Σ = 62 Punkte

Zum Bestehen: 31 Punkte

ulm university universität uulm

Prof. Dr. Evgeny Spodarev
Dipl.-Math. Stefan Roth

SS 2015
27. Juli 2015

Stochastik I - Klausur 1

Bearbeitungszeit: 120 Minuten.

Aufgabe 1

Betrachte den **R**-Quelltext weiter unten.

(a) Gib eine Formel für die Werte von `f(x)` an (Zeilen 1-8). Wie wird `f()` genannt?

(b) Welche Größe berechnet die Funktion `G()` (Zeilen 9-13)? Gib auch hier eine Formel an.

(c) Erkläre den Wert von `G(x)` anhand einer Skizze.

(d) Interpretiere `f()` und `G()` anhand eines Beispiels.

Quelltext zu Aufgabe 1:

```
1  f<-function(x){
2   y<-0
3   s<-sort(x)
4   for (j in 1:length(x)){
5      y<-c(y, sum(s[1:j]) / (length(x) * mean(x)))
6   }
7  return(y)
8  }

9  G<-function(x){
10   y<-f(x)
11   g<-(length(x) + 1) / length(x) - 2 * sum(y) / length(x)
12 return(g)
13 }

15 x<-c(12, 4, 1, 8, 22)
16 y<-rep(1,5)
17 z<-c(rep(0, 4), 1)

16 G(x)
17 [1] 0.4255319
18 G(y)
19 [1] 0
20 G(z)
21 [1] 0.8
```

Aufgabe 2

Seien $X_1, \ldots, X_n$ unabhängig und identisch $N(0, \sigma^2)$-verteilt, $\sigma^2 > 0$. Beobachtbar seien aber nur die Zufallsvariablen $Y_i = \log |X_i|$, $i = 1, \ldots, n$.

(a) Zeige, dass die Dichte von Y_1 gegeben ist durch

$$f(t) = \left(\frac{2}{\pi}\right)^{1/2} \exp\left\{-\frac{1}{2}e^{2(t-\log\sigma)} + t - \log\sigma\right\}, \quad t \in \mathbb{R}.$$

(b) Bestimme den ML-Schätzer $\hat{\sigma}$ für σ (basierend auf $Y_1, \ldots, Y_n$).

(c) Ist $\hat{\sigma}$ ein stark konsistenter Schätzer für σ? Begründe deine Antwort!

Aufgabe 3

Es seien $X_1, \ldots, X_n$ unabhängig und identisch verteilt mit Dichte $f(x) = 3\lambda x^2 e^{-\lambda x^3} \mathbb{1}_{(0,\infty)}(x)$, $x \in \mathbb{R}$.

(a) Bestimme einen M-Schätzer $\tilde{\lambda}$ für λ.

(b) Bestimme den ML-Schätzer $\hat{\lambda}$ für λ und zeige, dass $\bar{\lambda} = (1 - 1/n)\hat{\lambda}$ erwartungstreu für λ ist.

(c) Nimm nun an, dass $\bar{\lambda}$ ein vollständiger Schätzer ist. Ist $\bar{\lambda}$ bester erwartungstreuer Schätzer für λ? Begründe deine Antwort.

Aufgabe 4

Es seien $X_1, \ldots, X_n$ unabhängig und identisch $N(0, \sigma^2)$-verteilt, $\sigma^2 > 0$. Beobachtbar seien aber nur $Y_i = X_i^2$, $i = 1, \ldots, n$.

(a) Bestimme ein exaktes Konfidenzintervall für σ^2 zum Niveau $1 - \alpha$, welches nur von $Y_1, \ldots, Y_n$ abhängt.

(b) Bestimme ein asymptotisches Konfidenzintervall für σ^2 zum Niveau $1 - \alpha$ welches nur von $Y_1, \ldots, Y_n$ abhängt.

Aufgabe 5

Sei $X_1, \ldots, X_n$ eine i.i.d. Zufallsstichprobe mit differenzierbarer Verteilungsfunktion F und Dichte $f = F' = dF/dt$. Sei $\hat{f}_n$ der Kerndichteschätzer für f mit Bandbreite $h > 0$ und Rechteckskern K, d.h. $K(x) = 1/2\mathbb{1}_{[-1,1]}(x)$, $x \in \mathbb{R}$.

(a) Bestimme $\mathbb{E}\hat{f}_n(x)$ für $x \in \mathbb{R}$.

(b) Verwende (a) um zu zeigen, dass

$$\lim_{h \downarrow 0} \mathbb{E}\hat{f}_n(x) = f(x),$$

für jedes $x \in \mathbb{R}$.

