Maksimum likelihood-estimering

- Eksempel: $X_1, \dots, X_n \overset{uif}{\sim}$ Eksponensiell(λ)

Pga uavhengighet har vi:

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

$$= \prod_{i=1}^n \lambda e^{-\lambda x_i} = \lambda^n e^{\sum_{i=1}^n (-\lambda x_i)} = \underline{\lambda^n e^{-\lambda \sum_{i=1}^n x_i}}$$

Log-likelihooden er da:

$$\log\left(f(x_1, \dots, x_n; \lambda)\right) = \log\left(\lambda^n e^{-\lambda \sum_{i=1}^n x_i}\right) = \log(\lambda^n) + \log\left(e^{-\lambda \sum_{i=1}^n x_i}\right)$$

$$= \underline{n \log(\lambda) - \lambda \sum_{i=1}^n x_i}$$

Vi deriverer og setter lik 0:

$$\frac{d}{d\lambda} \log\left(f(x_1, \dots, x_n; \lambda)\right) = n \cdot \frac{1}{\lambda} - \sum_{i=1}^n x_i = 0$$

$$\longrightarrow \frac{n}{\lambda} = \sum_{i=1}^n x_i$$

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

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

• Eksempel: $X_1, \dots, X_n \overset{uif}{\sim} N(\mu, \sigma^2)$

Da er likelihood-funksjonen:

$$f(x_1,\dots,x_n;\mu,\sigma^2) = \prod_{i=1}^{n} f(x_i;\mu,\sigma^2)$$

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

$$= \frac{1}{(2\pi)^{n/2}\underbrace{\sigma^n}_{(\sigma^2)^{n/2}}} e^{-\frac{1}{2\sigma^2}\sum_{i=1}^{n}(x_i-\mu)^2}$$

$$\longrightarrow \log\left(f(x_1,\dots,x_n;\mu,\sigma^2)\right) = -\frac{n}{2}\log(2\pi) - \frac{n}{2}\log(\sigma^2) - \frac{1}{2\sigma^2}\sum_{i=1}^{n}(x_i-\mu)^2$$

$$\longrightarrow \frac{\partial}{\partial\mu}\log\left(f(x_1,\dots,x_n;\mu,\sigma^2)\right) = \frac{\partial}{\partial\mu}\left(-\frac{n}{2}\log(2\pi) - \frac{n}{2}\log(\sigma^2) - \frac{1}{2\sigma^2}\sum_{i=1}^{n}(x_i-\mu)^2\right)$$

$$= \cancel{-}\frac{1}{\cancel{2}\sigma^2}\sum_{i=1}^{n}\cancel{2}\cdot(x_i-\mu)\cdot\underbrace{\frac{\partial}{\partial\mu}(x_i-\mu)}_{=\cancel{-}1}$$

$$= \frac{1}{\sigma^2}\sum_{i=1}^{n}(x_i-\mu) = 0$$

$$\frac{\partial}{\partial\sigma^2}\log\left(f(x_1,\dots,x_n;\mu,\sigma^2)\right) = \frac{\partial}{\partial\sigma^2}\left(-\frac{n}{2}\log(2\pi) - \frac{n}{2}\log(\sigma^2) - \frac{1}{2\sigma^2}\sum_{i=1}^{n}(x_i-\mu)^2\right)$$

$$= -\frac{n}{2}\cdot\frac{1}{\sigma^2} + \frac{1}{2}\sum_{i=1}^{n}(x_i-\mu)^2\left(+\frac{1}{(\sigma^2)^2}\right)$$

$$= -\frac{n}{2\sigma^2} + \frac{1}{2\sigma^4}\sum_{i=1}^{n}(x_i-\mu)^2 = 0$$

$$\frac{1}{\sigma^2}\sum_{i=1}^{n}(x_i-\mu) = 0 \longrightarrow \sum_{i=1}^{n}x_i - n\mu = 0 \longrightarrow \sum_{i=1}^{n}x_i = n\mu$$

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

$$-\frac{n}{2\sigma^2} + \frac{1}{2\sigma^4}\sum_{i=1}^{n}(x_i-\mu)^2 = 0 \longrightarrow n\sigma^2 = \sum_{i=1}^{n}(x_i-\mu)^2$$

$$\longrightarrow \sigma^2 = \frac{1}{n}\sum_{i=1}^{n}(x_i-\mu)^2 = \frac{1}{n}\sum_{i=1}^{n}(x_i-\bar{x})^2$$

