ADVANCED MATHEMATICAL STATISTICS.

INSTRUCTOR:- Dr. Masood Anwar

STATISTICAL INFERENCE:

* Course Objectives:-

To provide a grounding in the theoratical foundations of statistical inference and, in particular, to introduce the theory underlying statistical and hypothesis testing.

On successful completion of the course students should be able to underpin statistical design in estimation and hypothesis testing which has fundamental applications to all fields in which statistical investigations are planned or data are analyzed. The important area including engineering, industry, physical sciences, medicine, biology, economics, finance, psychology and social sciences.

* Course Contents:-

Estimation: Criteria and methods for estimators, Classical and newer methods of Estimation, Deriving Estimators (Bayes Methods, MLE), Cramer-Rao and its extension, Bias reduction by Jack Knifing, Rao-Balckwellization, Baser's theorem, Estimation in parametric and Non Parametric Methods. Testing Hypothesis: Parametric methods. Neyman-Pearson Lemma. Uniformly most powerful tests. Unbiased tests. Locally optimal tests. Large simple theory, asymptotically best procedures. Testing under Nuisance Parameters, Review of tests for Normal Distribution, Confidence tests including Construction properties, Asymptotic confidence tests, Boot-strap, Confidence sets, Simultaneous confidence Intervals.

★ **Books:-**

1:- Kendall's Advanced theory of statistics by J. Keithy ord, Maurice Kendall, Steven Arnold.
2:- The foundations of statistics by Leonard J-Savage
3:- Introduction to Mathematical statistics (6th Edition) by Rober V. Hogg and Craig and Mckean.

★★ ——— ★★

MUHAMMAD TAHIR WATTOO★★★
FA15-RMT-00
SP18-PMT-005

PhD. MATHEMATICS CU ISLAMABAD
03448563284

Statistical Inference

- Estimation of Parameters
  - Point Estimation
    - Methods of Moments
    - Bay's Estimation
    - Methods of Maximum Likelihood Estimation
  - Interval Estimation
- Testing of Hypothesis
  - Neyman Pearson
  - Optimal Test

★ Statistics:-

The term Statistics refers to, "Listing of facts to a systematic methods of arranging and describing the data and finally the science of Inferring generalities from specific observations."

★ Purpoose:-

We will take decisions about the population parameters on the basis of informations contained in the sample

★ Population:-

Agrigate of material relavant to some characteristics of interest is called Population. Size of Population is Denoted by N.

★ Sample:-

The most representive part of the population is called sample. (The subset of Population) the size of sample is denoted by 'n'.

★ A population can be finite or infinite

Parameter	Statistics
Relates to Population	Relates to sample
μ	$\bar{x}$
σ	s
P	$\hat{p}$
R	$\hat{R}$

★ Parameter is a characteristic of Population.

★ Statistics is a characteristic of Sample.

★ Random Experiment:-

An experiment for which the outcome can not be determined or predicted with certanity, but the experiment is of such a nature that the collection of every possible outcome can be described prior to its performance is called a Random Experiment. e.g.

★ Throwing a die.

★ Tossing a coin.

★ Drawing a card from a standard deck of playing cards.

★ Sample Space:-

A set S of all possible outcomes of a random experiment is called a sample space and is denoted by 'S'.

The elements of S are called sample points.

★ Event:-

Certain subsets of S are called events.

★ Random Variable:-

A random variable (r.v) is a function from a sample space S into real numbers.

Let the random experiment be the toss of a coin and let the sample space associated with experiment be

$S = \{s : s \text{ is } T \text{ or } H\}$, where

$T = \text{Tail}$, $H = \text{Head}$

Let X be a function s.t

$X(s) = 0$ if s is tail.

& $X(s) = 1$ if s is head.

then X is a real valued function denoted on sample space S which takes us from S to space of real numbers

$$X = \{0, 1\}$$

Examples:-

1:- If two coins ar tossed then sample space S is

$$S = \{HH, HT, TH, TT\}$$

Let random variable X = no. of tail

$\Rightarrow X = \{0, 1, 2\}$

2:- If two dice are thrown then

$$S = \begin{Bmatrix} 1,1 & 1,2 & 1,3 & 1,4 & 1,5 & 1,6 \\ 2,1 & 2,2 & 2,3 & 2,4 & 2,5 & 2,6 \\ 3,1 & 3,2 & 3,3 & 3,4 & 3,5 & 3,6 \\ 4,1 & 4,2 & 4,3 & 4,4 & 4,5 & 4,6 \\ 5,1 & 5,2 & 5,3 & 5,4 & 5,5 & 5,6 \\ 6,1 & 6,2 & 6,3 & 6,4 & 6,5 & 6,6 \end{Bmatrix}$$

Let X = sum of dots that may turn up

$\Rightarrow X = \{2, 3, 4, \cdots, 12\}$

3:- If a die is thrown

$S = \{1, 2, 3, 4, 5, 6\}$

X = face of die $= \{1, 2, 3, 5, 6\}$

Let A = event (even number appears)

$\Rightarrow A = \{2, 4, 6\}$

★ Probability Distribution:-

If we associate the probability together with random variable and write in tabular form is called probability Distribution (P.d).

Note:- Here if we define r.v, then we assining probability against each value.

Example:- If three coins are tossed find the probability distribution of number of heads.

Solution

$S = \{HHH, HHT, HTH, THH, TTH, THT, HTT, TTT\}$

Let r.v X = number of heads $= \{0, 1, 2, 3\}$

$P(X=0) = \frac{1}{8}$, $P(X=1) = \frac{3}{8}$

$P(X=2) = \frac{3}{8}$; $P(X=3) = 1/8$

Prob. Distribution

X	P(X)
0	1/8
1	3/8
2	3/8
3	1/8

→ This tabular form is called probability distribution.

★ **The Distribution Function:-** (c.d.f Comulative Dis. fn)

The c.d.f of a r.v X is denoted and defined as

$$F_X(x) = P(X \leq x)$$

Note:- It can be defined by using probability distribution.

Example:- From last example find c.d.f of X.

Solution

X = number of heads

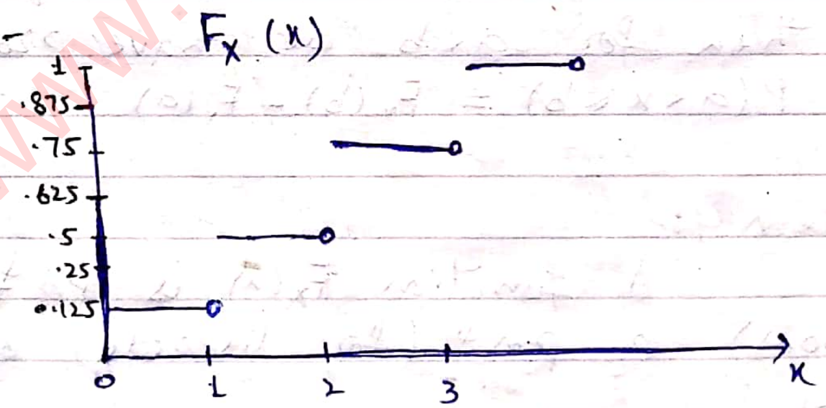
Prob. Distribution

X	$P(x)$	C.F
$x=0$	$1/8$	$1/8$
$x=1$	$3/8$	$1/8 + 3/8 = 4/8$
$x=2$	$3/8$	$4/8 + 3/8 = 7/8$
$x=3$	$1/8$	$7/8 + 1/8 = 1$

As $F_X(x) = P(X \leq x)$

$$= \begin{cases} 0 & x < 0 \text{ or } 0 > x \\ 1/8 & 0 \leq x < 1 \\ 4/8 & 1 \leq x < 2 \\ 7/8 & 2 \leq x < 3 \\ 1 & x \geq 3 \end{cases}$$

Plot :-



Theorem 1:-
The function $F_X(x)$ is a c.d.f iff the following conditions holds

1:- $\lim_{x \to -\infty} F(x) = 0$ (lowest limit)

and $\lim_{x \to +\infty} F(x) = 1$ (upper limit).

2:- $F(x)$ is a non-decreasing function of x

3:- $F(x)$ is a right continuous. i.e

$$\lim_{x \to x_0} F(x) = F(x_0)$$

Theorem 2:-
For any random variable X,

$$P(X = x) = F_X(x) - F_X(x^-) \quad , \forall x \in \mathbb{R}$$

Theorem 3:-
Let X = r.v with c.d.f $F_X(x)$, then for $a < b$ we have prob that $P(a < x \leq b) = F_X(b) - F_X(a)$

Theorem 4:-
A function $F_X(x)$ is pdf (for continuous) or pmf (for discrete) of r.v X iff

1:- $f_X(x) \geq 0 \quad \forall x$

2:- $\sum f_X(x) = 1$ for Discrete

or $\int_{-\infty}^{\infty} f_X(x)\, dx = 1$ for continuous

Question:- For each of the following find the constant c so that P(x) satisfies the condition of being a p.m.f of 1 r.v. X

a) $$P(x) = c\left(\frac{2}{3}\right)^x \qquad x = 1,2,3,\dots$$

Solution

Since P(x) is p.m.f

So, $$\sum_x P(x) = 1$$

$$\Rightarrow c\sum\left(\frac{2}{3}\right)^x = 1$$

$$\Rightarrow c\left[\frac{2}{3} + \left(\frac{2}{3}\right)^2 + \left(\frac{2}{3}\right)^3 + \cdots\right] = 1$$

$$\Rightarrow \frac{2}{3}c\left[1 + \frac{2}{3} + \left(\frac{2}{3}\right)^2 + \cdots\right] = 1$$

$$\Rightarrow \frac{2}{3}c\left[\frac{1}{1-2/3}\right] = 1$$

$$\Rightarrow c(3) = \frac{3}{2} \qquad \Rightarrow \boxed{c = \frac{1}{2}}$$

(b) If $P(x) = \frac{1}{3}$ $\quad x = -1, 0, 1$

find c.d.f $F_X(x)$

Solution

Probability Distribution

x	P(x)	c.f
−1	1/3	1/3
0	1/3	2/3
1	1/3	1

$$F_X(x) = \begin{cases} 0 & x < -1 \\ 1/3 & -1 \leq x < 0 \\ 2/3 & 0 \leq x < 1 \\ 1 & 1 \leq x \end{cases}$$

Question:- Given

$$F_X(x) = \begin{cases} 0 & x < -1 \\ \frac{x+2}{4} & -1 < x < 1 \\ 1 & x \geq 1 \end{cases}$$

find a) $P(-\frac{1}{2} < x < \frac{1}{2})$

b) $P(X = 0)$

c) $P(X = 1)$

d) $P(X : 2 \leq x \leq 3)$

Solution

a) By theorem 3:

$$P(-\tfrac{1}{2} < x < \tfrac{1}{2}) = F_X(\tfrac{1}{2}) - F_X(-\tfrac{1}{2})$$

$$= \frac{\frac{1}{2}+2}{4} - \frac{-\frac{1}{2}+2}{4}$$

$$= \frac{1}{4}$$

b) By theorem 2:

$$P(X=0) = F_X(0) - F_X(0^-)$$

$$= \frac{2}{4} - \frac{2}{4}$$

$$= 0$$

c) By therem 2:

$$P(X=1) = F_X(1) - F_X(1^-)$$

$$= 1 - \frac{1+2}{4}$$

$$= 1/4$$

d) By Theorem 3:

$$P(2 \leq x \leq 3) = F_X(3) - F_X(2)$$

$$P(2 \le x \le 3) = 1 - 1$$
$$= 0$$

* Discrete Random Variable:-

A r.v is said to be discrete random variable if its space is either finite or countable.

A set D is said to be countable if its elements can be listed; i.e there is one-to-one correspondence b/w D and positive integers.

* Transformations:-

We have a random variable X and we know its distribution. We are interested in another variable $Y = g(X)$. In particular, we want to determine the distribution of Y. Assume that X is a discrete variable with space D_X. Then the space of Y is

$$D_Y = \{ g(x) ; \ x \in D_X \}$$

$$P_Y(y) = P(Y = y) = P(g(X) = y) = P(X = g^{-1}(y))$$
$$= P_X(g^{-1}(y))$$

Example:- Let X have pmf

$$P_X(x) = \frac{3!}{x!(3-x)!} \left(\frac{2}{3}\right)^x \left(\frac{1}{3}\right)^{3-x} ; \ x = 0, 1, 2, 3$$

find the p.m.f of $Y = X^2$

Solution

$$P_Y(y) = P(Y=y) = P(X^2=y)$$
$$= P(X=\sqrt{y})$$
$$\Rightarrow P_Y(y) = \frac{3!}{\sqrt{y}!\,(3-\sqrt{y})!}\left(\frac{2}{3}\right)^{\sqrt{y}}\left(\frac{1}{3}\right)^{3-\sqrt{y}} \quad ; y = 0, 1, 4, 9$$

★ Continuous Random Variable:-

We say a r.v is a continuous random variable if its cumulative distribution function (c.d.f) $F_X(x)$ is a continuous function for $x \in \mathbb{R}$

$$\frac{d}{dx}F_X(x) = f_X(x)$$

$f_X(x)$ is the p.d.f and it satisfies the two properties

i) $f_X(x) \geq 0$ ii) $\int_{-\infty}^{\infty} f_X(x)\,dx = 1$

★ Theorem:-

Let X be a continuous random variable with p.d.f $f_X(x)$ and support S_X. Let $Y = g(X)$, where $g(X)$ is a one-to-one differentiable function, on the support of X, S_X. Denote the inverse of g by $x = g^{-1}(y)$ and let $dx/dy = d[g^{-1}(y)]$. Then the p.d.f of Y is given by

$$F_Y(y) = f_X(g^{-1}(y))\left|\frac{dx}{dy}\right|, \text{ for } y \in S_Y$$

Example:- Let X have p.d.f

$$f(x) = \begin{cases} 1 & 0 < x < 1 \\ 0 & \text{otherwise} \end{cases}$$

find the p.d.f of $Y = -2\log X$

Solution

Since $y = g(x) = -2\log x$

$$\Rightarrow x = e^{-y/2}$$

$$\Rightarrow \frac{dx}{dy} = e^{-y/2}\left(-\frac{1}{2}\right)$$

$|J| = \frac{dx}{dy}$
↳ Jacobian

$$\Rightarrow f_Y(y) = f_X(e^{-y/2})\,|J|$$

$$= 1 \cdot \frac{1}{2} e^{-y/2} \,; \quad 0 < y < \infty$$

★ Mode :-

A mode of a distribution of one random variable X is the value of x that maximizes the p.d.f or p.m.f. For X of continuous type, f(x) must be continuous. If there is only one such x, it is called the mode of the distribution.

★ Median :-

A median 'm' of a continuous distribution is obtained by $F(m) = 1/2$

Question:- Find the mode of the following distribution

$$f_X(x) = \frac{1}{2} x^2 e^{-x} \,; \quad 0 < x < \infty$$

Solution $f(x) = \frac{1}{2} x^2 e^{-x}$

$$f'(x) = \frac{1}{2}[2x e^{-x} + x^2(-1)e^{-x}]$$

$$= \frac{e^{-x}}{2}[2x - x^2]$$

$$f''(x) = \frac{(-1)e^{-x}}{2}(2x - x^2) + \frac{e^{-x}}{2}[2 - 2x]$$

$$\leq 0$$

$F(Q_3) = \frac{3}{4}$
$F(Q_1) = 1/4$
$F(P_{20}) = 0.20$
$F(P_{90}) = 0.90$
$F(\xi_p) = P$

Put $f'(x) = 0$

$$\Rightarrow \frac{x e^{-x}}{2}[2 - x] = 0$$

$$\Rightarrow x = 0 \quad \text{or} \quad x = 2$$

Therefore Mode $= x = 2$

Question:- Find the median of the following distribution function
$$f(x) = 3x^2, \quad 0 < x < 1$$

Solution

According to the definition of median $F(m) = 1/2$

$$F(x) = P(X \leq x) = \int_{-\infty}^{x} f(x)\,dx$$

$$\Rightarrow F(m) = 3\int_{-0}^{m} x^2\,dx$$

$$= \cancel{3}\,\frac{x^3}{\cancel{3}}\Big|_{-0}^{m} = [m^3 - 0]$$

$$= m^3 = \frac{1}{2}$$

$$\Rightarrow m = \left(\frac{1}{2}\right)^{1/3}$$

* Expectation of a Random Variable OR Mean:-

Let X be a random variable.

If X is a continuous random variable with p.d.f $f(x)$ and $\int_{-\infty}^{\infty} |x| f(x) < \infty$ then expectation of X is

$$E(X) = \int_{-\infty}^{\infty} x\, f(x)\, dx$$

If X is a discrete random variable with p.m.f $P(x)$ and $\sum |x| P(x) < \infty$ then

$$E(X) = \sum_{x} x\, P(x)$$

If $Y = g(x)$ then $E(Y) = \sum g(x) P(x)$

$$= \int_{-\infty}^{\infty} g(x) f_X(x)\, dx$$

⇒ Properties of Expectation:-

i) $E(aX+b) = a\,E(X) + b$

ii) $E(X \pm Y) = E(X) \pm E(Y)$

iii) $E(XY) = E(X)\,E(Y)$ provided X and Y are independent r.Vs.

* Variance of X:-

$$Var(X) = V(X) = E(X^2) - (E(X))^2$$

* Moment Generating Function:- (M.G.F)

$$M_X(t) = E(e^{tX}) \;;\; -h < t < h \,,\; h > 0$$

⇒ Ordinary or Raw Moments:-

$$\mu_r' = E(X^r) \;;\; r = 1, 2, 3, 4$$

⇒ Mean Moments:-

$$\mu_r = E(X-\mu)^r \; ; \; r = 1,2,3,4$$

where $\mu = E(X)$

Question:- Show that the M.G.F of the random variable X having the p.d.f $f(x) = 1/3 \; ; \; -1 < x < 2$ is

$$M_X(t) = \begin{cases} \dfrac{e^{2t} - e^{-t}}{3t} & ; \; t \neq 0 \\ 1 & ; \; t = 0 \end{cases}$$

Solution

$$M_X(t) = E(e^{tX})$$

$$= \frac{1}{3}\int_{-1}^{2} e^{tx}\, dx$$

$$= \frac{1}{3} \left. \frac{e^{tx}}{t} \right|_{-1}^{2} = \frac{1}{3t}\left[e^{2t} - e^{-t}\right]$$

$$\Rightarrow M_X(t) = \frac{e^{2t} - e^{-t}}{3t}$$

★ $\left.\frac{d}{dt} M_X(t)\right|_{t=0} = \mu_1'$, $\left.\frac{d^2}{dt^2} M_X(t)\right|_{t=0} = \mu_2'$

& so on

Question:- Let X_1 and X_2 have the joint p.d.f $f_{X_1,X_2}(x_1, x_2) = 10\, x_1 x_2^2 \; ; \; 0 < x_1, x_2 < 1$
$= 0$ else

Let $Y_1 = X_1/X_2$, $Y_2 = X_2$. Find the p.d.f of Y_1 and Y_2.

Solution The transformation is

$$y_1 = \frac{x_1}{x_2} \Rightarrow x_1 = y_1 y_2$$

$$y_2 = x_2 \Rightarrow x_2 = y_2$$

$$J = \begin{vmatrix} \frac{\partial x_1}{\partial y_1} & \frac{\partial x_1}{\partial y_2} \\ \frac{\partial x_2}{\partial y_1} & \frac{\partial x_2}{\partial y_2} \end{vmatrix} = \begin{vmatrix} y_2 & y_1 \\ 0 & 1 \end{vmatrix} = y_2$$

Then $0 < y_1 y_2 < 1 \Rightarrow 0 < y_2 < 1$

$0 < y_1 < 1 \Rightarrow 0 < y_1 < 1$

The joint p.d.f of Y_1 & Y_2

$$f_{Y_1,Y_2} = 10(y_1 y_2)(y_2)^2 |J|$$

$$= 10 y_1 y_2^4 \qquad 0 < y_1, y_2 < 1$$

By definition $f_{Y_1}(y_1) = \int_{y_2} f(y_1, y_2)\, dy_2$

$$\Rightarrow f_{Y_1}(y_1) = 10 \int_0^1 y_1 y_2^4\, dy_2$$

$$= 10 y_1 \left. \frac{y_2^5}{5} \right|_0^1 = 10 y_1 \left(\frac{1}{5} - 0\right)$$

$$= 2 y_1 \quad ; \quad 0 < y_1 < 1$$

& $f_{Y_2}(y_2) = 10 \int_{y_1} y_1 y_2^4\, dy_1$

$$= 10 y_2^4 \int_0^1 y_1\, dy_1 = 10 y_2^4 \left. \frac{y_1^2}{2} \right|_0^1$$

$$= 10 y_2^4 \left(\frac{1}{2} - 0\right)$$

$$= 5 y_2^4 \quad ; \quad 0 < y_2 < 1$$

* Some Discrete Probability Dist :-

1) Bernoulli Distribution :- A r.v is said to have Bernoulli Distribution if its p.m.f is

$$P_X(x) = P(1-P)^{1-x} \quad ; \quad x = 0, 1 \quad 0 < P < 1$$

2) Binomial Distribution :- A r.v. X is said to have Binomial Distribution if its p.m.f is

$$P_X(x) = \binom{n}{x} p^x q^{n-x} \quad ; \quad x = 0, 1, 2, \ldots, n \quad 0 < p < 1, \; q = 1 - p$$

3) Poisson Distribution :-

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

4) Negative Binomial Distribution :-

$$P_X(x) = \binom{k+x-1}{x} p^k q^x \quad ; \quad x = 0, 1, 2, \ldots, \infty \quad 0 < p < 1, \; q = 1 - p$$

5) Geometric Distribution :-

$$P_X(x) = p q^x \quad ; \quad x = 0, 1, 2, \ldots, \infty$$

* Some Continuous Probability Dist :-

1) Uniform Distribution :-

$$X \sim U(0, 1)$$

$\boxed{\sim \; = \text{is distributed as}}$

$$f(x) = \begin{cases} 1 & ; \; 0 < x < 1 \\ 0 & ; \; \text{otherwise} \end{cases}$$

$$X \sim U(\alpha,\beta)$$

$$f(x) = \frac{1}{\beta - \alpha} \quad ; \quad \alpha < x < \beta$$

2) Exponential Distribution:-

$$X \sim Exp(\lambda)$$

$$f(x) = \lambda e^{-x\lambda} \quad ; \quad 0 < x < \infty ; \quad \lambda > 0$$

3) Gemma Distribution:-

$$X \sim G(\alpha,\beta)$$

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

where $\Gamma\alpha = \int_0^{\infty} x^{\alpha-1} e^{-x} dx$

4) Normal Distribution:-

$$X \sim N(\mu, \sigma^2)$$

$$f(x) = \frac{1}{\sqrt{2\pi}\,\sigma} e^{-\frac{1}{2}\left(\frac{x-\mu}{\sigma}\right)^2} \; ; \quad \begin{array}{c} -\infty < x < \infty \\ -\infty < \mu < \infty \\ \sigma^2 > 0 \end{array}$$

5) Chi-Square Distribution:-

$$f(x) = \frac{1}{\Gamma\frac{n}{2} \, 2^{n/2}} x^{\frac{n}{2}-1} e^{-\frac{x}{2}} \quad ; \quad 0 < x < \infty$$

6) Beta Distribution:-

$$f(x) = \frac{\Gamma\alpha+\beta}{\Gamma\alpha \, \Gamma\beta} x^{\alpha-1} (1-x)^{\beta-1} \; ; \quad \begin{array}{c} \alpha,\beta > 0 \\ 0 < x < 1 \end{array}$$

7) Cauchy Distribution:-

$$f(x) = \frac{\sigma}{\pi} \cdot \frac{1}{\sigma^2 + (x-\mu)} \; ; \quad \begin{array}{c} x \in R \\ \mu \in R \end{array} , \quad \sigma > 0$$

* Random Sampling:-

The random variables $X_1, X_2, \cdots, X_n$ are called a random sample of size n from the population $f(x)$ if $X_1, \cdots, X_n$ are mutually indipendent random variables and marginal pdf or pmf of each X_i is the same function $f(x)$. Alternatively, $X_1, \cdots, X_n$ are called independent and identically distributed (iid) random variables with pdf or pmf $f(x)$.

