Lösungen zu den Übungsaufgaben zur Vorlesung

Aufbaumodul Statistik (AM Statistik)

— Kurzfassung —

1 Zufallsvariablen und deren Verteilung

Aufgabe 1.1

Wahrscheinlichkeitsfunktion:

$$f_X(x) = P(X = x) = \begin{cases} \frac{11}{36} & \text{für } x = 1, \\ \frac{9}{36} & \text{für } x = 2, \\ \frac{7}{36} & \text{für } x = 3, \\ \frac{5}{36} & \text{für } x = 4, \\ \frac{3}{36} & \text{für } x = 5, \\ \frac{1}{36} & \text{für } x = 6, \end{cases}$$

Verteilungsfunktion:

$$F_X(x) = P(X \leq x) = \begin{cases} 0 & \text{für } x < 1, \\ \frac{11}{36} & \text{für } 1 \leq x < 2, \\ \frac{20}{36} & \text{für } 2 \leq x < 3, \\ \frac{27}{36} & \text{für } 3 \leq x < 4, \\ \frac{32}{36} & \text{für } 4 \leq x < 5, \\ \frac{35}{36} & \text{für } 5 \leq x < 6, \\ 1 & \text{für } 6 \leq x. \end{cases}$$

Aufgabe 1.2

a) Verteilungsfunktion:

$$F_X(x) = \begin{cases} 0 & \text{für} \quad x < 1, \\ 0,1 & \text{für} \quad 1 \leq x < 2, \\ 0,2 & \text{für} \quad 2 \leq x < 3, \\ 0,3 & \text{für} \quad 3 \leq x < 4, \\ 0,4 & \text{für} \quad 4 \leq x < 5, \\ 0,5 & \text{für} \quad 5 \leq x < 6, \\ 1 & \text{für} \quad 6 \leq x. \end{cases}$$

b) $E[X] = 4,5$
$E[X^2] = 23,5$
$V[X] = 3,25$

c) Quantilfunktion:

$$Q_X(p) = \begin{cases} 1 & \text{für} \quad 0 < p \leq 0,1, \\ 2 & \text{für} \quad 0,1 < p \leq 0,2, \\ 3 & \text{für} \quad 0,2 < p \leq 0,3, \\ 4 & \text{für} \quad 0,3 < p \leq 0,4, \\ 5 & \text{für} \quad 0,4 < p \leq 0,5, \\ 6 & \text{für} \quad 0,5 < p < 1. \end{cases}$$

Aufgabe 1.3

a) Dichtefunktion:

$$f(x) = F'(x) = \begin{cases} \frac{1}{2}x & \text{für } 0 \leq x \leq 2, \\ 0 & \text{sonst.} \end{cases}$$

Träger: $[0;2]$.

b) $Q(p) = 2\sqrt{p} \qquad \text{für } 0 < p < 1.$

c) $x_{0,5} \approx 1,4142$
$E[X] = \frac{4}{3}$
$V[X] = \frac{2}{9}$

d) $0,0625$

Aufgabe 1.4

a) Für $c = 2$ ist $f(x) := \begin{cases} 2x-2 & \text{für } 1 \leq x \leq 2, \\ 0 & \text{sonst,} \end{cases}$ eine Dichtefunktion.

b) Verteilungsfunktion:

$$F(x) = \begin{cases} 0 & \text{für } x < 1, \\ x^2 - 2x + 1 & \text{für } 1 \leq x \leq 2, \\ 1 & \text{für } x > 2. \end{cases}$$

Quantilfunktion:

$$Q(p) = 1 + \sqrt{p} \qquad \text{für } 0 < p < 1.$$

c) $p = 0,2 : 1,4472$
$p = 0,9 : 1,9487$

d) $0,8889$

Aufgabe 1.5

a) $E[Y] \quad = 12,5$
$\sqrt{V[Y]} \quad = 20$

b) $P\left(|X - E[X]| \geq 10\right) = 0,25$

Aufgabe 1.6

X sei die Differenz zwischen der Augenzahl beim zweiten Wurf und der Augenzahl beim ersten Wurf. Dann ist $P(X > 0) = 0,4167$.

Aufgabe 1.7

a) $\Omega = \{(Z), (W,Z), (W,W,Z), (W,W,W,Z), (W,W,W,W,Z), (W,W,W,W,W)\}$

b) Wahrscheinlichkeitsfunktion

$$\begin{aligned} P(X=1) &= \frac{1}{2} \\ P(X=2) &= \frac{1}{2}\cdot\frac{1}{2} = \frac{1}{4} \\ P(X=3) &= \frac{1}{2}\cdot\frac{1}{2}\cdot\frac{1}{2} = \frac{1}{8} \\ P(X=4) &= \frac{1}{2}\cdot\frac{1}{2}\cdot\frac{1}{2}\cdot\frac{1}{2} = \frac{1}{16} \\ P(X=5) &= 1 - \left(\frac{1}{2}+\frac{1}{4}+\frac{1}{8}+\frac{1}{16}\right) = \frac{1}{16}\,. \end{aligned}$$

Verteilungsfunktion

$$F(x) = \begin{cases} 0, & \text{falls } x < 1, \\ \frac{1}{2}, & \text{falls } 1 \le x < 2, \\ \frac{3}{4}, & \text{falls } 2 \le x < 3, \\ \frac{7}{8}, & \text{falls } 3 \le x < 4, \\ \frac{15}{16}, & \text{falls } 4 \le x < 5, \\ 1, & \text{falls } x \ge 5. \end{cases}$$

c)

$$Q(p) = \begin{cases} 1, & \text{falls } 0 < p \le \frac{1}{2}, \\ 2, & \text{falls } \frac{1}{2} < p \le \frac{3}{4}, \\ 3, & \text{falls } \frac{3}{4} < p \le \frac{7}{8}, \\ 4, & \text{falls } \frac{7}{8} < p \le \frac{15}{16}, \\ 5, & \text{falls } \frac{15}{16} < p < 1. \end{cases}$$

