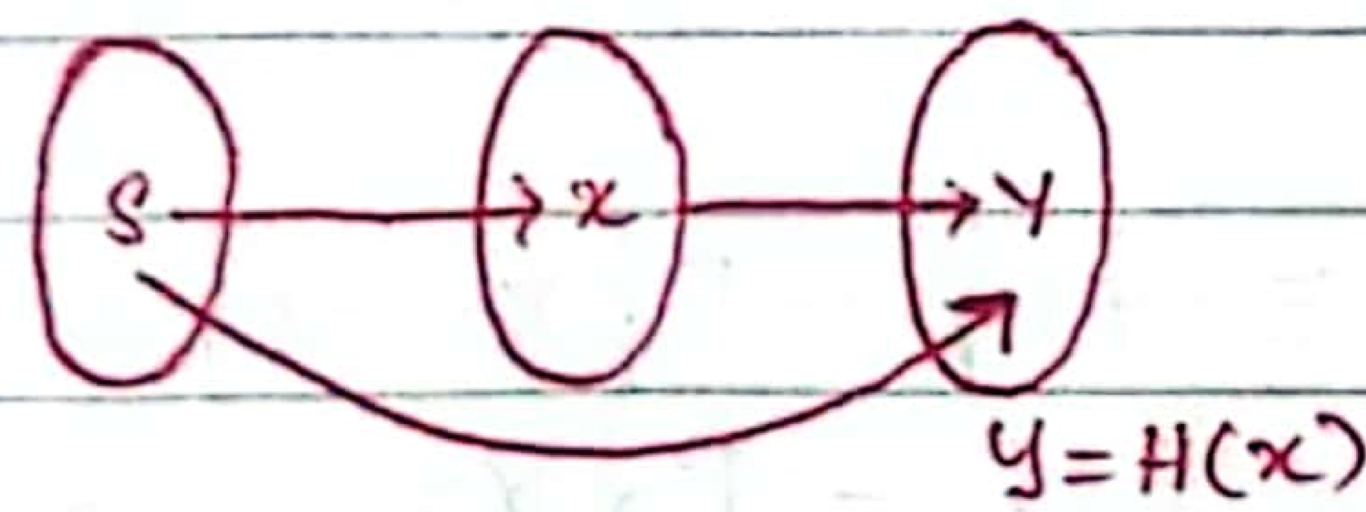
FUNCTIONS OF RANDOM VARIABLE.

1- Dimensional.

Let 'S' be the sample space associated with a random [variable] experiment 'E'. Let 'X' be the random variable defined on 'S'.

Let $Y = H(x)$ be a real valued function of X, then $Y = H(x)$ also defines a random variable on 'S' for every $s \in S$, there is a unique x and for every x there is a ~~unique~~ $Y = H(x)$.



Discrete Random Variable:

Suppose X is DRV, we define $Y = H(x)$, then it follows that Y is also a DRV with $H(x_1)$, $H(x_2)$ . . . .

→ X DRV Y(X) DRV

→ X CRV Y(X) DRV, Y(X) CRV.

Continuous Random Variable:

If X is a CRV, and let $Y = H(X)$, then Y is also CRV. Then pdf of Y is given by,

(i) $g(y) = \frac{d}{dy} G(y)$ where $G(Y)$ is the cdf of

ii) $g(y) \geq 0$.

If X – DRV $P(x_j) = F(x_j) - F(x_{j-1})$

X – CRV $f(x) = \frac{d}{dx} F(x) \ \forall x$.

classmate

Date
Page

Some examples:

1. If X takes the values $-1, 0, 1$ with probabili $1/3, 1/2, 1/6$ respectively. Let $Y = 3X+1$, then Y takes the values, $-2, 1, 4$.

Solution:

X:	-1	0	1
p(x):	1/3	1/2	1/6

Y:	-2	1	4
p(y):	1/3	1/2	1/6

$$P\{Y=3X+1\} = P\{X=-1\}$$
$$= P\{Y=-2\} = \underline{\underline{1/3}}$$

$$P\{Y=3(0)+1\} = P\{Y=1\} = P\{X=0\} = \underline{\underline{1/2}}$$

2. If X is a DRV which takes the values $-1, 0, 1$. with probabilities, $1/3, 1/2, 1/6$

Now suppose we have one random variable $Y = X^2$ which is interms of function of Then pmf of Y is given by,

$Y = X^2$:	1	0	1 OR we can write 0, 1
p(y) :	Neglect	1/2	$1/3 + 1/6 = \underline{\underline{1/2}}$

3. If X is a DRV which takes the values, $X = 1, 2, 3 \ldots n, \ldots$ and $P\{X=n\} = \frac{1}{2^n}$

Let $Y = \begin{cases} 1, & X \text{ is even} \\ -1, & X \text{ is odd} \end{cases}$

Then $P(Y=1) = P(X=2) + P(X=4) + \cdots$

$$= \frac{1}{2^2} + \frac{1}{2^4} + \frac{1}{2^6} + \cdots$$

$$= \frac{1/4}{1-1/4} = \underline{\underline{\frac{1}{3}}}$$

$P(Y=-1) = 1 - P(Y=1)$

$$= \underline{\underline{2/3}}$$

Procedure to find pdf of function of a random variable

If X is a random variable with pdf $f(x)$, Let $Y = H(X)$ be a random variable which is function of X. Then to find pdf of Y

→ We find cdf of Y i.e, $G(y) = P\{Y \leq y\}$

→ Then find pdf of Y by using the concept.

$$\Rightarrow g(y) = \frac{d}{dy}(G(y))$$

→ Find the range of Y.

[cdf of X is $F(x) = P\{X \leq x\}$
pdf of X is $\frac{d}{dx}F(x) = f(x)$] given in cumulative distribution function.

Problems

1\. Let X be a random variable having pdf

$$f(x) = \begin{cases} 2x, & 0 \le x \le 1 \\ 0, & \text{elsewhere.} \end{cases}$$

then find pdf of $Y = 3X+1$.

Solution:

$$\rightarrow G(y) = P(Y \le y)$$

$$= P(3X+1 \le y)$$

$$= P\left(X \le \frac{y-1}{3}\right)$$

* For continuous random variable its given by $P\{X \le x\} = \int_{-\infty}^{x} f(x)\,dx$. *

$$= \int_{-\infty}^{y-1/3} f(x)\,dx$$

$$= \int_{0}^{y-1/3} 2x\,dx = x^2 \Big|_0^{\frac{y-1}{3}}$$

$$G(y) = \left(\frac{y-1}{3}\right)^2 = \boxed{\frac{(y-1)^2}{9}}$$

$\therefore$ pdf is represented as $g(y)$

$$\Rightarrow g(y) = \frac{d}{dy}(G(y))$$

$$= \frac{d}{dy}\frac{(y-1)^2}{9} = \boxed{\frac{1}{9}\,2(y-1)}$$

$0 \le x \le 1 \rightarrow$	0	0.2	0.4	0.6	0.8	1
$Y = 3x+1 =$	1	1.6	2.2	2.8	3.4	4
$x = \frac{y-1}{3} =$	0	0.2	0.4	0.6	0.8	1

Problems

1. Let X be a random variable having pdf

$$f(x) = \begin{cases} 2x, & 0 \le x \le 1 \\ 0, & \text{elsewhere.} \end{cases}$$

then find pdf of $Y = 3X+1$.

Solution:

$$\rightarrow G(y) = P(Y \le y)$$

$$= P(3X+1 \le y)$$

$$= P\left(X \le \frac{y-1}{3}\right)$$

* For continuous random variable its given by $P\{X \le x\} = \int_{-\infty}^{x} f(x)\,dx$. *

$$= \int_{-\infty}^{y-1/3} f(x)\,dx$$

$$= \int_{0}^{y-1/3} 2x\,dx = x^2 \Big|_{0}^{\frac{y-1}{3}}$$

$$G(y) = \left(\frac{y-1}{3}\right)^2 = \boxed{\frac{(y-1)^2}{9}}$$

$\therefore$ pdf is represented as $g(y)$

$$\Rightarrow g(y) = \frac{d}{dy}(G(y))$$

$$= \frac{d}{dy}\frac{(y-1)^2}{9} = \boxed{\frac{1}{9}\,2(y-1)}$$

$0 \le x \le 1 \rightarrow$	0	0.2	0.4	0.6	0.8	1
$Y = 3X+1 =$	1	1.6	2.2	2.8	3.4	4
$x = \frac{y-1}{3} =$	0	0.2	0.4	0.6	0.8	1

To find the range,
As x varies from 0 to 1
$Y = 3X+1$ varies from 1 to 4.

$$\boxed{\therefore\ g(y) = \frac{2(y-1)}{9}\ ;\ 1 \le y \le 4}$$

2. Suppose X is a random variable which is uniformly distributed over (1, 3). Obtain pdf of the following (i) $Y = 3X+4$, (ii) $Z = e^{x}$

Solution:

X is uniformly distributed

$$\therefore f(x) = \begin{cases} \frac{1}{3-1}, & 1 \le x \le 3 \\ 0, & \text{else.} \end{cases}$$

i) $Y = 3X+4$,

cdf $G(y) = P(Y \le y)$

$= P(3X+4 \le y)$

$= P\left(X \le \frac{y-4}{3}\right) = \int_1^{\frac{y-4}{3}} f(x)\,dx$

$$G(y) = \frac{1}{2}\left(\frac{y-4}{3} - 1\right) = \frac{1}{2}\left(\frac{y-7}{3}\right)$$

$$\therefore \text{pdf } g(y) = \frac{1}{6}\frac{d}{dy}(y-7) = \frac{1}{6}$$

Range $x \to 1$ to 3

$Y = 3x+4 \to 7$ to 13

$\therefore$ Required pdf $\boxed{g(y) = \frac{1}{6},\ 7 \le y \le 13}$

Date
Page

ii) $Z = e^x$

cdf $H(z) = P(Z \le z)$

$= P(e^x \le z)$

$= P(X \le \log_e z)$

$$H(z) = \int_1^{\log_e z} \frac{1}{2}\, dx = \frac{1}{2}(\log_e z - 1)$$

$$h(z) = \frac{d}{dz} \frac{1}{2}(\log_e z - 1)$$

$$= \frac{1}{2}\left\{\frac{1}{z}\right\} = \frac{1}{2z}.$$

Range : X : 1 to 3

$z = e^x$: e to e^3

$$\boxed{\therefore h(z) = \frac{1}{2z}, \quad e \le z \le e^3}$$

3. Find the pdf of $Y = 8X^3$ if X has the pd of X i.e, $f(x) = \begin{cases} 2x, & 0 \le x \le 1 \\ 0, & \text{else} \end{cases}$

