No. of Printed Pages : 4 **MSTL-015(Set-I)**

M. SC. (APPLIED STATISTICS)
(MSCAST)
Term-End Practical Examination
June, 2025

MSTL-015(Set-I) : STATISTICAL COMPUTING USING R–IV

Time : 2 Hours *Maximum Marks : 25*

Note : *(i) Attempt any **one** question.*

(ii) Solve the question in R-software and create script file.

(iii) Mention necessary steps, hypotheses, interpretation, etc.

(iv) Symbols have their usual meanings.

1. (a) A clinical trial is conducted to test whether a new drug helps in curing a disease. Twenty patients are randomly assigned to either the treatment group or the control group. The obtained results are given as follows :

	Cured	Not-cured	Total
Treatment	8	2	10
Control	4	6	10
Total	12	8	20

Using Fisher's exact test, check whether there is a significant association between treatment and recovery at 5% level of significance. 10

(b) A medical researcher follows up 20 patients after a new cancer treatment. The obtained data are as follows : 15

Survival Time (in months)	**Status (1 = death, 0 = censored)**
3	1
4	1
5	0
6	1
6	0
7	1
8	0

9	1
10	1
10	0
11	1
12	0
13	1
14	0
15	1
16	1
17	0
18	1
19	0
20	1

(i) Estimate the survival time using the Kaplan-Meier method.

(ii) Also prepare Kaplan-Meier survival curve.

2. First install and load the package "dslabs". There is a dataset "movielens" in this package. Using this dataset, answer the following questions :

(a) Compute the number of ratings for each movie and then plot it against the year the movie released. Use the square root transformation on the counts. 10

(b) List down the top rated 25 movies for each year since 1993. Also, compute their average rating. 5

(c) Generally, the most rated movies tend to have above average ratings because more people watch popular movies. To confirm this, stratify the post-1993 movies by ratings per year and compute their average ratings. Make a plot of average rating *vs.* ratings per year and show an estimate of the trend. 10

× × × × ×