7.2 Momentestimator og Maksimumlikelihood estimator

Moment. $X_i \sim f(x; \bar{\theta})$, $\bar{\theta} = (\theta_1, \dots, \theta_m)$

La $h_j(\bar{\theta}) = E X_i^j$, $j = 1, \dots, m$

Estimerer $\bar{\theta}$ ved å løse likninger

$$h_j(\bar{\theta}) = \frac{1}{n}\sum_{i=1}^{n} X_i^j \quad , \quad j = 1, \dots, m$$

$$\bar{\theta} = \hat{\underset{\sim}{\theta}}$$

Løsningen $\hat{\underset{\sim}{\theta}} = (\hat{\underset{\sim}{\theta}}_1, \dots, \hat{\underset{\sim}{\theta}}_m)$ kalles momentestimator

Så at hvis $X_i \sim N(\mu, \sigma^2)$ ble momentest.

$$\mu = \bar{x} \quad \text{og} \quad \hat{\sigma}^3 = \frac{1}{n}\sum (x_i - \bar{x})^3$$

Videre hvis $X_i \sim \text{gamma}(\alpha, \beta)$ blir m. est.

$$\hat{\alpha} = \dots$$

$\vdots$

$\vdots$

Eks: $X_i \sim U[0, \theta]$ uif, $i = 1, \dots, n$

$E(X_i) = \frac{\theta}{2} \Rightarrow$ Moment. est $\hat{\theta} = 2\bar{X}$, forventningsrett

$Var(\hat{\theta}) = 4\, Var(\bar{X}) = 4 \frac{\theta^2}{12n} = \frac{\theta^2}{3n} \to 0$ når $n \to \infty$

$= O(\frac{1}{n})$

Men, som utledet i STK1100, gjelder for

$M = \max(X_1, X_2, \dots, X_n)$

at. $E(M) = \frac{n}{n+1}\theta$ og $V(M) = \frac{n}{(n+2)(n+1)^2}\theta^2$

så $\theta^* = \frac{n+1}{n} M$ er forventningsrett med

$V(\theta^*) = \frac{1}{n(n+2)}\theta^2 \to 0$ som $\frac{1}{n^2}$ når $n \to \infty$

$= O(\frac{1}{n^2})$

Så θ^* er en bedre estimator enn $\hat{\theta}$

Det gjelder ganske generelt at moment. est.

- Konsistente $\hat{\theta} \to \theta$, $n \to \infty$
- Tilnærmet normalt fordelte

Men de er ikke "effisiente", gir ikke nødvendigvis ikke estimatorer eksakt med minst mulig varians

Da har parvise av transformasjoner av X-er

f. eks. $Y_1 = \ln(X_1)$

5 ulike egenskaper tilfredsstilles av

Maximum Likelihood estimator (MLE)

Anta at $\bar{X} = (X_1, \dots, X_n)$ har en simultanfordeling

$f(x_1, \dots, x_n; \theta)$, $\bar{\theta} = (\theta_1, \dots, \theta_n)$

Vi observerer $X_1 = x_1$, $X_n = x_n$ og setter disse inn i

$f(x_1, \dots, x_n, \bar{\theta})$. Da er "L" funksjonen definert som

$$L(\bar{\theta}) = f(x_1, \dots, x_n; \theta) \text{ som funksjon av } \bar{\theta}$$

MLE $\hat{\theta}$ er den verdi som maksimerer $L(\bar{\theta})$.

man tenker ofte på MLE $\hat{\theta}$ som den verdien

av $\bar{\theta}$ som gir det ~~mest~~ mest mulig ...

