Cvič. 1.2.4 (Laplace 1786), BCh str. 12

Pohlaví dětí narozených v Paříži:

251 527 kluků

241 945 holek

$\pi = P(\text{narození kluka})$

DATA: $Y_1, \dots, Y_n$, $n = 493472$

$Y_i = \mathbb{I}(\text{i-té dítě je kluk})$

„MODEL": $Y_i | \pi \overset{iid}{\sim} A(\pi)$

$L_i(\pi) = p(y_i|\pi) = \pi^{y_i}(1-\pi)^{1-y_i}$

$L(\pi) = p(y|\pi) = \prod_{i=1}^{n} L_i(\pi) = \pi^{\sum y_i}(1-\pi)^{n-\sum y_i}$

označ: $N_1 = \sum_{i=1}^{n} Y_i$, $n_1 = \sum_{i=1}^{n} y_i$

$N_0 = \sum_{i=1}^{n}(1-Y_i)$, $n_0 = \sum_{i=1}^{n}(1-y_i) = n - n_1$

$\Rightarrow L(\pi) = \pi^{n_1}(1-\pi)^{n-n_1}$

NA OKRAJ: MLE

$\hat{\pi}_{ML} = \operatorname{argmax} L(\pi) = \frac{n_1}{n}$

1)

Bayes

apriorní rozdělení $p(\pi)$? , $0<\pi<1$

→ o π nic nevíme, apriori $\pi \sim U(0,1)$

$$p(\pi) = 1 \cdot \mathbb{I}_{(0,1)}(\pi)$$

→ aposteriorní : $p(\pi|y) \propto \underbrace{p(y|\pi)}_{L(\pi)} \cdot p(\pi) =$

$$= \pi^{n_1}(1-\pi)^{n-n_1} \mathbb{I}, \quad 0<\pi<1$$

↙ poznáváme jádro Beta hustoty, netřeba se trápit s normující konstantou

$$\to \int_0^1 \pi^{n_1}(1-\pi)^{n-n_1} d\pi = B(n_1+1, n-n_1+1)$$

tedy $\pi | Y \sim Be\left(\sum_{i=1}^n Y_i + 1, n - \sum_{i=1}^n Y_i + 1\right)$

$$= Be(N_1+1, N_0+1)$$

→ už Laplace spočítal

$$P\left(\pi \leq 0{,}5 \,\middle|\, \sum_{i=1}^n Y_i = 251\,527\right) =$$

$$= F_{Be(\dots)}(0{,}5) = 1{,}146 \cdot 10^{-42}$$

$$(= \text{pbeta}(0{,}5, n_1+1, n_0+1))$$

Bayes odhad

$$\hat{\pi}_B := \mathbb{E}(\pi / Y) = \frac{\sum Y_i + 1}{n+2} = \frac{N_1 + 1}{n+2}$$

→ v klasickém smyslu
- není nestranný
- je konzistentní

jako bych do dat při MLE přidal 1 kluka a 1 holku

Věrohodn. interval:

ET interval : $\left(Q_y(\alpha/2),\ Q_y(1-\frac{\alpha}{2}) \right)$

Q_y = kvantil fce $Be(n_1+1, n_0+1)$

zde vyjde: $(0{,}5083,\ 0{,}5111)$
($\alpha = 0{,}05$)

(od klasického asympt. intervalu spolehlivosti se liší na nevím kolikátém deset. místě)

Při práci s 0/1 veličinou nás často zajímá šance (odds) $\rho = \frac{\pi}{1-\pi}$

$p(\rho|y) = ?$ Potřeba znovu počítat nebo se jedná o úlohu pro 2./3. ročník?

