Sannsynlighetsmaksimering

Anta at $X_1, X_2, \dots, X_n$ er et tilfeldig utvalg fra eksponentialfordeling med tetthet

$$f_X(x) = \lambda e^{-\lambda x}, \quad x > 0 \qquad (1)$$

- Finn SMEen $\hat{\lambda}$ av λ. Er $\hat{\lambda}$ forventningsrett?

Likelihoodfunksjonen (rimelighetsfunksjonen)

$$L(\lambda) = f_{X_1, X_2, \dots, X_n}(x_1, \dots, x_n)$$

$$= \prod_{i=1}^{n} f_X(x_i)$$

$$= \prod_{i=1}^{n} \lambda e^{-\lambda x_i}$$

$$= \lambda^n e^{-\lambda \sum_{i=1}^{n} x_i}$$

Log-likelihoodet blir

$$\ell(\lambda) = \ln L(\lambda) = n \ln \lambda - \lambda \sum_{i=1}^{n} x_i$$

Denne har sitt maksimum da

$$\frac{d\ell}{d\lambda} = 0$$

$$\frac{n}{\lambda} - \sum x_i = 0$$

$$\lambda = \frac{n}{\sum x_i}$$

SME av λ er dermed

$$\hat{\lambda} = \frac{n}{\sum X_i}.$$

Er $\hat{\lambda}$ forventningsrett. Siden $Y = \sum X_i \sim \text{Gamma}(n, \lambda)$ er

$$E(\hat{\lambda}) = E\left(\frac{n}{\sum X_i}\right) = E\Big(\underbrace{\frac{n}{Y}}_{= g(Y)}\Big)$$

$$= \int_0^\infty g(y) f_Y(y)\, dy$$

$$= \int_0^\infty \frac{n}{y} \frac{\lambda^n}{\Gamma(n)} y^{n-1} e^{-\lambda y}\, dy$$

$$= \frac{n\lambda\Gamma(n-1)}{\Gamma(n)} \underbrace{\int_0^\infty \underbrace{\frac{\lambda^{n-1}}{\Gamma(n-1)} y^{(n-1)-1} e^{-\lambda y}}_{\text{tetthet til Gamma}(n-1,\lambda)} dy}_{=1}$$

$\Gamma(n) = (n-1)! = (n-1)(n-2)! = (n-1)\Gamma(n-1)$

$\frac{\Gamma(n-1)}{\Gamma(n)} = \frac{1}{n-1}$)

$$= \frac{n}{n-1}\lambda \neq \lambda$$

m.a.o. er $\hat{\lambda}$ ikke forventningsrett. Derimot har den (bias-) korrigerte estimatoren

$$\lambda^* = \frac{n-1}{n}\hat{\lambda} = \frac{n-1}{n} \cdot \frac{n}{\Sigma X_i} = \frac{n-1}{\Sigma X_i}$$

forventning

$$E(\lambda^*) = \frac{n-1}{n} E(\hat{\lambda}) = \frac{n-1}{n} \cdot \frac{n}{n-1}\lambda = \lambda.$$

- Hva om vi i stedet bruker parameteriseringen

$$f_X(x) = \frac{1}{\beta} e^{-x/\beta} ? \qquad (2)$$

Sammenheng mellom λ og β i parameteriseringene (1) og (2) er da at

$$\lambda = \frac{1}{\beta} \Leftrightarrow \beta = h(\lambda) = \frac{1}{\lambda}$$

Hva blir SME av β og er SME av β forventningsrett?

Likelihood

$$L(\beta) = \dots = \frac{1}{\beta^n} e^{-\frac{1}{\beta}\sum_{i=1}^n x_i}$$

Log-likelihood

$$\ell(\beta) = -n\ln\beta - \frac{1}{\beta}\sum_{i=1}^n x_i$$

I maksimum er

$$\frac{\partial \ell}{\partial \beta} = 0$$

$$-\frac{n}{\beta} + \frac{1}{\beta^2}\sum x_i = 0$$

$$\beta = \frac{\sum x_i}{n}$$

d.v.s SME av β er

$$\hat{\beta} = \frac{\sum X_i}{n} = h(\hat{\lambda}) = \frac{1}{\hat{\lambda}} = \frac{1}{n/\sum x_i}$$

