# Statistical Inference

**Department of Statistics & Informative**

**Fourth Stage**

**First Semester (2022-2023)**

**Assistant Professor**

**Dr. *Luceen Immanuel Kework***

# 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 ($\theta$)

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

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

### 3) Binomial Distribution X ~ Bin ($n$, $\theta$)

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

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

### 4) Negative Binomial Distribution X ~ N.Bin ($r$, $\theta$)

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

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

### 5) Geometric Distribution X ~ Geo ($\theta$)

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

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

## 6) The Poisson Distribution X ~ Poi ($\theta$)

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

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

# Second: The Continuous Distributions

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

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

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

## 2. Beta Distribution X ~ *Beta* (α, $\beta$)

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

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

## 3. Gamma Distribution

## a) Gamma, Distribution

**1.** X ~ Γ(α, $\theta$)

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

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

**2.** X ~ Γ(α, 1/$\theta$)

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

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

## b) Exponential Distribution

**1.** X ~ Exp($\theta$)

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

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

**2.** X ~ Exp(1/$\theta$)

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

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

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

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

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

## 4. Normal (Gaussian) Distribution X ~ N($\theta$, $\sigma^2$)

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

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

# 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;

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

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

**Sol:**

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

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

$$\because Y = 4X$$

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

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

**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:**

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

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

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

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 $\theta_1$ , $\theta_2$ 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:**

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

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

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

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

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

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

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

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

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

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

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

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

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

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

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

**Ex:** Let $X_1$ and $X_2$ have a joint p.d.f as follows;

$$f(x_1,x_2)=\begin{cases}\left(\frac{2}{3}\right)^{x_1+x_2}\left(\frac{1}{3}\right)^{2-x_1-x_2} & , \quad (x_1,x_2)=(0,0),(0,1),(1,0),(1,1)\\ 0 & o.w\end{cases}$$

**Find** the joint p.d.f of $Y_1 = X_1 - X_2$ and $Y_2 = X_1 + X_2$ and the marginal p.d.f of $Y_1$ and $Y_2$.

**Sol:**

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

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

*because*; $y_1 = x_1 - x_2$ *and* $y_2 = x_1 + x_2$

$\because y_1 = x_1 - x_2$

$\underline{y_2 = x_1 + x_2}$

$y_1 + y_2 = 2x_1 \quad \Rightarrow \quad x_1 = \frac{y_1+y_2}{2}\ (summation) \quad \Rightarrow \quad \therefore x_2 = \frac{y_2-y_1}{2} \quad (subtraction)$

$\Rightarrow x_1 + x_2 = \frac{y_1+y_2}{2} + \frac{y_2-y_1}{2} = \frac{2y_2}{2} = y_2$

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

$$f(y_1,y_2)=p(Y_1=y_1\ ,Y_2=y_2)=p(X_1=\frac{y_1+y_2}{2}\ ,X_2=\frac{y_2-y_1}{2})$$

$$f(y_1,y_2)=\begin{cases}\left(\frac{2}{3}\right)^{y_2}\left(\frac{1}{3}\right)^{2-y_2} & ,\ (y_1,y_2)\in B\\ 0 & o.w\end{cases}$$

The marginal p.d.f. of $Y_1$ and $Y_2$ are given by;

| $y_2$ \ $y_1$ | -1 | 0 | 1 | $g(y_2)$ |
|---|---|---|---|---|
| 0 | 0 | 1/9 | 0 | 1/9 |
| 1 | 2/9 | 0 | 2/9 | 4/9 |
| 2 | 0 | 4/9 | 0 | 4/9 |
| $g(y_1)$ | 2/9 | 5/9 | 2/9 | 1 |

## Second: Transformations of the Continuous Random Variables

**Definition:** Let X be continuous r.v. having a p.d.f of $f(x)$. Let $A$ be the one-dimension space $A=\{x:\ x\in R(x)\}$, where $f(x) > 0$. Consider the r.v. $Y = g(X)$, where $y = g(x)$ define a one-to-one transformation that maps the set $A$ onto the set $B$. Let the inverse of $y = g(x)$ be denoted by $x = w(y)$, then;

$$y = g(x) \quad \Rightarrow \quad x = w(y)$$

$$|w'(y)| = \left|\frac{dx}{dy}\right| = |J| \qquad \textit{is called the Jacobian}$$

Then the p.d.f. of the r.v. $Y = g(X)$ is given by;

$$f(y)=\begin{cases}f\left(w(y)\right)|J| & ,\quad y\in B\\ 0 & o.w\end{cases}$$

When we have two r.v.'s $X_1$ and $X_2$. 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 (two dimensional) set *A* in the $x_1x_2$-plane onto a (two dimensional) set *B* in the $y_1y_2$-plane. Then;

$$y_1 = g_1(x_1, x_2) \quad \Rightarrow \quad x_1 = w_1(y_1, y_2)$$
$$y_2 = g_2(x_1, x_2) \quad \Rightarrow \quad x_2 = w_2(y_1, y_2)$$

Then the j.p.d.f. of the r.v. $Y_1$ and $Y_2$ is given by;

$$f(y_1, y_2) = \begin{cases} f\big(w_1(y_1, y_2), w_2(y_1, y_2)\big)|J| & , \quad (y_1, y_2) \in B \\ 0 & o.w \end{cases}$$

Then the Jacobian |J| is the determinant of order (2); $|J| = \begin{vmatrix} \frac{\partial x_1}{\partial y_1} & \frac{\partial x_1}{\partial y_2} \\ \frac{\partial x_2}{\partial y_1} & \frac{\partial x_2}{\partial y_2} \end{vmatrix}$

**Ex:** Let X have the p.d.f.;

$f(x) = \begin{cases} 1 & , \quad 0 < x < 1 \\ 0 & o.w \end{cases}$ , where $Y = -2\ln X$, **find** the p.d.f. of Y.

**Sol.:**

$$A = \{x : 0 < x < 1\} \quad f(x) > 0$$
$$B = \{y : \ 0 < y < \infty\} \quad f(y) > 0 \qquad , \quad \ln(0) = -\infty$$
$$when \quad y = -2\ln x \quad \Rightarrow -\frac{y}{2} = \ln x \quad \Rightarrow e^{-y/2} = e^{\ln x} \quad \Rightarrow x = e^{-y/2}$$
$$|J| = \left|\frac{dx}{dy}\right| = \left|-\frac{1}{2}e^{-y/2}\right| = \frac{1}{2}e^{-y/2}$$
$$f(y) = f\big(w(y)\big)\,|J| \qquad ,$$
$$\therefore f(y) = \begin{cases} (1)\frac{1}{2}e^{-y/2} & , \quad 0 < y < \infty \\ 0 & o.w \end{cases} = \begin{cases} \frac{1}{2}e^{-y/2} & , \quad 0 < y < \infty \\ 0 & o.w \end{cases}$$

**Ex:** Let X is a uniform random variable on the interval (-2, 2), find the p.d.f. of Y;

**1.** $Y = 4X + 3$. **2.** $Y = |X|$.

**Sol:**

$$1) \ \because X \sim C.U(-2,2) \quad \Rightarrow f(x) = \begin{cases} \frac{1}{4} & , -2 < x < 2 \\ 0 & o.w \end{cases}$$
$$A = \{x : -2 < x < 2\} \quad f(x) > 0$$
$$\because \quad y = 4x + 3$$
$$B = \{y : \ -5 < y < 11\} \quad f(y) > 0$$
$$when \ y = 4x + 3 \Rightarrow x = \frac{y-3}{4} \qquad \Rightarrow |J| = \left|\frac{dx}{dy}\right| = \frac{1}{4}$$

$f(y) = f(w(y))\,|J|$ ,

$$\therefore f(y) = \begin{cases} \frac{1}{4} \times \frac{1}{4} & , -5 < y < 11 \\ 0 & o.w \end{cases} = \begin{cases} \frac{1}{16} & , -5 < y < 11 \\ 0 & o.w \end{cases}$$

2) $A = \{x : -2 \le x \le 2\}$ $f(x) > 0$

$\because \quad Y = |X|$

$B = \{y : \ 0 \le y \le 2\}$ $\quad f(y) > 0$

$$Y = |X| \Rightarrow y = \begin{cases} -x & , \ x < 0 \\ x & , \ x \ge 0 \\ 0 & o.w \end{cases}$$

$x = -y = (w_1(y))$ *and* $x = y = (w_2(y))$ , $|J_1| = |-1| = 1$ , $|J_2| = |1| = 1$

$$f(y) = \underset{x<0}{f(w_1(y))\,|J_1|} + \underset{x \ge 0}{f(w_2(y))\,|J_2|}$$

$$\therefore f(y) = \begin{cases} \frac{1}{4} \times 1 + \frac{1}{4} \times 1 & , \ 0 \le y \le 2 \\ 0 & o.w \end{cases} = \begin{cases} \frac{1}{2} & , \ 0 \le y \le 2 \\ 0 & o.w \end{cases}$$

**Ex:** Let X ~ Γ(*r*/2, *θ*), Y = $\frac{2X}{\theta}$, find; the p.d.f. of Y.

**Sol**

$\because$ X ~ Γ(*r*/2, *θ*) $\qquad \Rightarrow f(x, r/2, \theta) = \dfrac{1}{\Gamma(\frac{r}{2})\,\theta^{r/2}}\, x^{\frac{r}{2}-1}\, e^{-x/\theta} \qquad 0 < x < \infty$

$A = \{x : \ 0 < x < \infty\}$

$\because Y = \dfrac{2X}{\theta}$

$B = \{y : \ 0 < y < \infty)$

$y = \dfrac{2x}{\theta} \Rightarrow x = \dfrac{\theta y}{2}$ , $\quad |J| = \left|\dfrac{dx}{dy}\right| = \dfrac{\theta}{2}$

$f(y) = f\big(w(y)\big)|J|$

$$= \frac{1}{\Gamma(r/2)\,\theta^{r/2}} \left(\frac{\theta y}{2}\right)^{\frac{r}{2}-1} e^{-\frac{\frac{\theta y}{2}}{\theta}} \times \frac{\theta}{2}$$

$$= \frac{1}{\Gamma(r/2)\,\theta^{r/2}} \left(\frac{\theta y}{2}\right)^{\frac{r}{2}-1} e^{-y/2} \times \frac{\theta}{2}$$

$$= \frac{1}{\Gamma(r/2)\,\theta^{r/2}} \left(\frac{\theta}{2}\right)^{\frac{r}{2}-1} y^{\frac{r}{2}-1} e^{-y/2} \left(\frac{\theta}{2}\right)^{1} = \frac{1}{\Gamma(r/2)\,\theta^{r/2}} \left(\frac{\theta}{2}\right)^{\frac{r}{2}} y^{\frac{r}{2}-1} e^{-y/2}$$

$$\therefore f(y) = \begin{cases} \dfrac{1}{\Gamma(r/2)\,2^{r/2}} (y)^{\frac{r}{2}-1} e^{-y/2} & , \ 0 < y < \infty \\ 0 & o.w \end{cases} \quad , \ Y \sim \chi^2_{(r)}$$

**Ex:** Let X have the p.d.f.;

$$f(x)=\begin{cases}2xe^{-x^2} & , \quad 0<x<\infty \\ 0 & o.w\end{cases}$$

, where $Y = X^2$, **find** the p.d.f. of Y.

