AP Statistics
Unit 07 – HW #2
Confidence Intervals: 2-Sample Proportions

Name Key
Period

1. Do younger people use Twitter more often than older people? In a random sample of 316 adult Internet users aged 18-29, 26% used Twitter. In a separate random sample of 532 adult Internet users aged 30-49, 14% used Twitter.

a. Calculate the standard error of the sampling distribution of the difference in the sample proportions (younger adults – older adults). What information does this value provide?

$$SE = \sqrt{\frac{(.26)(.74)}{316} + \frac{(.14)(.86)}{532}} = \boxed{0.02890}$$

b. Construct and interpret a 90% confidence interval for the difference between the true proportions of adult Internet users in these age groups who use Twitter.

State: p_1 = true proportion of young adults who use twitter.
p_2 = true proportion of old adults who use twitter.
$\hat{p}_1 = 0.26$ we want to estimate the true difference
$\hat{p}_2 = 0.14$ in proportion of $p_1 - p_2$ with 90% confidence.

Plan: Random: the data come from 2 independent random samples ✓
10% Condition: 3160 < all young adult internet users ✓
$n_1 = 316$
$n_2 = 532$ 5320 < all old adult internet users ✓
Large Counts: $n_1\hat{p}_1 = 82 \geq 10$ ✓ Because our conditions
$n_1\hat{q}_1 = 233 \geq 10$ ✓ are met, we will use
$n_2\hat{p}_2 = 74 \geq 10$ ✓ a 2-sample z-interval
$n_2\hat{q}_2 = 457 \geq 10$ ✓ for difference in 2 proportions $p_1 - p_2$.

DO: $(0.26 - 0.14) \pm 1.645\sqrt{\frac{(0.26)(1-0.26)}{316} + \frac{(0.14)(1-0.14)}{532}}$

$z^* = 1.645$

$= (0.0729, 0.1679)$

Conclude: we are 90% confident that the interval from 0.0729 to 0.1679 captures the true difference in proportion of adult internet users in these age groups who use Twitter $(p_1 - p_2)$.

2. A surprising number of young adults (ages 19-25) still live in their parents' homes. A random sample by the National Institutes of Health included 2253 men and 2629 women in this age group. The survey found that 986 of the men and 923 of the women lived with their parents.

a. Construct and interpret a 99% confidence interval for the difference in the true proportions of men and women aged 19-25 who live in their parents' homes.

State: p_1 = true proportion of men who live with parents (ages 19-25)
p_2 = true proportion of women who live with parents
$\hat{p}_1 = 986/2253 = 0.4376$
$\hat{p}_2 = 923/2629 = 0.3511$

we want to estimate the true difference in proportion $p_1 - p_2$ with 99% confidence

Plan: Random: the data comes from 2 independent random samples ✓
10% condition: 22530 < all men 19-25 years old ✓
26290 < all women 19-25 years old
$n_1 = 2253$
$n_2 = 2629$
Large Counts: $n_1\hat{p}_1 = 986$
$n_1\hat{q}_1 = 1267$
$n_2\hat{p}_2 = 923$
$n_2\hat{q}_2 = 1706$

Because our conditions are met, we will use a 2-sample z interval for difference of 2 proportions $p_1 - p_2$.

DO: $$(0.4376 - 0.3511) \pm 2.576\sqrt{\frac{(0.4376)(1-0.4376)}{2253} + \frac{(0.3511)(1-0.3511)}{2629}}$$

$z^* = 2.576$

$= (0.0505, 0.1226)$

Conclude. We are 99% confident that the interval from 0.0505 to 0.1226 captures the true difference in proportion of men and women aged 19-25 who live in their parents' homes $(p_1 - p_2)$.

b. Does your interval from part (a) give convincing evidence of a difference between the population proportions? Explain.

Yes, because 0 is not in the interval and zero represents no difference. We believe there is a difference (specifically in the direction where ($p_1 - p_2 > 0$ so $p_1 > p_2$) there is a higher proportion of men living w/ their parents.