Recall that we will use boldface ~~the~~ letters to denote multiple variates, so $\mathbf{X}$ denotes the random variables $X_1, X_2, \cdots, X_n$ & $\mathbf{x}$ denotes the sample $x_1, x_2, \cdots, x_n$.

* Point Estimation:-

⇒ **Introduction:-** Suppose we have random variable X which has pdf (or pmf), $f(x;\theta)$, $(P(x;\theta))$, where θ is either a real number or a vector of real number. Assume that $\theta \in \Omega$ which is a subset of $\mathbb{R}^r$, for $r \geq 1$ (Ω stands for set of all possible values of θ and call it the parameter space). For example, θ could be the vector (μ, σ^2) when X has a $N(\mu, \sigma^2)$ distribution or θ could be the probability of success p when X has a binomial distribution. Because θ is unknown, we want to estimate it our information about θ comes from a sample $X_1, X_2, \cdots, X_n$. We often assume that this is a random sample which means that the random variables $X_1, X_2, \cdots, X_n$ are independent and have the same distribution as X; i.e $X_1, X_2, \cdots, X_n$ are iid.

A statistic T is a function of the sample, i.e. $T = T(X_1, X_2, \cdots, X_n)$. We may use T to estimate θ. In which case, we would say that T is a point estimation of θ. For example, suppose $X_1, X_2, \cdots, X_n$ is a random sample from a distribution with mean μ and variance σ^2. Then $\overline{X}$ and S^2 are point estimation of μ and σ^2, respectively.

⇒ The Difference b/w an estimate & estimator:-

An estimator is a function of the sample, while an estimate is the realized value of an estimator that is obtained when a sample is actually taken. Notationally, when a sample is taken, an estimator is a function of the random variables $X_1, X_2, \cdots,$

X_n, while an estimate is a function of the realized values $x_1, x_2, \cdots, x_n$

⇒ Methods of Finding Estimators:-

(i) Method of Moments.
(ii) Maximum Likelihood Estimators.
(iii) Bayes Estimators.

1) Method of Moments:- Let $X_1, X_2, \cdots, X_n$ be a sample from a population with pdf or pmf $f(x|\theta_1, \cdots, \theta_k)$. Method of Moments Estimators are found by equating the first K sample moments to the corresponding K population moments and solving the resulting system of equations simultaneously i.e

$$\underbrace{\begin{aligned} m_1 &= \frac{1}{n}\sum_{i=1}^{n} X_i \\ m_2 &= \frac{1}{n}\sum_{i=1}^{n} X_i^2 \\ &\vdots \\ m_K &= \frac{1}{n}\sum_{i=1}^{n} X_i^K \end{aligned}}_{\text{For sample}} \quad \underbrace{\begin{aligned} \mu_1' &= E(X) \\ \mu_2' &= E(X^2) \\ &\vdots \\ \mu_K' &= E(X^K) \end{aligned}}_{\text{For Population}}$$

M.TAHIR WATTOO

Example 1:- (Normal Method of Moments)
Suppose $X_1, X_2, \ldots, X_n$ are iid $N(\mu, \sigma^2)$.
Find MM estimators of parameters.

Solution

We have $m_1 = \bar{X}$, $m_2 = \frac{1}{n}\sum_{i=1}^{n} X_i^2$

Also $\mu_1' = \mu$ & $\mu_2' = \mu^2 + \sigma^2$ (Population)

Therefore after equating sample & population moments we have

$$\bar{X} = \mu \quad , \quad \frac{1}{n}\sum_{i=1}^{n} X_i^2 = \mu^2 + \sigma^2$$

Solving for μ & σ^2 yields the MM estimators

$$\hat{\mu} = \bar{X} \quad \& \quad \hat{\sigma}^2 = \frac{1}{n}\sum_{i=1}^{n} X_i^2 - \bar{X}^2$$

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

Since Variance $= \sigma^2 = E(X^2) - [E(X)]^2$

$$\Rightarrow \sigma^2 = E(X^2) - \mu^2$$

$$\Rightarrow \sigma^2 + \mu^2 = E(X^2)$$

$$\Rightarrow \sigma^2 + \mu^2 = \mu_2'$$

Example 2:- (Binomial Method of Moments)
Let $X_1, X_2, \ldots, X_n$ be iid binomial (K, P); Find MM estimators of K & P

Solution

$$m_1 = \bar{X} \quad , \quad m_2 = \frac{1}{n}\sum_{i=1}^{n} X_i^2$$

$$\mu_1' = KP \quad , \quad \mu_2' = KP(1-P) + K^2P^2$$

After equating we get

$$\bar{X} = kp \quad , \quad \frac{1}{n}\sum X_i^2 = kp(1-p)+k^2p^2$$

$$\hat{k} = \frac{\bar{X}^2}{\bar{X} - \frac{1}{n}\sum_{i=1}^{n}(X_i-\bar{X})^2} \quad , \quad \hat{p} = \frac{\bar{X}}{\hat{k}}$$

$$M_X(t) = E(e^{tX})$$

$$= \sum_{x=0}^{n} e^{tx}\binom{n}{x}p^x q^{n-x}$$

$$= \sum_{x=0}^{n}\binom{n}{x}(pe^t)^x q^{n-x}$$

$$= (q+pe^t)^n$$

$$\Rightarrow \frac{d}{dt}M_X(t) = n(q+pe^t)^{n-1}(pe^t)$$

$$\Rightarrow \frac{d^2}{dt^2}M_X(t) = npe^t(q+pe^t)^n + n(n-1)p^2e^{2t}(q+pe^t)^{n-2}$$

$$\Rightarrow \frac{d^2}{dt^2}M_X(t)\Big|_{t=0} = np(q+p)+n(n-1)p^2(q+p)^{n-2}$$

$$= np + n(n-1)p^2 \quad \because q+p=1$$

$$= np + n^2p^2 - np^2$$

$$= np(1-p) + n^2p^2$$

Replace n by k implies

$$\frac{d^2}{dt^2}M_X(t)\Big|_{t=0} = kp(1-p)+k^2p^2$$

Proved

Example No 3:- Let $X_1, X_2, \ldots, X_n$ be iid $U(0,\theta)$. Find an estimator of θ by MM.

Solution

As $m_1 = \bar{X}$ & $\mu_1' = E(X) = \int_0^\theta x \frac{1}{\theta}\, dx$

$$\Rightarrow \mu_1' = \theta/2$$

Matching the moments we get

$$\bar{X} = \frac{\theta}{2} \quad \text{or} \quad \hat{\theta} = 2\bar{X}$$

Example No. 4:- Let $X_1, \ldots, X_n$ be a random sample from $f(x|a,b) = \frac{1}{b-a}$; $a \le x \le b$

$= 0$ otherwise

Find the method of moments estimators of a and b.

Solution

$$E(X) = \int_a^b x\left(\frac{1}{b-a}\right)dx = \frac{1}{b-a}\left(\frac{b^2-a^2}{2}\right) = \frac{a+b}{2}$$

$$E(X^2) = \int_a^b x^2\left(\frac{1}{b-a}\right)dx = \frac{1}{b-a}\left(\frac{b^3-a^3}{2}\right)$$

$$= \frac{a^2+ab+b^2}{3}$$

Setting $\bar{X} = \frac{a+b}{2}$, $\frac{1}{n}\sum X^2 = \frac{a^2+ab+b^2}{3}$

$\Rightarrow \hat{a} = 2\bar{X} - \hat{b}$, $\hat{b} = \bar{X} + \sqrt{3S^2}$

$\Rightarrow \hat{a} = \bar{X} - \sqrt{3S^2}$

Example No 5:- Let $X_1, X_2, \ldots, X_n$ be a random sample from

$$f(x;\alpha,\beta) = \frac{\Gamma\alpha\,\Gamma\beta}{\Gamma\alpha\,\Gamma\beta}\, x^{\alpha-1}(1-x)^{\beta-1} \text{ if } 0 \le x \le 1, \quad \alpha>0, \beta>0$$

Find the method of Moment estimator of α and β.

Solution

Recall that

$$E(X) = \frac{\alpha}{\alpha+\beta}, \quad E(X^2) = \frac{\alpha(\alpha+1)}{(\alpha+\beta)(\alpha+\beta+1)}$$

$$S^2 = Var(X) = E(X^2) - [E(X)]^2$$

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

Now $\bar{X} = \frac{\hat{\alpha}}{\hat{\alpha}+\hat{\beta}} \Rightarrow \frac{\hat{\beta}}{\hat{\alpha}+\hat{\beta}} = 1-\bar{X}$

& $S^2 = \frac{\hat{\alpha}}{(\hat{\alpha}+\hat{\beta})} \cdot \frac{\hat{\beta}}{(\hat{\alpha}+\hat{\beta})} \cdot \frac{1}{(\hat{\alpha}+\hat{\beta}+1)}$

$$\Rightarrow S^2 = \bar{X}(1-\bar{X}) \frac{1}{(1+\hat{\alpha}/\bar{X})}$$

$$\Rightarrow 1 + \frac{\hat{\alpha}}{\bar{X}} = \frac{\bar{X}(1-\bar{X})}{S^2}$$

$$\Rightarrow \hat{\alpha} = \bar{X}\left[\frac{\bar{X}(1-\bar{X})}{S^2} - 1\right]$$

As $\frac{\hat{\alpha}}{\hat{\alpha}+\hat{\beta}} = \bar{X} \Rightarrow \hat{\beta} = \frac{\hat{\alpha}}{\bar{X}} - \hat{\alpha} = \hat{\alpha}\left[\frac{1}{\bar{X}} - 1\right]$

$$\Rightarrow \hat{\beta} = (1-\bar{X})\left[\frac{\bar{X}(1-\bar{X})}{S^2} - 1\right]$$

⇒ Maximum Likelihood Estimators:-

Consider a random sample $X_1, X_2, \dots, X_n$ from a distribution having pdf or pmf $f(x;\theta)$, $\theta \in \Omega$. The joint density of $X_1, X_2, \dots, X_n$ is $f(x_1;\theta) f(x_2;\theta) \cdots f(x_n;\theta)$. This joint pdf may be regarded as a function of θ, it is called the likelihood L

$$L(\theta; x_1, x_2, \dots, x_n) = f(x_1;\theta) \cdot f(x_2;\theta) \cdots f(x_n;\theta)$$

$$= \prod_{i=1}^{n} f(x_i;\theta) \; ; \; \theta \in \Omega$$

Def. (MLE) :- The method of Maximum likelihood consists of maximizing the likelihood function with respect to the parameter θ. The resulting value at which the maximum occur is called maximum likelihood estimator, usually denoted by $\hat{\theta}$.

that is $L(\hat{\theta}; x_1, x_2, \dots, x_n) = \sup\limits_{\theta} L(\theta, x_1, x_2, \dots, x_n)$

If $\hat{\theta}$ satisfies the above condition, then $\hat{\theta}$ maximum likelihood estimator.

Actually the Log of this function L is usually more convenient to work with mathematically, since the logarithm is a monotone function

that is $\ln L(\hat{\theta}; x_1, x_2, \dots, x_n) = \sup\limits_{\theta} \ln L(\theta; x_1, \dots, x_n)$

Suppose that $f(x)$ is a +ve and differentiable function of θ, then MLE can be found by taking the derivative of the likelihood function with respect to θ,

Example 1:- Let $X_1, X_2, \ldots, X_n$ denote a random sample from the distribution with pmf $p(x) = \theta^x (1-\theta)^{1-x}$; $x = 0, 1$; $0 \le \theta \le 1$

Find MLE of θ

Solution: The joint pmf of $(X_1 = x_1, X_2 = x_2, \ldots, X_n = x_n)$ is given by

$$L(\theta; x_1, x_2, \ldots, x_n) = \theta^{\sum x_i} (1-\theta)^{n - \sum x_i},$$

where $x_i = 0, 1$, $i = 1, 2, \ldots, n$

We might ask what value of θ would maximize the $L(\theta)$ of obtaining this parameter observed sample $x_1, x_2, \ldots, x_n$

Since the likelihood function $L(\theta)$ and its Logarithm $l(\theta) = \log L(\theta)$ are maximized for the same value of θ, either $L(\theta)$ or $l(\theta)$ can be used. Here

$$l(\theta) = \log L(\theta) = \left(\sum_{i=1}^{n} x_i\right) \log \theta + \left(n - \sum_{i=1}^{n} x_i\right) \log(1-\theta)$$

So we have $\dfrac{d\, l(\theta)}{d\theta} = \dfrac{\sum x_i}{\theta} - \dfrac{x - \sum x_i}{1-\theta} = 0$

Provided that $\theta \neq 0$ or 1

$$(1-\theta) \sum_{i=1}^{n} x_i = \theta \left(n - \sum_{i=1}^{n} x_i\right)$$

whose solution for θ is $\sum_{i=1}^{n} x_i / n$.

That $\sum^{n} x_i / n$ actually maximizes $L(\theta)$ and $\log L(\theta)$.

The corresponding statistic

$\hat{\theta} = \frac{1}{n} \sum^{n} X_i = \bar{X}$ is called the maximum likelihood estimator of θ and $\sum_{i=1}^{n} x_i/n$ the maximum likelihood estimate of θ

$$L(x_1, x_2, \cdots, x_n; \theta) = f(x_1, x_2, \cdots, x_n)$$
$$= f(x_1) \times f(x_2) \times \cdots \times f(x_n)$$
$$= \theta^{x_1}(1-\theta)^{1-x_1} \times \theta^{x_2}(1-\theta)^{1-x_2} \times \cdots \times \theta^{x_n}(1-\theta)^{1-x_n}$$
$$= \theta^{x_1+x_2+\cdots+x_n}(1-\theta)^{n-(x_1+x_2+\cdots+x_n)}$$

$\ell = \log L$

Example 2:- Let $X_1, X_2, \cdots, X_n$ be a random sample from $N(\mu, \sigma^2)$. Find the MLEs of μ and σ^2

Solution The likelihood function is given by

$$L(\mu, \sigma^2 / x_1, x_2, \cdots, x_n) = \prod_{i=1}^{n} f(x_i | \mu, \sigma^2)$$
$$= \prod_{i=1}^{n} (2\pi\sigma^2)^{-\frac{1}{2}} \exp\left[\frac{-1}{2\sigma^2}(x_i - \mu)^2\right]$$
$$= (2\pi\sigma^2)^{-\frac{n}{2}} \exp\left[\frac{-1}{2\sigma^2}\sum_{i=1}^{n}(x_i - \mu)^2\right]$$

Let $x = (x_1, x_2, \cdots, x_n)$. The log likelihood function is

$$\ell(\mu, \sigma^2/x) = \ln\left[L(\mu, \sigma^2/x)\right]$$
$$= -\frac{n}{2}\ln(2\pi\sigma^2) - \frac{1}{2\sigma^2}\sum_{i=1}^{n}(x_i - \mu)^2$$

To maximize the likelihood function we need to differentiate log likelihood $l(\mu, \sigma^2/x)$ with respect to μ and σ^2.

$$\frac{\partial l}{\partial \mu} = \frac{1}{\sigma^2}\sum_{i=1}^{n}(x_i - \mu), \; \& \; \frac{\partial l}{\partial \sigma^2} = \frac{-n}{2\sigma^2} + \frac{1}{4\sigma^2}\sum_{i=1}^{n}(x_i - \mu)^2$$

Setting these partial derivatives equal to zero and solving yields

$$\hat{\mu} = \bar{X} \quad \& \quad \hat{\sigma}^2 = n^{-1}\sum_{i=1}^{n}(x_i - \bar{x})^2$$

Example 3:- Suppose $X_1, X_2, \ldots, X_n$ is a random sample from uniform distribution $[0, \theta]$. Find the MLE for θ.

Solution The pdf of uniform distribution is

$$f(x/\theta) = \frac{1}{\theta} \; ; \text{ if } 0 \le x \le \theta$$
$$= 0 \quad \text{elsewhere}$$

because θ is in the support, differentiation is not helpful here.

The likelihood function

$$L(\theta/x) = \prod_{i=1}^{n} f(x_i/\theta) = \prod_{i=1}^{n} 1/\theta = \theta^{-n} I(\max\{x_i\}, \theta) \quad \forall \; \theta > 0$$

Where $I(a,b)$ is 1 or 0 if $a \le b$ or $a > b$ respectively.

This function is decreasing function of θ for all $\theta \ge \max\{x_i\}$

and 0 otherwise. Hence the maximum occurs at the smallest value of θ i.e the MLE is $\hat{\theta} = \max\{X_i\}$

Example 4:- Let $X_1, X_2, \ldots, X_n$ be iid random variables from $U(\alpha, \beta)$.
Here $\theta = (\alpha, \beta) \in \Omega$

Solution

$$L(\theta | x_1, x_2, \cdots, x_n) = \frac{1}{(\beta-\alpha)^n}$$

Here the likelihood function is not differentiable with respect to α, β. But it is clearly maximized when $(\beta - \alpha)$ is minimized subject to the conditions that $\alpha \leq \min\{x_i\}$ and $\beta \geq \max\{x_i\}$.

This happens when $\hat{\alpha} = \text{Min}\{X_i\}$ & $\hat{\beta} = \max\{X_i\}$. Thus $\hat{\alpha}$ & $\hat{\beta}$ are MLEs of α and β respectively.

* Invariance Property of MLE:-

In practice, there are many situations when the experiment is merely interested in the estimation of a function of the model parameter, $g(\theta)$ instead of the model parameter itself. For example, the model parameter for a Poisson distribution is λ, but the experiment is interested in estimating $P(X=0) = e^{-\lambda}$.

Thus, the parameter of interest is a function of λ that is $g(\lambda) = e^{-\lambda}$.
The invariance principle of MLE allows us to replace the parameter in the function with its MLE.
If $\hat{\lambda}$ is the MLE of λ, Then $g(\hat{\lambda}) = e^{-\hat{\lambda}}$ is the MLE for $g(\lambda)$

ASSIGNMENT # 1

Q1:- Let $X_1, X_2, \dots, X_n$ be a random sample from $G(\alpha, \beta)$. Find MM estimator of (α, β)

Q2:- A random sample of size n is taken from the log normal pdf

$$f(x; \mu, \sigma) = \frac{1}{x\sigma\sqrt{2\pi}} \exp\left[-\frac{1}{2\sigma^2}(\log x - \mu)^2\right]; x>0$$

Find MM estimator of μ & σ^2

Q3:- Let $X_1, X_2, \dots, X_n$ be a random sample of size n from a gamma distribution with parameter $\alpha = 2, \beta$. Compute the MLE for β.

Q4:- Let $X_1, X_2, \dots, X_n$ be a random sample of size n from a Weibull distribution of the form

$$f(x) = \frac{1}{\beta} x^{\alpha-1}, e^{-x^{\alpha}/\beta} \quad ; x>0$$

Suppose α is known. Compute MLE for β.

Q 5:- $f(x;\theta) = \dfrac{\theta}{(1+x)^{\theta+1}}$; $x > 0$, $\theta > 0$

Find MLE for θ.

Q 6:- Let $X_1, X_2, \ldots, X_n$ be iid random variables from Bernoulli population with probability of success p. Find the MLE of $V(X) = P(1-P)$

Solution As MLE of P is $\hat{P} = \dfrac{\sum x_i}{n} = \bar{x}$

The $V(X) = g(P) = P(1-P)$

We put MLEs into the function using the invariance property of MLE; that is

$$g(\hat{P}) = \hat{P}(1-\hat{P}) = \bar{x}(1-\bar{x})$$

Solutions :-

Ans 1 As $f(x) = \dfrac{1}{\Gamma\alpha \, \beta^{\alpha}} x^{\alpha-1} e^{-x/\beta}$, $0 < x < \infty$

where $\Gamma\alpha = \int_0^\infty x^{\alpha-1} e^{-x} dx$

Here $m_1 = \mu_1'$ & $m_2 = \mu_2'$

$\dfrac{\sum x_i}{n} = E(X) \longrightarrow ①$ & $\dfrac{\sum x_i^2}{n} = E(X^2) \longrightarrow ②$

Now

$$E(X) = \int_0^\infty x f(x) \, dx$$

$$\Rightarrow E(X) = \int_0^\infty x \cdot \frac{1}{\Gamma\alpha\, \beta^\alpha} x^{\alpha-1} e^{-x/\beta} \, dx$$

$$= \frac{1}{\Gamma\alpha\, \beta^\alpha} \int_0^\infty x^\alpha e^{-x/\beta} \, dx$$

Put $x/\beta = y \Rightarrow dx = \beta\, dy$

Then $$E(X) = \frac{1}{\Gamma\alpha\, \beta^\alpha} \int_0^\infty (y\beta)^\alpha e^{-y} \beta \, dy$$

$$= \frac{\beta \beta^\alpha}{\Gamma\alpha\, \beta^\alpha} \int_0^\infty y^{(n+1)-1} e^{-y} \, dy$$

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

$$\Rightarrow E(X) = \alpha\beta$$

Now
$$E(X^2) = \int_0^\infty x^2 \cdot \frac{1}{\Gamma\alpha\, \beta^\alpha} e^{-x/\beta} \cdot x^{\alpha-1} \, dx$$

$$= \frac{1}{\Gamma\alpha\, \beta^\alpha} \int_0^\infty x^{\alpha+1} e^{-x/\beta} \, dx$$

Put $y = x/\beta \Rightarrow \beta\, dy = dx$

So $$E(X^2) = \frac{1}{\Gamma\alpha\, \beta^\alpha} \int_0^\infty (\beta y)^{\alpha+1} e^{-y} \beta \, dy$$

$$= \frac{\beta^{\alpha+2}}{\Gamma\alpha\, \beta^\alpha} \int_0^\infty e^{-y} y^{\alpha+1} \, dy$$

$$\Rightarrow E(X^2) = \frac{\beta^2}{\Gamma\alpha} \Gamma\alpha+2$$

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

$$\Rightarrow E(X^2) = \alpha(\alpha+1)\beta^2$$

$$① \Rightarrow \bar{X} = \alpha\beta \longrightarrow ②$$

$$② \Rightarrow \frac{\sum x_i^2}{n} = \alpha\beta^2(\alpha+1)$$

$$= \alpha\beta \cdot \beta(\alpha+1) = \bar{X}(\alpha\beta+\beta)$$

$$= \bar{X}(\bar{X}+\beta) = \bar{X}^2 + \bar{X}\beta$$

$$\Rightarrow \frac{\sum x_i^2}{n} - \bar{X}^2 = \bar{X}\beta$$

$$\boxed{\because \frac{1}{n}\sum x_i^2 - \bar{X}^2 = \frac{1}{n}\sum(x-\bar{X})^2 = S^2}$$

$$S^2 = \bar{X}\beta$$

$$\Rightarrow \boxed{\hat{\beta} = \frac{S^2}{\bar{X}}}$$

$$③ \Rightarrow \hat{\alpha} = \frac{\bar{X}^2}{S^2} \quad \because \bar{X} = \alpha\frac{S^2}{\bar{X}}$$

Ans 3

As $f(x) = \frac{1}{\Gamma\alpha\, \beta^\alpha} e^{-x/\beta} x^{\alpha-1}$, $0<x<\infty$, $\alpha>0, \beta>0$

Put $\alpha = 2$ implies

$$f(x) = \frac{1}{\Gamma 2\, \beta^2} x \cdot e^{-x/\beta} \quad \because \Gamma 2 = 1$$

The joint density function of given distribution is

$$L(2, \beta, x_1, x_2, \cdots, x_n) = \prod_{i=1}^{n} f(x_i, 2, \beta)$$

$$= \frac{1}{\beta^{2n}} \cdot \prod_{i=1}^{n} x_i \cdot e^{-1/\beta \sum_{i=1}^{n} x_i}$$

