Lecture 28

A few more MLE examples

Example 1

Suppose we take a sample of size n from a distribution with pdf

$$f(x;\theta) = \begin{cases} 2\theta^2 x^{-3} & \theta \le x \\ 0 & x < \theta;\ 0 < \theta \end{cases}$$

Likelihood

$$L(\theta) = \prod_{i=1}^{n} 2\theta^2 x_i^{-3} \qquad = 2^n \theta^{2n} \prod_{i=1}^{n} x_i^{-3}$$

Log likelihood

$$l(\theta) = \log L(\theta) = n\log 2 + 2n\log\theta - 3\sum_{i=1}^{n} \log x_i$$

$$\frac{\partial l(\theta)}{\partial\theta} = \frac{2n}{\theta} = 0 ?$$

Calculus doesn't help us here. However, by reasoning it seems that since $\theta \le x$ and the pdf is 0 when $x < \theta$

and $f(x_i;\theta) = 2\theta^2 x^{-3}$

it makes sense to choose $\hat{\theta} = \min(x_i) = X_{[1]}$

notation for order statistic.

Example 2

A random sample of size n from

$$f(x;\theta_1,\theta_2) = \frac{1}{\theta_2 - \theta_1} \qquad \theta_1 \le x \le \theta_2$$

Again calculus will not help us. But reasoning says that making the difference between $\hat{\theta}_2$ and $\hat{\theta}_1$ as small as possible will maximize the likelihood. Since all x_i must satisfy

$$\theta_1 \le x_i \le \theta_2$$

It stands to reason that we should take

$\hat{\theta}_1 = \min X_i = X_{[1]}$ and $\hat{\theta}_2 = \max X_i = X_{[n]}$

Example 3

Consider independent random samples $X_1, \ldots, X_n$ and $Y_1, \ldots, Y_m$ from normal distributions with common mean μ but different variances σ_1^2 and σ_2^2. What are the MLE of μ, σ_1^2 and σ_2^2?

Likelihood

$$L(\theta) = \prod_{i=1}^{n} \frac{1}{\sqrt{2\pi\sigma_1^2}} \exp\left(-\frac{1}{2\sigma^2}(X_i-\mu)^2\right) \prod_{j=1}^{m} \frac{1}{\sqrt{2\pi\sigma_2^2}} \exp\left(-\frac{1}{2\sigma_2^2}(Y_j-\mu)^2\right)$$

$$= \frac{1}{(\sqrt{2\pi})^n (\sigma_1^2)^{\frac{n}{2}}} \exp\left(-\frac{1}{2\sigma_1^2}\sum (X_i-\mu)^2\right) \frac{1}{(\sqrt{2\pi})^{\frac{m}{2}} (\sigma_2^2)^n} \exp\left(-\frac{1}{2\sigma_2^2}\sum (Y_j-\mu)^2\right)$$

Log likelihood

$$l(\theta) = \frac{-(n+m)}{2}\log(2\pi) - \frac{n}{2}\log\sigma_1^2 - \frac{m}{2}\log\sigma_2^2 - \frac{1}{2\sigma_1^2}\sum_{i=1}^{n}(X_i-\mu)^2$$

$$- \frac{1}{2\sigma_2^2}\sum_{j=1}^{m}(Y_j-\mu)^2$$

$$\frac{\partial l}{\partial \mu} = \frac{\sum_{i=1}^{n}(X_i - \hat{\mu})}{\hat{\sigma}_1^2} + \frac{\sum_{j=1}^{n}(Y_j - \hat{\mu})}{\hat{\sigma}_2^2} = 0 \quad (1)$$

$$\frac{\partial l}{\partial (\sigma_1^2)} = -\frac{n}{2\hat{\sigma}_1^2} + \frac{1}{2(\hat{\sigma}_1^2)^2}\sum_{i=1}^{n}(X_i - \hat{\mu})^2 = 0 \quad (2)$$

$$\frac{\partial l}{\partial (\sigma_2^2)} = -\frac{m}{2\hat{\sigma}_2^2} + \frac{1}{2(\hat{\sigma}_2^2)^2}\sum_{j=1}^{m}(\cancel{X_i}\, Y_j - \hat{\mu})^2 = 0 \quad (3)$$

from (1)

$$\Rightarrow \frac{\sum X_i - n\hat{\mu}}{\hat{\sigma}_1^2} + \frac{\sum Y_j - m\hat{\mu}}{\hat{\sigma}_2^2} = 0$$

$$\hat{\sigma}_2^2\left(\sum X_i - n\hat{\mu}\right) + \hat{\sigma}_1^2\left(\sum Y_j - m\hat{\mu}\right) = 0$$

$$\Rightarrow \hat{\mu} = \frac{\hat{\sigma}_2^2 \sum X_i + \hat{\sigma}_1^2 \sum Y_j}{n\hat{\sigma}_2^2 + m\hat{\sigma}_1^2} \quad (4)$$

from (2)

$$\hat{\sigma}_1^2 = \frac{\sum_{i=1}^{n} (X_i - \hat{\mu})^2}{n} \quad (5)$$

from (3)

$$\hat{\sigma}_2^2 = \frac{\sum_{j=1}^{m} (Y_j - \hat{\mu})^2}{m} \quad (6)$$

so replacing these into (4) gives

$$\hat{\mu} = \frac{\dfrac{\sum_{j=1}^{m} (Y_j - \hat{\mu})^2}{m} \sum_{j=1}^{n} X_i + \dfrac{\sum_{j=1}^{n} (X_i - \hat{\mu})^2}{n} \sum_{j=1}^{m} Y_j}{n \dfrac{\sum (Y_j - \hat{\mu})^2}{m} + m \dfrac{\sum (X_j - \hat{\mu})^2}{n}}$$

Solve this for $\hat{\mu}$ and ~~re~~substitute into (5) and (6) for the solution.