# Question Bank for Statistics part-II

## Chapter # 10

1- Write p.d.f. of standard normal distribution.

Ans: The p.d.f. (probability density function) of the standard normal distribution $N \sim Z(0,1)$ is:

$$f(z) = \frac{1}{\sqrt{2\pi}} e^{-\frac{z^2}{2}} \ , -\infty < z < +\infty$$

2- Find the value of maximum ordinate for a normal distribution with mean 25 and variance 16.

Variance = 16, S.D. = 4, $f(x) = \frac{1}{\sigma\sqrt{2\pi}} = \frac{1}{4\sqrt{2\left(\frac{22}{7}\right)}} = 0.0997$

3- If $Z \sim N(0,1)$, then find the value of "a" such that $P(Z > a) = 0.025$
P(z<a) = 1-0.025=0.975, ϕ(a)=0.975, a=$\phi^{-1}$(0.975) = 1.96 (using table 10a)

4- If $X \sim N(40,49)$, find median and standard deviation?

$X \sim N(\mu, \sigma^2)$, μ=Median=40, S.D.= $\sqrt{49} = 7$

5- In a normal probability distribution, what are the first four moments about mean?
$\mu_{1=}0$ , $\mu_{2=}\sigma^2$ , $\mu_{3=}0$ , $\mu_{4=}3\sigma^4$

6- In a normal distribution $Q_1 = 8$ , $Q_3 = 17$. Find the value of mean and mode.

Question Bank part-I With Answers.docx

$\mu = \frac{(Q_1+Q_3)}{2} = \frac{(8+17)}{2} = \frac{25}{2} = 12.5$ = Mode

7- If $Z \sim N(0,1)$, then find $P[|z| < 1.64]$.
$P[|z| < 1.64] = 2\ \phi\ (1.64) - 1 = 2\ (0.9495) - 1 = 0.899$ (using table 9)

8- Write four properties of standard normal distribution.

i. The distribution is bell shaped, unimodal and symmetrical.
ii. In a normal probability distribution, the first four moments about mean are:
$\mu_{1=}0$, $\mu_{2=}\sigma^2$, $\mu_{3=}0$ and $\mu_{4=}3\sigma^4$
iii. The total area under the normal curve is unity.
iv. In a normal probability distribution: $Q_1 = \mu - 0.6745\ \sigma$ and $Q_3 = \mu + 0.6745\ \sigma$

9- Find the ordinate of the standard normal curve at z = -0.84
$Z = \frac{x-\mu}{\sigma}$ , $0.84 = \frac{x-0}{1}$ , x = 0.84.

10- Define standard normal distribution.
If the random variable X has a normal distribution with mean $\mu$ and variance $\sigma^2$, then the random variable $z = (x - \mu)/\sigma$ has a standard normal distribution with mean 0 and variance 1.

11- Define the point of inflexion in a normal distribution.
The points of inflexion of the normal probability density function are equidistant from mean $\mu$, they are at $x = \mu - \sigma$ and $x = \mu + \sigma$.

12- In a normal distribution $\mu = 24$ and $\sigma = 4$. Find the fourth moment about mean?

$\mu_{4=}3\sigma^4 = 3\ (4)^4 = 3(256) = 768$

13- Find the standard deviation, if Q.D. = 3.3725 for a normal distribution.
Q.D.= 0.6745 $\sigma$, then $\sigma$= Q.D. /0.6745 = 3.3725/0.6745 = 5

14- What is the relationship between mean, median and mode in a normal distribution?
The normal probability distribution is a symmetrical distribution. Thus the mean, median and mode coincide at $\mu$.

Mean = Median = Mode = $\mu$

15- In a normal distribution $\mu = 103$ and $Q_3 = 171.094$, find the standard deviation.
$Q_3 = \mu + 0.6745\ \sigma$
$171.094 = 103 + 0.6745\sigma$,
$$\sigma = \frac{171.094 - 103}{0.6745} = 100.95$$

16- If $Z \sim N(0\ ,1)$, find $P(Z < -1.645)$?
$P(Z < -1.645) = .05$ (using table 9)

17- Define normal frequency distribution.
It is a continuous symmetrical, bell shaped distribution. Its mean, mode and median are all the same; and both the tails of the normal curve are infinitely long.