**Sol.:**

$A=\{x:\ 0<x<\infty\}\quad f(x)>0$

$B=\{y:\ 0<y<\infty\}\quad g(y)>0$

$when\quad y=x^2 \quad\Rightarrow\quad x=\mp\sqrt{y}=y^{1/2}\quad (only)(one-to-one\ Transformation)$

$$|J|=\left|\frac{dx}{dy}\right|=\frac{1}{2}y^{-1/2}=\frac{1}{2y^{1/2}}$$

$$f(y)=f(w(y))|J|=2y^{1/2}e^{-y}\,\frac{1}{2y^{1/2}}=e^{-y}$$

$$\therefore\ f(y)=\begin{cases}e^{-y} & , \quad 0<y<\infty \\ 0 & o.w\end{cases}$$

**Ex:** Let $X_1$ and $X_2$ be two stochastically independent r.v.'s, which have gamma distribution, with parameters $(\alpha, \theta)$ and $(\beta, \theta)$ respectively, and $Y_1 = X_1 + X_2$ , $Y_2=\dfrac{X_1}{X_1+X_2}$ , find the j.p.d.f. of $Y_1$ and $Y_2$, $f(y_1,y_2)$ , $f(y_1)$ and $f(y_2)$ .

**Sol.:**

$$X_1\sim\Gamma(\alpha,\theta)\quad\Rightarrow\quad f(x_1)=\begin{cases}\dfrac{1}{\Gamma(\alpha)\ \theta^{\alpha}}x_1^{\alpha-1}e^{-\frac{x_1}{\theta}} & , \quad 0<x_1<\infty \\ 0 & o.w\end{cases}$$

$$X_2\sim\Gamma(\beta,\theta)\quad\Rightarrow\quad f(x_2)=\begin{cases}\dfrac{1}{\Gamma(\beta)\ \theta^{\beta}}x_2^{\beta-1}e^{-\frac{x_2}{\theta}} & , \quad 0<x_2<\infty \\ 0 & o.w\end{cases}$$

$\because\ X_1\ and\ X_2\ are\ independent$

$f(x_1,x_2)=f(x_1)\,f(x_2)$

$$f(x_1,x_2)=\begin{cases}\dfrac{1}{\Gamma(\alpha)\,\Gamma(\beta)\,\theta^{\alpha+\beta}}x_1^{\alpha-1}\ x_2^{\beta-1}e^{-\frac{x_1+x_2}{\theta}} & , \quad 0<x_1,x_2<\infty \\ 0 & o.w\end{cases}$$

$A=\{(x_1,x_2):\ 0<x_1,x_2<\infty\}\qquad , f(x_1,x_2)>0$

$B=\{(y_1,y_2):\ 0<y_1<\infty\ and\ \ 0<y_2<1,\ \}\quad , f(y_1,y_2)>0$

$y_1=x_1+x_2\quad\Rightarrow x_2=y_1-x_1 ....(1)\quad\Rightarrow x_1=y_1y_2\ ...(2)$

$\therefore\ x_2=y_1-y_1y_2=y_1(1-y_2)$

$$|J|=\begin{vmatrix}\frac{dx_1}{dy_1} & \frac{dx_1}{dy_2}\\ \frac{dx_2}{dy_1} & \frac{dx_2}{dy_2}\end{vmatrix}=\begin{vmatrix}y_2 & y_1\\ 1-y_2 & -y_1\end{vmatrix}=|-y_1|=y_1$$

$$f(y_1,y_2) = f\big(w_1(y_1,y_2),\, w_2(y_1,\, y_2)\big)\left|J\right|$$

$$= \frac{1}{\Gamma(\alpha)\,\Gamma(\beta)\,\theta^{\alpha+\beta}}\left(y_1\, y_2\right)^{\alpha-1}\left(y_1(1-y_2)\right)^{\beta-1} e^{-\frac{(y_1 y_2+(y_1(1-y_2))}{\theta}} \times y_1$$

$$\therefore\ f(y_1,y_2) = \begin{cases} \frac{1}{\Gamma(\alpha)\,\Gamma(\beta)\,\theta^{\alpha+\beta}}\, y_1^{\alpha+\beta-1}\, y_2^{\alpha-1}\left(1-y_2\right)^{\beta-1} e^{-y_{1/\theta}} & ,\ (y_1,y_2)\in B \\ 0 & o.w \end{cases}$$

$$f(y_1) = \int_{R_{y_2}} f(y_1,y_2)\, dy_2$$

$$= \frac{y_1^{\alpha+\beta-1}\, e^{-\frac{y_1}{\theta}}}{\Gamma(\alpha)\,\Gamma(\beta)\,\theta^{\alpha+\beta}} \int_0^1 y_2^{\alpha-1}\,(1-y_2)^{\beta-1} dy_2$$

$$= \frac{y_1^{\alpha+\beta-1}\, e^{-\frac{y_1}{\theta}}}{\Gamma(\alpha)\,\Gamma(\beta)\,\theta^{\alpha+\beta}}\ \frac{\Gamma(\alpha)\,\Gamma(\beta)}{\Gamma(\alpha+\beta)} = \frac{1}{\Gamma(\alpha+\beta)\ \theta^{(\alpha+\beta)}}\ \ y_1^{(\alpha+\beta)-1}\, e^{-\frac{y_1}{\theta}}\ \ ,\ 0 < y_1 < \infty$$

$$\therefore\ Y_1 \sim \Gamma(\alpha+\beta,\, \theta)$$

$$f(y_2) = \int_{R_{y_1}} f(y_1,y_2)\, dy_1 = \frac{y_2^{\alpha-1}\,(1-y_2)^{\beta-1}}{\Gamma(\alpha)\,\Gamma(\beta)\,\theta^{\alpha+\beta}} \int_0^{\infty} y_1^{(\alpha+\beta)-1}\, e^{-\frac{y_1}{\theta}} dy_1$$

$$= \frac{y_2^{\alpha-1}\,(1-y_2)^{\beta-1}}{\Gamma(\alpha)\,\Gamma(\beta)\,\theta^{\alpha+\beta}}\ \Gamma(\alpha+\beta)\,\theta^{\alpha+\beta} = \frac{\Gamma(\alpha+\beta)}{\Gamma(\alpha)\,\Gamma(\beta)}\ y_2^{\alpha-1}\left(1-y_2\right)^{\beta-1}\ \ ,\ 0< y_2 <1$$

$$\therefore\ Y_2 \sim Beta(\alpha\,,\beta)$$

## 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;**

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

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

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

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

**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)

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

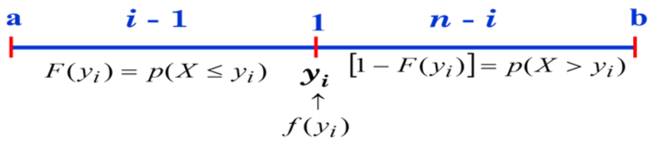
**Note: Always J equal to one because be continuous**

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

## 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:

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

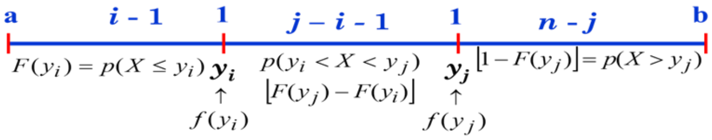
$$g(y_n) = \frac{n(n-1)!}{(n-1)!} f(y_n)\,[F(y_n)]^{n-1}$$

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

## 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:

$$g(y_i, y_j) = \begin{cases} \dfrac{n!}{(i-1)!(j-i-1)!(n-j)!} \times [F(y_i)]^{i-1}[F(y_j)-F(y_i)]^{j-i-1}[1-F(y_j)]^{n-j} f(y_i) f(y_j) & , \quad a < y_i < y_j < b \\ 0 & \quad o.w \end{cases}$$




**Ex:** let $X_1, X_2, \dots, X_n$ be a random sample of size ($n$) rssn taken from C.U(0,1). let $Y_1 < Y_2 < \dots < Y_n$ 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.:** $X \sim C.U(0,1)$

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

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

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

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

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

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

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

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

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

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

$$g(y_i, y_j) = \begin{cases} \dfrac{n!}{(i-1)!(j-i-1)!(n-j)!} \times [F(y_i)]^{i-1}[F(y_j)-F(y_i)]^{j-i-1}[1-F(y_j)]^{n-j} f(y_i) f(y_j) & , \quad a < y_i < y_j < b \\ 0 & \quad o.w \end{cases}$$

*When;* $i=1 \quad , \; j=n$

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

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

**Ex:** let $X_1, X_2, \dots, X_n$ be a rssn taken from $\text{Exp}(1/\theta)$, let $Y_1 < Y_2 < \dots < Y_n$ be the order statistics of this sample. **Find** $g(y_2)$ , $g(y_{n-1})$ and when ($n = 4$) Find $g(y_1, y_3)$.

**Sol.:** $X \sim \text{Exp}(1/\theta)$

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

$$F(x)=p(X\le x)=\int_0^x \theta e^{-\theta x}\,dx=-e^{-\theta x}\Big|_0^x=\begin{cases}0 & , \; x\le 0 \\ 1-e^{-\theta x} & , \; 0<x<\infty \\ 1 & , \; x\to\infty\end{cases}$$

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

$$\Rightarrow \; when \; i=2 \;\; \Rightarrow \therefore \; g(y_2)=n(n-1)\,f(y_2)[F(y_2)][1-F(y_2)]^{n-2} \qquad , \quad 0<y_2<\infty$$

$$and \; when \; x=y_2 \;\; \Rightarrow \therefore \;\; f(y_2)=\theta e^{-\theta y_2} \quad , and \;\; F(y_2)=1-e^{-\theta y_2}$$

$$\therefore g(y_2)=n\,(n-1)\left[\theta e^{-\theta y_2}\right]\left[1-e^{-\theta y_2}\right]\left[1-\left(1-e^{-\theta y_2}\right)\right]^{n-2}$$

$$=(n^2-n)\,\theta e^{-\theta y_2}(1-e^{-\theta y_2})\left(e^{-n\theta y_2+2\theta y_2}\right)$$

$$=\begin{cases}(n^2-n)\,\theta\,(1-e^{-\theta y_2})\left(e^{-n\theta y_2+\theta y_2}\right) & , \quad 0<y_2<\infty \\ 0 & o.w\end{cases}$$

*when* $i=n-1$

$$\therefore \; g(y_{n-1})=\frac{n!}{(n-1-1)!(n-n+1)!}f(y_{n-1})\left[F(y_{n-1})\right]^{n-1-1}\left[1-F(y_{n-1})\right]^{n-n+1}$$

$$=n(n-1)\,f(y_{n-1})\left[F(y_{n-1})\right]^{n-2}\left[1-F(y_{n-1})\right] \qquad , \; 0<y_{n-1}<\infty$$

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

$$=n(n-1)\left(\theta e^{-2\theta y_{n-1}}\right)\left(1-e^{-\theta y_{n-1}}\right)^{n-2} \quad , \quad 0<y_{n-1}<\infty$$

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

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

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

*When;* $i=1 \quad , \; j=3 \quad , \; n=4$

$$\therefore g(y_1,y_3)=\frac{4!}{0!\;1!\;1!}\left(F(y_3)-F(y_1)\right)\left(1-F(y_3)\right)f(y_1)\,f(y_3)$$

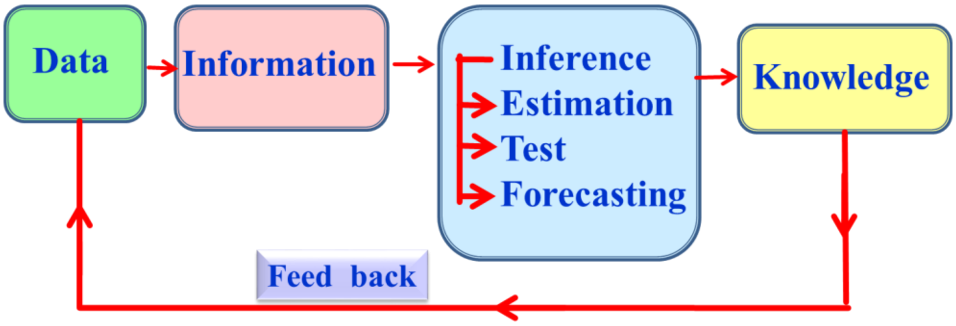
$$=24\left((1-e^{-\theta y_3})-(1-e^{-\theta y_1})\right)\left(1-(1-e^{-\theta y_3})\right)\theta\,e^{-\theta y_1}\,\theta\,e^{-\theta y_3}$$

$$=24\,\theta^2\,e^{-\theta y_1}\,e^{-2\theta y_3}\,(e^{-\theta y_1}-e^{-\theta y_3}) \qquad , \; 0<y_1<y_3<\infty$$