Taking ln on both sides

$$l = \ln L = \ln \frac{1}{\beta^{2n}} + \ln \sum x_i - \frac{1}{\beta} \sum x_i \quad (\because \ln e = 1)$$

differentiate w.r.t β

$$\frac{dl}{d\beta} = \frac{d}{d\beta}\left(-\ln \beta^{2n}\right) + 0 + \frac{1}{\beta^2} \sum x_i$$

$$= \frac{-2n}{\beta} + \frac{1}{\beta^2} \sum x_i$$

Put $\frac{dl}{d\beta} > 0$

$$\Rightarrow \frac{1}{\beta^2} \sum x_i = \frac{2n}{\beta}$$

$$\Rightarrow \frac{1}{\beta}\left[\frac{1}{\beta} \sum x_i - 2n\right] = 0 \quad \because \beta \neq 0$$

$$\Rightarrow \frac{1}{\beta} \sum x_i = 2n \Rightarrow \beta = \frac{\sum x_i}{2n}$$

$$\Rightarrow \boxed{\hat{\beta} = \frac{1}{2} \overline{X}}$$

* ——————— *

Ans 4 $L(\alpha, \beta, x_1, x_2, \cdots, x_n) = \prod_{i=1}^{n} f(\alpha, \beta, x_i)$

$$= \frac{1}{\beta} x_1^{\alpha-1} e^{-\frac{1}{\beta} x_1} \cdot \frac{1}{\beta} x_2^{\alpha-1} e^{-\frac{1}{\beta} x_2} \cdot$$

$$\cdots \frac{1}{\beta} x_n^{\alpha-1} e^{-\frac{1}{\beta} x_n}$$

$$= \frac{1}{\beta^n} \prod_{i=1}^{n} x_i^{\alpha-1} e^{-\frac{1}{\beta} \sum_{i=1}^{n} x_i^{\alpha}}$$

Taking ln on both sides

$$l = \ln L = \ln 1 - \ln \beta^n + (\alpha-1) \ln \sum x_i$$

$$- \frac{1}{\beta} \sum x_i^{\alpha} \ln e$$

differentiate w.r.t β

$$\frac{dl}{d\beta} = 0 - \frac{n}{\beta} + 0 + \frac{1}{\beta^2} \sum x_i^{\alpha}$$

Put $\frac{dl}{d\beta} = 0$

$$\Rightarrow \frac{-n}{\beta} + \frac{1}{\beta^2} \sum x_i^{\alpha} = 0$$

$$\Rightarrow \frac{1}{\beta}\left[-n + \frac{1}{\beta} \sum x_i^{\alpha}\right] = 0$$

$$\Rightarrow \frac{1}{\beta} \sum x_i^{\alpha} = n \qquad \because \frac{1}{\beta} \neq 0$$

$$\Rightarrow \boxed{\hat{\beta} = \frac{\sum x_i^{\alpha}}{n}}$$

Remaining at Page 30(back)

Example:- Let $X_1, X_2, \cdots, X_n$ is a random sample from a distribution specified by the probability density function

$$f(x) = \frac{1}{2\alpha} \exp\left(-\frac{|x|}{\alpha}\right); \quad -\infty < x < \infty, \quad \alpha > 0 \text{ (unknown)}$$

(i) Show that the MLE of α is $\hat{\alpha} = \frac{1}{n}\sum_{i=1}^{n}|X_i|$

(ii) Derive the expected value of $\hat{\alpha}$ in (i)

(iii) Derive the variance of $\hat{\alpha}$ in (i)

Solution

The joint density function of $(x_1, x_2, \cdots, x_n)$ is given by

$$L(\alpha) = \frac{1}{2\alpha}\exp\left(-\frac{|x_1|}{\alpha}\right) \times \frac{1}{2\alpha}\exp\left(-\frac{|x_2|}{\alpha}\right) \times \cdots \times \frac{1}{2\alpha}\exp\left(-\frac{|x_n|}{\alpha}\right)$$

$$\Rightarrow L(\alpha) = \frac{1}{(2\alpha)^n}\exp\left(-\frac{\sum_{i=1}^{n}|x_i|}{\alpha}\right)$$

$$\Rightarrow l = \log L(\alpha) = -n\log_e(2\alpha) - \frac{\sum_{i=1}^{n}|x_i|}{\alpha}$$

$$\Rightarrow \frac{\partial l}{\partial \alpha} = \frac{-n}{\cancel{2}\alpha}(\cancel{2}) + \frac{\sum_{i=1}^{n}|x_i|}{\alpha^2}$$

Put $\frac{\partial l}{\partial \alpha} = 0 \Rightarrow \frac{n}{\alpha} = \frac{\sum_{i=1}^{n}|x_i|}{\alpha^2}$

$$\Rightarrow \alpha = \frac{\sum_{i=1}^{n}|x_i|}{n}$$

Therefore $\hat{\alpha} = \frac{1}{n}\sum_{i=1}^{n}|X_i|$

(ii) $$E(\hat{\alpha}) = \int \frac{\sum_{i=1}^{n}|X_i|}{n} \times \left(\frac{1}{2\alpha} e^{-\frac{|X_i|}{\alpha}}\right) dx_i$$

$$= \frac{\sum_{i=1}^{n}}{n} \int_{-\infty}^{\infty} |x_i| \left(\frac{1}{2\alpha} e^{-\frac{|x_i|}{\alpha}}\right) dx_i$$

$$= \frac{2\sum_{i=1}^{n}}{2n\alpha} \int_{0}^{\infty} x_i\, e^{-\frac{x_i}{\alpha}}\, dx_i$$

Put $\frac{x_i}{\alpha} = y_i \Rightarrow dx_i = \alpha\, dy_i$

$\Rightarrow x_i = \alpha y_i$ $\quad$ $x_i \to 0 \Rightarrow y_i \to 0$, $x_i \to \infty \Rightarrow y_i \to \infty$

$$\Rightarrow E(\hat{\alpha}) = \frac{\sum_{i=1}^{n}}{n\alpha} \int_{0}^{\infty} \alpha y_i\, e^{-y_i} \alpha\, dy_i$$

$$= \frac{\alpha \sum_{i=1}^{n}}{n} \int_{0}^{\infty} y_i^{2-1} e^{-y_i}\, dy_i$$

$$= \frac{\alpha \sum_{i=1}^{n}}{n} \Gamma 2 \qquad \because \int_0^\infty x^{\alpha-1} e^{-x} dx = \Gamma\alpha$$

$$= \frac{\alpha\, \cancel{[n]}}{\cancel{n}}\, 1! = \alpha$$

$\boxed{\begin{array}{l} Var(aX) = a^2 Var(X) \\ Var(X+Y) = V(X) + V(Y) + 2CoVar(X,Y) \\ \text{If } X \text{ \& } Y \text{ are indepen then} \\ Cov(X,Y) = 0 \end{array}}$

(iii) $$Var(\hat{\alpha}) = Var\left(\frac{1}{n}\sum_{i=1}^{n}|X_i|\right)$$

$$= \frac{1}{n^2}\sum_{i=1}^{n} Var|x_i|$$

$\because$ all x_i are independent

Since $Var(X_i) = E(X_i^2) - [E(X_i)]^2$

$$\Rightarrow Var(X_i) = \int_{-\infty}^{\infty} x_i^2 \left(\frac{1}{2\alpha} e^{-\frac{|x_i|}{\alpha}}\right) dx - \left\{\int_{-\infty}^{\infty} x_i \left(\frac{1}{2\alpha} e^{-\frac{|x_i|}{\alpha}}\right) dx\right\}^2$$

$$\Rightarrow Var(X_i) = \frac{\cancel{2}}{\cancel{2}\alpha}\int_0^\infty x_i^2 e^{\frac{-x_i}{\alpha}} dx_i - \left(\frac{2}{2\alpha}\int_0^\infty x_i e^{\frac{-x_i}{\alpha}} dx_i\right)^2$$

$$= \frac{1}{\alpha}\int_0^\infty x_i^2 e^{\frac{-x_i}{\alpha}} dx_i - \left(\frac{1}{\alpha}\int_0^\infty x_i e^{\frac{-x_i}{\alpha}} dx_i\right)^2$$

Put $\frac{x_i}{\alpha} = y_i \Rightarrow dx_i = \alpha dy_i$

& $x_i = \alpha y_i$

$$\Rightarrow Var(X_i) = \frac{1}{\alpha}\int_0^\infty \alpha^2 y_i^2 e^{-y_i} \alpha dy_i - \left[\frac{1}{\alpha}\int_0^\infty \alpha y_i e^{-y_i} \alpha dy_i\right]^2$$

$$= \frac{\alpha^3}{\alpha}\int_0^\infty y_i^{3-1} e^{-y_i} dy_i - \alpha^2$$

$$= \alpha^2 \Gamma 3 - \alpha^2 = \alpha^2(2!) - \alpha^2$$

$$= 2\alpha^2 - \alpha^2$$

$$\Rightarrow Var(X_i) = \alpha^2$$

Example:- Suppose $X_1, X_2, \ldots, X_n$ is a random sample from $N(\mu, \sigma^2)$. Suppose $Y_1, Y_2, \ldots, Y_m$ is a random sample from $LN(\mu, \sigma^2)$. Assume both μ and σ^2 are unknown

i) Show that MLE of μ is

$$\frac{1}{m+n}\left[\sum_{i=1}^{n} X_i + \sum_{i=1}^{m} \log Y_i\right]$$

(ii) Show that MLE of σ^2 is

$$\frac{1}{m+n}\left[\sum_{i=1}^{n}(x_i-\hat{u})^2+\sum_{i=1}^{n}(\log y_i-\hat{u})^2\right]$$

Solution

The joint density function is

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

$$\Rightarrow L(u,\sigma^2)=\frac{1}{\sigma^{n+m}(2\pi)^{\frac{n+m}{2}}\,y_i^{m}}\exp\left[\frac{-1}{2\sigma^2}\left\{\sum_{i=1}^{n}(x_i-u)^2+\sum_{i=1}^{m}(\log y_i-u)^2\right\}\right]$$

$$\Rightarrow \log L(u,\sigma^2)=-\left(\frac{m+n}{2}\right)\log_e 2\pi-\sum_{i=1}^{m}\log y_i-(m+n)\log\sigma-\frac{1}{2\sigma^2}\left\{\sum_{i=1}^{n}(x_i-u)^2+\sum_{i=1}^{m}(\log y_i-u)^2\right\}$$

Now

$$\frac{\partial l}{\partial u}=\frac{-1}{2\sigma^2}\left[-2\sum_{i=1}^{n}(x_i-u)-2\sum_{i=1}^{m}(\log y_i-u)\right]$$

$$=\frac{1}{\sigma^2}\left[\sum_{i=1}^{n}x_i-nu+\sum_{i=1}^{m}\log y_i-mu\right]$$

$$=\frac{1}{\sigma^2}\left[-u(m+n)+\sum_{i=1}^{n}x_i+\sum_{i=1}^{m}\log y_i\right]$$

Put $\dfrac{\partial l}{\partial u}=0$

$$\Rightarrow +u(m+n)=\sum_{i=1}^{n}x_i+\sum_{i=1}^{m}\log y_i$$

$$\Rightarrow \mu = \frac{1}{m+n}\left[\sum_{i=1}^{n} x_i + \sum_{i=1}^{m} \log y_i\right]$$

$$\Rightarrow \hat{\mu} = \frac{1}{m+n}\left[\sum_{i=1}^{n} X_i + \sum_{i=1}^{m} \log Y_i\right]$$

Now Put $\frac{\partial l}{\partial \sigma} = 0$

$$\Rightarrow 0 = \frac{1}{\sigma^2}\left[-\sigma(m+n) + \frac{1}{\sigma}\left\{\sum_{i=1}^{n}(x_i-\mu)^2 + \sum_{i=1}^{m}(\log x_i-\mu)^2\right\}\right]$$

$\times$ by σ^3

$$\Rightarrow \sigma^2(m+n) = \sum_{i=1}^{n}(x_i-\mu)^2 + \sum_{i=1}^{m}(\log y_i-\mu)^2$$

$$\Rightarrow \sigma^2 = \frac{1}{m+n}\left[\sum_{i=1}^{n}(x_i-\mu)^2 + \sum_{i=1}^{m}(\log y_i-\mu)^2\right]$$

$$\Rightarrow \hat{\sigma}^2 = \frac{1}{m+n}\left[\sum_{i=1}^{n}(X_i-\hat{\mu})^2 + \sum_{i=1}^{m}(\log Y_i-\hat{\mu})^2\right]$$

Conditional Probability Density fn

$$f(x/y) = \frac{f(x,y)}{h(y)} \quad ; \quad h(y) = \int_x f(x,y)\,dx$$

$$f(y/x) = \frac{f(x,y)}{g(x)} \quad ; \quad g(x) = \int_y f(x,y)\,dy$$

$X \sim f(x;\theta) = f(x/\theta)$

$\theta \sim \pi(\theta) \longrightarrow$ Prior Distribution

$\pi(x/\theta) = \frac{f(x/\theta)\pi(\theta)}{\int f(x/\theta)\pi(\theta)\,d\theta}$ Posterior Distn

3:- **Bayes Estimators**:- The Bayesian approach to statistics is different from the classical approach that we have been taking. In the classical approach the parameter, θ, is thought to be an unknown, but fixed quantity. A random sample $X_1, X_2, \ldots, X_n$ is drawn from population indexed by θ and based on the observed values in the sample, knowledge about the value of θ obtained.

In the Bayesian approach, θ is considered to be a quantity whose variation can be described by a probability distribution (called prior distribution). This is a subjective distributor based on the experimenter's belief and is formulated before the data are seen (hence the name prior distributed). A sample is then taken from a population indexed by θ and the prior distribution is updated with the sample information. The updated prior is called the Posterior Distribution.

If we denote the prior distribution by $\pi(\theta)$ and the sampling distribution by $f(x/\theta)$, Then the posterior distribution, the

Conditional distribution of θ given the sample, $X = x$ is

$$\pi(\theta/x) = f(x/\theta)\cdot\pi(\theta)/m(x)\,; \quad f(x/\theta)\,\pi(\theta) = f(x,\theta)$$

where $m(x) = \int f(x/\theta)\,\pi(\theta)\,d\theta$

Notice that $\pi(\theta/x)$ is conditional distribution of a random variable θ while $X = x$, is held constant. The posterior distribution is now used to make statements about θ. The mean of the posterior distribution can be used as a point estimate of θ

Example 1:- Let $X_1, X_2, \ldots, X_n$ denote a random sample from the Bernoulli (θ). Assume that the prior distribution of θ is $U(0,1)$. Find the Bayes estimator of θ and $\tau(\theta) = \theta(1-\theta)$.

Solution

Since $X_1, X_2, \ldots, X_n$ be iid Bernoulli (θ)

i.e $\quad f(x/\theta) = \theta^x (1-\theta)^{1-x}$ for $x = 0, 1$

Then $Y = \sum_{i=1}^{n} X_i$ is binomial (n, θ).

We assume the prior distribution on θ is

$$\pi(\theta) = I_{(0,1)}(\theta)$$

The joint distribution of Y & θ is

$$f(y,\theta) = \binom{n}{y}\theta^y (1-\theta)^{n-y} \cdot 1$$

The joint pmf of Y is

$$f(Y) = \int_0^1 \binom{n}{y} \theta^y (1-\theta)^{n-y} d\theta = \binom{n}{y} \int_0^1 \theta^{y+1-1} (1-\theta)^{n-y+1-1} d\theta$$

$$\Rightarrow f(Y) = \binom{n}{y} \beta(y+1, n-y+1)$$

$$\boxed{\begin{array}{l} \beta(\alpha,\beta) = \dfrac{\Gamma\alpha\,\Gamma\beta}{\Gamma(\alpha+\beta)} \\ \beta(\alpha,\beta) = \int_0^1 x^{\alpha-1}(1-x)^{\beta-1} dx \\ \text{beta function} \end{array}}$$

Now The posterior distribution is

$$f(\theta/y) = \frac{f(y,\theta)}{f(y)}$$

$$= \frac{\cancel{\binom{n}{y}} \theta^y (1-\theta)^{n-y}}{\cancel{\binom{n}{y}} \beta(y+1, n-y+1)}$$

(i) The Posterior Bayes estimator of θ with respect to uniform prior distribution, given the data, is given by

$$E(\theta / X_1 = x_1, \dots, X_n = x_n) = \int_0^1 \theta \cdot f(\theta/x_1, \dots, x_n) d\theta$$

$$= \frac{1}{\beta(y+1, n-y+1)} \int_0^1 \theta \cdot \theta^y (1-\theta)^{n-y} d\theta$$

$$= \frac{\beta(y+2, n-y+1)}{\beta(y+1, n-y+1)} = \frac{\Gamma(y+2)\,\Gamma(n-y+1)}{\Gamma(n+3)} \cdot \frac{\Gamma(n+2)}{\Gamma(y+1) \cdot \Gamma(n-y+1)}$$

$$= \frac{y+1}{n+2} = \frac{\sum_{i=1}^{n} x_i + 1}{n+2}$$

Hence the posterior bayes estimator of ~~w.r.t~~ θ w.r.t uniform prior distribution is given by $\dfrac{\sum X_i + 1}{n+2}$

(ii) To obtain the posterior Bay's estimator of $\tau(\theta) = \theta(1-\theta)$ we have

$$E[\tau(\theta) | X_1 = x_1, \cdots, X_n = x_n] = \int_0^1 \theta(1-\theta) f(\theta / x_1, \cdots, x_n) \, d\theta$$

$$= \int_0^1 \frac{\theta(1-\theta)\theta^{y}(1-\theta)^{n-y}}{\beta(y+1, n-y+1)} \, d\theta = \frac{\beta(y+2, n-y+2)}{\beta(y+1, n-y+1)}$$

$$= \frac{(y+1)(n-y+1)}{(n+2)(n+3)} = \frac{(\sum x_i + 1)(n - \sum x_i + 1)}{(n+2)(n+3)}$$

So the posterior Bayes estimator of $\theta(1-\theta)$ w.r.t uniform prior distribution is given by

$$\frac{(\sum X_i + 1)(n - \sum X_i + 1)}{(n+2)(n+3)}$$

Example 2:- Let $X_1, X_2, \cdots, X_n$ be iid Bernoulli (θ). Then $Y = \sum X_i$ is binomial (n, p). We assume the prior distribution of θ is beta(α, β). Find the Bayes estimator of θ.

Solution: The joint distribution of Y & θ is

$$f(y, \theta) = f(y/\theta) \times \pi(\theta)$$

$$f(y, \theta) = \binom{n}{y} \theta^{y} (1-\theta)^{n-y} \times \frac{\Gamma(\alpha+\beta)}{\Gamma(\alpha)\Gamma(\beta)} \theta^{\alpha-1} (1-\theta)^{\beta-1}$$

$$= \binom{n}{y} \theta^{y+\alpha-1} (1-\theta)^{n-y+\beta-1} \frac{\Gamma(\alpha+\beta)}{\Gamma(\alpha)\Gamma(\beta)}$$

And

$$f(y) = \int_0^1 f(y, \theta) \, d\theta = \binom{n}{y} \frac{\Gamma(\alpha+\beta)}{\Gamma(\alpha)\Gamma(\beta)} \frac{\Gamma(y+\alpha)\Gamma(n-y+\beta)}{\Gamma(n+\alpha+\beta)}$$

$$\therefore \; f(\theta/y) = \frac{f(y,\theta)}{f(y)} = \frac{\Gamma(n+\alpha+\beta)}{\Gamma(y+\alpha)\,\Gamma(n-y+\beta)}\,\theta^{y+\alpha-1}(1-\theta)^{n-y+\beta-1}$$

which is $B(y+\alpha, n-y+\beta)$

The mean of posterior distribution is

$$E(\theta/X_1=x_1,\dots,X_n=x_n) = \int \theta f(\theta/y)\,d\theta$$

$$= \frac{(y+\alpha)}{\alpha+\beta+n}$$

Similarly we can solve for $\tau(\theta) = \theta(1-\theta)$

$$E[\tau(\theta)/X_1=x_1,\dots,X_n=x_n] = \int \theta(1-\theta) f(\theta/x_1,\dots,x_n)\,d\theta$$

?

Example 3:- Let $X_1, X_2, \dots, X_n$ be a random sample from Poisson distribution with mean θ, $0<\theta<\infty$. Let $Y = \sum_{i=1}^{n} X_i$. Assume prior distribution of θ is $G(\alpha,\beta)$. Find Bayes estimator of θ.

Solution

We have the following model

$X_i/\theta \sim$ iid. Poisson(θ)

$\theta \sim \Gamma_{\alpha,\beta}$ α & β are known

Thus, the joint distribution of X and θ is given by

$$f(x,\theta) = L(x/\theta \times \pi(\theta))$$

$$= \left(\frac{\theta^{x_1}e^{-\theta}}{x_1!}\cdots\frac{\theta^{x_n}e^{-\theta}}{x_n!}\right)\cdot\frac{\theta^{\alpha-1}e^{-\theta/\beta}}{\Gamma\alpha\,\beta^{\alpha}}$$

$$= \frac{\theta^{\sum x_i+\alpha-1}\,e^{-(n+\frac{1}{\beta})\theta}}{x_1!\cdots x_n!\,\Gamma\alpha\,\beta^{\alpha}} \qquad 0<\theta<\infty$$

The marignal distribution of sample is

$$f(x) = \int_0^\infty \frac{\theta^{\sum x_i + \alpha - 1} e^{-(n+\frac{1}{\beta})\theta}}{x_1! \cdots x_n! \Gamma\alpha \, \beta^\alpha} d\theta$$

$$= \frac{\Gamma(\sum x_i + \alpha)}{x_1! \cdots, x_n! \Gamma\alpha \, \beta^\alpha (n+\frac{1}{\beta})^{\sum x_i + \alpha}}$$

Hence the posterior pdf of θ, given $X = x$ is

$$f(\theta/x) = \frac{f(x,\theta)}{f(x)} = \frac{\theta^{\sum x_i + \alpha + 1} e^{-\frac{\theta}{\beta/(n\beta+1)}}}{\Gamma(\sum x_i + \alpha) \left[\frac{\beta}{(n\beta+1)}\right]^{\sum x_i + \alpha}}$$

$0 < \theta < \infty$

Thus $\theta/X \sim \Gamma\left(\sum_{i=1}^n X_i + \alpha, \frac{\beta}{n\beta+1}\right)$

$$\sim \Gamma\left(Y + \alpha = \alpha^*(\text{say}), \frac{\beta}{n\beta+1} = \beta^*(\text{say})\right) \text{ with } Y = \sum_{i=1}^n X_i$$

The Bayes Estimator is the conditional mean of the posterior distribution, given the data. That is

$$E(\theta / X_1 = x_1, \cdots, X_n = x_n) = \int_0^\theta \theta \cdot f(\theta/x) \, d\theta$$

$$= \alpha^* \beta^* = \frac{(Y+\alpha)\beta}{(n\beta+1)}$$

$$= \frac{(\sum x_i + \alpha)\beta}{n\beta+1}$$

Example 4:- Let $X \sim N(\mu,1)$ and let the prior pdf of μ is $N(0,1)$. Find the Bayes estimator of μ.

