Point Estimation

Basic definitions

- **Population:** The group of individual under study is called Population.
- **Sample:** A finite subset of statistical individuals in a population is called Sample.
- **Sample size:** The number of individuals in a sample is called sample size.
- **Simple Random Sampling:** A random sample is one in which each unit of population has an equal chance (say p) of being included in it and this probability is independent of the previous drawing.
- **Parameter:** The statistical constant of population is called parameter e.g. Mean μ, Variance σ^2 etc.

Basic definitions

- **Statistics:** Statistical measures computed from the sample observations e.g. sample mean $\bar{x}$, sample variance s^2 etc.
- **Estimator:** Any function of the random sample$x_1, x_2, \dots, x_n$that are being observed, say $T_n(x_1, x_2, \dots, x_n)$ is called a statistic, it is a random variable and it is used to estimate an unknown parameter θ of the distribution, it is called an estimator. A particular value of the estimator say $T_n(x_1, x_2, \dots, x_n)$ is called an estimate of θ.

Characteristics of Estimators

- Unbiasedness
- Consistency
- Efficiency
- Sufficiency

Unbiasedness

- **Unbiased Estimator:** An estimator $T_n = T(x_1, x_2, \dots, x_n)$ is said to be an unbiased estimator of parameter $\gamma(\theta)$, if
$$E(T_n) = \gamma(\theta)$$

Example: Sample mean $\overline{\boldsymbol{x}}$ is an unbiased estimator of population mean μ, but sample variance s^2 is biased estimator of population variance σ^2.

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

Q: If T is an unbiased estimator for θ, show that T^2 is a biased estimator for θ^2.

Consistency

- An estimator $T_n = T(x_1, x_2, \ldots, x_n)$ based on random sample of size n, is said to be consistent estimator of $\gamma(\theta)$, if T_n converges to $\gamma(\theta)$ in probability, i.e. $T_n \xrightarrow{p} \gamma(\theta)$ as $n \to \infty$.
- In other words, T_n is consistent estimator of $\gamma(\theta)$ if for every $\epsilon > 0$, $\eta > 0$, there exists a positive integer $n \geq m\ (\epsilon, \eta)$ such that

$$P\{|T_n - \gamma(\theta)| < \epsilon\} \to 1 \text{ as } n \to \infty$$
$$\Rightarrow P\{|T_n - \gamma(\theta)| < \epsilon\} > 1 - \eta\ \forall n \geq m$$

Where m is very large value of n.

Invariance property of consistent estimator

- If T_n is a consistent estimator of $\gamma(\theta)$ and $\psi\{\gamma(\theta)\}$ is a continuous function of $\gamma(\theta)$, then $\psi(T_n)$ is a consistent estimator of $\psi\{\gamma(\theta)\}$.

Sufficient conditions for consistency

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

(i) $E(T_n) \to \gamma(\theta)$, as $n \to \infty$

(ii) $V(T_n) \to 0$, as $n \to \infty$.

Then T_n is a consistent estimator of $\gamma(\theta)$.

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

Q: If $X_1, X_2, \dots, X_n$ are random observations on a Bernoulli variate X taking the value 1 with probability p and the value 0 with probability $(1-p)$, show that: $\frac{\sum x_i}{n}\left(1-\frac{\sum x_i}{n}\right)$ is a consistent estimator of $p(1-p)$.

Efficient estimator

- 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 sizes.
- **Most efficient estimator:** If in a class of consistent estimators or a parameter, there exists one whose sampling variance is less than that of any such estimator, it is called the most efficient estimator.
- **Efficiency:** 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 η of T_2 is defined as : $\eta = \frac{\mathrm{V}_1}{\mathrm{V}_2}$, $0 \leq \eta < 1$.

