Lösungen zur Vorlesung Statistik I + II

Humboldt-Universität zu Berlin

Wirtschaftswissenschaftliche Fakultät
Lehrstuhl für Statistik

8. April 2022

Inhaltsverzeichnis

1 Mathematik

Lösung 1-1: *Vereinfachung*

a) 59

b) 0

c) 111

d) $abd - abde + ab^2f + cd - cde + cfb$

e) $abd + cd - de - bdf$

f) 0

g) $2^3 \cdot 3^2 \cdot 2^{-1} \cdot 3^{-3} = 2^2 \cdot 3^{-1} = \frac{4}{3}$

h) $\left(\frac{2}{\cancel{3}}\right)^4 \cdot \cancel{2}^4 \cdot \cancel{3}^4 \cdot \left(\frac{1}{\cancel{2}}\right)^{\cancel{4}} = 2^4 = 16$

i) $(3 \cdot 4 \cdot 3)^{0.5} = (3 \cdot 4 \cdot 3)^{\frac{1}{2}} = \sqrt{3 \cdot 4 \cdot 3} = \sqrt{36} = 6$

j) $(3 \cdot 4 \cdot 3)^{-0.5} = \frac{1}{\sqrt{3 \cdot 4 \cdot 3}} = \frac{1}{\sqrt{36}} = \frac{1}{6}$

k) $a^2b^{-3}a^4c^{-2}b^{-1}c = \frac{a^2 \cdot a^4 \cdot \cancel{c}}{b^3 \cdot c^{\cancel{2}} \cdot b} = \frac{a^6}{b^4c} = a^6b^{-4}c^{-1}$

l) $(a + b + c)^0 = 1$

Lösung 1-2: *Potenzen und Wurzeln*

a) $\dfrac{3a+b}{a^2-b^2}+\dfrac{2a}{a-b}\cdot\dfrac{a+b}{a+b}=\dfrac{3a+b+2a^2+2ab}{a^2-b^2}$

b) $\dfrac{6xy^2-12xy^3}{9x^3y-33x^2y^2}=\dfrac{1-2y}{3x-11y}\cdot\dfrac{6xy^2}{3x^2y}=\dfrac{1-2y}{3x-11y}\cdot\dfrac{2y}{x}$

c) $\dfrac{x^2-2xy}{x^2-4y^2}=\dfrac{x\cancel{(x-2y)}}{\cancel{(x-2y)}(x+2y)}=\dfrac{x}{x+2y}$

d)

$$\frac{\frac{x}{y}-\frac{y}{x}}{\frac{1}{x}+\frac{1}{y}}=\frac{\frac{x}{y}\cdot\frac{x}{x}-\frac{y}{x}\cdot\frac{y}{y}}{\frac{1}{x}\cdot\frac{y}{y}+\frac{1}{y}\cdot\frac{x}{x}}=\frac{\frac{x^2-y^2}{yx}}{\frac{y+x}{xy}}=\frac{x^2-y^2}{\cancel{xy}}\cdot\frac{\cancel{yx}}{y+x}$$
$$=\frac{(x-y)\cancel{(x+y)}}{\cancel{(x+y)}}=x-y$$

Lösung 1-3: *Binomische Formeln*

a)

$$\begin{aligned}(3xy-2)^2+(3xy+2)^2&=9x^2y^2-\cancel{12xy}+4+9x^2y^2+\cancel{12xy}+4\\&=18x^2y^2+8\end{aligned}$$

b) $102^2=(100+2)^2=10404$

Lösung 1-4: *Wurzel*

A) Fasse zusammen bzw. kürze:

a) $\left(\dfrac{1}{x^{-2}}\right)^{-3}=x^{-6}$

b) $\left(\dfrac{x^4y^{-2}z^3}{a^3b}\right)^2=\dfrac{x^8y^{-4}z^6}{a^6b^2}$

B) Schreibe unter eine gemeinsame Wurzel:

a) $x\sqrt[3]{y}=\sqrt[3]{x^3y}$

b) $\sqrt[4]{x}\sqrt[6]{y}=\sqrt[12]{x^3y^2}$

c) $\dfrac{\sqrt{xy^2}}{x\sqrt{y}}=\sqrt{\dfrac{xy^2}{x^2y}}=\sqrt{\dfrac{y}{x}}$

C) Schreibe als Wurzel:

a) $x^{0,5}=x^{\frac{1}{2}}=\sqrt{x}$

b) $x^{\frac{4}{5}}=\sqrt[5]{x^4}$

c) $x^{0,1}=x^{\frac{1}{10}}=\sqrt[10]{x}$

d) $x^{-\frac{2}{3}}=\dfrac{1}{x^{\frac{2}{3}}}=\dfrac{1}{\sqrt[3]{x^2}}$

Lösung 1-5: *Logarithmus*

A) a) $2^3=8\Leftrightarrow\log_2 8=3$

b) $a^{0,25}=c\Leftrightarrow\log_a c=0,25$

c) $5^x=5\Leftrightarrow x=\log_5 5=1$

B) a) $7\log x-2\log x=0\Leftrightarrow 5\log x=0\Leftrightarrow\log x=0\Leftrightarrow x=1$

b) Hier gibt es keine Lösung: $\dfrac{2\log x-2\log x}{\log 10}=5\Leftrightarrow\dfrac{0}{\log 10}=5$

c) $\log x=3\Leftrightarrow x=e^3\approx 20,0855$

C) a) $\log\left(\prod_{i=1}^{n}a_ib_i\right)=\sum_{i=1}^{n}\log a_i+\sum_{i=1}^{n}\log b_i$

b) $\log\left(\prod_{i=1}^{n} a_i^{b_i}\right) = \sum_{i=1}^{n} b_i \log a_i$

Lösung 1-6: *Produktzeichen*

a) $\prod_{i=1}^{5}(i-3) = (1-3)\cdot(2-3)\cdot(3-3)\cdot(4-3)\cdot(5-3) = 0$

b) $\prod_{i=1}^{5}(-1)^i = (-1)\cdot 1\cdot(-1)\cdot 1\cdot(-1) = -1$

c) $(a+b)\cdot(a+b)^2 = (a+b)^3 \rightarrow$ b)

Lösung 1-7: *Summenzeichen*

A) a) Falsch, richtig wäre $1+2+3+4+5+6+7+8+9+10$

b) Richtig

c) Richtig

d) Richtig

B) a) Falsch, richtig wäre $\sum_{i=2}^{n+1} a_{i-1}$

b) Falsch, kann nicht weiter zerlegt werden

c) Falsch, richtig wäre $2a_n + 1$

d) Richtig

C) a)

$$\begin{aligned}\sum_{i=1}^{n} x_i^2 &= \sum_{i=1}^{n} ({x_i}^2) \\ &= x_1^2 + x_2^2 + x_3^2 + ... + x_{n-1}^2 + x_n^2\end{aligned}$$

b)

$$\begin{aligned}\sum_{i=1}^{n} a_i^{i-1} &= \sum_{i=1}^{n} (a_i^{(i-1)}) \\ &= a_1^0 + a_2^1 + a_3^2 + ... + a_{n-1}^{n-2} + a_n^{n-1}\end{aligned}$$

c)

$$\begin{aligned}\sum_{i=1}^{n} x_i - 1 &= \left(\sum_{i=1}^{n} (x_i)\right) - 1 \\ &= x_1 + x_2 + x_3 + ... + x_{n-1} + x_n - 1\end{aligned}$$

d)

$$\begin{aligned}\sum_{i=1}^{n}(a_i+3) &= \sum_{i=1}^{n} (a_i+3) \\ &= a_1 + 3 + a_2 + 3 + a_3 + 3 + ... + a_{n-1} + 3 + a_n + 3 \\ &= a_1 + a_2 + a_3 + ... + a_{n-1} + a_n + 3n\end{aligned}$$

e)

$$\begin{aligned}\sum_{i=1}^{n} |x_i - y_i| &= \sum_{i=1}^{n} |x_i - y_i| \\ &= |x_1 - y_1| + |x_2 - y_2| + ... + |x_n - y_n|\end{aligned}$$

f)

$$\begin{aligned}\frac{1}{n}\sum_{i=1}^{n} ax_i &= \frac{1}{n}\sum_{i=1}^{n} (a\cdot x_i) \\ &= \frac{1}{n}(ax_1 + ax_2 + ax_3 + ... + ax_{n-1} + ax_n) \\ &= \frac{a}{n}(x_1 + x_2 + x_3 + ... + x_{n-1} + x_n) \\ &= \frac{a}{n}\sum_{i=1}^{n} x_i\end{aligned}$$

g)

$$\begin{aligned}\sum_{i=1}^{n} x_i y_i &= \sum_{i=1}^{n} (x_i y_i) \\ &= x_1y_1 + x_2y_2 + x_3y_3 + ... + x_{n-1}y_{n-1} + x_ny_n\end{aligned}$$

h)

$$\sum_{i=1}^{n} x_i \sum_{j=1}^{m} y_j = \left(\sum_{i=1}^{n} (x_i)\right) \cdot \left(\sum_{j=1}^{m} (y_j)\right)$$
$$= (x_1 + x_2 + \dots + x_{n-1} + x_n) \cdot (y_1 + y_2 + \dots + y_{m-1} + y_m)$$

i)

$$\sum_{i=1}^{n}\sum_{j=1}^{m} x_i y_j = \sum_{i=1}^{n}\sum_{j=1}^{m} x_i y_j$$
$$= (x_1 y_1 + \dots + x_1 y_m) + \dots + (x_n y_1 + \dots + x_n y_m)$$

j)

$$\left(\sum_{i=1}^{n} x_i\right)^2 = \left(\sum_{i=1}^{n} (x_i)\right) \cdot \left(\sum_{j=1}^{n} (x_j)\right)$$
$$= (x_1 + x_2 + \dots + x_{n-1} + x_n) \cdot (x_1 + x_2 + \dots + x_{n-1} + x_n)$$
$$= \sum_{i=1}^{n} x_i^2 + 2 \sum_{\substack{i,j=1 \\ i<j}}^{n} x_i x_j$$

D)

a) $\sum_{i=1}^{n}(x_i^2 + y_i^2) = \sum_{i=1}^{n} x_i^2 + \sum_{i=1}^{n} y_i^2$

b) $\sum_{i=1}^{n}(x_i + y_i)^2 = \sum_{i=1}^{n} x_i^2 + 2\sum_{i=1}^{n} x_i y_i + \sum_{i=1}^{n} y_i^2$

c) $\sum_{i=1}^{n}(x_i - y_i)^2 = \sum_{i=1}^{n} x_i^2 - 2\sum_{i=1}^{n} x_i y_i + \sum_{i=1}^{n} y_i^2$

d) $\sum_{i=1}^{n}(x_i - \mu)^2 = \sum_{i=1}^{n} x_i^2 - 2\sum_{i=1}^{n} x_i \mu + \sum_{i=1}^{n} \mu^2 = \sum_{i=1}^{n} x_i^2 - 2\mu \sum_{i=1}^{n} x_i + n\mu^2$

e) $\sum_{i=1}^{n}(x_i + y_i)(x_i - y_i) = \sum_{i=1}^{n} x_i^2 - \sum_{i=1}^{n} y_i^2$

f) $\sum_{i=1}^{n}(ax_i + bx_i + ay_i - by_i) = (a+b)\sum_{i=1}^{n} x_i + (a-b)\sum_{i=1}^{n} y_i$

E)

a) $a_1 + 2a_2 + 4a_3 + 8a_4 + \dots + 2^n a_n = \sum_{i=1}^{n} 2^i a_i$

b) $x_1 + y_1 + x_2 + y_2 + \dots + x_n + y_n = \sum_{i=1}^{n} (x_i + y_i)$

c) $\frac{1}{x_1} + \frac{y_1}{x_2} + \frac{y_2}{x_3} + \dots + \frac{y_{n-1}}{x_n} = \frac{1}{x_1} + \sum_{i=2}^{n} \frac{y_{i-1}}{x_i}$

d) $x_1 + \sqrt{x_2} + \sqrt[3]{x_3} + \sqrt[4]{x_4} + + \dots + \sqrt[n]{x_n} = \sum_{i=1}^{n} \sqrt[i]{x_i} = \sum_{i=1}^{n} x_i^{\frac{1}{i}}$

e) $a_1^2 + 2a_1b_1 + b_1^2 + a_2^2 + 2a_2b_2 + b_2^2 + \dots + a_n^2 + 2a_nb_n + b_n^2 = \sum_{i=1}^{n} (a_i + b_i)^2$

f) $a_1^2 - 2a_1a_2 + 2a_2^2 - 2a_2a_3 + 2a_3^2 - 2a_3a_4 + \dots + a_n^2$
$= \sum_{i=2}^{n} (a_i - a_{i-1})^2 = \sum_{i=1}^{n-1} (a_i - a_{i+1})^2$

g) $\ln x_1 + \ln x_2 + \dots + \ln x_n = \sum_{i=1}^{n} \log x_i = \log\left(\prod_{i=1}^{n} x_i\right)$

Lösung 1-8: *Binomialkoeffizient*

a) $\binom{24}{21} = \frac{24!}{(24-21)! \cdot 21!} = \frac{\cancel{24!}}{3! \cdot \cancel{21!}} = \frac{22 \cdot 23 \cdot 24}{1 \cdot 2 \cdot 3} = 2024$

b) $\binom{6}{0} = \frac{6!}{0! \cdot 6!} = 1$

c) $\binom{10}{2} = \frac{10!}{8! \cdot 2!} = 45$

Lösung 1-9: *Funktion*

A) a) $\mathbb{D}(f) = \mathbb{R}, \quad \mathbb{W}(f) = \{y | y \in \mathbb{R} \wedge y > 2\}$

b) $\mathbb{D}(f) = \mathbb{R}, \quad \mathbb{W}(f) = \{y | y \in \mathbb{R} \wedge 0 < y \leq 1\}$

B) a) $$\lim_{x\to -1} \frac{x^2+2x+1}{x^2-x-2} = \lim_{x\to -1} \frac{\cancel{(x+1)}(x+1)}{\cancel{(x+1)}(x-2)} = \lim_{x\to -1} \frac{x+1}{x-2} = 0$$

b) $$\lim_{x\to 3} \frac{x^2-9}{x^2-2x-3} = \lim_{x\to 3} \frac{(x+3)\cancel{(x-3)}}{\cancel{(x-3)}(x+1)} = \lim_{x\to 3} \frac{x+3}{x+1} = \frac{3}{2}$$

c) $\lim\limits_{x\to 0} x^{-1} \log x \Rightarrow$ Grenzwert existiert nicht!

Lösung 1-10: *Ableitung*

A) a) Summenregel: $f'(x) = 16x + e^x$

b) Kettenregel: $f'(x) = e^{x^3+1} \cdot (3x^2)$

c) Produktregel: $f'(x) = 3 \cdot \log(x) + 3x \cdot \frac{1}{x} = 3\log(x) + 3$

d) Quotientenregel:

$$f(x) = \frac{(x+1)^2}{1-x^2} = \frac{(1+x)^2}{1-x^2} = \frac{(1+x)^{\cancel{2}}}{(1-x)\cancel{(1+x)}} = \frac{1+x}{1-x}$$

$$f'(x) = \frac{(1-x)(+1)-(1+x)(-1)}{(1-x)^2} = \frac{1-x+1+x}{(1-x)^2} = \frac{2}{(1-x)^2}$$

e) $f'(x) = e^{x^2} + xe^{x^2} \cdot 2x = e^{x^2}(1+2x^2)$

f) Kettenregel:

$$f'(x) = \frac{1}{2}\left(x + x^{\frac{1}{2}}\right)^{-\frac{1}{2}} \cdot \left(1 + \frac{1}{2}x^{-\frac{1}{2}}\right) = \frac{1}{2\sqrt{x+\sqrt{x}}} \cdot \left(1 + \frac{1}{2\sqrt{x}}\right)$$

g) Quotientenregel:

$$\begin{aligned}
f'(x) &= \frac{(e^x+e^{-x})(e^x-e^{-x}\cdot(-1)) - (e^x-e^{-x})(e^x+e^{-x}\cdot(-1))}{(e^x+e^{-x})^2} \\
&= \frac{(e^x+e^{-x})(e^x+e^{-x}) - (e^x-e^{-x})(e^x-e^{-x})}{(e^x+e^{-x})^2} \\
&= \frac{(e^{2x}+e^0+e^{-2x}+e^0) - (e^{2x}-e^0+e^{-2x}-e^0)}{(e^x+e^{-x})^2} \\
&= \frac{e^{2x}+e^{-2x}+2-e^{2x}+2-e^{-2x}}{(e^x+e^{-x})^2} = \frac{4}{(e^x+e^{-x})^2}
\end{aligned}$$

h) $$f(x) = \frac{(x+1)^{\frac{2}{5}}}{(x+1)^{-\frac{3}{5}}} = (x+1)^{\frac{2}{5}} \cdot (x+1)^{\frac{3}{5}} = (x+1)^{\frac{2}{5}+\frac{3}{5}} = x+1$$

$f'(x) = 1$

B) a) $$\frac{\partial}{\partial a}\left(\sum_{i=1}^n x_i + a\sum_{i=1}^n x_i^2 - a^2\sum_{i=1}^n x_i y_i\right) = \sum_{i=1}^n x_i^2 - 2a\sum_{i=1}^n x_i y_i$$

b) $\frac{\partial}{\partial \alpha}\sum\limits_{i=1}^{n}(x_i^2-2x_i\alpha+\alpha^2)=\sum\limits_{i=1}^{n}(-2x_i+2\alpha)=-2\sum\limits_{i=1}^{n}(x_i-\alpha)$

c) $\frac{\partial}{\partial \alpha}\sum\limits_{i=1}^{n}(\alpha^2x_i^2+2\alpha\beta x_i+\beta^2)=\sum\limits_{i=1}^{n}(2\alpha x_i^2+2\beta x_i)$

d) $\frac{\partial}{\partial \beta}\sum\limits_{i=1}^{n}(\alpha^2x_i^2+2\alpha\beta x_i+\beta^2)=2\sum\limits_{i=1}^{n}(\alpha x_i+\beta)$

e) $\sum\limits_{i=1}^{n}\frac{1}{i!}\frac{\partial}{\partial x}x^i=\sum\limits_{i=1}^{n}\frac{i}{i!}x^{i-1}=\sum\limits_{i=1}^{n}\frac{1}{(i-1)!}x^{i-1}=\sum\limits_{i=0}^{n-1}\frac{1}{i!}x^i$

f) $\frac{\partial}{\partial x}\sum\limits_{i=1}^{n}i\log x=\left(\sum\limits_{i=1}^{n}i\right)\frac{\partial}{\partial x}\log x=\frac{n(n+1)}{2x}$

Lösung 1-11: *Integration*

A) a) $\int\limits_0^5 x^4dx=\frac{1}{5}x^5\Big|_0^5=\frac{1}{5}5^5-\frac{1}{5}0^5=625$

b)

$$\int\limits_0^1 e^x\cdot(e^{-x}-1)\,dx=\int\limits_0^1(1-e^x)\,dx=x-e^x|_0^1$$
$$=(1-e^1)-(0-e^0)=2-e\approx-0.72$$

c)

$$\int\limits_e^{e^2}\left(\frac{1}{x}+1\right)dx=\log x+x|_e^{e^2}$$
$$=\left(\log(e^2)+e^2\right)-(\log(e)+e)=2+e^2-1-e$$
$$=1+e^2-e\approx 5.67$$

B) a) $f(x)=e^{\ln x}=x\Rightarrow F(x)=\frac{1}{2}x^2+C$

b) $F(x)=-\frac{1}{2}x^{-2}+\frac{4}{6}x^6+\frac{1}{8}x^8$

c) $F(x)=\frac{1}{3}e^{3x}$

2 Deskriptive Statistik

Lösung 2-1: *Berliner Bühnen*

a) Gesamtheit: Menge der Berliner Bühnen im Jahre 2022
Einheit: die einzelne Bühne (z.B. Deutsche Oper, Deutsches Theater, Komische Oper, Schauspielhaus, Hansa–Theater, Schiller–Theater, Theater am Kurfürstendamm, Volksbühne, Theater des Westens).

b)
- Identifikationsmerkmale:
  Ort: festgelegt, Berlin.
  Zeit: festgelegt, 2022.
  Sachlich: Bühne, festgelegt
- Erhebungsmerkmale:
  Bühne (festgelegt)
  Höhe der Einnahmen, Höhe der Ausgaben (variabel)

c) Anzahl der Schauspieler, Anzahl der Besucher pro Jahr, Art der Bühne, Anzahl der Gastspiele usw.

Grundbegriffe-Berliner Buehnen.mp4

Lösung 2-2: *Merkmalsausprägungen*

- sachliches Kriterium: Beteiligung am Erwerbsleben - auf Erwerbstätige (festgelegt).
- örtliches Kriterium: Großstadt K in der Bundesrepublik Deutschland.
- zeitliches Kriterium: März 2022.

Lösung 2-3: *Skalierung*

nominalskaliert	Geschlecht, Postleitzahl, abonnierten Zeitungen, Nationalität , Freizeitbeschäftigung, Rückennummern von Fußsballspielern, Familienstand, erlernter Beruf, Todesursache, Studienfach, Augenfarbe, Wohnsitz, Telefonnummer, Rechtsform einer Unternehmung , Finanzierung des Studiums
ordinalskaliert	Schulnote, Betriebsgrössenklasse, Militärdienstgrad, Windstärke, Schwierigkeitsgrad einer Klettertour, Tarifklassen bei der Kfz-Haftpflicht, Güteklasse, Handelsklasse bei Obst, Aggressivität, Intelligenz, sozialer Status
metrisch skaliert/ Intervallskala	Temperatur in Celsius, Normabweichung, Geburtsjahrgang, Breitengrade der Erde
metrisch skaliert/ Verhältnisskala	Körpergröße, Länge eines Werkstückes, Wahlergebnis einer Partei, Fahrpreise, Geschwindigkeit, Kraftstoffverbrauch eines PKW auf 100 km, Preis einer Ware, Lebensalter, Einkommen, Jahresumsatz, Grundstücksgröße, Produktionsdauer
metrisch skaliert/ Absolutskala	Kinderzahl, Bücherbestand einer Bibliothek, Seitenzahl eines Buches, Semesterzahl, Klausurpunkte

Lösung 2-4: *Familienstand*
Grundgesamtheit:
Menge der Studierenden des Fachbereiches Wirtschaftswissenschaften der Humboldt-Universität zu Berlin im Sommersemester 2012.
Identifikationskriterien:
örtlich – HUB
zeitlich – Sommersemester 2022
sachlich - Studierende, Wirtschaftswissenschaften
statistische Einheit:
Studierender des Fachbereiches Wirtschaftswissenschaften
der Humboldt-Universität zu Berlin im Sommersemester 2022.
untersuchtes statistisches Merkmal:
Familienstand.
Merkmalsausprägungen:
ledig, verheiratet, geschieden, verwitwet
Lösung 2-5: *Wählerverhalten*

sachlich: wahlberechtigte Personen (Deutsche im Sinne von Artikel 116 Abs. 1 des Grundgesetzes, 18. Lebensjahr vollendet, ...)

räumlich: Personen welche seit mindestens drei Monaten Ihre Wohnung in der Bundesrepublik Deutschland haben oder sich sonst gewöhnlich dort aufhalten

zeitlich: Zeitpunkt der Bundestagswahl

Lösung 2-6: *Versicherungsunternehmen*

a) **Statistische Einheit:** jeder Kfz-Haftpflichtversicherter, der bei diesem Versicherungsunternehmen in Vertrag steht.
Grundgesamtheit: Menge der Kfz-Haftpflichtversicherten, die bei diesem Versicherungsunternehmen in Vertrag stehen.

b) **Alter:** sachlich, verhältnisskaliert, stetig.
Geschlecht: sachlich, nominalskaliert, nicht häufbar.
Beruf: sachlich, nominalskaliert, häufbar.
Wohnort: örtlich, nominalskaliert, häufbar.
Dauer des Versicherungsvertrages: sachlich, verhältnisskaliert, stetig.
Anzahl der bisher eingetretenen Schadensfälle: sachlich, absolutskaliert, diskret.
Gesamtschadenshöhe: sachlich, verhältnisskaliert, diskret (bzw. quasistetig).

Grundbegriffe-Versicherungsunternehmen.mp4

2.1 Univariate Statistik

Lösung 2-7: *Eine Befragung von Studierenden - Teil I*

a)
- Grundgesamtheit: Die Menge aller Studierenden der deutschen Hochschule im Juni 2022
- Stichprobe: Die Menge der 25 befragten Studierenden der deutschen Hochschule im Juni 2022
- Statistische Einheiten: Studierende der deutschen Hochschule im Juni 2022
- Identifikationskriterien:
- örtlich: Deutschland
- zeitlich: Juni 2022
- sachlich: immatrikulierte Person an der Hochschule

b) X: "Studiengang"; nominalskaliert

Studiengang x_j	$h(x_j)$	$f(x_j)$
VWL	5	0,2
BWL	10	0,4
Polit.Wiss.	5	0,2
Sozialwiss.	5	0,2
Summe	25	1,0

Graphische Darstellung als Säulendiagramm, Kreisdiagramm oder Flächendiagramm. Siehe dazu Folien in der Vorlesung "Deskriptive Statistik", beispielsweise:

- Grafische Darstellung der Häufigkeit II

c) Y: "Anzahl der Geschwister"; kardinalskaliert, diskret

Anzahl der Geschwister y_j	$h(y_j)$	$f(y_j)$	$F(y)$
0	8	0,32	0,32
1	11	0,44	0,76
2	4	0,16	0,92
3	2	0,08	1,00
Summe	25	1,00	

Graphische Darstellung der Häufigkeitsverteilung als Stabdiagramm, der empirischen Verteilungsfunktion als Treppenfunktion

d)

- 23 Studierende haben höchstens 2 Geschwister.
- Viermal 2 Geschwister + zweimal 3 Geschwister = 6 Personen mit mindestens 2 Geschwistern, d.h.
  $\frac{6}{25} = 0,24 = 24\%$ haben mindestens zwei Geschwister.
  $\frac{15}{25} = 0,6 = 60\%$ haben ein oder zwei Geschwister.

e) Z: "Einkommen"; kardinalskaliert (metrisch), quasi-stetig

Einkommen $z_j^u \le Z < z_j^o$	$h(z_j)$	$f(z_j)$	$F(z)$	$f_K(z_j)$
600 – 650	6	0,24	0,24	0,0048
650 – 700	6	0,24	0,48	0,0048
700 – 900	8	0,32	0,80	0,0016
900 – 1200	2	0,08	0,88	0,0003
1200 – 1450	3	0,12	1,00	0,0005
Summe	25	1,00		

Graphische Darstellung der Häufigkeitsverteilung als Histogramm, der empirischen Verteilungsfunktion als stückweise linearer Funktion

f) – $P[\{750 \le Z \le 1300\}] = F(1300) - F(750)$

$$\begin{aligned} F(1300) &= 0,88 + \frac{1300-1200}{1450-1200} \cdot 0,12 = 0,928 \\ F(750) &= 0,48 + \frac{750-700}{900-700} \cdot 0,32 = 0,56 \\ P[\{750 \le Z \le 1300\}] &= F(1300) - F(750) = 0,928 - 0,56 \end{aligned}$$

– $P[\{Z > 800\}] = 1 - F(800)$

$$\begin{aligned} &= 1 - [0,24 + 0,24 + \frac{(800-700)}{(900-700)} \cdot 0,32] \\ &= 1 - [0,48 + \frac{100}{200} \cdot 0,32] \\ &= 0,36 \end{aligned}$$

– $0,5 = 0,48 + \frac{z-700}{900-700} \cdot 0,32 \rightarrow z = 712,5\text{EUR}$

– $0,12 + 0,08 = 0,2 \Rightarrow 20\% \Rightarrow 900EUR$

Grundbegriffe-Befragung von Studierenden Teil I.mp4

Lösung 2-8: *Zugfolge - Teil I*

a) X: "Zugfolgeabstand"; kardinalskaliert, stetig

b) und c) Zugfolgeabstand (Von ... bis unter ...)

x_j	$h(x_j)$	$f(x_j)$	$F(x_j^o)$	$f_K(x_j)$
0 – 30	14	0,3784	0,3784	0,0126
30 – 60	17	0,4595	0,8379	0,0153
60 – 90	5	0,1351	0,9730	0,0045
90 – 120	1	0,0270	1,0000	0,0009
Summe	37	1,0000		

Häufigkeitsverteilung: Histogramm
empirische Verteilungsfunktion: stückweise lineare Funktion

Grundbegriffe-Zugfolge Teil I.mp4

Lösung 2-9: *Bibliotheken - Teil I*

a) alle Merkmale sind kardinalskaliert
   - Öffnungszeiten und Ausleihzeiten: stetig
   - Etat für Neuerwerb: quasi-stetig
   - Planstellen: diskret

b) X: "Öffnungszeiten"

Öffnungszeiten x_j	$h(x_j)$	$f(x_j)$	$F(x)$	$f_K(x_j)$
40 – 50	2	0,0323	0,0323	0,00323
50 – 60	19	0,3065	0,3388	0,03065
60 – 70	25	0,4032	0,7420	0,04032
70 – 80	11	0,1774	0,9194	0,01774
80 – 90	4	0,0645	0,9839	0,00645
90 – 115	1	0,0161	1,0000	0,00064
Summe	62	1,0000		

Häufigkeitsverteilung: Histogramm
empirische Verteilungsfunktion: stückweise lineare Funktion

Y: "Ausleihzeiten"

Ausleihzeiten y_j	$h(y_j)$	$f(y_j)$	$F(y)$	$f_K(y_j)$
25 – 30	5	0,0806	0,0806	0,01612
30 – 40	14	0,2258	0,3064	0,02258
40 – 50	19	0,3065	0,6129	0,03065
50 – 60	12	0,1935	0,8064	0,01935
60 – 70	6	0,0968	0,9032	0,00968
70 – 80	3	0,0484	0,9516	0,00484
80 – 100	3	0,0484	1,0000	0,00242
Summe	62	1,0000		

Häufigkeitsverteilung: Histogramm
empirische Verteilungsfunktion: stückweise lineare Funktion

Z: "Etat für Neuerwerb"

Etat für Neuerwerb z_j	$h(z_j)$	$f(z_j)$	$F(z)$	$f_K(z_j)$
0 – 1	1	0,0161	0,0161	0,0161
1 – 2	14	0,2258	0,2419	0,2258
2 – 3	10	0,1613	0,4032	0,1613
3 – 4	17	0,2742	0,6774	0,2742
4 – 5	13	0,2097	0,8871	0,2097
5 – 6	6	0,0968	0,9839	0,0968
6 – 7	1	0,0161	1,0000	0,0161
Summe	62	1,0000		

Häufigkeitsverteilung: Histogramm
empirische Verteilungsfunktion: stückweise lineare Funktion

W: "Planstellen"

Planstellen w_j	$h(w_j)$	$f(w_j)$	$F(w)$	$f_K(w_j)$
30 – 70	6	0,0968	0,0968	0,00242
70 – 80	12	0,1935	0,2903	0,01935
80 – 100	10	0,1613	0,4516	0,00807
100 – 150	19	0,3065	0,7581	0,00613
150 – 200	12	0,1935	0,9516	0,00387
200 – 270	3	0,0484	1,0000	0,00069
Summe	62	1,0000		

Häufigkeitsverteilung: Histogramm
empirische Verteilungsfunktion: stückweise lineare Funktion

c)
   - 0,7872
   - 0,1452
   - 0,3064
   - 0,5484
   - 77 Planstellen
   - 76,088 Std./Wo.

Lösung 2-10: *Erdbeerplantage - Teil I*

a) X: "Anzahl der Sonnenstunden pro Tag in dieser Saison"; kardinalskaliert, stetig klassiert

b)

$x_j^u \leq x_j < x_j^o$	$h(x_j)$	$f(x_j)$	$F(x)$	$f_K(x_j)$
0 - 2	20	0,20	0,20	0,10
2 - 3	15	0,15	0,35	0,15
3 - 5	20	0,20	0,55	0,10
5 - 8	35	0,35	0,90	0,12
8 - 12	10	0,10	1,00	0,03
Summe	100	1,0000		

Häufigkeitsverteilung: Histogramm

c) empirische Verteilungsfunktion: stückweise lineare Funktion

d) an 55 Tagen

e) 3,5 Std./Tag

f) 47,5 %

Lösung 2-11: *Berliner Luftqualität*

a) X: "Stickstoffmonoxydgehalt der Berliner Luft im März 2022"; kardinalskaliert, stetig

b)

$x_j^u \leq X < x_j^o$	$h(x_j)$	$f(x_j)$	$F(x)$	$f_K(x_j)$
19,5 - 29,5	5	0,3333	0,3333	0,03333
29,5 - 34,5	4	0,2667	0,6000	0,05334
34,5 - 39,5	5	0,3333	0,9333	0,06666
39,5 - 44,5	1	0,0667	1,0000	0,01334
Summe	15	1,0000		

c) Häufigkeitsverteilung: Histogramm
empirische Verteilungsfunktion: stückweise lineare Funktion

d) 6 Tage

e) 37,5 mg/m^3

Lösung 2-12: *Zigaretten*

a) empirische Verteilungsfunktion,
Annahme: Gleichverteilung der Merkmalswerte innerhalb jeder Klasse

b) X: "Anzahl der gerauchten Zigaretten pro Tag"

$x_j^u \leq X < x_j^o$	$h(x_j)$	$f(x_j)$	$F(x)$	$f_K(x_j)$
0 - 5	10	0,05	0,05	0,01
5 - 10	40	0,20	0,25	0,04
10 - 20	90	0,45	0,70	0,045
20 - 40	60	0,30	1,00	0,015
Summe	200	1,0000		

c) 0,3

Lösung 2-13: *Eiskugelkonsum*

a) X: "Eiskonsum in Kugeln"; kardinalskaliert, diskret

b)

x_j	$h(x_j)$	$f(x_j)$	$F(x)$
1	40	0,20	0,20
3	50	0,25	0,45
4	70	0,35	0,80
5	10	0,05	0,85
6	30	0,15	1,00
Summe	200	1,00	

c) Häufigkeitsverteilung: Stabdiagramm
empirische Verteilungsfunktion: Treppenfunktion

d) $F(5) = 0,85$

e) 3 Kugeln

f) 4 Kugeln

Lösung 2-14: *WM–Berichterstattung*

a) X: "Fernsehkonsum (in Std.) während der letzten Fußball-WM" kardinalskaliert, diskret

b)

x_j	$h(x_j)$	$f(x_j)$	$F(x)$
0	1	0,05	0,05
1	2	0,10	0,15
2	8	0,40	0,55
3	4	0,20	0,75
4	5	0,25	1,00
Summe	20	1,00	

c) 1 Std.

d) 2 Std.

e) 3 Std.

Lösung 2-15: *Auswirkung der Regelstudienzeit*

a) X: "Semesterzahl"; kardinalskaliert, diskret

b) und c)

x_j	$h(x_j)$	$f(x_j)$	$F(x)$
10	20	0,10	0,10
11	20	0,10	0,20
12	80	0,40	0,60
13	40	0,20	0,80
14	30	0,15	0,95
15	10	0,05	1,00
Summe	200	1,00	

d) 10 Semester

e) 12 Semester

f) 13 Semester

g) ordinalskaliert

h) Säulendiagramm

i) Nein. Eine Verteilungsfunktion erfordert wegen

$$F(x) = \sum_{x_j \leq x} f(x)$$

eine Ordnungsrelation ("kleiner-gleich"-Beziehung) zwischen den Merkmalsausprägungen. Derartige Vergleichszeichen sind jedoch auf nominalskalierte Merkmale nicht anwendbar, da bei Nominalskalierung keine Anordnung möglich ist.

Lösung 2-16: *Tennis Turniere*

a) X: "Anzahl der pro Turnier bis zum Ausscheiden gespielten Runden"; kardinalskaliert, diskret

b)

x_j	$h(x_j)$	$f(x_j)$	$F(x)$
1	10	0,25	0,25
2	16	0,40	0,65
4	6	0,15	0,80
5	6	0,15	0,95
6	2	0,05	1,00
Summe	40	1,00	

c) Treppenfunktion

d) 65 %

e) 20 %

f) 30 Turnieren

g) 4. Runde

h) 2. Runde

i) $F(7) = 1$; d.h. jedes Turnier ging nur über 6 Runden; daher konnte B.B. niemals in eine 7. Runde gelangen, er war also in 100 % aller Turniere in Runde 7 (=x) ausgeschieden.

Lösung 2-17: *Eine Befragung von Studierenden - Teil II*

a) x_D = BWL, da BWL die höchste absolute Häufigkeit beinhaltet. Der Modus ist ein geeigneter Lageparameter, da es sich um nominalskalierte Variablenausprägungen handelt, die lediglich eine Verschiedenartigkeit zum Ausdruck bringen.

b) Die drei geeigneten Lageparameter entsprechen hier dem Modus, dem Median und dem Arithmetischen Mittel.

- Der Modus $y_D = 1$, da die Ausprägung eins die höchste absolute Häufigkeit besitzt.
- Der Median $y_{0,5} = 1$, da eins ungerade ist und daher die Formel

$$\begin{aligned} x_{0,5} &= x_{\left(\frac{n+1}{2}\right)} \\ &= x_{\left(\frac{25+1}{2}\right)} \\ &= x_{(13)} \end{aligned}$$

für den Median gilt.

- Das Arithmetische Mittel $\overline{y} = 1$, da sich mit der Formel

$$\begin{aligned} \overline{x} &= \frac{1}{n} \cdot \sum_{i=1}^{n} \cdot x_i \\ &= \frac{1}{25} \cdot (8 \cdot 0 + 11 \cdot 1 + 4 \cdot 2 + 2 \cdot 3) \\ &= 1 \text{ ergibt.} \end{aligned}$$

c) Das durchschnittliche Einkommen eines/einer Studierenden ergibt sich durch Berechnen des Arithmetischen Mittels wie folgt:

$$\begin{aligned} \overline{z} &= \frac{1}{n} \cdot \sum_{i=1}^{n} \cdot z_i \\ \overline{z} &= \frac{1}{25} \cdot (924 + ... + 640) \\ &= \frac{1}{25} \cdot 20200 \\ &= 808 \text{ EUR} \end{aligned}$$

d) Als aussagekräftige Lageparameter werden das Arithmetische Mittel und der Median berechnet.

- Der Median berechnet sich für klassierte Daten mit der folgenden Formel:

$$\begin{aligned} z_{0,5} &= x_j^u + \frac{(0,5 - F(z_j^u))}{f(z_j^m)} \cdot (z_j^o - z_j^u) \\ &= 700 + \frac{(0,5 - 0,48)}{0,32} \cdot 200 \\ &= 712,5 \text{ EUR} \end{aligned}$$

- Das Arithmetische Mittel wird hier für klassierte Daten mit der folgenden Formel berechnet, wobei x_j^m die Klassenmitte und n_j die Anzahl der Beobachtungen in Klasse j angibt:

$$\begin{aligned} \overline{z} &= \frac{1}{n} \cdot \sum_{j=1}^{k} \cdot z_j^m \cdot n_j \\ &= \frac{1}{25} \cdot (625 \cdot 6 + 675 \cdot 6 + 800 \cdot 8 + 1050 \cdot 2 + 1325 \cdot 3) \\ &= \frac{20275}{25} \\ &= 811 \text{ EUR.} \end{aligned}$$

Für die Berechnung der Quantile von klassierten Daten wird folgende Formel verwendet:

$$F(x_p) = p \Leftrightarrow x_p = x_j^u + \frac{(p - F(x_j^u))}{f(x_j)} \cdot (x_j^o - x_j^u)$$

- Die Quartile werden somit wie folgt für klassierte Daten berechnet:

$$\begin{aligned} z_{0,25} &= 650 + \frac{(0,25 - 0,24)}{0,24} \cdot 50 \\ &= 652,08 \text{ EUR} \\ z_{0,75} &= 700 + \frac{(0,75 - 0,48)}{0,32} \cdot 200 \\ &= 868,75 \text{ EUR} \end{aligned}$$

- Das 90%-Quantil ergibt somit:

$$\begin{aligned} z_{0,9} &= 1200 + \frac{(0,9 - 0,88)}{0,12} \cdot 250 \\ &= 1241,67 \text{ EUR} \end{aligned}$$

e) Minimum und Maximum sind hier wie folgt:

- Das Minimum ist hier $z_{(1)} = 616$, da es die kleinste Merkmalsausprägung des Merkmals Einkommen ist
- Das Maximum entspricht $z_{(25)} = 1440$, da es demgegenüber die größte Ausprägung des Merkmals Einkommen ist

Die Quantile ergeben sich wie folgt:

- $z_{0,25} = 660$, da $n \cdot p$ keine ganze Zahl ist, sondern hier $25 \cdot 0,25 = 6,25$ woraufhin die Ausprägung an der auf $n \cdot p$ folgenden Zahlenstelle gilt, das heißt 660 ist die Ausprägung an der siebten Stelle.
- $z_{0,5} = 700$ mit $n \cdot p = 25 \cdot 0,5 = 12,5 \Rightarrow 700$ ist das Ergebnis und steht als Ausprägung an dreizehnter Stelle

- $z_{0,75} = 850$ mit $n \cdot p = 25 \cdot 0,75 = 18,75 \Rightarrow$ Das entspricht der Ausprägung an neunzehnter Stelle

▶ Univariate Statistik-Befragung von Studierenden Teil II (14 min).mp4

Lösung 2-18: *Zugfolge - Teil II*

a) Wir wählen den Mittelwert, den Median und den Modus als Lageparameter für die klassierten Daten.
Für den Modus ist zunächst die Modalklasse zu berechnen; das bedeutet die Klasse mit der höchsten Häufigkeitsdichte. Da hier die Klassenbreite für alle Klassen 30 beträgt, genügt es, die Klasse mit der höchsten Häufigkeit zu wählen. Das entspricht der Klasse 30-60. Für die Feinbestimmung gilt folgende Formel für den Modus bei klassierten Daten:

$$\begin{aligned} x_D &= x_2^u + \frac{f_K(x_2) - f_K(x_{2-1})}{2f_K(x_2) - f_K(x_{2-1}) - f_K(x_{2+1})} \cdot (x_2^o - x_2^u) \\ &= 30 + \frac{\frac{0,4595}{30} - \frac{0,3784}{30}}{2 \cdot \frac{0,4595}{30} - \frac{0,3784}{30} - \frac{0,1351}{30}} \cdot (60 - 30) \\ &= 30 + \frac{1}{5} \cdot 30 = 36 \end{aligned}$$

Das Ergebnis entspricht hier 36 Minuten.
Für den Median von klassierten Daten gilt die folgende Formel:

$$\begin{aligned} x_{0,5} &= F(x_0,5) \Leftrightarrow \\ x_{0,5} &= x_j^u + \frac{(0,5 - F(x_j^u))}{f(x_j^m)} \cdot (x_j^o - x_j^u) \\ x_{0,5} &= 30 + \frac{(0,5 - 0,3784)}{0,4595} \cdot 30 \\ &= 37,94 \text{Minuten.} \end{aligned}$$

Der Mittelwert für klassierte Daten berechnet sich wie folgt, wobei x_j^m die Klassenmitte und n_j die Anzahl der Beobachtungen in Klasse j beschreibt:

$$\begin{aligned}\overline{x} &= \frac{1}{n} \cdot \sum_{j=1}^{k} \cdot x_j^m \cdot n_j \\ &= \frac{1}{37} \cdot (15 \cdot 14 + 45 \cdot 17 + 75 \cdot 5 + 105 \cdot 1) \\ &= 39{,}324 \text{ Minuten}\end{aligned}$$

b) Der Unterschied zum Ergebnis aus Teilaufgabe a) ergibt sich aufgrund von Informationsverlust durch die Klassierung der Daten.
Man berechnet das Arithmetische Mittel $\overline{x}$ von nicht-klassierten Daten wie folgt:

$$\begin{aligned}\overline{x} &= \frac{1}{n} \cdot \sum_{i=1}^{n} \cdot x_i \\ &= \frac{1}{37} \cdot \sum_{i=1}^{n} \cdot x_i \\ &= \frac{1}{37} \cdot 1440 \\ &= 38{,}92 \text{ Minuten}\end{aligned}$$

Univariate Statistik-Zugfolge Teil II.mp4

Lösung 2-19: *Gartenzwerg–Großhandel*

A: a) X: "Auftragshöhe (in EUR) pro Auftrag"; kardinalskaliert

b) Das Merkmal wurde an den einzelnen Aufträgen erhoben

c) Die durchschnittliche Auftragshöhe für klassierte Daten wird wie folgt berechnet:

$$\begin{aligned}\overline{x} &= \frac{1}{n} \cdot \sum_{j=1}^{k} \cdot x_j^m \cdot n_j, \text{ wobei } n = \sum_{j=1}^{k} \cdot u_j \\ \overline{x} &= \frac{1}{100} \cdot (10000 \cdot 15 + 35000 \cdot 30 + 100000 \cdot 45 + 225000 \cdot 10) \\ \overline{x} &= 79500 \text{ EUR/Auftrag}\end{aligned}$$

B: $\overline{x}$= a + b · $\overline{y}$; $\overline{x}$ = 0 + 1,5·60 000 = 90 000 EUR/Auftrag

C: a)

- Arithmetisches Mittel $\overline{x}$ nicht sinnvoll, da Ausreißer vorkommt
- Der Modus x_D nicht sinnvoll, da eventuell multimodale Häufigkeitsverteilung vorliegt
- Sinnvoll ist hier der Median mit dem Ergebnis $x_{0,5}$ = 12 000 EUR/Auftrag. Da n ungerade ist gilt:

$$\begin{aligned}x_{0,5} &= x_{\left(\frac{n+1}{2}\right)} \\ &= x_{\left(\frac{5+1}{2}\right)} \\ &= x_{(3)}\end{aligned}$$

b)

$$\begin{aligned}\overline{x} &= \frac{1}{n} \cdot \sum_{i=1}^{n} \cdot x_i \\ &= \frac{1}{5} \cdot 45000 \\ &= 9000 \text{ EUR/Auftrag}\end{aligned}$$

D: Die durchschnittliche Auftragshöhe für das gesamte Unternehmen beträgt:

$$\begin{aligned}\overline{x} &= \frac{1}{n} \cdot \sum_{i=1}^{n} \cdot x_i \\ &= \frac{1}{(95+5+100)} \cdot (100 \cdot 79500 + 95 \cdot 90000 + 5 \cdot 9000) \\ &= 82725 \text{ EUR/Auftrag}\end{aligned}$$

Univariate Statistik-Gartenzwerg-Grosshandel (11 min).mp4

Lösung 2-20: *Bibliotheken - Teil II*

a) $x_D = 62{,}998$ Std./Wo.; $x_{0,5} = 63{,}998$ Std./Wo.; $\overline{x} = 64{,}959$ Std./Wo.

b) $y_D = 44{,}166$ Std./Wo.; $y_{0,5} = 46{,}3165$ Std./Wo.; $\overline{y} = 48{,}831$ Std./Wo.

c) nein, da keine unimodale Häufigkeitsverteilung vorliegt

d) $z_{0,25} = 2{,}0502$ Mill.EUR/Jahr; $z_{0,5} = 3{,}353$ Mill.EUR/Jahr; $z_{0,75} = 4{,}346$ Mill.EUR/Jahr;
$z_{(1)} = 0{,}85$ Mill.EUR/Jahr; $z_{(62)} = 6{,}56$ Mill.EUR/Jahr

Lösung 2-21: *Erdbeerplantage - Teil II*

a) $\overline{x} = 4{,}65$ Std./Tag

b) $x_{0,5} = 4{,}5$ Std./Tag

Lösung 2-22: *Ladekabel*
$\overline{x} = 15{,}52$ EUR/Ladekabel

Lösung 2-23: *Tägliche Arbeitswege - Teil I*

a) Beschäftigter; Anfahrtsweg (in km) pro Beschäftigter; kardinalskaliert

b)

Anfahrtsweg	$h(x_j)$	$f(x_j)$	$F(x_j)$	$f_K(x_j)$
0 - 1	7	0,07	0,07	0,07
1 - 5	24	0,24	0,31	0,06
5 - 15	35	0,35	0,66	0,035
15 - 30	18	0,18	0,84	0,012
30 - 50	16	0,16	1,00	0,008

c) $\overline{x} = 14{,}705$ km

d) $x_D = 0{,}875$ km; Modus

e) $x_{0,5} = 10{,}4286$ km

f) $x_{0,05} = 0{,}7143$ km; $x_{0,25} = 4$ km; $x_{0,75} = 22{,}5$ km; $x_{0,9} = 37{,}5$ km

Lösung 2-24: *Reinigungsunternehmen - Teil I*

(1) $\overline{x} = 2488$ EUR; $x_{0,5} = 2470{,}5$ EUR;

(2) $\overline{x} = 2822{,}5$ EUR; $x_{0,5} = 2798$ EUR;

(3) $\overline{x} = 2470$ EUR; $x_{0,5} = 2454$ EUR;

(U) $\overline{x} = 2571{,}2$ EUR; $x_{0,5} = 2546$ EUR.

