**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?
3. Have a bimodal distribution? (One way to recognize a bimodal shape is a “gap” in the spacing of adjacent data values.)
4. Are skewed (i.e. not symmetric) ?
5. Have outliers on both sides of the center?



SOLN:

I. Are nearly normal? Plot C.

II. Have a bimodal distribution? Plot B.

III. Are skewed (i.e. not symmetric) ? Plot A,C,D.

IV. Have outliers on both sides of the center? Plot A.

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.

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

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.

**True.** The Central Limit Theorem (CLT) states that the sampling distribution of the sample mean is approximately normally distributed if the sample size is large enough. In this case, the sample size is 25, which is considered to be large enough. However, it is still important to check if the weights of individual packages are normally distributed. If they are not normally distributed, then the sampling distribution of the sample mean may not be normally distributed either.

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

**True**.

((*σ/251/2*) = 5/5 = 1)

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%

The answer is D. 21.1%.

The auditors will not initiate an investigation if the mean transaction amount of the sample is between $45 and $55. In other words, the auditors will initiate an investigation if the mean transaction amount of the sample is less than $45 or greater than $55.

The z-scores for the lower and upper bounds of the investigation range are:

z\_lower = (45 - 50) / 40 = -1.25

z\_upper = (55 - 50) / 40 = 1.25

The probability that the sample mean will be in the investigation range is:

P(-1.25 < z < 1.25) = 0.7887

Subtracting 1 from this value gives the probability that the sample mean will be outside the investigation range, which is the probability that an investigation will be initiated:

1 - 0.7887 = 0.2113 = 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

E. Not enough information.

To determine the minimum sample size, we need to know the following:

* The population mean, μ
* The population standard deviation, σ
* The desired confidence level, 1 - α
* The desired margin of error, ε

In the above example, we only know the desired confidence level and the desired margin of error. We do not know the population mean, σ, or the number of transactions in the population.

Therefore, we cannot determine the minimum sample size without more 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

Option D is correct. The average of the means across several random samples will approach the population mean 720.