Theory of Estimation

Dr Ashutosh Pandey

Estimator

An estimator is any quantity calculated from the sample data which is used to give information about an unknown quantity in the population. For example, the sample mean is an estimator of the population mean .

Estimate:

Results of estimation can be expressed as a single value; known as a point estimate, or a range of values, referred to as a confidence interval. Whenever we use point estimation, we calculate the margin of error associated with that point estimation.

Properties of a Good Estimator

A "Good" estimator is the one which provides an estimate with the following qualities:

(i) **Unbiasedness**

(ii) **Consistency:**

(iii) **Efficiency:**

(iv) **Sufficiency**

Unbiasedness

Definition

An estimator $T_n = T(x_1, x_2, x_3 \ldots \ldots . x_n)$is said to be unbiased estimator of $\gamma(\theta)$ if:

$$E(T_n) = \gamma(\theta), for\ all\ \theta \in \Theta$$

Remark:

If $E(T_n) > \gamma(\theta)$, T_n is said to be positively biased and if $E(T_n) < \gamma(\theta)$, it is said to be negatively biased.

Problem

$x_1, x_2, x_3 \ldots \ldots . x_n$ is a random sample from a normal population $N(\mu, 1)$. Show that $t = \frac{1}{n}\sum {x_i}^2$ is an unbiased estimator of $\mu^2 + 1$.

Problem:

If T is an unbiased estimator for θ, show that T^2

Is a biased estimator for θ^2

Problem:

Show that $\frac{[\sum x_i(\sum x_i - 1]}{n(n-1)}$ is an unbiased estimator of θ^2 for the sample $x_1, x_2, x_3 \ldots \ldots . x_n$ drawn on X which takes the values 1 or 0 with respective probabilities $\theta \; and \; (1 - \theta)$

Example 17·1. $x_1, x_2, \ldots, x_n$ is a random sample from a normal population $N(\mu, 1)$. Show that $t = \frac{1}{n}\sum_{i=1}^{n} x_i^2$, is an unbiased estimator of $\mu^2 + 1$.

Solution. (a) We are given : $\qquad E(x_i) = \mu,\ V(x_i) = 1\ \forall\ i = 1, 2, \ldots, n$ $\qquad$ ...(*)

Now $\qquad E(x_i^2) = V(x_i) + \{E(x_i)\}^2 = 1 + \mu^2$ $\qquad$ [From (*)]

$$\therefore \quad E(t) = E\left(\frac{1}{n}\sum_{i=1}^{n} x_i^2\right) = \frac{1}{n}\sum_{i=1}^{n} E(x_i^2) = \frac{1}{n}\sum_{i=1}^{n}(1+\mu^2) = 1 + \mu^2$$

Hence t is an unbiased estimator of $1 + \mu^2$.

Example 17·2. *If T is an unbiased estimator for θ, show that T^2 is a biased estimator for*

Solution. Since T is an unbiased estimator for θ, we have $\qquad E(T) = \theta$

Also $\quad \text{Var}(T) = E(T^2) - \{E(T)\}^2 = E(T^2) - \theta^2 \Rightarrow E(T^2) = \theta^2 + \text{Var}(T),\ (\text{Var}\ T > 0).$

Since $E(T^2) \neq \theta^2$, T^2 is a biased estimator for θ^2.

Example 17·3. *Show that $\frac{[\Sigma x_i(\Sigma x_i - 1)]}{n(n-1)}$ is an unbiased estimate of θ^2, for the sample x ... x_n drawn on X which takes the values 1 or 0 with respective probabilities θ and $(1 - \theta$*

Solution. Since $x_1, x_2, \ldots, x_n$ is a random sample from Bernoulli population w... ...meter θ, $T = \sum_{i=1}^{n} x_i \sim B(n, \theta) \quad \Rightarrow \quad E(T) = n\theta$ and $\text{Var}(T) = n\theta(1-\theta)$

$$\therefore \quad E\left\{\frac{\Sigma x_i(\Sigma x_i - 1)}{n(n-1)}\right\} = E\left\{\frac{T(T-1)}{n(n-1)}\right\} = \frac{1}{n(n-1)}\{E(T^2) - E(T)\}$$

$$= \frac{1}{n(n-1)} \left[\mathrm{Var}(T) + \{E(T)\}^2 - E(T) \right]$$

$$= \frac{1}{n(n-1)} \left\{ n\,\theta\,(1-\theta) + n^2\theta^2 - n\,\theta \right\} = \frac{n\,\theta^2\,(n-1)}{n(n-1)} = \theta^2$$

$\{\sum x_i\,(\sum x_i - 1)\} \,/\, \{n(n-1)\}$ is an unbiased estimator of θ^2.

