Statistical Inference

Dr Venkataramana B

Vellore Institute of Technology, Vellore

January 4, 2025

Overview

The Theory of Estimation was founded by ***Prof. R A Fisher*** in 1930.

Keywords

- Population
- Sample
- Parameter
- Statistic

Notations

Statistical Measure	**Parameter**	**Statistic**
Mean	μ	$\bar{x}$
Median	M	m
Variance	σ^2	s^2
Standard deviation	σ	s
Proportion	P	p
Correlation Coefficient	ρ	r
Regression Coefficient	β	b

Let us consider a random variable $\boldsymbol{X}$ with probability density function (*p.d.f*) $f(x, \theta)$. The functional form of the population distribution is assumed to be known except for the value of some unknown parameter(s) θ which may take any value on a set Θ.

Parameter Space

The set Θ, which is the set of all possible values of $\boldsymbol{\theta}$ is called the **Parameter space** *i.e.* the *p.d.f* can be written as $f(x, \theta), \boldsymbol{\theta} \in \Theta$.

Estimate & Estimator

Definition 1.1

Any function of the random sample $x_1, x_2, \ldots, x_n$ that are being observed, say $T_n(x_1, x_2, \ldots, x_n)$ is called a statistic. Clearly, a statistic is a random variable. If it is used used to estimate an unknown parameter $\boldsymbol{\theta}$ of the distribution, it is called an ***estimator***.

Definition 1.2

A particular value of the estimator, say $T_n(x_1, x_2, \ldots, x_n)$ is called an ***estimate*** of $\boldsymbol{\theta}$.

Example

- How to assess the performance of estimators and to choose the best one?

Example

- How to assess the performance of estimators and to choose the best one?
- Are there methods for obtaining estimators other than "ad-hoc" methods.

Characteristics of Estimators

- Unbiasedness
- Consistency
- Efficiency
- Sufficiency

Unbiasedness

Definition 2.1

An estimator $T_n = T(x_1, x_2, \ldots, x_n)$ is said to be an unbiased estimator of $\gamma(\theta)$ if $E(T_n) = \gamma(\theta)$, for all $\theta \in \Theta$.

$$\boldsymbol{E(Statistic) = Parameter}$$

Remark. If $E(T_n) > \theta$, T_n is said to be positively biased and
if $E(T_n) < \theta$, it is said to be negatively biased,
the amount of bias $b(\theta)$ is given by $b(\theta) = E(T_n) - \gamma(\theta), \theta \in \Theta$.

Example

Suppose that $\boldsymbol{X}$ is a random variable with mean μ and variance σ^2. Let $X_1, X_2, \ldots, X_n$ be a random sample of size n from the population represented by X. Show that the sample mean $\bar{X}$ and sample variance S^2 are unbiased estimators of μ and σ^2, respectively.

Solution: Since $E(X_i) = \mu$ and $V(X_i) = \sigma^2; i = 1, 2, \ldots, n$

Sample Mean, $\bar{X} = \frac{\sum_{i=1}^{n} X_i}{n}$

$$
\begin{aligned}
E(\bar{X}) &= E\left(\frac{\sum_{i=1}^{n} X_i}{n}\right) \\
&= \frac{E(X_1) + E(X_2) + \ldots + E(X_n)}{n} \\
&= \frac{\mu + \mu + \ldots + \mu}{n} \\
&= \frac{n\mu}{n} \\
E(\bar{X}) &= \mu
\end{aligned}
$$

Therefore, the sample mean $\bar{X}$ is an unbiased estimator of the population mean μ.

$$Sample\,Variance, \quad s^2 = \frac{\sum_{i=1}^{n}(X_i - \bar{X})^2}{n}$$

$$S^2 = \frac{\sum_{i=1}^{n}(X_i - \bar{X})^2}{n-1}$$

$$\begin{aligned}
E(S^2) &= E\left[\frac{\sum_{i=1}^{n}(X_i - \bar{X})^2}{n-1}\right] = \frac{1}{n-1}E\left[\sum_{i=1}^{n}(X_i - \bar{X})^2\right] \\
&= \frac{1}{n-1}E\left[\sum_{i=1}^{n}(X_i^2 + \bar{X}^2 - 2\bar{X}X_i)\right] \\
&= \frac{1}{n-1}E\left[\sum_{i=1}^{n}X_i^2 + \sum_{i=1}^{n}\bar{X}^2 - 2\bar{X}\sum_{i=1}^{n}X_i\right] \\
&= \frac{1}{n-1}E\left[\sum_{i=1}^{n}X_i^2 + n\bar{X}^2 - 2n\bar{X}^2\right] \\
&= \frac{1}{n-1}E\left[\sum_{i=1}^{n}X_i^2 - n\bar{X}^2\right] = \frac{1}{n-1}\left[\sum_{i=1}^{n}E(X_i^2) - nE(\bar{X}^2)\right]
\end{aligned}$$

However, since $E(X_i^2) = \mu^2 + \sigma^2 \quad and \quad E(\bar{X}^2) = \mu^2 + \sigma^2/n$

$$\begin{aligned}
E(S^2) &= \frac{1}{n-1}\left[\sum_{i=1}^{n}(\mu^2 + \sigma^2) - n(\mu^2 + \sigma^2/n)\right] \\
&= \frac{1}{n-1}\left(n\mu^2 + n\sigma^2 - n\mu^2 - \sigma^2\right) \\
E(S^2) &= \sigma^2
\end{aligned}$$

Therefore, the sample variance S^2 is an unbiased estimator of the population variance σ^2.

Example

Let $X_1, X_2, \ldots, X_n$ 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 $(1+\mu^2)$.

Solution:

$$Given, \quad E(X_i) = \mu, V(X_i) = 1, \forall \ i = 1, 2, \ldots, n$$

$$\implies E(X_i^2) = V(X_i) + \{E(X_i)\}^2 = 1 + \mu^2 \{\because V(x) = E(X^2) - (E(X))^2$$

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

$$= \frac{n(1+\mu^2)}{n}$$

$$E(t) = 1 + \mu^2$$

Hence, $\boldsymbol{t}$ is an unbiased estimator of the parameter $(1+\mu^2)$.

