AP Statistics
Chapter 7 Practice Test

Name Key
Date ________ Hour ________

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.

$\hat{p} = .72$ statistic
$p = .56$ parameter

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) have no effect since the population size is the same.

Sample size has no effect on bias of an estimate, but larger samples will reduce variability of an estimate

T7.3. Suppose we select an SRS of size $n = 100$ from a large population having proportion p of successes. Let ~~ôop~~ $\hat{p}$ 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 ~~ôop~~ $\hat{p}$?
(a) 0.01
(b) 1/11
(c) 0.85
(d) 0.975
(e) 0.999

C is the only one where $.1 \le p \le .9$

$np \ge 10$ and $n(1-p) \ge 10$
$100(p) \ge 10$ $\quad 100(1-p) \ge 10$
$p \ge \frac{10}{100}$ $\quad 1-p \ge \frac{10}{100}$
$p \ge \frac{1}{10}$ $\quad -p \ge -\frac{9}{10}$
$p \ge .1$ $\quad p \le .9$
so $.1 \le p \le .9$

T7.4. The central limit theorem is important in statistics because it allows us to use the Normal distribution to make inferences concerning the population 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 variance is known (for any sample size).
(e) if the population size is reasonably large (whether the population distribution is known or not).

CLT is not needed when original distribution is Normal: distribution of sample means is always Normal in this case

T7.5. The number of undergraduates at Johns Hopkins University is approximately 2000, while the number at Ohio State University is approximately 40,000. 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 since the students surveyed were different.
(e) None of the above.

The larger the sample size the smaller the variability

John Hopkins
$N = 2000$
SRS size $n = .03(2000)$
$n = 60$
$p = .80$
more variability

Ohio State
$N = 40,000$
SRS size $n = .03(40,000)$
$n = 1200$

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) $\frac{1}{\sqrt{2}}$

(c) 2

(d) $\frac{1}{2}$

(e) $\frac{1}{\sqrt{2n}}$

$$\sigma_{\bar{x}} = \frac{\sigma}{\sqrt{n}} = \frac{\sigma}{\sqrt{2n}} = \frac{\sigma}{\sqrt{2}\cdot\sqrt{n}} = \frac{1}{\sqrt{2}}\cdot\frac{\sigma}{\sqrt{n}}$$

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 ~~ôop~~ $\hat{p}$ 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.

$p = .55 \quad n = 250$

$\mu_{\hat{p}} = p = .55$

$\sigma_{\hat{p}} = \sqrt{\frac{(.55)(.45)}{250}} \approx .03$

✓ Normal
$np \geq 10$ and $n(1-p) \geq 10$

$250(.55) \geq 10$ and $250(.45) \geq 10$

$137.5 \geq 10 \qquad 112.5 \geq 10$

10% $\quad 10(250) \leq$ pop undergrads ✓

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. True

(b) The standard deviation of the sampling distribution is a measure of the variability of the sample mean among repeated samples. True

(c) The sample mean is an unbiased estimator of the true population mean. True

(d) The sampling distribution shows how the sample mean will vary in repeated samples. True

(e) The sampling distribution shows how the sample was distributed around the sample mean. False

The sampling distribution has information about how the sample mean varies from sample to sample Not what any sample itself looks like

T7.9. A machine is designed to fill 16-ounce bottles of shampoo. When the machine is working properly, the mean amount poured into the bottles is 16.05 ounces with a standard deviation of 0.1 ounce. Assume that the machine is working properly. If four bottles are randomly selected each hour and the number of ounces in each bottle is measured, then 95% of the observations should occur in which interval?

(a) 16.05 to 16.15 ounces

(b) −0.30 to +0.30 ounces

(c) 15.95 to 16.15 ounces

(d) 15.90 to 16.20 ounces

(e) None of the above

We are not told the distribution is Normal, and since $n = 4$, the CLT fails

Remember in a Normal Distribution 95% of data falls within 2 standard deviations

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 X. The shape of the sampling distribution of X will be

(a) Normal.

(b) approximately Normal.

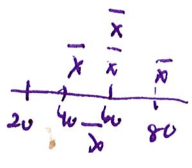
(c) right-skewed.

