Koordinator
Konstantin Rusch

Wahrscheinlichkeit & Statistik

Serie 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.

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

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

[1] Basierend auf n Beobachtungen von X.

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 $y \geq 0$.

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)**.

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?

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

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

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 $x_1, \ldots, x_n$ 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 $x_1, \ldots, x_n$). 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 a) gefundenen Likelihood-Funktion? Warum erhältst du den gleichen Maximum-Likelihood-Schätzer?