8.2 Estimating a Population Proportion using Confidence Intervals

Confidence Intervals: A Four-Step Process

State: What *parameter* do you want to estimate, and at what *confidence level*?

Plan: Identify the appropriate inference *method*. Check *conditions*.

Do: If the conditions are met, perform *calculations*.

Conclude: *Interpret* your interval in the context of the problem.

AP EXAM TIP: If a free-response question asks you to construct and interpret a confidence interval, you are expected to do the entire four step process. That includes clearly defining the parameter and checking conditions.

Jan 4-11:31 AM

Conditions for constructing a confidence interval about a proportion

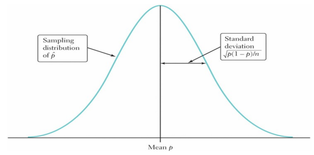
1. ***Random:** The data come from a well-designed random sample or randomized experiment.

2. ∘ **10%:** When sampling without replacement, check that $n \le 0.10N$.

3. ***Large Counts:** Both $n\hat{p}$ and $n(1-\hat{p})$ are at least 10.

4. **statistic ± (critical value) · (standard deviation of statistic)**

When the conditions are met, the sampling distribution of $\hat{p}$ will be approximately Normal with mean $\mu_{\hat{p}} = p$ and standard deviation $\sigma_{\hat{p}} = \sqrt{\frac{p(1-p)}{n}}$.



Jan 4-11:32 AM

Example 1: Mr. Vignolini's class took an SRS of beads from the container and got 107 red beads and 144 white beads.

a) Calculate and interpret a 90% confidence interval for getting a red bead.

Check that the conditions for constructing a confidence interval for p are met.

1. ***Random:** The class took an SRS of 251 beads from the container.

2. *10%: Because the class sampled without replacement, they need to check that there are at least 10(251) = 2510 beads in the population. Mr. Vignolini reveals that there are 3000 beads in the container.

3. **Large Counts:** To use a Normal approximation for the sampling distribution of $\hat{p}$, we need both np and $n(1-p)$ to be at least 10. Because we don't know p, we check

$$n\hat{p} = 251\left(\frac{107}{251}\right) = 107 \quad \text{and} \quad n(1-\hat{p}) = 251\left(1-\frac{107}{251}\right) = 144$$

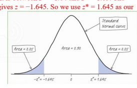
4. statistic ± (critical value) • (standard deviation of statistic)

$$= \hat{p} \pm z^* \sqrt{\frac{\hat{p}(1-\hat{p})}{n}}$$

The sample statistic $\hat{p} = \frac{107}{251} = 0.426$.

Now let's find the critical value. From Table A, we look for the point with area 0.05 to its left. As the excerpt from Table A shows, this point is between $z = -1.64$ and $z = -1.65$. The calculator's invNorm(area: 0.05, μ: 0, σ: 1) gives $z = -1.645$. So we use $z^* = 1.645$ as our critical value.

z	.03	.04	.05
−1.7	.0418	.0409	.0401
−1.6	.0516	.0505	.0495
−1.5	.0630	.0618	.0606



The resulting 90% confidence interval is

$$= 0.426 \pm 1.645\sqrt{\frac{(0.426)(1-0.426)}{251}}$$
$$= 0.426 \pm 0.051$$
$$= (0.375, 0.477)$$

We are 90% confident that the interval from 0.375 to 0.477 captures the true proportion of red beads in Mr. Vignolini's container.

Jan 4-11:35 AM

b) Mr. Vignolini claims that exactly half of the beads in the container are red. Use your result from part (a) to comment on this claim.

$$(0.375, 0.477)$$

The confidence interval in part (a) gives a set of plausible values for the population proportion of red beads. Because 0.5 is not contained in the interval, it is not a plausible value for p. We have good reason to doubt Mr. Vignolini's claim.

Jan 4-12:01 PM

Example 2: The Gallup Youth Survey asked a random sample of 439 U.S. teens aged 13 to 17 whether they thought young people should wait to have sex until marriage. Of the sample, 246 said "Yes." Construct and interpret a 95% confidence interval for the proportion of all teens who would say "Yes" if asked this question.

1. **Random:** Gallup surveyed a random sample of U.S. teens.

2. **10%:** Because Gallup is sampling without replacement, we need to check the 10% condition: there are at least 10(439) = 4390 U.S. teens aged 13 to 17.

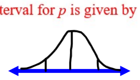
3. **Large Counts:** We check the counts of "successes" and "failures":

$$n\hat{p} = 246 \ge 10 \text{ and } n(1-\hat{p}) = 193 \ge 10$$

Therefore the sampling distribution of sample proportions is approximately normal.

4. The sample statistic is $\hat{p} = \frac{246}{439} = 0.56$. A 95% confidence interval for p is given by

$$\hat{p} \pm z^* \sqrt{\frac{\hat{p}(1-\hat{p})}{n}} = 0.56 \pm 1.96\sqrt{\frac{(0.56)(0.44)}{439}}$$
$$= 0.56 \pm 0.046$$
$$= (0.514, 0.606)$$



CONCLUDE: We are 95% confident that the interval from 0.514 to 0.606 captures the true proportion of 13- to 17-year-olds in the United States who would say that teens should wait until marriage to have sex.

