Sufficiency

Definition:

Let $X_1, X_2, X_3, \ldots, X_n$ be a random sample from the density $f(x,\theta)$. A statistic / estimator $s(\underline{x})$ is sufficient for θ, if $f(x,\theta)$ can be factorized as follows.

$f(x,\theta) = g(s(\underline{x}),\theta)h(\underline{x})$

Where $h(\underline{x})$ is nonnegative and does not involve parameter.

Procedure for finding sufficient estimator:

1. Find the joint p.d.f i.e. likelihood function.
2. P.d.f of statistic (estimator) has to be proved sufficient.
3. Find the conditional p.d.f such as

$$p\left(\frac{\underline{x}}{S=s}\right) = \frac{h(\underline{x},\theta)}{g(S=s)} = \frac{h(\underline{x},\theta)}{g(s)}$$

If conditional p.d.f is independent from the parameter then estimator (statistic) is said to be sufficient.

OR

Nyman Fisher Criteria for Sufficiency:

1. Take the likelihood function of p.d.f.
2. If it can be written as the product of (function of statistic & parameter) and a function which is independent of parameter. Then statistic is said to be sufficient.

Therefore

$$L(\underline{x}) = g(s,\theta)h(\underline{x})$$

Bernoulli distribution:

Let we have Bernoulli distribution with parameter 'θ'.

$f(x) = \binom{1}{x}\ \theta^x\ (1-\theta)^{1-x}$

Applying Likelihood function

$L(\underline{x}) = \theta^{\sum x}\ (1-\theta)^{n-\sum x}\ \prod_{i=1}^{n}\binom{1}{x}$

$L(\underline{x}) \quad = g(\sum x,\theta)h(\underline{x})$

Hence $\sum x$ *is sufficient estimator of* 'θ'.

Pission distribution:

Let we have Pission distribution with parameter 'θ'.

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

Applying Likelihood function

$L(\underline{x}) = \dfrac{e^{-n\theta}\ \theta^{\sum x}}{\prod_{i=1}^{n} x!}$

$L(\underline{x}) = e^{-n\theta}\ \theta^{\sum x}\ \dfrac{1}{\prod_{i=1}^{n} x!}$

$L(\underline{x}) = g(\sum x,\theta)h(\underline{x})$

Hence $\sum x$ *is sufficient estimator of* 'θ'.

Normal distribution:

$$f(x)=\frac{1}{\delta\sqrt{2\Pi}}\ e^{-\frac{1}{2}(\frac{x-\mu}{\delta})^2}$$

Applying Likelihood function

$$L(\underline{x})=(\frac{1}{\delta\sqrt{2\Pi}})^n\ e^{-\frac{1}{2\delta^2}\sum(x-\mu)^2}$$

$$L(\underline{x})=(\frac{1}{\delta^2})^{\frac{n}{2}}\ e^{-\frac{\sum(x-\mu)^2}{2\delta^2}}\ (\frac{1}{\sqrt{2\Pi}})^n$$

$$L(\underline{x})=(\frac{1}{\delta})^n\ e^{-\frac{\sum(x-\mu)^2}{2\delta^2}}\prod_{i=1}^{n}(\frac{1}{\sqrt{2\Pi}})$$

$$L(\underline{x})=g(\Sigma(x-\mu),\delta^2)h(\underline{x})$$

Hence $\sum(x-\mu)^2$ *is sufficient estimator of* $'\delta^2'.$

Q.1:

Let $X_1, X_2, X_3,..., X_n$ be a random sample from the Bernoulli distribution with parameter θ. Then show that S= $\sum x_i$ is sufficient for θ.

Solution:

As x ~ Bernoulli (θ)

$$f(x)=\binom{1}{x}\ \theta^x\ (1-\theta)^{1-x}$$

Applying Likelihood function

$$L(\underline{x})=\theta^{\Sigma x}\ (1-\theta)^{n-\Sigma x}\ \prod_{i=1}^{n}\binom{1}{x}$$

$$L(\underline{x})\ =g(\Sigma x,\theta)h(\underline{x})$$

Hence $h(\underline{x})$ *is independent from parameter* $'\theta'$ *therefore for nyman fisher factorization* $\sum x$ *is sufficient estimator of* $'\theta'.$

Q. 2:

Show that for Possion distribution $\overline{X}$ is sufficient of θ.

Solution:

As x~P(θ)

Let we have Possion distribution with parameter ‘θ’.

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

Applying Likelihood function

$$L(\underline{x})=\frac{e^{-n\theta}\ \theta^{\Sigma x}}{\prod_{i=1}^{n}x!}$$

$$L(\underline{x})=e^{-n\theta}\ \theta^{\Sigma x}\ \frac{1}{\prod_{i=1}^{n}x!}$$

$$L(\underline{x})=g(\Sigma x,\theta)h(\underline{x})$$

Hence $h(\underline{x})$ *is independent of parameter* $'\theta'$ *therefore* $\sum x$ & $\overline{x}$ *is one to one function so* $\overline{x}$ *is sufficient of* $\theta.$

Question no 3:

Let $X_1, X_2, X_3,..., X_n$ be a random sample from a Normal distribution with population mean μ and variance δ^2.Then show that $\Sigma(x-\mu)^2$ is sufficient for δ^2.

Solution:

As x ~ N $((\mu, \delta^2)$

$$f(x) = \frac{1}{\delta\sqrt{2\Pi}} \quad e^{-\frac{1}{2}(\frac{x-\mu}{\delta})^2}$$

Applying Likelihood function

$$L(\underline{x}) = (\frac{1}{\delta\sqrt{2\Pi}})^n \quad e^{-\frac{1}{2\delta^2}\sum(x-\mu)^2}$$

$$L(\underline{x}) = (\frac{1}{\delta^2})^{\frac{n}{2}} \quad e^{-\frac{\sum(x-\mu)^2}{2\delta^2}} \quad (\frac{1}{\sqrt{2\Pi}})^n$$

$$L(\underline{x}) = (\frac{1}{\delta})^n \quad e^{-\frac{\sum(x-\mu)^2}{2\delta^2}} \prod_{i=1}^{n}(\frac{1}{\sqrt{2\Pi}})$$

$$L(\underline{x}) = g(\Sigma(x-\mu), \delta^2)h(\underline{x})$$

Hence $\sum(x-\mu)^2$ *is sufficient estimator of* $'\delta^2'$.

