AP Statistics **Name:** Key
Chapter 10 Review (Test Only) **Hour:** ______

1. A bank surveyed all of its 60 employees to determine the proportion who participate in volunteer activities. Which of the following statements is true?

(A) The bank should not use the date from this survey because this is an observational study.
(B) The bank can use the result of this survey to prove that working for the bank causes employees to participate in volunteer activities.
(C) The bank did not select a random sample of employees, so the survey will not provide the bank with useful information.
(D) The bank would have to use the survey data to construct a confidence interval in order to estimate the proportion of employees who participate in volunteer activities.
(E) The bank does not need to use an inference procedure to determine the proportion of employees who participate in volunteer activities because the survey was a census of all employees.

2. Based on a survey of a random sample of 900 adults in the United States, a journalist reports that 60 percent of adults in the United States are in favor of increasing the minimum hourly wage. If the reported percent has a margin of error of 2.7 percentage points, which of the following is closest to the level of confidence?

(A) 80.0% (B) 90.0% (C) 95.0% (D) 95.5% (E) 99.0%

$ME = z^* \sqrt{\frac{\hat{p}(1-\hat{p})}{n}}$ $.027 = z \cdot \sqrt{\frac{.60(.40)}{900}}$ $z = 1.65 \Rightarrow 90\%$

3. A large-sample 98 percent confidence interval for the proportion of hotel reservations that are canceled on the intended arrival day is (0.048, 0.112). What is the point estimate for the proportion of hotel reservations that are canceled on the intended arrival day from which this interval was constructed?

(A) 0.032
(B) 0.064
(C) 0.080
(D) 0.160
(E) It cannot be determined from the information given.

$\frac{.048 + .112}{2} = .08$

4. A random sample of 432 voters revealed that 100 are in favor of a certain bond issue. A 95 percent confidence interval for the proportion of the population of voters who are in favor of the bond issue is

(A) $\cancel{100} \pm 1.96\sqrt{\frac{0.5(0.5)}{432}}$

(B) $\cancel{100} \pm \cancel{1.645}\sqrt{\frac{0.5(0.5)}{432}}$

(C) $\cancel{100} \pm 1.96\sqrt{\frac{0.231(0.769)}{432}}$

(D) $0.231 \pm 1.96\sqrt{\frac{0.231(0.769)}{432}}$

(E) $0.231 \pm \cancel{1.645}\sqrt{\frac{0.231(0.769)}{432}}$

$\hat{p} = \frac{100}{432} = .2315$

$95\% \Rightarrow 1.96$

5. In 2009 a survey of Internet usage found that 70 percent of adults age 18 years and older in the United States use the Internet. A broadband company believes that the percent is greater now than it was in 2009 and will conduct a survey. The company plans to construct a 98 percent confidence interval to estimate the current percent and wants the margin of error to be no more than 2.5 percentage points. Assuming that at least 79 percent of adults use the Internet, which of the following should be used to find the sample size (n) needed?

(A) $\cancel{1.96}\sqrt{\frac{0.5}{n}} \leq 0.025$

98% CI ⇒ 2.33

(B) $\cancel{1.96}\sqrt{\frac{(0.5)(0.5)}{n}} \leq 0.025$

(C) $2.33\sqrt{\frac{(0.5)(0.5)}{n}} \leq \cancel{0.5}$

(D) $2.33\sqrt{\frac{(0.79)(0.21)}{n}} \leq 0.025$

(E) $2.33\sqrt{\frac{(0.79)(0.21)}{n}} \leq \cancel{0.5}$

6. Based on a random sample of 50 students, the 90 percent confidence interval for the mean amount of money students spend on lunch at a certain high school is found to be ($3.45, $4.15). Which of the following statements is true?

