

## 9.1 Estimating a Population Proportion

A **point estimate** is the value of a statistic that estimates the value of a parameter.

For example, the point estimate for the population proportion is  $\hat{p} = \frac{x}{n}$ , where  $x$  is the number of individuals in the sample with a specified characteristic and  $n$  is the sample size (note:  $\hat{p}$  is read “p-hat”).

**Example:** *Calculating a Point Estimate for the Population Proportion*

In July of 2008, a Quinnipiac University Poll asked 1783 registered voters nationwide whether they favored or opposed the death penalty for persons convicted of murder. 1123 were in favor.

Obtain a point estimate for the proportion of registered voters nationwide who are in favor of the death penalty for persons convicted of murder.

A **confidence interval** for an unknown parameter consists of an interval of numbers based on a point estimate.

The **level of confidence** represents the expected proportion of intervals that will contain the parameter if a large number of different samples is obtained.

The level of confidence is denoted  $(1 - \alpha) \cdot 100\%$ .

**For example**, a 95% level of confidence ( $\alpha = 0.05$ ) implies that if 100 different confidence intervals are constructed, each based on a different sample from the same population, **we will expect 95 of the intervals to contain the parameter and 5 not to include the parameter.**

Confidence interval estimates for the population proportion are of the form

**Point estimate  $\pm$  margin of error.**

The margin of error of a confidence interval estimate of a parameter is a measure of how accurate the point estimate is.

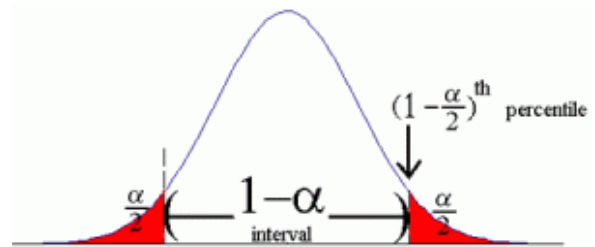
The **margin of error** depends on three factors:

- **Level of confidence:** As the level of confidence increases, the margin of error also increases.
- **Sample size:** As the size of the random sample increases, the margin of error decreases.
- **Standard deviation of the population:** The more spread there is in the population, the wider our interval will be for a given level of confidence.

## What is this $\alpha$ (alpha level) you speak of?

If we want an interval of a given width  $(1 - \alpha)$  on a standard normal distribution curve, the number  $\alpha$  is the area on the outside of this interval.

$$100\% - \text{Confidence Level}\% = \alpha$$



How do you find  $z_{\alpha/2}$ ?

1. You can use this table of common z-scores below

| Confidence Level | Alpha | Alpha/2 | z alpha/2 |
|------------------|-------|---------|-----------|
| 90%              | 10%   | 5.0%    | 1.645     |
| 95%              | 5%    | 2.5%    | 1.96      |
| 98%              | 2%    | 1.0%    | 2.326     |
| 99%              | 1%    | 0.5%    | 2.576     |

2. You can  $z_{\alpha/2}$  by looking up the z-score that has an area of  $\frac{\alpha}{2}$  **to the right** of it. To do this, look up the z-score for  $50\% - \frac{\alpha}{2}$  in a z-table. So, to find a 90% CL, look up the z-score for
3. You can use your TI-84 with the command like “invNorm(.05)” and drop the negative, or find the right hand z-score with the command “invNorm(.95)”. (The .95 is the right hand percentile because it is  $100\% - 5\% = 95\%$  which leaves 5% of the area to the right of 1.645.)

## Sampling Distribution of $\hat{p}$

For a simple random sample of size  $n$ , the sampling distribution of  $\hat{p}$  is approximately normal with mean  $\mu_{\hat{p}} = p$  and standard deviation  $\sigma_{\hat{p}} = \sqrt{\frac{p(1-p)}{n}}$ , provided that  $np(1-p) \geq 10$ .

**NOTE:** We also require that each trial be independent; when sampling from finite populations ( $n \leq 0.05N$ ).

## Constructing a $(1 - \alpha) \cdot 100\%$ Confidence Interval for a Population Proportion

### **Constructing a Confidence Interval for a Population Proportion**

Suppose that a simple random sample of size  $n$  is taken from a population. A  $(1 - \alpha) \cdot 100\%$  confidence interval for  $p$  is given by the following quantities

$$\text{Lower bound: } \hat{p} - z_{\frac{\alpha}{2}} \cdot \sqrt{\frac{\hat{p}(1-\hat{p})}{n}}$$

$$\text{Upper bound: } \hat{p} + z_{\frac{\alpha}{2}} \cdot \sqrt{\frac{\hat{p}(1-\hat{p})}{n}}$$

**Note:** It must be the case that  $n\hat{p}(1 - \hat{p}) \geq 10$  and  $n \leq 0.05N$  to construct this interval.

### Interpretation of a Confidence Interval

A  $(1 - \alpha) \cdot 100\%$  confidence interval indicates that  $(1 - \alpha) \cdot 100\%$  of all simple random samples of size  $n$  from the population whose parameter is unknown will result in an interval that contains the parameter.

### **Example: Constructing a Confidence Interval from a Population Proportion**

In July of 2008, a Quinnipiac University Poll asked 1783 registered voters nationwide whether they favored or opposed the death penalty for persons convicted of murder. 1123 were in favor.

Obtain a 90% confidence interval for the proportion of registered voters nationwide who are in favor of the death penalty for persons convicted of murder.

- Step 1: Find  $\hat{p}$
- Step 2: Is  $n\hat{p}(1 - \hat{p}) \geq 10$ ?
- Step 3: Define  $\alpha$  and  $z_{\alpha/2}$
- Step 4: Calculate the Lower and Upper Bounds

- Step 5: State your solution...

We are 90% confident that the proportion of registered voters who are in favor of the death penalty for those convicted of murder is between \_\_\_\_\_ and \_\_\_\_\_.

**On TI-84, use:** [STAT] – [Tests] – [1PropZinterval]

## Margin of Error

The **margin of error**,  $E$ , in a  $(1 - \alpha)$  100% confidence interval for a population proportion is given by

$$E = z_{\frac{\alpha}{2}} \cdot \sqrt{\frac{\hat{p}(1 - \hat{p})}{n}}$$

### Example

In July of 2008, a Quinnipiac University Poll asked 1783 registered voters nationwide whether they favored or opposed the death penalty for persons convicted of murder. 1123 were in favor. Find the margin of error for a 90% confidence interval.

Note: From above, we found  $\hat{p} = 0.63$ , and  $z_{0.05} = 1.645$

## Finding the Sample Size for a Given Error

The **Sample Size** needed for a specified margin of error,  $E$ , and level of confidence  $(1 - \alpha)$ :

If we know  $\hat{p}$ :

$$n = \hat{p}(1 - \hat{p}) \left( \frac{z_{\frac{\alpha}{2}}}{E} \right)^2$$

If we don't know  $\hat{p}$ , let  $\hat{p} = .5$ :

$$n = .25 \left( \frac{z_{\frac{\alpha}{2}}}{E} \right)^2$$

### Example: Determining Sample Size

A sociologist wanted to determine the percentage of residents of America that only speak English at home. What size sample should be obtained if she wishes her estimate to be within 3 percentage points with 90% confidence assuming she uses the 2000 estimate obtained from the Census 2000 Supplementary Survey of 82.4%?

- Step 1: Find  $E$
- Step 2: Find  $z_{\frac{\alpha}{2}}$
- Step 3: Find  $\hat{p}$
- Step 4: Find  $n$

Solution: The sociologist must survey \_\_\_\_\_ randomly selected American residents.