# Chapter Three

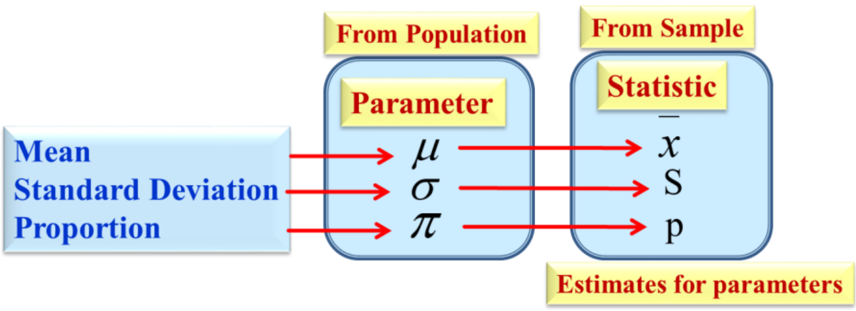
# Statistical Inference

**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 $\mu$ (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.**

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

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

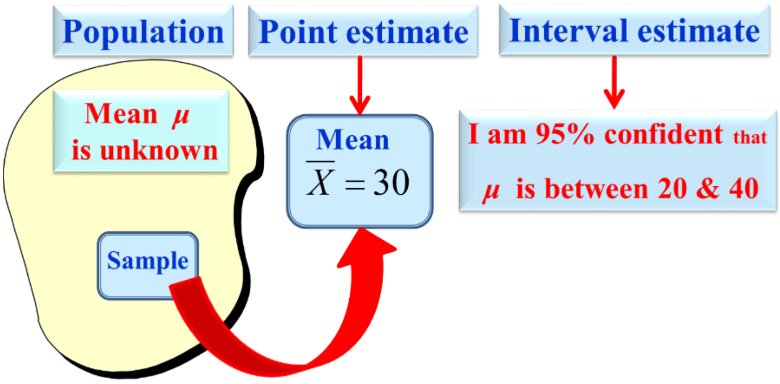
There are two steps to make inference:

## 1. Estimation of the population parameters

**a)** Point Estimation.
**b)** Intervals Estimation.

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

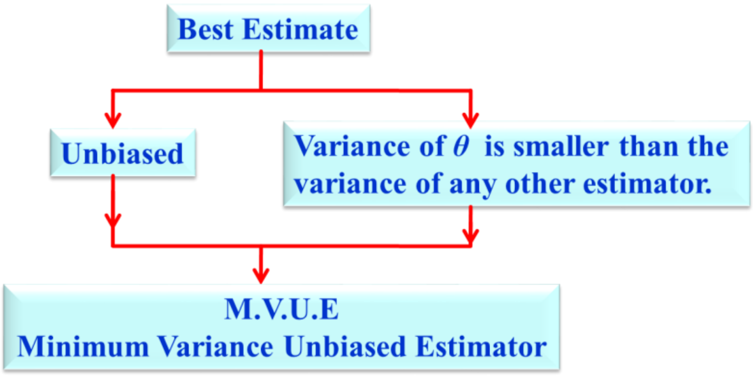
## Estimation of Parameters




## First: Point Estimation

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

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

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

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

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

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

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

**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 $\theta$.

**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 $\theta^2$.

**Ex:** In a random sample of size ($n$) from normal dist$^n$ N($\theta$, $\sigma^2$). Show that;

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

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

**Sol: 1)**

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

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

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

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

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

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

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

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

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

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

**2)**

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

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

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

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

Then; what is to be unbiased estimator for $\theta^2$.

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

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

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

**Ex:** In a random sample of size ($n$). Is $T = \overline{X}$ unbiased estimator for $\phi(\theta) = \theta$ of;

**1.** Ber($\theta$). **2.** Poisson($\theta$).

**Sol:**

$$E(T) \overset{?}{=} \phi(\theta) = \theta$$

1) $X \sim Ber(\theta) \quad \Rightarrow \quad f(x) = \theta^x (1-\theta)^{1-\theta} \quad x = 0,1 \quad , E(X) = \theta$

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

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

2) $X \sim Poi(\theta) \quad \Rightarrow \quad f(x) = \frac{e^{-\theta}\, \theta^x}{x!} \quad x = 0,1,2,.... \quad , E(X) = \theta$

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

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

## Unbiased in Limit

An estimator $\hat{\theta}$ for known parameter $\theta$ 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, $\theta$).

**1)** Is $Y_n$ unbiased in limit estimator for $\theta$; (Note: $Y_n$ estimator $\theta$).

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

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

**Sol: 1)**

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

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

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

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

$$= \frac{n}{\theta^n}\frac{y_n^{n+1}}{n+1}\Big|_0^{\theta} = \frac{n}{\theta^n}\frac{\theta^{n+1}}{n+1} = \frac{n}{n+1}\theta \ne \theta \rightarrow \therefore Y_n \textit{ is not unbiased est. for } \theta$$

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

**2)**

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

$$E(X) = \frac{a+b}{2} = \frac{\theta}{2}$$

$$E(\overline{X}) = E\left(\frac{\sum X_i}{n}\right) = \frac{n}{n} E(X) = \frac{\theta}{2} \neq \theta \quad \rightarrow \therefore \overline{X} \text{ is not unbiased estimator for } \theta$$

$$\lim_{n\to\infty} E(\hat{\theta} = \overline{X}) = \lim_{n\to\infty} \frac{\theta}{2} = \frac{\theta}{2} \neq \theta \rightarrow \therefore \overline{X} \text{ is not unbiased in limit estimator for } \theta$$

**3)**

$$E(\overline{X}) = E\left(\frac{\sum X_i}{n}\right) = \frac{n}{n} E(X) = \frac{\theta}{2} \quad \rightarrow \therefore \overline{X} \text{ is unbiased estimator for } \frac{\theta}{2}.$$

$$\lim_{n\to\infty} E(\overline{X}) = \lim_{n\to\infty} \frac{\theta}{2} = \frac{\theta}{2} \quad \rightarrow \therefore \overline{X} \text{ is unbiased in limit estimator for } \frac{\theta}{2}.$$

**Ex:** In a random sample of size ($n$) from normal dist$^n$ N($\theta$, $\sigma^2$). Is $S^2 = \frac{1}{n}\sum(X_i - \overline{X})^2$ unbiased estimator for the parameter ($\sigma^2$).

**Sol:**

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

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

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

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

$$= \sigma^2 - \frac{\sigma^2}{n} = \frac{n\sigma^2 - \sigma^2}{n} = \frac{(n-1)}{n}\sigma^2 \neq \sigma^2$$

$$\therefore S^2 = \frac{1}{n}\sum(X_i - \overline{X})^2 \text{ is not unbiased estimator for } \sigma^2$$

$$\lim_{n\to\infty} E(S^2) = \lim_{n\to\infty} \frac{(n-1)}{n}\sigma^2 = \lim_{n\to\infty}\left(\frac{n}{n}\sigma^2 - \frac{\sigma^2}{n}\right)$$

$$= \lim_{n\to\infty} \frac{n}{n}\sigma^2 - \lim_{n\to\infty} \frac{\sigma^2}{n}$$

$$= \sigma^2 - 0 = \sigma^2 \quad \rightarrow \therefore S^2 \text{ is unbiased in limit estimator for } \sigma^2.$$

## 2. Consistency Estimator

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

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

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

**Note:** Consistency means the estimator equal to the parameter or converges stochastically to the parameter $\theta$.

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

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

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

**Theorem:** Let $\hat{\theta}$ be an estimator for the population parameter $\theta$ of $f(x;\theta)$, then $\hat{\theta}$ is said to be consistent estimator for $\theta$ 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 rssn from Poisson dist$^n$, show that $\hat{\theta} = \overline{X}$ is consistent estimator for $\theta$.

**Sol:**

*First Method;*

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

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

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

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

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

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

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

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

*Second Method;*

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

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

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

**Ex:** Let $X_1, X_2, \ldots, X_n$ be a rss$n$ from normal dist$^n$ N($\theta$, $\sigma^2$), show that $S^2 = \frac{1}{n}\sum(X_i - \overline{X})^2$ is consistent estimator for $\sigma^2$.

**Sol:**

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

$$V(S^2) = ?$$

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

$$\frac{X-\theta}{\sigma} \sim N(0,1)$$

$$\frac{(X-\theta)^2}{\sigma^2} \sim \chi^2_{(1)}$$

$$\frac{\sum_{i=1}^{n}(X_i-\theta)^2}{\sigma^2} \sim \chi^2_{(n)} \qquad , \theta = \overline{X} \Rightarrow S^2 = \frac{\sum_{i=1}^{n}(X_i-\overline{X})^2}{n} \quad \Rightarrow \quad nS^2 = \sum_{i=1}^{n}(X_i-\overline{X})^2$$

$$\frac{nS^2}{\sigma^2} \sim \chi^2_{(n)}$$

$$V\left(\frac{nS^2}{\sigma^2}\right) = V\left(\chi^2_{(n)}\right)$$

$$\frac{n^2 V(S^2)}{\sigma^4} = 2n \qquad \Rightarrow V(S^2) = \frac{2n\sigma^4}{n^2} = \frac{2\sigma^4}{n}$$

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

$$\lim_{n\to\infty} p\left(\left|S^2 - \sigma^2\right| < \varepsilon\right) \geq \lim_{n\to\infty}\left(1 - \frac{2\sigma^4}{n\varepsilon^2}\right)$$

$$\lim_{n\to\infty} p\left(\left|S^2 - \sigma^2\right| < \varepsilon\right) = 1$$

$$\therefore\ S^2 \text{ is consistent estimator for } \sigma^2$$

**Ex:** Show that $\hat{\theta} = Y_n$ is consistent estimator for $\theta$ from C.U(0 , $\theta$), (by theorem).

**Sol:**

1) $\lim_{n\to\infty} E(\hat{\theta} = Y_n) \overset{?}{=} \theta$

As previous;

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

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

$$\lim_{n\to\infty} E(Y_n) = \lim_{n\to\infty} \frac{n}{n+1}\theta = (1)\,\theta = \theta$$

2) $\lim_{n\to\infty} V(Y_n) = 0$

$$V(Y_n) = E(Y_n^2) - (E(Y_n))^2$$

$$E(Y_n^2) = \int_0^\theta y_n^2 \frac{n\, y_n^{n-1}}{\theta^n}\, dy_n = \frac{n}{\theta^n}\int_0^\theta y_n^{n+1}\, dy_n = \frac{n}{\theta^n} \left.\frac{y_n^{n+2}}{n+2}\right|_0^\theta = \frac{n}{n+2}\,\theta^2$$

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

$$\lim_{n\to\infty} V(y_n) = \lim_{n\to\infty}\frac{n}{n+2}\,\theta^2 - \lim_{n\to\infty}\frac{n^2}{(n+1)^2}\,\theta^2 = \theta^2 - \theta^2 = 0$$

$\hat{\theta} = Y_n$ *is consistent estimator for* $\theta$.

**Ex:** In a rss$n$, show that $\overline{X}$ is consistent estimator for the parameter $\theta$, from;
**1)** N($\theta$, $\sigma^2$). **2)** Geo($\theta$).

