Prof. Dr. A. Steland
N. Mause

RWTH Aachen, SS2023

Ausgabe: 27. Juni 2023 — Besprechung: 3. Juli 2023

Einführung in die angewandte Stochastik

Übungsblatt 9

Aufgabe 36

An einer Klausur haben 15 Studierende teilgenommen. Es konnten nur ganzzahlige Punkte erreicht werden, die Maximalpunktzahl betrug 20 Punkte und es ergaben sich folgende Punktzahlen:

$$5, 10, 17, 12, 10, 13, 20, 13, 5, 13, 17, 17, 10, 20, 17$$

(a) Geben Sie die Ordnungsstatistik an und bestimmen Sie die absoluten und relativen Häufigkeiten.

(b) Bestimmen und skizzieren Sie die zugehörige empirische Verteilungsfunktion.

(c) Zum Bestehen der Klausur waren 8 Punkte nötig, ab 17 Punkten gab es die Note *gut*. Bestimmen Sie mit Hilfe der empirischen Verteilungsfunktion die relative Häufigkeit dafür, dass ein Student

(i) die Klausur nicht bestanden hat,
(ii) mindestens die Note *gut* erhielt,
(iii) bestanden hat, aber eine schlechtere Note als *gut* erhielt.

Lösung: a)

i	1	2	3	4	5	6	7	8	9	10	11	12	13
x_i	5	10	17	12	10	13	20	13	5	13	17	17	10
$x_{(i)}$	5	5	10	10	10	12	13	13	13	17	17	17	17

i	14	15
x_i	20	17
$x_{(i)}$	20	20

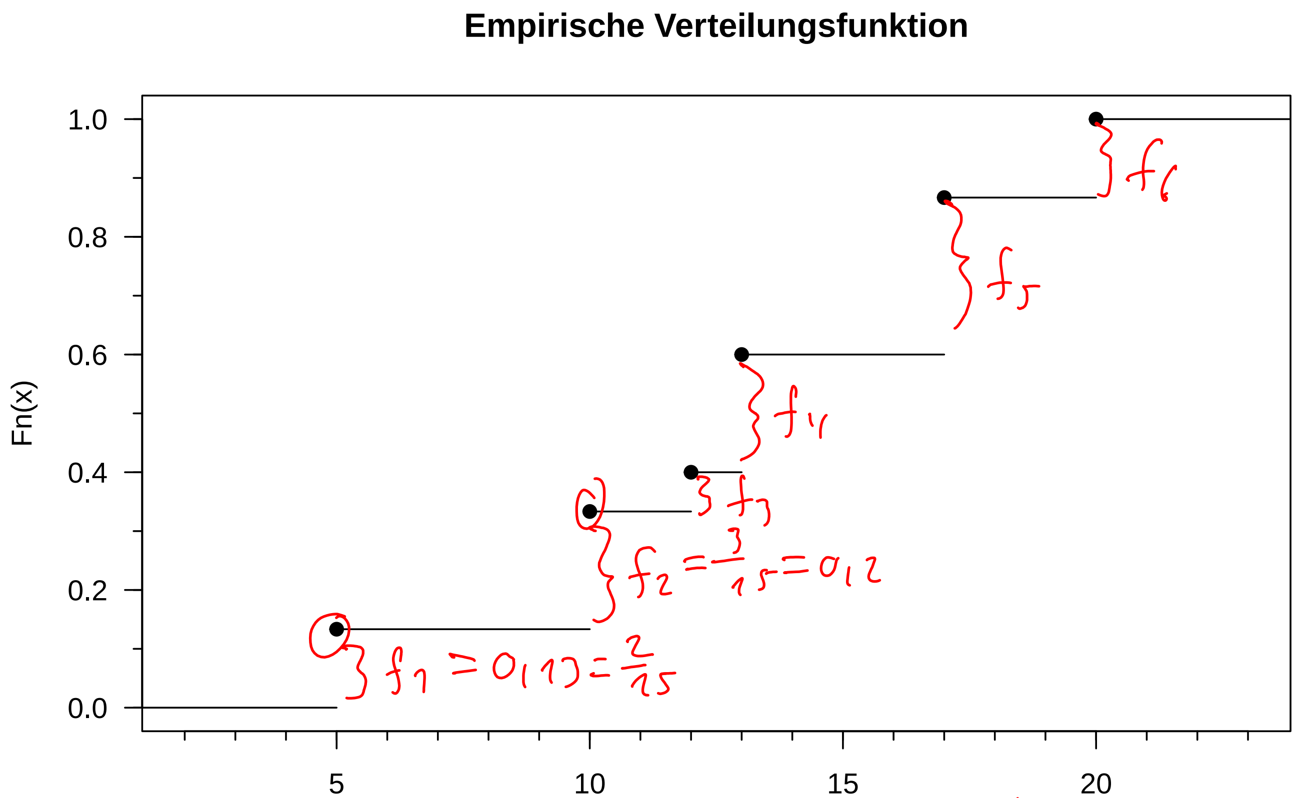
Ordnungsstatistik:
Sortiere die Daten aufsteigend

