# Tentamen Kres 4, 220609, 14:00–17:00

*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 steekproef $X_1, \ldots, X_n$ uit de verdeling met pdf

$$f_X(x;\theta) = \frac{\theta^3}{2x^4}\mathrm{e}^{-\theta/x} I_{[0,\infty)}(x), \qquad \theta > 0$$

   (a) Bereken de CRLB voor zuivere schatters van $\theta$.

   (b) Bepaal de MME van $\theta$. Is deze schatter MSE consistent? Bereikt de variantie van de schatter de CRLB voor zuivere schatters van $\theta$?

   (c) Geef de UMVUE van $\theta^{-1}$.

   (d) Laat zien dat de joint pdf de MLR eigenschap bezit en geef de UMP, grootte $\alpha$, toets voor $H_0$: $\theta \leq 1$ tegen $\theta > 1$.

   (e) Geef een $100\gamma\%$ gelijke staarten betrouwbaarheidsinterval voor $\theta$, gebaseerd op de complete & sufficient statistic voor $\theta$.

2. Beschouw een onafhankelijke steekproef $X_1, \ldots, X_n$ uit de volgende verdeling:

$$f_X(x;\theta) = \frac{5\theta^5}{x^6} I_{[\theta,\infty)}(x), \qquad \theta > 0$$

   (a) Toon aan dat $S = X_{1:n}$ 'sufficent' is voor $\theta$. Is $S$ tevens 'complete'?

   (b) Geef de UMVUE van $\theta^2$.

   (c) Bepaal de MLE van $\theta$. Wat is de MSE van deze schatter?

   (d) Toon aan dat $X_{1:n}/\theta$ een spilfunctie is, en gebruik dit om een $100\gamma\%$ betrouwbaarheidsbovengrens voor $\theta$ af te leiden.

   (e) Geef de GLR, grootte $\alpha$, toets voor $H_0$: $\theta \leq 2$ tegen $H_a$: $\theta > 2$.

## Succes!!!

**Oplossingen**

1.

$$f_X(x;\theta) = \frac{\theta^3}{2x^4}\mathrm{e}^{-\theta/x}I_{[0,\infty)}(x), \qquad \theta > 0$$

(a) Gevraagd is de CRLB voor zuivere schatters van $\tau(\theta) = \theta$ (zodat $\tau'(\theta) = 1$). De CRLB wordt hiermee

$$\text{CRLB}(\theta) = \frac{1}{nE\left(\frac{\partial}{\partial\theta}\log f_X(X;\theta)\right)^2}.$$

Er geldt

$$\log f_X(X;\theta) = 3\log\theta - \log 2 - 4\log X - \frac{\theta}{X},$$

zodat

$$\frac{\partial}{\partial\theta}\log f_X(X;\theta) = \frac{3}{\theta} - \frac{1}{X}.$$

Er volgt

$$E\left(\frac{\partial}{\partial\theta}\log f_X(X;\theta)\right)^2 = E\left(\frac{3}{\theta} - \frac{1}{X}\right)^2 = \text{Var}\left(\frac{1}{X}\right) = \frac{3}{\theta^2}.$$

Hiervoor is het nodig om in te zien (af te leiden) dat $Y = 1/X \sim \text{GAM}(\theta^{-1}, 3)$. [Bewijs: transformeer $y = \frac{1}{x}$, inverse $x = \frac{1}{y}$, $J = -\frac{1}{y^2}$, $f_Y(y;\theta) = \frac{\theta^3}{2}x^2\mathrm{e}^{-\theta x}$, $x > 0$, de pdf van een $\text{GAM}(\theta^{-1}, 3)$-stochast.]
Veel eenvoudiger is het om de andere vorm te gebruiken:

$$\text{CRLB}(\theta) = \frac{1}{-nE\left(\frac{\partial^2}{\partial\theta^2}\log f_X(X;\theta)\right)}.$$

