RAJIV GANDHI PROUDYOGIKI VISHWAVIDYALAYA, BHOPAL

New Scheme Based On AICTE Flexible Curricula

CSE-Artificial Intelligence and Machine Learning/ Artificial Intelligence and Machine Learning, III-Semester

AL302 Introduction to Probability and Statistics

Objective: The objective of this course is to familiarize the students with statistical techniques. It aims to equip the students with standard concepts and tools at an intermediate to advanced level that will serve them well towards tackling various problems in the discipline.

Unit 1: Basic Probability-

Probability spaces, conditional probability, independence; Discrete random variables, Independent random variables, the multinomial distribution, Poisson approximation to the binomial distribution, infinite sequences of Bernoulli trials, sums of independent random variables; Expectation of Discrete Random Variables, Moments, Variance of a sum, Correlation coefficient, Chebyshev's Inequality.

Unit 2: Continuous Probability Distributions-

Continuous random varibales and their properties, distribution functions and densities, normal, exponential and gamma densities.

Unit 3: Bivariate Distributions-

Bivariate distributions and their properties, distribution of sums and quotients, conditional densities, Bayes' rule.

Unit 4: Basic Statistics-

Measures of Central tendency: Moments, skewness and Kurtosis - Probability distributions: Binomial, Poisson and Normal - evaluation of statistical parameters for these three distributions, Correlation and regression – Rank correlation.

Unit 5: Applied Statistics-

Curve fitting by the method of least squares- fitting of straight lines, second degree parabolas and more general curves. Test of significance: Large sample test for single proportion, difference of proportions, single mean, difference of means, and difference of standard deviations.

Unit 6: Small samples-

Test for single mean, difference of means and correlation coefficients, test for ratio of variances - Chi-square test for goodness of fit and independence of attributes.

Suggested Text/Reference Books :

1. Erwin Kreyszig, Advanced Engineering Mathematics, 9th Edition, John Wiley & Sons, 2006.
2. P. G. Hoel, S. C. Port and C. J. Stone, Introduction to Probability Theory, Universal Book Stall, 2003 (Reprint).
3. S. Ross, A First Course in Probability, 6th Ed., Pearson Education India, 2002.

4. W. Feller, An Introduction to Probability Theory and its Applications, Vol. 1, 3rd Ed., Wiley, 1968.
5. N.P. Bali and Manish Goyal, A text book of Engineering Mathematics, Laxmi Publications, Reprint, 2010.
6. B.S. Grewal, Higher Engineering Mathematics, Khanna Publishers, 35th Edition, 2000.
7. Veerarajan T., Engineering Mathematics (for semester III), Tata McGraw-Hill, New Delhi, 2010.

- Random variable : If a real variable 'x' be associated with the outcome of a random experiment then since the value taken depends on chance is called random variable.

- Discrete Variate : If a random variable 'x' take a finite set of value then it is known as discrete variate

- Continuous Variate : If a random variable 'x' c assumes an infinite number of uncountable values, It is called as Continuous variate.

Que:1 If a die is tossed thrice, Success is getting on 1 or 6. Find the mean and variance of the number of Success.

Sol: Probability of Success $= \frac{2}{6} = \frac{1}{3}$

Probability of failure $= \frac{4}{6} = \frac{2}{3}$

Probability of No Success $= 3c_0 \times \left(\frac{2}{3}\right)^3$

$= \frac{3}{3} \times \frac{8}{27}$

$= \underline{\underline{\frac{8}{27}}}$

Probability of One success and 2 failure

$$= {}^nC_r\, p^{n-r} q^r$$

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

$$= 3 \times \frac{1}{3} \times \frac{4}{9} = \frac{4}{9}$$

Probability of two success and 1 failure

$$= 3C_2 \times \left(\frac{1}{3}\right)^2 \times \left(\frac{2}{3}\right)^1$$

$$= \frac{3 \times 2}{2 \times 1} \times \frac{1}{9} \times \frac{2}{3}$$

$$= \frac{2}{9}$$

Probability of 3 Success =

$$3C_3 \times \left(\frac{1}{3}\right)^3 = \frac{1}{27}$$

x_i	:	0	1	2	3
P_i	:	$\frac{8}{27}$	$\frac{4}{9}$	$\frac{2}{9}$	$\frac{1}{27}$

$$\text{Mean } (m) = \sum p_i x_i$$

$$= \left[\frac{4}{9} + \frac{4}{9} + \frac{1}{9}\right] = \underline{\underline{1}}$$

also, $\sum p_i x_i^2 = \left[\frac{4}{9} + \frac{8}{9} + \frac{9}{27}\right] = \underline{\underline{\frac{5}{3}}}$

$$\text{Variance } (\sigma^2) = \sum p_i x_i^2 - (m)^2$$

$$= \frac{5}{3} - 1$$

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

Q:2 The Probability density function of a variate x is

x:	0	1	2	3	4	5	6
$P(x)$:	k	$3k$	$5k$	$7k$	$9k$	$11k$	$13k$

find (i) $P(x \leq 4)$; $P(x \geq 5)$; $P(3 < x < 6)$

(ii) What will be the minimum value of k if $P(x \leq 2) > 3$

$$\therefore \sum_{i=0}^{6} p_i = 1$$

$$k + 3k + 5k + 7k + 9k + 11k + 13k = 1$$

$$49k = 1$$

Page No. :
Date :

$$\boxed{k = \frac{1}{49}}$$

(i) ⓐ $P(x \leq 4) = P(0) + P(1) + P(2) + P(3) + P(4)$

$= k + 3k + 5k + 7k + 9k$

$= 25k \quad \left[\because k = \frac{1}{49}\right]$

$= \frac{25}{49}$

(b) $P(x \geq 5) = P(5) + P(6)$

$= 11k + 13k$

$= 24k \quad \left[\because k = \frac{1}{49}\right]$

$= \frac{24}{49}$

ⓒ $P(3 < x < 6) = P(4) + P(5)$

$= 9k + 11k$

$= 20k \quad \left[k = \frac{1}{49}\right]$

$= \frac{20}{49}$

(ii) $P(x \leq 2) > 3$

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

$k + 3k + 5k > 3$

$9k > 3$

$k > \frac{3}{9}$

$$\boxed{k > \frac{1}{3}}$$

Page No. :
Date :

Que:3 A random variable 'x' has a following Probability

x :	0	1	2	3	4	5	6	7
$P(x)$:	0	k	$2k$	$2k$	$3k$	k^2	$2k^2$	$7k^2+k$

Find (i) Value of k
(ii) $p(x<6)$; $P(x \geq 6)$
(iii) $P(0<x<6)$

Sol: (i) $\because \sum_{i=0}^{7} P_i = 1$

$k+2k+2k+3k+k^2+2k^2+7k^2+k=1$

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

$k = \frac{1}{10}, -1$

$\therefore \boxed{k = \frac{1}{10}}$

(ii) $P(x<6) = P(0)+P(1)+P(2)+P(3)+P(4)+P(5)$

$= k+2k+2k+3k+k^2$

$= k(k+8)$

$= 0.81$

$P(x \geq 6) = P(6)+P(7)$

$= 2k^2+7k^2+k$

$= 9k^2+k$

$= 0.19$

Date :

(iii) $P(0<x<6) = P(1)+P(2)+P(3)+P(4)+P(5)$
$= 2K+2K+3K+K^2+K$
$= K(K+8)$
$= \underline{0.81}$

Distribution function

if $F(x) = P(x \le x) = \int_{-\infty}^{\infty} f(x)\,dx$.

then $F(x)$ is defined as the Cumulative distribution function.

Properties of the Distribution function.

(i) If $F(x) = f(x) \geq 0$; $F(x)$ is a increasing function.

(ii) $f(-\infty) = 0$
(iii) $f(\infty) = 1$
(iv) $P(a \le x \le b) = \int_a^b f(x)\,dx$.

Que:1 Is the function define as follows:

$$f(x) = \begin{cases} e^{-x} & ; \quad x \geq 0 \\ 0 & ; \quad x < 0 \end{cases}$$

a distribution function.

(ii) Determine the Prob. that the variate having this density in the interval (1, 2)

(iii) Also find the Cummulative Probability function F(2).

Sol: $\because \int_{-\infty}^{\infty} f(x)\,dx = \int_{-\infty}^{0} f(x)\,dx + \int_{0}^{\infty} f(x)\,dx$

$$= 0 + \int_{0}^{\infty} e^{-x}\,dx$$

$$= \left| -e^{-x} \right|_{0}^{\infty}$$

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

$$= 1$$

$\therefore$ It is a distribution function.

(ii) $P(1, 2) = \int_{1}^{2} f(x)\,dx$

$$= \int_{1}^{2} e^{-x} = \left| -e^{-x} \right|_{1}^{2}$$

$$= -e^{-2} + e^{-1}$$

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

$$= \int_{-\infty}^{0} f(x)\,dx + \int_{0}^{2} f(x)\,dx$$

$$= 0 + \left[-e^{-x} \right]_{0}^{2}$$

$$= -e^{-2} + 1$$

Que: 2 'X' is a Continuous random variable with Probability density function given by

$$f(x) = \begin{cases} kx & ; \quad 0 \le x \le 2 \\ 2k & ; \quad 2 \le x \le 4 \\ -kx + 6k & ; \quad 4 \le x \le 6 \end{cases}$$

Find k and mean value of x

Sol: W.K.T $\int_0^6 f(x)\,dx = 1$

$$\int_0^2 f(x)\,dx + \int_2^4 f(x)\,dx + \int_4^6 f(x)\,dx = 1$$

$$\int_0^2 kx + \int_2^4 2k + \int_4^6 -kx + 6k = 1$$

$$\left[\frac{kx^2}{2}\right]_0^2 + \left[2kx\right]_2^4 + \left[\frac{-kx^2}{2} + 6kx\right]_4^6 = 1$$

$$2k + [8k - 4k] + [-18k + 36k + 8k - 24k = 1$$

$$8k = 1$$

$$\boxed{k = \frac{1}{8}}$$

Mean of x $= \int_0^6 x \cdot f(x)\,dx$

Page No. :
Date :

$$= \int_0^2 kx^2\,dx + \int_2^4 2kx + \int_4^6 -kx^2 + 6kx$$

$$= \frac{1}{8}\left[\frac{x^3}{3}\right]_0^2 + \frac{1}{4}\left[\frac{x^2}{2}\right]_2^4 + \frac{1}{8}\left[\frac{-x^3}{3} + 3x^2\right]_4^6$$

$$= \frac{1}{3} + \frac{6}{4} + \frac{1}{8}\left[36 - 48 + \frac{64}{3}\right]$$

$$= \frac{22}{12} + \frac{1}{8}\left[\frac{28}{3}\right]$$

$$= \frac{22}{12} + \frac{14}{12} = \frac{36}{12} = 3$$

Mean of x is 3

Que:3 A Variate x has the probability distribution

x :	-3	6	9
$P(x)$:	$\frac{1}{6}$	$\frac{1}{2}$	$\frac{1}{3}$

Find $E(x)$ and $E(x^2)$ hence Evaluate $E(2x+1)^2$.

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

$$= -\frac{1}{2} + 3 + 3 = \frac{11}{2}$$

$$E(x^2) = \Sigma x^2 . p(x)$$

$$= -\frac{9}{6} + \frac{36}{2} + \frac{81}{3}$$

$$= \frac{93}{2}$$

$$E(2x+1)^2 = E[4x^2 + 1 + 4x]$$

$$= 4E(x^2) + 4E(x) + 1$$

$$= 4 \times \frac{93}{2} + 4 \times \frac{11}{2} + 1$$

$$= 209$$

Que:4 A random variable x has the P.d.f

$$f(x) = \begin{cases} ax & ; \quad 0 \le x \le 1 \\ a & ; \quad 1 \le x \le 2 \\ -ax + 3a & ; \quad 2 \le x \le 3 \\ 0 & ; \quad \text{Otherwise} \end{cases}$$

Find (i) Value of a

(ii) $P(x \le 1.5)$

Sol: W.K.T $\int_0^3 f(x)\,dx = 1$

Page No. :
Date :

$$\int_0^1 ax\,dx + \int_1^2 a\,dx + \int_2^3 -ax + 3a\,dx = 1$$

$$\left|\frac{ax^2}{2}\right|_0^1 + \left|ax\right|_1^2 + \left|-\frac{ax^2}{2} + 3ax\right|_2^3 = 1$$

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

$$\frac{3a - 9a + 18a + 4a - 12a}{2} = 1$$

$$2a = 1$$

$$\boxed{a = \frac{1}{2}}$$

(ii) $P(x \leq 1.5)$

$$\int_0^1 ax\,dx + \int_1^{1.5} a\,dx$$

$$\left|\frac{ax^2}{2}\right|_0^1 + \left|ax\right|_1^{1.5}$$

$$\frac{a}{2} + \frac{a}{2} = a =$$

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

Date :

Que: 5 Find the Value of k for P.d.f

$$f(x) = \begin{cases} kx^2 & : 0 \le x \le 3 \\ 0 & : \text{otherwise} \end{cases}$$

and Compute $P(1 \le x \le 2)$ also find the distribution function.

Sol: W.K.T $\int_0^3 kx^2 \, dx = 1$

$$\left[\frac{kx^3}{3}\right]_0^3 = 1$$

$$9k = 1 \qquad \boxed{k = \frac{1}{9}}$$

$$P(1 \le x \le 2) = \int_1^2 kx^2 = \left[\frac{kx^3}{3}\right]_1^2$$

$$= \frac{1}{9}\left[\frac{8}{3} - \frac{1}{3}\right] = \frac{7}{27} //$$

Que:6 A frequency distribution f^n is defined by

$$f(x) = \begin{cases} x^3 & ; \quad 0 \le x \le 1 \\ 3(2-x)^3 & ; \quad 1 \le x \le 2 \end{cases}$$

PT $f(x)$ is a P.d.f & find ~~an~~. S.D.

Sol: To prove : $f(x)$ is P.d.f

i.e $\int_0^2 f(x)\,dx = 1$

$$\int_0^1 x^3\,dx + \int_1^2 3(2-x)^3\,dx = 1$$

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

$$\frac{1}{4} + \frac{3}{4} = \frac{4}{4} = 1$$

Hence Proved

Now Standard deviation $= \int_0^2 x^2 f(x)\cdot dx$

$$= \int_0^1 x^2 \cdot x^3 dx + 3\int_1^2 x^2(2-x)^3\,dx$$

$$= \int_0^1 x^5\,dx + \int_1^2 24x^2 + 3x^5 + 36x^3 - 18x^4\,dx$$

Date :

$$\left[\frac{x^6}{6}\right]_0^1 + \left[8x^3 + \frac{x^6}{2} + 9x^4 - \frac{18x^5}{5}\right]_1^2$$

$$\frac{1}{6} + \left[64 + 32 + 144 - \frac{576}{5}\right] - \left[8 + \frac{1}{2} + 9 - \frac{18}{5}\right]$$

$$\frac{1}{6} + \frac{624}{5} - \frac{139}{10}$$

$$\frac{3170}{30} = \frac{317}{3}$$

Que: 7 Find the value of K so that the following function represents p.d.f

$$f(x) = \begin{cases} 0 & ; \quad x \le -1 \\ K(x+1) & ; \quad -1 < x < 3 \\ 4K & ; \quad 3 \le x \le 4 \\ 0 & ; \quad \text{Otherwise} \end{cases}$$

Also find Median

Sol: W.K.T $\int_{-\phi}^{4} f(x) = 1$

$$\int_{-1}^{3} Kx + K + \int_{3}^{4} 4K = 1$$

$$\left[\frac{kx^2}{2} + kx\right]_{-1}^{3} + \left[4kx\right]_{3}^{4} = 1$$

$$\frac{9k}{2} + 3k - \frac{k}{2} + k + 16k - 12k = 1$$

$$12k = 1$$

$$\boxed{k = \frac{1}{12}}$$

By defination of Median

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

$$\int_{-1}^{3} kx + k \quad dx + \int_{3}^{m} 4k\, dx = \frac{1}{2}$$

$$\left[\frac{kx^2}{2} + kx\right]_{-1}^{3} + \left[4kx\right]_{3}^{m} = \frac{1}{2}$$

$$\frac{9k}{2} + 3k - \frac{k}{2} + k + 4mk - 12mk = \frac{1}{2}$$

$$\cancel{\frac{9}{2} \times \frac{1}{12}} \cancel{+3}$$

$$8k - 8mk = \frac{1}{2}$$

$$8k[1-m] = \frac{1}{2}$$

$$1 - m = \frac{1}{2 \times 8k} \qquad \left[k = \frac{1}{12}\right]$$

$$1 - m = \frac{1 \times \cancel{12}^{\cancel{6}\,3}}{\cancel{2 \times 8}_{4}}$$

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

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

$$\boxed{m = \frac{1}{4}}$$

Que: 8 If the function $f(x)$ is defined as

$$f(x) = \begin{cases} 0 & ; \; x < 2 \\ \frac{1}{18}(3+2x) & ; \; 2 \leq x \leq 4 \\ 0 & ; \; x > 4 \end{cases}$$

To Prove - $f(x)$ is a Pdf
also find $P(2 \leq x \leq 3)$.

