AL/CD - 302 (GS) (1)

B.Tech. III Semester
Examination, June - 2023
Introduction to Probability and Statistics

Prof. Samrat Singh Baby - 9893210626
College Code - 0177, IES College of Technology BPL

1(a) Define continuous random variable. Also show that $E(X_1 X_2 \ldots X_n) = E(X_1) E(X_2) \ldots E(X_n)$

Solution:- A random variable X is said to be continuous if it can take all possible values between certain limits. When we deal with variates like weights and temperature then we know that these variates can take on infinite number of values in a given interval. Such type of variate is known as continuous random variable.

Let X be a continuous random variable and let the probability of X falling in very small interval $(x - \frac{1}{2}dx, x + \frac{1}{2}dx)$ be expressed by $f(x)dx$.

i.e. $P[x - \frac{1}{2}dx < X < x + \frac{1}{2}dx] = f(x)dx$

where $f(x)$ is a continuous function of x and satisfies the following two conditions:

(i) $f(x) \geq 0$ (ii) $\int_a^b f(x)\,dx = 1$; $a \leq x \leq b$.

1(B) Define non-central and central moments.
Also show that $V(2X+3) = 4V(X)$
CS-central moment 2.

Sol/ In Probability theory and statistics a central moment is a moment of a probability distribution of a random variable about the random variable's mean; that is, it is the expected value of a specified integer power of the deviation of the random variable from the mean. The various moments from one set of values by which the properties of a probability distribution can usefully characterized. Central moments are use in preference to ordinary moments, computed in terms of deviations the mean instead of from zero, because the higher order central moments relate only to the spread and shape of the distribution, rather than also to its location.

$$\mu_n = E[[X - E(X)]^n] = \int_{-\infty}^{\infty} (x-\mu)^n f(x)\, dx$$

<u>Non Central moments</u>: Let $(X_1, X_2, \ldots X_k)$ be k independence, normally distributed random variables with means μ_i and unit variances. Then the random variable

$$\sum_{i=1}^{k} X_i^2$$

is distributed according to non-central chi-squared distribution, is has two parameters k which specifies the number of degrees of freedom and λ which is related to the mean of random variables X_i by $\lambda = \sum_{i=1}^{k} \mu_i^2$

λ is sometimes called the noncentrality parameter.

— ... distribution is a limiting case of binomial distribution under the case $n \to \infty$, $p \to 0$ and $np = m$.

Sol. Binomial distribution

$$P(X=r) = {}^nC_r\, q^{n-r} p^r = \frac{n!}{r!(n-r)!}(1-p)^{n-r} p^r \quad \because q = 1-p$$

$$= \frac{n(n-1)(n-2)\cdots(n-r+1)(n-r)!}{r!\,(n-r)!}(1-p)^{n-r}\times p^r$$

$$P(X=r) = \frac{n(n-1)(n-2)\cdots(n-r+1)}{r!}\left(1-\frac{m}{n}\right)^{n-r}\times\left(\frac{m}{n}\right)^r$$

$\because np = m \quad \therefore p = \frac{m}{n}$

$$= \frac{m^r}{r!}\times\frac{n(n-1)(n-2)\cdots(n-r+1)}{n^r}\times\frac{\left(1-\frac{m}{n}\right)^n}{\left(1-\frac{m}{n}\right)^r}$$

$$= \frac{m^r}{r!}\left(\frac{n}{n}\right)\left(\frac{n-1}{n}\right)\left(\frac{n-2}{n}\right)\cdots\left(\frac{n-r+1}{n}\right)\cdot\frac{\left(1-\frac{m}{n}\right)^n}{\left(1-\frac{m}{n}\right)^r}$$

$$P(X=r) = \frac{m^r}{r!}\left(1-\frac{1}{n}\right)\left(1-\frac{2}{n}\right)\cdots\left(1-\frac{r-1}{n}\right)\times\frac{\left[\left(1-\frac{m}{n}\right)^{-n/m}\right]^{-m}}{\left(1-\frac{m}{n}\right)^r} \quad \text{--- (1)}$$