Det betyr at MLE for μ og σ^2 er

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

- Eksempel: $X_1, \dots, X_n \overset{uif}{\sim} \text{Gamma}(\alpha, \beta)$

Vi får:

$$f(x_1,\dots,x_n;\alpha,\beta) = \prod_{i=1}^{n} f(x_i;\alpha,\beta)$$

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

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

$$\rightarrow \log\left(f(x_1,\dots,x_n;\alpha,\beta)\right) = -n\alpha\log(\beta) - n\log(\Gamma(\alpha)) + \sum_{i=1}^{n} \underbrace{\log(x_i^{\alpha-1})}_{(\alpha-1)\log(x_i)} - \frac{1}{\beta}\sum_{i=1}^{n} x_i$$

$$= -n\alpha\log(\beta) - n\log(\Gamma(\alpha)) + (\alpha-1)\sum_{i=1}^{n}\log(x_i) - \frac{1}{\beta}\sum_{i=1}^{n} x_i$$

$$\rightarrow \frac{\partial}{\partial\alpha}\log\left(f(x_1,\dots,x_n;\alpha,\beta)\right) = -n\log(\beta) - n\underbrace{\frac{\Gamma'(\alpha)}{\Gamma(\alpha)}}_{\text{digammafunksjonen } \psi(\alpha)} + \sum_{i=1}^{n}\log(x_i)$$

$$\rightarrow \frac{\partial}{\partial\beta}\log\left(f(x_1,\dots,x_n;\alpha,\beta)\right) = -\frac{n\alpha}{\beta} + \frac{1}{\beta^2}\sum_{i=1}^{n} x_i$$

Vi får:

$$-\frac{n\hat{\alpha}}{\hat{\beta}} + \frac{1}{\hat{\beta}^2}\sum_{i=1}^{n} x_i = 0 \longrightarrow \frac{1}{\hat{\beta}^2}\sum_{i=1}^{n} x_i = \frac{n\hat{\alpha}}{\hat{\beta}}$$

$$\longrightarrow \sum_{i=1}^{n} x_i = n\hat{\alpha}\hat{\beta}$$

$$\longrightarrow \hat{\beta} = \frac{\sum_{i=1}^{n} x_i}{n\hat{\alpha}} = \frac{1}{\hat{\alpha}} \cdot \underbrace{\frac{1}{n}\sum_{i=1}^{n} x_i}_{\bar{x}} = \frac{\bar{x}}{\hat{\alpha}}$$

$$-n\log(\hat{\beta}) - n\psi(\hat{\alpha}) + \sum_{i=1}^{n}\log(x_i) = 0$$

$$\rightarrow -n\log\left(\frac{\bar{x}}{\hat{\alpha}}\right) - n\psi(\hat{\alpha}) + \sum_{i=1}^{n}\log(x_i) = 0$$

$$\rightarrow n\log(\hat{\alpha}) - n\log(\bar{x}) - n\psi(\hat{\alpha}) + \sum_{i=1}^{n}\log(x_i) = 0$$

Denne ligningen må løses numerisk for å finne $\hat{\alpha}$.

MLE for en transformasjon $\phi(\theta_1, \dots, \theta_m)$

- Eksempel : $X_1, \dots, X_n \overset{uif}{\sim}$ Eksponensiell (λ)

  Vi så at $\hat{\lambda} = \frac{1}{\overline{X}}$ er MLE for λ. Da er

  MLE for $\phi = \frac{1}{\lambda} = E(X) = \phi(\lambda)$ $\quad \hat{\phi} = \phi(\hat{\lambda}) = \frac{1}{\hat{\lambda}} = \frac{1}{1/\overline{X}} = \overline{X}$.

- Eksempel : $X_1, \dots, X_n \overset{uif}{\sim} N(\mu, \sigma^2)$

  Vi vet at $\hat{\mu} = \overline{X}$ og $\hat{\sigma}^2 = \frac{1}{n}\sum_{i=1}^{n}(X_i - \overline{X})^2$ er MLE for μ og σ^2. Hva er så MLE for standardavvik σ?

  Vi har :

  $$\sigma = \sigma(\mu, \sigma^2) = \sqrt{\sigma^2}$$

  Da er MLE for σ gitt ved :

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

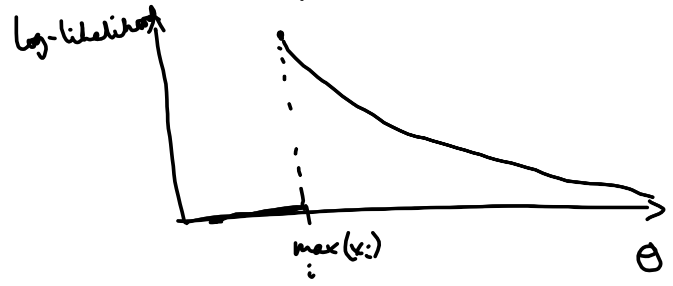
- Eksempel: $X_1, \dots, X_n \overset{\text{iid}}{\sim} U[0, \theta]$

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

slik

$$f(x_1, \dots, x_n; \theta) = \prod_{i=1}^{n} f(x_i; \theta) = \begin{cases} \frac{1}{\theta^n}, & 0 \le x_1, \dots, x_n \le \theta \\ 0, & \text{ellers} \end{cases}.$$

Da er likelihooden $1/\theta^n$ så lenge $\max_i(x_i) \le \theta$, men blir 0 straks $\theta < \max_i(x_i)$. Det er altså en diskontinuitet i log-likelihooden og den befinner seg akkurat i makspunktet



Det betyr at $\hat{\theta} = \max_i(X_i)$ er MLE for θ, men det nytter ikke å derivere for å finne den.