Tilastollinen päättely I
4. harjoitukset, 6. vko 2004

4.1. Olkoon $Y_1, \ldots, Y_n$ otos Poissonin jakaumasta $\mathrm{Poi}(\lambda)$. Lopulta pystytään havainnoimaan vain, onko $Y_i = 0$ vai $Y_i > 0$. Määritellään satunnaismuuttujat X_i siten, että

$$X_i = \begin{cases} 0, & \text{kun } Y_i = 0 \\ 1, & \text{kun } Y_i > 0. \end{cases}$$

(a) Totea, että $P(X_i = 0) = e^{-\lambda}$ ja $P(X_i = 1) = 1 - e^{-\lambda}$.

(b) Määritä otokseen $x_1, \ldots, x_n$ perustuva uskottavuusfunktio.

(c) Määritä parametrin λ SUE $\hat{\lambda}$. Mitä on $\hat{\lambda}$, jos $x_1 = \cdots = x_n = 1$?

4.2. Oletetaan yhden selittävän muuttujan regressiomalli, missä selitettävät muuttujat $Y_1, \ldots, Y_n$ ovat toisistaan riippumattomat ja $Y_i \sim \mathrm{N}(\alpha + \beta x_i, 1)$, missä selittävän muuttujan arvot x_i, $0 \leq i \leq n$ ovat annettuja vakioita. Määritä parametrien α ja β suurimman uskottavuuden estimaatit.

4.3. Oletetaan tehtävän 2 tilanteessa, että $\alpha = 0$ eli $Y_i \sim \mathrm{N}(\beta x_i, 1)$.

(a) Määritä pistefunktio $u(\beta)$ ja informaatiofunktio $\mathcal{I}(\beta)$.

(b) Laske SUE $\hat{\beta}$ ja $\mathcal{I}(\hat{\beta})$.

(c) Laske $\mathrm{Var}(\hat{\beta})$ ja vertaa varianssia informaatiofunktion arvoon.

(Ks. 7.11.1 s.254, JU, E1.2, s.3 ja s.16)

4.4. Perheiden tuloja x mitataan sellaisella asteikolla, että arvo $x = 1$ on minimipalkka. Oletetaan, että tulojakauman tiheysfunktio on

$$f(x) = \frac{\theta}{x^{\theta+1}}, \qquad \text{kun} \quad x \geq 1$$

ja $\theta > 0$. Satunnaisesti valittujen 10 perheen tulot olivat:

1.02, 1.41, 1.75, 2.31, 3.42, 4.31, 9.21, 17.4, 38.6, 392.8

(a) Määritä θ:n suurimman uskottavuuden estimaatti ja (b) 10%:n uskottavuusväli.

4.5. Kaupungin taksit on numeroitu juoksevasti $1, 2, \ldots, N$. Olet kahden viikon ajan jokaisena päivänä merkinnyt muistiin ensimmäisenä havaitsemasi taksin numeron ja saanut aineiston:

$$137, 24, 86, 33, 92, 129, 17, 11, 55, 61, 8, 29, 103, 71$$

Oletetaan, että havainnot ovat toisistaan riippumattomat ja jokaisen taksin todennäköisyys tulla havaituksi on $1/N$. Perustele, miksi N:n SUE $\hat{N} = 137$.

4.6. Oletetaan, että tuotteen kestoaika $X \sim \mathrm{Exp}(\theta)$. Tehtiin 10 kestokoetta ja saatiin seuraavat kestoajat: $70, 11, 66, 5, 20, 4, 35, 40, 29, 8$.

(a) Määritä pistefunktio $u(\theta)$ ja informaatiofunktio $\mathcal{I}(\theta)$.

(b) Laske SUE $\hat{\theta} = \bar{x}, u(\hat{\theta})$ ja $\mathcal{I}(\hat{\theta})$.

(c) Laske yhden havainnon $x_1 = 70$ informaatio.

4.7. Kun tietyn metallin pitoisuus liuoksessa mitataan, noudattaa mittauksessa syntyvä virhe normaalijakaumaa $\mathrm{N}(0, \sigma_0^2)$. Jos oikea pitoisuus on μ, niin havaittu pitoisuus $X \sim \mathrm{N}(\mu, \sigma_0^2)$. Varianssi σ_0^2 tunnetaan aikaisemmasta kokemuksesta.

(a) Olkoot $x_1, x_2, \ldots, x_n$ pitoisuuden μ toisistaan riippumattomat mittaukset. Mikä on SUE $\hat{\mu}$?

(b) Laimennetaan sitten alkuperäisen liuoksen pitoisuus puoleen niin, että pitoisuus on $\mu/2$. Tehdään lisämittaukset $y_1, y_2, \ldots, y_m$. Määritä m:n lisämittauksen perusteella $\hat{\mu}$.

(c) Määritä kaikkien $n + m$:n mittauksen perusteella $\hat{\mu}$.

4.8. Olkoot 1. kertaluvun autoregressiivisen aikasarjan oletukset seuraavat:

$$Y_t = \beta Y_{t-1} + E_t \quad \text{ja} \quad E_t \sim \mathrm{N}(0, \sigma^2), \quad 0 \leq t \leq n.$$

Sysäykset $E_1, \ldots, E_n$ ovat keskenään riippumattomat. Oletetaan lisäksi, että havainto $Y_0 = 0$.

(a) Osoita, että $E(Y_t) = 0, \quad 0 \leq t \leq n$.

(b) Laske $\mathrm{Var}(Y_t)$.

(c) Esitä logaritmoidun uskottavuusfunktion lauseke $l(\beta, \sigma^2)$.

(Ks. JU, E1.3, s.4 ja s.12)