Since n is very large so that

$$\therefore \lim_{n\to\infty}\left(1-\frac{1}{n}\right)\cdot\left(1-\frac{2}{n}\right)\cdots\left(1-\frac{r-1}{n}\right) = (1-0)(1-0)\cdots(1-0) = 1$$

and since $\left[\lim_{n\to\infty}\left(1-\frac{m}{n}\right)^{-n/m}\right]^{-m} = e^{-m}$ and $\lim_{n\to\infty}\left(1-\frac{m}{n}\right)^r = 1$

Then eq (1) becomes;

$$P(x=r) = \frac{e^{-m} m^r}{r!}, \quad r = 0,1,2,\ldots\infty$$

Which is probability function of Poisson distribution

https://www.rgpvonline.com

Q.1(B) Define norm - distribution. (+)
Q(b) Define gamma distribution with parameter θ, exponential distribution with parameter θ, obtain its mean and variance.

Sol. ① A continuous random variable X is said to be gamma distribution with parameter $\lambda > 0$, if its p.d.f. is given by:

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

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

(ii) The p.d.f. a gamma distribution with two parameters $a > 0$ and $\lambda > 0$, is given by

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

Exponential Distribution:

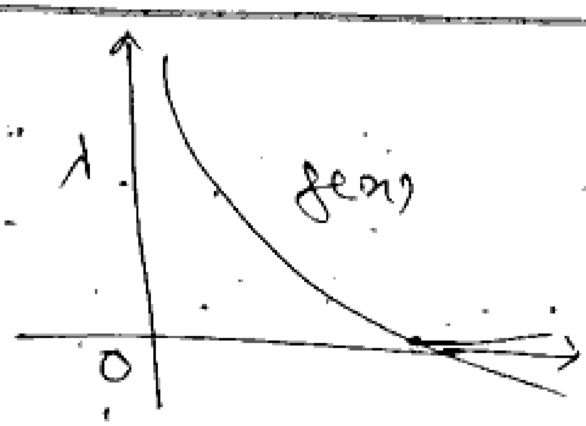
A random variable X is said to be an exponential distribution with parameter $\lambda > 0$, if its p.d.f. is given by

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

The graph of exponential distribution is:

Mean $= \mu_1' = \frac{1}{\lambda}$, $\mu_2' = \frac{2}{\lambda^2}$, Variance $= \mu_2 = \mu_2' - (\mu_1')^2 = \frac{1}{\lambda^2}$

https://www.rgpvonline.com



We have total Area $= \int_{-\infty}^{\infty} f(x)\,dx$

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

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

Mean, S.D. and Variance of Exponential Distribution

The p.d.f. of exponential distribution is given by:

$$f(x) = \begin{cases} \lambda e^{-\lambda x} & x \geq 0 \quad \lambda > 0 \\ 0 & \text{otherwise} \end{cases} \quad \text{— (k)}$$

To find mean:

we have mean $= \mu_1' = E(X)$

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

$$= \int_0^\infty x \cdot \lambda e^{-\lambda x}\,dx \quad \text{By eq (k)}$$

$$= \lambda \left[\frac{x e^{-\lambda x}}{-\lambda} \Big|_0^\infty - \int_0^\infty \frac{e^{-\lambda x}}{-\lambda}\,dx \right]$$

https://www.rgpvonline.com

$$\phi_x(t) = e^{\mu t}\int_{-\infty}^{\infty} \frac{1}{\sqrt{2\pi}} e^{t\sigma z} e^{-\frac{z^2}{2}} dz$$

$$= e^{\mu t}\int_{-\infty}^{\infty} \frac{1}{\sqrt{2\pi}} e^{-\frac{1}{2}(z^2 - 2t\sigma z + t^2\sigma^2) + \frac{1}{2}t^2\sigma^2}$$

$$= e^{\mu t + \frac{1}{2}t^2\sigma^2}\int_{-\infty}^{\infty} \frac{1}{\sqrt{2\pi}} e^{(z - t\sigma)^2} dz$$

