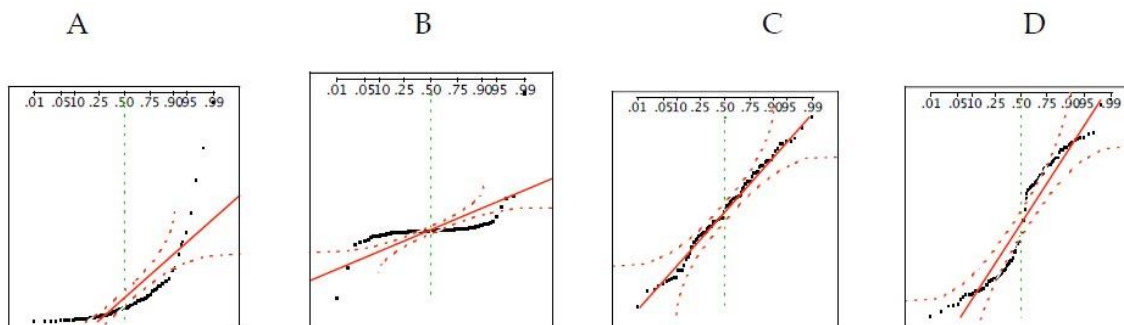


## CBA: Practice Problem Set 2

### Topics: Sampling Distributions and Central Limit Theorem

1 Examine the following normal Quantile plots carefully. Which of these plots indicates that the data.

- I. Are nearly normal? Ans => C
- II. Have a bimodal distribution? (One way to recognize a bimodal shape is a “gap” in the spacing of adjacent data values.) => B and D
- III. Are skewed (i.e. not symmetric)? => A, B and D
- IV. Have outliers on both sides of the center? => A and B



2. 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 are selected and weighed every day. Based on current contracts with customers, the weights should have  $\mu = 22$  lbs. and  $\sigma = 5$  lbs.

- (i) 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: False

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.

(ii) The standard error of the daily average  $SE(\bar{x}) = 1$ .

Ans: True :

As  $SE(\text{Standard Error}) = \text{sample standard deviation} / \text{Square root of (number of sample)}$

$$SE = 5 / (25)^{1/2}$$

$$SE = 1$$

5 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?

- A. The standard deviation of the scores within any sample will be 120.
- B. The standard deviation of the mean of across several samples will be 120.
- C. The mean score in any sample will be 720.
- D. The average of the mean across several samples will be 720.
- E. The standard deviation of the mean across several samples will be 0.60

Ans:

Option B is true because the standard deviation of the mean of several samples will be smaller than the standard deviation of the individual scores. This is known as the standard error of the mean, which is calculated as the population standard deviation divided by the square root of the sample size. Therefore, as the sample size increases, the standard error of the mean decreases, indicating that the mean of several samples will be more precise than the mean of a single sample.