**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:** **C**

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

**Answer**: **D**

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

**Answer**: **A**

1. Have outliers on both sides of the center?

**Answer**: **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.

**Answer**: **FALSE.**

The statement is false because the individual values do not need to be normally distributed as long as the sample size condition is satisfied.

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

**Answer**: TRUE.

The standard error of the mean is given by the formula:

SE(X) = σ/

For this example, this value is equal to SE(X) = 5/25 =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%

**Answer**:

n = 100,

Population mean = 50

Population Standard Deviation = 40. As no. of samples is more than 30, we can consider it normal distribution.

Probability that investigation will not take place = P(45<X<55)

Probability that investigation will take place = 1 - P(45<X<55)

Z = (Sample mean – Population mean) / [Population Stddev / Sqrt (n)]

P(45<X<55) = { P where Z= = 1.25} - { P where Z= = -1.25 }

= 0.89435 - 0.10565

P(45<X<55) = 0.7887

Probability that investigation will take place = 1 - P(45<X<55)

= 1 – 0.7887

**Probability that investigation will take place** = 0.2113 = **21.1%**

* **Python Code:**



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**:

Probability of investigation to occur should be 5%

So, Probability to not investigate = 1-0.05 = 0.95= 95%

Z at [(1-0.95)/2]=1.96

Sample Standard Deviation= 45-50/1.96=-5/1.96=-2.55

Sample Standard Deviation

=-15.68

**n=246.06 250**

**Python Code:**

****

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**: **FALSE**.

Since, sample size ‘n’ is not known, so sample standard deviation within any sample will not be 120.

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

**Answer**: **FALSE**.

Standard Deviation of mean is 0.6 so it cannot be 120.

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

**Answer**: **FALSE**.

The mean score can vary as it depends on sample data, and there can be chances of mean score other than 720 as sample mean is a random value and varies from sample to sample.

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

**Answer**: **TRUE**.

Average of means of samples can be around 720.

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

**Answer**: **TRUE**.

Standard Deviation of mean for various samples = Standard Deviation/

= 120/