Lösung 2-25: *Besuche pro Woche*

a)

x_j	$h(x_j)$	$f(x_j)$	$F(x_j)$
0	7	0,14	0,14
1	11	0,22	0,36
2	5	0,10	0,46
3	7	0,14	0,60
4	7	0,14	0,74
5	4	0,08	0,82
6	5	0,10	0,92
7	3	0,06	0,98
9	1	0,02	1,00
	50	1,00	

b) $\overline{x} = 3$ Besuche

Lösung 2-26: *Kurzarbeiter*

a) $\overline{x} = 44{,}328$ %

b) $x_{0,5} = 40{,}36$ %

Lösung 2-27: *Perlenkette*

a)

Durchmesser	$h(x_j)$	$f(x_j)$	$F(x_j)$
3 - 4	20	0,278	0,278
4 - 5	21	0,292	0,570
5 - 6	19	0,264	0,834
6 - 7	9	0,125	0,959
7 - 8	3	0,041	1,000

b) $\overline{x} = 4{,}861$ mm

c) $x_D = 4{,}33$ mm

d) $x_{0,5} = 4{,}76$ mm

Lösung 2-28: *Zuckergewicht*

Füllgewicht	$h(x_j)$	$f(x_j)$	$F(x_j)$	$f_K(x_j)$
980 - 990	5	0,067	0,067	0,0067
990 - 995	12	0,160	0,227	0,032
995 - 1000	23	0,307	0,534	0,0614
1000 - 1005	22	0,293	0,827	0,0586
1005 - 1010	11	0,147	0,974	0,0294
1010 - 1020	2	0,026	1,000	0,0026

$x_D = 999{,}565$ g; $x_{0,5} = 999{,}446$ g; $\overline{x} = 999{,}27$ g

Lösung 2-29: *Walzabteilung*

X = “Bearbeitungszeit in sec./Stück)”;

Y = “hergestellte Stück pro Stunde”

I a) Man berechnet das harmonische Mittel wie folgt:

$$\begin{aligned}\overline{x}_H &= \frac{n}{\sum_{i=1}^{n} \cdot \frac{1}{x_i}} \\ &= \frac{4}{\frac{1}{20} + \frac{1}{30} + \frac{1}{60} + \frac{1}{60}} \\ &= 34,286 \text{ sec./Stück}\end{aligned}$$

b) Man berechnet das arithmetische Mittel wie folgt:

$$\begin{aligned}\overline{x} &= \frac{1}{n} \cdot \sum_{i=1}^{n} \cdot x_i \\ &= \frac{1}{2000} \cdot (1000 \cdot 20 + 500 \cdot 30 + 300 \cdot 60 + 200 \cdot 60) \\ &= 32,5 \text{ sec./Stück}\end{aligned}$$

II a)

- Zuerst errechnet man die jeweils pro Stunde produzierten Einheiten pro

Arbeiter wie folgt:

$$\begin{aligned}
\text{A produziert} \frac{3600 \text{ sec}}{20 \text{ sec}} &= 180 \text{ Stück pro Stunde} \\
\text{B produziert} \frac{3600 \text{ sec}}{30 \text{ sec}} &= 120 \text{ Stück pro Stunde} \\
\text{C und D produzieren jeweils} \frac{3600 \text{ sec}}{60 \text{ sec}} &= 60 \text{ Stück pro Stunde}
\end{aligned}$$

- Anschließend wird das arithmetische Mittel berechnet:

$$\begin{aligned}
\overline{y} &= \frac{1}{n} \cdot \sum_{i=1}^{n} \cdot x_i \\
&= \frac{1}{4} \cdot (180 + 120 + 60 + 60) \\
&= 105 \text{ Stück/h}
\end{aligned}$$

b) Das harmonische Mittel wird errechnet:

$$\begin{aligned}
\overline{y}_H &= \frac{n}{\sum_{i=1}^{n} \cdot \frac{1}{x_i}} \\
&= \frac{2000}{\frac{1000}{180} + \frac{500}{120} + \frac{300}{60} + \frac{200}{60}} \\
&= 110,77 \text{ Stück/h}
\end{aligned}$$

Lösung 2-30: *Schafzucht - Teil I*
Wir berechnen das harmonische Mittel. Von den in der Vorlesung präsentierten Formeln wählen wir die folgende, weil der Lohn im **Verhältnis** zur Leistung gegeben ist:

$$\overline{x}_H = \frac{\sum_{j=1}^{k} g_j}{\sum_{j=1}^{k} \frac{g_j}{x_j}},$$

wobei g_1 die Lohnsumme für die Irischen Arbeiter und g_2 die Lohnsumme für die einheimischen Arbeiter darstellt. Das bedeutet hier, dass $g_1 = 285$ und $g_2 = 260$, und $x_1 = 15$ den Lohn für die Irischen Arbeiter in Pfund pro Kilogramm und $x_2 = 20$ den Lohn für die einheimischen Arbeiter darstellt. Damit ergibt sich für das harmonische Mittel:

$$\begin{aligned}
\overline{x}_H &= \frac{\sum_{j=1}^{k} g_j}{\sum_{j=1}^{k} \frac{g_j}{x_j}} \\
&= \frac{285 + 260}{\frac{285}{15} + \frac{260}{20}} \\
&= \frac{545}{19 + 13} = 17,03 \frac{\text{Pfund}}{\text{kg}}
\end{aligned}$$

Univariate Statistik-Schafszucht Teil I.mp4

Lösung 2-31: *Intercity – Zug*
$\overline{x}_H = 112{,}5$ km/h

Lösung 2-32: *Bevölkerungsdichte und Ärztedichte*
Bevölkerungsdichte: $\overline{x} = 150$ Einwohner/km^2
Ärztedichte: $\overline{y}_H = 1$ Arzt/1000 Einwohner

Lösung 2-33: *Anstieg der Produktion*

a) $i_G = \sqrt[4]{1,1^2 \cdot 1,2^2} = 1{,}1489$

b) $i_{2007/2011} = x_{2011}/x_{2007} = 1,2^4 = 2{,}0736$

Lösung 2-34: *Drei Betriebe*

a) $\overline{x} = 740$ EUR Materialverbrauch/1000EUR Produktion

b) Materialverbrauch pro 1000EUR Produktion, kardinalskaliert

c) $\mathrm{i}_G = 0{,}97$; $x_{2013} = 112\,908\,000$ EUR Materialverbrauch

Lösung 2-35: *Kartoffeln*

a) X:"Kosten je verladene Tonne Kartoffeln"; $\overline{x}_H = 0{,}83$ EUR/Tonne

b) Y:"Verladene Tonnen Kartoffeln je Stunde"; $\overline{y} = 52$ t/h

Lösung 2-36: *Nelkenstrauß*
$\overline{x} = 0{,}95$ EUR/Nelke
Lösung 2-37: *Neubauwohnungen*

$\overline{x} = 9{,}15$ Monate/Wohnungseinheit
$x_{0,5} = 8{,}5403$ Monate/Wohnungseinheit
$x_D = 7{,}3846$ Monate/Wohnungseinheit

Lösung 2-38: *Wanderer*
$\overline{x}_H = 4{,}8$ km/h
Lösung 2-39: *Gleisbaubetrieb*
Berechnet wurde das einfache arithmetische Mittel. Dies ist jedoch falsch, da das statistische Merkmal Bauzeit/Gleis eine Verhältniszahl ist und die gegebenen Zusatzinformationen über die Bauzeit jedes Bauzuges (8 Stunden) sich inhaltlich auf den Zähler der Verhältniszahl beziehen (In dieser Zusatzinformation steckt die unterschiedliche Anzahl von Gleisen, die von jedem Bauzug in dieser Schicht verlegt wurden und die bei der Berechnung des einfachen arithmetischen Mittels nicht berücksichtigt wurden). Es ist deshalb das einfache harmonische Mittel anzuwenden: $\overline{x}_H = 166{,}67$ Minuten/Gleis.

Lösung 2-40: *Brutto- und Nettoeinkommen*
Bruttoeinkommen/Beschäftigten und Monat: $i_G = 1{,}0352$
Nettoeinkommen/Beschäftigten und Monat: $i_G = 1{,}0279$

Lösung 2-41: *Lernzeit*

a) X:"Aufwand für das Studium (in Stunden)"; kardinalskaliert

b)

$x_j^u \leq x_j < x_j^o$	x_j	$h(x_j)$	$f(x_j)$	$F(x_j)$	$f_K(x_j)$
0 - 3	1,5	30	0,3	0,3	0,1
3 - 6	4,5	48	0,48	0,78	0,16
6 - 8	7	17	0,17	0,95	0,085
8 - 12	10	5	0,05	1,00	0,0125

c) $x_{0,5} = 4{,}25$ Stunden

d) 38 Studenten

e) $\overline{x} = 4{,}3$ Stunden

f) $x_D = 4{,}333$ Stunden

Lösung 2-42: *Körpergröße*

a) X: "Körpergröße in cm"; $\overline{x} = 128$ cm; $s_x^2 = 26$ cm^2; $s_x = 5{,}1$ cm;
$v_x = 0{,}0398$
Y:"Körpergröße in Zoll"; $\overline{y} = 51{,}2$ Zoll; $s_y^2 = 4{,}16$ Zoll^2; $s_y = 2{,}04$ Zoll;
$v_y = 0{,}0398$

b) $x_i = a + by_i$ mit $a = 0$ und $b = 2,5$; $\overline{x} = a + b\overline{y}$
$s_x = |b|s_y$; $v_x = s_x/\overline{x} = (|b|s_y)/b\overline{y} = s_y/\overline{y} = v_y$

Lösung 2-43: *Maschinen*
$v_{Fiat} = 0{,}05$; $v_{Mercedes} = 0{,}02$
Lösung 2-44: *Leichtathletikabteilung*

a) Histogramm

b) 10,8 sec.

c) $MQ(\overline{x}) = s^2 = 0,0601\text{sec}^2$; $x_{0,5} = 10{,}52$ sec.; $MQ(x_Z) = 0{,}0602$

d) $MQ(\overline{x}) < MQ(x_Z)$; ja; Beweis über den Verschiebungssatz

$$MQ(c) = \frac{1}{n}\sum(x_i - c)^2 = \frac{1}{n}\sum(x_i - \overline{x} + \overline{x} - c)^2 = s^2 + (\overline{x} - c)^2; c = x_{0,5}$$

Lösung 2-45: *Schafzucht Teil - II*
Pfund = Gebühr + Wechselkurs · Gulden, $y_i = a + b \cdot x_i$
$\overline{x} = 7{,}8$ Tsd. Gulden; $s_x^2 = 37{,}76$ (Tsd. Gulden)^2;
$\overline{y} = 3{,}1$ Tsd. Pfund; $s_y^2 = 9{,}44$ (Tsd. Pfund)^2;
$\overline{y} = 3{,}1 = \text{a} + \text{b}\overline{x}$; $s_y^2 = b^2 \cdot s_x^2$
$a = 3,1 - 7,8b$; $b^2 = s_y^2/s_x^2 = 0{,}25$; $b = 0,5$
$a = 3,1 - 0,5 \cdot 7{,}8 = -\,0{,}8$
Gebühr: 0,8 Tsd. Pfund; Wechselkurs: 1 Pfund : 2 Gulden

Lösung 2-46: *Reinigungsunternehmen - Teil II*
$s_1^2 = 45\,693$; $s_1 = 213{,}759$; $v_1 = 0{,}086$
$s_2^2 = 4848{,}75$; $s_2 = 69{,}633$; $v_2 = 0{,}02467$
$s_3^2 = 12\,032{,}4$; $s_3 = 109{,}692$; $v_3 = 0{,}04441$

$$\begin{aligned} s^2 &= \sum_{l=1}^{3}\frac{n_l}{n}s_l^2 + \sum_{l=1}^{3}\frac{n_l}{n}(\overline{x}_l - \overline{x})^2 \\ &= \frac{6}{15}\cdot 45693 + \frac{4}{15}\cdot 4848,75 \\ &+ \frac{5}{15}\cdot 12032,4 + \frac{6}{15}(-83,2)^2 + \frac{4}{15}\cdot 251,3^2 + \frac{5}{15}(-101,2)^2 \\ &= 23581 + 23023,16 \\ &= 46604,16 \end{aligned}$$

$$\begin{aligned} s &= 215,88 \text{ EUR}; \\ v &= 0,08396 \end{aligned}$$

Lösung 2-47: *Tägliche Arbeitswege - Teil II*

a) $x_{0,25} = 4$; $x_{0,75} = 22{,}5$; $QA = 18{,}5$ km

b) $x_Z = 10{,}4286$ km; $d = 9{,}53$ km

Lösung 2-48: *CDs*

a) $\overline{x}_H = 4{,}27$ EUR/CD

b) $s^2 = 0{,}1519$; $s = 0{,}39$ EUR/CD

Lösung 2-49: *Führerschein–Entziehungen*

a) $\overline{x} = 35{,}65$ Jahre ≈ 36 Jahre

b) $s = 10{,}39$ Jahre; $v = 0{,}2914$

c) $x_{0,25} = 26{,}44$ Jahre; $x_{0,75} = 43{,}868$ Jahre; $QA = 17{,}428$ Jahre

d) $x_Z = 34$ Jahre

e) $d = 8{,}6$ Jahre

Lösung 2-50: *Bibliotheken - Teil III*
$s_x = 10{,}4775$ Std./Wo.; $s_y = 15{,}2382$ Std./Wo.
$v_x = 0{,}1612$; $v_y = 0{,}312$

Lösung 2-51: *Produktionsleistung einer Maschine*

a) $\overline{x} = 296{,}75$ Stück/Tag; $296{,}75/300 = 0{,}9892$

Die Maschine ist durchschnittlich zu 98,92 % pro Tag ausgelastet; die Behauptung ist richtig.

b) $x_Z = 296{,}5$ Stück/Tag; $d = 1{,}15$ Stück/Tag

Lösung 2-52: *Lineares Streuungsmaß*

$$\begin{aligned} X_Z &= 17; d = (1/n)\sum_{i=1}^{n}|x_i - x_Z|; \\ 8 &= (1/5)(|3-17| + |7-17| + |17-17| + |19-17| + |x_5 - 17|) \\ \Leftrightarrow 40 &= 14 + 10 + 0 + 2 + x_5 - 17 \\ \Leftrightarrow x_5 &= 31 \end{aligned}$$

Lösung 2-53: *Tarifverhandlungen*
X: „Jahresbruttolohn vor der Tarifverhandlung“, Y:„Jahresbruttolohn nach der Tarifverhandlung“; $\overline{x} = 21200$ EUR

$$\overline{y} = \frac{\sum_{i=1}^{n} y_i}{n} = \frac{\sum_{i=1}^{n_1}(a_1 + bx_i) + \sum_{i=n_1+1}^{n_2}(a_2 + bx_i)}{n} = \frac{n_1 a_1 + n_2 a_2}{n} + b\overline{x}$$

$n_1 = 4000, n_2 = 16000, n = 20000, a_1 = 200, a_2 = 0, b = 1,05; \overline{y} = 4000 \cdot 200/20000 + 1,05 \cdot 21200 = 40 + 22260 = 22300$ EUR

Lösung 2-54: *Kontrollzeiten*
X: Kontrollzeit pro Stück; Verhältniszahl. Die Bestandsmasse, über die zu mitteln ist, sind somit die produzierten Stück und nicht die Arbeiterinnen. Als zusätzliche Informationen sind jedoch nicht die produzierten Stück (Zusatzinformation zum Nenner des Merkmals), sondern die gesamte Kontrollzeit von 8 Stunden (Zusatzinformation zum Zähler von X) gegeben, die für alle Arbeiterinnen gleich ist. Anwendung des einfachen harmonischen Mittels:

$$\begin{aligned} \overline{x}_H &= \frac{n}{\sum_{i=1}^{n}\frac{1}{x_i}} = \frac{6}{\frac{1}{0,2} + \frac{1}{0,4} + \frac{1}{0,8} + \frac{1}{0,5} + \frac{1}{0,5} + \frac{1}{0,8}} \\ &= \frac{6}{5 + 2,5 + 1,25 + 2 + 2 + 1,25} \\ &= \frac{6}{14} = 0,4285 \approx 0,429 \text{ min/Stück} \end{aligned}$$

Es können jedoch auch die von den einzelnen Arbeiterinnen in den acht Stunden kontrollierten Stückzahlen aus der Tabelle ermittelt werden, die die absoluten Häufigkeiten für die Berechnung des arithmetischen Mittels darstellen (2400,1200,600,960,960,600). Dann Anwendung des gewogenen arithmetischen Mittels:

$$\overline{x} = (6 \cdot 480)/6720 = 2880/6720 = 0,4285 \approx 0,429 \text{ min/Stück}$$

Lösung 2-55: *Körperschaftssteueraufkommen*
Berechnung der gepoolten Varianz; dafür ist der Gesamtmittelwert erforderlich

$$\begin{aligned}
\overline{x} &= \frac{1}{n}\sum_{p=1}^{r}\overline{x}_p n_p \\
&= \frac{1}{10000}(100\cdot 9500 + 9500\cdot 100 + 300\cdot 200 + 100\cdot 400) = \frac{2000000}{10000} = 200
\end{aligned}$$

$$\begin{aligned}
s^2 &= \sum_{l=1}^{r}\frac{n_l}{n}s_l^2 + \sum_{l=1}^{r}\frac{n_l}{n}(\overline{x}_l - \overline{x})^2 \\
&= \left[\frac{100}{10000}\cdot(1000)^2 + \frac{9500}{10000}\cdot(80)^2 + \frac{300}{10000}\cdot(100)^2 + \frac{100}{10000}\cdot(250)^2\right] \\
&+ \left[\frac{100}{10000}\cdot(9500-200)^2 + \frac{9500}{10000}\cdot(100-200)^2 + \frac{300}{10000}\cdot(200-200)^2\right. \\
&+ \left.\frac{100}{10000}\cdot(400-200)^2\right] \\
&= (10000 + 6080 + 300 + 625) + (864900 + 9500 + 0 + 400) = 891805
\end{aligned}$$

$$\sqrt{s^2} = \sqrt{891805} = 944,3543 \approx 944$$

Lösung 2-56: *Benzinverbrauch*
Bei Verbrauchsdaten in Liter/100km: $v = 0,41/8,2 = 0,05$.
Der Variationskoeffizient verändert sich durch die Umrechnung auf Gallonen/100 Meilen nicht, da diese Umrechnung sowohl beim arithmetischen Mittel als auch bei der Standardabweichung vorgenommen werden muss.

Lösung 2-57: *Einkommen der Beamten*
Gegeben:
$x_{0,5} = 2590$; $x_{0,75} = 3590$; $R = 8000$; $IQR = 1770$; $x_{min} = 500$; $s = 1620$; $v = 0,54$
Gesucht:
$x_{0,25} = x_{0,75} - IQR = 3590 - 1770 = 1820$

Lösung 2-58: *Telefonanbieter*
$\overline{x}_H = \frac{\text{Gesamtsumme}}{\text{Gesamtdauer}} = \frac{78,75}{525} = 0,15$

Lösung 2-59: *Kaltmieten*
Medianklasse ist 10-13.

$$x_z = x_j^u + \frac{0,5 - F(x_j^u)}{f(x_j)}\cdot(x_j^o - x_j^u) = 10 + \frac{0,5-0,45}{0,3}\cdot(13-10) = 10,5$$

Wohnfläche (m^2) von ... bis unter ...	Anzahl	$f(x_j)$	$F(x_j)$
0-6	5	0,05	0,05
6-8	10	0,10	0,15
8-10	30	0,30	0,45
10-13	30	0,30	0,75
13-16	20	0,20	0,95
16-20	5	0,05	1,00
	100	1,00	

Lösung 2-60: *Telefon–Interviews*
10 Minuten|Freizeit $= 0,48$ EUR, 18 Interviews
$10\cdot 60$ Sek. $= 600/150 = 4$ Einheiten $\cdot\, 12 = 48$ Cent
10 Minuten|Tag $= 0,84$ EUR, 10 Interviews
$10\cdot 60$ Sek. $= 600/90 = 6,\overline{66}$ Einheiten $\to 7\cdot 12 = 84$ Cent
20 Minuten|Freizeit $= 0,96$ EUR, 11 Interviews
$20\cdot 60$ Sek. $= 1200/150 = 8$ Einheiten $\cdot\, 12 = 96$ Cent
20 Minuten|Tag $= 1,68$ EUR, 20 Interviews
$20\cdot 60$ Sek. $= 1200/90 = 13,\overline{33}$ Einheiten $\to 14\cdot 12 = 168$ Cent
$\overline{X} = (18\cdot 0,48 + 10\cdot 0,84 + 11\cdot 0,96 + 20\cdot 1,68)/59 = 1,037$

Lösung 2-61: *Glücksspielautomaten*
$\overline{x}(A) = 2,0 \quad \overline{x}(B) = 1,893$
$x_z(A) = 1,75$
Lösung 2-62: *Einkommensgleichheit*

Person	Atoll A Einkommen (in Kauri Schnecken)	x_i^2	$x_i - \overline{x}$	$(x_i - \overline{x})^2$
1	225	50625	32,625	1064,390625
2	185	34225	−7,375	54,390625
3	250	62500	57,625	3320,640625
4	150	22500	−42,375	1795,640625
5	237	56169	44,625	1991,390625
6	100	10000	−92,375	8533,140625
7	87	7569	−105,375	11103,890625
8	305	93025	112,625	12684,390625
$\sum$	1539	336613		40547,875

$\overline{x} = 1539/8 = 192,375$
Lösung 2-63: *Festgeldkonten*
Zuerst muss der Median der Zinssätze für Einlagen im jeweiligen Land $x_{0,5}(L)$ berechnet werden. Hierzu wird jeweils die Klasse $x_j, j \in \{1,2,3,4\}$ bestimmt, in welcher der Median liegt. Bezeichnet x_j^u die untere Klassengrenze, so ist für klassierte Daten

$$\begin{aligned}
x_{0,5}(L) &= x_j^u + \frac{0,5 - F(x_j^u)}{f(x_j)} \cdot (x_j^o - x_j^u) \\
x_{0,5}(1) &= 5,0 + \frac{0,5 - 0,44}{0,30} \cdot 0,5 = 5,1 \\
x_{0,5}(2) &= 5,0 + \frac{0,5 - 0,16}{0,34} \cdot 0,5 = 5,5 \\
x_{0,5}(3) &= 5,5 + \frac{0,5 - 0,375}{0,625} \cdot 0,5 = 5,6 \\
x_{0,5}(4) &= 4,5 + \frac{0,5 - 0,15}{0,35} \cdot 0,5 = 5,0 \\
x_{0,5}(5) &= 4,5 + \frac{0,5 - 0,275}{0,375} \cdot 0,5 = 4,8
\end{aligned}$$

Ermittlung der länderspezifischen variablen Zinssätze $z(L) = 0,7z + 0,3x_{0,5}(L)$ mit $z = 5,2$:
Land 1: $z(1) = 0,7 \cdot 5,2 + 0,3 \cdot 5,1 = 5,17$
Land 2: $z(2) = 0,7 \cdot 5,2 + 0,3 \cdot 5,5 = 5,29$
Land 3: $z(3) = 0,7 \cdot 5,2 + 0,3 \cdot 5,6 = 5,32$
Land 4: $z(4) = 0,7 \cdot 5,2 + 0,3 \cdot 5,0 = 5,14$
Land 5: $z(5) = 0,7 \cdot 5,2 + 0,3 \cdot 4,8 = 5,08$
Damit folgt für den durchschnittlichen variablen Zinssatz der Bank:
$\overline{z}(L) = 26/5 = 5,2$
Lösung 2-64: *Einwohnerzahlen*
Da wir zwei große Ausreißer in den Daten haben (China und Indien) muss der Median verwendet werden:

$$x_{0,5} = 0,5 \cdot (x_{(5)} + x_{(6)}) = 0,5 \cdot (122 + 139) = 131$$

Lösung 2-65: *Histogramm*

j	Klassen	H.-dichte	$f(x_j)$			$F(X)$
1	0 – 2	0,05	0,1	=	$0,05 \cdot 2$	0,1
2	2 – 6	0,10	0,4	=	$0,1 \cdot 4$	0,5
3	6 – 12	0,05	0,3	=	$0,05 \cdot 6$	0,8
4	12 – 20	0,025	0,2	=	$0,025 \cdot 8$	1,0
			1,0			

$$\begin{aligned}
F(X) &= F(x_j^u) + \frac{x - x_j^u}{x_j^o - x_j^u} \cdot f(x_j) \\
F(5) &= 0,1 + \frac{5-2}{6-2} \cdot 0,4 = 0,1 + 0,3 = 0,4 \\
F(16) &= 0,8 + \frac{16-12}{20-12} \cdot 0,2 = 0,8 + 0,1 = 0,9
\end{aligned}$$

$F(5 \leq X \leq 16) = F(16) - F(15) = 0,9 - 0,4 = 0,5$
$H(5 \leq X \leq 16) = F(5 \leq X \leq 16) \cdot 110 = 0,5 \cdot 110 = 55$

55 Gemeinden dieses Landkreises haben eine Einwohnerzahl von mindestens 5000 und höchstens 16000.

Lösung 2-66: *Internetstunden*
Berechnung von $f(.)$:

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

Internetstunden von ... bis unter ...	$\hat{f}(x_j)$	$f(x_j)$	$F(x_j)$
0 – 4	0,05	0,20	0,20
4 – 8	0,06	0,24	0,44
8 – 12	0,09	0,36	0,80
12 – 16	0,03	0,12	0,92
16 – 20	0,02	0,08	1,00
		1,00	

Medianklasse ist 8 – 12.

$$x_z = x_j^u + \frac{0,5 - F(x_j^u)}{f(x_j)} \cdot (x_j^o - x_j^u) = 8 + \frac{0,5 - 0,44}{0,36} \cdot (12 - 8) = 8,6666$$

Lösung 2-67: *Fließband*
Die mittlere Stückzeit kann also harmonisches Mittel berechnet werden, da es sich bei den angebenen Daten um Verhältniszahlen (Zeit/Stück) handelt, deren Nenner (die Stückzahlen) nicht explizit gegeben sind.

$$\begin{aligned}\overline{x}_H &= \frac{n}{\sum_{i=1}^{n} \frac{1}{x_i}} = \frac{6}{\frac{1}{2} + \frac{1}{4} + \frac{1}{8} + \frac{1}{10} + \frac{1}{8} + \frac{1}{5}} \\ &= \frac{6}{0,5 + 0,25 + 0,125 + 0,1 + 0,125 + 0,2} = \frac{6}{1,3} = 4,61538\end{aligned}$$

Lösung 2-68: *Schulbezirke*
Für Variable X: X = nominal, nur Modus
Für Variable Y: Y = metrisch, alle außer Modus
Lösung 2-69: *Tarifverhandlungen*
X: „Jahresbruttolohn vor der Tarifverhandlung“, Y:„Jahresbruttolohn nach der Tarifverhandlung“; $\overline{x} = 21200$ EUR

$$\overline{y} = \frac{\sum_{i=1}^{n} y_i}{n} = \frac{\sum_{i=1}^{n_1}(a_1 + bx_i) + \sum_{i=n_1+1}^{n_2}(a_2 + bx_i)}{n} = \frac{n_1 a_1 + n_2 a_2}{n} + b\overline{x}$$

$n_1 = 4000, n_2 = 16000, n = 20000, a_1 = 200, a_2 = 0, b = 1,05; \overline{y} = 4000 \cdot 200/20000 + 1,05 \cdot 21200 = 40 + 22260 = 22300$ EUR
Lösung 2-70: *Kaufkurs der Aktien*

$$\begin{aligned}\overline{x}_H = \frac{\sum_{j=1}^{k} g_j}{\sum_{j=1}^{k} \frac{g_j}{x_j}} &= \frac{45000 + 84000 + 3600 + 14000}{\frac{45000}{500} + \frac{84000}{600} + \frac{3600}{400} + \frac{14000}{700}} \\ &= \frac{146600}{90 + 140 + 9 + 20} \\ &= \frac{146600}{259} = 566,02\end{aligned}$$

oder

$$\overline{x} = \frac{1}{n}\sum_{j=1}^{k} x_j h(x_j) = \frac{500 \cdot 90 + 600 \cdot 140 + 400 \cdot 9 + 700 \cdot 20}{90 + 140 + 9 + 20} = \frac{146600}{259} = 566,02$$

Lösung 2-71: *Drei Stichproben*
gepoolter Datensatz

$$\overline{x} = \frac{1}{n}\sum_{p=1}^{r}\overline{x}_p n_p, \quad n = \sum_{p=1}^{r} n_p; \quad s^2 = \sum_{l=1}^{r}\frac{n_l}{n}s_l^2 + \sum_{l=1}^{r}\frac{n_l}{n}\left(\overline{x}_l - \overline{x}\right)^2$$

$$\begin{aligned}\overline{x} &= 0,1 \cdot 51 + 0,4 \cdot 53 + 0,5 \cdot 54 = 53,3\\ s^2 &= (0,1 \cdot 21 + 0,4 \cdot 22 + 0,5 \cdot 23)\\ &\quad + 0,1(51-53,3)^2\\ &\quad + 0,4(53-53,3)^2 + 0,5(54-53,3)^2\\ &\quad = 22,4 + 0,529 + 0,036 + 0,245 = 22,4 + 0,81 = 23,21\end{aligned}$$

Lösung 2-72: *Grafische Darstellung*
Die beiden Balkendiagramme scheiden aus, da die Variable metrisch stetig ist. Da die Daten klassiert mit unterschiedlicher Klassenbreite vorliegen, muss auf der Ordinatenachse die Häufigkeitsdichte abgetragen werden. Damit scheiden die Histogramme aus, bei denen auf der Ordinatenachse die relativen Häufigkeiten abgetragen wurden.

Alter	unter 15	15-18	18-25	25-65	65-90	$\sum$
Anzahl der Getöteten	126	76	258	808	638	1906
rel. Häufigkeit	0,06611	0,03987	0,13536	0,42392	0,33473	
Häufigkeitsdichte	0,00472	0,01329	0,01934	0,01060	0,01339	

Aufgrund dieser Häufigkeitsdichten gibt das Histogramm D die Daten korrekt wieder.

Lösung 2-73: *Das erste Tor*
Klassierte Häufigkeitsverteilung

Klasse	$h(x)$	$H(x)$	$f(x)$	$F(x)$
0 – 15	9	9	0,36	0,36
15 – 30	5	14	0,20	0,56
30 – 45	9	23	0,36	0,92
45 – 90	2	25	0,08	1,00

$f(20 < x \le 60) = F(60) - F(20) = ?$
$F(x) = F(x_j^u) + f(x_j) \cdot (x - x_j^u)/(x_j^o - x_j^u)$ (Interpolation)
$F(60) = F(45) + 0,08 \cdot (60-45)/(90-45) = 0,92 + 0,0267 = 0,9467$
$F(20) = F(15) + 0,2 \cdot (20-15)/(30-15) = 0,36 + 0,067 = 0,4267$
$f(20 < x \le 60) = F(60) - F(20) = 0,9467 - 0,4267 = 0,52$

2.2 Bivariate Statistik

Lösung 2-74: *Stellung im Beruf*

a)

Geschlecht	Stellung im Beruf			RV
	Beamte(r)	Angestellte(r)	Arbeiter(in)	Geschlecht
weiblich	15	20	5	40
männlich	10	30	20	60
RV Beruf	25	50	25	n=100

b) Bedingte Verteilung $f(y_j|x_1)$

	Beamte	Angestellte	Arbeiter
w	0,375	0,5	0,125

c) Bedingte Verteilung $f(x_i|y_2)$

	Angestellte
w	0,4
m	0,6

d) Die Merkmale sind nicht unabhängig, da z.B. $h_{11} \neq \frac{h_{1.} \cdot h_{.1}}{n}$ ist.

Bivariate Statistik-Stellung im Beruf.mp4

Lösung 2-75: *Einkommen und Alter*

a) und b)

Einkommen	Alter					RV X
	20-30	30-40	40-50	50-60	60-70	
0-1000	0,02	0,04	0,02	0,02	0,02	0,12
1000-1500	0,04	0,08	0,08	0,06	0,02	0,28
1500-2000	0,06	0,12	0,12	0,06	0,04	0,40
2000-3000	0,02	0,06	0,04	0,04	0,04	0,20
RV Y	0,14	0,30	0,26	0,18	0,12	1,00

c)

Einkommen	Alter				
	20-30	30-40	40-50	50-60	60-70
0-1000	0,143	0,133	0,077	0,111	0,167
1000-1500	0,286	0,267	0,308	0,333	0,167
1500-2000	0,429	0,400	0,462	0,333	0,333
2000-3000	0,143	0,200	0,154	0,222	0,333
	1,000	1,000	1,000	1,000	1,000

d)

Einkommen	Alter					
	20-30	30-40	40-50	50-60	60-70	
0-1000	0,167	0,333	0,167	0,167	0,167	1,000
1000-1500	0,143	0,286	0,286	0,214	0,071	1,000
1500-2000	0,150	0,300	0,300	0,150	0,100	1,000
2000-3000	0,100	0,300	0,200	0,200	0,200	1,000

e) $\overline{x} = 1610$ EUR

f) $\overline{y} = 43{,}4$ Jahre

g) $x_D = 1642{,}86$ EUR

h) $x_Z = 1625$ EUR

i) 1535,71 EUR; 1600 EUR; 1615,38 EUR; 1611,11 EUR; 1708,33 EUR

j) $s_x = 591{,}95$ EUR

k) $s_{xy} = 476$

Lösung 2-76: *Teesorten*
r_S = 0,5714
Lösung 2-77: *Sportveranstaltungen*

a) $\chi^2 = 14{,}4797$; $C = 0{,}2146$; $C_{korr} = 0{,}3035$

b) $\chi^2 = 0$

c) $\chi^2 = 0$

d) Zusammenhang unter a) nur scheinbar; er wird durch den Einfluss des Lebensalters vorgetäuscht. Bei der Ausschaltung dieses Einflusses durch die Untersuchung altersspezifischer Teilgesamtheiten zeigt sich, dass in Wirklichkeit Unabhängigkeit besteht.

Bivariate Statistik-Sportveranstaltungen (15 min).mp4
Lösung 2-78: *Verspätungen*
r_S = - 0,8
Lösung 2-79: *Sanatorium*

a) Spearman'scher Rangkorrelationsoeffizient

$$r_S = 1 - \frac{6\sum_{i=1}^{n} d_i^2}{n(n^2-1)}, \quad \text{wobei:} \quad d_i = \text{Rang } x_i - \text{Rang } y_i$$

Zusammenhang zwischen Gewicht und Laufleistung:

Platzierung Y	Gewicht X	Rang Gewicht	Differenz^2
1	70	2	1
2	60	1	1
3	80	6	9
4	77	3	1
5	82	8	9
6	81	7	1
7	78	4	9
8	100	10	4
9	83	9	0
10	110	11	1
11	79	5	36

$$\sum_{i=1}^{n} d_i^2 = 1+1+9+1+9+1+9+4+0+1+36$$
$$= 72$$

$$r_S = 1 - \frac{6 \cdot 72}{11(11^2-1)} = 0,6727$$

b) a) Median für nicht klassierte Daten,n ungerade

$$x_Z = x_{0.5} = x_{(\frac{n+1}{2})} = x_6 = 80 \text{ kg}$$

c) Quadratisches Streuungsmaßim Bezug auf Median

$$\begin{aligned} MQ(x_Z) = MQ(x_{0,5}) &= \frac{1}{n}\sum_{i=1}^{n}(x_i - x_{0,5})^2 \\ &= ((70-80)^2 + (60-80)^2 + (80-80)^2 \\ &+ (77-80)^2 + (82-80)^2 + (81-80)^2 \\ &+ (78-80)^2 + (100-80)^2 + (83-80)^2 \\ &+ (110-80)^2 + (79-80)^2) \cdot \frac{1}{11} \\ &= \frac{1828}{11} = 166,18 \end{aligned}$$

d) Kleiner, da die Varianz ein Streuungsmaß in Bezug auf das arithmetische Mittel ist. Das arithmetische Mittel minimiert die mittlere quadratische Abweichung.

e) Durchschnittsgeschindigkeit der Frauen?

- Eine Frau läuft den Lauf mit 2m/s

- Die andere Frau läuft den Lauf mit 4m/s

Insgesamt wurden 100m gelaufen, 50m mit 2m/s und 50m mit 4m/s. Kann auch wie eine Frau betrachtet werden die erst 50m langsamer und dann 50m schneller läuft.
Harmonisches Mittel:

$$\begin{aligned}\overline{x}_H &= \frac{50m+50m}{\frac{50m}{2m/s}+\frac{50m}{4m/s}} = \frac{100m}{25s+12,5s}\\ &= \frac{100m}{37,5s} = 2,667m/s\end{aligned}$$

Bivariate Statistik-Sanatorium.mp4

Lösung 2-80: *Außentemperatur und Dauer eines Weges*

$$r_{xy} = \frac{5\cdot(-1000)}{\sqrt{(5*1000)(5*7225-175^2)}} = \frac{-5000}{5224} = -0,953$$

Bivariate Statistik-Aussentemperatur und Dauer eines Weges (27 min).mp4

Lösung 2-81: *Buttersorten*

Spearman'scher Rangkorrelationskoeffizient:
$r_S = 1-(6\cdot\sum d_i^2)/[n\cdot(n^2-1)]$
$r_S = 1-(6\cdot 8)/(7\cdot 48) = 1-48/336 = 1-0,1429 = 08571 \approx 0,857$

Lösung 2-82: *Relationen der Merkmalsausprägungen*

Da Relationen angegeben sind, sind die beiden Merkmale X und Y ordinal skaliert; Der Spearman'sche Rangkorrelationskoeffizient ist ein geeignetes Maß

i	1	2	3	4	5
x_i	1	3	5	2	4
y_i	3	1	4	2	5
d_i	2	2	1	0	1

$$r_s = 1-\frac{6\sum_{i=1}^n d_i^2}{n(n^2-1)} = 1-\frac{6\cdot 10}{5\cdot(25-1)} = 0,5$$

Lösung 2-83: *Tekolom und IBBM - Teil I*

X = Kurs der Tekolom–Aktie, Y = Kurs der IBBM–Aktie,
relative Häufigkeiten: $f_1 = 73/365 = 0,2$; $f_2 = 146/365 = 0,4$; $f_3 = 146/365 = 0,4$
$\overline{x} = \sum f_i x_i$, $\overline{y} = \sum f_i y_i$, $s_X^2 = \sum f_i(x_i-\overline{x})^2$, $s_Y^2 = \sum f_i(y_i-\overline{y})^2$,
$s_{XY} = \sum f_i(x_i-\overline{x})(y_i-\overline{y})$
$\overline{x} = 41$, $\overline{y} = 126$, $s_X^2 = 14$, $s_Y^2 = 14$, $s_{XY} = 4$, $\rho = \frac{s_{XY}}{s_X s_Y} = 0,28571$

Lösung 2-84: *Mensaessen*

Es sei R_E die Preis/Leistungs–Rangzahl von Eintopf und R_1 die von Essen 1.
Fall A: $R_E = 1$, $R_1 = 4$

$$r_s = 1-\frac{6\sum d_i^2}{n(n^2-1)} = 1-\frac{6\cdot(0+1+9+9+9+4)}{6\cdot 35} = 1-\frac{32}{35} = 0,085714 > 0$$

Fall B: $R_E = 4$, $R_1 = 1$

$$r_s = 1-\frac{6\sum d_i^2}{n(n^2-1)} = 1-\frac{6\cdot(9+16+9+9+9+4)}{6\cdot 35} = 1-\frac{56}{35} = -0,6 < 0$$

Lösung 2-85: *Miete und Wohnfläche*

$s_{xy} = \frac{1}{n}\sum x_i y_i - \overline{x}\cdot\overline{y}$

$\overline{x} = \frac{3\cdot 40 + 60 + 2\cdot 80 + 4\cdot 90}{10} = \frac{700}{10} = 70$

$\overline{y} = \frac{3\cdot 12 + 15 + 9 + 5\cdot 10}{10} = \frac{110}{10} = 11$

$\overline{x} = 70,\ \overline{y} = 11,\ \sum x_i y_i = 7390$

$\rightarrow s_{xy} = 7390/10 - 70\cdot 11 = -31$

Lösung 2-86: *Cafeteria*

	Frauen	Männer	
Mensa	h_{11}	h_{12}	$h_{1\bullet}$
Cafeteria	h_{21}	h_{22}	$h_{2\bullet}$
	$h_{\bullet 1}$	$h_{\bullet 2}$	n

$n = 200, \quad h_{\bullet 1} = 0,375\cdot 200 = 75, \quad h_{21} = 45 \quad \Rightarrow h_{11} = 30$

Unabhängigkeit:

$\frac{1}{n} h_{\bullet 1}\cdot h_{1\bullet} = h_{11} \quad \Leftrightarrow \quad h_{1\bullet} = 30\cdot 200/75 = 80$

Lösung 2-87: *Old Faithful*

Variable X: Dauer einer Eruption (in Minuten)
Variable Y: Zeit zwischen zwei Eruptionen (in Minuten)
Beide Variablen sind metrischen Skalenniveaus
→ Bravais–Pearson–Korrelationskoeffizient.

$$\sum_i x_i = 26,9 \qquad \sum_i x_i^2 = 100,53$$
$$\sum_i y_i = 587 \qquad \sum_i y_i^2 = 45131$$
$$\sum_i x_i y_i = 2114,6$$

$$r_{yx} = \frac{n\sum_i^n x_i y_i - \sum_{i=1}^n x_i \sum_i^n y_i}{\sqrt{\left(n\sum_{i=1}^n x_i^2 - \sum_{i=1}^n x_i \sum_{i=1}^n x_i\right)\left(n\sum_{i=1}^n y_i^2 - \sum_{i=1}^n y_i \sum_{i=1}^n y_i\right)}} = 0,97727$$

Lösung 2-88: *Gefahrene Strecke*

Gesucht ist der Quartilsabstand:

$$QA = x_{0,75} - x_{0,25}$$
$$x_p = x_j^u + (p - F(x_j^u))(x_j^o - x_j^u)/f(x_j)$$
$$x_{0,25} = 50 + (0,25 - 0,15)50/0,25 = 70 \text{ km}$$
$$x_{0,75} = 300 + (0,75 - 0,7)200/0,2 = 350 \text{ km}$$
$$QA = 350 - 70 = 280 \text{ km}$$

Lösung 2-89: *Streuungsmaß*

a) Begründung:
Gegeben $n = 50$ Elemente, $k = 2$
$K(50; 2) = 1225$

b) metrisch (kardinal) skalierte Merkmale

Lösung 2-90: *Arbeitslose*
X: „Arbeitslosenquote"→ Verhältnis von Arbeitslosen zu Erwerbspersonen
Möglichkeit 1
Zusätzliche Informationen $g(x_j)$ sind Arbeitslosenzahlen, beziehen sich inhaltlich auf den Zähler des Merkmals X → harmonisches Mittel

$$\bar{x}_H = \frac{\sum_{j=1}^{k} g_j}{\sum_{j=1}^{k} \frac{g_j}{x_j}} = \frac{3000 + 4000}{\frac{3000}{5} + \frac{4000}{20}} = \frac{7000}{800} = 8,75$$

Die Arbeitslosenquote für das Bundesland beträgt 8,75%.

Möglichkeit 2
Anwendung des arithmetischen Mittels nach vorheriger Berechnung der Informationen für den Nenner des Merkmals X. Diese sind $3000/5 = 600$ und $4000/20 = 200$.

$$\begin{aligned} \bar{x} &= \frac{1}{n}\sum_{j=1}^{k} x_j h(x_j) \\ &= \frac{1}{800}(5 \cdot 600 + 20 \cdot 200) = \frac{3000 + 4000}{800} = \frac{7000}{800} \\ &= 8,75 \end{aligned}$$

Lösung 2-91: *Minimale Summe*
$c = \bar{x} = 10$
Begründung:
Quadratische Minimumseigenschaft des arithmetischen Mittels

Lösung 2-92: *Alter und Preis eines PKWs*
Gegeben: $s_{xy} = -5,4$ $\quad s_y^2 = 4$ $\quad R_{yx}^2 = 0,81$
Es ist $r_{yx} = s_{yx}/s_x s_y$. Daraus folgt: $s_x = s_{yx}/(r_{yx} s_y)$
Ferner ist: $r_{yx} = -0,9$ (r_{yx} und die Kovarianz haben das gleiche Vorzeichen);
$s_y = 2$
$s_x = s_{yx}/r_{yx}s_y = -5,4/(-0,9 \cdot 2) = 3$

Lösung 2-93: *Tekolom und IBBM - Teil II*
Tekolom–Aktie $var(X) = 16$, IBBM–Aktie $var(Y) = 1$, Portfolio $Z = 100X + 200Y$

$$\begin{aligned} corr(X, Y) &= \frac{cov(X, Y)}{\sigma_X \cdot \sigma_Y} \to cov(X, Y) = 0,2 \cdot 4 \cdot 1 = 0,8 \\ var(Z) &= var(100X) + var(200Y) + 2 \cdot 100 \cdot 200 cov(X, Y) \\ &= 10000 \cdot 16 + 40000 \cdot 1 + 40000 \cdot 0,8 = 232000 \end{aligned}$$

Lösung 2-94: *GM*
X – Wert der Aktie
K_i – Kurs der Aktie $\quad E_j$ – Wechselkurs
Dann ist

$$\bar{X} = \sum_{j=1}^{5}\sum_{i=1}^{5} E_j K_i f_{ji}$$

zu bestimmen. Da die Kovarianz Null ist, folgt aus der Kovarianzzerlegung

$$Cov(E, K) = \sum_{j=1}^{5}\sum_{i=1}^{5} e_j K_i f_{ji} - \overline{E} \cdot \overline{K},$$

dass der obige Wert dem Produkt der Mittelwerte von E und K entspricht. Mit den marginalen Häufigkeiten berechnet man:

Randverteilung von E: $0,2;\ 0,4;\ 0,1;\ 0,1;\ 0,2$
Randverteilung von K: $0,1;\ 0,3;\ 0,05;\ 0,2;\ 0,35$
$\overline{e} = 1,97$ EUR/\$ $\quad \overline{k} = 118,0175$ \$
damit resultiert der durchschnittliche Wert der GM-Aktie zu $\bar{x} = 232,5$ EUR.

Lösung 2-95: *Koeffizienten Vergleich*

1. H) Median
2. F) Korr. Kontingenzkoeffizient, K) Quadratische Kontingenz χ^2
3. D) Interquartilsabstand
4. B) Bravais–Pearson KK, D) IQR, G) Kovarianz, L) Spannweite, O) Standardabweichung, P) Varianz

3 Kombinatorik

Lösung 3-1: *Bunte Häuser*
$P(5) = 120$
Lösung 3-2: *TEA*

a) $P(3) = 6$

b) $P(3;2) = 3$

Lösung 3-3: *Bücher*
$V(9,6) = 60480$

Lösung 3-4: *Geburtstagsparty*

a) Kombinaiton ohne Wiederholung für $n = 12$, $k = 6$:

$$\begin{aligned} K(12,6) &= \binom{12}{6} = \frac{12!}{(12-6)!6!} = \frac{12!}{6!6!} \\ &= \frac{7\cdot \cancel{8}^2\cdot \cancel{9}^3\cdot \cancel{10}^2\cdot 11\cdot \cancel{12}^1}{1\cdot \cancel{2}\cdot \cancel{3}\cdot 4\cdot \cancel{5}\cdot \cancel{6}} \\ &= 7\cdot 2\cdot 3\cdot 2\cdot 11\cdot 1 = 924 \end{aligned}$$

b) Permutation für $n = 6$:

$$P(6) = 6! = 1\cdot 2\cdot 3\cdot 4\cdot 5\cdot 6 = 720$$

c) Permutation mit Wiederholung für $n = 6$ und $k_1 = 3$, $k_2 = 3$:

$$P(6;3,3) = \frac{6!}{3!3!} = \frac{4\cdot 5\ \cancel{6}}{\cancel{6}} = 20$$

Kombinatorik-Geburtstagsparty.mp4

Lösung 3-5: *Zwei Würfel*
$K^W(6,3) = 56$

Lösung 3-6: *Lotto Toto*
$V^W(3,13) = 1594323$

Lösung 3-7: *6 aus 49*
$K(49,6) =$13 983 816

Lösung 3-8: *Skatspieler*
Nein, denn es gibt $K(32;10) = 64\ 512\ 240$ mögliche Spiele. Der Skatspieler spielt 73000 Spiele/Jahr. Somit müsste er knapp 884 Jahre spielen.

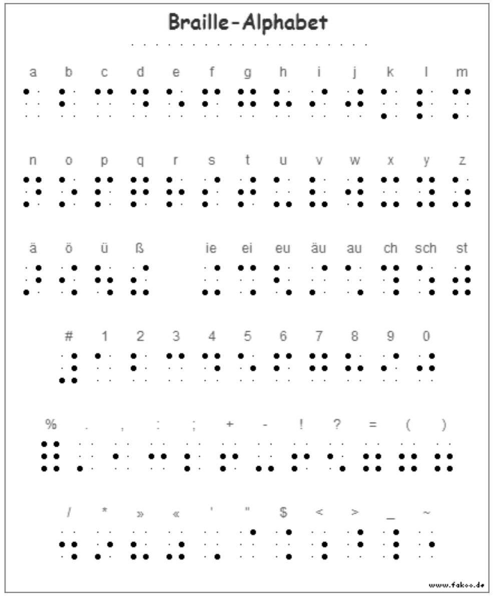
Lösung 3-9: *Einmaleins*
$K^W(9,2) = 45$

Lösung 3-10: *Hallenschwimmbad*
$P(13;1,2,10) = 858$

Lösung 3-11: *Schließfach*
$P(4,2) \cdot 3 = 12 \cdot 3 = 36$ Schließfächer $\rightarrow$ falls "3" bzw. "5" bzw. "7" doppelt sind.

Lösung 3-12: *Pferdelotto*
$V(23,4) = 212520$

Lösung 3-13: *Blindenschrift*
$V^W(2,6) = 64$



Quelle: http://www.siljakorn.de/braille-info.shtml

Lösung 3-14: *Wanderwege*

a) Variation mit Wiederholung für $n = 6$ Farben:
$$V^W(n,2) = 6^2 = 36$$

b) Kombinaiton ohne Wiederholung für $n = 7$ Farben:
$$K(n,2) = \binom{7}{2} = \frac{7!}{(7-2)!2!} = \frac{7!}{5!2!} = \frac{6 \cdot 7}{2} = \frac{42}{2} = 21$$

c) Kombination mit Wiederholung für $n = 5$ Farben:
$$K^W(n,2) = \binom{5+2-1}{2} = \binom{6}{2} = \frac{6!}{(6-2)!2!} = \frac{6!}{4!2!} = \frac{5 \cdot 6}{2} = 15$$

Lösung 3-15: *Genua Wahl*
Nein, denn $K(100, 5) = 75287520$ Es hätte ungefähr der 75 millionenfache Betrag des Einsatzes gezahlt werden müssen.