Example

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

Solution: Since, $\boldsymbol{T}$ is an unbiased estimator for θ,

we have, $E(T) = \theta$

$$Also, \quad Var(T) = E(T^2) - [E(T)]^2 = E(T^2) - \theta^2$$

$$\Longrightarrow \qquad E(T^2) = \theta^2 + Var(T)$$

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

Example

Show that $\frac{\left[\sum x_i(\sum x_i - 1)\right]}{n(n-1)}$ is an unbiased estimator of θ^2, for the sample $x_1, x_2, \ldots, x_n$ 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 with parameter θ,

$$Let\ T = \sum_{i=1}^{n} x_i \sim B(n,\theta)$$

$$\implies E(T) = n\theta \quad and \quad Var(T) = n\theta(1-\theta)$$

$$\Big\{X \sim B(1,\theta) \implies E(X) = \theta \quad \& \quad Var(X) = \theta(1-\theta)\Big\}$$

$$\begin{aligned}
E\left[\frac{\left[\sum x_i\left(\sum x_i - 1\right)\right]}{n(n-1)}\right] &= E\left[\frac{T(T-1)}{n(n-1)}\right] \\
&= \frac{1}{n(n-1)}\Big[E(T^2) - E(T)\Big] \\
&= \frac{1}{n(n-1)}\Big[Var(T) + \{E(T)\}^2) - E(T)\Big] \\
&= \frac{1}{n(n-1)}\Big[n\theta(1-\theta) + n^2\theta^2 - n\theta\Big] \\
&= \frac{n\theta^2(n-1)}{n(n-1)} = \theta^2
\end{aligned}$$

$\implies \frac{\left[\sum x_i(\sum x_i - 1)\right]}{n(n-1)}$ is an unbiased estimator of θ^2.

Example

Let X be distributed in the Poisson form with parameter θ. Show that only unbiased estimator of $\mathbf{exp}\{-(\boldsymbol{k}+\mathbf{1})\boldsymbol{\theta}\}$,$\boldsymbol{k} > \mathbf{0}$, is $T(X) = (-k)^X$ so that $T(X) > 0$ if x is even and $T(X) < 0$ if x is odd.

Solution:

Since $X \sim Poisson(\theta) \implies P(X = x) = \frac{e^{-\theta}\theta^x}{x!};\ X = 0, 1, \ldots, \infty$

$$\begin{aligned}
E\{T(X)\} &= E[(-k)^X],\ k > 0 \\
&= \sum_{x=0}^{\infty} (-k)^X \Big\{\frac{e^{-\theta}\theta^x}{x!}\Big\} \\
&= e^{-\theta} \sum_{x=0}^{\infty} \Big[\frac{(-k\theta)^X}{x!}\Big] \\
&= e^{-\theta} \cdot e^{-k\theta} \\
E\{T(X)\} &= e^{-(k+1)\theta}
\end{aligned}$$

$\implies T(X) = (-k)^X$ is an unbiased estimator for $\exp\{-(k+1)\theta\}, k > 0$.

Consistency

Definition 2.2

An estimator $T_n = T(x_1, x_2, \ldots, x_n)$ 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 $T_n \xrightarrow{p} \gamma(\theta)$ as $n \longrightarrow \infty$.

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

$$
\begin{aligned}
&P\{|T_n - \gamma(\theta)| < \varepsilon\} \to 1 \, as \, n \longrightarrow \infty \\
\Rightarrow &P\{|T_n - \gamma(\theta)| < \varepsilon\} > 1 - \eta; \forall \;\; n \geq m(\varepsilon, \eta)
\end{aligned}
$$

where ε, η are arbitrarily small positive numbers and m is some large value of n.

Remarks

1. If $X_1, X_2, \ldots, X_n$ is a random sample from population with finite mean, $E(X_i) = \mu$, then by Khinchine's weak law of large numbers (W.L.L.N), we have

$$\bar{X}_n = \frac{1}{n}\sum_{i=1}^{n} X_i \xrightarrow{p} E(X_i) = \mu, \text{ as } n \to \infty.$$

Hence the *sample mean* ($\bar{X}_n$) is always a consistent estimator of the population mean (μ).

2. Consistency is a property concerning the behavior of an estimator for indefinitely large values of the sample size n, *i.e.*, as $n \to \infty$.

In-variance Property

Theorem 2.3

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

Theorem 2.4

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

$$(i) E_\theta(T_n) \to \gamma(\theta), n \longrightarrow \infty \quad \textit{and} \quad (ii) Var_\theta(T_n) \to 0, as \ \ n \longrightarrow \infty.$$

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

Example

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

Solution: Given $X \sim N(\mu, \sigma^2) \implies E(X) = \mu$ and $V(X) = \sigma^2$

$$\begin{aligned} E(\bar{X}) &= E\Big(\frac{X_1 + \ldots + X_n}{n}\Big) \\ &= \frac{1}{n}\{E(X_1) + \ldots + E(X_n)\} \\ &= \frac{1}{n}(\mu + \ldots + \mu) \\ &= \frac{1}{n}n\mu \\ E(\bar{X}) &= \mu \end{aligned}$$

$$\begin{aligned} Var(\bar{X}) &= Var\Big(\frac{X_1 + \ldots + X_n}{n}\Big) \\ &= \frac{1}{n^2}V(X_1 + \ldots + X_n) \\ &\quad \{\text{ Since } X_i \text{ 's are independent RV's}\} \\ &= \frac{1}{n^2}\{V(X_1) + \ldots + V(X_n)\} \\ &= \frac{1}{n^2}(\sigma^2 + \ldots + \sigma^2) \\ &= \frac{1}{n^2}n\sigma^2 \\ V(\bar{X}) &= \frac{\sigma^2}{n} \quad \Big\{\because sd(\bar{X}) = \frac{\sigma}{\sqrt{n}}\Big\} \end{aligned}$$

$$\implies \quad E(\bar{X}) = \mu \quad and \quad V(\bar{X}) = \frac{\sigma^2}{n}$$

Thus as $n \to \infty$,

$$\implies \quad E(\bar{X}) = \mu \quad and \quad V(\bar{X}) = 0$$

Hence by *theorem 2.4*, $\bar{X}$ is a consistent estimator of μ.

Example

If $X_1, X_2, \ldots, X_n$ is a random observations on a Bernoulli variate X taking the value 1 with probability p and the value 0 with probability $(1-p)$, show that

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

Solution:

$\dot{\text{S}}$ince $X_1, \ldots, X_n$ are *i.i.d* Bernoulli random variables with the parameter 'p',

$$\begin{aligned} T &= \sum_{i=1}^{n} X_i \sim B(n,p) \\ E(T) = np \quad and \quad Var(T) = npq \\ E(\bar{X}) &= \\ Var(\bar{X}) &= \end{aligned}$$

Example

Suppose that $X_1, \ldots, X_n$ are an iid random sample from the distribution

$$f(x;\theta) = \frac{1}{2}(1+\theta x), -1 < X < 1, -1 < \theta < 1$$

Show that $3\bar{X}_n$ is a consistent estimator of the parameter θ.