Der Median hat den Wert $x_{0.5} = 1$. Für das erste und dritte Quartil erhält man $x_{0.25} = 1$ bzw. $x_{0.75} = 2$.

d) $E[X] = 1,9375$
$V[X] = 1,4336$

Aufgabe 1.8

a) $c = 2$

b) Verteilungsfunktion

$$F(x) = \begin{cases} 1 - \frac{8}{x^3}, & \text{falls } x \ge 2, \\ 0 & \text{sonst} \end{cases}$$

Quantilfunktion

$$Q(p) = \sqrt[3]{\frac{8}{1-p}}, \quad 0 < p < 1\,.$$

c) $E[X] = 3$; $V[X] = 3$; $x_{0,5} \approx 2,5198$

d)

$$\begin{aligned} P(X \leq 3) &= 0,7037 \\ P(2,5 \leq X \leq 3,5) &= 0,3254 \\ P(X = 3) &= 0 \\ P(X \geq 4) &= 0,125. \end{aligned}$$

Aufgabe 1.9

$0,4444$

Aufgabe 1.10

a) $0,25$

b) $0,3056$

2 Spezielle diskrete Verteilungen

Aufgabe 2.1

a) $X \sim B\left(n=5; \pi=\frac{1}{6}\right)$

b) $P(X=2)=0,16075$

c) $P(X>1)=0,19624$

Aufgabe 2.2

a) $P(X \geq 5)=0,14484$

b) $E[X] = 3$
$V[X] = 2$

Aufgabe 2.3

a) Annahme: Urlauber infizieren sich unabhängig voneinander.

$P(X \geq 1)=0,8647$

Möglich ist hier auch die approximative Berechnung der gesuchten Wahrscheinlichkeit mittels der Poisson-Verteilung (Bedingungen sind erfüllt).

b) $E[X] = 2$
$\sqrt{V[X]} = 1,4141$

Aufgabe 2.4

Ohne Zurücklegen: $P(X \geq 2)=0,04496$

Mit Zurücklegen: $P(X \geq 2)=0,0523$

Aufgabe 2.5

a) 0,0315

b) 0,6302

c) mindestens 60 €

Aufgabe 2.6

a) 0,0183

b) 0,7149

c) X = Anzahl Anrufe zwischen 17.15 Uhr und 17.45 Uhr
Z = Anzahl Anrufe zwischen 17.15 Uhr und 18.15 Uhr

$$P(X \leq 5 \mid Z = 10) = 0,6230$$

Aufgabe 2.7
X sei die Anzahl defekter Chips beim Kauf von fünf dieser Chips. Dann ist $X \sim B(5, 0.05)$.

a) $P(X \leq 2) = 0.9988$

b) $P(X < 2) = 0.9774$

c) $P(X \geq 3) = 0.0012$

d) $P(X \geq 1) = 0.2262$

Aufgabe 2.8
X sei die Anzahl gezogener Weißwein-Fässer.

a) $P(X = 3) = 0.2195$

b) $Y = 6 - X$ ist Anzahl entnommener Rotwein-Fässer.
$P(Y \geq 1) = 0.9986$

c) $P(X \leq 4) = 0.9822$

Aufgabe 2.9
X = Anzahl Versuche bis zum ersten Erfolg $\sim G(\pi = 0,6)$.

a) $E(X) = 1,6667$

b) $P(X > 6 \mid X > 2) = 0,0256$

Aufgabe 2.10
Sei X_1 = Anzahl Motorroller pro Stunde $\sim Po(\mu_1 = 9)$.

a) Sei X_2 = Anzahl Motorroller pro halbe Stunde.
$E(X_2) = 4,5$

b) Sei X_3 = Anzahl Motorroller pro 20 Minuten.
$P(X_3 > 4) = 0,1848$

3 Spezielle stetige Verteilungen

Aufgabe 3.1
$0,209$

Aufgabe 3.2

a) $0,9251$

b) $0,2611$

c) $0,0341$

d) 0

Aufgabe 3.3
$\sigma_X = 6$, $\mu_X = 10$.

Aufgabe 3.4
$E[X] = 28037\text{h} \approx 3,2$ Jahre

$x_{0,5} = 19432\text{h} \approx 2,2$ Jahre

Aufgabe 3.5

a) $0,3679$

b) $0,3679$

c) Die Beschreibung der Lebensdauer durch eine Exponentialverteilung erscheint nicht angemessen. Die Ergebnisse aus a) und b) zeigen, dass es theoretisch keinen Einfluss hat, wie lange z.B. das technische Gerät bereits im Einsatz ist. Praktisch jedoch treten Abnutzungserscheinungen auf, die auf diese Weise unberücksichtigt bleiben.

