INSTITUT FÜR STOCHASTIK
UNIVERSITÄT KARLSRUHE
Dr. B. Klar

WS 2007/08

Klausur

Wahrscheinlichkeitstheorie und Statistik für Maschineningenieure

vom 05.03.2008

Musterlösungen

Aufgabe A1

Gegeben sei eine Urliste mit den Paaren $(x_1, y_1), \ldots, (x_{12}, y_{12})$

j	1	2	3	4	5	6	7	8	9	10	11	12
x_j	1	2.1	2.8	4.1	5.1	6.1	7	8.2	9	10.1	10.8	12.1
y_j	11.1	14.5	15.8	18.9	18.8	21.2	21.4	22.5	26.3	25.3	26	30

a) Berechnen Sie die Stichprobenmittel $\bar{x}$, $\bar{y}$, die Stichproben-Standardabweichungen s_x, s_y und den empirischen Korrelationskoeffizienten r_{xy}.
Lösung: Direkt aus den Daten ergibt sich gemäß Definition 1.8 und Paragraph 1.5 unter Ausnützung der Beziehung

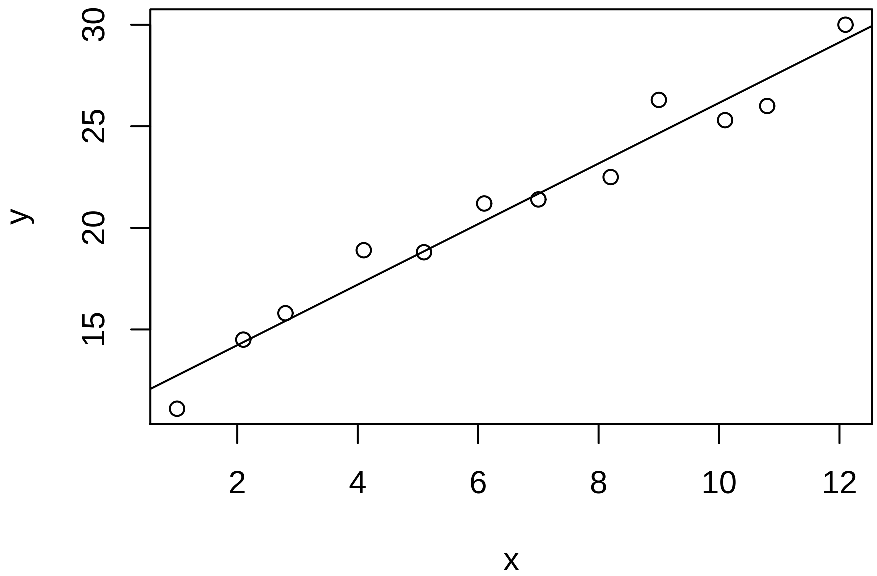
$$\sum_{j=1}^{n}(x_j - \bar{x}) \cdot (y_j - \bar{y}) = \sum_{j=1}^{n} x_j \cdot y_j - n \cdot \bar{x} \cdot \bar{y}$$

$$\bar{x} = 6.53 \qquad s_x = 3.61$$
$$\bar{y} = 20.98 \qquad s_y = 5.49$$
$$r_{xy} = 0.9802$$

b) Bestimmen Sie die zugehörige Regressionsgerade $y = a^* + b^* \cdot x$ von y auf x.
Lösung: Nach Paragraph 1.5 ist $b^* = r_{xy} \cdot \dfrac{s_y}{s_x}$ und $a^* = \bar{y} - b^* \cdot \bar{x}$, also

$$\begin{aligned} b^* &= 1.491 \\ a^* &= 11.24 \end{aligned}$$

und die Regressionsgerade $y = 11.24 + 1.491 \cdot x$.



