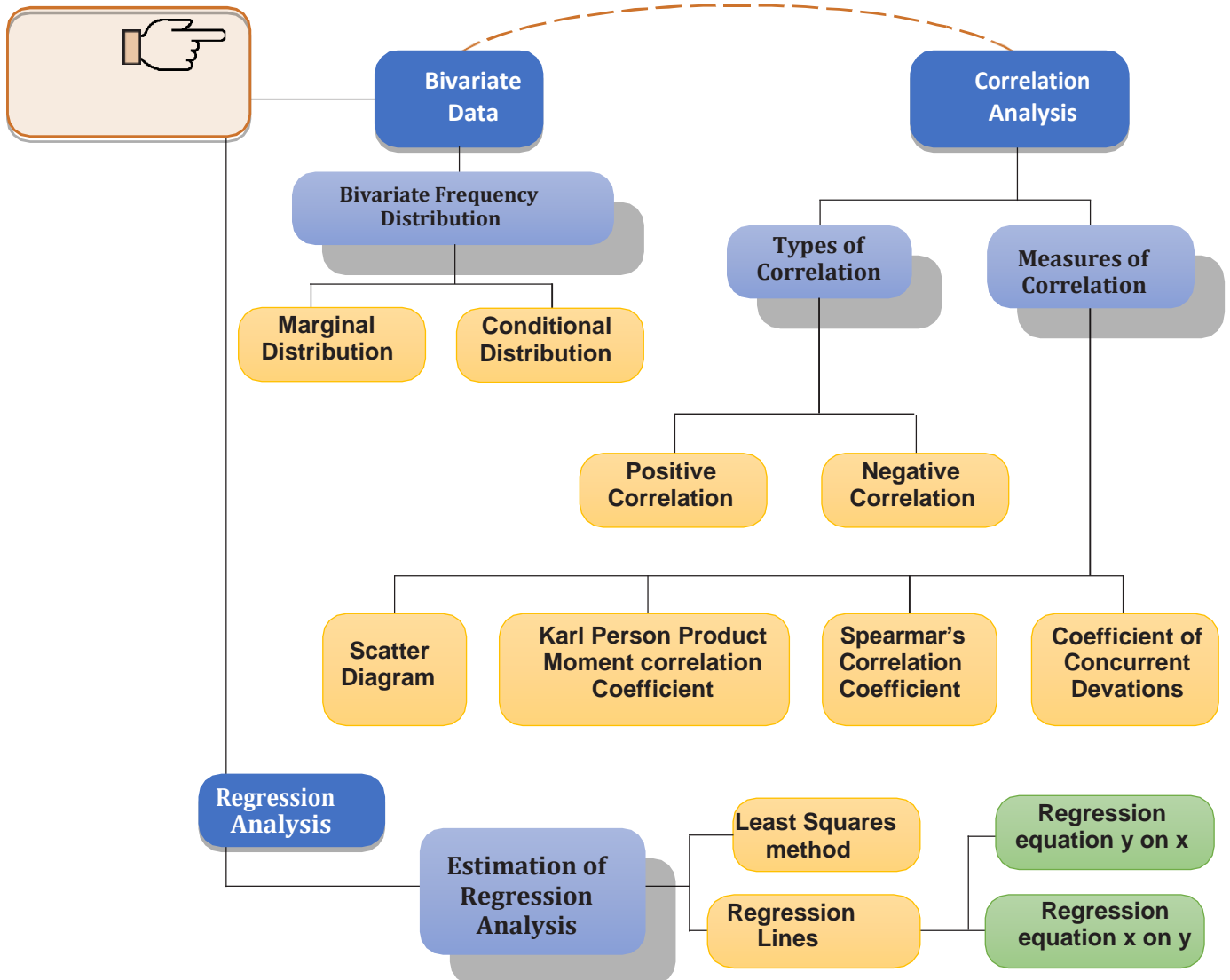# CHAPTER - 17
# CORRELATION AND REGRESSION

**Bivariate Data**

**Correlation Analysis**

**Bivariate Frequency Distribution**

**Marginal Distribution**

**Conditional Distribution**

**Types of Correlation**

**Measures of Correlation**

**Positive Correlation**

**Negative Correlation**

**Scatter Diagram**

**Karl Person Product Moment correlation Coefficient**

**Spearmar's Correlation Coefficient**

**Coefficient of Concurrent Devations**

**Regression Analysis**

**Estimation of Regression Analysis**

**Least Squares method**

**Regression Lines**

**Regression equation y on x**

**Regression equation x on y**
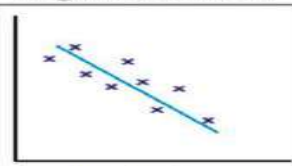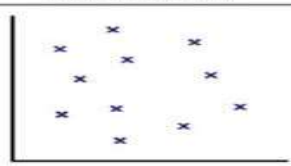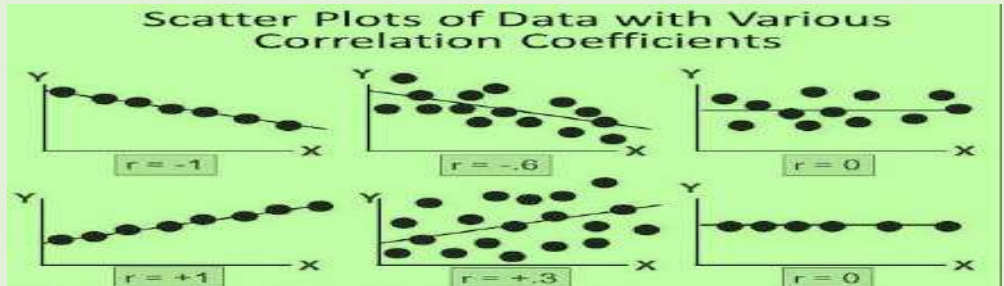
| CORRELATION | The change in one variable is reciprocated by a corresponding change in the other variable either directly or inversely, then the two variables are known to be associated or correlated |
|---|---|

| TYPES OF CORRELATION |  |
|---|---|

**Positive correlation**

The points lie close to a straight line, which has a positive gradient.

This shows that as one variable **increases** the other **increases**.

**Negative correlation**

The points lie close to a straight line, which has a negative gradient.

This shows that as one variable **increases**, the other **decreases**.

**No correlation**

There is no pattern to the points.

This shows that there is **no connection** between the two variables.

## MEASURES OF CORRELATION

### Scatter Plots of Data with Various Correlation Coefficients



$r = -1$  $r = -.6$  $r = 0$

$r = +1$  $r = +.3$  $r = 0$

## Rank Correlation

☐ The formula of Spearman's Rank correlation coefficient, is given as:

$$r_s = 1 - \frac{6 \sum d_i^2}{n(n^2 - 1)}$$

$r_s$ : is the coefficient of rank correlation.

$d_i$ : is the difference in rank between paired values of $X$ and $Y$, it can be calculated as ($rank\ of\ X_i - rank\ of\ Y_i$).

$n$ : is a sample size or the number of pairs values $X$ and $Y$ in the selected sample.

## Correlation Coefficient

**Correlation coefficient**

- A measure of the strength and the direction of a linear relationship between two variables.
- The symbol $r$ represents the sample correlation coefficient.
- A formula for $r$ is

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

$n$ is the number of data pairs

- The population correlation coefficient is represented by $\rho$ (rho).

