Tilastollinen päättely I Harj 7 vko 9 2012

7.1

Olk. $X_1,\dots,X_n$ otos jakaumasta $Ber(\theta)$

yhteisjakauman todennäköisyysfunktio

$$f(x_1,\dots,x_n;\theta) = \prod_{i=1}^{n} f(x_i;\theta) = \prod_{i=1}^{n} \theta^{x_i}(1-\theta)^{1-x_i} = \theta^{\sum_{i=1}^{n} x_i}(1-\theta)^{n-\sum_{i=1}^{n} x_i}, \quad x_i = 0 \text{ tai } 1$$

uskottavuusfunktio

$$L(\theta) = f(x_1,\dots,x_n;\theta) = \theta^{\sum_{i=1}^{n} x_i}(1-\theta)^{n-\sum_{i=1}^{n} x_i}$$

logaritmoitu uskottavuusfunktio

$$\ell(\theta) = \log L(\theta) = \sum_{i=1}^{n} x_i \log\theta + \left(n - \sum_{i=1}^{n} x_i\right)\log(1-\theta)$$

pistefunktio

$$S(\theta) = \ell'(\theta) = \frac{\sum_{i=1}^{n} x_i}{\theta} - \frac{n-\sum_{i=1}^{n} x_i}{1-\theta} = \frac{(1-\theta)\sum x_i - \theta(n-\sum x_i)}{\theta(1-\theta)}$$

SUE

$S(\theta)=0$, kun $\sum x_i - \theta\sum x_i - \theta n + \theta\sum x_i = 0$ eli $\theta = \frac{\sum x_i}{n} = \hat{\theta}$

Odotettu informaatio $\mathcal{I}(\theta) = E[I(\theta)] = E[-\ell''(\theta)]$

$$E[-\ell''(\theta)] = E\left[-\left(-\frac{\sum x_i}{\theta^2} - \frac{n-\sum x_i}{(1-\theta)^2}\right)\right] = E\left[\frac{\sum x_i}{\theta^2} + \frac{n-\sum x_i}{(1-\theta)^2}\right]$$

Nyt $\sum x_i \sim Bin(n,\theta)$
$E(\sum x_i) = n\theta$

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

Cramerin ja Raon alaraja $\quad Var(\hat{\theta}) \geq \frac{1}{\mathcal{I}(\theta)}$

$\sum x_i \sim Bin(n,\theta)$
$Var(\sum x_i) = n\theta(1-\theta)$

Nyt $Var(\hat{\theta}) = Var\left(\frac{\sum x_i}{n}\right) = \frac{1}{n^2} Var(\sum x_i) = \frac{n\theta(1-\theta)}{n^2}$

$= \frac{\theta(1-\theta)}{n} = \frac{1}{\mathcal{I}(\theta)}$, joten $Var(\hat{\theta})$ saavuttaa

Cramerin ja Raon alarajan $\longrightarrow$ SUE on tehokas.

7.2.

Olk. $X \sim Exp(\theta)$ $\quad f(x) = \frac{1}{\theta}e^{-\frac{x}{\theta}}$, $x>0$, $\theta>0$ $\quad E(X)=\theta$, $Var(X)=\theta^2$

a) pistefunktion varianssi:

$$S(\theta;x) = \ell'(\theta;x) = \frac{d}{d\theta}\log f(x;\theta) = \frac{d}{d\theta}\left[-\log\theta - \frac{x}{\theta}\right] = -\frac{1}{\theta} + \frac{x}{\theta^2}$$

$$Var(S(\theta;X)) = Var\left[-\frac{1}{\theta} + \frac{X}{\theta^2}\right] = Var\left(\frac{X}{\theta^2}\right) = \frac{1}{(\theta^2)^2}Var(X) = \frac{\theta^2}{(\theta^2)^2} = \underline{\underline{\frac{1}{\theta^2}}}$$

b) odotettu informaatio $\mathcal{I}(\theta) = E[I(\theta;x)] = E[-S'(\theta;x)]$

