MAILAM Engineering College
Mailam, Villupuram(Dt), Pin: 604 304
(Approved by AICTE, New Delhi, Affiliated to Anna University, Chennai,
Accredited by NBA, NAAC with 'A' Grade & TATA Consultancy Services)

DEPARTMENT OF MATHEMATICS

SUB CODE/NAME: MA3391/ Probability and Statistics

YEAR/SEM: II / IV **CLASS: AI&DS, CSBS**

Unit	Part – A				Part – B				Month/Year
	Q. No	Page No.	Q. No	Page No.	Q. No	Page No.	Q. No	Page No.	
01	63	20	64	21	75,76	85,86	77,78	87,88	Apr/May24
	65	21	66	21	27,79	48,89	2,37	23,60	Nov/Dec24
02	53	21	54	21	60	89	61	91	Apr/May24
	55	22	56	22	49,62	80,94	22,36,63	51,70,94	Nov/Dec24
03	28	6	29	6	25,26	28,29	16,27	20,29	Apr/May24
	30	7	10	2	8,17	14,22	28,28	30,31	Nov/Dec24
04	3	1	17	3	12,16	16,20	17,22	22,28	Apr/May24
	18	4	5	1	6,29	9,35	23,30	29,36	Nov/Dec24
05	22	3	23	4	9	13	16,21	21,38	Apr/May24
	25	4	4,6	1	26,10	28,14	22,17	25,22	Nov/Dec24

STAFF IN-CHARGE
M.BALAMURUGAN, AP/Mathematics

HOD/MATHS

PRINCIPAL

MA3391 PROBABILITY AND STATISTICS L T P C 3 1 0 4

COURSE OBJECTIVES

- This course aims at providing the required skill to apply the statistical tools in engineering problems.
- To introduce the basic concepts of probability and random variables.
- To introduce the basic concepts of two dimensional random variables.
- To acquaint the knowledge of testing of hypothesis for small and large samples which plays an important role in real life problems.
- To introduce the basic concepts of classifications of design of experiments which plays very important roles in the field of agriculture and statistical quality control.

UNIT I PROBABILITY AND RANDOM VARIABLES 9 + 3

Axioms of probability – Conditional probability – Baye's theorem - Discrete and continuous random variables – Moments – Moment generating functions – Binomial, Poisson, Geometric, Uniform, Exponential and Normal distributions – Functions of a random variable.

UNIT II TWO- DIMENSIONAL RANDOM VARIABLES 9 + 3

Joint distributions – Marginal and conditional distributions – Covariance – Correlation and linear regression – Transformation of random variables – Central limit theorem (for independent and identically distributed random variables).

UNIT III ESTIMATION THEORY 9 + 3

Unbiased estimators - Efficiency - Consistency - Sufficiency - Robustness - Method of moments - Method of maximum Likelihood - Interval estimation of Means - Differences between means, variations and ratio of two variances.

UNIT IV NON- PARAMETRIC TESTS 9 + 3

Introduction - The Sign test - The Signed - Rank test - Rank - sum tests - The U test - The H test - Tests based on Runs - Test of randomness - The Kolmogorov Tests .

UNIT V STATISTICAL QUALITY CONTROL 9 + 3

Control charts for measurements (¯ and R charts) – Control charts for attributes (p, c and np charts) – Tolerance limits - Acceptance sampling.

TOTAL: 60

PERIODS COURSE OUTCOMES:

Upon successful completion of the course, students will be able to:

CO1: Understand the fundamental knowledge of the concepts of probability and have knowledge of standard distributions which can describe real life phenomenon.

CO2: Understand the basic concepts of one and two dimensional random variables and apply in engineering applications.

CO3: Apply the concept of testing of hypothesis for small and large samples in real life problems.

CO4: Apply the basic concepts of classifications of design of experiments in the field of agriculture and Statistical quality control.

CO5: Have the notion of sampling distributions and statistical techniques used in engineering and management problems.

TEXT BOOKS

1. Johnson. R.A., Miller. I.R and Freund . J.E, " Miller and Freund's Probability and Statistics for Engineers", Pearson Education, Asia, 9th Edition, 2016.
2. Milton. J. S. and Arnold. J.C., "Introduction to Probability and Statistics", Tata Mc Graw Hill, 4th Edition, 2007.
3. John E. Freund, "Mathematical Statistics", Prentice Hall, 5th Edition, 1992.

REFERENCES:

1. Gupta. S.C. and Kapoor. V. K., "Fundamentals of Mathematical Statistics", Sultan Chand & Sons, New Delhi, 12th Edition, 2020.

2. Devore. J.L., "Probability and Statistics for Engineering and the Sciences", Cengage Learning, New Delhi, 8th Edition, 2014.

3. Ross. S.M., "Introduction to Probability and Statistics for Engineers and Scientists", 5 th Edition, Elsevier, 2014.

4. Spiegel. M.R., Schiller. J. and Srinivasan. R.A., "Schaum's Outline of Theory and Problems of Probability and Statistics", Tata McGraw Hill Edition, 4th Edition, 2012.

5. Walpole. R.E., Myers. R.H., Myers. S.L. and Ye. K., "Probability and Statistics for Engineers and Scientists", Pearson Education, Asia, 9th Edition, 2010.

Unit-I (Probability and Random Variables)

PART-A

1. Define Random variable. **[AU N/D 2013]**

Solution:

A real-valued function defined on the outcome of a probability experiment is called a random variable.

2. Obtain the mean for a Geometric distribution. **[AU A/M 2010]**

Solution:

We know that $M_X(t) = \dfrac{pe^t}{(1-qe^t)}$

Mean,
$$E(X) = \left[\frac{d}{dt}\left[M_X(t)\right]_{t=0}\right]$$
$$= \left[\frac{d}{dt}\left[\frac{pe^t}{(1-qe^t)}\right]_{t=0}\right]$$
$$= \left[\frac{d}{dt}\left[\frac{p}{(e^{-t}-q)}\right]_{t=0}\right]$$
$$= \left[\frac{pe^{-t}}{(e^{-t}-q)^2}\right]_{t=0}$$
$$= \frac{p}{(1-q)^2} = \frac{p}{p^2} = \frac{1}{p}$$

3. What is meant by memory less property? Which continuous distribution follows this property?

Solution: **[AU A/M 2010]**

If X is a random variable (discrete or continuous), then for any two positive integers m and n. $P[X > m+n / X > m] = P[X > n]$ which is the memory less property.

Exponential distribution follows this property.

4. If a random variable X has the distribution function $F(X) = \begin{cases} 1-e^{-\lambda x} & for\ x > 0 \\ 0 & for\ x \le 0 \end{cases}$

Where λ is the parameter, then find $P(1 \le X \le 2)$ **[AU N/D 2010]**

Solution:

We know that $P(a \le X \le b) = F(b) - F(a)$
$$P(1 \le X \le 2) = F(2) - F(1)$$
$$= \left(1-e^{-2\lambda}\right) - \left(1-e^{-\lambda}\right)$$
$$= e^{-\lambda} - e^{-2\lambda}$$

5. Every week the average number of wrong-number phone calls received by a certain mail order house is seven. What is the probability that they will receive two wrong calls tomorrow?

Solution: **[AU N/D 2010]**

Given : $\lambda = \frac{7}{7} = 1$

$$P[X = x] = \frac{e^{-\lambda}\lambda^x}{x!}$$

$$P[X = 2] = \frac{e^{-1}1^2}{2!} = \frac{e^{-1}}{2}$$

6. A continuous random variable X that can assume any value between $x = 2$ *and* $x = 5$ has a density function given by $f(x) = k(1+x)$. Find $P(X < 4)$.

Solution: **[AU N/D 2011]**

Formula $\int_{-\infty}^{\infty} f(x)dx = 1$

$$\int_2^5 k(1+x)dx = 1$$

$$k\left[x + \frac{x^2}{2}\right]_2^5 = 1$$

$$k\left[\left(5 + \frac{25}{2}\right) - \left(2 + \frac{4}{2}\right)\right] = 1$$

$$k\left[\frac{27}{2}\right] = 1$$

$$k = \frac{2}{27}$$

$$P[X < 4] = \int_2^4 f(x)dx$$

$$= \int_2^4 k(1+x)dx$$

$$= \int_2^4 \frac{2}{27}(1+x)dx$$

$$= \frac{2}{27}\left[x + \frac{x^2}{2}\right]_2^4$$

$$= \frac{2}{27}[12 - 4]$$

$$= \frac{16}{27}$$

7. Give the Probability law of Poisson distribution and also its mean and variance.

Solution: **[AU N/D 2011]**

The Probability law of Poisson distribution is given by

$$P(X = x) = p(x) = \frac{e^{-\lambda}\lambda^x}{x!}$$

Mean, $E(X) = \lambda$ and Variance $= Var(X) = \lambda$

8. The cumulative distribution function of the random variable X is given by

$$F_X(x) = \begin{cases} 0 \ , \ x < 0 \\ x + \frac{1}{2} \ , \ 0 \le x \le \frac{1}{2} \\ 1 \ , \ x > \frac{1}{2} \end{cases}$$; **Compute** $P\left[X > \frac{1}{4}\right]$ **[AU A/M 2011]**

Solution:

$$P\left[X > \frac{1}{4}\right] = 1 - P\left[X \le \frac{1}{4}\right] = 1 - F\left[\frac{1}{4}\right]$$

$$= 1 - \left(\frac{1}{4} + \frac{1}{2}\right) = 1 - \frac{3}{4} = \frac{1}{4}$$

9. Let the random variable X denotes the sum of obtained in rolling a pair of fair dice. Determine the probability mass function of X. **[AU A/M 2011]**

Solution:

x	**2**	**3**	**4**	**5**	**6**	**7**	**8**	**9**	**10**	**11**	**12**
$p(x)$	$\frac{1}{36}$	$\frac{2}{36}$	$\frac{3}{36}$	$\frac{4}{36}$	$\frac{5}{36}$	$\frac{6}{36}$	$\frac{5}{36}$	$\frac{4}{36}$	$\frac{3}{36}$	$\frac{2}{36}$	$\frac{1}{36}$

10. Check whether the following is a probability density function or not.

$$f(x) = \begin{cases} \lambda e^{-\lambda x}, x \ge 0, \lambda > 0 \\ 0 \ , \ elsewhere \end{cases}$$

[AU M/J 2012]

Solution:

$$To\ \ prove \int_{-\infty}^{\infty} f(x)dx = 1$$

$$\int_{-\infty}^{\infty} f(x)dx = \int_{0}^{\infty} \lambda e^{-\lambda x} dx \ = \lambda\left[\frac{e^{-\lambda x}}{-\lambda}\right]_0^{\infty}$$

$$= -\left[e^{-\lambda x}\right]_0^{\infty} = -(0-1) = 1$$

$\therefore$ The given function is a p.d.f.

11. If a random variable has the moment generating function given by $M_X(t) = \dfrac{2}{2-t}$**, determine the variance of X.** **[AU M/J 2011]**

Solution:

Given $\quad M_X(t) = \dfrac{2}{2-t} = 2(2-t)^{-1}$

$$M_X^{!}(t) = 2(2-t)^{-2}$$

$$M_X^{!!}(t) = 4(2-t)^{-3}$$

$$E(X) = M_X^{!}(0) = 2(2-0)^{-2}$$

$$= 2 \times 2^{-2}$$

$$= \frac{2}{4} = \frac{1}{2}$$

$$E(X^2) = M_X''(0) = 4(2-0)^{-3}$$
$$= 4\times 2^{-3} = \frac{4}{2^3} = \frac{4}{8} = \frac{1}{2}$$
$$Var(X) = E(X^2) - [E(X)]^2$$
$$= \frac{1}{2} - \frac{1}{4}$$
$$= \frac{1}{4}$$

12. Identify the random variable and name the distribution it follows, from the following statements: "A realtor claims that only 30% of the houses in a certain neighborhood, are appraised at less than Rs. 20 lakhs. A random sample of 10 houses from that neighborhood is selected and appraised to check the realtor's claim is acceptable are not". [AU M/J 2011]

Solution:

X is a binomial random variable with parameter (n, p)

Here $n = 10,\ p = \frac{30}{100} = \frac{3}{10}$.

13. A random variable 'X' has the following probability function.

x	0	1	2	3	4	5	6	7	8
$p(x)$	a	$3a$	$5a$	$7a$	$9a$	$11a$	$13a$	$15a$	$17a$

Determine the value of a. [AU M/J 2004]

Solution:

We know that if $P(x)$ is the probability mass function, then

$$\sum_{i=1}^{\infty} p(x_i) = 1$$

$$a + 3a + 5a + 7a + 9a + 11a + 13a + 15a + 17a = 1$$
$$81a = 1$$
$$a = \frac{1}{81}$$

14. A random variable 'X' has the following probability function

$x:$	0	1	2	3	4
$p(x):$	K	3K	5K	7K	9K

Find the value of K. [AU M/J 2006]

Solution:

We know that if $P(x)$ is the probability mass function, then

$$\sum_{i=1}^{\infty} p(x_i) = 1$$

$$K+3K+5K+7K+9K=1 \Rightarrow 25K=1 \Rightarrow K=\frac{1}{25}$$

15. Show that the function

$$f(x)=\begin{cases} \frac{x^2}{3}, & -1<x<2 \\ 0, & otherwise \end{cases} \quad is \ a \ pdf.$$

[AU M/J 2006,2009]

Solution:

$$\int_{-\infty}^{\infty} f(x)dx = \int_{-1}^{2} \frac{x^2}{3}dx = \frac{1}{3}\left[\frac{x^3}{3}\right]_{-1}^{2} = \frac{1}{9}[8+1]=1$$

Therefore the given function is a pdf.

16. If 'X' is a continuous random variable whose probability density function is given by $f(x)=\begin{cases} c(4x-2x^2), & 0<x<2 \\ 0, & otherwise \end{cases}$ **What is the value of 'c'?**

Solution:

We know that $\int_{-\infty}^{\infty} f(x)dx = 1$

$$\int_{0}^{2} c(4x-2x^2)dx = 1$$

$$2c\int_{0}^{2} (2x-x^2)dx = 1$$

$$2c\left[\frac{2x^2}{2}-\frac{x^3}{3}\right]_0^2 = 1 \Rightarrow \quad 2c\left[4-\frac{8}{3}\right]=1$$

$$2c\left[\frac{4}{3}\right]=1 \Rightarrow c=\frac{3}{8}$$

17. Given that the p.d.f of a R.V 'X' is $f(x)=Kx,\ 0<x<1$**, Find K and** $P(X>0.5)$

Solution: **[AU M/J 2005]**

We know that $\int_{-\infty}^{\infty} f(x)dx = 1$ $\qquad P(X>0.5)=\int_{0.5}^{\infty} f(x)dx$

$$\int_{0}^{1} Kxdx = 1 \qquad\qquad = \int_{0.5}^{1} 2xdx$$

$$K\left[\frac{x^2}{2}\right]_0^1 = 1 \qquad\qquad = 2\left[\frac{x^2}{2}\right]_{\frac{1}{2}}^{1}$$

$$K\left[\frac{1}{2}\right]=1 \qquad\qquad =1-\frac{1}{4}$$

$$K=2 \qquad\qquad =\frac{3}{4}$$

18. A random variable 'X' has a p.d.f $f(x)=K,\ 0<x<1$**, Find K . [AU N/D 2005]**

Solution:

Since $f(x)$ *is a pdf*

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

$$\int_{0}^{1} Kdx=1 \Rightarrow K(1-0)=1 \Rightarrow K=1$$

19. A random variable 'X' has a p.d.f $f(x)=\begin{cases}\frac{1}{4}, & |X|<2\\ 0, & otherwise\end{cases}$ **, Find** $P(X<1)$**.**

Solution: [AU A/M 2017]

Given $f(x)=\begin{cases}\frac{1}{4}, & |X|<2\\ 0, & otherwise\end{cases}$

$$P(X<1)=\int_{-\infty}^{1} f(x)dx=\int_{-2}^{1}\frac{1}{4}dx=\frac{1}{4}[x]_{-2}^{1}=\frac{1}{4}[1+2]=\frac{3}{4}$$

$$P(X>1)=1-P(X<1)=1-\frac{3}{4}=\frac{1}{4}$$

20. A random variable 'X' has a p.d.f $f(x)=\begin{cases}2x, & 0<x<1\\ 0, & otherwise\end{cases}$**, Find** $P\left(X<\frac{1}{2}\right)$**.**

Solution: [AU M/J 2007]

Given $f(x)=2x,\ 0<x<1$

$$P\left(X<\frac{1}{2}\right)=\int_{0}^{\frac{1}{2}} 2xdx=2\left(\frac{x^2}{2}\right)_0^{\frac{1}{2}}=\frac{1}{4}-0=\frac{1}{4}$$

21. A random variable 'X' has a p.d.f $f(x)=3x^2,\ 0<x<1$**, Find 'b' such that** $P(X>b)=0.05$ **[AU M/J 2005]**

Solution:

Given that $f(x)=3x^2,\ 0<x<1$

when $P(X>b)=0.05$

$$\int_b^1 f(x)dx = 0.05$$

$$\int_b^1 3x^2 dx = 0.05$$

$$3\left(\frac{x^3}{3}\right)_b^1 = 0.05$$

$$1 - b^3 = \frac{1}{20} \Rightarrow b^3 = 1 - \frac{1}{20} = \frac{19}{20} \Rightarrow b = 0.9830$$

22. For the following density function $f(x) = ae^{-|x|},\ -\infty < x < \infty$ **, Find the value of 'a' and E(X).** **[AU M/J 2006, 2021]**

Solution:

Given that $f(x) = ae^{-|x|},\ -\infty < x < \infty$

$f(x)$ is p.d.f

$$\therefore \int_{-\infty}^{\infty} f(x)dx = 1 \qquad \Rightarrow \int_{-\infty}^{\infty} ae^{-|x|}dx = 1$$

$$2a\int_0^{\infty} e^{-x}dx = 1 \Rightarrow 2a\left[-e^{-x}\right]_0^{\infty} = 1$$

$$2a[0+1] = 1 \Rightarrow a = \frac{1}{2}$$

$$E(x) = \int_{-\infty}^{\infty} xf(x)dx = \int_0^{\infty} \frac{1}{2}xe^{-x}dx = \frac{1}{2}\left\{x\left(-e^{-x}\right) - (1)(e^{-x})\right\}_0^{\infty} = \frac{1}{2}(0+1) = \frac{1}{2}$$

23. In a continuous random variable 'X' having the p.d.f

$$f(x) = \begin{cases} \dfrac{x^2}{3}, & -1 < x < 2 \\ 0, & otherwise \end{cases} \qquad Find \quad P(0 < x \le 1)$$

Solution:

$$P(0 < x \le 1) = \int_0^1 f(x)dx = \int_0^1 \frac{x^2}{3}dx$$

$$= \frac{1}{3}\left[\frac{x^3}{3}\right] = \frac{1}{3}\left[\frac{1}{3} - 0\right] = \frac{1}{9}$$

24. A random variable 'X' has the density function $f(x) = K\dfrac{1}{1+x^2}\ in\ -\infty < x < \infty$

Find 'K'. **[AU N/D 2007,M/J 14]]**

Solution:

Since $f(x)$ is a p.d.f., we've

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

$$K\int_{-\infty}^{\infty} \frac{1}{1+x^2}\, dx = 1 \Rightarrow K\left(\tan^{-1} x\right)_{-\infty}^{\infty} = 1$$

$$K\left(\tan^{-1}(\infty) - \tan^{-1}(-\infty)\right) = 1$$

$$K\left(\frac{\pi}{2}+\frac{\pi}{2}\right) = 1 \Rightarrow K\pi = 1 \Rightarrow K = \frac{1}{\pi}$$

25. For the following c.d.f. **[AU N/D 2004]**

$$F(x) = \begin{cases} 0\ ,\ \ x<0 \\ x\ , 0 \le x \le 1 \\ 1\ ,\ x>1 \end{cases}\ ,\ \textit{find}\ \ (i)\ P(X>0.2)\ (ii)\ \ P(0.2<X<0.5)$$

Solution:

$$(i)\quad P(X>0.2) = 1 - P(X \le 0.2)$$

$$= 1 - F(0.2) = 1 - 0.2 = 0.8$$

$$(ii)\quad P(0.2 < X \le 0.5) = F(0.5) - F(0.2)$$

$$= 0.5 - 0.2 = 0.3$$

26. The density function of a random variable 'X' is given by
$f(x) = Kx(2-x), 0 \le x \le 2$ **Find K.**

Solution: **[AU N/D 2008]**

Given $f(x) = Kx(2-x), 0 \le x \le 2$

$f(x)$ *is a* $p.d.f$

$$\therefore \int_{-\infty}^{\infty} f(x)dx = 1 \qquad \Rightarrow \int_0^2 Kx(2-x)dx = 1$$

$$K\int_0^2 (2x - x^2)dx = 1 \Rightarrow K\left[\frac{2x^2}{2} - \frac{x^3}{3}\right]_0^2 = 1$$

$$K\left[\left(4 - \frac{8}{3}\right) - (0-0)\right] = 1 \Rightarrow K\left[4 - \frac{8}{3}\right] = 1$$

$$K\left[\frac{12-8}{3}\right] = 1 \Rightarrow K\left(\frac{4}{3}\right) = 1 \quad \Rightarrow K = \frac{3}{4}$$

27. Find the MGF for the distribution where $f(x) = \begin{cases} \frac{2}{3} \quad at \quad x=1 \\ \frac{1}{3} \quad at \quad x=2 \\ 0,\ otherwise \end{cases}$

Solution:

Given : $f(1)=\frac{2}{3};\ f(2)=\frac{1}{3}; f(3)=f(4)=............=0$

MGF of a R.V 'X' is given by

$$M_X(t)=E\left[e^{tx}\right]$$
$$=\sum_{x=0}^{\infty}e^{tx}f(x)$$
$$=e^0f(0)+e^tf(1)+e^{2t}f(2)+...........$$
$$=e^t\left(\frac{2}{3}\right)+e^{2t}\left(\frac{1}{3}\right)+0...........$$
$$=\frac{2}{3}e^t+\frac{1}{3}e^{2t}$$

$$\therefore\ MGF\ is\ M_X(t)=\frac{e^t}{3}\left[2+e^t\right]$$

28. If a random variable 'X' has the MGF, $M_X(t)=\frac{2}{2-t}$ **find the variance of 'X'.**

Solution: **[AU M/J 2007]**

$$\text{Given } M_X(t)=\frac{2}{2-t}=2(2-t)^{-1}$$
$$M_X^{!}(t)=-2(2-t)^{-2}(-1)=2(2-t)^{-2}$$
$$M_X^{!!}(t)=-4(2-t)^{-3}(-1)$$
$$=4(2-t)^{-3}$$
$$E[X]=M_X^{!}(0)=2(2-0)^{-2}$$
$$=2\times\frac{1}{4}=\frac{1}{2}$$
$$E\left[X^2\right]=M_X^{!!}(0)=4(2-0)^{-3}$$
$$=\frac{4}{2^3}=\frac{4}{8}=\frac{1}{2}$$
$$Var(X)=E(X^2)-[E(X)]^2$$
$$=\frac{1}{2}-\frac{1}{4}$$
$$=\frac{1}{4}$$

29. Find the MGF of the distribution given by $f(x)=\begin{cases}\lambda e^{-\lambda x}, & x>0\\ 0, elsewhere\end{cases}$

Solution: **[AU N/D 2007]**

$$M_X(t)=\int_{-\infty}^{\infty}e^{tx}f(x)dx$$

$$= \int_0^\infty e^{tx} \lambda e^{-\lambda x} dx$$

$$= \lambda \int_0^\infty e^{-(\lambda - t)x} dx = \lambda \left[\frac{e^{-(\lambda-t)x}}{-(\lambda - t)} \right]_0^\infty$$

$$= -\frac{\lambda}{\lambda - t}\left[-e^{-\infty} - e^{-0}\right]$$

$$= -\frac{\lambda}{\lambda - t}[0 - 1]$$

$$\therefore The\ MGF\ is\ \ M_X(t) = \frac{\lambda}{\lambda - t}$$

30. State and prove additive property of binomial distribution.

Solution:

The sum of two binomial variates is not a binomial variate.

Let X and Y be two independent binomial variates with parameter (n_1, p_1) and (n_2, p_2) respectively.

Then $M_X(t) = \left(q_1 + p_1 e^t\right)^{n_1}$

$$M_Y(t) = \left(q_2 + p_2 e^t\right)^{n_2}$$

$$\therefore M_{X+Y}(t) = M_X(t) + M_Y(t) \quad [\because X\ and\ Y\ are\ independent\ R.V's]$$

$$= \left(q_1 + p_1 e^t\right)^{n_1}\left(q_2 + p_2 e^t\right)^{n_2}$$

The R.H.S cannot be expressed in the form $\left(q + pe^t\right)^n$. Hence by uniqueness theorem of MGF X+Y is not a binomial variate.

31. Check whether the following data follow a binomial distribution or not. Mean=3; Variance=4.

Solution: **[AU M/J 2004]**

Given that

$$Mean = np = 3 \,..................(1)$$

$$Variance = npq = 3 \,..................(2)$$

$$\frac{(2)}{(1)} \Rightarrow \frac{npq}{np} = \frac{4}{3} = 1\frac{1}{3}$$

$$q = 1\frac{1}{3}\ \ which\ is\ > 1$$

Since q>1 which is not possible (0<q<1). The given data does not follow binomial distribution.

32. If 'X' is a random variate following binomial distribution with mean 2.4 and Variance 1.44, find $P(X \geq 5)$. **[AU M/J 2003]**

Solution:

For a binomial distribution,

$$Mean = np = 2.4 \,..................(1)$$

$$Variance = npq = 1.44 \,..................(2)$$

$$\frac{(2)}{(1)} \Rightarrow \frac{npq}{np} = \frac{1.44}{2.4} = 0.6$$

Substituting

$$p = 0.4 \quad in \quad (1),$$

$$n(0.4) = 2.4 \Rightarrow n = \frac{2.4}{0.4} = 6$$

Therefore the distribution function is

$$P(X \geq 5) = 1 - P(X < 5)$$

$$= 1 - [P(X = 0) + P(X = 1) + P(X = 2) + P(X = 3) + P(X = 4)]$$

$$= 1 - \begin{Bmatrix} 6C_0(0.4)^0(0.6)^6 + 6C_1(0.4)^1(0.6)^5 + 6C_2(0.4)^2(0.6)^4 \\ + 6C_3(0.4)^3(0.6)^3 + 6C_4(0.4)^4(0.6)^2 \end{Bmatrix}$$

$$= 0.04096$$

33. With the usual notation find 'p' for a binomial random variate 'X' if n=6 and if $9P(X = 4) = P(X = 2)$ **[AU M/J 2004]**

Solution:

We know that

Given $9P(X = 4) = P(X = 2)$

$$9 \times 6C_4 p^4 q^2 = 6C_2 p^2 q^4$$

$$9p^2 = q^2$$

$$= (1 - p)^2$$

$$9p^2 = 1 + p^2 - 2p$$

$$8p^2 + 2p - 1 = 0$$

$$p = \frac{-2 \pm \sqrt{4 + 32}}{16} = \frac{-2 \pm 6}{16} = \frac{1}{4} \, or \, -\frac{1}{2}$$

$$p = -\frac{1}{2} \quad (not \;\; possible)$$

$$p = 0.25, q = 0.75$$

34. If the MGF of a r.v. X is of the form $(0.4e^t + 0.6)^8$ **. What is the MGF of 3X+2.**

Solution: **[AU N/D 2007]**

Given $M_X(t) = (0.4e^t + 0.6)^8 = E[e^{tX}]$

$$\therefore \;\; MGF \;\; of \;\; 3X + 2 \;\; is \;\; given \;\; by$$

$$M_{3X+2}(t) = E[e^{(3X+2)t}]$$

$$= e^{2t} E[e^{(3Xt)}]$$

$$= e^{2t} E[e^{X(3t)}] \qquad [\because E(e^{3Xt}) = M_X(t)]$$

$$= e^{2t}\left(0.4e^{3t} + 0.6\right)^8$$

35. If is a Poisson variate $P(X = 2) = 9P(X = 4) + 90P(X = 6)$, find (i) mean of X, (ii) variance of X.

Solution:

$$P(X = x) = \frac{e^{-\lambda}\lambda^x}{x!} \quad , \quad x = 0,1,2,.........$$

Given that $P(X = 2) = 9P(X = 4) + 90P(X = 6)$

$$\frac{e^{-\lambda}\lambda^2}{2!} = 9\frac{e^{-\lambda}\lambda^4}{4!} + 90\frac{e^{-\lambda}\lambda^6}{6!}$$

$$= e^{-\lambda}\lambda\left(\frac{9\lambda^2}{4!} + \frac{90\lambda^4}{6!}\right)$$

$$\frac{1}{2} = \frac{3\lambda^2}{8} + \frac{\lambda^4}{8}$$

$$\lambda^4 + 3\lambda^2 - 4 = 0$$

$$\lambda^2 = \frac{-3 \pm \sqrt{9+16}}{2} = \frac{-3 \pm 5}{2} = 1 \ or \ -4$$

$$\lambda^2 = 1 \ or \ \lambda^2 = -4$$

$$\lambda = \pm 1 \ or \ \lambda = \pm 2i$$

$$\therefore \ Mean \ = \lambda = 1 \ [\lambda - cannot \ be \ imaginary]$$

$$\therefore \ Variance = \lambda = 1$$

36. If 3% of the electric bulbs manufactured by a company are defective, find the probability that in a sample of 100 bulbs exactly 5 bulbs are defective $(e^{-3} = 0.0498)$.

Solution: **[AU N/D 2002]**

Let X be the R.V denoting the number of defective electric bulbs.

Given $P(a \ bulb \ is \ defective) = \dfrac{3}{100}$

$$p = 0.03$$

$$n = 100$$

$$\lambda = np = 100 \times 0.03 = 3$$

$$P('x' \ bulbs \ are \ defective) = P(X = x) = \frac{e^{-\lambda}\lambda^x}{x!}$$

$$P(exactly\ 5\ bulbs\ are\ defective) = P(X=5)$$
$$= \frac{e^{-3}3^5}{5!} = \frac{0.0498\times 243}{120}$$
$$= 0.1008$$

37. If the MGF of X is $(5-4e^t)^{-1}$, find the distribution of X and c.

Solution: **[AU N/D 2004]**

Let the geometric distribution be

$$P(X=x) = q^x p\ ,\ x = 0,1,2,\ldots\ldots\ldots$$

The MGF of geometric distribution is given by

$$M_X(t) = \frac{p}{1-qe^t}\ \ldots\ldots\ldots\ldots(1)$$

$$M_X(t) = (5-4e^t)^{-1} = 5^{-1}\left[1-\frac{4}{5}e^t\right]^{-1}\ \ldots\ldots\ldots\ldots(2)$$

Comparing (1) and (2), we get

$$q = \frac{4}{5} \Rightarrow p = \frac{1}{5} \qquad [\because p+q=1]$$

$$\therefore\ P(X=x) = pq^x\ ,\ x = 0,1,2,\ldots\ldots\ldots$$
$$= \left(\frac{1}{5}\right)\left(\frac{4}{5}\right)^x$$
$$P(X=5) = \left(\frac{1}{5}\right)\left(\frac{4}{5}\right)^5$$
$$P(X=5) = \frac{4^5}{5^6}$$

38. If the probability is 0.05 that a certain kind of measuring device will show excessive drift, what is the probability that the sixth of these measuring devices tested will be the first to show excessive drift?

Solution:

Let 'X' be the R.V denoting the no. of measuring devices to show excessive drift.

Here

$$p = 0.05 \Rightarrow q = 1-0.05 = 0.95$$
$$x = 6$$
$$We\ know\ that\ P(X=x) = q^{x-1}p$$
$$= (0.05)(0.95)^5$$
$$= 0.0387$$

39. Find the moment generating function of uniform distribution.

Solution:

$$M_X(t) = \int_{-\infty}^{\infty} e^{tx} f(x)dx$$

$$= \int_a^b e^{tx} f(x)dx$$

$$= \int_a^b e^{tx} \frac{1}{b-a} dx$$

$$= \frac{1}{b-a}\left[\frac{e^{tx}}{t}\right]_a^b$$

$$= \frac{1}{(b-a)t}\left[e^{bt} - e^{at}\right]$$

$\therefore$ *The MGF of uniform distribution is* $M_X(t) = \dfrac{\left[e^{bt} - e^{at}\right]}{(b-a)t}$

40. The time (in hours) required to repair a machine is exponentially distributed with parameter $\lambda = \frac{1}{2}$. What is the probability that a repair takes atleast 10 hrs given that its duration exceeds 9 hours.

Solution:

Let X be the R.V which represents the time to repair the machine.

The density function of X is given by

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

$$= \frac{1}{2} e^{-\frac{1}{2}x}, \quad x>0$$

$$\therefore P(X > 10 / X > 9) = P(X > 9 + 1 / X > 9)$$

$$= P(X > 1)$$

$$P(X > t) = e^{-\lambda t} = e^{-\frac{1}{2}t}$$

$$P(X > 1) = e^{-\frac{1}{2}} = 0.6065$$

41. The time in hours required a machine is exponentially distributed with parameter $\lambda = \frac{1}{2}$, what is the probability that the required time (i) exceeds 2 hours (ii) exceeds 5 hours.

Solution:

Let X be the R.V which represents the time to repair the machine. Then the density function of X is given by

$$f(x) = \lambda e^{-\lambda x} = \frac{1}{2} e^{-\frac{1}{2}x}, \quad x > 0$$

$$(i)\ \ P(X>k)=e^{-\lambda k} \Rightarrow P(X>2)=e^{-\frac{1}{2}\times 2}=e^{-1}$$

$$(ii)\ \ P(X>k)=e^{-\lambda k} \Rightarrow P(X>5)=e^{-\frac{1}{2}\times 5}=e^{-\frac{5}{2}}$$

42. A normal distribution has mean $\mu = 20\ \ and\ \ S.D\ \ \sigma = 10$**. Find** $P(15 \le X \le 40)$

Solution:

Given $\mu = 20\ \ and\ \ S.D\ \ \sigma = 10$

The normal variate $z=\dfrac{X-\mu}{\sigma}=\dfrac{X-20}{10}$

When $X=15,\ z=\dfrac{X-20}{10}=\dfrac{15-20}{10}=-0.5$

$X=40,\ z=\dfrac{X-20}{10}=\dfrac{40-20}{10}=2$

$$\begin{aligned}\therefore\ \ P(15\le X\le 40) &= P(-0.5\le z\le 2)\\ &= P(-0.5\le z\le 0)+P(0\le z\le 2)\\ &= P(0\le z\le 0.5)+P(0\le z\le 2)\\ &= 0.1915+0.4772\\ &= 0.6687\end{aligned}$$

43. If X is a R.V normally distributed with mean zero and variance σ^2**, Find** $E(|X|)$**.**

Solution:

We know that $E(|X|)$=Mean deviation about origin

= Mean deviation about mean (0)

Mean deviation about mean= $\dfrac{4}{5}\sigma$

[$\because$ Mean deviation about mean of the Normal distribution $=\dfrac{4}{5}\sigma$ **]**

$$E(|X|)=\frac{4}{5}\sigma$$

44. A random variable 'X' has a p.d.f $f(x)=3x^2\ ,\ 0<x<1$**, Find 'k' such that** $P(X>k)=0.5$

Solution: **[AU M/J 2005, 2014]**

Given that

$f(x)=3x^2\ ,\ 0<x<1$

when $P(X>k)=0.5$

$$\int_b^1 f(x)dx=0.5$$

$$\int_b^1 3x^2dx=0.5 \quad \Rightarrow 3\left(\frac{x^3}{3}\right)_b^1=0.5$$

$$1-k^3=\frac{1}{2}\Rightarrow k^3=1-\frac{1}{2}=0.5 \quad \Rightarrow k=(0.5)^{\frac{1}{3}}$$

45. If X and Y are i.i.d with variance 2 and 3 .Find the variance of 3X+4Y [A.U M/J 2014]

Solution:

$$Var(3X+4Y)=9Var(X)+16\ Var(X)=9(2)+16(3)=66$$

46. A Continuous random variable X has the random variable X has pdf is given by

$f(x)=\begin{cases} a(1+x^2) \\ 0 \end{cases}$ Find a and $P(X<4)$ **[A.U M/J 2014]**

Solution:

(i) Since $f(x)$ is the pdf, we have

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

$$ie., \int_0^2 a(1+x^2)dx=1$$

$$a\left[x+\frac{x^3}{3}\right]_2^5=1$$

$$a\left\{\left[5+\frac{125}{3}\right]-\left[2+\frac{8}{3}\right]\right\}=1$$

$$a\left[3+\frac{117}{3}\right]=1$$

$$a=\frac{1}{42}$$

(ii) To find $P(X<4)$

$$P(X\le 4)=\frac{1}{42}\int_2^4 f(x)\,dx$$

$$=\frac{1}{42}\left[x+\frac{x^3}{2}\right]_2^4$$

$$=\frac{1}{42}\left[4+\frac{64}{3}\right]-\left[2+\frac{8}{3}\right]$$

$$=\frac{31}{63}$$

47. Test whether $f(x)=\begin{cases} |x| & -1\le x\le 1 \\ 0 \end{cases}$, can be the pdf of a continuous random variable

Solution: **[A.U N/D '14 A/M '15]**

Since $f(x)$ is the pdf, we have

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

$$\int_{-1}^{0}(-x)dx+\int_0^1 xdx=1$$

$$-\left(\frac{x^2}{2}\right)_{-1}^{0}+\left(\frac{x^2}{2}\right)_0^1=\frac{1}{2}+\frac{1}{2}=1$$

48. For a binomial distribution with mean 6 and npq=2,find the first two terms of the distribution.

Solution:

$$Mean=np=6\ldots\ldots\ldots\ldots\ldots\ldots(1)$$

[A.U M/J 2014]

$$Variance = npq = 2..................(2) \qquad \frac{(2)}{(1)} \Rightarrow \frac{npq}{np} = \frac{2}{3}$$

49. What is meant by MGF? **[AU N/D 2014]**

Solution:

The moment generating function of a random variable whose pdf discrete probability distributionis given by

$$M_X(t) = \int_{-\infty}^{\infty} e^{tx} f(x)dx \qquad E(X) = \sum_{x=1}^{\infty} xp(x)$$

Where t is real parameter and the integration or summation being extended to the entire range of x.

50. What are the limitations of Poisson distribution? **[AU A/M '15]**

Solution:

(i) The number of trails $'n'$ should be indefinitely large.

$i.e., \quad n \to \infty$

(ii) The probability of successes $'p'$ for each trail is indefinitely small.

(iii) $np = \lambda$, should be finite where λ is a constant.

51. Let X be a discrete R.V. with probability mass function $P(X = x) = \begin{cases} \frac{x}{10}, & x = 1,2,3,4 \\ 0, & otherwise \end{cases}$

Compute $P(X < 3)$ and $E\left(\frac{1}{2}X\right)$

Solution: **[AU M/J 2016]**

Given that

$x:$	1	2	3	4
$p(x):$	$\frac{1}{10}$	$\frac{2}{10}$	$\frac{3}{10}$	$\frac{4}{10}$

$$P(X < 3) = P(X = 1) + P(X = 2) = \frac{1}{10} + \frac{2}{10} = \frac{3}{10}$$

$$E\left(\frac{1}{2}X\right) = \frac{1}{2}\sum_{x=1}^{4} xP(X = x) = \frac{1}{2}\left[\frac{1+4+9+16}{10}\right] = \frac{3}{2}$$

52. If a random variable has the moment generating function given by $M_X(t) = \frac{3}{3-t}$**,**

Compute $E(X^2)$ **[AU M/J 2016]**

Solution:

Given $\quad M_X(t) = \frac{3}{3-t} = 3(3-t)^{-1}$

$$M'_X(t) = 3(-1)(3-t)^{-2}(-1) = 3(3-t)^{-2}$$
$$M''_X(t) = 3(-2)(3-t)^{-3}(-1) = 6(3-t)^{-3}$$
$$E(X) = M'_X(0) = 3(3-0)^{-2}$$
$$= 3\times 3^{-2}$$
$$= \frac{3}{9} = \frac{1}{3}$$
$$E(X^2) = M''_X(0) = 6(3-0)^{-3}$$
$$= 6\times 3^{-3} = \frac{6}{3^3} = \frac{6}{27} = \frac{2}{9}$$

53. A continuous random variable X has the probability density function given by

$f(x) = \begin{cases} \lambda(1+x^2), & 1\le x\le 5 \\ 0, & otherwise \end{cases}$ **Find** λ *and* $P(X<4)$**.** **[AU N/D 2015]**

Solution:

We know that, $\int_{-\infty}^{\infty} f(x)dx = 1$

$$\int_1^5 \lambda(1+x^2)dx = 1$$
$$\lambda\left[x+\frac{x^3}{3}\right] = 1$$
$$\lambda\left[\left(5+\frac{5^3}{3}\right)-\left(1+\frac{1}{3}\right)\right] = 1$$
$$\lambda\left[\left(5+\frac{125}{3}\right)-\left(\frac{4}{3}\right)\right] = 1$$
$$\lambda\left(\frac{136}{3}\right) = 1$$
$$\lambda = \frac{3}{136}$$

$\therefore\ P(X<4) = P(1<X<4)$

$$= \int_1^4 f(x)dx = \int_1^4 \lambda(1+x^2)dx$$
$$= \int_1^4 \frac{3}{136}(1+x^2)dx$$
$$= \frac{3}{136}\left[x+\frac{x^3}{3}\right]_1^4$$
$$= \frac{3}{136}\left[4+\frac{64}{3}-\left(1+\frac{1}{3}\right)\right]$$
$$= \frac{9}{17}$$

54. What is meant by memory less property? Which discrete distribution follows this property?

Solution: **[AU N/D 2015]**

A non-negative random variable X is memory less if $P[X>m+n/X>m] = P[X>n]$ for all m *and* $n\ge 0$.

Geometric distribution (discrete distribution) follows this property.

55. If a fair coin is tossed twice, find $P(X\le 1)$ **, where X denotes the number of heads in each experiment.** **[AU N/D 2016]**

Solution:

$$P(X \le 1) = P\{TT\} + P\{TH\} + P\{HT\}$$
$$= P\{X = 0\} + P\{X = 1\}$$
$$= \frac{1}{4} + \frac{1}{2} = \frac{3}{4}$$

56. Balls are tossed at random into 50 boxes. Find the expected number of tosses required to get the first ball in the fourth box. [AU N/D 2016]

Solution:

Let X denote the number of tosses until the first ball goes into the fourth box. Then X has a geometric distribution with $p = \frac{1}{2}$

$$\therefore \quad E(X) = \frac{1}{p} = 50$$

57. A test engineer discovered that the cumulative distribution function of the lifetime of an equipment (in years) is given by $F_X(x) = 1 - e^{-\frac{x}{5}}, x \ge 0$. What is the expected lifetime of the equipment? [AU N/D 2017]

Solution:

Given $F_X(x) = 1 - e^{-\frac{x}{5}}, x \ge 0$

$$f(x) = F'(x) = \frac{1}{5}e^{-\frac{x}{5}}, x \ge 0$$

$$E(x) = \int_0^\infty x\frac{1}{5}e^{-\frac{x}{5}}dx = \frac{1}{5}\int_0^\infty xe^{-\frac{x}{5}}dx = \frac{1}{5}\left\{x\left(\frac{e^{-\frac{x}{5}}}{-\frac{1}{5}}\right) - \left(\frac{e^{-\frac{x}{5}}}{\frac{1}{25}}\right)\right\}_0^\infty = \frac{1}{5}(25) = 5$$

58. If X is a normal random variable with mean 3 and variance 9, find the probability that X is between 2 and 5. [AU N/D 2017]

Solution:

Given $\mu = 3, \sigma^2 = 9 \Rightarrow \sigma = 3$

$$P(2 \le X \le 5) = P\left(\frac{2-3}{3} \le \frac{X-3}{3} \le \frac{5-3}{3}\right) = P\left(\frac{-1}{3} \le Z \le \frac{2}{3}\right)$$
$$= P(-0.333 \le Z \le 0.666) = P(0 \le Z \le 0.333) + P(0 \le Z \le 0.666)$$
$$= 0.1293 + 0.2454 = 0.3747$$

59. The mean of binomial distribution is 20 and standard deviation is 4. Find the parameters of this distribution. [AU M/J 2018]

Solution:

Given that $np = 20 ......(1)\ and\ \sqrt{npq} = 4$

$$npq = 16\text{.........}(2)$$
$$\frac{(2)}{(1)} \Rightarrow q = \frac{4}{5}, \quad p = 1 - q = 1 - \frac{4}{5} = \frac{1}{5}$$
$$(1) \Rightarrow n\left(\frac{1}{5}\right) = 20 \Rightarrow n = 100$$

$\therefore$ *The parameters are* $n = 100$ *and* $p = \frac{1}{5}$

60. Check whether the function given by $f(x) = \frac{x+2}{25}$ $for\ x = 1,2,3,4,5$ can serve as the probability distribution of a discrete random variable. **[AU M/J 2018]**

Solution:

$$f(1) = \frac{3}{25}; f(2) = \frac{4}{25}; f(3) = \frac{5}{25}; f(4) = \frac{6}{25} \ and\ f(5) = \frac{7}{25}$$
$$f(1) + f(2) + f(3) + f(4) + f(5) = 1$$

Therefore given $f(x)$ *is a probability distribution.*

61. State Baye's theorem.

Solution:

Let $A_1, A_2, \ldots, A_n$ be $'n'$ mutually exclusive and exhaustive events with $P(A_i) \neq 0$ for $i = 1,2,\ldots\ldots,n$. Let B be an event such that $B \subset \bigcup_{i=1}^{n} A_i$, $P(B) \neq 0$. Then

$$P(A_i / B) = \frac{P(A_i)P(B/A_i)}{\sum_{i=1}^{n} P(A_i)P(B/A_i)}$$

62. Let A and B be two events such that $P(A) = \frac{1}{2}$, $P(B) = \frac{1}{3}$ and $P(A \cap B) = \frac{1}{6}$ Compute $P(B/A)$ *and* $P(\overline{A} \cap B)$. **[AU A/M 2021]**

Solution:

Given that $P(A) = \frac{1}{2}$, $P(B) = \frac{1}{3}$ and $P(A \cap B) = \frac{1}{6}$

$$P(B/A) = \frac{P(A \cap B)}{P(A)} = \frac{1/6}{1/2} = \frac{1}{3}$$
$$P(\overline{A} \cap B) = P(B) - P(A \cap B) = \frac{1}{3} - \frac{1}{6} = \frac{1}{6}$$

63. During off-peak hours a customer train has five cars. Suppose a customer is twice as likely o select the middle car (#3) as to select either adjacent car (#2 or #4), and is twice as likely to select either adjacent car as to select either end car (#1 or #5). Find the probability that one of the three middle cars is selected.

Solution: **[AU A/M 2024]**

Let P_i be probability that i^{th} car selected

We've $P_3 = 2P_2 = 2P_4$ *and* $P_3 = 2P_1 = 2P_5 = P_4$

$$Total\ probability = 1 \Rightarrow 10P_1 = 1 \Rightarrow P_1 = \frac{1}{10}$$

$$\Pr obability\ that\ one\ of\ \ he\ three\ middle\ cars\ selected = P_2 + P_3 + P_4 = 0.8$$

64. The probability that a certain kind of component will survive a shock test is 3/4. Find the probability that exactly 2 of the next 4 components tested survive.

Solution: **[AU A/M 2024]**

$$p = \frac{3}{4}$$

$$P(X = 2) = 4C_2\left(\frac{3}{4}\right)^2\left(\frac{1}{4}\right)^2 = \frac{27}{128}$$

65. What is the probability of getting an even number in a single throw with a dice?

Solution: **[AU N/D 2024]**

The probability of even number $= \frac{3}{6} = \frac{1}{2}$

66. Suppose a book of 585 pages contains 43 typographical errors. If these errors are randomly distributed throughout the book, Using Poisson distribution, find the probability that 10 pages, selected as random, will be free from errors?

Solution: **[AU N/D 2024]**

Given $n = 10,\ p = \frac{43}{585} = 0.0735, \quad mean = np = o.735$

$$p(0) = \Pr obability\ of\ \ zero\ error$$

$$= \frac{e^{-0.735}(0.735)^0}{0!} = 0.4795$$

PART-B
ONE DIMENSIONAL DISCRETE RANDOM VARIABLE

1. A random variable X has the following function:

X:	0	1	2	3	4	5	6	7
P(X) :	0	K	2K	2K	3K	K^2	$2K^2$	$7K^2+K$

Find (a) K,

(b) Evaluate P[X< 6], $P[X \geq 6]$

(c) If $P[X \leq C] > \frac{1}{2}$ **find the minimum value of C.**

(d) Evaluate $P[1.5 < X < 4.5 | X > 2]$

(e) Find $P[X < 2], P[X > 3], P[1 < X < 5]$. **[M/J 2012, M/J 2014]**

Solution:

(a) To Find K: We know that $\sum_{i=1}^{\infty} P(x_i) = 1.$

$$0 + K + 2K + 2K + 3K + K^2 + 7K^2 + K = 1$$

$$10K^2 + 9K - 1 = 0$$

$$K = -1, or\ K = \frac{1}{10}$$

Since $P(X) \geq 0$ the value K=-1 is not permissible

Result: Hence we have $K = \frac{1}{10}$.

(b) To Evaluate P[X< 6], $P[X \geq 6]$ **:**

$$P[X \geq 6] = P[X = 6] + P[X = 7] = 2K^2 + 7K^2 + K$$

$$= \frac{2}{100} + \left(\frac{7}{10} + \frac{1}{10}\right) = \frac{19}{100}.$$

Result: $P[X < 6] = 1 - P[X \geq 6] = 1 - \frac{19}{100} = \frac{81}{100}$

© New table:

X	0	1	2	3	4	5	6	7
P(x)	0	$\frac{1}{10}$	$\frac{2}{10}$	$\frac{2}{10}$	$\frac{3}{10}$	$\frac{1}{100}$	$\frac{2}{100}$	$\frac{17}{100}$
$P[X \leq x]$	0	$\frac{1}{10}$	$\frac{3}{10}$	$\frac{5}{10}$=0.5	$\frac{8}{10} =$ $0.8 > 0.5$	$\frac{81}{100}$	$\frac{83}{100}$	$\frac{100}{100} = 1$

Result: The minimum value of c=4.

(d) To Find $P[1.5 < X < 4.5 | X > 2]$

$$P[1.5 < X < 4.5 | X > 2] = \frac{P[(1.5 < X < 4.5) \cap x > 2]}{P(X > 2)} \quad \because P\left(\frac{W_1}{W_2}\right) = \frac{P(W_1 \cap W_2)}{P(W_2)}$$

$$= \frac{P[2 < X < 4.5]}{P(X > 2)}$$

$$= \frac{P(3) + P(4)}{1 - [P(0) + P(1) + P(2)]}$$

$$= \frac{\frac{2}{10} + \frac{3}{10}}{1 - [0 + \frac{1}{10} + \frac{2}{10}]}$$

Result:

$$= \frac{\frac{5}{10}}{1 - \frac{3}{7}} = \frac{\frac{5}{10}}{\frac{7}{10}} = \frac{5}{7}.$$

$$(e)\ (i)\ p(X < 2) = P(X = 0) + P(X = 1) = 0 + \frac{1}{10} = \frac{1}{10}$$

$$(ii)\ p(X > 3) = 1 - P(X \le 3) = 1 - [P(X = 0) + P(X = 1) + P(X = 2) + P(X = 3)]$$

$$= 1 - \left(\frac{1}{10} + \frac{2}{10} + \frac{2}{10}\right) = 1 - \frac{5}{10} = 1 - \frac{1}{2} = \frac{1}{2}$$

$$p(1 < X < 5) = P(X = 2) + P(X = 3) + P(X = 4) = \frac{2}{10} + \frac{2}{10} + \frac{3}{10} = \frac{7}{10}$$

2. A random variable X has the following probability distribution .

X	-2	-1	0	1	2	3
P(x)	0.1	k	0.2	2k	0.3	k

Find (1) the value of k. (2) Evaluate P(X<1) and P(-1<X<2) (3) find the cumulative distribution of x. (4) evaluate mean of x. [AU M/J 2016, N/D 2024]

Solution:

We know that

$$\sum_{i=1}^{\infty} P(x_i) = 1.$$

0.1+k+0.2+2k+0.3+k=1

4k+0.6=1

6k=1-0.6

Result:

k=0.4/4=0.1

New table

x	-2	-1	0	1	2	3
P(X=x)	$\frac{1}{10}$	$\frac{1}{10}$	$\frac{2}{10}$	$\frac{2}{10}$	$\frac{3}{10}$	$\frac{1}{10}$

(2) (i) P[X<1]=P[X=-2]+P[X=-1]+P[X=0]

$$= \frac{1}{10} + \frac{1}{10} + \frac{2}{10} = \frac{4}{10}$$

$$P[-1 < X \le 2] = P(X=0) + P(X=1) + P(X=2)$$

$$= \frac{2}{10} + \frac{2}{10} + \frac{3}{10} = \frac{2+2+3}{10} = \frac{7}{10}$$

X	-2	-1	0	1	2	3
P(X=x)	$\frac{1}{10}$	$\frac{1}{10}$	$\frac{2}{10}$	$\frac{2}{10}$	$\frac{3}{10}$	$\frac{1}{10}$
xp(X=x)	$\frac{-2}{10}$	$\frac{-1}{10}$	0	$\frac{2}{10}$	$\frac{6}{10}$	$\frac{3}{10}$

Result:

$$Mean = E(X) = \sum x\ p(x) = \frac{-2}{10} + \frac{-1}{10} + \frac{2}{10} + \frac{6}{10} + \frac{3}{10} = \frac{8}{10}.$$

3. The probability function function of random variable X is defined as

$P(X=0) = 3C^2$, $P(X=1) = 4C - 10C^2$, $P(X=2) = 5C - 1$, *where* $C > 0$, *and* $P(X=r) = 0$ *if* $r \ne 0,1,2$.

find (i) the value of C (iii) The distribution function of X. (iii) The distribution function of X.

(iv) The largest value of X for which F(x) < 1/2.

Solution:

$$\sum_{x=0}^{r} p(x) = 1$$

$$P(X=0) + P(X=1) + P(X=2) = 1$$

$$3C^2 + 4C - 10C^2 + 5C - 1 = 1$$

$$-7C^2 + 9C - 1 - 1 = 0$$

$$7C^2 - 9C + 2 = 0$$

$$(7C-2)\ (C-1) = 0$$

Result: $C = \frac{2}{7}$

$$(ii) \quad P(0 < X < 2 \mid X > 0) = \frac{P(0 < X < 2 \mid X > 0)}{P(X > 0)}$$

$$= \frac{4C - 10C^2}{9C - 10C^2 - 1}$$

$$=\frac{4\left(\frac{2}{7}\right)-10\left(\frac{4}{49}\right)}{9\left(\frac{2}{7}\right)-10\left(\frac{4}{49}\right)-1}$$

Result: $= 0.4323$

(iii) New table:

X	$F(x)=P(X\le x)$
0	$F(0)=P(X=0)=3C^2=3\left(\frac{4}{49}\right)=0.2449<0.5$
1	$F(1)=P(X=0)+p(X=1)=-7C^2+4C=\frac{4}{7}>0.5$
2	$F(1)=P(X=0)+p(X=1)+P(X=2)=1$

(iv) From new table

$$F(x)<\frac{1}{2} \quad for\ x=0.$$

4. The probability function of an infinite discrete distribution is given by P (X = j) = $\frac{1}{2^j}$; j=1,2,…∞. Verify that the total probability is 1 and find the mean and variance of the distribution. Find also P(X is even), P(X≥5) and P(X is divisible by 3).

Solution:

$$E(x)=\sum_{j=1}^{\infty} x_j p(x_j)$$

$$=(1)\left(\frac{1}{2}\right)+(2)\left(\frac{1}{2}\right)^2+(3)\left(\frac{1}{2}\right)^3+\ldots$$

$$=\frac{1}{2}\left[1+(2)\left(\frac{1}{2}\right)+(3)\left(\frac{1}{2}\right)^2+\ldots\right]$$

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

$$E[X^2]=\sum_{j=1}^{\infty} x^2{}_j p(x_j)=\sum_{j=1}^{\infty}(x_j)(x_j+1)p(x_j)-\sum_{j=1}^{\infty} x_j p(x_j)$$

$$=\left[(1)(2)\frac{1}{2}+(2)(3)\left(\frac{1}{2}\right)^2+(3)(4)\left(\frac{1}{2}\right)^3+\ldots\right]-2$$

$$=\frac{1}{2}\left[(1)(2)+(2)(3)\left(\frac{1}{2}\right)^1+(3)(4)\left(\frac{1}{2}\right)^2+\ldots\right]-2$$

$$= \frac{1}{2}\left[2\left(1-\frac{1}{2}\right)^{-3}\right] = 6$$

$$Var(X) = E(X^2)-[E(X)]^2$$
$$= 6-4 = 2$$

P[X is even] = P [X= 2]+ P [X= 4]+…

$$= \left(\frac{1}{2}\right)^2 + \left(\frac{1}{2}\right)^4 + \left(\frac{1}{2}\right)^6 + \dots$$

$$= \left(1-\frac{1}{4}\right)^{-1} - 1 = \frac{1}{3}$$

P[X ≥ 5] = P [X= 5]+ P [X= 6]+…

$$= \left(\frac{1}{2}\right)^5 + \left(\frac{1}{2}\right)^6 + \dots$$

$$= \left(\frac{1}{2}\right)^5\left[1+\left(\frac{1}{2}\right)^2 + \dots\right]$$

$$= \left(\frac{1}{2}\right)^5\left[1-\frac{1}{2}\right]^{-1} = \frac{1}{16}$$

P[X is divisible by 3] or P[X is multiple of 3]

= P [X= 3]+ P [X= 6]+…

$$= \left(\frac{1}{3}\right)^3 + \left(\frac{1}{3}\right)^6 + \dots$$

$$= \frac{1}{8} + \left(\frac{1}{8}\right)^2 + \left(\frac{1}{8}\right)^3 + \dots$$

$$= 1 + \frac{1}{8} + \left(\frac{1}{8}\right)^2 + \left(\frac{1}{8}\right)^3 + \dots\dots\dots - 1$$

$$= \left(1-\frac{1}{8}\right)^{-1} - 1 = \frac{1}{7}$$

5. If the random variable X takes the values 1,2,3 and 4 such that 2P(X=1) = 3P(X=2) = P(X=3) = 5P(X=4) find the probability distribution and cumulative distribution function of X. [AU N/D 2012]

Solution:

X is a discrete random variable.

Let 2P(X=1) = 3P(X=2) = P(X=3) = 5P(X=4)= k

$2P(X=1) = k, \Rightarrow P(X=1) = \frac{k}{2}$

3P(X=2) = k, $\Rightarrow$ P(X=2) = $\frac{k}{3}$

P(X=3) = k, $\Rightarrow$ P(X=3) = k

5P(X=4)= k , $\Rightarrow$ P(X=4) = $\frac{k}{5}$

We know that $\sum_{i=1}^{n} p(x_i) = 1$

$$\frac{k}{2}+\frac{k}{3}+k+\frac{k}{5}=1$$

Result: $\frac{61k}{30}=1$

$$k=\frac{30}{61}$$

x_i	p(x_i)	F(X)
1	p(1)= $\frac{k}{2}=\frac{15}{61}$	F(1)= p(1)= $\frac{15}{61}$
2	p(2)= $\frac{k}{3}=\frac{10}{61}$	F(2)=F(1)+p(2)= $\frac{15}{61}+\frac{10}{61}=\frac{25}{61}$
3	p(3) = k = $\frac{30}{61}$	F(3)=F(2)+ p(3)= $\frac{25}{61}+\frac{30}{61}=\frac{55}{61}$
4	p(4) = $\frac{k}{5}=\frac{6}{61}$	F(4)=F(3)+p(4)= $\frac{55}{61}+\frac{6}{61}=$ $\frac{61}{61}=1$

6. A random variable X takes the values -2,-1,0 and 1 with probabilities $\frac{1}{8},\frac{1}{8},\frac{1}{4}$ *and* $\frac{1}{2}$ respectively. Find and draw the probability distribution function. [AU N/D 2014]

Solution:

Distribution function F(x) of X.

x_i	p(x_i)	F(X)
-2	p(1) = $\frac{1}{8}$	F(1)= p(1)= $\frac{1}{8}$
1	p(2) = $\frac{1}{8}$	F(2)=F(1)+p(2)= $\frac{1}{8}+\frac{1}{8}=\frac{2}{8}$
0	p(3) = $\frac{1}{4}$	F(3)=F(2)+ p(3)= $\frac{2}{8}+\frac{1}{4}=\frac{4}{8}$
1	p(4) = $\frac{1}{2}$	F(4)=F(3)+p(4)= $\frac{4}{8}+\frac{1}{2}=\frac{8}{8}=1$

7. If X has the distribution function $F[x]=\begin{cases}0 & x<1\\ \frac{1}{3} & for\ \ 1\le x<4.\\ \frac{1}{2} & for\ \ 4\le x<6\\ \frac{5}{6} & for\ 6\le x<10\\ 1 & x\ge 10.\end{cases}$

Find (1) the probability distribution of X (2) p(2<x<6) (3) mean of X 4) Variance of X.

Solution:

(1) To find the Probability distribution of X (ie) P(x)

X	0	1	4	6	10
F(x)	0	$\frac{1}{3}$	$\frac{1}{2}$	$\frac{5}{6}$	1
P(X=x)	0	$\frac{1}{3}$	$\frac{1}{2}-\frac{1}{3}=\frac{1}{6}$	$\frac{5}{6}-\frac{1}{2}=\frac{2}{6}$	$\frac{1}{6}$
$x\ p(x)$	0	$\frac{1}{3}$	$\frac{4}{6}$	2	$\frac{10}{6}$
$x^2 p(x)$	0	$\frac{1}{3}$	$\frac{16}{6}$	12	$\frac{100}{6}$

(2) From above table probability distribution p(x) is

X	0	1	4	6	10
P(x)	0	$\frac{1}{3}$	$\frac{1}{6}$	$\frac{2}{6}$	$\frac{1}{6}$

(3) To find Mean of X

$$E[X]=\sum x\ p(x) = 0+\frac{1}{3}+\frac{4}{6}+2+\frac{10}{6}=\frac{28}{6}=\frac{14}{3}$$

(4) To find Variance of X.

$$E[X^2]=\sum x^2\ \ p(x) = 0+\frac{1}{3}+\frac{16}{6}+12+\frac{100}{6}=\frac{190}{6}$$

$$Var[X]=E[X^2]-[E[X]]^2$$

$$=\left(\frac{190}{6}\right)-\left(\frac{14}{3}\right)^2$$

$$=\frac{190}{6}-\frac{784}{36}$$

$$=\frac{356}{36}=\frac{89}{9.}.$$

P(2<x<6)=P(x=4)

$$P(X=4)=\frac{1}{6}$$

8. Consider a discrete r. v. 'X' with probability function

$$p(X=x)=\begin{cases}\dfrac{1}{x(x+1)}, & x=1,2,3\ldots\\ 0 & othewise\end{cases}$$

Show that E(X) does not exist even though MGF exist.

Solution:

$$Given\ p(x)=\frac{1}{x(x+1)}, \qquad x=1,2,3,\ldots$$

$$E(X)=\sum_{x=1}^{\infty}xp(x)=\sum_{x=1}^{\infty}x\frac{1}{x(x+1)}$$

$$=\sum_{x=1}^{\infty}\frac{1}{x+1}=\frac{1}{2}+\frac{1}{3}+\frac{1}{4}+\ldots$$

$$=-1+1+\frac{1}{2}+\frac{1}{3}+\frac{1}{4}+\ldots$$

$$=-1+\sum_{x=1}^{\infty}\frac{1}{x}$$

But $\sum_{x=1}^{\infty}\frac{1}{x}$ is a divergent series.

$\therefore$ E(X) does not exist and hence no moment exists.

Now, MGF of X is given by

$$M_X(t)=\sum_{x=1}^{\infty}e^{tx}\,p(x)$$

$$=\sum_{x=1}^{\infty}\frac{e^{tx}}{x(x+1)} \qquad \text{put } z=e^t$$

$$=\sum_{x=1}^{\infty}\frac{z^x}{x(x+1)}$$

$$=\frac{z}{1\cdot 2}+\frac{z^2}{2\cdot 3}+\frac{z^3}{3\cdot 4}+\ldots$$

$$=z\left(1-\frac{1}{2}\right)+z^2\left(\frac{1}{2}-\frac{1}{3}\right)+z^3\left(\frac{1}{3}+\frac{1}{4}\right)+\ldots$$

$$=\left[z+\frac{z^2}{2}+\frac{z^3}{3}+\ldots\right]-\frac{z}{2}-\frac{z^2}{3}-\frac{z^3}{4}-\ldots$$

$$=-\log(1-z)-\frac{1}{z}\left[\frac{z^2}{2}+\frac{z^3}{3}+\frac{z^4}{4}+\ldots\right]$$

$$\left[Formula: \quad x+\frac{x^2}{2}+\frac{x^3}{3}...=-\log(1-x) \right]$$

$$=-\log(1-z)-\frac{1}{z}\left[-z+z+\frac{z^2}{2}+\frac{z^3}{3}+...\right]$$

$$=-\log(1-z)+1-\frac{1}{z}\left[z+\frac{z^2}{2}+\frac{z^3}{3}+...\right]$$

$$=-\log(1-z)+1-\frac{1}{z}\left[-\log(1-z)\right]$$

$$=-\log(1-z)+1+\frac{1}{z}\log(1-z), |z|<1$$

$$=1+\left(\frac{1}{z}-1\right)\log(1-z), \ |z|<1$$

$$\therefore M_X(t)=1+(e^{-t}-1)\ \log(1-e^t)\ , t<0 \qquad \left[\because z=e^t\right]$$

$$M_X(t)=1, for\ t=0 \text{ does not exist for t > 0}$$

. **COTINUOUS RANDOM VARIABLE**

9. If the density function of a continuous random variable X is given by

$$f(x)=\begin{cases} ax & ,0\le x\le 1 \\ a & ,1\le x\le 2 \\ 3a-ax & ,2\le x\le 3 \\ 0 & ,otherwise \end{cases}$$

(i) Find the value of a

(ii) The cumulative distribution function of X

(iii) If x_1,x_2 and x_3 are 3 independent observations of X. What is the probability that exactly one of these 3 is greater than 1.5? [A.U. N/D 2007,'08 A/M '17]

Solution:

(i)Since f(x) is a p.d.f, then

$$\int_{-\infty}^{\infty} f(x)dx=1 \quad i.e., \int_0^3 f(x)dx=1$$

$$\int_0^1 axdx+\int_1^2 adx+\int_2^3 (3a-ax)dx=1$$

$$a\left[\frac{x^2}{2}\right]_0^1+a[x]_1^2+\left[3ax-\frac{ax^2}{2}\right]_2^3=1$$

$$a\left[\frac{1}{2}-0\right]+a[2-1]+\left[9a-\frac{9a}{2}\right]-[6a-2a]=1$$

$$\frac{a}{2}+a+\frac{9a}{2}-4a=1$$

Result: $$6a-4a=1\Rightarrow 2a=1\Rightarrow a=\frac{1}{2}$$

(ii) (a)If x<0 then F(x)=0

(b) $If\ \ 0\le x\le 1\ then\ F[x]=\int_{-\infty}^{x}f(x)dx=\int_{0}^{x}axdx=\int_{0}^{x}\frac{x}{2}dx$

$$=\frac{1}{2}\left[\frac{x^2}{2}\right]_0^x=\frac{1}{4}\left[x^2\right]_0^x=\frac{x^2}{4}$$

(c) $If\ \ 1\le x\le 2\ then\ F[x]=\int_{0}^{1}axdx+\int_{1}^{x}adx=a\left[\frac{x^2}{2}\right]_0^1+a[x]_1^x$

$$=\frac{1}{2}\left[\frac{x^2}{2}\right]_0^1+\frac{1}{2}[x]_1^x=\frac{x}{2}-\frac{1}{4}$$

$(d)If\ \ 2\le x\le 3\ then\ F[x]=\int_{0}^{1}axdx+\int_{1}^{2}adx+\int_{2}^{x}(3a-ax)dx$

$$=\int_{0}^{1}\frac{x}{2}dx+\int_{1}^{2}\frac{1}{2}dx+\int_{2}^{x}\left(\frac{3}{2}-\frac{x}{2}\right)dx$$

$$=\frac{1}{2}\left[\frac{x^2}{2}\right]_0^1+\frac{1}{2}[x]_1^2+\left[\frac{3}{2}x-\frac{1}{2}\frac{x^2}{2}\right]_2^x=\frac{3}{2}x-\frac{x^2}{4}-\frac{5}{4}$$

$(e)If\ \ x>3\ then\ F[x]=\int_{0}^{1}axdx+\int_{1}^{2}adx+\int_{2}^{3}(3a-ax)dx+\int_{3}^{x}f(x)dx$

$$=\int_{0}^{1}\frac{x}{2}dx+\int_{1}^{2}\frac{1}{2}dx+\int_{2}^{3}\left(\frac{3}{2}-\frac{x}{2}\right)dx+0$$

$$=\frac{1}{2}\left[\frac{x^2}{2}\right]_0^1+\frac{1}{2}[x]_1^2+\left[\frac{3}{2}x-\frac{1}{2}\frac{x^2}{2}\right]_2^3=1$$

(iii) $P(x>1.5)=\int_{1.5}^{3}f(x)dx+\int_{1.5}^{2}\frac{1}{2}dx+\int_{2}^{3}\left(\frac{3}{2}-\frac{x}{2}\right)dx=\frac{1}{2}$

Choosing an X and observing its value can be considered as a trial and X>1.5 can be considered as a success.

i.e., p=P[X>1.5]=1/2

$$\therefore\ p=\frac{1}{2},q=\frac{1}{2}$$

As we choose 3 independent observations of X ,n=3.

By Bernoulli's theorem $P(x)=nC_xp^xq^{n-x}$]

P(exactly one value>1.5)=P(1 success)

$$=3C_1\left(\frac{1}{2}\right)^1\left(\frac{1}{2}\right)^{3-1}=\frac{3}{8}$$

10. The Distribution F of a continuous random variable X is given by

$$F(x)=\begin{cases}0 & ,x<0\\ x^2 & ,0\le x<\frac{1}{2}\\ 1-\frac{3}{25}(3-x)^2 & ,\frac{1}{2}\le x<3\\ 1 & ,x\ge 3\end{cases}$$

Find the pdf of X and evaluate P(|X|≤1) and P(1/3<X<4) usin g both PDF and CDF.

Solution: **[AU N/D 2011]**

$$f(x)=\frac{d}{dx}F(x)=\begin{cases}2x & ,0\le x<\frac{1}{2}\\ \frac{6}{25}(3-x), & \frac{1}{2}\le x<3\\ 0 & ,otherwise\end{cases}$$

P(|X|≤1)=F[1]-F[-1]= $\left[1-\frac{3}{25}(3-1)^2\right]$-0= $\frac{13}{25}$

$$\int_{-1}^{1}f(x)\,dx=0+\int_0^{\frac{1}{2}}2x\,dx+\int_{\frac{1}{2}}^{1}\frac{6}{25}(3-x)\;dx=\left[x^2\right]_0^{\frac{1}{2}}+\frac{6}{25}\left[3x-\frac{x^2}{2}\right]_{\frac{1}{2}}^{1}=\frac{1}{4}+\frac{6}{25}\left[\left(3-\frac{1}{2}\right)-\left(\frac{3}{2}-\frac{1}{8}\right)\right]=\frac{13}{25}$$

P(1/3<X<4)= F[4]-F[1/3]= $\left[1-\left(\frac{1}{3}\right)^2\right]=\frac{8}{9}$

(or)

$$P\left(\frac{1}{3}<x<4\right)=\int_{\frac{1}{3}}^{4}f(x)\,dx=\int_{\frac{1}{3}}^{\frac{1}{2}}2x\,dx+\int_{\frac{1}{2}}^{3}\frac{6}{25}(3-x)\;dx+0$$

$$=\left(\frac{1}{4}-\frac{1}{9}\right)+\frac{6}{25}\left[\left(9-\frac{9}{2}\right)-\left(\frac{3}{2}-\frac{1}{8}\right)\right]=\frac{8}{9}$$

11. $If\ f(x)=\begin{cases}xe^{-\frac{x^2}{2}} & ;x\ge 0\\ 0 & ;x>0\end{cases}$ **, then show that f(x) is a pdf and find F(x).**

Solution: **[AU N/D 2014]**

To prove: $\int_{-\infty}^{\infty}f(x)\,dx=1$

$$\int_0^\infty xe^{-\frac{x^2}{2}}\,dx = \int_0^\infty e^{-t}\,dt = \left[-e^{-t}\right]_0^\infty = -\left[e^{-\infty} - e^{-0}\right] = 1$$

Therefore given f(x) is PDF.

$$F(x)=P(X \le x)= \int_0^x xe^{-\frac{x^2}{2}}\,dx = \int_0^{\frac{x^2}{2}} e^{-t}\,dt = 1 - e^{-\frac{x^2}{2}},\ x \ge 0.$$

12. The cumulative distribution function (cdf)of a random variable X is $F(x)=1-(1+x)e^{-x}$,x>0. Find the probability density function of X. Mean and Variance of X.

Solution: **[AU M/J 2006, N/D 2010]**

Given $F(x)=1-(1+x)e^{-x}$,x>0

$=1-e^{-x}-xe^{-x}$,x>0

p.d.f f(x) $= \frac{d}{dx}[F(x)]$

$$= \frac{d}{dx}[1-e^{-x} - xe^{-x}] = e^{-x} - [x(-e^{-x}) + e^{-x}] = xe^{-x},\ x > 0$$

$$E[x]=\int_0^\infty xf(x)\,dx = \int_0^\infty x^2 e^{-x}\,dx$$

$$= \left[x^2\left[\frac{e^{-x}}{-1}\right] - (2x)\left[\frac{e^{-x}}{(-1)^2}\right] + 2\left[\frac{e^{-x}}{(-1)^3}\right]\right]_0^\infty = 2$$

$$E[x^2]=\int_0^\infty x^2 f(x)\,dx = \int_0^\infty x^3 e^{-x}\,dx$$

$$= \left[x^3\left[\frac{e^{-x}}{-1}\right] - (3x^2)\left[\frac{e^{-x}}{(-1)^2}\right] + (6x)\left[\frac{e^{-x}}{(-1)^3}\right] - (6)\left[\frac{e^{-x}}{(-1)^3}\right]\right]_0^\infty = 6$$

Result:

$Var(X)=E(x^2) - [E(x)]^2 = 6 - 4 = 2$

13. If X is a random variable with a continuous distribution function F(x),prove that Y=F(x) has a uniform distribution in (0,1).Further if

$$f(x) = \begin{cases} \frac{1}{2}(x-1); 1 \le x \le 3 \\ 0 \quad ; otherwise \end{cases},\ find\ the\ range\ of\ Y\ corresponding\ to\ the\ range\ 1.1 \le x \le 2.9.$$

Solution: **[AU N/D 2010]**

The distribution function of Y

$G_Y(Y)=P[Y \le y]$

$=P[F(x) \le y]$

$=P[X \le F^{-1}(y)]$

$=F[F^{-1}(y)] = y$ [Since $P[X \le x]=F(x)$]

$\therefore$ The density function of Y is given by

$$g_y(Y)=\frac{d}{dy}[G_y(y)]$$

$$=\frac{d}{dy}[y]=1$$

The range F[x]=(0,1)

Therefore, Y follows uniform distribution in (0,1)

$$\text{Given } f(x)=\begin{cases}\frac{1}{2}(x-1)\,;1\le x\le 3\\ 0 \quad ;otherwise\end{cases}$$

$$F[x]=\int_{-\infty}^{x} f(x)dx=\int_{-\infty}^{x} 0\,dx=0 \text{ when } x<1$$

$$F[x]=\int_{-\infty}^{1} f(x)dx+\int_{1}^{x} f(x)dx=0+\int_{1}^{x}\frac{1}{2}(x-1)\,dx=\left[\frac{1}{2}\left(\frac{x^2}{2}-x\right)\right]_1^x=\frac{1}{4}(x-1)^2 \text{ when } 1<x<3$$

$$F[x]=\int_{-\infty}^{1} f(x)dx+\int_{1}^{3} f(x)dx+\int_{3}^{x} f(x)dx=0+\int_{1}^{3}\frac{1}{2}(x-1)\,dx+0=\left[\frac{1}{2}\left(\frac{x^2}{2}-x\right)\right]_1^3=1 \text{ when } x>3$$

Result:

$$F(x)=\begin{cases}0 & if\ \ x<1\\ \frac{1}{4}(x-1)^2\,;\ 1\le x\le 3\\ 1 & if\ \ x>3\end{cases}$$

$$Y=F(x)=\begin{cases}0 & if\ \ x<1\\ \frac{1}{4}(x-1)^2\,;\ 1\le x\le 3\\ 1 & if\ \ x>3\end{cases}$$

If $1.1\le X\le 2.9$ then $\frac{1}{4}(1.1-1)^2\le Y\le\frac{1}{4}(2.9-1)^2$

i.e., $0.0025\le Y\le 0.9025$.

14. If the density function of X equals $f(x)=\begin{cases}Ce^{-2x} & ;x\ge 0\\ 0 & ;x<0\end{cases}$ **Find C . What is P[X>2]?**

Solution: **[AU A/M 2010]**

We know that $\int_{-\infty}^{\infty} f(x)dx=1$

$$\int_0^{\infty} Ce^{-2x}dx=1\Rightarrow C\int_0^{\infty} e^{-2x}dx=1\Rightarrow C\left[\frac{e^{-2x}}{-2}\right]_0^{\infty}=1\Rightarrow C=2$$

Result: $$P[X > 2] = \int_2^\infty 2e^{-2x}dx = 2\left[\frac{e^{-2x}}{-2}\right]_2^\infty = e^{-4}$$

15. A continuous random variable X has a pdf f(x)=kx²e⁻ˣ , x≥0. Find K , mean and variance.

Solution:

To find K: **[A.U M/J 2013 R.P]**

We know that $\int_{-\infty}^{\infty} f(x)dx = 1$

$$\int_0^\infty kx^2e^{-x}dx = 1 \Rightarrow k\int_0^\infty x^2e^{-x}dx = 1 \Rightarrow 2k = 1 \Rightarrow k = \frac{1}{2}$$

To find Mean:

$$\text{Mean}=E[x]=\int_0^\infty xf(x)\,dx = \int_0^\infty \frac{1}{2}x^3e^{-x}dx$$

$$= \frac{1}{2}\left[x^3\left[\frac{e^{-x}}{-1}\right] - (3x^2)\left[\frac{e^{-x}}{(-1)^2}\right] + (6x)\left[\frac{e^{-x}}{(-1)^3}\right] - (6)\left[\frac{e^{-x}}{(-1)^3}\right]\right]_0^\infty = \frac{1}{2}(6) = 3$$

To find Variance:

$$E[x^2]=\int_0^\infty x^2 f(x)\,dx = \int_0^\infty \frac{1}{2}x^4e^{-x}dx$$

$$= \frac{1}{2}\left[x^4\left[\frac{e^{-x}}{-1}\right] - (4x^3)\left[\frac{e^{-x}}{(-1)^2}\right] + (12x^2)\left[\frac{e^{-x}}{(-1)^3}\right] - (24x)\left[\frac{e^{-x}}{(-1)^3}\right] + 24\left[\frac{e^{-x}}{(-1)^4}\right]\right]_0^\infty = 12$$

Result:

$\text{Var}(X)=E(x^2)-[E(x)]^2=12 - 9 = 3.$

16. A continuous random variable X has the pdf

$f(x) = \frac{K}{1+x^2}$; $-\infty < x < \infty$ **find (i) the value of K (ii) the distribution function of X (iii) P[X≥0]** **[AU N/D 2011]**

Solution:

(i) We know that $\int_{-\infty}^{\infty} f(x)dx = 1$

$$\int_{-\infty}^{\infty}\frac{K}{1+x^2}dx = 1 \Rightarrow k\left[\tan^{-1}(x)\right]_{-\infty}^{\infty} = 1 \Rightarrow k\left[\tan^{-1}(\infty) - \tan^{-1}(-\infty)\right] = 1 \Rightarrow K\left[\frac{\pi}{2}+\frac{\pi}{2}\right] = 1$$

Result : $K(\pi) = 1 \Rightarrow K = \frac{1}{\pi}$

(ii) To find Distribution function of X

$$F[x] = \int_{-\infty}^{x} f(x)dx = \int_{-\infty}^{x}\frac{1}{\pi}\frac{1}{1+x^2}dx = \frac{1}{\pi}\left[\tan^{-1}x\right]_{-\infty}^{x} = \frac{1}{\pi}\left[\tan^{-1}(x) - \tan^{-1}(-\infty)\right] = \frac{1}{\pi}\left[\tan^{-1}(x) + \frac{\pi}{2}\right]$$

(iii)To find P[X≥0]

$$P[X \geq 0] = \int_0^\infty f(x)dx = \int_0^\infty \frac{1}{\pi}\frac{1}{1+x^2}dx = \frac{1}{\pi}\left[\tan^{-1} x\right]_0^\infty = \frac{1}{\pi}\left[\tan^{-1}(\infty) - \tan^{-1}(0)\right] = \frac{1}{\pi}\left[\frac{\pi}{2} - 0\right] = \frac{1}{2}$$

17. If the density function of X equals $f(x) = \begin{cases} Ce^{-2x}, & 0 < x < \infty \\ 0, & x < 0 \end{cases}$,

Find 'C' . What is $P[X > 2]$?

Solution:

Since $f(x)$ is a p.d.f. then $\int_{-\infty}^{\infty} f(x)dx = 1 \Rightarrow \int_0^\infty Ce^{-2x}dx = 1$

$$C\int_{-\infty}^{\infty} e^{-2x}dx = 1 \Rightarrow C\left[\frac{e^{-2x}}{-2}\right]_0^\infty = 1 \Rightarrow C\left[0 + \frac{1}{2}\right] = 1$$

Result: $C = 2$

To find P(X>2):

$$P[X > 2] = \int_2^\infty f(x)dx = \int_2^\infty 2e^{-2x}dx = 2\int_2^\infty e^{-2x}dx$$

Result: $= 2\left[\frac{e^{-2x}}{-2}\right]_2^\infty = -\left[0 - e^{-4}\right] = e^{-4}$

18. The probability density function of a random variable X is given by

$f_X(X) = \begin{cases} x & ; 0 < X < 1 \\ k(2-x); & 1 \leq X \leq 2 \\ 0 & ; otherwise \end{cases}$ **(i) Find the value of K? (ii) Find** $P(0.2 < X < 1.2)$

(iii) What is $P(0.5 < X < 1.5 / X \geq 1)$

(iv) Find the distribution function of f(x) **[A.U. A/M 2011]**

Solution:

Since $f(x)$ is the pdf, we have

$$\int_{-\infty}^{\infty} f(x)dx = 1 \Rightarrow \int_0^2 f(x)dx = 1$$

$$\int_0^1 x\, dx + \int_1^2 k(2-x)\, dx = 1$$

$$\left[\frac{x^2}{2}\right]_0^1 + k\left[2x - \frac{x^2}{2}\right]_1^2 = 1$$

$$\left[\frac{1}{2} - 0\right] + k\left[(4-2) - (2 - \frac{1}{2})\right] = 1$$

$$k\left[\frac{1}{2}\right]=1-\frac{1}{2} \quad \Rightarrow k\left[\frac{1}{2}\right]=\frac{1}{2}$$

Result: $k=1$

(ii) To find $P(0.2 < X < 1.2)$

$$P(0.2 < X < 1.2)=\int_{0.2}^{1.2} f(x)\,dx$$

$$=\int_{0.2}^{1} x\,dx+\int_{1}^{1.2}(2-x)\,dx$$

$$=\left[\frac{x^2}{2}\right]_{0.2}^{1}+\left[2x-\frac{x^2}{2}\right]_{1}^{1.2}$$

$$=\left[\frac{1}{2}-\frac{0.04}{2}\right]+\left[(2.4-0.72)-(2-\frac{1}{2})\right]$$

$$=0.48+1.68-1.5$$

Result: $=0.66$

(iii) To find $P(0.5 < X < 1.5 / X \geq 1)$

$$P(0.5 < X < 1.5 / X \geq 1)=\frac{P[(0.5 < X < 1.5)\cap(X \geq 1)]}{P(X \geq 1)}$$

$$=\frac{P[1 \leq X \leq 1.5]}{P(X \geq 1)}$$

$$P(1 \leq X < 1.5)=\int_{1}^{1.5}(2-x)\,dx=\left[2x-\frac{x^2}{2}\right]_{1}^{1.5}$$

$$=\left[2(1.5)-\frac{(1.5)^2}{2}-(2-\frac{1}{2})\right]$$

$$=3-1.125-1.5$$

$$=0.375$$

$$P(X \geq 1)=\int_{1}^{2}(2-x)\,dx=\left[2x-\frac{x^2}{2}\right]_{1}^{2}$$

$$=\left[4-2-(2-\frac{1}{2})\right]$$

$$=0.5$$

$$P(0.5 < X < 1.5 / X \geq 1)=\frac{0.375}{0.5}=0.75$$

(iv) To find $F(x)=P(X \leq x)=\int_{0}^{x} f(x)\,dx$

(i) if $x \leq 0$ *then* $F(x)=0$

(ii) If $0 < x \le 1$ then $F(x) = \int_0^x x\,dx$

$$= \left[\frac{x^2}{2}\right]_0^x = \frac{x^2}{2}$$

(iii) If $1 \le x < 2$ then $F(x) = \int_0^1 x\,dx + \int_1^x (2-x)\,dx$

$$= \left[\frac{x^2}{2}\right]_0^1 + \left[2x - \frac{x^2}{2}\right]_1^x$$

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

$$= \frac{1}{2} + 2x - \frac{x^2}{2} - 2 + \frac{1}{2} = 2x - \frac{x^2}{2} - 1$$

(iv) If $x > 2$ then $F(x) = \int_{-\infty}^{x} f(x)\,dx$

$$= \int_0^1 x\;dx + \int_1^2 (2-x)\,dx + \int_2^x 0\,dx$$

$$= \left[\frac{x^2}{2}\right]_0^1 + \left[2x - \frac{x^2}{2}\right]_1^2$$

$$= \frac{1}{2} + \left[4 - 2 - (2 - \frac{1}{2})\right]$$

$$= 1$$

MOMENT GENERATING FUNCTION

19. The probability function of an infinite discrete distribution is given by P(X=x)= $\frac{1}{2^x}$,

x=1,2,….. ∞ Find the mean and variance of the distribution. Also find P(X is even).

Solution: [AU N/D 2011,'18]

We know that $M_X(t) = \sum_{x=1}^{\infty} e^{tx}\,p(x)$

$$= \sum_{x=1}^{\infty} e^{tx} \cdot \frac{1}{2^x} = \sum_{x=1}^{\infty} \left(\frac{e^t}{2}\right)^x$$

$$= \frac{e^t}{2} + \left(\frac{e^t}{2}\right)^2 + ...$$

$$= \frac{e^t}{2}\left[1 + \frac{e^t}{2} + \left(\frac{e^t}{2}\right)^2 + ...\right] \text{[Using } (1-x)^{-1} = 1 + x + x^2 + ...]$$

$$= \frac{e^t}{2}\left[1 - \frac{e^t}{2}\right]^{-1}$$

$$= \frac{e^t}{2}\frac{\left(2 - e^t\right)^{-1}}{2^{-1}}$$

$$M_X(t) = \frac{e^t}{2 - e^t}$$

$$= (2 - e^t)^{-1} e^t \qquad (1)$$

$$M'_X(t) = -e^t (2 - e^t)^{-2}(-e^t) + (2 - e^t)^{-1} e^t$$

$$= e^{2t}(2 - e^t)^{-2} + (2 - e^t)^{-1} e^t \qquad (2)$$

$$M''_X(t) = 2\left(2 - e^t\right)^{-2} e^{2t} + e^{2t}(-2)\left(2 - e^t\right)^{-3}\left(-e^t\right)$$

$$+ \left(2 - e^t\right)^{-1} e^t + e^t(-1)\left(2 - e^t\right)^{-2}\left(-e^t\right) \qquad .(3)$$

$$Now \quad E(X) = Mean = M_X{}'(0)$$

$$= 1 + 1 = 2 \qquad [put\ t = 0\ in\ (2)]$$

$$E(X^2) = M_X{}''(0) = 6 \qquad [Put\ t = 0\ in\ (3)]$$

Mean $\mu_1' = 2$

$$Variance = E\left[X^2\right] - \left[E(X)\right]^2 = 6 - 4 = 2$$

Now P(X=even)=P(X=2)+P(X=4)+…....

$$= \left(\frac{1}{2}\right)^2 + \left(\frac{1}{2}\right)^4 + ...$$

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

$$\left[\because Sum\ to\ \infty\ of\ a\ goemetric\ series\ is\ \frac{a}{1-r},\ a - 1^{st}\ term,\ r - common\ ratio\right]$$

$$= \frac{\frac{1}{4}}{1-\frac{1}{4}} = \frac{1}{4} \times \frac{4}{4-1} = \frac{1}{3}$$

Result:

MGF	Mean	Variance	P(X=even)
$e^t\left(2-e^t\right)^{-1}$	2	2	$\frac{1}{3}$

20. Find the MGF of the random variable with the probability law $P(X = x) = q^{x-1}\ p,\ x = 1,2,3....$ **Find the Mean and Variance and also find P(X is odd).** **[AU A/M 2017,'18]**

Solution:

We know that,

$$M_X(t) = E\left(e^{tx}\right)$$

$$= \sum_{x=1}^{\infty} e^{tx} \cdot p(x) \qquad [By\ defination]$$

$$= \sum_{x=1}^{\infty} e^{tx} \cdot q^{x-1} \cdot p \qquad \left[\because p(x) = q^{x-1}\ p\right]$$

$$= \sum_{x=1}^{\infty} e^{tx} \cdot q^{x} \cdot q^{-1}\ p$$

$$= \sum_{x=1}^{\infty} \left(e^t q\right)^x \frac{p}{q}$$

$$= \frac{p}{q} \cdot q\, e^t \sum_{x=1}^{\infty} \left(q\, e^t\right)^{x-1}$$

$$= p\ e^t \left[1 + \left(q\, e^t\right) + \left(q\, e^t\right)^2 + ...\right]$$

$$= p\ e^t \left(1 - q\ e^t\right)^{-1}$$

$$MGF\ is\ M_x(t) = \frac{p\ e^t}{1 - q\ e^t} \qquad \left[\begin{array}{l}\underline{Formula}: \\ (1-x)^{-1} = 1 + x + x^2 + ...\end{array}\right]$$

Differentiating

$$\frac{d}{dt}\{M_X(t)\} = \frac{(1-qe^t)pe^t - pe^t(-qe^t)}{(1-qe^t)^2}$$

$$= \frac{pe^t - pqe^{2t} + pqe^{2t}}{(1-qe^t)^2}$$

$$M'_X(t) = \frac{pe^t}{(1-qe^t)^2} \qquad ...(2)$$

$$E(X) = M'_X(0)$$

$$= \frac{p}{(1-q)^2} \qquad [Put\ t = 0\ in\ (2)]$$

$$= \frac{p}{p^2} \qquad [\because p + q = 1]$$

Result : $E(X) = Mean = \dfrac{1}{p}$ ....(A)

Differentiating (2) w.r.t. 't', we get

$$M''_X(t) = \frac{(1-qe^t)^2\ pe^t - pe^t\,2(1-qe^t)(-qe^t)}{(1-qe^t)^4}$$

$$= \frac{(1-qe^t)\{(1-qe^t)pe^t + 2pqe^{2t}\}}{(1-qe^t)^4} = \frac{pe^t + pqe^{2t}}{(1-qe^t)^3}$$

$$M''_X(t) = \frac{pe^t(1+qe^t)}{(1-qe^t)^3} ......(3)$$

$$\therefore\ E(X^2) = M''_X(0) \Rightarrow E(X^2) = \frac{p(1+q)}{(1-q)^3} = \frac{1+q}{p^2}$$

$$Mean = \frac{1}{p}$$

$$Variance = E(X^2) - [E(X)]^2 = \frac{1+q}{p^2} - \left(\frac{1}{p}\right)^2 = \frac{1+q-1}{p^2} = \frac{q}{p^2}$$

Result: $\therefore\ Variance = \dfrac{q}{p^2}$

$$P(X = odd) = P(X=1) + P(X=3) + P(X=5) + ...$$

$$= p + pq^2 + pq^4 + ...$$

$$= p\left(1 + q^2 + q^4 + ...\right) = p\left(1 - q^2\right)^{-1} = \frac{p}{\left(1 - q^2\right)}$$

21. Let 'X' be a random variable with p.d.f $f(x) = \begin{cases} \frac{1}{3} e^{-\frac{X}{3}} , & x > 0 \\ 0, & othewise \end{cases}$

Find (a) P(X>3) (b) MGF of 'X' (c) E(X) and Var(X)

Solution:

Given $f(x) = \begin{cases} \frac{1}{3} e^{-\frac{1}{3}} , & x > 0 \\ 0, & othewise \end{cases}$

(a) To find P(X>3)

$$\text{P}(\text{X} > 3) = \int_3^\infty f(x)\, dx = \int_3^\infty \frac{1}{3} e^{\frac{x}{3}}\, dx = \frac{1}{3}\left[\frac{e^{-\frac{x}{3}}}{-\frac{1}{3}}\right]_3^\infty$$

$$= -\left[0 - e^{-1}\right] = e^{-1} = \frac{1}{e} = 0.3679$$

(b) To find MGF of 'X'

$$M_X(t) = E(e^{tX}) = \int_{-\infty}^{\infty} e^{tx} f(x) dx$$

$$= \int_o^\infty e^{tx} \frac{1}{3} e^{-\frac{x}{3}} dx \qquad \left[\because x > 0, f(x) = \frac{1}{3} e^{-\frac{x}{3}}\right]$$

$$= \frac{1}{3}\int_0^\infty e^{\left(t - \frac{1}{3}\right)x}\, dx$$

$$= \frac{1}{3}\int_0^\infty e^{-\left(\frac{1}{3} - t\right)x}\, dx = \frac{1}{3}\left[\frac{e^{-\left(\frac{1}{3} - t\right)x}}{-\left(\frac{1}{3} - t\right)}\right]_0^\infty$$

$$= \frac{1}{3}\left[\frac{1}{\left(\frac{1-3t}{3}\right)}\right]$$

Result: $MGF\ of\ X = M_X(t) = \dfrac{1}{1-3t}$

(c) To find E(X), Var(X):

$$M_X(t) = \frac{1}{1-3t} \quad = (1-3t)^{-1}$$

$$M'_X(t) = -(1-3t)^{-2}(-3)$$

$$= 3(1-3t)^{-2}$$

$$E(X) = Mean \ = M'_X(0) = 3 \qquad .....(A)$$

$$M''_X(t) = -6(1-3t)^{-3}(-3) = 18(1-3t)^{-3}$$

$$M''_X(0) = 18$$

$$E(X^2) \ = M''_X(0) = 18 \qquad ....(B)$$

$$\therefore \quad Var(X) = E(X^2) - [E(X)]^2$$

$$= 18 - 3^2 \quad = 18 - 9 \qquad [U\sin g\,(A)\, and (B)]$$

$$\therefore \quad Var(X) = 9$$

Result:

P(X>3)	*MGF*	*Mean*	*Variance*
$\frac{1}{e} = 0.3679$	$\frac{1}{1-3t}$	3	9

22. Let 'X' be a random variable with p.d.f $f(x) = \begin{cases} \frac{1}{2}e^{-\frac{x}{2}}, & x > 0 \\ 0, & otherwise \end{cases}$

Find mean and variance **[A.U N/D 2013 R.P]**

Solution:

To find MGF of 'X'

$$M_X(t) = E(e^{tX}) = \int_{-\infty}^{\infty} e^{tx} f(x)dx$$

$$= \int_o^{\infty} e^{tx}\ \frac{1}{2}\ e^{-\frac{x}{2}} d\theta \quad \left[\because x > 0, f(x) = \frac{1}{3}e^{-\frac{x}{3}}\right]$$

$$= \frac{1}{2}\int_0^\infty e^{\left(t-\frac{1}{2}\right)x}\ dx \ = \frac{1}{2}\int_0^\infty e^{-\left(\frac{1}{2}-t\right)x}\ dx$$

$$= \frac{1}{2}\left[\frac{e^{-\left(\frac{1}{2}-t\right)x}}{-\left(\frac{1}{2}-t\right)}\right]_0^\infty = \frac{1}{2}\left[\frac{1}{\left(\frac{1-2t}{2}\right)}\right]$$

Result: $MGF\ of\ X = M_X(t) = \dfrac{1}{1-2t}$

(c) To find E(X), Var(X):

$$M_X(t) = \frac{1}{1-2t} \quad = (1-2t)^{-1}$$

$$M_X'(t) = -(1-2t)^{-2}(-2) = 2(1-2t)^{-2}$$

$$E(X) = Mean \ = M_X'(0) = 2..................(A)$$

$$M_X''(t) = -4(1-2t)^{-3}(-2) = 8(1-2t)^{-3} \Rightarrow M_X''(0) = 8$$

$$E(X^2) = M_X''(0) = 8............(B)$$

$$\therefore\ Var(X) = E(X^2) - [E(X)]^2 = 8-2^2 = 8-4 = 4 \qquad [U\sin g\,(A)\,and\,(B)]$$

$$\therefore\ Var(X) = 4$$

23. Find the MGF of the random variable X with p.d.f $f(x) = \begin{cases} \dfrac{x}{4}e^{-\frac{x}{2}}, & x > 0 \\ 0, & otherwise \end{cases}$

Also find the first four moments about the origin. **[A.U N/D 2013, M/J 2014,'17]**

Solution:

To find MGF of 'X'

$$M_X(t) = E(e^{tX}) = \int_{-\infty}^{\infty} e^{tx} f(x)dx$$

$$= \int_o^\infty e^{tx}\ \frac{1}{4}\ e^{-\frac{x}{2}}dx \qquad \left[\because x > 0, f(x) = \frac{1}{3}e^{-\frac{x}{3}}\right]$$

$$= \frac{1}{4}\int_0^\infty e^{\left(\frac{1}{2}-t\right)x}\ x\,dx$$

$$=\frac{1}{4}\left[x\frac{e^{-\left(\frac{1}{2}-t\right)}}{-\left(\frac{1}{2}-t\right)}-(1)\frac{e^{--\left(\frac{1}{2}-t\right)x}}{-\left(\frac{1}{2}-t\right)^{2}}\right]_{0}^{\infty}$$

$$MGF\ of\ X=M_X(t)=\left(\frac{1}{1-2t}\right)^{-2}$$

$$MGF\ of\ X=M_X(t)=1+2(2t)+3(2t)^2+4(2t)^4+.....$$

$$MGF\ of\ X=M_X(t)=1+\frac{t}{1!}(4)+\frac{t^2}{2!}(24)+\frac{t^3}{3!}(192)+.\frac{t^4}{4!}.(1920)...$$

Result: First 4 moments are 4, 24, 192 and 1920

24. A random variable has the p.d.f given by $f(x)=\begin{cases}2e^{-2x}, & x\geq 0\\ 0\ , & x<0\end{cases}$

Find (a) the moment generating function (b) the first four moment about the origin.

