Tilastollinen päättely I Harj. 5 vko 7 2012

(5.1) $X_i \sim Ber(\theta)$ $i=1,\dots,5$ $f(x;\theta)=\theta^{x_i}(1-\theta)^{1-x_i}$, $x_i=0$ tai $x_i=1$

Tulos 10011

Uskottavuusfunktio

$$L_1(\theta)=\prod_{i=1}^{5}\theta^{x_i}(1-\theta)^{1-x_i} = \theta^1(1-\theta)^{1-1}\cdot\theta^0(1-\theta)^{1-0}\cdot\theta^0(1-\theta)^{1-0}\cdot\theta^1(1-\theta)^{1-1}\cdot\theta^1(1-\theta)^{1-1}$$
$$=\theta^3(1-\theta)^2$$

Onnistumisten lkm $X=X_1+X_2+\dots+X_5 \sim Bin(5,\theta)$ $f(x;\theta)=\binom{5}{x}\theta^x(1-\theta)^{5-x}$

Tiedetään, että onnistumisia on 3.

$L_2(\theta)=\binom{5}{3}\theta^3(1-\theta)^2$

SUE (suurimman uskottavuuden estimaatti)

Molemmista ($L_1(\theta)$ ja $L_2(\theta)$) saadaan sama SUE

$$L_1'(\theta)=3\theta^2(1-\theta)^2-2(1-\theta)\theta^3$$
$$=\theta^2(1-\theta)[3(1-\theta)-2\theta]$$
$$=\theta^2(1-\theta)(3-5\theta)=0$$ Ratkaisut $\theta=0$ $\theta=1$ ja $\theta=\frac{3}{5}$

$\hat\theta=\frac{3}{5}$ antaa maksimin eli $\hat\theta=\frac{3}{5}$ on SUE.

(5.2.) $X_i\sim N(\theta,1)$ $i=1,2,3$

$\bar X\sim N(\theta,\frac{1}{3})$ silloin $f(\bar x;\theta)=\frac{1}{\sqrt{2\pi/3}}e^{-\frac{1}{2}\left(\frac{\bar x-\theta}{\sqrt{1/3}}\right)^2}=\frac{1}{\sqrt{2\pi/3}}e^{-\frac{3}{2}(\bar x-\theta)^2}$

Otoskeskiarvo saatiin 4, jolloin θ:in uskottavuusfunktio

$L(\theta)=\frac{1}{\sqrt{2\pi/3}}e^{-\frac{3}{2}(4-\theta)^2}$, joka maksimoituu kun $e^{-\frac{3}{2}(4-\theta)^2}$ maksimoituu

eli kun $-\frac{3}{2}(4-\theta)^2$ maksimoituu

Derivoidaan: $-\frac{3}{2}\cdot 2(4-\theta)(-1)=3(4-\theta)=0$ kun $\theta=4$

2. derivaatta <0 θ:n arvolla 4, joten kyseessä maksimi.

$\therefore$ SUE $\hat\theta=4$

(5.3.) $f(x;\theta)=\frac{1}{\theta}e^{-\frac{x}{\theta}}$ $x\geq 0$; $F(x;\theta)=\int_0^x f(x;\theta)dx=-\int_0^x -\frac{1}{\theta}e^{-\frac{x}{\theta}}dx=-\Big/_0^x e^{-\frac{x}{\theta}}=-(e^{-\frac{x}{\theta}}-\underbrace{e^{-\frac{0}{\theta}}}_{1})$
$=1-e^{-\frac{x}{\theta}}$

$\boxed{\int f'e^f\,dx=e^f}$

Hav. $X>1000$

a) $L(\theta)=P(X>1000;\theta)=1-P(\leq 1000)=1-F(1000)=1-(1-e^{-\frac{x}{\theta}})=e^{-\frac{1000}{\theta}}$, $\theta>0$

$L(\theta)$ on kasvava, joten sillä ei ole maksimia.

Hav. $X=1000$

b) $L(\theta)=f(1000;\theta)=\frac{1}{\theta}e^{-\frac{1000}{\theta}}$

Maksimikohta:

$$L'(\theta)=-\frac{1}{\theta^2}e^{-\frac{1000}{\theta}}+\frac{1}{\theta}e^{-\frac{1000}{\theta}}\cdot\frac{1000}{\theta^2}$$
$$=\frac{1}{\theta^2}e^{-\frac{1000}{\theta}}\left(\frac{1000}{\theta}-1\right)=0 \quad \theta>0$$

Toteutuu, kun $\frac{1000}{\theta}-1=0$ eli $\frac{1000}{\theta}=1$ eli $\theta=1000$

Siis $\hat\theta=1000$

5.4. Vrt. Esim. 9.11 s. 260 Eksponenttijakauma (kts. tehtävä 5.3)

$f(x;\theta) = \frac{1}{\theta} e^{-\frac{x}{\theta}}$ $x \geq 0$

$F(x;\theta) = 1 - e^{-\frac{x}{\theta}}$ $x \geq 0$

minimin tiheysfunktio:

$f_{(1)}(x;\theta) = n\,[1-F(x;\theta)]^{n-1}\, f(x;\theta)$

$= 3\,[1-(1-e^{-\frac{x}{\theta}})]^{3-1} \cdot \frac{1}{\theta} e^{-\frac{x}{\theta}}$ $\theta > 0$

$= 3(e^{-\frac{x}{\theta}})^2 \cdot \frac{1}{\theta} e^{-\frac{x}{\theta}}$

Joten θ:n uskottavuusfunktio:

