MAHATMA GANDHI UNIVERSITY, KOTTAYAM

MGU-UGP (HONOURS)
SECOND SEMESTER EXAMINATION
(2024 ADMISSION ONWARDS)

STATISTICS

MG2DSCSTA100 – Introduction to Statistical Modelling

Duration: 1Hour Maximum: 30 Marks

MODEL QUESTION PAPER

PRACTICAL EXAMINATION

Answer any five questions. Each question carries 6 marks.

Solve the questions using spreadsheet in computer lab. The data in the question along with their answer must be written in the answer paper

1. 6 unbiased coins are tossed together. Assuming binomial distribution, find the probability of getting (a) exactly 4 heads (b) 4 or more heads (c) at most 2 heads. [A] [4]
2. The following data show the number of seeds germinating in different sets. Fit a binomial distribution to the data and obtain the expected frequencies

x	0	1	2	3	4	5
freq.	6	20	28	12	8	6

   [An] [4]
3. Between the hours 2 pm and 4 pm, the average number of phone calls received per minute coming into the switch board of a

company is 3. Assuming Poisson distribution, find the probability that during one particular minute, there will be (a) no phone calls (b) exactly 2 calls (c) at least 2 calls. [A] [4]

4. The following mistakes per page were observed in a book. Fit a Poisson distribution to this data and obtain the expected frequencies

x	0	1	2	3	4	5	6	7	8
freq.	56	156	132	92	37	22	3	1	1

[An] [4]

5. The variable X denotes the marks of students in a test which is normally distributed with mean 45 and SD 10. Find
(1) $P(X > 60)$ (2) $P(40 < X < 56)$ [A] [4]

6. The following data shows the marks of 250 students in an entrance examination. Fit a normal distribution to this data and obtain the expected frequencies.

marks	60-65	65-70	70-75	75-80	80-85	85-90
freq.	21	45	99	64	16	5

[An][4]

7. Generate 25 random numbers from continuous uniform distribution U (0,10). [An] [4]

8. Generate 12 random numbers from a binomial distribution with parameters n = 29, p = 0.65. [An] [4]

MAHATMA GANDHI UNIVERSITY, KOTTAYAM

MGU-UGP (HONOURS)

FIRST SEMESTER EXAMINATION

(2024 ADMISSION ONWARDS)

MG2MDCSTA100 – Time Series Methods and their Applications

PRACTICAL EXAMINATION

Duration: 1Hour Maximum: 30 Marks

*From the following eight questions, students must answer any **five** questions. Each question carries **6** marks.*

[Solve the questions using R Studio in the computer lab. The R code and outputs (except graphs) must be written in the answer paper. Ensure that the graphs generated are evaluated by the invigilator before leaving the lab. Document your analysis and findings.]

1. Create a time series object starting from Jan 2020 with monthly frequency and plot all the time series plots for the following time series data.

   [Analyse][CO1, CO4]

   1.06, 2.63, 1.50, 0.01, -0.43, -0.38, 0.93, 1.54, -2.83, , -2.26, -2.43, -0.11, -0.88, -0.30, -3.70, -1.46, 0.12, 1.25, 0.60, 0.22, 1.54, 4.46, 3.72, 8.78, 2.30, 2.58, 4.15, 0.73, -2.63, 4.83

2. Create a time series object starting from April 2020 with monthly frequency and plot all the time series plots for the following time series data.

   [Analyse][CO1, CO4]

   0.76, 0.74, 0.77, 0.50, -0.31, 1.63, 1.31, -1.31, 0.05, -0.45, -1.29, -0.86, -1.46, -1.46, -1.35, -2.36, -0.34 , -0.02, -1.15, 0.68, 0.76, 0.09, 0.94, 1.35, 1.50

3. Create a time series object starting from June 2020 with monthly frequency and plot all the time series plots for the following time series data.