$p(\rho|y) = p(t(\pi)|y)$

$t(\pi) = \frac{\pi}{1-\pi} = \rho \in (0,\infty)$

$\tau(\rho) = \frac{\rho}{1+\rho} = \pi \in (0,1)$

$\tau'(\rho) = \frac{1}{(1+\rho)^2}$

$$p(\rho|y) = \frac{1}{B(n_1+1, n_0+1)} \cdot \left(\frac{\rho}{1+\rho}\right)^{n_1} \left(\frac{1}{1+\rho}\right)^{n-n_1} \cdot \frac{1}{(1+\rho)^2} =$$

$$= \frac{1}{B(n_1+1, n_0+1)} \cdot \frac{\rho^{n_1}}{(1+\rho)^{n+2}}, \quad \rho \in (0,\infty)$$

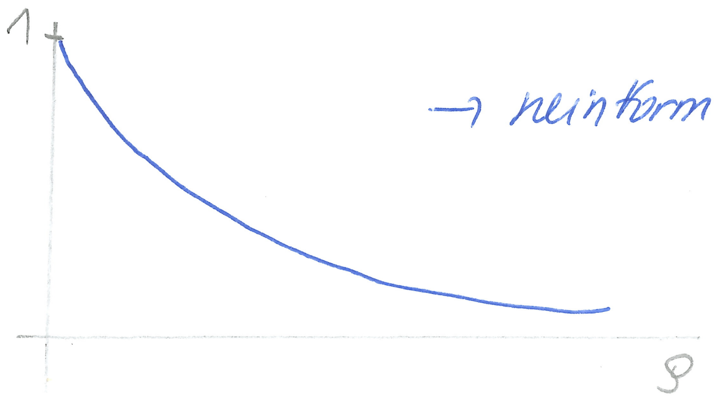
nezapomenout!

toto aposteriorní rozdělení odpovídá

$p(\pi) = \mathbb{I}_{(0,1)}(\pi)$ (neinformativní)

Jaké vlastně bylo apriorní rozdělení pro ρ?

$p(\rho) = 1 \cdot \frac{1}{(1+\rho)^2} \mathbb{I}_{(0,\infty)}(\rho)$



→ neinformativní?

Co takhle uvažovat „neinformativní" apriorní rozdělení pro ϱ:

$$p(\varrho) \propto 1 \,, \quad 0<\varrho<\infty \qquad \text{(nevlastní)}$$

$$\Rightarrow \quad p(\varrho|y) \propto p(y|\varrho)\cdot p(\varrho) \quad =$$

$\equiv p(y|\pi(\varrho))$ bez Jakobiánu!

$$= \left(\frac{\varrho}{1+\varrho}\right)^{n_1}\left(\frac{1}{1+\varrho}\right)^{n-n_1}\cdot 1 = \frac{\varrho^{n_1}}{(1+\varrho)^n}$$

SROVNEJ:

$p(\pi)=1 \,, \quad 0<\pi<1$ $\qquad$ $p(\pi) =$ snadné D.Cvič.

$p(\varrho) = \dfrac{1}{1+\varrho^2} \,, \quad 0<\varrho<\infty$ $\qquad$ $p(\varrho) \propto 1 \,, \quad 0<\varrho<\infty$

$p(\varrho|y) \propto \dfrac{\varrho^{n_1}}{(1+\varrho)^{n+2}}$ $\qquad$ $p(\varrho|y) \propto \dfrac{\varrho^{n_1}}{(1+\varrho)^n}$

$0<\varrho<\infty$

Jeffreys : „neinformativní" apriorní rozdělení
= invariance aposteriorního rozdělení
vůči počáteční parametrizaci
→ týká se principu

princip $p(\theta) \propto 1$ → potenciálně různá aposteriorní rozdělení při změně parametrizace

princip $p(\theta) \equiv$ Jeffreys → stejná aposteriorní rozdělení při změně parametrizace

