Lösung zur Selbstkontrollarbeit 1

Vertiefung der Wirtschaftsmathematik und Statistik (Teil Statistik)

10. Januar 2019

Aufgaben

Aufgabe 1

1.1

$$
\begin{aligned}
E(X) &= \sum_{x=0}^{N} \binom{N}{x} \cdot \pi^x \cdot (1-\pi)^{N-x} \cdot x \\
&= \sum_{x=1}^{N} \frac{N!}{x!(N-x)!} \cdot x \cdot \pi^x \cdot (1-\pi)^{N-x} \\
&= N\pi \sum_{x=1}^{N} \frac{(N-1)!}{(x-1)!(N-x)!} \cdot \pi^{x-1} \cdot (1-\pi)^{(N-1)-(x-1)} \\
&= N\pi \sum_{x=1}^{N} \binom{N-1}{x-1} \cdot \pi^{x-1} \cdot (1-\pi)^{(N-1)-(x-1)} \\
&= N\pi \sum_{y=0}^{N-1} \binom{N-1}{y} \cdot \pi^{y} \cdot (1-\pi)^{N-1-y} \\
&= N\pi.
\end{aligned}
$$

1.2

Zugrundegelegt werden N unabhängig bernoulliverteilte Zufallsvariablen X_n, $n = 1, ..., N$, wobei mit der Wahrscheinlichkeit π das Ereignis 1 und mit Wahrscheinlichkeit $1 - \pi$ das Ereignis 0 eintritt. Die binomialverteilte Zufallsvariable X ergibt sich durch Addition der einzelnen unabhängigen bernoulliverteilten Zufallsvariablen X_n, d.h. $X = \sum_{n=1}^{N} X_n$. Somit gilt:

$$
\begin{aligned}
E(X_n) &= \sum x_{n_i} f_{X_n}(x_{n_i}) = 1 \cdot \pi + 0 \cdot (1-\pi) = \pi \\
E(X) &= \sum_{n=1}^{N} E(X_n) = N\pi \\
\mathrm{Var}(X_n) &= \sum x_{n_i}^2 f_{X_n}(x_{n_i}) - [E(X_n)]^2 \\
&= 0^2 \cdot (1-\pi) + 1^2 \cdot \pi - \pi^2 = \pi - \pi^2 = \pi(1-\pi) \\
\mathrm{Var}(X) &= \sum_{n=1}^{N} \mathrm{Var}(X_n) = N\pi(1-\pi)
\end{aligned}
$$

Aufgabe 2

$X_1 \sim N(100, 156.25)$ für Test 1,
$X_2 \sim N(120, 204.75)$ für Test 2,
$X_3 \sim N(\mu_3, 243.36)$ für Test 3.

Testperson T_1 erzielt im Test 1 eine Punktzahl von 75, im Test 2 eine Punktzahl von 150 Punkten und im Test 3 eine Punktzahl von 110 Punkten. Es ist von der Unabhängigkeit der Testergebnisse auszugehen. Verwenden Sie bei der Bearbeitung dieser Aufgabe die Verteilungstabelle.

2.1

$$\begin{aligned} P(X_1 \geq 75) &= 1 - P(X_1 \leq 75) = 1 - P(Z \leq \tfrac{75-100}{12.5}) \\ &= 1 - P(Z \leq -2) = 1 - 1 + P(Z \leq 2) = 0.9772 \end{aligned}$$

2.2

$X = X_1 + X_2 \sim N(220, 361)$

$$P(X \leq 225) = P(Z \leq \tfrac{225-220}{19}) = P(Z \leq 0.26) = 0.6026$$

2.3

$$\begin{aligned} & P(X_3 \leq 110) &=& \quad P(Z \leq \tfrac{110-\mu_3}{15.6}) = 0.209 \\ \iff & \tfrac{110-\mu_3}{15.6} &=& \quad = z(0.209) = -z(0.791) = -0.81 \\ \iff & \mu_3 &=& \quad 122.636 \end{aligned}$$

2.4

$$X_1 + X_2 + X_3 \sim N(342.636, 604.36)$$

Aufgabe 3

Für unabhängig standardnormalverteilte Zufallsvariablen X_n gilt, dass $\sum_{n=1}^{N} X_n^2$ χ^2-verteilt ist mit N Freiheitsgraden.

3.1

$P(X \leq 13.12) = 0.025$

3.2

$P(4.66 \leq X \leq 7.79) = F(7.79) - F(4.66) = 0.1 - 0.01 = 0.09$

3.3

Für $N > 30$ kann durch die Normalverteilung approximiert werden.

$$\begin{aligned} P(24.5 \leq X \leq 40.5) &= P(\sqrt{49} - \sqrt{2 \cdot 41 - 1} \leq Z \leq \sqrt{81} - \sqrt{2 \cdot 41 - 1} \\ &= P(-2 \leq Z \leq 0) = F(0) - F(-2) = \Phi(0) - \Phi(-2) \\ &= 0.5 - 0.0228 = 0.4772 \end{aligned}$$

3.4

$P(X \leq -7) = 0$, da χ^2 nur im positiven Bereich definiert ist.

Aufgabe 4

4.1

$P(2.06 \leq X \leq 3.45) = F(3.45) - F(2.06) = 0.999 - 0.975 = 0.024$

4.2

$P(X \leq 1.708) = F(1.708) = 0.95$

4.3

$P(1.316 \geq X) = 1 - F(1.316) = 1 - 0.9 = 0.1$

4.4

$P(-0.684 \leq X \leq 0.684) = 0.5.$

Aufgabe 5

5.1

$$\begin{aligned} P(X \leq 2.54) &= 0.95 \\ f(0.05, 15, 10) &= \frac{1}{f(0.95, 10, 15)} = \frac{1}{2.54} = 0.394 \end{aligned}$$

5.2

Ist die Zufallsvariable X nach t-verteilt mit 20 Freiheitsgraden, so ist X^2 nach F-verteilt mit 1 und 20 Freiheitsgraden.

$P(X^2 \geq 8.1) = 1 - P(X^2 \leq 8.1) = 0.01$

5.3

Die Zufallsvariable $\frac{X_5^2}{\sum_{n=1}^{5} X_n/5}$ ist F-verteilt mit 1 und 5 Freiheitsgraden.