18- In a normal distribution $\mu_4 = 768$, find $\mu_2$.
Using $\mu_4 = 3\sigma^4$, $\sigma^4 = \frac{768}{3}$ = 256, taking square root $\sigma^2$= 16

19- If $X \sim N(25{,}25)$, find the value of maximum ordinate?
$\sigma^2$ = 25, $\sigma$ = 5, $f(x) = \frac{1}{\sigma\sqrt{2\pi}} = \frac{1}{5\sqrt{2\left(\frac{22}{7}\right)}}$ =0.079

20- In a normal distribution the $Q_1 = 18\ and\ Q_3 = 26$. Find its mean and standard deviation.
$$\mu = \frac{(Q_1 + Q_3)}{2} = \frac{(18 + 26)}{2} = \frac{44}{2} = 22$$
$\sigma = \frac{(Q_3 - Q_1)}{1.349}$ = 8/1.349 = 5.93

21- Define normal distribution?
The normal distribution is the limiting from of the binomial distribution. when the number of trials in a binomial distribution is large and neither *p* (the probability of success) nor *q* (the probability of failure) is very small, then this distribution approaches a continuous distribution known as the normal distribution.

22- Write down the equation of normal curve.
$f(x) = \frac{N}{\sigma\sqrt{2\pi}} e^{-\frac{1}{2}\left(\frac{x-\mu}{\sigma}\right)^2}$, where $-\infty < x < +\infty$

23- What is range of normal variable?
$-\infty < x < +\infty$

24- What is the shape of the normal curve?
Continuous, unimodal, bell shaped and symmetrical

25- In a normal distribution what are the values of $\mu_2 and\ \mu_3$?
$\mu_2 = \sigma^2$ , $\mu_3 = 0$

26- Write down the formulas for mean deviation, lower and upper quartiles in normal distribution.
M.D. = 0.7979$\sigma$, $Q_1 = \mu - 0.6745\ \sigma\ and Q_3 = \mu + 0.6745\ \sigma$

27- If $Z \sim N(0\ ,1)$, find $P(Z < -1.96)$?

$P(Z < -1.96)$ = 0.025. (using table 9)

## Chapter # 11

1- Define sample and sampling?
**Sample** is a part of population which is selected with the expectation that it will represent the characteristics of the population.
**Sampling** is a procedure of selecting a representative sample from a given population.

2- Explain sampling with replacement and without replacement?
Sampling is said to be **with replacement** when the unit selected at random is returned to the population before the next unit is selected.
Sampling is said to be **without replacement** when the unit selected at random is not returned to the population before the next unit is selected.

3- Define sampling distribution of means.
The sampling distribution of sample means is the probability distribution of the means of all possible simple random samples of 'n' observation that can be drawn from a given population with mean $\boldsymbol{\mu}$and variance $\boldsymbol{\sigma^2}$.

4- Find $\sigma_{\bar{x}}$if $\sigma^2 = 2.25$ and n = 4.

$$\sigma_{\bar{x}} = \frac{\sigma}{\sqrt{n}} = \frac{1.5}{2} = 0.75$$

5- If $n_1 = n_2 = 2$ and $P_1 = \frac{1}{3}$, $P_2 = \frac{2}{3}$, find $E(\hat{P}_1 - \hat{P}_2)$ and $\sigma_{\hat{P}_1-\hat{P}_2}$? Here $P_1\ and\ P_2$ are population proportions.

$E(\hat{P}_1 - \hat{P}_2)$ = $P_1 - P_2$ = 1/3 – 2/3 = - 1/3.

$\sigma_{\hat{P}_1-\hat{P}_2} = \sqrt{\frac{P_1(1-P_2)}{n_1} + \frac{P_1(1-P_2)}{n_1}} = \sqrt{\frac{1/3(1-1/3)}{2} + \frac{2/3(1-2/3)}{2}} = \sqrt{2/9} = 0.471$

6- What is population?
It is the totality of the observations made on all the objects under consideration, possessing some common characteristics.

7- What is non-sampling error?
The errors that are caused by sampling the wrong population of interest. These errors includes things such as faulty interviewing, incomplete or inaccurate responses, mistakes in recording or coding the data, errors made in processing the results, and the like.

8- Explain the properties of the sampling distribution of a mean?

