Name Matr.-Nr. Platznr.

Aufgabe 1 *(20 Punkte)*

Die Aufgabe besteht aus sieben Behautungen gegliedert in vier Unteraufgaben.

Überprüfen Sie die folgenden Behauptungen indem Sie eindeutig RICHTIG oder FALSCH zuordnen. Begründen Sie die Antworten! (Für unbegründete korrekte Antworten gibt es keine Punkte)

1. Die Schätzfunktion $\hat{b} = \frac{2}{n}\sum_{i=1}^{n} x_i = 2\overline{x}$ ist ein erwartungstreuer Schätzer für den Parameter b einer stetigen Gleichverteilung auf dem Intervall $[0, b]$, deren Verteilungsfunktion durch
$$F(x) = \begin{cases} 0 & fuer\, x < 0 \\ \frac{x}{b} & fuer\, 0 \leq x \leq b \\ 1 & x > b \end{cases}$$
gegeben ist.

Lösung: RICHTIG.
Erwartungswert einer stetigen Gleichverteilung mit $a = 0$:
$$E(X) = \frac{1}{2}(a+b) = \frac{1}{2}b$$
Beweis der Erwartungstreue:
$$E\left(\hat{b}\right) = E(2 \cdot \bar{x}) = 2\,E(\bar{x}) = 2\frac{1}{2}b = b$$

2. Die Schätzfunktion $\hat{b} = max\{x_1, \ldots, x_n\}$ ist der Maximum–Likelihood–Schätzer für den Parameter b einer stetigen Gleichverteilung auf dem Intervall $[0, b]$.

Lösung: RICHTIG.
Die Likelihood ist also immer Null, wenn eine Beobachtung nicht zwischen 0 und b liegt. Damit muß gelten: $\hat{b} \leq max\{x_1, \ldots x_n\}$. Die Likelihood ist dann maximal, wenn $\hat{b} = max\{x_1, \ldots x_n\}$ ist, da die Likelihoodfunktion proportional zu $\frac{1}{b}$ ist.

3. Das Prognose–Intervall für die Anzahl der Erfolge bei n unabhängigen Wiederholungen eines Versuchs wird umso breiter, je größer n wird.

Lösung: JA
Es gilt:

$$n\,\pi - \tau\sqrt{n\,\pi\,(1-\pi)} \leq y \leq n\,\pi - \tau\sqrt{n\,\pi\,(1-\pi)}$$

Je größer n ist, desto größer ist $n\,\pi$ und auch $\tau\sqrt{n\pi\,(1-\pi)}$

4. Sei $\hat{\Theta}_1$ ein erwartungstreuer ML-Schätzer für den Parameter θ_1. Dann ist auch

   (a) $\hat{\Theta}_2 = 5 \cdot \hat{\Theta}_1$ eine erwartungstreue Schätzfunktion für den Parameter $\theta_2 = 5 \cdot \theta_1$;
   Lösung: RICHTIG, weil

   $$E\left(\Theta_2\right) = E\left(5 \cdot \Theta_1\right) = 5 \cdot E\left(\Theta_1\right) = 5 \cdot \theta_1 = \theta_2$$

   (b) $\hat{\Theta}_3 = \hat{\Theta}_2 - 4$ eine erwartungstreue Schätzfunktion für den Parameter $\theta_3 = 5 \cdot \theta_1 - 4$;
   Lösung: RICHTIG, weil

   $$E\left(\Theta_3\right) = E\left(\Theta_2 - 4\right) = E\left(\Theta_2\right) - E\left(4\right) = 5 \cdot \theta_1 - 4 = \theta_3$$