(or)

Let X be an exponential random variable with $E(X^2)=\frac{1}{2}$**. Obtain E(X), Var(X), MGF and** $P(X>3/X>1)$

Solution: **[A.U N/D 2014, A/M 2021]**

Given $f(x)=2e^{-2x}, \quad x\geq 0$ ....(1)

$$M_X(t)=\int_{-\infty}^{\infty}e^{tx}f(x)dx=\int_{o}^{\infty}e^{tx}\cdot 2e^{-2x}dx \qquad [U\sin g\,(1), x\geq 0]$$

$$=2\int_0^{\infty}e^{-(2-t)x}\,dx=2\left[\frac{e^{-(2-t)x}}{-(2-t)}\right]_0^{\infty}$$

$$=-\frac{2}{2-t}\left[-e^{-\infty}-e^{-0}\right]=-\frac{2}{2-t}[0-1]=\frac{2}{2-t}$$

$$\therefore MGF\ \ is\ M_X(t)=2(2-t)^{-1}$$

$$M_X'(t)=-2(2-t)^{-2}(-1)$$
$$=2(2-t)^{-2} \qquad .....(1A)$$

$\therefore 1^{s}\ moment\ \ M_X^{'}(0)=\dfrac{2}{4}=\dfrac{1}{2}$ $\quad [Put\ t=0\ in\ (1)]$

$$M_x^{''}(t)=2(-2)(2-t)^{-3}(-1)$$
$$=4(2-t)^{-3} \qquad .(2)$$

$2^{nd}\ moment\ \ M_X^{''}(t)=\dfrac{4}{8}=\dfrac{1}{2}$ $\quad [Put\,t=0\ in\ (2)]$

$$Var(X)=\frac{1}{2}-\frac{1}{4}=\frac{1}{4}$$

$$M_X^{'''}(t)=-12(2-t)^{-4}(-1)$$
$$=12(2-t)^{-4}\ \ldots\ldots\ldots\ldots\ldots\ldots(3)$$

$3^{rd}\ moment\ \ M_X^{'''}(0)=\dfrac{12}{16}=\dfrac{3}{4}$ $\quad [Put\,t=0\ in\ (3)]$

$$M_X^{IV}(t)=-48(2-t)^{-5}(-1)$$
$$=48(2-t)^{-5} \qquad \ldots(4)$$

$4^{th}\ moment\ \ M_X^{IV}(t)=\dfrac{48}{32}=\dfrac{3}{2}$ $\quad [Put\ t=0\ in(4)]$

Result:

MGF	$1^{st}\ moment$	$2^{nd}\ moment$	$3^{rd}\ moment$	$4^{th}\ moment$
$\dfrac{2}{2-t}$	$\dfrac{1}{2}$	$\dfrac{1}{2}$	$\dfrac{3}{4}$	$\dfrac{3}{2}$

$$P(X>3/X>1)=\frac{P(X>3\cap X>1)}{P(X>1)}=\frac{P(X>3)}{P(X>1)}=\frac{\int_3^{\infty}2e^{-2x}dx}{\int_1^{\infty}2e^{-2x}dx}=\frac{e^{-6}}{e^{-2}}=e^{-4}$$

25. Find the mean, variance and MGF of the random variable X having the pdf

$$f(x)=\begin{cases} x & ,0<x\le 1 \\ 2-x & ,1<x\le 2 \\ 0 & ,otherwise \end{cases}$$

[AU N/D 2013]

Solution:

Mean= E(x)= $\int_{-\infty}^{\infty} x f(x)dx$

$$= \int_0^1 x(x)dx + \int_1^2 x(2-x)dx = \left[\frac{x^3}{3}\right]_0^1 + \left[\frac{2x^2}{2} - \frac{x^3}{3}\right]_1^2$$

$$= \left[\frac{1^3}{3} - 0\right] + \left[\left(4 - \frac{8}{3}\right) - \left(\frac{2}{2} - \frac{1}{3}\right)\right] = 1$$

E(x^2) = $\int_{-\infty}^{\infty} x^2 f(x)dx$

$$= \int_0^1 x^2(x)dx + \int_1^2 x^2(2-x)dx = \left[\frac{x^4}{4}\right]_0^1 + \left[\frac{2x^3}{3} - \frac{x^4}{4}\right]_1^2$$

$$= \left[\frac{1}{4} - 0\right] + \left[\left(\frac{16}{3} - \frac{16}{4}\right) - \left(\frac{2}{3} - \frac{1}{4}\right)\right] = \frac{7}{6}$$

Var(X)=E(x^2)-[E(x)]^2=(7/6)-1=1/6.

MGF= $\int_{-\infty}^{\infty} e^{tx} f(x)dx$

$$= \int_0^1 e^{tx} x\,dx + \int_1^2 e^{tx}(2-x)\,dx = \left[\frac{xe^{tx}}{t} - \frac{e^{tx}}{t^2}\right]_0^1 + \left[\frac{(2-x)e^{tx}}{t} + \frac{e^{tx}}{t^2}\right]_1^2$$

$$= \left[\left(\frac{e^t}{t} - \frac{e^t}{t^2}\right) - \left(\frac{0}{t} - \frac{1}{t^2}\right)\right] + \left[\left(\frac{0}{t} + \frac{e^{2t}}{t^2}\right) - \left(\frac{e^t}{t} + \frac{e^t}{t^2}\right)\right] = \frac{1}{t^2}\left(e^t - 1\right)^2$$

26. Find the mean, variance and MGF of the random variable X having the pdf

$$f(x) = \begin{cases} \lambda e^{-\lambda(X-a)}, & x \geq a \\ 0, & otherwise \end{cases}$$

Solution:

To find MGF:

$$M_X(t) = \int_0^{\infty} e^{tx} f(x)dx$$

$$= \int_0^{\infty} e^{tx}\lambda e^{-\lambda(x-a)}dx = \lambda\, e^{a\lambda}\int_0^{\infty} e^{-(\lambda-t)x}dx$$

$$= \lambda e^{at}\left[\frac{e^{-(\lambda-t)x}}{-(\lambda-t)}\right]_0^{\infty} = \frac{\lambda}{-(\lambda-t)}e^{at}\left[e^{-\infty} - e^{-0}\right]$$

$$= e^{at}\,\frac{\lambda}{(\lambda-t)}$$

$$\therefore \quad M_X(t) = e^{at}\frac{\lambda}{(\lambda - t)}$$

To find Mean and Variance:

$$E(X) = \left[\frac{d}{dt}e^{at}\frac{\lambda}{(\lambda - t)}\right]_{t=0}$$

$$= \left[\left(\frac{\lambda}{\lambda - t}\right)ae^{at} + e^{at}\frac{-\lambda}{(\lambda - t)^2}(-1)\right]_{t=0}$$

$$= \left[a + \frac{1}{\lambda}\right]$$

$$E(X^2) = \left[\frac{d^2}{dt^2}e^{at}\frac{\lambda}{(\lambda - t)}\right]_{t=0} \quad \Rightarrow E(X^2) = a^2 + \frac{2a}{\lambda} + \frac{2}{\lambda^2}$$

$$Var(X) = E(X^2) - [E(X)]^2$$

$$Var(X) = \left[a^2 + \frac{2a}{\lambda} + \frac{2}{\lambda^2}\right]^2 - \left[a + \frac{1}{\lambda}\right] = \frac{1}{\lambda^2}$$

Discrete Distributions

BINOMIAL DISTRIBUTION

27. Define binomial distribution, Find the moment generating function of binomial distribution And also find its mean and variance. [AU N/D 2024]

Solution:

A random variable X is said to follow binomial distribution if it assumes only non-negative values and its probability mass function p(x) is given by

$$P(X = x) = p(x) = \begin{cases} nC_x p^x q^{n-x}, & x = 0,1,2,3,....n,\ q = 1 - p \\ 0 & , \ otherwise \end{cases}$$

Where n and p are parameters.

To find MGF

We know that the moment generating function of a random variable X about origin whose probability function p(x) is given by

$$M_X(t) = E\left[e^{tX}\right] = \sum_{x=0}^{n} e^{tx} p(x) = \sum_{x=0}^{n} e^{tx} nC_x p^x q^{n-x}$$

$$= \sum_{x=0}^{n} (pe^t)^x nC_x q^{n-x}$$

$$= \sum_{x=0}^{n} nC_x (pe^t)^x q^{n-x}$$

$$= q^n + nC_1 q^{n-1}\left(pe^t\right)^1 + nC_2 q^{n-2}\left(pe^t\right)^2 + \ldots\ldots\ldots\ldots\ldots$$

$$= \left(q + pe^t\right)^n$$

$$\therefore M_X(t) = \left(q + pe^t\right)^n$$

To find Mean and Variance:

$$M_X(t) = \left(q + pe^t\right)^n$$

$$\therefore \; M_X^1(t) = n\left(q + pe^t\right)^{n-1} pe^t$$

Put t=0, we get

$$M_X^1(0) = n\left(q + p\right)^{n-1} p$$

$$M_X^{11}(t) = np\left[\left(q + pe^t\right)^{n-1} e^t + e^t(n-1)(q + pe^t)^{n-2} pe^t\right]$$

Put t=0, we get

$$M_X^{11}(0) = np\left[\left(q + p\right)^{n-1} + (n-1)(q + p)^{n-2} p\right]$$

$$M_X^{11}(0) = np\left[1 + (n-1)p\right] = np + n^2 p^2 - np^2$$

$$= n^2 p^2 + np(1-p)$$

$$= n^2 p^2 + npq$$

$$Variance = M_X^{11}(0) - \left(M_X^1(0)\right)^2$$

$$= npq + n^2 p^2 - n^2 p^2$$

$$Var(X) = npq$$

POISSON DISTRIBUTION

28. By Calculating the moment generating function of Poisson distribution with parameter λ, Prove that the mean and variance of the Poisson distribution are equal.

Solution: **[AU N/D 2014, 15]**

$$P(X = x) = p(x) = \frac{e^{-\lambda}\lambda^x}{x!}$$

The m.g.f $M_X(t) = \sum_{x=0}^{\infty} e^{tx} p(x) = \sum_{x=0}^{\infty} e^{tx} \frac{e^{-\lambda}\lambda^x}{x!}$

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

$$= e^{-\lambda}\left[1 + \frac{\lambda e^t}{1!} + \frac{(\lambda e^t)^2}{2!} + \ldots\ldots\right]$$

$$= e^{-\lambda} e^{\lambda e^t} = e^{\lambda\left[e^t - 1\right]}$$

Mean, $$E(X) = \left[\frac{d}{dt}\left[M_X(t)\right]\right]_{t=0}$$

$$= \left[\frac{d}{dt}\left[e^{\lambda\left[e^t-1\right]}\right]\right]_{t=0} = \left[\frac{d}{dt}\left[e^{\lambda e^t}e^{-\lambda}\right]\right]_{t=0}$$

$$= \left[e^{-\lambda}\frac{d}{dt}\left[e^{\lambda e^t}\right]\right]_{t=0} = \left[e^{-\lambda}e^{\lambda e^t}\lambda e^t\right]_{t=0}$$

$$= e^{-\lambda}e^{\lambda}\lambda = \lambda$$

$$E\left(X^2\right) = \left[\frac{d}{dt}M_X(t)\right]_{t=0} = \left[\frac{d}{dt}\left[\lambda e^{-\lambda}e^{\lambda e^t}e^t\right]\right]_{t=0}$$

$$= \left[\lambda e^{-\lambda}\frac{d}{dt}\left[e^{\lambda e^t}e^t\right]\right]_{t=0}$$

$$= \left[\lambda e^{-\lambda}[e^{\lambda e^t}e^t + e^t\lambda e^t e^{\lambda e^t}]\right]_{t=0} = \lambda e^{-\lambda}[e^{\lambda} + \lambda e^{\lambda}]$$

$$= \lambda + \lambda^2$$

$$\text{Var(X)} = E(X^2) - \left[E(X)\right]^2 = \lambda + \lambda^2 - \lambda^2 = \lambda$$

GEOMETRIC DISTRIBUTION

29. Define Geometric distribution and Find the moment generating function, mean and variance of geometric distribution?

Solution: **[AU A/M '15]**

A discrete random variable 'X' is said to follow geometric distribution, if it assumes only non-negative values and its probability mass function is given by

$$P(X = x) = p(x) = q^{x-1}p, \qquad x = 1,2,3,... \ 0 < p \le 1 \ where \quad q = 1 - p$$

The moment generating function of geometric distribution is given by

$$M_X(t) = E(e^{tx}) = \sum_{x=0}^{\infty} e^{tx}p(x) = \sum_{x=1}^{\infty} e^{tx}q^{x-1}p$$

$$= \frac{p}{q}\sum_{x=1}^{\infty} e^{tx}q^x \qquad \because q^{x-1} = \frac{q^x}{q}$$

$$= \frac{p}{q}\sum_{x=1}^{\infty}(qe^t)^x$$

$$= \frac{p}{q}\left[qe^t + \left(qe^t\right)^2 + \left(qe^t\right)^3 + ...\right]$$

$$= \frac{p}{q}qe^t\left[1 + \left(qe^t\right)^1 + \left(qe^t\right)^2 + ...\right]$$

$$= pe^t\left(1 - qe^t\right)^{-1} \qquad \because \left(1 - x\right)^{-1} = 1 + x + x^2 + ...$$

$$M_X(t) = \frac{pe^t}{(1 - qe^t)}$$

Mean:

WKT, $M_X(t) = \frac{pe^t}{(1 - qe^t)}$

$$M_X'(t) = \frac{d}{dt}(M_X(t))$$

$$= \frac{d}{dt}\frac{pe^t}{(1 - qe^t)}$$

$$= p\left[\frac{(1-qe^t)e^t - e^t(-qe^t)}{(1-qe^t)^2}\right]$$

$$= p\left[\frac{e^t - qe^{2t} + qe^{2t}}{(1-qe^t)^2}\right]$$

$$= p\left[\frac{e^t}{(1-qe^t)^2}\right] \text{ ------------------ (1)}$$

Putting t=0 in (1)

$$\therefore E(X) = M_X'(0) = \left[\frac{p}{(1-q)^2}\right] = \frac{p}{p^2} = \frac{1}{p} \quad \because 1-q = p$$

$$M_X''(t) = \frac{d^2}{dt^2}(M_X(t)) = \frac{d}{dt}(M_X^1(t))$$

$$= \frac{d}{dt}\left[\frac{pe^t}{(1-qe^t)^2}\right]$$

$$= p\left[\frac{(1-qe^t)^2 e^t - e^t 2(1-qe^t)(-qe^t)}{(1-qe^t)^4}\right]$$

$$E(X^2) = M_X''(0) = p\left[\frac{(1-qe^0)^2 e^0 - e^0 2(1-qe^0)(-qe^0)}{(1-qe^0)^4}\right]$$

$$= p\left[\frac{(1-q)^2 - 2(1-q)(-q)}{(1-q)^4}\right]$$

$$= p\left[\frac{p^2 + 2pq}{p^4}\right] \qquad \because 1-q = p$$

$$= \left[\frac{p(p+2q)}{p^3}\right]$$

$$= \left[\frac{p+2q}{p^2}\right]$$

$$= \left[\frac{p+q+q}{p^2}\right] = \left[\frac{1+q}{p^2}\right] \qquad \because p+q=1$$

$$E(X^2) = \frac{1}{p^2} + \frac{q}{p^2}$$

$$Var(X) = E(X^2) - [E(X)]^2$$

$$= \frac{1}{p^2} + \frac{q}{p^2} - \frac{1}{p^2} = \frac{q}{p^2}$$

CONTINOUS DISTRIBUTION
UNIFORM DISTRIBUTION

30. Find the moment generating function of Uniform distribution. Hence find its mean and variance.

Solution:

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

$$The\ mgf\ M_X(t) = \int_a^b e^{tx} f(x)dx = \int_a^b e^{tx} \frac{1}{b-a} dx = \frac{1}{b-a}\left[\frac{e^{tx}}{t}\right]_a^b$$

$$= \frac{1}{b-a}\left[\frac{e^{bx} - e^{ax}}{t}\right] = \frac{e^{bt} - e^{at}}{(b-a)t}$$

$$\mu_r^1 = \int_a^b x^r f(x)dx \quad where\ f(x)\ is\ p.d.f\ of\ 'X'............(1)$$

$$\mu_1^1 = \int_a^b xf(x)dx \qquad [put\ r=1\ in\ (1)]$$

$$= \int_a^b x\frac{1}{b-a}dx = \frac{1}{b-a}\left[\frac{x^2}{2}\right]_a^b$$

$$= \frac{1}{2(b-a)}\left[b^2 - a^2\right] = \frac{(b+a)(b-a)}{2(b-a)}$$

$$Mean = \frac{(b+a)}{2}$$

$$\mu_2^1 = \int_a^b x^2 f(x)dx \qquad [put\ r = 2\ in\ (1)]$$

$$= \int_a^b x^2 \frac{1}{b-a} dx = \frac{1}{b-a}\left[\frac{x^3}{3}\right]_a^b$$

$$= \frac{1}{3(b-a)}\left[b^3 - a^3\right] = \frac{(b-a)(b^2 + ab + a^2)}{3(b-a)}$$

$$= \frac{(b^2 + ab + a^2)}{3}$$

$$\therefore Variance = \mu_2^1 - (\mu_1^1)^2$$

$$= \frac{(b^2 + ab + a^2)}{3} - \left(\frac{b+a}{2}\right)^2$$

$$= \frac{(b^2 + ab + a^2)}{3} - \frac{(b^2 + 2ab + a^2)}{4}$$

$$= \frac{4(b^2 + ab + a^2) - 3(b^2 + ab + a^2)}{12}$$

$$= \frac{4b^2 + 4ab + 4a^2 - 3b^2 - 6ab - 3a^2}{12}$$

$$= \frac{(b^2 - 2ab + a^2)}{12} = \frac{(b-a)^2}{12}$$

$$\therefore \quad Variance = \frac{(b-a)^2}{12}$$

31. Define exponential distribution, Find the moment generating function of exponential distribution and also find its mean and variance.

Solution:

A continuous random variable 'X' is said to follow a exponential distribution with parameter $\lambda > 0$ if its probability density function is given by

$$f(x) = \begin{cases} \lambda e^{-\lambda x}, & x > 0 \\ 0, & otherwise \end{cases}$$

To find MGF:

$$M_X(t) = \int_0^\infty e^{tx} f(x)dx$$

$$= \int_0^\infty e^{tx} \lambda e^{-\lambda x} dx = \lambda \int_0^\infty e^{-(\lambda - t)x} dx$$

$$= \lambda \left[\frac{e^{-(\lambda - t)x}}{-(\lambda - t)} \right]_0^\infty = \frac{\lambda}{-(\lambda - t)} \left[e^{-\infty} - e^{-0} \right]$$

$$= \frac{\lambda}{(\lambda - t)}$$

$$\therefore \; M_X(t) = \frac{\lambda}{(\lambda - t)}$$

To find Mean and Variance:

$$M_X(t) = \frac{\lambda}{(\lambda - t)} = \frac{1}{1 - \frac{t}{\lambda}}$$

$$= \left(1 - \frac{t}{\lambda}\right)^{-1}$$

$$= 1 + \frac{t}{\lambda} + \frac{t^2}{\lambda^2} + \ldots\ldots\ldots + \frac{t^r}{\lambda^r} + \ldots\ldots\ldots$$

$$= 1 + \frac{t}{\lambda} + \left(\frac{t^2}{2!}\right)\left(\frac{2!}{\lambda^2}\right) + \ldots\ldots\ldots + \left(\frac{t^r}{r!}\right)\left(\frac{r!}{\lambda^r}\right) + \ldots\ldots\ldots$$

$$\therefore \; Mean \;\; \mu_1^1 = coefficient \;\; of \;\; \frac{t}{1!} = \frac{1}{\lambda}$$

$$\mu_2^1 = coefficient \;\; of \;\; \frac{t^2}{2!} = \frac{2}{\lambda^2}$$

$$Variance = \mu_2^1 - \left(\mu_1^1\right)^2$$

$$= \frac{2}{\lambda^2} - \frac{1}{\lambda^2} = \frac{1}{\lambda^2}$$

$$Variance = \frac{1}{\lambda^2}$$

GAMMA DISTRIBUTION

32. Define Gamma distribution, Find the MGF of Gamma distribution and also find mean and variance. [A.U N/D 2014, ’17 A/M ‘11,’13,’14,’18]

Solution:

The continuous random variable ‘X’ is said to follow a Gamma distribution with parameter λ if its probability density function is given by

$$f(x)=\begin{cases}\dfrac{e^{-x}x^{\lambda-1}}{\Gamma(\lambda)}, & \lambda>0, 0<x<\infty\\ 0, & otherwise\end{cases}$$

To find MGF:

$$M_X(t)=E\left(e^{tx}\right)$$
$$=\int_0^\infty e^{tx}f(x)dx$$
$$=\int_0^\infty e^{tx}\frac{e^{-x}x^{\lambda-1}}{\Gamma(\lambda)}dx$$
$$=\frac{1}{\Gamma(\lambda)}\int_0^\infty e^{tx}e^{-x}x^{\lambda-1}dx$$
$$=\frac{1}{\Gamma(\lambda)}\int_0^\infty e^{-(1-t)x}x^{\lambda-1}dx$$

$put\ \ (1-t)x=u \qquad If\ \ x=0,\ u=0$

$(1-t)dx=du \qquad If\ \ x=\infty,\ u=\infty$

$$=\frac{1}{\Gamma(\lambda)}\int_0^\infty e^{-u}\left(\frac{u}{1-t}\right)^{\lambda-1}\left(\frac{du}{1-t}\right)$$
$$=\frac{1}{\Gamma(\lambda)}\int_0^\infty \frac{u^{\lambda-1}e^{-u}}{(1-t)^{\lambda-1}(1-t)}du$$
$$=\frac{1}{\Gamma(\lambda)}\int_0^\infty \frac{u^{\lambda-1}e^{-u}}{(1-t)^{\lambda}}du$$
$$=\frac{1}{(1-t)^{\lambda}\Gamma(\lambda)}\int_0^\infty u^{\lambda-1}e^{-u}du$$
$$=\frac{1}{(1-t)^{\lambda}\Gamma(\lambda)}\Gamma(\lambda) \qquad \Gamma(n)=\int_0^\infty x^{n-1}e^{-x}dx$$
$$=\frac{1}{(1-t)^{\lambda}}$$
$$M_X(t)=(1-t)^{-\lambda},\ |t|<1$$

To find Mean and Variance:

$$M_X(t)=(1-t)^{-\lambda}$$
$$M_X^1(t)=-\lambda(1-t)^{-\lambda-1}(-1)$$
$$Mean=\mu_1^1=M_X^1(0)=\lambda$$

$$M_X^{11}(t) = \lambda(-\lambda-1)(1-t)^{-\lambda-2}(-1)$$
$$\mu_2^1 = M_X^{11}(0) = \lambda(\lambda+1)$$
$$Variance = \mu_2^1 - \left(\mu_1^1\right)^2$$
$$= \lambda(\lambda+1) - \lambda^2 = \lambda^2 + \lambda - \lambda^2$$
$$Variance = \lambda$$

NORMAL DISTRIBUTION

33. Define normal distribution and Find the moment generating function, mean and variance of normal distribution?

Solution: **[A.U N/D 2014, '17,'18]**

A normal distribution is a continuous distribution given by $y = \frac{1}{\sigma\sqrt{2\pi}} e^{\frac{-1}{2}\left[\frac{x-\mu}{\sigma}\right]^2}$ where X is a continuous normal variate distributed with density function $f(x) = \frac{1}{\sigma\sqrt{2\pi}} e^{\frac{-1}{2}\left[\frac{x-\mu}{\sigma}\right]^2}$ with mean μ and standard deviation σ

$$M_X(t) = E(e^{tX}) = \int_{-\infty}^{\infty} e^{tX} f(x)dx$$

$$= \int_{-\infty}^{\infty} e^{tX} \frac{1}{\sigma\sqrt{2\pi}} e^{\frac{-(x-\mu)^2}{2\sigma^2}} dx$$

$$= \frac{1}{\sigma\sqrt{2\pi}} \int_{-\infty}^{\infty} e^{tX} e^{\frac{-(x-\mu)^2}{2\sigma^2}} dx$$

Put $z = \frac{x-\mu}{\sigma}$, $\sigma\, dz = dx$

$$= \frac{1}{\sigma\sqrt{2\pi}} \int_{-\infty}^{\infty} e^{t(\sigma z+\mu)} e^{\frac{-(z)^2}{2}\sigma} dz$$

$$= \frac{e^{\mu t}}{\sqrt{2\pi}} \int_{-\infty}^{\infty} e^{\frac{-(z^2-2t\sigma z)}{2}} dz$$

$$= \frac{e^{\mu t}}{\sqrt{2\pi}} \int_{-\infty}^{\infty} e^{\frac{-1}{2}(z-\sigma t)^2+\frac{1}{2}(\sigma t)^2} dz$$

$$= \frac{e^{\mu t + \frac{t^2\sigma^2}{2}}}{\sqrt{2\pi}} \int_{-\infty}^{\infty} e^{\frac{-1}{2}(z-\sigma t)^2} dz$$

Put $u = z - \sigma t, \quad du = dz$

$$= \frac{e^{\mu t + \frac{t^2\sigma^2}{2}}}{\sqrt{2\pi}} \int_{-\infty}^{\infty} e^{\frac{-1}{2}(u)^2} du$$

$$= \frac{e^{\mu t + \frac{t^2\sigma^2}{2}}}{\sqrt{2\pi}} \sqrt{2\pi} \qquad \because \int_{-\infty}^{\infty} e^{\frac{-1}{2}(u)^2} du = \sqrt{2\pi}$$

$$= e^{\mu t + \frac{t^2\sigma^2}{2}}$$

To find mean:

$$E(X) = \left[\frac{d}{dt}(M_X(t))\right]_{t=0} = \frac{d}{dt}\left[e^{\mu t + \frac{t^2\sigma^2}{2}}\right]_{t=0}$$

$$= \left[e^{\mu t + \frac{t^2\sigma^2}{2}}\left(\mu + t\sigma^2\right)\right]_{t=0} = \mu$$

$$E(X^2) = \left[\frac{d^2}{dt^2}(M_X(t))\right]_{t=0} = \frac{d}{dt}\left[e^{\mu t + \frac{t^2\sigma^2}{2}}\left(\mu + t\sigma^2\right)\right]_{t=0}$$

$$= \left[\left(\mu + t\sigma^2\right)^2\left[e^{\mu t + \frac{t^2\sigma^2}{2}}\right] + e^{\mu t + \frac{t^2\sigma^2}{2}}\left(\sigma^2\right)\right]_{t=0} = \mu^2 + \sigma^2$$

$$Var(X) = E(X^2) - [E(X)]^2$$
$$= \mu^2 + \sigma^2 - \mu^2 = \sigma^2$$

34. State and Prove Memory less Property of Exponential Distribution. Using this property, Solve the following problem:

The length of the shower on a tropical island during rainy season has an exponential distribution with parameter 2, time being measured in minutes. What is the probability that a shower will last more than 3 minutes? [AU N/D 2016]

Solution:

If X is exponentially distributed then $P(X > S+t / X > S) = P(X > t), \;\; for \;\; any \;\; S, t > 0$

Proof:

$$P(X > K) = \int_K^\infty \lambda \, e^{-\lambda x} \, dx$$

$$P(X > K) = \lambda \left[-\frac{e^{-\lambda x}}{\lambda} \right]_k^\infty$$

$$= e^{-\lambda K} \qquad --------------- (1)$$

$$P(X > S+t / X > S) = \frac{P(X > S+t \;\; and \;\; X > S)}{P(X > S)}$$

$$= \frac{P(X > S+t)}{P(X > S)} = \frac{e^{-\lambda (S+t)}}{e^{-\lambda S}} \qquad by \;\; (1)$$

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

$$P(X > S+t / X > S) = P(X > t)$$

The converse of this result is also true.

i.e., If $P(X > S+t / X > S) = P(X > t)$ then X follows an exponential distribution.

$$Here \;\; \lambda = 2 \,, \Rightarrow \; P(X > 3) = e^{-3\lambda} = e^{-6} = 0.0025$$

35. State and prove the memory less property of the geometric distribution?

Solution:

If X has a geometric distribution, then for any two positive integers ‘s’ and ‘t’,

$$P[X > s+t / X > s] = P(X > t)$$

Proof:

Consider $P[X > s+t] = \sum_{x \,=\, s+t+1}^{\infty} p\, q^{x-1}$

$$= p\left[q^{s+t+1-1} + q^{s+t+2-1} + ...\right]$$

$$= p\left[q^{s+t} + q^{s+t+1} + ...\right]$$

$$= p \; q^{s+t}\left[1 + q + .q^2 + ...\right]$$

$$= p \; q^{s+t}\left[1 - q\right]^{-1}$$

$$= \frac{p\, q^{s+t}}{1-q} \quad ---------- \quad (1)$$

$$\therefore\ P(X > s) = \frac{p\, q^{s}}{1-q} ------------- \quad (2)$$

$$and \qquad P(X > t) = \frac{p\, q^{t}}{1-q} \quad --------- \quad (3)$$

Hence

$$\therefore\ P(X > s+t / X > s) = \frac{P(X > s+t \quad and \quad X > s)}{P(X > s)}$$

$$= \frac{P(X > s+t)}{P(X > s)}$$

$$= \frac{p\, q^{s+t}}{1-q} \times \frac{1-q}{pq^{s}} \qquad \text{by (1) and (2)}$$

$$= \frac{q^{s+t}}{q^{s}}$$

$$= q^{t}$$

$$= \frac{p\, q^{t}}{1-q} \qquad \because \frac{p}{1-q} = \frac{p}{p} = 1$$

$$= P(X > t) \qquad \text{by (3)}$$

36. Derive probability mass function of Poisson distribution as a limiting case of binomial distribution. **[AU N/D '14,'13 RP]**

Solution:

(i) The number of trails 'n' should be indefinitely large. [i.e., n tends to infinity]

(ii) The probability of successes 'p' for each trail is indefinitely small.

$(iii)\ np = \lambda$, *should be finite where* λ *is a constant*

We know that the binomial distribution is

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

$$= \frac{n!}{(n-x)!x!} p^{x} (1-p)^{n-x} \qquad [\because q = 1-p]$$

$$= \frac{1.2.3........(n-x)(n-x+1)..........n}{1.2.3............(n-x)x!} p^{x} \frac{(1-p)^{n}}{(1-p)^{x}}$$

$$= \frac{1.2.3........(n-x)(n-x+1)..........n}{1.2.3............(n-x)x!} \left(\frac{p}{(1-p)}\right)^{x} (1-p)^{n}$$

$$= \frac{(n-1)(n-2)........(n-x+1)}{x!}\left(\frac{\frac{\lambda}{n}}{1-\frac{\lambda}{n}}\right)^x\left(1-\frac{\lambda}{n}\right)^n \qquad \left[\because p = \frac{\lambda}{n}\right]$$

$$= \frac{(n-1)(n-2)........(n-x+1)}{x!}\frac{\lambda^x}{n^x}\frac{1}{\left(1-\frac{\lambda}{n}\right)^x}\left(1-\frac{\lambda}{n}\right)^n$$

$$= \frac{(n-1)(n-2)........(n-x+1)}{x!}\frac{\lambda^x}{n^x}\left(1-\frac{\lambda}{n}\right)^{n-x}$$

$$= \frac{1\left(1-\frac{1}{n}\right)\left(1-\frac{2}{n}\right)...............\left\{1-\left(\frac{x-1}{n}\right)\right\}}{x!}\lambda^x\left(1-\frac{\lambda}{n}\right)^{n-x}$$

When $n \to \infty$

we know that $\underset{n\to\infty}{Lt}\left(1-\frac{\lambda}{n}\right)^{n-x} = e^{-\lambda}$

$$\underset{n\to\infty}{Lt}\left(1-\frac{1}{n}\right) = \underset{n\to\infty}{Lt}\left(1-\frac{2}{n}\right) = ... = \underset{n\to\infty}{Lt}\left(1-\left(\frac{x-1}{n}\right)\right) = 1$$

$$\therefore P(X=x) = \frac{e^{-\lambda}\lambda^x}{x!}, \qquad x = 0,1,2,..........,\infty$$

Hence the probability mass function of a random variable 'X' which follows Poisson distribution is given by

$$P(X=x) = p(x) = \begin{cases} \frac{e^{-\lambda}\lambda^x}{x!}, & x = 0,1,2,..........,\infty \\ 0, & otherwise \end{cases}$$

PROBLEMS BASED ON DISTRIBUTIONS

37. The time (in hours) required to repairs a machine is exponential distributed with Parameter $\lambda = \frac{1}{2}$. **[AU N/D 2009, 2024]**

(i) What is the probability that the repair time exceeds 2 2hours?

(ii) What is the conditional probability that a repair takes at least 10h given that its duration exceeds 9h?

Solution:

If X represents the time to repair the machine , the density function of X is given by

$$f(x) = \lambda e^{-\lambda x} = \frac{1}{2}e^{\frac{-x}{2}}, \quad x > 0$$

$$(i)\ p(X > 2) = \int_2^\infty \frac{1}{2}e^{\frac{-x}{2}}dx = \frac{1}{2}\int_2^\infty e^{\frac{-x}{2}}dx = \frac{1}{2}\left[\frac{e^{\frac{-x}{2}}}{\frac{-1}{2}}\right]_2^\infty$$

$$= \left[-e^{\frac{-x}{2}}\right]_2^\infty = e^{-1} = 0.3679$$

$$(ii)\ p(X \geq 10 / X > 9) = p(X > 1) \quad (by\ the\ \ memoryless\ \ property)$$

$$= \int_1^\infty \frac{1}{2}e^{\frac{-x}{2}}dx = \frac{1}{2}\int_1^\infty e^{\frac{-x}{2}}dx = \frac{1}{2}\left[\frac{e^{\frac{-x}{2}}}{\frac{-1}{2}}\right]_1^\infty$$

$$= \left[-e^{\frac{-x}{2}}\right]_1^\infty = e^{-0.5} = 0.6065$$

38. Suppose that telephone calls arriving at a particular switchboard follow a Poisson process with an average of 5 calls coming per minute . What is the probability that up to a minute will elapse unit 2 calls have come into the switch board? [AU A/M 2011]

Solution:

The Poisson process applies with time until 2 Poisson events following a Gamma distribution with $\beta = \frac{1}{5}$ $and\ \ \alpha = 2$.

Let the random variable X be the time in minutes that transpires before 2 calls come.

$$P[x \leq 1] = 25\int_0^1 xe^{-5x}dx$$

$$= 25\left\{x\left[\frac{e^{-5x}}{-5}\right] - (1)\left[\frac{e^{-5x}}{25}\right]\right\}_0^1 = 25\left[-\frac{x}{5}e^{-5x} - \frac{1}{25}e^{-5x}\right]_0^1$$

$$= 25\left[\left(-\frac{1}{5}e^{-5} - \frac{1}{25}e^{-5}\right) - \left(-0 - \frac{1}{25}\right)\right] = 25\left[-\frac{6}{25}e^{-5} + \frac{1}{25}\right]$$

$$= 1 - 6e^{-5} = 0.96$$

39. If X is a Poisson variate such that $P(X = 2) = 9P(X = 4) + 90P(X = 6)$ Find Mean, $E(X^2)\ and\ \ Variance$

Solution:

The probability distribution for the Poisson R.V X is given by

$$P(X=x)=\frac{e^{-\lambda}\lambda^{x}}{x!},\ x=0,1,2,......n\ and\ \lambda>0$$

$$P(X=2)=9P(X=4)+90P(X=6)$$

$$\frac{e^{-\lambda}\lambda^{2}}{2!}=9\frac{e^{-\lambda}\lambda^{4}}{4!}+90\frac{e^{-\lambda}\lambda^{6}}{6!}$$

$\div$ *by* $\lambda^{2}e^{-\lambda}$

$$\frac{e^{-\lambda}\lambda^{2}}{2!}=9\frac{e^{-\lambda}\lambda^{4}}{4!}+90\frac{e^{-\lambda}\lambda^{6}}{6!}$$

$$\frac{9\lambda^{2}}{1.2.3.4}+\frac{90\lambda^{4}}{1.2.3.4.5.6}=\frac{1}{2}$$

$$\frac{3\lambda^{2}}{4}+\frac{\lambda^{4}}{4}=1\Rightarrow\lambda^{4}+3\lambda^{2}-4=0$$

$$(\lambda^{2}+4)(\lambda^{2}-1)=0$$

Result: $\lambda^{2}=-4\ (or)\ \lambda^{2}=1\Rightarrow\lambda=\pm1$

40. If X is a normal variate with mean 30 and standard deviation 5. Find the probabilities that

(i) $26\le X\le 40,$ $(ii)\ X\ge 45,$ $(iii)\left|X-30\right|>5$ **[A.U. 2009]**

Solution:

Given $\mu=30,\quad \sigma=5$

WKT, $Z=\dfrac{X-\mu}{\sigma}$

(i) When $X=26,\ z=\dfrac{26-30}{5}=0.8$ and

$X=40,\ Z=\dfrac{40-30}{5}=2$

$\therefore P(26\le X\le 40)=P(0.8\le Z\le 2)$

$=P(0.8\le Z\le 0)+P(0\le Z\le 2)$

$=P(0\le Z\le 0.8)+P(0\le Z\le 2)$

=0.2881+0.4772

=0.7653

(ii) When $X=45,\ z=\dfrac{45-30}{5}=3$

$\therefore P(X\ge 45)=P(Z\ge 3)$

$=0.5-P(0\le Z\le 3)$

= 0.5-0.4987=0.0013

$(iii)P(\left|X-30\right|\le 5)=P(25\le X\le 35)$

When $X = 25, \quad z = \dfrac{25-30}{5} = -1$ and

$$X = 35, \quad Z = \frac{35-30}{5} = 1$$

$$P(-1 \le Z \le 1) = 2P(0 \le Z \le 1)$$
$$= 2(0.3413) = 0.6826$$
$$P(|X-30| > 5) = 1 - P(|X-30| \le 5)$$
$$=1-0.6826=0.3174$$

41. Suppose that a trainee soldier shoots a target in an independent fashion. If the probability that the target is shot on any one shot is 0.7,
1) What is the probability that the target would be hit on tenth attempt?
2) What is the probability that it takes him less than 4 shots?
3) What is the probability that it takes him an even number of shots? [A.U. N/D 2014]
Solution:

Geometric distribution:

$$P(X = x) = q^{n-1} p$$

Here $p = 0.7$ *and* $q = 0.3$

1) $$P(X = 10) = q^{10-1} p$$
$$= (0.3)^9 (0.7)$$
$$= 0.000014$$

2) $$P(X < 4) = p(x=1) + p(x=2) + p(x=3)$$
$$= p + q' p + q^2 p$$
$$= (0.7) + (0.3)(0.7) + (0.3)^2 (0.7))$$
$$= (0.7) + (0.21) + 0.063$$
$$= 0.973$$

3) $$P(X = Even) = q' p + q^3 p + q^5 p + \ldots$$
$$= pq[1 + q^2 + q^4 + \ldots]$$
$$= pq[1 - q^2]^{-1}$$
$$= (0.7)(0.3)[1 - (0.3)^2]^{-1}$$
$$= (0.21)[1 - 0.09]^{-1}$$
$$= (0.21)[0.91]^{-1}$$
$$= \frac{(0.21)}{[0.91]} = 0.2308$$

42. Trains arrive at a station at 15 minutes intervals starting at 4a.m.If a passenger arrive at a station at a time that uniformly distributed between 9.00 and 9.30, find the probability that he has to wait for the train for (1) less than 6 minutes (2) more than 10 minutes.

Solution: **[A.U. N/D 2014 P.Q.T]**

Given that f(x) is a Uniformly distributed in (0,30)

$$f(x)=\frac{1}{b-a} \quad 0<x<30$$

$$f(x)=\frac{1}{30-0} \quad \Rightarrow f(x)=\frac{1}{30}$$

1) $p(x<6)$ *Starting time* $x=4$

$$p(4<x<6)=\int_4^6 \frac{1}{30}dx = \frac{1}{30}[x]_4^6 = \frac{1}{30}[6-4] = \frac{1}{30}[2] = \frac{1}{15}$$

2) $p(x>10)=p(10<x<30)$

$$=\int_{10}^{30}\left(\frac{1}{30}\right)dx = \frac{1}{30}[x]_{10}^{30} = \frac{1}{30}[30-10] = \frac{1}{30}[20] = \frac{2}{3}$$

43. The marks obtained by a number of students for a certain subject is assumed to be normally distributed with mean 65 and S.D of 5.If 3 students are taken at random from this set, what is the probability that exactly 2 of them will have marks over 70. **[A.U A/M 2010]**

Solution:

Let X denote marks obtained by the given set of students

Given $\mu = 65, \quad \sigma = 5$

WKT, $Z = \dfrac{X-\mu}{\sigma}$

(i)

When $X = 70, \quad z = \dfrac{70-65}{5} = 1$

$\therefore P(X \geq 70) = P(0 < Z < \infty)$

$= 0.5 - P(0 \leq Z \leq 1)$

=0.5-0.0.3413=0.1587

44. Assume that the reduction of a person's oxygen consumption during a period of Transcendental Meditation(TM) is a continuous random variable X normally distributed with mean 37.6 cc/mm and S.D 4.6 cc/min. Determine the probability that during a period of T.M a person's oxygen consumption will be reduced by (i) at least 44.5 cc/min (ii) at most 35.0 cc/min (iii) Anywhere from 30.0 to 40.0 cc/min **[A.U. N/D 2012]**

Solution:

Given $\mu = 37.6, \ \sigma = 4.6, \ Z = \dfrac{X-\mu}{\sigma}$

$= \dfrac{X-37.6}{4.6}$

(i) For atleast 44.5 cc/min

(i.e)., $X = 44.5$

$$\therefore \quad Z = \frac{44.5 - 37.6}{4.6} = 1.5$$

$$P[X \geq 44.5] = P[Z \geq 1.5]$$
$$= 0.5 - P\ [Z < 1.5]$$
$$= 0.5 - 0.4332$$
$$= 0.068$$

(ii) For atmost 35.0 cc/min

(i.e)., $X = 35$

$$\therefore \quad Z = \frac{35 - 37.6}{4.6} = -0.5652$$

$$P[X \leq 35] = P[-Z \leq -0.5652]$$
$$= 0.5 - 0.2157$$
$$= 0.2843$$

(iii) Anywhere from 30.0 to 40.0 cc/min

(i.e)., $X_1 = 30 \qquad \therefore \quad Z_1 = \frac{30 - 37.6}{4.6} = -1.6521$

$X_2 = 40 \qquad \therefore \quad Z_2 = \frac{40 - 37.6}{4.6} = 0.52173$

$$P[30 \leq X \leq 40] = P[-1.6521 \leq Z \leq -0.52173]$$
$$= 0.4505 + 0.1985$$
$$= 0.6490$$
$$\cong 1$$

45. Let X and Y be independent normal variates with mean 45 and 44 and standard deviation 2 and 1.5 respectively. What is the probability that randomly chosen values of X and Y differ by 1.5 or more? [A.U N/D 2011]

Solution:

Let X is $N(45,2)$ and Y is $N(44,1.5)$)

Therefore by property of additive $U = X - Y$ follows the distribution $N\left(1, \sqrt{4 + 2.25}\right)$

$$i.e., \quad N(1,2.5)$$

$$P[X \text{ and } Y \text{ differ by } 1.5 \text{ or more}] = P\left[\,|X - Y| \geq 1.5\right] = P\left[\,|U| \geq 1.5\right]$$
$$= 1 - P\left[\,|U| \leq 1.5\right]$$
$$= 1 - P[-1.5 \leq U \leq 1.5]$$
$$= 1 - P\left[\frac{-1.5 - 1}{2.5} \leq \frac{U - 1}{2.5} \leq \frac{1.5 - 1}{2.5}\right]$$
$$= 1 - P[-1 \leq Z < 0.2]$$
$$= 1 - [0.3413 + 0.0793]$$
$$= 0.5794$$

46. If X is a uniform random variable in the interval (-2,2) find the probability density function $Y = |X| \text{ and } E[Y]$ **[A.U N/D 2011]**

Solution:

$X \sim U[-2,2]$

The pdf $f(x) = \begin{cases} \frac{1}{b-a}; -2 \le x \le 2 \\ 0 \quad ; \; otherwise \end{cases}$

$$f(x) = \begin{cases} \frac{1}{4}; -2 \le x \le 2 \\ 0 \quad ; \; otherwise \end{cases}$$

$Given \quad Y = |X|$

C.D.F of $G(Y) = P[Y \le y] = P[|x| \le y]$

$$= F[y] - F(-y)$$

$$= \int_{-2}^{y} \frac{1}{4} dx - \int_{-2}^{-y} \frac{1}{4} dx$$

$$= \frac{1}{4}[x]_{-2}^{y} - \frac{1}{4}[x]_{-2}^{-y}$$

$$= \frac{1}{4}[y+2] - \frac{1}{4}[-y+2]$$

$$= \frac{1}{2} y$$

$$p.d.f. \; g(y) = \frac{dG(y)}{dy} = \begin{cases} \frac{1}{2}; 0 \le y \le 2 \\ 0 \; ; \; otherwise \end{cases}$$

$$E(Y) = \int_{0}^{2} y \frac{1}{2} dy$$

$$= \frac{1}{2}\left[\frac{y^2}{2}\right]_0^2 = \frac{4}{4} = 1$$

47. If the probability that an applicant for a driver's license will pass the road test on any given trial is 0.8. what is the probability that he will finally pass the test (1) on the fourth trial and (2) In less than 4 trials? [A.U A/M 2010]

Solution:

Let X denotes the number of trials required to achieve the first success. Then X is a geometric distribution given by $P(X = r) = q^{r-1} p; \; r = 1,2,3,...\infty$

Here $p = 0.8 \; \& \; q = 0.2$

(1) $P(X = 4) = 0.8(0.2)^{4-1} = 0.8(0.008) = 0.0064$

(2) $P(X < 4) = \sum_{r=1}^{3} 0.8(0.2)^{r-1} = 0.8(1 + 0.2 + 0.04) = 0.992$

48. Find the MGF of the two parameter exponential distribution whose density function is given by $f(x) = \lambda e^{-\lambda(x-a)}$, $x \geq a$ and hence find the mean and variance?

Solution: **[A.U A/M 2010]**

Given:

$$f(x) = \lambda e^{-\lambda(x-a)} \ , \ x \geq a$$

To find the MGF:

$$M_X(t) = \int_{-\infty}^{\infty} e^{tx} f(x)\, dx$$

$$M_X(t) = \int_a^{\infty} e^{tx} \lambda\, e^{-\lambda(x-a)}\, dx$$

$$= \lambda\, e^{a\lambda} \int_a^{\infty} e^{tx} \ e^{-\lambda x}\, dx$$

$$= \lambda\, e^{a\lambda} \int_a^{\infty} e^{-(\lambda - t)x}\, dx$$

$$= \lambda\, e^{a\lambda} \left[\frac{e^{-(\lambda-t)x}}{-(\lambda - t)} \right]_a^{\infty}$$

$$= \lambda\, e^{a\lambda} \left[0 - \frac{e^{-(\lambda-t)\,a}}{-(\lambda-t)} \right] = \frac{\lambda\, e^{at}}{\lambda - t}$$

$$Mean \ \ E(X) = \left[\frac{d}{dt} [M_X(t)] \right]_{t=0}$$

$$= \left[\frac{d}{dt} \left[\frac{\lambda\, e^{at}}{\lambda - t} \right] \right]_{t=0}$$

$$= \left[\left(\frac{\lambda}{\lambda - t} \right) a e^{at} + e^{at} \left(\frac{-\lambda}{(\lambda - t)^2} \right)(-1) \right]_{t=0}$$

$$= \left(\frac{\lambda a e^{at}}{\lambda - t} + \frac{\lambda e^{at}}{(\lambda - t)^2} \right)_{t=0} = \left(\frac{\lambda a}{\lambda} + \frac{\lambda}{(\lambda)^2} \right) = a + \frac{1}{\lambda}$$

$$E(X^2) = \left[\frac{d}{dt^2}^2 [M_X(t)] \right]_{t=0}$$

$$= \frac{d}{dt} \left[\left(\frac{\lambda}{\lambda - t} \right) a e^{at} + e^{at} \left(\frac{\lambda}{(\lambda - t)^2} \right) \right]_{t=0}$$

$$= \left[\left(\frac{\lambda}{\lambda - t} \right) a e^{at}(a) + e^{at} \left(\frac{-\lambda a}{(\lambda - t)^2} \right)(-1) + a e^{at} \left(\frac{\lambda}{(\lambda - t)^2} \right) + e^{at} \left(\frac{-2\lambda}{(\lambda - t)^3} \right)(-1) \right]_{t=0}$$

$$= \left[\left(\frac{\lambda}{\lambda} \right) a^2 + \left(\frac{\lambda a}{(\lambda)^2} \right) + a \left(\frac{\lambda}{(\lambda)^2} \right) + \left(\frac{2\lambda}{(\lambda)^3} \right) \right]$$

$$= \left[a^2 + \frac{a}{\lambda} + \frac{a}{\lambda} + \frac{2}{\lambda^2} \right]$$

$$Var(X) = E(X^2) - [E(X)]^2$$

$$= a^2 + \frac{2a}{\lambda} + \frac{2}{\lambda^2} - \left[a + \frac{1}{\lambda} \right]^2$$

$$= a^2 + \frac{2a}{\lambda} + \frac{2}{\lambda^2} - \left[a^2 + \frac{1}{\lambda^2} + \frac{2a}{\lambda} \right]$$

$$= \frac{1}{\lambda^2}$$

49. The marks obtained by a number of students in a certain subject are assumed to be normally distributed with mean 65 and standard deviation 5. If 3 students are selected at random from this group, what is the probability that two of them will have marks over 70? [A.U A/M 2010]

Solution:

Given $\mu = 65,\ \sigma = 5,\ Z = \dfrac{X-\mu}{\sigma}$

$$= \frac{X-65}{5}$$

When $X = 70$

$$\therefore\ Z = \frac{70-65}{5} = 1$$

$$P[X > 70] = P[1 < Z < \infty]$$

$$= 0.5 - P\ [0 < Z < 1]$$

$$= 0.5 - 0.3413$$

$$= 0.1587 \quad N = 3,\ P = 0.1587,\ q = 1 - p = 0.8413 \quad P[Y = 2] = 3C_2 p^2 q$$

Result: $= 3(0.1587)^2(0.8413) = 0.0636$

50. The CDF of the random variable of X is given by

$$F_x(x) = \begin{cases} 0 & ;\ x < 0 \\ x + \frac{1}{2} & ;\ 0 \le x \le \frac{1}{2} \\ 1 & ;\ x > \frac{1}{2} \end{cases}$$

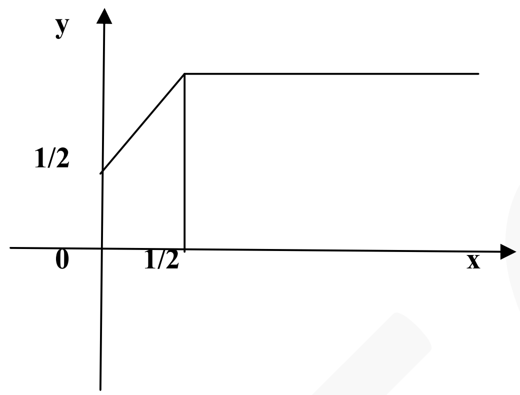
Draw the graph of the CDF. Compute $P\left(X > \frac{1}{4}\right),\ P\left(\frac{1}{3} < X < \frac{1}{2}\right)$ **[AU A/M '15]**

Solution:

Given that $F_x(x) = \begin{cases} 0 & ;\ x < 0 \\ x + \frac{1}{2} & ;\ 0 \le x \le \frac{1}{2} \\ 1 & ;\ x > \frac{1}{2} \end{cases}$

To Draw the graph of the CDF

$$f(x)=\begin{cases}1\ , & 0\le x\le \frac{1}{2}\\ 0\ , & x>\frac{1}{2}\end{cases}$$



To Find $P\left(X>\frac{1}{4}\right),\ P\left(\frac{1}{3}<X<\frac{1}{2}\right)$:

$$P\left(X>\frac{1}{4}\right)=\int_{\frac{1}{4}}^{\frac{1}{2}}1dx=[x]_{\frac{1}{4}}^{\frac{1}{2}}=\frac{1}{2}-\frac{1}{4}=\frac{1}{4}$$

$$P\left(\frac{1}{3}<X<\frac{1}{2}\right)=\int_{\frac{1}{3}}^{\frac{1}{2}}1dx=[x]_{\frac{1}{3}}^{\frac{1}{2}}=\frac{1}{2}-\frac{1}{3}=\frac{1}{6}$$

51. Messages arrive at a switch board in a Poisson manner at an average rate of six per hour. Find the probability for each of the following events: [AU A/M 15, N/D 16,'17]

(1) exactly two messages arrive within one hour

(2) no message arrives within one hour

(3) at least three messages arrive within one hour.

Solution:

Given that $\lambda=6$

Let X denotes the no. of messages arrive within a hour.

Therefore the probability distribution is $P(X=x)=\dfrac{e^{-6}6^x}{x!}$

(1) P (exactly two messages arrive within 1 hr) $=P(X=2)$

$$=\frac{e^{-6}6^2}{2!}$$

$$=\frac{e^{-6}(36)}{2}$$

$$=18e^{-6}=0.0446$$

(2) P (no messages arrive within 1 hr) $=P(X=0)$

$$= \frac{e^{-6}6^0}{0!}$$
$$= e^{-6}$$
$$= 0.0024$$

(3) P (atleast three messages arrive within 1 hr) $= P(X \geq 3)$

$$= 1 - P(X < 3)$$
$$= 1 - [P(X = 0) + P(X = 1) + P(X = 2)]$$
$$= 1 - \left[0.0024 + \frac{e^{-6}(6)}{1} + 0.0446\right]$$
$$= 1 - 0.0024 - 0.0148 - 0.0446$$
$$= 1 - 0.0618$$
$$= 0.9382$$

52. The peak temperature T, as measured in degrees Fahrenheit, on a particular day is the Gaussian (85,10) random variable. What is P(T>100),P(T<60) and P(70<T<100)?

Solution: **[AU A/M '15]**

We know that $Z = \frac{T - \mu}{\sigma}$

Given that $\mu = 85, \sigma = 10$

When T=100, $Z = \frac{100 - 85}{10} = 1.5$

$\therefore \quad P(T > 100) = P(z > 1.5)$

$$= 0.5 - P(0 < z < 1.5)$$
$$= 0.5 - 0..4332$$
$$= 0.0668$$

When T=60, $Z = \frac{60 - 85}{10} = -2.5$

$\therefore \quad P(T < 60) = P(z < -2.5)$

$$= 0.5 - P(0 < z < -2.5)$$
$$= 0.5 - 0..4938$$
$$= 0.0062$$

When T=70, $Z = \frac{70 - 85}{10} = -1.5$

When T=100, $Z = \frac{100 - 85}{10} = 1.5$

$\therefore \quad P(70 < T < 100) = P(-1.5. < z < 1.5)$

$$= 2P(0 < z < 1.5)$$
$$= 2(0.4332)$$
$$= 0.8664$$

53. A component has an exponential time to failure distribution with mean of 10,000 hours.
(1) The component has already been in operation for its mean life. What is the probability that it will fail by 15,000 hours?
(2) At 15,000 hours the component is still in operation. What is the probability that it will operate for another 5000 hours?

Solution: **[AU N/D 2015]**

Let a random variable X denote the exponential time of failure rate, it follows an exponential distribution with mean

$$\frac{1}{\lambda} = 10,000 \ hours \Rightarrow \lambda = \frac{1}{10,000}$$

$$\therefore \ The \ pdf \ of \ X \ is \ f(x) = \begin{cases} \frac{1}{10,000} e^{-x/10,000}; \ x \geq 0 \\ 0 \quad ; \ otherwise \end{cases}$$

$$(1) \quad P(X < 15,000 / X > 10,000) = \frac{P[X < 15000 \cap X > 10000]}{P[X > 10000]}$$

$$= \frac{P[10000 < X < 15000]}{P[X > 10000]} \ \ldots\ldots\ldots\ldots\ldots\ldots\ldots(1)$$

$$P[10000 < X < 15000] = \int_{10000}^{15000} f(x)dx = \int_{10000}^{15000} \frac{1}{10000} e^{\frac{-x}{10000}} dx$$

$$= \frac{1}{10000} \left[\frac{e^{\frac{-x}{10000}}}{\left(\frac{-1}{10000}\right)} \right]_{10000}^{15000} = -\left[e^{\frac{-x}{10000}} \right]_{10000}^{15000}$$

$$= -\left[e^{-\frac{3}{2}} - e^{-1} \right] = e^{-1} - e^{-\frac{3}{2}}$$

$$P[X > 15000] = \int_{10000}^{\infty} f(x)dx = \int_{10000}^{\infty} \frac{1}{10000} e^{\frac{-x}{10000}} dx$$

$$= \frac{1}{10000} \left[\frac{e^{\frac{-x}{10000}}}{\left(\frac{-1}{10000}\right)} \right]_{10000}^{\infty} = -\left[e^{\frac{-x}{10000}} \right]_{10000}^{\infty}$$

$$= -\left[0 - e^{-1}\right] = e^{-1}$$

$$(1) \Rightarrow \quad = \frac{e^{-1} - e^{-\frac{3}{2}}}{e^{-1}} = 1 - e^{-\frac{3}{2}} e^{1} = 1 - e^{-\frac{1}{2}} = 0.3935$$

$$(2) \quad P(X > 15,000 + 5000 \ / \ X > 15,000) = P[X > 5000]$$

$$= \int_{5000}^{\infty} f(x)dx = \int_{5000}^{\infty} \frac{1}{10000} e^{\frac{-x}{10000}} dx$$

$$= \frac{1}{10000}\left[\frac{e^{\frac{-x}{10000}}}{\left(\frac{-1}{10000}\right)}\right]_{5000}^{\infty} = -\left[e^{\frac{-x}{10000}}\right]_{5000}^{\infty}$$

$$= 0 + e^{-\frac{1}{2}} = e^{-0.5} = 0.6065$$

54. The average percentage of marks of candidates in an examination is 42 with a standard deviation of 10. If the minimum mark for pass is 50% and 1000 candidates appear for the examination, how many candidates can be expected to get the pass mark follow normal distribution? If it is required, that double the number of the candidates should pass, what should be the minimum mark for pass?

Solution: **[AU N/D 2016]**

Let X denote the marks of the candidates.

$$Let \ \ Z = \frac{X-42}{10} \qquad \left[\because \ \bar{x} = 42, \ \sigma = 10\right]$$

$$P\left[X \geq 50\right] = P\left[Z \geq 0.8\right]$$

$$= 0.5 - P\left[0 < Z \geq 0.8\right]$$

$$= 0.5 - 0.2881 = 0.2119$$

If 1000 students write the test.

$$1000\,P\left[X \geq 50\right] = (1000)(0.2119) = 212\,app$$

Candidates would pass the examination.

If double that number should pass, then the number of pass should be 424.

$$Find \ \ Z, \ \ s.t \ \ P[Z \geq Z_1] = 0.424$$

$$\therefore \ \ P\left[0 < Z \geq Z_1\right] = 0.5 - 0.424 = 0.076$$

$$From \ \ the \ \ table, \ \ we \ \ get \ \ Z_1 = 0.19$$

$$Z_1 = \frac{50 - x_1}{10} \qquad \Rightarrow x_1 = 50 - 10Z_1$$

$$= 50 - 10(0.19) = 50 - 1.9 = 48.1$$

The pass mark should be 48 nearly.

55. A continuous random variable X has the pdf $f(x) = kx^3 e^{-x}, \ x \geq 0$

Find the rth order moment about the origin, moment generating function, mean and variance of X.

Solution: **[AU N/D 2015]**

We know that, $\int_{-\infty}^{\infty} f(x)dx = 1$

Here, $\int_{0}^{\infty} kx^3 e^{-x} dx = 1 \qquad [\because \ x > 0]$

$$k\left[x^3\left(\frac{e^{-x}}{-1}\right) - 3x^2\left(\frac{e^{-x}}{(-1)^2}\right) + 6x\left(\frac{e^{-x}}{(-1)^3}\right) - 6\left(\frac{e^{-x}}{(-1)^4}\right)\right]_0^{\infty} = 1$$

$$k\left[-x^3e^{-x}-3x^2e^{-x}-6xe^{-x}-6e^{-x}\right]_0^\infty=1$$
$$k\left[0-(-6e^0)\right]=1$$
$$6k=1$$
$$k=\frac{1}{6}$$

$$f(x)=\frac{1}{6}x^3e^{-x},\ \ x\ge 0$$

$$M_X(t)=E\left[e^{tX}\right]$$
$$=\int_0^\infty e^{tx}f(x)dx=\int_0^\infty e^{tx}\frac{1}{6}x^3e^{-x}dx=\frac{1}{6}\int_0^\infty x^3e^{-(1-t)x}dx$$
$$=\frac{1}{6}\left[x^3\left(\frac{e^{-(1-t)x}}{-(1-t)}\right)-3x^2\left(\frac{e^{-(1-t)x}}{(1-t)^2}\right)+6x\left(\frac{e^{-(1-t)x}}{-(1-t)^3}\right)-6\left(\frac{e^{-(1-t)x}}{(1-t)^4}\right)\right]_0^\infty$$
$$=\frac{1}{6}\left[6\left(\frac{1}{(1-t)^4}\right)\right]=\frac{1}{(1-t)^4}$$

$$\mu_r^1=E\left[x^r\right]=\int_{-\infty}^\infty x^rf(x)dx=\int_0^\infty x^r\frac{1}{6}x^3e^{-x}dx=\frac{1}{6}\int_0^\infty x^{r+3}e^{-x}dx$$
$$=\frac{1}{6}\int_0^\infty e^{-x}x^{(r+3+1)-1}dx=\frac{1}{6}\Gamma_{(r+4)}=\frac{(r+3)!}{6}$$

$$put\quad r=1,\quad \mu_1^1=\frac{(1+3)!}{6}=\frac{4!}{6}=4$$
$$r=2,\quad \mu_2^1=\frac{(2+3)!}{6}=\frac{5!}{6}=20$$
$$Var(X)=E\left[X^2\right]-\left[E(X)\right]^2=20-(4)^2=4$$

56. Let X be a continuous R.V with probability density $f(x)=\begin{cases}xe^{-x}, & x>0\\ 0, & otherwise\end{cases}$ **, Find**

(1) the cumulative distribution function of X
(2) Moment Generating Function of X
(3) P(X<2)
(4) E(X)

Solution: **[AU M/J 2016]**

The probability density function is

$$f(x)=\begin{cases}xe^{-x}, & x>0\\ 0, & otherwise\end{cases}$$

(1) To find cdf:

$$F(x)=\int_0^x te^{-t}dt=\left[t\left(\frac{e^{-t}}{-1}\right)-(1)\left(\frac{e^{-t}}{(-1)^2}\right)\right]_0^x$$
$$=1-(1+x)e^{-x}\quad x>0$$

(2) To find MGF:

$$M_X(t) = E\left[e^{tX}\right]$$

$$= \int_0^\infty e^{tx} f(x)dx = \int_0^\infty e^{tx} te^{-t} dx = \int_0^\infty te^{-(1-t)x} dx$$

$$= \left[t\left(\frac{e^{-(1-t)x}}{-(1-t)}\right) - (1)\left(\frac{e^{-(1-t)x}}{(1-t)^2}\right)\right]_0^\infty = \frac{1}{(1-t)^2},\ if\ |t| < 1$$

(3) $P(X < 2) = F(2) = 1 - 3e^{-2}$

(4) $E(X) = \int_0^\infty x^2 e^{-x} dx = \left[x^2\left(\frac{e^{-x}}{-1}\right) - 2x\left(\frac{e^{-x}}{1}\right) + 2\left(\frac{e^{-x}}{-1}\right)\right]_0^\infty = 2$

57. Let $P(X = x) = \left(\frac{3}{4}\right)\left(\frac{1}{4}\right)^{x-1}$, $x = 1,2,3.....$ **be the probability mass function of a R.V X**

Compute

(1) $P(X > 4),\ P(X > 5)$

(2) $P(X > 4 / X > 2)$

(3) $MGF,\ E(X)$

(4) $Var(X)$

Solution: **[AU M/J 2016, 2021]**

Given that probability mass function $P(X = x) = \left(\frac{3}{4}\right)\left(\frac{1}{4}\right)^{x-1}$, $x = 1,2,3.....$

(1) $P(X > 4) = \sum_{x=5}^{\infty}\left(\frac{3}{4}\right)\left(\frac{1}{4}\right)^{x-1} = \left(\frac{3}{4}\right)\left(\frac{1}{4}\right)^4 + \left(\frac{3}{4}\right)\left(\frac{1}{4}\right)^5 + \left(\frac{3}{4}\right)\left(\frac{1}{4}\right)^6 + ...$

$$= \left(\frac{3}{4}\right)\left(\frac{1}{4}\right)^4\left[1 + \left(\frac{1}{4}\right) + \left(\frac{1}{4}\right)^2 + ...\right]$$

$$= \left(\frac{3}{4}\right)\left(\frac{1}{4}\right)^4\left[1 - \frac{1}{4}\right]^{-1} = \left(\frac{1}{4}\right)^4$$

$Similarly,\ P(X > 5) = \left(\frac{1}{4}\right)^5$

(2) $P(X > 4 / X > 2) = P(X > 2) = \left(\frac{3}{4}\right)\left(\frac{1}{4}\right)^2 + \left(\frac{3}{4}\right)\left(\frac{1}{4}\right)^3 + \left(\frac{3}{4}\right)\left(\frac{1}{4}\right)^4 + ...$

$$= \left(\frac{3}{4}\right)\left(\frac{1}{4}\right)^2\left[1 + \left(\frac{1}{4}\right) + \left(\frac{1}{4}\right)^2 + ...\right]$$

$$= \left(\frac{3}{4}\right)\left(\frac{1}{4}\right)^2\left[1 - \frac{1}{4}\right]^{-1} = \left(\frac{1}{4}\right)^2$$

(3) $MGF = M_X(t) = E(e^{tX}) = \sum_{x=1}^{\infty} e^{tx} = \sum_{x=1}^{\infty} e^{tx}\left(\frac{3}{4}\right)\left(\frac{1}{4}\right)^{x-1}$

$$= 3\sum_{x=1}^{\infty}\left(\frac{e^t}{4}\right)^x$$

$$= \frac{3e^t}{4}\left\{1 + \frac{e^t}{4} + \left(\frac{e^t}{4}\right)^2 + \ldots\ldots\right\}$$

$$= \frac{3e^t}{4}\left\{1 - \frac{e^t}{4}\right\}^{-1} = \frac{3e^t}{4 - e^t}$$

$$E(X) = \sum_{x=1}^{\infty} xP(X = x) = \sum_{x=1}^{\infty} x\left(\frac{3}{4}\right)\left(\frac{1}{4}\right)^{x-1}$$

$$= \left(\frac{3}{4}\right) + 2\left(\frac{1}{4}\right)\left(\frac{3}{4}\right) + 3\left(\frac{1}{4}\right)^2\left(\frac{3}{4}\right) + \ldots\ldots\ldots\ldots\ldots$$

$$= \left(\frac{3}{4}\right)\left[1 + 2\left(\frac{1}{4}\right) + 3\left(\frac{1}{4}\right)^2 + \ldots\ldots\ldots\ldots\ldots\right]$$

$$= \left(\frac{3}{4}\right)\left[1 - \frac{1}{4}\right]^{-2} = \left(\frac{3}{4}\right)\left(\frac{4}{3}\right)^2 = \frac{4}{3}$$

(4) $Var(X) = \dfrac{1-p}{p^2} = \dfrac{1 - \frac{3}{4}}{\left(\frac{3}{4}\right)^2} = \dfrac{\frac{1}{4}}{\left(\frac{9}{16}\right)} = \dfrac{4}{9}$

58. Let X be a uniformly distributed R.V over [-5,5]. Determine

(1) $P(X \le 2)$

(2) $P(|X| > 2)$

(3) *Cumulative distribution function of X*

(4) $Var(X)$

Solution: **[AU M/J 2016]**

A R.V X is uniformly distributed in [-5,5];

The pdf is

$$f(x) = \frac{1}{b-a}, \quad a < x < b \qquad i.e., f(x) = \frac{1}{10}, \quad -5 \le x \le 5$$

(1) $P(X \le 2) = \int_{-5}^{2} \frac{1}{10}\,dx = \frac{1}{10}[2+5] = \frac{7}{10}$

(2) $P(|X| > 2) = 1 - P(|X| \le 2) = 1 - P(-2 \le X \le 2) = 1 - \int_{-2}^{2}\frac{1}{10}\,dx = 1 - \frac{4}{10} = \frac{3}{5}$

(3) $For \ \ x < -5, \ \ f(x) = 0$

$$\therefore \ F(x) = \int_{-\infty}^{x} f(x)dx = 0$$

$For \ \ x \in [-5,5], \ \ f(x) = \frac{1}{10}$

$$\therefore\ F(x) = \int_{-\infty}^{-5} f(x)dx + \int_{-5}^{x} f(x)dx$$

$$= 0 + \int_{-5}^{x} \frac{1}{10} dx = \frac{1}{10}(x)_{-5}^{x} = \frac{1}{10}(x+5)$$

$For\ \ x > 5,\ \ f(x) = 0$

$$\therefore\ F(x) = \int_{-\infty}^{-5} f(x)dx + \int_{-5}^{5} f(x)dx + \int_{5}^{x} f(x)dx$$

$$= 0 + \int_{-5}^{5} \frac{1}{10} dx + 0 = 1$$

$$\therefore\ The\ \ CDF\ \ \ F(x) = \begin{cases} 0, & if\ \ x < -5 \\ \frac{1}{10}[x+5], & if\ \ -5 \le x \le 5 \\ 1, & if\ \ x > 5 \end{cases}$$

(4) $Var(X) = \dfrac{(b-a)^2}{12} = \dfrac{(5+5)^2}{12} = \dfrac{100}{12} = \dfrac{25}{3}$

59. The annual rainfall in inches in a certain region has a normal distribution with mean 40 and variance 16. What is the probability that the rainfall in a given year is between 30 and 48 inches?

Solution: **[AU N/D 2016]**

Given that $\mu = 40\ \ and\ \ \sigma^2 = 16\ \ X \sim N(\mu, \sigma^2)$

$$P(30 < X < 48) = P\left\{\frac{30-40}{4} < \frac{X-40}{4} < \frac{48-40}{4}\right\}$$

$$P(30 < X < 48) = P\{-2.5 < Z < 2\}$$

$$= P\{0 < Z < 2.5\} + P\{0 < Z < 2\}$$

$$= 0.4938 + 0.4772$$

$$= 0.9710$$

60. The probability distribution function of a random variable X is given by

$f(x) = \dfrac{4x(9-x^2)}{81},\ 0 \le x \le 3$ **. Find the mean and variance and third moments about origin.**

Solution: **[AU N/D 2016]**

$$E(x) = \int_{0}^{3} xf(x)dx$$

$$E(x) = \int_{0}^{3} \frac{4x^2(9-x^2)}{81}\ dx$$

$$E(x) = \int_0^3 \frac{36x^2 - 4x^4)}{81}\,dx = \left(\frac{36x^3}{3(81)} - \frac{4x^5}{5(81)}\right)_0^3 = 4 - \frac{12}{5} = \frac{8}{5}$$

$$E(x^2) = \int_0^3 \frac{36x^3 - 4x^5)}{81}\,dx = \left(\frac{36x^4}{4(81)} - \frac{4x^6}{6(81)}\right)_0^3 = 9 - 3 = 6$$

$$\therefore \sigma^2 = E(x^2) - [E(x)]^2 = 3 - \frac{64}{5} = \frac{11}{25}$$

Third moments about origin :

$$E(x^3) = \int_0^3 \frac{36x^4 - 4x^6)}{81}\,dx = \left(\frac{36x^5}{5(81)} - \frac{4x^7}{7(81)}\right)_0^3 = \frac{108}{5} - \frac{108}{7} = \frac{216}{35}$$

61. Let X be a continuous random variab le with the probability density function $f(x) = \frac{1}{4},\ 2 \le x \le 6$. Find the expected value and variance of X. **[AU N/D '17]**

Solution:

$$f(x) = \frac{1}{4},\ 2 \le x \le 6$$

$$E(X) = \int_2^6 xf(x)dx = \frac{1}{4}\int_2^6 xdx = \frac{1}{4}\left(\frac{x^2}{2}\right)_2^6 = \frac{1}{4}\left(\frac{36}{2} - \frac{4}{2}\right) = 4$$

$$E(X^2) = \int_2^6 x^2 f(x)dx = \frac{1}{4}\int_2^6 x^2 dx = \frac{1}{4}\left(\frac{x^3}{3}\right)_2^6 = \frac{1}{4}\left(\frac{216}{3} - \frac{8}{2}\right) = \frac{52}{3}$$

$$Var(X) = E(X^2) - (E(X))^2 = \frac{52}{3} - 16 = \frac{4}{3}$$

62. The probability of a man hitting a target is $\frac{1}{4}$.

(i) If he fires 7 times, what is the probability of his hitting the target at least twice?

(ii) How many times must he fire so that the probability of his hitting the target at least once is greater than $\frac{2}{3}$?. **[AU A/M 2017]**

Solution:

p = Pobability of the man hitting the target = $\frac{1}{4}$

$$\Rightarrow q = 1 - p = \frac{3}{4}.$$

p(x) = Probability of getting x hits in 7 shots = $7_{C_X}.\left(\frac{1}{4}\right)^x\left(\frac{3}{4}\right)^{7-x}$; $x = 0,1,\dots 7$

(i) Probability of at least two hits $= 1 - \{\, p(0) + p(1) \,\}$

$$= 1 - \left\{7_{C_0}.\left(\frac{1}{4}\right)^0\left(\frac{3}{4}\right)^{7-0} + 7_{C_1}.\left(\frac{1}{4}\right)^1\left(\frac{3}{4}\right)^{7-1}\right\} = \frac{4547}{8192}$$

(ii) Probability of at least one hit in n shots $= 1 - p(0)$

$$= 1 - \left(\frac{3}{4}\right)^x.$$

It is required to find n, so that $1 - \left(\frac{3}{4}\right)^n > \frac{2}{3} \Rightarrow \frac{1}{3} > \left(\frac{3}{4}\right)^n$

Taking logarithms of each side , $\log\frac{1}{3} > n\log\frac{3}{4} \Rightarrow \log 1 - \log 3 > n(\log 3 - \log 4)$

$$0 - 0.4771 > n(0.4771 - 0.6021) \Rightarrow 0.4771 > 0.1250n$$

$$n > \frac{0.4771}{0.1250} = 3.8$$

Result: Hence n cannot be fractional, the required number of shots is 4.

63. In a normal population with mean 15 and standard deviation 3.5, it is found that 647 observation exceed 16.25. What is the total number of observation in the population?

Solution: **[AU A/M 2017]**

$$Let\ \ X \sim N(\mu, \sigma^2)\ \ where\ \ \mu = 15,\ \sigma = 3.5$$

If N is the total number of observations in the population, then we have to find N such that

$$N \times P(X > 16.25) = 647 \ldots\ldots\ldots\ldots(1)$$

$$Now\ \ Z = \frac{X - \mu}{\sigma}$$

$$\begin{aligned} P(X > 16.25) &= P(Z > 0.3571) \\ &= 0.5 - P(0 < Z < 0.3571) \\ &= 0.5 - 0.3752 \\ &= 0.1248 \end{aligned}$$

Sub. in (1), *we get*

$$N \times 0.1248 = 647$$

$$N = 5184$$

64. Let X be uniformly distributed random variable in the interval (a,9) and $P(3 < x > 5) = \frac{2}{7}$.

Find the constant 'a' and compute $P[|x - 5| < 2]$.

Solution: **[A.U. A/M 2018]**

Since X is a Uniformly distributed in (a,9)

$$f(x) = \frac{1}{9 - a}\ ,\ a < x < 9$$

$$\int_3^5 \frac{1}{9 - a}dx = \frac{2}{7} \Rightarrow \frac{1}{9 - a}[x]_3^5 = \frac{2}{7} \Rightarrow \frac{2}{9 - a} = \frac{2}{7} \Rightarrow a = 2$$

$$\begin{aligned} p(|x - 5| < 2) &= p(-2 < x - 5 < 2) \\ &= p(3 < x < 7) \end{aligned}$$

$$= \int_3^7 \frac{1}{7} dx = \frac{1}{7}[x]_3^7 = \frac{4}{7}$$

65. The scores of an achievement test given to 1,00,000 students are normally distributed with mean 500 and standard deviation 100. What should the score of a student be to place him among the top 10% of all students?.

Solution: **[A.U. A/M 2018]**

Let X $\rightarrow$ scores of student in an achievement

Here $\mu = 500;\ \sigma = 100$

Let minimum mark of top 10% students be m.

$$P(X > \mu) = 10\% = 0.1(Area)$$

$$P(500 < X < m) = 0.4$$

$$z = \frac{x - mean}{\sigma} = \frac{m - 500}{100}$$

$$z = 1.28 \ \ (from\ table)$$

$$\frac{m - 500}{100} = 1.28$$

$$m - 500 = 128$$

$$m = 628 \rightarrow \min imum\ mark\ of\ top\ 10\%\ students$$

66. A radar system has a probability 0.1 of detecting a certain target during a single scan. Use binomial distribution to find the probability that the target will be detected at least 2 times in four consecutive scans. Also compute the probability that the target will be detected at least once in twenty scans.

Solution: **[A.U. N/D 2018]**

$$\Pr\{atleast\ 2\ times\ in\ four\ scans\} = 4C_2 p^2 q^{4-2} + 4C_3 p^3 q^{4-3} + 4C_4 p^4 q^{4-4}$$

$$= 0.0523$$

$$\Pr\{atleast\ once\ in\ twenty\} = 0.8784$$

67. An electrical firm manufactures light bulbs that have a length of life which is normally distributed with mean $\mu = 800$ hours and standard deviation $\sigma = 40$ hours. Find the probability that a bulb burns between 778 and 834 hours.

Solution: **[A.U. N/D 2018]**

$$\Pr\{778 < X < 834\} = \Pr\left\{\frac{778 - 800}{40} < X < \frac{834 - 800}{40}\right\}$$

$$= \Pr\{-0.55 < X < 0.85\}$$

$$= 0.5111$$

68. The probability of an individual suffering a bad reaction from an injection of a certain antibiotic is 0.001, Out of 2000 individuals, use Poisson distribution to find that exactly three suffer. Also, find the probability of more than two suffer from bad reaction.

Solution: **[A.U. N/D 2018]**

Given that $p = 0.001;\ n = 2000$

$$Mean = \lambda = np = (0.001)(2000)$$

$$\lambda = 2$$

$$Poisson\ distribution\ , P(X = x) = \frac{e^{-\lambda}\lambda^x}{x!}$$

$$P(X = 3) = \frac{e^{-2}2^3}{3!} = 0.1804$$

$$\begin{aligned} P(X > 2) &= 1 - P(X \le 2) \\ &= 1 - \{P(X = 0) + P(X = 1) + P(X = 2)\} \\ &= 1 - 0.6767 \\ P(X > 2) &= 0.3233 \end{aligned}$$

69. Electric trains in a particular route run every half an hour between 12. Midnight and 6 a.m. Using uniform distribution, find the probability that a passanger entering the station at any time between 1.00a.m. and 1.30 a.m. will have to wait at least twenty minutes.

Solution: **[A.U. N/D 2018]**

Using Uniform distribution,

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

$$f(x) = \frac{1}{30},\ 0 < x < 30$$

$$P(0 < x < 10) = \int_0^{10} f(x)dx = \frac{1}{30}[x]_0^{10} = \frac{10}{30} = \frac{1}{3}$$

Which is a probability of he arrives with in 10 minutes.

Which implies it is probability of 20 mins. Atleast 20 mins. wait

BAYE'S THEOREM:

70. The contents of urns I, II, III are as follows:

Balls / urns	white	black	red
I	**1**	**2**	**3**
II	**2**	**1**	**1**
III	**4**	**5**	**3**

One urn is chosen at random and two balls are drawn. They happen to be white and red. What is the probability that they come from urns I, II, III ? [AU M/J 2006, A/M 2008]

Solution:

Let B_1, B_2, B_3 denote the events that the urns I, II, III are chosen respectively and let A be the event that the two balls taken from the selected urn are white and red.

$$Then\ \ P(B_1) = P(B_2) = P(B_3) = \frac{1}{3}$$

$$P\left(A/B_1\right) = \frac{1C_1 \times 3C_1}{6C_2} = \frac{1\times 3}{15} = \frac{1}{5}$$

$$P\left(A/B_2\right) = \frac{2C_1 \times 1C_1}{4C_2} = \frac{1}{3}$$

$$P\left(A/B_3\right) = \frac{4C_1 \times 3C_1}{12C_2} = \frac{2}{11}$$

Baye's theorem

$$P\left(B_i/A\right) = \frac{P(B_i)P\left(A/B_i\right)}{\sum_{i=1}^{n} P(B_i)P\left(A/B_i\right)}$$

Hence,

$$P\left(B_2/A\right) = \frac{P(B_2)P\left(A/B_2\right)}{\sum_{i=1}^{3} P(B_i)P\left(A/B_i\right)}$$

$$= \frac{P(B_2)P\left(A/B_2\right)}{P(B_1)P\left(A/B_1\right) + P(B_2)P\left(A/B_2\right) + P(B_3)P\left(A/B_3\right)}$$

$$P\left(B_2/A\right) = \frac{\left(\frac{1}{3}\right)\left(\frac{1}{3}\right)}{\left(\frac{1}{3}\right)\left(\frac{1}{5}\right) + \left(\frac{1}{3}\right)\left(\frac{1}{3}\right) + \left(\frac{1}{3}\right)\left(\frac{2}{11}\right)} = \frac{55}{118}$$

$$P\left(B_3/A\right) = \frac{P(B_3)P\left(A/B_3\right)}{\sum_{i=1}^{3} P(B_i)P\left(A/B_i\right)}$$

$$P\left(B_3/A\right) = \frac{\left(\frac{1}{3}\right)\left(\frac{2}{11}\right)}{\left(\frac{1}{3}\right)\left(\frac{1}{5}\right) + \left(\frac{1}{3}\right)\left(\frac{1}{3}\right) + \left(\frac{1}{3}\right)\left(\frac{2}{11}\right)} = \frac{30}{118}$$

$$P\left(B_1/A\right) = \frac{P(B_1)P\left(A/B_1\right)}{\sum_{i=1}^{3} P(B_i)P\left(A/B_i\right)}$$

$$P\left(B_1/A\right) = 1 - P\left(B_2/A\right) - P\left(B_3/A\right)$$

$$= 1 - \frac{55}{118} - \frac{30}{118}$$

$$P\left(B_1/A\right) = \frac{33}{118}$$

71. A bag A contains 2 white and 3 red balls and a bag B contains 4 white and 5 red balls. One ball is drawn at random from one of the bags and is found to be red. Find the probability that it was

drawn from bag B. **[AU N/D 2006]**

Solution :

Let B_1 the event that the ball is drawn from the bag A

B_2 the event that the ball is drawn from the bag B

A be the event that the drawn ball is red.

$$P(B_1) = P(B_2) = \frac{1}{2}$$

$$P\left(A/B_1\right) = \frac{3C_1}{5C_1} = \frac{3}{5}$$

$$P\left(A/B_2\right) = \frac{5C_1}{9C_1} = \frac{5}{9}$$

Baye's theorem

$$P\left(B_i/A\right) = \frac{P(B_i)P\left(A/B_i\right)}{\sum_{i=1}^{n} P(B_i)P\left(A/B_i\right)}$$

$$P\left(B_2/A\right) = \frac{P(B_2)P\left(A/B_2\right)}{P(B_1)P\left(A/B_1\right) + P(B_2)P\left(A/B_2\right) + P(B_3)P\left(A/B_3\right)}$$

$$P\left(B_2/A\right) = \frac{\left(\frac{1}{2}\right)\left(\frac{5}{9}\right)}{\left(\frac{1}{2}\right)\left(\frac{3}{5}\right) + \left(\frac{1}{2}\right)\left(\frac{5}{9}\right)} = \frac{25}{52}$$

72. In a bolt factory, machines A, B and C manufacture respectively 25%, 35% and 40% of total output. Also out of these output of A,B,C. 5,4,2 percent respectively are defective. A bolt is drawn at random from the total output and it is found to be defective. What is the probability that it was manufactured by the machine B ?

Solution:

Let E_1 , E_2 , E_3 be the events that the bolts are manufactured by A,B,C respectively.

Given: In a bolt factory, machine A manufacture 25% of total output.

$$P(E_1) = 25\% = \frac{25}{100} = 0.25$$

Given: Machine B manufacture 35% of total output.

$$P(E_1) = 35\% = \frac{35}{100} = 0.35$$

Given: Machine C manufacture 40% of total output.

$$P(E_1) = 40\% = \frac{40}{100} = 0.40$$

Let X be the event of drawing defective bolt.

Given: In a bolt factory, machine A manufacture 25% of total output.

$$P\left(X/E_1\right)=\frac{5}{100}=0.05$$

$$P\left(X/E_2\right)=\frac{4}{100}=0.04$$

$$P\left(X/E_3\right)=\frac{2}{100}=0.02$$

To find the probability that the defective bolt selected at random is manufactured from the machine B.

To find $P\left(E_2/X\right)$

Baye's theorem

$$P\left(B_i/A\right)=\frac{P(B_i)P\left(A/B_i\right)}{\sum_{i=1}^{n}P(B_i)P\left(A/B_i\right)}$$

$$P\left(E_2/X\right)=\frac{P(E_2)P\left(X/E_2\right)}{P(E_1)P\left(X/E_1\right)+P(E_2)P\left(X/E_2\right)+P(E_3)P\left(X/E_3\right)}$$

$$=\frac{(0.35)(0.04)}{(0.25)(0.05)+(0.35)(0.04)+(0.4)(0.02)}$$

$$P\left(E_2/X\right)=0.406$$

73. Of these types of spark plugs, 6% of Type A plugs are defective, 4% of Type B spark plugs are defective, and 2% of Type C spark plugs are defective. A spark plug is selected at random from a batch of sparkplugs containing 50 Type A plugs, 30 Type B plugs, and 20 Type C plugs. The selected plug is found to be defective. What is the probability that the selected plug was of Type A? [AU A/M 2021]

Solution:

Let E_1 , E_2 , E_3 be the events that the plugs manufactured by Type A,B,C respectively.

Given: Type A containing 50 plugs of total.

$$P(E_1)=\frac{50}{100}=0.5$$

Given Type B containing 30 plugs of total.

$$P(E_2)=\frac{30}{100}=0.3$$

Given Type C containing 20 plugs of total.

$$P(E_3)=\frac{20}{100}=0.2$$

Let X be the event of drawing defective plugs.

$$P\left(X/E_1\right)=\frac{6}{100}=0.06$$
$$P\left(X/E_2\right)=\frac{4}{100}=0.04$$
$$P\left(X/E_3\right)=\frac{2}{100}=0.02$$

To find the probability that the defective bolt selected at random is manufactured from the Type A.

To find $P\left(E_1/X\right)$

Baye's theorem

$$P\left(B_i/A\right)=\frac{P(B_i)P\left(A/B_i\right)}{\sum_{i=1}^{n}P(B_i)P\left(A/B_i\right)}$$

$$P\left(E_2/X\right)=\frac{P(E_1)P\left(X/E_1\right)}{P(E_1)P\left(X/E_1\right)+P(E_2)P\left(X/E_2\right)+P(E_3)P\left(X/E_3\right)}$$
$$=\frac{(0.5)(0.06)}{(0.5)(0.06)+(0.3)(0.04)+(0.2)(0.02)}$$
$$P\left(E_2/X\right)=0.6522$$

74. The p.d.f of a continuous random variable X is given as

$$f(x)=\begin{cases}\frac{1}{6} & ,-3\le x\le 3\\ 0 & ,otherwise\end{cases}$$

Find (A) $P(-2<X<0)$**, (B) Cumulative distribution function F(x) and (C)** $E(X)\ and\ Var(X)$ **[AU A/M 2021]**

Solution:

(A) To find $P(-2<X<0)$

$$P(-2<X<0)=\int_{-2}^{0}\frac{1}{6}dx=\frac{1}{6}(x)_{-2}^{0}=\frac{1}{6}(2)=\frac{1}{3}$$

(B) To find $CDF, F(x)$

$Let\ x<-3,\ f(x)=0$

$$F(x)=\int_{-\infty}^{x}f(x)dx=0$$

$$Let\ -3 \le x \le 3,\ f(x) = \frac{1}{6}$$

$$F(x) = \int_{-\infty}^{-3} f(x)dx + \int_{-3}^{x} f(x)dx = 0 + \int_{-3}^{x} \frac{1}{6}dx = \frac{1}{6}(x+3)$$

$$Let\ x > 3,\ f(x) = 0$$

$$F(x) = \int_{-\infty}^{-3} f(x)dx + \int_{-3}^{3} f(x)dx + \int_{3}^{x} f(x)dx = 0 + \int_{-3}^{3} \frac{1}{6}dx + 0 = \frac{1}{6}(3+3) = 1$$

$$\therefore\ F(x) = \begin{cases} 0, & x < -3 \\ \frac{1}{6}(x+3) & ,-3 \le x \le 3 \\ 1, & x > 3 \end{cases}$$

© To find $E(X)\ and\ Var(X)$:

$$E(x) = \int_{-\infty}^{\infty} xf(x)dx = \int_{-3}^{3} x\frac{1}{6}dx = \frac{1}{6}\left(\frac{x^2}{2}\right)_{-3}^{3} = \frac{1}{6}\left(\frac{9}{2} - \frac{9}{2}\right) = 0$$

$$E(x^2) = \int_{-\infty}^{\infty} x^2 f(x)dx = \int_{-3}^{3} x^2 \frac{1}{6}dx = \frac{1}{6}\left(\frac{x^3}{3}\right)_{-3}^{3} = \frac{1}{18}(27+27) = 3$$

$$Var(X) = E(x^2) - [E(x)]^2 = 3 - 0 = 3$$

75. Customers are used to evaluate preliminary product designs. In the past, 95% of highly successful products received good reviews, 60% of moderately successful products received good reviews, and 10% of poor products received good reviews. In addition, 40% of products have been highly successful, 35% have been moderately successful, and 25% have been poor products.

(1) What is the probability that a product attains a good review?

(2) If a new design attains a good review, what is the probability that it will be a highly successful product?

(3) If a product does not attain a good review, what is the probability that it will be a highly successful product?

Solution: **[AU A/M 2024]**

P (Highly successful products) = P(H) = $\frac{40}{100}$ = 0.04

P (Moderately successful products) = P(M) = $\frac{35}{100}$ = 0.35

P (Poor products) = P(P) = $\frac{25}{100}$ = 0.25

P (Highly successful products received good reviews) = P(R/H) =95%

P (Moderately successful products received good reviews) = P(R/M) = 60%

P (Poor products received good reviews) = P(R/P) =10%

P (Products attains a good review) = P(R)

$$P(R) = P\left(R/_H\right)P(H) + P\left(R/_M\right)P(M) + P\left(R/_P\right)P(P)$$

$$= \left[\frac{40}{100}\right]\left[\frac{95}{100}\right] + \left[\frac{35}{100}\right]\left[\frac{60}{100}\right] + \left[\frac{25}{100}\right]\left[\frac{10}{100}\right] = 0.615$$

$$P(R^C) = 1 - P(R)$$

$$= 1 - 0.615 = 0.385 \quad (Bad\ \ review)$$

$$P\left(H/_R\right) = \frac{P(R)P(H)}{P(R)}$$

$$= \frac{\frac{95}{100} * \frac{40}{100}}{0.615} = 0.6178$$

$$P\left(H/_{R^C}\right) = \frac{P\left(R^C\right)P(H)}{P(R^C)}$$

$$= \frac{\frac{95}{100} * \frac{45}{100}}{0.385} = 0.0519$$

76. Assume that each of your calls to a popular radio station has a probability of 0.02 of connecting, that is, of not obtaining a busy signal. Assume that your calls are independent.

(1) What is the probability that your first call that connects is your tenth call?

(2) What is the probability that it requires more than five calls for you to connect?

(3) What is the mean number of calls needed to connect?

Solution: **[AU A/M 2024]**

Given $p = 0.02,\ q = 1 - p = 1 - 0.02 = 0.98$

$$P(X = r) = q^{r-1}\ p$$

$$P(X = 10) = (0.98)^{10-1}\ (0.02)$$

$$= (0.98)^{9}\ (0.02) = 0.01667$$

$$P(X > 5) = 1 - P(X \le 5)$$

$$= 1 - \{\ P(X = 1) + P(X = 2) + P(X = 3) + P(X = 4) + P(X = 5)\}$$

$$= 1 - \left\{(0.98)^0(0.02) + (0.98)^1(0.02) + (0.98)^2(0.02) + (0.98)^3(0.02) + (0.98)^4(0.02)\right\}$$

$$= 1 - 0.0961 = 0.9039$$

$$E(X) = Mean = \frac{1}{p} = \frac{1}{0.02} = 50$$

77. The thickness of a flange on an aircraft component is uniformly distributed between 0.95 and 1.05 millimeters.

(1) Determine the cumulative distribution function of flange thickness.

(2) Determine the proportion of flanges that exceeds 1.02 millimeters.

(3) What thickness is exceeded by 90% of the flanges?

(4) Determine the mean and variance of flange thickness.

Solution: **[AU A/M 2024]**

Given f(x) is uniformly distributed between 0.95 and 1.05 mm.

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

$$= \frac{1}{1.05 - 0.95},\ 0.95 < x < 1.05$$

$$= \frac{1}{0.1} = 10,\ 0.95 < x < 1.05$$

The CDF is

1. $$F(x) = \begin{cases} 0, & if\ -\infty < x < a \\ \dfrac{x-a}{b-a}, & if\ a \le x \le b \\ 1 & ,if\ x \succ b \end{cases}$$

$$F(x) = \begin{cases} 0, & if\ -\infty < x < 0.95 \\ \dfrac{x-0.95}{0.1}, & if\ 0.95 \le x \le 1.05 \\ 1 & ,if\ x \succ 1.05 \end{cases}$$

2. $$P(x > 1.02) = \int_{1.02}^{1.05} f(x)dx = 10[x]_{1.02}^{1.05}$$

$$= 10(1.05 - 1.02) = 10(0.003) = 0.3$$

3. $$P(X > x) = 0.9$$

$$\int_{x}^{1.05} f(x)dx = 0.9$$

$$10[x]_{x}^{1.05} = 0.9$$

$$10[1.05] - 10x = 0.9$$

$$10.5 - 0.9 = 10x$$

$$9.6 = 10x$$

$$x = 0.96$$

4. $E(X) = \frac{a+b}{2} = \frac{1.05+0.95}{2} = 2/2 = 1$

$$Var(X) = \frac{(b-a)^2}{12}$$

$$= \left[\frac{0.1^2}{12}\right] = 0.0008$$

78. The life of a semiconductor laser at a constant power is normally distributed with a mean of 7000 hours and a standard deviation of 600 hours.

(1) What is the probability that a laser fails before 5000 hours?

(2) What is the life in hours that 95% of the lasers exceed?

(3) If three lasers are used in a product and they are assumed to fail independently, what is the probability that all three are still operating after 7000 hours?

Solution: **[AU A/M 2024]**

Given $\mu = 7000$, $\sigma = 600$

$$N(\mu,\sigma) = N(7000,600)$$

$$z = \frac{x-\mu}{\sigma}$$

a) $P(X < 5000) \Rightarrow z = \frac{x-\mu}{\sigma}$

when $x = 5000$

$$z = \frac{5000-7000}{600} = \frac{-2000}{600} = -3.333$$

$$P(X < 5000) = P(z < -3.333) = P(0 < z < 3.333) = 0.0004$$

b) $P(X < x_0) = 0.95$

$$P\left[\frac{x-7000}{600} \le \frac{x_0-7000}{600}\right] = 0.95$$

$$P\left[z \le \frac{x_0-7000}{600}\right] = 0.95$$

$$P\left[z \ge \frac{7000-x_0}{600}\right] = 0.95$$

$$P\left[0 \le z \le \frac{7000-x_0}{600}\right] = 0.5 - 0.95 = -0.45$$

$$\frac{7000-x_0}{600} = 1.64$$

$$7000 - x_0 = 984$$

$$7000 - 984 = x_0$$

$$x_0 = 6016$$

c) $P(X > 7000) = P\left(\frac{X-7000}{600} > \frac{7000 > 7000}{600}\right)$

$$= P(Z > 0)$$
$$= 0.5 - P(Z < 0)$$
$$0.5 = P$$

P(All 3 lasers still operate)= p^3=1/8.

