**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 – C

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

Ans – B

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

Ans – A,C,D

1. Have outliers on both sides of the center?

Ans - 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.

Ans - False. The statement is false. The manager does not need to confirm that the weights of individual packages are normally distributed in order to use a normal model for the sampling distribution of the average package weights. According to the Central Limit Theorem, when the sample size is sufficiently large (typically, greater than or equal to 30), the sampling distribution of the sample mean will be approximately normal, regardless of the shape of the population distribution. In this case, the sample size is 25, which is reasonably large enough to assume that the sampling distribution of the sample mean will be approximately normal, even if the weights of individual packages are not normally distributed.

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

Ans - False. The statement is false. The standard error of the daily average, denoted by SE(x̅), is not necessarily equal to 1. The standard error represents the standard deviation of the sampling distribution of the sample mean and is calculated as the population standard deviation divided by the square root of the sample size (n). In this case, the population standard deviation (σ) is given as 5 lbs, but the sample size (n) is not specified. Without the specific value of the sample size, we cannot determine the standard error of the daily average. The formula for the standard error is SE(x̅) = σ / sqrt(n), where σ is the population standard deviation and n is the sample size.

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 - To calculate the probability of an investigation, we need to determine the probability that the mean transaction amount falls outside the range of $45 to $55.

Given that the population mean withdrawal amount is $50 and the standard deviation is $40, we can use the Central Limit Theorem to approximate the sampling distribution of the sample mean. The Central Limit Theorem states that for a large enough sample size, the sampling distribution of the sample mean will be approximately normally distributed, regardless of the shape of the population distribution.

In this case, the sample size is 100, which is reasonably large. Therefore, we can assume that the sampling distribution of the sample mean will be approximately normally distributed.

To calculate the probability of an investigation, we need to find the probability that the sample mean falls outside the range of $45 to $55. We can do this by calculating the z-scores for these values and then using the standard normal distribution (Z-distribution) to find the probabilities.

First, let's calculate the z-scores for $45 and $55 using the formula:

z = (x - μ) / (σ / sqrt(n))

where:

x = sample mean ($45 or $55)

μ = population mean ($50)

σ = population standard deviation ($40)

n = sample size (100)

For $45:

z = (45 - 50) / (40 / sqrt(100)) = -0.5

For $55:

z = (55 - 50) / (40 / sqrt(100)) = 0.5

Next, we can use a standard normal distribution table or a calculator to find the probabilities associated with these z-scores. The probability of an investigation can be calculated as the sum of the probabilities of falling outside the range of $45 to $55.

P(outside range) = P(z < -0.5) + P(z > 0.5)

Using a standard normal distribution table or calculator, we find:

P(z < -0.5) ≈ 0.3085

P(z > 0.5) ≈ 0.3085

Therefore,

P(outside range) = 0.3085 + 0.3085 = 0.617

So, the probability of an investigation in any given week is approximately 0.617, or 61.7%.

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 a probability of investigation at 5% and keep the thresholds of $45 and $55 without changing the sample statistics (mean and standard deviation), we need to find the minimum sample size that ensures the probability falls below or equal to 5%.

To determine the minimum sample size, we can use the z-score corresponding to a 5% probability from the standard normal distribution, which is approximately 1.645. This z-score represents the number of standard deviations away from the mean that captures the desired probability.

The formula to calculate the minimum sample size (n) is:

n = (z^2 \* σ^2) / (E^2)

Where:

z = z-score corresponding to the desired probability (5% = 0.05), approximately 1.645

σ = population standard deviation

E = maximum allowable error (half the width of the confidence interval)

In this case, the maximum allowable error is half the width of the interval between $45 and $55, which is ($55 - $45) / 2 = $5.

Plugging the values into the formula:

n = (1.645^2 \* 40^2) / (5^2)

n = (2.706025 \* 1600) / 25

n = 43.2964

Since we cannot have a fraction of a transaction, we need to round up to the nearest whole number. Therefore, the minimum number of transactions that should be sampled to maintain a 5% probability of investigation is 44.

Hence, the auditors should sample at least 44 transactions.

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.

Ans - True. The standard deviation within any sample will be 120. This is because the standard deviation represents the variability or spread of the individual data points within a sample. In this case, the individual GMAT scores have a standard deviation of 120, and this value will remain the same within any randomly chosen sample.

1. The standard deviation of the mean of across several samples will be 120.

Ans - False. The standard deviation of the mean across several samples, also known as the standard error, will be smaller than 120. The standard error represents the variability or spread of the sample means. It is calculated as the population standard deviation divided by the square root of the sample size. As the sample size increases, the standard error decreases. Therefore, the standard deviation of the mean across several samples will be smaller than 120.

1. The mean score in any sample will be 720.

Ans - False. The mean score in any sample is expected to be close to 720 but may not be exactly 720. The sample mean is an estimate of the population mean, and due to random sampling variability, it is unlikely to be exactly equal to the population mean. However, as the sample size increases, the sample mean becomes a more accurate estimate of the population mean.

1. The average of the mean across several samples will be 720.

Ans - True. The average of the mean across several samples is expected to be very close to 720. When randomly sampling from a population, the sample means will vary around the population mean. However, on average, the means of different samples will converge to the population mean. Therefore, the average of the mean across several samples is expected to be approximately 720.

1. The standard deviation of the mean across several samples will be 0.60

Ans - False. The standard deviation of the mean across several samples, also known as the standard error, will not be 0.60. The standard error is influenced by the sample size and the variability of the population. It is not a fixed value and cannot be determined without knowing the sample size and the population standard deviation. Therefore, we cannot conclude that the standard deviation of the mean across several samples will be 0.60 based on the given information.