Aufgabe 6 (Multiple Choice)

Jede richtig bewertete Aussage +1 Punkt, jede falsch bewertete -1 Punkt. Minimal 0 Punkte.

1.) Sind $X_1, \ldots, X_n$ unabhängig und identisch verteilte Zufallsvariablen mit Verteilungsfunktion F, so gilt stets $\lim_{n\to\infty} \hat{F}_n(x) = F(x)$, gleichmäßig in $x \in \mathbb{R}$, wobei $\hat{F}_n$ die empirische Verteilungsfunktion von $X_1, \ldots, X_n$ bezeichne.

2.) Seien $X_1, \ldots, X_n$ unabhängig und identisch verteilte Zufallsvariablen mit endlichem zweiten Moment und bezeichne $\sigma^2 = VarX_1$. Dann ist $\tilde{S}_n^2 = 1/n \sum_{k=1}^n (X_k - \bar{X}_n)^2$ ein erwartungstreuer Schätzer für σ^2.

3.) Sind $X_1, \ldots, X_n \overset{\text{i.i.d.}}{\sim} Exp(\lambda)$, $\lambda > 0$, so gilt: $n\tilde{S}_n^2/\sigma^2 \sim \chi_n^2$, wobei $\tilde{S}_n^2$ und σ^2 wie in 2.) definiert seien.

4.) Für zwei unabhängige Zufallsvariablen $X \sim \Gamma(\lambda, p_1)$ und $Y \sim \Gamma(\lambda, p_2)$ gilt $X+Y \sim \Gamma(2\lambda, p_1+p_2)$.

5.) Für $X_1, \ldots, X_n \overset{\text{i.i.d.}}{\sim} N(\mu, \sigma^2)$ gilt: $\bar{X}_n$ und $S_n^2 = 1/(n-1) \sum_{k=1}^n (X_k - \bar{X}_n)^2$ sind unabhängig.

6.) Für einen erwartungstreuen Schätzer $\hat{\theta}(X_1 \ldots, X_n)$, der die Bedingungen für die Ungleichung von Cramér-Rao erfüllt, gilt stets:

$$Var_\theta \hat{\theta}(X_1, \ldots, X_n) \geq \frac{1}{n^2 I(\theta)},$$

wobei $I(\theta)$ die Fisher-Information bezeichne.

7.) Jeder erwartungstreue, vollständige und suffiziente Schätzer ist bester erwartungstreuer Schätzer.

8.) Sei $X_1, \ldots, X_n$ eine i.i.d. Zufallsstichprobe. Dann ist die Verteilungsfunktion $F_{X_{(i)}}$ der i-ten Ordnungsstatistik $X_{(i)}$ gegeben durch

$$F_{X_{(i)}}(x) = \sum_{k=i}^{n} \binom{n}{k} F^k(x)(1-F(x))^{n-k}, \quad x \in \mathbb{R},$$

wobei F die Verteilungsfunktion von X_1 bezeichne.

9.) Sind $X_1, \ldots, X_n$ unabhängig und identisch verteilt mit Verteilungsfunktion F, so gilt $n\hat{F}_n(x) \sim Bin(n, F(x))$, für jedes $x \in \mathbb{R}$.

10.) Ist $\hat{\theta}$ ein vollständiger Schätzer für $\theta \in \Theta \subseteq \mathbb{R}^m$ und $g : \mathbb{R}^m \to \mathbb{R}^m$ eine messbare Abbildung, so ist auch $g(\hat{\theta})$ ein vollständiger Schätzer für θ.

Stochastik I

14.07.2015

1. Klausur

Musterlösung

(a) Stichprobe $x = (x_1, \dots, x_n)$, $s = (x_{(1)}, \dots, x_{(n)})$

$$y = (y_1, \dots, y_n) = f(x)$$

$$y_j = \frac{\sum_{k=1}^{j} x_{(k)}}{n \cdot \bar{x}_n}, \quad j = 1, \dots, n \qquad (1)$$

Lorenzkurve (1)

