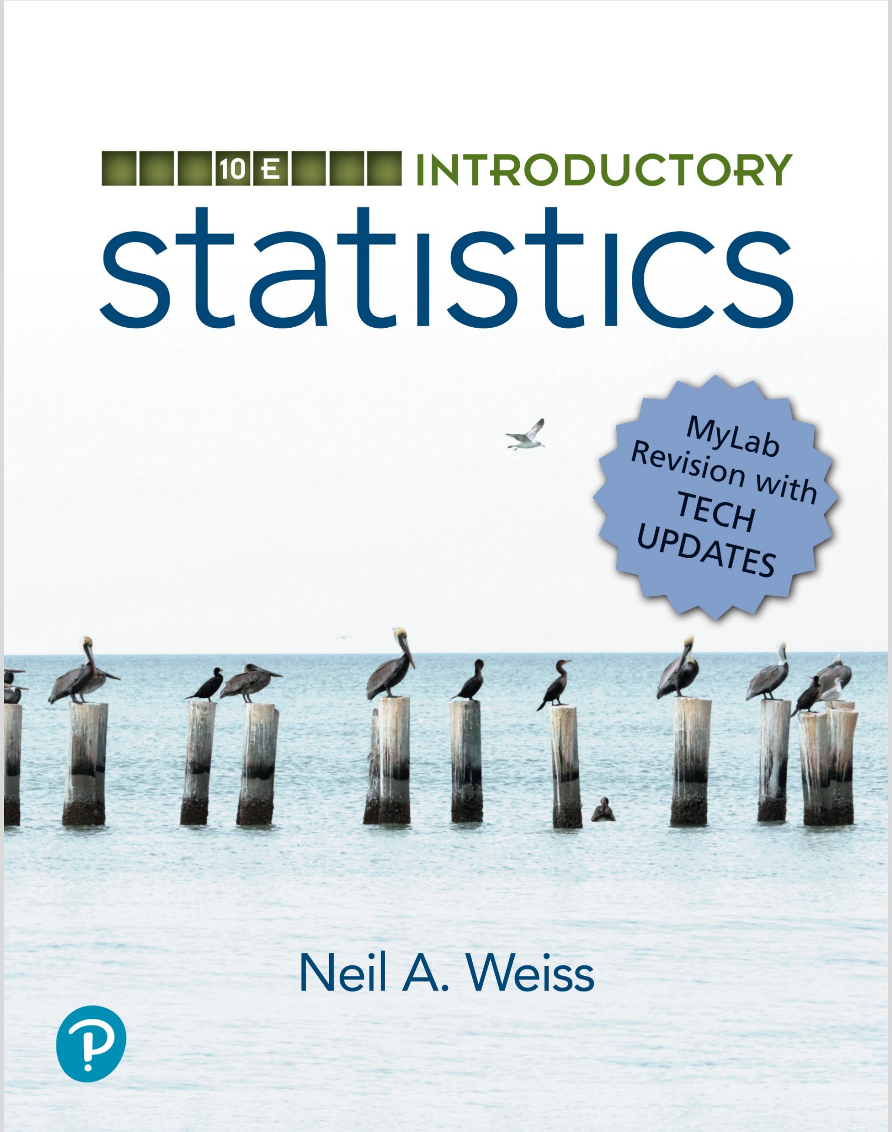
Chapter 12

Inferences for Population Proportions



Section 12.1

Confidence Intervals for One Population Proportion

Definition 12.1

Population Proportion and Sample Proportion

Consider a population in which each member either has or does not have a specified attribute. Then we use the following notation and terminology.

Population proportion, p: The proportion (percentage) of the entire population that has the specified attribute.

Sample proportion, $\hat{p}$: The proportion (percentage) of a sample from the population that has the specified attribute.

Formula 12.1

Sample Proportion

A sample proportion, $\hat{p}$, is computed by using the formula

$$\hat{p} = \frac{x}{n},$$

where x denotes the number of members in the sample that have the specified attribute and, as usual, n denotes the sample size.

Key Fact 12.1

The Sampling Distribution of the Sample Proportion

For samples of size n,

- the mean of $\hat{p}$ equals the population proportion: $\mu_{\hat{p}} = p$ (i.e., the sample proportion is an unbiased estimator of the population proportion);
- the standard deviation of $\hat{p}$ equals the square root of the product of the population proportion and one minus the population proportion divided by the sample size: $\sigma_{\hat{p}} = \sqrt{p(1 - p)/n}$; and
- $\hat{p}$ is approximately normally distributed for large n.