Lösung 3-16: *Unfallstation*

a) $V^W(3, 2) = 9$

b) $K^W(3, 2) = 6$

c) $V(3, 2) = 6$

d) $K(3, 2) = 3$

▶ Kombinatorik-Unfallstation.mp4

Lösung 3-17: *Orientierungsrundgang*

a) Permutation $n = 5$:

$$P(5) = 5! = 1 \cdot 2 \cdot 3 \cdot 4 \cdot 5 = 120$$

b) Permutation mit Wiederholung $n = 5$, $k_1 = 2$:

$$P(5; 2) = \frac{5!}{2!} = \frac{\cancel{1} \cdot \cancel{2} \cdot 3 \cdot 4 \cdot 5}{\cancel{1} \cdot \cancel{2}} = 60$$

c) 60 - 4·3! = 36
[4 Möglichkeiten für Dienst an aufeinanderfolgenden Tagen; für jede der 4 Möglichkeiten haben die anderen 3 Studenten $P(3) = 3!$ Möglichkeiten, sich auf die restlichen Tage zu verteilen]

▶ Kombinatorik-Orientierungsrundgang.mp4

Lösung 3-18: *Bridge*
$K(52, 13) = 635013559600$

Lösung 3-19: *Wagenreihungen*
$P(6; 2, 4) = 15$

Lösung 3-20: *Schiffsignale*
$K^W(6, 2) = 21$

Lösung 3-21: *Camel Cup*

a) $K^W(50, 3) = 22100$

b) Unter den 22 100 Möglichkeiten sind 50 Möglichkeiten, die Testausritte mit genau einem Kamel zu machen; $K(50, 3) = 19600$ Möglichkeiten, die Testausritte mit 3 unterschiedliche Kamelen zu machen und folglich $22100 - 50 - 19600 = 2450$ Möglichkeiten, die Testausritte mit zwei Kamelen zu machen.

Lösung 3-22: *Zahlenschlösser*
$V(5; 3) = 60$

Lösung 3-23: *Arbeitsgänge*
$P(9; 3, 2, 4) = 1260$

Lösung 3-24: *Pferderennen*
$V(10; 3) = 720$

Lösung 3-25: *Angebotsmöglichkeiten*
1. Unternehmen:

$$K(6,2) = \binom{6}{2} = 15$$

Kombination Preis/Menge nicht erlaubt: $15 - 1 = 14$

2. Unternehmen:

$$K(7,2) = \binom{7}{2} = 21$$

Kombination Preis/Menge nicht erlaubt: $21 - 1 = 20$
Insgesamt $14 \cdot 20 = 280$ Angebotsmöglichkeiten.

Lösung 3-26: *Schachturnier*
$K(12,2) = 66$

Lösung 3-27: *Geschenke für die Abteilungsleiter*
$P(5;1,1,1,2) = \dfrac{5!}{1!1!1!2!} = 120/2 = 60$

Lösung 3-28: *Computerraum-Code*
Anzahl aller Codes $55**$ $= 9 \cdot 10 = 90$ (dritte Stelle $\neq 5$)
Anzahl aller Codes $**55$ $= 9 \cdot 10 = 90$ (zweite Stelle $\neq 5$)
Anzahl aller Codes $*55*$ $= 9 \cdot 9 = 81$ (erste & vierte Stelle $\neq 5$)
$\rightarrow 261$

Lösung 3-29: *Code-Schlösser*
Anzahl solcher 5–stelligen Codes $= 2 \cdot 9^4 = 13122$
Anzahl aller Codes mit 4 gleichen Ziffern $= 2 \cdot 9 = 18$
$\rightarrow 13122 - 18 = 13104$

Lösung 3-30: *Anzahl der Abweichungen*
$K(50;2) = 1225$

Lösung 3-31: *Parkplätze*
$K(9,5) = 126$

Lösung 3-32: *Hemden*
$6 + 6 + 6 + 4 = 22$

4 Grundzüge der Wahrscheinlichkeitsrechnung

Lösung 4-1: *Wurf eines Würfels*

a) Wir fassen das zweimalige Würfeln als *ein* Zufallsexperiment auf. Ein Ergebnis ist dann durch das geordnete Paar bzw. 2-Tupel (a, b) gegeben, wobei a das Ergebnis im ersten Wurf und b das Ergebnis im zweiten Wurf angibt. Die Menge aller Ergebnisse lässt sich dann schreiben als:

$$\begin{aligned}S = \{1,2,3,4,5,6\}^2 &= \{(a,b) \mid a, b \in \{1,2,3,4,5,6\}\}\\ &= \{(1,1),(1,2),(1,3),(1,4),(1,5),(1,6),\\ &\quad (2,1),(2,2),(2,3),(2,4),(2,5),(2,6),\\ &\quad (3,1),(3,2),(3,3),(3,4),(3,5),(3,6),\\ &\quad (4,1),(4,2),(4,3),(4,4),(4,5),(4,6),\\ &\quad (5,1),(5,2),(5,3),(5,4),(5,5),(5,6),\\ &\quad (6,1),(6,2),(6,3),(6,4),(6,5),(6,6)\}.\end{aligned}$$

b) Nun wollen wir das zweimalige Würfeln als das zweimalige Ziehen aus einer Urne mit 6 "Kugeln", die für die Ergebnisse $1, \ldots, 6$ stehen, betrachten. Wir wählen also $k = 2$ Kugeln aus $n = 6$ vielen aus, wobei die Reihenfolge wichtig ist (1. oder 2. Wurf) und Wiederholung möglich ist (beispielsweise ist $(1,1) \in S$). Daher gilt:

$$\#S = V^W(6,2) = 6^2 = 6 \cdot 6 = 36.$$

c) Die Ereignisse A und B lassen sich dann folgendermaßen darstellen:

$$\begin{aligned}A &= \{6\} \times \{1,2,3,4,5,6\}\\ &= \{(6,b) \mid b \in \{1,2,3,4,5,6\}\}\\ &= \{(6,1),(6,2),(6,3),(6,4),(6,5),(6,6)\},\end{aligned}$$

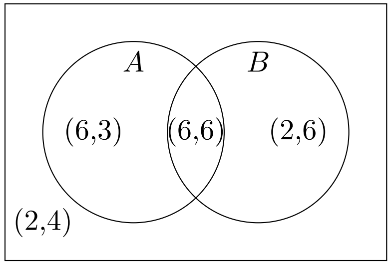
$$\begin{aligned}B &= \{1,2,3,4,5,6\} \times \{6\}\\ &= \{(a,6) \mid a \in \{1,2,3,4,5,6\}\}\\ &= \{(1,6),(2,6),(3,6),(4,6),(5,6),(6,6)\}.\end{aligned}$$

d) Die Vereinigung und der Durchschnitt der Ereignisse A und B sind gegeben durch:

$$\begin{aligned}A \cup B &= \{(a,b) \in S \mid (a,b) \in A \text{ oder } (a,b) \in B\}\\ &= \{(a,b) \in S \mid a = 6 \text{ oder } b = 6\}\\ &= \{(6,1),(6,2),(6,3),(6,4),(6,5),(6,6),\\ &\qquad (1,6),(2,6),(3,6),(4,6),(5,6)\},\end{aligned}$$

$$\begin{aligned}A \cap B &= \{(a,b) \in S \mid (a,b) \in A \text{ und } (a,b) \in B\}\\ &= \{(a,b) \in S \mid a = 6 \text{ und } b = 6\}\\ &= \{(6,6)\}.\end{aligned}$$

e) Veranschaulichung durch Venn-Diagramm:



f) Die leere Menge $\emptyset$ ist ein unmögliches Ereignis in jedem Zufallsexperiment.

g) Die Komplementärereignisse $\bar{A}$ und $\bar{B}$ lassen sich darstellen als:

$$\begin{aligned}\bar{A} = S \setminus A &= \{(a,b) \in S \mid (a,b) \notin A\}\\ &= \{(a,b) \in S \mid a \neq 6\}\\ &= \{1,2,3,4,5\} \times \{1,2,3,4,5,6\},\end{aligned}$$

$$\begin{aligned}\bar{B} = S \setminus B &= \{(a,b) \in S \mid (a,b) \notin B\}\\ &= \{(a,b) \in S \mid b \neq 6\}\\ &= \{1,2,3,4,5,6\} \times \{1,2,3,4,5\}.\end{aligned}$$

h) Nein. Denn beispielsweise $(6,3) \in A$, aber $(6,3) \notin B$.

i) Nein, weil $A \cap B = \{(6,6)\} \neq \emptyset$.

Lösung 4-2: *Ereignisoperationen*
$A \cup A = A$; $A \cup \emptyset = A$; $A \cap A = A$ $A \cap \emptyset = \emptyset$; $\emptyset \cap S = \emptyset$
$A \cup S = S$; $A \cup \overline{A} = S$; $A \cap S = A$; $A \cap \overline{A} = \emptyset$

Lösung 4-3: *Augenzahl eines Würfels*
$A = \emptyset$, $B = \emptyset \quad \to A = B$

Lösung 4-4: *Bauernwirtschaft*

a) Wir interessieren uns dafür, ob ein Bauernhof 0, 1 oder 2 Traktoren und, ob er 0, 1 oder 2 Pflüge zur Verfügung hat. Gegeben die Interpretation des Ereignisraum als Menge der Ereignisse, die wir unterscheiden, definieren wir daher

$$S = \{e_1, ..., e_9\},$$

mit $e_1 = \{0{,}0\}$, $e_2 = \{0{,}1\}$, $e_3 = \{0{,}2\}$, $e_4 = \{1{,}0\}$, $e_5 = \{1{,}1\}$, $e_6 = \{1{,}2\}$, $e_7 = \{2{,}0\}$, $e_8 = \{2{,}1\}$, $e_9 = \{2,2\}$.

b) Da wir ein karteisches Produkt zweier jeweils 3-elementiger Wahrscheinlichkeitsräume betrachten,und dementsprechend Reihenfolge und Wiederholung möglich ist, können wir die folgende Formel verwenden:

$$V^W(3,2) = 3^2 = 9.$$

c) $A \cap B = \{1,1\}$ Es sind genau ein Traktor und ein Pflug vorhanden.

Wahrscheinlichkeitsrechnung-Bauernwirtschaft.mp4

Lösung 4-5: *Zwei Würfel*
$A = \{(1,6), (2,5), (3,4), (4,3), (5,2), (6,1)\}$

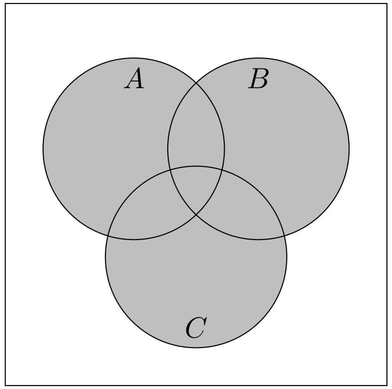
Lösung 4-6: *Kartenspiel*
$A \cap C = \emptyset$; $A \cap D = \emptyset$
$B \cap D = \emptyset$; $C \cap D = \emptyset$

Lösung 4-7: *Nicht–disjunkte Teilmengen*
Wir sagen: Ein Ereignis A tritt genau dann ein, wenn das Ergebnis $s \in S$ des Zufallsexperiments (bzw. einer konkreten Durchführung des Zufallsexperiments) in A liegt, falls also $s \in A$ gilt.
Sei nun $s \in S$ ein beliebiges Ergebnis des Zufallsexperiments. Dann gilt:

a) wenigstens eines tritt ein

$$\begin{aligned}
&\Leftrightarrow A \text{ tritt ein oder } B \text{ tritt ein oder } C \text{ tritt ein} \\
&\Leftrightarrow s \in A \text{ oder } s \in B \text{ oder } s \in C \\
&\Leftrightarrow s \in A \cup B \cup C. && \text{(Definiton von “}\cup\text{”)}
\end{aligned}$$

Veranschaulichung:



b) nur A eintritt (d.h. B und C treten nicht ein)

$$\begin{aligned}
&\Leftrightarrow A \text{ tritt ein und } B \text{ tritt nicht ein und } C \text{ tritt nicht ein} \\
&\Leftrightarrow s \in A \text{ und } s \notin B \text{ und } s \notin C \\
&\Leftrightarrow s \in A \text{ und } s \in \bar{B} \text{ und } s \in \bar{C} && \text{(Definition von Komplement)} \\
&\Leftrightarrow s \in A \cap \bar{B} \cap \bar{C}. && \text{(Definiton von “}\cap\text{”)}
\end{aligned}$$

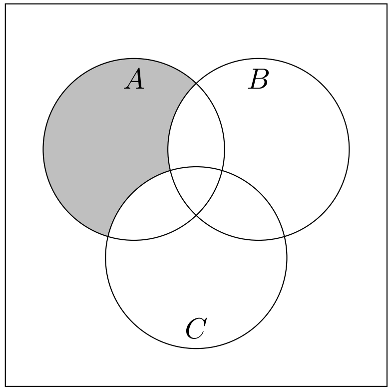
Alternativ:

$$\begin{aligned}
&\Leftrightarrow s \in A \text{ und } s \notin B \text{ und } s \notin C \\
&\Leftrightarrow s \in A \text{ und } s \notin B \cup C \\
&\Leftrightarrow s \in A \setminus (B \cup C). && \text{(Definiton von “}\backslash\text{”)}
\end{aligned}$$

Oder:

$$\begin{aligned}
&\Leftrightarrow s \in A \text{ und } s \notin B \text{ und } s \notin C \\
&\Leftrightarrow s \in A \text{ und } s \notin B \cup C \\
&\Leftrightarrow s \in A \text{ und } s \in \overline{B \cup C} \\
&\Leftrightarrow s \in A \cap \overline{B \cup C}.
\end{aligned}$$

Veranschaulichung:



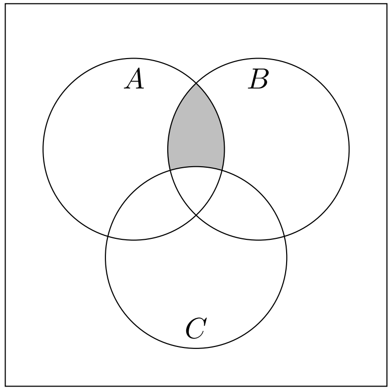
c) A und B treten ein, aber nicht C

$$\begin{aligned}
&\Leftrightarrow A \text{ tritt ein und } B \text{ tritt ein und } C \text{ tritt nicht ein} \\
&\Leftrightarrow s \in A \text{ und } s \in B \text{ und } s \notin C \\
&\Leftrightarrow s \in A \text{ und } s \in B \text{ und } s \in \bar{C} \\
&\Leftrightarrow s \in A \cap B \cap \bar{C}.
\end{aligned}$$

Alternativ:

$$\begin{aligned}
&\Leftrightarrow s \in A \text{ und } s \in B \text{ und } s \notin C \\
&\Leftrightarrow s \in (A \cap B) \setminus C.
\end{aligned}$$

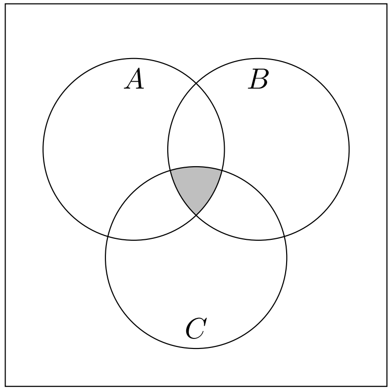
Veranschaulichung:



d) alle drei treten ein

$$\begin{aligned}
&\Leftrightarrow A \text{ tritt ein und } B \text{ tritt ein und } C \text{ tritt ein} \\
&\Leftrightarrow s \in A \text{ und } s \in B \text{ und } s \in C \\
&\Leftrightarrow s \in A \cap B \cap C.
\end{aligned}$$

Veranschaulichung:



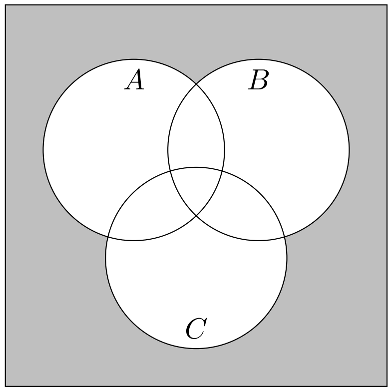
e) keines tritt ein

$$\begin{aligned}
&\Leftrightarrow A \text{ tritt nicht ein und } B \text{ tritt nicht ein und } C \text{ tritt nicht ein}\\
&\Leftrightarrow s \notin A \text{ und } s \notin B \text{ und } s \notin C\\
&\Leftrightarrow s \in \bar{A} \text{ und } s \in \bar{B} \text{ und } s \in \bar{C}\\
&\Leftrightarrow s \in \bar{A} \cap \bar{B} \cap \bar{C}.
\end{aligned}$$

Alternativ:

$$\begin{aligned}
&\Leftrightarrow s \notin A \text{ und } s \notin B \text{ und } s \notin C\\
&\Leftrightarrow s \notin A \cup B \cup C\\
&\Leftrightarrow s \in \overline{A \cup B \cup C}.
\end{aligned}$$

Veranschaulichung:

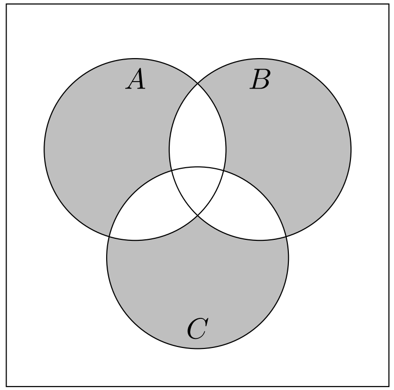


f) es tritt genau eines ein $\Leftrightarrow$ es tritt wenigstens eines ein (siehe a)) aber nicht zwei (vgl. c)) und auch nicht alle drei (siehe d))

$$\begin{aligned}
\Leftrightarrow s \in A \cup B \cup C &\text{ und } s \notin A \cap B \text{ und } s \notin B \cap C \text{ und}\\
&\text{ und } s \notin A \cap C \text{ und } s \notin A \cap B \cap C\\
\Leftrightarrow s \in A \cup B \cup C &\text{ und } s \notin A \cap B \text{ und } s \notin B \cap C \text{ und}\\
&\text{ und } s \notin A \cap C\\
\Leftrightarrow s \in A \cup B \cup C &\text{ und } s \notin (A \cap B) \cup (B \cap C) \cup (A \cap C)\\
\Leftrightarrow s \in [A \cup B \cup C] &\setminus [(A \cap B) \cup (B \cap C) \cup (A \cap C)]\,.
\end{aligned}$$

Bemerkung. Dieses Ereignis lässt sich auch noch auf andere Arten darstellen, die Herleitung kann aber recht mühsam werden.

Veranschaulichung:



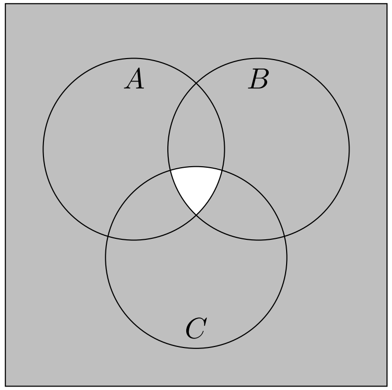
g) es treten höchstens zwei eintreten $\Leftrightarrow$ alle drei treten nicht gleichzeitig ein (siehe d))

$$\begin{aligned}
&\Leftrightarrow s \notin A \cap B \cap C\\
&\Leftrightarrow s \in \overline{A \cap B \cap C}.
\end{aligned}$$

Alternativ:

$$\begin{aligned}
&\Leftrightarrow s \notin A \cap B \cap C\\
&\Leftrightarrow s \in \overline{A \cap B \cap C}\\
&\Leftrightarrow s \in S \setminus (A \cap B \cap C)\\
&\Leftrightarrow s \in S \setminus A \text{ oder } s \in S \setminus B \text{ oder } s \in S \setminus C\\
&\Leftrightarrow s \in \bar{A} \text{ oder } s \in \bar{B} \text{ oder } s \in \bar{C}\\
&\Leftrightarrow s \in \bar{A} \cup \bar{B} \cup \bar{C}.
\end{aligned}$$

Veranschaulichung:



Lösung 4-8: *Felgen*

a) $P(4) = 24$

b) $A = \{(2,1,4,3), (2,3,4,1), (2,4,1,3), (3,1,4,2), (3,4,1,2), (3,4,2,1), (4,1,2,3), (4,3,1,2), (4,3,2,1)\}$

Lösung 4-9: *Produktionshalle*

a) $A = A_1 \cup A_2 \cup A_3 = (A_1 \cap \overline{A}_2 \cap \overline{A}_3) \cup (\overline{A}_1 \cap A_2 \cap \overline{A}_3) \cup (\overline{A}_1 \cap \overline{A}_2 \cap A_3) \cup (A_1 \cap A_2 \cap A_3)$
$\cup (\overline{A}_1 \cap A_2 \cap A_3) \cup (A_1 \cap \overline{A}_2 \cap A_3) \cup (A_1 \cap A_2 \cap \overline{A}_3)$,
$B = A_1 \cap A_2 \cap A_3$,
$C = A_1 \cap A_2 \cap \overline{A}_3$,
$D = \overline{A}_1 \cap \overline{A}_2 \cap \overline{A}_3$,
$E = (\overline{A}_1 \cap A_2 \cap A_3) \cup (A_1 \cap \overline{A}_2 \cap A_3) \cup (A_1 \cap A_2 \cap \overline{A}_3)$

b) $A = \overline{D}$ oder $D = \overline{A}$,
$B \subset A$, $C \subset A$, $E \subset A$, $D \subset \overline{B}$, $C \subset E$

Lösung 4-10: *1950–2000*
$E_0 = \{\text{keine Person erlebt das Jahr 2000}\}$
$E_1 = \{\text{eine Person erlebt das Jahr 2000}\}$
...
$E_{10} = \{\text{alle 10 Personen erleben das Jahr 2000}\}$
$A = E_2$, $B_1 = \{E_2, E_3, \ldots, E_{10}\}$, $B_2 = E_1$
$A \cap B_2 = \emptyset$ und $B_1 \cap B_2 = \emptyset$

Lösung 4-11: *Zufallsexperiment*

a) $A = \{3\}$; $P(A) = 1/6$

b) $B = \{1, 5\}$; $P(B) = 1/3$

c) $C = \{2, 4, 6\}$; $P(C) = 1/2$

d) klassische Definition der Wahrscheinlichkeit nach Laplace

Wahrscheinlichkeitsrechnung-Zufallsexperiment.mp4

Lösung 4-12: *Bus*
$F = \{\text{Besuch bei der Freundin}\}$, $U = \{\text{Erscheinen in der Universität}\}$
$F = \overline{U}$, $P(U) = 1/10$, $P(F) = 1 - P(U) = 9/10$
$U = \{$Ankunft an der Bushaltestelle zu einer Minute, so dass der Bus B_U zur Universität als erster kommt$\}$
$F = \{$Ankunft an der Bushaltestelle zu einer Minute, so dass der Bus B_F zur Freundin als erster kommt$\}$
$U = \{m_1, m_2, \ldots, m_k\}$ mit m_i für B_U günstige Minute
$P(U) = \frac{\text{Zahl der für Uni. günstigen Minuten}}{\text{Gesamtzahl der Minuten}} = \frac{k}{20} = \frac{1}{10} = \frac{2}{20}$

Abfahrtszeiten zur Universität sind somit $\underline{^{02}}$, $\underline{^{22}}$ und $\underline{^{42}}$
Die Universität hat somit nicht die gleiche Chance, da nur 2 Minuten Wartezeit auf den Bus zur Universität und bei allen anderen Minuten kommt der Bus zur Freundin zuerst.
Gleiche Chance wäre bei Abfahrtszeiten des Bus B_U $\underline{^{10}}$, $\underline{^{30}}$ und $\underline{^{50}}$ gegeben.

Lösung 4-13: *Schachbrett*
$A = \{$Plazieren von 8 Türmen, so dass keiner den anderen schlagen kann$\}$
Anzahl der möglichen Fälle: $K(n,k) = K(64,8)$
Anzahl der für A günstigen Fälle: $P(k) = P(8) = 8!$
$P(A) = \frac{P(k)}{K(n,k)} = \frac{8!8!56!}{64!} = 9,1 \cdot 10^{-6}$

Lösung 4-14: *Zufällige Ziehung einer Karte*

a) $A = \{$Pique-Karte$\}$, $P(A) = 8/32$

b) $B = \{$As$\}$, $P(B) = 4/32$

c) $C = \{$Pique-Karte oder As$\}$, $C = A \cup B$, $P(C) = P(A \cup B)$
$P(A \cup B) = P(A) + P(B) - P(A \cap B) = 11/32$

Lösung 4-15: *Antriebswellen*

a) Für die Überprüfung der i-ten Welle, mit $i = 1, \ldots, 10000$, bezeichne

$$x_i = \begin{cases} 0 & \text{falls Welle } i \text{ kein Ausschuss} \\ 1 & \text{falls Welle } i \text{ Ausschuss} \end{cases}$$

das Ergebnis des i-ten Durchgangs des Zufallsexperiments. Damit ist die Menge aller Ergebnisse des Zufallsexperiments für die Überprüfung einer Welle gegeben durch

$$S = \{0, 1\}$$

mit den Elementarereignissen

$$\{0\} = \text{“Kein Ausschuss wird produziert”},$$
$$\{1\} = \text{“Ausschuss wird produziert”}.$$

Nun bestimmen wir mithilfe der absoluten Häufigkeiten h die relativen Häufigkeiten $\hat{f}$ für die beiden Ergebnisse “kein Auschuss” und “Ausschuss” in unserer Stichprobe vom Umfang 10000. Die relativen Häufigkeiten ziehen wir heran, um die jeweiligen Wahrscheinlichkeiten zu

schätzen:

$$\begin{aligned} h(0) &= \sum_{i=1}^{10000} I(x_i = 0) = 4500 + 5200 = 9700, \\ h(1) &= 10000 - h(0) = 300, \qquad \text{(da alle } x_i \text{ binär)} \\ \hat{f}(0) &= \frac{h(0)}{10000} = 0.97 \approx P(\{0\}), \\ \hat{f}(1) &= \frac{h(1)}{10000} = 0.03 \approx P(\{1\}) = P(\text{“Ausschuss wird produziert”}). \end{aligned}$$

b) Von Mises, Pearson, Fisher, u. a., fassten Wahrscheinlichkeiten als den Grenzwert der relativen Häufigkeiten auf, wenn die Anzahl unabhängiger Wiederholungen des Zufallsexperiments gegen unendlich strebt (frequentistischer Wahrscheinlichkeitsbegriff). In unserem Beispiel bedeutet dies:

$$P(\{0\}) = \lim_{n\to\infty} \hat{f}_n(0),$$
$$P(\{1\}) = \lim_{n\to\infty} \hat{f}_n(1),$$

wobei $\hat{f}_n$ die relative Häufigkeit in Bezug auf eine Stichprobe vom Umfang n angibt. Da hier $n = 10000$ relativ groß ist, gehen wir davon aus, dass die relativen Häufigkeiten annähernd den Wahrscheinlichkeiten entsprechen.

Lösung 4-16: *15 Cent*

a) $A = \{$weniger als 15 Cent bei zweimaligem Ziehen mit Zurücklegen$\}$
$P(A) = 3/8$

b) $B = \{$weniger als 15 Cent bei zweimaligem Ziehen ohne Zurücklegen$\}$
$P(B) = 0,333$

Lösung 4-17: *Blumen*
R ={Rose, Rose}, N ={Narzisse, Narzisse}, L ={Lilie, Lilie}
A ={zwei Blumen gleicher Art}= $R \cup N \cup L$
$P(A) = P(R \cup N \cup L) = P(R) + P(N) + P(L) = 6/66 + 15/66 + 1/66 = 1/3$
Lösung 4-18: *Garderobe*
Es gibt 5! Möglichkeiten, jedem Mann einen Hut zuzuordnen. Eine davon ist im Sinne der Aufgabe nur günstig.
A ={jeder Mann bekommt seinen Hut}, $P(A) = 1/120$

Lösung 4-19: *Fachbereichsrat*
A ={drei Professoren werden gewählt}
Anzahl mögliche Fälle: $K(n, k) = K(8, 3) = 56$
Anzahl günstiger Fälle für A: $K(4, 3) = 4$
$P(A) = 4/56 = 1/14$

Lösung 4-20: *Hörer/innen einer Statistik–Vorlesung*
M ={männlicher Hörer}, F ={weiblicher Hörer},
V ={Hörer ist Student VWL}, B ={Hörer ist Student BWL}

a) $P(M) = 0,4$

b) $P(W) = 0,6$

c) $P(M \cup V) = 0,71$

d) $P(W \cap B) = 0,29$

e) $P(B|W) = 0,483$

f) $P(M|V) = 0,31$

Lösung 4-21: *Kfz–Händler*
K ={PKW weist Karosseriemängel auf}, M ={PKW weist Motormängel auf}
$M \cap K$ ={PKW weist Karosserie- und Motormängel auf}
Gesucht $P(\overline{M} \cap \overline{K})$
$P(\overline{M} \cap \overline{K}) = 1 - P(M \cup K)$
$P(M \cup K) = P(M) + P(K) - P(M \cap K)$ =0,6 + 0,8 - 0,4 = 1
$P(\overline{M} \cap \overline{K}) = 1 - 1 = 0$
Lösung 4-22: *Öffentliche Verkehrsmittel*

a) $P(\overline{U}) = 1 - P(U) = 1 - 0,4 = 0,6$

b) $P(S \cup B) = 0,4$

c) $P(\overline{U \cup S \cup B}) = 0,29$

d) $P(\overline{U \cap S \cap B}) = 0,99$

Lösung 4-23: *Aufzug*

- Ergebnisse: (i, j, k) mit Person 1 steigt in Etage i aus, Person 2 in Etage j und Person 3 in Etage k
- Elementarereignisse: $\{(i, j, k)\}$ mit $2 \leq i, j, k \leq 7$
- Ereignisraum: $S = \{(2, 2, 2), (2, 2, 3), \ldots, (7, 7, 7)\}$
- Anzahl der Elementarereignisse: $6 \cdot 6 \cdot 6 = 216$
  Formal: Ziehe dreimal aus einer Urne mit sechs Kugeln mit Wiederholung und die Anordnung spielt eine Rolle: $V^W(6; 3) = 6^3 = 216$
- $A = \{(4, 4, 4)\} \rightarrow P(A) = 1/216$
- $B = \{(2, 2, 2), \ldots, (7, 7, 7)\} \Rightarrow P(B) = 6/216 = 1/36$
- $C = \{(2, 3, 4), (2, 3, 5), \ldots, (7, 6, 5)\}$
  Formal: Ziehe dreimal aus einer Urne mit sechs Kugeln ohne Wiederholung und die Anordnung spielt eine Rolle: $V(6; 3) = \frac{6!}{3!} = 120 \Rightarrow P(C) = 120/216 = 5/9$

Lösung 4-24: *Tageszeitungen*
$Z_i = \{\text{Bewohner liest Zeitung } i\}$, $i = 1, 2$
$P(Z_1) = 0,6$; $P(Z_2) = 0,8$; $P(\overline{Z}_1 \cap \overline{Z}_2) = 0,1 = P(\overline{Z_1 \cup Z_2}) \Rightarrow P(Z_1 \cup Z_2) = 0,9$

a) $P(Z_1 \cap Z_2) = 0,5$; allgemeiner Additionssatz

b) $P(Z_1 \cap \overline{Z}_2) = P(Z_1) - P(Z_1 \cap Z_2) = 0,1$

c) $P(Z_2|\overline{Z}_1) = 0,75$; Formel für bedingte Wahrscheinlichkeit

d) $P[(\overline{Z}_1 \cap \overline{Z}_2) \cup (\overline{Z}_1 \cap Z_2) \cup (Z_1 \cap \overline{Z}_2)] = 1 - P(Z_1 \cap Z_2) = 0,5$

e) $P(\overline{Z}_1) = 0,4$

Lösung 4-25: *Geburtstag*

a) Dafür muss man die Gegenwahrscheinlichkeit benutzen:
- A: Min. ein Gast hat an meinem Geburtstag Geburtstag
- $\overline{A}$: Kein Gast hat an meinem Geburtstag Geburtstag

Wenn die Geburtstage unabhängig sind, dann gilt:

$$\begin{aligned} P(A) &= 1 - P(\overline{A}) \\ &= 1 - \left(\frac{364}{365}\right)^3 = 0,0082 \end{aligned}$$

b) Dafür muss man die Gegenwahrscheinlichkeit benutzen:
- B: Min. zwei Gäste haben am gleichten Tag Geburtstag
- $\overline{B}$: Alle Gäste haben an unterschiedlichen Tagen Geburtstag

Die Anzahl aller möglichen Geburtstagskombination berechnet sich als Variation mit Wiederholung: $V^W(365; 3) = 365^3$. Die Anzahl der für $\overline{B}$ günstige Ereignisse berechnet sich als Variation ohne Wiederholung: $V(365; 3) = \frac{365!}{362!} = 365 \cdot 364 \cdot 363$.

$$P(B) = 1 - P(\overline{B}) = 1 - \frac{365}{365}\frac{364}{365}\frac{363}{365} = 0,0082$$

c)

$$\begin{aligned} P(A) &= 1 - \left(\frac{364}{365}\right)^{23} = 0,0651 \\ P(B) &= 1 - \frac{365}{365} \cdots \frac{343}{365} = 0,5073 \end{aligned}$$

d) Der Unterschied liegt darin, dass bei A einer der Gäste am gleichen Tag wie ich Geburtstag haben muss. Bei B müssen zwei Gäste am gleichen Tag im Jahr Geburtstag haben, jedoch nicht notwendigerweise am gleichen Tag wie ich.

Lösung 4-26: *Altbauwohnung*

a) Wir betrachten die Ereignisse:
- F: Wasserzufuhr friert ein
- S: Strom fällt aus
- W: Es ist Winterzeit.

Aus dem Aufgabentext lassen sich folgende Informationen entnehmen:

1. *Sowohl im Winter (i.e. gegeben Winter) als auch im Sommer treten die beiden Missstände unabhängig voneinander auf.*

$$P(F \cap S|W) = P(F|W) \cdot P(S|W) \text{ und } P(F \cap S|\overline{W}) = P(F|\overline{W}) \cdot P(S|\overline{W})$$

2. *So friert natürlich das Wasser nur ein, wenn es Winter ist, und zwar mit 80%iger Wahrscheinlichkeit.*

$$P(F|W) = 0,8 \text{ und } P(F|\overline{W}) = 0$$

3. *Der Strom fällt aber, selbst wenn es nicht Winter ist, mit 40%iger Wahrscheinlichkeit aus. Das entspricht der gleichen Wahrscheinlichkeit, mit der der Strom, wenn es Winter ist, nicht ausfällt.*

$$P(S|\overline{W}) = 0,4 \text{ und } P(\overline{S}|W) = 0,4$$

4. *Gehen Sie davon aus, dass die Winterzeit 30% der gesamten Jahreszeit ausmacht.*

$$P(W) = 0,3$$

b) Gesucht ist $P(F)$. Um diese unbekannte Wahrscheinlichkeit auf die bekannten zurückzuführen, verwenden wir den Satz von der totalen Wahrscheinlichkeit:

$$\begin{aligned} P(F) &= P(F \cap W) + P(F \cap \overline{W}) \\ &= P(F|W)P(W) + P(F|\overline{W})P(\overline{W}) \\ &= 0,8 \cdot 0,3 + 0 \cdot 0,7 \\ &= 0,24. \end{aligned}$$

c) Gesucht ist $P(S)$. Wieder verwenden wir den Satz von der totalen Wahrscheinlichkeit

$$\begin{aligned} P(S) &= P(S \cap W) + P(S \cap \overline{W}) \\ &= P(S|W)P(W) + P(S|\overline{W})P(\overline{W}) \\ &= 0,6 \cdot 0,3 + 0,4 \cdot 0,7 \\ &= 0,46. \end{aligned}$$

d) Gesucht ist $P(F \cap S)$:

$$\begin{aligned} P(F \cap S) &= P(F \cap S|W)P(W) + P(F \cap S|\overline{W})P(\overline{W}) \\ &= P(F|W) \cdot P(S|W) \cdot P(W) + P(F|\overline{W}) \cdot P(S|\overline{W})P(\overline{W}) \\ &= 0,8 \cdot 0,6 \cdot 0,3 + 0 \cdot 0,6 \cdot 0,7 \\ &= 0,144 \end{aligned}$$

e) Gesucht ist die Wahrscheinlichkeit $P(F|S)$. Mit der Formel für die bedingte Wahrscheinlichkeit können wir die gesuchte Wahrscheinlichkeit auf die gerade berechneten Wahrscheinlichkeiten zurückführen:

$$\begin{aligned} P(F|S) &= \frac{P(F \cap S)}{P(S)} \\ &= \frac{0,144}{0,46} \approx 0,313. \end{aligned}$$

f) Gesucht ist die Wahrscheinlichkeit $P(S|F)$. Wieder mit der Definition der bedingten Wahrscheinlichkeit gilt

$$\begin{aligned} P(S|F)] &= \frac{P(F \cap S)}{P(F)} \\ &= \frac{0,144}{0,24} = 0,6. \end{aligned}$$

g) Gesucht ist die Wahrscheinlichkeit $P(S \cup F)$. Mit der Formel für die Vereinigung zweier Ereignisse gilt unter Verwendung der Ergebnisse der vorherigen Teilaufgaben

$$\begin{aligned} P(S \cup F) &= P(S) + P(F) - P(S \cap F) \\ &= 0,24 + 0,46 - 0,144 \\ &= 0,556. \end{aligned}$$

h) Gesucht ist die Wahrscheinlichkeit $P(\overline{S \cap F})$. Wir nutzen die Berechnung über die Gegenwahrscheinlichkeit und erhalten

$$\begin{aligned} P(\overline{S \cap F}) &= 1 - P(S \cap F) \\ &= 1 - 0,144 \\ &= 0,856. \end{aligned}$$

Lösung 4-27: *Last*
A_1 = {Seil 1 hält}; A_2 = {Seil 2 hält}; $P(A_1) = P(A_2) = 0,99$
$P(A_1 \cup A_2) = 0,9999$; allgemeiner Additionssatz

Lösung 4-28: *Alter*
A ={ein Zehnjähriger wird 40 Jahre alt}; $P(A) = 0,82277$
B ={ein Zehnjähriger wird 70 Jahre alt}; $P(B) = 0,37977$
$P(B|A) = 0,4616$; ($B \subset A$!); Formel für bedingte Wahrscheinlichkeit

Lösung 4-29: *Kugeln*
$A_1 = \{$die 1. Ziehung liefert eine rote Kugel$\}$
$A_2 = \{$die 2. Ziehung liefert eine weiße Kugel$\}$
$A_3 = \{$die 3. Ziehung liefert eine rote Kugel$\}$
$P(A_1) = 2/10$; $P(A_2|A_1) = 3/9$; $P(A_3|A_1 \cap A_2) = 1/8$
$P(A_1 \cap A_2 \cap A_3) = 1/120$; Multiplikationssatz

Lösung 4-30: *Papierstreifen*
$P(A_1) = P(A_2) = P(A_3) = 1/2$

a) $A = A_1 \cap A_2 \cap A_3 = \emptyset \rightarrow P(A) = 0$
$P(A_1)P(A_2)P(A_3) = \frac{1}{2} \cdot \frac{1}{2} \cdot \frac{1}{2} = 1/8 \neq 0$
Die 3 Ereignisse sind nicht voneinander unabhängig.

b) Die Ereignisse sind paarweise unabhängig. z.B. für A_1, A_2: $P(A_2|A_1) = \frac{1}{2} = P(A_1)$ und $P(A_1|A_2) = \frac{1}{2} = P(A_2)$

Lösung 4-31: *Webstühle*
Wir definieren zunächst den Ereignisraum $\{A, \overline{A}\}^3$ mit der Potenzmenge als Menge der Ereignisse. Die Zufallsvariable A_i gibt an, ob Webstuhl i Aufmerksamkeit beansprucht ($A_i = A$), oder nicht ($A_i = \overline{A}$), $i = 1, 2, 3$. Nach Aufgabenstellung gilt:

$$P[\{A_1 = \overline{A}\}] = 0,9,$$
$$P[\{A_2 = \overline{A}\}] = 0,8,$$
$$P[\{A_3 = \overline{A}\}] = 0,85.$$

a) Gesucht ist die Wahrscheinlichkeit des Ereignisses $\{A_1 = \overline{A}\} \cap \{A_2 = \overline{A}\} \cap \{A_3 = \overline{A}\}$. Wir nutzen die Unabhängigkeit dieser Ereignisse, um den Multiplikationssatz für unabhängige Ereignisse anzuwenden:

$$\begin{aligned}
&P[\{A_1 = \overline{A}\} \cap \{A_2 = \overline{A}\} \cap \{A_3 = \overline{A}\}] \\
&= P[\{A_1 = \overline{A}\}] \cdot P[\{A_2 = \overline{A}\}] \cdot P[\{A_3 = \overline{A}\}] \\
&= 0,9 \cdot 0,8 \cdot 0,85 \\
&= 0,612
\end{aligned}$$

b) Gesucht ist die Wahrscheinlichkeit des Ereignisses $\{A_1 = \overline{A}\} \cup \{A_2 = \overline{A}\} \cup \{A_3 = \overline{A}\}$. Wir verwenden den Satz von de Morgan, um die Wahrscheinlichkeit des Ereignisses auf die Wahrscheinlichkeit des Gegenereignisses zurückzuführen und nutzen die Unabhängigkeit der $A_i's$, und erhalten:

$$\begin{aligned}
&P[\{A_1 = \overline{A}\} \cup \{A_2 = \overline{A}\} \cup \{A_3 = \overline{A}\}] \\
&= 1 - P[\{A_1 = A\} \cap \{A_2 = A\} \cap \{A_3 = A\}] \\
&= 1 - P[\{A_1 = A\}] \cdot P[\{A_2 = A\}] \cdot P[\{A_3 = A\}] \\
&= 1 - 0,1 \cdot 0,2 \cdot 0,15 = 0,997.
\end{aligned}$$

Wahrscheinlichkeitsrechnung-Webstuehle.mp4

Lösung 4-32: *Angler*
$A_i = \{$Angeln am See i$\}$; $i = 1, 2, 3$; $P(A_i) = 1/3$
$B = \{$Angler hat etwas gefangen$\}$; $P(B|A_1) = 2/3$; $P(B|A_2) = 3/4$;
$P(B|A_3) = 4/5 \rightarrow P(B) = 133/180$; Formel für totale Wahrscheinlichkeit
$P(A_2|B) = 0,3383$; Satz von Bayes

Lösung 4-33: *Fernschreiben*
$A_1 = \{$Fehler bei 1. übertragung$\}$; $A_2 = \{$Fehler bei 2. übertragung$\}$
$A_1 \cap A_2 = \{$Fehler bei 1. und 2. übertragung$\}$; $P(A_1) = 0,01$; $P(A_2|A_1) = 0,1$
$P(A_1 \cap A_2) = P(A_1) \cdot P(A_2|A_1) = 0,001$

Lösung 4-34: *Fußballmannschaft*
$A = \{$Gewinn beim 1. Spiel$\}$; $B = \{$Gewinn beim 2. Spiel$\}$; $C = \{$Gewinn beim 3. Spiel$\}$;
$P(A) = P(B) = P(C) = 0,7$
$D = \{$Gewinnspiele überwiegen$\} = [(A \cap B \cap \overline{C}) \cup (A \cap \overline{B} \cap C) \cup (\overline{A} \cap B \cap C) \cup (A \cap B \cap C)]$
$P(D) = 0,784$

Lösung 4-35: *Waldbrand*
$A =${Förster entdeckt Brand zu spät};
$B =${Feuermelder funktioniert bei Betätigung nicht};
$P(A) = 0,05$; $P(B) = 0,03$
$A \cup B =${Förster entdeckt Brand zu spät oder Feuermelder funktioniert nicht}
$P(A \cup B) = P(A) + P(B) - P(A \cap B) = 0,0785$; allgemeiner Additionssatz

Lösung 4-36: *Regal*
$A =${Bände stehen in richtiger Reihenfolge von rechts nach links oder von links nach rechts}
$P(4) = 4! = 24$ mögliche Reihenfolgen der 4 Bände, zwei davon günstig
$P(A) = 2/24 = 1/12$

Lösung 4-37: *Tagesproduktion*
$M_i =${Stück stammt von Maschine i}; $i = 1, 2, 3$
$P(M_1) = 0,1$; $P(M_2) = 0,4$; $P(M_3) = 0,5$
$A =${Ausschussstück}
$P(A) = 0,031$; Formel für totale Wahrscheinlichkeit
$P(A|M_1) = 0,05$; $P(A|M_2) = 0,04$; $P(A|M_3) = 0,02$
Satz von Bayes: $P(M_1|A) = 0,1613$; $P(M_2|A) = 0,5161$; $P(M_3|A) = 0,3226$

Lösung 4-38: *Schummelei*

a) Ereignisse: S - Student schummelt, V - Maschine sagt Student schummelt (Schummelverdacht)

$$\begin{aligned} P(V|S) &= 0,9 \Rightarrow P(\overline{V}|S) = 0,1 \\ P(\overline{V}|\overline{S}) &= 0,9 \Rightarrow P(V|\overline{S}) = 0,1 \\ P(S) &= 0,1 \Rightarrow P(\overline{S}) = 0,9 \end{aligned}$$

b) Gesucht ist $P(V)$. Wir verwenden den Satz von der totalen Wahrscheinlichkeit und erhalten

$$\begin{aligned} P(V) &= P(V \cap S) + P(V \cap \overline{S}) \\ &= P(V|S) \cdot P(S) + P(V|\overline{S}) \cdot P(\overline{S}) \\ &= 0,9 \cdot 0,1 + 0,1 \cdot 0,9 \\ &= 0,18 \end{aligned}$$

c) Gesucht ist die Wahrscheinlichkeit $P(S|V)$. Wir verwenden den Satz von Bayes, der sich aus der Definition der bedingten Wahrscheinlichkeit und dem Satz von der totalen Wahrscheinlichkeit ergibt. Wir haben den Satz von der totalen Wahrscheinlichkeit aber bereits in der vorherigen Teilaufgabe angewandt und erhalten daher:

$$\begin{aligned} P(S|V) &= \frac{P(S \cap V)}{P(V)} \\ &= \frac{P(V|S) \cdot P(S)}{P(V)} \\ &= \frac{0,9 \cdot 0,1}{0,18} \\ &= 0,5 \end{aligned}$$

Wahrscheinlichkeitsrechnung-Schummelei.mp4

Lösung 4-39: *Reitturnier*
Ereignisse:

- S: Hindernis ist Steilsprung
- O: Hindernis ist Oxer
- G: Hindernis ist Graben
- F: Pferd macht Fehler am Hindernis

Wahrscheinlichkeiten:

$$P(S) = 0,6 \Rightarrow P(\overline{S}) = 0,4$$
$$P(O) = 0,3 \Rightarrow P(\overline{O}) = 0,7$$
$$P(G) = 0,1 \Rightarrow P(\overline{G}) = 0,1$$
$$P(F|S) = 0,03 \Rightarrow P(\overline{F}|S) = 0,97$$
$$P(F|O) = 0,04 \Rightarrow P(\overline{F}|O) = 0,96$$
$$P(F|G) = 0,05 \Rightarrow P(\overline{F}|G) = 0,95$$

a) Gesucht ist die Wahrscheinlichkeit $P(F)$, die wir mit dem Satz von der totalen Wahrscheinlichkeit berechnen:

$$\begin{aligned} P(F) &= P((F \cap S) \cup (F \cap O) \cup (F \cap G)) \\ &= P(F \cap S) + P(F \cap O) + P(F \cap G) \\ &= P(F|S)P(S) + P(F|O)P(O) + P(F|G)P(G) \\ &= 0,03 \cdot 0,6 + 0,04 \cdot 0,3 + 0,05 \cdot 0,1 \\ &= 0,035 \end{aligned}$$

b) Wir betrachten die Ereignisse: O_i Fehler am $i$ten Oxer mit $P(O_i) = 0,96$. Gesucht ist die Wahrscheinlichkeit $P(\overline{O}_1 \cap \overline{O}_2 \cap O_3)$. Wir verwenden die Unabhängigkeit der O_i und erhalten

$$\begin{aligned} P(\overline{O}_1 \cap \overline{O}_2 \cap O_3) &= P(\overline{O}_1) \cdot P(\overline{O}_2) \cdot P(O_3) \\ &= 0,96 \cdot 0,96 \cdot 0,04 \\ &= \frac{576}{15625} \approx 0,04. \end{aligned}$$

Wahrscheinlichkeitsrechnung-Reitturnier.mp4

Lösung 4-40: *Entwicklungsabteilung*

a) $A_1 = \{$Entwicklungsabteilung ist für Markteinführung des neuen Produkts$\}$
$A_2 = \{$Marketingabteilung ist für Markteinführung des neuen Produkts$\}$
$A_3 = \{$Geschäftsleitung ist für Markteinführung des neuen Produkts$\}$
$P(A_1) = 0,9;$
$P(A_2|A_1) = 0,7;$
$P(\overline{A}_3|A_1 \cap A_2) = 0,2;$
$P(A_3|A_1 \cap \overline{A}_2) = 0,4$

b) $P(A_1 \cap A_2 \cap A_3) = P(A_1) \cdot P(A_2|A_1) \cdot [1 - P(\overline{A}_3|A_1 \cap A_2)] = 0,504$

c) $P(A_1 \cap A_3) = P(A_1 \cap A_2 \cap A_3) + P(A_1 \cap \overline{A}_2 \cap A_3) = 0,612$

Lösung 4-41: *Eigener PKW*
$A_1 =$ „Bürokraft ist weiblich"
$A_2 =$ „Bürokraft ist männlich"
$B =$ „Bürokraft kommt mit dem PKW zur Arbeit"
$P(A_1) = 0,6, \quad P(A_2) = 0,4, \quad P(B|A_1) = 0,7, \quad P(B|A_2) = 0,8;$
$A_1 \cap A_2 = \varnothing, \quad A_1 \cup A_2 = S;$
$P(A_1|B) = [P(B|A_1)P(A_1)]/[P(B|A_1)P(A_1) + P(B|A_2)P(A_2)] = 0,42/0,74$
$= 05676 \approx 0,57$

Lösung 4-42: *Ereignisraum*
Allgemein gilt: $P(A \cup B) = P(A) + P(B) - P(A \cap B) \Rightarrow P(A \cap B) = \frac{1}{4}$

a) falsch: $P(A \cup B) \neq 1$ und $P(A \cap B) \neq \varnothing$

b) falsch: $P(\overline{A}) = P(B)$ aber $P(A \cap B) \neq \varnothing$

c) falsch: $P(A \cap B) \neq \varnothing$

d) richtig: $P(A|B) = \frac{P(A \cap B)}{P(B)} = \frac{1/4}{1/2} = \frac{1}{2} = P(A)$

Lösung 4-43: *Zerlegung*
Voraussetzung für eine vollständige Zerlegung:
$A_i \cap A_j = \varnothing$ (Paarweise disjunkt)
$A_1 \cup A_2 \cup \ldots \cup A_n = S$ (Vereinigung ergibt S)
$P(A_i) \neq 0 \Rightarrow A_i \neq \varnothing$

a) Nein: $A_3 \cap A_5 \neq \varnothing$

b) Nein: $A_4 \cap A_6 \neq \varnothing$

c) Nein: $A_1 \cup A_2 \cup A_4 \neq S$

d) Nein: $A_2 \cap A_5 \neq \varnothing$

e) Nein: $A_5 \cap A_6 \neq \varnothing$

f) Ja: $A_2 \cap A_3 = \varnothing, A_3 \cap A_4 = \varnothing, A_2 \cap A_4 = \varnothing, A_2 \cup A_3 \cup A_4 = S$

Lösung 4-44: *Musikkassette*
$A_1 =$ „Kassette wird im Auto benutzt“, $P(A_1) = 0,7$;
$A_2 =$ „Kassette wird im Haus benutzt“, $P(A_2) = 0,3$;
$B =$ „Kassette besitzt eine Lebensdauer von mehr als 500 Stunden“;
$P(B|A_1) = 0,75, P(B|A_2) = 0,95; A_1 \cap A_2 = \varnothing, A_1 \cup A_2 = S$;
$P(B) = \sum_i^n P(B|A_i)P(A_i) = 0,81$ (Satz von Bayes)

Lösung 4-45: *Erregertest*
I: „Mit Bazillus infiziert“; E: „Erkennen des Bazillus durch Test (positiver Befund)“
Gegeben:
$P(E|I) = 0,95$; $\qquad P(E|\overline{I}) = 0,03$ $\qquad P(I) = 0,02$
Daraus ergeben sich:
$P(\overline{E}|I) = 0,05$; $\qquad P(\overline{E}|\overline{I}) = 0,97$; $\qquad P(\overline{I}) = 0,98$
Gesucht: $P(I|E) = P(E \cap I)/P(E)$
$P(E \cap I) = P(E|I)P(I) = 0,019$
$P(E) = P(E \cap I) + P(E \cap \overline{I}) = P(E|I)P(I) + P(E|\overline{I})P(\overline{I}) = 0,0484$
$P(I|E) = 0,39256 \approx 0,3926$

Lösung 4-46: *Banknoten*
Es sei E: „Bankangestellter erkennt gefälschte Banknote “ und B: „Die Banknote ist echt“.
Gegeben: $P(E|\overline{B}) = 0,9$; $P(E|B) = 0,05$; $P(\overline{B}) = 0,002$.
Gesucht: $P(B|E)$
Anwendung des Satzes von Bayes:

$$P(B|E) = \frac{P(B \cap B)}{P(E)} = \frac{P(E|B) \cdot P(B)}{P(E|B) \cdot P(B) + P(E|\overline{B}) \cdot P(\overline{B})}$$
$$= 0,9652$$

Lösung 4-47: *Tennis*
A: „Assistent geht spielen, d.h. kann nicht angetroffen werden“; $P(A) = 8/20 = 0,4$;
B: „Chef sucht ihn auf oder ruft an“; $P(B) = 0,1$
Da A und B unabhängig voneinander sind, folgt $P(A \cap B) = P(A) \cdot P(B)$.
Wenn $P(A \cap B) = P(A) \cdot P(B) \leq 0,01$ sein soll, und $P(B) = 0,1$ ist, muss $P(A) \leq 0,1 = 2/20$ sein, so dass der Assistent nur an höchstens 2 von 20 Arbeitstagen spielen gehen darf.