Punkte und Regressionsgerade $y = a^* + b^* \cdot x$

Für die Lösung der nächsten drei Aufgabenteile benötigen wir die aufsteigend sortierten y-Werte. Es ist

$$y_{()} = (11.1,\ 14.5,\ 15.8,\ 18.8,\ 18.9,\ 21.2,\ 21.4,\ 22.5,\ 25.3,\ 26,\ 26.3,\ 30)$$

c) Berechnen Sie das 0.15-getrimmte Stichprobenmittel $\bar{y}_{0.15}$ von $(y_1, \ldots, y_{12})$.
Lösung: Mit $k = [12 \cdot 0.15] = 1$ ergibt sich

$$\bar{y}_{0.15} = \frac{1}{12 - 2 \cdot 1} \cdot (y_{(2)} + \ldots + y_{(11)}) = 21.07$$

d) Bestimmen Sie das Stichproben-0.3-Quantil $\tilde{y}_{0.3}$ von $(y_1, \ldots, y_{12})$.
Lösung: Da $12 \cdot 0.3 = 3.6$ nicht ganzzahlig ist, ist mit $k = [3.6] = 3$

$$\tilde{y}_{0.3} = y_{(k+1)} = y_{(4)} = 18.8$$

e) Berechnen Sie den Quartilsabstand von $(y_1, \ldots, y_{12})$.
Lösung: Da $0.25 \cdot 12 = 3$ und $0.75 \cdot 12 = 9$ beide ganzzahlig sind, ergibt sich mit $k_1 = 3$ und $k_2 = 9$

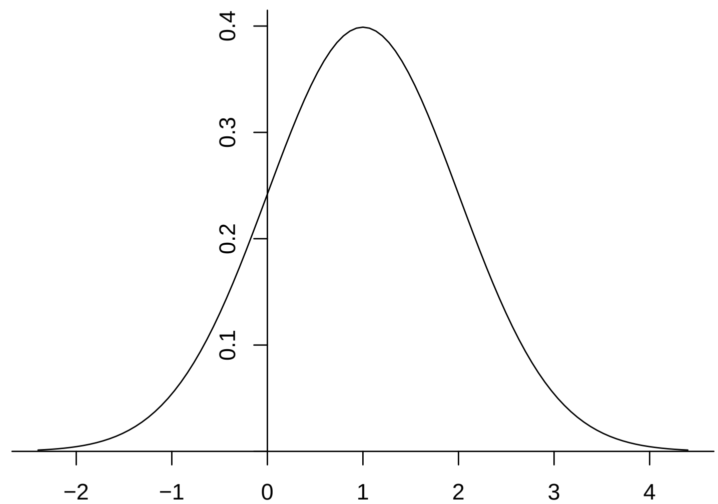
$$\begin{aligned}
\tilde{y}_{0.25} &= \frac{y_{(k_1)} + y_{(k_1+1)}}{2} = \frac{y_{(3)} + y_{(4)}}{2} = 17.3 \\
\tilde{y}_{0.75} &= \frac{y_{(k_2)} + y_{(k_2+1)}}{2} = \frac{y_{(9)} + y_{(10)}}{2} = 25.65
\end{aligned}$$

und damit der Quartilsabstand zu $\tilde{y}_{0.75} - \tilde{y}_{0.25} = 8.35$.

Aufgabe A2

Die Zufallsvariable X besitze die Verteilung $\mathcal{N}(1,1)$. Weiter sei $Y := 1 - 2X$.

a) Skizzieren Sie die Dichte $f(x)$ von X für $-2 \le x \le 4$.
Lösung: Die Dichte von X ist $f(x) = \frac{1}{\sqrt{2\pi}} \cdot \exp\left(-\frac{(x-1)^2}{2}\right)$:



b) Berechnen Sie den Erwartungswert $\mathbb{E}Y$ und die Varianz $V(Y)$.
Lösung: Mit $\mathbb{E}X = 1$ und $V(X) = 1$ folgt

$$\begin{aligned} \mathbb{E}Y &= 1 - 2\mathbb{E}X = -1, \\ V(Y) &= (-2)^2 V(X) = 4. \end{aligned}$$

c) Welche Verteilung besitzt Y?
Lösung: Y ist $\mathcal{N}(-1,4)$-verteilt.

