Tilastollinen päättely, syksy 2013 - kevät 2014

Harjoitus 3

Esimerkkiratkaisut

1. Tarkastellaan gammajakaumamallia $Y_1, \ldots, Y_n \sim G(\alpha, 1/\beta) \perp\!\!\!\perp$, jossa $\alpha, \beta > 0$.

   a) Parametrina on (α, β). Totea, että uskottavuusyhtälöt voidaan saattaa muotoon
   $$\begin{cases} \log\alpha - \psi(\alpha) = \log\overline{y} - n^{-1}\sum_{i=1}^n \log y_i, \\ \beta = \overline{y}/\alpha, \end{cases}$$
   jossa $\psi(\alpha) = \Gamma'(\alpha)/\Gamma(\alpha)$, ja järkeile, ettei niitä voi ratkaista suljetussa muodossa.

   b) Parametrina on vain β, ja α on tunnettu luku. Laske havaittu informaatio $j(\beta; \boldsymbol{y})$ ja Fisherin informaatio $i(\beta)$. Mikä on β:n su-estimaatti?

Ratk. a) Gammajakauman $G(\alpha, 1/\beta)$ tiheysfunktio on (ks. luentomonisteen liite)
$$f(y; \alpha, \beta) = \frac{(1/\beta)^\alpha}{\Gamma(\alpha)} y^{\alpha-1} e^{-y/\beta} = \frac{1}{\beta^\alpha \Gamma(\alpha)} y^{\alpha-1} e^{-y/\beta}.$$
Koska satunnaismuuttujat $Y_1, \ldots, Y_n$ ovat riippumattomia, niin yhteistiheysfunktioksi saadaan
$$\begin{aligned} f_{\boldsymbol{Y}}(\boldsymbol{y}; \alpha, \beta) &= \prod_{i=1}^n \frac{1}{\beta^\alpha \Gamma(\alpha)} y_i^{\alpha-1} e^{-y_i/\beta} = \frac{1}{\beta^{n\alpha}\Gamma^n(\alpha)} \left(\prod_{i=1}^n y_i\right)^{\alpha-1} e^{-\sum_{i=1}^n y_i/\beta} \\ &= \left(\prod_{i=1}^n y_i\right)^{-1} \frac{1}{\beta^{n\alpha}\Gamma^n(\alpha)} \left(\prod_{i=1}^n y_i\right)^{\alpha} e^{-\sum_{i=1}^n y_i/\beta} \end{aligned}$$
Asettamalla $c(\boldsymbol{y}) = \prod_{i=1}^n y_i$, uskottavuusfunktioksi saadaan
$$L(\alpha, \beta; \boldsymbol{y}) = c(\boldsymbol{y}) f_{\boldsymbol{Y}}(\boldsymbol{y}; \alpha, \beta) = \frac{1}{\beta^{n\alpha}\Gamma^n(\alpha)} \left(\prod_{i=1}^n y_i\right)^{\alpha} e^{-\sum_{i=1}^n y_i/\beta}$$
ja log-uskottavuusfunktioksi tulee siis
$$\begin{aligned} l(\alpha, \beta; \boldsymbol{y}) &= \log L(\alpha, \beta; \boldsymbol{y}) = -n\alpha\log(\beta) - n\log(\Gamma(\alpha)) + \alpha\sum_{i=1}^n \log(y_i) - \sum_{i=1}^n y_i/\beta \\ &= -n\alpha\log(\beta) - n\log(\Gamma(\alpha)) + \alpha\sum_{i=1}^n \log(y_i) - n\bar{y}\beta^{-1}. \end{aligned}$$

Lasketaan log-uskottavuusfunktion derivaatat parametrien α ja β suhteen:

$$\begin{aligned}
\frac{\partial}{\partial \alpha} l(\alpha, \beta; \boldsymbol{y}) &= -n\log(\beta) - n\frac{\Gamma'(\alpha)}{\Gamma(\alpha)} + \sum_{i=1}^{n} \log(y_i) \\
&= -n\log(\beta) - n\psi(\alpha) + \sum_{i=1}^{n} \log(y_i) \\
\frac{\partial}{\partial \beta} l(\alpha, \beta; \boldsymbol{y}) &= -\frac{n\alpha}{\beta} + n\bar{y}\beta^{-2}.
\end{aligned}$$