79. The probabilities of X ,Y and Z becoming managers are 4/9, 2/9 and 1/3 respectively. The probabilities that the Bonus scheme will be introduced if X, Y, Z becomes managers are 3/10, 1/2, and 4/5 respectively.

(1) What is the probability that the Bonus scheme will be introduced?

(2) If the Bonus scheme has been introduced, what is the probability that the manager appointed was X?

Solution: **[AU N/D 2024]**

Let P(x) be the probability that X becomes manager = 4/9

Let P(y) be the probability that Y becomes manager = 2/9

Let P(z) be the probability that Z becomes manager = 1/3

Let P(B/X) = Probability that bonus scheme is introduced when X becomes manager.

P(B/X) = 3/10

P(B/Y) = Probability that bonus scheme is introduced when Y becomes manager.

P(B/Y) = 1/2

P(B/Z) = Probability that bonus scheme is introduced when Z becomes manager.

P(B/Z) = 4/5

Let B denote event that the bonus scheme is introduced. Total probability

$$P(B) = P\left(B/X\right)P(X) + P\left(B/Y\right)P(Y) + P\left(B/Z\right)P(Z)$$
$$= \left(\frac{4}{9}\right)\left(\frac{3}{10}\right) + \left(\frac{2}{9}\right)\left(\frac{1}{2}\right) + \left(\frac{1}{3}\right)\left(\frac{4}{5}\right)$$
$$= \left(\frac{4}{30}\right) + \left(\frac{1}{9}\right) + \left(\frac{4}{15}\right) = \left(\frac{6+5+12}{45}\right) = \left(\frac{23}{45}\right)$$

$$P\left(X/B\right) = \frac{P(X \cap B)}{P(B)} = \frac{P(X)P\left(B/X\right)}{P(B)} = \frac{\frac{12}{90}}{\frac{23}{45}} = \frac{6}{23}$$

ANNA UNIVERSITY QUESTIONS

1. A random variable X has the following function:

X:	0	1	2	3	4	5	6	7
P(X):	0	K	2K	2K	3K	K^2	$2K^2$	$7K^2+K$

(a) Find K, (b) Evaluate P[X< 6], $P[X \geq 6]$ (c) If $P[X \leq C] > \frac{1}{2}$ find the minimum value of C.

(d) Evaluate $P[1.5 < X < 4.5 | X > 2]$ (e) Find $P[X < 2], P[X > 3], P[1 < X < 5]$.

[P.g no 22] [M/J 2012,M/J 2014]

2. A random variable X has the following probability distribution .

X	-2	-1	0	1	2	3
P(x)	0.1	k	0.2	2k	0.3	k

Find (1) the value of k. (2) Evaluate P(X<1) and P(-1<X<2) (3) find the cumulative distribution of x. (4) evaluate mean of x. [Pg.No. 23][AU M/J 2016, N/D 2024]

3. The probability function function of random variable X is defined as
$P(X=0)=3C^2, P(X=1)=4C-10C^2, P(X=2)=5C-1, where\ C>0, and\ P(X=r)=0\ if\ r \neq 0,1,2.$
find (i) the value of C (iii) The distribution function of X. (iii) The distribution function of X.
(iv) The largest value of X for which F(x) < 1/2. [Pg.No. 24]

4. The probability function of an infinite discrete distribution is given by P (X = j) = $\frac{1}{2^j}$;
j=1,2,…∞. Verify that the total probability is 1 and find the mean and variance of the distribution. Find also P(X is even), P(X≥5) and P(X is divisible by 3). [Pg No. 25]

5. If the random variable X takes the values 1,2,3 and 4 such that 2P(X=1) = 3P(X=2) = P(X=3) = 5P(X=4) find the probability distribution and cumulative distribution function of X. [P.g no. 26][AU N/D 2012]

6. A random variable X takes the values -2,-1,0 and 1 with probabilities $\frac{1}{8}, \frac{1}{8}, \frac{1}{4}\ and\ \frac{1}{2}$ respectively. Find and draw the probability distribution function.
Find (1) the probability distribution of X (2) p(2<x<6) (3) mean of X 4) Variance of X.
[P.g no 27][AU N/D 2014]

7. If X has the distribution function $F[x] = \begin{cases} 0 & x < 1 \\ \frac{1}{3} & for\ 1 \le x < 4. \\ \frac{1}{2} & for\ 4 \le x < 6 \\ \frac{5}{6} & for\ 6 \le x < 10 \\ 1 & x \ge 10. \end{cases}$

Find (1) the probability distribution of X (2) p(2<x<6) (3) mean of X 4) Variance of X. [Pg.no.28]

8. Consider a discrete r. v. 'X' with probability function

$$p(X = x) = \begin{cases} \dfrac{1}{x(x+1)} & , \quad x = 1,2,3\ldots \\ 0 & othewise \end{cases}$$

Show that E(X) does not exist even though MGF exist. [P.g no 29] [AU N/D 2012]

9. If the density function of a continuous random variable X is given by

$$f(x) = \begin{cases} ax & ,0 \le x \le 1 \\ a & ,1 \le x \le 2 \\ 3a - ax & ,2 \le x \le 3 \\ 0 & ,otherwise \end{cases}$$

[P.g no 30][AU N/D 2008, A/M '17]

(i)Find the value of a (ii)The cumulative distribution function of X

(iii)If x_1,x_2 and x_3 are 3 independent observations of X. What is the probability that exactly one of these 3 is greater than 1.5?

10. The Distribution F of a continuous random variable X is given by

$$F(x) = \begin{cases} 0 & , x < 0 \\ x^2 & ,0 \le x < \dfrac{1}{2} \\ 1 - \dfrac{3}{25}(3-x)^2 & , \dfrac{1}{2} \le x < 3 \\ 1 & , x \ge 3 \end{cases}$$

[AU N/D '31][Pg no. 32]

Find the pdf of X and evaluate P(|X|≤1) and P(1/3<X<4) using both PDF and CDF.]

11. $If\ f(x) = \begin{cases} xe^{-\frac{x^2}{2}} & ; x \ge 0 \\ 0 & ; x > 0 \end{cases}$, then show that f(x) is a pdf and find F(x).

[P.g no 32] [AU N/D 2014]

12. The cumulative distribution function (cdf)of a random variable X is $F(x)=1-(1+x)e^{-x}$,x>0. Find the probability density function of X.Mean and Variance of X. [P.g no 33][AU N/D 10]

13. If X is a random variable with a continuous distribution function F(x),prove that Y=F(x) has a

uniform distribution in (0,1).Further if

$$f(x)=\begin{cases}\frac{1}{2}(x-1);1\le x\le 3\\ 0\quad ;otherwise\end{cases}, find\ the\ range\ of\ Y\ corresponding\ to\ the\ range\ 1.1\le x\le 2.9.$$

[Pg.No. 33][AU N/D 2010]

14. If the density function of X equals $f(x)=\begin{cases}Ce^{-2x} & ;x\ge 0\\ 0 & ;x<0\end{cases}$ Find C . What is P[X>2]?

[P.g no 34] [AU A/M 2010]

15. Continuous random variable X has a pdf $f(x)=kx^2e^{-x}$, $x\geq0$. Find K , mean and variance.

random variable X has the pdf [AU M/J 2013] [P.g No 35]

16. Continuous random variable X has a pdf $f(x)=\frac{K}{1+x^2}$; $-\infty<x<\infty$ (i) find The value of K

(ii)Distribution function of X P[X≥0] [P.g no 35] [AU N/D 2011]

17. If the density function of X equals $f(x)=\begin{cases}Ce^{-2x}\ ,\ 0<x<\infty\\ 0\ ,\ x<0\end{cases}$ $Find\ \ P(X>2)$,

[P.g no 36]

18. The probability density function of a random variable X is given by

$f_X(X)=\begin{cases}x & ;0<X<1\\ k(2-x);1\le X\le 2\\ 0 & ;otherwise\end{cases}$ (i) Find the value of K? (ii) Find $P(0.2<X<1.2)$

(iii) What is $P(0.5<X<1.5/X\ge 1)$(iv) Find the distribution function of f(x)[A.U. A/M 2011] [Pg 36]

19. The probability function of an infinite discrete distribution is given by P(X=x)=$\frac{1}{2^x}$, x=1,2,…... ∞

Find the mean and variance of the distribution. Also find P(X is even).[Pg.no 38][AU N/D 2011]

20. Find the MGF of the random variable with the probability law

$P(X=x)=q^{x-1}\ p,\ x=1,2,3$ $[Pg\ no\ 40]$

21. Let 'X' be a random variable with p.d.f

$f(x)=\begin{cases}\frac{1}{3}e^{-\frac{X}{3}}\ , & x>0\\ 0,\ othewise\end{cases}$ $[Pg\ no\ 42][AU\ A/M\ 2007]$

Find (a) P(X>3) (b) MGF of 'X' (c) E(X) and Var(X)

22. Let 'X' be a random variable with p.d.f $f(x) = \begin{cases} \frac{1}{2}e^{-\frac{x}{2}} , & x > 0 \\ 0, & otherwise \end{cases}$

    Find mean and variance. [Pg .no 43][A.U N/D 2013 R.P]

23. Find the MGF of the random variable X with p.d.f $f(x) = \begin{cases} \frac{x}{4}e^{-\frac{x}{2}} , & x > 0 \\ 0, & otherwise \end{cases}$

    Also find the first four moments about the origin. [Pg .no 44] [A.U N/D 2013, M/J 2014,'17]

24. A random variable has the p.d.f given by $f(x) = \begin{cases} 2e^{-2x}, & x \geq 0 \\ 0 , & x < 0 \end{cases}$

    Find (a) the moment generating function (b) the first four moment about the origin.

    (or)

    Let X be an exponential random variable with $E(X^2) = \frac{1}{2}$. Obtain E(X), Var(X), MGF and $P(X > 3 / X > 1)$

    [Pg .no 45] [A.U N/D 2014, A/M 2021]

25. Find the mean, variance and MGF of the random variable X having the pdf

    $$f(x) = \begin{cases} x & ,0 < x \leq 1 \\ 2 - x & ,1 < x \leq 2 \\ 0 & ,otherwise \end{cases}$$

    [Pg no 46] [AU N/D 2013]

26. Find the mean, variance and MGF of the random variable X having the pdf

    $$f(x) = \begin{cases} \lambda e^{-\lambda(X-a)}, & x \geq a \\ 0, & otherwise \end{cases}$$

    [Pg .no 47][AU A/M 2010]

27. Define binomial distribution, Find the moment generating function of binomial distribution and also find its mean and variance. [Pg.no 48] [A.U M/J 14, N/D 2024]

28. By Calculating the moment generating function of Poisson distribution with parameter λ, prove that the mean and variance of the Poisson distribution are equal.

    [A.U N/D 2014, 15][Pg. no. 49]

29. Find the moment generating function, mean and variance of geometric distribution?

    [Pg.no. 50][AU A/M 15]

30. Find the moment generating function of Uniform distribution. Hence find its mean and variance. [Pg no 52][AU N/D 2006]

31. Define exponential distribution, Find the moment generating function of exponential distribution and also find its mean and variance. [Pg.no 53][AU M/J 2012]

32. Define Gamma distribution, Find the MGF of Gamma distribution and also find mean and variance. [Pg.no 54] [A.U N/D 2014, '17 M/J 14,0 2013,2011]

33. Find the moment generating function, mean and variance of Normal distribution? [Pg.no 56 AU N/D 2017]

34. State and Prove Memory less Property of Exponential Distribution. Using this property, Solve the following problem:
    The length of the shower on a tropical island during rainy season has an exponential distribution with parameter 2, time being measured in minutes. What is the probability that a shower will last more than 3 minutes? [AU N/D 2016][Pg.no.57]

35. State and prove the memory less property of the geometric distribution? [AU M/J 2010][Pg .no 58]

36. Derive probability mass function of Poisson distribution as a limiting case of binomial distribution. [Pg .no 59] [A.U N/D 2014, 2013 R.P]

37. The time (in hours) required to repairs a machine is exponential distributed with Parameter $\lambda = \frac{1}{2}$. [Pg.no 60] [AU N/D 2009, 2024]
    (i) What is the probability that the repair time exceeds 2 2hours?
    (ii) What is the conditional probability that a repair takes at least 10h given that its duration exceeds 9h?

38. Suppose that telephone calls arriving at a particular switchboard follow a Poisson process with an average of 5 calls coming per minute . What is the probability that up to a minute will elapse unit 2 calls have come into the switch board? [Pg.no 61] [AU A/M 2011]

39. X is a Poisson variate such that $P(X=2) = 9P(X=4) + 90P(X=6)$ Find Mean, [Pg.No. 61] $E(X^2)$ *and Variance*

40. .If X is a normal variate with mean 30 and standard deviation 5. Find the probabilities that
    (i) $26 \le X \le 40$, (ii) $X \ge 45$, $(iii)\ |X - 30| > 5$ [Pg.No. 62]

41. Suppose that a trainee soldier shoots a target in an independent fashion. If the probability that the target is shot on any one shot is 0.7
    1) What is the probability that the target would be hit on tenth attempt?
    2) What is the probability that it takes him less than 4 shots?
    3) What is the probability that it takes him an even number of shots?[Pg no.63][A.U. N/D 2014 P.Q.T]

42. Trains arrive at a station at 15 minutes intervals starting at 4a.m.If a passenger arrive at a station at a time that uniformly distributed between 9.00 and 9.30, find the probability that he has to wait for the train for (1) less than 6 minutes (2) more than 10 minutes. [A.U. N/D 2014 P.Q.T][Pg no. 63]

43. The marks obtained by a number of students for a certain subject is assumed to be normally distributed with mean 65 and S.D of 5.If 3 students are taken at random from this set, what is the

probability that exactly 2 of them will have marks over 70 [Pg. no. 64] [A.U A/M 2010]

44. Assume that the reduction of a person's oxygen consumption during a period of Transcendental meditation(TM) is a continuous random variable X normally distributed with mean 37.6 cc/mm and S.D 4.6 cc/min. Determine the probability that during a period of T.M a person's oxygen consumption will be reduced by (i) at least 44.5 cc/min (ii) at most 35.0 cc/min (iii) Anywhere from 30.0 to 40.0 cc/min [Pg. no. 64][A.U. N/D 2012]

45. Let X and Y be independent normal variates with mean 45 and 44 and standard deviation 2 and 1.5 respectively. What is the probability that randomly chosen values of X and Y differ by 1.5 or more? [A.U N/D 2011] [Pg. no. 65]

46. If X is a uniform random variable in the interval (-2,2) find the probability density function $Y = |X| \; and \; E[Y]$ [Pg. no. 65][A.U N/D 2011]

47. If the probability that an applicant for a driver's license will pass the road test on any given trial is 0.8. what is the probability that he will finally pass the test (1) on the fourth trial and (2) In less tha 4 trials? [Pg no 66][A.U A/M 2010]

48. Find the MGF of the two parameter exponential distribution whose density function is given by $f(x) = \lambda e^{-\lambda(x-a)}$, $x \geq a$ and hence find the mean and variance? [A.U A/M 2010][Pg no 67]

49. The marks obtained by a number of students in a certain subject are assumed to be normally distributed with mean 65 and standard deviation 5. If 3 students are selected at random from this group, what is the probability that two of them will have marks over 70? [A.U A/M 2010] [Pg no 68]

50. The CDF of the random variable of X is given by

$$F_x(x) = \begin{cases} 0 & ; \; x < 0 \\ x + \frac{1}{2} & ; \; 0 \leq x \leq \frac{1}{2} \\ 1 & ; \; x > \frac{1}{2} \end{cases}$$

Draw the graph of the CDF. Compute $P\left(X > \frac{1}{4}\right)$, $P\left(\frac{1}{3} < X < \frac{1}{2}\right)$ [AU A/M '15] [Pg no 68]

51. Messages arrive at a switch board in a Poisson manner at an average rate of six per hour. Find the probability for each of the following events: [AU A/M '15 N/D '16][Pg no 69]
(1) exactly two messages arrive within one hour
(2) no message arrives within one hour
(3) at least three messages arrive within one hour.

52. The peak temperature T, as measured in degrees Fahrenheit, on a particular day is the Gaussian (85,10) random variable. What is P(T>100),P(T<60) and P(70<T<100)? [AU A/M '15] [Pg no 70]

53. A component has an exponential time to failure distribution with mean of 10,000 hours.

(1) The component has already been in operation for its mean life. What is the probability that it will fail by 15,000 hours?

(2) At 15,000 hours the component is still in operation. What is the probability that it will operate for another 5000 hours? [Pg. no. 71][AU N/D 2015]

54. The average percentage of marks of candidates in an examination is 42 with a standard deviation of 10. If the minimum mark for pass is 50% and 1000 candidates appear for the examination, how many candidates can be expected to get the pass mark follow normal distribution? If it is required, that double the number of the candidates should pass, what should be the minimum mark for pass? [Pg. no. 72][AU N/D 2016]

55. A continuous random variable X has the pdf $f(x) = kx^3 e^{-x},\ x \ge 0$
Find the rth order moment about the origin, moment generating function, mean and variance of X. [Pg.no. 72][AU N/D 2015]

56. Let X be a continuous R.V with probability density function $f(x) = \begin{cases} xe^{-x}, & x > 0 \\ 0, & otherwise \end{cases}$

Find (1) the cumulative distribution function of X
(2) Moment Generating Function of X
(3) P(X<2)
(4) E(X) [Pg. no. 73][AU M/J 2016]

57. Let $P(X = x) = \left(\frac{3}{4}\right)\left(\frac{1}{4}\right)^{x-1}$, $x = 1,2,3.....$ be the probability mass function of a R.V X

Compute
(1) $P(X > 4),\ P(X > 5)$
(2) $P(X > 4 / X > 2)$
(3) $MGF,\ E(X)$
(4) $Var(X)$

[Pg. no. 74][AU M/J 2016, 2021]

58. Let X be a uniformly distributed R.V over [-5,5]. Determine
(1) $P(X \le 2)$
(2) $P(|X| > 2)$
(3) $Cumulative\ \ distributi\ on\ \ function\ \ of\ \ X$
(4) $Var(X)$ [Pg.no. 75][AU M/J 2016]

59. The annual rainfall in inches in a certain region has a normal distribution with mean 40 and variance 16. What is the probability that the rainfall in a given year is between 30 and 48 inches? [Pg.no. 76] [AU N/D 2016]

60. The probability distribution function of a random variable X is given by

$f(x) = \dfrac{4x(9 - x^2)}{81},\ 0 \le x \le 3$. Find the mean and variance and third moments about origin.

[Pg.no. 76][AU N/D 2016]

61. Let X be a continuous random variab le with the probability density function

$f(x) = \dfrac{1}{4},\ 2 \le x \le 6$. Find the expected value and variance of X. [pg.no77] [AU N/D ‘17]

62. The probability of a man hitting a target is $\dfrac{1}{4}$.

(i) If he fires 7 times, what is the probability of his hitting the target at least twice?

(ii) How many times must he fire so that the probability of his hitting the target at least once is greater than $\frac{2}{3}$?. [pg.no.77] [AU A/M 2017]

63. In a normal population with mean 15 and standard deviation 3.5, it is found that 647 observation exceed 16.25. What is the total number of observation in the population? [pg.no.78] [AU A/M 2017]

64. Let X be uniformly distributed random variable in the interval (a,9) and $P(3 < x > 5) = \frac{2}{7}$. Find the constant 'a' and compute $P[|x-5| < 2]$. [pg.no.78] [A.U. A/M 2018]

65. The scores of an achievement test given to 1,00,000 students are normally distributed with mean 500 and standard deviation 100. What should the score of a student be to place him among the top 10% of all students?. [pg.no.79] [A.U A/M 2018]

66. A radar system has a probability 0.1 of detecting a certain target during a single scan. Use binomial distribution to find the probability that the target will be detected at least 2 times in four consecutive scans. Also compute the probability that the target will be detected at least once in twenty scans. [pg.no.79] [A.U. N/D 2018]

67. An electrical firm manufactures light bulbs that have a length of life which is normally distributed with mean $\mu = 800$ hours and standard deviation $\sigma = 40$ hours. Find the probability that a bulb burns between 778 and 834 hours. [pg.no.79] [A.U. N/D 2018]

68. The probability of an individual suffering a bad reaction from an injection of a certain antibiotic is 0.001, Out of 2000 individuals, use Poisson distribution to find that exactly three suffer. Also, find the probability of more than two suffer from bad reaction. [pg.no.79][A.U. N/D 2018]

69. Electric trains in a particular route run every half an hour between 12. Midnight and 6 a.m. Using uniform distribution, find the probability that a passenger entering the station at any time between 1.00a.m. and 1.30 a.m. will have to wait at least twenty minutes. [pg.no.80][A.U. N/D 2018]

70. The contents of urns I, II, III are as follows:

Balls / urns	white	black	red
I	1	2	3
II	2	1	1
III	4	5	3

One urn is chosen at random and two balls are drawn. They happen to be white and red. What is the probability that they come from urns I, II, III ? [pg.no.80] [AU M/J 2006, A/M 2008]

71. A bag A contains 2 white and 3 red balls and a bag B contains 4 white and 5 red balls. One ball is drawn at random from one of the bags and is found to be red. Find the probability that it was drawn from bag B. [pg.no.81][AU N/D 2006]

72. In a bolt factory, machines A, B and C manufacture respectively 25%, 35% and 40% of total output. Also out of these output of A,B,C. 5,4,2 percent respectively are defective. A bolt is drawn at random from the total output and it is found to be defective. What is the probability that it was manufactured by the machine B ? [pg.no.82]

73. Of these types of spark plugs, 6% of Type A plugs are defective, 4% of Type B spark plugs are defective, and 2% of Type C spark plugs are defective. A spark plug is selected at random from a batch of sparkplugs containing 50 Type A plugs, 30 Type B plugs, and 20 Type C plugs. The selected plug is found to be defective. What is the probability that the selected plug was of Type A? [AU A/M 2021][pg. no. 83]

74. The p.d.f of a continuous random variable X is given as

$$f(x)=\begin{cases}\frac{1}{6} & ,-3\le x\le 3\\ 0 & ,otherwise\end{cases}$$

Find (A) $P(-2<X<0)$, (B) Cumulative distribution function F(x) and (C) $E(X)\ and\ Var(X)$ [AU A/M 2021] [pg. no. 84]

75. Customers are used to evaluate preliminary product designs. In the past, 95% of highly successful products received good reviews, 60% of moderately successful products received good reviews, and 10% of poor products received good reviews. In addition, 40% of products have been highly successful, 35% have been moderately successful, and 25% have been poor products.
    (4) What is the probability that a product attains a good review?
    (5) If a new design attains a good review, what is the probability that it will be a highly successful product?
    (6) If a product does not attain a good review, what is the probability that it will be a highly successful product?

[AU A/M 2024] [pg. no. 85]

76. Assume that each of your calls to a popular radio station has a probability of 0.02 of connecting, that is, of not obtaining a busy signal. Assume that your calls are independent.
    (4) What is the probability that your first call that connects is your tenth call?
    (5) What is the probability that it requires more than five calls for you to connect?
    (6) What is the mean number of calls needed to connect?

[AU A/M 2024] [pg. no. 86]

77. The thickness of a flange on an aircraft component is uniformly distributed between 0.95 and 1.05 millimeters.
    (1) Determine the cumulative distribution function of flange thickness.
    (2) Determine the proportion of flanges that exceeds 1.02 millimeters.
    (3) What thickness is exceeded by 90% of the flanges?
    (4) Determine the mean and variance of flange thickness.

[pg. no. 87][AU A/M 2024]

78. The life of a semiconductor laser at a constant power is normally distributed with a mean of 7000 hours and a standard deviation of 600 hours.
    (4) What is the probability that a laser fails before 5000 hours?
    (5) What is the life in hours that 95% of the lasers exceed?
    (6) If three lasers are used in a product and they are assumed to fail independently, what is the probability that all three are still operating after 7000 hours?

[pg. no. 88] [AU A/M 2024]

79. The probabilities of X ,Y and Z becoming managers are 4/9, 2/9 and 1/3 respectively. The probabilities that the Bonus scheme will be introduced if X, Y, Z becomes managers are 3/10,

1/2, and 4/5 respectively.

(3) What is the probability that the Bonus scheme will be introduced?

(4) If the Bonus scheme has been introduced, what is the probability that the manager appointed was X?

[pg. no. 89] [AU N/D 2024]

UNIT-II (TWO DIMENSIONAL RANDOM VARIABLES)

Part-A

1. The bivariate random variable X and Y has the p.d.f **[AU M/J 2009]**

$$f(x,y)=\begin{cases} Kx^2(8-y), x<y<2x \\ \qquad\qquad 0\le x<2 \end{cases} \qquad Find \quad K.$$

Solution:

We know that if $f(x,y)$ is a p.d.f. then

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

$$\int_0^2\int_x^{2x} Kx^2(8-y)dydx=1$$

$$K\int_0^2 x^2\left(8y-\frac{y^2}{2}\right)_x^{2x}dx=1$$

$$K\int_0^2 x^2\left(16x-\frac{4x^2}{2}-8x+\frac{x^2}{2}\right)dx=1$$

$$K\int_0^2\left(16x^3-2x^4-8x^3+\frac{x^4}{2}\right)dx=1$$

$$K\int_0^2\left(8x^3-\frac{3}{2}x^4\right)dx=1\Rightarrow K\left[\frac{8x^4}{4}-\frac{3}{2}\frac{x^5}{5}\right]_0^2=1$$

$$K\left[32-\frac{48}{5}\right]=1 \quad\Rightarrow K\left[\frac{160-48}{5}\right]=1$$

$$K\left[\frac{112}{5}\right]=1 \quad\Rightarrow K=\frac{5}{112}$$

2. The joint p.d.f of R.V X and Y is given by $f(x,y)=kxye^{-\left(x^2+y^2\right)}$, $x>0, y>0$ **. Find the value of k.** **[AU N/D 2013]**

Solution:

We know that if $f(x,y)$ is a p.d.f, then

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

$$\int_0^{\infty}\int_0^{\infty} kxye^{-\left(x^2+y^2\right)}dydx=1 \qquad [\because x>0,\ y>0]$$

$$\int_0^{\infty}\int_0^{\infty} kxye^{-x^2}e^{-y^2}dydx=1$$

$$k\int_0^\infty ye^{-y^2}dy\int_0^\infty xe^{-x^2}dx = 1$$

$$k\frac{1}{2}\frac{1}{2} = 1 \quad \because \left[\int_0^\infty xe^{-x^2}dx = \frac{1}{2}\right]$$

$$k = 4$$

3. If X and Y have joint p.d.f $f(x,y) = \begin{cases} x+y, & 0<x<1,\ 0<y<1 \\ 0, & otherwise \end{cases}$ **Check whether X and Y are independent.**

Solution:

The marginal probability function of 'X' is

$$f_X(x) = f(x) = \int_{-\infty}^{\infty} f(x,y)dy$$

$$= \int_0^1 (x+y)dy$$

$$= \left[xy + \frac{y^2}{2}\right]_0^1 = x + \frac{1}{2}$$

$$\therefore f(x) = x + \frac{1}{2}.........(1)$$

The marginal probability function of 'Y' is

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

$$= \int_0^1 (x+y)dx$$

$$= \left[xy + \frac{x^2}{2}\right]_0^1 = y + \frac{1}{2}$$

$$\therefore f(y) = y + \frac{1}{2}.............(2)$$

Now $f(x).f(y) = \left(x+\frac{1}{2}\right)\left(\frac{1}{2}+y\right) \neq f(x,y)$ $[Using\ (1)\ and\ (2)]$

$$\therefore X\ and\ Y\ are\ not\ independent$$

4. Let X and Y have j.d.f $f(x,y) = 2,\ 0<x<y<1$**. Find the marginal density function.**

Solution: **[AU M/J 2013]**

Marginal density function of X is given by

$$f_X(x) = f(x) = \int_{-\infty}^{\infty} f(x,y)dy$$

$$= \int_x^1 2dy$$

$$= 2(y)_x^1 = 2(1-x)$$

$$\therefore M.d.f\ of\ 'X' is\ f(x) = 2(1-x),\ 0<x<1$$

Marginal density function of Y is given by

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

$$= \int_x^1 2dx$$

$$= 2(x)_0^y = 2(y-0)$$

$$\therefore M.d.f\ of\ 'Y' is\ f(y) = 2y,\ 0<y<1$$

5. The j.d.f of the random variables X and Y is given by $f(x,y) = \begin{cases} 8xy, & 0<x<1, 0<y<x \\ 0, & otherwise \end{cases}$**.**

find $f_X(x)$**.** **[AU N/D 2017]**

Solution:

The M.d.f. of 'X' is

$$f_X(x) = f(x) = \int_{-\infty}^{\infty} f(x,y)dy$$

$$= \int_0^x 8xydy$$

$$= 8x\left(\frac{y^2}{2}\right)_0^x = 4x\left(x^2 - 0\right)$$

$$\therefore M.d.f \text{ of } 'X' \text{ is } f(x) = 4x^3 \ , 0 < x < 1$$

6. Given $f(x,y) = \begin{cases} Cx(x-y), \ 0 < x < 2, -x < y < x \\ 0, \quad otherwise \end{cases}$ **, find C.**

Solution:

We know that $f(x,y)$ should satisfy

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

$$\int_0^2\int_{-x}^{x} Cx(x-y)dydx = 1$$

$$C\int_0^2\left[x^2y - x\frac{y^2}{2}\right]_{-x}^{x} dx = 1$$

$$C\int_0^2\left[\left(x^3 - \frac{x^3}{2}\right) - \left(-x^3 - \frac{x^3}{2}\right)\right]dx = 1$$

$$C\int_0^2\left[2x^3\right]dx = 1$$

$$2C\left[\frac{x^4}{4}\right]_0^2 = 1$$

$$2C\left[\frac{16}{4} - 0\right] = 1 \ , \Rightarrow 8C = 1 \Rightarrow C = \frac{1}{8}$$

7. The joint p.d.f a bivariate R.V (X,Y) is given by

$f(x,y) = \begin{cases} Kxy, \ 0 < x < 1, 0 < y < 1 \\ 0, \quad otherwise \end{cases}$ **Find K.** **[AU M/J 2013]**

Solution:

Since $f(x,y)$ is a p.d.f, we have

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

$$\int_0^1\int_0^1 Kxydxdy = 1$$

$$\int_0^1 Ky\left(\frac{x^2}{2}\right)_0^1 dy = 1 \Rightarrow \int_0^1 K\frac{y}{2}dy = 1$$

$$\frac{K}{2}\int_0^1 ydy = 1 \Rightarrow \frac{K}{2}\left(\frac{y^2}{2}\right)_0^1 = 1$$

$$\frac{K}{2}\left(\frac{1}{2}\right) = 1$$

$$K = 4$$

8. The joint p.d.f a bivariate R.V (X,Y) is given by

$f(x,y) = \frac{1}{4},\ 0 < x, y < 2$ **Find** $P(X+Y \le 1)$. **[AU M/J 2007]**

Solution:

$$P(X+Y \le 1) = \int_0^1\int_0^{1-y} f(x,y)dxdy$$

$$= \int_0^1\int_0^{1-y} \frac{1}{4}dxdy$$

$$= \frac{1}{4}\int_0^1 (x)_0^{1-y} dy = \frac{1}{4}\int_0^1 (1-y)dy$$

$$= \frac{1}{4}\left[y - \frac{y^2}{2}\right]_0^1 = \frac{1}{4}\left[1 - \frac{1}{2}\right]$$

$$= \frac{1}{4}\left[\frac{2-1}{2}\right] = \frac{1}{8}$$

$$\therefore \quad P(X+Y \le 1) = \frac{1}{8}$$

9. The joint p.d.f a bivariate R.V (X,Y) is given by

$$f(x,y) = \begin{cases} \frac{xy}{96}, & 0 < x < 4, 1 < y < 5 \\ 0, & otherwise \end{cases}$$ **Find E(X).**

Solution:

We know that

$$E(X) = \int_{-\infty}^{\infty}\int_{-\infty}^{\infty} xf(x,y)dxdy$$

$$= \int_{1}^{5}\int_{0}^{4} x\left(\frac{xy}{96}\right)dxdy$$

$$= \int_{1}^{5}\int_{0}^{4} \left(\frac{x^2y}{96}\right)dxdy = \frac{1}{96}\int_{1}^{5} y\left(\frac{x^3}{3}\right)_0^4 dy$$

$$= \frac{64}{96}\int_{1}^{5}\left(\frac{y}{3}\right)dy = \frac{64}{288}\int_{1}^{5} ydy$$

$$\therefore \quad E(X) = \frac{8}{3}$$

10. Let X be a random variable with p.d.f

$f(x) = \frac{1}{2}, \ -1 \le x \le 1$ **and let** $Y = X^2$ **Find E(Y).** **[AU JAN 2008]**

Solution:

$$E(X) = \int_{-\infty}^{\infty} xf(x)dx = \int_{-1}^{1} x\frac{1}{2}dx$$

$$= \frac{1}{2}\left(\frac{x^2}{2}\right)_{-1}^{1} = \frac{1}{2}\left[\frac{1}{2} - \frac{1}{2}\right]$$

$$E(X) = 0$$

$$E(Y) = E(X^2) = \int_{-\infty}^{\infty} x^2 f(x)dx$$

$$= \int_{-1}^{1} x^2 \frac{1}{2} dx$$

$$= \frac{1}{2}\left(\frac{x^3}{3}\right)_{-1}^{1} = \frac{1}{2}\left(\frac{1}{3} + \frac{1}{3}\right) = \frac{1}{3}$$

11. If the joint pdf of (X,Y) is given by $f(x,y) = x + y, \ 0 \le x, y \le 1$**, find E(XY).**

Solution:

Given $f(x,y) = x + y, \ 0 \le x, y \le 1$

Now, $E(XY) = \int_{-\infty}^{\infty}\int_{-\infty}^{\infty} xyf(x,y)dxdy$

$$= \int_0^1 \int_0^1 xy(x+y)dxdy$$

$$= \int_0^1 \int_0^1 (x^2 y + xy^2)dxdy$$

$$= \int_0^1 \left[\frac{x^3 y}{3} + \frac{x^2 y^2}{2} \right]_0^1 dy = \int_0^1 \left[\frac{y}{3} + \frac{y^2}{2} \right] dy$$

$$= \left[\frac{y^2}{6} + \frac{y^3}{6} \right]_0^1 = \frac{1}{6} + \frac{1}{6} = \frac{2}{6} = \frac{1}{3}$$

12. Let X and Y be random variables having joint density function

$$f(x,y) = \begin{cases} \frac{3}{2}\left(x^2 + y^2\right), & 0 < x < 1, 0 < y < 1 \\ 0, & otherwise \end{cases},$$ **Find** $E(XY)$**.**

Solution:

Given $f(x,y) = \begin{cases} \frac{3}{2}\left(x^2 + y^2\right), & 0 < x < 1, 0 < y < 1 \\ 0, & otherwise \end{cases}$

$$E(XY) = \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} xyf(x,y)dxdy \quad = \int_0^1 \int_0^1 xy \frac{3}{2}\left(x^2 + y^2\right)dxdy$$

$$= \frac{3}{2} \int_0^1 \int_0^1 \left(x^3 y + xy^3\right)dxdy$$

$$= \frac{3}{2} \int_0^1 \left(\frac{x^4 y}{4} + \frac{x^2 y^3}{2} \right)_0^1 dy$$

$$= \frac{3}{2} \int_0^1 \left(\frac{y}{4} + \frac{y^3}{2} \right) dy = \frac{3}{2} \left(\frac{y^2}{8} + \frac{y^4}{8} \right)_0^1$$

$$= \frac{3}{2} \left[\frac{1}{8} + \frac{1}{8} - 0 - 0 \right] = \frac{3}{16}[1+1]$$

$$\therefore \quad E(XY) = \frac{3}{8}$$

13. Find the rank correlation coefficient from the following data.

Rank in X	**1**	**2**	**3**	**4**	**5**	**6**	**7**
Rank in Y	**4**	**3**	**1**	**2**	**6**	**5**	**7**

Solution:

X	Y	$d_i = x_i - y_i$	d_i^2
1	4	-3	9
2	3	-1	1
3	1	2	4
4	2	2	4
5	6	-1	1
6	5	1	1
7	7	0	0
		$\sum d_i = 0$	$\sum d_i^2 = 20$

Rank Correlation coefficient, $r(X,Y) = 1 - \dfrac{6\sum d_i^2}{n(n^2-1)}$

$$= 1 - \frac{6 \times 20}{7(49-1)} = 1 - \frac{120}{336}$$

$$= 0.6429$$

14. Find the acute angle between the two lines of regression.

Solution: **[AU N/D 2012]**

Angle between the lines, is given by

$$\tan\theta = \frac{1-r^2}{r}\left(\frac{\sigma_X \sigma_Y}{\sigma_X^2 + \sigma_Y^2}\right)$$

15. State the equation of the two regression lines, what is the angle between them?

Solution:

The line of regression of Y on X is given by

$$y - \bar{y} = r\frac{\sigma_Y}{\sigma_X}\left(x - \bar{x}\right) \text{......................(1)}$$

Where r is the correlation coefficient

σ_X, σ_Y are standard deviations

The line of regression of Y on X is given by

$$x - \bar{x} = r\frac{\sigma_X}{\sigma_Y}\left(y - \bar{y}\right) \text{........................(2)}$$

The angle between the two lines of regression (1) & (2) is given by

$$\tan\theta = \frac{1-r^2}{r}\left(\frac{\sigma_X \sigma_Y}{\sigma_X^2 + \sigma_Y^2}\right)$$

16. If X and Y are independently related, find the angle between the two regression lines.

Solution:

Let

$$Y = a + b_{YX} X \ , \quad X = a + b_{XY} Y$$

$$\therefore \qquad \theta = \frac{1 - b_{YX} b_{XY}}{b_{YX} + b_{XY}}$$

17. If X and Y are independent random variables with variance 2 and 3.

Find the variance of 3X+4Y.

Solution:

Given X and Y are independent RVs with Variance 2 and 3.

$$Var(X) = 2 \quad and \quad Var(Y) = 3$$

Consider

$$\begin{aligned} Var(3X+4Y) &= 3^2 Var(X) + 4^2 Var(Y) \\ &= 9 \times 2 + 16 \times 3 \\ &= 18 + 48 = 66 \end{aligned}$$

18. Prove that the correlation coefficient ρ_{xy} takes value in the range -1 to 1. [AU A/M 2021]

Solution:

We know that $r(X,Y) = \dfrac{Cov(X,Y)}{\sigma_X . \sigma_Y}$

$$= \frac{\frac{1}{n}\sum_i \left(x_i - \overline{x}\right)\left(y_i - \overline{y}\right)}{\sqrt{\frac{1}{n}\sum_i \left(x_i - \overline{x}\right)^2 \sum_i \left(y_i - \overline{y}\right)^2}}$$

$$= \frac{\sum_i a_i b_i}{\sqrt{\sum_i a_i^2}\sqrt{\sum_i b_i^2}}$$

Where $a_i = x_i - \overline{x}, b_i = y_i - \overline{y}$

Squaring on both sides, we get

$$r^2(X,Y) = \frac{\left(\sum_i a_i b_i\right)^2}{\left(\sum_i a_i^2\right)\left(\sum_i b_i^2\right)} \text{......................(1)}$$

By Schwartz inequality,

$$\left(\sum_i a_i b_i\right)^2 \le \left(\sum_i a_i^2\right)\left(\sum_i b_i^2\right) \ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots(2)$$

Substituting (2) in (1), we get

$$\frac{\left(\sum_i a_i b_i\right)^2}{\left(\sum_i a_i^2\right)\left(\sum_i b_i^2\right)} \le 1 \ldots\ldots\ldots\ldots\ldots\ldots\ldots(3)$$

Using (3) in (1), we get

$$r^2(X,Y) \le 1$$

$$|r| \le 1$$

$$-1 \le r \le 1$$

19. The tangent of the angle between the lines of regression Y on X and X on Y is 0.6 and $\sigma_X = \frac{1}{2}\sigma_Y$, find the correlation coefficient between X and Y. [AU A/M 2005]

Solution:

$$\tan\theta = \frac{1-r^2}{r}\left(\frac{\sigma_X\sigma_Y}{\sigma_X^2+\sigma_Y^2}\right)$$

$$0.6 = \frac{1-r^2}{r}\left(\frac{\frac{1}{2}\sigma_Y}{\frac{1}{4}\sigma_Y^2+\sigma_Y^2}\right)$$

$$= \frac{1-r^2}{r}\left(\frac{\frac{1}{2}\sigma_Y}{\sigma_Y^2\left(\frac{1}{4}+1\right)}\right)$$

$$= \frac{1-r^2}{r}\left(\frac{\frac{1}{2}}{\left(\frac{5}{4}\right)}\right) = \frac{2(1-r^2)}{5r}$$

$$\frac{3}{2} = \frac{1-r^2}{r} \Rightarrow 2r^2+3r-2=0$$

$$(2r-1)(2+1)=0 \quad \Rightarrow \quad r=\frac{1}{2} \ or \ -2$$

$$r=-2 \ is \ not \ possible$$

$$\therefore \qquad r=\frac{1}{2}$$

20. If the pdf of X is $f_X(x) = e^{-x}$, $x > 0$, find the pdf of $Y = 2X + 1$.

Solution:

Given pdf of 'X' is $f_X(x) = e^{-x}$, $x > 0$..........................(1)

Given $Y = 2X + 1$...(2)

$$y = 2x + 1$$

$$\Rightarrow \quad x = \frac{y-1}{2} = f(y) \text{..........................(3)}$$

$$(3) \Rightarrow \quad \frac{dx}{dy} = \frac{1}{2}$$

$$\therefore \quad \left|\frac{dx}{dy}\right| = \frac{1}{2} \text{............................(4)}$$

$$f_Y(y) = f_X(x)\left|\frac{dx}{dy}\right|$$

$$= e^{-x}\frac{1}{2},\ x > 0 \qquad [u\sin g\ \ (1)\ \ and\ \ (4)]$$

$$= e^{-\left(\frac{y-1}{2}\right)} \qquad \left[\therefore\ \ x = \frac{y-1}{2}\right] \text{....................(5)}$$

Since $x > 0 \Rightarrow \dfrac{y-1}{2} > 0$

$$\Rightarrow y - 1 > 0$$

$$\Rightarrow y > 0 \text{..................................(6)}$$

Using (5) and (6) we get

$$f_Y(y) = e^{-\left(\frac{y-1}{2}\right)},\ y > 1$$

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

21. If X is a uniformly distributed RV in $\left(\frac{-\pi}{2}, \frac{\pi}{2}\right)$, find the pdf of $Y = \tan X$.

Solution: **[AU M/J 2006]**

Since X is uniformly distributed in $\left(\frac{-\pi}{2}, \frac{\pi}{2}\right)$, the p.d.f of 'X' is

$$f_X(x) = \frac{1}{\frac{\pi}{2} + \frac{\pi}{2}},\quad -\frac{\pi}{2} < x < \frac{\pi}{2}$$

$$= \frac{1}{\frac{\pi}{2}},\ -\frac{\pi}{2} < x < \frac{\pi}{2} \text{.......................(1)}$$

$$Given \quad Y = \tan X \text{..................................}(2)$$

$$y = \tan x \Rightarrow x = \tan^{-1}(y) \text{.........................}(3)$$

$$\Rightarrow \frac{dx}{dy} = \frac{1}{1+y^2} \quad [using\ (3)]$$

$$\therefore \quad \left|\frac{dx}{dy}\right| = \frac{1}{1+y^2} \text{..................................}(4)$$

$$f_Y(y) = f_X(x)\left|\frac{dx}{dy}\right|$$

$$= f_X(x)\left|\frac{dx}{dy}\right| \text{..................................}(5)$$

$$Since \quad x \in \left(\frac{-\pi}{2}, \frac{\pi}{2}\right), \tan x \in (-\infty, \infty)$$

$$\Rightarrow \quad y \in (-\infty, \infty) \text{..................................}(6)$$

Using (5) and (6), we get

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

22. Give a real life example each for positive correlation and negative correlation.

Solution: **[AU A/M 2010]**

Positive correlation: If the demand increases, then the price, will also increase which is an example for positive correlation.

Negative correlation: If the availability increases then the demand will decrease which is an example for negative correlation.

23. If there is no linear correlation between two random variable X and Y, then what can you say about the regression lines?

Solution:

When r=0, that is, when there is no correlation between x and y.

$\tan\theta = \infty \quad (or) \quad \theta = \frac{\pi}{2}$ and so the regression lines are perpendicular.

24. Given the random variable X with density function $f(x) = \begin{cases} 2x, & 0 < x < 1 \\ 0, & elsewhere \end{cases}$.

Find the probability density function of $Y = 8X^3$**.** **[AU N/D 2013, '14]**

Solution:

Given: $f(x) = \begin{cases} 2x, & 0 < x < 1 \\ 0, & elsewhere \end{cases}$

$$y = 8x^3, \quad x = \frac{1}{2}y^{1/3}$$

$$f_Y(y) = f_X(x)\left|\frac{dx}{dy}\right|$$
$$= 2\left(\frac{1}{2}y^{1/3}\right)\left|\frac{1}{2}\frac{1}{3}y^{1/3-1}\right|$$
$$= y^{1/3}\frac{1}{6}y^{-2/3} = \frac{1}{6}y^{-1/3},\quad 0<y<8$$

25. Given the two regression lines 3X+12Y=19,3Y+9X=46, find the coefficient of correlation between X and Y. [AU M/J 2013, N/D 2016]

Solution:

$$3x+12y=19 \qquad\qquad 3x+9y=46$$
$$\Rightarrow 12y = 19-3x \qquad\qquad \Rightarrow 9x = 46-3y$$
$$\Rightarrow y = \frac{19}{12}-\frac{1}{4}x \qquad\qquad \Rightarrow x = \frac{46}{9}-\frac{1}{3}y$$
$$\Rightarrow b_{yx} = -\frac{1}{4} \qquad\qquad \Rightarrow b_{xy} = -\frac{1}{3}$$
$$r^2 = b_{xy}b_{yx} = \left(-\frac{1}{3}\right)\left(-\frac{1}{4}\right) = \frac{1}{12}$$
$$r = 0.29$$

26. The regression equation of X on Y and Y on X are respectively 5x-y=22 and 64x-45y=24. Find the means of X and Y. [AU M/J 2012]

Solution:

Since both the regression equations pass through $(\bar{x},\bar{y})$. We get

$$5\bar{x}-\bar{y} = 22 \ldots\ldots\ldots\ldots\ldots\ldots\ldots(1)$$
$$64\bar{x}-45\bar{y} = 24 \ldots\ldots\ldots\ldots\ldots\ldots\ldots(2)$$
$$(1)\times 45 \Rightarrow \quad 225\bar{x}-45\bar{y} = 990 \ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots(3)$$
$$(3)-(2) \Rightarrow \quad 161\bar{x} = 966 \Rightarrow \bar{x} = \frac{966}{161} = 6$$
$$\therefore \quad The \quad Mean \quad value \quad of \quad X = 6$$
$$Put \ \bar{x} = 6 \ \ in \ (1)$$
$$(1) \Rightarrow 5(6)-\bar{y} = 22$$
$$\bar{y} = 30-22 = 8$$
$$\therefore \ the \ \ mean \ \ value \ \ of \ Y = 8$$

27. When will the two regression lines be (i) a right angles (ii) coincident?

Solution: **[AU N/D 2012]**

(i) If r=0, the regression lines are at right angles.

(ii) If r=$\pm$1, the regression lines coincide.

28. A small college has 90 male and 30 female professors. An ad-hoc committee of 5 is selected at random to unite the vision and mission of the college. If X and y are the number of men and women in the committee, respectively, what is the joint probability mass function of X and Y?

Solution: **[AU N/D 2012]**

$$p(x,y)=\frac{90C_x 30C_y}{120C_5},\quad 0<x<90 \quad and \quad 0<y<30$$

29. Let X and Y have joint pdf $f(x,y)=\begin{cases}\frac{1}{\pi}, & x^2+y^2\le 1\\ 0, & otherwise\end{cases}$ **Are X and Y independent? Why?**

Solution:

Given $f(x,y)=\begin{cases}\frac{1}{\pi}, & x^2+y^2\le 1\\ 0, & otherwise\end{cases}$

The marginal density functions $f(x)$ and $f(y)$ of X and Y are given by,

$$f(x)=\int_{-\infty}^{\infty} f(x,y)dy=\int_{-\sqrt{1-x^2}}^{\sqrt{1-x^2}}\frac{1}{\pi}dy=\frac{2}{\pi}\sqrt{1-x^2}, 0\le x\le 1$$

Similarly,

$$f(y)=\int_{-\infty}^{\infty} f(x,y)dx=\int_{-\sqrt{1-y^2}}^{\sqrt{1-y^2}}\frac{1}{\pi}dx=\frac{2}{\pi}\sqrt{1-y^2}, 0\le y\le 1$$

Since $f(x).f(y)=\frac{4}{\pi^2}\sqrt{1-x^2}\sqrt{1-y^2}\ne f(x,y)$

Therefore X and Y are not independent.

30. The j.p.d.f of the RV's X and Y is given by

$$f(x,y)=\begin{cases}25, & 0.95<x<1.15, 0.95<y<1.15\\ 0\ , & otherwise\end{cases} \qquad .Find\ P(XY<1)$$

Solution:

$$P(XY<1)=\int_{0.95}^{1.15}\int_{0.95}^{1/y}25\,dx\,dy$$

$$=\int_{0.95}^{1.15}\left[25x\right]_{0.95}^{1/y}dy=\int_{0.95}^{1.15}\left[\frac{25}{y}-25(0.95)\right]dy$$

31. State the basic properties of joint distribution of (X,Y) when X and Y are random variable.
Solution: **[A.U. A/M 2005, M/J '14 PRP]**

1. $F_{X,Y}(-\infty,-\infty)=0;\quad F_{X,Y}(-\infty,y)=0;$ and $F_{X,Y}(x,-\infty)=0.$
2. $F_{X,Y}(\infty,\infty)=1$
3. $F_{X,Y}(x_2,y_2)-F_{X,Y}(x_1,y_1)+F_{X,Y}(x_1,y_1)-F_{XY}(x_2,y_1)$

$$=P\{x_1<X\le x_2;\ y_1<Y\le y_2\}\ge 0.$$

32. If two random variables X and Y have probability density function $f(x,y)=ke^{-(2x+y)}$ *for* $x>0, y>0$, **Evaluate 'k'.**

Solution:

By the property of joint pdf ,

$$\int_{x>0}\int_{y>0} f(x,y)\,dx\,dy=1$$

$$\text{(ie)}\ \int_0^\infty\int_0^\infty ke^{-(2x+y)}dx\,dy=1 \Rightarrow k\int_0^\infty\int_0^\infty e^{-2x}e^{-2y}dx\,dy=1$$

$$k\left[\int_0^\infty e^{-2x}\,dx\right]\left[\int_0^\infty e^{-y}\,dy\right]=1\Rightarrow k\left[\frac{e^{-2x}}{2}\right]_0^\infty\left[\frac{e^{-y}}{-1}\right]_0^\infty=1$$

$$k\left[0-\frac{1}{2}\right]\left[0-\frac{1}{-1}\right]=1\quad\Rightarrow k\left[-\frac{1}{2}\right][1]=1$$

$$k=-2$$

33. If the joint pdf of (X,Y) is $f(x,y)=\begin{cases}\frac{1}{4} & \textit{for } 0\le x,\ y\le 2,\\ 0 & \textit{otherwise}\end{cases}$ *find* $P(x+y\le 1).$

Solution:

$$P(x+y\le 1).=\int_0^1\int_0^{1-y} f(x,y)\,dx\,dy=\int_0^1\int_0^{1-y}\frac{1}{4}dx\,dy$$

$$=\frac{1}{4}\int_0^1[x]_0^{1-y}dy=\frac{1}{4}\int_0^1(1-y)dy=\frac{1}{4}\left[y-\frac{y^2}{2}\right]_0^1$$

$$=\frac{1}{4}[(1-\frac{1}{2})]-(0-0)]=\frac{1}{4}\left(\frac{1}{2}\right)=\frac{1}{8}.$$

34. Define the joint pmf of two dimensional discrete random variables. **[AU N/D 14 (PRP)]**
Solution:

For two discrete random variable X and Y, we write the probability that X will take the values x_i and Y will take the values y_j as $P\lfloor X=x_i, Y=y_j\rfloor$. Consequently $P\lfloor X=x_i, Y=y_j\rfloor$ is the probability of the intersection of the events $X=x_i$ *and* $Y=y_j$

$$P[X = x_i, Y = y_j] = P[(X = x_i) \cap (Y = y_j)]$$

The function $P[X = x_i, Y = y_j] = p(x_i, y_j)$ is called the joint probability mass function for discrete random variables X and Y is denoted by p_{ij}

35. The joint pdf of two dimensional random variable X,Y is given by $f(x,y) = \{k\, x\, e^{-y} \quad 0 < x < 2, y > 0$ Find (i) the value of K (ii) the marginal p.d.f of X.

Solution: **[AU N/D 14' A/M 2015, 2021]**

Given $f(x,y) = \{k\, x\, e^{-y} \quad 0 < x < 2, y > 0$

$$\int_{-\infty}^{\infty}\int_{-\infty}^{\infty} f(x,y)\,dx\,dy = 1 \quad \Rightarrow \int_{0}^{\infty}\int_{0}^{2} k\,x\,e^{-y}\,dx\,dy = 1$$

$$k\int_{0}^{\infty} e^{-y}\,dy \int_{0}^{2} x\,dx = 1 \Rightarrow k\left[\frac{e^{-y}}{-1}\right]_0^{\infty}\left[\frac{x^2}{2}\right]_0^2 = 1$$

$$k(1)\left[\frac{4^2}{2}\right] = 1 \Rightarrow k = \frac{1}{2}$$

$$f(x) = \int_{0}^{\infty} f(x,y)dy = \int_{0}^{\infty}\frac{1}{2}xe^{-y}dy = \frac{1}{2}x\left[-e^{-y}\right]_0^{\infty} = \frac{1}{2}x[0+1] = \frac{x}{2}$$

36. Comment on the statement "If $Cov(X,Y) = 0$ then X and Y are uncorrelated".

Solution: **[AU N/D 14 A/M 2018]**

If $Cov(X,Y) = 0$

$\Rightarrow X \; and \; Y$ are independent

$$\rho(x,y) = \frac{Cov(X,Y)}{\sigma_x\ \sigma_y} = 0$$

When $\rho(x,y) = 0$ then X and Y are uncorrelated.

37. Find the value of k, if the joint density function of (X,Y) is given by
$f(x,y) = \{k(1-x)(1-y) \quad 0 < x < 4, 1 < y < 5$ **[AU M/J 14 (PQT)]**

Solution:

Given $f(x,y) = \{k(1-x)(1-y) \quad 0 < x < 4, 1 < y < 5$

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

$$\int_{0}^{4}\int_{1}^{5} k\,(1-x)(1-y)\;dx\,dy = 1$$

$$k\int_{0}^{4}(1-x)dx\int_{1}^{5}(1-y)\,dy = 1$$

$$k\left[\frac{(1-x)^2}{-1(2)}\right]_0^4\left[\frac{(1-y)^2}{-1(2)}\right]_1^5=1$$

$$k\left[\frac{9}{-2}+\frac{1}{2}\right]\left[\frac{(-4)^2}{-2}-0\right]=1$$

$$k\left[\frac{-9+1}{2}\right]\left[\frac{16}{2}\right]=1$$

$$k\left[\frac{-8}{2}\right]\left[\frac{16}{2}\right]=1$$

$$k=\frac{1}{32}$$

38. Given that joint probability density function of (X,Y) is given by

$f(x,y)=\begin{cases}\frac{1}{6}, & 0<x<2, 0<y<3\end{cases}$ **Determine the marginal density? [AU M/J 14 (PQT)]**

Solution:

The marginal density function of X is $f_x(X)=\int_{-\infty}^{\infty} f(x,y)\,dy$

$$f_x(X)=\int_0^3 \frac{1}{6}\,dy$$

$$=\frac{1}{6}[y]_0^3$$

$$=\frac{1}{2}$$

The marginal density function of Y is $f_y(Y)=\int_{-\infty}^{\infty} f(x,y)\,dx$

$$f_y(Y)=\int_0^2 \frac{1}{6}\,dx$$

$$=\frac{1}{6}[x]_0^2$$

$$=\frac{1}{3}$$

39. The joint p.d.f a bivariate R.V (X,Y) is given by

$f(x,y)=\begin{cases}4xy, & 0<x<1, 0<y<1\\ 0, & elsewhere\end{cases}$ **Find $P(X+Y\le 1)$. [AU N/D 2015]**

Solution:

$$P(X+Y\le 1)=\int_0^1\int_0^{1-y} f(x,y)dxdy$$

$$=\int_0^1\int_0^{1-y} 4xydxdy$$

$$=\int_0^1\left(4y\frac{x^2}{2}\right)_0^{1-y}dy=2\int_0^1(1-y)^2\,ydy=2\int_0^1\left(1-2y+y^2\right)ydy$$

$$=2\left[\frac{y^2}{2}-\frac{2y^3}{3}+\frac{y^4}{4}\right]_0^1=2\left[\frac{1}{2}-\frac{2}{3}+\frac{1}{4}\right]=\frac{1}{6}$$

$$\therefore\quad P(X+Y\le 1)=\frac{1}{6}$$

40. The joint probability density function of a random variable (X,Y) is $f(x,y)=ke^{-(2x+3y)},\ x\ge 0,\ y\ge 0$. Find the value of k. **[AU N/D 2016]**

Solution:

We know that $\int_{-\infty}^{\infty} f(x,y)dxdy=1$

$$\int_0^{\infty} ke^{-(2x+3y)}dxdy=1 \Rightarrow k\int_0^{\infty} e^{-2x}e^{-3x}dxdy=1$$

$$k\left(\frac{e^{-2x}}{-2}\right)_0^{\infty}\left(\frac{e^{-3x}}{-3}\right)_0^{\infty}=1$$

$$k\left(\frac{1}{2}\right)\left(\frac{1}{3}\right)=1 \quad\Rightarrow k=6$$

41. Write any two properties of joint cumulative distribution function.[AU N/D 2016]

Solution:

$(i)\quad 0\le F(x,y)\le 1,\ -\infty<x<\infty,\ -\infty<y<\infty$

$(ii)\quad \lim_{x,y\to\infty} F(x,y)=1$

$(iii)\quad F(-\infty,y)=F(x,-\infty)=0$

$(iv)\quad F(x_1,y_1)\le F(x_1,y_1)\quad for\quad x_1\le x_2\ \&\ y_1\le y_2$

42. The joint p.d.f of R.V (X,Y)is given as $f(x,y)=\begin{cases}\dfrac{1}{x}, & 0<y<x\le 1\\ 0, & otherwise\end{cases}$ **. Find the marginal p.d.f of Y.**

Solution: **[AU M/J 2016]**

The Marginal density function of Y is given by

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

$$= \int_{y}^{1} \frac{1}{x} dx$$

$$= (\log x)_y^1 = 0 - \log y$$

$$\therefore M.d.f \ of \ 'Y' is \ f(y) = -\log y, \ 0 < y < 1$$

43. Let X and Y be two independent random variables with Var(X)=9 and Var(Y)=3.

Find the Var(4X-2Y+6)

Solution: **[AU M/J 2016]**

Given X and Y are independent RVs with Variance 9 and 3.

$$Var(X) = 9 \quad and \quad Var(Y) = 3$$

Consider

$$Var(4X - 2Y + 6) = 16Var(X) + 4Var(Y)$$

$$= 16 \times 9 + 4 \times 3$$

$$= 144 + 12 = 156$$

44. Two random variables X and Y have the following joint probability density function

$f(x, y) = \dfrac{6 - x - y}{8}, \ 0 \le x \le 2, 2 \le y \le 4$, **find** $P[x + y < 3]$

Solution: **[AU N/D 2017]**

Given $f(x, y) = \dfrac{6 - x - y}{8}, \ 0 \le x \le 2, 2 \le y \le 4$

$$P[x + y < 3] = \int_2^3 \int_0^{3-y} f(x, y)dxdy$$

$$= \int_2^3 \int_0^{3-y} \frac{1}{8}(6 - x - y)dxdy$$

$$= \frac{1}{8}\int_2^3 \left[6x - \frac{x^2}{2} - xy\right]_0^{3-y} dy = \frac{1}{8}\int_2^3 \left[6(3 - y) - \frac{(3 - y)^2}{2} - y(3 - y)\right]dy$$

$$= \frac{1}{8}\int_2^3 \left[18 - 9y + y^2 - \frac{(3 - y)^2}{2}\right]dy = \frac{1}{8}\left[18y - 9\frac{y^2}{2} + \frac{y^3}{3} + \frac{(3 - y)^3}{3}\right]_2^3$$

$$= \frac{1}{8}\left[18 - 9\frac{5}{2} + 9 - \frac{8}{3} - \frac{1}{6}\right] = \frac{1}{8}\left[\frac{162 - 135 - 16 - 1}{6}\right] = \frac{1}{8}\left(\frac{10}{6}\right) = \frac{5}{24}$$

45. Define conditional distribution for two-dimensional discrete and continuous random variables.

Solution: **[AU A/M 2017, '18]**

Discrete Case:

$$P(X/Y)=\frac{P(X\cap Y)}{P(Y)}$$

$$P(Y/X)=\frac{P(X\cap Y)}{P(X)}$$

Continuous Case:

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

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

46. If $X = R\cos\phi$ *and* $Y = R\sin\phi$**, how are the joint probability density function of** (X,Y) *and* (R,ϕ) **are related?** **[AU A/M 2017]**

Solution:

The joint pdf of (R,Ø) is given by

$$f_{r\theta}(r,\theta)=f_{xy}(x,y)|J|$$

47. Let X and Y have the joint probability mass function

Y \ **X**	**0**	**1**	**2**
0	**0.1**	**0.4**	**0.1**
1	**0.2**	**0.2**	**0**

Find $P(X+Y>1)$ *and* $E(XY)$. **[AU N/D 2018]**

Solution:

$$P(X+Y>1)=P(1,1)+P(2,0)+P(2,1)$$
$$=0.2+0.1+0$$
$$=0.3$$

$$E(XY)=\sum_{x=0}^{2}\sum_{y=0}^{1}xyp(x,y)=(1)(1)(0.2)=0.2$$

48. The joint probability distribution function of the random variable (X,Y) is given by $f(x,y)=k(x^3y+xy^3),\ 0\le x\le 2,\ 0\le x\le 2$**, Find the value of** k. **[AU N/D 2018]**

Solution:

Given that $f(x,y)=k(x^3y+xy^3),\ 0\le x\le 2,\ 0\le x\le 2$

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

$$\int_0^2\int_0^2 k(x^3y+xy^3)dxdy=1$$

$$k\int_0^2\left(\frac{x^4}{4}y+\frac{x^2y^3}{2}\right)_0^2 dy=1$$

$$k\int_0^2(4y+2y^3)dy=1$$

$$k\left(2y^2+\frac{y^4}{2}\right)_0^2=1$$
$$k(8+8)=1$$
$$k=\frac{1}{16}$$

49. Given two random variables X and Y with the joint CDF $F_{XY}(x,y)$ and marginal CDFs $F_X(x)$ and $F_Y(y)$ respectively, compute the joint probability that X is greater then a and Y is greater than b. **[AU A/M '23]**

Solution:

We can obtain the desired probability as follows. From the de Morgan's second law, we know that $A \cap B = A \cup B$. Thus,

$$P[X>a,Y>b]=P[(X>a)\cap(Y>b)]$$
$$=1-\overline{P[(X>a)\cap(Y>b)]}$$
$$=1-\left\{\overline{P[(X>a)]}\cup\overline{P[(Y>b)]}\right\}$$
$$=1-\{P[(X\le a)]\cup P[(Y\le b)]\}$$
$$=1-\{P[(X\le a)]+P[(Y\le b)]-P[X\le a,Y\le b]\}$$
$$=1-F_X(a)-F_Y(b)+F_{XY}(a,b)$$

50. The joint PMF of two random variables X and Y is given by

$$P_{XY}(x,y)=\begin{cases}\frac{1}{18}(2x+y), & x=1,2;\ y=1,2\\ 0 & ,otherwise\end{cases}$$ **. What is the marginal PMF of X?**

Solution: **[AU A/M '23]**

Given, $P_{XY}(x,y)=\begin{cases}\frac{1}{18}(2x+y), & x=1,2;\ y=1,2\\ 0 & ,otherwise\end{cases}$

$$P_X(x)=\sum_y P_{XY}(x,y)$$
$$P_X(x)=\sum_x\sum_y\frac{1}{18}(2x+y)$$
$$P_X(x)=\sum_x\frac{1}{18}[(2x+1)+(2x+2)]$$
$$P_X(x)=\sum_x\frac{1}{18}[(4x+3)]\ ,\quad x=1,2$$

51. For a bi-variate random variable (XY), prove that if X and Y are independent, then every event $a < X \le b$ is independent of the other event $c < X \le d$ [AU N/D '22]

Solution:

$$P(a < X \le b, c < X \le d) = F_{XY}(b,d) - F_{XY}(a,d) - F_{XY}(b,c) + F_{XY}(a,c)$$

$$= [F_X(b) - F_X(a)][F_Y(d) - F_Y(c)]$$

$$= P(a < X \le b)P(c < X \le d)$$

52. Let the joint probability mass function of (XY), be given by

$P_{xy}(x,y) = \{k(x+y), x = 1,2,3 : y = 1,2.$ **Find the value of k? [AU N/D '22]**

Solution:

$$P_{xy}(x,y) = \{k(x+y), x = 1,2,3 : y = 1,2.$$

$$\sum_{x=1}^{3}\sum_{y=1}^{2} k(x+y) = 1 \qquad \Rightarrow k = \frac{1}{21}$$

53. A privately owned business operates both a drive –in facility and a walk-in facility. On a randomly selected day, let X and Y, respectively, be the proportions of the time that the drive-in and the walk-in facilities are in use, and suppose that the joint density function of these random variables is $f(x,y) = k(2x+3y);\ 0 \le x \le 1,\ 0 \le y \le 1$. Find the value of k.

Solution: [AU A/M 2024]

$$\text{Total probability } = 1$$

$$\Rightarrow \int_0^1\int_0^1 k(2x+3y)\,dxdy = 1 \Rightarrow k = \frac{2}{5}$$

54. An electrical firm manufactures light bulbs that have a length of life that is approximately normally distributed, with mean equal to 800 hours and a standard deviation of 40 hours. What is the probability that a random sample of 16 bulbs will have an average life of less than 775 hours?

Solution: [AU A/M 2024]

$$P(\overline{X} < 775) = P\left(Z < \frac{775-800}{\frac{40}{\sqrt{16}}}\right) = P(Z < -2.5) = 0.0062$$

55. Find the marginal distribution of X and Y from the following probability distribution function of (X,Y)

X	**Y**		
		1	**2**
	1	**0.1**	**0.2**
	2	**0.3**	**0.4**

Solution: **[AU N/D 2024]**

X	Y			
		1	2	P(X=x)
	1	0.1	0.2	0.3
	2	0.3	0.4	0.7
P(Y=y)		0.4	0.6	1

The marginal distribution of X:

$$P(X=1)=0.3 \;\; and \;\; P(X=2)=0.7$$

The marginal distribution of Y:

$$P(Y=1)=0.4 \;\; and \;\; P(Y=2)=0.2$$

56. If $U = X+Y$ and $V = X-Y$, how are the joint probability density functions of (X,Y) and (U,V) related?

Solution: **[AU N/D 2024]**

Given $U = X+Y$ and $V = X-Y$

$$X=\frac{U+V}{2} \;\&\; Y=\frac{U-V}{2}$$

$$f(U,V)=|J|f(X,Y)$$

Where $|J|=\begin{vmatrix} x_u & x_v \\ y_u & y_v \end{vmatrix}=\begin{vmatrix} \frac{1}{2} & \frac{1}{2} \\ \frac{1}{2} & -\frac{1}{2} \end{vmatrix}=-\frac{1}{4}-\frac{1}{4}=-\frac{1}{2} \Rightarrow |J|=\frac{1}{2}$

$$f(U,V)=\frac{1}{2}f(X,Y)$$

Part-B

1. The joint probability mass function of (X,Y) is given by p(x,y)=k(2x+3y) ,x=0,1,2; y=1,2,3. Find the all marginal and conditional distribution. Also find the probability distribution and find Probability distribution of (X+Y) and P[X+Y>3]. [AU N/D 2011, '14 A/M 2017]

Solution:

Given:

X \ Y	1	2	3	Total
0	3k	6k	9k	18k
1	5k	8k	11k	24k
2	7k	10k	13k	30k
Total	15k	24k	33k	72k

Total probability =1.

$$72k = 1 \Rightarrow \quad k = \frac{1}{72.}$$

X \ Y	1	2	3	$P(X=x)$
0	$\frac{3}{72.}$	$\frac{6}{72.}$	$\frac{9}{72.}$	$P(X=0)=\frac{18}{72}$
1	$\frac{5}{72.}$	$\frac{8}{72.}$	$\frac{11}{72.}$	$P(X=1)=\frac{24}{72}$
2	$\frac{7}{72.}$	$\frac{10}{72.}$	$\frac{13}{72.}$	$P(X=2)=\frac{30}{72}$
$P(Y=y)$	$P(Y=1)=\frac{15}{72}$	$P(Y=2)=\frac{24}{72}$	$P(Y=3)=\frac{33}{72}$	$\frac{72}{72.}=1$

To find the marginal distribution of X:

$$P(X=0)=\frac{18}{72}, \quad P(X=1)=\frac{24}{72}, \quad P(X=2)=\frac{30}{72}$$

To find the marginal distribution of Y:

$$P(Y=1)=\frac{15}{72}, \quad P(Y=2)=\frac{24}{72}, \quad P(Y=3)=\frac{33}{72}$$

To find the Conditional distribution of X is given Y is $P(X=x_i \mid Y=y_j)$

$(i)\ P(X=x_i \mid Y=1)$

$$P(X=0 \mid Y=1)=\frac{P(X=0,Y=1)}{p(Y=1)}=\frac{\frac{3}{72}}{\frac{15}{72}}=\frac{3}{15}=\frac{1}{5}$$

$$P(X=1 \mid Y=1)=\frac{P(X=1,Y=1)}{p(Y=1)}=\frac{\frac{5}{72}}{\frac{15}{72}}=\frac{5}{15}=\frac{1}{3}$$

$$P(X=2\,|\,Y=1)=\frac{P(X=2,Y=1)}{p(Y=1)}=\frac{\frac{7}{72}}{\frac{15}{72}}=\frac{7}{15}$$

$(ii)\ P(X=x_i\,|\,Y=2)$

$$P(X=0\,|\,Y=2)=\frac{P(X=0,Y=2)}{p(Y=2)}=\frac{\frac{6}{72}}{\frac{24}{72}}=\frac{6}{24}=\frac{1}{4}$$

$$P(X=1\,|\,Y=2)=\frac{P(X=1,Y=2)}{p(Y=2)}=\frac{\frac{8}{72}}{\frac{24}{72}}=\frac{8}{24}=\frac{1}{3}$$

$$P(X=2\,|\,Y=2)=\frac{P(X=2,Y=2)}{p(Y=2)}=\frac{\frac{10}{72}}{\frac{24}{72}}=\frac{10}{24}=\frac{5}{12}$$

$(iii)\ P(X=x_i\,|\,Y=3)$

$$P(X=0\,|\,Y=3)=\frac{P(X=0,Y=3)}{p(Y=3)}=\frac{\frac{9}{72}}{\frac{33}{72}}=\frac{9}{33}$$

$$P(X=1\,|\,Y=3)=\frac{P(X=1,Y=3)}{p(Y=3)}=\frac{\frac{11}{72}}{\frac{33}{72}}=\frac{11}{33}=\frac{1}{3}$$

$$P(X=2\,|\,Y=3)=\frac{P(X=2,Y=3)}{p(Y=3)}=\frac{\frac{13}{72}}{\frac{33}{72}}=\frac{13}{33}$$

To find the Conditional distribution of Y is given X is $P(y=y_j\,|\,X=x_i)$

$(i)\quad P(y=y_j\,|\,X=0)$, $P(Y=1\,|\,X=0)=\frac{P(X=0,Y=1)}{P(X=0)}=\frac{\frac{3}{72}}{\frac{18}{72}}=\frac{3}{18}=\frac{1}{6}$

$$P(Y=2\,|\,X=0)=\frac{P(X=0,Y=2)}{P(X=0)}=\frac{\frac{6}{72}}{\frac{18}{72}}=\frac{6}{18}=\frac{1}{6}$$

$$P(Y=3\mid X=0)=\frac{P(X=0,Y=3)}{p(X=0)}=\frac{\frac{9}{72}}{\frac{18}{72}}=\frac{9}{18}=\frac{1}{2}$$

$(ii)\ P(Y=y_j\mid X=1)$, $$P(Y=1\mid X=1)=\frac{P(X=1,Y=1)}{p(X=1)}=\frac{\frac{5}{72}}{\frac{24}{72}}=\frac{5}{24}$$

$$P(Y=2\mid X=1)=\frac{P(X=1,Y=2)}{p(X=1)}=\frac{\frac{8}{72}}{\frac{24}{72}}=\frac{8}{24}=\frac{1}{3}$$

$$P(Y=3\mid X=1)=\frac{P(X=1,Y=3)}{p(X=1)}=\frac{\frac{11}{72}}{\frac{24}{72}}=\frac{11}{24}$$

$(iii)\ P(Y=y_j\mid X=2)$, $$P(Y=1\mid X=2)=\frac{P(X=2,Y=1)}{p(X=2)}=\frac{\frac{7}{72}}{\frac{30}{72}}=\frac{7}{30}$$

$$P(Y=2\mid X=2)=\frac{P(X=2,Y=2)}{p(X=2)}=\frac{\frac{10}{72}}{\frac{30}{72}}=\frac{1}{30}$$

$$P(Y=3\mid X=2)=\frac{P(X=2,Y=3)}{p(X=2)}=\frac{\frac{13}{72}}{\frac{30}{72}}=\frac{13}{30}$$

Probability distribution of (X+Y)	
(X+Y)	P
1, P(0,1)	$\frac{3}{72}$
2, P(0,2)+P(1,1)	$\frac{6}{72}+\frac{5}{72}=\frac{11}{72}$
3, P(0,3)+ P(1,2)+P(2,1)	$\frac{9}{72}+\frac{8}{72}+\frac{7}{72}=\frac{24}{72}$
4, P(1,3)+P(2,2)	$\frac{11}{72}+\frac{10}{72}=\frac{21}{72}$
5, P(2,3)	$\frac{13}{72}$
Total	1

$$P[X+Y>3] = P[X+Y=4]+P[X+Y=5]$$
$$= \frac{21}{72}+\frac{13}{72}=\frac{34}{72}.$$

2. Let X and Y be two random variables having the joint probability function f(x,y) =k(x+2y), where x and y can assume only the integer values 0, 1 and 2.Find the marginal and conditional distribution.

Solution: **[AU M/J 2012]**

X \ Y	0	1	2	Total
0	0	2k	4k	6k
1	k	3k	5k	9k
2	2k	4k	6k	12k
Total	3k	9k	15k	27k

Total probability =1.

$$27k = 1 \quad \Rightarrow k = \frac{1}{27}$$

X \ Y	0	1	2	$P(X = x)$
0	0	$\frac{2}{27}$	$\frac{4}{27}$	$P(X=0)=\frac{6}{27}$
1	$\frac{1}{27}$	$\frac{3}{27}$	$\frac{5}{27}$	$P(X=1)=\frac{9}{27}$
2	$\frac{2}{27}$	$\frac{4}{27}$	$\frac{6}{27}$	$P(X=2)=\frac{12}{27}$
$P(Y=y)$	$P(Y=0)=\frac{3}{27}$	$P(Y=1)=\frac{9}{27}$	$P(Y=2)=\frac{15}{27}$	$\frac{27}{27}=1$

To find the marginal distribution of X:

$$P(X=0)=\frac{6}{27}, \quad P(X=1)=\frac{9}{27}, \quad P(X=2)=\frac{12}{27}$$

To find the marginal distribution of Y:

$$P(Y=1)=\frac{3}{27}, \quad P(Y=2)=\frac{9}{27}, \quad P(Y=3)=\frac{15}{27}$$

To find the Conditional distribution of X is given Y is $P(X = x_i \mid Y = y_j)$

$(i)\ P(X = x_i \mid Y = 0)$

$$P(X=0 \mid Y=0) = \frac{P(X=0, Y=0)}{p(Y=0)} = \frac{0}{\frac{3}{27}} = 0$$

$$P(X=1\,|\,Y=0)=\frac{P(X=1,Y=0)}{p(Y=0)}=\frac{\frac{1}{27}}{\frac{3}{27}}=\frac{1}{3}$$

$$P(X=2\,|\,Y=0)=\frac{P(X=2,Y=0)}{p(Y=0)}=\frac{\frac{2}{27}}{\frac{3}{27}}=\frac{2}{3}$$

(ii) $P(X=x_i\,|\,Y=1)$

$$P(X=0\,|\,Y=1)=\frac{P(X=0,Y=1)}{p(Y=1)}=\frac{\frac{2}{27}}{\frac{9}{27}}=\frac{2}{9}$$

$$P(X=1\,|\,Y=1)=\frac{P(X=1,Y=1)}{p(Y=1)}=\frac{\frac{3}{27}}{\frac{9}{27}}=\frac{3}{9}=\frac{1}{3}$$

$$P(X=2\,|\,Y=1)=\frac{P(X=2,Y=1)}{p(Y=1)}=\frac{\frac{4}{27}}{\frac{9}{27}}=\frac{4}{9}$$

(iii) $P(X=x_i\,|\,Y=2)$

$$P(X=0\,|\,Y=2)=\frac{P(X=0,Y=2)}{p(Y=2)}=\frac{\frac{4}{27}}{\frac{15}{27}}=\frac{4}{15}$$

$$P(X=1\,|\,Y=2)=\frac{P(X=1,Y=2)}{p(Y=2)}=\frac{\frac{5}{27}}{\frac{15}{27}}=\frac{5}{15}=\frac{1}{3}$$

$$P(X=2\,|\,Y=2)=\frac{P(X=2,Y=2)}{p(Y=2)}=\frac{\frac{6}{27}}{\frac{15}{27}}=\frac{6}{15}$$

To find the Conditional distribution of Y is given X is $P(y=y_j\,|\,X=x_i)$

(i) $P(y=y_j\,|\,X=0)$

$$P(Y=0\,|\,X=0)=\frac{P(X=0,Y=0)}{P(X=0)}=\frac{0}{\frac{6}{27}}=0$$

$$P(Y=1 \mid X=0)=\frac{P(X=0,Y=1)}{P(X=0)}=\frac{\frac{2}{27}}{\frac{6}{27}}=\frac{2}{6}=\frac{1}{3}$$

$$P(Y=2 \mid X=0)=\frac{P(X=0,Y=2)}{p(X=0)}=\frac{\frac{4}{27}}{\frac{6}{27}}=\frac{4}{6}=\frac{2}{3}$$

$(ii)\ P(Y=y_j \mid X=1)$

$$P(Y=0 \mid X=1)=\frac{P(X=1,Y=0)}{p(X=1)}=\frac{\frac{1}{27}}{\frac{9}{27}}=\frac{1}{9}$$

$$P(Y=1 \mid X=1)=\frac{P(X=1,Y=1)}{p(X=1)}=\frac{\frac{3}{27}}{\frac{9}{27}}=\frac{3}{9}=\frac{1}{3}$$

$$P(Y=2 \mid X=1)=\frac{P(X=1,Y=2)}{p(X=1)}=\frac{\frac{5}{27}}{\frac{9}{27}}=\frac{5}{9}$$

$(iii)\ P(Y=y_j \mid X=2)$

$$P(Y=0 \mid X=2)=\frac{P(X=2,Y=0)}{p(X=2)}=\frac{\frac{2}{27}}{\frac{12}{27}}=\frac{2}{12}$$

$$P(Y=1 \mid X=2)=\frac{P(X=2,Y=1)}{p(X=2)}=\frac{\frac{4}{27}}{\frac{12}{27}}=\frac{4}{12}$$

$$P(Y=2 \mid X=2)=\frac{P(X=2,Y=2)}{p(X=2)}=\frac{\frac{6}{27}}{\frac{12}{27}}=\frac{6}{12}$$

3. The joint distribution of X and Y is given by f(x,y) = x+y/21, x = 1,2,3, y =1,2. Find the marginal distribution. **[AU N/D ’03, A/M ‘05]**

Solution:

Given: $f(x,y)=\frac{x+y}{21},\ x=1,2,3\ and\ \ y=1,2$

Y \ X	1	2	3	$P(Y=y)$
1	$\frac{2}{21}$	$\frac{3}{21}$	$\frac{4}{21}$	$P(Y=1)=\frac{9}{21}$
2	$\frac{3}{21}$	$\frac{4}{21}$	$\frac{5}{21}$	$P(Y=2)=\frac{12}{21}$
$P(X=x)$	$P(X=1)=\frac{5}{21}$	$P(X=2)=\frac{7}{21}$	$P(X=3)=\frac{9}{21}$	$\frac{21}{21}=1$

Result: To find the marginal distribution of X:

$$P(X=1)=\frac{5}{21},\quad P(X=1)=\frac{7}{21},\quad P(X=2)=\frac{9}{21}$$

To find the marginal distribution of Y:

$$P(Y=1)=\frac{9}{21},\quad P(Y=2)=\frac{12}{21},$$

4. The two dimensional random variable (X,Y) has the joint density function f(x,y)=x+2y/27, x =0,1,2; y =0,1,2.Find the conditional distribution of Y given X = x. Also find the conditional distribution of X given Y=1. [AU M/J 07, A/M '08]

Solution:

Given: f(x,y)=x+2y/27,

Y \ X	0	1	2	$P(Y=y)$
0	0	$\frac{1}{27}$	$\frac{2}{27}$	$P(Y=0)=\frac{3}{27}$
1	$\frac{2}{27}$	$\frac{3}{27}$	$\frac{4}{27}$	$P(Y=1)=\frac{9}{27}$
2	$\frac{4}{27}$	$\frac{5}{27}$	$\frac{6}{27}$	$P(Y=2)=\frac{15}{27}$
$P(X=x)$	$P(X=0)=\frac{6}{27}$	$P(X=1)=\frac{9}{27}$	$P(X=2)=\frac{12}{27}$	$\frac{27}{27}=1$

To find the marginal distribution of X:

$$P(X=0)=\frac{6}{27},\quad P(X=1)=\frac{9}{27},\quad P(X=2)=\frac{12}{27}$$

To find the marginal distribution of Y:

$$P(Y=0)=\frac{3}{27},\quad P(Y=1)=\frac{9}{27},\quad P(Y=2)=\frac{15}{27}$$

To find the Conditional distribution of X is given Y =1 is

$$P(X=0\mid Y=1)=\frac{P(X=0,Y=1)}{p(Y=1)}=\frac{\frac{2}{27}}{\frac{9}{27}}=\frac{2}{9}$$

$$P(X=1\mid Y=1)=\frac{P(X=1,Y=1)}{p(Y=1)}=\frac{\frac{3}{27}}{\frac{9}{27}}=\frac{3}{9}$$

$$P(X=2\mid Y=1)=\frac{P(X=2,Y=1)}{p(Y=1)}=\frac{\frac{4}{27}}{\frac{9}{27}}=\frac{4}{9}$$

To find the Conditional distribution of Y is given X is $P(y=y_j\mid X=x_i)$

$(i)\quad P(y=y_j\mid X=0)$

$$P(Y=0\mid X=0)=\frac{P(X=0,Y=0)}{P(X=0)}=\frac{0}{\frac{6}{27}}=0$$

$$P(Y=1\mid X=0)=\frac{P(X=0,Y=1)}{P(X=0)}=\frac{\frac{2}{27}}{\frac{6}{27}}=\frac{2}{6}=\frac{1}{3}$$

$$P(Y=2\mid X=0)=\frac{P(X=0,Y=2)}{p(X=0)}=\frac{\frac{4}{27}}{\frac{6}{27}}=\frac{4}{6}=\frac{2}{3}$$

$(ii)\ P(Y=y_j\mid X=1)$

$$P(Y=0\mid X=1)=\frac{P(X=1,Y=0)}{p(X=1)}=\frac{\frac{1}{27}}{\frac{9}{27}}=\frac{1}{9}$$

$$P(Y=1\mid X=1)=\frac{P(X=1,Y=1)}{p(X=1)}=\frac{\frac{3}{27}}{\frac{9}{27}}=\frac{3}{9}=\frac{1}{3}$$

$$P(Y=2\mid X=1)=\frac{P(X=1,Y=2)}{p(X=1)}=\frac{\frac{5}{27}}{\frac{9}{27}}=\frac{5}{9}$$

$(iii)\ P(Y=y_j\mid X=2)$

$$P(Y=0\mid X=2)=\frac{P(X=2,Y=0)}{p(X=2)}=\frac{\frac{2}{27}}{\frac{12}{27}}=\frac{2}{12}$$

$$P(Y=1\mid X=2)=\frac{P(X=2,Y=1)}{p(X=2)}=\frac{\frac{4}{27}}{\frac{12}{27}}=\frac{4}{12}$$

Result:

$$P(Y=2\mid X=2)=\frac{P(X=2,Y=2)}{p(X=2)}=\frac{\frac{6}{27}}{\frac{12}{27}}=\frac{6}{12}$$

5. Calculate the correlation coefficient for the following heights (in inches) of fathers X and their sons Y.

X	**65**	**66**	**67**	**67**	**68**	**69**	**70**	**72**
Y	**67**	**68**	**65**	**68**	**72**	**72**	**69**	**71**

Solution: **[AU Nov 2007, 09]**

X	Y	XY	X^2	Y^2
65	67	4355	4225	4489
66	68	4488	4356	4624
67	65	4355	4489	4225
67	68	4556	4489	4624
68	72	4896	4624	5184
69	72	4968	4761	5184
70	69	4830	4900	4761
72	71	5112	5184	5041
$\sum X=544$	$\sum Y=552$	$\sum XY=37560$	$\sum X^2=37028$	$\sum Y^2=38132$

Now $\overline{X}=\frac{544}{8}=68 \quad \overline{Y}=\frac{552}{8}=69$

$\overline{Y}\,\overline{X}=68\times 69=4692$

Formula:

$$\sigma_X=\sqrt{\frac{1}{n}\sum X^2-\overline{X}^2}$$

$$=\sqrt{\frac{37028}{8}-4624}$$

$$=2.121$$

$$\sigma_Y=\sqrt{\frac{1}{n}\sum Y^2-\overline{Y}^2}$$

$$= \sqrt{\frac{38132}{8} - 4761}$$
$$= 2.345$$

Result: $\therefore \quad r(X,Y) = \dfrac{Cov(X,y)}{\sigma_X . \sigma_Y}$

$$= \frac{\frac{1}{n}\sum XY - \overline{X}\,\overline{Y}}{\sigma_X . \sigma_Y} = \frac{\frac{1}{8} \times 37560 - 4692}{2.121 \times 2.345}$$
$$= 0.6030$$

6. Calculate the correlation coefficient for the following data.

X:	**9**	**8**	**7**	**6**	**5**	**4**	**3**	**2**	**1**
Y:	**15**	**16**	**14**	**13**	**11**	**12**	**10**	**8**	**9**

Solution: **[AU Nov 2007, '09 '14]**

X	Y	XY	X^2	Y^2
9	15	135	81	225
8	16	128	64	256
7	14	98	49	196
6	13	78	36	169
5	11	55	25	121
4	12	48	16	144
3	10	30	9	100
2	8	16	4	64
1	9	9	1	81
$\sum X = 45$	$\sum Y = 108$	$\sum XY = 597$	$\sum X^2 = 285$	$\sum Y^2 = 1356$

Now $\overline{X} = \dfrac{45}{9} = 5 \quad \overline{Y} = \dfrac{108}{9} = 12$

$\overline{Y}\,\overline{X} = 5 \times 12 = 60$

Formula:

$$\sigma_X = \sqrt{\frac{1}{n}\sum X^2 - \overline{X}^2}$$
$$= \sqrt{\frac{285}{9} - 25}$$
$$= 2.582$$

$$\sigma_Y = \sqrt{\frac{1}{n}\sum Y^2 - \overline{Y}^2}$$

$$= \sqrt{\frac{1356}{9} - 12}$$

$$= 11.7757$$

Result: $\therefore \quad r(X,Y) = \dfrac{Cov(X,y)}{\sigma_X . \sigma_Y}$

$$= \frac{\frac{1}{n}\sum XY - \overline{X}\overline{Y}}{\sigma_X . \sigma_Y} = \frac{\frac{1}{9}\times 597 - 60}{2.582 \times 11.7757} = \frac{6.3333}{30.4049}$$

$$= 0.2083$$

7. Calculate the correlation coefficient between industrial production and export using the following data.

Production(X)	**55**	**56**	**58**	**59**	**60**	**60**	**62**
Export(Y)	**35**	**38**	**37**	**39**	**44**	**43**	**44**

Solution: **[AU Nov 2007, 09]**

X	Y	U=X-58	V=Y-40	UV	U^2	V^2
55	35	-3	-5	15	9	25
56	38	-2	-2	4	4	4
58	37	0	-3	0	0	9
59	39	1	-1	-1	1	1
60	44	2	4	8	4	16
60	43	2	3	6	4	9
62	44	4	4	16	16	16
		$\sum U = 4$	$\sum V = 0$	$\sum UV = 48$	$\sum U^2 = 38$	$\sum V^2 = 80$

Now $\overline{U} = \dfrac{\sum U}{n} = \dfrac{4}{7} = 0.5714 \quad \overline{V} = \dfrac{\sum V}{n} = 0$

Formula: $\sigma_U = \sqrt{\frac{1}{n}\sum U^2 - \overline{U}^2}$

$$= \sqrt{\frac{38}{7} - (0.5714)^2}$$
$$= 2.2588$$
$$\sigma_V = \sqrt{\frac{1}{n}\sum V^2 - \overline{V}^2}$$
$$= \sqrt{\frac{80}{7} - 0}$$
$$= 3.38$$

Result: $\therefore \quad r(U,V) = \dfrac{Cov(U,V)}{\sigma_U . \sigma_V}$

$$= \frac{\frac{1}{n}\sum UV - \overline{U}\overline{V}}{\sigma_U . \sigma_V} = \frac{\frac{1}{7}\times 48 - 0}{2.258 \times 3.38}$$
$$= 0.898$$

8. Calculate the correlation coefficient from the following data.

(X)	**10**	**14**	**18**	**22**	**26**	**30**
(Y)	**18**	**12**	**24**	**6**	**30**	**36**

Solution: **[AU Nov. 2008]**

X	Y	$U = \frac{X-22}{4}$	$V = \frac{Y-24}{6}$	UV	U^2	V^2
10	18	-3	-1	3	9	1
14	12	-2	-2	4	4	4
18	24	-1	0	0	1	0
22	6	0	-3	0	0	9
26	30	1	1	1	1	1
30	36	2	2	4	4	4
		$\sum U = -3$	$\sum V = -3$	$\sum UV = 12$	$\sum U^2 = 19$	$\sum V^2 = 19$

Now $\overline{U} = \dfrac{\sum U}{n} = \dfrac{-3}{6} = -0.5 \quad \overline{V} = \dfrac{\sum V}{n} = \dfrac{-3}{6} = -0.5$

Formula:

$$\sigma_U = \sqrt{\frac{1}{n}\sum U^2 - \overline{U}^2}$$
$$= \sqrt{\frac{19}{6} - (-0.5)^2}$$
$$= 1.708$$
$$\sigma_V = \sqrt{\frac{1}{n}\sum V^2 - \overline{V}^2}$$
$$= \sqrt{\frac{19}{6} - (-0.5)^2}$$
$$= 1.708$$

Result: $\therefore \quad r(U,V) = \dfrac{Cov(U,V)}{\sigma_U . \sigma_V}$

$$= \frac{\frac{1}{n}\sum UV - \overline{U}\overline{V}}{\sigma_U . \sigma_V} = \frac{\frac{1}{6} \times 12 - (-0.5)(-0.5)}{1.708 \times 1.708}$$
$$= 0.6$$

9. If the independent random variables X and Y have the variances 36 and 16 respectively, ffind the Correlation coefficient between $U = X + Y$ and $V = X - Y$ [AU May '09, '14]

Solution:

Given: that $Var(X) = 36, Var(Y) = 16$ Since X and Y are independent,

$$E(XY) = E(X).E(Y)$$

$$Let \;\; U = X + Y \;\; and \;\; V = X - Y$$
$$Var(U) = Var(X + Y)$$
$$= 1^2 Var(X) + 1^2 Var(Y)$$
$$= 36 + 16 = 52$$
$$\sigma_U = \sqrt{52} \ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots (A)$$
$$Var(V) = Var(X - Y)$$
$$= 1^2 Var(X) + (-1)^2 Var(Y)$$
$$= 36 + 16 = 52$$
$$\sigma_V = \sqrt{52} \ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots (B)$$
$$Cov(U,V) = E(UV) - E(U).E(V) \ldots\ldots\ldots\ldots\ldots\ldots (1)$$
$$E(UV) = E[(X + Y)(X - Y)]$$
$$= E[X^2 - Y^2]$$
$$= E[X^2] - E[Y^2] \ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots (2)$$

$$E(U) = E[X+Y] = E(X)+E(Y)......................(3)$$
$$E(V) = E[X-Y] = E(X)-E(Y)......................(4)$$

Substituting (2),(3),(4) in (1), we get,

Formula: $Cov(U,V) = E(X^2)-E(Y^2)-[E(X)+E(Y)][E(X)-E(Y)]$

$$= E(X^2)-E(Y^2)-[E(X)]^2+[E(Y)]^2-[E(X)][E(Y)]+[E(X)][E(Y)]$$
$$= \{E(X^2)-[E(X)]^2\}-\{E(Y^2)-[E(Y)]^2\}$$
$$= Var(X)-Var(Y)$$

$$Cov(U,V) = 36-16 = 20..(C)$$

Result: $\rho(U,V) = \dfrac{Cov(U,V)}{\sigma_U \sigma_V}$

$$= \frac{20}{\sqrt{52}\sqrt{52}} = \frac{20}{52} = \frac{5}{13} \qquad [U\sin g\,(A),(B)\,\&\,(C)]$$

10. From the following data, find

(i) The two regression equations

(ii) The coefficient of correlation between the marks in Economics and Statistics.

(iii) The most likely marks in statistics when marks in Economics are 30.

Maks in Economics	**25**	**28**	**35**	**32**	**31**	**36**	**29**	**38**	**34**	**32**
Marks in Statistics	**43**	**46**	**49**	**41**	**36**	**32**	**31**	**30**	**33**	**39**

Solution: **[AU RP May '07, '09]**

X	Y	$X-\overline{X} = X-32$	$Y-\overline{Y} = Y-38$	$(X-\overline{X})^2$	$(Y-\overline{Y})^2$	$(X-\overline{X})(Y-\overline{Y})$
25	43	-7	5	49	25	-35
28	46	-4	8	16	64	-32
35	49	3	11	9	121	33
32	41	0	3	0	9	0
31	36	-1	-2	1	4	2
36	32	4	-6	16	36	-24
29	31	-3	-7	9	49	21
38	30	6	-8	36	64	-48
34	33	2	-5	4	25	-10
32	39	0	1	0	1	0
320	380	0	0	140	398	-93

$$\overline{X} = \frac{\sum X}{n} = \frac{320}{10} = 32 \quad \text{and} \quad \overline{Y} = \frac{\sum Y}{n} = \frac{380}{10} = 38$$

Formula: Coefficient of regression of Y on X is

$$b_{YX} = \frac{\sum \left(X - \overline{X}\right)\left(Y - \overline{Y}\right)}{\sum \left(X - \overline{X}\right)^2}$$

$$= \frac{-93}{140} = -0.6643$$

Coefficient of regression of X on Y is

$$b_{XY} = \frac{\sum \left(X - \overline{X}\right)\left(Y - \overline{Y}\right)}{\sum \left(Y - \overline{Y}\right)^2}$$

$$= \frac{-93}{398} = -0.2337$$

Equation of the line of regression of X on Y is

$$x - \overline{x} = b_{XY}\left(y - \overline{y}\right)$$

$$x - 32 = -0.2337(y - 38)$$

i.e., $x = -0.2337y + 40.8806$

Equation of the line of regression of X on Y is

$$y - \overline{y} = b_{YX}\left(x - \overline{x}\right)$$

$$y - 38 = -0.6643(x - 32)$$

i.e., $y = -0.6642x + 59.2576$

Coefficient of correlation

$$r^2 = b_{XY} \times b_{YX}$$

$$= -0.6643 \times (-0.2337) = 0.1552$$

$$r = \pm\sqrt{0.1552}$$

$$= \pm 0.394$$

Now we have to find the most likely marks in Statistics (Y) when marks in Economics (X) are 30. We use the line of regression of Y on X

Result: *i.e.,* $y = -0.6642x + 59.2576$

Put $x = 30$, *we get*

$$y = -0.6643 \times 30 + 59.2576 = 39.3286$$

$$\cong 39$$

11. The joint p.d.f of the RV (X,Y) is given by $f(x, y) = Kxye^{-\left(x^2+y^2\right)}$ $x > 0, y > 0$. Find the value of K and prove also that X and Y are independent. **[AUM/J 2000,04.09]**

Solution:

Given: function $f(x, y) = Kxye^{-\left(x^2+y^2\right)}$ $x > 0, y > 0$ is a p.d.f

$$\therefore \quad \int_0^\infty\int_0^\infty Kxye^{-\left(x^2+y^2\right)}dxdy = 1$$

$$K\int_0^\infty\int_0^\infty xye^{-x^2}e^{-y^2}dxdy = 1$$

$$K\left[\int_0^\infty xe^{-x^2}dx\right]\left[\int_0^\infty ye^{-y^2}dx\right] = 1$$

$$Put \quad x^2 = t,\ 2xdx = dt, \qquad i.e., \quad xdx = \frac{dt}{2}$$

$$as \quad x \to 0 \Rightarrow t \to 0, \quad x \to \infty \Rightarrow t \to \infty$$

$$Put \quad y^2 = v,\ 2ydy = dv, \qquad i.e., \quad ydy = \frac{dv}{2}$$

$$as \quad y \to 0 \Rightarrow v \to 0, \quad y \to \infty \Rightarrow v \to \infty$$

$$K\left[\int_0^\infty e^{-t}\frac{dt}{2}\right]\left[\int_0^\infty e^{-v}\frac{dv}{2}\right] = 1$$

$$K\left[\frac{e^{-t}}{-2}\right]_0^\infty\left[\frac{e^{-v}}{-2}\right]_0^\infty = 1$$

$$K\left[0-\left(\frac{1}{-2}\right)\right]\left[0-\left(\frac{1}{-2}\right)\right] = 1$$

$$K\left[\frac{1}{2}\right]\left[\frac{1}{2}\right] = 1$$

$$K = 4$$

The marginal density of X is given by

$$f_X(x) = \int_0^\infty f(x,y)dy = \int_0^\infty 4xye^{-\left(x^2+y^2\right)}dy$$

$$= \int_0^\infty 4xye^{-x^2}e^{-y^2}dy$$

$$= 4xe^{-x^2}\int_0^\infty ye^{-y^2}dy$$

$$= 4xe^{-x^2}\left[\frac{1}{2}\right]$$

$$= 2xe^{-x^2},\ x > 0$$

The marginal density of Y is given by

$$f_Y(y) = \int_0^\infty f(x,y)dx = \int_0^\infty 4xye^{-\left(x^2+y^2\right)}dx$$

$$= \int_0^\infty 4xye^{-x^2}e^{-y^2}dx$$

$$= 4ye^{-y^2}\int_0^\infty xe^{-x^2}dx$$

$$= 4ye^{-y^2}\left[\frac{1}{2}\right]$$

$$= 2ye^{-y^2}, \quad y > 0$$

If $f_X(x)f_Y(y) = f(x,y)$ then X and Y are independent.

$$f_X(x)f_Y(y) = \left(2xe^{-x^2}\right)\left(2ye^{-y^2}\right) = 4xye^{-\left(x^2+y^2\right)} = f(x,y)$$

Result: Therefore X and Y are independent.

