Statistical Inference

Department of Statistics & Informative

Fourth Stage

First Semester (2023-2024)

Lecturer: ***Zainab A. M***

References

1. Introduction to Mathematical Statistics, 5th edition; By Robert V. Hogg and Craig,1995.

2. Introduction to Probability Theory and Statistical Inference, 3rd edition; By Harold J. Larson, 1982.

3. Statistical inference / George Casella, Roger L. Berger.-2nd edition 2002.

4. Principles of Statistical Inference, D.R. Cox, 2006.

5. An introduction to Probability and Mathematical Statistics, Rohatgi, V.K. , 1976.

6. Theory of Point Estimation, E.L. Lehmann George Casella 2nd edition 1998.

7. Statistical Distributions. Merran Evans, Nicholas Hastings, Brian Peacock, 3rd Edition, 2000.

7. Mathematical Statistics. Ferguson, T.S. 1968.

8. Statistical inference. Silvey 1973.

9. Bayesian Inference in Statistical Analysis. Box and Tiro 1973.

10. The Theory of Statistical Inference. Zacks, S.

11. Introduction to Probability and Statistical Inference. George Roussas 2003.

12. Probability and Mathematical Statistics. Prasanna Sahoo 2013.

Contents of The First Semester

Chapter one

First review all subjects and laws of Mathematical Statistics

Statistical Distributions

First: The Discrete Distributions:

1) Discrete Uniform Distribution X ~ D.U (n)

$$f(x;n)=\begin{cases}\dfrac{1}{n} & x=1,2,\ldots,n\\ 0 & o.w\end{cases}$$

$$mean=\frac{n+1}{2} \quad , \quad \operatorname{var}(X)=\frac{n^2-1}{12}$$

2) Bernoulli Distribution X ~ Ber (θ)

$$f(x;\theta)=p(x)=\begin{cases}\theta^x(1-\theta)^{1-x} & , \quad x=0,1\\ 0 & o.w\end{cases}$$

$$mean=E(X)=\theta \qquad \operatorname{var}(X)=\theta(1-\theta)$$

3) Binomial Distribution X ~ Bin (n, θ)

$$f(x;n,\theta)=\begin{cases}C_x^n\,\theta^x(1-\theta)^{n-x} & , \quad x=0,1,2,\ldots.,n\\ 0 & o.w\end{cases}$$

$$mean=E(X)=n\theta \qquad \operatorname{var}(X)=n\theta(1-\theta)$$

4) Negative Binomial Distribution X ~ N.Bin (r, θ)

$$f(x;r,\theta)=\begin{cases}C_x^{x+r-1}\theta^r(1-\theta)^x & , \quad x=0,1,2,\ldots..\\ 0 & o.w\end{cases}$$

$$mean=E(X)=\frac{r(1-\theta)}{\theta} \quad , \quad v(X)=\frac{r(1-\theta)}{\theta^2}$$

5) Geometric Distribution X ~ Geo (θ)

$$f(x;\theta)=\begin{cases}\theta(1-\theta)^x & , \quad x=0,1,2,\ldots\ldots..\\ 0 & o.w\end{cases}$$

$$mean=E(X)=\frac{(1-\theta)}{\theta} \quad , \quad v(X)=\frac{(1-\theta)}{\theta^2}$$

6) The Poisson Distribution X ~ Poi (θ)

$$f(x;\theta) = \begin{cases} \dfrac{e^{-\theta}\,\theta^{x}}{x!} & , \quad x = 0,1,2,\ldots. \\ 0 & o.w \end{cases}$$

$$mean = E(X) = \theta \qquad , \quad v(X) = \theta$$

Second: The Continuous Distributions:

1. Continuous Uniform Distribution X ~ C.U(a , b)

$$f(x;a,b) = \begin{cases} \dfrac{1}{b-a} & , \quad a \le x \le b \\ 0 & o.w \end{cases}$$

$$mean = E(X) = \frac{a+b}{2} \qquad , \quad v(X) = \frac{(b-a)^2}{12}$$

2. Beta Distribution X ~ *Beta* (α, β)

$$f(x;\alpha,\beta) = \begin{cases} \dfrac{1}{B(\alpha,\beta)}\, x^{\alpha-1}(1-x)^{\beta-1} & 0 < x < 1 \\ 0 & o.w \end{cases}$$

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

3. Gamma Distribution:

a) Gamma, Distribution

1. X ~ Γ(α, θ)

$$f(x;\alpha,\theta) = \begin{cases} \dfrac{1}{\Gamma(\alpha)\theta^{\alpha}}\, x^{\alpha-1} e^{-x/\theta} & , \quad x > 0 \\ 0 & o.w \end{cases}$$

$$mean = E(X) = \alpha\theta \qquad , \quad v(X) = \alpha\theta^2$$

2. X ~ Γ(α, 1/θ)

$$f(x;\alpha,\theta) = \begin{cases} \dfrac{\theta^{\alpha}}{\Gamma(\alpha)}\, x^{\alpha-1} e^{-\theta x} & , \quad x > 0 \\ 0 & o.w \end{cases}$$

$$mean = E(X) = \alpha/\theta \qquad , \quad v(X) = \alpha/\theta^2$$

b) Exponential Distribution

1. X ~ Exp(θ)

$$f(x;\,\theta) = \begin{cases} \dfrac{1}{\theta}\, e^{-x/\theta} & ,\; x > 0 \\ 0 & o.w \end{cases}$$

$$mean = E(X) = \theta \quad , \quad v(X) = \theta^2$$

2. X ~ Exp(1/θ)

$$f(x;\,\theta) = \begin{cases} \theta\, e^{-\theta x} & ,\; x > 0 \\ 0 & o.w \end{cases}$$

$$mean = E(X) = 1/\theta \quad , \quad v(X) = 1/\theta^2$$

c) Chi-Square Distribution X ~ $\chi^2_{(r)}$

$$f(x;r) = \begin{cases} \dfrac{1}{\Gamma(r/2)\, 2^{r/2}}\, x^{\frac{r}{2}-1} e^{-x/2} & ,\; x > 0 \quad ,\; r > 0 \\ 0 & o.w \end{cases}$$

$$mean = E(X) = r \quad , \quad v(X) = 2r$$

4. Normal (Gaussian) Distribution X ~ N(θ, σ^2)

$$f(x;\,\theta,\sigma) = \begin{cases} \dfrac{1}{\sqrt{2\pi}\,\sigma}\, e^{-\frac{1}{2\sigma^2}(x-\theta)^2} & ,\; -\infty < x < \infty \\ 0 & o.w \end{cases}$$

$$mean = E(X) = \theta \quad , \quad v(X) = \sigma^2 \quad , \; -\infty < \theta < \infty \quad , \quad \sigma > 0$$

Chapter Two

Distributions of Functions of Random Variables

First: Transformations of the Discrete Random Variables

If X is a discrete r.v., having p.d.f. $f(x)$, taking values in sample space $\mathcal{S}$, $A = \{x;\ x = x_1, x_2, \ldots, x_n\}$, at each of which $f(x) > 0$, and let a r.v. $y = g(x)$ define a **one-to-one transformation** that maps A onto B, $B = \{y;\ y = y_1, y_2, \ldots, y_n\}$.
If we solve $y = g(x)$ for x in terms of y, say, $x = w(y)$, then for each $y \in B$, we have $x = w(y) \in A$.
Then to find the p.d.f. of Y, is given as follows;

$$f(y) = p(Y = y) = p(X = w(y)) = \begin{cases} f[w(y)] & , \quad y \in B \\ 0 & o.w \end{cases}$$

Ex: Let $X \sim poi(\theta)$ and $Y = 4X$ by using transformation technique, find the p.d.f. of Y.
Sol:

$$\because X \sim poi(\theta) \quad \therefore p.d.f.\ of\ X = f(x) = \frac{e^{-\theta}\theta^{x}}{x!} \quad , \ x = 0, 1, 2, \ldots\ldots$$

$$A = \{x : x = 0, 1, 2, \ldots\ldots\} \quad , \ f(x) > 0$$

$$\because Y = 4X$$

$$\therefore B = \{y : y = 0, 4, 8, \ldots\ldots\} \quad , \ f(y) > 0$$

$$f(y) = p(Y = y) = p(4X = y) = p(X = \tfrac{y}{4}) = \begin{cases} \dfrac{e^{-\theta}\theta^{\frac{y}{4}}}{\frac{y}{4}!} & , \ y = 0, 4, 8, \ldots\ldots \\ 0 & o.w \end{cases}$$

Ex: Let X have the binomial p.d.f. . $X \sim Bin(3, 2/3)$, where $Y = X^2$, by using one-to-one transformation, find the p.d.f. of Y.
Sol:

$$\because X \sim b(3, 2/3) \quad \Rightarrow \therefore f(x) = \begin{cases} C_x^3 (\frac{2}{3})^x (\frac{1}{3})^{3-x} & x = 0,1,2,3 \\ 0 & o.w \end{cases}$$

$$A = \{x : x \in R_X = 0, 1, 2, 3\} \quad , \ f(x) > 0$$

$$\because Y = X^2 \quad \Rightarrow \quad \therefore B = \{y : y \in R_Y = 0, 1, 4, 9\} \quad , \ f(y) > 0$$

In general, $Y = X^2$ does not define a one-to-one transformation, but here there are not negative values of x in $A = \{x;\ x = 0, 1, 2, 3\}$, then $x = w(y) = \sqrt{y}$ $(not\ -\sqrt{y})$, and so;

$$f(y) = p(Y = y) = p(X^2 = y) = p(X = \mp\sqrt{y}) = p(X = \sqrt{y})$$

$$= \begin{cases} \dfrac{3!}{(\sqrt{y})!\,(3-\sqrt{y})!} (\frac{2}{3})^{\sqrt{y}} (\frac{1}{3})^{3-\sqrt{y}} & y = 0,1,4,9 \\ 0 & o.w \end{cases}$$

Definition for the J.P.D.F.

