Wahrscheinlichkeit und Statistik

Musterlösung Serie 11

1. Seien $X_1, X_2, \ldots$ i.i.d. Zufallsvariablen mit Dichte

$$f_\vartheta(x) = \begin{cases} (\vartheta - 1)x^{-\vartheta} & \text{falls } x \geq 1 \\ 0 & \text{sonst} \end{cases},$$

wobei $\vartheta \in (1, \infty)$ ein unbekannter Parameter ist.

a) Berechnen Sie den Maximum Likelihood Schätzer T_n für ϑ, basierend auf n Beobachtungen.

b) Zeigen Sie, dass $(T_n)_n$ konsistent ist.

Lösung:

a) Seien $X = (X_1, \ldots, X_n)$ und $x = (x_1, \ldots, x_n)$. Da alle X_i unabhängig sind, ist die gemeinsame Dichte von X das Produkt der Dichten der X_i, und die Log-Likelihood-Funktion ist dann

$$\log L(\vartheta; x_1, \ldots, x_n) = \log \prod_{i=1}^{n} (\vartheta - 1)x_i^{-\vartheta} = n \log(\vartheta - 1) - \vartheta \sum_{i=1}^{n} \log x_i.$$

Ableiten und Null setzen liefert

$$\frac{n}{\vartheta - 1} = \sum_{i=1}^{n} \log x_i.$$

Der Maximum-Likelihood- Schätzer ist somit

$$T_n(X_1, \ldots, X_n) = 1 + \frac{n}{\sum_{i=1}^{n} \log X_i} = 1 + \frac{1}{\frac{1}{n}\sum_{i=1}^{n} \log X_i}.$$

b) Wir zeigen, dass für jedes $\varepsilon > 0$:

$$\mathbb{P}_\vartheta(\, |T_n(X_1, \ldots, X_n) - \vartheta\,| > \varepsilon\,) = \mathbb{P}_\vartheta\left(\left| \frac{1}{\frac{1}{n}\sum_{i=1}^{n} \log X_i} - (\vartheta - 1) \right| > \varepsilon \right) \longrightarrow 0 \ \ \forall \vartheta,$$

für $n \to \infty$. Aus dem Transformationssatz folgt, dass $\log X_i \sim \mathrm{Exp}(\vartheta - 1)$ mit Erwartungswert $\frac{1}{\vartheta - 1}$. Bei der Exponentialverteilung existieren wegen des starken Gesetzes der grossen Zahlen alle notwendigen Momente. Also konvergiert $\frac{1}{n}\sum_{i=1}^{n} \log X_i$ $\mathbb{P}_\vartheta$-fast sicher gegen $\frac{1}{\vartheta-1}$. Da fastsichere Konvergenz stochastische Konvergenz impliziert, folgt unmittelbar die Behauptung.

Bemerkung: Es ist auch möglich mit dem schwachen Gesetzt der grossen Zahlen zu argumentieren, jedoch benötig dies den sogenannten Satz von der stetigen Abbildung (siehe z.B. Wikipedia), welcher in dieser Vorlesung nicht behandelt wird. Der Satz von der stetigen Abbildung besagt, dass Konvergenz in Wahrscheinlichkeit unter stetigen Transformationen erhalten bleibt.

In unserem Fall kriegt man also mit dem schwachen Gesetz der grossen Zahlen, dass $\frac{1}{n}\sum_{i=1}^{n} \log X_i$ in Wahrscheinlichkeit gegen $\frac{1}{\vartheta-1}$ konvergiert. Dies impliziert für $f(x) = x^{-1}$ mit dem Satz von der stetigen Abbildung, dass $f(\frac{1}{n}\sum_{i=1}^{n} \log X_i) = \frac{1}{\frac{1}{n}\sum_{i=1}^{n} \log X_i}$ gegen $f(\frac{1}{\vartheta-1}) = \vartheta - 1$ konvergiert.

2. Eine Tankstelle veranschlagt für einen Ölwechsel mindestens α Minuten. Die tatsächlich benötigte Zeit X variiert natürlich im Bereich $X \geq \alpha$ und ist von Kunde zu Kunde verschieden. Man kann jedoch annehmen, dass diese Zeit durch eine exponential-verteilte Zufallsvariable gut wiedergegeben wird. Die Zufallsvariable X besitze somit die Dichtefunktion

$$f(t) = \begin{cases} e^{\alpha - t} & \text{falls } t \geq \alpha \\ 0 & \text{sonst} \end{cases},$$

d.h. $X = \alpha + Z$, wobei $Z \sim \text{Exp}(1)$. Um α schätzen zu können, wurde von 10 zufällig ausgewählten Kunden die für den Ölwechsel benötigte Arbeitszeit in Minuten notiert:

$$4.2, 3.1, 3.6, 4.5, 5.1, 7.6, 4.4, 3.5, 3.8, 4.3.$$

Bestimmen Sie den Maximum-Likelihood Schätzer für den unbekannten Parameter α.

Lösung: Die Likelihood Funktion ist durch

$$\begin{aligned} L(\alpha; X_1, \ldots, X_n) = \prod_{i=1}^n f(X_i) &= \exp\left(n\alpha - \sum_{i=1}^n X_i\right) \cdot \prod_{i=1}^n \mathbf{1}_{[\alpha,\infty)}(X_i) \\ &= \exp\left(n\alpha - \sum_{i=1}^n X_i\right) \cdot \mathbf{1}_{\min_{1\leq i\leq n} X_i \geq \alpha} \end{aligned}$$

gegeben. Diese Funktion ist genau dann positiv, wenn alle x_i grösser oder gleich α sind. Unter dieser Nebenbedingungen ist nun $n\alpha - \sum_{i=1}^n x_i$ zu maximieren. Daraus folgt sofort, dass

$$T(X_1, \ldots, X_n) = \min(X_i)$$

der Maximum-Likelihood Schätzer ist.