12. If the joint distribution function of X and Y is given by $F(x,y) = \left(1-e^{-x}\right)\left(1-e^{-y}\right), \quad x > 0, y > 0$

(i) Find the marginal densities of X and Y

(ii) Are X and Y independent?

(iii) Find $P(1 < X < 3, 1 < Y < 2)$

Solution: **[AU N/D '06, M/J '05]**

Given: $F(x,y) = \left(1-e^{-x}\right)\left(1-e^{-y}\right) = 1 - e^{-x} - e^{-y} + e^{-(x+y)}$

The joint p.d.f is given by $f(x,y) = \dfrac{\partial^2 F(x,y)}{\partial x \partial y}$

$$f(x,y) = \frac{\partial^2 F(x,y)}{\partial x \partial y}\left[1 - e^{-x} - e^{-x} + e^{-(x+y)}\right]$$

$$f(x,y) = \frac{\partial}{\partial x}\left[e^{-y} - e^{-(x+y)}\right] = e^{-(x+y)}$$

(i) The marginal density function of X is $f(x) = \int_{-\infty}^{\infty} f(x,y)dy$

$$f(x) = \int_{-\infty}^{\infty} e^{-(x+y)}dy = e^{-x}\left(e^{-y}\right)_0^\infty = e^{-x}, \quad x \ge 0$$

Similarly, The marginal density function of Y is $f(y) = e^{-y}$, $y \ge 0$

(ii) $f(x).f(y) = e^{-x}.e^{-y} = e^{-(x+y)} = f(x,y)$

Therefore X and Y are independent.

(iii) $P(1 < X < 3, 1 < Y < 2) = P(1 < X < 3)P(1 < Y < 2)$ [since X and Y are independent]

$$= \int_1^3 f(x)dx \int_1^2 f(y)dy$$

$$= \int_1^3 e^{-x}dx \int_1^2 e^{-y}dy = \left[-e^{-x}\right]_1^3 \left[-e^{-x}\right]_1^2 = \left(e^{-3} - e^{-1}\right)\left(e^{-2} - e^{-1}\right)$$

$$= \left(\frac{1}{e^3} - \frac{1}{e}\right)\left(\frac{1}{e^2} - \frac{1}{e}\right)$$

$$= \left(\frac{1-e^2}{e^3}\right)\left(\frac{1-e}{e^2}\right)$$

13. The joint p.d.f of X and Y is given by $f(x,y) = xy^2 + \frac{x^2}{8}, 0 \le x \le 2, 0 \le y \le 1$

Compute (i) $P\left(X > 1 / Y < \frac{1}{2}\right)$ **[AU Dec '09, Apr '08]**

(ii) $P\left(Y < \frac{1}{2} / X > 1\right)$

(iii) $P(X < Y)$ (iv) $P(X + Y \le 1)$

Solution:

Given: $f(x,y) = xy^2 + \frac{x^2}{8}, 0 \le x \le 2, 0 \le y \le 1$

Now $P\left(X > 1 / Y < \frac{1}{2}\right) = \dfrac{P\left(X > 1, Y < \frac{1}{2}\right)}{P\left(Y < \frac{1}{2}\right)}$

$$P\left(X > 1, Y < \frac{1}{2}\right) = \int_0^{\frac{1}{2}} \int_1^2 \left(xy^2 + \frac{x^2}{8}\right) dxdy$$

$$= \int_0^{\frac{1}{2}} \left(\frac{xy^2}{2} + \frac{x^3}{24}\right)_1^2 dy$$

$$= \int_0^{\frac{1}{2}} \left[\left(2y^2 + \frac{1}{3}\right) - \left(\frac{y^2}{2} + \frac{1}{24}\right)\right] dy$$

$$= \int_0^{\frac{1}{2}} \left(\frac{3y^2}{2} + \frac{7}{24}\right) dy$$

$$= \left(\frac{3}{2}\frac{y^3}{3} + \frac{7y}{24}\right)_0^{\frac{1}{2}}$$

$$= \left[\frac{3}{6}\left(\frac{1}{8}\right) + \frac{7}{24}\left(\frac{1}{2}\right)\right]$$

$$= \frac{1}{16} + \frac{7}{48} = \frac{5}{24}$$

$$P\left(Y < \frac{1}{2}\right) = \int_0^{\frac{1}{2}}\int_0^2 \left(xy^2 + \frac{x^2}{8}\right)dxdy$$

$$= \int_0^{\frac{1}{2}} \left(\frac{xy^2}{2} + \frac{x^3}{24}\right)_0^2 dy$$

$$= \int_0^{\frac{1}{2}} \left(\frac{4y^2}{2} + \frac{8}{24}\right)dy$$

$$= \int_0^{\frac{1}{2}} \left(2y^2 + \frac{1}{3}\right)dy$$

$$= \left(\frac{2y^3}{3} + \frac{y}{3}\right)_0^{\frac{1}{2}}$$

$$= \left[\frac{2}{3}\left(\frac{1}{8}\right) + \frac{1}{3}\left(\frac{1}{2}\right)\right]$$

$$= \frac{2}{24} + \frac{1}{6} = \frac{6}{24} = \frac{1}{4}$$

$$P(X > 1) = \int_0^1\int_1^2 \left(xy^2 + \frac{x^2}{8}\right)dxdy$$

$$= \int_0^1 \left(\frac{xy^2}{2} + \frac{x^3}{24}\right)_0^2 dy$$

$$= \int_0^1 \left[\left(\frac{4y^2}{2} + \frac{8}{24}\right) - \left(\frac{y^2}{2} + \frac{1}{24}\right)\right]dy$$

$$= \int_0^1 \left(2y^2 + \frac{1}{3} - \frac{1}{2}y^2 - \frac{1}{24}\right)dy$$

$$= \int_0^1 \left(\frac{3y^2}{2} + \frac{7}{24}\right)dy$$

$$= \left(\frac{3}{2}\frac{y^2}{2} + \frac{7y}{24} \right)_0^1$$

$$= \frac{1}{2} + \frac{7}{24} = \frac{19}{24}$$

(i) $$P\left(X > 1 / Y < \frac{1}{2} \right) = \frac{P\left(X > 1, Y < \frac{1}{2} \right)}{P\left(Y < \frac{1}{2} \right)}$$

$$= \frac{5}{24} \times 4 = \frac{5}{6}$$

(ii) $$P\left(Y < \frac{1}{2} / X > 1 \right) = \frac{P\left(Y < \frac{1}{2}, X > 1 \right)}{P(X > 1)}$$

$$= \frac{5}{24} \times \frac{24}{19} = \frac{5}{19}$$

(iii) $$P(X < Y) = \int_0^1 \int_0^y \left(xy^2 + \frac{x^2}{8} \right) dxdy$$

$$= \int_0^1 \left(\frac{xy^2}{2} + \frac{x^3}{24} \right)_0^y dy$$

$$= \int_0^1 \left(\frac{y^4}{2} + \frac{y^3}{24} \right) dy$$

$$= \left(\frac{y^5}{10} + \frac{y^4}{96} \right)_0^1$$

$$= \frac{1}{10} + \frac{1}{96} = \frac{106}{960}$$

$$= \frac{53}{480}$$

(iv) $$P(X + Y \le 1) = \int_0^1 \int_0^{1-y} \left(xy^2 + \frac{x^2}{8} \right) dxdy$$

$$= \int_0^1 \left(\frac{xy^2}{2} + \frac{x^3}{24} \right)_0^{1-y} dy$$

$$= \int_0^1 \left(\frac{(1-y)^2 y^2}{2} + \frac{(1-y)^3}{24} \right) dy$$

$$= \int_0^1 \left(\frac{(1 - y^2 - 2y) y^2}{2} + \frac{(1-y)^3}{24} \right) dy$$

$$= \int_0^1 \left(\frac{\left(y^2 + y^4 - 2y^3\right)}{2} + \frac{(1-y)^3}{24} \right) dy$$

$$= \frac{1}{2}\left(\frac{y^3}{3} + \frac{y^5}{5} - \frac{2y^4}{4} \right)_0^1 + \frac{1}{24}\left(\frac{(1-y)^4}{-4} \right)_0^1$$

$$= \frac{1}{2}\left(\frac{1}{3} + \frac{1}{5} - \frac{1}{2} \right) + \frac{1}{24}\left(\frac{1}{4} \right)$$

$$= \frac{1}{2}\left(\frac{10+6-15}{30} \right) + \frac{1}{96}$$

$$= \frac{1}{60} + \frac{1}{96} = \frac{156}{5760}$$

$$= \frac{13}{480}$$

14. The joint p.d.f of X and Y is given by $f(x,y) = \begin{cases} \frac{6}{7}\left(x^2 + \frac{xy}{2} \right), 0 \le x \le 1, 0 \le y \le 2 \\ 0, \quad otherwse \end{cases}$

Find the conditional density function of X given Y and the conditional density function of Y given X. **[AU May 2014]**

Solution:

Given: $f(x,y) = \begin{cases} \frac{6}{7}\left(x^2 + \frac{xy}{2} \right), 0 \le x \le 1, 0 \le y \le 2 \\ 0, \quad otherwse \end{cases}$

Formula: The marginal density function of X is

$$f_X(x) = \int_0^2 f(x,y)\,dy$$

$$= \int_0^2 \frac{6}{7}\left(x^2 + \frac{xy}{2} \right) dy$$

$$= \frac{6}{7}\left(yx^2 + \frac{xy^2}{4} \right)_0^2$$

$$= \frac{6}{7}\left(2x^2 + x\right)$$

The marginal density function of Y is

$$f_Y(y) = \int_0^1 f(x,y)\,dx$$

$$= \int_0^1 \frac{6}{7}\left(x^2 + \frac{xy}{2} \right) dx$$

$$= \frac{6}{7}\left(\frac{x^3}{3} + \frac{x^2 y}{4}\right)_0^1$$

$$= \frac{6}{7}\left(\frac{1}{3} + \frac{3y}{4}\right)$$

$$= \frac{1}{14}(4+3y)$$

The conditional density function of X given Y=y is

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

$$= \frac{\frac{6}{7}\left(x^2 + \frac{xy}{2}\right)}{\frac{1}{14}(4+3y)}$$

$$= \frac{12\left(x^2 + \frac{xy}{2}\right)}{(4+3y)}$$

The conditional density function of Y given X=x is

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

$$= \frac{\frac{6}{7}\left(x^2 + \frac{xy}{2}\right)}{\frac{6}{7}(2x^2+x)}$$

$$= \frac{\left(x^2 + \frac{xy}{2}\right)}{(2x^2+x)} = \frac{x + \frac{y}{2}}{2x+1}$$

15. The two solution of regression are $8x - 10y + 66 = 0, \quad 40x - 18y - 214 = 0$

The variance of X is 9. Find (i) mean value of X and Y

(ii) correlation coefficient between X and Y.[AU N/D 11'A/M 15']

Solution:

(i) mean value of X and Y at the point $(\bar{x}, \bar{y})$ must satisfied the two given regression

$$8\bar{x} - 10\bar{y} = -66 \text{...............(1)}$$

$$40\bar{x} - 18\bar{y} = 214 \text{.............(2)}$$

(1)x5 => $$40\bar{x} - 50\bar{y} = -330 \text{ ....(3)}$$

$$40\bar{x} - 18\bar{y} = 214$$

(2)-(3)=> $$32\bar{y} = 544$$

$$\overline{y} = 17.$$

Sub $\overline{y} = 17.$ in (1) we get

$$8\overline{x} - 10(17) = -66$$

$$8\overline{x} = -66 + 170$$

$$8\overline{x} = 104.$$

$$\overline{x} = 13.$$

Mean value give by $\overline{x} = 13., \overline{y} = 17.$

(ii) from the equation 1

$$8\overline{x} - 10\overline{y} = -66 \text{..............}(1)$$

$$10\overline{y} = 8x + 66$$

$$\overline{y} = \frac{8}{10}x + \frac{66}{10}$$

$$b_{YX} = \frac{8}{10}$$

from the equation 2. $40x - 18y = 214$

$$40x = 18y + 214$$

$$x = \frac{18}{40}y + \frac{214}{40}$$

$$b_{XY} = \frac{18}{40}$$

$$r = \pm\sqrt{b_{XY} \times b_{YX}}$$

$$= \pm\sqrt{\frac{18}{40} \times \frac{8}{10}}$$

$$r = \pm\frac{3}{5}$$

$$r = \pm 0.6.$$

Result: Both regression coefficient are positive "r" must be positive .

16. The regression equation of equation X and Y is $3y - 5x + 108 = 0$. if the mean value of y is 44 and the variance of X where $\frac{9}{16.}$ th of variance of a Y . find the mean value of X correlation Coefficient . **[AU A/M 2011]**

Solution:

Since the line of regression passes through $(\overline{x}, \overline{y})$. we get

$$3\overline{y} - 5\overline{x} + 108 = 0$$

$$3(44) - 5\overline{x} = 108$$

$$35\overline{x} = 108 + 132 = 240$$

$$\overline{x} = 48$$

The mean value of X is 48.

Given $3y-5x+108=0$

$$5x=3y+108$$

$$x=\frac{3}{5}y+\frac{108}{5}.$$

Which is line of recurrence $b_{XY}=\frac{3}{5}$

$$r\frac{\sigma_X}{\sigma_Y}=\frac{3}{5} \quad \ldots\ldots\ldots\ldots(1)$$

Given $\sigma_X{}^2=\frac{9}{16}\sigma_Y{}^2$

$$\sigma_X=\frac{3}{4}\sigma_Y \ldots\ldots\ldots\ldots(2)$$

Sub (2) in (1) we get

$$r\frac{3}{4}\frac{\sigma_X}{\sigma_Y}=\frac{3}{5}$$

$$r=\frac{12}{15}=0.8$$

Result: Correlation coefficient is r=0.8.

17. The tangent of angle between the lines of regression of Y on X and X on Y is 0.6 and $\sigma_X=\frac{1}{2}\sigma_Y$ Find the correlation coefficient between X and Y. [AU N/D '06, M/J '07]

Solution:

Formula: Angle between the lines of regression is

$$\tan\theta=\frac{1-r^2}{r}\frac{\sigma_X\sigma_Y}{\left(\sigma_X{}^2+\sigma_Y{}^2\right)}$$

$$0.6=\frac{1-r^2}{r}\frac{0.5\sigma_Y{}^2}{\left((.5\sigma_Y)^2+\sigma_Y{}^2\right)}$$

$$0.6=\frac{1-r^2}{r}\frac{0.5\sigma_Y{}^2}{\left((.25\sigma_Y{}^2)\ +\sigma_Y{}^2\right)}$$

$$\frac{3}{5}=\frac{1-r^2}{r}\left(\frac{\frac{1}{2}}{\frac{1}{4}+1}\right)$$

$$\frac{3}{5}=\frac{1-r^2}{r}\left(\frac{2}{5}\right)$$

$$3r = 2 - 2r^2$$
$$2r^2 + 3r - 2 = 0$$

Result: Solve this equation we get $r = \frac{1}{2}$ *or* $-\frac{1}{2}$

$$r = \frac{1}{2}$$

18. Two random variables X and Y have the following joint probability density function $f(x,y) = \begin{cases} 2-x-y, & 0 \le x \le 1, 0 \le y \le 1 \\ 0, & Otherwise \end{cases}$ **,Find Var(x),Var(y) and also the covariance between X and Y.Also find ρ_{XY}.**

Solution: **[A.U. N/D '18]**

Given: $f(x,y) = \begin{cases} 2-x-y, & 0 \le x \le 1, 0 \le y \le 1 \\ 0, & Otherwise \end{cases}$

Marginal density function of X is f(x)= $\int_{-\infty}^{\infty} f(x,y)dy$

$$f(x) = \int_0^1 (2-x-y)dy = \left[2y - xy - \frac{y^2}{2}\right]_0^1 = \frac{3}{2} - x$$

Marginal density function of Y is f(y)= $\int_{-\infty}^{\infty} f(x,y)dx$

$$f(y) = \int_0^1 (2-x-y)dx = \left[2x - \frac{x^2}{2} - xy\right]_0^1 = \frac{3}{2} - y$$

$$E(X) = \int_{-\infty}^{\infty} xf(x)dx = \int_0^1 x\left(\frac{3}{2} - x\right)dx = \left[\frac{3}{2}\left(\frac{x^2}{2}\right) - \left(\frac{x^3}{3}\right)\right]_0^1 = \frac{5}{12}$$

Similarly, $E(Y) = \frac{5}{12}$

$$E(X^2) = \int_{-\infty}^{\infty} x^2 f(x)dx = \int_0^1 x^2\left(\frac{3}{2} - x\right)dx = \left[\frac{3}{2}\left(\frac{x^3}{3}\right) - \left(\frac{x^4}{4}\right)\right]_0^1 = \frac{1}{4}$$

Similarly, $E(Y^2) = \frac{1}{4}$

Var(X)=E[X^2]-[E(X)]^2= $\frac{1}{4} - \left(\frac{5}{12}\right)^2 = \frac{11}{144}$

$$\therefore \sigma_X^2 = \frac{11}{144} \Rightarrow \sigma_X = \frac{\sqrt{11}}{12}$$

Similarly, $\sigma_Y^2 = \frac{11}{144} \Rightarrow \sigma_Y = \frac{\sqrt{11}}{12}$

$$E(XY)=\int_{-\infty}^{\infty}\int_{-\infty}^{\infty}xyf(x,y)dxdy=\int_0^1\int_0^1 xy(2-x-y)dxdy=\int_0^1\left[\frac{2x^2y}{2}-\frac{x^3y}{3}-\frac{x^2y^2}{2}\right]_0^1 dy$$

$$=\int_0^1\left[y-\frac{y}{3}-\frac{y^2}{2}\right]dy=\left[\frac{y^2}{2}-\frac{y^2}{6}-\frac{y^3}{6}\right]_0^1=\frac{1}{6}$$

Formula: Cov(X,Y)=E(XY) - E(X)E(Y)= $\frac{1}{6}-\left(\frac{5}{12}\right)\left(\frac{5}{12}\right)=-\frac{1}{144}$

Result:

$$\rho(X,Y)=\frac{Cov(X,Y)}{\sqrt{Var(X)\sqrt{Var(X)}}}=\frac{-\frac{1}{144}}{\frac{\sqrt{11}}{12}\frac{\sqrt{11}}{12}}=-\frac{1}{11}$$

.

19. If the joint pdf of (X,Y) is given by f(x,y)=x+y, 0≤x,y≤1. Find the correlation coefficient between X and Y. [AU M/J ’03, N/D ‘04(PQT)]

Solution:

Given: f(x,y)=x+y, 0≤x,y≤1.

Marginal density function of X is f(x)= $\int_{-\infty}^{\infty}f(x,y)dy$

$$f(x)=\int_0^1(x+y)dy=\left[xy+\frac{y^2}{2}\right]_0^1=x+\frac{1}{2}$$

Marginal density function of Y is f(y)= $\int_{-\infty}^{\infty}f(x,y)dx$

$$f(y)=\int_0^1(x+y)dx=\left[\frac{x^2}{2}+xy\right]_0^1=y+\frac{1}{2}$$

$$E(X)=\int_{-\infty}^{\infty}xf(x)dx=\int_0^1 x\left(x+\frac{1}{2}\right)dx=\left[\left(\frac{x^3}{3}\right)+\left(\frac{x^2}{4}\right)\right]_0^1=\frac{7}{12}$$

Similarly, $E(Y)=\frac{7}{12}$

$$E(X^2)=\int_{-\infty}^{\infty}x^2f(x)dx=\int_0^1 x^2\left(x+\frac{1}{2}\right)dx=\left[\left(\frac{x^4}{4}\right)+\left(\frac{x^3}{6}\right)\right]_0^1=\frac{5}{12}$$

Similarly, $E(Y^2)=\frac{5}{12}$

Var(X)=E[X^2]-[E(X)]^2=$\frac{5}{12}-\left(\frac{7}{12}\right)^2=\frac{11}{144}$

$$\therefore\sigma_X^2=\frac{11}{144}\Rightarrow\sigma_X=\frac{\sqrt{11}}{12}$$

Similarly, $\sigma_Y^2 = \frac{11}{144} \Rightarrow \sigma_Y = \frac{\sqrt{11}}{12}$

$$E(XY) = \int_{-\infty}^{\infty}\int_{-\infty}^{\infty} xyf(x,y)dxdy = \int_0^1\int_0^1 xy(x+y)dxdy = \int_0^1\left[\frac{x^3y}{3} + \frac{x^2y^2}{2}\right]_0^1 dy$$

$$= \int_0^1\left[\frac{y}{3} + \frac{y^2}{2}\right]dy = \left[\frac{y^2}{6} + \frac{y^3}{6}\right]_0^1 = \frac{1}{3}$$

Cov(X,Y)=E(XY) - E(X)E(Y)= $\frac{1}{3} - \left(\frac{7}{12}\right)\left(\frac{7}{12}\right) = -\frac{1}{144}$

Result:

$$\rho(X,Y) = \frac{Cov(X,Y)}{\sqrt{Var(X)}\sqrt{Var(X)}} = \frac{-\frac{1}{144}}{\frac{\sqrt{11}}{12}\frac{\sqrt{11}}{12}} = -\frac{1}{11}$$

20. Two random variables X and Y have the following joint probability density function $f(x,y) = \frac{6-x-y}{8},\ 0 \le x \le 2, 2 \le y \le 4$ **,Find the correlation coefficient between X and Y.**

Solution: **[A.U. N/D'08]**

Given: $f(x,y) = \frac{6-x-y}{8},\ 0 \le x \le 2, 2 \le y \le 4$

Marginal density function of X is f(x)= $\int_{-\infty}^{\infty} f(x,y)dy$

$$f(x) = \int_2^4\left(\frac{6-x-y}{8}\right)dy = \frac{1}{8}\left[6y - xy - \frac{y^2}{2}\right]_2^4 = \frac{1}{8}[6-2x]\ ,0 \le x \le 2$$

Marginal density function of Y is f(y)= $\int_{-\infty}^{\infty} f(x,y)dx$

$$f(y) = \int_0^2\left(\frac{6-x-y}{8}\right)dx = \frac{1}{8}\left[6x - \frac{x^2}{2} - xy\right]_0^2 = \frac{1}{8}[10-2y]\ ,2 \le y \le 4$$

$$E(X) = \int_{-\infty}^{\infty} xf(x)dx = \frac{1}{8}\int_0^1 x[6-2x]dx = \frac{1}{8}\left[6\left(\frac{x^2}{2}\right) - 2\left(\frac{x^3}{3}\right)\right]_0^2 = \frac{5}{6}$$

$$E(Y) = \int_{-\infty}^{\infty} yf(y)dy = \frac{1}{8}\int_2^4 y[10-2y]dy = \frac{1}{8}\left[10\left(\frac{y^2}{2}\right) - 2\left(\frac{x^3}{3}\right)\right]_2^4 = \frac{17}{6}$$

$$E(X^2) = \int_{-\infty}^{\infty} x^2f(x)dx = \frac{1}{8}\int_0^1 x^2[6-2x]dx = \frac{1}{8}\left[6\left(\frac{x^3}{3}\right) - 2\left(\frac{x^4}{4}\right)\right]_0^2 = 1$$

$$E(Y^2) = \int_{-\infty}^{\infty} y^2f(y)dy = \frac{1}{8}\int_2^4 y^2[10-2y]dy = \frac{1}{8}\left[10\left(\frac{y^3}{3}\right) - 2\left(\frac{y^4}{4}\right)\right]_2^4 = \frac{25}{3}$$

$$\text{Var(X)=E[X}^2\text{]-[E(X)]}^2=1-\left(\frac{5}{6}\right)^2=\frac{11}{36}$$

$$\therefore \sigma_X^2=\frac{11}{36}\Rightarrow \sigma_X=\frac{\sqrt{11}}{6}$$

$$\text{Var(Y)=E[Y}^2\text{]-[E(Y)]}^2=\frac{25}{3}-\left(\frac{17}{6}\right)^2=\frac{11}{36}$$

$$\therefore \sigma_Y^2=\frac{11}{36}\Rightarrow \sigma_Y=\frac{\sqrt{11}}{6}$$

$$E(XY)=\int_{-\infty}^{\infty}\int_{-\infty}^{\infty} xyf(x,y)dxdy=\int_2^4\int_0^2 xy\left(\frac{6-x-y}{8}\right)dxdy=\frac{1}{8}\int_2^4\left[6y\left(\frac{x^2}{2}\right)-\left(\frac{x^3}{3}\right)y-\frac{x^2y^2}{2}\right]_0^2 dy$$

$$=\frac{1}{8}\int_2^4\left[12y-\frac{8y}{3}-2y^2\right]dy=\frac{1}{8}\left[\frac{28}{3}\left(\frac{y^2}{2}\right)-2\left(\frac{y^3}{3}\right)\right]_2^4=\frac{7}{3}$$

$$\text{Cov(X,Y)=E(XY) - E(X)E(Y)}=\frac{7}{3}-\left(\frac{5}{6}\right)\left(\frac{17}{6}\right)=-\frac{1}{36}$$

Result:

$$\rho(X,Y)=\frac{Cov(X,Y)}{\sqrt{Var(X)}\sqrt{Var(X)}}=\frac{-\frac{1}{36}}{\frac{\sqrt{11}}{6}\frac{\sqrt{11}}{6}}=-\frac{1}{11}.$$

21. Let (X,Y) be the two-dimensional non-negative continuous random variable having the joint density.

$$f(x,y)=\begin{cases}4xy\,e^{-(x^2+y^2)}, & x\ge 0, y\ge 0\\ 0\ , & elsewhere\end{cases}$$ **Find the density function of U=$\sqrt{X^2+Y^2}$.**

Solution: **[AU N/D ’05, ‘08]**

Let $u=\sqrt{x^2+y^2}$, v=x

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

$$J=\begin{vmatrix}\frac{\partial u}{\partial x} & \frac{\partial u}{\partial y}\\ \frac{\partial v}{\partial x} & \frac{\partial v}{\partial y}\end{vmatrix}=\begin{vmatrix}\frac{x}{\sqrt{x^2+y^2}} & \frac{y}{\sqrt{x^2+y^2}}\\ 1 & 0\end{vmatrix}$$

$$=-\frac{y}{\sqrt{x^2+y^2}}$$

$$|J|=\frac{y}{\sqrt{x^2+y^2}}$$

$$f(u,v)=\frac{f(x,y)}{\frac{y}{\sqrt{x^2+y^2}}}$$

$$=4\,xy\,e^{-(x^2+y^2)}\left(\frac{\sqrt{x^2+y^2}}{y}\right)$$

$$=4x\sqrt{x^2+y^2}\;e^{-(x^2+y^2)}$$

$$=\begin{cases} 4\,uv\,e^{-u^2} & ,u\ge 0,\ \ 0\le v\le u \\ 0, & otherwise \end{cases}$$

Result: Hence the marginal density function of U is

$$f(u)=\int_0^u f(u\,,v)dv \quad =\int_0^u 4\,uv\,e^{-u^2}\;dv$$

$$=4\,u\,e^{-u^2}\int_0^u v\,dv=4\,u\;\;e^{-u^2}\left[\frac{v^2}{2}\right]_0^u$$

$$=4\,u\;e^{-u^2}\left[\frac{u^2}{2}\right]\;=2\;u^3e^{-u^2} \qquad ,u\ge 0$$

22. **If X and Y are independent R.V's with p.d.f** $e^{-x},\ x>o\ and\ e^{-y}\ ,\ y>0$ **, Find the density function of** $U=\frac{X}{X+Y}\ \ and\ \ V=X+Y$ **. Are** $U\ \ and\ \ V\ \ independen\,t\,?$

Solution: **[AU N/D 2024, A/M 2021]**

Given: Since X and Y are independent,

$$f(x,y)=e^{-x}\;\cdot e^{-y}\;=e^{-(x+y)}\quad ,x,y\ge 0$$

$$solving\ the\ equation\ \ u=\frac{x}{x+y}\ \ and\ v=x+y\ \ ,\ \ we\ \ get$$

$u=\frac{x}{x+y}$ . .....(1)	$v=x+y$ ......(2)
$u=\frac{x}{v}$ $by(2)$	$\Rightarrow y=v-x$
$\Rightarrow\ \ x=uv$ ......(3)	$\Rightarrow y=v-uv$ $by(3)$
	$i.e.,\ \ y=\ \ v(1-u)$

$\frac{\partial x}{\partial u} = v$ $\frac{\partial x}{\partial v} = u$	$\frac{\partial y}{\partial u} = -v$ $\frac{\partial y}{\partial v} = 1-u$

$$J = \begin{vmatrix} \frac{\partial x}{\partial u} & \frac{\partial x}{\partial v} \\ \frac{\partial y}{\partial v} & \frac{\partial y}{\partial v} \end{vmatrix} = \begin{vmatrix} v & u \\ -v & 1-u \end{vmatrix} = v(1-u)+uv \quad = v$$

The joint pdf of (U,V) is given by,

$$f(u,v) = \quad |J| \quad f(x,y)$$
$$= v \ e^{-(x+y)}$$
$$= v \ e^{-v}$$

$$Given: \quad x \ge 0 \Rightarrow \ uv \ \ge 0 \qquad \Rightarrow u \ge 0 \ \& \ v \ge 0$$
$$y \ge 0 \Rightarrow \ v(1-u) \ge 0 \qquad \Rightarrow v \ge 0 \ \& \ 1-u \ge 0$$
$$\Rightarrow 1 \ge u$$
$$\Rightarrow u \le 1$$

$$The\ pdf\ of\ U\ is\ given\ by,\ f(u) = \int_{-\infty}^{\infty} f(u,v)dv$$

$$= \int_0^{\infty} ve^{-v}\,dv \quad = \left[v\left(\frac{e^{-v}}{-1}\right) - (1)\left(\frac{e^{-v}}{(-1)^2}\right)\right]_0^{\infty}$$

$$= \left[-ve^{-v} \quad -e^{v}\right]_0^{\infty} \quad = [(0-0)-(-0-1)] \quad = 1$$

'U' is uniformly distributed in (0,1)

To find the pdf of v is given by

$$f(v) = \int_{-\infty}^{\infty} f(u,v)\ du$$
$$= \int_0^1 v\ e^{-v}\ du$$
$$= v\ e^{-v}[u]_0^1$$
$$= v\ e^{-v},\ v \ge 0$$

$$Now\ \ f(u).f(v) = v\ e^{-v} = f(u,v) \qquad by\ (2)$$

Result: $\Rightarrow U \ \ and \ \ V \ \ are\ independent \ \ RVs.$

23. If X and Y are independent R.V's with p.d.f

$f_X(x) = 1 \ in \ 1 \le x \le 2 \ \ and \ \ f_Y(y) = \frac{y}{6} \ in \ 2 \le y \le 4$, **Find the density function of** $Z = XY$ **.**

Solution

Given: X and Y are independent, $f_{xy}(x,y)=f_X(x).f_Y(y)$

$$f_{XY}(x,y) = \frac{y}{6}, \qquad 1 \le x \le 2 \ and \ 2 \le y \le 4$$

Let as consider the auxiliary random variable W = Y

$$\therefore \ x = \frac{z}{w} \ \ and \ \ y = w$$

$x = \frac{z}{w}$	Y = m
$\frac{\partial x}{\partial z} = \frac{1}{w}$ $\frac{\partial x}{\partial w} = \frac{-z}{w^2}$	$\frac{\partial y}{\partial z} = 0$ $\frac{\partial y}{\partial w} = 1$

$$J = \begin{vmatrix} \frac{\partial x}{\partial z} & \frac{\partial x}{\partial w} \\ \frac{\partial y}{\partial z} & \frac{\partial y}{\partial w} \end{vmatrix} = \begin{vmatrix} \frac{1}{w} & \frac{-z}{w^2} \\ 0 & 1 \end{vmatrix} = \frac{1}{w}$$

The joint pdf of (Z,W) is givn by,

$$f_{zw}(z,w) = |J| \ f_{XY}(x,y) = \frac{1}{w}\frac{y}{6} = \frac{1}{w}\frac{w}{6} = \frac{1}{6} \quad [\because y = w]$$

$Given: \quad 1 \le x \le 2$	$2 \le y \le 4$
$\Rightarrow 1 \le \frac{z}{w} \le 2$ $\Rightarrow 1 \le \frac{z}{w} \ and \ \frac{z}{w} \le 2$	$\Rightarrow \ 2 \le w \le 4 \ .....(3)$

$\Rightarrow w \le z \quad ....(1)$	and	$z \le 2w$	$....(2)$
$w = 2 \Rightarrow 2 \le z$	and	$z \le 4$	$i.e., 2 \le z \le 4$
$w = 4 \Rightarrow 4 \le z$	and	$z \le 8$	$i.e., 4 \le z \le 8$

$$z = 2w \qquad \Rightarrow w = \frac{1}{2}z$$

z	0	2	4	6	8
w	0	1	2	3	4

Result: The pdf of z is given by

$$f(z) = \int f(z,w)\; dw$$

$In\ 2 \le z \le 4,$
$$f(z) = \int_2^z \frac{1}{6}\, dw \;= \frac{1}{6}\int_2^w dw \;= \frac{1}{6}[w]_2^z \;= \frac{1}{6}[z-2]$$

$In 4 \le z < 8,$
$$f(z) = \int_{\frac{z}{2}}^4 \frac{1}{6}\, dw \;= \frac{1}{6}\int_{\frac{z}{2}}^4 dw$$

$$= \frac{1}{6}[w]_{z/2}^4 \;= \frac{1}{6}\left[4 - \frac{z}{2}\right] = \frac{1}{12}[8 - z]$$

24. The join p.d.f. of X and Y is given by f(x,y) = $e^{-(x+y)}$,x > 0 , y > 0, Find the probability density function of $U = \dfrac{X+Y}{2}$ **[AU N/D '06]**

Solution:

Given: $f(x,y) = e^{-(x+y)}$

$Let \qquad u = \dfrac{x+y}{2} \qquad ....(1)$ $u = \dfrac{x+y}{2} \quad by\ (2)$ $2u = x + v$ $x = 2u - v$	$v = y \qquad ....(2)$ $y = v$

$\frac{\partial x}{\partial u} = 2$ $\frac{\partial x}{\partial v} = -1$	$\frac{\partial y}{\partial u} = 0$ $\frac{\partial y}{\partial v} = 1$

$$J = \begin{vmatrix} \frac{\partial x}{\partial u} & \frac{\partial y}{\partial u} \\ \frac{\partial x}{\partial v} & \frac{\partial y}{\partial v} \end{vmatrix} = \begin{vmatrix} 2 & 0 \\ -1 & 1 \end{vmatrix} = 2$$

$$f(u,v) = f(x,y)|J|$$
$$= e^{-2u}\ (2) \qquad [\because x+y=2u \ \ by \ \ (1)]$$
$$= 2e^{-2u}$$

$$Given: \quad x>0 \quad \Rightarrow 2u-v>0 \Rightarrow 2u>v \qquad ....(1)$$
$$y>0 \quad \Rightarrow \quad v>0 \qquad ....(2)$$

$$\therefore f(u) = \int_0^{2u} f(u,v)\,dv$$
$$= \int_0^{2u} 2\,e^{-2u}\ dv$$
$$= 2\,e^{-2u}\int_0^{2u} dv$$
$$= 2\,e^{-2u}[v]_0^{2u}$$
$$= 2\,e^{-2u}[2u-0]$$
$$= 4u\ e^{-2u}, \qquad u>0$$

U	0	1	2	3
V	0	2	4	6

25. If X and Y are independent random variables each normally distributed with mean zero and Variance σ^2 ,find the density functions of $R = \sqrt{X^2 + Y^2}$ $and\ \theta = \tan^{-1}\left(\frac{Y}{X}\right)$

Solution: **[AU M/J '04 N/D '17]**

Given: Since X and Y are independent random variables normally distributed with mean zero and Variance σ^2,

The joint pdf of X and Y is given by

$$f_{xy}(x,y) = \frac{1}{2\pi\sigma^2} e^{\frac{-(x^2+y^2)}{2\sigma^2}}, \qquad -\infty < x, y < \infty$$

$$Given \ that \quad r = \sqrt{x^2+y^2} \quad and \quad \theta = \tan^{-1}\left(\frac{y}{x}\right)$$

Formula: $$\frac{\partial(x,y)}{\partial(r,\theta)} = \begin{vmatrix} \frac{\partial x}{\partial r} & \frac{\partial x}{\partial \theta} \\ \frac{\partial y}{\partial r} & \frac{\partial y}{\partial \theta} \end{vmatrix}$$

Since the given transformation is a polar form of (x,y), we have,

$$x = r\cos\theta \ ; \qquad y = r\sin\theta$$

$$Hence, \quad J = \frac{\partial(x,y)}{\partial(r,\theta)} = \begin{vmatrix} \cos\theta & -r\sin\theta \\ \sin\theta & r\cos\theta \end{vmatrix} = r$$

The joint pdf of (R,Ø) is given by

$$f_{r\theta}(r,\theta) = f_{xy}(x,y)\ |J|$$

$$= r\cdot\frac{1}{2\pi\sigma^2} e^{\frac{-(x^2+y^2)}{2\sigma^2}} = \frac{r}{2\pi\sigma^2} e^{\frac{-r^2}{e^2\sigma^2}}$$

Since $-\infty < x, y,$ we have $0 \le \theta \le 2\pi$ and $o \le r < \infty$.

Hence $f_{r\theta}(r,\theta) = \frac{r}{2\pi\sigma^2} e^{\frac{-r^2}{2\sigma^2}}, \ 0 \le \theta \le 2\pi, \quad 0 \le r < \infty$

The pdf of R is given by

$$f_r(r) = \int_0^{2\Pi} f_{r\theta}(r,\theta)\, d\theta = \frac{r}{2\pi\sigma^2} e^{\frac{-r^2}{2\sigma^2}} [\theta]_0^{2\Pi}$$

$$= \frac{r}{\sigma^2} e^{\frac{-r^2}{2\sigma^2}} \qquad 0 \le r \le \infty$$

$$f_\theta\theta = \int_0^\infty \frac{r}{2\pi\sigma^2} e^{\frac{-r^2}{2\sigma^2}} dr = \frac{1}{2\pi\sigma^2} \int_0^\infty r \frac{-r^2}{2\sigma^2} dr$$

$$Take \qquad u = \frac{r^2}{2\sigma^2} \Rightarrow du = \frac{2r\,dr}{2\sigma^2}$$

Result:

$$\therefore f_\theta\theta = \frac{1}{2\pi\sigma^2} \int_0^\infty \sigma^2 e^{-u} du = \frac{1}{2\pi}\int_0^\infty e^{-u} du$$

$$= \frac{1}{2\pi}\left[-e^{-u}\right]_0^\infty = \frac{-1}{2\pi}[-1] = \frac{1}{2\pi}, \qquad 0 \le \theta \le 2\pi$$

26. If X and y are independent random variables having density functions $f(x)=\begin{cases} 2e^{-2x}, & x\geq 0 \\ 0, & x<0\end{cases}$; $f(y)=\begin{cases} 3e^{-3y}, & y\geq 0 \\ 0, & y<0\end{cases}$ **respectively, find the density functions of Z =X - Y.** **[AU M/J '08]**

Solution:

$$f_{XY}(x,y) = f_X(x)\cdot f_Y(y) \qquad [\because X \text{ and } Y \text{ are independent}]$$
$$= \left(2e^{-2x}\right)\left(3e^{-3y}\right)$$
$$= 6\,e^{-(2x+3y)} \qquad x \,\&\, y \geq 0$$

Given: $U = X - Y$. Let W=Y.

The transformation functions are u = x - y

W = y

Solving we get x = u + w

Y = w

$$J = \frac{\partial(x,y)}{\partial(u,w)} = \begin{vmatrix} \frac{\partial x}{\partial u} & \frac{\partial x}{\partial w} \\ \frac{\partial y}{\partial u} & \frac{\partial y}{\partial w}\end{vmatrix} = \begin{vmatrix} 1 & 1 \\ 0 & 1\end{vmatrix} = 1$$

$$\therefore \quad f_{UW}(u,w) = |J| \; f_{XY}(x,y)$$
$$= (1)\,6\,e^{-(2x+3y)}$$
$$= 6\,e^{-(2(u+w)+3w)}$$
$$= 6\,e^{-(2u+2w+3w)} = 6\,e^{-(2u+5w)}$$

The range space (U,W) is obtained from the range space of (X,Y) using transformations

X = u + w and

Y = w

Since y≥0, w≥0 and

$$\therefore \quad f_{UW}(u,w) = 6\,e^{-(2u+5w)}, \quad w\geq 0, w\geq -u$$

$\therefore$ The marginal density of U is

$$f_U(u) = \int_{-\infty}^{\infty} f(u,w)\,dw$$

$$= \begin{cases} \int_{-u}^{\infty} f(u,w)\,dw & ,if\ \ u<0 \\ \int_{0}^{\infty} f(u,w)\ dw & if\ \ u \ge 0 \end{cases},$$

$$But \int_{-u}^{\infty} f(u\,,\,w)\,dw = \int_{-u}^{\infty} 6\,e^{-(2u+5w)}\ dw$$

$$= 6\,e^{-2u}\int_{-u}^{\infty} e^{-5w}\ dw$$

$$= 6\,e^{-2u}\cdot\left[\frac{e^{-5w}}{-5}\right]_{-u}^{\infty}$$

$$= \frac{6}{5}\ e^{3u}$$

$$\int_{0}^{\infty} f(u,w)\,dw = \int_{0}^{\infty} 6\ e^{-(2u+5w)}\ dw$$

$$= 6e^{-2u}\int_{0}^{\infty} e^{-5w}\ dw$$

$$= 6e^{-2u}\left[\frac{e^{-5w}}{-5}\right]_{0}^{\infty}$$

$$= \frac{6}{5}\ e^{-2u}$$

Result:

$$\therefore f_U(u) = \begin{cases} \frac{6}{5}\,e^{3u} & ,\ if\ \ u<0 \\ \frac{6}{5}\,e^{-2u} & ,\ \ if\ \ u \le 0 \end{cases}$$

27. If the p.d.f of two dimensional R.V (X,Y) is given by $f(x,y) = x+y,\ 0 \le x\ ,y \le 1.$

Find the p.d.f. of U =XY. **[AU N/D '06, '07, '18]**

Solution:

Given: u = xy and v = y

$$\therefore x = \frac{u}{v}\ \ and\ \ y = v$$

$$J = \begin{vmatrix} \frac{\partial x}{\partial u} & \frac{\partial x}{\partial v} \\ \frac{\partial y}{\partial u} & \frac{\partial y}{\partial v} \end{vmatrix} = \begin{vmatrix} \frac{1}{v} & \frac{-u}{v^2} \\ 0 & 1 \end{vmatrix} = \frac{1}{v}$$

The joint p.d.f. of u and v is given by

$$g(u,v) = f(x,y)\,|J| = (x+y)\,\frac{1}{v} = \left(\frac{u}{v}+v\right)\frac{1}{v}$$

$$= \frac{u}{v^2}+1 \qquad = 1+\frac{u}{v^2}$$

Since $0 \le y \le 1,\ 0 \le v \le 1$ *and* $0 \le x \le 1 \Rightarrow 0 \le u \le v$

v varies from v = u to v = 1

$\therefore$ The p.d.f of U is given by

$$f(u) = \int_{-\infty}^{\infty} g(u,v)\,dv$$

$$= \int_{u}^{1}\left[1+\frac{u}{v^2}\right] dv$$

$$= (v)_u^1 - \left(\frac{u}{v}\right)_u^1 = 1-u-u+1$$

$$\text{Result}: \qquad \therefore f(u) = 2(1-u), \quad 0 \le u \le 1$$

28. The joint probability density function of two random variables X and Y is given by

$$f(x,y) = \begin{cases} k\left[(x+y)-(x^2+y^2)\right], & 0<(x,y)<1 \\ 0\ , & otherwise \end{cases}$$ **Show that X and Y are uncorrelated but not independent.** **[AU M/J 2014]**

Solution:

To find k:

Given: $f(x,y) = \begin{cases} k\left[(x+y)-(x^2+y^2)\right], & 0<(x,y)<1 \\ 0\ , & otherwise \end{cases}$ is a pdf

$$\therefore \qquad \int_0^{\infty}\int_0^{\infty} f(x,y)\,dxdy = 1$$

$$\int_0^1\int_0^1 k\left[(x+y)-(x^2+y^2)\right]dxdy = 1$$

$$k\int_0^1\int_0^1 \left[(x+y-x^2-y^2)\right]dxdy = 1$$

$$k\int_0^1\left[\frac{x^2}{2}+xy-\frac{x^3}{3}-xy^2\right]_0^1 dy = 1 \Rightarrow k\int_0^1\left[\frac{1}{2}+y-\frac{1}{3}-y^2\right]dy = 1$$

$$k\int_0^1\left[y-y^2+\frac{1}{6}\right]dy=1\Rightarrow k\left(\frac{y^2}{2}-\frac{y^3}{3}+\frac{y}{6}\right)_0^1=1$$

$$\Rightarrow k\left(\frac{1}{2}-\frac{1}{3}+\frac{1}{6}\right)=1\Rightarrow k\left(\frac{3-2+1}{6}\right)=1$$

$$\Rightarrow k\left(\frac{1}{3}\right)=1 \Rightarrow k=3$$

To find $E(X)$ and $E(Y)$:

$$E(X)=\int_0^1\int_0^1 xf(x,y)dxdy$$

$$=3\int_0^1\int_0^1 x\left[x+y-x^2+y^2\right]dxdy$$

$$=3\int_0^1\int_0^1\left[x^2+xy-x^3+xy^2\right]dxdy$$

$$=3\int_0^1\left[\frac{x^3}{3}+\frac{x^2y}{2}-\frac{x^4}{4}+\frac{x^2y^2}{2}\right]_0^1 dy=3\int_0^1\left[\frac{1}{3}+\frac{y}{2}-\frac{1}{4}+\frac{y^2}{2}\right]dy$$

$$=3\int_0^1\left[\frac{y}{2}-\frac{y^2}{2}+\frac{1}{12}\right]dy=3\left[\frac{y^2}{4}-\frac{y^3}{6}+\frac{y}{12}\right]_0^1$$

$$=3\left[\frac{1}{4}-\frac{1}{6}+\frac{1}{12}\right]=3\left[\frac{2}{12}\right]=\frac{1}{2}$$

$$E(X)=\frac{1}{2}$$

Similarly, $E(Y)=\frac{1}{2}$

To find $E(XY)$:

$$E(XY)=\int_0^1\int_0^1 xyf(x,y)dxdy$$

$$=3\int_0^1\int_0^1 xy\left[x+y-x^2+y^2\right]dxdy$$

$$=3\int_0^1\int_0^1\left[x^2y+xy^2-x^3y+xy^3\right]dxdy$$

$$=3\int_0^1\left[\frac{x^3y}{3}+\frac{x^2y^2}{2}-\frac{x^4y}{4}+\frac{x^2y^3}{2}\right]_0^1 dy=3\int_0^1\left[\frac{y}{3}+\frac{y^2}{2}-\frac{y}{4}+\frac{y^3}{2}\right]dy$$

$$= 3\left[\frac{y^2}{6} + \frac{y^3}{6} - \frac{y^2}{8} - \frac{y^4}{8}\right]_0^1 = 3\left[\frac{1}{6} + \frac{1}{6} - \frac{1}{8} - \frac{1}{8}\right]$$

$$= 3\left[\frac{1}{12}\right] = \frac{1}{4}$$

$$E(XY) = \frac{1}{4}$$

To find $f_X(x)\, and\, f_Y(y):$

$$f_X(x) = \int_0^\infty f(x,y)dy$$

$$= \int_0^1 3\left[x + y - x^2 - y^2\right]dy$$

$$= 3\left[xy + \frac{y^2}{2} - x^2y - \frac{y^3}{3}\right]_0^1$$

$$= 3\left[x + \frac{1}{2} - x^2 - \frac{1}{3}\right] = 3\left[x - x^2 - \frac{1}{6}\right]$$

$$= \left[\frac{6x - 6x^2 + 1}{2}\right]$$

Similarly, $f_Y(y) = \left[\dfrac{6y - 6y^2 + 1}{2}\right]$

$$f_X(x)f_Y(y) = \left[\frac{6x - 6x^2 + 1}{2}\right]\left[\frac{6y - 6y^2 + 1}{2}\right] \neq f(x,y)$$

Result: Therefore X and Y are not independent.

$Cov(X,Y) = E(XY) - E(X)E(Y) = 0$

Therefore X and Y are uncorrelated.

29. Marks are obtained by 10 students in Mathematics (x) and Statistics (y) are given below.

x	**60**	**34**	**40**	**50**	**45**	**40**	**22**	**43**	**42**	**64**
y	**75**	**32**	**33**	**40**	**45**	**33**	**12**	**30**	**34**	**51**

Find the two regression lines. Also find y when x=55. **[AU M/J 2014]**

Solution:

X	Y	$X-\overline{X}$	$Y-\overline{Y}$	$\left(X-\overline{X}\right)^2$	$\left(Y-\overline{Y}\right)^2$	$\left(X-\overline{X}\right)\left(Y-\overline{Y}\right)^2$
60	75	16	36.5	256	1332.25	584
34	32	-10	-6.5	100	42.5	65
40	33	-4	-5.5	16	30.25	22
50	40	6	1.5	36	2.25	9
45	45	1	6.5	1	42.25	6.5
40	33	-4	-5.5	16	30.25	22
22	12	-22	-26.5	484	702.25	583
43	30	-1	-8.5	1	72.25	8.5
42	34	-2	-4.5	4	20.25	9
64	51	20	12.5	400	156.25	250
				1314	2430.50	1559

$$\overline{X} = \frac{\sum X}{10} = \frac{440}{10} = 44$$

$$\overline{Y} = \frac{\sum Y}{10} = \frac{385}{10} = 38.5$$

Formula:

$$b_{YX} = \frac{\sum \left(X-\overline{X}\right)\left(Y-\overline{Y}\right)^2}{\sum\left(X-\overline{X}\right)^2} = \frac{1559}{1314} = 1.1865$$

$$b_{YX} = \frac{\sum \left(X-\overline{X}\right)\left(Y-\overline{Y}\right)^2}{\sum\left(Y-\overline{Y}\right)^2} = \frac{1559}{2430.5} = 0.0005$$

Result: Equation of the line of regression of x on y is

$$\left(X-\overline{X}\right) = b_{XY}\,(Y-\overline{Y})$$

$$X - 44 = (0.0005)(Y - 38.5)$$

$$X = (0.0005)(Y) + 44 - 0.0188$$

$$X = (0.0005)(Y) + 43.9812.$$

Equation of the line of regression of y on x is

$$\left(Y-\overline{Y}\right) = b_{YX}\,(X-\overline{X})$$

$$Y - 38.5 = 1.1865(X - 44)$$

$$Y = 1.1865X + 38.5 - 52.2060$$

$$Y = 1.1865X - 13.7.60$$

When x=55, $Y = 1.1865(55) - 13.7.60$

$$Y = 51.5515.$$

30. If the joint pdf of a two dimensional RV(XY) is given by

$$f(x,y)=\begin{cases} x^2+\dfrac{xy}{3} & 0<x<1\ ,\quad 0<y<2 \\ 0 & Otherwise \end{cases}$$

Find $(i)\ P\left(X>\dfrac{1}{2}\right)$; $(ii)\ P(Y<X)$ *and* $(iii)\ P\left(Y<\dfrac{1}{2}\ \middle|X<\dfrac{1}{2}\right)$

Check Whether the Conditional density function are valid. **[AU M/J 2014]**

Solution:

Formula: The marginal density of X is given by, $f(x)=\int_{-\infty}^{\infty} f(x,y)\ dy$

$$=\int_0^2\left(x^2+\frac{xy}{3}\right)dy$$

$$=\left(x^2y+\frac{xy^2}{6}\right)_0^2$$

$$=\left(2x^2+\frac{2x}{3}\right),\ 0<x<1$$

The marginal density of Y is $f(y)=\int_{-\infty}^{\infty} f(x,y)\ dx\ =\int_0^1\left(x^2+\dfrac{xy}{3}\right)dx$

$$=\left[\frac{x^3}{3}+\frac{yx^2}{6}\right]_0^1$$

$$=\left(\frac{1}{3}+\frac{y}{6}\right)-(0+0)$$

$$=\frac{2+y}{3}$$

$(i)\ P\left(X>\dfrac{1}{2}\right)\ =\int_{1/2}^{1} f(x)\ dx\ =\int_{1/2}^{1}\left(2x^2+\dfrac{2x}{3}\right)dx=\left[\dfrac{2x^3}{3}+\dfrac{2x^2}{3}\right]_{1/2}^{1}\ =\dfrac{5}{6}$

$(ii)\ P(Y<X)\ \int_0^1\int_0^x\left(x^2+\dfrac{xy}{3}\right)dy\,dx=\int_0^1\left(yx^2+\dfrac{xy^2}{6}\right)_0^x dx=\int_0^1\left(x^3+\dfrac{x^3}{6}\right)dx=\dfrac{7}{6}\left(\dfrac{x^4}{4}\right)_0^1=\dfrac{7}{24}$

$$(iii)\ P\left(Y<\frac{1}{2}/X<\frac{1}{2}\right)=\frac{P\left(X<\frac{1}{2};Y<\frac{1}{2}\right)}{P\left(X<\frac{1}{2}\right)} \quad \ldots\ldots\ldots\ldots\ldots\ldots(1)$$

$$P\left(X<\frac{1}{2};Y<\frac{1}{2}\right)=\int_0^{1/2}\int_0^{1/2} f(x,y)\,dx\,dy=\int_0^{1/2}\int_0^{1/2}\left(x^2+\frac{xy}{3}\right)dx\,dy=\int_0^{1/2}\left[x^2y+\frac{xy^2}{6}\right]_0^{1/2}dx=\int_0^{1/2}\left[\frac{x^2}{2}+\frac{x}{24}\right]dx$$

$$=\left[\frac{x^3}{6}+\frac{x^2}{48}\right]_0^{1/2}=\frac{1}{48}+\frac{1}{192}=\frac{5}{192}$$

$$P\left(X<\frac{1}{2}\right)=\int_0^{1/2} f(x)\,dx=\int_0^{1/2}\left(2x^2+\frac{2x}{3}\right)dx=\left(\frac{2x^3}{3}+\frac{x^2}{3}\right)_0^{1/2}=\frac{1}{12}+\frac{1}{12}=\frac{1}{6}$$

$$(1)\Rightarrow P\left(Y<\frac{1}{2}/X<\frac{1}{2}\right)=\frac{5/192}{1/6}=\frac{5}{192}\times 6=\frac{5}{32}$$

Checking the conditional density functions are valid

$$\int_0^1 f(x/y)\,dx=\int_0^1\left(\frac{6x^2+2xy}{2+y}\right)dx=\frac{1}{2+y}\left(6\frac{x^3}{3}+2y\frac{x^2}{2}\right)_{x=0}^{x=1}$$

$$=\frac{1}{2+y}\left(\left(2+y-(0+0)\right)\right)=\frac{1}{2+y}(2+y)=1$$

$$\int_0^2 f(y/x)\,dy=\int_0^2\left(\frac{3x+y}{6x+2}\right)dy$$

$$=\frac{1}{6x+2}\int_0^2(3x+y)dy$$

$$=\frac{1}{6x+2}\left[3xy+\frac{y^2}{2}\right]_{y=0}^{y=2}$$

$$=\frac{1}{6x+2}\left[(6x+2)-(0+0)\right]$$

$$=1$$

Result: Hence The Conditional density function are valid.

31. The joint p.d.f of the random variable (X,Y) *is* $f(x,y)=3(x+y)$ $0\le x\le 1,\ 0\le y\le 1.,\ x+y\le 1.$ **Find** $Cov(X,Y)$.

Solution: **[AU M/J 2014]**

Given: $f(x,y)=3(x+y)$

$$f_X(x)=\int f(x,y)dy.$$

$$= \int_0^{1-x} (3x+3y)dy.$$

$$= \left[3xy + \frac{3y^2}{2}\right]_0^{1-x}$$

$$= \left[3x(1-x) + 3\frac{(1-x)^2}{2}\right]$$

$$= 3\left[x - x^2 + \frac{1}{2} - x + \frac{x^2}{2}\right]$$

$$= 3\left[\frac{1}{2} - \frac{x^2}{2}\right]$$

$$f_X(x) = \frac{3}{2}\left[1 - x^2\right]$$

$$E[X] = \int_0^1 x\ f(x)dx = \frac{3}{2}\int_0^1 (x - x^3)dx$$

$$= \frac{3}{2}\left[\frac{x^2}{2} - \frac{x^4}{4}\right]_0^1$$

$$E[X] = \frac{3}{8},$$

Similarly, $E[Y] = \dfrac{3}{8}$

Mean, $E[XY] = \int_0^1\int_0^{1-x} 3(x+y)xy\ dydx$

$$= 3\int_0^1\int_0^{1-x} (x^2y + xy^2)\ dydx$$

$$= 3\int_0^1 \left[\frac{x^2y^2}{2} + \frac{xy^3}{3}\right]_0^{1-x} dx$$

$$= 3\int_0^1 \left[\frac{(1-x)^2y^2}{2} + \frac{(1-x)^3x}{3}\right] dx$$

$$= 3\int_0^1 \left[\frac{x^2(1+x^2-2x)}{2} + \frac{x\left(1-3x+3x^2-x^3\right)}{3}\right] dx$$

$$= 3\int_0^1 \left[\frac{x^2+x^4-2x^3}{2} + \frac{x-3x^2+3x^3-x^4}{3}\right]dx$$

$$= \frac{3}{6}\int_0^1 \left[3x^2 + 3x^4 - 6x^3 + 2x - 6x^2 + 6x^3 - 2x^4\right]dx$$

$$= \frac{1}{2}\int_0^1 (x^4 - 3x^2 + 2x)\,dx.$$

$$= \frac{1}{2}\left[\frac{x^5}{5} - \frac{3x^3}{3} + \frac{2x^2}{2}\right]_0^1$$

$$= \frac{1}{2}\left[\frac{1}{5} - 1 + 1\right].$$

$$E[XY] = \frac{1}{10}$$

Result: $Cov(X,Y) = E(XY) - E(X)E(Y)$

$$= \frac{1}{10} - \frac{9}{64} = \frac{64-90}{640} = -\frac{26}{640} = -\frac{13}{320.}$$

32. Find the equation of the regression line Y on X from the following data: [AU A/M 15']

X	**3**	**5**	**6**	**8**	**9**	**11**
Y	**2**	**3**	**4**	**6**	**5**	**8**

Solution:

x	y	$x-\bar{x}$	$y-\bar{y}$	$(x-\bar{x})^2$	$(y-\bar{y})^2$	$(x-\bar{x})(y-\bar{y})$
3	2	-4	-2.67	16	7.1289	10.68
5	3	-2	0.33	4	0.1089	-0.66
6	4	-1	1.33	1	1.7689	-1.33
8	6	1	3.33	1	11.0889	1.33
9	5	2	4.33	4	18.7489	8.66
11	8	4	6.33	16	40.0689	25.32
$\sum x = 42$	$\sum y = 28$			$\sum(x-\bar{x})^2 = 42$	$\sum(y-\bar{y})^2 = 78.9134$	$\sum(x-\bar{x})(y-\bar{y}) = 44$

$$\bar{x} = \frac{\sum x}{n} = \frac{42}{6} = 7$$

$$\bar{y} = \frac{\sum y}{n} = \frac{42}{6} = 7$$

Formula: $$b_{yx} = \frac{\sum(x-\bar{x})(y-\bar{y})}{\sum(x-\bar{x})^2} = \frac{44}{42} = 1.0476$$

Equation of the line of regression of y on x is

$$(y-\bar{y}) = b_{yx}(x-\bar{x})$$

$$(y-4.67)=1.0476(x-7)$$
$$y=1.0476x-7.3332+4.67$$
$$y=1.0476x-2.6632$$

33. Assume that the random variables X and Y have the joint PDF $f(x,y)=\frac{1}{2}x^3y\ ;\ 0\le x\le 2, 0\le y\le 1$**. Determine if X and Y are independent.[AU A/M 15']**

Solution:

Given: $f(x,y)=\frac{1}{2}x^3y\ ;\ 0\le x\le 2, 0\le y\le 1$

To find the marginal density of X:

The marginal density of X is given by

$$f(x)=\int_{-\infty}^{\infty}f(x,y)dy$$
$$=\int_{-\infty}^{\infty}\frac{1}{2}x^3ydy$$
$$=\frac{1}{2}x^3\left(\frac{y^2}{2}\right)_0^1$$
$$=\frac{1}{4}x^3$$

To find the marginal density of Y:

The marginal density of Y is given by

$$f(y)=\int_{-\infty}^{\infty}f(x,y)dx$$
$$=\int_0^2\frac{1}{2}x^3ydx$$
$$=\frac{1}{2}y\left(\frac{x^4}{4}\right)_0^2$$
$$=\frac{1}{2}y(4)=2y$$

$$f(x)f(y)=\frac{x^3}{4}\times 2y=\frac{yx^3}{2}=f(x,y)$$

Result: $\therefore\ X\ \ and\ \ Y\ \ are\ \ independent$

34. The joint PDF of the random variables X and Y is defined as

$$f(x,y)=\begin{cases}25e^{-5y}; 0<x<0.2, y>0\\ 0, \quad otherwise\end{cases}$$

[AU A/M 15']

(i) Find the marginal PDFs of X and Y.
(ii) What is the covariance of X and Y?
Solution:

To Find the marginal PDFs of X and Y:

The marginal density function of X is given by

$$f(x) = \int_{-\infty}^{\infty} f(x,y)dy = \int_{0}^{\infty} 25e^{-5y}dy$$

$$= 25\left(\frac{e^{-5y}}{-5}\right)_0^{\infty} = 25\left(0+\frac{1}{5}\right)$$

$$f(x) = 5$$

The marginal density function of Y is given by

$$f(y) = \int_{-\infty}^{\infty} f(x,y)dx = \int_{0}^{0.2} 25e^{-5y}dx$$

$$= 25e^{-5y}(x)_0^{0.2} = 25e^{-5y}(0.2)$$

$$f(y) = 5e^{-5y}$$

To Find the covariance of X and Y:

$$E(X) = \int xf(x)dx$$

$$= \int_0^{0.2} 5x dx$$

$$= 5\left(\frac{x^2}{2}\right)_0^{0.2}$$

$$= 5(0.02)$$

$$= 0.1$$

$$E(Y) = \int yf(y)dy$$

$$= \int_0^{\infty} 5ye^{-5y}dy$$

$$= 5\left\{y\frac{e^{-5y}}{-5} - (1)\frac{e^{-5y}}{25}\right\}_0^{\infty}$$

$$= 5\left\{0+\frac{1}{25}\right\}$$

$$= \frac{1}{5}$$

$$E(XY) = \int_{-\infty}^{\infty}\int_{-\infty}^{\infty} xyf(x,y)dxdy$$

$$= \int_0^{\infty}\int_0^{0.2} xy25e^{-5y}dxdy$$

$$= 25\int_0^{\infty} ye^{-5y}\left(\frac{x^2}{2}\right)_0^{0.2}dy$$

$$= 25\int_0^{\infty} ye^{-5y}\left(\frac{0.04}{2}\right)dy$$

$$= 0.5\int_0^\infty ye^{-5y}\,dy$$

$$= 0.5\left\{\frac{ye^{-5y}}{-5} - (1)\frac{e^{-5y}}{25}\right\}_0^\infty$$

$$= 0.5\left\{0 + \frac{1}{25}\right\} = \frac{1}{50}$$

$$Cov(X,Y) = E(XY) - E(X)E(Y)$$

$$= \frac{1}{50} - (0.1)(0.2)$$

$$= 0.02 - 0.02$$

$$= 0$$

Result: $\therefore$ X *and* Y *are independent*

35. Find the constant k such that $f(x,y) = \begin{cases} k(x+1)e^{-y} & ,0 < x < 1, y > 0 \\ 0 & ,otherwise \end{cases}$ **is a joint probability density function of the continuous Random Variables (X,Y). Are X and Y independent Random Variables? Explain.** **[AU M/J 2016]**

Solution:

(1) The joint p.d.f $f(x,y) = k(1+x)e^{-y},\ 0 < x < 1, y > 0$

$$\int_{y=0}^{\infty}\int_{x=0}^{1} k(1+x)e^{-y}\,dx\,dy = 1$$

$$k\int_{x=0}^{1}(1+x)dx \int_{y=0}^{\infty} e^{-y}\,dy = 1$$

$$k\left(\frac{3}{2}\right) = 1$$

$$k = \frac{2}{3}$$

(2)The marginal p.d.f of X is $f(x) = \int_{y=0}^{\infty}\frac{2}{3}(1+x)e^{-y}\,dy$

$$= \begin{cases} \frac{2}{3}(x+1) & ,0 < x < 1 \\ 0 & ,otherwise \end{cases}$$

The marginal p.d.f of Y is

$$h(y) = \int_{x=0}^{1}\frac{2}{3}(1+x)e^{-y}\,dx$$

$$= \begin{cases} e^{-y} & , y > 0 \\ 0 & , otherwise \end{cases}$$

Since $f(x,y) = \frac{2}{3}(x+1)e^{-y} = g(x)h(y)$

$$\Rightarrow X \text{ and } Y \text{ are independent Random Variables}$$

36. The joint p.d.f of the continuous R.V (X,Y) is given as $f(x,y) = \begin{cases} e^{-(x+y)} & , x, y > 0 \\ 0 & , otherwise \end{cases}$

Find the p.d.f of the R.V $U = \frac{X}{Y}$

Solution: **[AU M/J 2016, N/D 2024]**

The joint p.d.f is $f(x,y) = \begin{cases} e^{-(x+y)} & , x, y > 0 \\ 0 & , otherwise \end{cases}$

Given that $U = \frac{X}{Y}$ $\quad Let\ V = Y \Rightarrow Y = V,\ X = UV$

$$J = \begin{vmatrix} \frac{\partial x}{\partial u} & \frac{\partial x}{\partial v} \\ \frac{\partial y}{\partial u} & \frac{\partial y}{\partial v} \end{vmatrix} = \begin{vmatrix} v & u \\ 0 & 1 \end{vmatrix} = v$$

$\therefore$ *The joint p.d.f of U and V is* $g(u,v) = ve^{-v(1+u)},\ u, v > o$

The marginal p.d.f of $U = \frac{X}{Y}$ *is*

$$h(u) = \int_{v=0}^{\infty} ve^{-v(1+u)} dv$$

$$= \left\{ v\left(\frac{e^{-(1+u)v}}{-(1+u)}\right) - (1)\left(\frac{e^{-(1+u)v}}{(1+u)^2}\right) \right\}_0^{\infty} = \frac{1}{(1+u)^2},\ u > 0$$

37. Let the joint p.d.f of R.V (X,Y) be given as $f(x,y) = \begin{cases} Cxy^2 & , 0 \le x \le y \le 1 \\ 0 & , otherwise \end{cases}$ **Determine (1) the value of C (2) the marginal p.d.fs of X and Y (3) the conditional p.d.f** $f(x/y)$ **of X given Y=y**

Solution: **[AU M/J 2016]**

The joint p.d.f of R.V (X,Y) be given as $f(x,y) = \begin{cases} Cxy^2 & , 0 \le x \le y \le 1 \\ 0 & , otherwise \end{cases}$

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

$$\int_{x=0}^{1}\int_{y=x}^{1} Cxy^2\,dy\,dx = 1 \Rightarrow C\int_{x=0}^{1}\left(\frac{xy^3}{3}\right)_x^1 dx = 1 \Rightarrow \frac{C}{3}\int_{x=0}^{1} x\left(1-x^3\right)dx = 1$$

$$\frac{C}{3}\int_{x=0}^{1}\left(x-x^4\right)dx = 1 \Rightarrow \frac{C}{3}\left[\frac{x^2}{2}-\frac{x^5}{5}\right]_0^1 = 1 \Rightarrow \frac{C}{3}\left[\frac{1}{2}-\frac{1}{5}\right] = 1$$

$$\frac{C}{3}\left[\frac{3}{10}\right] = 1 \Rightarrow C = 10$$

$$\therefore\quad f(x,y) = \begin{cases} 10xy^2 & ,0 \le x \le y \le 1 \\ 0 & ,otherwise \end{cases}$$

(2) $The\ \ marginal\ \ p.d.f\ \ of\ \ X\ \ is\ \ g(x) = \int_x^1 10xy^2\,dy = 10x\left(\frac{y^3}{3}\right)_x^1 = \frac{10}{3}x\left(1-x^3\right)$

$The\ \ marginal\ \ p.d.f\ \ of\ \ Y\ \ is\ \ h(y) = \int_0^y 10xy^2\,dx = 10y^2\left(\frac{x^2}{2}\right)_0^y = 5y^4$

(3) $The\ \ conditional\ \ p.d.f\ \ f(x/y) = \frac{f(x,y)}{h(y)} = \frac{10xy^2}{5y^4} = \frac{2x}{y^2}$

38. A joint probability mass function of the discrete R.Vs X and Y is given as

$$P(X = x, Y = y) = \begin{cases} \frac{x+y}{32} & , x = 1,2;\ y = 1,2,3,4 \\ 0 & ,otherwise \end{cases}$$ **Compute the covariance of X and Y.**

Solution: **[AU M/J 2016]**

The joint probability mass function is

$$P(X = x, Y = y) = \begin{cases} \frac{x+y}{32} & , x = 1,2;\ y = 1,2,3,4 \\ 0 & ,otherwise \end{cases}$$

$The\ \ marginal\ \ p.m.f\ \ of\ \ X\ \ is\ \ P(X = x) = \left(\frac{2x+5}{16}\right),\ x = 1,2$

$The\ \ marginal\ \ p.m.f\ \ of\ \ Y\ \ is\ \ P(X = y) = \left(\frac{2y+3}{32}\right),\ y = 1,2,3,4$

$X \setminus Y$	1	2	3	4	$P(X=x)$
1	$\frac{2}{32}$	$\frac{3}{32}$	$\frac{4}{32}$	$\frac{5}{32}$	$\frac{7}{16}=P(X=1)$
2	$\frac{2}{32}$	$\frac{4}{32}$	$\frac{5}{32}$	$\frac{6}{32}$	$\frac{9}{16}=P(X=2)$
$P(Y=y)$	$P(Y=1)=\frac{5}{32}$	$P(Y=2)=\frac{7}{32}$	$P(Y=3)=\frac{9}{32}$	$P(Y=4)=\frac{11}{32}$	**1**

$$E(X)=(1)\left(\frac{7}{16}\right)+(2)\left(\frac{9}{16}\right)=\frac{25}{16} \quad E(Y)=(1)\left(\frac{5}{32}\right)+(2)\left(\frac{7}{32}\right)+(3)\left(\frac{9}{32}\right)+(4)\left(\frac{11}{32}\right)$$

$$=\left(\frac{5+14+27+44}{32}\right)=\frac{90}{32}=\frac{45}{16}$$

$$E(XY)=(1)(1)\left(\frac{2}{32}\right)+(1)(2)\left(\frac{3}{32}\right)+(1)(3)\left(\frac{4}{32}\right)+(1)(4)\left(\frac{5}{32}\right)+(2)(1)\left(\frac{3}{32}\right)+(2)(2)\left(\frac{4}{32}\right)+(2)(3)\left(\frac{5}{32}\right)+(2)(4)\left(\frac{6}{32}\right)$$

$$=\frac{140}{32}=\frac{70}{16}$$

$$Cov(X,Y)=E(XY)-E(X)E(Y)=\frac{70}{16}-\left(\frac{25}{16}\right)\left(\frac{45}{16}\right)=4.375-(1.5625)(2.8125)=-0.01953$$

39. Two random variables X and Y have the following joint probability density function $f(x,y)=\begin{cases} c(4-x-y), & 0\le x\le 2, 0\le y\le 2 \\ 0, & elsewhere \end{cases}$**, Find the equations of two lines of regression.**

Solution: **[A.U N/D 2015]**

Given: $f(x,y)=\begin{cases} c(4-x-y), & 0\le x\le 2, 0\le y\le 2 \\ 0, & elsewhere \end{cases}$

We know that $f(x,y)$ should satisfy

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

$$\int_0^2\int_0^2 C(4-x-y)\,dydx=1$$

$$C\int_0^2\left[4y-xy-\frac{y^2}{2}\right]_0^2 dx=1$$

$$C\int_0^2[6-2x]dx=1 \Rightarrow C\left[6x-\frac{2x^2}{2}\right]_0^2=1 \Rightarrow C[(12-4)-(0-0)]=1$$

$$\Rightarrow 8C=1 \Rightarrow C=\frac{1}{8}$$

Marginal density function of X is $f(x)=\int_{-\infty}^{\infty} f(x,y)dy$

$$f(x)=\int_0^2\left(\frac{4-x-y}{8}\right)dy=\frac{1}{8}\left[4y-xy-\frac{y^2}{2}\right]_0^2=\frac{1}{4}[3-x]\ ,0\le x\le 2$$

Marginal density function of Y is $f(y)=\int_{-\infty}^{\infty} f(x,y)dx$

$$f(y)=\int_0^2\left(\frac{4-x-y}{8}\right)dx=\frac{1}{8}\left[4x-\frac{x^2}{2}-xy\right]_0^2=\frac{1}{4}[3-y]\ ,0\le y\le 2$$

$$E(X)=\int_{-\infty}^{\infty} xf(x)dx=\frac{1}{4}\int_0^2 x[3-x]dx=\frac{1}{4}\left[3\left(\frac{x^2}{2}\right)-\left(\frac{x^3}{3}\right)\right]_0^2=\frac{1}{4}\left[\left(6-\frac{8}{3}\right)-0\right]=\frac{5}{6}$$

$$E(Y)=\int_{-\infty}^{\infty} yf(y)dy=\frac{1}{4}\int_0^2 y[3-y]dy=\frac{1}{4}\left[3\left(\frac{y^2}{2}\right)-\left(\frac{x^3}{3}\right)\right]_0^2=\frac{1}{4}\left[6-\frac{8}{3}\right]=\frac{5}{6}$$

$$E(X^2)=\int_{-\infty}^{\infty} x^2 f(x)dx=\frac{1}{4}\int_0^1 x^2[3-x]dx=\frac{1}{4}\left[3\left(\frac{x^3}{3}\right)-\left(\frac{x^4}{4}\right)\right]_0^2=1$$

$$E(Y^2)=\int_{-\infty}^{\infty} y^2 f(y)dy=\frac{1}{4}\int_0^2 y^2[3-y]dy=\frac{1}{4}\left[3\left(\frac{y^3}{3}\right)-\left(\frac{y^4}{4}\right)\right]_0^2=1$$

$$Var(X)=E\left(X^2\right)-[E(X)]^2=1-\left(\frac{5}{6}\right)^2=\frac{11}{36}\Rightarrow\sigma_x=\frac{\sqrt{11}}{36}$$

$$Var(Y)=E\left(Y^2\right)-[E(Y)]^2=1-\left(\frac{5}{6}\right)^2=\frac{11}{36}\Rightarrow\sigma_y=\frac{\sqrt{11}}{36}$$

$$E(XY)=\int_{-\infty}^{\infty}\int_{-\infty}^{\infty} xyf(x,y)dxdy=\int_0^2\int_0^2 xy\left(\frac{4-x-y}{8}\right)dxdy=\frac{1}{8}\int_0^2\left[4y\left(\frac{x^2}{2}\right)-\left(\frac{x^3}{3}\right)y-\frac{x^2y^2}{2}\right]_0^2 dy$$

$$=\frac{1}{8}\int_0^2\left[8y-\frac{8y}{3}-2y^2\right]dy=\frac{1}{8}\left[\frac{16}{3}\left(\frac{y^2}{2}\right)-2\left(\frac{y^3}{3}\right)\right]_0^2=\frac{1}{8}\left[\frac{32}{3}-\left(\frac{16}{3}\right)\right]=\frac{2}{3}$$

$$Cov(X,Y)=E(XY)-E(X)E(Y)=\frac{2}{3}-\left(\frac{5}{6}\right)\left(\frac{5}{6}\right)=-\frac{1}{36}$$

$$r(X,Y)=\frac{Cov(X,Y)}{\sqrt{Var(X)}\sqrt{Var(X)}}=\frac{-\frac{1}{36}}{\frac{\sqrt{11}}{6}\frac{\sqrt{11}}{6}}=-\frac{1}{11}$$

The regression line of X on Y is *The regression line of Y on X is*

$$(X-\overline{X}) = r\frac{\sigma_x}{\sigma_Y}(Y-\overline{Y}) \qquad (Y-\overline{Y}) = r\frac{\sigma_y}{\sigma_x}(X-\overline{X})$$

$$\left(X-\frac{5}{6}\right) = (1)\left(\frac{-1}{11}\right)\left(Y-\frac{5}{6}\right) \qquad \left(Y-\frac{5}{6}\right) = (1)\left(\frac{-1}{11}\right)\left(X-\frac{5}{6}\right)$$

$$X = \frac{-1}{11}Y + \frac{10}{11} \qquad Y = \frac{-1}{11}X + \frac{10}{11}$$

Result:

$$\rho(X,Y) = \frac{Cov(X,Y)}{\sqrt{Var(X)}\sqrt{Var(X)}} = \frac{-\frac{1}{36}}{\frac{\sqrt{11}}{6}\frac{\sqrt{11}}{6}} = -\frac{1}{11}$$

.

40. The joint CDF of two discrete random variable X and Y is given by

$$F(x,y) = \begin{cases} 1/8 & x=1, y=1 \\ 5/8 & x=1, y=2 \\ 1/4 & x=2, y=1 \\ 1 & x=2, y=2 \end{cases}$$

Find the joint probability mass function and the marginal probability mass function of X and Y. **[AU N/D 2016]**

Solution:

$$P_{xy}(x,y) = \begin{cases} 1/8 & x=1, y=1 \\ 5/8 & x=1, y=2 \\ 1/4 & x=2, y=1 \\ 1 & x=2, y=2 \end{cases}$$

$$p_x(x) = \begin{cases} 1/8+1/2=5/8 & when\ x=1 \\ 1/8+1/4=3/8 & when\ x=2 \end{cases}$$

$$p_y(y) = \begin{cases} 1/8+1/8=1/4 & when\ y=1 \\ 1/2+1/4=3/4 & when\ y=2 \end{cases}$$

41. The joint probability density function of a two dimensional random variable (X,Y) is given by $f(x,y) = \frac{1}{8}x(x-y)$, $0<x<2, -x<y<x\ and\ 0\ elsewhere.$ **Find the marginal distributions of X and Y and the conditional distribution of Y=y given that X=x.**

Solution: **[AU N/D 2016]**

$$f(x) = \int_{-x}^{x}\frac{1}{8}(x^2-xy)\,dy = \frac{x^3}{4}\ in\ 0<x<2$$

$$f(y)=\int_{-y}^{2}\frac{1}{8}(x^2-xy)\,dx \quad in\ -2\le y\le 0\ and$$

$$\int_{y}^{2}\frac{1}{8}(x^2-xy)\,dx \quad in\ 0\le y\le 2$$

$$\therefore f(y)=\begin{cases}1/3-y/4+5/48y^3 & in\ -2\le y\le 0\\ 1/3-y/4+5/48y^3 & in\ 0\le y\le 2\end{cases}$$

$$f\left(\frac{y}{x}\right)=\frac{f(x,y)}{f(x)}=\frac{1}{2x^2}(x-y),-x<y<x$$

42. If X and Y each follow an exponential distribution with parameter 1 and are independent, find the probability density function of U=X-Y.

Solution: **[AU N/D 16]**

Given:

The pdf of X and y are

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

Since X and Y are independent,

$$f(x,y)=f(x)f(y)=e^{-x}\cdot e^{-y}=e^{-(x+y)}\quad ,x,y>0$$

$u=x-y$ . .....(1) $u=x-v$ $by(2)$ $\Rightarrow\ x=u+v$ ......(3)	$v=y$ ......(2) $\Rightarrow y=v$
$\frac{\partial x}{\partial u}=1$ $\frac{\partial x}{\partial v}=1$	$\frac{\partial y}{\partial u}=0$ $\frac{\partial y}{\partial v}=1$

$$J=\begin{vmatrix}\frac{\partial x}{\partial u} & \frac{\partial x}{\partial v}\\ \frac{\partial y}{\partial v} & \frac{\partial y}{\partial v}\end{vmatrix}=\begin{vmatrix}1 & 1\\ 0 & 1\end{vmatrix}=1$$

The joint pdf of (U,V) is given by,

$$f(u,v)=|J|\ f(x,y)$$
$$=e^{-(u+2v)}$$

$$Given:\ x>0\Rightarrow u+v>0\ \Rightarrow u>-v$$
$$y\ge 0\Rightarrow v>0$$

The pdf of U is given by,

$$\text{If } u<0, \text{ then } \quad f(u)=\int_{-\infty}^{\infty} f(u,v)\,dv$$

$$=\int_{-u}^{\infty} e^{-(u+2v)}\,dv = e^{-u}\int_{-u}^{\infty} e^{-2v}\,dv = e^{-u}\left[\frac{e^{-2v}}{-2}\right]_{-u}^{\infty}$$

$$=e^{-u}\left[0-\left(\frac{e^{2u}}{-2}\right)\right]=e^{-u}\,\frac{e^{2u}}{2}=\frac{e^{u}}{2}$$

$$\text{If } u\geq 0, \text{ then } \quad f(u)=\int_{0}^{\infty} f(u,v)\,dv$$

$$=\int_{0}^{\infty} e^{-(u+2v)}\,dv = e^{-u}\int_{0}^{\infty} e^{-2v}\,dv = e^{-u}\left[\frac{e^{-2v}}{-2}\right]_{0}^{\infty}$$

$$=e^{-u}\left[0-\left(\frac{1}{-2}\right)\right]=\frac{e^{-u}}{2}$$

43. If X and Y follows an exponential distribution with parameter 2 and 3 respectively and are independent, find the probability density function of U=X+Y. [AU A/M'17]

Solution:

$$\text{Given that } f(x)=2e^{-2x},\ x>0 \ \text{ and } \ g(y)=2e^{-2y},\ y>0$$

$$f_U(u)=\int f_x(u-y)g_y(y)dy$$

$$=\int_0^u 2e^{-2(u-y)}3e^{-3y}dy=6\int_0^u e^{-2u}e^{-2y}e^{-3y}dy=6e^{-2u}\int_0^u e^{-y}dy$$

$$=6e^{-2u}\left[-e^{-y}\right]_0^u=6e^{-2u}\left[1-e^{-u}\right],\ u>0 \qquad [\because x>0, y>0 \Rightarrow u>0]$$

44. Two random variables X and Y have the following joint probability density function $f(x,y)=xe^{-x(y+1)},\ x\geq 0, y\geq 0$. Determine the conditional probability density function of X given Y and the conditional probability density function of Y given X. [AU N/D 2017]

Solution:

Marginal density function of X is given by

$$f_X(x)=f(x)=\int_{-\infty}^{\infty} f(x,y)dy$$

$$=\int_0^{\infty} xe^{-x(y+1)}dy$$

Marginal density function of Y is given by

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

$$=\int_0^{\infty} xe^{-x(y+1)}dx$$

$$= x\left(\frac{e^{-x(y+1)}}{-x}\right)_0^\infty$$

$$= -\left(e^{-\infty} - e^{-x}\right)$$

$$= e^{-x}$$

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

$$= 0 - 0 + 0 + \frac{1}{(y+1)^2}$$

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

$\therefore M.d.f \ of \ 'X' is \ f(x) = e^{-x}$ $\therefore M.d.f \ of \ 'Y' is \ f(y) = \frac{1}{(y+1)^2}$

The conditional density function of X given Y is

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

$$= xe^{-x(y+1)}(y+1)^2$$

The conditional density function of Y given X is

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

$$= \frac{xe^{-x(y+1)}}{e^{-x}} = xe^{-xy}$$

45. The probability density function of (X,Y) is given by $f(x,y) = \begin{cases} 8xy, & 0 < x < y < 1 \\ 0 & otherwise \end{cases}$.

Find $P\left(x < \frac{1}{2} \cap y < \frac{1}{4}\right)$ **. Are X and Y independent? Justify your answer.**

Solution: **[AU A/M 2017]**

The marginal probability function of 'X' is

$$f_X(x) = f(x) = \int_{-\infty}^{\infty} f(x,y)dy$$

$$= \int_X^1 8xydy$$

$$= 8y\left[\frac{y^2}{2}\right]_x^1 = 4x\left(1 - x^2\right)$$

$$\therefore f(x) = 4x\left(1 - x^2\right) \text{..............(1)}$$

The marginal probability function of 'Y' is

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

$$= \int_0^y 8xydx$$

$$= 8y\left[\frac{x^2}{2}\right]_0^y = 4y^3$$

$$\therefore f(y) = 4y^3 \text{..............(2)}$$

Now $f(x).f(y) = 4x(1 - x \ )4y^3 \neq f(x,y)$ $[U \sin g \ (1) \ and \ (2)]$

$\therefore X \ and \ Y \ are \ not \ independent$

$$P\left(x < \frac{1}{2} \cap y < \frac{1}{4}\right) = \int_0^{\frac{1}{2}}\int_0^{\frac{1}{4}} 8xydydx = 8\int_0^{\frac{1}{2}} x\left(\frac{y^2}{2}\right)_0^{\frac{1}{4}} dx = 8\int_0^{\frac{1}{2}} \frac{x}{32} dx$$

$$= \frac{1}{4}\left(\frac{x^2}{2}\right)_0^{\frac{1}{2}} = \frac{1}{4}\left(\frac{1}{8}\right) = \frac{1}{32}$$

46. If the joint probability density function of (X,Y) is given by $f(x,y) = e^{-(x+y)}$ $x > 0,\ y > 0$ **.**

Prove that X and Y are uncorrelated. [AU A/M 2017]

Solution:

The marginal probability function of 'X' is

$$f_X(x) = f(x) = \int_{-\infty}^{\infty} f(x,y)dy$$

$$= \int_0^{\infty} e^{-(x+y)}dy$$

$$= e^{-x}\left[-e^{-y}\right]_0^{\infty} = e^{-x}(0+1)$$

$$\therefore f(x) = e^{-x}\ldots\ldots\ldots\ldots\ldots(1)$$

The marginal probability function of 'Y' is

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

$$= \int_0^{\infty} e^{-(x+y)}dx$$

$$= e^{-y}\left[-e^{-x}\right]_0^{\infty} = e^{-y}(0+1)$$

$$\therefore f(y) = 4y^3\ldots\ldots\ldots\ldots\ldots(2)$$

$$f(x).f(y) = e^{-x}e^{-y} = e^{-(x+y)} = f(x,y) \quad [U\sin g\ (1)\ and\ (2)]$$

Therefore X and Y are independent.

$$E(x) = \int_{-\infty}^{\infty} xf(x)dx$$

$$= \int_0^{\infty} xe^{-x}dx$$

$$= \left\{xe^{-x} - \left[-e^{-x}\right]\right\}_0^{\infty} = (0+1)$$

$$\therefore E(x) = 1\ldots\ldots\ldots\ldots\ldots(1)$$

$$E(y) = \int_{-\infty}^{\infty} yf(y)dy$$

$$= \int_0^{\infty} ye^{-y}dy$$

$$= \left\{ye^{-y} - \left[-e^{-y}\right]\right\}_0^{\infty} = (0+1)$$

$$\therefore E(y) = 1\ldots\ldots\ldots\ldots\ldots(1)$$

$$E(xy) = \int_{-\infty}^{\infty} xyf(x,y)dxdy = \int_0^{\infty} xye^{-(x+y)}dxdy$$

$$E(xy) = \int_0^{\infty} xe^{-x}dx\int_0^{\infty} ye^{-y}dx = 1\times 1 = 1$$

$Cov(X,Y) = E(XY) - E(X)E(Y) = 1 - 1\times 1 = 0$. Therefore X and Y are uncorrelated.

47. If the joint probability density function of the two dimensional random variable (X,Y) is given by $f(x,y) = \frac{x}{4}\left(1+3y^2\right),\ 0 < x < 2,\ 0 < y < 1$ **. Find**

(i) Conditional probability density functions of X given Y=y and Y given X=x.

(ii) $P\left[0.25 < X < 0.5 / Y = 0.33\right]$

Solution: [AU A/M 2018]

The marginal probability function of 'X' is

$$f_X(x) = f(x) = \int_{-\infty}^{\infty} f(x, y)dy$$

$$= \frac{1}{4}\int_0^1 \left(x + 3xy^2\right)dy$$

$$= \frac{x}{2},\ 0 < x < 2$$

The marginal probability function of 'Y' is

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

$$= \frac{1}{4}\int_0^2 \left(x + 3xy^2\right)dy$$

$$= \frac{1}{2}\left(1 + 3y^2\right),\ 0 < x < 1$$

$$f(x).f(y) = \frac{1}{4}\left(x + 3xy^2\right) = f(x, y)$$

$$\Rightarrow X \text{ and } Y \text{ are independent}$$

Therefore, $P(x / y) = P(x)$

$$P(0.25 < x < 0.5 / y = 0.33) = P(0.25 < x < 0.5)$$

$$= \int_{0.25}^{0.5} \frac{x}{2} dx = \frac{3}{64}$$

$$f(x / y) = \frac{f(x, y)}{f(y)} = \frac{x}{2},\ 0 < x < 2$$

$$f(x / y) = \frac{f(x, y)}{f(y)} = \frac{1}{2}\left(1 + 3y^2\right),\ 0 < y < 1$$

48. Given that $X = 4Y + 5$ and $Y = kX + 4$ are regression lines of X on Y and Y on X respectively. Show that $0 \le k \le \frac{1}{4}$. If $k = \frac{1}{16}$. Find the means of X and Y and the correlation coefficient r_{XY}.

Solution: **[AU A/M 2018]**

From the regression line of X on Y, $b_{XY} = 4$

From the regression line of Y on X, $b_{YX} = k$

$$r = \pm\sqrt{b_{xy}.b_{yx}} = \pm\sqrt{4k}$$

Since $-1 \le r \le 1,\ -1 \le \sqrt{4k} \le 1$

$$\Rightarrow 4k \le 1,\ 1 \le 4k \Rightarrow 1 \le 4k \le 1$$

$$\Rightarrow k \le \frac{1}{4},\ \frac{1}{4} \le k \ \text{ or } \ k \ge \frac{1}{4} \ \text{ i.e., } k \ge 0$$

$$\therefore\ 0 \le k \le 4$$

If $k = \frac{1}{16} \Rightarrow x - 4y = 5\text{.....(1)}$ and $\frac{1}{16}x - y = -4\text{......(2)}$

Solving (i) & (ii), we get mean values of x and y

$$\bar{x} = 28 \text{ and } \bar{y} = 5.75$$

49. The joint probability distribution of a two dimensional random variables X and Y is given by $f(x,y)=\begin{cases}\frac{1}{3}(x+y), & 0\le x\le 1, 0\le y\le 2\\ 0, & elsewhere\end{cases}$ **, Find the correlation coefficient. Also, find the equations of two lines of regression and conditional density function of X given Y and Y given X.**

Solution: **[A.U N/D 2018, 2024]**

Given: $f(x,y)=\begin{cases}\frac{1}{3}(x+y), & 0\le x\le 1, 0\le y\le 2\\ 0, & elsewhere\end{cases}$

Marginal density function of X is $f(x)=\int_{-\infty}^{\infty} f(x,y)dy$

$$f(x)=\int_0^2 \frac{1}{3}(x+y)dy=\frac{1}{3}\left[xy+\frac{y^2}{2}\right]_0^2=\frac{1}{3}[2x+2]=\frac{2(x+1)}{3} \quad ,0\le x\le 1$$

Marginal density function of Y is $f(y)=\int_{-\infty}^{\infty} f(x,y)dx$

$$f(y)=\int_0^1 \frac{1}{3}(x+y)dx=\frac{1}{3}\left[\frac{x^2}{2}+xy\right]_0^1=\frac{1}{3}\left[\frac{1}{2}+y\right]=\frac{1+2y}{6} \quad ,0\le y\le 2$$

$$E(X)=\int_{-\infty}^{\infty} xf(x)dx=\int_0^1 x\frac{2(x+1)}{3}dx=\frac{2}{3}\left[\left(\frac{x^3}{3}\right)+\left(\frac{x^2}{2}\right)\right]_0^1=\frac{2}{3}\left[\left(\frac{1}{3}\right)+\left(\frac{1}{2}\right)\right]=\frac{5}{9}$$

$$E(Y)=\int_{-\infty}^{\infty} yf(y)dy=\frac{1}{6}\int_0^2 y[1+2y]dy=\frac{1}{6}\left[\left(\frac{y^2}{2}\right)+2\left(\frac{y^3}{3}\right)\right]_0^2=\frac{1}{6}\left[2+\frac{16}{3}\right]=\frac{11}{9}$$

$$E(X^2)=\int_{-\infty}^{\infty} x^2 f(x)dx=\frac{2}{3}\int_0^1 x^2[x+1]dx=\frac{2}{3}\left[\left(\frac{x^4}{4}\right)+\left(\frac{x^3}{3}\right)\right]_0^1=\frac{2}{3}\left(\frac{7}{12}\right)=\frac{7}{18}$$

$$E(Y^2)=\int_{-\infty}^{\infty} y^2 f(y)dy=\frac{1}{6}\int_0^2 y^2[1+2y]dy=\frac{1}{6}\left[\left(\frac{y^3}{3}\right)+2\left(\frac{y^4}{4}\right)\right]_0^2=\frac{1}{6}\left(\frac{32}{3}\right)=\frac{16}{9}$$

$$Var(X)=E(X^2)-[E(X)]^2=\frac{7}{18}-\left(\frac{5}{9}\right)^2=\frac{13}{162}$$

$$Var(Y)=E(Y^2)-[E(Y)]^2=\frac{16}{9}-\left(\frac{11}{9}\right)^2=\frac{23}{81}$$

$$E(XY)=\int_{-\infty}^{\infty}\int_{-\infty}^{\infty} xyf(x,y)dxdy=\frac{1}{3}\int_0^2\int_0^1 xy(x+y)dxdy=\frac{1}{3}\int_0^2\left[y\left(\frac{x^3}{3}\right)+\left(\frac{x^2}{2}\right)y^2\right]_0^1 dy$$

$$=\frac{1}{3}\int_0^2\left[\frac{y}{3}+\frac{y^2}{2}\right]dy=\frac{1}{3}\left[\left(\frac{y^2}{6}\right)+\left(\frac{y^3}{6}\right)\right]_0^2=\frac{1}{3}\left[\frac{4}{6}+\left(\frac{8}{6}\right)\right]=\frac{2}{3}$$

$$Cov(X,Y)=E(XY)-E(X)E(Y)=\frac{2}{3}-\left(\frac{5}{9}\right)\left(\frac{11}{9}\right)=-\frac{1}{81}$$

$$r(X,Y)=\frac{Cov(X,Y)}{\sqrt{Var(X)}\sqrt{Var(X)}}=\frac{-\frac{1}{81}}{\sqrt{\frac{13}{162}}\sqrt{\frac{23}{81}}}=-\sqrt{\frac{2}{299}}$$

The regression line of X on Y is

$$(X-\overline{X})=r\frac{\sigma_x}{\sigma_Y}(Y-\overline{Y})$$

$$\left(X-\frac{5}{9}\right)=-\sqrt{\frac{2}{299}}\left(\sqrt{\frac{13}{162}}\sqrt{\frac{81}{23}}\right)\left(Y-\frac{11}{9}\right)$$

$$X=\frac{-1}{23}Y+\frac{14}{23}$$

The regression line of Y on X is

$$(Y-\overline{Y})=r\frac{\sigma_y}{\sigma_x}(X-\overline{X})$$

$$\left(Y-\frac{11}{9}\right)=-\sqrt{\frac{2}{299}}\left(\sqrt{\frac{23}{81}}\sqrt{\frac{162}{13}}\right)\left(X-\frac{5}{9}\right)$$

$$Y=\frac{-2}{13}X+\frac{17}{13}$$

Result:

$$r(X,Y)=\frac{Cov(X,Y)}{\sqrt{Var(X)}\sqrt{Var(X)}}=\frac{-\frac{1}{81}}{\sqrt{\frac{13}{162}}\sqrt{\frac{23}{81}}}=-\sqrt{\frac{2}{299}}$$

The conditional density function of X given Y=y is

$$f(x/y)=\frac{f(x,y)}{f(y)}=\frac{\frac{1}{3}(x+y)}{\frac{1+2y}{6}}$$

$$=\frac{2(x+y)}{1+2y}\qquad 0\le x\le 1$$

The conditional density function of Y given X=x is

$$f(y/x)=\frac{f(x,y)}{f(x)}=\frac{\frac{1}{3}(x+y)}{\frac{2}{3}(x+1)}$$

$$=\frac{(x+y)}{2(x+1)}\quad 0\le y\le 2$$

Central Limit Theorem

50. If $X_1, X_2, X_3, ... X_n$ are Poisson variables with parameter $\lambda = 2$, use the central limit theorem to estimate $P(120 < S_n < 160)$ where $S_n = X_1 + X_2 + X_3 + ... + X_n$: and n=75

Solution:

To find mean and variance $s^2 = 2, m = 2$

To find nm ns^2

$$nm = 75 \times 2 = 150$$

$$ns^2 = 75 \times 2 = 150$$

$$\sqrt{n}\, s = \sqrt{150}$$

Application of central limit theorem

$$S_n \sim N\left(\text{n}\, m, s\sqrt{n}\right)$$

$$S_n \sim N\left(150, \sqrt{150}\right)$$

To find $P(120 < \text{S}_n < 160)$

$$z = \frac{s_n - nm}{s\sqrt{n}}$$

$$z = \frac{s_n - 150}{\sqrt{150}}$$

If $S_n = 120$ $\quad z = \dfrac{120 - 150}{\sqrt{150}} = \dfrac{-30}{\sqrt{150}} = -2.45$

If $S_n = 160$ $\quad z = \dfrac{160 - 150}{\sqrt{150}} = \dfrac{10}{\sqrt{150}} = 0.85$

$$P(120 < \text{S}_n < 160) = P(-2.45 < Z < 0.85)$$

$= P(-2.45 < Z < 0) + P(0 < Z < 0.85)$ From Normal table

$$= 0.4927 + 0.2939 = 0.7866$$

RESULT: $P(120 < \text{S}_n < 160) = 0.7866$

51 .If $X_1, \text{X}_2, \text{X}_3, \ldots\ldots\ldots\ldots.\text{X}_{100}$ **be independent identically distributed random variables** $m = 2$ **,**

$s^2 = \dfrac{1}{4}$ **find** $P(192 < X_1 + \text{X}_2 + \text{X}_3 + \ldots\ldots\ldots\ldots. + \text{X}_{100} < 210)$.