Fordelingen er lik hverandre når X_i-ene er diskrete

Nesten alltid: dette huset er X_i-ene uavhengige

i Kap 6-9 er de gjennomgående uif

Så $X_i \sim f(x, \bar{\theta})$ og $(X_1, \dots, X_n)$

Simultan tetthet

$$f_n(x_1, \dots, x_n, \bar{\theta}) = \prod_{i=1}^{n} f(x_i; \bar{\theta})$$

For x_i uif $f(x_i\theta)$ blir L

$$L(\theta) = \prod_{i=1}^{n} f(x_i, \theta)$$

Det er vanligvis bekvemt å jobbe med log (L)

$$l(\vec{\theta}) = \ln L(\vec{\theta}) = \sum_{i=1}^{n} \ln f(x_i, \theta)$$

Maksimering av $l(\theta)$ gir også MLE $\hat{\theta}$ siden $\ln(u)$ er en voksende funksjon

Eks.

$$x_i = \begin{cases} 1, & p \\ 0, & 1-p \end{cases} \quad \text{uavhengig}$$

Da er $f(x;p) = p^x(1-p)^{1-x}$, $x = 0, 1$

Dette gir L

$$L(p) = \prod_{i=1}^{n} f(x_i, p) = \prod_{i=1}^{n} p^{x_i}(1-p)^{1-x_i}$$

$$= p^{\sum x_i}(1-p)^{\sum(1-x_i)} = p^y(1-p)^{n-y}$$

der $y = \sum_{i=1}^{n} x_i$

Dette gir log-L

$$l(p) = y\ln(p) + (n-y)\ln(1-p)$$

og finner MLE $\hat{p}$ ved å løse lig.

$$l'(p) = \frac{y}{p} - \frac{n-y}{1-p} = 0 \Rightarrow \hat{p} = \frac{y}{n} = \text{andelen } x_i = 1$$

Har vet vi $E(\hat{p}) = p$, $V(\hat{p}) = \frac{p(1-p)}{n}$, $S_e(\hat{p}) = \sqrt{\frac{\hat{p}(1-\hat{p})}{n}}$

Vi kunne alternativt basert oss på y direkte

$Y \sim Bin(n, p)$, $P(Y = y) = \binom{n}{y} p^y (1-p)^{n-y}$

$L(p) = \binom{n}{y} p^y (1-p)^{n-y}$

$$l(p) = \ln\left(\binom{n}{y}\right) + y\ln(p) + (n-y)\ln(1-p)$$

$$\Rightarrow l'(p) = \frac{y}{p} - \frac{n-y}{1-p} = 0 \Rightarrow \hat{p} = \frac{y}{n}$$

Ex) $X_i \sim N(\mu, 1)$ uavhengig, $\sigma^2 = 1$

$$f(x_i; \mu) = \frac{1}{\sqrt{2\pi}} e^{\frac{1}{2}(x-\mu)^2}$$

$$\Rightarrow L(\mu) = \prod_{i=1}^{n} f(x_i; \mu) = (2\pi)^{-n/2} e^{-\frac{1}{2}\sum(x_i-\mu)^2}$$

$$\Rightarrow l(\mu) = \underbrace{-\frac{n}{2}\ln 2\pi}_{\text{Konstant}} - \frac{1}{2}\sum(x_i-\mu)^2$$

$$l'(\mu) = -\frac{1}{2}\sum -2(x_i-\mu) = \sum(x_i-\mu)\Big|_{\mu=\hat{\mu}} = 0$$

$$\Rightarrow \hat{\mu} = \bar{x} = \frac{1}{n}\sum_{i=1}^{n} x_i$$

Generaliser må til $x_i \sim N(\mu, \sigma^2)$ analysis

Setter $\Psi = \sigma^2$ s. a. $\underline{\theta} = (\mu, \Psi)$

$$f(x ; \underline{\theta}) = \frac{1}{\sqrt{2\pi}\sqrt{\Psi}} e^{-\frac{1}{2\Psi}(x-\mu)^2}$$

$$L(\mu, \Psi) = 2\pi^{-n/2} \Psi^{-n/2} e^{-(1/2\Psi)\sum_{i=1}^{n}(x_i - \mu)^2}$$

$$l(\mu, \Psi) = -\frac{n}{2}\ln(2\pi) - \frac{n}{2}\ln\Psi - \frac{1}{2\Psi^2}\sum(x_i - \mu)^2$$