Q: A random sample X_1, X_2, X_3, X_4, X_5 of size 5 is drawn from a normal population with unknown mean μ. Consider the following estimators to estimate μ:
(i) $t_1 = \frac{X_1+X_2+X_3+X_4+X_5}{5}$ (ii) $t_2 = \frac{X_1+X_2}{2} + X_3$ (iii) $t_3 = \frac{(2X_1+X_2+\lambda X_3)}{3}$, where λ is such that t_3 is an unbiased estimator of μ . Find λ. Are t_1 and t_2 unbiased? State giving reasons, the estimator which is best among t_1, t_2 and t_3.

Q: If X_1, X_2 and X_3 is a random sample of size 3 from a population with mean value μ and variance σ^2. If t_1, t_2 and t_3 are the estimators used to estimate the mean value μ, where $t_1 = X_1 + X_2 - X_3$, $t_2 = 2X_1 + 3X_3 - 4X_2$, and $t_3 = \frac{\lambda X_1 + X_2 + X_3}{3}$.

(i) Are t_1 and t_2 unbiased estimators?

(ii) Find the value of λ such that t_3 is unbiased estimator for μ.

(iii) With the value of λ, is t_3 a consistent estimator?

(iv)Which is the best estimator?

Sufficiency

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

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

i.e. $P(x_1 \cap x_2 \cap \cdots \cap x_n | T)$ is independent of parameter θ, then T is sufficient estimator of θ.

Condition for sufficient estimator

- **Factorization Theorem:** $T = t(x)$ is sufficient for θ, iff the joint density function say L, of the sample values can be expressed in the form: $L = g_\theta[t(x)]h(x)$, where $g_\theta[t(x)]$ depends on θ and x only through the value f $t(x)$ and $h(x)$ is independent of θ.
- **Fisher-Neymann Criteria:** A statistic $t_1 = t(x_1, x_2, \dots, x_n)$ is a sufficient estimator of parameter θ if and only if the likelihood function can be expressed as:

$$L = \prod_{i=1}^{n} f(x_i, \theta) = g(t_1, \theta)h(x_1, x_2, \dots, x_n)$$

Where $g(t_1, \theta)$ is thee pdf of the statistic t_1 and$h(x_1, x_2, \dots, x_n)$

is a function of sample observations only, independent of θ.

Q: Let $X_1, X_2, \dots, X_n$ be a random sample from the population with pdf $f(x,\theta) = \theta x^{\theta-1}; 0 < x < 1,\ \theta > 0$. Show that $\prod_{i=1}^{n} X_i$ is sufficient for θ.

Order statistics

- Let $X_1, X_2, \dots, X_n$ be n independent and identically distributed variates, each with cumulative distribution function $F(x)$ and pdf $f(x)$. If these variables are arranges in ascending order of magnitude and then written as $X_{(1)} \le X_{(2)} \le X_{(3)} \le \cdots \le X_{(n)}$, we call $X_{(r)}$ as the r^{th} order statistics, $r = 1,2,3, \dots, n$.

Here, $X_{(1)}$= the smallest of $X_1, X_2, \dots, X_n$.

$X_{(n)}$= the largest of $X_1, X_2, \dots, X_n$.

Cumulative and pdf of order statistics

- The cumulative distribution function of $X_{(n)}$, the largest order statistics is given by

$$F_n(x) = [F(x)]^n$$

The pdf $f_n(x)$ of$X_{(n)}$ is given by

$$f_n(x) = n\,[F(x)]^{n-1} f(x)$$

- The cumulative distribution function of $X_{(1)}$, the largest order statistics is given by

$$F_1(x) = 1 - [1 - F(x)]^n$$

The pdf $f_n(x)$ of$X_{(n)}$ is given by

$$f_n(x) = n\,[1 - F(x)]^{n-1} f(x)$$

Q: Let $x_1, x_2, \ldots, x_n$ be a random sample from a uniform population on $[0, \theta]$. Find a sufficient estimator for θ.

Q: Let $x_1, x_2, \dots, x_n$ be a random sample from a distribution with p.d.f $f(x) = e^{-(x-\theta)}, \theta < x < \infty, -\infty < \theta < \infty$. Obtain a sufficient statistic for θ.

