**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. **False**

A statistical probability distribution derived from a greater number of samples taken from a given population is known as a sampling distribution. In our instance, there are 25 packages in each sample, and there are more samples (25+25+25+25, etc.) that contain each of the 25 packages. These samples have a mean weight of 22 pounds and a standard deviation of 5 pounds, meaning that the weight of each package varies by + or – 5 pounds relative to the mean of 22 pounds. Therefore, before using a normal model for the sampling distribution, it is invalid to take the weight of each individual package and verify that it follows a normal distribution. According to the Sample Central Limit Theorem, when the sample size is sufficiently large, the sampling distribution of the sample mean approaches normal distribution.

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

Since sample standard deviation / square root of (number of sample) equals SE(Standard Error), SE is equal to 5 / (25)^1/2. SE is equal to 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% Ans**

Since the standard deviation is not provided for the long term, use the formula

t=(x - μ)/σ/√n; t-test = (45-50) or (55-50)/40/√100 = ±5/40/√100 = ±1.25.

Since there is a 0.7857 chance of z falling between those two numbers, the likelihood of an investigation is 1-0.7887, or 21.1%

1. 50%
2. 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.
3. 144
4. 150
5. 196
6. **250 Ans**

T-value = (x\_bar – μ)/(s / √n) is ±1.96 for 5%.

√n = (40 \* t-value) / (5)

**n=245**

1. Not enough information
2. 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?
3. The standard deviation of the scores within any sample will be 120.
4. The standard deviation of the mean of across several samples will be 120.
5. The mean score in any sample will be 720.
6. The average of the mean across several samples will be 720.
7. **The standard deviation of the mean across several samples will be 0.60**

Standard error = σ / (n)^0.5 = σ/ (sample size)^0.5 = 120 / (40000)^0.5 = 0.6