**Sol:**

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

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

$$\hat{\theta} = \overline{X} \quad , \quad V(\hat{\theta}) = V(\overline{X}) = \frac{v(X)}{n} = \frac{\sigma^2}{n}$$

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

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

$\therefore\ \overline{X}$ *is a consistent est for* $\theta$

2) $X \sim Geo(\theta)$ , $E(X) = \frac{(1-\theta)}{\theta}$ , $V(X) = \frac{(1-\theta)}{\theta^2}$

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

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

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

$$\lim_{n\to\infty} p\left(\left|\overline{X} - \theta\right| < \varepsilon\right) = 1 \qquad \therefore\ \overline{X} \text{ is a consistent est for } \theta.$$

## The Score Function

The score function is the partial derivative of Log the function $f(x;\theta)$ with respect to the parameter $\theta$, 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; $\theta$)) = 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 $\theta$, and can be written as;

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

*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 $\theta$.

## 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 $\theta$ if the Fisher information in $\hat{\theta}$ is equal to the Fisher information in a rss($n$).

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

**Sol:**

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

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

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

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

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

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

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

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

$$\hat{\theta} = \sum X_i = x_1 + x_2 + \ldots + x_n$$

$$X \sim Ber(\theta) \qquad , \Rightarrow \sum X_i \sim Bin(n,\theta)$$

$$f(\sum x_i\,;\,\theta) = C^n_{\sum x_i}\, \theta^{\sum x_i} (1-\theta)^{n-\sum x_i}$$

$$\ln f(\sum x_i\,;\,\theta) = \ln C^n_{\sum x_i} + \sum x_i \ln(\theta) + (n - \sum x_i)\ln(1-\theta)$$

$$\frac{\partial \ln f(\sum x_i\,;\,\theta)}{\partial \theta} = zero + \frac{\sum x_i}{\theta} - \frac{(n-\sum x_i)}{(1-\theta)}$$

$$\frac{\partial^2 \ln f(\sum x_i\,;\,\theta)}{\partial \theta^2} = -\frac{\sum x_i}{\theta^2} - \frac{(n-\sum x_i)}{(1-\theta)^2}$$

$$-E\left(\frac{\partial^2 \ln f(\sum x_i\,;\,\theta)}{\partial \theta^2}\right) = \frac{n\,E(X)}{\theta^2} + \frac{(n-n\,E(X))}{(1-\theta)^2}$$

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

$$F.I\ in\ a\,rssn = F.I\ in\ \hat{\theta} = \sum X_i$$

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

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

**Sol:**

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

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

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

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

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

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

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

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

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

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

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

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

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

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

$$\because F.I \text{ in } arssn = F.I \text{ in } (\hat{\theta} = \bar{x})$$

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

**Ex:** In a rss$n$ from Poisson dist$^n$ Poi($\theta$), is $\sum X_i$ sufficient estimator for $\theta$?

**Sol:**

$$F.I.\ in\ a\ rss(n) \overset{?}{=} F.I.(\hat{\theta})$$

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

$$f(x;\theta) = \frac{e^{-\theta}\theta^x}{x!} \quad , x = 0,1,2,......$$

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

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

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

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

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

$$\therefore \hat{\theta} = \sum X_i = x_1 + x_2 + .... + x_n \sim Poi(n\theta)$$

$$f(\sum x_i\,;\,n\theta) = \frac{e^{-n\theta}(n\theta)^{\sum x_i}}{(\sum x_i)!}$$

$$\ln f(\sum x_i\,;\,n\theta) = \sum x_i \ln(n\theta) - n\theta - \ln((\sum x_i)!)$$

$$\frac{\partial \ln f(\sum x_i\,;\,n\theta)}{\partial\theta} = \frac{n\sum x_i}{n\theta} - n$$

$$\frac{\partial^2 \ln f(\sum x_i\,;\,n\theta)}{\partial\theta^2} = -\frac{\sum x_i}{\theta^2}$$

$$-E\left(\frac{\partial^2 \ln f(\sum x_i\,;\,n\theta)}{\partial\theta^2}\right) = -E\left(\frac{-\sum x_i}{\theta^2}\right) = \frac{nE(X)}{\theta^2} = \frac{n\theta}{\theta^2} = \frac{n}{\theta} \quad is\ F.I\ in\ \hat{\theta} = \sum X_i$$

$$\because F.I\ in\ a\,rssn = F.I\ in\ \hat{\theta} = \sum X_i$$

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

## 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 $\theta$, an estimator $\hat{\theta}$ is sufficient estimator for the parameter $\theta$ if the conditional p.d.f. of ($X_1$, $X_2$, …, $X_n$) given $\hat{\theta}$ does not contain the parameter $\theta$:

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

**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 $\theta$.

**Sol:**

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

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

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

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

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

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

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

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

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

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

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

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

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

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

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

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

$$\therefore\ Y_1 \ \ is\ \ suff\ \ est\ \ for\ \ \theta$$

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

**Sol:**

$$f(x_1,x_2,\dots,x_n \mid \hat\theta = \sum X_i) = \frac{f(x_1,x_2,\dots,x_n;\theta)}{g(\sum x_i)}$$

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

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

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

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

$X \sim Poi(\theta)$

$X_1 + X_2 + \dots + X_n = \sum X_i \sim Poi(n\theta)$

$$f(\sum x_i ; n\theta) = \frac{e^{-n\theta}(n\theta)^{\sum x_i}}{(\sum x_i)!}$$

$$f(x_1, x_2, \dots, x_n \mid \hat{\theta} = \sum x_i) = \frac{e^{-n\theta}\,\theta^{\sum x_i} / \prod (x_i)!}{e^{-n\theta}(n\theta)^{\sum x_i} / (\sum x_i)!}$$

$$= \frac{(\sum x_i)!}{\prod (x_i)!\, n^{\sum x_i}} \qquad \text{does not contain } \theta$$

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

**Ex:** Let $X_1, X_2, \dots, X_n$ be a rssn. Is $\overline{X}$ sufficient estimator for $\theta$? of; **1)** $Exp(1/\theta)$. **2)** $N(\theta, \sigma^2)$.

**Sol: 1)**

$$f(x_1, x_2, \dots, x_n \mid \hat{\theta} = \bar{x}) = \frac{f(x_1, x_2, \dots, x_n ; \theta)}{f(\hat{\theta} = \bar{x})}$$

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

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

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

$$= \theta^n e^{-\theta \sum x_i} \qquad \text{joint p.d.f.}$$

$$= \theta^n e^{-\theta n \bar{x}}$$

$$n\overline{X} \sim \Gamma(n, \theta) \qquad , \; f(x;\alpha,\beta) = \frac{\beta^\alpha}{\Gamma(\alpha)} x^{\alpha-1} e^{-\beta x}$$

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

$$\therefore f(x_1, x_2, \dots, x_n \mid \bar{x}) = \frac{\theta^n e^{-\theta n \bar{x}}}{\frac{\theta^n}{\Gamma(n)}(n\bar{x})^{n-1} e^{-n\theta\bar{x}}}$$

$$= \frac{\Gamma(n)}{(n\bar{x})^{n-1}} \qquad \text{does not depend on } \theta$$

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

**2)** $N(\theta, \sigma^2)$.

$$f(x_1, x_2, \dots, x_n \mid \hat{\theta} = \bar{x}) = \frac{f(x_1, x_2, \dots, x_n ; \theta)}{f(\hat{\theta} = \bar{x})}$$

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

$\because$ Xs are independent

$$f(x_1, x_2, \dots, x_n ; \theta) = \prod_{i=1}^{n} f(x_i ;\theta,\sigma^2) = f(x_1;\theta)\cdot f(x_2;\theta)\cdots f(x_n;\theta)$$

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

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

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

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

$$but;\ \sum(x_i-\theta)^2 = \sum(x_i-\bar{x})^2 + n(\bar{x}-\theta)^2$$

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

$$= \frac{\left(\frac{1}{\sqrt{2\pi\sigma^2}}\right)^n e^{-\frac{1}{2\sigma^2}\sum(x_i-\bar{x})^2}}{\frac{\sqrt{n}}{\sqrt{2\pi\sigma^2}}} \quad does\ not\ depend\ on\ \theta$$

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

## 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 $\theta$. Then the necessary and sufficient condition for an estimator $\hat{\theta}$ to be sufficient estimator, if there are two non-negative functions, such that:

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

**Theorem:**

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

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

**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 rss$n$ from Bernoulli dist$^n$ Ber($\theta$). Show that $\hat{\theta} = \sum X_i$ is sufficient estimator for the parameter $\theta$.

**Sol:**

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

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

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

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

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

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

**Sol:**

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

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

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

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

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

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

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

For $\overline{X}$ we get;

$$f(x_1, x_2, \dots, x_n; \theta) = e^{-n\theta}\,\theta^{n\bar{x}} \cdot \frac{1}{\prod (x_i)!}$$

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

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

**Sol:**

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

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

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

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

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

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

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

**Ex:** From Beta($\theta, \beta = 1$), $f(x;\theta) = \theta x^{\theta - 1} \quad , \quad 0 < x < 1$

Is $\prod_{i=1}^{n} X_i$ sufficient estimator for $\theta$? (By factorization theorem).

**Sol:**

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

$$\because Xs \text{ are independent}$$

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

$$=\theta x_1^{\theta-1}\times\theta x_2^{\theta-1}\times\dots\times\theta x_n^{\theta-1}$$

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

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

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

$$\therefore\ \hat{\theta}=\prod_{i=1}^{n} X_i \text{ is suff est for } \theta.$$

**Note:** $x_1^{\alpha-1}\cdot x_2^{\alpha-1}\cdot\ldots\cdot x_n^{\alpha-1}=\left(\prod_{i=1}^{n} x_i\right)^{\alpha-1}.$

**Ex:** Let $X_1, X_2, \ldots, X_n$ be a rss$n$. Is $\sum X_i^2$ sufficient estimator for $\theta$? From N(0, $\theta$).

$\because X\sim N(0,\theta)$

$$f(x;\theta)=\frac{1}{\sqrt{2\pi\theta}}\,e^{-\frac{x^2}{2\theta}}\qquad,\quad\because Xs \text{ are independent}$$

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

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

$$\therefore\ \hat{\theta}=\sum X_i^2 \text{ is suff est for } \theta.$$

## 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:

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

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 $(\alpha, \theta)$.

**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 ($\theta$, $\sigma^2$) 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 ($\theta_1$, $\theta_2$) respectively.

**Sol:**

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

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

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

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

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

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

$$0 \qquad \text{o.w}$$

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

$$= n(n-1)\left(\frac{1}{2\theta_2}\left(y_n - (\theta_1 - \theta_2)\right) - \frac{1}{2\theta_2}\left(y_1 - (\theta_1 - \theta_2)\right)\right)^{n-2} \frac{1}{2\theta_2} \times \frac{1}{2\theta_2}$$

$$= n(n-1)\frac{1}{(2\theta_2)^n}(y_n - y_1)^{n-2} \quad , \quad \theta_1 - \theta_2 < y_1 < y_n < \theta_1 + \theta_2$$

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

$$f(x_1, x_2, \dots, x_n; \theta_1, \theta_2 \mid y_1, y_n) = \frac{f(x_1, x_2, \dots, x_n; \theta_1, \theta_2)}{g(y_1, y_n)}$$

$$= \frac{\left(\frac{1}{2\theta_2}\right)^n}{n(n-1)\frac{1}{(2\theta_2)^n}(y_n - y_1)^{n-2}}$$

$$= \frac{1}{n(n-1)(y_n - y_1)^{n-2}} = \frac{1}{n(n-1)(X_{\max} - X_{\min})^{n-2}}$$

$\because$ *which not depend on* $\theta_1, \theta_2$

$\therefore Y_1$ *and* $Y_n$ *are jointly suff est for* $\theta_1$ *and* $\theta_2$ *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:

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

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

**Ex:** Let $X_1, X_2, \dots, X_n$ be a rss$n$ from Bernoulli dist[n] Ber($\theta$), show that if the dist[n] of X can be written in exponential form?

