

College: S. S. College, Jehanabad

Department: Zoology

Class: M.Sc. Semester III

Subject: Zoology / Assignment

Topic: Biostatistics & Bioinformatics (Paper – 1)

Teacher: Praveen Deepak

Last date of assignment submission: 18.05.2020

Mode of submission: E-mail or WhatsApp or Google Classroom

E-mail: deepakprav@gmail.com

WhatsApp No.: +91 75360 68068

Note: Students must submit their detail viz. Name, Class, and College Roll Number during the course of assignment submission. Students can access by going to link <https://classroom.google.com/c/MTAzMzaA0NzcwMTg3/p/MTAzMzaA0NzcwMjky/detail> or joining Google Classroom with code nmbvoje.

*To join Department's group, students can use following link
<https://chat.whatsapp.com/EHuHNfQzoAzJBMFNjsjQx>
or scan QR Code*





S.S. COLLEGE, JEHANABAD
(NAAC Accredited- Grade 'B')

Department's Internal Assessment

Name: _____ Class: _____
Class Roll No.: _____ Total Marks: 40
Assignment: May 18, 2020 Submission: May 19, 2020

M.Sc. Zoology Sem III Paper 1 - Assignment

This assignment is for evaluation of students with respect to online classes and e-contents. It has a total of 40 questions and a total of 40 marks. Each question carries 1 mark. There is no minus marking.

- The stages of a malignant disease (cancer) is recorded using the symbols 0, I, II, III, IV. We say that the scale used is _____
 (A) alphanumeric (B) numerical
 (C) ordinal (D) nominal
- The fundamental statistical indicators are _____
 (A) mean (B) median
 (C) mode (D) variance
- Standard deviation is _____
 (A) the sum of the values divided by their number (B) lower than the maximum value in the series
 (C) an indicator of central tendency for the values of the series (D) All of the above
- If the average of a series of values is 10 and their variance is 4, then the coefficient of variation (= the ratio standard deviation / average) is _____
 (A) 10% (B) 20%
 (C) 40% (D) 80%
- Having two sets of data, we can compare their scattering as follows _____
 (A) for approximately equal average values, the one with a higher standard deviation is more scattered (B) or approximately equal standard deviation values, the one with a lower average is more scattered
 (C) if both the averages and standard deviations differ much between the series, we can compare scattering using the coefficient of variation (D) All of these

6. The median of a series of numerical values is _____
- (A) equal to the average (B) a graph or hart
(C) a number (D) a frequency table
7. The median of a series of numerical values is _____
- (A) a value for which half of the values are higher and half of the values are lower (B) the value located exactly midway between the minimum and maximum of the series
(C) the most commonly encountered values among the series (D) a measure of the eccentricity of the series
8. If a series of values consists of 21 numbers, then, for finding the median, we ordered the series ascending and we use _____
- (A) the 11th value in the ordered series (B) the mean between the 10th and 11th values
(C) the mean between the 11th and 12th values (D) the 10th value in the ordered series
9. The first quartile of a series of values is _____
- (A) the value in the ordered series located at 25% of the number of values in the series (B) the numeric value for which a quarter of the series' values are lower
(C) Both (a) & (b) (D) None of these
10. Relative risk _____
- (A) shows the relationship between a factor assumed to influence the occurrence of disease, and the disease (B) is the ratio of the risk of disease for those exposed and those not exposed to that risk factor
(C) Both (a) & (b) (D) None of these
11. If on a group of 457 patients, for a risk factor we calculated a Relative Risk $RR = 12.74$, the possibility of developing the disease being investigated is _____
- (A) very high when exposed to the factor (B) very small when exposed to the factor
(C) the same in the case of exposure in the case of non-exposure (D) lower in the exposed than in the unexposed, RR being less than 100

