

**QUESTION BANK FOR
STATISTICAL
METHODS**

Classification, Tabulation, Frequency distribution and Graphical representation

Fill in the Blanks

1. Classification can be done according to _____.
2. Yearwise recording of data of food production will be called _____ classification.
3. The difference between the upper and lower limit of a class is called _____.
4. The average of upper and lower limits of a class is known as _____
5. Formula for determining the number of classes was given by _____
6. Sturges formula for determining the number of classes is _____
7. An arrangement of data in rows and columns is known as _____
8. The graphs of less than and more than ogive intersect at _____
9. The distribution of frequencies according to individual variate values is called _____ distribution.
10. Frequency distributions are often constructed with the help of _____.

Answer

- | | |
|--------------------------------|---------------------------|
| 1. Characteristics | 6. $1+3.322\log N$ |
| 2. Geographical classification | 7. Tabulation of data |
| 3. Class interval | 8. Median |
| 4. Mid value | 9. Frequency distribution |
| 5. H.A. Sturges | 10. Tally marks |

MULTIPLE CHOICES

1. Numerical data presented in descriptive form is called
 - a. Classified presentation
 - b. Tabular presentation
 - c. Graphical presentation
 - d. Textual presentationAns. d
2. For the mid value given 25, 34, 43, 53, 61,70. The first class of the distribution is
 - a. 24.5-34.5
 - b. 25-35
 - c. 20-30
 - d. 20.5-29.5Ans. d
3. In an exclusive type distribution, the limits excluded are
 - a. Lower limit
 - b. Upper limit
 - c. Either of lower or upper limit
 - d. Both a and bAns. c

4. If the lower limit and upper limit of a class are 10 and 40 respectively, the midpoint of the class is:
- 25
 - 12.5
 - 15
 - 30

Ans. a

5. A frequency distribution can be :
- Discrete
 - Continuous
 - Both a and b
 - None of these

Ans. c

6. Classification is applicable in case of :
- Quantitative classification
 - Qualitative classification
 - Both a and b
 - None of these

Ans. C

7. Graphs and charts facilitate:
- Comparison of values
 - To know the trend
 - To know the relationship
 - All the above

Ans.d

8. Ogive curve occur for
- More than type distribution
 - Less than type distribution
 - Both a and b
 - None of these

Ans. C

9. Mode can be calculated from
- Ogive
 - Histogram
 - Bar diagram
 - Pie-chart

Ans.a

10. In bar diagram, bars are:
- Horizontal
 - Vertical
 - Slanting
 - None of above

Ans. vertical

11. In a component bar diagram the length of the bar

- a. Will be same for all
- b. will not be same
- c. Depends on the total
- d. none of these

Ans: Depends on the total

12. The length of the bar will be same for all categories in
- a. Multiple bar diagram
 - b. component bar diagram
 - c. Percentage bar diagram
 - d. none of these

Ans: Percentage bar diagram

13. Sub-divided bar diagram are also called Component bar diagram.

Ans: True

14. The multiple bar diagram is the extension of simple bar diagram.

Ans: True

15. In a bar the width of the bars should be equal.

Ans: True

16. In a percentage bar diagram the length of the bars will not be equal.

Ans: False

Short notes and long question

1. Define primary and secondary data
2. Give the advantages of tabulation.
3. Write a detail note on the types of classification
4. What are the essential characteristics of a good table?
5. Write the limitations of Statistics.
6. Difference between qualitative and quantitative data.
7. Construct a frequency distribution table for the following data
32,45,8, 24, 42, 22, 12, 9, 15, 26, 35, 23, 41, 47, 18, 44, 37, 27, 46, 38, 24,
43,10,21,36, 45, 22, 18.
8. How diagrams are useful in representing statistical data?
9. How to draw a pie chart?
10. Explain how to draw simple and multiple bar diagrams.
11. Explain how to draw Component and percentage bar diagrams.
12. Draw frequency curve for the following data

Seed Yield (gms)	No. of Plants
2.5-3.5	4
3.5-4.5	6
4.5-5.5	10
5.5-6.5	26
6.5-7.5	24
7.5-8.5	15
8.5-9.5	10
9.5-10.5	5

MEASURES OF CENTRAL TENDENCY

Fill in the blanks

1. Arithmetic mean is very much affected by _____.
2. In open end distribution _____ cannot be determined.
3. The mean of the values 11,12,13,14, and 15 is _____.
4. The sum of deviations from mean is _____
5. Sum of square of the deviation from mean is always _____
6. Geometric mean cannot be calculated if any value of the set is _____
7. If the AM of certain value is 9 and GM is 6, the HM is _____.
8. Relation between mean median and mode _____
9. 2nd quartile is same as _____
10. For measuring average shoe size in a market _____ is used.

