Research Methodology

Unit - IV

Sampling theorey

Definition:

A function of one or more random variable that does not depend upon any unknown parameter is called statistic

Example

Let $X_1, X_2, \ldots, X_n$ be n random variables having the joined p.d.f $f(x_1, x_2, \ldots, x_n)$

Let Y be the random variable defined by

$$Y = \frac{X_1, X_2, \ldots X_n}{n}$$

i.e, $$Y = \frac{\sum_{i=1}^{n} X_i}{n}$$

$\therefore$ Y is statistic ($\because$ Y does not depend upon any unknown parameter

Example

Let X be random variable and let $Y = \left(\frac{X-\mu}{\sigma}\right)^2$

Then Y is not satistic ($\because$ Y dependen upon μ and σ)

Random sample

Let $X_1, X_2, \ldots X_n$ be n mutually stochastically independent random variable each of which has the same but possible unknown p.d.f $f(x)$.

i.e, p.d.f of $X_1, X_2, \ldots X_n$ are respectively

$$f_1(x_1) = f(x_1)$$
$$f_2(x_2) = f(x_2)$$
$$\vdots$$
$$f_n(x_n) = f(x_n)$$

$\therefore$ The joined p.d.f of $X_1, X_2, \ldots X_n$ is

$$f(x_1, x_2, \ldots, x_n) = f_1(x_1)\, f_2(x_2) \ldots f_n(x_n)$$
$$= f(x_1)\, f(x_2) \ldots f(x_n)$$

Defn

The random variable $X_1, X_2, \ldots X_n$ are said to form a random sample of size n and $X_1, X_2, \ldots, X_n$ are called the items of random sample.

Defn

Let $X_1, X_2, \ldots, X_n$ be random sample of size n then $\bar{X} = \dfrac{X_1 + X_2 + \ldots + X_n}{n}$ is a statistic and it is called the mean of the random sample $S^2 = \dfrac{\sum_1^n (X_i - \bar{X})^2}{n}$ is also a statistic and it is called variance of the random sample

Show that $s^2 = \frac{1}{n}\sum_{i=1}^{n}(x_i-\bar{x})^2 = \left(\frac{1}{n}\sum_{i=1}^{n}x_i^2\right)-\bar{x}^2$

where $\bar{x} = \frac{1}{n}\sum_{i=1}^{n}x_i$

Soln

$$s^2 = \frac{1}{n}\sum_{i=1}^{n}(x_i-\bar{x})^2$$

$$= \frac{1}{n}\sum_{i=1}^{n}(x_i^2+\bar{x}^2-2x_i\bar{x})$$

$$= 1/n\sum_{i=1}^{n}x_i^2 + 1/n\sum_{i=1}^{n}\bar{x}^2 - 2/n\sum_{i=1}^{n}x_i\bar{x}$$

$$= 1/n\sum_{i=1}^{n}x_i^2 + \frac{\bar{x}^2}{n}\cdot n - 2/n\,\bar{x}\sum_{i=1}^{n}x_i$$

$$= 1/n\sum_{i=1}^{n}x_i^2 + \bar{x}^2 - 2\bar{x}^2$$

$$= 1/n\sum_{i=1}^{n}x_i^2 - \bar{x}^2$$

Distribution function technique

The method of finding the distribution of a function of one & more random variable is called the distribution function technique

ie, If $x_1, x_2, \ldots, x_n$ are n random variables and if $Y = g(x_1, x_2, \ldots, x_n)$ is function of the random variables $x_1, x_2, \ldots, x_n$ then the distribution function of y is obtained by $G(y) = P_r(Y \le y)$

$$= P_r(g(x_1, x_2, \ldots x_n) \le y)$$

Let x_1 and x_2 denote a random variable sample of size 2 from a distribution is $n(0,1)$ find the p.d.f of $Y = X_1^2 + X_2^2$ (Hint: In the douple integral representing $P(Y<y)$ use polar co-ordinates)

soln

Given X_1 and X_2 are $n(0,1)$ variables

The p.d.f of X_1 is $f_1(x_1) = \frac{1}{\sqrt{2\pi}} e^{-x_1^2/2}$, $-\infty < x < \infty$

The p.d.f of X_2 is $f_2(x_2) = \frac{1}{\sqrt{2\pi}} e^{-x_2^2/2}$, $-\infty < x < \infty$

Since X_1 and X_2 are stochastically independent we have $f(x_1, x_2) = f_1(x_1) \cdot f_2(x_2)$.