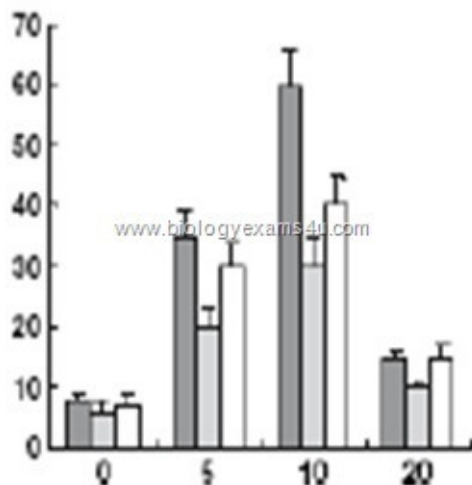
12. The Sensitivity (SN) of a clinical trial is _____
- (A) is the ratio of sick patients, diagnosed as positive, and the total number of sick patients.
- (B) is the ratio of healthy subjects, diagnosed as negative, and the total number of healthy subjects
- (C) is the ratio of sick patients, diagnosed as negative, and the total number of patients.
- (D) is the ratio of sick patients, diagnosed as negative, and the total number of healthy persons
13. A clinical trial is more valuable when _____
- (A) Sensitivity and Specificity have higher values
- (B) Sensitivity is higher than specificity
- (C) Specificity is higher than Sensitivity
- (D) The sensitivity and specificity values are close, even equal, regardless of their values
14. In a contingency table that shows data from a clinical trial is good to have high values for _____
- (A) sick subjects, diagnosed as positive
- (B) healthy subjects, diagnosed as negative
- (C) Both (a) & (b)
- (D) Neither (a) nor (b)
15. For a clinical trial, the Sensitivity is $S_n = 0.562$ and Specificity is $S_p = 0.893$. This means that _____
- (A) the test is a valuable test because both indicators are more than 50%
- (B) the test is a worthless test, since it gives errors when detecting both sick and healthy subjects
- (C) the test is a worthless test, because the sensitivity is too low (lower than 75%)
- (D) None of these
16. A regression line is a straight line which _____
- (A) is located as close as possible to all the points of a scatter chart
- (B) is defined by an equation having 2 parameters: the slope and the intercept
- (C) provides an approximate relationship between the values of two parameters
- (D) All of the above

17. Pearson correlation coefficient, denoted by r , measures _____
- (A) the scattering strength of data for a statistical series
- (B) the strength of the correlation between the mean and median
- (C) the tendency of simultaneous increase or decrease, or inverse evolution, for two numerical parameters
- (D) None of these
18. The correlation coefficient computed for two parameters measured in 429 patients is $r = 0.829$. This means that _____
- (A) the two parameters are directly correlated, and the link is weak - r is positive and close to 0
- (B) the two parameters are inversely correlated, and the link is strong - r is negative and close to 1
- (C) the two parameters are directly correlated, and the link is strong - r is positive and close to 1
- (D) there are too few cases (under 30) and we do not trust this coefficient's value
19. For a Histogram chart, which statement is true?
- (A) The height of the bars is proportional to that class's absolute frequency (number of individuals in the class)
- (B) The width of the bars (classes) is obtained by dividing the difference between the maximum and the minimum values in the series we represent to the number of desired class.
- (C) Both (a) & (b)
- (D) Neither (a) nor (b)
20. The Student's t test is _____
- (A) a parametric test
- (B) a test for comparing averages
- (C) Both (a) & (b)
- (D) Neither (a) nor (b)
21. The result of a statistical test, denoted p , shall be interpreted as follows _____
- (A) the null hypothesis H_0 is rejected if $p < 0.05$
- (B) the null hypothesis H_0 is rejected if $p > 0.05$
- (C) the alternate hypothesis H_1 is rejected if $p > 0.05$
- (D) the null hypothesis H_0 is accepted if $p < 0.05$

22. The null hypothesis (H_0) when comparing two means shall be interpreted as _____

- (A) data do not support the Hypothesis that the populations' means are different
- (B) the two sampling averages do not differ significantly
- (C) the two populations, from which the compared values were sampled, do not differ
- (D) All of above

23.



The graph shown above generally represents _____

- (A) mean and standard error
- (B) mean and mode
- (C) mean and standard deviation
- (D) mean and standard error

24. Chi square test X^2 _____

- (A) measure the degree of deviation of the experimental result from the expected result
- (B) to test the closeness of observed and expected frequency
- (C) to test the population variance and sample variance
- (D) All of the above

25. Find the mode in the following data set:

{11, 12, 13, 14, 14}.

