Chapter 7 AP® Statistics Practice Test

Section I: Multiple Choice *Select the best answer for each question.*

T7.1 A study of voting chose 663 registered voters at random shortly after an election. Of these, 72% said they had voted in the election. Election records show that only 56% of registered voters voted in the election. Which of the following statements is true about the boldface numbers?

(a) 72% is a sample; 56% is a population.
(b) 72% and 56% are both statistics.
(c) 72% is a statistic and 56% is a parameter.
(d) 72% is a parameter and 56% is a statistic.
(e) 72% and 56% are both parameters.

T7.2 The Gallup Poll has decided to increase the size of its random sample of voters from about 1500 people to about 4000 people right before an election. The poll is designed to estimate the proportion of voters who favor a new law banning smoking in public buildings. The effect of this increase is to

(a) reduce the bias of the estimate.
(b) increase the bias of the estimate.
(c) reduce the variability of the estimate.
(d) increase the variability of the estimate.
(e) reduce the bias and variability of the estimate.

T7.3 Suppose we select an SRS of size $n = 100$ from a large population having proportion p of successes. Let $\hat{p}$ be the proportion of successes in the sample. For which value of p would it be safe to use the Normal approximation to the sampling distribution of $\hat{p}$?

(a) 0.01 (b) 0.09 (c) 0.85 (d) 0.975 (e) 0.999

T7.4 The central limit theorem is important in statistics because it allows us to use the Normal distribution to find probabilities involving the sample mean

(a) if the sample size is reasonably large (for any population).
(b) if the population is Normally distributed and the sample size is reasonably large.
(c) if the population is Normally distributed (for any sample size).
(d) if the population is Normally distributed and the population standard deviation is known (for any sample size).
(e) if the population size is reasonably large (whether the population distribution is known or not).

T7.5 The number of undergraduates at Johns Hopkins University is approximately 2000, while the number at Ohio State University is approximately 60,000. At both schools, a simple random sample of about 3% of the undergraduates is taken. Each sample is used to estimate the proportion p of all students at that university who own an iPod. Suppose that, in fact, $p = 0.80$ at both schools. Which of the following is the best conclusion?

(a) The estimate from Johns Hopkins has less sampling variability than that from Ohio State.
(b) The estimate from Johns Hopkins has more sampling variability than that from Ohio State.
(c) The two estimates have about the same amount of sampling variability.
(d) It is impossible to make any statement about the sampling variability of the two estimates because the students surveyed were different.
(e) None of the above.

T7.6 A researcher initially plans to take an SRS of size n from a population that has mean 80 and standard deviation 20. If he were to double his sample size (to $2n$), the standard deviation of the sampling distribution of the sample mean would be multiplied by

(a) $\sqrt{2}$. (b) $1/\sqrt{2}$. (c) 2. (d) 1/2. (e) $1/\sqrt{2n}$.

T7.7 The student newspaper at a large university asks an SRS of 250 undergraduates, "Do you favor eliminating the carnival from the term-end celebration?" All in all, 150 of the 250 are in favor. Suppose that (unknown to you) 55% of all undergraduates favor eliminating the carnival. If you took a very large number of SRSs of size $n = 250$ from this population, the sampling distribution of the sample proportion $\hat{p}$ would be

(a) exactly Normal with mean 0.55 and standard deviation 0.03.
(b) approximately Normal with mean 0.55 and standard deviation 0.03.
(c) exactly Normal with mean 0.60 and standard deviation 0.03.
(d) approximately Normal with mean 0.60 and standard deviation 0.03.
(e) heavily skewed with mean 0.55 and standard deviation 0.03.

T7.8 Which of the following statements about the sampling distribution of the sample mean is *incorrect?*

(a) The standard deviation of the sampling distribution will decrease as the sample size increases.
(b) The standard deviation of the sampling distribution is a measure of the variability of the sample mean among repeated samples.
(c) The sample mean is an unbiased estimator of the population mean.
(d) The sampling distribution shows how the sample mean will vary in repeated samples.
(e) The sampling distribution shows how the sample was distributed around the sample mean.