$P(\frac{X_5^2}{\sum_{n=1}^{5} X_n/5} \leq 6.61) = 0.95$

Aufgabe 6

6.1

Die Aussage ist falsch. Eine effiziente Schätzfunktion ist stets erwartungstreu; aus der Erwartungstreue folgt aber nicht notwendigerweise die Effizienz.

6.2

Die Schätzfunktion X_u ist erwartungstreu, während X_g verzerrt ist.

$$\begin{aligned} E(X_g) &= \frac{1}{2N} \sum_{n=1}^{N/2} E(X_{2n}) = \frac{1}{2N} \cdot \frac{N}{2} \mu = \frac{\mu}{4} \\ E(X_u) &= \frac{2}{N-1} \sum_{n=1}^{(N-1)/2} E(X_{2n}) = \frac{2}{N-1} \cdot \frac{(N-1)\mu}{2} = \mu \end{aligned}$$

6.3

Gegeben ist eine normalverteilte Zufallsvariable X mit Erwartungswert μ und Varianz σ^2. Für eine lineare Transformation der Form $Y = aX + b$ ist Y eine normalverteilte Zufallsvariable mit $E(Y) = aE(X) + b$ und $\mathrm{Var}(Y) = a^2\mathrm{Var}(X)$. Somit gilt hier für $Y = \frac{2}{a}X + 3b^2$:

$$\begin{aligned} E(Y) &= \frac{2}{a}E(X) + 3b^2 = \frac{2}{a}\mu + 3b^2 \\ \mathrm{Var}(Y) &= \frac{4}{a^2}\mathrm{Var}(X) = \frac{4}{a^2}\sigma^2 \end{aligned}$$

6.4

Es ist $\sigma^2 = \mathrm{Var}(X_n) = E(X_n^2) - [E(X_n)]^2 = E(X_n^2) - \mu^2$, d.h. $E(X_n^2) = \sigma^2 + \mu^2$. Weiter ist $\mathrm{Var}(\overline{X}) = \frac{\sigma^2}{N}$ Mit diesen Beziehungen ergibt sich:

$$\begin{aligned} E(S^2) &= E\left(\frac{1}{N-1}\sum(X_n - \overline{X})^2\right) \\ &= E\left(\frac{1}{N-1}\sum X_n^2 - \frac{2}{N-1}\overline{X}\sum X_n + \frac{1}{N-1}\sum \overline{X}^2\right) \\ &= E\left(\frac{1}{N-1}\sum X_n^2 - \frac{2}{N-1}\overline{X}\cdot N\overline{X} + \frac{1}{N-1}N\overline{X}^2\right) \\ &= \frac{1}{N-1}\left[\sum E(X_n^2) - NE(\overline{X}^2)\right] \\ &= \frac{1}{N-1}\left[N(\sigma^2 + \mu^2) - N(\sigma_{\overline{X}}^2 + \mu^2)\right] \\ &= \frac{1}{N-1}(N\sigma^2 + N\mu^2 - N\frac{\sigma^2}{N} - N\mu^2) \\ &= \frac{1}{N-1}(N-1)\sigma^2 = \sigma^2 \end{aligned}$$

Aufgabe 7

7.1

Die Erwartungstreue von $\overline{X}$ ist klar. Die minimale Varianz wird nachgewiesen, in dem die Varianz der allgemeinen linearen Schätzfunktion $\hat{X} = \sum_n a_n X_n$ mit $\sum_n a_n = 1$ mittels der Lagrange-Funktion minimiert wird.

Es gilt $\text{Var}(\hat{X}) = \sigma^2 \sum_n a_n^2$. Somit wird $\text{Var}(\hat{X})$ minimiert, indem $\sum_n a_n^2$ unter der Nebenbedingung $\sum_n a_n = 1$ minimiert wird. Zugrundegelegt wird die Lagrange-Funktion $L = \sum_n a_n^2 - \lambda(\sum a_n - 1)$. Es liegen somit folgende $N+1$ Gleichungen vor:

$$
\begin{aligned}
\frac{\partial L}{\partial a_n} &= 2a_n - \lambda = 0 \quad \text{für } n = 1, ..., N \\
\frac{\partial L}{\partial \lambda} &= -\sum_n a_n + 1 = 0
\end{aligned}
$$

Aufsummieren der ersten N Gleichungen ergibt

$$2\sum_n a_n - N\lambda = 0 \Longleftrightarrow \lambda = \frac{2}{N}$$

d.h.

$$2a_n - \lambda = 0 \Longleftrightarrow 2a_n - \frac{2}{N} = 0 \Longleftrightarrow a_n = \frac{1}{N}$$

Somit ist $\text{Var}(\hat{X}) = \sigma^2 \sum_n a_n^2$ minimal für $a_n = \frac{1}{N}$, so dass hier der Schätzer mit der minimalen Varianz $\hat{X} = \frac{1}{N}\sum_n X_n = \overline{X}$ lautet.

7.2

Gezeigt werden muss lediglich die Erwartungstreue der Schätzfunktion S_*^2. Die Erwartungstreue von S^2 wurde bereits in der vorherigen Aufgabe gezeigt.

$$
\begin{aligned}
E(S_*^2) &= \frac{1}{N}\sum_{n=1}^{N} E(X_n - \mu)^2 \\
&= \frac{1}{N}\sum_{n=1}^{N}(E(X_n^2) - 2\mu E(X_n) + \mu^2) \\
&= \frac{1}{N}(N\sigma^2 + N\mu^2 - 2N\mu^2 + N\mu^2) \\
&= \frac{1}{N}N\sigma^2 = \sigma^2
\end{aligned}
$$

Aufgrund der Normalverteilungsannahme gilt, dass die Statistik $\frac{(N-1)S^2}{\sigma^2}$ χ^2-verteilt ist mit $(N-1)$ Freiheitsgraden. Die Statistik $\frac{NS_*^2}{\sigma^2}$ ist dagegen χ^2-verteilt mit N Freiheitsgraden. Somit gilt

$$\begin{aligned}
\mathrm{Var}\left(\frac{(N-1)S^2}{\sigma^2}\right) &= 2(N-1) \\
\Leftrightarrow \quad \frac{(N-1)^2}{\sigma^4}\mathrm{Var}(S^2) &= 2(N-1) \\
\Leftrightarrow \quad \mathrm{Var}(S^2) &= \frac{2\sigma^4}{(N-1)}
\end{aligned}$$

$$\begin{aligned}
\mathrm{Var}\left(\frac{NS^{*2}}{\sigma^2}\right) &= 2N \\
\Leftrightarrow \quad \frac{N^2}{\sigma^4}\mathrm{Var}(S^{*2}) &= 2N \\
\Leftrightarrow \quad \mathrm{Var}(S^{*2}) &= \frac{2\sigma^4}{N}
\end{aligned}$$

Der Schätzer S_*^2 ist aufgrund der geringeren Varianz effizienter (wirksamer) als S^2.