Jan 4-12:04 PM

AP EXAM TIPS:

1. If a free-response question asks you to construct and interpret a confidence interval, you are expected to do the entire four step process. That includes clearly defining the parameter and checking conditions.

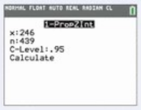
2. You may use your calculator to compute a confidence interval on the AP Exam. But there's a risk involved. If you just give the calculator answer with no work, you'll get either full credit for the "Do" step (if the interval is correct) or no credit (if it's wrong). If you opt for the calculator-only method, be sure to name the procedure (e.g., one-proportion z interval) and to give the interval (e.g., 0.514 to 0.607).

Jan 4-12:10 PM

The TI-83/84 and TI-89 can be used to construct a confidence interval for an unknown population proportion. We'll demonstrate using the previous example. Of $n = 439$ teens surveyed, $X = 246$ said they thought that young people should wait to have sex until after marriage. To construct a confidence interval:

TI-83/84

- Press STAT, then choose TESTS and 1-PropZInt.
- When the 1-PropZInt screen appears, enter $x = 246$,



Jan 5-12:18 PM

Day 2: Determine n

Choosing the Sample Size

In planning a study, we may want to choose a sample size that allows us to estimate a population proportion within a given margin of error.

statistic ± (critical value) · (standard deviation of statistic)   margin of error

The margin of error (ME) in the confidence interval for p is

$$ME = z^* \sqrt{\frac{\hat{p}(1-\hat{p})}{n}}$$

✓ z^* is the standard Normal critical value for the level of confidence we want.

Because the margin of error involves the sample proportion $\hat{p}$, we have to guess the latter value when choosing n. There are two ways to do this:

- Use a guess for $\hat{p}$ based on a pilot study or on past experience with similar studies. You should do several calculations that cover the range of $\hat{p}$-values you might get.
- ★ Use $\hat{p} = 0.5$ as the guess. The margin of error ME is largest when $\hat{p} = 0.5$, so this guess is conservative in the sense that if we get any other $\hat{p}$ when we do our study, we will get a margin of error smaller than planned.

Jan 5-12:21 PM

SAMPLE SIZE FOR DESIRED MARGIN OF ERROR

To determine the sample size n that will yield a C% confidence interval for a population proportion p with a maximum margin of error ME, solve the following inequality for n:

$$z^* \sqrt{\frac{\hat{p}(1-\hat{p})}{n}} \le ME$$

where $\hat{p}$ is a guessed value for the sample proportion. The margin of error will always be less than or equal to ME if you use $\hat{p} = 0.5$.

solve for n:

$n \ge (z^*/\text{error})^2 (p(1-p))$

Jan 5-12:23 PM

Example 10 **Customer Satisfaction**

Determining sample size

A company has received complaints about its customer service. The managers intend to hire a consultant to carry out a survey of customers. Before contacting the consultant, the company president wants some idea of the sample size that she will be required to pay for. One critical question is the degree of satisfaction with the company's customer service, measured on a 5-point scale. The president wants to estimate the proportion p of customers who are satisfied (that is, who choose either "somewhat satisfied" or "very satisfied," the two highest levels on the 5-point scale). She decides that she wants the estimate to be within 3% (0.03) at a 95% confidence level. How large a sample is needed?

PROBLEM: Determine the sample size needed to estimate p within 0.03 with 95% confidence.

$$z^* \sqrt{\frac{\hat{p}(1-\hat{p})}{n}} \le ME$$

$$1.96 \sqrt{\frac{0.5(1-0.5)}{n}} \le 0.03$$

or $n \ge (z^*/\text{error})^2 (p(1-p))$

$$n \ge \left(\frac{1.96}{0.03}\right)^2 (0.5)(1-0.5)$$

answer:

[redacted]

Jan 5-12:26 PM

If you want a 2.5% margin of error rather than 3%, then

$$n \ge \left(\frac{1.96}{0.025}\right)^2 (0.5)(1-0.5) = 1536.64 \Rightarrow n = 1537$$

For a 2% margin of error, the sample size you need is

$$n \ge \left(\frac{1.96}{0.02}\right)^2 (0.5)(1-0.5) = 2401$$

smaller margins of error call for larger samples.

Jan 5-3:24 PM

What if the $\hat{p}$ is given?

1. In the company's prior-year survey, 80% of customers surveyed said they were "somewhat satisfied" or "very satisfied." Using this value as a guess for $\hat{p}$, find the sample size needed for a margin of error of 3% at a 95% confidence level.

Show Answer

2. What if the company president demands 99% confidence instead? Determine how this would affect your answer to Question 1.

Show Answer

Jan 5-3:30 PM