$$-S'(\theta;x) = -\left[\frac{1}{\theta^2} - \frac{2x}{\theta^3}\right] = -\frac{1}{\theta^2} + \frac{2x}{\theta^3}$$

$$\mathcal{I}(\theta) = E[-S'(\theta;X)] = E\left[-\frac{1}{\theta^2} + \frac{2X}{\theta^3}\right] = -\frac{1}{\theta^2} + \frac{2}{\theta^3}\overbrace{E(X)}^{\theta} = -\frac{1}{\theta^2} + \frac{2}{\theta^2} = \underline{\underline{\frac{1}{\theta^2}}}$$

(s.303)
(11.1.7)

c) $\mathcal{I}(\theta;x_1,\dots,x_n) = n\,\mathcal{I}(\theta;x) = \frac{n}{\theta^2}$

Cramerin ja Raon raja $\quad \frac{1}{\mathcal{I}(\theta;x_1,\dots,x_n)} = \frac{1}{n/\theta^2} = \frac{\theta^2}{n}$

$$Var(\bar{X}) = Var\left(\frac{\sum x_i}{n}\right) = \frac{1}{n^2}\underbrace{\sum Var(x_i)}_{\theta^2} = \frac{1}{n^2}\cdot n\cdot\theta^2 = \frac{\theta^2}{n}$$

$Var(\bar{X})$ saavuttaa C-R -rajan $\Longrightarrow$ $\bar{X}$ on minimivarianssinen

(7.3.) Val. n:n alkion otos jakaumasta, jonka tf $f(x;\theta)=\frac{1}{\theta^2}xe^{-\frac{x}{\theta}}$, $x>0$, $\theta>0$

$\boxed{\text{Gamma}(2,\theta) \quad E(X)=2\theta \quad Var(X)=2\theta^2}$

us. $L(\theta;x_1,\dots,x_n)=\prod_{i=1}^{n}\frac{1}{\theta^2}x_ie^{-\frac{x_i}{\theta}}=\frac{1}{\theta^{2n}}\cdot\Sigma x_i\cdot e^{-\frac{\Sigma x_i}{\theta}}$

log.uf. $\ell(\theta)=-2n\log\theta+\sum_{i=1}^{n}\log x_i-\frac{\Sigma x_i}{\theta}$

pistef. $S(\theta)=\ell'(\theta)=-\frac{2n}{\theta}+\frac{\Sigma x_i}{\theta^2}=0 \iff \frac{\Sigma x_i}{\theta}=2n \quad \theta>0$

$\therefore \hat{\theta}=\frac{\Sigma x_i}{2n}=\frac{\bar{X}}{2}$

Harhattomuus:

$E(\hat{\theta})=E\left(\frac{\Sigma x_i}{2n}\right)=\frac{1}{2n}\Sigma \underbrace{E(x_i)}^{2\theta}=\frac{1}{2n}n\cdot 2\theta=\theta \implies \hat{\theta}$ on harhaton

Cramerin ja Raon alaraja:

$-S'(\theta)=-\left[\frac{2n}{\theta^2}-\frac{2\Sigma x_i}{\theta^3}\right]=-\frac{2n}{\theta^2}+\frac{2\Sigma x_i}{\theta^3}=I(\theta)$

$J(\theta)=E[I(\theta)]=E\left[-\frac{2n}{\theta^2}+\frac{2\Sigma x_i}{\theta^3}\right]=-\frac{2n}{\theta^2}+\frac{2\Sigma \overbrace{E(x_i)}^{2\theta}}{\theta^3}$

$=-\frac{2n}{\theta^2}+\frac{2n\cdot 2\theta}{\theta^3}=\frac{4n}{\theta^2}-\frac{2n}{\theta^2}=\frac{2n}{\theta^2}$

$Var(\hat{\theta})=Var\left(\frac{\Sigma x_i}{2n}\right)=\frac{1}{(2n)^2}\Sigma \overbrace{Var(x_i)}^{2\theta^2}=\frac{1}{4n^2}n2\theta^2=\frac{\theta^2}{2n}$