(d) left-skewed.

(e) symmetric but not Normal.

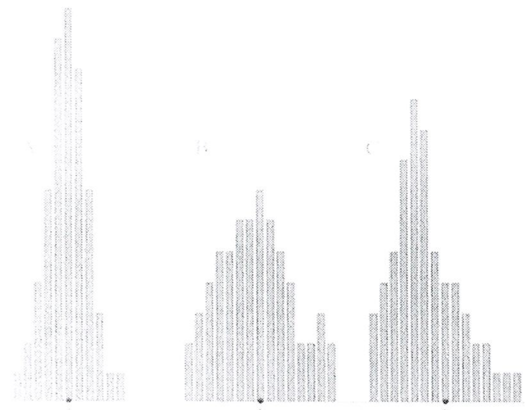
$n = 2 \qquad H = \$80 \qquad T = \$40$

			$\bar{x}$
H \$80	H	\$80	→ \$80
	T	\$40	→ \$60
T \$40	H	\$80	→ \$60
	T	\$40	→ \$40



Since there are only 3 possible outcomes, will not be Normal.

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.

Statistic A provides the best estimate of the parameter. Its center is lined up with the true parameter, making it unbiased, and it has the lowest amount of variability.

T7.12. The amount that households pay service providers for access to the Internet varies quite a bit, but the mean monthly fee is $28 and the standard deviation is $10. The distribution is not Normal: many households pay about $10 for limited dial-up access or about $30 for unlimited dial-up 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 X be the mean amount paid. X = mean amount paid

$\bar{x}$ of 500

a) Explain why you can't determine the probability that the amount a randomly selected household pays for access to the Internet exceeds $29.

We cannot calculate this probability because we do not know the shape of the amount that individual households pay for internet services. (For instance, we know that "many households pay about $10" but we don't know what % of households are in this category.) So we cannot find probability of a randomly selected household. * This question is asking about a selected household, not the sampling distribution. *

b) What are the mean and standard deviation of the sampling distribution of X? SRS of n = 500

$\mu_{\bar{x}} = \mu = \$28.$

✓ 10%

$10(500) \leq$ pop households internet access

$5000 \leq$ pop reasonable to assume

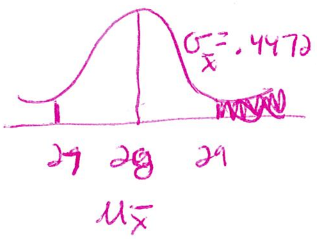
$\sigma_{\bar{x}} = \frac{\$10}{\sqrt{500}} \approx .4472$

c) What is the shape of the sampling distribution of X? Justify your answer.

Since $n=500$ is ≥ 30, by CLT, sampling distribution of X is ~ Normal.

d) Find the probability that the average fee paid by the sample of households exceeds \$29. Show your work. $\sim N(28, .4472)$

$P(\bar{x} > 29)$ = normcdf $(29, \infty, 28, .4472) \approx .0127$



There is about a 1.27% chance of getting a sample of 500 in which the average amount paid for the Internet is more than \$29.

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 households. Be sure to check that you can use the Normal approximation.

Since SRS $n=300$ with $p=.22$

$\mu_{\hat{p}} = p = .22$

✓ 10% condition

$10(300) \leq$ pop children under 6

$3000 \leq$ pop

reasonable to assume

$\sigma_{\hat{p}} = \sqrt{\frac{(.22)(.78)}{300}} \approx .0239$

✓ Normal Conditions

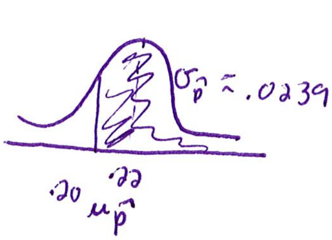
$300(.22) \geq 10$ and $300(.78) \geq 10$

$66 \geq 10$ $\quad 234 \geq 10$

∴ sampling distribution of sample proportion is $\sim N(.22, .0239)$

$P(\hat{p} > .20)$ = normcdf $(.20, \infty, .22, .0239)$

$\approx .7987$



There is about an 80% chance that a sample of 300 children will yield more than 20% who live in households with incomes less than the official poverty level.