Q-1. The first four moments of a distribution about 30.2 of the variable are 0.255, 6.222, 30.211, and 400.25. Find central moments and β_1 and β_2. Also comment on skewness and kurtosis of the distribution.

[May 2017]

→ Given:- $\mu_1' = 0.255$

$\mu_2' = 6.222$

$\mu_3' = 30.211$

$\mu_4' = 400.25$

$A = 30.2.$

We know,

$$\mu_r = \mu_r' - {}^rC_1\, \mu_{r-1}' . \mu_1' + {}^rC_2\, \mu_{r-2}' . (\mu_1')^2 - \ldots$$

$$\therefore \mu_1 = \mu_1' - {}^1C_1\, \mu_0' . \mu_1'$$

$$= \mu_1' - \mu_1' \quad \ldots \left[\because \mu_0' = 1\right]$$

$$\boxed{\mu_1 = 0}$$

$$\mu_2 = \mu_2' - {}^2C_1\, \mu_1' . \mu_1' + {}^2C_2\, \mu_0' . (\mu_1')^2$$
$$= \mu_2' - 2(\mu_1')^2 + (\mu_1')^2 \quad \ldots\ldots [\because \mu_0' = 1]$$
$$= \mu_2' - (\mu_1')^2$$
$$= 6.222 - (0.255)^2$$
$$= 6.222 - (0.0650)$$
$$\boxed{\mu_2 = 6.157}$$

$$\mu_3 = \mu_3' - {}^3C_1\, \mu_2' . \mu_1' + {}^3C_2\, \mu_1' . (\mu_1')^2 - {}^3C_3\, \mu_0' (\mu_1')^3$$
$$= \mu_3' - 3\mu_2' . \mu_1' + 3(\mu_1')^3 - (\mu_1')^3 \quad \ldots\ldots [\because \mu_0' = 1]$$
$$= \mu_3' - 3\mu_2' \mu_1' + 2(\mu_1')^3$$
$$= 30.211 - 3(6.222)(0.255) + 2(0.255)^3$$
$$= 30.211 - 3(1.5866) + 2(0.0165)$$
$$= 30.211 - 4.7598 + 0.033$$
$$\boxed{\mu_3 = 25.4842}$$

$$\mu_4 = \mu_4' - {}^4C_1\, \mu_3' . \mu_1' + {}^4C_2\, \mu_2' . (\mu_1')^2 - {}^4C_3\, \mu_1' . (\mu_1')^3 + {}^4C_4\, \mu_0' . (\mu_1')^4$$
$$= \mu_4' - 4\mu_3' . \mu_1' + 6\mu_2' . (\mu_1')^2 - 4(\mu_1')^4 + (\mu_1')^4$$
$$= 400.25 - 4(30.211)(0.255) + 6(6.222)(0.255)^2 - 3(0.255)^4$$
$$= 400.25 - 30.8152 + 2.427 - 0.0126$$
$$\boxed{\mu_4 = 371.8492.}$$

$$\beta_1 = \frac{\mu_3^2}{\mu_2^3} = \frac{(25.4842)^2}{(6.157)^3} = \frac{649.4444}{233.4035} = 2.7824.$$
$$\gamma_1 = \sqrt{\beta_1} = \sqrt{2.7824} = 1.6680$$
$\therefore$ It indicates considerable positive skewness.

$$\beta_2 = \frac{\mu_4}{\mu_2^2} = \frac{371.8492}{(6.157)^2} = \frac{371.8492}{37.9086} = 9.8090$$
$\therefore$ As $\beta_2 > 3$, Distribution is Leptokurtic.

Q-2. Find the regression equation of y on x for a bivariate data with the following details: $n = 25$, $\sum_{i=1}^{n} x_i = 75$, $\sum_{i=1}^{n} y_i = 100$

$\sum_{i=1}^{n} x_i^2 = 250$, $\sum_{i=1}^{n} y_i^2 = 500$, $\sum_{i=1}^{n} x_i y_i = 325$. [Nov. 2019]

→ Given :- $n = 25$, $\sum x_i = 75$

$\sum y_i = 100$

$\sum x^2 = 250$

$\sum y^2 = 500$

$\sum xy = 325$

To find :- regression equation of y on x.

∴ The line of regression of y on x is

$$y - \bar{y} = r \cdot \frac{\sigma_y}{\sigma_x}(x - \bar{x})$$

∴ we need to find :- $\bar{y}$, $\bar{x}$, σ_y, σ_x, r.

$$\therefore \quad \bar{y} = \frac{\sum y_i}{n} = \frac{100}{25} = 4$$

$$\therefore \boxed{\bar{y} = 4}$$

$$\bar{x} = \frac{\sum x_i}{n} = \frac{75}{25} = 3$$

$$\boxed{\bar{x} = 3}$$

$$\sigma_x = \sqrt{\frac{1}{n}(\Sigma x^2) - (\bar{x})^2}$$

$$= \sqrt{\frac{1}{25} \cdot 250 - (3)^2} \quad \ldots \left[\bar{x} = 3,\ \Sigma x^2 = 250,\ n = 25\right]$$

$$= \sqrt{10 - 9}$$

$$\boxed{\sigma_x = 1}$$

$$\sigma_y = \sqrt{\frac{1}{n}(\Sigma y_i^2) - (\bar{y})^2}$$

$$= \sqrt{\frac{1}{25}(500) - (4)^2} \quad \ldots \left[\bar{y} = 4,\ \Sigma y^2 = 500,\ n = 25\right]$$