(i) $\boldsymbol{\mu_{\bar{x}} = \mu}$

(ii) $\sigma_{\bar{x}} = \frac{\sigma}{\sqrt{n}}$, when the sampling is done with replacement.

(iii) $\sigma_{\bar{x}} = \frac{\sigma}{\sqrt{n}}\sqrt{\frac{N-n}{N-1}}$ , when the sampling is done without replacement.

9- Distinguish between finite and infinite population.
A population is said to be **finite** if it includes a limited number of elementary unites.
A population is said to be **infinite** if it consists of unlimited number of elementary unites.

10- Given $N_1 = 3, n_1 = 2, and\ N_2 = 4, n_2 = 2$. If ${\sigma_1}^2 = \frac{8}{3}$and ${\sigma_2}^2 = \frac{5}{4}$. Find $var(\bar{X}_1 - \bar{X}_2)$ when sampling is done without replacement?

$$Var(\bar{X}_1 - \bar{X}_2) = \frac{\sigma_1^2}{n_1}\frac{N_1 - n_1}{N_1 - 1} + \frac{\sigma_2^2}{n_2}\frac{N_2 - n_2}{N_2 - 1} = \frac{8/3}{2}\left(\frac{3-2}{3-1}\right) + \frac{5/4}{2}\left(\frac{4-2}{4-1}\right) = \frac{13}{12}$$

11- Define sample survey?
The collection of information from a part of population is called a sample survey.

12- Define meaning of census?
The collection of information from all elements in a population is called census.

13- Define sampling?
**Sampling** is a procedure of selecting a representative sample from a given population.

14- What do you mean by non-probability sampling?
It is a procedure in which we cannot assign to an element in population, the probability of its being included in the sample.

15- If $\sigma = 4$ , $N = 2$ , $n = 10$, find $\sigma_{\bar{x}}$, if sampling is done with replacement?
$\sigma_{\bar{x}} = \frac{\sigma}{\sqrt{n}} = \frac{4}{\sqrt{10}} = 1.265$

16- Define random digits in sampling?
Random digits consist of a sequence of digits designed to represent the result of a simple random sampling with replacement from 0, 1, 2, …., 9.

17- What is meant by parameter?
The numerical quantities such as mean, S.D., etc., calculated from population data are called parameters. These are usually denoted by Greek letters.

18- Define sampling error?
It is the difference between the value of a statistic obtained from an observed random sample and the value of corresponding parameter being estimated.

19- Describe sampling units?
Before selecting the sample, the population must be divided into parts that parts are called sampling units.

20- Write two advantages of sampling?
Time Saving and economic.

21- Given $n_1 = 30$ , $n_2 = 25$ , ${\sigma_1}^2$= 300 and ${\sigma_2}^2 = 150$. Find ${\sigma^2}_{\bar{X}_1 - \bar{X}_2}$
${\sigma^2}_{\bar{X}_1 - \bar{X}_2} = \frac{\sigma_1^2}{n_1} + \frac{\sigma_2^2}{n_2} = 4$

22- Define parameter and statistic?
The numerical quantities such as mean, S.D., etc., calculated from population data are called parameters. These are usually denoted by Greek letters.
The numerical quantities such as mean, S.D., etc., calculated from sample data are called statistic. These are usually denoted by Roman letters.

23- What is meant by bias?
It is the difference between the expected value of a statistic and the true value of the parameter being estimated.

24- Take all possible sample of size of 2, without replacement from the following population: 2,4,6,8,10.
Number of samples without replacement = ${}^{n}Cr = {}^{5}C2 = 10$
(2,4), (2,6), (2,8), (2,10), (4,6), (4,8), (4,10), (6,8), (6,10), (8,10).

25- Take all possible sample of size 2 n = 2, with replacement from the following population: 2,4,6,8,10.
Number of samples with replacement = $N^n = 5^2 = 25$
(2,2), (2,4), ……(2,10), (4,2), (4,4), ……(10,10), (10,2), ……., (10,10)

26- What do you know about sampling frame?
It is a complete list of sampling units.

27- What is sample design?
It is a procedure or plan for obtaining a sample from a given population prior to collecting any data.

28- Explain the term non-sampling error. How it is reduced?