$$\frac{\partial l}{\partial \mu} = \frac{1}{\Psi}\sum_{i=1}^{n}(x_i - \mu) = 0 \Rightarrow \hat{\mu} = \bar{x} \quad \text{uavhengig } \Psi$$

$$\frac{\partial l}{\partial \Psi} = \frac{n}{2}\frac{1}{\Psi} + \frac{1}{2\Psi^2}\sum_{i=1}^{n}(x_i - \mu)^2 \Big| = 0 \Rightarrow \mu^2 = \bar{x}$$

$$\hat{\Psi} = \frac{1}{n}\sum(x_i - \bar{x})^2$$

$$= \tilde{\sigma}^2$$

Så MLE (μ, σ^2) blir $(\bar{x}, \tilde{\sigma}^2)$,

$x_i \sim N(\mu, 1)$ $L(\mu) = (2\pi)^{-1/2} \sum(x_i - \mu)^2$

$$l(\mu) = -\frac{n}{2}\ln(2\pi) - \frac{1}{2}\sum(x_i - \mu)^2$$

$= \text{Parabel} : \mu$

Ex) $x_i \sim \text{Gamma}(\alpha, \beta)$ uavh

$$f(x_i; \alpha, \beta) = \frac{x^{\alpha-1}}{\beta^{\alpha}\Gamma(\alpha)} e^{-\frac{x}{\beta}}$$

$$= l(\alpha, \beta) = \sum_{i=1}^{n} \ln f(x_i; \alpha, \beta)$$

$$= \sum_{i=1}^{n}\left[(\alpha - 1)\ln(x_i) - \alpha\ln\beta - \ln\Gamma\alpha - \frac{x_i}{\beta}\right]$$

$$\frac{\partial l(\alpha,\beta)}{\partial \beta} = -\frac{n\alpha}{\beta} + \frac{1}{\beta^2}\sum_{i=1}^{n} x_i \quad , \quad \text{Gitt } \alpha \text{ blir da } \hat\beta = \frac{\bar{x}}{\hat\alpha}$$

$$\frac{\partial l(\alpha,\beta)}{\partial \alpha} = \sum_{i=1}^{n}\left[\ln(x_i) - \ln(\beta) - \frac{\Gamma'(\alpha)}{\Gamma(\alpha)}\right]$$

$$= \sum_{i=1}^{n}\left[\ln(x_i) - n\ln(\beta) - n\frac{\Gamma'(\alpha)}{\Gamma(\alpha)}\right]$$

og vi finner ingen analytisk løsning på ligningene

Må å løse (α, β) ved å maksimere $l(\alpha, \beta)$ numerisk

f. eks. ved optim i R (nlm, nlminb)

Ex) $x_i \sim U[0, \theta]$

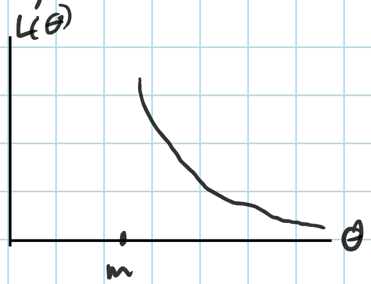
$$f(x;\theta) = \begin{cases} 1/\theta & 0 \le x \le \theta \\ 0 & \text{ellers} \end{cases} = \frac{1}{\theta} I(0 \le x \le \theta)$$

$$\Rightarrow L(\theta) = \prod_{i=1}^{n} \frac{1}{\theta} I(0 \le x_i \le \theta)$$

$$= \frac{1}{\theta^n} I(0 \le \min(x_i) < \text{maks } x_i \le \theta)$$

Ser at hvis $\theta < m$ så er $L(\theta) = 0$

og hvis $\theta > m$ så er $L(\theta) = \frac{1}{\theta^n}$



MLE $\hat\theta = M = \max(X_1, \dots, X_n)$