Homework Key

SECTION 8.1

In Exercises 1 to 4, determine the point estimator you would use and calculate the value of the point estimate.

1. Got shoes? How many pairs of shoes, on average, do female teens have? To find out, an AP Statistics class conducted a survey. They selected an SRS of 20 female students from their school. Then they recorded the number of pairs of shoes that each student reported having. Here are the data:

50	26	26	31	57	19	24	22	23	38
13	50	13	34	23	30	49	13	15	51

TODO
① create L1
② Stat-1VAR
$\bar{x} = 30.35$ $s_x = 13.88$

2. Got shoes? The class in Exercise 1 wants to estimate the variability in the number of pairs of shoes that female students have by estimating the population variance σ^2.

1 POINT ESTIMATOR: Mean number of shoes $\bar{x} = 30.35$
POINT ESTIMATE

2 $s^2 = (13.88)^2 = 202.77$
Sample variance of the pairs of shoes

15. Shoes The AP Statistics class in Exercise 1 also asked an SRS of 20 boys at their school how many shoes they have. A 95% confidence interval for the difference in the population means (girls – boys) is 10.9 to 26.5. Interpret the confidence interval and the confidence level.

15 CI: We are 95% confident that the interval from 10.9 to 26.5 captures the true difference in the ~~average~~ average # of pairs of shoes owned by girls and boys (girls − boys).

CL: IF THIS SAMPLING METHOD WERE EMPLOYED MANY TIMES, APPROX. 95% of the resulting confidence interval would capture the true difference avg pairs shoes between boys & g

Multiple choice: Select the best answer for Exercises 21 to 24.

21. A researcher plans to use a random sample of n = 500 families to estimate the mean monthly family income for a large population. A 99% confidence interval based on the sample would be ______ than a 90% confidence interval.

(a) narrower and would involve a larger risk of being incorrect   99% wider

(b) wider and would involve a smaller risk of being incorrect   90%

(c) narrower and would involve a smaller risk of being incorrect

(d) wider and would involve a larger risk of being incorrect

(e) wider, but it cannot be determined whether the risk of being incorrect would be larger or smaller

22. In a poll, * * IMPORTANT POINT ON ME *

I. Some people refused to answer questions.

II. People without telephones could not be in the sample.

III. Some people never answered the phone in several calls.

Which of these sources is included in the ±2% margin of error announced for the poll?

(a) I only (b) II only (c) III only (d) I, II, and III (e) None of these

ME accounts variability due to random selection/assignment
ME does NOT compensate for any bias in the data collection process

23. You have measured the systolic blood pressure of an SRS of 25 company employees. A 95% confidence interval for the mean systolic blood pressure for the employees of this company is (122, 138). Which of the following statements gives a valid interpretation of this interval?

(a) 95% of the sample of employees have a systolic blood pressure between 122 and 138.

(b) 95% of the population of employees have a systolic blood pressure between 122 and 138.

(c) If the procedure were repeated many times, 95% of the resulting confidence intervals would contain the population mean systolic blood pressure.

(d) The ~~probability~~ that the population mean blood pressure is between 122 and 138 is 0.95.

(e) If the procedure were repeated many times, 95% of the ~~sample means~~ would be between 122 and 138.

24. A polling organization announces that the proportion of American voters who favor congressional term limits is 64%, with a 95% confidence margin of error of 3%. If the opinion poll had announced the margin of error for 80% confidence rather than 95% confidence, this margin of error would be

ME = critical value * SD (statistic)

(a) 3%, because the same sample is used.

(b) less than 3%, because we require less confidence.

(c) less than 3%, because the sample size is smaller.

(d) greater than 3%, because we require less confidence.

(e) greater than 3%, because the sample size is smaller.

SECTION 8.2 Exercises

For Exercises 27 to 30, check whether each of the conditions is met for calculating a confidence interval for the population proportion p.

27. **Rating dorm food** Latoya wants to estimate what proportion of the seniors at her high school like the cafeteria food. She interviews an SRS of 50 of the 175 seniors living in the dormitory. She finds that 14 think the cafeteria food is good.