Solution

$$X \sim N(\mu,1) \Rightarrow f(x/\mu) = \frac{1}{\sqrt{2\pi}} e^{-\frac{1}{2}(x_i-\mu)^2}$$

$$\mu \sim N(0,1) \Rightarrow \pi(\mu) = \frac{1}{\sqrt{2\pi}} e^{-\frac{1}{2}\mu^2}$$

$$f(x,\mu) = f(x/\mu) \times \pi(\mu)$$

$$= \frac{1}{\sqrt{2\pi}} e^{-\frac{1}{2}(x_1-\mu)^2} \cdots \frac{1}{\sqrt{2\pi}} e^{-\frac{1}{2}(x_n-\mu)^2} \times \frac{1}{\sqrt{2\pi}} e^{-\frac{1}{2}\mu^2}$$

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

Now $$f(x) = \int_{-\infty}^{\infty} f(x,\mu)\, d\mu$$

$$= \frac{1}{(2\pi)^{\frac{n+1}{2}}} \exp\left(-\frac{1}{2}\sum_{i=1}^{n} x_i^2\right) \int_{-\infty}^{\infty} \exp\left[-\frac{n+1}{2}\left(\mu^2 - 2\mu\frac{n\bar{x}}{n+1}\right)\right] d\mu$$

$$= \frac{(n+1)^{-1/2}}{(2\pi)^{n/2}} \exp\left[-\frac{1}{2}\sum_{i=1}^{n} x_i^2 + \frac{n^2\bar{x}^2}{2(n+1)}\right]$$

It follows that

$$f(\mu/x) = \frac{1}{\sqrt{2\pi/(n+1)}} \exp\left[-\frac{(n+1)}{2}\left(\mu - \frac{n\bar{x}}{n+1}\right)^2\right]$$

& the Bayes Estimator is

$$E[\mu/X_1=x_1,\ldots,X_n=x_n] = \frac{n\bar{x}}{n+1} = \frac{\sum_{i=1}^{n} X_i}{(n+1)}$$

Example 5:- Let $X \sim N(\theta, \sigma^2)$ and suppose that the prior distribution on θ is $N(\mu, \tau^2)$. (Here we assume that σ^2, μ, τ^2 are all known) Find the Bayes Estimator of θ.

Solution

$$E(\theta / X_1 = x_1, \ldots, X_n = x_n) = \frac{\tau^2}{\tau^2 + \sigma^2} x + \frac{\sigma^2}{\sigma^2 + \tau^2} \mu$$

Assignment # 1 (Part II)

Q1:- Let $X_1, \ldots, X_n$ be a random sample from $P(\lambda)$. For estimating λ, using prior distribution $\pi(\lambda) = e^{-\lambda}$ if $\lambda > 0$. Find the Bayes Estimator for λ and $\phi(\lambda) = e^{-\lambda}$

Q2:- Let $X_1, \ldots, X_n$ be iid from $U(0, \theta)$. Suppose that the prior distribution for θ is $\pi(\theta) = \alpha a^{\alpha} / \theta^{\alpha+1}$ for $\theta \geq a$. Find the Bayes Estimator of θ.

Q3:- Let $X_1, \ldots, X_n$ be iid from $G(1, 1/\lambda)$. To estimate λ Let the prior distribution of λ be $\pi(\lambda) = e^{-\lambda}$, $\lambda > 0$. Find the Bayes estimator of λ.

Q4:- Let $X_1, \ldots, X_n$ be iid with pdf $f(x/\theta) = e^{-(x-\theta)}$, $x > \theta$. Take $\pi(\theta) = e^{-\theta}$, $\theta > 0$. Find Bayes estimator of θ.

Solutions:-

Q1:-

$$P(\lambda) = \frac{\lambda^{x} e^{-\lambda}}{x!}$$

$$f(x/\lambda) = \frac{\lambda^{x_1} e^{-\lambda}}{x_1!} \cdot \frac{\lambda^{x_2} e^{-\lambda}}{x!} \cdots \frac{\lambda^{x_n} e^{-\lambda}}{x_n!}$$

Prior distribution λ is $\pi(\lambda) = e^{-\lambda}$

The joint distribution of x and λ is

$$f(x, \lambda) = \frac{\lambda^{x_1} e^{-\lambda}}{x_1!} \cdot \frac{\lambda^{x_2} e^{-\lambda}}{x_2!} \cdots \frac{\lambda^{x_n} e^{-\lambda}}{x_n!} \times e^{-\lambda}$$

$$\Rightarrow f(x, \lambda) = \frac{\lambda^{\sum x_i} e^{-(n+1)\lambda}}{x_1!\, x_2! \cdots x_n!}$$

$$f(x) = \int_0^\infty \frac{\lambda^{\sum x_i} e^{-(n+1)\lambda}}{x_1!\, x_2! \cdots x_n!} \, d\lambda$$

Let $(n+1)\lambda = y \quad \Rightarrow \lambda = y/n+1$

$$d\lambda = \frac{dy}{n+1}$$

$$\Rightarrow f(x) = \frac{\int_0^\infty \left(\frac{y}{n+1}\right)^{\sum x_i} e^{-y} \left(\frac{1}{n+1}\right) dy}{x_1!\, x_2! \cdots x_n!}$$

$$= \left(\frac{1}{n+1}\right)^{\sum x_i + 1} \cdot \frac{\int_0^\infty y^{(\sum x_i + 1) - 1} e^{-y} \, dy}{x_1!\, x_2! \cdots x_n!}$$

$$= \frac{\boxed{\sum x_i + 1}}{x_1!\, x_2! \cdots x_n!\, (n+1)^{\sum x_i + 1}}$$

The posterior distribution is

$$f(\lambda/x) = \frac{f(x, \lambda)}{f(x)}$$

$$\Rightarrow f(\lambda/x) = \frac{\lambda^{\sum x_i} e^{-(n+1)\lambda}}{x_1!\, x_2! \dots x_n!} \times \frac{x_1!\, x_2! \dots x_n!\, (n+1)^{\sum x_i + 1}}{\Gamma(\sum x_i + 1)}$$

$$= \frac{\lambda^{\sum x_i} e^{-(n+1)\lambda}}{\Gamma(\sum x_i + 1) \left(\frac{1}{n+1}\right)^{\sum x_i + 1}}$$

$$\lambda/x \sim \Gamma\left(\sum x_i + 1, \frac{1}{n+1}\right)$$

$$\sim \Gamma\left(y+1 = \alpha^*, \frac{1}{n+1} = \beta^*\right)$$

$$E(\lambda/x_1, \dots x_n) = \int_0^\infty \lambda\, f(\lambda/x)\, dx$$

$$= \alpha^* \beta^*$$

$$= (y+1)\left(\frac{1}{n+1}\right)$$

$$\Rightarrow E(\lambda) = \frac{\left(\sum_{i=1}^n x_i + 1\right)}{n+1}$$

$$P(\lambda) = e^{-\lambda}$$

$$E(e^{-\lambda}) = \frac{\int_0^\infty e^{-\lambda} \lambda^{\sum x_i} e^{-(n+1)\lambda}\, d\lambda}{\Gamma(\sum x_i + 1) \left(\frac{1}{n+1}\right)^{\sum x_i + 1}}$$

$$= \frac{\int_0^\infty \lambda^{\sum x_i} e^{-(n+2)\lambda}\, d\lambda}{\Gamma(\sum x_i + 1) \left(\frac{1}{n+1}\right)^{\sum x_i + 1}}$$

Let $y' = (n+2)\lambda \Rightarrow \lambda = y'/_{n+2}$

& $d\lambda = dy'/_{n+2}$

$$\Rightarrow E(e^{-\lambda}) = \frac{\int_0^\infty e^{-y'} (y'/_{n+2})^{\Sigma x_i} \cdot \frac{1}{n+2} \, dy'}{\Gamma(\Sigma x_i + 1) \left(\frac{1}{n+1}\right)^{\Sigma x_i + 1}}$$

$$= \frac{\Gamma(\Sigma x_i + 1) \left(\frac{1}{n+2}\right)^{\Sigma x_i + 1}}{\Gamma(\Sigma x_i + 1) \left(\frac{1}{n+1}\right)^{\Sigma x_i + 1}}$$

$$\Rightarrow E(e^{-\lambda}) = \left(\frac{n+1}{n+2}\right)^{\Sigma x_i + 1}$$

Q2:-

$$f(x/_\theta) = 1/_\theta$$

$$\pi(\theta) = \alpha a^\alpha / \theta^{\alpha+1}$$

$$f(x/_\theta) = L(x/_\theta) \times \pi(\theta)$$

$$= \frac{1}{\theta} \cdot \frac{1}{\theta} \cdots \frac{1}{\theta} \times \frac{\alpha a^\alpha}{\theta^{\alpha+1}}$$

$$= \left(\frac{1}{\theta}\right)^n \frac{\alpha a^\alpha}{\theta^{\alpha+1}} = \frac{1}{\theta^{n+\alpha+1}} \alpha a^\alpha$$

$$\Rightarrow f(x) = \int_a^\infty \theta^{-n-\alpha-1} \alpha a^\alpha \, d\theta$$

$$= \alpha a^\alpha \left. \frac{\theta^{-n-\alpha}}{-n-\alpha} \right|_a^\infty$$

$$= \alpha a^\alpha \left[\frac{a^{-n-\alpha}}{n+\alpha} \right]$$

$$= \frac{\alpha a^\alpha}{n+\alpha} \cdot \frac{1}{a^{n+\alpha}}$$

$$= \frac{\alpha}{a^n (n+\alpha)}$$

Posterior distribution is

$$f(\theta/x) = \frac{f(x,\theta)}{f(x)} = \frac{1}{\theta^{n+\alpha+1}} \alpha a^{\alpha} \times \frac{a^{n}(n+\alpha)}{\alpha}$$

$$\Rightarrow f(\theta/x) = \frac{a^{\alpha+n}(n+\alpha)}{\theta^{n+\alpha+1}}$$

$$\Rightarrow E(\theta/x_1, x_2, \ldots, x_n) = \int_{a}^{\infty} \theta \frac{a^{\alpha+n}(n+\alpha)}{\theta^{n+\alpha+1}} d\theta$$

$$= a^{\alpha+n}(n+\alpha)\int_{a}^{\infty} \theta^{-n-\alpha} d\theta$$

$$= a^{\alpha+n}(n+\alpha)\left|\frac{\theta^{-n-\alpha+1}}{-n-\alpha+1}\right|_{a}^{\infty}$$

$$= \frac{(n+\alpha)a^{\alpha+n}}{-(n+\alpha-1)}\left[-a^{-n-\alpha+1}\right]$$

$$\Rightarrow E(\theta) = \frac{n+\alpha}{n+\alpha-1} \cdot a$$

Q3:- $$f(x/\lambda) = \frac{1}{\Gamma(1/\lambda)} \cdot x^{1-1} e^{-x/\frac{1}{\lambda}}$$

$$= \lambda e^{-x\lambda}$$

$$\Rightarrow f(x,\lambda) = L(x/\lambda) \times \pi(x)$$

$$= \left[\lambda e^{-x_1\lambda} \cdot \lambda e^{-x_2\lambda} \cdots \lambda e^{-\lambda x_n}\right] e^{-\lambda}$$

$$= \left[\lambda^{n} e^{-\lambda\Sigma x_i}\right] e^{-\lambda}$$

$$\Rightarrow f(x,\lambda) = \lambda^n e^{-\lambda(\sum x_i + 1)}$$

$$\Rightarrow f(x) = \int_0^\infty \lambda^n e^{-\lambda(\sum x_i + 1)} \, d\lambda$$

Let $\lambda(\sum x_i + 1) = y \quad \Rightarrow \lambda = y/(\sum x_i + 1)$

$$\Rightarrow d\lambda = \frac{dy}{\sum x_i + 1}$$

$$\Rightarrow f(x) = \int_0^\infty \left(\frac{y}{\sum x_i + 1}\right)^n \cdot e^{-y} \frac{dy}{\sum x_i + 1}$$

$$= \left[\frac{1}{(\sum x_i + 1)}\right]^{n+1} \int_0^\infty y^n e^{-y} \, dy$$

$$= \frac{\Gamma(n+1)}{(\sum x_i + 1)^{n+1}}$$

$$\Rightarrow f(\lambda/x) = \frac{f(x,\lambda)}{f(x)}$$

$$= \frac{\lambda^n e^{-[\sum x_i + 1]\lambda}}{\Gamma(n+1)} (\sum x_i + 1)^{n+1}$$

$$\Rightarrow E(\lambda) = \frac{\int_0^\infty \lambda^{n+1} e^{-(\sum x_i + 1)\lambda} \cdot (\sum x_i + 1)^{n+1} \, d\lambda}{\Gamma(n+1)}$$

$$= \frac{(\sum x_i + 1)^{n+1}}{\Gamma(n+1)} \int_0^\infty \lambda^{n+1} e^{-(\sum x_i + 1)\lambda} \, d\lambda$$

Let $(\sum x_i + 1)\lambda = y' \quad \Rightarrow \lambda = y'/\sum x_i + 1$

$$\Rightarrow d\lambda = dy'/\sum x_i + 1$$

$$\Rightarrow E(\lambda) = \frac{(\Sigma x_i + 1)^{n+1}}{\Gamma(n+1)} \int_0^\infty \left(\frac{y'}{\Sigma x_i + 1}\right)^{n+1} \cdot e^{-y'} \frac{dy'}{\Sigma x_i + 1}$$

$$= \frac{(\Sigma x_i + 1)^{n+1}}{\Gamma(n+1)\,(\Sigma x_i + 1)^{n+2}} \int_0^\infty y'^{(n+1)} e^{-y'} dy'$$

$$= \frac{(\Sigma x_i + 1)^{-1}\,\Gamma(n+2)}{\Gamma(n+1)}$$

$$= \frac{(n+1)\,\cancel{\Gamma(n+1)}}{\cancel{\Gamma(n+1)}\,(\Sigma x_i + 1)}$$

$$\Rightarrow E(\lambda) = \frac{n+1}{\Sigma x_i + 1}$$

Q 4:-

$$f(x/\theta) = e^{-(x-\theta)}$$

$$\pi(\theta) = e^{-\theta}$$

$$f(x,\theta) = L(x/\theta) \times \pi(\theta)$$

$$= e^{-(x_1-\theta)} e^{-(x_2-\theta)} \cdots e^{-(x_n-\theta)} \times e^{-\theta}$$

$$\Rightarrow f(x,\theta) = e^{-\Sigma x_i} e^{(n-1)\theta}$$

$$\Rightarrow f(x) = \int_0^x e^{(n-1)\theta} d\theta$$

$$= e^{-\Sigma x_i} \int_0^x e^{(n-1)\theta} d\theta$$

$$\Rightarrow f(x) = e^{-\sum x_i} \left. \frac{e^{(n-1)\theta}}{n-1} \right|_0^x$$

$$= \frac{e^{-\sum x_i}}{n-1} \left(e^{(n-1)x} - e^{(n-1)0} \right).$$

$$\Rightarrow f(x) = \frac{e^{-\sum x_i}}{n-1} \left[e^{(n-1)x} - 1 \right]$$

Now

$$f(\theta/x) = \frac{f(x,\theta)}{f(x)} = e^{-\sum x_i} e^{(n-1)\theta} \times \frac{n-1}{e^{-\sum x_i}\left(e^{(n-1)x} - 1\right)}$$

$$= \frac{e^{(n-1)\theta} (n-1)}{e^{(n-1)x} - 1}$$

$$E(\theta/x_1, x_2, \dots, x_n) = \int_0^x \frac{\theta \cdot e^{(n-1)\theta}\, n-1}{e^{(n-1)x} - 1} d\theta$$

$$= \frac{(n-1)}{e^{(n-1)x} - 1} \int_0^x \theta \cdot e^{(n-1)\theta} d\theta$$

$$= \frac{(n-1)}{e^{(n-1)x} - 1} \left[\theta \cdot \frac{e^{(n-1)\theta}}{n-1} \Big|_0^x - \int_0^x e^{(n-1)\theta} d\theta \right.$$

$$= \frac{(n-1)}{e^{(n-1)x} - 1} \left[\frac{x e^{(n-1)x}}{n-1} - \frac{e^{(n-1)x}}{n-1} + \frac{1}{n-1} \right]$$

$$= \frac{1}{e^{(n-1)x} - 1} \left[x e^{(n-1)x} - e^{(n-1)x} + 1 \right]$$

$$\Rightarrow E(\theta) = \frac{e^{(n-1)x}}{e^{(n-1)x}-1}\left(x-1+\frac{1}{e^{(n-1)x}}\right)$$

Ans 5 joint density function of given function is

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

$$= \frac{\theta}{(1+x_1)^{\theta+1}} \cdot \frac{\theta}{(1+x_2)^{\theta+1}} \cdots \frac{\theta}{(1+x_n)^{\theta+1}}$$

$$= \frac{\theta^n}{[(1+x_1)(1+x_2)\cdots(1+x_n)]^{\theta+1}}$$

$$\Rightarrow L(\theta, x_i) = \frac{\theta^n}{\prod_{i=1}^{n}(1+x_i)^{\theta+1}}$$

Taking ln on both sides

$$l = \ln L(\theta, x_i) = \ln \theta^n - (\theta+1)\ln \sum_{i=1}^{n}(1+x_i)$$

differentiate w.r.t θ & Put $\frac{dl}{d\theta} = 0$

$$\frac{dl}{d\theta} = \frac{n}{\theta} - \ln \sum(1+x_i)$$

$$0 = \frac{n}{\theta} - \ln \sum(1+x_i)$$

$$\Rightarrow \ln \sum(1+x_i) = \frac{n}{\theta}$$

$$\Rightarrow \boxed{\hat{\theta} = \frac{n}{\ln \sum_{i=1}^{n}(1+x_i)}}$$

Ans 2 Here

$$E(X) = \exp\left(\mu + \frac{1}{2}\sigma^2\right)$$

& $$E(X^2) = \exp(2(\mu+\sigma^2)) - \exp(2\mu+\sigma^2) + \exp\left(\mu + \frac{1}{2}\sigma^2\right)^2$$

Now

$$m_1 = \mu_1'$$

$$\frac{\sum x_i}{n} = E(X)$$

$$\bar{X} = \exp\left(\mu + \frac{1}{2}\sigma^2\right)$$

$$\ln \bar{X} = \mu + \frac{1}{2}\sigma^2 \longrightarrow (1)$$

& $$m_2 = E(X^2)$$

$$\frac{\sum x_i^2}{n} = e^{2\mu+2\sigma^2} - e^{2\mu+\sigma^2} + \left(e^{(\mu+\frac{1}{2}\sigma^2)}\right)^2$$

$$= e^{2\mu} \cdot e^{2\sigma^2} - e^{2\mu} e^{\sigma^2} + \left(e^{\ln \bar{X}}\right)^2 \quad \text{by (1)}$$

$$= e^{2\mu} \cdot e^{\sigma^2} (e^{\sigma^2} - 1) + \bar{X}^2$$

$$\frac{\sum x_i^2}{n} - \bar{X}^2 = e^{2\mu+\sigma^2}(e^{\sigma^2} - 1)$$

$$s^2 = e^{2\ln \bar{X}} (e^{\sigma^2} - 1) \quad \text{by (1)}$$

$$s^2 = \bar{X}^2 (e^{\sigma^2} - 1)$$

$$\frac{s^2}{\bar{X}^2} + 1 = e^{\sigma^2}$$

$$\Rightarrow \boxed{\hat{\sigma}^2 = \ln\left(1+\frac{S^2}{\bar{X}^2}\right)}$$

① $\Rightarrow \ln \bar{X} = \mu + \frac{1}{2}\ln\left(1+\frac{S^2}{\bar{X}^2}\right)$

$$\ln \bar{X} - \ln\sqrt{\left(1+\frac{S^2}{\bar{X}^2}\right)} = \mu$$

$$\Rightarrow \boxed{\hat{\mu} = \ln\left(\bar{X}\Big/\sqrt{1+\frac{S^2}{\bar{X}^2}}\right)}$$

MUHAMMAD TAHIR WATTOO ***

M.S. MATH CIIT ISLAMABAD

Ph.D MATH CU ISLAMABAD
SP18-PMT-005

0344 8563284

★ Criteria to Compare Estimators

Suppose we have several Estimators of the unknown parameter θ or $g(\theta)$ (real valued parameteric function of θ). We compare the performance of rival estimators on the basis of

1) Unbiasedness
2) Variance / Mean Square Error (MSE).
3) Consistency
4) Efficiency
5) Sufficiency

⇒ **Unbiasedness:-** A real values statistic or estimator $\hat{\theta}$ is said to be an unbiased estimator of a unknown parameter θ, iff

$$E(\hat{\theta}) = \theta \quad \forall\ \theta \in \Omega$$

Bias:- The Bias of an estimator $\hat{\theta}$ of a parameter θ is defined as

$$Bias(\theta) = E(\hat{\theta}) - \theta \quad , \theta \in \Omega$$

⇒ **Mean Square Error:-** The Mean Square error (MSE) of an estimator $\hat{\theta}$ of a parameter θ is defined by $E(\hat{\theta}-\theta)^2$

$$E(\hat{\theta}-\theta)^2 = E[\hat{\theta} - E(\hat{\theta}) + E(\hat{\theta}) - \theta]^2$$

$$= Var(\hat{\theta}) + [Bias(\theta)]^2$$

Example 1:- Let $X_1, X_2, \cdots, X_n$ iid RVs from $N(\mu, \sigma^2)$, where μ is unknown and $\sigma^2 > 0$ is assumed known. Let us consider several rival estimators of μ defined as below

$$T_1 = X_1 + X_4 \quad , \quad T_2 = \frac{1}{2}(X_1 + X_3)$$

$$T_3 = \bar{X} \quad , \quad T_4 = \frac{1}{3}(X_1 + X_3)$$

$$T_5 = X_1 + T_2 - X_4 \quad , \quad T_6 = \frac{1}{10}\sum_{i=1}^{4} iX_i$$

Based on $X_1, \cdots, X_4$, one can certainly form many other estimators for μ. observe that

$$E(T_1) = 2\mu \, , \, E(T_2) = \mu \, , \, E(T_3) = \mu$$

$$E(T_4) = \frac{2}{3}\mu \, , \, E(T_5) = \mu \;\&\; E(T_6) = \mu$$

Thus T_1 and T_4 are biased estimators and T_2, T_3, T_6, T_5 are unbiased estimators of μ.

Example 2:- Let $X_1, X_2, \cdots, X_n$ are iid with finite mean μ & variance $\sigma^2 > 0$ then

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

are unbiased estimators of μ and σ^2

Solution Given $E(X_i) = \mu$ & $V(X_i) = \sigma^2$

We have to show that

$$E(\bar{X}) = \mu \quad \& \quad E(S^2) = \sigma^2$$

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

$$= \mu \quad \text{Proved}$$

Thus $\bar{X}$ is an unbiased estimator of μ.

$$E[(n-1)S^2] = \sum_{i=1}^{n} E(X_i - \bar{X})^2 = \sum_{i=1}^{n} E(X_i^2) - nE[\bar{X}^2]$$

Now $Var(\bar{X}) = E(\bar{X}^2) - [E(\bar{X})]^2$

$$\because Var(\bar{X}) = Var\left(\frac{1}{n}\sum_{i=1}^{n} X_i\right) = \frac{1}{n^2} Var(\sum X_i)$$

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

$$V(X_i) = E(X_i^2) - [E(X_i)]^2$$

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

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

&

$$E(\bar{X}^2) = V(\bar{X}) + [E(\bar{X})]^2$$

$$= \frac{\sigma^2}{n} + \sigma^2$$

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

$$= n\sigma^2 + \cancel{n\mu^2} - \sigma^2 - \cancel{n\mu^2}$$

$$= (n-1)\sigma^2$$

$$\Rightarrow E(s^2) = \frac{(n-1)\sigma^2}{(n-1)}$$

$$\Rightarrow E(s^2) = \sigma^2 \quad \text{Proved}$$

Thus s^2 is an unbiased estimator of σ^2.