Aufgabe 3.6

a) $P(100 \leq X \leq 200) = 0,2778$ und $x_{0,5} = 180$

b) $P(10n \leq X \leq 10(n+1)) = 0,0278$

c) $E[Y] = 17,5140$, $\sqrt{V[Y]} = 10,3883$, $P(Y < 15) = 0,4167$ und $y_{0,5} = 17$

Aufgabe 3.7

a) $P(X > 200) = 0,6703$

b) $P(200 \leq X \leq 300) = 0,1215$

c) Mit mindestens 90% Wahrscheinlichkeit überlebt das Bauteil also alle Zeitpunkte vor $t_0 = 52,68$.

d) $\lambda = 0,002107$

Aufgabe 3.8

a) $P(|X - \mu_X| < 1) = 0,9544$ und Tschebyscheff ($\epsilon = 1$): $P(|X - \mu_X| < 1) \geq 0,75$

b) $P(|X - \mu_X| < 1) = 0,9933$ und Tschebyscheff ($\epsilon = 1$): $P(|X - \mu_X| < 1) \geq 0,9375$

c) $P(|X - \mu_X| < 1) = \frac{1}{\alpha}$ und Tschebyscheff ($\epsilon = 1$): $P(|X - \mu_X| < 1) \geq \frac{3-\alpha^2}{3}$

Aufgabe 3.9

a) Die Lösung dieses Systems ist $\mu = 1$ und $\sigma^2 = 1$.

b) $P(X < 0) = 0,1587$

4 Gemeinsame Verteilung von Zufallsvariablen

Aufgabe 4.1

a) $E[Z_1] = 172$
$V[Z_1] = 678$

b) $E[Z_2] = -10$
$V[Z_2] = 678$

c) $E[Z] = 162$
$V[Z] = 360$

Aufgabe 4.2

a) $E[G] = 300.000$ €
$\sqrt{V[G]} = 101.612$ €

b) $E[G] = 300.000$ €
$\sqrt{V[G]} = 106.701$ €

Aufgabe 4.3

$E[X+Y] = 150$
$V[X+Y] = 12.500$
$Cov(X,Y) = 0$
$\rho(X,Y) = 0$
$E[X \cdot Y] = 5.000$

Aufgabe 4.4

a) Sei Y die Gesamteinnahme des Taxiunternehmers an einem Tag.
$E[Y] = 750$ € und $\sqrt{V[Y]} = 180,28$ €

b) Sei Z die Jahreseinnahme des Taxiunternehmers.
$E[Z] = 270.000$ € und $\sqrt{V[Z]} = 3.420,53$ €

Aufgabe 4.5

a) $\rho_{R_B,R_G} = 0,5669$

b) Anteil der Aktie B-Bank: 20%
Anteil der Aktie G-Versicherung: 80%
$E[R_{ges}] = 8,4\%$
$V[R_{ges}] = 0,0228$

Aufgabe 4.6

a) $0,6700$

b) (i) Sei W der Wert des Portefeuilles nach einem Jahr. Dann ist $E[W] = 14.688$ und $V[W] = 1.849.600$.

(ii) $\leq 0,25$

(iii) $\geq 0,5$

5 Grenzwertsätze und Approximationen

Aufgabe 5.1

a) $P(X < 50) = P(X \leq 49) \approx \Phi\left(\frac{49 + 0,5 - 70}{\sqrt{21}}\right) = \Phi(-4,4735) \approx 0$

b) $0,9596$

Aufgabe 5.2
0,6472

Aufgabe 5.3
0,0431

Aufgabe 5.4

a) $E[X] = 4,5$
$V[X] = 8,25$

b) $E[Z] = 225$
$V[Z] = 412,5$

c) Z ist annähernd normalverteilt mit Erwartungswert und Varianz wie in Teil b), also
$$Z \simeq N(225; 412,5).$$

Aufgabe 5.5
$X_i =$ Gehalt des i-ten Fahrers

a) $M =$ monatliche Nettogehaltssumme aller FahrerInnen
$$E[M] = 120 \cdot 1.800 = 216.000$$
$$V[M] = 120 \cdot 200^2 = 4.800.000$$

b) $P(M > 219\,000) \approx 0,0853$

Aufgabe 5.6
$X =$ Anzahl Doppelsechser bei 100 Würfen. $P(X \geq 3) = 0,5278$.

Aufgabe 5.7

a) $X =$ Anzahl Tippfehler pro Seite.

   (i) $P(X \geq 1) = 0,8653$
   (ii) $P(X \geq 2) = 0,5946$

b) $Y =$Anzahl Tippfehler im gesamten Manuskript.

   (i) $E[Y] = 400, V[Y] = 398.$
   (ii) $P(300 \leq Y \leq 500) \approx 1$

6 Stichproben und Statistiken

Aufgabe 6.1

a) Summe/Anzahl aller $n = 1.100$ Wähler, welche die Partei A gewählt haben.

b) $\sum_{i=1}^{1.100} X_i$ ist binomialverteilt mit Erwartungswert 1100π und Varianz $1.100\pi(1-\pi)$, also

$$\sum_{i=1}^{1.100} X_i \sim B(1100;\pi) \simeq N(1.100\pi; 1.100\pi(1-\pi)) .$$

c) $1 - \Phi(6,80) \approx 0$

Aufgabe 6.2

a)

$$\sum_{i=1}^{50} X_i \simeq N\left(\frac{50}{\lambda} = 250; \frac{50}{\lambda^2} = 1.250\right)$$
$$\frac{1}{50}\sum_{i=1}^{50} X_i \simeq N\left(\frac{1}{50}\cdot\frac{50}{\lambda} = 5; \frac{1}{50^2}\cdot\frac{50}{\lambda^2} = 0,5\right)$$

b) Die ZVen Y_i sind Bernoulli-verteilt mit $\pi = P(Y_i = 1) = P(X_i \geq 3) = 1 - P(X_i \leq 3) = 1 - (1 - e^{-3\lambda}) = e^{-3\lambda}$.

(i) $\sum_{i=1}^{50} Y_i$ ist die Anzahl Glühlampen, deren Brenndauer mindestens 3.000 Stunden beträgt; $\frac{1}{50}\sum_{i=1}^{50} Y_i$ ist der Anteil Glühlampen, deren Brenndauer mindestens 3.000 Stunden beträgt.

(ii) Als Summe Bernoulli-verteilter ZVen ist $\sum_{i=1}^{50} Y_i$ binomialverteilt mit $n = 50$ und (mit $\lambda = 0,2$) $\pi = e^{-3\cdot 0,2} = e^{-0,6} = 0,5488$.

