A. Textbook

- 8.52
  p_1 = proportion of high-tech companies which offer stock options
  p_2 = proportion of non-high-tech companies which offer stock options

  Sample Data
  $n_1 = 91 \quad x_1 = 73 \quad \Rightarrow \widehat{p}_1 = 73/91 = .8022$
  $n_2 = 109 \quad x_2 = 75 \quad \Rightarrow \widehat{p}_2 = 75/109 = .6881$

  (a) Find a 95% confidence interval for $(p_1 - p_2)$ by formula:

$$\sigma_{\widehat{p}_1-\widehat{p}_2} = \sqrt{\frac{\widehat{p}_1\,(1-\widehat{p}_1)}{n_1} + \frac{\widehat{p}_2\,(1-\widehat{p}_2)}{n_2}}$$

$$= \sqrt{\frac{(.8022)(.1978)}{91} + \frac{(.6881)(.3119)}{109}}$$

$$= .0609$$

  So a 95% confidence interval for $(p_1 - p_2)$ is

$$(\widehat{p}_1 - \widehat{p}_2) \pm z^* \,\sigma_{\widehat{p}_1-\widehat{p}_2} = (.8022 - .6881) \pm (1.96)(.0609)$$

$$= .1141 \pm .1194$$

$$= \boxed{(-.0053,\ .2335)}$$

  (b) Interpret:

  We are 95% confident that between 0.53% fewer and 23.36% more high-tech companies than non-high-tech-companies offer stock options.

- 8.56

  Four Steps:

  1. (The proportions were defined in the previous exercise.)

$$H_A\colon\ p_1 \neq p_2 \ \text{ or } \ (p_1 - p_2) \neq 0$$
$$H_0\colon\ p_1 = p_2 \ \text{ or } \ (p_1 - p_2) = 0$$

  2. P-value by formula:

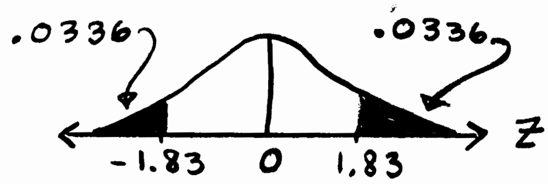
$$\widehat{p} = \frac{x_1 + x_2}{n_1 + n_2} = \frac{73 + 75}{91 + 109} = \frac{148}{200} = 0.74$$

$$Z = \frac{(\widehat{p}_1 - \widehat{p}_2)}{\sqrt{\left(\frac{1}{n_1} + \frac{1}{n_2}\right)\widehat{p}\,(1 - \widehat{p}\,)}}$$

$$= \frac{(.8022 - .6881)}{\sqrt{\left(\frac{1}{91} + \frac{1}{109}\right)(0.74)(0.26)}}$$

$$= \frac{.1141}{.0623}$$

$$= \boxed{1.83}$$



P-value $= P(Z < -1.83) + P(Z > 1.83) = 2(.0336) = \boxed{.0672}$

3. Fail to Reject H_0 since P-value $= 0.0672 > .05 = \alpha$.

4. There is insufficient evidence to show that high-tech firms differ from non-high tech firms in the provision of stock options.

- MINITAB 8.52: $(-0.00530103,\ 0.233550)$

- MINITAB 8.56: P-value $= \boxed{0.067}$

- 8.53

(a)

p_1 = proportion of complainers who leave HMO
p_2 = proportion of non-complainers who leave HMO

Sample Data
$n_1 = 639 \quad x_1 = 54 \quad \Rightarrow \widehat{p}_1 = 54/639 = .0845$
$n_2 = 743 \quad x_2 = 22 \quad \Rightarrow \widehat{p}_2 = 22/743 = .0296$

$$\sigma_{\widehat{p}_1 - \widehat{p}_2} = \sqrt{\frac{\widehat{p}_1\,(1 - \widehat{p}_1)}{n_1} + \frac{\widehat{p}_2\,(1 - \widehat{p}_2)}{n_2}}$$

$$= \sqrt{\frac{(.0845)(.9155)}{639} + \frac{(.0296)(.9704)}{743}}$$

$$= .0126$$

A 95% confidence interval for $(p_1 - p_2)$ is

$$(\widehat{p}_1 - \widehat{p}_2) \pm z^*\,\sigma_{\widehat{p}_1 - \widehat{p}_2} = (.0845 - .0296) \pm (1.96)(.0126)$$

$$= .0549 \pm .0247$$

$$= \boxed{(.0302,\ .0796)}$$

(b) Interpret:

We are 95% confident that between 3.01% and 7.97% more complaining patients than non-complaining patients leave the HMO.