   [Analyse][CO1, CO4]

   5.89, 3.32, -0.33, 3.51, 3.64, 4.70, 2.59, 6.28, 8.03, 8.18, 6.41, 3.70, 1.23, -0.84, -0.18, -1.60, 0.15, 1.25, 3.71, 3.23, 4.81, 1.59, 2.55, 2.51, 2.01, 5.05, 5.91

4. Create a time series object starting from Jan 2021 with monthly frequency and plot all the time series plots for the following time series data.

   [Analyse][CO1, CO4]

   -2.43, 1.09, -6.26, -12.41, -8.40, -5.17, 4.40, 2.70, 0.78, 3.90, 3.26, 10.81, 7.19, -1.63, 2.29, 1.89, -6.35, -13.31, -11.94, -9.61, -4.85, 1.82, -1.01, -3.14, -2.93

5. Create a time series object starting from Jan 2000 with monthly frequency and plot all the time series plots for the following time series data.

[Analyse][CO1, CO4]

48.37, 50.79, 51.83, 50.64, 48.93, 49.98, 50.41, 50.55, 51.61, 51.43, 48.39, 48.05, 50.78, 50.68, 49.25, 49.44, 49.97, 49.44, 49.42, 50.46, 52.09, 50.45, 48.22, 47.64, 51.32, 51.75, 50.64, 48.31

6. Plot all the time series plots for the **AirPassengers** data (a built-in time series data set on monthly airline passenger totals from 1949 to 1960) in R.

[Analyse][CO1, CO4]

7. Create a time series object and fit an AR model for the following time series data.

[Skill][CO1, CO2, CO4]

1.06, 2.63, 1.50, 0.01, -0.43, -0.38, 0.93, 1.54, -2.83, , -2.26, -2.43, -0.11, -0.88, -0.30, -3.70, -1.46, 0.12, 1.25, 0.60, 0.22, 1.54, 4.46, 3.72, 8.78, 2.30, 2.58, 4.15, 0.73, -2.63, 4.83

8. Create a time series object and fit an AR model for the following time series data.

[Skill][CO1, CO2, CO4]

0.76, 0.74, 0.77, 0.50, -0.31, 1.63, 1.31, -1.31, 0.05, -0.45, -1.29, -0.86, -1.46, -1.46, -1.35, -2.36, -0.34 , -0.02, -1.15, 0.68, 0.76, 0.09, 0.94, 1.35, 1.50

9. Create a time series object and fit an AR model for the following time series data.

[Skill][CO1, CO2, CO4]

5.89, 3.32, -0.33, 3.51, 3.64, 4.70, 2.59, 6.28, 8.03, 8.18, 6.41, 3.70, 1.23, -0.84, -0.18, -1.60, 0.15, 1.25, 3.71, 3.23, 4.81, 1.59, 2.55, 2.51, 2.01, 5.05, 5.91

10. Create a time series object and fit an AR model for the following time series data.

[Skill][CO1, CO2, CO4]

-2.43, 1.09, -6.26, -12.41, -8.40, -5.17, 4.40, 2.70, 0.78, 3.90, 3.26, 10.81, 7.19, -1.63, 2.29, 1.89, -6.35, -13.31, -11.94, -9.61, -4.85, 1.82, -1.01, -3.14, -2.93

11. Create a time series object and fit an AR model for the following time series data.

[Skill][CO1, CO2, CO4]

48.37, 50.79, 51.83, 50.64, 48.93, 49.98, 50.41, 50.55, 51.61, 51.43, 48.39, 48.05, 50.78, 50.68, 49.25, 49.44, 49.97, 49.44, 49.42, 50.46, 52.09, 50.45, 48.22, 47.64, 51.32, 51.75, 50.64, 48.31

12. Fit an AR model for the **AirPassengers** data (a built-in time series data set on monthly airline passenger totals from 1949 to 1960) in R.