$$= e^{\mu t + \frac{1}{2}t^2\sigma^2} \cdot 2\int_{0}^{\infty} \frac{1}{\sqrt{2\pi}} e^{-\frac{1}{2}(z - t\sigma)^2} dz$$

$$= e^{\mu t + \frac{1}{2}t^2\sigma^2} \sqrt{\frac{2}{\pi}} \int_0^{\infty} e^{-y} \frac{dy}{\sqrt{2y}}$$

where $\frac{1}{2}(z - t\sigma)^2 = y$

$(z - t\sigma)dz = dy \Rightarrow \sqrt{2y}\, dz = dy$

$$\phi_x(t) = e^{\mu t + \frac{1}{2}t^2\sigma^2} \sqrt{\frac{1}{\pi}} \int_0^{\infty} e^{-y} y^{1/2 - 1} dy$$

$$= \text{——} \cdot \frac{1}{\sqrt{\pi}} \Gamma\tfrac{1}{2} \quad \because \Gamma n = \int_0^{\infty} e^{-t} t^{n-1} dt$$

$$= \text{——} \cdot \frac{1}{\sqrt{\pi}} \cdot \sqrt{\pi} \qquad \boxed{\Gamma\tfrac{1}{2} = \sqrt{\pi}}$$

$$\phi_x(t) = e^{\mu t + \frac{1}{2}t^2\sigma^2}$$

$$\phi_x(t) = e^{\mu t + \frac{1}{2}t^2\sigma^2}$$

Expt. No. .... Date

Page No. (28)

Q4 (a) Write down merits and demerits of measure of central tendency. Find the median for the following data

Wages	2000-3000	3000-4000	4000-5000	5000-6000	6000-7000
No. of workers	03	5	20	10	5

Ans An average is computed to meet one or more of the following aspects

(i) To give a concise picture of a large group

(ii) To compare different groups by means of these simple figures

(iii) To obtain a picture of a complete group with the help of sample

(iv) To give a mathematical concept to the relationship between different groups

Advantages of measurement of central tendency.

(i) An average form convenient medium for representing a particular aspect or aspects of large group or mass of data.

(ii) An average taken from sample data is in the majority of cases, as accurate as a average calculated from the complete data

(iii) The influence of extremes or abnormalities is minimised.

(iv) An average forms a more convenient method of comparisons than a complete group or table

(v) An average is a useful basis for comparing past and present items and for forming the basis of future comparisons.

Teacher's Signature : ____________

https://www.rgpvonline.com

Disadvantages of measurement of Central tendency

(i) An average is characteristics of the data as a whole and does not show the trend.

(ii) An average may not be an actual item.

(iii) An average not supported by actual figures lead to fallacious conclusions.

(iv) Arithmetical inaccuracies are often made when calculating averages.

(v) the incorrect use of averages is prevalent. i.e. the wrong type of a average might be

Wages	No. of workers	Cumulative frequency, f
2000-3000	3	3
3000-4000	5	8
4000-5000	20	28
5000-6000	10	38
6000-7000	5	43

Here $N = 43$

$$\text{Median} = \left(\frac{N+1}{2}\right)^{th} \text{term}$$

$$= \left(\frac{43+1}{2}\right)^{th} = 22$$

$l_1 = 4000$, $l_2 = 5000$

$N = 43$, $c = 8$, $f = 20$

$i = l_2 - l_1 = 1000$

$$\text{Median} = l_1 + \frac{\frac{N}{2} - c}{f} \times i$$

$$= 4000 + \frac{\frac{43}{2} - 8}{20} \times 1000$$

$$\boxed{M_d = 4678}$$

https://www.rgpvonline.com

5(a) Using method of least squares, find the curve $y = ax + ax^2$ that best fit the following data

x	1	2	3	4	5
y	1.8	5.1	8.9	14.1	19.8

Sol. Given $y = ax + bx^2$

Let $(x_1, y_1)(x_2, y_2)(x_3, y_3) \ldots (x_n, y_n)$ be the points (1)

