Oplossingen tentamen Kres 4, 300802

1a. De joint pdf is

$$f(x_1,\ldots,x_n;\theta) = \left(\frac{1}{2\theta}\right)^n \prod_{i=1}^n I_{(-\theta,\theta)}(x_i) = \left(\frac{1}{2\theta}\right)^n I_{(-\theta,\theta)}(x_{1:n}) I_{[x_{1:n},\theta)}(x_{n:n}).$$

De factorisatiestelling kan nu worden toegepast, met

$$g(s;\theta) = \left(\frac{1}{2\theta}\right)^n I_{(-\theta,\theta)}(x_{1:n}) I_{[x_{1:n},\theta)}(x_{n:n}) \qquad \text{en} \qquad h(x_1,\ldots,x_n) = 1.$$

Dit geeft $S = \{X_{1:n}, X_{n:n}\}$. S is niet compleet want we kunnen makkelijk een functie van S vinden die niet identiek nul is, maar wel verwachting nul heeft. Neem b.v. $U = X_{1:n} + X_{n:n}$. Er geldt (via symmetrie) $E[U] = 0$, dus $E[U(S)] = 0$ impliceert *niet* dat $U(S)$ nul is.

1b. We kunnen $f(x_1,\ldots,x_n;\theta)$ als volgt omschrijven:

$$\begin{aligned} f(x_1,\ldots,x_n;\theta) &= \left(\tfrac{1}{2\theta}\right)^n \textstyle\prod_{i=1}^n I_{(-\theta,\theta)}(x_i) \\ &= \left(\tfrac{1}{2\theta}\right)^n \textstyle\prod_{i=1}^n I_{(0,\theta)}(|x_i|) \\ &= \left(\tfrac{1}{2\theta}\right)^n \textstyle\prod_{i=1}^n I_{(0,\theta)}(\max(|x_{1:n}|, |x_{n:n}|)). \end{aligned}$$

M.b.v. het factorisatiecriterium volgt dat $\max(|X_{1:n}|, |X_{n:n}|)$ sufficient is.

Voor het aantonen van compleetheid hebben we de pdf van $\max(|X_{1:n}|, |X_{n:n}|)$ nodig. Transformeer eerst via $Y_i = |X_i|$. $F_Y(y) = P[Y \le y] = P[-y \le X \le y] = F_X(y) - F_X(-y)$. Via $F_X(x) = \int_{-\theta}^x \frac{1}{2\theta} ds = \frac{\theta + x}{2\theta}$ (voor $-\theta < x < \theta$) vind je dan: $F_Y(y) = \frac{y}{\theta}$ voor $0 < y < \theta$. De CDF van $\widehat{\theta} = \max(|X_{1:n}|, |X_{n:n}|) = Y_{n:n}$ is $P[Y_{n:n} \le y] = (F_Y(y))^n = \left(\frac{y}{\theta}\right)^n$. Als de verwachting $E[U] = E[u(Y_{n:n})]$ van een functie $u(Y_{n:n})$ nul is, impliceert dat:

$$\int_0^\theta u(y) \frac{y^{n-1}}{\theta^n} dy = 0 \qquad \text{oftewel} \qquad \int_0^\theta u(y) y^{n-1} dy = 0$$

Differentieren naar θ geeft $u(\theta)\theta^{n-1} = 0$. Daar $\theta > 0$ impliceert dat $u(\theta) = 0$.

1c. $\tau(\theta) = \theta$, $\tau'(\theta) = 1$, $(\tau'(\theta))^2 = 1$. $\log f(x;\theta) = \log \frac{1}{2\theta}$, $\frac{d}{d\theta} \log f(x;\theta) = -\frac{1}{\theta}$, zodat $E[\frac{d}{d\theta} \log f(x;\theta)]^2 = \frac{1}{\theta^2}$. Toepassen van de formule voor de CRLB geeft nu:

$$\text{CRLB} = \frac{(\tau'(\theta))^2}{nE[\frac{d}{d\theta} \log f(x;\theta)]^2} = \frac{\theta^2}{n}.$$

Je zou verwachten dat zuivere schatters geen kleinere variantie kunnen hebben dan deze ondergrens. Echter, omdat er aan een regulariteitseis niet voldaan is is dat hier niet zo. Bij de beoordeling is het bovenstaande antwoord wel goed gerekend.