[Skill][CO1, CO2, CO4]

13. Fit an AR model for the **sunspot.year** data (a built-in time series data set on yearly sunspot counts from 1749 to 1983) in R. [Skill][CO1, CO2, CO4]

14. Fit an AR model for the **Nile** data (a built-in time series data set on annual river flow data from 1871 to 1970) in R. [Skill][CO1, CO2, CO4]

15. Fit an AR model for the **LakeHuron** data (a built-in time series data set on annual water levels of Lake Huron from 1875 to 1972) in R. [Skill][CO1, CO2, CO4]

16. Simulate 30 values of the specified time series model and plot all the time series plots. Also note down the first 5 and last 5 values generated.

[Interest][CO1, CO3, CO4]

- AR(1) model using $\varphi_1 = 0.70$ and seed value 13
- AR(2) model using $\varphi_1 = 0.50$, $\varphi_2 = 0.33$, intercept 55 and seed value 7

17. Simulate 30 values of the specified time series model and plot all the time series plots. Also note down the first 5 and last 5 values generated.

[Interest][CO1, CO3, CO4]

- AR(1) model using $\varphi_1 = -0.20$, intercept 75 and seed value 99
- AR(2) model using $\varphi_1 = 0.60$, $\varphi_2 = 0.20$ and seed value 345

18. Simulate 30 values of the specified time series model and plot all the time series plots. Also note down the first 5 and last 5 values generated.

[Interest][CO1, CO3, CO4]

- AR(1) model using $\varphi_1 = 0.80$, intercept 50 and seed value 8
- AR(2) model using $\varphi_1 = 1.50$, $\varphi_2 = -0.75$ and seed value 888

19. Simulate 30 values of the specified time series model and plot all the time series plots. Also note down the first 5 and last 5 values generated.

[Interest][CO1, CO3, CO4]

- AR(1) model using $\varphi_1 = 0.70$ and seed value 13
- AR(2) model using $\varphi_1 = 0.60$, $\varphi_2 = 0.20$, intercept 55 and seed value 345

20. Simulate 30 values of the specified time series model and plot all the time series plots. Also note down the first 5 and last 5 values generated.

[Interest][CO1, CO3, CO4]

- AR(1) model using $\varphi_1 = -0.20$ and seed value 99

- AR(2) model using $\varphi_1 = 1.50$, $\varphi_2 = -0.75$, intercept 50 and seed value 888

21. Simulate 30 values of the specified time series model and plot all the time series plots. Also note down the first 5 and last 5 values generated.

[Interest][CO1, CO3, CO4]

- AR(1) model using $\varphi_1 = 0.80$, intercept 60 and seed value 8
- AR(2) model using $\varphi_1 = 0.50$, $\varphi_2 = 0.33$ and seed value 7

22. Simulate 30 values of the specified time series model and plot all the time series plots. Also note down the first 5 and last 5 values generated.

[Interest][CO1, CO3, CO4]

- AR(1) model using $\varphi_1 = 0.70$, intercept 55 and seed value 13
- AR(2) model using $\varphi_1 = 1.50$, $\varphi_2 = -0.75$ and seed value 888

23. Simulate 30 values of the specified time series model and plot all the time series plots. Also note down the first 5 and last 5 values generated.

[Interest][CO1, CO3, CO4]

- AR(1) model using $\varphi_1 = -0.20$ and seed value 99
- AR(2) model using $\varphi_1 = 0.50$, $\varphi_2 = 0.33$, intercept 45 and seed value 7

24. Simulate 30 values of the specified time series model and plot all the time series plots. Also note down the first 5 and last 5 values generated.

[Interest][CO1, CO3, CO4]

- AR(1) model using $\varphi_1 = 0.80$ and seed value 8
- AR(2) model using $\varphi_1 = 0.60$, $\varphi_2 = 0.20$, intercept 40 and seed value 345

25. Simulate 30 values of the specified time series model and plot all the time series plots. Also note down the first 5 and last 5 values generated.