Consistency

Definition

An estimator $T_n = T(x_1, x_2, x_3 \dots \dots . x_n)$, based on a random sample of size n, is said to be consistent estimator of $\gamma(\theta)$, $\theta \in \Theta$ the parameter space, if T_n converges to $\gamma(\theta)$ in probability, i.e. $if\ T_n \rightarrow \gamma(\theta) as\ n \rightarrow \infty$.

In otherworld ,

T_n is a consistent estimator of $\gamma(\theta)$ if for every $\varepsilon > 0$, there exist a positive integer $n \geq m(\varepsilon, \eta)\ such\ that$

$P\{|T_n - \gamma(\theta)| < \varepsilon\} \rightarrow 1\ as\ n \leftarrow \infty$

$\Rightarrow P\{|T_n - \gamma(\theta)| < \varepsilon\} > 1 - \eta; \qquad for\ every\ n \geq m,$ where m is some very large value of n.

Invariance property of consistent estimators:

If
$\mathrm{T_n}$ is a consistent estimator of $\gamma(\theta)$ and $\vartheta\{\gamma(\theta)\}$ is continuous
Function of $\gamma(\theta)$, then
$\vartheta(\mathrm{T_n})$is a consistent estimator of$\vartheta\{\gamma(\theta)\}$

Sufficient Conditions for consistency:

Let $\{T_n\}$ be a sequence of estimators such that for all $\theta \in \Theta$,

(i) $E_\theta(T_n) \rightarrow \gamma(\theta), n \rightarrow \infty$

(ii) $V_\theta(T_n) \rightarrow 0, as\ \ n \rightarrow \infty$

Q.1

Prove that in sampling form a $N(\mu, \sigma^2)$ population, the sample mean is a consistent estimator of μ.

Q.2.

If $x_1, x_2, x_3 \ldots \ldots . x_n$ are random observation on a Bernoulli variate X taking the value 1 with probability p and the value 0 with probability (1-p), show that:

$$\frac{\sum x_i}{n}\left(1-\frac{\sum x_i}{n}\right) \text{ is a consistent estimator of } p(1-p).$$

Solution:

Since $x_1, x_2, x_3 \ldots \ldots . x_n$ are i.i.d. Bernoulli variates with parameter 'p',

$$T = \sum x_i \sim B(n,p) \Rightarrow E(T) = np \qquad and \;\; V(T) = npq$$

$$\bar{X} = \frac{1}{n}\sum x_i = \frac{T}{n} \qquad \Rightarrow E(\bar{X}) = \frac{1}{n}E(T) = \frac{1}{n}.\mathrm{np} = \mathrm{p}$$

$$V(\bar{X}) = V\left(\frac{T}{n}\right) == \frac{1}{n^2}V(T) = \frac{1}{n^2}.npq = \frac{pq}{n}$$

Hence $as\ n \to \infty$ $\quad E(\bar{X}) = \mathrm{p}$ and $V(\bar{X}) = 0$, $\bar{X}$ is a consistent estimator of p.

Also $\frac{\sum x_i}{n}\left(1-\frac{\sum x_i}{n}\right) = \bar{X}(\bar{X}-1)$ being a polynomial in $\bar{X}$, is a continuous function of $\bar{X}$.

Since $\bar{X}$ is a consistent estimator of p, by the invariance property of consistent estimators $\bar{X}(\bar{X}-1)$ is consistent estimator of p(1-p).

Efficient estimator

If we confine to unbiased estimates, in general there exist more than one consistent estimator of a parameter. For example,

In a sampling from a normal population $N(\mu, \sigma^2)$ sample mean $\bar{X}$ is an unbiased and consistent estimator of μ.

