ETH Zurich FS 2020
D-INFK
Prof. J. Teichmann

Koordinator
Konstantin Rusch

Wahrscheinlichkeit & Statistik

Lösungen 11

1. 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 exponentialverteilte Zufallsvariable gut wiedergegeben wird. Die Zufallsvariable X besitze somit die Dichtefunktion

$$f_X(x) := \begin{cases} e^{\alpha - x} & x \geq \alpha, \\ 0 & \text{sonst}, \end{cases}$$

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

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

woraus sich der empirische Mittelwert von $\overline{x}_{10} = 4.41$ ergibt. Bestimme (zunächst allgemein[1]) den

a) Maximum-Likelihood-Schätzer (**Vorsicht:** Die Likelihood-Funktion ist nicht differenzierbar. Das Maximum der Likelihood-Funktion muss daher auf eine andere Art bestimmt werden.),

b) Momentenschätzer

für α, und werte diesen aus für die angegebene Stichprobe.

Lösung:

a) Die Likelihoodfunktion ist gegeben durch

$$\begin{aligned} L(\alpha; x_1, \ldots, x_n) &= \prod_{i=1}^n f_\alpha(x_i) = \prod_{i=1}^n e^{\alpha - x_i} 1_{[\alpha,\infty)}(x_i) = \exp\left(n\alpha - \sum_{i=1}^n x_i\right) \cdot \prod_{i=1}^n 1_{[\alpha,\infty)}(x_i) \\ &= \exp\left(n\alpha - \sum_{i=1}^n x_i\right) \cdot 1_{[\alpha,\infty)}\left(\min_{1\leq i\leq n} x_i\right) \\ &= \begin{cases} \exp\left(n\alpha - \sum_{i=1}^n x_i\right) & \alpha \leq \min_{1\leq i\leq n} x_i, \\ 0 & \text{sonst} \end{cases} \end{aligned}$$

[1]Basierend auf n Beobachtungen von X.

wobei wir zur Vereinfachung der Notation die Indikatorfunktion benutzt haben: Für eine Menge A ist 1_A die Funktion

$$1_A(x) = \begin{cases} 1, & x \in A\,, \\ 0, & \text{sonst.} \end{cases}$$

Um die obige Likelihoodfunktion zu maximieren, sollte α so gross wie möglich sein, darf aber nicht grösser als $\min_{1\le i\le n} x_i$ werden. Der Maximum-Likelihood-Schätzer für α ist daher

$$\widehat{\alpha} = \min_{1\le i\le n} x_i \qquad \text{bzw. als Zufallsvariable} \qquad \widehat{\alpha} = \min_{1\le i\le n} X_i.$$

b) Es gilt $\mu = \mathbb{E}\,(X) = \alpha + \mathbb{E}\,(Z) = \alpha + 1$; somit ist $\alpha = \mu - 1$ und der Momentenschätzer ist gegeben durch

$$\widehat{\alpha} = \widehat{\mu} - 1 = \overline{X}_n - 1.$$

Die Realisierung für die gegebenen Werte ergibt $\widehat{\alpha} = \overline{x}_{10} - 1 = 3.41$.

2. Wir betrachten Pegelstände bei Hochwasser im Zürichsee. Hochwasser bedeute dabei, dass der Pegelstand die kritische Marke von 140 cm über Normalniveau überschreitet. Die Zufallsvariable X messe die Wasserhöhe in cm über der kritischen Marke. Zur Modellierung von X können wir eine sogenannte verallgemeinerte Pareto-Verteilung mit Dichte

$$f_X(x;\vartheta) = \begin{cases} \frac{1}{\vartheta}(1+x)^{-(1+\frac{1}{\vartheta})} & \text{falls } x > 0, \\ 0 & \text{falls } x \le 0 \end{cases}$$

verwenden. Dabei ist ϑ ein unbekannter Parameter, der auf Basis von Daten $x_1, \ldots, x_n$ geschätzt werden soll; diese Daten werden wie üblich als Realisierungen von Zufallsvariablen $X_1, \ldots, X_n$ aufgefasst, die für jede Wahl des Parameters ϑ unter P_ϑ i.i.d. sind mit Dichte $f_X(x;\vartheta)$.

