Likelihood funktion
Givet data $y = (y_1 \dots y_n)^T$ og en parametrisk model $f_Y(y;\theta)$, $\theta \in \Theta^k$, så er likelihoodfunktionen enhver funktion på formen

$$L(\theta, y) = c(y) f_Y(y, \theta) \tag{1}$$

hvor $c(y) > 0$ ikke afhænger af θ.

Normaliseret likelihood
$L_{norm}(\theta, y) = \frac{f_Y(y;\theta)}{\sup_{\tilde{\theta}} f_Y(y;\theta)}$

Log-likelihood
$l(\theta; y) = \log L(\theta; y)$

Maksimum likelihood
Givet $Y = y$, så siges maksimum likelihood estimatet (MLE) $\hat{\theta}(y) = \hat{\theta}$ at eksistere, hvis det er det entydige maksimum til $L(\theta, y)$.

Eksistens af MLE
Givet $E = \{y : \hat{\theta}(y) \text{ eksisterer}\}$. Hvis $P(Y \in E) = 1$ for alle $\theta \in \Theta^h$, så er $\hat{\theta}(Y)$ maksimum likelihood estimatoren.

Bemærk

- Hvis $\hat{\theta} \in \text{int } \Theta^k$ (indre punkt), så kan MLE findes som en løsning til ligningen $\frac{\partial}{\partial \theta} l(\theta; y) = 0_{k \times 1}$.
- Husk at tjekke om det rent faktisk er et maksimum, dvs $\frac{\partial^2}{\partial \theta \partial \theta^T} l(\theta; y) = 0_{k \times k}$.

Sufficient statistik
$t(y_1, \dots, y_n)$ er en sufficient statistik for θ, hvis

$$f_Y(y; \theta) = h(y) g(t(y); \theta) \tag{2}$$

hvis h ikke afhænger af θ og g kun afhænger af y gennem $t(y)$.

Eksempel 2.4 + 2.5
Tre flasker fyldes til 700ml.
$y =$ for meget fyldt i: $(4.6; 6.3; 5.0)$.
$\mu =$ hvor meget overfyldes de i middel
Opstil model:

$$Y_i = \mu + \varepsilon_i \tag{3}$$

Antag Y_i uafhængige af hinanden, $Y_i \sim N(\mu, \sigma^2)$, og $\sigma^2 = 1$, μ ukendt, $\theta = \mu$.

Tæthed af model:

$$f_Y(y;\mu) = \prod_{i=1}^{3} \frac{1}{\sqrt{2\pi}} \exp\left(-\frac{(y_i-\mu)^2}{2}\right) \tag{4}$$

Likelihood:

$$L(\mu;y) = \left(\frac{1}{\sqrt{2\pi}}\right)^3 \exp\left(-\frac{\sum(y_i-\bar{y})}{2}\right) \exp\left(-\frac{n(\bar{y}-\mu)^2}{2}\right) \tag{5}$$

hvor $\bar{y} = \frac{\sum y_i}{n}$.

Ved brug af ovenstående metode til at finde MLE, så fås

$$\text{MLE} = \mu = \bar{y} = 5.3 \tag{6}$$

Dette er det besdste gæt vi har efter tre opfyldninger.

Der haves desuden en sifficient statistik:

$$t(y) = \bar{y} \tag{7}$$

for μ.

Skore funktion
Lad $\theta = (\theta_1, \dots, \theta_k) \in \Theta^k$, antag Θ^k er en åben mængde, og at log-likelihood funktionen er kontinuert differentiabel. Skore funktionen er da

$$S(\theta, y) = \frac{\partial}{\partial\theta} L(\theta;y) = \begin{pmatrix} \frac{\partial}{\partial\theta_1} l(\theta;y) \\ \vdots \\ \frac{\partial}{\partial\theta_k} l(\theta;y) \end{pmatrix} \tag{8}$$

Under regularitetsantagelser

$$E_\theta\left[\frac{\partial}{\partial\theta} l(\theta;y)\right] = 0_{k\times 1} \tag{9}$$

Informationsmatrix
Observeret information:

$$j(\theta;y) = -\frac{\partial^2}{\partial\theta\partial\theta^T} l(\theta;y) \tag{10}$$

Forventet information:

$$i(\theta) = E_\theta\left[j(\theta; Y)\right] \tag{11}$$

Under regularitetsantagelser er den forventede information og Fisher iformationen den samme, det vil sige

$$\begin{aligned} i(\theta) &= E_\theta\left[\frac{\partial^2}{\partial\theta\partial\theta^T} l(\theta; Y)\right] \\ &= E_\theta\left[\frac{\partial}{\partial\theta} l(\theta; Y)\left(\frac{\partial}{\partial\theta} l(\theta; Y)\right)^T\right] \end{aligned}$$

Bevis
Skriv en nulmatrix på en besværlig måde:

$$\begin{aligned} 0_{k\times k} &= \frac{\partial^2}{\partial\theta\partial\theta^T}\int f_Y(y;\theta)\,dy = \int \frac{\partial^2}{\partial\theta\partial\theta^T} f_Y(y;\theta)\,dy \\ &= \int \frac{\partial}{\partial\theta}\left(\frac{\partial \log f_Y(y;\theta)}{\partial\theta} f_Y(y;\theta)\right) dy \\ &= \int\left(\frac{\partial^2}{\partial\theta\partial\theta^T}\log f_Y(y;\theta)\right) f_Y(y;\theta)\,dy + \int \frac{\partial \log f_Y(y;\theta)}{\partial\theta}\frac{\partial \log f_Y(y;\theta)}{\partial\theta^T} f_Y(y;\theta)\,dy \\ &= E_\theta\left[\frac{\partial^2}{\partial\theta\partial\theta^T} l(\theta, Y)\right] + E_\theta\left[\frac{\partial l(\theta; Y)}{\partial\theta}\left(\frac{\partial l(\theta; Y)}{\partial\theta}\right)^T\right] \end{aligned}$$

Bemærk: $j(\theta; y) \geq 0_{k\times k}$ for MLE

Eksempel 2.6
Likelihood funktion:

$$L(\mu; y) = c(y)\exp\left(-\frac{n(\bar{y}-\mu)^2}{2}\right) \tag{12}$$

Log-lik funktion:

$$l(\mu; y) = -\frac{n(\bar{y}-\mu)^2}{2} + \text{ konstant} \tag{13}$$

Skore funktion:

$$S(\mu; y) = n(\bar{y}-\mu) \tag{14}$$

Obersveret information:

$$j(\mu; y) = n \tag{15}$$

Forventet information:

$$i(\mu) = E\left[j(\mu; y)\right] = E[n] = n \tag{16}$$

Generelt er $j(\theta; y)$, $j(\theta; Y)$ og $i(\theta)$ ikke ens.

Reparametrisering
$\psi = h(\theta)$ er en bijektiv funktion fra Θ^k til ψ^k. Y er en omparametrisering af θ.

$$L_\Psi(\psi; y) = L(h^{-1}(\psi); y) \tag{17}$$

Invarians
Antag $\hat{\theta}$ er MLE for θ, og $\psi = h(\theta)$ være en bijektiv afbildning fra Θ^k til Ψ^k. Så er $\hat{\psi} = h(\hat{\theta})$ MLE for $\psi = h(\theta)$.

Bevis

$$\sup_\psi L_\Psi(\psi; y) = \sup_\psi L(h^{-1}(Y); y) = \sup_\theta L(\theta; y)$$