   (c) $\hat{\Theta}_4 = \hat{\Theta}_1^2$ eine erwartungstreue Schätzfunktion für den Parameter $\theta_4 = \theta_1^2$.
   Lösung: FALSCH, weil

   $$\begin{aligned} Var\left(\Theta_1\right) &= E\left(\Theta_1^2\right) - \left(E\left(\Theta_1\right)\right)^2 \\ E\left(\Theta_1^2\right) &= Var\left(\Theta_1\right) + \left(E\left(\Theta_1\right)\right)^2 \\ &= Var\left(\Theta_1\right) + \theta_1^2 \\ E\left(\Theta_1^2\right) &\geq \theta_1^2 \\ E\left(\Theta_4\right) &= E\left(\Theta_1^2\right) \geq \theta_1^2 \\ E\left(\Theta_4\right) &\geq \theta_4 \end{aligned}$$

   (d) Alle drei Schätzfunktionen $\hat{\Theta}_2$, $\hat{\Theta}_3$ oder $\hat{\Theta}_4$ sind asymptotisch erwartungstreu.

   Lösung: RICHTIG.
   Alle drei Schätzfunktionen sind asymptotisch erwartungstreu. Die Funktionen $\hat{\Theta}_2$ und $\hat{\Theta}_3$ sind erwartungstreu und damit auch asymptotisch erwartungstreu. Die Funktion $\hat{\Theta}_4$ ist asymptotisch erwartungstreu (unter gewissen Regularitätsbedingungen), weil sie eine ML-Schätzfunktion ist.

Aufgabe 2 (20 Punkte)

Die Aufgabe besteht aus zwei Teilaufgaben.

Ein Monitorhersteller produziert zwei Modelle, den Standardmonitor *Plusminus* und den High End Monitor *Frontal*. Das Modell *Plusminus* macht achzig Prozent der verkauften Stückzahl aus.

Seine Geräte vertreibt der Hersteller über das Internet. Er versendet siebzig Prozent der Monitore mit dem Paketdienst *Schnell&Billig*. Der Rest wird mit dem Paketdienst *Lamm&Sanft* verschickt. Die Wahl des Paketdienstes ist dabei unabhängig vom Monitormodell.

Insgesamt wird jeder zehnte versandte Monitor von Kunden beanstandet, wobei neun von zehn Reklamationen bei mit *Schnell&Billig* versandten Monitoren registiert werden. Die Reklamationen können alleine auf den Versand zurückgeführt werden, da ein Fünftel der defekten Monitore aus der Modellreihe *Frontal* stammen.

1. Wie groß ist die Wahrscheinlichkeit, daß ein Monitor, der mit *Schnell&Billig* verschickt wird, nicht vom Kunden beanstandet wird?

   Lösung: Gegeben:

$$\begin{array}{lclcl}
P\left(\text{Monitor wird beanstandet}\right) & = & P\left(A\right) & = & 0.1 \\
P\left(\text{Monitor wird nicht beanstandet}\right) & = & P\left(\bar{A}\right) & = & 0.9 \\
P\left(\text{Schnell \& Billig} \mid \text{Monitor wird beanstandet}\right) & = & P\left(S|A\right) & = & 0.9 \\
P\left(\text{Frontal} \mid \text{Monitor wird beanstandet}\right) & = & P\left(\overline{PM}|A\right) & = & 0.2 \\
P\left(\text{Plusminus}\right) & = & P\left(PM\right) & = & 0.8 \\
P\left(\text{Frontal}\right) & = & P\left(\overline{PM}\right) & = & 0.2 \\
P\left(\text{Schnell \& Billig}\right) & = & P\left(S\right) & = & 0.7 \\
P\left(\text{Lamm \& Soft}\right) & = & P\left(\bar{S}\right) & = & 0.3
\end{array}$$