Answer

- | | |
|-------------------|------------------------|
| 1. Extreme values | 6.0 |
| 2. Mean | 7.4 |
| 3. 13 | 8. Mode=3median -2mean |
| 4. zero | 9. Median |
| 5. minimum | 10. Mode |

Multiple choices

1. mean is a measure of

- a. central values
- b. dispersion
- c. correlation
- d. none of above

Ans. a

2. which of the following represents median?

- a. First quartile
- b. Fiftieth percentiles
- c. Sixth decile
- d. None of above

Ans. b

3. if each observation of a set is divided by 2, then the mean of new values:

- a. is two times the original mean.
- b. Is decreased by 2
- c. Is half of the original mean
- d. Remain the same

Ans. c

4. Harmonic mean is better than other means if the data are for:

- a. Speed or rate
- b. Heights or lengths
- c. Binary values like 0 and 1
- d. None of above

Ans. a

5. Extreme value have no effect on:

- a. Average
- b. Median
- c. Geometric mean
- d. Harmonic mean

Ans. b

6. Correct relationship between A.M. , G.M., and H.M. is :

- a. $A.M.=G.M.=H.M.$
- b. $G.M \geq A.M \geq H.M.$
- c. $A.M \geq G.M. \geq H.M.$
- d. None of above

Ans. c

7. What percentage of values is greater than 3rd quartile?

- a. 75percent
- b. 50percent
- c. 25percent
- d. 0percent

Ans. c

8. A frequency distribution having two modes is said to be:

- a. Unimodal
- b. Bimodal

- c. Trimodal
- d. Without mode

Ans. a

9. The average of n natural numbers is

- a. $n(n+1)/2$
- b. $n+1/2$
- c. $n^2(n+1)/2$
- d. none of above

Ans. b

10. for deciles, the total number of pertition values are

- a. 5
- b. 8
- c. 9
- d. 10

Ans. c

Short notes and long question

1. Given the following frequency distribution, calculate the arithmetic mean

Marks : 64 63 62 61 60 59

Number of students: 8 18 12 9 7 6

- 2. If the sorghum ear- heads are 5,48, 60, 65, 65, 100 gms, calculate the median.
- 3. The following data pertaining to the number of insects per plant. Find median number of insects per plant.

Number of insects per plant (x)	1	2	3	4	5	6	7	8	9	10	11	12
No. of plants(f)	2	3	5	6	10	13	9	5	3	2	2	1

- 4. Find the mode for the following
- 5.

Weight of sorghum in gms (x)	No. of ear head(f)
50	6
65	8
75	16
80	8
95	12

100	4
-----	---

6. If the weights of sorghum ear heads are 45, 60, 48, 100, 65 gms. Find the Geometric mean.
7. Compute quartiles for the data given below (grains/panicles) 25, 18, 30, 8, 15, 5, 20, 40, 45
8. For which type of data mode can be calculated.
9. Write down the merit and demerit of Arithmetic mean.
10. Write the characteristics of good measures of central tendency.

MEASURES OF DISPERSION

Fill the blanks

1. Range is a _____ measures of dispersion.
2. Inter-quartile range is equal to _____
3. Standard deviation is the _____ of variance.
4. Best measures of dispersion is _____
5. _____ deviation can be obtained in case of open end intervalas.
6. _____ moment represents variances.
7. The moments about mean are called _____ moments.
8. Moments about origin are called _____ moments.
9. For symmetric distribution coefficient of skewness is _____
10. Formula for coefficient of variation is _____
11. For a mesokurtic curve, β_2 value is _____
12. Unitless measures of dispersion is _____
13. Peakness of a frequency curve is measured by _____ curve.

Answer

- | | |
|--------------------|-------------------------------------|
| 1. Absolute | 7. Central moment |
| 2. $Q_3 - Q_1$ | 8. Raw moment |
| 3. Square root | 9. 0 |
| 4. Variance | 10. $(SD / \text{mean}) \times 100$ |
| 5. Median | 11. 3 |
| 6. 2 nd | 12. Coefficient of variation |
| | 13. kurtosis |

Multiple choices

1. Which of the following is not a measure of dispersion?
 - a. Mean deviation
 - b. Quartile deviation
 - c. Standard deviation

d. average deviation from mean

Ans. d

2. Which of the following is a unitless measure of dispersion?

- a. Standard deviation
- b. Mean deviation
- c. Coefficient of variance
- d. Range

Ans. c

3. Correct formula for mean deviation from a constant A of a Series in which the variate values $x_1, x_2, x_3, \dots, x_k$ have frequencies f_1, f_2, \dots, f_k respectively is:

- a. $1/N \sum_i (f_i x_i - A)$
- b. $1/N \sum_i (f_i (x_i - A))$
- c. $1/N \sum_i |f_i (x_i - A)|$
- d. None of above

