PAGE No.

DATE / /

PROBABILITY & STATISTICS - I
ASSIGNMENT

Q.1. Calculate the mean, median, mode, standard deviation and inter quartile range of following data

i) 4, 7, 9, 9, 11, 15, 18, 18, 18, 25.

- Mean $(\bar{x}) = \dfrac{\sum x_i}{n}$

Where x is the data and n is no. of x.

$$\therefore \bar{x} = \frac{4+7+9+9+11+15+18+18+18+25}{10}$$

$$\bar{x} = 13.4$$

$\therefore$ Mean = 13.4.

- Median:

$n = 10$.

Median = $\dfrac{n+1}{2}$ $\therefore \dfrac{n+1}{2} = \dfrac{10+1}{2} = \dfrac{11}{2} = (5.5)^{th}$ term $= 13$

$\therefore$ Median = 13.

- Mode: Most common is 18

$\therefore$ Mode = 18.

- Standard Deviation -

x_i	$x_i - \bar{x}$	$(x_i - \bar{x})^2$
4	-9.4	88.36
7	-6.4	40.96
9	-4.4	19.36
9	-4.4	19.36
11	-2.4	5.76
15	1.6	2.56
18	4.6	21.16
18	4.6	21.16
18	4.6	21.16
25	11.6	134.56

$\sum x_i = 134$

PAGE No.

DATE / /

$$\sum (x_i - \bar{x})^2 = 374.4$$

Here, $n = 10$

$$\bar{x} = \frac{\sum x_i}{n} = \frac{134}{10} = 13.4$$

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

$$\therefore \sigma_x^2 = \frac{1}{10}(374.4)$$

$$= 37.44$$

Now, Standard deviation $= \sigma_x = \sqrt{37.44}$

$$= 6.11$$

$\therefore$ Standard deviation $= 6.11$

- Interquartile Range :

$$\cancel{Q_1 = \frac{n+1}{4} = \frac{10+1}{4} = \frac{11}{4}}$$

Q_1 = Value of $\left(\frac{n+1}{4}\right)^{th}$ Observation = value of $\left(\frac{10+1}{4}\right)^{th}$ Observation

= value of $(2.75)^{th}$ observation

= Value of 2^{nd} observation + 0.75 (value of 3^{rd} observation - value of 2^{nd} observation)

$= 7 + 0.75(9-7)$

$= 7 + 1.50$

$= 8.15$

Q_3 = Value of $\left[3\left(\frac{n+1}{4}\right)\right]^{th}$ observation

= Value of $\left[3\left(\frac{10+1}{4}\right)\right]^{th}$ observation

= Value of $(8.25)^{th}$ observation

= Value of 8^{th} observation + 0.25 (value of 9^{th} observation - Value of 8^{th} observation)

PAGE No.	
DATE	/ /

$= 18 + 0.285\,(18-18)$

$= 18 + 0$

$= 18$

$\therefore$ Interquartile Range $= Q_3 - Q_1$

$= 18 - 8.15$

$= 10.85$

$\therefore$ **Interquartile Range $= 10.85$**

ii) Find Mode

Number of households	0	1	2	3	4
Number of people	5	11	26	15	3

Number of households (x_i)	Number of people (f_i)	$x_i f_i$
0	5	0
1	11	11
2	26	52
3	15	45
4	3	12
	$\sum f_i = 60$	$\sum x_i f_i = 120$

$$\text{Mean } (\bar{x}) = \frac{\sum x_i f_i}{\sum f_i} = \frac{120}{60} = 2$$

$\therefore$ **Mean $= 2$**

Median:

Number of households (x_i)	Number of people (f_i)	C.f
0	5	5
1	11	16
2	26	42

3	15	57
4	3	60
	$\Sum f_i = 60$	

$\Sigma f_i = N = 60$

$\frac{N+1}{2} = \frac{60+1}{2} = \frac{61}{2} = 30.5$

$\therefore$ Median = 2.

- Mode :

Most frequent

$\therefore$ Mode = 3.