INDEPENDENT: NOT MET BECAUSE sample does not meet the 10% condition.

$\boxed{N=175}$ $n=50*10=500$ minimum population (N)

28. **High tuition costs** Glenn wonders what proportion of the students at his school think that tuition is too high. He interviews an SRS of 50 of the 2400 students at his college. Thirty-eight of those interviewed think tuition is too high. phat=38/50=.76

Conditions met Random: SRS

Independent: $50*10=500<2400$ 10% condition ✓

Normal: np=50*.76=38>10 and nq=50*.24=12>10

29. **AIDS and risk factors** In the National AIDS Behavioral Surveys sample of 2673 adult heterosexuals, 0.2% had both received a blood transfusion and had a sexual partner from a group at high risk of AIDS. We want to estimate the proportion p in the population who share these two risk factors.

2 Conditions Not met:

1. Random - may not be met since not told how sample was gotten
2. Normal: $np=.002*2673=5.3<10$ ✗ $nq=2668$ ✓

30. **Whelks and mussels** The small round holes you often see in sea shells were drilled by other sea creatures, who ate the former dwellers of the shells. Whelks often drill into mussels, but this behavior appears to be more or less common in different locations. Researchers collected whelk eggs from the coast of Oregon, raised the whelks in the laboratory, then put each whelk in a container with some delicious mussels. Only 9 of 98 whelks drilled into a mussel.[11] The researchers want to estimate the proportion p of Oregon whelks that will spontaneously drill into mussels.

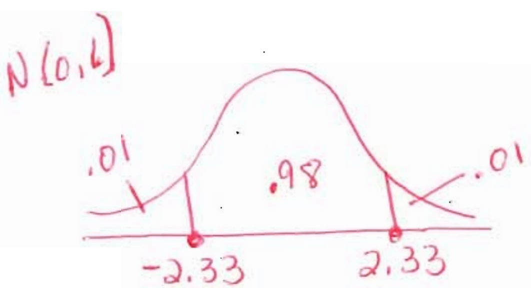
Normal not met:

$\hat{p}=9/98=.092$

$n=98$

$n\hat{p}=.092(98)=9$ is NOT at least 10

31. **98% confidence** Find z^* for a 98% confidence interval using Table A or your calculator. Show your method.



invNorm(.01, 0, 1)

↓

$\boxed{z^*=2.33}$

33 Going to the prom Tonya wants to estimate what proportion of her school's seniors plan to attend the prom. She interviews an SRS of 50 of the 750 seniors in her school and finds that 36 plan to go to the prom.

(a) Identify the population and parameter of interest.

(b) Check conditions for constructing a confidence interval for the parameter.

$\hat{p} = 36/50 = .72$

A Population: the seniors at Tonya's HS
Parameter: true proportion (p) who attend prom.

B Random: an SRS ✓
Independent: $n = 50 \times 10 = 500 < 750$ seniors ✓
Normal: $n\hat{p} = 50(.72) = 36 \geq 10$ ✓
$n\hat{q} = 50(.28) = 14 \geq 10$ ✓

36 Teens' online profiles Over half of all American teens (ages 12 to 17 years) have an online profile, mainly on Facebook. A random sample of 487 teens with profiles found that 385 included photos of themselves.[13] $\hat{p} = 385/487 = .791$ $n = 487$

(a) Construct and interpret a 95% confidence interval for p. Follow the four-step process.

(b) Is it plausible that the true proportion of American teens with profiles who have posted photos of themselves is 0.75? Use your result from part (a) to support your answer.

