

1.

a. De score functie heeft de vorm

$$\begin{aligned}\dot{\ell}_\theta(x) &= \frac{\partial}{\partial \theta} \log p_\theta(x) = \frac{x}{\theta} - \frac{2-x}{1-\theta} - \frac{1}{\theta} + \frac{1}{2-\theta} \\ &= \frac{-x\theta + 2x - 2}{\theta(1-\theta)(2-\theta)}.\end{aligned}$$

De aannemelijkheidsvergelijking $\sum_{i=1}^n \dot{\ell}_\theta(X_i) = 0$ is equivalent aan $-\bar{X}\theta + 2\bar{X} - 2 = 0$, en heeft de unieke oplossing $\hat{\theta} = 2 - 2/\bar{X}$. De log likelihood is proportioneel aan $\theta \mapsto \bar{X} \log \theta + (2 - \bar{X}) \log(1 - \theta) - \log \theta - \log(2 - \theta)$. Als $1 < \bar{X} < 2$, dan divergeert deze functie naar $-\infty$ als $\theta \downarrow 0$ of $\theta \uparrow 1$. Omdat de functie continu is op $(0, 1)$ moet een maximum worden aangenomen, en dit moet een nulpunt van de afgeleide zijn. Het punt $\hat{\theta}$ correspondeert derhalve met een maximum.

b. De grootheid X_1 is discreet verdeeld op $\{1, 2\}$.

$$E_\theta X_1 = 1 \binom{2}{1} \frac{\theta(1-\theta)}{\theta(2-\theta)} + 2 \binom{2}{2} \frac{\theta^2}{\theta(2-\theta)} = \frac{2}{2-\theta}.$$

c. De momentenschatter is de oplossing van de vergelijking $\bar{X} = 2/(2 - \theta)$ naar θ . Dit is precies $\hat{\theta}$ als in a).

d. De likelihood functie heeft de vorm

$$\theta \mapsto \prod_{i=1}^n p_\theta(X_i) = \prod_{i=1}^n \binom{2}{X_i} \frac{\theta^{\sum_{i=1}^n X_i} (1-\theta)^{\sum_{i=1}^n (2-X_i)}}{\theta^n (2-\theta)^n}.$$

De a-posteriori dichtheid is proportioneel aan deze functie keer $\pi(\theta)$, dat wil zeggen proportioneel aan

$$\begin{aligned}\frac{\theta^{\sum_{i=1}^n X_i} (1-\theta)^{\sum_{i=1}^n (2-X_i)}}{\theta^n (2-\theta)^n} \theta^{\alpha-1} (1-\theta)^{\beta-1} (2-\theta)^n \\ = \theta^{\sum_{i=1}^n X_i - n + \alpha - 1} (1-\theta)^{2n - \sum_{i=1}^n X_i + \beta - 1}.\end{aligned}$$

We herkennen hierin de vorm van de beta-dichtheid met parameters $\sum_{i=1}^n X_i - n + \alpha$ en $2n - \sum_{i=1}^n X_i + \beta$.

e. De Bayes schatter is de verwachting van de a-posteriori verdeling. De verwachting van de Beta verdeling met parameters α en β is $\alpha/(\alpha + \beta)$ (zoals ook gemakkelijk te berekenen is uit het gegeven). De verwachting van de a-posteriori verdeling is dus

$$\frac{\sum_{i=1}^n X_i - n + \alpha}{\sum_{i=1}^n X_i - n + \alpha + 2n - \sum_{i=1}^n X_i + \beta} = \frac{\bar{X} - 1 + \alpha/n}{\alpha/n + 1 + \beta/n}.$$

2. We berekenen dat $E_\theta X_1 = (4/3)\theta$ en $E_\theta X_1^2 = 2\theta^2$, zodat $\text{var}_\theta X_1 = (2/9)\theta^2$.a. $E_\theta \bar{X} = E_\theta X_1 = (4/3)\theta$. De schatter $\bar{X}$ is dus niet zuiver voor θ .

b.

$$\text{MSE}(\bar{X}, \theta) = \text{var}_\theta \bar{X} + (E_\theta \bar{X} - \theta)^2 = (2/9n)\theta^2 + (1/9)\theta^2.$$

$$\text{MSE}((3/4)\bar{X}, \theta) = (3/4)^2 \text{var}_\theta \bar{X} + ((3/4)E_\theta \bar{X} - \theta)^2 = (1/8n)\theta^2 + 0.$$

c. De mean square error van $(3/4)\bar{X}$ is kleiner dan de mean square error van $\bar{X}$. We geven daarom de voorkeur aan $(3/4)\bar{X}$.

3.

Laat $Z_1, \dots, Z_9$ de verschillen zijn tussen de temperatuur in De Bilt en Eelde in de 9 jaren.

- Er geldt $\bar{z} = 0.3$, $s_z^2 = 0.0475$ en $\sqrt{9}\bar{x}/s_z = 4.129483$. Dit geeft een eenzijdige overschrijdingskans van 0.001650402 bij de t -verdeling met 8 vrijheidsgraden, en een tweezijdige overschrijdingskans van ongeveer 0.003. Omdat de laatste kleiner is dan 5% verwerpen we de nulhypothese.
- Een betrouwbaarheidsinterval is $\bar{Z} \pm qt_{0.025} * S_z/\sqrt{9}$, voor qt_α het α bovenkwantiel van de t -verdeling met 8 vrijheidsgraden. Dit wordt $0.3 \pm 2.306 * 0.2179449/3 = [0.1324727, 0.4675273]$.
- We hebben voor a) en b) aangenomen dat $Z_1, \dots, Z_9$ een steekproef uit een normale verdeling is, d.w.z. $Z_1, \dots, Z_9$ onafhankelijk en normaal verdeeld met zekere verwachting en zekere variantie.