Aufgabe 8

Tschebyscheff-Ungleichung $P(|X-\mu| \geq c\sigma) \leq \frac{1}{c^2}$

$\sigma = 0.1 \qquad c\sigma = 1 \qquad 0.1c = 1$

$c = 10 \qquad c^2 = 100 \qquad \frac{1}{c^2} = \frac{1}{100} = 0.01$

$\mu = 100$

$P(|X - 100\ mm| \geq 1\ mm) \leq 0.01$

d.h. die Wahrscheinlichkeit, dass die Wellenlänge größer als 101 mm oder kleiner als 99 mm ist (Wahrscheinlichkeit für den Ausschuss), ist kleiner als 0.01, d.h. der Ausschussanteil beträgt höchstens 1%.

Aufgabe 9

9.1

Die unbekannten Parameter können oft als Funktion von Momenten von X dargestellt werden. Werden die Momente durch die Stichprobenmomente ersetzt, ergeben sich die sogenannten Momentenschätzer als Lösungen eines Gleichungssystems. Bekannte Momentenschätzer sind die Schiefe und die Kurtosis.

9.2

Es ist $\mu_1 = E(X_n)$ und $\mu_2 = E(X_n^2) = \mathrm{Var}(X_n) + [E(X_n)]^2 = \sigma^2 + \mu^2$, so dass sich die Parmeter μ und σ^2 mittels der Momente darstellen lassen.

$$\mu = \mu_1 \qquad \sigma^2 = \mu_2 - \mu^2 = \mu_2 - \mu_1^2$$

Einsetzen der Stichprobenmomente $\hat{\mu}_1 = \overline{X}$ und $\hat{\mu}_2 = \frac{1}{N}\sum_n X_n^2$ ergibt die gesuchten Momentenschätzer.

$$\hat{\mu} = \overline{X} \qquad \widehat{\sigma^2} = \frac{1}{N}\sum_n X_n^2 - \overline{X}^2 = \frac{1}{N}\sum_n (X_n - \overline{X})^2$$

9.3

$$\begin{aligned}
\mu_1 &= \tfrac{a+b}{2} \\
\Longleftrightarrow a &= 2\mu_1 - b \\
\mu_2 &= E(X^2) = \mathrm{Var}(X) + [E(X)]^2 \\
&= \tfrac{(b-a)^2}{12} + \tfrac{(a+b)^2}{4} \\
&= \tfrac{b^2+ab+a^2}{3} \\
\Longleftrightarrow \mu_2 &= \tfrac{b^2-2\mu_1 b+4\mu_1^2}{3}
\end{aligned}$$

Mit Hilfe der pq-Formel und $a = 2\mu_1 - b$ ergibt sich:

$$\begin{aligned}
b_{1,2} &= \mu_1 +/- \sqrt{3(\mu_2 - \mu_1^2)} \\
a_{1,2} &= \mu_1 -/+ \sqrt{3(\mu_2 - \mu_1^2)}
\end{aligned}$$

Da $a < b$ gilt folgt

$$\begin{aligned} b &= \mu_1 + \sqrt{3(\mu_2 - \mu_1^2)} \\ a &= \mu_1 - \sqrt{3(\mu_2 - \mu_1^2)} \end{aligned}$$

Einsetzen der Stichprobenmomente $\hat{\mu}_1 = \overline{X}$ und $\hat{\mu}_2 = \frac{1}{N}\sum_n X_n^2$ ergibt die gesuchten Momentenschätzer.

$$\hat{a} = \overline{X} - \sqrt{3}S_* \qquad \hat{b} = \overline{X} + \sqrt{3}S_*$$

Aufgabe 10

10.1

Aus $F(x) = 1 - e^{-\lambda x}$ ergibt sich $f(x) = \lambda e^{-\lambda x}$ und somit die Likelihoodfunktion

$$L(x_1, ..., x_N; \lambda) = \prod_{n=1}^{N} \lambda e^{-\lambda x_n}$$

durch Logarithmieren (monotone Transformation) ergibt sich die sogenannte Loglikelihoodfunktion LL, die leichter zu maximieren ist.

$$LL(x_1, ..., x_N; \lambda) = \sum_{n=1}^{N}(ln\lambda - \lambda x_n)$$

Differentiation von $LL(x_1, ..., x_N; \lambda)$ nach λ und Nullsetzen ergibt

$$\frac{\partial LL(x_1, ..., x_N; \lambda)}{\partial \lambda} = \frac{1}{\lambda}N - \sum_{i=1}^{N} x_n = 0$$

und somit $\hat{\lambda} = \frac{N}{\sum_{n=1}^{N} X_n}$.

10.2

Für die Poisson-Verteilung gilt:

$$Ps(x|\mu) = f_X(x|\mu) = \frac{1}{x!}\mu^x e^{-\mu}$$

Daraus ergibt sich die Likelihood-Funktion

$$L(x_1, x_2, ..., x_N;\ \mu) = \prod_{n=1}^{N} \frac{1}{x_n!} \cdot \mu^{x_n} \cdot e^{-\mu}$$

Durch Umordnung und Zusammenfassung ergibt sich

$$L(x_1, x_2, ..., x_N;\ \mu) = \frac{1}{x_1!\ x_2!\ x_3!\ ...\ x_N!} \cdot e^{-N\mu} \cdot \mu^{N\overline{x}}$$

Die logarithmierte Likelihood-Funktion lautet

$$LL(x_1, x_2, ..., x_N;\ \mu) = -\ln(x_1! \cdot x_2! \cdot x_3! \cdot ... \cdot x_N!) - N\mu + N\overline{x} \cdot \ln \mu$$

Differentiation von $LL(x_1, ..., x_N; \mu)$ nach μ und Nullsetzen ergibt

$$\frac{\partial LL(x_1, ..., x_N; \mu)}{\partial \mu} = -N + N\overline{x}\frac{1}{\mu} = 0$$

und somit $\hat{\mu} = \overline{x}$.