Solution:

Efficiency

Definition 2.5

If T_n 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 :

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

E cannot exceed *unity*.

Example: In sampling from a Normal population $N(\mu, \sigma^2)$, when σ^2 is known, sample mean $\bar{x}$ is an unbiased and consistent estimator of population mean μ.
From symmetry it follows immediately that sample median (*Md*) is an unbiased estimate of μ, which is the same as the population median. Also for large *n*,

$$V(Md) = \frac{1}{4nf_1^2}$$

$$\begin{aligned} Here \quad f_1 &= \text{Median ordinate of the parent distribution} \\ &= \text{Modal ordinate of the parent distribution} \\ &= \Big[\frac{1}{\sigma\sqrt{2\pi}} e^{-(x-\mu)^2/2\sigma^2}\Big]_{x=\mu} \\ &= \frac{1}{\sigma\sqrt{2\pi}} \\ \therefore V(Md) &= \frac{1}{4n} \cdot 2\pi\sigma^2 = \frac{\pi\sigma^2}{2n} \end{aligned}$$

$$\text{Since} \quad E(Md) = \mu \quad \text{and} \quad V(Md) = 0, as \quad n \to \infty$$

$$\text{For all n}, V(\bar{x}) = \frac{\sigma^2}{n}$$

$$\text{For large n}, \quad V(Md) = \frac{\pi\sigma^2}{2n} = 1.57\frac{\sigma^2}{n}$$

Since $V(\bar{x}) < V(Md)$, we conclude that for normal distribution, sample mean is more efficient estimator for μ than the sample median, for large samples at least.

Example

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.

(a) Find λ. Are t_1 and t_2 are unbiased?

(b) Also find the estimator which is best among t_1, t_2 and t_3.

Example

Let the random sample (X_1, X_2) and X_3 of size 3 is drawn from a population with mean μ and variance σ^2. T_1, T_2 and T_3 are the estimators used to estimate mean value μ, where
$T_1 = X_1 + X_2 - X_3, T_2 = 2X_1 + 3X_3 - 4X_2$ and $T_3 = \frac{1}{3}(\lambda X_1 + X_2 + X_3)$.

1. Are T_1 and T_2 unbiased estimators?
2. FInd the value of λ such that T_3 is unbiased estimator for μ.
3. With this value of λ is T_3 a consistent estimator?
4. Which is the best estimator?

MVUE

Definition 2.6 (*MVUE*)

If a statistic $T = T(x_1, x_2, \ldots, x_n)$, based on sample of size n is such that:
(i) T is unbiased for $\gamma(\theta)$, for all $\theta \in \Theta$ and
(ii) It has the smallest variance among the class of all unbiased estimators of $\gamma(\theta)$, then T is called minimum variance unbiased estimator (***MVUE***) of $\gamma(\theta)$.

T is ***MVUE*** of $\gamma(\theta)$ if,

$$E_\theta(T) = \gamma(\theta) \text{ for all } \theta \in \Theta$$

$$\text{and } Var_\theta(T) \leq Var_\theta(T') \text{ for all } \theta \in \Theta$$

where T' is any other unbiased estimator of $\gamma(\theta)$.

MVUE

Theorem 2.7

An M.V.U.E is unique in the sense that if T_1 and T_2 are M.V.U. estimators for $\gamma(\theta)$, then $T_1 = T_2$ almost surely.

Theorem 2.8

Let T_1 and T_2 be unbiased estimators of $\gamma(\theta)$ with efficiencies e_1 and e_2 respectively and $\rho = \rho_\theta$ be the correlation coefficient between them. Then $\sqrt{e_1 e_2} - \sqrt{(1-e_1)(1-e_2)} \leq \rho \leq \sqrt{e_1 e_2} + \sqrt{(1-e_1)(1-e_2)}$

Corollary 2.9

If we take $e_1 = 1$ and $e_2 = e$, we get $\sqrt{e} \leq \rho \leq \sqrt{e} \implies \rho = \sqrt{e}$

Example

Data on pull-off force (pounds) for connectors used in an automobile engine application are as follows: 79.3, 75.1, 78.2, 74.1, 73.9, 75.0, 77.6, 77.3, 74.6, 75.5, 74.0, 74.7, 75.9, 72.9, 73.8, 74.2, 78.1, 75.4, 76.3, 75.3, 76.2, 74.9, 78.0, 75.1, 76.8.

1. Calculate the point estimate of the mean pull-off force of all connectors in the population. State which estimator you used and why.
2. Calculate the point estimates of the population variance and population standard deviation.

Sufficiency

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

Definition 2.10 (***Sufficiency***)

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

Example

Let $x_1, x_2, \ldots, x_n$ be random sample from a Bernoulli population with the parameter 'p', $0 < p < 1$. Find the sufficient estimator of p.

Solution:

Let $x_1, x_2, \ldots, x_n$ be random sample from a Bernoulli population with the parameter 'p', $0 < p < 1$, *i.e.*,

$$x_i = \left\{ \begin{array}{l} 1, \text{with probability } p \\ 0, \text{with probability } q=(1\text{-}p) \end{array} \right\}$$

$$\text{Then} \quad T = t(x_1, x_2, \ldots, x_n) = x_1 + x_2 + \ldots + x_n \sim B(n,p)$$

$$\therefore \quad P(T = k) = \begin{pmatrix} n \\ k \end{pmatrix} p^k (1-p)^{n-k}$$

The conditional distribution of $(x_1, x_2, \ldots, x_n)$ given T is

$$P[x_1 \cap x_2 \cap \ldots \cap x_n | T = k] = \frac{P[x_1 \cap x_2 \cap \ldots \cap x_n \cap T = k]}{P(T = k)}$$

$$= \left\{ \begin{array}{c} \frac{p^k(1-p)^{n-k}}{\begin{pmatrix} n \\ k \end{pmatrix} p^k(1-p)^{n-k}} = \frac{1}{\begin{pmatrix} n \\ k \end{pmatrix}} \\ 0, \quad \text{if } \sum_{i=1}^n x_i \neq k \end{array} \right\}$$

$$\text{It does not depend on } 'p', T = \sum_{i=1}^{n} x_i, \text{ is sufficient for } 'p'.$$

Fisher-Neyman Factorization Theorem

$T = t(x)$ is sufficient for θ if and only if the joint density function L, of the sample values can be expressed in the form:

$$L = g_\theta[t(x)].h(x)$$

where $g_\theta[t(x)]$ is depends on θ and x only through the value of $t(x)$ and $h(x)$ is independent of θ.