[Interest][CO1, CO3, CO4]

- AR(1) model using $\varphi_1 = 0.70$, intercept 50 and seed value 13
- AR(2) model using $\varphi_1 = 0.50$, $\varphi_2 = 0.33$ and seed value 7

26. Simulate 30 values of the specified time series model and plot all the time series plots. Also note down the first 5 and last 5 values generated.

[Interest][CO1, CO3, CO4]

- AR(1) model using $\varphi_1 = -0.20$ and seed value 99
- AR(2) model using $\varphi_1 = 0.60$, $\varphi_2 = 0.20$, intercept 50 and seed value 345

27. Simulate 30 values of the specified time series model and plot all the time series plots. Also note down the first 5 and last 5 values generated.

[Interest][CO1, CO3, CO4]

- AR(1) model using $\varphi_1 = 0.80$ and seed value 8
- AR(2) model using $\varphi_1 = 1.50$, $\varphi_2 = -0.75$, intercept 100 and seed value 888

28. Simulate 30 values of the specified time series model and plot all the time series plots. Also note down the first 5 and last 5 values generated.

[Interest][CO1, CO3, CO4]

- AR(1) model using $\varphi_1 = 0.70$, intercept 125 and seed value 13
- AR(2) model using $\varphi_1 = 0.60$, $\varphi_2 = 0.20$ and seed value 345

29. Simulate 30 values of the specified time series model and plot all the time series plots. Also note down the first 5 and last 5 values generated.
[Interest][CO1, CO3, CO4]

- AR(1) model using $\varphi_1 = -0.20$, intercept 55 and seed value 99
- AR(2) model using $\varphi_1 = 1.50$, $\varphi_2 = -0.75$ and seed value 888

30. Simulate 30 values of the specified time series model and plot all the time series plots. Also note down the first 5 and last 5 values generated.
[Interest][CO1, CO3, CO4]

- AR(1) model using $\varphi_1 = 0.80$ and seed value 8
- AR(2) model using $\varphi_1 = 0.50$, $\varphi_2 = 0.33$, intercept 75 and seed value 7

31. Simulate 30 values of the specified time series model and plot all the time series plots. Also note down the first 5 and last 5 values generated.
[Interest][CO1, CO3, CO4]

- AR(1) model using $\varphi_1 = 0.70$ and seed value 13
- AR(2) model using $\varphi_1 = 1.50$, $\varphi_2 = -0.75$, intercept 25 and seed value 888

32. Simulate 30 values of the specified time series model and plot all the time series plots. Also note down the first 5 and last 5 values generated.
[Interest][CO1, CO3, CO4]

- AR(1) model using $\varphi_1 = -0.20$, intercept 10 and seed value 99
- AR(2) model using $\varphi_1 = 0.50$, $\varphi_2 = 0.33$ and seed value 7

33. Simulate 30 values of the specified time series model and plot all the time series plots. Also note down the first 5 and last 5 values generated.
[Interest][CO1, CO3, CO4]

- AR(1) model using $\varphi_1 = 0.80$, intercept 50 and seed value 8
- AR(2) model using $\varphi_1 = 0.60$, $\varphi_2 = 0.20$ and seed value 345

34. Simulate 15 white noise values using seed values 1 and 100 and plot them. Also note down all the 15 values generated. [Skill][CO2, CO4]

35. Simulate 15 white noise values using seed values 234 and 143 and plot them. Also note down all the 15 values generated. [Skill][CO2, CO4]

36. Simulate 15 white noise values using seed values 888 and 88 and plot them. Also note down all the 15 values generated. [Skill][CO2, CO4]

37. Simulate 15 white noise values using seed values 7 and 18 and plot them. Also note down all the 15 values generated. [Skill][CO2, CO4]

38. Simulate 15 white noise values using seed values 45 and 99 and plot them. Also note down all the 15 values generated. [Skill][CO2, CO4]

39. Simulate 15 white noise values using seed values 16325 and 88 and plot them. Also note down all the 15 values generated. [Skill][CO2, CO4]

40. Simulate 15 white noise values using seed values 888 and 5433 and plot them. Also note down all the 15 values generated. [Skill][CO2, CO4]

(5 x 6 = 30)

MAHATMA GANDHI UNIVERSITY, KOTTAYAM

MGU-UGP (HONOURS)
SECOND SEMESTER EXAMINATION
(2024 ADMISSION ONWARDS)

MG2MDCSTA101 - Data Analysis using JAMOVI and Introduction to R

Duration: 1Hour Maximum: 35 Marks

PRACTICAL EXAMINATION

From the following **eight** questions, students must answer any **five** questions.
Each question carries **seven** marks.