d) Bestimmen Sie die Wahrscheinlichkeit $\mathbb{P}(-1 \le Y \le 3)$.
Lösung: Es ist

$$\begin{aligned} \mathbb{P}(-1 \le Y \le 3) &= \mathbb{P}(Y \le 3) - \mathbb{P}(Y \le -1) \\ &= \Phi\left(\frac{3-(-1)}{2}\right) - \Phi\left(\frac{-1-(-1)}{2}\right) = \Phi(2) - \Phi(0). \end{aligned}$$

Dabei ist Φ die Verteilungsfunktion von $\mathcal{N}(0,1)$. Wegen $\Phi(2) = 0.9772$ und $\Phi(0) = 0.5$ (aus Tabelle A.1) folgt

$$\mathbb{P}(-1 < Y \le 3) = 0.9772 - 0.5 = 0.4772.$$

e) Bestimmen Sie das 0.1-Quantil $q_{0.1}$ der Zufallsvariablen Y.
Lösung: Die Verteilungsfunktion von Y sei mit F_Y bezeichnet. Dann ist $q_{0.1}$ die Lösung q der Gleichung

$$F_Y(q) = \Phi\left(\frac{q-(-1)}{2}\right) = 0.1.$$

Wegen $\Phi(1.28) = 0.9$ und $\Phi(-x) = 1 - \Phi(x)$ gilt $\Phi(-1.28) = 0.1$ und somit

$$\frac{q+1}{2} = -1.28$$

Hieraus folgt $q_{0.1} = q = 2 \cdot (-1.28) - 1 = -3.56$.

f) Bestimmen Sie die Kovarianz $C(X,Y)$ und den Korrelationskoeffizienten $\rho(X,Y)$.
Lösung: Es gilt

$$\begin{aligned} C(X,Y) &= C(X, 1-2X) = C(X,-2X) = -2\,C(X,X) = -2\,V(X) = -2, \\ \rho(X,Y) &= C(X,Y)/\sqrt{V(X)V(Y)} = (-2)/\sqrt{1 \cdot 4} = -1. \end{aligned}$$

Aufgabe A3

X sei eine Zufallsvariable mit Werten in $\{1, 2\}$ und Y eine Zufallsvariable mit Werten in $\{0, 1, c\}$. Hierbei sei c eine feste Zahl aus $\{2, 3, 4, \ldots\}$. Die folgende Tabelle gibt die gemeinsame Verteilung $\mathbb{P}(X = i, Y = j)$ des Zufallsvektors (X, Y) für die Werte $i = 1, 2$ und $j = 0, 1, c$ an.

	$j = 0$	$j = 1$	$j = c$
$i = 1$	$\frac{1}{4}$	$\frac{1}{8}$	$\frac{1}{8}$
$i = 2$	$\frac{1}{8}$	$\frac{1}{8}$	$\frac{1}{4}$

a) Berechnen Sie die Randverteilung von Y, d.h. $P(Y = 0), P(Y = 1)$ und $P(Y = c)$, und den Erwartungswert $\mathbb{E}Y$. Für welche c gilt $\mathbb{E}Y = 4$?
Lösung:

$$\mathbb{P}(Y=0) = \frac{3}{8}, \quad \mathbb{P}(Y=1) = \frac{2}{8}, \quad \mathbb{P}(Y=c) = \frac{3}{8}$$

$$\mathbb{E}Y = 0 \cdot \frac{3}{8} + 1 \cdot \frac{2}{8} + c \cdot \frac{3}{8} = (3c+2)/8$$

Für $c = 10$ ist $\mathbb{E}X = 4$.