Eerst de UMVUE bepalen (een functie van $\widehat{\theta}$, met verwachting θ):

$$E[\widehat{\theta}] = \int_0^\theta \frac{y^{n-1}}{\theta^n} y\, dy = \frac{n}{n+1}\theta.$$

Dus $T = \frac{n+1}{n}\widehat{\theta}$ is UMVUE voor θ.

$$E[\widehat{\theta}^2] = \int_0^\theta \frac{n}{\theta}\left(\frac{y}{\theta}\right)^{n-1} y^2\, dy = \frac{n}{\theta^n}\frac{y^{n+2}}{n+2}|_{y=0}^n = \frac{n}{n+2}\theta^2,$$

zodat

$$\text{Var}\widehat{\theta} = \left(\frac{n}{n+2} - \frac{n^2}{(n+1)^2}\right)\theta^2 = \frac{n}{(n+2)(n+1)^2}\theta^2.$$

Er volgt dat

$$\text{Var}T = \frac{(n+1)^2}{n^2}\text{Var}\widehat{\theta} = \frac{1}{n(n+2)}\theta^2.$$

Voor $n \geq 1$ is dit kleiner dan de berekende CRLB. Dat komt doordat het hier om een niet regulier geval gaat.

1d.

$$E[\frac{1}{\widehat{\theta}}] = E[\frac{1}{Y_{n:n}}] = \int_0^\theta \frac{n}{\theta^2}\left(\frac{y}{\theta}\right)^{n-2} dy = \frac{n}{\theta^n}\frac{y^{n-1}}{n-1}\;|_{y=0}^\theta = \frac{n}{n-2}\theta.$$

De UMVUE voor $\frac{1}{\theta}$ is $T' = \frac{n-1}{n}\frac{1}{Y_{n:n}}$

1e. $\widehat{\theta} = Y_{n:n}$, $F_{Y_{n:n}}(y) = \frac{y^n}{\theta^n}$ voor $0 < y < \theta$, $T = \frac{\widehat{\theta}}{\theta}$ is pivotal (θ is een schaalparameter). $F_T(t) = t^n$ $(0 < t < 1)$. We zoeken k zodanig dat

$$P[\frac{\widehat{\theta}}{\theta} \geq k] = 0.95 \qquad \text{oftewel} \qquad P[\frac{\widehat{\theta}}{\theta} < k] = F_T(k) = 0.05$$

Dit geeft $k = (0.05)^{\frac{1}{n}}$. Hieruit volgt dat $P\left[\theta < \frac{\widehat{\theta}}{(0.05)^{\frac{1}{n}}}\right] = 0.95$. Omdat $\widehat{\theta}$ altijd kleiner of gelijk θ is, is $\left(\widehat{\theta}, \frac{\widehat{\theta}}{(0.05)^{\frac{1}{n}}}\right)$ een 95% BI voor θ.

2a.

$$\log f(x_i;\theta) = \log(3) - \log(\theta) - 4\log(x_i) - \frac{1}{\theta x_i^3}$$

$$LogL(\theta) = n\log(3) - n\log(\theta) - 4\sum\log(x_i) - \frac{1}{\theta}\sum\frac{1}{x_i^3}$$

$$\begin{aligned}\log L(\theta)' &= -\tfrac{n}{\theta} + \tfrac{1}{\theta^2}\textstyle\sum\tfrac{1}{x_i^3} \\ &= \tfrac{n}{\theta^2}\left(\tfrac{1}{n}\textstyle\sum\tfrac{1}{x_i^3} - \theta\right) \\ &= \tfrac{n}{\theta^2}\left(\hat{\theta} - \theta\right) \qquad \text{met } \hat{\theta} = \tfrac{1}{n}\textstyle\sum\tfrac{1}{x_i^3}\end{aligned}$$

Uit de laatste vergelijking volgt tevens dat $\hat{\theta}$ UMVUE is. Aangezien $f(x)$ REC is, kun je ook laten zien dat $\hat{\theta}$: (i) een (lineaire) functie van de sufficient en complete statistic $\sum x_i^{-3}$ is en (ii) $E[\hat{\theta}] = \theta$.

