AP Stats Ch 20 - Required Practice **Name:** Solutions **Per:** ____

#1. **Congenital abnormalities.** In the 1980s it was generally believed that congenital abnormalities affected about 5% of the nation's children. Some people believe that the increase in the number of chemicals in the environment has lead to an increase in the incidence of abnormalities. A recent study examined 384 children and found that 46 of them showed signs of an abnormality. Is this evidence that the risk has increased?

a) State the hypotheses:

H_0: $p = .05$ The percentage of children with abnormalities is 5%.

H_A: $p > .05$ The percentage of children with abnormalities has increased beyond 5%.

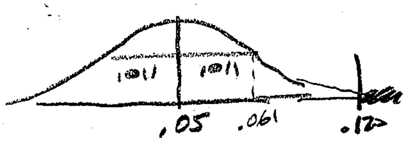
b) Check conditions

- ✓ SRS (assuming representative sample)
- ✓ indep. (assumed)
- ✓ $n < 10\%$ pop (384 < 10% of all children)
- ✓ $np = (384)(.05) = 19.2$; $nq = (384)(.95) = 364.8$ ≥10

c) Conduct the test (find the p-value). Do this by hand first, and then check your answer by using an appropriate hypothesis test in your calculator:

by hand

$\hat{p} = \frac{46}{384} = .120$ $\quad \sigma_{\hat{p}} = \sqrt{\frac{(.05)(.95)}{384}} = .011$



p-value $= P(\hat{p} > .120)$

$=$ normalcdf$(.120, 999, .05, .011)$

$= 9.896 \times 10^{-11} \approx 0$

by calculator

Perform a 1 PropZTest in a TI-84 using:

p_0: .05

x: 46

$n = 384$

prop > p_0

$z = 6.275$

p-value $= 1.754 \times 10^{-10} \approx 0$

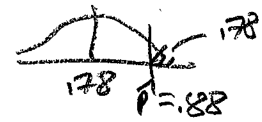
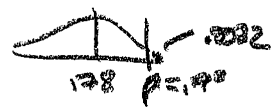
d) Write your conclusion paragraph.

With a significance level of .05, p-value ≈ 0 is low, so we reject H_0. We do have sufficient statistical evidence to conclude that the risk of abnormalities in children has increased.

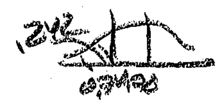
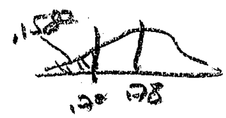
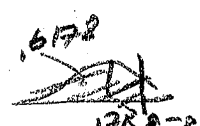
#2. **DV Seniors (group investigation)**

a) In your group, each of you separately use your phones' internet browsers to browse to www.mrfelling.com/sa3 .
Select 'One-sided, upper (>)' and click 'sample'.

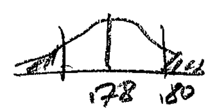
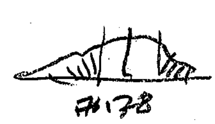
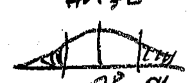
Read the whole page on your phone, and press 'sample' at least 3 times. Each time write down the following:

Null hypothesis value p_0	sample's $\hat{p}$	p-value	sketch the distribution and shading
$p_0 = 0.78$	$\hat{p} = 0.88$	0.0446	
$p_0 = 0.78$	$\hat{p} = 0.92$	0.0082	
$p_0 = 0.78$	$\hat{p} = 0.8$	0.3821	

b) Now click 'reset', and select 'One-sided, lower (<)' and click 'sample' a few times. Record the same things for these trials:

Null hypothesis value p_0	sample's $\hat{p}$	p-value	sketch the distribution and shading
$p_0 = 0.78$	$\hat{p} = .71$	0.242	
$p_0 = 0.78$	$\hat{p} = .72$	0.1582	
$p_0 = .78$	$\hat{p} = .8$	.6178	

c) Click 'reset', and select 'Two-sided, ($\neq$)' and click 'sample' a few times. Record the same things for these trials:

