Post Graduate Diploma in Management (Hospital & Health Management)

PGDM – 2024-26 Batch

1ˢᵗ Year – 1ˢᵗ Semester End Examination

Subject & Code	: **Bio Statistics (BS)-CC 604**		**Reg. No.**	:
Semester & Batch	: **I, 2024-26**		**Date**	: **23-12-2024**
Time & Duration	: **10:30 A.M.-01:30 P.M. (3 Hrs.)**		**Max. Marks**	: **70**

Instructions:

- Budget your time as per the marks given for each question and write your answer accordingly.
- Don't write anything on the Question Paper except writing your Registration No.
- Mobile Phones are not allowed even for computations.

Part A: Q.1 to Q.10 all questions are compulsory (10 X 2 Marks = 20 Marks)

1. **What type of data is required for the Wilcoxon Signed-Rank Test?**

   a) Nominal data
   b) Ordinal or continuous data
   c) Interval data only
   d) Ratio data only

2. **Which of the following is NOT an assumption of the Wilcoxon Signed-Rank Test?**

   a) The data must be paired
   b) The differences between pairs are independent
   c) The scale of measurement must be at least ordinal
   d) The data must follow a normal distribution

3. **In the Wilcoxon Signed-Rank Test, how are zero differences between pairs handled?**

   a) They are ignored
   b) They are given a rank of zero
   c) They are considered outliers
   d) They are ranked as the average of their positions

4. **For normal distribution which statement is false?**

   a) For normal distribution, mean = median = mode
   b) It is symmetry about the center
   c) 50% of values less than the mean and 50% greater than the mean
   d) None of the above

5. **Which is not probability sampling technique?**

   a) Simple random sampling
   b) Stratified random sampling
   c) Cluster random sampling
   d) Quota Sampling

6. **An event that cannot occur has a probability of x_1, and an event that is certain to occur has a probability of x_2, therefore $x_1 \; and \; x_2$ is**

   a) 1, 0
   b) $\frac{1}{2}$, 0
   c) 0, 1
   d) 0, $\frac{1}{2}$

7. **What is the level of significance (α)?**

   a) The probability of rejecting the null hypothesis when it is true.
   b) The probability of accepting the null hypothesis when it is false.
   c) The probability of rejecting the null hypothesis when it is false.
   d) The probability of accepting the null hypothesis when it is true.

8. **Which of the following is the correct relationship between Type I and Type II errors?**

   a) As the probability of Type I error decreases, the probability of Type II error increases.
   b) As the probability of Type I error increases, the probability of Type II error decreases.
   c) There is no relationship between Type I and Type II errors.
   d) The probability of Type I error is always equal to the probability of Type II error.

9. **Which test is commonly used to compare two nested logistic regression models?**

   a) Hosmer-Lemeshow test
   b) Likelihood Ratio test
   c) Breusch-Pagan test
   d) Chi-square test

10. **Which assumption of regression is often more challenging to satisfy in curvilinear regression compared to linear regression?**

   a) Linearity
   b) Normality of residuals
   c) Homoscedasticity (constant variance of residuals)
   d) Independence of observations

Part B: Q.11 to Q.15 attempt any four questions (4 X 5 Marks = 20 Marks)
Short Notes

11. What is Correlation? Explain Pearson's correlation coefficient with appropriate formula?
OR What is Mann's Whitney test? Write short notes on it's assumptions.

12. What is sampling? What are the probability and non-probability sampling?
OR What is chi-square test. Describe the step to calculate goodness-of-fit.

13. What is measurement scale? Write the characteristic of types of measurement scale.
OR What is hypothesis testing? Discuss the importance of power of test in hypothesis testing.

14. What is variance? Discuss its suitability in understanding the population distribution.
OR
Suppose that each child born is equally likely to be a boy or a girl. Consider a family with exactly three children. List elements in the sample space whose outcomes are all possible genders of the three children. Find the probability to have exactly one child is a girl.

15. What is correlation? Define Spearman Rank Correlation with formula for tied rank.
OR What is coefficient of variation? How it is different from coefficient of dispersion?

Part C: Q.16 to Q.20 attempt any three questions (3 X 10 Marks = 30 Marks)
Long Notes

16. Give the number and name of characteristics used for below ANOVA output. Also, from the given output formulate hypotheses, discuss the decision making on hypotheses using critical value approach.

SUMMARY

Groups	Count	Sum	Average	Variance
Cricket A	10	58.9	5.89	0.027667
Hockey B	12	69.8	5.816667	0.02697
Volley ball C	11	67.8	6.163636	0.018545

ANOVA

Source of Variation	SS	df	MS	F	P-value	F crit
Between Groups	0.750697	2	0.375348	15.40162	0.0000	5.390345863
Within Groups	0.731121	30	0.024371			
Total	1.481818	32				

OR

What is stratified random sampling? How it is different from cluster sampling? Give importance of stratified and cluster sampling in research?

Contd…4..

17. Using below given data, measure and test the association between two groups.

Table 3: Data of smoking and respiratory tract infections

Respiratory Tract Infections	Smoker	Non-Smoker	Total
Yes	35	22	57
No	25	38	63
Total	60	60	120

Formulate the hypotheses, discuss the decision making using critical value approach at 98% level of significance [$\chi^2_{(1),\, 0.01} = 5.41189$].

OR

The below table show the number of females (*f*) of reproductive life span by their age-category. Compute Standard Deviation and Standard Error. Interpret the result of Standard Deviation and Error.

Age-groups	*f*
15-20	7
20-25	3
25-30	3
30-35	6
35-40	6
40-45	5
45-50	3

18. What is hypothesis? Explain the step involved in testing a hypothesis.

OR

An investigator wants to estimate the proportion of new men at his company who currently smoke cigarettes (i.e. the prevalence of smoking). How many new men should be involved in the study to ensure that a 95% confidence interval estimate of the proportion of new men who smoke is within 5% of the true proportion?

19. What is regression. Explain Simple linear regression and Logistic regression with general formula and its underlying assumptions.

OR

A researcher want to analyse the association of cholesterol (mg/dL) and Systolic Blood Pressure (mm/Hg). Using below given output, formulate hypotheses, give line of fit, and interpret the regression equation with appropriate statistical notations.

SUMMARY OUTPUT

Regression Statistics	
Multiple R	0.94888
R Square	0.9003732
Adjusted R Square	0.8977515
Standard Error	23.056181
Observations	40

ANOVA

	df	SS	MS	F	Significance F
Regression	1	182559.68	182559.7	343.4236	0.0000
Residual	38	20200.324	531.5875		
Total	39	202760			

	Coefficients	Standard Error	t Stat	P-value	Lower 95%	Upper 95%	Lower 95.0%	Upper 95.0%
Intercept	-77.289	16.767	-4.610	0.000	-111.232	-43.346	-111.232	-43.346
Cholesterol (mg/dL) (X)	0.762	0.041	18.532	0.000	0.678	0.845	0.678	0.845

20. What is ANOVA? Explain the statistics associated with ANOVA, its assumption and formula.