang von der Gruppe der 45- bis 55-Jährigen in die Gruppe der 55- bis 65-Jährigen liegt bei ~~fast~~ -22.8% . Der tota-

S. 197

Eine differenzierte Untersuchung in homogeneren Untergruppen liefert ein zwispältiges Resultat. Beschränkt man sich nur auf die beiden älteren Altersgruppen, so stellt man fest, dass der direkte negative Alterseffekt unter den Personen ohne Abitur deutlich stärker ist (-24.5%) als unter den Personen mit Abitur (-15.3%) und sogar den totalen Effekt von -22.8% übersteigt. Umgekehrt ist in der Gruppe der 45- bis 55-Jährigen der direkte Abitureffekt mit $+8.9\%$ deutlich schwächer als in der Gruppe der 55- bis 65-Jährigen mit $+18.1\%$. Dies deutet darauf hin, dass die Effektstärke

-24.6%

$+8.8\%$

S. 210

• Vereinigungsmenge • Die *Vereinigungsmenge* zweier Mengen ist die Menge aller Elemente, die ~~entweder~~ in der einen oder in der anderen Menge enthalten sind. Das „oder“ ist hierbei als „und/oder“ zu verstehen. Dies bedeutet, dass jedes Element der

S. 228

$$P(A \cup C|B) = \frac{P((A \cap C) \cap B)}{P(B)} = \frac{P(\{2, 3, 6\})}{P(B)} = \frac{3/6}{3/6} = 1.$$

∪ „vereinigt“

S. 231

Insgesamt erhalten wir damit

$$\begin{aligned}
P(A_1 \cap A_3) &= P(A_1 \cap A_2 \cap A_3) + P(A_1 \cap \bar{A}_2 \cap A_3) \\
&= 0.504 + P(A_1)P(\bar{A}_2|A_1)[1 - P(\bar{A}_3|A_1 \cap A_2)] \\
&= 0.504 + 0.9 \cdot (1 - 0.7) \cdot 0.4 = 0.504 + 0.108 = 0.612.
\end{aligned}$$

$P(A_3|A_1 \cap \bar{A}_2)$ (hier unnötig über das Gegenereignis formuliert)

S. 259

Damit wird zum Ausdruck gebracht, dass nicht mehr nur Teilmengen von Wahrscheinlichkeiten zugeordnet werden können, sondern sich der Additionskalkül auf beliebige Teilmengen von \mathbb{R} fortsetzen lässt. Im stetigen Fall müsste man sich aufgrund des mit Ω

S. 257

- Beispiel S1-a fortgesetzt • Die Dichtefunktion von Beispiel S1-a ist gegeben durch

$$f_X(x) = 1.2I_{(0,0.2)}(x) + 0.5I_{(0.2,1]}(x) + 0.15I_{(1,2]}(x).$$

Dann lautet die zu Abbildung 7.1.5 (links) gehörende Verteilungsfunktion

0.5

S. 260

für abzählbare Werte $a_1, a_2, \dots, a_N, \dots$ gilt:

$$(i) f(a_i) > 0 \text{ für alle } i \text{ und } (ii) \sum_{i=1}^{\infty} f(a_i) = 1.$$

Damit lässt sich jede solche Funktion natürlich auch als Wahrscheinlichkeitsfunktion (irgendeiner) diskreten Zufallsvariable interpretieren.

Analog definiert jede Funktion $f : \mathbb{R} \rightarrow [0, 1]$ eindeutig ein *stetiges Wahrscheinlichkeitsmaß* auf \mathbb{R} , falls gilt:

$$(i) f(x) \geq 0 \text{ und } (ii) \int_{-\infty}^{\infty} f(x) dx = 1.$$

$[0, \infty)$

\geq

S. 261

oder auch

$$P(X_1 \in A | X_2 \in B) \text{ und } P(X_1 = x_1 | X_2 = x_2, X_3 = x_3).$$

$A_1 A_2$

S. 263

$X \setminus Y$	b_1	b_2	\dots	b_j	\dots	b_l	$P(Y = y)$
a_1	p_{11}	p_{12}	\dots	p_{1j}	\dots	p_{1l}	$p_{1\bullet}$
a_2	p_{21}	p_{22}	\dots	p_{2j}	\dots	p_{2l}	$p_{2\bullet}$
\vdots	\vdots	\vdots	\ddots			\vdots	\vdots
a_i	p_{i1}	p_{i2}	\dots	p_{ij}	\dots	p_{il}	$p_{i\bullet}$
\vdots	\vdots	\vdots		\ddots		\vdots	\vdots
a_k	p_{k1}	p_{k2}	\dots	p_{kj}	\dots	p_{kl}	$p_{k\bullet}$
$P(X = x)$	$p_{\bullet 1}$	$p_{\bullet 2}$	\dots	$p_{\bullet j}$	\dots	$p_{\bullet l}$	1

vertauscht: links unten: $Y = y$, rechts oben: $X = x$

S. 263

$$f_Y(y) = P(Y = y) = \sum_j f_{XY}(a_i, y).$$

i

S. 264

Somit implizieren die Werte

$$f_{X|Y}(0|0) = 0.25 \text{ und } f_{X|Y}(1|0) = 0.25$$

0.75

S. 271

Die bedingte Dichte von X gegeben $Y = y$ lautet

$$f_{X|Y}(x|y) = \frac{f(x, y)}{f_Y(y)} = \frac{(0.5x + 0.25y)I_{[0,1]}(x)I_{[0,2]}(y)}{(0.25y + 0.25)I_{[0,2]}(y)} = \frac{2x + y}{y + 1}I_{[0,1]}(x),$$

für $y \in [0, 2]$. Für $y \notin [0, 1]$ ist die bedingte Dichte nicht definiert. Die bedingte Dichte

[0, 2]

S. 274

und für $y > 2$ und $x \in [0, 1]$ entsprechend

$$F_{XY}(x, y) = P(X \leq x, Y \leq 2) = 0.5x^2 + 0.5x.$$

2

S. 282

• **Beispiel D3-a** • Tabelle 7.1.6 zeigt ein Beispiel einer dreidimensionalen diskreten Verteilung. Die Realisationsmöglichkeiten von Y sind 0, 1 und 2, die von X und Z jeweils 0 und 1. Gemäß Lesart der Tabelle gilt dann beispielsweise

$$P(X = 0, Y = 0, Z = 0) = 0.04, \quad P(X = 0, Y = 0, Z = 1) = 0.07$$

oder $P(X = 1, Y = 2, Z = 1) = 0.04$.

