Tilastollinen päättely II, syksy 2015 – kevät 2016
Harjoitus 3
Malliratkaisut

Tehtävissä 1–3 käsitellään vielä uskottavuusfunktiota ja suurimman uskottavuuden estimointia. Luentomonisteen jaksot 2.1–2.3.

1. (Monisteen teht. 2.5.) Olkoon θ positiivinen parametri, ja asetetaan

$$f(y;\theta) = 2\theta^{-1} y \exp(-y^2/\theta), \qquad \text{kun } y > 0,$$

ja $f(y;\theta) = 0$, kun $y \leq 0$.

a) Tarkista integroimalla, että tämä kelpaa erään jatkuvan jakauman tiheysfunktioksi.

b) Oletetaan, että $Y_1, \ldots, Y_n$ ovat riippumattomia ja noudattavat kukin em. jakaumaa. Muodosta tämän mallin uskottavuusfunktio ja log-uskottavuusfunktio sekä määritä suurimman uskottavuuden estimaatti $\hat{\theta}$, kun aineisto on $\mathbf{y} = (y_1, \ldots, y_n)$.

Ratkaisu 1

a) Tarkistetaan, että tiheysfunktion integraali koko reaaliakselin yli on 1.

$$\begin{aligned}
\int_{-\infty}^{\infty} f(y;\theta)\,dy &= \int_0^{\infty} 2\theta^{-1} y \exp(-y^2/\theta)\,dy \\
&= \lim_{a\to\infty} -\int_0^a -2\theta^{-1} y \exp(-y^2/\theta)\,dy \\
&= \lim_{a\to\infty} - \Big/_{0}^{a} \exp(-y^2/\theta) \\
&= \lim_{a\to\infty} -(\exp(-a^2/\theta) - 1) \\
&= 0 + 1 \\
&= 1
\end{aligned}$$

$f(y;\theta)$ kelpaa siis tiheysfunktioksi.

b) Muodostetaan ensin yhteistiheysfunktio.

$$f_{\mathbf{Y}}(\mathbf{y};\theta) = \prod_{i=1}^{n} 2\theta^{-1} y_i \exp(-y_i^2/\theta) = 2^n \theta^{-n} \left(\prod_{i=1}^{n} y_i \right) \exp(-\sum_{i=1}^{n} y_i^2/\theta)$$

Kerrotaan parametrista θ riippumattomat tekijät pois ja saadaan uskottavuusfunktioksi:

$$L(\theta;\mathbf{y}) = \theta^{-n} \exp(-\sum_{i=1}^{n} y_i^2/\theta)$$

Ja log-uskottavuusfunktioksi saadaan:

$$l(\theta;\mathbf{y}) = -n \log\theta - \sum_{i=1}^{n} y_i^2/\theta$$

Derivoidaan ja etsitään nollakohdat:

$$l'(\theta; \mathbf{y}) = -\frac{n}{\theta} + \frac{\sum_{i=1}^n y_i^2}{\theta^2}$$

$$l'(\theta; \mathbf{y}) = 0 \Leftrightarrow -\frac{n}{\theta} + \frac{\sum_{i=1}^n y_i^2}{\theta^2} = 0 \Leftrightarrow \theta = \frac{\sum_{i=1}^n y_i^2}{n}$$

Ja koska

$$l''(\theta; \mathbf{y}) = \frac{n}{\theta^2} - \frac{2\sum_{i=1}^n y_i^2}{\theta^3} < 0, \text{ kun } \theta = \frac{\sum_{i=1}^n y_i^2}{n}$$

niin ensimmäisen derivaatan nollakohta on uskottavuusfunktion maksimikohta, eli suurimman uskottavuuden estimaatti $\hat{\theta} = \frac{\sum_{i=1}^n y_i^2}{n}$.

2. (Vrt. monisteen teht. 2.15.) Tarkastellaan parametrin $\theta \in (0,1)$ *logit-muunnosta* $\phi = \phi(\theta) = \log\{\theta/(1-\theta)\}$.

a) Totea, että kyseessä on kääntäen yksikäsitteinen eli bijektiivinen funktio $(0,1) \to \mathbb{R}$, ja määritä käänteisfunktio $\theta = \theta(\phi)$. Hahmottele myös kuvaaja (θ, ϕ)-koordinaatistoon.

b) Toistokoemalli $Y_1, \ldots, Y_n \sim B(\theta) \perp\!\!\!\perp$ tai $K \sim \mathrm{Bin}(n, \theta)$ uudelleenparametroidaan logit-muunnoksella. Johda syntyvän mallin log-uskottavuusfunktio $l^*(\phi)$ ja su-estimaatti $\hat{\phi}$.

