13. Übungsblatt

Grundbegriffe der Statistik, empirische Verteilungsfunktion, Punktschätzungen, Maximum-Likelihood-Schätzung

1. In einer mündlichen Prüfung prüfte ein Prüfer 13 Studierende an einem Tag und vergab die folgende Noten (in dieser Reihenfolge): $3, 2, 4, 5, 2, 1, 3, 3, 4, 5, 4, 1, 2$. Betrachten wir die obigen Daten als die Realisierung einer einfachen Stichprobe $\mathbf{X} = (X_1, \ldots, X_{13})$. Berechnen wir das empirische Mittel, die empirische Standardabweichung, den Stichprobenmedian, den Modus und den der Variablen X_8^* entsprechenden Wert für diese Realisierung.

2. Eine Mine hat 8 Mitarbeiter: Schneewittchen (die Direktorin der Mine), Chef (der Schichtleiter), ferner Happy, Brummbär, Hatschi, Pimpel, Schlafmütz und Seppel (die Bergmänner).

   a) Das Monatsgehalt der Direktorin ist 2000 Taler, der Schichtleiter verdient 300 Taler pro Monat und jeder Bergmann verdient 200 Taler monatlich. Bestimmen wir das Mittel, den Median und den Modus der Gehälter der Arbeitnehmer der Mine.

   b) In seiner Freizeit beschäftigt sich Happy auch mit offenen Problemen der Informatik. Nehmen wir an, dass er das P-NP-Problem löst, für das er 1 Million Dollar erhält, und nach Umwechslung erhöht sich sein Gehalt für diesen Monat auf genau 5 Millionen Taler. Wie ändern sich das Mittel, der Median und der Modus der Gehälter in diesem Monat?

   Was können wir anhand der Aufgabe über das Verhalten des Mittels und des Medians sagen?

---

3. Ein regulärer Würfel wird 6 Mal geworfen, und wir erhalten die Augenzahlen $1, 3, 4, 5, 6, 1$ in dieser Reihenfolge.

   (a) Stellen wir den Graph der zu dieser Realisierung gehörigen empirischen Verteilungsfunktion F_6^* dar. $F_6^*(1)$=?

   (b) Sei $(X_1, \ldots, X_n)$ der Stichprobenvektor, die das Ergebniss der n-maligen Würfelns angibt. Bestimmen wir die Werte $\lim_{n\to\infty} F_n^*(1)$, $\lim_{n\to\infty} F_n^*(3,5)$, und $\lim_{n\to\infty} F_n^*(6)$ (genauer gesagt, die Werte, mit denen diese Grenzwerte mit Wahrscheinlichkeit 1 übereinstimmen).

4. Zeigen wir, dass die emprische Verteilungsfunktion F_n^*, die zu einer Realisierung einer beliebigen Stichproben vom Umfang n gehört, für alle $n \in \mathbb{N}^+$ wirklich eine Verteilungsfunktion ist. Ist es möglich, dass eine Zufallsvariable mit Verteilungsfunktion F_n^* auch eine Dichtefunktion hat?

5. Sei $(X_1, X_2, \ldots)$ eine Folge von unabhängigen und identisch verteilten Stichprobenvariablen mit Verteilnugsfunktion F. Zeigen wir, dass der Wert der entsprechenden empirischen Verteilungsfunktion $F_n^*(t)$ für alle $t \in \mathbb{R}$ eine erwartungstreue und stark konsistente Schätzfunktion für den Wert $F(t)$ ist.

---

6. Seien $X_1, \ldots, X_n, \ldots$ (unabhängige und identisch verteilte) Stichprobenvariablen mit Dichtefunktion

$$f_\vartheta(t) = \begin{cases} \frac{1}{2}(1 + \vartheta t), & \text{falls } -1 < t < 1 \\ 0, & \text{sonst,} \end{cases}$$

   wo $\vartheta \in (-1, 1)$ ein unbekannter Parameter ist (überprüfen wir, dass f_ϑ für alle $\vartheta \in (-1, 1)$ wirklich eine Dichtefunktion ist).

   (a) Geben wir eine erwartungstreue und stark konsistente Schätzfunktion $\vartheta_n(X_1, \ldots, X_n)$ für den Parameter ϑ.
   Hinweis: Betrachten wir das empirische Mittel.

   (b) Bei einer Beobachtung erhielten wir die folgende Realisierung der Stichprobe $(X_1, \ldots, X_6)$:

$$x_1 = -0{,}58 \qquad x_2 = 0{,}93 \qquad x_3 = 0{,}01 \qquad x_4 = -0{,}23 \qquad x_5 = 0{,}55 \qquad x_6 = 0{,}01$$

   Berechnen wir den Wert der im Teil a) bestimmten Schäztfunktion für den unbekannten Parameter ϑ bei dieser Realisierung.