Then error or estimate for i^{th} the points (x_i, y_i) is

$$E_i = (y_i - ax_i - bx_i^2)$$

By Principle of least squares the values of a and b are such that

$$S = \Sigma E_i^2 = \Sigma (y_i - ax_i - bx_i^2)^2 \text{ is minimum}$$

∴ Normal equation are given by

$$\frac{\partial S}{\partial a} = 0 \quad \text{and} \quad \frac{\partial S}{\partial b} = 0$$

$$\Rightarrow \Sigma x_i y_i = a \Sigma x_i^2 + b \Sigma x_i^3 \text{ and}$$

$$\Sigma x_i^2 y_i = a \Sigma x_i^3 = b \Sigma x_i^4$$

or $$\Sigma xy = a \Sigma x^2 + b \Sigma x^3 \quad \text{—— (2)}$$

$$\Sigma x^2 y = a \Sigma x^3 + b \Sigma x^4 \quad \text{—— (3)}$$

Teacher's Signature :

https://www.rgpvonline.com

(11)

x	y	xy	x^2	x^3	x^4	x^2y
1	1.8	1.8	1	1	1	1.8
2	5.1	10.2	4	8	16	20.4
3	8.9	26.7	9	27	81	80.1
4	14.1	56.4	16	64	256	225.6
5	19.8	99	25	125	625	495
		$\Sigma xy = 194.1$	$\Sigma x^2 = 55$	$\Sigma x^3 = 225$	$\Sigma x^4 = 979$	$\Sigma x^2y = 822.9$

Putting all values from the table in the normal equation, we get

$$55a + 225b = 194.1 \quad \text{--- (4)}$$

$$225a + 979b = 822.9 \quad \text{--- (5)}$$

On solving equation (4) and (5), we get

$$a = 1.52 \text{ and } b = 0.49$$

Hence required curve eq (1) becomes

$$y = 1.52x + 0.49x^2$$ Ans

https://www.rgpvonline.com

5(b) Out of 8000 graduates in a town, 800 are female, out of 1600 graduate employes, 120 are females use χ^2 test to determine if any distinction is made in appointment on the basis of sex. The value of χ^2 for 1 degree of freedom at 5% level is 3.841.

Sol.

	Graduate employee	Graduate non-employee	Total
Female	120	800	920 B_1
Male	1480	7200	8680 B_2
Total	1600 A_1	8000 A_2	9600

H_0: There is no distinction is made in appointment on the basis of sex.

H_1: There is distinction is made in appointment on the basis of sex.

Now, we find the expected frequencies with the help of given contingency table of observed frequencies

$$E_{11} = \frac{A_1 \times B_1}{N} = \frac{1600 \times 920}{9600} = 153.33$$

$$E_{12} = \frac{A_1 \times B_2}{N} = \frac{1600 \times 8680}{9600} = 1446.67$$

$$E_{21} = \frac{A_2 \times B_1}{N} = \frac{8000 \times 920}{9600} = 766.67$$

https://www.rgpvonline.com

(17) $E_{22} = \frac{A_2 \times B_2}{N} = \frac{8000}{9600}$

Now the contigency table is:

	Graduate Employee	Graduate Non-Employee	Total
Female	153.33	766.67	920
Male	1446.67	7233.33	8680
Total	1600	800	9600

Elementary Theory of Testing of Hypothesis

" Small Sample "

O	E	O−E	$(O-E)^2$	$(O-E)^2/E$
120	153.33	−33.33	1110.889	7.25
7480	1446.67	33.33	1110.889	0.768
800	766.67	33.33	1110.889	1.45
~~7200~~	7233.33	−33.33	1110.889	0.153
				9.621
Total				

$\therefore \chi^2 = \sum_{i=1}^{n} \frac{(O_i - E_i)^2}{E_i} = 9.621$

$d.f. = (2-1)(2-1) = 1$

from table $\chi_1^2 = (0.05) = 3.84$, Since calculated values of χ^2 statistic at 5% level of significance for 1 degree freedom is greater than the tabulated value of χ^2. some reject our null hypothesis and conclude the ~~basis of sex, distinction is made to appoint on the~~

0 otherwise

The graph of exponential distribution is

6(a) Q) Define Poisson distribution and obtain its mean and variance.