Solution:

To find mean and variance $s^2 = \frac{1}{4}, m = 2$, n=100

To find nm ns^2

$$nm = 100 \times 2 = 200$$

$$ns^2 = 100 \times \frac{1}{4} = 25$$

$$\sqrt{n}\, s = 5$$

Application of central limit theorem

$$S_n \sim N\left(\mathrm{n}\, m, s\sqrt{n}\right)$$

$$S_n \sim N(200, 5)$$

To find $P(192 < \mathrm{S}_n < 210)$

$$z = \frac{s_n - nm}{s\sqrt{n}}$$

$$z = \frac{s_n - 200}{5}$$

If $S_n = 192$ $\quad z = \frac{192 - 200}{5} = -1.6$

If $S_n = 210$ $\quad z = \frac{210 - 200}{5} = 2$

$$P(192 < \mathrm{S}_n < 210) = P(-1.6 < Z < 2)$$

$$= P(-1.6 < Z < 0) + P(0 < Z < 2) \text{ From Normal table}$$

$$= 0.4452 + 0.4772 = 0.9224$$

RESULT: $P(192 < \mathrm{S}_n < 210) = 0.9224$

52 .The resisters $r_1, \mathrm{r}_2, \mathrm{r}_3$ and r_4 are independent identically distributed random variables and is uniform in the interval (450,550) , using the central limit theorem find $P\left(1900 \le r_1 + r_2 + r_3 + r_4 \le 2100\right)$.

Solution:

To find mean and variance $s^2 = 833.33, m = 500$, n=4

To find nm ns^2

$$nm = 4 \times 500 = 2000$$

$$ns^2 = 4 \times 833.3$$

$$\sqrt{n}\ s = 57.73$$

Application of central limit theorem

$$S_n : N\left(nm, s\sqrt{n}\right)$$

$$S_n : N(2000, 57.73)$$

To find $P(1900 < S_n < 2100)$

$$z = \frac{s_n - nm}{s\sqrt{n}}$$

$$z = \frac{s_n - 2000}{57.73}$$

If $S_n = 1900$ $\quad z = \frac{1900 - 2000}{57.73} = -1.73$

If $S_n = 210$ $\quad z = \frac{2100 - 2000}{57.73} = 1.73$

$$P(1900 < S_n < 2100) = P(-1.73 < Z < 1.73)$$

$$= P(-1.73 < Z < 0) + P(0 < Z < 1.73) \text{ From Normal table}$$

$$= 2(0.4582) = 0.9164$$

RESULT: $P(1900 < S_n < 2100) = 0.9164$

53. If $X_{i,} i = 1,2,3.....50$ **are independent identically distributed random variables having a Poisson distribution with parameter** $\lambda = 0.03$ **, use the central limit theorem to estimate** $P(S_n \geq 3)$ **where** $S_n = X_1 + X_2 + X_3 + ...+ X_n$ **: and n=75**

Solution:

To find mean and variance $\quad s^2 = 0.03, m = 0.03$, n=50

To find $\quad nm \quad ns^2$

$$nm = 50 \times 0.03 = 1.5$$

$$ns^2 = 50 \times 0.03$$

$$\sqrt{n}\ s = \sqrt{1.5}$$

Application of central limit theorem

$$S_n : N\left(nm, s\sqrt{n}\right)$$

$$S_n : N\left(1.5, \sqrt{1.5}\right)$$

To find $P(S_n \geq 3)$

$$z = \frac{s_n - nm}{s\sqrt{n}}$$

$$z = \frac{s_n - 1.5}{\sqrt{1.5}}$$

If $S_n = 3$ $\quad z = \dfrac{3 - 1.5}{\sqrt{1.5}} = \sqrt{1.5}$

$P(S_n \geq 3) = P(Z \geq \sqrt{1.5}) = 0.1112$ From Normal table

RESULT: $P(S_n \geq 3) = P(Z \geq \sqrt{1.5}) = 0.1112$

54. A coin is tossed 300 times .what is the probability that heads will appear more than 140 times and less than 150 times

Solution:

To find mean and variance

$$p = 0.5, q = 0.5$$

$$s^2 = npq = 300 \times 0.5 \times 0.5 = 75,$$

$$m = np = 300 \times 0.5 = 150, n=300$$

To find $P(140 < S_n < 150)$

$$z = \frac{X - m}{s}$$

$$z = \frac{X - 150}{\sqrt{75}}$$

If $X = 140$ $\quad z = \dfrac{140 - 150}{\sqrt{75}} = -1.15$

If $X = 150$ $\quad z = \dfrac{150 - 150}{\sqrt{75}} = 0$

$P(140 < S_n < 150) = P(-1.15 < Z < 0)$

$= P(0 < Z < 1.15) = 0.3749$ From Normal table

RESULT: $P(140 < X < 150) = 0.3749$

55. The life time of a certain brand of a tube light may be considered as a Random variable with mean 1200 h and standard deviation 250 h. Find the probability, using central limit theorem, that the average life time of 60 lights exceeds 1250 h.

Solution:

Given: mean $\mu = 1200$

S.D $\sigma^2 = 250$

$$\frac{\sigma}{\sqrt{n}} = \frac{250}{\sqrt{60}}$$

To find $P\left(\overline{X} \succ 1250\right)$

Let $z = \dfrac{\overline{X} - \mu}{\dfrac{\sigma}{\sqrt{n}}}$

$$= \frac{\overline{X} - 1200}{\frac{250}{\sqrt{60}}}$$

Now,

$$P(\overline{X} \succ 1250) = P\left(\frac{\overline{X} - 1200}{\frac{250}{\sqrt{60}}} \succ \frac{1250 - 1200}{\frac{250}{\sqrt{60}}}\right)$$

$$= P\left(z \succ \frac{50 \times \sqrt{60}}{250}\right)$$

$$= P(z \succ 1.55)$$

$$= P(0 \prec z \prec 3) - P(0 \prec z \prec 1.55)$$

$$= 0.5 - (Area\ from\ 0\ to\ 1.55)$$

$$= 0.5 - 0.4394$$

$$P(\overline{X} \succ 1250) = 0.0606$$

56. A random sample of size 100 is taken from a population whose mean is 60 and variance is 400. Using central limit theorem, with what probability can we assert that the mean of the sample will not differ from $\mu = 60$ by more than 4.

Solution:

Given: mean $\mu = 60$

S.D $\sigma^2 = 400$, n=100

$$\frac{\sigma}{\sqrt{n}} = \frac{250}{\sqrt{60}} = 2$$

To find $P(|\overline{X} - \mu| \le 4)$

Now,

$$P(|\overline{X} - \mu| \le 4) = P(-4 \le \underline{X} - \mu \le 4)$$

$$= P(-4 \le \underline{X} - 60 \le 4)$$

$$= P(60 - 4 \le \overline{X} \prec 60 + 4)$$

$$= P(56 \le \overline{X} \le 64)$$

Let $z = \dfrac{\overline{X} - \mu}{\frac{\sigma}{\sqrt{n}}}$

$$= \frac{\overline{X} - 60}{\frac{20}{\sqrt{100}}}$$

$$= \frac{\overline{X} - 60}{2}$$

$$P\left(56 \le \overline{X} \le 64\right) = P\left(\frac{56-60}{2} \le \frac{\overline{X}-60}{2} \le \frac{64-60}{2}\right)$$

$$= P(-2 \le z \le 2)$$

$$= P(0 \le z \le 2) + P(0 \le z \le 2)$$

$$= 2\{P(0 \le z \le 2)\}$$

$$= 2(0.4773)$$

$$P\left(\left|\overline{X} - \mu\right| \le 4\right) = 0.9546$$

57. The joint PDF of the random variables X and Y is defined as

$$f(x,y) = \begin{cases} ke^{-(x+y)}, & 0 \le x, y < \infty \\ 0, & otherwise \end{cases}$$

[AU A/M 2021]

(i) Find the value of k.

(ii) Find the marginal p.d.f of the random variables X and Y.

(iii) Find the conditional p.d.f of X given Y=y

Solution:

To find k:

$$\int\int f(x,y)dxdy = 1$$

$$\int_0^\infty\int_0^\infty ke^{-(x+y)}dxdy = 1$$

$$k\int_0^\infty e^{-x}dx\int_0^\infty e^{-y}dy = 1 \Rightarrow k\left(\frac{e^{-x}}{-1}\right)_0^\infty\left(\frac{e^{-y}}{-1}\right)_0^\infty = 1 \Rightarrow k(0+1)(0+1) = 1 \Rightarrow k = 1$$

The marginal probability density function of X is

$$f(x) = \int_0^\infty f(x,y)dy = \int_0^\infty e^{-(x+y)}dy = e^{-x}\left(\frac{e^{-y}}{-1}\right)_0^\infty = e^{-x}(0+1) = e^{-x}$$

The marginal probability density function of Y is

$$f(y) = \int_0^\infty f(x,y)dx = \int_0^\infty e^{-(x+y)}dx = e^{-y}\left(\frac{e^{-x}}{-1}\right)_0^\infty = e^{-y}(0+1) = e^{-y}$$

The conditional p.d.f X given Y is

$$f(x/y) = \frac{f(x,y)}{f(y)} = \frac{e^{-(x+y)}}{e^{-y}} = e^{-x}$$

58. The joint PMF of discrete random variables (X,Y) is given as

$$P(X=-1,Y=0)=\frac{1}{8}, P(X=-1,Y=1)=\frac{2}{8},\ P(X=1,Y=0)=\frac{3}{8}\ and\ P(X=1,Y=1)=\frac{2}{8}$$

Compute the correlation coefficient of X and Y.

Solution:

y \ x	-1	1	$p(y)$
0	$\frac{1}{8}$	$\frac{3}{8}$	$\frac{4}{8}$
1	$\frac{2}{8}$	$\frac{2}{8}$	$\frac{4}{8}$
$p(y)$	$\frac{3}{8}$	$\frac{5}{8}$	1

$$E(X)=\sum xp(x)=(-1)\left(\frac{3}{8}\right)+(1)\left(\frac{5}{8}\right)=\frac{2}{8}=\frac{1}{4}$$

$$E(X^2)=\sum x^2p(x)=(1)\left(\frac{3}{8}\right)+(1)\left(\frac{5}{8}\right)=\frac{8}{8}=1$$

$$E(Y)=\sum yp(y)=(0)\left(\frac{4}{8}\right)+(1)\left(\frac{4}{8}\right)=\frac{4}{8}=\frac{1}{2}$$

$$E(Y^2)=\sum y^2p(y)=(0)\left(\frac{4}{8}\right)+(1)\left(\frac{4}{8}\right)=\frac{4}{8}=\frac{1}{2}$$

$$E(XY)=\sum_i\sum_j x_iy_jp(x_iy_j)=(0)(-1)\left(\frac{1}{8}\right)+(0)(1)\left(\frac{3}{8}\right)+(1)(-1)\left(\frac{2}{8}\right)+(1)(1)\left(\frac{2}{8}\right)$$

$$=0+0-\frac{2}{8}+\frac{2}{8}=0$$

$$\sigma_X^2=E(X^2)-E(X)^2=1-\left(\frac{1}{4}\right)^2=\frac{15}{16}\Rightarrow\sigma_X=\frac{\sqrt{15}}{4}$$

$$\sigma_Y^2=E(Y^2)-E(Y)^2=\frac{1}{2}-\left(\frac{1}{2}\right)^2=\frac{1}{4}\Rightarrow\sigma_Y=\frac{1}{2}$$

$$r_{XY}=\frac{Cov(X,Y)}{\sigma_X\sigma_Y}=\frac{E(XY)-E(X)E(Y)}{\sigma_X\sigma_Y}=\frac{0-\left(\frac{1}{4}\right)\left(\frac{1}{2}\right)}{\left(\frac{\sqrt{15}}{4}\right)\left(\frac{1}{2}\right)}=\frac{-1}{\sqrt{15}}=-0.258$$

59. Two random variables X and Y have joint PDF

$$f(x,y)=\begin{cases}\frac{5}{16}x^2y, & 0<y<x<2\\ 0, & otherwise\end{cases}$$

[AU A/M 2021]

(i) Find the marginal p.d.f of the random variables X and Y.

(ii) Find the conditional p.d.f of X given Y=y.

(iii) Are the random variables X and Y independent? Justify

Solution:

The marginal probability density function of X is

$$f(x)=\int f(x,y)dy=\int_0^x \frac{5}{16}x^2y\,dy=\frac{5}{16}\left(\frac{x^2y^2}{2}\right)_0^{x}=\frac{5}{16}\left(\frac{x^4}{2}\right)=\frac{5x^4}{32}$$

The marginal probability density function of Y is

$$f(y)=\int f(x,y)dx=\int_y^2 \frac{5}{16}x^2y\,dy=\frac{5}{16}\left(\frac{x^3y}{3}\right)_y^2=\frac{5}{16}\left(\frac{8y}{3}-\frac{y^4}{3}\right)=\frac{5y(8-y^3)}{48}$$

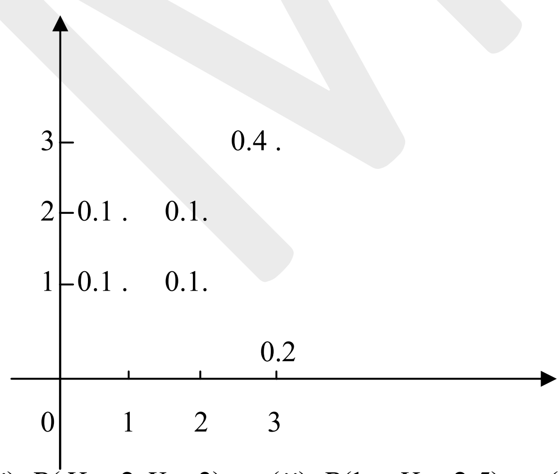
The conditional p.d.f X given Y is

$$f(x/y)=\frac{f(x,y)}{f(y)}=\frac{\frac{5}{16}x^2y}{\frac{5}{48}y(8-y^3)}=\frac{3x^2}{(8-y^3)}$$

$$f(x).f(y)=\frac{5}{32}x^4\times\frac{5}{48}y(8-y^3)=\frac{25x^4y}{1536}\neq f(x,y)$$

Therefore, X and Y are not independent.

60. For the discrete random variables X and Y with the joint distribution shown in the following figure: Determine the following:



$(i)\ P(X<2, Y<3)$ $(ii)\ P(1<X<2,5)$ $(iii)\ P(0<Y<2.5)$

$(iv)\ E(X),\ E(Y), V(X)\ and\ V(Y)$

(v) Marginal probability distribution of the random variable X and Y.

(vi) Conditional probability distribution of Y given that X=1.

(vii) Covariance and Correlation

(viii) Are X and Y independent?

Solution: **[AU A/M 2024]**

Y \ X	0	1	2	3	P(Y=y)
0	0	0	0	0.2	**P(y=0)=0.2**
1	0	0.1	0.1	0	**P(y=1)=0.2**
2	0	0.1	0.1	0	**P(y=2)=0.2**
3	0	0	0	0.4	**P(y=3)=0.4**
P(X=x)	**P(X=0)=0**	**P(X=1)=0.2**	**P(X=2)=0.2**	**P(X=3)=0.6**	**1**

(i) $P(X<2,Y<3)=P(0,0)+P(0,1)+P(0,2)+P(1,0)+P(1,1)+P(1,2)$

$P(X<2,Y<3)=0+0+0+0+0.1+0.1=0.2$

(ii) $P(1<X<2.5)=P(X=2)=0.2$

(iii) $P(0<Y<2.5)=P(Y=1)+P(Y=2)=0.2+0.2=0.4$

$$E(X)=\sum xp(x)=0(0)+1(0.2)+2(0.2)+3(0.6)$$

$$=0.2+0.4+1.8=2.4$$

$$E(Y)=\sum yp(y)=0(0.2)+1(0.2)+2(0.2)+3(0.4)$$

$$=0.2+0.4+1.2=1.8$$

$$E(X^2)=\sum x^2p(x)=0(0)+1(0.2)+4(0.2)+9(0.6)$$

$$=0.2+0.8+5.4=6.4$$

$$E(Y^2)=\sum y^2p(y)=0(0.2)+1(0.2)+4(0.2)+9(0.4)$$

$$=0.2+0.8+3.6=4.6$$

$$Var(x)=E(x^2)-\left(E(x)\right)^2=6.4-(2.4)^2=0.64$$

$$Var(y)=E(y^2)-\left(E(y)\right)^2=4.6-(1.8)^2=1.36$$

(v) Marginal distribution of X:

$P(X=0)=0$, $P(X=1)=0.2$ $P(X=2)=0.2$, $P(X=3)=0.6$

Marginal distribution of Y:

$P(Y=0)=0.2$, $P(Y=1)=0.2$, $P(Y=2)=0.2$ $P(Y=3)=0.4$

(vi) Conditional distribution of Y is given X is $P(Y=y\mid X=1)$

,
$$P(Y=0\mid X=1)=\frac{P(X=1,Y=0)}{P(X=1)}=0$$

$$P(Y=1\mid X=1)=\frac{P(X=1,Y=1)}{P(X=1)}=\frac{0.1}{0.2}=0.5$$

$$P(Y=2\mid X=1)=\frac{P(X=1,Y=2)}{p(X=1)}=\frac{0.1}{0.2}=0.5$$

$$P(Y=3\mid X=1)=\frac{P(X=1,Y=3)}{p(X=1)}=0$$

(vii) $E(XY)=\sum xyp(xy)$

$$=(1)(1)(0.1)+(1)(2)(0.1)+(2)((1)(0.1)+(2)(2)(0.1)+(3)(3)(0.4)=4.9$$

$$Cov(X,Y)=E(XY)-E(X)E(Y)$$

$$=4.9-(2.4)(1.8)=0.58$$

(viii) $P(1,,1)\neq P(X=1).P(Y=1)$

$0.1\neq(0.2)\ (0.2)$

$0.1\neq(0.4)$

Therefore X and Y are not independent

61. The joint pdf of X amount of almonds and Y amount of cashews were

$f(x,y)=\begin{cases}24xy, & 0\le x\le 1, 0\le y\le 1,\ x+y\le 1\\ 0 & otherwise\end{cases}$ **Find the covariance and correlation between the random variables X and Y.**

Solution: **[AU A/M 2024]**

Given $f(x.y)=\begin{cases}24xy & ,0\le x\le 1\ ,0\le y\le 1\ ,x+y\le 1\\ 0. & otherwise\end{cases}$

$$f_x(x)=\int_0^{1-x}24xy\,dy$$

$$f_x(x) = 24x\left(\frac{y^2}{2}\right)_0^{1-x}$$

$$= 12x(1-x)^2$$

$$= 12x(1-2x+x^2)$$

$$= 12(x-2x^2+x^3)$$

$$f_y(y) = 12y(1-y)^2$$

$$= 12y(1-2y+y^2)$$

$$= 12(y-2y^2+y^3)$$

$$E[x] = \int xf(x)dx = \int_0^1 x\ .12\left(x-2x^2+x^3\right)dx$$

$$= 12\int_0^1 \left(x^2-2x^3+x^4\right)dx$$

$$= 12\left(\frac{x^3}{3} - 2\frac{x^4}{4} + \frac{x^5}{5}\right)_0^1$$

$$= 12\left(\frac{1}{3} - \frac{1}{2} + \frac{1}{5}\right)$$

$$= 12\left(\frac{10-15+6}{30}\right)$$

$$= 12\left(\frac{1}{30}\right) = \frac{2}{5}$$

Similarly,

$$E[y] = 12\left(\frac{1}{30}\right) = \frac{2}{5}$$

$$E[xy] = \int_0^1 \int_0^{1-x} xy\ f(xy)\, dydx$$

$$= \int_0^1 \int_0^{1-x} 24x^2y^2 dydx$$

$$= \int_0^1 24x^2\left(\frac{y^3}{3}\right)_0^{1-x} dx$$

$$= 8\int_0^1 x^2(1-x)^3\, dx$$

$$= 8\int_0^1 \left(x^2-3x^3+3x^4-x^5\right)dx$$

$$= \left(\frac{x^3}{3} - 3\frac{x^4}{4} + 3\frac{x^5}{5} - \frac{x^6}{6} \right)_0^1$$

$$= \left(\frac{1}{3} - \frac{3}{4} + \frac{3}{5} - \frac{1}{6} \right)$$

$$= \frac{2}{15}$$

$$E[x^2] = \int_0^1 x^2 f(x)dx = \int_0^1 x^2 \ .12\left(x - 2x^2 + x^3\right)dx$$

$$= 12\int_0^1 \left(x^3 - 2x^4 + x^5\right)dx$$

$$= 12\left(\frac{x^4}{4} - 2\frac{x^5}{5} + \frac{x^6}{6} \right)_0^1$$

$$= 12\left(\frac{1}{4} - \frac{2}{5} + \frac{1}{6} \right)$$

$$= 12\left(\frac{15 - 24 + 10}{60} \right)$$

$$= 12\left(\frac{1}{60} \right) = \frac{1}{5}$$

$$Similarly,\ E[y^2] = \frac{1}{5}$$

$$Cov(x, y) = E(xy) - E(x)E(y)$$

$$= \frac{2}{15} - \frac{2}{5} . \frac{2}{5}$$

$$= \frac{2}{15} - \frac{4}{25} = -\frac{2}{15}$$

$$Var(x) = E(x^2) - [E(x)]^2 = \frac{1}{5} - \frac{4}{25} = \frac{1}{25}$$

$$Var(y) = E(y^2) - [E(y)]^2 = \frac{1}{5} - \frac{4}{25} = \frac{1}{25}$$

$$\rho(x, y) = \frac{Cov(x, y)}{\sqrt{Var(x)Var(y)}} = \frac{-\frac{2}{75}}{\frac{1}{5}.\frac{1}{5}} = \frac{-\frac{2}{75}}{\frac{1}{25}} = -\frac{2}{3}$$

62. A study of prices of rice a Chennai and Madurai gave the following data. Also the coefficient of correlation between the two is 0.8.

	Chennai	Madurai
Mean	19.5	17.75
Standard Deviation	1.75	2.5

Estimate he most likely price of rice

(1) At Chennai corresponding to the price of 18
(2) At Madurai corresponding to the price of 17 at Chennai.

Solution: **[AU N/D 2024]**

Given $E(X) = \vec{x} = 19.5$. $E(y) = \vec{y} = 17.75$

$\sigma_x = 1.75$ $\sigma_y = 2.5$, $r(x, y) = 0.8$

Equation of the line of regression of X on Y is

$$x - \bar{x} = r\frac{\sigma_x}{\sigma_y}\left(y - \bar{y}\right)$$

$$x - 19.5 = (0.8)\frac{1.75}{2.5}\left(y - 17.75\right)$$

$$x = 0.56y - 9.94 + 19.5$$

$$x = 0.56y + 9.56$$

$when \ \ y = 18,$

$$x = 0.56(18) + 9.56 = 19.64$$

Equation of the line of regression of Y on X is

$$y - \bar{y} = r\frac{\sigma_y}{\sigma_x}\left(x - \bar{x}\right)$$

$$y - 17.5 = (0.8)\frac{2.5}{1.75}\left(x - 19.5\right)$$

$$y = 1.428x - 4.5357$$

$when \ \ y = 17,$

$$y = 1.428(17) - 4.5357 = 19.74$$

63. Ten students got the following percentage of marks in Mathematical and Physical sciences. Calculate the rank correlation coefficient.

Maths	78	36	98	25	75	82	90	62	65	39
Physics	84	51	91	60	68	62	86	58	63	47

Solution: **[AU N/D 2024]**

Maths	Physics	Rank of Maths	Rank of Physics	d	d^2
78	84	4	3	-1	1
36	51	9	9	0	0
98	91	1	1	0	0
25	60	10	7	3	9
75	68	5	4	1	1
82	62	3	6	-3	9
90	86	2	2	0	0
62	58	7	8	-1	1
65	63	6	5	1	1
39	47	8	10	-2	4
					$\sum d^2 = 26$

The Rank Correlation

$$r = 1 - \frac{6\sum d^2}{n(n^2-1)} = 1 - \frac{6(26)}{10(99)} = 1 - \frac{156}{990} = 1 - 0.1575$$

$$r = 0.8424$$

Anna University Questions

1. The joint probability mass function of (X,Y) is given by p(x,y)=k(2x+3y) ,x=0,1,2; y=1,2,3. Find the all marginal and conditional distribution .als find the probability distribution .also find Probability distribution of (X+Y) and P[X+Y>3]. [pg. no. 23] [AU N/D 2011, '14]

2. Let X and Y be two random variables having the joint probability function f(x,y) =k(x+2y), where x and y can assume only the integer values 0, 1 and 2.Find the marginal and conditional distribution. [pg. no. 26] [AU M/J 2012]

3. The joint distribution of X and Y is given by f(x,y) = x+y/21, x = 1,2,3, y =1,2. Find the marginal distribution. [pg. no. 28] [AU N/D 2003,15 A/M '05,]

4. The two dimensional random variable (X,Y) has the joint density function f(x,y)=x+2y/27, x =0,1,2; y =0,1,2.Find the conditional distribution of Y given X = x. Also find the conditional distribution of X given Y=1. [pg. no. 29] [AU A/M '05, '08]

5. Calculate the correlation coefficient for the following heights (in inches) of fathers X and their sons Y. [pg. no. 31] [AU N/D 2007, 09,16]

X	65	66	67	67	68	69	70	72
Y	67	68	65	68	72	72	69	71

6. Calculate the correlation coefficient for the following data. [pg. no. 32] [AU Nov 2007, '09 '14]

X:	9	8	7	6	5	4	3	2	1
Y:	15	16	14	13	11	12	10	8	9

7. Calculate the correlation coefficient between industrial production and export using the following data. [pg. no. 33] [AU Nov 2007, 09]

Production(X)	55	56	58	59	60	60	62
Export(Y)	35	38	37	39	44	43	44

8. Calculate the correlation coefficient from the following data. [pg. no. 34] [AU Nov. 2008]

(X)	10	14	18	22	26	30
(Y)	18	12	24	6	30	36

9. If the independent random variables X and Y have the variances 36 and 16 respectively, find the Correlation coefficient between $U = X + Y \; and \; V = X - Y$ [pg. no. 35] [AU May '09, '14]

10. From the following data, find [pg. no. 36] [AU RP May '07, '09]
    (i) The two regression equations
    (ii) The coefficient of correlation between the marks in Economics and Statistics.
    (iii) The most likely marks in statistics when marks in Economics are 30.

Maks in Economics	25	28	35	32	31	36	29	38	34	32
Marks in Statistics	43	46	49	41	36	32	31	30	33	39

11. The joint p.d.f of the RV (X,Y) is given by $f(x,y) = Kxye^{-(x^2+y^2)}$ $x > 0, y > 0$. Find the value of K and prove also that X and Y are independent. [pg. no. 37] [AUM/J 2000,04.09]

12. If the joint distribution function of X and Y is given by $F(x,y) = (1-e^{-x})(1-e^{-y}),\ x > 0, y > 0$
 (i) Find the marginal densities of X and Y [pg. no. 39] [AU N/D '06, M/J '05]
 (ii) Are X and Y independent?
 (iii) Find $P(1 < X < 3, 1 < Y < 2)$

13. The joint p.d.f of X and Y is given by $f(x,y) = xy^2 + \frac{x^2}{8}, 0 \le x \le 2, 0 \le y \le 1$

 Compute (i) $P\left(X > 1 / Y < \frac{1}{2}\right)$ [pg. no. 40][AU Dec '09, Apr '08]
 (ii) $P\left(Y < \frac{1}{2} / X > 1\right)$
 (iii) $P(X < Y)$ (iv) $P(X + Y \le 1)$

14. The joint p.d.f of X and Y is given by $f(x,y) = \begin{cases} \frac{6}{7}\left(x^2 + \frac{xy}{2}\right), 0 \le x \le 1, 0 \le y \le 2 \\ 0, \quad otherwse \end{cases}$

 Find the conditional density function of X given Y and the conditional density function of Y given X. [pg. no. 43] [AU May 2014]

15. The two solution of regression are $8x - 10y + 66 = 0, \quad 40x - 18y - 214 = 0$
 The variance of X is 9. Find (i) mean value of X and Y
 (ii) correlation coefficient between X and Y. [pg. no. 44] [AU N/D '08, '11]

16. The regression equation of equation X and Y is $3y - 5x + 108 = 0$. if the mean value of y is 44 and the variance of X where $\frac{9}{16.}$ th of variance of a Y . find the mean value of X correlation Coefficient . [pg. no. 45] [AU A/M 2011]

17. The tangent of angle between the lines of regression of Y on X and X on Y is 0.6 and $\sigma_X = \frac{1}{2}\sigma_Y$ Find the correlation coefficient between X and Y. [pg. no. 46] [AU N/D '06, M/J '07]

18. Two random variables X and Y have the following joint probability density function
 $f(x,y) = \begin{cases} 2 - x - y, & 0 \le x \le 1, 0 \le y \le 1 \\ 0, & Otherwise \end{cases}$,Find Var(x),Var(y) and also the covariance between X and Y.Also find ρ_{XY} [pg. no. 47][A.U. N/D'18]

19. If the joint pdf of (X,Y) is given by f(x,y)=x+y, $0 \le x,y \le 1$. Find the correlation coefficient between X and Y. [pg. no. 48] [AU M/J '03, N/D '04(PQT)]

20. Two random variables X and Y have the following joint probability density function $f(x,y)=\dfrac{6-x-y}{8},\ 0\le x\le 2, 2\le y\le 4$,Find the correlation coefficient between X and Y.

[pg. no. 49] [A.U. N/D'08]

21. Let (X,Y) be the two-dimensional non-negative continuous random variable having the joint density. [pg. no. 50] [AU N/D '05, '08]

$$f(x,y)=\begin{cases}4xy\,e^{-(x^2+y^2)}, & x\ge 0, y\ge 0\\ 0\ , & elsewhere\end{cases}$$ Find the density function of U= $\sqrt{X^2+Y^2}$.

22. If X and Y are independent R.V's with p.d.f $e^{-x},\ x>o\ and\ e^{-y}\ ,\ y>0$, Find the density function of $U=\dfrac{X}{X+Y}\ and\ V=X+Y$. Are $U\ and\ V\ independen\ t$?

[pg. no. 51][AU N/D 2024, A/M 2021]

23. If X and Y are independent R.V's with p.d.f

$f_X(x)=1\ in\ 1\le x\le 2\ \ and\ \ f_Y(y)=\dfrac{y}{6}\ in\ 2\le y\le 4$, Find the density function of $Z=XY$. [pg. no. 53]

24. The join p.d.f. of X and Y is given by f(x,y) = $e^{-(x+y)}$,x > 0 , y > 0, Find the probability density function of $U=\dfrac{X+Y}{2}$ [pg. no. 54][AU N/D '06]

25. If X and Y are independent random variables each normally distributed with mean zero and Variance σ^2 ,find the density functions of $R=\sqrt{X^2+Y^2}\ \ and\ \theta=\tan^{-1}\left(\dfrac{Y}{X}\right)$.

[pg. no. 55] [AU M/J '04 '17]

26. If X and y are independent random variables having density functions

$$f(x)=\begin{cases}2e^{-2x}, & x\ge 0\\ 0 & , x<0\end{cases}\ ;\ f(y)=\begin{cases}3\,e^{-3y}, & y\ge 0\\ 0\ , & y<0\end{cases}$$ respectively, find the density functions of Z =X - Y. [pg. no. 57] [AU M/J '08]

27. If the p.d.f of two dimensional R.V (X,Y) is given by $f(x,y)\ =\ x+y,\ 0\le x\ ,y\le 1$.

Find the p.d.f. of U =XY. [pg. no. 58] [AU N/D '06, A/M '18]

28. The joint probability density function of two random variables X and Y is given by

$$f(x,y)=\begin{cases}k\left[(x+y)-(x^2+y^2)\right], & 0<(x,y)<1\\ 0\ , & otherwise\end{cases}$$ Show that X and Y are uncorrelated but not independent. [pg. no. 59] [AU M/J 2014]

29. Marks are obtained by 10 students in Mathematics (x) and Statistics (y) are given below.

x	60	34	40	50	45	40	22	43	42	64
y	75	32	33	40	45	33	12	30	34	51

Find the two regression lines. Also find y when x=55. [pg. no. 61] [AU M/J 2014]

30. If the joint pdf of a two dimensional RV(XY) is given by

$$f(x,y)=\begin{cases} x^2+\dfrac{xy}{3} & 0<x<1\ ,\quad 0<y<2 \\ 0 & Otherwise \end{cases}$$

Find $(i)\ P\left(X>\dfrac{1}{2}\right)$; $(ii)\ P(Y<X)$ and $(iii)\ P\left(Y<\dfrac{1}{2}\ \middle|X<\dfrac{1}{2}\right)$

Check Whether the Conditional density function are valid. [pg. no. 63] [AU M/J 2014]

31. The joint p.d.f of the random variable $(X,Y)\ is\ f(x,y)=3(x+y)$

$0\le x\le 1, 0\le y\le 1., x+y\le 1.$ Find $Cov(X,Y).$ [pg. no. 64] [AU M/J 2014]

32. Find the equation of the regression line Y on X from the following data:[Pg.no. 66] [AU A/M '15]

X	3	5	6	8	9	11
Y	2	3	4	6	5	8

33. Assume that the random variables X and Y have the joint PDF [Pg.no. 67][AU A/M '15]

$f(x,y)=\dfrac{1}{2}x^3y$; $0\le x\le 2, 0\le y\le 1$. Determine if X and Y are independent.

34. The joint PDF of the random variables X and Y is defined as

$$f(x,y)=\begin{cases} 25e^{-5y}; 0<x<0.2, y>0 \\ 0, \quad otherwise \end{cases}$$

[Pg. no. 67][AU A/M 15']

(i) Find the marginal PDFs of X and Y.

(ii) What is the covariance of X and Y?

35. Find the constant k such that $f(x,y)=\begin{cases} k(x+1)e^{-y} & ,0<x<1, y>0 \\ 0 & ,otherwise \end{cases}$ is a joint probability density function of the continuous Random Variables (X,Y). Are X and Y independent Random Variables? Explain. [Pg. no. 69][AU M/J 2016]

36. The joint p.d.f of the continuous R.V (X,Y) is given as $f(x,y)=\begin{cases} e^{-(x+y)} & ,x,y>0 \\ 0 & ,otherwise \end{cases}$

Find the p.d.f of the R.V $U=\dfrac{X}{Y}$ [Pg. no. 70] [AU M/J 2016, N/D 2024]

37. Let the joint p.d.f of R.V (X,Y) be given as $f(x,y)=\begin{cases} Cxy^2 & ,0 \le x \le y \le 1 \\ 0 & ,otherwise \end{cases}$ Determine (1) the value of C (2) the marginal p.d.fs of X and Y (3) the conditional p.d.f $f(x/y)$ of X given Y=y

[Pg. no. 70] [AU M/J 2016]

38. A joint probability mass function of the discrete R.Vs X and Y is given as

$$P(X=x,Y=y)=\begin{cases} \dfrac{x+y}{32} & ,x=1,2;\ y=1,2,3,4 \\ 0 & ,otherwise \end{cases}$$ Compute the covariance of X and Y.

[Pg. no. 71] [AU M/J 2016]

39. Two random variables X and Y have the following joint probability density function

$f(x,y)=\begin{cases} c(4-x-y), & 0 \le x \le 2, 0 \le y \le 2 \\ 0, & elsewhere \end{cases}$, Find the equations of two lines of regression.

[Pg. no. 72] [A.U N/D 2015]

40. The joint CDF of two discrete random variable X and Y is given by

$$F(x,y)=\begin{cases} 1/8 & x=1, y=1 \\ 5/8 & x=1, y=2 \\ 1/4 & x=2, y=1 \\ 1 & x=2, y=2 \end{cases}$$ Find the joint probability mass function and the marginal probability mass function of X and Y. [Pg. no. 74] [AU N/D 2016]

41. The joint probability density function of a two dimensional random variable (X,Y) is given by $f(x,y)=\frac{1}{8}x(x-y)$, $0<x<2, -x<y<x$ and 0 elsewhere. Find the marginal distributions of X and Y and the conditional distribution of Y=y given that X=x.

[Pg. no. 74] [AU N/D 2016]

42. If X and Y each follow an exponential distribution with parameter 1 and are independent, find the probability density function of U=X-Y. [Pg.no. 75][AU N/D 2016]

43. If X and Y follows an exponential distribution with parameter 2 and 3 respectively and are independent, find the probability density function of U=X+Y. [pg. no. 76][AU A/M'17]

44. Two random variables X and Y have the following joint probability density function $f(x,y)=xe^{-x(y+1)}$, $x \ge 0, y \ge 0$. Determine the conditional probability density function of X given Y and the conditional probability density function of Y given X.

[pg. no. 76] [AU N/D 2017]

45. The probability density function of (X,Y) is given by $f(x,y)=\begin{cases}8xy, & 0<x<y<1\\ 0 & otherwise\end{cases}$.

Find $P\left(x<\frac{1}{2}\cap y<\frac{1}{4}\right)$. Are X and Y independent? Justify your answer.

[pg. no. 77] [AU A/M 2017]

46. If the joint probability density function of (X,Y) is given by $f(x,y)=e^{-(x+y)} \quad x>0,\ y>0$. Prove that X and Y are uncorrelated. [pg. no. 78] [AU A/M 2017]

47. If the joint probability density function of the two dimensional random variable (X,Y) is given by $f(x,y)=\frac{x}{4}\left(1+3y^2\right),\ 0<x<2,\ 0<y<1$. Find

(i) Conditional probability density functions of X given Y=y and Y given X=x.

(ii) $P[0.25<X<0.5/Y=0.33]$ [pg. no. 78] [AU A/M 2018]

48. Given that $X=4Y+5 \;\; and \;\; Y=kX+4$ are regression lines of X on Y and Y on X respectively. Show that $0\le k\le\frac{1}{4}$. If $k=\frac{1}{16}$. Find the means of X and Y and the correlation coefficient r_{XY}. [pg. no. 79][AU A/M 2018]

49. The joint probability distribution of a two dimensional random variables X and Y is given by

$f(x,y)=\begin{cases}\frac{1}{3}(x+y), & 0\le x\le 1, 0\le y\le 2\\ 0, & elsewhere\end{cases}$, Find the correlation coefficient. Also, find the equations of two lines of regression and conditional density function of X given Y and Y given X.

[pg. no. 80] [A.U N/D 2018, 2024]

50. If $X_1, X_2, X_3, \ldots\ldots X_n$ are Poisson variables with parameter $\lambda=2$, use the central limit theorem to estimate $P(120<S_n<160)$ where $S_n = X_1+X_2+X_3+\ldots\ldots+X_n$: and n=75

[pg. no. 81]

51. If $X_1, X_2, X_3, \ldots\ldots X_{100}$ be independent identically distributed random variables $m=2$,

$s^2=\frac{1}{4}$ find $P(192<X_1+X_2+X_3+\ldots\ldots+X_{100}<210)$. [pg. no. 82]

52. The resisters r_1, r_2, r_3 and r_4 are independent identically distributed random variables and is uniform in the interval (450,550), using the central limit theorem find

$P(1900\le r_1+r_2+r_3+r_4\le 2100)$. [pg. no. 83]

53. If $X_{i,} i = 1,2,3.....50$ are independent identically distributed random variables having a Poisson distribution with parameter $\lambda = 0.03$, use the central limit theorem to estimate $P(S_n \geq 3)$ where $S_n = X_1 + X_2 + X_3 + ...+ X_n$: and n=75 [pg. no. 84]

54. A coin is tossed 300 times .what is the probability that heads will appear more than 140 times and less than 150 times. [pg. no. 85]

55. The life time of a certain brand of a tube light may be considered as a Random variable with mean 1200 h and standard deviation 250 h. Find the probability, using central limit theorem, that the average life time of 60 lights exceeds 1250 h. [pg. no. 85

56. A random sample of size 100 is taken from a population whose mean is 60 and variance is 400. Using central limit theorem, with what probability can we assert that the mean of the sample will not differ from $\mu = 60$ by more than 4. [pg. no. 86

57. The joint PDF of the random variables X and Y is defined as

$$f(x,y) = \begin{cases} ke^{-(x+y)}, & 0 \leq x, y < \infty \\ 0, & otherwise \end{cases}$$

[pg. no. 87][AU A/M 2021]

(i) Find the value of k.

(ii) Find the marginal p.d.f of the random variables X and Y.

(iii) Find the conditional p.d.f of X given Y=y

58. The joint PMF of discrete random variables (X,Y) is given as

$$P(X=-1,Y=0)=\frac{1}{8}, P(X=-1,Y=1)=\frac{2}{8}, \; P(X=1,Y=0)=\frac{3}{8} \; and \; P(X=1,Y=1)=\frac{2}{8}$$

Compute the correlation coefficient of X and Y. [pg. no. 88] [AU A/M 2021]

59. Two random variables X and Y have joint PDF

$$f(x,y) = \begin{cases} \frac{5}{16}x^2 y, & 0 < y < x < 2 \\ 0, & otherwise \end{cases}$$

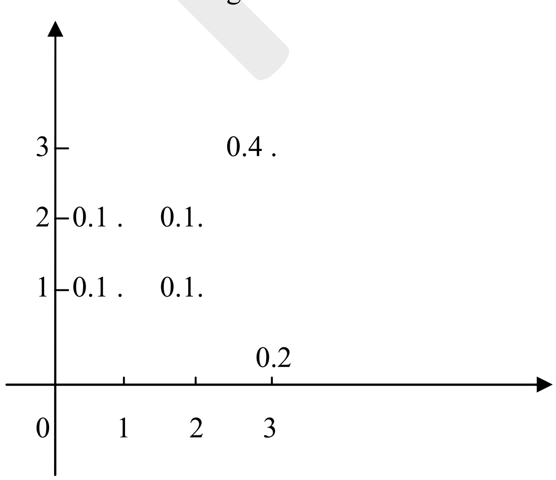
[pg. no. 88][AU A/M 2021]

(i) Find the marginal p.d.f of the random variables X and Y.

(ii) Find the conditional p.d.f of X given Y=y.

(iii) Are the random variables X and Y independent? Justify

60. For the discrete random variables X and Y with the joint distribution shown in the following figure: Determine the following:



$(i)\ P(X<2,Y<3)$ $(ii)\ P(1<X<2,5)$ $(iii)\ P(0<Y<2.5)$

$(iv)\ E(X),\ E(Y), V(X)\ and\ V(Y)$

(v) Marginal probability distribution of the random variable X and Y.

(vi) Conditional probability distribution of Y given that X=1.

(vii) Covariance and Correlation

(viii) Are X and Y independent? [AU A/M 2024][pg.no.89]

61. The joint pdf of X amount of almonds and Y amount of cashews were

$$f(x,y)=\begin{cases} 24xy, & 0\le x\le 1, 0\le y\le 1,\ x+y\le 1 \\ 0 & otherwise \end{cases}$$ Find the covariance and correlation between the random variables X and Y. [AU A/M 2024] [pg.no.91]

62. A study of prices of rice a Chennai and Madurai gave the following data. Also the coefficient of correlation between the two is 0.8.

	Chennai	Madurai
Mean	19.5	17.75
Standard Deviation	1.75	2.5

Estimate he most likely price of rice
(1) At Chennai corresponding to the price of 18
(2) At Madurai corresponding to the price of 17 at Chennai. [AU N/D 2024] [pg.no.94]

63. Ten students got the following percentage of marks in Mathematical and Physical sciences. Calculate the rank correlation coefficient.

Maths	78	36	98	25	75	82	90	62	65	39
Physics	84	51	91	60	68	62	86	58	63	47

[AU N/D 2024] [pg.no.94]

UNIT III
ESTIMATION THEORY
PART-A

1. **What is an Estimator?**
Solution:
Estimator (or) Point estimator is a procedure for producing an estimate of a parameter of interest. An estimator is usually a function of only sample data values, and when these data values are available, it results in an estimator of the parameter of interest.

2. **Define a Point Estimator.**
Solution:
A point estimator of some population parameter θ is single numerical values $\overline{\theta}$ of a statistic $\overline{\theta}$. The statistic $\overline{\theta}$ is called the point estimator.
For example, the sample mean $\left(\overline{X}\right)$ is a point estimate of the population mean μ. Similarly, the sample proportion p is a point estimate of the population proportion P.

3. **Define an Interval Estimator.**
Solution:
An interval estimate is defined by two numbers, between which a population parameter is said to lie.
For example, $a < X < b$ is an interval estimate of the population mean μ. It indicates that the population mean is greater than a but less than b.

4. **What are the characteristics that should be satisfied by a good estimator?**
Solution:
(i) Unbiasedness (ii) Consistency (iii) Efficiency (iv) Sufficiency

5. **What are the commonly used methods of point estimation?**
Solution:
(i) Methods of maximum likelihood estimator.
(ii) Method of minimum variance.
(iii) Method of moments.
(iv) Method of least squares.
(v) Method of minimum chi-square.
(vi) Method of inverse probability.

6. **Define an Unbiased estimator.**
Solution:
An estimator $T_n = T(x_1, x_2, x_3, ..., x_n)$ is said to be an unbiased estimator of $\gamma(\theta)$ if $E(T_n) = \gamma(\theta)$, $\forall\, \theta \in \Theta$.
If $E(T_n) > \theta$ then T_n is said to be positively biased.
If $E(T_n) < \theta$ then T_n is said to be negatively biased.

7. Define a Consistency estimator.
Solution:

An estimator $T_n = T(x_1, x_2, x_3, ..., x_n)$ based on a random sample of size n, is said to be consistent estimator of $\gamma(\theta)$, $\forall\, \theta \in \Theta$, the parameter space, if T_n converges to $\gamma(\theta)$ in probability .

8. Define an Efficiency estimator.
Solution:

If T_1 is the most efficient estimator with variance V_1 and T_2 is any other estimator with variance V_1 , then the efficiency E of T_2 is defined as:

$$E = \frac{V_1}{V_2}$$

Obviously, E cannot exceed unity.

9. Define a Sufficiency estimator.
Solution:

An estimator is said to be sufficient for a parameter θ, if it contains all the information in the sample regarding the parameter.

10. Define a maximum likelihood estimator of θ. **[AU N/D 2024]**
Solution:

Let $x_1, x_2, x_3, ..., x_n$ be a random sample of size n from a population with density function $f(x,\theta)$. Then the likelihood function of the sample values $x_1, x_2, x_3, ..., x_n$, usually denoted by $L = L(\theta)$ is their joint density function, given by:

$$L = f(x_1,\theta) f(x_1,\theta) ....f(x_1,\theta) = \prod_{i=1}^{n} f(x_i,\theta)$$

11. Define minimum variance Unbiased (M.V.U) estimators.
Solution:

If a statistic $T_n = T(x_1, x_2, x_3, ..., x_n)$, based on sample of size n is such that:

(i) T is unbiased for $\gamma(\theta)$, for all $\theta \in \Theta$ and

(ii) It has the smallest variance among the class of all unbiased estimators of $\gamma(\theta)$, then T is called the minimum variance unbiased estimator of $\gamma(\theta)$

12. State any two properties of maximum likelihood estimators.
Solution:

(i) The first and second order derivatives, viz., $\frac{\partial \log L}{\partial \theta}$ *and* $\frac{\partial^2 \log L}{\partial \theta^2}$ exist and are continuous functions of θ in a range R for almost all x . For every θ *in* R,

$\left|\frac{\partial \log L}{\partial \theta}\right| < F_1(x)$ *and* $\left|\frac{\partial^2 \log L}{\partial \theta^2}\right| < F_2(x)$ where $F_1(x)$ *and* $F_2(x)$ are integrable functions over $(-\infty, \infty)$

(ii) The third order derivative $\dfrac{\partial^3 \log L}{\partial\theta^3}$ exists such that $\left|\dfrac{\partial^3 \log L}{\partial\theta^3}\right| < M(x)$

Where $E[M(x)] < K,\ a\ positive\ quantity$

13. State Factorization theorem (Neymann).

Solution:

$T = t(x)$ is sufficient for θ if and only if the joint density function L (say), of the sample values can be expressed in the form:

$$L = g_\theta[t(x)]h(x)$$

Where $g_\theta[t(x)]$ depends on $\theta\ and\ x$ only through the value of $t(x)\ and\ h(x)$ is independent of θ .

14. State Invariance property of sufficient estimator.

Solution:

If T is a sufficient estimator for the parameter θ and if $\Psi(T)$ is a one to one function of T, then $\Psi(T)$ is sufficient for $\Psi(\theta)$.

15. State Fisher-Neymann Criterion.

Solution:

A statistic $t_1 = t(x_1, x_2, \ldots\ldots, x_n)$ is sufficient estimator of parameter θ if and only if the likelihood function (joint p.d.f of the sample) can be expressed as:

$$L = \prod_{i=1}^{n} f(x_i, \theta) = g_1(t_1, \theta).k(x_1, x_2, \ldots\ldots, x_n)$$

Where $g_1(t_1, \theta)$ the p.d.f of the statistic is $t_1\ and\ k(x_1, x_2, \ldots\ldots, x_n)$ is a function of sample observations only, independent of θ.

16. State Cramer-Rao Inequality.

Solution:

If t is an unbiased estimator for $\gamma(\theta)$, a function of parameter θ, then

$$Var(t) \geq \frac{\left\{\dfrac{d}{d\theta}.\gamma(\theta)\right\}^2}{E\left(\dfrac{\partial}{\partial\theta}\log L\right)^2} = \frac{\{\gamma'(\theta)\}^2}{I(\theta)}$$

where $I(\theta)$ is the information on θ, supplied by the sample.

17. State Cramer-Rao Theorem.

Solution:

"With probability approaching unity as $n \to \infty$ the likelihood equation $\dfrac{\partial}{\partial\theta}\log L = 0$, has a solution which converges in probability to the true value θ_0". In other words M.L.E's are consistent.

18. State Hazoor Bazar's Theorem.

Solution:

Any consistent solution of the likelihood equation provides a maximum of the likelihood with probability tending to unity as the sample size (n) tends to infinity.

19. State Asymptotic Normality of MLE's.

Solution:

Any consistent solution of the likelihood equation is asymptotically normally distributed about the true value θ_0 . Thus $\hat{\theta}$ is asymptotically $N\left(\theta_0, \dfrac{I}{I(\theta_0)}\right)$, as $n \to \theta$.

20. State Invariance property of MLE.

Solution:

If T is the MLE of θ and $\Psi(\theta)$ is a one to one function of θ, then $\Psi(T)$ is the MLE of $\Psi(\theta)$.

21. If 'X' is a binomial variate with parameters n and p , then show that $\dfrac{X}{n}$, the observed proportion of successes is an unbiased estimator of the parameter p .

Solution:

Given 'X' is a binomial variate with parameters n and p , then the mean of the binomial distribution is np

$$i.e., \; E(X) = np$$

To prove: $E\left(\dfrac{X}{n}\right) = p$

Let $E\left(\dfrac{X}{n}\right) = \dfrac{1}{n}E(X) = \dfrac{1}{n}np = p$ a

$\therefore \dfrac{X}{n}$ is an unbiased estimator of the parameter p .

22. If T is an unbiased estimator for θ, show that T^2 is a biased estimator for μ^2

Solution:

Since T is an unbiased estimator for θ, we have $E(T) = \theta$

Also $Var(T) = E(T^2) - [E(T)]^2 = E(T^2) - \theta^2$

$$\Rightarrow E(T^2) = \theta^2 + Var(T), \qquad (VarT > 0)$$

Since $E(T^2) \neq \theta^2$

Therefore, T^2 is an unbiased estimator of θ^2

23. The time to failure of an electronic component follows an exponential distribution with parameter λ . Eight units are randomly selected and tested resulting in the following failure time (in hours):

13.03, 6.07, 68.44, 17.11, 32.54, 8.77, 12.14, 23.42

Find the moment estimate of λ

Solution:

Here $\bar{x} = \frac{1}{8}(13.03 + 6.07 + 68.44 + 77.11 + 32.54 + 8.77 + 12.14 + 23.42)$

$= \frac{1}{8}(181.52) = 22.69$

Therefore the moment estimate of λ is

$$\lambda = \frac{1}{\bar{x}} = \frac{1}{22.69} = 0.04407$$

24. A random sample of size 100 has a standard deviation of 5. What can you say about the maximum error with 95% confidence?

Solution:

Given: $n = 100$; $\sigma = 5$

The maximum error (E) $= z_{\frac{\alpha}{2}} \frac{\sigma}{\sqrt{n}}$, *where* $z_{\frac{\alpha}{2}}$ *at* $95\% = 1.96$

Therefore the maximum error (E) $= (1.96)\frac{5}{\sqrt{100}} = (1.96)\frac{1}{2} = 0.98$

25. Assuming that $\sigma = 20.0$, how large a random sample be taken to assert with probability 0.95 that the sample mean will not differ from the true mean by more than 3.0 points?

Solution:

Given: $\sigma = 20,\ z_{\alpha/2} = 1.96,\ E = 3,$

To find n:

We know that,

$$E = z_{\frac{\alpha}{2}} \frac{\sigma}{\sqrt{n}}$$

$$\Rightarrow \quad 3 = (1.96)\,\frac{20}{\sqrt{n}}$$

$$\Rightarrow \quad \sqrt{n} = \frac{(1.96)20}{3} = 13.067$$

$$\therefore \quad n = 170.729 \approx 171$$

26. In 16 test runs the gasoline consumption of an experimental engine had a standard deviation of 2.2 gallons. Construct a 99% confidence interval for σ^2, which measures the true variability of the gasoline consumption of the engine.

Solution:

Given: $n = 16,\ s = 2.2,\ \chi^2_{0.005,15} = 32.801\ \ and\ \ \chi^2_{0.995,15} = 4.601$

$$\frac{(n-1)s^2}{\chi^2_{\alpha/2,\,n-1}} < \sigma^2 < \frac{(n-1)s^2}{\chi^2_{1-\alpha/2,\,n-1}}$$

$$\frac{15(2.2)^2}{32.801} < \sigma^2 < \frac{15(2.2)^2}{4.601} \Rightarrow 2.21 < \sigma^2 < 15.78$$

27. A study has been made to compare the nicotine contents of two brands of cigarettes. Ten cigarettes of Brand A and an average nicotine content of 3.1 milligrams with a standard deviation of 0.5 milligram, while eight cigarettes of brands B had an average nicotine content of 2.7 milligrams with a standard deviation of 0.7 milligrams. Assuming that the two sets of data are independent random samples from normal populations with equal variances, construct a 95% confidence interval for the difference between the mean nicotine contents of the two brans of cigarettes. Find a 98% confidence interval for $\frac{\sigma_1^2}{\sigma_2^2}$

Solution:

Given: $n_1 = 10,\ s_1 = 12,\ n_2 = 8,\ s_2 = 14$

$f_{0.01,9,7} = 6.72,\ \ f_{0.01,7,9} = 5.61$

$$\frac{s_1^2}{s_2^2}\frac{1}{f_{\alpha/2}n_1-1,n_1-1} < \frac{\sigma_1^2}{\sigma_2^2} < \frac{s_1^2}{s_2^2} f_{\alpha/2}n_1-1,n_1-1$$

$$\Rightarrow \frac{0.25}{0.49}\frac{1}{6.72} < \frac{\sigma_1^2}{\sigma_2^2} < \frac{0.25}{0.49}5.61$$

$$\Rightarrow 0.076 < \frac{\sigma_1^2}{\sigma_2^2} < 2.862$$

28. Data pertaining to height of 5 school students are given as 149, 150, 151, 138, 148 cms. Obtain a point estimate for the mean μ.

Solution: **[AU A/M 2024]**

$$\hat{\mu} = \bar{x} = \frac{149+150+151+138+148}{5} = 147.2$$

29. A random sample of size $n = 100$ is taken from a population with $\sigma = 5.1$. Given that the sample mean is $\bar{x} = 21.6$ construct a 95% confidence interval for the population mean μ.

Solution: **[AU A/M 2024]**

Confidence interval :

$$\left(\bar{x} - z_{\alpha/2}\frac{\sigma}{\sqrt{n}} < \mu < \bar{x} + z_{\alpha/2}\frac{\sigma}{\sqrt{n}}\right)$$

$$\left(21.6-1.96\frac{5.1}{\sqrt{100}}<\mu<21.6+1.96\frac{5.1}{\sqrt{100}}\right)$$
$$(20.6<\mu<22.6)$$

30. If t_1 and t_2 are both most efficient estimators with equal variance V and if t_3 is the average of t_1 and t_2. Prove that $Var(t_2)=\frac{1}{2}V(1+\rho)$ where ρ is the coefficient of correlation between t_1 and t_2.

Solution: **[AU N/D 2024]**

$$Let \quad t_3=\frac{1}{2}(t_1+t_2)$$
$$Var(at_1+bt_2)=a^2V(t_1)+b^2V(t_2)+2abCov(t_1,t_2)$$
$$\therefore \quad V(t_3)=Var\left(\frac{1}{2}t_1+\frac{1}{2}t_2\right)$$
$$=\frac{1}{4}V+\frac{1}{4}V+2.\frac{1}{2}.\frac{1}{2}.Cov(t_1,t_2)$$
$$=\frac{1}{2}V+\frac{1}{2}\rho\sqrt{V(t_1).V(t_2)}$$
$$=\frac{1}{2}V(1+\rho)$$

PART-B

1. **Show that the sample mean $\bar{x}$ is an unbiased estimator for the population mean μ.**

Solution:

To Prove that $E(\bar{x}) = \mu$:

If $x_1, x_2, ..., x_n$ be a random sample of size n

$$\bar{x} = \frac{\sum x_i}{n}$$

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

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

$$= \frac{1}{n}\left[\mu + \mu + ... + \mu\right] \qquad \left[\because\ E(x_i) = \mu\right]$$

$$= \frac{1}{n}\left[n\mu\right]$$

$$= \mu$$

$\Rightarrow \bar{x}$ is an unbiased estimator of μ

2. **If $x_1, x_2, ..., x_n$ is a random sample from a normal population $N(\mu, 1)$, show that $t = \frac{1}{n}\sum_{i=1}^{n} x_i^2$ is an unbiased estimator for $\mu^2 + 1$.**

Solution:

Given that $E(x_i) = \mu$, $Var(x_i) = 1 \ \forall\ i = 1,2,3,....,n$ ....(1)

Now $E(x_i^2) = Var(x_i) + \{E(x_i)\}^2 = 1 + \mu^2$ *from* (1)

$$E(t) = E\left(\frac{1}{n}\sum_{i=1}^{n} x_i^2\right) = \frac{1}{n}\sum_{i=1}^{n} x_i^2 = \frac{1}{n}\sum_{i=1}^{n}\left(1 + \mu^2\right) = 1 + \mu^2$$

Hence is an unbiased estimator of $1 + \mu^2$.

3. **Prove that for a random sample $(x_1, x_2, ..., x_n)$ of size n drawn from a given large population (μ, σ^2), $s^2 = \frac{1}{n}\sum_{i=1}^{n}\left(x_i - \bar{x}\right)^2$ is not an unbiased estimator of the parameter σ^2, but $\frac{ns^2}{n-1} = \frac{1}{n-1}\sum_{i=1}^{n}\left(x_i - \bar{x}\right)^2$ is an unbiased estimator of σ^2.**

Solution:

Here, we have assumed that the sample has been drawn from a normal population.

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

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

$$= \sum_{i=1}^{n}(x_i-\mu)^2 - 2(\bar{x}-\mu)\sum_{i=1}^{n}(x_i-\mu) + (\bar{x}-\mu)^2\sum_{i=1}^{n}1 \quad \text{.............(1)}$$

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

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

$$\left[\because \sum_{i=1}^{n}(x_i-\mu) = n(\bar{x}-\mu) \text{ and } \sum_{i=1}^{n}1 = n\right]$$

$$i.e., \quad \sum_{i=1}^{n}(x_i-\bar{x})^2 = \sum_{i=1}^{n}(x_i-\mu)^2 - n(\bar{x}-\mu)^2 \quad \text{.............(2)}$$

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

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

$$\Rightarrow \quad E(s^2) = \frac{1}{n}\left[\sum_{i=1}^{n}E(x_i-\mu)^2 - nE(\bar{x}-\mu)^2\right] \quad \text{................(3)}$$

We know that, $E(x_i-\mu)^2 = \sigma^2$, $E(\bar{x}-\mu)^2 = \frac{\sigma^2}{n}$

$$(3) \Rightarrow \quad E(s^2) = \frac{1}{n}\left[\sum_{i=1}^{n}\sigma^2 - n\frac{\sigma^2}{n}\right] = \frac{1}{n}\left[\sigma^2\sum_{i=1}^{n}1 - \sigma^2\right]$$

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

$$= \frac{n-1}{n}\sigma^2 < \sigma^2 \quad \text{...........(4)}$$

$\therefore \quad s^2 = \frac{1}{n}\sum_{i=1}^{n}(x_i-\bar{x})^2$ is not an unbiased estimator of σ^2

Let us consider $E\left[\frac{n}{n-1}s^2\right] = \frac{n}{n-1}E[s^2]$

$$= \left(\frac{n}{n-1}\right)\left(\frac{n-1}{n}\right)\sigma^2 \quad by\ (4)$$

$$= \sigma^2$$

$\therefore \quad \frac{n}{n-1}s^2$ is an unbiased estimator of σ^2 .

4. Show that $s^2 = \frac{1}{n-1}\sum_{i=1}^{n}\left(x_i - \bar{x}\right)^2$ **is not an unbiased estimator of the parameter** σ^2 **.**

Solution:

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

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

$$= \sum_{i=1}^{n}(x_i - \mu)^2 - 2(\bar{x} - \mu)\sum_{i=1}^{n}(x_i - \mu) + (\bar{x} - \mu)^2\sum_{i=1}^{n}1 \quad \text{.............(1)}$$

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

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

$$\left[\because \sum_{i=1}^{n}(x_i - \mu) = n(\bar{x} - \mu) \text{ and } \sum_{i=1}^{n}1 = n\right]$$

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

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

$$\Rightarrow \quad E(s^2) = \frac{1}{n-1}\left[\sum_{i=1}^{n}E(x_i - \mu)^2 - nE(\bar{x} - \mu)^2\right] \quad \text{................(3)}$$

We know that, $E(x_i - \mu)^2 = \sigma^2$, $E(\bar{x} - \mu)^2 = \frac{\sigma^2}{n}$

$$(3) \Rightarrow \quad E(s^2) = \frac{1}{n-1}\left[\sum_{i=1}^{n}\sigma^2 - n\frac{\sigma^2}{n}\right] = \frac{1}{n-1}\left[\sigma^2\sum_{i=1}^{n}1 - \sigma^2\right]$$

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

$$= \frac{n-1}{n-1}\sigma^2 = \sigma^2 \quad \text{...........(4)}$$

Hence, $s^2 = \frac{1}{n-1}\sum_{i=1}^{n}\left(x_i - \bar{x}\right)^2$ is an unbiased estimator of the parameter σ^2

5. Below you are given the values obtained from a random of observations taken from an infinite population

32 34 35 39

(a) Find a point estimator for μ. Is this an unbiased estimate of μ ? Explain

(b) Find a point estimator for σ^2. Is this an unbiased estimate of σ^2

(c) Find a point estimator for σ^2

(d) What can be said about the sampling distribution of $\overline{x}$? Be sure to discuss the expected value, the standard deviation, and the shape of the sampling distribution of $\overline{x}$?

Solution:

(a) Point estimator of $\mu = \dfrac{32+34+35+39}{4} = 35$

It is an unbiased estimator of μ .

(b) Point estimator of $\sigma^2 = \dfrac{\sum\left(x-\overline{x}\right)^2}{n-1}$

$$= \frac{(32-35)^2+(34-35)^2+(35-35)^2+(39-35)^2}{4-1}$$

$$= \frac{9+1+0+16}{3} = 8.6667$$

It is an unbiased estimator of σ^2 .

© Point estimator of $\sigma = \sqrt{\sigma^2} = \sqrt{8.6667} = 2.9439$

$$Z = \frac{\overline{x}-\mu}{\frac{\sigma}{\sqrt{n}}} \qquad \because \mu = \overline{x} \pm Z\frac{\sigma}{\sqrt{n}}$$

(d) The sampling distribution has a mean equal to the population mean.

The sampling distribution has a standard deviation equal to the population standard deviation of divide by the square root of the sample size.

The shape of the sampling distribution of $\overline{x}$ will be a normal curve.

6. A random sample $(X_1, X_2, X_3, X_4, X_5)$ of size 5 is drawn from a normal population with unknown mean μ. Consider the following estimators to estimate μ :

$$(i)\ t_1 = \frac{X_1+X_2+X_3+X_4+X_5}{5},\ (ii)\ t_2 = \frac{X_1+X_2}{2}+X_3\ \ (iii)\ t_3 = \frac{2X_1+X_2+\lambda X_3}{3}$$

where λ is such that t_3 is an unbiased estimator of μ.

Find λ. Are t_1 and t_2 unbiased? State giving reasons, the estimator which is best among t_1, t_2 and t_3.

Solution:

We have $E(X_i) = \mu, Var(X_i) = \sigma^2$; $Cov(X_i, X_j) = 0$, $(i \neq j = 1,2,....,n)$ ....(1)

$$(i)\ E(t_1) = E\left[\frac{X_1 + X_2 + X_3 + X_4 + X_5}{5}\right]$$

$$= \frac{1}{5}\left[E(X_1) + E(X_2) + E(X_3) + E(X_4) + E(X_5)\right]$$

$$= \frac{1}{5}\left[\mu + \mu + \mu + \mu + \mu\right] = \frac{5\mu}{5} = \mu$$

$\Rightarrow$ t_1 is an unbiased estimator of μ

$$(ii)\ E(t_2) = E\left[\frac{X_1 + X_2}{2} + X_3\right]$$

$$= \frac{1}{2}\left[E(X_1) + E(X_2)\right] + E(X_3)$$

$$= \frac{1}{2}\left[\mu + \mu\right] + \mu = 2\mu$$

$\Rightarrow$ t_2 is not an unbiased estimator of μ

(iii) *Given* $E(t_3) = \mu$ $\quad [\because\ t_3$ *is unbiased estimator of* $\mu\ ]$

$$\Rightarrow\quad E\left[\frac{2X_1 + X_2 + \lambda X_3}{3}\right] = \mu$$

$$\Rightarrow\quad \frac{1}{3}E\left[2X_1 + X_2 + \lambda X_3\right] = \mu$$

$$\Rightarrow\quad \frac{1}{3}\left[2E(X_1) + E(X_2) + \lambda E(X_3)\right] = \mu$$

$$\Rightarrow\quad \frac{1}{3}\left[2\mu + \mu + \lambda\mu\right] = \mu$$

$$\Rightarrow\quad \left[3\mu + \lambda\mu\right] = 3\mu$$

$$\Rightarrow\quad \lambda = 0$$

Using (1), we get

$$Var(t_1) = Var\left[\frac{X_1 + X_2 + X_3 + X_4 + X_5}{5}\right]$$

$$= \frac{1}{25}\left[Var(X_1) + Var(X_2) + Var(X_3) + Var(X_4) + Var(X_5)\right] = \frac{1}{25}\left(5\sigma^2\right)$$

$$= \frac{\sigma^2}{5}$$

$$Var(t_2) = Var\left[\frac{X_1 + X_2}{2} + X_3\right]$$

$$= \frac{1}{4}\left[Var(X_1) + Var(X_2)\right] + Var(X_3)$$

$$= \frac{1}{4}\left[\sigma^2 + \sigma^2\right] + \sigma^2 = \frac{3}{2}\sigma^2$$

$$Var(t_3) = Var\left[\frac{2X_1 + X_2 + \lambda X_3}{3}\right]$$

$$= \frac{1}{9}\left[4Var(X_1) + Var(X_2)\right] \qquad (\because \lambda = 0)$$

$$= \frac{1}{9}\left[4\sigma^2 + \sigma^2\right]$$

$$= \frac{5}{9}\sigma^2$$

Since $Var(t_1)$ is least t_1 is the best estimator (in the same of least variance) of μ.

7. Let X_1, X_2 *and* X_3 is a random sample of size 3 from a normal population with mean value μ and variance σ^2. T_1, T_2 *and* T_3 are the estimators used to estimate mean μ, where

$$T_1 = X_1 + X_2 - X_3,\ T_2 = 2X_1 + 3X_3 - 4X_2 \ \ and \ \ T_3 = \frac{1}{3}(\lambda X_1 + X_2 + X_3)$$

(i) Are T_1 *and* T_2 unbiased estimators?

(ii) Find the value of λ such that T_3 is unbiased estimator for μ.

(iii) With this value of λ is T_3 a consistent estimator?

(iv) Which is the best estimator?

Solution:

We have $E(X_i) = \mu, Var(X_i) = \sigma^2$; $Cov(X_i, X_j) = 0$, $(i \neq j = 1,2,....,n)$ ....(1)

(i) $E(T_1) = E(X_1 + X_2 - X_3) = E(X_1) + E(X_2) - E(X_3) = \mu + \mu - \mu = \mu$

$E(T_2) = E(2X_1 + 3X_3 - 4X_2) = 2E(X_1) + 3E(X_3) - 4E(X_2) = 2\mu + 3\mu - 4\mu = \mu$

$\therefore$ T_1 *and* T_2 are unbiased estimators of μ

(ii) *Given* $E(t_3) = \mu$ $\left[\because t_3 \text{ is unbiased estimator of } \mu\right]$

$$\Rightarrow E\left[\frac{\lambda X_1 + X_2 + X_3}{3}\right] = \mu$$

$$\Rightarrow \frac{1}{3}\left[\lambda E(X_1) + E(X_2) + E(X_3)\right] = \mu$$

$$\Rightarrow \frac{1}{3}\left[\lambda\mu + \mu + \mu\right] = \mu$$

$$\Rightarrow \left[\lambda\mu + 2\mu\right] = 3\mu$$

$$\Rightarrow \lambda = 1$$

(iii) With $\lambda = 1,\ T_3 = \frac{1}{3}(X_1 + X_2 + X_3) = \overline{X}$

Since sample mean is a consistent estimator of population mean μ , by weak law of large numbers, T_3 is a consistent estimator of μ .

(iv) We have [on using (1)]

$$Var(T_1) = Var(X_1) + Var(X_1) + Var(X_1) = 3\sigma^2$$

$$Var(T_2) = 4Var(X_1) + 9Var(X_1) + 16Var(X_1) = 29\sigma^2$$

$$Var(T_3) = \frac{1}{9}\left[Var(X_1) + Var(X_1) + Var(X_1)\right] = \frac{1}{3}\sigma^2$$

Since $Var(T_3)$ is minimum, T_3 is the best estimator of μ in the sense of minimum variance.

8. If $x_1, x_2, \ldots, x_n$ are random observations of a Bernoulli's variate x which assumes values 1 and 0 with probabilities $p\ and\ (1-p)$ respectively. Show that $\frac{T(n-T)}{n(n-1)}$ is an unbiased estimator of $p(1-p)$, where $T = x_1 + x_2 + \ldots + x_n$

Solution: **[AU N/D 2024]**

$P(x=1) = p\ \ and\ \ P(x=0) = 1-p = q$

$\therefore\ \ T = (x_1 + x_2 + \ldots + x_n)$ follows a binomial distribution with mean np and variance npq .

$\therefore\ \ E(T) = np\ \ and\ \ Var(T) = E(T^2) - [E(T)]^2 = npq$

$E(T^2) = npq + n^2p^2$

$$\text{Now } E\left[\frac{T(n-T)}{n(n-1)}\right] = \frac{1}{n(n-1)}\left[nE(T) - E(T^2)\right]$$

$$= \frac{1}{n(n-1)}\left[n^2p - npq - n^2p^2\right]$$

$$= \frac{1}{n(n-1)}\left[n^2pq - npq\right]$$

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

$$= pq$$

$$= p(1-p)$$

Hence the result follows.

9. If t_1 is a most efficient estimator and t_2 is an unbiased estimator (of some population parameter) with efficiency e, and if the correlation coefficient between t_1 and t_2 is ρ. Show that $\rho = \sqrt{e}$.

Solution:

Let V_1 and V_2 be the variance of t_1 and t_1 respectively. Then

$$E = \frac{V_1}{V_2} \Rightarrow V_2 = \frac{V_1}{E} \quad ....(1)$$

Let $t_3 = pt_1 + qt_2$, where $p+q=1$ and let V_3 be the variance of t_3.

Then $V_3 = Var(pt_1 + qt_2)$

$$= p^2V_1 + q^2V_2 + 2pqCov(t_1,t_2)$$

$$= p^2V_1 + q^2\frac{V_1}{E} + 2pq.\rho\sqrt{V_1.\frac{V_1}{E}} \qquad \sin ce \ \ \rho = \frac{Cov(t_1,t_2)}{\sqrt{V_1.V_2}} \ and \ (1)$$

$$V_3 = \left(p^2 + \frac{q^2}{E} + \frac{2pq\rho}{\sqrt{E}}\right)V_1 \quad ....(2)$$

Since V_3 is a less efficient estimator, $V_3 \geq V_1$

$$i.e., \left(p^2 + \frac{q^2}{E} + \frac{2pq\rho}{\sqrt{E}}\right)V_1 \geq (p+q)^2V_1 \qquad \sin ce \ p+q=1 \ and \ \ from \ (2)$$

$$\Rightarrow \quad p^2 + \frac{q^2}{E} + \frac{2pq\rho}{\sqrt{E}} \geq \left(p^2 + q^2 + 2pq\right)$$

$$\Rightarrow \quad q^2\left(\frac{1}{E} - 1\right) + 2pq\left(\frac{\rho}{\sqrt{E}} - 1\right) \geq 0 \quad ....(3)$$

In (3), the equality holds good, when $q = 0 \ and \ \ p = 1$

$$(3) \quad \Rightarrow \quad 2\left(\frac{\rho}{\sqrt{E}} - 1\right) = 0$$

$$\Rightarrow \quad \frac{\rho}{\sqrt{E}} - 1 = 0$$

$$\Rightarrow \quad \rho = \sqrt{E}$$

10. Find the estimator of θ in the population with density function $f(x,\theta) = \theta x^{\theta-1}; 0 < x < 1; \theta > 0$, by method of moments.

Solution:

The first order moment (about the origin) of the population is given by

$$\mu_1^{!} = \int_0^1 x.\theta x^{\theta-1}dx$$

$$= \theta \int_0^1 x.\, x^{\theta-1} dx$$

$$= \theta \int_0^1 x^{\theta} dx$$

$$= \theta \left(\frac{x^{\theta+1}}{\theta+1} \right)_0^1 = \frac{\theta}{\theta+1}$$

The first order moment of the sample $(x_1, x_2, ..., x_n)$ about the origin is given by

$$m_1^! = \frac{1}{n} \sum_{i=1}^{n} x_i = \bar{x}$$

By the method of moments,

$$\bar{x} = \frac{\theta}{\theta+1}$$

$$(\theta+1)\bar{x} = \theta$$

$$(\theta\bar{x} + \bar{x}) = \theta$$

$$\bar{x} = \theta - \theta\bar{x}$$

$$i.e., \quad \theta = \frac{\bar{x}}{1-\bar{x}}$$

11. If $(x_1, x_2, ..., x_n)$ a random sample from the uniform population with the density function $f(x, a, b) = \frac{1}{b-a}; a < x < b$. Find the estimators of a and b by method of moments.

Solution:

The first order moment (about the origin) of the population is given by

$$\mu_1^! = \int_a^b x.\frac{1}{b-a} dx = \frac{1}{b-a} \int_a^b x.dx = \frac{1}{b-a} \left(\frac{x^2}{2} \right)_a^b$$

$$= \frac{1}{b-a} \left(\frac{b^2 - a^2}{2} \right) = \frac{a+b}{2}$$

$$\mu_2^! = \int_a^b x^2.\frac{1}{b-a} dx = \frac{1}{b-a} \int_a^b x^2 \, dx = \frac{1}{b-a} \left(\frac{x^3}{3} \right)_a^b$$

$$= \frac{1}{b-a} \left(\frac{b^3 - a^3}{3} \right) = \frac{a^2 + ab + b^2}{3}$$

If $m_1^! = \frac{1}{n} \sum_{i=1}^{n} x_i = \bar{x}$ and $m_2^! = \frac{1}{n} \sum_{i=1}^{n} x_i^2 = s^2$ are the first and second order moments of the sample about the origin, then the method of moments gives

$$a+b=2\bar{x} \quad \Rightarrow \quad b=2\bar{x}-a \qquad \text{................(1)}$$

And $$a^2+ab+b^2=3s^2 \qquad \text{................(2)}$$

$$(2) \Rightarrow \quad a^2+a\left(2\bar{x}-a\right)+\left(2\bar{x}-a\right)^2=3s^2 \qquad by \ (1)$$

$$i.e., \quad a^2-2\bar{x}a+\left(4\bar{x}^2-3s^2\right)=0$$

$$\therefore \quad a=\frac{2\bar{x}\pm\sqrt{4\bar{x}^2-4\left(4\bar{x}^2-3s^2\right)}}{2}$$

$$a=\bar{x}\pm\sqrt{3\left(s^2-\bar{x}^2\right)}$$

Again from (1) $$b=\bar{x}\mp\sqrt{3\left(s^2-\bar{x}^2\right)}$$

Since $a<b$, we have $\bar{a}=\bar{x}-\sqrt{3\left(s^2-\bar{x}^2\right)}$ and $\bar{b}=\bar{x}+\sqrt{3\left(s^2-\bar{x}^2\right)}$

12. For the probability mass function $f(x,p)=3C_x\dfrac{p^x(1-p)^{3-x}}{1-(1-p)^3}$; $x=1,2,3$. Obtain the estimators of p by method of moments, if the frequencies at $x=1,2$ *and* 3 are respectively 22, 20 and 18.

Solution:

$$f(x,p)=3C_x\frac{p^x(1-p)^{3-x}}{1-(1-p)^3} \ ; \ x=1,2,3$$

$$f(x,p)=\frac{1}{1-(1-p)^3}B(3;p)$$

Therefore, the first order moment about the origin, viz., the mean of the given distribution is given by

$$\mu_1^{!}=\frac{1}{1-(1-p)^3}.3p$$

The men of the observed sample is given by

$$\bar{x}=\frac{(1)(22)+(2)(20)+(3)(18)}{22+20+18}=\frac{116}{60}=\frac{29}{15}$$

By the method of moments,

$$\mu_1^{!}=\bar{x}$$

$$i.e., \quad \frac{3p}{3p-3p^2+p^3}=\frac{29}{15}$$

$$i.e., \quad 29p^2-87p+42=0$$

$$\Rightarrow \qquad p=\frac{87\pm 51.93}{58}=2.395 \ (or) \ 0.605$$

Since2.395 is inadmissible, $\bar{p}=0.605$

13. A random variable X takes the values 0, 1, 2 with probabilities $\frac{1}{2}-\theta, \frac{\alpha}{2}+2(1-\alpha)\theta \ and \ \left(\frac{(1-\alpha)}{2}\right)+(2\alpha-1)\theta$, **where** $\alpha \ and \ \theta$ **are the parameters. If a sample of size 75 drawn from the population yielded the values 0, 1, 2 with respective frequencies 27, 38, 10 respectively, find the estimators of** $\alpha \ and \ \theta$ **by the method of moments.**

Solution:

$$\mu_1^{!} = E(X) = (0)\left(\frac{1}{2}-\theta\right)+(1)\left\{\frac{\alpha}{2}+2(1-\alpha)\theta\right\}+(2)\left\{\frac{1-\alpha}{2}+(2\alpha-1)\theta\right\}$$

$$=1-\frac{\alpha}{2}+2\alpha\theta$$

$$\mu_2^{!} = E(X^2) = (0^2)\left(\frac{1}{2}-\theta\right)+(1^2)\left\{\frac{\alpha}{2}+2(1-\alpha)\theta\right\}+(2^2)\left\{\frac{1-\alpha}{2}+(2\alpha-1)\theta\right\}$$

$$=2-\frac{3}{2}\alpha+(6\alpha-2)\theta$$

The mean of the observed sample

$$m_1^{!} = \frac{(38)(1)+(10)(2)}{75} = \frac{58}{75}$$

The second order moment about the origin is given by

$$m_2^{!} = s^2 = \frac{(38)(1^2)+(10)(2^2)}{75} = \frac{78}{75}$$

By the method of moments, $\mu_1^{!} = \overline{x} \ and \ \mu_2^{!} = s^2$

$$1-\frac{\alpha}{2}+2\alpha\theta = \frac{58}{75} \quad \text{.....................(1)}$$

$$2-\frac{3}{2}\alpha+(6\alpha-2)\theta = \frac{78}{75} \quad \text{.................(2)}$$

Solving the equations (1) and (2), we get

$$\overline{\alpha} = \frac{34}{33} \ and \ \overline{\theta} = \frac{7}{50}$$

14. Consider a characteristic that occurs in proportion p **of a population. Let** $X_1, X_2, ..., X_n$ **be a random sample of size** n **so** $P[X_i = 0] = 1-p \ and \ P[X_i = 1] = p \ for \ i = 1,2,...,n$ **where** $0 \le p \le 1$ **.Obtain the maximum likelihood estimator of** p **.**

(or)

Find the maximum likelihood estimator for the parameter p **of the binomial distribution** $B(N, P)$ **where** N **is very large but finite, on the basis of sample of size** n **, Also find its inverse.**

Solution:

The probability mass function of the binomial distribution is

$$P[X=x]=p[x;N,P]=NC_x P^x (1-P)^{N-x} \;;\; x=0,1,2,.......,N$$

Therefore, the likelihood function of the random sample $(x_1,x_2,......,x_n)$ is given by

$$L(x_1,x_2,......,x_n,P)=\prod_{i=1}^{n} nC_{x_i} P^{\sum x_i}(1-P)^{n^2-\sum x_i}$$

$$=\prod_{i=1}^{n} nC_{x_i} P^{n\bar{x}}(1-P)^{n^2-n\bar{x}}$$

$$\log L=\sum_{i=1}^{n}\log\left(nC_{x_i}\right)+n\bar{x}\log P+n\left(n-\bar{x}\right)\log(1-P)$$

The likelihood equation is

$$\frac{\partial}{\partial P}\log L=0$$

$$i.e.,\quad \frac{n\bar{x}}{P}-\frac{n\left(n-\bar{x}\right)}{1-P}=0$$

$$i.e.,\quad (1-P)\bar{x}-P\left(n-\bar{x}\right)=0$$

$$\Rightarrow\quad \bar{x}-P\bar{x}-Pn+P\bar{x}=0$$

$$\bar{x}-Pn=0\quad\Rightarrow\quad P=\frac{\bar{x}}{n}$$

By Rao-Cramer's formula,

$$\left\{Var(\bar{P})\right\}^{-1}=-E\left\{\frac{\partial^2}{\partial P^2}\log L\right\}$$

$$=-E\left[\frac{\partial}{\partial P}\left\{\frac{n\bar{x}}{P}-\frac{n\left(n-\bar{x}\right)}{1-P}\right\}\right]$$

$$=-E\left[-\frac{n\bar{x}}{P^2}-\frac{n\left(n-\bar{x}\right)}{(1-P)^2}\right]$$

$$=n\left[\frac{1}{P^2}E(\bar{x})+\frac{1}{(1-P)^2}\left[\left(n-E\left(\bar{x}\right)\right)\right]\right]$$

$$=n\left[\frac{1}{P^2}nP+\frac{1}{(1-P)^2}(n-nP)\right]$$

$$=n^2\left[\frac{1}{P}+\frac{1}{(1-P)}\right]$$

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

$$\left\{Var(\bar{P})\right\}=\frac{PQ}{n^2}$$

15. Let $X_1, X_2, ..., X_n$ be a random sample of size n from the Poisson distribution $f(x/\lambda) = \frac{e^{-\lambda}\lambda^x}{x!}$ where $0 \le \lambda < \infty$. Obtain the maximum likelihood estimator of λ.

Solution:

The p.d.f is $P[X = x] = \frac{e^{-\lambda}\lambda^x}{x!}$

$$L(\lambda : x_1, x_2, ..., x_n) = \prod_{i=1}^{n} \frac{e^{-\lambda}\lambda^{x_i}}{x_i!}$$

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

$$= \sum_{i=1}^{n} \left[\log\left(e^{-\lambda}\lambda^{x_i}\right) - \log(x_i!)\right]$$

$$= \sum_{i=1}^{n} \left[\log\left(e^{-\lambda}\right) + \log\left(\lambda^{x_i}\right) - \log(x_i!)\right]$$

$$= \sum_{i=1}^{n} \left[x_i \log(\lambda) - \lambda - \log(x_i!)\right]$$

$$= \log(\lambda)\sum_{i=1}^{n} x_i - n\lambda - \sum_{i=1}^{n} \log(x_i!)$$

$$\frac{d}{d\lambda}\log L = 0$$

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

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

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

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

16. For random sampling from a normal population $N(\mu, \sigma^2)$, find the maximum likelihood estimators for

(i) μ, when σ^2 is known.

(ii) σ^2, when μ is known and

(iii) The simultaneous estimation of μ and σ^2. [AU A/M 2024]

Solution:

For the normal distribution $N(\mu, \sigma^2)$, the likelihood function is given by

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

$$= \prod_{i=1}^{n} \frac{1}{\sigma\sqrt{2\pi}} \exp\left\{-\frac{1}{2\sigma^2}(x-\mu)^2\right\}$$

$$= \left(\frac{1}{\sigma\sqrt{2\pi}}\right)^n \exp\left\{-\frac{1}{2\sigma^2}\sum_{i=1}^{n}(x-\mu)^2\right\}$$

$$\therefore \quad \log L = -\frac{n}{2}\log(\sigma^2) - \frac{n}{2}\log(2\pi) - \frac{1}{2\sigma^2}\sum_{i=1}^{n}(x-\mu)^2 \quad ....(1)$$

$$(i) \quad \frac{\partial}{\partial\mu}(\log L) = \frac{1}{\sigma^2}\sum_{i=1}^{n}(x-\mu) = \frac{1}{\sigma^2}(n\overline{x} - n\mu)$$

The MLE of μ is given by

$$\frac{\partial}{\partial\mu}(\log L) = 0$$

$$i.e., \quad \overline{x} - \mu = 0$$

$$\therefore \quad \overline{\mu} = MLE \ of \ \mu = \overline{x}, \ \text{since it is seen that}$$

$$\frac{\partial^2}{\partial\mu^2}(\log L) = \frac{-n}{\sigma^2} < 0$$

(ii) *Differentiating both sides of* (i) *w.r.t* σ^2, *we get*

$$\frac{\partial}{\partial\sigma^2}(\log L) = \frac{n}{2\sigma^2} + \frac{1}{2\sigma^4}\sum_{i=1}^{n}(x_i-\mu)^2$$

The MLE of σ^2 is given by

$$\frac{\partial}{\partial\sigma^2}(\log L) = 0$$

$$\therefore \quad \overline{\sigma}^2 = \frac{1}{n}\sum_{i=1}^{n}(x_i-\mu)^2$$

(iii) The likelihood equations for the simultaneous estimation of μ and σ^2 are

$$\frac{\partial}{\partial\mu}(\log L) = 0$$

and
$$\frac{\partial}{\partial\sigma^2}(\log L) = 0$$

$$i.e., \quad \overline{\mu} = \overline{x} \quad ....(2)$$

$$and \quad \overline{\sigma}^2 = \frac{1}{n}\sum_{i=1}^{n}(x_i-\mu)^2 \quad ....(3)$$

Using (2) in (3), we get

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

17. Obtain the maximum likelihood estimators a and b in terms of the sample observations $x_1, x_2, \ldots\ldots, x_n$ taken from the exponential population with density function

$f(x,a,b) = ke^{-b(x-a)}$; $x \geq a,\ b > 0$

Solution: **[AU N/D 2024]**

To find the unknown constant k, we have

$$\int_a^{\infty} f(x,a,b)dx = 1$$

$$i.e.,\quad k\int_a^{\infty} e^{-b(x-a)}dx = 1$$

$$i.e.,\quad ke^{ab}\left(\frac{e^{-bx}}{-b}\right)_a^{\infty} = 1$$

$$i.e.,\quad \frac{k}{b} = 1 \ \Rightarrow\ k = b$$

$$\therefore\quad f(x,a,b) = be^{-b(x-a)}\ ;\ x \geq a,\ b > 0$$

$$L(x_1, x_2, \ldots\ldots, x_n) = b^n e^{-b\sum_{i=1}^{n}(x_i - a)}$$

$$= b^n e^{-b(n\bar{x} - na)} \quad \ldots\ldots\ldots\ldots\ldots\ldots(1)$$

The likelihood equations for the simultaneous estimation of a and b are

$$\frac{\partial}{\partial a}(\log L) = 0 \quad \ldots\ldots\ldots\ldots\ldots(2)$$

and

$$\frac{\partial}{\partial b}(\log L) = 0 \quad \ldots\ldots\ldots\ldots(3)$$

From (1), we get

$$\log L = n\log b - nb(\bar{x} - a)$$

$\therefore$ (2) $\Rightarrow nb = 0$, *which means* $b = 0$, *which is absard as* b *is given to be positive*

$$(3)\ \Rightarrow\ \frac{n}{b} - n(\bar{x} - a) = 0$$

$$\therefore\quad b = \frac{1}{\bar{x} - a}$$

Viz., b is not unique as a is not definitely known.

Again from (1), we not that L is maximum, for a given positive value of b, when $e^{-nb(\bar{x}-a)}$ is maximum.

i.e., when $(\bar{x} - a)$ is minimum.

i.e., when $\frac{1}{n}\left[(x_1 - a) + (x_2 - a) + \ldots\ldots + (x_n - a)\right]$ is minimum.

i.e., when $a =$ *smallest among* $x_1, x_2, \ldots\ldots, x_n$, *say* x_S, For this value of

$$\bar{a} = x_S \;;\quad \bar{b} = \frac{1}{\bar{x} - x_S}$$

18. Find the maximum likelihood estimators θ in population with density function $f(x,\theta) = (1+\theta)x^{\theta};\ 0 \le x \le 1\ ;\ \theta > 0;$ based on random sample of size n. Test whether this estimator is a sufficient estimator of θ.

Solution:

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

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

$$\therefore\ \log L = n\log(1+\theta) + \theta \sum_{i=1}^{n} \log x_i$$

The likelihood equation is

$$\frac{\partial}{\partial\theta}(\log L) = 0$$

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

$$\Rightarrow \quad n + \theta\sum_{i=1}^{n} \log x_i + \sum_{i=1}^{n} \log x_i = 0$$

Where G is the geometric mean of the sample,

$$\therefore \quad \hat{\theta} = \frac{-n}{\sum_{i=1}^{n} \log x_i} - 1 = \frac{-n}{\log\left(\prod_{i=1}^{n} x_i\right)} - 1$$

$$Now \quad L(x_1, x_2, \ldots\ldots, x_n; \theta) = \left\{(1+\theta)^n \left(\prod_{i=1}^{n} x_i\right)^{\theta-1}\right\}\left\{\prod_{i=1}^{n} x_i\right\}$$

Hence by Factorization theorem,

$T = \left\{\prod_{i=1}^{n} x_i\right\}$ *is a sufficient statistic for* θ *and* $\hat{\theta}$ *being a one to one function of sufficient statistic* $\left\{\prod_{i=1}^{n} x_i\right\}$, *is also sufficient for* θ.

19. For the double Poisson distribution:

$$p(x) = P(X = x) = \frac{1}{2}\frac{e^{-m_1} m_1^x}{x!} + \frac{1}{2}\frac{e^{-m_2} m_2^x}{x!}\ ;\ x = 0,1,2,\ldots..$$

Show that the estimates for m_1 and m_2 by the method of moments are: $\mu_1^{'} \pm \sqrt{\mu_2^{'} - \mu_1^{'} - \mu_1^{'2}}$

Solution:

$$\mu_1^! = \sum_{x=0}^{\infty} x.p(x) = \frac{1}{2}\sum_{x=0}^{\infty} x.\frac{e^{-m_1}m_1^x}{x!} + \frac{1}{2}\sum_{x=0}^{\infty} x.\frac{e^{-m_2}m_2^x}{x!} = \frac{1}{2}m_1 + \frac{1}{2}m_2 \text{ ....(1)}$$

(since the first and second summations are the means of Poisson distributions with parameters m_1 and m_2 respectively)

$$\mu_2^! = \sum_{x=0}^{\infty} x^2.p(x) = \frac{1}{2}\sum_{x=0}^{\infty} x^2\frac{e^{-m_1}m_1^x}{x!} + \frac{1}{2}\sum_{x=0}^{\infty} x^2\frac{e^{-m_2}m_2^x}{x!}$$

$$= \frac{1}{2}\left\{\left(m_1^2 + m_1\right) + \left(m_2^2 + m_2\right)\right\}$$

$$\Rightarrow \quad \mu_2^! = \frac{1}{2}\left\{\left(m_1 + m_2\right) + \left(m_1^2 + m_2^2\right)\right\} \text{ ....(2)}$$

$$= \frac{1}{2}\left\{2\mu_1^! + m_1^2 + \left(2\mu_1^! - m_1\right)^2\right\} \qquad u\sin g\ (1)$$

$$= \frac{1}{2}\left\{2\mu_1^! + m_1^2 + 4\mu_1^{!2} + m_1^2 - 4m_1\mu_1^!\right\}$$

$$\Rightarrow \quad \mu_2^! = \mu_1^! + m_1^2 + 2\mu_1^{!2} - 2\mu_1^! m_1$$

$$\Rightarrow \quad m_1^2 - 2m_1\mu_1^! + \left(2\mu_1^{!2} + \mu_1^! - \mu_2^!\right) = 0$$

$$\therefore \quad \hat{m}_1 = \frac{2\mu_1^! \pm \sqrt{4\mu_1^{!2} - 4\left(2\mu_1^{!2} + \mu_1^! - \mu_2^!\right)}}{2} = \mu_1^! \pm \sqrt{\mu_2^! - \mu_1^! - \mu_1^{!2}}$$

Similarly on substituting for m_1 in terms of m_2 from (1) in (2), we get

$$m_2^2 - 2m_2\mu_1^! + \left(2\mu_1^{!2} + \mu_1^! - \mu_2^!\right) = 0$$

$$Solving\ \ for\ m_2, we\ get\ \ \hat{m}_2 = \mu_1^! \pm \sqrt{\mu_2^! - \mu_1^! - \mu_1^{!2}}$$

20. Let $x_1, x_2, ..., x_n$ be a random sample from the uniform distribution with p.d.f :

$$f(x,\theta) = \begin{cases} \dfrac{1}{\theta}, & 0 < x < \infty,\ \theta > 0 \\ 0\ , & otherwise \end{cases}$$ **Obtain the maximum likelihood estimator for θ**

Solution:

Here $L = \prod_{i=1}^{n} f(x_i, \theta)$

$$= \frac{1}{\theta}.\frac{1}{\theta}.\frac{1}{\theta}..........\frac{1}{\theta}\ (n\ times)$$

$$= \left(\frac{1}{\theta}\right)^n \quad \text{.................(1)}$$

Likelihood equation, viz., $\frac{\partial}{\partial\theta}\log L = 0,\ gives$

$$\frac{\partial}{\partial\theta}(-n\log\theta)=0 \Rightarrow \frac{-n}{\theta}=0 \quad or \quad \hat{\theta}=\infty, \; obviously\ an\ absurd\ result$$

In this case we locate M.L.E as follows: We have to choose θ so that L in (1) is maximum. Now L is maximum if θ is minimum.