HW

Solution:

$G(y) = P\{Y \le y\}$

$= P\{8X^3 \le y\}$

$= P\{X^3 \le \frac{y}{8}\}$

$= P\{X \le \frac{y^{1/3}}{2}\}$

$$= \int_0^{y^{1/3}/2} f(x)\,dx = \int_0^{y^{1/3}/2} 2x\,dx$$

CS CamScanner

$$= x^2 \Big|_0^{y^{1/3}/2} = \frac{(y^{1/3})^2}{4}$$

$$\text{pdf} \Rightarrow g(y) = \frac{d}{dy}\left(\frac{y^{2/3}}{4}\right) = \frac{1}{4}\cdot\frac{2}{3}y^{-1/3}$$

$$\boxed{g(y) = \frac{y^{-1/3}}{6}, \quad 0 \le y \le 8}$$

$x \to 0$ to 1

$y = x^3 \to 0$ to 8

4. Find the pdf of $Y = -2\log_e X$, where,

$$f(x) = \begin{cases} 1, & 0 \le x \le 1 \\ 0, & \text{elsewhere.} \end{cases}$$

MONDAY

Solution:

$Y = -2\log X$ is monotonic, $(0,1)$

$$g(y) = f(x)\left|\frac{dx}{dy}\right|$$

$-y/2 = \log x$

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

$$= 1\left|\frac{dx}{dy}\right|$$

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

$$\boxed{g(y) = 1\cdot e^{-y/2}\left(\frac{1}{2}\right)} \quad 0 \le y \le \infty.$$

$0 \le e^{-y/2} \le 1$

$\ln(0) \le -y/2 \le \ln(1)$

$2\ln(0) \le -y \le$

II Method: cdf $G(y) = P(-2\log X \le y)$

$= P(\log X \ge -y/2)$

$= P(X \ge e^{-y/2})$

$\ln(0) = -\infty$

$\because e^{-\infty} = 0$

$-\infty = \ln(0)$

undefined

$$= \int_{e^{-y/2}}^{1} dx = -e^{-y/2} + 1$$

$$g(y) = \frac{d}{dy}G(y) = e^{-y/2}\left(+\frac{1}{2}\right) = \boxed{\frac{1}{2}e^{-y/2}}$$

$0 \le y \le \infty$

5. Suppose X is a random variable, has pdf,

$$f(x) = \begin{cases} 2x, & 0 \le x \le 1 \\ 0, & \text{else,} \end{cases}$$ then find pdf of $Y = e^{-X}$.

MONDAY

Solution: $Y = e^{-x}$ is monotonic,

$$G(y) = P(e^{-X} \le y)$$
$$= P(-X \le \log y) \qquad x = -\log y$$
$$= P(X \ge -\log y)$$
$$\int_{-\log y}^{1} 2x\,dx \qquad 0 \le x \le 1$$
$$0 \le -\log y \le 1$$
$$= 1 - (\log y)^2 \qquad -1 \le \log y \le 0$$

$$\boxed{g(y) = -2\log y \cdot \frac{1}{y} \qquad 1 \le y \le e^{-1}}$$

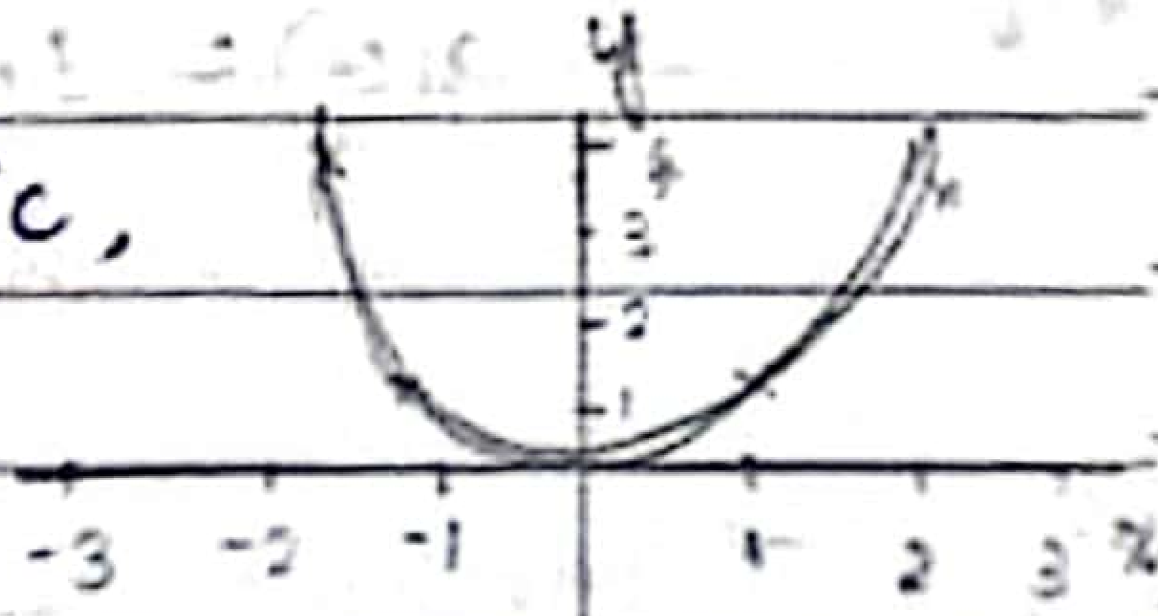
$$\frac{1}{e} \le y \le 1$$

6. If 'X' is CRV with $f(x) = \begin{cases} 1/2, & -1 \le X \le 1 \\ 0, & \text{else,} \end{cases}$ find ... of $Y=$

MONDAY

Solution:

$y = x^2$ is not monotonic,



$$G(y) = P(Y \le y)$$
$$= P(X^2 \le y)$$
$$= P(\pm X \le \sqrt{y}) \qquad P\{X \le \sqrt{y}\} \quad P\{-X \le \sqrt{y}\}$$
$$P\{-\sqrt{y} \le X\}$$
$$= P\{-\sqrt{y} \le X \le \sqrt{y}\}$$
$$= \int_{-\sqrt{y}}^{\sqrt{y}} \frac{1}{2}\,dx$$

$G(y) = \sqrt{y}$. $\quad g(y) = \dfrac{1}{2\sqrt{y}}$

$X \to Y$
P

$G(y) = F(\sqrt{y}) - F(-\sqrt{y})$ using cdf,

$$g(y) = \frac{d}{dy} G(y) = \frac{d}{dy}\left(F(\sqrt{y}) - F(-\sqrt{y})\right)$$

$$= f(\sqrt{y}) \frac{1}{2\sqrt{y}} - f(-\sqrt{y}) \cdot \frac{1}{-2\sqrt{y}}$$

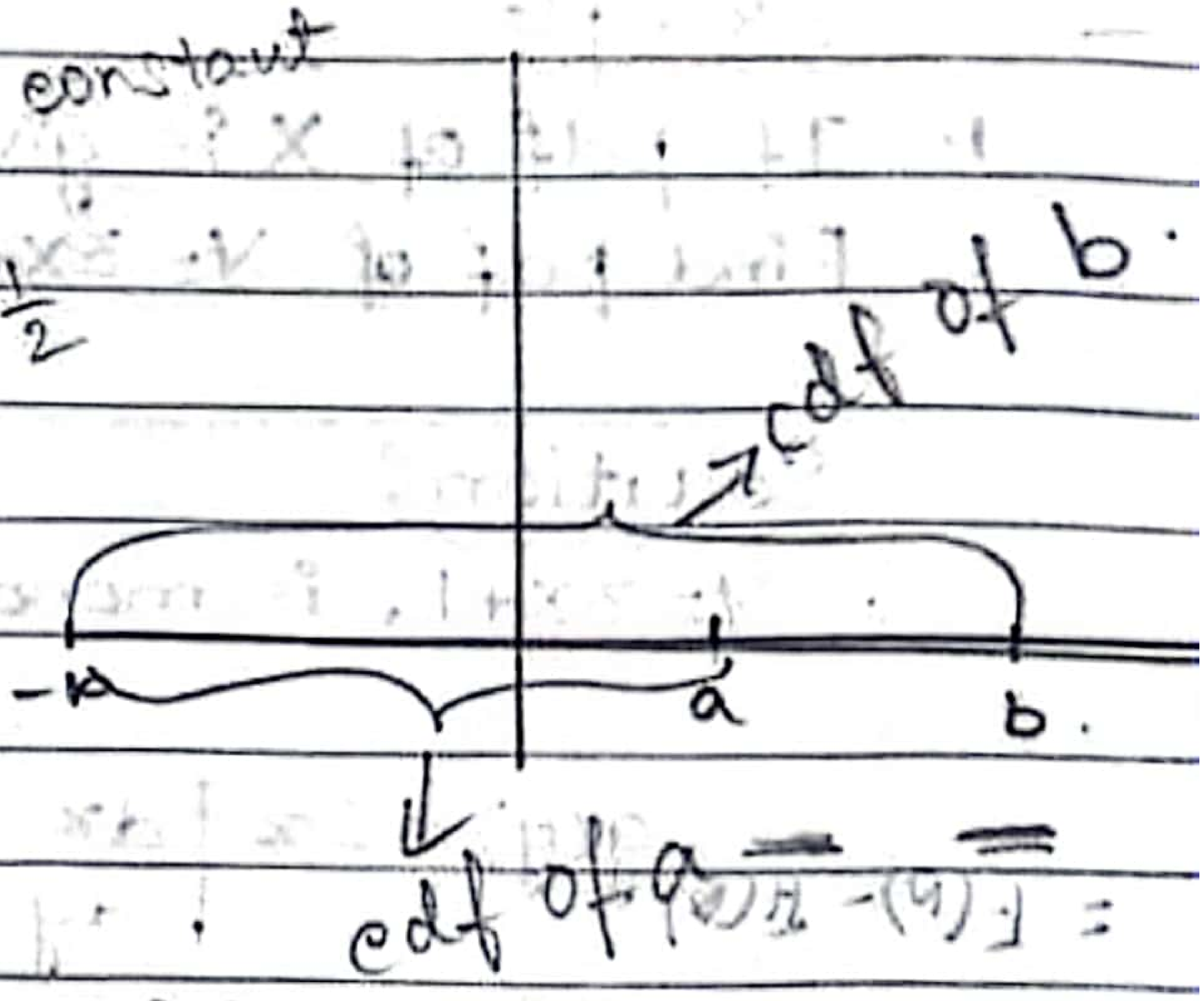
$$= \frac{1}{2\sqrt{y}}\left(f\sqrt{y} + f(-\sqrt{y})\right)$$

$$\therefore g(y) = \frac{1}{2\sqrt{y}}\left(\frac{1}{2} + \frac{1}{2}\right) \rightarrow \text{doubt}$$

$$\boxed{\therefore g(y) = \frac{1}{2\sqrt{y}}} \qquad 0 \le y \le 1.$$

$f(x) = \frac{1}{2} = \text{constant}$

$x^2 = y$
$x = \sqrt{y}$ $\quad f(\sqrt{y}) = \frac{1}{2}$



$P\{a \le X \le b\} = F(b) - F(a)$

$G(y) = F(b) - F(a)$

$F(x) = \begin{cases} & x < 0 \\ & 0 \le x \end{cases}$

classmate

Date ____
Page ____

Theorem:

Define

Let X be a continuous random variable with pdf $f(x)$, where $f(x)>0$ for $a<x<b$. Suppose $y=H(x)$, is strictly monotonic (As x increases, the function either increases or decreases) function of x, (continuous) assume that the function is differentiable for all x. Then the random variable $Y=H(X)$ has pdf given by.

$$g(y)=f(x)\left|\frac{dx}{dy}\right|$$

Example.

