AP Statistics Concept Review: Sampling Distributions

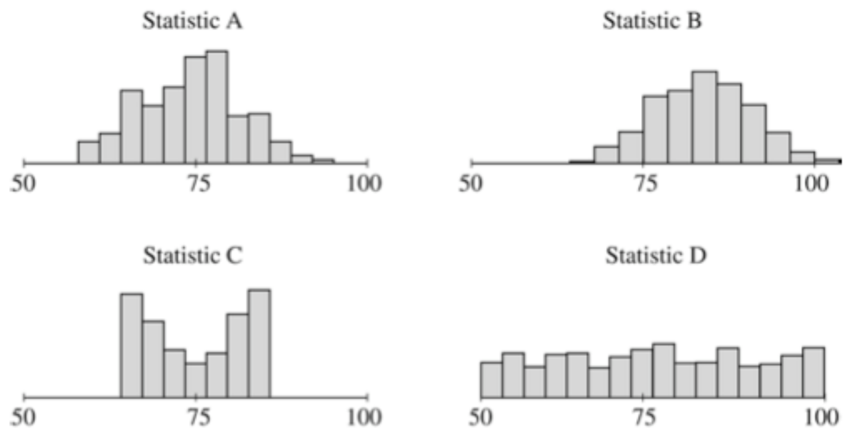
1. A local radio station plays 40 rock-and-roll songs during each 4-hour show. The program director at the station needs to know the total amount of airtime for the 40 songs so that time can also be programmed during the show for news and advertisements. The distribution of the lengths of rock-and-roll songs, in minutes, is roughly symmetric with a mean length of 3.9 minutes and a standard deviation of 1.1 minutes.

   (a) Describe the sampling distribution of the sample mean song lengths for random samples of 40 rock-and-roll songs.

   (b) If the program manager schedules 80 minutes of news and advertisements for the 4-hour (240-minute) show, only 160 minutes are available for music. Approximately what is the probability that the total amount of time needed to play 40 randomly selected rock-and-roll songs exceeds the available airtime?

2. The distribution of the number of gallons of water used for taking a shower in the United States is skewed right. Some people take long showers using as much as 63 gallons compared to the average of 17.2 gallons. A conservationist is interested in the mean number of gallons of water used in college apartments. Assuming a standard deviation of 4.2 gallons for all US showers, what is the probability of a random sample of 30 college student showers with a mean of more than 19 gallons of water?

3. Four different statistics have been proposed as estimators of a population parameter. To investigate the behavior of these estimators, 500 random samples are selected from a known population and each statistic is calculated for each sample. The true value of the population parameter is 75. The graphs below show the distribution of values for each statistic.



(a) Which of the statistics appear to be unbiased estimators of the population parameter? How can you tell?

(b) Which of statistics A or B would be a better estimator of the population parameter? Explain your choice.

(c) Which of statistics C or D would be a better estimator of the population parameter? Explain your choice.

ANSWERS:

1.

Part (a):

The sampling distribution of the sample mean song length has mean $\mu_{\bar{X}} = \mu = 3.9$ minutes and standard deviation $\sigma_{\bar{X}} = \frac{\sigma}{\sqrt{n}} = \frac{1.1}{\sqrt{40}} \approx 0.174$ minutes. The central limit theorem (CLT) applies in this case because the sample size ($n = 40$) is fairly large, especially with the population of song lengths having a roughly symmetric distribution. Thus, the sampling distribution of the sample mean song length is approximately normal.

Part (b):

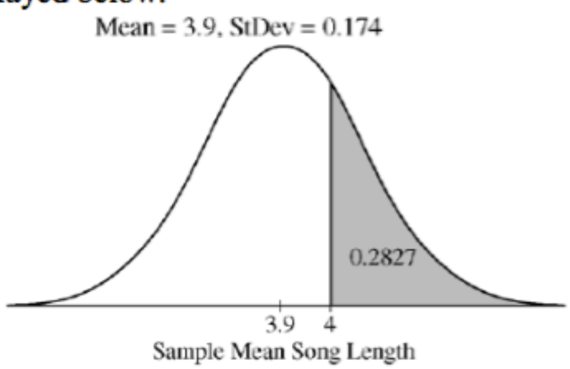
The probability that the total airtime of 40 randomly selected songs exceeds the available time (that is, the probability that the total airtime of 40 randomly selected songs is greater than 160 minutes) is equivalent to the probability that the sample mean length of the 40 songs is greater than $\frac{160}{40} = 4.0$ minutes.

According to part (a), the distribution of the sample mean length of $\bar{X}$ is approximately normal. Therefore,

$$P(\bar{X} > 4.0) \approx P\left(Z > \frac{4.0 - 3.9}{0.174}\right) = P(Z > 0.57) = 1 - 0.7157 = 0.2843.$$

(The calculator gives the answer as 0.2827.)

The approximate sampling distribution of the sample mean song length and the desired probability are displayed below.



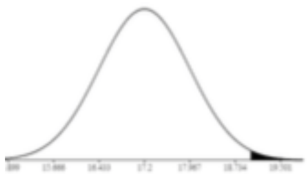
2.

Conditions:

1) Random sample of showers stated.

2) 10% Rule: There's certainly more than 10(30)=300 showers taken.

3) Normal Condition: The sample is large enough since n = 30 ≥ 30



$$\sigma_{\bar{x}} = \frac{4.2}{\sqrt{30}} = 0.767 \qquad z = \frac{19-17.2}{.767} = 2.347$$

$$P(\bar{x} > 19) = 0.00947$$

3. a) A, C and D are unbiased since their means are all ≈ 75.

b) A since it is an unbiased estimator (its mean is ≈75). Statistic B is biased (its mean is ≈85)

c) C is better. Both C and D are unbiased (centered ≈75), but C has less variability than D.