Gesucht: $P\left(\bar{A}|S\right)$

1. Möglichkeit:

Satz der totalen Wahrscheinlichkeit:

$$\begin{aligned}
P\left(\bar{A}|S\right) &= \frac{P\left(S|\bar{A}\right)\cdot P\left(\bar{A}\right)}{P\left(S\right)} \\
P\left(S\right) &= P\left(S|A\right)\cdot P\left(A\right) + P\left(S|\bar{A}\right)\cdot P\left(\bar{A}\right) \\
P\left(S|\bar{A}\right) &= \frac{P\left(S\right) - P\left(S|A\right)\cdot P\left(A\right)}{P\left(\bar{A}\right)} \\
&= \frac{0.7 - 0.9\cdot 0.1}{0.9} \\
&= 0.678 \\
P\left(\bar{A}|S\right) &= \frac{0.678\cdot 0.9}{0.7} \\
P\left(\bar{A}|S\right) &= \underline{\underline{0.871}}
\end{aligned}$$

2. Möglichkeit:

$$\begin{aligned} P\left(\bar{A}|S\right) &= 1 - P\left(A|S\right) \\ &= 1 - \tfrac{P(S|A)\cdot P(A)}{P(S)} \\ &= 1 - \tfrac{0.9\cdot 0.1}{0.7} \\ &= \underline{\underline{0.871}} \end{aligned}$$

Die Wahrscheinlichkeit beträgt 0.871.

2. Wie groß ist die Wahrscheinlichkeit, daß ein beanstandeter Monitor aus der Modellreihe *Plusminus* stammt und mit *Schnell&Billig* versandt wurde?

   Lösung: gesucht: $P\left(PM \bigcap S|A\right)$
   Da eine Unabhängigkeit zwischen Modellreihe und Paketdienst besteht, gilt:

$$\begin{aligned} P\left(PM \bigcap S|A\right) &= P\left(PM|A\right)\cdot P\left(S|A\right) \\ &= \left(1 - P\left(\overline{PM}|A\right)\right)\cdot P\left(S|A\right) \\ &= (1-0.2)\cdot 0.9 \\ &= 0.72 \end{aligned}$$

Die Wahrscheinlichkeit beträgt 0.72.

Aufgabe 3 (30 Punkte)

Die Aufgabe besteht aus zwei Teilaufgaben mit insgesamt fünf Unteraufgaben.

In einem Unternehmen werden Geschäftsergebnisse mittels Standarddokumenten von jeweils 10 000 Zeichen von den Filialen zur Zentrale gesendet. Zur Übertragung werde bisher eine Methode A(lt) verwandt. Die Unternehmensleitung überlegt nun, ob sie ein neues Übertragungsverfahren, Methode N, einführen soll, wobei sich an den Standarddokumenten nichts ändert.

1. Bei der Übertragung mit Methode A ist das Auftreten eines Fehlers für jedes Zeichen unabhängig und identisch verteilt mit einer Fehlerwahrscheinlichkeit von einem Prozent.

   (a) Wie ist die Zufallsvariable X 'Anzahl der Übertragungsfehler pro Dokument' verteilt?
   Lösung: **1. Möglichkeit:** Binomialverteilung

   $$X \sim B(n;\, \pi) \qquad \text{mit}\, n = 10\,000;\ \pi = 0.01$$

   2. Möglichkeit: Poissonveteilung

   $$X \sim PV(\lambda) \qquad \text{mit}\, \lambda = 100$$

   (b) Wie groß ist der Erwartungswert und die Varianz von X?
   Lösung: **1. Möglichkeit:** Binomialverteilung

   $$\begin{aligned} E(X) &= n\,\pi = 10\,000 \cdot 0.01 \\ &= 100 \\ Var(X) &= n\,\pi\,(1-\pi) = 100 \cdot 0.99 \\ &= 99 \end{aligned}$$

   2. Möglichkeit: Poissonverteilung

   $$\begin{aligned} E(X) &= \lambda = n\,\pi = 10\,000 \cdot 0.01 \\ &= 100 \\ Var(X) &= \lambda = n\,\pi \\ &= 100 \end{aligned}$$

   Der Erwartungwert liegt bei 100 und die Varianz ist 99.