**Sol:**

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

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

**Ex:** Let $X_1, X_2, \dots, X_n$ be a rss$n$ from Poisson dist[n] Poi($\theta$), show that if the dist[n] of X can be written in exponential form?

**Sol:**

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

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

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

**H.W:** Let $X_1$, $X_2$, …, $X_n$ be a rss$n$ from normal dist$^n$ N(0, $\theta$), 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 ($X_1$, $X_2$, …, $X_n$) is:

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

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

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

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

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

**1)** Poisson($\theta$). **2)** Beta(α, $\beta$).

**Sol:**

1) $Poisson(\theta)$

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

*In a rssn*;

$$\begin{aligned} f(x_1,x_2,\dots,x_n;\theta) &= \exp\big(-n\,\theta + \ln(\theta)\sum x_i - \sum\ln(x!)\big) \\ &= \exp\big(-n\,\theta + \ln(\theta)\sum x_i\big)\cdot\exp\big(-\sum\ln(x!)\big) \end{aligned}$$

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

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

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

$$\begin{aligned} f(x;\theta) &= \frac{\Gamma(\alpha+\beta)}{\Gamma(\alpha)\Gamma(\beta)}\,x^{\alpha-1}(1-x)^{\beta-1} \\ &= \exp\big(-\ln\beta(\alpha,\beta) + (\alpha-1)\ln(x) + (\beta-1)\ln(1-x)\big) \\ &= \exp\big(-\ln\beta(\alpha,\beta) + \alpha\ln(x) - \ln(x) + \beta\ln(1-x) - \ln(1-x)\big) \end{aligned}$$

*In a rssn*;

$$\begin{aligned} f(x_1,x_2,\dots,x_n;\theta) &= \exp\begin{pmatrix} -n\ln\beta(\alpha,\beta) + \alpha\sum\ln(x_i) - \sum\ln(x_i) + \\ + \beta\sum\ln(1-x_i) - \sum\ln(1-x_i) \end{pmatrix} \\ &= \exp\big(-n\ln\beta(\alpha,\beta) + \alpha\sum\ln(x_i) + \beta\sum\ln(1-x_i)\big)\times \\ &\quad\times\exp\big(-\sum\ln(x_i) - \sum\ln(1-x_i)\big) \\ &= \exp\big(p_1(\alpha)\sum K_1(x_i) + p_2(\beta)\sum K_2(x_i) + n\,q(\alpha,\beta)\big)\times\exp\big(\sum S(x_i)\big) \end{aligned}$$

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

$$are\ \ minimal\ \ suff\ \ est\ \ for\ \ \alpha\ \ and\ \ \beta\ \ respectively$$

**Ex:** In a rss$n$. Find minimal sufficient estimators for $\theta$ from $\Gamma(2, \theta)$.

**Sol:**

$X \sim \Gamma(2,\theta)$

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

$$f(x;\alpha = 2, \beta = \theta) = \frac{1}{\Gamma(2)\theta^2} x^{2-1} e^{-x/\theta} = \frac{1}{\theta^2} x\, e^{-x/\theta}$$

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

$$= \exp\left(\ln(1) - 2\ln(\theta) + \ln(x) - \frac{x}{\theta}\right)$$

$$= \exp\left(-2\ln(\theta) + \ln(x) - \frac{x}{\theta}\right)$$

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

*In a rssn;*

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

$$= \exp\left(p(\theta)\sum K(x_i) + n\,q(\theta)\right) \times \exp\left(\sum S(x_i)\right)$$

$$\Rightarrow \therefore \hat{\theta} = \sum K(X_i) = \sum X_i \quad \text{is minimal suff est for } \theta$$

**Ex:** In a rss$n$. Find minimal sufficient estimators for $\theta$, $\sigma^2$ from $N(\theta, \sigma^2)$.

**Sol:**

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

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

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

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

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

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

*In a rssn;*

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

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

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

## The Rao-Blackwell Theorem

Let X has a p.d.f. $f(x;\theta)$, and $u$ be an unbiased estimator for parameter $\theta$, and $T$ be sufficient estimator, then;

1) $E(U) = E\big(E(U|T)\big).$

2) $Var(U) \geq Var\big(E(U|T)\big)$

**Proof(1):**

$Let;\quad E(U|T) = W$

$$E(W) \overset{?}{=} E(U)$$

$$E(W) = \int w\, f(t)\, dt = \int E(U|T)\, f(t)\, dt = \int \Big(\int u\, f(u|t)\, du\Big) f(t)\, dt$$

$$= \int\int u\, \frac{f(u,t)}{f(t)}\, f(t)\, dt\, du \;=\; \int u \int f(u,t)\, dt\, du = \int u\, f(u)\, du = E(U)$$

$$\therefore E(W) = E(U) = \theta$$

**Proof(2):**

$$Var(U) = E(U-\theta)^2 \qquad \{\mp W$$

$$= E\big((U-W)+(W-\theta)\big)^2$$

$$= E(U-W)^2 + 2E(U-W)(W-\theta) + E(W-\theta)^2$$

$$\therefore Var(U) = E(U-W)^2 + Var(W)$$

$$\Rightarrow \therefore\; Var(U) \geq Var(W)$$

**Proof:** $E(U-W)(W-\theta) \overset{?}{=} 0$

$$E(U-W)(W-\theta) = \big(E(U)-E(W)\big)(W-\theta)$$

$$= (\theta-\theta)(W-\theta)$$

$$= zero$$

**Ex:** Let X and Y be two random variables with j.p.d.f.;

$$f(x,y) = \frac{2}{\theta^2}\, e^{-(x+y)/\theta} \qquad ,\; 0 < x < y < \infty$$

Show that; 1) $E(Y) = E\big(E(Y|X)\big).$

2) $Var(Y) \geq Var\big(E(Y|X)\big).$

**Sol:**

1) $$f(y) = \int_0^y f(x,y)\, dx = \int_0^y \frac{2}{\theta^2}\, e^{-(x+y)/\theta} dx = \frac{2}{\theta}\, e^{-y/\theta} \int_0^y \frac{1}{\theta} e^{-x/\theta} dx$$

$$= \frac{2}{\theta}\, e^{-y/\theta} \Big(-e^{-x/\theta}\Big)_0^y = \frac{2}{\theta}\, e^{-y/\theta}\, (1-e^{-y/\theta})$$

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

$$E(Y) = \int_0^\infty y\, f(y)\, dy = \int_0^\infty y \left( \frac{2}{\theta} e^{-y/\theta} - \frac{2}{\theta} e^{-2y/\theta} \right) dy$$

$$= \int_0^\infty \frac{2}{\theta}\, y\, e^{-y/\theta}\, dy - \int_0^\infty \frac{2}{\theta}\, y\, e^{-2y/\theta}\, dy$$

$$= \frac{2}{\theta} \int_0^\infty y\, e^{-y/\theta}\, dy - \frac{2}{\theta} \int_0^\infty y\, e^{-2y/\theta}\, dy$$

$$= \frac{2}{\theta}\Gamma(2)\theta^2 - \frac{2}{\theta}\Gamma(2)\left(\frac{\theta}{2}\right)^2 \qquad , \ \frac{1}{\beta} = \frac{2}{\theta} \Rightarrow \beta = \frac{\theta}{2}$$

$$= 2\theta - \frac{\theta}{2} = \frac{4\theta - \theta}{2} = \frac{3\theta}{2}$$

$$E(Y|X) = \int_x^\infty y\, f(y|x)\, dy$$

$$f(y|x) = \frac{f(x,y)}{f(x)}$$

$$f(x) = \int_x^\infty f(x,y)\, dy = \frac{2}{\theta} e^{-x/\theta} \int_x^\infty \frac{1}{\theta} e^{-y/\theta} dy = \frac{2}{\theta} e^{-x/\theta} \left( -e^{-y/\theta} \right)_x^\infty$$

$$= \frac{2}{\theta} e^{-x/\theta} e^{-x/\theta} = \frac{2}{\theta} e^{-2x/\theta} \qquad , \quad 0 < x < \infty$$

$$\therefore f(y|x) = \frac{\frac{2}{\theta^2} e^{-(x+y)/\theta}}{\frac{2}{\theta} e^{-2x/\theta}} = \frac{2}{\theta^2} e^{-(x+y)/\theta} \times \frac{\theta}{2} e^{2x/\theta} = \frac{1}{\theta} e^{(-x+2x-y)/\theta}$$

$$= \frac{1}{\theta} e^{-(y-x)/\theta} \qquad , \quad 0 < x < y < \infty$$

$$\therefore E(Y|X) = \int_x^\infty y \frac{1}{\theta} e^{-(y-x)/\theta}\, dy = \frac{1}{\theta} \int_x^\infty y\, e^{-(y-x)/\theta}\, dy$$

$$Let; \quad u = y - x \qquad \Rightarrow y = u + x \quad , dy = du$$

$$\therefore E(Y|X) = \frac{1}{\theta} \int_0^\infty (u + x)\, e^{-u/\theta}\, du$$

$$= \frac{1}{\theta} \int_0^\infty u\ e^{-u/\theta}\, du + \frac{1}{\theta} \int_0^\infty x\, e^{-u/\theta}\, du$$

$$= \frac{1}{\theta}\Gamma(2)\theta^2 + \frac{x}{\theta}\Gamma(1)\theta$$

$$= \theta + x = w$$

$$E(Y|X) = \theta + x$$

$$E(W) = E\big(E(Y|X)\big) = \theta + E(X)$$

$$E(X) = \frac{2}{\theta} \int_0^\infty x\, e^{-2x/\theta}\, dx = \frac{2}{\theta}\Gamma(2)\left(\frac{\theta}{2}\right)^2 = \frac{\theta}{2}$$

$$E\big(E(Y|X)\big) = \theta + \frac{\theta}{2} = \frac{3\theta}{2} \quad \textit{this is the first condition}$$

$$2)\ E(Y^2) = \int_0^\infty y^2\, f(y)\, dy = \int_0^\infty y^2 \left( \frac{2}{\theta} e^{-y/\theta} - \frac{2}{\theta} e^{-2y/\theta} \right) dy$$

$$= \int_0^\infty \frac{2}{\theta} y^2 e^{-y/\theta}\, dy - \int_0^\infty \frac{2}{\theta} y^2 e^{-2y/\theta}\, dy$$

$$= \frac{2}{\theta} \int_0^\infty y^2 e^{-y/\theta}\, dy - \frac{2}{\theta} \int_0^\infty y^2 e^{-2y/\theta}\, dy$$

$$= \frac{2}{\theta} \Gamma(3) \theta^3 - \frac{2}{\theta} \Gamma(3) \left(\frac{\theta}{2}\right)^3 \qquad , \frac{1}{\beta} = \frac{2}{\theta} \Rightarrow \beta = \frac{\theta}{2}$$

$$= 4\theta^2 - \frac{\theta^2}{2} = \frac{8\theta^2 - \theta^2}{2} = \frac{7\theta^2}{2}$$

$$V(Y) = E(Y^2) - (E(Y))^2 = \frac{7\theta^2}{2} - \left(\frac{3\theta}{2}\right)^2 = \frac{7\theta^2}{2} - \frac{9\theta^2}{4} = \frac{14\theta^2 - 9\theta^2}{4} = \frac{5\theta^2}{4}$$

$$V(E(Y|X)) = V(\theta + X) = V(\theta) + V(X) = 0 + V(X) = V(X)$$

$$E(X^2) = \frac{2}{\theta} \int_0^\infty x^2 e^{-2x/\theta}\, dx = \frac{2}{\theta} \Gamma(3) \left(\frac{\theta}{2}\right)^3 = \frac{\theta^2}{2}$$

$$V(X) = E(X^2) - (E(X))^2 = \frac{\theta^2}{2} - \left(\frac{\theta}{2}\right)^2 = \frac{\theta^2}{2} - \frac{\theta^2}{4} = \frac{\theta^2}{4} = V(E(Y|X))$$

$$\because \frac{5\theta^2}{4} > \frac{\theta^2}{4}$$

$$\therefore V(Y) > V(E(Y|X)) \qquad \textit{this is the second condition}$$

**Ex:** In a rss3 from C.U(0, $\theta$). Show that [E($2Y_2$) = E{E($2Y_2$ | $Y_3$)}], and compare the variances of ($2Y_2$) and [E($2Y_2$ |$Y_3$)].