Asettamalla derivaatat nolliksi saadaan uskottavuusyhtälöiksi

$$\begin{aligned}
&\begin{cases} -n\log(\beta) - n\psi(\alpha) + \sum_{i=1}^{n} \log(y_i) = 0 & \| \cdot (1/n) \\ -\frac{n\alpha}{\beta} + n\bar{y}\beta^{-2} = 0 & \| \cdot (\beta^2/n) \end{cases} \\
\Leftrightarrow &\begin{cases} -\log(\beta) - \psi(\alpha) + \frac{1}{n}\sum_{i=1}^{n} \log(y_i) = 0 & \\ -\alpha\beta + \bar{y} = 0 & \| \text{ Ratk. param. } \beta \text{ suhteen} \end{cases} \\
\Leftrightarrow &\begin{cases} -\log(\beta) - \psi(\alpha) + \frac{1}{n}\sum_{i=1}^{n} \log(y_i) = 0 & \\ \beta = \bar{y}/\alpha & \| \text{ Sij. ed. yhtälöön} \end{cases} \\
\Leftrightarrow &\begin{cases} -\log(\bar{y}/\alpha) - \psi(\alpha) + \frac{1}{n}\sum_{i=1}^{n} \log(y_i) = 0 \\ \beta = \bar{y}/\alpha \end{cases} \\
\Leftrightarrow &\begin{cases} \log\alpha - \psi(\alpha) = \log\overline{y} - n^{-1}\sum_{i=1}^{n} \log y_i \\ \beta = \bar{y}/\alpha \end{cases}
\end{aligned}$$

Huomataan, että eo. yhtälöparin ensimmäistä yhtälöä ei selvästikään pysty ratkaisemaan parametrin α suhteen suljetussa muodossa eli yhtälön ratkaisua ei ole mahdollista kirjoittaa alkeisfunktioiden avulla.

b) Kun α on tunnettu vakio, niin voidaan valita $c(\boldsymbol{y}) = \left(\prod_{i=1}^{n} y_i\right) \Gamma^n(\alpha) \left(\prod_{i=1}^{n} y_i\right)^{-\alpha}$. Tällöin uskottavuusfunktioksi saadaan

$$L(\beta; \boldsymbol{y}) = c(\boldsymbol{y}) f_{\boldsymbol{Y}}(\boldsymbol{y}; \beta) = \frac{1}{\beta^{n\alpha}} e^{-\sum_{i=1}^{n} y_i/\beta}$$

ja log-uskottavuusfunktioksi tulee

$$l(\beta; \boldsymbol{y}) = -n\alpha\log(\beta) - \sum_{i=1}^{n} y_i/\beta = -n\alpha\log(\beta) - n\bar{y}\beta^{-1}.$$

Lasketaan log-uskottavuusfunktion ensimmäinen ja toinen derivaatta:

$$\begin{aligned} l'(\beta;\boldsymbol{y}) &= \frac{d}{d\beta} l(\beta;\boldsymbol{y}) = -\frac{n\alpha}{\beta} + n\bar{y}\beta^{-2}. \\ l''(\beta;\boldsymbol{y}) &= \frac{d^2}{d\beta^2} l(\beta;\boldsymbol{y}) = \frac{n\alpha}{\beta^2} - 2n\bar{y}\beta^{-3}. \end{aligned}$$

Havaittu informaatio on nyt

$$j(\beta;\boldsymbol{y}) = -l''(\beta;\boldsymbol{y}) = -\frac{n\alpha}{\beta^2} + 2n\bar{y}\beta^{-3},$$

joten Fisherin informaatioksi saadaan

$$\begin{aligned} i(\beta) &= E\left(j(\beta;\boldsymbol{Y})\right) = -\frac{n\alpha}{\beta^2} + 2nE(\bar{Y})\beta^{-3} = -\frac{n\alpha}{\beta^2} + 2nE(Y_i)\beta^{-3} \\ &= -\frac{n\alpha}{\beta^2} + 2n\left(\frac{\alpha}{1/\beta}\right)\beta^{-3} = -\frac{n\alpha}{\beta^2} + \frac{2n\alpha}{\beta^2} = \frac{n\alpha}{\beta^2}. \end{aligned}$$