⇒ Best Unbiased Estimator:-

For comparing the performance of unbiased estimators is done by comparing the variance of the rival estimators. So if T_1 and T_2 are two unbiased estimators of $g(\theta)$ then T_1 is preferable to (or better than) T_2 if

$$V(T_1) \leq V(T_2) \text{ for all } \theta \in \Omega$$

Now in the class of unbiased estimators of $g(\theta)$, the one having smallest variance is called the best unbiased estimator of $g(\theta)$

⇒ Uniformly Minimum Variance Unbiased Estimator (UMVUE):-

Assume that there is atleast one unbiased estimator of the unknown real valued parameter of $g(\theta)$. An estimator $T \in C$ is called the UMVUE of $g(\theta)$ iff for all estimators $T^* \in C$ we have

$$V(T) \leq V(T^*) \quad \forall \; \theta \in \Omega$$

There are several approaches to locate the UMVUE. In searching for an optimal estimate, we might ask whether there is a lower bound for the MSE of any estimate. If such lower bound existed, it would function as a benchmark against which estimates could be compared. If an estimate achieved this lower bound, we would know that it could not be improved upon.

In the case in which the estimate is unbiased, the Cramer-Rao Inequality provides such a lower bound.

⇒ Cramer-Rao Inequality:-

Let $X_1, X_2, \ldots, X_n$ be iid with density function $f(x/\theta)$. Let $T = t(X_1, \cdots, X_n)$ be an unbiased estimate of $g(\theta)$. Then, under smoothness assumptions on $f(x/\theta)$,

$$Var_\theta(T) \geq \frac{[g'(\theta)]^2}{n I_1(\theta)}$$

where $I_1(\theta) = E_\theta\left[\frac{\partial \ln f(X_1)}{\partial \theta}\right]^2 = -E_\theta\left(\frac{\partial^2 \ln f(X_1)}{\partial \theta^2}\right)$

and is called Fisher information[I] in X_1

and $I_n(\theta) = E_\theta\left[\frac{\partial \ln f(X)}{\partial \theta}\right]^2 = n I_1(\theta)$ is →

[I]:- Information about the parameter θ provided by the sample.

known as Fisher estimation in the random sample $X_1, \ldots, X_n$

Example:- Let $X_1, X_2, \ldots, X_n$ be iid random variables from $B(1,\theta)$

Solution We have $f(x/\theta) = \theta^x (1-\theta)^{1-x}$; $x = 0,1$

$$\ln f(x/\theta) = x \ln\theta + (1-x)\ln(1-\theta)$$

$$\frac{\partial \ln f(x/\theta)}{\partial\theta} = \frac{x}{\theta} - \frac{(1-x)}{1-\theta}$$

$$\frac{\partial^2 \ln f(x/\theta)}{\partial\theta^2} = \frac{-x}{\theta^2} - \frac{(1-x)}{(1-\theta)^2}$$

$$= \frac{-[x(1-2\theta+\theta^2)+\theta^2(1-x)]}{\theta^2(1-\theta)^2}$$

$$= \frac{-[x-2\theta x+\cancel{x\theta^2}+\theta^2-\cancel{x\theta^2}]}{\theta^2(1-\theta)^2}$$

$$= \frac{-[x(1-2\theta)+\theta^2]}{\theta^2(1-\theta)^2}$$

$E(X) = \sum_0^1 x\theta^x(1-\theta)^{1-x}$

$\therefore E(X) = \theta$

$E(X^2) = \theta$

$V(X) = \theta - \theta^2$

$= \theta(1-\theta)$

$\bar{X} = \frac{\sum X_i}{n}$

$V(\bar{X}) = \frac{1}{n^2}\sum V(X_i)$

$= \frac{n \cdot \theta(1-\theta)}{n^2}$

$V(\bar{X}) = \frac{\theta(1-\theta)}{n}$

$$I_1(\theta) = -E\left[\frac{\partial^2 \ln f(x/\theta)}{\partial\theta^2}\right] = \frac{(1-2\theta)\theta+\theta^2}{\theta^2(1-\theta)^2}$$

$$= \frac{\theta-\theta^2}{\theta^2(1-\theta)^2} = \frac{1}{\theta(1-\theta)}$$

So the Cramer-Rao lower bound is equal to $\frac{\theta(1-\theta)}{n}$. Now $\bar{X}$ is an Unbiased estimator of θ and its variance is $V(\bar{X}) = \frac{\theta(1-\theta)}{n}$, that is equal to Cramer-Rao bound. Therefore $\bar{X}$ is UMVUE of θ.

Example:- Let $X_1, X_2, \dots, X_n$ be iid random variables from $P(\lambda)$; $\lambda > 0$, by setting $\lambda = \theta$ we have

$$f(x;\theta) = e^{-\theta} \theta^x / x!$$

$$\ln f(x;\theta) = -\theta + x \ln\theta - \ln x!$$

$$\frac{\partial \ln f(x;\theta)}{\partial\theta} = -1 + x/\theta$$

$$\frac{\partial^2 \ln f(x;\theta)}{\partial\theta^2} = -x/\theta^2$$

$$I_1(\theta) = -E_\theta\left[\frac{\partial^2 \ln f(x;\theta)}{\partial\theta^2}\right] = \frac{E(x)}{\theta^2}$$

$$= \frac{\theta}{\theta^2} = 1/\theta$$

$$\therefore I_n(\theta) = n I_1(\theta) = \frac{n}{\theta}$$

So the Cramer-Rao bound is equal to θ/n. Since $\bar{X}$ is an Unbiased estimator of θ with variance θ/n. We have that $\bar{X}$ is a UMVUE estimator of θ.

Example:- Let $X_1, X_2, \ldots, X_n$ be iid random variables from $N(\mu, \sigma^2)$. Assume that σ^2 is known and set $\mu = \theta$, then

$$f(x;\theta) = \frac{1}{\sqrt{2\pi}\,\sigma} e^{-\frac{1}{2\sigma^2}(x-\theta)^2}; \; x \in \mathbb{R}$$

$$\therefore \ln f(x;\theta) = \ln\left(\frac{1}{\sqrt{2\pi}\,\sigma}\right) - \frac{1}{2\sigma^2}(x- \ldots$$

$$\frac{\partial \ln f(x;\theta)}{\partial \theta} = \frac{(x-\theta)}{\sigma^2}$$

$$\frac{\partial^2 \ln f(x;\theta)}{\partial \theta^2} = -\frac{1}{\sigma^2}$$

$$-E_\theta\left[\frac{\partial^2 \ln f(x;\theta)}{\partial \theta^2}\right] = \frac{1}{\sigma^2} = I_1(\theta)$$

$$\text{\&} \quad I_n(\theta) = n I_1(\theta) = \frac{n}{\sigma^2}$$

Thus the CR bound is σ^2/n. Once again $\bar{X}$ is an unbiased estimator of θ and its Variance is equal to σ^2/n, i.e. the CR bound. Therefore $\bar{X}$ is UMVUE.

Show that if μ is known & set $\sigma^2 = \theta$, then

$$f(x;\theta) = \frac{1}{\sqrt{2\pi\theta}} \exp\left[\frac{(x-\mu)^2}{2\theta}\right]$$

$$\ln f(x;\theta) = -\frac{1}{2}\ln 2\pi - \frac{1}{2}\ln\theta - \frac{(x-\mu)^2}{2\theta}$$

$$\frac{\partial \ln f(x;\theta)}{\partial \theta} = \frac{-1}{2\theta} + \frac{(x-\mu)^2}{2\theta^2}$$

$$\frac{\partial^2 \ln f(x;\theta)}{\partial \theta^2} = \frac{1}{2\theta^2} - \frac{2(x-\mu)^2}{2\theta^3}$$

$$-E_\theta\left[\frac{\partial^2 \ln f(x;\theta)}{\partial \theta^2}\right] = \frac{-1}{2\theta^2} + \frac{E(x-\mu)^2}{\theta^3}$$

$$I_1(\theta) = \frac{1}{\theta^2} - \frac{1}{2\theta^2} = \frac{1}{2\theta^2} \quad \& \quad I_n(\theta) = nI_1(\theta) = \frac{n}{2\theta^2}$$

& CR bound is $2\theta^2/n$. Next

$$\sum_{j=1}^{n} \frac{(X_j-\mu)^2}{(\sqrt{\theta})^2} \sim \psi^2(n)$$

$Z = \frac{X-\mu}{\sigma} \to$ S.N.V

$X \sim N(\mu, \sigma^2)$

$\Rightarrow Z \sim N(0,1)$

$f(x) = \frac{1}{\sqrt{2\pi}\,\sigma} e^{-\frac{1}{2}\left(\frac{x-\mu}{\sigma}\right)^2}$

$\Rightarrow f(z) = \frac{1}{\sqrt{2\pi}} e^{-\frac{1}{2}z^2}$

So $$E\left[\sum_{j=1}^{n} \frac{(X_j-\mu)^2}{(\sqrt{\theta})^2}\right] = n$$

$$\Rightarrow E\left[\frac{1}{n}\sum_{j=1}^{n} (x_j-\mu)^2\right] = \theta$$

Therefore, $\frac{1}{n}\sum_{j=1}^{n}(x_j-\mu)^2$ is an unbiased estimator of θ and its variance is $2\theta^2/n$ equal to the CR bound.

Thus $\frac{1}{n}\sum_{j=1}^{n}(X_j-\mu)^2$ is UMVUE of θ

★ Properties of a good Estimator ($\hat{\theta}$):-

1) $\hat{\theta}$ is an unbiased estimator of θ if $E(\hat{\theta}) = \theta$
2) $\hat{\theta}$ is an asymptotically unbiased estimator of θ if

$$\lim_{n\to\infty} E(\hat{\theta}) = \theta$$

3) The bias of $\hat{\theta}$ is $E(\hat{\theta}) - \theta$
4) The Mean Square error of $\hat{\theta}$ is $E(\hat{\theta}-\theta)^2$
5) $\hat{\theta}$ is a consistent estimator of θ if

$$\lim_{n\to\infty} E(\hat{\theta}-\theta)^2 = 0$$

6) Let $\hat{\theta}_1$ & $\hat{\theta}_2$ be two unbiased estimators of θ and if $V(\hat{\theta}_1) < V(\hat{\theta}_2)$, we say that $\hat{\theta}_1$ is relatively more (best unbiased) efficient estimator than $\hat{\theta}_2$.

★ $MSE(\hat{\alpha}) = V(\hat{\alpha}) + [Bias(\hat{\alpha})]^2$

if Bias $= 0$

$\Rightarrow MSE(\hat{\alpha}) = V(\hat{\alpha})$

★ $X \sim exp(\beta) \Rightarrow x$ follows exponential distribution

$X \sim Neg\ exp(\beta) \Rightarrow x$ follows negative exponential

Example:- An electrical circuit consists of four batteries conneting in series to a light bulb. We model the batteries life time x_1, x_2, x_3 & x_4 as independent and identically distributed $U(0, \theta)$ random variables. Our experiment to measure the operating time of the circuit is stopped when any one of the batteries fails. hence, the only random variable we obtain is

$Y = \min(X_1, X_2, X_3, X_4)$

i) Determine the cdf of r.v. Y

ii) Write down the likelihood function of θ based on a single observation Y.

iii) Derive the MLE of θ

iv) Find the bias of estimator in (iii). Is the estimator unbiased

v) Find the Mean Square Error of the estimator in (iii)

Solution

i) $P(Y \leq y) = 1 - P[Y > y]$

$$= 1 - P[\min(X_1, X_2, X_3, X_4) > y]$$

$$\Rightarrow P(Y \leq y) = 1 - [P(X_1 > y) \times P(X_2 > y) \times P(X_3 > y) \times P(X_4 > y)]$$

$$= 1 - [P(X > y)]^4 \quad \because X_1, X_2, X_3 \text{ \& } X_4 \text{ are iid}$$

$$= 1 - [1 - P(X \leq y)]^4$$

$$\Rightarrow F_Y(y) = 1 - \left(1 - \frac{y}{\theta}\right)^4$$

$$\Rightarrow f_Y(y) = \frac{d}{dy} F_Y(y)$$

$X \sim U(0, \theta)$

$\Rightarrow f(x) = 1/\theta$

$\Rightarrow F(x) = P(X \leq x)$

$= \int_0^x \frac{1}{\theta}\, dx = \frac{1}{\theta} |x|$

$\Rightarrow F(x) = \frac{x}{\theta}$

P.d.f C.d.f

$$f(x) = \frac{d}{dx} F(x)$$

$$\Rightarrow f_Y(y) = \frac{4}{\theta^4}(\theta - y)^3$$

ii) $$\mathcal{L}(\theta) = \frac{4}{\theta^4}(\theta - y)^3$$

iii) $$\log \mathcal{L} = \log 4 + 3\log(\theta - y) - 4\log\theta$$

$$\Rightarrow \frac{\partial \log \mathcal{L}}{\partial \theta} = 0 + \frac{3}{(\theta - y)} - 4/\theta$$

$$\frac{\partial \log \mathcal{L}}{\partial \theta} = 0 \quad \Rightarrow \hat{\theta} = 4y$$

iv)

$$\text{Bias}(\hat{\theta}) = E(\hat{\theta}) - \theta$$

$$\Rightarrow \text{Bias}(\hat{\theta}) = 4\int_0^{\theta} y f_Y(y)\,dy - \theta = \frac{16}{\theta^4}\int_0^{\infty} y(\theta - y)^3\,dy - \theta$$

$$= 16\theta \int_0^1 y(1-y)^3\,dy - \theta$$

$$= \frac{4\theta}{5} - \theta$$

$$\Rightarrow \text{Bias}(\hat{\theta}) = \frac{-\theta}{5}$$

v) $$V(\hat{\theta}) = V(4y) = 16V(y)$$

Now $$V(y) = E(y^2) - [E(y)]^2$$

$$\Rightarrow V(\hat{\theta}) = 16[E(y^2) - (E(y))^2]$$

$$= \frac{64}{\theta^4}\int^{\theta} y^2(\theta - y)^3\,dy - \frac{16\theta^2}{25}$$

$$\Rightarrow V(\hat{\theta}) = \frac{16\theta^2}{15} - \frac{16\theta^2}{25}$$

$$= \frac{32\theta^2}{75}$$

$$\Rightarrow MSE(\hat{\theta}) = V(\hat{\theta}) + [Bias(\hat{\theta})]^2$$

$$= \frac{32\theta^2}{75} + \frac{\theta^2}{25} = \frac{32\theta^2 + 3\theta^2}{75}$$

$$= \frac{35\theta^2}{75} = \frac{7}{15}\theta^2$$

Now $\lim\limits_{n\to\infty} MSE(\hat{\theta}) = \lim\limits_{n\to\infty} \frac{7}{15}\theta^2 = \frac{7}{15}\theta^2$

$\Rightarrow \hat{\theta}$ is not a consistent estimator of θ (Also Biased)

⇒ Efficient Estimation:- Let $\hat{\theta}_1$ & $\hat{\theta}_2$ be two unbiased estimator of the population parameter θ of the population and if $V(\hat{\theta}_1) < V(\hat{\theta}_2)$, we say that $\hat{\theta}_1$ is relatively more efficient than $\hat{\theta}_2$. So we can define Relative efficiency (RE) of $\hat{\theta}_2$ with respect to $\hat{\theta}_1$ as follows
if $RE > 1$ indicates the superiority of the estimator $\hat{\theta}_2$ over $\hat{\theta}_1$.

⇒ Sufficient Estimation:-

Suppose the distribution of a random variable X is a function of an unknown parameter θ. For example, the distribution of a binomial random variable is a function of parameter p, the probability of success. If $X_1, \dots, X_n$ is a random sample with probability distribution $f(x, \theta)$. Now a random sample can be reduced via a function of sample information without discarding any information about the parameter θ. Consider the sample mean $\bar{X}$, which can be represented as a function of a random sample. That is

$$T = u(X_1, X_2, \cdots, X_n) = \frac{1}{n} \sum_{i=1}^{n} X_i$$

may contain all the information about the parameter θ. Thus it provides the relavent information in just one number instead displaying the entire random sample $X_1, \dots, X_n$ (in other words we can use one number versus n numbers). This argument leads to the definition of a sufficient statistic.

Def:- Let $X_1, X_2, \dots, X_n$ denote a random sample of size n from a distribution that has a pdf or pmf $f(x; \theta)$, $\theta \in \Omega$. Let $T = u(X_1, X_2, \dots, X_n)$ be a statistic whose pdf or pmf is $f_T[u(x_1, x_2, \dots, x_n); \theta]$.

Then T is a sufficient statistic for θ iff
$$\frac{f(x_1;\theta)\cdot f(x_2;\theta)\cdots f(x_n,\theta)}{f_T[u(x_1,x_2,\cdots,x_n);\theta]} = h(x_1,\cdots,x_n)$$

where $h(x_1,\cdots,x_n)$ does not depend on $\theta\in\Omega$. To avoid the compotational difficulty, it is much easier to use the following Factorization Theorem, to find ~~&~~ verify the sufficient statistic.

Theorem:- Let $X_1, X_2, \cdots, X_n$ be a random sample with joint density $f(x_1, x_2, \cdots, x_n;\theta)$. A statistic $T = u(X_1, X_2, \cdots, X_n)$ is said to be a sufficient statistic iff the joint can be factorized as follows
$$f(x_1, x_2, \cdots, x_n;\theta) = g(T,\theta)\, h(x_1, x_2, \cdots, x_n)$$

g and h are non-negative functions. Furthermore the function $g(T,\theta)$ can depend on θ and can depend on the random sample only through the value of $u(x_1,\cdots,x_n)$. On the other hand, the function h can depend on the unknown parameter θ.

Example 1:- Let $X_1, X_2, \cdots, X_n$ be a random sample from a Bernoulli distribution $f(x/p) = p^x(1-p)^{1-x}$ if $x = 0, 1$ & $0 \le p \le 1$

$= 0$ elsewhere

Find the sufficient statistic for P.

Solution The joint probability mass function

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

$$= p^{\sum_1^n x_i}(1-p)^{n-\sum_1^n x_i}$$

$$= p^t(1-p)^{n-t} \quad ; \text{where } t = \sum_{i=1}^{n} x_i$$

$$= g(t, p) \cdot h(x_1, x_2, \cdots, x_n)$$

where $g(t, p) = p^t(1-p)^{n-t}$ &

$h(x_1, x_2, \cdots, x_n) = 1$

Hence $T = \sum_{i=1}^{n} X_i$ is a sufficient statistic.

Example 2:- Let $X_1, X_2, \cdots, X_n$ be a random sample from probability density function

$$f(x/\theta) = \frac{1}{\theta} e^{-x/\theta} \quad 0 \le x \le \infty$$

$= 0$ elsewhere

Show that $\bar{X}$ is a sufficient statistic for θ.

Solution: The joint density function is given by

$$f(x_1, x_2, \cdots, x_n | \theta) = \prod_{i=1}^{n} \frac{1}{\theta} e^{-\frac{x_i}{\theta}}$$

$$= \frac{e^{-\sum_1^n x_i/\theta}}{\theta^n} = \frac{e^{-n\bar{x}/\theta}}{\theta^n}$$

where $g(\bar{x}, \theta) = \frac{e^{-n\bar{x}/\theta}}{\theta^n}$ & $h(x_1, x_2, \cdots, x_n) = 1$

so $\bar{X}$ is a sufficient statistic for θ

Example 3:- Let $X_1, X_2, \cdots, X_n$ be a random sample from $N(\mu, \sigma^2)$. Find the sufficient statistics for μ and σ^2.

Solution: We have $\bar{X} = \frac{1}{n}\sum_1^n X_i$ & $S^2 = \frac{1}{n-1}\sum_1^n (X_i - \bar{X})^2$

as estimators of μ & σ^2 respectively.

The joint density function is given by

$$f(x_1, x_2, \cdots, x_n | \mu, \sigma^2) = \prod_{i=1}^{n} (2\pi\sigma^2)^{-\frac{1}{2}} \exp\left[\frac{-1}{2\sigma^2}(x_i - \mu)^2\right]$$

$$= (2\pi\sigma^2)^{-\frac{n}{2}} \cdot \exp\left[\frac{-1}{2\sigma^2}\sum_1^n (x_i - \bar{x} + \bar{x} - \mu)^2\right]$$

$$= (2\pi\sigma^2)^{-\frac{n}{2}} \cdot \exp\left[\frac{-1}{2\sigma^2}\left\{\sum_1^n (x_i - \bar{x})^2 + n(\bar{x} - \mu)^2 + 2(\bar{x} - \mu)\sum_1^n (x_i - \bar{x})\right\}\right]$$

$$= (2\pi\sigma^2)^{-\frac{n}{2}} \exp\left[\frac{-1}{2\sigma^2}\left\{(n-1)s^2 + n(\bar{x} - \mu)^2\right\}\right] \quad \because \Sigma(x_i - \bar{x}) = 0$$

$$\Rightarrow f(x_1, x_2, \ldots, x_n / \mu, \sigma^2) = (2\pi\sigma^2)^{-\frac{n}{2}} \cdot \exp\left[\frac{-1}{2\sigma^2}\left\{(n-1)s^2 + n(\bar{x}-\mu)^2\right\}\right]$$

$$= g(\bar{x}, s^2, \mu, \sigma^2) \cdot h(x_1, x_2, \ldots, x_n)$$

where

$$g(\bar{x}, s^2, \mu, \sigma^2) = (2\pi\sigma^2)^{-\frac{n}{2}} \exp(\cdot) \quad \&$$

$$h(x_1, x_2, \ldots, x_n) = 1$$

In this case $(\bar{X}, S^2)$ is a sufficient statistic for (μ, σ^2). i.e. sufficient statistic $(\bar{X}, S^2)$ contains all the information about (μ, σ^2) that is available in the sample.

Example 4:- Let $X_1, X_2, \ldots, X_n$ be a random sample from a distribution with pdf $f(x/\theta) = \theta x^{\theta-1}$, $0<x<1$, $\theta>0$. Show that $\prod_{i=1}^{n} x_i$ is sufficient statistic for θ.

Solution The joint p.d.f is

$$f(x_1, x_2, \ldots, x_n / \theta) = \theta^n \left(\prod_{i=1}^{n} x_i\right)^{\theta-1}$$

$$= \theta^n \left(\prod^n x_i\right)^{\theta} \left(\frac{1}{\prod_{i=1}^{n}(x_i)}\right)$$

$$= g(t, \theta) \cdot h(x_1, x_2, \ldots x_n)$$

where $t = \prod_{i=1}^{n} x_i$

And $g(t,\theta) = \theta^n \left(\prod_{i=1}^{n} x_i\right)^\theta$ & $h(x_1,\dots,x_n) = \frac{1}{\prod_{i=1}^{n} x_i}$

Since $h(x_1, x_2, \dots, x_n)$ does not depend upon θ, the product $\prod_{i=1}^{n} X_i$ is a sufficient statistic for θ.

Assignment # 2

Q1:- Let $X_1, X_2, \dots, X_n$ be iid r.vs from Gamma distribution with α known and $\beta = \theta \in \Omega(0, \infty)$ unknown. Then show that UMVUE of θ is

$$T(X_1, X_2, \dots, X_n) = \frac{1}{n\alpha} \sum_{j}^{n} X_j = \bar{X}/\alpha$$

Q2:- Let $X_1, X_2, \dots, X_n$ iid r.vs from negative exponetial distribution with parameter $\theta \in \Omega(0,\infty)$ & investigate whether the CR bound is attained.

Q3:- For negative binomial distribution with parameter $\theta \in \Omega(0,1)$

Q4:- Let $X_1, X_2, \dots, X_n$ be iid $N(0,\theta)$, $0<\theta<\infty$ show that $T(X) = \sum_{i=1}^{n} X_i^2$ is a sufficient statistic for θ.

Q5:- Let $X_1, X_2, \dots, X_n$ be a random sample of size n from normal distribution with same mean & variance, say $\theta > 0$. Find a sufficient statistic for the parameter.

Q6:- Let $X_1, X_2, \ldots, X_n$ is a random sample of size n from $G(\alpha, \beta)$

(a) Find a sufficient statistic for α when the value of β is known.

(b) If $\alpha = 1$ find sufficient statistic for β.

Q7:- Let $X_1, X_2, \ldots, X_n$ be a random sample of size n from geometric distribution with pmf

$$f(x|\theta) = \theta(1-\theta)^x \; ; \; x = 0, 1, 2, \ldots \quad 0 < x < 1$$

Show that $\sum_{i=1}^{n} X_i$ is sufficient statistic for θ.