## Pearson's Correlation Coeff.

Pearson's correlation coefficient between two variables is defined as the <u>covariance</u> of the two variables divided by the product of their <u>standard deviations</u>:

$$\rho_{X,Y} = \frac{\text{cov}(X,Y)}{\sigma_X \sigma_Y} = \frac{E[(X - \mu_X)(Y - \mu_Y)]}{\sigma_X \sigma_Y},$$

| | The above formula defines the *population* correlation coefficient, commonly represented by the Greek letter *ρ* (rho). Substituting estimates of the covariances and variances based on a <u>sample</u> gives the *sample correlation coefficient*, commonly denoted *r* : |
|---|---|

$$r = \frac{\sum_{i=1}^{n}(X_i - \bar{X})(Y_i - \bar{Y})}{\sqrt{\sum_{i=1}^{n}(X_i - \bar{X})^2}\sqrt{\sum_{i=1}^{n}(Y_i - \bar{Y})^2}}.$$

$$r_{xy} = \frac{\sum x_i y_i - n\bar{x}\bar{y}}{n s_x s_y} = \frac{n\sum x_i y_i - \sum x_i \sum y_i}{\sqrt{n\sum x_i^2 - (\sum x_i)^2}\sqrt{n\sum y_i^2 - (\sum y_i)^2}}.$$

| | |
|---|---|
| **POINT TO SIGNIFY** | The 'coefficient of non-determination' is given by $(1-r^2)$ and can be interpreted as the ratio of unexplained variance to the total variance. <br><br> The two lines of regression coincide i.e. become identical when r = –1 or 1 or in other words, there is a perfect negative or positive correlation between the two variables under discussion. If r = 0 Regression lines are perpendicular to each other <br> The two lines of regression intersect at the point, where x and y are the variables under consideration <br><br> There aggression coefficients remain unchanged due to a shift of origin but change due to a shift of scale. |



### Question 1
**The table below show the height, x, in inches and the pulse rate y, per minute for 9 people find the correlation coefficient and interpret your result.**

| X | 68 | 72 | 65 | 70 | 62 | 75 | 78 | 64 | 68 |
|---|----|----|----|----|----|----|----|----|----|
| Y | 90 | 85 | 88 | 100 | 105 | 98 | 70 | 65 | 72 |

(a) 0.69                                                    (b) 0.56

(c) 0.15                                    (d) None

**Answer: c**

**Explanation:**

You may the use the fact that (double check this for practice)

$\sum x = 622'$ $\sum y = 773$, $\sum x^2 = \sum y^2 = 68.007$, $\sum x\,y = 53,336$

Calculate the numerator

$x^2 = 43206$

$y^2 = 68007$

$n\sum (xy) - (\sum x)(\sum y) = 9.53336 – 622.773 = -782$

$$\sqrt{n\sum x^2 - \left(\sum x\right)^2}\ \sqrt{n\sum y^2 - \left(\sum y\right)^2}$$

$= \sqrt{9.43206 - (622)^2} \cdot \sqrt{9.68007 - (773)^2}$

$= \sqrt{1970}.\sqrt{14534} = 5350.89$

Now, divide to get r $= \dfrac{-782}{5350.89} = 0.15$

## Question 2

**In the previous problem the researcher decides to use data only for adults age 21 to 60 to compute a correlation coefficient what value of r should he expect?**

(a) r = 0                                    (b) r ≠ 0

(c) r < 0                                    (d) r > 0

**Answer: a**

**Explanation:**

r = 0. It is unexpected that mathematical ability and shoe size varies together

## Question 3

**The following data relate to the test scores obtained eight salesmen in an aptitude test and their daily sales in thousands of rupees:**

|        | 1  | 2  | 3  | 4  | 5  | 6  | 7  | 8  |
|--------|----|----|----|----|----|----|----|----|
| Scores | 60 | 55 | 62 | 56 | 62 | 64 | 70 | 54 |
| Sales  | 31 | 28 | 26 | 24 | 30 | 35 | 28 | 24 |

(a) 48                                       (b) 56

(c) 4.5                                       (d) 0.48

**Answer: d**

**Explanation:**

Asb $= \dfrac{24+35}{2} = 30$

| Scores (xi) | Sales in 1000(yi) | ui=xi =62 (3) | Vi = yi – 30 | Ui vi (5)=(3)×( | (6) = (U i)² | (7) (Vi)² |
|-------------|-------------------|---------------|--------------|-----------------|--------------|-----------|

| (i) | (2) | | (4) | 4) | | |
|---|---|---|---|---|---|---|
| 60 | 31 | -2 | 1 | -2 | 4 | 1 |
| 55 | 28 | -7 | -2 | 14 | 49 | 4 |
| 62 | 26 | 0 | 4 | 0 | 0 | 16 |
| 56 | 24 | -6 | -6 | 36 | 36 | 36 |
| 62 | 30 | 0 | 0 | 0 | 0 | 0 |
| 64 | 35 | 2 | 5 | 10 | 4 | 25 |
| 70 | 28 | 8 | -2 | -16 | 64 | 4 |
| 54 | 24 | -8 | 6 | 48 | 64 | 36 |
| Total | _ | -13 | -14 | 90 | 221 | 122 |

Since correlation coefficient remain unchanged due to change of origin we have

## Question 4
**If r = 0.7; and n = 64 find out the probable error of the coefficient of correlation.**

(a) 0.043                                         (b) 0.43

(b) 0.747, 0.657                              (d) 0.7

**Answer: a**

**Explanation:**

r = 0.7: n = 64

P.E. = $0.67745 \times \left[\frac{1-r^2}{\sqrt{n}}\right]$

Probable Error (P.E) = $0.6745 \times \frac{1-(0.7)^2}{\sqrt{64}}$

= $(0.6745) \times (0.06375)$

= 0.043

## Question 5
**Compute the probable error assuming the correlation coefficient of 0.8 from a sample of 25 pairs of item**

(a) 0.0486                                       (b) 0.0456

(c) 0.0567                                       (d) 0.0789

**Answer: a**

**Explanation:**

r = 0.8, n = 25

P.E. = 0. 6745

$\frac{1 - (0.8)^2}{\sqrt{25}}$

= 0.6745 ×0.07 = 0.0486

## Question 6
**Difference between Correlation and Causation**

(a) The variable mutually influence each

   other so that neither can be called the

   causes of other

(b) The correlated variables are influenced

   By one or more variables.

(c) Pure change correlation                     (d) All

**Answer: d**

**Explanation:**

The term correlation should not be misunderstood as causation if correlation exists between two variables it must not be assumed that a changed in one variable is the cause of a change in over variable.


**Question 7**

**For some bivarilate data the following result were obtained the two variable X and Y:**

**x = 53.2, y = 27.9 bvx = - 1.5 bxy = -0.2**

**The most probable value of y when x = 60 is**

(a) 15.6                                        (b) 13.4

(c) 19.7                                        (d) 17.7

**Answer: d**

**Explanation:**

**The regression equation of y of x is:**

y – y = by x (x-x)

= y – 27.9 = (-1.5) (x – 53.2)

Or y = 107.7 – 1.5x

When x = 60 then

y = 107.7 – 1.5 ×60= 17.7


**Question 9**

**If the sum of square of the rank difference in mathematics and physics marks of 10 students is 22, then the coefficient of rank correlation is:**

(a) 0.267                                       (b) 0.867

(c) 0.92                                        (d) None

**Answer: b**

**Explanation:**

Co. efficient of rank correlation

$1 - \dfrac{6 \sum d^2}{n(n^2-1)}$

$1 - \dfrac{6 \times 22}{10(10^2-1)}$

$$1 - \frac{6 \times 2}{10 \times 9}$$

$\frac{13}{15}$ = 0. 867 (Approx.)