Aufgabe 6.3

a) $E\left[\overline{X}\right] = 100$

b) $V\left[\overline{X}\right] = 3,6$

c) Aus den Teilen a) und b) folgt, dass $\overline{X} \sim N(100; 3,6)$.
Damit ist

$$\frac{\overline{X} - \mu_{\overline{X}}}{\sigma_{\overline{X}}} = \frac{\overline{X} - 100}{\sqrt{3,6}} \sim N(0;1)$$

d) 0,1469

e) $x_{0,9} = 102,43$ und $x_{0,1} = 97,57$
Das p-Quantil ist derjenige Wert, der mit Wahrscheinlichkeit p höchstens erreicht wird.

7 Punkt- und Intervallschätzung

Aufgabe 7.1

a) $E[T_1] = \mu$
$\Rightarrow T_1 = \overline{X}$ ist erwartungstreu für μ

$E[T_2] = \mu$
$\Rightarrow T_2 = \frac{1}{2}(X_2 + X_n)$ ist erwartungstreu für μ

$E[T_3] = \frac{4}{5}\mu$
$\Rightarrow T_3 = \frac{4}{5}\overline{X}$ ist nicht erwartungstreu für μ

b) $V[T_1] = \dfrac{V[X]}{n}$

$V[T_2] = \dfrac{1}{2}V[X]$

$V[T_3] = \dfrac{16 \cdot V[X]}{25 \cdot n}$

c) Schätzer $T_1 = \overline{X}$.

Aufgabe 7.2
$\widehat{\theta}_{u,o} = [35,21; 43,13]$

Aufgabe 7.3

a) $\widehat{\mu} = 187,4$ (Gramm)

b) $\widehat{\theta}_{u,o} = [181,905;\ 192,895]$

Aufgabe 7.4

a) $\widehat{\mu} = 496$ (Gramm)

b) $\widehat{\theta}_{u,o} = [489,8.316;\ 502,1.684]$

c) Mindeststichprobenumfang 68

Aufgabe 7.5

a) $[0,3.079; 0,4.421]$

b) Wähle $\pi^* = 0,375$ (siehe Teil a)). Dann hätten mindestens 1.001 Personen befragt werden müssen. (Fordert man lediglich $0,1 < \pi^* < 0,9$, so erhält einen Mindeststichprobenumfang von 1.068 Personen.

Aufgabe 7.6

a) $[0,7892; 0,8908]$

b) 207

Aufgabe 7.7

a) Konfidenzintervall für μ: $[0,057121; 0,082879]$.

b) Der Stichprobenumfang muss mindestens $n = 60$ sein.

Aufgabe 7.8

a) $\widehat{\pi}_A = 0,41$. Approximatives Konfidenzintervall:

$$\widehat{\theta}_{u,o} \approx \widehat{\pi}_A \mp u_{0,975}\sqrt{\frac{\widehat{\pi}_A(1-\widehat{\pi}_A)}{n}} = 0,41 \mp 1,96 \cdot 0,01555$$

b) Konfidenzintervall lautet $[45; 53]$.

c) Nein.

Aufgabe 7.9
2377

8 Hypothesentests für Erwartungswerte

Aufgabe 8.1
$H_0 : \mu \geq 1000$ wird wegen $t = -2,07 < -t_{6;0,95} = -1,9432$ abgelehnt (t-Test).

Aufgabe 8.2
$H_0 : \mu \geq 3000$ wird wegen $t = -4,375 < -1,6449$ abgelehnt (approximativer t-Test).

Aufgabe 8.3

a) 0,1336

b) $H_0 : \mu_X = 60$ wird wegen $|t| = 2,5 > u_{1-0,025} = u_{0,975} = 1,96$ abgelehnt (Gauß-Test).

Aufgabe 8.4

a) $H_0 : \mu_X = \mu_Y$ wird wegen $|t| = 2,4495 \not> u_{1-\frac{\alpha}{2}} = u_{0,995} = 2,5758$ nicht abgelehnt (Approximativer Zweistichproben -Gauß-Test).

b) Nein, denn es führt zum gleichen Test bzw. zur gleichen Teststatistik.

Aufgabe 8.5
Stichprobe bei den Männern: X
Stichprobe bei den Frauen: Y
$H_0 : \mu_X \leq \mu_Y$ wird wegen $t = 2,4946 > u_{1-\alpha} = u_{0,95} = 1,6449$ abgelehnt (Approximativer Zwei-Stichproben- Gauß-Test).

Aufgabe 8.6

a) $H_0 : \mu_X \geq 38$ wird wegen $t = -1.8 < -u_{0,95} = -1,6449$ abgelehnt (Gauß-Test).

b) $\beta = P(H_0 \text{ nicht ablehnen} \mid H_0 \text{ falsch}) = 1 - \Phi(0,1551) = 0,4384$

Aufgabe 8.7

a) (i) $H_0 : \mu = 330,\ H_1 : \mu \neq 330,\ H_0$ ablehnen, falls $|T| > 2,5758$,
(ii) $H_0 : \mu \geq 330,\ H_1 : \mu < 330,\ H_0$ ablehnen, falls $T < -2,3263$,
(iii) $H_0 : \mu \leq 330,\ H_1 : \mu > 330,\ H_0$ ablehnen, falls $T > 2,3263$.
Testgröße ist jeweils
$$T = \frac{\overline{x} - 330}{1.5}\sqrt{30}\,.$$

b) H_0 wird abgelehnt, falls

(i) $\overline{x} < 329,29$ oder $\overline{x} > 330,71$.