Solutions

Q1:- We know that

$$f(x/\theta) = \frac{1}{\Gamma\alpha \, \beta^\alpha} e^{-x/\beta} \cdot x^{\alpha-1} \quad , \text{given } \beta = \theta$$

and α is known

$$f(x/\theta) = \frac{1}{\Gamma\alpha \, \theta^\alpha} \cdot e^{-x/\theta} \, x^{\alpha-1}$$

$$\ln f = \ln 1 - \ln\Gamma\alpha - \alpha \ln\theta - \frac{x}{\theta} + (\alpha-1)\ln x$$

$$\frac{\partial \ln f}{\partial\theta} = -\frac{\alpha}{\theta} + \frac{x}{\theta^2}$$

$$\frac{\partial^2 \ln f}{\partial\theta^2} = \frac{\alpha}{\theta^2} - \frac{2x}{\theta^3}$$

$$I_1(\theta) = -E\left(\frac{\partial^2 \ln f}{\partial \theta^2}\right)$$

$$= -E\left(\frac{\alpha}{\theta^2} - \frac{2x}{\theta^3}\right)$$

$$= -E\left(\frac{\alpha}{\theta^2}\right) + \frac{2}{\theta^3} E(x)$$

$$= -\frac{\alpha}{\theta^2} + \frac{2}{\theta^3}(\alpha\theta)$$

∵ from Gamma Distribution
$E(X) = \alpha\beta$
$E(X) = \alpha\theta$

$$\Rightarrow I_1(\theta) = -\frac{\alpha}{\theta^2} + \frac{2\alpha}{\theta^2}$$

$$= \frac{\alpha}{\theta^2}$$

$$I_n(\theta) = nI_1(\theta) = \frac{n\alpha}{\theta^2}$$

$$\Rightarrow \boxed{I_n(\theta) = \frac{n\alpha}{\theta^2}}$$

C.R lower bound:-

$$Var_\theta(T) \geq \frac{[g'(\theta)]^2}{I_n(\theta)}$$

Here $g(\theta) = \theta$
$g'(\theta) = 1$

$$\Rightarrow Var_\theta(T) \geq \frac{1}{n\alpha/\theta^2} = 1 \times \frac{\theta^2}{n\alpha}$$

Now taking variance of $\bar{X}/\alpha$

$$V\left(\frac{\bar{X}}{\alpha}\right) = \frac{1}{\alpha^2} V(\bar{X}) = \frac{1}{\alpha^2} V\left[\frac{1}{n}\sum_{i=1}^{n} X_i\right]$$

$$= \frac{1}{\alpha^2 n^2}\sum_{i=1}^{n} V(X_i)$$

$$= \frac{1}{\alpha^2 n^2}\sum_{i=1}^{n} (\alpha\theta^2)$$

$$= \frac{1}{\alpha^2 n^2} \times n\alpha\theta^2 = \frac{\theta^2}{n\alpha}$$

$$\Rightarrow V\left(\frac{\bar{X}}{\alpha}\right) = \frac{\theta^2}{n\alpha}$$

Since $V\left(\frac{\bar{X}}{\alpha}\right)$ and C.R lower bound is same

$\Rightarrow$ UMVUE of θ is $\bar{X}/\alpha$.

Q2:- We know that

$$f(x/\theta) = 1/\theta \, e^{-x/\theta}$$

$$\ln f = \ln 1 - \ln \theta - x/\theta$$

$$\frac{\partial \ln f}{\partial \theta} = -\frac{1}{\theta} + \frac{x}{\theta^2}$$

$$\frac{\partial^2 \ln f}{\partial \theta^2} = 1/\theta^2 - \frac{2x}{\theta^3}$$

$$I_1(\theta) = -E\left(\frac{\partial^2 \ln f}{\partial \theta^2}\right) = -E\left(\frac{1}{\theta^2} - \frac{2x}{\theta^3}\right)$$

$$= -E\left(\frac{1}{\theta^2}\right) + \frac{2}{\theta^3} E(x)$$

For -ve Binom $E(x) = \theta$

$$= \frac{-1}{\theta^2} + \frac{2}{\theta^3} \cdot \theta$$

$$\Rightarrow I_1(\theta) = 1/\theta^2$$

So $I_n(\theta) = n I_1(\theta) = \frac{n}{\theta^2}$

C.R lower bound

$$Var_\theta(T) \geq \frac{[g'(\theta)]^2}{I_n(\theta)} = 1/\frac{n}{\theta^2}$$

$$\Rightarrow Var_\theta(T) = \frac{\theta^2}{n}$$

As $\bar{X}$ is unbiased estimator of θ so,

$$Var(\bar{X}) = V\left[\frac{1}{n}\sum_{i=1}^{n} X_i\right]$$

$$\Rightarrow Var(\bar{X}) = \frac{1}{n^2} \sum_{i=1}^{n} V(X_i)$$

$$= \frac{1}{n^2} \cdot n \cdot \theta^2$$

$$= \frac{\theta^2}{n}$$

As C.R lower bound and $V(\bar{x})$ is same so $\bar{X}$ is UMVUE.

Q4:-

We know that $X \sim N(\ , \sigma^2)$

$$f(x) = \frac{1}{\sqrt{2\pi}\sigma} e^{-\frac{1}{2}\left(\frac{x-\mu}{\sigma}\right)^2}$$

Given $X \sim N(0, \theta)$

$$\Rightarrow f(x) = \frac{1}{\sqrt{2\pi\theta}} e^{-\frac{x^2}{2\theta^2}}$$

The joint density function is given by

$$f(x_1, \dots, x_n/\theta) = \frac{1}{\sqrt{2\pi\theta}} e^{-x_1^2/2\theta^2} \cdots \cdots \frac{1}{\sqrt{2\pi\theta}} e^{-x_n^2/2\theta^2}$$

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

$$= g(T, \theta) \cdot h(x_1, \dots, x_n)$$

where $g(T,\theta) = \left(\frac{1}{\sqrt{2\pi\theta}}\right)^n e^{-\frac{1}{2\theta^2}\sum_{i=1}^{n} x_i^2}$ &

$h(x_1, x_2, \dots, x_n) = 1$

So $T = \sum_{i=1}^{n} x_i^2$ is sufficient statistic for θ.

$$\Rightarrow f(x_1,\cdots,x_n/\theta) = \left(\frac{1}{\sqrt{\alpha}\,\beta^{\alpha}}\right)^n e^{-\frac{1}{\beta}\sum_{i=1}^{n} x_i} \cdot e^{\alpha \sum_{i=1}^{n} \ln x_i} \cdot e^{-\sum_{i=1}^{n} \ln x_i}$$

$$= \left(\frac{1}{\sqrt{\alpha}\,\beta^{\alpha}}\right)^n e^{-\frac{1}{\beta}\sum_{i=1}^{n} x_i} \cdot e^{-\sum_{i=1}^{n} \ln x_i} \cdot e^{\alpha \sum_{i=1}^{n} \ln x_i}$$

$$= g(T,\alpha)\, h(x_1, x_2, \cdots, x_n)$$

Where

$$g(T,\alpha) = \left(\frac{1}{\sqrt{\alpha}\,\beta^{\alpha}}\right)^n e^{-\frac{1}{\beta}\sum_{i=1}^{n} x_i} \cdot e^{\alpha \sum_{i=1}^{n} \ln x_i}$$

& $h(x_1,\cdots,x_n) = e^{-\sum_{i=1}^{n} \ln x_i}$

With $\boxed{T = \sum_{i=1}^{n} \ln x_i}$

(b) $f(x) = \frac{1}{\beta} e^{-x/\beta}$

The joint p.d.f is

$$f(x_1,\cdots,x_n/\beta) = \left(\frac{1}{\beta}\right)^n e^{-1/\beta \sum_{i=1}^{n} x_i}$$

$$= g(T,\beta)\, h(x_1,\cdots,x_n)$$

Where $g(T,\beta) = \left(\frac{1}{\beta}\right)^n e^{-1/\beta \sum_{i=1}^{n} x_i}$

Here $T = \sum_{i=1}^{n} x_i$

MUHAMMAD TAHIR WATTOO

SP18-PMT-005

0344-8563284

⇒ Consistent Estimation:-

As sample size increases, the sample data should behave more like population data. So it is natural to suspect that the estimator of a parameter of a given population should approach the parameter as $n \to \infty$. In other words, a good estimator should be consistent – that it approaches the parameter for large values of n.

We first define "Convergence in probability".

Def:- Let $X_1, X_2, \ldots, X_n$ be a sequence of jointly distributed random variables on some probability space S. Also, X is another random variable on sample space S. We say that the sequence $X_1, X_2, \ldots, X_n$ converges to X in probability, if for every $\epsilon > 0$ we have

$$\lim_{n\to\infty} P[|X_n - X| > \epsilon] \to 0 \text{ as } n \to \infty$$

Symbolically $X_n \xrightarrow{P} X$.

Theorem:- Let $\hat{\theta}$ be an estimator of θ. If $V(\hat{\theta}) \to 0$ as $n \to \infty$, then $\hat{\theta} \xrightarrow{P} \theta$. In other words $\hat{\theta}$ is a consistent estimator of θ. OR $\hat{\theta}$ is a consistent estimator of θ if

$$\lim_{n\to\infty} E(\hat{\theta} - \theta)^2 = 0.$$

Example 1:- Let $X_1, X_2, \dots X_n$ be a random sample with finite mean μ and variance σ^2. Show that the sample mean $\bar{X}_n$ is a consistent estimator of μ.

Solution

$$V(\bar{X}_n) = \frac{\sigma^2}{n} \to 0 \text{ as } n \to \infty$$

Hence $\bar{X}_n \xrightarrow{P} \mu$, so it is a consistent estimator of μ.

According to the Chebyshev's Theorem

$$P[|\bar{X}_n - \mu| > \epsilon] \le \frac{V(\bar{X}_n)}{\epsilon^2}$$

we have

$$P[|\bar{X}_n - \mu| > \epsilon] \le \frac{\sigma^2}{n\epsilon^2}$$

$$\lim_{n\to\infty} P[|\bar{X}_n - \mu| > \epsilon] \to 0 \quad \text{as } n \to \infty$$

$$\Rightarrow \bar{X}_n \xrightarrow{P} \mu$$

Example 2:- Let $X_1, X_2, \dots X_n$ be a sequence of iid Random variables with common pdf

$$f(x) = e^{-x+\theta} \quad \text{if } x \ge \theta$$
$$= 0 \quad \text{if } x < \theta$$

write $\bar{X}_n = \frac{1}{n}\sum_{i=1}^{n} X_i$

(a) Show that $\bar{X}_n \xrightarrow{P} (1+\theta)$

(b) Show that $\min\{X_1, X_2, \dots, X_n\} \xrightarrow{P} \theta$

Solution

$$E(X) = \int_\theta^\infty x e^{-x+\theta} dx = \int_\theta^\infty x e^{-(x-\theta)} dx$$

put $x - \theta = t \Rightarrow x = t + \theta$

$\Rightarrow dx = dt$

$$\Rightarrow E(X) = \int_0^\infty (t+\theta) e^{-t} dt$$

$$= \int_0^\infty t e^{-t} dt + \theta \int_0^\infty e^{-t} dt$$

$$= \Gamma 2 + \theta(1) = (1+\theta)$$

$$E(X^2) = \int_\theta^\infty x^2 e^{-(x-\theta)} dx$$

$$= \int_0^\infty (t+\theta)^2 e^{-t} dt$$

$$= \int_0^\infty (t^2 + 2\theta t + \theta^2) e^{-t} dt$$

$$= \int_0^\infty t^{3-1} e^{-t} dt + 2\theta \int_0^\infty t^{2-1} e^{-t} dt + \theta^2 \int_0^\infty t^{1-1} e^{-t} dt$$

$$= \Gamma 3 + 2\theta \Gamma 2 + \theta^2 \Gamma 1$$

$$= 2 + 2\theta + \theta^2 = 1 + (1+\theta)^2$$

$$V(X) = (1+\theta)^2 + 1 - (1+\theta)^2 = 1$$

$$V(\bar{X}_n) = \frac{V(X)}{n} = \frac{1}{n}$$

According to chebyshev's Inequality

$$P[|\bar{X}_n - \mu| > \epsilon] \leq \frac{V(\bar{X}_n)}{\epsilon^2}$$

$$\Rightarrow P[|\bar{X}_n - \mu| > \epsilon] \leq \frac{1}{n\epsilon^2}$$

$$\lim_{n\to\infty} P[|\bar{X}_n - \mu| > \epsilon] \to 0 \quad \text{as } n \to \infty$$

or

$$\lim_{n\to\infty} P[|\bar{X}_n - (1+\theta)| > \epsilon] \to 0 \quad \text{as } n \to \infty$$

$\Rightarrow \bar{X}_n \xrightarrow{P} (1+\theta)$; In other words $\bar{X}_n$ is a consistent estimator of $(\theta+1)$

(b) $$E(X_{(1)}) = \int_0^\infty x_{(1)}\, g(x_{(1)})\, dx_{(1)}$$

$$g(x_{(1)}) = n[1 - F(x_{(1)})]^{n-1} f(x_{(1)})$$

As $f(x) = e^{-(x-\theta)} \Rightarrow f(x_{(1)}) = e^{-(x_{(1)}-\theta)}$

$$\Rightarrow F(x) = \int_0^x e^{-(x-\theta)}\, dx = e^{\theta} \int_0^x e^{-x}\, dx$$

$$= e^{\theta} \left| \frac{e^{-x}}{-1} \right|_0^x$$

$$F(x) = 1 - e^{-(x-\theta)}$$

$$\Rightarrow F(x_{(1)}) = 1 - e^{-((x_{(1)})-\theta)}$$

$$g(x_{(1)}) = n\left[e^{-(x_{(1)}-\theta)}\right]^{n-1} e^{-(x_{(1)}-\theta)}$$

$$= n e^{-n(x_{(1)}-\theta)} \qquad \theta \le x_{(1)} < \infty$$

$$E(X_{(1)}) = n\int_{\theta}^{\infty} x_{(1)} e^{-n(x_{(1)}-\theta)} \, dx_{(1)}$$

Put $n(x_{(1)}-\theta) = t \qquad x_{(1)} = \frac{t}{n} + \theta$

$$dx_{(1)} = \frac{1}{n} dt$$

$$\Rightarrow E(X_{(1)}) = \cancel{n}\int_0^{\infty}\left(\theta + \frac{t}{n}\right)e^{-t}\frac{dt}{\cancel{n}}$$

$$= \theta\int_0^{\infty} e^{-t} dt + \frac{1}{n}\int_0^{\infty} t e^{-t} dt$$

$$= \theta + \frac{1}{n}\Gamma 2 = \left(\theta + \frac{1}{n}\right)$$

$$E(X_{(1)}^2) = \cancel{n}\int_0^{\infty}\left(\theta + \frac{t}{n}\right)^2 e^{-t}\frac{dt}{\cancel{n}}$$

$$= \theta^2\int_0^{\infty} e^{-t} dt + \frac{1}{n^2}\int_0^{\infty} t^2 e^{-t} dt$$

$$+ \frac{2\theta}{n}\int_0^{\infty} t e^{-t} dt$$

$$= \theta^2 + \frac{2\theta}{n} + \frac{2}{n^2}$$

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

$$= 1/n^2$$

⇒ Order Statistics:-

Theorem:- Let $X_{(1)}, X_{(2)}, \dots, X_{(n)}$ denote the order statistics of a random sample $X_1, X_2, \dots, X_n$ from a continuous population with cdf $F_X(x)$ and pdf $f_X(x)$. Then the pdf of $X_{(j)}$ is

$$f_{X_{(j)}}(x) = \frac{n!}{(j-1)!\,(n-j)!} f_X(x) \left[F_X(x)\right]^{j-1} \left[1 - F_X(x)\right]^{n-j}$$

where $X_{(1)} = \min(X_1, X_2, \dots, X_n)$

$\vdots$

$X_{(n)} = \max(X_1, X_2, \dots, X_n)$

Exercise *

Question 1:- Let $X_1, X_2, \dots, X_n$ be a random sample from a Gamma (α, β) population

(a) Find the MLE of β, assumin α is known.

(b) If α and β are unknown, there is no explicit formula for the MLEs of α & β. Compute them umerically for data 20.9, 22.0, 23.9, 23.8, 25.0, 24.0, 21.7, 23.8, 22.8, 23.1, 23.5, 23.0, 23.0

Solution (a) The joint density function of

$x_1, x_2, \ldots, x_n$ is given by

$$L(\beta | x_1, x_2, \cdots, x_n) = \prod_{i=1}^{n} \frac{1}{\Gamma\alpha \, \beta^{\alpha}} x_i^{\alpha-1} e^{-x_i/\beta}$$

$$= \frac{1}{(\Gamma\alpha)^n \beta^{n\alpha}} \left(\prod_{i=1}^{n} x_i\right)^{\alpha-1} \times e^{-\sum_{i=1}^{n} x_i/\beta}$$

$$\log L(\beta / x_1, \cdots, x_n) = -n\log\Gamma\alpha - n\alpha\log\beta + (\alpha-1)\log\left(\prod_{i=1}^{n} x_i\right) - \frac{\sum x_i}{\beta}$$

$$\Rightarrow \frac{\partial \log L}{\partial \beta} = \frac{-n\alpha}{\beta} + \frac{\sum x_i}{\beta^2}$$

Setting $\left.\dfrac{\partial \log L}{\partial \beta}\right|_{\beta=\hat\beta} = 0 \Rightarrow \hat\beta = \dfrac{\sum X_i}{\alpha}$ is

The required MLE of β, when α is known.

(b) When α and β are unknowns, there is no analytic formula for the MLEs. One approach to finding $\hat\alpha$ & $\hat\beta$ is to numerically maximize the likelihood function. From part (a), for fixed value of α, the value of β that maximizes L is $\sum_{i=1}^{n} x_i / n\alpha$. Putting this into L we get

$$= \frac{1}{(\Gamma\alpha)^n \left(\sum x_i / n\alpha\right)^{n\alpha}} \left(\prod^{n} x_i\right)^{\alpha-1} e^{-\frac{\sum x_i}{\sum x_i / n\alpha}}$$

$$= \frac{1}{(\Gamma\alpha)^n \left(\sum_{1}^{n} x_i / n\alpha\right)^{n\alpha}} \left(\prod_{i=1}^{n}\right)^{\alpha-1} e^{-n\alpha}$$

Now the function can be maximized using computer program and for the above data $(n=14, \sum_{i=1}^{n} x_i = 323.6)$ we obtain

$$\hat{\alpha} = 514.219 \text{ \& hence } \hat{\beta} = \frac{323.6}{14 \times 514.219}$$

$$= 0.045$$

Question No 2:- Let $X_1, X_2, \cdots, X_n$ be a random sample from pdf

This is Pareto Dist with $\alpha = \theta$ & $\beta = 1$

$$f(x/\theta) = \theta x^{-2} \; ; \; 0 < \theta \leq x < \infty$$

a) What is sufficient statistic for θ?

b) Find the MLE of θ

c) Find the MM estimator of θ.

Solution

(a) The joint density function is given by

$$f(x_1, x_2, \cdots x_n/\theta) = \prod_{i=1}^{n} \theta x_i^{-2} I_{[\theta,\infty)}^{x_i}$$

$$= \theta^n \left(\prod_{i=1}^{n} x_i^{-2}\right) I_{[\theta,\infty)}(x_{(1)})$$

$$= \theta^n I_{[\theta,\infty)}(x_{(1)}) \times \left(\prod_{i=1}^{n} x_i^{-2}\right)$$

$t = x_{(1)} = \min(x_1, \cdots, x_n)$

$$= g(t,\theta) \cdot h(x_1, x_2, \cdots, x_n) \text{ where}$$

$$g(t,\theta) = \theta^n I_{[\theta,\infty)}(x_{(1)}) \quad \&$$

$$h(x_1, \cdots, x_n) = \prod_{i=1}^{n} x_i^{-2}$$

Scanned by CamScanner

$$f(x) = \frac{\theta}{x^2} \quad \theta < x < \infty$$

$$\Rightarrow f(x) = \frac{\theta}{x^2} \, I_{[\theta,\infty)}(x_i)$$

Thus $T = X_{(1)}$ is sufficient statistic for θ.

(b) $$L(\theta / x_1, x_2, \cdots, x_n) = \theta^n \left(\prod^n x_i^{-2}\right) I_{[\theta,\infty)}(x_{(1)})$$

To maximize $L(\theta / x_1, x_2, \cdots, x_n)$, we want to make θ as large as possible. Because the indicator function,

$L(\theta / x_1, x_2, \cdots, x_n) =$

$\theta > x_{(1)}$. Thus $\hat{\theta} = X_{(1)}$

(c) As $E(X) = \int_\theta^\infty \theta x^{-1} \, dx = \theta \log x \Big|_\theta^\infty = \infty$

Thus MM estimator of θ does not exist.

Question 3:- One observation, X_1 is taken from $N(0, \sigma^2)$ population.

a) Find an unbiased estimator of σ^2

b) Find the MLE of σ

c) Find MM estimator of σ

Solution (a) Given

$$f(x_1 | \sigma^2) = \frac{1}{\sqrt{2\pi}\,\sigma} e^{-\frac{1}{2}\frac{x_1^2}{\sigma^2}} ; \quad -\infty < x_1 < \infty, \ \sigma > 0$$

As

$$\sigma^2 = E(X_1^2) - (\mu)^2$$

$$\Rightarrow E(X_1^2) = \sigma^2 \quad \because \mu = 0$$

$\Rightarrow X_1^2$ is an unbiased estimator of σ^2

(b) $L(\sigma/x_1) = \frac{1}{\sqrt{2\pi}\,\sigma} e^{-x_1^2/2\sigma^2}$

$\log L = \log(2\pi)^{-\frac{1}{2}} - \log\sigma - \frac{x_1^2}{2\sigma^2}$

$= 0 - \frac{1}{\sigma} + \frac{x_1^2}{\sigma^3}$

$\left.\frac{\partial \log L}{\partial \sigma}\right|_{\sigma=\hat{\sigma}} = 0 \Rightarrow \hat{\sigma}^2 = X_1^2$

$\Rightarrow \hat{\sigma} = \sqrt{X_1^2} = |X_1|$

As $\frac{\partial^2 \log L}{\partial \sigma^2} = \frac{-3x^2\sigma^2}{\sigma^6} + \frac{1}{\sigma^2}$, which is $-$ve at $\hat{\sigma} = |X_1|$

Thus $\hat{\sigma} = |X_1|$ is a local maximum. Because it is the only place where the 1st derivative is zero, It is a global maximum.

(c) Because $E[X_1] = 0$ is known, just equate

$E(X_1^2) = \sigma^2 = \frac{1}{1}\sum_{i=1}^{1} X_i^2 = X_1^2$

$\Rightarrow \hat{\sigma} = \sqrt{X_1^2} = |X_1|$

Question 4:- Let $X_1, X_2, \ldots, X_n$ be i.i.d with pdf

$f(x/\theta) = \frac{1}{\theta} \; ; \; 0 \leq x \leq \theta \; , \; \theta > 0$

estimate θ using both the method

of moments & MLE. Calculate Means and variance of the two estimators. Which one should be preffered & why?