Lösung 4-48: *Systemausfallrisiko*
A: „Ausfall des Computers A im Verlaufe eines Arbeitstages“; $P(A) = 0,05$
B: „Ausfall des Computers B im Verlaufe eines Arbeitstages“; $P(B) = 0,04$
C: $A \cap B$ = „Ausfall beider Computer im Verlaufe eines Arbeitstages“ = „Systemausfall“
$\overline{C}$ = „Computer A und B fallen nicht gleichzeitig aus“ = „kein Systemausfall“
A und B unabhängige Ereignisse; Multiplikationssatz für unabhängige Ereignisse:
$P(C) = P(A \cap B) = P(A) \cdot P(B) = 0,002$
Gesucht: $P(\overline{C})$
Möglichkeit 1:
$P(\overline{C}) = 1 - P(C) = 0,998$
Möglichkeit 2:
$\overline{C} = \overline{A \cap B} = (\overline{A} \cap \overline{B}) \cup (\overline{A} \cap B) \cup (A \cap \overline{B})$;

$$\begin{aligned} P(\overline{C}) &= P(\overline{A}) \cdot P(\overline{B}) + P(\overline{A}) \cdot P(B) + P(A) \cdot P(\overline{B}) \\ &= 0,95 \cdot 0,96 + 0,95 \cdot 0,04 + 0,05 \cdot 0,96 \\ &= 0,912 + 0,038 + 0,048 \\ &= 0,998 \end{aligned}$$

Lösung 4-49: *Kundenbesuche*
Ereignisse:
G = „M. nimmt Übergang Guben“;
F = „M. nimmt Übergang Forst“
W = „M. muss am Übergang sehr lange warten“
Gegeben:
$P(G) = 0,6 \qquad P(F) = 0,4 \qquad P(G \cap W) = 0,3$
Gesucht:
$P(W|G) = \dfrac{P(G \cap W)}{P(G)} = \dfrac{0,3}{0,6} = 0,5$

Lösung 4-50: *RealProfit*
Gegeben:
$P(X \leq 0) = 1 - 0.75 = 0.25$
Gesucht:

$$\begin{aligned} P(X \leq 0) &= P(X = -10000) + P(X = 0) \\ &= P(X = -10000) + 0,2 = 0.25 \\ &\Leftrightarrow P(X = -10000) = 0.05 \\ P(X = 20000) &= 1 - P(X = -10000) - P(X = 0) - P(X = 5000) \\ &- P(X = 10000) - P(X = 15000) \\ &= 0,2 \end{aligned}$$

Lösung 4-51: *Jeeps*
A = „Jeep 1 startet“, B = „Jeep 2 startet“
$P(\text{kein Jeep startet}) = P(\overline{A} \cap \overline{B}) = P(\overline{A}) \cdot P(\overline{B}) = (1 - P(A)) \cdot (1 - P(B)) = 0,2 \cdot 0,5 = 0,1$

Lösung 4-52: *Summe von Augenzahlen*
Summe der Augenzahlen für alle möglichen Kombinationen:

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

$$P(\{2,3\}|\text{Summe} = 5) = \frac{P(\{2,3\})}{P(\sum = 5)} = \frac{2/36}{4/36} = \frac{1}{2}$$

Lösung 4-53: *Fahrrad oder Straßenbahn*
Ereignisse:
$A =$ „Fritzi braucht mehr als 30 Min. bis in die Uni“
$B_1 =$ „Fritzi nimmt das Fahrrad“
$B_2 =$ „Fritzi nimmt die Straßenbahn“
Gegeben:

$$P(B_1) = 0,8; \qquad P(B_2) = 0,2; \qquad P(A|B_1) = 0,3 \qquad P(A|B_2) = 0,6$$

Gesucht:

$$\begin{aligned} P(B_2|A) &= \frac{P(A|B_2) \cdot P(B_2)}{P(A|B_1) \cdot P(B_1) + P(A|B_2) \cdot P(B_2)} \\ &= 0,\overline{333} \end{aligned}$$

Lösung 4-54: *Münzwurf*
Wir definieren die Ereignisse:

- H: Hinz fährt mit der Strassenbahn
- K: Kunz fährt mit der Strassenbahn
- M: Die Münze von Hinz zeigt Kopf
- M_1: Die Münze 1 von Kunz zeigt Zahl
- M_2: Die Münze 2 von Kunz zeigt Zahl

Die Wahrscheinlichkeit, dass beide mit der Straßenbahn fahren ist damit gegeben durch

$$\begin{aligned} P(H \cap K) &= P(H) \cdot P(K) \\ &= P(M) \cdot (1 - P(M_1) \cdot P(M_2)) \\ &= 0,5 \cdot (1 - 0,5 \cdot 0,5) = \frac{3}{8} \\ &= 37,5\% \end{aligned}$$

Im ersten Schritt haben wir die Unabhängigkeit der Entscheidungen von Hinz und Kunz verwendet, im zweiten Schritt die Gegenwahrscheinlichkeit (Wk. Strassenbahn zu fahren = 1- Wk. Fahrrad zu fahren), im dritten Schritt noch einmal die Unabhängigkeit der beiden Münzwürfe von Kunz.

Lösung 4-55: *Gangsterbande*
Ereignisse:
$D =$ „Donnerstag“; $\overline{D} =$ „nicht Donnerstag“
$Y =$ „Scotland Yard fasst Täter am selben Tag“,
$H =$ „Sherlock Holmes fasst Täter am selben Tag“
$G =$ „Täter am selben Tag im Gefängnis“
Gegeben:
$P(Y) = P(Y|D) = P(Y|\overline{D}) = 0,25; \quad P(H|D) = 0,00; \quad P(H|\overline{D}) = 0,35;$
$P(D) = 1/6; \quad P(\overline{D}) = 5/6$
Gesucht:

$$\begin{aligned} P(G|D) &= P(Y|D) = 0,25 \\ P(G|\overline{D}) &= P(Y \cup H|\overline{D}) \\ &= P(Y|\overline{D}) + P(H|\overline{D}) - P(Y \cap H|\overline{D}) \\ &= 0,25 + 0,35 - 0,25 \cdot 0,35 = 0,5125 \\ P(G) &= P(G|D) \cdot P(D) + P(G|\overline{D}) \cdot P(\overline{D}) \\ &= 0,25 \cdot 1/6 + 0,5125 \cdot 5/6 = 0,46875 \approx 0,47 \end{aligned}$$

Lösung 4-56: *Wochenendgrundstück*
Wir betrachten den Ereignisraum $\{D, F\} \times \{R, \overline{R}\} \times \{S, \overline{S}\}$ mit der Potenzmenge als Menge der Ereignisse.
Die Zufallsvariable M (Münze) gibt an, ob Familie Sommer die Fähre ($M = F$) oder den Damm ($M = D$) nutzt. Die Zufallsvariable W (Wetter) gibt an, ob es regnet ($W = R$) oder nicht ($W = \overline{R}$). Die Zufallsvariable S gibt an, ob es bei der Fahrt zu einem Stau kommt ($S = St$) oder
nicht ($S = \overline{St}$).
Familie Sommer steht genau dann nicht im Stau, wenn sie gar nicht losfährt (bei Regen) oder bei gutem Wetter ohne Stau über den Damm bzw. die Fähre nach Rügen fahren kann. D.h. die folgende Wahrscheinlichkeit ist gesucht:

$$P[\{W = R \text{ oder } (W = \overline{R} \text{ und } S = \overline{St})\}]$$

Wir berechnen:

$$\begin{aligned}
&P[\{W = R \text{ oder } (W = \overline{R} \text{ und } S = \overline{St})\}] \\
&= P[\{W = R\}] + P[\{W = \overline{R} \text{ und } S = \overline{St}\}] \\
&- P[\{W = R \text{ und } W = \overline{R} \text{ und } S = \overline{St}\}].
\end{aligned}$$

Offenbar ist der Schnitt in der letzten Wahrscheinlichkeit leer, denn es kann nicht gleichzeitig regnen und nicht regnen. Also vereinfacht sich die gesuchte Wahrscheinlichkeit zu

$$= P[\{W = R\}] + P[\{W = \overline{R} \text{ und } S = \overline{St}\}].$$

Jetzt benutzen wir die Unabhängigkeit von Wetter W und Stau S, um diesen Ausdruck noch weiter zu vereinfachen:

$$= P[\{W = R\}] + P[\{W = \overline{R}\}] \cdot P[\{S = \overline{St}\}].$$

Zuletzt wenden wir noch den Satz von der totalen Wahrscheinlichkeit an, um die Wahrscheinlichkeit von Stau auszurechnen und erhalten:

$$\begin{aligned}
&= P[\{W = R\}] \\
&P[\{W = \overline{R}\}] \cdot (P[\{S = \overline{St}\} \mid \{M = D\}] \\
&\cdot P[\{M = D\}] + P[\{S = \overline{St}\} \mid \{M = F\}] \cdot P[\{M = F\}]) \\
&= 0.2 + 0.8 \cdot (0.5 \cdot 0,75 + 0.5 \cdot 0,9) \\
&= 0,86.
\end{aligned}$$

Lösung 4-57: *Elemente eines Ereignisraumes*

a) Für eine Zerlegung von S muss u.a. gelten: $A_1 \cup \ldots \cup A_n = S$, d.h. es müsste $A \cup B = S$ sein und somit $P(A \cup B) = 1$. Da $P(A \cup B) = 3/4$ ist, gilt diese Behauptung nicht. Außerdem müssten die Ereignisse A und B disjunkt sein, was nicht der Fall ist (siehe c).

b) Wenn A und B komplementär wären, müsste gelten: $A \cup B = S$ und somit $P(A \cup B) = 1$. Da $P(A \cup B) = 3/4$ ist, gilt diese Behauptung nicht.

c) Für disjunkte Ereignisse gilt $A \cap B = \varnothing$ und somit $P(A \cap B) = 0$ (Berechnung unter d).

d) $P(A \cup B) = P(A) + P(B) - P(A \cap B)$
$P(A \cap B) = P(A) + P(B) - P(A \cup B) = 1/2 + 1/2 - 3/4 = 1/4$
$P(A \cap B) = P(A) \cdot P(B) = 1/2 \cdot 1/2 = 1/4$
oder
$P(A|B) = P(A|\overline{B})$ mit $P(A|B) = P(A \cap B)/P(B) = (1/4)/(1/2) = 1/2$
und $P(A|\overline{B}) = P(A \cap B)/P(\overline{B}) = (1/4)/(1/2) = 1/2$
Die Ereignisse A und B sind unabhängig.

Lösung 4-58: *Fernsehshow*
$W = \{\text{weiße Kugel}\}; \quad U_i = \{\text{Urne i}\} \quad i = 1, 2$
Anwendung des Satzes der totalen Wahrscheinlichkeit auf jedes Verfahren:
$P(W) = P(W|U_1) \cdot P(U_1) + P(W|U_2) \cdot P(U_2)$
$P(U_1) = P(U_2) = 1/2$
1. Verfahren:
U_1: $6W$, $2S \quad \Rightarrow P(W|U_1) = 6/8$
U_2: $6W$, $10S \quad \Rightarrow P(W|U_2) = 6/16$
$P(W) = 6/8 \cdot 1/2 + 6/16 \cdot 1/2 = 9/16 = 0,5625$
2. Verfahren:
U_1: $6W$, $6S \quad \Rightarrow P(W|U_1) = 6/12$
U_2: $6W$, $6S \quad \Rightarrow P(W|U_2) = 6/12$
$P(W) = 6/12 \cdot 1/2 + 6/12 \cdot 1/2 = 1/2 = 0,5$

Lösung 4-59: *Verkehrsunfälle*
Wir betrachten den Ereignisraum $S := \{K, \overline{K}\} \times \{G, \overline{G}\}$ mit der Potenzmenge als Ereignismenge. Die Zufallsvariable V (Verletzung) gibt an, ob eine schwere Kopfverletzung eingetreten ist ($V = K$), oder nicht ($V = \overline{K}$). Die Zufallsvariable S (Sicherheit) gibt an, ob ein Gurt angelegt worden ist ($S = G$), oder nicht ($S = \overline{G}$). Gesucht ist die Wahrscheinlichkeit

$$P[\{S = \overline{G}\} \mid \{V = K\}] = \frac{P[\{S = \overline{G}\} \cap \{V = K\}]}{P[\{V = K\}]}.$$

Gegeben sind die Wahrscheinlichkeiten

$$\begin{aligned} P[\{S = G\}] &= 0,85 \\ P[\{V = K\} \mid \{S = \overline{G}\}] &= 0,62 \\ P[\{V = K\} \mid \{S = G\}] &= 0,08 \end{aligned}$$

Anwenden des Satzes von Bayes im ersten und der Definition der bedingten Wahrscheinlichkeit im zweiten Schritt ergibt:

$$P[\{S = \overline{G}\} \mid \{V = K\}]$$
$$= \frac{P[\{V = K\} \cap \{S = \overline{G}\}] \cdot P[\{S = \overline{G}\}]}{P[\{V = K\} \mid \{S = \overline{G}\}] \cdot P[\{S = \overline{G}\}] + P[\{V = K\} \mid \{S = G\}] \cdot P[\{S = G\}]}$$

Um die Wahrscheinlichkeit im Zähler mit den gegebenen Informationen zu berechnen, verwenden wir die Definition der bedingten Wahrscheinlichkeit:

$$\begin{aligned} &= \frac{P[\{V = K\} \mid \{S = \overline{G}\}] \cdot P[\{S = \overline{G}\}]}{P[\{V = K\} \mid \{S = \overline{G}\}] \cdot P[\{S = \overline{G}\}] + P[\{V = K\} \mid \{S = G\}] \cdot P[\{S = G\}]} \\ &= \frac{0,62 \cdot 0,15}{0,62 \cdot 0,15 + 0,08 \cdot 0,85} \\ &= \frac{93}{161} \approx 0,578. \end{aligned}$$

Wahrscheinlichkeitsrechnung-Verkehrsunfaelle.mp4
Lösung 4-60: *Lebenserwartung der US–Bürger*
X: „Alter“
Gegeben: $P(X \geq 25) = 0,98$ und $P(X \geq 65) = 0,8$
Gesucht: $P(X < 65|X \geq 25)$
Mit den gegebenen Wahrscheinlichkeiten kann berechnet werden:
$P(X \geq 65|X \geq 25) = P(\{X \geq 65\} \cap \{X \geq 25\})/P(X \geq 25)$
Da das Ergeignis $\{X \geq 65\}$ in dem Ereignis $\{X \geq 25\}$ enthalten ist, folgt für den Durchschnitt der beiden Ereignisse $\{X \geq 65\} \cap \{X \geq 25\} = \{X \geq 65\}$ und somit
$P(X \geq 65|X \geq 25) = P(\{X \geq 65\} \cap \{X \geq 25\})/P(X \geq 25) = P(X \geq 65)/P(X \geq 25) = 0,8/0,98 = 0,81632653.$
Es folgt: $P(X < 65|X \geq 25) = 1 - P(X \geq 65|X \geq 25) = 1 - 0,81632653 = 0,18367347$

Lösung 4-61: *Ausschussteile*
Anwendung des Satzes der totalen Wahrscheinlichkeit:

$$\begin{aligned} P(B) &= P(B|C)P(C) + P(B|\overline{C})(1 - P(C)) \\ &= 0,95 \cdot 0,05 + 0,1 \cdot 0,95 \\ &= 0,1425 \end{aligned}$$

$$\begin{aligned} P(B) &= P(B|A_1)P(A_1) + P(B|A_2)P(A_2) + P(B|A_3)P(A_3) \\ \Leftrightarrow P(B|A_2)P(A_2) &= P(B) - P(B|A_1)P(A_1) - P(B|A_3)P(A_3) \\ \Leftrightarrow P(B|A_2) &= [P(B) - P(B|A_1)P(A_1) - P(B|A_3)P(A_3)]/P(A_2) \\ &= (0,1425 - 0,8 \cdot 0,1 - 0,6 \cdot 0,05)/0,0422 \\ &= 0,0325/0,0422 = 0,77014 \end{aligned}$$

Lösung 4-62: *Unabhängige Ereignisse*
Wir betrachten den Ereignisraum S mit der Borelalgebra als Menge der Ereignisse. Entsprechend der Abbildung gilt: $P[A] > 0$ und $P[B_i] > 0$ für $i = 1, 2, 3, 4$. Nach Definition sind zwei Ereignisse C und D unabhängig, wenn gilt

$$P[C \cap D] = P[C] \cdot P[B].$$

- Zwei Ereignisse C und D mit $P[C] > 0$ und $P[D] > 0$, die disjunkt sind, sind immer abhängig:

  $$P[C \cap D] = 0 < P[C] \cdot P[D]$$

  .Es fallen also als Lösungen heraus: a) – g).

- Zwei Ereignisse, von denen eines das andere enthält, sind immer abhängig, wenn das „kleinere“ Ereignis eine positive Wahrscheinlichkeit hat und das grössere nicht Wahrscheinlichkeit 1: Für $B_3 \subset A$ folgt:

  $$P[A \cap B_3] = P[B_3] > P[B_3] \cdot P[A].$$

  Damit fällt i) als Lösung weg.

- Das Ereignis B_2 liegt mit seiner halben Fläche in A. Auf A liegt $\frac{1}{4}$ der Wahrscheinlichkeitsmasse, insbesondere ist die Wahrscheinlichkeitsmasse kleiner als $\frac{1}{2}$. Deshalb gilt

  $$P[A \cap B_2] = \frac{P[A] \cdot P[B_2]}{2} < P[A] \cdot P[B_2].$$

  Damit fällt h) als Lösung weg.

- B_4 und der Ereignisraum S sind durch konzentrische Kreise dargestellt. Der Anteil von $A \cap B_4$ an B_4 ist somit gleich dem Anteil von A an S. Folglich gilt:

  $$P[A|B_4] = P[A \cap B_4]/P[B_4] = P[A]/P[S] = P[A].$$

  A und B_4 sind also unabhängig (Fall j).

Lösung 4-63: *Eignungstest*
Die Ereignisse A_1 = „Bewerber besteht Eignungstest“ und A_2 = „Bewerber besteht Eignungstest nicht“ bilden eine vollständige Zerlegung des Ereignisraums S. Gegeben ist $P(A_1) = 0,25$. Aufgrund von $A_2 = \overline{A_1}$ folgt $P(A_2) = 1 - P(A_1) = 0,75$.
Ferner sind ein zufälliges Ereignis B = „Bewerber ist für die Tätigkeit geeignet“ und die bedingten Wahrscheinlichkeiten dieses Ereignisses $P(B|A_1) = 0,95$ und $P(B|A_2)$=0,10 gegeben.
Gesucht wird die bedingte Wahrscheinlichkeit $P(A_1|B)$. Diese lässt sich nach dem Theorem von Bayes

$$P(A_1|B) = \frac{P(A_1 \cap B)}{P(B)} = \frac{P(B|A_1)P(A_1)}{P(B|A_1)P(A_1) + P(B|A_2)P(A_2)}$$

berechnen. Für $P(B)$ resultiert:

$$P(B) = P(B|A_1) \cdot P(A_1) + P(B|A_2) \cdot P(A_2) = 0,3125$$

$\rightarrow P(A_1|B) = 0,76$

Lösung 4-64: *Spiel 4 aus 20*

Es gibt $\binom{4}{2}$ Möglichkeiten 2 aus den 4 geratenen Kugeln zu ziehen und $\binom{16}{2}$ für die übrigen. Da jede Kombination mit jeder verknüpft werden kann gibt es $\binom{4}{2}\binom{16}{2}$ günstige Möglichkeiten. Also ist die Wahrscheinlichkeit für 2 richtige:

$$\frac{\binom{4}{2}\binom{16}{2}}{\binom{20}{4}} = 14,86\%$$

Lösung 4-65: *Biergärten*

Gegeben:

$A = \{\text{Gast aus Biergarten } A\}, \quad P(A) = 0,6$
$B = \{\text{Gast aus Biergarten } B\}, \quad P(B) = 0,3$
$C = \{\text{Gast aus Biergarten } C\}, \quad P(C) = 0,1$
$U = \{\text{unzufriedener Gast}\}$ mit
$P(U|A) = 0,1;\ P(U|B) = 0,4;\ P(U|C) = 0,7$

Gesucht:

$P(B|U)$

Theorem von Bayes:

$$P(B|U) = \frac{P(U|B)P(B)}{P(U|A)P(A) + P(U|B)P(B) + P(U|C)P(C)}$$

$$\begin{aligned} P(U|A)P(A) &= 0,1 \cdot 0,6 = 0,06 \\ P(U|B)P(B) &= 0,4 \cdot 0,3 = 0,12 \\ P(U|C)P(C) &= 0,7 \cdot 0,1 = 0,07 \\ \sum P(U|i)P(i) &= 0,25 \\ P(B|U) &= 0,12/0,25 = 0,48 \end{aligned}$$

5 Zufallsvariablen

5.1 Univariate Zufallsvariablen

Lösung 5-1: *Kinder*

a) Zuerst werden die möglichen Merkmalsausprägungen der Zufallsvariable bestimmt. Die kleinste Summe an Kinder bei drei Ziehungen ohne Zurücklegen ist zwei, z.B. (P_3, P_4, P_6), und die größte Summe ist zehn, z.B. (P_1, P_2, P_5).

   Für jede Merkmalsausprägung zwischen zwei und zehn können wir die Wahrscheinlichkeit des Auftretens mit Hilfe der Wk. nach Laplace bestimmen. Zunächst halten wir fest, dass es

   $$K(6;3) = \binom{6}{3} = \frac{6!}{3! \cdot 3!} = \frac{4 \cdot 5 \cdot 6}{1 \cdot 2 \cdot 3} = 4 \cdot 5 = 20$$

   Möglichkeiten gibt (Ziehen ohne Zurücklegen und ohne Beachtung der Reihenfolge, da die Addition kommutativ ist), 3 Personen aus 6 Personen zu ziehen.

   Für jede Merkmalsausprägung lässt sich nun angeben (ohne Berücksichtigung der Reihenfolge) wieviele Möglichkeiten es gibt sie zu erhalten:

x	Ereignisse	$P(X = x)$
2	(P_3, P_4, P_6)	1/20
3	(P_2, P_3, P_4), (P_2, P_3, P_6), (P_3, P_4, P_5), (P_3, P_5, P_6)	4/20
4	(P_2, P_3, P_5), (P_2, P_4, P_6), (P_4, P_5, P_6)	3/20
5	(P_2, P_4, P_5), (P_2, P_5, P_6)	2/20
6	–	0/20
7	(P_1, P_3, P_4), (P_1, P_5, P_6)	2/20
8	(P_1, P_2, P_3), (P_1, P_3, P_4), (P_1, P_3, P_6)	3/20
9	(P_1, P_2, P_4), (P_1, P_2, P_6), (P_1, P_4, P_5), (P_1, P_5, P_6)	4/20
10	(P_1, P_2, P_5)	1/20

b) Um den Wert der Verteilungsfunktion an einer Stelle s zu erhalten, genügt es, die Werte der Wahrscheinlichkeitsfunktion bis zu dieser Stelle s zu addieren, da es sich um eine diskrete Wahrscheinlichkeitsverteilung handelt. Z.B. ist der Wert der Verteilungsfunktion F an der Stelle 5 gegeben durch

$$F(5) = f(2) + f(3) + f(4) + f(5) = \frac{1}{2}.$$

Es gilt ebenfalls

$$F(6) = f(2) + f(3) + f(4) + f(5) = \frac{1}{2},$$

da sich zwischen 5 und 6 kein Ergebnis befindet, das mit positiver Wahrscheinlichkeit auftritt. Die Verteilungsfunktion hat also die für diskrete Verteilungen typische „Sprungstellen" an den Stellen, an denen Ergebnisse mit positiver Wahrscheinlichkeit auftreten. Die Ergebnisse für die anderen Stellen ergeben sich mit analoger Rechnung zu

$$F(x) = \begin{cases} 0 & \text{für} \quad x < 2, \\ 1/20 & \text{für} \quad 2 \leq x < 3, \\ 5/20 & \text{für} \quad 3 \leq x < 4, \\ 8/20 & \text{für} \quad 4 \leq x < 5, \\ 10/20 & \text{für} \quad 5 \leq x < 7, \\ 12/20 & \text{für} \quad 7 \leq x < 8, \\ 15/20 & \text{für} \quad 8 \leq x < 9, \\ 19/20 & \text{für} \quad 9 \leq x < 10, \\ 1 & \text{für} \quad 10 \leq x. \end{cases}$$

d) Die gesuchten Wahrscheinlichkeiten werden wir nun unter Rückgriff auf die Verteilungsfunktion berechnen. Wir erinnern uns daher nocheinmal an die Definition der Verteilungsfunktion F als

$$F(x) = P[\{X \leq 4\}].$$

Gesucht ist die Wahrscheinlichkeit $P[\{X \leq 4\}]$, die wir direkt aus der Tabelle der Verteilungsfunktion in Aufgabenteil b) als

$$P[\{X \leq 4\}] = F(4) = \frac{8}{20} = \frac{2}{5}$$

abgelesen werden kann. Nun wird die Wahrscheinlichkeit $P[\{X > 8\}]$ gesucht. Um diese Wahrscheinlichkeit auf die Verteilungsfunktion zurückzuführen, verwenden wir den Trick der Berechnung über die Gegenwahrscheinlichkeit:

$$P[\{X > 8\}] = 1 - P[\{X \leq 8\}] = 1 - \frac{15}{20} = \frac{5}{20} = \frac{1}{4}.$$

Um die Wahrscheinlichkeit $P[\{3 < X < 9\}]$ zu berechnen, greifen wir wieder auf die Verteilungsfunktion zurück:

$$\begin{aligned} P[\{3 < X < 9\}] &= P[\{X < 9\}] - P[\{X \leq 3\}] \\ &= P[\{X \leq 8\}] - P[\{X \leq 3\}] \\ &= F[8] - F[3] \\ &= \frac{15}{20} - \frac{5}{20} = \frac{10}{20} = \frac{1}{2}. \end{aligned}$$

Lösung 5-2: *Lostrommel*

a) X: "Gewinn"; $P(X = 5) = 5 \cdot 10/1000 = 0,05$; $P(X = 2) = 4 \cdot 100/1000 = 0,4$; $P(X = 0) = 1 - 0,05 - 0,4 = 0,55$

x	0	2	5
$f(x)$	0,55	0,4	0,05

b)

$$F(x) = \begin{cases} 0 & \text{für} \quad x < 0 \\ 0,55 & \text{für} \quad 0 \leq x < 2 \\ 0,95 & \text{für} \quad 2 \leq x < 5 \\ 1 & \text{für} \quad 5 \leq x \end{cases}$$

Lösung 5-3: *Ampeln*

A_k ={Die k–te Ampel steht auf grün}; $P(A_k) = 0,5$; Ereignisse sind unabhängig;

$\overline{A}_k$ ={Die k–te Ampel steht auf rot}; $P(\overline{A}_k) = 0,5$; $k = 1,2,3,4$

a) Auto fährt an keiner Ampel vorbei: $P(\overline{A}_1) = 0,5$; $A_1 \cap \overline{A}_2$ ={Auto fährt an 1. Ampel vorbei und muss an 2. Ampel halten}, $P(A_1 \cap \overline{A}_2) = 0,5 \cdot 0,5 = 0,25$; analog folgt: $P(A_1 \cap A_2 \cap \overline{A}_3) = 0,125$; $P(A_1 \cap A_2 \cap A_3 \cap \overline{A}_4) = 0,0625$; $P(A_1 \cap A_2 \cap A_3 \cap A_4) = 0,0625$;

b) X: "Anzahl der Verkehrsampeln, an denen das Auto bis zum ersten Halt vorbeifährt"; diskrete Zufallsvariable

x	0	1	2	3	4
$f(x)$	0,5	0,25	0,125	0,0625	0,0625

Lösung 5-4: *Würfelspiel*

Z ={Erscheinen der gesetzten Zahl beim Werfen eines Würfels};

$P(Z) = 1/6; P(\overline{Z}) = 5/6$

Ereignis	$P(E_i)$	Spielgewinn $X = x$	$f(x)$
$E_1 = \overline{Z} \cap \overline{Z} \cap \overline{Z}$	125/216	$x_1 = -1$	125/216
$E_2 = Z \cap \overline{Z} \cap \overline{Z}$	25/216		
$E_3 = \overline{Z} \cap Z \cap \overline{Z}$	25/216	$x_2 = 1$	75/216
$E_4 = \overline{Z} \cap \overline{Z} \cap Z$	25/216		
$E_5 = Z \cap Z \cap \overline{Z}$	5/216		
$E_6 = Z \cap \overline{Z} \cap Z$	5/216	$x_3 = 2$	15/216
$E_7 = \overline{Z} \cap Z \cap Z$	5/216		
$E_8 = Z \cap Z \cap Z$	1/216	$x_4 = 3$	1/216

Wegen der Unabhängigkeit der einzelnen Würfelergebnisse Anwendung des Multiplikationssatzes für unabhängige Ereignisse zur Berechnung der $P(E_j)$. Die E_j sind disjunkt; Anwendung des Axioms 3 zur Berechnung der $f(x_i)$. X ist eine diskrete Zufallsvariable.

$E(X) = -0,079 EUR$

Lösung 5-5: *Diskrete Zufallsvariable*

Zunächst prüfen wir, ob die Aufgabe wohlgestellt ist, d.h. ob überhaupt eine Wahrscheinlichkeitsfunktion vorliegt. Dazu ist zu prüfen, ob sie nichtnegativ ist und ihre Werte sich zu 1 aufsummieren. Aus der Definition ist klar, dass die Funktion, von der zu prüfen ist, ob sie eine Wahrscheinlichkeitsfunktion ist, nur nichtnegative Werte annimt. Wir berechnen

$$f(0) + f(1) + f(2) + f(3) + f(4) = \frac{4+5+8+13+20}{50} = 1.$$

Es liegt also in der Tat eine Wahrscheinlichkeitsfunktion vor.

a) Gesucht ist die Wahrscheinlichkeit $P[\{X = 2\}]$. Mit der Definition der Wahrscheinlichkeitsfunktion gilt

$$P[\{X = 2\}] = f(2) = \frac{8}{50} = \frac{4}{25}0,16.$$

b) Gesucht ist die Wahrscheinlichkeit $P[\{X < 2\}]$. Mit der Definition der Wahrscheinlichkeitsfunktion gilt

$$P[\{X < 2\}] = P[\{X = 0\}]+P[\{X = 1\}] = f(0)+f(1) = \frac{4+5}{50} = 0,18.$$

c) Gesucht ist die Wahrscheinlichkeit

$$P[\{X \leq 2\}] = P[\{X = 0\}]+P[\{X = 1\}]+P[\{X = 2\}] = \frac{4+5+8}{50} = 0,34.$$

d) Gesucht ist die Wahrscheinlichkeit

$$P[\{X > 3\}] = P[\{X = 4\}] = f(4) = \frac{20}{50} = 0,4$$

e) Es gilt

$$P[\{X < 5\}] = f(0) + f(1) + f(2) + f(3) + f(4) = 1.$$

Zufallsvariablen-Diskrete Zufallsvariable.mp4

Lösung 5-6: *Intervall–Bestimmung*

a) $F(a) = a/6 - 1/3 \doteq 0$; $a = 2$; $F(b) = b/6 - 1/3 \doteq 1$; $b = 8$

b)

$$\frac{dF(x)}{dx} = f(x) = \begin{cases} 1/6 & \text{für} \quad 2 \leq x \leq 8 \\ 0 & \text{sonst} \end{cases}$$

c) $P(6 \leq X \leq 8) = \int_6^8 (1/6)dt = 1/3$; $P(X = 5) = P(5 \leq X \leq 5) = 0$

Lösung 5-7: *Fachliteratur*

a) Der Abbildung entnehmen wir, dass die Dichtefunktion im Intervall $[1, 4]$ als affine Funktion dargestellt werden kann, d.h. für $x \in [1, 4]$ gilt

$$f(x) = m \cdot x + b,$$

wobei $m \in \mathbb{R}$ die Steigung und $b \in \mathbb{R}$ den Achsenabschnitt der affinen Funktion angibt. Wir verwenden die Formel für die Steigung m der Geraden, indem wir die beiden bekannten Punkte $(1/a)$ und $(4/0)$ verwenden ($a \in \mathbb{R}_+$ ist unbekannt):

$$m = \frac{0 - a}{4 - 1} = -\frac{a}{3}.$$

Für den Achsenabschnitt gilt damit

$$b = f(4) + \frac{a}{3} \cdot 4 = \frac{4}{3} \cdot a.$$

In Abhängigkeit vom unbekannten Parameter $a \in \mathbb{R}_+$ gilt damit für $x \in [1, 4]$,

$$f(x) = -\frac{a}{3} \cdot x + \frac{4}{3}a.$$

Aus der Skizze ist ebenfalls klar, dass

$$f(x) = 0$$

für alle $x \notin [1, 4]$. Um a zu bestimmen, müssen wir es so wählen, dass f eine Wahrscheinlichkeitsdichte wird, d.h. nichtnegativ ist und sich zu

1 integriert. Die Nichtnegativität ist klar. Für die Normierung rechnen wir

$$\begin{aligned} 1 &= \int_{-\infty}^{\infty} f(x)dx \\ &= \int_1^4 -\frac{a}{3} \cdot x + \frac{4}{3}adx \\ &= \left[-\frac{a}{6} \cdot x^2 + \frac{4}{3}ax + c\right]_{x=1}^4 \\ &= -\frac{a}{6} \cdot (4^2 - 1^2) + \frac{12}{3}a \\ &= -\frac{5}{2}a + 4a = \frac{3}{2}a. \end{aligned}$$

Damit ergibt sich

$$a = \frac{2}{3}.$$

Zusammengefasst erhalten wir also

$$f(x) \quad = \quad \begin{cases} \frac{8}{9} - \frac{2}{9}x & \text{für} \quad 1 \leq x \leq 4 \\ 0 & \text{sonst} \end{cases}$$

In der obigen Rechnung haben wir zudem die allg. Stammfunktion F von f berechnet, die noch von einer Konstante c abhängt. Diese wollen wir nun so wählen, dass die Verteilungsfunktion stetig ist (d.h., dass sie keine „Sprungstellen“ hat). Offenbar muss gelten

$$F(x) = 0$$

für $x \leq 1$ und

$$F(x) = 1$$

für $x \geq 4$. Damit muss auch für die Stammfunktion $S_c(x)$ der Dichte gelten, dass $S_c(1) = 0$. Wir berechnen

$$\begin{aligned} 0 &= S_c(1) \\ &= -\frac{2}{18} \cdot 1^2 + \frac{8}{9} + c \\ &= \frac{7}{9} + c. \end{aligned}$$

Also muss gelten

$$c = -\frac{7}{9},$$

damit F an der Stelle 1 stetig ist. Die Stetigkeit an der Stelle 4 ergibt sich aus der Wahl von a, die die Normierung der Dichtefunktion garantiert. Zusammengefasst ergibt sich

$$F(x) = \begin{cases} 0 & \text{für} \quad x < 1, \\ -\frac{1}{9}x^2 + \frac{8}{9}x - \frac{7}{9} & \text{für} \quad 1 \leq x < 4, \\ 1 & \text{für} \quad 4 \leq x. \end{cases}$$

b) Um den Erwartungswert zu berechnen, verwenden wir die Definition aus dem Foliensatz zu Zufallsvariablen

$$\begin{aligned} E[X] &= \int_{-\infty}^{\infty} x \cdot f(x)dx \\ &= \int_1^4 x \cdot (\frac{8}{9} - \frac{2}{9}x)dx \\ &= \int_1^4 \frac{8}{9}x - \frac{2}{9}x^2 dx \\ &= [\frac{4}{9}x^2 - \frac{2}{27}x^3 + c]_{x=1}^4 \\ &= \frac{64}{9} - \frac{128}{27} - \frac{4}{9} + \frac{2}{27} \\ &= \frac{60}{9} - \frac{126}{27} \\ &= \frac{20}{3} - \frac{14}{3} = \frac{6}{3} = 2. \end{aligned}$$

Für die Varianz von X gilt mit der Definition aus dem Foliensatz

$$\begin{aligned} Var[X] &= \int_{-\infty}^{\infty} x^2 \cdot f(x)dx - E[X]^2 \\ &= \int_1^4 x^2 \cdot \left(\frac{8}{9} - \frac{2}{9}x\right) dx - 2^2 \\ &= \int_1^4 \frac{8}{9}x^2 - \frac{2}{9}x^3 dx - 2^2 \\ &= [\frac{8}{27}x^3 - \frac{1}{18}x^4]_{x=1}^4 - 2^2 \\ &= \frac{8 \cdot 4^3}{27} - \frac{4^4}{18} - \frac{8 \cdot 1^3}{27} + \frac{1^4}{18} - 2^2 \\ &= \frac{8 \cdot (4^3 - 1)}{27} - \frac{4^4 - 1}{18} - 2^2 \\ &= \frac{8 \cdot 63}{27} - \frac{255}{18} - 2^2 \\ &= \frac{56}{3} - \frac{85}{6} - 2^2 \\ &= 4,5 - 4 = 0,5. \end{aligned}$$

c) Um die gesuchten Wahrscheinlichkeiten zu berechnen, verwenden wir unsere Kenntnis der Verteilungsfunktion. Zunächst berechnen wir

$$P[\{X \leq 2\}] = F(2) = -\frac{4}{9} + \frac{16}{9} - \frac{7}{9} = \frac{5}{9}.$$

Für die nächste gesuchte Wahrscheinlichkeit gilt

$$P[\{2 \leq X \leq 3\}] = F(3) - F(2) = \frac{1}{3}.$$

Für die letzte gesuchte Wahrscheinlichkeit gilt

$$\begin{aligned} P[\{X \geq 3\}] &= 1 - P[\{X < 3\}] \\ &= 1 - P[\{X \leq 3\}] \\ &= 1 - F(3) = \frac{1}{9}, \end{aligned}$$

wobei wir im zweiten Schritt verwendet haben, dass eine stetige Verteilung vorliegt und daher jeder Punkt eine Masse von 0 besitzt.

Lösung 5-8: *Rechteckverteilung*

a)

$$f(x) = \begin{cases} \frac{1}{8} & \text{für} \quad -2 \leq x \leq 6 \\ 0 & \text{sonst} \end{cases}$$

$$F(x) = \begin{cases} 0 & \text{für} \quad x < -2 \\ \frac{x+2}{8} & \text{für} \quad -2 \leq x < 6 \\ 1 & \text{für} \quad 6 \leq x \end{cases}$$

b) $E(X) = 2$; $Var(X) = 5,333$

c) $P(X \leq 0) = 1/4$; $P(X \leq |1|) = 1/4$; $P(X \leq 2|X \text{ positiv}) = 1/3$

Lösung 5-9: *Dichtefunktion*

a) Für eine Dichtefunktion muss gelten

$$f(x) \geq 0, \quad \int_{-\infty}^{+\infty} f(x)\, dx = 1.$$

Da $f(x) = 0$ für $x \notin [0, 2]$, genügt es für die erste Eigenschaft, zu zeigen, die Nullstellen der nach oben geöffnete Parabel beide links oder beide rechts des Intervalls $(0, 2)$ liegen. Wir verwenden die $p - q$ Formel, um diese Nullstellen zu berechnen und erhalten eine doppelte Nullstelle bei

$$x = 2 + \sqrt{2^2 - 4} = 2.$$

Damit gilt also sogar, dass die Parabel global nichtnegativ ist. Für die Normierung berechnen wir

$$\int_{-\infty}^{\infty} f(x)dx = \int_0^2 \frac{3}{8}(4 - 4x + x^2)dx$$
$$= \frac{3}{8}\left[4x - 2x^2 + \frac{1}{3}x^3 + c\right]_{x=0}^2$$
$$\frac{3}{8}\left(8 - 8 + \frac{8}{3}\right) = 1.$$

Also sind beide Eigenschaften erfüllt.

b) Wir berechnen die Verteilungsfunktion mithilfe der Definition im Foliensatz „Zufallsvariablen“für $x \in [0, 2]$:

$$\begin{aligned} F(x) &= \int_{-\infty}^{x} f(s)ds \\ &= \int_0^x \frac{3}{8}(4 - 4s + s^2)ds \\ &= \frac{3}{8}\left[4s - 2s^2 + \frac{1}{3}s^3 + c\right]_{s=0}^x \\ &= \frac{3}{8} \cdot \left(4x - 2x^2 + \frac{1}{3}x^3\right). \end{aligned}$$

Für $x \leq 0$ gilt $F(x) = 0$, während $F(x) = 1$ für $x \geq 2$. Zusammengefasst gilt also:

$$F(x) = \begin{cases} 0 & \text{für} \quad x < 0, \\ \frac{3}{8}\left(4x - 2x^2 + \frac{1}{3}x^3\right) & \text{für} \quad 0 \leq x < 2, \\ 1 & \text{für} \quad 2 \leq x. \end{cases}$$

c) Wir berechnen den Erwartungswert mit der Definition aus dem Foliensatz „Zufallsvariablen“:

$$\begin{aligned} E[X] &= \int_{-\infty}^{\infty} x \cdot f(x)dx \\ &= \int_0^2 x \cdot \left(\frac{3}{8}(4 - 4x + x^2)\right) dx \\ &= \frac{3}{8}\int_0^2 (4x - 4x^2 + x^3)dx \\ &= \frac{3}{8}[2x^2 - \frac{4}{3}x^3 + \frac{1}{4}x^4]_{x=0}^2 \\ &= \frac{3}{8} \cdot (8 - \frac{32}{3} + \frac{16}{4}) \\ &= \frac{1}{2}. \end{aligned}$$

Analog gilt für die Varianz

$$\begin{aligned}
Var(X) &= \int_{-\infty}^{\infty} x^2 \cdot f(x)dx - E[X]^2 \\
&= \int_0^2 x^2 \cdot \left(\frac{3}{8}(4-4x+x^2)\right) dx - \frac{1}{4} \\
&= \frac{3}{8}\int_0^2 (4x^2-4x^3+x^4)dx - \frac{1}{4} \\
&= \frac{3}{8}[\frac{4}{3}x^3 - x^4 + \frac{1}{5}x^5]_{x=0}^2 - \frac{1}{4} \\
&= \frac{3}{8}\cdot(\frac{4}{3}2^3 - 2^4 + \frac{1}{5}2^5) - \frac{1}{4} = 0,15.
\end{aligned}$$

Zufallsvariablen-Dichtefunktion.mp4

Lösung 5-10: *Konstanten*

a) $a = 7/117$; $b = 2/13$

b)

$$F(x) = \begin{cases} 0 & \text{für} \quad x < 0 \\ \frac{7}{351}x^3 & \text{für} \quad 0 \leq x < 3 \\ \frac{-1}{13}x^2 + x - \frac{23}{13} & \text{für} \quad 3 \leq x < 4 \\ 1 & \text{für} \quad 4 \leq x \end{cases}$$

Lösung 5-11: *Herstellung eines Gutes*

a) $E(Z) = 6000$ EUR; $Var(Z) = 18000$ [EUR]^2

b) $E(Y) = 3250$ EUR; $Var(Y) = 4500$ [EUR]^2

c) G: "Gewinn"; $E(G) = 2750$ EUR; $Var(G) = 4500$ [EUR]^2

Lösung 5-12: *Platten*

X: "Länge einer Platte"; Y: "Breite einer Platte"; $X \cdot Y$: "Fläche einer Platte"

$X\backslash Y$	5	6	f(x)
8	0,1	0,1	0,2
10	0,6	0,2	0,8
$f(y)$	0,7	0,3	1,0

$E(X \cdot Y) = 50,8\text{mm}^2$

Lösung 5-13: *Umweltschützer*

a) $P(X \geq 4) = 3/4$

b) $E(X) = 4$ Fässer;
$Var(X) = \frac{1}{2}$ Fässer^2

c)

$X\backslash Y$	0	1	2	$f(x)$
3	2/16	1/16	1/16	1/4
4	4/16	4/16	0	2/4
5	3/16	1/16	0	1/4
$f(y)$	9/16	6/16	1/16	1,0

d) $Cov(x, Y) = -1/8$

e) nein

f) Z: "Anzahl der an zwei aufeinanderfolgenden Tagen in einer Region gefundenen Fässer";
$Z = X + Y$; $E(Z) = 4,5$ U: "Erlös von 2 Fahrten an zwei aufeinanderfolgenden Tagen in einer Region";
$U = 20 + 5 \cdot Z$; $E(U) = 42,50$ EUR

Lösung 5-14: *Maschinenbauunternehmen*
Bezeichnet X die zufällige Anzahl der abgesetzten Anlagen, so ergibt sich die Zufallsvariable G, die den Gewinn in Mio. EUR (bzw. den Verlust im Fall von Realisationen kleiner als Null) beschreibt, zu

$$G(X) = 1 \cdot X - (1 + 0,5 \cdot X) = 0,5 \cdot X - 1$$

mit der Wahrscheinlichkeitsverteilung

Anlagenzahl x	0	1	2	3	4	5
$P(X = x)$, $P(G = g)$	0,05	0,15	0,25	0,30	0,15	0,10
Verlust/Gewinn $g = G(x)$	-1,0	-0,5	0,0	0,5	1,0	1,5

$E(G) = -1 \cdot 0,05 + (-0,5) \cdot 0,15 + 0,0 \cdot 0,25 + 0,5 \cdot 0,3 + 1,0 \cdot 0,15 + 1,5 \cdot 0,1 = 0,325$

oder

$$E(X) = 1 \cdot 0,15 + 2 \cdot 0,25 + 3 \cdot 0,3 + 4 \cdot 0,15 + 5 \cdot 0,1 = 2,65$$

$$E(G) = 0,5 \cdot 2,65 - 1 = 0,325$$

Der erwartete Gewinn der Abteilung beträgt 325.000 EUR.