$\therefore$ The joint p.d.f of X_1 and X_2 is given by

$$f(x_1, x_2) = \frac{1}{2\pi} e^{-\left(\frac{x_1^2 + x_2^2}{2}\right)}, \quad -\infty < x_1 < \infty, \ -\infty < x_2 < \infty$$

Let $Y = X_1^2 + X_2^2$

Let $G(y)$ be the distribution function of Y

then $G(y) = Pr(Y \leq y)$

$= Pr(X_1^2 + X_2^2 \leq y)$

$= \iint_A f(x_1, x_2)\, dx_1\, dx_2$ where $A = \{(x_1, x_2) / x_1^2 + x_2^2 \leq y\}$

$= \iint_A \frac{1}{2\pi} e^{-\frac{1}{2}(x_1^2 + x_2^2)}\, dx_1\, dx_2$

changing into polar co-ordinates putting $x_1 = r\cos\theta$, $x_2 = r\sin\theta$.

$dx_1\, dx_2 = r\, dr\, d\theta$ where $|J| = r$

$x_1^2 + x_2^2 = r^2$

Then $0 \le r \le \sqrt{y}$, $0 \le \theta \le 2\pi$

$$G(y) = \int_0^{2\pi}\int_0^{\sqrt{y}} \frac{1}{2\pi} e^{-r^2/2}\, r\, dr\, d\theta.$$

$$= \int_0^{2\pi} d\theta \int_0^{\sqrt{y}} \frac{1}{2\pi} e^{-r^2/2}\, r.dr$$

$$= 2\pi \frac{1}{2\pi} \int_0^{\sqrt{y}} e^{-r^2/2}\, r\, dr.$$

put $t = r^2 \Rightarrow dt = 2r dr$

when $r = 0 \Rightarrow t = 0$

when $r = \sqrt{y} \Rightarrow t = y$.

$$\therefore G(y) = \int_0^y e^{-t/2} \frac{dt}{2}$$

$$= \frac{1}{2} \int_0^y e^{-t/2} dt$$

$$= \int_0^y \frac{1}{\Gamma 2/2\ 2^{2/2}} t^{2/2 - 1} e^{-t/2} dt$$

the p.d.f of Y is $g(y) = G'(y)$

$$g(y) = \frac{1}{\Gamma 2/2\ 2^{2/2}} t^{2/2-1} e^{-t/2}$$

Thus Y is $\chi^2(2)$

Note

Let $X_1, X_2, \ldots, X_n$ denote the random same of size n from a distribution is $n(0,1)$. Then $Y = X_1^2 + X_2^2 + \ldots + X_n^2$ is a $\chi^2(n)$ variable

Transformation of vareable of the discrete type

An alternating method of finding the

distribution function of one or more random variable & the change variable technique

Let X be a random variable of the discrete type having p.d.f of $f(x)$

Let $\mathcal{A}$ denote the set of discrete points at each of which $f(x)>0$ and let $y=u(x)$. The space $\mathcal{B}$ is obtained to transforming each point in $\mathcal{A}$ in according with define a 1-1 transformation that maps $\mathcal{A}$ onto $\mathcal{B}$. If we slove $y=u(x)$ for x interms of y say $x=w(y)$ then for each $y\in\mathcal{B}$ we have $x=w(y)\in\mathcal{A}$

Consider the random variable $y=u(x)$. If $y\in\mathcal{B}$ then $x=w(y)\in\mathcal{A}$ and the events $Y=y$ or $u(x)=y$ and $x=w(y)$ are equivalent

According the p.d.f of Y is $g(y)=P_r(Y\leq y)$

$$= P_r(X=w(y))$$

$$g(y)=f(w(y)),\ y\in\mathcal{B}$$

Let x have a binomial p.d.f $f(x)=\begin{cases}\frac{3!}{x!(3-x)!}\left(\frac{2}{3}\right)^x\left(\frac{1}{3}\right)^{3-x}, & x=0,1,2,3\\ 0 & \text{elsewhere}\end{cases}$

Find the p.d.f of $Y=X^2$.

