TU Berlin - Institut für Mathematik
Sommersemester 2024
Dozent: Dr. Nikolas Tapia
Assistentin: M.Sc. Claudia Drygala

Stochastik für Informatik(er) – Übung 10

Abgabe: Keine Abgabe

Hinweise zur Bearbeitung des Übungsblattes:

- Das Übungsblatt enthält Haus- und Tutoriumsaufgaben.
- Die Tutoriumsaufgaben werden in den Tutorien der KW 26 besprochen (24.06.-28.06.).
- Die Hausaufgaben werden selbstständig bearbeitet. Lösungsvorschläge werden von uns hochgeladen. Bei Fragen wenden Sie sich an die Tutor*innen direkt im Tutorium oder in den Sprechstunden.

Tutoriumsaufgaben

Tutoriumsaufgabe 10.1

Die Kapazitäten von 10 Kondensatoren sind

Kondensator	1	2	3	4	5	6	7	8	9	10
Kapazität	101	100	102	102	101	99	99	101	102	98

(i) Bestimmen Sie alle absoluten Häufigkeiten.

(ii) Bestimmen Sie das empirische Mittel und die empirische Varianz.

(iii) Zeichnen Sie das zugehörige Histogramm mit den Klassen $\{98, 99, 100\}$, $\{101\}$ und $\{102\}$.

Tutoriumsaufgabe 10.2

Seien $X_1, ..., X_n$ Zufallsvariablen. Zeigen Sie für $\bar{\mu} := \frac{1}{n}\sum_{j=1}^{n} X_j$, dass

(i) $\sum_{j=1}^{n}(X_j - \bar{\mu}) = 0$;

(ii) $\sum_{j=1}^{n}(X_j - \bar{\mu})^2 = \sum_{j=1}^{n}(X_j^2 - \bar{\mu}^2)$;

(iii) $\sum_{j=1}^{n}(X_j - M)^2$ ist minimal genau dann, wenn $M = \bar{\mu}$.

Tutoriumsaufgabe 10.3

Eine unfaire Münze, die mit unbekannter Wahrscheinlichkeit $p \in (0, 1)$ Kopf zeigt, wird wiederholt geworfen. Bei 10 Wiederholungen wurden 7 Köpfe und 3 Zahlen beobachtet.

(i) Ermitteln Sie einen Schätzer $\overline{p}_n$ für den Parameter p mittels der Maximum-Likelihood-Methode.

(ii) Geben Sie die Schätzung $\overline{p}_{10}$ für das Münzwurfexperiment an.

Tutoriumsaufgabe 10.4

Sei $y_1, y_2, \ldots, y_n$ eine Stichprobe von unabhängigen, identisch verteilten Zufallsvariablen $Y_1, Y_2, \ldots, Y_n$ mit Dichte

$$f(y) = \begin{cases} \frac{\alpha y^{\alpha-1}}{\theta^\alpha} & \text{für } 0 \leq y \leq \theta \\ 0 & \text{sonst} \end{cases}$$

mit unbekanntem $\theta > 0$ und bekanntem $\alpha > 0$.

(i) Zeigen Sie, dass der Maximum-Likelihood-Schätzer $\bar{\theta} := \max\{y_1, \ldots, y_n\}$ ist.

(ii) Zeigen Sie, dass der Maximum-Likelihood-Schätzer nicht erwartungstreu ist.

(iii) Finden Sie eine Konstante $c > 0$, sodass der skalierte Schätzer $c\bar{\theta}$ erwartungstreu ist.

Tutoriumsaufgabe 10.1

Die Kapazitäten von 10 Kondensatoren sind

Kondensator	1	2	3	4	5	6	7	8	9	10
x = Kapazität	101	100	102	102	101	99	99	101	102	98

(i) Bestimmen Sie alle absoluten Häufigkeiten.

(ii) Bestimmen Sie das empirische Mittel und die empirische Varianz.

(iii) Zeichnen Sie das zugehörige Histogramm mit den Klassen $\{98, 99, 100\}$, $\{101\}$ und $\{102\}$.