   (c) Wie groß ist die Wahrscheinlichkeit, daß ein Dokument fehlerfrei übertragen wird? Runden Sie dabei auf die erste Stelle, die nicht Null ist!
   Lösung: **1. Möglichkeit:** Binomialverteilung

   $$\begin{aligned} P(X=k) &= \begin{pmatrix} n \\ k \end{pmatrix} \cdot \pi^k \cdot (1-\pi)^{n-k} \\ P(X=0) &= \begin{pmatrix} 10\,000 \\ 0 \end{pmatrix} \cdot 0.01^0 \cdot 0.99^{10\,000} \\ &= 1 \cdot 1 \cdot 2 \cdot 10^{-44} \\ &= 2 \cdot 10^{-44} \end{aligned}$$

Die Wahrscheinlichkeit beträgt $2 \cdot 10^{-44}$.
2. Möglichkeit: Poissonverteilung

$$\begin{aligned} P(X = k) &= \tfrac{\lambda^k}{k!} e^{-\lambda} \\ P(X = 0) &= \tfrac{100^0}{0!} e^{-100} \\ &= 3.7 \cdot 10^{-44} \end{aligned}$$

Die Wahrscheinlichkeit beträgt $4 \cdot 10^{-44}$.

2. Es soll eine neue Übertragungstechnik N(eu) eingeführt werden, wobei die Zeichenanzahl in den Standarddokumente gleich bleibt. Vorher soll jedoch sichergestellt werden, daß in Bezug auf die Anzahl der Übertragungsfehler eine Verbesserung erzielt wird. Dazu werden versuchsweise zehn Dokumente mit der neuen Übertragungstechnik verarbeitet und die Anzahl der Übertragungsfehler bestimmt:

$$91 \quad 97 \quad 96 \quad 101 \quad 90 \quad 88 \quad 95 \quad 94 \quad 102 \quad 91$$

   (a) Berechnen Sie ein zweiseitiges approximatives 95%–Konfidenzintervall für die erwartete Anzahl von Übertragungsfehlern mittels der neuen Übertragungstechnik. Formulieren Sie die Aussage des Konfidenzintervalls.

   Lösung: **1. Möglichkeit:** Approximation der Binomialverteilung durch die Normalverteilung mit

$$N(n\pi;\, n\pi(1-\pi)) = N(100;\, 99)$$

   Berechnung von $\bar{X}$:

$$\begin{aligned} \bar{x} &= \tfrac{1}{n}\textstyle\sum_{i=1}^{n} x_i \\ &= \tfrac{1}{10}(91+97+96+101+90+88+95+94+102+91) \\ &= 94.5 \end{aligned}$$

   Berechnung von $\hat{\sigma}$:

$$\sigma = \mu(1-\pi) = 94.5 \cdot 0.99 = 93.555$$

   Bestimmung von τ:

$$\tau_{1-\frac{\alpha}{2}} = \tau_{1-\frac{0.05}{2}} = \tau_{0.975} = 1.96$$

   Bestimmung des Konfidenzintervalls:

$$\begin{aligned} \bar{x} - \tau_{1-\frac{\alpha}{2}}\tfrac{\sigma}{\sqrt{n}} \leq\ & \mu \leq \bar{x} + \tau_{\frac{\alpha}{2}}\tfrac{\sigma}{\sqrt{n}} \\ 94.5 - 1.96 \cdot \tfrac{\sqrt{93.555}}{\sqrt{10}} \leq\ & \mu \leq 94.5 + 1.96 \cdot \tfrac{\sqrt{93.555}}{\sqrt{10}} \\ 94.5 - 1.96 \cdot 3.059 \leq\ & \mu \leq 94.5 + 1.96 \cdot 3.059 \\ 88.500 \leq\ & \mu \leq 100.495 \end{aligned}$$

   95% der Werte liegen im Intervall [88; 101].