Soln

Given $f(x)=\begin{cases}\frac{3!}{x!(3-x)!}(2/3)^x(1/3)^{3-x}, & x=0,1,2,3\ldots\\ 0 & \text{elsewhere}\end{cases}$

Here $A = \{x / x = 0, 1, 2, 3\}$

$B = \{y / y = 0, 1, 4, 9\}$

Since there is no negative value of x in A

$y = x^2$

Define 1-1 transformation from A onto B

$\therefore$ The inverse function of $y = x^2 \Rightarrow x = \sqrt{y}$.

$\therefore$ The p.d.f of Y is $g(y) = f(\omega(y)) = f(\sqrt{y})$

$$g(y) = \begin{cases} \dfrac{3!}{\sqrt{y}!(3-\sqrt{y})!} (2/3)^{\sqrt{y}} (1/3)^{3-\sqrt{y}} & , y = 0, 1, 4, 9 \\ 0 & \text{elsewhere} \end{cases}$$

Let X have the p.d.f $f(x) = \begin{cases} 1/3 & , x = 1, 2, 3. \\ 0 & \text{elsewhere} \end{cases}$. Find the p.d.f of $Y = 2X+1$

Soln

Here $A = \{x / x = 1, 2, 3\}$

$B = \{y / y = 3, 5, 7\}$

$\therefore y = 2x+1$ defines the 1-1 transformation A onto B.

Now, $y = 2x+1 \Rightarrow x = \dfrac{y-1}{2} = \omega(y)$

$\therefore$ the p.d.f of Y is $g(y) = f(\omega(y)) = f\left(\frac{y-1}{2}\right)$

$$= \begin{cases} 1/3 & , y = 3, 5, 7 \\ 0 & \text{elsewhere} \end{cases}$$

Let X have a poison p.d.f $f(x) = \begin{cases} e^{-\mu} \mu^x & , x = 0, 1, 2, \ldots \\ 0 & \text{elsewhere} \end{cases}$

Find the p.d.f of $Y = 4X$.

Given $f(x) = \begin{cases} \dfrac{e^{-\mu} \mu^{x}}{x!} & , x = 0,1,2,\ldots \\ 0 & \text{elsewhere} \end{cases}$

Here $A = \{x \mid x = 0,1,2,\ldots\}$

$B = \{y \mid y = 0,4,8,\ldots\}$

$Y = 4X$ defines the 1-1 transformation A onto B

Now. $y = 4x \Rightarrow x = y/4 = w(y)$

The p.d.f of Y is $g(y) = f(w(y)) = f(y/4)$

$$= \begin{cases} \dfrac{e^{-\mu} \mu^{y/4}}{(y/4)!} & , y = 0,4,8 \\ 0 & \text{elsewhere} \end{cases}$$

Let X have the p.d.f $f(x) = \begin{cases} (1/2)^{x} & , x = 1,2,3,\ldots \\ 0 & \text{elsewhere} \end{cases}$

Find the p.d.f of $Y = X^3$.

Soln Given $f(x) = \begin{cases} (1/2)^{x} & , x = 1,2,3,\ldots \\ 0 & \text{elsewhere} \end{cases}$

Here $A = \{x \mid x = 1,2,3,\ldots\}$

and $B = \{y \mid y = 1,8,27,\ldots\}$

$y = x^3$ defines a 1-1 transformation A onto B

Now, $y = x^3 \Rightarrow x = y^{1/3} = w(y)$

The p.d.f of Y is $g(y) = f(w(y))$

$= f(y^{1/3})$

$$g(y) = \begin{cases} (1/2)^{y^{1/3}} & , y = 1, 8, 27 \\ 0 & \text{elsewhere} \end{cases}$$

Transformation of two demensional random variable of discrete type

Let X_1, X_2 be two random variables of discrete type with joint p.d.f $f(x_1, x_2)$

Let $y_1 = u_1(x_1, x_2)$ and $y_2 = u_2(x_1, x_2)$

Define a 1-1 transformation that maps $\mathcal{A}$ onto $\mathcal{B}$.