Lösung 5-15: *Auslastung der Schiffe*
$X =$ Kosten, $E(X) = T_1 \cdot K_1 + T_2 \cdot K_2 + T_3 \cdot K_3$
$T_1 = 65, T_2 = 45; T_3 = 95, K_1 = 1000, K_2 = 1200, K_3 = 700, E(X) = 185500$

Lösung 5-16: *MegaShop*
$E(X) = 1000 \cdot \dfrac{1}{x} + 500 \cdot \dfrac{4}{x} + 20 \cdot \dfrac{100}{x} + 0 \cdot \dfrac{x - 105}{x} = \dfrac{5000}{x}$
$E(X) = 5 \rightarrow x = 1000$

Lösung 5-17: *Dichtefunktion einer Zufallsvariablen*
Die Verteilung ist eine Gleichverteilung auf $[a; b] = [-1; 3]$:

$$1 = \int_{-1}^{3} f(x)dx = [ax]_{-1}^{3} = 4a \rightarrow a = 1/4 = 0,25$$

$$P(X > 0) = \int_{0}^{3} f(x)dx = [ax]_0^3 = 3a = 3/4 = 0,75$$

.

Lösung 5-18: *Konstante a*

$$\begin{aligned} 1 = \int_0^1 f(x)dx &= \int_0^1 \{ax^2(1-x)\}dx = a\int_0^1 \{x^2 - x^3\} \\ &= a\left[\frac{x^3}{3} - \frac{x^4}{4}\right]_0^1 = a\left[\frac{1}{3} - \frac{1}{4}\right] = \frac{a}{12} \\ a &= 12 \end{aligned}$$

Lösung 5-19: *Fernsehsendung*
$G =$ Gewinn; R : richtige Antwort F : falsche Antwort
Für jede Runde gilt: $P(R) = 0,2$; $P(F) = 0,8$.

Runde	Antwort	Gewinn	Wahrscheinlichkeit P_i	$G \cdot P_i$
1	F	0	$P(F_1) = 0,8$	0
2	F	100	$P(R_1 \cap F_2) = 0,2 \cdot 0,8$ $= 0,16$	16
3	F	200	$P(R_1 \cap R_2 \cap F_3)$ $= 0,2 \cdot 0,2 \cdot 0,8 = 0,032$	6,4
4a	F	300	$P(R_1 \cap R_2 \cap R_3 \cap F_4)$ $= 0,2 \cdot 0,2 \cdot 0,2 \cdot 0,8 = 0,0064$	1,92
4b	R	400	$P(R_1 \cap R_2 \cap R_3 \cap R_4)$ $= 0,2 \cdot 0,2 \cdot 0,2 \cdot 0,2 = 0,0016$	0,64

$E(G) = \sum_i G_i \cdot P_i = 24,96$ EUR

Lösung 5-20: *Zurückgelegte Strecke*
Zufallsvariable $X =$ „täglich zurückgelegte Strecke"
Erwartungswert $\mu = 140$ km; Varianz $\sigma^2 = 144$ (km^2); $\sigma = 12$ (km)
Da die Verteilung der Zufallsvariablen nicht bekannt ist, kann die gesuchte Wahrscheinlichkeit nicht exakt berechnet werden, sondern nur mittels der Tschebyschev–Ungleichung grob abgeschätzt werden:
$P(|X-\mu| > a) \leq \sigma^2/a^2$ mit $a > 0$ bzw. für $a = k\sigma$ folgt: $P|X-\mu| > k\sigma) \leq 1/k^2$ mit $k > 0$.
$a = 24$; $k = a/\sigma = 24/12 = 2$
$P(|X-140| > 24) \leq 144/576 = 0,25$ bzw. $P(|X-140| > 2\cdot 12) \leq 1/2^2 = 0,25$
Dies ist eine Abschätzung für die Wahrscheinlichkeit, dass die Zufallsvariable X einen Wert außerhalb des 2–fachen zentralen Schwankungsintervalls annimmt. Die Formulierung in der Frage „um höchstens 24 (km) vom Erwartungswert abweicht" impliziert jedoch, dass X Werte innerhalb eines zentralen Schwankungsintervalls annimmt:
$[\mu - a; \mu + a] = [\mu - k\sigma; \mu + k\sigma]$ mit $a = k\sigma$.
Dies ist das Komplementärereignis zu $|X - \mu| > a$, so dass das gesuchte Ergebnis wie folgt lautet:
$P(\mu - a \leq X \leq \mu + a) \geq 1 - \sigma^2/a^2$ bzw. $P(\mu - k\sigma \leq X \leq \mu + k\sigma) \geq 1 - 1/k^2$
$P(140 - 24 \leq X \leq 140 + 24) \geq 1 - 144/576 = 0,75$ bzw.
$P(140 - 2\cdot 12 \leq X \leq 140 + 2\cdot 12) \geq 1 - 1/2^2 = 0,75$

Lösung 5-21: *Glücksrad*
Zufallsvariable X: Punkt, an dem der Zeiger des Glücksrades stehen bleibt. Die stetige Zufallsvariable kann alle Werte des Intervalls $[0; 60]$ annehmen. X folgt der Rechteckverteilung:

$$f(x) \begin{cases} 1/60 & \text{für } 0 \leq x \leq 60 \\ 0 & \text{sonst} \end{cases}$$

$P(X = 14,08) = 0$, da die Wahrscheinlichkeit, dass eine stetige Zufallsvariable genau einen Wert annimmt, stets Null ist.

Lösung 5-22: *Zufallsvariable X*

$$\int_3^x \frac{1}{8}(t-3) = \frac{1}{8}\left[\frac{1}{2}t^2 - 3t\right]_3^x = \left(\frac{1}{16}x^2 - \frac{3}{8}x\right) - \left(\frac{9}{16} - \frac{9}{8}\right)$$
$$= \frac{1}{16}x^2 - \frac{3}{8}x + \frac{9}{16}$$

Die Verteilungsfunktion von X ist:

$$F(x) = \begin{cases} 0 & \text{für } x < 3 \\ \frac{1}{16}x^2 - \frac{3}{8}x + \frac{9}{16} & \text{für } 3 \leq x \leq 7 \\ 1 & \text{für } x > 7 \end{cases}$$

$P(X > 5) = 1 - P(X \leq 5) = 1 - F(5) = 1 - (25/16 - 15/8 + 9/16) = 1 - 4/16 = 0,75$

Lösung 5-23: *Feuerwehr*
Die erwartete quadrierte Fahrstrecke ist minimal, wenn sich die Feuerwehr an der Stelle aufstellt, die dem Erwartungswert der Zufallsvariablen $X = \{\text{Ort des nächsten Feuers}\}$ entspricht.
Herleitung:
$\min_c E[(x-c)^2]$
1. Ableitung: $-2\sum_x (x-c)P(x) = 0; c = \sum_x xP(x) = E(X)$

Punkt x	Wahrscheinlichkeit $P(x)$	$x \cdot P(x)$
-3	0,2	-0,6
-1	0,1	-0,1
0	0,1	0
1	0,4	0,4
2	0,2	0,4
$\sum$	1	0,1

Die Feuerwehr sollte sich an der Stelle $x = 0,1$ aufstellen.

Lösung 5-24: *Mautpflichtige Brücke*

$$\begin{aligned} P_{(a)}(X) &= 0,5 \cdot (0,05 + 2 \cdot 0,43 + 3 \cdot 0,27 + 4 \cdot 0,12 + 5 \cdot 0,09 + 6 \cdot 0,04) \\ &= 1,445 \\ P_{(b)}(X) &= 2 \cdot 0,5 + 0,35 \cdot (0,43 + 2 \cdot 0,27 + 3 \cdot 0,12 + 4 \cdot 0,09 + 5 \cdot 0,04) \\ &= 1,6615 \end{aligned}$$

Lösung 5-25: *ICE*

$$\begin{aligned} V &= \frac{\text{Gesamtstrecke}}{\text{Gesamtzeit}} \\ &= \frac{23 + 81 + 90 + 44 + 78 + 43 + 32 + 169}{14 + 39 + 31 + 17 + 29 + 25 + 17 + 75} \\ &= \frac{560 \text{ km}}{247 \text{ Min.}} = \frac{560 \text{ km}}{4,1166 \text{ h}} = 136,032 \text{ km/h} \end{aligned}$$

Lösung 5-26: *Dichtefunktion und Erwartungswert*

$$E(X) = \int_0^2 x \cdot f(x) dx$$

$$E(X) = \int_0^2 x \left(\frac{1}{4}x + \frac{1}{4}\right) dx = \int_0^2 \left(\frac{1}{4}x^2 + \frac{1}{4}x\right) dx$$

$$= \left[\frac{1}{12}x^3 + \frac{1}{8}x^2\right]_0^2 = \frac{8}{12} + \frac{4}{8} = \frac{7}{6} = 1,16667$$

Lösung 5-27: *Bahnstrecke Berlin – Nauen*

Klasse	0 - 30	30 - 60	60 - 90	90 - 120
$f(x_i)$	0,35	0,45	0,15	0,05
$F(x_j)$	0,35	0,8	0,95	1,00

$$x_{0,5} = x_j^u + \frac{0,5 - F(x_j^u)}{f(x_j)} \cdot (x_j^o - x_j^u) = 30 + \frac{0,5 - 0,35}{0,45} \cdot 30 = 30 + 10 = 40$$

Lösung 5-28: *Qualitätskontrolle*
Anwendung des allgemeinen Multiplikationssatzes,
Erwartungswert einer diskreten Zufallsvariablen: $E(X) = \sum_i x_i \cdot f(x_i)$
A_i : ein fehlerhaftes Stück wird bei der i-ten Kontrolle entdeckt; die Ereignisse A_i sind unvereinbar (disjunkt);
gegeben: $P(A_1) = 0,8;\ P(A_2|\overline{A}_1) = 0,6;\ P(A_3|\overline{A}_1 \cap \overline{A}_2) = 0,3$
$f(X = 5) = P(A_1) = 0,8$
$f(X = 10) = P(A_2) = P(A_2|\overline{A}_1) \cdot P(\overline{A}_1) = 0,6 \cdot 0,2 = 0,12$
$f(X = 20) = P(A_3) = P(A_3|\overline{A}_1 \cap \overline{A}_2) \cdot P(\overline{A}_2|\overline{A}_1) \cdot P(\overline{A}_1) = 0,3 \cdot 0,4 \cdot 0,2 = 0,024$
$f(X = 50) = P(\overline{A}_1 \cap \overline{A}_2 \cap \overline{A}_3) = P(\overline{A}_3|\overline{A}_1 \cap \overline{A}_2) \cdot P(\overline{A}_2|\overline{A}_1) \cdot P(\overline{A}_1) = 0,7 \cdot 0,4 \cdot 0,2 = 0,056$
$E(X) = \sum_i x_i \cdot f(x_i) = 5 \cdot 0,8 + 10 \cdot 0,12 + 20 \cdot 0,024 + 50 \cdot 0,056 = 8,48$

5.2 Bivariate Zufallsvariablen

Lösung 5-29: *Zweidimensionale Zufallsvariable*
Wir charakterisieren zunächst die Wahrscheinlichkeitsfunktion der Summe. Dazu stellen wir zunächst fest, dass 2,3,4 und 5 als Summe von X_1 und X_2 dargestellt werden können. Nun müssen wir noch die jeweiligen Wahrscheinlichkeiten berechnen. Die Summe nimmt nur dann den Wert 2 an, wenn X_1 und X_2 beide den Wert 1 annehmen. Wir entnehmen der gemeinsamen Wahrscheinlichkeitsfunktion, dass dieser Fall mit einer Wahrscheinlichkeit von $0,1$ eintritt.
Die Summe kann den Wert 3 annehmen, wenn $X_1 = 1$ und $X_2 = 2$ (Wahrscheinlichkeit ist $0,3$) oder wenn $X_1 = 2$ und $X_2 = 1$ (Wahrscheinlichkeit ist 0,1). Damit ist die Wahrscheinlichkeit, dass die Summe den Wert 4 annimmt, $0,4$. Analog werden die übrigen Wahrscheinlichkeiten berechnet. Die folgende Tabelle fasst die Ergebnisse zusammen:

x_3	2	3	4	5
$f(x_3)$	0,1	0,4	0,3	0,2

Damit gilt für den Erwartungswert:

$$\begin{aligned} E[X_3] &= 2 \cdot 0,1 + 3 \cdot 0,4 + 4 \cdot 0,3 + 5 \cdot 0,2 \\ &= 0,2 + 1,2 + 1,2 + 1 = 3,6. \end{aligned}$$

Lösung 5-30: *Zweidimensionale Zufallsvariable und Erwartungswert*
Wir beginnen wieder mit dem Aufstellen der Wahrscheinlichkeitsfunktion von Y. Dazu stellen wir zunächst fest, dass das Produkt der Faktoren 1,2 und 3 mit 1,2 und 4 die Werte 1,2,3,4,6,8 und 12 annehmen kann. Nun müssen wir noch die jeweiligen Wahrscheinlichkeiten berechnen. Damit das Produkt den Wert 1 annimmt, müssen sowohl der erste als auch der zweite Faktor den Wert 1 annehmen. Da dies mit einer Wahrscheinlichkeit von 0,1 passiert, ist die Wahrscheinlichkeit, dass das Produkt den Wert 1 annimmt, 0,1 [1]. Es gibt zwei Ereignisse, die dazu führen, dass das Produkt den Wert 2 annimmt: Wenn $X_1 = 2$ und $X_2 = 1$ (Wahrscheinlichkeit ist $0,2$) oder $X_1 = 1 = 1$ und $X_2 = 2$ (Wahrscheinlichkeit ist 0,1) eintritt. Damit ist die Wahrscheinlichkeit dafür, dass das Produkt den Wert 2 annimmt, $0,3$. Die anderen Wahrscheinlichkeiten werden analog berechnet und sind in der folgenden Tabelle aufgetragen.

y	1	2	3	4	6	8	12
f(y)	0,1	0,3	0,2	0,2	0,1	0	0,1

Damit ergibt sich für den Erwartungswert von Y

$$E(Y) = 1 \cdot 0,1 + 2 \cdot 0,3 + 3 \cdot 0,2 + 4 \cdot 0,2 + 6 \cdot 0,1 + 12 \cdot 0,1 = 3,9.$$

Lösung 5-31: *Gemeinsame Verteilung*
a) $P(X = Y) = 0,02 + 0,28 = 0,3$
b) $P(X + Y = 2) = 0,34 + 0,28 = 0,62$
c) $P(Y - X = 1) = 0,04 + 0,3 = 0,34$
d) $P(X \cdot Y = 1) = 0,28$
e) $E(X + Y) = E(X) + E(Y) = 0,6 + 1,6 = 2,2$
f) $E(X - Y) = E(X) - E(Y) = 0,6 - 1,6 = -1$
g) $Var(X) = (-0,6)^2 \cdot 0,4 + (0,4)^2 \cdot 0,6 = 0,144 + 0,096 = 0,24$
h) $Var(Y) = (-1,6)^2 \cdot 0,04 + (-0,6)^2 \cdot 0,32 + (0,4)^2 \cdot 0,64 = 0,1024 + 0,1152 + 0,1024 = 0,32$

[1] Das gleiche Argument führt auf die Wahrscheinlichkeit, dass das Produkt den Wert 6 bzw. 12 annimmt.

Lösung 5-32: *Bauteile*

a) Antwort: nein
Begründung:
Wenn X und Y unabhängig voneinander $\rightarrow f(x_i, y_j) = f(x_i) \cdot f(y_j)$ $\forall i, j$
Ist nicht erfüllt, da z.B.
$f(x_2, y_3) = 0,115 \neq 0,3 \cdot 0,4 = 0,12$

b) $f(x_1, y_1) + f(x_1, y_2) + f(x_2, y_1) = 0,015 + 0,03 + 0,04 = 0,085$

Lösung 5-33: *Spielkasino*
1. Durchgang:
Es gibt vier mögliche Ereignisse (Z, Z), (W, Z), (Z, W) und (W, W), wobei die grüne Münze an 1. Stelle und die rote Münze an 2. Stelle genannt wird, mit jeweils der Wahrscheinlichkeit von 0,25.
$X = 1$ tritt ein, wenn (Z,Z), (W,Z), (Z,W) eintritt, damit ist $P(X = 1) = 0,75$.
$X = 0$ tritt ein, wenn (W,W) eintritt, damit ist $P(X = 0) = 0,25$

2. Durchgang:
Es gibt zwei mögliche Ereignisse: die rote Münze zeigt Z bzw. W jeweils mit der Wahrscheinlichkeit $P(Z) = P(W) = 0,5$.
$P(Y = y \cap X = x) = P(Y = y|X = x) \cdot P(X = x)$
$P(Y = 1 \cap X = 0) = 0 \cdot 0,25 = 0$, da $(Y = 1 \cap X = 0)$ ein unmögliches Ereignis ist.
$P(Y = 0 \cap X = 0) = 1 \cdot 0,25 = 0,25$, da $(Y = 0|X = 0)$ ein sicheres Ereignis ist.
$P(Y = 1 \cap X = 1) = 0,5 \cdot 0,75 = 0,375$ da $P(Y = 1|X = 1) = P(Z|X = 1) = 0,5$
$P(Y = 0 \cap X = 1) = 0,5 \cdot 0,75 = 0,375$, da $P(Y = 0|X = 1) = P(W|X = 1) = 0,5$

Ergebnis:
$P(Y = 1, X = 1) = 0,375 \quad P(Y = 1, X = 0) = 0$
$P(Y = 0, X = 1) = 0,375 \quad P(Y = 0, X = 0) = 0,25$

6 Wichtige Verteilungsmodelle

Lösung 6-1: *Wartungen*
X_A = „Anzahl Wartungen für Maschine A in 8 Stunden" $\sim$ Poisson(1)
X_B = „Anzahl Wartungen für Maschine B in 8 Stunden" $\sim$ Poisson(2)

$$\begin{aligned} P(X_A = 0) &= e^{-1 \cdot 1} = e^{-1} = 0,3679 \\ P(X_B = 0) &= e^{-2 \cdot 1} = e^{-2} = 0,1353 \\ P(\text{keine Wartung}) &= P(X_A = 0 \cap X_B = 0) \\ &= 0,3679 \cdot 0,1353 \\ &= 0,04977687 = 4,98\% \end{aligned}$$

Lösung 6-2: *Gemeindegröße*
Allgemein gilt $q_\alpha = (x_\alpha - \mu)/\sigma$ mit q_α dem α–Quantil der Standardnormalverteilung und x_α dem α–Quantil der Normalverteilung der Gemeindegröße.
Aus den Angaben ergeben sich zwei Gleichungen:
$q_\alpha = (x_\alpha - \mu)/\sigma$
$q_{1-\beta} = (x_{1-\beta} - \mu)/\sigma$
mit $\alpha = 0,033626; \quad \beta = 0,008894; \quad 1 - \beta = 1 - 0,008894 = 0,991106$
Durch Auflösen der beiden Gleichungen nach σ ergibt sich:

$$\sigma = \frac{x_{1-\beta} - x_\alpha}{q_{1-\beta} - q_\alpha}$$

$x_{1-\beta} - x_\alpha = 100 - 1 = 99; \quad q_{1-\beta} - q_\alpha = 2,37 - (-1,83) = 4,2; \quad \sigma = 99/4,2 = 23,57143$

Lösung 6-3: *XXmega*

X = Anzahl der Gewinner, $P(\text{Gewinn}) = \pi = 0,6$ $X \sim B(30;\pi) = B(30;0,6)$

$$\begin{aligned} P\left(X < \frac{30}{3}\right) &= P(X < 10) = P(x \leq 9) \\ &= \sum_{k=0}^{9} \binom{30}{k} \pi^k (1-\pi)^{30-k} = \sum_{k=21}^{30} \binom{30}{k} \pi^{30-k}(1-\pi)^k \\ &= P(\tilde{X} \geq 21), \quad \tilde{X} \sim B(30; 1-\pi) \\ &= 1 - P(\tilde{X} \leq 20), \quad \tilde{X} \sim B(30; 0,4) \\ &= 1 - 0,9991 = 0,0009 \end{aligned}$$

Lösung 6-4: *Jahresrendite*

Zur Berechnung des Jahresendvermögens ist der Anlagewert mit dem zufälligen Jahreswachstumsfaktor zu multiplizieren. Letzterer ergibt sich aus der in Prozent angegebenen Rendite, indem sie durch 100 dividiert und anschließend zur Zahl 1 addiert wird:

Jahresendvermögen: $J = 150000(1 + R/100)$

$Var(J) = Var[150000(1 + R/100)] = Var(150000 + 1500 \cdot R) = Var(1500 \cdot R) = 1500^2 \cdot Var(R)$

Varianz der Rendite R:

da Rendite als gleichverteilt zwischen 6 und 8% angenommen wurde

$Var(R) = (b-a)^2/12 = (8-6)^2/12 = 1/3$

Damit resultiert:

$Var(J) = 1500^2 \cdot Var(R) = 1500^2 \cdot 1/3 = 750000$

$\sigma = 866,0254 \approx 866$ EUR

Lösung 6-5: *Varianz*

Die Verteilung ist eine Gleichverteilung auf $[a,b] = [-1,1]$;

$$f(x) = F'(x) = \begin{cases} 0 & -\infty < x \leq -1, \\ \frac{1}{2} & -1 < x < 1, \\ 0 & 1 \leq x < \infty. \end{cases}$$

Damit ist

$$Var(X) = \frac{(b-a)^2}{12} = 4/12$$

Lösung 6-6: *Formfehler*

X: “Anzahl der Formfehler bei $n = 10$ Belegen”

a) $X \sim B(10; 0,1)$

b) $P(X > 1) = 0,2639$

Lösung 6-7: *Vier Kinder*

X: “Anzahl der Jungen in einer vierköpfigen Familie”, $X \sim B(4; 0,5)$

a) $P(X = 2) = 0,375$

b) $P(X = 3) = 0,25$

c) $P(X = 4) = 0,0625$

Verteilungsmodelle-Vier Kinder.mp4

Lösung 6-8: *Kornflakes*

a) $X \sim B(26; 0,75)$

b) genau 3 Poster: $P(12 \leq X \leq 15) = 0,0397$;
höchstens 4 Poster: $P(X \leq 19) = 0,4846$;
genau 6 Poster: $P(X \geq 24) = 0,0258$;
höchstens 1 Poster: $P(X \leq 7) = 0$

c) $E(X) = 19,5$ Packungen mit Coupons

Lösung 6-9: *Prüfungsfragen*
X: "Anzahl des Auftretens einer beantwortbaren Frage bei $n = 3$ abhängigen Ziehungen"; $X \sim H(10; 4; 3)$

a) $P(X = 3) = 0,0333$

b) $P(X \geq 1) = 0,8333$

Verteilungsmodelle-Pruefungsfragen.mp4

Lösung 6-10: *Eier*
X: "Anzahl der faulen Eier bei n=3 abhängigen Ziehungen";
$X \sim H(6; 2; 3)$

a) $P(X = 1) = 0,6$

b) $P(X \leq 1) = 0,8$

c) $P(X = 3) = 0$

d) $E(X) = 1$

Y: "Anzahl der guten Eier in Lieferung von 20 Eiern"
Y ist approximativ $(20n \leq N)$ $B(20; 0,8)$ verteilt.
Z: "Anzahl der faulen Eier in Lieferung von 20 Eiern"
Z ist approximativ $(20n \leq N)$ $B(20; 0,2)$ verteilt.

e) $P(Z > 2) = 0,7939$

f) $E(Y) = 16$

g) $P(Z = 16) = 0$

Verteilungsmodelle-Eier.mp4

Lösung 6-11: *Telefongespräche*

a) $X \sim PO(2,5)$

b) $P(X = 0) = 0,0821$; $P(X < 3) = 0,5438$; $P(X \geq 4) = 0,2424$

Verteilungsmodelle-Telefongespraeche.mp4

Lösung 6-12: *Betriebe der chemischen Industrie*

a) $Y \sim B(750; 0,01)$

b) $E(Y) = 7,5$

c) Y ist approximativ $(n > 50; p \leq 0,1; np < 10)$ $PO(7,5)$–verteilt.

d) $P(Y < 8) = 0,5246$

e) $P(Y \leq 5) = 0,2414$

Lösung 6-13: *Serum*
X: "Anzahl der Impfschäden bei n=20 000 Impfungen";
$X \sim B(20000; 0,0001)$
X ist approximativ $(n > 50; p \leq 0,1; np < 10)$ $PO(2)$–verteilt.

a) $P(X = 0) = 0,1353$

b) $P(X = 1) = 0,2707$

c) $P(X = 6) = 0,0121$

d) $P(X > 4) = 0,0527$

Lösung 6-14: *Elektronisches Bauteil*
X: "Anzahl der Ausfälle pro Stunde"; $X \sim PO(2)$

a) Y: "Wartezeit auf den nächsten Ausfall (in Std.)"; $Y \sim EX(2)$

b) $P(Y > 2) = 0,01832$

c) Wahrscheinlichkeit, dass bis zum nächsten Ausfall mehr als eine, aber höchstens zwei Stunden vergehen.

d) $P[(Y_1 > 2) \cap (Y_2 > 2)] = 0,000335$

Lösung 6-15: *Telefonzentrale*

1. X: "Anzahl der pro Dienstzeit ankommenden Alarmmeldungen"
   $E(X) = 0,5 \cdot 6 = 3$; $X \sim PO(3)$
   a) $P(X = 0) = 0,0498$
   b) $P(X \geq 3) = 0,5768$
   c) $P(X \leq 7) = 0,9881$
2. T: "Wartezeit bis zum ersten Alarm" [$t = 1$ Std.]
   $T \sim EX(0,5)$
   a) $P(T \leq 1) = 0,3935$
   b) $P(T > 2) = 0,3679$
   c) $P(T \leq 5 + 1 | T > 5) = P(T \leq 1) = 0,3935$
3. $P(T \leq t) = 0,95 \quad \Rightarrow \quad t = 5,99$

Verteilungsmodelle-Telefonzentrale (13 min).mp4

Lösung 6-16: *Stahlstifte*

$X_1 \sim N(6; 0,4)$

a) $P[(X_1 < \mu_1 - 0,12) \cup (X_1 > \mu_1 + 0,12) = 0,764178$

b) $P(X_1 = 6) = 0$

c) $P(X_1 \leq x_1) \leq 0,85 \Rightarrow x_1 = 6,412$ mm

d) $X_2 \sim N(6,05; 0,3)$; $P(X_2 < 6) = 0,432505$

e) $Y = X_2 - X_1; Y \sim N(0,05; 0,5)$

f) $P(Y \leq 0) = 0,460172$

Lösung 6-17: *Taschenrechner*

$X \sim N(30; 3)$; $Y \sim N(35; 4)$

a) $P(15 \leq X \leq 27) = 0,158655$

b) $P(Y \leq y) = 0,853141 \quad \Rightarrow \quad y = 39,2$ Std.

c) $P[(X > 24) \cap (Y > 24)] = 0,9743378$

d) $P(X < Y) = P(X - Y < 0) =?$
   $X - Y \sim N(-5; 5) \Rightarrow P(X - Y < 0) = 0,841345$

Lösung 6-18: *Bäcker Backfrisch*

$X \sim N(150; 4)$

a) $Y \sim N(600; 8)$

b) $P(Y = 600) = 0$; $P(594 \leq Y \leq 606) = 0,546746$

Verteilungsmodelle-Baecker Backfrisch.mp4

Lösung 6-19: *Mittagszeit*

A. U: "Anzahl der zwischen 13 und 15 Uhr eintreffenden Kunden"; $U \sim PO(8)$
   a) $U_1 \sim EX(8)$
   b) $P(U_1) > 1/4) = 0,1353$
   c) $P(U_1 > 1/4 + 1/8 | U_1 > 1/4) = 0,3679$

B. a) U_2 ist stetig gleichverteilt in $[0, 3]$
   b) $P(U_2 \leq 1 + 1/2 | U_2 > 1) = 0,25$

Lösung 6-20: *Briefmarkenschalter*
$X \sim PO(\lambda)$, $\lambda = 4$. Vor dem Schalter hat sich nach einer Minute eine Schlange gebildet, wenn in diesem Zeitraum mehr als 5 Kunden eingetroffen sind.
$P(X > 5) = 1 - P(X \leq 5) = 1 - 0,7851$ (aus Tafel) $= 0,2149$

Lösung 6-21: *Bogenschütze*

a) $X_i \sim$ Bernoulli(p) bzw. $B(1;p) = B(1;0,6)$
$Y = \sum_i X_i \sim B(n;p) = B(8;0,6)$
X_i: „Anzahl der Treffer bei einem Schuß“ kann nur die Werte 1 (Treffer) oder 0 (kein Treffer) annehmen; dichotome Grundgesamtheit
$P(X_i = 1) = 3/5 = 0,6$; $P(X_i = 0) = 1 - p = 1 - 0,6 = 0,4$; Wahrscheinlichkeiten konstant.
Da Grundgesamtheit unendlich groß ist, kann Modell mit Zurücklegen als Stichprobentechnik unterstellt werden $\rightarrow X_i (i = 1, \ldots, 8)$ sind unabhängig voneinander. Die Bedingungen eines Bernoulli–Experiments sind erfüllt.

b) $E(Y) = n \cdot p = 8 \cdot 0,6 = 4,8$ Treffer

c) $P(Y = 3) = \binom{8}{3} \cdot 0,6^3 \cdot 0,4^5 = 56 \cdot 0,216 \cdot 0,01024 = 0,1239$
oder $P(Y = y) = F(n - y; n; 1 - p) - F(n - y - 1; n; 1 - p)$;
$P(Y = 3) = F(5;8;0,4) - F(4;8;0,4) = 0,9502 - 0,8263 = 0,1239$

Verteilungsmodelle-Bogenschuetze.mp4

Lösung 6-22: *Straßenmusikant*

a) $X \sim PO(\lambda)$ mit $\lambda = 1/5$
X beinhaltet das Auftreten eines Ereignisses in einem Kontinuum.

b) $E(X) \cdot 4 \cdot 60 = 48$ Geldstücke
$\rightarrow (48/5) \cdot 3$ EUR $= 28,80$ EUR

c) $X \sim PO(\lambda)$ mit $\lambda = 1/5$
$\rightarrow T$: „Wartezeit zwischen zwei Ereignissen“; $T \sim EX(\lambda)$
$E(T) = 1/\lambda = 1/(1/5) = 5$ Minuten Wartezeit

d) $P(T > t) = 1 - F_{EX}(t;\lambda) = 1 - (1 - e^{-\lambda t}) = e^{-\lambda t}$
$P(T > 3) = 1 - F_{EX}(3;0,2) = e^{-0,2 \cdot 3} = 0,5488$

Verteilungsmodelle-Strassenmusikant.mp4

Lösung 6-23: *Gaststätte*

A.

a) X: „Anzahl der am Sonntagabend pro Stunde kommenden Gäste“ [Auftreten von unabhängigen Ereignissen in einem Kontinuum]
$E(X) = \lambda = 25$(Gäste)/5(Stunden) $= 5$ Gäste/Stunde; $X \sim PO(5)$
$P(X = 1) = (\lambda^X \cdot e^{-\lambda})/x! = 5^1 \cdot e^{-5}/1! = 0,03369 \approx 0,0337$
(Oder unter Verwendung der Tabelle der Verteilungsfunktion der Poisson–Verteilung:
$P(X = 1) = F_{PO}(1;5) - F_{PO}(0;5) = 0,0404 - 0,0067 = 0,0337$)
Die Wahrscheinlichkeit dafür, dass in der ersten Stunde genau ein Gast erscheint, beträgt 3,37%.

b) Y: „Wartezeit auf den nächsten Gast am Sonntagabend“; $Y \sim EX(\lambda) = EX(25)$
$E(Y) = 1/\lambda = 1/25$ am Sonntagabend (5 Stunden$\cdot$60 Minuten) d.h. $300/25 = 12$ Minuten $= 0,2$ Stunden
Im Mittel vergehen am Sonntagabend 12 Minuten zwischen der Ankunft zweier Gäste.

B.

a) $X \sim B(n;p) = B(25;0,3)$
(2 mögliche Ereignisse: $A =$ „Nachbestellung“;
$\overline{A} =$ „keine Nachbestellung“;
$P(\overline{A} = 0,7 \rightarrow P(A) = 0,3$; unabhängige Versuche)

b) $P(X > x) = 0,0005$; $P(X \leq x) = 0,9995$
$x = 15$ (aus Tabelle der Verteilungsfunktion der $B(25;0,3)$)
Die Kapatitätsgrenze ist bei 15 Nachbestellungen erreicht.

Lösung 6-24: *Tulpenzwiebeln*
$X =$ „Anzahl der nicht blühenden Tulpenzwiebeln“; $X \sim B(10; 0,05)$

$$\begin{aligned} P(X > 1) &= 1 - P(X \leq 1) = 1 - P(X = 0) - P(X = 1) \\ &= 1 - \binom{10}{0} \cdot 0,05^0 \cdot 0,95^{10} - \binom{10}{1} \cdot 0,05^1 \cdot 0,95^9 \\ &= 1 - 0,95^{10} - 10 \cdot 0,05^1 \cdot 0,95^9 \\ &= 1 - 0,5987 - 0,3151 = 0,0862 \end{aligned}$$

Verteilungsmodelle-Tulpenzwiebeln.mp4

Lösung 6-25: *Supermarkt*
$T_B =$ „Wartezeit am Backstand“ $\sim$ Exp(1/5)
$T_K =$ „Wartezeit am Käsestand“ $\sim$ Exp(1/4)

$P(T_B > 10) = 1 - P(T_B \leq 10) = 1 - \{1 - \exp(-10/5)\} = \exp(-10/5) = 0,135$
$P(T_K > 10) = 1 - P(T_K \leq 10) = 1 - \{1 - \exp(-10/4)\} = \exp(-10/4) = 0,082$

$P(T_B > 10 \cup T_K > 10) = P(T_B > 10) + P(T_K > 10) - P(T_B > 10 \cap T_K > 10)$
$P(T_B > 10 \cap T_K > 10) = P(T_B > 10) \cdot P(T_K > 10)$
$P(T_B > 10 \cup T_K > 10) = 0,135 + 0,082 - 0,135 \cdot 0,082 = 0,206$

Lösung 6-26: *Abendessen*
$X \sim N(1000; 20)$ Gewicht der Apfel-Schale
$Y \sim N(1000, 15)$ Gewicht des Mandarinen-Netzes
Gesamtgewicht: $G = X + Y \sim N(2000; 25)$, da $Var(G) = 400 + 225 = 625 = 25^2$.
Also:

$$\begin{aligned} P(G > 1950) &= P\Big(\frac{G - 2000}{25} > \frac{1950 - 2000}{25}\Big) \\ &= 1 - \Phi(-50/25) = 1 - \{1 - \Phi(50/25)\} \\ &= \Phi(2) = 0,97725 \approx 0,977 \end{aligned}$$

Lösung 6-27: *Pizza– und Kuchenverkauf*
$X_A =$ „Anzahl der Kuchen-Kunden in 20 Minuten“ $\sim$ Poisson(4)
$X_B =$ „Anzahl der Pizza-Kunden in 20 Minuten“ $\sim$ Poisson(1)
$T_A =$ „Wartezeit auf ersten Kuchen-Kunden“ $\sim$ Exp(4)
$T_B =$ „Wartezeit auf ersten Pizza-Kunden“ $\sim$ Exp(1)
Da:

$$\begin{aligned} P(T_A > 0,5) &= 1 - P(T_A \leq 0,5) = 1 - \{1 - \exp(-4 \cdot 0,5)\} \\ &= \exp(-2) = 0,1353 \\ P(T_B \leq 0,5) &= 1 - \exp(-1 \cdot 0,5) = 1 - \exp(-0,5) = 0,3935 \end{aligned}$$

folgt

$$P(T_A > 0,5 \cap T_B \leq 0,5) = P(T_A > 0,5) \cdot P(T_B \leq 0,5) = 0,1353 \cdot 0,3935 = 0,0532$$

Lösung 6-28: *Wertpapierkurse*
X : „Wartezeit auf die nächste Anfrage“,
$E(X) = 1/\lambda = 20$; $\quad X \sim \text{EX}(\lambda = 1/20)$
$P(X > 30) = 1 - P(X \leq 30) = 1 - F(30)$
$= 1 - [1 - e^{-\lambda x}] = e^{-30/20} = e^{-1,5} = 0,2231$

Lösung 6-29: *Zug nach Brandenburg*
X : „Wartezeit auf den Regionalzug nach Brandenburg“
$x = 1$ (Stunde) $\quad a = 0 \quad b = 3$ (Stunden)
Anwendung der stetigen Gleichverteilung (regelmäßig im 3-Stunden-Takt, nicht im Mittel alle 3 Stunden)

$$F(X) = \begin{cases} 0 & \text{für } x < a \\ \dfrac{x - a}{b - a} & \text{für } a \leq x < b \\ 1 & \text{für } b \leq x \end{cases}$$

$F(x \leq 1) = 1/3 = 0,3333$
mindestens 1 Stunde warten: $1 - F(x \leq 1) = 1 - 0,3333 = 0,6667$

Lösung 6-30: *Radrennen*
Geamtzahl der Fahrer: $10 \cdot 3 = 30$
Auswahl von 4 Fahrern aus 30 ohne Berücksichtigung der Reihenfolge und ohne Wiederholung:

$$K(30, 4) = 27405$$

Auswahl von 4 Fahrern aus 30, wobei 3 vom eigenen Team sind = Auswahl von einem Fahrer aus 27:

$$K(27, 1) = 27$$

Wahrscheinlichkeit, dass alle drei Fahrer vom eigenen Team ausgewählt werden:

$$\frac{K(27,1)}{K(30,4)} = 0,000985221 \approx 0,0010$$

Oder über die hypergeometrische Verteilung:

$$f_H(x; N, M, n) = \frac{\binom{27}{1} \cdot \binom{30-27}{4-1}}{\binom{30}{4}} = \frac{\frac{27!}{1!26!} \cdot \frac{3!}{3!0!}}{\frac{30!}{4!26!}} = 0,000985221 \approx 0,0010$$

Lösung 6-31: *Rückversicherungsgesellschaft*
Bei den beschriebenen Großschäden handelt es sich um zufällige Ereignisse, die in einem Kontinuum (Zeit) vorgegebener Größe (4 Monate) auftreten. Der Parameter $\lambda = 1$ gibt die mittlere Anzahl von Großschäden in diesem Intervall an. Frage richtet sich auf ein Intervall von 1 Jahr = 12 Monate = $3 \cdot 4$ Monate.
Für das Intervall von 1 Jahr ist $\lambda = 1 \cdot 3 = 3$.
X = Anzahl von Großschäden in einem Jahr $\sim$ PO(3)
Wahrscheinlichkeit, dass in einem Jahr mindestens 5 derartige Großschadensfälle auftreten:
$P(X \geq 5) = 1 - P(X \leq 4) = 1 - F_{PO}(4) = 1 - 0,8153 = 0,1847$
mit $F_{PO}(4)$ aus der Tabelle der Verteilungsfunktion der PO(3)

Lösung 6-32: *Landwirtschaftsexperte*
Ereignis A : BSE–verseuchtes Rind Ereignis $\overline{A}$: BSE–freies Rind
$P(A) = p = 0,10$ und $P(\overline{A}) = 1 - p = 0,90$
sehr große Gesamtheit (europäischer Rinderbestand), so dass mit oder ohne Zurücklegen keine Rolle spielt
$X = \{$Anzahl des Auftretens BSE–verseuchter Rinder bei n Ziehungen$\}$
Wertebereich: $0, 1, 2, \ldots, n$
$X \sim B(n; p)$ mit $p = 0,10$ und unbekanntem n
Gegeben: $P(X \geq 1) \geq 0,95$
Ermittlung von n:
$P(X \geq 1) = 1 - P(X = 0) = 0,95$
$P(X \geq 1) = 1 - P(X = 0) = 1 - \binom{n}{0} \cdot 0,10^0 \cdot (0,90)^{n-0} = 1 - 0,90^n = 0,95$
$\rightarrow 0,90^n = 0,05;\qquad n = \ln 0,05 / \ln 0,9 = 28,4332$
Es muss also mindestens $n = 29$ gewählt werden.
Kontrolle:
$X \sim B(28; 0,10):\qquad P(x \geq 1) = 1 - P(X = 0) = 1 - 0,0523 = 0,9477$
$\rightarrow P(X \geq 1) \geq 0,95$ wird nicht eingehalten
$X \sim B(29; 0,10):\qquad P(X \geq 1) = 1 - P(X = 0) = 1 - 0,0471 = 0,9529$
$\rightarrow P(X \geq 1) \geq 0,95$ eingehalten.

Lösung 6-33: *Unfallmeldungen*

$$X = (\text{Zeitspanne, die zwischen zwei Unfallmeldungen in einer Polizeistation vergeht})$$

Wegen $E(X) = 1/\lambda = 160$ folgt $\lambda = 1/160$ und somit $X \sim \mathrm{EX}(1/160)$.

$$\begin{aligned} P(60 < X \leq 160) &= F_{EX}(160; 1/160) - F_{EX}(60; 1/160) \\ &= 1 - e^{-160/160} - (1 - e^{-60/160}) \\ &= 1 - 0,3679 - (1 - 0,6873) = 0,3194 \end{aligned}$$

Lösung 6-34: *Radrennfahrer*
Ereignisse: A = Unfall bei Anton, $P(A) = 1/12000$;
B = Unfall bei Bertram, $P(B) = 1/10000$
X = Anzahl der Unfälle von Anton in zwei Wochen
Y = Anzahl der Unfälle von Bertram in zwei Wochen
In beiden Fällen:
Wahrscheinlichkeiten für das Eintreten des Ereignisses (Unfall) sehr klein, d.h. seltene Ereignisse und Anzahl n der unabhängigen Versuche (gefahrene Kilometer in zwei Wochen) sehr groß:
$n(\text{Anton}) = 14(\text{Tage}) \cdot 180(\text{km/Tag}) = 2520$ km; $n(\text{Bertram}) = 14(\text{Tage}) \cdot 210(\text{km/Tag}) = 2940$ km
mittlere Anzahl der Unfälle in zwei Wochen:
$\lambda_x = 2520/12000 = 0,21 \quad \lambda_y = 2940/10000 = 0,294$
$X \sim \text{PO}(\lambda = 0,21) = f_{PO}(x;\lambda) = (\lambda^x/x!)e^{-\lambda}$; $Y \sim \text{PO}(\lambda = 0,294)$
Da X und Y unabhängig voneinander sind, gilt aufgrund der Reproduktivitätseigenschaft der Poisson-Verteilung:
$Z = X + Y \sim \text{PO}(\lambda_x + \lambda_y) = \text{PO}(0,21 + 0,294) = \text{PO}(0,504)$
Gesuchte Wahrscheinlichkeit:
$P(Z \leq 1|\lambda_z = 0,504) = P(Z = 0|\lambda_z = 0,504) + P(Z = 1|\lambda_z = 0,504)$
$= \frac{0,504^0}{0!}e^{-0,504} + \frac{0,504^1}{1!}e^{-0,504} = e^{-0,504} + 0,504 \cdot e^{-0,504} = 1,504 \cdot e^{-0,504}$
$= 0,90858 \approx 0,909$

Lösung 6-35: *Polizeistation*
$X = \{\text{Zeit bis zur ersten Unfallmeldung}\} \sim \text{EX}(0,5)$
$F(x) = 1 - e^{-\lambda x}$ für $x \geq 0$ und $\lambda = 0,5$
$P(X > 2) = 1 - P(X \leq 2) = 1 - (1 - e^{-0,5\cdot 2}) = e^{-0,5\cdot 2} = 0,36788$

Lösung 6-36: *Fahrtkostenzuschuss*
$X_i = \{\text{täglicher Arbeitsweg eines Mitarbeiters}\}; X_i \sim N(\mu = 50; \sigma^2 = 32); i = 1, \ldots, 50$
$Y = \{\text{Zahlung des Unternehmens an die Mitarbeiter je Tag}\}$; Reproduktivitätseigenschaft der Normalverteilung, $a = 0,1$ für alle $i = 1, \ldots, 50$

$$Y = \sum_{i=1}^{n} aX_i \sim N\left(\sum_{i=1}^{n} a\mu, \sqrt{\sum_{i=1}^{n} a^2\sigma^2}\right)$$

$\mu_y = na\mu = 50 \cdot 0,1 \cdot 50 = 250,\ \sigma_y^2 = na^2\sigma^2 = 50 \cdot 0,1^2 \cdot 32 = 16$
$P(Y > 255) = 1 - P(Y \leq 255) = 1 - P(Z \leq (255 - 250)/4)$
$= 1 - P(Z \leq 1,25) = 1 - 0,89435 = 0,10565$

Lösung 6-37: *Gleichverteilung*
stetige Gleichverteilung
$E(X) = (a + b)/2 = 16;\ a + b = 32$
$Var(X) = (b-a)^2/12 = 12;\ (b-a)^2 = 12^2;\ b-a = 12;\ a = b-12;\ b = 12+a$
$b - 12 + b = 32;\ b = 22;\ a + 12 + a = 32;\ a = 10$

Lösung 6-38: *Parkplaketten*
X : „Anzahl der gewonnenen Parkplaketten bei $n = 3$ Versuchen";
$X \sim B(n;p) = B(3;0,4)$
$P(X \geq 2) = 1 - P(X < 2) = 1 - P(X \leq 1) = 1 - F_B(1;3;0,4) = 0,352$

Lösung 6-39: *Traineeprogramm*
Es bezeichne X die Anzahl der geeigneten Bewerber. Die Frage lautet $P(X \geq 20)$. Die Zufallsvariable X genügt einer B(23;0,9) Verteilung.

$$\binom{23}{20} 0,9^{20} \cdot 0,1^3 + \binom{23}{21} 0,9^{21} \cdot 0,1^2 + \binom{23}{22} 0,9^{22} \cdot 0,1^1 + \binom{23}{23} 0,9^{23} \cdot 0,1^0$$

$0,21531 + 0,27683 + 0,2265 + 0,08863 = 0,80727 \approx 0,8073$
Um jedoch die gesuchte Wahrscheinlichkeit den Tabellen der Verteilungsfunktion der Binomialverteilung entnehmen zu können, kann auch die Zufallsvariable $Y = n - X$(Anzahl der ungeeigneten Trainees), die B(23;0,1)–verteilt ist, verwendet werden:

$$P(X \geq 20) = P(-X \leq -20) = P(23 - X \leq 23 - 20)$$
$$= P(Y \leq 3) = F_{B(23;0,1)}(3) = 0,8073$$

Lösung 6-40: *Suppe mit Fleischeinlage*
125 l $\widehat{=}$ 500 Portionen. Die Fleischstückchen sind in der Suppe zufällig verteilt. 1/4 l Suppe (1 Portion) enthält im Mittel $\lambda = 400/500 = 0,8$ Fleischstückchen.
X: „Anzahl der Fleischstückchen je Portion“ ; $X \sim PO(\lambda = 0,8)$
$P(X > 2) = 1 - P(X \leq 2) = 1 - 0,9526$ (aus Tabelle der Poisson-Verteilung) $= 0,0474$

Lösung 6-41: *Miss–Wahl*
Es bezeichne X die Anzahl der geeigneten Kandidatinnen. Die Zufallsvariable X genügt einer $B(25;0,55)$. Gesucht ist $P(X = 12)$.

$$\binom{25}{15} \cdot 0,55^{12} \cdot 0,45^{13} = 0,1236$$

Oder aus der Tabelle der Verteilungsfunktion der $B(25;0,45)$:
$f_B(x;n;p) = F_B(n - x;n;1 - p) - F_B(n - x - 1;n;1 - p)$

$$B(13;25;0,45) - B(12;25;0,45) = 0,8173 - 0,6937 = 0,1236$$

Lösung 6-42: *Computernetzwerk*
X:„Wartezeit bis zum nächsten Defekt“; X ist exponentialverteilt mit $E(X) = 1/\lambda = 10$ (Tage bis zum nächsten Defekt); $\lambda = 1/10$. $P(X > 21) = \exp(-21/10) = \exp(-2,1) = 0,122$

Lösung 6-43: *Produktionsanlage*
X: „Anzahl der Ausschußstücke“
Wegen p klein und n groß ist $X \sim PO(x;\lambda)$ mit $\lambda = np = 500 \cdot 0,002 = 1$
Mindestens 499 Stück normgerecht entspricht $(X = 0) \cup (X = 1) = (X \leq 1)$
$P(X \leq 1) = 0,7358$ (aus Tabelle der Poisson–Verteilung)
Die Wahrscheinlichkeit dafür, dass unter $n = 500$ Stück mindestens 499 Stück normgerecht sind, beträgt 73,58%.