a) Sei $\vec{X} = (X_1, \ldots, X_n)$. Zeige, dass

$$T^{(n)} = t^{(n)}(\vec{X}) := \sum_{i=1}^{n} \frac{\log(1+X_i)}{n}$$

der Maximum-Likelihood-Schätzer von ϑ ist.

b) Berechne den Erwartungswert und die Varianz von $T^{(n)}$ in jedem Modell P_ϑ.
Hinweis: Benutze, dass $Y_i := \log(1+X_i)$ $Exp(\frac{1}{\vartheta})$-verteilt ist, d.h.
$f_Y(y) = \frac{1}{\vartheta}e^{-\frac{1}{\vartheta}y}$ für $y \ge 0$.

c) Ist $T^{(n)}$ ein erwartungstreuer Schätzer für ϑ?

d) Ist die Folge von Schätzern $T^{(n)}, n \in \mathbb{N}$, konsistent?

e) Beweise den Hinweis in **b)**.

Lösung:

a) Die log-Likelihood-Funktion ist gegeben durch

$$\log L(x_1,\dots,x_n;\vartheta) = \log\left(\frac{1}{\vartheta^n}\prod_{i=1}^n (1+x_i)^{-(1+\frac{1}{\vartheta})}\right) = -n\log\vartheta - (1+\frac{1}{\vartheta})\sum_{i=1}^n \log(1+x_i).$$

Die Ableitung nach ϑ ist

$$\frac{\partial}{\partial\vartheta}\log L(x_1,\dots,x_n;\vartheta) = -\frac{n}{\vartheta} + \frac{1}{\vartheta^2}\sum_{i=1}^n \log(1+x_i),$$

und das ist 0 für

$$\vartheta = \frac{1}{n}\sum_{i=1}^n \log(1+x_i).$$

Der ML-Schätzer für ϑ ist also gegeben durch

$$T^{(n)} = \frac{1}{n}\sum_{i=1}^n \log(1+X_i).$$

b) Die Linearität des Erwartungswertes liefert

$$E_\vartheta[T^{(n)}] = E_\vartheta\left[\frac{1}{n}\sum_{i=1}^n \log(1+X_i)\right] = \frac{1}{n}\sum_{i=1}^n E_\vartheta[\log(1+X_i)] = E_\vartheta[\log(1+X_1)].$$

Nach dem Hinweis ist $Y_1 := \log(1+X_1)$ $Exp(\frac{1}{\vartheta})$-verteilt. Also ist $E[Y_1] = \vartheta$ und $\mathrm{Var}[Y_1] = \vartheta^2$. Daher gilt $E_\vartheta[T^{(n)}] = E_\vartheta[Y_1] = \frac{1}{\frac{1}{\vartheta}} = \vartheta$.

Ähnlich rechnen wir für die Varianz

$$\begin{aligned}\mathrm{Var}_\vartheta[T^{(n)}] &= \mathrm{Var}_\vartheta\Big[\frac{1}{n}\sum_{i=1}^n \log(1+X_i)\Big] = \frac{1}{n^2}\sum_{i=1}^n \mathrm{Var}_\vartheta[\log(1+X_i)] \\ &= \frac{1}{n}\mathrm{Var}_\vartheta[\log(1+X_1)] = \frac{1}{n}\mathrm{Var}_\vartheta[Y_1] = \frac{1}{n}\frac{1}{\frac{1}{\vartheta^2}} = \frac{1}{n}\vartheta^2.\end{aligned}$$

c) Aus **b)** folgt, dass $E_\vartheta[T^{(n)}] = \vartheta$, d.h. $T^{(n)}$ ist erwartungstreu für ϑ.

d) Die Chebyshev-Ungleichung liefert für jedes $n \in \mathbb{N}$ und für beliebiges $\varepsilon > 0$

$$P_\vartheta[|T^{(n)} - \vartheta| > \varepsilon] \le \frac{1}{\varepsilon^2}\mathrm{Var}_\vartheta[T^{(n)}] = \frac{1}{\varepsilon^2}\frac{1}{n}\vartheta^2 \overset{n\to\infty}{\longrightarrow} 0.$$

Dies beweist die Konsistenz.

e) Aus $Y_i = \log(1+X_i)$ folgt $F_Y(y) = P[\log(1+X) \le y] = F_X(\exp(y)-1)$. Dies impliziert, dass $f_Y(y) = f_X(\exp(y)-1))\exp(y) = \frac{1}{\vartheta}\exp(y)^{-1/\vartheta} = \frac{1}{\vartheta}\exp(-\frac{1}{\vartheta}y)$.