Ans. b

4. Quartile deviation is given by the formula:

- a. $Q.D. = Q_3 + Q_1 / 2$
- b. $Q.D. = Q_3 - Q_1$
- c. $Q.D. = (Q_3 - Q_1) / 2$
- d. None of above

Ans. c

5. Sum of squares of the deviations is minimum when deviations are taken from

- a. Mean
- b. Median
- c. Mode
- d. Zero

Ans. a

6. Which measures of dispersion ensures highest degree of reliability?

- a. Standard deviation
- b. Mean deviation
- c. Coefficient of variance
- d. Range

Ans. a

7. Average wages of workers of a factory are Rs. 550 per month and SD of wages is 110. The coefficient of variation is:

- a. 30 percent
- b. 15percent
- c. 500 percent
- d. 20 percent

Ans. d

8. Quartile deviation is equal to :
- a. Interquartile range
 - b. Double the Interquartile range
 - c. Half of Interquartile range
 - d. None of above

Ans. c

9. For a positive skewed distribution, which of the following inequalities holds?
- a. Median > mode
 - b. Mode > mean
 - c. Mean > median
 - d. Mean > mode

Ans. d

10. For a symmetrical distribution, the coefficient of skewness:
- a. $B_1 = 1$
 - b. $B_2 = 3$
 - c. $B_3 = 0$
 - d. $B_4 = - 1$

Ans. c

11. For a leptokurtic frequency curve, the measure of kurtosis.
- a. $\beta_1 = 0$
 - b. $\beta_2 = - 3$
 - c. $\beta_3 < 1$
 - d. $\beta_2 > 3$

Ans. d

12. in case of positive skewed distribution, the relation between mean, median and mode that holds is :
- a. median > mean > mode
 - b. mean > median > mode
 - c. mean = median = mode
 - d. none of the above

Ans. b

13. in case of positive skewed distribution, the extreme values lie in the
- a. left tail
 - b. right tail

- c. middle
- d. anywhere

Ans. b

14. the standard deviation of a set of values will be:

- a. positive when the values are positive
- b. positive when the values are negative
- c. always positive
- d. all the above

Ans. c

15. All values in a sample are same. Then their variance is:

- a. Zero
- b. One
- c. Not calculable
- d. All the above

Ans. a

Short notes and long question

1. What do you mean by measures of dispersion? Briefly explain different measures of dispersion
2. Write down the characteristics of good measures of dispersion.
3. The weights of 5 ear-heads of sorghum are 100, 102,118,124,126 gms. Find the standard deviation.
4. The yield and plant height of a paddy variety are given, the mean and standard deviation for yield are 50 kg and 10 kg respectively. The mean and standard deviation for plant height are 55 cm and 5 cm respectively. Find the variability among yield and plant height.
5. Define C.V. What are its uses?
6. What are the differences between absolute measure and relative measure of dispersion?
7. Find Variance of the marks from the following table.

Length of pods	0-10	10-20	20-30	30-40	40-50	50-60
No of pods	12	18	27	20	17	6

8. Find the variance.

Seed yield

in gms (x) 2.5-35 3.5-4.5 4.5-5.5 5.5-6.5 6.5-7.5

No. of

plants (f) 4 6 15 15 10

PROBABILITY AND DISTRIBUTIONS

Fill in the blanks

1. Two events are mutually exclusive if there is _____ in between them.
2. Intersection of two mutually exclusive events is a _____ event.
3. Range of probability is _____
4. If two events are A and B , not mutually exclusive event then $P(A \cup B) =$
5. The mean of binomial distribution $b(n,p)$ is _____ and variance _____
6. Normal distribution curve is _____ and _____
7. In poisson distribution the relation between mean and variance _____

Answer

1. No common points
2. Null event
3. 0 to 1
4. $P(A)+P(B)-P(A \cap B)$
5. np and npq
6. bell shaped and symmetrical
7. mean= variance

Multiple choices

1. Probability can take values
 - a. $-\infty$ to ∞
 - b. $-\infty$ to 1
 - c. -1 to 1
 - d. 0 to 1Ans. d
2. Probability is expressed as:
 - a. Ratio
 - b. Proportion
 - c. Percentage
 - d. All the aboveAns. d
3. Two events are said to be independent if:
 - a. Each outcome has equal chance of occurrence
 - b. There is no common point in between them
 - c. One does not affect the occurrence of the other
 - d. Both the events have only one pointAns. c
4. If A and B are two events, the probability of occurrence of either A or B is given as:
 - a. $P(A) + P(B)$

- b. $P(A \cup B)$
- c. $P(A \cap B)$
- d. $P(A)P(B)$

Ans. b

5. The probability of the intersection of two mutually exclusive events is always:
- a. Infinity
 - b. Zero
 - c. One
 - d. None of the above