Let $f(x_1, x_2)$ be the j.p.d.f. of two discrete r.v.'s X_1 and X_2 with A the (two dimensional) set of points. Which $f(x_1, x_2) > 0$, let $y_1 = g_1(x_1, x_2)$ and $y_2 = g_2(x_1, x_2)$ define a one-to-one transformation that maps A onto B (two dimensional), then the j.p.d.f. of the two new r.v.'s $Y_1 = g_1(X_1, X_2)$, $Y_2 = g_2(X_1, X_2)$ is given;

$$f(y_1, y_2) = p(Y_1 = y_1, Y_2 = y_2) = \begin{cases} f[w_1(y_1, y_2), w_2(y_1, y_2)] & , (y_1, y_2) \in B \\ 0 & o.w \end{cases}$$

Ex: Let X_1 and X_2 be two stochastically independent r.v.'s that have Poisson distribution with means θ_1, θ_2 respectively, the j.p.d.f. of X_1 and X_2 is;

$$f(x_1, x_2) = \begin{cases} \dfrac{\theta_1^{x_1}\, \theta_2^{x_2}\, e^{-\theta_1 - \theta_2}}{x_1!\, x_2!} & , x_1 = 0,1,2,3,\ldots.. \quad , x_2 = 0,1,2,3,\ldots.. \\ 0 & o.w \end{cases}$$

Where $Y_1 = X_1 + X_2$, $Y_2 = X_2$. **Find:** the j.p.d.f. of Y_1 and Y_2. and $f_1(y_1)$. , **HW**: $f_2(y_2)$

Sol:

$\because X_1 \text{ and } X_2 \sim Poi(\theta_1, \theta_2)$

$A = \{(x_1, x_2) : x_1 = 0,1,2,\ldots.. \quad , x_2 = 0,1,2,\ldots..\} \ , \ f(x_1, x_2) > 0$

$B = \{(y_1, y_2) : y_1 = 0,1,2,\ldots.. \quad , y_2 = 0,1,2,\ldots., y_1\} \ , \ f(y_1, y_2) > 0$

$because \ ; \ y_2 = x_2 \Rightarrow \because x_2 = y_1 - x_1 \Rightarrow \therefore y_2 = y_1 - x_1$

$when \ x_1 = 0 \Rightarrow y_2 = y_1 \ (\max) \ldots (y_1 - 1, y_1 - 2, \ldots.) \ldots \ when \ x_1 = \infty \Rightarrow y_2 = \infty - \infty = 0 \ (\min)$

$\because y_1 = x_1 + x_2 \Rightarrow x_1 = y_1 - x_2 \Rightarrow x_1 = y_1 - y_2$

$\quad y_2 = x_2 \qquad\qquad\qquad \Rightarrow x_2 = y_2$

$\therefore the\ j.p.d.f.\ of\ Y_1\ and\ Y_2\ is;$

$$f(y_1, y_2) = p(Y_1 = y_1, Y_2 = y_2) = p(X_1 = y_1 - y_2, X_2 = y_2)$$

$$= \begin{cases} \dfrac{\theta_1^{y_1 - y_2}\, \theta_2^{y_2}\, e^{-\theta_1 - \theta_2}}{(y_1 - y_2)!\, y_2!} & , (y_1, y_2) \in B \\ 0 & o.w \end{cases}$$

The marginal p.d.f. of Y_1 is given by;

$$f_1(y_1) = \sum_{y_2=0}^{y_1} f(y_1, y_2) = e^{-\theta_1 - \theta_2} \sum_{y_2=0}^{y_1} \frac{\theta_1^{y_1 - y_2}\, \theta_2^{y_2}}{(y_1 - y_2)!\, y_2!} \} \times \frac{y_1!}{y_1!}$$

$$= \frac{e^{-\theta_1 - \theta_2}}{y_1!} \sum_{y_2=0}^{y_1} \frac{y_1!}{(y_1 - y_2)!\, y_2!} \theta_1^{y_1 - y_2}\, \theta_2^{y_2} = \frac{e^{-\theta_1 - \theta_2}}{y_1!} \sum_{y_2=0}^{y_1} C_{y_2}^{y_1}\, \theta_1^{y_1 - y_2}\, \theta_2^{y_2}$$

$$= \frac{e^{-\theta_1 - \theta_2} (\theta_1 + \theta_2)^{y_1}}{y_1!} \quad , \quad y_1 = 0,1,2,\ldots...$$

That is, $Y_1 = X_1 + X_2$ has a Poisson distribution with parameter $(\theta_1 + \theta_2)$.

Distribution of Order Statistics

Let $X_1, X_2, \ldots, X_n$ denote a random sample and be independent identically distributed r.v's with a p.d.f. $f(x)$, and let $Y_1 < Y_2 < \ldots < Y_n$ be their ascending ordered values, i.e.;

Y_1: is a smallest value of $(X_1, X_2, \ldots, X_n)$ (min).
Y_2: is the second smallest value of $(X_1, X_2, \ldots, X_n)$.
.
.
Y_n: the largest value of $(X_1, X_2, \ldots, X_n)$ (max).

Then Y_i $(i = 1, 2, \ldots, n)$ is called the i-th order statistic of the random sample $X_1, X_2, \ldots, X_n$. and $Y_1 < Y_2 < \ldots < Y_n$ are called the order statistics corresponding of the random sample $X_1, X_2, \ldots, X_n$.

Then the j.p.d.f. of $X_1, X_2, \ldots, X_n$ is given by;

$$f(x_1, x_2, \ldots, x_n) = f(x_1).\, f(x_2) \ldots f(x_n)$$

The j.p.d.f. of the order statistics $Y_1, Y_2, \ldots, Y_n$ is given by;

$$g(y_1, y_2, \ldots, y_n) = (n!)\; g(y_1)\; g(y_2) \ldots g(y_n)$$

$$= \begin{cases} (n!) \prod_{i=1}^{n} g(y_i) & , \quad a < y_1 < y_2 < \ldots < y_n < b \\ 0 & o.w \end{cases}$$

Explain:

Let $(n = 2)$, then we have two probabilities;

$X_1 > X_2$	or	$X_1 < X_2$
$Y_1 = X_2$		$Y_1 = X_1$
$Y_2 = X_1$		$Y_2 = X_2$

Discrete

$$g(y_1, y_2) = g(y_1 = x_2)\; g(y_2 = x_1) + g(y_1 = x_1)\; g(y_2 = x_2)$$

$$= (2!)\; g(y_1)\; g(y_2)$$

$$= \begin{cases} (2!) \prod_{i=1}^{2} g(y_i) & , \quad a < y_1 < y_2 < b \\ 0 & o.w \end{cases}$$

When $(n = 3)$

$$g(y_1, y_2, y_3) = (3!)\; g(y_1)\; g(y_2)\; g(y_3)$$

$$= \begin{cases} (3!) \prod_{i=1}^{3} g(y_i) & , \quad a < y_1 < y_2 < y_3 < b \\ 0 & o.w \end{cases}$$

Continuous

When $(n = 2)$

$$J_1 = \begin{vmatrix} \frac{dx_1}{dy_1} & \frac{dx_1}{dy_2} \\ \frac{dx_2}{dy_1} & \frac{dx_2}{dy_2} \end{vmatrix} = \begin{vmatrix} 0 & 1 \\ 1 & 0 \end{vmatrix} = 1 \quad , \quad J_2 = \begin{vmatrix} \frac{dx_1}{dy_1} & \frac{dx_1}{dy_2} \\ \frac{dx_2}{dy_1} & \frac{dx_2}{dy_2} \end{vmatrix} = \begin{vmatrix} 1 & 0 \\ 0 & 1 \end{vmatrix} = 1$$

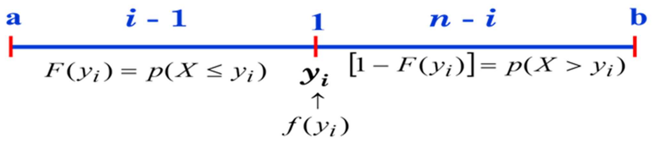
Note: Always J equal to one because be continuous

$$g(y_1, y_2) = g(y_1 = x_2)\ g(y_2 = x_1)\ J_1 + g(y_1 = x_1)\ g(y_2 = x_2)\ J_2$$

The Marginal P.D.F. of an Individual Order Statistics

The marginal p.d.f. of the *i*-th order statistics is given by:

$$g(y_i) = \frac{n!}{(i-1)!(n-i)!} f(y_i)[F(y_i)]^{i-1}[1-F(y_i)]^{n-i} \quad , \quad a < y_i < b$$