**Sol:1)**

$$f(x;\theta) = \frac{1}{\theta} \quad , \ 0 < x < \theta \quad , \ F(x) = \int_0^x \frac{1}{\theta}\, dx = \frac{x}{\theta} \ , \ f(y_2) = \frac{1}{\theta} \quad , \ F(y_2) = \frac{y_2}{\theta}$$

$$E(2Y_2) = 2\ E(Y_2)$$

$$E(Y_2) = \int_{R_{y_2}} y_2\, g(y_2)\, dy_2$$

$$g(y_i) = \frac{n!}{(i-1)!(n-i)!} f(y_i) (F(y_i))^{i-1} (1 - F(y_i))^{n-i}$$

$$g(y_2) = \frac{3!}{(2-1)!(3-2)!} f(y_2) (F(y_2)) (1 - F(y_2))$$

$$= 6 \frac{1}{\theta} \frac{y_2}{\theta} \left(1 - \frac{y_2}{\theta}\right) = \frac{6y_2}{\theta^2} \left(1 - \frac{y_2}{\theta}\right) \qquad , \ 0 < y_2 < \theta$$

$$E(2Y_2) = 2 \int_0^\theta y_2 \frac{6y_2}{\theta^2} \left(1 - \frac{y_2}{\theta}\right) dy_2 = 12 \int_0^\theta \left(\frac{y_2^2}{\theta^2} - \frac{y_2^3}{\theta^3}\right) dy_2$$

$$= 12\left(\frac{y_2^3}{3\theta^2}\Big|_0^\theta - \frac{y_2^4}{4\theta^3}\Big|_0^\theta\right) = 12\left(\frac{\theta^3}{3\theta^2} - \frac{\theta^4}{4\theta^3}\right) = 12\left(\frac{\theta}{3} - \frac{\theta}{4}\right) = 12\,\frac{4\theta - 3\theta}{12} = \theta$$

$\therefore 2y_2$ *is unbiased est for* $\theta$.

$$E(2Y_2|Y_3) = \int_{R_{y_2}} 2y_2\, g(y_2|\,y_3)\, dy_2$$

$$g(y_2|\,y_3) = \frac{g(y_2,\,y_3)}{g(y_3)}$$

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

$$g(y_3) = 3\, f(y_3)\,(F(y_3))^2 = 3\,\frac{1}{\theta}\left(\frac{y_3}{\theta}\right)^2 = \frac{3\,y_3^2}{\theta^3} \qquad , \quad 0 < y_3 < \theta$$

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

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

$$g(y_2,\,y_3) = 3!\,F(y_2)\, f(y_2)\, f(y_3) = 6\,\frac{y_2}{\theta}\,\frac{1}{\theta}\,\frac{1}{\theta} = \frac{6\,y_2}{\theta^3} \quad , \quad 0 < y_2 < y_3 < \theta$$

$$g(y_2|\,y_3) = \frac{6\,y_2\ /\theta^3}{3\,y_3^2\ /\ \theta^3} = \frac{2\,y_2}{y_3^2} \qquad , \quad 0 < y_2 < y_3 < \theta$$

$$E(2Y_2|Y_3) = \int_0^{y_3} 2\,y_2\,\frac{2\,y_2}{y_3^2}\, dy_2 = \frac{4}{y_3^2}\int_0^{y_3} y_2^2\, dy_2 = \frac{4}{y_3^2}\,\frac{y_2^3}{3}\Big|_0^{y_3} = \frac{4}{3}\,y_3$$

$$E(E(2\,Y_2\,|\,Y_3)) = E\left(\frac{4}{3}\,Y_3\right)$$

$$E\left(\frac{4}{3}Y_3\right) = \int_{R_{y_3}} \frac{4}{3}\,y_3\, g(y_3)\, dy_3 = \int_0^\theta \frac{4}{3}\,y_3\,\frac{3\,y_3^2}{\theta^3}\, dy_3 = \frac{4}{\theta^3}\int_0^\theta y_3^3\, dy_3$$

$$= \frac{4}{\theta^3}\,\frac{y_3^4}{4}\Big|_0^\theta = \theta$$

$$\therefore\ E(2\,Y_2) = E(E(2\,Y_2\,|\,Y_3))$$

**2)**

$$V(2Y_2) = 4\,V(Y_2)$$

$$V(Y_2) = E(Y_2^2) - (E(Y_2))^2$$

$$E(Y_2) = \int_{R_{y_2}} y_2\, g(y_2)\, dy_2 = 6\left(\frac{\theta}{12}\right) = \frac{\theta}{2}$$

$$E(Y_2^2) = \int_{R_{y_2}} y_2^2\, g(y_2)\, dy_2 = \int_0^\theta y_2^2\,\frac{6y_2}{\theta^2}\left(1 - \frac{y_2}{\theta}\right) dy_2$$

$$= 6\int_0^\theta \left(\frac{y_2^3}{\theta^2} - \frac{y_2^4}{\theta^3}\right) dy_2 = 6\left(\frac{y_2^4}{4\theta^2}\Big|_0^\theta - \frac{y_2^5}{5\theta^3}\Big|_0^\theta\right)$$

$$= 6\left(\frac{\theta^2}{4} - \frac{\theta^2}{5}\right) = 6\left(\frac{5\theta^2 - 4\theta^2}{20}\right) = \frac{3\theta^2}{10}$$

$$V(Y_2)=\frac{3\theta^2}{10}-\left(\frac{\theta}{2}\right)^2=\frac{3\theta^2}{10}-\frac{\theta^2}{4}=\frac{6\theta^2-5\theta^2}{20}=\frac{\theta^2}{20}$$

$$V(2Y_2)=4V(Y_2)=4\,\frac{\theta^2}{20}=\frac{\theta^2}{5}$$

$$V\big(E(2Y_2|Y_3)\big)=V\left(\frac{4}{3}Y_3\right)=\frac{16}{9}V(Y_3)$$

$$V(Y_3)=E(Y_3^2)-\big(E(Y_3)\big)^2$$

$$E(Y_3)=\int_{R_{Y_3}} y_3\,g(y_3)\,dy_3=\int_0^\theta y_3\left(\frac{3y_3^2}{\theta^3}\right)dy_3=\frac{3}{\theta^3}\int_0^\theta y_3^3\,dy_3=\frac{3}{\theta^3}\left.\frac{y_3^4}{4}\right|_0^\theta=\frac{3}{4}\theta$$

$$E(Y_3^2)=\int_{R_{Y_3}} y_3^2\,g(y_3)\,dy_3=\int_0^\theta y_3^2\left(\frac{3y_3^2}{\theta^3}\right)dy_3=\frac{3}{\theta^3}\int_0^\theta y_3^4\,dy_3=\frac{3}{\theta^3}\left.\frac{y_3^5}{5}\right|_0^\theta=\frac{3}{5}\theta^2$$

$$\therefore V(Y_3)=\frac{3}{5}\theta^2-\left(\frac{3}{4}\theta\right)^2=\frac{3\theta^2}{5}-\frac{9\theta^2}{16}=\frac{48\theta^2-45\theta^2}{80}=\frac{3\theta^2}{80}$$

$$\Rightarrow\therefore\ V\big(E(2Y_2|Y_3)\big)=\frac{16}{9}\,\frac{3\theta^2}{80}=\frac{\theta^2}{15}$$

$$\therefore\ \frac{\theta^2}{15}<\frac{\theta^2}{5}$$

$$\therefore V\big(E(2Y_2|Y_3)\big)<V(2Y_2)$$

## 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 does depend on $\theta$, then we use the general rule to derivative of integral;

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

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

**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;*

$$E\big(u(X)\big)=0$$

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

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

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

$\because \theta \neq 0$

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

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

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

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

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

$E(u(X)) = 0$

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

$$= \sum_{x=0}^{\infty} u(x)\ \frac{e^{-\theta}\theta^x}{x!} = 0$$

$$\Rightarrow e^{-\theta} \sum_{x=0}^{\infty} u(x)\ \frac{\theta^x}{x!} = 0 \qquad \} \div e^{-\theta}$$

$$\Rightarrow \sum_{x=0}^{\infty} u(x)\ \frac{\theta^x}{x!} = 0$$

$$\Rightarrow u(0) + u(1)\theta + u(2)\ \frac{\theta^2}{2!} + u(3)\ \frac{\theta^3}{3!} + ...... = 0$$

$\because \theta > 0$

$\therefore u(0) = u(1) = u(2) = u(3) = ..... = 0 \qquad \Rightarrow u(x) = 0 \ \ \forall x$

$\therefore f(x\,;\theta)$ *of Poisson dist is complete*

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

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

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

$E(u(X)) = 0$

$$E(u(X)) = \int_{-\infty}^{\infty} u(x)\ f(x;\theta, \sigma^2)\,dx = 0$$

$$= \frac{1}{\sqrt{2\pi\sigma^2}} \int_{-\infty}^{\infty} u(x)\ e^{-\frac{1}{2\sigma^2}(x-\theta)^2}\,dx = 0$$

$\because e^{-\frac{1}{2\sigma^2}(x-\theta)^2} \neq 0 \qquad \therefore u(x) = 0$

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

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

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

Show that $f(x;\theta)$ is complete?

**Sol:**

$\because$ *the range depend on* $\theta$.

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

$$E\big(u(X)\big) = 0$$

$$E\big(u(X)\big) = \int_0^{\theta} u(x)\ f(x;\theta)\ dx = 0$$

$$\Rightarrow \int_0^{\theta} u(x)\ \frac{1}{\theta}\ dx = 0 \qquad \} \times \theta$$

$$\Rightarrow \int_0^{\theta} u(x)\ dx = 0$$

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

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

$$\Rightarrow \frac{\partial}{\partial \theta} \int_0^{\theta} u(x)\ dx = 0$$

$$\Rightarrow \int_0^{\theta} \frac{\partial\, u(x)}{\partial \theta}\ dx + u(\theta)(1) - u(0)\ (0) = 0$$

$$\Rightarrow \int_0^{\theta} (0)\ dx + u(\theta)\ - 0 = 0$$

$$\Rightarrow\ 0 + u(\theta)\ - 0 = 0 \qquad , \therefore u(\theta) = 0 \quad , \quad \because\ \theta > 0 \quad , \quad i.e., \theta \neq 0$$

$$\therefore u(x) = 0 \quad , x > 0$$

$$\therefore u(x) = 0 \quad , 0 < x < \theta$$

$$\therefore\ E\big(u(X)\big) = 0 \qquad \forall\ x$$

$$\therefore\ f(x\ ;\ \theta)\quad is\ \ complete$$

## 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 $\theta$, and let $g(y_1;\theta)$ be **complete** if there is a continuous function of $Y_1$ which is an **unbiased** estimator for θ, $\phi(\theta)$ such that $E(\phi(\theta)) = \theta$, **then** $\phi(\theta)$ **is the unique best estimator for $\theta$ (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 $\theta$.

**Sol:**

$\because$ the range depend on $\theta$.

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

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

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

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

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

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

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

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

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

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

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

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 $X_1, X_2, \ldots, X_n$ be a rss$n$ from Poisson dist$^n$ poi($\theta$). Show that $Y = \sum X_i$ is a complete sufficient estimator for $\theta$. Find the unique continuous function of Y which is the best estimator for $\theta$ (M.V.U.E).