0.00

Tabelle 7.1.6: Gemeinsame Verteilung von Beispiel D3-a

$Z = 0$				
$X \setminus Y$	0	1	2	Summe
0	0.04	0.07	0.26	0.37
1	0.11	0.10	0.32	0.63
Summe	0.04	0.08	0.28	0.90
$Z = 1$				
$X \setminus Y$	0	1	2	Summe
0	0.00	0.01	0.02	0.03
1	0.01	0.02	0.04	0.07
Summe	0.01	0.03	0.06	0.10

0.15, 0.17, 0.58 und 0.53

S. 283

Tabelle 7.1.8: Zweidimensionale Randverteilungen von Beispiel D3-a

$X \setminus Z$	0	1	$P(Z = z)$	$Z \setminus Y$	0	1	2	$P(Z = z)$
0	0.37	0.03	0.40	0	0.15	0.17	0.58	0.40
1	0.53	0.07	0.60	1	0.01	0.03	0.06	0.60
$P(Y = y)$	0.90	0.10	1.00	$P(Y = y)$	0.16	0.20	0.64	1.00

links unten: $Z = z$, rechts oben: $X = x$

0.9
0.1

S. 284

bestimmt. Beispielsweise gilt dann für $Z = 0$:

$$f_{XY|Z}(0, 0|0) = \frac{0.04}{0.9} \approx 0.04, \quad f_{XY|Z}(0, 1|0) = \frac{0.07}{0.9} \approx 0.08,$$

$$f_{XY|Z}(0, 2|0) = \frac{0.26}{0.9} \approx 0.29, \quad f_{XY|Z}(0, 0|1) = \frac{0.11}{0.9} \approx 0.12,$$

$$f_{XY|Z}(0, 1|0) = \frac{0.10}{0.9} \approx 0.11, \quad f_{XY|Z}(0, 2|0) = \frac{0.32}{0.9} \approx 0.36.$$

links unten: 1, 1 | 0

rechts oben: 1, 0 | 0

rechts unten: 1, 2 | 0

S. 286

für $x \in [0, 1]$ und $y \in (0, 2]$ bestimmt. Beispielsweise gilt dann

$$f_{Z|XY}(z|0, 2) = 2zI_{[0,1]}(z) \text{ oder } f_{Z|XY}(z|1, 1) = \left(\frac{2}{3}z + \frac{2}{3}\right)I_{[0,1]}(z).$$

[0, 2] (links abgeschlossen)

$$\begin{aligned} f_{XY|Z}(x, y|z) &= \frac{f_{XYZ}(x, y, z)}{f_Z(z)} = \frac{0.5x + 0.5yz}{z + 0.5} I_{[0,1]}(x) I_{[0,2]}(y) \\ &= \frac{x + yz}{2z + 1} I_{[0,1]}(x) I_{[0,2]}(y) \end{aligned}$$

für $z \in [0, 1]$ und $y \in [0, 2]$ bestimmt. Abbildung 7.1.14 zeigt dann beispielsweise die **entfällt**

S. 288

- Beispiele: Funktionen einer Zufallsvariable •

0-1-Variable:

Betrachten wir als einfaches Einstiegsbeispiel zunächst eine sog. Bernoulli-Variable X , die lediglich die Werte 0 und 1 annehmen kann. Dabei gelte

$$P(X = 0) = 0.9 \text{ und } P(X = 1) = 0.1.$$

Dann wäre $U = X^2$ objekttechnisch eine Zufallsvariable, welche gerade die quadrierten Realisationen von X angibt. Für $X = 0$ gilt dann also $U = 0^2 = 0$ und für $U = 1$ gilt

X = 1

S. 289

$$P(U = 1) = P(U = 4) = \dots = P(U = 36) = 1/36.$$

1/6

S. 295

Tabelle 7.1.12: Zweimaliges Würfeln – Verteilung der Summe Augenzahlen

u	1	2	3	4	5	6	8	9	10
$P(U = u)$	1/36	2/36	2/36	3/36	2/36	4/36	2/36	1/36	2/36
u	12	15	16	18	20	24	25	30	36
$P(U = u)$	4/36	2/36	1/36	2/36	2/36	2/36	1/36	2/36	1/36

des Produkts

S. 304

$$(i^*) E(Y) = \sum_j b_j P(Y = b_j) = \sum_j b_j f_Y(a_j)$$

wobei $b_1, b_2, \dots, b_l, \dots$ die Realisationsmöglichkeiten von Y sind bzw.

 b_j **S. 306**

$$E(U) = \sum_{j=1}^6 b_j^2 P(U = b_j)$$

 b_j

$$E(U) = 1 + 2E(X) = 8.$$

Z

~~Nach den Varianten (i) und (ii)~~ werden die Erwartungswerte von U und Z jeweils über die Dichte von X berechnet. Daraus ergibt sich dann

Gemäß Variante (ii) werden...

S. 307

~~Die Varianten (i*) und (ii*)~~ verwenden dagegen die Dichten von U bzw. Z , woraus sich **Gemäß Variante (ii*)...**

bzw.

$$\begin{aligned} E(Z) &= \int_{-\infty}^{\infty} z f_Z(z) dz = \int_{-\infty}^{\infty} 0.5z I_{[1,3]}(z) dz \\ &= 0.5 \int_1^3 z dz = 0.5 [0.5z^2]_1^3 = 2 \end{aligned}$$

Grenzen lauten: 1 und 3

S. 316

$$\sigma^2 = \sqrt{0.2294} \approx 0.48.$$

 σ **S. 319**

Man beachte, dass durch

$$F_X^{-1}(\alpha) = -20 \ln(1 - \alpha) \text{ für } \alpha \in (0, 1)$$

Y**S. 321**

$$Var(Y|X=0) = E(Y^2|X=0) - E(Y|X=0)^2 = 3 - 1.6^2 = 0.44$$

$E(Y|X=0)^2$

S. 323

• Beispiel S2-b fortgesetzt • Dem gleichnamigen Beispiel aus Abschnitt 7.1.2 liegt die gemeinsame stetige Verteilung zweier Zufallsvariablen X und Y gemäß Dichte

$$f_{XY}(x, y) = I_{[0,1]}(x)I_{[0,1]}(y)$$

zugrunde. Da X und Y unabhängig sind, stimmen die bedingten Verteilungen mit den jeweiligen Randverteilungen überein, d.h. es gilt

$$f_{Y|X}(y|x) = I_{[0,1]}(y) \text{ für } x \in [0, 1] \text{ bzw. } f_{X|Y}(x|y) = I_{[0,1]}(x) \text{ für } y \in [0, 1].$$

$f_{X|Y}(x|y)$ (ein X im Index zu viel)

S. 324

Der einfacheren Notation wegen setzen wir $U := E(Y|X)$ und erhalten schließlich

$$P(U = 1.4) = 0.6 \text{ und } P(U = 1.6) = 0.4,$$

$P(U = 1.4) = 0.6$

S. 332

$$Var(X) = E(X^2) - (E(X))^2 = \frac{7}{12} - \left(\frac{5}{12}\right)^2 = \frac{59}{144}$$

5, 7, 11/144

$$\rho_{XY} = \frac{Cov(X, Y)}{\sqrt{Var(X)Var(Y)}} = \frac{-1/72}{\sqrt{(59/144) \cdot (11/36)}} = -\frac{1}{\sqrt{649}} \approx -0.0393$$

$-1/11 \approx -0.0909$

S. 334

• Beispiel S2-a fortgesetzt • Setzen wir das vorhergehende gleichnamige Beispiel mit den stetigen Zufallsvariablen X und Y fort, erhalten wir für die Koeffizienten

$$\tilde{\beta}_1 = \frac{Cov(X, Y)}{Var(X)} = \frac{-1/72}{59/144} = -\frac{2}{59} \approx -0.0339 \text{ und}$$

$$\tilde{\beta}_0 = E(Y) - \tilde{\beta}_1 E(X) = \frac{7}{6} - \frac{2}{59} \cdot \frac{7}{12} \approx 1.1469$$