Ans. b

6. In tossing three coins at a time, the probability of getting at most one head is :
- a. $3/8$
 - b. $7/8$
 - c. $1/2$
 - d. $1/8$

Ans. c

7. The probability of two persons being born on the same day (ignoring date) is :
- a. $1/49$
 - b. $1/365$
 - c. $1/7$
 - d. None of the above

Ans. c

8. Three dice are rolled simultaneously. The probability of getting 12 spots is:
- a. $1/8$
 - b. $25/216$
 - c. $1/12$
 - d. None of the above

Ans. b

9. Given that $P(A) = 1/3$, $P(B) = 1/4$, $P(A/B) = 1/6$, the probability $P(B/A)$ is equal to:
- a. $1/4$
 - b. $3/4$
 - c. $1/8$
 - d. None of the above

Ans. c

10. If for a binomial distribution, $b(n,p)$, $n=4$ and also $P(X=2)=3P(X=3)$, the value of p is:
- a. $9/11$
 - b. 1
 - c. $1/3$
 - d. None of these

Ans.c

11. An experiment succeeds twice as often as it fails. The chance that in the next six trials, there shall be at least four successes is:

- a. $240/729$
- b. $289/729$
- c. $496/729$
- d. None of above

Ans.c

12. The relation between mean and variance in binomial distribution is:

- a. Mean=variance
- b. Mean>variance
- c. Mean< variance
- d. None of above

Ans.b

13. Mean and variance are equal in

- a. Poisson distribution
- b. Binomial distribution
- c. Normal distribution
- d. Chi square distribution

Ans.a

14. Mean ,median and mode all are equal in

- a. Poisson distribution
- b. Binomial distribution
- c. Normal distribution
- d. Chi square distribution

Ans. c

15. A manufacture produces switches and experiences that 2percent switches are defective. The probability that in a box of 50 switches, there are at most two defective is:

- a. $2.5e^{-1}$
- b. e^{-1}
- c. $2e^{-1}$
- d. none of the above

Ans.a

Short notes and long question

1. Two dice are tossed. What is the probability of getting (i) Sum 10 (ii) Sum 8?
2. A card is drawn from a pack of cards. What is a probability of getting (i) a king (ii) a spade (iii) a red card (iv) a numbered card?
3. What is the probability of getting 53 Sundays when a leap year selected at random?
4. Finding the probability of drawing a 4 of hearts or a 6 or any suit from of a pack of cards.
5. Define an independent event.

6. What is conditional probability?
7. State the addition and multiplication laws..
8. Additional theorem of probability.
9. Eight coins are tossed simultaneously. Find the probability of getting atleast six heads.
10. Ten coins are tossed simultaneously. Find the probability of getting (i) atleast 6 heads (ii) exactly 6 heads (iii) atleast 6 heads
11. 20 wrist watches in a box of 100 are defective. If 10 watches are selected at random, find the probability that (i) 10 are defective (ii) 10 are good (iii) at least one watch is defective (iv) at most 3 are defective.
12. Suppose on an average 1 house in 1000 in a certain district has a fire during a year. If there are 2000 houses in that district, what is the probability that exactly 5 houses will have a fire during the year? [given that $e^{-2} = 0.13534$]
13. Write down the properties of normal distribution.
14. The variance of a Poisson distribution is 0.5. Find $P(x = 3)$.
[$e^{-0.5} = 0.6065$]
15. The customer accounts of a certain departmental store have an average balance of Rs.1200 and a standard deviation of Rs.400. Assuming that the account balances are normally distributed. (i) what percentage of the accounts is over Rs.1500? (ii) What percentage of the accounts is between Rs.1000 and Rs.1500? (iii) What percentage of the accounts is below Rs.1500?
16. Give any two examples of Poisson distribution

SAMPLING METHODS

Fill in the blanks

1. _____ is a representative of population.
2. If the numbers of units in a population are limited, it is known as _____ population.
3. Stratified sampling is appropriate when population is _____
4. The probability of selection of any one sample out of ${}^N C_n$ sample is _____
5. The probability of drawing of each sample is equal, and then the method is called _____ method.
6. Sampling error may arise due to _____ selection of sample.
7. Census data is free from _____ error.
8. _____ error is more in census data.