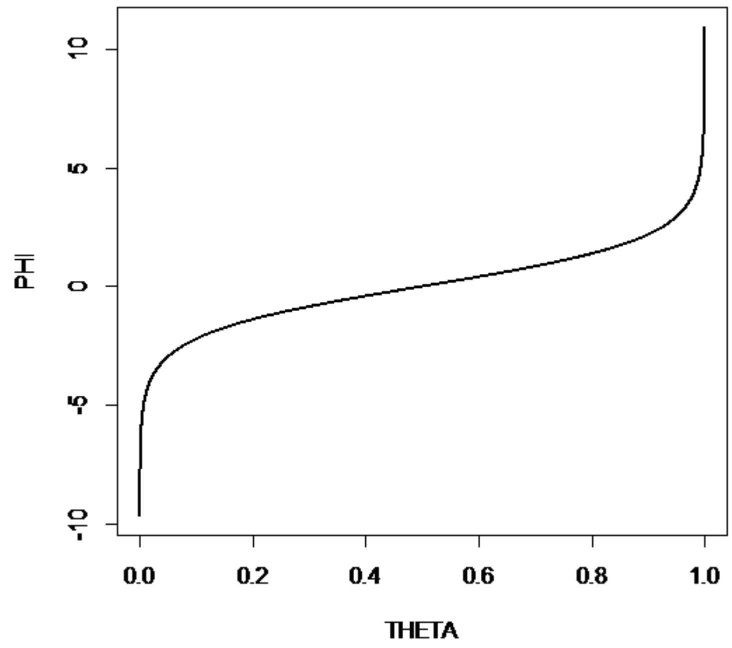
Lisätieto. Ns. logistisessa regressiossa mallitetaan toistokokeen onnistumistodennäköisyyden logit-muunnosta joidenkin selittävien muuttujien avulla. Esimerkiksi $Y_1, \ldots, Y_n \perp\!\!\!\perp$, jossa $Y_i \sim B(\theta_i)$ ja $\phi(\theta_i) = \alpha + \beta x_i$, jolloin mallin parametri olisi pari (α, β). Tällaisia malleja tutkitaan yleistettyjen lineaaristen mallien kurssilla.

Ratkaisu 2

a) Muunnos $\phi(\theta)$ on bijektiivinen funktio, sillä saadaan ratkaistua (kun $\phi \in \mathbb{R}$)

$$\begin{aligned} \phi &= \log(\theta/(1-\theta)) \\ e^{\phi} &= \frac{\theta}{1-\theta} \\ \theta &= \frac{e^{\phi}}{1+e^{\phi}}, \text{ jossa selvästi pätee } 0 < \theta < 1 \end{aligned}$$

Ja käänteisfunktio $\theta = \theta(\phi) = e^{\phi}/(1+e^{\phi})$.



b) Alkuperäinen tilastollinen malli on muotoa

$$f_K(k;\theta) = \binom{n}{k}\theta^k(1-\theta)^{n-k}, \text{ kun } y = 0, 1, \ldots, n$$

joten sen uskottavuusfunktio on

$$L(\theta;k) = \theta^k(1-\theta)^{n-k}, \text{ kun } 0 < \theta < 1.$$

Tästä saadaan log-uskottavuusfunktioksi

$$l(\theta;k) = k\log(\theta) + (n-k)\log(1-\theta)$$

Uudelleen parametroidaan log-uskottavuusfunktio

$$\begin{aligned}
l^*(\phi) = l(\theta(\phi)) &= k\log\left(\frac{e^\phi}{1+e^\phi}\right) + (n-k)\log\left(1 - \frac{e^\phi}{1+e^\phi}\right) \\
&= k\phi - k\log(1+e^\phi) + (n-k)\log\left(\frac{1}{1+e^\phi}\right) \\
&= k\phi - k\log(1+e^\phi) - (n-k)\log(1+e^\phi) \\
&= k\phi - n\log(1+e^\phi)
\end{aligned}$$

Su-estimaatin invarianssiominaisuuden perusteella voidaan selvittää su-estimaatti $\hat{\phi}$ alkuperäisen mallin su-estimaatin avulla. Monisteen esimerkin 2.2.5 perusteella tiedetään, että $\hat{\theta} = \frac{k}{n}$.

Nyt saadaan invarianssiominaisuuden perusteella parametrin ϕ su-estimaatti $\hat{\phi}$

$$\hat{\phi} = \phi(\hat{\theta}) = \log\left(\frac{k/n}{1-k/n}\right) = \log\left(\frac{k}{n-k}\right)$$

Su-estimaatin olisi voinut etsiä myös uudelleenparametroidun log-uskottavuusfunktion avulla tuttuun tapaan.