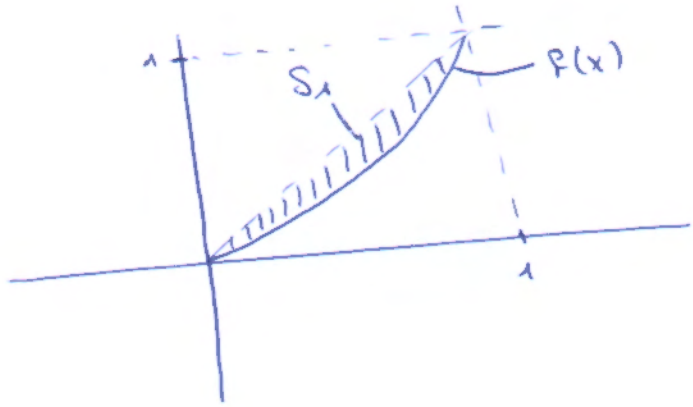
(b) x wie in (a).

$$g(x) = \frac{n+1}{n} - 2 \cdot \bar{y}_n \qquad (1)$$

mit y wie in (a)

Gini - Koeffizient (1)

(c)



(2)

$$G = \frac{S_1}{1/2} = 2S_1 \qquad (1)$$

Bei Gleichverteilung $f(x)$ = Diagonale und $G = 1$.

(d) Bsp. Einkommensverteilung:

$x_1, \dots, x_n$ Einkommen von n Personen

Gesamteinkommen $n \cdot \bar{x}_n$ (1)

y_j - Anteil der j kleinsten Einkommen

am Gesamteinkommen (1)

Je kleiner G, desto gleichmäßiger ist

das Gesamteinkommen auf alle Arbeit-

nehmer verteilt. (1)

(2) $X_1, \dots, X_n \overset{i.i.d.}{\sim} N(0, \sigma^2)$, $\sigma^2 > 0$, $Y_i = \log|X_i|$

(a) $F(t) = P(Y_i \leq t) = P\left(\log \frac{|X_i|}{\sigma} \leq t - \log\sigma\right)$

$= P\left(\frac{|X_i|}{\sigma} \leq e^{t - \log\sigma}\right)$

$\frac{X_i}{\sigma} \sim N(0,1) \ \forall i$

$= 2\Phi\left(e^{t-\log\sigma}\right) - 1$ (2)

$\Rightarrow f(t) = e^{t-\log\sigma} \cdot \Phi'\left(e^{t-\log\sigma}\right)$

$= 2 \frac{1}{\sqrt{2\pi}} e^{t-\log\sigma} \cdot \exp\left(-\frac{1}{2} e^{2(t-\log\sigma)}\right)$

$= \left(\frac{2}{\pi}\right)^{1/2} \cdot \frac{1}{\sigma} \cdot \exp\left(-\frac{1}{2\sigma^2} e^{2t} + t\right)$

$= \left(\frac{2}{\sigma^2\pi}\right)^{1/2} \exp\left(-\frac{1}{2\sigma^2} e^{2t} + t\right)$ (2)

(b) $L(x_1, \dots, x_n; \sigma^2) = \left(\frac{2}{\pi}\right)^{n/2} \cdot \frac{1}{\sigma^n} \cdot \exp\left(-\frac{1}{2\sigma^2} \sum_{i=1}^n e^{2x_i} + n\bar{x}_n\right)$ (1)

$$\Rightarrow \log L = \frac{n}{2} \log \frac{2}{\pi} - n \log \sigma + n\bar{x}_n - \frac{1}{2\sigma^2} \sum_{i=1}^n e^{2x_i} \quad (1)$$

$$\Rightarrow \frac{\partial}{\partial \sigma} \log L = -\frac{n}{\sigma} + \frac{1}{\sigma^3} \sum_{i=1}^n e^{2x_i}$$

$$\overset{!}{=} 0 \overset{\sigma \neq 0}{\Leftrightarrow} \hat{\sigma} = \left(\frac{1}{n} \sum_{i=1}^n e^{2x_i}\right)^{1/2} \quad (1)$$

$$\Rightarrow \frac{\partial^2}{\partial^2 \sigma} \log L \Big|_{\sigma = \hat{\sigma}} = \frac{n}{\sigma^2} - 3 \frac{1}{\sigma^4} \sum_{i=1}^n e^{2x_i} \Big|_{\sigma = \hat{\sigma}}$$

$$= n^2 \frac{1}{\sum_{i=1}^n e^{2x_i}} - 3n^2 \frac{1}{\sum_{i=1}^n e^{2x_i}}$$

$$= -\frac{2n^2}{\sum_{i=1}^n e^{2x_i}} < 0 \quad (1)$$

ML-Schätzer: $\hat{\sigma}(Y_1, \dots, Y_n) = \left(\frac{1}{n} \sum_{i=1}^n e^{2Y_i}\right)^{1/2}$

(c) Nach dem SGGZ gilt:

$$\hat{\sigma} \xrightarrow{f.s.} \left(\mathbb{E}\, e^{2Y_1}\right)^{1/2} \; (n \to \infty) \quad (1)$$

$$\mathbb{E}\, e^{2Y_1} = \mathbb{E}\, X_1^2 = \sigma^2 \quad (1)$$

$$\Rightarrow \hat{\sigma} \xrightarrow{f.s.} \sigma \quad (1)$$

$\square$

③ $X_1, \dots, X_n$ i.i.d. mit Dichte $f(x) = 3\lambda x^2 e^{-\lambda x^3}$, $x > 0$.