- Standard Deviation:

Number of household	Midpoint (x_i)	No. of students (f_i)	$f_i x_i$	x_i^2	$f_i x_i^2$
0 - 1	0.5	5	2.5	0.25	1.25
1 - 2	1.5	11	16.5	2.25	24.75
2 - 3	2.5	26	65	6.25	162.5
3 - 4	3.5	15	52.5	12.25	183.75
4 - 5	4.5	3	13.5	20.25	60.75
		N = 60	150		$\Sigma f_i x_i^2 = 433$

$\bar{x} = \frac{\Sigma f_i x_i}{N} = \frac{150}{60} = 2.5$

$\sigma_x^2 = \frac{1}{N} \Sigma f_i x_i^2 - (\bar{x})^2 = \frac{1}{60}(433) - (2.5)^2$

$\sigma_x^2 = 0.9666$

$S.D(x) = \sigma_x = \sqrt{0.9666} = 0.98319$

- Interquartile Range: (referring to 1st table)

$N = 60$

$\therefore Q_1 =$ Value of $\left(\frac{N+1}{4}\right)^{th}$ observation

PAGE No.
DATE / /

$= \text{Value of } \left(\frac{60+1}{4}\right)^{th}$ observation

$= \text{Value of } (15.25)^{th}$ observation

Now, 15.25 is close to 16 in cumulative frequency table.

$\therefore Q_1 = 1.$

$Q_3 = \text{Value of } \left[3\left(\frac{N+1}{4}\right)\right]^{th}$ Observation.

$= \text{Value of } \left[3\left(\frac{60+1}{4}\right)\right]^{th}$ observation

$= \text{Value of } [3 \times 15.25]^{th}$ observation

$= \text{Value of } [45.75]^{th}$ observation

Now, 45.75 is close to 57 in cumulative frequency table.

$\therefore Q_3 = 3.$

$\therefore$ Interquartile Range $= Q_3 - Q_1 = 3 - 1 = 2$

$\therefore$ Interquartile Range $= 2$

Q. 2. It is assumed that claims arising on an industrial policy can be modelled as a Poisson Process at a rate of 0.5 per year.

i) $X \sim Poi(0.5)$

$P(X \leq 0) = 0.60653$

$\boxed{\therefore \text{Probability that no claims arise in a single year} = 0.60653}$

ii)

Q.3 An analyst is interested in using gamma distribution with parameters $\alpha = 2$ and $\lambda = \frac{1}{2}$. The range of x is defined as $0 < x < \infty$

→ i) State the Mean & SD of this distribution

$\alpha = 2$ and $\lambda = \frac{1}{2}$

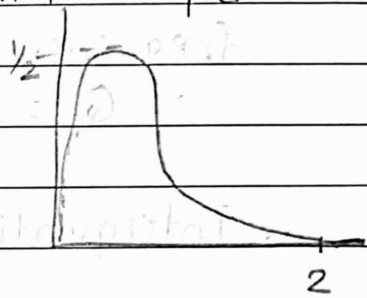
$X \sim Gamma(2, \frac{1}{2})$

$$Mean\ (E(X)) = \frac{\alpha}{\lambda} = \frac{2}{\frac{1}{2}} = \underline{\underline{4}}$$

$$SD\ (V(x)) = \frac{\alpha}{\lambda^2} = \frac{2}{(\frac{1}{2^2})} = \frac{2}{\frac{1}{4}} = \underline{\underline{8}}$$

ii) Hence comment briefly on its shape

It touches on 2 on x-axis and $\frac{1}{\alpha}$ on y-axis.



iii) Calculate the cumulative distribution function

$\therefore$ ~~P(α<x<∞) e-of~~

$$\cancel{\Gamma(\alpha) = \int_0^\infty y^{\alpha-1} e^{-y}\,dy}$$

$$\Gamma(\alpha) = (\alpha - 1)!$$
$$= (2-1)!$$
$$= \underline{\underline{1}}$$

$$\cancel{= \int_0^\infty y^1 e^{-y}\,dy}$$

$$\cancel{= \left[\frac{y^2}{2} \cdot e^{y} \cdot (-1)\right]_0^\infty = \left[\frac{-y^2}{2} \cdot e^{y}\right]_0^\infty}$$