3. Palautetaan mieleen JTP-kurssilta, miten parametria θ koskevaa päättelyä tehdään bayesläisittäin: θ:lle määritellään priorijakauma (ptf/tf) $p(\theta)$ ja jos $L(\theta) = L(\theta; \mathbf{y})$ on havaittua aineistoa vastaava uskottavuusfunktio, niin θ:n posteriorijakauma saadaan Bayesin kaavasta

$$p(\theta|\mathbf{y}) = c(\mathbf{y})p(\theta)L(\theta; \mathbf{y}).$$

Tässä vakio $c(\mathbf{y})$ on käänteisluku funktion $p(\theta)L(\theta; \mathbf{y})$ summasta tai integraalista, mutta emme nyt tarvitse tätä tietoa.

Pohditaan θ:n piste-estimointia bayesläisestä näkökulmasta toistokoemallissa $K \sim \text{Bin}(n, \theta)$. Valitaan priorijakaumaksi (beetajakauman) tiheysfunktio

$$p(\theta) = c \cdot \theta^a (1-\theta)^b, \qquad 0 \leq \theta \leq 1,$$

jossa $a \geq 0$, $b \geq 0$ ja $c > 0$ on eräs a:sta ja b:stä riippuva vakio, jonka tarkkaa arvoa emme myöskään tarvitse.

a) Totea, että priorijakauman moodi eli piste, jossa se saa suurimman arvonsa, on kohdassa $\theta_0 = a/(a+b)$. Tapauksessa $a = b = 0$ sopikaamme, että $\theta_0 = \frac{1}{2}$.

b) Muodosta posteriorijakauma $p(\theta|k)$, kun olemme havainneet k onnistumista. (Ei tarvitse laskea vakioiden arvoja.)

c) Etsi posteriorijakauman moodi θ^*. Totea, että se voidaan kirjoittaa painotettuna keskiarvona θ_0:stä ja su-estimaatista $\hat{\theta} = k/n$:

$$\theta^* = \frac{a+b}{a+b+n}\theta_0 + \frac{n}{a+b+n}\hat{\theta}.$$

d) Millä a:n ja b:n arvoilla on $\theta^* = \hat{\theta}$? Mikä priorijakauma on tällöin?

e) Totea, että $\theta^* \to \hat{\theta}$, kun $n \to \infty$ (olettaen, että $\hat{\theta}$:n arvo pysyy muuttumattomana). Priorijakauman vaikutus siis ”häviää” kun havaintojen lukumäärä kasvaa.

Vihje. a- ja c-kohdissa maksimoitavat funktiot ovat samaa muotoa kuin toistokoemallin uskottavuusfunktio, joten tarvittavat laskut on olennaisesti jo tehty (ks. monisteen esimerkit 2.1.5 ja 2.2.5).

Lisätietoa. Posteriorijakauman moodi on eräs bayesläinen piste-estimointimenetelmä; tasaisen priorin tapauksessa se yhtyy su-estimaattiin. Toinen paljon käytetty vaihtoehto on posteriorijakauman odotusarvo.

Ratkaisu 3

Huom! Priorijakauma on beetajakauman tiheysfunktio parametrein $a + 1$ ja $b + 1$.

a) Etsitään priorijakauman moodi, eli selvitetään sen maksimikohdat derivaatan nollakohtien avulla. Logaritmoidaan ensin priorijakauma:

$$\log p(\theta) = \log c + a \log \theta + b \log(1-\theta)$$

Derivoidaan:

$$\frac{\partial}{\partial\theta}\log p(\theta) = \frac{a}{\theta} - \frac{b}{1-\theta}$$

Etsitään seuraavaksi derivaatan nollakohta.

$$\begin{aligned}
\frac{a}{\theta} - \frac{b}{1-\theta} &= 0 \\
a(1-\theta) &= b\theta \\
\frac{1-\theta}{\theta} &= \frac{b}{a} \\
\frac{1}{\theta} &= \frac{a+b}{a} \\
\theta &= \frac{a}{a+b}
\end{aligned}$$