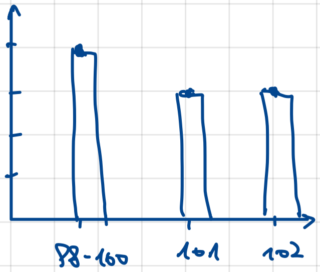
(i)

x	98	99	100	101	102	
$H(x)$	1	2	1	3	3	absolute Häufigkeiten
$h(x)$	$\frac{1}{10}$	$\frac{1}{5}$	$\frac{1}{10}$	$\frac{3}{10}$	$\frac{3}{10}$	relative Häufigkeiten

(ii) $\bar{\mu}_{10} = \frac{1}{10} \sum_{i=1}^{10} x_i = \frac{1005}{10} = 100{,}5$

$\bar{\sigma}_{10} = \frac{1}{9} \sum_{i=1}^{10} (x_i - \bar{\mu}_{10})^2 = 2{,}056 = \frac{37}{18}$

(iii)



Tutoriumsaufgabe 10.2

Seien $X_1, ..., X_n$ Zufallsvariablen. Zeigen Sie für $\bar{\mu} := \frac{1}{n} \sum_{j=1}^{n} X_j$, dass

(i) $\sum_{j=1}^{n} (X_j - \bar{\mu}) = 0$;

(ii) $\sum_{j=1}^{n} (X_j - \bar{\mu})^2 = \sum_{j=1}^{n} (X_j^2 - \bar{\mu}^2)$;

(iii) $\sum_{j=1}^{n} (X_j - M)^2$ ist minimal genau dann, wenn $M = \bar{\mu}$.

(i)

Wir zeigen $\sum_{j=1}^{n} (X_j - M) = 0 \iff M = \bar{\mu}$ sei $M \in \mathbb{R}$

"$\Rightarrow$" $0 = \sum_{j=1}^{n} (X_j - M) = \sum_{j=1}^{n} X_j - n \cdot M = n\bar{\mu} - nM$

$\Rightarrow n\bar{\mu} = nM$

$\Rightarrow M = \bar{\mu}$

"$\Leftarrow$" es ist $\sum_{j=1}^{n} (X_i - \bar{\mu}) = \sum_{j=1}^{n} X_i - n\bar{\mu} = n\bar{\mu} - n\bar{\mu} = 0$

(ii) $\sum_{j=1}^{n} (X_i - \bar{\mu})^2 = \sum_{j=1}^{n} (X_i^2 - 2X_i\bar{\mu} + \bar{\mu}^2) = \sum_{j=1}^{n} (X_i^2 - \bar{\mu}^2 - 2X_i\bar{\mu} + 2\bar{\mu}^2)$

$= \sum_{j=1}^{n} (X_i^2 - \bar{\mu}^2) - 2\bar{\mu} \cdot \underbrace{\sum_{j=1}^{n} (X_i - \bar{\mu})}_{=0}$

$= \sum_{j=1}^{n} (X_j^2 - \bar{\mu}^2)$

(iii) Variante 1: $f: \mathbb{R} \to \mathbb{R},\ M \mapsto \sum_{j=1}^{n} (x_j - M)^2$

$$f'(M) = \sum_{j=1}^{n} 2(M - x_j) = 2nM - \sum_{j=1}^{n} 2x_j = 2nM - 2n\bar{\mu} = 0$$

$$M = \bar{\mu}$$

$$f''(M) = \sum_{j=1}^{n} 2 = 2n > 0$$

Nach i) $\{f' = 0\} = \{\bar{\mu}\}$

Außerdem ist $f''(\bar{\mu}) = 2n > 0$

Nach Extremwertkriterium ist $\bar{\mu}$ das einzige strikte lokale Minimum von f.