Punkte	abs. Häufigkeit	rel. Häufigkeiten
5	2	$2/15 \approx 0.13$
10	3	$3/15 = \frac{1}{5} = 0,2$
12	1	$\frac{1}{15} \approx 0.07$
13	3	$3/15 = \frac{1}{5} = 0,2$
17	4	$4/15 \approx 0,27$
20	2	$2/15 \approx 0.13$
Σ	$n = 15$	1

b) Die emp. VF ist die Funktion, die jeden $x \in \mathbb{R}$ den Anteil der Stichprobenwerte zuordnet, die kleiner oder gleich x sind.

$$\hat{F}(x) = \frac{1}{15} \sum_{i=1}^{15} \mathbb{1}_{(-\infty, x]}(x_i) \qquad \mathbb{1}_{(-\infty, x]}(x_i) = \begin{cases} 1, & x_i \leq x \\ 0, & x_i > x \end{cases}$$

$$= \begin{cases} 0, & x < 5 \\ \frac{2}{15} \approx 0{,}13, & 5 \leq x < 10 \\ \frac{5}{15} \approx 0{,}33, & 10 \leq x < 12 \\ 6/15 \approx 0{,}4, & 12 \leq x < 13 \\ 9/15 = 0{,}6, & 13 \leq x < 17 \\ 13/15 \approx 0{,}87, & 17 \leq x < 20 \\ \frac{15}{15} = 1, & 20 \leq x \end{cases}$$



c) Relative Häufigkeit dafür, dass ein Student

(i) die Klausur nicht bestanden hat ("<8" $=$ "≤ 7")

$\hat{F}(7) = \frac{2}{15} \approx 0.13$

(ii) mindestens die Note gut erhalten hat, d.h. mindestens 17 Punkte hat.

$\boxed{\text{Überlegung}: P(X \geq 17) = 1 - P(X < 17) = 1 - P(X \leq 16) = 1 - F_X(16)}$

$1 - \hat{F}(16) = 1 - \frac{9}{15} = 1 - \frac{3}{5} = \frac{2}{5}$

(iii) bestanden hat, aber Note schlechter als gut hat

Überlegung: $(-\infty, 16]$ ist disjunkt zerlegbar in $(-\infty, 7]$ und $(7, 16]$

$\Rightarrow P(X \in (-\infty, 16]) = P(X \in (-\infty, 7]) + P(X \in (7, 16])$
$= P(X \leq 7) + P(7 < X \leq 16)$

$$\Leftrightarrow \quad P(7 < X \leq 16) = P(X \leq 16) - P(X \leq 7)$$
$$= F_X(16) - F_X(7)$$

$$\hat{F}(16) - \hat{F}(7) = \frac{9}{15} - \frac{2}{15} = \frac{7}{15}$$

Aufgabe 37

Die Zufallsvariablen $X_1, \ldots, X_n$ seien stochastisch unabhängig und jeweils Pareto–verteilt mit (unbekanntem) Parameter $\alpha > 0$. Die zugehörige Dichtefunktion f_α und die zugehörige Verteilungsfunktion F_α der Zufallsvariablen X_i für $i \in \{1, \ldots, n\}$ in Abhängigkeit vom Parameter α sind gegeben durch

$$f_\alpha(x) = \begin{cases} \frac{\alpha}{x^{\alpha+1}} & , \quad x \geq 1 \,, \\ 0 & , \quad x < 1 \end{cases} \qquad \text{bzw.} \qquad F_\alpha(x) = \begin{cases} 1 - \frac{1}{x^{\alpha}} & , \quad x \geq 1 \,, \\ 0 & , \quad x < 1 \,. \end{cases}$$

Bestimmen Sie eine Maximum–Likelihood–Schätzung $\widehat{\alpha}$ für den Parameter α.