1. If pdf of X is given by, $f(x)=\begin{cases} 2x, & 0\le x\le 1 \\ 0, & \text{else} \end{cases}$

Find pdf of $Y=3X+1$.

Solution:

$Y=3X+1$, is monotonic function.

$$\therefore g(y)=2x\left|\frac{dx}{dy}\right| = 2x\cdot\frac{1}{3}$$

$y=3x+1$

$dy=3dx \Rightarrow \frac{dx}{dy}=\frac{1}{3}$

$$\boxed{\therefore g(y)=\frac{2}{3}\left\{\frac{y-1}{3}\right\}} \qquad 1\le y\le 4.$$

Date ______
Page ______

Theorem:

Define

Let X be a continuous random variable with pdf $f(x)$, where $f(x)>0$ for $a<x<b$. Suppose $y=H(x)$, is strictly monotonic (As x increases, the function either increases or decreases) function of x, (continuous) assume that the function is differentiable for all x. Then the random variable $Y=H(X)$ has pdf given by.

$$g(y)=f(x)\left|\frac{dx}{dy}\right|$$

Example.

1\. If pdf of X is given by, $f(x)=\begin{cases}2x, & 0\le x\le 1\\ 0, & \text{else}\end{cases}$

Find pdf of $Y=3X+1$.

Solution:

$Y=3X+1$, is monotonic function.

$$\therefore g(y)=2x\left|\frac{dx}{dy}\right| = 2x\cdot\frac{1}{3}$$

$y=3x+1$

$dy=3\,dx \Rightarrow \dfrac{dx}{dy}=\dfrac{1}{3}$

$$\boxed{\therefore g(y)=\frac{2}{3}\left\{\frac{y-1}{3}\right\}} \quad 1\le y\le 4.$$

2. The random variable X has uniform distribution in $(-\pi/2, \pi/2)$ find the pdf of $Y = \tan X$.

Solution:

pdf of X is $f(x) = \begin{cases} 1/\pi, & -\pi/2 \leq x \leq \pi/2 \\ 0, & \text{else.} \end{cases}$

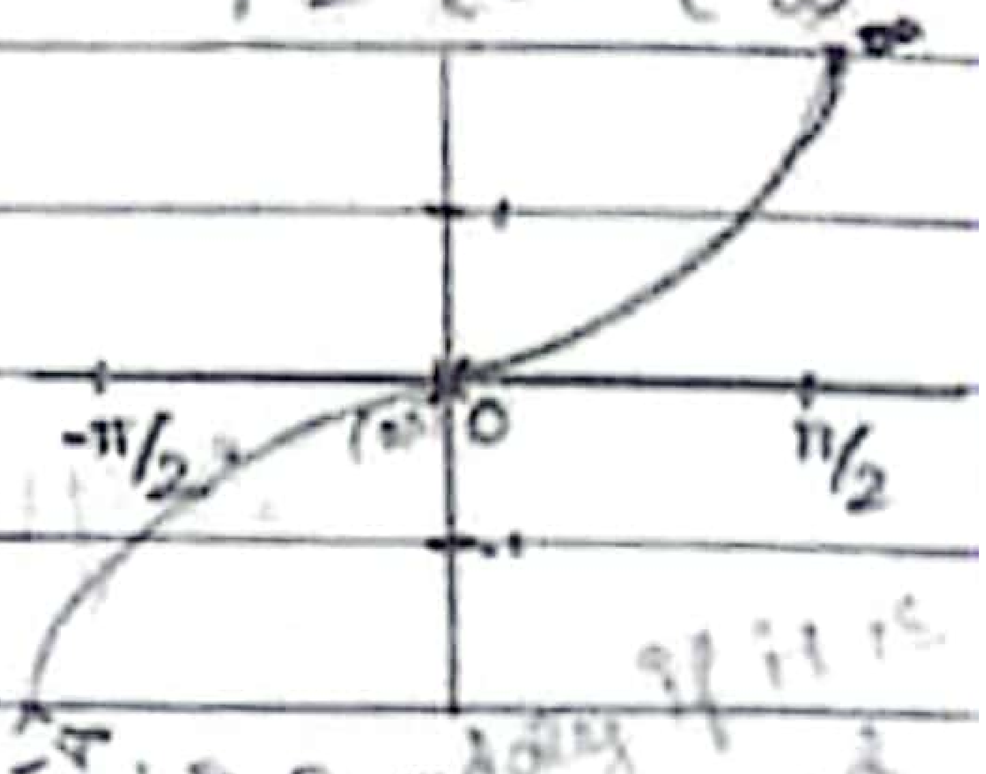
$Y = \tan X$ is monotonic function.

$\therefore g(y) = f(x)\left|\frac{dx}{dy}\right|$

$= \frac{1}{\pi}\left|\frac{dx}{dy}\right|$

$Y = \tan x$

$Y = \tan^{-1}(\pi/2)$



We have to find. $X = \tan^{-1} Y$

$\frac{dx}{dy} = \frac{1}{1+y^2}$

Boundary if it is not defined

$$g(y) = \frac{1}{\pi}\boxed{\frac{1}{1+y^2}}, \quad -\infty < y < \infty$$

This is called Cauchy's distribution.

[Its one-one & onto → monotonic]

---

3. If X has Cauchy's distribution then S.T pdf of $Y = 1/X$ also has Cauchy's distribution.

MONDAY

Solution:

$Y = \frac{1}{X}$ is not defined at $X = 0$, $Y = \infty$

Given X has Cauchy's distribution,

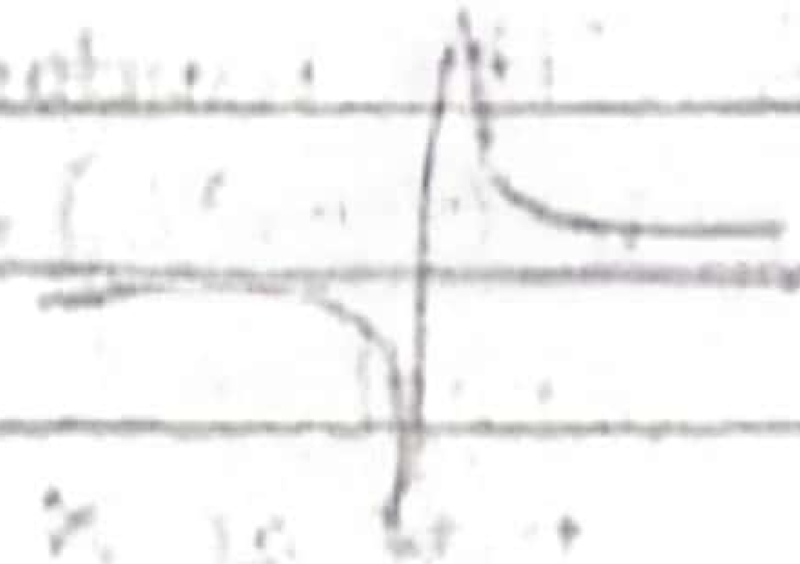
$$f(x) = \frac{1}{\pi}\left(\frac{1}{1+x^2}\right) \qquad -\infty \leq x \leq \infty$$

$$G(y) = P\{Y \leq y\} = P\left\{\frac{1}{X} \leq y\right\}$$

$$= P\{X \geq 1/y\}$$

$$= \int_{1/y}^{\infty} f(x)\,dx$$

$$= \frac{1}{\pi}\int_{1/y}^{\infty} \frac{1}{1+x^2}\,dx$$



$$= \frac{1}{\pi}\tan^{-1}(x)\Big|_{1/y}^{\infty}$$

$$= \frac{1}{\pi}\left\{\tan^{-1}(\infty)-\tan^{-1}\left(\frac{1}{y}\right)\right\}$$

$$= \frac{1}{\pi}\left\{\frac{\pi}{2}-\tan^{-1}\left(\frac{1}{y}\right)\right\}$$

$$\therefore g(y) = \frac{-1}{\pi}\frac{d}{dy}\tan^{-1}\left(\frac{1}{y}\right)$$

$$= \frac{-1}{-\pi}\left\{\frac{1}{1+1/y^2}\right\}$$

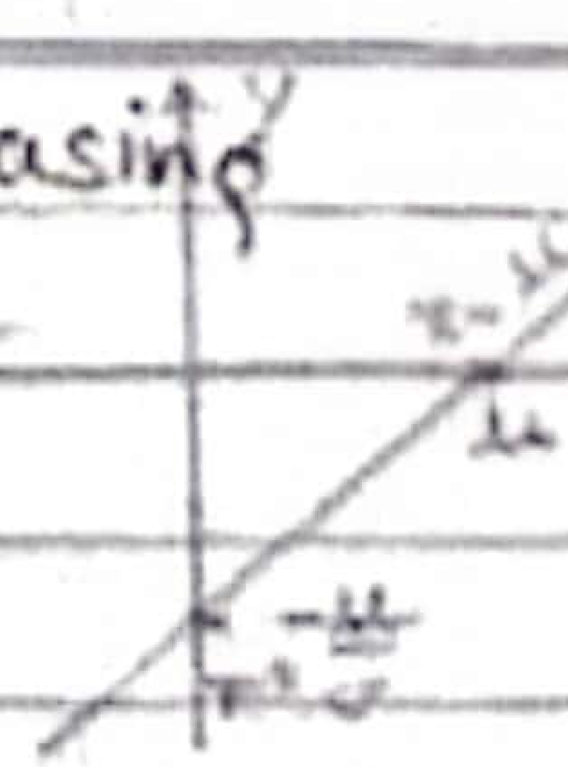
$\tan^{-1}\left(\frac{1}{y}\right) d\left(\frac{1}{y}\right)$

$$= \frac{-1}{\pi}\left\{\frac{y^2}{1+y^2}\right\}\cdot\left(\frac{-1}{y^2}\right)$$

$$\boxed{g(y) = \frac{1}{\pi}\left\{\frac{1}{1+y^2}\right\}}, \quad -\infty<y<\infty$$

which is Cauchy's distibution

$Y = \frac{X-\mu}{\sigma}$ is monotonically increasing



$$g(y) = \frac{1}{\sqrt{2\pi}\sigma} e^{-\frac{1}{2}\left(\frac{x-\mu}{\sigma}\right)^2} \left|\frac{dx}{dy}\right|$$

$$= \frac{1}{\sqrt{2\pi}} e^{-y^2/2} \quad -\infty\le y\le\infty$$

$x = \sigma y+\mu$

$dx = \sigma\,dy$

Theorem:

If $X \sim N(\mu, \sigma^2)$ then $Y = \frac{X-\mu}{\sigma}$ has $N(0,1)$.