For Symmetry it follows immediately that sample median (Md) is an unbiased estimator of μ , which is same as the population median. Also for large n,

$$V(M_d) = \frac{1}{4nf1^2}$$

f_1 = Median ordinate of the parent distribution

= Median ordinate of the parent distribution

Hence

$$f_1 = \left[\frac{1}{\sigma\sqrt{2\pi}} e^{-\frac{1}{2}\left(\frac{x-\mu}{\sigma}\right)^2}\right] = \frac{1}{\sigma\sqrt{2\pi}} \; at \; x = \mu.$$

$$V(M_d) = \frac{1}{4n} . 2\pi\sigma^2 = \frac{\pi\sigma^2}{2n}$$

Since $\qquad E(M_d) = \mu$

and $\qquad \mathrm{V}(M_d) = 0 \qquad as \; n \to \infty.$

Hence Median is also an unbiased and consistent estimator of μ

“ If of the two consistent estimators T_1, T_2 of a certain parameter θ, we have

$$V(T_1) < V(T_2), for\ all\ n$$

Then T_1 is more efficient than T_2 for all sample size.”

For Example in case of normal distribution:

For all n $V(\bar{X}) = \frac{\sigma^2}{n}$ $\ and\ for\ large\ n, V(M_d) = 1.57\frac{\sigma^2}{n}$

Since $V(\bar{X}) < V(M_d)$, we conclude that for normal distribution, sample mean is more efficient estimator for μ than the sample median , for large sample atleast.

Most Efficient Estimator

Efficiency (Definition):

If T_1 is the most efficient estimator with variance V_1 and T_2 is any other estimator with variance V_2 then the efficiency E of T_2 is defined as:

$$E = \frac{V_1}{V_2}$$

Obviously, E can not exceed unity.

Question:1

A random sample $(X_1, X_2, X_3, X_4, X_5)$ of size five is drawn from a normal population with unknown mean μ. Consider the following estimators to estimate μ:

(a). $t_1 = \frac{X_1+X_2+X_3+X_4+X_5}{5}$

(b). $t_2 = \frac{X_1+X_2}{2} + X_3$

(c). $t_3 = \frac{2X_1+X_2+\lambda X_3}{3}$

(i) Find λ such that t_3 is unbiased estimator.

(ii) Which is the best estimator among t_1, t_2 and t_3.

Question 2:

A random sample (X_1, X_2, X_3) of size three is drawn from a population with mean value μ and variance σ^2. Consider the following estimators to estimate μ:

(a). $t_1 = X_1 + X_2 - X_3$

(b). $t_2 = 2X_1 + 3X_3 - 4X_2$

(c). $t_3 = \frac{\lambda X_1 + X_2 + X_3}{3}$

(i) Find λ such that t_3 is unbiased estimator for μ.

(ii) Which is the best estimator among t_1, t_2 and t_3.

Sufficiency

An estimator is said to be sufficient for a parameter, if it contains all the information in the sample regarding the parameter.

" if $T = t(x_1, x_2, x_3 \dots \dots . x_n)$ is an estimator of a parameter θ, based on a sample $x_1, x_2, x_3 \dots \dots . x_n$ of size n from the population with density $f(x, \theta)$ such that the conditional distribution of $x_1, x_2, x_3 \dots \dots . x_n$ given T, is independent of θ, then T is sufficient estimator for θ."

Maximum Likelihood Estimation[M.L.E.]

Likelihood function:

Let $x_1, x_2, x_3 \ldots \ldots . x_n$ be a random sample of size n from a population with density function $f(x, \theta)$. Then the likelihood function of the sample values $x_1, x_2, x_3 \ldots \ldots . x_n$, usually denoted by $L = L(\theta)$ is their joint density function, given by

$$L = f(x_1, \theta). f(x_2, \theta). f(x_3, \theta) \ldots . f(x_n, \theta) = \prod f(x_i, \theta)$$

L gives the relative likelihood that the random variables assume a particular set of values $x_1, x_2, x_3 \ldots \ldots . x_n$. For a given sample $x_1, x_2, x_3 \ldots \ldots . x_n$, L becomes a function of the variable θ, the parameter.

Maximum Likelihood Estimator:

Step I: Obtain the likelihood function i.e. $L(\theta)$

Step II: Take the logarithms i.e. $\log[L(\theta)]$

Step III : Now equate

$$\frac{\partial}{\partial\theta} Log[L(\theta)] = 0$$

This gives the Maximum Likelihood Estimator of θ.

Q. Find the maximum likelihood estimate for the parameter λ of a Poisson distribution on the basis of a sample of size n.

