

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 ...
- Are nearly normal?

ANS: C

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

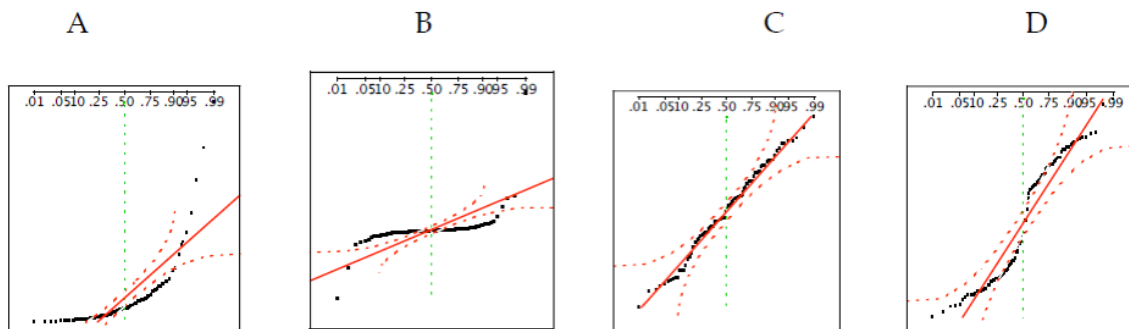
ANS: D

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

ANS: B

- Have outliers on both sides of the center?

ANS: A



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

- 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.

In this case, the sample size is 25, which can be considered enough for the CLT to apply. Therefore, even if the weights of individual packages are not normally distributed, the distribution of the average weights of samples of 25 packages will be approximately normal. This means that the manager can use a normal model to make inferences about the population mean based on the sample mean.

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

ANS: True.

3. 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?
- A. 1.25%
 - B. 2.5%
 - C. 10.55%
 - D. 21.1%
 - E. 50%

ANS: D. 21.1%, refer python file.

4. 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.
- A. 144
 - B. 150
 - C. 196
 - D. 250
 - E. Not enough information

ANS: D.250, refer python file.

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.

ANS: False. The standard deviation of a sample is a measure of how spread out the scores are in that sample. The standard deviation of the population is 120, but the standard deviation of any given sample will depend on the size of the sample and the variability of the scores within that sample.

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

ANS: False. The standard deviation of the mean is a measure of how much the means of different samples are likely to vary from the population mean. In this case, the standard deviation of the population is 120, so the standard deviation of the mean for a sample size of n would be $120/\sqrt{n}$.

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

ANS: False. The mean score in any sample will not necessarily be 720. The mean of a sample is the average of the scores in that sample. The mean of the population is 720, but the mean of any given sample will depend on the scores in that sample.

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

ANS: True.

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

ANS: False. The standard deviation of the mean is a measure of how much the means of different samples are likely to vary from the population mean. In this case, the standard deviation of the population is 120, so the standard deviation of the mean for a sample size of n would be $120/\sqrt{n}$.