Solve the questions using JAMOVI and R in computer lab. The data in the question along with their answer must be written in the answer paper.

1. The following data gives the weights in kgs of 20 college students recorded correct to the first decimal place. Obtain various descriptive measures using JAMOVI

40.2 26.3 43.6 56.6 46.2 48.3 49.7 28.3 35.4 40.3 38.4 39.6 42.6
50.0 48.3 58.2 57.2 58.3 62.4 31.4 [A][2]

2. The monthly sales of items in a shop is as follows. Draw a simple bar plot in JAMOVI for the given data of monthly sales of items in the shop.

Month	Jan	Feb	Mar	Apr	May	June	Jul
Sales	1700	2000	2500	1200	3000	1000	1300

[A][1]

3. From following information shows the advertisement expenses and sales volume of 10 firms in India.

Firm	1	2	3	4	5	6	7	8	9	10
Advertisement expenses(Rs.in lakhs)	11	13	14	16	16	15	15	14	13	13
Sales Volume (Rs.in lakhs)	50	50	55	60	65	67	66	60	61	57

Using JAMOVI

(a) Construct a scatter plot to study the relationship between advertisement expenditure and sales volume and the nature of correlation.

(b) Calculate Pearson's correlation coefficient between the variables and comment on the result. [An][3]

4. A research study was conducted to examine the differences between older and younger adults on perceived life satisfaction. Ten older adults (over the age of 70) and ten younger adults (between 20 and 30) were given a life satisfaction test. Scores on the measure range from 0 to 60 with high scores indicative of high life satisfaction; low scores indicative of low life satisfaction. The data are presented below. Using JAMOVI, test whether the average score of life satisfaction of the two groups are the same?

Older Adults	45	38	52	48	25	39	51	46	55	46
Younger Adults	34	22	15	27	37	41	24	19	26	36

[An][3]

5. Manures M1, M2, M3, M4 and M5 are applied to 4 plots each in 20 identical plots in which the same variety of wheat was cultivated. The following table gives the yields per plot in quintals. Use Analysis of Variance, test whether the manures have significantly different effects at 5% level of significance in JAMOVI

M1	6	8	4	2
M2	7	5	5	3
M3	4	7	3	2
M4	8	4	6	7
M5	2	5	4	4

[An][3]

6. A company has to choose among three pensions plans. The company wishes to test the hypothesis 'preference for plans is independent of job classification'. It asks the opinion of a sample of 600 employees and obtained the information presented below. Use chi-square, test the hypothesis which the company wishes to do at 5% level of significance in JAMOVI.

	No. of employees favouring			
Job Classification		Plan A	Plan B	Plan C
	Factory Employees	72	10	10
	Clerical Employees	70	20	10
	Supervisors	160	30	10
	Executives	148	40	20

[An][3]

7. The following data on the marks of 30 students is stored in a data vector named data: Using basic R programming, find the arithmetic mean, median and range of the marks of these 30 students. [A][4]

31, 31, 36, 33, 36, 37, 36, 35, 30, 35, 36, 33, 30, 36, 33, 37, 34, 30, 37, 35, 37, 31, 34, 35, 31, 37, 36, 34, 31, 34.

8.Suppose we have four vectors: x = (2,1,3,7), y = (4,2,3,5), z = (10,20) and w = (15). Write the R script for creating following data frames. Also write the output you obtained. [A][4]

a) (x, z)
b) (x, w)
c) (x, y)