Sol: W.K.T $\int_2^4 f(x) = 1$

$$\frac{1}{18}\int_2^4 3 + 2x = 1$$

$$\left[3x + x^2\right]_2^4 = 18$$

$$\frac{12 + 16 - [6 + 4]}{18} = 1$$

$$\frac{18}{18}$$

$\therefore$ It is a Pdf

Now $P(2 \le x \le 3) = \int_2^3 f(x)\,dx$

$= \frac{1}{18}\int_2^3 3 + 2x \; dx$

$= \frac{1}{18}\left[3x + x^2\right]_2^3$

$= \frac{1}{18}\left[[9+9] - (6+4)\right]$

$= \frac{1}{18}[18 - 10] = \frac{8}{18}$

$= \frac{4}{9}$

Que:9 A random variable x has a following probability function.

$X = x$:	0	1	2	3
$P(x)$:	0	$\frac{1}{5}$	$\frac{2}{5}$	$\frac{2}{5}$

WKT for P.d.f

$F(x) = P(X \le x) = \sum_{X \le x_i} P(x_i)$

Putting $X = 0, 1, 2, 3 \ldots$

1. $F(0) = P(x \leq 0) = P(0) = 0$

2. $F(1) = P(x \leq 1) = P(0) + P(1) = \frac{1}{5}$

3. $F(2) = P(x \leq 2) = P(0) + P(1) + P(2)$

$= \frac{3}{5}$

4. $F(3) = P(x \leq 3) = P(0) + P(1) + P(2) + P(3)$

$= 0 + \frac{1}{5} + \frac{2}{5} + \frac{2}{5} = 1$

Que: 10 Find the Mean, Variance & Standard deviation for the discrete random Variate 'X' having the following probability mass function.

Sol: $P(x) = \begin{cases} 1/3 & x = -2 \\ 1/3 & x = -1 \\ 1/3 & x = 1 \\ 0 & \text{Otherwise} \end{cases}$

Sol. $X = x$: -2 -1 1 Otherwise

$P(x)$: $1/3$ $1/3$ $1/3$ 0

Mean $E(x) = \sum x_i E(x_i$

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

S.D $E(x^2)$ $= E(x^2) = x_i^2 (E x_i) - x^2$

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

Variance $= E(x^2) - [E(x)]^2$

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

$= 2 - \frac{4}{9} = \frac{14}{9}$

Standard deviation $= \sqrt{\text{Variance}}$

$= \frac{\sqrt{14}}{3}$

Binomial distribution

$P(r) = {}^nC_r \, p^r \, q^{n-r}$

$P(r) = (p+q)^n$

where $p+q = 1$

${}^nC_r = \frac{\lfloor n}{\lfloor r \lfloor n-r}$

Page No. :
Date :

Que:1 Find the mean & variance of binomial distribution $r = 0, 1, 2, 3, \dots n$.

Sol: For mean we fill find the first moment about origin

Proof: $\mu_1' = r \cdot P(r)$

$= \Sigma r \cdot {}^nC_r \cdot p^r q^{n-r}$

$= \Sigma r \dfrac{\lfloor n}{\lfloor n-r \lfloor r} p^r q^{n-r}$

$= \Sigma r \dfrac{n \lfloor n-1}{r \lfloor r-1 \lfloor (n-1)-(r-1)} \cdot p \cdot p^{r-1} q^{(n-1)-(r-1)}$

$\mu_1' = np \, \Sigma^{n-1} C_{r-2} \, p^{r-1} q^{(n-1)-(r-1)}$

$= np(p+q)^{n-1}$

$= np(1)^{n-1}$ $[(1)^n = 1]$

$= np$

$\therefore$ $\boxed{\text{mean} = np}$

Similarly, we can show the 2nd moment about origin

$\mu_2' = r^2 \cdot P(r)$

$= npq + n^2p^2$

Que:1 Find the mean & variance of binomial distribution $x = 0, 1, 2, 3, \ldots n$.

Sol: For mean we will find the first moment about origin

Proof: $\mu_1' = r \cdot p(r)$

$= \sum r \, {}^nC_r \, p^r q^{n-r}$

$= \sum r \frac{\lfloor n}{\lfloor n-r \, \lfloor r} p^r q^{n-r}$

$= \sum r \frac{n \lfloor n-1}{r \lfloor r-1 \, \lfloor (n-1)-(r-1)} \cdot p \cdot p^{r-1} q^{(n-1)-(r-1)}$

$\mu_1' = np \sum {}^{n-1}C_{r-2} \, p^{r-1} q^{(n-1)-(r-1)}$

$= np (p+q)^{n-1}$

$= np (1)^{n-1} \quad [(1)^n = 1]$

$= np$

$\therefore \boxed{\text{mean} = np}$

Similarly, we can show the 2nd moment about origin

$\mu_2' = r^2 \cdot p(r)$

$= npq + n^2p^2$

Page No. :
Date :

$$\text{Variance} = \mu_2' - (\mu_1)^2$$
$$= npq - n^2p^2 - n^2p^2$$
$$= \underline{npq}$$

Que:2 In a binomial distribution the mean & S.D are 12 and 2 respectively. Find n & P.

$$m = np = 12$$
$$S.D = \sqrt{npq} = 2$$
$$q = \frac{4}{12} = \frac{1}{3} \qquad \boxed{P = \frac{1}{3}}$$
$$P + q = 1 \qquad \boxed{P = \frac{2}{3}}$$
$$n = \frac{12 \times 3}{2} = 18 \qquad \boxed{n = 18}$$

Que:3 Using the binomial distribution find the probabilit[illegible] of rolling

Atmost 2.

Sol: $n = 5 \qquad P = \frac{1}{6} \qquad q = \frac{5}{6}$

$P(\text{atmost } 2) = P(0) + P(1) + P(2)$

$= {}^nC_r \, p^r \, q^{n-r}$

$= 5C_0 \left(\frac{1}{6}\right)^0 \left(\frac{5}{6}\right)^5 + 5C_1 \left(\frac{1}{6}\right)^1 \left(\frac{5}{6}\right)^4 +$

$5C_2 \left(\frac{1}{6}\right)^2 \left(\frac{5}{6}\right)^3$

$= \left(\frac{5}{6}\right)^5 + \cancel{5} \times \left(\frac{5}{6}\right)^{5} + \frac{5 \times 4}{2 \times 1} \times \left(\frac{5}{6}\right)^3 \times \frac{1}{36}$

$= \left(\frac{5}{6}\right)^3 \left[1 + 1 + \frac{\cancel{10}\,5}{\cancel{36}\,18} \right]$

$= \frac{125}{216} \times 2.27$

$= \underline{\underline{1.3136}}$

Que:4 A Coin is tossed 4 times. What is the Probability of getting
① two heads
② atleast two heads.

Sol: $n = 4$, $P = \frac{1}{2}$, $q = \frac{1}{2}$

$P(r) = {}^nC_r \, p^r \, q^{n-r}$

$P(2 \text{ heads}) = 4C_2 \, p^2 \, q^2$

$= \frac{4 \times 3}{2 \times 1} \times \frac{1}{4} \times \frac{1}{4} = \underline{\underline{\frac{3}{8}}}$

Page No. :
Date :

P(atleast two heads) = P(2) + P(3) + P(4)
$= \frac{3}{8} + \frac{1}{4} + \frac{1}{16} = \frac{11}{16}$ //

Que 5 A cubical die is thrown in set of 8. The occurance of 5 or 6 is called a Success in what proportion of the set you expect 3 success.

Sol: $n = 8$ $\quad p = \frac{2}{6} = \frac{1}{3}$ $\quad q = \frac{2}{3}$

$P(3) = 8c_3 \left(\frac{1}{3}\right)^3 \left(\frac{2}{3}\right)^{8-3}$

$= \underline{\underline{0.2781}}$

Que: 6 The 10 percent of Screws produces in a certain factory turn out to be defective. find Probability that in a Sample of 10 Screw of random exactly 2 will be defective.

Sol: $P = \frac{1}{10}$ $\quad q = \frac{9}{10}$ $\quad n = 10$

$P(2) = 10c_2 (P)^2 (q)^8$

$= \frac{10 \times 9}{2 \times 1} \left(\frac{1}{10}\right)^2 \left(\frac{9}{10}\right)^8$

$= \underline{\underline{0.1937}}$

Que. 7 The probability of entering students in CA with graduate is 0.5 Determine the probability that out of 10 students (I) none (II) one (III) at least 1 will graduate.

Sol:

$$n = 10 \qquad P_1 = 0.5$$
$$q = 1 - P = 0.5$$

$$P(r) = {}^nC_r \cdot p^r \cdot q^{n-r}$$

$$P(0) = {}^nC_0 \cdot p^0 \, q^n$$
$$= \frac{|10}{|0 \, |10} \cdot (0.5)^0 \, (0.5)^{10}$$
$$= 0.0009765$$

$$P(1) = {}^nC_1 \, p^1 \, q^9$$
$$= 10 \times 0.5 \times (0.5)^9$$
$$= 0.009765$$

$$P(\text{at least one}) = P(1) + P(2) + \cdots + P(10)$$
$$= 1 - P(0)$$
$$= 0.9990235$$

Que: 8 Find the binomial distribution whose mean $m = 4$ and variance is 3 also find its mode.

Sol

$$\text{mean} = np = 4$$

$$\text{Variance} = npq = 3$$

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

$p + q = 1$

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

$\Rightarrow p = \frac{1}{4}$

$n = 4 \times 4$

$n = 16$

Mode = Integral part of $(np + p)$

$= \text{Integral part of } \left(\frac{16}{4} + \frac{1}{4}\right)$

$= \text{integral part of } \left[\frac{17}{4}\right]$

$= 4$

$\boxed{\text{Mode} = 4}$

Que: 9 ~~with~~ Fit a Binomial distribution to the following data and compare a theoratical frequencies with actual ones.

(1)

x :	0	1	2	3	4	5
f :	2~~4~~	14 ~~20~~	28	34	22	8

Sol: $n = 5$

$m = \frac{\Sigma f(x)}{\Sigma f}$

Date :

$$= \frac{2\times0+14\times1+20\times2+34\times3+22\times4+8\times5}{2+14+20+34+22+8}$$

$$= \frac{284}{100}$$

$$= 2.84$$

$$\text{Mean} = np = 2.84$$

$$5p = 2.84$$

$$p = \frac{2.84}{5}$$

$$p = 0.568$$

$$q = 1 - p$$
$$= 0.432$$

$$\text{Fit} = (p+q)^n$$

$$= (0.568 + 0.432)^5$$

or $P(r) = {}^nC_r\, p^r\, q^{n-r}$

Que: 10 A bag contain 3 red and 4 black balls, one ball is drawn and then replaced in the bag and the process is repeated. Getting a red ball in draw is considered a success. find the distribution of capital X, where X denote the number of success in 3 drawn, Assuming that in each draw each ball is exactly like to be selected.

Sol: Total balls = 7

red balls = $\frac{3}{7}$

black balls = 4/7

$P(r) = {}^nC_r \, p^r \, q^{n-r}$ $\boxed{n=3}$

$P(0) = {}^3C_0 \left(\frac{3}{7}\right)^0 \left(\frac{4}{7}\right)^3 = \frac{64}{343}$

$= \left(\frac{4}{7}\right)^3$

$P(1) = {}^3C_1 \left(\frac{3}{7}\right)^1 \left(\frac{4}{7}\right)^2 = \frac{9 \times 16}{343} = \frac{14}{34}$

$P(2) = {}^3C_2 \left(\frac{3}{7}\right)^2 \left(\frac{4}{7}\right)^1 = \frac{27 \times 4}{343} = \frac{10}{34}$

$P(3) = {}^3C_3 \left(\frac{3}{7}\right)^3 \left(\frac{4}{7}\right)^0 = \frac{27}{343}$

Que:11 Out of 800 families with 4 Children each How many families would be expected to have (i) Two boys & Two girls.
(ii) Atleast one boy.
(iii) No girl
(iv) at most two girls
Assume equal Probabilities for boy & girls.

Date :

Que : 12 The probability that a bomb drop from a plane will strike the target is 1/5, If 6 bomb are dropped find the probability that (i) exactly two will strike the target.

(ii) atleast two will strike the target

Sol: $P = \frac{1}{5}$ $q = 1 - P = \frac{4}{5}$ $n = 6$

$$P(r) = {}^nC_r \, p^r q^{n-r}$$

$$P(2) = {}^6C_2 \left(\frac{1}{5}\right)^2 \left(\frac{4}{5}\right)^4$$

$$= \frac{\cancel{6}^{3} \times 5}{\cancel{2} \times 1} \times \frac{1}{\cancel{25}\,5} \times \frac{256}{625}$$

$$= \frac{768}{3125} = 0.2457$$

$$P(\text{at least } 2) = 1 - \{P(0) + P(1)\}$$

$$= 1 - \left\{ {}^6C_0 \left(\frac{1}{5}\right)^0 \left(\frac{4}{5}\right)^6 + {}^6C_1 \left(\frac{1}{5}\right)^1 \left(\frac{4}{5}\right)^5 \right.$$

$$= 1 - \left\{ \left(\frac{4}{5}\right)^6 + \frac{6}{5} \times \left(\frac{4}{5}\right)^5 \right]$$

$$= 1 - \left[\frac{4096}{15625} + \frac{6144}{15625} \right]$$

Page No. :
Date :

$$= 1 - \frac{10240}{15625} = \frac{5385}{15625}$$

$$= 0.34464$$

Que: 13 The following data are the number of seeds germinating out of 10 on dam filter for 80 sets of seeds. Fit a Binomial distribution to these data.

x:	0	1	2	3	4	5	6	7	8	9	10
f:	6	20	28	12	8	6	0	0	0	0	0

Sol. Total $\Sigma f = 80$

$$\text{Mean} = \frac{\Sigma f(x)}{\Sigma f}$$

$$= \frac{20 + 56 + 36 + 32 + 30}{80}$$

$$= \frac{174}{80} = \frac{87}{40}$$

$$m = np = \frac{174}{80} \quad [n = 10]$$

$$10\,p = \frac{174}{80} \qquad p = \frac{174}{800} = 0.2175$$

$q = 1 - p = 0.7825$

$P(x) = (0.2175 + 0.7825)^{10}$

Also: Expected frequency $= N \cdot P(x)$

$= 80(0.2175 + 0.7825)^{10}$

Que: 14 In litters of 4 mice. The no. of ~~cat~~ litters, which contains 0, 1, 2, 3, 4 female were noted. The data are given below.

No of female mice :	0	1	2	3	4	Total
No. of litters :	8	32	34	24	5	103

If the chance of obtaining a female in a single trial is assumed constant, Estimate this const of unknown probability find also the expected frequency.

$$\text{mean} = \frac{\sum fx}{\sum f}$$

$$= \frac{32 + 68 + 72 + 20}{103} = \frac{192}{103}$$

$$= 1.86$$

mean $= np = \frac{192}{103}$

$p = \frac{192}{103 \times 4} = \frac{48}{103} = 0.466$

$q = 1 - p = \frac{55}{103} = 0.533$

$P(r) = (0.466 + 0.533)^4$

Expected frequency $= N \cdot P(r)$

$= 103 \cdot (0.466 + 0.533)^4$

Page No. :
Date :

Poisson - distribution :

$n \longrightarrow$ infinite

$$P(r) = \sum_{r=0}^{\infty} \frac{e^{-m} m^{r}}{\lfloor r}$$

where $m \rightarrow$ mean

Que: Find the mean & Variance of Poisson distribution
For mo

Sol: For mean, we will find first moment about origin.

$$\mu_1' = r . P(r)$$

$$= \sum_{r=0}^{\infty} r . \frac{e^{-m} m^{r}}{\lfloor r}$$

$$= \sum_{r=0}^{\infty} \cancel{r} . \frac{e^{-m} m^{r-1} m}{\cancel{r} \lfloor r-1}$$

$$= m \sum_{r=0}^{\infty} e^{-m} \frac{m^{r-1}}{\lfloor r-1}$$

$$= m \sum_{r=1}^{\infty} e^{-m} \frac{m^{r-1}}{\lfloor r-1}$$

$$= m \cdot e^{-m} \left[1 + m + \frac{m^2}{\lfloor 2} + \cdots \cdots \right]$$

$$= m \cdot e^{-m} \cdot e^{m}$$

$$= \underline{\underline{m}}$$

$$\mu_1' = m = \text{mean} = \underline{\underline{m}}$$

Similarly $\mu_2' = x^2 \cdot P(x)$

Second moment about origin

$$\mu_2 = m^2 + m$$

$$\text{Variance} = \mu_2' - (\mu_1')^2$$

$$= m^2 + m - m^2$$

$$= \underline{m}$$

In poisson distribution mean & variance are equal.

$$S.D = \sqrt{m}$$

Que: 2 Find the probability that atmost 5 defective fuses will be found in a ~~box~~ box of 200 fuses. If experience shows that 2 % of such fuses are defective.