## Question 10
**The coefficient of correlation r between x and y when: Cov (x, y) = - 16.5, Var (x) = 2.89, Var (y) = 100is:**

(a) -0.97                                              (b) 0.97

(c) 0.89                                              (d) – 0.89

**Answer: a**

**Explanation:**

r = $\frac{Cov\ (x,y)}{\sigma_x \sigma_y}$

Or r = $\frac{Cov\ (x,y)}{\sqrt{vary\ (x)-\ vary\ (y)}}$

$\frac{-16.5}{\sqrt{2.89\ \times\ 100}}$

= - 0.97

## Question11
**Two random variable have the regression line 3x + 2y = 26 and 6x + y = 31. The coefficient of correlation between x and y is:**

(a) -0.25                                              (b) -0.5

(c) 0.5                                                (d) 0.25

**Answer: c**

**Explanation:**

The regression lines 3x + 2y = 26 and 6x + y = 31 are given

Let first equation be y on x sand second be x only respectively therefore, 3x + 2y = 26

=$\left(\frac{-3}{2}\right)x + 26$

∴byx = -3/2

and 6x + y = 31

= x = $\left(\frac{-1}{6}\right)$x + $\left(\frac{31}{6}\right)$

by = -1/6 Now

$r^2\ byx.bxy$

= $\left(\frac{-3}{2}\right) \times \left(\frac{-1}{6}\right)$

= 0.25

r= 0.5

Hence, our assumption hold true hold and r = 0.5 (- 1 r 1)

Note r is negative because byx and bxy = 0

## Question 12

**The coefficient of correlation between X and Y is 0.6 U and V are two variable defined as U = $\frac{x-3}{2}$, V = $\frac{y-2}{3}$, then the coefficient of correlation between U and V is:**

(a) 0.6                                      (b) 0.8
(c) 0.4                                      (d) 1

**Answer: a**

**Explanation:**

**Since correlation coefficient (Karl Pearson`s) is independence of both scale and origin therefore,**

p(u,v) = p(x,y) = 0.6

it may be noted that if

**μ**, = ax, + b and V; = CY; + d then

r (u,v) = P (x,y) if an and care of same signs

P (x,y) if a and c are of opposite sing

## Question 13

**For the following data the coefficient of rank correlation is:**

| Rank in Botany | 1 | 2 | 3 | 4 | 5 |
|---|---|---|---|---|---|
| Rank in Chemistry | 2 | 3 | 1 | 5 | 4 |

(a) 0.93                                      (b) 0.4
(c) 0.6                                      (d) None

**Answer: c**

**Explanation:**

| S No. | Rank in Botany(xi) | Rank in Chem (yi) | d = (xi) – (yi) | d² |
|---|---|---|---|---|
| 1 | 1 | 2 | -1 | 1 |
| 2 | 2 | 3 | -1 | 1 |
| 3 | 3 | 1 | 2 | 4 |
| 4 | 4 | 5 | -1 | 1 |
| 5 | 5 | 4 | 1 | 1 |
| Total | | | 0 | 8 |

Hence coefficient of rank correction

$1 - \frac{6 \times 8}{5 (5^2 - 1)}$

$S = 1 - \frac{2}{5} = 0.6$

## Question 14

**The following data is given on 450 students for marks is statistics and Economic at a certain examination**

**Mean marks in statistics          = 40**
**Mean marks in economics          = 48**
**S.D. of marks (statistics)          = 12**
**Variance of marks (Economics)      = 256**
**Sum of the products of deviation of**
**Marks from their respective mean      = 42075**
**The average marks in economics of candidates who obtained 50 marks in statistics is:**
(a) 45                                    (b) 54
(c) 54.5                                  (d) 47.5
**Answer: c**
**Explanation:**
Let x = marks statistics
and y = marks in Economics
we know that

$$r_{xy} = \sum \frac{(\sum dx \times dy)}{n \times \sigma_x \sigma_y}$$

Where dx = $x_1 - \bar{x}$ and dy = $y_1 - \bar{y}$

$$r_{xy} = \frac{(42075)}{450 \times 12 \times 16} = 0.49$$

Now regression equation of y on x

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

$$= y - 48 = \times \frac{0.49 \times 16}{12} (x-40)$$

= y = 0.65 then
y = 0.65x – 26 + 48
y = 0.65 x + 22
x = 50

**Question 15**
**For 10 pair of observation, number of concurrent deviation was found to be 4. what is the value of the coefficient deviation?**
(a) $\sqrt{0.2}$                              (b) -$\sqrt{0.2}$
(c) 1/3                                    (d) – 1/3
**Answer: d**
**Explanation:**
Here C = 4, N= 10, So
n = N -1 = 10 -1 = 9

$$rc = \pm\sqrt{\frac{\pm (2_{c-n})}{n}}$$

$$rc \pm\sqrt{\pm \frac{(2 \times 4-9)}{9}}$$

Here $(2_c – n)$ is negative so negative sign is take at both the place so rc = (-1)/3

## Question16
## Karl Pearson`s formula:

(a) $\dfrac{[N \sum XY - (\sum X)[(\sum Y)]}{\sqrt{[N \sum X^2 - (\sum X^2)]}}$       (b) r = $\dfrac{n(\sum xy) - (\sum X)(\sum Y)}{\sqrt{[n \sum x^2 - (\sum x)^2][n \sum y^2 - (\sum y)^2]}}$

(c) Either a or  b                      (d) None

**Answer: b**

**Explanation:**

r = $\dfrac{n(\sum xy) - (\sum X)(\sum Y)}{\sqrt{[n \sum x^2 - (\sum x)^2][n \sum y^2 - (\sum y)^2]}}$

Where

N = number of pairs of scores

∑xy = sum or the products or paired scores

∑x = sum of x scores

∑y = sum of y scores

∑x² = sum of squared x scores

∑y² = sum of squared y scores

## Question 17
## If the coefficient of correlation between x and Y variable is +0. 90 then what will be coefficient of determination?

(a) 0.39                               (b) 0.81

(c) 0.94                               (d) None of these

**Answer: b**

**Explanation:**

If Coeff. of Correlation ® = 0.90

Coeff .of Determination = $r^2$

= $(0.90)^2$

= 0. 81

## Question 18
## The two lines of regression become identical when

(a) 0.4                                (b) 0.6

(c) 0.36                               (d) 0.64

**Answer: c**

**Explanation:**

If r = 0.6

Then Coeff, of determination = $r^2$

= $(0.6)^2$

= 0.36

## Question 19
**There two regression lines passing through**
(a) Represent means                         (b) Represent  S. Ds
(c) (a) and (b)                                   (d) None of these
**Answer: a**
**Explanation:**
The two regressions lines passing through or (Intersect) at their means.

## Question 20
**The regression equation x and y is 3x + 2y = 100 value of $b_{xy}$**

(a) $-\dfrac{2}{3}$                                      (b) $-\dfrac{3}{2}$

(c) $\dfrac{2}{3}$                                        (d) $\dfrac{100}{2}$

**Answer: a**
**Explanation:**
The regression equation of x & y is
3x + 2y = 100
3x + 2y – 100 = 0
$b_{xy} = -\dfrac{Cofficient\ of\ y}{coefficient\ of\ x} = -\dfrac{2}{3}$

## Question 21
**In beauty contest there were 10 competitors of these candidates assigned by two judge A and b the sum of squares of difference of rank is 44. The value of rank correlation is:**
(a) 0.70                              (b) 0.73
(c) 0.80                              (d) 0.60
**Answer: b**
**Explanation:**
Sum of square of difference of rank ($\sum d^2$) = 44
$r_R$=?