Die resultierende theoretische Geradengleichung lautet dann

$$\tilde{y}(x) = 1.1469 - 0.0339x.$$

11/144, $-2/11 \approx -0.1818$,

$-2/11$, 1.0606

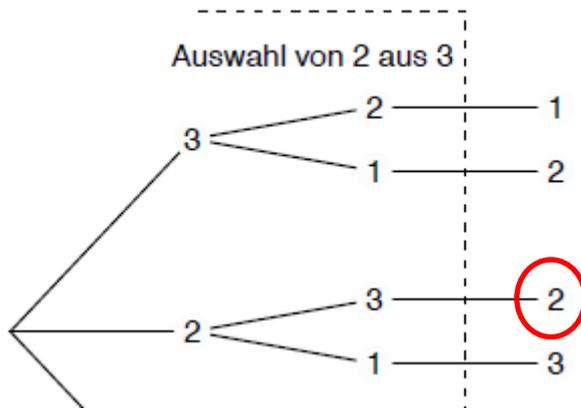
1.0606 – 0.1818x

S. 337

$$(9) \text{ } Corr(c_X X, c_Y Y) = Corr(X, Y) \text{ für alle } c_X, c_Y \neq 0.$$

>

S. 342



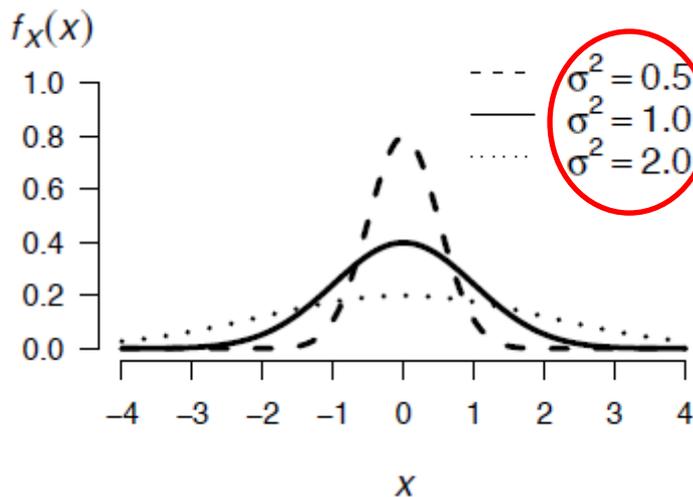
1

S. 359

$$\begin{aligned}
 P(X \leq t + s | X > t) &= \frac{P(t < X \leq t + s)}{P(X > t)} = \frac{P(X \leq t + s) - P(X \leq t)}{1 - P(X \leq t)} \\
 &= \frac{1 - e^{-\lambda(t+s)} - (1 - e^{-\lambda t})}{1 - (1 - e^{-\lambda t})} = \frac{-e^{-\lambda t}e^{-\lambda s} + e^{-\lambda t}}{e^{-\lambda t}} \\
 &= \frac{-e^{-\lambda t}(1 - e^{-\lambda s})}{e^{-\lambda t}} = 1 - e^{-\lambda s} = F_X(s).
 \end{aligned}$$

plus (Minuszeichen ist falsch)

S. 363



$\sigma = 0.5$

$\sigma = 1.0$

$\sigma = 2.0$

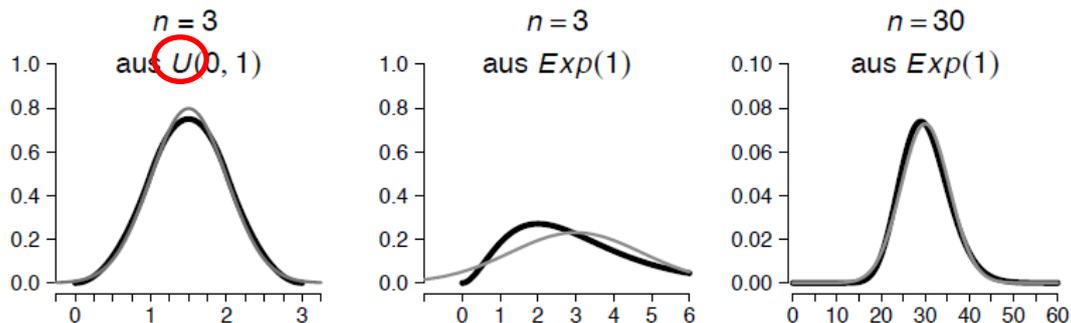
S. 375

$$f_{\bar{X}_2}(x) = 2f_{S_2}(2x) = 2 \cdot [2xI_{[0,1]}(2x) + (2 - 2x)I_{(1,2]}(2x)] \\ = 4xI_{[0,0.5]}(s) + (4 - 4x)I_{(0.5,1]}(x).$$

x

S. 388

Abb. 7.4.11: Approximationsgüte des ZGWS bei verschiedenen stetigen Verteilungen



G

S. 389

Mit einer Wahrscheinlichkeit von 95% sollte die Anzahl mitreisender Kinder bei 50 gebuchten Reisen unter diesen Annahmen also irgendwo zwischen 43 und 67 liegen.

90%

S. 396

$$P(X = 18) = \binom{20}{18} \cdot \left(\frac{1}{6}\right)^3 \cdot \left(\frac{1}{6}\right)^{18} \approx 0.0330.$$

5

S. 398

Der Erwartungswert von X ist aufgrund der Symmetrie der Dichtefunktion bzgl. 0 für $n > 1$ gleich 0. Allerdings existiert er für $n \leq 1$ nicht. Die Varianz ist für $n \leq 2$ unendlich und für $n > 2$ gegeben durch

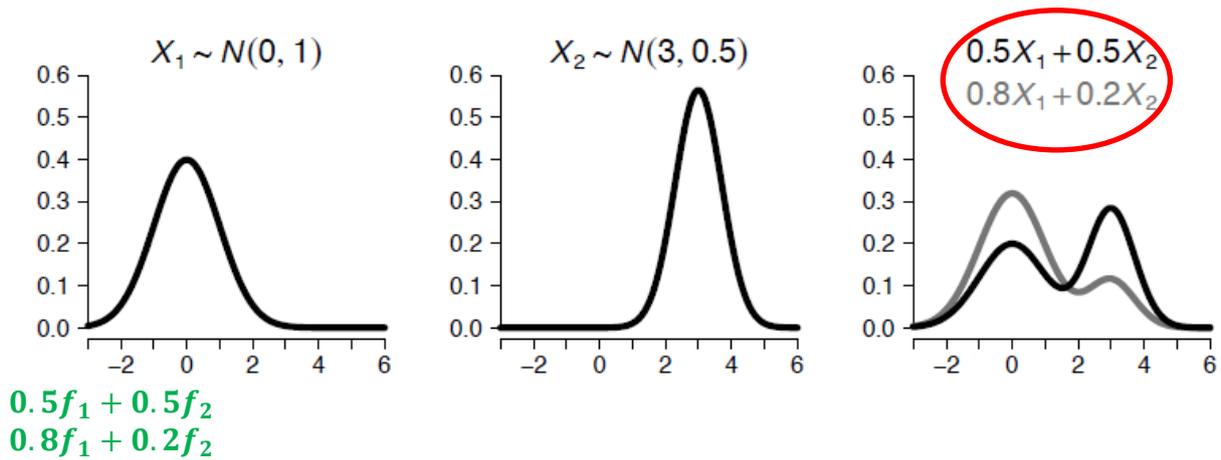
Allerdings existiert er für $n = 1$ nicht. Die Varianz ist für $n < 3$ unendlich und für $n \geq 3$ gegeben durch