Etsitään suurimman uskottavuuden estimaattia log-uskottavuusfunktion derivaatan nollakohdasta:

$$-\frac{n\alpha}{\beta} + n\bar{y}\beta^{-2} \doteq 0 \Leftrightarrow \beta = \bar{y}/\alpha.$$

Koska log-uskottavuusfunktion derivaatta kohdassa $\bar{y}/\alpha$,

$$l''(\bar{y}/\alpha;\boldsymbol{y}) = \frac{n\alpha}{(\bar{y}/\alpha)^2} - 2n\bar{y}(\bar{y}/\alpha)^{-3} = \frac{n\alpha^3}{\bar{y}^2} - \frac{2n\alpha^3}{\bar{y}^2} = -\frac{n\alpha^3}{\bar{y}^2},$$

on negatiivinen ($\bar{y}^2 > 0$ ja $\alpha^3 > 0$), niin $\bar{y}/\alpha$ on uskottavuusfunktion lokaali maksimikohta. Koska parametriavaruus on avoin väli ($\beta \in \Omega = (0,\infty)$) ja $\bar{y}/\alpha$ on ainoa nollakohta, niin parametrin β suurimman uskottavuuden estimaatti on

$$\hat{\beta} = \frac{\bar{y}}{\alpha}.$$

2. Olkoot $Y_1, \ldots, Y_n \sim Exp(\lambda) \perp\!\!\!\perp$. Laske havaittu informaatio $j(\hat{\lambda}; \boldsymbol{y})$, Fisherin informaatio $i(\lambda)$ ja odotusarvo $E[l'(\lambda; \boldsymbol{Y})^2]$. [Monisteen harjoitustehtävä 2.11]

Ratk. Tilastollinen malli on

$$f_{\boldsymbol{Y}}(\boldsymbol{y}; \lambda) = \prod_{i=1}^{n} \lambda e^{-\lambda y_i} = \lambda^n e^{-\lambda \sum_{i=1}^{n} y_i} = \lambda^n e^{-\lambda n \bar{y}},$$

joten uskottavuusfunktioksi saadaan

$$L(\lambda; \boldsymbol{y}) = f_{\boldsymbol{Y}}(\boldsymbol{y}; \lambda) = \lambda^n e^{-\lambda n \bar{y}}, \quad \lambda > 0.$$

Log-uskottavuusfunktio on nyt

$$l(\lambda; \boldsymbol{y}) = \log L(\lambda; \boldsymbol{y}) = n \log(\lambda) - \lambda n \bar{y}, \quad \lambda > 0.$$

Etsitään suurimman uskottavuuden estimaatti asettamalla log-uskottavuusfunktion derivaatta nollaksi ja ratkaisemalla yhtälö λ:n suhteen.

$$l'(\lambda; \boldsymbol{y}) = \frac{n}{\lambda} - n\bar{y} = 0 \;\Leftrightarrow\; \lambda = \frac{1}{\bar{y}}.$$

Tutkitaan toisen derivaatan avulla onko kyseessä maksimikohta. Havaitaan, että

$$l''(\lambda; \boldsymbol{y}) = -\frac{n}{\lambda^2} < 0 \text{ kaikilla } \lambda > 0,$$

joten $\hat{\lambda} = 1/\bar{y}$ on uskottavuusfunktion lokaali maksimikohta. Koska $1/\bar{y}$ on uskottavuusyhtälön ainoa ratkaisu ja parametriavaruus on avoin väli ($\lambda \in \Omega = (0, \infty)$), niin $\hat{\lambda} = 1/\bar{y}$ on parametrin λ suurimman uskottavuuden estimaatti. Havaittu informaatio on

$$j(\hat{\lambda}; \boldsymbol{y}) = -l''(\hat{\lambda}; \boldsymbol{y}) = \frac{n}{\hat{\lambda}^2} = n\bar{y}^2.$$

Fisherin informaatio on

$$i(\lambda) = \mathrm{E}[-l''(\lambda; \boldsymbol{Y})] = \mathrm{E}\left[\frac{n}{\lambda^2}\right] = \frac{n}{\lambda^2}.$$

Lasketaan vielä odotusarvo $\mathrm{E}[l'(\lambda; \boldsymbol{Y})^2]$.