3. The elderly fear crime more than younger people, even though they are less likely to be victims of crime. One study recruited separate random samples of 56 black women and 63 black men over the age of 65 from Atlantic City, New Jersey. Of the women, 27 said they "felt vulnerable" to crime; 46 of the men said this.

   a. Construct and interpret a 90% confidence interval for the difference in the true proportions of black women and black men in Atlantic City who would say they felt vulnerable to crime.

State: p_1 = true proportion of black men 65+ who feel vulnerable to crime.
p_2 = true proportion of black women 65+ who feel vulnerable to crime.
$\hat{p}_1 = 46/63 = 0.7302$ we want to estimate the true difference in proportion
$\hat{p}_2 = 27/56 = 0.4821$ $p_1 - p_2$ with 90% confidence.

Plan: Random: the data come from 2 independent random samples
10% condition: 56 < all black women in Atlantic City, NJ (65+) ✓
$n_1 = 63$ 630 < all black men in Atlantic City, NJ (65+) ✓
$n_2 = 56$
Large counts: $n\hat{p}_1 = 46 \geq 10$ ✓
$n\hat{q}_1 = 17 \geq 10$ ✓
$n\hat{p}_2 = 27 \geq 10$ ✓
$n\hat{q}_2 = 29 \geq 10$ ✓
Because our conditions are met, we will use a 2-sample z-interval for difference in 2 proportions $p_1 - p_2$.

DO: $(0.7302 - 0.4821) \pm 1.645\sqrt{\frac{(0.7302)(1-0.7302)}{63} + \frac{(0.4821)(1-0.4821)}{56}}$
$z^* = 1.645$
$= (0.1048, 0.3913)$

Conclude: We are 90% confident that the interval from 0.1048 to 0.3913 captures the true difference in proportion of black men and women 65+ years old in Atlantic City, NJ who say they felt vulnerable to crime.

   b. Does your interval from part (a) give convincing evidence of a difference between the population proportions? Explain. Because 0 is not in the interval, we believe there is a difference. ($p_1 - p_2 > 0$ so $p_1 > p_2$) we believe a higher proportion of black men 65+ in Atlantic City feel vulnerable to crime.

4. Are teens or adults more likely to go to McDonalds weekly? The Pew Internet and American Life Project asked a random sample of 799 teens and a separate random sample of 2253 adults how often they go to McDonalds. In these two surveys, 63% of teens and 68% of adults said that they go to McDonalds weekly. Construct and interpret a 90% confidence interval for the difference between adults and teens.

State: p_1 = true proportion of adults who go to McDonald's weekly.
p_2 = the proportion of teens who go to McDonald's weekly.
$\hat{p}_1 = 0.68$ we want to estimate the true difference
$\hat{p}_2 = 0.63$ in proportion $p_1 - p_2$ with 90% confidence.

Plan: Random: the data come from 2 independent random samples. ✓
10% condition: 7990 < all teens ✓
$n_1 = 799$ 22530 < all adults ✓
$n_2 = 2253$
Large counts: $n\hat{p}_1 = 1532 \geq 10$ ✓
$n\hat{q}_1 = 720 \geq 10$ ✓
$n_2\hat{p}_2 = 503 \geq 10$ ✓
$n\hat{q}_2 = 295 \geq 10$ ✓

Because our conditions are met, we will use a 2-sample z-interval for difference of 2 proportions $p_1 - p_2$.

Do: $$(0.68 - 0.63) \pm 1.645\sqrt{\frac{(0.68)(1-0.68)}{799} + \frac{(0.63)(1-0.63)}{2253}}$$

$z^* = 1.645$
(0.01788, 0.08167)

Conclude: we are 90% confident that the interval from 0.018 to 0.0829 captures the true difference in proportion between adults and teens who visit McDonald's weekly.