Übungsblatt 1

Abgabe in Zweiergruppen (Namen nicht vergessen) per mail an `hoepfner@mathematik.uni-mainz.de`

bis **MO 20.05.19** , Besprechung am FR 24.05.19

Aufgabe 1.1 (ein einfacher Minimum-Distanz-Schätzer): Die Laplace-Transformierte

$$t \longrightarrow E_\lambda\left(e^{-tX}\right) \quad , \quad X \sim \mathcal{E}xp(\lambda)$$

der Exponentialverteilung mit Parameter $\lambda > 0$ ist auf einer Obermenge D von $[0,\infty)$ definiert; wir betrachten sie auf $[0,\infty)$ und schreiben

$$\text{(1)} \qquad \kappa_\lambda(\cdot): \quad [0,\infty) \ni t \longrightarrow \left(\frac{\lambda}{\lambda+t}\right) \in (0,1] \, .$$

Die empirische Laplace-Transformierte nach Beobachtung von $X_1,\ldots,X_n$ ist

$$\text{(2)} \qquad \widehat{\kappa}_n(\cdot): \quad [0,\infty) \ni t \longrightarrow \frac{1}{n}\sum_{i=1}^{n} e^{-\lambda X_i} \in (0,1] \, .$$

Für ein geeignetes Intervall $0 < T < \infty$ und ein geeignetes auf $[0,T]$ konzentriertes endliches Mass μ betrachte man die Funktionen

$$\kappa_\lambda \, , \; \lambda > 0 \quad , \quad \widehat{\kappa}_n \, , \; n \geq 1$$

als Elemente des Hilbertraums $H := L^2([0,T],\mu)$, und schätze den unbekannten Parameter $\lambda > 0$ im statistischen Modell

$$(\star) \qquad \mathcal{E} := \left(I\!R^n, \mathcal{B}(I\!R^n), \left\{ P_\lambda^n := \bigotimes_{i=1}^{n} \mathcal{E}xp(\lambda) \; : \; \lambda > 0 \right\} \right)$$

durch einen Minimum-Distanz-Schätzer

$$\text{(3)} \qquad \widehat{\lambda}_n \; := \; \text{Minimalstelle der Abbildung} \quad \lambda \longrightarrow \|\widehat{\kappa}_n - \kappa(\lambda)\|_H \, .$$

a) Graphik 1: Wähle $\widetilde{T} = 5$, plotte einen Rahmen $[0,\widetilde{T}] \times [0,1]$ für die Darstellung von Laplace-Transformierten von Exponentialverteilungen auf $[0,\widetilde{T}]$, und zeichne in diesen Rahmen für 15 aussagekräftig gewählte λ-Werte die Funktionen $t \to \kappa_\lambda(t)$ in unterscheidbaren Farben ein. Eine Legende soll den Parameter λ für jede der dargestellten Funktionen sichtbar machen.

b) Für $\lambda_0 := 0.643$ und $n = 117$ simuliere man iid ZV $X_1, \ldots, X_n$ iid $\sim \mathcal{E}xp(\lambda_0)$ und bilde die empirische Laplace-Transformierte (2). Wie fügt sich diese in das Tableau aus a) ein?

c) Graphik 2: Mit der empirischen Laplace-Transformierten aus b) berechne und plotte man die Funktion aus (3)

$$(4) \qquad [0,T] \ni \lambda \quad \longrightarrow \quad \|\widehat{\kappa}_n - \kappa(\lambda)\|_H \in I\!R^+$$

über einem geeigneten Intervall $[0,T]$. Man überzeuge sich graphisch, dass die Funktion (4) eine eindeutige Minimalstelle $\widehat{\lambda}_n$ besitzt, und ermittle diese aus der Graphik.
Man experimentiere dabei mit verschiedenen Massen μ auf $[0,T]$, insbesondere solchen, die Umgebungen von 0^+ im Vergleich zu anderen Stellen in $[0,T]$ mit höherem Gewicht belegen (Frage: warum sollte das Mass μ bevorzugt auf die Stelle 0^+ schauen?).

d) Graphik 3: Hat man sich in c) für ein geeignetes Paar (μ, T) entschieden, so halte man dieses fest und wiederhole $N = 500$ mal die Schritte b) und c): in jedem Simulationslauf ermittle man den Wert $\widehat{\lambda}_n$ des Minimum-Distanz-Schätzers.
Man speichere die N Schätzwerte als Datensatz `meineMDEschaetzwerte` ab. Als Approximation an die wahre Verteilung des gewählten Minimum-Distanz-Schätzers an der Stelle λ_0 betrachte man ein Histogramm von `meineMDEschaetzwerte` in geeigneter Zelleinteilung und geeigneter Beschriftung.