Aufgabe 11

Gegeben sei eine GG mit Umfang N, in der jedes Element die Eigenschaft A oder $\overline{A}$ besitzt mit den Wahrscheinlichkeiten $P(A) = P(\overline{A}) = 1/2$. Die relative Häufigkeit $f_N(A)$ ist ein Punktschätzer für $P(A)$.

11.1

Mit $N = 8$ liegt eine Binomialverteilung $B(8, 0.5)$ vor. Die relative Häufigkeit $f_8(A)$ trifft die Wahrscheinlichkeit $P(A) = 1/2$ genau dann, wenn in der Stichprobe 4 Elemente mit der Eigenschaft A sind, d.h. $P(X = 4) = 0.2734$ ist die gesuchte Wahrscheinlichkeit.

Für $N = 15$ nimmt die gesuchte Wahrscheinlichkeit den Wert 0 an, da für eine ungerade Anzahl die relative Häufigkeit die Werte 7/15 oder 8/15 annehmen kann aber nie genau 1/2.

11.2

$\hat{\pi} = \frac{18}{40}$, $N\pi = 20 \geq 5$, $N(1-\pi) = 20 \geq 5$, $\alpha = 0.05$.

Als Schätzer für die Varianz ergibt sich $\frac{1}{40}\hat{\pi}(1-\hat{\pi}) = 0.00625$.

$$
\begin{aligned}
KI_{zweiseitig} &= \left[\ \hat{\pi} - z\sqrt{\frac{1}{N}\hat{\pi}(1-\hat{\pi})}\ ,\ \hat{\pi} + z\sqrt{\frac{1}{N}\hat{\pi}(1-\hat{\pi})}\ \right] \\
&= [0.45 - 1.96 \cdot 0.0787\ ,\ 0.45 + 1.96 \cdot 0.0787] \\
&= [0.45 - 0.154\ ,\ 0.45 + 0.154] = [0.296\ ,\ 0.604] \\
&\quad (z = z(1-\alpha/2) = z(0.975))
\end{aligned}
$$

$$
\begin{aligned}
KI_{einseitig} &= \left[\ \hat{\pi} - z\sqrt{\frac{1}{N}\hat{\pi}(1-\hat{\pi})}\ ,\ \infty\ \right] \\
&= \left[0.45 - 1.65 \cdot 0.0787\ ,\ \infty\right] \\
&= [0.32\ ,\ \infty] \\
&\quad (z = z(1-\alpha) = z(0.95))
\end{aligned}
$$

$$
\begin{aligned}
KI_{einseitig} &= \left[\ -\infty\ ,\ \hat{\pi} + z\sqrt{\frac{1}{N}\hat{\pi}(1-\hat{\pi})}\ \right] \\
&= \left[-\infty\ ,\ 0.45 + 1.65 \cdot 0.0787\right] \\
&= [-\infty\ ,\ 0.58] \\
&\quad (z = z(1-\alpha) = z(0.95))
\end{aligned}
$$

Aufgabe 12

Überprüfen sie folgende Aussagen über ein zweiseitiges Konfidenzintervall für den Erwartungswert μ einer Grundgesamtheit auf ihre Richtigkeit. Begründen Sie Ihre Antworten.

12.1

Die Grenzen eines 95%-Konfidenzintervalls sind Zufallsvariablen.

Diese Aussage ist richtig. Die Grenzen eines Konfidenzintervalls sind Funktionen von Zufallsvariablen und somit selbst zufällig.

12.2

Mit wachsendem Stichprobenumfang wird das 95%-Konfidenzintervall im Allgemeinen breiter.

Diese Aussage ist falsch. Bei steigendem Stichprobenumfang nimmt die Genauigkeit der Schätzung zu und das Intervall wird schmaler.

12.3

Bei gleichem Datensatz ist das 95%-Konfidenzintervall breiter als das 90%-Konfidenzintervall.

Diese Aussage ist richtig. Bei zunehmendem α nimmt die Breite des Konfidenzintervalls ab, d.h. bei abnehmendem Konfidenzniveau $1-\alpha$ wird das Konfidenzintervall schmaler. Somit ist das 90%-Konfidenzintervall schmaler als das 95%-Konfidenzintervall.

12.4

Die Wahrscheinlichkeit, dass die obere Grenze des zweiseitigen 95%-Konfidenzintervalls kleiner als μ ist, beträgt 0.05 (stetige Zufallsvariable).

Diese Aussage ist falsch.

$$P(O < \mu) = P(\mu > O) = \frac{\alpha}{2} = 0.025$$

Aufgabe 13

13.1

Aus der Tabelle der Student-Verteilung wird für $N-1 = 15$ Freiheitsgraden und für $\alpha = 0.01$ der Wert $t(1-\alpha/2, N-1) = t(0.995, 15) = 2.947$ abgelesen.

$$\overline{x} \pm t\frac{s}{\sqrt{N}} = 5 \pm 2.947\frac{5}{\sqrt{16}} = [1.32, 8.68]$$

13.2

Wegen

$$\hat{\pi} = \frac{45}{225} = 0.2 \quad \text{und} \quad N\hat{\pi} = 45 \geq 5, N(1-\hat{\pi}) = 180 \geq 5$$

kann die vorliegende Binomialverteilung durch die Normalverteilung approximiert werden. Die Grenzen der approximativen einseitigen Konfidenzintervalle für π mit $z = z(1-\alpha) = z(0.95) = 1.65$ lauten

$$\left[\hat{\pi} - z\sqrt{\tfrac{\hat{\pi}(1-\hat{\pi})}{N}}, \infty\right] = \left[0.2 - 1.65 \cdot \sqrt{\tfrac{0.16}{225}}, \infty\right] = [0.156, \infty]$$

$$\left[-\infty, \hat{\pi} + z\sqrt{\tfrac{\hat{\pi}(1-\hat{\pi})}{N}}\right] = [-\infty, 0.2 + 0.044] = [-\infty, 0.244]$$

13.3

Mit den Quantilen $\chi^2(0.025, 99) = 73.361$ und $\chi^2(0.975, 99) = 128.422$ ergeben sich die Grenzen

$$\sigma_o^2 = \frac{99 \cdot 950625}{73.361} = 1282859.762,$$

$$\sigma_u^2 = \frac{99 \cdot 950625}{128.422} = 732832.965.$$

Zu beachten ist, dass das Konfidenzintervall [732832.965,1282859.762] nicht symmetrisch bezüglich $s^2 = 950625$ ist.