---

Jeffreys $p_J(\theta) \propto (\det J(\theta))^{1/2}$

$J(\theta)$ = očekávaná Fisherova informace

$Y_1, \dots, Y_n \overset{iid}{\sim} A(\pi)$

$$L(\pi) = p(y|\pi) = \pi^{n_1}(1-\pi)^{n-n_1}, \quad n_1 = \sum_{i=1}^{n} y_i$$

$$l(\pi) = \log L(\pi) = n_1 \log \pi + (n-n_1)\log(1-\pi)$$

$$U(\pi) = \frac{\partial l}{\partial \pi}(\pi) = \frac{n_1}{\pi} - \frac{n-n_1}{1-\pi}$$

$$I(\pi) = -\frac{\partial}{\partial \pi}U(\pi) = -\frac{\partial^2 l}{\partial \pi^2} l(\pi) = \frac{n_1}{\pi^2} + \frac{(n-n_1)}{(1-\pi)^2}$$

≡ pozorovaná informace

pozorovaná informace:

$$I(Y_i;\pi) = \frac{\sum Y_i}{\pi^2} + \frac{n-\sum Y_i}{(1-\pi)^2}$$

očekávaná informace:

$$J(\pi) \underset{\substack{\text{za podmínek} \\ \text{regularity}}}{=} \mathbb{E}_\pi I(Y_i;\pi) =$$

$$= \mathbb{E}_\pi \left(\frac{\sum Y_i}{\pi^2} + \frac{n-\sum Y_i}{(1-\pi)^2} \right) =$$

$$= \frac{n\pi}{\pi^2} + \frac{n-n\pi}{(1-\pi)^2} = \frac{n}{\pi} + \frac{n}{1-\pi} = \frac{n}{\pi(1-\pi)}$$

ODBOČKA 1: $\hat{\pi}_{ML} = \frac{n_1}{n}$ (resp. jako náh. veličina $\frac{N_1}{n}$)

$$\underbrace{(J(\pi))^{1/2}}_{= n J_1(\pi)} (\hat{\pi}_{ML} - \pi) \xrightarrow[n\to\infty]{\mathcal{D}} N(0,1)$$

v případě iid, kde $J_1(\pi)$ = oček. informace od 1 pozorování

$$\hat{\pi}_{ML} \overset{\cdot}{\sim} N(\pi, (J(\pi))^{-1})$$

ODBOČKA 2: též $J(\pi) = E_\pi U(\pi) U^T(\pi)$

$$= \mathrm{var}_\pi U(\pi)$$

Jeffreys: $p_J(\pi) \propto (\det J(\pi))^{1/2}$ = preferuje π s velkou informační hodnotou

- preferuje oblasti, kde skór má velkou variabilitu, tj. kde i jenom „málo odlišná" data vedou k „hodně odlišným" hodnotám skóru
  $\Rightarrow$ i „málo odlišná" data dobře rozlišují různé hodnoty parametru
- preferuje oblasti, kde MLE má malý rozptyl

Zde: $p_J(\pi) \propto \frac{n^{1/2}}{\sqrt{\pi(1-\pi)}} \propto \pi^{-1/2}(1-\pi)^{-1/2}$, $0<\pi<1$

v iid situaci je jedno, zda se vše odvozuje
z informace pro n nebo 1 pozorování

$$J(\pi) = n \cdot J_1(\pi)$$

$$p_J(\pi) = \frac{1}{\pi^*}\sqrt{\frac{1}{\pi(1-\pi)}} \equiv Be\left(\tfrac{1}{2}, \tfrac{1}{2}\right)$$

3,1415...