Example

Let $x_1, x_2, \ldots, x_n$ be random sample from a distribution with the *p.d.f*:

$$f(x,\theta) = \theta x^{\theta-1}; 0 < X < 1; \theta > 0$$

Show that $t = \prod_{i=1}^{n} X_i$ is sufficient estimator for θ.

Solution:

$$\begin{aligned}
L(x,\theta) = \prod_{i=1}^{n} f(x_i,\theta) &= \prod_{i=1}^{n} \theta\left(x_i^{\theta-1}\right) \\
&= \theta^n \prod_{i=1}^{n} \left({x_i}^{\theta-1}\right) \\
&= \theta^n \left(\prod_{i=1}^{n} x_i\right)^{\theta} \cdot \frac{1}{\left(\prod_{i=1}^{n} x_i\right)} \\
&= g(t,\theta) \cdot h(x_1, x_2, \ldots, x_n), (say)
\end{aligned}$$

Hence by Factorization theorem,

$$t = \prod_{i=1}^{n} x_i, \quad \text{is sufficient for} \quad \theta.$$

Example

Let $x_1, x_2, \ldots, x_n$ be random sample from $N(\mu, \sigma^2)$ population. Find the sufficient estimator for μ and σ^2.

Solution:

Let us write $\theta = (\mu, \sigma^2)$; $-\infty < \mu < \infty$, $0 < \sigma^2 < \infty$

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

Then

$$=\Big(\frac{1}{\sigma\sqrt{2\pi}}\Big)^n \cdot e^{\left[-\frac{1}{2\sigma^2}\left(\sum_{i=1}^{n} {x_i}^2 - 2\mu\sum x_i + n\mu^2\right)\right]}$$

$$=g_\theta[t(x)] \cdot h(x)$$

$$g_\theta[t(x)] =\Big(\frac{1}{\sigma\sqrt{2\pi}}\Big)^n e^{\left[-\frac{1}{2\sigma^2}\{t_2(x) - 2\mu t_1(x) + n\mu^2\}\right]}$$

$$t(x) =[t_1(x), t_2(x)] = \Big(\sum x_i, \sum {x_i}^2\Big) \ \ and \ \ h(x) = 1$$

Thus $\quad t_1(x) = \sum x_i$ is sufficient for μ *and*

$t_2(x) = \sum {x_i}^2$ is sufficient for σ^2.

Example

Let $Y_1, Y_2, \ldots, Y_n$ be iid random sample from a Poisson distribution with parameter λ. Show that $U = \sum_{i=1}^{n} Y_i$ is a sufficient statistic for λ

Solution:

- Method of Maximum Likelihood Estimation
- Method of Minimum Variance
- Method of Moments
- Method of Least Squares
- Method of Minimum Chi-square
- Method of Inverse Probability

Maximum Likelihood Estimation(MLE)

Definition 3.1 (Likelihood function)

Let $x_1, x_2, \cdots, x_n$ be a random sample of size n from a population with density function $f(x, \theta)$. Then the likelihood function of the sample values $x_1, x_2, \cdots, x_n$, 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 maximizes 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

$$L(\hat{\theta}) > L(\theta) \qquad \forall\, \theta \in \Theta$$
$$\text{i.e., } L(\hat{\theta}) = SupL(\theta) \quad \forall\, \theta \in \Theta$$

$\hat{\theta}$ usually called ***Maximum likelihood Estimator (M.L.E.)***.

Thus $\hat{\theta}$ is the solution, if any of $\dfrac{\mathrm{d}L}{\mathrm{d}\theta} = 0$ and $\dfrac{\mathrm{d}^2 L}{\mathrm{d}\theta^2} < 0$

Since $L > 0$, and $logL$ is a non-decreasing function of L;
L and $logL$ attain their extreme values (maxima or minima) at the same value of $\hat{\theta}$. Then

$$\tfrac{1}{L} \cdot \frac{\mathrm{d}L}{\mathrm{d}\theta} = 0 \implies \frac{\mathrm{d}logL}{\mathrm{d}\theta} = 0$$

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

Example

1. 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.
2. Show that the sample mean $\bar{x}$, is sufficient for estimating the parameter λ of the Poisson distribution.

Solution: 1) The probability function of the Poisson distribution with the parameter λ is given by

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

Likelihood function of a random sample $x_1, x_2, \cdots, x_n$ of n observations from Poisson population is

$$\begin{aligned} L(x|\lambda) &= \prod_{i=1}^{n} f(x_i, \lambda) = \prod_{i=1}^{n} \frac{e^{-\lambda}\lambda^{x_i}}{x_i!} \\ L &= \frac{e^{-n\lambda}\lambda^{\sum_{i=1}^{n} x_i}}{x_1!\, x_2! \cdots x_n!} \end{aligned} \tag{1}$$

$$\begin{aligned} \therefore \quad logL &= -n\lambda + \Big(\sum_{i=1}^{n} x_i\Big)\, log\lambda - \sum_{i=1}^{n} log(x_i!) \\ &= -n\lambda + n\bar{x}\, log\lambda - \sum_{i=1}^{n} log(x_i!) \end{aligned} \tag{2}$$

The likelihood equation for estimating λ is

$$\frac{\mathrm{d}logL}{\mathrm{d}\lambda} = 0 \implies -n + \frac{n\bar{x}}{\lambda} = 0 \implies \lambda = \bar{x}$$

Thus the M.L.E. for λ is the sample mean $\bar{x}$.

The variance of the estimate is given by

$$
\begin{aligned}
\frac{1}{V(\hat{\lambda})} &= E\Big[-\frac{\mathrm{d}^2 logL}{\mathrm{d}\lambda^2}\Big] \\
&= E\Big[-\frac{\mathrm{d}}{\mathrm{d}\lambda}\Big(-n+\frac{n\bar{x}}{\lambda}\Big)\Big] \\
&= E\Big[-\Big(-\frac{n\bar{x}}{\lambda^2}\Big)\Big] = \frac{n}{\lambda^2}E(\bar{x}) \\
&= \frac{n}{\lambda} \\
\therefore \quad V(\hat{\lambda}) &= \frac{\lambda}{n}
\end{aligned} \tag{3}
$$

2) For the Poisson distribution with parameter λ, we have

$$\begin{aligned} \frac{\mathrm{d}}{\mathrm{d}\lambda} logL &= -n + \frac{n\bar{x}}{\lambda} \\ &= n\Big(\frac{\bar{x}}{\lambda} - 1\Big) = \psi(\bar{x}, \lambda), \text{a function of } \bar{x} \text{ and } \lambda \text{ only} \end{aligned} \tag{4}$$

$\bar{x}$ is sufficient for estimating λ.