$$\cancel{= 0 - 0 = 0}$$

$$\therefore f_x(x) = \frac{\lambda^\alpha e^{-\lambda x} x^{\alpha-1}}{\Gamma(\alpha)} = \frac{(\frac{1}{2})^2 e^{-\frac{1}{2}x} x^{2-1}}{1}$$

$$= \underline{\frac{e^{-\frac{1}{2}x}\, x}{4}}$$

Q.4. In an actuarial office, there are 10 male & 8 female staff. 6 of the males and 7 of the females are fully qualified actuaries. Calculate the probability that a team of two workers randomly selected consists of two qualified females given that it contains at least one female.

→ The staff consists of 10 ~~boys~~ males & 8 females

i.e. $6 + 7 = 13$ persons

A team of 2 is to be formed from this group.

∴ 2 persons can be selected from 13 persons in $^{13}C_2$ ways

$$\therefore n(S) = {}^{13}C_2 = 78$$

Now,

Let A be the event that the team contains at least one female

(i.e., one female and one male or two females)

$$\therefore n(A) = {}^6C_1 \cdot {}^7C_1 + {}^6C_0 \, {}^7C_2$$

$$\therefore n(A) = (6)(7) + (1)(21)$$

$$n(A) = 63$$

$$\therefore P(A) = \frac{n(A)}{n(S)} = \frac{63}{78} = 0.8076$$

$\boxed{\therefore \text{ Probability that team consists of atleast 1 female} = 0.8076}$

Q.5. Let the probability that a late package was sent via Level 2 service be $P(2/L)$

$$\therefore P(2/L) = \frac{P(L/2) \cdot P(2)}{P(L/1) \cdot P(1) + P(L/2) \cdot P(2) + P(L/3) \cdot P(3)}$$

$$= \frac{(0.5)(0.4)}{(0.4)(0.5) + (0.5)(0.4) + (0.1)(0.1)}$$

$$= \underline{\underline{0.49723}}$$

$\boxed{\text{Probability that a late package was sent via level 2} = 0.49723}$

Q.6. If $X \sim U(1,2)$ and $Y = \dfrac{3}{6-X}$ determine the PDF of Y.

$\rightarrow$ $\cancel{P(g) \le p(Y \le y)}$

$$f_Y(y) = P(Y \le y)$$

$$= P\left(\frac{3}{6-X} \le y\right).$$

$$= P\left(\frac{-3-6Y}{Y} \le X \le \frac{+3+6Y}{Y}\right).$$

$$= \frac{-3+6(2)}{2} = \underline{\underline{\frac{-15}{2}}}$$

Q.7. A discrete random variable X has the foll. disⁿ.

X	1	2	3	4	5
$P(X=x)$	0.05	0.15	0.35	0.4	0.05

Calculate.

i) $F(3) = P(X \le 3) = P(X = 1 \text{ or } 2 \text{ or } 3)$

$= p(1) + p(2) + p(3)$

$= 0.55$

$\therefore F(3) = 0.55$

ii) $E(X) = \sum P(X = x) \cdot x$

$\underline{E(X) = 3.25}$

PAGE No
DATE / /

iii) $V(x) = (x-\mu)^2 P(X=x)$

$\mu = 1\times0.05 + 2\times0.15 + 3\times0.35 + 0.4\times4 + 0.05\times5$

$\mu = 3.25$

$Var(x) = (1-3.25)^2 \times 0.05 + (2-3.25)\times0.15 + (3-3.25)\times0.35 + (4-3.25)\times0.4 + (5-3.25)\times0.05$

$= 0.253125 + 0.234375 + 0.021875 + 0.225 + 0.153125$

$= 0.8875$

2nd method

Using $Var(x) = E(x^2) - [E(x)]^2$

$E(x) = 3.25$

$E(x^2) = 1^2\times0.05 + 2^2 * 0.15 + 3^2\times0.35 + 4^2\times0.4 + 5^2\times0.05$

$= 0.05 + 0.6 + 3.15 + 6.4 + 1.25$

$= 11.45$

$Var(x) = 11.45 - (3.25)^2$

$= 0.8875.$

iv) Coefficient of skewness $= \dfrac{3\,(\text{mean}-\text{median})}{\sigma}$

Mean $(E(x)) = 3.25$

Median $=$ ~~0.05~~ ?