(A) 90% of the time, the mean amount of money that all students spend on lunch at this high school will be between $3.45 and $4.15.
(B) 90% of all students spend between $3.45 and $4.15 on lunch at this high school.
(C) 90% of all random samples of 50 students obtained at this high school would result in a sample mean amount of money student spend on lunch between $3.45 and $4.15.
(D) 90% of all random samples of 50 students obtained at this high school would result in a 90% confidence interval that contains the true mean amount of money students spend on lunch.
(E) Approximately 45 of the 50 students in the random sample will spend between $3.45 and $4.15 on lunch at this high school.

7. A planning board in Elm County is interested in estimating the proportion of its residents that are in favor of offering incentives to high-tech industries to build plants in that county. A random sample of Elm County residents was selected. All of the residents were asked, "Are you in favor of offering incentives to high-tech industries to build plants in your county?" A 95 percent confidence interval for the proportion of residents in favor of offering incentives was calculated to be 0.54 ± 0.05. Which of the following statements is correct?

(A) At the 95% confidence level, the estimate of 0.54 is within 0.05 of the true proportion of county residents in favor of offering incentives to high-tech industries to build plants in the county.
(B) At the 95% confidence level, the majority of area residents are in favor of offering incentives to high-tech industries to build plants in the county.
(C) In repeated sampling, 95% of sample proportions will fall in the interval (0.49, 0.59).
(D) In repeated sampling, the true proportion of county residents in favor of offering incentives to high-tech industries to build plants in the county will fall in the interval (0.49, 0.59). Does not incl. 95% Conf.
(E) In repeated sampling, 95% of the time the true proportion of county residents in favor of offering incentives to high-tech industries to build plants in the county will be equal ~~to 0.54~~.

8. A researcher has conducted a survey using a simple random sample of 50 registered voters to create a confidence interval to estimate the proportion of registered voters favoring the election of a certain candidate for mayor. Assume that the sample proportion does not change. Which of the following best describes the anticipated effect on the width of the confidence interval if the researcher were to survey a random sample of 200, rather than 50, registered voters?

(A) The width of the new interval would be about one-fourth the width of the original interval.
(B) The width of the new interval would be about one-half the width of the original interval.
(C) The width of the new interval would be about the same width as the original interval.
(D) The width of the new interval would be about twice the width of the original interval.
(E) The width of the new interval would be about four times the width of the original interval.

$\frac{200}{50} = 4 \quad \sqrt{4} = 2 \qquad \sqrt{\frac{p(1-p)}{200}} \qquad \sqrt{\frac{p(1-p)}{50}}$

200 ⇒ smaller width ($\frac{1}{2}$)

9. A marketing company wants to estimate the proportion of consumers in a certain region of the country who would react favorably to a new marketing campaign. Further, the company wants the estimate to have a margin of error of no more than 5 percent with 90 percent confidence. Of the following, which is closest to the minimum number of consumers needed to obtain the estimate with the desired precision?

(A) 136
(B) 271
(C) 385
(D) 542
(E) 769

$\frac{.05}{1.645} \geq \frac{1.645\sqrt{\frac{.5(.5)}{n}}}{1.645}$

$\left(\frac{.05}{1.645}\right)^2 = \sqrt{\frac{.5^2}{n}}^{\,2}$

$n\left(\frac{.05}{1.645}\right)^2 = .25$

$n = \frac{.25}{\left(\frac{.05}{1.645}\right)^2} = 270.6025$

10. A random sample has been taken from a population. A statistician, using this sample, needs to decide whether to construct a 90 percent confidence interval for the population mean or a 95 percent confidence interval for the population mean. How will these intervals differ?

(A) The 90 percent confidence interval will not be as wide as the 95 percent confidence interval.
(B) The 90 percent confidence interval will be wider than the 95 percent confidence interval.
(C) Which interval is wider will depend on how large the sample is.
(D) Which interval is wider will depend on whether the sample is unbiased.
(E) Which interval is wider will depend on whether a z-statistic or a t-statistic is used.