(ii) $\overline{x} < 329,36$,

(iii) $\overline{x} > 330,64$.

c) Es ist $t = -2,4465$. Somit wird im Fall 2. H_0 abgelehnt, in den Fällen 1. und 3. jedoch nicht.

Aufgabe 8.8

a) Folgende Annahmen sind zu treffen:

- Es handelt sich um einfache Zufallsstichproben aus normalverteilten Grundgesamtheiten.
- Die beiden gezogenen Stichproben sind unabhängig.
- Die Varianzen der beiden Grundgesamtheiten sind gleich.

Sei X der Preis bei einem deutschen und Y der Preis bei einem europäischen Autohändler. Hypothesen: $H_0 : \mu_X = \mu_Y$ (wenn man nur testen will, ob ein Unterschied besteht) bzw. $H_0 : \mu_X \leq \mu_Y$ (wenn man testen will, ob die Preise in Deutschland höher sind als im benachbarten Ausland).

b) $H_0 : \mu_X = \mu_Y$ wird wegen $|t| = 1.3624 \not> t_{15;0,975} = 2.1314$ nicht abgelehnt (Zweistichproben-t-Test).

Testet man $H_0 : \mu_X \leq \mu_Y$, so wird diese Nullhypothese ebenfalls nicht abgelehnt.

9 Lineare Regression

Aufgabe 9.1

a) $\widehat{\beta_1} = 1,0742$, $\widehat{\beta_0} = 39,5506$.

b) $0,11987$

c) $H_0 : \beta_1 = 0$ wird abgelehnt, da $|T| = 8,96 > t_{n-2;1-\frac{\alpha}{2}} = t_{8;0,975} = 2,306$

d) $H_0 : \beta_1 = 1,1$ kann nicht abgelehnt werden, da $|T| = 0,2152 \leq t_{n-2;1-\frac{\alpha}{2}} = t_{8;0,975} = 2,306$

Aufgabe 9.2

a) $\hat{\beta}_0 = -118,1250$, $\hat{\beta}_1 = 1,1042$, $\hat{\sigma}^2 = 21,35$.

b) $\widehat{\sigma_{\hat{\beta}_0}} = 23,7252$, $\widehat{\sigma_{\hat{\beta}_1}} = 0,1361$.

c) $[-170,9871 \quad ; \quad -65,2629]$.

d) Die Nullhypothese $H_0 : \beta_1 = 1$ wird wegen $|t| = 0,7656 > t_{10;0.975} = 2,2281$ nicht abgelehnt.

Aufgabe 9.3

a) $\hat{\beta}_1 = 0,80557$, $\hat{\beta}_0 = 112,042$.
Regressionsgleichung: $\widehat{Y_i} = 112,042 + 0,80557 \cdot x_i$.

b) $[-87,0692 \quad ; \quad 311,1524]$

c) $H_0 : \beta_0 \geq 500$ muss abgelehnt werden, da $T = -3,7862 < -t_{n-2;1-\alpha} = -t_{6;0,95} = -1,9432$

Aufgabe 9.4

a) $\hat{\beta}_1 = 0,1785$, $\hat{\beta}_0 = -5,8609$.

b) Konfidenzintervall für β_1: $[0,1488; 0,2083]$.

c) $H_0 : \beta_1 \leq 0,15$ wird wegen $T = 1,8793 > 1,6449 = t_{48;0,95}$ abgelehnt.

Aufgabe 9.5

a) $H_0 : \beta_2 \geq 0$ wird wegen $t = -2,0029 < -1,6449 = u_{0,95}$ abgelehnt.

b) $H_0 : \beta_1 = 0$ wird wegen p-Wert < 0,0001 oder $|t| = 4,6569 > -1,96 = u_{0,975}$ abgelehnt.

c) $\hat{\beta}_3$ ist nicht signifikant von 0 verschieden.

d) $R^2 = 0,3835 =$ Anteil der durch die Regression erklärten Varianz.

e) $H_0 : \beta_1 = \beta_2 = \beta_3 = 0$ wird wegen $t = 9,538 > 2,81 = F^{-1}_{F_{3,46}}(0,95)$ abgelehnt.

Aufgabe 9.6

a) $H_0 : \beta_2 = 0$ wird wegen $|t| = 4,7363 > 2,11 = F^{-1}_{t_{17}}(0,975)$ abgelehnt.

b) $H_0 : \beta_1 = \beta_2 = 0$ wird wegen $t = 52,21 > 3,59 = F^{-1}_{F_{2,17}}(0,95)$ abgelehnt.

c) β_1: $x_1 \uparrow +1 \Rightarrow Y \uparrow +0,454$
β_2: $x_2 \uparrow +1 \Rightarrow Y \uparrow +0,431$
β_0: Achsenabschnitt = 5,4
R^2: Regression erklärt 86 % der Varianz von Y.

d) $r_{\hat{\beta}_1;\hat{\beta}_2} = -0,4995 \Rightarrow$ Schätzer sind tendenziell gegenläufig.

Aufgabe 9.7

a) Logarithmieren der Produktionsfunktion und Einführung von Störtermen führt auf den Regressionsansatz $\ln(Y_i) = \beta_0 + \beta_1 \ln K_i + \beta_2 \ln L_i + U_i$, $i = 1, \ldots 6$, mit $\ln b_0 = \beta_0$ und Störtermen $U_1, \ldots, U_6$.