$r_R = 1 - 6\dfrac{\sum d^2}{n(n^2 - 1)}$

$1 - \dfrac{6 \times 4}{10\ (10^2 - 1)}$

$1 - \dfrac{6 \times 44}{10 \times 99}$

= 1 – 0.267
= 0. 733
So answer be 0.73

## Question 22
**If two regression lines are x + y = 1 and x – y = 1 then mean value of x and y will be:**

(a) 0 and 1                                    (b) 1 and 1
(c) 1 and 0                                    (d) None
**Answer: c**
**Explanation:**
Given Regression line

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

x = 1 in equation (1) we get

1 – y = 1

y = 0

Mean of x =x = 1

Mean of y = y= 0

Hence 1 and 0

## Question 23
**The coefficient of correlation between x and y is 0.6 If x and y value are multiplied by 1 then the coefficient will be**
(a) 0.6                                        (b) 1-0.6
(c) 1/0.6                                      (d) -0.6
**Answer: a**
**Explanation:**
The coefficient of correlation between X and Y is 0.6 If x and y values are multiplied by 1 then coefficient remains unchanged then are coefficient of correlation will be 0.6

## Question 24
**The coefficient of correlation between the temperature of environment and power consumption is always:**
(a) +ve                                        (b) - ve
(c) 0                                          (d) = 1
**Answer: a**
**Explanation:**
The coefficient of correlation between the temperature of environment and power consumption is always positive.

## Question 25
**Out of the following the one which the regression coefficient is**
(a) Change origin only                         (b) Change of scale only
(c) Change of scale and origin both            (d) Neither a nor b
**Answer: b**
**Explanation:**
By shifting the scale, coefficient of regression is changed.

## Question 26
**When the correlation coefficient r is equal to + 1 all the point in a scatter diagram would be**

(a) On a straight line direct from upper left
   to lower right

(b) On a straight direction from lower left
   upper right

(c ) On a straight line

(d) Both (a) and (b)

**Answer: b**

**Explanation:**

When the correlation coefficient r is equal to + 1 all the point in a scatter diagram on a straight line directed from lower left to upper right.

## Question 27
**In case of ``Insurance companies" profit and the number of claim they have to pay there is -----correlation.**

(a) +ve

(b) -ve

(c) No relation

(d) None

**Answer: b**

**Explanation:**

In case of ``Insurance companies" profit and the number of claim they have to pay there pay there is Negative correlation:

## Question 28
**If the correlation coefficient between two variables is zero then the lines of regression are**

(a) Parallel

(b) Perpendicular

(c) Coincide

(d) None

**Answer: b**

**Explanation:**

If the correlation coefficient between two variables is zero then the lines of regression are perpendicular

## Question 29
**Their competitors in a contest are ranked by two judges in the order 1,2,3 and 2,3,1 respective Calculate the spearman`s rank correlation coefficient.**

(a) -0.5

(b) -0.8

(c) 0.8

(d) 0.5

**Answer: a**

**Explanation:**

| Rank by 1st judge $R_1$ | Rank by 2nd Judge $R_2$ | Diff D= $R_1 - R_2$ | $D^2$ |
|---|---|---|---|

| 1 | 2 | -1 | 1 |
|---|---|---|---|
| 2 | 3 | -1 | 1 |
| 3 | 1 | +2 | 4 |
|   |   |   | $\sum d^2$ = 6 |

Here n = 3

Spearman`s Rank Correlation Coefficient = $1 - 6\frac{\sum d^2}{n(n^2-1)}$

$= 1 - \frac{6\times 6}{3(3^2-1)}$

= -0.5

## Question 30
**The strength (degree) of the correlation between a set of independent variables X and dependent variable Y is measured by**
(a) Coefficient of Correlation      (b) Standard error of estimate
(c) Coefficient Determination      (d) All of these
**Answer: d**
**Explanation:**
The strength (degree) of the correlation between a set of independent variables X and dependent variable Y is measured through
- Coefficient of Correlation
- Standard error of estimate
- Coefficient Determination

## Question 31
**The percent of told variation of the dependent variable Y explained by the set of independent variables X is measured by:**
(a) Coefficient of Correlation      (b) Standard error of estimate
(c) Coefficient Determination      (d) Coefficient of skewness
**Answer: c**
**Explanation:**
The coefficient of determination (denoted by R?) is a key output of regression analysis an $R^2$ of 0 means that the dependent variable cannot be predicted from the independent variable An $R^2$ of 1 means the dependents variable can be predicted without error from the independent variable

## Question 32
**A coefficient of correlation is computed to be -0.95 means that**
(a) The relationship between two variables    (b) The relationship two variables is strong
   is weak                                 and positive
(c) The relationship between two      (d) Correlation coefficient cannot

variables                                            have this
    is strong and but negative                       value
**Answer: c**
**Explanation:**
A coefficient of correlation is compute to be -0.95 means that relationship between two variables is strong and but negative

## Question 33
**Let the coefficient of determination computed to be 0.39 in a problem involving one independent variable and one dependent variable this result means that**

(a) The relationship between two variables
    is negative

(b) The correlation coefficient is 0.39 also

(c) 39% of the total variation is explained by
    the independent variable

(d) 39% of the total variation explained by
    the dependent variables

**Answer: c**
**Explanation:**
The coefficient of determination computed to be 0.39 in a problem involving one independent variable and one dependent variable. 39% of the total variation is explained by the independent variable.

## Question 34
**Relationship between correlation coefficient and coefficient of determination is that:**

(a) The coefficient of determination is the
    square of coefficient of correlation

(b) The coefficient determination is the
    square root of the coefficient of correction

(c) Both are unrelated

(d) Both are equal

**Answer: a**
**Explanation:**
Coefficient of correlation is ``R" value which given in the summary table in the regression output.  R square is called coefficient of determination multiply R times R to get the R value. In other word coefficient of correlation R square or Coeff. of determination shows percentage variation and in y which is explained by all the x variable together higher the better it is always between 0 and 1. it can never be negative – since is a squared value.

It is easy to explain the R square in term of regression it is not so easy to explain the R in terms of regression.

## Question 35
**For a bivariate data two tines of re regression are 40 x – 18y = 214 and 8x – 10y + 66 = 0 then find the value of x and y**

(a) 17 and 13                               (b) 13 and 17

(c) 15 and 17                               (d) None

**Answer: b**

**Explanation:**

Given: 40 x −18y = 214

8x: - 10y = - 66

on solving (1) and (2) we get

x =13 and y = 17

∴ x = 13 and y = 17

## Question 36
**In multiple regression when the global test of significance is rejected we can conclude that:**

(a) All of the net sample regression coefficient are equal to zero

(b) All of the sample regression coefficient are not equal to zero

(c) At least one sample regression coefficient is not equal to zero

(d) The regression equation intersects the y – axis at zero

**Answer: c**

**Explanation:**

In multiple regression when the global test of significance is rejected we can conclude that at least one simple regression coefficient is not equal to zero.

## Question 37
**Correlation Coefficient value lies between**

(a) – 1 and + 1                               (b) 0 and 1

(c) -1 and 0                                   (d) None

**Answer: a**

**Explanation:**

The strength of the linear association between two variables is qualified by the correlation coefficient the correlation coefficient always takes a value between -1 and 1 with 1 or -1 indicating perfect correlation (all point would lie along a straight line in this case)

## Question 38
**In correlation both variables are always**

(a) Random                                    (b) Non Random

(c) Same                                       (d) None

**Answer: a**
**Explanation:**
Complete correlation between two variables is expressed by either + 1 or -1 when one variable increases the correlation is positive when on decrease as the order increases it is negative complete absence of correlation is represented by 0.