Example

In a random sampling from a normal population $N(\mu, \sigma^2)$, find the maximum likelihood estimators for

1. μ when σ^2 is known,
2. σ^2 when μ is known

Solution:

$$\text{Given,} \quad X \sim N(\mu, \sigma^2) \quad \text{then}$$

$$\begin{aligned} L &= \prod_{i=1}^{n} \Big[\frac{1}{\sigma\sqrt{2\pi}} e^{-\frac{1}{2\sigma^2}(x_i-\mu)^2}\Big] \\ &= \Big(\frac{1}{\sigma\sqrt{2\pi}}\Big)^n e^{\left\{-\sum_{i=1}^{n}(x_i-\mu)^2/2\sigma^2\right\}} \qquad (5) \\ logL &= \frac{-n}{2} log(2\pi) - \frac{n}{2} log(\sigma^2) - \frac{1}{2\sigma^2}\sum_{i=1}^{n}(x_i-\mu)^2 \end{aligned}$$

Case (i). When σ^2 is known, the likelihood equation for estimating μ is

$$\begin{aligned} \frac{\partial}{\partial\mu} logL = 0 &\implies -\frac{1}{2\sigma^2}\sum_{i=1}^{n} 2(x_i-\mu)(-1) = 0 \\ \sum_{i=1}^{n}(x_i-\mu) = 0 &\implies \sum_{i=1}^{n} x_i - n\mu = 0 \qquad (6) \\ \implies \qquad & \hat{\mu} = \frac{1}{n}\sum_{i=1}^{n} x_i = \bar{x} \end{aligned}$$

Hence the M.L.E. for μ is the sample mean $\bar{x}$.

Case(*ii*): when μ is known, the likelihood equation for estimating σ^2 is

$$\begin{aligned} Since \quad logL &= \frac{-n}{2} log(2\pi) - \frac{n}{2} log(\sigma^2) - \frac{1}{2\sigma^2}\sum_{i=1}^{n}(x_i-\mu)^2 \\ \frac{\partial}{\partial\sigma^2} logL = 0 &\implies \frac{-n}{2}\times\frac{1}{\sigma^2} + \frac{1}{2\sigma^4}\sum_{i=1}^{n}(x_i-\mu)^2 = 0 \\ n - \frac{1}{\sigma^2}\sum_{i=1}^{n}(x_i-\mu)^2 &= 0, \\ i.e., \quad \hat{\sigma^2} &= \frac{1}{n}\sum_{i=1}^{n}(x_i-\mu)^2 \ = \ s^2 \ = \ \text{sample variance} \end{aligned} \tag{7}$$

Hence the M.L.E. for σ^2 is the sample variance s^2.

Important Note:

$$\begin{aligned} E(\hat{\mu}) &= E(\bar{x}) = \mu \\ E(\hat{\sigma^2}) &= E(s^2) \neq \sigma^2 \end{aligned} \tag{8}$$

Hence the maximum likelihood estimators (M.L.E.) need not necessarily be *unbiased*.

Example

Obtain the maximum likelihood estimate of θ in

$$f(x,\theta) = (1+\theta)x^{\theta}\,; \quad 0 < x < 1$$

based on an independent sample of size n. Examine whether this estimate is sufficient for θ.

Solution:

$$\begin{aligned}
L(x,\theta) &= \prod_{i=1}^{n} f(x_i,\theta) = (1+\theta)^n \cdot \Big(\prod_{i=1}^{n} x_i\Big)^{\theta} \\
\implies logL &= n\,log(1+\theta) + \theta\sum_{i=1}^{n} log\,x_i \\
\frac{\mathrm{d}logL}{\mathrm{d}\theta} &= \frac{n}{1+\theta} + \sum_{i=1}^{n} log\,x_i = 0 \\
\implies n + \theta\sum_{i=1}^{n} log\,x_i &+ \sum_{i=1}^{n} log\,x_i = 0 \\
\therefore \quad \hat{\theta} &= \frac{-n}{\sum_{i=1}^{n} log\,x_i} - 1 = \frac{-n}{log\Big(\prod_{i=1}^{n} x_i\Big)} - 1 \\
\text{Also} \quad L(x,\theta) &= \Big\{(1+\theta)^n \cdot \Big(\prod_{i=1}^{n} x_i\Big)^{\theta-1}\Big\} \cdot \Big(\prod_{i=1}^{n} x_i\Big)
\end{aligned} \tag{9}$$

Hence by the factorization theorem, $T = \Big(\prod_{i=1}^{n} x_i\Big)$ is a sufficient statistic for θ

and $\hat{\theta}$ being a one to one function of sufficient statistic $\Big(\prod_{i=1}^{n} x_i\Big)$ is also sufficient for θ.

Example

Let X be a Bernoulli random variable with the probability mass function

$$P(X = x) = \begin{cases} p^x(1-p)^{1-x}, & x = 0, 1 \\ 0, & Otherwise \end{cases}$$

Estimate the Maximum Likelihood Estimator of the parameter p.

MVUE

Definition 3.2 (*Method of Minimum Variance (MVUE)*)

If a statistic $T = T(x_1, x_2, \ldots, x_n)$, based on sample of size n is such that:
(i) T is unbiased for $\gamma(\theta)$, for all $\theta \in \Theta$ and
(ii) It has the smallest variance among the class of all unbiased estimators of $\gamma(\theta)$, then T is called minimum variance unbiased estimator (***MVUE***) of $\gamma(\theta)$.

T is ***MVUE*** of $\gamma(\theta)$ if,

$$E_\theta(T) = \gamma(\theta) \text{ for all } \theta \in \Theta$$

$$\text{and } Var_\theta(T) \leq Var_\theta(T') \text{ for all } \theta \in \Theta$$

where T' is any other unbiased estimator of $\gamma(\theta)$.