   2. Möglichkeit: Approximation der Poissonverteilung durch die Normalverteilung mit $N(\lambda;\, \lambda)$

   Berechnung von $\bar{x}$: siehe 1. Möglichkeit

   Berechnung von σ:

$$\sigma = 94.5$$

$$\begin{aligned} \bar{x} - \tau_{1-\frac{\alpha}{2}}\tfrac{\sigma}{\sqrt{n}} \leq\ & \mu \leq \bar{x} + \tau_{\frac{\alpha}{2}}\tfrac{\sigma}{\sqrt{n}} \\ 94.5 - 1.96\sqrt{\tfrac{94.5}{10}} \leq\ & \mu \leq 94.5 + 1.96\sqrt{\tfrac{94.5}{10}} \\ 94.5 - 1.96 \cdot 3.074 \leq\ & \mu \leq 94.5 + 1.96 \cdot 3.074 \\ 88.475 \leq\ & \mu \leq 100.525 \end{aligned}$$

   Mit einer Sicherheit von mindestens 95% liegt μ im Intervall [88; 101].

   (b) Testen Sie zum 5 Prozent Niveau, ob die neue Übertragungsmethode in Bezug auf Übertragungsfehler besser ist.

i. Wie lautet die Hypothesen?
Lösung: Die Hypothesen lauten:

$$\begin{aligned} H_0 &: \mu \geq \mu_0 \\ H_1 &: \mu < \mu_0 \qquad \text{(Das neue Verfahren ist besser.)} \end{aligned}$$

ii. Geben Sie allgemein die Prüfgröße an.
Lösung: $T_{PG} = \frac{\sqrt{n}\left(\bar{X}-\mu_0\right)}{\sigma_0}$oder $T_{PG} = \overline{X}$

iii. Wie ist die Prüfgröße verteilt?
Lösung: Da $X \sim B\left(n;\, \pi\right)$ verteilt ist und duch die Normalverteilung $NV\left(n\pi;\, n\pi\left(1-\pi\right)\right)$ approximiert werden kann, ist die Prüfgröße standardnormalverteilt. Oder normalverteilt mit Parametern $\mu = n\pi \, und \, \sigma^2 = \frac{\pi(1-\pi)}{n}$.

iv. Berechnen Sie die Prüfgröße.
Lösung: Gegeben aus Aufgabenteil a):

$$\begin{aligned} n &= 10 \\ \bar{X} &= 94.5 \\ \mu_0 &= 100 \\ \sigma_0^2 &= 99 \\ \tau_{0.95} &= 1.645 \end{aligned}$$

Berechnung der Prüfgröße:

$$T_{PG} = \frac{\sqrt{10}\,(94.5-100)}{\sqrt{99}} = -1.748$$

v. Gegen Sie den Ablehnbereich an.
Lösung: Ablehnbereich: $\{T < -1.645\}$

vi. Wie lautet die Testentscheidung?

Lösung: Die Prüfgröße liegt im Ablehnbereich. Aus diesem Grunde kann die Nullhypothese verworfen werden, d.h. die Hypothese, daß das neue Verfahren besser ist, wird angenommen.

Aufgabe 4 *(30 Punkte)*

Die Aufgabe besteht aus vier Teilaufgaben.

Im Oktober 2000 wurden an dreißig Tagen folgende Temperaturen in einem Ort im Sauerland gemessen (in $C°$):

-0.9	1.6	6.6	1.6	1.1	6.6	1.6	3.6	0.1	10.1
7.6	8.6	9.6	7.6	-0.9	8.1	-0.4	11.1	0.1	-0.9
10.6	7.1	-0.4	0.1	0.1	7.6	0.1	-0.4	-0.4	-1.3

Dabei spielt die Reihenfolge keine Rolle, da die Temperaturen auf einzelne Zettel notiert wurden, die leider durcheinander geraten sind.

Aus langjährigen Messung ist bekannt, daß die Standardabweichung der Tagestemperaturen im Oktober $4.3\,C°$ beträgt.