(a) $$\mathbb{E}X_1 = \int_0^\infty 3\lambda x^3 e^{-\lambda x^3} dx$$

$$\overset{y = \lambda x^3}{=} \int_0^\infty 3y \cdot \frac{1}{3\lambda x^2} e^{-y} dy$$

$$\frac{dy}{dx} = 3\lambda x^2$$

$$dx = \frac{1}{3\lambda x^2} dy$$

$$\overset{x = (y/\lambda)^{1/3}}{=} \frac{1}{\lambda} \int_0^\infty \frac{\lambda^{2/3}}{y^{2/3}} \cdot y \, e^{-y} dy$$

$$= \lambda^{-1/3} \int_0^\infty y^{1/3} e^{-y} dy$$

$$= \frac{\Gamma(4/3)}{\lambda^{1/3}} \qquad (2)$$

$$\bar{X}_n \overset{!}{=} \frac{\Gamma(4/3)}{\lambda^{1/3}} \iff \lambda = \left(\frac{\Gamma(4/3)}{\bar{X}_n}\right)^3 \qquad (2)$$

$\Rightarrow$ M-Schätzer: $\tilde{\lambda}(X_1, \dots, X_n) = \left(\frac{\Gamma(4/3)}{\bar{X}_n}\right)^3$

(b) $$L(x_1, \dots, x_n; \lambda) = (3\lambda)^n x_1^2 \cdot \ldots \cdot x_n^2 e^{-\lambda \sum_{i=1}^n x_i^3} \qquad (1)$$

$$\Rightarrow \log L = n \log(3\lambda) + 2 \sum_{i=1}^n \log x_i - \lambda \sum_{i=1}^n x_i^3 \qquad (1)$$

$$\frac{\partial}{\partial \lambda} \log L = \frac{n}{\lambda} - \sum_{i=1}^n x_i^3 \overset{!}{=} 0$$

$$\iff \lambda = \frac{n}{\sum_{i=1}^n x_i^3} \qquad (1)$$

$$\frac{\partial^2}{\partial^2 \lambda} \log L = -\frac{n}{\lambda^2} < 0 \qquad (1)$$

ML - Schätzer: $\hat{\lambda}(X_1, \dots, X_n) = \dfrac{n}{\sum_{i=1}^{n} X_i^3}$

$$Y_i = X_i^3 \sim Exp(\lambda)$$

$$\Big[P(Y_i \le t) = P(X_i \le \sqrt[3]{t})$$

$$\Rightarrow f_{Y_i}(t) = \frac{1}{3\, t^{2/3}}\, 3\, \sqrt[3]{t}^{\,2} \lambda\, e^{-\lambda \sqrt[3]{t}^{\,3}}$$

$$= \lambda e^{-\lambda t} \Big]$$

$$\Rightarrow \sum_{i=1}^{n} X_i^3 = n \cdot \overline{Y}_n \sim Erl(\lambda, n) \qquad (1)$$

$$\Rightarrow \mathbb{E} \frac{n}{\sum_{i=1}^{n} X_i^3} = n \int_0^\infty \frac{\lambda^n x^{n-2}}{(n-1)!} e^{-\lambda x}\, dx \qquad (2)$$

$$= \frac{n}{n-1} \lambda \quad \Rightarrow \quad \frac{n-1}{n} \hat{\lambda} = \left(1 - \frac{1}{n}\right) \hat{\lambda} = \overline{\lambda}$$

ist erwartungstreu!

(c) $\overline{\lambda}$ ist suffizient für $\underline{n \ge 2}$:

$$L(x_1, \dots, x_n; \lambda) = \underbrace{\left(\prod_{i=1}^{n} x_i^2 \right)}_{\substack{= h(x_1, \dots, x_n) \\ (1)}} \cdot \underbrace{(3\lambda)^n \exp\left(- \frac{\lambda (n-1)}{\overline{\lambda}} \right)}_{\substack{= g(\lambda, \overline{\lambda}) \\ (1)}}$$

N.-F.
$\Rightarrow$ $\overline{\lambda}$ suffizient. (1)

Da $\overline{\lambda}$ e-treu und vollst. ist (1)

L.-S.
$\Rightarrow$ $\overline{\lambda}$ bester e-treuer Schätzer für λ.

