MAHATMA GANDHI UNIVERSITY, KOTTAYAM

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

MODEL QUESTION PAPER - THEORY

STATISTICS

MG2DSCSTA100 – Introduction to Statistical Modelling

Duration: 1.5 Hours Maximum: 50 Marks

Students should attempt at least one question from each course outcome to enhance their overall outcome attainability.

Part A

Short answer questions

Answer any seven questions. Each question carries 2 marks.

1. Mention any two examples for continuous random variable. [U] [1]
2. Given the pmf as f(x) = $\frac{x+2}{12}$; x = 1, 2, 3 and 0 elsewhere, find pmf of y = 2x – 3. [U] [1]
3. Define distribution function of a discrete random variable. [K] [1]
4. If it rains, a taxi driver can earn Rs 100 per day. If the climate is fair, he will lose Rs 40 per day. If the probability of raining is 0.4, what is his expected earning? [A] [2]
5. Define variance and harmonic mean in terms of expectation. [K] [2]
6. Given the pdf as f(x) = 2x; 0 < x < 1 and 0 elsewhere, find E (3X). [An] [2]
7. Define discrete uniform distribution. [K] [3]
8. Mention any two properties of exponential distribution. [U] [3]
9. If Z is a standard normal variable, find P (Z > 1.8). [An] [3]
10. Mention a context where Bernoulli distribution can be applied. [A] [3]

Part B

Short Essay questions

Answer any four questions. Each question carries 6 marks.

11. In a factory, 28% of tune ups are done on 4 cylinders, 36% on 5 cylinders, 17% on 6 cylinders and 19% on 8 cylinders. Let X be the random variable denoting the number of cylinders undergoing tune ups. Find pmf of X. Also find P(X > 4). [U] [1]
12. Define moment generating function and characteristic function of a random variable. Mention the properties of moment generating function. [K] [2]
13. A company forecasts the number of units of a product it will sell in a month as follows. Find the mean number of units sold and its variance in sales.

x	100	200	300
probability	0.4	0.3	0.3

[An] [2]

14. Mention the properties of mathematical expectation. [U] [2]
15. The incidence of occupational disease in an industry is such that the workers have a 20% chance of suffering from it. Five workers are chosen at random. Assuming binomial distribution, find the probability that (a) none of the chosen workers would be suffering from the disease (b) at least four of them would be suffering from the disease. [A] [3]
16. Define continuous uniform distribution. Mention its properties. [U] [3]

Part C

Essay questions

Answer any one question. Each question carries 12 marks.

17. The following data is related to number of wrong calls received at a call centre during a year. Fit a Poisson distribution to this data and obtain the expected frequencies

x	0	1	2	3	4
freq.	182	88	66	24	5

[An] [3]

18. Given f(x) = $\frac{x}{15}$; x = 1, 2, 3, 4, 5 and 0 elsewhere. Check whether it is a pmf. If yes, find (a) P ($2 \leq x < 4$) (b) P(x takes an odd number) (c) P ($\frac{1}{2} < x \leq \frac{5}{2}$) [U] [1]

MAHATMA GANDHI UNIVERSITY, KOTTAYAM

MGU-UGP (HONOURS)

SECOND SEMESTER EXAMINATION

(2024 ADMISSION ONWARDS)

MG2MDCSTA100 – Time Series Methods and their Applications

Duration: 1Hour Maximum Marks: 35

Students should attempt at least one question from each course outcome to enhance their overall outcome attainability.

Part A

Multiple Choice Questions

*Answer **All** Questions.*

*Each question carries **2** marks*

1. What is the primary characteristic of a time series data set? [Remember][CO1]
   (a) Data points are not related (b) Random data
   (c) Data is collected over time (d) None of these

2. Which of the following is not a time series plot? [Analyse][CO1]
   (a) Scatter plot (b) Month plot
   (c) Histogram (d) Season plot

3. Which of the following is a component of a time series? [Understand][CO1]
   (a) Trend (b) Seasonality
   (c) Irregularity (d) All of the above

4. The component of a time series that reflects long-term growth or decline is known as: [Remember][CO1]
   (a) Seasonal component (b) Trend component
   (c) Irregular component (d) Cyclical component

5. Identify the properties of a white noise model. [Skill][CO2]
   (a) Independent (b) Identical
   (c) Pattern free (d) All of the above

6. How would you make a random walk time series stationary? [Skill][CO2]
   (a) By taking the square of each observation.
   (b) By taking the first difference of the series.
   (c) By adding a constant to each observation.
   (d) By multiplying each observation by a constant.