Question no 4:

Let $X_1, X_2, ..., X_n$ *be a random sample from a normal distribution with* $(0, \theta^2)$.

then show that $\frac{\sum x^2}{n}$ *is sufficient for* θ.

Solution:

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

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

Taking likelihood function

$$L(\underline{x}) = (\frac{1}{\sqrt{\theta}\sqrt{2\pi}})^n \quad e^{-\frac{1}{2\theta}\Sigma X^2}$$

$$L(\underline{x}) = (\frac{1}{\sqrt{\theta}})^n \quad e^{-\frac{\Sigma X^2}{2\theta}} \quad (\frac{1}{\sqrt{2\pi}})^n$$

$$L(\underline{x}) = g(\Sigma X^2, \ \theta) \ h(\underline{x})$$

Hence $h(\underline{x})$ *is independent of parameter* $'\theta'$ *therefore for Neyman Fisher factorization criteria* $\frac{\sum x^2}{n}$ *is sufficient estimator of* θ.

Question no 5:

Show that sample of $'n'$ *from a normal distribution with mean* $'\mu'$ *and standard deviation* $'\delta'$. *then estimator* $\delta^* = \sqrt{\frac{\sum(x-\mu)^2}{n}}$ *is sufficient estimator for* $'\delta'$

Solution:

As we know that

$$\delta^* = \sqrt{\frac{\sum(x-\mu)^2}{n}}$$

$$\delta^{2*} = \frac{\sum(x-\mu)^2}{n}$$

$$n\delta^{2*} = \sum(x-\mu)^2$$

$$f(x) = \frac{1}{\delta\sqrt{2\pi}} \quad e^{-\frac{1}{2\delta^2}(X-\mu^2)}$$

Applying likelihood function.

$$L(\underline{x}) = (\frac{1}{\delta\sqrt{2\pi}})^n \ e^{-\frac{1}{2\delta^2}\Sigma(X-\mu^2)}$$

$$L(\underline{x}) = (\frac{1}{\delta^2})^{\frac{n}{2}} \; e^{-\frac{n\delta^{2*}}{2\delta^2}} \; (\frac{1}{\sqrt{2\pi}})^n$$

$$L(\underline{x}) = g(\delta^{2*},\, \delta^2)\; h(\underline{x})$$

Hence $h(\underline{x})$ *is independent from parameter therefore for neyman fisher factorization criteria*

$\delta^* = \sqrt{\frac{\sum (x-\mu)^2}{n}}$ *is sufficient estimator for* 'δ'.

Question no 6:

Let $X_1 \,\&\, X_2$ *be a random sample of size* '$n=2$'. *From the following pission distribution with mean* 'θ'.*Show that* $S = X_1 + X_2$ *is sufficient estimator for* 'θ'.

Solution:

As $x \sim P(\theta)$

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

Applying likelihood function.

$$L(\underline{x}) = \frac{e^{-n\theta}\; \theta^{\Sigma x}}{\prod_{i=1}^{n} x!}$$

Therefore $n = 2$

$$L(\underline{x}) = \frac{e^{-2\theta}\; \theta^{x_1+x_2}}{x_1!\; x_2!} \qquad \therefore\; S = x_1 + x_2$$

$$L(\underline{x}) = e^{-2\theta}\; \theta^S\; \frac{1}{x_1!\; x_2!}$$

$$L(\underline{x}) = g(s,\theta)\; h(\underline{x})$$

As $h(\underline{x})$ *is independent from parameter* 'θ'. *Therefore for Neyman Fisher Factorization criteria* $S = x_1 + x_2$ *is sufficient estimator for* 'θ'.

Question no 7:

Let $X_1, X_2, ..., X_n$ *be a random sample of size* 'n' *from a Gamma distribution with* (P,δ). *then show that* $\prod_{i=1}^{n} x!$ & $\sum x$ *is sufficient for* P & δ.

Solution:

As $x \sim G(P,\delta)$

$$f(x) = \frac{1}{\Gamma P\, \delta^P}\; X^{P-1}\; e^{-\frac{x}{\delta}}$$

Applying likelihood function

$$L(\underline{x}) = (\frac{1}{\Gamma P\, \delta^P})^n \prod_{i=1}^{n} X^{P-1}\; e^{-\frac{\Sigma x}{\delta}}$$

$$L(\underline{x}) = (\frac{1}{\Gamma P\, \delta^P})^n\; e^{-\frac{\Sigma x}{\delta}} \left[\prod_{i=1}^{n} X_i\right]^P \left[\prod_{i=1}^{n} X_i\right]^{-1}$$

$$L(\underline{x}) = (\frac{1}{\Gamma P\, \delta^P})^n\; e^{-\frac{\Sigma x}{\delta}} \left[\prod_{i=1}^{n} X\right]^P \frac{1}{\left[\prod_{i=1}^{n} X\right]}$$

$$L(\underline{x}) = g\,(\prod_{i=1}^{n} X_i \;\&\; \sum x;\; P \;\&\; \delta)\; h(\underline{x})$$

Hence $h(\underline{x})$ *is independent from the paameter* $P \,\&\, \delta$.*Therefore by Neyman Fisher Factorization* $\prod_{i=1}^{n} x!$ & $\sum x$ *is joint sufficient for* $P \,\&\, \delta$.

