

(8 pages)

Reg. No. :

Code No. : 8878

Sub. Code : KZOM 32

M.Sc. (CBCS) DEGREE EXAMINATION,
NOVEMBER 2017.

Third Semester

Zoology

BIostatistics and Bioinformatics

(For those who joined in July 2016 and afterwards)

Time : Three hours

Maximum : 75 marks

PART A — (10 × 1 = 10 marks)

Answer ALL questions.

Choose the correct answer :

1. When each member of a population has an equal likely chance of being selected this is called
 - (a) Non-random sampling method
 - (b) Quota sample
 - (c) Snowball sample
 - (d) Equal probability selection method

2. Which of the following is the most effective random sampling technique?
 - (a) Simple random sampling
 - (b) Stratified sampling
 - (c) Cluster random sampling
 - (d) Systemic sampling
3. Which of the following statistics measures the most frequently occurring value in a set of data?
 - (a) Mean
 - (b) Range
 - (c) Mode
 - (d) Median
4. A measure of variability tells us that
 - (a) The high and low extreme scores cancel each other out
 - (b) The mean is greater than mode
 - (c) How well the measure of central tendency represents the entire set of scores
 - (d) Whether or not to compute percentile
5. If the Pearson correlation coefficient R is equal to one then
 - (a) There is perfect positive relationship between the two variables
 - (b) There is positive relationship between the two variables
 - (c) There is no relationship between the two variables
 - (d) There is negative relationship between the two variables



6. Which of the following are likely to be dependent events
- (a) The weather and the number of books on your shelf
 - (b) The colour of your car and its mileage
 - (c) The weight of your car and its mileage
 - (d) The size of your house and size of your shoes
7. One of the following is a non parametric test
- (a) t-test
 - (b) F-test
 - (c) Z-test
 - (d) Chi square test
8. When using the chi-square test for differences of proportion with a contingency table that has r rows and c columns, the degrees of freedom for the test statistics will be
- (a) $n_1 + n_2 - 2$
 - (b) $(r - 1) + (c - 1)$
 - (c) $(r - 1)(c - 1)$
 - (d) $n - 1$

9. What does it indicate 500— 506(3) CGACG
- (a) Deletions
 - (b) Duplication
 - (c) Short sequence repeat
 - (d) Insertion
10. All are nucleotide database except
- (a) SWISS-PROT (b) EMBL
 - (c) DDBJ (d) Gen Bank

PART B — (5 × 5 = 25 marks)

Answer ALL questions choosing either (a) or (b).

Each answer should not exceed 250 words.

11. (a) Define primary and secondary data- Explain their role with suitable examples.

Or

- (b) What is random sampling? Explain any one method of random sampling.

12. (a) What do you understand by dispersion? What purpose does it serve?

Or

- (b) What is skewness? Explain the main types of skewness.



13. (a) State and prove the multiplication theorem of probability. How is the result modified if the events are dependent?

Or

- (b) Define normal distribution. What are the main characteristics of a normal distribution?
14. (a) What is 'Analysis of variance' and where is it used?

Or

- (b) The following table gives the height of students in a class, Find out the quartile deviation.

Height (inches)	50-53	53-56	56-59	59-62	62-65	65-68
No. of students	2	7	24	27	13	3

15. (a) Give a short account about the types of sequences used in bioinformatics.

Or

- (b) What is the nomenclature rule for substitutions and deletions in reporting DNA sequence?

Page 5

Code No. : 8878

PART C — (5 × 8 = 40 marks)

Answer ALL questions choosing either (a) or (b).

Each answer should not exceed 600 words.

16. (a) What is classification of data? Explain the method of classification based on class limit, class interval and number of classes.

Or

- (b) Explain with the help of sketches the construction of the following (i) Bar diagram (ii) Histogram (iii) Frequency polygon (iv) circular diagram.

17. (a) Explain with suitable example the term variation. What purpose does a measure of variation serve? Comment on some of the well known measures of variation along with their respective merits and demerits.

Or

- (b) Calculate Karl Pearson's coefficient of correlation from the following data.

x : 48 35 17 23 47

y : 45 20 40 25 45

Page 6

Code No. : 8878



18. (a) Write in detail about how a binomial distribution arises. Mention the properties and constants of binomial distribution.

Or

- (b) Below are given the gains in weight (lb) of persons on two diets X and Y.

Diet X: 25 32 30 32 24 14 32 35 24 32

Diet Y: 24 34 22 30 42 31 40 30 32 35

Test at 5% level whether the two diets differ significantly with regard to increase in weight (Table value of 't' for 5 degrees of freedom at 5% = 1.573).

19. (a) 1000 students at college level are graded according to their IQ and their economic conditions. Use chi square test to find out whether there is any association between economic conditions and level of IQ.

Economic conditions :	High	Medium	Low	Total
Rich	160	300	140	600
Poor	140	100	160	400
Total	300	400	300	1000

(The table value for degree of freedom 2 at 5% level of significance is 5.99)

Or

Page 7

Code No. : 8878

- (b) Describe the technique of analysis of variance with an illustration for a one way classification.

20. (a) Write an essay on the applications of Bioinformatics.

Or

- (b) Write short notes on

(i) SWISS – PROT (ii) Protein data base.

Page 8

Code No. : 8878

