



WEST BENGAL STATE UNIVERSITY
B.Sc. Honours 5th Semester Examination, 2021-22

STSADSE01T-STATISTICS (DSE1/2)

Time Allotted: 2 Hours

Full Marks: 40

*The figures in the margin indicate full marks.
Candidates should answer in their own words and adhere to the word limit as practicable.
All symbols are of usual significance.*

Answer any four questions from Question No. 1-6 and any two questions from Question No. 7-9

Wherever required, m stands for 100 plus the number represented by the last two digits of your roll number; i.e. $m = 100 + 10x + y$ if the last digit of your registration number is y and the second last digit is x .

You should first put the value of m before solving each question. **Do not** work with general m and then replace it subsequently by the actual value.

1. Which scale of measurement is most appropriate for the following variables – nominal, or ordinal? Answer with proper reasons. 1×5=5
 - (a) Favorite type of music (classical, folk, jazz, rock, other).
 - (b) Marital status (married, divorced, widowed, never married).
 - (c) Diagnostic rating based on mammogram (definitely normal, equivocal, definitely abnormal).
 - (d) Frequency of going out to have fun (never, rarely, occasionally, often).
 - (e) Severity of injury (mild, moderate, severe, death).

2. For the following 2×2 table, find the values of Yule's coefficient of colligation and Pearson's coefficient of absolute association and interpret your results. $2\frac{1}{2} + 2\frac{1}{2}$

	Studied more	Studied less
No. of students Passed	m	$m + 10$
No. of students Failed	$m - 10$	$m + 20$

3. Define odds ratio, risk ratio and risk difference with appropriate examples. 5

4. Describe any model used for contingency table with example. 5

5. Briefly describe the following with practical examples: 5
 - (a) Derived categories, (b) Dichotomy (c) Manifold

6. If you are given a dataset with 1000 patients who are either suffering from novel coronavirus Covid-19 or not along with their age and cholesterol level, is it justified fitting a linear regression model in this data? Give reasons. Also discuss if there is any better regression model in this regard. 5

7. For the following table of religious survey (belief in heaven), find the values of $2\frac{1}{2} \times 4 = 10$ Pearson's chi square, Pearson's coefficient of contingency, Tchuprow's measure and Cramer's V and interpret the results.

	Hindu	Muslim	Christian	Buddhist
Have belief in existence of heaven	$m + 50$	$m + 10$	$m - 20$	$m + 5$
Uncertain belief in existence of heaven	$m - 40$	$m + 20$	$m + 15$	$m - 50$
Does not believe in heaven	$m + 2$	$m - 30$	$m - 5$	$m + 10$

8. For the following table of dose response study, find the values of Goodman- $2\frac{1}{2} \times 4 = 10$ Kruskal Gamma, Kendall's τ_A, τ_B and Somer's D and interpret the results.

	Mild	Moderate	Severe	Death
Mild dose	$m + 50$	$m + 10$	$m - 20$	$m + 5$
Medium dose	$m - 40$	$m + 20$	$m + 15$	$m - 50$
High dose	$m + 2$	$m - 30$	$m - 5$	$m + 10$

9. Describe different types of observational studies with examples. Give an example of non-observational study. 8+2

N.B. : *Students have to complete submission of their Answer Scripts through E-mail / Whatsapp to their own respective colleges on the same day / date of examination within 1 hour after end of exam. University / College authorities will not be held responsible for wrong submission (at in proper address). Students are strongly advised not to submit multiple copies of the same answer script.*

—×—