QP CODE: 25022631

25022631

Reg No : .....

Name : .....

M.Sc DEGREE (CSS) SPECIAL REAPPEARANCE EXAMINATION, APRIL 2025

Third Semester

M.Sc STATISTICS with DATA SCIENCE

CORE - ST040303 - STATISTICAL COMPUTING 3 USING R AND PYTHON

2020 ADMISSION ONWARDS

67FE688A

Time: 3 Hours Weightage: 30

(Answer any **THREE** questions. Each question carries a weight of **10**)

1. A experiment was carried out on wheat with three treatments in four randomized blocks. the plan and yield per plot in kgm are given below:

BLOCKS							
I		II		III		IV	
A		C		A		B	
8		10		6		10	
C		B		B		A	
12		8		9		8	
B		A		C		C	
10		8		10		9	

Analyze the data and state the conclusions.

2. In an experiment on cotton with 5 manurial treatments, it was observed that the number of plants per plot is varying from plot to plot. The yields of cotton along with the number of plants per plot are given in the following table Analyse the yield data removing the effect of variation in plant population on the yield by analysis of covariance technique and draw the conclusions. The design adopted was an RBD with 4 replications.

Treatments: 5 levels of Nitrogen: $N_0 = 0,\ N_1 = 20,\ N_2 = 40,\ N_3 = 60$ and $N_4 = 80\,kg/ha$

Replicate I	N_1 12 (24)	N_0 10.5 (30)	N_4 27 (30)	N_2 16.5 (28)	N_3 25 (35)
Replicate II	N_3 26 (40)	N_2 20 (25)	N_0 12 (25)	N_4 26 (22)	N_1 15.5 (28)
Replicate III	N_2 22 (32)	N_4 30 (35)	N_3 20 (24)	N_1 20 (35)	N_0 14.5 (30)
Replicate IV	N_1 19 (26)	N_3 18.5 (16)	N_0 8.5 (24)	N_4 29 (30)	N_2 25 (35)

Also, obtain (1) Average variance for the comparisons of treatment means and (2) the gain in precision obtained on using ANCOVA over RBD

3. Consider the seasonal time series data shown in the following table

Year	Q1	Q2	Q3	Q4
2011	318	380	358	423
2012	379	394	412	439
2013	413	458	492	493
2014	461	468	529	575
2015	441	548	561	620

(a) Calculate seasonal indices by the "Ratio to Moving Average "method.

(b) Use Holt-Winters' method to develop a forecasting method for this data $(\alpha = 0.1, \beta = 0.2, \gamma = 0.1)$.

4. Consider the time series data shown in the following table.

Period	y_t	Period	y_t	Period	y_t
1	29	11	29	21	31
2	20	12	28	22	30
3	25	13	28	23	37
4	29	14	26	24	30
5	31	15	27	25	33
6	33	16	26	26	31
7	34	17	30	27	27
8	27	18	28	28	33
9	26	19	26	29	37
10	30	20	30	30	29

a. Calculate and plot the sample autocorrelation and partial autocorrelation functions (maximum 12 lags).

b. What process would you tentatively suggest could represent the most appropriate model for this series?

c. Fit the model identified in part (b).

5. *The following data gives the summary statistics of 3 multivariate normal populations. The sample sizes are* $N_1 = 271, N_2 = 138, \quad N_3 = 117$

The sample mean vectors are $\begin{bmatrix} 2.006 \\ 0.480 \\ 0.082 \end{bmatrix}$, $\begin{bmatrix} 2.167 \\ 0.596 \\ 0.124 \end{bmatrix}$, $\begin{bmatrix} 2.273 \\ 0.521 \\ 0.125 \end{bmatrix}$

The sample dispersion matrices are

$$S_1 = \begin{bmatrix} 0.291 & -0.001 & 0.002 \\ & 0.011 & 0 \\ & & 0.001 \end{bmatrix}, \quad S_2 = \begin{bmatrix} 0.561 & 0.011 & 0.001 \\ & 0.025 & 0.004 \\ & & 0.005 \end{bmatrix}$$

$$S_3 = \begin{bmatrix} 0.261 & 0.030 & 0.003 \\ & 0.017 & 0 \\ & & 0.004 \end{bmatrix}$$

a. *Test the hypothesis that* $H0 : \Sigma_1 = \Sigma_2 = \Sigma_3$

b. *Construct the MANOVA table and test the hypothesis that* $H0 : \mu_1 = \mu_2 = \mu_3$ *assuming that the population have same covariance matrix.*

c. *Test the sphericity of the given pooled population matrix*

6. *The covariance matrix of a 4 variate normal population is given below:*

$$S = \begin{bmatrix} 13.86 & 6.969 & 1.077 & 2.318 \\ & 4.735 & 0.562 & 1.469 \\ & & 0.143 & 0.217 \\ & & & 0.570 \end{bmatrix}$$

Grouping the first two variables and the last two variables into 2 groups, obtain the canonical correlation between the groups. Also, compute the first pair of canonical variates.