Name: **Enrolment No:**	

UPES

End Semester Examination, December 2024

Course: Applied Biostatistics **Semester : I**

Program: M.Sc. (Microbiology/ Food, Nutrition & Dietetics)

Duration : 3 Hours

Course Code: HSCC7019 **Marks: 100**

Instructions:

S. No.	Section A Short answer questions/ MCQ/T&F (20Qx1.5M= 30 Marks)	Marks	COs
Q 1	Which measure of central tendency is most affected by extreme values? a. Mean b. Median c. Mode d. Standard deviation	**1.5**	**CO1**
Q 2	In a normal distribution, what percentage of data falls within one standard deviation of the mean? a. 68% b. 75% c. 95% d. 99%	**1.5**	**CO1**
Q 3	What statistical test is used to compare the means of two independent groups? a. t-test b. Chi-square test c. ANOVA d. Mann-Whitney U test	**1.5**	**CO1**

Q 4	Which of the following is not a measure of dispersion? a. Range b. Standard deviation c. Interquartile range (IQR) d. Mode	**1.5**	**CO2**
Q 5	A researcher is interested in studying the relationship between smoking and lung cancer. What types of variables will be used to do the study smoking status and lung cancer. a. Both quantitative b. Both Categorical or Qualitative c. Smoking Categorical and Lung Cancer Quantitative d. Smoking Quantitative and Lung Cancer Categorical	**1.5**	**CO1**
Q 6	The five number summary measure is also called? a. Pie Chart b. Bar Plot c. Box Plot d. Histogram	**1.5**	**CO1**
Q 7	Categorical Proportions can be plot as ? a. Pie Chart b. Bar Plot c. Box Plot d. Histogram	**1.5**	**CO3**
Q 8	Continuous Quantitative variables can be visualized using? a. Pie Chart b. Bar Plot c. Box Plot d. Histogram	**1.5**	**CO3**
Q 9	Define Sample and Population.	**1.5**	**CO1**
Q 10	Define Null Hypothesis	**1.5**	**CO3**
Q 11	Define p-value.	**1.5**	**CO4**
Q 12	Define IQR.	**1.5**	**CO4**
Q 13	Define 95% Confidence Interval for mean	**1.5**	**CO2**
Q 14	State one difference between descriptive statistics and inferential statistics	**1.5**	**CO2**
Q 15	Comment briefly why standard error came into play.	**1.5**	**CO2**
Q 16	State the statistical test to use if one wants to test between bivariate association between categorical variables?	**1.5**	**CO2**
Q 17	Identify the correct formula for calculating 95% Confidence interval for sample mean ($\overline{X}$) if S.D. and S.E. are standard deviation and standard error of the sample respectively? a. ($\overline{X}$ - 1.96 × S.D., $\overline{X}$ + 1.96 × S.D.) b. ($\overline{X}$ + 1.96 × S.D., $\overline{X}$ - 1.96 × S.D.)	**1.5**	**CO4**

	c. ($\overline{X}$ - 1.96 × S.E., $\overline{X}$ - 1.96 × S.E.) d. ($\overline{X}$ + 1.96 × S.E., $\overline{X}$ - 1.96 × S.E.)		
Q 18	Identify which of following is not correct in context of test statistic? a. It captures the deviation of sample estimate from the population. b. It captures the standard error of the sample taken. c. It provides the basis for rejection or not rejection of null hypothesis. d. It provides the level of significance.	**1.5**	**CO4**
Q 19	Identify which is true in context of standard error? a. Standard deviation of the different samples of a same population b. Standard deviation of different observations in a sample c. Standard deviation of different samples from different populations d. Standard deviation of different observations from different samples.	**1.5**	**CO3**
Q 20.	Identify which of the following is true for a null hypothesis to be rejected if p represents the p-value and α is the level of significance? a. $p < \alpha$ b. $p = \alpha$ c. $p > \alpha$ d. $p \sim \alpha$	**1.5**	**CO4**
Section B (4Qx5M=20 Marks)			
Q 1	Discuss difference between standard error and how it differs from standard deviation?	**5**	**CO1**
Q 2	Discuss Box plot and its elements. Also, discuss about outliers.	**5**	**CO2**
Q 3	Explain Correlation Coefficient, mention its range. Explain difference between Pearson Correlation Coefficient and Spearman Correlation Coefficient	**5**	**CO2**
Q 4	Explain Type I error and Type II error in hypothesis testing	**5**	**CO3**
Section C (2Qx15M=30 Marks)			
Q 1	Identify which of the following variables are outcome and exposures. Also, link the variables through possible outcome exposure relationship	**15**	**CO1**

	i) Baby born with low birth weight (yes, no) ii) Mother smoked during pregnancy (yes, no) iii) Number of diarrhea episodes experienced in a year iv) Access to clean water supply (yes, no) v) Child develops leukaemia (yes, no) vi) Duration of exclusive breastfeeding (weeks)		
Q 2	Apply t-test for the following data set to test whether Drug A was effective in reducing the Diastolic Blood Pressure for the 5 patients. Also, state clearly which of the two available t-test will be used and clearly explain all the steps. DBP before taking Drug: 140, 150, 160, 150, 140 DBP after taking Drug: 130, 140, 150, 140, 131 I) State Null and Alternative Hypothesis II) Calculate Test-Statistic III) Conclude the result about effectiveness of Drug A [P(x < 49) = 0.9999 where x follows a t-distribution)]	**15**	**CO4**
Section D **(2Qx10M=20 Marks)**			
Q 1	For the following dataset, calculate mean, median, mode, standard deviation and variance 10,10,10,10,10 10,20,30,40,50	**10**	**CO2**
Q 2	Explain with suitable examples the application of following statistical tests. Provide all necessary steps. a. ANOVA b. Chi-square test	**10**	**CO3**