The P.D.F. of the Smallest Order Statistics

If (i = 1) then the distribution of y_1 is given by:

$$g(y_1) = \frac{n!}{(1-1)!(n-1)!} f(y_1)[F(y_1)]^{1-1}[1-F(y_1)]^{n-1} \quad , \quad a < y_1 < b$$

$$g(y_1) = \frac{n(n-1)!}{(n-1)!} f(y_1)\,[1-F(y_1)]^{n-1}$$

$$g(y_1) = n\, f(y_1)\,[1-F(y_1)]^{n-1} \quad , \quad a < y_1 < b$$

The P.D.F. of the Largest Order Statistics

If ($i = n$) then the distribution of y_n is given by:

$$g(y_n) = \frac{n!}{(n-1)!(n-n)!} f(y_n)[F(y_n)]^{n-1}[1-F(y_n)]^{n-n} \quad , \quad a < y_n < b$$

$$g(y_n) = \frac{n(n-1)!}{(n-1)!} f(y_n)\ [F(y_n)]^{n-1}$$

$$g(y_n) = n\, f(y_n)\,[F(y_n)]^{n-1} \quad , \quad a < y_n < b$$

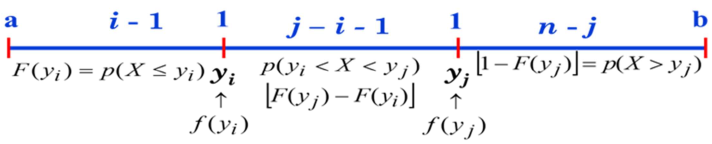
The Joint Probability Density Fun. of Two Order Statistics

The joint p.d.f. of any two order statistics Y_i and Y_j ($i < j$) is given by:

$$g(y_i, y_j) = \frac{n!}{(i-1)!(j-i-1)!(n-j)!} \times$$

$$\times\ [F(y_i)]^{i-1}[F(y_j) - F(y_i)]^{j-i-1}[1-F(y_j)]^{n-j} f(y_i)\, f(y_j) \quad , \quad a < y_i < y_j < b$$

$$0 \qquad\qquad o.w$$



Ex: let $X_1, X_2, \ldots, X_n$ be a random sample of size (n) rssn taken from C.U(0,1). let $Y_1 < Y_2 < \ldots < Y_n$ be the order statistics of this sample. **Find** the p.d.f. of Y_1 and Y_n, the j.p.d.f. of Y_1 and Y_n

Sol.: $X \sim C.U(0,1)$

$$\therefore f(x) = \begin{cases} \frac{1}{b-a} & , a \le x \le b \\ 0 & o.w \end{cases} = \begin{cases} 1 & , 0 \le x \le 1 \\ 0 & o.w \end{cases}$$

$$F(x) = p(X \le x) = \int_0^x 1\,dx = \begin{cases} 0 & , x \le 0 \\ x & , 0 < x < 1 \\ 1 & , x \ge 1 \end{cases}$$

$$g(y_1) = n\ f(y_1)\,[1 - F(y_1)]^{n-1} \quad , \quad a \le y_1 \le b$$

$$when\ \ x = y_1 \ \Rightarrow\ \therefore\ f(y_1) = 1 \ ,\ F(y_1) = y_1$$

$$\therefore g(y_1) = n\,(1)\,[1 - y_1]^{n-1} \quad , \quad 0 \le y_1 \le 1$$

$$= \begin{cases} n(1-y_1)^{n-1} & , 0 \le y_1 \le 1 \\ 0 & o.w \end{cases}$$

$$g(y_n) = n\ f(y_n)\,[F(y_n)]^{n-1} \quad , \quad a \le y_n \le b$$

$$when\ \ x = y_n \ \Rightarrow\ \therefore\ f(y_n) = 1$$

$$\therefore\ g(y_n) = n\,(1)\,[y_n]^{n-1} \quad , \quad 0 \le y_n \le 1$$

$$= \begin{cases} n\,y_n^{n-1} & , 0 \le y_n \le 1 \\ 0 & o.w \end{cases}$$

$$g(y_i, y_j) = \frac{n!}{(i-1)!(j-i-1)!(n-j)!} \times$$

$$\times\ [F(y_i)]^{i-1}\,[F(y_j) - F(y_i)]^{j-i-1}\,[1 - F(y_j)]^{n-j} f(y_i)\,f(y_j) \quad , \quad a < y_i < y_j < b$$

$$0 \qquad o.w$$

When, $i = 1\ ,\ j = n$

$$\therefore g(y_1, y_n) = \frac{n!}{0!(n-2)!0!}(y_n - y_1)^{n-2} = \frac{n(n-1)(n-2)!}{(n-2)!}(y_n - y_1)^{n-2}$$

$$= n(n-1)\,(y_n - y_1)^{n-2} \quad , \ 0 < y_1 < y_n < 1$$

Chapter Three

Statistical Inference

Statistical Inference: making conclusions about the whole population on the basis of a sample, i.e., use a random sample to learn something about a large population.
Precondition for statistical inference: A sample is randomly selected from the population.

Concepts and Important Definitions about Stat. Inference

1. $\underline{X} = (X_1, X_2, \ldots, X_n) \equiv$ rssn $\equiv$ Data
2. Statistic: is a function of the random variable (r.v.) only in the sample data.
3. Parameter: It is a characteristic or a measure that is calculated from the population under study. **Ex:** The unemployment rate in Erbil. The average of assumption life for a particular device. **[Parameter = Statistic ± It's Error]**.
4. Population parameters are denoted using Greek letters μ (mean), σ (standard deviation), π (proportion). Sample values are denoted $\overline{x}$ (mean), S (standard deviation), p (proportion).

5. Estimator: is a function.
6. Estimate: is a value of the estimator.

$$\overline{X} = \underbrace{\frac{\sum X_i}{n}}_{Estimator} = \underbrace{15}_{Estimate}$$

7.

$$Quantitative\ Variable \Rightarrow Standard\ \ Error = SE(Mean) = S / \sqrt{n}$$

$$Qualitative\ Variable \Rightarrow Standard\ \ Error = SE(p) = \sqrt{p(1-p)/n}$$

There are two steps to make inference:

1. Estimation of the population parameters

a) Point Estimation.
b) Intervals Estimation.

2. Testing of Hypotheses about the right values of population parameters.

Estimation of Parameters

First: Point Estimation

Let $X_1, X_2, \ldots, X_n$ be a rssn from the p.d.f. $f(x;\theta)$, θ is unknown. We want to estimate θ from the information in the data.

$$\hat{\theta} = estimator\ \ of\ \ \theta$$

Properties of Estimator:

1. Unbiased Estimator

An estimator $\left(\hat{\theta} = t(x_1, \ldots, x_n)\right)$ from a sample of size (n) with p.d.f. $f(x;\theta)$ is said to be an unbiased estimator for a population parameter θ if:

$$E(\hat{\theta}) = \theta$$

The quantity $\left(E(\hat{\theta}) - \theta\right)$ is called bias of an estimator $\hat{\theta} = t(X)$ of θ.

$$bias\ (\hat{\theta}) = E(\hat{\theta}) - \theta$$

Ex: In a random sample of size (n) taken from exponential dist^n Exp(θ). Show that;

1. $T_1 = \overline{X}$ is unbiased estimator for the parameter (θ).

2. $T_2 = \frac{n}{n+1} \overline{X}^2$ is unbiased estimator for the parameter (θ^2).

Sol: 1)

$$E(T_1) \overset{?}{=} \theta$$

$$E(T_1) = E(\overline{X}) = \frac{1}{n} E(\sum X_i) = \frac{1}{n}\, n\, E(X) = \theta$$

$\therefore\ \overline{X}$ is unbiased estimator for θ.

2)

$$E(T_2) = \frac{n}{n+1} E(\overline{X})^2$$

$$= \frac{n}{n+1}\left(V(\overline{X}) + (E(\overline{X}))^2\right) = \frac{n}{n+1}\left(\frac{V(X)}{n} + \theta^2\right)$$

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

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

$\therefore\ T_2 = \frac{n}{n+1} \overline{X}^2$ is unbiased estimator for θ^2.

Ex: In a random sample of size (n) from normal dist^n N(θ, σ^2). Show that;

1) $S^2 = \frac{1}{n-1} \sum (X_i - \overline{X})^2$ is unbiased estimator for the parameter (σ^2).

2) Is $T = \overline{X}^2$ unbiased estimator for θ^2.

Sol: 1)

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

$$E(\overline{X}) = E\left(\frac{\sum X_i}{n}\right) = \frac{n}{n} E(X) = \theta$$

$$V(\overline{X}) = V\left(\frac{\sum X_i}{n}\right) = \frac{1}{n^2} \sum V(X_i) = \frac{n}{n^2} \sigma^2 = \frac{\sigma^2}{n}$$

$$\overline{X} \sim N(\theta, \frac{\sigma^2}{n})$$

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

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

$$= \frac{n}{n-1}\left(\left(V(X) + (E(X))^2\right) - \left(V(\overline{X}) + (E(\overline{X}))^2\right)\right)$$

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

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

$$\therefore\ S^2 = \frac{1}{n-1}\sum (X_i - \overline{X})^2 \ \textit{is unbiased estimator for } \sigma^2$$

2)

$$E(\overline{X})^2 \overset{?}{=} \theta^2$$

$$V(\overline{X}) = E(\overline{X})^2 - (E(\overline{X}))^2$$

$$E(\overline{X})^2 = V(\overline{X}) + (E(\overline{X}))^2 = \frac{\sigma^2}{n} + \theta^2 \neq \theta^2$$

$$\therefore\ \hat{\theta} = \overline{X}^2 \ \textit{is not unbiased estimator for } \theta^2.$$

Then; what is to be unbiased estimator for θ^2.

Now from both sides we subtract $\frac{\sigma^2}{n}$;

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

$$\therefore\ \hat{\theta} = \overline{X}^2 - \frac{\sigma^2}{n} \ \textit{is unbiased estimator for } \theta^2$$

Unbiased in Limit

An estimator $\hat{\theta}$ for known parameter θ of p.d.f. $f(x;\theta)$ is unbiased in limit if:

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

Ex: In a rss(n) from uniform dist^n C.U(0, θ).

1) Is Y_n unbiased in limit estimator for θ; (Note: Y_n estimator θ).

2) Is $\overline{X}$ unbiased in limit estimator for θ.