④ $X_1, \dots, X_n \overset{i.i.d.}{\sim} N(0, \sigma^2)$, $\sigma^2 > 0$, $Y_i = X_i^2$, $i = 1, \dots, n$

(a) Es gilt: $\sum_{i=1}^{n} \frac{Y_i}{\sigma^2} = \sum_{i=1}^{n} \left(\frac{X_i}{\sigma}\right)^2 \sim \chi_n^2$ (1)

$$\Rightarrow 1 - \alpha = P\left(\frac{1}{\sigma^2} \sum_{i=1}^{n} Y_i \geq \chi^2_{n,\alpha} \right)$$

$$= P\left(\sigma^2 \leq \frac{n \overline{Y}_n}{\chi^2_{n,\alpha}} \right), \quad (2)$$

wobei $\chi^2_{n,\alpha}$ das α-Quantil der χ^2-Verteilung mit n Freiheitsgraden bezeichnet.

$$\Rightarrow KI = \left[0, \frac{n \overline{Y}_n}{\chi^2_{n,\alpha}}\right] \quad (1)$$

(b) $EY_i = EX_i^2 = \sigma^2$

$$\operatorname{Var} Y_i = \sigma^4 \operatorname{Var} \underbrace{\frac{X_i^2}{\sigma^2}}_{\sim \chi^2_1} = 2\sigma^4 \quad (1)$$

$$\Rightarrow 1 - \alpha \approx P\left(\frac{\overline{Y}_n - \sigma^2}{\sigma^2 \sqrt{2}} \sqrt{n} \geq z_\alpha \right)$$

$$= P\left(\frac{1}{\sigma^2} \overline{Y}_n \geq 1 + \frac{\sqrt{2}}{\sqrt{n}} z_\alpha \right) \quad (2)$$

$$= P\left(\sigma^2 \leq \frac{\overline{Y}_n}{1 + \frac{\sqrt{2}}{\sqrt{n}} z_\alpha} \right)$$

wo. $z_\alpha = \Phi^{-1}(\alpha)$.

$\Rightarrow$ asympt. $K = \left[0, \dfrac{\overline{Y}_n}{1+\frac{\sqrt{2}}{\sqrt{n}} z_\alpha}\right]$ (1)

<u>Bem.:</u> Für $n > 2z_\alpha^2$ ist $0 < 1 + \frac{\sqrt{2}}{\sqrt{n}} z_\alpha$.

(5) $X_1, \dots, X_n$ i.i.d., $K(x) = \frac{1}{2} \mathbb{1}_{[-1,1]}(x)$

(a) $\hat{f}_n(x) = \dfrac{1}{nh} \sum_{i=1}^{n} K\left(\dfrac{x-X_i}{h}\right)$

$= \dfrac{1}{2nh} \cdot \#\{i : x-h \le X_i \le x+h\}$

$= \dfrac{1}{2nh} \sum_{i=1}^{n} \mathbb{1}_{[x-h, x+h]}(X_i)$ (2)

$\Rightarrow \mathbb{E}\hat{f}_n(x) = \int_{\mathbb{R}} \dfrac{1}{2nh} \sum_{i=1}^{n} \mathbb{1}_{[x-h,x+h]}(y) f(y)\,dy$

$= \dfrac{1}{2nh} \sum_{i=1}^{n} \int_{x-h}^{x+h} f(y)\,dy$ (3)

$= \dfrac{1}{2h} \int_{x-h}^{x+h} f(y)\,dy = \dfrac{F(x+h)-F(x-h)}{2h}$

(b) Wg. (a) gilt

$\lim_{h \downarrow 0} \dfrac{F(x+h)-F(x-h)}{2h}$

$= \frac{1}{2} \lim_{h \downarrow 0} \left(\dfrac{F(x+h)-F(x)}{h} + \dfrac{F(x)-F(x-h)}{h} \right)$ (2)

$\frac{1}{2}(F'(x)+F'(x)) = F'(x) = f(x)$. (1)

6

1.) richtig

2.) falsch

3.) falsch

4.) falsch

5.) richtig

6.) richtig

7.) falsch

8.) richtig

9.) richtig

10.) richtig

je +1 bei richtiger Beantwortung & -1 bei falscher Beantw.

Minimal 0 Punkte