Aufgabe 14

14.1

Ein Signifikanzniveau von α bedeutet, dass die Nullhypothese mit einer Wahrscheinlichkeit von α falsch ist.

Diese Aussage ist falsch. Das Signifikanzniveau α gibt die Wahrscheinlichkeit an, die Nullhypothese abzulehnen, wenn diese zutrifft. Über das Zutreffen der Hypothese wird keine Aussage gemacht.

14.2

Wenn die Nullhypothese abgelehnt wird, so kann sie mit einer Wahrscheinlichkeit von höchstens α trotzdem richtig sein.

Diese Aussage ist falsch. Eine Hypothese ist entweder richtig oder falsch.

14.3

Bei einem einseitigen Test ist die Alternativhypothese genau dieselbe wie bei einem zweiseitigen Test.

Diese Aussage ist falsch. Bei einem zweiseitigen Test lautet die Alternativhypothese $H_1: \ \theta \neq \theta_0$ und bei einem einseitigen Test $H_1: \ \theta > \theta_0$ bzw. $H_1: \ \theta < \theta_0$.

14.4

Liegt die Prüfgröße im Annahmebereich der Verteilung der Prüfgröße unter der Alternativhypothese, so kann H_0 nicht verworfen werden.

Diese Aussage ist falsch. Die Verteilung der Prüfgröße wird unter der Nullhypothese berechnet.

Aufgabe 15

$$H_0: \ \mu \geq \mu_0 = 100000 \ (m/h) \qquad\qquad H_1: \ \mu < \mu_0$$

Die Standardabweichung σ ist bekannt und $\overline{X}$ ist wegen $N > 30$ approximativ normalverteilt. Als Prüfgröße wird $Z = \frac{\overline{X}-\mu_0}{\sigma_{\overline{X}}}$ verwendet, welche approximativ standardnormalverteilt ist ($N = 36, \alpha = 0.05$). Mit $\sigma_{\overline{X}} = \frac{\sigma}{\sqrt{N}} = \frac{30000}{6} = 5000$ ergibt sich:

$$z = \frac{90000 - 100000}{5000} = -2.$$

Wegen $-2 < -z(0.95) = -1.65$ ist die Behauptung des Automobilherstellers A statistisch widerlegt.

Aufgabe 16

$$H_0: \ \mu = \mu_0 = 500 \ (ml) \qquad\qquad H_1: \ \mu \neq \mu_0$$

Die Standardabweichung σ ist unbekannt und X wird als approximativ normalverteilt vorausgesetzt. Als Prüfgröße kann $T = \frac{\overline{X}-\mu_0}{\hat{\sigma}_{\overline{X}}} = \sqrt{N}\frac{\overline{X}-\mu_0}{S}$ verwendet werden, welche approximativ t-verteilt ist mit $N-1$ Freiheitsgraden ($N = 25, \alpha = 0.05$).

$$t = \sqrt{25}\,\frac{490-500}{5} = -2$$

Die Nullhypothese wird nicht verworfen, da $t(0.025, 24) = -2.064 \leq -2 \leq t(0.975, 24) = 2.064$ gilt.

Dass die Nullhypothese zutrifft, ist jedoch nicht statistisch gesichert.

Aufgabe 17

$$H_0: \ \mu = \mu_0 = 38 \qquad\qquad H_1: \ \mu \neq \mu_0$$

Die Standardabweichung σ ist bekannt und wegen $N > 30$ ist $\overline{X}$ approximativ $N(38,\ \sigma^2_{\overline{X}})$-verteilt. Es wird die approximativ standardnormalverteilte Prüfgröße $Z = \frac{\overline{X}-\mu_0}{\sigma_{\overline{X}}}$ verwendet mit $\sigma_{\overline{X}} = \frac{8}{\sqrt{100}} = 0.8$ $(N = 100, \alpha = 0.05)$.

$$z = \frac{39-38}{0.8} = 1.25$$

Es ist $z(0.025) = -z(0.975) = -1.96 \leq 1.25 = z \leq z(0.975) = 1.96$, d.h. die Realisation z der Prüfgröße Z liegt zwischen den gegebenen kritischen Werte $-z(0.975)$ und $z(0.975)$. Die Nullhypothese kann somit nicht abgelehnt werden.

Aufgabe 18

$$H_0: \ \sigma^2 \leq \sigma_0^2 = 0.01 \qquad\qquad H_1: \ \sigma^2 > \sigma_0^2 = 0.01$$

Die Prüfgröße $\chi^2 = \frac{(N-1)S^2}{\sigma_0^2}$ ist χ^2-verteilt mit $N - 1 = 31$ Freiheitsgraden. Mit $s^2 = 0.012$ ergibt sich die Realisation

$$\chi^2 = \frac{31 \cdot 0.012}{0.01} = 37.2.$$

Der obere kritische Wert entspricht dem Quantil $\chi^2(1-\alpha, 31)$ mit $\alpha = 0.05$. Da $\chi^2 = 37.2 \leq \chi^2(0.95, 31) = 44.99$, kann H_0 nicht abgelehnt werden. Es besteht kein hinreichender Grund, an der Aussage des Herstellers zu zweifeln.

Aufgabe 19

$$H_0: \ \sigma \leq \sigma_0 = 10 \qquad\qquad H_1: \ \sigma > \sigma_0$$

Die Prüfgröße $\chi^2 = \frac{(N-1)S^2}{\sigma_0^2}$ ist χ^2-verteilt mit $N-1 = 49$ Freiheitsgraden. Mit $s^2 = 125$ ergibt sich die Realisation

$$\chi^2 = \frac{125 \cdot 49}{100} = 61.25.$$

Der kritische Wert lautet $\chi^2(0.9, 49) = 62.038$, so dass die Nullhypothese nicht abgelehnt werden kann ($61.25 < 62.038$).

