STA255: Statistical Theory

Chapter 9: Properties of Point Estimators and Methods of Estimation

Summer 2017

Relative Efficiency

If $\hat{\theta}_1$ and $\hat{\theta}_2$ denote two unbiased estimators for the same parameter θ, then we say that $\hat{\theta}_1$ is relatively more efficient than $\hat{\theta}_2$ if $V(\hat{\theta}_2) > V(\hat{\theta}_1)$.

Definition

Given two unbiased estimators $\hat{\theta}_1$ and $\hat{\theta}_2$ of a parameter θ, with variances $V(\hat{\theta}_1)$ and $V(\hat{\theta}_2)$, respectively, then the efficiency of $\hat{\theta}_2$ relative to $\hat{\theta}_1$, denoted $\mathit{eff}(\hat{\theta}_1, \hat{\theta}_2)$, is defined to be the ratio

$$\mathit{eff}(\hat{\theta}_1, \hat{\theta}_2) = \frac{V(\hat{\theta}_2)}{V(\hat{\theta}_1)}$$

- $\mathit{eff}(\hat{\theta}_1, \hat{\theta}_2) > 1 \Rightarrow$ choose $\hat{\theta}_1$.
- $\mathit{eff}(\hat{\theta}_1, \hat{\theta}_2) < 1 \Rightarrow$ choose $\hat{\theta}_2$.

Example: #9.2

Consider the following three estimators for μ:

$$\hat{\mu}_1 = \frac{Y_1 + Y_2}{2},\ \hat{\mu}_2 = \frac{Y_1}{4} + \frac{Y_2 + \ldots + Y_{n-1}}{2(n-2)} + \frac{Y_n}{4},\ \hat{\mu}_3 = \bar{Y}.$$

(a) Show that each of the three estimators is unbiased.

(b) Find the efficiency of $\hat{\mu}_3$ relative to $\hat{\mu}_2$ and $\hat{\mu}_1$, respectively.

Consistency

Definition

The estimator $\hat{\theta}_n$ is said to be a consistent estimator of θ if θ_n converges in probability to θ. That is, for any positive number ϵ,

$$\lim_{n \to \infty} P(|\hat{\theta}_n - \theta| > \epsilon) = 0.$$

Theorem

An unbiased estimator $\hat{\theta}_n$ for θ is consistent if

$$\lim_{n \to \infty} V(\hat{\theta}_n) = 0.$$

Example

Let $Y_1, Y_2, \ldots, Y_n$ be a random sample of size n from a population with mean μ and variance $\sigma^2 < \infty$. Show that

$$\bar{Y} = \frac{1}{n}\sum_{i=1}^{n} Y_i$$

is a consistent estimator of θ.

Sufficiency

Definition

Let $Y_1, Y_2, \ldots, Y_n$ denote a random sample from a probability distribution with unknown parameter θ. Then the statistic $U = g(Y_1, Y_2, \ldots, Y_n)$ is said to be sufficient for θ if the conditional distribution of $Y_1, Y_2, \ldots, Y_n$, given U, does not depend on θ.

- All information about the the value of the parameter θ is contained in the value of the sufficient statistic.
- A statistic is sufficient if no other statistic that can be calculated from the same sample provides any additional information as to the value of the parameter.
- A statistic is sufficient if the sample from which it is calculated gives no additional information than does the statistic.
- How to find a sufficient statistic for θ?

