

Part 1: Multiple Choice. Circle the letter corresponding to the best answer.

1. In a study of the effects of acid rain, a random sample of 100 trees from a particular forest is examined. Forty percent of the trees show some signs of damage. Which of the following statements is correct?
 - (a) 40% is a parameter
 - (b) 40% is a statistic
 - (c) 40% of all trees in the forest show some signs of damage
 - (d) More than 40% of the trees in the forest show some signs of damage
 - (e) Less than 40% of the trees in the forest show some signs of damage

2. The sampling distribution of a statistic is
 - (a) the probability that we obtain the statistic in repeated random samples.
 - (b) the mechanism that determines whether randomization was effective.
 - (c) the distribution of values taken by a statistic in all possible samples of the same sample size from the same population.
 - (d) the extent to which the sample results differ systematically from the truth.
 - (e) the distribution of values in a sample of size n from the population

3. A statistic is said to be *unbiased* if
 - (a) the survey used to obtain the statistic was designed so as to avoid even the hint of racial or sexual prejudice.
 - (b) the mean of its sampling distribution is equal to the true value of the parameter being estimated.
 - (c) both the person who calculated the statistic and the subjects whose responses make up the statistic were truthful.
 - (d) the value of the statistic from any sample is equal to the parameter being estimated.
 - (e) it is used for honest purposes only.

4. Which of the following distributions has a mean that varies from sample to sample?
 - I. The population distribution
 - II. The distribution of sample data
 - III. The sampling distribution
 - (a) I only
 - (b) II only
 - (c) III only
 - (d) II and III
 - (e) all three distributions

5. You take a simple random sample of size 25 from a very large population in which the true proportion is $p = 0.1$, thus violating the condition that $np \geq 10$ and $n(1 - p) \geq 10$. Which statement below best describes what you know about the sampling distribution of \hat{p} ?

(a) $\mu_{\hat{p}} = 0.1; \sigma_{\hat{p}} = \sqrt{\frac{(0.1)(0.9)}{25}}$; the distribution is not approximately Normal.

(b) $\mu_{\hat{p}} = 0.1; \sigma_{\hat{p}} = \sqrt{\frac{(0.1)(0.9)}{25}}$; the distribution is approximately Normal.

(c) $\mu_{\hat{p}} = 0.1$; we cannot use the formula $\sigma_{\hat{p}} = \sqrt{\frac{(0.1)(0.9)}{25}}$; the distribution is approximately Normal.

(d) $\mu_{\hat{p}} = 0.1$; we cannot use the formula $\sigma_{\hat{p}} = \sqrt{\frac{(0.1)(0.9)}{25}}$; the distribution is not approximately Normal.

(e) We do not know $\mu_{\hat{p}}$; $\sigma_{\hat{p}} = \sqrt{\frac{(0.1)(0.9)}{25}}$; the distribution is not approximately Normal.

6. The number of hours a light bulb burns before failing varies from bulb to bulb. The distribution of burnout times is strongly skewed to the right. The central limit theorem says that

(a) as we look at more and more bulbs, their mean burnout time gets ever closer to the mean μ for all bulbs of this type.

(b) the mean burnout time for any number of bulbs has a distribution of the same shape (strongly skewed) as the distribution for individual bulbs.

(c) the mean burnout time for any number of bulbs has a distribution that is close to Normal.

(d) the mean burnout time for a large number of bulbs has a distribution of the same shape (strongly skewed) as the distribution for individual bulbs.

(e) the mean burnout time for a large number of bulbs has a distribution that is close to Normal.

7. You take an SRS of size 500 from the 37,000 students at Purdue University and measure individual's heights. You then take an SRS of size 500 from the 4,400,000 adults in the state of Indiana and measure their heights. Assuming the standard deviation of individual heights in the two populations is the same, the standard deviation of the sampling distribution of mean heights for the Indiana sample is

(a) approximately the same as for the Purdue sample because both are samples of size 500.

(b) smaller than for the Purdue sample because the population of Indiana is much larger.

(c) larger than for the Purdue sample because the population of Indiana is much larger.

(d) larger, because the Indiana sample is smaller relative to the population from which it's been taken.

(e) either larger or smaller than for the Purdue sample because it varies from sample to sample.

8. The chipmunk population in a certain area is known to have a mean weight of 84 gm and a standard deviation of 18 gm. A wildlife biologist weighs 9 chipmunks that have been caught in live traps before releasing them. Which of the following best describes what we know about the sampling distribution of means for the biologist's sample? [Assume the 9 chipmunks represent a simple random sample of chipmunks in the area.]

- (a) $\mu_{\bar{x}} = 84$; $\sigma_{\bar{x}} = 18$; distribution approximately Normal
- (b) $\mu_{\bar{x}} = 84$; $\sigma_{\bar{x}} = 6$; shape of distribution unknown
- (c) $\mu_{\bar{x}} = 84$; $\sigma_{\bar{x}} = 6$; distribution approximately Normal
- (d) $\mu_{\bar{x}} = 84$; $\sigma_{\bar{x}}$ unknown; distribution approximately Normal
- (e) $\mu_{\bar{x}} = 84$; $\sigma_{\bar{x}}$ unknown; shape of distribution unknown

9. Interpupillary distance (IPD) is the distance between the centers of the pupils of a person's left and right eyes. In adult males IPD is approximately Normally distributed with a mean of 62.5 mm and a standard deviation of 6 mm. Suppose you take a simple random sample of 5 adult males. What is the probability that their mean IPD is greater than 60 mm?