Nogmaals differentiëren van bovenstaande afgeleide geeft

$$\frac{\partial^2}{\partial\theta^2}\log f_X(X;\theta) = -\frac{3}{\theta^2}.$$

De CRLB is dus

$$\text{CRLB}(\theta) = \frac{1}{n\frac{3}{\theta^2}} = \frac{\theta^2}{3n}.$$

(b) Voor de MME moeten we eerst de verwachting van $X$ bepalen.

$$E(X) = \frac{\theta}{3}\int_0^\infty x^{-3}\mathrm{e}^{-\theta/x}\,\mathrm{d}x \overset{y=\theta/x}{=} \frac{\theta^3}{2}\int_0^\infty \frac{y}{\theta^2}\mathrm{e}^{-y}\,\mathrm{d}y = \frac{\theta}{2}.$$

De MME is de oplossing van de vergelijking

$$\bar{X} = \frac{\theta}{2},$$

die wordt gegeven door

$$\tilde{\theta} = 2\bar{X}.$$

De MME is zuiver voor $\theta$ omdat $E(\tilde{\theta}) = E(2\bar{X}) = 2E(X) = \theta$. Voor het bepalen van de MSE consistentie hebben we de variantie van de MME nodig: $\mathrm{Var}(\tilde{\theta}) = \mathrm{Var}(2\bar{X}) = 4\mathrm{Var}(\bar{X}) = \frac{4}{n}\mathrm{Var}X$. Op soortgelijke wijze als voor de verwachting van $X$ volgt

$$E(X^2) = \frac{\theta}{3}\int_0^\infty x^{-2}\mathrm{e}^{-\theta/x}\,\mathrm{d}x \overset{y=\theta/x}{=} \frac{\theta^3}{2}\int_0^\infty \frac{1}{\theta^2}\mathrm{e}^{-y}\,\mathrm{d}y = \frac{\theta^2}{2}.$$

Hiermee kunnen we de variantie van $X$ berekenen:

$$\mathrm{Var}(X) = E(X^2) - (E(X))^2 = \frac{\theta^2}{2} - \frac{\theta^2}{4} = \frac{\theta^2}{4},$$

zodat

$$\mathrm{Var}(\tilde{\theta}) = \frac{4}{n}\mathrm{Var}X = \frac{\theta^2}{n}.$$

Omdat de MME zuiver is en de variantie asymptotisch naar nul gaat is de MME MSE consistent. De variantie is groter dan de CRLB voor zuivere schatters van $\theta$. Mocht de CRLB onder 1(a) niet zijn gevonden, dan had men dit laatste ook kunnen concluderen m.b.v. Opmerking 9.3.II uit het collegedictaat.

(c) Het betreft een exponentiële familie. De C&S statistic is

$$S = \frac{1}{n}\sum_{i=1}^n X_i^{-1}.$$

Onder 1(a) was al opgemerkt dat $X_i^{-1} \sim \mathrm{GAM}(\theta^{-1}, 3)$, dus $S \sim \mathrm{GAM}(\theta^{-1}, 3n)$. Hieruit volgt $E(S) = 3\theta^{-1}n$, dus is $\frac{S}{3n}$ de gevraagde UMVUE.

(d) Beschouw de likelihood ratio voor $\theta_2 > \theta_1$

$$\frac{f_{X_1,\ldots,X_n}(x_1,\ldots,x_n;\theta_2)}{f_{X_1,\ldots,X_n}(x_1,\ldots,x_n;\theta_1)} = \left(\frac{\theta_2}{\theta_1}\right)^{3n} \mathrm{e}^{-(\theta_2-\theta_1)\sum_{i=1}^n \frac{1}{x_i}}.$$

Omdat $\theta_2 - \theta_1 > 0$ is deze functie stijgend in $t(x_1,\ldots,x_n) = -\sum_{i=1}^n \frac{1}{x_i}$. De joint pdf heeft dus de MLR eigenschap in $T = -S$. De UMP toets verwerpt als $-S$ te groot is, oftewal als $S$ te klein is. Onder het grensgeval $\theta = 1$ geldt $S \sim \mathrm{GAM}(1, 3n)$, dus $2S/1 = 2S \sim \chi^2(6n)$. De UMP toets bestaat er dus uit te verwerpen als $S \leq \frac{1}{2}\chi^2_\alpha(6n)$.