**Sol:**

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

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

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

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

$$f(x_1, x_2, \ldots, x_n; \theta) = \frac{e^{-n\theta}\,\theta^{\sum x_i}}{\prod (x_i)!} = e^{-n\theta}\,\theta^{\sum x_i} \times \frac{1}{\prod (x_i)!}$$

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

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

2) Let $u(y)$ is a fun of $Y$. then; $X \sim Poi(\theta) \Rightarrow Y = \sum X_i \sim Poi(n\theta)$

$$\text{Let; } n\theta = \lambda \qquad \Rightarrow g(y;\lambda) = \frac{e^{-\lambda}\lambda^{y}}{y!} \quad , \; y = 0, 1, \ldots.$$

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

$$E(u(Y)) = \sum_{y=0}^{\infty} u(y)\; g(y;\lambda)$$

$$\Rightarrow \sum_{y=0}^{\infty} u(y)\,\frac{e^{-\lambda}\lambda^{y}}{y!} = 0$$

$$\Rightarrow e^{-\lambda} \sum_{y=0}^{\infty} u(y)\,\frac{\lambda^{y}}{y!} = 0 \qquad \} \div e^{-\lambda}$$

$$\Rightarrow \sum_{y=0}^{\infty} u(y)\,\frac{\lambda^{y}}{y!} = 0$$

$$\Rightarrow u(0) + u(1)\lambda + u(2)\frac{\lambda^2}{2!} + u(3)\frac{\lambda^3}{3!} + \ldots\ldots = 0$$

$\because \lambda > 0 \Rightarrow \lambda \neq 0$

$\therefore u(0) = u(1) = u(2) = u(3) = \ldots.. = 0 \qquad \Rightarrow u(y) = 0 \quad \forall\; y$

$\therefore g(y;\lambda)$ is complete

3) $X \sim poi(\theta)$ , $E(X) = \theta$

$Y \sim poi(n\theta)$ , $E(Y) = E(\sum X_i) = n\,E(X) = n\theta$

$E(Y) = n\theta$

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

**Ex:** Let $X_1, X_2, \ldots, X_n$ be a rssn from Bernoulli dist$^n$ Ber($\theta$). Show that $Y = \sum X_i$ is a complete sufficient estimator for $\theta$. Find the unique continuous function of Y which is the best estimator for $\theta$ (M.V.U.E).

**Sol:**

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

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

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

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

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

$$= g(y;\theta) \times H(x)$$

$\therefore \hat{\theta} = \sum X_i = Y$ *is suff est for* $\theta$

2) *Let* $u(y)$ *be a fun of* $Y$. *then;* $X \sim Ber(1,\theta) \quad \Rightarrow Y = \sum X_i \sim Bin(n,\theta)$

$$g(y;\theta) = C^n_y\, \theta^y (1-\theta)^{n-y} \quad , y = 0,1,2,\ldots,n$$

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

$$E(u(Y)) = \sum_{x=0}^{n} u(y)\; g(y;\theta) = 0$$

$$\Rightarrow \sum_{x=0}^{n} u(y)\; C^n_y\, \theta^y (1-\theta)^{n-y} = 0$$

$$\Rightarrow u(0)\theta^0(1-\theta)^{n-0} + u(1)n\theta^1(1-\theta)^{n-1} + \ldots + u(n)\,\theta^n(1-\theta)^{n-n} = 0$$

$$\Rightarrow u(0)(1-\theta)^n + u(1)n\theta(1-\theta)^{n-1} + \ldots + u(n)\,\theta^n = 0$$

$\because \theta > 0 \quad \Rightarrow \theta \neq 0 \qquad , n \neq 0$

$\therefore u(0) = u(1) = u(2) = \ldots. = u(n) = 0 \quad \Rightarrow u(y) = 0 \ \ \forall y$

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

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

$$g(y;\theta) = C^n_y\, \theta^y (1-\theta)^{n-y} \quad , y = 0,1,2,\ldots,n \ , E(Binomial) = n\theta$$

$$E(Y) = \sum_{y=0}^{n} y\, C^n_y\, \theta^y (1-\theta)^{n-y} = n\theta$$

*or;* $E(Y) = E(\sum X_i) = n\,E(X) = n\theta \quad$ , *because* $X \sim Bernoulli$ , $E(X) = \theta$

$E(Y) = n\theta$

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

**Ex:** Let $X_1, X_2, \ldots, X_n$ is a rssn from Gamma distⁿ $\Gamma(4, \theta)$, $0 < \theta < \infty$. **1)** Show that $Y = \sum X_i$ is a complete sufficient estimator for $\theta$. **2)** Find the unique continuous function of Y which is the best estimator for $\theta$ (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 $\theta$, 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 a rssn*

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

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

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

$$\therefore Y = \sum K(x_i) = \sum X_i \ \text{ is sufficient estimator for } \theta.$$

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

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

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

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

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

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

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

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

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

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

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

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

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

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

**Ex:** Let $X_1$, $X_2$, …, $X_n$ denote a random sample of size $n > 2$ from a dist$^n$ with p.d.f. $f(x;\theta) = \theta e^{-\theta x} \ \ 0 < x < \infty$ , and $\theta > 0$. **1)** Show that $Y = \sum X_i$ is a complete sufficient estimator for $\theta$. **2)** Prove that (n – 1)/Y is the unique continuous function of Y which is the best estimator for $\theta$ (M.V.U.E).

**Sol:**

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

$\because$ *the range of X does not depend on* $\theta$,

*then we use exponential family to prove suff.*

$f(x;\theta) = \exp(\ln\{f(x;\theta)\}) = \exp(\ln(\theta) - \theta\,x)$

$\qquad = \exp(q(\theta) + S(x) + p(\theta)K(x)) \qquad , S(x) = 0$

*In assn;*

$f(x_1,\dots,x_n;\theta) = \exp(n\ln(\theta) - \theta \sum x_i)$

$\qquad = \exp(n\ln(\theta) - \theta \sum x_i).\exp(0)$

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

$\sum K(x_i) = \sum X_i \ , \ p(\theta) = -\theta \ , \ \sum S(x_i) = 0 \ , \ n\,q(\theta) = n\ln(\theta)$

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

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

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

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

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

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

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

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

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

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

$$X \sim Exp(1/\theta) \quad \Rightarrow E(X) = \frac{1}{\theta} \quad , V(X) = \frac{1}{\theta^2}$$

$$Y = \sum X_i \sim \Gamma(n, 1/\theta)$$

$$E\left(\frac{1}{Y}\right) = \int_0^\infty \frac{1}{y} \frac{1}{\Gamma(n)\left(\frac{1}{\theta}\right)^n} y^{n-1} e^{-\theta y} \, dy = \int_0^\infty \frac{1}{\Gamma(n)\left(\frac{1}{\theta}\right)^n} y^{(n-1)-1} e^{-\theta y} \, dy$$

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

$$= \frac{\theta^n}{\Gamma(n)} \frac{\Gamma(n-1)}{\theta^{n-1}} = \frac{\theta^n (n-2)!}{(n-1)! \, \theta^{n-1}} = \frac{\theta (n-2)!}{(n-1)(n-2)!} = \frac{\theta}{(n-1)}$$

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

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

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

**Sol:**

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

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

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

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

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

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

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

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

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

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

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

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

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

**Ex: (Functions of Parameter):** Let $X_1, X_2, \ldots, X_n$ denote a random sample from a dist[n] which is $N(0, \theta)$. Then $Y = \sum X_i^2$ is a sufficient estimator for $\theta$. Find the best estimator for $\theta^2$ (M.V.U.E).

**Sol:**

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

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

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

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

$$f(x\,;\theta) = \exp\big(p(\theta)K(x) + q(\theta) + S(x)\big)$$

*In a rssn;*

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

$$= \exp\left( -\frac{n}{2}\ln(2\pi\theta) - \frac{\sum x_i^2}{2\theta} \right)\exp(0)$$

$$= \exp\big(p(\theta)\sum K(x_i) + n\,q(\theta)\big) \times \exp\big(\sum S(x_i)\big)$$

$\Rightarrow \therefore\ Y = \sum K(x_i) = \sum X_i^2$ *is suff est for* $\theta$.

2) $X \sim N(0, \theta)$

$$Let; \quad Z = \frac{X - \mu}{\sigma} = \frac{X - 0}{\sqrt{\theta}} = \frac{X}{\sqrt{\theta}}$$

$$Z^2 = \frac{X^2}{\theta} \sim \chi^2_{(1)}$$

$$\sum Z^2 = \frac{\sum X_i^2}{\theta} = \frac{Y}{\theta} \sim \chi^2_{(n)}$$

$$g(y; n, \theta) = \frac{1}{\Gamma(n/2)2^{n/2}} (y/\theta)^{\frac{n}{2}-1} e^{-y/2\theta} \ , \ y > 0$$

*Let* $u(y)$ *be a con. fun. of* $Y$, *then;*

$E(u(Y)) = 0$

$$E(u(Y)) = \int_0^\infty u(y) \frac{1}{\Gamma(n/2)2^{n/2}} (y/\theta)^{\frac{n}{2}-1} e^{-y/2\theta}\, dy = 0 \qquad \} \times \Gamma(n/2)2^{n/2}$$

$$\Rightarrow \int_0^\infty u(y)\ (y/\theta)^{\frac{n}{2}-1} e^{-y/2\theta}\, dy = 0$$

$\because\ \theta, n > 0\ ,\ (y/\theta)^{\frac{n}{2}-1} > 0\ ,\ e^{-y/2\theta} > 0 \ne 0$

$\therefore\ u(y) = 0 \quad y > 0$

$\therefore\ g(y; n, \theta)$ *is complete*

3) $X \sim N(0,\theta)$

$$Let; \quad Z = \frac{X-\mu}{\sigma} = \frac{X-0}{\sqrt{\theta}} = \frac{X}{\sqrt{\theta}}$$

$$Z^2 = \frac{X^2}{\theta} \sim \chi^2_{(1)}$$

$$\sum Z^2 = \frac{\sum X_i^2}{\theta} = \frac{Y}{\theta} \sim \chi^2_{(n)}$$

$$V(Y) = E(Y^2) - (E(Y))^2$$

$$E(Y^2) = V(Y) + (E(Y))^2 \qquad , \quad E\left(\frac{Y}{\theta}\right) = n \quad \Rightarrow E(Y) = n\theta$$

$$, \quad V\left(\frac{Y}{\theta}\right) = 2n \quad \Rightarrow \frac{V(Y)}{\theta^2} = 2n \Rightarrow V(Y) = 2n\theta^2$$

$$\therefore E(Y^2) = 2n\theta^2 + n^2\theta^2$$

$$= \theta^2(2n + n^2)$$

$$\Rightarrow E\left(\frac{Y^2}{2n+n^2}\right) = \theta^2 \qquad \Rightarrow \quad \hat{\theta}^2 = \frac{Y^2}{2n+n^2} \quad is\ M.V.U.E.\ for\ \theta^2.$$

## The Rao- Cramer Inequality

Let $X_1, X_2, \ldots, X_n$ be a rss$n$ 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 $\theta$, E($T$) = $\theta$;

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

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

**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 $\theta$ iff;

$$V(T) = RCLB$$

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

**Sol:**

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

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

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

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

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

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

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

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

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

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

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

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

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

**Ex:** Let $X_1, X_2, \ldots, X_n$ be a rss$n$ from exponential dist$^n$ Exp($\theta$);

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

**2)** Find RCLB for each of [$\phi(\theta) = \ln\theta, \phi(\theta) = 2\theta$].

**Sol:**

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

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

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

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

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

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

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

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

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

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

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

$$\therefore RCLB = V(\overline{X})$$

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

2) $\phi(\theta) = \ln\theta \;\rightarrow\; \phi'(\theta) = \frac{1}{\theta}$

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

$$\phi(\theta) = 2\theta \;\rightarrow\; \phi'(\theta) = 2$$

$$RCLB\,(2\theta) = \frac{(\phi'(\theta))^2}{V(S)} = \frac{4}{n/\theta^2} = \frac{4\theta^2}{n}$$