Conditions
Random – random sample
Independent – sample less than 10% of all teens
Normal – $n\hat{p} = 487(.791) = 385 \geq 10$ ✓
$n\hat{q} = 487(.209) = 102 \geq 10$ ✓

A $z^* = 1.96$ $\hat{p} \pm z^* \sqrt{\frac{\hat{p}\hat{q}}{n}}$

$.791 \pm 1.96 \sqrt{\frac{(.791)(.209)}{487}}$

$.791 \pm$ .036 (.755, .827)

WE ARE 95% CONFIDENT THAT THE INTERVAL .755 TO .827 OF TEENS CONTAINS THE TRUE PROPORTION OF TEENS WHO HAVE ONLINE PROFILES THAT CONTAIN PHOTO'S

B THE VALUE .75 DOES NOT APPEAR IN OUR 95% CONFIDENCE INTERVAL, SO IT WOULD BE SURPRISING IF THE TRUE PROPORTION WAS .75

38 Teens' online profiles Describe a possible source of error that is not included in the margin of error for the 95% confidence interval in Exercise 36.

The margin of error was .018 does not include bias that occurs from any bias in the data collection process. In this example, sources of bias could result from under coverage and non response.

41. A TV poll A television news program conducts a call-in poll about a proposed city ban on handgun ownership. Of the 2372 calls, 1921 oppose the ban. The station, following recommended practice, makes a confidence statement: "81% of the Channel 13 Pulse Poll sample opposed the ban. We can be 95% confident that the true proportion of citizens opposing a handgun ban is within 1.6% of the sample result."

(a) Is the station's quoted 1.6% margin of error correct? Explain.

(b) Is the station's conclusion justified? Explain.

$\hat{p} = \frac{1921}{2372} = .8099$ $n = 2,372$

FIRST—CHECK CONDITIONS:

X ① RANDOM – NO – NOT A random sample a voluntary sample

✓ ② NORMAL – $n\hat{p} = 1,921$ $n\hat{q} = 451$ ✓

X ③ INDEPENDENT – 10% condition not met

ⓐ NO, THEY SHOULD NOT CALCULATE A MARGIN OF ERROR, BECAUSE THE RANDOM Condition was NOT met

ⓑ No. This was a voluntary response sample, so no inference should be made about the population.

---

44. School vouchers A national opinion poll found that 44% of all American adults agree that parents should be given vouchers that are good for education at any public or private school of their choice. The result was based on a small sample.

(a) How large an SRS is required to obtain a margin of error of 0.03 (that is, ±3%) in a 99% confidence interval? Answer this question using the previous poll's result as the guessed value for $\hat{p}$.

(b) Answer the question in part (a) again, but this time use the conservative guess $\hat{p} = 0.5$. By how much do the two sample sizes differ?

(A) $\hat{p} = .44$

$ME = .03$

99% CI $z^* = 2.58$

invNorm(.005, 0, 1)

To Calculate sample size

$z^*\sqrt{\frac{\hat{p}\hat{q}}{n}} \le ME$

$2.58\sqrt{\frac{(.44)(.56)}{n}} \le .03$

$\left(\frac{2.58\sqrt{(.44)(.56)}}{.03}\right)^2 \le (\sqrt{n})^2$

$\boxed{n \ge 1,823}$

(B) $\hat{p} = .5$

$\left(\frac{2.58\sqrt{(.5)(.5)}}{.03}\right) \le (\sqrt{n})^2$

$\boxed{n \ge 1,849}$

The difference in sample size is 26.

Multiple choice: Select the best answer for Exercises 49 to 52.

49. A Gallup Poll found that only 28% of American adults expect to inherit money or valuable possessions from a relative. The poll's margin of error was ±3 percentage points at a 95% confidence level. This means that

(a) the poll used a method that gets an answer within 3% of the truth about the population 95% of the time.

(b) the percent of all adults who expect an inheritance is between 25% and 31%.

(c) if Gallup takes another poll on this issue, the results of the second poll will lie between 25% and 31%.

(d) there's a 95% chance that the percent of all adults who expect an inheritance is between 25% and 31%.

(e) Gallup can be 95% confident that between 25% and 31% of the sample expect an inheritance.

50. Most people can roll their tongues, but many can't. The ability to roll the tongue is genetically determined. Suppose we are interested in determining what proportion of students can roll their tongues. We test a simple random sample of 400 students and find that 317 can roll their tongues. The margin of error for a 95% confidence interval for the true proportion of tongue rollers among students is closest to

(a) 0.008. (b) 0.02. (c) 0.03. (d) 0.04. (e) 0.208.