S. 400

Während (8.1.13) dem Resultat (7.4.22) bei einer Normalverteilung sehr ähnlich ist, weicht (8.1.14) fundamental von den bisherigen Resultaten für stochastische Mittel ab. Resultat (8.1.14) besagt, dass der Mittelwert identisch verteilt ist wie die ~~Verteilung~~ der Ausgangsvariablen selbst. Die Verteilung zieht sich also nicht über dem Lagewert

S. 401

Abb. 8.1.3: Mixturverteilung aus zwei Normalverteilungen



S. 404

$X_1 \sim U(0, 1)$, $X_2 \sim U(0, 2)$, $X_3 \sim U(0, 3)$, $X_4 \sim U(0, 4)$, $X_5 \sim U(0, 5), \dots$
G

S. 413

Tabelle 8.3.1: Gemeinsame Verteilung von $(X, Y, Z)^T$ von Beispiel D3-b

Z = 0				
X \ Y	0	1	2	Summe
0	0.06	0.03	0.00	0.09
1	0.15	0.21	0.15	0.41
Summe	0.21	0.24	0.15	0.60
Z = 1				
X \ Y	0	1	2	Summe
0	0.04	0.02	0.00	0.06
1	0.10	0.14	0.10	0.34
Summe	0.14	0.16	0.10	0.40

0.51

S. 414

Tabelle 8.3.2: Randverteilung von $(X, Y)^T$ von Beispiel D3-b

$Z = 0$				
$X \setminus Y$	0	1	2	$P(X = x)$
0	0.10	0.05	0.00	0.15
1	0.25	0.35	0.25	0.85
$P(Y = y)$	0.35	0.40	0.25	1.00

S. 415

Tabelle 8.3.4: Beispiel einer 3-dimensionalen Verteilung mit bedingter Unabhängigkeit

$Z = 0$				$Z = 1$			
$X \setminus Y$	0	1	Summe	$X \setminus Y$	0	1	Summe
0	0.05	0.05	0.10	0	0.05	0.20	0.25
1	0.20	0.20	0.40	0	0.05	0.20	0.25
Summe	0.25	0.25	0.50	Summe	0.10	0.40	0.50

1

S. 418

schen X und Y). Für den Ausdruck (8.3.4) erhalten wir

$$E(XYZ|Z = 0) = 0 \cdot E(XY|Z = 0) = 0 \text{ bzw.}$$

$$E(XYZ|Z = 1) = 1 \cdot E(XY|Z = 0) = 1.$$

Hier verzichten wir wiederum auf eine inhaltliche Deutung und beachten nur die Berechnungsweise.

1

S. 422

$$Y|X = x \sim N_q(\mu_Y, \Sigma_{YY}) \text{ bzw. } X|Y = \mathbf{x} \sim N_p(\mu_X, \Sigma_{XX}).$$

y

S. 423

• Beispiel 8.3.4 • Sei $(X, Y)^T$ zweidimensional (bivariat) normalverteilt, d.h.

$$\begin{pmatrix} X \\ Y \end{pmatrix} \sim N_2\left(\begin{pmatrix} \mu_X \\ \mu_Y \end{pmatrix}, \begin{pmatrix} \sigma_X^2 & \sigma_{XY} \\ \sigma_{XY} & \sigma_Y^2 \end{pmatrix}\right).$$

Gemäß (8.3.8) und (8.3.9) sind dann X_1 und X_2 jeweils normalverteilt mit

$$X \sim N(\mu_X, \sigma_X^2) \text{ bzw. } Y \sim N(\mu_Y, \sigma_Y^2).$$

Aus (8.3.6) lässt sich beispielsweise die Verteilung der gewichteten Summe $c_0 + c_1X + c_2Y$ ableiten. Mit entsprechender Wahl von

$$c_0 = c_0 \text{ und } C = (c_1 \quad c_2)$$

erhält man nach einigen Vereinfachungsschritten

$$c_0 + c_1X + c_2Y \sim N(c_0 + c_1\mu_X + c_2\mu_Y, c_1^2\sigma_X^2 + c_2^2\sigma_Y^2 + 2c_1c_2\sigma_{XY}). \quad (8.3.14)$$

X und Y

$$c_2^2$$

S. 429

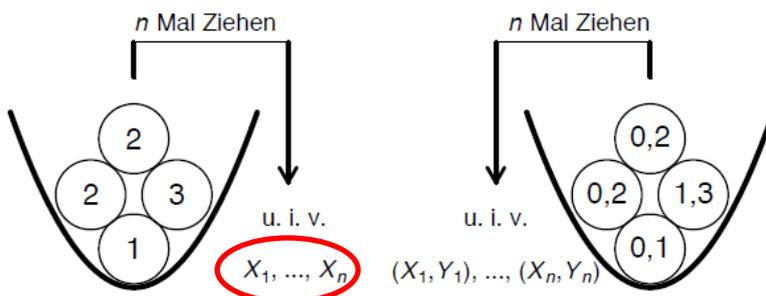
$$P(S_{50} \geq 31) = \sum_{i=31}^{50} \binom{50}{i} 0.5^{30} \approx 0.0595 \text{ oder}$$

$$P(S_{50} \geq 36) = \sum_{i=36}^{50} \binom{50}{i} 0.5^{30} \approx 0.0013.$$

50

S. 434

Abb. 9.2.1: Illustration realer Stichproben



Y_1, \dots, Y_n

S. 435

Tabelle 9.2.2: Gemeinsame Verteilung von $(X_1, Y_1)^T$ und $(X_2, Y_2)^T$ beim Ziehen „mit Zurücklegen“

Tupel	Wahrschein.	Tupel	Wahrschein.
(0, 1, 0, 1)	1/16	(1, 3, 0, 1)	1/16
(0, 1, 0, 2)	1/8	(0, 2, 1, 3)	1/8
(0, 2, 0, 1)	1/8	(1, 3, 0, 2)	1/8
(0, 1, 1, 3)	1/16	(1, 3, 1, 3)	1/16

(0, 2, 0, 2) fehlt noch mit Wahrscheinlichkeit 1/4

S. 436

Tabelle 9.2.3: Gemeinsame Verteilung von $(X_1, Y_1)^T$ und $(X_2, Y_2)^T$ bei Ziehen „ohne Zurücklegen“

Tupel	Wahrschein.	Tupel	Wahrschein.
(0, 1, 0, 1)	0	(1, 3, 0, 1)	1/12
(0, 1, 0, 2)	1/6	(0, 2, 1, 3)	1/6
(0, 2, 0, 1)	1/6	(1, 3, 0, 2)	1/6
(0, 1, 1, 3)	1/12	(1, 3, 1, 3)	0