Sol:
$$P = \frac{2}{100} = 0.02$$

$$m = np$$
$$= 200 \times 0.02$$
$$= 4.$$
$$\boxed{mean = 4}$$

$$P(r) = \sum \frac{e^{-m} \cdot m^{r}}{\lfloor r}$$

$$= \sum \frac{e^{-4} \cdot 4^{r}}{\lfloor r}$$

∴ at most 5 defective fuse

$$P(0) + P(1) + P(2) + P(3) + P(4) + P(5).$$

$$e^{-4} + \left[4e^{-4}!\right] + \left[\frac{e^{-4}\, \overset{8}{\cancel{16}}}{\cancel{2}}\right] + \left[\frac{\overset{32}{\cancel{64}}\, e^{-4}}{\lfloor 3}\right] + \left[\frac{e^{-4}\, 256}{24}\right]$$

$$+ \left[\frac{e^{-4}\ 1024}{120}\right]$$

$$5e^{-4} + 8e^{-4} + \frac{32e^{-4}}{3} + \frac{32e^{-4}}{3} + \frac{256e^{-4}}{30}$$

$$13e^{-4} + \frac{64e^{-4}}{3} + \frac{256e^{-4}}{30}$$

$$\frac{390e^{-4} + 640e^{-4} + 256e^{-4}}{30}$$

$$= \frac{1286e^{-4}}{30}$$

$$= \frac{643e^{-4}}{15} = \frac{643 \times 0.018}{15}$$

$$= 0.7849$$

Que.3 In a certain factory turning razor blades. There is a small chance of 0.002 for ~~f~~ any blade to be defective. The blades are in packets of 10. Use ~~Pis~~ Poisson distribution to calculate the appropriate number of packets.

Containing, (i) No defective
(ii) 1 defective
(iii) Two defective blades

respectively. In a Consinment of 50000 packets.

Sol:

$P = 0.002$

$m = np = m = 10 \times 0.002 = 0.02$

$\text{Mean} = 0.02$

$$P(r) = \sum \frac{e^{-m} m^r}{\lfloor r}$$

$$= \sum \frac{e^{-0.02} (0.02)^r}{\lfloor r}$$

(i) No defective

$$P(0) = \frac{e^{-0.02} \times (0.02)^0}{\lfloor 0} = e^{-0.02}$$

$$= 0.9802$$

$$P(1) = \frac{e^{-0.02} (0.02)^1}{1} = 0.01960$$

$$P(2) = \frac{e^{-0.02} (0.02)^2}{\lfloor 2} = 0.00019604$$

Now In Consinment = Expected frequency

$$\therefore N\ P(0) = 50{,}000 \times 0.9802$$
$$= \underline{\underline{49010}}$$

$$N(P(1)) = 50000 \times 0.01960$$
$$= \underline{\underline{980}}$$

$$N\ P(2) = 50000 \times 0.00019604$$
$$= \underline{\underline{9.802}}$$

Page No. :
Date :

Que: 4 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.

$$p = \frac{3}{100} = 0.03$$

$$m = n \times p = 3$$

$$n = 100.$$

$$P(r) = \sum \frac{e^{-m} m^r}{\lfloor r}$$

$$P(5) = \frac{e^{-3} 3^5}{120} = \underline{\underline{0.10081}}$$

Que: 5 A car hire firm two cars which it hires out day by day. The number of demands for a car on each day is distributed as a poisson distribution. If both mean = 1.5, Calculate the proportion of days on which neither car is used and the proportion of days on which some demand is refused.

Sol:

$$P(r) = \sum \frac{e^{-m} m^r}{\lfloor r}$$

$$P(r) = \sum \frac{e^{-1.5} (1.5)^0}{\lfloor 0}$$

Neither car is used

$$P(0) = \frac{e^{-1.5}(1.5)^0}{L0} = 0.2231$$

(ii) Some demand is refused :

$x > 2$

$P(3) + P(4) + \cdots$

$= 1 - \{P(0) + P(1) + P(2)\}$

$= 1 - \{0.223 + 0.33465 + 0.2509875\}$

$= 1 - \underline{0.8086375}$ ~~0.1385~~

$= \underline{\underline{0.1913625}}$ ~~0.8615~~

Que: 6 Fit a Poission distribution to the following Calculate theoritical frequencies

x :	0	1	2	3	4
f :	122	60	15	2	1

$$m = \frac{\Sigma f(x)}{\Sigma f} = 60 + 30 + 6 + 4$$

$$= \frac{100}{200} = \frac{1}{2}$$

$$= \underline{\underline{0.5}}$$

Page No.:
Date:

$$P(x) = \sum \frac{e^{-m} \cdot m^{x}}{\lfloor x}$$

$$P(0i) = \sum \frac{e^{-0.5} \cdot (0.5)^{r}}{\lfloor r}$$

$$P(0) = \sum \frac{e^{-0.5} \cdot (0.5)^{0}}{\lfloor 0} = 0.6065$$

$$P(1) = \sum \frac{e^{-0.5} (0.5)^{1}}{\lfloor 1} = 0.3032$$

$$P(2) = \sum \frac{e^{-0.5} (0.5)^{2}}{\lfloor 2} = 0.07581$$

$$P(3) = \sum \frac{e^{-0.5} (0.5)^{3}}{\lfloor 3} = 0.01263$$

$$P(4) = \sum \frac{e^{-0.5} (0.5)^{4}}{\lfloor 4} = 0.0015794$$

$N \cdot P(0) = 121.3$

$N \cdot P(1) = 60.64$

$N P(2) = 15.162$

$N P(3) = 2.526$

$N P(4) = 0.31588$

Page No. :
Date :

derivation

x:	0	1	2	3	4	5	6
f:	123	110	49	14	3	1	0

Qu: 7 A Skilled typist on routine work kept a record of mistakes made per day during 300 working days

Mistakes per day:	0	1	2	3	4	5	6
No. of days:	143	90	42	12	9	3	1

Fit a poission distribution to the above data & calculate expected frequencies

$$\text{Mean} = \frac{\Sigma f_i x_i}{\Sigma f_i} = \frac{90+84+36+36+15+6}{143+90+42+12+9+3+1}$$

$$= \frac{267}{300} = 0.89$$

$$P(r) = \Sigma \frac{e^{-m} m^r}{\lfloor r}$$

$$P(r) = \Sigma \frac{e^{-(0.89)} (0.89)^r}{\lfloor r}$$

$$P(0) = \frac{e^{-(0.89)} \cdot (0.89)^0}{\lfloor 0} = 0.41065$$

$$P(1) = \frac{e^{-(0.89)} (0.89)^1}{\lfloor 1} = 0.36548$$

Page No. :
Date :

$$P(2) = \frac{e^{-(0.89)}(0.89)^2}{\lfloor 2} = 0.16264$$

$$P(3) = \frac{e^{-(0.89)}(0.89)^3}{\lfloor 3} = 0.04824$$

$$P(4) = \frac{e^{-(0.89)}(0.89)^4}{\lfloor 4} = 0.010735$$

$$P(5) = \frac{e^{-(0.89)}(0.89)^5}{\lfloor 5} = 0.001910$$

$$P(6) = \frac{e^{-(0.89)}(0.89)^6}{\lfloor 6} = 0.000283$$

$N(P(0)) = 123.195$

$NP(1) = 109.643$

$NP(2) = 48.792$

$NP(3) = 14.4747$

$NP(4) = 3.2206$

$NP(5) = 0.5732$

$NP(6) = 0.08503557$

Date :

Que: 8 Prove that Poisson distribution has a limiting case of Binomial distribution when

1. n is large i.e $n \to \infty$
2. p is small.
3. $np = m$ is finite.

Sol: By defination of Binomial distribution

$$P(r) = {}^nC_r \cdot p^r \cdot q^{n-r}$$

$$= \frac{\lfloor n}{\lfloor r \lfloor n-r} \cdot p^r \cdot (1-p)^{n-r} \quad \because p+q=1$$

$$= \frac{\{n(n-1)(n-2)\ldots\{n-(r-1)\}\} \cdot \cancel{\lfloor n-r}}{\lfloor r \cdot \cancel{\lfloor n-r}}$$

$$\cdot \left(\frac{m}{n}\right)^r \left\{1 - \frac{m}{n}\right\}^{n-r}$$

$$= \frac{n^r \left\{\left(1-\frac{1}{n}\right)\left(1-\frac{2}{n}\right)\ldots\ldots\left[1-\frac{r-1}{n}\right]\right.}{\lfloor r \cdot n^r}$$

$$\cdot m^r \cdot \frac{(1-m/n)^n}{(1-m/n)^r}$$

$$= \lim_{n\to\infty} \frac{\left(1-\frac{1}{n}\right)\left(1-\frac{2}{n}\right)\ldots\ldots\left\{1-\frac{r-1}{n}\right\}}{\lfloor r}$$

$$\cdot m^r \cdot \frac{\left\{(1-m/n)^{-n/m}\right\}^{-m}}{(1-m/n)^r}$$

Page No. :
Date :

Normal distribution

$$y = \frac{1}{\sigma\sqrt{2\pi}} \cdot e^{-(x-m)^2/2\sigma^2}$$

where $m \rightarrow$ mean
$\sigma \rightarrow$ S.D

★ Area under the normal curve :- $Z = \frac{x-m}{\sigma}$

Que: 1 If X is a normal variate with mean = 30 & std deviation = 5
Find the probability that

(i) $26 \leq x \leq 40$

(ii) $X \geq 45$

(iii) $|X - 30| > 5$

Sol: W.K.T (i) $Z = \frac{x-m}{\sigma}$

$$Z = \frac{x-30}{5} = \frac{26-30}{5} = \frac{-4}{5}$$

$$= -0.8$$

also $Z = \frac{40-30}{5} = \frac{10}{5} = 2$

$P(26 \le X \le 40) = P(-0.8 \le Z \le 2)$

$= P(-0.8 \le z \le 0) + P(0 \le z \le 2)$.

$= 0.2881 + 0.4772$

$= 0.7653$

(ii) $Z = \frac{x - m}{\sigma} = \frac{x - 30}{6}$

$Z = \frac{45 - 30}{5} = 3$

$P(x \geq 45) = P(z \geq 3)$

$= 0.5 - P(0 \le z \le 3)$

$= 0.5 - 0.4986$

$= \underline{\underline{0.0014}}$

(iii) $|X - 30| > 5$

$P(|X - 30| \cancel{>} 5) = P(25 \le x \le 35)$

$= P(-1 \le X \le 1)$

$= 2P(0 \le X \le 1)$

$= 2 \times 0.3413$

$= \underline{\underline{0.6826}}$

Date :

$= 1 - 0.6826$

$= 0.3174$

Que. 2 A Certain type of wooden beam as a mean breaking strength 1500 kg and Standard deviation of $\sigma = 100$ kg Find the relative frequencies of all such beam where breaking strength are between 1450 and 1600 kg

$$Z = \frac{x - m}{\sigma} \qquad (1450 \le x \le 1600)$$

$$Z = \frac{1450 - 1500}{100} = \frac{-50}{100} = -0.5$$

$$Z = \frac{1600 - 1500}{100} = \frac{100}{100} = 1$$

$$P(1450 \le x \le 1600) = P(-0.5 \le Z \le 1)$$

$$= P(-0.5 \le Z \le 0) + P(0 \le x \le 1)$$

$$= 0.1915 + 0.3413$$

$$= 0.5328$$

Date :

Que:3 The mean height of 500 Students is 151 cm and the Standard deviation is 15 cm assuming that the heights are normally distributed, Find how many Student have height between $P(122 < x < 155)$ cm

$$Z = \frac{x - m}{\sigma}$$

$$(122 \leq x \leq 155)$$

$$Z = \frac{122 - 151}{15} = \frac{-29}{15} = -1.93$$

$$Z = \frac{155 - 151}{15} = 0.26$$

$$P(122 < x < 155) = P(-1.93 < z < 0.26)$$

$$= P(-1.93 < z < 0) + P(0 \leq z < 0.2$$

$$= 0.4713 + 0.0793$$

$$= \underline{\underline{0.5506}}$$

Expected freq N $= 500 \times 0.5506$

$$= \underline{\underline{275.3}}$$

Que:4 The distribution of weekly wages of 500 workers in a factory is approximately Normal with mean and S.D of

Rupees 75 and 15 respectively. Find the No. of workers who receive weekly wages

1. more than 90 (>90)
2. Less than 45 (<45)

Sol:

$$z = \frac{x-m}{\sigma}$$

$$z = \frac{90-75}{15} = \frac{15}{15} = 1$$

$$z = \frac{45-75}{15} = -2$$

$$P(x>90) = P(z>1)$$

$$= 0.5 - P(0<x<1)$$

$$= 0.5 - 0.3413$$

$$= \underline{0.1587}$$

$$P(x<45) = P(z<-2)$$

$$= 0.5 - P(-2>x>0)$$

$$= 0.5 - 0.4772$$

$$= \underline{0.0228}$$

Expected freq $(x>90) = 0.1587 \times 500$
$= \underline{\underline{79.35}}$

Expected freq $(x<45) = 0.228 \times 500$
$= \underline{\underline{11.4}}$

Que:5 For the normal (curve) Equation

$$y = f(x) = \frac{1}{\sigma\sqrt{2\pi}} \cdot e^{-(x-m)^2/2\sigma^2}$$

Find the mean and Standard deviation.

Sol: $\because$ $f(x)$ is P.d.f

Then mean $= \int_{-\infty}^{\infty} x \cdot f(x)$

$$= \int_{-\infty}^{\infty} x \cdot \frac{1}{\sigma\sqrt{2\pi}} \cdot e^{-(x-m)^2/2\sigma^2} \, dx$$

put $\frac{(x-m)}{\sqrt{2}\sigma} = t$

$$\frac{dx}{\sqrt{2}\sigma} = dt$$

$\therefore$ $x - m = \sqrt{2}\sigma t$ $\quad\therefore \boxed{x = m + \sqrt{2}\sigma t}$

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

Date :

$$= \frac{1}{\sqrt{\pi}} \int_{-\infty}^{\infty} m \cdot e^{-t^2} dt + \int_{-\infty}^{\infty} \sqrt{2}\, \sigma\, t\, e^{-t^2} dt$$

$$= \frac{2m}{\sqrt{\pi}} \int_{0}^{\infty} e^{-t^2} dt + 0$$

$$= \frac{2m}{\sqrt{\pi}} \cdot \frac{1}{2} \int_{0}^{\infty} e^{-x} \cdot x^{-1/2} dx$$

putting $t^2 = x$

$2t\,dt = dx$

$dt = \frac{dx}{2\sqrt{x}}$

$$= \frac{m}{\sqrt{\pi}} \int_{0}^{\infty} e^{x} \cdot x^{1/2 - 1} dx$$

$$\boxed{\Gamma n = \int_{0}^{\infty} e^{-x} x^{n-1} dx \qquad \Gamma 1/2 = \sqrt{\pi}}$$

$$= \frac{m}{\sqrt{\pi}} \Gamma 1/2$$

$$= \frac{m}{\cancel{\sqrt{\pi}}} \cancel{\sqrt{\pi}}$$

$$= m$$

Similarly

$$\text{Variance} = \int_{-\infty}^{\infty} (x-m)^2 \cdot f(x)\, dx = \sigma^2$$

$$S.D = \sigma$$

Que:6 The life time of a certain kind of battery has a mean of 300 hours and Standard deviation 35 hours.
Assuming that the distribution of life time which are measured to the nearest hour is normal. Find the percentage of battery of more than 370 hour

$m = 300$
$\sigma = 35$

$Z = \frac{x - m}{\sigma} = \frac{370 - 300}{35}$

$Z = 2$

$(P > 370) = (z > 2)$
$= 0.5 - (0 < z < 2)$
$= 0.5 - 0.4772$
$= 0.0228$

Percentage $= 2.28\%$

Que: Find the expected value of discrete random Variable x ppm is

$f(x) = P(x) = \frac{2}{3}\left(\frac{1}{3}\right)^x, \; x = 0, 1, 2, \ldots$

Page No. :
Date :

Que: Calculate mean & variance of the throwing & an unbiased dice.

$$E(x) = \Sigma x \cdot P(x)$$

$$= \{1 \cdot P(1) + 2 \cdot P(2) + 3P(3) + 4(P(4)) + 5P(5) + 6P(6))\}$$

$$= \left\{1 \cdot \frac{1}{6} + 2 \cdot \frac{1}{6} + 3 \cdot \frac{1}{6} + 4 \cdot \frac{1}{6} + 5 \cdot \frac{1}{6} + 6 \cdot \frac{1}{6}\right\}$$

$$= \frac{1+2+3+4+5+6}{6}$$

$$= \frac{6(6+1)}{2 \times 6} = \frac{7}{2} = 3.5$$

$$\Sigma(x^2) = \Sigma x^2 P(x)$$

$$= \{1^2 \cdot P(1) + 2^2 P(2) + 3^2 P(3) + 4^2 P(4) + 5^2 P(5) + 6^2 P(6)\}$$

$$= \frac{1^2 + 2^2 + 3^2 + 4^2 + 5^2 + 6^2}{6}$$

$$= \frac{6(6+1)(2 \times 6 + 1)}{6 \times 6}$$

$$= \frac{7 \times 13}{6}$$

$$= 15.6$$

$$Var(x) = E(x^2) - \{E(x)\}^2$$

$$= 15.16 - (3.5)^2$$

$$= \underline{\underline{2.91}}$$

The bernoulli's distribution or trials is introduced by Jacob Bernoulli, is one of the simplest yet most important random process in probability

There are 3 assumptions in bernoulli's trial distribution :-

1) Each trial has 2 possible outcomes called success & failure.

2) The trials are independent, the outcome of 1 trial has no influence over the outcome of other trial.

3) On each trial, the probability of success is small p, the probability of failure is (1 - p).