b) Berechnen Sie die Wahrscheinlichkeit $\mathbb{P}(X = 2, Y > 0)$.
Lösung:

$$\mathbb{P}(X=2, Y>0) = \mathbb{P}(X=2, Y=1) + \mathbb{P}(X=2, Y=c) = \frac{1}{8} + \frac{1}{4} = \frac{3}{8}$$

c) Berechnen Sie die bedingte Wahrscheinlichkeit $\mathbb{P}(X = 2|Y > 0)$.
Lösung:

$$\mathbb{P}(X=2|Y>0) = \frac{\mathbb{P}(X=2, Y>0)}{\mathbb{P}(Y>0)} = \frac{\mathbb{P}(X=2, Y>0)}{\mathbb{P}(Y=1) + \mathbb{P}(Y=c)}$$
$$= \frac{\frac{1}{8} + \frac{2}{8}}{\frac{2}{8} + \frac{3}{8}} = \frac{3}{5}$$

d) Es sei $c = 2$ sowie $Z := X \cdot Y$. Berechnen Sie den Erwartungswert $\mathbb{E}Z$, das zweite Moment $\mathbb{E}Z^2$ und die Varianz $V(Z)$.
Lösung: Es sei $f(i, j) := \mathbb{P}(X = i, Y = j)$. Dann gilt

$$\mathbb{E}Z = \mathbb{E}[X \cdot Y] = \sum_{i,j:\, f(i,j)>0} i \cdot j \cdot f(i,j) = 1 \cdot \frac{1}{8} + 2 \cdot \frac{1}{8} + 2 \cdot \frac{1}{8} + 4 \cdot \frac{1}{4} = \frac{13}{8},$$

$$\mathbb{E}Z^2 = \mathbb{E}[X^2 \cdot Y^2] = \sum_{i,j:\, f(i,j)>0} i^2 \cdot j^2 \cdot f(i,j) = 1 \cdot \frac{1}{8} + 4 \cdot \frac{1}{8} + 4 \cdot \frac{1}{8} + 16 \cdot \frac{1}{4}$$
$$= \frac{41}{8},$$

$$V(Z) = \mathbb{E}Z^2 - (\mathbb{E}Z)^2 = \frac{41}{8} - \frac{169}{64} = \frac{159}{64}.$$

e) Es sei wieder $c = 2$. Sind X und Y unabhängig? Begründen Sie Ihre Antwort!
Lösung: X und Y sind nicht unabhängig; z.B. gilt

$$\mathbb{P}(X=1, Y=0) = \frac{1}{4} \neq \mathbb{P}(X=1) \cdot \mathbb{P}(Y=0) = \frac{1}{2} \cdot \frac{3}{8}.$$

Aufgabe A4

X und Y seien die zufälligen Wartezeiten [in Minuten] von zwei Kunden A und B, die an unterschiedlichen Kassen stehen. Wir nehmen an, dass X und Y stochastisch unabhängig und jeweils exponentialverteilt mit Parameter 0.5 sind.

a) Mit welcher Wahrscheinlichkeit wartet Kunde A länger als 4 Minuten?
Lösung: Die Verteilungsfunktion der Wartezeit X von Kunde A ist $F_X(t) = 1 - \exp(-0.5t)$ für $t > 0$. Somit gilt:

$$\mathbb{P}(X > 4) = 1 - F_X(4) \;=\; e^{-2} \;=\; 0.135.$$

b) Wie groß ist die Wahrscheinlichkeit, dass beide Kunden länger als 4 Minuten warten müssen?
Lösung: Die Unabhängigkeit von X, Y liefert

$$\mathbb{P}(X > 4, Y > 4) = \mathbb{P}(X > 4) \cdot \mathbb{P}(Y > 4) \;=\; e^{-4} \;=\; 0.018.$$

c) Wie groß ist die Wahrscheinlichkeit, dass mindestens eine Kunde länger als 4 Minuten warten muss?
Lösung: Es ist

$$\begin{aligned}\mathbb{P}(\max(X,Y) > 4) &= 1 - \mathbb{P}(X \le 4, Y \le 4) \;=\; 1 - \mathbb{P}(X \le 4) \cdot \mathbb{P}(Y \le 4) \\ &= 1 - \left(1 - e^{-2}\right)^2 \;=\; e^{-2}\left(2 - e^{-2}\right) \;=\; 0.252.\end{aligned}$$

d) Geben Sie die gemeinsame Dichte von X und Y an.
Lösung: Die Dichte von X bzw. Y ist $f_X(t) = f_Y(t) = 0.5\exp(-0.5t)$ für $t > 0$. Wegen der Unabhängigkeit von X und Y ist

$$f(x,y) \;=\; f_X(x) \cdot f_Y(y) \;=\; \begin{cases} 0.25e^{-0.5(x+y)}, & x, y > 0 \\ 0, & \text{sonst} \end{cases}$$