(0, 2, 0, 2) fehlt noch mit Wahrscheinlichkeit 1/12

S. 459

$$MSE(\hat{\mu}_C) = MSE\left(\bar{X} + \frac{1}{n}\right) = \text{Var}\left(\bar{X} + \frac{1}{n}\right) + \text{Bias}\left(\bar{X} + \frac{1}{n}\right) = \frac{\sigma^2}{n} + \frac{1}{n^2}.$$

muss noch quadriert werden

S. 480

• **Beispiel 10.2.2: Gepäckabfertigungszeiten** • Wir betrachten zunächst die gleiche Situation wie in Beispiel 10.2.1 mit $n = 10$ Beobachtungswerten. Unter der Annahme einer Normalverteilung und einer bekannten Varianz von $\sigma^2 = 2$ ergab sich als 0.95-Konfidenzintervall für μ

$$[9.52 - 1.96 \cdot \sqrt{2/10}, 9.52 + 1.96 \cdot \sqrt{2/10}] \approx [8.64, 10.40].$$

Ist die Varianz hingegen unbekannt, müssen wir diese durch die korrigierte Stichprobenvarianz schätzen. Im vorliegenden Fall erhalten wir dafür als Schätzwert

$$\hat{\sigma}^2 = s^2 = 1.50$$

Mit $t_{9,0.975} \approx 2.26$ (vgl. Tab. A.2) resultiert daraus mit

$$[9.52 - 2.26 \cdot \sqrt{1.5/10}, 9.52 + 2.26 \cdot \sqrt{1.5/10}] \approx [8.64, 10.40]$$

~~fast genau das gleiche Intervall (gleich bis zur 2. Nachkommastelle). Der kleinere Wert der geschätzten Varianz wird durch das größere Quantil gerade wieder kompensiert. Dennoch beachte man, dass bei einer geschätzten Varianz von genau 2 das Intervall länger wäre. Jedenfalls sollte die langfristige durchschnittliche Abfertigungszeit bei einem Niveau von 95% irgendwo zwischen 8.64 und 10.4 Minuten liegen.~~

Rechenfehler

$\hat{\sigma}^2 = s^2 = 2.24$, Intervall: [8.45, 10.59]

Bei einem Niveau von 95% sollte die langfristige durchschnittliche Abfertigungszeit irgendwo zwischen 8.45 und 10.59 Minuten liegen.

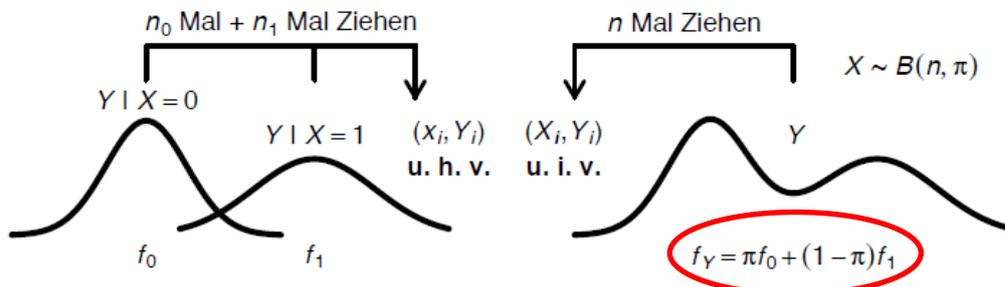
S. 483

$$\bar{Y}_1 - \bar{Y}_0 = \frac{1}{n_1} \sum_{i=1}^{n_1} Y_{1i} - \frac{1}{n_0} \sum_{i=1}^{n_0} Y_{1i} = \frac{1}{n} \left[\sum_{i=1}^{n_1} \frac{n}{n_1} Y_{1i} - \sum_{i=1}^{n_0} \frac{n}{n_0} Y_{0i} \right] \quad (10.2.23)$$

Y_{0i}

S. 500

Abb. 10.2.11: Nichtstochastische vs. stochastische Gruppenumfänge



f_0 und f_1 hier vertauscht: $(1-\pi) f_0 + \pi f_1$

S. 508

$$\hat{\mu}_{ML} = 1/\bar{X}.$$

λ statt $\mu \Rightarrow \hat{\lambda}_{ML}$

S. 519

Betrachten wir die im vorhergehenden Punkt beschriebene Entscheidungsregel zu Beispiel 11.1.1 (Raten vs. Wissen) nochmals genauer. Das Testproblem hierzu lautete

$$H_0 : \pi = 0.5 \text{ vs. } H_1 : \pi \neq 0.5.$$

>

S. 520

$$Z_{20} = \frac{\bar{X}_{20} - 0.8}{\sqrt{0.8 \cdot 0.2/20}} \text{ bzw. } Z_{30} = \frac{\bar{X}_{500} - 0.6}{\sqrt{0.6 \cdot 0.4/500}}.$$

Z₅₀₀**S. 539**

(c) $H_0 : \mu \leq \mu_0$ vs. $H_1 : \mu > \mu_0$ gegeben durch

$$g(\mu) = \Phi\left(-z_{1-\alpha/2} - \frac{\mu - \mu_0}{\sqrt{\sigma^2/n}}\right) + \Phi\left(-z_{1-\alpha/2} + \frac{\mu - \mu_0}{\sqrt{\sigma^2/n}}\right).$$

$H_0: \mu = \mu_0$ vs. $H_1: \mu \neq \mu_0$

S. 545

Denn um mit dem Wert 19 gerade noch verwerfen zu können, müsste der kritische Wert gleich 20 sein. Analog erhält man für 15 korrekte Antworten als p -Wert

$$P(S_{30} < 16) = P(S_{30} \leq 15) \approx 0.571.$$

0.572**S. 549**

• **Beispiel 11.3.2: t -Test über μ** • Gegeben seien folgende 10 Realisationen einer Stichprobe aus einer Normalverteilung:

$$7.8, 10.1, 9.0, 8.0, 11.6, 10.7, 8.1, 8.6, 9.4, 11.9.$$

Hierbei gilt:

$$\bar{x} = 9.52 \text{ und } s^2 = 1.50.$$

2.24

Testen wir nun beispielsweise

$$H_0 : \mu \geq 10 \text{ vs. } H_1 : \mu < 10$$

zum Niveau 5%, erhalten wir als Wert der Teststatistik

$$t = \frac{9.52 - 10}{\sqrt{1.5} \cdot 10} \approx -1.24$$

Der kritische Wert ergibt sich aus dem 0.05-Quantil der $t(9)$ -Verteilung (Tab. A.2):

$$t_{9,0.05} = -t_{9,0.95} \approx -1.83.$$

Wegen $t \approx -1.24 > -1.83$ wird H_0 nicht verworfen. Das arithmetische Mittel ist bei einem Niveau von 5% folglich nicht signifikant kleiner als 10. Der korrespondierende p -Wert beträgt 12.3%. Dieser Wert lässt sich beispielsweise mit statistischer Software wie R exakt berechnen. Alternativ lässt sich der p -Wert auch mithilfe der Quantiltabelle A.2 abschätzen. Da der Wert +1.24 zwischen dem 0.8- und dem 0.9-Quantil liegt, muss der Wert -1.24 aufgrund der Symmetrie der Verteilung zwischen dem 0.1- und dem 0.2-Quantil liegen. Damit liegt der p -Wert irgendwo zwischen 10% und 20%.

2.24, -1.01,
-1.01,
16.9%
+1.01
-1.01

S. 550

Der Wert der Teststatistik bleibt gleich, falls wir alternativ

$$H_0 : \mu = 10 \text{ vs. } H_1 : \mu \neq 10$$

testen. Die kritischen Werte zum Niveau 5% lauten dann

$$t_{9,0.025} \approx -2.26 \text{ und } t_{9,0.975} \approx 2.26.$$

Die Nullhypothese kann hierbei (erst recht) nicht verworfen werden. Der korrespondierende exakte p -Wert beträgt dann das Zweifache des einseitigen Tests von zuvor, also 24.6%.