Aufgabe 20

$$H_0: \ \mu_X - \mu_Y = 0 \qquad\qquad \mu_X - \mu_Y \neq 0$$

Unter den getroffenen Annahmen (Normalverteilungsannahme, Varianzhomogenität, σ_X und σ_Y unbekannt, $N, M < 30$, $\alpha = 0.05$) wird zur Berechnung der Prüfgröße $T = \frac{\overline{X}_X - \overline{X}_Y}{S_D}$ die Varianz

$$s_d^2 = \left(\frac{1}{N} + \frac{1}{M}\right) \frac{(N-1)s_x^2 + (M-1)s_y^2}{N+M-2} = \frac{1}{6} \cdot \frac{14 \cdot 0.6 + 9 \cdot 0.4}{23} = \frac{2}{23}$$

benötigt. Da sich der Wert der Prüfgröße T zu

$$t = \frac{\overline{x}_x - \overline{x}_y}{s_d} = \frac{3.15 - 2.55}{\sqrt{\frac{2}{23}}} = 2.035$$

ergibt mit $-2.069 = t(0.025, 23) \leq t \leq t(0.975, 23) = 2.069$, kann die Nullhypothese nicht abgelehnt werden.

Aufgabe 21

$$H_0: \ \frac{\sigma_Y^2}{\sigma_X^2} = 1 \qquad\qquad H_1: \ \frac{\sigma_Y^2}{\sigma_X^2} \neq 1$$

Für die Prüfgröße $F = \frac{S_Y^2}{S_X^2}$ ergibt sich der Wert $\frac{12.25}{28.09} = 0.436$. Die kritischen Werte lauten ($\alpha = 0.1$):

$$\begin{aligned}
F(0.95, 19, 14) &= 2.4 \\
F(0.05, 19, 14) &= \frac{1}{F(0.95, 14, 19)} = \frac{1}{2.26} = 0.442
\end{aligned}$$

Da $0.436 < F(0.05, 19, 14) = 0.442$ gilt, wird H_0 abgelehnt.

Aufgabe 22

$$H_0: \ \pi \leq \pi_0 = 0.15 \qquad\qquad H_1: \ \pi > \pi_0$$

In der Stichprobe befinden sich $x = 15$ Kreditnehmer mit Rückzahlungsschwierigkeiten.

Es gilt $50\pi_0 = 7.5 \geq 5$, $50(1-\pi_0) = 42.5 \geq 5$, so dass eine Approximation durch die Normalverteilung möglich ist.

$$z = \frac{\hat{\pi} - \pi_0}{\sqrt{\frac{1}{N}\pi_0(1-\pi_0)}} = \frac{0.3 - 0.15}{0.0505} = 2.97$$

Da $2.97 > 2.33 = z(0.99)$ gilt, wird die Nullhypothese abgelehnt.

Aufgabe 23

$$H_0: \ \pi \geq \pi_0 = 0.2 \quad \text{vs.} \quad H_1: \ \pi < \pi_0$$

Mit $N = 100$ gilt $N\pi_0 = 20 \geq 5$, $N(1-\pi_0) = 80 \geq 5$, so dass eine Approximation durch die Normalverteilung möglich ist.

$$z = \frac{\hat{\pi} - \pi_0}{\sqrt{\frac{1}{N}\pi_0(1-\pi_0)}} = \frac{0.15 - 0.2}{0.04} = -1.25$$

Die Nullhypothese wird zum Signifikanzniveau $\alpha = 0.05$ nicht abgelehnt, da $-1.25 \geq -1.65 = z(0.05)$ gilt.

Aufgabe 24

$$\begin{aligned} H_0: &\quad P(X \in \text{Klasse i}) = P(X = x_i) = \tfrac{1}{6} \text{ für alle } i = 1, ..., 6 \\ H_1: &\quad P(X = x_i) \neq \tfrac{1}{6} \text{ für mindestens ein } i = 1, ..., 6 \end{aligned}$$

Augenzahl x_i	beobachtete Häufigkeit n_i	erwartete Häufigkeit $N\pi_i$	$(n_i - N\pi_i)^2$	$\frac{(n_i - N\pi_i)^2}{N\pi_i}$
1	20	20	0	0
2	22	20	4	$\frac{4}{20}$
3	17	20	9	$\frac{9}{20}$
4	18	20	4	$\frac{4}{20}$
5	19	20	1	$\frac{1}{20}$
6	24	20	16	$\frac{16}{20}$
				$\chi^2 = \frac{34}{20} = 1.7$

Nach dem χ^2-Anpassungstest mit der Prüfgröße $\chi^2 = \sum_{i=1}^{l} \frac{(n_i - N\pi_i)^2}{N\pi_i}$ ergibt sich zum Signifikanzniveau $\alpha = 0.1$ das Quantil $\chi^2(l-1, 1-\alpha) = \chi^2(5, 0.9) = 9.236$. Mit $9.236 \geq 1.7 = \chi^2$ kann die Hypothese „Der Würfel ist ideal“ nicht abgelehnt werden.

Aufgabe 25

Als Prüfgröße für den χ^2-Anpassungstest wird $\chi^2 = \sum_{i=1}^{l} \frac{(n_i - N\pi_i)^2}{N\pi_i}$ verwendet. Die erwarteten absoluten Häufigkeiten werden hier über die Wahrscheinlichkeiten der Normalverteilung π_i für die entsprechenden Intervalle berechnet. Für die erste und die letzte Klasse gilt

$$\pi = \pi_1 = P(X < 140) = \pi_5 = P(X \geq 200) = P(Z < -3) = 0.00135$$

und somit $n\pi = 0.675 < 5$. Die ersten beiden und die letzten beiden Klassen werden daher zusammengefasst, um die Bedingung $N\pi_i \geq 1, N\pi_i \geq 5$ für mindestens 80% der Klassen zu erfüllen.

$$\begin{aligned} H_0: &\quad P(X \in \text{Klasse i}) = \pi_i \text{ für alle } i = 1, 2, 3 \\ H_1: &\quad P(X \in \text{Klasse i}) \neq \pi_i \text{ für mindestens ein } i = 1, 2, 3 \end{aligned}$$