Null hypothesis value p_0	sample's $\hat{p}$	p-value	sketch the distribution and shading
$p_0 = .78$	$\hat{p} = .86$	.1616	
$p_0 = .78$	$\hat{p} = .76$	.7642	
$p_0 = .78$	$\hat{p} = .84$	.3174	

d) Compare with others in your group. Write about what you notice is changing and staying the same in these cases: p_0 stays constant, but $\hat{p}$ varies for each sample. The farther $\hat{p}$ is away from p_0 the smaller the p-value is. One-sided shading matches the <, > of H_A, but two-sided shading is always shaded away from center, then doubled to find p-value.

#3. **Frogs** A certain species of poison-dart from always has a body with two yellow stripes on its back, but a percentage of frogs also have a blue stripe between the two yellow stripes. The blue stripe is historically found in 34% of the frogs. Industrial pollution has been found in rivers containing these frogs and researchers wonder if the pollution is causing a change in the percentage of frogs with the blue stripe. A random sample of 60 frogs was collected and 29 had the blue stripe. Does this data provide evidence that the proportion of frogs with the blue stripe has changed?

a) Conduct a hypothesis test (by hand) to answer the question (use p=0.34 when computing $\sigma_{\hat{p}}$):

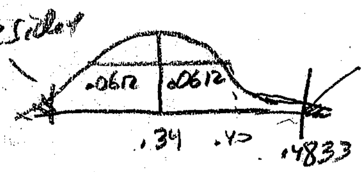
$H_0: p = .34$ The proportion of frogs w/blue stripe is 34%.

$H_a: p \neq .34$ The proportion of frogs w/blue stripe is no longer 34%.

$\hat{p} = \frac{29}{60} = .4833$ $\quad \sigma_{\hat{p}} = \sqrt{\frac{(.34)(.66)}{60}} = .0612$

Conditions
- ✓ $np = (60)(.34) = 20.4 \geq 10$
- ✓ $nq = (60)(.66) = 39.6 \geq 10$
- ✓ SRS "random"
- ✓ indep (implied in SRS)
- 60 < 10% of all frogs

2 sided



normalcdf(.4833, 999, .34, .0612)
= .0096

p-value = 2(.0096) = .0192

With significance level of .05, p-value = .0192 is low so we reject H_0.
We do have sufficient statistical evidence to conclude that the proportion of frogs with blue stripe is no longer 34%.

b) Your analysis should have concluded that the percentage of frogs with the blue stripe has changed…but to what percentage? To establish a percentage for the population, now compute a 95% confidence interval around your statistic. (use $\hat{p}$ when computing $SE_{\hat{p}}$):

(conditions checked above)

$CI = \hat{p} \pm z^* \sqrt{\frac{\hat{p}\hat{q}}{n}}$

$= .4833 \pm (1.96)\sqrt{\frac{(.4833)(.5167)}{60}}$

$= .4833 \pm .126447\ldots$

$= (.3569, .6097)$

z^* for 95%:

$z^* = \text{invNorm}(.025, 0, 1)$
$= -1.96$
(ignore $-$)

We are 95% confident that the true percentage of all frogs with blue stripe is between 35.69% and 60.97%.

c) Is the null hypothesis value, 0.34, in your confidence interval?

No (which makes sense… we are saying it changed, so it shouldn't be one of the likely values)

2. More hypotheses. Write the null and alternative hypotheses you would use to test each of the following situations.

a) In the 1950s only about 40% of high school graduates went on to college. Has the percentage changed?

b) 20% of cars of a certain model have needed costly transmission work after being driven between 50,000 and 100,000 miles. The manufacturer hopes that redesign of a transmission component has solved this problem.

c) We field test a new flavor soft drink, planning to market it only if we are sure that over 60% of the people like the flavor.

a) H_0: $p = .40$ The percentage of HS going to college has not changed.
H_A: $p \neq .40$ The percentage of HS going to college has changed,

b) H_0: $p = .20$ The proportion of cars needing transmission repair is still 20%.
H_A: $p < .20$ the proportion of cars needing transmission repair has decreased.

c) H_0: $p = .60$ The percentage of people who like the new flavor is 60%
H_A: $p > .60$ the percentage of people who like the new flavor is greater than 60%.