Solution

The Probability function of the poisson distribution with parameter λ is given by:

$$P(X = x) = f(x, \lambda) = \frac{e^{-\lambda}.\lambda^x}{x!}, x = 0,1,2,3 \ldots..$$

Likelihood function of the random sample $x_1, x_2, x_3 \ldots \ldots. x_n$ of n observation from this population is

$$L = \prod f(x_i, \theta) = \frac{e^{-n\lambda}.\lambda^{\sum x_i}}{x_1!x_2!x_3!\ldots\ldots\ldots x_n!}$$

Hence

$$logL = -n\lambda + n\bar{x}\, log\lambda - \sum \log(x_i!)$$

The likelihood equation for estimating λ is:

$$\frac{\partial}{\partial\lambda} logL = 0 \Rightarrow \lambda = \bar{x}$$

Q. In a Random sampling from normal population $N(\mu, \sigma^2)$ find the maximum likelihood estimator for μ *and* σ^2.

Solution

- Suppose $X_i \sim N(\mu, \sigma^2)$ and i.i.d.
- What is the likelihood function?

$$lik(\mu, \sigma^2) = \prod_{i=1}^{n} \left[\frac{1}{\sigma\sqrt{2\pi}} \exp\left(-\frac{(X_i - \mu)^2}{2\sigma^2} \right) \right]$$

- What is the log-likelihood function?

$$l(\mu, \sigma^2) = \sum_{i=1}^{n} \log\left[\frac{1}{\sigma\sqrt{2\pi}} \exp\left(-\frac{(X_i - \mu)^2}{2\sigma^2} \right) \right]$$

$$= -\sum_{i=1}^{n} \log(\sigma) - \sum_{i=1}^{n} \log\left(\sqrt{2\pi}\right) - \sum_{i=1}^{n} \left[\frac{(X_i - \mu)^2}{2\sigma^2} \right]$$

$$= -n\log(\sigma) - n\log\left(\sqrt{2\pi}\right) - \frac{1}{2\sigma^2} \sum_{i=1}^{n} (X_i - \mu)^2$$

Cont.

$$\frac{\partial l(\mu,\sigma^2)}{\partial \mu} = \frac{\partial}{\partial \mu}\left(-n\log(\sigma) - n\log\left(\sqrt{2\pi}\right) - \frac{1}{2\sigma^2}\sum_{i=1}^{n}(X_i-\mu)^2\right)$$

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

$$\frac{\partial l(\mu,\sigma^2)}{\partial \sigma^2} = \frac{\partial}{\partial \sigma^2}\left(-\frac{n}{2}\log(\sigma^2) - n\log\left(\sqrt{2\pi}\right) - \frac{1}{2}\left(\sigma^2\right)^{-1}\sum_{i=1}^{n}(X_i-\mu)^2\right)$$

$$= -\frac{n}{2\sigma^2} + \frac{1}{2}\left(\sigma^2\right)^{-2}\sum_{i=1}^{n}(X_i-\mu)^2$$

Set the separate partial derivatives to zero and solve for the specific parameter:

$$\frac{\partial l(\mu,\sigma^2)}{\partial \mu} = \frac{-1}{\sigma^2}\sum_{i=1}^{n}(X_i - \hat{\mu}) \equiv 0$$

$$\hat{\mu} = \frac{1}{n}\sum_{i=1}^{n} X_i = \overline{X}$$

$$\frac{\partial l(\mu,\sigma^2)}{\partial \sigma^2} = -\frac{n}{2\hat{\sigma}^2} + \frac{1}{2\left(\hat{\sigma}^2\right)^2}\sum_{i=1}^{n}(X_i - \mu)^2 \equiv 0$$

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

Example

Suppose we have a random sample X1, X2,..., Xn where:

Xi = 0 if a randomly selected student does not own a sports

car, and

Xi = 1 if a randomly selected student does own a sports car.

Assuming that the Xi are independent Bernoulli random variables with unknown parameter p, find the maximum likelihood estimator of p, the proportion of students who own a sports car.

Solution. If the X_i are independent Bernoulli random variables with unknown parameter p, then the probability mass function of each X_i is:

$$f(x_i; p) = p^{x_i}(1-p)^{1-x_i}$$

for $x_i = 0$ or 1 and $0 < p < 1$. Therefore, the likelihood function $L(p)$ is, by definition:

$$L(p) = \prod_{i=1}^{n} f(x_i; p) = p^{x_1}(1-p)^{1-x_1} \times p^{x_2}(1-p)^{1-x_2} \times \cdots \times p^{x_n}(1-p)^{1-x_n}$$

for $0 < p < 1$. Simplifying, by summing up the exponents, we get :

$$L(p) = p^{\sum x_i}(1-p)^{n-\sum x_i}$$