3) Is $\overline{X}$ unbiased in limit estimator for $\theta/2$.

Sol: 1)

$$f(x) = \frac{1}{b-a} = \frac{1}{\theta - 0} = \frac{1}{\theta} \quad , 0 < x < \theta$$

$$F(y_i) = p(X \le y_i) = \int_0^{y_i} \frac{1}{\theta}\,dx = \frac{y_i}{\theta}$$

$$g(y_n) = n\ f(y_n)\,(F(y_n))^{n-1} = n\,\frac{1}{\theta}\left(\frac{y_n}{\theta}\right)^{n-1} = \frac{n\,y_n^{n-1}}{\theta^n}\quad ,0 < y_n < \theta$$

$$E(Y_n) = \int_{R_{y_n}} y_n\ g(y_n)\ dy_n = \int_0^\theta y_n\ \frac{n\,y_n^{n-1}}{\theta^n}\ dy_n = \frac{n}{\theta^n}\int_0^\theta y_n^n\ dy_n$$

$$= \frac{n}{\theta^n}\ \frac{y_n^{n+1}}{n+1}\Bigg|_0^\theta = \frac{n}{\theta^n}\ \frac{\theta^{n+1}}{n+1} = \frac{n}{n+1}\ \theta \neq \theta\ \rightarrow \therefore Y_n \text{ is not unbiased est. for } \theta$$

$$\lim_{n\to\infty} E(Y_n) = \lim_{n\to\infty} \frac{n}{n+1}\ \theta = (1)\ \theta = \theta\ \rightarrow \therefore Y_n \text{ is unbiased in limit est. for } \theta$$

2. Consistency Estimator

Definition: An estimator $\hat{\theta}$ of the parameter θ of $f(x;\theta)$ is called consistent estimator for θ if;

$$\lim_{n\to\infty} p\left(\left|\hat{\theta} - \theta\right| < \varepsilon\right) = 1 \quad , \quad \forall\ \varepsilon > 0$$

$$or\ ;\ \lim_{n\to\infty} p\left(\left|\hat{\theta} - \theta\right| \geq \varepsilon\right) = 0$$

Note: Consistency means the estimator equal to the parameter or converges stochastically to the parameter θ.

A consistent estimator: That the estimator gets closer to the parameter value as n increases without limit.

$$|\,\hat{\theta} - \theta\,| \ \Rightarrow \text{called estimated error}$$

$$\left.\begin{aligned} p\left(\left|\hat{\theta} - \theta\right| < \varepsilon\right) &\geq 1 - \frac{v(\hat{\theta})}{\varepsilon^2} \\ p\left(\left|\hat{\theta} - \theta\right| \geq \varepsilon\right) &< \frac{v(\hat{\theta})}{\varepsilon^2} \end{aligned}\right\} \rightarrow (\text{Chebycheve inequality})$$

Theorem: Let $\hat{\theta}$ be an estimator for the population parameter θ of $f(x;\theta)$, then $\hat{\theta}$ is said to be consistent estimator for θ if:

$$1)\ \lim_{n\to\infty} E(\hat{\theta}) = \theta \qquad\qquad 2)\ \lim_{n\to\infty} V(\hat{\theta}) = 0$$

Ex: Let $X_1, X_2, \ldots, X_n$ be a rss$n$ from Poisson dist[n], show that $\hat{\theta} = \overline{X}$ is consistent estimator for θ.

Sol:

First Method ;

$$1)\quad p\left(\left|\hat{\theta}-\theta\right|<\varepsilon\right)\ge 1-\frac{v(\hat{\theta})}{\varepsilon^{2}}$$

$$\lim_{n\to\infty} p\left(\left|\hat{\theta}-\theta\right|<\varepsilon\right)\ge \lim_{n\to\infty}\left(1-\frac{v(\hat{\theta})}{\varepsilon^{2}}\right)$$

$$\hat{\theta}=\overline{X}\ ,\ v(\hat{\theta})=v(\overline{X})=\frac{v(X)}{n}=\frac{\theta}{n}$$

$$\lim_{n\to\infty} p\left(\left|\overline{X}-\theta\right|<\varepsilon\right)\ge \lim_{n\to\infty}\left(1-\frac{\theta}{n\,\varepsilon^{2}}\right)$$

$$\lim_{n\to\infty} p\left(\left|\overline{X}-\theta\right|<\varepsilon\right)=1$$

$$2)\quad p\left(\left|\hat{\theta}-\theta\right|\ge\varepsilon\right)<\frac{v(\hat{\theta})}{\varepsilon^{2}}$$

$$\lim_{n\to\infty} p\left(\left|\hat{\theta}-\theta\right|\ge\varepsilon\right)<\lim_{n\to\infty}\frac{\theta}{n\,\varepsilon^{2}}$$

$$\lim_{n\to\infty} p\left(\left|\overline{X}-\theta\right|\ge\varepsilon\right)=0$$

Second Method ;

$$1)\quad \lim_{n\to\infty} E(\hat{\theta})=\lim_{n\to\infty} E(\overline{X})=\theta$$

$$2)\quad \lim_{n\to\infty} v(\hat{\theta})=\lim_{n\to\infty}\frac{\theta}{n}=0$$

$\therefore\ \hat{\theta}=\overline{X}$ *is consistent estimator for* θ.

The Score Function

The score function is the partial derivative of Log the function $f(x;\theta)$ with respect to the parameter θ, is defined as;

$$S(x;\theta)=\frac{\partial}{\partial\theta}\ln f(x;\theta)=\frac{1}{f(x;\theta)}\frac{\partial}{\partial\theta}f(x;\theta)$$

Properties

1) The mean of the score is zero, E(S(X; θ)) = zero

Proof:

$$E\big(S(X;\theta)\big)=\int_{R_X} s(x;\theta)\,f(x;\theta)\,dx=\int_{R_X}\frac{1}{f(x;\theta)}\frac{\partial}{\partial\theta}f(x;\theta)\,f(x;\theta)\,dx$$

$$=\int_{R_X}\frac{\partial}{\partial\theta}f(x;\theta)\,dx=\frac{\partial}{\partial\theta}\int_{R_X}f(x;\theta)\,dx=\frac{\partial}{\partial\theta}(1)=zero$$

2) The variance of the score is known as the Fisher Information (F.I), which is measure the information in the sample $\mathcal{S}$ about the parameter *θ*, and can be written as;

$$F.I = I(\theta) = E\left(\frac{\partial}{\partial\theta}\ln f(x;\theta)\right)^2 \qquad , \text{because mean} = \text{zero}$$

Or;

$$F.I = I(\theta) = -E\left(\frac{\partial^2}{\partial\theta^2}\ln f(x;\theta)\right)$$

If Fisher Information multiply by (n), *we get;*

$$n\,I(\theta) = F.I \text{ in a } rss(n)$$

Ex: Let $X_1, \ldots, X_n$ be a rss$n$ from exponential dist^n Exp($1/\theta$). Find the F.I. of X.

Sol:

$$f(x;\theta) = \theta e^{-\theta x} \qquad , x > 0$$

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

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

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

$$F.I = -E\left(\frac{\partial^2}{\partial\theta^2}\ln f(x;\theta)\right) = \frac{1}{\theta^2} \qquad \Rightarrow \qquad \therefore n\,I(\theta) = F.I. \text{ in a } rss(n) = \frac{n}{\theta^2}$$

3. Sufficiency Estimator

Sufficiency estimator is containing all the information in the data about the parameter *θ*.

First Method (Fisher Information)

Definition 1: Let $X_1, X_2, \ldots, X_n$ be a rss$n$ from the dist^n with p,d.f. $f(x\,;\,\theta)$, an estimator $\hat{\theta}$ is sufficient estimator for the parameter *θ* if the Fisher information in $\hat{\theta}$ is equal to the Fisher information in a rss(n).

Ex: Show that $\overline{X}$ is sufficient estimator for the mean of $N(\theta, \sigma^2)$.

Sol:

$$F.I \text{ in a } rssn = F.I \text{ in } \hat{\theta}$$

$$F.I \text{ in a } rssn = -n\,E\left(\frac{\partial^2 \ln f(x;\theta)}{\partial\theta^2}\right)$$

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

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

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

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

$$-nE\left(\frac{\partial^2 \ln f(x;\theta,\sigma^2)}{\partial \theta^2}\right) = \frac{n}{\sigma^2} \quad is\ F.I.\ in\ a\ rssn$$

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

$$\overline{X} \sim N(\theta, \frac{\sigma^2}{n})$$

$$g(\overline{x};\theta,\frac{\sigma^2}{n}) = \frac{1}{\sqrt{2\pi\frac{\sigma^2}{n}}} e^{-\frac{1}{2\frac{\sigma^2}{n}}(\overline{x}-\theta)^2}$$

$$\ln\left(g(\overline{x};\theta,\frac{\sigma^2}{n})\right) = \ln\left(\frac{1}{\sqrt{2\pi\frac{\sigma^2}{n}}}\right) - \frac{n}{2\sigma^2}(\overline{x}-\theta)^2$$

$$\frac{\partial \ln g(\overline{x};\theta,\frac{\sigma^2}{n})}{\partial \theta} = zero - \frac{2n(\overline{x}-\theta)(-1)}{2\sigma^2} = \frac{n(\overline{x}-\theta)}{\sigma^2}$$

$$\frac{\partial^2 \ln g(\overline{x};\theta,\frac{\sigma^2}{n})}{\partial \theta^2} = -\frac{n}{\sigma^2}$$

$$-E\left(\frac{\partial^2 \ln g(\overline{x};\theta,\frac{\sigma^2}{n})}{\partial \theta^2}\right) = -E\left(\frac{-n}{\sigma^2}\right) = \frac{n}{\sigma^2} \quad is\ F.I.\ (\hat{\theta} = \overline{x})$$

$$\because F.I\ in\ a\ rssn = F.I\ in\ (\hat{\theta} = \overline{x})$$

$$\therefore \hat{\theta} = \overline{x}\ is\ suff\ est\ for\ \theta$$

Second Method (Conditional)

Definition 2: Let $X_1, X_2, \ldots, X_n$ be a rss$n$ from the dist^n with p,d.f. $f(x\ ;\ \theta)$, and $\hat{\theta}$ be an estimator for θ, an estimator $\hat{\theta}$ is sufficient estimator for the parameter θ if the conditional p.d.f. of $(X_1, X_2, \ldots, X_n)$ given $\hat{\theta}$ does not contain the parameter θ:

$$f(x_1, x_2, \ldots, x_n \mid \hat{\theta}) = \frac{f(x_1, x_2, \ldots, x_n; \theta)}{g(\hat{\theta})}$$

Note: If the range depends on the parameter, in this case we can't find F.I; therefore, we use the second method (Conditional).

