Tentamen Inleiding Statistiek 2018-2019

Toegestane hulpmiddelen: een eenvoudige rekenmachine. Geen boeken, aantekeningen, grafische rekenmachines, telefoons, smart watches of andere hulpmiddelen.

Licht al je antwoorden toe.

1. De stochastische grootheden $X_1, X_2, \ldots, X_n$ zijn onderling onafhankelijk en continu verdeeld volgens de kansdichtheid
$$p_\theta(x) = c\theta^c x^{-(c+1)}\mathbf{1}\{x \geq \theta\}.$$
Hierin is $\theta > 0$ een onbekende parameter, en $c > 2$ een bekende constante.
   (a) Bepaal een momentenmethodeschatter van θ.
   (b) Bepaal de meest aannemelijke schatter (MLE) voor θ.
   (c) Bepaal de verwachte kwadratische fout (MSE) van de momentenmethodeschatter uit vraag 1(a).
   (d) Bepaal een ééndimensionale voldoende statistiek voor θ.
   (e) Is de momentenschatter UMVZ voor θ?

2. Zij X een stochast met dichtheid
$$p_\theta(x) = \theta x^{\theta-1}\mathbf{1}\{0 < x < 1\},$$
met $\theta \geq 2$ onbekend. We beschouwen de hypothesen $H_0 : \theta = 2$ en $H_1 : \theta = 3$.
   (a) Geef het kritieke gebied van de uniform meest onderscheidende toets met onbetrouwbaarheid α voor deze hypothesen.
   (b) Wat is het onderscheidend vermogen (power) van de toets uit de vorige deelvraag, uitgedrukt in α?
   (c) Stel dat we $x = 0.4$ waarnemen. Wat is de p-waarde behorende bij deze waarneming?
   (d) Wat is de likelihoodratiostatistiek voor de hypothesen $H_0 : \theta = 2$ en $H_1 : \theta > 2$?

3. Zij $X_1, \ldots, X_n$ i.i.d. inverse-Gamma verdeeld, met dichtheid
$$p_{\alpha,\beta}(x) = \frac{\beta^\alpha}{\Gamma(\alpha)} x^{-(\alpha+1)} e^{-\frac{\beta}{x}}, \quad x > 0,$$
waarbij $\alpha > 0$ bekend is, en $\beta > 0$ een onbekende parameter. Voor β selecteren we de Gamma(γ, δ)-verdeling als a-prioriverdeling. De dichtheid van de a-prioriverdeling is dus $\pi(\beta) = \frac{\delta^\gamma}{\Gamma(\gamma)}\beta^{\gamma-1}e^{-\delta\beta}$, voor $\beta > 0$.
   (a) Toon aan dat de a-posterioriverdeling een Gamma(γ', δ')-verdeling is, met $\gamma' = \gamma + n\alpha$ en $\delta' = \delta + \sum_{i=1}^n \frac{1}{x_i}$.
   (b) We kiezen $\alpha = \gamma = \delta = 1$, en doen twee waarnemingen, $x_1 = 0.5$ en $x_2 = 1$. Bepaal een 95%-Bayesiaans overdekkingsgebied voor β. Je kunt daarbij gebruikmaken van Tabel 1 **(op het tweede blad)** met kwantielen van de Gamma(γ, δ)-verdeling.

4. Zij $Y_i = \theta + \frac{\theta}{2}x_i^2 + \varepsilon_i$, $i = 1, \ldots, n$, waarbij de ε_i i.i.d. variabelen met verwachtingswaarde 0 en variantie 1 zijn, en de x_i bekende constanten. Hierbij is $\theta \in \mathbb{R}$ een onbekende parameter.
   (a) Bepaal de kleinste kwadratenschatter van θ.

   We nemen nu aan dat de ε_i i.i.d $\mathcal{N}(0, 1)$ verdeeld zijn. Het kan in het vervolg van pas komen om te weten dat de dichtheid van een $\mathcal{N}(\mu, \sigma^2)$ verdeelde variabele X gegeven wordt door $p_{\mu,\sigma^2}(x) = \frac{1}{\sqrt{2\pi\sigma^2}}e^{-\frac{1}{2\sigma^2}(x-\mu)^2}$.
   (b) Bereken de Fisher-informatie in de gehele waarneming $(Y_1, \ldots, Y_n)$.
   (c) Bepaal de Cramér-Rao ondergrens voor de variantie van een zuivere schatter van θ, en laat zien dat de variantie van de kleinste kwadratenschatter gelijk is aan de Cramér-Rao ondergrens.
   (d) Voor welke keuze van de x_i, onder de beperking $0 \leq x_i \leq 1$ voor alle i, is de variantie van de kleinste kwadratenschatter minimaal?

Tabel 1: Waarden a zodanig dat $\mathbb{P}(X \leq a) = b$, voor $X \sim Gamma(\gamma, \delta)$.

b	$\gamma = 1, \delta = 4$	$\gamma = 2, \delta = 4$	$\gamma = 3, \delta = 4$	$\gamma = 4, \delta = 4$	$\gamma = 4, \delta = 3$	$\gamma = 4, \delta = 2$	$\gamma = 4, \delta = 1$
0.025	0.006	0.061	0.155	0.272	0.363	0.545	1.090
0.05	0.013	0.089	0.204	0.342	0.455	0.683	1.366
0.95	0.749	1.186	1.574	1.938	2.585	3.877	7.754
0.975	0.922	1.393	1.806	2.192	2.922	4.384	8.767