$L(\theta) = 3 \cdot e^{-\frac{2x}{\theta}} \cdot \frac{1}{\theta} e^{-\frac{x}{\theta}} = 3 \cdot \frac{1}{\theta} e^{-\frac{3x}{\theta}}$

$L(\theta)$ maksimoituu, kun $l(\theta)$ maksimoituu:

$l(\theta) = \log L(\theta) = \log 3 - \log\theta - \frac{3x}{\theta}$

$l'(\theta) = -\frac{1}{\theta} + \frac{3x}{\theta^2} = \frac{1}{\theta}\left(\frac{3x}{\theta} - 1\right)$

$l'(\theta) = 0$, kun $\frac{3x}{\theta} - 1 = 0$ eli $\frac{3x}{\theta} = 1$

$\theta = 3x$

eli $\hat{\theta} = 3x$ Tässä koe kesti 1000 aikayksikköä eli $x = 1000$

$\underline{\underline{\hat{\theta} = 3000}}$

5.5. Infarktien lkm $X \sim Bin(\theta, n)$

$f(x;n;\theta) = \binom{n}{x} \theta^x (1-\theta)^{n-x}$

$L(\theta) = \binom{n}{x} \theta^x (1-\theta)^{n-x}$

$\underline{\underline{\hat{\theta} = \frac{x}{n}}}$ $\boxed{\text{s. 273}}$

a) $x = 139$ $n_A = 11037$

$\hat{\theta}_A = \frac{139}{11037} \approx 0.01259$

b) $x = 239$ $n_L = 11034$

$\hat{\theta}_L = \frac{239}{11034} \approx 0.02166$

Kuvaajat:

```
curve(dbinom(139, 11037, x))
curve(dbinom(239, 11034, x))
```

$P(X \leq 139; \theta) = ?$

$n = n_A + n_L = 11037 + 11034$

$\hat{\theta} = \frac{n_A}{n}\hat{\theta}_A + \frac{n_L}{n}\hat{\theta}_L = \frac{139 + 239}{n_A + n_L}$ (vrt. Esim. 10.3)

$P(X \leq 139; \theta) = \text{pbinom}\left(139, 11037 + 11034, \frac{139+239}{11037+11034}\right) \approx 0{,}000,\ldots$

(5.6.) Merk. $\frac{n(A)}{n} = \frac{X_1 + \dots + X_n}{n} = \bar{X}_n$

$$E(X_i) = P(X_i = 1)\cdot 1 + P(X_i = 0)\,0 = P(X_i = 1) = P(A)$$
$$E(\bar{X}_n) = \frac{\sum E(X_i)}{n} = \frac{n P(A)}{n} = P(A)$$

$$Var(X_i) = P(A)[1 - P(A)] \quad i = 1, \dots, n$$
$$Var(\bar{X}_n) = \frac{P(A)[1-P(A)]}{n}$$

Lause 9.15 (HSLL) $P(|\bar{X}_n - \mu| \geq \varepsilon) \to 0$, kaikilla $\varepsilon > 0$ kun $n \to \infty$

Eli tässä Lauseen 9.15 nojalla

$$P\left(\left|\frac{n(A)}{n} - P(A)\right| \geq \varepsilon\right) \to 0 \text{ kaikilla } \varepsilon > 0 \text{ kun } n \to \infty$$

Määritelmän 9.3 mielessä tämä tarkoittaa, että

$\frac{n(A)}{n} \xrightarrow[n\to\infty]{P} P(A)$ eli $\frac{n(A)}{n}$ lähenee todennäköisyyden mielessä todennäköisyyttä $P(A)$.

(5.7.) Määritellään

$$Y_i = \begin{cases} 1, & \text{kun } X_i < x \\ 0, & \text{muutoin} \end{cases}$$

$n(x) = Y_1 + \dots + Y_n$ on niiden otosarvojen lkm, jotka $< x$

$$\frac{n(x)}{n} = \frac{Y_1 + \dots + Y_n}{n} = \bar{Y}_n$$

$$P(Y_i = 1) = P(X_i \leq x) = F(x)$$
$$Var(Y_i) = F(x)[1 - F(x)] \quad i = 1, \dots, n$$
$$Var(\bar{Y}_n) = \frac{var(Y_i)}{n}$$
$$E(\bar{Y}) = F(x)$$

Lause 9.15 (HSLL): $\bar{Y}_n = \frac{n(x)}{n} \xrightarrow{P} F(x)$ (Määritelmä 9.3)

(5.8.) Merkitään $A = \{X_1 \leq x_1, X_2 \leq x_2\}$, $B = \{X_2 \leq x_1, X_1 \leq x_2\}$

Silloin $P(A \cup B) = P(A) + P(B) - P(A \cap B)$, missä

$$A \cap B = \{X_1 \leq x_1, X_2 \leq x_1\}$$
$$P(A) = P(X_1 \leq x_1, X_2 \leq x_2) \overset{X_1 \perp\!\!\!\perp X_2}{=} P(X_1 \leq x_1)P(X_2 \leq x_2) = F(x_1)F(x_2)$$

Samalla tavalla

$$P(B) = F(x_2)F(x_1)$$
$$P(A \cap B) = P(X_1 \leq x_1, X_2 \leq x_1) = F(x_1)F(x_1) = [F(x_1)]^2$$

$$\therefore P(A \cup B) = P(X_{(1)} \leq x_1, X_{(2)} \leq x_2) = 2F(x_1)F(x_2) - [F(x_1)]^2$$