## Question 39
**The table below shows the number of absence x, in a calculsis course and the final exam grade y for 7 student find the correlation coefficient.**

| x | 1 | 0 | 2 | 6 | 4 | 3 | 3 |
|---|---|---|---|---|---|---|---|
| y | 95 | 90 | 90 | 55 | 70 | 80 | 85 |

(a) 0.38                                    (b) -0.38
(c) 0.62                                    (d) -0.93
**Answer: d**
**Explanation:**
You may use the facts that (double check this for practice)
$\sum x = 19$, $\sum y = 565$, $\sum x^2 = 75$, $\sum y^2 = 46,775$, $\sum xy = 1,380$
Calculate the numerator:
$n \sum(xy) - (\sum x)(\sum y) = 7.1380 - 19 \times 565 = -1075$
Then the calculate the denominator;
$$\sqrt{[n(x^2) - (x)^2]}\sqrt{[n(y^2) - (y)^2]}$$
$$(525 - 369).[327425 - 319225]$$

## Question 40
**Two regression lines are parallel to each other if their slope is**
(a) Random                                    (b) Non Random
(c) Same                                    (d) None
**Answer: c**
**Explanation:**
When there is a reasonable amount of scatter we can draw to different regression lines depending upon which variable we consider to be the most accurate The first is a line of regression of y on x which can be used to estimate y given x the other is a line of regression of x on y used to estimate x given y Hence two regression lines are parallel to each other if their slope is same

## Question 41
**When regression line passes through the origin then**
(a) Regression coefficient is zero            (b) Correlation is zero
(c) Intercept is zero                         (d) Association is zero
**Answer: c**

**Explanation:**
Prism linear regression analysis fits a straight line through your data and lets you force the line to go through the origin this is useful when you are sure that the line must begin at the origin (x = 0and y = 0) Prism`s nonlinear regression offers the equation line through origin.

**Question 42**
**The table below shows the number of absence, x in a calculate course and the final exam grand, y for 7 student find the correlation coefficient.**

| x | 1 | 0 | 2 | 6 | 4 | 3 | 3 |
|---|---|---|---|---|---|---|---|
| y | 85 | 80 | 70 | 55 | 90 | 90 | 95 |

(a) 0.38                                    (b) 0.6
(c) -0.38                                   (d) 0.62
**Answer: c**
**Explanation:**
There are 7 ordered pairs (x, y) so n = 7 Calculate the needed sums:

| X | Y | $X^2$ | $Y^2$ | xy |
|---|---|---|---|---|
| 1 | 85 | 1 | 7225 | 85 |
| 0 | 80 | 0 | 6400 | 0 |
| 2 | 70 | 4 | 4900 | 140 |
| 6 | 55 | 36 | 3025 | 330 |
| 4 | 90 | 16 | 8100 | 360 |
| 3 | 90 | 9 | 8100 | 270 |
| 3 | 95 | 9 | 9025 | 285 |
| X = 19 | Y = 565 | $X^2$ = 75 | $Y^2$ = 46775 | Xy= 1470 |

**Calculation the numerator:** $n \sum (xy) - (\sum x)(\sum y)$
= 7. 1470 -19. 565
= - 445
Then calculate the denominator:

$$\sqrt{n \sum x^2 - \left(\sum x\right)^2} \sqrt{n \sum y^2 - \left(\sum y\right)^2}$$

= $\sqrt{7.75 - (19)^2}$. $\sqrt{746775 - (565)^2}$
= $\sqrt{164}$. $\sqrt{8200}$ = 1159.66
Now, divide to get r = $\frac{-445}{1159.66}$ = -0.38

**Question 43**
**If two variables oppose each other than the correlation will be**

(a) Positive Correlation          (b) Negative Correlation
(c) Perfect Correlation           (d) None
**Answer: b**
**Explanation:**
A correlation of zero means there is no relationship between the two variables, when there is a negative correlation between two variables as the value of one variable increase the value of the other variable decrease and *vise- versa*

## Question 44
**The time x in years that an employee spent at a company and the employee`s hourly pay, y for 5 employees are listed in the table below. Calculate and interpret the correlation coefficient r. Include a plot of the data in your discussion**
(a) 0.38                          (b) -097
(c) 0.62                          (d) None
**Answer: d**
**Explanation:**

| X | Y | X² | Y² | xy |
|---|---|-----|------|------|
| 5 | 25 | 25 | 625 | 125 |
| 3 | 20 | 9 | 400 | 60 |
| 4 | 21 | 16 | 441 | 84 |
| 10 | 35 | 100 | 1225 | 350 |
| 15 | 38 | 225 | 1444 | 570 |
| X = 37 | Y=139 | X² 375 | Y² = 4135 | XY = 1189 |

Hint Calculate the numerator:

$n \sum (xy) - (\sum x)(\sum y)$ = 5. 1189− 37. 139= 802

Then calculate the denominator

$$= \sqrt{n \sum x^2 - (\sum x)^2} \sqrt{n \sum y^2 - (\sum y)^2}$$

$$=\sqrt{5 \times 375 - (37)^2}.\sqrt{5.4135 - (37)^2}$$
$$= \sqrt{506}.\sqrt{1354} = 827.72$$
Now divide to get r = $\frac{802}{827.72}$ = 0.97

## Question45
**Identify the true correlation**
(a) -1 & 1                        (b) -1 & 0
(c) 0& 1                          (d) All are true
**Answer: a**
**Explanation:**

This will always be a number between -1 and 1 (inclusive)

- If is close to 1 we say that the variables are positively correlated. This means there is likely a strong linear relationship between the two variables with a positive slope.
- If is close to-1 we say that the variable are negatively correlated this means there is likely a strong linear relationship between the two variable with a negative slope.
- If r is close to 0, we say that the variables are not correlation this means that variables may still be related some other way.

## Question 46
**A researcher carefully computes the correlation coefficient between two variables and gets r = 1.23 what does this value mean?**
(a) $-1 \leq r \leq 1$                                   (b) $-1 \geq r \geq 1$
(c) Both                                                   (d) None
**Answer: a**
**Explanation:**
A error was made all correlation coefficient $-1 \leq r \leq 1$

## Question 47
**If $R^2$ is zero that is no collinearly / Multi collinearity the variance inflation factor (VIF) will be**
(a) 1                                                      (b) 2
(c) 3                                                      (d) None
**Answer: a**
**Explanation:**
$VIF = \frac{1}{1-R^2}$

## Question 48
**If the equation of regression line is y = 5, then what result will you take out from it?**
(a) The line is parallel to x- axis           (b)_ The line passes through (5.0)
(c) The line passes through origin          (d) The line passes through origin
**Answer: a**
**Explanation:**
y = k for one value of y there are infinite value of x

## Question 49
**The method of least squares finds the best fit line that the error between observed and estimated point on the line**

(a) Reduces to zero                          (b) Approaches to infinity
(c) Minimize                                 (d) Maximize
**Answer: c**
**Explanation:**
The method f least squares finds the best fit line that minimize the error between observed and estimate points on the line.