Ex: Let $X_1, X_2, \ldots, X_n$ be a rss$n$ from a dist^n with p.d.f.:

$$f(x;\theta) = e^{2\theta - x} \quad , \quad x \ge 2\theta$$

Show that Y_1 is sufficient estimator for the parameter θ.

Sol:

$$f(x_1, x_2, ..., x_n \mid \hat{\theta} = y_1) = \frac{f(x_1, x_2, ..., x_n ; \theta)}{g(y_1)}$$

$\because\ f(x;\theta) = e^{2\theta - x}$, Xs are independent

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

$$= e^{2\theta - x_1} \times e^{2\theta - x_2} \times ..... \times e^{2\theta - x_n}$$

$$= e^{\sum(2\theta - x_i)}$$

$$= e^{2n\theta - \sum x_i}$$

$$g(y_1) = n\, f(y_1)(1 - F(y_1))^{n-1}$$

$$f(y_1) = e^{2\theta - y_1}$$

$$F(y_1) = p(X \le y_1) = \int_{2\theta}^{y_1} e^{2\theta - x}\, dx = e^{2\theta} \int_{2\theta}^{y_1} e^{-x}\, dx$$

$$= e^{2\theta} \left(-e^{-x}\right)_{2\theta}^{y_1} = e^{2\theta}\,(e^{-2\theta} - e^{-y_1})$$

$$= 1 - e^{2\theta - y_1}$$

$$\therefore\ g(y_1) = n\left(e^{2\theta - y_1}\right)\left(1 - (1 - e^{2\theta - y_1})\right)^{n-1}$$

$$= n\left(e^{2\theta - y_1}\right)\left(e^{2\theta - y_1}\right)^{n-1}$$

$$= n\, e^{2n\theta - n y_1} \qquad , y_1 \ge 2\theta$$

$$\therefore\ f(x_1, x_2, ..., x_n \mid \hat{\theta} = Y_1) = \frac{e^{2n\theta - \sum x_i}}{n\, e^{2n\theta - n y_1}}$$

$$= \frac{e^{-\sum x_i}}{n\, e^{-n y_1}} = \frac{1}{n} e^{-\sum x_i + n y_1} \quad \text{does not contain } \theta$$

$\therefore\ Y_1$ is suff est for θ

Third Method: Factorization Theorem

Definition 3: Let $\hat{\theta}$ be an estimator for the parameter of $f(x;\theta)$ such that the range does not depend on θ. Then the necessary and sufficient condition for an estimator $\hat{\theta}$ to be sufficient estimator, if there are two non-negative functions, such that:

$$f(x_1, x_2, ..., x_n ; \theta) = g(\hat{\theta}; \theta) \cdot H(x)$$

Theorem:

Let $\hat{\theta}$ be sufficient estimator for the parameter θ, and $u(\hat{\theta})$ be a one-to-one transformation, then $u(\hat{\theta})$ is sufficient estimator for θ.

Note: 1) $\bar{x}$ is one to one transformation to $\sum x_i$. $\Rightarrow\ \sum X_i = n\bar{X}$.

2) If we have more than one parameter, we use factorization theorem (third method) for sufficiency.

Ex: Let $X_1, X_2, \ldots, X_n$ be a rssn from Bernoulli dist^n Ber(θ). Show that $\hat{\theta} = \sum X_i$ is sufficient estimator for the parameter θ.

Sol:

$$\because X \sim Ber(\theta)$$

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

$$f(x_1, x_2, \ldots, x_n; \theta) = \prod_{i=1}^{n} f(x_i;\theta) = \theta^{\sum x_i}(1-\theta)^{n-\sum x_i} \quad \} \times \frac{C^n_{\sum x_i}}{C^n_{\sum x_i}}$$

$$= C^n_{\sum x_i}\, \theta^{\sum x_i}(1-\theta)^{n-\sum x_i} \times \frac{1}{C^n_{\sum x_i}} \quad , \text{ free of } \theta$$

$$= g(\hat{\theta} = \sum x_i; \theta) \times H(x) \qquad \Rightarrow \therefore \hat{\theta} = \sum x_i \text{ is suff est for } \theta$$

Ex: Let $X_1, X_2, \ldots, X_n$ be a rssn from Poisson dist^n Poi(θ), show that $\hat{\theta} = \sum X_i$ is sufficient estimator for θ?

Sol:

$$X \sim Poi(\theta) \quad \Rightarrow f(x;\theta) = \frac{e^{-\theta}\theta^x}{x!} \quad , \because Xs \text{ are independent}$$

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

$$= \frac{e^{-\theta}\theta^{x_1}}{x_1!} \times \frac{e^{-\theta}\theta^{x_2}}{x_2!} \times \ldots\ldots \times \frac{e^{-\theta}\theta^{x_n}}{x_n!}$$

$$= \frac{e^{-n\theta}\theta^{\sum x_i}}{\prod (x_i)!} \qquad \text{joint p.d.f.}$$

$$= e^{-n\theta}\theta^{\sum x_i} \times \frac{1}{\prod (x_i)!}$$

$$= g(\hat{\theta} = \sum x_i; \theta) \times H(x)$$

$$\therefore \hat{\theta} = \sum x_i \text{ is suff est for } \theta.$$

Ex: from Exp($1/\theta$). Is $\sum_{i=1}^{n} X_i$ sufficient estimator for θ? (by factorization theorem).

Sol:

$$f(x_1, x_2, \ldots, x_n; \theta) = g(\hat{\theta}; \theta)\cdot H(x)$$

$$X \sim Exp(1/\theta) \quad \Rightarrow f(x;\theta) = \theta e^{-\theta x} \quad , \because Xs \text{ are independent}$$

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

$$= \theta e^{-\theta x_1} \times \theta e^{-\theta x_2} \times \ldots\ldots \times \theta e^{-\theta x_n}$$

$$= \theta^n e^{-\theta \sum x_i} \times 1$$

$$= g(\hat{\theta} = \sum x_i; \theta) \cdot H(x)$$

$$\therefore \hat{\theta} = \sum X_i \text{ is suff est for } \theta.$$

Multi-Parameters Case (Joint Sufficient Estimator)

Let $X_1, X_2, \ldots, X_n$ be a rss$n$ from a (k) parameters dist^n $f(x;\theta_1,\theta_2,\ldots,\theta_k)$, then $\hat{\theta}_1, \hat{\theta}_2, \ldots, \hat{\theta}_k$ are jointly sufficient estimators for parameters $(\theta_1,\theta_2,\ldots,\theta_k)$ respectively if the j.p.d.f. of $(X_1, X_2, \ldots, X_n)$ can be expressed as:

$$f(x_1, x_2, \ldots, x_n; \theta_1, \theta_2, \ldots, \theta_k) = g(\hat{\theta}_1, \hat{\theta}_2, \ldots, \hat{\theta}_k; \theta_1, \theta_2, \ldots, \theta_k) \cdot H(x)$$

Where; H(x) independent of the parameters $(\theta_1,\theta_2,\ldots,\theta_k)$.

Ex: Let $X_1, X_2, \ldots, X_n$ be a rss$n$ from Gamma dist^n $\Gamma(\alpha, 1/\theta)$, find the jointly sufficient estimators for the parameters (α, θ).

Sol:

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

$$f(x_1, x_2, \ldots, x_n; \alpha, \theta) = \left(\frac{\theta^\alpha}{\Gamma(\alpha)}\right)^n \left(\prod_{i=1}^n x_i\right)^{\alpha-1} e^{-\theta \sum x_i}$$

$$= \left(\frac{\theta^\alpha}{\Gamma(\alpha)}\right)^n \left(\prod_{i=1}^n x_i\right)^{\alpha} e^{-\theta \sum x_i} \times \frac{1}{\prod_{i=1}^n x_i}$$

$$= g(\hat{\alpha} = \prod_{i=1}^n x_i, \hat{\theta} = \sum x_i; \alpha, \theta) \cdot H(x)$$

$$\therefore \hat{\alpha} = \prod_{i=1}^n X_i \text{ and } \hat{\theta} = \sum X_i \text{ are jointly suff est for } \alpha \text{ and } \theta.$$

Ex: Let $X_1, X_2, \ldots, X_n$ be a rss$n$ from normal dist^n $N(\theta, \sigma^2)$, show that $\sum X_i$, $\sum X_i^2$ are the jointly sufficient estimators for the parameters (θ, σ^2) respectively.