Sol.

Poisson distribution:- The poisson distribution was discovered by a French mathematics Simen Denis Poisson in 1837.

It is a discrete distribution and is very widely used. Poisson distribution is a limiting form of the Binomial distribution in which n, the number of trials, becomes very large and p, the probability of the success of the event, is very very small such that mean m = np is a finite quantity.

Probability of r - Successes

$$P(X=r) = \frac{e^{-m} m^r}{r!}, \quad r = 0, 1, 2, 3, \ldots$$

Also theoretical or expected frequencies is

$$f(X=r) = \frac{N \cdot e^{-m} m^r}{r!}, \text{ where } m = \text{mean and } N \text{ is number of trials.}$$

The following are the statistical measures of the poisson distribution

(i) mean = np or m (ii) Variance m or np.

(iii) Standard Deviation $\sigma = \sqrt{m}$

(iv) Moment measure of Skewness $(\gamma_1) = \frac{1}{\sqrt{m}}$

(v) moment measure of kurtosis $(\gamma_2) = \frac{1}{m}$.

https://www.rgpvonline.com

(15) For the Poisson distribution $P(x=r) = \frac{e^{-m} m^r}{r!}, r = 0, 1, \ldots$

We know that. mean $= \mu_1'$ —(1)

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

Where μ_1' and μ_2' are first and second moment about origin. for $r = 0, 1, 2, \ldots \infty$ first moment about origin.

$$\mu_1' = \sum_{r=1}^{\infty} r P(r) = \sum_{r=0}^{\infty} r \frac{e^{-m} m^r}{r!} \quad \text{—(2)} \quad (m = \text{mean})$$

$$= \sum_{r=0}^{m} \frac{r m^r}{r(r-1)!} = e^{-m} \sum_{r=1}^{\infty} \frac{m^r}{(r-1)!} = e^{-m}\left[m + \frac{m^2}{1!} + \frac{m^3}{2!} + \ldots\right]$$

$$\mu_1' = m e^{-m}\left[1 + \frac{m}{1!} + \frac{m^2}{2!} + \ldots\right] = m e^{-m} e^{m}$$

$$\text{mean} = \boxed{\mu_1' = m} \text{—(3)}$$

Now, second moment about origin

$$\mu_2' = \sum_{r=0}^{\infty} r^2 P(r) \qquad \text{since } r^2 = r(r-1) + r$$

$$= \sum_{r=0}^{\infty} r(r-1) \cdot \frac{e^{-m} m^r}{r!} + \sum_{r=0}^{\infty} \frac{r e^{-m} m^r}{r!} \quad (m = \text{mean})$$

$$= e^{-m} \sum_{r=2}^{\infty} \frac{m^r}{(r-2)!} + \mu_1' = e^{-m}\left[m^2 + \frac{m^3}{1!} + \frac{m^4}{2!} + \ldots\right] + m$$

$$= m^2 e^{-m}\left[1 + \frac{m}{1!} + \frac{m^2}{2!} + \ldots\right] + m = m^2 e^{-m} e^{m} + m$$

$$\boxed{\mu_2' = m^2 + m}$$

variance $= \mu_2' - (\mu_1')^2 = m^2 + m - m^2 = m$.

Variance $= m$

otherwise

The graph of exponential distribution is

Mean $= \mu_1' = \frac{1}{\lambda}$, $\mu_2' = \frac{2}{\lambda^2}$, Variance $= \mu_2 = \mu_2' - (\mu_1')^2 = \frac{1}{\lambda^2}$

https://www.rgpvonline.com

6(b) Q) Define gamma distribution with parameter λ and obtain its mean, variance and characteristic function.