### Question50
**A regression model may be:**
(a) Linear                                   (b) Non - linear
(c) Both (a) and (b)                         (d) Neither (a) and (b)
**Answer: a**
**Explanation:**
In the regression it appears on the left side of the equal sign, while your can use regression to predict the dependent variable your always start with a set of known y value and use these be build (or to calibrate) the regression model may be linear and nonlinear both

# PAST EXAMINATION QUESTION

## MAY 2018

### Question1
**If the model points are a scatter diagram is evenly distributed then the correlation is:**
(a) 0                                        (b)  -ve
(c) +ve                                      (d) a or b
**Answer: a**
**Explanation:**
In the case of a positive correlation, the plotted points are distributed from lower left corner to upper right corner (in the general pattern of being evenly spread about a straight line with a positive slope), and in the case of a negative correlation, the plotted points are spread out about a straight line of a ...

### Question2
**If the plotted points in a scatter are evenly distributed, then the correlations zero.**

**The covariance between variable is**
  (a) Strictly positive                              (b) Strictly negative
  (c) Always zero.                                    (d) Either positive or negative zero.
**Answer: d**
**Explanation:**
The Co- variance between two variables is either positive or negative or zero.

## Question3
**The coefficient of determination is defined by the formula.**
  (a) $r^2 = \dfrac{1 - unexpalained\ variance}{Total\ variance}$           (b) $r^2 = \dfrac{expalained\ variance}{Total\ variance}$
  (c) Both (a) and (b)                               (d) None
**Answer: c**
**Explanation:**
The coefficient of determination
$r^2 = \dfrac{1 - unexpalained\ variance}{Total\ variance}$
$r^2 = \dfrac{expalained\ variance}{Total\ variance}$

## Question4
**In the method of concurrent deviations only the directions of change (positive direction/ Negative direction) in the variable are taken into account for calculation of**
  (a) Coefficient of SD                              (b) Coefficient of regression
  (c) Coefficient of correlation                     (d) None
**Answer: c**
**Explanation:**
The method of concurrent deviation only the direction of change (positive direction/ Negative direction) in the variables are taken into account for calculation of coefficient of correlation

## Question5
**Correlation coefficient is----of the units of measurement**
  (a) Dependent                                      (b) Independent
  (c) Both                                           (d) None
**Answer: b**
**Explanation:**
Correlation coefficient is Independent of the units of measurement.

## Question 6
**In case speed of an automatic and the distance required to stop the car after applying correlation is**

(a) +ve                                        (b) -ve
(c) 0                                          (d) None
**Answer: a**
**Explanation:**
In case speed of an automatic and the distance required to stop the car after applying correlation is positive

## Question7
**A relationship $r^2 = 1 - \frac{500}{300}$ is possible**
(a) True                                       (b) False
(c) Both                                       (d) None
**Answer: a**
**Explanation:**
$r^2 = 1 - \frac{500}{300}$ is possible
$r^2 = 1 - \frac{-200}{300}$ is not possible
So it is true

## Question8
**Rank correlation coefficient lies between**
(a) -1 to + 1                                  (b) 0 to 1
(c) -1 to 0                                    (d) Both
**Answer: a**
**Explanation:**
Rank correlation coefficient lies between -1 to + 1 inclusive of both value.

# NOV 2018

## Question1
**The two lines of regression intersect at the point**
(a) Mean                                       (b) Mode
(c) Median                                     (d) None
**Answer: a**
**Explanation:**
The two lines of regression intersect at the point is Mean.

## Question2
**If the two line of regression are x + 2y -5 = 0 and 0, then the regression line of y on x s:**
(a) x + 2y -5 = 0                              (b) 2x +3y -8 = 0
(c) x + 2y = 0                                 (d) 2x + 3y = 0

**Answer: a**
**Explanation:**
Given two regression line are
x + 2y 5 = 0 and 2 x +3y -8 =0
byx = $\frac{-coff.of\ x}{coff\ of\ y}$ = $\frac{-1}{2}$ and bxy $\frac{-coff.of\ y}{coff\ of\ x}$ = $\frac{-3}{2}$
Here, bxy× $bxy \leq 1$ which is satisfied
So.1st equation x+2y-5 =0 is the regression equation y on x

**Question3**
**If the two regression line lines are 3x=y and 8y = 6x the value of correlation coefficient is:**
 (a) 0.5                                          (b) -0.5
 (c) 0.75                                         (d) -0.80
**Answer: a**
**Explanation:**
**Given**
Regression line
3x=y and 8y=6x
3x-y =0 and 6x-8y =0
bxy =$\frac{-coff.of\ y}{coff\ of\ x}$ $and\ byx\ \frac{-coff.of\ x}{coff\ of\ y}$
$\frac{-(-1)}{3}$ = $\frac{-6}{-8}$ = $\frac{3}{4}$
bxy = $\frac{1}{3}$bxy = 3/4
Coff. of correlation is given by
r = $\pm\sqrt{byx \times bxy}$
 = $\pm\sqrt{\frac{3}{4} \times \frac{1}{3}}$
 = $+\frac{1}{4}$
+1/2
 = 0.5

**Question4**
**The regression coefficient is independent of the change of**
 (a) Scale                                       (b) Origin
 (c) Scale and Origin both                        (d) None
**Answer: b**
**Explanation:**
The regression coefficient is independent of the change of `Origin'

**Question5**

**If the correlation coefficient between the variable X and Y is 0.5, then the correlation between the variable 2x-4 and 3-2y**

**Answer:**

2x-u -4 =0        and    2y + v-3 = 0

b= $\frac{-coff.of\ u}{coff\ of\ x}$ and  $d = \frac{-coff.of\ v}{coff\ of\ y}$

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

Here, b and d both have different sign so $r_{uv}$ = $-r_{xy}$

= -0.5

# MAY 2019

## Question1
**Given that**

| x | -3 | -3/2 | 0 | 3/2 | 3 |
|---|----|------|---|-----|---|
| y | 9  | 9/4  | 0 | 9/4 | 9 |

 (a) Positive                          (b) Zero
 (c) Negative                         (d) None

**Answer: b**

**Explanation:**

| x | y | $X^2$ | $Y^2$ | xy |
|---|---|-------|-------|-----|
| -3 | 9 | 9 | 81 | -27 |
| -3/2 | 9/4 | 9/4 | 81/16 | -27/8 |
| 0 | 0 | 0 | 0 | 0 |
| 3/2 | 9/4 | 9/4 | 81/16 | 27/8 |
| 3 | 9 | 9 | 81 | 27 |
| 0 | $\frac{90}{4}$ $=\frac{45}{2}$ | $\frac{90}{4}$ $=\frac{45}{2}$ | $\frac{2754}{16}$ $\frac{1377}{8}$ | 0 |
| | | Hence Answer 0 | | |

## Question2
**Given the following series:**

| x | 10 | 13 | 12 | 15 | 8 | 15 |
|---|----|----|----|----|---|----|
| y | 12 | 6  | 18 | 16 | 7 | 18 |

The rank correction coefficient r =

 (a) $1 - \frac{6\sum d^3 + \sum_{i=d}^{2}\frac{m_2(m_2^3-1)}{12}}{m(n^2-1)}$                (b) $1 - \frac{6\left[\sum d^2 + \sum_{i=1}^{3}\frac{m_2(m_1^2-1)}{12}\right]}{n(n^2-1)}$

 (c) $1 - \frac{6\left[\sum d^2 + \sum_{i=1}^{3} m_2(m_1^9-1)\right]}{n(n^2-1)}$                (d) None

**Answer: b**
**Explanation:**

$$1 - \frac{6\left[\sum d^2 + \sum_{i=1}^3 \frac{m_2\left(m_1^2-1\right)}{12}\right]}{n\left(n^2-1\right)}$$

