Tentamen Kres 4, 050109, 13:00–16:00

Antwoorden motiveren; kritieke gebieden en betrouwbaarheidsintervallen zo uitvoerig mogelijk beschrijven. Ieder onderdeel is 10 punten waard. Bij het tentamen mag alleen gebruik worden gemaakt van een (zelf mee te nemen) onbeschreven kopie van Appendix B van Bain en Engelhardt!

1. Gegeven is een onafhankelijke steekproef $X_1, \ldots, X_n$ van grootte n uit een verdeling met pdf
$$f_X(x;\theta) = \frac{1}{2\theta^3} x^2 \mathrm{e}^{-x/\theta} I_{[0,\infty)}(x), \qquad \theta > 0$$
   (a) Bepaal de MLE van θ. Is deze MSE consistent?
   (b) Geef de UMVUE van θ^{-1}.
   (c) Construeer een "gelijke staarten" $100\gamma\%$ betrouwbaarheidsinterval voor $e^{-\theta}$, gebaseerd op $S = \sum_{i=1}^n X_i$.
   (d) Laat zien dat de simultane pdf een MLR eigenschap bezit (in welke statistische grootheid?) en geef de UMP, grootte α, toets voor H_0: $\theta \geq 2$ tegen $\theta < 2$.
   (e) Geef de GLR toets voor H_0: $\theta = 1$ tegen H_a: $\theta \neq 1$.

2. Beschouw een onafhankelijke steekproef $X_1, \ldots, X_n$ uit de volgende verdeling (met $\theta > 0$):
$$f_X(x;\theta) = \frac{5x^4}{\theta^5} I_{[0,\theta]}(x)$$
   (a) Toon aan dat $S = X_{n:n}$ 'sufficent' is voor θ. Is S tevens 'complete'?
   (b) Geef de UMVUE van θ^2.
   (c) Bepaal de MLE van θ. Wat is de MSE van deze schatter?
   (d) Toon aan dat $X_{n:n}/\theta$ een spilfunctie is, en gebruik dit om een $100\gamma\%$ betrouwbaarheidsbovengrens voor θ af te leiden.
   (e) Geef de UMP, grootte α, toets voor H_0: $\theta \geq 3$ tegen H_a: $\theta < 3$. Schets de powerfunctie $\pi(\theta)$ van deze toets.

Succes!!!

Oplossingen

1.

$$f_X(x;\theta) = \frac{1}{2\theta^3} x^2 \mathrm{e}^{-x/\theta} I_{[0,\infty)}(x), \qquad \theta > 0$$

(a) De log-likelihoodvergelijking (eerste orde conditie) is

$$\frac{d}{d\theta}\ell(\theta) = \frac{d}{d\theta} \ln L(\theta) = -\frac{3n}{\theta} + \frac{\sum_{i=1}^n x_i}{\theta^2} = 0$$

dit geeft $\hat{\theta} = \frac{1}{3}\bar{x}$. Het betreft een maximum omdat $L(\theta) > 0$ en $\lim_{\theta\to 0} L(\theta) = \lim_{\theta\to\infty} L(\theta) = 0$. Deze schatter is zuiver: $E(\hat{\theta}) = \theta$ en z'n variantie gaat naar nul als $n \to \infty$: $\mathrm{Var}\hat{\theta} = \frac{1}{9n}\mathrm{Var}X = \frac{\theta^2}{3n} \underset{n\to\infty}{\longrightarrow} 0$, dus de MLE is inderdaad MSE consistent.

(b) We weten (REC) dat $S = \sum_{i=1}^n X_i$ C&S is voor θ. Verder geldt dat θ een schaalparameter is, dus een kandidaat UMVUE voor θ^{-1} is een veelvoud van S^{-1}. Omdat $S \sim \mathrm{GAM}(\theta, 3n)$, geldt

$$\begin{aligned} E(S^{-1}) &= \frac{1}{\Gamma(3n)\theta^{3n}} \int_0^\infty s^{3n-2} e^{-s/\theta} ds \\ &= \frac{1}{\Gamma(3n)\theta^{3n}} \frac{\theta^{3n-1}}{\Gamma(3n-1)} \int_0^\infty \frac{1}{\Gamma(3n-1)\theta^{3n-1}} s^{3n-2} e^{-s/\theta} ds \\ &= \frac{1}{(3n-1)\theta}, \end{aligned}$$

waaruit blijkt dat $\frac{3n-1}{S}$ UMVUE van θ^{-1} is.