$\hat{p} = \frac{317}{400}$

$\hat{p} = .79$

$z^* = 1.96$

$ME = 1.96 * \sqrt{\frac{.79(.21)}{400}}$

$ME = .0399$

51. You want to design a study to estimate the proportion of students at your school who agree with the statement, "The student government is an effective organization for expressing the needs of students to the administration." You will use a 95% confidence interval, and you would like the margin of error to be 0.05 or less. The minimum sample size required is

(a) 22. (b) 271. (c) 385. (d) 769. (e) 1795.

$1.96\sqrt{\frac{.5(.5)}{n}} \le .05$

remember to round up

52. I collect an SRS of size n from a population and compute a 95% confidence interval for the population proportion. Which of the following would produce a new confidence interval with larger width (larger margin of error) based on these same data?

(a) Use a larger confidence level.

(b) Use a smaller confidence level.

(c) Increase the sample size.

(d) Use the same confidence level, but compute the interval n times. Approximately 5% of these intervals will be larger.

(e) Nothing can guarantee absolutely that you will get a larger interval. One can only say that the chance of obtaining a larger interval is 0.05.

$ME = z^*\sqrt{\frac{\hat{p}\hat{q}}{n}}$

This doesn't change

SECTION 8.3 Exercises

Find sample size when population SD known

$\sigma = 50$

$ME = .02$

$z^* = 1.96$

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

$1.96 \cdot \frac{50}{\sqrt{n}} \le 2$ ← Note 2 pts NOT 2%

$\left(\frac{1.96(50)}{2}\right)^2 \le (\sqrt{n})^2$

$n \ge 2{,}401$

56. **The SAT again** High school students who take the SAT Math exam a second time generally score higher than on their first try. Past data suggest that the score increase has a standard deviation of about 50 points. How large a sample of high school students would be needed to estimate the mean change in SAT score to within 2 points ~~with 95% confidence? Show your~~ work. → Minimum sample size of 2,401.

57. **Critical values** What critical value t^* from Table B (or calc) would you use for a confidence interval for the population mean in each of the following situations?

(a) A 95% confidence interval based on $n = 10$ observations.

(b) A 99% confidence interval from an SRS of 20 observations.

(a) 95% $n = 10$ $df = 10 - 1 = 9$

$T^* = \text{invT}(.025, 9) \rightarrow 2.26 = t^*$

(b) 99% $n = 20$ $df = 20 - 1 = 19$

$t^* = \text{invT}(.005, 19)$

$t^* = 2.861$

60. **Travel time to work** A study of commuting times reports the travel times to work of a random sample of 20 employed adults in New York State. The mean is $\bar{x} = 31.25$ minutes, and the standard deviation is $s_x = 21.88$ minutes. What is the standard error of the mean? Interpret this value in context.

$\bar{x} = 31.25$

$s_x = 21.88$

$n = 20$

$SE(\bar{x}) = \frac{s_x}{\sqrt{n}} = \frac{21.88}{\sqrt{20}} = 4.8925$

Conditions Checked

Random: SRS ✓

Independent: 20 * 10 = 200 NY employees ✓

Normal – must assume

IN REPEATED SAMPLING, THE AVERAGE DISTANCE BETWEEN THE SAMPLE MEANS AND THE POPULATION MEAN WILL BE ABOUT 4.89 minutes.

63. **Give it some gas!** Computers in some vehicles calculate various quantities related to performance. One of these is fuel efficiency, or gas mileage, usually expressed as miles per gallon (mpg). For one vehicle equipped in this way, the miles per gallon were recorded each time the gas tank was filled and the computer was then reset.[24] Here are the mpg values for a random sample of 20 of these records:

Create L1 / 1VAR Stat

$\bar{x} = 18.48$

$s_x = 3.116$

$n = 20$

$df = 19$

$t^* = 2.093$

15.8	13.6	15.6	19.1	22.4	15.6	22.5	17.2	19.4	22.6
19.4	18.0	14.6	18.7	21.0	14.8	22.6	21.5	14.3	20.9

Construct and interpret a 95% confidence interval for the mean fuel efficiency μ for this vehicle.