$$\begin{aligned}
\mathrm{E}[l'(\lambda; \boldsymbol{Y})^2] &= \mathrm{Var}[l'(\lambda; \boldsymbol{Y})] + (\mathrm{E}[l'(\lambda; \boldsymbol{Y})])^2 = \mathrm{Var}\left[\frac{n}{\lambda} - n\bar{Y}\right] + \left(\mathrm{E}\left[\frac{n}{\lambda} - n\bar{Y}\right]\right)^2 \\
&= \mathrm{Var}\left[\frac{n}{\lambda} - \sum_{i=1}^{n} Y_i\right] + \left(\frac{n}{\lambda} - n\mathrm{E}(\bar{Y})\right)^2 = \mathrm{Var}\left(\sum_{i=1}^{n} Y_i\right) + \left(\frac{n}{\lambda} - n\frac{1}{\lambda}\right)^2 \\
&\overset{\perp\!\!\!\perp}{=} \sum_{i=1}^{n} \mathrm{Var}\,(Y_i) + 0 = \sum_{i=1}^{n} \frac{1}{\lambda^2} = \frac{n}{\lambda^2}.
\end{aligned}$$

3. Tarkastellaan mallia, jossa havaintoja vastaavat satunnaismuuttujat $Y_1, \ldots, Y_n$ ovat riippumattomat. Mallin parametri on yksiulotteinen θ. Totea, että mallin havaittu informaatio ja Fisherin informaatio ovat

$$j(\theta; \boldsymbol{y}) = j_1(\theta; y_1) + \cdots + j_n(\theta; y_n) \qquad \text{ja} \qquad i(\theta) = i_1(\theta) + \cdots + i_n(\theta),$$

jossa $j_k(\theta; y_k)$ on pelkästään yhteen havaintoon y_k perustuva havaittu informaatio ja $i_k(\theta) = E[j_k(\theta; Y_k)]$ on vastaava Fisherin informaatio. Miten tulkitset tämän tuloksen? [Monisteen harjoitustehtävä 2.13]

Ratk. Koska havainnot ovat riippumattomia, niin tilastolliseksi malliksi saadaan

$$f_{\boldsymbol{Y}}(\boldsymbol{y}; \theta) = \prod_{i=1}^{n} f_{Y_i}(y_i; \theta).$$

Uskottavuusfunktio on siis

$$L(\theta; \boldsymbol{y}) = f_{\boldsymbol{Y}}(\boldsymbol{y}; \theta) = \prod_{i=1}^{n} f_{Y_i}(y_i; \theta) = \prod_{i=1}^{n} L_i(\theta; y_i),$$

jossa $L_i(\theta; y_i)$ on i:nnettä havaintoa vastaava uskottavuusfunktio. Log-uskottavuusfunktioksi tulee näinollen

$$l(\theta; \boldsymbol{y}) = \log L(\theta; \boldsymbol{y}) = \sum_{i=1}^{n} \log L_i(\theta; y_i) = \sum_{i=1}^{n} l_i(\theta; y_i),$$

jossa $l_i(\theta; y_i)$ on i:nnettä havaintoa vastaava log-uskottavuusfunktio. Havaituksi informaatioksi saadaan

$$\begin{aligned} j(\theta; \boldsymbol{y}) &= -l''(\theta; \boldsymbol{y}) = -\frac{d^2}{d\theta^2} \sum_{k=1}^{n} l_k(\theta; y_k) \\ &= -\sum_{k=1}^{n} \frac{d^2}{d\theta^2} l_k(\theta; y_k) = \sum_{k=1}^{n} (-l_k''(\theta; y_k)) = \sum_{k=1}^{n} j_k(\theta; y_k), \end{aligned}$$

jossa $j_k(\theta; y_k)$ on k:nnetta havaintoa vastaava havaittu informaatio. Fisherin informaatioksi saadaan

$$i(\theta) = \mathrm{E}[j(\theta; \boldsymbol{Y})] = \mathrm{E}\left[\sum_{k=1}^{n} j_k(\theta; Y_k)\right] = \sum_{k=1}^{n} \mathrm{E}\left(j_k(\theta; Y_k)\right) = \sum_{k=1}^{n} i_k(\theta),$$

jossa $i_k(\theta)$ on k:nnetta havaintoa vastaava Fisherin informaatio.

