Name:	
Enrolment No:	UPES UNIVERSITY OF TOMORROW

UPES

End Semester Examination, May 2025

Course: **Fundamentals of Biostatistics**   Semester : **II**

Program: **B.Sc. (Microbiology/ Clinical Research/Food, Nutrition & Dietetics)**

Duration : **3 Hours**

Course Code: **HSCC1029**   **Marks: 100**

Instructions:

S. No.	Section A Short answer questions/ MCQ/T&F (20Qx1.5M= 30 Marks)	Marks	COs
Q 1	Which of the following is the primary condition to use a Z-test instead of a t-test? a. Sample size less than 30 b. Population variance is unknown c. Population standard deviation is known d. Population follows a normal distribution	**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?	**1.5**	**CO2**

	a. Range b. Standard deviation c. Interquartile range (IQR) d. Mode		
Q 5	In a paired sample t-test, the data must: a. Be from two unrelated groups b. Be categorical c. Be from the same group tested at two different times d. Have equal sample sizes	1.5	CO1
Q 6	A researcher wants to compare the mean blood pressure of two independent groups. Which test should be used? a. Paired t-test b. Z-test c. Independent samples t-test d. ANOVA	1.5	CO1
Q 7	In a two-tailed Z-test, if the calculated Z-value is 2.6 and the critical value at 5% significance is 1.96, then: a. Accept the null hypothesis b. Reject the null hypothesis c. Cannot conclude d. Increase the sample size	1.5	CO3
Q 8	If calculated value of test statistic is -3.56, then comment about the p-value.	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.) 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.)	1.5	CO4
Q 18	Identify which of following is not correct in context of test statistic?	1.5	CO4

	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.		
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 outcomes and which are exposures.
	 Then describe at least two plausible exposure outcome relationships that connect them.
	    1. Adult diagnosed with hypertension (yes / no)
	    2. Average daily sodium intake (milligrams per day)
	    3. Minutes of moderate to vigorous physical activity per week (continuous)
	    4. Body mass index (BMI) ($kg\,m^{-2}$, continuous)
		15	**CO1**

	5. Family history of cardiovascular disease (yes / no)		

Q 2	Apply t-test for the following data set to test whether Drug A was effective in reducing the Diastolic Blood Pressure (DBP) 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$)]$
	
	OR
	
	A pharmaceutical company produces 500 mg of medication tablets. The quality control department claims the average active ingredient in the tablets is 500 mg. A regulatory audit tests a random sample of 40 tablets and finds a sample mean of 497.5 mg.
	The population standard deviation (σ) is known to be 5 mg.
	At $\alpha = 0.05$, test whether the tablets are under-dosed.
	
	    I) State Null and Alternative Hypothesis
	    II) Calculate Test-Statistic
	    III) Conclude the result
		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**