6. Cars. A survey investigating whether the proportion of today's high school seniors who own their own cars is higher than it was a decade ago finds a P-value of 0.017. Is it reasonable to conclude that more high schoolers have cars? Explain.

Yes, because there is only a 1.7% chance of the observed higher proportion of seniors who own their own cars occurring randomly from natural sampling variation.

19. Twins. In 2001 a national vital statistics report indicated that about 3% of all births produced twins. Is the rate of twin births the same among very young mothers? Data from a large city hospital found only 7 sets of twins were born to 469 teenage girls. Test an appropriate hypothesis and state your conclusion. Be sure the appropriate assumptions and conditions are satisfied before you proceed.

$H_0: p = .03$ Twin birth rate is 3%.

$H_A: p \neq .03$ Twin birth rate is not 3%.

Conditions

✓ – $np = (469)(.03) = 14$ ≥ 10
$nq = (469)(.97) = 455$

✓ – SRS (assume this sample is representative)

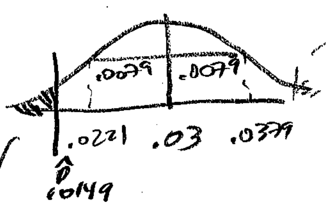
✓ – independent (assumed)

✓ – 469 < 10% of all teenage mothers (unfortunately, true)

by hand

$\hat{p} = \frac{7}{469} = .0149$

$\sigma_{\hat{p}} = \sqrt{\frac{(.03)(.97)}{469}} = .0079$



2 sided

so p-value = 2(.028)

$= \boxed{.056}$

normalcdf(−999, .0149, .03, .0079)

= .028

by calculator

Perform a 1PropZtest in TI 84

using: $p_0 = .03$

$x = 7$

$n = 469$

$prop \neq p_0$

$z = -1.9138$

p-value = $\boxed{.0557}$

With significance level .05, p-value = .056 is high, so we fail to reject H_0. We do not have sufficient statistical evidence to conclude that the rate of twin births is different than 3% for teenage mothers.

16. Satisfaction. A company hopes to improve customer satisfaction, setting as a goal no more than 5% negative comments. A random survey of 350 customers found only 10 with complaints.

a) Create a 95% confidence interval for the true level of dissatisfaction among customers.

b) Does this provide evidence that the company has reached its goal? Using your confidence interval, test an appropriate hypothesis and state your conclusion.

a) check conditions (confidence intervals are inference analyses which require conditions)

✓ — $np = (350)(.05) = 17.5$
$nq = (350)(.95) = 332.5$

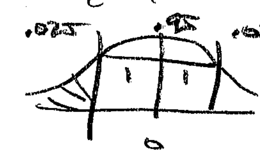
✓ — SRS (assume the sample is representative)

✓ — indep (assume the customers are indep)

✓ — 350 < 10% of all customers

$\hat{p} = \frac{10}{350} = .0286$ $\qquad SE_{\hat{p}} = \sqrt{\frac{(.0286)(.9714)}{350}} = .0089$

z^* for 95%



$CI = \hat{p} \pm z^* SE_{\hat{p}}$

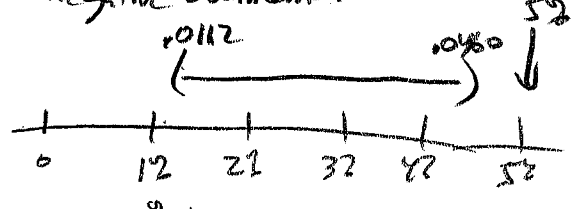
$z^* = \text{invNorm}(.025, 0, 1) = \pm 1.96$

$= .0286 \pm (1.96)(.0089)$

$= (.0112, .0460)$

we are 95% confident that between 1.12% and 4.6% of customers would leave negative comments.

b)



$H_0: p = .05$ 5% of customers leave negative comments

$H_A: p < .05$ Less than 5% of customers leave negative comments.