1. Stellen Sie mittels eines Histogramms die gruppierten Daten dar. Erstellen Sie dazu zuerst eine geeignete Tabelle. Wählen Sie dazu eine der beiden Klasseneinteilungen aus und begründen Sie Ihre Wahl.

	1. Einteilung	2. Einteilung
1	$-20 < x \leq -5$	$x \leq -1$
2	$-5 < x \leq 0$	$-1 < x \leq 1$
3	$0 < x \leq 5$	$1 < x \leq 3$
4	$5 < x \leq 10$	$3 < x \leq 5$
5	$10 < x \leq 15$	$5 < x \leq 7$
6	$15 < x \leq 30$	$7 < x$

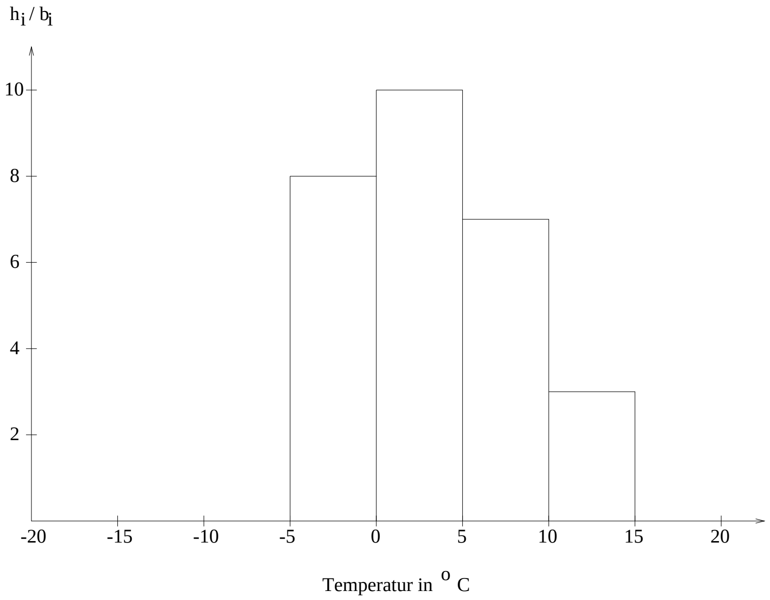
Lösung: Da bei der 2. Einteilung offene Klassen existieren und in diesen auch Werte liegen, wird es bei der Darstellung der Werte im sowohl im Histogramm als auch bei der Verteilungsfunktion Schwierigkeiten erwartet. Aus diesem Grunde wurde die 1. Einteilung gewählt.

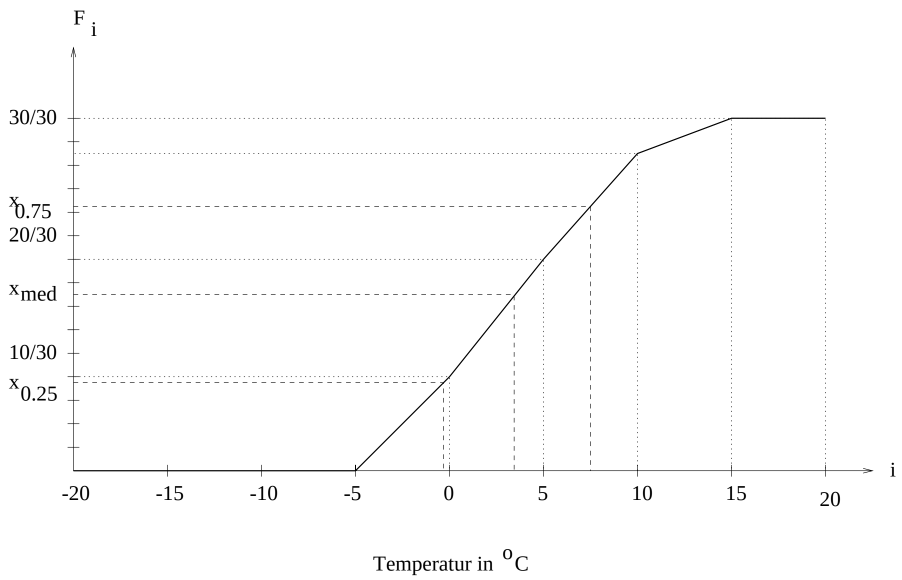
Klasse i	Grenzen	h_i	b_i	h_i/b_i	f_i	F_i
1	$-20 < x \leq -5$	0	3	0	0	0
2	$-5 < x \leq 0$	8	1	8	$\frac{8}{30}$	$\frac{8}{30}$
3	$0 < x \leq 5$	10	1	10	$\frac{10}{30}$	$\frac{18}{30}$
4	$5 < x \leq 10$	9	1	9	$\frac{9}{30}$	$\frac{27}{30}$
5	$10 < x \leq 15$	3	1	3	$\frac{3}{30}$	$\frac{30}{30}$
6	$15 < x \leq 20$	0	1	0	0	$\frac{30}{30}$