Tuloksen tulkinta: Kun havainnot ovat riippumattomia, niin lisähavainnot kasvattavat informaatiota additiivisesti.

4. Olkoot $Y_1, \ldots, Y_n \perp\!\!\!\perp$ ja $Y_i \sim N(\beta x_i, \sigma_0^2)$, jossa $x_1, \ldots, x_n$ ovat tunnettuja lukuja ja $\sigma_0^2 > 0$ on tunnettu. Johda parametrin β suurimman uskottavuuden estimaatti ja mallin havaittu informaatio $j(\beta; \boldsymbol{y})$. Vertaa havaitun informaation arvoja tapauksissa a) $x_i \geq c > 0$ ja b) $x_i = 1/i$, kun $n \to \infty$.

   Ratk. Koska havainnot ovat riippumattomia, niin tilastolliseksi malliksi saadaan

$$
\begin{aligned}
f_{\boldsymbol{Y}}(\boldsymbol{y}; \beta) &= \prod_{i=1}^{n} \frac{1}{\sqrt{2\pi\sigma_0^2}} \exp\left\{-\frac{1}{2\sigma_0^2}(y_i - \beta x_i)^2\right\} \\
&= (2\pi\sigma_0^2)^{-n/2} \exp\left\{-\frac{1}{2\sigma_0^2}\sum_{i=1}^{n}(y_i - \beta x_i)^2\right\}.
\end{aligned}
$$

   Olkoon $c(\boldsymbol{y}) = (2\pi\sigma_0^2)^{n/2}$. Tällöin uskottavuusfunktioksi saadaan

$$
L(\beta; \boldsymbol{y}) = c(\boldsymbol{y}) f_{\boldsymbol{Y}}(\boldsymbol{y}; \beta) = \exp\left\{-\frac{1}{2\sigma_0^2}\sum_{i=1}^{n}(y_i - \beta x_i)^2\right\}.
$$

   Logaritminen uskottavuusfunktio on nyt

$$
l(\beta; \boldsymbol{y}) = \log L(\beta; \boldsymbol{y}) = -\frac{1}{2\sigma_0^2}\sum_{i=1}^{n}(y_i - \beta x_i)^2.
$$

   Etsitään suurimman uskottavuuden estimaatti derivointimenetelmän avulla. Derivoidaan log-uskottavuusfunktio, asetetetaan derivaatta nollaksi ja ratkaistaan β:n suhteen.

$$
\begin{aligned}
l'(\beta; \boldsymbol{y}) &= \frac{d}{d\beta}\left[-\frac{1}{2\sigma_0^2}\sum_{i=1}^{n}\frac{d}{d\beta}(y_i - \beta x_i)^2\right] = -\frac{1}{2\sigma_0^2}\sum_{i=1}^{n}\frac{d}{d\beta}(y_i - \beta x_i)^2 \\
&= -\frac{1}{2\sigma_0^2}\sum_{i=1}^{n} 2(y_i - \beta x_i)(-x_i) = \frac{1}{\sigma_0^2}\sum_{i=1}^{n} x_i(y_i - \beta x_i) \\
&= \frac{1}{\sigma_0^2}\left(\sum_{i=1}^{n} x_i y_i - \beta\sum_{i=1}^{n} x_i^2\right) = 0 \quad \Rightarrow \quad \beta = \frac{\sum_{i=1}^{n} x_i y_i}{\sum_{i=1}^{n} x_i^2}
\end{aligned}
$$

   Tutkitaan toisen derivaatan avulla onko kyseessä maksimikohta.

$$
l''(\beta; \boldsymbol{y}) = -\frac{\sum_{i=1}^{n} x_i^2}{\sigma_0^2} < 0 \text{ kaikilla } \beta \in \mathbb{R},
$$

   joten

$$
\hat{\beta} = \frac{\sum_{i=1}^{n} x_i y_i}{\sum_{i=1}^{n} x_i^2}
$$

on parametrin β suurimman uskottavuuden estimaatti.