2b.

$$y_i = \frac{1}{y_i^3} \rightarrow x_i = \frac{1}{y_i^{1/3}} \text{en } |J| = \frac{1}{3y_i^{4/3}}$$

$$f(y_i) = \frac{3}{\theta y_i^{-4/3}} \exp\left(-\frac{1}{\theta y_i^{-1}}\right) \frac{1}{3y_i^{4/3}} = \frac{1}{\theta} \exp\left(-\frac{1}{\theta} y_i\right)$$

$$\frac{1}{X_i^3} \sim Exp(\theta), \sum \frac{1}{X_i^3} \sim Gam(\theta, n), \frac{2}{\theta} \sum \frac{1}{X_i^2} = 2n\frac{\hat{\theta}}{\theta} \sim \chi^2_{(2n)}$$

$$P\left[\chi^2_{\alpha/2}(2n) \leq 2n\frac{\hat{\theta}}{\theta} \leq \chi^2_{1-\alpha/2}(2n)\right] = 1 - \alpha$$

$$\left(\frac{2n\hat{\theta}}{\chi^2_{1-\alpha/2}(2n)}, \frac{2n\hat{\theta}}{\chi^2_{\alpha/2}(2n)}\right)$$

2c.

$$\frac{f(x_1, ..., x_n, \theta_2)}{f(x_1, ..., x_n, \theta_1)} = \frac{\theta_2^{-n} \exp(-\frac{1}{\theta_2} \sum \frac{1}{x_i^3})}{\theta_1^{-n} \exp(-\frac{1}{\theta_1} \sum \frac{1}{x_i^3})} = \left(\frac{\theta_1}{\theta_2}\right)^n \exp\left((\frac{1}{\theta_1} - \frac{1}{\theta_2}) \sum \frac{1}{x_i^3}\right)$$

Als $\theta_1 < \theta_2$, dan monotoon niet-dalend in $\sum \frac{1}{x_i^3}$. Verwerp $H_0 : \theta \geq \theta_0$ versus $H_a : \theta < \theta_0$ als $\sum x_i^{-3} \leq k$.

$$P[\sum \frac{1}{X_i^3} \leq k|\theta_0] = P[\frac{2}{\theta_0} \sum \frac{1}{X_i^3} \leq \chi^2_\alpha(2n)|\theta_0] = \alpha$$

Verwerp H_0 als $\sum x_i^{-3} \leq \frac{\theta_0}{2}\chi^2_\alpha(2n)$.

2e.

$$\begin{aligned}
\lambda &= \frac{f(x_1,...,x_n,\hat{\theta}_0)}{f(x_1,...,x_n,\hat{\theta})} = \frac{\hat{\theta}_0^{-n} \exp(-\frac{1}{\theta_0} \sum \frac{1}{x_i^3})}{\hat{\theta}^{-n} \exp(-\frac{1}{\hat{\theta}} \sum \frac{1}{x_i^3})} \\
&= \left(\frac{\hat{\theta}}{\hat{\theta}_0}\right)^n \exp\left((\frac{1}{\hat{\theta}} - \frac{1}{\hat{\theta}_0}) \sum \frac{1}{x_i^3}\right) |\hat{\theta}_0 = 2, n = 10 \\
&= (\hat{\theta}/2)^{10} \exp\left((\frac{1}{\hat{\theta}} - \frac{1}{2}) \sum \frac{1}{x_i^3}\right) |\hat{\theta} = 1 \\
&= (1/2)^{10} \exp\left(\frac{1}{2} 10\right) = 0,1449
\end{aligned}$$

$$-2\log(0,1449) = 3,86 > 3.84 = \chi^2_{0.9}(1)$$

Dus er is voldoende statistisch bewijs gevonden bij $\alpha = 0,1$ om de nulhypothese $H_0 : \theta = 2$ te verwerpen.