Let $x_{(1)}, x_{(2)}, \ldots\ldots, x_{(n)}$ be the ordered random sample of n independent observations from the given population so that $0 \le x_{(1)} \le x_{(2)} \le \ldots\ldots \le x_{(n)} \le \theta \quad \Rightarrow \theta \ge x_{(n)}$

Since the minimum value of θ consistent with the sample is $x_{(n)}$, the largest sample observation, $\hat{\theta} = x_{(n)}$.

21. Obtain the maximum likelihood estimators α *and* β for the rectangular population:

$$f(x,\alpha,\beta)=\begin{cases}\frac{1}{\beta-\alpha}, & \alpha < x < \beta \\ 0 \;, & otherwise\end{cases}$$

Solution:

Here $L=\prod_{i=1}^{n} f(x_i,\alpha,\beta)$

$$=\frac{1}{\beta-\alpha}\cdot\frac{1}{\beta-\alpha}\ldots\ldots\ldots\ldots\frac{1}{\beta-\alpha} \; (n\ times)$$

$$L=\left(\frac{1}{\beta-\alpha}\right)^n$$

$$\log L = -n\log(\beta-\alpha) \;\ldots\ldots\ldots\ldots\ldots\ldots(1)$$

Likelihood equations for α *and* β give

$$\frac{\partial}{\partial\alpha}\log L = 0 \Rightarrow \frac{n}{\beta-\alpha}=0$$

$$and \quad \frac{\partial}{\partial\beta}\log L = 0 \Rightarrow \frac{-n}{\beta-\alpha}=0$$

Each of these equations gives $\beta-\alpha=\infty$, an obviously negative result. So we find M.L.Es for $\beta-\alpha=\infty$ by some other means.

Now L in (1) is maximum if $(\beta-\alpha)$ is minimum, i.e., if β takes the minimum possible value and α takes the maximum possible value.

If $x_{(1)}, x_{(2)}, \ldots\ldots, x_{(n)}$ is an ordered random sample from this population, then

$\alpha \le x_{(1)} \le x_{(2)} \le \ldots\ldots \le x_{(n)} \le \beta$. Thus $\beta \ge x_{(n)}$ *and* $\alpha \le x_{(1)}$.

Hence the minimum possible value of β consistent with the sample is $x_{(n)}$ and the maximum possible value of α consistent with the sample is $x_{(n)}$. Hence L is maximum if

$\beta = x_{(n)}$ *and* $\alpha = x_{(1)}$

Therefore M.L.E for α *and* β are given by:

$\hat{\alpha} = x_{(1)} =$ The smallest sample observation

and $\hat{\beta} = x_{(n)} =$ The largest sample observation.

Confidence interval for the population mean for large samples

$(\sigma \text{ is known})$ is

$$\bar{x} \pm z_{\alpha/2}\frac{\sigma}{\sqrt{n}}$$

22. In order to introduce some incentive for higher balance in savings accounts, a random sample size of 64 savings accounts at a bank's branch was studied to estimate the average monthly balance in saving bank accounts. The mean and standard deviation were found to be Rs. 8,500 and Rs. 2,000, respectively.

Find (i) 90%

(ii) 95%

(iii) 99% confidence intervals for the population mean.

Solution:

Confidence limits with confidence level $(100-\alpha)\%$ for average monthly balance in savings accounts are given as:

(i) 90% confidence limits	(ii) 95% confidence limits	(iii) 99% confidence limits
$\bar{x} \pm z_{\alpha/2}\frac{\sigma}{\sqrt{n}}$	$\bar{x} \pm z_{\alpha/2}\frac{\sigma}{\sqrt{n}}$	$\bar{x} \pm z_{\alpha/2}\frac{\sigma}{\sqrt{n}}$
$8500 \pm 1.645\frac{2000}{\sqrt{64}}$	$8500 \pm 1.96\frac{2000}{\sqrt{64}}$	$8500 \pm 2.575\frac{2000}{\sqrt{64}}$
$\Rightarrow\ 8500 \pm \frac{3290}{8}$	$\Rightarrow\ 8500 \pm \frac{3920}{8}$	$\Rightarrow\ 8500 \pm \frac{5150}{8}$
$\Rightarrow\ 8500 \pm 411.25$	$\Rightarrow\ 8500 \pm 490$	$\Rightarrow\ 8500 \pm 644$

It may be noted that the interval or limits gets wider as the desired level of confidence is increased.

Confidence interval for the difference between two population means for large samples

$(\sigma \text{ is known})$

$$\left(\bar{x}_1 - \bar{x}_2\right) \pm z_{\alpha/2}\sqrt{\frac{\sigma_1^2}{n_1} + \frac{\sigma_2^2}{n_2}}$$

23. In a certain factory there are two independent processes manufacturing the same item. The average weight in a sample of 250 items produced from one process is found to be 120 O_{zs}, with a standard deviation of 12 O_{zs} . While the corresponding figures in a sample of 400 items from the other process are 124 O_{zs} and 14 O_{zs} . Find the 995 confidence limits for the difference in the average weight of items produced by the two processes respectively.

Solution:

Given: $n_1 = 250,\ \overline{x_1} = 120,\ S_1 = 12,$

$n_2 = 400,\ \overline{x_2} = 124,\ S_2 = 14$

$z_{\frac{\alpha}{2}} = 2.58$

Therefore 99% confidence limits for $(\mu_1 - \mu_2)$ are

$$\left(\overline{x_1} - \overline{x_2}\right) \pm z_{\alpha/2}\sqrt{\frac{\sigma_1^2}{n_1} + \frac{\sigma_2^2}{n_2}} = \left(\overline{x_1} - \overline{x_2}\right) \pm z_{\alpha/2}\sqrt{\frac{S_1^2}{n_1} + \frac{S_2^2}{n_2}}$$

$$= (120 - 124) \pm 2.58\sqrt{\frac{12^2}{250} + \frac{14^2}{400}}$$

$$= 4 \pm 2.58(1.0324)$$

$$= 4 \pm 2.6635$$

$$(\mu_1 - \mu_2) = (1.34, 6.66)$$

<u>Confidence interval for the difference between two population means for small samples</u>

$$\left(\overline{x_1} - \overline{x_2}\right) \pm t_{\alpha/2} S\sqrt{\frac{1}{n_1} + \frac{1}{n_2}}$$

Where $t_{\alpha/2}$ is the tabulated value of t with $n_1 + n_2 - 2$ degrees of freedom at α level of significance.

24. In a test given to two groups of students the marks obtained were as follows:

First group:	**18**	**20**	**36**	**50**	**49**	**36**	**34**	**49**	**61**
Second group:	**29**	**28**	**26**	**35**	**30**	**44**	**46**		

Construct a 95% confidence interval on the mean marks secured by students of the above two groups.

Solution:

Given: $n_1 = 9,\ \overline{x_1} = \frac{333}{9} = 37$

$n_2 = 7,\ \overline{x_2} = \frac{238}{7} = 34$

$$S^2 = \frac{1}{n_1 + n_2 - 2}\left[\sum_i \left(x_i - \overline{x_1}\right)^2 + \sum_j \left(x_j - \overline{x_2}\right)^2\right]$$

$$= \frac{1}{14}[1134 + 386] = 108.57$$

$$S = 10.42 \text{ and } t_{\alpha/2} \text{ with } n_1 + n_2 - 2 = 14, \ d.f \text{ is } 1.76$$

Therefore the 95% confidence limits for $(\mu_1 - \mu_2)$ are

$$(\overline{x_1} - \overline{x_2}) \pm t_{\alpha/2} S \sqrt{\frac{1}{n_1} + \frac{1}{n_2}} = (37 - 34) \pm (1.76)(10.42)\sqrt{\frac{1}{9} + \frac{1}{7}}$$

$$= 3 \pm (1.76)(5.25) = 3 \pm 9.24$$

$$= (6.24, 12.24)$$

25. Let X be exponentially distributed with parameter λ. Using maximum likelihood estimation, find an estimate for the parameter λ.

Solution: **[AU A/M 2024]**

Let X be an exponential distribution with parameter λ, $f(x) = \lambda e^{-\lambda x}, x > 0$.

Let $x_1, x_2, ..., x_n$ be a sample of size n.

Then the likelihood function is

$$L(x_1, x_2, ..., x_n) = \prod_{i=1}^{n} \lambda e^{-\lambda xi}$$

$$= \lambda e^{-\lambda x_1} . \lambda e^{-\lambda x 2} ... \lambda e^{-\lambda x_n}$$

$$= \lambda^n e^{-\lambda(x_1 + x_2 + ... + x_n)}$$

$$L = \lambda^n e^{-\lambda \sum_{i=1}^{n} x_i}$$

Taking logarithm,

$$\log L = \log\left(\lambda^n e^{-\lambda \sum_{i=1}^{n} x_i}\right)$$

$$= \log \lambda^n + \log e^{-\lambda \sum_{i=1}^{n} x_i}$$

$$\log L = n \log \lambda - \lambda \sum_{i=1}^{n} x_i$$

The likelihood equation is

$$\frac{d}{d\lambda} \log L = 0$$

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

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

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

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

$$\hat{\lambda} = \frac{1}{\overline{x}}$$

26. If 83 male students are randomly chosen and yield an average of 6.6 hours of sleep with a standard deviation of 1.8 and 65 females are randomly selected with an average of 6.9 hours of sleep with a standard deviation of 1.5. Construct a 95 % confidence interval for the difference between the two mean sleep hours for males and females.

Solution: **[AU A/M 2024]**

Given: $n_1 = 83,\ n_2 = 65$

$\overline{x_1} = 6.6,\ \overline{x_2} = 6.9$

$\sigma_1 = 1.8,\ \sigma_2 = 1.5$

$Z_{\alpha/2} = 1.96$

$\therefore$ 95% confidence limits for $\mu_1 - \mu_2$ are

$$\left(\overline{x_1} - \overline{x_2}\right) \pm Z_{\alpha/2}\sqrt{\frac{\sigma_1^2}{n_1} + \frac{\sigma_2^2}{n_2}}$$

$$= (6.6 - 6.9) \pm 1.96\sqrt{\frac{(1.8)^2}{83} + \frac{(1.5)^2}{65}}$$

$$= (-0.3) \pm 1.96\sqrt{0.039 + 0.034}$$

$$= -0.3 \pm 1.96\sqrt{0.0736}$$

$$= -0.3 \pm 0.5235$$

$$= (-0.3 + 0.5235,\ -0.3 - 0.5235)$$

$$= (0.2035,\ -0.8235)$$

$$\left(\overline{x_1} - \overline{x_2}\right) \pm Z_{\alpha/2}\sqrt{\frac{\sigma_1^2}{n_1} + \frac{\sigma_2^2}{n_2}} = (-0.8235,\ 0.2035).$$

27. Find the estimator for λ by the method of moments for the exponential distribution whose probability density function is given by $f(x,\lambda) = \frac{1}{\lambda} e^{-\frac{x}{\lambda}}$, $x > 0,\ \lambda > 0$.

Solution: **[AU A/M 2024]**

Given: $f(x,\lambda) = \frac{1}{\lambda} e^{-x/\lambda}, x \succ 0, \lambda \succ 0$

The first order moment (about the origin) of the population is given by

$$\mu_1^1 = \int_0^\infty x\frac{1}{\lambda}e^{-x/\lambda}dx$$

$$= \frac{1}{\lambda}\int_0^\infty xe^{-x/\lambda}dx$$

$$= \frac{1}{\lambda}\left\{x\frac{e^{-x/\lambda}}{-1/\lambda} - (1)\frac{e^{-x/\lambda}}{1/\lambda^2}\right\}_0^\infty$$

$$= \frac{1}{\lambda}\left\{x\frac{e^{-x/\lambda}}{-1/\lambda} - (1)\frac{e^{-x/\lambda}}{1/\lambda^2}\right\}_0^\infty$$

$$= \frac{1}{\lambda}\left(\lambda^2\right)$$

$$\mu_1^1 = \lambda$$

The first order moment of the sample $(x_1, x_2, ..., x_n)$ about the orgin is given by

$$m_1^1 = \frac{1}{n}\sum x_i = \overline{x}$$

$$\overline{x} = \lambda$$

$$\hat{\lambda} = \overline{x}$$

28. If $\{x_1, x_2, ..., x_n\}$ is a random sample of size n, drawn from a geometric distribution, then the probability mass function of which is given by $P(x = r) = pq^{r-1};\ r = 1,2,3,.....\infty$. Prove that the mean of the sample is a consistent estimator of the population mean.

Solution: **[AU N/D 2024]**

We know that the mean and variance of the given geometric population are $\frac{1}{p}$ and $\frac{q}{p^2}$

$$E(\overline{x}) = E\left(\frac{1}{n}\sum x_i\right)$$

$$= \frac{1}{n}\frac{n}{p}$$

$$E(\overline{x}) = \frac{1}{p}$$

$$Var(\overline{x}) = Var\left(\frac{1}{n}\sum x_i\right)$$

$$= \frac{1}{n^2}\sum_{i=1}^{n} Var(x_i). \quad \text{Since } x_1, x_2, ..., x_n \text{ are independent}$$

$$= \frac{1}{n^2}n\frac{q}{p^2}$$

$$Var(\bar{x}) = \frac{q}{np^2} \quad \text{as } n \to \infty,\ E(\bar{x}) \to \frac{1}{p}\ and\ Var(\bar{x}) = 0$$

$\therefore\ \bar{x}$ is a consistent estimator of the population mean $\frac{1}{p}$.

29. Let $\{x_1, x_2, ..., x_n\}$ is a random sample from a population density function $f(x;\ \theta, \mu) = \theta\, e^{-\theta(x-\mu)}\ ; x > \mu.$. Find the method of moments estimators of θ and μ.

Solution: **[AU N/D 2024]**

$$f(x;\theta,\mu) = \theta e^{-(x-\mu)};\ x \succ \mu$$

The first order moment (about the origin) of the population is given by

$$\mu_1^1 = \int x f(x;\theta,\mu)dx$$

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

$$= \theta\int_{\mu}^{\infty} x e^{-\theta x} e^{\theta\mu} dx$$

$$= \theta e^{\theta\mu}\int_{\mu}^{\infty} x e^{-\theta x} dx$$

$$= \theta e^{\theta\mu}\left\{ x\left(\frac{e^{-\theta x}}{-\theta}\right) - (1)\left(\frac{e^{-\theta x}}{\theta^2}\right)\right\}_{\mu}^{\infty}$$

$$= \theta e^{\theta\mu}\left\{ 0 + \mu\frac{e^{-\theta\mu}}{\theta} + \frac{e^{-\theta\mu}}{\theta^2}\right\}$$

$$= \theta e^{\theta\mu}\left\{\frac{\mu e^{-\theta\mu} + e^{-\theta\mu}}{\theta^2}\right\}$$

$$= \theta\left\{\frac{\theta\mu + 1}{\theta^2}\right\}$$

$$= \frac{1 + \theta\mu}{\theta}$$

$$\mu_1^1 = \frac{1}{\theta} + \mu \quad ...(1)$$

$$\mu_2^1 = \int x^2 f(x;\theta,\mu)dx$$

$$= \theta\int_{\mu}^{\infty} x^2 e^{-\theta(x-\mu)}dx$$

$$= \theta e^{\theta\mu}\int_{\mu}^{\infty} x^2 e^{-\theta x} dx$$

$$= \theta e^{\theta\mu}\left\{x^2\left(\frac{e^{-\theta x}}{-\theta}\right) - (2x)\left(\frac{e^{-\theta x}}{\theta^2}\right) + (2)\left(\frac{e^{-\theta x}}{-\theta^3}\right)\right\}_{\mu}^{\infty}$$

$$= \theta e^{\theta\mu}\left\{0 + \mu^2\frac{e^{-\theta\mu}}{\theta} + 2\mu\frac{e^{-\theta\mu}}{\theta^2} + 2\frac{e^{-\theta\mu}}{\theta^3}\right\}$$

$$= \mu^2 + \frac{2\mu}{\theta} + \frac{2}{\theta^2}$$

$$\mu_2^1 = \left(\mu + \frac{1}{\theta}\right)^2 + \frac{1}{\theta^2} \quad ...(2)$$

The first order moment of the sample $(x_1, x_2, ..., x_n)$ about the orgin is given by

$$m_1^1 = \frac{1}{n}\sum x_i = \bar{x}$$

The second order moment about the orgin is given by

$$m_2^1 = S^2 = \frac{1}{n}\sum x_i^2$$

By the method of moments,

$$\mu_1^1 = m_1^1 \ \ and \ \ \mu_2^1 = m_2^1$$

Using (1) *and* (2)

$$\mu_2^1 = \mu_1^{1^2} + \frac{1}{\theta^2}$$

$$\frac{1}{\theta^2} = \mu_2^1 - \mu_1^{1^2} \quad ...(3)$$

$$\hat{\theta} = \frac{1}{\sqrt{\mu_2^1 - \mu_1^{1^2}}}$$

Sub. (3) *in* (1)

$$\mu_1^1 = \frac{1}{\theta} + \mu$$

$$\mu = \mu_1^1 - \frac{1}{\theta}$$

$$\hat{\mu} = \mu_1^1 - \sqrt{\mu_2^1 - \mu_1^{1^2}}$$

ANNA UNIVERSITY QUESTIONS

1. Show that the sample mean $\bar{x}$ is an unbiased estimator for the population mean μ. [Pg. no. 8]
2. If $x_1, x_2, ..., x_n$ is a random sample from a normal population $N(\mu, 1)$, show that $t = \frac{1}{n}\sum_{i=1}^{n} x_i^2$ is an unbiased estimator for $\mu^2 + 1$. [Pg. no. 8]
3. Prove that for a random sample $(x_1, x_2, ..., x_n)$ of size n drawn from a given large population (μ, σ^2), $s^2 = \frac{1}{n}\sum_{i=1}^{n}\left(x_i - \bar{x}\right)^2$ is not an unbiased estimator of the parameter σ^2, but $\frac{ns^2}{n-1} = \frac{1}{n-1}\sum_{i=1}^{n}\left(x_i - \bar{x}\right)^2$ is an unbiased estimator of σ^2. [Pg. no. 8]
4. Show that $s^2 = \frac{1}{n-1}\sum_{i=1}^{n}\left(x_i - \bar{x}\right)^2$ is not an unbiased estimator of the parameter σ^2. [Pg. no. 10]
5. Below you are given the values obtained from a random of observations taken from an infinite population

   33 34 35 39

   (a) Find a point estimator for μ. Is this an unbiased estimate of μ ? Explain [Pg. no. 11]

   (b) Find a point estimator for σ^2. Is this an unbiased estimate of σ^2

   (c) Find a point estimator for σ^2

   (d) What can be said about the sampling distribution of $\bar{x}$? Be sure to discuss the expected value, the standard deviation, and the shape of the sampling distribution of $\bar{x}$?
6. A random sample $(X_1, X_2, X_3, X_4, X_5)$ of size 5 is drawn from a normal population with unknown mean μ. Consider the following estimators to estimate μ:

   $(i)\ t_1 = \frac{X_1 + X_2 + X_3 + X_4 + X_5}{5}$, $(ii)\ t_2 = \frac{X_1 + X_2}{2} + X_3$ $(iii)\ t_3 = \frac{2X_1 + X_2 + \lambda X_3}{3}$

   where λ is such that t_3 is an unbiased estimator of μ. Find λ. Are t_1 and t_2 unbiased? State giving reasons, the estimator which is best among t_1, t_2 and t_3. [Pg. no. 11]
7. Let X_1, X_2 *and* X_3 is a random sample of size 3 from a normal population with mean value μ and variance σ^2. T_1, T_2 *and* T_3 are the estimators used to estimate mean μ, where

   $$T_1 = X_1 + X_2 - X_3,\ T_2 = 2X_1 + 3X_3 - 4X_2\ \ and\ \ T_3 = \frac{1}{3}\left(\lambda X_1 + X_2 + X_3\right)$$

(i) Are T_1 *and* T_2 unbiased estimators?

(ii) Find the value of λ such that T_3 is unbiased estimator for μ.

(iii) With this value of λ is T_3 a consistent estimator?

(iv) Which is the best estimator? [Pg. no. 13]

8. If $x_1, x_2, \ldots, x_n$ are random observations of a Bernoulli's variate x which assumes values 1 and 0 with probabilities p *and* $(1-p)$ respectively. Show that $\frac{T(n-T)}{n(n-1)}$ is an unbiased estimator of $p(1-p)$, where $T = x_1 + x_2 + \ldots + x_n$ [AU N/D 2024][Pg. no. 14]

9. If t_1 is a most efficient estimator and t_2 is an unbiased estimator (of some population parameter) with efficiency e, and if the correlation coefficient between t_1 and t_2 is ρ. Show that $\rho = \sqrt{e}$. [Pg. no. 15]

10. Find the estimator of θ in the population with density function $f(x,\theta) = \theta x^{\theta-1}; 0 < x < 1; \theta > 0$, by method of moments. [Pg. no. 15]

11. If $(x_1, x_2, \ldots, x_n)$ a random sample from the uniform population with the density function $f(x,a,b) = \frac{1}{b-a}; a < x < b$. Find the estimators of a and b by method of moments. [Pg. no. 16]

12. For the probability mass function $f(x,p) = 3C_x \frac{p^x(1-p)^{3-x}}{1-(1-p)^3}$; $x = 1,2,3$. Obtain the estimators of p by method of moments, if the frequencies at $x = 1, 2$ *and* 3 are respectively 22, 20 and 18. [Pg. no. 17]

13. A random variable X takes the values 0, 1, 2 with probabilities $\frac{1}{2} - \theta, \frac{\alpha}{2} + 2(1-\alpha)\theta$ *and* $\left(\frac{(1-\alpha)}{2}\right) + (2\alpha - 1)\theta$, where α *and* θ are the parameters. If a sample of size 75 drawn from the population yielded the values 0, 1, 2 with respective frequencies 27, 38, 10 respectively, find the estimators of α *and* θ by the method of moments. [Pg. no. 18]

14. Consider a characteristic that occurs in proportion p of a population. Let $X_1, X_2, \ldots, X_n$ be a random sample of size n so $P[X_i = 0] = 1 - p$ *and* $P[X_i = 1] = p$ *for* $i = 1,2,\ldots,n$ where $0 \le p \le 1$ .Obtain the maximum likelihood estimator of p.

(or)

Find the maximum likelihood estimator for the parameter p of the binomial distribution $B(N,P)$ where N is very large but finite, on the basis of sample of size n , Also find its inverse. [Pg. no. 18]

15. Let $X_1, X_2, ..., X_n$ be a random sample of size n from the Poisson distribution [Pg. no. 20]
$f(x/\lambda) = \dfrac{e^{-\lambda}\lambda^x}{x!}$ where $0 \le \lambda < \infty$. Obtain the maximum likelihood estimator of λ.

16. For random sampling from a normal population $N(\mu, \sigma^2)$, find the maximum likelihood estimators for
(i) μ, when σ^2 is known.
(ii) σ^2, when μ is known and
(iii) The simultaneous estimation of μ and σ^2 . [AU A/M 2024] [Pg. no. 20]

17. Obtain the maximum likelihood estimators a *and* b in terms of the sample observations $x_1, x_2, ..., x_n$ taken from the exponential population with density function
$f(x,a,b) = ke^{-b(x-a)}$; $x \ge a,\ b > 0$ [AU N/D 2024] [Pg. no. 22]

18. Find the maximum likelihood estimators θ in population with density function
$f(x,\theta) = (1+\theta)x^{\theta};\ 0 \le x \le 1;\ \theta > 0;$ based on random sample of size $n.$ Test whether this estimator is a sufficient estimator of θ. [Pg. no. 22]

19. For the double Poisson distribution: [Pg. no. 23]
$$p(x) = P(X = x) = \frac{1}{2}\frac{e^{-m_1}m_1^x}{x!} + \frac{1}{2}\frac{e^{-m_2}m_2^x}{x!}\ ;\ x = 0,1,2,.....$$
Show that the estimates for m_1 *and* m_2 by the method of moments are: $\mu_1' \pm \sqrt{\mu_2' - \mu_1' - \mu_1'^2}$

20. Let $x_1, x_2, ..., x_n$ be a random sample from the uniform distribution with p.d.f : [Pg. no. 24]
$f(x,\theta) = \begin{cases} \dfrac{1}{\theta},\ 0 < x < \infty,\ \theta > 0 \\ 0 \quad ,\quad otherwise \end{cases}$ Obtain the maximum likelihood estimator for θ

21. Obtain the maximum likelihood estimators α *and* β for the rectangular population:
$$f(x,\alpha,\beta) = \begin{cases} \dfrac{1}{\beta-\alpha},\ \alpha < x < \beta \\ 0 \quad ,\quad otherwise \end{cases}$$ [Pg. no. 25]

22. In order to introduce some incentive for higher balance in savings accounts, a random sample size of 64 savings accounts at a bank's branch was studied to estimate the average monthly balance in saving bank accounts. The mean and standard deviation were found to be Rs. 8,500 and Rs. 2,000, respectively.
Find (i) 90%
(ii) 95%
(iii) 99% confidence intervals for the population mean. [Pg. no. 26]

23. In a certain factory there are two independent processes manufacturing the same item. The average weight in a sample of 250 items produced from one process is found to be 120 O_{zs}, with a standard deviation of 12 O_{zs} . While the corresponding figures in a sample of 400 items from the other process are 124 O_{zs} and 14 O_{zs} . Find the 995 confidence limits for the difference in the average weight of items produced by the two processes respectively. [Pg. no. 27]

24. In a test given to two groups of students the marks obtained were as follows:

First group:	18	20	36	50	49	36	34	49	61
Second group:	29	28	26	35	30	44	46		

Construct a 95% confidence interval on the mean marks secured by students of the above two groups. [Pg. no. 27]

25. Let X be exponentially distributed with parameter λ . Using maximum likelihood estimation, find an estimate for the parameter λ. [AU A/M 2024] [Pg. no. 28]

26. If 83 male students are randomly chosen and yield an average of 6.6 hours of sleep with a standard deviation of 1.8 and 65 females are randomly selected with an average of 6.9 hours of sleep with a standard deviation of 1.5. Construct a 95 % confidence interval for the difference between the two mean sleep hours for males and females. [AU A/M 2024] [Pg. no. 29]

27. Find the estimator for λ by the method of moments for the exponential distribution whose probability density function is given by $f(x,\lambda)=\frac{1}{\lambda}e^{-\frac{x}{\lambda}}$, $x>0,\ \lambda>0$. [AU A/M 2024] [Pg. no. 29]

28. If $\{x_1, x_2, \ldots\ldots, x_n\}$ is a random sample of size n , drawn from a geometric distribution, then the probability mass function of which is given by $P(x=r)=pq^{r-1};\ r=1,2,3,\ldots\ldots\infty$ Prove that the mean of the sample is a consistent estimator of the population mean. [AU N/D 2024] [Pg. no. 30]

29. Let $\{x_1, x_2, \ldots\ldots, x_n\}$ is a random sample from a population density function $f(x;\ \theta,\mu)=\theta\, e^{-\theta(x-\mu)}$ $;x>\mu.$. Find the method of moments estimators of θ and μ. [AU N/D 2024] [Pg. no. 31]

UNIT-IV
NON-PARAMETRIC TESTS
PART-A

1. **What is meant by non-parametric test?**
Solution:
The tests which are not based on any assumption are called as non-parametric test. Non-parametric tests are sometimes called distribution-free tests.

2. **Name four non-parametric tests used in statistical study.**
Solution:
(i) Chi-square test (ii) Run test (iii) U-test (iv) H-test (v) Rank test (vi) Sign test (vii) Spearman's rank correlation coefficient.

3. **Write any two advantages of non-parametric methods over parametric methods.**
Solution: **[AU A/M 2024]**

	Parametric test	**Non- Parametric test**
1.	Specific assumptions are made regarding the population.	No assumptions are made regarding the population.
2.	Null hypothesis is made on parameters of the population distribution	The null hypothesis is free from parameters

4. **What are the primary short comings of non-parametric tests?**
Solution:
The disadvantages of the non-parametric test are less efficient as compared to parametric test. The results may or may not provide an accurate answer because they are distribution free.

5. **State the limitations of non-parametric test. (or) Disadvantages of non-parametric test.**
Solution: **[AU N/D 2024]**
(i) Cannot easily use confidence intervals or effect sizes.
(ii) Have less statistical power than parametric tests.
(iii) Nominal and ordinal data provide less information.
(iv) More likely to commit type II error.

6. **Explain the sign test stating clearly assumption made.**
Solution:
The sign test is based on + and – signs.
There are two types of sign tests.
(i) One sample sign test.
(ii) Two sample sign test.
We replace the given data by two symbols + and -, if the given values are greater than μ by + sign and if given values are less than μ by – sign.
This test is based on the assumption that the population is continuous and symmetric.

7. What do you meant by rank test?

Solution:

In statistics – rank test is any test involving ranks.

Examples:

(i) Wilcoxon signed – rank test.

(ii) Kruskal-wallis one-way analysis of variance.

(iii) Mann-Whitney U-test.

8. What is Mann U Whitney U-test used for?

Solution:

Then Mann Whitney U-test, sometimes called the Mann Whitney Wilcoxon test or the Wilcoxon rank sum test, is used to test whether two samples are likely to drive from the same population.

9. Distinguish between the Mann-Whitney U-test and the Kruskal–Wallis test.

Solution:

The major difference between the Mann-Whitney U and the Kruskal-Wallis H is simply that the latter can accommodate more than two groups. Both tests require independent (between – subjects) designs and use summed rank scores to determine the results.

10. Write the formula for Kruskal Wallis test.

Solution:

$$H = \frac{12}{n(n+1)} \sum_{i=1}^{k} \frac{R_i^2}{n_i} - 3(n+1)$$

Where $H \to Kruskal - Wallis\ test$

$R_i \to Rank\ of\ the\ sample$

$n \to Total\ number\ of\ observtions\ in\ all\ samples.$

11. Give an example of run-test.

Solution:

LL WWW LLLL W LLL

Here LL is one run, WWW is one run, LLLL is one run

W is one run, LLL is one run

Total number of runs $= 5$.

12. Where do we use run-test?

Solution:

This test is used to test the randomness of the given data.

13. What are the uses run-test?

Solution:

(i) Can be used to decide if a data set is from a random process.

(ii) Traders who focus on technical analysis can we use a runs test to help analyze the price of a Security.

14. What is the use of KW (Kruskal-Wallis) tests?
Solution:

A KW test is used to determine whether or not there is a statistically significant difference between the medians of three or more independent groups.

15. What are the assumptions of the Mann-Whitney?
Solution:

(i) The sample drawn from the population is random.
(ii) Independence within the samples and mutual independence is assumed.
i.e., That means that an observation is in one group or the other(it cannot be in both).
(iii) Ordinal measurement scale is assumed.

16. Explain stating the necessary assumptions the two sample Mann-Whitney U-test
Solution:

Formula: $$Z = \frac{U - \mu_U}{\sigma_U}$$

$$U = n_1 n_2 + \left[\frac{n_1(n_1+1)}{2}\right] - R_1$$

$$\mu_U = \frac{n_1 n_2}{2}$$

$$\sigma_U = \sqrt{\frac{n_1 n_2 (n_1 + n_2 + 1)}{12}}$$

$n_1 \rightarrow$ Size of the first sample.
$n_2 \rightarrow$ Size of the second sample.
$R_1 \rightarrow$ Sum of the ranks of the first sample.
$R_2 \rightarrow$ Sum of the ranks of the second sample.

We have to consider sum of the ranks of the smaller sample.
When some ranks are repeated, we have to follow the method as shown below.

In case, third rank is repeated two times the corresponding ranks will be $\frac{3+4}{2} = \frac{7}{2}$

i.e., 3.5, 3.5

In case, fourth rank is repeated three times the corresponding ranks will be $\frac{4+5+6}{2} = 5$

i.e., 5,5,5

17. What is the Kolmogorov-Smirnov Test? Is it non parametric?
Solution: **[AU A/M 2024]**

The Kolmogorov-test compares your data with a known distribution and lets you know of they have the same distribution . Commonly used as test for normally. It is non-parametric test.

18. What is the use of median test?

Solution: **[AU N/D 2024]**

Median test is used to test if two samples differ in their central tendencies.

PART-B

SIGN TEST TYPE: 1-A $[n \leq 20]$

1. The following data constitute a random sample of 15 measurements of the octane rating of a certain kind of gasoline:

99.0	**102.3**	**99.8**	**100.5**	**99.7**	**96.2**	**99.1**	**102.5**
103.3	**97.4**	**100.4**	**98.9**	**98.3**	**98.0**	**101.6**	

Test the null hypothesis $\mu = 98.0$ **against the alternative hypothesis** $\mu > 98.0$ **at the 0.01 level of significance.**

Solution:

Given: $\mu = 98.0$

One of the sample value equal to 98 and must be discarded.

So, $n = 15 - 1 = 14$

1. $H_0 : \mu = 98 \quad (p = \frac{1}{2})$
2. $H_1 : \mu > 98 \quad (p > \frac{1}{2})$
3. $\alpha = 0.01 \quad (Given)$
4. Reject H_0 if $p(x) < 0.01$
5. Let $x = number\ of\ +ve\ signs$

Subtract 98 from the given data we get

+ + + + + - + + + - + + + +

$\therefore\ x = 12,\ n = 14,\ p = 0.5$

$$p(x) = P[X \geq 12] = 1 - P[X < 12] = 1 - P[X \leq 11]$$
$$= 1 - 0.9935$$
$$= 0.0065$$

6. Conclusion:

Here, 0.0065 < 0.01

$\therefore$ Reject H_0 at 1% level

We conclude that the median octane rating of the given kind of gasoline exceeds 98.0

2. In a factory, 20 observations of the factors that could heat up a conveyor belt yielded the following results: 0.36, 0.41, 0.25, 0.34, 0.28, 0.26, 0.39, 0.28, 0.40, 0.26, 0.35, 0.38, 0.29, 0.42, 0.37, 0.37, 0.39, 0.32, 0.29 and 0.36. Use the sign test at the 0.01 level of significance to test the null hypothesis $\mu = 0.34$ **against the alternative hypothesis** $\mu \neq 0.34$

Solution:

Given: $\mu = 0.34$

One of the sample value equal to 0.34 and must be discarded.

So, $n = 20 - 1 = 19$

1. $H_0 : \mu = 0.34 \quad (p = \frac{1}{2} = 0.5)$
2. $H_1 : \mu \neq 0.34 \quad (p \neq \frac{1}{2} \neq 0.5)$
3. $\alpha = 0.01 \quad (Given)$
4. Reject H_0 if $p(x) < 0.01$
5. Let $x = number\ of\ +ve\ signs$

Subtract 0.34 from the given data we get

+ + - - - + - + - + + - + + + + + - - +

$\therefore\ x = 11,\ n = 19,\ p = 0.5$

$$p(x) = P[X \geq 11] = 1 - P[X < 11] = 1 - P[X \leq 10]$$
$$= 1 - 0.6762$$
$$= 0.3238$$

6. Conclusion:

   Here, 0.3238 > 0.01

   $\therefore$ we accept H_0 at 1% level, (or) we cannot reject H_0

SIGN TEST TYPE: 1-B $[n > 20]$

3. The time sheet of a factory showed the following sample data (in hours) on the time spent by a worker operating a hydraulic gear lift : 1.0, 0.8, 0.5, 0.9, 1.2, 0.9, 1.4, 10, 1.3, 0.8, 1.5, 1.2, 1.9, 1.1, 0.7, 0.8, 1.1, 1.2, 1.5, 1.1, 1.8, 0.5, 0.8, 0.9 and 1.6. Use the sign test at the 0.05 level of significance to test the null hypothesis $\mu = 1.1$ **against the alternative hypothesis** $\mu > 1.1$

Solution:

Given: $\mu = 1.1$

Three of the sample values are equal to 1.1 and must be discarded.

So, $n = 25 - 3 = 22$

1. $H_0 : \mu = 1.1 \quad (p = \frac{1}{2} = 0.5)$
2. $H_1 : \mu > 1.1 \quad (p > \frac{1}{2} > 0.5)[One - table\, test]$
3. $\alpha = 0.05 \quad (Given)$
4. Reject H_0 if $|Z| > 1.645$
5. Let $u = number\ of\ +ve\ signs$

Subtract 1.1 from the given data we get

- - - - + - + - + - + + + - - + + + - - - +

$\therefore\ u = 10,\ n = 22,\ p = 0.5$

$$|Z| = \frac{u - np}{\sqrt{npq}} = \frac{10 - (22)(0.5)}{\sqrt{(22)(0.5)(0.5)}}$$

$$= \frac{10-11}{\sqrt{5.5}} = -0.4264$$

$$|Z| = 0.4264$$

6. Conclusion:

Here, 0.4264 < 1.645

$\therefore$ Accept H_0 at 5% level.

4. The following are the sizes of particles of cement dust (given to the nearest hundredth of a Micron) in a cement factory.

16.12	**10.48**	**11.12**	**16.18**	**18.13**	**19.10**	**13.21**	**10.12**
21.18	**15.12**	**10.11**	**13.31**	**18.61**	**11.43**	**18.26**	**13.77**
13.24	**12.16**	**17.19**	**11.36**	**12.53**	**13.25**	**10.67**	**15.45**
14.28	**14.32**	**15.18**	**14.21**	**10.20**	**15.64**	**11.68**	**18.76**
19.32	**17.50**	**11.46**	**20.59**	**16.38**	**21.42**	**16.27**	**21.30**
16.12	**10.55**	**11.49**	**15.48**	**11.62**	**13.54**	**13.69**	**16.72**
15.11	**14.33**	**17.23**	**17.22**	**19.37**	**10.41**	**18.28**	**19.29**
21.23	**12.56**	**12.57**	**11.60**	**15.24**	**21.65**	**20.70**	**11.44**
12.22	**19.34**	**20.35**	**19.47**	**21.63**	**19.40**	**19.75**	**21.71**
15.19	**18.51**	**10.58**	**13.52**	**11.39**	**13.66**	**21.73**	**11.74**

Which pertained to the particle size of cement dust in a factory producing cement, use the sign? test at the 0.05 level of significance to test the null hypothesis $\mu = 15.13$ hundredth of a micron against the alternative hypothesis $\mu < 15.13$ hundredth of a micron.

Solution:

Given: $\mu = 15.13$

$n = 80$

1. $H_0 : \mu = 15.13 \quad (p = \frac{1}{2} = 0.5)$
2. $H_1 : \mu < 15.13$
3. $\alpha = 0.05 \quad (Given)$
4. Reject H_0 if $|Z| > 1.645$
5. Let $u = number\ of\ +ve\ signs$

Subtract 15.13 from the given data

$\therefore u = 41,\ n = 80,\ p = 0.5$

$$|Z| = \frac{u - np}{\sqrt{npq}} = \frac{41-(80)(0.5)}{\sqrt{(22)(0.5)(0.5)}}$$

$$= \frac{41-40}{\sqrt{20}} = 0.2236$$

$$|Z| = 0.2236$$

6. Conclusion:
   Here, 0.2236 < 1.645
   So, Accept H_0 at 5% level.

TYPE: 2

5. The following are the average weekly losses of worker-hours due to accidents in 10 industrial plants before and after a certain safety program was into operation

Before	**45**	**73**	**46**	**124**	**33**	**57**	**83**	**34**	**26**	**17**
After	**36**	**60**	**44**	**119**	**35**	**51**	**77**	**29**	**24**	**11**

Use the 0.05 level of significance to test whether the safety program is effective.
Solution:

1. $H_0: \ p = \frac{1}{2} = 0.5$
2. $H_1: \ p > \frac{1}{2}$
3. $\alpha = 0.05 \quad (Given)$
4. Reject H_0 if $p(x) \le 0.05$
5. Let $x = number\ of\ +ve\ signs$
   Now, the sign of scores are (before – after)
   $+ + + +\ - + + + + +$
   $\therefore x = 9,\ n = 10,\ p = 0.5$
   $$p(x) = P[X \ge 9] = 1 - P[X < 9] = 1 - P[X \le 8]$$
   $$= 1 - 0.9893$$
   $$= 0.0107$$
6. Conclusion:
   Here, 0.0107 < 0.05
   $\therefore$ Reject H_0 at 5% level of significance.
   We conclude that the safety program is effective.

6. Use the sign test t see if there is a difference between the number of days required to collect an account receivable before and after a new collection policy. Use the 0.05 significance level.

Before:	**33**	**36**	**41**	**32**	**39**	**47**	**34**	**29**	**32**	**34**	**40**	**42**	**33**	**36**	**27**
After:	**35**	**29**	**38**	**34**	**37**	**47**	**36**	**32**	**30**	**34**	**41**	**38**	**37**	**35**	**28**

Solution: **[AU N/D 2024]**

Null hypothesis $H_0 : p = 0.5$.

Alternate hypothesis $H_1 : p \neq 0.5$.

$$\alpha = 0.05$$

Reject H_0 *if* $p(x) \leq 0.5$

Let x = number of positive signs.

Now, the sign of scores are (before - after)

\- + + - + 0 - - + 0 - + - + -

Number of positive signs = 6

Number of negative signs = 7

Number of zeros = 2

Total sample size = 15

We exclude the evaluation of 0's.

$\therefore Total = n = 13,\ x = 6,\ p = 0.5$

At $\alpha = 0.05$ level of significance. $Z_\alpha = 1.96$

$$P(x) = P(X \geq 6) = 1 - P(X < 6)$$
$$= 1 - P(X \leq 5)$$
$$= 1 - 0.2905$$
$$= 0.7095$$

Conclusion:

0.7095 > 0.05

H_0 is rejected at 5% level of significance

7. A consumer panel includes 14 individuals. It is asked to rate two brands of co-co cola according to a point evaluation system based on several criteria. The table gives below reports the points assigned. Test the null hypothesis that there is no difference in the level of ratings for the two brands of cola at 5% level of significance using the sign test.

Panel member	**1**	**2**	**3**	**4**	**5**	**6**	**7**	**8**	**9**	**10**	**11**	**12**	**13**	**14**
Brand I:	**20**	**24**	**28**	**24**	**20**	**29**	**19**	**27**	**20**	**30**	**18**	**28**	**26**	**24**
Brand II:	**16**	**26**	**18**	**17**	**20**	**21**	**23**	**22**	**23**	**20**	**18**	**21**	**17**	**26**

Solution:

1. $H_0: \ p = \frac{1}{2} = 0.5$
2. $H_1: \ p \neq \frac{1}{2}$
3. $\alpha = 0.05 \quad (Given)$
4. Reject H_0 if $p(x) \leq 0.05$
5. Let $x = number\ of\ +ve\ signs$

Now, the sign of scores are (Brand II-Brand I)

- + - - 0 - + - + - 0 - - +

$\therefore\ x = 4,\ n = 12,\ p = 0.5$

$$p(x) = P[X \geq 4] = 1 - P[X < 4] = 1 - P[X \leq 3]$$
$$= 1 - 0.0730$$
$$= 0.927$$

6. Conclusion:
   Here, 0.927 < 0.05
   $\therefore$ We accept H_0 at 5% level of significance.

THE SIGNED – RANK TEST

8. The following are 15 measurements of the octane rating of a certain kind of gasoline. 97.5, 95.2, 97.3, 96.0, 96.8, 100.3, 97.4, 95.3, 93.2, 99.1, 96.1, 97.6, 98.2, 98.5 and 94.9. Use the signed-rank Test at the 0.05 level of significance to test the mean octane rating of the given kind of gasoline is 98.5.

Solution:

Given: $n = 14$

1. $H_0 : \mu = 98.5$
2. $H_1 : \mu \neq 98.5$
3. $\alpha = 0.05$
4. Reject H_0 if $T \leq T_{0.05}$, where $T_{0.05}$ must be read from table 2 for the appropriate value of n
   $T_{0.05} = 21$ for $n = 14$
5. Subtracting 98.5 from each value and ranking the differences without regard to their sign, we get

Measurement	Difference	Rank	T^+	T^-
97.5	-1.0	4		4
95.2	-3.3	12		12
97.3	-1.2	6		6
96.0	-2.5	10		10

96.8	-1.7	7		7
100.3	1.8	8		
97.4	-1.1	5	8	
95.3	-3.2	11		5
93.2	-5.3	14		11
99.1	0.6	2		14
96.1	-2.4	9	2	
97.6	-0.9	3		9
98.2	-0.3	1		3
98.5	0.0	-		1
94.9	-3.6	13		-
				13
			10	95

$T^{+} = 8 + 2 = 10$

$T^{-} = 4 + 12 + 6 + 10 + 7 + 5 + 11 + 14 + 9 + 3 + 1 + 13 = 95$

T = Smaller of T^{-} and T^{+}

$= 10$

6. Conclusion:

Here, $10 < 21$, so, we reject H_0

The mean octane rating the given kind of gasoline is not 98.5.

9. Drop in diastolic blood pressure (in mm mercury)

Drug D_1	**10**	**16**	**10**	**4**	**2**	**14**	**4**
Drug D_2	**33**	**34**	**41**	**36**	**42**	**42**	**32**

Test whether there is any difference in the effectiveness of drugs at $\alpha = 0.05$. Using Wilcoxon's signed rank test. (Table value for $n = 7$ is 2)

Solution:

Given: $n = 7$

1. $H_0 : \mu_1 = \mu_2$

2. $H_1 : \mu_1 \neq \mu_2$
3. $\alpha = 0.05$
4. Reject H_0 if $T \leq T_{0.05}$ =2
5. Here 4 are repeated twice.

$\therefore$ Second and third ranks are same.

$\therefore$ This corresponds to rank $\frac{2+3}{2} = 2.5$

Drug D_1	Drug D_2	Difference $D = D_1 - D_2$	Absolute Difference $\lvert D \rvert$	Signed rank	T^+	T^-
10	6	4	4	2.5	2.5	
16	20	-4	4	-2.5		2.5
10	8	2	2	1	1	
4	12	-8	8	-5		5
2	8	-6	6	-4		4
14	4	10	10	6	6	
4	15	-11	11	-7		7
					$T^+ = 9.5$	$T^- = 18.5$

$T^+ = 2.5 + 1 + 6 = 9.5$

$T^- = 2.5 + 5 + 4 + 7 = 18.5$

T = Smaller of T^- and T^+

= 9.5

6. Conclusion:

Here, 9.5 > 2, so, we accept H_0

$\therefore$ There is no significant difference, between the given scores and median.

10. In a market research, it was decided to examine the effect of brand name on quality perception.16 subjects are recruited for the purpose and are asked to taste and compare two sample of product on a set of scale items judged to be ordinal. The following data are obtained (table value for $n = 15$ is 25)

Pair	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16
Item A:	**73**	**43**	**47**	**52**	**58**	**47**	**52**	**58**	**38**	**61**	**56**	**56**	**34**	**55**	**65**	**75**
Item B:	**51**	**41**	**43**	**41**	**47**	**32**	**24**	**58**	**43**	**63**	**52**	**57**	**44**	**57**	**40**	**68**

Test the hypothesis, using Wilcoxon matched-pairs test, that there is no difference between the perceived qualities of the two samples. Use 5% level of significance.

Solution:

Given: $n = 16 - 1 = 15$

1. $H_0 : \mu_1 = \mu_2$
2. $H_1 : \mu_1 \neq \mu_2$
3. $\alpha = 0.05$
4. Reject H_0 if $T \leq T_{0.05}$ =25
5. Since some rank is repeated the calculation for entering rank is shown below.

$$\left.\begin{matrix} 2^{nd}\ rank \\ and\ 3^{rd}\ rank \end{matrix}\right\} \Rightarrow \frac{2+3}{2} = 2.5$$

$$\left.\begin{matrix} 4^{th}\ rank \\ and\ 5^{th}\ rank \end{matrix}\right\} \Rightarrow \frac{4+5}{2} = 4.5$$

$$\left.\begin{matrix} 10^{thd}\ rank \\ and\ 11^{th}\ rank \end{matrix}\right\} \Rightarrow \frac{10+11}{2} = 10.5$$

Pair	**Product A**	**Product B**	**D= A-B**	$\lvert D\rvert$	**Signed rank**
1	73	51	22	22	13
2	43	41	2	2	2.5
3	47	43	4	4	4.5
4	52	41	11	11	10.5
5	58	47	11	11	10.5
6	47	32	15	15	12
7	52	24	28	28	15
8	58	58	0	0	-
9	38	43	-5	5	6
10	61	53	8	8	8

11	56	52	4	4	4.5
12	56	57	-1	1	1
13	34	44	-10	10	9
14	55	57	-2	2	2.5
15	65	40	25	25	14
16	75	68	7	7	7

The pair 8 can be dropped as D value for zero and

$\therefore\ n = 16 - 1 = 15$

T^+ = Sum of +ve Rank values
= 13 + 2.5 + 4.5 + 10.5 + 10.5 + 12 + 15 + 8 + 4.5 + 14 + 7
= 101.5

T^- = Sum of –ve Rank values
= 6 + 1 + 9 + 2.5 = 18.5

T = Smaller of T^- and T^+
= 18.5

6. Conclusion:

Here, 18.5 < 25, so, we reject H_0

There is significant difference.

11. The following are the weights in pounds before and after of 16 persons who stayed on a certain reducing diet for four weeks.

Before:	**147.0**	**183.5**	**232.1**	**161.6**	**197.5**	**206.3**	**177.0**	**215.4**	**147.7**	**208.1**	**166.8**	**131.9**	**150.3**	**197.2**	**159.8**	**171.7**
After:	**137.9**	**176.2**	**219.0**	**163.8**	**193.5**	**201.4**	**180.6**	**203.2**	**149.0**	**195.4**	**158.5**	**134.4**	**149.3**	**189.1**	**159.1**	**173.2**

Use the signed rank test to test at the 0.05 level of significance whether the weight reducing diet is effective.

Solution:

Given: $n = 16$

1. $H_0 : \mu_1 = \mu_2$
2. $H_1 : \mu_1 \neq \mu_2$
3. $\alpha = 0.05$
4. Reject H_0 if $Z \geq t_{0.05} = 1.96$
5. $Z = \dfrac{T^+ - \mu}{\sigma}$, where $\mu = \dfrac{n(n+1)}{4}$, $\sigma^2 = \dfrac{n(n+1)(2n+1)}{24}$

Before	After	Difference	Rank
147.0	137.9	9.1	13
183.5	176.2	7.3	10
232.1	219.0	13.1	16
161.6	163.8	-2.2	-
197.5	193.5	4	8
206.3	201.4	4.9	9
177.0	180.6	-3.6	-
215.4	203.2	12.2	14
147.7	149.0	-1.3	-
208.1	195.4	12.7	15
166.8	158.5	8.3	12
131.9	134.4	-2.5	-
150.3	149.3	1	2
197.2	189.1	8.1	11
159.8	159.1	0.7	1
171.7	173.2	-1.5	-

$$\mu = \frac{n(n+1)}{4} = \frac{16(16+1)}{4} = 68, \quad \sigma^2 = \frac{n(n+1)(2n+1)}{24} = \frac{16(16+1)(33)}{24} = 374$$

$T^+ = 13 + 10 + 16 + 8 + 9 + 14 + 15 + 12 + 2 + 11 + 1 = 111$

$$Z = \frac{T^+ - \mu}{\sigma} = \frac{111-68}{\sqrt{374}} = 2.22$$

6. Conclusion:

Here, 2.22 > 1.96, so, we reject H_0.

We conclude that the diet is, indeed, effective in reducing weight.

12. A manufacturer of electric irons, wishing to test the accuracy of the thermostat control at the 500^o F setting, instructs a test engineer to obtain actual temperature at that setting for 15 irons

using a thermocouple. The resulting measurements are as follows:

494.6	**510.8**	**487.5**	**493.2**	**502.6**	**485.0**	**495.9**	**498.2**
501.6	**497.3**	**492.0**	**504.3**	**499.2**	**493.5**	**505.8**	

The engineer believes it is reasonable to assume that a temperature deviation from 500^o of any particular magnitude is just as likely to be positive as negative (the assumption of symmetry) but wants to protect against possible non normality of the actual temperature distribution. Use signed-rank test to see whether the data strongly suggests incorrect calibration of the iron.

Solution: **[AU A/M 2024]**

Given n=15

Null hypothesis $H_0 : \mu = 500$.

Alternate hypothesis $H_1 : \mu \neq 500$.

$\alpha = 0.05$

Reject H_0 *if* $T \leq T_{0.05}$, where $T_{0.05}$ is table value.

$T_{0.05} = 21$

Subtracting 500 from each value and ranking the differences without regard to their sign, we get

Temperature	Difference	Rank	T^+	T^-
494.6	-5.4	8		8
510.8	10.8	13	13	
487.5	-12.5	14		14
493.2	-6.8	11		11
502.6	2.6	5	5	
485.0	-15	15		15
495.9	-0.1	1		1
498.2	-1.8	4		4
501.6	1.6	3	3	
497.3	-2.7	6		6
492.0	-8.0	12		12
504.3	4.3	7	7	
499.2	-0.8	2		2
493.5	-6.5	10		10
505.8	5.8	9	9	
			37	54

$T^+ = 37$, $T^- = 54$.

T = Smaller of T^- and $T^+ = 37$

Conclusion:

Here 37>21

So, we accept H_0.

RANK-SUM TESTS – THE U-TEST

13. The nicotine contents of two brands of cigarettes, measured in milligrams, was found to be as follows:

Brand A	**2.1**	**4.0**	**6.3**	**5.4**	**4.8**	**3.7**	**6.1**	**3.3**		
Brand B	**4.1**	**0.6**	**3.1**	**2.5**	**4.0**	**6.2**	**1.6**	**2.2**	**1.9**	**5.4**

Test the hypothesis, at the 0.05 level of significance, that the average nicotine contents of the two brands are equal against the alternative that they are unequal.

Solution:

1. $H_0 : \mu_1 = \mu_2$
2. $H_1 : \mu_1 \neq \mu_2$
3. $\alpha = 0.05 \quad (Given)$
4. Reject H_0 if $Z > 1.96$
5. Computation of U-statistic :
   The observations are arranged in ascending order and ranks from 1 to 18 are assigned.

Original data:	0.6	1.6	1.9	2.1	2.2	2.5	3.1	3.3	3.7
Rank:	1	2	3	4	5	6	7	8	9

Original data:	4.0	4.0	4.1	4.8	5.4	5.4	6.1	6.2	6.3
Rank:	10.5	10.5	12	13	14.5	14.5	16	17	18

The ranks of the observation belonging to the Brand A samples are put in bold form.

R_1 = **4+8+9+10.5+13+14.5+16+18=93**

R_2 = 1+2+3+4+5+7+10.5+12+14.5+17=78

$n_1 = 8,\ n_2 = 10,$

$$\therefore \text{ U-statistic : U } = n_1 n_2 + \frac{n_1(n_1+1)}{2} - R_1$$

$$= (8)(10) + \frac{8(8+1)}{2} - 93 \quad = 80 + 36 - 93$$

$$= 23$$

$$\text{Mean of U} = \mu_U \ = \frac{n_1 n_2}{2} = \frac{(10)(8)}{2} = 40$$

$$\text{Variance of U} = \mu_U^2 \ = \frac{n_1 n_2 (n_1 + n_2 + 1)}{12}$$

$$= \frac{(10)(8)(10+8+1)}{12} = 126.67$$

$$\alpha_U = \sqrt{126.67} = 11.25$$

Here $n_2 = 11$, so we use the statistic

$$Z = \frac{U-\mu_U}{\sigma_U} = \frac{23-40}{11.25} = -1.51$$

$$|Z| = 1.51$$

The tabled value of Z_α at $\alpha = 0.05$ 1.96 [From normal table]

6. Conclusion:
   Here, 1.51< 1.96
   $\therefore$ Accept H_0 at 5% level
   There is no significant difference in the average nicotine contents of the two brands of cigarettes.

14. Two classes of students are tested using a certain competitive exam. The scores off a sample of students from each class is given below:

Class A	**45**	**44**	**47**	**48**	**55**	**53**	**55**	**63**			
Class B	**65**	**67**	**77**	**65**	**56**	**67**	**78**	**55**	**66**	**65**	**58**

Use Mann Whitney-U test to whether both classes have similar scholastic levels.

Solution:

1. $H_0 : \mu_1 = \mu_2$
2. $H_2 : \mu_1 \neq \mu_2$
3. $\alpha = 0.05$ $(Given)$
4. Reject H_0 if $Z > 1.64$
5. Computation of U-statistic :
   The observations are arranged in ascending order and ranks from 1 to 19 are assigned. The ranks of the observation belonging to the Class A samples are put in bold form.

Original data:	44	45	47	48	53	55	55	55	56	58
Rank:	1	2	3	4	5	7	7	7	9	10

Original data:	63	65	65	65	66	67	67	77	78
Rank:	11	13	13	13	15	16.5	16.5	18	19

R_1 = 1+2+3+4+5+7+7+11=40

R_2 = 7+9+10+13+13+13+15+16.5+16.5+18+19 = 150

Also, $n_1 = 8$, $n_2 = 10$

$\therefore$ U-statistic : U $= n_1 n_2 + \dfrac{n_1(n_1+1)}{2} - R_1$

$$= (8)(11) + \frac{8(8+1)}{2} - 40 = 88 + 36 - 40 = 84$$

Mean of U $= \mu_U = \frac{n_1 n_2}{2} = \frac{88}{2} = 44$

Variance of U $= \mu_U^2 = \frac{n_1 n_2 (n_1 + n_2 + 1)}{12}$

$$= \frac{(8)(11)(8+11+1)}{12} = 146.67$$

$$\alpha_U = \sqrt{146.67} = 12.11$$

Here $n_2 = 11$, so we can use the statistic

$$Z = \frac{U - \mu_U}{\sigma_U} = \frac{84 - 44}{12.11} = 3.3$$

6. Conclusion:
   Here, 3.3 > 1.64
   $\therefore$ Reject H_0 at 5% level

15. Independent random samples of ten day students and ten evening students at a University Showed the following age distributions. We want to use the Mann-Whitney-Wilcoxon test to determine if there is a significant difference in the age distribution of the two groups.

Day	**26**	**18**	**25**	**27**	**19**	**30**	**34**	**21**	**33**	**31**
Evening	**32**	**24**	**23**	**30**	**40**	**41**	**42**	**39**	**45**	**35**

Use Mann Whitney-U test to whether both classes have similar scholastic levels.

(i) Compute the sum of the ranks (T) for the day students.

(ii) Compute the mean μ_T

(iii) Compute σ_T

(iv) Use $\alpha = 0.05$ and test for any significant differences in the age distribution of the two populations.

Original Data	Rank
18	**1**
19	**2**
21	**3**

Solution:

1. $H_0 : \mu_1 = \mu_2$
2. $H_2 : \mu_1 \neq \mu_2$
3. $\alpha = 0.05 \quad (Given)$
4. Reject H_0 if $Z > 1.96$
5. Computation of U-statistic :

The observations are arranged in ascending order and ranks from 1 to 20 are assigned.
The ranks of the observation belonging to the Day are put in bold form.

23	4
24	5
25	**6**
26	**7**
27	**8**
30	**9.5**
30	9.5
31	**11**
32	12
33	**13**
34	**14**
35	15
39	16
40	17
41	18
42	19
45	20

(i) T = R_1 = 1+2+3+6+7+8+9.5+11+13+14 = 74.5
R_2 = 4+5+9.5+12+15+16+17+18+19+20 = 135.5

Also, $n_1 = 10, \; n_2 = 10$

$$\therefore \text{ U-statistic : U } = n_1 n_2 + \frac{n_1(n_1+1)}{2} - R_1$$

$$= (10)(10) + \frac{10(10+1)}{2} - 74.5$$

$$= 100 + 55 - 74.5$$

$$= 80.5$$

(ii) Mean of U $= \mu_T = \dfrac{n_1 n_2}{2} = \dfrac{(10)(10)}{2} = 50$

(iii) Variance of U $= \mu_T^2 = \dfrac{n_1 n_2 (n_1 + n_2 + 1)}{12}$

$$= \frac{(10)(10)(10+10+1)}{12} = 175$$

$$\alpha_T = \sqrt{175} = 13.23$$

Here $n_2 = 10$, so we can use the statistic

$$Z = \frac{U - \mu_U}{\sigma_U} = \frac{80.5 - 50}{13.23} = 2.31$$

(iv) Conclusion:

Here, 2.31 > 1.96

$\therefore$ Reject H_0 at 5% level

16. The urinary fluoride concentration (pairs per million) was measured both for a sample of livestock grazing in an area previously exposed to fluoride pollution and for a similar sample grazing in an unpolluted region:

Polluted	**21.3**	**18.7**	**23.0**	**17.1**	**16.8**	**20.9**	**19.7**
Unpolluted	**14.2**	**18.3**	**17.2**	**18.4**	**20.0**		

Does the data indicate strongly that the true average fluoride concentration for livestock grazing in the polluted region is larger than for the unpolluted region? Use the Wilcoxon rank-sum test at level $\alpha = 0.01$.

Solution: **[AU A/M 2024]**

Null hypothesis $H_0 : \mu_1 = \mu_2$.

Alternate hypothesis $H_1 : \mu_1 < \mu_2$.

$$\alpha = 0.01$$

Reject H_0 *if* $Z > 2.33$

Computation of Wilcoxon rank-sum text

The observations are arranged in ascending order and ranks from 1 to 12 are assigned.

The ranks of the observations belonging to the polluted are put in bold form.

Original Data	14.2	16.8	17.1	17.2	18.3	18.4	18.7	19.7	20	20.9	21.3	23
Rank	1	**2**	**3**	4	5	6	**7**	**8**	9	**10**	**11**	**12**

$R_1 = 2 + 3 + 7 + 8 + 10 + 11 + 12 = 53$

$R_2 = 1 + 4 + 5 + 6 + 9 = 25$

Also, $n_1 = 7 ; n_2 = 5$

U – Statistic,

$$U = n_1 n_2 + \frac{n_1(n_1+1)}{2} - R$$

$$= (7)(5) + \frac{7(8)}{2} - 53$$

$$= 35 + 28 - 53 = 10$$

Mean of U= $\mu_U = \frac{n_1 n_2}{2} = \frac{35}{2} = 17.5$

Variance of U= $\sigma_U^2 = \frac{n_1 n_2 (n_1 + n_2 + 1)}{12} = \frac{(7)(5)(13)}{12} = 37.916$

$$\sigma_U = 6.1575.$$

Here $n_2 = 5,$ so we use the statistic

$$Z = \frac{U - \mu_U}{\sigma_U} = \frac{10 - 17.5}{6.1575} = -1.218$$

$$|z| = 1.218.$$

Conclusion: Here $1.218 < 2.33$ So, we accept H_0.

17. The effectiveness of advertising for two rival products (Brand X and Brand Y) was compared. Market research at a local shopping centre was carried out, with the participants being shown adverts for two rival brands of coffee, which they then rated on the overall likelihood of them buying the product (out of 10, with 10 being "definitely going to buy the product"). Half of the participants gave ratings for one of the products, the other half gave ratings for the other product. Is there is a highly significant difference between the ratings given to each brand in terms of the likelihood of buying the product. Use U-test (take $\alpha = 0.01.$)

Brand X		**Brand Y**	
Participant	**Rating**	**Participant**	**Rating**
1	**3**	**1**	**9**
2	**4**	**2**	**7**
3	**2**	**3**	**5**
4	**6**	**4**	**10**
5	**2**	**5**	**6**
6	**5**	**6**	**8**

Solution: **[AU A/M 2024]**

Brand X			Brand Y		
Participant	Rating	Rank	Participant	Rating	Rank
1	3	3	1	9	11
2	4	4	2	7	9
3	2	1.5	3	5	5.5
4	6	7.5	4	10	12
5	2	1.5	5	6	7.5
6	5	5.5	6	8	10

Null hypothesis $H_0 : \mu_1 = \mu_2$.

Alternate hypothesis $H_1 : \mu_1 \neq \mu_2$.

$$\alpha = 0.05$$

Calculation:

$$n_1 = 6; n_2 = 6$$

$$R_1 = 3 + 4 + 1.5 + 7.5 + 1.5 + 5.5 = 23$$

$$R_2 = 11 + 9 + 5.5 + 12 + 7.5 + 10 = 55$$

U – Statistic,

$$U = R - \frac{n_1(n_1+1)}{2}$$

$$= 23 - \frac{6(6+1)}{2} = 2$$

At 0.05 level of significance.

$$\alpha_{0.05} = 1.96.$$

$$\mu_U = \frac{n_1 n_2}{2} = \frac{36}{2} = 18.$$

Variance of U,

$$\sigma_U^2 = \frac{n_1 n_2 (n_1 + n_2 + 1)}{12} = \frac{(36)(13)}{12} = 39$$

$$\sigma_U = 6.244.$$

Here $n_2 = 6$, so we use the statistic

$$Z = \frac{U - \mu_U}{\sigma_U} = \frac{2-18}{6.244} = -2.562$$

$$|z| = 2.562.$$

Conclusion:

Here 2.562 > 1.96.

Calculated value > Table value.

H_0 is rejected

$\therefore \quad \mu_1 \neq \mu_2$

<u>The H-test (or) Kruskal-Wallis test</u>

18. Use the Kruskal-Wallis test to test for differences in mean among the 3 samples. If $\alpha = 0.01$, what are your conclusions.

Sample I :	95	97	99	98	99	99	99	94	95	98
Sample II :	104	102	102	105	99	102	111	103	100	103
Sample III :	119	130	132	136	141	172	145	150	144	135

Solution:

Here, $n_1 = 10,\ n_2 = 10,\ n_3 = 10,\ n = n_1 + n_2 + n_3 = 10 + 10 + 10 = 30$

$k = 3,\ d.f = k - 1 = 3 - 1 = 2$

$H_0 : \mu_1 = \mu_2 = \mu_3$

$H_1 : \mu_1 \neq \mu_2 \neq \mu_3$

$\alpha = 0.01$

Criterion: Reject the null hypothesis $H > 9.21$ *from table*

$$\left.\begin{array}{r}\chi^2_{0.05} \\ k-1 = 2\ d.f\end{array}\right] = 9.21$$

Calculation:

Sample I	R_1	**Sample II**	R_2	**Sample III**	R_3
95	2.5	104	18	119	21
97	4	102	14	180	22

99	9	102	14	132	23
98	5.5	105	19	136	25
99	9	99	9	141	26
99	9	102	14	172	30
99	9	111	20	145	28
94	1	103	16.5	150	29
95	2.5	100	12	144	27
98	5.5	103	16.5	135	24
	$R_1 = 57$ $n_1 = 10$		$R_2 = 153$ $n_2 = 10$		$R_2 = 255$ $n_3 = 10$

$$H = \frac{12}{n(n+1)} \sum_{i=1}^{k} \frac{R_i^2}{n_i} - 3(n+1)$$

$$= \frac{12}{(30)(31)} \left[\frac{57^2}{10} + \frac{153^2}{10} + \frac{255^2}{10} \right] - 3(31) = 25.30$$

Conclusion:

Since $25.30 > 9.21$, we reject the null hypothesis H_0 and conclude that $\mu_1 \neq \mu_2 \neq \mu_3$

19. The following are the number of misprints counted on pages selected at random from a Newspaper

Day 1 :	4	10	2	6	4	12
Day 2 :	8	5	13	8	8	10
Day 3 :	7	9	11	2	14	7

Use H-test to test whether the samples come from same population.

Solution:

Here, $n_1 = 6,\ n_2 = 6\ , n_3 = 6,\ n = n_1 + n_2 + n_3 = 6+6+6 = 18$

$k = 3,\ d.f = k-1 = 3-1 = 2$

$H_0 : \mu_1 = \mu_2 = \mu_3$

$H_1 : \mu_1 \neq \mu_2 \neq \mu_3$

$\alpha = 0.05$

Criterion: Reject the null hypothesis $H > 5.991$ *from table*

$$\left.\begin{array}{c}\chi^2_{0.05}\\ k-1 = 2\ d.f\end{array}\right] = 5.991$$

Calculation:

Sample I	R_1	Sample II	R_2	Sample III	R_3
4	3.5	8	10	7	7.5
10	13.5	5	5	9	12
2	1.5	13	17	11	15
6	6	8	10	2	1.5
4	3.5	8	10	14	18
12	16	10	13.5	7	7.5
	$R_1 = 44$		$R_2 = 65.5$		$R_2 = 61.5$

$$H = \frac{12}{n(n+1)} \sum_{i=1}^{k} \frac{R_i^2}{n_i} - 3(n+1)$$

$$= \frac{12}{(18)(19)} \left[\frac{44^2}{6} + \frac{65.5^2}{6} + \frac{61.5^2}{6} \right] - 3(19) = 1.53$$

Conclusion:

Since $1.53 < 5.991$, we accept the null hypothesis H_0 and conclude that $\mu_1 = \mu_2 = \mu_3$

20. An experiment designed to compare three preventive methods against corrosion yielded the following maximum depths of pits (in thousands of an inch) in pieces of wire subjected to the respective treatments.

Method A :	77	54	67	74	71	66	-
Method B :	60	41	59	65	62	64	52
Method C :	49	52	69	47	56	-	-

Use the Kruskal-Wallis test at the 5% level of significance to test the null hypothesis that the three samples come from identical populations.

Solution:

Here, $n_1 = 6,\ n_2 = 7\ , n_3 = 5,\ n = n_1 + n_2 + n_3 = 6 + 7 + 5 = 18$

$k = 3,\ d.f = k - 1 = 3 - 1 = 2$

$H_0 : \mu_1 = \mu_2 = \mu_3$

$H_1 : \mu_1 \neq \mu_2 \neq \mu_3$

$\alpha = 0.05$

Criterion: Reject the null hypothesis $H > 5.991\ \ from\ \ table$

$$\left.\begin{matrix} \chi^2_{0.05} \\ k - 1 = 2\ d.f \end{matrix}\right] = 5.991$$

Calculation:

Method A	R_1	**Method B**	R_2	**Method C**	R_3
77	18	60	9	49	3
54	6	41	1	52	4.5
67	14	59	8	69	15
74	17	65	12	47	2
71	16	62	10	56	7
66	13	64	11		
		52	4.5		
	$R_1 = 84$ $n_1 = 6$		$R_2 = 55.$ $n_1 = 7$		$R_2 = 31.5$ $n_1 = 5$

$$H = \frac{12}{n(n+1)}\sum_{i=1}^{k}\frac{R_i^2}{n_i} - 3(n+1)$$

$$= \frac{12}{(18)(19)}\left[\frac{84^2}{6} + \frac{55.5^2}{7} + \frac{31.5^2}{5}\right] - 3(19) = 6.66$$

Conclusion:

Since $6.66 > 5.991$, we reject the null hypothesis H_0 and conclude that $\mu_1 \neq \mu_2 \neq \mu_3$

21. A company's trainees are randomly assigned to groups which are taught a certain industrial inspection procedure by 3-different methods. At the end of the instruction period they are tested for inspection performance quality. The following are their scores.

Method A :	80	83	79	85	90	68	-
Method B :	82	84	60	72	86	67	91
Method C :	93	65	77	78	88	-	-

Use H test to determine at the 0.05 level of significance whether the 3-methods are equally effective.

Solution:

Here, $n_1 = 6,\ n_2 = 7\ , n_3 = 5,\ n = n_1 + n_2 + n_3 = 6 + 7 + 5 = 18$

$k = 3,\ d.f = k - 1 = 3 - 1 = 2$

$H_0 : \mu_1 = \mu_2 = \mu_3$

$H_1 : \mu_1 \neq \mu_2 \neq \mu_3$

$\alpha = 0.05$

Criterion: Reject the null hypothesis $H > 5.991$ *from table*

$$\left.\begin{matrix}\chi^2_{0.05} \\ k-1 = 2\ d.f\end{matrix}\right] = 5.991$$

Calculation:

Method A	R_1	**Method B**	R_2	**Method C**	R_3
80	9	82	10	93	18
83	11	84	12	65	2
79	8	60	1	77	6
85	13	72	5	78	7
90	16	86	14	88	15

68	4	67 91	3 17		
	$R_1 = 61$ $n_1 = 6$		$R_2 = 62$ $n_1 = 7$		$R_2 = 48$ $n_1 = 5$

$$H = \frac{12}{n(n+1)} \sum_{i=1}^{k} \frac{R_i^2}{n_i} - 3(n+1)$$

$$= \frac{12}{(18)(19)} \left[\frac{61^2}{6} + \frac{62^2}{7} + \frac{48^2}{5} \right] - (3)(19)$$

$$= (0.035)(1630.1096) - (57)$$

$$= 57.2169 - 57$$

$$= 0.2169$$

Conclusion:

Since $0.2169 < 5.991$, we accept the null hypothesis H_0 and conclude that $\mu_1 = \mu_2 = \mu_3$

22. Four group of students were randomly assigned to taught with four different techniques and their achievement test scores were recorded. Are the distributions of test scores the same or do they differ in location? (take $\alpha = 0.01.$)

1	2	3	4
65	75	59	94
87	69	78	89
73	83	67	80
79	81	62	88

Solution: **[AU A/M 2024]**

Here $k = 4,\ n_1 = 4; n_2 = 4,\ n_3 = 4; n_4 = 4$

Null hypothesis $H_0 : \mu_1 = \mu_2 = \mu_3 = \mu_4$.

Alternate hypothesis $H_1 : \mu_1 \neq \mu_2 \neq \mu_3 \neq \mu_4$.

$\alpha = 0.05$; d.f = k-1=3.

Reject H_0 if H > 7.815.

$\chi^2_{0.05} = 7.815$ [See table values of χ^2]

1	2	3	4
65(3)	75(7)	59(1)	94(16)
87(13)	69(5)	78(8)	89(15)
73(6)	83(12)	67(4)	80(10)
79(9)	81(11)	62(2)	88(14)

$R_1 = 3+13+6+9 = 31$,

$R_2 = 7+5+12+11 = 35$

$R_3 = 1+8+4+2 = 15$

$R_4 = 16+15+10+14 = 55$

$n = n_1 + n_2 + n_3 + n_4 = 16$

H – Statistic is

$$H = \frac{12}{n(n+1)}\sum_{i=1}^{k}\frac{R_i^2}{n_i} - 3(n+1)$$

$$= \frac{12}{16(17)}\left\{\frac{31^2}{4}+\frac{35^2}{4}+\frac{15^2}{4}+\frac{55^2}{4}\right\} - 3(17)$$

$$= 0.044\{240.25+306.25+56.25+756.25\} - 51$$

$$H = 8.796$$

Conclusion:

H= 8.796 > 7.815 = $\chi^2_{0.05}$.

Calculated value > Table value at 5% level of significance.

H_0 is rejected.

23. The following are the final examination of marks of three groups of students who were taught computer by three different methods.

First method:	**94**	**88**	**91**	**74**	**87**	**97**	
Second method:	**85**	**82**	**79**	**84**	**61**	**72**	**80**
Third method:	**89**	**67**	**72**	**76**	**69**		

Use the H-test at the 0.05 level of significance to test the null hypothesis tht the three methods are equally effective.

Solution: **[AU N/D 2024]**

Here $k = 3,\ n_1 = 6; n_2 = 7,\ n_3 = 5$

Null hypothesis $H_0 : \mu_1 = \mu_2 = \mu_3$.

Alternate hypothesis $H_1 : \mu_1 \neq \mu_2 \neq \mu_3$.

$\alpha = 0.05$; d.f = k-1=2.

$\chi^2_{0.05} = 5.991$ [See table values of χ^2]

First method:	94(17)	88(14)	91(16)	74(6)	87(13)	97(18)	-
Second method:	85(12)	82(10)	79(8)	84(11)	61(1)	72(4.5)	80(9)
Third method:	89(15)	67(2)	72(4.5)	69(3)	-	-	-

$R_1 = 17+14+16+6+13+18 = 84$,

$R_2 = 12+10+8+11+1+4.5+9 = 55.5$

$R_3 = 15+2+4.5+7+3 = 31.5$

$n = n_1 + n_2 + n_3 = 84$

H – Statistic is

$$H = \frac{12}{n(n+1)} \sum_{i=1}^{k} \frac{R_i^2}{n_i} - 3(n+1)$$

$$= \frac{12}{18(19)} \left\{ \frac{84^2}{6} + \frac{55.5^2}{7} + \frac{31.5^2}{5} \right\} - 3(19)$$

$$= 0.035\{1176 + 440.03 + 198.45\} - 57$$

$$H = 6.5068$$

Conclusion:

$H = 6.5068 > 5.991 = \chi^2_{0.05}$.

Calculated value > Table value at 5% level of significance.

H_0 is rejected.

Hence the three methods are no equally effective.

<u>TESTS BASED ON RUNS-TEST OF RANDOMNESS</u>

24. In 30 tosses of a coin the following sequence of heads (H) and tails (T) is obtained

H TT H T HHH T HH TT H T H T HH T H TT H T HH T H T

(a) Determine the number of runs.

(b) Test at the 0.05 significance level whether the sequence is random.

Solution:

(a) Let us find the number of runs

$$\frac{1}{H}\ \frac{2}{TT}\ \frac{3}{H}\ \frac{4}{T}\ \frac{5}{HHH}\ \frac{6}{T}\ \frac{7}{HH}\ \frac{8}{TT}\ \frac{9}{H}\ \frac{10}{T}\ \frac{11}{H}\ \frac{12}{T}\ \frac{13}{HH}\ \frac{14}{T}\ \frac{15}{H}\ \frac{16}{TT}$$

$$\frac{17}{H}\ \frac{18}{T}\ \frac{19}{HH}\ \frac{20}{T}\ \frac{21}{H}\ \frac{22}{T}$$

Therefore the number of runs $R = 22$

(b) H_0 : The sequence is random.

H_1 : The sequence is not random.

$\alpha = 0.05$

Calculation:

Here, $n_1 = 16,\ n_2 = 14,\ R = 22$

$$\mu = \frac{2n_1n_2}{n_1+n_2}+1 = \frac{2(16)(14)}{16+14}+1 = 15.93$$

$$\sigma = \sqrt{\frac{2n_1n_2(2n_1n_2-n_1-n_2)}{(n_1+n_2)^2(n_1+n_2-1)}}$$

$$= \sqrt{\frac{2(16)(14)(2\times16\times14-16-14)}{(16+14)^2(16+14-1)}}$$

$$= \sqrt{7.175} = 2.679$$

$$\therefore\ Z = \frac{R-\mu}{\sigma}$$

$$Z = \frac{22-15.93}{2.679} = 2.27$$

Conclusion:

Since $2.27 > 1.96$, we reject the null hypothesis H_0 and conclude that the sequence is not random.

25. The following is the arrangement of defective, and non-defective, pieces produced in the given order by a certain machine:

$\underline{n\,n\,n\,n}\ \underline{d\,d\,d\,d}\ \ \underline{n\,n\,n\,n\,n\,n\,n\,n\,n\,n}\ \ \underline{d\,d}\ \ \underline{n\,n}\ \ \underline{d\,d\,d\,d}$

Test for randomness at the 0.01 level of significance.

Solution:

Let us find the number of runs

$\underline{n\,n\,n\,n}\ \underline{d\,d\,d\,d}\ \ \underline{n\,n\,n\,n\,n\,n\,n\,n\,n\,n}\ \ \underline{d\,d}\ \ \underline{n\,n}\ \ \underline{d\,d\,d\,d}$

Therefore the number of runs $R = 6$

H_0 : Arrangement is random.

H_1 : Arrangement is not random.

$\alpha = 0.01$

Criterion: Reject the null hypothesis $Z > 2.575$

Calculation:

Here, $n_1 = 10,\ n_2 = 17\ \ and\ \ R = 22$

$$\mu = \frac{2n_1n_2}{n_1+n_2}+1 = \frac{2(10)(17)}{10+17}+1 = 13.59$$

$$\sigma = \sqrt{\frac{2n_1n_2(2n_1n_2-n_1-n_2)}{(n_1+n_2)^2(n_1+n_2-1)}}$$

$$= \sqrt{\frac{2(10)(17)(2\times10\times17-10-17)}{(10+17)^2(10+17-1)}}$$

$$= 2.37$$

$$\therefore\ Z = \frac{R-\mu}{\sigma}$$

$$Z = \frac{6-13.59}{2.37} = -3.20 \quad \Rightarrow |Z| = 3.20$$

Conclusion:

Since $3.20 > 2.575$, we reject the null hypothesis H_0 and conclude that Arrangement is not random.

26. The production manager of a large undertaking randomly paid 10 visits to the work site in a month. The number of workers who reported late for duty was found to be 2, 4, 5, 1, 6, 3, 2, 1, 7 and 8 respectively. Use the run test for randomness at $\alpha = 0.05$ to check the claim of the production superintendent that on an average not more than 3 workers report late for duty.

Solution:

Let A be the average above 3

B be the average below 3

2	4	5	1	6	3	2	1	7	8
B	A	A	B	A	-	B	B	A	A

The above sequence can be written as $\frac{B}{1}\ \frac{AA}{2}\ \frac{B}{3}\ \frac{A}{4}\ \frac{BB}{5}\ \frac{AA}{6}$

$\therefore\quad R = 6\ \ (the\ number\ of\ runs)$

$n_1 = 5\ \ (the\ number\ of\ occurrences\ of\ 'A')$

$n_2 = 4\ \ (the\ number\ of\ occurrences\ of\ 'B')$

H_0 : The sample is randomly chosen.

H_1 : The sample is not randomly chosen.

$\alpha = 0.05$

Criterion: Reject the null hypothesis $Z > 1.96$

Calculation:

Here, $n_1 = 5,\ n_2 = 4\ and\ \ R = 6$

$$\mu = \frac{2n_1n_2}{n_1+n_2} + 1 = \frac{2(5)(4)}{5+4} + 1 = 5.4444$$

$$\sigma = \sqrt{\frac{2n_1n_2(2n_1n_2 - n_1 - n_2)}{(n_1+n_2)^2(n_1+n_2-1)}}$$

$$= \sqrt{\frac{2(5)(4)(2\times5\times4-5-4)}{(5+4)^2(5+4-1)}} = \sqrt{\frac{1240}{648}}$$

$$=1.3833$$

$$\therefore\ Z=\frac{R-\mu}{\sigma}\approx N(0,1)$$

$$Z=\frac{6-5.4444}{1.3833}=0.4016$$

Conclusion:

Since $0.4016>1.96$, we accept the null hypothesis H_0 and conclude that the sample is randomly chosen.

27. A technician is asked to analyze the results of 22items made in a preparation run. Each item has been measured and compared to engineering specifications. The order of acceptance $'a'$ and rejections $'r'$ is

$$aa\ \ rrr\ \ a\ \ rr\ \ aaaaa\ \ rr\ \ a\ \ rr\ aa\ \ r\ \ a$$

Determine whether it is a random sample or not.

Solution:

Given: $\underset{1}{\underline{aa}}\ \underset{2}{\underline{rrr}}\ \underset{2}{\underline{a}}\ \underset{4}{\underline{rr}}\ \underset{5}{\underline{aaaaa}}\ \underset{6}{\underline{rr}}\ \underset{7}{\underline{a}}\ \underset{8}{\underline{rr}}\ \underset{9}{\underline{aa}}\ \underset{10}{\underline{r}}\ \underset{11}{\underline{a}}$

Here, $n_1=12,\ n_2=10\ and\ \ R=11$

H_0 : The sample is randomly chosen.

H_1 : The sample is not randomly chosen.

$\alpha=0.05$

Criterion: Reject the null hypothesis $Z>1.96$

Calculation:

$$\mu=\frac{2n_1n_2}{n_1+n_2}+1=\frac{2(12)(10)}{12+10}+1=\frac{240}{22}+1=11.9091$$

$$\sigma=\sqrt{\frac{2n_1n_2\left(2n_1n_2-n_1-n_2\right)}{\left(n_1+n_2\right)^2\left(n_1+n_2-1\right)}}$$

$$=\sqrt{\frac{2(12)(10)\left(2\times12\times10-12-10\right)}{(12+10)^2(12+10-1)}}=\sqrt{\frac{52320}{10163}}$$

$$=2.2688$$

$$\therefore\ Z=\frac{R-\mu}{\sigma}\approx N(0,1)$$

$$Z=\frac{11-11.9091}{2.2688}=-0.4007$$

$$|Z|=|-0.4007|=0.4007$$

Conclusion:

Since $0.4007<1.96$, we accept the null hypothesis H_0 and conclude that the sample is randomly chosen.

28. The following are the prices in Rupees. 1Kg of a commodity from 2 random samples of shops from 2 cities A & B.

City A:	2.73	3.82	4.35	3.23	4.74	3.65	3.8	4.15	
	2.76	2.85	3.25	3.45	3.85	4.45	4.95	3.95	4.72
City B:	3.75	5.37	4.78	3.69	4.75	4.85	6.0	4.8	4.9
	3.84	3.9	4.8	5.23	6.1	3.6	3.83		

Apply the run test to examine whether the distribution of prices of commodity in the two cities is the same.

Solution:

H_0 : The distribution of prices of commodity in the 2 cities is same.

H_0 : The distribution of prices of commodity in the 2 cities is not same.

$Level\ of\ significance\ \ \alpha = 5\% = 0.05$

Criterion: Reject the null hypothesis $Z > 1.96$

Calculation:

Let us combine the observation from both cities and arrange them in ascending order. Now assign the letter a to city A and letter B to city B as follows:

$$\underbrace{\begin{matrix}2.73 & 2.76 & 2.85 & 3.23 & 3.25 & 3.45\\ A & A & A & A & A & A\end{matrix}}_{1}\ \underbrace{\begin{matrix}3.6\\ B\end{matrix}}_{2}\ \underbrace{\begin{matrix}3.65\\ A\end{matrix}}_{3}\ \underbrace{\begin{matrix}3.69 & 3.75\\ B & B\end{matrix}}_{4}\ \underbrace{\begin{matrix}3.8 & 3.82\\ A & A\end{matrix}}_{5}\ \underbrace{\begin{matrix}3.83 & 3.84\\ B & B\end{matrix}}_{6}\ \underbrace{\begin{matrix}3.85\\ A\end{matrix}}_{7}$$

$$\underbrace{\begin{matrix}3.9\\ B\end{matrix}}_{8}\ \underbrace{\begin{matrix}3.95 & 4.15 & 4.35 & 4.45 & 4.72 & 4.74\\ A & A & A & A & A & A\end{matrix}}_{9}\ \underbrace{\begin{matrix}4.75 & 4.78 & 4.8 & 4.85 & 4.9\\ B & B & B & B & B\end{matrix}}_{10}\ \underbrace{\begin{matrix}4.95\\ A\end{matrix}}_{11}\ \underbrace{\begin{matrix}5.23 & 5.27 & 6.0 & 6.1\\ B & B & B & B\end{matrix}}_{12}$$

Here, $n_1 = 17,\ n_2 = 16\ and\ \ R = 12$

$$\mu = \frac{2n_1n_2}{n_1+n_2} + 1 = \frac{2(17)(16)}{17+16} + 1 = 17.485$$

$$\sigma = \sqrt{\frac{2n_1n_2(2n_1n_2 - n_1 - n_2)}{(n_1+n_2)^2(n_1+n_2-1)}}$$

$$= \sqrt{\frac{2(17)(16)(2\times17\times16-17-16)}{(17+16)^2(17+16-1)}} = \sqrt{7.977}$$

$$= 2.824$$

$$\therefore\ Z = \frac{R-\mu}{\sigma} \approx N(0,1)$$

$$Z = \frac{12 - 17.485}{2.824} = \frac{5.485}{2.824} = 1.94$$

$$|Z| = |-0.4007| = 0.4007$$

Conclusion:

Since $0.4007 < 1.96$, we accept the null hypothesis H_0 and conclude that the sample is randomly chosen.

29. Twenty three individuals were sampled as to whether they or did not like a product indicated by Y and N respectively, the resulting sample is shown by the following sequence:

YY NNNN YYY N Y NN Y NNNN YYY NN

(i) Determine the number of runs, V.

(ii) Test at 0.05 significance level whether the responses are random.

Solution: **[AU N/D 2024]**

H_0 : Arrangement is random

H_1 : Arrangement is not random

$\alpha = 0.05$

Reject H_0 if $Z > 1.96$

Calculation:

YY	NNNN	YYY	N	Y	NN	Y	NNNN	YYY	NN
1	2	3	4	5	6	7	8	9	10

$R = 10, \quad n_1 = 10, \quad n_2 = 13$

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

$$where \quad \mu = \frac{2n_1 n_2}{n_1 + n_2} + 1 = \frac{2(10)(13)}{23} + 1 = 12.3043$$

$$\sigma = \sqrt{\frac{2n_1 n_2 (2n_1 n_2 - n_1 - n_2)}{(n_1 + n_2)^2 (n_1 + n_2 - 1)}} = \sqrt{\frac{2(10)(13)(2.10.13 - 10 - 13)}{(23)^2 (22)}}$$

$$= \sqrt{\frac{(260)(237)}{11638}} = \sqrt{5.2947}$$

$$= 2.3010$$

$$Z = \frac{R - \mu}{\sigma} = \frac{10 - 12.3043}{2.3010} = -1.0014$$

$$|Z| = 1.0014$$

Conclusion:

Here, $1.0014 < 1.96$ we accept H_0 and we may conclude that sample is randomly chosen.

30. The following are the speeds (in kilometer per hour) at which every fifth passenger car was timed at a certain checkpoint: 46, 58, 60, 56, 70, 66, 48, 48, 54, 62, 41, 39, 52, 45, 62, 53, 69, 65, 65, 67, 76, 52, 52, 59, 59, 67, 51, 46, 61, 40, 43, 42, 77, 67, 63, 59, 63, 63, 72, 57, 59, 42, 56, 47, 62, 67, 70, 63, 66, 69, and 73.
Test the null hypothesis of randomness at the 0.05 level of significance. (Given median speed = 59.5 km per hour).
Solution: **[AU N/D 2024]**

$H_0:$ Sample is random

$H_1:$ Sample is random

$\alpha = 5\%, \quad Z_\alpha = 1.96$

Median speed = 59.5

Arrangements of $a's \ and \ b's$

$b\ b\ |a|\ b\ |aa|\ \ b\ b\ \ |a|\ b\ b\ b\ b\ \ |a|\ b\ \ |aaaaa|\ b\ b\ b\ b\ \ |a|\ b\ b\ \ |a|\ b\ b\ b$

$|aaa|\ b\ \ |aaa|\ \ b\ b\ b\ b\ b\ |aaaaaaa|$

$n_1 = 25, \quad n_2 = 25$

$\therefore \quad R = no.\ of \ \ runs = 20$

$$\mu = \frac{2n_1n_2}{n_1+n_2} + 1 = \frac{2(25)(25)}{50} + 1 = 26$$

$$\sigma = \sqrt{\frac{2n_1n_2(2n_1n_2 - n_1 - n_2)}{(n_1+n_2)^2(n_1+n_2-1)}} = \sqrt{\frac{2(25)(25)(2.25.25 - 25 - 125)}{(50)^2(49)}}$$

$$= \sqrt{\frac{15000}{1225}} = 3.4992$$

$$Z = \frac{R-\mu}{\sigma} = \frac{20-26}{3.4992} = -1.714$$

$$|Z| = 1.714$$

Conclusion:

Here, $1.714 < 1.96$ we accept H_0 at 5% level of significance

we may conclude that sample is randomly chosen

KOLMOGOROV-SMIRNOV GOODNESS OF FIT TEST

31. Test whether the following numbers 0.44, 0.81, 0.14, 0.05, and 0.93 are uniformly distributed using Kolmogorov-Smirnov test.
Solution:

Given numbers are $0.44, 0.81, 0.14, 0.05, 0.93$

Here $N = 5,$ we arrange the numbers must be ranked from smallest to largest. $N = 5$

Prepare the table for Kolmogorov-Smirnov test as given below:

	i	1	2	3	4	5
	$R_{(i)}$	0.05	0.14	0.44	0.81	0.93
	i/N	0.20	0.40	0.60	0.80	1.00
D^+	$i/N - R_{(i)}$	0.15	0.26	0.16	-	0.07
D^-	$R_{(i)} - (i-1)/N$	0.05	-	0.04	0.21	0.13

The second row from top lists the numbers from smallest $\left(R_{(1)}\right)$ to largest$\left(R_{(5)}\right)$,

The computations for D^+, namely $i/N - R_{(i)}$ and for D^-, namely $R_{(i)} - (i-1)/N$ are performed.

The statistics are computed as $D^+ = 0.26$ and $D^- = 0.21$

Therefore, $D = \max(0.26.0.21) = 0.26$

The critical value of D for $\alpha = 0.05$ and $N = 5$ is 0.565 $(given)$

Since, the computed value, 0.26 is less than the tabulated critical value, 0.565, the hypothesis of no difference between the distribution of the generated numbers and the uniform distribution is not rejected.

32. Below is the table of observed frequencies along with the frequency to the observed under a normal distribution.

(a) Calculate the K-S statistic.

(b) Can we conclude that this distribution does in fact follow a normal distribution? Use 0.10 level of significance.

Test Score	**51-60**	**61-70**	**71-80**	**81-90**	**91-100**
Observed Frequency	**30**	**100**	**440**	**500**	**130**
Expected Frequency	**40**	**170**	**500**	**390**	**100**

Solution:

H_0 : This distribution follows a normal distribution

H_1 : This distribution follows not a normal distribution

$\alpha = 0.10$

To Calculate : $D_n = \max\left|F_e - F_0\right|$

Calculation of K-S Statistic

Observed frequency	Observed Cumulative frequency	Observed relative Frequency F_0	Expected frequency	Expected Cumulative Frequency	Expected relative Frequency F_e	$D_n = \lvert F_e - F_0 \rvert$
30	30	$\frac{30}{1200} = 0.025$	40	40	$\frac{40}{1200} = 0.033$	**0**.008
100	130	$\frac{130}{1200} = 0.108$	170	210	$\frac{210}{1200} = 0.175$	0.067
440	570	$\frac{570}{1200} = 0.475$	500	710	$\frac{710}{1200} = 0.592$	0.117
500	1070	$\frac{1070}{1200} = 0.891$	390	1100	$\frac{1100}{1200} = 0.920$	0.029
130	1200	$\frac{1200}{1200} = 1$	100	1200	$\frac{1200}{1200} = 1$	0

K-S statistic is

$$D_n = \max\lvert F_e - F_0 \rvert = 0.117$$

The tabulated value of D_n for $n = 5$ and $\alpha = 0.05$ is 0.510 (from table).

Since, the table value of $D_n = 0.510$ is greater than the calculated value of $D_n = 0.117$,

we accept the null hypothesis.

i.e., the distribution follows a normal distribution.

ANNA UNIVERSITTY QUESTIONS

SIGN TEST TYPE: 1-A $[n \leq 20]$

1. The following data constitute a random sample of 15 measurements of the octane rating of a certain kind of gasoline:

99.0	102.3	99.8	100.5	99.7	96.2	99.1	102.5
103.3	97.4	100.4	98.9	98.3	98.0	101.6	

Test the null hypothesis $\mu = 98.0$ against the alternative hypothesis $\mu > 98.0$ at the 0.01 level of significance. [pg. no. 5]

2. In a factory, 20 observations of the factors that could heat up a conveyor belt yielded the following results: 0.36, 0.41, 0.25, 0.34, 0.28, 0.26, 0.39, 0.28, 0.40, 0.26, 0.35, 0.38, 0.29, 0.42, 0.37, 0.37, 0.39, 0.32, 0.29 and 0.36. Use the sign test at the 0.01 level of significance to test the null hypothesis $\mu = 0.34$ against the alternative hypothesis $\mu \neq 0.34$ [pg. no. 5]

SIGN TEST TYPE: 1-B $[n > 20]$

3. The time sheet of a factory showed the following sample data (in hours) on the time spent by a worker operating a hydraulic gear lift : 1.0, 0.8, 0.5, 0.9, 1.2, 0.9, 1.4, 10, 1.3, 0.8, 1.5, 1.2, 1.9, 1.1, 0.7, 0.8, 1.1, 1.2, 1.5, 1.1, 1.8, 0.5, 0.8, 0.9 and 1.6. Use the sign test at the 0.05 level of significance to test the null hypothesis $\mu = 1.1$ against the alternative hypothesis $\mu > 1.1$ [pg. no. 6]

4. The following are the sizes of particles of cement dust (given to the nearest hundredth of a Micron) in a cement factory.

16.12	10.48	11.12	16.18	18.13	19.10	13.21	10.12
21.18	15.12	10.11	13.31	18.61	11.43	18.26	13.77
13.24	12.16	17.19	11.36	12.53	13.25	10.67	15.45
14.28	14.32	15.18	14.21	10.20	15.64	11.68	18.76
19.32	17.50	11.46	20.59	16.38	21.42	16.27	21.30
16.12	10.55	11.49	15.48	11.62	13.54	13.69	16.72
15.11	14.33	17.23	17.22	19.37	10.41	18.28	19.29
21.23	12.56	12.57	11.60	15.24	21.65	20.70	11.44
12.22	19.34	20.35	19.47	21.63	19.40	19.75	21.71
15.19	18.51	10.58	13.52	11.39	13.66	21.73	11.74

Which pertained to the particle size of cement dust in a factory producing cement, use the sign? test at the 0.05 level of significance to test the null hypothesis $\mu = 15.13$ hundredth of a micron against the alternative hypothesis $\mu < 15.13$ hundredth of a micron. [pg. no. 7]

TYPE: 2

5. The following are the average weekly losses of worker-hours due to accidents in 10 industrial plants before and after a certain safety program was into operation

Before	45	73	46	124	33	57	83	34	26	17
After	36	60	44	119	35	51	77	29	24	11

Use the 0.05 level of significance to test whether the safety program is effective. [pg. no. 8]

6. Use the sign test t see if there is a difference between the number of days required to collect an account receivable before and after a new collection policy. Use the 0.05 significance level.