Für die zusammengefassten Klassen gilt

$$\begin{aligned} \pi_1 &= P(X < 170) = P\big(Z < \tfrac{170-\mu}{\sigma}\big) = P(Z < 0) = 0.5, \\ \pi_2 &= P(170 \leq X < 190) = P(0 \leq Z < 2) = 0.4772, \\ \pi_3 &= P(190 \leq X) = P(2 \leq Z) = 0.0228. \end{aligned}$$

i	Körpergröße X in Klasse i	π_i	$N\pi_i$	n_i	$n_i - N\pi_i$	$(n_i - N\pi_i)^2$	$\frac{(n_i - N\pi_i)^2}{N\pi_i}$
1	$x < 170$	0.5	250	202	-48	2304	9.216
2	$170 \leq x < 190$	0.4772	238.6	265	26.4	696.96	2.921
3	$190 \leq x$	0.0228	11.4	33	21.6	466.56	40.926
		1.0000	500	500	0		$53.063 = \chi^2$

Zum Signifikanzniveau $\alpha = 0.05$ ergibt sich das Quantil $\chi^2(l-1, 1-\alpha) = \chi^2(2, 0.95) = 5.99$. Mit $\chi^2(2, 0.95) < 53.063$ wird die Nullhypothese abgelehnt.

Aufgabe 26

$$H_0: \ x(0.5) = 0 \qquad\qquad H_1: \ x(0.5) \neq 0$$

Zugrundegelegt werden die Zufallsvariablen $D_n = X_n - Y_n$ mit $D_n = 1$ falls X besser wirkt und $D_n = -1$ falls Y besser wirkt. Als Prüfgröße des Vorzeichentests (Signifikanzniveau $\alpha = 0.05$) wird

$$Z = \frac{T - N/2}{\sqrt{N/4}}$$

mit T=„Anzahl der Personen, bei denen X besser wirkt" verwendet, welche approximativ standardnormalverteilt ist.

$$z = \frac{38 - 32}{4} = 1.5$$

Die kritischen Werte $z(\alpha/2)$, $z(1-\alpha/2)$ lauten -1.96, 1.96, so dass die Nullhypothese nicht abgelehnt werden kann.

Aufgabe 27

$$H_0: \ x(0.5) = 0 \qquad\qquad H_1: \ x(0.5) \neq 0$$

Betrachtet werden die Zufallsvariablen $D_n = X_n - Y_n$.

i	1	2	3	4	5	6	7	8	9	10	$\sum$
x_n	8.3	10.2	8.1	7.4	10.8	12.5	12.2	9.8	10.2	8.4	
y_n	8.8	10.6	8.7	7.9	10.7	13.2	12.9	9.6	10.1	8.9	
$x_n - y_n$	-0.5	-0.4	-0.6	-0.5	0.1	-0.7	-0.7	0.2	0.1	-0.5	
r_n^+					1.5			3	1.5		6
r_n^-	6	4	8	6		9.5	9.5			6	49

Die Prüfgröße des Wilcoxon-Vorzeichen-Rang-Tests

$$W^+ = \sum_{n=1}^{N} R_n Z_n \quad \text{mit} \quad Z_n = \begin{cases} 1 & \text{falls } D_n > 0 \\ \\ 0 & \text{falls } D_n < 0 \end{cases} \qquad n = 1, ..., N$$

und $R_n = rg|D_n| = rg|X_n - Y_n|$ nimmt den Wert 6 an. Für $N = 10$ und $\alpha = 0.05$ ergibt sich der untere kritische Wert zu $w(\alpha/2) = w(0.025) = 9$. Die Nullhypothese muss abgelehnt werden, denn es gilt $W^+ = 6 < 9$.

Aufgabe 28

$$H_0: \ \mu_x = \mu_y \qquad\qquad H_1: \ \mu_x \neq \mu_y$$

Als Prüfgröße des Wilcoxon-Rangsummen-Tests wird

$$\begin{aligned} T_W &= \sum_{n=1}^{N} \mathrm{rg}(X_n) = \sum_{n=1}^{N+M} rg(Z_k) V_k \\ & \quad V_k = \begin{cases} 1 & \text{falls } Z_k \in \text{Stichprobe 1} \\ 0 & \text{sonst} \end{cases} \end{aligned}$$

verwendet, wobei $Z = \{X_1, ..., X_N, Y_1, ..., Y_M\}$ ist. Die Rangzahlen werden für die gepoolte Stichprobe vergeben, wobei Bindungen innerhalb derselben Stichprobe nicht von Interesse sind. Liegen Bindungen zwischen beiden Stichproben vor, werden Durchschnittsränge gebildet ($N + M = 20$).

	# Schuhpaare pro Tag										$\sum$
X_n	14	13	18	17	19	25	21	11	8	14	
$\mathrm{rg}(X_n)$	5	4	11	9.5	12	18.5	14	2	1	6	**83**
Y_n	24	15	20	17	12	24	26	22	25	16	
$\mathrm{rg}(Y_n)$	16	7	13	9.5	3	17	20	15	18.5	8	127

Mit $\alpha = 0.1$ ergeben sich die kritischen Werte zu $w(0.05) = 83$ und $w(0.95) = 127$. Da $83 \leq T_W = 83 \leq 127$ gilt, kann die Nullhypothese, beide Verkäufer weisen die gleichen Umsatzzahlen auf, soeben nicht abgelehnt werden.

Aufgabe 29

Zur Berechnung von β muß bestimmt werden, mit welcher Wahrscheinlichkeit die Prüfgröße $\overline{X}$ für $\mu = 102$ in das Intervall [98.7; 101.3] fällt, d.h. in den Beibehaltungsbereich. $\overline{X}$ ist unter der getroffenen Annahme normalverteilt mit dem Erwartungswert $\mu = 102$ und Standardabweichung $\sigma_{\overline{X}} = \frac{3}{\sqrt{20}} = 0.67$.

$$\begin{aligned}
\beta(\mu = 102) &= P(98.7 \leq \overline{X} \leq 101.3|\mu = 102) \\
&= P\left(\frac{98.7 - 102}{0.67} \leq Z \leq \frac{101.3 - 102}{0.67}\right) \\
&= P(-4.9 \leq Z \leq -1.04) \\
&= F_Z(-1.04) - F_Z(-4.9) = F_Z(4.9) - F_Z(1.04) \\
&= 1 - 0.8508 = 0.1492 = 0.15.
\end{aligned}$$

Für den Fall, dass der tatsächliche Erwartungswert $\mu = 102$ ist, beträgt die Wahrscheinlichkeit für die Nichtablehnung der Nullhypothese ($H_0 : \mu = \mu_0 = 100$)

$$\beta(\mu = 102) = 0.15.$$

Folgende Abbildung veranschaulich den Fehler 2. Art $\beta(\mu = 102)$ mit den kritischen Werten c_u und c_o.