Question no 8:

$Show\ that\ in\ estimator\ of\ '\theta'\ in\ f(x)=\frac{1}{\theta}\quad 0\le x\le\theta.\ Then\ the\ largest\ observation\ x_{(n)}\ is\ sufficient\ for\ '\theta'.$

Solution:

$As\ f(x)=\frac{1}{\theta}$

$Applying\ likelihood\ function.$

$$L(\underline{x})=\prod_{i=1}^{n} f(x;\theta)$$

$$L(\underline{x})=(\frac{1}{\theta})^{n}$$

$We\ know\ that\ P.d.f\ of\ rth\ order\ statistic$

$$g(x_{(r)})=\frac{n!}{(r-1)!\,(n-r)!}\left[F(x_{(r)})\right]^{r-1}\left[1-F(x_{(r)})\right]^{n-r} f(x_{(r)})$$

$Put\ n=n\ ,\ r=n$

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

$$g(x_{(n)})=n\left[F(x_{(n)})\right]^{n-1} f(x_{(n)}) \qquad A$$

Now

$$f(x)=\frac{1}{\theta}$$

$$F(x)=\int_0^x\frac{1}{\theta}dx$$

$$F(x)=\frac{1}{\theta}\int_0^x 1.dx$$

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

$$F(x_{(n)})=\frac{x_{(n)}}{\theta}$$

$$f(x_{(n)})=\frac{1}{\theta}$$

$$g(x_{(n)})=n\left[\frac{x_{(n)}}{\theta}\right]^{n-1}\frac{1}{\theta}$$

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

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

$As\ Conditional\ p.d.f$

$$f\left(\frac{\underline{x}}{x_{(n)}}\right)=\frac{f(x;\theta)}{g(x_{(n)})}$$

$$f\left(\frac{\underline{x}}{x_{(n)}}\right)=\frac{1/\theta^{n}}{n(x_{(n)})^{n-1}/\theta^{n}}=\frac{1}{n(x_{(n)})^{n-1}}$$

$Hence\ the\ conditional\ p.d.f\ is\ independent\ for\ parameter\ '\theta'.\ Therefore\ x_{(n)}\ is\ sufficient\ for\ '\theta'.$

Question no: 9:

Show that in estimator of $'\theta'$ *in* $f(x)=\frac{1}{\theta}\quad 0\le x\le\theta$. *Then the smallest observation* $x_{(1)}$ *is sufficient for* $'\theta'$.

Solution:

$$As\ \ f(x)=\frac{1}{\theta}$$

Applying likelihood function.

$$L(\underline{x})=\prod_{i=1}^{n} f(x;\,\theta)$$

$$L(\underline{x})=(\frac{1}{\theta})^{n}$$

We know that P.d.f of rth order statistic

$$g(x_{(r)})=\frac{n!}{(r-1)!\,(n-r)!}\left[F(x_{(r)})\right]^{r-1}\left[1-F(x_{(r)})\right]^{n-r} f(x_{(r)})$$

$$Put\ n=n\ ,\ r=1$$

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

$$g(x_{(n)})=(n-1)\left[1-F(x_{(n)})\right]^{n-1} f(x_{(n)}) \qquad A$$

Now

$$f(x)=\frac{1}{\theta}$$

$$F(x)=\int_0^x \frac{1}{\theta}dx$$

$$F(x)=\frac{1}{\theta}\int_0^x 1.dx$$

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

$$F(x_{(1)})=\frac{x_{(1)}}{\theta}$$

$$f(x_{(1)})=\frac{1}{\theta}$$

$$g(x_{(1)})=(n-1)\left[1-\frac{x_{(1)}}{\theta}\right]^{n-1}\frac{1}{\theta}$$

As Conditional p.d.f

$$f\left(\frac{\underline{x}}{x_{(1)}}\right)=\frac{f(x;\theta)}{g(x_{(1)})}$$

$$f\left(\frac{\underline{x}}{x_{(1)}}\right)=\frac{1/\theta^{n}}{(n-1)\left[1-\frac{x_{(1)}}{\theta}\right]^{n-1}\frac{1}{\theta}}$$

Hence the conditional p.d.f is not independent for parameter $'\theta'$. *Therefore* $x_{(1)}$ *is not sufficient for* $'\theta'$.

Question no: 10

Let $x_1, x_2, ..., x_n$ be a random sample from the normal density with mean $'\mu'$ & variance δ^2.

Then show that $\hat{\mu} = \overline{x}$ & $\delta^2 = \dfrac{\sum(x-\overline{x})^2}{n}$ are the set of two sufficient estimators for $'\mu'$ & δ^2

.

Solution:

As we know that

$x \sim N(\mu, \delta^2)$

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

Taking likelihood function

$$L(\underline{x}) = (\frac{1}{\delta\sqrt{2\pi}})^n\, e^{-\frac{1}{2\delta^2}\Sigma(x-\mu)^2} \qquad A$$

$$\therefore\ \Sigma(x-\mu)^2 = \Sigma(x-\overline{x}+\overline{x}-\mu)^2 = \Sigma[(x-\overline{x})^2+(\overline{x}-\mu)^2] = \Sigma[(x-\overline{x})^2+(\overline{x}-\mu)^2+2(x-\overline{x})(\overline{x}-\mu)]$$

$$\Sigma(x-\mu)^2 = \Sigma(x-\overline{x})^2 + n(\overline{x}-\mu)^2 + 2(\overline{x}-\mu)\Sigma(x-\overline{x})$$

$$\Sigma(x-\mu)^2 = \Sigma(x-\overline{x})^2 + n(\overline{x}-\mu)^2$$

$$\therefore\ \frac{\Sigma(x-\overline{x})^2}{n} = \hat{\delta}^2 = \Sigma(x-\overline{x})^2 = n\hat{\delta}^2$$

$$\Sigma(x-\mu)^2 = n\hat{\delta}^2 + n(\overline{x}-\mu)^2$$

$$L(\underline{x}) = (\frac{1}{\delta\sqrt{2\pi}})^n\, e^{-\frac{1}{2\delta^2}[n\hat{\delta}^2+n\,(\overline{x}-\mu)^2]}$$

$$L(\underline{x}) = (\frac{1}{\delta})^n\, e^{-\frac{n\hat{\delta}^2}{2\delta^2}}\, e^{-\frac{n\,(\overline{x}-\mu)^2}{2\delta^2}}\, (\frac{1}{\sqrt{2\pi}})^n$$

$$L(\underline{x}) = g\,(\hat{\delta}^2, \overline{x}\,;\,\delta^2, \mu).h(\underline{x})$$

As $h(\underline{x})$ *is independent from parameter* μ, δ^2 *therefore by Neyman Fisher Factorization* $\overline{x}$ & δ^2 *is a set of sufficient estimator for* μ & δ^2.