$P(X\le x) = 0.05,\ 0.2,\ 0.55,\ 0.95,\ 1$

Median $= M = \dfrac{x_1 + x_2}{2} = \dfrac{2+3}{2} = \dfrac{5}{2}$

$\therefore$ Median $= 2.5$

$\sigma = 0.8875$

$\therefore$ Coefficient of skewness $= \dfrac{3\,(3.25 - 2.5)}{0.8875}$

$\therefore$ Coefficient of skewness $= 2.5352$

PAGE No. 5

DATE / /

Q.8. Insurance claim amounts are independent and exceed 1000 with probability 0.2. In a sample of 15 claims, calculate the probability that fewer than 3 of them exceed 1000.

→ Let $p = 0.2$ $\therefore q = 0.8$ $(\because p+q=1)$

We know that,

$P(X=x) = {}^{n}C_{x}\, p^{x} q^{n-x}$

But we have $P(X<3)$

$\therefore P(X<3) = P(X=0) + P(X=1) + P(X=2)$

$= {}^{15}C_{0}(0.2)^{0}(0.8)^{15} +$

${}^{15}C_{1}(0.2)^{1}(0.8)^{14} +$ ... $(n=15$ given$)$

${}^{15}C_{2}(0.2)^{2}(0.8)^{13}$

$= (1)(1)(0.035184) + (15)(0.2)(0.0439804)$

$+ (105)(0.04)(0.054975)$

$= 0.035184 + 0.1319412 + 0.2308974$

$= 0.398022$

$\therefore$ Probability that fewer than 3 of them exceed 1000 is 0.398022

Q.9 A man enters a raffle each week where the probability of winning prize is 0.01. Calculate the probability that he takes seven weeks up to and including the week where he wins his third prize.

→ $p = 0.01$ (winning), $q = 0.09$ (losing)

$$\begin{array}{r} 1 \\ 0.01 \\ 0.09 \\ \hline 0.10 \end{array}$$

$n = 7$

$P(X=x) = {}^{n}C_{x}\, p^{x} q^{n-x}$

$P(X=3) = P(X=\cancel{0}) + P(X=2) + P(X=3)$

$= {}^{7}C_{1}(0.01)^{1}(0.09)^{6} + {}^{7}C_{2}(0.01)^{2}(0.09)^{5}$

${}^{7}C_{3}(0.01)^{3}(0.09)^{4}$

PAGE No

DATE / /

$$= (7)(0.01)(0.00000000531441) + (21)(0.01)^2(0.0000005904A) + 35(0.00001)(0.00006561)$$

$$= 3.7200 + \cancel{0.0000}1.240029 + 2.2963$$

$$= 7.25629$$

Q.11 If the probability of having a male and female child is equal and a woman has 5 boys and two girls, Calculate the probability that her next three children are boys

→ Given that,

Prob. of having a male and female is equal.

$\therefore P(M) = P(F) = 0.5 = \frac{1}{2}$

Let $P(B)$ be the probability that her next three children are boys.

$\therefore P(B) = 0.5 \times 0.5 \times 0.5$

$= \underline{\underline{0.125}}$

$\boxed{\therefore \text{Probability that her next three children are boys is } 0.125}$

Q.10. The number of claims submitted to an insurance office averages two per working week (of 5 days). Calculate the probability that more than two claims arrive in one day.

→ ~~Let X~~

Let X be rv → Number of claims

$X \sim Exp(\lambda)$

Here, $\lambda = 2$

$X \sim Exp(2)$

$P(X > 2) = e^{-\lambda(2)} = e^{-2(2)}$

$= e^{-4} = 0.0183156$

$\therefore$ Probability that more than two claim arrive in one day $= 0.0183156$

Q.12. Claims follow a Poisson process with rate 10 per day. Calculate the probability that the time between two claims is less than 0.1 days.