Od začátku vše v ρ parametrizaci

$$\rho = \frac{\pi}{1-\pi} \qquad \pi(\rho) = \frac{\rho}{1+\rho}$$

$$L(\rho) = p(y|\rho) = \left(\frac{\rho}{1+\rho}\right)^{n_1}\left(\frac{1}{1+\rho}\right)^{n-n_1}$$

$$l(\rho) = n_1 \log(\rho) - n \log(1+\rho)$$

$$U(\rho) = \frac{n_1}{\rho} - \frac{n}{1+\rho} \implies \hat{\rho}_{ML} = \frac{n_1}{n_0} \left(= \rho(\hat{\pi}_{ML})\right)$$

invariance MLE
na parametrizaci

$$I(\rho) = \frac{n_1}{\rho^2} - \frac{n}{(1+\rho)^2}$$

$$J(\rho) = E_\rho\, I(Y, \rho) = \frac{n}{\rho(1+\rho)^2}$$

Jeffreys: $p_J(\rho) \propto \frac{1}{\sqrt{\rho}\,(1+\rho)}$, $0<\rho<\infty$

$$= \frac{1}{\pi^*}\,\frac{1}{\sqrt{\rho}\,(1+\rho)}, \quad 0<\rho<\infty$$

BYLO: $p_J(\pi) \propto \pi^{-1/2}(1-\pi)^{-1/2}$, $0<\pi<1$

- Co nám vyjde, když teď provedeme transformaci do $\rho = \frac{\pi}{1-\pi}$?

$\rho = \frac{\pi}{1-\pi}$

$\pi = \frac{\rho}{1+\rho}$ $\quad \pi' = \frac{1}{(1+\rho)^2}$

$1-\pi = \frac{1}{1+\rho}$

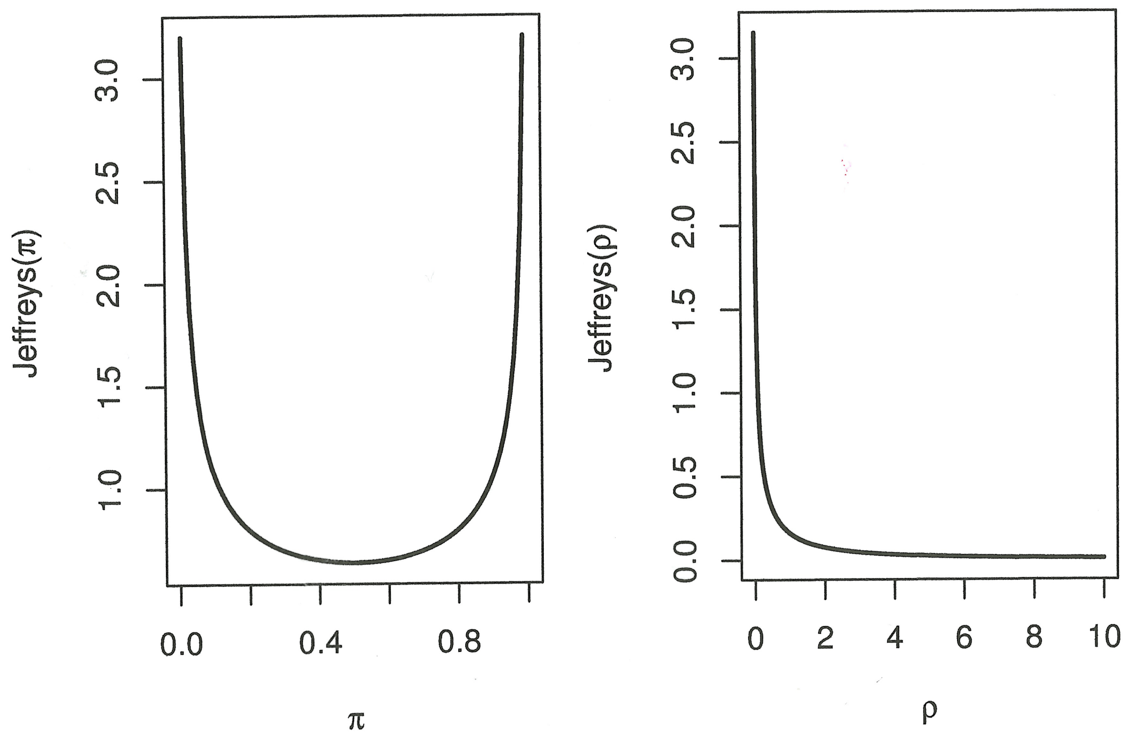
$\Big\downarrow$ transformace

$$p_J^*(\rho) \propto \cancel{\rho^{-1/2}(1+\rho)} \cdot \left(\frac{\rho}{1+\rho}\right)^{-1/2}\left(\frac{1}{1+\rho}\right)^{-1/2} \cdot \frac{1}{(1+\rho)^2}$$