11. *USA Today* reported that speed skater Bonnie Blair had "won the USA's heart," according to a *USA Today*/CNN/Gallop poll conducted on the final Thursday of the 1994 Winter Olympics. When asked who was the hero of the Olympics, 65 percent of the respondents chose Blair, who won five gold medals. The poll of 615 adults, done by telephone, had a margin of error of 4 percent. Which of the following statements best describes what is meant by the 4 percent margin of error?

(A) About 4 percent of adults were expected to change their minds between the time of the poll and its publication in *USA Today*.
(B) About 4 percent of adults did not have telephones.
(C) About 4 percent of the 615 adults polled refused to answer.
(D) Not all of the 615 adults knew anything about the Olympics.
(E) The difference between the sample percentage and the population percentage is likely to be less than 4 percent.

CI: 2nd Vars 4: invT (area, df).

area $\frac{1 \pm c}{2}$

12. A quality control inspector must verify whether a machine that packages snack foods is working correctly. The inspector will randomly select a sample of packages and weigh the amount of snack food in each. Assume that the weights of food in packages filled by this machine have a standard deviation of 0.30 ounce. An estimate of the mean amount of snack food in each package must be reported with 99.6 percent confidence and a margin of error of no more than 0.12 ounce. What would be the minimum sample size for the number of packages the inspector must select?

(A) 8 (B) 15 (C) 25 (D) 52 (E) 60

$.12 = 2.878\left(\frac{.3}{\sqrt{n}}\right)$

99.6CI ⇒

13. You want to estimate the mean number of hours of sleep of students who hold part-time jobs. Assume that the distribution of hours of sleep is roughly normal with $\sigma = 2$ hours. What size SRS would give an interval with a margin of error of 10 minutes with 90% confidence?

120 minutes

$10 = 1.645\left(\frac{120}{\sqrt{n}}\right)$

$\sqrt{n} \cdot \frac{10}{1.645} = \frac{120}{\sqrt{n}} \cdot \sqrt{n}$

$\sqrt{n}\left(\frac{10}{1.645}\right) = 120$

$\sqrt{n}^2 = \left(\frac{120}{\left(\frac{10}{1.645}\right)}\right)^2$

$n = 389.6676$

$\boxed{n = 390 \text{ students}}$

14. Will a CI always capture the true population parameter?

no

15. How do you decrease the width of a confidence interval?

increase sample size or decrease the confidence level

16. Which interval is narrower, a 98% confidence interval or a 92% confidence interval?

92%

17. To cut the margin of error in half, what must you do to the sample size?

$\left(\frac{1}{2}\right)^2 = \frac{1}{4}$

Quadruple it (x4)

(2003B #6)

18. Researchers at a large health maintenance organization (HMO) are planning a study of a certain mild illness. They will select a random sample of patients who are ages 35 to 54 and see if they contract the illness in the next year. The researchers are interested in estimating the proportions of men and of women who are likely to develop the illness in each of 4 age-groups: 35-39, 40-44, 45-49, and 50-54.

The researchers plan to include 2,000 patients in the study. Suppose the researchers draw a random sample from all of the patients at this HMO who are ages 35 to 54 and find the following numbers within each gender and age-group.

	Age-Group			
	35-39	40-44	45-49	50-54
Male	350	230	150	60
Female	445	370	245	150

(a) Suppose that at the end of the study, 10 percent of the females in the 40-44 age group contracted the illness. Calculate a 95 percent confidence interval to estimate the population proportion of females in this age-group that contracted the illness. Interpret this confidence interval in the context of this situation.

① One sample z CI for p, the proportion of all 40-44 yo female HMO patients who contracted the illness

② An SRS of HMO patients is stated, so we can assume the 370 female sample is also SRS. $N \geq 10n = 10(370) = 3700$

It is described as a large HMO, so we can assume there is more than 3700 female patients aged 40-44, so the condition of independence is satisfied.

$n\hat{p} = 370(.10) = 37$ and $n(1-\hat{p}) = 370(.90) = 333$. Since both are ≥ 10 we can assume normality and use normal approximation.