Sol Gama Distribution:-

(i) A continuous random variable X is said to be a gama distribution with parameter $\lambda > 0$, if its p.d.f. is given by

$$f(x) = \begin{cases} \dfrac{e^{-x} x^{\lambda-1}}{\Gamma\lambda}, & 0 < x < \infty, \ \lambda > 0 \\ 0 & \text{otherwise} \end{cases} \quad \text{—(i)}$$

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

To find mean:- We have Mean $= \mu_1' = E(X)$

$$\mu_1' = \int_{-\infty}^{\infty} x f(x) dx = \int_0^\infty x \frac{e^{-x} x^{\lambda-1}}{\Gamma\lambda} dx \quad \text{By eq (i)}$$

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

$\left[\because \Gamma n = \int_0^\infty e^{-t} t^{n-1} dt \right.$

$\Gamma n+1 = n\Gamma n$

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

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

mean $= \mu_1' = \lambda$

https://www.rgpvonline.com

(i) The second moment about origin!

$$\mu_2' = \int_{-\infty}^{\infty} x^2 f(x)\,dx = \int_0^{\infty} x^2 \cdot \frac{e^{-x} x^{\lambda-1}}{\Gamma\lambda}\,dx \quad \text{by } (1)$$

$$\mu_2' = \frac{1}{\Gamma\lambda}\int_0^{\infty} e^{-x} x^{\lambda+2-1}\,dx = \frac{1}{\Gamma\lambda}\Gamma\lambda+2$$ [By def. of Gamma fn.]

$$\Gamma n = \int_0^{\infty} e^{-t} t^{n-1}\,dt,\ n>0$$

$$= \frac{(\lambda+1)(\lambda)\Gamma\lambda}{\Gamma\lambda} \qquad \left\{ \Gamma n+1 = n\Gamma n \right.$$

$$\boxed{\mu_2' = \lambda^2+\lambda}$$

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

$$\mu_2 = \lambda^2+\lambda-(\lambda)^2 = \lambda$$

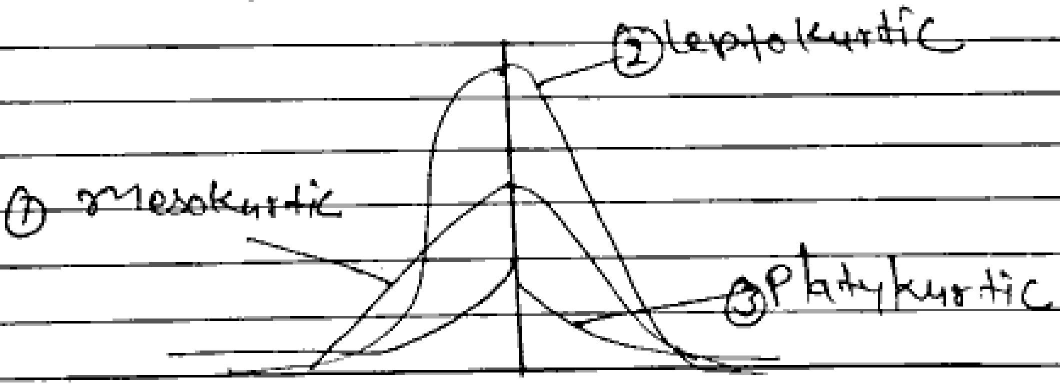
$$\boxed{\text{Variance} = (\mu_2) = \lambda}$$

7/Q1 What do you mean by measure of kurtosis? obtain β_1 and β_2 for the following function

$$f(x) = y_0 x(2-x) \quad ; \quad 0 \le x \le 2$$

Sol. Measure of kurtosis:- In statistics, kurtosis refers to the degree of flatness or peakedness in the region about the mode of frequency curve. "The degree of kurtosis of a distribution is measured relative to the peakedness of normal curve".

Karl Pearson in 1905 defined following three types of curves



Second and fourth moments are use to measure kurtosis. ~~Karl~~ Karl Pearson gave the following formula to measure kurtosis:

$$\text{Kurtosis or } \beta_2 = \mu_4 / \mu_2^2$$

to measure kurtosis, γ_2 is used and it is given by following formula

$$\gamma_2 = \beta_2 - 3 = \frac{\mu_4 - 3\mu_2^2}{\mu_2^2}$$

Teacher's Signature : ....

https://www.rgpvonline.com

7(b) Karl Pearson's Correlation coefficient:

It the best mathematical method to find correlation since it is based on mean and standard deviation. By this method we can find not only the direction and magnitude of correlation but also it's positive measure. Karl Pearson (1862 - 1936) developed a formula (in 1890) called correlation coefficient.