Question no 11:

Let 'n' be a fixed quantity and $X_{(1)}, X_{(2)}, X_{(3)}, ..., X_{(n)}$ *are n order statistic which are obtained from the* $f(x;\theta) = e^{-(x-\theta)}$ $\theta < x < \infty$. *Then the first order statistic is sufficient for* $'\theta'$.

Solution:

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

Taking likelihood function

$$L(\underline{x}) = \prod_{i=1}^{n} f(x;\,\theta)$$

$$L(\underline{x}) = e^{-\Sigma(x-\theta)} = e^{-(\Sigma x - n\theta)}$$

$$L(\underline{x}) = e^{-\Sigma x}\, e^{n\theta)}$$

We know that p.d.f of rth order statistic.

$$g\,(x_{(r)}) = \frac{n!}{(r-1)!\,(n-r)!}\left[F(x_{(r)})\right]^{r-1}\left[1-F(x_{(r)})\right]^{n-r} f\,(x_{(r)})$$

Put $n = n$, $r = 1$

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

$$g\,(x_{(1)}) = n\left[1-F(x_{(1)})\right]^{n-1}\, f\,(x_{(1)}) \qquad \text{A}$$

Now

$$\text{f(x)} = \text{e}^{\text{-x}}\, e^{\theta} \qquad \theta < x < \infty$$

$$F(x)=\int_{\theta}^{x} e^{-x}\ e^{\theta}\ dx$$

$$F(x)=e^{\theta}\int_{\theta}^{x} e^{-x}\ \ dx$$

$$F(x)=e^{\theta}\,\frac{e^{-x}}{-1}\Big|_{\theta}^{x}$$

$$F(x)=-e^{\theta}[\ e^{-x}\ -e^{\theta}\ ]$$

$$F(x)=-e^{\theta}e^{-x}+e^{-\theta+\theta}$$

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

$$F(x_{(1)})=1-e^{\theta}e^{-x(1)}$$

$$f(x_{(1)})=e^{-x_{(1)}}\ e^{\theta}$$

$$g(x_{(1)})=n\left[1-(1-e^{\theta}e^{-x_{(1)}})\right]^{n-1}\ (e^{-x_{(1)}}\ e^{\theta})$$

$$g(x_{(1)})=n\left[1-1+e^{\theta}e^{-x_{(1)}})\right]^{n-1}\ (e^{-x_{(1)}}\ e^{\theta})$$

$$g(x_{(1)})=n\left[e^{\theta}e^{-x_{(1)}})\right]^{n-1}\ (e^{-x_{(1)}}\ e^{\theta})$$

$$g(x_{(1)})=n e^{n\theta}\ e^{-xn}$$

As conditional density

$$f(\underline{x}/x_{(1)})=\frac{f(x;\theta)}{g(x_{(1)})}=\frac{e^{-\Sigma X}\ e^{n\theta}}{n e^{n\theta}\ e^{-xn}}=\frac{e^{-\Sigma X}}{ne^{-xn}}$$

As conditional p.d.f is independent of parameter 'θ' so first order statistic is sufficient for θ.

Question no 12:

Let 'n' be a fixed quantity and $X_{(1)}, X_{(2)}, X_{(3)}, ..., X_{(n)}$ *are n order statistic which are obtained from the* $f(x;\theta)=e^{-(x-\theta)}\quad \theta<x<\infty.$ *Then the largest order statistic is sufficient for 'θ'.*

Solution:

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

Taking likelihood function

$$L(\underline{x})=\prod_{i=1}^{n} f(x;\ \theta)$$

$$L(\underline{x})=e^{-\Sigma(x-\theta)}=e^{-(\Sigma x-n\theta)}$$

$$L(\underline{x})=e^{-\Sigma x}\ e^{n\theta)}$$

We know that p.d.f of rth order statistic.

$$g(x_{(r)})=\frac{n!}{(r-1)!\,(n-r)!}\left[F(x_{(r)})\right]^{r-1}\left[1-F(x_{(r)})\right]^{n-r}\ f(x_{(r)})$$

Put $n=n,\ r=n$

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

$$g(x_{(n)})=(n-1)\left[F(x_{(n)})\right]^{n-1}\ f(x_{(n)}) \qquad \text{(A)}$$

Now

$$f(x)=e^{-x}\ e^{\theta} \qquad \theta<x<\infty$$

$$F(x)=\int_{\theta}^{x} e^{-x}\ e^{\theta}\ dx$$

$$F(x)=e^{\theta}\int_{\theta}^{x} e^{-x}\ \ dx$$

$$F(x)=e^{\theta}\,\frac{e^{-x}}{-1}\Big|_{\theta}^{x}$$

$$F(x) = -e^{\theta}[\,e^{-x} - e^{\theta}\,]$$
$$F(x) = -e^{\theta}e^{-x} + e^{-\theta+\theta}$$
$$F(x) = 1 - e^{\theta}e^{-x}$$
$$F(x_{(n)}) = 1 - e^{\theta}e^{-x(n)}$$
$$f(x_{(n)}) = e^{-x_{(n)}}\ e^{\theta}$$
$$g(x_{(n)}) = n\left[1 - e^{\theta}e^{-x_{(n)}})\right]^{n-1}\ (e^{-x_{(n)}}\ e^{\theta})$$
$$g(x_{(n)}) = n\left[e^{\theta}e^{-x_{(n)}})\right]^{n-1}\ (e^{-x_{(n)}}\ e^{\theta})$$
$$g(x_{(n)}) = n\left[1 - e^{\theta}e^{-x_{(n)}})\right]^{n-1}\ (e^{-x_{(n)}}\ e^{\theta})$$

