Lecture 5

Casella and Berger

Sections 7.1, 7.2.1, 7.2.2

Point Estimation

Suppose $X_1, ..., X_n$ i.i.d. $\sim f(x|\theta)$. We are interested in estimating the unknown parameters θ based on the random samples.

An estimator of unknown parameters θ is a function of the random variable $X_1, \cdots, X_n$.

An estimate of θ is a function of the realized values $x_1, \cdots, x_n$.

Methods of Finding Estimators

- *Methods of Moments*
- *Maximum Likelihood Estimators*
- *Bayes Estimators*
- *The EM Algorithm*

Methods of Moments

The methods of moments estimator $(\tilde{\theta}_1, \cdots, \tilde{\theta}_k)$ of $(\theta_1, \cdots, \theta_k)$ is obtained by solving the following system of equations for $(\theta_1, \cdots, \theta_k)$ in terms of $(m_1, \cdots, m_k)$:

$$m_i = \mu_i'(\theta_1, \cdots, \theta_k), i = 1, \cdots, k$$

where m_i and μ_i' are the sample and population moments, respectively.

Suppose $X_1, ..., X_n$ are i.i.d $\sim f(x|\theta)$.

Example 1: $X \sim N(\mu, \sigma^2)$, both μ and σ^2 are unknown.

Example 2: $X \sim Binomial(k, p)$, p is unknown.

Example 3: $X \sim Binomial(k, p)$, both k and p are unknown.

Example 4: $X \sim Gamma(\alpha, \beta)$, both α and β are unknown.

Maximum Likelihood Estimators

Suppose $X_1, ..., X_n$ are an i.i.d. sample from a population with pdf or pmf $f(x|\theta_1, ..., \theta_k)$, the likelihood function is defined by

$$L(\theta|\mathbf{x}) = L(\theta_1, ..., \theta_k|x_1, ..., x_n) = \prod_{i=1}^{n} f(x_i|\theta_1, ..., \theta_k).$$

A *maximum likelihood estimator (MLE)* of the parameter θ is a function of the sample, denoted by $\hat{\theta}(\mathbf{X})$, that maximizes the likelihood function based on the sample $L(\theta|\mathbf{X})$.

The MLE is the parameter point for which the observed sample is most likely.

If the likelihood function is differentiable (in θ), *possible candidates* for the MLE are the values of θ that solve

$$\frac{\partial L(\theta|\mathbf{x})}{\partial \theta} = 0 \text{ or } \frac{\partial \log L(\theta|\mathbf{x})}{\partial \theta} = 0(\text{score equations}).$$

To verify the solution is in fact a global maximum of the likelihood function, for univariate unknown parameter, we need to verify

- the solution is unique.
- $$\frac{d^2 \log L(\theta|\mathbf{x})}{d\theta^2}|_{\theta=\hat{\theta}} < 0.$$

Example 1. $X_1, ..., X_n$ i.i.d. $\sim N(\mu, 1)$.

Example 2. $X_1, ..., X_n$ i.i.d. $\sim Binomial(k, p)$, p is unknown.

Example 3. $X_1, ..., X_n$ i.i.d. $\sim EXP(\beta)$.

If there are more than one unknown parameters, we need to verify that the solution of the equations satisfies that $-\partial^2 \log L(\theta|\mathbf{X})/\partial\theta^2$ is positive definite.

Example. $X_1, ..., X_n$ i.i.d. $\sim N(\mu, \sigma^2)$, both μ and σ^2 are unknown.

In some situations, the solutions of the score equations may not be the MLE.

Example 1. $X_1, ..., X_n$ i.i.d. $\sim N(\mu, 1)$, and μ is nonnegative.

Example 2. $X_1, ..., X_n$ i.i.d. $\sim Binomial(k, p)$, both k and p are unknown.

Invariance property of MLEs If $\hat{\theta}$ is the MLE of θ, then for any function $\tau(\theta)$, the MLE of $\tau(\theta)$ is $\tau(\hat{\theta})$.

Example 1. $X_1, ..., X_n$ i.i.d. $\sim N(\mu, \sigma^2)$, MLE of μ^2?

Example 2. $X_1, ..., X_n$ i.i.d. $\sim Binomial(k, p)$, p is unknown, MLE of odds $p/(1-p)$?

Newton-Raphson Algorithm

A root-finding algorithm based on Taylor series expansion for equation(s) $f(x) = 0$.

$$x^{(n+1)} = x^{(n)} - [f'(x^{(n)})]^{-1} f(x^{(n)}).$$