Procedure 12.1

One-Proportion z-Interval Procedure

Purpose To find a confidence interval for a population proportion, p

Assumptions

1. Simple random sample
2. The number of successes, x, and the number of failures, $n - x$, are both 5 or greater.

Step 1 **For a confidence level of $1 - \alpha$, use Table II to find $z_{\alpha/2}$.**

Step 2 **The confidence interval for p is from**

$$\hat{p} - z_{\alpha/2} \cdot \sqrt{\hat{p}(1-\hat{p})/n} \quad \text{to} \quad \hat{p} + z_{\alpha/2} \cdot \sqrt{\hat{p}(1-\hat{p})/n},$$

where $z_{\alpha/2}$ is found in Step 1, n is the sample size, and $\hat{p} = x/n$ is the sample proportion.

Step 3 **Interpret the confidence interval.**

Formula 12.2

Margin of Error for the Estimate of *p*

The **margin of error** for the estimate of *p* is

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

Formula 12.3

Sample Size for Estimating p

- A $(1 - \alpha)$-level confidence interval for a population proportion that has a margin of error of at most E can be obtained by choosing

$$n = 0.25\left(\frac{z_{\alpha/2}}{E}\right)^2$$

rounded up to the nearest whole number.
- If you can make an educated guess, $\hat{p}_g$ (g for guess), for the observed value of $\hat{p}$, then you should instead choose

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

rounded up to the nearest whole number.
- If you have in mind a likely range for the observed value of $\hat{p}$, then you should apply the preceding formula with your educated guess for the observed value of $\hat{p}$ being the value in the range closest to 0.5.

Section 12.2

Hypothesis Tests for One Population Proportion

Procedure 12.2

One-Proportion z-Test

Purpose To perform a hypothesis test for a population proportion, p

Assumptions

1. Simple random sample

2. Both np_0 and $n(1 - p_0)$ are 5 or greater

Step 1 **The null hypothesis is H_0: $p = p_0$, and the alternative hypothesis is**

H_a: $p \neq p_0$ (Two tailed)	**or**	**H_a: $p < p_0$ (Left tailed)**	**or**	**H_a: $p > p_0$ (Right tailed)**

Step 2 **Decide on the significance level, α.**

Step 3 **Compute the value of the test statistic**

$$z = \frac{\hat{p} - p_0}{\sqrt{p_0(1 - p_0)/n}}$$

and denote that value z_0.

Procedure 12.2 (cont.)

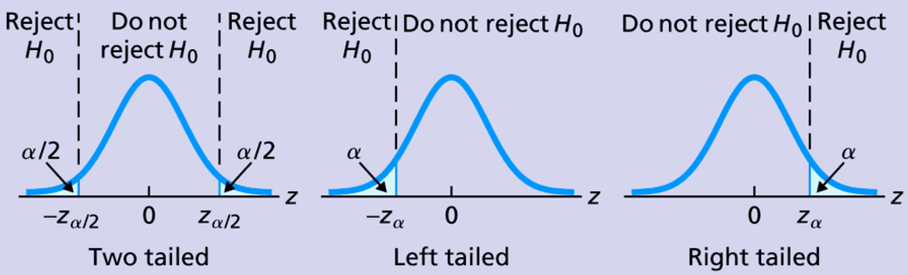
CRITICAL-VALUE APPROACH	OR	*P*-VALUE APPROACH

CRITICAL-VALUE APPROACH

Step 4 The critical value(s) are

$\pm z_{\alpha/2}$ (Two tailed) or $-z_{\alpha}$ (Left tailed) or z_{α} (Right tailed)

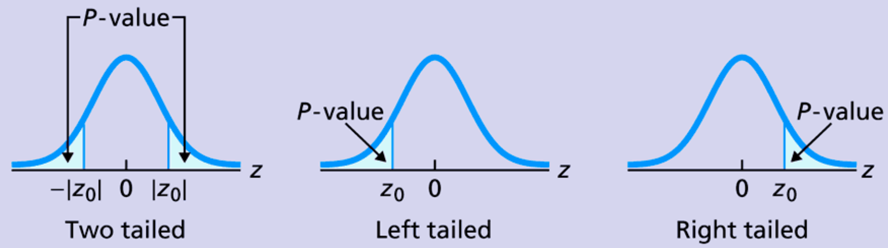
Use Table II to find the critical value(s).



Step 5 If the value of the test statistic falls in the rejection region, reject H_0; otherwise, do not reject H_0.

P-VALUE APPROACH

Step 4 Use Table II to obtain the *P*-value.



Step 5 If $P \leq \alpha$, reject H_0; otherwise, do not reject H_0.

Step 6 Interpret the results of the hypothesis test.