Sol: $$f(x;\theta,\sigma^2) = \frac{1}{\sqrt{2\pi\sigma^2}} e^{-\frac{1}{2\sigma^2}(x_i - \theta)^2}$$

$$f(x_1, x_2, \ldots, x_n; \theta, \sigma^2) = \prod_{i=1}^n f(x_i;\theta) = \left(\sqrt{2\pi\sigma^2}\right)^{-n} e^{-\frac{1}{2\sigma^2}\sum(x_i - \theta)^2}$$

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

$$= g(\hat{\theta} = \sum x_i, \hat{\sigma}^2 = \sum x_i^2; \theta, \sigma^2) \cdot H(x)$$

$$\therefore \sum X_i \text{ and } \sum X_i^2 \text{ are jointly suff est for } \theta \text{ and } \sigma^2.$$

Ex: Let $X_1, X_2, \ldots, X_n$ be a rss$n$ from C.U$(\theta_1 - \theta_2, \theta_1 + \theta_2)$, and $Y_1 < Y_2 < \ldots < Y_n$ be the order statistics, show that Y_1 and Y_n are the jointly sufficient estimators for the parameters (θ_1, θ_2) respectively.

Sol:

$$f(x;a,b)=\frac{1}{b-a}$$

$$f(x;\theta_1,\theta_2)=\frac{1}{\theta_1+\theta_2-(\theta_1-\theta_2)}=\frac{1}{2\theta_2} \quad , \ \theta_1-\theta_2<x<\theta_1+\theta_2$$

$$f(y_1)=\frac{1}{2\theta_2}$$

$$F(y_1)=p(Y_1\le y_1)=\int_{\theta_1-\theta_2}^{y_1}\frac{1}{2\theta_2}\,dy_1=\frac{1}{2\theta_2}\,y_1\Big|_{\theta_1-\theta_2}^{y_1}=\frac{1}{2\theta_2}\left(y_1-(\theta_1-\theta_2)\right)$$

$$g(y_i,y_j)=\frac{n!}{(i-1)!(j-i-1)!(n-j)!}\times$$
$$\times\ [F(y_i)]^{i-1}\left[F(y_j)-F(y_i)\right]^{j-i-1}\left[1-F(y_j)\right]^{n-j} f(y_i)\,f(y_j) \quad , \quad a<y_i<y_j<b$$
$$0 \qquad\qquad o.w$$

$$g(y_1,y_n)=n\,(n-1)\left[F(y_j)-F(y_i)\right]^{n-2} f(y_1)\,f(y_n) \quad , \ \theta_1-\theta_2<y_1<y_n<\theta_1+\theta_2$$
$$=n\,(n-1)\left(\frac{1}{2\theta_2}(y_n-(\theta_1-\theta_2))-\frac{1}{2\theta_2}(y_1-(\theta_1-\theta_2))\right)^{n-2}\frac{1}{2\theta_2}\times\frac{1}{2\theta_2}$$
$$=n\,(n-1)\frac{1}{(2\theta_2)^n}\,(y_n-y_1)^{n-2} \quad , \ \theta_1-\theta_2<y_1<y_n<\theta_1+\theta_2$$

$$f(x_1,x_2,\dots,x_n;\theta_1,\theta_2)=\prod_{i=1}^{n} f(x_i;\theta_1,\theta_2)=\left(\frac{1}{2\theta_2}\right)^n$$

$$f(x_1,x_2,\dots,x_n;\theta_1,\theta_2\mid y_1,y_n)=\frac{f(x_1,x_2,\dots,x_n;\theta_1,\theta_2)}{g(y_1,y_n)}$$
$$=\frac{\left(\frac{1}{2\theta_2}\right)^n}{n\,(n-1)\frac{1}{(2\theta_2)^n}\,(y_n-y_1)^{n-2}}$$
$$=\frac{1}{n\,(n-1)\,(y_n-y_1)^{n-2}}=\frac{1}{n\,(n-1)\,(X_{\max}-X_{\min})^{n-2}}$$

$\because$ *which not depend on* θ_1,θ_2

$\therefore$ Y_1 *and* Y_n *are jointly suff est for* θ_1 *and* θ_2 *respectively.*

The Exponential Class of Probability Density Functions

Let X has a p.d.f. $f(x;\theta)$, then the family of $f(x;\theta)$ is belong to exponential class of distribution if it can be expressed as:

$$f(x;\theta)=Exp(\ln f(x;\theta))$$
$$=Exp\left(p(\theta)\,K(x)+S(x)+q(\theta)\right)$$

Such that: p(θ) K(x) must have to be for exponential class.

Ex: Let $X_1, X_2, \dots, X_n$ be a rssn from Bernoulli dist^n Ber(θ), show that if the dist^n of X can be written in exponential form?

Sol:

$$f(x;\theta) = \theta^x (1-\theta)^{1-x}$$
$$= \exp\left(x\ln(\theta) + (1-x)\ln(1-\theta)\right)$$
$$= \exp\left(x\ln\left(\frac{\theta}{1-\theta}\right) + \ln(1-\theta)\right)$$
$$= \exp\left(p(\theta)K(x) + \underset{=0}{S(x)} + q(\theta)\right)$$

$\therefore$ *the family of X is brlongs to the* exp. *class of distribution*

Ex: Let $X_1, X_2, \ldots, X_n$ be a rss$n$ from Poisson dist^n Poi(θ), show that if the dist^n of X can be written in exponential form?

Sol:

$$f(x;\theta) = \frac{e^{-\theta}\theta^x}{x!}$$
$$= \exp\left(\ln f(x;\theta)\right)$$
$$= \exp\left(-\theta + x\ln(\theta) + \ln(x!)\right)$$
$$= \exp\left(q(\theta) + p(\theta)K(x) + S(x)\right)$$

$\therefore$ *the family of X is brlongs to the* exp. *class of distribution*

H.W: Let $X_1, X_2, \ldots, X_n$ be a rss$n$ from exponential dist^n Exp(θ), show that if the exponential dist^n belongs to the exponential family?

H.W: Let $X_1, X_2, \ldots, X_n$ be a rss$n$ from normal dist^n N(0, θ), show that if the normal dist^n belongs to the exponential family?

Theorem

Let $f(x;\theta)$ belongs to exponential class of distributions, then the j.p.d.f. of $(X_1, X_2, \ldots, X_n)$ is:

$$f(x_1, x_2, \ldots, x_n;\theta) = Exp\left(p(\theta)\sum K(x_i) + \sum S(x_i) + n\,q(\theta)\right)$$

Using factorization theorem then the j.p.d.f. can be written as;

$$f(x_1, x_2, \ldots, x_n;\theta) = Exp\left(p(\theta)\sum K(x_i) + n\,q(\theta)\right)\cdot Exp\left(\sum S(x_i)\right)$$

Then we say that $\sum K(X_i)$ is minimal sufficient estimator for θ.

Ex: In a rss$n$. Find minimal sufficient estimators for parameters of:

1) Poisson(θ). **2)** Beta(α, β).

Sol:

1) $Poisson(\theta)$

$$f(x;\theta)=\frac{e^{-\theta}\,\theta^{x}}{x!}=\exp\left(-\,\theta+x\ln(\theta)-\ln(x!)\right)$$

In a rssn;

$$f(x_1,x_2,\dots,x_n;\theta)\ =\exp\left(-\,n\,\theta+\ln(\theta)\sum x_i\ -\sum\ln(x!)\right)$$

$$=\exp\left(-\,n\,\theta+\ln(\theta)\sum x_i\right)\cdot\exp\left(-\sum\ln(x!)\right)$$

$$\therefore\ f(x_1,x_2,\dots,x_n;\theta)=\exp\left(p(\theta)\sum K(x_i)+n\,q(\theta)\right)\cdot\exp\left(\sum S(x_i)\right)$$

$$\Rightarrow\ \sum K(X_i)=\sum X_i\ \ \textit{is}\ \ \textit{minimal}\ \ \textit{suff}\ \ \textit{est}\ \ \textit{for}\ \ \theta$$

2) $Beta(\alpha,\beta)$

$$f(x;\theta)=\frac{\Gamma(\alpha+\beta)}{\Gamma(\alpha)\Gamma(\beta)}\,x^{\alpha-1}(1-x)^{\beta-1}$$

$$=\exp\left(-\ln\beta(\alpha,\beta)+(\alpha-1)\ln(x)\ +(\beta-1)\ln(1-x)\right)$$

$$=\exp\left(-\ln\beta(\alpha,\beta)+\alpha\ln(x)-\ln(x)+\beta\ln(1-x)-\ln(1-x)\right)$$

In a rssn;

$$f(x_1,x_2,\dots,x_n;\theta)\ =\exp\begin{pmatrix}-\,n\ln\beta(\alpha,\beta)+\alpha\sum\ln(x_i)-\sum\ln(x_i)+\\ +\ \beta\sum\ln(1-x_i)-\sum\ln(1-x_i\ )\end{pmatrix}$$

$$=\exp\left(-\,n\ln\beta(\alpha,\beta)+\alpha\sum\ln(x_i)+\beta\sum\ln(1-x_i)\ \right)\times$$

$$\times\exp\left(-\sum\ln(x_i)-\sum\ln(1-x_i\ )\right)$$

$$=\exp\left(p_1(\alpha)\sum K_1(x_i)+\ p_2(\beta)\sum K_2(x_i)+n\,q(\alpha,\ \beta)\right)\times\exp\left(\sum S(x_i)\right)$$

$$\Rightarrow\ \sum K_1(X_i)=\sum\ln(X_i)\ \ \textit{and}\ \ \sum K_2(X_i)=\sum\ln(1-\ X_i)$$

are min *imal* *suff* *est* *for* α *and* β *respective ly*

Ex: In a rss*n*. Find minimal sufficient estimators for θ, σ^2 from N(θ, σ^2).