- 8.57

Four Steps:

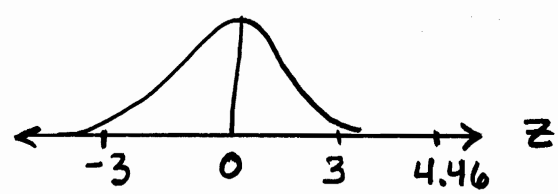
1. (Proportions are already defined.)

$$H_A\colon p_1 > p_2 \text{ or } (p_1 - p_2) > 0$$
$$H_0\colon p_1 \le p_2 \text{ or } (p_1 - p_2) \le 0$$

2. Z statistic and P-value:

$$\widehat{p} = \frac{x_1 + x_2}{n_1 + n_2} = \frac{54 + 22}{639 + 743} = \frac{76}{1382} = 0.0550$$

$$Z = \frac{(\widehat{p}_1 - \widehat{p}_2)}{\sqrt{\left(\frac{1}{n_1} + \frac{1}{n_2}\right)\widehat{p}\,(1 - \widehat{p}\,)}}$$
$$= \frac{(.0845 - .0296)}{\sqrt{\left(\frac{1}{639} + \frac{1}{743}\right)(0.0550)(0.9450)}}$$
$$= \frac{.0549}{.0123}$$
$$= \boxed{4.46}$$



P-value $= P(Z > 4.46) \approx \boxed{0}$

3. Reject H_0 since P-value $= 0 < .05 = \alpha$.

4. There is sufficient evidence to show that complaining patients are more likely to leave the HMO than non-complaining patients.

- MINITAB 8.53: (0.0301254, 0.0796693)

- MINITAB 8.57: P-value $= \boxed{0.000}$

(continued)

- 8.68

  (a)
  1. p_1 = proportion of men employed in summer
     p_2 = proportion of women employed in summer

     $$H_A\colon\ p_1 \neq p_2$$
     $$H_0\colon\ p_1 = p_2$$

  2. From MINITAB, $Z = 2.59$, P-value= 0.010
  3. Reject H_0 since $P\text{-value} = 0.010 < 0.05 = \alpha$
  4. There is enough evidence to confirm a difference in the summer employment rate between men and women.

  (b) From MINITAB, a 95% CI is

  $$(.0100, .0730)$$

  (c) With 95% confidence, between 1% and 7.30% more undergrad men than women work during the summer.

  (d)

  * If the low figure of 1% is correct:

  $$1\% \times 15,000 = 150 \text{ fewer women}$$

  * If the high figure of 7.3% is correct:

  $$7.3\% \times 15,000 = 1095 \text{ fewer women}$$

  So between 150 and 1095 fewer UI women than UI men undergrads work during the summer, with 95% confidence.

(continued)

B. Additional

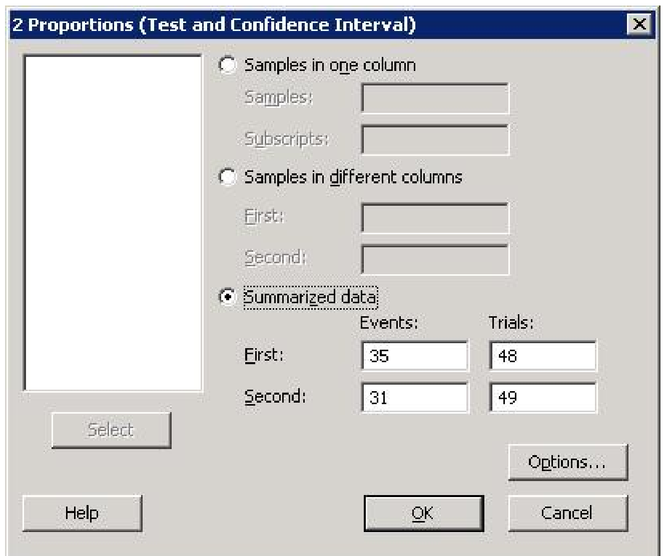
(1) (a) The analysis is not correct. The standard deviation for each of the two groups should be *combined* to produce a standard deviation for the *difference* in proportions. (See Example 1 in the Topic 3 Notes for more details.)

(2)

(a) Credit card users make between 8.79% fewer and 28.10% more impulse buys than non-credit-card users.

(b) A careful look at the output shows that the friend apparently entered the number of planned purchases (not impulse buys) into MINITAB. (See the screenshot below.)

Therefore the interpretation in (a) is incorrect.



(c) The friend analyzed the proportion of planned purchases:

$1 - p_1$ = proportion of planned purchases by credit-card users
$1 - p_2$ = proportion of planned purchases by non-credit-card users

so MINITAB actually calculates for: $\big[(1 - p_1) - (1 - p_2)\big] = (p_2 - p_1)$:

$$-0.0879 < p_2 - p_1 < 0.2810$$

Multiply all three sides by (-1):

$$-0.2810 < p_1 - p_2 < 0.0879$$

Interpret:

Credit card users make between 28.10% fewer and 8.79% more impulse buys than non-credit-card users.

Or alternatively:

Non-credit-card users make between 8.79% fewer and 28.10% more impulse buys than credit-card users.