7. Sei $\mathbf{X} = (X_1, \dots, X_n)$ eine (einfache) Stichprobe vom Umfang n, für die $X_i \sim U(\vartheta; \vartheta + 1)$ für alle $1 \le i \le n$ gilt, wo $\vartheta \in \Theta = (0, \infty)$ ein unbbekannter Parameter ist.
   (a) Berechnen wir $\mathbb{E}_\vartheta(1/X_1)$ und $\mathbb{D}^2_\vartheta(1/X_1)$,
   (b) und geben wir einen erwartungstreuen und stark konsistenten Schätzer für die Parameterfunktion $\psi(\vartheta) = \ln \frac{\vartheta+1}{\vartheta}$.

8. Sei $\mathbf{X} = (X_1, \dots, X_n)$ eine (einfache) Stichprobe vom Umfang n, für die $X_i \sim U(a; b)$ gilt, wo $\vartheta = (a; b) \in \Theta \subset \mathbb{R}^2$ ein zweidimensionaler Parameter ist und
$$\Theta = \{(a, b) \in \mathbb{R}^2 \colon 0 < a < b\}.$$
   Berechnen wir $\mathbb{D}^2_\vartheta(X_1)$ für alle $\vartheta = (a, b) \in \Theta$. Ist die Stichprobenfunktion $T(\mathbf{X}) = S_n^{*2}$ ein erwartungstreuer Schätzer der Parameterfunktion $\psi(\vartheta) = \psi(a, b) = \frac{(a+b)^2}{4}$?

9. Sei $\mathbf{X} = (X_1, \dots, X_n)$ eine (einfache) Stichprobe vom Umfang n, für die $X_i \sim U(0; \vartheta)$ für einen Parameter $\vartheta \ge 0$ gilt.
   (a) Zeigen wir, dass $2\overline{X}$ ein erwartungstreuer Schätzer für ϑ ist.
   (b) Bestimmen wir die Konstante c_n, für die $c_n X_n^*$ ein erwartungstreuer Schätzer für ϑ ist.
   (c) Entscheiden wir, welcher der obigen zwei erwartungstreuen Schätzer effizienter ist.

---

10. Sei $\mathbf{X} = (X_1, \dots, X_n)$ eine (einfache) Stichprobe vom Umfang n, für die $X_i \sim \mathrm{Geo}(\vartheta)$ für einen unbekannten Parameter $\vartheta \in (0; 1)$ gilt. Bestimmen wir den Wert der Likelihood-Funktion $L_\vartheta(x_1, \dots, x_n)$ für alle Realisierungen der Stichprobe und für alle $\vartheta \in (0; 1)$ und geben wir die Maximum-Likelikhood-Schätzung für ϑ an.

11. Sei $\mathbf{X} = (X_1, \dots, X_n)$ eine (einfache) Stichprobe vom Umfang n, wo die Dichtefunktion von X_i für einen Parameter $\vartheta > 0$ durch
$$f_\vartheta(t) = \begin{cases} \vartheta^2 t e^{-\vartheta t}, & \text{falls } t > 0, \\ 0 & \text{sonst} \end{cases}$$
   angegeben wird. Geben wir die Likelihood-Funktion $L_\vartheta(x_1, \dots, x_n)$ für jede Realisierung $(x_1, \dots, x_n)$ an und bestimmen wir die Likelihood-Schätzung für ϑ.

12. Sei $\mathbf{X} = (X_1, \dots, X_n)$ eine (einfache) Stichprobe vom Umfang n, wo die Verteilungsfunktion von X_i für einen Parameter $\vartheta > 1$ durch
$$F_\vartheta(x) = \begin{cases} 1 - x^{-\vartheta}, & \text{falls } x > 1, \\ 0, & \text{falls } x \le 1 \end{cases}$$
   angegeben wird. Geben wir die Likelihood-Funktion $L_\vartheta(x_1, \dots, x_n)$ für jede Realisierung $(x_1, \dots, x_n)$ an und bestimmen wir die Likelihood-Schätzung für ϑ.

*13. Welche Beziehung besteht zwischen der folgenden Aufgabe und der Maximum-Likelihood-Schätzung?

   Zwei Würfel werden geworfen, und falls k Sechse gewürfelt werden, dann legen wir k rote und $2 - k$ gelbe Bälle in einer (am Anfang leeren) Urne. Dann ziehen wir zweimal mit Zurücklegen, und bei beiden Ziehungen ziehen wir einen roten Ball. Was wäre unser Tipp für den Wert von k? Mit welcher Wahrscheinlichkeit wird unser Tipp richtig?