Section 5.5
Sample Size estimates for a One Sample Dichotomous Outcome
Calculate and interpret sample size estimates for a one sample dichotomous outcome

  Estimate for a confidence interval

  Estimate for a hypothesis test
Dichotomous Outcome – One Sample

Sample Size to Estimate C.I.

C.I. for \( p \)

\[
n = p(1-p) \left( \frac{Z}{E} \right)^2
\]

Sample Size for Hypothesis Test

Hypothesis Test

\[
H_0: p = p_0
\]

\[
n = \left( \frac{Z_{1-\alpha/2} + Z_{1-\beta}}{ES} \right)^2
\]

Effect Size (ES)

\[
ES = \frac{|p_1 - p_0|}{\sqrt{p_0(1-p_0)}}
\]
Dichotomous Outcome – One Sample (C.I.)

Sample Size to Estimate C.I.

Estimate required sample size for 95% C.I. of the proportion of freshman at a college who currently smoke cigarettes.

C.I. for \( p \)

\[
n = p(1-p) \left( \frac{Z}{E} \right)^2
\]

Parameters:
- Margin of error: 5%
- Assumed prevalence (i): 50% (i.e. no data available)
- Assumed prevalence (ii): 27% (i.e. national data)
- Desired C.I.: 95% (i.e. \( z = 1.96 \))

\[
\begin{align*}
n &= 0.5(1-0.5) \left( \frac{1.96}{0.05} \right)^2 = 384.2 \\
n &= 0.27(1-0.27) \left( \frac{1.96}{0.05} \right)^2 = 302.9
\end{align*}
\]
Estimate required sample size for 95% C.I. of the proportion of male freshman at a college who report binge drinking at least once a month.

\[ n = p(1-p) \left( \frac{Z}{E} \right)^2 \]

**Parameters:**
- Margin of error: 5%
- Assumed prevalence (ii): 14% (i.e. national data)
- Desired C.I.: 95% (i.e. \( z = 1.96 \))

\[ n = \left( \frac{1.96}{0.05} \right)^2 \times 0.14 \times (1-0.14) \]

\[ n = \left( \frac{1.96}{0.05} \right)^2 \times 0.14 \times 0.86 \]

\[ n = 3.84^2 \times 0.14 \times 0.86 \]

\[ n = 14.736 \times 0.14 \times 0.86 \]

\[ n = 1.84 \]

\[ n = \text{round up to the nearest whole number} \]

\[ n = 2 \]
Estimate required sample size for 95% C.I. of the proportion of male freshman at a college who report binge drinking at least once a month.

**Dichotomous Outcome – One Sample (C.I.)**

**Practice Exercise**

**Sample Size to Estimate C.I.**

Estimate required sample size for 95% C.I. of the proportion of male freshman at a college who report binge drinking at least once a month.

**C.I. for \( p \)**

\[
 n = p(1-p) \left( \frac{Z}{E} \right)^2
\]

**Parameters:**

- Margin of error: 5%
- Assumed prevalence (ii) 14% (i.e. national data)
- Desired C.I.: 95% (i.e. \( z = 1.96 \))

\[
 n = 0.14(1-0.14) \left( \frac{1.96}{0.05} \right)^2 = 185.0
\]
Dichotomous Outcome – One Sample (H₀ Test)

Sample Size for Hypothesis Test

Compare prevalence of high LDL cholesterol (>159 mg/dL) between patients with a history of CVD and a population estimate of those without CVD.

**H₀: p = p₀**

\[
n = \left( \frac{Z_{1-\alpha/2} + Z_{1-\beta}}{ES} \right)^2
\]

**Parameters/Assumptions:**

- General population proportion: \( p_0 = 0.26 \) (26%)
- Clinically important difference: 0.05 (5%)
- 2-sided type I error rate (\( \alpha \)): 0.05
- Desired power (1-\( \beta \)): 0.90

\[
ES = \frac{|p_1 - p_0|}{\sqrt{p_0(1-p_0)}}
\]

\[
ES = \frac{|0.31 - 0.26|}{\sqrt{0.26(1-0.26)}} = 0.114
\]

\[
n = \frac{1.96 + 1.282}{0.114}^2 = 808.9
\]

A sample size of \( n = 809 \) will ensure that a 2-sided test with \( \alpha = 0.05 \) has 90% power to detect a difference of 0.05, or 5%, in the proportion of patients with a history of CVD who have an elevated LDL cholesterol level.
Dichotomous Outcome – One Sample (H₀ Test) Practice Exercise

Sample Size for Hypothesis Test

Compare the prevalence of the metabolic syndrome between patients with a history of alcohol abuse and a population estimate of those without.

\[ H₀: \ p = p₀ \]

\[ n = \left[ \frac{Z_{1-\alpha/2} + Z_{1-\beta}}{ES} \right]² \]

Parameters/Assumptions:
- General population proportion: \( p₀ = 0.22 \) (22%)
- Clinically important difference: 0.06 (6%)
- 2-sided type I error rate (\( \alpha \)): 0.05
- Desired power (1-\( \beta \)): 0.80

\[ ES = \frac{|p₁ - p₀|}{\sqrt{p₀(1-p₀)}} \]

\[ ES = \sqrt{\frac{|p₁ - p₀|}{p₀(1-p₀)}} = \text{_______} \]

\[ n = \left[ \frac{Z_{1-\alpha/2} + Z_{1-\beta}}{ES} \right]² \]

\[ n = \text{_______} \]
Dichotomous Outcome – One Sample (H₀ Test)(Practice)

Sample Size for Hypothesis Test

Compare the prevalence of the metabolic syndrome between patients with a history of alcohol abuse and a population estimate of those without.

H₀: \( p = p₀ \)

\[
n = \left( \frac{Z_{1-\alpha/2} + Z_{1-\beta}}{\text{ES}} \right)^2
\]

\[
\text{ES} = \left| \frac{p₁ - p₀}{\sqrt{p₀(1-p₀)}} \right|
\]

\[
\text{ES} = \left| \frac{0.28 - 0.22}{\sqrt{0.22(1-0.22)}} \right| = 0.145
\]

\[
n = \left( \frac{1.96 + 0.84}{0.145} \right)^2 = 373.7
\]

A sample size of \( n = 374 \) will ensure that a 2-sided test with \( \alpha = 0.05 \) has 80% power to detect a difference of 0.06, or 6%, in the proportion of patients with a history of alcohol abuse who have the metabolic syndrome.