$$= \sqrt{20 - 16}$$

$$\boxed{\sigma_y = 2}$$

$$r = \frac{cov.(x,y)}{\sigma_x \cdot \sigma_y} = \left[\left[\frac{1}{n} \Sigma x_i y_i\right] - \bar{x}\bar{y}\right] \frac{1}{\sigma_x \cdot \sigma_y}$$

$$= \left[\frac{325}{25} - 3 \times 4\right] \times \frac{1}{1 \times 2}$$

$$= \left[13 - 12\right] \times \frac{1}{2}$$

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

$$y - 4 = \frac{1}{2} \times \frac{2}{1}(x - 3)$$

$\therefore$ The line of regression of y on x

$$\boxed{y = x + 1}$$

Q-3. For a bivariate data, the regression equation of y on x is $8x - 10y + 66 = 0$ and the regression equation x on y is $40x - 18y = 214$. Find the mean values of x and y. Also, find the correlation coefficient between x and y.

[May 2019]

→ Given:- Line of regressions as $8x - 10y + 66 = 0$ and $40x - 18y = 214$.

To find:- ① Mean of $x = \bar{x}$

② Mean of $y = \bar{y}$

③ Correlation coefficient $= r$.

We know, both mean of x and y satisfy line of regressions.

$$\therefore \quad 8\bar{x} - 10\bar{y} = -66 \quad — ①$$

$$40\bar{x} - 18\bar{y} = 214 \quad — ②$$

After solving 2 linear equation, we get

$$\bar{x} = 13 \quad \text{and} \quad \bar{y} = 17$$

$$y = \frac{8x}{10} + \frac{66}{10} \quad — ③$$

$$x = \frac{18y}{40} + \frac{214}{40} \quad — ④$$

We know, b_{yx} = coefficient of x in line of regression of y on x

b_{xy} = coefficient of y in line of regression of x on y.

$\therefore \ b_{yx} = \frac{8}{10}$, $b_{xy} = \frac{18}{40}$

And $b_{yx} \cdot b_{xy} = r^2$

$$\frac{8}{10} \times \frac{18}{40} = r^2$$

$$\boxed{r = \frac{3}{5}}$$

$\therefore$ Mean of x is 13, Mean of $y = 17$, Correlation Coefficient $(r) = \frac{3}{5}$.

Q-4. The first four moments of a distribution about the value 4 are -1.5, 17, -30 and 108. Obtain the first four central moments, β_1 and β_2.

[Nov. 2018]

$\rightarrow$ Given:- $\mu_1' = -1.5$, $\mu_2' = 17$, $\mu_3' = -30$, $\mu_4' = 108$ and $A = 4$.

We know,

$$\mu_r = \mu_r' - {}^rC_1\, \mu_{r-1}' \cdot \mu_1' + {}^rC_2\, \mu_{r-2}' (\mu_1')^2 - {}^rC_3\, \mu_{r-3}' (\mu_1')^3 \ldots + (-1)^r (\mu_1')^r$$

$$\therefore \quad \mu_1 = \mu_1' + (-1)^1 (\mu_1')^1$$

$$\boxed{\mu_1 = 0}$$

$$\mu_2 = \mu_2' - {}^2C_1\, \mu_1' \cdot \mu_1' + (-1)^2 (\mu_1')^2$$

$$\boxed{\mu_2 = 14.75}$$

$$\mu_3 = \mu_3' - {}^3C_1 \, \mu_2' . \mu_1' + {}^3C_2 \, \mu_1' (\mu_1')^2 + (-1)^3 . (\mu_1')^3$$

$$= (-30) - 3(17)(-1.5) + 2(-1.5)^3$$

$$\boxed{\mu_3 = 39.75}$$

$$\mu_4 = \mu_4' - {}^4C_1 \, \mu_3' . \mu_1' + {}^4C_2 \, \mu_2' (\mu_1')^2 - {}^4C_3 \, \mu_1' . (\mu_1')^3 + (-1)^4 (\mu_1')^4$$

$$= (108) - 4(-30)(-1.5) + 6(17)(-1.5)^2 - 3(-1.5)^4$$

$$\boxed{\mu_4 = 157.49}$$

$\therefore$ The first four moments about central are 0, 14.75, 39.79, 157.49.

Q-5. For a bivariate data, the regression equation of y on x is $4x + y = \mu$ and the regression equation of x on y is $9x + y = \lambda$. Find the value of μ and λ. Also, find the correlation coefficient between x and y, if the mean of x and y are 2 and -3 respectively.
[Nov. 2018]

$\rightarrow$ We know, mean of x and y satisfy the lines of regregression.

$$\therefore \quad 4\bar{x} + \bar{y} = \mu.$$

$$\mu = 4 \times 2 + (-3)$$

$$\boxed{\mu = 5}$$

$9\bar{x} + \bar{y} = \lambda$

$\therefore \lambda = 9 \times 2 + (-3)$

$\boxed{\lambda = 15}$

$\therefore$ Lines of regression are

$y = -9x + 15$ and $x = -\frac{y}{4} + \frac{5}{4}$.

We know, b_{yx} = coefficient of x in line of regression of y on x.

b_{xy} = coffient of y in line of regression of x on y

$\therefore b_{yx} = -9$ and $b_{xy} = -\frac{1}{4}$.

And $b_{yx} \cdot b_{xy} = -9 \times \frac{-1}{4} = \frac{9}{4} = r^2$

$\therefore \boxed{r = \frac{3}{2}}$

$\therefore \mu = 3$, $\lambda = 15$ and correllation of coffient $(r) = \frac{3}{2}$.