STAT 417 Lecture Note 7.

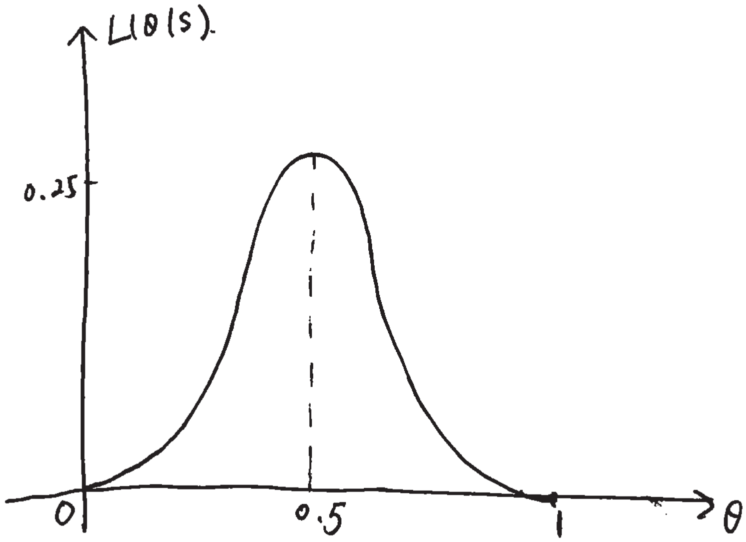
§ 6.2. Maximum likelihood principle.

In the eight-patient example. $n = 8$ patients. $\theta = P(\text{death})$. (Intuitively, we estimate $\hat{\theta} = \frac{1}{2}$)

$s = 4$ is the number of dying patients. Since this is a binomial model, i.e., Binomial$(8, \theta)$, the likelihood function is

$$L(\theta|s) = \binom{8}{4} \theta^4 (1-\theta)^{8-4} = 70\, \theta^4 (1-\theta)^4. \quad \theta \in [0,1].$$

the plot of $L(\theta|s)$ is



For each θ, $L(\theta|s)$ is considered as the likelihood. a "good" θ should give us a "big" likelihood, the best θ should give us the maximum likelihood. To find $\hat{\theta}$, we only have to maximize $L(\theta|s)$. Since ~~the~~ $L(\theta|s)$ and $\log L(\theta|s)$ have the same maximizer, we maximize instead the log-likelihood. i.e.

$$\ell(\theta|s) = \log L(\theta|s).$$

In calculus, we know that maximizer of $l(\theta|s)$ is the solution of

$$\frac{\partial}{\partial\theta} l(\theta|s) = 0. \quad \cdots\cdots (**).$$

We define the solution to $(**)$ to be the maximum likelihood estimation, denoted as $\hat{\theta}_{ML}$.

E.g. 1. $L(\theta|s) = 70\,\theta^4(1-\theta)^4$. $\quad l(\theta|s) = \log 70 + 4\log\theta + 4\log(1-\theta)$.

so solving $\frac{\partial}{\partial\theta} l(\theta|s) = \frac{4}{\theta} - \frac{4}{1-\theta} = 0$. we have

$\hat{\theta}_{ML} = 0.5$.

Ex 1. $L(\theta|s) = \binom{10}{4}\theta^4(1-\theta)^6$. Find $\hat{\theta}_{ML}$. $(=0.4)$.

E.g. 2. (Normal model).

$$L(\theta|x_1,x_2) = \left(\frac{1}{\sqrt{2\pi}\,\sigma}\right)^2 \exp\left(-\frac{(x_1-\mu)^2+(x_2-\mu)^2}{2\sigma^2}\right).$$

then $l(\theta|x_1,x_2) = -2\log(\sqrt{2\pi}\,\sigma) - \frac{(x_1-\mu)^2+(x_2-\mu)^2}{2\sigma^2}$.

$$\begin{cases} \frac{\partial}{\partial\mu} l(\theta|x_1,x_2) = -\frac{2(\mu-x_1)+2(\mu-x_2)}{2\sigma^2} = 0 & ① \\ \frac{\partial}{\partial\sigma} l(\theta|x_1,x_2) = -\frac{2}{\sigma} + \frac{(x_1-\mu)^2+(x_2-\mu)^2}{\sigma^3} = 0. & ② \end{cases}$$

From ①, $\mu = \frac{x_1+x_2}{2} = \bar{x}$.

From ② $\hat{\sigma}^2 = \frac{(x_1-\bar{x})^2+(x_2-\bar{x})^2}{2}$.

so. $\hat{\theta}_{ML} = \left(\bar{x}, \frac{(x_1-\bar{x})^2+(x_2-\bar{x})^2}{2}\right)$.

E.g. 3 (Normal model, more general).

$$L(\theta \mid x_1, x_2, \ldots, x_n) = \left(\frac{1}{\sqrt{2\pi}\,\sigma}\right)^n \exp\left(-\frac{\sum_{i=1}^n (x_i-\mu)^2}{2\sigma^2}\right).$$

$$l(\theta \mid x_1, \ldots, x_n) = -n\log(\sqrt{2\pi}\,\sigma) - \frac{\sum_{i=1}^n (x_i-\mu)^2}{2\sigma^2}.$$

so
$$\frac{\partial}{\partial\mu} l(\theta \mid x_1, \ldots, x_n) = -\frac{2\sum_{i=1}^n(\mu - x_i)}{2\sigma^2} = 0 \qquad ①$$

$$\frac{\partial}{\partial\sigma} l(\theta \mid x_1, \ldots, x_n) = -\frac{n}{\sigma} + \frac{\sum_{i=1}^n (x_i-\mu)^2}{\sigma^3} = 0. \qquad ②.$$

From ①. $0 = \sum_{i=1}^n (\mu - x_i) = n\mu - \sum_{i=1}^n x_i$. So $\mu = \frac{\sum_{i=1}^n x_i}{n} = \bar{x}$.

From ②. $\frac{\sum_{i=1}^n (x_i - \bar{x})^2}{\sigma^3} = \frac{n}{\sigma}$.

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

Define $s^2 = \frac{1}{n-1}\sum_{i=1}^n (x_i - \bar{x})^2$ as the sample variance.

then $\sigma^2 = \frac{n-1}{n} s^2$. so $\hat{\theta}_{ML} = (\bar{x}, \frac{n-1}{n} s^2)$.

E.X. 2. ~~X_1, X_2 are iid from Gamma(α, β) so.~~

$X_1, X_2, \ldots, X_n$ are iid samples from Exponential(θ). Find $\hat{\theta}_{ML}$.

Sol: $L(\theta \mid x_1, \ldots, x_n) = f_\theta(x_1, \ldots, x_n) = \theta^n e^{-\theta(x_1 + \cdots + x_n)}$.

so. $l(\theta \mid x_1, \ldots, x_n) = n\log\theta - \theta(x_1 + \cdots + x_n)$.

$\frac{\partial}{\partial\theta} l(\theta \mid x_1, \ldots, x_n) = \frac{n}{\theta} - (x_1 + \cdots + x_n) = 0$ so $\hat{\theta}_{ML} = \frac{x_1 + \cdots + x_n}{n} = \bar{x}$.