**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?

**Answer: Only the C quantile plot shows that the data is nearly normal.**

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

**Answer: B and D quantile plot have a normal distribution.**

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

**Answer: A,B and D are skewed.**

1. Have outliers on both sides of the center?

**Answer: A and B 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.

**Answer:**

**A sampling distribution is a probability distribution of a statistic obtained from a larger number of samples drawn from a specific population. In our case the samples contain 25 packages and the larger number of samples contain of each such 25 packages taken into different samples (25+25+25+25…and so on). The mean for one these samples is 22lbs and standard deviation of 5lbs which means each individual package is having a weight varying between + or – 5lbs with respect to mean(22lbs). Hence it is invalid to take a weight of individual packages and confirm that it follows normal distribution before using a normal model for the sampling distribution. The Sample Central Limit Theorem states that the sampling distribution of the samples mean approaches normal distribution as the sample size is large enough.**

**Therefore, the statement given above is False.**

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

**Answer:**

**The standard error (SE) of the daily average is calculated as the standard deviation of the sample() divided by the square root of the sample size(n).**

**SE(x̅) = σ / √n**

**In this case the standard deviation of the population () is given as 5 Lbs. and the sample size (n) is 25 packages.**

**Calculation of standard error :**

**SE(x̅) = 5/25**

**= 5/5**

**= 1**

**Therefore, the statement given above is True.**

1. Auditors at a small community bank randomly sampled 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%

**Answer:**

**Given :**

**population mean (µ) = $50**

**population standard deviation () = $40**

**sample size (n) = 100**

**Sample mean follows a normal distribution.**

**The standard error of the sample mean (SE) is given by :**

**SE = /√n**

**= 40/√100**

**= 40/10**

**= 4**

**Now calculate the Z-scores for the lower and upper bounds of the range:**

**Z lower = (lower bound - µ)/ SE**

**= (45-50)/4**

**= - 1.25**

**Z upper = (upper bound - µ)/SE**

**= (55-40)/4**

**= 1.25**

**Finding the probabilities associated with these Z-scores using a standard normal distribution table or calculator :**

**P(Z < -1.25) ≈ 0.1056 (rounded to for decimal places)**

**P(Z < 1.25) ≈ 0.1056 (rounded to for decimal places)**

**Finally, we find the probability that the sample mean falls outside the range $45 to $55:**

**P(outside range) = P(Z < -1.25) + P(Z > 1.25)**

**= 0.1056 + 0.1056**

**= 0.2112**

**The probability that the sample mean falls outside the range $45 to $55 is 0.2112, which is equivalent to 21.12%.**

**Therefore, the correct answer is option 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

**Answer:**

**The correct answer is option E.**

**As provided answer option doesn’t include the calculated minimum sample size of 171.**

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.

**Answer: The statement given above is False.**

**Here the standard deviation of the population is 120, the standard deviation of individual samples will likely be different from 120. When you take a sample from a population, the standard deviation of that sample will depend on the variability of the scores within that specific sample, which can vary from one sample to another.**

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

**Answer: The statement given above is False.**

**Here the standard deviation of the mean across several samples, often referred to as the standard error of the mean (SE), is given by SE = σ / √n, where σ is the population standard deviation (120) and n is the sample size. As you take more samples, the standard error of the mean will decrease as the square root of the sample size. The larger the sample size, the smaller the standard error, and the more precise the estimates of the population mean will be.**

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

**Answer:** **The statement given above is False.**

**The standard deviation of the mean across several samples, often referred to as the standard error of the mean (SE), is given by SE = σ / √n, where σ is the population standard deviation (120) and n is the sample size. As you take more samples, the standard error of the mean will decrease as the square root of the sample size. The larger the sample size, the smaller the standard error, and the more precise the estimates of the population mean will be**

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

**Answer: The statement given above is True.**

**The average (or mean) of the means across several samples, also known as the grand mean, is expected to be close to the population mean (720). As the sample size increases and the number of samples becomes larger, the grand mean will tend to converge to the population mean**.

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

**Answer: The statement given above is False.**

**The standard deviation of the mean across several samples, as mentioned in option B, is the standard error of the mean (SE), not 0.60. The value of the standard error depends on the population standard deviation (120) and the sample size (n), as given by SE = σ / √n. The standard error will be different from 0.60, which appears to be an arbitrary value.**