Man beachte, dass die obigen Daten bereits in Beispiel 10.2.1 (Gepäckabfertigungszeiten) bei der Konstruktion eines Konfidenzintervalls für μ verwendet wurden. Das 0.95-Konfidenzintervall ergab sich dabei als

$$[\bar{x} - t_{9,0.975} \sqrt{s^2/n}, \bar{x} + t_{9,0.975} \sqrt{s^2/n}] \approx [8.64, 10.40]$$

Dabei ist der Wert 10 im Intervall enthalten, was mit dem bestehenden Zusammenhang zwischen Intervallschätzung und zweiseitigem Testen konform ist (Abschnitt 11.1.3).

33.9%
[8.45, 10.59]

S. 553

zu verwenden. Allerdings genügt T_D , wie bereits in Abschnitt 10.2.3 ausgeführt, keiner einfach zu spezifizierenden Verteilung. Lediglich für großes n_0 und n_1 schafft hier der ZGWS wieder Abhilfe (siehe späterer Punkt). Eine einfache Lösung gibt es nur, falls man die Zusatzannahme einer *homoskedastischen Varianz*

T_Ddach (Dach darüber!)

S. 555

Formel für die „gepoolte“ Varianzschätzung sollte im Kasten ergänzt werden:

$$S_p^2 = \frac{1}{n_0 + n_1 - 2} (n_0 \tilde{S}_0^2 + n_1 \tilde{S}_1^2) = \frac{1}{n_0 + n_1 - 2} ((n_0 - 1)S_0^2 + (n_1 - 1)S_1^2)$$

S. 557

$$\chi_A^2 = \sum_{i=1}^k \frac{(N_i - n\pi_i)^2}{n\pi_i} \stackrel{a}{\sim} \chi^2(k-1). \quad (11.3.1)$$

k

S. 559

$$\begin{aligned} \chi_A^2 &= \sum_{i=1}^k \frac{(n_i - n\pi_i)^2}{n\pi_i} \\ &\approx \frac{(548 - 568.1)^2}{568.1} + \frac{(370 - 351.8)^2}{351.8} + \dots + \frac{(82 - 86.2)^2}{86.2} \approx 5.10. \end{aligned}$$

k

S. 561

Damit wird H_0 zum Niveau 5% klar verworfen. Gemäß Tabelle A.3 ist der p -Wert sogar **noch kleiner als 1%**. Die empirische Verteilung für die einzelnen Größenklassen weicht

knapp über

S. 567

Die statistische Modellierung erfolgt analog zum vorhergehenden Beispiel, wobei X mit $k = 2$ die Würfelfarbe angibt (1=weiß, 2=schwarz) und Y mit $l = 3$ die geworfene Zahl. Mit $(k-1) \times (l-1) = 5$ genügt die Teststatistik unter H_0 approximativ einer

6

S. 568

Würfel Zahl	1	2	3	4	5	6	$n_{i\bullet}$
Weiß	9 (9) 0	6 (11) -5	14 (12) 2	10 (8) 2	8 (7.5) 0.5	13 (12.5) 0.5	60
Schwarz	9 (9) 0	16 (11) -5	10 (12) 2	6 (8) 2	7 (7.5) 0.5	12 (12.5) 0.5	60
$n_{\bullet j}$	18	22	24	16	15	25	120

Minuszeichen fehlen davor: -2, -2, -0.5, -0.5

S. 569

$$\chi^2 = \frac{n(N_{11}N_{22} - N_{12}N_{21})^2}{N_{\bullet 1}N_{\bullet 2}N_{1\bullet}N_{2\bullet}}$$

muss quadriert werden

$$\chi_{links}^2 = \frac{200(45 \cdot 62 - 45 \cdot 48)^2}{88 \cdot 107 \cdot 85 \cdot 110} \approx 0.2268.$$

40, 0.2326

S. 571

• **Approximativer Binomialtest** • Binomialtests lassen sich in großen Stichproben auch als approximative Gauß-Tests sowohl in einer nichtstandardisierten als auch in einer standardisierten Variante durchführen. Grundlage hierfür bilden die Verteilungsergebnisse (11.2.7) und (11.2.8) bzw. (11.2.9). Dies wurde anhand der Einführungsbeispiele in Abschnitt 11.1.1 bereits ausführlich dargestellt. Im Folgenden sei nur die standardisierte Variante auf Basis von Resultat (11.2.9) d.h. des Stichprobenmittels, zusammengefasst.

(11.1.8) und (11.1.9) bzw. (11.1.10) (11.1.10)

S. 576

$$\tilde{T}_D^{links} = \frac{0.0142}{\sqrt{\frac{\hat{\pi}_0(1-\hat{\pi}_0)}{85} + \frac{\hat{\pi}_1(1-\hat{\pi}_1)}{110}}} \approx 0.1975 \quad \text{und}$$

$$\tilde{T}_D^{rechts} = \frac{0.0142}{\sqrt{\frac{\hat{\pi}_0(1-\hat{\pi}_0)}{8500} + \frac{\hat{\pi}_1(1-\hat{\pi}_1)}{11000}}} \approx 1.9755.$$

$$\hat{t}_D^{links} = \dots$$

$$\hat{t}_D^{rechts} = \dots$$

S. 592

$$\frac{1}{n} \sum_{i=1}^n \hat{y}_i \hat{u}_i - \hat{y} \cdot \hat{u} = \frac{1}{n} \sum_{i=1}^n \hat{y}_i \hat{u}_i,$$

$$\hat{y}_i$$

S. 593

Aufsummieren über alle Beobachtungswerte ergibt

$$\sum_{i=1}^n (y_i - \bar{y})^2 = \sum_{i=1}^n (\hat{y}_i - \bar{y})^2 + \sum_{i=1}^n \hat{u}_i^2 + 2 \sum_{i=1}^n (\hat{y}_i - \bar{y}) \hat{u}_i. \quad (12.1.11)$$

^2 (muss ins Quadrat)

S. 600

$$E(Y_i | X_i = x_i) = E(\beta_0 + \beta_1 X_i + U_i) = \beta_0 + \beta_1 x_i + E(U_i | X_i = x_i),$$

$E[\beta_0 + \beta_1 X_i + U_i | X_i = x_i]$

S. 605

• **Hintergrund** • Nach der Definition der *theoretischen KQ-Regression* und der Beschreibung ihrer Eigenschaften wird gezeigt, dass die zu schätzende *theoretische Regressionsgerade* als *theoretische KQ-Gerade* interpretierbar ist. Diese Tatsache erweist sich später etwa beim Nachweis von Konsistenzeigenschaften und anderen theoretischen Betrachtungen als nützlich.