As conditional density

$$f(\underline{x}/x_{(n)}) = \frac{f(x;\theta)}{g(x_{(n)})} = \frac{e^{-\Sigma X}\ e^{n\theta}}{n\left[1 - e^{\theta}e^{-x_{(n)}})\right]^{n-1}\ (e^{-x_{(n)}}\ e^{\theta})}$$

As conditional p.d.f is not independent of parameter 'θ' so $X_{(n)}$ order statistic is not sufficient for θ.

Question no: 13

Let $X_1, X_2, ..., X_n$ be a random sample of size 'n' from the distribution $f(x;\theta).=\theta X^{\theta-1}$ then show that $\prod_{i=1}^{n} X_i$ is sufficient for θ.

As

$$f(x;\theta).=\theta X^{\theta-1}$$

Taking likelihood function

$$L(\underline{x}) = (\theta)^n\ (\prod_{i=1}^{n} X_i\ )^{\theta-1}$$

$$L(\underline{x}) = (\theta)^n \left[\prod_{i=1}^{n} X_i\right]^{\theta} \frac{1}{\left[\prod_{i=1}^{n} X_i\right]}$$

$$L(\underline{x}) = g(\prod_{i=1}^{n} X_i, \theta)\ h(\underline{x})$$

Hence $h(\underline{x})$ is independent from the parameter θ. Therefore by Nyman Fisher Factorization $\prod_{i=1}^{n} x_i$ is sufficient for θ . And $\prod_{i=1}^{n} x_i$ & G.M is one to one function so G.M is also sufficient for θ.

Question no 14:

Let $X_1 \& X_2$ be a random sample of size 'n=2'. From $N(\theta,1)$. Show that $Y_i = X_1 + X_2$ is sufficient estimator for 'θ'.

As $x \sim N(0,1)$

$$f(x) = \frac{1}{\delta\sqrt{2\pi}}\ e^{-\frac{1}{2}(x-\mu)^2} \qquad -\infty < x < +\infty$$

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

Applying likelihood function.

$$L(\underline{x})=(\frac{1}{\sqrt{2\pi}})^n\;e^{-\frac{1}{2}\Sigma(x-\theta)^2}$$

$$L(\underline{x})=(\frac{1}{\sqrt{2\pi}})^2\;e^{-\frac{1}{2}\Sigma(x-\theta)^2} \qquad \therefore n=2$$

$$\therefore\;\Sigma(x-\theta)^2=(x_1-\theta)^2+(x_2-\theta)^2$$

$$\Sigma(x-\theta)^2=x_1^{\;2}+\theta^{\,2}-2x_1\theta+x_2^{\;2}+\theta^{\,2}-2x_2\theta$$

$$\Sigma(x-\theta)^2=x_1^{\;2}+x_2^{\;2}+2\theta^{\,2}-2\theta(x_1+x_2)$$

$$\Sigma(x-\theta)^2=x_1^{\;2}+x_2^{\;2}+2\theta^{\,2}-2\theta Y$$

$$L(\underline{x})=(\frac{1}{\sqrt{2\pi}})^2\;e^{-\frac{1}{2}(x_1^{\;2}+x_2^{\;2}+2\theta^{\,2}-2\theta Y)}$$

$$L(\underline{x})=(\frac{1}{\sqrt{2\pi}})^2\;e^{-\frac{1}{2}(x_1^{\;2}+x_2^{\;2})}\;e^{-\theta^2}e^{\theta Y}$$

$$L(\underline{x})=(\frac{1}{\sqrt{2\pi}})^2\;e^{-\frac{1}{2}(x_1^{\;2}+x_2^{\;2})}\;e^{-\theta^2}e^{\theta Y}$$

$$L(\underline{x})=e^{-\theta^2}e^{\theta Y}(\frac{1}{\sqrt{2\pi}})^2\;e^{-\frac{1}{2}(x_1^{\;2}+x_2^{\;2})}$$

$$L(\underline{x})=g(Y=x_1+x_2\,;\,\theta)\;h(\underline{x})$$

Hence $h(\underline{x})$ is independent from parameter $'\theta'$. Therefore $Y=x_1+x_2$ is sufficient for $'\theta'$.

Minimal Sufficient Statistic:

A set of jointly sufficient statistic is said to be minimal if and only if it is the function of any other sufficient statistic

Example:

If $Y_{(1)}$ is sufficient statistic and $Y_{(n)}$ is also sufficient statistic and joint $(Y_{(1)}, Y_{(n)})$ is the sufficient for both $Y_{(1)}\,\&\,Y_{(n)}$.