A discrete random variable X is said to bernoulli distribution with parameter p with its mass function, is given by;

Page No. :
Date :

$$P(X=x) = \begin{cases} p^x (1-p)^{1-x} & ; x = 0, 1 \\ 0 & ; \text{otherwise.} \end{cases}$$

$P(X=0) = 1-p.$

$P(X=1) = p.$

Chebyshev inequality :-

$\mu \pm 2\sigma \longrightarrow 75\%.$

$\mu \pm 3\sigma \longrightarrow 88.9\%$

The variance of random variable X gives us an idea about the variability of the observations about mean.

Chebyshev inequality gives us bound on probability that how a random variable is deviated when both mean & variance are known / given.

<u>Statement</u> $\rightarrow$ If x is random variable with mean μ and variance σ^2, then for any value of $k > 0$.

$$\boxed{P\{|x-\mu| \geq k\} \leq \frac{\sigma^2}{k^2}}$$

Date :

Proof :- We prove the above inequality for continuous random variable

$$E(x) = \int_0^\infty x \cdot P(x)\, dx \quad \text{—— (1)}$$

$$= \int_0^a x \cdot P(x)\, dx + \int_a^\infty x \cdot P(x)\, dx .$$

$$E(x) \geq a \int_a^\infty P(x)\, dx$$

$$\frac{E(x)}{a} \geq \int_a^\infty P(x)\, dx$$

$$\frac{E(x)}{a} \geq P(x \geq a) \quad \text{—— (2)}$$

Eq (2) is called Markov inequality

In eq (2) we put $x = (x-\mu)^2$
$a = k^2$

$$P\left((x-\mu)^2 \geq k^2\right) \leq \frac{E(x-\mu)^2}{k^2}$$

$$P(|x-\mu| \geq k) \leq \frac{\text{Var } x}{k^2} = \frac{\sigma^2}{k^2}$$

$$\boxed{P(|x-\mu| \geq k) \leq \frac{\text{Var } x}{k^2} = \frac{\sigma^2}{k^2}}$$

Hence proved.

Page No.:
Date:

Que: A random variable X with unknown distribution has mean $\mu = 8$ and Varian[ce] $\sigma^2 = 9$

(i) $P(-4 < x < 20)$
(ii) $P(|x-8| \geq 6)$

Sol: W.K.T $P(|x-\mu| \geq k) \leq \frac{\sigma^2}{k^2}$

(ii) $P(|x-8| \geq k) \leq \frac{9}{k^2}$

$\because k = 6$

$P(|x-8| \geq 6) \leq \frac{9}{36} = \frac{1}{4}$.

eg: $\because k = 3\lambda$

$P(|x-8| \geq 3\lambda) \leq \frac{1}{\lambda^2}$

$P(|x-8| < 3\lambda) \leq 1 - \frac{1}{\lambda^2}$

$P(8-3\lambda < x < 8+3\lambda) \leq 1 - \frac{1}{\lambda^2}$

$P(-4 < x < 20) \leq \frac{15}{16}$

[Q]ue: $f(x) = k(1-x^2) \qquad 0 < x < 1$

(1) Find k
(2) Mean
(3) Variance.

Date :

Sol: As we know that

Total probability = 1

$$\int_0^1 f(x)\, dx = 1$$

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

$$k\left[x - \frac{x^3}{3}\right]_0^1 = 1$$

$$k\left[1 - \frac{1}{3}\right] = 1$$

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

$$\boxed{k = \frac{3}{2}}$$

$$\text{Mean} = \int_0^1 x \cdot f(x)\, dx$$

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

$$= \frac{3}{2}\int_0^1 x - x^3\, dx$$

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

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

Page No. :
Date :

Variance $= E(x^2) - [E(x)]^2$

Now. $E(x^2) = \frac{3}{2}\int_0^1 x^2 - x^4 \, dx.$

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

$= \frac{3}{2}\left[\frac{1}{3} - \frac{1}{5}\right] = \frac{\cancel{3}}{\cancel{2}} \times \frac{\cancel{2}}{\cancel{15}} = \frac{1}{5}$

$\therefore$ Variance $= \frac{1}{5} - \left(\frac{3}{8}\right)^2$

$= \frac{1}{5} - \frac{9}{64} = \frac{64-45}{320} = \frac{19}{320}$

Median of Normal Distribution
Median (Md)

Let Median be Md.

$\int_0^{md} f(x)\,dx = \int_{Md}^1 f(x)\,dx = \frac{1}{2}$

$\int_0^{md} \frac{3}{\cancel{2}}(1-x^2)\,dx = \frac{1}{\cancel{2}}$

$\left[x - \frac{x^3}{3}\right]_0^{md} = \frac{1}{3}$

$md - \frac{md^3}{3} = \frac{1}{3}$

Date :

$md^3 - 3md + 1 = 0$

Que: Fit a Normal Curve to the following data

Length of Line Lin(cm): 8.60 8.59 ~~8.58~~

Length of line (in cm)	8.60	8.59	8.58	8.57	8.56	8.55	8.54	8.53	8.
Frequency	2	3	4	9	10	8	4	1	1

Sol: Let the normal curve by $y = \frac{N}{\sigma\sqrt{2\pi}} e^{-(x-m)^2/2\sigma^2}$

$$N = \Sigma f = 42$$

$$\text{mean } (m) = A + \frac{\Sigma f u}{\Sigma f}$$

$$S.D\ (\sigma) = \sqrt{\frac{\Sigma f u^2}{\Sigma f} - \left(\frac{\Sigma f u}{\Sigma f}\right)^2}$$

$$= \sqrt{\frac{0.0132}{42} - \left(\frac{0.11}{42}\right)^2}$$

$$= \sqrt{\frac{0.133}{42} - \frac{0.0121}{1764}}$$

$$\sigma = \sqrt{\frac{234.1038}{74088}} = 0.0175$$

$$= \cancel{0.0}562$$

$$\sigma = \sqrt{0.0081598}$$

x	f	$u = x - A$, $A = 8.56$	u^2	fu	fu^2
8.60	2	0.04	0.0016	0.08	0.0032
8.59	3	0.03	0.0009	0.09	0.0027
8.58	4	0.02	0.0004	0.08	0.0016
8.57	9	0.01	0.0001	0.09	0.0009
8.56	10	0	0	0	0
8.55	8	−0.01	0.0001	−0.08	0.0008
8.54	4	−0.02	0.0004	−0.08	0.0016
8.53	1	−0.03	0.0009	−0.03	0.0009
8.52	1	−0.04	0.0016	−0.04	0.0016
				$\Sigma fu = 0.11$	$\Sigma fu^2 = 0.0133$

$$\text{Mean} = 8.56 + \frac{0.11}{42}$$

$$= \underline{\underline{8.5626}}$$

$$y = \frac{42}{0.0562 \times \sqrt{2\pi}} e^{-(x-m)^2/2(0.0175)^2}$$

Date :

Gamma distribution

A continuous random variable X is said to be gamma distribution with parameter λ if its p.d.f is given by

$$\Gamma\lambda = \int_0^\infty e^{-x} . x^{\lambda-1} dx$$

$$f(x) = \begin{cases} \dfrac{e^{-x} . x^{\lambda-1}}{\Gamma\lambda}, & 0 < x < \infty, \text{ where } \lambda > 0 \\ 0 & \text{Otherwise} \end{cases}$$

Que:1 Find the mean and variance of the gamma distribution.

(1) mean $\mu_1' = E(x) = \int_{-\infty}^{\infty} x . f(x) dx$

$$\mu_1' = \int_0^\infty \frac{x . e^{-x} . x^{\lambda-1}}{\Gamma\lambda} dx$$

$$= \frac{1}{\Gamma\lambda} \int_0^\infty e^{-x} . x^{\lambda} dx$$

$$= \frac{1}{\Gamma\lambda} \int_0^\infty e^{-x} x^{(\lambda+1)-1} dx$$

By defination of gamma

$$= \frac{1}{\Gamma\lambda} \cdot \Gamma\lambda+1$$

$$= \frac{1}{\cancel{\Gamma\lambda}} \cdot \lambda \cancel{\Gamma\lambda} \qquad \Gamma\lambda+1 = \lambda\Gamma\lambda$$

$\boxed{\text{Mean} = \lambda}$

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

$$u_2^2 = \int_0^{\infty} \frac{x^2 . e^{-x} . x^{\lambda-1}}{\Gamma\lambda} dx$$

$$= \frac{1}{\Gamma\lambda} \int_0^{\infty} e^{-x} x^{\lambda+1} dx$$

$$= \frac{1}{\Gamma\lambda} \int_0^{\infty} e^{-x} . x^{(\lambda+2)-1} dx$$

$$= \frac{1}{\Gamma\lambda} \Gamma\lambda+2 \qquad = \frac{1}{\Gamma\lambda} \cdot \Gamma(\lambda+1)+1$$

$$= \frac{1}{\Gamma\lambda} (\lambda+1) \Gamma\lambda+1$$

$$= \frac{1}{\cancel{\Gamma\lambda}} (\lambda+1) \lambda \cancel{\Gamma\lambda}$$

$$= \lambda(\lambda+1)$$

Date :

Exponential Distribution:

A random variable 'x' is said to be exponential distribution with parameter $\lambda > 0$ if x its p.d.f is given by

$$f(x) = \begin{cases} \lambda e^{-\lambda x} & ; x \geq 0 \\ 0 & ; \text{otherwise.} \end{cases}$$

Que: 1 Show that the total P.d.f of exponential distribution is unity.

Proof : Total area $= \int_{-\infty}^{\infty} f(x)\, dx$

$$= \int_{0}^{\infty} f(x) \cdot dx$$

$$= \int_{0}^{\infty} \lambda e^{-\lambda x}\, dx$$

$$= \left[\frac{\cancel{\lambda} \cdot -e^{\lambda x}}{\cancel{\lambda}} \right]_0^{\infty}$$

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

$$= 1$$

Hence proved

Que 2 Find the mean and variance of exponential distribution.

Mean $\rightarrow$ first moment about origin.

$$\mu_1' = \int_{-\infty}^{\infty} x \cdot f(x)\, dx$$

$$= \int_{-\infty}^{0} \underset{0}{f(x)\, dx} + \int_{0}^{\infty} x \cdot f(x)\, dx$$

$$= \int_{0}^{\infty} \lambda \cdot e^{-\lambda x} \cdot x\, dx$$

[U.V.

$$= \lambda \int_{0}^{\infty} e^{-\lambda x} \cdot x\, dx$$

$$= \lambda \left[\frac{\cancel{x}\,x \cdot e^{-\lambda x}}{\lambda} + \frac{\cancel{x^2}\, e^{-\lambda x}}{\lambda^2} \right]_0^{\infty}$$

$$= \frac{1}{\lambda} \quad //$$

$$\text{Mean} = \frac{1}{\lambda}$$

$$\mu_2' = \int_{0}^{\infty} x^2 \cdot f(x)\, dx$$

$$= \int_{0}^{\infty} x^2 \cdot \lambda e^{-\lambda x}\, dx$$

$$= \lambda \left[x^2 \cdot e^{-\lambda x}\, dx \right]_0^{\infty}$$

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

$$= 0 + \lambda \left[2 \cdot \frac{1}{\lambda^3} \right]$$

$$\mu_2' = \frac{2}{\lambda^2}$$

$$\text{Variance} = \mu_2' - (\mu_1')^2$$

$$= \frac{2}{\lambda^2} - \frac{1}{\lambda^2}$$

$$= \frac{1}{\lambda^2}$$

$$S.D = \sqrt{\text{Variance}}$$

$$= \frac{1}{\lambda}$$

SHEPPARD'S CORRECTION OF MOMENTS:

When moments are calculated in case of class interval, It is supposed that the frequency of each class interval is centered at the midpoint of the respective class. There is probability of some error in the values of the moment. To remove these error is called the Sheppard's Correction of moments.

Page No. :
Date :

μ_1 (corrected) $= \mu_1 = 0$

μ_2 (—"—) $= \mu_2 - \frac{1}{12}h^2$

μ_3 (—"—) $= \mu_3$

μ_4 (—"—) $= \mu_4 - \frac{1}{2}h^2\mu_2 + \frac{7}{240}h^4$

Note

σ_1^2 is the value of variance which is obtained from the grouped data and σ^2 be the corrected value then

$$\sigma^2 = \sigma_1^2 - \frac{1}{12}h^2$$

Que: The value of 2nd, 3rd & 4th moment are $\mu_2 = 88.75$

$\mu_3 = -131.25$

$\mu_4 = 25445.3125$

Calculate the corrected moment when the class interval is 10

μ_2(corrected) $= 88.75 - \frac{1}{12} \times 100$

$= 80.417$

Date :

μ_3 (corrected) $= -131.25$

μ_4 (corrected) $= 25445.3125 - \frac{100}{2} \times 80.41$

$+ \frac{7}{240} \times 10000$

$= 25445.3125 - 4020.85 + 291.6667$

$= 21716.1295$

KARL PEARSON'S, α, β, γ Coefficients.

ALPHA coeff	Beta coeff	gamma coeff
$\alpha_1 = \frac{\mu_1}{\sigma} = 0$	$\beta_1 = \frac{\mu_3^2}{\mu_2^3}$	$\gamma_1 = \pm\sqrt{\beta_1}$
$\alpha_2 = \frac{\mu_2}{\sigma^2} = 0$	$\beta_2 = \frac{\mu_4}{\mu_2^2}$	$\gamma_2 = \beta_2 - 3 = \frac{\mu_4 - 3\mu_2^2}{\mu_2^3}$
$\alpha_3 = \frac{\mu_3}{\sigma^3} = \frac{\mu_3}{(\mu_2)^{3/2}}$		
$\alpha_4 = \frac{\mu_4}{\sigma^4} = \mu_4/(\mu_2)^2$		

Que: The sign of gamma 1 depends on μ_3.
If μ_3 is positive then gamma 1 is +ve
If μ_3 is -ve then gamma 1 is -ve.

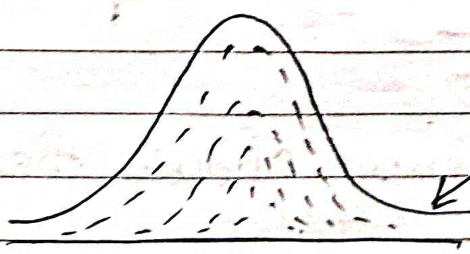
* Coefficient of Skewness based on moments

When there is Symmetrical distribution, all the moment of odd order about the arithmatic mean i.e. ($\mu_1, \mu_3, \mu_5 \dots$) Vanish.
If the value of those Coefficient do not Vanish then there is Skewness in the frequency distribution.

- First Coefficient of Skewness $= \upsilon$

$$= \frac{\mu_3}{\sqrt{\mu_2^3}} = \sqrt{\beta_1} = \alpha_3 = \gamma_1$$

- Second $= \dfrac{\sqrt{\beta_1}\,(\beta_2 + 3)}{2\{5\beta_2 - 6\beta_1 - 9\}}$



Leptokurtosis

Platukurto-sis.

Mesokurtists

Date :

Kurtosis measures the degree of peakedness of a distribution and is given by

$$B_2 = \frac{\mu_4}{\mu_2^2}$$

$$\mu_2 = B_2 - 3$$

gives the access of Kurtosis

if $B_2 > 3$ is Leptokurtosis

if $B_2 = 3$ is Mesokurtosis.

if $B_2 < 3$ is Platukurtosis.

Que:1 The first 4 moment about the working mean i.e 28.5 of a distribution are 0.294, 7.144, 42.409 and 454.98 Calculate the moment about the mean. also Evaluate B_1, B_2 & Comments upon the Skewness and Kurtosis of the distribution.

Sol: The first 4 moments about the arbitrary

$m = 28.5$ $\qquad$ $\mu_4' = 454.98$

$\mu_1' = 0.294$.

$\mu_2' = 7.144$.