(slik funksjonal invarians til SME gjelder generelt)

Vi blir $\hat{\beta}$ (i motsetning til $\hat{\lambda}$) forventningsrett siden

$$E(\hat{\beta}) = E\left(\frac{\sum X_i}{n}\right) = \frac{1}{n}\sum EX_i = \frac{1}{n}\sum\frac{1}{\lambda} = \frac{1}{\lambda} = \beta$$

Siden $\hat{\lambda} = \frac{1}{\hat{\beta}}$ er en konveks funksjon av $\hat{\beta}$ blir

$$E(\hat{\lambda}) = E\left(\frac{1}{\hat{\beta}}\right) > \frac{1}{E\hat{\beta}} = \frac{1}{\beta} = \lambda$$

Fra forelesning (16. mars)

Gammafordeling

$$f_Y(y) = \frac{1}{\beta^\alpha \Gamma(\alpha)} y^{\alpha-1} e^{-\frac{y}{\beta}}, \quad y \geq 0$$

$E(Y) = \alpha\beta$

Kji-kvadrat fordeling ν frihetsgrader

spesialtilfelle $\alpha = \frac{\nu}{2}$, $\beta = 2$

$E(\chi^2_\nu) = \nu$

Teorem

Dersom $Y \sim \text{Gamma}(\alpha, \beta)$ er $W = \frac{2Y}{\beta} \sim$ Kji-kvadrat (2α)

Bevis

$$f_W(w) = f_Y(y(w)) \left| \frac{dy}{dw} \right| , \quad y = \frac{\beta w}{2} , \quad \frac{dy}{dw} = \frac{\beta}{2}$$

$$= \frac{1}{\beta^\alpha \Gamma(\alpha)} \left(\frac{\beta w}{2} \right)^{\alpha - 1} e^{-\left(\frac{\beta w}{2}\right)/\beta} \cdot \frac{\beta}{2}$$

$$= \frac{1}{\Gamma(\alpha) 2^\alpha} w^{\alpha-1} e^{-\frac{w}{2}}$$

$$= \frac{1}{\Gamma(\nu/2) 2^{\nu/2}} w^{\frac{\nu}{2}-1} e^{-\frac{w}{2}}$$

d.v.s. kji-kvadrat med $\nu = 2\alpha$ frihetsgrader.

- Konfidensintervall for β (fortsettelse av eksempel)

SME av β var $\hat{\beta} = \frac{\sum X_i}{n} = \frac{Y}{n}$

hvor $Y \sim \text{Gamma}(n, \beta)$. Har at

$$\left. \frac{2Y}{\beta} = \frac{2 \sum X_i}{\beta} = \frac{2n \sum X_i}{\beta n} = \frac{2n\hat{\beta}}{\beta} \right\} \text{pivotal størrelse}$$

er kji-kvadrat med $\nu = 2n$ frihetsgrader og

$$P\left(\chi^2_{1-\alpha/2, 2n} < \frac{2n\hat{\beta}}{\beta} < \chi^2_{\alpha/2, 2n} \right) = 1 - \alpha$$

$$P\left(\frac{1}{\chi^2_{1-\alpha/2, 2n}} > \frac{\beta}{2n\hat{\beta}} > \frac{1}{\chi^2_{\alpha/2, 2n}} \right) = 1 - \alpha$$

$$P\left(\frac{2n\hat{\beta}}{\chi^2_{1-\alpha/2,\,2n}} > \beta > \frac{2n\hat{\beta}}{\chi^2_{\alpha/2,\,2n}}\right) = 1-\alpha$$

slik at $(\hat{\beta}_L, \hat{\beta}_U)$

$$= \left(\frac{2n\hat{\beta}}{\chi^2_{\alpha/2,\,2n}},\ \frac{2n\hat{\beta}}{\chi_{1-\alpha/2,\,2n}}\right)$$

er et $(1-\alpha)$–konfidensintervall for β.

