Kurs 40601 Grundlagen der Statistik

Kurseinheit 2
Lösungskommentare

Aufgabe 1 B,C und E sind richtig.

zu A:

$$\begin{aligned} f_Y(y) &= \int_{-1}^{1} \frac{3}{2}(1-x^2)(1-y)\ dx = \frac{3}{2}(1-y) \cdot \left[x - \frac{x^3}{3}\right]_{-1}^{1} \\ &= \frac{3}{2}\left(1-\frac{1}{3}+1-\frac{1}{3}\right)(1-y) = \frac{3}{2}\cdot\frac{4}{3}(1-y) = 2(1-y) \end{aligned}$$

zu B:

$$\begin{aligned} f_X(x) &= \int_{0}^{1} \frac{3}{2}(1-x^2)(1-y)\ dy = \frac{3}{2}(1-x^2) \cdot \left[y - \frac{y^2}{2}\right]_{0}^{1} \\ &= \frac{3}{2}\left(1-\frac{1}{2}\right)(1-x^2) = \frac{3}{4}(1-x^2) \end{aligned}$$

zu C: $f_X(x) \cdot f_Y(y) = \frac{3}{4}(1-x^2)\cdot 2(1-y) = \frac{3}{2}(1-x^2)(1-y) = f_{XY}(x,y)$

zu D:

$$E(Z) = \int_{-\infty}^{\infty} z f_Z(z)\ dz = \int_0^2 z \cdot \frac{1}{4}z^3\ dz = \left[\frac{1}{20}z^5\right]_0^2 = \frac{32}{20} - 0 = 1.6$$

zu E:

$$\begin{aligned} \mathrm{Var}(Z) &= \int_{-\infty}^{\infty} [z-E(Z)]^2 f_Z(z)\ dz \\ &= E(Z^2) - [E(Z)]^2 = \int_{-\infty}^{\infty} z^2 f_Z(z)\ dz - [E(Z)]^2 \\ &= \int_0^2 z^2 \cdot \frac{1}{4}z^3\ dz - (1.6)^2 \\ &= \left[\frac{1}{24}z^6\right]_0^2 - (1.6)^2 = \frac{64}{24} - (1.6)^2 = 2.67 - 2.56 = 0.11 \end{aligned}$$

Aufgabe 2 A und D sind richtig.

X	Y		$f_X(x_i)$
	$y_1 = 2$	$y_2 = 3$	
$x_1 = 1$	$f_{XY}(x_1, y_1) = 0.2$	$f_{XY}(x_1, y_2) = 0.4$	$f_X(x_1) = 0.6$
$x_2 = 3$	$f_{XY}(x_2, y_1) = 0.2$	$f_{XY}(x_2, y_2) = 0.2$	$f_X(x_2) = 0.4$
$f_Y(y_i)$	$f_Y(y_1) = 0.4$	$f_Y(y_2) = 0.6$	1

zu A:

$$\begin{aligned} E(X) &= 0.6 \cdot 1 + 0.4 \cdot 3 = 1.8 \\ E(Y) &= 0.4 \cdot 2 + 0.6 \cdot 3 = 2.6 \end{aligned}$$

zu B:

$$\begin{aligned} &\mathrm{Cov}(X, Y) \\ &= (1 - 1.8) \cdot (2 - 2.6) \cdot 0.2 + (1 - 1.8) \cdot (3 - 2.6) \cdot 0.4 \\ &\quad + (3 - 1.8) \cdot (2 - 2.6) \cdot 0.2 + (3 - 1.8) \cdot (3 - 2.6) \cdot 0.2 \\ &= 0.8 \cdot 0.6 \cdot 0.2 - 0.8 \cdot 0.4 \cdot 0.4 - 1.2 \cdot 0.6 \cdot 0.2 + 1.2 \cdot 0.4 \cdot 0.2 \\ &= 0.096 - 0.128 - 0.144 + 0.096 = -0.08 \end{aligned}$$

zu C: Da Cov $\neq 0$, sind X und Y stochastisch abhängig. Desweiteren ist die Bedingung $f_{XY}(x_i, y_j) = f_X(x_i) f_Y(y_j)$ für alle i und j nicht erfüllt, z.B. ist $f_{XY}(x_1, y_1) = 0.2 \neq 0.24 = f_X(x_1) f_Y(y_1)$.

zu D: $f_X(x_2|Y = 2) = f_X(x_2|y_1) = \frac{f_{XY}(x_2,y_1)}{f_Y(y_1)} = \frac{0.2}{0.4} = 0.5$