Abgabe: Programm und Graphiken (bitte die drei Graphiken zu einem pdf zusammenbinden)

Aufgabe 1.2 (Maximum-Likelihood-Schätzer und Bayes-Schätzer):

Betrachte ein parametrisches statistisches Modell

$$\mathcal{E} := \left(I\!R^n, \mathcal{B}(I\!R^n), \left\{ P_\vartheta^n := \bigotimes_{i=1}^n P_\vartheta \,:\, \vartheta \in \Theta \right\} \right)$$

und bezeichne die Beobachtungen in diesem Modell als $X = (X_1, \ldots, X_n)$. Setze voraus, dass für jedes $\vartheta \in \Theta$ eine produktmessbare Lebesgue-Dichte $p(\vartheta, y)dy = P_\vartheta(dy)$ gegeben ist. Bei vorliegender Beobachtung X betrachte man die Likelihoodfunktion

$$\Theta \ni \vartheta \quad \longrightarrow \quad L_n(\vartheta, X) := \prod_{i=1}^n p(\vartheta, X_i) \in [0, \infty) \,.$$

Ein Maximum-Likelihood-Schätzer (MLE) für den unbekannten Parameter (falls ein solcher existiert) ist ein Schätzer $T_{n,1}$ mit der Eigenschaft

$$T_{n,1} \text{ liefert eine Maximalstelle der Likelihoodfunktion } \vartheta \longrightarrow L_n(\vartheta, X).$$

Ein Bayes-Schätzer (BE) für den unbekannten Parameter (falls ein solcher existiert) ist ein Schätzer $T_{n,2}$ mit der Eigenschaft

$$T_{n,2} \;=\; \frac{\int_\Theta \vartheta \, L_n(\vartheta, X)\, d\vartheta}{\int_\Theta L_n(\vartheta, X)\, d\vartheta} \;;$$

dazu muss insbesondere ein Quotient von Integralen wie auf der rechten Seite wohldefiniert sein. Ist dies der Fall, sieht man

$$\vartheta \quad \longrightarrow \quad \frac{L_n(\vartheta, X)}{\int_\Theta L_n(\vartheta, X)\, d\vartheta}$$

als Dichte einer Wahrscheinlichkeitsverteilung auf Θ an und nennt diese die a posteriori- Verteilung auf Θ bei vorliegender Beabachtung X: ein Bayes-Schätzer ist also Mittelwert einer a-posteriori Verteilung.

a) Für das Exponentialmodell $(\star)$ aus Aufgabe 1.1 beweise man:

i) der Maximum-Likelihood-Schätzer existiert im Modell $(\star)$ und hat die explizite Form

$$(5) \qquad T_{n,1} \;=\; \left(\overline{X}_n\right)^{-1} = \frac{n}{\sum_{i=1}^n X_i} \;.$$

ii) der Bayes-Schätzer existiert im Modell $(\star)$ und hat die explizite Form

$$(6) \qquad T_{n,2} \;=\; \frac{n+1}{\sum_{i=1}^n X_i} \;.$$

b) Die Verteilung $\mathcal{L}(T_{n,i}|P^n_\lambda)$ der Schätzer (5) und (6) im Modell $(\star)$ unter wahrem Parameter λ –und sogar eine explizite Darstellung der Lebesgue-Dichte durch Ableiten nach t– erhält man aus

$$\begin{aligned} P_\lambda\left(T_{n,1} \le t\right) \;&=\; 1 - \Gamma(n,\lambda)\left(\left(0\,,\,\frac{n}{t}\right]\right) \quad , \quad t > 0 \\ P_\lambda\left(T_{n,2} \le t\right) \;&=\; 1 - \Gamma(n,\lambda)\left(\left(0\,,\,\frac{n+1}{t}\right]\right) \quad , \quad t > 0\,. \end{aligned}$$

Für $\lambda = \lambda_0$ aus Aufgabe 1.1 b) plotte man die Dichte von $\mathcal{L}(T_{n,1}|P^n_{\lambda_0})$ in einem Masstab, der einen direkten Vergleich mit dem Histogramm aus Aufgabe 1.1 d) erlaubt. Welcher der beiden Schätzer führt zu geringeren Schätzfehlern ?

Abgabe: nur Teil b): Programm zur Berechnung der Dichte, Graphik.