Da man sogar $f''(M) > 0$ für alle $M \in \mathbb{R}$ ist f strikt konvex und das lokale Minimum sogar global

Variante 2: (Zu Folge) Für $M \in \mathbb{R}$ mit $\sum_{i=1}^{n} (x_i - M)^2 = \sum_{i=1}^{n} \left((x_i - \bar{\mu}) + (\bar{\mu} - M)\right)^2$

$$= \sum_{i=1}^{n} (x_i - \bar{\mu})^2 + 2(\bar{\mu} - M)\underbrace{\sum_{i=1}^{n} (x_i - M)}_{=0} + \underbrace{\sum_{i=1}^{n} (\bar{\mu} - M)^2}_{= n(\bar{\mu} - M)^2}$$

$$= \sum_{i=1}^{n} (x_i - \bar{\mu})^2 + \underbrace{n(\bar{\mu} - M)^2}_{=0}$$

$$\Rightarrow M = \bar{\mu}$$

Damit „$\in$" gezeigt, „$\Rightarrow$" sei $M \in \mathbb{R}$ ein Minimum

D.h. $\sum_{i=1}^{n} (x_i - M)^2 \leq \sum_{i=1}^{n} (x_i - \tilde{M})^2$

f.a. $\tilde{M} \in \mathbb{R}$, z.z. $M = \bar{\mu}$ Aus der Rechnung von quadrat. daan

$$\sum_{i=1}^{n} (x_i - \tilde{M}) \overset{\tilde{M} = M}{=} \sum_{i=1}^{n} (x_i - M)^2 = \sum_{i=1}^{n} (x_i - \bar{\mu}) + n(\bar{\mu} - M)^2$$

$$\Rightarrow 0 \geq n(\bar{\mu} - M)^2$$

$$\Rightarrow M = \bar{\mu}$$

Tutoriumsaufgabe 10.3

Eine unfaire Münze, die mit unbekannter Wahrscheinlichkeit $p \in (0, 1)$ Kopf zeigt, wird wiederholt geworfen. Bei 10 Wiederholungen wurden 7 Köpfe und 3 Zahlen beobachtet.

(i) Ermitteln Sie einen Schätzer $\overline{p}_n$ für den Parameter p mittels der Maximum-Likelihood-Methode.

(ii) Geben Sie die Schätzung $\overline{p}_{10}$ für das Münzwurfexperiment an.

(ii) $\bar{p}_{10} = \frac{7}{10}$

(i) $X_1, \dots, X_{10}$ mit $X_i \sim \mathrm{Ber}(p)$ $p \in (0,1)$ $\{X_i = 1\} =$ „i-te Wurf ist Kopf"

Sei $(x_1, \dots, x_{10}) \in \{0,1\}^n$ eine Stichprobe von $(X_1, \dots, X_{10})$ mit $\sum_{i=1}^{10} x_i = 7$

Schritt 1: Stelle Max-Likelihood-Fkt mit

$$L(x_1, \dots, x_{10}, p) = \prod_{i=1}^{10} \mathbb{P}(X_i = x_i) = \prod_{i=1}^{10} p^{x_i} (1-p)^{1-x_i}$$

$$\mathbb{P}(X_i = x_i) = \begin{cases} p & x_i = 1 \\ 1-p & x_i = 0 \end{cases} \quad X_i \sim \mathrm{Bin}(1, p) \quad \mathbb{P}(X_i = x_i) = p^{x_i}(1-p)^{1-x_i}$$

Schritt 2: Maximiere L bzgl p

Hin: Extremwertkriterium in Analysis

Dann stelle die log-Max-L.-h.-Fkt auf $\ell = \log \circ L$

d.h. $\ell(x_1, \dots, x_n, p) = \sum_{i=1}^{10} \left(x_i \log(p) + (1-x_i) \log(1-p) \right)$

$$\frac{d}{dp} \ell(x_1, \dots, x_n, p) = \sum_{i=1}^{10} \left(\frac{x_i}{p} - \frac{1-x_i}{1-p} \right)$$

$$\frac{d}{d^2p^2} \ell(x_1, \dots, x_n, p) = \sum_{i=1}^{10} \left(-\frac{x_i}{p^2} - \frac{1-x_i}{(1-p)^2} \right)$$