Answer

1. Sample

2. Finite population
3. Heterogeneous
4. $1/{}^N C_n$
5. Simple random sampling
6. Wornng
7. Sampling error
8. Non sampling error

Multiple choices

1. Probability of any one sample of size n being drawn out of N units is:
 - a. $1/N$
 - b. n/N
 - c. $1/n!$
 - d. $1/{}^N C_n$
2. A selection procedure of a sample having no involvement of probability is known as:
 - a. Purposive sampling
 - b. Judgement sampling
 - c. Subjective sampling
 - d. All the above
3. Simple random sample can be drawn with the help of :
 - a. Random number table
 - b. Chit method
 - c. Lottery method
 - d. All the above
4. If the observations recorded on five sampled items are 3, 4,5 6,7 the sample variance is:
 - a. 1
 - b. 0
 - c. 2
 - d. 2.5
5. Probability of including a specified unit in a sample of size n selected out of N units is:
 - a. $1/n$
 - b. $1/N$
 - c. n/N
 - d. N/n

Answer

- 1.d
- 2.d
- 3.d
- 4.d
- 5.b

Short notes and long question

1. Define the following:
population, sample, statistics, parameter
2. Write down the need for sampling.
3. What do you mean by sampling? Write different methods of sampling.
4. Explain the Lottery method.
5. Explain the method of selection of samples in simple random sampling.
6. Explain the method of selection of samples in Stratified random sampling.
7. What do you mean by stratified sampling?

Testing of hypothesis

Fill the blanks

1. The hypothesis which is under test for possible rejection is called _____ hypothesis.
2. A hypothesis contrary to null hypothesis is known as _____ hypothesis.
3. There can be only _____ types of errors in taking a decision about H_0 .
4. Probability of type one error is called _____
5. Rejecting null hypothesis when it is true is _____ error.
6. Accepting null hypothesis when it is false is _____ error.
7. Whether a test is one-sided or two-sided depends on _____ hypothesis.
8. Critical region is also known as _____
9. If β is the probability of type 2 error, the power of test is _____
10. Student's t test is applicable in case of _____ sample
11. The formula for student's t statistics is _____
12. The error degree of freedom for student t test is _____
13. Paired t-test is applicable only when the observations are _____
14. When the population SD is known, the hypothesis about population mean is tested by _____
15. To test whether the data follow an assumed distribution or not is known as a test of _____
16. The value of chi-square statistics depends on the difference between _____ and _____ frequencies.
17. The value of chi-square varies from _____ to _____
18. Chi square statistics value in Yates correction method _____
19. In working with a contingency table of order 4×5 , the degree of freedom for chi-square is _____
20. Homogeneity of several population variances can be tested by _____ test
21. Equality of two population variances can be tested by _____

Answer

1. Null hypothesis
2. Alternate hypothesis
3. Two types
4. α
5. type 1 error
6. typr 2 error
7. alternate hypothesis
8. rejection region
9. $1-\beta$
10. Small sample
11. $\frac{\bar{x}-\mu}{\sqrt{\frac{s^2}{n}}}$
12. $n - 1$
13. Independent but same population
14. Z test
15. Chisquare test
16. Observed and expected
17. 0 to 1
18.
$$\chi^2 = \frac{n \left(|ad - bc| - \frac{n}{2} \right)^2}{c1.c2.r1.r2}$$
19. 12
20. Barlett's test
21. F test

Multiple choices

1. A hypothesis may be classified as:
 - a. Simple
 - b. Composite
 - c. Null
 - d. All the above

Ans. d

2. The hypothesis under test is:
 - a. Simple hypothesis
 - b. Alternative hypothesis
 - c. Null hypothesis
 - d. None of the above

Ans . c

3. Whether a test is one-sided or two-sided depends on:
 - a. Alternative hypothesis
 - b. Composite hypothesis

- c. Null hypothesis
- d. Simple hypothesis

Ans. a

4. Power of a test is related to:
- a. Type I error
 - b. Type II error
 - c. Type I and Type II error both
 - d. None of the above

Ans. b

5. Level of significance is the probability of:
- a. Type I error
 - b. Type II error
 - c. Not committing error
 - d. Any of the above

Ans.a

6. Area of the critical region depends on:
- a. Size of type I error
 - b. Size of type II error
 - c. Value of the statistics
 - d. Number of observation

Ans.a

7. Size of critical region is known as:
- a. Power of the test
 - b. Size of type II error
 - c. Critical value of the test statistics
 - d. Size of the test

Ans.d

8. Degrees of freedom is related to:
- a. No. of observations in a set
 - b. Hypothesis under test
 - c. No. of independent observations in a set
 - d. None of the above

Ans.c

9. Students t-test was invented by:
- a. R.A Fisher
 - b. G.W. Snedecor
 - c. W.S. Gosset
 - d. W.G.Cochran

Ans. c

10. Students t-test is applicable in case of:
- a. Small samples
 - b. For samples of size between 5 and 30
 - c. Large samples
 - d. None of the above

Ans.b

11. The degrees of freedom for statistics-t for paired t-test based on “n”pairs of observations is:
- 2 (n-1)
 - n-1
 - 2n-1
 - None of the above

Ans.b

12. The mean difference between 9 paired observations is 15.0 and the standard deviation of differences is 5.0. the values of statistic t is:
- 27
 - 9
 - 3
 - Zero