Siis $Var(\hat{\theta})=\frac{1}{J(\theta)}$ (saavuttaa C-R alarajan)

(7.4.) Olk. $X\sim Bin(n,\pi)$ $\quad E(X)=n\pi \quad Var(X)=n\pi(1-\pi)$

s. 299 (11.1.1) a) $MSE(\hat{\pi})=E(\hat{\pi}-\pi)^2=Var(\hat{\pi})+[E(\hat{\pi})-\pi]^2$

$E(T_1)=E(X/n)=\frac{E(X)}{n}=\frac{n\pi}{n}=\pi$, $Var(T_1)=Var\left(\frac{X}{n}\right)=\frac{n\pi(1-\pi)}{n^2}=\frac{\pi(1-\pi)}{n}$

$MSE(T_1)=\frac{\pi(1-\pi)}{n}+[\pi-\pi]^2=\frac{\pi(1-\pi)}{n}$

$E(T_2)=E\left(\frac{X+1}{n+2}\right)=\frac{E(X)+1}{n+2}=\frac{n\pi+1}{n+2}$, $Var(T_2)=Var\left(\frac{X+1}{n+2}\right)=\frac{Var(X)}{(n+2)^2}=\frac{n\pi(1-\pi)}{(n+2)^2}$

$MSE(T_2)=\frac{n\pi(1-\pi)}{(n+2)^2}+\left[\frac{n\pi+1}{n+2}-\pi\right]^2=\frac{n\pi(1-\pi)}{(n+2)^2}+\left[\frac{n\pi+1}{n+2}-\frac{(n+2)\pi}{n+2}\right]^2$

$=\frac{n\pi(1-\pi)}{(n+2)^2}+\left[\frac{n\pi+1-n\pi-2\pi}{n+2}\right]^2=\frac{n\pi(1-\pi)+(1-2\pi)^2}{(n+2)^2}$

b) Kun $n=100$ ja $\pi=0.4$

$MSE(T_1)=\frac{0.4(1-0.4)}{100}=\frac{0.4\cdot 0.6}{100}\approx 0.0024$

$MSE(T_2)=\frac{100\cdot 0.4(1-0.4)+(1-2\cdot 0.4)^2}{(100+2)^2}=\frac{100\cdot 0.4\cdot 0.6+0.2^2}{102^2}\approx 0.0023$

$\therefore MSE(T_2)<MSE(T_1)$, kun $n=100$ ja $\pi=0.4$

Mitä pienempi keskineliövirhe, sitä lähempänä π:tä estimaattorin arvot ovat keskimäärin.

7.5. Otos $X_1,\dots,X_n$ $Exp(\theta)$ -jakaumasta, $n \geq 3$.

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

a) $E(\hat{\theta}_1) = E(X_1) = \theta$

$E(\hat{\theta}_2) = E\left(\frac{X_1+X_2}{2}\right) = \frac{E(X_1)+E(X_2)}{2} = \frac{\theta+\theta}{2} = \theta$

$E(\hat{\theta}_3) = E(\bar{X}) = \frac{1}{n}\sum_{i=1}^{n} E(X_i) = \frac{1}{n} n \theta = \theta$

$Var(\hat{\theta}_1) = Var(X_1) = \theta^2$

$Var(\hat{\theta}_2) = Var\left(\frac{X_1+X_2}{2}\right) = \frac{1}{4}\underbrace{(Var(X_1)+Var(X_2))}_{2\theta^2} = \frac{\theta^2}{2}$

$Var(\hat{\theta}_3) = Var\left(\frac{\sum X_i}{n}\right) = \frac{1}{n^2} n \cdot Var(x) = \frac{\theta^2}{n}$