- Konfidensintervall for $\lambda = g(\beta) = \frac{1}{\beta}$:

$$P(\hat{\beta}_L < \beta < \hat{\beta}_U) = 1-\alpha$$

Når g er strengt avtagende blir

$$P(g(\hat{\beta}_U) < g(\beta) < g(\hat{\beta}_L)) = 1-\alpha$$

$$P\left(\frac{1}{\hat{\beta}_U} < \underset{\lambda}{\underset{\shortparallel}{\frac{1}{\beta}}} < \frac{1}{\hat{\beta}_L}\right) = 1-\alpha$$

d.v.s. $\left(g(\hat{\beta}_U), g(\hat{\beta}_L)\right)$

$$= \left(\frac{\chi^2_{1-\alpha/2,\,2n}\hat{\lambda}}{2n},\ \frac{\chi_{\alpha/2,\,2n}\hat{\lambda}}{2n}\right)$$

er et $(1-\alpha)$–konfidensintervall for λ.

Eksempel 2: Genetikk.

Populasjon bestående av tre genotyper
AA, Aa, aa med ukjente frekvenser

$$p_{AA} = p^2$$
$$p_{Aa} = 2p(1-p)$$
$$p_{aa} = (1-p)^2$$

Trekker utvalg på $n=100$ individer og observerer $x_{AA} = 2$, $X_{Aa} = 16$, $x_{aa} = 82$

av genotype AA, Aa, aa.
Finn SME av p (p = frekvens av genvarianten A i populasjon)

$$(X_{AA}, X_{Aa}, X_{aa}) \sim \text{Multinom}(n, p^2, 2p(1-p), (1-p)^2)$$

Dermed er likelihoodfunksjonen

$$L(p) = \frac{n!}{x_{AA}!\, x_{Aa}!\, x_{aa}!} (p^2)^{x_{AA}} (2p(1-p))^{x_{Aa}} ((1-p)^2)^{x_{aa}}$$

$$= C\, p^{\overbrace{2x_{AA} + x_{Aa}}^{= x_A}} (1-p)^{\overbrace{x_{Aa} + 2x_{aa}}^{x_a}},$$

log-likelihoodet

$$\ell(p) = \ln C + x_A \ln p + x_a \ln(1-p)$$

I maksimum er

$$\frac{d\ell}{dp} = 0$$

$$\frac{x_A}{p} - \frac{x_a}{1-p} = 0$$

$$x_A(1-p) - x_a p = 0$$

SME av p blir

$$\hat{p} = \frac{x_A}{x_A + x_a} = \frac{2x_{AA} + x_{Aa}}{2x_{AA} + 2x_{Aa} + 2x_{aa}} = \frac{2x_{AA} + x_{Aa}}{2n}$$

$$= \frac{2 \cdot 2 + 16}{2 \cdot 100} = \frac{20}{200} = 0.1.$$

Konfidensintervall for p:

Vi innser at $X_A \sim \text{bin}(2n, p)$

Dermed er

$$E(\hat{p}) = E\left(\frac{X_A}{2n}\right) = \frac{1}{2n} \cdot 2n \cdot p = p$$

og

$$Var(\hat{p}) = Var\left(\frac{X_A}{2n}\right) = \frac{1}{4n^2} Var(X_A)$$

$$= \frac{1}{4n^2} 2n\, p(1-p)$$

$$= \frac{p(1-p)}{2n}$$

og

$$P\left(-z_{\alpha/2} < \underbrace{\frac{\hat{p} - p}{\sqrt{\hat{p}(1-\hat{p})/(2n)}}}_{\approx N(0,1)} < z_{\alpha/2}\right) \approx 1 - \alpha$$

$$P\left(\hat{p} - z_{\alpha/2}\sqrt{} < p < \hat{p} + z_{\alpha/2}\sqrt{}\right) \approx 1-\alpha$$

slik at

$$\hat{p} \pm z_{\alpha/2}\sqrt{\hat{p}(1-\hat{p})/(2n)} = 0.1 \pm 1.96$$

$$= (0.058,\ 0.142)$$

er et $(1-\alpha)^2$ konfidensintervall for $\hat{p}$

Konfidens-intervall for $P_{AA} = g(p) = p^2$ blir

$$(0.058^2,\ 0.142^2)$$

$$= (0.0034,\ 0.020)$$