Proof:

To prove the result we need to show,

$$g(y) = \frac{1}{\sqrt{2\pi}} e^{-\frac{1}{2}y^2} \quad -\infty \le y \le \infty.$$

$$\rightarrow \quad G(y) = P\{Y \le y\} = P\left\{\frac{X-\mu}{\sigma} \le y\right\}$$

$$= P\{X \le y\sigma + \mu\}.$$

$$= \int_{-\infty}^{y\sigma+\mu} f(x)\,dx = \int_{-\infty}^{\mu+y\sigma} \frac{1}{\sqrt{2\pi}\,\sigma} e^{-\frac{1}{2}\left(\frac{x-\mu}{\sigma}\right)^2} \cdot dx$$

Substituting $\frac{x-\mu}{\sigma} = t$, $dx = \sigma dt$.

$$= \int_{-\infty}^{y}$$

$x = \mu + t\sigma$, $x \to -\infty$, $t \to -\infty$

$x \to \mu + y\sigma$ $t \to y$.

$$= \int_{-\infty}^{y} \frac{1}{\sqrt{2\pi}\,\sigma} e^{-\frac{1}{2}(t)^2} \sigma\, dt.$$

$$G(y) = \frac{1}{\sqrt{2\pi}} \int_{-\infty}^{y} e^{-t^2/2} \cdot dt \quad \leftarrow$$

Again substitute

$t = y$

$dt = dy$

$t \to -\infty$

$y \to -\infty$

$t \to y$

$y \to y$

$$g(y) = \frac{d}{dy} G(y) = \frac{1}{\sqrt{2\pi}} \frac{d}{dy} \int_{-\infty}^{y} e^{-y^2/2}\,dy$$

$$= \frac{1}{\sqrt{2\pi}} e^{-y^2/2} \Big|_{-\infty}^{y} = \frac{1}{\sqrt{2\pi}} \left(e^{-y^2/2} - e^{-(-\infty)^2} \right)$$

$$\boxed{g(y) = \frac{1}{\sqrt{2\pi}} e^{-y^2/2}} \quad -\infty \le y \le \infty$$

classmate

Date
Page

Theorem:

If X has $N(\mu, \sigma^2)$ & if $Z = \left(\frac{X-\mu}{\sigma}\right)^2$ then Z has χ^2 distribution with one degree of freedom.

Proof:

$$\chi^2(n) = \frac{e^{-x/2} x^{n/2-1}}{2^{n/2}\Gamma(n/2)}, \quad x>0$$

$n \rightarrow$ degree of freedom.

$$\therefore \chi^2(1) = \frac{e^{-x/2} x^{-1/2}}{2^{1/2}\Gamma(1/2)} = \frac{e^{-x/2}}{\sqrt{2\pi x}}$$

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

$$Y = \frac{x-\mu}{\sigma} \sim N(0,1)$$

$$\therefore Z = Y^2$$

To find pdf

$$H(z) = P(Z \leq z) = P(Y^2 \leq z)$$
$$= P(-\sqrt{z} \leq y \leq \sqrt{z})$$
$$= \int_{-\sqrt{z}}^{\sqrt{z}} f(y)\,dy$$

$(\because Y \sim N(0,1))$

$$= \int_{-\sqrt{z}}^{\sqrt{z}} \frac{e^{-y^2/2}}{\sqrt{2\pi}} \cdot dy$$

$$= \frac{2}{\sqrt{2\pi}} \int_0^{\sqrt{z}} e^{-y^2/2} \cdot dy$$

$y^2 = t$

$2y\,dy = dt$

$\cdot dy = \frac{dt}{2\sqrt{t}}$

$$= \frac{1}{\sqrt{2\pi}} \int_0^{z} e^{-t/2} t^{-1/2} dt$$

$\therefore g(z) = \frac{d}{dz} G(z)$ Substitute $t = z$,

$$= \frac{d}{dz} \int_0^z \frac{e^{-z/2} z^{-1/2}}{\sqrt{2\pi}} dz$$

$$= \frac{e^{-z/2} \cdot z^{-1/2}}{\sqrt{2\pi}} \Big|_0^z = \frac{1}{\sqrt{2\pi}} e^{-z/2} \cdot z^{-1/2}$$

$$\boxed{\therefore g(y) = \frac{e^{-z/2}}{\sqrt{2\pi z}} = \chi^2(1)}$$

$\therefore z = \left(\frac{x-\mu}{\sigma}\right)^2$ follows $\chi^2(1)$.

Note:-

Let X be a continuous random variable, with pdf f. Let $Y = X^2$, then random variable Y pdf is given by,

give

$$g(y) = \frac{1}{2\sqrt{y}} \left\{ f(\sqrt{y}) + f(-\sqrt{y}) \right\}.$$

* Find the pdf $y = -\log_e x$ pdf of $x \to f(x) = 2x$ $0 < x$

Monotonic $-y = \log_e x$

$y = \log x^{-1}$

$e^{y} = \frac{1}{x} \Rightarrow x = \frac{1}{e^{y}} = e^{-y}$

$\frac{dx}{dy} = e^{-y}(-1) = -e^{-y}$

pdf of $g(y) = 2x e^{-y} = -2 \frac{1}{e^{y}} (e^{-y}) = -2e^{-2y}$

By the theorem,

$$g(y) = \frac{1}{2\sqrt{y}} \left\{ f(\sqrt{y}) + f(-\sqrt{y}) \right\}$$

$Z = Y^2$,

$$f(y) = \frac{1}{\sqrt{2\pi}} e^{-y^2/2} \quad . \quad -\infty < y < \infty$$

$$g(z) = \frac{1}{2\sqrt{z}} \left\{ f(\sqrt{z}) + f(-\sqrt{z}) \right\}$$

$$= \frac{1}{2\sqrt{z}} \left\{ \frac{e^{-z/2}}{\sqrt{2\pi}} + \frac{e^{-z/2}}{\sqrt{2\pi}} \right\}$$

$$= \frac{1}{\sqrt{z 2\pi}} \left(e^{-z/2} \right)$$

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

Date
Page

2-Dimensional.

Let (X,Y) be a 2-dimensional continuous random variable & if $Z = H_1(X,Y)$ be a continuo-us function of (X,Y), then Z will be a continuous random variable (1-dimensional). To find pdf of Z, we will introduce a new random variable, $W = H_2(X,Y)$ & obtain the jpdf of Z, W. Then we can find the desired pdf of Z, say, $g(z)$ by integrating w.r.to w. (mpdf)

Theorem :

Suppose that (X,Y) is a 2D random varia-ble with joint pdf f, defined over the region R. Let $Z = H_1(X,Y)$ & $W = H_2(X,Y)$. If H_1 & H_2 satisfy the follows conditions,

i) Equations $Z = H_1(X,Y)$ & $W = H_2(X,Y)$ may be uniquely solved for X & Y interms of Z & W

ii) The partial derivatives $\frac{\partial x}{\partial z}, \frac{\partial x}{\partial w}, \frac{\partial y}{\partial z}$ & $\frac{\partial y}{\partial w}$ exists & are (continuous)

Then the joint pdf is given by,

$$\begin{cases} f(x,y)\,|J|_{(z,w)} & , \text{ if } (z,w) \in R \\ 0 & , \text{ else} \end{cases}$$

where J is given by, $J = \begin{vmatrix} \partial x/\partial z & \partial x/\partial w \\ \partial y/\partial z & \partial y/\partial w \end{vmatrix}$

Jacobian of transformation $(x,y) \to (z,w)$

Problems

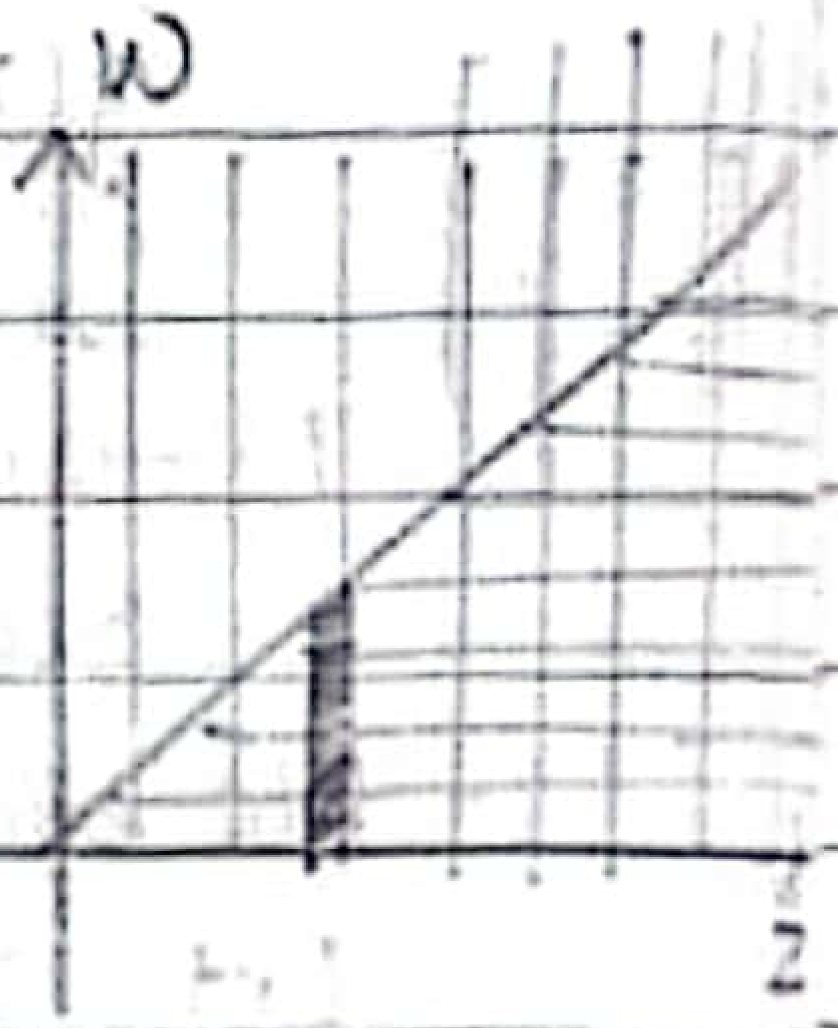
1. 2 independent random variable 'X' and 'Y' having pdf $f(x) = e^{-x}$, $g(y) = 2e^{-2y}$, $0 \le x, y \le \infty$. Find pdf of $X+Y$.