Bzw. ohne Approximation:
$X \sim B(500;0,002)$
$P(X \leq 1) = f(0) + f(1) = 0,998^{500} + 0,998^{499} = 0,998^{499}(1 + 0,998) = 0,7358$

Lösung 6-44: *Samstagslotto*
Hypergeometrische Verteilung mit $N = 49$, $M = 24$, $n = 6$;
Wert der Wahrscheinlichkeitsfunktion für $x = 3$

$$f_H(x;N,M,n) = 0,3328991$$

Lösung 6-45: *Kommode*
Hypergeometrische Verteilung mit $N = 20$, $M = 10$ und $n = 2$;
Wert der Wahrscheinlichkeitsfunktion für $x = 2$

$$f_H(x;N,M,n) = \frac{9}{38} \approx 0,2368$$

Lösung 6-46: *Geschirr*
Zwei mögliche Ereignisse:
A : „Geschirr geht an einem bestimmten Tag kaputt“ und
$\overline{A}$: „Geschirr geht an einem bestimmten Tag nicht kaputt“;
Unabhängigkeit des Eintretens der Ereignisse von Tag zu Tag;
X : „Anzahl der Tage, an denen Geschirr kaputt geht bei insgesamt 5 Tagen“;
$X \sim B(n; p) \quad n = 5; \quad p = 0,7; \quad$ gefragt: $P(X = 2) = ?$

$$P(X = 2) = \binom{5}{2} 0,7^2 \cdot 0,3^3 = 10 \cdot 0,49 \cdot 0,027 = 0,1323$$

oder unter Verwendung der Tabellen der Verteilungsfunktion der Binomialverteilung:

$n = 5; \quad p^* = 1 - p = 0,3; \quad y = n - x = 5 - 2 = 3$
$f(y; n; p^*) = f(3; 5; 0,3) = F(3; 5; 0,3) - F(2; 5; 0,3) = 0,9692 - 0,8369 = 0,1323$

Lösung 6-47: *Prüfgebiete*
Hypergeometrische Verteilung mit $N = 6$, $M = 4$ und $n = 3$;
Wert der Wahrscheinlichkeitsfunktion für $x = 2$

$$f_H(x; N, M, n) = 0,6$$

Lösung 6-48: *Dichtefunktion*
Die Verteilung ist eine Exponentialverteilung mit λ–Parameter 2. Da der Erwartungswert von X 0,5 ist, ist die Varianz von X gesucht, $1/\lambda^2$, also 0,25.

7 Stichprobentheorie

Lösung 7-1: *Drei Personen*

(a) Mittelwert μ und Varianz σ^2 der Grundgesamtheit:

$$\mu = \frac{20 + 22 + 24}{3} = 22$$
$$\sigma^2 = \frac{(20-22)^2 + (22-22)^2 + (24-22)^2}{3} = \frac{4+0+4}{3} = \frac{8}{3}$$

(b) Anordnung spielt eine Rolle $\Rightarrow$ Variation, Wiederholung möglich

$$V^W(3; 2) = 3^2 = 9$$

(c),(d)

Nr.		1	2	3	4	5	6	7	8	9
x_1		20	20	20	22	22	22	24	24	24
x_2		20	22	24	20	22	24	20	22	24
$P(X_1 = x_1, X_2 = x_2)$		$\frac{1}{9}$	$\frac{1}{9}$	$\frac{1}{9}$	$\frac{1}{9}$	$\frac{1}{9}$	$\frac{1}{9}$	$\frac{1}{9}$	$\frac{1}{9}$	$\frac{1}{9}$
y	$20 + 20 =$	40	42	44	42	44	46	44	46	48
$\bar{x}$	$\frac{20+20}{2} =$	20	21	22	21	22	23	22	23	24
$s^{2\prime}$	$\frac{(20-20)^2+(20-20)^2}{2} =$	0	1	4	2	0	1	4	1	0
s^2	$\frac{(20-20)^2+(20-20)^2}{1} =$	0	2	8	2	0	2	8	2	0

(i)

y	40	42	44	46	48
$P(Y = y)$	1/9	2/9	3/9	2/9	1/9

$$E(Y) = \frac{40}{9} + \frac{42 \cdot 2}{9} + \frac{44 \cdot 3}{9} + \frac{46 \cdot 2}{9} + \frac{48 \cdot 1}{9} = 44$$
$$Var(Y) = \frac{(40-44)^2}{9} + \frac{(42-44)^2 \cdot 2}{9} + \frac{(44-44)^2 \cdot 3}{9} + \frac{(46-44)^2 \cdot 2}{9} + \frac{(48-44)^2}{9}$$
$$= \frac{16}{9} + \frac{8}{9} + \frac{0}{9} + \frac{8}{9} + \frac{16}{9} = \frac{48}{9} = \frac{16}{3}$$

(ii)

$\overline{x}$	20	21	22	23	24
$P(\bar{X}=\bar{x})$	1/9	2/9	3/9	2/9	1/9

$$E(\bar{X}) = \frac{20}{9} + \frac{21\cdot 2}{9} + \frac{22\cdot 3}{9} + \frac{23\cdot 2}{9} + \frac{24\cdot 1}{9} = 22$$
$$E(\bar{X}) = E\left(\frac{Y}{2}\right) = \frac{1}{2}E(Y) = 22 \text{ (alternativ)}$$
$$Var(\bar{X}) = \frac{(20-22)^2}{9} + \frac{(21-22)^2\cdot 2}{9} + \frac{(22-22)^2\cdot 3}{9} + \frac{(23-22)^2\cdot 2}{9} + \frac{(24-22)^2}{9}$$
$$= \frac{4}{9} + \frac{2}{9} + \frac{0}{9} + \frac{2}{9} + \frac{4}{9} = \frac{12}{9} = \frac{4}{3}$$
$$Var(\bar{X}) = Var\left(\frac{Y}{2}\right) = \frac{1}{4}Var(Y) = \frac{4}{3} \text{ (alternativ)}$$

(iii)

s'^2	0	1	4
$P(S'^2 = s'^2)$	3/9	4/9	2/9

$$E(S'^2) = \frac{0\cdot 3}{9} + \frac{1\cdot 4}{9} + \frac{4\cdot 2}{9} = \frac{12}{9} = \frac{4}{3}$$

(iv)

s^2	0	2	8
$P(S^2 = s^2)$	3/9	4/9	2/9

$$E(S^2) = \frac{0\cdot 3}{9} + \frac{2\cdot 4}{9} + \frac{8\cdot 2}{9} = \frac{24}{9} = \frac{8}{3}$$
$$E(S^2) = E\left(S'^2 \cdot \frac{n}{n-1}\right) = E(2\cdot S'^2) = 2\cdot\frac{4}{3} = \frac{8}{3} \text{ (alternativ)}$$

e) (i)

$$44 = E(Y) = n\cdot\mu = 2\cdot 22 = 44$$
$$\frac{16}{3} = Var(Y) = n\cdot\sigma^2 = 2\cdot\frac{8}{3} = \frac{16}{3}$$

(ii)

$$22 = E(\overline{X}) = \mu = 22$$
$$\frac{4}{3} = Var(\overline{X}) = \sigma^2/n = \frac{8}{3\cdot 2} = \frac{4}{3}$$

(iii)

$$\frac{4}{3} = E(S'^2) = \sigma^2\cdot(n-1)/n = \frac{8}{3}\frac{1}{2} = \frac{4}{3}$$

(iv)

$$\frac{8}{3} = E(S^2) = \sigma^2 = \frac{8}{3}$$

Lösung 7-2: *Normal–Verteilung Approximation*

(a) (i) $n > 30$

(ii) $E(X_i) \neq \pm\infty$

(iii) $0 < Var(X_i) < \infty$

(iv) X_i identisch verteilt

(v) X_i unabhängig

(b) Zentraler Grenzwertsatz

(c)

$$\begin{aligned}
Y = \sum_{i=1}^{n} X_i \quad &\approx \quad N(\bullet;\bullet) \text{ (ZGWS)} \\
E(Y) \quad &= \quad E\left(\sum_{i=1}^{n} X_i\right) = \sum_{i=1}^{n} E(X_i) = \sum_{i=1}^{n} \mu = n \cdot \mu \\
Var(Y) \quad &= \quad Var\left(\sum_{i=1}^{n} X_i\right) = \sum_{i=1}^{n} Var(X_i) = \sum_{i=1}^{n} \sigma^2 = n \cdot \sigma^2
\end{aligned}$$

Lösung 7-3: *Tabletten gegen Kopfschmerzen*
X_i: "Wirkstoffmenge in einer Tablette"; $X_i \sim N(\mu;\sigma)$
$\overline{X}$: "Durchschnittliche Wirkstoffmenge in einer Tablette"; $\overline{X} \sim N(\mu;\sigma/\sqrt{n})$

$$P(\overline{X} > \mu + \frac{1}{2}) = P\left(\frac{\overline{X}-\mu}{\sigma/\sqrt{n}} > \frac{\sqrt{n}}{2\sigma}\right) = 1 - \Phi\left(\frac{\sqrt{n}}{2\sigma}\right)$$

(i) $P(Z > 2) = 0,02275$

(ii) $P(Z > 4) = 0,000032$

(iii) $P(Z > 2) = 0,02275$

Lösung 7-4: *Erwartungswert und Varianz*

a) $\sum_{i=1}^{n} X_i \sim B(n;p)$; $E(\sum_{i=1}^{n} X_i) = np$; $Var(\sum_{i=1}^{n} X_i) = np(1-p)$

b) $\sum_{i=1}^{n} X_i \sim N(n\mu;\sqrt{n\sigma^2})$; $E(\sum_{i=1}^{n} X_i) = n\mu$; $Var(\sum_{i=1}^{n} X_i) = n\sigma^2$

c) $\sum_{i=1}^{n} X_i \sim$ Einpunkt$(n\mu)$; $E(\sum_{i=1}^{n} X_i) = n\mu$; $Var(\sum_{i=1}^{n} X_i) = 0$

Lösung 7-5: *Spielautomat*
X_i: "Gewinn beim Spiel i"; $X_i \sim N(0;1)$
$\sum_{i=1}^{16} X_i \sim N(0;4)$
$P(\sum_{i=1}^{n} X_i > 16) = 0,000032$

Lösung 7-6: *Anteil der Studentinnen an allen Studierenden*
Aus dem Aufgabentext: $\pi = 0,4$ und $n = 30$
Y: "Anzahl der Studentinnen bei einer einfachen Zufallsstichprobe vom Umfang $n = 30$"
$Y \sim B(n;\pi) = B(30;0,4)$

$$\begin{aligned}
P(Y < 0,3 \cdot 30) + P(Y > 0,5 \cdot 30) \quad &= \quad P(Y < 9) + P(Y > 15) \\
&= \quad P(Y \leq 8) + [1 - P(Y \leq 15)] \\
&= \quad 0,0940 + (1 - 0,9029) \\
&= \quad 0,1911
\end{aligned}$$

Lösung 7-7: *Lift*
X_i: "Gewicht von Person i"
$E(X_i) = 72$; $Var(X_i) = 24^2$
$\sum_{i=1}^{36} X_i$: "Gesamtgewicht der 36 Personen"
$\sum_{i=1}^{36} X_i$ ist approximativ [Z.G.S., $n > 30$] $N(n\mu;\sqrt{n\sigma^2})$
$P(\sum_{i=1}^{36} > 2800) = 0,074934$

Lösung 7-8: *Tippfehler*
Y: "Anzahl der Tippfehler im Manuskript"; $n = 1100$

a) $Y = \sum_{i=1}^{n} X_i$ ist approximativ (ZGWS.; $n > 30$) $N(n\mu; \sqrt{n\sigma^2})$,
$E(Y) = n\mu = 1540$; $Var(Y) = n\sigma^2 = 484$

b) (i) $P(Y \geq 1600) = 0,0032$,
(ii) $P(1500 \leq Y \leq 1650) = 0,9656$,
(iii)
   - ohne Stetigkeitskorrektur: $P(Y = 1540) = 0$,
   - mit Stetigkeitskorrektur (nicht zulässig, da $\sigma^2 > 9$):
     $P(1539,5 \leq Y \leq 1540,5) \approx \Phi(0,02) - \Phi(-0,02) = 0,015956$
     bzw. mit $\Phi(0,02272727) - \Phi(-0,02272727) = 0,01813218$

(iv) $P(Y \leq 1540) = 0,5$

Stichprobentheorie-Tippfehler (13 min).mp4

Lösung 7-9: *Tennislehrer*
X_i: "Anzahl der über den Zaun geschlagenen Bälle pro Schüler in einer Trainingsstunde"

x	0	1	2	3	4	5	6	7	$\sum$
$P(X_i = x)$	0	0,3	0,1	0	0	**0,2**	0,3	0,1	1

$$\begin{aligned} E(X_i) &= 1 \cdot 0,3 + 2 \cdot 0,1 + 5 \cdot 0,2 + 6 \cdot 0,3 + 7 \cdot 0,1 = 4 \\ Var(X_i) &= (1-4)^2 \cdot 0,3 + (2-4)^2 \cdot 0,1 + (5-4)^2 \cdot 0,2 \\ &+ (6-4)^2 \cdot 0,3 + (7-4)^2 \cdot 0,1 = 5,4 \end{aligned}$$

Y: "Anzahl der über den Zaun geschlagenen Bälle pro Monat"
1 Monat = 30 Tage mit 8 Trainungsstunden

$$\begin{aligned} Y &= \sum_{i=1}^{8\cdot 30} X_i \\ Y &\approx N(\bullet; \bullet) \text{ wegen ZGS} \\ E(Y) &= E\left(\sum_{i=1}^{240} X_i\right) = \sum_{i=1}^{240} E(X_i) = 240 \cdot 4 = 960 \\ Var(Y) &= Var\left(\sum_{i=1}^{240} X_i\right) = \sum_{i=1}^{240} Var(X_i) = 240 \cdot 5,4 = 1296 = 36^2 \end{aligned}$$

Die Kovarianzterme $Cov(X_i, X_j)$ fallen weg, da X_i und X_j unabhängig sind.

(a) Standardisierung: $Z = \dfrac{Y - E(Y)}{\sqrt{Var(Y)}} \approx N(0;1) \Rightarrow Z = \dfrac{Y - 960}{36}$

$$\begin{aligned} P(900 \leq Y \leq 1000) &= P\left(\frac{900-960}{36} \leq \frac{Y-960}{36} \leq \frac{1000-960}{36}\right) \\ &= P(-1,67 \leq Z \leq 1,11) \\ &= \Phi(1,11) - \Phi(-1,67) \\ &\quad \text{Eigenschaft der Standardnormalverteilung: } \Phi(-z) \\ &= \Phi(1,11) - (1 - \Phi(1,67)) \\ &= 0,81904 \end{aligned}$$

(b)

$$\begin{aligned} P(Y > 1050) &= P\left(\frac{Y-960}{36} > \frac{1050-960}{36}\right) \\ &= P(Z > 2,5) = 1 - \Phi(2,5) = 0,00621 \end{aligned}$$

(c)

$$\begin{aligned} 0,99 &= P(y_u \leq Y \leq y_o) \\ &= P\left(\frac{y_u-960}{36} \leq Z \leq \frac{y_o-960}{36}\right) \\ &= \Phi\left(\frac{y_o-960}{36}\right) - \Phi\left(\frac{y_u-960}{36}\right) \end{aligned}$$

$$\Phi\left(\frac{y_o-960}{36}\right) = 0,995 \quad \Rightarrow \quad \frac{y_o-960}{36} = 2,58 \quad (\text{da } \Phi(2,58) = 0,995)$$

$$\begin{aligned} y_u &= 960 - 2,58 \cdot 36 = 867,12 \\ y_o &= 960 + 2,58 \cdot 36 = 1052,88 \end{aligned}$$

(d) (i) $n = 240 > 30$

(ii) $E(X_i) = 4 \neq \pm\infty$

(iii) $0 < Var(X_i) = 5,4 < \infty$

(iv) X_i identisch verteilt, da nur ein Verteilungsmodell für alle Schüler angesetzt

(v) X_i unabhängig, da alles unterschiedliche Schüler

Hinweis: Es handelt sich nur um ein Modell wegen der letzen beiden Bedingungen

Lösung 7-10: *Ausschussanteil*

a) Menge der produzierten Teil $\approx$ unendlich groß e Grundgesamtheit; Binomialverteilung; $n = 50$, $\pi = 0,02 \Rightarrow$ Approximation mittels Poisson-Verteilung

b) X: "Anzahl defekter Stücke in Zufallsstichprobe $n = 50$"

x	0	1	2	3	4	5
$P(X = x)$	0,3679	0,3679	0,1839	0,0613	0,0153	0,0031

c) $P(X \leq k) = 0,95$, k ganzzahlig $\Rightarrow k = 3$

Stichprobentheorie-Ausschussanteil.mp4

Lösung 7-11: *Betriebsunfälle*

X: "Anzahl der Schwerverletzten bei 60 Betriebsunfällen"; $X \sim B(60; 0,2)$
$np = 12 > 5$; $n(1-\pi) = 48 > 5 \Rightarrow$ Approximation durch Normalverteilung
$\hat{\Pi}$: "Anteil der Schwerverletzten bei 60 Betriebsunfällen"; $X \approx N(0,2; 0,0516)$

a) $P(\hat{\pi} < 0,3) = 0,97381$

b) $P(\hat{\pi} > 0,2) = 0,5$

c) $P(0,15 \leq \hat{\pi} < 0,25) = 0,667954$

Lösung 7-12: *Urne*

$\hat{\pi}$: "Anteil roter Kugeln in der Stichprobe"; $\hat{\pi} = X/n$; X: "Anzahl roter Kugeln in der Stichprobe"; $X \sim H(N; M; n)$

a) $X \sim H(5; 2; 3)$; $P(p_1 \leq \hat{\pi} \leq p_2) = P(x_1 \leq X \leq x_2) = P(1 \leq X \leq 2) = 0,9$

b) $n/N < 0,05 \Rightarrow$ Approximation durch $B(4; 0,2)$; $P(1 \leq X \leq 3) = 0,5888$

c) $n\pi(1-\pi) = 16 > 9$; $0,1 < M/N < 0,9$; $n > 30 \Rightarrow$ Approximation durch N$(0,2; 0,031)$; bei Berücksichtigung der Stetigkeitskorrektur: $P(0,095 \leq \hat{\pi} \leq 0,305) = 0,999276$

d) Approximation durch $N(0,2; 0,04)$, wegen $n/N < 0,05$ kann Korrekturfaktor vernachlässigt werden; $P(0,14 \leq \hat{\pi} \leq 0,3) = 0,926983$

Lösung 7-13: *Tabletten gegen Kopfschmerzen, weiter*

a) $P(S^{*2} > 2\sigma^2) = P(nS^{*2}/\sigma^2 > 2n) = P(nS^{*2}/\sigma^2 > 14) \approx 0,05$
(weil $nS^{*2}/\sigma^2 \sim \chi^2_n$)

b) $P(S^{*2} > 2\sigma^2) = P(nS^{*2}/\sigma^2 > 2n) = P(nS^{*2}/\sigma^2 > 32) \approx 0,01$
(weil $nS^{*2}/\sigma^2 \sim \chi^2_n$)

c) $P(S^2 > 2\sigma^2) = P[(n-1)S^2/\sigma^2 > 2(n-1)] = P[(n-1)S^2/\sigma^2 > 30) \approx 0,01$
(weil $nS^2/\sigma^2 \sim \chi^2_{n-1}$)

latex

8 Statistische Schätzverfahren

Lösung 8-1: *Erwartungstreue*

- X_1, X_2, X_3 einfache Zufallsstichprobe
- $X_i \sim (\mu; \sigma^2)$
- X_i unabhängig

(a)

$$\begin{aligned}
E(\widehat{\theta}_1) &= E\left(\frac{1}{3}(X_1 + X_2 + X_3)\right) \\
&= \frac{1}{3}\left(E(X_1) + E(X_2) + E(X_3)\right) = \frac{1}{3} \cdot 3\mu = \mu \\
E(\widehat{\theta}_2) &= E\left(\frac{1}{4}(2X_1 + 2X_2)\right) = \\
&= \frac{1}{4}\left(2E(X_1) + 2E(X_2)\right) = \frac{1}{4} \cdot 4\mu = \mu \\
E(\widehat{\theta}_3) &= E\left(\frac{1}{3}(2X_1 + X_3)\right) \\
&= \frac{1}{3}\left(2E(X_1) + E(X_3)\right) = \frac{1}{3} \cdot 3\mu = \mu
\end{aligned}$$

(b)

$$\begin{aligned}
Var(\widehat{\theta}_1) &= Var\left(\frac{1}{3}(X_1 + X_2 + X_3)\right) \\
&= \frac{1}{9}(Var(X_1) + Var(X_2) + Var(X_3)) = \frac{3}{9}\sigma^2 = \frac{\sigma^2}{3} \\
Var(\widehat{\theta}_2) &= Var\left(\frac{1}{4}(2X_1 + 2X_3)\right) \\
&= \frac{1}{16}(4Var(X_1) + 4Var(X_2)) = \frac{8}{16}\sigma^2 = \frac{\sigma^2}{2} \\
Var(\widehat{\theta}_3) &= Var\left(\frac{1}{3}(2X_1 + X_3)\right) \\
&= \frac{1}{9}(4Var(X_1) + Var(X_3)) = \frac{5}{9}\sigma^2
\end{aligned}$$

$$Var(\widehat{\theta}_1) < Var(\widehat{\theta}_2) < Var(\widehat{\theta}_3)$$

Lösung 8-2: *Mittelwert und Varianz*
$\overline{x} = 3$ als Schätzwert für μ; $s^2 = 2,22$ als Schätzwert für σ^2

Schaetztheorie-Mittelwert und Varianz.mp4

Lösung 8-3: *Lampen*

a) X: "Anzahl der defekten Lampen in der Stichprobe"
$X \sim B(n;\pi)$ mit $\pi = d/N$
"Intuitive" Schätzfunktion: $\widehat{\theta} = N/n \cdot X = 50X$
$E(\widehat{\theta}) = N/n \cdot E(X) = N/n \cdot n \cdot d/N = d$

b) $\vartheta = 150$

c) Die einfache Zufallsstichprobe wird durch ein Ziehen mit Zurücklegen realisiert, d.h. mit geringer Wahrscheinlichkeit wird die gleiche Lampe zweimal gezogen. Sinnvoller wäre hier ein Ziehen ohne Zurücklegen, d.h. eine uneingeschränkte Zufallsstichprobe.

An dem Ergebnis in a) ändert sich nichts, da $X \sim Hyp(N,d,n) \approx B(n,\pi = d/n)$ ist $(n/N = 0,002 < 0,05)$.

Lösung 8-4: *Studienmotivation*

x	0	1	2	$\sum$
$P(X_i = x)$	π_0	π_1	π_2	$\pi_0 + \pi_1 + \pi_2 = 1$

(a)

$$\begin{aligned} E(X_i) &= 0 \cdot \pi_0 + 1 \cdot \pi_1 + 2 \cdot \pi_2 = \pi_1 + 2 \cdot \pi_2 \\ E(\widehat{\pi}_1) &= E\left(\frac{1}{10}\sum_{i=1}^{10}(2X_i - X_i^2)\right) = \frac{1}{10}\sum_{i=1}^{10}\Big(2E(X_i) - E(X_i^2)\Big) \end{aligned}$$

x^2	0	1	4	$\sum$
$P(X_i^2 = x^2)$	π_0	π_1	π_2	$\pi_0 + \pi_1 + \pi_2 = 1$

$$\begin{aligned} E(X_i^2) &= 0 \cdot \pi_0 + 1 \cdot \pi_1 + 4 \cdot \pi_2 = \pi_1 + 4 \cdot \pi_2 \\ E(\widehat{\pi}_1) &= \frac{1}{10}\sum_{i=1}^{10}(2(\pi_1 + 2\pi_2) - (\pi_1 + 4\pi_2)) = \frac{1}{10}\sum_{i=1}^{10}\pi_1 = \pi_1 \\ E(\widehat{\pi}_2) &= E\left(\frac{1}{20}\sum_{i=1}^{10}(X_i^2 - X_i)\right) = \frac{1}{20}\sum_{i=1}^{10}\Big(E(X_i^2) - E(X_i)\Big) \\ &= \frac{1}{20}\sum_{i=1}^{10}((\pi_1 + 4\pi_2) - (\pi_1 + 2\pi_2)) = \frac{1}{20}\sum_{i=1}^{10}2\pi_2 = \pi_2 \end{aligned}$$

(b) Schätzfunktion: $\widehat{\pi}_0 = 1 - \widehat{\pi}_1 - \widehat{\pi}_2$

$$\begin{aligned} E(\widehat{\pi}_0) &= E(1 - \widehat{\pi}_1 - \widehat{\pi}_2) \\ &= 1 - E(\widehat{\pi}_1) - E(\widehat{\pi}_2) \\ &= 1 - \pi_1 - \pi_2 = \pi_0 \end{aligned}$$

(c) Stichprobe:

i	1	2	3	4	5	6	7	8	9	10
x_i	0	2	1	0	0	2	2	1	0	2

x_i	$2x_i - x_i^2$	$x_i^2 - x_i$
0	0	0
1	1	0
2	0	2

$$\begin{aligned} \widehat{\pi}_1 &= \frac{1}{10}\sum_{i=1}^{10}(2x_i - x_i^2) = \frac{0+0+1+0+0+0+0+1+0+0}{10} = 0,2 \\ \widehat{\pi}_2 &= \frac{1}{20}\sum_{i=1}^{10}(x_i^2 - x_i) = \frac{0+2+0+0+0+2+2+0+0+2}{20} = 0,4 \\ \widehat{\pi}_0 &= 1 - 0,2 - 0,4 = 0,4 \end{aligned}$$

Lösung 8-5: *Spielautomat*

a) Beim 1. Spiel Verlust von 1 EUR; beim 2. Spiel Gewinn von 1 EUR; ...

b) $P(X=0)=p$; $P(X=-1)=p$; $\mathrm{P}(X=1)=1-2p$

c) $f(X=0)=1/6$; $f(X=-1)=2/6$; $f(X=1)=3/6$

d) $L(x_1,...,x_6|p)=p^3\cdot(1-2p)^3$

e) v - Anzahl der verlorenen Spiele;
u - Anzahl der unentschiedenen Spiele;
g - Anzahl der gewonnenen Spiele;
$n=v+u+g$ $\widehat{p}=(v+u)/2n$

f) $\widehat{p}=1/4$

g) $\widehat{p}=5/18$

Lösung 8-6: *Finanzamt*

a) $L(\lambda)=\lambda^3\cdot e{-}^{9\lambda}$

b) $\widehat{\lambda}=1/3 \Rightarrow$ Frau Hurtig

Schaetztheorie-Finanzamt.mp4

Lösung 8-7: *Startprobleme*

a) $L(p)=(1-p)^8\cdot p^2$

b) $\widehat{p}=0,2$

Lösung 8-8: *Dichotome Grundgesamtheit*
$Q(\pi)=3\pi^2-2\pi+1$; $\widehat{\pi}=1/3$

Lösung 8-9: *Unfallhäufigkeit*

a) $L_{Vers.}(0,2,0,2,1)=0,00196$; $L_{Pol.}(0,2,0,2,1)=0,001$
$\Rightarrow$ Versicherungsgesellschaft

b) $Q_{Vers.}(0,2,0,2,1)=5,25$; $Q_{Pol.}(0,2,0,2,1)=4,8$
$\Rightarrow$ Polizei

Schaetztheorie-Unfallhaeufigkeit.mp4

Lösung 8-10: *Likelihood-Funktion*

a) Likelihood-Funktion:

$$\begin{aligned} L(\lambda;x_1,x_2,x_3,x_4) &= f(x_1|\lambda)\cdot f(x_2|\lambda)\cdot f(x_3|\lambda)\cdot f(x_4|\lambda) \\ &= \frac{\lambda^2}{2!}e^{-\lambda}\cdot\frac{\lambda^4}{4!}e^{-\lambda}\cdot\frac{\lambda^6}{6!}e^{-\lambda}\cdot\frac{\lambda^3}{3!}e^{-\lambda}=\frac{\lambda^{15}}{2!4!6!3!}e^{-4\lambda} \end{aligned}$$

b) ML-Schätzwert für λ:

Poisson-Verteilung: $\quad$ Log-Likelihood-Funktion:

$$f_{PO}(x;\lambda)=\frac{\lambda^x}{x!}e^{-\lambda} \qquad \ln L(\lambda)=15\ln\lambda-\ln(2!4!6!3!)-4\lambda$$

$$\frac{\delta\ln L(\lambda)}{\delta\lambda}=\frac{15}{\hat{\lambda}}-4=0$$
$$\hat{\lambda}=\frac{15}{4}=3,75$$

Lösung 8-11: *Fahrradschläuche*
X: "Durchmesser eines Fahrradschlauches"; $X \sim N(\mu; \sigma)$
$\overline{X}$: "Mittlerer Durchmesser eines Fahrradschlauches bei einer Zufallsstichprobe vom Umfang $n = 25$"; $\overline{X} \sim N(\mu; \sigma/\sqrt{n})$

a)
$$P(\overline{X} - c\frac{S}{\sqrt{n}} \leq \mu \leq \overline{X} + c \cdot \frac{S}{\sqrt{n}}) = 1 - \alpha$$
c aus t–Verteilung mit $f = 24$: $c = 1,711$

b) $[39,9734; 42,0266]$

c) $n \geq 293$

Lösung 8-12: *Schweinemäster*
X: "Gewicht eines Schweins"; $X \sim N(\mu; \sigma)$
$\overline{X}$: "Durchschnittliches Gewicht eines Schweins bei einer Zufallsstichprobe vom Umfang $n = 6$"; $\overline{X} \sim N(\mu; \sigma/\sqrt{n})$

a) $\overline{X} = 100$; $s^2 = 6$

b)
$$P(\overline{X} - c \cdot \frac{S}{\sqrt{n}} \leq \mu \leq \overline{X} + c \cdot \frac{S}{\sqrt{n}}) = 1 - \alpha$$
c aus t–Verteilung mit $f = 5$: $c = 2,571$

c) $[97,429; 102,571]$

d) Wird dieses Verfahren der Intervallschätzung "unendlich oft" durchgeführt, so kann man mit (durchschnittlich) 95% richtigen Ergebnissen rechnen; d.h. Schätzintervallen, in denen der unbekannte Wert enthalten ist.

e) durch die Wahl von n bzw. α

f) Intervall wird kleiner: statt $c = 2,571$ aus t–Verteilung ist $c = 1,96$ aus $N(0; 1)$ zu verwenden.

Konfidenzintervall-Schweinmaester (14 min).mp4

Lösung 8-13: *Milchfettgehalt*
X:„Milchfettgehalt", $\mu = 3,7352$, $\sigma^2 = 0,0081$, $X \sim N(3,7352; 0,09)$

$$P(X > x) = P\Big(\frac{X - 3,7352}{0,09} > \frac{x - 3,7352}{0,09}\Big) = P(Z > z) = 0,61$$

Aus der vorliegenden Tabelle der Standardnormalverteilung findet man für $P(Z \leq z) = 0,61$ den Wert $z = 0,28$, so dass der gesuchte Wert $z = -0,28$ ist.
$(x - 3,7352)/0,09 = -0,28$; **x=3,71**

Lösung 8-14: *Apfelsinen*

- X : "Gewicht der Apfelsinen" $\sim N(\mu; \sigma = 20g)$
- Einfache Zufallsstichprobe mit $n = 25$
- Summe des Gewichts: $7500g \Rightarrow \bar{x} = \frac{7500}{25} = 300g$

Allgemeines Konfidenzintervall für den Mittelwert μ der Grundgesamtheit:

$$P\left(\bar{X} - c\frac{\sigma}{\sqrt{n}} \leq \mu \leq \bar{X} + c\frac{\sigma}{\sqrt{n}}\right) = 1 - \alpha$$

c aus $N(0; 1)$, da σ bekannt $\Longrightarrow c = z_{1-\alpha/2}$
Schätzintervall für den Mittelwert μ der Grundgesamtheit:

$1 - \alpha = 80,64\% \Rightarrow 1 - \alpha/2 = 90,32\% \Rightarrow \Phi(z_{1-\alpha/2}) = 0,9032 \Rightarrow z_{1-\alpha/2} = 1,3$

$$\left[\bar{x} - z_{1-\alpha/2}\frac{\sigma}{\sqrt{n}}; \bar{x} + z_{1-\alpha/2}\frac{\sigma}{\sqrt{n}}\right]$$
$$\left[300 - 1,3\frac{20}{\sqrt{25}}; 300 + 1,3\frac{20}{\sqrt{25}}\right] = [294,8; 305,2]$$

Konfidenzintervall-Apfelsinen.mp4

Lösung 8-15: *PKWs in Berlin*
Da keine Informationen über π in Form einer Vorstichprobe gegeben sind, wird der ungünstigste Fall angenommen und $\hat{\pi}$ so gewählt, dass die Varianz $\sigma^2_{\hat{\pi}} = \hat{\pi} \cdot (1 - \hat{\pi})$ maximal ist. Dieser ungünstigste Fall tritt bei $\hat{\pi} = 0,5$ ein.
$1 - \alpha = 0,95$; $z_{0,975} = 1,96$; $I = 0,06$; $n = z^2_{1-\alpha/2}/I^2 = 1,96^2/0,06^2 = 1067,11; \rightarrow n \geq 1068$.

Lösung 8-16: *Konfidenzniveau*
$\overline{x} - t_{1-\alpha/2;f} \cdot s/\sqrt{n} = 1,6$; $s = 10$; $\overline{x} = (6,4 + 1,6)/2 = 4$; $4 - t_{1-\alpha/2;f} \cdot 10/10 = 1,6$; $f = 99$;
$t_{1-\alpha/2;f} = z_{1-\alpha/2} = 2,4$; $1 - \alpha/2 = 0,991802$; $\alpha/2 = 0,008198$; $\alpha = 0,016396$; $1 - \alpha = 0,983604$

Lösung 8-17: *Gasverbrauch*
$P(\mu - t_{1-\alpha/2;f} \cdot s/\sqrt{n} \leq \bar{X} \leq \mu + t_{1-\alpha/2;f} \cdot s/\sqrt{n}) = 1 - \alpha$;
$[\mu - t_{1-\alpha/2;f} \cdot s/\sqrt{n}; \mu + t_{1-\alpha/2;f} \cdot s/\sqrt{n}]$
$n = 36$; $\frac{\bar{X}-\mu}{S/\sqrt{n}} \sim t_f \approx N(0;1)$
$\mu = 5$; $s = 2$; $1 - \alpha = 0,95$; $1 - \alpha/2 = 0,975$; $z_{0,975} = 1,96$
$[5 - 1,96 \cdot 2/6; 5 + 1,96 \cdot 2/6] = [5 - 0,6533; 5 + 0,6533] = [4,3467; 5,5633]$
Konfidenzintervall-Gasverbrauch.mp4

Lösung 8-18: *Kugelschreiber*
Gewicht der Schreibminen: $M \sim N(\mu_M; \sigma_M) = N(\mu_M; 0,4)$; Gewicht der Metallfedern: $F \sim N(\mu_F; \sigma_F) = N(\mu_F; 0,2)$; Gewicht der Kunststoffhüllen: $H \sim N(\mu_H; \sigma_H) = N(\mu_H; 0,4)$
Gesamtgewicht eines Kugelschreibers: $X = M + F + H$; $X \sim N(\mu_X; \sigma_X) = N(\mu_X; 0,6)$
$\sigma^2_X = \sigma^2_M + \sigma^2_F + \sigma^2_H = 0,4^2 + 0,2^2 + 0,4^2 = 0,36$ (wegen Unabhängigkeit von M, F und H)
$\overline{X}$: „Durchschnittsgewicht eines Kugelschreibers in einer einfachen Zufallsstichprobe vom Umfang $n = 25$“; $\overline{X} \sim N(\mu; \sigma/\sqrt{n}) = N(\mu; 0,12)$; $\sigma/\sqrt{n} = 0,6/5 = 0,12$
Konfidenzniveau: $P(\overline{X} - z_{1-\alpha/2} \cdot \sigma/\sqrt{n} \leq \mu \leq \overline{X} + z_{1-\alpha/2} \cdot \sigma/\sqrt{n}) = 0,8064$; $z_{1-\alpha/2}$ aus $N(0;1)$
$z_{0,9032} = 1,3$; $\overline{x} = 375/25 = 15$
Schätzintervall:
$[\overline{x} - z_{1-\alpha/2} \cdot \sigma/\sqrt{n}; \overline{x} + z_{1-\alpha/2} \cdot \sigma/\sqrt{n}]$
$[15 - 1,3 \cdot 0,12; 15 + 1,3 \cdot 0,12] = [14,844; 15,156]$

Lösung 8-19: *Kaltwasserverbrauch*
X: „Kaltwasserverbrauch pro Spülgang“; $X \sim N(\mu; \sigma) = N(\mu; 2)$
$\overline{X}$: „Durchschnittlicher Kaltwasserverbrauch in einer einfachen Zufallsstichprobe vom Umfang n“; $\overline{X} \sim N(\mu; \sigma/\sqrt{n}) = N(\mu; 2/\sqrt{n})$
Konfidenzniveau: $P(\overline{X} - z_{1-\alpha/2} \cdot \sigma/\sqrt{n} \leq \mu \leq \overline{X} + z_{1-\alpha/2} \cdot \sigma/\sqrt{n}) = 0,95$; $z_{1-\alpha/2}$ aus $N(0;1)$;
$z_{0,975} = 1,96$; $e = z_{1-\alpha/2} \cdot \sigma/\sqrt{n}$; $e = 0,5 = 1,96 \cdot 2/\sqrt{n}$; $n = 2^2 \cdot 1,96^2/0,5^2 = 15,3664/0,25 = 61,4656$; $n \geq 62$

Lösung 8-20: *Konfidenzniveau 2*
$\overline{x}-t_{1-\alpha/2;f}\cdot s/\sqrt{n}=2,1$; $s=10$; $\overline{x}=(2,1+5,9)/2=4$; $4-t_{1-\alpha/2;f}\cdot 10/10=2,1$; $f=99$; $t_{1-\alpha/2;f}=z_{1-\alpha/2}=1,9$; $1-\alpha/2=0,971283$; $\alpha/2=0,028717$; $\alpha=0,057434$; $1-\alpha=0,942566$

Lösung 8-21: *Stichprobenmittelwert*
$P(\mu-t_{1-\alpha/2;f}\cdot s/\sqrt{n}\leq\overline{X}\leq\mu+t_{1-\alpha/2;f}\cdot s/\sqrt{n})=1-\alpha$; $[\mu-t_{1-\alpha/2;f}\cdot s/\sqrt{n};\mu+t_{1-\alpha/2;f}\cdot s/\sqrt{n}]$
$\mu=5$; $s=2$; $n=25$; $f=24$; $1-\alpha=0,99$; $1-\alpha/2=0,995$; $t_{0,995;24}=2,797$
$[5-2,797\cdot 2/5;5+2,797\cdot 2/5]=5-1,1188;5+1,1188]=[3,8812;6,1188]$

Lösung 8-22: *Konzentration des Stoffes E*
X: „Konzentration von E im Wasser"; $X\sim N(\mu;\sigma)$
$\overline{X}$:„Mittlere Konzentration von E im Wasser bei Zufallsstichprobe $n=9$";
$\overline{X}\sim N(\mu;\sigma/\sqrt{n})$
σ unbekannt, mittels $S^2=\sum(X_i-\overline{X})^2/(n-1)$ schätzen
$T=(\overline{X}-\mu)\sqrt{n}/S$ ist t-verteilt mit $f=n-1$ Freiheitsgraden
$P[\overline{X}-t_{f;1-\alpha/2}\cdot S/\sqrt{n}\leq\mu\leq\overline{X}+t_{f;1-\alpha/2}\cdot S/\sqrt{n}]=1-\alpha=0,9$
$\overline{x}=90/9=10$; $s^2=(0+4+0+4+81+25+1+9+4)/8=128/8=16$; $s=4$; $t_{8;0,95}=1,86$
$[10-1,86\cdot 4/3;10+1,86\cdot 4/3]=[7,520;12,480]$

Lösung 8-23: *Brikett*
X : „Gewicht eines Briketts"; $X\sim N(500;50)$
$\overline{X}$: „Durchschnittliches Gewicht eines Briketts bei einer Zufallsstichprobe $n=25$"
$\overline{X}\sim N(500;10)$; $z=(\overline{X}-\mu_0)\sqrt{n}/\sigma$; $z=(510-500)5/50=1$; $P(Z\leq 1)=0,841345$; $1-P(Z\leq 1)=1-0,841345=0,158655$

Lösung 8-24: *Mietverein*
$X\sim N(\mu;\sigma=180\text{ EUR})$; $\ell=120$; $1-\alpha=0,95$; $1-\alpha/2=0,975$; $z_{0,975}=1,96$
$n\geq(4\sigma^2z^2_{1-\alpha/2})\ell^2$
$n\geq(4\cdot 180^2\cdot 1,96^2)/120^2=(4\cdot 32400\cdot 3,8416)/14400=497871,36/14400=34,5744$
$n\geq 35$

Lösung 8-25: *Jährliche Fahrleistung*

a) $\left[\overline{X}-t_{1-\alpha/2;n-1}\cdot\frac{S}{\sqrt{n}};\overline{X}+t_{1-\alpha/2;n-1}\cdot\frac{S}{\sqrt{n}}\right]$, so dass
$P\left(\overline{X}-t_{1-\alpha/2;n-1}\cdot\frac{S}{\sqrt{n}}\leq\mu\leq\overline{X}+t_{1-\alpha/2;n-1}\cdot\frac{S}{\sqrt{n}}\right)=1-\alpha$

b) $\left[\overline{x}-t_{1-\alpha/2;n-1}\cdot\frac{s}{\sqrt{n}};\overline{x}+t_{1-\alpha/2;n-1}\cdot\frac{s}{\sqrt{n}}\right]$
$\overline{x}=25$; $s^2=80$; $n=20$; $1-\alpha=0,95$; $t_{0,975;19}=2,093$
$\left[25-2,093\sqrt{\frac{80}{20}};25+2,093\sqrt{\frac{80}{20}}\right]$
$=[25-4,186;25+4,186]$
$=[20,814\ (1000\text{ km});29,186\ (1000\text{ km})]$

Lösung 8-26: *Dioxinausstoß*
X: „Dioxinausstoß [kg/min]", $X\sim N(5;1)$
$\overline{X}$: „Durchschnittlicher Dioxinausstoß [kg/min]", $X\sim N(5;1/3)$

a) Berechnung der statistischen Sicherheit für ein gegebenes Schwankungsintervall
$P\left(\mu-c\cdot\frac{\sigma}{\sqrt{n}}\leq\overline{X}\leq\mu+c\cdot\frac{\sigma}{\sqrt{n}}\right)=1-\alpha$

$P(4\leq\overline{X}\leq 6)=1-\alpha=?$

$P\left(\frac{4-5}{1}\sqrt{9}\leq\overline{X}\leq\frac{6-5}{1}\sqrt{9}\right)=P(-3\leq Z\leq 3)=2\cdot P(Z\leq 3)-1$
$=2\cdot 0,99865-1=0,9973$

Mit einer Wahrscheinlichkeit von 99,73% liegt der Durchschnitt einer Stichprobe vom Umfang $n = 9$ zwischen 4 und 6 kg/min Dioxinausstoß.

b) symmetrisches Schwankungsintervall gesucht bei gegebener statistischer Sicherheit

$$P\left(\mu - c\frac{\sigma}{\sqrt{n}} \leq \overline{X} \leq \mu + c\frac{\sigma}{\sqrt{n}}\right) = 0,95$$

$$P\left(5 - \frac{c}{3} \leq \overline{X} \leq 5 + \frac{c}{3}\right) = 0,95$$

$$c_{0,975} = 1,96$$

$$\left[5 - \frac{1,96}{3}; 5 + \frac{1,96}{3}\right] = [4,347; 5,653]$$

c) $n \geq \dfrac{\sigma^2 \cdot z^2_{1-\alpha/2}}{e^2}$

$$\Leftrightarrow n \geq \frac{1 \cdot 1,96^2}{0,5^2} = 15,37 \approx 16$$

Um mit einer Sicherheit von 95% den durchschnittlichen Dioxinausstoß auf 0,5 kg/min genau schätzen zu können, benötigt man einen Stichprobenumfang von mindestens 16 Zeitintervallen.

d) $P\left(\overline{X} - c \cdot \dfrac{\sigma}{\sqrt{n}} \leq \mu \leq \overline{X} + c \cdot \dfrac{\sigma}{\sqrt{n}}\right) = 1 - \alpha$
c aus $N(0;1)$

e) $1 - \alpha = 0,98$; $1 - \alpha/2 = 0,99$; $c = 2,33$; $\overline{x} = 63/9 = 7$ kg/min; $\sigma = 1$
$[7 - 2,33/3; 7 + 2,33/3] = [6,22; 7,78]$

Lösung 8-27: *Mietverein 2*

a) X: „Mietpreis einer 80m^2–Altbauwohnung“
X ist beliebig verteilt mit $E(X) = \mu$ und $Var(X) = \sigma^2$
$\overline{X}$: „Durchschnittlicher Mietpreis einer 80m^2–Altbauwohnung bei einer Zufallsstichprobe von $n = 36$“
$\overline{X}$ ist approximativ (zentraler Grenzwertsatz; $n > 30$) $N(\mu; \sigma/\sqrt{n})$–verteilt.
σ^2 ist unbekannt und wird mittels der Stichprobenfunktion

$$S^2 = \frac{1}{n-1}\sum_{i=1}^{n}(X_i - \overline{X})^2$$

geschätzt.

$$P\left(\overline{X} - c \cdot \frac{S}{\sqrt{n}} \leq \mu \leq \overline{X} + c \cdot \frac{S}{\sqrt{n}}\right) = 1 - \alpha;\ c \text{ aus } N(0;1)$$

b) $1 - \alpha = 0,98 \rightarrow c = 2,33$; $\overline{x} = 950$; $s = 180$
$[950 - 2,33 \cdot 180/6; 950 + 2,33 \cdot 180/6] = [950 - 69,9; 950 + 69,9] = [880,1; 1019,9]$

Lösung 8-28: *Love–Parade*

X : „Ausgaben“, X ist beliebig verteilt mit $E(X) = \mu$ und $Var(X) = \sigma^2$
$\overline{X}$: „Durchschnittliche Ausgaben bei einer Zufallsstichprobe vom Umfang $n = 100$“
$\overline{X}$ ist approximativ (zentraler Grenzwertsatz; $n > 30$) $N(\mu; s/\sqrt{n})$–verteilt.
$\overline{x} = 350$, $s = 24$, $z_{1-\alpha/2} = z_{0,975} = 1,96$
$[\overline{x} - z_{1-\alpha/2} \cdot s/\sqrt{n} \leq \mu \leq \overline{x} + z_{1-\alpha/2} \cdot s/\sqrt{n}] = [350 - 1,96 \cdot 24/\sqrt{n}; 350 + 1,96 \cdot 24/\sqrt{n}] = [345,296; 354,704]$

Lösung 8-29: *Langlebensdauergarantie*
X : „Brenndauer", ist beliebig verteilt mit $E(X) = \mu$ und $Var(X) = \sigma^2$
$\overline{X}$: „Durchschnittliche Brenndauer bei einer Zufallsstichprobe vom Umfang $n = 100$"
$\overline{X}$ ist approximativ (zentraler Grenzwertsatz; $n > 30$) $N(\mu; s/\sqrt{n})$–verteilt.
$\overline{x} = 1300$, $s = 100$, $z_{1-\alpha/2} = z_{0,99} = 2,33$
$P(\overline{X} - z_{1-\alpha/2} \cdot S/\sqrt{n} \le \mu \le \overline{X} + z_{1-\alpha/2} \cdot S/\sqrt{n}) \approx 1 - \alpha = 0,98$;
$[\overline{x} - z_{1-\alpha/2} \cdot s/\sqrt{n} \le \mu \le \overline{x} + z_{1-\alpha/2} \cdot s/\sqrt{n}]$
$\approx [1300 - 2,33 \cdot 100/\sqrt{100}; 1300 + 2,33 \cdot 100/\sqrt{100}] \approx [1276,7; 1323,3]$

Lösung 8-30: *Trinkwasserverbrauch*
X : „Wasserverbrauch", ist beliebig verteilt mit $E(X) = \mu$ und $Var(X) = \sigma^2$
$\overline{X}$: „Durchschnittlicher Wasserverbrauch bei einer Zufallsstichprobe vom Umfang $n = 100$"
$\overline{X}$ ist approximativ (zentraler Grenzwertsatz; $n > 30$) $N(\mu; s/\sqrt{n})$–verteilt.
$\overline{x} = 12$, $s^2 = \sum(x_i - \overline{x})^2/(n-1) = 5$, $z_{1-\alpha/2} = z_{0,975} = 1,96$
$P(\overline{X} - z_{1-\alpha/2} \cdot S/\sqrt{n} \le \mu \le \overline{X} + z_{1-\alpha/2} \cdot S/\sqrt{n}) \approx 1 - \alpha = 0,95$;
$[\overline{x} - z_{1-\alpha/2} \cdot s/\sqrt{n} \le \mu \le \overline{x} + z_{1-\alpha/2} \cdot s/\sqrt{n}] \approx [12 - 1,96 \cdot 5/\sqrt{100}; 12 + 1,96 \cdot 5/\sqrt{100}] \approx [11,02; 12,98]$

Lösung 8-31: *500 Haushalte*
X : „Haushaltsgröße ", ist beliebig verteilt mit $E(X) = \mu$ und $Var(X) = \sigma^2$
$\overline{X}$: „Durchschnittliche Haushaltsgröße bei einer Zufallsstichprobe vom Umfang $n = 100$"
$\overline{X}$ ist approximativ (zentraler Grenzwertsatz; $n > 30$) $N(\mu; s/\sqrt{n})$–verteilt.
$\overline{x} = 2,73$, $s^2 = \sum(x_i - \overline{x})^2/(n-1) = 1,58$, $z_{1-\alpha/2} = z_{0,975} = 1,96$
$P(\overline{X} - z_{1-\alpha/2} \cdot S/\sqrt{n} \le \mu \le \overline{X} + z_{1-\alpha/2} \cdot S/\sqrt{n}) \approx 1 - \alpha = 0,95$
, $[\overline{x} - z_{1-\alpha/2} \cdot s/\sqrt{n} \le \mu \le \overline{x} + z_{1-\alpha/2} \cdot s/\sqrt{n}] \approx [2,73 - 1,96 \cdot \sqrt{1,58/500}; 2,73 + 1,96 \cdot \sqrt{1,58/500}] \approx [2,620; 2,840]$