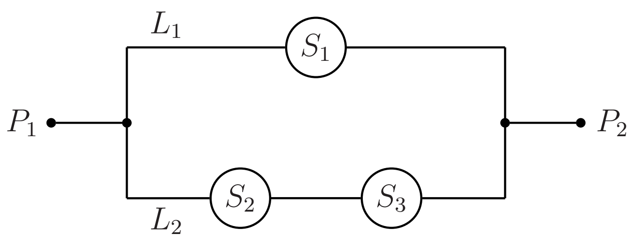
(eine) gemeinsame Dichte von X und Y.

e) Wie groß ist die Wahrscheinlichkeit, dass Kunde B mindestens doppelt so lange wie Kunde A wartet?
Hinweis: $\int ae^{-ax}dx = -e^{-ax} + c$.
Lösung: Es ist mit $B = \{(x,y) \in \mathcal{R}^2 : y \ge 2x\}$

$$\begin{aligned}\mathbb{P}(Y \ge 2X) &= \mathbb{P}((X,Y) \in B) \;=\; \int_B f(x,y)\, d(x,y) \\ &= \int_0^\infty \left(\int_{2x}^\infty f_Y(y)\, dy \right) f_X(x)\, dx \\ &= \int_0^\infty \left(\int_{2x}^\infty 0.5\; e^{-0.5y}\, dy \right) 0.5\; e^{-0.5x}\, dx \\ &= \int_0^\infty e^{-x}\; 0.5\; e^{-0.5x} dx = \left| -\frac{e^{-3x/2}}{3} \right|_0^\infty = \frac{1}{3}.\end{aligned}$$

Aufgabe A5

Zwischen zwei Punkten P_1 und P_2 verläuft folgendes Leitungsnetz:



Dabei sind S_1, S_2, S_3 störanfällige Stellen. Leitung L_1 läuft durch die störanfällige Stelle S_1, Leitung L_2 läuft durch die störanfälligen Stellen S_2 und S_3. Die Zufallsvariablen X_i seien definiert als

$$X_i = \begin{cases} 0, & S_i \text{ ist unterbrochen,} \\ 1, & S_i \text{ ist nicht unterbrochen,} \end{cases} \qquad i = 1, 2, 3.$$

Ferner seien X_1, X_2, X_3 stochastisch unabhängig mit $\mathbb{P}(X_i = 1) = p$, $i = 1, 2, 3$, für ein p mit $0 < p < 1$.
A sei das Ereignis, dass Leitung L_1 frei ist, d.h. dass S_1 nicht unterbrochen ist. B sei das Ereignis, dass Leitung L_2 frei ist.

a) Wie groß ist $\mathbb{P}(B)$, also die Wahrscheinlichkeit, dass L_2 frei ist?
Lösung: Wegen der Unabhängigkeit von X_2 und X_3 gilt

$$\mathbb{P}(B) \;=\; \mathbb{P}(\{X_2 = 1\} \cap \{X_3 = 1\}) \;=\; \mathbb{P}(X_2 = 1) \cdot \mathbb{P}(X_3 = 1) \;=\; p^2.$$

b) Wie groß ist die Wahrscheinlichkeit, dass L_2 frei ist, unter der Bedingung, dass L_1 frei ist?
Lösung: Wegen der Unabhängigkeit von X_1, X_2 und X_3 sind auch A und B unabhängig, und es gilt

$$\mathbb{P}(B|A) = \mathbb{P}(B) \;=\; p^2.$$

c) Wie groß ist die Wahrscheinlichkeit, dass L_2 frei ist, aber L_1 nicht?
Lösung:

