AP Stats Ch 22 - Required Practice Name: Solutions Per: ____

19. Twins. In 2001, one county reported that among 3132 white women who had babies, 94 were multiple births. There were also 20 multiple births to 606 black women. Does this indicate any racial difference in the likelihood of multiple births?

a) Test an appropriate hypothesis and state your conclusion.

b) If your conclusion is incorrect, which type of error did you commit?

a) H_0: $p_B = p_W$ multiple births for black and white women are the same.

H_A: $p_B \neq p_W$ multiple births for black and white women are different.

Conditions

✓ – $n_W p_W = 94$
$n_W q_W = 3132 - 94 = 3038$ ≥10
$n_B p_B = 20$
$n_B q_B = 606 - 20 = 586$

✓ – SRSs (assume representative)

✓ – groups indep (assumed)

✓ – 3132 < 10% of all white births
606 < 10% of all black births

with $\alpha = .05$, p-value = .7042 is high so we fail to reject H_0. We do not have SSE to conclude that multiple births for black and white women are different.

b)

	H_0 T	H_0 F
H_0 R	I	
H_0 NR		II

we failed to reject, so if this is in error, it would be a type II error.

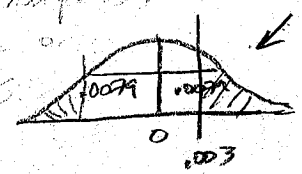
by hand (non-pooled) one

don't do this

$\hat{p}_B = \frac{20}{606} = .033$ $\hat{p}_W = \frac{94}{3132} = .030$

statistic $\hat{p}_B - \hat{p}_W = .033 - .030 = .003$

$SE_{\hat{p}_B - \hat{p}_W} = \sqrt{\frac{(.033)(.967)}{606} + \frac{(.030)(.97)}{3132}} = .0079$

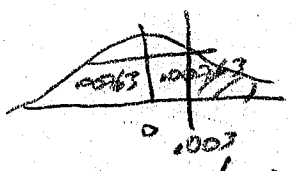


normalcdf(.003, 999, 0, .0079)
= .3521
p-value = 2(.3521) = .7042

by hand (pooled)

$\hat{p}_{pooled} = \frac{x_B + x_W}{n_B + n_W} = \frac{20 + 94}{606 + 3132} = \frac{114}{3738}$
= .0305

$SE_{pooled} = \sqrt{\frac{(.0305)(.9695)}{606} + \frac{(.0305)(.9695)}{3132}}$
= .00763

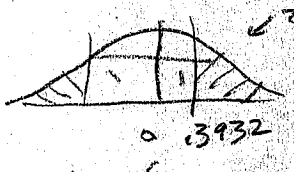


normalcdf(.003, 999, 0, .00763)
= .3470915 X2
p-val = .6942

using the "test-statistic"

$z = \frac{(\hat{p}_B - \hat{p}_W) - p_0}{SE_{\hat{p}_B - \hat{p}_W}}$

$z = \frac{.003 - 0}{.00763} = .3932$



normalcdf(., 999, 0, 1)
= .3470786
p-value = 2 · = .6942

by calculator

Perform 2-PropZTest
Using: x1: 20
n1: 606
x2: 94
n2: 3132
p1 ≠ p2

z = .5202
p-value = .6029

8. **Race and smoking.** In 1995, 24.8% of 550 white adults surveyed reported that they smoked cigarettes, while 25.7% of the 550 black adults surveyed were smokers.
   a) Create a 90% confidence interval for the difference in the percentages of smokers among black and white American adults.
   b) Does this survey indicate a race-based difference in smoking among American adults? Explain, using your confidence interval to test an appropriate hypothesis.

a) conditions

SRS ✓ (assumed)

nC10%pop ✓ 550+550 < 10% of all adults

Success/fail ✓
(550)(.248) = 136
(550)(.257) = 141
(550)(.752) = 413
(550)(.743) = 409

Indep obs ✓ (assumed)

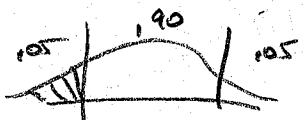
manually (unpooled)

