+

Unit 5: Estimating with Confidence

Section 8.2

Estimating a Population Proportion

The Practice of Statistics, 4th edition – For AP*

STARNES, YATES, MOORE

Unit 5
Estimating with Confidence

- 8.1 Confidence Intervals: The Basics
- **8.2 Estimating a Population Proportion**
- 8.3 Estimating a Population Mean

Section 8.2
Estimating a Population Proportion

Learning Objectives

After this section, you should be able to…

- ✓ CONSTRUCT and INTERPRET a confidence interval for a population proportion
- ✓ DETERMINE the sample size required to obtain a level *C* confidence interval for a population proportion with a specified margin of error
- ✓ DESCRIBE how the margin of error of a confidence interval changes with the sample size and the level of confidence *C*

Activity: The Beads

Your teacher has a container full of different colored beads. Your goal is to estimate the actual proportion of red beads in the container.

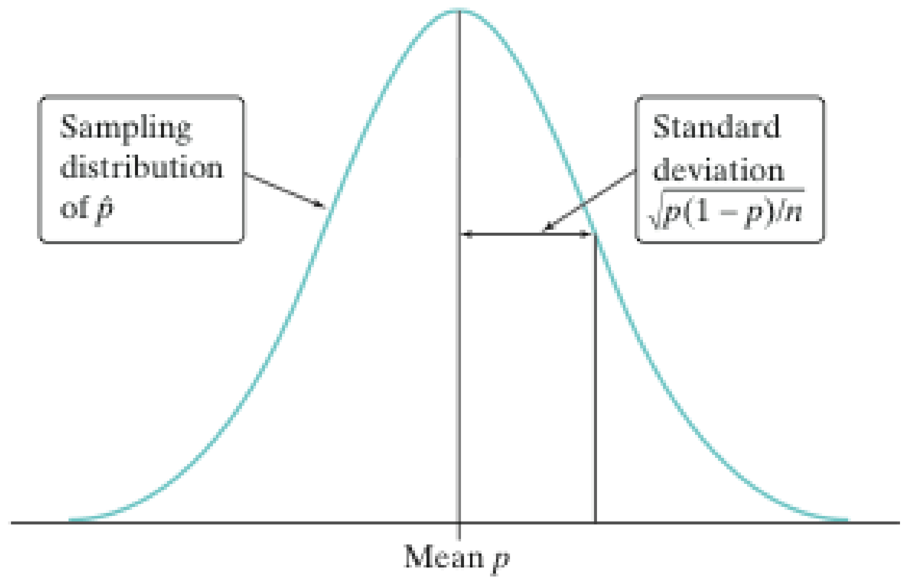
- ✓ Determine how to use a cup to get a simple random sample of beads from the container.
- ✓ Each team is to collect one SRS of beads.
- ✓ Determine a point estimate for the unknown population proportion.
- ✓ Find a 90% confidence interval for the parameter *p*. Consider any conditions that are required for the methods you use.

Data Collection Shows

- Suppose one SRS of beads resulted in 107 red beads and 144 beads of another color. The point estimate for the unknown proportion *p* of red beads in the population would be

$$\hat{p} = \frac{107}{251} = 0.426$$

How can we use this information to find a confidence interval for *p*?



In practice, we do not know the value of p. If we did, we would not need to construct a confidence interval for it! In large samples, $\hat{p}$ will be close to p, so we will replace p with $\hat{p}$ in checking the Normal condition.

Conditions for Estimating *p*

Check the conditions for estimating *p* from our sample. $\hat{p} = \frac{107}{251} = 0.426$

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

Normal: Both *np* and *n*(1 – *p*) must be greater than 10. Since we don't know *p*, we check that

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

The counts of successes (red beads) and failures (non-red) are both ≥ 10.

Independent: Since the class sampled without replacement, they need to check the 10% condition. At least 10(251) = 2510 beads need to be in the population. The teacher reveals there are 3000 beads in the container, so the condition is satisfied.