Question no 15:

Let $Y_{(1)}, Y_{(2)}, Y_{(3)}, ..., Y_{(n)}$ be a random sample of size ‘n’ from the p.d.f $f(Y)=\dfrac{1}{2\lambda_2}$ $\lambda_1-\lambda_2\le Y\le\lambda_1+\lambda_2$. Then show that $Y_{(1)}\;\&\,Y_{(n)}$ are joint sufficient estimator for $\lambda_1\,\&\,\lambda_2$.

$$As \quad f(Y) = \frac{1}{2\lambda_2}$$

Taking likelihood function

$$L(\underline{Y}) = (\frac{1}{2\lambda_2})^n$$

$$L(\underline{Y}) = (\frac{1}{2^n \lambda_2{}^n}) = f(\underline{Y}; \lambda_1, \lambda_2)$$

We know that p.d.f of i^{th} & j^{th} order statistic.

$$g(y_i, y_j) = \frac{n!}{(i-1)!\,(n-j)!(j-i-1)!}\left[F(y_i)\right]^{i-1}\left[1-F(y_j)\right]^{n-j}\left[F(y_j)-F(y_i)\right]^{j-i-1} f(y_i)\, f(y_j)$$

$$put \quad i=1\,,\ j=n, n=n$$

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

$$g(y_1, y_n) = \frac{n(n-1)(n-2)!}{(n-2)!}\left[F(y_n)-F(y_1)\right]^{n-2} f(y_1)\, f(y_n)$$

$$g(y_1, y_n) = n(n-1)\left[F(y_n)-F(y_1)\right]^{n-2} f(y_1)\, f(y_n) \qquad \rightarrow A$$

$$As \quad f(x) = \frac{1}{2\lambda_2}$$

$$f(y_1) = \frac{1}{2\lambda_2} \qquad\qquad f(y_n) = \frac{1}{2\lambda_2}$$

$$F(x) = \int_{\lambda_1-\lambda_2}^{x} \frac{1}{2\lambda_2}\, dx$$

$$F(x) = \frac{1}{2\lambda_2}\int_{\lambda_1-\lambda_2}^{x} 1.\ dx$$

$$F(x) = \frac{1}{2\lambda_2}\, x \Big|_{\lambda_1-\lambda_2}^{x}$$

$$F(x) = \frac{1}{2\lambda_2}\left[x-(\lambda_1-\lambda_2)\right]$$

$$F(x) = \frac{1}{2\lambda_2}\left[x-\lambda_1+\lambda_2\right]$$

$$F(y_{(1)}) = \frac{1}{2\lambda_2}\left[y_{(1)}-\lambda_1+\lambda_2\right]$$

$$F(y_{(n)}) = \frac{1}{2\lambda_2}\left[y_{(n)}-\lambda_1+\lambda_2\right]$$

Put these values in equation (A).

$$g(y_1, y_n) = n(n-1)\left[\frac{[y_{(n)} - \lambda_1 + \lambda_2]}{2\lambda_2} - \frac{[y_{(1)} - \lambda_1 + \lambda_2]}{2\lambda_2}\right]^{n-2} \frac{1}{2\lambda_2} \cdot \frac{1}{2\lambda_2}$$

$$g(y_1, y_n) = \frac{n(n-1)}{(2\lambda_2)^2}\left[\frac{y_{(n)} - \lambda_1 + \lambda_2 - y_{(1)} + \lambda_1 - \lambda_2}{2\lambda_2}\right]^{n-2}$$

$$g(y_1, y_n) = \frac{n(n-1)}{(2\lambda_2)^2}\left[\frac{y_{(n)} - y_{(1)}}{2\lambda_2}\right]^{n-2}$$

$$g(y_1, y_n) = \frac{n(n-1)}{(2\lambda_2)^2} \frac{1}{(2\lambda_2)^{n-2}} (y_{(n)} - y_{(1)})^{n-2}$$

$$g(y_1, y_n) = \frac{n(n-1)}{(2\lambda_2)^n} (y_{(n)} - y_{(1)})^{n-2}$$

Now we consider.

$$f(\underline{y} / y_1, y_n) = \frac{f(\underline{y})}{g(y_1, y_n)} = \frac{1/2^n \; 1/\lambda_2{}^n}{\frac{n(n-1)}{(2\lambda_2)^n}(y_{(n)} - y_{(1)})^{n-2}}$$

$$f(\underline{y} / y_1, y_n) = \frac{1}{n(n-1)(y_{(n)} - y_{(1)})^{n-2}}$$

As conditional p.d.f is independent from the parameter $\lambda_1 \& \lambda_2$ therefore $y_1 \& y_n$ is joint sufficient for $\lambda_1 \& \lambda_2$

Question no 16:

Let $X_{(1)} < X_{(2)} < X_{(3)} < ... < X_{(n)}$ be a random sample of size 'n' from the p.d.f $f(X) = \frac{1}{\beta - \alpha}$ $\alpha < X < \beta$. Then show that $X_{(1)} \; \& \; X_{(n)}$ are jointly sufficient estimator for $\alpha \& \beta$.

$$As \quad f(X) = \frac{1}{\beta - \alpha}$$

Taking likelihood function

$$L(\underline{x}) = (\frac{1}{\beta - \alpha})^n$$

$$L(\underline{Y}) = \frac{1}{(\beta - \alpha)^n} = f(\underline{Y};\; \alpha \;,\; \beta)$$

We know that p.d.f of i^{th} & j^{th} order statistic.

$$g(x_i, x_j) = \frac{n!}{(i-1)!\,(n-j)!(j-i-1)!}[F(x_i)]^{i-1}[1 - F(x_j)]^{n-j}[F(x_j) - F(x_i)]^{j-i-1} f(x_i)\, f(x_j)$$

$$Put \; i = 1\,,\; j = n, n = n$$

$$g(x_1, x_n) = \frac{n!}{(1-1)!\,(n-n)!(n-1-1)!}[F(x_1)]^{1-1}[1 - F(x_n)]^{n-n}[F(x_n) - F(x_1)]^{n-1-1} f(x_1)\, f(x_n)$$

$$g(x_1, x_n) = \frac{n(n-1)(n-2)!}{(n-2)!}\left[F(x_n) - F(x_1\right]^{n-2} f(x_1)\, f(x_n)$$

$$g(x_1, x_n) = n(n-1)\left[F(x_n) - F(x_1\right]^{n-2} f(x_1)\, f(x_n) \qquad \rightarrow A$$