4. Laat X binomiaal verdeeld zijn met parameters n en p . We verwerpen de nulhypothese $H_0: p \leq 0.25$ voor grote waarden van X , zeg $X \geq c$. De fout van de eerste soort wordt gegeven door

$$P_{0.25}(X \geq c) \approx 1 - \Phi\left(\frac{c - n0.25}{\sqrt{n0.25(1 - 0.25)}}\right).$$

Het onderscheidend vermogen in $p = 0.3$ is gelijk aan

$$P_{0.3}(X \geq c) \approx 1 - \Phi\left(\frac{c - n0.3}{\sqrt{n0.3(1 - 0.3)}}\right).$$

Deze kansen zijn kleiner gelijk 0.05 en, respectievelijk, groter gelijk 0.95 als

$$\frac{c - n0.25}{\sqrt{n0.25(1 - 0.25)}} \approx 1.65$$

en

$$\frac{c - n0.3}{\sqrt{n0.3(1 - 0.3)}} \approx -1.65.$$

Na eliminatie van c vinden we een kwadratische vergelijking in $\sqrt{n}$. Dit levert op $\sqrt{n} \approx 29.4119$ en $n \approx 865$.

5.

- De aannemelijkheidsfunctie is gelijk aan

$$L(\theta) = \prod_{i=1}^n p_\theta(X_i) = 2^n \prod_{i=1}^n X_i \frac{1}{\theta^{2n}} 1_{X_{(n)} \leq \theta}.$$

De schets is te moeilijk op de computer. Deze laat echter een plaatje zien met θ op de horizontale as en een functie die 0 is voor $\theta < X_{(n)} = U$ en positief en dalend is voor $\theta \geq U$. Dus de aannemelijke schatter is U .

- Voor $0 < u < \theta$,

$$P_\theta(U \leq u) = P_\theta(X_1 \leq u)^n = \left(\int_0^u \frac{2x}{\theta^2} dx\right)^2 = \frac{u^{2n}}{\theta^{2n}}.$$

Een kansdichtheid voor U is de $p_\theta^U(u) = 2nu^{2n-1}/\theta^{2n} 1_{0 < u < \theta}$.

- Een pivot is een functie $T(X, \theta)$ van de waarneming en de parameter zodanig dat de kansverdeling van $T(X, \theta)$ als θ de ware parameter is niet van θ afhangt (d.w.z. hetzelfde is voor iedere θ).
- Uit b) volgt dat $P_\theta(U/\theta \leq u) = u^{2n}$ voor $0 \leq u \leq 1$. De rechterkant is onafhankelijk van θ .
- Uit d) volgt dat $P_\theta(c \leq U/\theta \leq d) = d^{2n} - c^{2n}$ voor iedere $0 < c < d < 1$. Voor $d = (1 - \frac{1}{2}\alpha)^{1/2n}$ en $c = (\frac{1}{2}\alpha)^{1/2n}$ is de rechterkant gelijk aan $1 - \frac{1}{2}\alpha - \frac{1}{2}\alpha = 1 - \alpha$. Voor deze waarden van c en d is $\{\theta: c \leq U/\theta \leq d\} = [U/d, U/c]$ een $1 - \alpha$ -betrouwbaarheidsinterval.

6.

a. De likelihood functie is

$$\theta \mapsto L(\theta) = \prod_{i=1}^n p_{\theta}(X_i) = \theta^n \left(\prod_{i=1}^n \frac{2 \log X_i}{X_i} \right) e^{-\theta \sum_{i=1}^n (\log X_i)^2}.$$

De likelihood ratio statistiek voor de nulhypothese $H_0: \theta = \theta_0$ is $L(\hat{\theta})/L(\theta_0)$ voor $\hat{\theta}$ de meest aannemelijke schatter. Door differentiatie van de log likelihood functie vinden we dat $\hat{\theta} = 1/(\log X)^2$. Invullen levert voor de likelihood ratio statistiek dat

$$\left(\frac{\hat{\theta}}{\theta_0} \right)^n \frac{e^{-\sum_{i=1}^n (\log X_i)^2 \hat{\theta}}}{e^{-\sum_{i=1}^n (\log X_i)^2 \theta_0}} = \frac{1}{(\theta_0 (\log X)^2)^n} e^{-n + \sum_{i=1}^n (\log X_i)^2 \theta_0}.$$

b. Met $\Lambda(\theta_0) = \log(L(\hat{\theta})/L(\theta_0))$ en L de likelihood als in a) is het benaderend betrouwbaarheidsinterval gelijk aan $\{\theta_0: 2\Lambda(\theta_0) \leq \chi_{1,\alpha}^2\}$.

c. We berekenen

$$\begin{aligned} \dot{\ell}_{\theta}(x) &= \frac{\partial}{\partial \theta} \log p_{\theta}(x) = \frac{1}{\theta} - (\log x)^2, \\ \ddot{\ell}_{\theta}(x) &= \frac{\partial^2}{\partial \theta^2} \log p_{\theta}(x) = \frac{-1}{\theta^2}. \end{aligned}$$

De waargenomen informatie is daarom $-\frac{1}{n} \sum_{i=1}^n \ddot{\ell}_{\hat{\theta}}(X_i) = 1/\hat{\theta}^2 = ((\log X)^2)^2$.

d. Het interval is $\hat{\theta} \pm z_{\alpha/2}/\sqrt{n \hat{i}_{\theta}}$ met z_{α} het bovenkwantiel van de normale verdeling en $\hat{i}_{\theta}$ een schatter voor de Fisher informatie. Gebruikmakend van de waargenomen informatie als schatter voor de Fisher informatie vinden we $\hat{\theta} \pm \hat{\theta} z_{\alpha/2}/\sqrt{n}$.