Solution

MM $\quad E(X) = \frac{1}{\theta}\int_0^\theta x\,dx = \frac{1}{\theta}\left|\frac{x^2}{2}\right|_0^\theta$

$$= \frac{\theta}{2}$$

Now equating population moment with sample moment i.e

$$\frac{1}{n}\sum_1^n x_i = \frac{\theta}{2} \Rightarrow \tilde{\theta} = 2\bar{X}$$

Now for mean & variance of $\tilde{\theta}$ is given by

$$E(\tilde{\theta}) = E(2\bar{X}) = 2E\left[\frac{1}{n}\sum_1^n X_i\right]$$

$$= 2\frac{\cancel{n}E(X)}{\cancel{n}} = 2\times\frac{\theta}{2} = \theta$$

$$\& \; V(\tilde{\theta}) = 4V(\bar{X}) = 4\cdot\frac{\theta^2/12}{n}$$

$$= \frac{\theta^2}{3n}$$

as $X \sim U(0,\theta)$

$V(X) = E(x^2) - [E(X)]^2$

$E(x^2) = \frac{1}{\theta}\int_0^\theta x^2\,dx$

$= \frac{1}{\theta}\left|\frac{x^3}{3}\right|_0^\theta$

$= \frac{1}{3\theta}(\theta^3) = \frac{\theta^2}{3}$

$V(X) = \frac{\theta^2}{3} - \left(\frac{\theta}{2}\right)^2$

$= \frac{4\theta^2 - 3\theta^2}{12}$

$= \frac{\theta^2}{12}$

Now for MLE

$$L(\theta/x_1, x_2, \cdots x_n) = \prod_{i=1}^n \frac{1}{\theta} I_{[0,\theta]}(x_i)$$

$$= \frac{1}{\theta^n} I_{[0,\theta]}(x_{(n)})$$

For $\theta \geq x_{(n)}$, $L = \frac{1}{\theta^n}$, a decreasing function. So for $\theta \geq x_{(n)}$, L is maximized at $\hat{\theta} = x_{(n)}$. $L = 0$ for $\theta < x_{(n)}$. So the MLE is $\hat{\theta} = X_{(n)}$

The pdf of $\hat{\theta} = X_{(n)}$ is

$$f(x_{(n)}) = \frac{n}{\theta^n}(x_{(n)})^{n-1}, \quad 0 \leq x_{(n)} < \theta$$

Now

$$E(X_{(n)}) = \frac{n\theta}{n+1}$$

$$V(\hat{\theta}) = \frac{n\theta^2}{(n+2)(n+1)^2}$$

Now $\tilde{\theta}$ is an unbiased estimator of θ; $\hat{\theta}$ is a biased estimator of θ. If n is large, This bias is not large because $\frac{n}{n+1}$ is close to 1. on the other hand

$V(\hat{\theta}) < V(\tilde{\theta})$ $\forall$ θ, so if n is large, $\hat{\theta}$ is preferable.

Question 5:- The independent random variable $X_1, \dots, X_n$ have the common distribution

$$P(X_i \leq x \mid \alpha, \beta) = \begin{cases} 0 & \text{if } x < 0 \\ (x/\beta)^\alpha & \text{if } 0 \leq x \leq \beta \\ 1 & \text{if } x > \beta \end{cases} \quad \alpha, \beta > 0$$

a) Find a two dimensional sufficient statistic for (α, β)

b) Find the MLEs of α & β

c) The length (in milimeters) of cuckoo's eggs found in hedge sparrow nests can be modeled with this distribution for the data

22.0, 23.9, 23.8, 20.9, 25.0, 24.0, 21.7, 23.8, 22.8, 23.1, 23.1, 23.5, 23.0, 23.0 find the MLEs of α, β

Solution The pdf is given by

$$f(x|\alpha,\beta) = \frac{dP}{dx}(X_i \le x) = \alpha \frac{x_i^{\alpha-1}}{\beta^\alpha}, \quad 0 \le x_i \le \beta$$

(a) The joint density function is given by

$$f(x_1, x_2, \ldots, x_n|\alpha,\beta) = \prod_{i=1}^{n} \frac{\alpha}{\beta^\alpha}(x_i)^{\alpha-1} I_{[0,\beta]}(x_i)$$

$$= \left(\frac{\alpha}{\beta^\alpha}\right)^n \left(\prod_{i=1}^{n} x_i\right)^{\alpha-1} \cdot I_{(-\infty,\beta]}(x_{(n)})$$

$$= \left(\frac{\alpha}{\beta^\alpha}\right)^n \left(\prod_{i=1}^{n} x_i\right)^{\alpha} \cdot I_{(-\infty,\beta]}(x_{(n)}) \cdot \frac{1}{\prod_{i=1}^{n} x_i}$$

$$= g\left(\prod_{i=1}^{n} x_i, x_{(n)}; \alpha, \beta\right) * h(x_1, x_2, \ldots, x_n)$$

By Factorization theorem $\left(\prod_{i=1}^{n} X_i, X_{(n)}\right)$ are sufficient statistic for (α, β)

(b) For any fixed α,
$L(\alpha, \beta | x_1, \dots, x_n) = 0$ if $\beta < x_{(n)}$, and

$L(\alpha, \beta | x_1, \dots, x_n)$ is a decreasing function of β if $\beta \geq x_{(n)}$. Thus $X_{(n)}$ is MLE of β. for MLE of α calculate

$$\frac{\partial \log L}{\partial \alpha} = \frac{\partial}{\partial \alpha}\left[n \log \alpha - n\alpha \log \beta + (\alpha - 1) \log \prod x_i\right]$$

$$= \frac{n}{\alpha} - n \log \beta + \log \prod_{i=1}^{n} x_i$$

Also $$\frac{\partial^2 \log L}{\partial \alpha^2} = -\frac{n}{\alpha^2} < 0$$

So the derivative equal to zero & use $\hat{\beta} = X_{(n)}$ to obtain

$$\hat{\alpha} = \frac{n}{n \log X_{(n)} - \log \prod_{i=1}^{n} X_i}$$

(c) As $X_{(n)} = 25.0$, $\log \prod_i X_i = \sum \log X_i$

$$= 43.95$$

$$\Rightarrow \hat{\beta} = 25.0, \ \hat{\alpha} = 12.59$$

Question 6:- Let $X_1, X_2, \ldots, X_n$ be i.i.d with p.d.f

$$f(x/\theta) = \theta x^{\theta-1}, \quad 0 \le x \le 1, \ \theta > 0$$

(a) Find the MLE of θ and show that $V(\hat{\theta}) \to 0$ as $n \to \infty$

(b) Find the MM estimator of θ.

Solution The joint density function is given by

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

$$= \theta^n \left(\prod_{i=1}^{n} x_i\right)^{\theta-1}$$

$$\log L = n \log\theta + (\theta-1)\log\prod_{i=1}^{n} x_i$$

$$\frac{\partial \log L}{\partial\theta} = \frac{n}{\theta} + \log\prod^{n}(x_i) = \frac{n}{\theta} + \sum_{i=1}^{n}\log x_i$$

$$\left.\frac{\partial \log L}{\partial\theta}\right| = 0 \Rightarrow \hat{\theta} = \left(-\frac{1}{n}\sum_{i=1}^{n}\log x_i\right)^{-1}$$

$$\frac{\partial^2 \log L}{\partial\theta^2} < 0$$

Now for $V(\hat{\theta}) = ?$ we

As $-\log X_i \sim$ exponential $(1/\theta)$ and

$-\sum \log X_i \sim$ Gamma $(n, 1/\theta)$

$$\Rightarrow \hat{\theta} = \frac{n}{T} \quad \& \quad T \sim \text{Gamma}(n, 1/\theta)$$

Now $$E\left(\frac{1}{T}\right) = \frac{\theta^n}{\Gamma n}\int_0^\infty \frac{1}{t} t^{n-1} e^{-\theta t}\, dt$$

$$= \frac{\theta}{n-1}$$

$$E\left(\frac{1}{T^2}\right) = \frac{\theta^n}{\Gamma n}\int_0^\infty \frac{1}{t^2} t^{n-1} e^{-\theta t}\,dt$$

$$= \frac{\theta^2}{(n-1)(n-2)}$$

$$E(\hat{\theta}) = E\left(\frac{n}{T}\right) = n \cdot E\left(\frac{1}{T}\right) = \frac{n\theta}{(n-1)}$$

$$V(\hat{\theta}) = E(\hat{\theta})^2 - [E(\hat{\theta})]^2$$

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

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

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

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

$$\Rightarrow V(\hat{\theta}) = \frac{n\theta^2}{(n-1)^2(n-2)} \to 0 \quad \text{as } n \to \infty$$

$\Rightarrow \hat{\theta} \xrightarrow{P} \theta$ i.e $\hat{\theta}$ is an consistent estimator of θ

(b) $X \sim \text{beta}(\theta, 1)$

$E(X) = \dfrac{\theta}{\theta+1}$ and after equating

$$\frac{1}{n}\sum_1^n X_i = \frac{\theta}{\theta+1}$$

$$\Rightarrow \hat{\theta} = \frac{\sum_1^n X_i}{n - \sum_1^n X_i} = \frac{\bar{X}}{1-\bar{X}}$$

Question 7:- Let $X_1, X_2, \ldots, X_n$ be a random sample from a population with pmf

$$P_\theta(X=x) = \theta^x(1-\theta)^{1-x} \quad ; \; x=0,1$$
$$0 \leq \theta \leq \frac{1}{2}$$

(a) Find the MM estimator and MLE of θ

(b) Find the mean square errors of each of the estimator

(c) Which estimeter is prefffered? justify your choice.

Solution

(a) $E(X)=\theta$ & $\theta = \frac{1}{n}\sum^n X_i \Rightarrow \tilde{\theta} = \bar{X}$

(b) $L(\theta/x_1, \ldots, x_n) = \theta^{\sum x_i}(1-\theta)^{n-\sum x_i}$

$$= \theta^{n\bar{x}}(1-\theta)^{n(1-\bar{x})}$$

Remember that $0 \leq x \leq \frac{1}{2}$, therefore, when $\bar{X} \leq \frac{1}{2}$, $\bar{X}$ is the MLE of θ, therefore $\bar{X}$ is the overall maximum of $L(\theta/x_1, x_2, \ldots, x_n)$. When $\bar{X} > 1/2$, $L(\theta/x_1, x_2, \ldots, x_n)$ is an increasing function of θ on $[0, 1/2]$ and obtains its maximum at the upper bound of θ which is $1/2$. So the MLE is

$$\hat{\theta} = \min\{\bar{X}, 1/2\}$$

(C) The MSE of $\tilde{\theta}$ is

$$MSE(\tilde{\theta}) = V(\tilde{\theta}) + (bias(\tilde{\theta}))^2 \quad \because bias\, \tilde{\theta} = 0$$

$$= V(\tilde{\theta}) + \frac{\theta(1-\theta)}{n}$$

For the $MSE(\hat{\theta})$, we have

$V(\tilde{\theta}) = \frac{n}{n^2} V(X_i) - I$

$V(X_i) = E(X_i^2) - [E(X)]$

$E(X) = \sum_{i=0}^{1} x_i P(x_i)$

$= 0 P(X) + 1 \cdot \theta$

$= \theta$

$E(X^2) = \theta$

$V(X) = \theta - \theta^2$

$= \theta(1-\theta)$

Putting in ①

$V(\tilde{\theta}) = \frac{\theta(1-\theta)}{n}$

$\therefore \hat{\theta} = \min\{\bar{X}, 1/2\}$

$$MSE(\hat{\theta}) = E(\hat{\theta} - \theta)^2$$

$$= \sum_{y=0}^{n} (\hat{\theta} - \theta)^2 \binom{n}{y} \theta^y (1-\theta)^{n-y}$$

Corollary:- Let $X_1, X_2, \dots, X_n$ be i.i.d of $f(x/\theta)$, where $f(x/\theta)$ satisfies the conditions of the Cramer-Rao theorem. Let $L(\theta/x) = \prod_{i=1}^{n} f(x_i/\theta)$ denote the likelihood function. If $T(X) = t(X_1, \dots, X_n)$ is any unbiased estimator of $g(\theta)$, then $T(X)$ attains the Cramer-Rao lower bound iff

$$\frac{\partial}{\partial \theta} \log L(\theta/x) = a(\theta)[t(X) - g(\theta)]$$ for some function $a(\theta)$.

Example:- Let $X_1, X_2, \dots, X_n$ be i.i.d $N(\mu, \sigma^2)$, consider estimation of σ^2, where μ is known.

Solution

$$f(x|\mu, \sigma^2) = \frac{1}{\sqrt{2\pi\sigma^2}} e^{-\frac{1}{2\sigma^2}(x_i - \mu)^2}$$

$$L(\mu, \sigma^2 / x_1, \dots, x_n) = \frac{1}{(2\pi\sigma^2)^{n/2}} \cdot e^{-\frac{1}{2\sigma^2}\sum_{1}^{n}(x_i - \mu)^2}$$

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

$$\frac{\partial \log L}{\partial \sigma^2} = \frac{-n}{2\sigma^2} + \frac{1}{2\sigma^4}\sum_{1}^{n}(X_i - \mu)^2$$

$$\Rightarrow \frac{\partial \log L}{\partial \sigma^2} = \frac{n}{2\sigma^4}\left(\sum_1^n \frac{(x_i-\mu)^2}{n} - \sigma^2\right)$$

Thus taking $a(\sigma^2) = \frac{n}{2\sigma^4}$ shows that the best unbiased estimator of σ^2 is $\frac{1}{n}\sum_1^n (X_i-\mu)^2$, which is calculable only if μ is known. If μ is unknown, the bound can not be attained.

Question 8:- For each of the following distributions let $X_1, \ldots, X_n$ be random sample. Is there a function of θ, say $g(\theta)$, for which there exist an unbiased estimator whose variance attains the Cramer-Rao lower bound if so, find it, show why not

(a) $f(x|\theta) = \theta x^{\theta-1}$; $0<x<1$, $\theta>0$ i.e $X \sim beta(\theta,1)$

(b) $f(x/\theta) = \frac{\log\theta}{\theta-1}\theta^x$; $0<x<1$, $\theta>1$

Solution

(a) $$\frac{\partial}{\partial\theta}\log L(\theta/x_1,\ldots,x_n) = \frac{\partial}{\partial\theta}\log\prod_{i=1}^n \theta x_i^{\theta-1}$$

$$= \frac{\partial}{\partial\theta}\sum_i\left[\log\theta + (\theta-1)\log x_i\right]$$

$$= \sum_i^n\left[\frac{1}{\theta}+\log x_i\right] = \frac{n}{\theta}+\sum\log x_i$$

$$= -n\left[\frac{-\sum^n \log x_i}{n} - \frac{1}{\theta}\right]$$

Thus $-\sum^{n} \log X_i/n$ is the UMVUE of $1/\theta$ and attains the Cramer-Rao bound.

(b) $$\frac{\partial}{\partial\theta} \log(\theta/x_1,\ldots,x_n) = \frac{\partial}{\partial\theta} \log \prod_{i=1}^{n} \frac{\log\theta}{\theta-1}\theta^{x_i}$$

$$= \frac{\partial}{\partial\theta} \sum_i \left[\log\log\theta - \log(\theta-1) + x_i\log\theta\right]$$

$$= \frac{n}{\theta\log\theta} - \frac{n}{(\theta-1)} + \frac{n\bar{x}}{\theta}$$

$$= \frac{n}{\theta}\left[\bar{x} - \left(\frac{\theta}{\theta-1} - \frac{1}{\log\theta}\right)\right]$$

Thus $\bar{X}$ is the UMVUE of $\left(\frac{\theta}{\theta-1} - \frac{1}{\log\theta}\right)$ and attains Cramer-Rao Power bound.

Question 9:- Let $X_1,\ldots,X_n$ be i.i.d Bernoulli(θ). Show that the variance of $\bar{X}$ attains Cramer-Rao Power bound & hence $\bar{X}$ is UMVUE of θ.

$\because \log\prod_1^n X_i = \sum_1^n \log X_i$

Solution

$$\frac{\partial}{\partial\theta}\log L(\theta/x_1,\ldots,x_n) = \frac{\partial}{\partial\theta}\log\prod_i \theta^{x_i}(1-\theta)^{1-x_i}$$

$$= \frac{\partial}{\partial\theta}\left[\sum_i^n \left(x_i\log\theta + (1-x_i)\log(1-\theta)\right)\right]$$

$$\Rightarrow \frac{\partial}{\partial\theta}\log L(\theta/x) = \sum_i^n \left[\frac{x_i}{\theta} - \frac{(1-x_i)}{1-\theta}\right]$$

$$\Rightarrow \frac{\partial}{\partial \theta} \log L(\theta/x) = \left(\frac{n\bar{x}}{\theta} - \frac{n - n\bar{x}}{(1-\theta)}\right)$$

$$= \frac{n}{\theta(1-\theta)}\left[\bar{x} - \theta\right]$$

Thus, $\bar{X}$ is UMVUE of θ & attains the cramer-Rao lower bound.

Question 10:- Let $X_1, \ldots, X_n$ be i.i.d $N(\theta, 1)$. Show that $\bar{X}$ is the best unbiased estimator of θ.

Solution

$$\frac{\partial}{\partial \theta} \log L(\theta/x_1, x_2, \ldots, x_n) = \frac{\partial}{\partial \theta} \log \prod_i^n \frac{1}{\sqrt{2\pi}} e^{-\frac{1}{2}(x_i-\theta)^2}$$

$$= \frac{\partial}{\partial \theta}\left[-\frac{n}{2}\log 2\pi - \frac{1}{2}\sum_{i=1}^n (x_i - \theta)^2\right]$$

$$= 0 + \frac{2}{2}\sum_{i=1}^n (x_i - \theta) = \sum x_i - n\theta$$

$$= n\bar{x} - n\theta = n(\bar{x} - \theta)$$

This $\bar{X}$ is UMVUE of θ and attains Cramer-Rao lower bound.

MUHAMMAD TAHIR WATTOO

PhD MATHEMATICS CU ISBD

0344-8563284

*Hypothesis Testing:-

Def:- A hypothesis is a statement about a population.

The goal of hypothesis test is to decide based on a sample from the population, which of two complementary hypothesis is true. The two complementary hypothesis in a hypothesis testing problem are called the null hypothesis and the alternative hypothesis respectively. They are denoted by H_0 & H_1.

In a Hypothesis testing problem, after observing the sample the experimenter must decide either to reject H_0 as false and decide H_1 is true. The subset of the sample space for which H_0 will be rejected is called the rejection region or critical region. The compliment of the rejection region is called the acceptance region

***Region:-** Typically a Hypothesis testing is specified in terms of a test statistic

$W(X_1, \dots, X_n) = W(X)$, a function of the sample.

$H_0 : \mu = 60$ Null Hypothesis	$H_0 : p = 0.01$	$H_0 : \sigma^2 = 4$
$H_1 : \mu \neq 60$ alternative Hypothesis	$H_1 : p \neq 0.01$	$H_1 : \sigma^2 \neq 4$

If $H_0 : \mu \geq 60$
& $H_1 : \mu < 60$

Equality sign always appear in null Hypothesis.

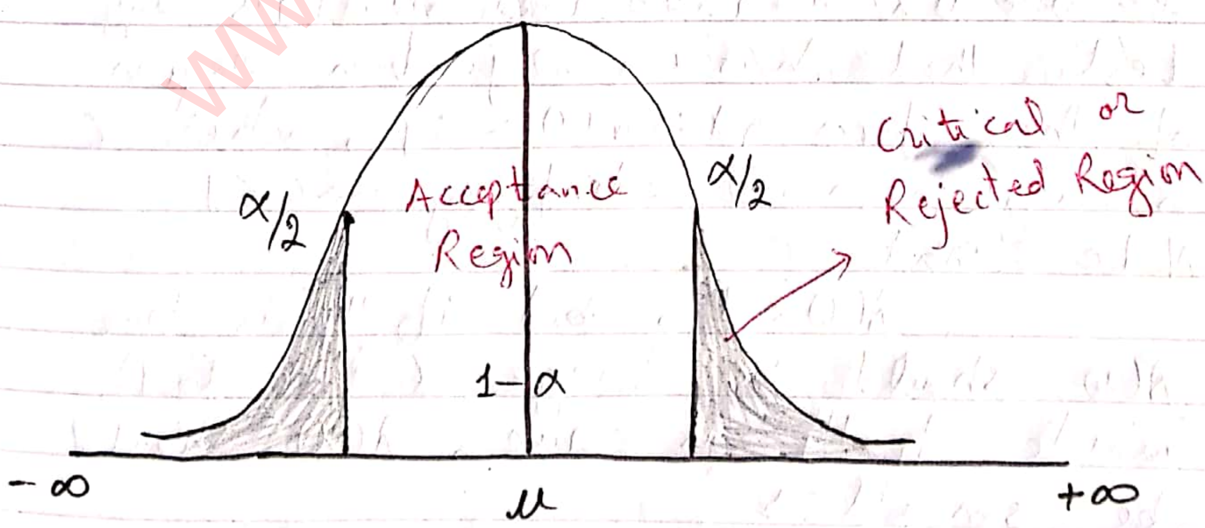
* Type I & Type II Errors:-

$\alpha = P(\text{Type I error}) = P(\text{Reject } H_0 / H_0 \text{ is true})$

&

$\beta = P(\text{Type II error}) = P(\text{Acceptance } H_0 / H_1 \text{ is true})$

$$f(x) = \frac{1}{\sqrt{2\pi}\,\sigma} e^{-\frac{1}{2}\left(\frac{x-\mu}{\sigma}\right)^2} \; ; \qquad H_0 : \mu = 60$$
$$H_1 : \mu \neq 60$$



* Likelihood Ratio Test :-

If $X_1, \ldots, X_n$ is a random sample from a population with pmf or pdf $f(x/\theta)$ (θ may be a vector). The likelihood function is defined as

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

Let Θ (Parametri space) denote the entire parameter space. Likelihood ratio tests defined as :

Definition:- The likelihood ratio test statistic for testing $H_0 : \theta \in \Theta$ Vs $H_1 : \theta \in \Theta$ is

$$\lambda(X) = \frac{\sup_{\Theta_0} L(\theta/x)}{\sup_{\Theta} L(\theta/x)} = \frac{\max_{H_0} L}{\max_{H_0 \cup H_1} L}$$

A likelihood ratio test (LRT) is any test that has a rejection region of the form $\{X : \lambda(x) \leq c\}$, where c is any number satisfying $0 \leq c \leq 1$

Note that

$\lambda(x) \leq 1$, but if H_0 is true $\lambda(x)$ should be large (close to 1), while if H_1 is true, $\lambda(x)$ should be smaller

For a specified significance level α this leads to the

decision rule i.e
Reject H_0 in favour of H_1 if $\lambda(x) \leq c$, where c is such that

$$\alpha = P[\lambda(x) \leq c]$$

This test is called likelihood ratio test of size α. Usually we take $\alpha = 5\% = 0.05$

Example:- Let $X_1, X_2, \dots, X_n$ be r.s from the exponential distribution with parameter λ.

(a) Derive the likelihood ratio test for $H_0 : \lambda = \lambda_0$ Vs $H_1 : \lambda \neq \lambda_0$

(b) Derive the critical region of this test.

Solution

(a) The likelihood function is given by

$$L(\lambda / x_1, \dots, x_n) = \prod_{i=1}^{n} \frac{1}{\lambda} e^{-x_i/\lambda}$$

$$= \frac{1}{\lambda^n} e^{-\sum x_i/\lambda}$$

$$l = \text{Log } L = -n \ln \lambda - \frac{\sum x_i}{\lambda}$$

$$\frac{\partial l}{\partial \lambda} = \frac{-n}{\lambda} + \frac{\sum x_i}{\lambda^2}$$

Put $\frac{\partial l}{\partial \lambda} = 0$

$$\Rightarrow \frac{\sum x_i}{\lambda^2} = \frac{n}{\lambda} \quad \Rightarrow \frac{\sum x_i}{n} = \lambda$$

$$\Rightarrow \quad \hat{\lambda} = \bar{X}$$

The likelihood ratio test statistic is

$$\frac{L(\lambda_0)}{L(\hat{\lambda})} = \lambda(x) = \frac{\max\limits_{H_0} L(\lambda|x_1,\dots,x_n)}{\max\limits_{H_0 \cup H_1} L(\lambda|x_1,\dots,x_n)}$$

$$= \frac{1/\lambda_0^n \cdot e^{-\frac{\sum x_i}{\lambda_0}}}{\frac{1}{(\bar{x})^n} \cdot e^{-\sum x_i/\bar{x}}}$$

$$= \frac{\frac{1}{\lambda_0^n} e^{-\frac{n\bar{x}}{\lambda_0}}}{\frac{1}{(\bar{x})^n} e^{-n}}$$

$$= \left[\frac{\bar{X}}{\lambda_0} e^{-\bar{x}/\lambda_0}\right]^n e^n$$

(b) The critical region for LRT is i-e $\lambda(x) \leq c$

$$\left[\left(\frac{\bar{X}}{\lambda_0}\right) e^{-\frac{\bar{x}}{\lambda_0}}\right]^n e^n \leq c$$

$$\Rightarrow \left\{\left(\frac{\bar{X}}{\lambda_0}\right) e^{-\bar{x}/\lambda_0}\right\}^n \leq c e^{-n}$$

$$\Rightarrow \frac{\bar{x}}{\lambda_0} e^{-\bar{x}/\lambda_0} \leq \left[e^{-n} c\right]^{1/n}$$

$$\Rightarrow \bar{x}\, e^{-\bar{x}/\lambda_0} \le \lambda_0 [e^{-n} c]^{1/n} = k$$

The constant k is selected s.t the size of the critical region is α.

