Fordeling af ML estimator
Husk mængden

$$E = \{y : \text{ MLE eksisterer}\}$$

Definér rod af matrix

$$A^{1/2} : A = A^{1/2}(A^{1/2})^T$$

Denne eksisterer for positive semidefinitte matricer og er ikke nødvendigvis entydige. Cholesky dekomposition er én metode til at finde sådan en matrix.

Sætning
Underregularitetsbetingelser, gælder at når $n \to \infty$

1. $P(Y \in E) \to 1$: Hvis vi har et tilstrækkeligt stort datasæt, så er der stor chance for at MLE eksisterer
2. $P(Y \in E, \|\hat{\theta} - \theta\| \leq \varepsilon) \to 1, \forall \varepsilon > 0$ kaldes asymptotisk konsistens
3. $\mathbf{1}[y \in E] i(\hat{\theta})^{1/2}(\hat{\theta} - \theta) \xrightarrow{D} N_k(0, I_k)$
   $\mathbf{1}[y \in E] j(\hat{\theta})^{1/2}(\hat{\theta} - \theta) \xrightarrow{D} N_k(0, I_k)$
   Disse er praktisk anvendelige
4. $\mathbf{1}[y \in E] i(\theta)(\hat{\theta} - \theta) \xrightarrow{D} N_k(0, T_k)$ – teoretisk

Bemærk: Hvis $y \in E$, så svarer 4. til $\hat{\theta} \xrightarrow{D} N_k(\theta, i^{-1}(\theta)$. Dette kaldes asymptotisk centralitet og har asymptotisk efficiens.

De tre ovenstående kvaliteter, asymptotisk konsistens, centralitet og efficiens, er ønskværdige for en estimator.

Hvis $Var_{i,i}(\hat{\theta})$ er det i'te diagonalelement af $j^{-1}(\hat{\theta})$, så

$$\hat{\theta}_i \xrightarrow{D} N(\theta_i, Var_{i,i}(\hat{\theta}))$$

Kvadratisk approksimation af log-likelihood
Lav en 2.-ordens Taylorapproksimation omkring $\hat{\theta}$ ($\hat{\theta}$ 1-dimensionel).

$$l(\theta) \approx l(\hat{\theta}) + l'(\hat{theta})(\theta - \hat{\theta}) - \frac{1}{2} j(\hat{\theta})(\theta - \hat{\theta})^2$$

Midterste led er 0 da det er sådan vi har fundet $\hat{\theta}$. Dvs.

$$j_{norm} \approx -\frac{1}{2} j(\hat{\theta})(\theta - \hat{\theta})^2$$

Taylorapproksimationen kan ses som en normalfordelingsapproksimation. **Vektorversion**

$$l_{norm} \approx -\frac{1}{2}(\theta - \hat{\theta})^T j(\hat{\theta})(\theta - \hat{\theta})$$

Hypotesetest
Antag model $f_Y(y;\theta)$ med $\theta \in \Omega$. To hypoteser:

- Nulhypotese:
  $H_0 : \theta \in \Omega_0$, $\Omega_0 \in \Omega$a
- $H_1 : \theta \in \Omega|\Omega_0$

Teststørrelse/teststatistik
$T(Y)$. T funktion af data med kendt fordeling under H_0, altså forudsat at den er rigtig, $T(y)$ konkret værdi.

p-værdi
Sandsynligheden for at observere en teststørrelse der ligger mere ekstremt end den observerede, givet at H_0 er sand.

- Små p-værdier er (statistisk) bevis mod H_0. Forkast H_0.
- Store p-værdier: acceptér H_0. Dette betyder ikke, at den er sand, men blot at data ikke siger at den er usand. Måske ikke nok data?

Fejl

- Sand H_0 men forkastes kaldes type 1 fejl
- Falsk H_0 men accepteres kaldes type 2 fejl

Sandsynligheden for en type 1 fejl kaldes signifikansniveau, $\alpha = P(\text{type 1})$. α vælges frit i $(0, 1)$, men typisk $\alpha = 0.05$. Den sætter grænsen for store/små p-værdier, dvs. forkest H_0 hvis $p < \alpha$.

Likelihood ratio test
Lav en nulhypotese

- $H_0 : \theta \in \Omega_0$, $\dim(\Omega_0) = m$
- $H_1 : \theta \in \Omega|\Omega_0$, $\dim(\Omega) = k$
- $m < k$

Likelihood ratio

$$\lambda(y) = \frac{\sup_{\theta \in \Omega_0} l(\theta; y)}{\sup_{\theta \in \Omega} l(\theta; y)} \in [0, 1]$$

Små værdier er kritiske for H_0 – H_0 forkastes hellere ved små værdier, mens store værdier bekræfter.

χ^2-fordeling
Hvis $X_1, \ldots, X_n \overset{i.i.d.}{\sim} N(0, 1)$, så er $X_1^2 + \ldots + X_n^2 \sim \chi^2(n)$. Hvor n er frihedsgrader.

Den er et specialtilfælde af gammafordelinge: $\chi^2(n) = \Gamma(\frac{n}{2}, 2)$.

Sætning
Givet regularitetsantagelser, så gælder under H_0, at

$$-2\ln(\lambda(Y)) \xrightarrow{D} \chi^2(k - m)$$

når datamængden går mod uendelig.

Bemærk: store værdier af $-2\ln(\lambda(Y))$ er kritiske for H_0. Testen kaldes højresidet, fordi det kritiske område ligger til højre i fordelingen.

p-værdi for likelihood ratio test

$$\begin{aligned} p(y) &= \sup_{\theta \in \Omega_0} P(\lambda(Y) \leq (\lambda(y)) \\ &= \sup_{\theta \in \Omega_o} P\left(-2\ln(\lambda(Y) \leq -2\ln(\lambda(y))\right) \end{aligned}$$