$\therefore$ The joint p.d.f of two new r random variables $Y_1 = u_1(X_1, X_2)$ and $Y_2 = u_2(X_1, X_2)$ is given by

$$g(y_1, y_2) = \begin{cases} f(w_1(y_1, y_2), w_2(y_1, y_2)), & y_1, y_2 \in \mathcal{B} \\ 0 & \text{elsewhere} \end{cases}$$

the marginal p.d.f of Y_1 is $g(y_1) = \sum_{y_2} g(y_1, y_2)$

The marginal p.d.f of Y_2 is $g(y_2) = \sum_{y_1} g(y_1, y_2)$

Let X_1, X_2 be two stochastically independent random variable that have poison distribution with mean μ_1 and μ_2 respectively. Find the distribution of the random variable $Y = X_1 + X_2$ (or) Let X_1 and X_2 are independent poison variable with mean μ_1 ~~mean variable~~ and μ_2. Then $X_1 + X_2$ is also a poison variable with mean $\mu_1 + \mu_2$.

Given X_1 is a poison random variable with parameter μ_1.

It's p.d.f is $f_1(x_1) = \begin{cases} \dfrac{e^{-\mu}\mu_1^{x_1}}{x!} & , x_1 = 0,1,2,\ldots \\ 0 & \text{elsewhere} \end{cases}$

Given X_2 is a poison random variable with parameter μ_2.

It's p.d.f is $f_2(x_2) = \begin{cases} \dfrac{e^{-\mu_2}\mu_2^{x_2}}{x_2!} & , x_2 = 0,1,2,\ldots \\ 0 & \text{elsewhere} \end{cases}$

since X_1, X_2 are stochastically independent

we have $f(x_1, x_2) = f(x_1) f(x_2)$

$$= \begin{cases} e^{-(\mu_1+\mu_2)} \mu_1^{x_1} \mu_2^{x_2} & , x_1 = 0,1,2,\ldots , x_2 = 0,1,2,\ldots \\ 0 & \text{elsewhere} \end{cases}$$

Let $Y_1 = X_1 + X_2$

choose $Y_2 = X_2$ such that this transformation is a 1-1 transformation from $\mathcal{A}$ onto $\mathcal{B}$.

here $\mathcal{A} = \left\{ (x_1, x_2) \middle| \begin{array}{l} x_1 = 0,1,2,\ldots \\ x_2 = 0,1,2,\ldots \end{array} \right\}$

$\mathcal{B} = \left\{ (y_1, y_2) \middle| \begin{array}{l} y_1 = 0,1,2,\ldots \\ y_2 = 0,1,2,\ldots \end{array} \right\}$

$x_2 = y_2 = \omega_2(y_1, y_2)$

$x_1 = y_1 - y_2 = y_1(y_1, y_2)$

$\therefore g(y_1, y_2) = f(w_1(y_1, y_2), w_2(y_1, y_2))$

$= f(y_1 - y_2, y_2)$

$$= \begin{cases} \dfrac{e^{-(\mu_1+\mu_2)} \mu_1^{y_1-y_2} \mu_2^{y_2}}{(y_1-y_2)!\, y_2!}, & y_1 = 0, 1, 2, \ldots \\ & y_2 = 0, 1, 2, \ldots, y_1 \\ 0 & \text{elsewhere} \end{cases}$$

The marginal p.d.f of Y_1 is $g(Y_1) = \sum_{y_2=0}^{y_1} g(y_1, y_2)$

$$= \sum_{y_2=0}^{y_1} \frac{e^{-(\mu_1+\mu_2)} \mu_1^{y_1-y_2} \mu_2^{y_2}}{(y_1-y_2)!\, y_2!}$$