b)
```
Call:
lm(formula = log(Y) ~ log(K) + log(L))

Residuals:
1         2         3         4         5         6
0.002571 -0.004332  0.003858 -0.009266  0.004093  0.003075

Coefficients:
Estimate Std. Error t value Pr(>|t|)
(Intercept)  4.34905    1.17611   3.698  0.03433 *
log(K)       0.28563    0.04058   7.039  0.00589 **
log(L)       0.06110    0.15000   0.407  0.71107
---
Signif. codes: 0 ‘***’ 0.001 ‘**’ 0.01 ‘*’ 0.05 ‘.’ 0.1 ‘ ’ 1

Residual standard error: 0.007126 on 3 degrees of freedom
Multiple R-squared:  0.9871,Adjusted R-squared:  0.9785
F-statistic:   115 on 2 and 3 DF,  p-value: 0.001462
```

c) β_1 und β_2 sind die partiellen Elastizitäten des Outputs bezüglich des Faktors Kapital bzw. Arbeit. Ein Anstieg des Kapitalstocks um 1% würde (unter Vernachlässigung des Störterms) also einen Anstieg des Outputs um ungefähr 0.2856% zur Folge haben. Entsprechend würde ein Anstieg des Faktors Arbeit um 1% einen Anstieg des Outputs um ungefähr 0.06110% zur Folge haben.

d) Der p-Wert des F-Tests ist 0.001462. Daher wird auf jedem gängigen Signifikanzniveau die Hypothese $H_0 : \beta_1 = \beta_2 = 0$ abgelehnt. Das Regressionsmodell liefert also einen signifikanten Beitrag zur Erklärung des Outputs.

e) Der p-Wert für den Test $H_0 : \beta_1 = 0$ hat den Wert 0.00589. Die Nullhypothese wird also auf den gängigen Signifikanzniveaus verworfen.

Der p-Wert für den Test $H_0 : \beta_2 = 0$ hat den Wert 0.7111. Die Nullhypothese wird also auf keinem gängigen Signifikanzniveaus verworfen.

„Kapital“ besitzt einen signifikanten Einfluss, „Arbeit“ nicht.

Aufgabe 9.8

a) Mit Y = Anzahl der Störfälle pro Jahr, X_1 = Reaktoralter [in Jahren] und X_2 = Reaktortyp (0: Siedewasserreaktor, 1: Druckwasserreaktor) lautet der Regressionsansatz $Y_i = \beta_0 + \beta_1 x_{i1} + \beta_2 x_{i2} + U_i$, $i = 1, \ldots, n$, wobei $U_1, \ldots, U_n \overset{iid}{\sim} N(0, \sigma^2)$.

b)
```
Call:
lm(formula = Störfälle.je.Jahr ~ Alter + Typ)

Residuals:
Min       1Q  Median      3Q     Max
-2.7726 -1.4289 -0.1516  1.1411  4.4682

Coefficients:
Estimate Std. Error t value Pr(>|t|)
(Intercept)  -9.4678     3.5328  -2.680 0.017946 *
Alter         0.5893     0.1124   5.245 0.000124 ***
TypDWR        0.4734     1.0823   0.437 0.668494
---
Signif. codes: 0 ‘***’ 0.001 ‘**’ 0.01 ‘*’ 0.05 ‘.’ 0.1 ‘ ’ 1

Residual standard error: 2.103 on 14 degrees of freedom
Multiple R-squared:  0.6642,Adjusted R-squared:  0.6163
F-statistic: 13.85 on 2 and 14 DF,  p-value: 0.0004811
```

c) Der p-Wert des F-Tests ist 0.0004811. Daher wird auf jedem gängigen Signifikanzniveau die Hypothese $H_0 : \beta_1 = \beta_2 = 0$ abgelehnt. Das Regressionsmodell liefert also einen signifikanten Beitrag zur Erklärung des Outputs.

d) Der p-Wert für den Test $H_0 : \beta_1 = 0$ hat den Wert 0.000124. Die Nullhypothese wird also auf den gängigen Signifikanzniveaus verworfen. Das Alter des Reaktors besitzt einen signifikanten Einfluss auf die Störanfälligkeit.

e) Der p-Wert für den Test $H_0 : \beta_2 = 0$ hat den Wert 0.6685. Die Nullhypothese wird also auf keinem gängigen Signifikanzniveau verworfen. Der Reaktortyp besitzt also keinen signifikanten Einfluss auf die Störanfälligkeit.

f) $R^2 = 0.6642$, d.h. 66.42% der Varianz der Störanfälligkeit wird durch die Regression auf das Alter und den Reaktortyp erklärt.

Aufgabe 9.9

Für lineare Einfachregression gilt $X = \begin{pmatrix} 1 & X_1 \\ \vdots & \vdots \\ 1 & X_n \end{pmatrix}, \quad Y = \begin{pmatrix} Y_1 \\ \vdots \\ Y_n \end{pmatrix}$.

Somit ist $X^T X = \begin{pmatrix} n & n\overline{X} \\ n\overline{X} & \sum_{i=1}^n X_i^2 \end{pmatrix}, \quad (X^T X)^{-1} = \frac{1}{n^2 s_X^2} \begin{pmatrix} \sum_{i=1}^n X_i^2 & -n\overline{X} \\ -n\overline{X} & n \end{pmatrix}$

und $X^T Y = \begin{pmatrix} n\overline{Y} \\ \sum_{i=1}^n X_i Y_i \end{pmatrix}$.

Daraus folgt $(X^T X)^{-1} X^T Y = \frac{1}{n^2 s_X^2} \begin{pmatrix} \sum_{i=1}^n X_i^2 \cdot n\overline{Y} - n\overline{X} \cdot \sum_{i=1}^n X_i Y_i \\ -n\overline{X} \cdot n\overline{Y} + n \cdot \sum_{i=1}^n X_i Y_i \end{pmatrix}$.

Daher ist $\hat{\beta}_1 = \frac{1}{n^2 s_X^2}(-n\overline{X} \cdot n\overline{Y} + n \cdot \sum_{i=1}^n X_i Y_i) = \frac{s_{XY}}{s_X^2}$

und $\hat{\beta}_0 = \frac{1}{n^2 s_X^2}(\sum_{i=1}^n X_i^2 \cdot n\overline{Y} - n\overline{X} \cdot \sum_{i=1}^n X_i Y_i) = \overline{Y} - \frac{s_{XY}}{s_X^2}\overline{X} = \overline{Y} - \hat{\beta}_1 \overline{X}$.