Correlation coefficient between two random variables X and Y, denoted by r, is a numerical measure of linear relationship between X and Y and is defined (by Karl Pearson) by

$$r = \frac{\Sigma xy}{\sqrt{\Sigma x^2 \cdot \Sigma y^2}} = \frac{\Sigma xy}{n \sigma_x \sigma_y} = \frac{P}{\sigma_x \sigma_y}$$

Where

$x = X - M_x$ = deviation of variable X measured from its mean M_x

$y = Y - M_y$

σ_x = Standard deviation of X - Series

σ_y = " " Y - "

$P = \{\Sigma xy / n\}$

n = number of pairs of two variables

r is called the product moment correlation coefficient.

https://www.rgpvonline.com

(20)

x	y	$x = X - M_x$	$y = Y - M_y$	x^2	y^2	xy
65	67	−3	−2	9	2	+6
66	68	−2	−1	4	1	+2
67	65	−1	−4	1	16	+4
67	68	−1	−1	1	1	+1
68	72	0	3	0	9	0
69	72	1	3	1	9	3
70	69	2	0	4	0	0
72	71	4	2	16	4	8
$\Sigma X = 544$	$\Sigma Y = 542$	0	0	36	42	24

$M_x = \frac{\Sigma X}{n}$, $M_y = \frac{\Sigma Y}{n} = \frac{542}{8} = 69$

$= \frac{544}{8} = 68$

$\sigma_x = \sqrt{\frac{\Sigma x^2}{n}} = \sqrt{\frac{36}{8}}$, $\sigma_y = \sqrt{\frac{\Sigma y}{n}} = \sqrt{\frac{42}{8}} =$

$= 2.121$ $\qquad \sigma_y = 2.291$

$r = \frac{\Sigma xy}{\sqrt{\Sigma x^2 \cdot \Sigma y^2}} = \frac{\Sigma xy}{n \sigma_x \sigma_y} = \frac{24}{8 \times 2.121 \times 2.91} = 0.486$

$\boxed{r = 0.486}$

0 otherwise

graph of exponential distribution is ... 2

https://www.rgpvonline.com

Q.3

	x	y	Rank in x	Rank in y	X-Y = d	d^2
1	65	68	4 4	9	-5	25
2	63	66	2 2	4	-2	04
3	~~64~~ 67	68	7 7	8	-1	01
4	~~68~~ 64	65	3 3	2	1	01
5	~~62~~ 68	69	8 8	10	-2	04
6	~~70~~ 62	66	1 1	3	-2	04
7	~~66~~ 70	68	11 11	7	+4	16
8	~~68~~ 66	65	5 5	1	4	16
9	~~67~~ 68	71	9 9	12	-3	09
10	~~69~~ 67	67	6 6	5	1	01
11	~~71~~ 69	68	10 10	6	4	16
12	71	70	12 12	11	1	01
	$\Sigma x =$	$\Sigma y =$			$\Sigma d = 0$	$\Sigma d^2 = 98$

$$r = 1 - \frac{6\Sigma d^2}{n(n^2-1)} = 1 - \frac{6 \times 98}{12(144-1)}$$

$$= 0.652 \quad \text{Ans}$$

8(b) Binomial Distribution:- Binomial distribution is a discrete probability distribution. Let an experiment consisting of n-trials be performed and let the occurrence of an event in any trial be called a success and its non-occurrence a failure. Let p be the probability of success and q be the probability of the failure in a single trial, when $q = 1-p$, so that $p+q=1$.

Teacher's Signature : ____________

https://www.rgpvonline.com

$\therefore$ Probability of r successes in n-trials

$$P(r) = {}^{n}C_{r}\, q^{n-r} p^{r}, \quad r = 0,1,2,\ldots n$$

$$\text{or } P[X=r] = {}^{n}C_{r}\, q^{n-r}, \quad r = 0,1,2,\ldots n$$

Where $P(X=r)$ or $P(r)$ is called binomial probability distribution of a random variable, X.