T7.9 A machine is designed to fill 16-ounce bottles of shampoo. When the machine is working properly, the amount poured into the bottles follows a Normal distribution with mean 16.05 ounces and standard deviation 0.1 ounce. Assume that the machine is working properly. If four bottles are randomly selected and the number of ounces in each bottle is measured, then there is about a 95% chance that the sample mean will fall in which of the following intervals?

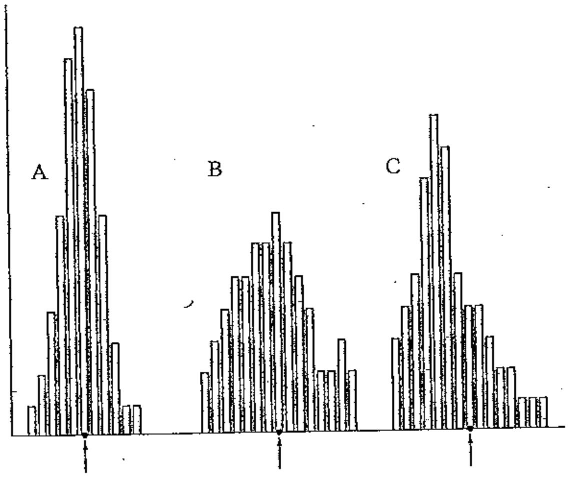
(a) 16.05 to 16.15 ounces
(b) 16.00 to 16.10 ounces
(c) 15.95 to 16.15 ounces
(d) 15.90 to 16.20 ounces
(e) 15.85 to 16.25 ounces

T7.10 Suppose that you are a student aide in the library and agree to be paid according to the "random pay" system. Each week, the librarian flips a coin. If the coin comes up heads, your pay for the week is \$80. If it comes up tails, your pay for the week is \$40. You work for the library for 100 weeks. Suppose we choose an SRS of 2 weeks and calculate your average earnings $\bar{x}$. The shape of the sampling distribution of $\bar{x}$ will be

(a) Normal.
(b) approximately Normal.
(c) right-skewed.
(d) left-skewed.
(e) symmetric but not Normal.

Section II: 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.*

T7.11 Below are histograms of the values taken by three sample statistics in several hundred samples from the same population. The true value of the population parameter is marked with an arrow on each histogram.



Which statistic would provide the best estimate of the parameter? Justify your answer.

T7.12 The amount that households pay service providers for access to the Internet varies quite a bit, but the mean monthly fee is \$38 and the standard deviation is \$10. The distribution is not Normal: many households pay a base rate for low-speed access, but some pay much more for faster connections. A sample survey asks an SRS of 500 households with Internet access how much they pay. Let $\bar{x}$ be the mean amount paid.

(a) Explain why you can't determine the probability that the amount a randomly selected household pays for access to the Internet exceeds \$39.
(b) What are the mean and standard deviation of the sampling distribution of $\bar{x}$?
(c) What is the shape of the sampling distribution of $\bar{x}$? Justify your answer.
(d) Find the probability that the average fee paid by the sample of households exceeds \$39. Show your work.

T7.13 According to government data, 22% of American children under the age of six live in households with incomes less than the official poverty level. A study of learning in early childhood chooses an SRS of 300 children. Find the probability that more than 20% of the sample are from poverty-level households. Be sure to check that you can use the Normal approximation.

AP Statistics Practice Test (page 468)

T7.1 c. The statistic is a measure of the sample and the parameter is a measure of the population.

T7.2 c. Sample size has no effect on the bias of an estimate, but larger samples will reduce the variability of an estimate.

T7.3 c. To use the Normal approximation, both np and $n(1-p)$ must be at least 10. In all options other than c, this condition is not met.

T7.4 a. The central limit theorem applies when the sample size is large ($n \geq 30$). The CLT is not needed when the original distribution is Normal; the distribution of the sample mean is always Normal in that case.

T7.5 b. Because a sample of 3% of the undergraduates from Ohio State University consists of approximately 1800 students whereas a sample of 3% of the undergraduates from Johns Hopkins consists of just 60 students, the estimate from Ohio State University will have less sampling variability.