5% negative comments is outside the 95% confidence interval of likely values, so we can reject H_0. This confidence interval suggests that the company has reached its goal of having less than 5% negative comments.

24. Jury. Census data for a certain county shows that 19% of the adult residents are Hispanic. Suppose 72 people are called for jury duty, and only 9 of them are Hispanic. Does this apparent underrepresentation of Hispanics call into question the fairness of the jury selection system? Explain.

H_0: $p = .19$ percentage of jurors who are Hispanic matches population 19%.

H_A: $p < .19$ percentage of jurors who are Hispanic is less than the population 19%.

conditions

✓ – $np = (72)(.19) = 13.7$
$nq = (72)(.81) = 58.3$ ≥ 10

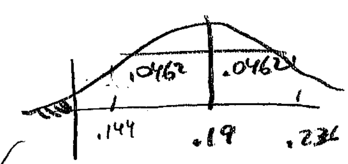
✓ – SRS (we are assuming these 72 people are representative of a typical jury)

✓ – indp (assume selection process means jurors are indp!)

✓ – 72 < 10% of all jury candidates

by hand

$\hat{p} = \frac{9}{72} = .125$ $\sigma_{\hat{p}} = \sqrt{\frac{(.19)(.81)}{72}} = .0462$



$\hat{p} = .125$

p-value = normalcdf(-999, .125, .19, .0462)
= .0797

by calculator

Perform a 1propZTest in Ti 84:
using: $p_0 = .19$
$x = 9$
$n = 72$
prop < p_0

$z = -1.4059$
p-value = .0799

With significance level .05, p-value = .0797 is high, so we fail to reject H_0.
We do not have sufficient statistical evidence to conclude that the percentage of jurors who are Hispanic is less than the general population's percentage Hispanic of 19%.

Chapter 20 Practice Quiz

AP Statistics Quiz B – Chapter 20 **Name**________________________

The International Olympic Committee states that the female participation in the 2004 Summer Olympic Games was 42%, even with new sports such as weight lifting, hammer throw, and modern pentathlon being added to the Games. Broadcasting and clothing companies want to change their advertising and marketing strategies if the female participation increases at the next games. An independent sports expert arranged for a random sample of pre-Olympic exhibitions. The sports expert reported that 202 of 454 athletes in the random sample were women. Is this strong evidence that the participation rate may increase?

1. Test an appropriate hypothesis and state your conclusion.

H_0: $p = .42$ Percentage of athletes who are women is 42%.

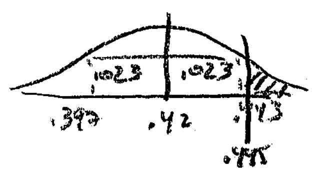
H_a $p > .42$ Percentage of athletes who are women is greater than 42%.

Conditions

- $np = (454)(.42) = 190.7$
- $nq = (454)(.58) = 263.3$ ≥ 10
- SRS "random sample"
- indep (implied in SRS)
- 454 < 10% of all Olympic athletes

by hand (do by hand on quiz/test)

$\hat{p} = \frac{202}{454} = .445$ $\sigma_{\hat{p}} = \sqrt{\frac{(.42)(.58)}{454}} = .023$



p-value = normalcdf(.445, 99, .42, .023)
= .139

with significance level of .05, p-value = .139 is high so we fail to reject H_0. We do not have sufficient statistical evidence to conclude that the percentage of athletes who are women has increased above 42%.

2. Was your test one-tail upper tail, lower tail, or two-tail? Explain why you choose that kind of test in this situation.

One-tail, upper-tail because of wording "evidence that participation rate may increase"

3. Explain what your *P*-value means in this context.

If the percentage of athletes who are women is really 42%, there is a 13.9% chance that a sample of 454 athletes will have percentage of women as high as in this study (44.5%) or higher, just due to random chance.