Maximum Likelihood estimator (MLE)

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

$$L = f(x_1, \theta) f(x_2, \theta) \ldots f(x_n, \theta) = \prod_{i=1}^{n} f(x_i, \theta).$$

The principle of maximum likelihood consists in finding an estimator for the unknown parameter $\theta = (\theta_1, \theta_2, \ldots, \theta_k)$, say, which maximises the likelihood function $L(\theta)$ for variations in parameter *i.e.*, we wish to find $\hat{\theta} = (\hat{\theta}_1, \hat{\theta}_2, \ldots, \hat{\theta}_k)$ so that

$$L(\hat{\theta}) > L(\theta) \qquad \forall \; \theta \in \Theta$$

i.e., $$L(\hat{\theta}) = \text{Sup } L(\theta) \; \forall \; \theta \in \Theta.$$

Thus if there exists a function $\hat{\theta} = \hat{\theta}\,(x_1, x_2, \ldots, x_n)$ of the sample values which maximises L for variations in θ, then $\hat{\theta}$ is to be taken as an estimator of θ. $\hat{\theta}$ is usually called *Maximum Likelihood Estimator (M.L.E.)*. Thus $\hat{\theta}$ is the solution, if any, of

$$\frac{\partial L}{\partial \theta} = 0 \quad \text{and} \quad \frac{\partial^2 L}{\partial \theta^2} < 0 \qquad \ldots(15{\cdot}54)$$

Since $L > 0$, and $\log L$ is a non-decreasing function of L; L and $\log L$ attain their extreme values (maxima or minima) at the same value of $\hat{\theta}$. The first of the two equations in (15·54) can be rewritten as

$$\frac{1}{L} \cdot \frac{\partial L}{\partial \theta} = 0 \quad \Rightarrow \quad \frac{\partial \log L}{\partial \theta} = 0, \qquad \ldots(15{\cdot}54a)$$

a form which is much more convenient from practical point of view.

Properties of MLE

Theorem 15·11. (*Cramer-Rao Theorem*). *"With probability approaching unity as $n \to \infty$, the likelihood equation $\frac{\partial}{\partial \theta} \log L = 0$, has a solution which converges in probability to the true value θ_0". In other words M.L.E.'s are consistent.*

Remark. *MLE's* are always consistent estimators but need not be unbiased. For example in sampling from $N(\mu, \sigma^2)$ population,

$\text{MLE}(\mu) = \bar{x}$ (sample mean), which is both unbiased and consistent estimator of μ.

$\text{MLE}(\sigma^2) = s^2$ (sample variance), which is consistent but not unbiased estimator of σ^2.

Theorem 15·13. (Asymptotic Normality of MLE's). *A consistent solution of the likelihood equation is asymptotically normally distributed about the true value θ_0.* Thus, $\hat{\theta}$ is asymptotically $N\left(\theta_0, \frac{I}{I(\theta_0)}\right)$, as $n \to \infty$.

Remark. Variance of M.L.E. is given by

$$V(\hat{\theta}) = \frac{1}{I(\theta)} = \frac{1}{\left[E\left(-\frac{\partial^2}{\partial\theta^2}\log L\right)\right]} \qquad \ldots(15\cdot55)$$

Theorem 15·14. *If M.L.E. exists, it is the most efficient in the class of such estimators.*

Theorem 15·15. *If a sufficient estimator exists, it is a function of the Maximum Likelihood Estimator.*

Theorem 15·17. (Invariance Property of MLE). *If T is the MLE of θ and $\psi(\theta)$ is one to one function of θ, then $\psi(T)$ is the MLE of $\psi(\theta)$.*

Maximum likelihood estimator

Q: In random sampling from normal population N(μ; σ^2), find the maximum likelihood estimators for

(i) μ when σ^2 is known. (ii) σ^2 when μ is known. and (iii) the simultaneous estimation of μ and σ^2

Q: Find the maximum likelihood estimate for the parameter λ of a Poisson distribution on the basis of a sample of size n. Also find its variance.