$\hat{\theta}_4 = \min(X_1, X_2, X_3\}$ (Minimin jakauma 9.5.1)

s. 261 minimin tiheysfunktio on $f_{(1)}(x) = n[1-F(x)]^{n-1} f(x)$

siis $\hat{\theta}_4$:n tf on $f(x;\theta) = 3 \cdot e^{-\frac{2x}{\theta}} \cdot \frac{1}{\theta} e^{-\frac{x}{\theta}} = \frac{3}{\theta} e^{-\frac{3x}{\theta}}$ ts. $\hat{\theta}_4 \sim Exp\left(\frac{\theta}{3}\right)$

$E(\hat{\theta}_4) = \frac{\theta}{3}$ ja $Var(\hat{\theta}_4) = \frac{\theta^2}{9}$

b) Tehtävässä 2 laskettiin Cramerin ja Raon alaraja $= \frac{\theta^2}{n}$
Estimaattori $\hat{\theta}_3$ on tehokas (se on harhaton ja saavuttaa C-R alarajan)
$\hat{\theta}_4$:llä on pienin varianssi, kun $n<9$, mutta se ei ole harhaton.

c) SUE: $s(\theta) = \ell'(\theta) = -\frac{n}{\theta} + \frac{\sum x_i}{\theta^2} = 0$

$\frac{\sum x_i}{\theta^2} = \frac{n}{\theta} \iff \frac{\sum x_i}{\theta} = n \iff \hat{\theta} = \bar{X}$

$\therefore$ SUE $= \hat{\theta}_3$

7.6. $X_1,\dots,X_n$ otos gammajakaumasta $\Gamma(\alpha,\beta)$, α:n ja β:n estimaattorit momenttimenetelmällä:

$E(X) = \alpha\beta$ $Var(X) = \alpha\beta^2$

1. ja 2. otosmomentti

$m_1 = \frac{1}{n}\sum_{i=1}^{n} x_i = \bar{x}$

$m_2 = \frac{1}{n}\sum_{i=1}^{n} x_i^2$

1. ja 2. populaatiomomentti

$(\alpha_1) = E(X^1) = E(X) = \alpha\beta$

$(\alpha_2) = E(X^2) = Var(x) + E(X)^2 = \alpha\beta^2 + \alpha^2\beta^2$

Asetetaan populaation momentit ja otoksen momentit yhtäsuuriksi ja ratkaistaan α ja β.

$m_1 = \alpha_1$ eli $\bar{X} = \alpha\beta \Rightarrow \alpha = \frac{\bar{X}}{\beta}$ sij.

$m_2 = \alpha_2$ eli $\frac{1}{n}\sum X_i^2 = \alpha\beta^2 + \alpha^2\beta^2$

saadaan

$\frac{1}{n}\sum X_i^2 = \frac{\bar{X}}{\beta}\beta^2 + \frac{\bar{X}^2}{\beta^2}\beta^2 = \bar{X}\beta + \bar{X}^2$

$\bar{X}\beta = \frac{1}{n}\sum X_i^2 - \bar{X}^2 = \frac{1}{n}\left(\sum X_i^2 - n\bar{X}^2\right) = \frac{1}{n}\sum_{i=1}^{n}(X_i - \bar{X})^2 = S_b^2$ popul. varianssi

$\therefore \tilde{\beta} = \frac{S_b^2}{\bar{X}}$

$\tilde{\alpha} = \frac{\bar{X}}{\tilde{\beta}} = \frac{\bar{X}}{S_b^2/\bar{X}} = \frac{\bar{X}^2}{S_b^2}$

(7.7.) $X \sim Bin(n, \pi)$ $n > 1$, $0 < \pi < 1$ $E(X) = n\pi$

$T(X)$ on π^2:n estimaattori $Var(X) = n\pi(1-\pi)$

a) $T(X) = \left(\frac{X}{n}\right)^2 = \frac{X^2}{n^2}$

$Harha(T(x)) = E(T(x)) - \pi^2$

$E(T(x)) = E\left(\frac{X^2}{n^2}\right) = \frac{1}{n^2} E(X^2) = \frac{1}{n^2}\left[Var(X) + E(X)^2\right] = \frac{1}{n^2}\left[n\pi(1-\pi) + (n\pi)^2\right]$