(c) We weten dat $2S/\theta \sim \chi^2(6n)$, dus

$$\begin{aligned} \gamma &= P\left(\chi^2_{\frac{1-\gamma}{2}}(6n) < \frac{2S}{\theta} < \chi^2_{\frac{1+\gamma}{2}}(6n)\right) \\ &= P\left(e^{-\frac{2S}{\chi^2_{\frac{1-\gamma}{2}}(6n)}} < e^{-\theta} < e^{-\frac{2S}{\chi^2_{\frac{1+\gamma}{2}}(6n)}}\right). \end{aligned}$$

Het gevraagde betrouwbaarheidsinterval is dus

$$\left(e^{-\frac{2s}{\chi^2_{\frac{1-\gamma}{2}}(6n)}}, e^{-\frac{2s}{\chi^2_{\frac{1+\gamma}{2}}(6n)}}\right).$$

(d) Voor $\theta_1 < \theta_2$ vinden we

$$\frac{L(\theta_2)}{L(\theta_1)} = \left(\frac{\theta_1}{\theta_2}\right)^{3n} e^{-\left(\frac{1}{\theta_2} - \frac{1}{\theta_1}\right)\sum_{i=1}^n x_i}.$$

Omdat $-\left(\frac{1}{\theta_2}-\frac{1}{\theta_1}\right)>0$ heeft de familie van pdf's de MLR eigenschap in $S=\sum_{i=1}^n X_i$. De UMP toets voor H_0: $\theta\geq 2$ tegen H_a: $\theta<2$ bestaat eruit te verwerpen als s te klein is. Een toets met grootte α wordt verkregen door de verwerpingskans voor $\theta=2$ gelijk aan α te stellen. Voor $\theta=2$ geldt $2S/2\sim\chi^2(6n)$. De UMP toets verwerpt dus als $S\leq\chi^2_\alpha(6n)$.

(e) Introduceer $\theta_0=1$. De GLR toets verwerpt als λ te klein is, met

$$\begin{aligned}\lambda=\frac{L(\theta_0)}{L(\hat\theta)} &= \left(\frac{\hat\theta}{\hat\theta_0}\right)^{3n} e^{-\left(\frac{1}{\theta_0}-\frac{1}{\hat\theta}\right)\sum_{i=1}^n x_i}\\ &= \left(\frac{\hat\theta}{\hat\theta_0}\right)^{3n} e^{3n-3n\hat\theta}.\end{aligned}$$

Dit is maximaal voor $\hat\theta=1$, zoals is na te gaan uit de eerste orde conditie (het maximum van $\ln\lambda$ ligt bij $\hat\theta=1$). De GLR toets verwerpt als λ te klein is. Men kan gebruik maken van de benadering $-2\ln\lambda\sim\chi^2(1)$, dus verwerpen als $-2\ln\lambda\leq\chi^2_\alpha(1)$. Men kan ook als volgt te werk gaan: verwerp als $\hat\theta=\frac{\bar x}{3}$ te groot danwel te klein is. Onder H_0 geldt $2S\sim\chi^2(6n)$, dus verwerp als $s\leq\frac{1}{2}\chi^2_{\frac{\alpha}{2}}(6n)$ of als $s\geq\frac{1}{2}\chi^2_{1-\frac{\alpha}{2}}(6n)$.

2.

$$f_X(x;\theta)=\frac{5x^4}{\theta^5}I_{[0,\theta]}(x)$$

(a) De joint pdf is

$$f(x_1,\dots,x_n;\theta)=\frac{5^n}{\theta^{5n}}\prod_{i=1}^n x_i^4 I_{[0,\theta]}(x_{n:n}).$$

Dit is het product van $c(\theta)=\frac{5^n}{\theta^{5n}}$, $g(x_1,\dots,x_n)=\prod_{i=1}^n x_i^4$ en $h(x<1,\dots,x_n;\theta)=I_{[0,\theta]}(x_{n:n})$. Met behulp van het factorisatiecriterium volgt hieruit dat $S=X_{n:n}$ sufficient is voor θ. Is S tevens compleet? De CDF van S is $F_S(s;\theta)=\left(\frac{s}{\theta}\right)^{5n}$, en

$$f_s(s;\theta)=\frac{5ns^{5n-1}}{\theta^5 n}.$$

De eis $E(u(S))=0$ geeft als voorwaarde op $u(s)$:

$$\begin{aligned}E(u(S))=\int_0^\theta f_S(s;\theta)u(s)\mathrm{d}s &= 0\\ \Leftrightarrow \frac{5n}{\theta^{5n}}\int_0^\theta s^{5n-1}u(s)\mathrm{d}s &= 0\\ \Leftrightarrow \int_0^\theta s^{5n-1}u(s)\mathrm{d}s &= 0\\ \Rightarrow \theta^{5n-1}u(\theta)\mathrm{d}s &= 0,\end{aligned}$$

waarbij de laatste implicatie is verkregen d.m.v. differentiëren naar θ (m.b.v. de regel van Leibniz). Omdat dit voor alle $\theta > 0$ geldt, vinden we dus dat $E(u(S)) = 0$, $u(\theta) = 0$ voor alle $\theta > 0$ impliceert. S is dus tevens compleet.