$$\mathbb{P}(A^c \cap B) \;=\; \mathbb{P}(A^c) \cdot \mathbb{P}(B) \;=\; (1-p)p^2.$$

d) Wie groß ist die Wahrscheinlichkeit, dass genau eine der beiden Leitungen frei ist?
Lösung: Mit $\mathbb{P}(A \cap B^c) = \mathbb{P}(A) \cdot \mathbb{P}(B^c) = p(1-p^2)$ folgt

$$\mathbb{P}((A \cap B^c) \cup (A^c \cap B)) \;=\; \mathbb{P}(A \cap B^c) + \mathbb{P}(A^c \cap B) \;=\; p(1-p^2) + (1-p)p^2 \;=\; p(1-p)(1+2p).$$

e) C sei das Ereignis {„P_1 ist mit P_2 verbunden“}. Bestimmen Sie $\mathbb{P}(C)$.
Lösung: Mit $\mathbb{P}(A \cap B) = \mathbb{P}(A) \cdot \mathbb{P}(B) = p \cdot p^2 = p^3$ folgt

$$\mathbb{P}(C) \;=\; \mathbb{P}(A \cup B) \;=\; \mathbb{P}(A) + \mathbb{P}(A) - \mathbb{P}(A \cap B) \;=\; p + p^2 - p^3.$$

f) Die Zufallsvariable $Y := \min\{X_2, X_3\}$ besitzt die Binomial-Verteilung $Bin(n, q)$. Bestimmen Sie die Parameter n und q.
Lösung: Die Zufallsvariable Y nimmt die Werte 0 und 1 an. Wegen

$$\mathbb{P}(Y = 1) \;=\; \mathbb{P}(\min\{X_2, X_3\} = 1) \;=\; \mathbb{P}(X_2 = 1, X_3 = 1) \;=\; \mathbb{P}(X_2 = 1) \cdot \mathbb{P}(X_3 = 1) \;=\; p^2$$

gilt $\mathbb{P}(Y = 0) = 1 - p^2$. Somit ist Y $Bin(1, p^2)$-verteilt.

Aufgabe A6

Es soll der unbekannte Parameter $\vartheta \in \Theta = (0, \infty)$ für die Verteilung mit der Dichte

$$f(x;\vartheta) = \frac{1}{\sqrt{2\pi\vartheta}} \exp\left(-\frac{x^2}{2\vartheta}\right), \qquad x \in \mathbb{R}.$$

bestimmt werden.

a) Geben Sie die zur Stichprobe $x = (x_1, \ldots, x_n)$ gehörende Likelihood-Funktion $L_x(\vartheta)$ an.
Lösung: Es ist

$$L_x(\vartheta) \;=\; \prod_{i=1}^{n} \frac{1}{\sqrt{2\pi\vartheta}} \exp\left(-\frac{x_i^2}{2\vartheta}\right) \;=\; \left(\frac{1}{2\pi\vartheta}\right)^{n/2} \exp\left(-\frac{1}{2\vartheta}\sum_{i=1}^{n} x_i^2\right).$$

b) Berechnen Sie die Loglikelihood-Funktion $M_x(\vartheta) := \log L_x(\vartheta)$.
Lösung: Folglich ist

$$M_x(\vartheta) \;=\; \log L_x(\vartheta) \;=\; -\frac{n}{2}\log(2\pi) - \frac{n}{2}\log\vartheta - \frac{1}{2\vartheta}\sum_{i=1}^{n} x_i^2$$

c) Berechnen Sie den Maximum-Likelihood-Schätzer $\hat{\vartheta}(x)$ für ϑ.
Lösung: Für die Ableitung von M_x nach ϑ gilt:

$$\frac{d}{d\vartheta} M_x(\vartheta) \;=\; -\frac{n}{2\vartheta} + \frac{1}{2\vartheta^2}\sum_{i=1}^{n} x_i^2 = 0 \quad \Leftrightarrow \quad \vartheta = \frac{1}{n}\sum_{i=1}^{n} x_i^2.$$

Weiter gilt mit $\hat{\vartheta}_n := \frac{1}{n}\sum_{i=1}^{n} x_i^2$

$$\frac{d^2}{d\vartheta^2} M_x(\hat{\vartheta}_n) \;=\; \frac{n}{2\hat{\vartheta}_n^2} - \frac{n}{\hat{\vartheta}_n^2} < 0.$$

Somit liegt eine lokale Maximalstelle vor, und

$$\hat{\vartheta}_n(x) = \frac{1}{n}\sum_{i=1}^{n} x_i^2$$

ist der gesuchte Maximum-Likelihood-Schätzer für ϑ.

