MATH F113
Probability and Statistics

BITS Pilani
Pilani Campus

Dr. Shivi Agarwal
Department of Mathematics

Chapter # 7

Estimation

- Point Estimation
- Interval Estimation
- Significance testing
- Hypothesis testing

- To estimate numerical value of a population parameter using a sample of some size n.
- We must device an appropriate statistic (on random sample of size n) so that we only need to consider its value on the observed sample.
- It is desirable that the statistic satisfies certain properties. We may want to have such a statistic.

Estimator and Estimate

- A statistic (which is a function on a random sample, and hence a random variable) used to estimate the population parameter θ is called a ***point estimator*** for θ and is denoted by $\hat{\theta}$
- The value of the point estimator on a particular sample of that size is called a point estimate for θ.

- Point Estimator
  - Unbiased Estimator
  - Method of Moments for estimator
  - Maximum Likelihood estimator

Desirable Properties

1. $\hat{\theta}$ to be unbiased for θ.

2. $\hat{\theta}$ to have a small variance for large sample size.

Unbiased estimator:

An estimator $\hat{\theta}$ is an unbiased estimator for a population parameter θ if and only if

$$E\left(\hat{\theta}\right) = \theta.$$

Unbiased estimator :

Let θ be the parameter of interest and $\hat{\theta}$ be a statistic. Then the statistic $\hat{\theta}$ is said to be an ***unbiased estimator***, or its value an **unbiased estimate**, if and only if the mean of the sampling distribution of the estimator equals θ, whatever the value of θ, viz. $E[\hat{\theta}]=\theta$.

Theorem :

The sample mean $\overline{X}$ of a random sample of size n from population X is an unbiased estimator of the population mean μ.

More efficient unbiased estimator :

A statistics $\hat{\theta}_1$ is said to be a more efficient unbiased estimator of the parameter θ than the statistics $\hat{\theta}_2$ if

1. $\hat{\theta}_1$ and $\hat{\theta}_2$ are both unbiased estimators of θ;

2. the variance of the sampling distribution of the first estimator is no larger than that of the second and is smaller for at least one value of θ.

Theorem:

$$\sigma_{\overline{X}}^2 = Var(\overline{X}) = \frac{\sigma^2}{n}$$

- From this theorem, it follows that larger the sample size, sample mean can be expected to lie close to population mean.
- Thus choosing large sample makes estimation more reliable.

Definition : Let $\bar{X}$ denote the sample mean of a (random) sample of size n from a distribution of standard deviation σ. Then

Standard error of mean =

$$\sigma_{\bar{X}} = \sigma / \sqrt{n}.$$

Unbiased estimator of variance

Theorem : The sample variance S^2 of a random sample of size n from a population X is an unbiased estimator for population variance σ^2, viz.

$E[S^2] = \sigma^2$.

$$E[S^2] = E\left[\frac{\sum_{i=1}^{n}(X_i - \overline{X})^2}{n-1}\right] = \frac{1}{n-1}E\left[\sum_{i=1}^{n}(X_i^2 + \overline{X}^2 - 2X_i\overline{X})\right]$$

$$= \frac{1}{n-1}\left[E(\sum_{i=1}^{n}X_i^2) + E(\sum_{i=1}^{n}\overline{X}^2) - 2E(\sum_{i=1}^{n}X_i\overline{X})\right]$$

$$= \frac{1}{n-1}\left[E(\sum_{i=1}^{n}X_i^2) + E(\overline{X}^2.n) - 2E(\overline{X}.n\overline{X})\right]$$

$$= \frac{1}{n-1}\left[E(\sum_{i=1}^{n}X_i^2) - nE(\overline{X}^2)\right] = \frac{1}{n-1}\left[n.E(X_i^2) - nE(\overline{X}^2)\right]$$

since

$$\mathrm{Var}(X_i) = E(X_i^2) - (E(X_i))^2 \Rightarrow \sigma^2 = E(X_i^2) - \mu^2$$

$$E(X_i^2) = \sigma^2 + \mu^2$$

$$and\ \mathrm{Var}(\overline{X}) = E(\overline{X}^2) - (E(\overline{X}))^2$$

$$Var\left(\frac{\sum_{i=1}^{n} X_i}{n}\right) = E(\overline{X}^2) - \mu^2 \Rightarrow \frac{Var(\sum_{i=1}^{n} X_i)}{n^2} = E(\overline{X}^2) - \mu^2$$

$$\frac{1}{n^2}\sum_{i=1}^{n} Var(X_i) = E(\overline{X}^2) - \mu^2 \Rightarrow \frac{1}{n^2} n\sigma^2 = E(\overline{X}^2) - \mu^2$$

$$E(\overline{X}^2) = \frac{1}{n}\sigma^2 + \mu^2$$

$$E[S^2] = \frac{n}{n-1}\left[E(X_i^2) - E(\overline{X}^2)\right]$$

$$= \frac{n}{n-1}\left[(\sigma^2 + \mu^2) - (\frac{1}{n}\sigma^2 + \mu^2)\right]$$

$$= \frac{n}{n-1}\left[\sigma^2 - \frac{1}{n}\sigma^2\right]$$

$$= \frac{n}{n-1} . \frac{n-1}{n}\sigma^2 = \sigma^2$$

Q.4 : An interactive computer system is available at a large installation. Let X denote the number of requests for this system received per hr. Assume X has Poisson dist with parameter λs. These data are obtained :

25 20 20 30 24 15 10 23 4.

(a)Find an unbiased estimate for λs.

(b)Find an unbiased estimate for the average number of requests per hr.