Cramer-Rao Inequality

If t is an unbiased estimator for $\gamma(\theta)$, a function of parameter θ, then

$$Var(t) \geq \frac{\left[\frac{d}{d\theta}(\gamma(\theta))\right]^2}{E\left[\frac{\partial}{\partial\theta}logL\right]^2} = \frac{\left[\gamma'(\theta)\right]^2}{-E\left(\frac{\partial^2}{\partial\theta^2}logL\right)} = \frac{\left[\gamma'(\theta)\right]^2}{I(\theta)}$$

where $I(\theta)$ is the information on θ, supplied by the sample.

In other words, Cramer-Rao inequality provides a lower bound $\frac{\left[\gamma'(\theta)\right]^2}{I(\theta)}$, to the variance of an unbiased estimator of $\gamma(\theta)$.

Assumptions & Regularity conditions

- We assume that there is only a single parameter θ which is unknown.
- Random variable is continuous.
- In case of discrete random variables can be dealt with similarly on replacing the multiple integrals by appropriate multiple sums.

Corollary

If t is an unbiased estimator for θ, *i.e.*

$$E(t) = \theta \implies \gamma(\theta) = \theta \implies \gamma'(\theta) = 1,$$

then by the cramer-rao in-equality, we get

$$Var(t) \geq \frac{1}{E\left[\frac{\partial}{\partial\theta} logL\right]^2} = \frac{1}{-E\left(\frac{\partial^2}{\partial\theta^2} logL\right)} = \frac{1}{I(\theta)}$$

Where $I(\theta) = E\left[\left(\frac{\partial}{\partial\theta} logL\right)^2\right]$ is called by R. A. Fisher as the amount of information on θ supplied by the sample $(x_1, x_2, \ldots, x_n)$ and its reciprocal $1/I(\theta)$, as the information limit to the variance of estimator $t = t(x_1, x_2, \ldots, x_n)$.

Example

Find the Cramer-Rao Lower Bound for unbiased estimators of the parameter λ of Poisson distribution.

Solution: Suppose that $X_1, X_2, \ldots, X_n$ is a set of independent random variables each arising from the same Poisson distribution with parameter λ.

Given, the p.m.f. of the Poisson distribution:

$$\begin{aligned} f(x|\lambda) &= \frac{\lambda^x e^{-\lambda}}{x!} \\ logf(x|\lambda) &= xlog\lambda - \lambda - log(x!) \end{aligned} \tag{10}$$

Differentiate with respect to the parameter λ:

$$\begin{aligned} \frac{\partial}{\partial\lambda} logf(x|\lambda) &= \frac{x}{\lambda} - 1 \\ \left\{\frac{\partial}{\partial\lambda} logf(x|\lambda)\right\}^2 &= \left(\frac{x}{\lambda} - 1\right)^2 \\ &= \frac{1}{\lambda^2}(x-\lambda)^2 \\ E\left[\left\{\frac{\partial}{\partial\lambda} logf(x|\lambda)\right\}^2\right] &= \frac{1}{\lambda^2} E\Big\{(X-\lambda)^2\Big\} \\ &= \frac{1}{\lambda^2} V(X) = \frac{1}{\lambda^2}\lambda \\ I(\theta) &= \frac{1}{\lambda} \end{aligned} \tag{11}$$

Since, $E(X) = V(X) = \lambda$ and hence, the CRLB for the parameter of Poisson distribution is the reciprocal of n times the value of $I_n(\theta)$, i.e., $\frac{1}{I_n(\theta)} = \frac{1}{n/\lambda} = \frac{\lambda}{n}$.

Example

Find the Cramer-Rao Lower Bound for unbiased estimators of the parameter p of Bernoulli distribution.

Minimum Variance Bound Estimator(*MVB*)

An unbiased estimator t of $\gamma(\theta)$ for which Cramer-Rao lower bound is attained is called a ***minimum variance bound(MVB)*** estimator.

$$I(\theta) = E\Big\{\Big(\frac{\partial}{\partial\theta} logL\Big)^2\Big\} = -E\Big(\frac{\partial^2}{\partial\theta^2} logL\Big)$$

$$I(\theta) = n\Big\{\frac{\partial}{\partial\theta} logf(x,\theta)\Big\}^2 = -n\Big(\frac{\partial^2}{\partial\theta^2} logf\Big)$$

A necessary and sufficient condition for an unbiased estimator t to attain the lower bound of its variance is given by

$$\frac{\partial}{\partial\theta}logL = \frac{t-\gamma(\theta)}{\lambda(\theta)}$$

where λ is a constant independent of $(x_1, x_2, \ldots, x_n)$ but may depend on θ.

Hence if the likelihood function L is expressible in the above form, then

(i) t is unbiased estimator of $\gamma(\theta)$
(ii) Minimum Variance Bound(MVB) estimator (t) for $\gamma(\theta)$ exists, and
(iii) $Var(t) = \left|\frac{\gamma^{'}(\theta)}{A(\theta)}\right| = \left|\gamma^{'}(\theta) \cdot \lambda(\theta)\right|$, where $A(\theta) = \frac{1}{\lambda(\theta)}$

Example

Obtain the Minimum Variance Bound estimator for μ in normal population $N(\mu, \sigma^2)$, where σ^2 is known.

Solution: If $x_1, x_2, \ldots, x_n$ is a random sample of size n drawn from the normal population, then

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

where k is a constant independent of μ, (σ being known).

$$\begin{aligned} \frac{\partial}{\partial\mu} logL &= -\frac{1}{2\sigma^2}\sum_{i=1}^{n} 2(x_i-\mu)(-1) \\ \frac{\sum_{i=1}^{n}(x_i-\mu)}{\sigma^2} &= \frac{\sum_{i=1}^{n} x_i - n\mu}{\sigma^2} \\ \implies \frac{\partial}{\partial\mu} logL &= \frac{\bar{x}-\mu}{\sigma^2/n} \end{aligned} \tag{12}$$

Hence $\bar{x}$ is a MVB unbiased estimator for μ and $V(\hat{\mu}) = V(\bar{x}) = \frac{\sigma^2}{n}$

Example

Find the Minimum Variance Bound estimator of θ , where the random sample is drawn from exponential population with the *p.d.f*

$$f(x,\theta) = \begin{cases} \frac{1}{\theta}\, e^{-x/\theta}, 0 < x < \infty, \theta > 0 \\ 0, otherwise \end{cases}$$