Berechnen von $\{ \frac{d}{dp} \ell = 0 \}$ für alle $p \in (0,1)$ gilt

$$\frac{d}{dp} \ell(x_1, \dots, x_n, p) = 0 \iff \sum_{i=1}^{10} \left(\frac{x_i}{p} - \frac{1-x_i}{1-p} \right) = 0$$

$$\iff \sum_{i=1}^{10} \left(x_i(1-p) - (1-x_i)p \right) = 0$$

$$\iff \sum_{i=1}^{10} (x_i - p) = 0$$

$$\iff \bar{p}_{10} = p = \bar{\mu}_{10} = \frac{1}{10} \sum_{i=1}^{10} x_i$$

d.h. $\{ \frac{d}{dp} \ell = 0 \} = \{ \bar{\mu}_{10} \}$

<u>Prüfung von 2. Ableitung</u>

$$\frac{d^2}{dp^2} \ell \iff \sum_{i=1}^{10} \left(\frac{x_i}{p^2} + \frac{1-x_i}{(1-p)^2} \right) <$$

Also $p \mapsto \ell(x_1, \dots, x_n, p)$ strikt konkav und insgesamt $\bar{p}_{10}$ das globale Maximum von ℓ

Da log strikt wachsend, ist $\bar{p}_{10}$ auch das globale Max von L

Tutoriumsaufgabe 10.4

Sei $y_1, y_2, \dots, y_n$ eine Stichprobe von unabhängigen, identisch verteilten Zufallsvariablen $Y_1, Y_2, \dots, Y_n$ mit Dichte

$$f(y) = \begin{cases} \frac{\alpha y^{\alpha-1}}{\theta^\alpha} & \text{für } 0 \le y \le \theta \\ 0 & \text{sonst} \end{cases}$$

mit unbekanntem $\theta > 0$ und bekanntem $\alpha > 0$.

(i) Zeigen Sie, dass der Maximum-Likelihood-Schätzer $\bar{\theta} := \max\{y_1, \dots, y_n\}$ ist.

(ii) Zeigen Sie, dass der Maximum-Likelihood-Schätzer nicht erwartungstreu ist.

(iii) Finden Sie eine Konstante $c > 0$, sodass der skalierte Schätzer $c\bar{\theta}$ erwartungstreu ist.

(i) ① Stelle Max-L-H-Fkt auf.

$$L(y_1, \dots, y_n, \theta) = \prod_{i=1}^{n} f(y_i) \begin{cases} \dfrac{\alpha^n \prod_{i=1}^{n} y_i^{\alpha-1}}{\theta^{\alpha n}} & \text{für } (y_1, \dots, y_n) \in [0,\theta]^n \\ 0 & \text{sonst} \end{cases}$$

② Maximale $\theta \mapsto L(y_1, \dots, y_n, \theta)$. Dieses Mal hin:
keine Differentialrechnung, sondern zu Fuß

f.a. $\theta < \max_{i=1}^{n} y_i$ gilt

$$L(y_1, \dots, y_n, \theta) = 0 < L(y_1, \dots, y_n, \max_{i=1}^{n} y_i)$$

und f.a. $\theta \geq \max_{i=1}^{n} y_i$

$$L(y_1, \dots, y_n, \theta) = \frac{\alpha^n \prod_{i=1}^{n} y_i^{\alpha-1}}{\theta^{\alpha n}} \overset{\theta > \max y_i}{\leq} \frac{\alpha^n \prod_{i=1}^{n} y^{\alpha-1}}{(\max_{i=1}^{n} y_i)^{\alpha n}} = L(y_1, \dots, y_n, \max_{i=1}^{n} y_i)$$

ii) $\bar{\theta}_i = \max_{i=1}^{n} y_i$ ist $\bar{\theta}$ erwartungstreu? d.h. $\mathbb{E}[\bar{\theta}] = \theta$?