Solution:

Given X & Y are independent

$$\therefore f(x,y) = f(x) \cdot g(y)$$

$$= 2e^{-x-2y}$$



Let $w = y$

$Z = X + Y \Rightarrow X = Z - Y$

$= z - w$

Solve $w = y$, $x = z - w$

$$|J| = \begin{vmatrix} \partial x/\partial z & \partial x/\partial w \\ \partial y/\partial z & \partial y/\partial w \end{vmatrix} = \begin{vmatrix} 1 & -1 \\ 0 & 1 \end{vmatrix} = 1$$

Joint pdf of z & w is given by, say,

$$k(z,w) = f(x,y)\,|J|$$

$0 \le x \le \infty$

$0 \le z - w \le \infty$

$w \le z \le \infty \rightarrow ①$

$$= 2e^{-x-2y} \cdot 1$$

$0 \le y \le \infty$

$0 \le w \le \infty$

$$= e^{-x-y-y} = e^{-(x+y)+y} = e^{-[z+w]}$$

$\therefore$ To find pdf of z, integrate w.r.to w,

$-(x+y+y)$, $z+w$

$$h(z) = \int_{w=0}^{z} 2e^{-x-2y}\,dw$$

$$= \int_0^z 2e^{-(z+w)}\,dw = -2e^{-z}(e^{-z}+1) = -2e^{-2z} - 2e^{-z}$$

2. If X_1 & X_2 are independent and have standard normal distribution i.e, $X_1, X_2 \sim N(0,1)$ Find pdf of $Z = X_1/X_2$

Solution :-

Given $X_1 \sim N(0,1)$

$$\therefore f(x_1) = \frac{1}{\sqrt{2\pi}} e^{-x_1^2/2}, \quad -\infty \le x_1 \le \infty$$

lll'y

$$f(x_2) = \frac{1}{\sqrt{2\pi}} e^{-x_2^2/2}, \quad -\infty \le x_2 \le \infty$$

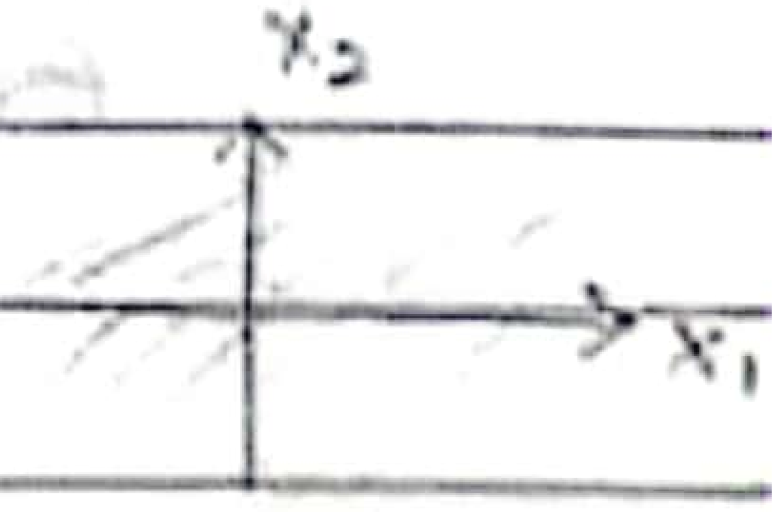
The jpdf is, $f(x_1, x_2) = \frac{1}{2\pi} e^{-\frac{(x_1^2+x_2^2)}{2}}$ $\quad -\infty \le x_1, x_2 \le \infty$

Let $W = X_2 \qquad Z = \frac{X_1}{X_2}$

$$X_1 = Z \cdot W$$

$$\therefore J = \begin{vmatrix} \partial x_1/\partial z & \partial x_1/\partial w \\ \partial x_2/\partial z & \partial x_2/\partial w \end{vmatrix} = \begin{vmatrix} w & z \\ 0 & 1 \end{vmatrix} = \underline{\underline{w}}$$

$\therefore$ Jpdf of z & w is given by,



$$k(z, w) = f(x_1, x_2)\,|J|$$

$-\infty \le wz \le \infty$

$$= \frac{1}{2\pi} e^{-\frac{(x_1^2+x_2^2)}{2}} \cdot |w|$$

$-\infty \le w \le \infty$

$-\infty \le z \le \infty$

$x_2^2\left(\frac{x_1^2}{x_2^2}+1\right)$

$$\boxed{= \frac{|w|}{2\pi} e^{-w^2/2\,(1+z^2)}}$$

$-\omega \le \infty$

$\infty \le w \le z$

$\infty \le w \le z \le \infty$

0 $\quad$ 7

$$\therefore f(z) = \int_{-\infty}^{\infty} \frac{|w|}{2\pi} e^{-w^2/2\,(1+z^2)}\, dw.$$

$0 \le z \le \infty$

By substituting $\frac{w^2}{2}(1+z^2) = t$

$$w(1+z^2)\,dw = dt$$

$$\therefore = \frac{1}{\pi}\int_0^{\infty} e^{-t} \cdot \frac{dt}{1+z^2}$$

$$\boxed{f(z) = \frac{1}{\pi(1+z^2)}}$$ This distribution called Cauchy's distri

Thus if $X_1, X_2 \sim N(0,1)$ then $\frac{X_1}{X_2} \sim$ Cauchy distribut

---

③ If $X \sim N(0,\sigma^2)$, $Y \sim N(0,\sigma^2)$ and X & Y are independent. Find pdf of $R = \sqrt{x^2+y^2}$

Solution:

$0 \le x, y \le \infty$

$\theta \to 0$ to 2π

Given $X \sim N(0,\sigma^2)$

$$\therefore f(x) = \frac{1}{\sqrt{2\pi}\sigma} e^{-\left(\frac{x}{\sigma}\right)^2}, \quad -\infty < x < \infty.$$

and

$Y \sim N(0,\sigma^2)$

$$f(y) = \frac{1}{\sqrt{2\pi}\sigma} e^{-\left(\frac{y}{\sigma}\right)^2}, \quad -\infty < y < \infty$$

$\because$ They are independent,

$$f(x,y) = f(x)\cdot f(y)$$

$$= \frac{1}{2\pi\sigma^2} e^{-\left(\frac{x^2+y^2}{\sigma^2}\right)} \quad -\infty \le x, y \le \infty.$$

Now, instead of z & w we introduce polar co-ordinates, $X = R\cos\theta$, $Y = R\sin\theta$, $\theta = \tan^{-1}\left(\frac{X}{Y}\right)$ & $R = \sqrt{X^2+Y^2}$

$$J = \begin{vmatrix} \partial X/\partial r & \partial X/\partial\theta \\ \partial Y/\partial r & \partial Y/\partial\theta \end{vmatrix} = \begin{vmatrix} \cos\theta & -R\sin\theta \\ \sin\theta & R\cos\theta \end{vmatrix}$$

$$= R(\cos^2\theta + \sin^2\theta) = \underline{R}$$

$$\therefore k(R,\theta) = \frac{R}{2\pi\sigma^2} e^{-\left(\frac{x^2+y^2}{\sigma^2}\right)},$$

$$= \frac{R}{2\pi\sigma^2} e^{-\left(\frac{R^2}{\sigma^2}\right)}, \quad R \geq 0, \; 0 \le \theta \le 2\pi$$

Now to find

$$h(R) = \int_0^{2\pi} \frac{R}{2\pi\sigma^2} e^{-(R^2/\sigma^2)} \cdot d\theta$$

$$= \frac{R\, e^{-(R^2/\sigma^2)}}{2\pi\sigma^2} (2\pi - 0)$$

$$\boxed{h(R) = \frac{R}{\sigma^2} e^{-(R^2/\sigma^2)}, \quad R > 0}$$

This is called Rayleigh density distribution

SAMPLING

* A finite subset of statistical individuals in a population (universe) is called as a sample.
* The number of individuals in the sample is 'sample size'
* If the sample units are selected at random from the population, then its called 'random sampling'.

Random Sample Space :

Let X be a random variable having pdf f(x
Let $X_1, X_2 \dots X_n$ be 'n' independent random variables having the same pdf (distribution) as that of X. Then we say that $(X_1, X_2 \dots X_n)$ is a random sample of size 'n'.

Statistics :

Is a function of one or more random variables which doesn't depend on any unknown parameter.

eg: $Y = \frac{X - \mu}{\sigma}$ is a statistic if μ, σ are know

Sample Mean : $\overline{X}$ Is a function

The statistic $\overline{X} = \frac{X_1 + X_2 + \dots + X_n}{n}$ is called sample mean & denoted by $\overline{X}$.
[Its again a function]

Sample Variance : (S^2) :

The statistic $\boxed{S^2 = \frac{\sum (x_i - \overline{X})^2}{n}} = \frac{\sum (x_i - \overline{X})^2}{n-1}$

SAMPLING DISTRIBUTION.

Distribution of a statistic of a random sample is called sampling distribution.

THEOREM:

Let X be the random variable with expectation $E(X) = \mu$ and variance $V(X) = \sigma^2$. Let $\overline{X}$ be the sample mean of a random sample of size 'n'. Then,

$E(\overline{X}) = \mu$, $V(\overline{X}) = \sigma^2/n \Rightarrow V(\overline{X}) = \frac{1}{n^2}\sum_{i=1}^{n} V(X_i) = \frac{\sigma^2}{n}$

$\sum_{i=1}^{n} E(X_i) = n\mu$, and $\overline{X} = \frac{X_1 + X_2 + \dots + X_n}{n} = \frac{n\mu}{n} = \mu$ (last page)

CENTRAL LIMIT THEOREM:

Let $X_1, X_2 \dots X_n$ denote a random sample (any) size 'n' from a distribution with mean 'μ' and variance σ^2. Let,

Sum $S = X_1 + X_2 + \dots + X_n = \sum_{i=1}^{n} X_i$

Then $Y = \frac{S - n\mu}{\sqrt{n}\,\sigma}$ has $N(0,1)$ as $n \to \infty$.