3. Wir betrachten eine stetige Verteilung mit der Dichte

$$f(x) = \begin{cases} \frac{\alpha}{x^{\alpha+1}} & x \geq 1\,, \\ 0 & x < 1\,, \end{cases}$$

wobei $\alpha > 0$ ein unbekannter Parameter ist. Wir wollen den Parameter α aus einer Stichprobe schätzen.

a) Bestimme die Likelihood- und die log-Likelihood-Funktion basierend auf Beobachtungen $x_1, \ldots, x_n$ von n unabhängigen identisch verteilten einer Zufallsvariablen mit obiger Dichte.

b) Bestimme den zugehörigen Maximum-Likelihood-Schätzer für α. Schreibe zuerst die allgemeine Formel für n Beobachtungen hin und berechne den Schätzwert dann für die folgende konkrete Stichprobe:

x_1	x_2	x_3	x_4	x_5
12.0	4.0	6.9	27.9	15.4

c) Bestimme den Momentenschätzer für α, wieder zuerst allgemein basierend auf n Beobachtungen $x_1, \ldots, x_n$ und dann für die obige Stichprobe.

Hinweis: Der Momentenschätzer setzt voraus, dass $\alpha > 1$ ist; warum?

d) Vergleiche den Maximum-Likelihood- und den Momentenschätzer für obige Stichprobe. Ist der Momentenschätzer hier sinnvoll?

Lösung:

a) Die Likelihood-Funktion ergibt sich aus dem Produkt der Dichten:

$$L(x_1, \ldots x_n; \alpha) = \prod_{i=1}^{n} f(x_i) = \prod_{i=1}^{n} \frac{\alpha}{x_i^{\alpha+1}} = \alpha^n \frac{1}{(\prod_{i=1}^{n} x_i)^{\alpha+1}}\ .$$

Beachte: wir können immer die Formel für $x_i \geq 1$ verwenden; der Fall $x_i < 1$ tritt nur mit Wahrscheinlichkeit 0 ein.

Die log-Likelihood-Funktion erhalten wir durch Logarithmieren obiger Formel als

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

b) Ableiten und Nullsetzen der log-Likelihood-Funktion gibt

$$\frac{\partial}{\partial \alpha} \ell(x_1, \ldots x_n; \alpha) = \frac{n}{\alpha} - \sum_{i=1}^{n} \log x_i = 0$$

$$\Rightarrow \widehat{\alpha}_{\text{MLE}} = \frac{n}{\sum_{i=1}^{n} \log X_i}.$$

Der Schätzwert ist somit

$$\widehat{\alpha}_{\text{MLE}}(\omega) = \frac{n}{\sum_{i=1}^{n} \log x_i} = 0.4214\,.$$

c) Der Erwartungswert einer Zufallsvariablen X mit obiger Dichte f ist

$$E(X) = \int_{x=-\infty}^{\infty} xf(x)dx = \int_{x=1}^{\infty} x\frac{\alpha}{x^{\alpha+1}}dx = -\left.\frac{\alpha}{\alpha-1}\frac{1}{x^{\alpha-1}}\right|_{x=1}^{\infty} = \begin{cases} \frac{\alpha}{\alpha-1}, & \text{falls } \alpha > 1, \\ \infty, & \text{falls } \alpha \leq 1. \end{cases}$$

Da der Erwartungswert für $\alpha \leq 1$ gleich ∞ ist, ist der Momentenschätzer für diesen Fall nicht definiert.