(e) We hebben eerder gezien dat $S = \frac{1}{n}\sum_{i=1}^n X_i^{-1}$ C&S is voor $\theta$, en dat $S \sim \mathrm{GAM}(\theta^{-1}, 3n)$, dus $2\theta S \sim \chi^2(6n)$. Dit geeft

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

het gevraagde BI is dus

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

2.
$$f_X(x;\theta) = \left(\frac{x}{\theta}\right)^\theta I_{[0,\theta]}(x), \qquad \theta > 0$$

(a) De joint pdf is te schrijven als

$$f_{\boldsymbol{X}}(\boldsymbol{X}) = \frac{5^n}{\prod_{i=1}^n x_i^6} \theta^{5n} I_{[\theta,\infty)}(x_{1:n}) = h(\boldsymbol{x}) g(x_{1:n};\theta)$$

met

$$g(x_{1:n};\theta) = \theta^{5n} I_{[\theta,\infty)}(x_{1:n})$$

en

$$h(\boldsymbol{x}) = \frac{5^n}{\prod_{i=1}^n x_i^6}.$$

Het volgt nu uit het factorisatiecriterium dat $S = X_{1:n}$ sufficient is voor $\theta$.

$$F_S(s) = P(X_{1:n} \leq s) = 1 - (1 - F_X(s))^n = 1 - \left(\frac{\theta}{s}\right)^{5n}, \quad s > \theta.$$

Differentiëren geeft

$$f_S(s) = 5n\theta^{5n} s^{-5n-1} I_{[\theta,\infty)}(s)$$

$$\begin{aligned} E(u(s)) &= 5n\theta^{5n} \textstyle\int_\theta^\infty s^{-5n-1} u(s) ds = 0 \\ \Leftrightarrow \textstyle\int_\theta^\infty s^{-5n-1} u(s) ds = 0 & \\ \Rightarrow -\theta^{-5n-1} u(\theta) = 0 \quad (\theta > 0). & \end{aligned}$$

Dus $E(u(S)) = 0 \Rightarrow u(\theta) = 0$ $(\theta > 0)$, waarmee is aangetoond dat de familie van verdelingen van $S$ 'complete' is.

(b)

$$E(X_{1:n}^k) = E(S^2) = 5n\theta^{5n} \int_\theta^\infty s^{-5n-1+k} = 5n\theta^{5n} \frac{1}{k-5n} \left[s^{k-5n}\right]_{s=\theta}^\infty = \frac{5n}{5n-k}\theta^k.$$

Voor $k = 2$ vinden we

$$E(S^2) = \frac{5n}{5n-2}\theta^2.$$

$\Rightarrow \frac{5n-2}{5n} S^2$ is de gevraagde UMVUE (zuiver voor $\theta^2$ en een functie van de C&S statistic).

(c) De aannemelijkheidsfunctie is

$$L(\theta) = \frac{5^n \theta^{5n}}{\prod_{i=1}^n x_i^6} I_{[\theta,\infty)}(x_{1:n})$$

Deze functie is positief en stijgend in $\theta$ voor $\theta \leq x_{1:n}$ maar wordt nul zodra $\theta > x_{1:n}$. Er is dus sprake van een randmaximum: $\hat{\theta} = x_{1:n}$. De MSE wordt gegeven door

$$\text{MSE}(\hat{\theta}) = \text{b}^2(\hat{\theta}) + \text{Var}(\hat{\theta}).$$

