National University of Computer and Emerging Sciences, Lahore Campus

Course Name:	Probability and Statistics	Course Code:	MT2005
Program:	BSE/BSCS/BDS	Semester:	Fall 2023
Duration:	180 Minutes	Total Marks:	100
Paper Date:	30-12-2023	Weight	50%
Section:	ALL SECTIONS	Page(s):	8
Exam Type:	FINAL	Moderator	Ms. Sarah Ahmad

Student : Name: Ahmed Abdullah Roll No. 22L-7503 Section: BDS-3 A

Instruction/Notes:

1. **All the questions are compulsory. Use answer book to solve the questions.**
2. Exchange of calculators is not allowed. You can only **use your own scientific calculator (programmable calculators are not allowed).**
3. **Statistical Tables are attached with the question paper. You are not allowed to bring it.**
4. Don't get panic. If you found any ambiguity in the data then do not ask anything to the invigilator, just make assumption and continue solving your paper.
5. **Pencil work would not be marked and if you attempt any question with all possibilities then only first solution will be considered for marking.**
6. Believe in yourself & do not waste your time by looking in answer sheets of your fellows and copying them.
7. If you are thinking that it's a revenge. No, it is not. It is just an exam. We want you to be a most successful person in life. All the Best!

Don't Hurry. Don't Worry. Do your Best and Let it rest.

Question 1:

Baysian

[CLO-2, Marks: 07]

Computer assembling company receives 24% of parts from supplier X, 36% of parts from supplier Y, and the remaining 40% of parts from supplier Z. Five percent of parts supplied by X, ten percent of parts supplied by Y, and six percent of parts supplied by Z are defective. If an assembled computer has a defective part in it, what is the probability that this part was received from supplier Z?

Question 2:

[CLO-5, Marks: 5+5=10]

In a tech company, an algorithm that models the data processing time (in minutes), say X, for a computer system tasked with analyzing and organizing information, operates in two distinct phases based on the amount of data received.

$$f(x) = \begin{cases} \frac{x}{25} & 0 < x < 5 \\ \frac{10 - x}{25} & 5 < x < 10 \end{cases}$$

pdf → cdf approach

a) Verify the given function of data processing time as a proper pdf.
b) Compute the probability that data processing time is less than 7 minutes.

0076

Question 3: **[CLO-4, Marks: 3+6+3=12]**

According to a June 2022 poll conducted by the Massachusetts Health Benchmarks project, approximately 55 percent of residents answered "serious problem" to the question: "Some people think that childhood obesity is a national health problem. What do you think? Is it a very serious problem, somewhat of a problem, not much of a problem, or not a problem at all?" Assuming that the probability of giving this answer to the question is 0.55 for any Massachusetts resident, find the probability that if 12 residents are chosen at random

a) Exactly seven will answer "serious problem."
b) Two or fewer households will answer "serious problem."
c) No one will answer "serious problem."

Question 4: **[CLO-8 &1, Marks: 5+5+(5+3)+(7+5)=30]**

In a software development firm, the management team is keen on understanding how different programming aptitude levels might influence the efficiency of code compilation, measured in time taken (in seconds). To optimize coding processes and identifying proficient programmers based on their aptitude scores, the team has collected aptitude test scores of 8 programmers alongside their corresponding code compilation times, and they aim to explore the following:

Programmer No.	Aptitude Test Scores	Code Compilation Time (seconds)
1	57	67
2	58	68
3	59	65
4	59	68
5	60	72
6	61	72
7	62	69
8	64	71

a) Whether there is a relationship between code compilation time and aptitude test scores? Calculate and Interpret.
b) Fit a linear regression model to predict code compilation time based on programmers' aptitude test scores.
c) What is the sum of squared residuals? Also assess the goodness of fit of the obtained regression model and interpret.
d) Show the five number summary and Draw a box & whisker plot to pinpoint any potential outliers within the dataset of aptitude test scores.

0076

Question 5:

[CLO-5, Marks: 6+3=9]

The finished inside diameter of a piston ring is normally distributed with a mean of 10 centimeters and a standard deviation of 0.03 centimeters.

a) What is the probability that a piston ring will not have an inside diameter between 9.97 and 10.03 centimeters?

b) Below what value of inside diameter will 15% of the piston rings fall?

Question 6:

[CLO-6, Marks: 15]

THE VIDEO GAME SATISFACTION RATING CASE STUDY: A company that produces and markets video game systems wishes to assess its customers' level of satisfaction with a relatively new model, the XYZ-Box. In the six months since the introduction of the model, the company has received many warranty registrations from purchasers. The company will select a random sample of 12 of these registrations and will conduct telephone interviews with the purchasers. Specifically, each purchaser will be asked to state his or her level of agreement with each of the seven statements listed on the survey instrument. The level of agreement for each statement is measured on a 7-point Likert scale.

The Video Game Satisfaction Survey Instrument

Statement	Strongly disagree						Strongly agree
The game console of the XYZ-Box is well designed.	1	2	3	4	5	6	7
The game controller of the XYZ-Box is easy to handle.	1	2	3	4	5	6	7
The XYZ-Box has high quality graphics capabilities.	1	2	3	4	5	6	7
The XYZ-Box has high quality audio capabilities.	1	2	3	4	5	6	7
The XYZ-Box serves as a complete entertainment center.	1	2	3	4	5	6	7
There is a large selection of XYZ-Box games to choose from.	1	2	3	4	5	6	7
I am totally satisfied with my XYZ-Box game system.	1	2	3	4	5	6	7

Purchaser satisfaction will be measured by adding the purchaser's responses to the seven statements stated above. Experience has shown that a purchaser of a video game system is "very satisfied" if his or her composite score is at least 42. The composite scores for the sample of 12 customers are given in the table below:

Composite Scores	39	38	42	38	46	40	45	42	41	42	44	39

Let μ denote the mean of all possible customer satisfaction ratings based composite scores for the XYZ-Box video game system.

a) Test at 5% significance level that mean of all possible customer satisfaction ratings based composite scores is at least 42.