29- The errors that are caused by sampling the wrong population of interest. These errors include things such as faulty interviewing, incomplete or inaccurate responses, mistakes in recording or coding the data, errors made in processing the results, and the like.
These errors can be controlled if the volume of data is small.

30- What is probability sampling?

It is a process in which the sample is selected in such a way that every element of population has a known non-zero probability of being included in the sample.

31- For finite population of size $N = 4$, find $\sigma_{\bar{x}}$ if $\mu = 6$, $\sigma = 5$ and $n = 2$.

$\sigma_{\bar{x}} = \frac{\sigma}{\sqrt{n}}\sqrt{\frac{N-n}{N-1}}$ = 2.887

32- A population consists of values 0, 3, 6, 9. How many possible samples should be drawn without replacement of size 3?

${}^{n}Cr = {}^{4}C_{3} = 4$ or ${}^{n}Pr = {}^{4}P_{3} = 24$


## Chapter # 12

1- Explain what is meant by statistical estimation.
It is a procedure of making judgment about the unknown value of a population parameter by using sample observations.

2- Given $n = 4$ , $\sum X = 120$ , $\sum (X - \bar{X})^2 = 303$ , compute the best unbiased estimates of the population mean $\mu$ and of variance $\sigma^2$.

$\mu = \frac{\sum X}{n} = 30,$ $\hat{S}^2 = \frac{\sum (X-\bar{X})^2}{n-1}$ = 101

3- Define interval estimation.
It is a procedure of constructing an interval from a random sample; it has a high specified probability of including the unknown true value of a parameter.

4- Differentiate between point estimate and point estimator.
A **point estimate** is a specific value which we obtain for an estimator is called point estimate or simply an estimator.
A statistic used to estimate a population parameter is called a **point estimator** or simply an estimator.

5- What is meant by point estimation?
The object of point estimation is to obtain a single number from the sample which will represent the unknown value of the parameter.

6- Write down only the name of the properties of good estimator.
Unbiasedness, consistency and efficiency.

7- Explain the statistical inference.
It is an art of drawing conclusions or inferences about the population from the limited information contained in the sample.

8- What is meant by unbiasedness?
It refers to the desirability of the sampling distribution of an estimator being centered at the parameter to be estimated.
An estimator is said to be unbiased estimator if its expected value equals the corresponding parameter.

9- If $\bar{X} = 100$ , $\sigma = 8$ $and$ $n = 64$ , set up a 95% confidence interval for $\mu$?

$\bar{X} \pm Z_{(1-\frac{\alpha}{2})} \frac{\sigma}{\sqrt{n}}$, 1 - α = 95%, α/2 = .025, using table 10a, $Z_{(1-\frac{\alpha}{2})} = 1.96$

$100 \pm 1.96 \frac{8}{\sqrt{64}}$ => (98.04, 101.96)

10- Distinguish between point estimate and interval estimate.
A **point estimate** is a specific value which we obtain for an estimator is called point estimate or simply an estimator.
An interval estimate is an interval calculated from a random sample, such that prior to sampling, it has a high specified probability of including the unknown true value of a parameter.


## Chapter # 13

1- Distinguish between null hypothesis and alternative hypothesis.
A null hypothesis, denoted by "$H_0$", is that hypothesis which is to be tested for possible rejection under the assumption that it is true.
Any hypothesis that is different from the null hypothesis and is set up in parallel to the null hypothesis, is called alternative hypothesis, denoted by $H_1$ or $H_a$.

2- Define simple hypothesis.
If a hypothesis completely specifies the form of the distribution as well as values of its parameters is called simple hypothesis.

3- If $\alpha = 0.05$, what will be the value of $Z_{\frac{\alpha}{2}}$?
$\alpha/2 = 0.025$, using table 10a we get $Z_{\frac{\alpha}{2}}$ = -1.96

4- What are the assumptions of student's t-statistics?
The population should be normal, whose variance is unknown and the sample size is small (n≤30).

5- Define level of significance.
The probability of making type-I error is called level of significance, denoted by "α".

6- Given $\bar{X} = 28$ , $\mu_0 = 28$. Find the value of z-score.
As Z = $\frac{\overline{X}-\mu}{\sigma/\sqrt{n}}$ = 0

7- Write a short note on critical region.
A rejection region or critical region specifies a set of values of the test statistic for which the null hypothesis is rejected.