Since all three conditions are met, it is safe to construct a confidence interval.

Constructing a Confidence Interval for *p*

We can use the general formula from Section 10.1 to construct a confidence interval for an unknown population proportion *p*:

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

The sample proportion $\hat{p}$ is the statistic we use to estimate p. When the Independent condition is met, the standard deviation of the sampling distibution of $\hat{p}$ is

$$\sigma_{\hat{p}} = \sqrt{\frac{p(1-p)}{n}}$$

Since we don't know p, we replace it with the sample proportion $\hat{p}$. This gives us the **standard error (SE)** of the sample proportion:

$$\sqrt{\frac{\hat{p}(1-\hat{p})}{n}}$$

Definition:

When the standard deviation of a statistic is estimated from data, the results is called the **standard error** of the statistic.

One-Sample z Interval for a Population Proportion

Once we find the critical value z^*, our confidence interval for the population proportion p is

statistic $\pm$ (critical value) $\cdot$ (standard deviation of statistic)

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

One-Sample z Interval for a Population Proportion

Choose an SRS of size n from a large population that contains an unknown proportion p of successes. An approximate level C **confidence interval for p is**

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

where z^* is the critical value for the standard Normal curve with area C between $-z^*$ and z^*.

Use this interval only when the numbers of successes and failures in the sample are both at least 10 and the population is at least 10 times as large as the sample.

The Four-Step Process

We can use the familiar four-step process whenever a problem asks us to construct and interpret a confidence interval.

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.

One-Sample *z* Interval for a Population Proportion

State: We want to calculate and interpret a 90% confidence interval for the proportion of red beads in the container.

Plan:

✓ sample proportion = 107/251 = 0.426

✓ We checked the conditions earlier.

For a 90% confidence level, $z^* = 1.645$

Do:

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

statistic ± (critical value) • (standard deviation of the statistic)

$$\hat{p} \pm z^* \sqrt{\frac{\hat{p}(1-\hat{p})}{n}} \qquad = 0.426 \pm 1.645\sqrt{\frac{(0.426)(1-0.426)}{251}} \qquad \begin{aligned} &= 0.426 \pm 0.051 \\ &= (0.375,\ 0.477) \end{aligned}$$

Conclude:

We are 90% confident that the interval from 0.375 to 0.477 captures the actual proportion of red beads in the container.

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. 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 past experience or a pilot study
- Use $\hat{p} = 0.5$ as the guess. ME is largest when $\hat{p} = 0.5$

Sample Size for Desired Margin of Error

To determine the sample size n that will yield a level 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 take the guess $\hat{p}$ to be 0.5.

Example: Customer Satisfaction

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 five-point scale. The president wants to estimate the proportion ***p*** of customers who are satisfied (that is, who choose either "satisfied" or "very satisfied", the two highest levels on the five-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?

Example: Customer Satisfaction

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

- ✓ The critical value for 95% confidence is $z^* = 1.96$.
- ✓ Since the company president wants a margin of error of no more than 0.03, we need to solve the equation

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

Multiply both sides by square root *n* and divide both sides by 0.03.

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

Square both sides.

$$\left(\frac{1.96}{0.03}\right)^2 \hat{p}(1-\hat{p}) \le n$$

Substitute 0.5 for the sample proportion to find the largest *ME* possible.

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

$$1067.111 \le n$$

We round up to 1068 respondents to ensure the margin of error is no more than 0.03 at 95% confidence.

Looking Ahead…

In the next Section…

We’ll learn how to estimate a population mean.

We’ll learn about

- ✓ **The one-sample *z* interval for a population mean when *σ* is known**
- ✓ **The *t* distributions when *σ* is unknown**
- ✓ **Constructing a confidence interval for *μ***
- ✓ **Using *t* procedures wisely**

Homework

Textbook – Chapter 12 #'s, 1-4, 6, 7, 10, 11

Additional notes packet – Read and answer all questions, pgs. 14-19