Überblick**S. 606**

das theoretische Pendant zu Formel (12.1.19). Man könnte diesen Kennwert als *theoretische Fehlervarianz der KQ-Regression* oder alternativ als theoretische Varianz zu Formel (12.1.19) bzw. (12.1.17)

S. 607

Es sei bemerkt, dass die für den empirischen Fall der KQ-Regression geltenden 8 Eigenschaften gemäß Satz 12.1.1 auch für die theoretische KQ-Regression in analoger Weise

S. 609

$$\frac{1}{n} \sum_{i=1}^n E(X_i^2) - \left(\frac{1}{n} \sum_{i=1}^n x_i \right)^2 \xrightarrow{n \rightarrow \infty} c_2. \quad (K2^*)$$

Gleichheitszeichen dazwischen „=“

hergeleitet werden kann. Entsprechend würde dann gelten:

$$\beta_1 = \sigma_{XY} / \sigma_X^2.$$

Für die Eigenschaft

$$\sigma_U^2 = \sigma_U^2 (1 - \rho_{XY}^2)$$

σ_Y^2

S. 611

$$\begin{aligned}
 &= \frac{1}{n} \sum_{i=1}^n (X_i - \bar{X})(\beta_0 + \beta_1 X_i + U_i - \beta_0 - \beta_1 \bar{X} + \bar{U}) \\
 &= \beta_1 \frac{1}{n} \sum_{i=1}^n (X_i - \bar{X})^2 + \frac{1}{n} \sum_{i=1}^n (X_i - \bar{X})(U_i - \bar{U}) \\
 &= \beta_1 \tilde{S}_X^2 + \frac{1}{n} \sum_{i=1}^n (X_i - \bar{X})U_i + \bar{U} \frac{1}{n} \sum_{i=1}^n (X_i - \bar{X}) \\
 &= \beta_1 \tilde{S}_X^2 + \frac{1}{n} \sum_{i=1}^n (X_i - \bar{X})U_i,
 \end{aligned}$$

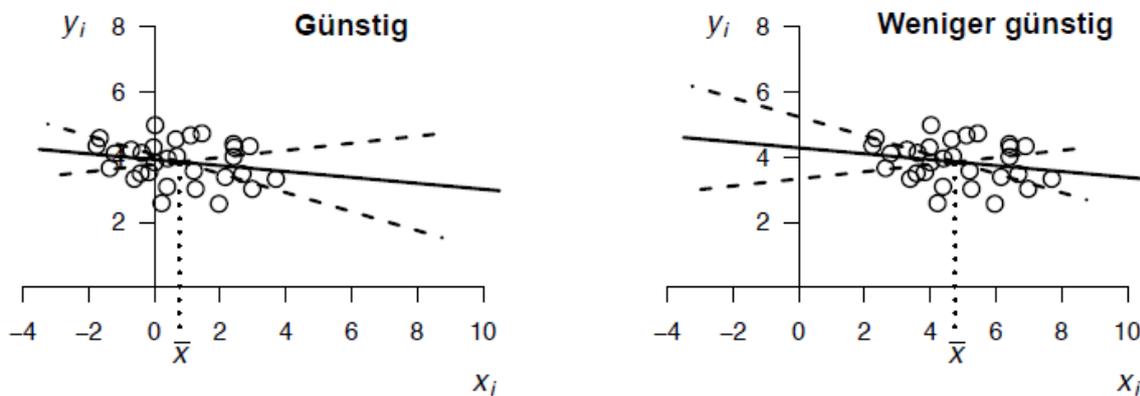
$-\bar{U}$

S. 613

$$\text{Var}(\hat{\beta}_0) = \frac{1}{n} \sum_{i=1}^n x_i^2 \text{Var}(\beta_1^2) \quad (12.1.58)$$

$\hat{\beta}_1$

Abb. 12.1.9: Schätzung des Steigungskoeffizienten



Achsenabschnitts

S. 614

ähnlich sieht sich jedoch heraus, dass ausgangsl. bereits die Bestimmung der bedingten Verteilungen genügt. Dabei wird auf alle X_i -Variablen bedingt. Somit erhält man dann

$$C_i U_i | X_1 = x_1, \dots, X_n = x_n \sim N\left(0, \frac{(X_i - \bar{X})^2 \sigma_U^2}{n \tilde{S}_X^2}\right) \text{ für } \tilde{S}_X^2 > 0.$$

Zur Abkürzung können wir hier auch die Großschreibweise

$$C_i U_i | X_1, \dots, X_n \sim N\left(0, \frac{(X_i - \bar{X})^2 \sigma_U^2}{n \tilde{S}_X^2}\right) \text{ für } \tilde{S}_X^2 > 0$$

$(x_i - \bar{x})^2$ (Kleinschreibweise)

$\tilde{S}_X^4, \tilde{S}_X^4$

S. 617

~~Mit dem GGZ für k -te Momente (Satz 8.2.4) gilt hierbei:~~

$$\tilde{S}_X^2 \xrightarrow{p} \sigma_X^2. \quad (12.1.73)$$

~~Man beachte hierzu Beispiel 8.3.6 mit Resultat (8.3.20).~~ Aus (12.1.71), (12.1.72) und **mit Satz 10.1.4**

S. 617

$$\begin{aligned} E(\hat{\beta}_1) &= \beta_1 + E\left[\frac{1}{\tilde{S}_X^2} \frac{1}{\sqrt{n}} \sum_{i=1}^n (X_i - \bar{X}) U_i\right] \\ &= \beta_1 + E\left[E\left(\frac{1}{\tilde{S}_X^2} \frac{1}{\sqrt{n}} \sum_{i=1}^n (X_i - \bar{X}) U_i \mid X_1, \dots, X_n\right)\right] \\ &= \beta_1 + E\left[\frac{1}{\tilde{S}_X^2} \frac{1}{\sqrt{n}} \sum_{i=1}^n (X_i - \bar{X}) E(U_i \mid X_1, \dots, X_n)\right] = \beta_1. \end{aligned}$$

jeweils n statt \sqrt{n}

S. 624

$$\sqrt{n}(\hat{\beta}_j - \beta_j) \stackrel{a}{\sim} N\left(0, \underbrace{\sigma_{\hat{\beta}_j}^2}_{n\sigma_{\beta_j}^2}\right) \quad \text{und} \quad \underbrace{\hat{\sigma}_{\hat{\beta}_j}^2}_{n\hat{\sigma}_{\hat{\beta}_j}^2} \xrightarrow{p} \underbrace{\sigma_{\hat{\beta}_j}^2}_{n\sigma_{\beta_j}^2}$$

S. 626

• Äquivalenz der Tests auf $\beta_1 = 0$ und $\rho_{XY} = 0$ • Die Teststatistik bezüglich β_1 in den Modellen KN und KS für $\beta_1 = 0$ ist mit derjenigen des Korrelationstests (Satz 11.3.8) identisch, d.h. es gilt:

$$\tilde{T}_{\beta_1} = T_R \quad \text{für} \quad \beta_1 = 0. \quad (12.1.91)$$

$\beta_{1,0}$

S. 628

$$\tilde{T}_{\beta_1} = \tilde{T}_D \quad \text{für} \quad \beta_1 = \mu_1 - \mu_0. \quad (12.1.92)$$

Weiter lautet das approximative Konfidenzintervall für β_1 in Modell BH

$$[\hat{\beta}_1 - z_{1-\alpha/2} \cdot \hat{\sigma}_{\hat{\beta}_1}, \hat{\beta}_1 + z_{1-\alpha/2} \cdot \hat{\sigma}_{\hat{\beta}_1}],$$

Dieses ist mit demjenigen für $\mu_1 - \mu_0$ von Fall (iv) in Satz 10.2.2 identisch ist. Die korrespondierende Teststatistik für Tests über β_1 ist mit derjenigen für $\mu_1 - \mu_0$ von Fall (ii) in Satz 11.3.2 identisch, d.h. es gilt:

$$\hat{T}_{\beta_1} = \hat{T}_D \quad \text{für} \quad \beta_1 = \mu_1 - \mu_0. \quad (12.1.93)$$

$\beta_{1,0} = \mu_1 - \mu_0 = \delta_0$
(iv)

S. 638

Man beachte hierzu auch Abbildung 5.2.18. Den Standardfehler der Regression erhält man aus den KQ-Residuen aus Tabelle 5.2.3:

...