Havaittu informaatio on

$$j(\beta;\boldsymbol{y}) = -l''(\beta;\boldsymbol{y}) = \frac{\sum_{i=1}^n x_i^2}{\sigma_0^2}.$$

(a) Olkoon $x_i \geq c > 0$. Tällöin

$$j(\beta;\boldsymbol{y}) = \frac{\sum_{i=1}^n x_i^2}{\sigma_0^2} \geq \frac{\sum_{i=1}^n c^2}{\sigma_0^2} = \frac{nc^2}{\sigma_0^2} \underset{n\to\infty}{\longrightarrow} \infty,$$

joten

$$\lim_{n\to\infty} j(\beta;\boldsymbol{y}) = \infty.$$

(b) Olkoon $x_i = 1/i$. Tällöin

$$j(\beta;\boldsymbol{y}) = \frac{\sum_{i=1}^n x_i^2}{\sigma_0^2} = \frac{\sum_{i=1}^n \frac{1}{i^2}}{\sigma_0^2}$$

Kuviosta 1 nähdään, että

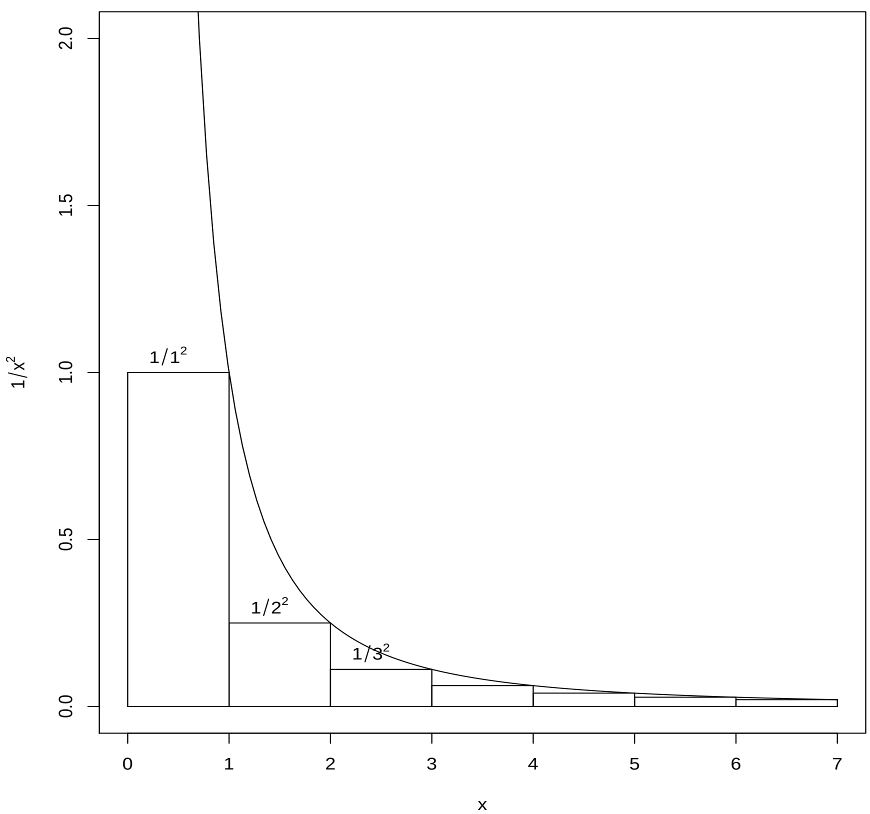
$$0 < \sum_{i=1}^{\infty} \frac{1}{i^2} < 1 + \int_1^{\infty} x^{-2}dx = 1 + \Big|_1^{\infty}(-x^{-1}) = 1 + 1 = 2,$$

joten

$$\lim_{n\to\infty} j(\beta;\boldsymbol{y}) = \frac{\sum_{i=1}^{\infty} \frac{1}{i^2}}{\sigma_0^2} < \frac{2}{\sigma_0^2},$$

Huom! Voidaan osoittaa, että $\sum_{i=1}^{\infty} \frac{1}{i^2} = \pi^2/6 \approx 1.645$. Tällöin tarkaksi raja-arvoksi saadaan

$$\lim_{n\to\infty} j(\beta;\boldsymbol{y}) = \frac{\pi^2/6}{\sigma_0^2} = \frac{\pi^2}{6\sigma_0^2}.$$



Kuva 1: Summan $\sum_{i=1}^{n} \frac{1}{i^2}$ arvioiminen funktion $f(x) = 1/x^2$ avulla.