Lösung 8-32: *Glücksspiel*
X : „Ertrag ", ist beliebig verteilt mit $E(X) = \mu$ und $Var(X) = \sigma^2$
$\overline{X}$: „Durchschnittlicher Ertrag bei einer Zufallsstichprobe vom Umfang $n = 50$"
$\overline{X}$ ist approximativ (zentraler Grenzwertsatz; $n > 30$) $N(\mu; s/\sqrt{n})$–verteilt.
$\overline{x} = -0,58$, $\quad s^2 = \sum(x_i - \overline{x})^2/(n-1) = 0,82$, $z_{1-\alpha/2} = z_{0,975} = 1,96$
$P(\overline{X} - z_{1-\alpha/2} \cdot S/\sqrt{n} \le \mu \le \overline{X} + z_{1-\alpha/2} \cdot S/\sqrt{n}) \approx 1 - \alpha = 0,95$
$[\overline{x} - z_{1-\alpha/2} \cdot s/\sqrt{n} \le \mu \le \overline{x} + z_{1-\alpha/2} \cdot s/\sqrt{n}] \approx [-0,58 - 1,96 \cdot \sqrt{0,0164}; -0,58 + 1,96 \cdot \sqrt{0,0164}] \approx [-0,831; -0,329]$

Lösung 8-33: *Jährliche Fahrleistung 2*
X : „Jährliche Fahrleistung ", X ist normalverteilt
$\mu = 25, \sigma$ unbekannt, $s^2 = 80$, $n = 20, f = 19, 1 - \alpha = 0,95, t_{1-0,975;19} = 2,093$
$P(\mu - t_{1-\alpha/2;f} \cdot s/\sqrt{n} \le \overline{X} \le \mu + t_{1-\alpha/2;f} \cdot s/\sqrt{n}) = 1 - \alpha$;
$[\mu - t_{1-\alpha/2;f} \cdot s/\sqrt{n}; \mu + t_{1-\alpha/2;f} \cdot s/\sqrt{n}]$;
$$\left[25 - 2,093 \cdot \sqrt{\frac{80}{20}}; 25 + 2,093 \cdot \sqrt{\frac{80}{20}}\right] = [25 - 4,186; 25 + 4,186] = [20,814; 29,186]$$

Lösung 8-34: *Antibiotikumtabletten*
Grundgesamtheit: X: „Wirkstoffgehalt je Tablette"; $X \sim N(\mu; 10)$
$\overline{X}$: „Durchschnittlicher Wirkstoffgehalt je Tablette bei einer einfachen Zufallsstichprobe vom Umfang n"; $\overline{X} \sim N(\mu; 10/\sqrt{n})$
$P(\overline{X} - 2 \le \mu \le \overline{X} + 2) = P(\overline{X} - z_{1-\alpha/2} \cdot 10/\sqrt{n} \le \mu \le \overline{X} + z_{1-\alpha/2} \cdot 10/\sqrt{n}) = 0,98$
$z_{0,99} = 2,33$; $\quad n \ge (\sigma z/e)^2 = (10 \cdot 2,33/2)^2 = 11,65^2 = 135,7225 \to n \ge 136$

Lösung 8-35: *Weizenhektarerträge*
Grundgesamtheit sind alle Hektarflächen in Deutschland, auf denen 1996 Weizen angebaut wurde; X: „Hektarertrag für Weizen“; Verteilung von X unbekannt; $\sigma^2 = 324$ [dt/ha]^2. $\overline{X}$: „Durchschnittlicher Hektarertrag für Weizen in Deutschland bei einer einfachen Zufallsstichprobe vom Umfang $n = 144$“
Verteilung von $\overline{X}$ unbekannt. Da aber der Stichprobenumfang $n > 30$ ist, kann aufgrund des Zentralen Grenzwertsatzes approximativ die Normalverteilung verwendet werden:
$\overline{X} \sim N(\mu; \sigma/\sqrt{n})$; $\quad P(\overline{X} - z_{1-\alpha/2} \cdot \sigma/\sqrt{n} \le \mu \le \overline{X} + z_{1-\alpha/2} \cdot \sigma/\sqrt{n}) = 0,8064$
Schätzintervall: $[\overline{x} - z_{1-\alpha/2} \cdot \sigma/\sqrt{n}; \overline{x} + z_{1-\alpha/2} \cdot \sigma/\sqrt{n}]$
$1 - \alpha/2 = 0,9032$; $\quad z_{0,9032} = 1,3$ aus $N(0;1)$; $\quad n = 144$; $\quad \sigma = 18$
Stichprobenmittelwert für Deutschland:
$\overline{x} = (60,4 \cdot 48 + 68,8 \cdot 96)/144 = (2899,2 + 6604,8)/144 = 66$ dt/ha;
Schätzintervall: $[66 - 1,3 \cdot 18/12; 66 + 1,3 \cdot 18/12] = [66 - 1,95; 66 + 1,95] = [64,05; 67,95]$

Lösung 8-36: *Fluggesellschaft*
$1 - \alpha/2 = 0,995$; $\quad z_{0,995} = 2,58$
$0,9 \pm 2,58\sqrt{\dfrac{0,9 \cdot 0,1}{200}}$
[84, 5%; 95, 5%]

Lösung 8-37: *Faktenmagazin*
Konfidenzintervall für den Erwartungswert μ:

$$\left[\overline{x} - t_{1-\alpha/2;f}\frac{s}{\sqrt{n}}; \overline{x} + t_{1-\alpha/2;f}\frac{s}{\sqrt{n}}\right]$$

$t_{0,995;24} = 2,797$
$[74000 - 2,797 \cdot 10000/5; 74000 + 2,797 \cdot 10000/5] = [68406; 79594]$

Lösung 8-38: *Handybesitzer*
Da keine Information über π in Form einer Vorstichprobe oder anderweitig gegeben ist, wird der ungünstigste Fall angenommen und $\hat{\pi}$ so gewählt, dass die Varianz $\sigma^2_{\hat{\pi}} = \hat{\pi}(1 - \hat{\pi})$ maximal wird. Dieser ungünstigste Fall tritt bei $\hat{\pi} = 0,5$ ein.
$1 - \alpha = 0,95$; $\quad z_{0,975} = 1,96 \quad \ell = 0,06$
$n = z^2_{1-\alpha/2}/\ell^2 = 1,96^2/0,06^2 = 3,8416/0,0036 = 1067,11 \to n = 1068$

Lösung 8-39: *Jährliche Fahrleistung 3*
X : „Jährliche Fahrleistung “, X ist normalverteilt
$\overline{x} = 25$,σ unbekannt, $s^2 = 90,25$, $n = 100$, $\overline{X}$ ist approximativ normalverteilt ($n \ge 30$)
Konfidenzniveau $1 - \alpha = 0,95$; $z_{0,975} = 1,96$
Schätzintervall:

$$\left[25 \pm 1,96 \cdot \sqrt{\frac{90,25}{100}}\right] = [25 \pm 1,96 \cdot 0,95] = [25 \pm 1,862] = [23,138; 26,862]$$

Lösung 8-40: *Notwendiger Stichprobenumfang*
Da keine Informationen über π in Form einer Vorstichprobe gegeben sind, wird der ungünstigste Fall angenommen und $\hat{\pi}$ so gewählt, dass die Varianz $\sigma^2_{\hat{\pi}} = \hat{\pi}(1 - \hat{\pi})$ maximal wird. Dieser ungünstigste Fall tritt bei $\hat{\pi} = 0,5$ ein.

$$1 - \alpha = 0,99; \quad z_{0,995} = 2,58; \quad \ell = 0,05;$$

$$n \ge z^2_{1-\alpha/2}/\ell^2 = 2,58^2/0,05^2 = 2662,56 \to n \ge 2663$$

Lösung 8-41: *Eintagsfliegen*
X : „Lebensdauer von Eintagsfliegen"$\sim N(\mu;\sigma^2)$, μ und σ^2 unbekannt
$n = 16$ (kleine Stichprobe); $\overline{x} = 1440$; $s^2 = 57600$, $s = 240$
Schätzintervall:
$\left[\overline{x} \pm t_{n-1;1-\alpha/2} \cdot \frac{s}{\sqrt{n}}\right] = \left[1440 \pm t_{15;1-\alpha/2} \cdot \frac{240}{\sqrt{16}}\right]$
$= \left[1440 \pm t_{15;1-\alpha/2} \cdot 60\right] = [1263,12; 1616,88]$ $1616,88 = 1440 + t_{15;1-\alpha/2} \cdot 60;$ $t_{15;1-\alpha/2}$
$= 2,948;$ $1-\alpha/2 = 0,995$ (aus t-Verteilung);
$1-\alpha = 0,99$

Lösung 8-42: *Absolventen der Fakultät*
$1-\alpha/2 = 0,975;$ $z_{0,975} = 1,96;$ $e = 0,2;$ $n \geq z^2_{1-\alpha/2}/4e^2;$ $n \geq 1,96^2/4\cdot 0,2^2 = 3,8416/0,16 = 24,01 \rightarrow n \geq 25$

Lösung 8-43: *Sportliche Betätigung*
$1-\alpha/2 = 0,99506$ $z_{0,99506} = 2,58$ $\hat{\pi} = 180/200 = 0,9$
Da n sehr groß, Schätzfunktion $\hat{\pi} = X/n$ approx. normalverteilt.

$$\left[\hat{\pi} \pm z_{1-\alpha/2} \cdot \sqrt{\frac{\hat{\pi}(1-\hat{\pi})}{n}}\right]$$

$$= 0,9 \pm 2,58\sqrt{\frac{0,9\cdot 0,1}{200}} = 0,9 \pm 2,58 \cdot 0,0212132 = 0,9 \pm 0,05473$$

$[0,845; 0,955]$

Lösung 8-44: *Schwankungsintervall*
Zentrales Schwankungsintervall:

$$\left[\mu - c\cdot\frac{\sigma}{\sqrt{n}}; \mu + c\cdot\frac{\sigma}{\sqrt{n}}\right]$$

Sicherheitswahrscheinlichkeit:

$$P\left(\mu - c\cdot\frac{\sigma}{\sqrt{n}} < \overline{X} < \mu + c\cdot\frac{\sigma}{\sqrt{n}}\right) = 1-\alpha$$

$1-\alpha/2 = 0,975;$ $c = 1,96$
$\mu - c\cdot\frac{\sigma}{\sqrt{n}} = 354 - 1,96\cdot\frac{22,5}{\sqrt{81}} = 354 - 1,96\cdot 2,5 = 354 - 4,9 = 349,1$
$\mu + c\cdot\frac{\sigma}{\sqrt{n}} = 354 + 1,96\cdot\frac{22,5}{\sqrt{81}} = 354 + 1,96\cdot 2,5 = 354 + 4,9 = 358,9$

Lösung 8-45: *Kilometerleistung*

A) $n = 49$, $\bar{x} = 50$ km, $\sigma = 7$ km bekannt

(a) Allgemeines Konfidenzintervall für den Mittelwert μ der Grundgesamtheit

$$P\left(\bar{X} - c\frac{\sigma}{\sqrt{n}} \leq \mu \leq \bar{X} + c\frac{\sigma}{\sqrt{n}}\right) = 1-\alpha$$

c aus $N(0;1)$, da σ bekannt $\Longrightarrow c = z_{1-\alpha/2}$

(b) Schätzintervall für den Mittelwert μ der Grundgesamtheit

$$\begin{aligned} 1-\alpha = 95\% &\Rightarrow 1-\alpha/2 = 97,5\% \\ &\Rightarrow \Phi(z_{1-\alpha/2}) = 0,975 \\ &\Rightarrow z_{1-\alpha/2} = 1,96 \end{aligned}$$

$$\left[\bar{x} - z_{1-\alpha/2}\frac{\sigma}{\sqrt{n}}; \bar{x} + z_{1-\alpha/2}\frac{\sigma}{\sqrt{n}}\right]$$

$$\left[50 - 1,96\frac{7}{\sqrt{49}}; 50 + 1,96\frac{7}{\sqrt{49}}\right] = [48,04; 51,96]$$

(c) Schätzintervall für den Mittelwert μ der Grundgesamtheit, Breite fix, n variabel

$$\left[\bar{x} - z_{1-\alpha/2}\frac{\sigma}{\sqrt{n}}; \bar{x} + z_{1-\alpha/2}\frac{\sigma}{\sqrt{n}}\right] = [v_u; v_o]$$

Breite:

$$\begin{aligned} v_o - v_u &= \left(50 + 1,96\frac{7}{\sqrt{n}}\right) - \left(50 - 1,96\frac{7}{\sqrt{n}}\right) \\ 2 \text{ km} &= \frac{2 \cdot 1,96 \cdot 7 \text{ km}}{\sqrt{n}} \\ \sqrt{n} &= 13,72 \Rightarrow n = 188,23 \text{ also } n \geq 189 \end{aligned}$$

B) Y : "Anzahl der ADAC Mitglieder" $\sim B(200;\pi)$

Approximationsbedingung: $\hat{\pi} = \frac{40}{200} = 0,2 \Rightarrow n\pi(1-\pi) \approx 200 \cdot 0,2 \cdot 0,8 = 32 > 9$

$$\begin{aligned} Y &\approx N\left(\mu = n\pi; \sigma = \sqrt{n\pi(1-\pi)}\right) \\ \hat{\pi} = \frac{Y}{n} &\approx N\left(\mu = \pi; \sigma = \sqrt{\frac{\pi(1-\pi)}{n}}\right) \end{aligned}$$

Konfidenzintervall

$$P\left(\frac{Y}{n} - c\sqrt{\frac{\frac{Y}{n}\left(1-\frac{Y}{n}\right)}{n}} \leq \pi \leq \frac{Y}{n} + c\sqrt{\frac{\frac{Y}{n}\left(1-\frac{Y}{n}\right)}{n}}\right) = 1 - \alpha$$

c aus $N(0;1)$, da σ bekannt $\Longrightarrow c = z_{1-\alpha/2}$

Schätzintervall

$$\begin{aligned} 1 - \alpha = 99\% &\Rightarrow 1 - \alpha/2 = 99,5\% \\ &\Rightarrow \Phi(z_{1-\alpha/2}) = 0,995 \\ &\Rightarrow z_{1-\alpha/2} = 2,58 \end{aligned}$$

$$\left[0,2 - 2,58\sqrt{\frac{0,2\cdot(1-0,2)}{200}}; 0,2 + 2,58\sqrt{\frac{0,2\cdot(1-0,2)}{200}}\right]$$
$$= [0,12703; 0,27297]$$

C) $X_i \sim N(\mu;\sigma) \Longrightarrow \bar{X} \sim N(\mu_{\bar{X}} = \mu; \sigma_{\bar{X}} = \sigma/\sqrt{n})$

$$\begin{aligned} n &= 5 \\ \bar{x} &= \frac{0,18 + 0,25 + 0,12 + 0,20 + 0,25}{5} = 0,2 \\ s^2 &= \frac{0,0004 + 0,0025 + 0,0064 + 0 + 0,0025}{4} = 0,00295 \end{aligned}$$

(a) Allgemeines Konfidenzintervall für den Mittelwert μ der Grundgesamtheit

$$P\left(\bar{X} - c\frac{S}{\sqrt{n}} \leq \mu \leq \bar{X} + c\frac{S}{\sqrt{n}}\right) = 1 - \alpha$$

c aus t_{n-1}, da σ unbekannt $\Longrightarrow c = t_{n-1;1-\alpha/2}$

(b) Schätzintervall

$$\begin{aligned} 1 - \alpha = 95\% &\Rightarrow 1 - \alpha/2 = 97,5\% \\ &\Rightarrow t_{n-1;1-\alpha/2} = 2,776 \end{aligned}$$

$$\left[0,2 - 2,776\sqrt{\frac{0,00295}{5}}; 0,2 + 2,776\sqrt{\frac{0,00295}{5}}\right] = [0,1326; 0,2674]$$

9 Statistische Testverfahren

Lösung 9-1: *Spezialgefrierschränke*

a) $H_0 : \mu \geq \mu_0 = -25^o$C, $H_1 : \mu < \mu_0 = -25^o$C
$P(''H_1''|H_0) = P($ "Kunden zufrieden?" | Ruin $) = \alpha$

b) $\overline{X}$: "Durchschnittliche Temperatur eines Spezialgefrierschrankes bei einer Zufallsstichprobe $n = 100$" X_i: "Temperatur des i–ten Spezialgefrierschrankes"; $i = 1, \ldots, 100$ $X_i \sim N(\mu; 2)$; $\overline{X}$ ist unter H_0 $N(-25; 0,2)$–verteilt

c) $V = (\overline{X} - \mu_0)/(\sigma/\sqrt{n})$ ist unter H_0 $N(0; 1)$-verteilt

d) Ablehnungsbereich: $\{v|v < -2\}$, Nicht–Ablehnungsbereich: $\{v|v \geq -2\}$

e) (i) $v = -5 \in$ Ablehnungsbereich $\Rightarrow ''H_1''$

(ii) Auf einem Signifikanzniveau von 2,275% und basierend auf einer Zufallsstichprobe vom Umfang n = 100 konnte statistisch bewiesen werden, dass die durchschnittliche Temperatur der Geräte unter -25^{o}C liegt. Somit keine Produktionsveränderung notwendig.

f) (i) $v = -1,5 \notin$ Ablehnungsbereich $\Rightarrow ''H_0''$

(ii) Auf einem Signifikanzniveau von 2,275% und basierend auf einer Zufallsstichprobe vom Umfang $n = 100$ konnte statistisch nicht bewiesen werden, dass die durchschnittliche Temperatur der Geräte unter -25^{o}C liegt. Somit Produktionsveränderung notwendig.

(iii) Fehler 2. Art

(iv) Frage kann nicht beantwortet werden; Fehler ist unterlaufen oder nicht.

(v) $P(''H_0''|H_1 : \mu = -29) = 0$

g) $P(''H_1''|H_0) = \alpha$ ist an der Nahtstelle der Hypothesen stets am größten

Lösung 9-2: *Spezialgefrierschränke (Gütefunktion)*

a) (i) $G(\mu_1 = -24,8) = 0,00135$

(ii) $G(\mu_2 = -25,8) = 0,97725$

(iii) $G(\mu_3 = -29) = 1$

b) Die Skizze ist in den Lösungen nicht enthalten.

Lösung 9-3: *Durchmesser von Wellen*

a) Ablehnungsbereich: $\{v|v < -1,96$ oder $v > 1,96\}$

b) $v = 0,8 \in$ Nicht–Ablehnungsbereich $\Rightarrow ''H_0''$

c) Fehler 2. Art

d) $v = 4 \in$ Ablehnungsbereich $\Rightarrow ''H_1''$

e) Fehler 1. Art

Lösung 9-4: *Phosphatgehalt der Waschmittel*
X_i: "Phosphatgehalt des i-ten Paketes"; $i = 1, ..., 36$
X_i ist beliebig verteilt mit $E(X) = \mu$; $Var(X_i) = 36\text{g}^2$

a) $H_0 : \mu \leq \mu_0 = 18$, $H_1 : \mu > \mu_0 = 18$
$''H_1''|H_0$ = "Phosphatgehalt zu hoch" | Phosphatgehalt stimmt; dies ist aus Sicht des Fabrikanten die schlimmere Fehlentscheidung.

b) $\overline{X}$: "Durchschnittlicher Phosphatgehalt eines Paketes bei einer Zufallsstichprobe vom Umfang $n = 36$"
$\overline{X}$ ist unter H_0 approximativ $N(\mu_0; \sigma/\sqrt{n}) = N(18; 1)$–verteilt wegen Zentralem Grenzwertsatz, $n > 30$

c) $V = (\overline{X} - \mu_0)/(\sigma/\sqrt{n})$ ist unter H_0 approximativ $N(0; 1)$–verteilt

d) Ablehnungsbereich: $\{v|v > 3,09\}$, Nicht–Ablehnungsbereich: $\{v|v \leq 3,09\}$

e) $v = 2 \in$ Nicht–Ablehnungsbereich $\Rightarrow ''H_0''$

f) Es konnte statistisch nicht bewiesen werden, dass der Richtwert überschritten wird. Die Firma spricht aber von einem statistischen Beweis, dass der Richtwert eingehalten wird (der H_0!). α ist sehr klein! Kommt bei dieser Hypothesenformulierung nur der Firma zugute, d.h. nur bei einem ganz extrem großen Stichprobenwert von $\overline{X}$ muss die Firma das Produkt vom Markt nehmen ($''H_1''$).

g) Wenn der wahre Wert des mittleren Phosphatgehalts 21,09g ist, würden 50% der Stichproben einen Mittelwert unter 21,09g und der Rest einen Mittelwert über 21,09g ergeben. Bei $\overline{X} = 21,09$ nimmt der Prüfwert den Wert $\frac{21,09-18}{\sqrt{36}/\sqrt{36}} = 3,09$ an, was genau der Grenze des Ablehnungsbereiches entspricht. Im Fall von 50% der möglichen Stichproben bekommt man also einen Prüfwert, der nicht zum Ablehnungsbereich gehört.

$\Rightarrow P(''H_0''|\mu = 21,09) = 0,5$

Testtheorie-Phosphatgehalt der Waschmittel (12 min).mp4

Lösung 9-5: *Phosphatgehalt der Waschmittel (Gütefunktion)*
Der Verlauf der Gütefunktion ist *nicht abhängig* vom Stichprobenergebnis, aber *abhängig* vom Stichprobenumfang.

Lösung 9-6: *Durchschnittsgewicht*
X_i: "Gewicht des i-ten Hähnchens"; $i = 1, ..., 25$; $X_i \sim N(\mu; \sigma)$

a) $H_0 : \mu \geq \mu_0 = 1400$, $H_1 : \mu < \mu_0 = 1400$
$''H_1''|H_0$ = "Angebot zurückweisen" | gutes Geschäft vermasselt

b) $\overline{X}$: "Durchschnittliches Gewicht eines Hähnchens bei einer Zufallsstichprobe vom Umfang $n = 25$"

c) $\overline{X}$ ist unter H_0 $N(1400; \sigma/\sqrt{n})$–verteilt

d) $V = (\overline{X} - \mu_0)/(S/\sqrt{n})$ ist unter H_0 t–verteilt mit $f = 24$ Freiheitsgraden

e) Ablehnungsbereich: $\{v|v < -1,711\}$, Nicht–Ablehnungsbereich: $\{v|v \geq -1,711\}$

f) $v = -0,9 \notin$ Ablehnungsbereich $\Rightarrow ''H_0''$

g) Fehler 2. Art

h) $v = -1,9 \in$ Ablehnungsbereich $\Rightarrow ''H_1''$

i) Fehler 1. Art

Lösung 9-7: *Sollwerte*

a) $H_0 : \mu = \mu_0(= 300)$, $H_1 : \mu \neq \mu_0(= 300)$

X_i: "Füllgewicht der i-ten Konserve"; $i = 1, ..., 100$; $X_i \sim N(\mu; \sigma)$

$\overline{X}$: "Durchschnittliches Füllgewicht einer Konserve bei einer Zufallsstichprobe vom Umfang $n = 100$"

$\overline{X}$ ist unter H_0 $N(300; \sigma/\sqrt{n})$–verteilt σ unbekannt, aber $n > 30 \Rightarrow$ Verwendung der Normalverteilung $V = (\overline{X} - \mu_0)/(S/\sqrt{n})$ ist unter H_0 approximativ $N(0; 1)$–verteilt

Ablehnungsbereich: $\{v|v < -1,96 oder v > 1,96\}$, Nicht–Ablehnungsbereich: $\{v| -1,96 \leq v \leq 1,96\}$
$v = 2 \in$ Ablehnungsbereich $\Rightarrow ''H_1''$; Produktionsprozeß stoppen.

b) $H_0 : \mu \geq \mu_0 = 300$, $H_1 : \mu < \mu_0 = 300$ (das will Abnehmer beweisen!)

Lösung 9-8: *Zigarettenpreis*
X_i: "Zigarettenkonsum des i–ten Rauchers pro Tag"; $i = 1, ..., 100$;
X_i ist beliebig verteilt mit $E(X_i) = \mu$ und $Var(X_i) = \sigma^2$

a) $H_0 : \mu \geq \mu_0 = 16$, $H_1 : \mu < \mu_0 = 16 \Rightarrow$ das will der Prokurist beweisen

b) $\overline{X}$: "Durchschnittlicher Zigarettenkonsum eines Rauchers pro Tag bei einer Zufallsstichprobe vom Umfang $n = 100$"

c) $\overline{X}$ ist unter H_0 approximativ $N(16; \sigma/\sqrt{n})$–verteilt wegen Zentralem Grenzwertsatz und $n > 30$

d) $V = (\overline{X} - \mu_0)/(S/\sqrt{n})$ ist unter H_0 approximativ $N(0; 1)$

e) Ablehnungsbereich: $\{v|v < -2,33\}$, Nicht–Ablehnungsbereich: $\{v|v \geq -2,33\}$,
$v = -2 \notin$ Ablehnungsbereich $\Rightarrow ''H_0''$

f) Fehler 2. Art

g) Auf einem Signifikanzniveau von $\alpha = 0,01$ und basierend auf einer Zufallsstichprobe vom Umfang $n = 100$ konnte statistisch nicht gezeigt werden, dass sich der durchschnittliche Zigarettenkonsum verringert hat.

Testtheorie-Zigarettenpreis.mp4

Lösung 9-9: *Schwergewichtsboxer*

a) $H_0 : \pi \leq \pi_0 = 0,5,\ H_1 : \pi > \pi_0 = 0,5 \Rightarrow$ das will er beweisen

b) X: "Anzahl der von J.Knockout gewonnenen Kämpfe bei einer Zufallsstichprobe vom Umfang $n = 11$"

c) X ist unter H_0 $B(11; 0,5)$–verteilt

d) Ablehnungsbereich: $\{x > 8\}$, Nicht–Ablehnungsbereich: $\{x \leq 8\}$

e) $x = 8 \notin$ Ablehnungsbereich $\Rightarrow {''}H_0''$

f) Fehler 2. Art

g) Auf einem Signifikanzniveau von $\alpha_{ex.} = 0,0327$ und basierend auf einer Zufallsstichprobe vom Umfang $n = 11$ Kämpfen konnte statistisch nicht gezeigt werden, dass J. Knockout der bessere Boxer ist.

Lösung 9-10: *Skirennen*

a) $H_0 : \pi \geq \pi_0 = 0,1,\ H_1 : \pi < \pi_0 = 0,1$
${''}H_1''|H_0$ = "Hang bleibt wie gesteckt" | Krankenhaus überfüllt

b) X: "Anzahl der Gäste, die ausscheiden, bei einer Zufallsstichprobe vom Umfang $n = 22$"
X ist unter H_0 $B(22; 0,1)$–verteilt

c) Ablehnungsbereich: $\{x < 1\}$, Nicht–Ablehnungsbereich: $\{x \geq 1\}$

d) $\alpha_{ex.} = 0,0985$

e) $x = 1 \notin$ Ablehnungsbereich $\Rightarrow {''}H_0''$

f) Auf einem Signifikanzniveau von $\alpha_{ex.} = 0,0985$ und basierend auf einer Zufallsstichprobe vom Umfang $n = 22$ konnte statistisch nicht bewiesen werden, dass die Ausfallquote kleiner als 10% ist.

Testtheorie-Skirennen (11 min).mp4

Lösung 9-11: *Skirennen (Gütefunktion)*

a) $G(\pi = 0) = 1$; $G(\pi = 0,1) = 0,0985$; $G(\pi = 0,2) = 0,0074$

b) Die Skizze ist in den Lösungen nicht enthalten.

Lösung 9-12: *Chininhaltige Limonade*

a) $H_0 : \pi \geq \pi_0 = 0,1,\ H_1 : \pi < \pi_0 = 0,1$
${''}H_1''|H_0$ = "Es wird importiert" | Kunden werden krank

b) X: "Anzahl der Flaschen, die den Vorschriften nicht entsprechen, bei einer Zufallsstichprobe vom Umfang $n = 30$"

c) X ist unter H_0 $B(30; 0,1)$–verteilt

d) Ablehnungsbereich: $\{x < 1\}$, Nicht–Ablehnungsbereich: $\{x \geq 1\}$

e) $x = 1 \notin$ Ablehnungsbereich $\Rightarrow {''}H_0''$

f) Auf einem Signifikanzniveau von $\alpha_{ex.} = 0,0424$ und basierend auf einer Zufallsstichprobe vom Umfang $n = 30$ konnte statistisch nicht bewiesen werden, dass der Anteil der Flaschen, die den Vorschriften nicht entsprechen, kleiner als 10% ist, d.h. der Großhändler sucht sich einen neuen Importeur.

g) $G(\pi = 0) = 1$; $G(\pi = 0,1) = 0,0424$; $G(\pi = 0,2) = 0,0012$

Lösung 9-13: *Schlampiges Gepäck-Handling*

i	x_i	h_i	p_i	np_i	$h_i - np_i$	$(h_i - np_i)^2$	$(h_i - np_i)^2/(np_i)$
1	0	460	0,449	449	11	121	0,269
2	1	350	0,360	360	-10	100	0,278
3	2	135	0,144	144	-9	81	0,563
4	3	40	0,038	38	2	4	0,105
5	4	15	0,008	8	7	49	5,125
6	>4	0	0,001	1	-1	1	1

a) H_0: Stichprobenverteilung des Gepäckverlustes entspricht einer Poisson-Verteilung,
H_1: Stichprobenverteilung des Gepäckverlustes entspricht nicht einer Poisson-Verteilung

b)
$$V = \sum_{i=1}^{I} \frac{(h_i - n \cdot p_i)^2}{n \cdot p_i}$$

V ist unter H_0 approximativ ($np_i \geq 1$ für alle i, $np_i \geq 5$ für 80% der i) χ^2–verteilt mit $f = I - 1 - k = 4$ Freiheitsgraden

c) Ablehnungsbereich: $\{v|v > 13,28\}$, Nicht–Ablehnungsbereich: $\{v|v \leq 13,28\}$

d)
$$L(x_1, ..., x_n|\lambda) = \frac{\lambda^{x_1+...+x_n}}{x_1! \cdot ... \cdot x_n!} e^{-n\lambda} \to \max \quad \Rightarrow \quad \widehat{\lambda} = 0,8$$

e) siehe obige Tabelle $v = 7,34 \notin$ Ablehnungsbereich $\Rightarrow {}''H_0''$

f) H_0 läßt sich statistisch nicht beweisen! Auf einem Signifikanzniveau von $\alpha = 0,01$ und basierend auf einer Zufallsstichprobe vom Umfang $n = 1000$ konnte lediglich statistisch <u>nicht</u> bewiesen werden, dass es sich <u>nicht</u> um eine Poisson-Verteilung handelt.

Lösung 9-14: *Münzen*
H_0: Stichprobenverteilung stimmt mit der vermuteten Verteilung überein
H_1: Stichprobenverteilung stimmt nicht mit der vermuteten Verteilung überein
8 mögliche Ereignisse: ZZZ; KZZ; ZKZ; ZZK; KKZ; KZK; ZKK; KKK

i	x_i	h_i	p_i	np_i	$h_i - np_i$	$(h_i - np_i)^2$	$(h_i - np_i)^2/(np_i)$
1	0	24	1/8	30	-6	36	1,2
2	1	108	3/8	90	18	324	3,6
3	2	85	3/8	90	-5	25	0,277
4	3	23	1/8	30	-7	49	1,633

Ablehnungsbereich: $\{v|v > 7,81\}$, Nicht–Ablehnungsbereich: $\{v|v \leq 7,81\}$
$v = 6,71 \notin$ Ablehnungsbereich $\Rightarrow {}''H_0''$

Lösung 9-15: *Torerfolge*
X: "Torerfolge pro Spiel"

i	x_i	h_i	p_i	np_i	$h_i - np_i$	$(h_i - np_i)^2$	$(h_i - np_i)^2/(np_i)$
1	0	18	0,0334	10	8	64	6,40
2	1	24	0,1134	34	-10	100	2,94
3	2	56	0,1929	58	-2	4	0,07
4	3	63	0,2187	66	-3	9	0,14
5	4	61	0,1858	56	5	25	0,45
6	5	39	0,1263	38	1	1	0,03
7	6	26	0,0716	21	5	25	1,19
8	7	6	0,0348	10	-4	16	1,60
9	8	5	0,0148	4	1	1	0,25
10	9	(2)2	0,0056	(3)2	-1	1	0,33
11	>9	0	0,0027	1			

Werte in Klammern, wenn alle Werte mit $x_i \geq 9$ in einer Klasse.

a) H_0: Stichprobenverteilung entspricht einer $PO(3,4)$
H_1: Stichprobenverteilung entspricht nicht einer $PO(3,4)$

b)

$$V=\sum_{i=1}^{I}\frac{(h_i-n\cdot p_i)^2}{n\cdot p_i}$$

V ist unter H_0 approximativ ($np_i \geq 1$ für alle i, $np_i \geq 5$ für mindestens 80% der i) χ^2–verteilt mit $f = I-1-k = 10-1-0 = 9$ Freiheitsgraden

c) Ablehnungsbereich: $\{v|v > 14,68\}$, Nicht–Ablehnungsbereich: $\{v|v \leq 14,68\}$

d) $v = 13,40 \notin$ Ablehnungsbereich $\Rightarrow$ $''H_0''$

e) Auf einem Signifikanzniveau von $\alpha = 0,1$ und basierend auf einer Zufallsstichprobe vom Umfang $n = 300$ konnte statistisch nicht bewiesen werden, dass die Stichprobenverteilung der Torfolge nicht einer $PO(3,4)$ entspricht.

Lösung 9-16: *Wetterlage und Geschäftslage*
X: “Wetterlage”; Y: “Geschäftslage”

a)

$X \setminus Y$	y_1=gut	y_2=normal	y_3=schlecht	
x_1=Regentag	5	10	5	20
x_2=Sonnentag	15	5	10	30
	20	15	15	50

b) H_0: Wetter und Geschäftslage sind stochastisch unabhängig
H_1: Wetter und Geschäftslage sind nicht stochastisch unabhängig

c) ja, da alle $\widetilde{h}_{ij} \geq 5$

$$V=\sum_{i=1}^{I}\sum_{j=1}^{J}\frac{(h_{ij}-\widetilde{h}_{ij})^2}{\widetilde{h}_{ij}}$$

ist unter H_0 approximativ χ^2–verteilt mit $f = 2$ Freiheitsgraden.

d) Tabelle mit $\widetilde{h}_{ij}$

$X \setminus Y$	y_1=gut	y_2=normal	y_3=schlecht	
x_1=Regentag	8	6	6	20
x_2=Sonnentag	12	9	9	30
	20	15	15	50

(i) Ablehnungsbereich: $\{v|v > 9,21\}$, Nicht–Ablehnungsbereich: $\{v|v \leq 9,21\}$
$v = 6,597 \notin$ Ablehnungsbereich $\Rightarrow$ $''H_0''$

(ii) Ablehnungsbereich: $\{v|v > 5,99\}$, Nicht–Ablehnungsbereich: $\{v|v \leq 5,99\}$
$v = 6,597 \in$ Ablehnungsbereich $\Rightarrow$ $''H_1''$

e) (i) Fehler 2. Art, (ii) Fehler 1. Art

Testtheorie-Wetterlage und Geschäftslage (14 min).mp4

Lösung 9-17: *Mietpreisbindung*

a) χ^2–Anpassungstest

b) X: “Mietpreissteigerung [in %]”
H_0: Stichprobenverteilung folgt einer Gleichverteilung in $[a, b]$
H_1: Stichprobenverteilung folgt nicht einer Gleichverteilung in $[a, b]$
$b = 5$ [%]; $(a + b)/2 = 2,5$ [%] $\Rightarrow$ $a = 0$ [%]

c)

$$V=\sum_{i=1}^{I}\frac{(h_i-n\cdot p_i)^2}{n\cdot p_i}$$

d) V ist unter H_0 approximativ ($np_i \geq 5$ für alle i) χ^2–verteilt mit $f = 4$ Freiheitsgraden

e) Ablehnungsbereich: $\{v|v > 14,86\}$, Nicht–Ablehnungsbereich: $\{v|v \leq 14,86\}$

f)

i	x_i	h_i	p_i	np_i	h_i-np_i	$(h_i-np_i)^2$	$(h_i-np_i)^2/(np_i)$
1	0-1	0	0,2	20	-20	400	20
2	1-2	0	0,2	20	-20	400	20
3	2-3	10	0,2	20	-10	100	5
4	3-4	10	0,2	20	-10	100	5
5	4-5	40	0,2	20 20	60	3600	180
6	5-	40	0	0			

$v = 230 \in$ Ablehnungsbereich $\Rightarrow "H_1''$

g) Auf einem Signifikanzniveau von 0,5% und basierend auf einer Zufallsstichprobe vom Umfang $n = 100$ konnte statistisch bewiesen werden, dass die Stichprobenverteilung keiner Gleichverteilung im Bereich $[0; 5]$ folgt.

Lösung 9-18: *Wocheneinkommen*
X:„Wocheneinkommen in diesem Stadtteil“, Verteilung unbekannt, $\sigma = 20$ EUR;
$\overline{X}$:„Durchschnittliches Wocheneinkommen in diesem Stadtteil“, $\overline{X}$ ist approximativ (Zentraler Grenzwertsatz, $n = 100 > 30$) $N \sim (\mu; \sigma/\sqrt{n})$ mit $\sigma/\sqrt{n} = 20/10 = 2$
$\mu_0 = 400$, $\alpha = 0,050503$, $z_{0,949497} = 1,64$, $\mu_1 = 406$, $H_0 : \mu \leq 400$ $\quad H_1 : \mu > 400$

$$G(\mu_1) = 1 - P\Big(V \leq c - \frac{\mu_1 - \mu_0}{\sigma/\sqrt{n}}\Big)$$

$G(\mu_1 = 406) = 1 - P(V \leq 1,64 - (406 - 400)/2) = 1 - P(V \leq -1,36) = 1 - (1 - P(V \leq 1,36)) = P(V \leq 1,36) = 0,913085$, $\beta = 1 - G(\mu_1) = 1 - 0,913085 = 0,086915 \approx 0,087$

Lösung 9-19: *Testfunktion*
Für den Ablehnungsbereich $\{v|v > c\}$ gilt $P(V > c) = \alpha$.
Für jedes $v \leq c$ ist $P(V > v) > P(V > c)$, d.h. das vorgegebene Signifikanzniveau wird nicht eingehalten.
Oder: $P(V > c) = \alpha$; $P(V > v) = \gamma$,
$P(V > v|v \leq c) = [P(V \leq c) - P(V \leq v)] + P(V > c)$
$\gamma = \delta + \alpha$
$\gamma = \alpha$ für $v = c$, $\delta = 0$; $\gamma > \alpha$ für $v < c$, $\delta > 0$.
$P(V > v|v > c) = P(V > c) - [P(V \leq v) - P(V \leq c)]$
$\gamma = \alpha - \delta$
$\gamma < \alpha$ für $v > c$, $\delta > 0$.

Lösung 9-20: *Zugkraft eines Drahtseiles*
$n = 49 > 30; \overline{X}$ approximativ normalverteilt
$\mu_0 = 15$; $\mu = 14,8$; $\sigma = 0,4964$; $\alpha = 0,07927$; $c_{0,92073} = 1,41$; $\beta = 1 - G(\mu)$

$$\begin{aligned} G(\mu) &= P\Big(V \leq -c - \frac{\mu - \mu_0}{\sigma}\sqrt{n}\Big) \\ G(14,8) &= P\Big(V \leq -1,41 - \frac{14,8 - 15}{0,4964}\sqrt{49}\Big) \\ &= P(V \leq -1,41 + 2,82) \\ &= P(V \leq 1,41) \\ &= 0,92073 \\ \beta &= 1 - 0,92073 = 0,07927 \end{aligned}$$

Lösung 9-21: *Paketversandfirma*

$$V = \frac{\hat{\pi} - \pi_0}{\sqrt{\dfrac{\pi_0(1 - \pi_0)}{n}}}$$

V ist unter H_0 approximativ $[n\pi_0 > 9; n(1 - \pi_0) > 9; n > 30]$ $N(0; 1)$
$\alpha = 0,0359$; $1 - \alpha = 0,9641$; $c = 1,8$
Ablehnungsbereich der $H_0 : \{v|v > 1,8\}$
$n = 900$; $p = 828/900 = 0,92$; $v = (0,92 - 0,9)/0,01 = 2$
$v = 2 \in$ Ablehnungsbereich $\rightarrow$ „H_1“
Auf einem Signifikanzniveau von $\alpha = 0,0359$ und basierend auf einer Stichprobe vom Umfang $n = 900$ konnte statistisch gezeigt werden, dass mehr als 90% der Pakete den Empfänger innerhalb einer Woche erreichen. Das Unternehmen beauftragt die Versandfirma mit dem Versand ihrer Pakete.

Lösung 9-22: *Dicke der Fahrbahndecke*
$H_0 : \mu \geq \mu_0 = 3,5$ $\qquad H_1 : \mu < \mu_0 = 3,5$
Der Bauunternehmer muss nachweisen, dass die Fahrbahndecke zu dünn ist, da er nur dann Abzüge hinnehmen muss.
Risikobetrachtung:
„H_1“$|H_0$ = „Fahrbahndecke zu dünn, muss Abzüge hinnehmen“|Fahrbahndecke o.k., müsste keine Abzüge hinnehmen
Dies ist für den Bauunternehmer das größere Risiko, das gleich dem Fehler 1. Art ist, für den die Wahrscheinlichkeit mit α vorgegeben ist.

Lösung 9-23: *Neues Präparat*

a) $H_0 : \pi \leq \pi_0\ (= 0,35)$ $\qquad H_1 : \pi > \pi_0\ (= 0,35)$
„H_1“$|H_0$ =„Einführung des Präparates“|Hersteller lügt; Krankenkassen zahlen, obwohl Heilungsquote minimal

b) $V = X$: „Anzahl der Patienten, bei denen Heilerfolg eintritt, bei einer Zufallsstichprobe vom Umfang $n = 19$“

$$V = \sum_{i=1}^{n} X_i$$

X_i = „Heilerfolg beim i-ten Patienten“

c) V ist unter H_0 $B.V.(n; \pi_0) \sim B.V.(19; 0,35)$

d) Nicht-Ablehnungsbereich: $\{v|v \leq 12\}$; Ablehnungsbereich: $\{v|v > 12\}$
$\alpha_{exakt} = 0,0031$

e)
1. $P(„H_0“|\pi_0 = 0,5 \in H_1) = \beta_{(\pi_0=0,5)} = 0,9165$
Die Wahrscheinlichkeit für einen Fehler 2. Art (unberechtigte Annahme von H_0) beträgt 91,65%, wenn die wahre Heilungsquote 50% beträgt.
2. 2. $P(„H_1“|\pi_0 = 0,4 \in H_1) = 1 - \beta_{(\pi_0=0,4)} = 1 - 0,9884 = 0,0116$
Die Wahrscheinlichkeit für eine berechtigte Annahme der H_1 beträgt 1,16%, wenn die wahre Heilungsquote 40% beträgt.

Lösung 9-24: *Batterien Lebensdauer*

a) χ^2–Anpassungstest

b) H_0: „Die Stichprobenverteilung der Lebensdauer der Batterien ist normalverteilt“
H_1: „Die Stichprobenverteilung der Lebensdauer der Batterien ist nicht normalverteilt“

c) X: „Lebensdauer einer Batterie“

$$V = \sum_{i=1}^{I} \frac{(h_i - np_i)^2}{np_i}$$

ist unter H_0 χ^2–verteilt mit $f = I - 1 - k$ Freiheitsgraden, wenn für alle i $np_i \geq 5$ gilt (I – Anzahl der Klassen, k – Anzahl der zu schätzenden Parameter)

i	Klassen	h_i	$\bar{x}_i$	$h_i\bar{x}_i$	p_i	np_i
1	– 300	10	160	1600	0,16	16
2	300 – 340	10	320	3200	0,12	12
3	340 – 460	60	400	24000	0,45	45
4	460 –	20	560	11200	0,27	27
		100		40000		

i	Klassen	$h_i - np_i$	$(h_i - np_i)^2$	$\frac{(h_i-np_i)^2}{np_i}$
1	– 300	-6	36	2,25
2	300 – 340	-2	4	0,33
3	340 – 460	15	225	5,00
4	460 –	-7	49	1,82
				$v = 9,40$

$$\bar{x} = \frac{1}{n}\sum_i x_i h_j = \frac{1}{100} \cdot 40000 = 400$$

$$s = 100$$

$$p_1 = P(V \le 300) = P\left(Z \le \frac{300-400}{100}\right)$$
$$= P(Z \le -1) = 1 - P(Z \le 1) = 1 - 0,841345 \approx 0,16$$

$$p_2 = P(300 \le V \le 340) = P\left(\frac{300-400}{100} \le Z \le \frac{340-400}{100}\right) =$$
$$= P(-1 \le Z \le -0,6) = P(Z \le 1) - P(Z \le 0,6)$$
$$= 0,841345 - 0,725747 \approx 0,12$$

$$p_3 = P(340 \le V \le 460) = P\left(\frac{340-400}{100} \le Z \le \frac{460-400}{100}\right)$$
$$= P(-0,6 \le Z \le 0,6) = 2 \cdot P(Z \le 0,6) - 1$$
$$= 2 \cdot 0,725747 - 1 \approx 0,45$$

$$p_4 = P(V \ge 460) = P\left(Z \ge \frac{460-400}{100}\right)$$
$$= 1 - P(Z \le 0,6) = 1 - 0,725747 \approx 0,27$$

Approximationsbedingung erfüllt; $f = 4 - 1 - 2 = 1$; $\alpha = 0,01$
Nicht–Ablehnungsbereich: $\{v | v \le 6,63\}$
Ablehnungsbereich: $\{v | v > 6,63\}$

d) $v = 9,4 \in$ Ablehnungsbereich $\to$ „H_1“
Auf einem Signifikanzniveau von 1% und basierend auf einer Zufallsstichprobe vom Umfang $n = 100$ konnte statistisch bewiesen werden, dass es sich bei der Stichprobenverteilung der Lebensdauer der Batterien nicht um eine Normalverteilung handelt.

e) Weiß man nicht; wir hoffen nicht!

Lösung 9-25: *Kaffee Packungen*

a) $H_0 : \mu \le \mu_0 = 500$ g $\qquad H_1 : \mu > \mu_0 = 500$ g
„H_1“$|H_0$ = „Abfüllmenge o.k.“|ärger mit dem Kunden
$P(„H_1“|H_0) = \alpha = 0,02275 \to$ klein halten

b) $\overline{X}$: „Durchschnittliche Füllmenge einer Kaffeepackung in einer einfachen Zufallsstichprobe vom Umfang $n = 25$“
$\overline{X} = \frac{1}{n}\sum_{i=1}^{n} X_i$
X_i: „Füllmenge der i–ten Kaffeepackung“; $i = 1, \ldots, 25$
$X_i \sim N(\mu; 10)$ für alle i, unabhängig
$\overline{X}$ ist unter H_0 $N(\mu_0; \sigma/\sqrt{n}) = N(500; 2)$–verteilt.

c) $V = \dfrac{\overline{X} - \mu_0}{\frac{\sigma}{\sqrt{n}}} = \dfrac{\overline{X} - 500}{2}$ ist unter H_0 $N(0;1)$–verteilt.

d) c für $1 - \alpha = 0,97725$ aus Tabelle der $N(0;1) \to c = 2$
Ablehnungsbereich: $\{v | v > 2\}$
Nicht–Ablehnungsbereich: $\{v | v \le 2\}$

e) $v = (504,5 - 500)/2 = 2,25 \in$ Ablehnungsbereiches $\to$ „H_1“

f) Auf einem Signifikanzniveau von $\alpha = 0,02275$ und basierend auf einem Stichprobenumfang von $n = 25$ konnte statistisch bewiesen werden, dass die wahre durchschnittliche Füllmenge einer Packung bei der neuen Kaffeebohnensorte der Norm entspricht.

g)

$$\begin{aligned}
\beta &= 1 - G(\mu = 501) = 1 - P(\overline{X} > \overline{x}_c | \mu = 501) \\
&= P(\overline{X} \le \overline{x}_c | \mu = 501) \\
&= P\Big(V \le c - \frac{\mu - \mu_0}{\sigma/\sqrt{n}}\Big) \\
&= P\Big(V \le 2 - \frac{501 - 500}{2}\Big) \\
&= P(V \le 1,5) = 0,933193
\end{aligned}$$

Die Wahrscheinlichkeit für einen Fehler 2. Art ist 93,32%, wenn in Wahrheit die mittlere Abfüllmenge $\mu = 501$ g beträgt.

h)

$$\begin{aligned} G(\mu = 499) &= 1 - P\Big(V \le c - \frac{\mu - \mu_0}{\sigma/\sqrt{n}}\Big) = 1 - P\Big(V \le 2 - \frac{499 - 500}{2}\Big) \\ &= 1 - P(V \le 2,5) = 1 - 0,99379 = 0,00621 \\ &= P(\text{„}H_1\text{“}|H_0) = \alpha(\mu = 499) \end{aligned}$$

Die Wahrscheinlichkeit für einen Fehler 1. Art (unberechtigte Annahme der H_1) beträgt $\alpha = 0,00621$, wenn das wahre $\mu = 499$ ist.

$$\begin{aligned} G(\mu = 502) &= 1 - P\Big(V \le 2 - \frac{502 - 500}{2}\Big) = 1 - P(V \le 1) \\ &= 1 - 0,841345 = 0,158655 \\ &= P(\text{„}H_1\text{“}|H_1) = 1 - \beta(\mu = 502) \end{aligned}$$

Die Wahrscheinlichkeit für die berechtigte Annahme der H_1, wenn das wahre $\mu = 502$ ist, beträgt 15,8655%.

