**Group #2**

**Group Leader and ID: Anubhav Shankar (01951462)**

**Member Names and IDs: N/A**

**Sectional Written Homework #1**: (**25 points**):

1. (6 points; 2 points \* 3) Consider two random variables X (0 = male; 1 = female) and Y (0= low risk; 1= medium risk; 2 = high risk) with a joint pmf given in the Table below.

Table Joint pmf of X and Y

|  |  |  |  |
| --- | --- | --- | --- |
|  | Y=0 | Y=1 | Y=2 |
| X=0 | 1/25 | 1/10 | 1/5 |
| X=1 | 2/5 | 4/25 | 1/10 |

Compute:

1. p (X = female, Y= high risk) = ?
2. p( X = female ) = ?
3. p (Y= high risk|X=female) = ?

Your answer:

1. **0.1**
2. 2/5 + 4/25 + 1/10 = **0.66**
3. 0.1/(2/5 + 4/25 + 1/10) = 0.1/0.66 = **0.15**
4. (3 points) Suppose a success of a medical trial, X (yes/no), follows a true binomial population distribution. We randomly draw samples of size, n=30, from this binomial population distribution, with the probability of success =0.2. Answer the following questions:
5. Given this sample size (n=30), will the CLT hold true? Show your steps to show if the CLT holds true.

Your answer:

n = 30 ; p = 0.2; 1-p = 0.8 Population follows a true binomial distribution

np -> 30 \* 0.2 = 6 -(1)

n(1-p) -> 30 \* 0.8 = 24 –(2)

For CLT to hold true -> min(np, n(1-p)) > 5 –(3)

Plugging (1) and (2) in (3) ->

Min(6,24) = 6

As 6>5, hence, CLT holds true

1. If the CLT holds true, what is the mean of these sample means? What is the standard deviation of these sample means?

Note: Show your solution steps and no scores are assigned if only answers are provided. Remember you can receive partial scores for your solution steps even if your answer is incorrect.

Your answer:

**For Binomial Distribution** -> Plugging in the values from problem 1

Mean of Sample Means: np = 6 [From (1)]

SD = sqrt((n\*p\*(1-p))/n) = sqrt((30\*0.2\*0.8)/30) = 0.4

1. (4 points) Let’s assume the number of spams follows the Poisson distribution and we randomly draw samples of size =40, with the mean of sample means of 6. Estimate the population mean and the population standard deviation based on CLT.

Note: Show your solution steps and no scores are assigned if only answers are provided. Remember you can receive partial scores for your solution steps even if your answer is incorrect.

Your Answer:

**For Poisson Distribution** -> CLT applies Ɐ n > 30

Mean of Sample means (X) = 6

N = 40

By CLT,

Population Mean = X = 6

SD = sqrt(X \* n) = sqrt(6 \* 40) = **15.5**

# (12 points; 4 points\*3) Use R to

1. Randomly generate 10 rows (n=10) for each of three random variables (RVs), x1, x2, and x3, using a random seed = **490**, from a multivariate **Gaussian** distribution, with a population mean vector of [2, 4, 6] (ie., mean of X1= 2, mean of X2= 4, mean of X3= 6) and a population covariance matrix of these RVs (shown below):

X1 X2 X3

X1 4 3 2

X2 3 9 5

X3 2 5 36

Note: **please use the exact random seed = 490 for reproducibility**;

Please past your R code and randomly generated data output below:

# Generate a pre-set covariance matrix

S <- matrix(c(4,3,2,3,9,5,2,5,36),3,3)

row\_names <- c("x1","x2","x3")

col\_names <- c("x1","x2","x3")

rownames(S) <- row\_names

colnames(S) <- col\_names

# Create a vector containing the means

m1 <- c(2,4,6)

x <- mvrnorm(10,m1,S) # A randomly generated set

Text

Description automatically generated

1. Use relevant R functions to compute Expected Value (ie., mean), Median, Skewness and Kurtosis using the data generated for x1, x2 and x3 in Problem (a), respectively. Please past your R code for computing each statistic, and corresponding R output below:

mean(x) # Calculate the Expected Value

median(x) # Calculate the median

skewness(x) # Calculate the skewness of the random variables

kurtosis(x) # Calculate the kurtosis of the random variables

Text

Description automatically generated

1. Use relevant R functions to compute the correlation and covariance matrix of x1, x2 and x3, with the data generated from Problem (a). Please paste your R code and output below:

cov(x) # Generate the covariance matrix of the random variables

Text

Description automatically generated