Gleichsetzen von Erwartungswert und Stichprobenmittel, $E(X) = \frac{1}{n}\sum_{i=1}^{n} X_i$, ergibt

$$\widehat{\alpha}_{\text{MoM}} = \frac{\overline{X}}{\overline{X}-1}$$

und

$$\widehat{\alpha}_{\text{MoM}}(\omega) = 1.0817.$$

d) Der Momentenschätzer setzt $\alpha > 1$ voraus, was unplausibel erscheint, da der MLE deutlich kleiner als 1 ist.

4. Wir betrachten die geometrische Verteilung. Zur Erinnerung: dies ist eine diskrete Verteilung mit Wahrscheinlichkeitsfunktion

$$P(X = x) = (1-p)^{x-1}p, \quad x = 1, 2, 3, \dots,$$

wobei $0 < p < 1$ die Erfolgswahrscheinlichkeit ist. Eine geometrisch verteilte Zufallsvariable beschreibt die Anzahl der Versuche bis zum ersten Erfolg bei unabhängigen Bernoulliversuchen mit Erfolgswahrscheinlichkeit p. Wir wollen den Parameter p aus einer Stichprobe schätzen.

a) Bestimme die Likelihood-Funktion basierend auf Beobachtungen $x_1, \dots, x_n$ von n unabhängigen identisch verteilten geometrisch verteilten Zufallsvariablen.

b) Bestimme den zugehörigen Maximum-Likelihood-Schätzer für p.

c) Bestimme den Momentenschätzer für p (wieder basierend auf n Beobachtungen $x_1, \dots, x_n$). Vergleiche mit dem Maximum-Likelihood-Schätzer.

d) Angenommen, du hast nur eine einzige Beobachtung x einer geometrisch verteilten Zufallsvariablen, also ist $n = 1$. Wir können das Experiment dann auch folgendermassen interpretieren: es wurden x unabhängige Bernoulli-Versuche durchgeführt, und dabei wurde genau ein Erfolg beobachtet. Schreibe die Likelihood-Funktion für dieses Experiment auf: was ist der Unterschied zu der in 4. gefundenen Likelihood-Funktion? Warum erhältst du den gleichen Maximum-Likelihood-Schätzer?

Lösung:

a) Die Likelihood-Funktion ist

$$L(x_1, \dots x_n; p) = \prod_{i=1}^{n}(1-p)^{x_i-1}p = (1-p)^{\sum_{i=1}^{n} x_i - n}p^n = (1-p)^{\sum_{i=1}^{n} x_i}(\frac{p}{1-p})^n.$$

b) Die log-Likelihood-Funktion ist

$$\ell(x_1, \ldots, x_n; p) = \log L(x_1, \ldots, x_n; p) = \left(\sum_{i=1}^{n} x_i\right) \log(1-p) + n(\log p - log(1-p)).$$

Ableiten und Nullsetzen gibt

$$\begin{aligned} \frac{\partial}{\partial p}\ell(x_1, \ldots, x_n; p) &= \frac{-\sum_{i=1}^n x_i}{1-p} + n\left(\frac{1}{p} + \frac{1}{1-p}\right) =: 0 \\ \iff \quad & \frac{-\sum_{i=1}^n x_i}{n} + \frac{1-p}{p} + 1 = 0 \end{aligned}$$

und somit

$$\hat{p} = \frac{n}{\sum_{i=1}^n X_i}.$$

c) Der Erwartungswert einer geometrisch verteilten Zufallsvariablen X ist $E(X) = \frac{1}{p}$, siehe Skript.
Gleichsetzen der Momente, $\mathbb{E}(X) = \frac{1}{n}\sum_{i=1}^n X_i$, ergibt

$$\widehat{p}_{\text{MoM}} = \frac{n}{\sum_{i=1}^n X_i} = \frac{1}{\overline{X}_n},$$

also $\widehat{p}_{\text{MLE}} = \widehat{p}_{\text{MoM}}$.

d) Die Likelihood-Funktion ist

$$L(p; x) = \binom{x}{1}(1-p)^{x-1}p = x(1-p)^{x-1}p.$$

Die Likelihood-Funktionen hier und in 4. unterscheiden sich nur durch eine Proportionalitätskonstante, die nicht von p abhängt, und daher werden die Likelihoods als Funktionen in p für das gleiche Argument maximiert.