$\mu_3' = 42.409$

Page No. :
Date :

Now $\mu_1' = \frac{1}{N} \Sigma f_i x_i - 28.5$

$0.294 = \bar{x} - 28.5$

$\bar{x} = 28.794$

$\mu_2 = \mu_2' - (\mu_1')^2$

$= 7.144 - 0.0864$

$= \underline{\underline{7.0576}}$

$\mu_3 = \mu_3' - 3\mu_2' . \mu_1' + 2\mu_1'^3$

$= 42.409 - 6.301008 + 0.05082$

$= \underline{\underline{36.158}}$

$\mu_4 = \mu_4' - 4\mu_3' . \mu_1' + 6\mu_2' . \mu_1'^2 - 3(\mu_1')^4$

$= 454.98 - 49.872984 + 3.70499 - 0.0224$

$= \underline{\underline{408.789}}$

$B_1 = \frac{\mu_3^2}{\mu_2^3} = \frac{1307.40}{49.81} = 26$

Page No. :
Date :

Que: Calculate the median Quartiles and the Quartile Coefficient of Skewness from the following data

Weight (Lbs):	70-80	80-90	90-100	100-110	110-120	120-130	130-140	140-150
No of persons:	12	18	35	42	50	45	20	8

$\Sigma f = 230$:

CF :- = 12, 30, 65, 107, 157, 202, 222, 230

Sol: Now $\frac{N}{2} = \frac{230}{2} = 115^{th}$ Lies between 110-120 group

$\therefore$ Median or $Q_2 = L + \frac{\frac{N}{2} - C^f}{f} \times h$

$= 110 + \frac{115 - 107}{50} \times 10$

$= 110 + \frac{8}{5} = 111.6$

also $\frac{N}{4} = \frac{230}{4} = 57.5$ i.e $Q_1 = 58^{th}$

which lies between 90-100

Page No.:
Date:

$$Q_1 = L + \frac{N/4 - c}{f} \times h = Q$$

$$= 90 + \frac{57.5 - 30}{35} \times 10 = 97.85$$

Similary $\frac{3N}{4} = 172.5$ Lies Q_3 in 120 - 130

$$Q_3 = L + \frac{\frac{3N}{4} - c}{f} \times h$$

$$= 120 + \frac{175.5 - 157}{45} \times 10$$

$$= 123.44$$

Hence quartile Coefficient of Skewness

$$= \frac{Q_1 + Q_3 - 2Q_2}{Q_3 - Q_1} = \frac{-1.91}{25.59}$$

$$= -0.0746$$

PROBLEMS RELATED TO EXPONENTIAL DISTRIBUTION

Que: 1 A random variable x has an expo. distribution with p.d.f is given by.

$$f(x) = \begin{cases} 2e^{-x} ; & \text{if } x > 0 \\ 0 ; & \text{if } x \leq 0. \end{cases}$$

Sol: Compute the probability that $x \nless 3$. also find the mean and Std dev. and PT Coefficient of variation is unity.

Sol:- By defination of exponential distribution

$$E(\text{dis}) = \int_0^{\infty} f(x)\, dx$$

$$= 2\int_3^{\infty} e^{-x}\, dx$$

$$= 2 - \left[e^{-x}\right]_3^{\infty}$$

$$= 2\left[-e^{-\infty} + e^{-3}\right]$$

$$= e^{-6}$$

(II) To find mean $\Rightarrow \mu_1' = E(x)$

$$= \int_{-\infty}^{\infty} x . f(x)\, dx = 2\int_0^{\infty} x . e^{-x}\, dx$$

$$= 2\left[-x . e^{-x} - 1 . e^{-x}\right]_0$$

Page No. :
Date :

$$= \left| 2e^{-x}\,[+x+1] \right|_0^{\infty}$$

$$= +2.e^{0}$$

$$= \underline{2}$$

$$\mu_2' = \int_0^{\infty} x^2 \cdot (2e^{-x})\,dx.$$

$$= 2\left[-x^2e^{-x} - 2xe^{-x} + 2e^{-x}\right]_0^{\infty}.$$

$$= +2\left[2\right]$$

$$= \underline{\underline{+4}}$$

Variance $= \mu_2' - (\mu_1')^2$
$= 4 - 4$
$= 0$

Que:2 The Income tax X, of a man has an exponential distribution with P.d.f is given by

$$f(x) = \begin{cases} \frac{1}{4}\,e^{-x/4} & ; x \geq 0 \\ 0 & ; x < 0 \end{cases}$$

If income tax is levied at the rate of 5%. What is the probability that his income exceed Rs 10,000.

Solⁿ: If Income exceed Rs 10000, then income tax will be income

$$\text{income tax} = 10000 \times \frac{5}{100} = 500$$

∴ required probability :-

$$P(x > 500) = \int_{500}^{\infty} f(x)\, dx.$$

$$= \frac{1}{4} \int (e^{-x/4})\, dx.$$

$$= \frac{1}{4} \left(\frac{e^{-x/4}}{-1/4} \right)_{500}^{\infty}$$

$$= \cancel{\frac{1}{4}} - (0 - e^{-125}$$

$$= e^{-125}$$

Que: The lifetime of a certain kind of battery is a random variable, which is as an exponential distribution with a mean of 200 hrs.

Find the probability that such a battery will last.

1) at most 100 hrs and

2) last any where from 400 to 600 hrs.

Page No. :
Date :

$$\text{exp. dis} = \int_0^\infty \lambda e^{-\lambda x}\, dx$$

$$m = \frac{1}{\lambda}$$

$$\frac{1}{\lambda} = 200$$

$$\lambda = \frac{1}{200}$$

① atmost 100 $= \frac{1}{200}\int_0^{100} e^{-\frac{1}{200}x}\, dx.$

$$= \frac{1}{\cancel{200}}\left[\frac{-\cancel{200}}{200} e^{-1/200\, x}\right]_0^{100}$$

$$= -e^{-1/2} + 1$$

② 400 - 600 $= \int_{400}^{600} \frac{1}{200} \times e^{-\frac{1}{200}x}\, dx$

CORRELATION:

If an increase (or decrease) in the values of one variable corresponds to an increase (or decrease) in the other. The Correlation is said to be positive.

If an increase (or decrease) in the values of one variable corresponds to or decrease (or increase) in the other. The Correlation is said to be negative.

If variables are independent then they are independent of correlation.

Coefficient of Co-relation:

$$r = \frac{\Sigma XY}{n \sigma x \cdot \sigma y}$$

where $X = x - \bar{x}$
$Y = y - \bar{y}$

σx = S.D of x series
σy = S.D of y series

n = No of values of the two variable

Direct method.

$$r = \frac{\Sigma xy}{\sqrt{\Sigma x^2 \cdot \Sigma y^2}}$$

Date :

another form

$$r = \frac{n\Sigma xy - \Sigma x \cdot \Sigma y}{\sqrt{\{n\Sigma x^2 - (\Sigma x)^2\} \times \{n\Sigma y^2 - (\Sigma y)^2\}}}$$

Que: Psychological tests of intelligence and of engineering ability were applied to 10 Students. Here is a record of ungrouped data showing intelligence ratio (I.R) and engineering ratio (E.R). Calculate the Co-efficient of Correlation.

Student:	A	B	C	D	E	F	G	H	I	J
IR:	105	104	102	101	100	99	98	96	93	92
ER:	101	103	100	98	95	96	104	92	97	94

$\bar{y} = 98$

Student's	Intelligence x; $x - \bar{x} = X$	X^2	Y^2	Engineering ratio y; $y - \bar{y} = Y$	XY
A	6	36	9	3	18
B	5	25	25	5	25
C	3	9	4	2	6
D	2	4	0	0	0
E	1	1	9	−3	−3
F	0	0	4	−2	0
G	−1	1	36	6	−6
H	−3	9	36	−6	18
I	−6	36	1	−1	6
J	−7	49	16	−4	28
		Σ 170	140		92
		ΣX^2	ΣY^2		ΣXY

$\Sigma x^2 = 170$ $\Sigma y^2 = 140$ $\Sigma xy = 92$

$$\therefore \frac{\Sigma xy}{\sqrt{\Sigma x^2 \cdot \Sigma y^2}}$$

$$\frac{92}{\sqrt{170 \cdot 140}} = \frac{92}{154.27} = 0.596$$

Q- A computer operator while calculating the coefficient between 2 variates x & y for 25 pairs of observations obtained the following constants :-

$n = 25$; $\Sigma x = 125$; $\Sigma x^2 = 650$;
$\Sigma xy = 508$; $\Sigma y = 100$; $\Sigma y^2 = 460$

It was however latter discovered at the time of checking that he had copied down 2 pairs as (6, 14) & (8, 6) while the correct pairs were (8, 12) & (6, 8) obtain the correct value of the correlation coefficient.

Sol: Given:

	Incorrect data	Correct data
x_1	6	8
x_2	8	6
y_1	14	12
y_2	6	8

Page No. :
Date :

Incorrect data are $\Sigma x = 125$; $\Sigma y = 100$; $\Sigma x^2 = 650$; $\Sigma y^2 = 460$; $\Sigma xy = 508$.

Then Corrected Σx = Incorrected $\Sigma x - (6+8) + (8+6)$

$= 125 - 14 + 14 = \underline{\underline{125}}$

Corrected Σy = Incorrected $\Sigma y - (14+6) + (12+8)$

$= 100 - 20 + 20 = \underline{\underline{100}}$

Corrected x^2 = Incorrected $\Sigma x^2 - (6^2 + 8^2) + (8^2 + 6^2)$

$= 650$

Corrected y^2 = Incorrected $\Sigma y^2 - (14^2 + 6^2) + (12^2 + 8^2)$

$= \underline{\underline{436}}$

Corrected Σxy = Incorrected $\Sigma xy - (6 \times 14 + 8 \times 6) + (8 \times 12 + 6 \times 8)$

$= \underline{\underline{520}}$.

$$r = \frac{n \Sigma xy - (\Sigma x) \cdot (\Sigma y)}{\sqrt{n \Sigma x^2 - (\Sigma x)^2} \times \sqrt{n \Sigma y^2 - (\Sigma y)^2}}$$

$$= \frac{25(520) - 125 \times 100}{\sqrt{16250 - 15625} \times \sqrt{10900 - 10000}}$$

25×30

$= \underline{\underline{0.6667}}$

Rank Correlation : Imp

$$r = 1 - \frac{6 \Sigma di^2}{n^3 - n}$$

1 Ten participant in a Contest are ranked by two Judges as follows

x :	1	6	5	10	3	2	4	9	7	8
y :	6	4	9	8	1	2	3	10	5	7

Calculate the rank Correlation Coefficient

Sol: $di = \Sigma xi - yi$

$di = -5, 2, -4, 2, 2, 0, 1, -1, 2, 1$

$\Sigma di^2 = 25 + 4 + 16 + 4 + 4 + 1 + 1 + 4 + 1$
$= 60$

$$r = 1 - \frac{6 \times 60}{10^3 - 10}$$

$$r = \frac{630}{990} = 0.636$$

① Three Judges A, B, C give the following rank, find which pair of Judges has common approach.

A : 1 6 5 10 3 2 4 9 7 8

B : 3 5 8 4 7 10 2 1 6 9

C : 6 4 9 8 1 2 3 10 5 9

$A = x$	Rank by $B = y$	Rank by $C = z$	d_1 $x-y$	d_2 $y-z$	d_3 $z-x$	d_1^2	d_2^2	d_2^2
1	3	6	−2	−3	5	4	9	25
6	5	4	1	1	−2	1	1	4
5	8	9	−3	−1	4	9	1	16
10	4	8	6	−4	−2	36	16	4
3	7	1	−4	6	−2	16	36	4
2	10	2	−8	8	0	64	64	0
4	2	3	2	−1	−1	4	1	1
9	1	10	8	−9	1	64	81	1
7	6	5	1	1	−2	1	1	4
8	9	7	−1	2	−1	1	04	1
						200	214	60

$$r = 1 - \frac{6\,[200 + 214 + 60]}{990}$$

$$= 1 - 2.8727$$

$$= -1.8727$$

For Comparision

$$\rho(x, y) = 1 - \frac{6 \sum d_1^2}{n^3 - n}$$

$$= 1 - \frac{6 \times 200}{990}$$

$$= -0.21$$

$$\rho(y, z) = 1 - \frac{6 \sum d_2^2}{n^3 - n}$$

$$= 1 - \frac{6 \times 214}{990}$$

$$= -0.29$$

$$\rho(z, x) = 1 - \frac{6 \sum d_3^2}{n^3 - n}$$

$$= 1 - \frac{6 \times 60}{990}$$

$$= 0.6$$

Que 2 The marks obtain by 9 Students in Chemistry and maths are given below:

Marks in Chemistry : 35 23 47 17 10 43 9 6 28

Marks in Maths : 30 33 45 23 8 49 12 4 31

Sol:

Chem (X)	Maths (Y)	rank R_x	rang R_y	d $R_x - R_y$	d^2
35	30	3	5	−2	4
23	33	5	3	2	4
47	45	1	2	−1	1
17	23	6	6	0	0
10	8	7	8	−1	1
43	49	2	1	1	1
9	12	8	7	1	1
6	4	9	9	0	0
28	31	4	4	0	0
					$\Sigma d^2 = 12$

Now $\Sigma d^2 = 12$

$$r = 1 - \frac{6 \times 12}{729 - 9} =$$

$= \underline{\underline{0.9}}$

$r =$

Date :

Que: 3 10 Student get the following percentage of marks in MATHS and PHYSICS.

Maths : 8 36 98 25 75 82 92 62 69 85

Physics : 84 51 91 60 68 62 86 58 35 49

Find rank Corelation.

Maths	Physics	Rx	Ry	d Rx - Ry	d^2
8	84	10	3	7	49
36	51	7	8	-1	1
98	91	1	1	0	0
25	60	9	6	3	9
75	68	4	4	0	0
82	62	3	5	-2	4
92	86	2	2	0	0
62	58	6	7	-1	1
69	35	5	10	-5	25
85	49	8	9	-1	1
					$\Sigma d^2 = 90$

$$r = 1 - \frac{6 \times 90}{10^3 - 10}$$

$$= 1 - \frac{540}{990}$$

$$= 0.454$$

Que: From the following table find the rank Correlation Coefficient

X : 48 33 40 9 16 16 65 24 16 57

Y : 13 13 24 6 15 4 20 9 6 19

Here in the Series X the value of

We can Say $m_1 = 3$

in Series Y the value of 6 and 13 are repeated to say $m_2 = 2$ and $m_3 = 2$

$\therefore$ the rank Correlation Coefficient is given by

$$r = 1 - \frac{6\left\{\sum d^2 + \frac{1}{12} m_1(m_1^2 - 1) + \frac{1}{12} m_2(m_2^2 - 1) + \dots\right\}}{n(n^2 - 1)}$$

Here $n = 10$, $m_1 = 3$ $m_2 = 2$, $m_3 = 2$

X	Y	Rank X in (RX)	Rank in Y (RY)	d = Rx - Ry	d2
48	13	8	5.5	2.5	6.25
33	13	6	5.5	0.5	0.25
40	24	7	10	-3.0	9.0
9	6	1	2.5	-1.5	2.25

Date :

16	4	3	1	2.0	4.0
65	20	10	9	1.0	1.0
24	9	5	4	1.0	1.0
16	6	3	2.5	0.5	0.25
57	19	9	8	1.0	1.0
Total				$\Sigma d = 0$	$\Sigma d^2 = 41$

$$r = 1 - \frac{6\left\{41 + \frac{1}{12}3(9-1) + \frac{1}{12}2(4-1)\right\}}{10(100-1)}$$

$$= \underline{\underline{0.733}}$$

Que: Find the rank-correlation Coefficient for the following data.

X : 68 64 75 50 64 80 75 40 55 64

Y : 62 58 68 45 81 60 68 48 50 70

X. $m_1 = 2$, $m_2 = 3$ $m_3 = 2$

$$r = 1 - \frac{6\left\{\Sigma d^2 + \frac{1}{12}m_1(m_1^2-1) + \frac{1}{12}m_2(m_2^2-1) + \ldots\right\}}{n(n^2-1)}$$

X	Y	Rx	Ry	$R_x - R_y$ d	d2
68	62	7	6	1	1
64	58	5	4	1	1
75	68	8.5	7.5	1	1
50	45	2	1	1	1
64	81	5	10	-5	25
80	60	10	5	5	25
75	68	8.5	7.5	1	1
40	48	1	2	-1	1
55	50	3	3	0	0
64	70	5	9	-4	16

$\sum d^2 = 72$

$$r = 1 - \frac{6\left\{72 + \frac{6}{12} + \frac{24}{12} + \frac{6}{12}\right\}}{10(99)}$$

$= 1 - 0.4545$

$= \underline{\underline{0.5455}}$

Date :

Assignment - 2

Que: 1 Find first four moments of Gamma distribution.

Sol: First moment about origin $= \int_0^\infty x \cdot f(x)\,dx$

$$= \int_0^\infty x \cdot \frac{e^{-x}}{\Gamma\lambda} x^{\lambda-1}\,dx$$

$$= \frac{1}{\Gamma\lambda}\int_0^\infty e^{-x} x^{\lambda}\,dx$$

$$= \frac{1}{\Gamma\lambda}\int_0^\infty e^{-x} x^{(\lambda+1)-1}\,dx$$

By defination of gamma function

$$= \frac{1}{\Gamma\lambda}\Gamma\lambda+1$$

$$= \frac{1}{\cancel{\Gamma\lambda}}\lambda\cancel{\Gamma\lambda} = \lambda$$

$$\boxed{\mu_1' = \lambda}$$

Second moment about origin $\mu_2' = \int_0^\infty x^2 f(x)\,dx$

$$= \int_0^\infty x^2 \cdot \frac{e^{-x}}{\Gamma\lambda} x^{\lambda-1}\,dx$$

$$= \frac{1}{\Gamma\lambda}\int_0^\infty e^{-x} x^{(\lambda+2)-1}\,dx$$

Date :

By defination of gamma function.

$$= \frac{1}{\Gamma\lambda} \Gamma\lambda+2$$

$$= \frac{1}{\cancel{\Gamma\lambda}} (\lambda+1)\lambda \cancel{\Gamma\lambda}$$

$$\boxed{\mu_2' = \lambda(\lambda+1)}$$

Third moment about origin. $\mu_3' = \int_0^\infty x^3 \cdot f(x) dx$

$$\mu_3' = \int_0^\infty x^3 \cdot f(x)\, dx$$

$$= \int_0^\infty x^3 \cdot \frac{e^{-x} x^{\lambda-1}}{\Gamma\lambda} dx$$

$$= \frac{1}{\Gamma\lambda} \int \cancel{x^3} \cdot e^{-x} \, x^{(\lambda+3)-1} dx.$$

By defination of gamma function

$$= \frac{1}{\Gamma\lambda} \Gamma\lambda+3$$

$$= \frac{1}{\cancel{\Gamma\lambda}} (\lambda+2)(\lambda+1)\lambda \cancel{\Gamma\lambda}$$

$$\boxed{\mu_3' = \lambda(\lambda+1)(\lambda+2)}$$

Page No. :
Date :

Fourth moment about origin $\mu_4' = \int_0^\infty x^4 \cdot f(x)$

$$\mu_4' = \int_0^\infty x^4 \cdot \frac{e^{-x} x^{\lambda-1}}{\Gamma\lambda} dx$$

$$= \frac{1}{\Gamma\lambda} \int_0^\infty e^{-x} x^{(\lambda+4)-1} dx$$

By defination of gamma function

$$= \frac{1}{\Gamma\lambda} \Gamma\lambda+4$$

$$= \frac{1}{\cancel{\Gamma\lambda}} (\lambda+3)(\lambda+2)(\lambda+1)(\lambda)\cancel{\Gamma\lambda}$$

$$\boxed{\mu_4' = \lambda(\lambda+1)(\lambda+2)(\lambda+3)}$$

Que: 3 In a Normal Distribution 7% of items are under 35 and 89% are under 63. determine mean & variance of Distribution

$\because$ 7% items are below 35, i.e $50-7 = 43\%$ items are between 35 & m &

$\because$ 89 items are below 63, $89-50 = 39$ items are between m & 63.

for area 0.43 ; $z = 1.48$.