8- Define acceptance region.
Acceptance region specifies a set of values of the test statistic for which the null hypothesis is not rejected.

9- Define a Type – I error.
A Type-I error is made by rejecting $H_0$ when it is actually; true.

10- What is meant by critical value?
The values of the test statistic which separate the rejection and nonrejection region for the test are called critical values.

11- Explain simple and composite hypothesis.
If a hypothesis completely specifies the form of the distribution as well as values of its parameters is called simple hypothesis. e.g. $\mu = 30$
A hypothesis is said to be composite if it does not completely specify the form of the distribution as well as values of its parameters is called composite hypothesis. e.g. $\mu \leq 30$

12- Write down the steps in testing hypothesis of population mean 4, when the sample size is large.
   i. $H_0: \mu = 4$
   ii. $H_1: \mu \neq 4$
   iii. Specify level of significance α
   iv. Test Statistic: Z = $\frac{\overline{X}-\mu}{\sigma/\sqrt{n}}$
   v. Calculation of the value of test statistic Z.
   vi. Critical region: Z < $Z_{\frac{\alpha}{2}}$ or Z > $Z_{1-\frac{\alpha}{2}}$
   vii. Conclusion: Reject $H_0$, if value of test statistics Z falls in the region of rejection.

13- Given $n_1 = 6\ and\ \sum(X_1 - \bar{X}_1)^2 = 6500, n_2 = 8\ and\ \sum(X_2 - \bar{X}_2)^2 = 1000$, find $S_p$
$S_p = \sqrt{\frac{\sum(X_1-\bar{X}_1)^2+\sum(X_2-\bar{X}_2)^2}{n_1+n_1-2}}$ = 25

14- Differentiate between acceptance region and rejection region.
A rejection region or critical region specifies a set of values of the test statistic for which the null hypothesis is rejected.

Acceptance region specifies a set of values of the test statistic for which the null hypothesis is not rejected.

## Chapter # 14

1- Express two properties of regression line.

(i) The sum of the residuals is zero, that is $\sum e_i = 0$ where $e_i = y_i - \hat{y}_i$

(ii) The regression line always passes through the point of means $(\overline{X} - \overline{Y})$, the centre of gravity of the observed data.

2- What is meant by intercept?

It is the expected mean value of dependent variable when all the values of independent variable are 0. Or A point where the regression line intersects the axis.

3- Differentiate between regressor and regressand?

The variable, that forms the basis of estimation or prediction, is called regressor or predictor or independent variable.

The variable, whose value depends upon the selected value of the independent variable, is called regressand or response or dependent variable.

4- Write any two formulae of correlation co-efficient.

$$r = \frac{\sum(X-\bar{X})(Y-\bar{Y})}{\sqrt{\sum(X-\bar{X})^2 \sum(Y-\bar{Y})^2}}, \qquad r = \frac{S_{xy}}{S_x S_y}$$

5- Given $S_{xy} = 16$ and $S_x . S_y = 81$, find $r_{xy}$

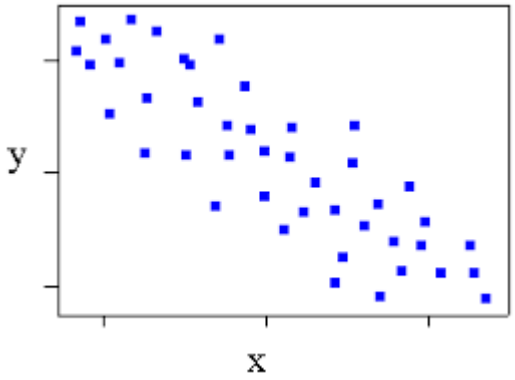
Using $r = \frac{S_{xy}}{S_x S_y}$ = 16/81 = 0.198

6- Write down any two properties of correlation co-efficient.

The range of 'r' is -1 and +1.

It is the geometric mean of the two regression coefficients, i.e. $\left(r = \sqrt{b_{yx} . b_{xy}}\right)$

7- Sketch scatter diagram indicating negative correlation.




8- Explain the term regression co-efficient.

It is the relative change in the expected value of the dependent random variable with respect to a unit increase in the independent non-random variable. It is denoted by β. It is also called slope of line.

9- Given x = 2,4,6 and y = 4,4,4. Find simple correlation co-efficient.