Sufficiency

Definition

Let $y_1, y_2, \ldots, y_n$ be sample observations taken on corresponding random variables $Y_1, Y_2, \ldots, Y_n$ whose distribution depends on a parameter θ. Then, if $Y_1, Y_2, \ldots, Y_n$ are discrete (continuous) random variables, the likelihood of the sample, $L(\theta) = L(y_1, y_2, \ldots, y_n|\theta)$, is defined to be the joint probability (density) of $y_1, y_2, \ldots, y_n$. That is

- Discrete case:

$$L(\theta) = L(y_1, y_2, \ldots, y_n|\theta) = p(y_1, y_2, \ldots, y_n) = p(y_1|\theta) \times \ldots \times p(y_n|\theta).$$

- Continuous case:

$$L(\theta) = L(y_1, y_2, \ldots, y_n|\theta) = f(y_1, y_2, \ldots, y_n) = f(y_1|\theta) \times \ldots \times f(y_n|\theta).$$

Sufficiency

Theorem (Factorization Theorem)

Let U be a statistic based on the random sample $Y_1, Y_2, \ldots, Y_n$. Then U is a sufficient statistic for the estimation of a parameter θ if and only if the likelihood $L(\theta) = L(y_1, y_2, \ldots, y_n|\theta)$ can be factored into two nonnegative functions,

$$L(y_1, y_2, \theta, y_n|\theta) = g(u, \theta) \times h(y_1, y_2, \ldots, y_n),$$

where $g(u, \theta)$ is a function only of u and θ and $h(y_1, y_2, \ldots, y_n)$ is not a function of θ.

Example

Let $Y_1, Y_2, \ldots, Y_n$ be a random sample in which Y_i possesses the probability density function $f(y_i|\theta) = \frac{1}{\theta}e^{-y_i/\theta}, y_i \geq 0$, where $\theta > 0$ and $i = 1, 2, \ldots, n$. Show that Y is a sufficient statistic for the parameter θ.

Example

Let $Y_1, Y_2, \ldots, Y_n$ denote a random sample from the uniform distribution over the interval $(0, \theta)$. Show that $Y_{(n)} = max(Y_1, Y_2, \ldots, Y_n)$ is sufficient for θ.

How to find estimators?

There are two main methods for finding estimators:

(1) Method of moments.

(2) The method of Maximum likelihood.

The Method of Moments (MoM)

- The method of moments is a very simple procedure for finding an estimator for one or more population parameters.
- **Recall that**: the kth moment of a random variable is

$$\mu_k = E(Y^k).$$

- The corresponding kth sample moment is

$$m_k = \frac{1}{n}\sum_{i=1}^{n} Y_i^k.$$

- **Method of Moments:** The estimator based on the method of moments will be the solution to the equation $\mu_k = m_k$. That is, set

$$E(Y^k) = \frac{1}{n}\sum_{i=1}^{n} Y_i^k$$

and solve for θ.

Example

Let $Y_1, Y_2, \ldots, Y_n$ denote a random sample from the uniform distribution over the interval $(0, \theta)$. Use the method of moments to estimate θ.

Example

Let $Y_1, Y_2, \ldots, Y_n \stackrel{i.i.d.}{\sim} N(\mu, \sigma^2)$. Use the method of moments to estimate μ, σ^2.

The Method of Maximum Likelihood (MLE)

- **Method of Maximum Likelihood**: Suppose that the likelihood function depends on k parameters $\theta_1, \theta_2, \ldots, \theta_k$. Choose as estimates those values of the parameters that maximize the likelihood

$$L(y_1, y_2, \ldots, y_n | \theta_1, \theta_2, \ldots, \theta_k)$$

- $l(\theta) = \ln(L(\theta))$ is the log likelihood function.
- Both the likelihood function and the log likelihood function have their maximums at the same value of $\hat{\theta}$.
- It is often easier to maximize $l(\theta)$.

Example

Let $Y_1, Y_2, \ldots, Y_n \overset{i.i.d.}{\sim} \mathit{Bernoulli}(p)$. Find the maximum likelihood estimator of p. Is it unbiased estimator for p?

Example

Let $Y_1, Y_2, \ldots, Y_n \overset{i.i.d.}{\sim} N(\mu, \sigma^2)$. Find the maximum likelihood estimator of μ and σ^2.

Example

Let $Y_1, Y_2, \ldots, Y_n \overset{i.i.d.}{\sim} Poisson(\lambda)$. Find the maximum likelihood estimator of λ.