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



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.

A sampling distribution is a probability distribution that represents the statistical outcomes of a specific sample statistic obtained from multiple samples drawn from a particular population. In our scenario, each sample consists of 25 packages, and there are numerous such samples, each comprising 25 different packages (25 packages in the first sample, 25 in the second, and so on). The mean weight for one of these samples is 22 pounds, with a standard deviation of 5 pounds. This implies that the individual package weights within each sample vary around the mean weight of 22 pounds, with fluctuations of up to ±5 pounds.

It's important to note that it may not be necessary to assume that the individual package weights follow a perfectly normal distribution before applying a normal model to the sampling distribution. The Sample Central Limit Theorem comes into play here. This theorem states that as the sample size becomes sufficiently large, the sampling distribution of the sample means will tend to follow a normal distribution, regardless of the distribution of individual package weights.

In summary, the key idea is that while individual package weights may not follow a normal distribution, when you aggregate these weights by taking the mean of each sample of 25 packages, the resulting sampling distribution of means will approach a normal distribution as the sample size grows. This property of the Sample Central Limit Theorem allows us to use normal distribution assumptions for making statistical inferences about the sample means, even when the underlying individual package weights may not be normally distributed.

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

As SE(Standard Error) = sample standard deviation / Square root of (number of sample) SE = 5 / (25)^1/2 SE = 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%
7. 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.
8. 144
9. 150
10. 196
11. **250**
12. Not enough information
13. 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?
14. The standard deviation of the scores within any sample will be 120.
15. The standard deviation of the mean of across several samples will be 120.
16. The mean score in any sample will be 720.
17. The average of the mean across several samples will be 720.
18. **The standard deviation of the mean across several samples will be 0.60**

Standard error = sigma / (n)^0.5 = standard deviation / (sample size)^0.5 = 120 / (40000)^0.5 = 0.6