③ $\hat{p} \pm z^* \sqrt{\frac{\hat{p}(1-\hat{p})}{n}} = .10 \pm 1.96\sqrt{\frac{.10(.90)}{370}} = (.06943, .13057)$

④ Based on this sample, I am 95% confident that the true proportion of all 40-44 yo female patients at this HMO who contracted the illness is between .069 and .131.

(b) Interpret the confidence level of 95 percent.

95% of the time the true proportion will be within the interval found using the one sample z CI.

(c) Suppose that at the end of the study, 10 percent of the males in the 40-44 age group contracted the illness. The corresponding 95 percent confidence interval to estimate the population proportion of males in this age-group that contracted the illness is (0.061, 0.139).

Note that this interval and the interval in part (a) are of different lengths even though the two sample proportions were identical. What would be an alternative way to allocate a sample of 2,000 subjects so that the 95 percent confidence interval widths for all male age-groups and for all female age-groups (i.e., for all 8 groups) would be the same when the sample proportions are the same? Justify your answer.

The difference is the margin of error $z^* \sqrt{\frac{\hat{p}(1-\hat{p})}{n}}$.

z is 1.96 at the 95% confidence level and $\hat{p}$ is the same for males and females at .10. The only differing variable is n. The same sample size would be needed for each study. $\frac{2000}{8} = 250$.

(2008B #3)

19. A car manufacturer is interested in conducting a study to estimate the mean stopping distance for a new type of brakes when used in a car that is traveling at 60 miles per hour. These new brakes will be installed on cars of the same model and the stopping distance will be observed. The cost of each observation is $100. A budget of $12,000 is available to conduct the study and the goal is to carry it out in the most economical way possible. Preliminary studies indicate that $\sigma = 12$ feet for stopping distances.

(a) Are sufficient funds available to estimate the mean stopping distance to within 2 feet of the true mean stopping distance with 95% confidence? Explain your answer.

$$ME = z^* \frac{\sigma}{\sqrt{n}}$$

$$\frac{2}{1.96} = \frac{1.96\left(\frac{12}{\sqrt{n}}\right)}{1.96}$$

$$\sqrt{n}\,\frac{2}{1.96} = \frac{12}{\sqrt{n}}\,\sqrt{n}$$

$$\sqrt{n}^2 = \left(\frac{12}{\left(\frac{2}{1.96}\right)}\right)^2$$

$$n = 138.2976$$

The manufacturer needs to test 139 cars with the new brakes. Each observation costs $100 bringing the study to 139(100) = $13900. A $12,000 budget will not be enough.

(b) A regulatory agency requires a 95% level of confidence for an estimate of mean stopping distance that is within 2 feet of the true mean stopping distance. The car manufacturer cannot exceed the budget of $12,000 for the study. Discuss the consequences of these constraints.

Because the regulatory agency requires a 95% confidence level, the manufacturer must test a minimum of 139 cars. The manufacturer must come up with an additional $1900 to complete the study.

(~1997 #5)

20. A company bakes computer chips in two ovens, oven A and oven B. The chips are randomly assigned to an oven and hundreds of chips are baked each hour. The percentage of defective chips coming from these ovens for each hour of production throughout a day is shown below.

Percentage of Defective Chips

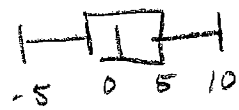
Hour	Oven A	Oven B
1	45	36
2	32	37
3	34	33
4	31	34
5	35	33
6	37	32
7	31	33
8	30	30
9	27	24

The hourly differences in percentages for oven A minus oven B have a mean of 1.11 and a standard deviation of 4.28.

(a) Construct a confidence interval for the mean difference in percentage of defective chips produced each hour from oven A and oven B. Be sure to interpret this interval.

① One sample t CI for the mean difference in percentage of defective chips produced each hour (A-B)

② The chips are randomly assigned to ovens so it is safe to assume it is a random sample of defective chips.
$N \geq 10(9) = 90$. Since there are more than 90 differences, independence is assumed. A graph of the differences is $\sim N$ so the population will be $\sim N$.