Das Histogramm ist in Abbildung 1 zu sehen.

2. Stellen Sie mittels einer Verteilungsfunktion die gruppierten Daten dar. Erweitern Sie dazu die Tabelle aus 1 geeignet.

Lösung: Die Werte für die Funktion sind in der Tabelle in Aufgabe 1 enthalten. Die Verteilungsfunktion ist in Abbildung 2 zu sehen.


h_i / b_i
10
8
6
4
2
-20
-15
-10
-5
0
5
10
15
20
Temperatur in ^{o} C


F_i
30/30
$x_{0.75}$
20/30
x_{med}
10/30
$x_{0.25}$
i
-20
-15
-10
-5
0
5
10
15
20
Temperatur in ^{o}C

3. Bestimmen Sie graphisch einen Lageparameter sowie einen Streuungsparameter für die Temperaturdaten. Verwenden Sie dazu die Verteilungsfunktion aus dem 2. Aufgabenteil.

   Lösung: Lageparameter: $x_{med} = 3.5°C$
   Streuungsparameter: $IQR = x_{0.75} - x_{0.25} = 7.5°C - (-0.5°C) = 8°C$

4. Berechnen Sie basierend auf den 30 Messungen und der bekannten Varianz ein 80%–Konfidenzintervall für die Tagestemperatur im Oktober in diesem Ort im Sauerland. Einzige zulässige Annahme ist, daß die Tagestemperaturen unabhängig identisch verteilt sind.

   Lösung: Die Abschätzung erfolgt nach Tschebyscheff:

$$\begin{aligned}
\sigma &= 4.3 \\
\bar{x} &= \tfrac{1}{n}\textstyle\sum_{i=1}^{n} x_i \\
\text{Gegeb} \quad &= \tfrac{1}{30}\left(3\,(-0.9) + 3\,(1.6) + 2\,(6.6) + 1.1 + 3.6 + 5\,(0.1) + 10.1 + 3\,(7.6) + 8.6 + 9.6 + 8.1 - \right. \\
&= \tfrac{1}{30} \cdot 105.6 = 3.52 \\
\mu &= \bar{x}
\end{aligned}$$

Berechnung von k:

$$\begin{aligned}
P\left\{|X - \mu| < k\sigma\right\} &\geq 1 - \tfrac{1}{k^2} \\
1 - \tfrac{1}{k^2} &= 0.8 \\
k &= \sqrt{5}
\end{aligned}$$

Konfidenzintervall:

$$[\mu - k\sigma;\ \mu - k\sigma] = \left[3.52 - \sqrt{5} \cdot 4.3;\ 3.52 + \sqrt{5} \cdot 4.3\right] = [-6.1;\ 13.1]$$

Name Matr.-Nr. Platznr.

Zu Aufgabe:

Name Matr.-Nr. Platznr.

Zu Aufgabe:

Name Matr.-Nr. Platznr.

Zu Aufgabe: