7.2 Estimeringsmetoder

Data (observasjoner) $x_1, \ldots, x_n$ betraktes som realiseringer av stok. var. $X_1, \ldots, X_n$ uif $\sim f(\cdot\,; \theta)$

θ kan være en vektor med m elementer

Vil finne en estimator $\hat{\theta}(X_1, \ldots, X_n)$

Får til slutt et estimat $\hat{\theta}(x_1, \ldots, x_n)$

Momentmetoden:

Finn teoretiske momenter $E(X_i^k)$ og sett (ik

empiriske momenter $\frac{1}{n}\sum_{i=1}^{n} X_i^k$

vil være funksjoner av θ

Løser ut θ fra ligningene

Eks. gamma(α, β)

$$\frac{1}{n}\sum X_i = \alpha\beta$$

$$\frac{1}{n}\sum X_i^2 = \alpha\beta^2(1+\alpha)$$

to ligninger med to ukjente α, β

Annen metode for å finne estimatorer:

Maximum likelihood estimering

Sannsynlighetsmaksimeringsestimering

Konsept - eksempel (eks. 7.16 i boken)

Passord: setning Hvor vanlig er det?

$p = P(\text{tilfeldig student bruker setning})$

10 studenter (uavhengige)

$$X_i = \begin{cases} 1 & \text{hvis student bruker setning} \\ 0 & \text{ellers} \end{cases}$$

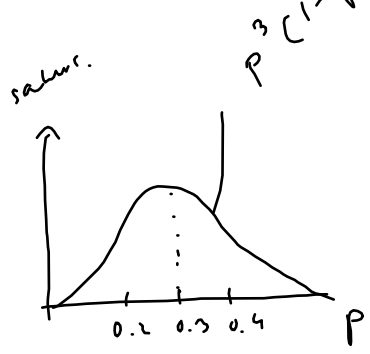
$P(X_i = 1) = p$

$P(X_i = 0) = 1 - p$

$x_1 = x_3 = x_{10} = 1$, alle andre $= 0$

$P(\text{observasjonene våre}) = p \cdot (1-p)\, p\, (1-p) \cdots p$

$= p^3 (1-p)^7$



For hvilken verdi av p er det utfallet (observasjonene) vi fikk, mest sannsynlige, mest rimelige?

Finner p som maksimerer sannsynligheten $p^3(1-p)^7$, som er ekvivalent med å maksimere $\ln(p^3(1-p)^7)$

$= 3\ln p + 7\ln(1-p)$

$$\frac{d}{dp}\left[\ln\left(p^3(1-p)^7\right)\right] = \frac{3}{p} - \frac{7}{1-p} = 0 \Rightarrow p = 0.3$$

Generisk versjon – maximum likelihood for p i Bernoulli(p)

Vi har observert data $x_1, \dots, x_n$

Disse representerer $X_1, \dots, X_n$ u.i.f. $\sim$ Bernoulli(p)

Sannsynligheten for å observere akkurat $x_1, \dots, x_n$ er da

$$L(p) = f(x_1, \dots, x_n; p) = P(X_1 = x_1, X_2 = x_2, \dots, X_n = x_n)$$

$L(p)$: likelihood-funksjonen

f: simultan-tetthet

$$\overset{\text{u.i.f.}}{=} \prod_{i=1}^{n} P(X_i = x_i)$$

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

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

$y = \sum x_i$

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

Vil finne den verdien av p som maksimerer

$L(p) = f(x_1, \dots, x_n; p) = p^y (1-p)^{n-y}$ m.h.p. p

Samme som å maksimere

$$l(p) = \ln L(p) = \ln f(x_1, \dots, x_n; p) = y \ln p + (n-y) \ln(1-p)$$

$l(p)$: log-likelihood-funksjonen

Deriverer $l(p)$ mhp. p og setter lik 0

Løser ut $p = \frac{y}{n}$; derfor $\hat{p} = \frac{Y}{n} = \frac{1}{n} \sum_{i=1}^{n} X_i$

MLE for p

MLE eksponensial fordeling

$X_1 \dots X_n$ u.i.f. $\sim \exp(\lambda)$

$x_1 \dots x_n$ observasjoner

Tetthet $f(x_i;\lambda) = \lambda e^{-\lambda x_i}$

$$L(\lambda) = f(x_1 \dots x_n;\lambda) \overset{\text{uavh.}}{=} \prod_{i=1}^{n} f(x_i;\lambda)$$

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

$$= \lambda^n e^{-\lambda \Sigma x_i}$$

$$\boxed{\ell(\lambda) = \ln L(\lambda) = n \ln \lambda - \lambda \Sigma x_i}$$

Vil finne den λ som maksimerer log likelihood:

$$\frac{d}{d\lambda}\ell(\lambda) = \frac{n}{\lambda} - \Sigma x_i = 0$$

$$\Rightarrow \quad \lambda = \frac{n}{\Sigma x_i} = \frac{1}{\bar{x}}$$

Så MLE for λ blir

$$\hat{\lambda} = \frac{1}{\bar{X}}$$

MLE normalfordeling

$x_1 \dots x_n$ obs.

$X_1 \dots X_n$ u.i.f. $\sim N(\mu, \sigma^2)$

$$L(\mu, \sigma^2) = f(x_1 \dots x_n \mid \mu, \sigma^2)$$

uavh.
$$= \prod_{i=1}^{n} f(x_i \mid \mu, \sigma^2)$$

$$= \prod_{i=1}^{n} \frac{1}{\sqrt{2\pi\sigma^2}} e^{-\frac{1}{2\sigma^2}(x_i-\mu)^2}$$

$$= (2\pi)^{-n/2} (\sigma^2)^{-n/2} \cdot e^{-\frac{1}{2\sigma^2}\sum (x_i-\mu)^2}$$

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

$$\frac{\partial l}{\partial \mu} = -\frac{2}{2\sigma^2}\sum (x_i-\mu)(-1) = \frac{1}{\sigma^2}\left(\sum x_i - n\mu\right) = 0$$

$$\Rightarrow \mu = \frac{1}{n}\sum x_i = \bar{x}$$

$$\frac{\partial l}{\partial \sigma^2} = -\frac{n}{2}\frac{1}{\sigma^2} + \frac{1}{2\sigma^4}\sum (x_i-\mu)^2 = 0$$

$$\sum (x_i-\mu)^2 = n\sigma^2 \quad \Rightarrow \quad \sigma^2 = \frac{1}{n}\sum (x_i-\mu)^2$$

Så MLE blir:

$$\hat{\mu} = \bar{X}$$
$$\hat{\sigma}^2 = \frac{1}{n}\sum_{i=1}^{n} (X_i - \bar{X})^2$$

NB!

$\hat{\sigma}^2 \neq S^2$, dvs. ikke forventningsrett (for små n)

MLE Gamma(α, β)

$$L(\alpha,\beta) = f(x_1,\dots,x_n;\alpha,\beta) \overset{\text{uavh.}}{=} \prod f(x_i;\alpha,\beta)$$

$$= \prod_{i=1}^{n} \frac{1}{\beta^\alpha \Gamma(\alpha)} x_i^{\alpha-1} e^{-\frac{x_i}{\beta}}$$

Kan ta ln, og partiellderivere mhp. α, β, sette lik 0, men vi klarer ikke å finne eksplisitte uttrykk for estimatorene.

Kan bruke numerisk optimering

Invariansprinsippet for MLE

- Vil finne MLE for $\varphi = \frac{1}{\lambda}$

  Har $\hat{\lambda} = \frac{1}{\bar{X}}$ er MLE for λ

  Har da at $\hat{\varphi} = \frac{1}{\hat{\lambda}} = \frac{1}{1/\bar{X}} = \bar{X}$

  er MLE for $\frac{1}{\lambda}$

- Vil finne MLE for $\varphi = \sigma = \sqrt{\sigma^2}$

  Har at $\widehat{\sigma^2} = \frac{1}{n}\Sigma(X_i - \bar{X})^2$ er MLE for σ^2

  Har da at $\hat{\varphi} = \sqrt{\widehat{\sigma^2}} = \sqrt{\frac{1}{n}\Sigma(X_i - \bar{X})^2}$

  er MLE for σ

MLE uniform fordeling

$X_1 \ldots X_n$ u.i.f. Uniform $[0, \theta]$

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

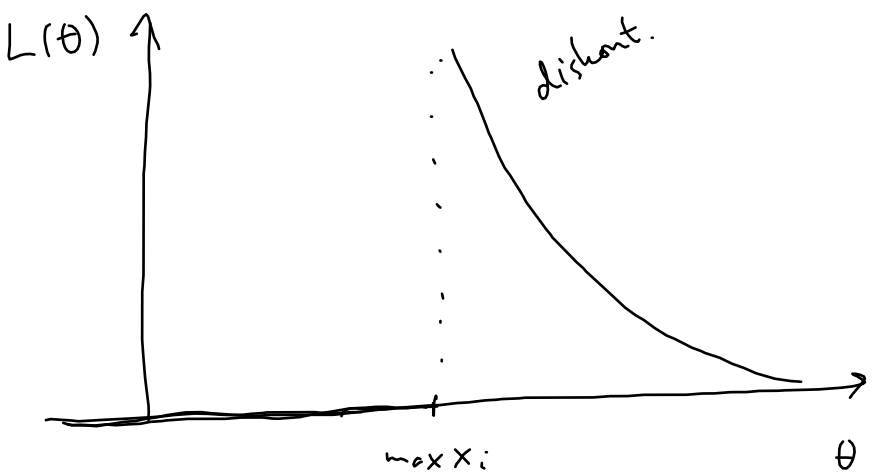
Hva er MLE for θ?

Studente tidligere to estimatorer for θ,

$\theta^* = 2\bar{X}$ og $\hat{\theta} = \frac{n+1}{n} \max X_i$

Finner likelihood

$$L(\theta) = f(x_1, \ldots x_n; \theta) = \begin{cases} \frac{1}{\theta^n} & 0 \le x_1 \le \theta,\ 0 \le x_2 \le \theta, \ldots 0 \le x_n \le \theta \\ 0 & \text{ellers} \end{cases}$$



$$\hat{\theta}_{MLE} = \max X_i$$

ikke forv. rett.