$\because$ it is less than m

$z = -1.48$ and for area 0.39 $z = 1.23$

Page No. :
Date :

$\therefore \dfrac{(35 - m)}{\sigma} = -1.48$ —— (1)

$\dfrac{63 - m}{\sigma} = 1.23$ —— (2)

from (1) & (2)

$m = 50.3$ $\qquad \sigma = 10.33$

$\boxed{\text{Mean} = 50.3}$

Lines of Regression

(I) Y on x : $Y - \bar{Y} = r \cdot \frac{\sigma y}{\sigma x}(x - \bar{x})$

(II) x on Y : $x - \bar{x} = r \cdot \frac{\sigma x}{\sigma y}(y - \bar{y})$.

Que: If there is the angle between the two regression line. then Show that

$$\tan\theta = \frac{1 - r^2}{r} \cdot \frac{\sigma x \cdot \sigma y}{\sigma x^2 + \sigma y^2}$$

Sol: W.K.T eq of the line of regression

① Y on x : $Y - \bar{Y} = r \cdot \frac{\sigma y}{\sigma x}(x - \bar{x})$

② x on y : $x - \bar{x} = r \cdot \frac{\sigma x}{\sigma y}(y - \bar{y})$

Slopes are :

$m_1 = r \cdot \frac{\sigma y}{\sigma x}$; $m_2 = \cancel{r} \cdot \cancel{\frac{\sigma x}{\sigma y}} \; \frac{\sigma y}{r \cdot \sigma x}$

$$\tan\theta = \frac{m_2 - m_1}{1 + m_1 m_2}$$

$$= \frac{r\left[\frac{\sigma x}{\sigma y} - \frac{\sigma y}{\sigma x}\right]}{1 + r^2}$$

Page No. :
Date :

$$= \frac{r\left[\frac{\sigma x^2 - \sigma y^2}{\sigma x \sigma y}\right]}{1 + r^2}$$

$$= \frac{\frac{\sigma y}{r \cdot \sigma x} - \frac{r \sigma y}{\sigma x}}{1 + \frac{\sigma y^2}{\sigma x^2}}$$

$$= \frac{\frac{\sigma y \sigma x - r^2 \sigma x \sigma y}{\cancel{\sigma x}(r \cdot \cancel{\sigma x})}}{\frac{\sigma x^2 + \sigma y^2}{\cancel{\sigma x^2}}}$$

$$= \frac{\sigma x \sigma y (1 - r^2)}{r (\sigma x^2 + \sigma y^2)}.$$

→ When $r = 0$; $\tan\theta = \infty$ or $\theta = \frac{\pi}{2}$.

i.e when the variables are independent the two line of regression are perpendiculer to each other.

→ When $r = \pm 1$ $\tan\theta = 0$ i.e $\theta = 0$ or π.

Thus the two lines of regression coincide i.e there is perfect co-relation between the two variables.

Que: 1 In a partially destroyed record, only the line of regression of y on x and x on y are available as $4x - 5y + 33 = 0$, $20x - 9y = 107$ respectively; calculate $\bar{x}, \bar{y}$ and the coefficient of co-relation between x and y.

Sol: Since the regression line passes through $(\bar{x}, \bar{y})$

then,
$$4\bar{x} - 5\bar{y} + 33 = 0 \quad \text{—— (1)}$$
$$20\bar{x} - 9\bar{y} - 107 = 0 \quad \text{—— (2)}$$

$$\bar{x} = \frac{107 + 9\bar{y}}{20}$$

$$\frac{107 + 9\bar{y}}{5} - 5y + 33 = 0$$

$$107 + 9\bar{y} - 25\bar{y} + 165 = 0$$

$$272 - 16\bar{y} = 0$$

$$\boxed{\bar{y} = 17}$$

$$\boxed{\bar{x} = 13}$$

Rewriting the line of regression of y on x

$$Y = \frac{4}{5}x + \frac{33}{5} \quad \text{we get}$$

$$b(y\,x) = r \cdot \frac{\sigma y}{\sigma x} = \frac{4}{5} \quad \text{—— (3)}$$

Rewriting the line of regression of x on y ;

$$x = \frac{9}{20} y + \frac{107}{9}$$

$$b_{xy} = r \cdot \frac{\sigma x}{\sigma y} = \frac{9}{20} \quad -(4)$$

$$r^2 = \frac{36}{100} = 0.36$$

$$r = 0.6$$

Coefficient of Co-relation b/w x and y $= 0.6$

Que 2 Find the regression equation of y and x and the coefficient of Co-relation from the following data $\Sigma x = 60$; $\Sigma y = 40$; $\Sigma xy = 1150$; $\Sigma x^2 = 4160$; $\Sigma y^2 = 1720$ and $n = 10$.

Sol: we have $\bar{x} = \frac{\Sigma x}{n} = \frac{60}{10} = 6$

$$\bar{y} = \frac{\Sigma y}{n} = \frac{40}{10} = 4$$

$$b_{xy} = \frac{\Sigma xy - \frac{\Sigma x \cdot \Sigma y}{n}}{\Sigma y^2 - \frac{(\Sigma y)^2}{n}}$$

Date :

$$= \frac{1150 - \frac{60 \times 4\cancel{0}}{1\cancel{0}}}{1720 - \frac{160\cancel{0}}{1\cancel{0}}}$$

$$= \frac{910}{1560}$$

$$b_{xy} = \underline{\underline{0.5833}}$$

$$r = \sqrt{b_{xy} \times b_{yx}}$$

$$b_{yx} = \frac{\Sigma xy - \frac{\Sigma x \cdot \Sigma y}{n}}{\Sigma x^2 - \frac{\Sigma(x)^2}{n}}$$

$$= \frac{1150 - \frac{60 \times 4\cancel{0}}{1\cancel{0}}}{4160 - \frac{360\cancel{0}}{1\cancel{0}}}$$

$$= \frac{910}{3800}$$

$$b_{yx} = \underline{\underline{0.239}}$$

$$r = \sqrt{b_{xy} \times b_{yx}}$$

$$= \sqrt{0.5833 \times 0.239}$$

$$= \underline{\underline{0.3733}}$$

Also regression eqn of Y on x :-

$$Y - \overline{Y} = byx\,(x - \overline{x})$$

$$Y - 4 = 0.239(x - 6)$$

regression eqn of x on y.

$$x - \overline{x} = bxy\,(y - \overline{y})$$

$$x - 6 = 0.5833\,(Y - 4)$$

Que: If regression eqn of x on y : $5x - y = 22$ and Y on x : $64x - 45y = 24$ are two lines of regression Find:

1. The mean value of x and y.
2. The regression coefficient
3. The Coefficient of Correlation b/w x and y.
4. The S.D of y ; if the variance of x is 25.

Sol: Since the regression line passes through $(\overline{x}, \overline{y})$

$$5\overline{x} - \overline{y} = 22 \quad \text{--- (1)}$$

$$64\overline{x} - 45\overline{y} = 24 \quad \text{--- (2)}$$

from (1) & (2)

$$64\overline{x} - 45(5\overline{x} - 22) = 24$$

$$64\overline{x} - 225\overline{x} + 990 = 24$$

$$+161\overline{x} = \cancel{1014} + 966$$

$$\overline{x} = \cancel{-6.29}\ \ 6$$

Page No. :
Date :

$\bar{y} = \cancel{5x = 6.24} - 22$

$\bar{y} = 5\bar{x} - 22$

$\bar{y} = 30 - 22 = \underline{\underline{8}}$

Rewriting the equation y on x:

$y = \frac{64\bar{x}}{\cancel{8}45} - \frac{24}{45}$

$byx = r \cdot \frac{\sigma y}{\sigma x} = \frac{64}{45}$

Rewriting the equation x on y

$x = \frac{y}{5} + \frac{22}{5}$

$bxy = r \cdot \frac{\sigma x}{\sigma y} = \frac{1}{5}$

$r = \sqrt{\frac{64}{45} \times \frac{1}{5}} = \frac{8}{15}$

(ii) Since Variance ~~bxy~~ $\boxed{\sigma x = 25}$

$byx = r \cdot \frac{\sigma y}{\sigma x}$

$\frac{64}{45} = \frac{8}{15} \cdot \frac{\sigma y}{25}$

$\sigma y = \underline{\underline{66.667}}$

Page No. :
Date :

Que:4 If $2x + 3y = 7$ and $5x + 4y = 9$ are 2 lines of regression. Find.

1. The mean value of x and y
2. The regression Coefficients
3. The Coefficient of Correlation b/w x and y.

Sol: Since the lines of regression passes through $(\bar{x}, \bar{y})$

$$10x + 15y = 35$$
$$10x + 8y = 18$$
$$- \quad - \quad -$$
$$7y = 17$$

$$\bar{y} = 2.42$$

$$x = \frac{7 - 3(2.42)}{2}$$

$$\bar{x} = -\cancel{0.565} = -0.13$$

Rewriting the equation y on x

$$y = \frac{7}{3} - \frac{2x}{3}$$

$$b_{yx} = r \cdot \frac{\sigma y}{\sigma x} = -\frac{2}{3}$$

Date :

Rewriting the equation x on y:

$$x = \frac{9}{5} - \frac{4y}{5}$$

$$b_{xy} = r \cdot \frac{\sigma x}{\sigma y} = \frac{-4}{5}$$

$$r = \sqrt{\frac{8}{15}}$$

$$r = -\sqrt{\frac{8}{15}}$$

Que:5 Find the two lines of regression from the following data.

Age of husband:	25	22	28	26	35	20	22	40	20	18
Age of wife:	18	15	20	17	22	14	16	21	15	14

Hence estimate: (1) The age of husband; when the age of wife 19

(2) The age of wife; when the age of husband. 30

(3) The correlation coefficient between them.

Sol: Let x = Age of husband

Y = Age of wife

Let $u = x - 26$ & $v = y - 17$
where 26 and 17 are assumed means of x and y series respectively;
here $n = 10$

Mean of $x = \bar{x} = \frac{\Sigma x}{n} = \frac{256}{10} = 25.6$

Mean of $y = \bar{y} = \frac{\Sigma y}{n} = \frac{172}{10} = 17.2$

x	$u = x - 26$	u^2	y	$v = y - 17$	v^2	uv
25	-1	-1	18	1	1	-1
22	-4	16	15	-2	4	-16
28	2	4	26	3	9	18
26	0	0	17	0	0	0
35	9	81	22	5	25	45
20	-6	36	14	-3	9	18
22	-4	16	16	-1	1	4
40	14	196	21	4	16	56
20	-6	36	15	-2	4	12
18	-8	64	14	-3	9	24
Σx 256	-4	450	172	2	78	340 172

y on $x \Rightarrow b_{yx} = \dfrac{\Sigma uv - \dfrac{\Sigma u \cdot \Sigma v}{n}}{\left\{ \Sigma u^2 - \dfrac{(\Sigma u)^2}{n} \right\}}$

$$= \frac{\cancel{340}\ 172 - \frac{(-4)\times 2}{10}}{450 - \frac{16}{10}}$$

$$= \frac{\cancel{170.8}}{\cancel{448.4}} \quad \frac{172.8}{448.4}$$

$$byx = \cancel{0.760} \quad \underline{\underline{0.385}}$$

$$\Rightarrow x \text{ on } y = bxy = \frac{\Sigma uv - \frac{\Sigma u \cdot \Sigma v}{n}}{\Sigma v^2 - \frac{(\Sigma v)^2}{n}}$$

$$= \frac{172 + \frac{8}{10}}{78 - 0.4}$$

$$= \underline{\underline{2.226}}$$

⇒ Regression line of Y on x:

$$Y - \overline{Y} = byx(x - \overline{x})$$

$$Y - 17.2 = 0.385(x - 25.6)$$

$$0.385x - y = \cancel{27.1456} = 0$$

$$y = 0.385x + 7.34 \quad \text{—} \ 1$$

Regression line of x on y:

$$x - \bar{x} = bxy(y - \bar{y}$$

$$x - 25.6 = 2.226(y - 17.2)$$

$$x - 25.6 = 2.226y - 38.28$$

$$x = 2.226y - 12.68. \quad \text{--- (2)}$$

When age of wife is 19

$$x = 42.294 - 12.68$$

Age of husband $x = \underline{29.614} \simeq 30$

Age of wife when age of husband is 30

$$y = 0.385x + 7.34$$

$$= 0.385 \times 30 + 7.34$$

$$= 18.89$$

$$\simeq 19$$

Coefficient of Correlation $= \sqrt{byx \times bxy}$

$$= \sqrt{0.384 \times 2.23}$$

$$= \underline{0.925}$$

Qu: Calculate the Coefficient of regression line and find the two lines of regression from the following data.

x : 78 89 97 69 59 79 68 61

y : 125 137 156 112 107 136 123 108

$$\Sigma x = \frac{600}{8} = 75$$

$$\Sigma y = \frac{1004}{8} = 125.5$$

x	$u = x - \bar{x}$	u^2	y	$v = y - \bar{y}$	v^2	uv
78	9	81	125	13	169	117
89	20	400	137	25	625	500
97	28	784	156	44	1936	1232
69	0	0	112	0	0	0
59	-10	100	107	-5	25	50
79	10	100	136	24	576	240
68	-1	1	123	11	121	-11
61	-8	64	108	-4	16	32
Total 600	48	1530	1004	108	3468	2160

y on x

$$b_{yx} = \frac{\Sigma uv - \frac{\Sigma u \cdot \Sigma v}{n}}{\left\{\Sigma u^2 - \frac{(\Sigma u)^2}{n}\right\}}$$

Date :

$$= \frac{2160 - \frac{48 \times 108}{8}}{1530 - \frac{2304}{8}}$$

$$= \frac{1512}{1242} = \underline{\underline{1.2173}}$$

$$\boxed{byx = 1.2173}$$

x on y $bxy = \dfrac{\Sigma uv - \frac{\Sigma u . \Sigma v}{n}}{\left\{\Sigma v^2 - \frac{(\Sigma v)^2}{n}\right\}}$

$$= \frac{2160 - \frac{48 \times 108}{8}}{3468 - \frac{11664}{8}}$$

$$= \frac{1512}{2010} = \underline{\underline{0.7522}}$$

$$r = \sqrt{bxy \times byx}$$

$$= \sqrt{0.7522 \times 1.2173}$$

$$= \underline{\underline{0.956}}$$

Date :

Ques. Calculation check — The following data represent rainfall x and yield of paddy per hectare y in a particular area. Find the line of regression of x on y

x :	113	102	95	120	140	130	125
y :	1.8	1.5	1.3	1.9	1.1	2.0	1.7

x	$u = x - \bar{x}$	u^2	y	$v = y - \bar{y}$	v^2	uv
113	−7	49	1.8	−0.1	0.01	0.7
102	−18	324	1.5	−0.4	0.16	7.2
95	−25	625	1.3	−0.6	0.36	9
100	0	0	1.9	0	0	0
140	20	400	1.1	−0.8	0.64	−16
130	10	100	2.0	0.1	0.01	1
125	5	25	1.7	−0.2	0.04	−1
Total 825	−15	1523	11.3	−2	1.22	0.9

y on x $\quad b_{yx} = \dfrac{\Sigma uv - \dfrac{\Sigma u \cdot \Sigma v}{n}}{\Sigma u^2 - \dfrac{(\Sigma u)^2}{n}}$

$$= \frac{0.9 - \dfrac{15 \times 2}{7}}{1523 - \dfrac{225}{7}}$$

$$= \frac{-3.385}{1490.85}$$

Page No. :
Date :

$= -2.270 \times 10^{-3}$

$b_{yx} = -0.002270$

~~b_{xy}~~ = x on y $\quad b_{xy} = \dfrac{\Sigma uv - \dfrac{\Sigma u \Sigma v}{n}}{\Sigma u^2 - \dfrac{(\Sigma u)^2}{n}}$

$= \dfrac{0.9 - \dfrac{30}{7}}{1.22 - \dfrac{4}{7}} = \dfrac{-3.385}{0.6485}$

$b_{xy} = -5.2197$

$r = \sqrt{-0.002270 \times -5.2197}$

19.68

$r = -0.10885$

CHI - SQUARE (χ^2) Test

Date :

Defination: If $O_1, O_2, \ldots\ldots, O_n$ be a set of observed frequencies and $E_1, E_2, \ldots\ldots, E_n$ be the corresponding set of expected frequencies then CHI-SQUARE is defined by the relation

$$\chi^2 = \frac{(O_1 - E_1)^2}{E_1} + \frac{(O_2 - E_2)^2}{E_2} + \ldots\ldots + \frac{(O_n - E_n)^2}{E_n}.$$

with $(n-1)$ degree of freedom

$\{ \Sigma O_i = \Sigma E_i = n$ (The Total frequencies).