7. Correlation coefficient is used to determine: [Interest][CO3]
   (a) A specific value of the y-variable given a specific value of the x-variable
   (b) A specific value of the x-variable given a specific value of the y-variable
   (c) The strength of the relationship between the x and y variables
   (d) None of these

8. What is the range of values that the auto correlation coefficient can take? [Interest][CO1, CO3]
   (a) -1 to 0 (b) 0 to 1
   (b) Any real number (d) -1 to 1

9. Which of the following is the possible cause of auto correlation in the time series data? [Interest] [CO1, CO3]

(a) Seasonality (b) Trend
(c) Measurement Error (d) All of the above

10. Identify the statistical tool for measuring the significance of correlation between a variable and the its lagged values. [Interest][CO1, CO3]

(a) PACF (b) Scatter plot
(b) Season plot (d) Month plot

(2 x 10 = 20)

Part B

Short Essay Type questions

*Answer any **3** questions.*

*Each question carries **5** marks.*

11. What are basic assumptions of time series? [Remember][CO1]
12. Briefly explain different methods of measurement of trend. [Skill][CO1, CO2]
13. Introduce the term stationary process. [Skill][CO2]
14. How would you construct a scatter plot? [Interest][CO3]
15. Comment on auto regressive models. [Interest][CO1, CO3]

(5 x 3 = 15)

MAHATMA GANDHI UNIVERSITY, KOTTAYAM
MGU-UGP (HONOURS)
SECOND SEMESTER EXAMINATION
(2024 ADMISION ONWARDS)
MG2MDCSTA101 - Data Analysis using JAMOVI and Introduction to R

Duration: 1 Hour Maximum Marks: 35

Students should attempt at least one question from each course outcome to enhance their overall outcome attainability.

Part A
Multiple Choice Questions
Answer **All** Questions
Each question carries **2** marks

1.A study measures the income levels of individuals in five different cities. The income data is collected in three categories: low income, middle income, and high income. The cities are then ranked based on their average income. What type of data is being used in this scenario? [An] [2]

a) Nominal b) Ordinal
c) Interval d) Ratio

2. A scatter diagram is plotted with the number of hours studied on the x-axis and exam scores on the y-axis. If the points form a pattern where exam scores increase as the number of hours studied increase, what kind of relationship does the scatter diagram show? [An] [3]

a) Negative correlation b) No correlation
c) Positive correlation d) U-shaped curve

3. Which measure of central tendency is most affected by outliers? [U] [2]

a) Mean b) Median
c) Mode d) Standard deviation

4. Which relative measure is used to compare the variability of two or more groups? [A] [2]

a) Range b) Variance
c) Coefficient of Variation d) Standard deviation

5. The assumption of linearity in linear regression means: [U] [3]

a) The relationship between the dependent and independent variables is linear
b) The residuals should be normally distributed.
c) The residuals should have a mean of zero.
d) The residuals should be uncorrelated with each other.

6. The output of logistic regression is always [U] [3]

a) 0 b) 1 c) 2 d) 0 or 1

7. Which one of the following is an arithmetic operator in R? [U] [4]

a) = = b) > = c) & d) +

8. In R, which statement is used to terminate the execution of a loop? [A] [4]

a) Next b) Skip c) Break d) Delete

9. Which function is used to create data frame in R? [K] [4]

a) create.data.frame ( ) b) data.frame ( )
c) make.data.frame ( ) d) df ( )

10. What is R Markdown primarily used for? [U] [4]

a) Writing HTML code b) Combining R code, text, and plots in a document
c) Managing databases d) Running Java programs

[2x10=20]

Part B

Short Essay Type Questions
Answer any **three** Questions
Each question carries **5** marks

11. Explain the different scales of measurement. [U] [2]
12. Describe a situation in which correlation is applied in data analysis. [A] [3]
13. What are the assumptions of linear regression? [U] [3]
14. List the various data types in R programming language. [K] [4]
15. Prepare a short note on R Markdown. [U] [4]

[5x3=15]