Aufgabe 3 A,C und E sind richtig.

zu A: Im Maximum muss die 1. Ableitung der Dichtefunktion den Wert 0 annehmen.

$$\frac{d}{dx}\frac{1}{\sigma\sqrt{2\pi}}\exp\left(-\frac{(x-\mu)^2}{2\sigma^2}\right)=\frac{1}{\sigma\sqrt{2\pi}}\exp\left(-\frac{(x-\mu)^2}{2\sigma^2}\right)\frac{(-1)(x-\mu)}{\sigma^2}\stackrel{!}{=}0$$

Dies ist für $x=\mu$ erfüllt. In den Wendepunkten muss die 2. Ableitung der Dichtefunktion den Wert 0 annehmen.

$$\begin{aligned}
&\frac{d^2}{(dx)^2}\frac{1}{\sigma\sqrt{2\pi}}\exp\left(-\frac{(x-\mu)^2}{2\sigma^2}\right)\\
=\;&\frac{1}{\sigma\sqrt{2\pi}}\left[\exp\left(-\frac{(x-\mu)^2}{2\sigma^2}\right)\cdot\frac{(-1)(x-\mu)}{\sigma^2}\cdot\frac{(-1)(x-\mu)}{\sigma^2}\right.\\
&\left.+\exp\left(-\frac{(x-\mu)^2}{2\sigma^2}\right)\cdot\frac{(-1)}{\sigma^2}\right]\\
=\;&\frac{1}{\sigma\sqrt{2\pi}}\exp\left(-\frac{(x-\mu)^2}{2\sigma^2}\right)\cdot\left(\frac{(x-\mu)^2}{\sigma^4}-\frac{1}{\sigma^2}\right)\stackrel{!}{=}0
\end{aligned}$$

Dies gilt, wenn $\frac{(x-\mu)^2}{\sigma^4}=\frac{1}{\sigma^2}$ bzw. $(x-\mu)^2=\sigma^2$ bzw. $x-\mu=\pm\sigma$ ist. Die Normalverteilung nimmt ihre Wendepunkte somit an den Stellen $x=\mu\pm\sigma$ an. Auf den Nachweis der hinreichenden Bedingungen wird an dieser Stelle verzichtet.

zu B:

$$\begin{aligned}
\boldsymbol{P}(2\mu-2\sigma\le X\le 2\mu+2\sigma) &= \boldsymbol{P}(-2\sigma\le X-2\mu\le 2\sigma)\\
&= \boldsymbol{P}\left(-1\le\frac{X-2\mu}{2\sigma}\le 1\right)
\end{aligned}$$

Dies lässt keine allgemeingültige Aussage zu, da $\frac{X-2\mu}{2\sigma}$ nicht $N(0,1)$ verteilt ist.

zu C: Für jede Normalverteilung gilt aufgrund der Symmetrieeigenschaft $\boldsymbol{P}(X\le\mu)=\boldsymbol{P}(X\ge\mu)=0.5$

zu D: $(aX_1+bX_2+c)\sim N(a\mu_1+b\mu_2+c,a^2\sigma_1^2+b^2\sigma_2^2)$

zu E: Dies gilt, da die Normalverteilung symmetrisch und unimodal ist.

Aufgabe 4 B und C sind richtig.

Die Anzahl der defekten Transistoren wird mit X bezeichnet. Die Zufallsvariable X ist somit binomialverteilt mit den Parametern n und π, $X \sim B(n, \pi)$.

zu A: Gesucht ist $\boldsymbol{P}(1700 \cdot 0.01 \leq X \leq 1700 \cdot 0.02)$. Wegen $1700 \cdot 0.01 = 17 \geq 5$ und $1700 \cdot 0.99 = 1683 \geq 5$ ist X approximativ $N(17, 4.102)$ verteilt. Unter Berücksichtigung der Stetigkeitskorrektur gilt:

$$\begin{aligned}
\boldsymbol{P}(17 - 0.5 \leq X \leq 34 + 0.5) &= \boldsymbol{P}\left(\frac{-0.5}{4.102} \leq Z \leq \frac{17.5}{4.102}\right) \\
&= \boldsymbol{P}(-0.12 \leq Z \leq 4.27) \\
&= F_Z(4.27) - F_Z(-0.12) \\
&= F_Z(4.27) - 1 + F_Z(0.12) \\
&= 1 - 1 + 0.5478 \\
&= 0.5478
\end{aligned}$$

zu B: s.A.