(b) Geef de UMVUE van θ^2. In het algemeen hebben we

$$E(S^k) = \frac{5n}{\theta^{5n}} \int_0^\theta s^{5n+k-1} \mathrm{d}s = \frac{5n}{\theta^{5n}} \frac{\theta^{5n+k}}{5n+k} = \frac{5n}{5n+k}\theta^k.$$

zodat $E(S^2) = \frac{5n}{5n+2}\theta^2$, dus is $\frac{5n+2}{5n}S$ UMVUE voor θ^2.

(c) Bepaal de MLE van θ. Wat is de MSE van deze schatter? De likelihoodfunctie is

$$L(\theta) = \prod_{i=1}^n \frac{5^n}{\theta^n} \left(\prod_{i=1}^n x_i^4\right) I_{[0,\theta]}(x_{n:n}).$$

Dit is dalend in θ zolang $\theta \geq x_{n:n}$ (daaronder is de likelihoodfunctie nul). Er is dus sprake van een randmaximum: $\hat{\theta} = X_{n:n} = S$. Eerder was aangetoond dat $E(S^k) = \frac{5n}{5n+k}\theta^k$. We hebben dus $E(S) = \frac{5n}{5n+1}\theta$ (zodat de bias van S gegeven is door $b(S) = E(S) - \theta = \frac{-1}{5n+1}\theta$) en

$$\mathrm{Var}S = E(S^2) - (E(S))^2 = \left(\frac{5n}{5n+2} - \left(\frac{5n}{5n+1}\right)^2\right)\theta^2.$$

De MSE is

$$(b(S))^2 + \mathrm{Var}(S) = \left(\frac{5n}{5n+2} - \left(\frac{5n}{5n+1}\right)^2 + \left(\frac{1}{5n+1}\right)^2\right)\theta^2.$$

(d) Laat $Y = X/\theta$. De pdf van Y is te vinden met de transformatiemethode. $X = \theta Y$, $J = \theta$,

$$f_Y(y;\theta) = 5y^4 I_{[0,1]}(y).$$

Deze hangt niet van θ af. Het gevolg is dat de geschaalde steekproef $Y_1, \ldots, Y_n$ ook een verdeling heeft die vrij is van θ, en dus heeft ook $Y_{n:n} = X_{n:n}/\theta$ een verdeling die niet van θ afhangt. De CDF van Y is

$$F_Y(y) = y^5$$

en we weten (probability integral transform van KRES 3) dat

$$F_Y(Y) \sim \mathrm{UNIF}(0,1),$$

dus

$$\begin{aligned} \gamma &= P\left(\left(\frac{X_{n:n}}{\theta}\right)^5 > 1-\gamma\right) \\ &= P\left(X_{n:n}(1-\gamma)^{-1/5} > \theta\right), \end{aligned}$$

zodat de gevraagde ondergrens $x_{n:n}(1-\gamma)^{-1/5}$ blijkt te zijn.

(e) Geef de UMP toets voor H_0: $\theta \geq 3$ tegen H_a: $\theta < 3$. Schets de powerfunctie $\pi(\theta)$ van deze toets. Ook hier is sprake van een MLR eigenschap: Voor $\theta_2 > \theta_1$ geldt

$$\frac{f(x_1,\ldots,x_n;\theta_2)}{f(x_1,\ldots,x_n;\theta_1)} = \left(\frac{\theta_1}{\theta_2}\right)^{5n} \begin{cases} 1 & \text{als } x_{n:n} \leq \theta_1 \\ \infty & \text{als } \theta_1 < x_{n:n} \leq \theta_2. \end{cases}$$

De UMP toets voor H_0: $\theta \geq 3$ tegen H_a: $\theta < 3$ verwerpt als $x_{n:n}$ te klein is (zeg als $s \leq k$). Voor $\theta = 3$ geldt $F_S(s;3) = \left(\frac{s}{3}\right)^{5n}$, zodat k de oplossing is van

$$P_{\theta=3}\left(S \leq k\right) = \alpha$$

ofwel van

$$F_S(s;3) = \left(\frac{s}{3}\right)^{5n} = \alpha$$

Dit geeft $k = 3\alpha^{\frac{1}{5n}}$.