Aufgabe A7

Eine Münze mit den Symbolen „Zahl" und „Wappen" soll auf ihre Echtheit überprüft werden, d.h es soll überprüft werden ob

$$\mathbb{P}(\{\text{„Zahl" geworfen}\}) \;=\; \mathbb{P}(\{\text{„Wappen" geworfen}\}) \;=\; \frac{1}{2}$$

gilt. Dem folgenden Zufallsexperiment werden die Zufallsvariablen

$$X_i \;=\; \begin{cases} 1, & \text{falls „Zahl" geworfen,} \\ 0, & \text{falls „Wappen" geworfen,} \end{cases} \qquad i = 1, \dots, n,$$

zugrunde gelegt. Dabei seien $X_1, X_2, \dots, X_n$ stochastisch unabhängig und jeweils $Bin(1, \vartheta)$-verteilt. Der Parameter $0 < \vartheta < 1$ sei unbekannt.

a) Geben Sie Erwartungswert, Varianz und Standardabweichung des Stichproben-Mittels $\bar{X}_n = \frac{1}{n}\sum_{i=1}^{n} X_i$ an.
Lösung: Es ist

$$\begin{aligned}
\mathbb{E}\left(\bar{X}_n\right) &= \frac{1}{n}\sum_{i=1}^{n} E(X_i) \;=\; E(X_1) \;=\; \vartheta, \\
V\left(\bar{X}_n\right) &= \frac{1}{n^2}\sum_{i=1}^{n} V(X_i) \;=\; \frac{V(X_1)}{n} \;=\; \frac{\vartheta(1-\vartheta)}{n}, \\
\sigma_{\bar{X}_n} &= \sqrt{V\left(\bar{X}_n\right)} \;=\; \sqrt{\frac{\vartheta(1-\vartheta)}{n}}.
\end{aligned}$$

b) Nun wird die Münze $n = 2000$ Mal geworfen und 965 Mal das Symbol „Zahl" beobachtet.

(i) Bestimmen Sie das Stichproben-Mittel $\bar{X}_n = \frac{1}{n}\sum_{i=1}^{n} X_i$.
Lösung: Mit $\sum_{i=1}^{n} X_i = 965$ erhält man $\bar{X}_n = \frac{965}{2000} = 0.4825$.

(ii) Geben Sie ein konkretes asymptotisches Konfidenzintervall für ϑ zum Konfidenzniveau 0.99 an.
Lösung: Mit 18.10 Beispiel und 18.11 Bemerkung 1. folgt, dass $[l_n^*, L_n^*]$ ein asymptotisches Konfidenzintervall zum Konfidenzniveau $1-\alpha$ ist, wobei

$$\begin{aligned}
l_n^* &= l_n^*(X_1, \dots, X_n) \;=\; \bar{X}_n - \frac{h}{\sqrt{n}}\sqrt{\bar{X}_n(1-\bar{X}_n)}\;, \\
L_n^* &= L_n^*(X_1, \dots, X_n) \;=\; \bar{X}_n + \frac{h}{\sqrt{n}}\sqrt{\bar{X}_n(1-\bar{X}_n)}\;,
\end{aligned}$$

mit $h = u_{1-\alpha/2} = \Phi^{-1}(1-\alpha/2)$. Für $\alpha = 0.01$ ist $h = u_{0.995} = 2.5758$ (aus Tabelle A.1). Mit $n = 2000$ und $\bar{X}_n = 0.4825$ erhält man $l_n^* = 0.454$ und $L_n^* = 0.511$.