Sol:

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

$$f(x;\theta,\sigma^2)=\frac{1}{\sqrt{2\pi\sigma^2}}\ e^{-\frac{1}{2\sigma^2}(x-\theta)^2}\qquad,\ -\infty<x<\infty$$

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

$$=\exp\left(-\frac{1}{2}\ln(2\pi\sigma^2)-\frac{(x_i-\theta)^2}{2\sigma^2}\right)$$

$$=\exp\left(-\frac{1}{2}\ln(2\pi\sigma^2)-\frac{x_i^2}{2\sigma^2}+\frac{\theta\,x}{\sigma^2}-\frac{\theta^2}{2\sigma^2}\right)$$

$$f(x;\ \theta,\sigma^2)=\exp\left(p_1(\theta)\,K_1(x)+\ p_2(\sigma^2)\,K_2(x)+q(\theta,\sigma^2)+\ S(x)\right)$$

In a rssn;

$$f(x_1,x_2,\ldots,x_n;\theta,\sigma^2) = \exp\left(-\frac{n}{2}\ln(2\pi\sigma^2) - \frac{\sum x_i^2}{2\sigma^2} + \frac{\theta\sum x_i}{\sigma^2} - \frac{n\theta^2}{2\sigma^2}\right)$$

$$= \exp\left(p_1(\theta)\sum K_1(x_i) + p_2(\sigma^2)\sum K_2(x_i) + n\,q(\theta,\sigma^2)\right)\times \exp\left(\sum S(x_i)\right)$$

$$\Rightarrow \therefore\ \sum X_i \text{ and } \sum X_i^2 \text{ are minimal jointly suff est for } \theta,\sigma^2 \text{ respectively.}$$

4. Completeness

Let $f(x\,;\theta)$ denote a family of probability density function, **let u(x) be a continuous function of (X), then if [E{u(X)}= 0] implies (u(x) = 0) at each point of (X)**, we say that the family of p.d.f. is complete.

Note: If the range depends on θ, then we use the general rule to derivative of integral;

$$Let;\quad G(\theta) = \int_{a(\theta)}^{b(\theta)} f(x;\theta)\,dx \qquad ,\ where\quad f: is\ any\ function$$

$$\frac{\partial G(\theta)}{\partial\theta} = \int_{a(\theta)}^{b(\theta)} \frac{\partial f(x;\theta)}{\partial\theta}\,dx + f\left(b(\theta),\theta\right)\times b'(\theta) - f\left(a(\theta),\theta\right)\times a'(\theta)$$

Ex: Let X be a random variable from; **1)** Bernoulli dist^n. **2)** Poisson dist^n. **3)** Normal dist^n. Show that the family of X is complete.

Sol: 1)

1) $X \sim Ber(\theta)$

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

Let $u(x)$ *be a continuous fun of* X*. then;*

$E\left(u(X)\right) = 0$

$$E\left(u(X)\right) = \sum_{x=0}^{1} u(x)\ f(u;\theta) = 0$$

$$= u(0)\theta^0(1-\theta)^{1-0} + u(1)\theta^1(1-\theta)^{1-1} = 0$$

$$= u(0)(1-\theta) + u(1)\theta = 0$$

$\because \theta \neq 0$

$\therefore u(0) = u(1) = 0 \quad \Rightarrow u(x) = 0\ \ \forall x$

$\therefore$ *the family of* X *is complete*

5) Uniqueness Estimator (M.V.U.E)

Th: Let $X_1, X_2, \ldots, X_n$ be a rss$n$ from a dist^n with p.d.f. $f(x\,;\theta)$, let Y_1 be a **sufficient** estimator for θ, and let g(y_1; θ) be **complete** if there is a continuous function of Y_1 which is an **unbiased** estimator for θ, $\phi(\theta)$ such that $E\left(\phi(\theta)\right) = \theta$, **then** $\phi(\theta)$ **is the unique best estimator for θ (M.V.U.E).**

Note: If an estimator does not complete then we do not find the unique and if have complete then we find a unique estimator.

Ex: Let X be a r.v. with p.d.f.;

$$f(x;\theta) = \frac{1}{2\theta} \quad , \quad -\theta < x < \theta \quad , \quad \theta > 0$$

Show that $f(x;\theta)$ is not complete? If it is then find the unique estimator for θ.

Sol:

$\because$ *the range depend on* θ.

Let $u(x)$ *be a continuous fun of* X. *then* ;

$$E(u(X)) = 0 \quad , \quad u(x) = 0$$

$$E(u(X)) = \int_{-\theta}^{\theta} u(x)\, f(x;\theta)\, dx = 0$$

$$\Rightarrow \int_{-\theta}^{\theta} u(x) \frac{1}{2\theta} dx = 0 \qquad \} \times 2\theta$$

$$\Rightarrow \int_{-\theta}^{\theta} u(x)\, dx = 0$$

Let ; $G(\theta) = \int_{a(\theta)}^{b(\theta)} f(x;\theta)\, dx$, *where* f : *is any function*

$$\frac{\partial G(\theta)}{\partial \theta} = \int_{a(\theta)}^{b(\theta)} \frac{\partial f(x;\theta)}{\partial \theta} dx + f(b'(\theta),\theta) \times b'(\theta) - f(a'(\theta),\theta) \times a'(\theta)$$

$$\Rightarrow \frac{\partial}{\partial \theta} \int_{-\theta}^{\theta} u(x)\, dx = 0$$

$$\Rightarrow \int_{-\theta}^{\theta} \frac{\partial\, u(x)}{\partial \theta} dx + u(\theta)(1) - u(-\theta)(-1) = 0$$

$$\Rightarrow \int_{-\theta}^{\theta} (0)\, dx + u(\theta) + u(-\theta) = 0$$

$$\Rightarrow u(\theta) + u(-\theta) = 0$$

If $u(\theta) = -u(-\theta)$ *is odd function*

If $u(\theta) = u(-\theta)$ *is even function*

$\because$ $u(\theta) \neq 0$

$\therefore$ $f(x;\theta)$ *is not complete*

$\because$ $f(x;\theta)$ *is not complete then there isn't has the unique estimator*

Ex: Let $X_1, X_2, \ldots, X_n$ is a rss$n$ from Gamma dist^n $\Gamma(4, \theta)$, $0 < \theta < \infty$. **1)** Show that $Y = \sum X_i$ is a complete sufficient estimator for θ. **2)** Find the unique continuous function of Y which is the best estimator for θ (M.V.U.E).

Sol:

$$\because X \sim \Gamma(4,\theta) \Rightarrow f(x;\theta) = \frac{1}{\Gamma(4)\theta^4} x^{4-1} e^{-x/\theta} \quad , x > 0 \quad , \theta > 0$$

$$= \frac{1}{6\theta^4} x^3 e^{-x/\theta}$$

$\because$ *the range of X does not depend on θ, then we use exponential family to prove suff.*

$$f(x;\theta) = \exp\left(\ln\left\{\frac{1}{6\theta^4} x^3 e^{-x/\theta}\right\}\right) = \exp\left(\ln(1) - \ln(6) - 4\ln(\theta) + 3\ln(x) - \frac{x}{\theta}\right)$$

$$= \exp\left(-\ln(6) - 4\ln(\theta) + 3\ln(x) - \frac{x}{\theta}\right)$$

In arssn

$$f(x_1,...,x_n;\theta) = \exp\left(-n\ln(6) - 4n\ln(\theta) + 3\sum \ln(x_i) - \frac{\sum x_i}{\theta}\right) \qquad , 3\sum \ln(x_i) = \ln\left(\prod_{i=1}^{n} x_i^3\right)$$

$$\therefore f(x_1,...,x_n;\theta) = \exp\left(n\,q(\theta) + p(\theta)\sum K(x_i) + \sum S(x_i)\right)$$

$$\sum K(x_i) = \sum X_i \ , \ p(\theta) = -\frac{1}{\theta} \ , \ \sum S(x_i) = \ln\left(\prod_{i=1}^{n} x_i^3\right) \ , \ n\,q(\theta) = -n\ln(6) - 4n\ln(\theta)$$

$\therefore Y = \sum K(x_i) = \sum X_i$ *is sufficient estimator for θ.*

$$\because X \sim \Gamma(4,\theta) \Rightarrow Y = \sum X_i \sim \Gamma(4n,\theta) \quad , g(y;\theta) = \frac{1}{\Gamma(4n)\,\theta^{4n}} y^{4n-1} e^{-y/\theta} \quad , \quad y > 0$$

Let $u(y)$ be a continuous fun of Y. then ;

$$E(u(Y)) = 0 \qquad , u(y) = 0$$

$$E(u(Y)) = \int_0^\infty u(y)\, g(y;\theta)\, dy = 0$$

$$\Rightarrow \int_0^\infty u(y) \frac{1}{\Gamma(4n)\,\theta^{4n}} y^{4n-1} e^{-y/\theta}\, dy = 0$$

$$\Rightarrow \frac{1}{\Gamma(4n)\,\theta^{4n}} \int_0^\infty u(y)\ y^{4n-1} e^{-y/\theta}\, dy = 0 \quad \} \times \Gamma(4n)\,\theta^{4n}$$

$$\Rightarrow \int_0^\infty u(y)\ y^{4n-1} e^{-y/\theta}\, dy = 0$$

$y^{4n-1} \neq 0$ *(never)* $\quad , e^{-y/\theta} \neq 0 \quad , \Rightarrow \therefore u(y) = 0$

$g(y;\theta)$ *is complete* .