$\bar{x} \pm t^* \frac{s_x}{\sqrt{n}} \rightarrow 18.48 \pm 2.093 \cdot \frac{(3.116)}{\sqrt{20}}$

18.48 ± 1.46 (17.02, 19.94)

WE ARE 95% CONFIDENT THAT THE INTERVAL FROM 17.02 TO 19.94 CAPTURES THE TRUE MEAN MILES PER GALLON FOR THIS TYPE OF CAR.

CONDITIONS CHECKED

✓ (1) Random SRS of 20

✓ (2) Independent Reasonable that there are more than 200 Cars records

✓ (3) Normal

stem	leaf
13	6
14	6 8 3
15	8 6 6
16	
17	2
18	0 7
19	1 4 4
20	9
21	0 5
22	4 5 6 6

USE CALC TO GRAPH ** A HISTOGRAM OR A STEM LEAF IS AN EASY WAY TO CONFIRM THERE ARE NO OUTLIERS OR ANY STRONG SKEWNESS.

Condition have been met to do a t statistic confidence interval.

Tip** WHEN USE TI84, OVERLAY HISTOGRAM (TO SEE SHAPE) AND BOX PLOT (Check SKEWNESS AND OUTLIERS).

65. Critical value What critical value t^* from Table B would you use for a 99% confidence interval for the population mean based on an SRS of size 58? If possible, use technology to find a more accurate value of t^*. What advantage does the more accurate df provide?

USING TABLE B $t^* = 2.678$

INV T (.005, 57) = $t^* = 2.665$

NOT IMPORTANT TO BE ABLE TO DO BOTH, BUT CALC WILL HAVE A SLIGHTLY SHORTER INTERVAL WITH THE SAME LEVEL OF CONFIDENCE.

67. Bone loss by nursing mothers Breast-feeding mothers secrete calcium into their milk. Some of the calcium may come from their bones, so mothers may lose bone mineral. Researchers measured the percent change in bone mineral content (BMC) of the spines of 47 randomly selected mothers during three months of breast-feeding.[26] The mean change in BMC was −3.587% and the standard deviation was 2.506%.

(a) Construct and interpret a 99% confidence interval to estimate the mean percent change in BMC in the population.

(b) Based on your interval from (a), do these data give good evidence that on the average nursing mothers lose bone mineral? Explain.

CHECK CONDITIONS

* we do not know the population mean or standard deviation so we check conditions for "t interval of μ"

1. Random – Random Sample
2. Independent – We have data from less than 10% of nursing mothers
3. Normal – $n \geq 30$

CONDITIONS MET

(A) $\bar{x} = -3.587\%$ $SD(\bar{x}) = 2.506\%$ $n = 47$ $df = 46$ 99% CI → $t^* = 2.69$ invT(.005, 46)

$\bar{x} \pm t^* \frac{s_x}{\sqrt{n}} \rightarrow -3.587 \pm 2.69 \left(\frac{2.506}{\sqrt{47}}\right)$

Table B $t^* = 2.704$

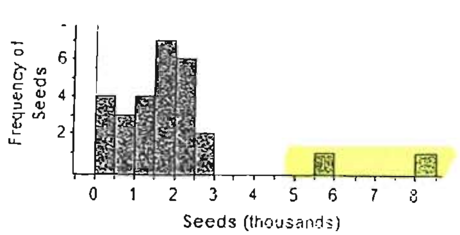
$-3.587 \pm .983$

$(-4.57, -2.604)$

WE ARE 99% Confident that the interval from −4.57 to −2.60 captures the true mean percent change in BMC.