Section 12.3

Inferences for Two Population Proportions

Key Fact 12.2

The Sampling Distribution of the Difference Between Two Sample Proportions for Independent Samples

For independent samples of sizes n_1 and n_2 from the two populations,

- $\mu_{\hat{p}_1-\hat{p}_2} = p_1 - p_2$ (i.e., the difference between sample proportions is an unbiased estimator of the difference between population proportions),
- $\sigma_{\hat{p}_1-\hat{p}_2} = \sqrt{p_1(1-p_1)/n_1 + p_2(1-p_2)/n_2}$, and
- $\hat{p}_1 - \hat{p}_2$ is approximately normally distributed for large n_1 and n_2.

Procedure 12.3

Two-Proportions z-Test

Purpose To perform a hypothesis test to compare two population proportions, p_1 and p_2

Assumptions

1. Simple random samples
2. Independent samples
3. x_1, $n_1 - x_1$, x_2, and $n_2 - x_2$ are all 5 or greater

Step 1 **The null hypothesis is H_0: $p_1 = p_2$, and the alternative hypothesis is**

$$H_a: p_1 \neq p_2 \text{ (Two tailed)} \quad \text{or} \quad H_a: p_1 < p_2 \text{ (Left tailed)} \quad \text{or} \quad H_a: p_1 > p_2 \text{ (Right tailed)}$$

Step 2 **Decide on the significance level, α.**

Step 3 **Compute the value of the test statistic**

$$z = \frac{\hat{p}_1 - \hat{p}_2}{\sqrt{\hat{p}_p(1 - \hat{p}_p)}\sqrt{(1/n_1) + (1/n_2)}},$$

where $\hat{p}_p = (x_1 + x_2)/(n_1 + n_2)$. Denote the value of the test statistic z_0.

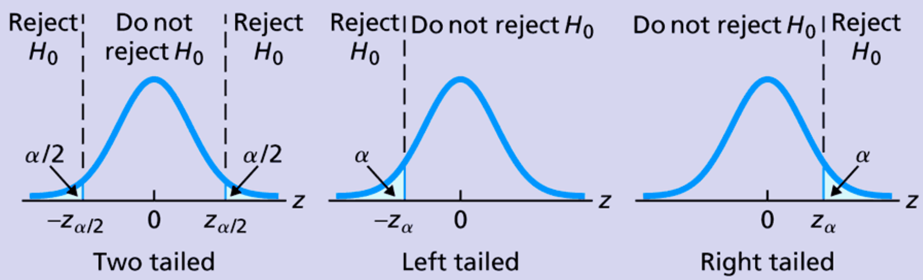
Procedure 12.3 (cont.)

CRITICAL-VALUE APPROACH	OR	*P*-VALUE APPROACH

Step 4 The critical value(s) are

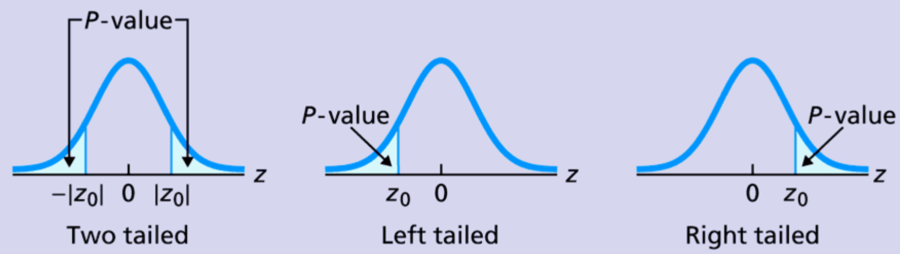
$\pm z_{\alpha/2}$ (Two tailed) or $-z_{\alpha}$ (Left tailed) or z_{α} (Right tailed)

Use Table II to find the critical value(s).



Step 5 If the value of the test statistic falls in the rejection region, reject H_0; otherwise, do not reject H_0.

Step 4 Use Table II to obtain the *P*-value.



Step 5 If $P \leq \alpha$, reject H_0; otherwise, do not reject H_0.

Step 6 Interpret the results of the hypothesis test.

Procedure 12.4

Two-Proportions z-Interval Procedure

Purpose To find a confidence interval for the difference between two population proportions, p_1 and p_2

Assumptions

1. Simple random samples
2. Independent samples
3. $x_1, n_1 - x_1, x_2$, and $n_2 - x_2$ are all 5 or greater

Step 1 **For a confidence level of $1 - \alpha$, use Table II to find $z_{\alpha/2}$.**

Step 2 **The endpoints of the confidence interval for $p_1 - p_2$ are**

$$(\hat{p}_1 - \hat{p}_2) \pm z_{\alpha/2} \cdot \sqrt{\hat{p}_1(1-\hat{p}_1)/n_1 + \hat{p}_2(1-\hat{p}_2)/n_2}.$$

Step 3 **Interpret the confidence interval.**