Hertentamen inleiding statistiek

25 maart 2009, 10.00–13.00

Het boek van Rice, en uw eigen aantekeningen, mag bij de tentamen worden gebruikt.

1. Stel dat $X_1, \ldots, X_n$ zijn n onafhankelijk en identiek verdeelde waarnemingen uit de uniforme verdeling op het interval $[0, \theta]$, waar $\theta > 0$ een onbekende parameter is.

a) Wat is de momenten schatter van θ en wat is zijn m.s.e. (*mean square error* oftewel verwachte kwadratische fout)?

b) Laat zien dat de aannemelijkheidheidsfunctie $\mathrm{lik}(\theta)$ geschreven kan worden als

$$\mathrm{lik}(\theta) = \begin{cases} \dfrac{1}{\theta^n} & \theta \geq \max X_i \\ 0 & \theta < \max X_i. \end{cases}$$

Laat hiermee zien dat $\widehat{\theta}_{\mathrm{ML}} = \max_i X_i$ de meest aannemelijke schatter van θ is.

c) Door $\Pr(\max X_i \leq x)$ te bekijken voor $0 \leq x \leq \theta$, laat zien dat $\max X_i$ kansdichtheid nx^{n-1}/θ^n heeft, $0 \leq x \leq \theta$, 0 elders.

d) Bereken de m.s.e. van $\widehat{\theta}_{\mathrm{ML}}$ en vergelijk (voor n groot) met die van de momenten schatter.

2. Stel dat X_i onafhankelijk binomiaal(n_i, p_i) verdeeld is, $i = 1, \ldots, m$, waar $p_1, \ldots, p_m$ onbekende parameters zijn; de n_i zijn bekende getallen.

a) Wat is de meest aannemelijke schatter van $(p_1, \ldots, p_m)$

(i) zonder enige beperking op de p_i,

(ii) onder de hypothese dat $p_1 = \ldots = p_m$ (de p_i zijn gelijk aan elkaar, maar hun waarde p is niet bekend)?

b) Wat is de gegeneraliseerde aanemelijkheidsquotient toets van de nul hypothese $p_1 = \ldots = p_m$ tegen de alternatief dat ze niet allemaal gelijk zijn? Ervan uitgaande dat de gebruikelijke grote steekproeftheorie van toepassing is, hoe zou u in de praktijk de toets uitvoeren?

3. De Weibull verdeling (veel gebruikt bij de analyse van levensduren van industriele producten) heeft verdelingsfunctie

$$\Pr(X \leq x) = \begin{cases} 1 - \exp(-\alpha x^{\beta}), & x \geq 0 \\ 0 & x \leq 0 \end{cases}$$

voor zekere parameters $\alpha > 0$ en $\beta > 0$.

a) Met X Weibull verdeeld als boven, definieer $U = \log \alpha + \beta \log X$. Laat zien dat de verdelingsfunctie van U is dan

$$\Pr(U \leq u) = 1 - e^{-e^{u}}, \quad -\infty < u < \infty.$$

b) Laat u_p de p-e kwantiel van de verdeling van U zijn, $0 < p < 1$. Wat is u_p? Laat zien dat de p-e kwantiel van de verdeling van $\log X$ gelijk is aan $(u_p - \log \alpha)/\beta$.

c) Gegeven een steekproef $X_1, \ldots, X_n$ uit een Weibull verdeling, hoe ziet een QQ plot van de geordende waardes van $\log X_i$ versus de bijbehorende kwantielen u_p uit?

4. We gaan verder met de steekproef uit Weibull verdeling die in opgave 3c) is beschreven.

a) Onderzoek de meest aannemelijke schatters van α en β, en beschrijf hoe U benaderende 95% betrouwbaarheidsinterval voor α en β zou berekenen.

b) Stel dat het bekend is dat $\beta = \beta_0$. We willen de nul-hypothese toetsen dat $\alpha = 1$ tegen het alternatief $\alpha > 1$. Leid een uniform meest onderscheidend toets af van niveau α.