zu C: Die Anzahl X der defekten Transistoren ist binomialverteilt mit $X \sim B(20, 0.05)$. Gesucht ist $\boldsymbol{P}(20 \cdot 0.05 \leq X \leq 20 \cdot 0.1)$. Wegen $20 \cdot 0.05 = 1 < 5$ ist keine Approximation durch die Normalverteilung möglich.

$$\begin{aligned}
\boldsymbol{P}(1 \leq X \leq 2) &= \boldsymbol{P}(X = 1) + \boldsymbol{P}(X = 2) \\
&= 0.3774 + 0.1887 = 0.5661
\end{aligned}$$

zu D: s.C. Hier ist $\boldsymbol{P}(X > 20 \cdot 0.05)$ gesucht.

$$\boldsymbol{P}(X > 1) = 1 - \boldsymbol{P}(X \leq 1) = 1 - F_X(1) = 1 - 0.7358 = 0.2642$$

Aufgabe 5 D ist richtig.

zu A: Eine effiziente Schätzfunktion ist auch erwartungstreu; aus der Erwartungstreue folgt aber nicht notwendig die Effizienz.

zu B:
$$E(X) = E\left(\frac{1}{n-1}\sum_{i=1}^{n} X_i\right) = \frac{1}{n-1}\sum_{i=1}^{n} E(X_i) = \frac{n}{n-1}\mu \neq \mu$$

zu C: Für eine erwartungstreue Schätzfunktion T des Parameters θ gilt: $E(T) = \theta$, aber nicht $T = \theta$.

zu D:
$$\begin{aligned} E(X) &= E\left(\frac{1}{2}(X_1 + X_n)\right) = \frac{1}{2}[E(X_1) + E(X_n)] \\ &= \frac{1}{2}(\mu + \mu) = \mu \end{aligned}$$

Aufgabe 6 B und D sind richtig.

zu A: Konfidenzintervalle zum Niveau $1-\alpha$ für einen Parameter θ beruhen auf der Wahrscheinlichkeitsaussage $\boldsymbol{P}(U \leq \theta \leq O) = 1 - \alpha$, wobei U und O die zufälligen Intervallgrenzen sind. Nach Realisation von U und O ist eine solche Wahrscheinlichkeitsaussage nicht mehr möglich, d.h. eine Aussage in der Form $\boldsymbol{P}(2650 \leq \theta \leq 4550) = 1-\alpha$ ist nicht mehr möglich. Entweder liegt θ im Intervall [2650; 4550] oder aber nicht. Dies hängt nur vom expliziten Wert von θ ab.

zu B: s. Kurseinheit 2.

zu C: Die Breite des Konfidenzintervalls steigt bei Erhöhung des Konfidenzniveaus.

zu D: s. A.

zu E: Dies gilt nur für den Fall, dass symmetrische Konfidenzintervalle konstruiert werden, nicht aber für den allgemeinen Fall, z. B. wenn Konfidenzintervalle mit Hilfe der F-Verteilung konstruiert werden.

Aufgabe 7 A und C sind richtig.

zu A: Als Alternativhypothese wird die zu überprüfende Vermutung aufgestellt, d.h. $H_0: \ \mu \geq \mu_0 = 5000$ und $H_1: \ \mu < \mu_0$.

zu B: Die Testgröße $\overline{X}$ ist unter der Nullhypothese $N(\mu_0, \sigma^2_{\overline{X}})$-verteilt, mit $\sigma_{\overline{X}} = \frac{\sigma}{\sqrt{n}} = \frac{100}{5} = 20$, d.h. $\overline{X} \sim N(5000, 400)$.

zu C: Für das Signifikanzniveau $\alpha = 0.05$ ergibt sich der untere kritische Wert zu

$$\begin{aligned} c_u &= \mu_0 - z_{1-\alpha}\sigma_{\overline{X}} \\ &= 5000 - 1.65 \cdot 20 = 4967. \end{aligned}$$

zu D: Wegen $\overline{x} = 4960 < c_u$ ist H_0 abzulehnen.