→ Here $p = 0.1$ days

$n = 10$

$X \sim Poi(2)$

$P(X < 0.1) = 0.99985$

$X \rightarrow 2$

$\mu \lambda$

$0.1 \rightarrow$

$\frac{e^{-\mu}\mu^x}{x!}$

$\therefore$ Probability that the time between two Clams is 0.99985

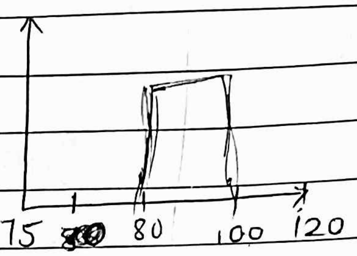
PAGE No.

DATE / /

Q.13. If x is following uniform distribution with parameters $a = 75$ and $b = 120$, calculate the probability that x lies between 80 & 100.

$\rightarrow$ It is a continuous uniform distribution

$\therefore X \sim U(a, b)$



$$F_X(x) = \frac{x-a}{b-a}$$

$$= \frac{97.5-75}{120-75}$$

$$= \frac{22.5}{45}$$

$$= 0.5$$

$$\therefore F_X(x) = 0.5.$$

Q.14. If $X \sim Gamma(5, 0.4)$ calculate $P(X < 22.89)$.

$\rightarrow$ $\alpha = 5, \quad \lambda = 0.4$

$$P(X < 22.89) = \int_0^{22.89} \frac{(0.4)^5 \cdot e^{-0.4x} \cdot x^4}{\Gamma(\alpha)}$$

$\Gamma(\alpha) = 4!$

$$= \frac{(0.4)^5}{4!}\left[\frac{-1}{0.4} e^{-0.4x} \cdot \frac{x^5}{5}\right]_0^{22.89}$$

$$= \frac{(0.4)^5}{4!}\left[\left(\frac{-1}{0.4} \cdot e^{-0.4(22.89)} \cdot \frac{(22.89)^5}{5}\right) - 0\right]$$

$$= 0.0004266 \left[0.000263960 \times 331.29\right]$$

$$= 0.141520035.$$

$$\therefore P(X < 22.89) = 0.141520035$$

0.01024

.1906761

PAGE No. | DATE / /

Q.15. A continuous rv, X has the following probability density function: $\frac{k}{(x+5)^4}$; $x > 0$

Calculate:

→ i) k

$$f_X(x) = \frac{k}{(x+5)^4}$$

Since it is p.d.f. it will be

1) $f(x) \geq 0$

2) $\int_{-\infty}^{\infty} f(x)\,dx = 1.$

$$\therefore \int_0^{\infty} \frac{k}{(x+5)^4}\,dx = 1.$$

$$k \int_0^{\infty} \frac{1}{(x+5)^4}\,dx = 1.$$

$-4+1$

-3

$$k \left[\frac{(x+5)^{-3}}{-3} \right]_0^{\infty} = 1.$$

$(0+5)^{-3}$

$\frac{\infty}{-3}$

$\frac{(5)^{-3}}{-3}$

$$k\,[\,0 - 0.0026666\,] = 1$$

$$-0.00266666\,k = 1$$

$$k = -375$$

ii) $F(10) = \frac{-35}{(10+5)^4} = 0.00069135$

iii) $E(X) = \int_0^{22.89} x \cdot \frac{22.89}{(x+5)^4}$

$$= 22.89 \int_0^{22.89} x \cdot \frac{1}{(x+5)^4}$$

$$= 22.89\left[x\cdot\int(x+5)^{-4}dx - \int\left[\int(x+5)^{-4}dx\right](1)\,dx\right]_0$$

$$= 22.89\left[x\cdot\frac{(x+5)^{-3}}{-3} - \int\frac{(x+5)^{-3}}{-3}dx\right]_0^{22.89}$$

$$= \frac{22.89}{6}\left[x\cdot(x+5)^{-3} - (x+5)^{-2}\right]_0^{22.89}$$

$$= 22.89\left[(22.89)(22.89+5)^{-3} - (22.89+5)^{-2}\right] - \left[(-5)^{-2}\right]$$

$$= 22.89\left[+105.4987 + 0.04\right]$$

$$= \underline{\underline{2415.810}}$$