Lösung (1) Aufgrund der Unabhängigkeit der ZV $X_1,..,X_n$ ist die gemeinsame Dichte $f_\alpha(x_1..x_n)$ als Produkt der einzelnen Randdichten $f_\alpha(x_i)$ darstellbar, d.h. es gilt

$$f_\alpha(x_1,..,x_n) \overset{\text{unabh.}}{=} \prod_{i=1}^{n} f_\alpha(x_i) \overset{\text{Def } f_\alpha}{=} \prod_{i=1}^{n} \frac{\alpha}{x^{\alpha+1}} \underbrace{\mathbb{1}(x_i \geq 1)}_{= \begin{cases} 1, & x_i \geq 1 \\ 0, & x_i < 1 \end{cases}}$$

$$= \alpha^n \prod_{i=1}^{n} x^{-(\alpha+1)} \mathbb{1}(x_i \geq 1) \,, \quad \text{für } x_1,..,x_n \in \mathbb{R}.$$

Seien nun $x_1,..,x_n$ Realisationen von den ZV $X_1,..,X_n$ für alle $i=1..n$ mit $x_i \geq 1$ $\forall i=1..n$ fest gegeben.
Die Likelihood-Funktion ist nun gegeben durch

$$L(\alpha|x_1..x_n) = f_\alpha(x_1..x_n) = \alpha^n \prod_{i=1}^{n} x_i^{-(\alpha+1)} \,, \quad \alpha > 0$$

[$L(\alpha|x_1..x_n)$: Funktion von α bei festen Daten $x_1..x_n$
$f_\alpha(x_1..x_n)$: Funktion von $x_1..x_n$ bei festem α]

(2) Ziel ist nun die globale Maximierung der Likelihood-Funktion.
Da der Logarithmus eine streng monoton wachsende Funktion ist, ist die Maximierung der Likelihood-Funktion äquivalent zur Maximierung der Log-Likelihood-Funktion. Diese ist gegeben durch

$$\ell(\alpha) = \log(L(\alpha|x_1..x_n)) = \log\left(\alpha^n \prod_{i=1}^{n} x_i^{-(\alpha+1)}\right)$$

[$\log(a \cdot b) = \log(a) + \log(b)$, $\log(a^n) = n \cdot \log(a)$
$\log(\prod a_i) = \sum \log(a_i)$]

$$= n \cdot \log(\alpha) + \sum_{i=1}^{n} -(\alpha+1)\log(x_i)$$

$$= n \cdot \log(\alpha) - (\alpha+1)\sum_{i=1}^{n} \log(x_i), \quad \alpha > 0$$

3) Die Log-Likelihood-Funktion ℓ ist differenzierbar auf $(0,\infty)$ mit

$$\ell'(\alpha) = \frac{d\,\ell(\alpha)}{d\alpha} = \frac{n}{\alpha} - \sum_{i=1}^{n} \log(x_i), \quad \alpha > 0$$

und es gilt

$$\ell'(\alpha) \overset{!}{=} 0 \iff \frac{n}{\alpha} = \sum_{i=1}^{n} \log(x_i) \iff \alpha = \underbrace{\frac{n}{\sum_{i=1}^{n} \log(x_i)}}_{=:\hat{\alpha}}.$$

$\Rightarrow$ $\hat{\alpha}$ ist ein stationärer Punkt der Log-Likelihood-Funktion ℓ

4) Weiterhin haben wir

$$\ell'(\alpha) > 0 \iff \alpha < \frac{n}{\sum_{i=1}^{n} \log(x_i)} \iff \alpha < \hat{\alpha}$$

$$\overset{\ell \text{ stetig}}{\underset{\text{Monotoniekriterium}}{\Longrightarrow}} \ell(\alpha) \text{ streng monoton wachsend auf } (0, \hat{\alpha})$$

und

$$\ell'(\alpha) < 0 \iff \alpha > \frac{n}{\sum_{i=1}^{n} \log(x_i)} \iff \alpha > \hat{\alpha}$$

$$\overset{\ell \text{ stetig}}{\underset{\text{Monotoniekriterium}}{\Longrightarrow}} \ell(\alpha) \text{ ist streng monoton fallend auf } (\hat{\alpha}, \infty)$$

$\Rightarrow$ $\hat{\alpha}$ ist globales Maximum von ℓ und L.

$\Rightarrow$ $\hat{\alpha}$ ist der Maximum-Likelihood-Schätzer für α bei gegebenen Daten $x_1, \dots, x_n$.

MLE-Schema

Gegeben: unabhängige ZV $X_1, \dots, X_n$ mit Dichtefunktion f_{ϑ}, ϑ unbekannte Parameter

(1) Likelihood-Funktion $L(\vartheta | x_1,\dots,x_n)$ aufstellen bei gegebenen Realisationen der ZV $X_1,\dots,X_n$ (dabei Unabh. ausnutzen!)