Ans.b

13. Degrees of freedom for statistic-chi square in case of contingency table of order (2×2) is :
- 3
 - 4
 - 2
 - 1

Ans.d

14. Paired t-test is applicable when the observations in the two samples are:
- Paired
 - Correlated
 - Equal in number
 - All the above

Ans. d

15. Coefficient of contingency is calculated when:
- The attributions are independent
 - The attributes are associated
 - Both (a) and (b)
 - Neither (a) and (b)

Ans.b

16. Homogeneity of several variances can be tests by:
- Bartlett’s test
 - Fishers’s exact test
 - F-test
 - t-test

Ans. a

17. Equality of several normal population means can be tested by:
- Bartlett’s test
 - F-test

- c. Chi-square test
 - d. t-test
- Ans. b

18. the ratio of between sample variance and within sample variance follows:

- a. F-distribution
 - b. Chi-square distribution
 - c. Z- distribution
 - d. t-distribution
- Ans.a

19. Analysis of variance utilises:

- a. F-test
 - b. Chi-square test
 - c. Z – test
 - d. t-test
- Ans. a

Short notes and long question

1. Write down the steps of testing of hypothesis.
2. Define standard error.
3. Define Type I and Type II error
4. Describe the procedure of comparing two group means.
5. Describe the procedure of comparing two proportions.
6. Define hypothesis. Write its type.
7. Write short notes on errors in hypothesis.
8. Difference between null hypothesis and alternate hypothesis.
9. Define the following:

Critical region, degree of freedom, one tail test and two tail test

10. Based on field experiments, a new variety of green gram is expected to give a yield of 12.0 quintals per hectare. The variety was tested on 10 randomly selected farmer's fields. The yield (quintals/hectare) were recorded as 14.3,12.6,13.7,10.9,13.7,12.0,11.4,12.0,12.6,13.1. Do the results conform to the expectation?
11. In a fertilizer trial the grain yield of paddy (Kg/plot) was observed as follows

Under ammonium chloride 40,35,38,60 &43 kgs

Under urea 37, 42, 52, 65, 60, 49,& 61 kgs.

Find whether there is any difference between the sources of nitrogen?

12. In an experiment the plots were divided into two equal parts. One part received soil treatment A and the second part received soil treatment B. each plot was planted with sorghum. The sorghum yield (kg/plot) was recorded. The results are given below. Test the effectiveness of soil treatments on sorghum yield.

Soil treatment A	49	53	51	52	47	50	52	53
------------------	----	----	----	----	----	----	----	----

Soil treatment B 52 55 52 53 50 54 54 53

13. Define t statistic.
14. Define F statistic.
15. Explain the procedure of testing the equality of two variances.
16. The number of yeast cells counted in a haemocytometer is compared to the theoretical value is given below. Does the experimental result support the theory?

No. of Yeast cells per square	Observed frequency	Expected frequency
0	103	106
1	143	141
2	98	93
3	42	41
4	8	14
5	6	5

17. In order to determine the possible effect of a chemical treatment on the rate of germination of cotton seeds a pot culture experiment was conducted. The results are given below

Chemical treatment and germination of cotton seeds

	Germinated	Not germinated	Total
Chemically Treated	118	22	140
Untreated	120	40	160
Total	238	62	300

Does the chemical treatment improve the germination rate of cotton seeds?

18. What is a contingency table?
19. When and how to apply Yates correction?
20. Explain the χ^2 test of goodness of fit?
21. Explain how to test the independence of attributes?
22. Write the uses of chi-square test.
23. In a cross between parents of the genetic constitution, AAbb & aaBB, the phenotypes in F₂ sample is classified as follows.

AB	Ab	aB	ab	Total
87	29	32	12	160

Correlation and Regression

Fill in the blanks

1. The correlation between two variables can be shown graphically by a _____
2. The spearman rank order correlation is used when the variables to be correlated are measured on _____ scale.
3. When increase in one variable is associated with decrease in other variable, the correlation between these variables is _____
4. The range of correlation coefficient is _____
5. Correlation coefficient is independent of _____
6. Correlation can be calculated when the variables have _____ unit.
7. If correlation coefficient value is +1 then it indicates _____
8. The regression line is also called a _____
9. The slope of the regression line is represented by _____
10. In regression, the independent variable is also called _____
11. The geometric mean of two regression coefficient is _____
12. If one regression coefficient is more than unity then other is _____
13. Regression coefficient is _____ on change of origin and _____ on change of scale.
14. If $r=0$, then the angle between two regression line is _____
15. _____ gave the term Regression.