Weiter ergibt sich aus Tabelle 5.2.3:

$$\bar{x} = \frac{15}{5} = 3, \quad \frac{1}{5} \sum_{i=1}^n x_i^2 = \frac{55}{5} = 11 \quad \text{und}$$

5.2.4

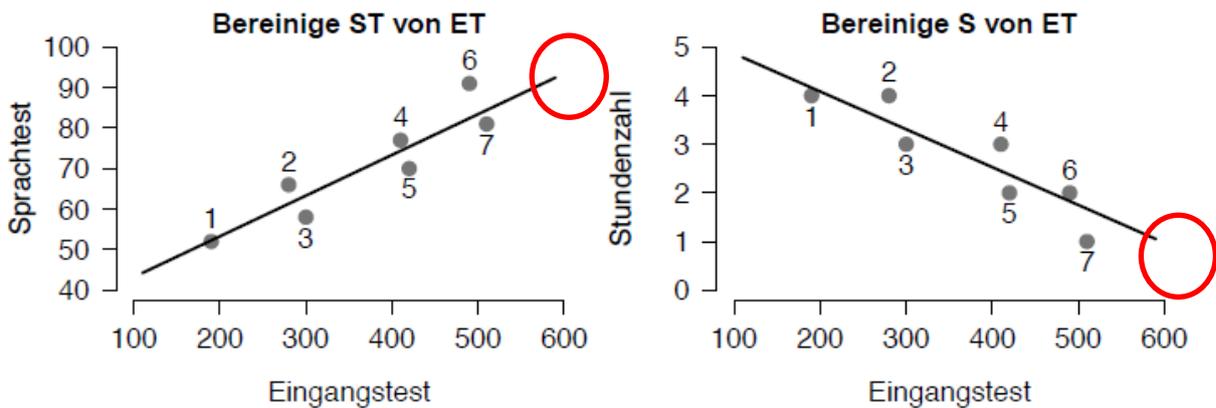
S. 639

Mit Blick auf Abbildung 12.1.4 erscheint das nichtsignifikante Ergebnis nicht sonderlich überraschend. Die Wirkungsverhältnisse erscheinen auch optisch noch nicht ein-

12.1.14

S. 649

Abb. 12.2.2: Lineares Herausrechnen von ET aus ST und S



In den beiden Diagrammen ist die 8. Beobachtung jeweils nicht mehr sichtbar
Koordinate (630, 95) bzw. (630, 1).

S. 651

$$r_{XY \cdot Z} = \frac{r_{XY} - r_{YZ}r_{XZ}}{(1 - r_{YZ}^2)(1 - r_{XZ}^2)}; \quad \text{sofern } \tilde{s}_Y^2 > 0. \quad (12.2.2)$$

Vom Nenner ist noch die Wurzel zu ziehen

S. 653

$$\rho_{XY \cdot Z} = \frac{\rho_{XY} - \rho_{YZ}\rho_{XZ}}{(1 - \rho_{YZ}^2)(1 - \rho_{XZ}^2)}, \quad \text{sofern } \sigma_Y^2 > 0. \quad (12.2.6)$$

Vom Nenner ist noch die Wurzel zu ziehen

$$\begin{aligned} \text{Corr}(Y - \hat{Y}, X - \hat{X}) &= \frac{\text{Cov}(Y - \hat{Y}, X - \hat{X})}{\sqrt{\text{Var}(Y - \hat{Y})\text{Var}(X - \hat{X})}} \\ &= \frac{\rho_{XY} - \rho_{YZ}\rho_{XZ}}{(1 - \rho_{YZ}^2)(1 - \rho_{XZ}^2)} \end{aligned}$$

Vom Nenner ist noch die Wurzel zu ziehen

S. 655

$$\hat{\rho}_{XY \cdot Z} = \frac{R_{XY} - R_{YZ}R_{XZ}}{(1 - R_{YZ}^2)(1 - R_{XZ}^2)}$$

Vom Nenner ist noch die Wurzel zu ziehen

S. 656

$$\hat{y}_i = \hat{y}(x_i) = \hat{b}_0 + \hat{b}_1 x_{i1} + \hat{b}_2 x_{i2} + \dots + \hat{b}_p x_{ip} \quad \text{für } i = 1, \dots, n$$

$\hat{y}(x_{i1}, x_{i2}, \dots, x_{ip})$

S. 667

Damit wird der wahre Wert β_X asymptotisch verzerrt geschätzt, falls wie zuvor

$$\beta_Z \neq 0 \quad \text{und} \quad \sigma_{XZ} = 0.$$

≠ (ungleich)

S. 668

Angenommen, X und Z üben jeweils einen direkten **kausalen** Effekt auf Y aus. Diese Effekte werden mit den Koeffizienten β_X und β_Z modelliert. ~~Außerdem übe X auf Z einen direkten kausalen Einfluss aus, der durch $\beta_{Z \sim X}$ modelliert wird (Abb. 12.2.5).~~ Ändert man nun X um Δx Einheiten, so ändert sich Z um $\beta_{Z \sim X} \Delta x$ Einheiten. Diese allein durch X bewirkte Änderung in Z führt wiederum zu einer Änderung in Y um

$$\beta_Z \beta_{Z \sim X} \Delta x$$

Einheiten. Es erscheint nun naheliegend den OVB von X als **indirekten **kausalen** Effekt** von X über Z auf Y zu interpretieren. Aufgrund von (12.2.7) gilt:

$$\beta_{Y \sim X} = \beta_X + \beta_Z \beta_{Z \sim X}. \quad (12.2.40)$$

Bezeichnen wir den im fehlspezifizierten Modell geschätzten Effekt $\beta_{Y \sim X}$ als **kausalen Gesamteffekt** oder **totalen **kausalen** Effekt** von X auf Y , so erhalten wir im Endeffekt die einfache Formel (vgl. hierzu etwa Urban und Mayerl [2011, Abschnitt 5.3]).

„**kausal**“ stets in Klammern setzen

Außerdem bestehe zwischen X und Z eine (kausale) Abhängigkeit, die über $\beta_{Z \sim X}$ modelliert wird.

Tabelle 12.2.3: Beobachtungen mit binären Regressoren

i	y_i	x_{1i}	x_{2i}
1	1.0	0	0
2	1.3	0	0
3	0.7	0	0
4	1.7	1	0
5	2.1	1	0
6	1.9	1	0
7	2.3	1	0
8	1.3	0	1
9	1.2	0	1
10	1.8	0	1
11	1.4	0	1
12	1.6	0	1

Linien zur Gruppierung: 3, 4 und 5 Beobachtungen innerhalb der Gruppen