(2) Log-Likelihood-Fkt aufstellen, $\ell(\vartheta) = \log(L(\vartheta | x_1,\dots,x_n))$

(3) stationäre Punkte berechnen, d.h. $\ell'(\vartheta) \overset{!}{=} 0$

$\hookrightarrow$ liefert Kandidaten für lokale Extrema

(4) Zeige, dass (einer der gefundenen) stationäre Punkte ein globales Maximum ist, z.B. mit dem Monotoniekriterium.

Aufgabe 38

Seien X_1, X_2 und X_3 i.i.d. Zufallsvariablen auf einem Wahrscheinlichkeitsraum $(\Omega, \mathcal{F}, P)$ mit $X_i \sim \mathcal{N}(\mu, 1), i = 1, 2, 3$, wobei $\mu \in \mathbb{R}$ unbekannt sei. Betrachten Sie die folgenden beiden Schätzer für μ:

$$\hat{\mu}_1 = \frac{1}{3}(X_1 + 2X_2), \quad \hat{\mu}_2 = \frac{1}{3}\sum_{i=1}^{3} X_i.$$

(a) Sind $\hat{\mu}_1$ und $\hat{\mu}_2$ erwartungstreu für μ?

(b) Berechnen Sie für $\hat{\mu}_1$ und $\hat{\mu}_2$ jeweils die Varianz und entscheiden Sie auf Basis Ihrer Ergebnisse, welcher der beiden Schätzer für eine Schätzung von μ verwendet werden sollte.

Lösung: a)

$$E_\mu(\hat{\mu}_1) = E_\mu\left(\frac{1}{3}(X_1 + 2X_2)\right) = \frac{1}{3}E_\mu(X_1) + \frac{2}{3}E_\mu(X_2)$$
$$= \frac{1}{3}\cdot\mu + \frac{2}{3}\cdot\mu = \mu$$

$$E_\mu(\hat{\mu}_2) = E_\mu\left(\frac{1}{3}\sum_{i=1}^{3} X_i\right) = \frac{1}{3}\sum_{i=1}^{3} E_\mu(X_i) = \frac{1}{3}\cdot 3\cdot \mu = \mu$$

$\Rightarrow$ $\hat{\mu}_1$ und $\hat{\mu}_2$ sind erwartungstreu

b) Nach Vor. gilt $Var(X_i) = 1 \ \forall\, i = 1, 2, 3$

$$\Rightarrow Var_\mu(\hat{\mu}_1) = Var_\mu\left(\frac{1}{3}(X_1 + 2X_2)\right) \underset{\text{unabh.}}{\overset{\text{stoch.}}{=}} \frac{1}{9} Var_\mu(X_1) + \frac{4}{9} Var_\mu(X_2)$$
$$= \frac{1}{9} + \frac{4}{9} = \frac{5}{9}$$

und

$$Var_\mu(\hat{\mu}_2) = Var_\mu\left(\frac{1}{3}\sum_{i=1}^{3} X_i\right) \overset{\text{unabh.}}{=} \frac{1}{9}\sum_{i=1}^{3} \underbrace{Var(X_i)}_{=1} = \frac{3}{9} = \frac{1}{3}$$

Da beide Schätzer erw.-treu sind, ist der Schätzer mit der kleineren Varianz effizienter. Es gilt

$$Var(\hat{\mu}_2) < Var(\hat{\mu}_1)$$

$\Rightarrow$ $\hat{\mu}_2$ ist effizienter und sollte verwendet werden.

Aufgabe 39

Seien X_1, X_2 i.i.d. Zufallsvariablen auf einem Wahrscheinlichkeitsraum $(\Omega, \mathcal{F}, P)$ mit $X_i \sim \mathrm{Exp}(\lambda), i = 1, 2$, wobei $\lambda > 0$ unbekannt sei. Betrachten Sie den folgenden Schätzer für λ:

$$\hat{\lambda} = \frac{1}{2}(X_1 + X_2)$$

Bestimmen Sie den mittleren quadratischen Fehler $\mathrm{MSE}(\hat{\lambda}, \lambda)$.