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

$Harha(T) = E(T) - \pi^2 = \frac{\pi(1-\pi)}{n}$

b) $T_1(X) = \frac{1}{n-1}(nX - X^2)$

$E(T_1) = \frac{1}{n-1}\left[n E(X) - \overbrace{E(X^2)}^{Var(X)+[E(X)]^2}\right] = \frac{1}{n-1}\left[n \cdot n\pi - n\pi(1-\pi) - (n\pi)^2\right]$

$= \frac{n\pi}{n-1}\left[n - (1-\pi) - n\pi\right] = \frac{n\pi}{n-1}\left[n(1-\pi) - (1-\pi)\right]$

$= \frac{n\pi}{\cancel{n-1}}\left[\cancel{(n-1)}(1-\pi)\right] = n\pi(1-\pi) = Var(X)$ eli $T_1(X)$ on

varianssin harhaton estimaattori.

c) $E\left(T(X) - \frac{T_1(X)}{n}\right) = E(T(X)) - \frac{E(T_1(X))}{n} = \frac{\pi(1-\pi)}{n} + \pi^2 - \frac{\cancel{n}\pi(1-\pi)}{\cancel{n}} = \pi^2$

siis $T(X) - \frac{T_1(X)}{n}$ on π^2:n harhaton estimaattori.

(7.8.) $X_1, \ldots, X_n$ otos $Tas(0, \theta)$. $T = X_{(n)}$ ← hav. max on θ:n estimaattori.

$F(x) = \frac{x-0}{\theta-0} = \frac{x}{\theta}$, $f(x) = \frac{1}{\theta}$

Luku 9.5.1

s. 261 Maksimin tiheysfunktio $f(t) = n[F(t)]^{n-1} \cdot f(t) = n\left(\frac{t}{\theta}\right)^{n-1} \cdot \frac{1}{\theta} = n \cdot \frac{1}{\theta} \cdot \frac{1}{\theta^{n-1}} \cdot t^{n-1}$

$= \frac{n}{\theta^n} \cdot t^{n-1}$ $0 \le t \le \theta$

$E(T) = \int_0^\theta t \frac{n}{\theta^n} t^{n-1}\, dt = \frac{n}{\theta^n}\int_0^\theta t^n\, dt = \frac{n}{\theta^n} \Big/_0^\theta \frac{t^{n+1}}{n+1} = \frac{n}{\theta^n} \cdot \frac{\theta^{n+1}}{n+1} = \frac{n}{n+1}\theta^{n+1-n} = \frac{n}{n+1}\theta$

<θ koska $\frac{n}{n+1} < 1$

Harha: $E(T) - \theta = \frac{n}{n+1}\theta - \theta = \theta\left(\frac{n}{n+1} - 1\right) = \theta\left(\frac{n}{n+1} - \frac{n+1}{n+1}\right) = \theta\left(-\frac{1}{n+1}\right) = -\frac{\theta}{n+1}$

Tarkentuvuus:

$Var(T) = E(T^2) - E(T)^2$

$E(T^2) = \frac{n}{\theta^n}\int_0^\theta t^2 t^{n-1}\, dt = \frac{n}{\theta^n}\int_0^\theta t^{n+1}\, dt = \frac{n}{\theta^n} \Big/_0^\theta \frac{t^{n+2}}{n+2} = \frac{n}{\theta^n} \cdot \frac{\theta^{n+2}}{n+2} = \frac{n}{n+2}\theta^2$

$\therefore Var(T) = \frac{n}{n+2}\theta^2 - \left(\frac{n}{n+1}\right)^2\theta^2 = \theta^2\left[\frac{n}{n+2} - \left(\frac{n}{n+1}\right)^2\right] \to 0$, kun $n \to \infty$

ja

$harha(T) = -\frac{\theta}{n+1} \to 0$, kun $n \to \infty$ $\Rightarrow$ siis T on tarkentuva