If an experiment consists of n-trials and this experiment is repeated N times, then Binomial frequency distribution or Expected frequency of r successes is

$$f(r) = N \cdot P(r)$$

or

$$f(r) = N \cdot {}^{n}C_{r}\, q^{n-r} p^{r}, \quad r = 0,1,2 \ldots n,$$

Mean, Variance of Binomial Distribution

We know that

$$\text{Mean} = \mu_1' \text{ and Variance} = \mu_2' - (\mu_1')^2 \quad (1)$$

Where μ_1' and μ_2' are first and second moment about origin, respectively.

We have first moment about origin,

$$\mu_1' = \sum_{r=0}^{n} r\,P(r) \quad \text{where } P(r) = {}^{n}C_{r}\, q^{n-r} p^{r}$$

is Binomial probability distribution.

$$\Rightarrow \mu_1' = \sum_{r=0}^{n} r \cdot {}^{n}C_{r}\, q^{n-r} p^{r} = \sum_{r=1}^{n} r \cdot \frac{n!}{r!\,(n-r)!}\, q^{n-r} p^{r}$$

$\because r\,P(r) = 0$ for $r=0$

https://www.rgpvonline.com

$$\mu' = \sum_{r=1} \frac{n!}{(r-1)!\,(n-r)!} q^{n-r} p^r$$

$$= p \sum_{r=1}^{n} \frac{n(n-1)!}{(r-1)!\,(n-1-\overline{r-1})} q^{n-1-\overline{r-1}} p^{r-1}$$

$$\left[\because {}^{n-1}C_{r-1} = \frac{(n-1)!}{(r-1)!\,(n-1-\overline{r-1})}\right]$$

$$\mu_1' = np \sum_{r=1}^{n} {}^{n-1}C_{r-1}\, q^{n-1-\overline{r-1}} \cdot p^{\overline{r-1}} = np(q+p)^{n-1}$$

$$\boxed{\mu_1' = np} \qquad \boxed{\because q+p=1}$$

Thus, $\boxed{\text{mean} = \mu_1' = np}$ (B) i.e. $\boxed{m = np}$

Now, second moment about origin is

$$\mu_2' = \sum_{r=0}^{n} r^2 p(r) = \sum_{r=0}^{n} r^2\, {}^nC_r q^{n-r} p^r$$

$$= \sum_{r=0}^{n} \{r(r-1)+r\}\, {}^nC_r q^{n-r} p^r \quad [\because r^2 = r(r-1)+r$$

$$= \sum_{r=0}^{n} r(r-1)\, {}^nC_r q^{n-r} p^r + \sum_{r=0}^{n} r \cdot {}^nC_r q^{n-r} p^r$$

$$= \sum_{r=2}^{n} r(r-1) \frac{n!}{r!\,(n-r)!} q^{n-r} p^r + np \quad \text{By eq (B)}$$

(27)

Expt. No. Page No. Date

$$\mu_2' = \sum_{r=2}^{n} \frac{r(r-1)\,n(n-1)(n-2)!}{r(r-1)(r-2)!\,(n-r)!} q^{n-r} p^r + np$$

$$\mu_2' = n(n-1)p^2 \sum_{r=2}^{n} \frac{(n-2)!}{(r-2)!\,(n-2-\overline{r-2})!} q^{\overline{n-2}-\overline{r-2}} p^{r-2} + np$$

$$= n(n-1)p^2 \sum_{r=2}^{n} {}^{n-2}C_{r-2}\, q^{n-2-(r-2)} p^{r-2} + np$$

$$= n(n-1)p^2 (q+p)^{n-2} + np \quad \boxed{\because q+p=1}$$

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

$$= n^2p^2 + np(1-p) \quad \boxed{\because q = 1-p}$$

$$\boxed{\mu_2' = n^2p^2 + npq}$$

Since, variance $= \mu_2' - (\mu_1')^2$

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

$$= npq$$

$$\boxed{\text{Variance}(\mu_2) = npq}$$

Teacher's Signature

https://www.rgpvonline.com