**CBA: Practice Problem Set 2**

**Topics: Sampling Distributions and Central Limit Theorem**

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: TRUE.**

at least 30 sample must be selected and weighed every day. Based on CLT, sampling distribution of the sample mean approach normal distribution as the sample size become more (i.e.>30).

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

**Ans: TRUE.** x ̅ = Mean

SE = Standard deviation divided by square root of sample size

SE = 5/sqrt(25)

SE (x ̅) = 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: D. 250** *(Approximately)*

Theprobability of Z between 45 & 55 was found 0.7887 and probability of an investigation is 1-0.7887 = 0.2113 or 21.1% that we had calculated in previous example.

Now we have to calculate it for **5%,** z-value will be + / -1.96

So, 1.96 = (5)\* Sqrt(n)/40

Sqrt(n) = 15.68

n = 246.

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?

SEM = SD/Sqrt(n) = 120/Sqrt(40,000) = 120/200 = 0.6

1. The standard deviation of the scores within any sample will be 120.

**Ans: FALSE.** SD won’t be 120 of scores in any one sample as we don't know the sample size.

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

**Ans: FALSE.** SD of the mean of across several samples won’t also be 120. It will be less, certainly, probably about 0.6.

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

**Ans: TURE,** but it may or may not be less/more.

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

**Ans: TRUE.** This is certainly possible, but it requires mean of all samples that sample size which will be the case.

Data is distributed with leptokurtic kurtosis mean more information at the center and lesser information at the tail. This mean there is higher chance that average of mean of an aspirant that randomly chosen will be 720 that fall in between 650 and 790 at the center.

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

**Ans: TRUE.** The SEM will be 0.60. This is likely, given the sample size, which even with a lot of skewness will tend towards normality given the sample size. The mean would have an expected value of 720, but in calculations, the SEM is 0.6