

Stat 324: Lecture 15

Maximum likelihood estimation

Moo K. Chung
mchung@stat.wisc.edu

March 17, 2005

Concepts

1. Given random sample $X_1, \dots, X_n$, the *likelihood function* is given as the product of probability or density functions, i.e.

$$L(\theta) = f(x_1; \theta) f(x_2; \theta) \cdots f(x_n; \theta).$$

The *maximum likelihood estimate* of θ is an estimate that maximizes $L(\theta)$. If we denote $\hat{\theta} = \hat{\theta}(x_1, \dots, x_n)$ to be the maximum likelihood estimate, the *maximum likelihood estimator* (MLE) of θ is $\hat{\theta} = \hat{\theta}(X_1, \dots, X_n)$.

2. *Exponential distribution.* X is an exponential distribution with parameter λ if the density function is $f(x) = \lambda e^{-\lambda x}$ for $x \geq 0$. It can be shown that $\mathbb{E}X = \frac{1}{\lambda}$, $\mathbb{V}X = \frac{1}{\lambda^2}$. Then MLE of λ is $\hat{\lambda} = \bar{X}$.

3. *Loglikelihood.* Maximizing $L(\theta)$ is equivalent to maximizing $\ln L(\theta)$ since $\ln$ is an increasing function.

Example. This technique is best illustrated by finding the MLE of parameters in $N(\mu, \sigma^2)$.

$$\hat{\mu} = \bar{X},$$

$$\hat{\sigma}^2 = \frac{1}{n} \sum_{i=1}^n (X_i - \bar{X})^2$$

are the MLE of μ and σ^2 respectively. Note that $\hat{\sigma}^2$ is not an unbiased estimator of σ^2 .

4. (Invariance Principle) If $\hat{\theta}$ is the MLE's of parameter θ then the MLE of $h(\theta)$ is $h(\hat{\theta})$ for some function h .

Proof. Consider likelihood function $L(\theta)$. $\hat{\theta}$ satisfies $\frac{dL(\theta)}{d\theta} = 0$. Let $\phi = h(\theta)$. Then the likelihood function for $\phi = h(\theta)$ is given by $L(h^{-1}(\phi))$. Differentiating the likelihood with respect to ϕ , we have

$$\frac{dL(h^{-1}(\phi))}{d\phi} = \frac{dL(\theta)}{d\theta} \frac{d\theta}{d\phi} = \frac{dL(\theta)}{d\theta} \frac{1}{h'(\theta)} = 0.$$

5. *Asymptotic unbiasedness.* When the sample size is large, the maximum likelihood estimator of θ is approximately unbiased. The MLE of θ is approximately the MVUE of θ . This is why it is the most widely used estimation technique in statistics. For the previous example,

$$\mathbb{E}\hat{\sigma}^2 = \mathbb{E}\frac{n-1}{n}S^2 = \frac{n-1}{n}\sigma^2 \rightarrow \sigma^2 \text{ as } n \rightarrow \infty$$

Problems

Let X_1, X_2 be a random sample from $N(0, 1/\theta)$. Note that the sample size is 2 and the density function for X_i is

$$f(x_i) = \frac{\sqrt{\theta}}{\sqrt{2\pi}} \exp(-\theta x_i^2/2).$$

Find the likelihood function and use it to obtain the maximum likelihood estimator of θ .

Solution. The likelihood function is

$$L(\theta) = \theta/(2\pi) \exp\left(-\theta(x_1^2 + x_2^2)/2\right).$$

Now get log-likelihood function

$$\ln L(\theta) = \text{const} + \ln \theta - \theta(x_1^2 + x_2^2)/2.$$

Differentiate with respect to θ we get

$$\frac{d \ln L(\theta)}{d\theta} = \frac{1}{\theta} - \frac{1}{2}(x_1^2 + x_2^2) = 0.$$

Solving the equation, we get

$$\hat{\theta} = 2/(x_1^2 + x_2^2).$$

Assigned problems. 6.20., 6.30. HW 5 due April 14th.