Alternative form:

$S = X_1 + X_2 + \dots + X_n$

$\frac{S}{n} = \frac{\sum_{i=1}^{n} X_i}{n} = \overline{X}$,

Then, $Y = \frac{S - n\mu}{\sqrt{n}\,\sigma} \Rightarrow \frac{n\left[S/n - \mu\right]}{\sqrt{n}\,\sigma}$

$= \frac{\sqrt{n}\,(\overline{X} - \mu)}{\sigma} \sim N(0,1)$

$\because \left[Y = \frac{S - n\mu}{\sqrt{n}\,\sigma} \sim N(0,1)\right]$

Note :

* If X has $N(\mu, \sigma^2)$ then $\overline{X} \sim N(\mu, \sigma^2/N)$
* [The sampling distribution of large samples is assumed to be normal distribution].
* Standard deviation $\Leftrightarrow$ standard error.

* The totality of units under consideration is called population [finite or infinite]
* Representative portion of the population selected for a study is called sample.
* Every unit in the population has a chance of selection is called random
* If the chance of selection is equal is called simple random sample.

* Population mean $= \mu$, variance $= \sigma^2$

* Sample mean $\overline{X}$ $= \mu$ & variance $= \sigma^2/N$

Result :

If s^2 is a sample variance then $ns^2/\sigma^2 \sim \chi^2_{(n-1)}$

Also $\overline{X}$ & s^2 are always independent.

$X_1, X_2 \ldots X_n$ is population

PROBLEMS:

1. Let $\overline{X}$ be mean of random sample of size 5 taken from normal distribution of $\mu=0, \sigma^2$=
Determine c so that $P\{\overline{X}<c\}=0.9$.

Solution:

1. $N=5,\ \mu=0,\ \sigma^2=125$

$\overline{X} \sim N(\mu, \sigma^2/N) = N(0, 125/5)$

$= N(0, 25)$

$\therefore P\{\overline{X}<c\} = 0.9$

$\Rightarrow$ Standardizing

$P\left\{Z< \frac{c}{5}\right\} = 0.9$

$\Rightarrow \Phi(c/5) = 0.9 \Rightarrow c/5 = \Phi^{-1}(0.9)$

$\Phi^{-1}(0.9) \to 1.29$

$\Rightarrow c = 5 \times 1.29$

$\boxed{c = 6.45}$

---

2. Let $\overline{X}$ be mean of random sample of size 25 from the distribution $X \sim N(75, 100)$. Find $P_r\{71<\overline{X}<79\}$.

H.W

Solution:

$\overline{X} \sim N(75, 4)$

$P\{71<\overline{X}<79\} = P_r\left\{\frac{71-\mu}{\sigma} < Z < \frac{79-\mu}{\sigma}\right\}$

$= \left\{\frac{-4}{2} < Z < \frac{4}{2}\right\}$

$= \Phi(2) - (1-\Phi(2)) = 2\Phi(2) - 1$

$= 2(0.9772) - 1$

$= \boxed{0.954}$

Date
Page

3

3. If $\overline{x}$ is mean of random sample of size N
// taken from $N(\mu, 100)$. Find N so that,
$P\{\mu-5 < \overline{x} < \mu+5\} = 0.954$.

Solution:

$$P\{\mu-5 < \overline{x} < \mu+5\} = 0.954$$

$$x \sim N(\mu, 100) \Rightarrow \overline{x} \sim N(\mu, 100/N)$$

2

$$P\left\{\frac{\mu-5-\mu}{10/\sqrt{N}} < Z < \frac{\mu+5-\mu}{10/\sqrt{N}}\right\} = 0.954$$

$$P\left\{\frac{-5\sqrt{N}}{10} < Z < \frac{5\sqrt{N}}{10}\right\} = 0.954$$

$$2\Phi\left(\frac{5\sqrt{N}}{10}\right) - 1 = 0.954$$

$$= \frac{1.954}{2} = 0.977.$$

$$\frac{\cancel{5}\sqrt{N}}{\cancel{10}\,2} = 2.00 \Rightarrow \sqrt{N} = 4 = \boxed{N=16}$$

4. A random sample if size 100 is taken from an
// infinite population with $\mu = 53$, $\sigma^2 = 400$, find
H.W $P\{50 < \overline{x} < 56\}$.

Solution:

$$\overline{x} \sim N(53, 400/100) \sim N(53, 4)$$

$$P\left\{\frac{50-53}{2} < Z < \frac{56-53}{2}\right\} \Rightarrow P\{-1.5 < z < 1.5\}$$

$$\Rightarrow 2\Phi(1.5) - 1 \Rightarrow 2 \times 0.9332 - 1.$$

$$= \boxed{0.8664}$$

5. A random sample if size 64 is taken from an
H.W infinite population with $\mu = 112$, $\sigma^2 = 144$. Find

the probability of getting sample mean greater than 114.5.

Solution :

$n = 64, \ \mu = 122, \ \sigma^2 = 144$

$P\{\bar{x} > 114.5\} = 1 - P\{\bar{x} \leq 114.5\}$

$= 1 - P\left\{Z \leq \left(\frac{114.5 - 122}{12}\right) 8\right\}$

$\Rightarrow 1 - \phi(+0.208) = \cancel{2 + \phi(0.625)}$

$= 1 - 0.6103 = 0.3897$

$\Rightarrow 1 - P\{Z \leq 1.666\}$

$= 0.048$

---

6. Let s^2 be variance of random sample of size 6 from $N(\mu, 12)$ find $P\{2.3 < s^2 < 22.2\}$

dont do

Solution :

$P\{2.3 < s^2 < 22.2\}$

$\because \frac{s^2 n}{\sigma^2} \sim$

$\Rightarrow P\left\{\frac{2.3 \times 6}{12} < \chi^2(5) < \frac{22.2 \times 6}{12}\right\}$

$= P\{1.15 < \chi^2(5) < 11.1\}$

$\Rightarrow \phi(11.1) - \phi(1.15)$

$= 0.95 - 0.05 = 0.90$

---

7. Let $\bar{x}$ & s^2 be mean & variance of random sample of size 25 from a distribution $N(3, 100)$. Find $P\{0 < \bar{x} < 6, \ 55.2 < s^2 < 145.6\}$ where $\bar{x}$ & s^2 are independent.

dont do

Solution :-

$n = 25, \quad X \sim N(3, 100)$

Given $\overline{x}$ & s^2 are independent,

$P\{0 < \overline{x} < 6,\ 55.2 < s^2 < 145.6\}$

$= P\{0 < \overline{x} < 6\} \cdot P\{55.2 < s^2 < 145.6\}$

Standardizing,

$$P\left\{\frac{0-3}{10} < z < \frac{6-3}{10}\right\} \cdot P\left\{\frac{55.2 \times 25}{100} < \frac{ns^2}{\sigma^2} < \frac{145.6 \times 25}{100}\right\}$$

$\Rightarrow P\{-0.3 < z_1 < 0.3\} \cdot P\{13.8 < Z_2 < 36.4\}$

$= \boxed{0.779}$

8. Suppose that X_j where $j = 1, 2 \ldots 50$ are independent random variable each having Poisson distribution $\alpha = 0.03$ & $S = X_1 + X_2 + \cdots + X_{50}$. Evaluate $P\{S \geq 3\}$

do CLT

Solution :-

For Poisson distribution $E(x) = V(x) = \boxed{0.03}$

$\therefore E(S) = V(S) = n \cdot \alpha = 50 \times 0.03 = \boxed{1.5}$

Converting

$P\{S \geq 3\} = 1 - P\{S < 3\}$

$$= 1 - P\left\{\frac{S - n\mu}{\sqrt{n}\,\sigma} < \frac{3 - (50)(0.03)}{\cancel{5\times}\sqrt{50 \times 0.03}}\right\}$$

$= 1 - P\{Y < 1.5/\sqrt{1.5}\}$

$= 1 - 0.8888 = \boxed{0.1112}$

Date
Page

Problems. (Central Limit theorem):

1. Let $\overline{x}$ denote the mean of the random sample of size 100 from a distribution which has $\chi^2(50)$. Compute $P\{49 \le \overline{x} \le 51\}$.

Solution:

If $X \sim \chi^2(n)$ then $\mu = n$, $\sigma^2 = 2n$

do $\therefore \mu = 50,\ \sigma^2 = 100 \quad \because X \sim \chi^2(50),\ N = 100$

$\overline{x}$ follows some distribution, which we will convert to standard normal distribution by,

$Y = \dfrac{S - n\mu}{\sqrt{n}\,\sigma}$ OR $Y = \dfrac{\sqrt{n}\,(\overline{x} - \mu)}{\sigma}$

when we know population mean & variance → sample mean define

$$P\{49 \le \overline{x} \le 51\} = P\left\{\frac{49-\mu}{\sigma/\sqrt{n}} \le Y \le \frac{51-\mu}{\sigma/\sqrt{n}}\right\}$$

$$\Rightarrow P\left\{\frac{49-50}{1} \le x \le \frac{51-50}{1}\right\}$$

$$\Rightarrow 2\phi(1) - 1 \Rightarrow 0.8413 \times 2 - 1$$

$$\Rightarrow \boxed{0.6826}$$

2. Let $\overline{x}$ denotes the sample mean of a random sample of size 128 from a gamma distribution with $\alpha = 2$, $\beta = 4$, Evaluate $P\{7 < \overline{x} < 9\}$

Solution:

2, 0.25

We know that, $G(r, \alpha)$, mean $= r/\alpha$, $v(x) = \frac{r}{\alpha^2}$

|||ly

$G(\alpha, \beta)$ distribution, by replacing $r = \alpha$,
$\alpha = 1/\beta$.

$$G(\alpha, \beta) = \frac{e^{-x/\beta}\, x^{\alpha - 1}}{\Gamma(\alpha)\,\beta^{\alpha}}, \quad E(x) = \alpha\beta, \quad V(x) = \alpha\beta^2$$

CamScanner

$N = 128$, $\alpha = 2$, $\beta = 4$, $\mu = 8$
$V(x) = 32$.