3. **Vergleich zweier Schätzmethoden.** Während der Vorbereitungszeit auf die mündliche Prüfung kommunizieren Arno und Benno per E-mail, um auftretende Probleme zu bereinigen. Aufgrund der Tageszeiten zu denen Benno Meldungen abschickt, hat Arno den Verdacht, dass Benno mehr arbeitet als er. Aus den Tageszeiten der letzten zehn Meldungen (umgerechnet in Stunden)

$$10.55, 14.9, 11.2, 18.85, 9.75, 11.5, 16.1, 14.4, 9.2, 12.95$$

will Arno Bennos Arbeitszeiten schätzen. Zu diesem Zweck nimmt er an, dass Benno jeden Tag zur Zeit a anfängt zu lernen und zur Zeit b aufhört, und dass die Zeiten, zu denen die E-mails abgeschickt werden, unabhängig und alle uniform-verteilt sind auf dem Intervall $[a, b]$. Schätzen Sie die Parameter a und b mit

a) der Maximum-Likelihood-Methode, und
b) der Momenten-Methode.

Sind diese Schätzungen vernünftig?

Lösung: Seien $X_1, \ldots, X_n$ iid Unif(a, b)-verteilt und $X \coloneqq (X_1, \ldots, X_n)$.

a) Die gemeinsame Dichte von $X_1, \dots, X_n$ ist

$$f_X(x_1, \dots, x_n) = \frac{1}{(b-a)^n} \prod_{i=1}^n \mathbf{1}_{[a,b]}(x_i). \tag{1}$$

Nun muss $f_X(x_1, \dots, x_n)$ für feste $(x_1, \dots, x_n)$ bezüglich a und b maximiert werden. Seien $x_* := \min_{1 \le i \le n} x_i$ und $x^* := \max_{1 \le i \le n} x_i$. Falls x_* oder x^* ausserhalb $[a,b]$ liegen, dann verschwindet die rechte Seite von (1). Also muss die Maximumstelle $(\hat{a}_{\mathrm{ML}}, \hat{b}_{\mathrm{ML}})$ die Bedingungen

$$\hat{b}_{\mathrm{ML}} \ge x^* \ge x_* \ge \hat{a}_{\mathrm{ML}}$$

erfüllen. Unter diesen Bedingungen wird die Differenz $\hat{b}_{\mathrm{ML}} - \hat{a}_{\mathrm{ML}}$ möglichst klein genau für $\hat{a}_{\mathrm{ML}} = x_*$ und $\hat{b}_{\mathrm{ML}} = x^*$. Somit sind x_* und x^* die Maximum-Likelihood Schätzungen für a und b. Numerisch

$$\hat{a}_{\mathrm{ML}} = 9.2 \quad \text{und} \quad \hat{b}_{\mathrm{ML}} = 18.85.$$

b) Es gilt:

$$\mathbb{E}[X_i] = \frac{a+b}{2} \quad \text{und} \quad \mathrm{Var}(X_i) = \frac{(b-a)^2}{12}.$$

Die Schätzungen mittels Momentenmethode erfüllen also

$$\bar{x}_n = \frac{\hat{a} + \hat{b}}{2} \quad \text{und} \quad s_n = \sqrt{\frac{(\hat{b} - \hat{a})^2}{12}}.$$

Es gilt $\bar{x}_n = 12.94$ und $s_n^2 = 9.54$, also:

$$\hat{a}_{\mathrm{MM}} = 7.59 \quad \text{und} \quad \hat{b}_{\mathrm{MM}} = 18.29.$$

Alternative Lösung: Beachten Sie, dass die obige Lösung eine Abänderung des Momentenmethodes ist. Man kann natürlich auch die Parameter a und b mit der klassischen Momentenmethode schätzen.
Es gilt:

$$\mathbb{E}[X_i] = \frac{a+b}{2}$$

und

$$\mathbb{E}[X_i^2] = \mathrm{Var}(X_i) + \mathbb{E}[X_i]^2 = \frac{(b-a)^2}{12} + \left(\frac{a+b}{2}\right)^2 = \frac{a^2+ab+b^2}{3}.$$

Wir erhalten das folgende Gleichungssystem

$$\begin{cases} \bar{x}_n = \frac{a+b}{2} \\ \frac{x_1^2 + \dots + x_n^2}{n} = \frac{a^2+ab+b^2}{3} \end{cases}.$$

Es gilt $\bar{x}_{10} = 12.94$ und $\frac{x_1^2 + \dots + x_{10}^2}{10} = 176.03$. Das Auflösen des Gleichungssystems ergibt:

$$a = 7.864 \quad \text{und} \quad b = 18.016.$$

Die Momentenmethode hat den Vorteil, dass sie $b - a$ (Arbeitsdauer) erwartungstreu schätzt, aber es kann vorkommen (wie hier für den Wert x_4), dass Daten ausserhalb des geschätzten Intervalls $[\hat{a}, \hat{b}]$ liegen. Nach Konstruktion kann dies bei der Maximum-Likelihood Methode nicht passieren, aber dafür unterschätzt man hier die Arbeitsdauer systematisch. Abgesehen davon müsste man in diesem Beispiel die Annahme der Unabhängigkeit und Gleichverteilung der Daten sehr in Frage stellen.