Answer

1. Scatter diagram
2. Ordinal
3. Negative correlation
4. -1 to +1
5. Change of origin and scale
6. Same
7. Perfect positive correlation
8. Line of best fit
9. B
10. Regressor
11. Correlation coefficient
12. Less than unity
13. Independent, dependent
14. 90°
15. Sir Francis Galton

Multiple choices

1. The term regression was introduced by:
 - a. R.A. Fisher

- b. Sir Francis Galton
- c. Karl Pearson
- d. None of above

Ans. b

2. In a regression line Y on X , the variable X is known as:
- a. Independent variable
 - b. Regressor
 - c. Explanatory variable
 - d. All the above

Ans. d

3. The formula for the estimate of β in the regression equation $Y = \alpha + \beta X + \epsilon$ is:
- a. $\text{Cov}(X,Y)/V(X)$
 - b. $r \frac{\sigma_x}{\sigma_y}$
 - c. $\sum(X_i - \bar{X})(Y_i - \bar{Y}) / \sum(X_i - \bar{X})^2$
 - d. *all the above*

Ans. d

4. If β_{xy} and β_{yx} are two regression coefficient, they have:
- a. Same sign
 - b. Opposite sign
 - c. Either same or opposite
 - d. Nothing can be said

Ans. a

5. If $\beta_{yx} > 1$, then β_{xy} is:
- a. Less than 1
 - b. Greater than 1
 - c. Equal to 1
 - d. Equal to 0

Ans. a

6. Scatter diagram of the variate values (X,Y) gives the idea about:
- a. Functional relationship
 - b. Regression model
 - c. Distribution of errors
 - d. None of the above

Ans. a

7. In the regression line $Y = \alpha + \beta X$, β is called the:
- a. slope of the line
 - b. intercept of the line
 - c. neither (a) nor (b)
 - d. both (a) and (b)

Ans. a

8. The idea of product moment correlation was given by:

- a. R.A. Fisher
- b. Sir Francis Galton
- c. Karl Pearson
- d. Spearman

Ans. c

9. The formula for simple correlation co-efficient between the variable X and Y with usual notations is:

- a. $\text{Cov}(X,Y) / \sqrt{V(X)V(Y)}$
- b. $\mu_{XY} / \sqrt{\mu_{XX} \mu_{YY}}$
- c. $\sigma_{xy} / \sigma_x \sigma_y$
- d. all the above

Ans. d

10. the unit of correlation coefficients is:

- a. Kg/cc
- b. Per cent
- c. Non-existing
- d. None of the above

Ans. c

11. The range of simple correlation coefficient is:

- a. 0 to ∞
- b. $-\infty$ to ∞
- c. 0 to 1
- d. -1 to 1

Ans. d

12. If $\rho=1$, the relation between the two variables X and Y is:

- a. Y is proportional to X
- b. Y is inversely proportional to X
- c. Y is equal to X
- d. None of the above

Ans. c

13. The geometric mean of the two regression coefficient b_{yx} is equal to:

- a. Correlation co-efficient
- b. Co-efficient determination
- c. Regression co-efficient
- d. None of the above

Ans. a

14. Homogeneity of three or more population correlation coefficients can be tested by:

- a. t- test
- b. Z- test

- c. Chi-square test
 - d. F- test
- Ans. c

15. Regression coefficient is independent of the change of:

- a. Scale
 - b. Origin
 - c. Both origin and scale
 - d. Neither origin nor scale
- Ans. b

16. A positive significant correlation between the number of shoes produced and the steel produced per year is:

- a. A nonsense correlation
 - b. A spurious correlation
 - c. A meaningless correlation
 - d. All the above
- Ans. d

17. If the correlation between the two variables X and Y is negative, the regression coefficient of Y on X is:

- a. Positive
 - b. Negative
 - c. Not certain
 - d. None of the above
- Ans. b

Short notes and long question

1. What is a scatter diagram? Mention its uses
2. Define correlation. Write its type.
3. Define regression.
4. Mention the properties of the correlation coefficient?
5. Find correlation coefficient between plant height and number of pods.

X	15	20	17	22	25	29	12
Y	18	17	21	23	20	19	22

6. Properties of regression coefficient.
7. Find Correlation coefficient.

Plant Height (in cm)	45	49	42	51	62	53	58
No. Of branches	30	25	20	32	38	30	29

8. Briefly write the significance of correlation coefficient.
9. From a paddy field, 36 plants were selected at random. The length of panicles(x) and the number of grains per panicle (y) of the selected plants were recorded. The results are given below. Fit a regression line y on x.