As

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

$$f(y_1) = \frac{1}{\beta - \alpha} \qquad\qquad f(y_n) = \frac{1}{\beta - \alpha}$$

$$F(x) = \int_\alpha^x \frac{1}{\beta - \alpha}\, dx$$

$$F(x) = \frac{1}{\beta - \alpha}\int_\alpha^x 1.\, dx$$

$$F(x) = \frac{1}{\beta - \alpha}\, x \Big|_\alpha^x$$

$$F(x) = \frac{1}{\beta - \alpha}\left[x - \alpha\right]$$

$$F(x_1) = \frac{x_{(n)} - \alpha}{\beta - \alpha}$$

$$F(x_{(n)}) = \frac{x_{(n)} - \alpha}{\beta - \alpha}$$

Put these values in equation (A).

$$g(x_1, x_n) = n(n-1)\left[F(x_n) - F(x_1)\right]^{n-2} f(x_1)\, f(x_n)$$

$$g(x_1, x_n) = n(n-1)\left[\frac{x_{(n)} - \alpha}{\beta - \alpha} - \frac{x_{(n)} - \alpha}{\beta - \alpha}\right]^{n-2} \frac{1}{\beta - \alpha} \cdot \frac{1}{\beta - \alpha}$$

$$g(x_1, x_n) = \frac{n(n-1)}{(\beta - \alpha)^2}\left[\frac{x_{(n)} - \alpha - x_{(1)} + \alpha}{\beta - \alpha}\right]^{n-2}$$

$$g(x_1, x_n) = \frac{n(n-1)}{(\beta - \alpha)^2}\left[\frac{x_{(n)} - x_{(1)}}{\beta - \alpha}\right]^{n-2}$$

$$g(x_1, x_n) = \frac{n(n-1)}{(\beta - \alpha)^2} \frac{1}{(\beta - \alpha)^{n-2}} (x_{(n)} - x_{(1)})^{n-2}$$

$$g(x_1, x_n) = \frac{n(n-1)}{(\beta - \alpha)^n} (x_{(n)} - x_{(1)})^{n-2}$$

Now we consider

$$f(\underline{x} \,/\, x_1, x_n) = \frac{f(\underline{x})}{g(x_1, x_n)}$$

$$f(\underline{x} / x_1, x_n) = \frac{1/(\beta-\alpha)^n}{\frac{n(n-1)}{(\beta-\alpha)^n}(x_{(n)} - x_{(1)})^{n-2}}$$

$$f(\underline{x} / x_1, x_n) = \frac{1}{n(n-1)(x_{(n)} - x_{(1)})^{n-2}}$$

As conditional p.d.f is independent from the parameter $\alpha \,\&\, \beta$ therefore $x_1 \,\&\, x_n$ is jointly sufficient estimator for $\alpha \,\&\, \beta$.

Question no 17:

Let $X_{(1)} < X_{(2)} < X_{(3)} < ... < X_{(n)}$ be a random sample of size 'n' from the uniform distribution with parameter (θ_1, θ_2) and $X_i = Mini(Y_1)$ and $X_n = Max(Y_n)$.

a. If $\theta_1 = 0$ show that Y_n is sufficient for θ_2.
b. If $\theta_1 \,\&\, \theta_2$ are unknown then show that (Y_1, Y_2) is sufficient for θ_1 *and* θ_2.

(a):

$$As \quad f(x, \theta_1, \theta_2) = \frac{1}{\theta_2 - \theta_1} \qquad \theta_1 < x < \theta_2$$

$$if \quad \theta_1 = 0$$

$$f(x, \theta_2) = \frac{1}{\theta_2} \qquad 0 < x < \theta_2$$

Taking likelihood function

$$L(\underline{x}) = (\frac{1}{\theta_2})^n$$

$$L(\underline{x}) = \frac{1}{(\theta_2)^n}$$

The p.d.f of i^{th} order statistic.

$$g(y_{(i)}) = \frac{n!}{(i-1)!\,(n-i)!}\left[F(y_{(i)})\right]^{r-1}\left[1 - F(y_{(i)})\right]^{n-r} f(y_{(i)})$$

$$Put \;\; n = n \;,\; i = n$$

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

$$g(y_{(n)}) = n\left[F(y_{(n)})\right]^{n-1} f(y_{(n)}) \qquad \rightarrow A$$

$$Now$$

$$\text{As } f(x) = \frac{1}{\theta^2}$$

$$F(x) = \int_0^x \frac{1}{\theta^2} dx$$

$$F(x) = \frac{1}{\theta^2} \int_0^x 1 \,.\, dx$$

$$F(x) = \frac{1}{\theta^2} x \Big|_0^x = \frac{x}{\theta^2}$$

$$F(y_n) = \frac{y_{(n)}}{\theta^2}$$

$$g(y_{(n)}) = n \left[\frac{y_{(n)}}{\theta_2}\right]^{n-1} \frac{1}{\theta_2}$$

$$g(y_{(n)}) = \left[\frac{(y_n)^{n-1}}{(\theta_2)^{n-1}}\right] \frac{n}{\theta_2}$$

$$g(y_{(n)}) = \frac{n(y_n)^{n-1}}{\theta_2{}^n}$$

Now we consider

$$f(\underline{x} / y_n) = \frac{L(\underline{x})}{g(y_n)}$$

$$f(\underline{x} / y_n) = \frac{1/\theta_2{}^n}{n(y_n)^{n-1}./\theta_2{}^n}$$

$$f(\underline{x} / y_n) = \frac{1}{n(y_n)^{n-1}}$$

As conditional p.d.f is independent from parameter θ_2 .therefore Y_n is sufficient for θ_2 .

