**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?



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.

Anns:-  True.

The standard error of the mean (SE) is given by the population standard deviation (σ) divided by the square root of the sample size (n),

or SE = σ/sqrt(n).

In this case, σ = 5 lbs and n = 25

so SE = 5/sqrt(25) = 1 lb.

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

Ans:-

False.

The statement is incomplete as it does not specify the units for the standard error. The standard error (SE) is the standard deviation of the sampling distribution of the sample mean and is a measure of the precision of the sample mean as an estimate of the population mean.

The standard error can be calculated using the formula:

SE = σ/√n

where σ is the population standard deviation, n is the sample size, and √n is the square root of the sample size.

Based on the given information, σ = 5 lbs and n = 25. Substituting these values in the formula, we get:

SE = 5/√25 = 1

So, if the units are indeed pounds, then the statement is true. However, if the units are something else, such as grams or kilograms, then the standard error would be a different value.

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%

To solve this problem, we need to calculate the probability that the sample mean falls outside the range of $45 to $55.

The sample mean follows a normal distribution with mean μ = $50 and standard deviation σ = $40/√100 = $4.

The probability that the sample mean is less than $45 is:

Z = (45 - 50) / 4 = -1.25

Using a standard normal distribution table or calculator, we can find that the probability of Z being less than -1.25 is 0.1056.

The probability that the sample mean is greater than $55 is:

Z = (55 - 50) / 4 = 1.25

Using the same standard normal distribution table or calculator, we can find that the probability of Z being greater than 1.25 is also 0.1056.

Therefore, the probability that there will be an investigation is the sum of these two probabilities:

P(investigation) = P(sample mean < $45) + P(sample mean > $55) = 0.1056 + 0.1056 = 0.2112

So the answer is 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:- To maintain the probability of investigation at 5%, we need to find the sample size such that the probability of the sample mean being outside the range of $45 to $55 is 2.5% on each tail.

Using a standard normal distribution table or calculator, we can find the Z-score for a tail probability of 2.5%:

Z = 1.96

We can then use the formula for the standard error of the sample mean:

SE = σ/√n

where σ = $40, n is the sample size, and SE = $5 (half the width of the range of $45 to $55).

Substituting these values, we get:

$5 = $40/√n \* 1.96

Solving for n, we get:

n = (40/5\*1.96)^2 = 152.1

Rounding up to the nearest integer, we get a sample size of 153.

Therefore, the answer is not listed, and E. Not enough information is the closest option.

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:- B. The standard deviation of the mean across several samples will be 120/√n.

The central limit theorem states that, for large enough sample sizes, the distribution of the sample means approaches a normal distribution regardless of the shape of the population distribution. Therefore, we can assume that the sample means of GMAT scores from several random samples will follow a normal distribution.

The standard deviation of the sample means (standard error) can be calculated using the formula:

SE = σ/√n

where σ is the population standard deviation, n is the sample size, and SE is the standard deviation of the sample means.

Substituting the given values, we get:

SE = 120/√n

Therefore, option B is likely to be true. The standard deviation of the mean across several samples will be 120/√n.