Lösung: $F_{\bar{\theta}}(x) = \mathbb{P}(\max_{i=1}^{n} Y_i \leq x) = \mathbb{P}(Y_1 \leq x, \dots, Y_n \leq x)$

$$= \prod_{i=1}^{n} \mathbb{P}(Y_i \leq x) = F_{Y_n}(x)^n$$

$Y_1, \dots, Y_n$ unabhängig

$$= \left(\int_0^x \frac{\alpha y^{\alpha-1}}{\theta^\alpha} dy \right)^n \qquad \frac{y^\alpha}{\theta^\alpha} \Big|_0^x$$

$$= \left(\frac{x}{\theta} \right)^{\alpha n}$$

(i) Stelle die Linkelihoodfunktion auf:

$$L(\theta; y_1, \dots, y_n) = \begin{cases} \frac{\alpha^n}{\theta^{n\alpha}} (y_1 \cdots y_n)^{\alpha-1} & \text{falls } (y_1, \dots, y_n) \in [0,\theta]^n \\ 0 & \text{sonst.} \end{cases}$$

Falls θ groß und wächst, so fällt die Funktion. Wird θ kleiner, so wächst die Funktion streng monoton so lange, bis sie das erste Mal einen der Werte $y_1, \dots, y_n$ trifft (das ist das Maximum derer) – dann fällt sie sofort auf 0. Es gilt also $\bar{\theta} = \max\{Y_1, \dots, Y_n\}$.

Dichte $f_{\bar{\theta}}(x) = F_{\bar{\theta}}'(x) = \dfrac{\alpha n x^{\alpha n - 1}}{\theta^{\alpha n}}$

$$\mathbb{E}[\bar{\theta}] = \int_0^\theta x \frac{\alpha n x^{\alpha - 1}}{\theta^{\alpha n}} dx = \frac{\alpha n}{\theta^{\alpha n}} \int_0^\theta x^{\alpha n} dx$$

$$= \frac{\alpha n}{1 + \alpha n} \theta \neq \theta$$

also $\bar{\theta}$ nicht E

iii) $\tilde{\theta} = \dfrac{n + \alpha n}{\alpha n} \bar{\theta}$

Hausaufgaben

Hausaufgabe 10.1 (0 Punkte)

Die Zahl der Verkehrsunfälle in einer Stadt lag im letzen Jahr an zehn zufällig gewählten, regenfreien Tagen bei

$$4 \quad 0 \quad 6 \quad 5 \quad 2 \quad 1 \quad 2 \quad 0 \quad 4 \quad 3$$

Die Anzahl der Unfälle an einem Tag sei poissonverteilt mit unbekanntem Parameter $\lambda > 0$.

(i) Ermitteln Sie einen Schätzwert für den Parameter λ mittels der Maximum-Likelihood-Methode, nehmen sie dabei an, dass jede beobachtete Stichprobe mindestens einen Tag mit einem Unfall enthält.

(ii) Verwenden Sie den ML-Schätzer zur Stichprobe aus dem ersten Aufgabenteil, um den den Anteil der regenfreien Tage, an denen im letzten Jahr zwei oder weniger Verkehrsunfälle passierten, zu schätzen.

Hausaufgabe 10.2 (0 Punkte)

Die Körpergröße der Studentinnen (in cm) beim Kurs Stochastik für Informatiker ist normalverteilt mit unbekanntem Erwartungswert $\mu > 0$ und bekannter Varianz $\sigma^2 > 0$.

(i) Bestimmen Sie den Maximum-Likelihood-Schätzer für μ und seinen Erwartungswert.

(ii) Wir beobachten die folgenden 6 Realisierungen von $X_1, \dots, X_6$:

$$x_1 = 176 \qquad x_2 = 164 \qquad x_3 = 169 \qquad x_4 = 158 \qquad x_5 = 192 \qquad x_6 = 155$$

Berechnen Sie auf Grundlage dieser Beobachtungen den Schätzwert $\overline{\mu}_6$ für den unbekannten Parameter μ, falls $\sigma^2 = 81$. Berechnen Sie auch den Median und die empirische Varianz von $(x_1, \dots, x_n)$. Welche von diesen Kenngrößen hängen von σ^2 ab?