Solution: If $x_1, x_2, \ldots, x_n$ is a random sample of size n drawn from the exponential population with *p.d.f.*, then

$$L = \prod_{i=1}^{n} f(x_i,\theta) = \frac{1}{\theta^n} exp\Big(\frac{-\sum_{i=1}^{n} x_i}{\theta}\Big)$$

$$\implies \quad logL = -nlog\theta - \frac{1}{\theta}\sum_{i=1}^{n} x_i$$

$$\therefore \quad \frac{\partial}{\partial\theta} logL = -\frac{n}{\theta} + \frac{1}{\theta^2}\sum_{i=1}^{n} x_i$$

$$\implies \quad \frac{\partial^2}{\partial\theta^2} logL = \frac{n}{\theta^2} - \frac{2}{\theta^3}\sum_{i=1}^{n} x_i$$

$$Thus \qquad I(\theta) = -E\Big(\frac{\partial^2}{\partial\theta^2} logL\Big) = -\frac{n}{\theta^2} + \frac{2}{\theta^3}\sum_{i=1}^{n} E(x_i)$$

In sampling from exponential population, we have

$$E(X) = \theta \quad \text{and} \quad Var(X) = \theta^2$$

$$\therefore \quad I(\theta) = -\frac{n}{\theta^2} + \frac{2}{\theta^3}\sum_{i=1}^{n}(\theta) = -\frac{n}{\theta^2} + \frac{2}{\theta^3}\, n\theta = \frac{n}{\theta^2}$$

$$\text{Also} \qquad \gamma(\theta) = \theta \implies \gamma^{'}(\theta) = 1$$

Hence Cramer Rao lower bound to the variance of an unbiased estimator of θ is:

$$\frac{[\gamma'(\theta)]^2}{I(\theta)} = \frac{1}{(n/\theta^2)} = \frac{\theta^2}{n}$$

Consider the estimator $\bar{X} = \frac{1}{n}\sum_{i=1}^{n} x_i$, we have

$$E(\bar{X}) = \frac{1}{n}\sum_{i=1}^{n} E(x_i) = \frac{1}{n}\sum_{i=1}^{n}(\theta) = \frac{n\theta}{n} = \theta$$

$$\implies \quad \bar{X} \text{ is an unbiased estimator of } \theta.$$

$$\text{Also} \quad Var(\bar{X}) = \frac{\sigma^2}{n} = \frac{Var(X)}{n} = \frac{\theta^2}{n}$$

Thus we see that $Var(\bar{X})$ coincided with the Cramer-Rao lower bound.

Hence $\bar{X}$, the sample mean is an MVB unbiased estimator, for θ.

Aliter: A more convenient way of doing this problem is as follows:

$$\frac{\partial}{\partial\theta} logL = -\frac{n}{\theta} + \frac{1}{\theta^2}\sum_{i=1}^{n} x_i = \frac{\sum_{i=1}^{n} x_i - n\theta}{\theta^2}$$
$$= \frac{\bar{X} - \theta}{(\theta^2/n)} = \frac{\bar{X} - \theta}{\lambda(\theta)}$$

Hence $\bar{X}$ is an MVB unbiased estimator of θ and $Var(\bar{X}) = \lambda(\theta) = \frac{1}{n}\theta^2$

Rao-Blackwell theorem

Let X and Y be random variables such that

$$E(Y) = \mu \quad \text{and} \quad Var(Y) = \sigma_Y^2 > 0$$

Let $\quad E(Y|X = x) = \phi(x)$,
Then

$$\begin{aligned} (i) &\quad E(\phi(X)) = \mu \\ (ii) &\quad Var[\phi(X)] \leq Var(Y) \end{aligned}$$

Method of Moments

Let $f(x;\theta_1,\theta_2,\ldots,\theta_k)$ be the probability density function of the parent population with k parameters $\theta_1,\theta_2,\ldots,\theta_k$. If $\mu_r^{'}$ denotes the r^{th} moment about origin then

$$\mu_r^{'} = \int_{-\infty}^{\infty} x^r f(x;\theta_1,\theta_2,\ldots,\theta_k)dx \quad ;(r=1,2,\ldots,k)$$

In general $\mu_1^{'},\mu_2^{'},\ldots,\mu_k^{'}$ will be the functions of the parameters $\theta_1,\theta_2,\ldots,\theta_k$.

Let $x_i, i = 1, 2, \ldots, n$ be a random sample of size n from the given population. The method of moments consists in solving the $k-$ equations for $\theta_1, \theta_2, \ldots, \theta_k$ in terms of $\mu_1^{'}, \mu_2^{'}, \ldots, \mu_k^{'}$ and then replacing these moments $\mu_r^{'}; r = 1, 2, \ldots, k$ by the sample moments.

$$\begin{aligned}\hat{\theta}_i &= \theta_i\left(\hat{\mu_1'}, \hat{\mu_2'}, \ldots, \hat{\mu_k'}\right)\\ &= \theta_i(m_1^{'}, m_2^{'}, \ldots, m_k^{'}) \quad ; \quad i = 1, 2, \ldots, k\end{aligned}$$

where m_i is the i^{th} moment about origin in the sample.

Remarks:

1. Let $x_1, x_2, \ldots, x_n$ be a random sample of size n from a population with p.d.f. $f(x,\theta)$. Then $X_i, (i = 1, 2, \ldots, n)$ are *i.i.d.* $\implies$ $X_i^r, i = (1, 2, \ldots, n)$ are *i.i.d.* random variables.
Hence if $E(X_i^r)$ exists, then by *W.L.L.N.*, we get

$$\frac{1}{n}\sum_{i=1}^{n} x_i^r \xrightarrow{P} E(X_i^r) \implies m_r^{'} \xrightarrow{P} \mu_r^{'} \tag{13}$$

Hence the sample moments are consistent estimators of the corresponding population moments.

2. It has been shown that under quite general conditions, the estimates obtained by the method of moments are asymptotically normal but not in general efficient.

3. Generally the method of moments yields less efficient estimators than those obtained from the principle of maximum likelihood.

Example

The sample values from population with probability density function

$$f(x) = (1+\theta)x^{\theta}; 0 < x < 1, \theta > 0,$$

are given below:
0.46, 0.38, 0.61, 0.82, 0.59, 0.53, 0.53, 0.72, 0.44, 0.59, 0.6
Find the estimate of θ by using (i) method of moments (ii) method of maximum likelihood estimation.

