

Probability & Statistics (MT2005)

Date: May 30th 2025

Course Instructor(s)

Ms. Sarah Ahmad

Ms. Kanwal Salaam

Ms. Huma Akbar

Dr. Muzahid Afzal

Final Exam

Total Time (Hrs): 3

Total Marks: 60

Total Questions: 4

Roll No

Section

Student Signature

Do not write below this line

Attempt all the questions on Answer Book. Recreate tables/graphs (if any) on the answer book. Statistical tables and a formula sheet are provided. Personal tables/formula sheets are not permitted.

Lead Pencil work wouldn't be marked or claimed for rechecking. Use permanent ink pen.

Ensure the final answer is expressed in decimal form to 4 significant figures (not as a ratio).

If the data is split or divided, make sure to display the resulting (new) dataset/variable clearly.

If data is ambiguous, make an assumption and proceed—do not ask the invigilator.

Please read carefully and choose the correct variables and units before attempting your solution.

If you select the wrong variable(s), incorrect values and wrong cases etc., your entire answer will be awarded zero marks, even if the working steps are shown.

DO NOT ATTACH QUESTION PAPER.

CLO 1: Statistical Data Interpretation: Analyze and interpret various data types by computing measures of central tendency and dispersion, constructing frequency distributions, and utilizing graphical techniques for precise data representation. [wt:15]

Q1:

Congratulations! You are part of the Customer Insights & Analytics Team at a FinTech Software Company, reporting to Mr. William, the Director of Data Strategy. You have been assigned a customer behavior analytics project using the company's digital banking platform, which serves users across various regions of New York City. The CEO has requested data-driven insights into how users interact with platform features and how these behaviors relate to financial patterns. To support this initiative, Mr. William has asked your team to carry out an analysis using a sample of users. As the team lead, you are responsible for data analysis and reporting. You have extracted a random sample of 10 users from the database (shown in Datasheet: 1 on Page no. 4).

Task: Compare the relative variation in account balances between users who make more than 10 ATM transactions in a month and those who make 10 or fewer transactions. Assess and interpret which group exhibits more stable or consistent account balances.

Datasheet No. 1 Information:

Gender (Male /Female)

User Age (Years)

User Category (Job Holder/ Entrepreneur/Homemaker/Student)

Account Type (Basic/Savings/Salary)

Account balance in \$(hundreds)

Number of ATM transactions in a month

Additional Platform Services Used

User's Primary banking account (1: Yes, 0: No)

Receives interest on the account (1: Yes, 0: No)

Service Region (1: Manhattan, 2: Brooklyn, 3: Queens, 4: Bronx)

CLO 2: Probability and Distribution Analysis: Apply foundational principles of probability to analyze experiments, including Bayes' theorem, evaluate discrete and continuous distributions, and explore applications in machine learning. [wt:15]

Q2:

[Marks=5+3]

- A company deploys an AI-powered chatbot to manage customer queries on its website. During peak hours, the system receives an average of 3 customer queries every 10 seconds. To ensure the chatbot remains responsive and does not get overwhelmed, the AI team is analyzing the incoming request pattern. What is the probability that the system receives at least 3 customer queries in a 5-second interval?
- A Machine Learning Operations (MLOps) team monitors the training time of a complex Convolutional Neural Network (CNN) model on their GPU cluster. The training time follows a normal distribution with a mean of 120 minutes and a standard deviation of 15 minutes. To optimize resource allocation, they aim to analyze the variability in training durations. Determine the cutoff training time below which only 2.5% of the training runs fall.

CLO 3: Inferential and Predictive Analysis: Estimate parameters, perform hypothesis testing, and employ regression analysis to model relationships between variables, assess coefficient significance, and confidently predict future outcomes. [wt:20]

Q3:

[Marks=10]

A software performance engineering team in a data science company is optimizing the training pipeline of a deep learning model to reduce the average training time per epoch. Historically, the mean training time on their cloud GPU cluster was 43 minutes. After implementing improvements in scheduling and memory handling, the team evaluates performance using a sample of 12 randomly selected training runs, which shows an average training time of 41.5 minutes. The new system shows consistent results with only minor variability of $(1.784)^2$ minutes². The team now seeks to statistically verify whether the new system has significantly reduced the mean training time per epoch compared to the historical benchmark.

National University of Computer and Emerging Sciences
Lahore Campus

Datasheet No.1 of Sample Users

Customer ID: 01	Customer ID: 02
Gender: Male	Gender: Female
Age: 38	Age: 32
User Category: Entrepreneur	User Category: Job
Account Type: Savings	Account Type: Salary
Account balance 17	Account balance 11
Number of ATM transactions (month) <u>13</u> >	Number of ATM transactions (month) <u>9</u> <
Additional Platform Services Used: 4	Additional Platform Services Used: 2
User's Primary banking account: 0	User's Primary banking account: 1 ✓
Receives interest on the account: 1	Receives interest on the account: 0
Service Region: 2	Service Region: 1
Customer ID: 03	Customer ID: 04
Gender: Male	Gender: Female
Age: 29	Age: 35
User Category: Job	User Category: Entrepreneur
Account Type: Salary	Account Type: Savings
Account balance 15	Account balance 18
Number of ATM transactions (month) <u>10</u> <-	Number of ATM transactions (month) <u>10</u> <-
Additional Platform Services Used: 1	Additional Platform Services Used: 4
User's Primary banking account: 0 ✓	User's Primary banking account: 0 ✓
Receives interest on the account: 0	Receives interest on the account: 1
Service Region: 1	Service Region: 3
Customer ID: 05	Customer ID: 06
Gender: Male	Gender: Male
Age: 40	Age: 42
User Category: Job	User Category: Entrepreneur
Account Type: Salary	Account Type: Savings
Account balance 16	Account balance 19
Number of ATM transactions (month) <u>14</u> >	Number of ATM transactions (month) <u>17</u> >
Additional Platform Services Used: 6	Additional Platform Services Used: 3
User's Primary banking account: 0	User's Primary banking account: 0
Receives interest on the account: 1	Receives interest on the account: 1
Service Region: 4	Service Region: 1
Customer ID: 07	Customer ID: 08
Gender: Male	Gender: Male
Age: 25	Age: 34
User Category: Student	User Category: Entrepreneur
Account Type: Basic	Account Type: Basic
Account balance 9	Account balance 15
Number of ATM transactions (month) <u>6</u> <-	Number of ATM transactions (month) <u>10</u> <-
Additional Platform Services Used: 3	Additional Platform Services Used: 8
User's Primary banking account: 0 ✓	User's Primary banking account: 1 ✓
Receives interest on the account: 0	Receives interest on the account: 0
Service Region: 3	Service Region: 1
Customer ID: 09	Customer ID: 10
Gender: Male	Gender: Female
Age: 26	Age: 36
User Category: Student	User Category: Entrepreneur
Account Type: Basic	Account Type: Basic
Account balance 6	Account balance 25
Number of ATM transactions (month) <u>5</u> <-	Number of ATM transactions (month) <u>20</u> >
Additional Platform Services Used: 4	Additional Platform Services Used: 7
User's Primary banking account: 0 ✓	User's Primary banking account: 1
Receives interest on the account: 0	Receives interest on the account: 0
Service Region: 4	Service Region: 4