$X \sim \Gamma(\alpha,\beta) \Rightarrow E(X) = \alpha\beta \quad , V(X) = \alpha\beta^2$

$X \sim \Gamma(4,\theta) \Rightarrow E(X) = 4\theta \quad , V(X) = 4\theta^2$

$Y = \sum X_i \sim \Gamma(4n,\theta)$

$E(Y) = 4n\theta \qquad \} \div 4n$

$E\left(\frac{Y}{4n}\right) = \theta \qquad \Rightarrow \therefore \hat{\theta} = \frac{Y}{4n}$ *is M.V.U.E. for θ.*

Ex: (Functions of Parameter): Let $X_1, X_2, \ldots, X_n$ denote a random sample from a dist^n which is Ber(1, θ), find the best estimator for the variance $n\theta(1-\theta)$ of $Y=\sum X_i$ (M.V.U.E).

Sol:

$$X \sim Ber(1,\theta)$$

$$Y = \sum X_i \sim Bin(n,\theta)$$

$$E(Y) = E\left(\sum X_i\right) = nE(X) = n\theta$$

$$\Rightarrow E\left(\frac{Y}{n}\right) = \theta \qquad \Rightarrow \therefore \hat{\theta} = \frac{Y}{n} \text{ is M.V.U.E. for } \theta.$$

But the required is $V(Y) = n\theta(1-\theta)$

$$E\left(V(Y)\right) = E\left(n\frac{Y}{n}\left\{1-\frac{Y}{n}\right\}\right) = E\left(Y-\frac{Y^2}{n}\right) = E(Y) - \frac{E(Y^2)}{n}$$

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

$$= n\theta - \frac{n\theta - n\theta^2 + n^2\theta^2}{n} = \frac{n^2\theta - n\theta + n\theta^2 - n^2\theta^2}{n}$$

$$= \frac{n\theta(n-1+\theta-n\theta)}{n} = \frac{n\theta(n-1-\theta(n-1))}{n}$$

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

$$\therefore E\left(Y\left\{1-\frac{Y}{n}\right\}\right) = n\theta(1-\theta)\frac{(n-1)}{n} \qquad \} \times \frac{n}{n-1}$$

$$E\left(\frac{nY\left(1-\frac{Y}{n}\right)}{n-1}\right) = n\theta(1-\theta)$$

$$\Rightarrow n\hat{\theta}(1-\hat{\theta}) = \frac{nY\left(1-\frac{Y}{n}\right)}{n-1} \text{ is M.V.U.E. for } Y = \sum X_i$$

The Rao- Cramer Inequality

Let $X_1, X_2, \ldots, X_n$ be a rssn from a dist^n with p.d.f. $f(x;\theta)$, and let $T = u(X_1, X_2, \ldots, X_n)$ be an unbiased estimator for $\phi(\theta)$, then the variance of T satisfies the inequality;

$$V(T) \geq \frac{(\phi'(\theta))^2}{nE\left(\frac{\partial \ln f(x;\theta)}{\partial\theta}\right)^2} = \frac{(\phi'(\theta))^2}{Var(S)} = \frac{(\phi'(\theta))^2}{-nE\left(\frac{\partial^2 \ln f(x;\theta)}{\partial\theta^2}\right)}$$

Notes:

1)

$\frac{(\phi'(\theta))^2}{V(S)}$ *is called Rao – Cramer Lower Bound (RCLB) (Minimum variance bound (MVB)*

2) If T unbiased estimator for θ, E(T) = θ;

$\phi(\theta) = \theta \quad \rightarrow \quad \phi'(\theta) = 1$

$$\therefore \left(RCLB = \frac{1}{V(S)} \right)$$

3) In normal distribution case, we apply the second law is easier.
4) We do not use (n) in case using the likelihood function in law.

6. Efficient Estimator

Def^n: The ratio of the RCLB to the actual variance of any unbiased estimator for θ is called the efficiency;

$$eff = \frac{RCLB}{V(T)} \quad , \quad 0 \le eff \le 1$$

$$if \;\; eff = 1 \quad \Rightarrow T \text{ is called efficient estimator for } \theta.$$

Def^n: Let T be an unbiased estimator for $\phi(\theta)$, then we say that T is an efficient estimator for θ iff;

$$V(T) = RCLB$$

Ex: Let $X_1, X_2, \ldots, X_n$ be a rss$n$ from Poison dist^n Poi(θ), if $T = \overline{X}$ is an efficient estimator for $\phi(\theta) = \theta$.

Sol:

$$X \sim Poi(\theta)$$

$$f(x;\theta) = \frac{e^{-\theta}\theta^{x}}{x!} \quad , \quad x = 0, 1, \ldots\ldots$$

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

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

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

$$-E\left(\frac{\partial^2 \ln f(x;\theta)}{\partial \theta^2}\right) = \frac{E(X)}{\theta^2} = \frac{\theta}{\theta^2} = \frac{1}{\theta}$$

$$-nE\left(\frac{\partial^2 \ln f(x;\theta)}{\partial \theta^2}\right) = \frac{n}{\theta} = V(S)$$

$$\phi(\theta) = \theta \quad \rightarrow \quad \phi'(\theta) = 1$$

$$RCLB = \frac{(\phi'(\theta))^2}{V(S)} = \frac{1}{\frac{n}{\theta}} = \frac{\theta}{n}$$

$$V(T) = V(\overline{X})$$

$$V(\overline{X}) = \frac{V(X)}{n} = \frac{\theta}{n}$$

$$\therefore RCLB = V(\overline{X}) \quad \Rightarrow \quad eff = 1$$

$$\therefore \overline{X} \text{ is an efficient estimator for } \phi(\theta).$$

Ex: In a rss$n$ from N(θ, σ^2). Show that;

1) If $T=\overline{X}$ is an efficient estimator for $\phi(\theta)=\theta$.

2) $S^2=\dfrac{\sum(x_i-\overline{x})^2}{n}$ or $S^2=\dfrac{\sum(x_i-\overline{x})^2}{n-1}$ is an efficient estimator for $\phi(\sigma^2)=\sigma^2$.

Sol:

1) $X\sim N(\theta,\sigma^2)$

$$f(x;\theta,\sigma^2)=\frac{1}{\sqrt{2\pi\sigma^2}}\,e^{-\frac{1}{2\sigma^2}(x-\theta)^2}\quad,\ -\infty<x<\infty$$

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

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

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

$$-n\,E\left(\frac{\partial^2\ln f(x;\theta,\sigma^2)}{\partial\theta^2}\right)=\frac{n}{\sigma^2}=V(S)=F.I$$

$$\phi(\theta)=\theta\ \rightarrow\ \phi'(\theta)=1$$

$$RCLB=\frac{(\phi'(\theta))^2}{V(S)}=\frac{1}{n/\sigma^2}=\frac{\sigma^2}{n}$$

$$V(T)=V(\overline{X})$$

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

$$\therefore RCLB=V(\overline{X})=\frac{\sigma^2}{n}\quad,\ \Rightarrow eff=1$$

$\therefore \overline{X}$ *is an efficient estimator for* $\phi(\theta)=\theta$.

Mean Square Error (MSE)

One way of measuring the accuracy of an estimator is via its mean square error. The mean square error of an estimator $\hat{\theta}$ is defined as:

$$MSE(\hat{\theta})=E(\hat{\theta}-\theta)^2=Var(\hat{\theta})+b^2(\hat{\theta})$$

Note: If $\hat{\theta}$ is unbiased estimator for θ then; $MSE(\hat{\theta})=Var(\hat{\theta})$

Relative Efficient Estimator

Let $\hat{\theta}_1$ and $\hat{\theta}_2$ be two estimators for parameter θ of $f(x\,;\,\theta)$, the relative efficient of $\hat{\theta}_1$ relative to $\hat{\theta}_2$ is given by:

$$R.Eff\,.(\hat{\theta}_1 \mid \hat{\theta}_2) = \frac{MSE\ (\hat{\theta}_1)}{MSE\ (\hat{\theta}_2)} < 1$$

$i.e.,\quad MSE\ (\hat{\theta}_1) < MSE\ (\hat{\theta}_2)$

$\therefore\ \hat{\theta}_1$ *is more efficient than* $\hat{\theta}_2$.

Ex: In a rss2 from Bernoulli dist^n Ber(θ), let $T_1 = X_1$ and $T_2 = \frac{\sum X_i}{n+1}$ be two estimators for parameter θ, show that which of them more efficient.

Sol:

$E(T_1) = E(X_1) = \theta \quad unbiased$

$$E(T_2) = E\left(\frac{\sum X_i}{n+1}\right) = \frac{n}{n+1}\,E(X) \quad , \ when \quad n = 2$$

$$E(T_2) = \frac{2}{3}\,\theta \quad biased$$

$$b(T_2) = E(T_2) - \theta = \frac{2}{3}\,\theta - \theta = \frac{-\,\theta}{3}$$

$V(T_1) = Var(X_1) = \theta(1-\theta) = MSE(T_1)$

$$V(T_2) = \frac{1}{(n+1)^2} Var(\sum X_i) = \frac{1}{(n+1)^2}\ n\,Var(X) = \frac{2}{9}\theta(1-\theta) = \frac{2\theta - 2\theta^2}{9}$$

$MSE(T_2) = V(T_2) + b^2(T_2)$

$$= \frac{2\theta - 2\theta^2}{9} + \frac{\theta^2}{9} = \frac{2\theta - \theta^2}{9}$$

$$\because\ \frac{2\theta - \theta^2}{9} < \theta(1-\theta)$$

$\therefore MSE(T_2) < MSE(T_1)$

$\Rightarrow\ \therefore\ T_2$ *is more efficient than* T_1.