$$= \rho^{-1/2}(1+\rho)^{-1} = \frac{1}{\sqrt{\rho}(1+\rho)} = p_J(\rho)$$

$0<\rho<\infty$

PŘEKVAPENÍ ?


Jeffreys(π)
3.0
2.5
2.0
1.5
1.0
0.0
0.4
0.8
π
Jeffreys(ρ)
3.0
2.5
2.0
1.5
1.0
0.5
0.0
0
2
4
6
8
10
ρ

Příklad:

$Y_1, \dots, Y_n \overset{iid}{\sim}$ DVOJITĚ EXPONENCIÁLNÍ (Laplace)

$$p(y_i|\theta) = \frac{\theta}{2}\exp(-\theta|y_i|) \quad , y_i \in \mathbb{R}$$

$\theta > 0$ parametr

$$E(Y_i|\theta) = 0$$

$$var(Y_i|\theta) = \frac{2}{\theta^2}$$

---

a) ML odhad + související veličiny

$$L(\theta) = \prod_{i=1}^{n}\left(\frac{\theta}{2}\exp(-\theta|y_i|)\right) = \left(\frac{\theta}{2}\right)^n \exp\left(-\theta\sum|y_i|\right)$$

$$l(\theta) = n\log\theta - n\log 2 - \theta\sum|y_i|$$

$$U(\theta) = \frac{n}{\theta} - \sum|y_i|$$

$$U(\theta) = 0$$

$$\frac{n}{\theta} = \sum|y_i|$$

$$\boxed{\hat{\theta}_{ML} = \frac{n}{\sum|y_i|}}$$

$$I(\theta) = -\left(-\frac{n}{\theta^2}\right) = \frac{n}{\theta^2}$$

$$J(\theta) = \frac{n}{\theta^2}$$

JEFFREYS: $p(\theta) \propto \frac{1}{\theta}$, $0 < \theta < \infty$

$\nearrow$

také nevlastní (i když ne $\propto 1$)

Aposteriorní pro Jeffreys:

$$p(\theta|y) \propto L(\theta)p(\theta) = \theta^n \exp(-\theta \sum|y_i|) \cdot \frac{1}{\theta} =$$

$$= \theta^{n-1} \exp(-\theta \sum|y_i|), \quad \theta > 0$$

$\equiv$ Gama rozdělení

$$\theta|Y \sim Ga(n, \sum|Y_i|)$$

opačně než v knihách prof. Anděla

$$\mathbb{E}(\theta|Y) = \frac{n}{\sum|Y_i|} = \hat{\theta}_{ML}$$

Jak by to dopadlo s „neinformativním"
$p(\theta) \propto 1, \theta > 0$?

$$\Rightarrow p(\theta|y) \propto \theta^n \exp(-\theta \sum|y_i|)$$

$$\equiv Ga(n+1, \sum|Y_i|)$$

Domácí úloha:

$Y_1, \ldots, Y_n$ nezávislé, $Y_i \sim N(x_i^T \beta, \sigma^2)$

$x_1, \ldots, x_n \in \mathbb{R}^k$ konstanty

$\Theta = (\beta^T, \sigma^2)^T$: parametry

Jak vypadá Fisherova matice $J(\Theta)$?

$$J(\Theta) = \begin{pmatrix} J_{\beta,\beta} & J_{\beta,\sigma^2} \\ J_{\sigma^2,\beta} & J_{\sigma^2,\sigma^2} \end{pmatrix}$$