$P\{7 < \bar{x} < 9\}$ when standardizing,

$$P\left\{\frac{7-\mu}{\sigma/\sqrt{N}} < z < \frac{9-\mu}{\sigma/\sqrt{N}}\right\}$$

$$\Rightarrow \left\{\frac{8\sqrt{2}\,(7-8)}{4\sqrt{2}} < z < \frac{8\sqrt{2}\,(9-8)}{4\sqrt{2}}\right\}$$

$$\Rightarrow P\{-2 < z < 2\} = 2\phi(2.0) - 1$$
$$= 2 \times 0.9772 - 1$$

$$= \boxed{0.9544}$$

---

8. Let $\bar{x}$ be a mean of a random sample of size 'n' from the distribution which has $N(\mu, 9)$. Find 'n' such that,
$P\{\bar{x}-1 < \mu < \bar{x}+1\} = 0.90$.

do

Solution:

$N = n$, $N(\mu, 9) \Rightarrow E(x) = \mu$, $V(x) = 9$.
$\bar{x} \sim N(\mu, \sigma^2/N) \sim N(\mu, 9/n)$

$$P\left\{\bar{x}-1 < \mu < \bar{x}+1\right\} \Rightarrow P\left\{-\bar{x}-1 < -\mu < 1-\bar{x}\right\}$$

$$= P\{-1 < \bar{x}-\mu < 1\} = 0.90$$
$\div\ \sigma/\sqrt{n}$

$$= P\left\{\frac{-1}{3/\sqrt{n}} < y < \frac{1}{3/\sqrt{n}}\right\} = 0.90$$

$$2\phi\left(\frac{\sqrt{n}}{3}\right) - 1 = 0.90$$

$$\Phi(\sqrt{n}/3) = 1.9/2 = 0.95$$

$$\sqrt{n} = 3 \times \Phi^{-1}(0.95) = 4.95 \simeq 25$$

$$\Rightarrow \boxed{n \simeq 25}$$

4\. Compute approximately the probability that the mean of the random sample space of size 15 from a distribution having pdf

$$f(x) = \begin{cases} 3x^2, & 0 < x < 1 \\ 0, & \text{else} \end{cases}$$ is between $3/5$ & $4/5$.

Solution:

$N = 15, \quad f(x) = 3x^2, \; 0 < x < 1$

$P\{3/5 < \overline{X} < 4/5\}$

To find the above, we need to find, μ,

For CRV,

$$\mu = \int_0^1 x \, 3x^2 \, dx = \underline{3/4}$$

$$E(X^2) = \int_0^1 3x^4 \, dx = \underline{3/5}$$

$$V(X) = 3/5 - 9/16 = \frac{48-45}{80} = \frac{3}{80} \Rightarrow \frac{3}{80 \times 15}$$

$$\overline{X} \sim N(\mu, \sigma^2/N) = (3/4, 1/400)$$

$$\therefore P\left\{\frac{3/5 - 3/4}{1/20} < Z < \frac{4/5 - 3/4}{1/20}\right\}$$

$$= P\{-3 < Z < 1\} = \Phi(1) - (1 - \Phi(3))$$

$$= \Phi(3) + \Phi(1) - 1$$

$$= 0.9987 + 0.8413 - 1$$

$$= \boxed{0.84}$$

classmate
Date
Page
6

do

5. A computer in adding numbers rounds off each number to the nearest integer. Suppose that all rounding errors are independent and uniformly distributed over the interval $(-0.5, 0.5)$

1) If 1500 numbers are added what is the probability that the magnitude of total error exceeds 15?

that

2) How many numbers are to be added together in order that the magnitude of the total error is less than 10 with probability 0.90?

Solution:

For uniform distribution, $\mu = \frac{a+b}{2} = 0$

$\sigma^2 = \frac{(b-a)^2}{12} = (1/\sqrt{12})^2 = 1/12$, $n = 1500$

$|S| = |x_1| + |x_2| + \cdots + |x_{1500}|$

(a) $P\{|S|>15\} = 1 - P\{|S|<15\} = 1 - P\{-15<S<15\}$

$= 0.1802$

$= 1 - P\left\{\frac{-15-0}{\sqrt{1/12}\,/\sqrt{1500}} < Y < \frac{15}{\sqrt{1/12}\,/\sqrt{1500}}\right\}$

$= 1 - P\{-1.3416 < Y < 1.3416\}$

(b) $P\{|S|<10\} = 0.90$

$n \simeq 441$

6. If X is a random variable having $B(72, \frac{1}{3})$ using CLT find $P\{22 < X < 28\}$

Solution:

$X \sim B(n, p) \Rightarrow n = 72,\ p = 1/3,\ q = 2/3$

$E(X) = np \quad V(X) = npq$

$P\{22 < X < 28\} = P\left\{\frac{22-24}{4} < Y < \frac{28-24}{4}\right\}$

$= P\{-0.5 < Y < 1\}$

$= \phi(1) - (1 - \phi(0.5))$

$= 0.5328$

A distribution with unknown mean μ has a variance 1.5. Find how large a sample should be taken from the distribution in order that the probability will be atleast 0.95, that the sample mean will be within 0.5 of the population mean.

Solution:

$\mu = \mu \ \& \ \sigma^2 = 1.5$

$P\{\mu - 0.5 < \bar{X} < \mu + 0.5\} \leq 0.95$

$P\left\{\frac{\mu - 0.5 - \mu}{\sqrt{1.5}/\sqrt{n}} < \frac{\bar{X} - \mu}{\sigma/\sqrt{n}} < \frac{\mu + 0.5 - \mu}{\sqrt{1.5}/\sqrt{n}}\right\} = 0.95$

$2\,\Phi\left\{\frac{0.5\sqrt{n}}{\sqrt{1.5}}\right\} - 1 = 0.95$

$\Phi\left\{\frac{0.5\sqrt{n}}{\sqrt{1.5}}\right\} = \frac{1.95}{2} = 0.975$

$$\sqrt{n} = \frac{\Phi^{-1}(0.975)}{0.4082} \Rightarrow n \simeq 23.05 \simeq \boxed{23}$$

ESTIMATION :

Consider a random variable X having the pdf which is in the known function form but it depends upon an unknown parameter θ.

The process of finding approximate value of θ is called estimation.

The representation of pdf with the unknown parameter θ is $f(x,\theta)$

Let $X_1, X_2 \ldots X_n$ be a sample of X & let $x_1, x_2 \ldots x_n$ are corresponding sample values. If $g(x_1, x_2 \ldots x_n)$ is a function of the sample to be used for estimating θ, the $\hat{g}$ is called estimator of θ.

The values $g(x_1, x_2 \ldots x_n)$ is called estimate of θ & represented by, (point estimation)

$$\boxed{\hat{\theta} = g(x_1, x_2 \ldots x_n)}$$

Unbiased estimate :

Let $\hat{\theta}$ be an estimate for θ based on a distribution, we say that $\hat{\theta}$ is an unbiased estimate of θ, if $\boxed{E(\hat{\theta}) = \theta}$

LIKELIHOOD FUNCTION :

Let $X_1, X_2 \ldots X_n$ be a random sample of size n from the distribution having pdf $f(x,\theta)$

Then joint pdf of $(X_1, X_2 \ldots X_n)$ & θ which is defined as likelihood function L, which represents the function of the sample fo

is, $L(x_1, x_2 \dots x_n; \theta) = f(x_1, \theta) \cdot f(x_2, \theta) \dots f(x_n, \theta)$
(RHS is the product of mpdf).

Maximum Likelihood Estimate (MLE):

The value θ for which $L(x_1, x_2 \dots x_n; \theta)$ is maximum. (Alternative representation is $L(\theta)$)

To find MLE for θ,

→ We differentiate $L(\theta)$ w.r.to θ & equal to zero.

i.e, $\frac{d}{d\theta}(L(\theta)) = 0$

(We know that at a maximum value of $L(\theta)$ its derivative w.r.t θ is zero)

* To find MLE for an unknown parameter θ, we solve for θ for which $\frac{d}{d\theta} L(\theta) = 0$. Since the function $L(\theta)$ & $\log(L(\theta))$ have same maximum value which is convinient to find maximum value of $\log(L(\theta))$.

* The concept of finding MLE may be extended when 2 parameters say μ & σ^2 both are unknown but the pdf is known in a functional form. Then we write the pdf of X as, $f(x, \theta_1, \theta_2)$. To find MLE for θ_1, we partially differentiate $f(x, \theta_1, \theta_2)$ w.r.to θ_1, θ_2 to find θ_1 & θ_2.

Date
Page

PROBLEMS.

1\. Let $X_1, X_2 \dots X_n$ be random sample of size n from a normal distribution which has mean θ & variance 1. Find MLE for θ.

Solution:

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

The pdf of X is given by,

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

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

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

$$= \left(\frac{1}{\sqrt{2\pi}}\right)^n e^{-1/2[(x_1-\theta)^2+(x_2-\theta)^2+\dots+(x_n-\theta)^2]}$$

$$= \left(\frac{1}{2\pi}\right)^{n/2} e^{-1/2\sum_{i=1}^{n}(x_i-\theta)^2}$$

Take $\log_e$ on both sides.

$$\log_e(L(\theta)) = \log_e\left\{\left(\frac{1}{2\pi}\right)^{n/2} e^{-1/2\left(\sum_{i=1}^{n}(x_i-\theta)^2\right)}\right\}$$

$$\log_e(L(\theta)) = \frac{n}{2}\log_e\frac{1}{2\pi} - \frac{1}{2}\left(\sum_{i=1}^{n}(x_i-\theta)^2\right)$$

diff w.r.to θ, partially.

$$\frac{d}{d\theta}\left\{\log(L(\theta))\right\} = 0 - \frac{1}{2} \times 2 \sum_{i=1}^{n} (x_i - \theta)(-1)$$

$$\Rightarrow 0 = \sum_{i=1}^{n} (x_i - \theta)$$

$$\sum_{i=1}^{n} x_i - n\theta = 0$$

$$\Rightarrow \sum_{i=1}^{n} x_i = n\theta \Rightarrow \theta = \sum_{i=1}^{n} x_i / n = \underline{\underline{\bar{X}}}$$

$\boxed{\therefore \theta = \bar{X}}$ it shows in the given random sample θ approximately estimate the sample mean.

2. Let $X_1, X_2 \ldots X_n$ be a random sample of size n from a distribution having the pdf $f(x,\theta) = \frac{1}{\theta} e^{-x/\theta}$, $0 < x < \infty$, find MLE for θ.