- (A) 11
- (B) 12.8
- (C) 13
- (D) 14

26. Which of the following is not a measure of central tendency?
- (A) Mean (B) Median
(C) Mode (D) Range
27. Correlation coefficient is a number between _____
- (A) +1 and +2 (B) 0 and +1
(C) -1 and 0 (D) -1 and +1
28. Chi square is zero when _____
- (A) expected frequency is lesser than the observed frequency (B) expected frequency is equal to the observed frequency
(C) expected frequency is double that of the observed frequency (D) expected frequency is greater than the observed frequency
29. A single of information in a database is called as _____
- (A) File (B) Field
(C) Record (D) Data set
30. Which of the following is a nucleotide sequence data base?
- (A) EMBL (B) SWISS PORT
(C) PROSITE (D) TREMBL
31. A database of current sequence map of the human genome is called as _____
- (A) OMIM (B) HGMD
(C) Golden path (D) GeneCards
32. Blast programme is used in _____
- (A) DNA sequencing (B) Amino acid sequencing
(C) DNA barcoding (D) Bioinformatics
33. SWISS PORT is related to _____
- (A) portable data (B) Swiss bank data
(C) sequence data bank (D) sequence sequence data

34. BLOSUM matrices are used for _____
- (A) multiple sequence alignment (B) pairwise sequence alignment
(C) phylogenetic analysis (D) All of the above
35. ClastalW is a _____
- (A) multiple sequence alignment tool (B) protein secondary structure predicting tool
(C) data retriving tool (D) nucleic acid sequence analysis tool
36. Phylogenetic relationship can be shown by _____
- (A) dendrogram (B) gene bank
(C) data retriving tool (D) data search tool
37. PRINTS are software used for _____
- (A) detection of genes from genome sequence (B) detection of tRNA genes
(C) prediction of function of a new gene (D) identification of functional domains/motifs of proteins
38. What would be a likely explanation for the existence of pseudogenes?
- (A) gene duplication (B) gene duplication and mutation events
(C) mutation events (D) unequal crossing over
39. CpG islands and codon bias are tools used in eukaryotic genomics to _____.
- (A) identify open reading frames (B) differentiate between eukaryotic and prokaryotic DNA sequences
(C) find regulatory sequences (D) look for DNA-binding domains
40. The identification of drugs through genomic study _____.
- (A) Genomics (B) Cheminformatics
(C) Pharmacogenomics (D) Pharmacogenetics

Name: _____

Class: _____

Saturday, May 18, 2020

Total Marks: 40

M.Sc. Zoology Sem III Paper 1 - Assignment Answer Sheet

This assignment is for evaluation of students with respect to online classes and e-contents. It has a total of 40 questions and a total of 40 marks. Each question carries 1 mark. There is no minus marking.



S.S. COLLEGE, JEHANABAD
(NAAC Accredited- Grade 'B')

Department's Internal Assessment

— 1	1. (A) (B) (C) (D)	— 1	19. (A) (B) (C) (D)	— 1	37. (A) (B) (C) (D)
— 1	2. (A) (B) (C) (D)	— 1	20. (A) (B) (C) (D)	— 1	38. (A) (B) (C) (D)
— 1	3. (A) (B) (C) (D)	— 1	21. (A) (B) (C) (D)	— 1	39. (A) (B) (C) (D)
— 1	4. (A) (B) (C) (D)	— 1	22. (A) (B) (C) (D)	— 1	40. (A) (B) (C) (D)
— 1	5. (A) (B) (C) (D)	— 1	23. (A) (B) (C) (D)		
— 1	6. (A) (B) (C) (D)	— 1	24. (A) (B) (C) (D)		
— 1	7. (A) (B) (C) (D)	— 1	25. (A) (B) (C) (D)		
— 1	8. (A) (B) (C) (D)	— 1	26. (A) (B) (C) (D)		
— 1	9. (A) (B) (C) (D)	— 1	27. (A) (B) (C) (D)		
— 1	10. (A) (B) (C) (D)	— 1	28. (A) (B) (C) (D)		
— 1	11. (A) (B) (C) (D)	— 1	29. (A) (B) (C) (D)		
— 1	12. (A) (B) (C) (D)	— 1	30. (A) (B) (C) (D)		
— 1	13. (A) (B) (C) (D)	— 1	31. (A) (B) (C) (D)		
— 1	14. (A) (B) (C) (D)	— 1	32. (A) (B) (C) (D)		
— 1	15. (A) (B) (C) (D)	— 1	33. (A) (B) (C) (D)		
— 1	16. (A) (B) (C) (D)	— 1	34. (A) (B) (C) (D)		
— 1	17. (A) (B) (C) (D)	— 1	35. (A) (B) (C) (D)		
— 1	18. (A) (B) (C) (D)	— 1	36. (A) (B) (C) (D)		