Aufgabe 9.10

a) $s_{XY} \approx 0$, also $\hat{\beta}_1 = \frac{s_{XY}}{s_X^2} \approx 0$

b) Sei $Z_t = X_t^2$, dann ist $\hat{\beta}_1 = \frac{\frac{1}{n}\sum_{i=1}^n Z_i Y_i - \overline{Z} \cdot \overline{Y}}{\frac{1}{n}\sum_{i=1}^n Z_i^2 - \overline{Z}^2} = \frac{\sum_{i=1}^n X_i^2 Y_i - \sum_{i=1}^n X_i^2 \cdot \overline{Y}}{\sum_{i=1}^n X_i^4 - \sum_{i=1}^n X_i^2}$ und

$\hat{\beta}_0 = \overline{Y} - \hat{\beta}\overline{Z} = \overline{Y} - \hat{\beta} . \frac{1}{n}\sum_{i=1}^n X_i^2$

Aufgabe 9.11

Regressormatrix: $X = \begin{pmatrix} 1 \\ \vdots \\ 1 \end{pmatrix}, \quad (n \times 1)$

$\hat{\beta}_0 = (X^T X)^{-1} X^T y = \frac{1}{n}\sum_{t=1}^n Y_t = \overline{Y}$

Aufgabe 9.12

a) $PY + MY = (X(X^T X)^{-1} X^T + I - X(X^T X)^{-1} X^T)Y = IY = Y$

b) $PY = X(X^T X)^{-1} X^T Y = X\hat{\beta} = \hat{Y}$

c) $MY = (I - P)Y = Y - PY = Y - \hat{Y} = \hat{U}$

Aufgabe 9.13

a) $\ln(y) = \beta_1 \ln(x_1) + x_2 \tilde{\beta}_2 + U \quad \text{mit } \tilde{\beta}_2 = \ln(\beta_2)$

b) Die Gleichung ist bereits linear in den Parametern.

c) $y^{-1/3} = \beta_0 + \beta_1 x_1 + U$

Aufgabe 9.14

(LIN): β_1 ist ein Steigungskoeffizient. Eine absolute Erhöhung der Einkommen um $\Delta X = 1$ führt ceteris paribus zu einer absoluten Änderung der Wohnungsausgaben von $\Delta Y \approx \beta_1$.

(LOG): β_1 ist eine Elastizität. Eine relative Erhöhung der Einkommen um $\frac{\Delta X}{X} = 0{,}01 = 1\%$ führt ceteris paribus zu einer relativen Änderung der Wohnungsausgaben von $\frac{\Delta Y}{Y} \approx \beta_1 \cdot 0{,}01$.

(SEMLOG1): Auch (LOGLIN) genannt. β_1 ist eine Semi-Elastizität. Eine absolute Erhöhung der Einkommen um $\Delta X = 1$ führt ceteris paribus zu einer relativen Änderung der Wohnungsausgaben von $\frac{\Delta Y}{Y} \approx \beta_1$.

(SEMLOG2): Auch (LINLOG) genannt. β_1 ist eine Semi-Elastizität. Eine relative Erhöhung der Einkommen um $\frac{\Delta X}{X} = 0{,}01 = 1\%$ führt ceteris paribus zu einer absoluten Änderung der Wohnungsausgaben von $\Delta Y \approx \beta_1 \cdot 0{,}01$.

Aufgabe 9.15

Diese Aufgabe muss in R gelöst werden.

Aufgabe 9.16

a) H_0 wird abgelehnt, da $1{,}6 \notin [0{,}5;\ 1{,}5]$.

b) H_0 wird nicht abgelehnt, da $0{,}6 \in [0{,}5;\ 1{,}5]$

c) Für größeres α wird das $1-\alpha$ Konfidenzintervall kleiner. D.h. 0,5 liegt nicht im Konfidenzintervall für $\alpha = 10\%$ und H_0 wird abgelehnt.

Aufgabe 9.17

a) $\hat{\beta} = (X^T X)^{-1} X^T y = \begin{pmatrix} 0{,}25 \\ 4 \\ 2{,}5 \end{pmatrix}$

b) $\hat{U} = y - X\hat{\beta} = \begin{pmatrix} 0{,}75 \\ -0{,}75 \\ -0{,}75 \\ 0{,}75 \end{pmatrix}, \quad \sum_{t=1}^{4} \hat{U}_t = 0, \quad \hat{U}^T \hat{U} = \sum_{t=1}^{4} \hat{U}_t^2 = 2{,}25$

c) SST = 176,75 SSR = 174,5 SSE = 2,25
Die Streuungszerlegung gilt, da SSE + SSR = 174,5 + 2,25 = 176,75 = SST.

d) $R^2 = 0{,}9873$
98,73% der Varianz in y_t kann durch die Variablen x_t und z_t linear erklärt werden.

Aufgabe 9.18

Diese Aufgabe muss in R gelöst werden.