T7.6 b. Because the standard deviation of the sampling distribution of the sample mean is $\sigma_{\bar{x}} = \frac{\sigma}{\sqrt{n}}$, doubling the sample size gives $\sigma_{\bar{x}} = \frac{\sigma}{\sqrt{2n}} = \frac{1}{\sqrt{2}}\frac{\sigma}{\sqrt{n}}$.

T7.7 b. The sampling distribution would be only approximately Normal with mean equal to the population proportion (0.55 in this case) and standard deviation equal to $\sqrt{\frac{(0.55)(0.45)}{250}} = 0.03$.

T7.8 e. The sampling distribution has information about how the sample mean varies from sample to sample, not what any sample itself looks like.

T7.9 c. $\mu_{\bar{x}} = \mu = 16.05$ and $\sigma_{\bar{x}} = \frac{\sigma}{\sqrt{n}} = \frac{0.1}{\sqrt{4}} = 0.05$. About 95% of the sample means should be within two standard deviations of the mean, which is $16.05 \pm 2(0.05) = 15.95$ to 16.15.

T7.10 e. The distribution of the average amount of pay will not be Normal because there are only three possible outcomes, $\bar{x} = 40$, 60, or 80. $\bar{x} = 40$ will occur 25% of the time, as will $\bar{x} = 80$. $\bar{x} = 60$ will happen 50% of the time.

T7.11 Sample statistic A provides the best estimate of the parameter. Both statistics A and B appear to be unbiased, while statistic C appears to be biased because the center of its sampling distribution is smaller than the value of the parameter. In addition, statistic A has lower variability than statistic B. In this situation, we want low bias and low variability, so statistic A is the best choice.

T7.12 (a) The probability that a single household pays more than \$39 cannot be calculated, because we do not know the shape of the population distribution of monthly fees.
(b) The mean of the sampling distribution of the sample mean is equal to the mean of the population distribution. Therefore $\mu_{\bar{x}} = \mu = \$38$. Also we know that $\sigma_{\bar{x}} = \frac{\sigma}{\sqrt{n}} = \frac{10}{\sqrt{500}} = \$0.4472$ because the sample of size 500 is less than 10% of all households with Internet access.
(c) Because the sample size is large ($n = 500 \geq 30$), the distribution of $\bar{x}$ will be approximately Normal.
(d) **Step 1: State the distribution and values of interest.** We want to find $P(\bar{x} > 39)$ using the $N(38, 0.4472)$ distribution. **Step 2: Perform calculations. Show your work.** The standardized score for the boundary value is $z = \frac{39-38}{0.4472} = 2.24$. The desired probability is $P(Z > 2.24) = 1 - 0.9875 = 0.0125$. *Using technology*: normalcdf(lower: 39, upper: 1000, μ : 38, σ : 0.4472) = 0.0127. **Step 3: Answer the question.** There is a 0.0127 probability that the mean monthly fee exceeds \$39.

T7.13 **Step 1: State the distribution and values of interest.** We have an SRS of size 300 drawn from a population in which the proportion who live in households with incomes under the poverty line is $p = 0.22$. Thus, $\mu_{\hat{p}} = p = 0.22$. Because 300 is less than 10% of children under the age of six, $\sigma_{\hat{p}} = \sqrt{\frac{0.22(0.78)}{300}} = 0.0239$. Because $np = 300(0.22) = 66$ and $n(1-p) = 300(0.78) = 234$ are both at least 10, the sampling distribution of $\hat{p}$ can be approximated by a Normal distribution. We want to find $P(\hat{p} > 0.20)$ using the $N(0.22, 0.0239)$ distribution. **Step 2: Perform calculations. Show your work.** The standardized score for the boundary value is $z = \frac{0.20-0.22}{0.0239} = -0.84$. The desired probability is $P(Z > -0.84) = 1 - 0.2005 = 0.7995$ *Using technology*: normalcdf (lower: 0.20, upper: 1000, μ : 0.22, σ : 0.0239) = 0.7987. **Step 3: Answer the question.** There is a 0.7987 probability that more than 20% of the sample are from poverty-level households.