CHI-SQUARE Distribution :

If $x_1, x_2 \ldots\ldots x_n$ be n independent normal variate betwee with mean 0 & std deviation unity. Then $x_1^2 + x_2^2 + \ldots + x_n^2$ is a random variate having CHI-Square distribution

i.e $Y = Y_0 \cdot e^{-(x)^2/2} \{\chi^2\}^{(v-1)/2}$

where $V = n-1$

χ^2 = CHI-SQUARE.

f χ^2 >, so it is rejected

ese: accpt.

Que: 1 In Experiment's on pea breeding the following frequency of seeds were obtained

Round & Yellow	wrinkled & yellow	Round & green	wrinkled & green	Total
315	101	108	32	556

Theory Predict's that the frequencies should be in preparations: 9:3:3:1. Examine the corresponding between theory and Experiments.

Sol: The corresponding frequencies are

$\frac{9}{16} \times 556 = 312.75$

$\frac{3}{16} \times 556 = 104.25$

$\frac{3}{16} \times 556 = 104.25$

$\frac{1}{16} \times 556 = 34.75$

$\chi^2 = \frac{(315 - 313)^2}{313} + \frac{(101 - 104)^2}{104} + \frac{(108-104)^2}{104}$

$+ \frac{(32-35)^2}{35}$

Date :

$$x^2 = \frac{4}{313} + \frac{9}{104} + \frac{16}{104} + \frac{9}{35}$$

$$= 0.012 + 0.086 + 0.15 + 0.25$$

$$= 0.51 \text{ (approx)}$$

Degree of freedom $= V - 1$ $V \rightarrow$ Total No of entries

$= 4 - 1$

$= 3$

$x^2_{0.05} = 7.818$

at 3

Hence, the Calculate value of CHI-SQUARE is much less than CHI-SQUARE 0.5

There is a very high degree of Agreement b/w ~~theorey~~ theory & experiment.

Que:2 A Set of five similar Coins is tossed 320 times and the result is

No of heads :	0	1	2	3	4	5
frequency :	6	27	72	112	71	32

Test the hypothesis that the data follow a binomial distribution.

Date :

~~Que.~~
Sol: For $v = 5$ we have $x^2_{0.05} = 11.07$

p: Probability of getting head $= \frac{1}{2}$

q: Probability of getting tail $= \frac{1}{2}$

hence the Theoretical frequency of getting 0, 1, 2, 3, 4, 5 heads are the successive terms of the binomial Expansion ~~320~~

$320(P+q)^n$

$= 320\{P^5 + 5P^4q + 10P^3.q^2 + 10P^2q^3 + 5Pq^4 + q^5\}$

$= 320\left\{\frac{1}{32} + \frac{5}{32} + \frac{10}{32} + \frac{10}{32} + \frac{5}{32} + \frac{1}{32}\right\}$

$= 10 + 50 + 100 + 100 + 50 + 10$

$=$

Thus the Theoritical frequencies are

10, 50, 100, 100, 50, 10

Hence

$X^2 = \frac{(6-10)^2}{10} + \frac{(27-50)^2}{50} + \frac{(72-100)^2}{100} +$

$\frac{(112-100)^2}{100} + \frac{(71-50)^2}{50} + \frac{(32-10)^2}{10}$

$=$

Page No. :
Date :

$= \frac{16}{10} + \frac{529}{50} + \frac{784}{100} + \frac{144}{100} + \frac{441}{50} + \frac{484}{10}$

$= 1.6 + 10.58 + 7.84 + 1.44 + 8.82 + 48.4$

$= \underline{\underline{78.68}}$

Que:3 Fit a Poisson distribution to the following data and test for its goodness of fit at least of significant 0.05.

x :	0	1	2	3	4
f :	419	352	154	58	19

$P(x) = \frac{1002 \times e^{-m} \cdot m^{x}}{\lfloor x}$

$\text{mean} = \frac{\sum f(x)}{f} = \frac{352 + 308 + 174 + 76}{1002}$

$= \frac{910}{1002} = 0.9$

$P(0) = \frac{1002 \times e^{-0.9} \cdot 0.9^{0}}{\lfloor 0} = 407.38$

$P(1) = \frac{1002 \times e^{-0.9} \cdot (0.9)^{1}}{1} = 366.64$

$P(2) = \frac{1002 \times e^{-0.9} (0.9)^{2}}{\lfloor 2} = 164.99$

$$P(3) = \frac{1002 \times e^{-0.9} \cdot (0.9)^3}{\underline{|3}} = 49.49.$$

$$P(4) = \frac{1002 \times e^{-0.9} (0.9)^4}{\underline{|4}} = 11.13.$$

Theoretical frequencies

407.38, 366.64, 164.99, 49.49, 11.13

Hence

$$\chi 2 = \frac{(419-407.38)^2}{407.38} + \frac{(352-366.64)^2}{366.64} + \frac{(19-11.13)^2}{11.13} + \frac{(154-164.99)^2}{164.99} + \frac{(58-49.49)^2}{49.49}$$

$$\chi 2 = 0.331 + 0.584 + 0.732 + 1.463 + 5.564$$

$$\chi 2 = \underline{8.674}$$

Que:4 In a Sample Survey of Public openion Answers to the question.

1. Do you drink
2. Are you in favour of local option on Sale of Liquor are tabulated below

	Yes	No	Total
Yes	56	31	87
No	18	6	24
Total	74	37	111

Can you peform proper whether or not the local option on the Sale of liquor is dependent on indusable drink, Given that the value of χ^2 at 5% level of Significat is 3.841

Sol: (I) Null hypothesis H_0 – The option on the Sale of liquor is not dependent with the indusable drinking.

(II) Calculation of expected / Theoretical frequence f_e

The expected frequencies Corresponds to the Theoretical frequencies are calculated as follow:

Cell 11

$$f_{e11} = \frac{87 \times 74}{111} = 58$$

Cell 12 is –

$$fe_{12} = \frac{87 \times 37}{111} = 29$$

Cell 21 is

$$fe_{21} = \frac{74 \times 24}{111} = 16$$

Cell 22

$$fe_{22} = \frac{24 \times 37}{111} = 8.$$

III) Calculation of χ^2 Statistic :

$$\chi^2 = \Sigma \frac{(fo - fe)^2}{fe}$$

fo	fe	$(fo-fe)^2/fe = \chi^2$
56	58	0.068
31	29	0.137
18	16	0.250
6	8	0.5
		$\Sigma\chi^2 = 0.955$

also d.o.f = deg(v) = (m–1)(n–1)
= (2–1)(2–1)
= 1

m → no. of row.
n → no of col.

(IV) The tabulated value of χ^2 at 5% label of significant and 1 degree of freedom.

$$\chi^2_{0.05;1} = 3.841$$

(V) Decision :
Clearly calculated value of $\chi^2 = 0.955$ is less than tabulated value i.e 3.841

The Null hypothesis is ~~expected~~. accepted.

Liquor is not dependent with the indusible drinking

Que:5 50 Students selected at random from 500 Students Enrolled in a Computer program were classified according to age and grade point. given the following data :

grade Points	Age (in years) 20 & under	21-30	above 30	Total
up to 5.0	3	5	2	10
5.1 to 7.5	8	~~7~~ 8	5	20
7.6 to 10	4	8	8	20
Total	15	20	15	N = 50

F - distribution

Let $x_1, x_2 \ldots \ldots x_n$ and $y_1, y_2 \ldots \ldots y_n$ be the values of 2 independent random sample drawn from the random & normal population σ^2 & having equal variance

Let $\bar{x}_1$, and $\bar{x}_2$ be the sample means and

$$S_1^2 = \frac{1}{n_1 - 1} \sum_{i=1}^{n_1} (x_i - \bar{x})^2$$

$$S_2^2 = \frac{1}{n_2 - 1} \sum_{i=1}^{n_2} (y_i - \bar{y})^2$$

be the sample variance then we define F

$$\boxed{F = \frac{S_1^2}{S_2^2}} \quad ; \quad S_1^2 > S_2^2.$$

with $\left.\begin{array}{l} \mu_1 = n_1 - 1 \\ \mu_2 = n_2 - 1 \end{array}\right\}$ degree of freedom

Que: Random sample are drawn from two population and the following result

Sample X :	20	16	26	27	23	22	18	24	25	19	
Sample Y :	27	33	42	35	32	34	38	28	41	43	
							39		37		

Find the Variance of two binomial and these Whether the two Sample has Same Variance
(given that $F_{0.05}$ for 11 and 9 d.o.f is 3.112.

Sol: Given that $n_1 = 10$ and $n_2 = 12$ with d.o.f $V = n_1 - 1 = 9$ and $V = n_2 - 1 = 11$

Step 1 : Null hypothesis H_0 : Let $\sigma_1^2 = \sigma_2^2$
i.e the two Samples have the Same Varience.

Step 2 : Calculation of F - Statistic

$\bar{x} = 22$
$\bar{y} = 35.75$

we have to find s_1^2 and s_2^2

Sample X			Sample Y		
X	$x - \bar{x}$	$(x - \bar{x})^2$	Y	$Y - \bar{Y}$	$(Y - \bar{Y})^2$
20	−2	4	27	−8.75	76.56
16	−6	36	33	−2.75	7.56
26	4	16	42	6.25	39.06
27	5	25	35	−0.75	0.5625
23	1	1	32	−3.75	14.062
22	0	0	34	−1.75	3.0625
18	−4	16	38	2.25	5.0625
24	2	4	28	−7.75	60.0625
25	3	9	41	5.25	27.5625 ~~11.819~~
19	−3	9	43	7.25	52.5625
			39	3.3	10.89
$\Sigma = 220$	$\Sigma = 0$	$\Sigma = 120$	37	1.3	1.69

$$\Sigma(y-\bar{y})^2 = 298.69$$

$$S_1^2 = \frac{120}{9} = 13.33$$

$$S_2^2 = \frac{298.69}{11} = 27.15$$

$$\Rightarrow S_2^2 > S_1^2$$

$$F = \frac{S_2^2}{S_1^2} = \frac{27.15}{13.33} = 2.03$$

Hence calculated value of F is 2.03

Step 3: The tabulated value of F at 5% level of significance for the degree of freedom at 11 & 9 is 3.112 i.e.

$$F_{0.05} = 3.112$$

Step 4: Decision

Calculated value of F is less than tabulated value of F

$$F_{cal} < F_{Tab.}$$

The Null hypothesis H_0 is accepted.

Date :

Que: **Student's t-distribution**

Ans 1 Consider a small sample of size n drawn from a normal population with mean μ and S.D σ
if $\bar{x}$ and σ_s be the sample mean and std deviation

Then the statistic "t" is defined as

$$\boxed{t = \frac{\bar{x} - \mu}{\sigma} \cdot \sqrt{n}}$$ Or

$$\boxed{t = \frac{\bar{x} - \mu}{\sigma_s} \sqrt{n-1}}$$

where $\nu = n-1$ is degree of freedom

The 9 item of a sample have the following value

45, 47, 50, 52, 48, 47, 49, 53, 51

does the mean of these differ significance from the assumened mean 47.5.

We will find the mean and S.D of sample as follows:

x	$d = x - 48$	d^2
45	-3	9
47	-1	1
50	2	4
52	4	16
48	0	0
47	-1	1
49	1	1
53	5	25
51	3	9
$\Sigma x = 442$	$\Sigma d = 10$	$\Sigma d^2 = 66$

$$\bar{x} = \text{mean} = 48 + \frac{\Sigma d}{9}$$

$$= 48 + \frac{10}{9}$$

$$= 49.11$$

$$\sigma_s^2 = \frac{\Sigma d^2}{9} - \left(\frac{\Sigma d}{9}\right)^2$$

$$= \frac{66}{9} - 1.23$$

$$= 7.33 - 1.23$$

$$\sigma_s^2 = 6.10$$

$$\sigma_s = 2.469$$

Date :

$$t = \frac{\bar{x} - u}{\sigma_s} \sqrt{n-1}$$

$$= \frac{49.11 - 47.5}{2.469} \times \sqrt{8}$$

$$= 0.652 \times \sqrt{8}$$

$$= \underline{\underline{1.844}}$$

degree of freedom $= n-1 = 9-1 = \underline{\underline{8}}$

$t_{0.05} = 2.31$

Calculated value of t is less than $t_{0.05}$
i.e $t < t_{0.05}$

The value of t is not significant.

Qu: The average number of articles produced by two machines per day are 200 & 250 with std deviation 20 & 25 respectively on the basis of record of 25 days production. Can you regard both the machines equally efficient at 1% level of significance.

Given $t_{0.01;48} = 2.58$

Sol: Given that $n_1 = 25$, $\bar{x}_1 = 200$, $S.d.\ \sigma_1 = 20$

$n_2 = 25$; $\bar{x}_2 = 250$; $S.D\ \sigma_2 = 25$

$$S^2 = \frac{n_1\sigma_1^2 + n_2\sigma_2^2}{n_1+n_2-2}$$

$$= \frac{25 \times 400 + 25 \times 625}{48}$$

$$= \underline{\underline{533.85}}$$

$$S = \underline{\underline{23.105}}$$

Degree of freedom $(v) = n_1 + n_2 - 2 = 48$.

Step I Null hypothesis H_0 - Both the machines are equally efficient.

$$H_0 : \mu_1 = \mu_2$$

Step II Calculation of 't' Statistics

$$|t| = \frac{\bar{x}_1 - \bar{x}_2}{S} \cdot \sqrt{\frac{n_1 \cdot n_2}{n_1+n_2}}$$

$$|t| = \frac{200-250}{23.105} \sqrt{\frac{25\times 25}{50}} = \cancel{189.77 \times 1.1767}$$

$$|t| = \underline{\underline{7.65}}$$

Fisher Z-distribution.

$Z = \frac{1}{2}\log_e F$ in the F-dis

i.e $\log_e F = 2Z$

$F = e^{2Z}$ in the F-dis we get Fisher-Z distribution

Test whether the two sets of observation

I : 17 27 18 25 27 29 27 23 17
II : 16 16 20 16 20 17 15 21

Indicate sample drawn from the same universe (the value of Z at 5% level for 8 and 7 degree of freedom is 0.6575)

Sol: Given $n_1 = 9$; $n_2 = 8$, $\bar{x} = 23.33$

with dof $\nu_1 = n_1 - 1 = 8$ $\bar{Y} = 17.62$
$\nu_2 = n_2 - 1 = 7$

First observation / Second observation

x	$x-\bar{x}$	$(x-\bar{x})^2$	Y	$(Y-\bar{Y})$	$(Y-\bar{Y})^2$
17	-6.33	40.06	16	-1.62	2.629
27	3.67	13.46	16	-1.62	2.629
18	-5.33	28.40	20	2.38	5.664
25	1.67	2.78	16	-1.62	2.629
27	3.67	13.46	20	2.38	5.664
29	5.67	32.148	17	0.62	0.384

27	3.67	13.46	15	-2.62	6.86
23	-0.33	0.108	21	3.38	11.42
17	-6.33	40.068			
$\Sigma x =$		183.94.			37.879

Calculation of Z - Statistics

we have

$$S_1^2 = \frac{\Sigma (x-\bar{x})^2}{n_1 - 1} = \frac{183.94}{8}$$

$$= 22.9925$$

$$\cancel{S_1} = \cancel{4.79}$$

$$S_2^2 = \cancel{37.} \frac{\Sigma (y-\bar{y})^2}{n_2 - 1}$$

$$= 5.41$$

$$S_1^2 > S_2^2$$

$$Z = \frac{1}{2} \log_e \left(\frac{22.99}{5.41} \right)$$

$$= \frac{1.4468}{2} = 0.7234$$

The tabulated of Z at 5% level of significance with dof 8 & 7 is 0.6525.

$Z_{0.05}$ ~~te~~ 0.6525

$\therefore$ Calculated value of Z is 0.7236

i.e $Z_{calculated} = 0.7236$

$0.7236 \nless$ tabulated 0.6575

$\therefore$ The null hypothesis is rejected.

The two variance are not same.

Curve Fitting

Method of Least Square.

Que: If P is the Pull required to lift a load W, by means of pulley block find a linear law of the form

$$P = mW + C \quad \text{—— (I)}$$

Connecting P & W using the following data.