Aufgabe 9.19

a) $$\underbrace{\begin{pmatrix} 0 & 1 & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \end{pmatrix}}_{=H} \underbrace{\begin{pmatrix} \beta_0 \\ \beta_1 \\ \beta_2 \\ \beta_3 \end{pmatrix}}_{=\beta} = \underbrace{\begin{pmatrix} 0 \\ 0 \\ 0 \end{pmatrix}}_{h}$$

b) $$\underbrace{\begin{pmatrix} 1 & 1 & -1 & 0 \\ 0 & 0 & 0 & 1 \end{pmatrix}}_{=H} \underbrace{\begin{pmatrix} \beta_0 \\ \beta_1 \\ \beta_2 \\ \beta_3 \end{pmatrix}}_{=\beta} = \underbrace{\begin{pmatrix} 0 \\ 1 \end{pmatrix}}_{h}$$

c) $$\underbrace{\begin{pmatrix} 1 & 0 & 1 & -2 \end{pmatrix}}_{=H} \underbrace{\begin{pmatrix} \beta_0 \\ \beta_1 \\ \beta_2 \\ \beta_3 \end{pmatrix}}_{=\beta} = \underbrace{\begin{pmatrix} -1 \end{pmatrix}}_{h}$$

Die Darstellungen sind nicht eindeutig, denn wir können jede Restriktionsgleichung beliebig äquivalent umformen.

Aufgabe 9.20

Diese Aufgabe muss in R gelöst werden.

Aufgabe 9.21

a) $\beta_2 \approx -0{,}8$

b) $\hat{\beta}_0^{neu} = -0{,}2062$; $\hat{\beta}_1^{neu} = 1{,}3$; $\hat{\beta}_2^{neu} = -1{,}5$

Aufgabe 9.22

a) Eine Erhöhung des deutschen Preisniveaus P_t um 1% führt im Durchschnitt c.p. zu einer Verringerung des Wechselkurses um 1,42 %.

b) $WK_{t+1} = 1{,}18 \cdot (1 + (-1{,}42 \cdot 0{,}05)) = 1{,}0962$

c) $\underbrace{\begin{pmatrix}1 & 0 & 0\\ 0 & 1 & 0\\ 0 & 0 & 1\end{pmatrix}}_{=H}\underbrace{\begin{pmatrix}\beta_0\\ \beta_1\\ \beta_2\end{pmatrix}}_{=\beta} = \underbrace{\begin{pmatrix}0\\ -1\\ 1\end{pmatrix}}_{h}$

d) F Test:

$$d = \begin{pmatrix}0{,}59\\ -0{,}42\\ 0{,}28\end{pmatrix}, \quad H = \begin{pmatrix}1 & 0 & 0\\ 0 & 1 & 0\\ 0 & 0 & 1\end{pmatrix} \Rightarrow \left(H\left(X^TX\right)^{-1}H^T\right)^{-1} = X^TX$$

$$F = \frac{1}{0{,}17^2 \times 3}\begin{pmatrix}0{,}59\\ -0{,}42\\ 0{,}28\end{pmatrix}^T \begin{pmatrix}0{,}08 & -0{,}02 & -0{,}03\\ -0{,}02 & 0{,}37 & 0{,}40\\ -0{,}03 & 0{,}40 & 0{,}77\end{pmatrix}\begin{pmatrix}0{,}59\\ -0{,}42\\ 0{,}28\end{pmatrix} = 0{,}6852$$

Da $F = 0{,}6852 < F_{3;294;0{,}95} \approx 2{,}7$ kann H_0 zum Signifikanzniveau $\alpha = 0.05$ nicht abgelehnt werden. Die Kaufkraftparitätentheorie wird nicht verworfen.

Aufgabe 9.23

a) Chow-Test: $F = 46{,}4269 \quad F_{4;3071;0{,}95} \approx 2{,}38$
Mit $46{,}4269 > 2{,}38$ wird H_0 abgelehnt. Es gibt einen signifikanten Unterschied für Männer und Frauen in der Regressionsbeziehung.

b) β_3: Ein zusätzliches Lebensjahr führt ceteris paribus im Durchschnitt zu einem um 1,4% höherem Einkommen.
β_4: Eine Frau hat ceteris paribus im Durchschnitt ein um 25,7% geringeres Einkommen als ein Mann.

$H_0 : \beta_4 = 0$ wird wegen $|T| = 12{,}85 > 1{,}96 = t_{3074;0{,}975}$ abgelehnt.

c) β_3: Ein zusätzliches Lebensjahr für einen Mann führt ceteris paribus im Durchschnitt zu einem um 1,6% höherem Einkommen.
β_4: Eine Hypothetische Frau im Alter von 0 Jahren hat ceteris paribus ein um 8,7% geringeres Einkommen als ein Mann im Alter von 0 Jahren. (keine nützliche/sinnvolle Interpretation)
β_5: Der Einkommenszuwachs durch ein zusätzliches Lebensjahr fällt bei Frauen ceteris paribus im Durchschnitt um 0,3% niedriger aus als bei Männern.
$H_0 : \beta_4 + 30\beta_5 = 0$ wird wegen $|T| = 5{,}9 > 1{,}96 = t_{3073;0{,}975}$ abgelehnt.
Signifikanter Einkommensunterschied bei 30-jährigen Frauen und Männern.

d) F_i und M_i sind perfekt negativ korreliert. Damit ist $rg(X) = k$ verletzt und der OLS-Schätzer nicht bestimmbar (Multikollinearität). Lösung: F_i oder M_i aus dem Modell entfernen.

Aufgabe 9.24

a) Chow-Test: Mit $F = 11{,}70 > 2{,}7 \approx F_{3;183;0{,}95}$ wird H_0 abgelehnt.
Zum Jahrtausendwechsel liegt ein signifikanter Strukturbruch in den Regressionskoeffizienten vor.

b) Goldfeld-Quandt-Test, Vergleich der Regime 1993(M01) - 1999(M12) und 2000(M01) - 2008(M09):

$GQ = \frac{200/81}{201{,}4/102} = 1{,}2505 \qquad F_{krit} = F_{81;102;0{,}95} \approx 1{,}41$

H_0 kann nicht abgelehnt werden und somit wird Heteroskedastizität nicht nachgewiesen.