Lösung 9-26: *FKK*
Anwendung des χ^2–Unabhängigkeitstests, weil die Beziehung zwischen zwei nominalskalierten Zufallsvariablen zu prüfen ist.
X: „Neigung zu FKK“; Y: „Region“
H_0: X und Y sind unabhängig; H_1: X und Y sind nicht unabhängig
$\alpha = 0,01$

X/Y	alt	neu	$h_{i\cdot}$
für	20 (26,7)	20 (13,3)	40
gegen	80 (73,3)	30 (36,7)	110
$h_{\cdot j}$	100	50	150

(in Klammern die erwarteten $\tilde{h}_{ij}$)
$V = \sum_{i=1}^{I=2} \sum_{j=1}^{J=2} \frac{(h_{ij} - \tilde{h}_{ij})^2}{\tilde{h}_{ij}}$ ist unter H_0 approximativ χ^2–verteilt mit $f = (I - 1)(J - 1) = 1$ Freiheitsgrad.
$c = \chi^2_{0,99;1} = 6,63$
Ablehnungsbereich der H_0:$\{v|v > 6,63\}$

$v = 1,7 + 3,4 + 0,6 + 1,2 = 6,9$
$v = 6,9 \in$ Ablehnungsbereich $\rightarrow$ „H_1“
Auf einem Signifikanzniveau von 1% und basierend auf einer Zufallsstichprobe vom Umfang $n = 150$ konnte statistisch bewiesen werden, dass die Neigung zu FKK von der Region der Befragten abhängig ist.
Lösung 9-27: *Gewinnspiel–Automat*
$U_i =$ „Ertrag pro Spiel“, $i = 1, \ldots, n = 50$, $n > 30$
$\overline{U} = (\sum_{i=1}^n U_i)/n = -0,58$, $S^2 = \sum(U_i - \overline{U})^2/(n-1) = 0,82$
$E(U_i) = \mu$, $Var(U_i) = \sigma^2$
$H_0 : \mu \ge 0$; $H_1 : \mu < 0$
asymptotisch

$$V = \frac{\overline{U} - \mu}{\sqrt{\sigma^2}}\sqrt{n} \approx \frac{\overline{U} - \mu}{\sqrt{S^2}}\sqrt{n} \sim N(0,1)$$

daher für $\mu_0 = 0$

$$\begin{aligned} 0,05 &= P(V \le c|H_0) = P\left(\frac{\overline{U} - \mu_0}{\sqrt{S^2}}\sqrt{n} \le \frac{c - \mu_0}{\sqrt{S^2}}\sqrt{n}\right) = \Phi\left(\frac{c - \mu_0}{\sqrt{S^2}}\sqrt{n}\right) \\ 0,95 &= 1 - \Phi\left(\frac{c - \mu_0}{\sqrt{S^2}}\sqrt{n}\right) = \Phi\left(-\frac{c - \mu_0}{\sqrt{S^2}}\sqrt{n}\right) \\ 1,64 &= -\frac{c - \mu_0}{\sqrt{S^2}}\sqrt{n} = -\frac{c - 0}{\sqrt{0,82}}\sqrt{50} \end{aligned}$$

$\rightarrow c = -1,64 \cdot \sqrt{0,82}/\sqrt{50} = -0,21$

Lösung 9-28: *Werbeaktion*
$U_i =$ Umsatz pro Kunde, $i = 1, \ldots, n = 900$, $n > 30$
$\overline{U} = (\sum_{i=1}^{n} U_i)/n$, $E(U_i) = \mu$, $Var(U_i) = \sigma^2$
$H_0 : \mu \geq 165$; $\qquad H_1 : \mu < 165$
asymptotisch:

$$V = \frac{\overline{U} - \mu}{\sqrt{\sigma^2}}\sqrt{n} \approx \frac{\overline{U} - \mu}{\sqrt{S^2}}\sqrt{n} \sim N(0;1)$$

daher für $\mu_0 = 165$

$$\begin{aligned}
0,05 &= P(V \leq c|H_0) = P\Big(\frac{\overline{U} - \mu_0}{\sqrt{S^2}}\sqrt{n} \leq \frac{c - \mu_0}{\sqrt{S^2}}\sqrt{n}\Big) = \Phi\Big(\frac{c - \mu_0}{\sqrt{S^2}}\sqrt{n}\Big) \\
0,95 &= 1 - \Phi\Big(\frac{c - \mu_0}{\sqrt{S^2}}\sqrt{n}\Big) = \Phi\Big(-\frac{c - \mu_0}{\sqrt{S^2}}\sqrt{n}\Big) \\
1,64 &= -\frac{c - \mu_0}{\sqrt{S^2}}\sqrt{n} = -\frac{c - 165}{\sqrt{900}}\sqrt{900}
\end{aligned}$$

$\rightarrow c = 165 - 1,64 = 163,36$
Lösung 9-29: *1000g–Portionen*
$X \sim N(1000; 25)$, $\quad \overline{X} \sim N(1000; 5)$, $\quad n = 25$
$\alpha = 0,05 = P(\overline{X} > 1000 + c$ oder $\overline{X} < 1000 - c) = 1 - P(1000 - c \leq \overline{X} \leq 1000 + c)$
$U = (\overline{X} - 1000)/5 \sim N(0;1)$
$0,05 = P(-c/5 \leq U \leq c/5) = \Phi(c/5) - \Phi(-c/5)$
d.h. $c/5$ ist das $1 - \alpha/2 = 0,975$ Quantil der $N(0;1)$ $\rightarrow c/5 = z_{0,975} = 1,96$; $\quad c = 9,8$

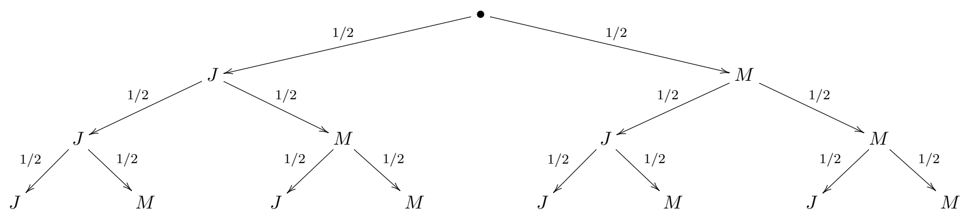
Lösung 9-30: *Arbeitsproduktivität*
X: „Arbeitsproduktivität“, Verteilung unbekannt, $\sigma = 0,8$ Stück/Stunde
$\overline{X}$: „Durchschnittliche Arbeitsproduktivität bei einer einfachen Zufallsstichprobe vom Umfang $n = 64$“ $\overline{X}$ ist approximativ $N(\mu; \sigma/\sqrt{n})$ (Begründung: Zentraler Grenzwertsatz, $n = 64 > 30$);
$\sigma/\sqrt{n} = 0,8/8 = 0,1$; $\quad \mu_0 = 5,5$; $\quad \alpha = 0,05 \quad z_{0,975} = 1,96$; $\quad H_0 : \mu = 5,5$; $\quad H_1 : \mu \neq 5,5$; $\quad \mu_1 = 5,1$

$$\begin{aligned}
\beta(\mu) &= 1 - G(\mu) \\
G(\mu) &= 1 - \Big[P\Big(V \leq c - \frac{\mu - \mu_0}{\sigma\sqrt{n}}\Big) - P\Big(V < -c - \frac{\mu - \mu_0}{\sigma\sqrt{n}}\Big)\Big] \\
\beta(\mu) &= P\Big(V \leq c - \frac{\mu - \mu_0}{\sigma\sqrt{n}}\Big) - P\Big(V < -c - \frac{\mu - \mu_0}{\sigma\sqrt{n}}\Big) \\
\beta(\mu_1 = 5,6) &= P(V \leq 1,96 - (5,6 - 5,5)/0,1) \\
&- P(V < -1,96 - (5,6 - 5,5)/0,1) \\
&= P(V \leq 0,96) - P(V \leq -2,96) \\
&= P(V \leq 0,96) - [1 - P(V \leq 2,96)] \\
&= 0,831472 - [1 - 0,998462] \\
&= 0,831472 - 0,001538 = 0,829934 \\
\beta(\mu_1 = 5,6) &= 0,8299
\end{aligned}$$

Lösung 9-31: *Anzahl der Kinder*
$H_0 : P(\text{Junge}) = P(\text{Mädchen}) \quad H_1 : P(\text{Junge}) \neq P(\text{Mädchen})$

$$V = \sum_{i=1}^{I} \frac{(h_i - np_i)^2}{np_i}$$

Unter H_0 gilt:



$P(3\text{J}, 0\text{M}) = 0,125 = 1/8$
$P(2\text{J}, 1\text{M}) = 0,125 \cdot 3 = 3/8$
$P(1\text{J}, 2\text{M}) = 0,125 \cdot 3 = 3/8$
$P(0\text{J}, 3\text{M}) = 0,125 = 1/8$
h_j – beobachtete absolute Häufigkeit np_i – unter H_0 erwartete absolute Häufigkeit
$np_i > 1$ für alle i und $np_i \geq 5$ für mindestens 80% der erwarteten Häufigkeiten erfüllt.

h_j	np_i	$h_j - np_i$	$(h_j - np_i)^2$	$\chi^2 = (h_j - np_i)^2/np_i$
16	25	−9	81	3,24
60	75	−15	225	3,00
92	75	17	289	3,853333
32	25	7	49	1,96

$v = 12,053333$ $f = I - 1 - k = 4 - 1 = 3;$ $k = 0$ (kein Parameter war zu schätzen)
aus Tabelle der Chi–Quadrat–Verteilung für $f = 3$:
$1 - \alpha : 0,99$ $\chi^2 = 11,35$ $1 - \alpha : 0,995$ $\chi^2 = 12,84$
signifikant zum 1%–Niveau

Lösung 9-32: *Ausfallsicherheit*
X = Ausfallzeit eines Servers in Stunden $\sim N(\mu, \sigma)$
Betriebszeit eines Servers: 365 Tage $\cdot$ 24 Stunden = 8760 Stunden
maximale mittlere Ausfallzeit lt. Hersteller: 1% von $8760 = 87,6$ Stunden
Der Hersteller will seine Behauptung statistisch untermauern, wobei er das Risiko einer Fehlentscheidung möglichst klein halten will. Da nur Abweichungen von μ_0 nach einer Seite von Bedeutung sind, wird ein einseitiger Test durchgeführt. Die Behauptung des Herstellers wird als Alternativhypothese formuliert, womit ein linksseitiger Test resultiert
$H_0 : \mu \geq \mu_0 = 87,6$ Stunden $\quad H_1 : \mu < \mu_0 = 87,6$ Stunden
Die Wahrscheinlichkeit eines Fehlers 1. Art $P(„H_1“|H_0)$ ist das Signifikanzniveau α, mit dessen Vorgabe das Risiko eines derartigen Fehlers gering gehalten werden kann. Damit wird die Zielstellung des Herstellers bei der Durchführung des Tests eingehalten. Da σ der Grundgesamtheit unbekannt ist, folgt die Teststatistik unter H_0 einer t–Verteilung mit $f = n - 1 = 24$ Freiheitsgraden. Kritischer Wert: $t_{0,95;24} = -1,711$

$$v = \frac{\overline{x} - \mu_0}{s}\sqrt{n} = \frac{84,2 - 87,6}{10}\sqrt{25} = -1,70$$

Da $v > t_{0,95;24}$ ist und damit in den Nichtablehnungsbereich von H_0 fällt, besteht keine Veranlassung H_0 abzulehnen.

Lösung 9-33: *Ausgaben für Urlaubsreisen*
Auswahlsatz $n/N = 10000/2500000 = 0,04 < 0,05 \to$ Endlichkeitskorrektur kann vernachlässigt werden; σ der Grundgesamtheit unbekannt; $N = 2500000$;
hypothetischer Wert der Gesamtausgaben:

$$10000000000 \to \mu_0 = 10000000000/2500000 = 4000$$

$n = 10000;$ $\overline{x} = 3780;$ $s = 2290$
Teststatistik:

$$V = \frac{\overline{X} - \mu_0}{S}\sqrt{n}$$

Wert der Teststatistik für die Stichprobe:

$$v = \frac{3780 - 4000}{2290}\sqrt{10000} = -9,606987 \approx -9,61$$

Lösung 9-34: *Kaffee Packungen 2*
Grundgesamtheit: X = Füllgewicht, Verteilung von X unbekannt, $\sigma = 15$, Grundgesamtheit kann als sehr groß angesehen werden, mittleres Füllgewicht μ unbekannt
hypothetischer Wert: $\mu_0 = 500$
einfache Zufallsstichprobe: $n = 100$, Stichprobenvariablen sind i.i.d.
linksseitiger Test auf μ : $H_0 : \mu \geq \mu_0$ und $H_1 : \mu < \mu_0$
Teststatistik V:

$$V = \frac{\overline{X} - \mu_0}{\sigma}\sqrt{n}$$

$\alpha = 0,05$; $\quad z_{1-\alpha} = z_{0,95} = 1,64$ aus Tabelle der Verteilungsfunktion $N(0;1)$, da aufgrund des großen Stichprobenumfangs und des ZGS die Verteilung von X approximativ normalverteilt ist; kritischer Wert: $-z_{1-\alpha} = -z_{0,95} = -1,64$ (wegen Symmetrie der Normalverteilung)
Ablehnungsbereich der H_0: $\{v|v < -z_{1-\alpha}\} = \{v|v < -1,64\}$
Nichtablehnungsbereich der H_0: $\{v|v \geq -z_{1-\alpha}\} = \{v|v \geq -1,64\}$
Fehler 2. Art: fälschliche Beibehaltung der H_0, d.h. „H_0“$|H_1$; $P($„H_0“$|H_1) = \beta$
Inhalt der Gütefunktion:

$$G(\mu) = \begin{cases} P(\text{„}H_1\text{“}|H_0) \leq \alpha & \text{für alle } \mu \geq \mu_0 \\ P(\text{„}H_1\text{“}|H_1) = 1-\beta & \text{für alle } \mu < \mu_0. \end{cases}$$

Es ist (wahr) $\mu = 497 < \mu_0 = 500$; es gilt in Wirklichkeit die Alternativhypothese und mit der Ablehnung von H_0 wird eine richtige Entscheidung getroffen. Es ist $P(V \in \text{Ablehnungsbereich der } H_0|\mu < \mu_0) = P(\text{„}H_1\text{“}|H_1) = 1-\beta$
Berechnung der Gütefunktion:

$$\begin{aligned} G(\mu) &= P\left(V \leq -z_{1-\alpha} - \frac{\mu-\mu_0}{\sigma/\sqrt{n}}\right) = P\left(V \leq -1,64 - \frac{497-500}{15/\sqrt{100}}\right) \\ &= P\left(V \leq -1,64 - \frac{-3}{1,5}\right) = P(V \leq -1,64+2) = P(V \leq 0,36) \\ &= 1-\beta = 0,64058 \end{aligned}$$

$\rightarrow \beta = 0,35942 \approx 0,36$

Lösung 9-35: *Lagerhaltungsprobleme*
X = Anzahl der nachgefragten Produkte pro Tag
Chi-Quadrat-Anpassungstest bei Wahl der hypothetischen Verteilung $F_0(x)$ = Poisson-Verteilung. Der Parameter $\lambda = E(X)$ ist unbekannt und muss aus der Stichprobe geschätzt werden: $\hat{\lambda} = 200/100 = 2,0$. Aus der Tabelle der Verteilungsfunktion der Poisson–Verteilung PO(2,0) lassen sich die unter H_0 gültigen Wahrscheinlichkeiten $p_i = P(X = x_i)$ ermitteln. (5. Spalte der folgenden Tabelle). Für alle Klassen ist die Voraussetzung $n \cdot p_i \geq 5$ erfüllt. Die Anzahl der Freiheitsgrade des Chi–Quadrat–Anpassungstests beträgt $f = I - 1 - k$ mit I der Anzahl der Klassen und k der Anzahl der aus der Stichprobe zu schätzenden Parameter. Damit resultiert: $f = 6 - 1 - 1 = 4$.

i	x_i	h_i	x_ih_i	p_i	np_i
1	0	17	0	0,1353	13,53
2	1	20	20	0,2707	27,07
3	2	27	54	0,2707	27,07
4	3	18	54	0,1804	18,04
5	4	18	72	0,0902	9,02
6	5 und mehr	0	0	0,0527	5,27
$\sum$		100	200	1,0000	100

Aus der Tabelle der Verteilungsfunktion der Chi–Quadrat–Verteilung findet man:
$\chi^2_{0,95;4} = 9,49$

Lösung 9-36: *Grönländische Bohrlochkerne*
Gegeben: $\mu_0 = -25$C; $\quad n = 100$; $\quad \alpha = 0,025$; $\quad \overline{x} = -24$C; $\quad s = 1,5$C (diese Stichprobenergebnisse werden nicht benötigt); $\quad \mu = -24,8$C
Da die Forscher nachweisen wollen, dass eine Erwärmung des Eises stattgefunden hat wird ein rechtsseitiger Test durchgeführt:
$H_0 : \mu \leq \mu_0$ $(= -25\text{C})$ gegen $H_1 : \mu > \mu_0$ $(= -25\text{C})$. Daher $z_{0,975} = 1,96$.
Es ist der Wert der Gütefunktion $G(\mu = -24,8\text{C})$ zu berechnen, denn

- die Gütefunktion $G(\mu)$ gibt die Wahrscheinlichkeit der Ablehnung von H_0 in Abhängigkeit vom Parameter μ an:

$$G(\mu) = P(V \in \text{Ablehnungsbereich der } H_0|\mu);$$

- für alle zulässigen Werte von $\mu > \mu_0$ gilt in Wirklichkeit die Alternativhypothese und mit der Ablehnung der Nullhypothese wird eine richtige Entscheidung getroffen; das ist hier wegen $\mu(= -24,8C) > \mu_0(= -25C)$ gegeben;
- es ist $P(V \in \text{Ablehnungsbereich der } H_0|\mu > \mu_0) = P(„H_1“|H_1) = 1-\beta$.

$$\begin{aligned} G(\mu = -24,8) &= 1 - P\left(V \leq z_{1-\alpha} - \frac{\mu - \mu_0}{\sigma/\sqrt{n}}\right) \\ &= 1 - P\left(V \leq 1,96 - \frac{-24,8-(-25)}{2}\sqrt{100}\right) \\ &= 1 - P(V \leq 0,96) = 1 - 0,831482 = 0,168518 \approx 0,17 \end{aligned}$$

Lösung 9-37: *Fachgebiete*
Anwendung des Chi-Quadrat-Anpassungstests zur Prüfung der Hypothese, ob die von Bärbel beobachtete Verteilung ($h_{Stat} = 5, h_{VWL} = 35, h_{BWL} = 50, h_{WI} = 10$) mit der theoretisch erwarteten Verteilung (Gerdas Behauptung: $nf_{Stat} = 10, nf_{VWL} = 30, nf_{BWL} = 40, nf_{WI} = 20$) übereinstimmt. Beide Approximationsbedingungen sind erfüllt.
Prüfwert:

$$\begin{aligned} v &= \sum_i [(h_i - np_i)^2/np_i] \\ &= (5-10)^2/10 + (35-30)^2/30 + (50-40)^2/40 + (10-20)^2/20 \\ &= 25/10 + 25/30 + 100/40 + 100/20 = (300+100+300+600)/120 \\ &= 1300/120 = 10,83 \approx 10,8 \end{aligned}$$

Lösung 9-38: *Benzinverbrauch Test*
$\mu_0 = 6;\ H_0 : \mu = 6;\ H_1 : \mu \neq 6$, zweiseitiger Test, da Abweichungen von der Behauptung, also nach beiden Seiten; $X \sim N(\mu_0 = 6; \sigma^2)$, σ^2 unbekannt;
$\overline{x} = \sum_i x_i/n = 97,6/16 = 6,1$
$s^2 = \sum_i (x_i - \overline{x})^2/(n-1) = 0,6615/15 = 0,0441;\ s = 0,21$
$v = \frac{\overline{x} - \mu_0}{s}\sqrt{n} = \frac{6,1-6}{0,21} \cdot 4 = 1,90476 \approx 1,905$
$t_{1-\alpha/2;n-1} = t_{0,975;15} = 2,132$

10 Regressionsanalyse

Lösung 10-1: *Gesamtkosten und Produktionsmenge*
$\widehat{y}_i = 17,253 + 4,039x_i$
Lösung 10-2: *Konsumausgaben und verfügbares Einkommen*
$\widehat{y}_i = 1,94 + 0,78x_i$
Lösung 10-3: *Querschnittsanalyse von 11 Unternehmen*

a) $\sum_{i=1}^{11} y_i = 191,\ \sum_{i=1}^{11} y_i^2 = 3446{,}92$

a.1) $\sum_{i=1}^{11} x_{i1} = 1671{,}90,\ \sum_{i=1}^{11} x_{i1}^2 = 259\,297{,}25,\ \sum_{i=1}^{11} x_{i1}y_i = 29\,829{,}70,$

$$b_1^{(1)} = \frac{11 \cdot 29\,829{,}70 - 1671{,}90 \cdot 191}{11 \cdot 259\,297{,}25 - 1671{,}90^2} = 0{,}1542,$$
$$b_0^{(1)} = \frac{191}{11} - 0{,}1542 \cdot \frac{1671{,}90}{11} = -6{,}0768,$$
$$R_{y1}^2 = \frac{(11 \cdot 29\,829{,}70 - 1671{,}90 \cdot 191)^2}{(11 \cdot 259\,297{,}25 - 1671{,}90^2) \cdot (11 \cdot 3446{,}92 - 191^2)} = 0{,}9450,$$
$$\hat{y}_1 = -6{,}0768 + 0{,}1542x_1$$

a.2) $\sum_{i=1}^{11} x_{i2} = 1197{,}40,\ \sum_{i=1}^{11} x_{i2}^2 = 132\,804{,}82,\ \sum_{i=1}^{11} x_{i2}y_i = 21\,344{,}04,$

$$b_1^{(2)} = \frac{11 \cdot 21\,344{,}04 - 1197{,}40 \cdot 191}{11 \cdot 132\,804{,}82 - 1197{,}40^2} = 0{,}2245,$$
$$b_0^{(2)} = \frac{191}{11} - 0{,}2245 \cdot \frac{1197{,}40}{11} = -7{,}0749,$$
$$R_{y2}^2 = \frac{(11 \cdot 21\,344{,}04 - 1197{,}40 \cdot 191)^2}{(11 \cdot 132\,804{,}82 - 1197{,}40^2) \cdot (11 \cdot 3446{,}92 - 191^2)} = 0{,}9513,$$
$$\hat{y}21 = -7{,}0749 + 0{,}2245x_2$$

a.3) $\sum_{i=1}^{11} x_{i3} = 29{,}20,\ \sum_{i=1}^{11} x_{i3}^2 = 88{,}44,\ \sum_{i=1}^{11} x_{i3}y_i = 519{,}52,$

$$b_1^{(3)} = \frac{11 \cdot 519{,}52 - 29{,}20 \cdot 191}{11 \cdot 88{,}44 - 29{,}20^2} = 1{,}1441,$$
$$b_0^{(3)} = \frac{191}{11} - 1{,}1441 \cdot \frac{29{,}20}{11} = 14{,}3266,$$

$$R_{y3}^2 = \frac{(11 \cdot 519{,}52 - 29{,}20 \cdot 191)^2}{(11 \cdot 88{,}44 - 29{,}20^2) \cdot (11 \cdot 3446{,}92 - 191^2)} = 0{,}1096,$$
$$\hat{y}_3 = 14{,}3266 + 1{,}1441 x_3$$

b.1) $r_{y1} = \sqrt{R_{y1}^2} = 0{,}9721$; $r_{y2} = \sqrt{R_{y2}^2} = 0{,}9753$; $r_{y3} = \sqrt{R_{y3}^2} = 0{,}3311$;

b.2) $\sum_{i=1}^{11} x_{i1} = 1671{,}90$, $\sum_{i=1}^{11} x_{i1}^2 = 259\,297{,}25$,
$\sum_{i=1}^{11} x_{i2} = 1197{,}40$, $\sum_{i=1}^{11} x_{i2}^2 = 132\,804{,}82$,
$\sum_{i=1}^{11} x_{i1} x_{i2} = 185\,557{,}02$,

$$r_{12} = \frac{11 \cdot 185\,557{,}02 - 1671{,}90 \cdot 1197{,}40}{\sqrt{(11 \cdot 259\,297{,}25 - 1671{,}90^2) \cdot (11 \cdot 132\,804{,}82 - 1197{,}40^2)}}$$
$$r_{12} = 0{,}9973$$

$\sum_{i=1}^{11} x_{i1} = 1671{,}90$, $\sum_{i=1}^{11} x_{i1}^2 = 259\,297{,}25$,
$\sum_{i=1}^{11} x_{i3} = 29{,}20$, $\sum_{i=1}^{11} x_{i3}^2 = 88{,}44$,
$\sum_{i=1}^{11} x_{i1} x_{i3} = 4478{,}28$,

$$r_{12} = \frac{11 \cdot 4478{,}28 - 1671{,}90 \cdot 29{,}20}{\sqrt{(11 \cdot 259\,297{,}25 - 1671{,}90^2) \cdot (11 \cdot 88{,}44 - 29{,}20^2)}}$$
$$r_{12} = 0{,}1687$$

$\sum_{i=1}^{11} x_{i2} = 1197{,}40$, $\sum_{i=1}^{11} x_{i2}^2 = 132\,804{,}82$,
$\sum_{i=1}^{11} x_{i3} = 29{,}20$, $\sum_{i=1}^{11} x_{i3}^2 = 88{,}44$,
$\sum_{i=1}^{11} x_{i2} x_{i3} = 3203{,}89$,

$$r_{12} = \frac{11 \cdot 3203{,}89 - 1197{,}40 \cdot 29{,}20}{\sqrt{(11 \cdot 132\,804{,}82 - 1197{,}40^2) \cdot (11 \cdot 88{,}44 - 29{,}20^2)}}$$
$$r_{12} = 0{,}1545$$

Lösung 10-4: *Hypothekenzinssatz*

	x_i	y_i	$x_i - \bar{x}$	$(x_i - \bar{x})^2$	$y_i - \bar{y}$	$(y_i - \bar{y})^2$	$(x_i - \bar{x})(y_i - \bar{y})$
1	6	3 000	-1	1,0	500	250 000,0	-500,0
2	5	3 200	-2	4,0	700	490 000,0	-1 400,0
3	7	2 500	0	0,0	0	0,0	0,0
4	7	2 300	0	0,0	-200	40 000,0	-0,0
5	8	2 000	1	1,0	-500	250 000,0	-500,0
6	9	2 000	2	4,0	-500	250 000,0	-1 000,0
Summe	42	15 000	0	10,0	0	1 280 000,0	-3 400,0
Mittel	7	2 500	0	1,7	0	213 333,3	-566,7

a) $r = \dfrac{-3400}{\sqrt{10 \cdot 1\,280\,000}} = -0{,}9503$

b) $b_1 = \dfrac{-3400}{10} = -340$, $b_0 = 2500 - (-340) \cdot 7 = 4880$
$\widehat{y} = 4880 - 340 \cdot x$

c) $R^2 = r^2 = -0{,}9503^2 = 0{,}9031$

d) $4880 - 340 \cdot 4 = 3520$ Mio EUR, $4880 - 340 \cdot 7,5 = 2330$ Mio EUR

Regression-Hypothekenzinssatz.mp4

Lösung 10-5: *Konsumausgaben*

a) $\widehat{y}_i = 211{,}82 + 0{,}813 x_i$

b) 2488,22 EUR Konsumausgaben

Lösung 10-6: *Kunstdünger*

a) ja

b) $\widehat{y}_i = 19{,}93 + 5{,}0526 x_i$

c) 75,5086 dt

d) $B = 0{,}9753$

Lösung 10-7: *Zusätzliche statistische Einheit*
Lösung g)
$b_1 = (10 \cdot 304 - 40 \cdot 70)/(10 \cdot 180 - 40^2) = 240/200 = 1,2$
$b_0 = (70 \cdot 180 - 40 \cdot 304)/10 \cdot 180 - 40^2) = 440/200 = 2,2$

Lösung 10-8: *Gewinn eines Unternehmens*

$$\hat{y}_i = a + bx_i$$

$$\begin{aligned} a &= \frac{\sum y_i \sum x_i^2 - \sum x_i \sum x_i y_i}{n \sum x_i^2 - \sum x_i \sum x_i} \\ &= \frac{0 - 55 \cdot 99}{10 \cdot 385 - 55^2} = -6,6 \end{aligned}$$

$$\begin{aligned} b &= \frac{n \sum x_i y_i - \sum x_i \sum y_i}{n \sum x_i^2 - \sum x_i \sum x_i} \\ &= \frac{10 \cdot 99 - 0}{10 \cdot 385 - 55^2} = 1,2 \end{aligned}$$

$$\hat{y}_i = -6,6 + 1,2x_i$$

Lösung 10-9: *Arbeitslosenquoten*

$$\sum_{t=0}^{3} t = 6 \quad \sum_{t=0}^{3} x_t = 45,2 \quad \sum_{t=0}^{3} tx_t = 71,5; \quad \sum_{t=0}^{3} t^2 = 14$$

$$\begin{aligned} b &= \frac{(T+1)\sum tx_t - \sum x_t \sum t}{(T+1)\sum t^2 - (\sum t)^2} \\ &= \frac{4 \cdot 71,5 - 45,2 \cdot 6}{4 \cdot 14 - 6^2} = \frac{286 - 271,2}{56 - 36} = \frac{14,8}{20} = 0,74 \\ a &= \frac{\sum x_t}{T+1} - b\frac{\sum t}{T+1} = \frac{45,2}{4} - 0,74 \cdot \frac{6}{4} = 11,3 - 1,11 = 10,19 \\ \hat{y}_i &= 10,19 + 0,74 \cdot x_i \\ \hat{y}_4 &= 10,19 + 0,74 \cdot x_4 = 10,19 + 0,74 \cdot 4 = 13,15 \end{aligned}$$

Lösung 10-10: *Quadratmetermiete*

	x_i	y_i	x_i^2	y_i^2	$x_i \cdot y_i$
1	40	12	1 600	144,0	480
2	40	12	1 600	144,0	480
3	40	15	1 600	225,0	600
4	60	12	3 600	144,0	720
5	80	10	6 400	100,0	800
6	80	10	6 400	100,0	800
7	90	9	8 100	81,0	810
8	90	10	8 100	100,0	900
9	90	10	8 100	100,0	900
10	90	10	8 100	100,0	900
Summe	700	110	53 600	1 238,0	7 390
Mittel	70	11	5 360	123,8	739

$$b_1 = \frac{10 \cdot 7390 - 700 \cdot 110}{10 \cdot 53\,600 - 700^2} = -0{,}0674,$$
$$b_0 = \frac{110}{10} - \frac{700}{10} \cdot -0{,}0674 = 15{,}7174,$$

Regression-Quadratmetermiete.mp4

Lösung 10-11: *Ökonomische Variablen*

$$b_1 = \frac{n\sum x_iy_i - \sum x_i \sum y_i}{n\sum x_i^2 - \sum x_i \sum x_i} = \frac{10\cdot 304 - 40\cdot 70}{10\cdot 180 - 40\cdot 40} = \frac{240}{200} = 1,2$$

$$b_0 = \frac{\sum y_i \sum x_i^2 - \sum x_i \sum x_iy_i}{n\sum x_i^2 - \sum x_i \sum x_i} = \frac{70\cdot 180 - 40\cdot 304}{10\cdot 180 - 40\cdot 40} = \frac{440}{200} = 2,2$$

$$b_0 = \overline{y} - b_1\overline{x} = 7 - 1,2\cdot 4 = 2,2$$

Lösung 10-12: *Immobiliensachverständiger*

Objekt i	Alter x_i	Preis y_i	x_iy_i	x_i^2
1	15	190	2850	225
2	12	210	2520	144
3	3	400	1200	9
4	17	125	2125	289
5	5	300	1500	25
6	8	197	1576	64
$\sum$	60	1422	11771	756

$$b_1 = \frac{n\sum x_iy_i - \sum x_i \sum y_i}{n\sum x_i^2 - \sum x_i \sum x_i} = \frac{6\cdot 11771 - 60\cdot 1422}{6\cdot 756 - 60^2} = \frac{-14694}{936} = -15,699$$

$$b_0 = \overline{y} - b_1\overline{x} = \frac{1422}{6} - (-15,699)\frac{60}{6} = 393,99$$

$$\hat{y}_i = b_0 + b_1x_i = 393,99 - 15,699\cdot 1 = 378,291$$

Lösung 10-13: *Umsatz und Werbeetat*
Schätzung des Parameters b_1 in der linearen Regressionsfunktion $\hat{y}_i = b_0 + b_1x_i$.

Filiale	1	2	3	4	5	6	$\sum$
y_i : Umsatz (1000 EUR)	20	16	18	17	12	13	96
x_i : Werbeetat (100 EUR)	29	25	28	26	20	22	150
x_i^2	841	625	784	676	400	484	3810
x_iy_i	580	400	504	442	240	286	2452

$$\begin{aligned} b_1 &= \frac{n\sum x_iy_i - \sum x_i \sum y_i}{n\sum x_i^2 - \sum x_i \sum x_i} \\ &= \frac{6\cdot 2452 - 150\cdot 96}{6\cdot 3810 - 150^2} = \frac{14712 - 14400}{22860 - 22500} = \frac{312}{360} = 0,866666 \\ &= 0,867 \end{aligned}$$

Lösung 10-14: *Kosten und Output*
$X =$ Output, $Y =$ Kosten
Gegeben: $s_y^2 = 1801,6$; $s_{yx} = 67,2$
Gesucht: $b_1 = s_{yx}/s_x^2$
$s_x^2 = \sum x_i^2/n - \overline{x}^2 = 1468/10 - 12^2 = 146,8 - 144 = 2,8$
$b_1 = s_{yx}/s_x^2 = 67,2/2,8 = 24$

Lösung 10-15: *Alter und Händlerverkaufspreis*
Gegeben: $s_{xy} = -5,4$ $\quad s_y^2 = 4$ $\quad R_{yx}^2 = 0,81$
Es ist $r_{yx} = s_{yx}/s_xs_y$. Daraus folgt: $s_x = s_{yx}/(r_{yx}s_y)$
Ferner ist: $r_{yx} = -0,9$ (r_{yx} und die Kovarianz haben das gleiche Vorzeichen); $s_y = 2$
$s_x = s_{yx}/r_{yx}s_y = -5,4/(-0,9\cdot 2) = 3$

11 Zeitreihenanalyse

Lösung 11-1: *Warenausfuhr*

a) $i_G = \sqrt[5]{\frac{74{,}237}{53{,}892}} = 1{,}0662$

b) $x_{92} = 84{,}391$ Mrd. EUR

c) T =4,65; also im Jahre 1995

Lösung 11-2: *Anzahl der Beschäftigten*

a) nein; Angabe des Nullpunktes fehlt; $t = 0 \mathrel{\hat{=}} 1987$; $t = 1 \mathrel{\hat{=}} 1988$

b) Durchschnittlich sinkt die Anzahl der Beschäftigten um 9 Beschäftige pro Jahr.

Lösung 11-3: *Mikroprozessoren*

a) Zeitreihe

b) $i_G = 1{,}19$

c) exponentieller Trend; weist eine kleinere Streuung als der lineare Trend auf;
$\widehat{x}_t = 92396{,}57 \cdot 1{,}1867^t$ mit $t = 0 \mathrel{\hat{=}} 1985$; $t = 1 \mathrel{\hat{=}} 1986$

d) Basis i_G: $\widehat{x}_{92} = 311\ 542$ Stück;
Basis exp. Trend: $\widehat{x}_{92} = 306216{,}26$ Stück

Lösung 11-4: *Speiseeis*

a) $\widehat{x}_t = 138{,}44 + 9{,}67\ t$ mit $t = 0 \mathrel{\hat{=}} 1.1.1989$; $t = 1 \mathrel{\hat{=}} 1.7.1989$

b) $\widehat{x}_{t,j}^{ZRM} = (140 + 10\ t)\overline{s}_j$; $\overline{s}_1 = 0{,}9$; $\overline{s}_2 = 1{,}1$ mit $t = 0 \mathrel{\hat{=}} 1.1.1989$; $t = 1 \mathrel{\hat{=}} 1.7.1989$

d) $\widehat{x}_{t,j}^{ZRM} = (60 + 20\ t)\overline{s}_j$; $\overline{s}_1 = 0{,}9$; $\overline{s}_2 = 1{,}1$ mit $t = 0 \mathrel{\hat{=}} 1.1.1985$; $t = 1 \mathrel{\hat{=}} 1.1.1986$

e) 209 kg

Lösung 11-5: *Quartalsproduktion*
Additives Zeitreihenmodell; 2 305 000; 2 520 000; 2 565 000; 2 610 000;
Jahresproduktion: 10 000 000
Lösung 11-6: *Quartalsproduktion 2*
Multiplikatives Zeitreihenmodell; 2 277 000; 5 528 050; 4 015 110; 1 923 906,875;
Jahresproduktion: 13 744 066,875
Lösung 11-7: *Trendfunktion*

b) 2740,047 Mio. Personen

Lösung 11-8: *Maschinenzeitfondsauslastungen*

a) $\widehat{x}_t = 68{,}1111 + 2{,}967t$ mit $t = 0 \mathrel{\hat{=}} 0.$Monat; $t = 1 \mathrel{\hat{=}} 1.$Monat
$\widehat{x}_t = 68{,}766 \cdot 1{,}03722^t$ mit $t = 0 \mathrel{\hat{=}} 0.$Monat; $t = 1 \mathrel{\hat{=}} 1.$Monat

b) R^2(lin. Trend) = 0,9377;
R^2(exp. Trend) = 0,9214

c) 103,715 % (!, Interpretation)

Lösung 11-9: *Eheschließungen und Ehescheidungen*

a) Eheschließungen:
$\widehat{x}_t = 367,25 + 1,25t$ mit $t = -1 \mathrel{\hat{=}} 1983$; $t = +1 \mathrel{\hat{=}} 1984$
$\widehat{x}_t = 446,25 - 12,774t$ mit $t = -1 \mathrel{\hat{=}} 1965$; $t = +1 \mathrel{\hat{=}} 1970$
Ehescheidungen:
$\widehat{x}_t = 119,5 + 2,012t$ mit $t = -1 \mathrel{\hat{=}} 1983$; $t = +1 \mathrel{\hat{=}} 1984$
$\widehat{x}_t = 81,375 + 4,3t$ mit $t = -1 \mathrel{\hat{=}} 1965$; $t = +1 \mathrel{\hat{=}} 1970$

b) Eheschließungen: s = 4,17, v = 0,01135; s = 26,89, v = 0,06
Ehescheidungen: s = 6, v = 0,05; s = 17,98, v = 0,221

c) Eheschließungen: Basis 1980-1987 383 500; Basis 1950-1985 331 284
Ehescheidungen: Basis 1980-1987 145 656; Basis 1950-1985 120 075

Lösung 11-10: *Bruttosozialprodukt von Deutschland*

a) $\widehat{x}_t = 1444,737 + 28,413t$ mit $t = 0 \,\widehat{=}\, 1980$; $t = +1 \,\widehat{=}\, 1981$

b) $\widehat{x}_{89} = 1700{,}454$ Mrd. EUR; 1990 Vorhersage auf der Basis dieses Trends sehr fragwürdig wegen Vereinigung Deutschlands

Lösung 11-11: *Hausschlachtungen von Schweinen*
$\widehat{x}^{ZRM}_{t,j}$ = (9,5 - 0,15 t) + $\overline{s}_j$; $\overline{s}_1$ = 3,58; $\overline{s}_2$ = -3,27; $\overline{s}_3$ = -4,45; $\overline{s}_4$ = 4,03;
mit $t = 0 \,\widehat{=}\,$ 4.Quartal 1989; $t = +1 \,\widehat{=}\,$ 1.Quartal 1990

Lösung 11-12: *Indizes der Aktienkurse*

a) -; 84,8; 85; 85,1; 84,9; 82,7; 80,9; 79,1; 77,6; 76,3; 77,3; -
-; -; 85,1; 84,9; 84; 82,8; 81,1; 79; 77,7; 77,3; -; -

b) $\widehat{x}_t = 87,3 - 0,89t$ mit $t = 0 \,\widehat{=}\,$ 0.Monat; $t = +1 \,\widehat{=}\,$ 1.Monat

Lösung 11-13: *Transportleistung*

a) Die Anpassungsunterschiede zwischen einem additiven Zeitreihenmodell mit linearem Trend und einem multiplikativen Zeitreihenmodell mit linearem Trend sind sehr gering, deshalb wird ersteres gewählt:
$\widehat{x}^{ZRM}_{t,j} = (10,6363 + 0,479t) + \overline{s}_j$, $\overline{s}_1$ = 1,637; $\overline{s}_2$ = -2,507; $\overline{s}_3$ = -1,65;
$\overline{s}_4$ = 2,54;
mit $t = 0 \,\widehat{=}\,$ 4.Quartal 1989; $t = +1 \,\widehat{=}\,$ 1.Quartal 1990

b) s = 0,50

c) 18,52; 14,85; 16,19; 20,86; gesamt: 70,42 10^3tkm

Lösung 11-14: *Souvenirhändler*

a) –; 50; 100; 200; 400; 800; – (3.Ordnung)

b) $\widehat{x}_t = 50{\cdot}2^t$ mit $t = 0 \,\widehat{=}\,$ Februar 1992; $t = +1 \,\widehat{=}\,$ März 1992

c) $\widehat{x}_{Januar}$ = 25; $\widehat{x}_{Juli}$ = 1600

Lösung 11-15: *Haushalte eines Landes*
$a = (213 \cdot 55 - 15 \cdot 678)/(5 \cdot 55 - 15^2) = 1545/50 = 30,9$
$b = (5 \cdot 678 - 213 \cdot 15)/(5 \cdot 55 - 15^2) = 195/50 = 3,9$

Lösung 11-16: *Haushalte eines Landes 2*
$x_{1993} = 50 \cdot 1,53 = 76,5$
$i_G = \sqrt[9]{\frac{76,5}{34}} = 1,094287 \approx 1,094$

Lösung 11-17: *Telefonkosten*
$\hat{x}_t = a + bt$; Zeitcodierung: $t = 0$ 1990; $t = 1$ 1991;$\ldots$ $t = 5$ 1995

$$\begin{aligned} a &= \frac{\sum x \sum t^2 - \sum t \sum x_t t}{T \sum t^2 - (\sum t)^2} \\ &= (213 \cdot 55 - 15 \cdot 678)/(5 \cdot 55 - 15^2) \\ &= 1545/50 \\ &= 30,9 \end{aligned}$$

$$\begin{aligned} b &= \frac{T \sum x_t t - \sum x_t \sum t}{T \sum t^2 - (\sum t)^2} \\ &= (5 \cdot 678 - 213 \cdot 15)/5 \cdot 55 - 15^2) \\ &= 195/50 \\ &= 3,9 \end{aligned}$$

$$\hat{x}_t = 30,9 + 3,9t$$

Lösung 11-18: *Telefonkosten 2*
Begründung für Funktionsform:
Lineare Trendfunktion, da absoluter Zuwachs (in EUR) gegeben ist:
$\hat{x}_t = a + bt$
Berechnungen:
Gegeben: $T = 10$; $\sum x_t = 28$; $b = 0,125$
Entweder:
jährlicher Zuwachs gegeben, T gerade $\rightarrow$ Zeitcodierung: $t = 0 = 1980$; $t = 1 = 1981, \ldots, t = 10 = 1990$

$$\begin{aligned} a &= \overline{x} - b\overline{t} \\ &= \frac{\sum x_t}{T} - b\frac{\sum t}{T} \\ &= \frac{28}{10} - 0,125\frac{55}{10} \\ &= 2,8 - 0,6875 \\ &= 2,1125 \text{ Mill. EUR} \end{aligned}$$

Trendfunktion:
$x_t = 2,1125 + 0,125t$ mit $t = 0 = 1980$, $t = 1 = 1981$
Oder:
Es wird der halbjährliche Zuwachs genommen: $b = 0,125/2 = 0,0625$
Zeitcodierung kann dann gewählt werden:
$t = -9 = 1981; \ldots; t = -1 = 1985; t = 1 = 1986; \ldots; t = 9 = 1990$

$$a = \overline{x} = \frac{\sum x_t}{T} = \frac{28}{10} = 2,8 \text{ Mill. EUR}$$

Trendfunktion:
$x_t = 2,8 + 0,0625t$ mit $t = -1 = 1985; t = 1 = 1986$

Lösung 11-19: *Gecrashte Festplatte*
Zu bestimmen ist $\hat{y}_{43}^{ZRM}$ anhand des geringsten Bestimmtheitsmaßes. Zu berechnen:
$b = \dfrac{T\sum tx_t - \sum x_t \sum t}{T\sum t^2 - (\sum t)^2}, \quad a = \dfrac{\sum x_t}{T} - b\dfrac{\sum t}{T}$
$R^2 = 1 - \dfrac{\sum_i (y_i - \hat{y}_i^{ZRM})^2}{\sum_i (y_i - \overline{y})^2}$
$\hat{y}_{43} = a + b \cdot 43$, bzw. $\hat{y}_{43}^{ZRM} = \hat{y}_{43} + \overline{s}_3$, bzw. $\hat{y}_{43}^{ZRM} = \hat{y}_{43} \cdot \overline{s}_3$

a	$-2,738462$		
b	0,530769		
	Saisonkomponente		
	keine	additiv	multiplikativ
R^2	0,80	0,81	0,48
$\hat{y}_{43}^{ZRM}$	20,4		

Lösung 11-20: *Arbeitslosenquoten*
Zeitpunkte $t = 0, \ldots, T = 3$

$$\sum_{t=0}^{3} t = 6; \qquad \sum_{t=0}^{4} x_t = 45,2; \qquad \sum_{t=0}^{4} tx_t = 71,5; \qquad \sum_{t=0}^{4} t^2 = 14$$

$$\begin{aligned} b &= \frac{(T+1)\sum tx_t - \sum x_t \sum t}{(T+1)\sum t^2 - (\sum t)^2} \\ &= \frac{4 \cdot 71,5 - 45,2 \cdot 6}{4 \cdot 14 - 6^2} = \frac{286 - 271,2}{56 - 36} = \frac{14,8}{20} = 0,74 \\ a &= \frac{\sum x_t}{T+1} - b\frac{\sum t}{T+1} = \frac{45,2}{4} - 0,74 \cdot \frac{6}{4} = 11,3 - 1,11 = 10,19 \\ \hat{x}_4 &= 10,19 + 0,74 \cdot 4 = 13,15 \end{aligned}$$

Lösung 11-21: *Wachstum des Bruttoinlandsprodukts*
Geometrisches Mittel:

$$i_G = \sqrt[4]{1,027 \cdot 1,018 \cdot 1,014 \cdot 1,022} = \sqrt[4]{1,083446} = 1,020239 = 102,0239\%$$

Lösung 11-22: *Benutzer des Dial-In-Service*
$t = 0, \ldots, T = 3$

$$\sum_{t=0}^{3} t = 6; \quad \sum_{t=0}^{4} x_t = 10097; \quad \sum_{t=0}^{4} tx_t = 17734; \quad \sum_{t=0}^{4} t^2 = 14$$

$$\begin{aligned} b &= \frac{(T+1)\sum tx_t - \sum x_t \sum t}{(T+1)\sum t^2 - (\sum t)^2} = \frac{4 \cdot 17734 - 10097 \cdot 6}{4 \cdot 14 - 6^2} = 517,7 \\ a &= \frac{\sum x_t}{T+1} - b\frac{\sum t}{T+1} = 2524,25 - 517,7 \cdot 1,5 = 1747,7 \\ \hat{x}_5 &= 1747,7 + 517,7 \cdot 5 = 4336,2 \approx 4337 \end{aligned}$$

Lösung 11-23: *Eheschließungen*
$t = 0, \ldots, T = 4$

$$\sum_{t=0}^{4} t = 10; \quad \sum_{t=0}^{4} x_t = 40,7; \quad \sum_{t=0}^{4} tx_t = 69,2; \quad \sum_{t=0}^{4} t^2 = 30$$

$$\begin{aligned} b &= \frac{(T+1)\sum tx_t - \sum x_t \sum t}{(T+1)\sum t^2 - (\sum t)^2} = \frac{5 \cdot 69,2 - 40,7 \cdot 10}{5 \cdot 30 - 10^2} = -1,22 \\ a &= \frac{\sum x_t}{T+1} - b\frac{\sum t}{T+1} = 8,14 + 1,22 \cdot 2 = 10,58 \\ \hat{x}_5 &= 10,58 - 1,22 \cdot 5 = 4,48 \end{aligned}$$

Lösung 11-24: *Bauhauptgewerbe*
Jahresumsätze:
$x_{1991} = x_0 = 260, \quad x_{1992} = x_1 = 410, \quad x_{1993} = x_2 = 580, \quad x_{1994} = x_3 = 700$
Anwendung des geometrischen Mittels in Form des mittleren Entwicklungstempos:

$$i_G = \sqrt[n]{\frac{x_n}{x_0}}; \quad i_G = \sqrt[3]{\frac{700}{260}} = 1,391$$

Lösung 11-25: *Abschreibung*
Zeitpunkt t (Beginn des Jahres) $t = 0, 1, 2, 3, 4, 5, 6, 7 \rightarrow$ Zeiträume: 7
Anwendung des geometrischen Mittels, da nach mittleren relativen Veränderungen gefragt:

$$x_5 = 1 \text{ EUR} \qquad x_0 = 50000 \text{ EUR}$$

$$\bar{i}_G = \sqrt[7]{\frac{x_n}{x_0}} = \sqrt[7]{\frac{1}{50000}} = \sqrt[7]{0,00002} = 0,213166$$

mittlerer Abschreibungssatz: 0,2132
zur Kontrolle:

t	Restbuchwert (EUR) zu Beginn des Jahres t			Abschreibung (EUR) im Jahre $t-1$
0	50000, 00			
1	50000, 00	$\cdot 0,2132 =$	10660, 00	39340,00
2	10660, 00	$\cdot 0,2132 =$	2272, 71	8387,29
3	2272, 71	$\cdot 0,2132 =$	484, 54	1788,17
4	484, 54	$\cdot 0,2132 =$	103, 30	381,24
5	103, 30	$\cdot 0,2132 =$	22, 02	81,28
6	22, 02	$\cdot 0,2132 =$	4, 70	17,32
7	4, 70	$\cdot 0,2132 =$	1, 00	3,70

Aufgaben Index