P : 12 15 21 25

W : 50 70 100 120

Sol: The Normal eq^n of (I) are

$$\Sigma P = m\Sigma W + \Sigma C$$

i.e $\Sigma P = m\Sigma W + 4C$ —— (1)

$\Sigma PW = m\Sigma W^2 + C\,\Sigma W$ —— (2)

P	W	PW	W^2
12	50	600	2500
15	70	1050	4900
21	100	2100	10000
25	120	3000	14,400
73	340	6750	31800

eq (1)

$73 = m(340) + 4c$ —— (iii)

$73 = 340m + 4c$

eq (2)

$\cancel{65}$

$6750 = 31800m + 340c.$

$m = \frac{109}{580}$ $c = \frac{66}{29}$

$m = \underline{\underline{0.187}}$

$c = \underline{\underline{2.27}}$

$P = 0.1870\,w + 2.2785$

when $w = 150$ kg

$P = 28.05 + 2.2785$

$= \underline{\underline{30.3285}}$

Page No. :
Date :

Que: Fit a second degree parabola to the following data

x:	0	1	2	3	4
y:	1	1.8	1.3	2.5	6.3

$$y = a + bx + cx^2 \quad \text{——— (I)}$$

Normal eqn of (I) are

$$\Sigma y = \Sigma a + b\Sigma x + c\Sigma x^2 \quad \text{——— (1)}$$
$$\Sigma xy = a\Sigma x + b\Sigma x^2 + c\Sigma x^3 \quad \text{——— (2)}$$
$$\Sigma x^2 y = a\Sigma x^2 + b\Sigma x^3 + c\Sigma x^4 \quad \text{——— (3)}$$

	x	y	x^2	xy	x^3	x^2y	x^4
	0	1	0	0	0	0	0
	1	1.8	1	1.8	1	1.8	1
	2	1.3	4	2.6	8	5.2	16
	3	2.5	9	7.5	27	22.5	81
	4	6.3	16	25.2	64	100.8	256
$\Sigma =$	10	12.9	30	37.1	100	130.3	354

Now substitute the following in eq (1), (2), (3)

$$12.9 = 5a + 10b + 30c$$
$$37.1 = 10a + 30b + 100c$$
$$130.3 = 30a + 100b + 354c$$

$$a = \frac{71}{50} \qquad b = \frac{-107}{100}, \quad c = \frac{11}{20}$$

$$a = 1.42$$
$$b = -1.07$$
$$c = 0.55$$

$$y = 1.42 - 1.07x + 0.55x^2$$

$$\boxed{y = 1.42 - 1.07x + 0.55x^2}$$

Que: Fit a Second degree parabola to the following data.

x :	1.0	1.5	2.0	2.5	3.0	3.5	4.0
y :	1.1	1.3	1.6	2.0	2.7	3.4	4.1

$$y = a + bx + cx^2 \quad \text{--- (I)}$$

Normal equation of (I) are.

$$\Sigma y = \Sigma a + b\Sigma x + c\Sigma x^2 \quad \text{--- (1)}$$
$$\Sigma xy = a\Sigma x + b\Sigma x^2 + c\Sigma x^3 \quad \text{--- (2)}$$
$$\Sigma x^2y = a\Sigma x^2 + b\Sigma x^3 + c\Sigma x^4 \quad \text{--- (3)}$$

from Table,

$$16.2 = 7a + 17.5b + 50.75c \quad \text{--- (4)}$$
$$47.65 = 17.5a + 50.75b + 157.966c \quad \text{--- (5)}$$
$$154.475 = 50.75a + 157.966b + 538.78c \quad \text{--- (6)}$$

Page No. :
Date :

x	y	x^2	xy	x^3	x^2y	x^4
1.0	1.1	1	1.1	1	1.1	1
1.5	1.3	2.25	1.95	3.375	2.925	5.06
2.0	1.6	4.0	3.2	4.096	6.4	6.55
2.5	2.0	6.25	5	15.625	12.5	39.06
3.0	2.7	9	8.1	27	24.3	81
3.5	3.4	12.25	11.9	42.87	41.65	150.06
4.0	4.1	16	16.4	64	65.6	256
17.5	16.2	50.75	47.65	157.966	154.475	538.73

$a = 0.381$
$b = 0.305$
$c = 0.16$

Substitute a b c in equation I.

$$\boxed{y = 0.381 + 0.305x + 0.16x^2}$$

Fitting of other curve's

Que: A $Y = ax^b$

$\log_{10} Y = \log_{10} a + b \log_{10} x$

i.e. $Y = A + bX$ where $X = \log_{10} x$, $Y = \log_{10} Y$

$\Sigma Y = nA + b\Sigma X$ — (1)

$\Sigma XY = A\Sigma X + b\Sigma X^2$ — (2)

Que:

V :	350	400	500	600
t :	61	26	7	26

Fitted the curve $V = at^b$.

(B) $y = a \cdot e^{bx}$

$\log_{10} Y = \log_{10} a + bx \cdot \log_{10} e$

$Y = A + Bx$ where $Y = \log_{10} Y$

$\Sigma Y = nA + B\Sigma x$ — (1) $A = \log_{10} a$

$\Sigma xy = A\Sigma x + B\Sigma x^2$ — (2) $B = b \cdot \log_{10} e$

Que:

V :	350	400	500	600
t :	61	26	7	26

Curve = $V = at^b$

1: $\log_{10} V = \log_{10} a + b \log_{10} t$

Page No. :
Date :

Let $T = \log_{10} t$ & $A = \log_{10} a$

$V = \log_{10} v$

$\therefore V = A + bT$ — I

$\Sigma V = nA + b\Sigma T$ — 1

$\Sigma VT = A\Sigma T + b\Sigma T^2$ — 2

	V	T	T^2	VT
	350	61	3721	21350
	400	26	676	10400
	500	7	49	3500
	600	26	676	15600
Total	1850	120	5122	50850

Substituting values

$1850 = 4A + 120b$ — 3

$50850 = 120A + 5122b.$ — 4

$A = 554.15$

$b = -3.055$

$V = 554.15 - 3.055T$

$$\boxed{V = 554.15 - 3.055T}$$

Bivariate Distribution

Let x_1, x_2 be a random variable with the real plane as the Sample Space i.e x_1 and x_2 can be take real values from $-\infty$ to ∞ then x_1 x_2 has a Continuous probability distribution. If there is a density function $f(x_1, x_2)$ such that

$$Pr\{a_1 < x_1 < a_2 ; b_1 < x_1 < b_2\}$$

$$= \int_{b_1}^{b_2} \int_{a_1}^{a_2} f(x_1, x_2)\, dx_2 \cdot dx_1$$

with $f(x_1, x_2) \geq 0$

and $$\int_{-\infty}^{\infty} \int_{-\infty}^{\infty} f(x_1, x_2)\, dx_1\, dx_2 = 1$$

also. $$\frac{\partial}{\partial x_2} \cdot \frac{\partial}{\partial x_1} f(x_1, x_2) = f(x_1, x_2)$$

and $$f(-\infty, \infty) = 1$$

Page No. :
Date :

Conditional Probability Density:

The Conditional probability Density of x_1 gives that x_2 assume the value of.
$$f(x_1/x_2) = \frac{f(x_1, x_2)}{f_2(x_2)} \quad \text{if } f_2(x_2) \neq 0$$

Ex: given the Joint density

$$f(x_1, x_2) = \begin{cases} \frac{k}{(1+x_1+x_2)^3}, & \text{for } x_1>0, x_2>0 \\ 0, & \text{Otherwise} \end{cases}$$

Que: Find K and the marginal density of x_1 & x_2, find the Conditional density of x_1 given that x_2 assume the value of x_2. Check wether the random variable are independent

Sol: W.K.T.
$$\int_0^\infty \int_0^\infty \frac{k}{(1+x_1+x_2)^3} \, dx_1 . dx_2 = 1$$

$$k \int_0^\infty \int_0^\infty (1+x_1+x_2)^{-3} \, dx_1 \, dx_2$$

$$\frac{k}{-2} \int_0^\infty \frac{(1+x_1+x_2)^{-2+}}{1+\cancel{x}_0} \, dx_2$$

$$\cancel{\theta}\frac{\cancel{k}}{\cancel{16}} \int_0^\infty \frac{k}{2y^2+4y+2} = \frac{k}{2} = 1$$

$$\boxed{k = 2}$$

Page No. :
Date :

Then the marginal density of x_1 is

$$f_1(x_1) = \int_0^\infty f(x_1, x_2)\, dx_2$$

$$= \int_0^\infty \frac{k}{(1+x_1+x_2)^3}\, dx_2$$

$$= \int_0^\infty 2(1+x_1+x_2)^{-3}\, dx_2$$

$$= \frac{2}{-2}\left[(1+x_1+x_2)^{-2}\right]$$

$$= \int_0^\infty -(1+x_1+x_2)^{-2}$$

$$= (1+x_1)^{-2} = \frac{1}{(1+x_1)^2}$$

for $x_1 > 0$

$$f_2(x_2) = \int_0^\infty f(x_1, x_2)\, dx_1$$

$$= \frac{1}{(1+x_2)^2} \quad ; \quad x_2 > 0$$

The Conditional density of x_1 gives x_2 assume

The value x_2 is $f(x_1/x_2) = \dfrac{f(x_1, x_2)}{f_2(x_2)}$

Page No. :
Date :

$$= \frac{\frac{2}{(1+x_1+x_2)^3}}{\frac{1}{(1+x_2)^2}}$$

$$= \frac{2(1+x_2)^2}{(1+x_1+x_2)^3}.$$

Test of Significance for large Sample:

We know that the binomial distribution tends to normal for large n. Suppose we wish to test the hypothesis that the probability of success in such trial is p. Suppose it to be true then mean (μ), and standard deviation (σ) of the sampling distribution of no. of success are

$$\mu = np$$
$$\sigma = \sqrt{npq}.$$

respectively.

For a normal distribution only 5% of the number lie outside $\mu \pm 1.96\sigma$ which only 1% of the numbers lie outside 2.58σ.

If x be the observed number of success in the sample z is the std no. curve.

$$\boxed{z = \frac{x-\mu}{\sigma}}$$

Date :

Thus, we have the following test of Significance

1. if $|Z| < 1.96$
difference between the observed and expected number of success is not significant.

2. if $|Z| > 1.96$
difference is significant at 5% level of significance.

3. if $|Z| > 2.58$ / (3)
difference is significant at 1% level of significance.

Que: A coin was tossed 400 times and the head turned up 216 times. Test the hypothesis that the coin is unbaised at 5% level of significance

Sol: Suppose the coin is unbiased

Then the probability of getting the head at toss $= \frac{1}{2}$

Expected No of Success $= \frac{1}{2} \times 400 = 200$
(i.e calculated).

and the observed value of Success $= 216$

Page No. :
Date :

Thus the excess of observed value over expected value

$$x - np = 216 - 200 = 16$$

$$\boxed{x - np = 16}$$

also S.D of Simple Sampling $= \sqrt{npq}$

$$= \sqrt{400 \times \frac{1}{2} \times \frac{1}{2}} = 10$$

$$Z = \frac{x - np}{\sigma} = \frac{16}{10} = 1.6$$

→ The null hypothesis is not Significant

Que: A die was thrown 9000 times and a throw of 5 or 6 was obtained 3240 times on the assumption of random throwing. Do the data indicate a unbaised die.

Suppose the die is unbaised

Then the Prob of Success $= \frac{2}{6} = \frac{1}{3}$
(~~Calculated~~)

Expected Success (calculated) $= \frac{1}{3} \times 9000$

$= 3000$

$1 - \frac{1}{3}$

Page No. :
Date :

Observed = 3240

Thus the excess of observed value over expected value

$x - np = 3240 - 3000$
$= 240$

also S.D of Sampling $= \sqrt{npq}$

$= \sqrt{9000 \times \frac{2}{6} \times \frac{4}{6}}$

$= \sqrt{2000}$

$= 44.72$

$Z = \frac{x - np}{\sigma} = \frac{240}{44.72}$

$Z = 5.36$

Thus, Null hypothesis is Significant at 1 % level of Significance.

$\frac{16 \times 15 \times 14}{3 \times 2 \times 1}$

Conditional Probability.

Que: 1 Find the Probability of a throwing a sum of 7 in a single throw with two dice.

Sol:

$$n(S) = 36$$

$$n(E) = \{(1,6), (2,5), (3,4), (4,3), (5,2), (6,1)\}$$

$$n(E) = 6$$

$$P(E) = \frac{n(E)}{n(S)} = \frac{1}{6} \;//$$

Que: 2 From a bag containing 5 white, 7 red, and 4 black balls a man draw 3 at random. find the probability of being all white

$$n(S) = 5 + 7 + 4 = 16$$

$$n(E) = 5$$

$$P(E) = {}^{16}C_3 \times \left(\frac{5}{16}\right)^3 \times \frac{7}{16}^0 \times \frac{4}{16}^0$$

$$P(\text{success}) = \frac{5}{16} \qquad = \frac{\cancel{16} \times \cancel{15}^{5} \times \cancel{14}^{7}}{\cancel{3} \times \cancel{2} \times 1} \times \frac{125}{\cancel{16} \times 16 \times 16}$$

$$P(\text{failure}) = \frac{11}{16} = \frac{35 \times 125}{16 \times 16} \qquad {}^{16}C_3 \times \left(\frac{5}{16}\right)^3 \times \left(\frac{15}{16}\right)^{13}$$

$$= \frac{\cancel{16} \times \cancel{15}^{5} \times \cancel{14}^{7}}{\cancel{3} \times \cancel{2} \times 1} \times 0.489$$

$$= 17.0$$

$$= 35 \times 0.489 \times 0.007$$

$$= 0.0131$$

Date :

Sol: Total No of balls = 5+7+4 = 16

The total No of ways in which 3 balls can be drawn = $^{16}C_3$ = 560

The Sample Space $n(S)$ = 560.

Let E be the event of all three balls are white = $5C_3$ = 10 = $n(E)$ = 100

$$P(E) = \frac{n(E)}{n(S)} = \frac{10}{560} = 0.017$$

Que: 3 from a pack of 52 cards two cards are drawn at random. Find the probability of following events:

① both cards are Spade.
② One card is of Spade and one card of diamond.

Total Cards = 52

Total No. of ways in which 2 cards can be drawn = $52C_2$ = $\frac{52 \times 51}{2 \times 1}$ = 1326

The Sample Space $n(S)$ = 1326.

Let E_1 be the event of 2 cards or Spade = $13C_2$ = $\frac{13 \times 12}{2 \times 1}$ = 78 = $n(E)$

Page No. :
Date :

P(both cards spade) $= \frac{78}{1326} = \underline{0.058}$

Let E_2 be the event of 1 card of spade & 1 of diamond $= \cancel{13}\ \cancel{20}_{C_2} \times \left(\frac{1}{13}\right)$

$= {}^{13}C_1 \times {}^{13}C_1 = 13 \times 13 = 169$

P(1 → Spade and 1 → diamond) $= \frac{169}{1326}$

$= \underline{0.1274}$

Properties:

I. If E is an event and $\bar{E}$ is its complementary event then

$$P(E) + P(\bar{E}) = 1$$

II. If E_1 and E_2 are any two events then

$$P(E_1 \cup E_2) = P(E_1) + P(E_2) - P(E_1 \cap E_2).$$

III. If E_1 and E_2 are mutually exclusive then

$$P(E_1 \cup E_2) = P(E_1) + P(E_2).$$

Date :

Que: Two cards are drawn at random from a pack of cards. Find the probability that both these cards are of black colour or both are aces.

Sample Space $= {}^{52}C_2$

$= 1326.$

E_1 (both cards are black) $= \cancel{13}C\ {}^{26}C_2$

$= \frac{26 \times 25}{2 \times 1} = 325.$

E_2 (both are aces) $= {}^4C_2 = \frac{4 \times 3}{2 \times 1} = 6$

$P(E_1) = \frac{325}{1326} = 0.245$

$P(E_2) = \frac{6}{1326} = 0.0045.$

$n(E_1 \cap E_2) = 1$

$P(E_1 \cap E_2) = \frac{1}{1326} = 0.00075$

$P(E_1 \cup E_2) = P(E_1) + P(E_2) - P(E_1 \cap E_2)$

$= 0.245 + 0.0045 - 0.0075$

$= 0.24$

Page No. :
Date :

$$= \frac{330}{1326} = \frac{165}{663} = \frac{55}{221}.$$

Ans: Conditional Probability.

When the happening of an event E_1 depend upon the happening of another event E_2. Then the probability of event is called Conditional probability. It is denoted by

$$P(E_1/E_2) -$$

i.e $P(E_1 \cap E_2) = P(E_1) . P(E_2/E_1)$

Que: If A & B are two events $P(A) = \frac{1}{2}$, $P(B) = \frac{1}{3}$ and $P(A \cap B) = \frac{1}{4}$.

Then Evaluate (I) $P(A/B)$ (II) $P(B/A)$ (III) $P(A \cup B)$

(I) $P(A/B) = \frac{P(A \cap B)}{P(B)} = \frac{1}{4} \times 3 = \frac{3}{4}$

(II) $P(B/A) = \frac{P(A \cap B)}{P(A)} = \frac{1}{4} \times \frac{2}{1} = \frac{1}{2}$

(III) $P(A \cup B) = P(A) + P(B) - P(A \cap B)$

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

$$= \frac{6+4-3}{12} = \frac{7}{12} //.$$