It is zero because y is constant.

10- Write the relationship between regression coefficient and correlation coefficient.

Correlation coefficient is the G.M. of the two regression coefficients. That is $r = \sqrt{b_{yx} . b_{xy}}$

11- What is curve fitting?

It is the process of estimating the parameters of the population regression function of the response variable on a regressor variable, from an observed sample.

12- If $\sum(X - \bar{X})\,(Y - \bar{Y}) = 8400$ and $\sum(X - \bar{X})^2 = 2800$, find $b_{yx}$ =?

$b_{yx} = \frac{\sum(x-\bar{x})(y-\bar{y})}{\sum(x-\bar{x})^2}$ = 3

13- What is difference between correlation and correlation coefficient?

**Correlation** is a statistical measure that indicates the extent to which two or more variables fluctuate together.

**Correlation coefficients** are used in statistics to measure how strong a relationship is between two variables.

14- Interpret the meaning when r = 0.
Both the variables are independent.

15- Define negative correlation and positive correlation.
Correlation is said to be negative if both the variables tend to move in opposite direction.
Correlation is said to be positive if both the variables tend to move in same direction.

16- Define regression analysis.
It is a process by which we estimate one of the variables on the basis of other variable.

17- The regression equations of $x\ on\ y$ is $x = 40.7 - 0.587y$ and of $y\ on\ x$ is $y = 20.8 - 0.912x$, find $r_{xy}$.
$\hat{y} = a + b_{yx}x; \qquad b_{yx} = -0.912$
$\hat{x} = a + b_{xy}y; \qquad b_{xy} = -0.587$
$r = \sqrt{b_{yx}.b_{xy}}$ = - 0.732

18- Given that $n = 15$, $S_y = 16.627$, $S_x = 7.933$ and $\sum(X - \bar{X})\,(Y - \bar{Y}) = 148$, find $b_{yx} = ?\ and\ b_{xy} = ?$
$b_{yx} = \frac{\sum(x-\bar{x})(y-\bar{y})}{S_x^2}$ = 2.35,
$b_{yx} = \frac{\sum(x-\bar{x})(y-\bar{y})}{S_x^2}$ = 0.535

19- Explain the difference between fixed variable and random variable.
Independent variable is called fixed variable whereas dependent variable is called random variable.

20- Given $\bar{X} = 40, \bar{Y} = 180$ and $b = 2$, find the value of intercept "a".
$a = \bar{y} - b_{yx}\bar{x}$ = 100

21- If $b_{yx} = 1.6\ and\ b_{xy} = 0.4$, find the value of $r_{xy}$.
$r = \sqrt{b_{yx}.b_{xy}}$ = 0.8

22- Given $b_{xy} = -0.86$, $b_{yx} = -0.85$, find $r_{xy}$
$r = \sqrt{b_{yx}.b_{xy}}$ = - 0.855

23- Define simple linear regression.
If the simple regression describes the dependence of the expected value of the dependent random variable Y as a linear function of the independent variable X, then the regression is called simple linear regression.

24- Find the slope and intercept of the line whose equation is 3X – 5$\hat{y}$ = 20.
$\hat{y} = \frac{-20 + 3X}{5}$ = - 4 + 0.6 X; Slope = 0.6, Intercept = -4

25- Given $\sum(X - \bar{X})\,(Y - \bar{Y}) = 0$ , $\sum(X - \bar{X})^2 = 10$ , $\sum(Y - \bar{Y})^2 = 10$ and $n = 5$, find the coefficient of correlation.
r = $\frac{\sum(X-\bar{X})(Y-\bar{Y})}{\sqrt{\sum(X-\bar{X})^2 \sum(Y-\bar{Y})^2}}$ = 0

26- Explain the method of least square.
The least squares method is a statistical procedure to find the best fit for a set of data points by minimizing the sum of the squares of residuals of points from the plotted curve.

27- What is the range of the correlation coefficient "r".
correlation coefficient, **r**, can take a range of values from +1 to -1.

28- If $r = 0.48$, $s_{xy} = 36$, $s_x^2 = 16$, find the value of $S_y$.
Using r = $\frac{S_{xy}}{S_x S_y}$, $S_y$= 18.75

29- Explain scatter diagram.
The scatter diagram is a set of points in a rectangular plane (where x measured horizontally and y measured vertically), where each pair represents an observed pair of values.