Example:- (Likelihood Ratio Test for mean of a Normal Population)

Let $X_1, X_2, \dots, X_n$ is a random sample from $N(\theta, \sigma^2)$ distribution. where $-\infty < x < \infty$, $\sigma^2 > 0$ is known. Test the Hypothesis $H_0: \theta = \theta_0$ Vs $H_1: \theta \neq \theta_0$ of size α.

Solution The likelihood function is

$$L(\theta, \sigma^2 / x_1, \dots, x_n) = \prod_{i=1}^{n} \frac{1}{\sqrt{2\pi\sigma^2}} e^{-\frac{1}{2}\left(\frac{x_i - \theta}{\sigma}\right)^2}$$

$$= \left(\frac{1}{2\pi\sigma^2}\right)^{\frac{n}{2}} e^{-\frac{1}{2\sigma^2}\sum_1^n (x_i - \theta)^2}$$

$$\because \sum_1^n (x_i - \theta)^2 = \sum_1^n (x_i - \bar{x} + \bar{x} - \theta)^2$$

$$= \sum_1^n [(x_i - \bar{x})^2 + (\bar{x} - \theta)^2 + 2(x_i - \bar{x})(\bar{x} - \theta)]$$

$$= \sum_1^n (x_i - \bar{x})^2 + n(\bar{x} - \theta)^2 + 2(\bar{x} - \theta)\sum_1^n (x_i - \bar{x})$$

$$\because \sum_1^n (x_i - \bar{x}) = 0$$

$$\Rightarrow \sum_1^n (x_i - \theta)^2 = \sum_1^n (x_i - \bar{x})^2 + n(\bar{x} - \theta)^2$$

So

$$L = \left(\frac{1}{2\pi\sigma^2}\right)^{\frac{n}{2}} \exp\left[-(2\sigma^2)^{-1}\sum_1^n (x_i - \bar{x})^2\right] \exp\left[-(2\sigma^2)^{-1} n(\bar{x} - \theta)^2\right]$$

$$\therefore \hat{\theta} = \bar{X}$$

As

$$\lambda(x) = \frac{L(\theta_0)}{L(\hat{\theta})}$$

After simplification

$$= \exp\left[-(2\sigma^2)^{-1} n(\bar{x} - \theta_0)^2\right]$$

Then $\lambda(x) \leq c$ is equivalent to

$$\exp\left[-(2\sigma^2)^{-1} n(\bar{x} - \theta_0)^2\right] \leq c$$

taking ln on both sides

$$\frac{-n(\bar{x} - \theta_0)^2}{\sigma^2} \leq 2 \ln c$$

$$\left(\frac{\bar{x} - \theta_0}{\sigma/\sqrt{n}}\right)^2 \geq -2 \ln c = c'$$

Let $z = \left(\frac{\bar{x} - \theta}{\sigma/\sqrt{n}}\right)$

Then descision rule is equivalent to Reject H_0 in favour of H_1 if

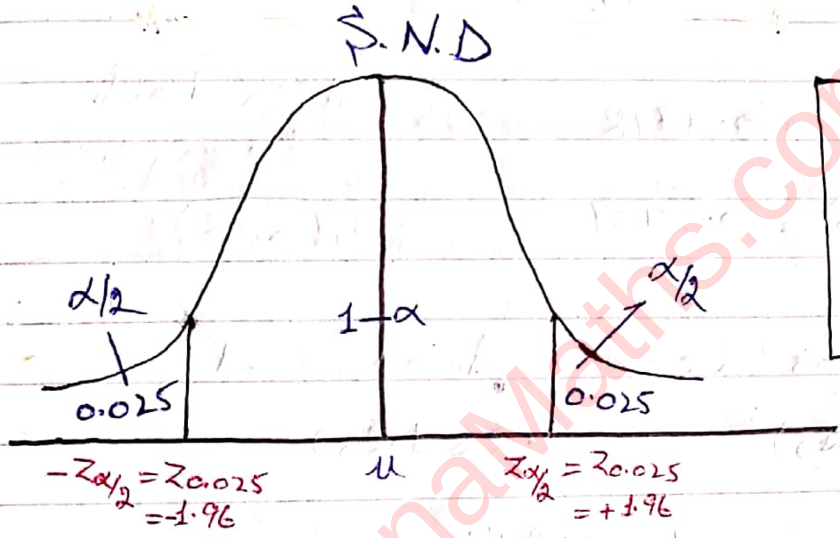
$$|z| \geq c^*$$

where $\alpha = P[|z| \geq c^*]$

If we take c^* to $Z_{\alpha/2}$

i.e the upper $\alpha/2$ critical value of a standard Normal dist. then our test will have exact level α i.e

$$\alpha = P\left[|Z| \geq Z_{\alpha/2}\right]$$



$1-\alpha = 95\%$
$1-\alpha = 0.95$
$\alpha = 0.05$
$\frac{\alpha}{2} = 0.025$

S.N.V $\quad Z = \frac{\bar{X}-\mu}{\sigma}$

$$Z = \frac{\bar{X}-\mu}{\sigma/\sqrt{n}} \Rightarrow$$

Note:- If $X \sim N(\mu, \sigma^2)$

$$f(x) = \frac{1}{\sqrt{2\pi}\,\sigma} e^{-\frac{1}{2}\left(\frac{x-\mu}{\sigma}\right)^2} \quad ; \quad -\infty < x < \infty$$

$$Z \sim N(0,1) \quad , \quad Z = \frac{X-\mu}{\sigma}$$

$$f(z) = \frac{1}{\sqrt{2\pi}} e^{-\frac{1}{2}z^2} \quad , \quad -3.4 < Z < 3.4$$

$$f(z \leq Z) = \frac{1}{\sqrt{2\pi}} \int_{-\infty}^{Z} e^{-\frac{1}{2}z^2}\, dz$$

$$P(Z<-1.96)=\frac{1}{\sqrt{2\pi}}\int_{-\infty}^{-1.96} e^{-\frac{1}{2}z^2}\,dz$$

$$=0.025$$

If mean $=\mu$, $\sigma=2$
$P(Z<40)=\frac{1}{\sqrt{2\pi}\cdot 2}\int_{-\infty}^{40} e^{\ldots}$

$$P(Z<1.96)=\frac{1}{\sqrt{2\pi}}\int_{-\infty}^{1.96} e^{-\frac{1}{2}z^2}\,dz$$

$$=0.975$$

$P(Z<2.0)=0.9772$

$P(Z<1.23)=0.8907$

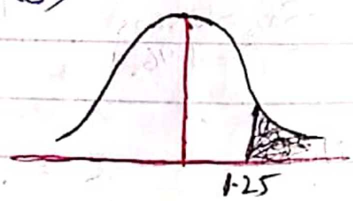
Area table always gives $P(Z\leq z)$

If it has been asked that

$P(Z>1.25)=1-P(Z\leq 1.25)$

$=1-0.8944$

$=0.1056$



$P(Z<?)=0.05$

$P(Z<-1.65)=0.05 \Rightarrow Z=-1.65$

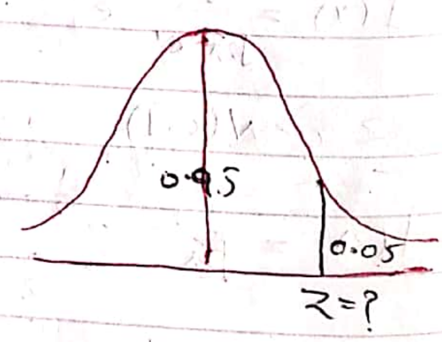
$P(Z>?)=0.05$

$P(Z>1.65)=0.05$

Alternate

$P(Z>?)=0.05$

$1-P(Z\leq ?)=0.05$



$1-0.05=P(Z\leq ?)$

$0.95=P(Z\leq ?) \Rightarrow Z=1.65$

Question:- (LRT for mean of Normal population, when σ is unknown)

Let $X_1, \ldots, X_n$ be a random sample from $N(\mu, \sigma^2)$, where σ is unknown. Show that the LRT of $H_0: \mu = \mu_0$ Vs $H_1: \mu \neq \mu_0$ can be based upon the statistic

$$T = \frac{(\bar{X} - \mu_0)\sqrt{n}}{\sqrt{\frac{\sum (X_i - \bar{X})^2}{(n-1)}}}$$

Determine the null distribution of T & give explicitly the rejection rule for a level α-test.

Solution The likelihood function is given by

$$L(\mu, \sigma^2) = \left(\frac{1}{2\pi\sigma^2}\right)^{\frac{n}{2}} \cdot e^{-\frac{1}{2\sigma^2}\sum_1^n (x_i - \mu)^2}$$

The MLEs of μ and σ^2 are

$$\hat{\mu} = \bar{X} \quad \& \quad \hat{\sigma}^2 = \frac{1}{n}\sum_1^n (X_i - \bar{X})^2$$

$$L(\hat{\mu}, \hat{\sigma}^2) = \left(\frac{1}{2\pi\hat{\sigma}^2}\right)^{\frac{n}{2}} \cdot e^{-\frac{1}{2\hat{\sigma}^2}\sum_1^n (X_i - \hat{\mu})^2}$$

$$= \left(\frac{1}{2\pi}\right)^{\frac{n}{2}} \cdot \left(\frac{1}{\hat{\sigma}^2}\right)^{\frac{n}{2}} \cdot e^{-\frac{1}{2\hat{\sigma}^2}\sum_1^n (X_i - \bar{X})^2}$$

$$= \left(\frac{1}{2\pi}\right)^{\frac{n}{2}} \cdot \frac{1}{(\hat{\sigma}^2)^{n/2}} \cdot e^{-\frac{n}{2}}$$

Now

$$L(\mu_0, \sigma_0^2) = \left(\frac{1}{2\pi}\right)^{\frac{n}{2}} \cdot \frac{1}{(\sigma_0^2)^{n/2}} \cdot e^{-\frac{1}{2\sigma^2}\sum_1^n (X_i - \mu_0)^2}$$

$$= \left(\frac{1}{2\pi}\right)^{\frac{n}{2}} \frac{1}{(\sigma_0^2)^{\frac{n}{2}}} e^{-\frac{n}{2}}$$

$$\boxed{\sigma_0^2 = \frac{1}{n}\sum_1^n (X_i - \mu_0)^2 \qquad n = \frac{\sum_1^n (X_i - \mu_0)^2}{\sigma_0^2}}$$

& $\lambda(x) = \dfrac{L(\mu_0, \sigma_0^2)}{L(\hat{\mu}, \hat{\sigma}^2)} = \left(\dfrac{\hat{\sigma}^2}{\sigma_0^2}\right)^{\frac{n}{2}}$

Put $\hat{\sigma}^2 = \frac{1}{n}\sum_1^n (X_i - \bar{X})^2$

So

$$\lambda(x) = \left[\frac{\sum_1^n (X_i - \bar{X})^2}{\sum_1^n (X_i - \mu_0)^2}\right]^{\frac{n}{2}}$$

The likelihood ratio test rejects H_0 if $\lambda(x) \leq c$ & is equivalent to

$$(\lambda(x))^{-\frac{2}{n}} = \left[\frac{\sum_1^n (X_i - \bar{X})^2}{\sum_1^n (X_i - \mu_0)^2}\right]^{-1} \leq (c)^{-2/n} = c'$$

$$(\lambda(x))^{-\frac{2}{n}} = \frac{\sum_1^n (X_i - \mu_0)^2}{\sum_1^n (X_i - \bar{X})^2} \leq c'$$

Since $\sum_1^n (X_i - \mu_0)^2 = \sum_1^n (X_i - \bar{X})^2 + n(\bar{X} - \mu_0)^2$

so

$$\frac{\sum_1^n (X_i - \bar{X})^2 + n(\bar{X} - \mu_0)^2}{\sum_1^n (X_i - \bar{X})^2} \leq c'$$

$$1 + \frac{n(\bar{X} - \mu_0)^2}{\sum_1^n (X_i - \bar{X})^2} \leq c'$$

$$\frac{n(\bar{X} - \mu_0)^2}{\sum_1^n (X_i - \bar{X})^2} \leq c' - 1$$

$$\Rightarrow \frac{n(\bar{X}-\mu_0)^2}{\frac{\sum_1^n (X_i-\bar{X})^2}{(n-1)}} \leq (c'-1)(n-1) = c''$$

$$\Rightarrow \left\{ \frac{\sqrt{n}\,(\bar{X}-\mu_0)}{\sqrt{\frac{\sum_1^n (X_i-\bar{X})^2}{(n-1)}}} \right\}^2 \leq c''$$

let $$T = \frac{\sqrt{n}\,(\bar{X}-\mu_0)}{\sqrt{\frac{\sum_1^n (X_i-\bar{X})^2}{(n-1)}}}$$

Then the decision rule is to reject H_0 in favour of H_1 if $|T| \geq c^*$

under H_0, T follows students T-dist. with $(n-1)$ degree of freedom. Based on this, the following decision rule results in a level α-test.

Reject H_0 if
$$|T| \geq t_{\alpha/2}(n-1)$$

i.e we take c^* to be $t_{\alpha/2}(n-1)$ where $t_{\alpha/2}(n-1)$ is the upper $\alpha/2$ critical value of a t-dist. with $(n-1)$ degree of freedom

$\alpha = P_{H_0}(|T| \geq c^*) \longrightarrow ①$

from table

$\alpha = P_{H_0}(|T| \geq t_{\alpha/2}(n-1)) \longrightarrow ②$

* ———— *

Question:- Let $X_1, \dots, X_n$ be a random sample from a $N(\mu_0, \sigma^2 = \theta)$ distribution, where $0 < \theta < \infty$ and μ_0 is known. Show that the likelihood ratio test of $H_0: \theta = \theta_0$ Vs $H_1: \theta \neq \theta_0$ can be based

upon the statistic $W = \sum_1^n (X_i - \mu_0)^2 / \theta_0$.
Determine the null
distribution of W and give explicitly
the rejection rule for a level α-test.

Solution

$$X \sim N(\mu, \sigma^2 = \theta)$$

$$f(x) = \frac{1}{\sqrt{2\pi\theta}} e^{-\frac{1}{2\theta}(x_i - \mu_0)^2}$$

The likelihood function is

$$L(\mu_0, \theta) = \frac{1}{\sqrt{2\pi\theta}} e^{-\frac{1}{2\theta}(x_1 - \mu_0)^2} \cdots \frac{1}{\sqrt{2\pi\theta}} e^{-\frac{1}{2\theta}(x_n - \mu_0)^2}$$

$$= \left(\frac{1}{2\pi\theta}\right)^{\frac{n}{2}} \cdot e^{-\frac{1}{2\theta}\sum_1^n (x_i - \mu_0)^2}$$

The MLE of θ is

$$\hat{\theta} = \frac{1}{n}\sum_1^n (X_i - \mu_0)^2$$

The LRT is

$$\lambda(x) = \frac{L(\theta_0)}{L(\theta)}$$

$$= \frac{\left(\frac{1}{2\pi\theta_0}\right)^{\frac{n}{2}} \cdot e^{-\frac{1}{2\theta_0}\sum_1^n (x_i - \mu_0)^2}}{\left(\frac{1}{2\pi\hat{\theta}}\right)^{\frac{n}{2}} \cdot e^{-\frac{1}{2\hat{\theta}}\sum_1^n (X_i - \mu_0)^2}}$$

$$= \left(\frac{\hat{\theta}}{\theta_0}\right)^{\frac{n}{2}} \frac{e^{-\frac{1}{2\theta_0}\sum_1^n (X_i - \mu_0)^2}}{e^{-n/2}}$$

$$= \left(\frac{\hat{\theta}}{\theta_0}\right)^{\frac{n}{2}} e^{-\frac{1}{2\theta_0}\sum_1^n (X_i - \mu_0)^2 + \frac{n}{2}}$$

$$\Rightarrow \lambda(x) = \left(\frac{\hat{\theta}}{\theta_0}\right)^{\frac{n}{2}} \cdot e^{-\frac{n}{2}\hat{\theta}/\theta_0} \cdot e^{\frac{n}{2}}$$

The decision rule is to reject H_0 if

$$\lambda(x) \leq C$$

But further simplification of the test is possible other than $e^{n/2}$, the test statistic is of the form

$$g(t) = t^{n/2} \exp\left(-\frac{n}{2} t\right), \quad t > 0$$

where $t = \dfrac{\sum_1^n (x_i - \mu_0)^2}{n\theta_0} = \dfrac{\hat{\theta}}{\theta_0}$

Reject H_0 when $g(t) \leq C$

iff $t \leq C_1$ or $t \geq C_2$

This leads to the

$\lambda(x) \leq c$ iff

$$\frac{\sum_1^n (X_i - \mu_0)^2}{\theta_0} \leq nC_1 = C_1' \quad \text{or} \quad \frac{\sum_1^n (X_i - \mu_0)^2}{\theta_0} \geq nC_2 = C_2'$$

under the null Hypothesis, H_0 The statistic $W = \dfrac{\sum_1^n (X_i - \mu_0)^2}{\theta_0}$ has a chi-square distribution

Reject H_0 if $\sum_1^n (X_i - \mu_0)^2/\theta_0 \leq \chi^2_{1-\alpha/2}(n)$

or $\sum_1^n (X_i - \mu_0)^2/\theta_0 \geq \chi^2_{\alpha/2}(n)$

Note:- Testing about mean

① $H_0 : \mu = \mu_0$ Vs $H_1 : \mu \neq \mu_0$, σ^2 is known

$$|Z| \geq c^* \sim \text{Normal distribution}$$

$$\alpha = P_{H_0}\left(|Z| \geq Z_{\alpha/2}\right)$$

② $H_0 : \mu = \mu_0$ Vs $H_1 : \mu \neq \mu_0$, σ^2 is unknown

$$|T| \geq c^* \sim t\text{-distribution}$$

$$\alpha = P_{H_0}\left(|T| \geq t_{\frac{\alpha}{2}}(n-1)\right)$$

③ $H_0 : \sigma^2 = \sigma_0^2$, $H_1 : \sigma^2 \neq \sigma_0^2$, μ_0 is known

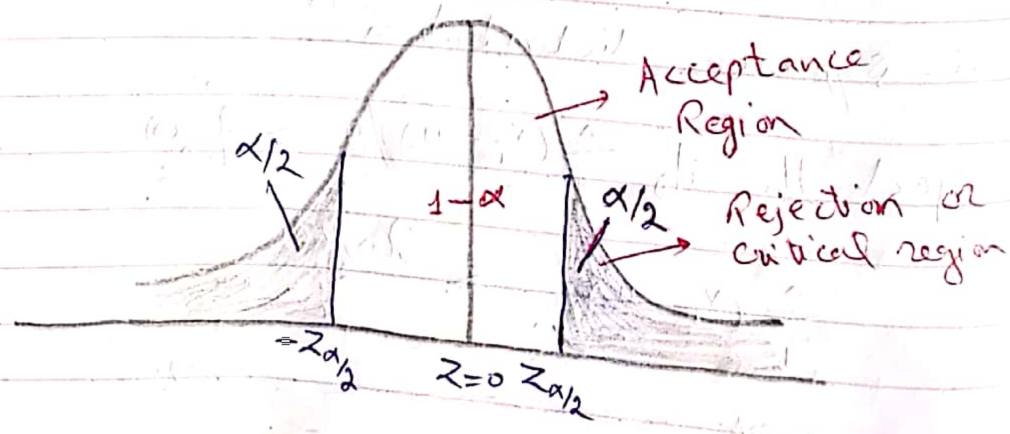
$H_0 : \theta = \theta_0$, $H_1 : \theta \neq \theta_0$

$$|\omega| \geq c^* \sim \chi^2\text{-distribution}$$

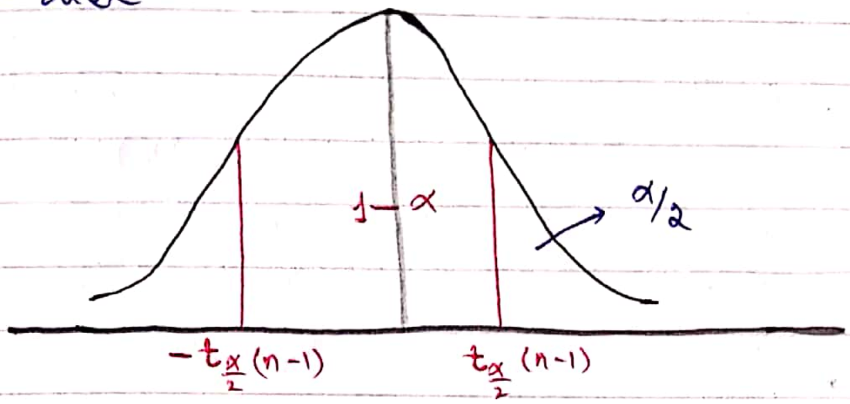
$$\alpha = P_{H_0}\left(\frac{\sum_1^n (X_i - \mu_0)^2}{\theta_0} \leq \chi^2_{1-\alpha/2}(n)\right)$$

or $$\alpha = P_{H_0}\left(\frac{\sum_1^n (X_i - \mu_0)^2}{\theta_0} \geq \chi^2_{\frac{\alpha}{2}}(n)\right)$$

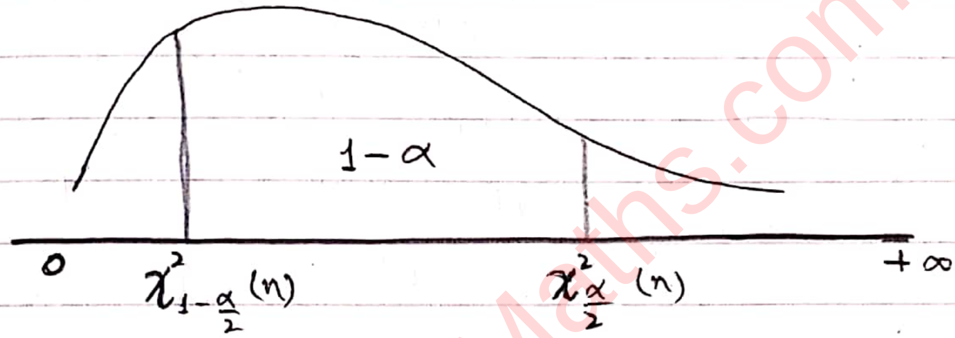
① Standard Normal Distribution



② t - dist



③



Also note that

If $\star$ $X_i \sim$ bernoulli

$\Sigma X_i \sim$ binomial

$\star$ $X_i \sim P(\lambda)$

$\Sigma X_i \sim P(n\lambda)$

$\star$ $X_i \sim \exp(\lambda)$

$\Sigma X_i \sim$ Gamma

MUHAMMAD TAHIR WATTOO

COMSATS UNIVERSITY ISLAMABAD

0344 8563284