Lösung:

$$\mathrm{MSE}(\hat\lambda, \lambda) = \mathrm{Var}(\hat\lambda) + \mathrm{Bias}^2(\hat\lambda, \lambda)$$
$$= \mathrm{Var}(\hat\lambda) + (E(\hat\lambda) - \lambda)^2.$$

Hier ist $X_i \sim \mathrm{Exp}(\lambda)$

$$\Rightarrow E(\hat\lambda) = E\left(\frac{1}{2}(X_1 + X_2)\right) = \frac{1}{2}\left(E(X_1) + E(X_2)\right)$$
$$= \frac{1}{2}\cdot\left(\frac{1}{\lambda} + \frac{1}{\lambda}\right) = \frac{1}{\lambda}$$

$\Rightarrow$ Schätzer ist <u>nicht</u> erwartungstreu, außer wenn $\lambda = 1$.

$\Rightarrow$ Schätzer hat Verzerrung (Bias)

$$\Rightarrow \mathrm{Bias}(\hat\lambda, \lambda) = E(\hat\lambda) - \lambda = \frac{1}{\lambda} - \lambda$$

Weiterhin gilt

$$\mathrm{Var}(\hat\lambda) = \mathrm{Var}\left(\frac{1}{2}(X_1 + X_2)\right) = \frac{1}{4}\cdot\left(\mathrm{Var}(X_1) + \mathrm{Var}(X_2)\right)$$
$$\overset{X_i \sim \mathrm{Exp}(\lambda)}{=} \frac{1}{4}\cdot\left(\frac{1}{\lambda^2} + \frac{1}{\lambda^2}\right) = \frac{1}{2\lambda^2}$$

$$\Rightarrow \mathrm{MSE}(\hat\lambda, \lambda) = \frac{1}{2\lambda^2} + \left(\frac{1}{\lambda} - \lambda\right)^2$$
$$= \frac{1}{2\lambda^2} + \frac{1}{\lambda^2} - 2 + \lambda^2$$
$$= \lambda^2 + \frac{3}{2\lambda^2} - 2$$

Aufgabe 40

Seien $X_1, \dots, X_n \overset{\text{i.i.d.}}{\sim} \mathcal{N}(\mu_1, \sigma^2)$ und $Y_1, \dots, Y_n \overset{\text{i.i.d.}}{\sim} \mathcal{N}(\mu_2, \sigma^2)$ zwei unabhängige Folgen von normalverteilten Zufallsvariablen auf einem Wahrscheinlichkeitsraum $(\Omega, \mathcal{F}, P)$ mit Erwartungswerten $\mu_1, \mu_2 \in \mathbb{R}$ und Varianz $\sigma^2 > 0$. Um zu überprüfen, ob die beiden Erwartungswerte μ_1 und μ_2 übereinstimmen oder voneinander abweichen, kann z.B. der Schätzer

$$T_n = T_n(X_1, \dots, X_n, Y_1, \dots, Y_n) := \frac{1}{n} \sum_{i=1}^{n} (X_i - Y_i)$$

verwendet werden.

(a) Ist der Schätzer T_n erwartungstreu für den Parameter $\vartheta := \mu_1 - \mu_2$?

(b) Ist der Schätzer T_n schwach bzw. stark konsistent für den Parameter ϑ?

Lösung: a)

$$E(T_n) = E\left(\frac{1}{n}\sum_{i=1}^{n}(X_i - Y_i)\right) = \frac{1}{n}\sum_{i=1}^{n}\left(\underbrace{E(X_i)}_{=\mu_1} - \underbrace{E(Y_i)}_{=\mu_2}\right)$$

$$= \frac{1}{n}\cdot\sum_{i=1}^{n}(\mu_1 - \mu_2) = \mu_1 - \mu_2$$

$\Rightarrow T_n$ erwartungstreu

b) Starke Gesetz großer Zahlen:

$$\frac{1}{n}\sum_{i=1}^{n} X_i \xrightarrow[n\to\infty]{f.s.} E(X_1) = \mu_1$$

und $$\frac{1}{n}\sum_{i=1}^{n} Y_i \xrightarrow[n\to\infty]{f.s.} E(Y_1) = \mu_2$$

$$\Rightarrow T_n \overset{\text{Def.}}{=} \frac{1}{n}\sum_{i=1}^{n}(X_i - Y_i) = \underbrace{\frac{1}{n}\sum_{i=1}^{n} X_i}_{\xrightarrow[n\to\infty]{f.s.} \mu_1} - \underbrace{\frac{1}{n}\sum_{i=1}^{n} Y_i}_{\xrightarrow[n\to\infty]{f.s.} \mu_2} \xrightarrow[n\to\infty]{f.s.} \mu_1 - \mu_2$$

$\Rightarrow T_n$ ist stark konsistent

Da fast sichere Konvergenz die stochastische Konvergenz impliziert, ist T_n auch schwach konsistent für ϑ.