30- What are the parameters of the simple linear regression model?
α: intercept; β: Slope

31- Explain the term residual.

In regression analysis, the difference between the observed value of the dependent variable (y) and the predicted value (ŷ) is called the residual (e). where e = y - ŷ

## Chapter # 15

1- Define contingency table.
A table consisting of 'r' rows and 'c' columns in which the data is classified according to two attributes A and B is called (r × c) contingency table.

2- Define rank correlation.
The correlation between ranks of individuals for both the variables x and y is called rank correlation

3- Give formulae for spearman's rank correlation.
$$r = 1 - \frac{6\sum d^2}{n(n^2-1)}$$

4- Give formulae for yule's co-efficient of association.
$$Q = \frac{(AB)(\alpha\beta) - (A\beta)(\alpha B)}{(AB)(\alpha\beta) + (A\beta)(\alpha B)}$$

5- Define class and class frequency in contingency table.
A class is a set of objects which are sharing a given characteristic.
A class frequency is the number of objects which are distributed in a class.

6- Define the term dichotomy for attributes.
Process of dividing the object into two mutually exclusive classes is called dichotomy.

7- What is positive and negative association?
Two attributes A and B are said to be positively associated, if $(AB) > \frac{(A)(B)}{n}$.
Two attributes A and B are said to be negatively associated, if $(AB) < \frac{(A)(B)}{n}$.

8- What is a contrary class?
The classes Aβ, αB, etc. represented by both positive as well as negative attributes are called contrary classes.

9- Define independence of attributes.
Two attributes A and B are said to be independent if there is no relationship between them.
i.e. $(AB) = \frac{(A)(B)}{n}$

10- Whether the two attributes are attributes are independent or associated for the given data:
N = 1024, (A) = 640, (B) = 384 and (AB) = 54
If $(AB) = \frac{(A)(B)}{n}$, here 54 ≠ 240, A and B are associated.

11- Given that $x^2 = 20.178$, if d.f. = 4 ,$\alpha = 0.01$. find table value of $x^2(chi\ square)$ would be.
Using table 11, $\chi^2_{(1-\alpha),\text{d.f.}}$=13.28

12- What is difference between attribute and variable?
A characteristic which varies only in quality form one individual to another is called an attribute.
A characteristic which varies form one individual to another is called variable.

13- Explain the term association of attributes.
Two attributes A and B are said to be associated if they are not independent. i.e., $(AB) \neq \frac{(A)(B)}{n}$

14- Interpret the meaning of Q = +1
The two attributes are completely associated.

15- Given (AB) = 30 , (A) =40, find $(A\beta)$.
$(A\beta)$ = (AB) – (A) = 10

16- What is perfect positive association?
If Q = +1, it is called perfect positive association.

17- Define chi-square.
A chi square ($\chi^2$) statistic is a test that measures how expectations compare to actual observed data. The data used in calculating a chi square statistic must be random, raw, mutually exclusive, drawn from independent variables, and drawn from a large enough sample.

18- Given (A) = 200, (B) = 800 and N = 1000, find (AB). Assuming A and B are independent.
Using $(AB) = \frac{(A)(B)}{n}$ =160

## Chapter # 16

1- Define analysis of time series.
The analysis of time series is the decomposition (isolating) of time series into its components for their separate study.

2- Write down main components of time series.
A time series has the four types of movements or variations are called components of time series.
Secular Trend, Cyclical movements, Seasonal Movements, Irregular Movements.

3- Write two merits of moving average method.
It is easy and simple.
The moving averages of an appropriate period eliminate the periodic fluctuations, so it may be used to eliminate cyclical and seasonal fluctuations.

4- Define seasonal variations.
The seasonal movements are short term movements that represent the regularly recurring changes. These variations indicate a repeated pattern of identical changes in the data within one year or less.

5- Define principle of least square.
It says that "the sum of squares of the deviations of the observed values from the corresponding expected values should be least."

6- Write down two properties of least square line.
The sum of squares of residuals is minimum. $\sum {e_i}^2 = \sum (y_i - \hat{y}_i)^2$
The least squares line always passes through the point of means $(\overline{X} - \overline{Y})$, the centre of gravity of the observed data.