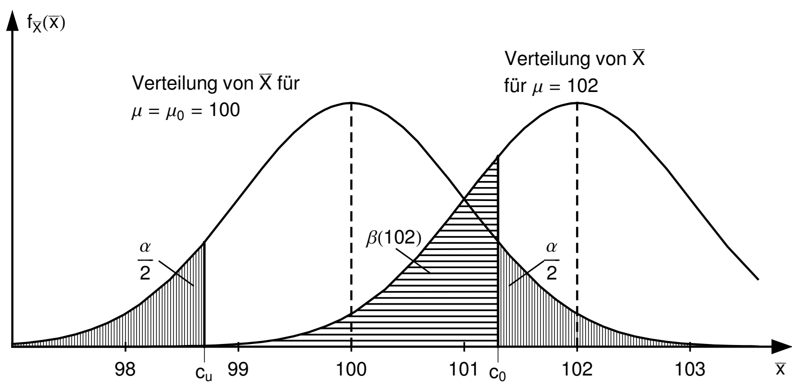

Abbildung 0.0.1: *Wahrscheinlichkeit für den Fehler 2. Art $\beta(\mu = 102)$*

Die Wahrscheinlichkeit β für den Fall, dass der tatsächliche Parameterwert $\mu = 100.5$ beträgt wird analog berechnet.

$$
\begin{aligned}
\beta(\mu = 100.5) &= P(98.7 \leq \overline{X} \leq 101.3 | \mu = 100.5) \\
&= P\left(\frac{98.7 - 100.5}{0.67} \leq Z \leq \frac{101.3 - 100.5}{0.67}\right) \\
&= P(-2.69 \leq Z \leq 1.19) \\
&= F_Z(1.19) - F_Z(-2.69) \\
&= F_Z(1.19) - 1 + F_Z(2.69) \\
&= 0.8830 - 1 + 0.9964 \\
&= 0.8794.
\end{aligned}
$$

Die Wahrscheinlichkeit für den Fehler 2. Art ist also im Fall $\mu = 100.5$ recht hoch.

In folgender Abbildung ist β für verschiedene μ-Werte veranschaulicht. Im Grenzfall $\mu = \mu_0$ ist die Wahrscheinlichkeit, H_0 nicht zu verwerfen, gleich

$$\beta(\mu = 100) = 1 - \alpha.$$

Das ist selbstverständlich, denn in diesem Fall ist H_0 tatsächlich richtig, und die Wahrscheinlichkeit, in diesem Fall eine richtige Entscheidung zu treffen, nämlich H_0 nicht zu verwerfen, wurde mit $1 - \alpha$ vorgegeben.

Die Wahrscheinlichkeit β entspricht der Fläche unterhalb der jeweiligen Dichtefunktion über der μ-Achse, zwischen den kritischen Werten.

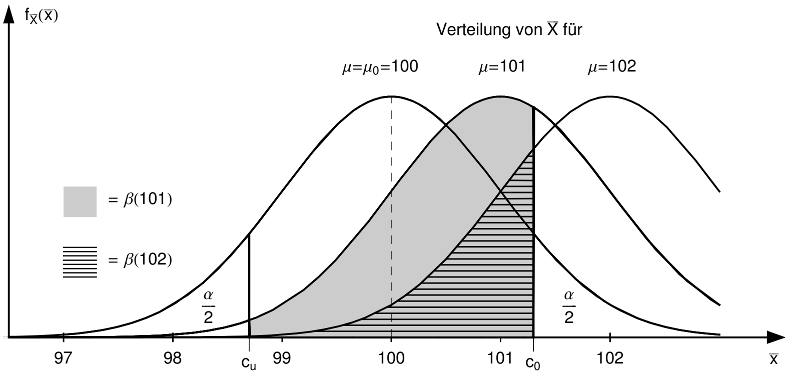

Abbildung 0.0.2: *β für verschiedene Werte von μ*

Aufgabe 30

$$H_0: \ \mu \geq \mu_0 \qquad\qquad H_1: \ \mu < \mu_0$$

30.1

Der kritische Wert c_u wird mit $\alpha = 0.01$ und $z = z(0.99) = 2.33$ zu

$$c_u = \mu_0 - z \cdot \sigma_{\overline{X}} = 1000 - 2.33 \cdot \frac{5}{8} = 998.54$$

berechnet. Die Nullhypothese wird abgelehnt, falls $\overline{x} < 998.54$ gilt.

30.2

Es muss die Wahrscheinlichkeit berechnet werden, die Nullhypothese nicht abzulehnen, obwohl sie falsch ist unter der Voraussetzung $\mu = 998$, d.h. gesucht ist

$$\beta(\mu = 998) = P(\overline{X} \geq 998.54 | \mu = 998).$$

Dies entspricht der Wahrscheinlichkeit, dass $\overline{X}$ größer als der kritische Wert ist unter der Annahme, dass $\overline{X}$ normalverteilt ist mit Erwartungswert 998 und Standardabweichung 5/8.

$$\beta(\mu = 998) = P(\overline{X} \geq 998.54 | \mu = 998) = P(Z \geq 0.864) = 1 - P(Z \leq 0.864)$$

Laut Tabelle gilt $1 - P(Z \leq 0.86) = 0.1949$.

30.3

Zu berechnen ist N mit

$$\begin{aligned} & P(\overline{X} \geq 998.54 | \mu = 998) \leq 0.05 \\ \Longleftrightarrow \quad & P(Z \geq \frac{0.54}{5/\sqrt{N}}) \leq 0.05 \\ \Longleftrightarrow \quad & P(Z \leq \frac{0.54}{5/\sqrt{N}}) \geq 0.95 \end{aligned}$$

Für die Gleichheit $P(Z \leq \frac{0.54}{5/\sqrt{N}}) = 0.95$ ergibt sich somit $\frac{0.54}{5/\sqrt{N}} = 1.65$. Dies nach N aufgelöst ergibt 233.41. Somit muss der Stichprobenumfang mindestens 234 lauten.