**Ex:** In a rss$n$ from N($\theta$, $\sigma^2$). Show that;

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

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

**Sol:**

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

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

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

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

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

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

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

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

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

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

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

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

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

$$= (2\pi)^{-1/2}(\sigma^2)^{-1/2}\, e^{-\frac{1}{2\sigma^2}(x-\theta)^2}$$

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

$$\frac{\partial \ln f(x;\theta,\sigma^2)}{\partial \sigma^2} = zero - \frac{1}{2\sigma^2} + \frac{2(x-\theta)^2}{4(\sigma^2)^2} = -\frac{1}{2\sigma^2} + \frac{(x-\theta)^2}{2(\sigma^2)^2}$$

$$\frac{\partial^2 \ln f(x;\theta,\sigma^2)}{\partial (\sigma^2)^2} = \frac{2}{4(\sigma^2)^2} - \frac{(x-\theta)^2\, 4\sigma^2}{4(\sigma^2)^4} = \frac{1}{2(\sigma^2)^2} - \frac{(x-\theta)^2}{(\sigma^2)^3}$$

$$-n\,E\left(\frac{\partial^2 \ln f(x;\theta,\sigma^2)}{\partial (\sigma^2)^2}\right) = -n\,E\left(\frac{1}{2(\sigma^2)^2} - \frac{(x-\theta)^2}{(\sigma^2)^3}\right)$$

$$= \frac{-n}{2(\sigma^2)^2} + \frac{n\,E(x-\theta)^2}{(\sigma^2)^3} = \frac{-n}{2(\sigma^2)^2} + \frac{n\,\sigma^2}{(\sigma^2)^3}$$

$$= \frac{-n}{2(\sigma^2)^2} + \frac{n}{(\sigma^2)^2} = \frac{-n+2n}{2(\sigma^2)^2} = \frac{n}{2(\sigma^2)^2} = V(S) \; = F.I \; in \; arssn$$

$$\phi(\sigma^2) = \sigma^2 \quad \rightarrow \quad \phi'(\sigma^2) = 1$$

$$RCLB = \frac{\left(\phi'(\sigma^2)\right)^2}{F.I} = \frac{1}{n/2(\sigma^2)^2} = \frac{2(\sigma^2)^2}{n}$$

$$S^2 = \frac{\sum(X_i - \overline{X})^2}{n} \qquad , \qquad V(T = S^2) \overset{?}{=} RCLB \qquad , \qquad V(S^2) = ?$$

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

$$\therefore \quad \frac{X-\theta}{\sigma} \sim N(0,1)$$

$$\frac{(X-\theta)^2}{\sigma^2} \sim \chi^2_{(1)}$$

$$\frac{\sum(X_i-\theta)^2}{\sigma^2} \sim \chi^2_{(n)}$$

$$\frac{n\,S^2}{\sigma^2} \sim \chi^2_{(n)} \qquad , \quad \sum(X_i - \overline{X})^2 = \; n\,S^2$$

$$V\left(\frac{n\,S^2}{\sigma^2}\right) = V\left(\chi^2_{(n)}\right)$$

$$\frac{n^2\,V(S^2)}{\sigma^4} = 2n \qquad \Rightarrow V(S^2) = \frac{2n\sigma^4}{n^2} = \frac{2\,\sigma^4}{n}$$

$$\because \frac{2\,\sigma^4}{n} \; = \frac{2\sigma^4}{n}$$

$$\Rightarrow V(S^2) = RCLB$$

$$\therefore \; S^2 \; is \; an \; efficient \; estimator \; for \; \phi(\sigma^2) = \sigma^2 \; .$$

For $\quad S^2 = \dfrac{\sum(x_i - \overline{x})^2}{n-1}$

$$\frac{\sum(X_i-\theta)^2}{\sigma^2} \sim \chi^2_{(n-1)}$$

$$\frac{(n-1)\,S^2}{\sigma^2} \sim \chi^2_{(n-1)} \qquad , \quad \sum(X_i - \overline{X})^2 = \; (n-1)\,S^2$$

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

$$\frac{(n-1)^2\,V(S^2)}{\sigma^4} = 2(n-1) \qquad \Rightarrow V(S^2) = \frac{2(n-1)\sigma^4}{(n-1)^2} = \frac{2\,\sigma^4}{(n-1)}$$

$$\therefore eff(S^2) = \frac{RCLB}{V(S^2)} = \frac{2\,\sigma^4/n}{2\sigma^4/(n-1)} \; = \frac{2\sigma^4}{n} \times \frac{(n-1)}{2\sigma^4} = \frac{(n-1)}{n}$$

$$\Rightarrow \; eff(S^2) \; = \frac{(n-1)}{n} = \left(\frac{n-1}{n}\right) = \left(1 - \frac{1}{n}\right)$$

$$When \; n \rightarrow \infty \Rightarrow 1 - \frac{1}{n} \rightarrow 1 \qquad\qquad \therefore eff(S^2) \rightarrow 1$$

## 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:

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

**Proof:**

$$\begin{aligned}
MSE(\hat{\theta}) &= E(\hat{\theta} - \theta)^2 \\
&= E\left(\hat{\theta} - \theta \mp E(\hat{\theta})\right)^2 \\
&= E\left(\{\hat{\theta} - E(\hat{\theta})\} + \{E(\hat{\theta}) - \theta\}\right)^2 \\
&= E\left(\hat{\theta} - E(\hat{\theta})\right)^2 + E\left(E(\hat{\theta}) - \theta\right)^2 + 2\,E\left(\hat{\theta} - E(\hat{\theta})\right)\left(E(\hat{\theta}) - \theta\right) \\
&= Var(\hat{\theta}) + b^2(\hat{\theta}) + zero
\end{aligned}$$

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

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

## Relative Efficient Estimator

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

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

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

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

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

**Sol:**

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

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

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

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

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

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

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

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

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

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

$$\Rightarrow \therefore T_2 \text{ is more efficient than } T_1.$$

**Ex:** In a rssn from normal dist^n^ N($\theta$, $\sigma^2$), let $S_1^2 = \frac{\sum(X_i - \overline{X})^2}{n}$ and $S_2^2 = \frac{\sum(X_i - \overline{X})^2}{n-1}$ be two estimators for parameter $\sigma^2$, show that which of them more efficient.

**Sol:**

$$E(S_1^2) = \frac{n-1}{n}\sigma^2 \qquad \textit{biased}$$

$$E(S_2^2) = \sigma^2 \qquad \textit{unbiased}$$

$$b(S_1^2) = E(S_1^2) - \sigma^2$$

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

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

$$V(S^2) = ?$$

$$\frac{X-\theta}{\sigma} \sim N(0,1)$$

$$\frac{(X-\theta)^2}{\sigma^2} \sim \chi^2_{(1)}$$

$$\frac{\sum(X_i-\theta)^2}{\sigma^2} \sim \chi^2_{(n)}$$

$$\frac{\sum(X_i-\overline{X})^2}{\sigma^2} \sim \chi^2_{(n-1)} \quad , \because S^2 = \sum(X_i - \overline{X})^2 / n-1 \Rightarrow \therefore (n-1)\, S^2 = \sum(X_i - \overline{X})^2$$

$$\frac{(n-1)\, S^2}{\sigma^2} \sim \chi^2_{(n-1)}$$

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

$$\frac{(n-1)^2\, V(S^2)}{\sigma^4} = 2(n-1)$$

$$\therefore V(S^2) = 2(n-1)\frac{\sigma^4}{(n-1)^2} = \frac{2\sigma^4}{n-1} = V(S_2^2)$$

$$V(S_1^2) = \frac{2\sigma^4}{n} \quad , \quad V(S_2^2) = \frac{2\sigma^4}{n-1} = MSE(S_2^2)$$

$$MSE(S_1^2) = V(S_1^2) + b^2(S_1^2)$$

$$= \frac{2\sigma^4}{n} + \left(\frac{-\sigma^2}{n}\right)^2 = \frac{2\sigma^4}{n} + \frac{\sigma^4}{n^2} = \frac{2n\sigma^4 + \sigma^4}{n^2}$$

*Let;*

$$\frac{2\sigma^4}{n-1} > \frac{2n\sigma^4 + \sigma^4}{n^2}$$

$$\Rightarrow 2n^2\sigma^4 > (2n\sigma^4 + \sigma^4)(n-1)$$

$$\Rightarrow 2n^2\sigma^4 > 2n^2\sigma^4 + n\sigma^4 - 2n\sigma^4 - \sigma^4$$

$$\Rightarrow 2n^2\sigma^4 > 2n^2\sigma^4 - n\sigma^4 - \sigma^4$$

$$\Rightarrow \because MSE(S_2^2) > MSE(S_1^2) \qquad \Rightarrow \therefore S_1^2 \text{ is more efficient than } S_2^2$$

**Ex:** Given $f(x;\theta) = 1/\theta$ , $0 < x < \theta$ , *with* $\theta > 0$, formally compute the reciprocal of;

$$n\,E\left(\left\{\frac{\partial \ln f(X;\theta)}{\partial \theta}\right\}^2\right)$$

Compare this with the variance of ($n$ + 1) $Y_n$ / $n$, where $Y_n$ is the largest item of a random sample of size ($n$) from this distribution (*nth* order statistic)

**Sol 1:**

$$\ln f(x;\theta) = \ln(1) - \ln(\theta) = -\ln(\theta)$$

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

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

$$RCLB = \frac{1}{F.I} = \frac{1}{nE\left(\frac{\partial \ln f(x;\theta)}{\partial \theta}\right)^2} = \frac{1}{\frac{n}{\theta^2}} = \frac{\theta^2}{n}$$

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

$$V(Y_n) = E(Y_n^2) - (E(Y_n))^2$$

$$g(y_n;\theta) = n\,f(y_n)(F(y_n))^{n-1} \quad , \quad F(y_n) = p(X \le y_n) = \int_0^{y_n} \frac{1}{\theta}\,dx = \frac{y_n}{\theta} \quad , \quad f(y_n) = \frac{1}{\theta}$$

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

$$E(Y_n) = \int_0^{\theta} y_n \frac{n\,y_n^{n-1}}{\theta^n}\,dy_n = \frac{n}{\theta^n}\int_0^{\theta} y_n^n\,dy_n = \frac{n}{\theta^n}\left.\frac{y_n^{n+1}}{n+1}\right|_0^{\theta} = \frac{n}{\theta^n}\frac{\theta^{n+1}}{n+1} = \frac{n}{n+1}\theta$$

$$E(Y_n^2) = \int_0^\theta y_n^2 \; \frac{n\, y_n^{n-1}}{\theta^n} \, dy_n = \frac{n}{\theta^n} \int_0^\theta y_n^{n+1} \, dy_n = \frac{n}{\theta^n} \left. \frac{y_n^{n+2}}{n+2} \right|_0^\theta = \frac{n}{\theta^n} \, \frac{\theta^{n+2}}{n+2} = \frac{n}{n+2} \, \theta^2$$

$$\therefore V(Y_n) = \frac{n}{n+2} \, \theta^2 - \left( \frac{n}{n+1} \, \theta \right)^2 = \frac{n}{n+2} \, \theta^2 - \frac{n^2}{(n+1)^2} \, \theta^2$$

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

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

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

$$eff = \frac{RCLB}{V\left( \frac{n+1}{n} \, Y_n \right)} = \frac{\theta^2 / n}{\theta^2 / n(n+2)} = (n+2) > 1$$

$\because$ *the range depend on* $\theta$*, then we can't find the efficient est.*