7- Enlist the different methods of measuring secular trend.
Free-Hand curve, Semi-average, Moving average, least squares

8- Define time series in short.
Arrangement of data by successive time periods is called a time series.

9- Write down four phases of a business cycle.
Trough (depression), Expansion (recovery), Peak (boom or prosperity), Recession(contraction)

10- Given that $\sum X = 0, \sum Y = 245, \sum X^2 = 28, \sum XY = 66$ and n = 7. Fit linear trend.
$\hat{y} = a + bx$
$\sum x = 0, b = \frac{\sum xy}{\sum x^2}$ = 2.36 and $a = \frac{\sum y}{n}$ = 35
$\hat{y} = 35 + 2.36x$

11- Explain irregular trend.
The irregular movements refer to fluctuations of irregular nature caused by chance such as wars, floods etc.

12- Give two examples of seasonal trend.
An increase in consumption of electricity in summer.
An increase in sales of cold drinks during summer.

13- Define signal.
It is a systematic component of variation in time series.

14- Define stationary time series.
It is one whose statistical properties such as mean, variance, correlation, etc. are all constant over time.

15- Define the term secular trend.
A secular trend is a long-term movement of time series in one direction. This movement is smooth, steady and regular in nature.

16- Define noise.
The noise is an irregular component of variation in a time series.

17- Write down two advantages of the semi-average method.
It is simple, easy and quick.
It smoothes out seasonal variations.

18- What is historigram?
Graph of time series is historigram.

19- Give any two examples of cyclical variations.
Business cycle, ……

20- Explain two models of time series.
Additive Model: Y = T + C + S + I
Multiplicative Model: Y = T x C x S x I
**T**: Secular Trend, **C**: Cyclical Fluctuation **S**: Seasonal Variation **I**: Irregular Movement

21- Define seasonal variation and give examples.
The seasonal movements are short term movements that represent the regularly recurring changes. These variations indicate a repeated pattern of identical changes in the data within one year or less.

i. The weekly statements of sales in a store.
ii. An after Eid sale in a department store.
iii. An increase in consumption of electricity in summer.
iv. An increase in sales of cold drinks during summer.

## Chapter # 17

1- Differentiate between hardware and software.
The physical parts of the computer are called hardware. Computer hardware consists of CPU, monitor, mouse, keyboard etc. Whereas, the set of instructions given to the computer to solve the problem or to control the operations of the computer is called software. e.g. Microsoft windows, Microsoft office etc.

2- What is secondary storage? Explain with examples.
Secondary storage refers to storage devices and media that are not constantly accessible by a computer system. Examples include external hard drives, portable flash drives, CDs, and DVDs. These devices and media must be either plugged in or inserted into a computer in order to be accessed by the system.

3- What is computer software?
Computer software, also called software, is a set of instructions and its associated documentations that tells a computer what to do or how to perform a task. Software includes all different software programs on a computer, such as applications and the operating system.

4- What is compiler?
A compiler is a translator that converts a program written in a high level language into machine code and operates the whole program at a time.

5- Define central processing unit (CPU).
The component of a computer system that performs the basic operations (such as processing data) of the system, that exchanges data with the system's memory or peripherals, and that manages the system's other components.

6- What is a super computer?
A supercomputer is a computer with a high level of performance compared to a general-purpose computer. These systems are built to process huge amount of data. The performance of a supercomputer is commonly measured in floating-point operations per second (FLOPS) instead of million instructions per second (MIPS).

7- What do you understand by ALU?
The arithmetic logic unit is a part of the processor in which all arithmetic and logical operations on the data are performed

8- What is mean by programming?
Programming is the process of creating a set of instructions that tell a computer how to perform a task. Programming can be done using a variety of computer 'languages,' such as SQL, Java, Python, C++, etc.

9- What is minicomputer?
The mini computers have large memory and faster input/output devices. They are more expensive and have more processing speed than microcomputers.

10- What do you know about DOS?
The term DOS can refer to any operating system, but it is most often used as a shorthand for MS-DOS (Microsoft disk operating system). Originally developed by Microsoft for IBM, MS-DOS was the standard operating system for IBM-compatible personal computers.

11- Write down the names of different computers.
The four basic types of computers are as under:

i. Supercomputer.
ii. Mainframe Computer.
iii. Minicomputer.
iv. Microcomputer.