Aufgabe 8 A,B und D sind richtig.

zu A: Es werden die Hypothesen $H_0: \ D_{med} \leq 0$ und $H_1: \ D_{med} > 0$ aufgestellt, wobei D_{med} den Median der Differenzen D_i bezeichnet mit $D_i = X_{Ai} - X_{Bi}$.

zu B:

D_i	-2	-1	0	-2.5	-1	-0.5	1	-1.5	-0.5	1.5	-1	-2	
Rang	9.5	4.5	-	11	4.5	1.5	4.5	7.5	1.5	7.5	4.5	9.5	
$\boldsymbol{W^+}$			-				4.5			7.5			$\boldsymbol{\sum = 12}$
W^-	9.5	4.5	-	11	4.5	1.5		7.5	1.5		4.5	9.5	$\sum = 54$

zu C: Da die Differenz $D_3 = 0$ ist, wird diese Beobachtung nicht betrachtet und n reduziert sich auf 11. Als kritischer Wert ergibt sich somit $c_o = w_{1-\alpha}(11) = w_{0.95}(11) = 52$.

zu D: Ist W^+ zu groß, d.h. liegt W^+ über $c_o = w_{1-\alpha}(11)$, so wird die Nullhypothese abgelehnt. Mit $W^+ = 12 < w_\alpha(11) = 52$ wird daher H_0 nicht abgelehnt.

Aufgabe 41 $p = [0.131; 0.1315]$

Es handelt sich um eine dichotome Grundgesamtheit, die als unendlich groß angesehen wird. Als Hypothesen werden $H_0: \ \pi \geq \pi_0 = 0.15$ und $H_1: \ \pi < \pi_0$ aufgestellt.

Die Testgröße X (Anzahl defekter Prozessoren) ist binomialverteilt. Wegen $n\pi_0 = 15 \geq 5$ und $n(1-\pi_0) = 85 \geq 5$ kann die Approximation durch die Normalverteilung $N(n\pi_0, \sigma_X^2)$ erfolgen.

Es ist $\sigma_X = \sqrt{n\pi_0(1-\pi_0)} = 3.57$. Mit der Anzahl von 11 defekten Prozessoren ist die Wahrscheinlichkeit $p = \boldsymbol{P}(X \leq 11)$ gesucht (ohne Stetigkeitskorrektur).

$$\begin{aligned} p = \boldsymbol{P}(X \leq 11) &= \boldsymbol{P}\left(Z \leq \frac{11-15}{3.57}\right) = \boldsymbol{P}(Z \leq -1.12) \\ &= 1 - \boldsymbol{P}(Z \leq 1.12) = 1 - 0.8686 = 0.1314 \end{aligned}$$

Anmerkung: Bei einem Signifikanzniveau von 0.05 kann die Nullhypothese nicht abgelehnt werden, da $p > 0.05$ gilt.

Aufgabe 42 $1 - \alpha = 0.98$

Der Parameter σ ist unbekannt und $n < 30$. Somit ist $\frac{\overline{X}-\mu}{\hat{\sigma}_{\overline{X}}}$ t-verteilt mit $n-1 = 9$ Freiheitsgraden. Es kann von einer unendlichen Grundgesamtheit ausgegangen werden, so dass $\frac{n}{N} < 0.05$ gilt. Das Konfidenzintervall für μ wird daher wie folgt berechnet:

$$\left[\overline{X} - t_{1-\frac{\alpha}{2}}(n-1)\frac{S}{\sqrt{n}} \ ; \ \overline{X} + t_{1-\frac{\alpha}{2}}(n-1)\frac{S}{\sqrt{n}}\right]$$

Weiter ist:

$$\begin{aligned} \overline{x} &= 40.4 \qquad \sum x_i^2 = 16322.488 \\ s &= \sqrt{\frac{1}{n-1}\sum x_i^2 - n\overline{x}^2} \\ &= \sqrt{\frac{1}{9}16322.488 - 10 \cdot 40.4^2} = 0.314 \end{aligned}$$

Die Irrtumswahrscheinlichkeit α berechnet sich aus:

$$
\begin{aligned}
&\overline{x} + t_{1-\frac{\alpha}{2}}(n-1)\frac{s}{\sqrt{n}} = 40.68 \\
&\Leftrightarrow \quad t_{1-\frac{\alpha}{2}}(9) \cdot \frac{0.314}{\sqrt{10}} = 0.28 \\
&\Leftrightarrow \quad t_{1-\frac{\alpha}{2}}(9) = \frac{0.28 \cdot \sqrt{10}}{0.314} = 2.8199 \\
&\Leftrightarrow \quad 1 - \frac{\alpha}{2} = 0.99 \\
&\Leftrightarrow \quad \alpha = 0.02 \quad \text{bzw.} \quad 1 - \alpha = 0.98
\end{aligned}
$$