Solution:

$$\mu_r^{'} = E(X^r) = = \int_0^1 x^r (1+\theta)x^{\theta}\, dx$$
$$= \int_0^1 (1+\theta)x^{r+\theta}\, dx$$
$$\mu_r^{'} = \frac{1+\theta}{r+\theta+1}$$
$$\mu_1^{'} = E(X^{'}) = \bar{X} = \frac{1+\theta}{2+\theta}$$
$$\bar{X} = 1 - \frac{1}{\theta+2}$$
$$\hat{\theta} = \frac{1}{1-\bar{X}} - 2$$

Example

Estimate α and β in the case of *Pearsons Type III* distribution by the method of moments.

$$f(x;\alpha,\beta) = \frac{\beta^\alpha}{\Gamma(\alpha)} x^{\alpha-1} e^{-\beta x} \quad ; \ 0 \leq x < \infty$$

Solution:

$$\mu_r^{'} = \int_{-\infty}^{\infty} x^r f(x;\theta)\, dx \quad ; \ 0 \leq x < \infty$$

$$\begin{aligned}
\mu_r^{'} &= \int_0^{\infty} x^r f(x;\alpha,\beta)\, dx \\
&= \frac{\beta^\alpha}{\Gamma(\alpha)} \int_0^{\infty} x^r x^{\alpha-1} e^{-\beta x}\, dx \\
&= \frac{\beta^\alpha}{\Gamma(\alpha)} \frac{\Gamma(\alpha+r)}{\beta^{\alpha+r}} \\
\mu_r^{'} &= \frac{\Gamma(\alpha+r)}{\Gamma(\alpha)\beta^r}
\end{aligned}$$

$$\mu_1' = \frac{\Gamma(\alpha+1)}{\Gamma(\alpha)\beta} = \frac{\alpha}{\beta}$$

$$\mu_2' = \frac{\Gamma(\alpha+2)}{\Gamma(\alpha)\beta^2} = \frac{(\alpha+1)\alpha}{\beta^2}$$

$$\frac{\mu_2'}{\mu_1'^2} = \frac{\alpha+1}{\alpha} = \frac{1}{\alpha} + 1$$

$$\implies \quad \alpha = \frac{\mu_1'^2}{\mu_2' - \mu_1'^2}$$

$$\beta = \frac{\alpha}{\mu_1'} = \frac{\mu_1'}{\mu_2' - \mu_1'^2}$$

$$Hence \quad \hat{\alpha} = \frac{m_1'^2}{m_2' - m_1'^2} \quad and$$

$$\hat{\beta} = \frac{m_1'}{m_2' - m_1'^2}$$

where m_1' and m_2' are the sample moments.

Example

Let X be a discrete random variable with the following probability mass function, with the population parameter $0 \leq \theta \leq 1$:

X	0	1	2	3
$P(X=x)$	$2\theta/3$	$\theta/3$	$2(1-\theta)/3$	$(1-\theta)/3$

The following 10 independent observations were taken from such a distribution: $(3, 0, 2, 1, 3, 2, 1, 0, 2, 1)$. Find the Maximum likelihood estimate of the parameter θ.

Example

Let $X_1, X_2, \ldots, X_n$ be a random sample from a population with the following probability density function:

$$f(x|\theta) = \frac{1}{\theta}\ x^{(1-\theta)/\theta} \quad ; \quad 0 < X < 1, \quad 0 < \theta < \infty$$

a) Find the MLE of the parameter θ. Also, Calculate an estimate using this estimator when $x_1 = 0.10, x_2 = 0.22, x_3 = 0.54, x_4 = 0.36$.
b) Obtain a method of moments estimator for θ. Calculate an estimate using this estimator when $x_1 = 0.10, x_2 = 0.22, x_3 = 0.54, x_4 = 0.36$..

Interval Estimation

Let $x_i, (i = 1{,}2, \cdots, n)$ be a random sample of n observations from a normal population involving a single unknown parameter θ.

Let $f(x, \theta)$ be the probability function of the parent distribution from which the sample is drawn and let us assume that the distribution is continuous.

Let $t = t(x_1, x_2, \cdots, x_n)$, a function of the sample values be an estimate of the population parameter θ with the sampling distribution given by $g(t, \theta)$.

Let us select a small value of $\boldsymbol{\alpha(5\%\ or\ 1\%)}$ and determine the two constants say, c_1 and c_2 such that:

$$P(c_1 < \theta < c_2 \mid t) = 1 - \alpha \tag{14}$$

is called a $100(1-\alpha)\%$ confidence interval for the unknown parameter θ and

$0 \leq \alpha \leq 1$. Where c_1 and c_2 are called the lower and upper ***Confidence limits*** or ***Fiducial limits*** and $\boldsymbol{(1-\alpha)}$ is called the ***Confidence coefficient***.

There is a probability of $(1-\alpha)$ selecting a sample for which the CI will contain the true value of the parameter θ. Once we have selected the sample, so that $X_1 = x_1, X_2 = x_2 \dots X_n = x_n$, and computed l and u, the resulting confidence interval for θ is

$$l \leq \theta \leq u$$

Thus if we take $\alpha = 0.05(or\ 0.01)$, we shall get ***95%(or 99%)*** confidence

limits.

How to find c_1 and c_2 ?

Let T_1 and T_2be two statistics such that

$$P(T_1 > \theta) = \alpha_1 \tag{15a}$$
$$P(T_2 < \theta) = \alpha_2 \tag{15b}$$

Where α_1 and α_2 are constants independent of θ.

$$\implies P(T_1 < \theta < T_2) = 1 - \alpha, \quad \text{where } \alpha = \alpha_1 + \alpha_2.$$

Statistics T_1 and T_2 are defined in $(15a)$ and $(15b)$ may be taken as c_1 and c_2 defined in (14).

Example

Let $X_1, X_2, \ldots, X_n$ be random sample of size n drawn from a normal population with mean $\boldsymbol{\mu}$ and variance $\boldsymbol{\sigma^2}$. Obtain the $100(1-\alpha)\%$ confidence interval for $(i)\mu$ and $(ii)\sigma^2$.

Choose a Sample Size

We can choose $\boldsymbol{n}$ to be $\mathbf{100(1-\boldsymbol{\alpha})}\%$ confidence that the error in estimating $\boldsymbol{\mu}$ is less than a specified error $\boldsymbol{E}$. The appropriate sample size is

$$n = \left(\frac{Z_{\alpha/2}\sigma}{E}\right)^2$$

Example

Using a temperature of $100^o F$ and a power input of 550W, the following 10 measurements of thermal conductivity were obtained.

41.60, 41.48, 42.34, 41.95, 41.86, 42.18, 41.72, 42.26, 41.81, 42.04

Obtain the point estimate and 95% and 99% interval estimate of sample mean.

Thank You