(B) YES. THE INTERVAL INCLUDES ONLY NEGATIVE NUMBERS, WHICH REPRESENTS BONE MINERAL LOSS, SO WE ARE QUITE CONFIDENT THAT NURSING MOTHERS LOSE BONE MINERAL

72. Weeds among the corn Velvetleaf is a particularly annoying weed in cornfields. It produces lots of seeds, and the seeds wait in the soil for years until conditions are right for sprouting. How many seeds do velvetleaf plants produce? The Fathom histogram below shows the counts from 28 plants that came up in a cornfield when no herbicide was used.[30] Explain why it would not be wise to use a one-sample t interval to estimate the mean number of seeds μ produced by velvetleaf plants.



THE SAMPLE SIZE IS SMALL ($n = 28$) AND THERE ARE SEVERAL OUTLIERS SO IT WOULD NOT BE APPROBRIATE TO USE A ONE-SAMPLE t interval to estimate a confidence interval.

73 Should we use t? In each of the following situations, discuss whether it would be appropriate to construct a one-sample t interval to estimate the population mean.

(a) We collect data from a random sample of adult residents in a state. Our goal is to estimate the overall percent of adults in the state who are college graduates.
^p

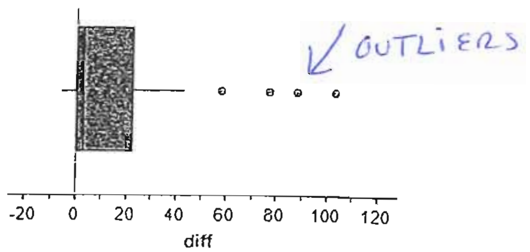
A. NO. THE GOAL IS TO ESTIMATE POPULATION PROPORTION AND NOT THE POPULATION MEAN

(b) The coach of a college men's basketball team records the resting heart rates of the 15 team members. We use these data to construct a confidence interval for the mean resting heart rate of all male students at this college.

B. No. The sample was not an SRS of ALL males at the college. It only included members of the team.

(c) Do teens text more than they call? To find out, an AP Statistics class at a large high school collected data on the number of text messages and calls sent or received by each of 25 randomly selected students. The Fathom boxplot below displays the difference (texts − calls) for each student.

C. NO. IT IS A SMALL SAMPLE ($n = 25$) AND THE GRAPH SHOWS SEVERAL OUTLIERS.



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

Multiple choice: Select the best answer for Exercises 75 to 78.

75 One reason for using a t distribution instead of the standard Normal curve to find critical values when calculating a level C confidence interval for a population mean is that

(a) z can be used only for large samples.

(b) z requires that you know the population standard deviation σ.

(c) z requires that you can regard your data as an SRS from the population.

(d) the standard Normal table doesn't include confidence levels at the bottom.

(e) a z critical value will lead to a wider interval than a t critical value.

Small sample and

76 You have an SRS of 23 observations from a Normally distributed population. What critical value would you use to obtain a 98% confidence interval for the mean μ of the population if σ is unknown?

(a) 2.508 (c) 2.326 (e) 2.177
(b) 2.500 (d) 2.183

$t^* = invT(.01, 22)$

77 A quality control inspector will measure the salt content (in milligrams) in a random sample of bags of potato chips from an hour of production. Which of the following would result in the smallest margin of error in estimating the mean salt content μ?

Lower CL will have smaller critical values, t

(a) 90% confidence; $n = 25$ $t^* = 1.71$

(b) 90% confidence; $n = 50$ $z^* = 1.65$

(c) 95% confidence; $n = 25$ $z^* = 2.06$

(d) 95% confidence; $n = 50$ $z^* = 1.96$

(e) $n = 100$ at any confidence level

For same CL t^* will always be larger than z^*

78 Scientists collect data on the blood cholesterol levels (milligrams per deciliter of blood) of a random ~~sample of 24 laboratory rats.~~ A 95% confidence interval for the mean blood cholesterol level μ is 80.2 to 89.8. Which of the following would cause the most worry about the validity of this interval?

Small sample so use t-statistic

(a) There is a clear outlier in the data. cannot use t^*

(b) A stemplot of the data shows a mild right-skew. OK

(c) You do not know the population standard deviation σ. OK w/ t^*

(d) The population distribution is not exactly Normal. OK

(e) None of these would be a problem because the t procedures are robust.