Solution:

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

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

$$= \frac{1}{\theta^n} e^{-1/\theta(x_1 + x_2 + \ldots + x_n)}$$

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

$$\log_e(L(\theta)) = n\log(1/\theta) - 1/\theta \sum_{i=1}^{n} x_i$$

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

Date
Page

$$\left\{\frac{\partial}{\partial\theta}\log_e(L(\theta))\right\} = \frac{-n}{\theta} + \frac{1}{\theta^2}\sum_{i=1}^{n} x_i$$

$$\Rightarrow \frac{-n}{\theta} + \frac{1}{\theta^2}\sum_{i=1}^{n} x_i = 0$$

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

$$\boxed{\theta = \bar{X}} \Rightarrow \text{sample mean.}$$

---

3. Let $X_1, X_2 \ldots X_n$ denote random sample of size n. from $N(\theta_1, \theta_2)$. Find the M.L.E for θ_1, θ_2.

Solution:

$X \sim N(\theta_1, \theta_2)$

$$f(x, \theta_1, \theta_2) = \frac{1}{\sqrt{2\pi}\sqrt{\theta_2}} e^{-1/2\left(\frac{x-\theta_1}{\sqrt{\theta_2}}\right)^2}$$

OR

$$= \frac{1}{\sqrt{2\pi\theta_2}} e^{-1/2 \frac{(x-\theta_1)^2}{\theta_2}}$$

$$\therefore L(\theta_1, \theta_2) = f(x_1, \theta_1, \theta_2) \times \ldots \times f(x_n, \theta_1, \theta_2)$$

$$= \left(\frac{1}{2\pi\theta_2}\right)^{n/2} e^{-1/2\theta_2\left[(x_1-\theta_1)^2 + (x_2-\theta_1)^2 + \ldots (x_n-\theta_1)^2\right]}$$

Taking log,

$$\log(L(\theta_1,\theta_2)) = \frac{n}{2}\log\frac{1}{2\pi\theta_2} - \frac{1}{2\theta_2}\sum_{i=1}^{n}(x_i-\theta_1)^2$$

differentiate w.r.to θ_1, θ_2 partially,

$$\Rightarrow \sum_{i=1}^{n}(x_i-\theta_1) = 0$$

$$\Rightarrow \sum_{i=1}^{n} x_i = n\theta_1 \Rightarrow \boxed{\theta_1 = \overline{X}} \text{ Sample mean}$$

$$\log(L(\theta_1,\theta_2)) = \frac{n}{2}\log 1 - \frac{n}{2}\log 2\pi\theta_2 - \frac{1}{2\theta_2}\sum_{i=1}^{n}$$

$$0 = -\frac{n}{2}\frac{1}{2\pi\theta_2}(2\pi) - \frac{1}{2}\sum_{i=1}^{n}(x_i-\theta_1)^2 \cdot \frac{-1}{\theta_2^2}$$

$$\sum_{i=1}^{n}(x_i-\theta_1)^2 \cdot \frac{1}{\theta_2^2} = \frac{n}{\theta_2}$$

$$\theta_2 = \frac{\sum_{i=1}^{n}(x_i-\theta_1)^2}{n} = \boxed{S^2} \text{ Sample Variance.}$$

---

4. Let $X_1, X_2 \ldots X_n$ denotes random sample of size n from distribution where pdf,

$$f(x,\theta) = \begin{cases} \theta^x e^{-\theta}/x! & , x = 0,1\ldots \quad 0 \le \theta \le 1 \\ 0 & , \text{otherwise} \end{cases}$$

Solution:

$$L(\theta) = f(x_1,\theta) \times f(x_2,\theta) \times \ldots \times f(x_n,\theta)$$

$$= \frac{\theta^{x_1} e^{-\theta}}{x_1!} \cdot \frac{\theta^{x_2} e^{-\theta}}{x_2!} \cdot \ldots \frac{\theta^{x_3} e^{-\theta}}{x_n!}$$

$$= \frac{\theta^{\sum_{i=1}^{n} x_i}\, e^{-\theta \cdot n}}{\prod_{i=1}^{n} x_i!} \rightarrow \log \prod_{mn} = \sum \quad = \log m + \log n$$

$$\log(L(\theta)) = \log\left(\theta^{\sum_{i=1}^{n} x_i} \cdot e^{-\theta \cdot n}\right) - \log\left(\prod_{i=1}^{n} x_i!\right)$$

Consistency :

Let $\hat{\theta}$ be an estimate of the parameter θ. then, we say that $\hat{\theta}$ is a consistent estimate of θ, if,

$$\lim_{n\to\infty} \text{prob}\left\{ |\hat{\theta} - \theta| \leq \epsilon \right\} = 1,$$ where ϵ

is a small +ve number.

Theorem :

Let $\hat{\theta}$ be an estimate of θ based on a sample of size n. If $\lim_{n\to\infty} E(\hat{\theta}) = \theta$ & $\lim_{n\to\infty} V(\hat{\theta}) = 0$,

then $\hat{\theta}$ is a consistent estimate of θ.

Example for unbiased estimator is,

we know that $E(\bar{x}) = \mu$ & population mean $\Rightarrow \mu$.

i.e, $E(\bar{x}) = E\left\{\frac{\sum x_i}{n}\right\} = \frac{1}{n} \cdot n\mu = \underline{\underline{\mu}}$

---

* Show that a sample variance s^2 is not un-biased estimator of σ^2.

// Solution: To prove $E(s^2) \neq \sigma^2$

By the defination,

$$s^2 = E(x_i - \bar{x})^2$$

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

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

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

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

$$-n(\mu - \bar{x})$$

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

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

Date
Page

$$E(s^2) = \frac{1}{n} E\left(\sum_{i=1}^{n}(x_i-\mu)^2\right) - \underbrace{E(\mu-\bar{x})^2}_{\text{variance of } \bar{x}}$$

$$= \frac{1}{n} \cdot n\sigma^2 - \frac{\sigma^2}{n}$$

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

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

After adding & subtracting μ

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

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

ut $(x_i-\mu)$ $\bar{x}.n - n\mu = n(\bar{x}-\mu) = -n(\mu-\bar{x})$

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

Show that the sample mean $\bar{x}$ is an unbiased and consistent estimate for the population mean μ.

Proof:

Sample mean $\bar{x} = \frac{x_1 + x_2 + \cdots + x_n}{n}$

$$E(\bar{x}) = \frac{1}{n}\left[E(x_1) + E(x_2) + \cdots + E(x_n)\right]$$

$$= \frac{n\mu}{n} = \mu$$

$$V(\bar{x}) = V\left[\frac{x_1 + x_2 + \cdots + x_n}{n}\right]$$

$$= \frac{1}{n^2}\,(V(x_1) + V(x_2) + \cdots + V(x_n)]$$

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

As $n \to \infty$, $\frac{\sigma^2}{n} \to 0$

$\therefore$

Let $X_1, X_2 \ldots X_n$ denote random sample with mean 0 & variance θ. $(0 < \theta < \infty)$. Show that $\frac{\sum x_i^2}{n}$ is unbiased estimator of θ & has variance $\frac{2\theta^2}{n}$.

from normal distribution

Solution:

population mean = 0, variance = θ.

$$V\left\{\frac{\sum x_i^2}{n}\right\} = \frac{2\theta^2}{n}$$

To prove $E\left\{\frac{\sum x_i^2}{n}\right\} = \theta$

$$\frac{1}{n} E\left(\sum_1 x_i^2\right) = \frac{1}{n} \sum E(x_i^2)$$

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

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

$$= \theta - 0$$

$$E\left\{\frac{\sum x_i^2}{n}\right\} = \theta$$

To prove : $V\left\{\frac{\sum x_i^2}{n}\right\} = \frac{2\theta^2}{n}$

We know that if $Z = \frac{X-\mu}{\sigma}$, then $Z \sim N(0,1)$ and $Z^2 \sim \chi^2(1)$

$$Z^2 = \left(\frac{X-\mu}{\sigma}\right)^2 = \frac{X^2}{\sigma^2} = \frac{X^2}{\theta} \sim \chi^2(0,1)$$

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

$$\therefore V\left(\frac{X^2}{\theta}\right) = \frac{1}{\theta^2} V(X^2)$$

$$V(X) = 2n \text{ for } \chi^2(n)$$

$$= 2 \text{ for } \chi^2(1)$$

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

Solution :

Given $X_1, X_2 \ldots X_n \sim N(\mu, \sigma^2) = N(0, \theta)$

i) $E\left\{\sum_{i=1}^{n} x_i^2/n\right\} = \theta$

$$\frac{1}{n}\sum_{i=1}^{n} E(X_i^2) = \frac{1}{n} \cancel{\sum_{i=1}^{n}} nE(X^2)$$

$$= V(X) + (E(X))^2$$

$$= \theta - 0 = \theta$$

ii) $V\left[\sum_{i=1}^{n}\frac{x_i^2}{n}\right] = \frac{2\theta^2}{n}$

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

w.k.T, $X \sim N(\mu, \sigma^2)$

$$Z = \frac{X-\mu}{\sigma} \sim N(0,1)$$

$$Z^2 = \frac{(X-\mu)^2}{\sigma^2} \sim \chi^2(1)$$ [from the resul

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

The mean of $\chi^2(1) \Rightarrow n = \underline{\underline{1}}$

variance $\chi^2(1) \Rightarrow 2n = \underline{\underline{2}}$

$$V\left(\frac{X^2}{\theta}\right) = \frac{1}{\theta^2} V(X^2) = V(X^2) = \underline{2\theta^2}$$

$$\Rightarrow V\left[\sum_{i=1}^{n}\frac{x_i^2}{n}\right] = \frac{1}{n^2} V\left[\sum_{i=1}^{n} x_i^2\right]$$

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

$$= \underline{\frac{1}{n} 2\theta^2}$$

---

* $$S^2 = E(X_i - \overline{X})^2 = \frac{1}{n}\sum_{i=1}^{n}(x_i - \overline{x})^2$$

$$= \frac{1}{n}\left[\sum_{i=1}^{n} x_i^2 + \sum_{i=1}^{n}\overline{x}^2 - 2\overline{x}\sum_{i=1}^{n} x_i\right]$$

$$= \frac{1}{n}\left[\sum_{i=1}^{n} x_i^2 + n\bar{x}^2 - 2\bar{x}\,n\bar{x}\right]$$

$$s^2 = \frac{1}{n}\left[\sum_{i=1}^{n} x_i^2 - n\bar{x}^2\right]$$

$$E(s^2) = \frac{1}{n}\left[E\left(\sum_{i=1}^{n} x_i^2\right) - n\,E(\bar{x}^2)\right]$$

$$X \sim (\mu, \sigma^2) \qquad \bar{X} \sim (\mu, \sigma^2/n)$$

$$= \frac{n}{n} E(X^2) - E(\bar{X}^2)$$

$$= V(x) + (E(X))^2 - \left(V(\bar{X}) + (E(\bar{X}))^2\right)$$

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

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

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