$\text{statistic} = \hat{p}_B - \hat{p}_W = .257 - .248 = .009$

$SE_{\hat{p}_B - \hat{p}_W} = \sqrt{\frac{(.257)(.743)}{550} + \frac{(.248)(.752)}{550}}$

$= .026$

for 90% CI:



$z^* = \text{invNorm}(.05, 0, 1) = -1.64$

$CI = (\hat{p}_B - \hat{p}_W) \pm z^* SE_{\hat{p}_B - \hat{p}_W}$

$= .009 \pm 1.64(.026)$

$= .009 \pm .0426$

$= (-.0336, .0516)$

calculator

2PropZInt
x1: 141
n1: 550
x2: 136
n2: 550
Clevel: .90
(-.034, .05214)

b) No Since 0 is within the CI there is not sufficient evidence to conclude there is a difference in smoking rates between black and white smokers.

16. Birthweight. In 2003 the *Journal of the American Medical Association* reported a study examining the possible impact of air pollution caused by the 9/11 attack on New York's World Trade Center on the weight of babies. Researchers found that 8% of 182 babies born to mothers who were exposed to heavy doses of soot and ash on September 11 were classified as having low birthweight. Only 4% of 2300 babies born in another New York City hospital whose mothers had not been near the site of the disaster were similarly classified. Does this indicate a possibility that air pollution might be linked to a significantly higher proportion of low weight babies?

a) Test an appropriate hypothesis and state your conclusion.

b) If you concluded there is a difference, estimate that difference with a confidence interval and interpret that interval in context.

$\hat{p}_{pool} = \frac{15+92}{182+2300} = \frac{107}{2482} = .0431$

a) H_0: $p_E = p_{NE}$ There is no difference in % low birthweights.

H_A: $p_E > p_{NE}$ Exposed babies have higher % low birthweights.

Conditions

✓ $n_E p_E = (182)(.08) = 14.56$

$n_E q_E = (182)(.92) = 167.44$ ≥ 10

$n_{NE} p_{NE} = (2300)(.04) = 92$

$n_{NE} q_{NE} = (2300)(.96) = 2208$

✓ SRSs (assume samples are representative)

✓ groups indep

✓ 2482 < 10% of all babies

With $\alpha = .05$, p-value = .0255 is low so we reject H_0.

We do have sufficient statistical evidence to conclude that air pollution might be linked to a significantly higher proportion of low weight babies.

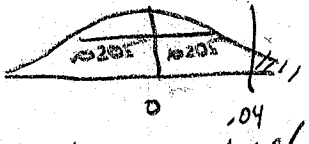
by hand (non-pooled)

$\hat{p}_E = .08$ $\hat{p}_{NE} = .04$

$\hat{p}_E - \hat{p}_{NE} = .08 - .04 = .04$

$SE_{\hat{p}_E - \hat{p}_{NE}} = \sqrt{\frac{(.08)(.92)}{182} + \frac{(.04)(.96)}{2300}}$

$= .0205$



p-value = normalcdf(.04, 999, 0, .0205)

$= .0255$

by calculator

Perform a 2PropZTest

using: $x_1 = 15$ (must be integer)

$n_1 = 182$

$x_2 = 92$

$n_2 = 2208$ 2300

$p_1 > p_2$

$z = 2.7177$

p-value = .00329

(Calc. does pooled)

b) (conditions already checked) by hand

z^* for 95%

$z^* = \text{invNorm}(.025, 0, 1)$

$= \pm 1.96$

$CI = (\hat{p}_E - \hat{p}_{NE}) \pm z^* SE_{\hat{p}_E - \hat{p}_{NE}}$

$= .04 \pm (1.96)(.0205)$

$= (-.00018, .0802)$

We are 95% confident that between .018% fewer and 8.02% more exposed babies have low birthweight.

by calculator

Perform a 2PropZInt

using $x_1 = 15$

$n_1 = 182$ (.00167, .08346)

$x_2 = 92$ (-6E-5, .08156)

$n_2 = 2208$ -.00006, .0850)

C-level = .95

We are 95% confident that between 0.1% and 8.3% more exposed babies have low birthweight.