Fractional Distribution

Question 7: **[CLO-2 & 4, Marks: 4+5+8=17]**

The response time is the speed of page downloads and it is critical for a mobile Web site. As the response time increases, customers become more frustrated and potentially abandon the site for a competitive one. Let X denote the number of bars of service, and let Y denote the response time (to the nearest second) for a particular user and site. The Joint Probability distribution of X and Y is given below.

	X: No. of Bars of Signal Strength		
Y: Response time, (nearest second)	1	2	3
4	0.15	0.1	0.05
3	0.02	0.1	0.05
2	0.04	0.03	0.2
1	0.01	0.02	0.25

a) Obtain the conditional probability of 2 bars of signal strength if the response time is 3 seconds.
b) For a particular user, what is the probability of 4 seconds response time or 1 bar of signal strength?
c) Calculate covariance i.e. Cov (X,Y).

X: 1-3

FORMULA SHEET FINAL EXAM FALL 2023

$$\bar{x} \pm z_{\alpha/2}\frac{\sigma}{\sqrt{n}}$$

$$z = \frac{\bar{x}-\mu}{\sigma/\sqrt{n}}$$

$$\bar{x} \pm t_{(\alpha/2,\nu)}\frac{s}{\sqrt{n}}; \quad t = \frac{\bar{x}-\mu}{s/\sqrt{n}}$$

$$b_1 = \frac{\sum(x-\bar{x})(y-\bar{y})}{\sum(x-\bar{x})^2} = \frac{n\sum xy - \sum x \sum y}{n\sum x^2 - (\sum x)^2}$$

$$b_0 = \bar{y} - b_1\bar{x}$$

$$r = \frac{\sum(x-\bar{x})(y-\bar{y})}{\sqrt{\sum(x-\bar{x})^2 \sum(y-\bar{y})^2}}$$

$$r = \frac{n\sum xy - \sum x \sum y}{\sqrt{(n\sum x^2 - (\sum x)^2)(n\sum y^2 - (\sum y)^2)}}$$

$$t = \frac{r\sqrt{n-2}}{\sqrt{1-r^2}}$$

$E(X) = \sum xP(x)$

Lower limit = Q1-1.5(IQR)

Upper limit = Q3+1.5(IQR)

$Var(X) = \frac{\sum fx^2}{\sum f} - (\frac{\sum fx}{\sum f})^2$

$s^2 = \frac{\sum(X-\bar{X})^2}{n-1}$

$P(A \cap B \cap C) = P(A)P(B)P(C)$

$P(A \cap B \cap C) = P(A)P(B|A)P(C|A \cap B)$

$P(AUB) = P(A)+P(B)-P(A \cap B)$

$$P(A|B) = \frac{P(A \cap B)}{P(B)}$$

$E(X^2) = \sum x^2 f(x)$

$E(X^2) = \int_{-\infty}^{\infty} x^2 f(x)dx$

$i = \frac{p}{100} * n$

$$F = \frac{s_1^2}{s_2^2} \text{ If } s_1^2 > s_2^2$$

$$P(B_r|A) = \frac{P(B_r \cap A)}{\sum_i^k P(B_i \cap A)} = \frac{P(B_r)P(A|B_r)}{\sum_i^k P(B_i)P(A|B_i)} \quad , r = 1,2,,k$$

$$b(x;n,p) = \binom{n}{x}p^x q^{n-x}, x = 0,1,\dots n\,, q = 1-p\,,$$

$$f(x_1,x_2,\dots,x_k;p_1,p_2,\dots,p_k;n) = \binom{n}{x_1,x_2,\dots,x_k}p_1^{x_1}p_2^{x_2},\dots,p_k^{x_k}$$

$$h(x;N,n,k) = \frac{\binom{k}{x}\binom{N-k}{n-x}}{\binom{N}{n}}$$

$$g(x;p) = pq^{x-1}, \qquad x = 1,2,3,\dots$$

$$p(x;\lambda t) = \frac{e^{-\lambda t}(\lambda t)^x}{x!}, \qquad x = 0,1,2,\dots, \qquad \mu = \lambda t$$

$$f(x) = \frac{1}{n}$$

$$z = \frac{x-np}{\sqrt{npq}}$$

$R^2 = 1 - \frac{SSE}{SST}$ or $R^2 = \frac{SSR}{SST}$

$e = Y - \hat{y}$

$SST = \sum(Y-\bar{Y})^2 = \sum Y^2 - (\sum y)^2/n$

$SSE = \sum(Y-\hat{y})^2 = \sum Y^2 - b_0\sum y - b_1\sum XY$

$SSR = \sum(\hat{y}-\bar{Y})^2$

$\mu = E(X)$

$\sigma^2 = E(X^2) - \mu^2$

$$t = \frac{\bar{d}-d_0}{s_d/\sqrt{n}}$$

$$\bar{d} \pm t_{(\alpha/2,\nu)}\, s_d/\sqrt{n}$$

$$z = \frac{x-\mu}{\sigma}$$

0076