Johdatus tilastolliseen päättelyyn, 2. harjoitus (Viikot 13–14: 26.–27.3. ja 4.–6.4.2013)

1. Käsittelemme uudestaan ensimmäisten harjoitusten ensimmäisen tehtävän tilastollista mallia. Parametrilla θ on kaksi vaihtoehtoista arvoa: 1 tai 2. Diskreetin satunnaismuuttujan Y pistetodennäköisyysfunktion $g(y;\theta)$ arvot on taulukoitu alla.

y	1	2	3	4	5
$g(y;1)$	0	0.3	0.4	0.2	0.1
$g(y;2)$	0.1	0.4	0.3	0.1	0.1

Luennoitsija valitsi näistä pistetodennäköisyysfunktioista yhden ja simuloi siitä tietokoneella kolme arvoa, jotka ovat

$$y_1 = 3, \quad y_2 = 2, \quad y_3 = 4.$$

(Käsittelemme niitä satunnaisotoksen Y_1, Y_2, Y_3 havaittuina arvoina.)

Muodosta uskottavuusfunktio, ja laske SU-estimaatti $\hat{\theta}$.

2. Olkoon $\mathbf{Y} = (Y_1, Y_2, Y_3)$ satunnaisotos normaalijakaumasta $N(\mu, 1)$, jossa normaalijakauman odotusarvo $\mu \in \mathbb{R}$ on tuntematon, ja varianssiparametrilla on arvona yksi. Tutkimme seuraavia parametrin μ estimaattoreita:

$$\hat{\mu}_1(\mathbf{Y}) = \frac{1}{2}(Y_1 + Y_2), \qquad \hat{\mu}_2(\mathbf{Y}) = \frac{1}{3}(Y_1 + Y_2 + Y_3)$$

$$\hat{\mu}_3(\mathbf{Y}) = \frac{1}{5}(Y_1 + 2\,Y_2 + Y_3) \qquad \hat{\mu}_4(\mathbf{Y}) = \frac{1}{4}(Y_1 + 2\,Y_2 + Y_3)$$

Laske kunkin estimaattorin $\hat{\mu}_i(\mathbf{Y})$ harha. Mitkä niistä ovat harhattomia?

3. Laske edellisen tehtävän tilanteessa kunkin estimaattorin $\hat{\mu}_i(\mathbf{Y})$ varianssi sekä keskineliövirhe (hajotelman (3.11) avulla). Mikä estimaattoreista on keskineliövirheen mielessä tarkin, jos $\mu = 0$? Mikä estimaattoreista on tarkin, jos tiedetään että $|\mu| \geq 2$?

4. Tällä kurssilla törmäämme monessa yhteydessä positiivisella reaaliakselilla määriteltyihin positiivisiin funktioihin $L(t)$, jotka ovat joko muotoa $t^a \exp(-bt)$ tai muotoa $t^{-a} \exp(-b/t)$, jossa a ja b ovat positiivisia vakioita. Näiden funktioiden maksimipisteet löytyvät kätevästi tarkastelemalla logaritmia $\ell(t) = \log L(t)$.

a) Olkoot $a > 0$ ja $b > 0$. Näytä, että funktiolla

$$\ell(t) = a \ln(t) - bt, \qquad t > 0$$

on yksikäsitteinen maksimipiste $t_0 = a/b$.

b) Olkoot $a > 0$ ja $b > 0$. Näytä, että funktiolla

$$\ell(t) = -a \ln(t) - \frac{b}{t}, \qquad t > 0$$

on yksikäsitteinen maksimipiste $t_0 = b/a$.

5. Eksponenttijakauma Exp(λ) on jatkuva jakauma, jonka tiheysfunktio on

$$g(y;\lambda) = \lambda\,\mathrm{e}^{-\lambda\,y}, \quad \text{kun } y > 0.$$

(Jakauman tiheysfunktio on nolla negatiivisella reaaliakselilla.) Parametriavaruus on $\lambda > 0$. Jos $Y \sim \mathrm{Exp}(\lambda)$, niin $EY = 1/\lambda$.

Luennoitsija simuloi tietystä eksponenttijakaumasta riippumattomasti kolme arvoa, jotka olivat

$$y_1 = 0.40, \quad y_2 = 0.16, \quad y_3 = 1.12.$$

a) Muodosta uskottavuusfunktio, ja laske SU-estimaatti $\hat{\lambda}$. (Voit käyttää hyväksesi edellisessä tehtävässä kerrottuja tietoja.)

b) Usein eksponenttijakauma parametroidaan sen odotusarvolla $\theta = 1/\lambda$. Koska $\lambda = 1/\theta$, niin odotusarvoparametrin θ avulla ilmaistuna eksponenttijakauman tiheysfunktio on

$$h(y;\theta) = g(y;\frac{1}{\theta}) = \frac{1}{\theta}\,\mathrm{e}^{-y/\theta}, \quad \text{kun } y > 0.$$

Muodosta uskottavuusfunktio odotusarvoparametrille θ sekä laske SU-estimaatti $\hat{\theta}$. (Voit käyttää hyväksesi edellisessä tehtävässä kerrottuja tietoja.)

6. Olkoot $Y_1, \ldots, Y_n$ riippumattomia ja samoin jakautuneita satunnaismuuttujia siten, että

$$\operatorname{var} Y_i = \sigma^2, \qquad \text{kaikilla } i.$$

Laske satunnaismuuttujan

$$Z = \sum_{i=1}^{n} (Y_i - \bar{Y})^2$$

odotusarvo. Huomaa, että Z voidaan (ensimmäisten harjoitusten perusteella) ilmaista myös muodossa

$$Z = \sum_{i=1}^{n} (Y_i - \mu)^2 - n(\bar{Y} - \mu)^2,$$

jossa satunnaismuuttuja $\bar{Y}$ on keskiarvo $\bar{Y} = \frac{1}{n}\sum_{i=1}^{n} Y_i$.