Multiply and divided by $y_1!$

$$\therefore g(y_1) = \frac{\sum_{y_2=0}^{y_1} e^{-(\mu_1+\mu_2)} \mu_1^{y_1-y_2} \mu_2^{y_2}\, y_1!}{(y_1-y_2)!\, y_2!\, y_1!}$$

$$= \frac{e^{-(\mu_1+\mu_2)} (\mu_1+\mu_2)^{y_1}}{y_1!}$$

which is the p.d.f of poison distribution with parameter μ_1 to μ_2

Hence $y_1 = x_1 + x_2$ has the poison distribution with parameter μ_1 and μ_2.

Transformation of variable of the continuous type

Let X be a random variable of the continuous type with p.d.f $f(x)$

Let $\mathcal{A} = \{x \mid f(x) > 0\}$ be a space of X.

Let $Y = u(X)$ be 1-1 transformation that maps $\mathcal{A}$ onto $\mathcal{B}$ where $\mathcal{B} = \{y \mid y = u(x), x \in \mathcal{A}\}$

Let $x = w(y)$ be its inverse $\frac{dx}{dy} = w'(y)$

Then the p.d.f of Y is given by $g(y) = P_r(y \in \mathcal{B})$

$$g(y) = P_r(x \in \mathcal{A})$$

$$= f(x, (y)) |J|$$

$$g(y) = f(w(y)) |w'(y)|$$

Note

$w'(y)$ is called the Jacobian of inverse transformation $X = w(y)$ and it is denoted by J.

Let X have p.d.f $f(x) = \begin{cases} 1 & \text{if } 0 < x < 1 \\ 0 & \text{elsewhere} \end{cases}$. Let $Y = -2 \log X$.

Find the p.d.f of Y.

Soln

Given the p.d.f of X is $f(x) = \begin{cases} 1 & \text{if } 0 < x < 1 \\ 0 & \text{elsewhere} \end{cases}$

here $\mathcal{A} = \{x \mid 0 < x < 1\}$ and $\mathcal{B} = \{y \mid 0 < y < \infty$

$$y = -2 \log x$$

$$x = e^{-y/2} = w(y)$$

$$\therefore w'(y) = (-1/2)\, e^{-y/2}$$

$\therefore$ The p.d.f of Y is $g(y) = f(w(y)) \, |w'(y)|$

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

$= f(e^{-y/2}) \, [(1/2) e^{-y/2}]$

$= (1) \, (1/2 \, e^{-y/2})$

$= \dfrac{e^{-y/2}}{2}$

$$g(y) = \frac{1}{\Gamma 2/2 \, 2^{2/2}} \, y^{2/2 - 1} \, e^{-y/2} \quad , 0 < y < \infty$$

which is the p.d.f of $\chi^2(2)$

Let X be a continuous random variable with p.d.f $f(x) = \begin{cases} x/12 & , 1 < x < 5 \\ 0 & \text{elsewhere} \end{cases}$. Find the p.d.f $Y = 2X - 3$

Soln.

Given the p.d.f of X is $f(x) = \begin{cases} x/12 & , 1 < x < 5 \\ 0 & \text{elsewhere} \end{cases}$

here $A = \{x \,/\, 1 < x < 5\}$ and $B = \{y \,/\, -1 < y < 7\}$

Now, $y = 2x - 3 \Rightarrow y + 3 = 2x$

$x = \dfrac{y+3}{2} = w(y)$

$\dfrac{dx}{dy} = 1/2 = w'(y)$

the p.d.f of Y is $g(y) = f(w(y)) \, |w'(y)|$

$= f\left(\dfrac{y+3}{2}\right) |1/2|$

$= (1/2) \, \dfrac{y+3}{2 \times 12}$

$$g(y) = \begin{cases} \dfrac{y+3}{48} & , -1 < y < 7 \\ 0 & \text{elsewhere} \end{cases}$$