- (a) $P\left(z > \frac{60 - 62.5}{6}\right)$
- (b) $P\left(z > \frac{62.5 - 60}{6}\right)$
- (c) $P\left(z > \frac{60 - 62.5}{\frac{6}{\sqrt{5}}}\right)$
- (d) $P\left(z < \frac{60 - 62.5}{\frac{6}{\sqrt{5}}}\right)$
- (e) $P\left(z > \frac{62.5 - 60}{\frac{6}{\sqrt{5}}}\right)$

10. A survey asks a simple random sample of 500 adults in Ohio if they support an increase in the state sales tax from 5% to 6%, with the additional revenue going to education. Let \hat{p} denote the proportion in the sample who say they support the increase. Suppose that 53% of *all* adults in Ohio support the increase. What is the probability that less than half the sample will say they support the increase?

- (a) $P\left(z < \frac{0.5 - 0.53}{\sqrt{(0.5)(0.5)(500)}}\right)$
- (b) $P\left(z > \frac{0.5 - 0.53}{\sqrt{(0.53)(0.47)(500)}}\right)$
- (c) $P\left(z > \frac{0.53 - 0.5}{\sqrt{\frac{(0.5)(0.5)}{500}}}\right)$
- (d) $P\left(z < \frac{0.5 - 0.53}{\sqrt{\frac{(0.53)(0.47)}{500}}}\right)$
- (e) $P\left(z > \frac{0.5 - 0.53}{\sqrt{\frac{(0.53)(0.47)}{500}}}\right)$

Part 2: Free Response

Show all your work. Indicate clearly the methods you use, because you will be graded on the correctness of your methods as well as on the accuracy and completeness of your results and explanations.

11. The weight of the eggs produced by a certain breed of hen is Normally distributed with mean 65 grams (g) and standard deviation 5 g.

(a) Calculate the probability that a randomly selected egg weighs between 62.5 g and 68.75 g. Show your work.

(b) Think of cartons of such eggs as SRSs of size 12 from the population of all eggs. Calculate the probability that the mean weight of the eggs in a carton falls between 62.5 g and 68.75 g. Show your work.

(c) Did you need to know that the population distribution of egg weights was Normal in order to complete parts (a) or (b)? Justify your answer.

12. Companies are interested in the demographics of those who listen to the radio programs they sponsor. A radio station has determined that only 20% of listeners phoning in to a morning talk program are male. The station management wonders if adding a male host to the program will increase the proportion of callers who are male. After adding the male host, the station records the gender of 200 people who phone in to the program during a particular week. The station is willing to view these 200 callers as an SRS from the population of all those who call in to this program.

(a) For the moment, assume that the addition of the male host has *no effect* on the proportion of callers who are male. If \hat{p} is the proportion of callers in the sample who are male, what are the mean and standard deviation of the sampling distribution of \hat{p} ?

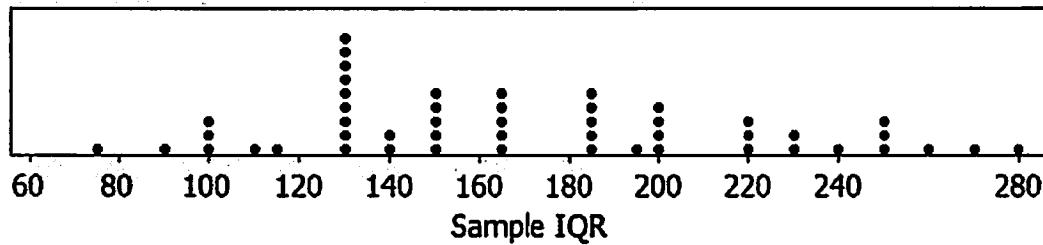
(b) What assumption are you making when you use the formula for the standard deviation of \hat{p} in this setting?

(c) In fact, during this particular week, 50 of the 200 callers were male. Does this provide sufficient evidence to suggest that the proportion of male callers has increased from 20%? Support your answer with an appropriate probability calculation. [Assume all necessary conditions have been met.]

13. Buying a year's worth of textbooks for college can be expensive! Consider a large population of college students for whom the distribution of the annual cost of textbooks is slightly skewed to the left. Here is the five-number summary for the annual cost of textbooks for this population:

Minimum = 80 Quartile 1 = 215 Median = 335 Quartile 3 = 380 Maximum = 440

Suppose we take random samples of size 32 from this population and calculate the interquartile range (IQR) for each of our samples. Below is a dotplot of the IQR from 50 such samples.



(a) Briefly explain what the dot at 240 represents.

(b) Is the sample IQR is an unbiased estimator of the population IQR? Justify your answer.