## Question3
**If the regression line of y on x is given by y = x+ 2 and Karlperson`s coefficient of correlation is 0.5 then $\frac{\sigma y^2}{\sigma x^2}$**

 (a) 3                                         (b) 2
 (c) 4                                         (d) None

**Answer: c**
**Explanation:**
y on x = > y = x + 2
R = 0.5
byx = r × $\frac{\sigma y}{\sigma x}$
byx = $\frac{2}{0.5}$

## Question4
**A .M. OF regression coefficient is:**
 (a) Equal to r                               (b) Great than or equal to r
 (c) Half of r                                (d) None of these

**Answer: b**
**Explanation:**
Regression coefficient is a statistical measure of the average functional relationship between two or more variable In regression analysis one variable is considered as dependent and other as independent, Thus it measure the degree of dependence of one variable on the order (s)

## Question5
**If the two regression lines are x + y = 1 and x – y = 1 then $\overline{x}$ and $\overline{y}$ are**
 (a) 1,0                                       (b) 0,1
 (c) 1,1                                       (d) None of these

**Answer: d**
**Explanation:**
Consider x-y =1 as equation (1) as equation (2)
Now add both (1) and (2)
You get 2x= 2 i.e. x=1
Now put x=1 in either of equation (1) or (2)
You get y=0

**Question6**
**Coefficient of correlation between X and Y is 0.6 if both X and Y are multiplied then resultant coefficient of correlation:**
(a) 0.6                                          (b) 1/0.6
(c) Both                                        (d) None of these
**Answer: a**
**Explanation:**
Value of Correlation will -0.6 as if you multiply x & y by -1 then this will show negative sign and as per property bxy , byx& r sigh should be equal / same

# NOV 2019

**Question1**
**If two of regression are x + 2y -5 = 0 and 2x + 3y -8 = 0 So x + 2y -5 0 is**
(a) y on x                                       (b) x on y
(c) Both                                         (d) None
**Answer: a**
**Explanation:**
x + 2y -5 =0 – Eq 1 2x + 3y -8 = 0- Eq2
Let Eq 1 be y on x from Eq$^2$
byx $=\frac{-cofficient\ of\ x}{cofficient\ of\ y}$  bxy $= \frac{-cofficient\ of\ x}{cofficient\ of\ y}$
byx $= \left(\frac{-1}{2}\right) \times \left(\frac{-3}{2}\right) = \frac{3}{4}$
So, byx × bxy< 1
So, ×+ 2y -5 = 0 is y on ×
and 2x + 3y -8 = 0 is x on y

**Question2**
**Find the coefficient of regression.**
**2x + 3y = 2**
**4x + 3y = 4**
(a) 0.5                                           (b) -0.5
(c) 0.25                                          (d) -0.25
**Answer: b**
**Explanation:**
2x + 3y = 2 – Eq 1 ax + 3y = 4 – Eq 2
Let Eq1 be y on x
From Eq1
byx $=\frac{-cofficient\ of\ x}{cofficient\ of\ y} = \frac{-2}{3}$
From Eq2

bxy = $\frac{-cofficient\ of\ x}{cofficient\ of\ y}$ = $\frac{-3}{4}$

So above assumption hold true.

r= $\pm\sqrt{b_{yx}\ x b_{xy}}$

r= $\pm\sqrt{\left(\frac{-2}{3}\right) x \left(\frac{-3}{4}\right)}$

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

r= -0.5

## Question3

## What is the coefficient of correlation from the following data?

| X: | 1 | 2 | 3 | 4 | 5 |
|---|---|---|---|---|---|
| Y: | 5 | 4 | 3 | 2 | 6 |

 (a) 0                                        (b) -0.75
 (c) -0.85                                    (d) 0.82

**Answer: a**

**Explanation:**

| X | Y | Xy |
|---|---|---|
| 1 | 5 | 5 |
| 2 | 4 | 8 |
| 3 | 3 | 9 |
| 4 | 2 | 8 |
| 5 | 6 | 30 |
| $\sum x = 15$ | $\sum x = 20$ | $\sum xy = 60$ |

cov (x,y) = $\frac{\sum xy}{n}$ - $\overline{x}, \overline{y}$

= $\frac{60}{5}$ - $\left(\frac{15}{5}\right)$ × $\left(\frac{20}{5}\right)$

= 12 – 12

cov (x,y) = 0

r = $\frac{cov(x,y)}{\sigma_x \sigma_x}$

r = 0

## Question4

If the plotted points in a scatter diagram lie from upper left to lower right, then correlation is:
 (a) Positive                                 (b) Negative
 (c) Zero                                      (d) None of these

**Answer: b**

**Explanation:**
If the points in a scatter diagram lie from upper to left lower right them correlation us negative.

# DEC 2020

## Question 1
**Which of the following is spurious correlation?**

(a) Correlation between two variables having no causal relationship

(b) Negative Correlation

(c) Bad relation between two variables

(d) Very low correlation between two variables

**Answer: a**
**Explanation:**
Correlation between two variables having no causal relationship

## Question 2
**Scatter diagram does not help us to**

(a) Find the type of correlation

(b) Identify whether variables correlated or not

(c) Determine the linear (or) non – linear correlation

(d) Find the numerical value of correlation coefficient

**Answer: d**
**Explanation:**
To Find the numerical value of correlation coefficient

## Question 3
**The Covariance between two variables is**

(a) Strictly Positive

(b) Strictly Negative

(c) Always Zero

(d) Either positive (or) Negative (or) Zero

**Answer: d**
**Explanation:**
Covariance can be positive, zero, or negative. ... If X and Y are independent variables, then their covariance is 0: Cov (X, Y ) = E(XY ) − μXμY = E(X)E(Y ) − μXμY = 0 The converse, however, is not always true. Cov (X, Y) can be 0 for variables that are not inde- pendent. Hence, either positive (or) Negative (or) Zero

# JAN 2021

## Question 1
**For the set of observations {(1, 2), (2, 5), (3, 7), (4, 8), (5, 10)}, the value of Karl-**

**person's coefficient is approximately given by**
 (a) 0.755                                      (b) 0.655
 (c) 0.525                                      (d) 0.985
**Answer:  d**
**Explanation:**

| x | y | xy | $x^2$ | $y^2$ |
|---|---|---|---|---|
| 1 | 2 | 2 | 1 | 4 |
| 2 | 5 | 10 | 4 | 25 |
| 3 | 7 | 21 | 9 | 49 |
| 4 | 8 | 32 | 16 | 64 |
| 5 | 10 | 50 | 25 | 100 |
| **Total 15** | **32** | **115** | **55** | **242** |

$$\frac{5 \times 115 - 15 \times 32}{\sqrt{5 \times 55 - (15)^2}\sqrt{5 \times 242 - (32)^2}}$$

$$\frac{95}{\sqrt{50}\sqrt{186}}$$

$$\frac{95}{7.07106 \times 13.638} = 0.985105$$

**Question 2**
**The coefficient of correlation between x and y is 0.5, the covariance is 16, and the standard deviation of x is 4. Then the standard deviation of y is**
 (a) 4                                          (b) 8
 (c) 16                                         (d) 64
**Answer:  b**
**Explanation:**

$$r_{xy} = \frac{cov(x, y)}{\sqrt{Var(x). Var(y)}}$$

$$r_{xy} = 8$$

**Question 3**
**The intersecting point of the two regression lines: y on x and x on y is**
 (a) $(0, 0)$                                   (b) $(\bar{x}, \bar{y})$
 (c) $(b_{yx}, b_{xy})$                         (d) (1, 1)
**Answer:  b**
**Explanation:**
Properties of Regression Lines There are two lines of regression. Both these lines are known to intersect at a specific point $(\bar{x}, \bar{y})$Here the variables under consideration are x and y.