③ $\bar{X}_D \pm t^* \frac{S_D}{\sqrt{n}} = 1.11 \pm 2.306\left(\frac{4.28}{\sqrt{9}}\right) = (-2.183, 4.405)$ $df = 9-1 = 8$

④ Based on this sample, I am 95% confident that the mean difference in the percentage of defective chips produced each hour is between -2.183 and 4.405.

(b) Based only on the confidence interval in part (a), is there sufficient evidence to conclude that there is a significant mean difference in percentage of defective chips produced each hour by oven A and oven B? Justify your conclusion.

There is not sufficient evidence since zero is in the 95% CI and a mean difference of zero would signify no difference.

(2010B #4)

21. A husband and wife, Mike and Lori, share a digital music player that has a feature that randomly selects which song to play. A total of 2,384 songs were loaded onto the player, some by Mike and the rest by Lori. Suppose that when the player was in the random-selection mode, 13 of the first 50 songs selected were songs loaded by Lori.

(a) Construct and interpret a 90 percent confidence interval for the proportion of songs on the player that were loaded by Lori.

① One sample z CI for p, proportion of songs on the player loaded by Lori.

② SRS is stated. $n \geq 10(50) = 500$. Since there are more than 500 songs loaded on the player, independence is satisfied. $n\hat{p} = 13$ $n(1-\hat{p}) = 37$. Since both are ≥ 10, normality is satisfied and I can use normal approximation

③ $\hat{p} \pm z^* \sqrt{\frac{\hat{p}(1-\hat{p})}{n}} = .26 \pm 1.64 \sqrt{\frac{.26(.74)}{50}} = (.158, .362)$

④ Based on this sample I am 90% confident the true proportion of songs loaded on the player by Lori is between

(b) Mike and Lori are unsure about whether the player samples the songs with replacement or without replacement when the player is in random-selection mode. Explain why this distinction is not important for the construction of the interval in part (a).

There are 2,384 songs on the player which is much greater than $10(50) = 500$ songs. Therefore replacement is not important.

(2011 #6)

22. Every year, each student in a nationally representative sample is given tests in various subjects. Recently, a random sample of 9,600 twelfth-grade students from the United States were administered a multiple-choice United States history exam. One of the multiple-choice questions is below. (The correct answer is C.)

> In 1935 and 1936 the Supreme Court declared that important parts of the New Deal were unconstitutional. President Roosevelt responded by threatening to
>
> (A) impeach several Supreme Court justices.
> (B) eliminate the Supreme Court.
> (C) appoint additional Supreme Court justices who shared his views.
> (D) override the Supreme Court's decisions by gaining three-fourths majorities in both houses of Congress.

Of the 9,600 students, 28 percent answered the multiple-choice question correctly.

(a) Let p be the proportion of all United States twelfth-grade students who would answer the question correctly. Construct and interpret a 99 percent confidence interval for p.

1. One sample z CI for the proportion of all US 12th graders who would answer the question correctly.

2. SRS is stated. $N \geq 10(9600) = 96000$. Since there are more than 96000 seniors in the U.S. independence is satisfied.

   $n\hat{p} = 9600(.28) = 2688$ $\quad n(1-\hat{p}) = 9600(.72) = 6912$

   Since both are ≥ 10, we may use normal approximation.

3. $\hat{p} \pm z^* \sqrt{\frac{\hat{p}(1-\hat{p})}{n}} = .28 \pm 2.58 \sqrt{\frac{.28(.72)}{9600}} = (.2682, .2918)$