(c) Find an unbiased estimate for the average number of requests per quarter of an hr.

Q.8

Note that S is a statistic, and unless X is constant, its value varies from sample to sample, thus Var[S]>0. Show S is not an unbiased estimator of σ by method of contradiction.

If $E[S]= \sigma$, $Var[S]=E[S^2] – E[S]^2 = \sigma^2 - \sigma^2=0$.

7.2. Methods to find estimators

To find 'good' estimators for other population parameters, we describe 2 methods :

1. Method of moments
2. Maximal likelihood method.

Methods of Moments

- Method of Moments:

  k^{th} moment = E (X^k)

- An estimator of E (X^k) based on a random sample size n = M_k

$$M_k = \sum_{i=1}^{n} (X^k_i / n)$$

$$M_1 = \sum_{i=1}^{n} (X_i / n) = \overline{X}$$

$$M_2 = \sum_{i=1}^{n} (X^2_i / n)$$

$$M_3 = \sum_{i=1}^{n} (X_i^3 / n)$$

Method of moments

1) Use estimators $M_k = \sum_{i=1}^{n} \frac{X_i^k}{n}$

for the moments $E[X^k], k = 1, 2, etc..$

2) Express $E[X^k]$ in terms of parameters of the distribution.

3) Set the equations by replacing the parameters in $E[X^k]$ by their estimators and equating to M_k.

4) Set as many (suitable) equations as number of parameters and solve them for the estimators of the parameters in terms of M_k's.

Ex. Show method of moments estimator for σ^2 is $(n-1)S^2/n$, hence not unbiased.

$\sigma^2 = E[X^2]-E[X]^2$, therefore its estimator is

Q. 7.2.16

Let $X_1, \ldots, X_m$ be a random sample of size m from a binomial distribution with parameters n, assumed to be known, and p (unknown). Show that the method of moments estimator of p is

$$\hat{p} = \frac{\bar{X}}{n}.$$

Example 7.2.2

Let $X_1, \ldots, X_n$ be a random sample from a gamma distribution with parameters α, β. Find method of moments estimators for α, β.

Using $E[X] = \alpha\beta$, $E[X^2]-(E[X])^2 = \alpha\beta^2$, we get

$$\hat{\beta} = \frac{M_2 - M_1^2}{M_1}, \hat{\alpha} = \frac{M_1^2}{M_2 - M_1^2}.$$

Maximum likelihood estimation :

Consider a random sample of size n from population $f(x;\theta)$ that depends on a parameter θ. The joint distribution is

$$f(x_1;\theta)f(x_2;\theta)......f(x_n;\theta).$$

$$L(\theta) = f(x_1;\theta)\, f(x_2;\theta) \ldots\ldots f(x_n;\theta)$$

is called the ***likelihood function***. The ***maximum likelihood estimator*** of θ is the random variable which equals the value for θ that maximizes the probability of the observed sample.

Maximum Likelihood Estimation

1. MLE is the most widely used parameter estimation method as on today.

2. The basic principle is to maximize the likelihood of the parameters, denoted by $L(\theta \mid x)$, as a function of the model parameters θ.

3. Note that the θ can be a single parameter or a vector of parameters; $\theta = (\theta_1, \theta_2, \cdots, \theta_p)$.

4. The likelihood function $L(\theta \mid x)$ is defined as $L(\theta \mid x) = \prod_{i=1}^{n} f(x_i; \theta)$

5. As log is a one – to – one function, maximization of log – likelihood $(\ln L)$ is often preferred for computational ease.

Method of Maximum Likelihood function for Estimating θ

1. Obtain a random sample $X_1, X_2, X_3, \ldots\ldots, X_n$ from the distribution of a random variable X with the density 'f ' and associated parameter θ

2. Define a function $L(\theta)$ by

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

This function is called the **likelihood** function for the sample.

3. Find the expression for θ that maximizes the likelihood function. This can be done directly or by maximizing $\ln L(\theta)$.

4. Replace θ by $\hat{\theta}$ to obtain an expression for the maximum likelihood estimator for θ

5. Find the observed value of this estimator for a given sample.

Example: Let X_1, X_2 , X_3 ,……., X_n be a random sample from a normal distribution with mean μ variance σ^2 .The density for X is

$$f(x)=(1/\sqrt{(2\pi)}\sigma)\, e^{-(1/2)[(x-\mu)/\sigma]^2}$$

The likelihood function for the sample is a function of both μ and σ . In particular,

$$L(\mu,\sigma) = \prod_{i=1}^{n} \frac{1}{\sqrt{2\pi}\sigma} e^{-\frac{1}{2}\left(\frac{x_i-\mu}{\sigma}\right)^2}$$

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

$$ln(L(\mu,\sigma)) = -n\ln(\sqrt{2\pi}) - n\ln(\sigma) - \frac{1}{2\sigma^2}\sum_{i=1}^{n}(x_i-\mu)^2$$

To maximize this function , we take the partial derivatives with respect to μ and σ , set these derivatives equal to 0, and solve the equations simultaneously for μ and σ :

The method of moments estimator for a parameter and the maximum likelihood estimator often agree. However, if they do not, the maximum likelihood estimator is usually preferred.

Q. 30. Let $X_1, X_2, X_3, \ldots\ldots\ldots, X_m$ be a random sample of size m from a Binomial distribution with parameters n (known) and p (unknown). Find the maximum likelihood estimator for p.

Q. 31. Let W be an exponential random variable with parameter β unknown. Find the maximum likelihood estimator for β based on a sample of size n.