Before:	33	36	41	32	39	47	34	29	32	34	40	42	33	36	27
After:	35	29	38	34	37	47	36	32	30	34	41	38	37	35	28

[AU N/D 2024] [pg. no. 9]

7. A consumer panel includes 14 individuals. It is asked to rate two brands of co-co cola according to a point evaluation system based on several criteria. The table gives below reports the points assigned. Test the null hypothesis that there is no difference in the level of ratings for the two brands of cola at 5% level of significance using the sign test. [pg. no. 9]

Panel member	1	2	3	4	5	6	7	8	9	10	11	12	13	14
Brand I:	20	24	28	24	20	29	19	27	20	30	18	28	26	24
Brand II:	16	26	18	17	20	21	23	22	23	20	18	21	17	26

THE SIGNED – RANK TEST

8. The following are 15 measurements of the octane rating of a certain kind of gasoline. 97.5, 95.2, 97.3, 96.0, 96.8, 100.3, 97.4, 95.3, 93.2, 99.1, 96.1, 97.6, 98.2, 98.5 and 94.9. Use the signed-rank Test at the 0.05 level of significance to test the mean octane rating of the given kind of gasoline is 98.5. [pg. no. 10]

9. Drop in diastolic blood pressure (in mm mercury)

Drug D_1	10	16	10	4	2	14	4
Drug D_2	33	34	41	36	42	42	32

Test whether there is any difference in the effectiveness of drugs at $\alpha = 0.05$. Using Wilcoxon's signed rank test. (Table value for $n = 7$ is 2) [pg. no. 11]

10. In a market research, it was decided to examine the effect of brand name on quality perception.16 subjects are recruited for the purpose and are asked to taste and compare two sample of product on a set of scale items judged to be ordinal. The following data are obtained (table value for $n = 15$ is 25) [pg. no. 12]

Pair	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16
Item A:	73	43	47	52	58	47	52	58	38	61	56	56	34	55	65	75
Item B:	51	41	43	41	47	32	24	58	43	63	52	57	44	57	40	68

Test the hypothesis, using Wilcoxon matched-pairs test, that there is no difference between the perceived qualities of the two samples. Use 5% level of significance.

11. The following are the weights in pounds before and after of 16 persons who stayed on a certain reducing diet for four weeks.

Before:	147.0	183.5	232.1	161.6	197.5	206.3	177.0	215.4	147.7	208.1	166.8	131.9	150.3	197.2	159.8	171.7
After:	137.9	176.2	219.0	163.8	193.5	201.4	180.6	203.2	149.0	195.4	158.5	134.4	149.3	189.1	159.1	173.2

Use the signed rank test to test at the 0.05 level of significance whether the weight reducing diet is effective. [pg. no. 14]

12. A manufacturer of electric irons, wishing to test the accuracy of the thermostat control at the 500^o F setting, instructs a test engineer to obtain actual temperature at that setting for 15 irons using a thermocouple. The resulting measurements are as follows:

494.6	510.8	487.5	493.2	502.6	485.0	495.9	498.2
501.6	497.3	492.0	504.3	499.2	493.5	505.8	

The engineer believes it is reasonable to assume that a temperature deviation from 500^o of any particular magnitude is just as likely to be positive as negative (the assumption of symmetry) but wants to protect against possible non normality of the actual temperature distribution. Use signed-rank test to see whether the data strongly suggests incorrect calibration of the iron.

[AU A/M 2024] [pg. no. 16]

RANK-SUM TESTS – THE U-TEST

13. The nicotine contents of two brands of cigarettes, measured in milligrams, was found to be as follows:

Brand A	2.1	4.0	6.3	5.4	4.8	3.7	6.1	3.3		
Brand B	4.1	0.6	3.1	2.5	4.0	6.2	1.6	2.2	1.9	5.4

Test the hypothesis, at the 0.05 level of significance, that the average nicotine contents of the two brands are equal against the alternative that they are unequal. [pg. no. 17]

14. Two classes of students are tested using a certain competitive exam. The scores off a sample of students from each class is given below:

Class A	45	44	47	48	55	53	55	63		
Class B	65	67	77	65	56	67	78	55	66	65	58

Use Mann Whitney-U test to whether both classes have similar scholastic levels. [pg. no. 18]

15. Independent random samples of ten day students and ten evening students at a University Showed the following age distributions. We want to use the Mann-Whitney-Wilcoxon test to determine if there is a significant difference in the age distribution of the two groups.

Day	26	18	25	27	19	30	34	21	33	31
Evening	32	24	23	30	40	41	42	39	45	35

Use Mann Whitney-U test to whether both classes have similar scholastic levels.

(i) Compute the sum of the ranks (T) for the day students.

(ii) Compute the mean μ_T

(iii) Compute σ_T

(iv) Use $\alpha = 0.05$ and test for any significant differences in the age distribution of the two populations. [pg. no. 19]

16. The urinary fluoride concentration (pairs per million) was measured both for a sample of livestock grazing in an area previously exposed to fluoride pollution and for a similar sample grazing in an unpolluted region:

Polluted	21.3	18.7	23.0	17.1	16.8	20.9	19.7
Unpolluted	14.2	18.3	17.2	18.4	20.0		

Does the data indicate strongly that the true average fluoride concentration for livestock grazing in the polluted region is larger than for the unpolluted region? Use the Wilcoxon rank-sum test at level $\alpha = 0.01$. [AU A/M 2024][pg. no. 20]

17. The effectiveness of advertising for two rival products (Brand X and Brand Y) was compared. Market research at a local shopping centre was carried out, with the participants being shown adverts for two rival brands of coffee, which they then rated on the overall likelihood of them buying the product (out of 10, with 10 being "definitely going to buy the product"). Half of the participants gave ratings for one of the products, the other half gave ratings for the other product. Is there is a highly significant difference between the ratings given to each brand in terms of the likelihood of buying the product. Use U-test (take $\alpha = 0.01$.)

[AU A/M 2024] [pg. no. 22]

Brand X		Brand Y	
Participant	Rating	Participant	Rating
1	3	1	9
2	4	2	7
3	2	3	5
4	6	4	10
5	2	5	6
6	5	6	8

<u>The H-test (or) Kruskal-Wallis test</u>

18. Use the Kruskal-Wallis test to test for differences in mean among the 3 samples. If $\alpha = 0.01$, what are your conclusions. [pg. no. 23]

Sample I :	95	97	99	98	99	99	99	94	95	98
Sample II :	104	102	102	105	99	102	111	103	100	103
Sample III :	119	130	132	136	141	172	145	150	144	135

19. The following are the number of misprints counted on pages selected at random from a Newspaper

Day 1 :	4	10	2	6	4	12
Day 2 :	8	5	13	8	8	10
Day 3 :	7	9	11	2	14	7

Use H-test to test whether the samples come from same population. [pg. no. 24]

20. An experiment designed to compare three preventive methods against corrosion yielded the following maximum depths of pits (in thousands of an inch) in pieces of wire subjected to the respective treatments.

Method A :	77	54	67	74	71	66	-
Method B :	60	41	59	65	62	64	52
Method C :	49	52	69	47	56	-	-

Use the Kruskal-Wallis test at the 5% level of significance to test the null hypothesis that the three samples come from identical populations. [pg. no. 26]

21. A company's trainees are randomly assigned to groups which are taught a certain industrial inspection procedure by 3-different methods. At the end of the instruction period they are tested for inspection performance quality. The following are their scores.

Method A :	80	83	79	85	90	68	-
Method B :	82	84	60	72	86	67	91
Method C :	93	65	77	78	88	-	-

Use H test to determine at the 0.05 level of significance whether the 3-methods are equally effective. [pg. no. 27]

22. Four group of students were randomly assigned to taught with four different techniques and their achievement test scores were recorded. Are the distributions of test scores the same or do they differ in location? (take $\alpha = 0.01.$) [AU A/M 2024] [pg. no. 28]

1	2	3	4
65	75	59	94
87	69	78	89
73	83	67	80
79	81	62	88

23. The following are the final examination of marks of three groups of students who were taught computer by three different methods.

First method:	94	88	91	74	87	97	
Second method:	85	82	79	84	61	72	80
Third method:	89	67	72	76	69		

Use the H-test at the 0.05 level of significance to test the null hypothesis that the three methods are equally effective. [AU N/D 2024][pg. no. 29]

TESTS BASED ON RUNS-TEST OF RANDOMNESS

24. In 30 tosses of a coin the following sequence of heads (H) and tails (T) is obtained

H TT H T HHH T HH TT H T H T HH T H TT H T HH T H T

(a) Determine the number of runs.

(b) Test at the 0.05 significance level whether the sequence is random. [pg. no. 30]

25. The following is the arrangement of defective, and non-defective, pieces produced in the given order by a certain machine:

$\underline{nnnn}\ \underline{dddd}\ \ \underline{nnnnnnnnnn}\ \ \underline{dd}\ \ \underline{nn}\ \ \underline{dddd}$

Test for randomness at the 0.01 level of significance. [pg. no. 31]

26. The production manager of a large undertaking randomly paid 10 visits to the work site in a month. The number of workers who reported late for duty was found to be 2, 4, 5, 1, 6, 3, 2, 1, 7 and 8 respectively. Use the run test for randomness at $\alpha = 0.05$ to check the claim of the production superintendent that on an average not more than 3 workers report late for duty.[pg. no. 32]

27. A technician is asked to analyze the results of 22items made in a preparation run. Each item has been measured and compared to engineering specifications. The order of acceptance $'a'$ and rejections $'r'$ is

$aa\ \ rrr\ \ a\ \ rr\ \ aaaaa\ \ rr\ \ a\ \ rr\ aa\ \ r\ \ a$

Determine whether it is a random sample or not. [pg. no. 33]

28. The following are the prices in Rupees. 1Kg of a commodity from 2 random samples of shops from 2 cities A & B.

City A:	2.73	3.82	4.35	3.23	4.74	3.65	3.8	4.15	
	2.76	2.85	3.25	3.45	3.85	4.45	4.95	3.95	4.72
City B:	3.75	5.37	4.78	3.69	4.75	4.85	6.0	4.8	4.9
	3.84	3.9	4.8	5.23	6.1	3.6	3.83		

Apply the run test to examine whether the distribution of prices of commodity in the two cities is the same. [pg. no. 34]

29. Twenty three individuals were sampled as to whether they or did not like a product indicated by Y and N respectively, the resulting sample is shown by the following sequence:

YY NNNN YYY N Y NN Y NNNN YYY NN

(i) Determine the number of runs, V.

(ii) Test at 0.05 significance level whether the responses are random.

[AU N/D 2024][pg. no. 35]

30. The following are the speeds (in kilometer per hour) at which every fifth passenger car was timed at a certain checkpoint: 46, 58, 60, 56, 70, 66, 48, 48, 54, 62, 41, 39, 52, 45, 62, 53, 69, 65, 65, 67, 76, 52, 52, 59, 59, 67, 51, 46, 61, 40, 43, 42, 77, 67, 63, 59, 63, 63, 72, 57, 59, 42, 56, 47, 62, 67, 70, 63, 66, 69, and 73. Test the null hypothesis of randomness at the 0.05 level of significance. (Given median speed = 59.5 km per hour). [AU N/D 2024][pg. no. 36]

KOLMOGOROV-SMIRNOV GOODNESS OF FIT TEST

31. Test whether the following numbers 0.44, 0.81, 0.14, 0.05, and 0.93 are uniformly distributed using Kolmogorov-Smirnov test. [pg. no. 36]

32. Below is the table of observed frequencies along with the frequency to the observed under a normal distribution.
    (a) Calculate the K-S statistic.
    (b) Can we conclude that this distribution does in fact follow a normal distribution? Use 0.10 level of significance. [pg. no. 37]

Test Score	51-60	61-70	71-80	81-90	91-100
Observed Frequency	30	100	440	500	130
Expected Frequency	40	170	500	390	100

PART-A

1. Define Variables.

Solution:

Variables are those quality characteristics of a product or item which are measurable and can be expressed in their respective units of measurements.

eg: Weights of packed materials, tensile strength of rods, diameter of wires, operating temp in degree, diameter of ball- bearing, life of e-bulb, settling time of concrete etc.

2. Define Attributes.

Solution:

Attributes are characteristics of product which are non-measurable. Such characteristics can be felt by their presence or absence.

eg: Blemishes in piece of cloth, spot in unit of same size threads in white cloth, bubbles of air in windscreen fault found during inspection.

3. Define Defective.

Solution:

Defective is an item or unit contain certain one or more error flow or faults, A defective is a individual flow, error or fault.

4. What are the Types of control charts? **[AU N/D 2020, 2024]**

Solution:

(i) Control charts for variables:

$\overline{X}$ chart and R – chart

(ii) Control charts for attributes

P- chart and np – chart , C – chart.

5. Write the Control chart for Attributes.

Solution:

(i) P – chart for proportion of defectives for sample of varying size.

(ii) np – chart for number of defectives for constant sample size.

(iii) C- chart for number of defectives per unit.

6. Define Control chart. **[AU N/D 2024]**

Solution:

The statistical tool applied in process control is the control chart.

7. Specify the Control chart consists of three horizontal lines.

Solution:

(i) A central line(CL) to indicate the desired standard or level of the process

(ii) Upper control limit(UCL)

(iii) Lower control limit(LCL).

8. Define Acceptance inspection (or) Sampling inspection.

Solution:

The inspection of materials to determine their acceptability, whether they be in raw, semi-finished or finished state. This is known as acceptance inspection or sampling inspection.

9. Write the Uses of C – chart.

Solution:

(i) The number of defects in a galvanized sheet or painted plate or enameled surface of given area.

(ii) The number of defects of all types in aircrafts, sub-assemblies or final assembly.
(iii) Number of defects observed in a roll of paper bale of cloth, sheet of photographic film.

10. What are the Advantages of statistical quality control?

Solution:

(i) Reduction in costs.
(ii) Easy to apply.
(iii) Greater efficiency.
(iv) Early detection of faults.
(v) Adherence to specifications
(vi) The only course.
(vii) To determine the effect of change in process.
(viii) Statistical quality control ensures overall coordination.

11. Define Acceptance sampling.

Solution:

The control charts described above cannot be applied to all types of problems. They are useful only for the regulation of the manufacturing process. Another important field of quality control is acceptance sampling.

12. What are the Role of Acceptance sampling?

Solution:

(i) Acceptance sampling is much less expensive than 100 percent inspection.
(ii) In many cases it provides better outgoing quality.
(iii) In modern manufacturing plants, acceptance sampling is used for evaluating the acceptability of incoming lots of raw materials and parts at various stages of manufacture and final inspection of finished product.
(iv) Where quality can be tested only by destroying items as in determining the strength of glass containers, 100 percent inspection is out of the question and sampling must be used.

13. Write down the Types of Acceptance sampling plans.

Solution:

(i) Single sampling plan.
(ii) Double sampling plan.

14. Define Single sampling plan.

Solution:

When the decision whether to accept a lot or reject a lot is made on the basis of only one sample, the acceptance plan is described as a single sampling plan.

15. Define Double sampling plan.

Solution:

Double sampling involves the possibility of putting off the decision on the lot until a second sample has been taken. A lot may be accepted at once if the first sample is good enough or rejected at once if the first sample is bad enough.

16. A garment was sampled on 10 consecutive hours of production. The number of defects found per garment is given below:
Defects: 5,1,7,0,2,3,4,0,3,2. Compute upper and lower control limits for monitoring number of defects. [AU A/M 2019]

Solution:

$$\bar{C} = 2.7 \ central \ line$$

$$UCL = \bar{C} + 3\sqrt{\bar{C}} = 7.6295$$

$$LCL = \bar{C} - 3\sqrt{\bar{C}} = -2.2295$$

17. Define tolerance limits: **[A. U A/M 2019]**

Solution:

Tolerance limits $\left(\bar{X} - kS, \bar{X} + kS\right)$

18. Explain upper control limit and lower control limit in quality control. **[A. U N/D 2019]**

Solution:

UCL : The line at $\bar{x} + 3\sigma$ is known as upper limit control .

LCL : The line at $\bar{x} - 3\sigma$ is known as lower limit control**.**

19. Compare c-chart with p-chart. **[A. U N/D 2019]**

Solution:

c-chart	p-chart
Actual number defects are plotted	The values of p fraction defective is plotted.
Units of same size is selected	Sample size may be consistent or even variable from sample to sample.

20. Write down the formula for UCL and LCL for np-charts. **[AU N/D 2020]**

Solution:

$$UCL = n\bar{p} + 3\sqrt{n\bar{p}\left(1 - \bar{p}\right)} \quad \text{(or)} \quad UCL = n\left[\bar{p} + 3\sqrt{\frac{\bar{p}\left(1 - \bar{p}\right)}{n}}\right]$$

$$LCL = n\bar{p} - 3\sqrt{n\bar{p}\left(1 - \bar{p}\right)} \quad \text{(or)} \quad LCL = n\left[\bar{p} - 3\sqrt{\frac{\bar{p}\left(1 - \bar{p}\right)}{n}}\right]$$

21. What is the probabilistic base behind *c* chart? **[AU N/D 2021]**

Solution:

The Poisson distribution is the basis for the chart and requires the following assumptions: The number of opportunities or potential locations for nonconformities is very large. The probability of nonconformity at any location is small and constant.

22. Which control chart is not influenced by the sample size? What probability distribution does it follows?

Solution: **[AU A/M 2024]**

C-chart : Underlying distribution Poisson.

23. What is the advantage of using control charts for attributes compared to variable control charts?

Solution: **[AU A/M 2024]**

If we've to check only whether the item is conform to the quality characteristic, then attributes chart will be beneficial.

24. What is the difference between control limits and tolerance limits? **[AU N/D 2021]**

Solution:

Control Limits	Tolerance Limits
Control limits, also known as natural process limits, are horizontal lines drawn on a statistical process control chart	Tolerance limits consist of the the upper and lower limits of a particular environmental condition which allows a certain species to survive
The **control limits** of your control chart represent your process variation and help indicate when your process is out of control.	Tolerance limits are limits that include a specific proportion of the population at a given confidence level.

25. When do you say that a process is out of control? **[AU N/D 2021,2024]**

Solution:

When one or more plots points lie outside the control lines, it is to be considered as a danger signal, indicating the variations between samples are caused by assignable causes and the process is out of control and that necessary corrective action should be taken at once.

PART-B

Problems based on $\overline{x}$ and R- chart

1. Given below are the values of sample mean $\overline{x}$ and sample range R for 10 samples, each of size 5. Draw the appropriate mean and range charts and comment on the state of control of the process.

Sample no.	1	2	3	4	5	6	7	8	9	10
Mean $\overline{x}_i$	43	49	37	44	45	37	51	46	43	47
Range R	5	6	5	7	7	4	8	6	4	6

Solution:

$$\overline{X} = \frac{1}{N}\sum \overline{x_i} \quad = \frac{1}{10}(43+49+37+\ldots 47)\overline{X} = 44.2$$

$$\overline{R} = \frac{1}{N}\sum \overline{R_i} \quad = \frac{1}{10}(5+6+5+\ldots+6)\ \overline{R} = 5.8$$

From the table of control chart, for sample size n=5, we have

$$A_2 = 0.577,\ D \ = 0 \ and \ D_4 = 2.115$$

(i) Control limits for $\overline{X}$ chart:

CL (Central line) = $\overline{X} = 44.2$

LCL = $\overline{X} - A_2\overline{R}$ = 44.2 - (0.577)(5.8) = 40.8534 ≈ 40.85

UCL = $\overline{X} + A_2\overline{R}$ = 44.2+ (0.577)(5.8) = 47.5466 ≈ 47.55

Conclusion:

Since 2nd,3rd 6th and 7th sample means fall outside the control limits the statistical process is out of control according to $\overline{X}$ chart.

(ii) Control limits for R chart:

CL (Central line) = $\overline{R} = 5.8$

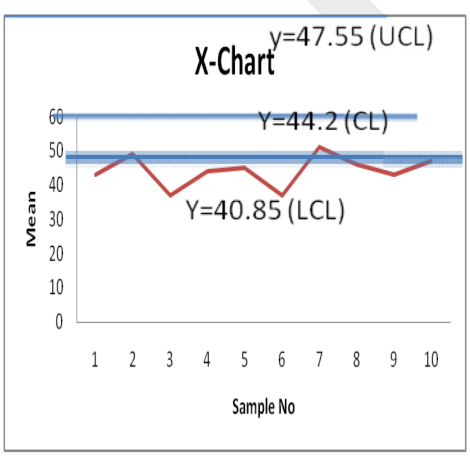
LCL = $D_3\overline{R} = 0$;

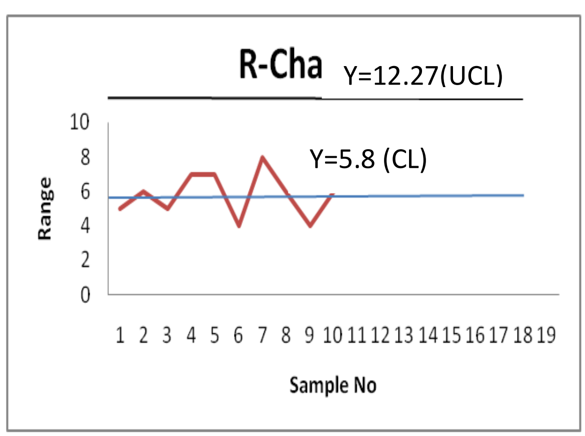
UCL = $D_4\overline{R}$ = (2.115)(5.8) = 12.267 ≈ 12.27

Conclusion:

Since all the sample mean fall within the control limits the statistical process is under control according to R chart.

Inference: From both $\overline{X}$ and R chart, we see that a point in $\overline{X}$ chart lies outside control limits while all points in R chart lie within control limits. Though the range variation is under control, we conclude that the process is out of statistical control.





2. The following are the sample means and ranges for ten samples, each of size 5. Construct the control chart for mean and range and comment on the nature of control.

Sample no.	1	2	3	4	5	6	7	8	9	10
Mean	12.8	13.1	13.5	12.9	13.2	14.1	12.1	15.5	13.9	14.2
Range	2.1	3.1	3.9	2.1	1.9	3.0	2.5	2.8	2.5	2.0

Solution:

$$\overline{X}=\frac{1}{N}\sum\overline{x_i}\quad=\frac{1}{10}(12.8+13.1+13.5+...+14.2)\Rightarrow\overline{X}=13.53$$

$$\overline{R}=\frac{1}{N}\sum\overline{R_i}\quad=\frac{1}{10}(2.1+3.1+3.9+...2.0)=2.59$$

From the table of control chart, for sample size n=5, we have

$$A_2=0.577,\ D_3=0\ and\ D_4=2.115$$

(i) Control limits for $\overline{X}$ chart:

CL (Central line) = $\overline{X}$ = 13.53

LCL = $\overline{X}$ - $A_2\overline{R}$ = 13.53 - (0.577)(2.59) = 12.03557 ≈ 12.04

UCL = $\overline{X}$ + $A_2\overline{R}$ = 13.53+ (0.577)(5.8) = 15.02443 ≈ 15.02

Conclusion:

Since 8th sample means fall outside the control limits the statistical process is out of control according to $\overline{X}$ chart.

(ii) Control limits for R chart:

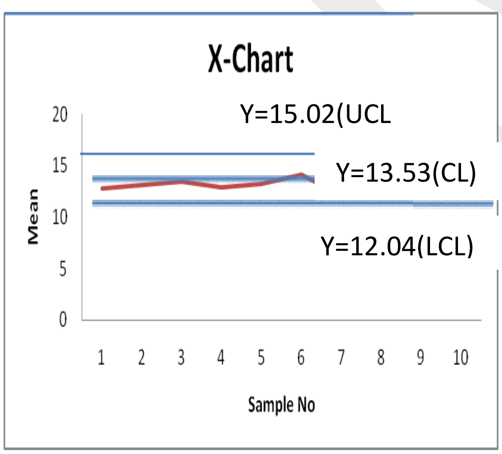
CL (Central line) = $\overline{R}$ = 2.59

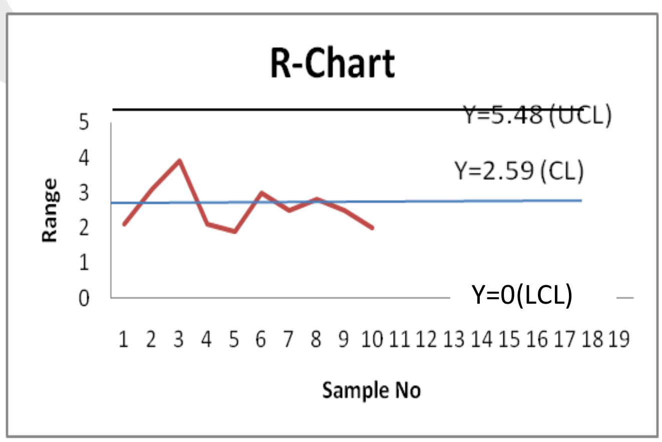
LCL = $D_3\overline{R}$ = 0 ;

UCL = $D_4\overline{R}$ = (2.115)(2.59) ≈ 5.48

Conclusion:

Since all the sample mean fall within the control limits the statistical process is under control according to R chart.





3. The following table gives the sample means and ranges for 10 samples, each of size 6, in the production of certain component. Construct the control charts for mean and range and comment on the nature of control.

Sample No:	**1**	**2**	**3**	**4**	**5**	**6**	**7**	**8**	**9**	**10**
Mean $\overline{X}$:	**37.3**	**49.8**	**51.5**	**59.2**	**54.7**	**34.7**	**51.4**	**61.4**	**70.7**	**75.3**
Range R:	**9.5**	**12.8**	**10.0**	**9.1**	**7.8**	**5.8**	**14.5**	**2.8**	**3.7**	**8.0**

Solution: **[A. U N/D 2019]**

$$\overline{\overline{X}} = \frac{1}{N}\sum \overline{x_i} \quad = \frac{1}{10}(37.3 + 49.8 + 51.5 + ... + 75.3) \Rightarrow \overline{\overline{X}} = 4.6$$

$$\overline{R} = \frac{1}{N}\sum \overline{R_i} \quad = \frac{1}{10}(9.5 + 12.8 + 10 + ... + 8.0) = 8.4$$

From the table of control chart, for sample size n=6, we have

$$A_2 = 0.483,\ D_3 = 0\ and\ D_4 = 2.004$$

(i) Control limits for $\overline{X}$ chart:

CL (Central line) = $\overline{\overline{X}}$ = 54.6

LCL = $\overline{\overline{X}}$ - $A_2\overline{R}$ =54.6- (0.483)(8.4) =50.543

UCL = $\overline{\overline{X}}$ + $A_2\overline{R}$ = 54.6+ (0.483)(8.4) =58.657

Conclusion:

Since 1st, 2nd, 4th, 6th, 8th, 9th and 10th sample means fall outside the control limits the statistical process is out of control according to $\overline{X}$ chart.

(ii) Control limits for R chart:

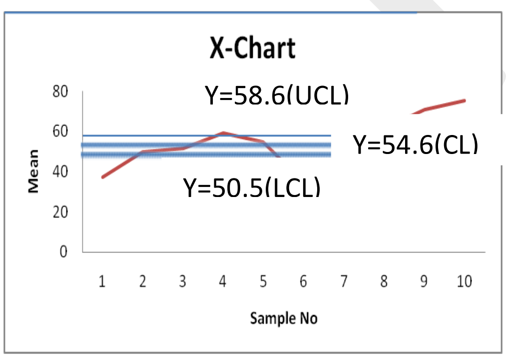
CL (Central line) = $\overline{R}$ = 8.4

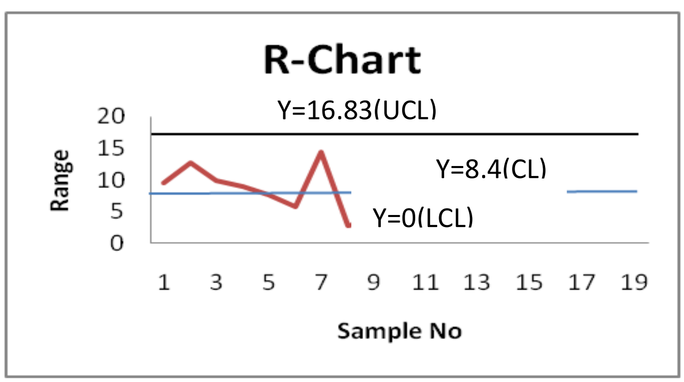
LCL = $D_3\overline{R}$ = 0

UCL = $D_4\overline{R}$ = (2.004)(8.4) =16.834

Conclusion:

Since all the sample mean fall within the control limits the statistical process is under control according to R chart.





4. The following data give the measurements of 10 samples each of size 5 in the production process taken in an interval of 2 hours. Calculate the sample mean and ranges and draw the control charts for mean and range.

	1	**2**	**3**	**4**	**5**	**6**	**7**	**8**	**9**	**10**
Observed measurements X	**49**	**50**	**50**	**48**	**47**	**52**	**49**	**55**	**53**	**54**
	55	**51**	**53**	**53**	**49**	**55**	**49**	**55**	**50**	**54**
	54	**53**	**48**	**51**	**50**	**47**	**49**	**50**	**54**	**52**
	49	**46**	**52**	**50**	**44**	**56**	**53**	**53**	**47**	**54**
	53	**50**	**47**	**53**	**45**	**50**	**45**	**57**	**51**	**56**

Solution:

We shall find Mean and Range for each sample.

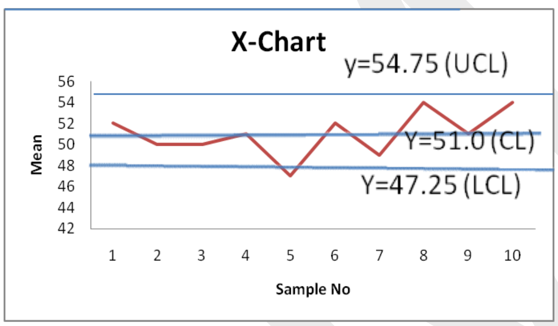
	1	2	3	4	5	6	7	8	9	10
$\sum X$	260	250	250	255	235	260	245	270	255	270
$\overline{X}$	52	50	50	51	47	52	49	54	51	54
R	6	7	6	5	6	9	8	7	7	4

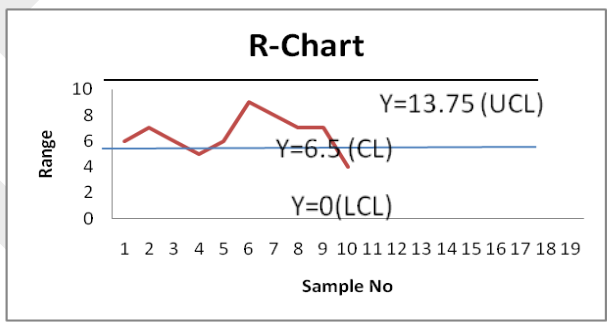
$$\overline{\overline{X}} = \frac{1}{N}\sum \overline{x_i} = \frac{1}{10}(52+50+....+54) = 51.0$$

$$\overline{R} = \frac{1}{N}\sum \overline{R_i} = \frac{1}{10}(6+7+6+...+4) = 6.5$$

From the table of control chart, for sample size n=5, we have

$$A_2 = 0.577,\ D_3 = 0\ and\ D_4 = 2.115$$





(i) Control limits for $\overline{X}$ chart:

CL (Central line) = $\overline{\overline{X}}$ = 51.0

LCL = $\overline{\overline{X}} - A_2\overline{R}$ =51.0- (0.577)(6.5) =47.2495

UCL = $\overline{\overline{X}} + A_2\overline{R}$ = 51.0+ (0.577)(6.5) =54.7505

Conclusion:

Since 5th sample means fall outside the control limits the statistical process is out of control according to $\overline{X}$ chart.

(ii) Control limits for R chart:

CL (Central line) = $\overline{R}$ =6.5 LCL = $D_3\overline{R}$ = 0 ;

UCL = $D_4\overline{R}$ = (2.115)(6.5) =13.7475

Conclusion: Since all the sample mean fall within the control limits the statistical process is under control according to R chart.

5. **The table given below gives the measurements obtained in 10 samples. Construct control charts for mean and the range. Discuss the nature of control**.

	1	2	3	4	5	6	7	8	9	10
Observed measurements X	62	50	67	64	49	63	61	63	48	70
	68	58	70	62	98	75	71	72	79	52
	66	52	68	57	65	62	66	61	53	62
	68	58	56	62	64	58	69	53	61	50
	73	65	61	63	66	68	77	55	49	66
	68	66	66	74	64	55	53	57	56	75

Solution: We shall find $\overline{X}$ and R for each sample.

	1	2	3	4	5	6	7	8	9	10
$\sum X$	405	349	388	382	406	381	397	361	346	395
$\overline{X}$	67.5	58.2	64.7	63.7	67.7	63.5	66.2	60.1	57.7	65.8
R	11	16	14	17	49	20	24	19	31	25

Solution:

$$\overline{\overline{X}} = \frac{1}{N}\sum \overline{x_i} \quad = \frac{1}{10}(67.5 + 58.2 + ... + 65.8) = 63.51$$

$$\overline{R} = \frac{1}{N}\sum \overline{R_i} \quad = \frac{1}{10}(11 + 16 + 14 + ... + 25) = 22.6$$

From the table of control chart, for sample size n=6, we have

$$A_2 = 0.483,\ D_3 = 0 \ and\ D_4 = 2.004$$

(i) Control limits for $\overline{X}$ chart:

CL (Central line) = $\overline{\overline{X}}$ = 63.51

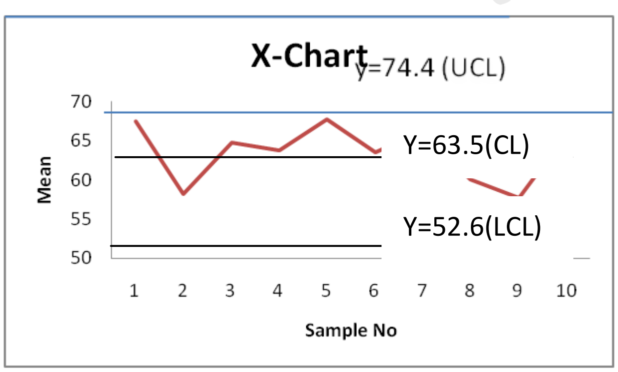
LCL = $\overline{\overline{X}}$ - $A_2\overline{R}$ =63.51- (0.483)(22.6) =52.59

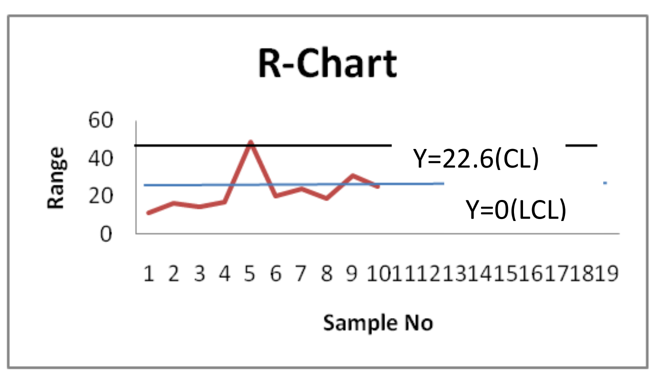
UCL = $\overline{\overline{X}}$ + $A_2\overline{R}$ = 63.51+ (0.483)(22.6) =74.43

Conclusion:

All the means of the sample lie between UCL and LCL 52.59 < all $\overline{X}$ <74.43

i.e., all sample points of means falling within 3 control limits. Hence, the process is in a state of Statistical control as far as means are concerned.





UCL = $D_4\overline{R}$ = (2.004)(22.6) =45.29

Conclusion:

The value of R corresponding to same[i.e., no. 5, namely 49], lies outside the control limits. Hence the variability is out of control.

6. **Control on measurements of pitch diameter of thread in air-craft fittings is checked with 5 samples each containing 5 items at equal intervals of time. The measurements are given below. Construct $\overline{X}$ and R charts and state your inference from the charts.**

	1	2	3	4	5
Measurements	46	41	40	42	43
	45	41	40	43	44
	44	44	42	43	47
	43	42	40	42	47
	42	40	42	45	45

Solution:

We shall find $\overline{X}$ and R for each sample.

Sample number

	1	2	3	4	5
$\sum X$	220	208	204	215	226
$\overline{X}$	44.0	41.6	40.8	43.0	45.2
R	4	4	2	3	4

$$\overline{\overline{X}} = \frac{1}{N}\sum \overline{x_i}$$

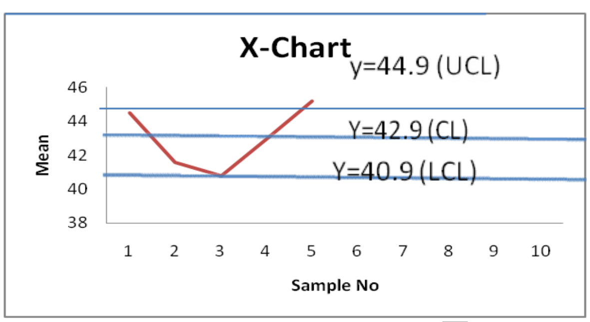
$$= \frac{1}{5}(44 + 41.6 + 40.8 + 43 + 45.2) = 42.92$$

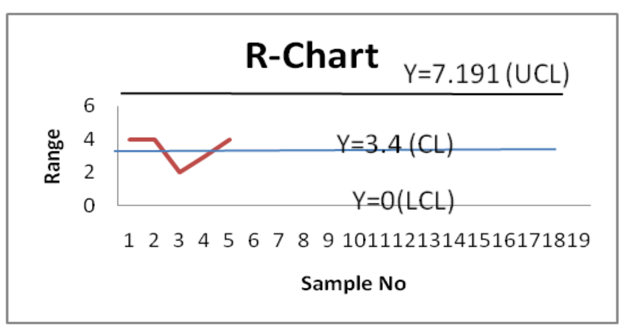
$$\overline{R} = \frac{1}{N}\sum \overline{R_i}$$

$$= \frac{1}{5}(17) = 3.4$$

From the table of control chart, for sample size n=5, we have

$$A_2 = 0.577,\ D_3 = 0\ and\ D_4 = 2.115$$





CL (Central line) = $\overline{\overline{X}}$ = 42.92

LCL = $\overline{X} - A_2\overline{R}$ =42.92- (0.577)(3.4) =40.96

UCL = $\overline{X} + A_2\overline{R}$ = 42.92+ (0.577)(3.4) =44.88

Conclusion:

$\overline{X}_5 = 45.2 > UCL = 44.88$ ie.) all sample do not lie between control limits. Hence, the process is out of control.

(ii) Control limits for R chart:

CL (Central line) = $\overline{R}$ =3.4

LCL = $D_3\overline{R}$ = 0 ;

UCL = $D_4\overline{R}$ = (2.115)(3.4) =7.191

Conclusion:

All sample points lie between control limits. Hence, the variability is under control. But process is out of control due to $\overline{X}$ charts.

7. Construct an $\overline{X} - R$ chart for the following data that give the heights of fragmentation bomb. Draw also the engineering specification tolerance limits of $0.830 \pm 0.010\,cm$ in the sample graph. Infer your conclusion

Group	ITEMS				
No.	1	2	3	4	5
1	0.831	0.829	0.836	0.840	0.826
2	0.834	0.826	0.831	0.831	0.831
3	0.836	0.826	0.831	0.822	0.816
4	0.833	0.831	0.835	0.831	0.833
5	0.830	0.831	0.831	0.833	0.820
6	8.829	0.828	0.828	0.832	0.841
7	0.835	0.833	0.829	0.830	0.841
8	0.818	0.838	0.835	0.834	0.830
9	0.841	0.831	0.831	0.833	0.832
10	0.832	0.828	0.836	0.832	0.825

Solution:

In the given problem there are 10 sample groups of 5 each; that is, N=10, n=5.

Group N0.	1	2	3	4	5	6	7	8	9	10
Sample total	4.162	4.153	4.131	4.163	4.145	4.158	4.168	4.155	4.153	4.153
Sample range	0.014	0.008	.020	.004	.013	.013	.012	.020	.010	0.011
Sample mean	.8324	.8306	.8262	.8326	.8290	.8316	.8310	.8310	.8336	0.8306

The control limits for $\overline{X} - chart$ are given $\overline{X} \pm A_2\overline{R}$ where A_2 is taken from statistical table for control chart for the sample size n=5. From the statistical table it is taken $A_2 = 0.577$. The values of $\overline{X}$ and $\overline{R}$ are calculated from the following table:

Grand total = Total of sample totals = 41.556

$$\overline{X} = \frac{grand\ total}{(Total\ no.\ of\ samples)(Sample\ size)} \qquad = \frac{41.446}{(10)(5)} = 0.83112$$

Total of sample ranges = 0.125

$$\overline{R} = \frac{Total\ of\ sample\ ranges}{no.\ of\ samples} = \frac{0.125}{10} = 0.0125$$

(i) Control limits for $\overline{X}$ chart:

CL (Central line) = $\overline{X}$ = 0.83112

LCL = $\overline{X} - A_2\overline{R}$ =0.83112- (0.577)(0.0125) =0.8239

UCL = $\overline{X} + A_2\overline{R}$ = 0.83112+ (0.577)(0.0125) =0.8383

(ii) Control limits for R chart:

LCL = $D_3\overline{R}$

UCL = $D_4\overline{R}$

Where D_3 and D_4 are constants taken from statistical tables for n. here n=5, and therefore theses values are $D_3 = 0$ and $D_4 = 2.115$

Hence, the LCL = 0 x 0.0125 = 0

UCL = 2.115 x 0.0125 = 0.0264

The process is under control.

8. The following data gives the average life in hours and range in hours of 12 samples each of 5 lamps. Construct $\overline{x}$ -chart and R-chart. Comment on the state of control of the process.

Sample no.:	**1**	**2**	**3**	**4**	**5**	**6**	**7**	**8**	**9**	**10**	**11**	**12**
Mean $\overline{x}_i$:	**120**	**127**	**152**	**157**	**160**	**134**	**137**	**123**	**140**	**144**	**120**	**127**
Range R:	**30**	**44**	**60**	**34**	**38**	**35**	**45**	**62**	**39**	**50**	**35**	**41**

Solution: **[A.U A/M 2019]**

$$\overline{X} = \frac{1}{N}\sum \overline{x}_i = \frac{1}{10}(120+127+\ldots+127) = 136.75$$

$$\overline{R} = \frac{1}{N}\sum \overline{R}_i = \frac{1}{10}(30+44+60+\ldots+41) = 42.75$$

From the table of control chart, for sample size n=5, we have

$$A_2 = 0.577,\ D\ = 0\ and\ D_4 = 2.115$$

(i) Control limits for $\overline{X}$ chart:

CL (Central line) = $\overline{X}$ = 136.75

LCL = $\overline{X} - A_2\overline{R}$ = 136.75 - (0.577)(42.75) = 112.08325

UCL = $\overline{X} + A_2\overline{R}$ = 136.75+ (0.577)(42.75) =161.41675

(ii) Control limits for R chart:

CL (Central line) = $\overline{R}$ = 42.75

LCL = $D_3\overline{R}$ = 0 ;

UCL = $D_4\overline{R}$ = (2.115)(42.75) =90.41625

Inference: From both $\overline{X}$ and R chart, we see that a point in $\overline{X}$ chart lies outside control limits while all points in R chart lie within control limits. Though the range variation is under control, we conclude that the process is out of statistical control.

9. A hard-bake process is used in conjunction with photolithography in semiconductor manufacturing. Fifteen samples, each of size five wafers, have been taken when we think the process is in control. The interval of time between samples or subgroups is one hour. The flow width measurement data (in x microns) from these samples are shown in Table. Establish statistical control of the flow width of the resist in this process using $\overline{x}$ and R charts.

	Wafers				
Sample No.	**1**	**2**	**3**	**4**	**5**
1	**1.3235**	**1.4128**	**1.6744**	**1.4573**	**1.6914**
2	**1.4314**	**1.3592**	**1.6075**	**1.4666**	**1.6109**
3	**1.4284**	**1.4871**	**1.4932**	**1.4324**	**1.5674**
4	**1.5028**	**1.6352**	**1.3841**	**1.2831**	**1.5507**
5	**1.5604**	**1.2735**	**1.5265**	**1.4363**	**1.6441**
6	**1.5955**	**1.5451**	**1.3574**	**1.3281**	**1.4198**
7	**1.6274**	**1.5064**	**1.8366**	**1.4177**	**1.5144**
8	**1.4190**	**1.4303**	**1.6637**	**1.6067**	**1.5519**
9	**1.3884**	**1.7277**	**1.5355**	**1.5176**	**1.3688**
10	**1.4039**	**1.6697**	**1.5089**	**1.4627**	**1.5220**
11	**1.4158**	**1.7667**	**1.4278**	**1.5928**	**1.4181**
12	**1.5821**	**1.3355**	**1.5777**	**1.3908**	**1.7559**
13	**1.2856**	**1.4106**	**1.4447**	**1.6398**	**1.1928**
14	**1.4951**	**1.4036**	**1.5893**	**1.6458**	**1.4969**
15	**1.3589**	**1.2863**	**1.5996**	**1.2497**	**1.5471**

Solution: **[AU A/M 2024]**

Sample No.	$\overline{x}$	Range(R)
1	1.5119	0.3679
2	1.4951	0.2517
3	1.4817	0.1390
4	1.4712	0.3521
5	1.4882	0.3706
6	1.4492	0.2674
7	1.5805	0.4189
8	1.5343	0.2447

9	1.5076	0.3589
10	1.5134	0.2658
11	1.5242	0.3509
12	1.5284	0.4204
13	1.3947	0.4470
14	1.5261	0.2422
15	1.4083	0.3499
	$\sum \bar{x} = 22.4148$	$\sum R = 4.8474$

$$\bar{\bar{x}} = \frac{\sum \bar{x}}{15} = \frac{22.4148}{15} = 1.49432$$

$$\bar{R} = \frac{\sum R}{15} = \frac{4.8474}{15} = 0.32316$$

$$D_3 = 0, \quad D_4 = 2.114$$

R – CHART

LCL = $\bar{R} \; D_3 = 0$

UCL = $\bar{R} \; D_4 = 0.68749$

CL = $\bar{R} = 0.32316$

$\bar{x}$ - CHART

UCL = $\bar{\bar{x}} + A_2 \bar{R} \approx 1.693$

LCL = $\bar{\bar{x}} - A_2 \bar{R}$

CL = $\bar{\bar{x}} = 1.49432$

10. Given below are the values of sample mean and sample range R for 10 samples, each of size 5. Draw the appropriate mean chart and comment on the state of control of the process.

Sample no.:	**1**	**2**	**3**	**4**	**5**	**6**	**7**	**8**	**9**	**10**
Mean:	**43**	**49**	**37**	**44**	**45**	**37**	**51**	**46**	**43**	**47**
Range:	**5**	**6**	**5**	**7**	**7**	**4**	**8**	**6**	**4**	**6**

Solution: **[AU N/D 2024]**

$$\overline{X} = \frac{43+49+37+44+45+37+51+46+43+47}{10}$$

$$= 44.2$$

$$\overline{R} = \frac{\sum R}{N}$$

$$= \frac{5+6+5+7+7+4+8+6+4+6}{10}$$

$$= 5.8$$

n = 5 , $A_2 = 0.577$, $D_3 = 0$, $D_4 = 2.45$

LCL = $\overline{\overline{x}} - A_2\overline{R}$ = 44.2 – (0.577) (5.8) = 40.8534

UCL = $\overline{\overline{x}} + A_2\overline{R}$ = 44.2 + (0.577)(5.8) = 47.5466

Hence two points are lie above the upper control and two points lie below the lower control line. Therefore, the process is not under control.

PROBLEMS BASED ON SAMPLE STANDARD DEVIATION (or) S-CHART

11. The following data give the coded measurements of 10 samples each of size 5, drawn from a process at intervals of 1 hour. Calculate the sample means and S.D's and draw the control charts for $\overline{X}$ and s.

Sample no.:	1	2	3	4	5	6	7	8	9	10
Coded measurements(X)	9	10	10	8	7	12	9	15	10	16
	15	11	13	13	9	15	9	15	13	14
	14	13	8	11	10	7	9	10	14	12
	9	6	12	10	4	16	13	13	7	14
	13	10	7	13	5	10	5	17	11	14

Solution:

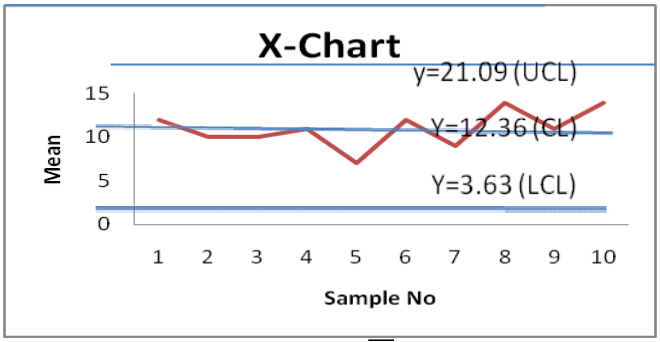
Sample no.	1	2	3	4	5	6	7	8	9	10
$\sum X$	60	50	50	55	35	60	45	70	55	70
$\overline{X}$	12	10	10	11	7	12	9	14	11	14
$\sum\left(X-\overline{X}\right)^2$	32	26	26	18	26	54	32	28	30	8
S	2.5	2.3	2.3	1.9	2.3	3.3	2.5	2.4	2.4	1.3

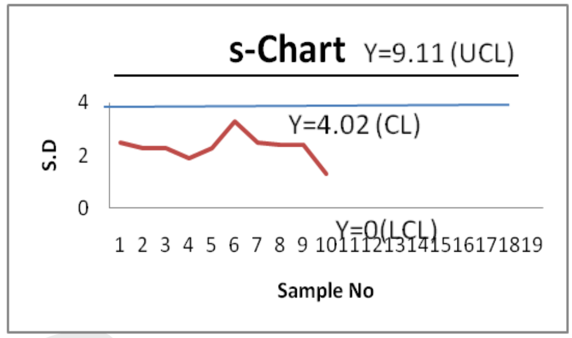
$$\overline{\overline{X}} = \frac{1}{N}\sum \overline{X}_i = \frac{1}{10}(12+10+\ldots+14) = \frac{110}{10} = 11$$

$$\bar{s} = \frac{1}{N}\sum s_i = \frac{1}{10}(2.5+2.3+\ldots+1.3) = \frac{23.2}{10} = 2.32$$

From the table of control chart constants, for sample size n=5, we have

$A_1 = 1.596,\ B_3 = 0\ and\ B_4 = 2.089$





(i) Control limits for $\overline{X}$ chart:

CL (Central line) = $\overline{\overline{X}} = 11$

LCL = $\overline{\overline{X}} - A_1\sqrt{\frac{n}{n-1}}\,\bar{s}$ = 11- (1.596) $\left(\sqrt{\frac{5}{4}}\right)$ (2.32) = 6.86

UCL = $\overline{\overline{X}} + A_1\sqrt{\frac{n}{n-1}}\,\bar{s}$ = 11+ (1.596) $\left(\sqrt{\frac{5}{4}}\right)$ (2.32) = 15.14

(ii) Control limits for s chart:

CL= $\bar{s} = 2.32$

LCL = $B_3\bar{s} = 0$ UCL = $B_4\bar{s} = (2.089)(2.32) = 4.85$

Conclusion:

The sample mean $\overline{X}$ values lie between 6.86 and 15.14 and the given S.D (s) values between 0 and 4.85. Hence the process is under control with respect to average and variability.

12. The values of sample mean $\overline{X}$ and sample S.D's for 15 samples, each of size 4, drawn from a production process are given below. Draw the appropriate control charts for the process average and process variability. Comment on the state of control.

Sample No.	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15
Mean	15.0	10.0	12.5	13.0	12.5	13.0	13.5	11.5	13.5	13.0	14.5	9.5	12.0	10.5	11.5
S.D	3.1	2.4	3.6	2.3	5.2	5.4	6.2	4.3	3.4	4.1	3.9	5.1	4.7	3.3	3.3

Solution:

$$\overline{\overline{X}} = \frac{1}{N}\sum \overline{X}_i = \frac{1}{15}(15+10+\ldots+11.5) = \frac{185.5}{15} = 12.36$$

$$\bar{s} = \frac{1}{N}\sum s_i = \frac{1}{15}(3.1+2.4+\ldots+3.3) = \frac{60.3}{15} = 4.02$$

From the table of control chart constants, for sample size n=4, we have

$A_1 = 1.880,\ B_3 = 0\ and\ B_4 = 2.266$

(i) **Control limits for $\overline{X}$ chart:**

CL (Central line) = $\overline{\overline{X}} = 12.36$

$$\text{LCL} = \overline{X} - A_1\sqrt{\frac{n}{n-1}}\,\overline{s} = 12.36 - (1.880)\left(\sqrt{\frac{4}{3}}\right)(4.02) = 3.63$$

$$\text{UCL} = \overline{X} + A_1\sqrt{\frac{n}{n-1}}\,\overline{s} = 12.36 + (1.880)\left(\sqrt{\frac{4}{3}}\right)(4.02) = 21.09$$

(ii) Control limits for s chart:

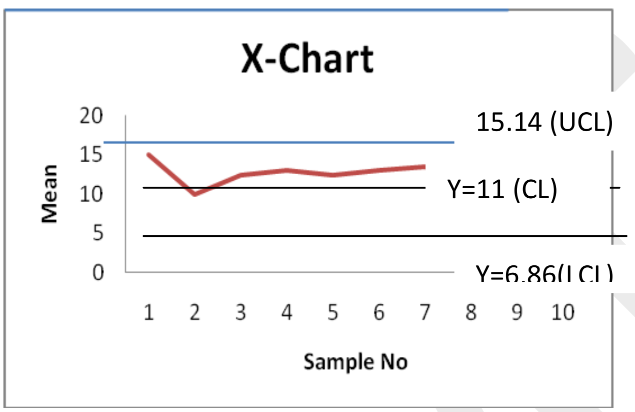
$\text{CL} = \overline{s} = 4.02$

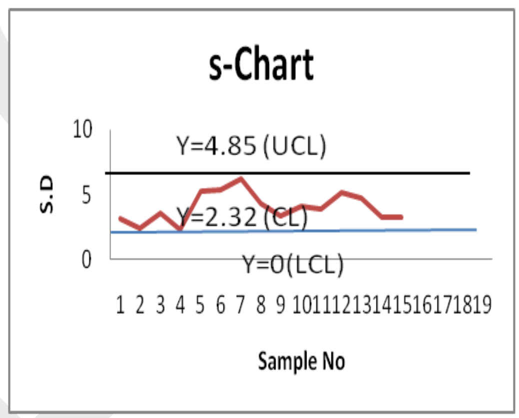
$\text{LCL} = B_3\overline{s} = 0$

$\text{UCL} = B_4\overline{s} = (2.266)(4.02) = 9.11$

Conclusion:

The sample mean $\overline{X}$ values lie between 3.63 and 21.09 and the given S.D (s) values between 0 and 9.11. Hence the process is under control with respect to average and variability.





PROBLEMS BASED ON P-CHART:

13. Thirty five successive samples of 100 castings each taken from a production line contained 3,3,5,3,5,0,3,2,3,5,6,5,9,1,2,4,5,2,0,10,3,6,3,2,5,6,3,3,2,5,1,0,7,4 and 3 rejectable castings. Construct a p-chart and state whether the process is under control or not.

Solution:

There are 35 samples each containing 100 castings.

Sample No:	No. of rejectable castings	Fraction defective	Sample No:	No. of rejectable castings (defectives)	Fraction defective
1	3	0.03	19	0	0.00
2	3	0.03	20	10	0.1
3	5	0.05	21	3	0.03
4	3	0.03	22	6	0.06
5	5	0.05	23	3	0.03
6	0	0.00	24	2	0.02
7	3	0.03	25	5	0.05
8	2	0.02	26	6	0.06

9	3	0.03	27	3	0.03
10	5	0.05	28	3	0.03
11	6	0.06	29	2	0.02
12	5	0.05	30	5	0.05
13	9	0.09	31	1	0.01
14	1	0.01	32	0	0.00
15	2	0.02	33	7	0.07
16	4	0.04	34	4	0.04
17	5	0.05	35	3	0.03
18	2	0.02	Total		1.29

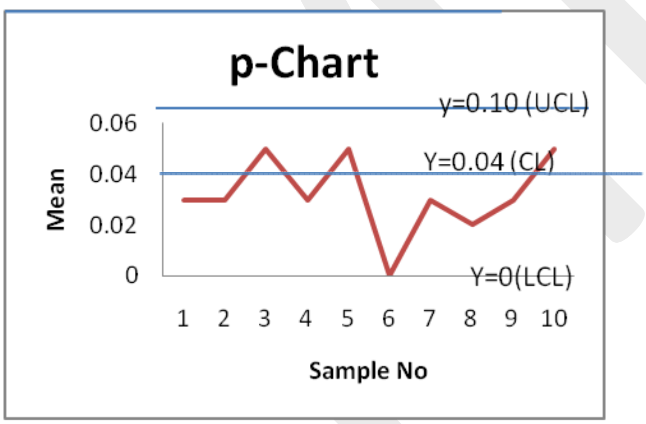
Average fraction defective $= \dfrac{1.29}{35} = 0.04$

$$UCL = \bar{p} + 3\sqrt{\frac{\bar{p}\left(1-\bar{p}\right)}{n}} = 0.04 + 3\sqrt{\frac{0.04(1-0.04)}{100}} = 0.04 + 0.06 = 0.10$$

$$LCL = \bar{p} - 3\sqrt{\frac{\bar{p}\left(1-\bar{p}\right)}{n}} = 0.04 - 3\sqrt{\frac{0.04(1-0.04)}{100}} = 0.04 - 0.06 = -0.02(Taken = 0)$$

(since negative value is meaningless)

Since no fraction defective exceeds 10 or below LCL the process is under control. The process is under control



14. 20 samples of each containing 100 items were taken at regular intervals of time. Construct a p-chart for the following observed data of those samples.

Sample No:	**1**	**2**	**3**	**4**	**5**	**6**	**7**	**8**	**9**	**10**
No. of defectives	**2**	**2**	**3**	**6**	**1**	**3**	**6**	**4**	**7**	**2**
Sample No:	**11**	**12**	**13**	**14**	**15**	**16**	**17**	**18**	**19**	**20**

No. of defectives	**5**	**0**	**3**	**2**	**4**	**5**	**3**	**8**	**1**	**4**

State whether the process is under control or not?

Solution:

Sample No:	No. of defects	Sample No:	No. of defects	Sample No:	No. of defects
1	2	0.02	11	5	0.05
2	2	0.02	12	0	0.00
3	3	0.03	13	3	0.03
4	6	0.06	14	2	0.02
5	1	0.01	16	5	0.05
6	3	0.03	17	3	0.03
7	6	0.06	18	8	0.08
8	4	0.04	19	1	0.01
9	7	0.07	20	4	0.04
10	2	0.02	Total	71	0.71

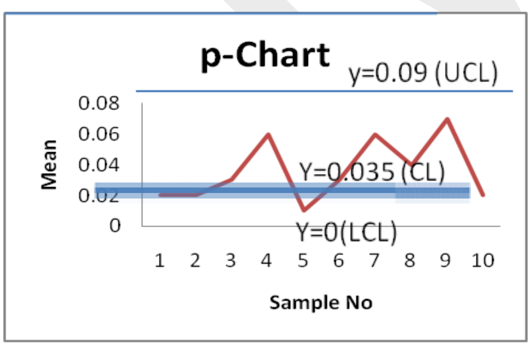
$$\bar{p} = \frac{0.71}{20*100} = 0.0355 \qquad UCL = \bar{p} + 3\sqrt{\frac{\bar{p}\left(1-\bar{p}\right)}{n}}$$

$$= 0.0355 + 3\sqrt{\frac{0.0355(1-0.0355)}{100}} = 0.0355 + 0.0545 = 0.09$$

$$LCL = \bar{p} - 3\sqrt{\frac{\bar{p}\left(1-\bar{p}\right)}{n}} = 0.0355 - 3\sqrt{\frac{0.0355(1-0.0355)}{100}} = 0.0355 - 0.0545 = -0.19(Taken = 0)$$

(since negative value is meaningless)

The process is under control.



15. Construct a control chart for defectives for the following data:

Sample No.	1	2	3	4	5	6	7	8	9	10
No. inspected	90	65	85	70	80	80	70	95	90	75
No. of defectives	9	7	3	2	9	5	3	9	6	7

Solution: **[A.U A/M 2019]**

We note that the size of the sample varies from sample to sample. We can construct p-chart, provided $0.75\overline{n} < n_i < 1.25\overline{n}$,

For all I,

Here, $\overline{n} = \frac{1}{N}\sum n_i = \frac{1}{10}(90+65+\ldots+75) = 80$

Hence $0.75\overline{n} = 60 \;\; and \; 1.25\overline{n} = 100$

The values 0f n_i be between 60 and 100. Hence p-chart, can be drawn by the method given below.

Now,

$$\bar{p} = \frac{total\,no.of\;defectives}{total\,no\,of\;items\,inspected} = \frac{60}{800} = 0.075$$

Hence p-chart to be constructed,

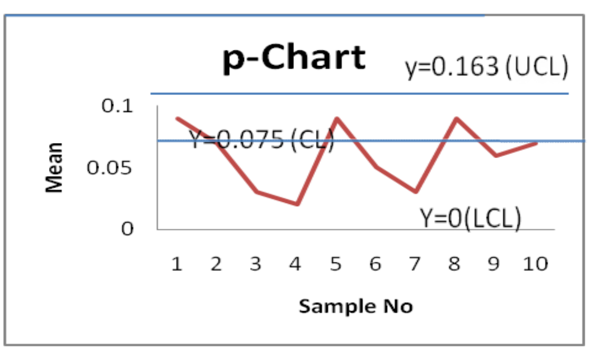
$$CL = \bar{p} = 0.075 \qquad UCL = \bar{p} + 3\sqrt{\frac{\bar{p}\left(1-\bar{p}\right)}{n}}$$

$$= 0.075 + 3\sqrt{\frac{0.075(1-0.075)}{80}} = 0.163$$

$$LCL = \bar{p} - 3\sqrt{\frac{\bar{p}\left(1-\bar{p}\right)}{n}} = 0.075 - 3\sqrt{\frac{0.075(1-0.075)}{80}} = -0.013(Taken = 0)$$

The values of p_i for the samples are 0.100,0.108,0.135,0.029,0.063,0.043,0.095, 0.067,0.093

Since all the sample point lie within the control lines, the process is under control.



16. Frozen orange juice concentrate is packed in 6-oz cardboard cans. These cans are formed on a machine by spinning them from cardboard stock and attaching a metal bottom panel. By inspection of a can, we may determine whether, when filled, it could possibly Leak either on the side seam or around the bottom joint. Such a nonconforming can has an improper seal on either the side seam or the bottom panel. To establish the control chart, 15 samples of $n = 50$ cans each were selected at half-hour intervals over a three-shift period in which the machine was in continuous operation.
The data are shown in Table. Set up a control chart for the fraction of nonconforming cans produced by this machine.

Sample Number	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15
Number of Nonconforming Cans, D_i	12	15	8	10	4	7	16	9	14	10	5	6	17	12	22

Solution: **[AU A/M 2024]**

Sample size is constant for all samples, n=50

Total no .of defectives

$$= 12+15+8+10+4+7+16+9+14+10+5+6+17+12+22$$
$$= 167$$

Total no .of Inspected = 50 x 15 = 750

Average fraction defection = $\overline{p}$

$$= \frac{Total\ no.of\ defectives}{Total\ no.of\ items\ inspected}$$
$$= \frac{167}{750}$$
$$= 0.2226$$

For p - chart

$$\text{UCL} = \overline{p} + 3\sqrt{\frac{(\overline{p})(1-\overline{p})}{n}}$$
$$= 0.2226 + 3\sqrt{\frac{0.2226(0.7773)}{50}}$$
$$= 0.2226 + 3\sqrt{0.00346}$$
$$= 0.2226 + 0.1764$$
$$= 0.3990$$

$$\text{LCL} = \overline{p} - 3\sqrt{\frac{(\overline{p})(1-\overline{p})}{n}}$$

$$= 0.2226 \ -3\sqrt{\frac{0.2226(0.7773)}{50}}$$

$$= 0.2226 \ -3\sqrt{0.00346}$$

$$= 0.2226 - 0.1764$$

$$= 0.0462$$

CL = 0.2226

17. 15 sample of 200 items each were drawn from the output of a process. The numbers of defectives items in the samples are given below. Prepare a control chart for the fraction defective and comment on the state of control.

Sample no.:	**1**	**2**	**3**	**4**	**5**	**6**	**7**	**8**
No. of defective:	**12**	**15**	**10**	**8**	**19**	**15**	**17**	**11**
Sample no.:	**9**	**10**	**11**	**12**	**13**	**14**	**15**	
No. of defective:	**13**	**20**	**10**	**8**	**9**	**5**	**8**	

Solution: **[AU N/D 2024]**

Sample size is constant for all samples , n = 200

Total no .of defectives

= 12+15+10+8+19+15+17+11+13+20+10+8+9+5+8

= 180

Total no .of inspected = 15 x 200 = 3000

$$\text{Average fraction defective} = \bar{p} = \frac{Total\ no.of\ defectives}{Total\ no.of\ items\ inspected}$$

$$= \frac{180}{3000}$$

$$= 0.06$$

For p – chart

$$\text{UCL} = \bar{p} + 3\sqrt{\frac{(\bar{p})(1-\bar{p})}{n}}$$

$$= 0.06 \ +3\sqrt{\frac{0.06(0.94)}{200}}$$

$$= 0.06 + 3\ (0.01679)$$

$$= 0.1103$$

$$\text{LCL} = \bar{p} - 3\sqrt{\frac{(\bar{p})(1-\bar{p})}{n}}$$

$$= 0.06 - 3\ (0.01679)$$

$= 0.00963 \approx 0.01$

$CL = \overline{p} = 0.06$

Fraction defectives are

0.06, 0.075, 0.05, 0.04, 0.095, 0.075, 0.085, 0.055, 0.065, 0.1, 0.05, 0.04, 0.045, 0.025, 0.04.

All points lies between LCL and UCL

Therefore, the process is under control.

PROBLEMS ON C-CHART

18. Samples of 40 articles are selected at regular intervals from the output of a stamping machine. The following are the number of non-confirming articles in each of the 20 samples 1,2,0,3,4,0,0,2,0,4,4,2,4,0,1,1,4,2,3,3. Construct c-chart and check whether the process is under control.

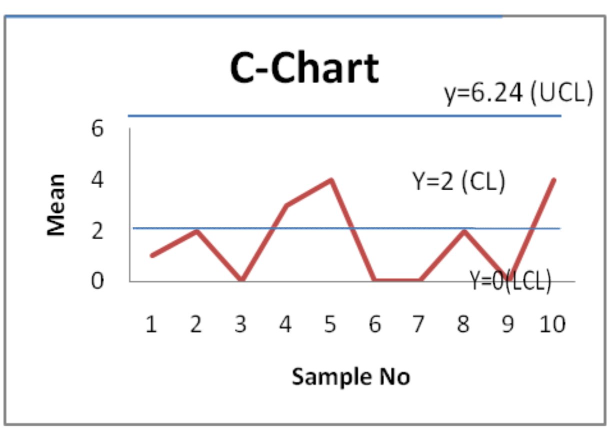
Solution:

Sample No:	No. of defects	Sample No:	No. of defects
1	1	11	4
2	2	12	2
3	0	13	4
4	3	14	0
5	4	15	1
6	0	16	1
7	0	17	4
8	2	18	2
9	0	19	3
10	4	20	3
		Total	40

$\bar{C} = \frac{40}{20} = 2 \ central \ \ line$ $UCL = \bar{C} + 3\sqrt{\bar{C}} = 2 + 3\sqrt{2} = 6.24$

$LCL = \bar{C} - 3\sqrt{\bar{C}} = -ve = 0$

The process is under control.



19. Ten units were inspected for non-confirming welds with the total number of defects as 360. Construct a c-chart for the number of non-conforming welds.

Solution:

$\bar{C} = \frac{360}{10} = 36 \; central \; line$

$UCL = \bar{C} + 3\sqrt{\bar{C}} = 36 + 3\sqrt{36} = 36 + 18 = 54$

$LCL = \bar{C} - 3\sqrt{\bar{C}} = 36 - 3\sqrt{36} = 36 - 18 = 18$

The process is under control.

20. **A plant produces paper for newsprint and rolls of paper are inspected for defects. The results of inspection of 20 rolls of paper are given below: draw the c-chart and comment on the state of control.**

Roll No.	**1**	**2**	**3**	**4**	**5**	**6**	**7**	**8**	**9**	**10**	**11**	**12**	**13**	**14**	**15**	**16**	**17**	**18**	**19**	**20**
No. of defects	**19**	**10**	**8**	**12**	**15**	**22**	**7**	**13**	**18**	**13**	**16**	**14**	**8**	**7**	**6**	**4**	**5**	**6**	**8**	**9**

Solution:

$\bar{C} = \frac{220}{20} = 11 \; central \; line$

$UCL = \bar{C} + 3\sqrt{\bar{C}} = 11 + 3\sqrt{11} = 20.95$

$LCL = \bar{C} - 3\sqrt{\bar{C}} = 11 - 3\sqrt{11} = 20.95$

Since one point falls outside the control line, the process is out of control.

21. Table presents the number of nonconformities observed in 26 successive samples of 100 printed circuit boards. Note that, for reasons of convenience, the inspection unit is defined as 100 boards. Set up a c-chart for these data.

Sample Number	**1**	**2**	**3**	**4**	**5**	**6**	**7**	**8**	**9**	**10**	**11**	**12**	**13**
Number of Nonconformities	**21**	**24**	**16**	**12**	**15**	**5**	**28**	**20**	**31**	**25**	**20**	**24**	**16**
Sample Number	**14**	**15**	**16**	**17**	**18**	**19**	**20**	**21**	**22**	**23**	**24**	**25**	**26**
Number of Nonconformities	**19**	**10**	**17**	**13**	**22**	**18**	**39**	**30**	**24**	**16**	**19**	**17**	**15**

Solution: **[AU A/M 2024]**

The number of non conformities observed in 26 successive samples of 100 printed circuit boards.

$$\overline{C} = \frac{1}{N}\sum C_i = \frac{1}{26} \text{ (21+24+16+12+15+5+28+20+31+25+20+24+16+19+10+17+13+22+18}$$
$$\text{+39+30+24+16+19+17+15)}$$
$$= \frac{1}{26} (516)$$
$$= 19.85$$

$$\text{CL} = \overline{C}$$
$$\text{LCL} = \overline{C} - 3\sqrt{C} = 19.85 - 3\sqrt{19.85}$$
$$= 6.48$$
$$\text{UCL} = \overline{C} + 3\sqrt{C} = 19.85 + 3\sqrt{19.85}$$
$$= 33.22$$

22. 15 Tape-recorders were examined for quality control test. The number of defects in each tape-recorder below. Draw the appropriate control chart and comment on the state control.

No. of units:	**1**	**2**	**3**	**4**	**5**	**6**	**7**	**8**	**9**
No. of defects:	**2**	**4**	**3**	**1**	**1**	**2**	**5**	**3**	**6**
No. of units:	**10**	**11**	**12**	**13**	**14**	**15**			
No. of defects:	**7**	**3**	**1**	**4**	**2**	**1**			

Solution: **[AU N/D 2024]**

Let c denote the number of defects in each tape – recorder is recorded.

$$\bar{c} = \frac{Total\ no.of\ defects}{Total\ sample\ inspected}$$
$$= \frac{\sum c}{n}$$
$$= \frac{45}{15}$$
$$= 3$$
$$\text{UCL} = \bar{c} + 3\sqrt{\bar{c}} = 3 + 3\sqrt{3}$$
$$= 8.196$$
$$\text{LCL} = \bar{c} - 3\sqrt{\bar{c}} = 3 - 3\sqrt{3}$$
$$= -2.196$$

$$CL = \bar{c} = 3$$

Therefore, the process is under control.

PROBLEMS ON np-chart

23. In a factory 1000 bolts are examined daily for defects. The following are the number of defects in 15 days: 9,10,12,8,7,15,10,8,7,13,14,15,16. Draw an np-chart and give your findings?

Solution:

Sample No:	No. of defects	Sample No:	No. of defects
1	9	9	10
2	10	10	8
3	12	11	7
4	8	12	13
5	7	13	14
6	15	14	15
7	10	15	16
8	12	Total	166

$$\bar{p} = average\ \ number\ \ of\ \ defective = \frac{166}{15*1000} = 0.01$$

$$n\bar{p} = 1000*0.01 = 10$$

$$UCL = n\bar{p} + 3\sqrt{n\bar{p}\left(1-\bar{p}\right)}$$

$$= 10 + 3\sqrt{10*(1-0.01)}$$

$$= 10 + 9.39 = 19.39$$

$$LCL = n\bar{p} - 3\sqrt{n\bar{p}\left(1-\bar{p}\right)}$$

$$= 1 - 9.39 = 0.61$$

The process is under control.

24. The following table gives the data on completed sparkplugs for 10 samples each of 100 plugs.

Sample No:	**1**	**2**	**3**	**4**	**5**	**6**	**7**	**8**	**9**	**10**
No: of defectives	**5**	**12**	**6**	**6**	**3**	**4**	**8**	**3**	**5**	**6**

Construct a suitable control chart and comment on your result:

Solution:

Sample No:	No. of defects
1	5
2	12
3	6
4	6
5	3
6	4
7	8
8	3
9	5
10	6
total	58

$$\bar{p} = average\ \ number\ \ of\ \ defective = \frac{58}{10*1000} = 0.058$$,n=100

$$n\bar{p} = 100*0.06 = 6$$

$$UCL = n\bar{p} + 3\sqrt{n\bar{p}\left(1-\bar{p}\right)} = 6 + 3\sqrt{6*(1-0.06)} = 12.90 = 13$$

$$LCL = n\bar{p} - 3\sqrt{n\bar{p}\left(1-\bar{p}\right)} = 6 - 0.91 = 5.09$$

The process is under control.

25. The following table gives the number of defectives in 10 samples, each of size 100. Construct a np-chart for these data and also determine whether the process is in control.

Sample No:	**1**	**2**	**3**	**4**	**5**	**6**	**7**	**8**	**9**	**10**
No: of defectives	**24**	**38**	**62**	**34**	**26**	**36**	**38**	**52**	**33**	**44**

Solution: **[A.U A/M 2019]**

Sample No:	No. of defects
1	24
2	38
3	62
4	34
5	26
6	36
7	38
8	52
9	33
10	44
total	387

$$\bar{p} = average\ \ number\ \ of\ \ defective = \frac{387}{10*100} = 0.387 \text{ ,n=100}$$

$$n\bar{p} = 100*0.387 = 38.7$$

$$UCL = n\bar{p} + 3\sqrt{n\bar{p}\left(1-\bar{p}\right)} = 38.7 + 3\sqrt{38.7*(1-0.387)} = 53.3119$$

$$LCL = n\bar{p} - 3\sqrt{n\bar{p}\left(1-\bar{p}\right)} = 24.0881$$

The process is not under control.

26. **10 Samples each of size 50 were inspected and the numbers of defectives in the inspection were: 2, 1, 1, 2, 3, 5, 5, 1, 2, 3. Draw the np-chart for defectives.**
Solution: **[AU N/D 2024]**

Given n = 50

Total no. of defectives = 2+1+1+2+3+5+5+1+2+3
= 25

Total no .of inspected = 10 x 50 =500

Average fraction defective = $\bar{p}$

$$= \frac{Total\ no.of\ defectives}{Total\ no.of\ items\ inspected}$$

$$= \frac{25}{500}$$

$$= 0.05$$

For np – chart

$$\text{UCL} = n\bar{p} + 3\sqrt{n\bar{p}}\left(1-\bar{p}\right)$$

$$= n\left[\bar{p} + 3\sqrt{\frac{\left(\bar{p}\right)\left(1-\bar{p}\right)}{n}}\right]$$

$$= 50\left[0.05 + 3\sqrt{\frac{0.05(0.95)}{50}}\right]$$

$$= 50\left[0.05 + 3\left(0.0308\right)\right] = 50\ (0.1425) = 7.1233$$

$$\text{LCL} = n\bar{p} - 3\sqrt{n\bar{p}}\left(1-\bar{p}\right)$$

$$= n\left[\bar{p} - 3\sqrt{\frac{\left(\bar{p}\right)\left(1-\bar{p}\right)}{n}}\right]$$

$$= 50\left[0.05 - 3\left(0.0308\right)\right] = -2.12$$

$$\text{CL} = n\bar{p} = 50\left(0.05\right) = 2.5$$

PROBLEMS BASED ON p-CHART & np-CHART:

27. The data given below are the number of defectives in 10 samples of 100 items each. Construct a p-chart and an np-chart and comment on the results:

Sample No:	1	2	3	4	5	6	7	8	9	10
No: of defectives	6	16	7	3	8	12	7	11	11	4

Solution: **[AU N/D 2019]**

Sample size is constant for all samples n=100.

Total no. of defectives = 6+16+7+3+8+12+7+11+11+4=85

Total no. of inspected =10 x 100=1000

Average fraction defective $=\bar{p}=\dfrac{Total\ no.\ of\ defectives}{Total\ no.\ of\ items\,inspected}=\dfrac{85}{1000}=0.085$

For p-chart:

$$UCL=\bar{p}+3\sqrt{\frac{\bar{p}\left(1-\bar{p}\right)}{n}}=0.085+3\sqrt{\frac{0.085(1-0.085)}{100}}=0.1687$$

$$LCL=\bar{p}-3\sqrt{\frac{\bar{p}\left(1-\bar{p}\right)}{n}}=0.085-3\sqrt{\frac{0.085(1-0.085)}{100}}=0.0013$$

CL corresponds to $\bar{p}=0.085$

Fraction defectives for samples are 0.06,0.16,0.07,0.03,0.8,0.12,0.07,0.11,0.04.

Conclusion: All these values are less than UCL=0.1687 and greater than LCL=0.0013. In the control chart, all sample point lie within the control limits. Hence, the process is under statistical control.

For np-chart:

$$UCL=n\bar{p}+3\sqrt{n\bar{p}\left(1-\bar{p}\right)} \qquad UCL=n\left[\bar{p}+3\sqrt{\frac{\bar{p}\left(1-\bar{p}\right)}{n}}\right]=100*0.1687=16.87$$

$$n\bar{p}=100*0.085=8.5$$

$$LCL=n\bar{p}-3\sqrt{n\bar{p}\left(1-\bar{p}\right)} \quad LCL=n\left[\bar{p}-3\sqrt{\frac{\bar{p}\left(1-\bar{p}\right)}{n}}\right]=100*0.0013=0.13$$

Conclusion:

All the values of number of defectives in the table lie between 16.87 and 0.13. hence, the process is under control even in np-chart.

28. The table below presents 10 subgroups of four measurements on the critical dimension of a part produced by a machining process. [A.U N/D 2021]

Sample No.	1	2	3	4	5	6	7	8	9	10
x1	13	14	11	11	10	10	12	13	13	13
x2	11	14	13	11	12	11	8	15	12	11
x3	13	10	12	13	11	13	11	12	15	14
x4	13	13	15	12	14	13	11	12	12	13

Construct an $X - R$ control chart. Is the process under control?

Solution:

We shall find $\overline{X}$ and R for each sample.

	1	2	3	4	5	6	7	8	9	10
$\sum X$	50	51	51	47	47	47	42	52	52	51
$\overline{X}$	18.5	12.75	12.75	11.75	11.75	11.75	10.5	13	13	12.75
R	2	4	4	2	4	3	4	3	3	3

$$\overline{X} = \frac{1}{N}\sum \overline{x_i} = \frac{1}{10}(122.5) = 12.25$$

$$\overline{R} = \frac{1}{N}\sum \overline{R_i} = \frac{1}{5}(32) = 3.2$$

From the table of control chart, for sample size n=5, we have

$$A_2 = 0.729,\ D_3 = 0\ and\ D_4 = 2.282$$

(i) Control limits for $\overline{X}$ chart:

CL (Central line) = $\overline{X}$ = 12.25

LCL = $\overline{X} - A_2\overline{R}$ =12.25- (0.729)(3.2) =9.9172

UCL = $\overline{X} + A_2\overline{R}$ = 12.25+ (0.729)(3.2) =14.5828

Conclusion:

All sample lie between control limits. Hence, the process is under the control.

(ii) Control limits for R chart:

CL (Central line) = $\overline{R}$ =3.2

LCL = $D_3\overline{R}$ = 0 ;

UCL = $D_4\overline{R}$ = (2.282)(3.2) =7.3024

Conclusion:

All sample points lie between control limits. Hence, the process is under the control.

29. Surface defects have been counted on 25 rectangular steel plates and the data are 1, 0, 4, 3, 1, 2, 5, 0, 2, 1, 1, 0, 8, 0, 2, 1, 3, 5, 4, 6, 3,1, 0, 2 and 4. Set up a control chart for number of conformities using these data. Is the process under control? [A.U N/D 2021]

Solution:

$$\bar{C} = \frac{55}{25} = 2.2\ \ central\ \ line$$

$$UCL = \bar{C} + 3\sqrt{\bar{C}} \;= 2.2 + 3\sqrt{2.2} = 6.64$$

$$LCL = \bar{C} - 3\sqrt{\bar{C}} = 2.2 - 3\sqrt{2.2} = 2.24$$

Since one point falls outside the control line, the process is out of control

30. The data below gives the number of non-conforming bearing and seal assemblies in samples of size 100. Construct a fraction non conforming control chat for these data. Check if the process is under control.

Sample no.: 1 2 3 4 5 6 7 8 9 10

No. of non-conforming assemblies: 7 4 1 3 6 8 10 5 2 7 **[A.U N/D 2021]**

Solution:

There are 10 samples each containing 100 castings.

Sample No:	No. of rejectable castings	Fraction defective
1	7	0.07
2	4	0.04
3	1	0.01
4	3	0.03
5	6	0.06
6	8	0.08
7	10	0.1
8	5	0.05
9	2	0.02
10	7	0.07

Average fraction defective $= \frac{0.53}{10} = 0.053$

$$UCL = \bar{p} + 3\sqrt{\frac{\bar{p}\left(1 - \bar{p}\right)}{n}}$$

$$= 0.053 + 3\sqrt{\frac{0.053(1 - 0.053)}{100}} = 0.053 + 0.066 = 0.1202$$

$$LCL = \bar{p} - 3\sqrt{\frac{\bar{p}\left(1 - \bar{p}\right)}{n}}$$

$$= 0.053 + 3\sqrt{\frac{0.053(1 - 0.053)}{100}} = 0.053 - 0.066 = -0.013(Taken = 0)$$

(since negative value is meaningless)

The process is under control

31. The following table data show the values of the sample mean $\bar{x}$ and the range R for the samples of size 5 each. Determine whether the process is in control. [A.U N/D 2020]

Sample No.	**1**	**2**	**3**	**4**	**5**	**6**	**7**	**8**	**9**	**10**
Mean $\bar{x}$	**11.2**	**11.8**	**10.8**	**11.6**	**11.0**	**9.6**	**10.4**	**9.6**	**10.6**	**10.0**
Range R	**7**	**4**	**8**	**5**	**7**	**4**	**8**	**4**	**7**	**9**

Given the conversion factors for $n = 5$ are $A2 = 0.577$, $D3 = 0$ and $D4 = 2.115$.

Solution:

$$\overline{X} = \frac{1}{N}\sum \overline{x_i} = \frac{106.6}{10} = 10.66$$

$$\overline{R} = \frac{1}{N}\sum \overline{R_i} = \frac{63}{10} = 6.3$$

From the table of control chart, for sample size n=5, we have

$$A_2 = 0.577,\ D_3 = 0\ and\ D_4 = 2.115$$

(i) Control limits for $\overline{X}$ chart:

CL (Central line) = $\overline{X}$ = 10.66

LCL = $\overline{X} - A_2\overline{R}$ = 10.66 - (0.577)(6.3) = 7.0249

UCL = $\overline{X} + A_2\overline{R}$ = 10.66 + (0.577)(6.3) = 14.295

Conclusion: Since 2nd,3rd 6th and 7th sample means fall outside the control limits the statistical process is out of control according to $\overline{X}$ chart.

(ii) Control limits for R chart:

CL (Central line) = $\overline{R}$ = 6.3

LCL = $D_3\overline{R}$ = 0 ;

UCL = $D_4\overline{R}$ = (2.115)(6.3) = 13.959

Conclusion:

Since all the sample mean fall within the control limits the statistical process is under control according to R chart.

Inference: From both $\overline{X}$ and R chart, we see that a point in $\overline{X}$ chart lies outside control limits while all points in R chart lie within control limits. Though the range variation is under control, we conclude that the process is out of statistical control.

32. The following values represent the number of defectives of 10 samples each containing 100 items. 8, 10, 9, 8, 10, 11, 7, 9, 6, 12 Draw control chart for fraction defective and comment on the state of control of the process [A.U N/D 2020]

Solution:

There are 10 samples each containing 100 castings.

Sample No.:	No. of rejectable castings	Fraction defective
1	8	0.08
2	10	0.1
3	9	0.09
4	8	0.08
5	10	0.1

6	11	0.11
7	7	0.07
8	9	0.09
9	6	0.06
10	12	0.12

Average fraction defective $=\dfrac{0.9}{10}=0.09$

$$UCL=\bar{p}+3\sqrt{\frac{\bar{p}\left(1-\bar{p}\right)}{n}}=0.09+3\sqrt{\frac{0.09\left(1-0.09\right)}{100}}=0.09+3\left[0.029\right]=0.177$$

$$LCL=\bar{p}-3\sqrt{\frac{\bar{p}\left(1-\bar{p}\right)}{n}}=0.09+3\sqrt{\frac{0.09\left(1-0.09\right)}{100}}=0.09-0.087=0.003$$

The process is under control.

ANNA UNIVERSITY QUESTIONS

Problems based on $\bar{x}$ and R- chart

1. Given below are the values of sample mean $\bar{x}$ and sample range R for 10 samples, each of size 5. Draw the appropriate mean and range charts and comment on the state of control of the process.[pg. no. 5]

Sample no.	1	2	3	4	5	6	7	8	9	10
Mean $\bar{x}_i$	43	49	37	44	45	37	51	46	43	47
Range R	5	6	5	7	7	4	8	6	4	6

2. The following are the sample means and ranges for ten samples, each of size 5. Construct the control chart for mean and range and comment on the nature of control. [pg. no. 6]

Sample no.	1	2	3	4	5	6	7	8	9	10
Mean	12.8	13.1	13.5	12.9	13.2	14.1	12.1	15.5	13.9	14.2
Range	2.1	3.1	3.9	2.1	1.9	3.0	2.5	2.8	2.5	2.0

3. The following table gives the sample means and ranges for 10 samples, each of size 6, in the production of certain component. Construct the control charts for mean and range and comment on the nature of control.

Sample No:	1	2	3	4	5	6	7	8	9	10
Mean $\bar{X}$:	37.3	49.8	51.5	59.2	54.7	34.7	51.4	61.4	70.7	75.3
Range R:	9.5	12.8	10.0	9.1	7.8	5.8	14.5	2.8	3.7	8.0

[A. U N/D 2019][pg.no.7]

4. The following data give the measurements of 10 samples each of size 5 in the production process taken in an interval of 2 hours. Calculate the sample mean and ranges and draw the control charts for mean and range. [pg.no.8]

	1	2	3	4	5	6	7	8	9	10
Observed measurements X	49	50	50	48	47	52	49	55	53	54
	55	51	53	53	49	55	49	55	50	54
	54	53	48	51	50	47	49	50	54	52
	49	46	52	50	44	56	53	53	47	54
	53	50	47	53	45	50	45	57	51	56

5. The table given below gives the measurements obtained in 10 samples. Construct control charts for mean and the range. Discuss the nature of control. [pg.no.9]

	1	2	3	4	5	6	7	8	9	10
Observed measurements X	62	50	67	64	49	63	61	63	48	70
	68	58	70	62	98	75	71	72	79	52
	66	52	68	57	65	62	66	61	53	62
	68	58	56	62	64	58	69	53	61	50
	73	65	61	63	66	68	77	55	49	66
	68	66	66	74	64	55	53	57	56	75

6. Control on measurements of pitch diameter of thread in air-craft fittings is checked with 5 samples each containing 5 items at equal intervals of time. The measurements are given below. Construct $\overline{X}-R$ charts and state your inference from the charts. [pg.no.10]

	1	2	3	4	5
Measurements	46	41	40	42	43
	45	41	40	43	44
	44	44	42	43	47
	43	42	40	42	47
	42	40	42	45	45

7. Construct an $\overline{X}-R$ chart for the following data that give the heights of fragmentation bomb. Draw also the engineering specification tolerance limits of $0.830 \pm 0.010\,cm$ in the sample graph. Infer your conclusion. [pg.no.11]

Group no.	Items				
	1	2	3	4	5
1	0.831	0.829	0.836	0.840	0.826
2	0.834	0.826	0.831	0.831	0.831
3	0.836	0.826	0.831	0.822	0.816
4	0.833	0.831	0.835	0.831	0.833
5	0.830	0.831	0.831	0.833	0.820
6	8.829	0.828	0.828	0.832	0.841
7	0.835	0.833	0.829	0.830	0.841
8	0.818	0.838	0.835	0.834	0.830
9	0.841	0.831	0.831	0.833	0.832
10	0.832	0.828	0.836	0.832	0.825

8. The following data gives the average life in hours and range in hours of 12 samples each of 5 lamps. Construct $\overline{x}$ -chart and R-chart. Comment on the state of control of the process. [pg.no.12]

Sample no.:	1	2	3	4	5	6	7	8	9	10	11	12
Mean $\overline{x}_i$:	120	127	152	157	160	134	137	123	140	144	120	127
Range R:	30	44	60	34	38	35	45	62	39	50	35	41

9. A hard-bake process is used in conjunction with photolithography in semiconductor manufacturing. Fifteen samples, each of size five wafers, have been taken when we think the process is in control. The interval of time between samples or subgroups is one hour. The flow width measurement data (in x microns) from these samples are shown in Table. Establish statistical control of the flow width of the resist in this process using $\overline{x}$ and R charts.

	Wafers				
Sample No.	1	2	3	4	5
1	1.3235	1.4128	1.6744	1.4573	1.6914
2	1.4314	1.3592	1.6075	1.4666	1.6109
3	1.4284	1.4871	1.4932	1.4324	1.5674
4	1.5028	1.6352	1.3841	1.2831	1.5507

5	1.5604	1.2735	1.5265	1.4363	1.6441
6	1.5955	1.5451	1.3574	1.3281	1.4198
7	1.6274	1.5064	1.8366	1.4177	1.5144
8	1.4190	1.4303	1.6637	1.6067	1.5519
9	1.3884	1.7277	1.5355	1.5176	1.3688
10	1.4039	1.6697	1.5089	1.4627	1.5220
11	1.4158	1.7667	1.4278	1.5928	1.4181
12	1.5821	1.3355	1.5777	1.3908	1.7559
13	1.2856	1.4106	1.4447	1.6398	1.1928
14	1.4951	1.4036	1.5893	1.6458	1.4969
15	1.3589	1.2863	1.5996	1.2497	1.5471

[AU A/M 2024][pg.no.13]

10. Given below are the values of sample mean and sample range R for 10 samples, each of size 5. Draw the appropriate mean chart and comment on the state of control of the process.

Sample no.:	1	2	3	4	5	6	7	8	9	10
Mean:	43	49	37	44	45	37	51	46	43	47
Range:	5	6	5	7	7	4	8	6	4	6

[AU N/D 2024][pg.no.14]

PROBLEMS BASED ON SAMPLE STANDARD DEVIATION (or) S-CHART

11. The following data give the coded measurements of 10 samples each of size 5, drawn from a process at intervals of 1 hour. Calculate the sample means and S.D's and draw the control charts for $\overline{X}$ and s. [pg.no.15]

Sample no.:	1	2	3	4	5	6	7	8	9	10
Coded measurements(X)	9	10	10	8	7	12	9	15	10	16
	15	11	13	13	9	15	9	15	13	14
	14	13	8	11	10	7	9	10	14	12
	9	6	12	10	4	16	13	13	7	14
	13	10	7	13	5	10	5	17	11	14

12. The values of sample mean $\overline{X}$ and sample S.D's for 15 samples, each of size 4, drawn from a production process are given below. Draw the appropriate control charts for the process average and process variability. Comment on the state of control. [pg.no.16]

Sample No.	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15
Mean	15.0	10.0	12.5	13.0	12.5	13.0	13.5	11.5	13.5	13.0	14.5	9.5	12.0	10.5	11.5
S.D	3.1	2.4	3.6	2.3	5.2	5.4	6.2	4.3	3.4	4.1	3.9	5.1	4.7	3.3	3.3

PROBLEMS BASED ON P-CHART:

13. Thirty five successive samples of 100 castings each taken from a production line contained 3,3,5,3,5,0,3,2,3,5,6,5,9,1,2,4,5,2,0,10,3,6,3,2,5,6,3,3,2,5,1,0,7,4 and 3 reject able castings. Construct a p-chart and state whether the process is under control or not. [pg.no.17]

14. 20 samples of each containing 100 items were taken at regular intervals of time. Construct a p-chart for the following observed data of those samples. [pg.no.18]

Sample No:	1	2	3	4	5	6	7	8	9	10
No. of defectives	2	2	3	6	1	3	6	4	7	2
Sample No:	11	12	13	14	15	16	17	18	19	20
No. of defectives	5	0	3	2	4	5	3	8	1	4

State whether the process is under control or not?

15. Construct a control chart for defectives for the following data:

Sample No.	1	2	3	4	5	6	7	8	9	10
No. inspected	90	65	85	70	80	80	70	95	90	75
No. of defectives	9	7	3	2	9	5	3	9	6	7

[A.U A/M 2019][pg.no.20]

16. Frozen orange juice concentrate is packed in 6-oz cardboard cans. These cans are formed on a machine by spinning them from cardboard stock and attaching a metal bottom panel. By inspection of a can, we may determine whether, when filled, it could possibly Leak either on the side seam or around the bottom joint. Such a nonconforming can has an improper seal on either the side seam or the bottom panel. To establish the control chart, 15 samples of $n = 50$ cans each were selected at half-hour intervals over a three-shift period in which the machine was in continuous operation.
The data are shown in Table. Set up a control chart for the fraction of nonconforming cans produced by this machine.

Sample Number	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15
Number of Nonconforming	12	15	8	10	4	7	16	9	14	10	5	6	17	12	22

Cans, D_i															

[AU A/M 2024][pg.no.21]

17. 15 sample of 200 items each were drawn from the output of a process. The numbers of defectives items in the samples are given below. Prepare a control chart for the fraction defective and comment on the state of control.

Sample no.:	1	2	3	4	5	6	7	8
No. of defective:	12	15	10	8	19	15	17	11
Sample no.:	9	10	11	12	13	14	15	
No. of defective:	13	20	10	8	9	5	8	

[AU N/D 2024][pg.no.22]

PROBLEMS ON C-CHART

18. Samples of 40 articles are selected at regular intervals from the output of a stamping machine. The following are the number of non-confirming articles in each of the 20 samples 1,2,0,3,4,0,0,2,0,4,4,2,4,0,1,1,4,2,3,3. Construct c-chart and check whether the process is under control. [pg.no.23]

19. Ten units were inspected for non-confirming welds with the total number of defects as 360. Construct a c-chart for the number of non-conforming welds. [pg.no.24]

20. A plant produces paper for newsprint and rolls of paper are inspected for defects. The results of inspection of 20 rolls of paper are given below: draw the c-chart and comment on the state of control. [pg.no.24]

Roll No.	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17	18	19	20
No. of defects	19	10	8	12	15	22	7	13	18	13	16	14	8	7	6	4	5	6	8	9

21. Table presents the number of nonconformities observed in 26 successive samples of 100 printed circuit boards. Note that, for reasons of convenience, the inspection unit is defined as 100 boards. Set up a c-chart for these data.

Sample Number	1	2	3	4	5	6	7	8	9	10	11	12	13
Number of Nonconformities	21	24	16	12	15	5	28	20	31	25	20	24	16
Sample Number	14	15	16	17	18	19	20	21	22	23	24	25	26
Number of Nonconformities	19	10	17	13	22	18	39	30	24	16	19	17	15

[AU A/M 2024] [pg.no.24]

22. 15 Tape-recorders were examined for quality control test. The number of defects in each tape-recorder below. Draw the appropriate control chart and comment on the state control.

No. of units:	1	2	3	4	5	6	7	8	9
No. of defects:	2	4	3	1	1	2	5	3	6
No. of units:	10	11	12	13	14	15			
No. of defects:	7	3	1	4	2	1			

[AU N/D 2024][pg.no.25]

Problems on np-chart:

23. In a factory 1000 bolts are examined daily for defects. The following are the number of defects in 15 days: 9,10,12,8,7,15,10,8,7,13,14,15,16. Draw an np-chart and give your findings? [pg.no.26]

24. The following table gives the data on completed sparkplugs for 10 samples each of 100 plugs.

Sample No:	1	2	3	4	5	6	7	8	9	10
No: of defectives	5	12	6	6	3	4	8	3	5	6

Construct a suitable control chart and comment on your result: [pg.no.26]

25. The following table gives the number of defectives in 10 samples, each of size 100. Construct a np-chart for these data and also determine whether the process is in control.

