Point Estimation

□ Estimator and Point Estimation

□ Confidence Interval

□ The effect of sample size for an estimate

Suppose that the random variables $X_1, X_2, \cdots, X_n$, whose joint distribution is assumed given excep for an unknown parameter θ, are to be observed. The problem of interest is to use the observed values to estimate θ.

♢ *parameter space* - $\Omega = \{all\ \theta\}$

♢ *estimator* - a random variable (or vector)

♢ *estimate* - a value (vector) derived from a realization

♢ *(log)-likelihood function* - $\prod_{i=1}^{n} f(x_i; \theta)$.

♢ *maximum likelihood estimator (estimate)*[mle]

Definition: If $E[u(X_1, X_2, \ldots, X_n)] = \theta$, the statistic, $u(X_1, X_2, \ldots, X_n)$ is called an *unbiased estimator* of θ. Otherwise, it is said to be *biased*.

♢ parameter estimation by *method of moments*.

Maximum Likelihood Estimation (MLE)

Let $X_1, X_2, \ldots, X_n$ be a random sample with parameter(s) $\theta \in \Omega$. Observing n independent results $x_1, x_2, \ldots, x_n$, one wants to find a *statistic* $u(X_1, X_2, \ldots, X_n)$ (called an **estimator**) to estimate θ such that $u(x_1, x_2, \ldots, x_n)$ is close to θ.

□ *Likelihood Function:* $L(\theta) = L(\theta|x_1, x_2, \cdots, x_n) = \Pi_{i=1}^{n} f(x_i; \theta$

Example 1: $X_1, X_2, \ldots, X_n \;\sim\; b(1,p) \quad \Rightarrow \quad \widehat{p} = \overline{X} = \frac{1}{n}\sum_{i=1}^{n} X_i$

Example 2: $X_1, X_2, \ldots, X_n \;\sim\; Geometric(p) \quad \Rightarrow \quad \widehat{p} = \frac{1}{\overline{X}} = \frac{n}{\sum_{i=1}^{n} X_i}$

Example 3: $X_1, X_2, \ldots, X_n \;\sim\;$ *Exponential distribution with parameter* θ, then

$$L(\theta) = \Pi_{i=1}^{n}\left[\frac{1}{\theta}e^{-x_i/\theta}\right] \quad and \quad ln(L(\theta)) = \theta^{-n} + e^{-[\sum_{i=1}^{n} x_i]/\theta}$$

Solving $\frac{d\{ln(L(\theta))\}}{d\{\theta\}} = 0$, we obtain the *MLE estimator*

$$\widehat{\theta} = \overline{X} = \frac{1}{n}\sum_{i=1}^{n} X_i$$

Example 4: $X_1, X_2, \ldots, X_n \;\sim\; N(\mu, \sigma^2)$, then

$$L(\mu, \sigma^2) = \Pi_{i=1}^{n}\left(\frac{1}{\sqrt{2\pi\sigma^2}}e^{-(x_i-\mu)^2/2\sigma^2}\right) \quad and \quad ln(L(\mu,\sigma^2)) = -\frac{n}{2}ln(2\pi) - \frac{n}{2}ln(\sigma^2) - \sum_{i=1}^{n}\frac{(x_i-\mu)^2}{2\sigma^2}$$

Solving $\frac{\partial ln(L(\mu,\sigma^2))}{\partial\mu} = 0$ and $\frac{\partial ln(L(\mu,\sigma^2))}{\partial\sigma^2} = 0$, we obtain the *MLE estimators*

$$\widehat{\mu} = \overline{X}, \quad \widehat{\sigma^2} = \frac{1}{n}\sum_{i=1}^{n}(X_i - \overline{X})^2$$

Confidence Intervals for Means

♣ Given a random sample $X_i \sim N(\mu, \sigma^2)$ of size n, we want to know how *close* of the *unbiased estimator*, $\overline{X}$, to the unknown mean μ.

$$P\left[\overline{X} - z_{\alpha/2} * (\frac{\sigma}{\sqrt{n}}) \leq \mu \leq \overline{X} + z_{\alpha/2} * (\frac{\sigma}{\sqrt{n}})\right] = 1 - \alpha \tag{1}$$

$$\left[\overline{X} - z_{\alpha/2} * (\frac{\sigma}{\sqrt{n}}), \quad \overline{X} + z_{\alpha/2} * (\frac{\sigma}{\sqrt{n}})\right] \tag{2}$$

$$S^2 = \frac{1}{n-1}\sum_{i=1}^{n}(X_i - \overline{X})^2, \quad then$$

$$(n-1)S^2 = \sum_{i=1}^{n}(X_i - \mu)^2 \; - \; n(\overline{X} - \mu)^2 \tag{3}$$

$$\frac{(n-1)S^2}{\sigma^2} = \sum_{i=1}^{n}\left(\frac{X_i - \mu}{\sigma}\right)^2 \; - \; \left[\frac{(\overline{X} - \mu)}{\sigma/\sqrt{n}}\right]^2 \tag{4}$$

$$\frac{(n-1)S^2}{\sigma^2} \sim \chi^2_{n-1} \tag{5}$$

♢ *confidence intervals*

♢ *confidence coefficient*

♢ *two-sided confidence intervals*

♢ *one-sided confidence intervals*

Effect of Sample Size:

Exercises: