**CBA: Practice Problem Set 2**

**Topics: Sampling Distributions and Central Limit Theorem**

1. Examine the following normal Quantile plots carefully. Which of these plots indicates that the data …
2. Are nearly normal?

Ans-A normal quantile plot with a straight line suggests that the data are nearly normal. If the points closely follow a straight line, it indicates that the data are approximately normally distributed.

1. Have a bimodal distribution? (One way to recognize a bimodal shape is a “gap” in the spacing of adjacent data values.)

Ans- A bimodal distribution would manifest as distinct peaks or modes in the data. One way to recognize a bimodal distribution in a normal quantile plot is to observe nonlinearity or curvature in the plot, suggesting two separate groups of data points

1. Are skewed (i.e. not symmetric) ?

Ans- Skewness in a distribution is reflected in a non-constant slope in the normal quantile plot. If one tail of the plot is longer or extends further than the other, it suggests skewness in the data.

1. Have outliers on both sides of the center?

Ans- Outliers in a dataset would appear as points that significantly deviate from the expected pattern of the normal quantile plot. If there are outliers on both sides of the center, it suggests that there are extreme values present in both the lower and upper tails of the distribution



1. For each of the following statements, indicate whether it is True/False. If false, explain why.

The manager of a warehouse monitors the volume of shipments made by the delivery team. The automated tracking system tracks every package as it moves through the facility. A sample of 25 packages is selected and weighed every day. Based on current contracts with customers, the weights should have *μ* = 22 lbs. and *σ* = 5 lbs.

Ans-True

1. Before using a normal model for the sampling distribution of the average package weights, the manager must confirm that weights of individual packages are normally distributed.

Ans=True. Before using a normal model for the sampling distribution of the average package weights . it's essential to confirm that the weights of individual packages are normally distributed

1. The standard error of the daily average SE() = 1.

Ans=False

1. Auditors at a small community bank randomly sample 100 withdrawal transactions made during the week at an ATM machine located near the bank’s main branch. Over the past 2 years, the average withdrawal amount has been $50 with a standard deviation of $40. Since audit investigations are typically expensive, the auditors decide to not initiate further investigations if the mean transaction amount of the sample is between $45 and $55. What is the probability that in any given week, there will be an investigation?
2. 1.25%
3. 2.5%
4. 10.55%
5. 21.1%
6. 50%

Ans= D) 21.1%

1. The auditors from the above example would like to maintain the probability of investigation to 5%. Which of the following represents the minimum number transactions that they should sample if they do not want to change the thresholds of 45 and 55? Assume that the sample statistics remain unchanged.
2. 144
3. 150
4. 196
5. 250
6. Not enough information

Ans=E) Not enough information

1. An educational startup that helps MBA aspirants write their essays is targeting individuals who have taken GMAT in 2012 and have expressed interest in applying to FT top 20 b-schools. There are 40000 such individuals with an average GMAT score of 720 and a standard deviation of 120. The scores are distributed between 650 and 790 with a very long and thin tail towards the higher end resulting in substantial skewness. Which of the following is likely to be true for randomly chosen samples of aspirants?
2. The standard deviation of the scores within any sample will be 120.
3. The standard deviation of the mean of across several samples will be 120.
4. The mean score in any sample will be 720.
5. The average of the mean across several samples will be 720.
6. The standard deviation of the mean across several samples will be 0.60

Ans- D) The average of the mean across several samples will be 720. The average of the means across several samples, if computed, is expected to be very close to the population mean of 720 due to the Central Limit Theorem.