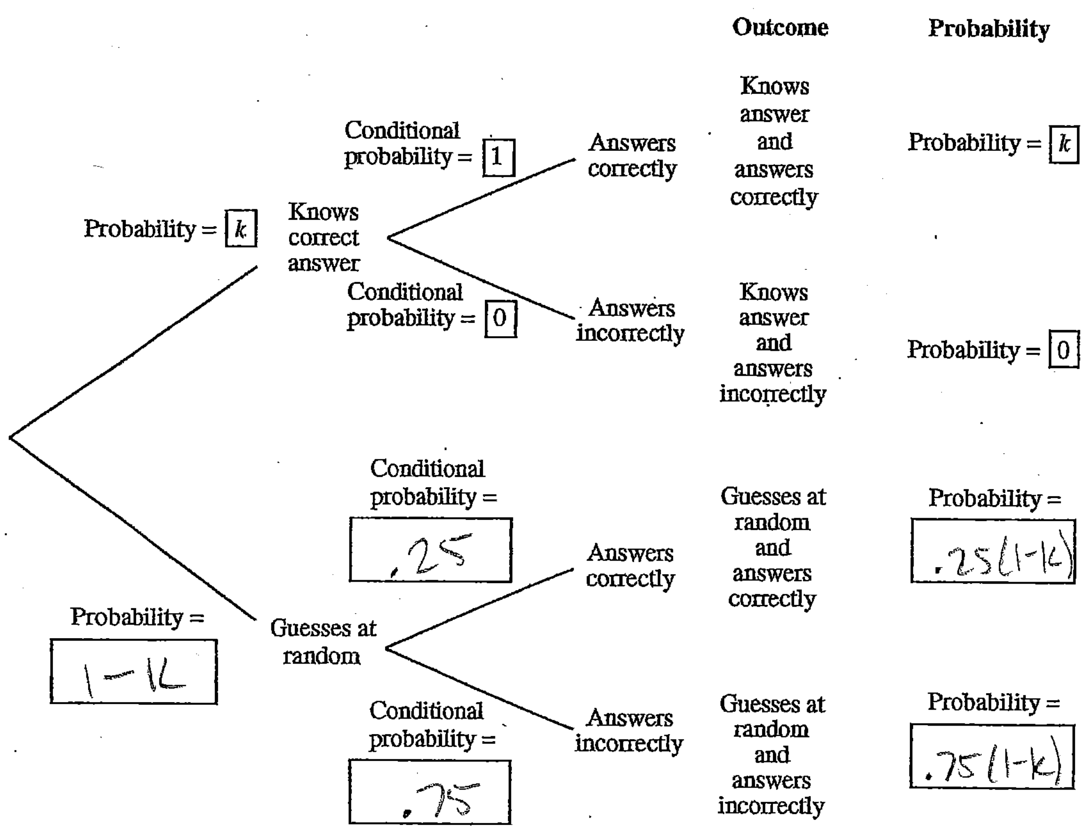
4. Based on this sample I am 99% confident the true proportion of all US 12th grade students who would answer the question correctly is between .2682 and .2918.

Assume that students who actually know the correct answer have a 100 percent chance of answering the question correctly, and students who do not know the correct answer to the question guess completely at random from among the four options.

Let k represent the proportion of all United States twelfth-grade students who actually know the correct answer to the question.

(b) A tree diagram of the possible outcomes for a randomly selected twelfth-grade student is provided on the following page. Write the correct probability in each of the five empty boxes. Some of the probabilities may be expressions in terms of k.

TREE DIAGRAM OF OUTCOMES FOR A RANDOMLY SELECTED TWELFTH-GRADE STUDENT



(c) Based on the completed tree diagram, express the probability, in terms of k, that a randomly selected twelfth-grade student would correctly answer the history question.

$$k + .25(1-k) = k + .25 - .25k = .75k + .25$$

(d) Using your interval from part (a) and your answer to part (c), calculate and interpret a 99 percent confidence interval for k, the proportion of all United States twelfth-grade students who actually know the answer to the history question. You may assume that the conditions for inference for the confidence interval have been checked and verified.

$.268 = .75k + .25$ $\quad\quad$ $.292 = .75k + .25$

$k = .024$ $\quad\quad$ $k = .056$ $\quad\quad$ $(.024, .056)$

Based on this sample, I am 99% confident that the true proportion of all US 12th graders who actually know the answer to the question is between .024 and .056.