## Question 4
**Given that the variance of x is equal to the square of standard deviation by and the regression line of y on x is y = 40 + 0.5 (x – 30). Then regression line of x on y is**

(a) y = 40 + 4(x – 30)                    (b) y = 40 + (x – 30)
(c) y  = 40 +2(x – 30)                    (d) x = 30 + 2 + 2(x – 40)

**Answer:  d**
**Explanation:**

## Question 5
**The regression coefficient remain unchanged due to**

(a) A shift of scale                    (b) A shift of origin

(c) Replacing x – values by $\frac{1}{x}$            (d) Replacing y values by $\frac{1}{y}$

**Answer:  b**
**Explanation:**
The regression coefficient remain unchanged due to A shift of origin
By properties of regression line we have, The regression coefficients remain unchanged due to a shift of origin but change due to a shift of scale.

# JULY 2021

## Question 1
**If the sum of the product of the deviation of and Y from their means is zero correlation coefficient between X and Y is:**

(a) Zero                                (b) Positive
(c) Negative                            (d) 10

**Answer: Options (a)**
**Explanation:**
Given: sum of the product of deviations of x and y series from their mean is zero,
To Find:  the coefficient of correlation
r = coefficient of correlation
r = Sxy  / (Sx . Sy)
Correlation coefficient = cov (x,y)/ (std deviation (x) × std deviation (y))
product of deviations of x and y series from their mean is zero
=> Sxy = 0
=> r = 0
Coefficient of correlation = 0
Ans: If the sum of the product of deviations of x and y series from their mean is zero, then the coefficient of correlation will be  ZERO

**Question 2**
**If the slope of the regression line is calculated to be 5.5 and the intercept 15 then the value of Y when X is 6 is**
  (a) 88                                      (b) 48
  (c) 18                                      (d) 78
**Answer: Options (b)**
**Explanation:**
The value of Y when X is 6
a+bX
15+5.5(6)
Ans: 48

**Question 3**
**If y = 9x and X = 0.01Y, then r is equal to:**
  (a) -0.1                                    (b) 0.1
  (c) 0.3                                     (d) -0.3
**Answer: Options (c)**

**Question 4**
**The straight – line graph of the linear equation Y = a +b X, slope is horizontal if:**
  (a) b =1                                    (b) b≠ 0
  (c) b =0                                    (d) a = b ≠ 0
**Answer: Options (c)**

**Question 5**
**If $b_{xy}$ = -1.6 and $b_{xy}$ = -0.4, then $r_{xy}$ will be**
  (a) 0.4                                     (b) -0.8
  (c) 0.64                                    (d) 0.8
**Answer: Options (b)**

# DEC 2021

**Question 1**
**I f the data points of (X,Y) series on a scatter diagram lie along a straight line that goes downwards as X-values move from left to right, then the data exhibit -------------------correlation.**
  (a) Direct                                  (b) Imperfect indirect
  (c) Indirect                                (d) Imperfect direct
**Answer: c**
**Explanation:**

This is a Perfect Negative correlation, or indirect correlation.

## Question 2
**For any two variables x and y the regression equations are given as 2x + 5y - 9 = 0 and 3x - y - 5 = 0. What are the A.M. of x and y?**
 (a) 2, 1                                        (b) 1, 2
 (c) 4, 2                                        (d) 2, 4
**Answer: a**
**Explanation:**
The regression lines intersect at the means of x and y. Therefore, the common point of intersection of these two lines will give the means of X and y. This means that the means of x and y will satisfy these two equations simultaneously.
We can either solve these two equations simultaneously or find out the values of x and y, which will give uS our means; or, we can simply try the options.
Option (a) - 2, 1
Putting the value of × = 2, and y = 1 in the equation
2x + 5y - 9 = 0, we get
LHS = 2(2) + 5(1) - 9 = 0 = RHS
Putting the value of x = 2, and y = in the equation
3x - y - 5 = 0, we get
LHS = 3(2) - 1 - 5 = 0 = RHS
Therefore, option (a) is the answer.

## Question 3
**The intersecting point of two regression lines falls at X-axis. If the mean of X-values is 16, the standard deviations of X and Y are respectively, 3 and 4, then the mean of Y-values is**
 (a) 16/3                                        (b) 4
 (c) 0                                           (d) 1
**Answer: c**
**Explanation:**
The intersecting point of two regression lines gives the means of x and y. Since the point of intersection falls on the x-axis, the value of y is 0. Therefore, the mean of y-values is zero.

## Question 4

**The regression coefficients remain unchanged due to**
 (a) Shift of origin                        (b) Shift of scale
 (c) Always                                  (d) Never
**Answer: a**
**Explanation:**
The regression coefficient remain uncharged due to shift of origin.

# JUNE 2022

**Question 1**
**If Coefficient of correlation for 3X+ 4y = 6 is 0.5. Find the coefficient of correlation for of 3u+ 9v = 7 for u and v.**
 (a) -(0.5)                                  (b) (0.5)
 (c) ‡ 0.5                                   (d) 0.25
**Answer: b**
**Explanation:**
We know that shift of scale coefficient of (under consideration) then $r_{xy}$ = $r_{uv}$ = 0.5 correlation is change

**Question 2**
**Karl Pearson Correlation Coefficient method is used for -**
 (a) Any data                               (b) Scattered data
 (c) Grouped data                           (d) Ungrouped data
**Answer: d**
**Explanation:**
Karl Pearson Correlation Coefficient method is used for ungrouped data.

**Question 3**
**If the plotted point in a scatter diagram lie from lower left to upper right then correction is:**
 (a) Positive                               (b) Negative
 (c) Perfectively Negative                  (d) Zero
**Answer: a**
**Explanation:**
If the plotted point in a scatter diagram lie from lower left to upper right then it is said to be positive correlation.

**Question 4**
**If concurrent coefficient is $\frac{1}{\sqrt{3}}$. If sum of deviation is n 6 for n pairs of data?**
 (a) 9                                       (b) 8
 (c) 10                                      (d) 11
**Answer: c**

**Explanation:**

Given $r_c = \frac{1}{\sqrt{3}}$, n= ?

C= 6

Coeff of concurrent deviation

$$r_c = \pm \sqrt{\frac{2c - m}{m}}$$

$$\frac{1}{\sqrt{3}} = \pm \sqrt{\frac{2 \times 6 - m}{m}}$$

On squaring both side

$$\left[\frac{1}{\sqrt{3}}\right]^2 = \left[\pm \sqrt{\frac{12 - m}{m}}\right]^2$$

$$\frac{1}{3} = \frac{12 - m}{m}$$

m = 36 - 3m

m + 3m = 36

4m = 36

m= $\frac{36}{4}$ = 9

n= m +1=9+1=10

## Question 5
**Which of the following is used he find correlation between two qualitative characteristics**

 (a) Karl Pearson                         (b) Spearman rank correlation
 (c) Concurrent deviation                 (d) Scatter diagram

**Answer: b**

**Explanation:**

Spearman's rank correlation coefficient is used to find correlation between two qualititative characteristics.

## Question 6
**Scattered diagram is used the plot**

 (a) Quantitative data                    (b) Qualitative data
 (c) Discrete data                        (d) Continuous data

**Answer: a**

**Explanation:**

Scattered diagram is used to plot quantitative data.