(b):

$$As \quad f(x,\theta_1,\theta_2) = \frac{1}{\theta_2 - \theta_1} \qquad \theta_1 < x < \theta_2$$

Taking likelihood function

$$L(\underline{x}) = (\frac{1}{\theta_2 - \theta_1})^n$$

$$L(\underline{x}) = \frac{1}{(\theta_2 - \theta_1)^n}$$

We know that p.d.f of i^{th} & j^{th} order statistic.

$$g(y_i, y_j) = \frac{n!}{(i-1)!\,(n-j)!(j-i-1)!} \left[F(y_i)\right]^{i-1} \left[1 - F(y_j)\right]^{n-j} \left[F(y_j) - F(y_i)\right]^{j-i-1} f(y_i)\, f(y_j)$$

$$put \quad i = 1\,,\; j = n\,,\; \; n = n$$

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

$$g(y_1, y_n) = \frac{n(n-1)(n-2)!}{(n-2)!}\left[F(y_n)-F(y_1)\right]^{n-2} f(y_1)\, f(y_n)$$

$$g(y_1, y_n) = n(n-1)\left[F(y_n)-F(y_1)\right]^{n-2} f(y_1)\, f(y_n) \qquad \rightarrow A$$

As

$$f(x) = \frac{1}{\theta_2-\theta_1} \qquad \theta_1 < x < \theta_2$$

$$f(y_1) = \frac{1}{\theta_2-\theta_1} \qquad f(y_n) = \frac{1}{\theta_2-\theta_1}$$

$$F(x) = \int_{\theta_1}^{x} \frac{1}{\theta_2-\theta_1}\, dx$$

$$F(x) = \frac{1}{\theta_2-\theta_1}\int_{\theta_1}^{x} 1.\, dx$$

$$F(x) = \frac{1}{\theta_2-\theta_1}\, x \Big|_{\theta_1}^{x}$$

$$F(x) = \frac{1}{\theta_2-\theta_1}(x-\theta_1)$$

$$F(x) = \frac{x-\theta_1}{\theta_2-\theta_1}$$

$$F(y_{(1)}) = \frac{y_1-\theta_1}{\theta_2-\theta_1}$$

$$F(y_{(n)}) = \frac{y_n-\theta_1}{\theta_2-\theta_1}$$

Put these values in equation (A).

$$g(y_1, y_n) = n(n-1)\left[\frac{y_1-\theta_1}{\theta_2-\theta_1} - \frac{y_1-\theta_1}{\theta_2-\theta_1}\right]^{n-2} \frac{1}{\theta_2-\theta_1} \cdot \frac{1}{\theta_2-\theta_1}$$

$$g(y_1, y_n) = \frac{n(n-1)}{(\theta_2-\theta_1)^2}\left[\frac{y_n-\theta_1-y_1+\theta_1}{\theta_2-\theta_1}\right]^{n-2}$$

$$g(y_1, y_n) = \frac{n(n-1)}{(\theta_2-\theta_1)^2}\left[\frac{(y_n-y_1)^{n-2}}{(\theta_2-\theta_1)^{n-2}}\right]$$

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

Now we consider

$$f(\underline{x} \,/\, y_1, y_n) = \frac{f(\underline{x})}{g(y_1, y_n)}$$

$$f(\underline{x}\ /\ y_1, y_n) = \frac{1/(\theta_2-\theta_1)^n}{\frac{n(n-1)}{(\theta_2-\theta_1)^n}(y_{(n)}-y_{(1)})^{n-2}}$$

$$f(\underline{x}\ /\ y_1, y_n) = \frac{1}{n(n-1)(y_{(n)}-y_{(1)})^{n-2}}$$

As conditional p.d.f is independent from the parameter $\theta_1 \& \theta_2$ therefore $y_1 \& y_n$ is joint sufficient estimator for $\theta_1 \& \theta_2$

Nyman Fisher Factorization Theorem:

OR

Theorem of Sufficient Estimator:

Proof:

Suppose S=s is sufficient for $'\theta'$

Then by definition

$$p\left[\frac{\underline{x}}{S=s}\right] = h(\underline{x})$$

Where $h(\underline{x})$ is independent of parameter $'\theta'$

But

$$p\left[\frac{\underline{x}}{S=s}\right] = \frac{f(\underline{x};\theta)}{g(S=s)} = h(\underline{x}) \rightarrow A$$

$$p(\underline{x};\theta) = g(S=s)h(\underline{x})$$

Hence it is proved that “S=s” is sufficient for parameter $'\theta'$.

Then

$$p(\underline{x};\theta) = g(s(\underline{x}),\theta)h(\underline{x})$$

Where $h(\underline{x})$ is independent for the parameter $'\theta'$.

Now we suppose equation A hold that

$$p\left[\frac{\underline{x}}{S=s}\right] = \frac{f(\underline{x};\theta)}{g(S=s)}$$

$$g(S=s;\theta) = \Sigma\, p(\underline{x};\theta)$$

$$= g(s(\underline{x}),\theta)\Sigma h(\underline{x})$$

$$p\left[\frac{\underline{x}}{S=s}\right] = \frac{g(s(\underline{x});\theta)h(\underline{x})}{g(s(\underline{x});\theta)\Sigma h(\underline{x})}$$

$$p\left[\frac{\underline{x}}{S=s}\right]=\frac{h(\underline{x})}{\sum h(\underline{x})}$$

Which is independent for $'\theta'$. Hence by definition "S=s" is sufficient for $'\theta'$.

Note:

If

Where $f(\underline{x};\theta)$ is joint P.d.f and $g(S=s)$ is a function of's' then s is called sufficient for $'\theta'$.

Minimal Sufficient Statistic:

A set of jointly sufficient statistic is said to be minimal if and only if it is the function of any other sufficient statistic

Example:

If $Y_{(1)}$ is sufficient statistic and $Y_{(n)}$ is also sufficient statistic and joint $(Y_{(1)}, Y_{(n)})$ is the sufficient for both $Y_{(1)} \& Y_{(n)}$.

Prepared by: Muhammad Riaz
Lecturer: Statistics
The Islamia University Bahawalpur
Sub Campus
RYK