S.No.	Y	X
1	95	22.4
2	109	23.3
3	133	24.1
4	132	24.3
5	136	23.5
6	116	22.3
7	126	23.9
8	124	24.0
9	137	23.9
10	90	20.0

10. From the following data, find the regression equation $\sum X = 21$, $\sum Y = 20$, $\sum X^2 = 91$, $\sum XY = 74$, $n = 7$

Design of experiment

Fill in the blanks

1. The variable whose change we wish to study is known as _____
2. Objects of comparison in an experiment are known as _____
3. Repeated application of the treatments is known as _____
4. Grouping of Heterogeneous materials into homogeneous material is _____
5. The object to which treatments are applied _____
6. In order to have valid estimate of experimental error the principles of _____ and _____ are used.
7. ANOVA was developed by _____
8. In order to reduce the experimental error, the principles of _____ and _____ are used.
9. For conducting an experiment with one factor _____ type ANOVA is used.
10. Error degree of freedom for CRD design with t number of treatments is _____
11. CRD is applicable if the experimental material is _____ in nature.

12. In case of RBD, each treatment will occur only once in each _____
13. _____ to the direction of fertility gradient of a field blocking should be done.
14. The most frequently used experimental design is _____
15. In LSD the number of treatments, rows and columns are _____
16. Error degree of freedom for a LSD design with 6 treatments is _____
17. In RBD, the randomization is restricted in _____ direction.
18. For a 6×6 Latin Square Design, the total number of observations are _____

Answer

1. Response variable
2. Treatment
3. Replication
4. Local control
5. Experimental unit
6. Replication, Randomization
7. R.A. Fisher
8. Replication, Local control
9. One way
10. $N-t$
11. Homogeneous
12. Block
13. Perpendicular
14. RBD
15. Equal
16. 20
17. One direction
18. 36

Multiple choices

1. Randomization in an experiment provide:
 - a. The estimate of experimental error
 - b. Impetus to the treatments
 - c. A check to the variation in soil fertility
 - d. None of the above

Ans. a
2. Number of replications in an experiment is based on:
 - a. The precision required

- b. Experimental material available
- c. Heterogeneity of experimental material
- d. All the above

Ans. d

3. The factors responsible for deciding the number of replications in an experiment are:

- a. The desire of the experimenter
- b. Minimum degrees of freedom required for experimental error
- c. Shape of experimental units
- d. All the above

Ans. d

4. The decision about the number of replications is taken in view of:

- a. Size of experimental units
- b. Competition among experimental units
- c. Fraction to be sampled
- d. All the above

Ans. a

5. Local control is a device to maintain:

- a. Homogeneity among blocks
- b. Homogeneity within blocks
- c. Both (a) and (b)
- d. Neither (a) and (b)

Ans. c

6. Local control in the field is maintained through:

- a. Infirmity trails
- b. Randomization
- c. Natural factors
- d. All the above

Ans. a

7. Experimental error is due to:

- a. Experimenter's mistakes
- b. Extraneous factors
- c. Variation in treatment effects
- d. None of the above

Ans. b

8. Errors in a statistical model are always taken to be:

- a. Independent
- b. Distributed as $N(0, \sigma^2)$

- c. Both (a) and(b)
- d. Neither (a) and (b)

Ans. b

9. The information from an experiment stabilizes when error degrees of freedom is at least:
- a. 6
 - b. 8
 - c. 10
 - d. 12

Ans. d

10. A randomized block design is a:
- a. Three restrictional design
 - b. Two restrictional design
 - c. One restrictional design
 - d. No restrictional design

Ans. c

11. In the analysis of data of a randomized block design with b block and v treatments, the error degrees of freedom are:
- a. $b(v-1)$
 - b. $v(b-1)$
 - c. $(b-1)(v-1)$
 - d. None of the above

Ans. c

12. Error sum of squares in RBD as compared to CRD using the same material is:
- a. More
 - b. Less
 - c. Equal
 - d. Not comparable

Ans. b

13. The F ratio in a completely randomized ANOVA is the ratio of
- a. MST/MSE
 - b. MSE/MSTR
 - c. MSE/MST
 - d. MSTR/MSE

Ans. d

14. The mean square is the ratio of sum of square divided by
- a. The total number of observation
 - b. Its corresponding degree of freedom -1

- c. Its corresponding degree of freedom
- d. None of above

Ans. c

Short notes and long questions

- 1. What is experimental error?**
- 2. Define treatment and experimental unit.**
- 3. What is meant by designing an experiment?**
- 4. What do you mean by replication? How it is useful for reducing experimental error?**
- 5. Write down the assumptions of ANOVA.**
- 6. Why ANOVA is used for comparisons of multiple means instead of multiple t test?**
- 7. What is meant by ANOVA? write its type.**
- 8. Briefly explain the principles of design of experiment.**
- 9. Write down the advantages of a RBD.**
- 10. Explain the Layout of the CRD?**
- 11. Give the ANOVA table for RBD with 6 treatments and 4 replications.**
- 12. Furnish the ANOVA model for RBD**
- 13. What is a Latin Square Design?**
- 14. Give the layout of 5×5 LSD design.**