Sample No:	1	2	3	4	5	6	7	8	9	10
No: of defectives	24	38	62	34	26	36	38	52	33	44

[A.U A/M 2019][pg.no.27]

26. 10 Samples each of size 50 were inspected and the numbers of defectives in the inspection were: 2, 1, 1, 2, 3, 5, 5, 1, 2, 3. Draw the np-chart for defectives. [AU N/D 2024][pg.no.28]

PROBLEMS BASED ON p-CHART & np-CHART:

27. The data given below are the number of defectives in 10 samples of 100 items each. Construct a p-chart and an np-chart and comment on the results:

Sample No:	1	2	3	4	5	6	7	8	9	10
No: of defectives	6	16	7	3	8	12	7	11	11	4

[AU N/D 2019] [pg.no.29]

28. The table below presents 10 subgroups of four measurements on the critical dimension of a part produced by a machining process. [A.U N/D 2021][pg.no.30]

Sample No.	1	2	3	4	5	6	7	8	9	10
x1	13	14	11	11	10	10	12	13	13	13
x2	11	14	13	11	12	11	8	15	12	11
x3	13	10	12	13	11	13	11	12	15	14
x4	13	13	15	12	14	13	11	12	12	13

Construct an $X - R$ control chart. Is the process under control?

29. Surface defects have been counted on 25 rectangular steel plates and the data are 1, 0, 4, 3, 1, 2, 5, 0, 2, 1, 1, 0, 8, 0, 2, 1, 3, 5, 4, 6, 3,1, 0, 2 and 4. Set up a control chart for number of conformities using these data. Is the process under control? [A.U N/D 2021][pg.no.30]

30. The data below gives the number of non-conforming bearing and seal assemblies in samples of size 100. Construct a fraction non conforming control chat for these data. Check if the process is under control.

Sample no.:	1	2	3	4	5	6	7	8	9	10
No. of non-conforming assemblies:	7	4	1	3	6	8	10	5	2	7

[A.U N/D 2021][pg.no.31]

31. The following table data show the values of the sample mean $\bar{x}$ and the range R for the samples of size 5 each. Determine whether the process is in control. [A.U N/D 2020][pg.no.32]

Sample No. :	1	2	3	4	5	6	7	8	9	10
Mean $\bar{x}$:	11.2	11.8	10.8	11.6	11.0	9.6	10.4	9.6	10.6	10.0
Range R :	7	4	8	5	7	4	8	4	7	9

Given the conversion factors for $n = 5$ are $A2 = 0.577$, $D3 = 0$ and $D4 = 2.115$.

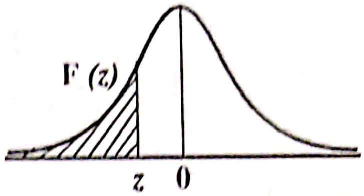
32. The following values represent the number of defectives of 10 samples each containing 100 items. 8, 10, 9, 8, 10, 11, 7, 9, 6, 12 Draw control chart for fraction defective and comment on the state of control of the process [A.U N/D 2020][pg.no.32]

Table : Quality Control - Chart Constants

Sample Size	Chart for average $\bar{X}$-chart			σ-chart — Chart for Standard Deviations					Chart for Ranges — R-chart				
	Factors for Control Limits			Factors for Central line	Factors for Control Limits				Factors for Central line	Factors for Control Limits			
n	A	A_1	A_2	C_2	B_1	B_2	B_3	B_4	d_2	D_1	D_2	D_3	D_4
2	2.121	3.760	1.880	0.5642	0	1.843	0	3.267	1.128	0	3.686	0	3.262
3	1.732	2.394	1.023	0.7236	0	1.858	0	2.568	1.663	0	4.358	0	2.575
4	1.500	1.880	0.729	0.7979	0	1.808	0	2.266	2.059	0	4.698	0	2.282
5	0.342	1.596	0.577	7.8407	0	1.756	0	2.089	2.326	0	4.918	0	2.115
6	1.225	1.410	0.483	0.8686	0.026	0.711	0.030	1.970	2.534	0	5.078	0	2.004
7	1.134	1.277	0.419	0.8882	0.105	1.672	0.118	1.882	2.704	0.205	5.203	0.076	1.924
8	1.061	1.175	0.373	0.9027	0.167	1.638	0.185	1.815	2.847	0.387	5.307	0.136	1.864
9	1.000	1.094	0.337	0.9139	0.219	1.609	0.239	1.760	2.970	0.546	5.394	0.184	1.816
10	0.949	1.028	0.308	0.9227	0.262	1.584	0.284	1.716	3.078	0.687	5.469	0.223	1.777
11	0.905	0.973	0.285	0.9300	0.299	1.561	0.321	1.679	3.173	0.812	5.534	0.256	1.744
12	0.866	0.925	0.266	0.9359	0.331	1.541	0.354	1.646	3.258	0.924	5.592	0.284	1.716
13	0.832	0.884	0.249	0.9410	0.359	1.523	0.382	1.618	3.336	1.026	5.646	0.308	1.692
14	0.802	0.848	0.235	0.9453	0.384	1.507	0.406	1.594	3.407	1.121	5.693	0.329	1.671
15	0.775	0.816	0.223	0.9490	0.406	1.492	0.428	1.572	3.472	1.207	5.737	0.348	1.652
16	0.750	0.788	0.212	0.9523	0.427	1.478	0.448	1.552	3.532	1.285	5.779	0.364	1.636
17	0.72[illegible]	0.762	0.203	0.9551	0.445	1.465	0.466	1534	3.588	1.359	5.817	0.379	1.621
18	0.707	0.738	0.194	0.9576	0.461	1.454	0.482	1.518	3.640	1.426	5.854	0.392	1.608
19	0.688	0.717	0.184	0.9599	0.477	1.443	0.497	1.503	3.689	1.490	5.888	0.404	1.596
20	0.671	0.697	0.110	0.9619	0.491	1.433	0.510	1.490	3.735	1.544	5.922	0.418	1.586

Standard Normal Distribution Function

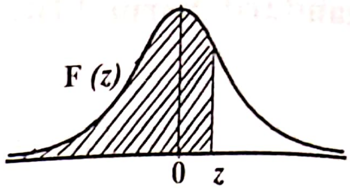
$$F(z) = \frac{1}{\sqrt{2\pi}} \int_{-\infty}^{z} e^{-t^2/2}\, dt$$



z	0.00	0.01	0.02	0.03	0.04	0.05	0.06	0.07	0.08	0.09
−5.0	.0000003									
−4.0	.00003									
−3.5	.0002									
−3.4	.0003	.0003	.0003	.0003	.0003	.0003	.0003	.0003	.0003	.0002
−3.3	.0005	.0005	.0005	.0004	.0004	.0004	.0004	.0004	.0004	.0004
−3.2	.0007	.0007	.0006	.0006	.0006	.0006	.0006	.0005	.0005	.0005
−3.1	.0010	.0009	.0009	.0009	.0008	.0008	.0008	.0008	.0007	.0007
−3.0	.0013	.0013	.0013	.0012	.0012	.0011	.0011	.0011	.0010	.0010
−2.9	.0019	.0018	.0018	.0017	.0016	.0016	.0015	.0015	.0014	.0014
−2.8	.0026	.0025	.0024	.0023	.0023	.0022	.0021	.0021	.0020	.0019
−2.7	.0035	.0034	.0033	.0032	.0031	.0030	.0029	.0028	.0027	.0026
−2.6	.0047	.0045	.0044	.0043	.0041	.0040	.0039	.0038	.0037	.0036
−2.5	.0062	.0060	.0059	.0057	.0055	.0054	.0052	.0051	.0049	.0048
−2.4	.0082	.0080	.0078	.0075	.0073	.0071	.0069	.0068	.0066	.0064
−2.3	.0107	.0104	.0102	.0099	.0096	.0094	.0091	.0089	.0087	.0084
−2.2	.0139	.0136	.0132	.0129	.0125	.0122	.0119	.0116	.0113	.0110
−2.1	.0179	.0174	.0170	.0166	.0162	.0158	.0154	.0150	.0146	.0143
−2.0	.0228	.0222	.0217	.0212	.0207	.0202	.0197	.0192	.0188	.0183
−1.9	.0287	.0281	.0274	.0268	.0262	.0256	.0250	.0244	.0239	.0233
−1.8	.0359	.0351	.0344	.0336	.0329	.0322	.0314	.0307	.0301	.0294
−1.7	.0446	.0436	.0427	.0418	.0409	.0401	.0392	.0384	.0375	.0367
−1.6	.0548	.0537	.0526	.0516	.0505	.0495	.0485	.0475	.0465	.0455
−1.5	.0668	.0655	.0643	.0630	.0618	.0606	.0594	.0582	.0571	.0559
−1.4	.0808	.0793	.0778	.0764	.0749	.0735	.0721	.0708	.0694	.0681
−1.3	.0968	.0951	.0934	.0918	.0901	.0885	.0869	.0853	.0838	.0823
−1.2	.1151	.1131	.1112	.1093	.1075	.1056	.1038	.1020	.1003	.0985
−1.1	.1357	.1335	.1314	.1292	.1271	.1251	.1230	.1210	.1190	.1170
−1.0	.1587	.1562	.1539	.1515	.1492	.1469	.1446	.1423	.1401	.1379
−0.9	.1841	.1814	.1788	.1762	.1736	.1711	.1685	.1660	.1635	.1611
−0.8	.2119	.2090	.2061	.2033	.2005	.1977	.1949	.1922	.1894	.1867
−0.7	.2420	.2389	.2358	.2327	.2296	.2265	.2236	.2206	.2177	.2148
−0.6	.2743	.2709	.2676	.2643	.2611	.2578	.2546	.2514	.2483	.2451
−0.5	.3085	.3050	.3015	.2981	.2946	.2912	.2877	.2843	.2810	.2776
−0.4	.3446	.3409	.3372	.3336	.3300	.3264	.3228	.3192	.3156	.3121
−0.3	.3821	.3783	.3745	.3707	.3669	.3632	.3594	.3557	.3520	.3483
−0.2	.4207	.4168	.4129	.4090	.4052	.4013	.3974	.3936	.3897	.3859
−0.1	.4602	.4562	.4522	.4483	.4443	.4404	.4364	.4325	.4286	.4247
−0.0	.5000	.4960	.4920	.4880	.4840	.4801	.4761	.4721	.4681	.4641

Standard Normal Distribution Function

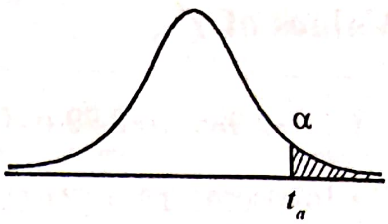
$$F(z) = \frac{1}{\sqrt{2\pi}} \int_{-\infty}^{z} e^{-t^2/2}\, dt$$



z	0.00	0.01	0.02	0.03	0.04	0.05	0.06	0.07	0.08	0.09
0.0	.5000	.5040	.5080	.5120	.5160	.5199	.5239	.5279	.5319	.5359
0.1	.5398	.5438	.5478	.5517	.5557	.5596	.5636	.5675	.5714	.5733
0.2	.5793	.5832	.5871	.5910	.5948	.5987	.6026	.6064	.6103	.6141
0.3	.6179	.6217	.6255	.6293	.6331	.6368	.6406	.6443	.6480	.6517
0.4	.6554	.6591	.6628	.6664	.6700	.6736	.6772	.6808	.6844	.6879
0.5	.6915	.6950	.6985	.7019	.7054	.7088	.7123	.7157	.7190	.7224
0.6	.7257	.7291	.7324	.7357	.7389	.7422	.7454	.7486	.7517	.7549
0.7	.7580	.7611	.7642	.7673	.7704	.7734	.7764	.7794	.7823	.7852
0.8	.7881	.7910	.7939	.7967	.7995	.8023	.8051	.8078	.8106	.8133
0.9	.8159	.8186	.8212	.8238	.8264	.8289	.8315	.8340	.8365	.8389
1.0	.8413	.8438	.8461	.8485	.8508	.8531	.8554	.8577	.8599	.8621
1.1	.8643	.8665	.8686	.8708	.8729	.8749	.8770	.8790	.8810	.8830
1.2	.8849	.8869	.8888	.8907	.8925	.8944	.8962	.8980	.8997	.9015
1.3	.9032	.9049	.9066	.9082	.9099	.9115	.9131	.9147	.9162	.9177
1.4	.9192	.9207	.9222	.9236	.9251	.9265	.9279	.9292	.9306	.9319
1.5	.9332	.9345	.9357	.9370	.9382	.9394	.9406	.9418	.9429	.9441
1.6	.9452	.9463	.9474	.9484	.9495	.9505	.9515	.9525	.9535	.9545
1.7	.9554	.9564	.9573	.9582	.9591	.9599	.9608	.9616	.9625	.9633
1.8	.9641	.9649	.9656	.9664	.9671	.9678	.9686	.9693	.9699	.9706
1.9	.9713	.9719	.9726	.9732	.9738	.9744	.9750	.9756	.9761	.9767
2.0	.9772	.9778	.9783	.9788	.9793	.9798	.9803	.9808	.9812	.9817
2.1	.9821	.9826	.9830	.9834	.9838	.9842	.9846	.9850	.9854	.9857
2.2	.9861	.9864	.9868	.9871	.9875	.9878	.9881	.9884	.9887	.9890
2.3	.9893	.9896	.9898	.9901	.9904	.9906	.9909	.9911	.9913	.9916
2.4	.9918	.9920	.9922	.9925	.9927	.9929	.9931	.9932	.9934	.9936
2.5	.9938	.9940	.9941	.9943	.9945	.9946	.9948	.9949	.9951	.9952
2.6	.9953	.9955	.9956	.9957	.9959	.9960	.9961	.9962	.9963	.9964
2.7	.9965	.9966	.9967	.9968	.9969	.9970	.9971	.9972	.9973	.9974
2.8	.9974	.9975	.9976	.9977	.9977	.9978	.9979	.9979	.9980	.9981
2.9	.9981	.9982	.9982	.9983	.9984	.9984	.9985	.9985	.9986	.9986
3.0	.9987	.9987	.9987	.9988	.9988	.9989	.9989	.9989	.9990	.9990
3.1	.9990	.9991	.9991	.9991	.9992	.9992	.9992	.9992	.9993	.9993
3.2	.9993	.9993	.9994	.9994	.9994	.9994	.9994	.9995	.9995	.9995
3.3	.9995	.9995	.9995	.9996	.9996	.9996	.9996	.9996	.9996	.9997
3.4	.9997	.9997	.9997	.9997	.9997	.9997	.9997	.9997	.9997	.9998
3.5	.9998									
4.0	.99997									
5.0	.9999997									

Values of t_α (one tail)

Note : 5% in one-tail	=	10% in two-tail
0.5% in one-tail	=	1% in two-tail



ν	α=0.10	α=0.05	α=0.025	α=0.01	α=0.00833	α=0.00625	α=0.005	ν
1	3.078	6.314	12.706	31.821	38.204	50.923	63.657	1
2	1.886	2.920	4.303	6.965	7.650	8.860	9.925	2
3	1.638	2.353	3.182	4.541	4.857	5.392	5.841	3
4	1.533	2.132	2.776	3.747	3.961	4.315	4.604	4
5	1.476	2.015	2.571	3.365	3.534	3.810	4.032	5
6	1.440	1.943	2.447	3.143	3.288	3.521	3.707	6
7	1.415	1.895	2.365	2.998	3.128	3.335	3.499	7
8	1.397	1.860	2.306	2.896	3.016	3.206	3.355	8
9	1.383	1.833	2.262	2.821	2.934	3.111	3.250	9
10	1.372	1.812	2.228	2.764	2.870	3.038	3.169	10
11	1.363	1.796	2.201	2.718	2.820	2.981	3.106	11
12	1.356	1.782	2.179	2.681	2.780	2.934	3.055	12
13	1.350	1.771	2.160	2.650	2.746	2.896	3.012	13
14	1.345	1.761	2.145	2.624	2.718	2.864	2.977	14
15	1.341	1.753	2.131	2.602	2.694	2.837	2.947	15
16	1.337	1.746	2.120	2.583	2.673	2.813	2.921	16
17	1.333	1.740	2.110	2.567	2.655	2.793	2.898	17
18	1.330	1.734	2.101	2.552	2.639	2.775	2.878	18
19	1.328	1.729	2.093	2.539	2.625	2.759	2.861	19
20	1.325	1.725	2.086	2.528	2.613	2.744	2.845	20
21	1.323	1.721	2.080	2.518	2.602	2.732	2.831	21
22	1.321	1.717	2.074	2.508	2.591	2.720	2.819	22
23	1.319	1.714	2.069	2.500	2.582	2.710	2.807	23
24	1.318	1.711	2.064	2.492	2.574	2.700	2.797	24
25	1.316	1.708	2.060	2.485	2.566	2.692	2.787	25
26	1.315	1.706	2.056	2.479	2.559	2.684	2.779	26
27	1.314	1.703	2.052	2.473	2.553	2.676	2.771	27
28	1.313	1.701	2.048	2.467	2.547	2.669	2.763	28
29	1.311	1.699	2.045	2.462	2.541	2.663	2.756	29
inf	1.282	1.645	1.960	2.326	2.394	2.493	2.576	inf.

Values of χ^2_α

ν	α=0.995	α=0.99	α=0.975	α=0.95	α=0.05	α=0.025	α=0.01	α=0.005	ν
1	0.0000393	0.000157	0.000982	0.00393	3.841	5.024	6.635	7.879	1
2	0.0100	0.0201	0.0506	0.103	5.991	7.378	9.210	10.597	2
3	0.0717	0.115	0.216	0.352	7.815	9.348	11.345	12.838	3
4	0.207	0.297	0.484	0.711	9.488	11.143	13.277	14.860	4
5	0.412	0.554	0.831	1.145	11.070	12.832	15.056	16.750	5
6	0.676	0.872	1.237	1.635	12.592	14.449	16.812	18.548	6
7	0.989	1.239	1.690	2.167	14.067	16.013	18.475	20.278	7
8	1.344	1.646	2.180	2.733	15.507	17.535	20.090	21.955	8
9	1.735	2.088	2.700	3.325	16.919	19.023	21.666	23.589	9
10	2.156	2.558	3.247	3.940	18.307	20.483	23.209	25.188	10
11	2.603	3.053	3.816	4.575	19.675	21.920	24.725	26.757	11
12	3.074	3.571	4.404	5.226	21.026	23.337	26.217	28.300	12
13	3.565	4.107	5.009	5.892	22.362	24.736	27.688	29.819	13
14	4.075	4.660	5.629	6.571	23.685	26.119	29.141	31.319	14
15	4.601	5.229	6.262	7.261	24.996	27.488	30.578	32.801	15
16	5.142	5.812	6.908	7.962	26.296	28.845	32.000	34.267	16
17	5.697	6.408	7.564	8.672	27.587	30.191	33.409	35.718	17
18	6.265	7.015	8.231	9.390	28.869	31.526	34.805	37.156	18
19	6.844	7.633	8.907	10.117	30.144	32.852	36.191	38.582	19
20	7.434	8.260	9.591	10.851	31.410	34.170	37.566	39.997	20
21	8.034	8.897	10.283	11.591	32.671	35.479	38.932	41.401	21
22	8.643	9.542	10.982	12.338	33.924	36.781	40.289	42.796	22
23	9.260	10.196	11.689	13.091	35.172	38.076	41.638	44.181	23
24	9.886	10.856	12.401	13.484	36.415	39.364	42.980	45.558	24
25	10.520	11.524	13.120	14.611	37.652	40.646	44.314	46.928	25
26	11.160	12.198	13.844	15.379	38.885	41.923	45.642	48.290	26
27	11.808	12.879	14.573	16.151	40.113	43.194	46.963	49.645	27
28	12.461	13.565	15.308	16.928	41.337	44.461	48.278	50.993	28
29	13.121	14.256	16.047	17.708	42.557	45.772	49.588	52.336	29
30	13.787	14.953	16.791	18.493	43.773	46.979	50.892	53.672	30
40	20.706	22.164	24.433	26.509	55.758	59.342	63.691	66.766	40
50	27.991	29.707	32.357	34.764	67.505	71.420	76.154	79.490	50
60	35.535	37.485	40.482	43.118	79.082	83.298	88.379	91.952	60
70	43.275	45.442	48.758	51.739	90.531	95.023	100.425	104.215	70
80	51.172	53.540	57.153	60.391	101.879	106.629	112.329	116.321	80
90	59.196	61.754	65.646	69.126	113.145	118.136	124.116	128.299	90
100	67.328	70.065	74.222	77.929	124.342	129.561	135.807	140.169	100

Values of $F_{0.05}$

v_2 = degree of freedom for denominator	v_1 = Degree of freedom for numerator																		
	1	2	3	4	5	6	7	8	9	10	12	15	20	25	30	40	60	120	∞
1	161	200	216	225	230	234	237	239	241	242	244	246	248	249	250	251	252	253	254
2	18.51	19.00	19.16	19.25	19.30	19.33	19.35	19.37	19.38	19.40	19.41	19.43	19.45	19.46	19.46	19.47	19.48	19.49	19.50
3	10.13	9.55	9.28	9.12	9.01	8.94	8.89	8.85	8.81	8.79	8.74	8.70	8.66	8.63	8.62	8.59	8.57	8.55	8.53
4	7.71	6.94	6.59	6.39	6.26	6.16	6.09	6.04	6.00	5.96	5.91	5.86	5.80	5.77	5.75	5.72	5.69	5.66	5.63
5	6.61	5.79	5.41	5.19	5.05	4.95	4.88	4.82	4.77	4.74	4.68	4.62	4.56	4.52	4.50	4.46	4.43	4.40	4.37
6	5.99	5.14	4.76	4.53	4.39	4.28	4.21	4.15	4.10	4.06	4.00	3.94	3.87	3.83	3.81	3.77	3.74	3.70	3.67
7	5.59	4.74	4.35	4.12	3.97	3.87	3.79	3.73	3.68	3.64	3.57	3.51	3.44	3.40	3.38	3.34	3.30	3.27	3.23
8	5.32	4.46	4.07	3.84	3.69	3.58	3.50	3.44	3.39	3.35	3.28	3.22	3.15	3.11	3.08	3.04	3.01	2.97	2.93
9	5.12	4.26	3.86	3.63	3.48	3.37	3.29	3.23	3.18	3.14	3.07	3.01	2.94	2.89	2.86	2.83	2.79	2.75	2.71
10	4.96	4.10	3.71	3.48	3.33	3.22	3.14	3.07	3.02	2.98	2.91	2.85	2.77	2.73	2.70	2.66	2.62	2.58	2.54
11	4.84	3.98	3.59	3.36	3.20	3.09	3.01	2.95	2.90	2.85	2.79	2.72	2.65	2.60	2.57	2.53	2.49	2.45	2.40
12	4.75	3.89	3.49	3.26	3.11	3.00	2.91	2.85	2.80	2.75	2.69	2.62	2.54	2.50	2.47	2.38	2.38	2.34	2.30
13	4.67	3.81	3.41	3.18	3.03	2.92	2.83	2.77	2.71	2.67	2.60	2.53	2.46	2.41	2.38	2.34	2.30	2.25	2.21
14	4.60	3.74	3.34	3.11	2.96	2.85	2.76	2.70	2.65	2.60	2.53	2.46	2.39	2.34	2.31	2.27	2.22	2.18	2.13
15	4.54	3.68	3.29	3.06	2.90	2.79	2.71	2.64	2.59	2.54	2.48	2.40	2.33	2.28	2.25	2.20	2.16	2.11	2.07
16	4.49	3.63	3.24	3.01	2.85	2.74	2.66	2.59	2.54	2.49	2.42	2.35	2.28	2.23	2.19	2.15	2.11	2.06	2.01
17	4.45	3.59	3.20	2.96	2.81	2.70	2.61	2.55	2.49	2.45	2.38	2.31	2.23	2.18	2.15	2.10	2.06	2.01	1.96
18	4.41	3.55	3.16	2.93	2.77	2.66	2.58	2.51	2.46	2.41	2.34	2.27	2.19	2.14	2.11	2.06	2.02	1.97	1.92
19	4.38	3.52	3.13	2.90	2.74	2.63	2.54	2.48	2.42	2.38	2.31	2.23	2.16	2.11	2.07	2.03	1.98	1.93	1.88
20	4.35	3.49	3.10	2.87	2.71	2.60	2.51	2.45	2.39	2.35	2.28	2.20	2.12	2.07	2.04	1.99	1.95	1.90	1.84
21	4.32	3.47	3.07	2.84	2.68	2.57	2.49	2.42	2.37	2.32	2.25	2.18	2.10	2.05	2.01	1.96	1.92	1.87	1.81
22	4.30	3.44	3.05	2.82	2.66	2.55	2.46	2.40	2.34	2.30	2.23	2.15	2.07	2.02	1.98	1.94	1.89	1.84	1.78
23	4.28	3.42	3.03	2.80	2.64	2.53	2.44	2.37	2.32	2.27	2.20	2.13	2.05	2.00	1.96	1.91	1.86	1.81	1.76
24	4.26	3.40	3.01	2.78	2.62	2.51	2.42	2.36	2.30	2.25	2.18	2.11	2.03	1.97	1.94	1.89	1.84	1.79	1.73
25	4.24	3.39	2.99	2.76	2.60	2.49	2.40	2.34	2.28	2.24	2.16	2.09	2.01	1.96	1.92	1.87	1.82	1.77	1.71
30	4.17	3.32	2.92	2.69	2.53	2.42	2.33	2.27	2.21	2.16	2.09	2.01	1.93	1.88	1.84	1.79	1.74	1.68	1.62
40	4.08	3.23	2.84	2.61	2.45	2.34	2.25	2.18	2.12	2.08	2.00	1.92	1.84	1.78	1.74	1.69	1.64	1.58	1.51
60	4.00	3.15	2.76	2.53	2.37	2.25	2.17	2.10	2.04	1.99	1.92	1.84	1.75	1.69	1.65	1.59	1.53	1.47	1.39
120	3.92	3.07	2.68	2.45	2.29	2.18	2.09	2.02	1.96	1.91	1.83	1.75	1.66	1.60	1.55	1.50	1.43	1.35	1.25
∞	3.84	3.00	2.60	2.37	2.21	2.10	2.01	1.94	1.88	1.83	1.75	1.67	1.57	1.51	1.46	1.39	1.32	1.22	1.00

Values of $F_{0.01}$

v_2 = degree of freedom for denominator	v_1 = Degree of freedom for numerator																		
	1	2	3	4	5	6	7	8	9	10	12	15	20	25	30	40	60	120	∞
1	4,052	5,000	5,403	5,625	5,764	5,859	5,928	5,982	6,023	6,056	6,106	6,157	6,209	6,240	6,261	6,287	6,313	6,339	6,366
2	98.50	99.00	99.17	99.25	99.30	99.33	99.36	99.37	99.39	99.40	99.42	99.43	99.45	99.46	99.57	99.47	99.48	99.49	99.50
3	34.12	30.82	29.46	28.71	28.24	27.91	27.67	27.49	27.35	27.23	27.05	26.87	26.69	26.58	26.50	26.41	26.32	26.22	26.13
4	21.20	18.00	16.69	15.98	15.52	15.21	14.98	14.80	14.66	14.55	14.37	14.20	14.02	13.91	13.84	13.75	13.65	13.56	13.46
5	16.26	13.27	12.06	11.39	10.97	10.67	10.46	10.29	10.16	10.05	9.86	9.72	9.55	9.45	9.38	9.29	9.20	9.11	9.02
6	13.75	10.92	9.78	9.15	8.75	8.47	8.26	8.10	7.98	7.87	7.72	7.56	7.40	7.30	7.23	7.14	7.06	6.97	6.88
7	12.25	9.55	8.45	7.85	7.46	7.19	6.99	6.84	6.72	6.62	6.47	6.31	6.16	6.06	5.99	5.91	5.82	5.74	5.65
8	11.26	8.65	7.59	7.01	6.63	6.37	6.18	6.03	5.91	5.81	5.67	5.52	5.36	5.26	5.20	5.12	5.03	4.95	4.86
9	10.56	8.02	6.99	6.42	6.06	5.80	5.61	5.47	5.35	5.26	5.11	4.96	4.81	4.71	4.65	4.57	4.48	4.40	4.31
10	10.04	7.56	6.55	5.99	5.64	5.39	5.20	5.06	4.94	4.85	4.71	4.56	4.41	4.31	4.25	4.17	4.08	4.00	3.91
11	9.65	7.21	6.22	5.67	5.32	5.07	4.89	4.74	4.63	4.54	4.40	4.25	4.10	4.01	3.94	3.86	3.78	3.69	3.60
12	9.33	6.93	5.95	5.41	5.06	4.82	4.64	4.50	4.39	4.30	4.16	4.01	3.86	3.76	3.70	3.62	3.54	3.45	3.36
13	9.07	6.70	5.74	5.21	4.86	4.62	4.44	4.30	4.19	4.10	3.96	3.82	3.66	3.57	3.51	3.43	3.34	3.25	3.17
14	8.86	6.51	5.56	5.04	4.69	4.46	4.28	4.14	4.03	3.94	3.80	3.66	3.51	3.41	3.35	3.27	3.18	3.09	3.00
15	8.68	6.36	5.42	4.89	4.56	4.32	4.14	4.00	3.89	3.80	3.67	3.52	3.37	3.28	3.21	3.13	3.05	2.96	2.87
16	8.53	6.23	5.29	4.77	4.44	4.20	4.03	3.89	3.78	3.69	3.55	3.41	3.26	3.16	3.10	3.02	2.93	2.84	2.75
17	8.40	6.11	5.18	4.67	4.34	4.10	3.93	3.79	3.68	3.59	3.46	3.31	3.16	3.07	3.00	2.92	2.83	2.75	2.65
18	8.29	6.01	5.09	4.58	4.25	4.01	3.84	3.71	3.60	3.51	3.37	3.23	3.08	2.98	2.92	2.84	2.75	2.66	2.57
19	8.18	5.93	5.01	4.50	4.17	3.94	3.77	3.63	3.52	3.43	3.30	3.15	3.00	2.91	2.84	2.76	2.67	2.58	2.49
20	8.10	5.85	4.94	4.43	4.10	3.87	3.70	3.56	3.46	3.37	3.23	3.09	2.94	2.84	2.78	2.69	2.61	2.52	2.42
21	8.02	5.78	4.87	4.37	4.04	3.81	3.64	3.51	3.40	3.31	3.17	3.03	2.88	2.79	2.72	2.64	2.55	2.46	2.36
22	7.95	5.72	4.82	4.31	3.99	3.76	3.59	3.45	3.35	3.26	3.12	2.98	2.83	2.73	2.67	2.58	2.50	2.40	2.31
23	7.88	5.66	4.76	4.26	3.94	3.71	3.54	3.41	3.30	3.21	3.07	2.93	2.78	2.69	2.62	2.54	2.45	2.35	2.26
24	7.82	5.61	4.72	4.22	3.90	3.67	3.50	3.36	3.26	3.17	3.03	2.89	2.74	2.64	2.58	2.49	2.40	2.31	2.21
25	7.77	5.57	4.68	4.18	3.85	3.63	3.46	3.32	3.22	3.13	2.99	2.85	2.70	2.60	2.54	2.45	2.36	2.27	2.17
30	7.56	5.39	4.51	4.02	3.70	3.47	3.30	3.17	3.07	2.98	2.84	2.70	2.55	2.45	2.39	2.30	2.21	2.11	2.01
40	7.31	5.18	4.31	3.83	3.51	3.29	3.12	2.99	2.89	2.80	2.66	2.52	2.37	2.27	2.20	2.11	2.02	1.92	1.80
60	7.08	4.98	4.13	3.65	3.34	3.12	2.95	2.82	2.72	2.63	2.50	2.35	2.20	2.10	2.03	1.94	1.84	1.73	1.60
120	6.85	4.79	3.95	3.48	3.17	2.96	2.79	2.66	2.56	2.47	2.34	2.19	2.03	1.93	1.86	1.76	1.66	1.53	1.38
∞	6.63	4.61	3.78	3.32	3.02	2.80	2.64	2.51	2.41	2.32	2.18	2.04	1.88	1.77	1.70	1.59	1.47	1.32	1.00

TABLE

Area under the Standard Normal Probability Distribution between the Mean and Positive Values of z^*

z	.00	.01	.02	.03	.04	.05	.06	.07	.08	.09
0.0	.0000	.0040	.0080	.1200	.0160	.0199	.0239	.0279	.0319	.0359
0.1	.0398	.0438	.0478	.0517	.0557	.0596	.0636	.0675	.0714	.0753
0.2	.0793	.0832	.0871	.0910	.0948	.0987	.1026	.1064	.1103	.1141
0.3	.1179	.1217	.1255	.1293	.1331	.1368	.1406	.1443	.1480	.1517
0.4	.1554	.1591	.1628	.1664	.1700	.1736	.1772	.1808	.1844	.1879
0.5	.1915	.1950	.1985	.2019	.2054	.2088	.2123	.2157	.2190	.2224
0.6	.2257	.2291	.2324	.2357	.2389	.2422	.2454	.2486	.2517	.2549
0.7	.2580	.2611	.2642	.2673	.2704	.2734	.2764	.2794	.2823	.2852
0.8	.2881	.2910	.2939	.2967	.2995	.3023	.3051	.3078	.3106	.3133
0.9	.3159	.3186	.3212	.3238	.3264	.3289	.3315	.3340	.3365	.3389
1.0	.3413	.3438	.3461	.3485	.3508	.3531	.3554	.3577	.3599	.3621
1.1	.3643	.3665	.3686	.3708	.3729	.3749	.3770	.3790	.3810	.3830
1.2	.3849	.3869	.3888	.3907	.3925	.3944	.3962	.3980	.3997	.4015
1.3	.4032	.4049	.4066	.4082	.4099	.4115	.4131	.4147	.4162	.4177
1.4	.4192	.4207	.4222	.4236	.4251	.4265	.4279	.4292	.4306	.4319
1.5	.4332	.4345	.4357	.4370	.4382	.4394	.4406	.4418	.4429	.4441
1.6	.4452	.4463	.4474	.4484	.4495	.4505	.4515	.4525	.4535	.4545
1.7	.4554	.4564	.4573	.4582	.4591	.4599	.4608	.4616	.4625	.4633
1.8	.4641	.4649	.4656	.4664	.4671	.4678	.4686	.4693	.4699	.4706
1.9	.4713	.4719	.4726	.4732	.4738	.4744	.4750	.4756	.4761	.4767
2.0	.4772	.4778	.4783	.4788	.4793	.4798	.4803	.4808	.4812	.4817
2.1	.4821	.4826	.4830	.4834	.4838	.4842	.4886	.4850	.4854	.4857
2.2	.4861	.4864	.4868	.4871	.4875	.4878	.4881	.4884	.4887	.4890
2.3	.4893	.4896	.4898	.4901	.4904	.4906	.4909	.4911	.4913	.4916
2.4	.4918	.4920	.4922	.4925	.4927	.4929	.4931	.4932	.4934	.4936
2.5	.4938	.4940	.4941	.4943	.4945	.4946	.4948	.4949	.4951	.4952
2.6	.4953	.4955	.4956	.4957	.4959	.4960	.4961	.4962	.4963	.4964
2.7	.4965	.4966	.4967	.4968	.4969	.4970	.4971	.4972	.4973	.4974
2.8	.4974	.4975	.4976	.4977	.4978	.4979	.4979	.4979	.4980	.4981
2.9	.4981	.4982	.4982	.4983	.4984	.4984	.4985	.4985	.4986	.4986
3.0	.4987	.4987	.4987	.4988	.4988	.4989	.4989	.4989	.4990	.4990

$$B(x;n,p) = \sum_{k=0}^{x} \binom{n}{k} p^k (1-p)^{n-k}$$

											p									
n	**x**	**0.05**	**0.10**	**0.15**	**0.20**	**0.25**	**0.30**	**0.35**	**0.40**	**0.45**	**0.50**	**0.55**	**0.60**	**0.65**	**0.70**	**0.75**	**0.80**	**0.85**	**0.90**	**0.95**
2	0	0.9025	0.8100	0.7225	0.6400	0.5625	0.4900	0.4225	0.3600	0.3025	0.2500	0.2025	0.1600	0.1225	0.0900	0.0625	0.0400	0.0225	0.0100	0.0025
	1	0.9975	0.9900	0.9775	0.9600	0.9375	0.9100	0.8775	0.8400	0.7975	0.7500	0.6975	0.6400	0.5775	0.5100	0.4375	0.3600	0.2775	0.1900	0.0975
3	0	0.8574	0.7290	0.6141	0.5120	0.4219	0.3430	0.2746	0.2160	0.1664	0.1250	0.0911	0.0640	0.0429	0.0270	0.0156	0.0080	0.0034	0.0010	0.0001
	1	0.9927	0.9720	0.9393	0.8960	0.8438	0.7840	0.7183	0.6480	0.5748	0.5000	0.4252	0.3520	0.2818	0.2160	0.1563	0.1040	0.0607	0.0280	0.0073
	2	0.9999	0.9990	0.9966	0.9920	0.9844	0.9730	0.9571	0.9360	0.9089	0.8750	0.8336	0.7840	0.7254	0.6570	0.5781	0.4880	0.3859	0.2710	0.1426
4	0	0.8145	0.6561	0.5220	0.4096	0.3164	0.2401	0.1785	0.1296	0.0915	0.0625	0.0410	0.0256	0.0150	0.0081	0.0039	0.0016	0.0005	0.0001	0.0000
	1	0.9860	0.9477	0.8905	0.8192	0.7383	0.6517	0.5630	0.4752	0.3910	0.3125	0.2415	0.1792	0.1265	0.0837	0.0508	0.0272	0.0120	0.0037	0.0005
	2	0.9995	0.9963	0.9880	0.9728	0.9492	0.9163	0.8735	0.8208	0.7585	0.6875	0.6090	0.5248	0.4370	0.3483	0.2617	0.1808	0.1095	0.0523	0.0140
	3	1.0000	0.9999	0.9995	0.9984	0.9961	0.9919	0.9850	0.9744	0.9590	0.9375	0.9085	0.8704	0.8215	0.7599	0.6836	0.5904	0.4780	0.3439	0.1855
5	0	0.7738	0.5905	0.4437	0.3277	0.2373	0.1681	0.1160	0.0778	0.0503	0.0313	0.0185	0.0102	0.0053	0.0024	0.0010	0.0003	0.0001	0.0000	0.0000
	1	0.9774	0.9185	0.8352	0.7373	0.6328	0.5282	0.4284	0.3370	0.2562	0.1875	0.1312	0.0870	0.0540	0.0308	0.0156	0.0067	0.0022	0.0005	0.0000
	2	0.9988	0.9914	0.9734	0.9421	0.8965	0.8369	0.7648	0.6826	0.5931	0.5000	0.4069	0.3174	0.2352	0.1631	0.1035	0.0579	0.0266	0.0086	0.0012
	3	1.0000	0.9995	0.9978	0.9933	0.9844	0.9692	0.9460	0.9130	0.8688	0.8125	0.7438	0.6630	0.5716	0.4718	0.3672	0.2627	0.1648	0.0815	0.0226
	4	1.0000	1.0000	0.9999	0.9997	0.9990	0.9976	0.9947	0.9898	0.9815	0.9688	0.9497	0.9222	0.8840	0.8319	0.7627	0.6723	0.5563	0.4095	0.2262
6	0	0.7351	0.5314	0.3771	0.2621	0.1780	0.1176	0.0754	0.0467	0.0277	0.0156	0.0083	0.0041	0.0018	0.0007	0.0002	0.0001	0.0000	0.0000	0.0000
	1	0.9672	0.8857	0.7765	0.6554	0.5339	0.4202	0.3191	0.2333	0.1636	0.1094	0.0692	0.0410	0.0223	0.0109	0.0046	0.0016	0.0004	0.0001	0.0000
	2	0.9978	0.9841	0.9527	0.9011	0.8306	0.7443	0.6471	0.5443	0.4415	0.3438	0.2553	0.1792	0.1174	0.0705	0.0376	0.0170	0.0059	0.0013	0.0001
	3	0.9999	0.9987	0.9941	0.9830	0.9624	0.9295	0.8826	0.8208	0.7447	0.6563	0.5585	0.4557	0.3529	0.2557	0.1694	0.0989	0.0473	0.0158	0.0022
	4	1.0000	0.9999	0.9996	0.9984	0.9954	0.9891	0.9777	0.9590	0.9308	0.8906	0.8364	0.7667	0.6809	0.5798	0.4661	0.3446	0.2235	0.1143	0.0328
	5	1.0000	1.0000	1.0000	0.9999	0.9998	0.9993	0.9982	0.9959	0.9917	0.9844	0.9723	0.9533	0.9246	0.8824	0.8220	0.7379	0.6229	0.4686	0.2649
7	0	0.6983	0.4783	0.3206	0.2097	0.1335	0.0824	0.0490	0.0280	0.0152	0.0078	0.0037	0.0016	0.0006	0.0002	0.0001	0.0000	0.0000	0.0000	0.0000
	1	0.9556	0.8503	0.7166	0.5767	0.4449	0.3294	0.2338	0.1586	0.1024	0.0625	0.0357	0.0188	0.0090	0.0038	0.0013	0.0004	0.0001	0.0000	0.0000
	2	0.9962	0.9743	0.9262	0.8520	0.7564	0.6471	0.5323	0.4199	0.3164	0.2266	0.1529	0.0963	0.0556	0.0288	0.0129	0.0047	0.0012	0.0002	0.0000
	3	0.9998	0.9973	0.9879	0.9667	0.9294	0.8740	0.8002	0.7102	0.6083	0.5000	0.3917	0.2898	0.1998	0.1260	0.0706	0.0333	0.0121	0.0027	0.0002
	4	1.0000	0.9998	0.9988	0.9953	0.9871	0.9712	0.9444	0.9037	0.8471	0.7734	0.6836	0.5801	0.4677	0.3529	0.2436	0.1480	0.0738	0.0257	0.0038
	5	1.0000	1.0000	0.9999	0.9996	0.9987	0.9962	0.9910	0.9812	0.9643	0.9375	0.8976	0.8414	0.7662	0.6706	0.5551	0.4233	0.2834	0.1497	0.0344
	6	1.0000	1.0000	1.0000	1.0000	0.9999	0.9998	0.9994	0.9984	0.9963	0.9922	0.9848	0.9720	0.9510	0.9176	0.8665	0.7903	0.6794	0.5217	0.3017
8	0	0.6634	0.4305	0.2775	0.1678	0.1001	0.0576	0.0319	0.0168	0.0084	0.0039	0.0017	0.0007	0.0002	0.0001	0.0000	0.0000	0.0000	0.0000	0.0000
	1	0.9428	0.8131	0.6572	0.5033	0.3671	0.2553	0.1691	0.1064	0.0632	0.0352	0.0181	0.0085	0.0036	0.0013	0.0004	0.0001	0.0000	0.0000	0.0000
	2	0.9942	0.9619	0.8948	0.7969	0.6785	0.5518	0.4278	0.3154	0.2201	0.1445	0.0885	0.0498	0.0253	0.0113	0.0042	0.0012	0.0002	0.0000	0.0000
	3	0.9996	0.9950	0.9786	0.9437	0.8862	0.8059	0.7064	0.5941	0.4770	0.3633	0.2604	0.1737	0.1061	0.0580	0.0273	0.0104	0.0029	0.0004	0.0000
	4	1.0000	0.9996	0.9971	0.9896	0.9727	0.9420	0.8939	0.8263	0.7396	0.6367	0.5230	0.4059	0.2936	0.1941	0.1138	0.0563	0.0214	0.0050	0.0004
	5	1.0000	1.0000	0.9998	0.9988	0.9958	0.9887	0.9747	0.9502	0.9115	0.8555	0.7799	0.6846	0.5722	0.4482	0.3215	0.2031	0.1052	0.0381	0.0058
	6	1.0000	1.0000	1.0000	0.9999	0.9996	0.9987	0.9964	0.9915	0.9819	0.9648	0.9368	0.8936	0.8309	0.7447	0.6329	0.4967	0.3428	0.1869	0.0572
	7	1.0000	1.0000	1.0000	1.0000	1.0000	0.9999	0.9998	0.9993	0.9983	0.9961	0.9916	0.9832	0.9681	0.9424	0.8999	0.8322	0.7275	0.5695	0.3366

n	x	p = 0.05	0.10	0.15	0.20	0.25	0.30	0.35	0.40	0.45	0.50	0.55	0.60	0.65	0.70	0.75	0.80	0.85	0.90	0.95
9	0	0.6302	0.3874	0.2316	0.1342	0.0751	0.0404	0.0207	0.0101	0.0046	0.0020	0.0008	0.0003	0.0001	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000
	1	0.9288	0.7748	0.5995	0.4362	0.3003	0.1960	0.1211	0.0705	0.0385	0.0195	0.0091	0.0038	0.0014	0.0004	0.0001	0.0000	0.0000	0.0000	0.0000
	2	0.9916	0.9470	0.8591	0.7382	0.6007	0.4628	0.3373	0.2318	0.1495	0.0898	0.0498	0.0250	0.0112	0.0043	0.0013	0.0003	0.0000	0.0000	0.0000
	3	0.9994	0.9917	0.9661	0.9144	0.8343	0.7297	0.6089	0.4826	0.3614	0.2539	0.1658	0.0994	0.0536	0.0253	0.0100	0.0031	0.0006	0.0001	0.0000
	4	1.0000	0.9991	0.9944	0.9804	0.9511	0.9012	0.8283	0.7334	0.6214	0.5000	0.3786	0.2666	0.1717	0.0988	0.0489	0.0196	0.0056	0.0009	0.0000
	5	1.0000	0.9999	0.9994	0.9969	0.9900	0.9747	0.9464	0.9006	0.8342	0.7461	0.6386	0.5174	0.3911	0.2703	0.1657	0.0856	0.0339	0.0083	0.0006
	6	1.0000	1.0000	1.0000	0.9997	0.9987	0.9957	0.9888	0.9750	0.9502	0.9102	0.8505	0.7682	0.6627	0.5372	0.3993	0.2618	0.1409	0.0530	0.0084
	7	1.0000	1.0000	1.0000	1.0000	0.9999	0.9996	0.9986	0.9962	0.9909	0.9805	0.9615	0.9295	0.8789	0.8040	0.6997	0.5638	0.4005	0.2252	0.0712
	8	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000	0.9999	0.9997	0.9992	0.9980	0.9954	0.9899	0.9793	0.9596	0.9249	0.8658	0.7684	0.6126	0.3698
10	0	0.5987	0.3487	0.1969	0.1074	0.0563	0.0282	0.0135	0.0060	0.0025	0.0010	0.0003	0.0001	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000
	1	0.9139	0.7361	0.5443	0.3758	0.2440	0.1493	0.0860	0.0464	0.0233	0.0107	0.0045	0.0017	0.0005	0.0001	0.0000	0.0000	0.0000	0.0000	0.0000
	2	0.9885	0.9298	0.8202	0.6778	0.5256	0.3828	0.2616	0.1673	0.0996	0.0547	0.0274	0.0123	0.0048	0.0016	0.0004	0.0001	0.0000	0.0000	0.0000
	3	0.9990	0.9872	0.9500	0.8791	0.7759	0.6496	0.5138	0.3823	0.2660	0.1719	0.1020	0.0548	0.0260	0.0106	0.0035	0.0009	0.0001	0.0000	0.0000
	4	0.9999	0.9984	0.9901	0.9672	0.9219	0.8497	0.7515	0.6331	0.5044	0.3770	0.2616	0.1662	0.0949	0.0473	0.0197	0.0064	0.0014	0.0001	0.0000
	5	1.0000	0.9999	0.9986	0.9936	0.9803	0.9527	0.9051	0.8338	0.7384	0.6230	0.4956	0.3669	0.2485	0.1503	0.0781	0.0328	0.0099	0.0016	0.0001
	6	1.0000	1.0000	0.9999	0.9991	0.9965	0.9894	0.9740	0.9452	0.8980	0.8281	0.7340	0.6177	0.4862	0.3504	0.2241	0.1209	0.0500	0.0128	0.0010
	7	1.0000	1.0000	1.0000	0.9999	0.9996	0.9984	0.9952	0.9877	0.9726	0.9453	0.9004	0.8327	0.7384	0.6172	0.4744	0.3222	0.1798	0.0702	0.0115
	8	1.0000	1.0000	1.0000	1.0000	1.0000	0.9999	0.9995	0.9983	0.9955	0.9893	0.9767	0.9536	0.9140	0.8507	0.7560	0.6242	0.4557	0.2639	0.0861
	9	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000	0.9999	0.9997	0.9990	0.9975	0.9940	0.9865	0.9718	0.9437	0.8926	0.8031	0.6513	0.4023
11	0	0.5688	0.3138	0.1673	0.0859	0.0422	0.0198	0.0088	0.0036	0.0014	0.0005	0.0002	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000
	1	0.8981	0.6974	0.4922	0.3221	0.1971	0.1130	0.0606	0.0302	0.0139	0.0059	0.0022	0.0007	0.0002	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000
	2	0.9848	0.9104	0.7788	0.6174	0.4552	0.3127	0.2001	0.1189	0.0652	0.0327	0.0148	0.0059	0.0020	0.0006	0.0001	0.0000	0.0000	0.0000	0.0000
	3	0.9984	0.9815	0.9306	0.8389	0.7133	0.5696	0.4256	0.2963	0.1911	0.1133	0.0610	0.0293	0.0122	0.0043	0.0012	0.0002	0.0000	0.0000	0.0000
	4	0.9999	0.9972	0.9841	0.9496	0.8854	0.7897	0.6683	0.5328	0.3971	0.2744	0.1738	0.0994	0.0501	0.0216	0.0076	0.0020	0.0003	0.0000	0.0000
	5	1.0000	0.9997	0.9973	0.9883	0.9657	0.9218	0.8513	0.7535	0.6331	0.5000	0.3669	0.2465	0.1487	0.0782	0.0343	0.0117	0.0027	0.0003	0.0000
	6	1.0000	1.0000	0.9997	0.9980	0.9924	0.9784	0.9499	0.9006	0.8262	0.7256	0.6029	0.4672	0.3317	0.2103	0.1146	0.0504	0.0159	0.0028	0.0001
	7	1.0000	1.0000	1.0000	0.9998	0.9988	0.9957	0.9878	0.9707	0.9390	0.8867	0.8089	0.7037	0.5744	0.4304	0.2867	0.1611	0.0694	0.0185	0.0016
	8	1.0000	1.0000	1.0000	1.0000	0.9999	0.9994	0.9980	0.9941	0.9852	0.9673	0.9348	0.8811	0.7999	0.6873	0.5448	0.3826	0.2212	0.0896	0.0152
	9	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000	0.9998	0.9993	0.9978	0.9941	0.9861	0.9698	0.9394	0.8870	0.8029	0.6779	0.5078	0.3026	0.1019
	10	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000	0.9998	0.9995	0.9986	0.9964	0.9912	0.9802	0.9578	0.9141	0.8327	0.6862	0.4312
12	0	0.5404	0.2824	0.1422	0.0687	0.0317	0.0138	0.0057	0.0022	0.0008	0.0002	0.0001	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000
	1	0.8816	0.6590	0.4435	0.2749	0.1584	0.0850	0.0424	0.0196	0.0083	0.0032	0.0011	0.0003	0.0001	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000
	2	0.9804	0.8891	0.7358	0.5583	0.3907	0.2528	0.1513	0.0834	0.0421	0.0193	0.0079	0.0028	0.0008	0.0002	0.0000	0.0000	0.0000	0.0000	0.0000
	3	0.9978	0.9744	0.9078	0.7946	0.6488	0.4925	0.3467	0.2253	0.1345	0.0730	0.0356	0.0153	0.0056	0.0017	0.0004	0.0001	0.0000	0.0000	0.0000
	4	0.9998	0.9957	0.9761	0.9274	0.8424	0.7237	0.5833	0.4382	0.3044	0.1938	0.1117	0.0573	0.0255	0.0095	0.0028	0.0006	0.0001	0.0000	0.0000
	5	1.0000	0.9995	0.9954	0.9806	0.9456	0.8822	0.7873	0.6652	0.5269	0.3872	0.2607	0.1582	0.0846	0.0386	0.0143	0.0039	0.0007	0.0001	0.0000
	6	1.0000	0.9999	0.9993	0.9961	0.9857	0.9614	0.9154	0.8418	0.7393	0.6128	0.4731	0.3348	0.2127	0.1178	0.0544	0.0194	0.0046	0.0005	0.0000
	7	1.0000	1.0000	0.9999	0.9994	0.9972	0.9905	0.9745	0.9427	0.8883	0.8062	0.6956	0.5618	0.4167	0.2763	0.1576	0.0726	0.0239	0.0043	0.0002
	8	1.0000	1.0000	1.0000	0.9999	0.9996	0.9983	0.9944	0.9847	0.9644	0.9270	0.8655	0.7747	0.6533	0.5075	0.3512	0.2054	0.0922	0.0256	0.0022
	9	1.0000	1.0000	1.0000	1.0000	1.0000	0.9998	0.9992	0.9972	0.9921	0.9807	0.9579	0.9166	0.8487	0.7472	0.6093	0.4417	0.2642	0.1109	0.0196
	10	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000	0.9999	0.9997	0.9989	0.9968	0.9917	0.9804	0.9576	0.9150	0.8416	0.7251	0.5565	0.3410	0.1184
	11	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000	0.9999	0.9998	0.9992	0.9978	0.9943	0.9862	0.9683	0.9313	0.8578	0.7176	0.4596

n	x	0.05	0.10	0.15	0.20	0.25	0.30	0.35	0.40	0.45	p 0.50	0.55	0.60	0.65	0.70	0.75	0.80	0.85	0.90	0.95
13	0	0.5133	0.2542	0.1209	0.0550	0.0238	0.0097	0.0037	0.0013	0.0004	0.0001	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000
	1	0.8646	0.6213	0.3983	0.2336	0.1267	0.0637	0.0296	0.0126	0.0049	0.0017	0.0005	0.0001	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000
	2	0.9755	0.8661	0.6920	0.5017	0.3326	0.2025	0.1132	0.0579	0.0269	0.0112	0.0041	0.0013	0.0003	0.0001	0.0000	0.0000	0.0000	0.0000	0.0000
	3	0.9969	0.9658	0.8820	0.7473	0.5843	0.4206	0.2783	0.1686	0.0929	0.0461	0.0203	0.0078	0.0025	0.0007	0.0001	0.0000	0.0000	0.0000	0.0000
	4	0.9997	0.9935	0.9658	0.9009	0.7940	0.6543	0.5005	0.3530	0.2279	0.1334	0.0698	0.0321	0.0126	0.0040	0.0010	0.0002	0.0000	0.0000	0.0000
	5	1.0000	0.9991	0.9925	0.9700	0.9198	0.8346	0.7159	0.5744	0.4268	0.2905	0.1788	0.0977	0.0462	0.0182	0.0056	0.0012	0.0002	0.0000	0.0000
	6	1.0000	0.9999	0.9987	0.9930	0.9757	0.9376	0.8705	0.7712	0.6437	0.5000	0.3563	0.2288	0.1295	0.0624	0.0243	0.0070	0.0013	0.0001	0.0000
	7	1.0000	1.0000	0.9998	0.9988	0.9944	0.9818	0.9538	0.9023	0.8212	0.7095	0.5732	0.4256	0.2841	0.1654	0.0802	0.0300	0.0075	0.0009	0.0000
	8	1.0000	1.0000	1.0000	0.9998	0.9990	0.9960	0.9874	0.9679	0.9302	0.8666	0.7721	0.6470	0.4995	0.3457	0.2060	0.0991	0.0342	0.0065	0.0003
	9	1.0000	1.0000	1.0000	1.0000	0.9999	0.9993	0.9975	0.9922	0.9797	0.9539	0.9071	0.8314	0.7217	0.5794	0.4157	0.2527	0.1180	0.0342	0.0031
	10	1.0000	1.0000	1.0000	1.0000	1.0000	0.9999	0.9997	0.9987	0.9959	0.9888	0.9731	0.9421	0.8868	0.7975	0.6674	0.4983	0.3080	0.1339	0.0245
	11	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000	0.9999	0.9995	0.9983	0.9951	0.9874	0.9704	0.9363	0.8733	0.7664	0.6017	0.3787	0.1354
	12	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000	0.9999	0.9996	0.9987	0.9963	0.9903	0.9762	0.9450	0.8791	0.7458	0.4867
14	0	0.4877	0.2288	0.1028	0.0440	0.0178	0.0068	0.0024	0.0008	0.0002	0.0001	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000
	1	0.8470	0.5846	0.3567	0.1979	0.1010	0.0475	0.0205	0.0081	0.0029	0.0009	0.0003	0.0001	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000
	2	0.9699	0.8416	0.6479	0.4481	0.2811	0.1608	0.0839	0.0398	0.0170	0.0065	0.0022	0.0006	0.0001	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000
	3	0.9958	0.9559	0.8535	0.6982	0.5213	0.3552	0.2205	0.1243	0.0632	0.0287	0.0114	0.0039	0.0011	0.0002	0.0000	0.0000	0.0000	0.0000	0.0000
	4	0.9996	0.9908	0.9533	0.8702	0.7415	0.5842	0.4227	0.2793	0.1672	0.0898	0.0426	0.0175	0.0060	0.0017	0.0003	0.0000	0.0000	0.0000	0.0000
	5	1.0000	0.9985	0.9885	0.9561	0.8883	0.7805	0.6405	0.4859	0.3373	0.2120	0.1189	0.0583	0.0243	0.0083	0.0022	0.0004	0.0000	0.0000	0.0000
	6	1.0000	0.9998	0.9978	0.9884	0.9617	0.9067	0.8164	0.6925	0.5461	0.3953	0.2586	0.1501	0.0753	0.0315	0.0103	0.0024	0.0003	0.0000	0.0000
	7	1.0000	1.0000	0.9997	0.9976	0.9897	0.9685	0.9247	0.8499	0.7414	0.6047	0.4539	0.3075	0.1836	0.0933	0.0383	0.0116	0.0022	0.0002	0.0000
	8	1.0000	1.0000	1.0000	0.9996	0.9978	0.9917	0.9757	0.9417	0.8811	0.7880	0.6627	0.5141	0.3595	0.2195	0.1117	0.0439	0.0115	0.0015	0.0000
	9	1.0000	1.0000	1.0000	1.0000	0.9997	0.9983	0.9940	0.9825	0.9574	0.9102	0.8328	0.7207	0.5773	0.4158	0.2585	0.1298	0.0467	0.0092	0.0004
	10	1.0000	1.0000	1.0000	1.0000	1.0000	0.9998	0.9989	0.9961	0.9886	0.9713	0.9368	0.8757	0.7795	0.6448	0.4787	0.3018	0.1465	0.0441	0.0042
	11	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000	0.9999	0.9994	0.9978	0.9935	0.9830	0.9602	0.9161	0.8392	0.7189	0.5519	0.3521	0.1584	0.0301
	12	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000	0.9999	0.9997	0.9991	0.9971	0.9919	0.9795	0.9525	0.8990	0.8021	0.6433	0.4154	0.1530
	13	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000	0.9999	0.9998	0.9992	0.9976	0.9932	0.9822	0.9560	0.8972	0.7712	0.5123
15	0	0.4633	0.2059	0.0874	0.0352	0.0134	0.0047	0.0016	0.0005	0.0001	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000
	1	0.8290	0.5490	0.3186	0.1671	0.0802	0.0353	0.0142	0.0052	0.0017	0.0005	0.0001	0.0000	0.0090	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000
	2	0.9638	0.8159	0.6042	0.3980	0.2361	0.1268	0.0617	0.0271	0.0107	0.0037	0.0011	0.0003	0.0001	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000
	3	0.9945	0.9444	0.8227	0.6482	0.4613	0.2969	0.1727	0.0905	0.0424	0.0176	0.0063	0.0019	0.0005	0.0001	0.0000	0.0000	0.0000	0.0000	0.0000
	4	0.9994	0.9873	0.9383	0.8358	0.6865	0.5155	0.3519	0.2173	0.1204	0.0592	0.0255	0.0093	0.0028	0.0007	0.0001	0.0000	0.0000	0.0000	0.0000
	5	0.9999	0.9977	0.9832	0.9389	0.8516	0.7216	0.5643	0.4032	0.2608	0.1509	0.0769	0.0338	0.0124	0.0037	0.0008	0.0001	0.0000	0.0000	0.0000
	6	1.0000	0.9997	0.9964	0.9819	0.9434	0.8689	0.7548	0.6098	0.4522	0.3036	0.1818	0.0950	0.0422	0.0152	0.0042	0.0008	0.0001	0.0000	0.0000
	7	1.0000	1.0000	0.9994	0.9958	0.9827	0.9500	0.8868	0.7869	0.6535	0.5000	0.3465	0.2131	0.1132	0.0500	0.0173	0.0042	0.0006	0.0000	0.0000
	8	1.0000	1.0000	0.9999	0.9992	0.9958	0.9848	0.9578	0.9050	0.8182	0.6964	0.5478	0.3902	0.2452	0.1311	0.0566	0.0181	0.0036	0.0003	0.0000
	9	1.0000	1.0000	1.0000	0.9999	0.9992	0.9963	0.9876	0.9662	0.9231	0.8491	0.7392	0.5968	0.4357	0.2784	0.1484	0.0611	0.0168	0.0022	0.0001
	10	1.0000	1.0000	1.0000	1.0000	0.9999	0.9993	0.9972	0.9907	0.9745	0.9408	0.8796	0.7827	0.6481	0.4845	0.3135	0.1642	0.0617	0.0127	0.0006
	11	1.0000	1.0000	1.0000	1.0000	1.0000	0.9999	0.9995	0.9981	0.9937	0.9824	0.9576	0.9095	0.8273	0.7031	0.5387	0.3518	0.1773	0.0556	0.0055
	12	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000	0.9999	0.9997	0.9989	0.9963	0.9893	0.9729	0.9383	0.8732	0.7639	0.6020	0.3958	0.1841	0.0362
	13	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000	0.9999	0.9995	0.9983	0.9948	0.9858	0.9647	0.9198	0.8329	0.6814	0.4510	0.1710
	14	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000	0.9999	0.9995	0.9984	0.9953	0.9866	0.9648	0.9126	0.7941	0.5367

Table-4

Binomial Distribution Function

											p									
n	x	0.05	0.10	0.15	0.20	0.25	0.30	0.35	0.40	0.45	0.50	0.55	0.60	0.65	0.70	0.75	0.80	0.85	0.90	0.95
16	0	0.4401	0.1853	0.0743	0.0281	0.0100	0.0033	0.0010	0.0003	0.0001	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000
	1	0.8108	0.5147	0.2839	0.1407	0.0635	0.0261	0.0098	0.0033	0.0010	0.0003	0.0001	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000
	2	0.9571	0.7892	0.5614	0.3518	0.1971	0.0994	0.0451	0.0183	0.0066	0.0021	0.0006	0.0001	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000
	3	0.9930	0.9316	0.7899	0.5981	0.4050	0.2459	0.1339	0.0651	0.0281	0.0106	0.0035	0.0009	0.0002	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000
	4	0.9991	0.9830	0.9209	0.7982	0.6302	0.4499	0.2892	0.1666	0.0853	0.0384	0.0149	0.0049	0.0013	0.0003	0.0000	0.0000	0.0000	0.0000	0.0000
	5	0.9999	0.9967	0.9765	0.9183	0.8103	0.6598	0.4900	0.3288	0.1976	0.1051	0.0486	0.0191	0.0062	0.0016	0.0003	0.0000	0.0000	0.0000	0.0000
	6	1.0000	0.9995	0.9944	0.9733	0.9204	0.8247	0.6881	0.5272	0.3660	0.2272	0.1241	0.0583	0.0229	0.0071	0.0016	0.0002	0.0000	0.0000	0.0000
	7	1.0000	0.9999	0.9989	0.9930	0.9729	0.9256	0.8406	0.7161	0.5629	0.4018	0.2559	0.1423	0.0671	0.0257	0.0075	0.0015	0.0002	0.0000	0.0000
	8	1.0000	1.0000	0.9998	0.9985	0.9925	0.9743	0.9329	0.8577	0.7441	0.5982	0.4371	0.2839	0.1594	0.0744	0.0271	0.0070	0.0011	0.0001	0.0000
	9	1.0000	1.0000	1.0000	0.9998	0.9984	0.9929	0.9771	0.9417	0.8759	0.7728	0.6340	0.4728	0.3119	0.1753	0.0796	0.0267	0.0056	0.0005	0.0000
	10	1.0000	1.0000	1.0000	1.0000	0.9997	0.9984	0.9938	0.9809	0.9514	0.8949	0.8024	0.6712	0.5100	0.3402	0.1897	0.0817	0.0235	0.0033	0.0001
	11	1.0000	1.0000	1.0000	1.0000	1.0000	0.9997	0.9987	0.9951	0.9851	0.9616	0.9147	0.8334	0.7108	0.5501	0.3698	0.2018	0.0791	0.0170	0.0009
	12	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000	0.9998	0.9991	0.9965	0.9894	0.9719	0.9349	0.8661	0.7541	0.5950	0.4019	0.2101	0.0684	0.0070
	13	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000	0.9999	0.9994	0.9979	0.9934	0.9817	0.9549	0.9006	0.8029	0.6482	0.4386	0.2108	0.0429
	14	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000	0.9999	0.9997	0.9990	0.9967	0.9902	0.9739	0.9365	0.8593	0.7161	0.4853	0.1892
	15	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000	0.9999	0.9997	0.9990	0.9967	0.9900	0.9719	0.9257	0.8147	0.5599
17	0	0.4181	0.1668	0.0631	0.0225	0.0075	0.0023	0.0007	0.0002	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000
	1	0.7922	0.4818	0.2525	0.1182	0.0501	0.0193	0.0067	0.0021	0.0006	0.0001	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000
	2	0.9497	0.7618	0.5198	0.3096	0.1637	0.0774	0.0327	0.0123	0.0041	0.0012	0.0003	0.0001	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000
	3	0.9912	0.9174	0.7556	0.5489	0.3530	0.2019	0.1028	0.0464	0.0184	0.0064	0.0019	0.0005	0.0001	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000
	4	0.9988	0.9779	0.9013	0.7582	0.5739	0.3887	0.2348	0.1260	0.0596	0.0245	0.0086	0.0025	0.0006	0.0001	0.0000	0.0000	0.0000	0.0000	0.0000
	5	0.9999	0.9953	0.9681	0.8943	0.7653	0.5968	0.4197	0.2639	0.1471	0.0717	0.0301	0.0106	0.0030	0.0007	0.0001	0.0000	0.0000	0.0000	0.0000
	6	1.0000	0.9992	0.9917	0.9623	0.8929	0.7752	0.6188	0.4478	0.2902	0.1662	0.0826	0.0348	0.0120	0.0032	0.0006	0.0001	0.0000	0.0000	0.0000
	7	1.0000	0.9999	0.9983	0.9891	0.9598	0.8954	0.7872	0.6405	0.4743	0.3145	0.1834	0.0919	0.0383	0.0127	0.0031	0.0005	0.0000	0.0000	0.0000
	8	1.0000	1.0000	0.9997	0.9974	0.9876	0.9597	0.9006	0.8011	0.6626	0.5000	0.3374	0.1989	0.0994	0.0403	0.0124	0.0026	0.0003	0.0000	0.0000
	9	1.0000	1.0000	1.0000	0.9995	0.9969	0.9873	0.9617	0.9081	0.8166	0.6855	0.5257	0.3595	0.2128	0.1046	0.0402	0.0109	0.0017	0.0001	0.0000
	10	1.0000	1.0000	1.0000	0.9999	0.9994	0.9968	0.9880	0.9652	0.9174	0.8338	0.7098	0.5522	0.3812	0.2248	0.1071	0.0377	0.0083	0.0008	0.0000
	11	1.0000	1.0000	1.0000	1.0000	0.9999	0.9993	0.9970	0.9894	0.9699	0.9283	0.8529	0.7361	0.5803	0.4032	0.2347	0.1057	0.0319	0.0047	0.0001
	12	1.0000	1.0000	1.0000	1.0000	1.0000	0.9999	0.9994	0.9975	0.9914	0.9755	0.9404	0.8740	0.7652	0.6113	0.4261	0.2418	0.0987	0.0221	0.0012
	13	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000	0.9999	0.9995	0.9981	0.9936	0.9816	0.9536	0.8972	0.7981	0.6470	0.4511	0.2444	0.0826	0.0088
	14	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000	0.9999	0.9997	0.9988	0.9959	0.9877	0.9673	0.9226	0.8363	0.6904	0.4802	0.2382	0.0503
	15	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000	0.9999	0.9994	0.9979	0.9933	0.9807	0.9499	0.8818	0.7475	0.5182	0.2078
	16	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000	0.9998	0.9993	0.9977	0.9925	0.9775	0.9369	0.8332	0.5819
18	0	0.3972	0.1501	0.0536	0.0180	0.0056	0.0016	0.0004	0.0001	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000
	1	0.7735	0.4503	0.2241	0.0991	0.0395	0.0142	0.0046	0.0013	0.0003	0.0001	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000
	2	0.9419	0.7338	0.4797	0.2713	0.1353	0.0600	0.0236	0.0082	0.0025	0.0007	0.0001	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000
	3	0.9891	0.9018	0.7202	0.5010	0.3057	0.1646	0.0783	0.0328	0.0120	0.0038	0.0010	0.0002	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000
	4	0.9985	0.9718	0.8794	0.7164	0.5187	0.3327	0.1886	0.0942	0.0411	0.0154	0.0049	0.0013	0.0003	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000
	5	0.9998	0.9936	0.9581	0.8671	0.7175	0.5344	0.3550	0.2088	0.1077	0.0481	0.0183	0.0058	0.0014	0.0003	0.0000	0.0000	0.0000	0.0000	0.0000
	6	1.0000	0.9988	0.9882	0.9487	0.8610	0.7217	0.5491	0.3743	0.2258	0.1189	0.0537	0.0203	0.0062	0.0014	0.0002	0.0000	0.0000	0.0000	0.0000
	7	1.0000	0.9998	0.9973	0.9837	0.9431	0.8593	0.7283	0.5634	0.3915	0.2403	0.1280	0.0576	0.0212	0.0061	0.0012	0.0002	0.0000	0.0000	0.0000
	8	1.0000	1.0000	0.9995	0.9957	0.9807	0.9404	0.8609	0.7368	0.5778	0.4073	0.2527	0.1347	0.0597	0.0210	0.0054	0.0009	0.0001	0.0000	0.0000
	9	1.0000	1.0000	0.9999	0.9991	0.9946	0.9790	0.9403	0.8653	0.7473	0.5927	0.4222	0.2632	0.1391	0.0596	0.0193	0.0043	0.0005	0.0000	0.0000
	10	1.0000	1.0000	1.0000	0.9998	0.9988	0.9939	0.9788	0.9424	0.8720	0.7597	0.6085	0.4366	0.2717	0.1407	0.0569	0.0163	0.0027	0.0002	0.0000

n	x	0.05	0.10	0.15	0.20	0.25	0.30	0.35	0.40	0.45	P 0.50	0.55	0.60	0.65	0.70	0.75	0.80	0.85	0.90	0.95
	11	1.0000	1.0000	1.0000	1.0000	0.9998	0.9986	0.9938	0.9797	0.9463	0.8811	0.7742	0.6257	0.4509	0.2783	0.1390	0.0513	0.0118	0.0012	0.0000
	12	1.0000	1.0000	1.0000	1.0000	1.0000	0.9997	0.9986	0.9942	0.9817	0.9519	0.8923	0.7912	0.6450	0.4656	0.2825	0.1329	0.0419	0.0064	0.0002
	13	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000	0.9997	0.9987	0.9951	0.9846	0.9589	0.9058	0.8114	0.6673	0.4813	0.2836	0.1206	0.0282	0.0015
	14	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000	0.9998	0.9990	0.9962	0.9880	0.9672	0.9217	0.8354	0.6943	0.4990	0.2798	0.0982	0.0109
	15	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000	0.9999	0.9993	0.9975	0.9918	0.9764	0.9400	0.8647	0.7287	0.5203	0.2662	0.0581
	16	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000	0.9999	0.9997	0.9987	0.9954	0.9858	0.9605	0.9009	0.7759	0.5497	0.2265
	17	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000	0.9999	0.9996	0.9984	0.9944	0.9820	0.9464	0.8499	0.6028
19	0	0.3774	0.1351	0.0456	0.0144	0.0042	0.0011	0.0003	0.0001	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000
	1	0.7547	0.4203	0.1985	0.0829	0.0310	0.0104	0.0031	0.0008	0.0002	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000
	2	0.9335	0.7054	0.4413	0.2369	0.1113	0.0462	0.0170	0.0055	0.0015	0.0004	0.0001	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000
	3	0.9868	0.8850	0.6841	0.4551	0.2631	0.1332	0.0591	0.0230	0.0077	0.0022	0.0005	0.0001	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000
	4	0.9980	0.9648	0.8556	0.6733	0.4654	0.2822	0.1500	0.0696	0.0280	0.0096	0.0028	0.0006	0.0001	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000
	5	0.9998	0.9914	0.9463	0.8369	0.6678	0.4739	0.2968	0.1629	0.0777	0.0318	0.0109	0.0031	0.0007	0.0001	0.0000	0.0000	0.0000	0.0000	0.0000
	6	1.0000	0.9983	0.9837	0.9324	0.8251	0.6655	0.4812	0.3081	0.1727	0.0835	0.0342	0.0116	0.0031	0.0006	0.0001	0.0000	0.0000	0.0000	0.0000
	7	1.0000	0.9997	0.9959	0.9767	0.9225	0.8180	0.6656	0.4878	0.3169	0.1796	0.0871	0.0352	0.0114	0.0028	0.0005	0.0000	0.0000	0.0000	0.0000
	8	1.0000	1.0000	0.9992	0.9933	0.9713	0.9161	0.8145	0.6675	0.4940	0.3238	0.1841	0.0885	0.0347	0.0105	0.0023	0.0003	0.0000	0.0000	0.0000
	9	1.0000	1.0000	0.9999	0.9984	0.9911	0.9674	0.9125	0.8139	0.6710	0.5000	0.3290	0.1861	0.0875	0.0326	0.0089	0.0016	0.0001	0.0000	0.0000
	10	1.0000	1.0000	1.0000	0.9997	0.9977	0.9895	0.9653	0.9115	0.8159	0.6762	0.5060	0.3325	0.1855	0.0839	0.0287	0.0067	0.0008	0.0000	0.0000
	11	1.0000	1.0000	1.0000	1.0000	0.9995	0.9972	0.9886	0.9648	0.9129	0.8204	0.6831	0.5122	0.3344	0.1820	0.0775	0.0233	0.0041	0.0003	0.0000
	12	1.0000	1.0000	1.0000	1.0000	0.9999	0.9994	0.9969	0.9884	0.9658	0.9165	0.8273	0.6919	0.5188	0.3345	0.1749	0.0676	0.0163	0.0017	0.0000
	13	1.0000	1.0000	1.0000	1.0000	1.0000	0.9999	0.9993	0.9969	0.9891	0.9682	0.9223	0.8371	0.7032	0.5261	0.3322	0.1631	0.0537	0.0086	0.0002
	14	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000	0.9999	0.9994	0.9972	0.9904	0.9720	0.9304	0.8500	0.7178	0.5346	0.3267	0.1444	0.0352	0.0020
	15	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000	0.9999	0.9995	0.9978	0.9923	0.9770	0.9409	0.8668	0.7369	0.5449	0.3159	0.1150	0.0132
	16	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000	0.9999	0.9996	0.9985	0.9945	0.9830	0.9538	0.8887	0.7631	0.5587	0.2946	0.0665
	17	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000	0.9998	0.9992	0.9969	0.9896	0.9690	0.9171	0.8015	0.5797	0.2453
	18	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000	0.9999	0.9997	0.9989	0.9958	0.9856	0.9544	0.8649	0.6226
20	0	0.3585	0.1216	0.0388	0.0115	0.0032	0.0008	0.0002	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000
	1	0.7358	0.3917	0.1756	0.0692	0.0243	0.0076	0.0021	0.0005	0.0001	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000
	2	0.9245	0.6769	0.4049	0.2061	0.0913	0.0355	0.0121	0.0036	0.0009	0.0002	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000
	3	0.9841	0.8670	0.6477	0.4114	0.2252	0.1071	0.0444	0.0160	0.0049	0.0013	0.0003	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000
	4	0.9974	0.9568	0.8298	0.6296	0.4148	0.2375	0.1182	0.0510	0.0189	0.0059	0.0015	0.0003	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000
	5	0.9997	0.9887	0.9327	0.8042	0.6172	0.4164	0.2454	0.1256	0.0553	0.0207	0.0064	0.0016	0.0003	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000
	6	1.0000	0.9976	0.9781	0.9133	0.7858	0.6080	0.4166	0.2500	0.1299	0.0577	0.0214	0.0065	0.0015	0.0003	0.0000	0.0000	0.0000	0.0000	0.0000
	7	1.0000	0.9996	0.9941	0.9679	0.8982	0.7723	0.6010	0.4159	0.2520	0.1316	0.0580	0.0210	0.0060	0.0013	0.0002	0.0000	0.0000	0.0000	0.0000
	8	1.0000	0.9999	0.9987	0.9900	0.9591	0.8867	0.7624	0.5956	0.4143	0.2517	0.1308	0.0565	0.0196	0.0051	0.0009	0.0001	0.0000	0.0000	0.0000
	9	1.0000	1.0000	0.9998	0.9974	0.9861	0.9520	0.8782	0.7553	0.5914	0.4119	0.2493	0.1275	0.0532	0.0171	0.0039	0.0006	0.0000	0.0000	0.0000
	10	1.0000	1.0000	1.0000	0.9994	0.9961	0.9829	0.9468	0.8725	0.7507	0.5881	0.4086	0.2447	0.1218	0.0480	0.0139	0.0026	0.0002	0.0000	0.0000
	11	1.0000	1.0000	1.0000	0.9999	0.9991	0.9949	0.9804	0.9435	0.8692	0.7483	0.5857	0.4044	0.2376	0.1133	0.0409	0.0100	0.0013	0.0001	0.0000
	12	1.0000	1.0000	1.0000	1.0000	0.9998	0.9987	0.9940	0.9790	0.9420	0.8684	0.7480	0.5841	0.3990	0.2277	0.1018	0.0321	0.0059	0.0004	0.0000
	13	1.0000	1.0000	1.0000	1.0000	1.0000	0.9997	0.9985	0.9935	0.9786	0.9423	0.8701	0.7500	0.5834	0.3920	0.2142	0.0867	0.0219	0.0024	0.0000
	14	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000	0.9997	0.9984	0.9936	0.9793	0.9447	0.8744	0.7546	0.5836	0.3828	0.1958	0.0673	0.0113	0.0003
	15	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000	0.9997	0.9985	0.9941	0.9811	0.9490	0.8818	0.7625	0.5852	0.3704	0.1702	0.0432	0.0026
	16	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000	0.9997	0.9987	0.9951	0.9840	0.9556	0.8929	0.7748	0.5886	0.3523	0.1330	0.0159
	17	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000	0.9998	0.9991	0.9964	0.9879	0.9645	0.9087	0.7939	0.5951	0.3231	0.0755
	18	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000	0.9999	0.9995	0.9979	0.9924	0.9757	0.9308	0.8244	0.6083	0.2642
	19	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000	0.9998	0.9992	0.9968	0.9885	0.9612	0.8784	0.6415

Reg. No. :

Question Paper Code : 51329

B.E./B.Tech. DEGREE EXAMINATIONS, APRIL/MAY 2024.

Third / Fourth Semester

Environmental Engineering

MA 3391 — PROBABILITY AND STATISTICS

(Common to : Artificial Intelligence and Data Science / Biotechnology and Biochemical Engineering / Computer Science and Business Systems / Plastic Technology)

(Regulations 2021)

Time : Three hours　　　　Maximum : 100 marks

(use of statistical table is pemitted)

Answer ALL questions.

PART A — (10 × 2 = 20 marks)

1. During off-peak hours a commuter train has five cars. Suppose a commuter is twice as likely to select the middle car (#3) as to select either adjacent car (#2 or #4), and is twice as likely to select either adjacent car as to select either end car (#1 or #5). Find the probability that one of the three middle cars is selected.

2. The probability that a certain kind of component will survive a shock test is 3/4. Find the probability that exactly 2 of the next 4 components tested survive.

3. A privately owned business operates both a drive-in facility and a walk-in facility. On a randomly selected day, let X and Y, respectively, be the proportions of the time that the drive-in and the walk-in facilities are in use, and suppose that the joint density function of these random variables is $f(x,y) = k(2x+3y)$; $0 \le x \le 1$, $0 \le y \le 1$. Find the value of k.

4. An electrical firm manufactures light bulbs that have a length of life that is approximately normally distributed, with mean equal to 800 hours and a standard deviation of 40 hours. What is the probability that a random sample of 16 bulbs will have an average life of less than 775 hours?

5. Data pertaining to height of 5 school students are given as 149, 150,151, 138,148 cms. Obtain a point estimate for the mean μ.

6. A random sample of size $n = 100$ is taken from a population with $\sigma = 5.1$. Given that the sample mean is $\bar{x} = 21.6$, construct a 95% confidence interval for the population mean μ.

7. Write the benefit of applying non-parametric tests compared to parametric tests.

8. What is the Kolmogorov-Smirnov Test? Is it non parametric?

9. Which control chart is not influenced by the sample size? What probability distribution does it follow?

10. What is the advantage of using control charts for attributes compared to variable control charts?

PART B — (5 × 16 = 80 marks)

11. (a) (i) Customers are used to evaluate preliminary product designs. In the past, 95% of highly successful products received good reviews, 60% of moderately successful products received good reviews, and 10% of poor products received good reviews. In addition, 40% of products have been highly successful, 35% have been moderately successful, and 25% have been poor products.

(1) What is the probability that a product attains a good review?

(2) If a new design attains a good review, what is the probability that it will be a highly successful product?

(3) If a product does not attain a good review, what is the probability that it will be a highly successful product? (8)

(ii) Assume that each of your calls to a popular radio station has a probability of 0.02 of connecting, that is, of not obtaining a busy signal. Assume that your calls are independent.

(1) What is the probability that your first call that connects is your tenth call?

(2) What is the probability that it requires more than five calls for you to connect?

(3) What is the mean number of calls needed to connect? (8)

Or

(b) (i) The thickness of a flange on an aircraft component is uniformly distributed between 0.95 and 1.05 millimeters.

(1) Determine the cumulative distribution function of flange thickness.

(2) Determine the proportion of flanges that exceeds 1.02 millimeters.

(3) What thickness is exceeded by 90% of the flanges?

(4) Determine the mean and variance of flange thickness. (8)

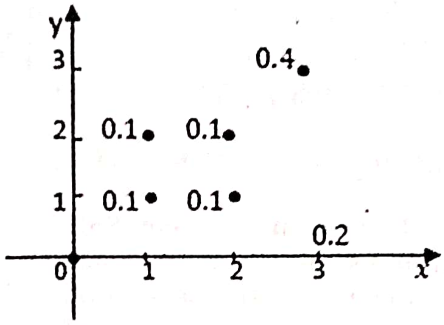
(ii) The life of a semiconductor laser at a constant power is normally distributed with a mean of 7000 hours and a standard deviation of 600 hours.

(1) What is the probability that a laser fails before 5000 hours?

(2) What is the life in hours that 95% of the lasers exceed?

(3) If three lasers are used in a product and they are assumed to fail independently, what is the probability that all three are still operating after 7000 hours? (8)

12. (a) For the discrete random variables X and Y with the joint distribution shown in the following figure: Determine the following :



(i) $P(X < 2, Y < 3)$

(ii) $P(1 < X < 2.5)$

(iii) $P(0 < Y < 2.5)$

(iv) $E(X)$, $E(Y)$, $V(X)$ and $V(Y)$

(v) Marginal probability distribution of the random variable X and Y.

(vi) Conditional probability distribution of Y given that $X = 1$.

(vii) Covariance and Correlation

(viii) Are X and Y independent? (16)

Or

(b) The joint pdf of X amount of almonds and Y amount of cashews were $f(x,y)=\begin{cases} 24xy, & 0 \le x \le 1, 0 \le y \le 1, x+y \le 1 \\ 0 & otherwise \end{cases}$. Find the covariance and correlation between the random variables X and Y. (16)

13. (a) (i) Let X be exponentially distributed with parameter λ. Using maximum likelihood estimation, find an estimate for the parameter λ. (8)

(ii) If 83 male students are randomly chosen and yield an average of 6.6 hours of sleep with a standard deviation of 1.8 and 65 females are randomly selected with an average of 6.9 hours of sleep with a standard deviation of 1.5. Construct a 95% confidence interval for the difference between the two mean sleep hours for males and females. (8)

Or

(b) (i) For random sampling from normal population $N(\mu, \sigma^2)$, find the maximum likelihood estimators for μ and σ^2. (8)

(ii) Find the estimator for λ by the method of moments for the exponential distribution whose probability density function is given by $f(x, \lambda) = \frac{1}{\lambda} e^{-x/\lambda}$, $x > 0, \lambda > 0$. (8)

14. (a) (i) A manufacturer of electric irons, wishing to test the accuracy of the thermostat control at the 500°F setting, instructs a test engineer to obtain actual temperatures at that setting for 15 irons using a thermocouple. The resulting measurements are as follows :

494.6	510.8	487.5	493.2	502.6	485.0	495.9	498.2
501.6	497.3	492.0	504.3	499.2	493.5	505.8	

The engineer believes it is reasonable to assume that a temperature deviation from 500° of any particular magnitude is just as likely to be positive as negative (the assumption of symmetry) but wants to protect against possible non normality of the actual temperature distribution. Use signed-rank test to see whether the data strongly suggests incorrect calibration of the iron. (8)

(ii) The urinary fluoride concentration (parts per million) was measured both for a sample of livestock grazing in an area previously exposed to fluoride pollution and for a similar sample grazing in an unpolluted region :

Pollutted	21.3	18.7	23.0	17.1	16.8	20.9	19.7
Unpollutted	14.2	18.3	17.2	18.4	20.0		

Does the data indicate strongly that the true average fluoride concentration for livestock grazing in the polluted region is larger than for the unpolluted region? Use the Wilcoxon rank-sum test at level $\alpha = 0.01$. (8)

Or

(b) (i) The effectiveness of advertising for two rival products (Brand X and Brand Y) was compared. Market research at a local shopping centre was carried out, with the participants being shown adverts for two rival brands of coffee, which they then rated on the overall likelihood of them buying the product (out of 10, with 10 being "definitely going to buy the product"). Half of the participants gave ratings for one of the products, the other half gave ratings for the other product. Is there is a highly significant difference between the ratings given to each brand in terms of the likelihood of buying the product. Use U-test (take $\alpha = 0.05$) (8)

Brand X		Brand Y	
Participant	Rating	Participant	Rating
1	3	1	9
2	4	2	7
3	2	3	5
4	6	4	10
5	2	5	6
6	5	6	8

(ii) Four group of students were randomly assigned to taught with four different techniques and their achievement test scores were recorded. Are the distributions of test scores the same or do they differ in location? (take $\alpha = 0.05$). (8)

1	2	3	4
65	75	59	94
87	69	78	89
73	83	67	80
79	81	62	88

15. (a) A hard-bake process is used in conjunction with photolithography in semiconductor manufacturing. Fifteen samples, each of size five wafers, have been taken when we think the process is in control. The interval of time between samples or subgroups is one hour. The flow width measurement data (in x microns) from these samples are shown in Table. Establish statistical control of the flow width of the resist in this process using $\bar{x}$ and R charts. (16)

	Wafers				
Sample No.	1	2	3	4	5
1	1.3235	1.4128	1.6744	1.4573	1.6914
2	1.4314	1.3592	1.6075	1.4666	1.6109
3	1.4284	1.4871	1.4932	1.4324	1.5674
4	1.5028	1.6352	1.3841	1.2831	1.5507
5	1.5604	1.2735	1.5265	1.4363	1.6441
6	1.5955	1.5451	1.3574	1.3281	1.4198
7	1.6274	1.5064	1.8366	1.4177	1.5144
8	1.4190	1.4303	1.6637	1.6067	1.5519
9	1.3884	1.7277	1.5355	1.5176	1.3688
10	1.4039	1.6697	1.5089	1.4627	1.5220
11	1.4158	1.7667	1.4278	1.5928	1.4181

Wafers

Sample No.	1	2	3	4	5
12	1.5821	1.3355	1.5777	1.3908	1.7559
13	1.2856	1.4106	1.4447	1.6398	1.1928
14	1.4951	1.4036	1.5893	1.6458	1.4969
15	1.3589	1.2863	1.5996	1.2497	1.5471

Or

(b) (i) Frozen orange juice concentrate is packed in 6-oz cardboard cans. These cans are formed on a machine by spinning them from cardboard stock and attaching a metal bottom panel. By inspection of a can, we may determine whether, when filled, it could possibly Leak either on the side seam or around the bottom joint. Such a nonconforming can has an improper seal on either the side seam or the bottom panel. To establish the control chart, 15 samples of $n = 50$ cans each were selected at half-hour intervals over a three-shift period in which the machine was in continuous operation.

The data are shown in Table. Set up a control chart for the fraction of nonconforming cans produced by this machine. (8)

Sample Number	Number of Nonconforming Cans, D_i
1	12
2	15
3	8
4	10
5	4
6	7
7	16
8	9
9	14
10	10
11	5
12	6
13	17
14	12
15	22

(ii) Table presents the number of nonconformities observed in 26 successive samples of 100 printed circuit boards. Note that, for reasons of convenience, the inspection unit is defined as 100 boards. Set up a c chart for these data. (8)

Sample Number	Number of Nonconformities	Sample Number	Number of Nonconformities
1	21	14	19
2	24	15	10
3	16	16	17
4	12	17	13
5	15	18	22
6	5	19	18
7	28	20	39
8	20	21	30
9	31	22	24
10	25	23	16
11	20	24	19
12	24	25	17
13	16	26	15

Reg. No. : | | | | | | | | | | | | |

Question Paper Code : 41366

B.E./B.Tech. DEGREE EXAMINATIONS, NOVEMBER/DECEMBER 2024.

Third/Fourth Semester

Environmental Engineering

MA 3391 — PROBABILITY AND STATISTICS

(Common to: Artificial Intelligence and Data Science/Biotechnology and Biochemical Engineering/Computer Science and Business Systems/Plastic Technology)

(Regulations 2021)

Time : Three hours Maximum : 100 marks

(Use of statistics table is permitted)

Answer ALL questions.

PART A — ($10 \times 2 = 20$ marks)

1. What is the probability of getting an even number in a single throw with a dice?

2. Suppose a book of 585 pages contains 43 typographical errors. If these errors are randomly distributed throughout the book, Using Poisson distribution, find the probability that 10 pages, selected as random, will be free from errors?

3. Find the marginal distribution of X and Y from the following probability distribution function of (X, Y).

		Y	
		1	2
X	1	0.1	0.2
	2	0.3	0.4

4. If $U = X + Y$ and $V = X - Y$, how are the joint probability density functions of (X, Y) and (U, V) related?

5. If t_1 and t_2 are both most efficient estimators with equal variance V and if t_3 is the average of t_1 *and* t_2. Prove that $\operatorname{var}(t_3)=\frac{1}{2}V(1+\rho)$, where ρ is the coefficient of correlation between t_1 and t_2.

6. Define the maximum likelihood estimator.

7. What is the use of median test?

8. What are the disadvantages of non-parametric methods of testing hypothesis?

9. When do you say that a process is out of control?

10. What is control chart? Name the types of control charts.

PART B — (5 × 16 = 80 marks)

11. (a) (i) Find the mean and variance of Binomial distribution. (8)

(ii) The probabilities of X, Y and Z becoming managers are 4/9, 2/9 and 1/3 respectively. The probabilities that the Bonus scheme will be introduced if X, Y, Z becomes managers are 3/10, 1/2 and 4/5 respectively.

(1) What is the probability that the Bonus scheme will be introduced? (4)

(2) If the Bonus scheme has been introduced, what is the probability that the manager appointed was X? (4)

Or

(b) (i) The random variable X has the following probability distribution:

$X=x$:	-2	-1	0	1	2	3
$P(x)$:	0.1	k	0.2	$2k$	0.3	$3k$

Find

(1) The value of 'k' (2)

(2) $P(X<2)$ (2)

(3) $P(-2<X<2)$ (2)

(4) The mean of X. (2)

(ii) The time (in hours) required to repair a machine is exponentially distributed with parameter $\lambda = 1/2$.

(1) Find the probability that the repair time exceeds 2 hours? (4)

(2) What is the conditional probability that a repair takes at least 10 hours, given that its duration exceeds 9 hours? (4)

12. (a) (i) If the joint density function of (X, Y) is given by $f(x, y) = \frac{1}{3}(x + y)$, $0 \le x \le 1, 0 \le y \le 2$, find

(1) The marginal density function of X and Y (4)

(2) The conditional density function of X given Y and Y given X. (4)

(ii) A study of prices of rice at Chennai and Madurai gave the following data. Also the coefficient of correlation between the two is 0.8.

	Chennai	Madurai
Mean	19.5	17.75
Standard Deviation	1.75	2.5

Estimate the most likely price of rice

(1) At Chennai corresponding to the price of 18

(2) At Madurai corresponding to the price of 17 at Chennai. (8)

Or

(b) (i) The probability density function of (X, Y) is given by

$f_{XY}(x, y) = \begin{cases} e^{-(x+y)}; & x > 0, y > 0 \\ 0; & \text{otherwise} \end{cases}$. Find the joint probability density function of (U, V), where $U = \frac{X}{Y}$ and $V = X + Y$. Also find marginal density function of U and V. (8)

(ii) Ten students got the following percentage of marks in Mathematical and Physical sciences. Calculate the rank correlation coefficient. (8)

Maths	78	36	98	25	75	82	90	62	65	39
Physics	84	51	91	60	68	62	86	58	63	47

13. (a) (i) Let $x_1, x_2, \ldots, x_n$ are random observations of a Bernoulli's variate x which assumes values 1 and 0 with probabilities θ and $1-\theta$ respectively. Show that $\frac{T(n-T)}{n(n-1)}$ is an unbiased estimator of $\theta(1-\theta)$, where $T = x_1 + x_2 + x_3 + \ldots + x_n$. (8)

(ii) Find the maximum likelihood estimators of 'a' and 'b' in terms of the sample observations $x_1, x_2, \ldots, x_n$ taken from the exponential population with density function $f(x, a, b) = ke^{-b(x-b)}; x \geq a, b > 0$. (8)

Or

(b) (i) If $\{x_1, x_2, \ldots, x_n\}$ is a random sample of size n, drawn from a geometric distribution, then the probability mass function of which is given by $P(x = r) = pq^{r-1}; r = 1, 2, 3, \ldots, \infty$. Prove that the mean of the sample is a consistent estimator of the population mean. (8)

(ii) Let $(x_1, x_2, \ldots, x_n)$ is a random sample from a population density function $f(x; \theta, \mu) = \theta e^{-\theta(x-\mu)}; x > \mu$. Find the method of moments estimators of θ and μ. (8)

14. (a) (i) Use the sign test to see if there is a difference between the number of days required to collect an account receivable before and after a new collection policy. Use the 0.05 significance level.

Before : 33 36 41 32 39 47 34 29 32 34 40 42 33 36 27

After : 35 29 38 34 37 47 36 32 30 34 41 38 37 35 28 (8)

(ii) Twenty five individuals were sampled as to whether they or did not like a product indicated by Y and N respectively, the resulting sample is shown by the following sequence:

YY NNNN YYY N Y NN Y NNNN YYY NN

(1) Determine the number of runs, V.

(2) Test at 0.05 significance level whether the responses are random. (8)

Or

(b) (i) The following are the final examination of marks of three groups of students who were taught computer by three different methods.

First method :	94	88	91	74	87	97	
Second method :	85	82	79	84	61	72	80
Third method :	89	67	72	76	69		

Use the H-test at the 0.05 level of significance to test the null hypothesis that the three methods are equally effective. (8)

(ii) The following are the speeds (in kilometer per hour) at which every fifth passenger car was timed at a certain checkpoint: 46, 58, 60, 56, 70, 66, 48, 48, 54, 62, 41, 39, 52, 45, 62, 53, 69, 65, 65, 67, 76, 52, 52, 59, 59, 67, 51, 46, 61, 40, 43, 42, 77, 67, 63, 59, 63, 63, 72, 57, 59, 42, 56, 47, 62, 67, 70, 63, 66, 69 and 73.

Test the null hypothesis of randomness at the 0.05 level of significance. (Given median speed = 59.5 km per hour). (8)

15. (a) (i) 10 samples each of size 50 were inspected and the number of defectives in the inspection were: 2, 1, 1, 2, 3, 5, 5, 1, 2, 3. Draw the np-chart for defectives. (8)

(ii) Given below are the values of sample mean $\overline{X}$ and sample range R for 10 samples, each of size 5. Draw the appropriate mean chart and comment on the state of control of the process. (8)

Sample no :	1	2	3	4	5	6	7	8	9	10
Mean :	43	49	37	44	45	37	51	46	43	47
Range :	5	6	5	7	7	4	8	6	4	6

Or

(b) (i) 15 tape-recorders were examined for quality control test. The number of defects in each tape-recorder is recorded below. Draw the appropriate control chart and comment on the state of control. (8)

No of units :	1	2	3	4	5	6	7	8	9
No of defects :	2	4	3	1	1	2	5	3	6
No of units :	10	11	12	13	14	15			
No of defects :	7	3	1	4	2	1			

(ii) 15 samples of 200 items each were drawn from the output of a process. The number of defective items in the samples are given below. Prepare a control chart for the fraction defective and comment on the state of control. (8)

Sample no :	1	2	3	4	5	6	7	8
No of defective :	12	15	10	8	19	15	17	11
Sample no :	9	10	11	12	13	14	15	
No of defective :	13	20	10	8	9	5	8	