Joten priorijakauman moodi on kohdassa $\theta_0 = \frac{a}{a+b}$, sillä

$$p'(\theta) > 0, \text{ kun } \theta < \theta_0$$

ja

$$p'(\theta) < 0, \text{ kun } \theta > \theta_0$$

b) Selvitetään uskottavuusfunktio posteriorijakauman muodostamista varten. Koska $K \sim Bin(n, \theta)$ saadaan uskottavuusfunktioksi

$$L(\theta; k) = \theta^k (1-\theta)^{n-k}$$

Nyt saadaan posteriorijakaumaksi

$$p(\theta|k) = c(k)\theta^a(1-\theta)^b\theta^k(1-\theta)^{n-k} \propto \theta^{a+k}(1-\theta)^{b+n-k}$$

c)Posteriorijakauma on samaa muotoa kuin priorijakauma, mutta korvataan $a \to a+k$ ja $b \to b+n-k$. Posteriorijakauman moodi löytyy siis kohdasta

$$\begin{aligned}
\theta^* &= \frac{a+k}{a+k+b+n-k} = \frac{a+k}{a+b+n} = \frac{a}{a+b+n} + \frac{k}{a+b+n} \\
&= \frac{a(a+b)}{(a+b+n)(a+b)} + \frac{kn}{(a+b+n)n} = \frac{a+b}{a+b+n}\theta_0 + \frac{n}{a+b+n}\hat{\theta}
\end{aligned}$$

Siis $\theta^* = \frac{a+b}{a+b+n}\theta_0 + \frac{n}{a+b+n}\hat{\theta}$ missä $\theta_0 = \frac{a}{a+b}$ ja $\hat{\theta} = \frac{k}{n}$

d) $\theta^* = \frac{a+b}{a+b+n}\theta_0 + \frac{n}{a+b+n}\hat{\theta} = \hat{\theta}$, jos ja vain jos $a+b=0$ eli $a=b=0$. Nyt priorijakauma on muotoa $p(\theta) = c\theta^0(1-\theta)^0 = c$, jossa $c = B(1,1) = 1$, eli tasajakauma $\text{Tas}(0,1)$.

e) $\theta^* = \frac{a+b}{a+b+n}\theta_0 + \frac{n}{a+b+n}\hat{\theta} = \frac{a+b}{n\left(\frac{a}{n}+\frac{b}{n}+1\right)}\theta_0 + \frac{n}{n\left(\frac{a}{n}+\frac{b}{n}+1\right)}\hat{\theta} \to 0\cdot\theta_0 + 1\cdot\hat{\theta} = \hat{\theta}$ kun $n \to \infty$

Tehtävät 4 ja 5 liittyvät havaitun informaation ja Fisherin informaation käsitteisiin. Monisteen jakso 2.4.

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

Ratkaisu 4

Harjoituksista 2 tiedetään, että tilastollinen malli on

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

uskottavuusfunktio on

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

ja log-uskottavuusfunktio on

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

Suurimman uskottavuuden estimaatti löydettiin log-uskottavuusfunktion derivaatan nollakohdasta

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

Tämä on maksimikohta, koska

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

joten $\hat{\lambda} = \frac{1}{\bar{y}}$ on suurimman uskottavuuden estimaatti.
Nyt saadaan havaituksi informaatioksi

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

Fisherin informaatio on

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

Lasketaan odotusarvo $E[l'(\lambda; \mathbf{Y})^2]$

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

5. (Monisteen teht. 2.13.) Tarkastellaan mallia, jossa havaintoja vastaavat satunnaismuuttujat $Y_1, \ldots, Y_n$ ovat riippumattomat. Mallin parametri on θ. Totea, että mallin havaittu informaatio ja Fisherin informaatio ovat

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

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?

Ratkaisu 5

Satunnaismuuttujat ovat riippumattomia, joten tilastolliseksi malliksi saadaan

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

joten uskottavuusfunktio on

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

jossa $L_i(\theta; y_i)$ on i:nnettä havaintoa vastaava uskottavuusfunktio.

Nyt saadaan log-uskottavuusfunktioksi

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

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

Havaituksi informaatioksi saadaan

$$\begin{aligned} j(\theta; \mathbf{y}) &= -l''(\theta; \mathbf{y}) \\ &= -\frac{d^2}{d\theta^2} \sum_{i=1}^{n} l_i(\theta; y_i) \\ &= -\sum_{i=1}^{n} \frac{d^2}{d\theta^2} l_i(\theta; y_i) \\ &= \sum_{i=1}^{n} (-l_i''(\theta; y_i)) \\ &= \sum_{i=1}^{n} j_i(\theta; y_i), \end{aligned}$$

jossa $j_i(\theta; y_i)$ on i:nnetta havaintoa vastaava havaittu informaatio.

Nyt saadaan Fisherin informaatioksi

$$
\begin{aligned}
i(\theta) &= E[j(\theta; \mathbf{Y})] \\
&= E\left[\sum_{i=1}^{n} j_i(\theta; Y_i)\right] \\
&= \sum_{i=1}^{n} E\left(j_i(\theta; Y_i)\right) \\
&= \sum_{i=1}^{n} i_i(\theta),
\end{aligned}
$$

jossa $i_i(\theta)$ on i:nnetta havaintoa vastaava Fisherin informaatio.

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