4. Graduation. In October 2000 the U.S. Department of Commerce reported the results of a large-scale survey on high school graduation. Researchers contacted more than 25,000 Americans aged 24 years to see if they had finished high school; 84.9% of the 12,460 males and 88.1% of the 12,678 females indicated that they had high school diplomas.

a) Are the assumptions and conditions necessary for inference satisfied? Explain.

b) Create a 95% confidence interval for the difference in graduation rates between males and females.

c) Interpret your confidence interval.

d) Does this provide strong evidence that girls are more likely than boys to complete high school? Explain.

a) Conditions

SRS ✓ Indep ✓ nc≤10% ✓ success/fail ✓

assumed likely 12460 ≤ 10% of male (12460)(.849) = 105

12678 ≤ 10% of female (12460)(.151) = 182

(12678)(.881) = 111

(12678)(.119) = 15

Statistic: $\hat{p}_F - \hat{p}_M = .881 - .849 = .032$

manually, non-pooled

$SE_{\hat{p}_F - \hat{p}_M} = \sqrt{\frac{(.849)(.151)}{12460} + \frac{(.881)(.119)}{12678}}$

$= .0043$

for 95% CI, $z^* = 1.96$

$CI = (\hat{p}_F - \hat{p}_M) \pm z^* SE_{\hat{p}_F - \hat{p}_M}$

$= .032 \pm (1.96)(.0043)$

$= .032 \pm .0084$

(.0236, .0404)

w/ calculator

perform 2-PropZInt

using $x_1 = 10579$

$n_1 = 12460$

$x_2 = 11169$

$n_2 = 12678$

C-level = .95

(−.0404, −.0235) ← these are negative b/c I put lower proportion as x1/n1 (arbitrary)

* which is equivalent to...

(.0235, .0404)

c) We are 95% confident that the true difference in graduation rates (F−M) is between 2.4% and 4.0%.

—or—

We are 95% confident that between 2.4% and 4.0% more females graduate than males.

c) Yes, because 0 is not within the interval.

(all of the likely values show a positive difference)

Chapter 22 Practice Quiz

AP Statistics Quiz A – Chapter 22 ***Name*** _____

Great Britain has a great literary tradition that spans centuries. One might assume, then, that Britons read more than citizens of other countries. Some Canadians, however, feel that a higher percentage of Canadians than Britons read. A recent Gallup Poll reported that 86% of 1004 randomly sampled Canadians read at least one book in the past year, compared to 81% of 1009 randomly sampled Britons. Do these results confirm a higher reading rate in Canada?

1. Test an appropriate hypothesis and state your conclusions.

H_0: $p_C = p_B$ percentage of Canadians and Britons who've read a book is the same,

H_A: $p_C > p_B$ percentage of Canadians who've read a book is higher than Britons.

Conditions

✓ $n_C p_C = (1004)(.86) = 863.44$

$n_C q_C = (1004)(.14) = 140.56 \geq 10$

$n_B p_B = (1009)(.81) = 817.29$

$n_B q_B = (1009)(.19) = 191.71$

✓ – SRS, "randomly sampled"

– 2 pop'l indep (different countries)

✓ – 1004 < 10% Canadians

1009 < 10% of Britons

Perform a 2PropZTest in TI84

using: $x1 = 863$, $n1 = 1004$, $x2 = 817$, $n2 = 1009$, $p_1 > p_2$

$z = 3.0096$

$p\text{-value} = .0013$

With $\alpha = .05$, p-value = .0013 is low so we reject H_0.

We do have sufficient statistical evidence to conclude that the reading rate in Canada is higher than Britain.

2. Find a 99% confidence interval for the difference in the proportion of Britons and Canadians who read at least one book in the last year. Interpret your interval.

Perform a 2PropZInt in TI84

using $x1 = 863$, $n1 = 1004$, $x2 = 817$, $n2 = 1009$, C-level = .99

(.00729, .0922)

We are 99% confident that between 0.73% and 9.2% more Canadians read at least 1 book compared to Britons.

We are 99% confident that between 0.73% and 9.2% more Canadians read at least 1 book compared to Britons.