De bias is

$$\text{b}(\hat{\theta}) = E(\hat{\theta}) - \theta = \frac{5n}{5n-1}\theta - \theta = -\frac{1}{5n-1}\theta.$$

De variantie is

$$\begin{aligned} \text{Var}(\hat{\theta}) &= E(\hat{\theta}^2) - (E(\hat{\theta}))^2 = \tfrac{5n}{5n-2}\theta^2 - \left(\tfrac{5n}{5n-1}\right)^2 \theta^2 \\ &= \tfrac{5n}{(5n-1)^2(5n-2)}\theta^2. \end{aligned}$$

Dit geeft

$$\text{MSE}(\hat{\theta}) = \text{b}^2(\hat{\theta}) + \text{Var}(\hat{\theta}) = \ldots = \frac{2}{25n^2 - 15n + 2}\theta^2.$$

(d) Laat $Y = \frac{X_{1:n}}{\theta}$.

$$F_Y(y) = P(Y \leq y) = P\left(\frac{X_{1:n}}{\theta} \leq y\right) = P(X_{1:n} \leq \theta y) = 1 - \left(\frac{1}{y}\right)^{5n}, y > 1.$$

Differentiëren geeft

$$f_Y(y) = 5ny^{-5n-1}, \qquad y > 1$$

de pdf van $Y$ hangt niet van $\theta$ af, waarmee is aangetoond dat $Y$ een spilfunctie is. Volgens stelling 6.3.3 (Probability Integral Transform) geldt

$$F_Y(Y) = 1 - \left(\frac{1}{Y}\right)^{5n} \sim \text{UNIF}(0,1),$$

zodat geldt

$$\begin{aligned} P\left(1 - \left(\tfrac{1}{Y}\right)^{5n} \geq 1 - \gamma\right) &= \gamma \\ \Leftrightarrow P\left(1 - \left(\tfrac{\theta}{S}\right)^{5n} \geq 1 - \gamma\right) &= \gamma \\ \Leftrightarrow P\left(\left(\tfrac{\theta}{S}\right)^{5n} \leq \gamma\right) &= \gamma \\ \Leftrightarrow P\left(\left(\tfrac{\theta}{S}\right) \leq \gamma^{\frac{1}{5n}}\right) &= \gamma \\ \Leftrightarrow P\left(\theta \leq S\gamma^{\frac{1}{5n}}\right) &= \gamma \end{aligned}$$

De gevraagde betrouwbaarheidsbovengrens is dus $S\gamma^{\frac{1}{5n}}$.

(e) De GLR toets verwerpt als de likelihood ratio

$$\frac{f(\boldsymbol{x};\hat{\theta}_0)}{f(\boldsymbol{x};\hat{\theta})}$$

te klein is, waarbij $\hat{\theta}_0$ de gehinderde MLE voorstelt. Het is eenvoudig na te gaan dat

$$\hat{\theta}_0 = \begin{cases} \hat{\theta}_0 & \text{als } \hat{\theta} < 2 \\ 2 & \text{als } \hat{\theta} \geq 2. \end{cases}$$

Dit geeft

$$\frac{f(\boldsymbol{x};\hat{\theta}_0)}{f(\boldsymbol{x};\hat{\theta})} = \begin{cases} 1 & \text{als } \hat{\theta} < 2 \\ \frac{\hat{\theta}_0}{\hat{\theta}} = \frac{2}{\hat{\theta}} & \text{als } \hat{\theta} \geq 2. \end{cases}$$

De GLR toets verwerpt dus als $\hat{\theta}$ te groot is, zeg als $\hat{\theta} \geq k$. De waarde van $k$ volgt uit de eis dat de grootte van de toets gelijk moet zijn aan $\alpha$, d.w.z.

$$P(\hat{\theta} \geq k|\theta = 2) = 1 - F_{X_{1:n}}(k;\theta = 2) = \left(\frac{2}{k}\right)^{5n} = \alpha,$$

met als oplossing $k = 2\alpha^{-\frac{1}{5n}}$.