AP Stats Ch 21 - Required Practice **Name:** Solutions **Per:** ____

#1. **Breast Cancer.** A course of treatment for breast cancer has a 5-year survival rate of 0.86. A new treatment has been tested on a sample of 50 volunteers and it was found that the 5-year survival rate for that sample was 0.94.

a) At the 0.05 significance level, do we have evidence that this new treatment is better than the old treatment?

H_0: $p = .86$ New treatment same effectiveness as old treatment.

H_A: $p > .86$ New treatment more effective than old treatment.

conditions

X – $np = (50)(.86) = 43$ not ≥ 10

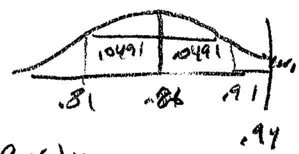
$nq = (50)(.14) = 7$ ←

~ – SRS (assume representative?)

✓ – indep (assumed)

✓ – 50 < 10% of all breast cancer patients

$\hat{p} = .94$ $\sigma_{\hat{p}} = \sqrt{\frac{(.86)(.14)}{50}} = .0491$



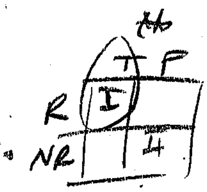
p-value = normalcdf(.94, 999, .86, .0491)

$= \boxed{.0516}$

With $\alpha = .05$, p-value = .0516 is high so we fail to reject H_0. We do not have sufficient statistical evidence to conclude that the new treatment is better than the old treatment.

b) What would be a type I error?

If the new treatment is the same effectiveness as the old treatment, but we conclude it is better.

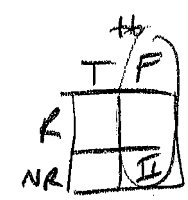


c) What is the probability of making a type I error?

$P(\text{type I}) = \alpha = \boxed{.05}$

d) What would be a type II error?

If the new treatment is actually better but we conclude it isn't better than the old treatment.



e) Which error is more serious? Type I or Type II? Explain.

Type I consequence: New treatment is not better but is used, possibly more expense or less safe than old treatment.

Type II consequence: new treatment is actually better, but we don't use it — result is some people die on old treatment that could have been saved.

I would say the <u>type II</u> error is more serious.

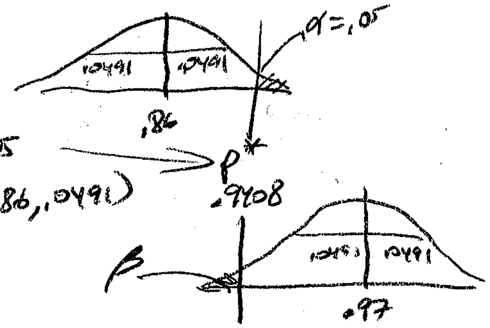
f) In Europe, this new treatment has been used far longer and they have found a 5-year success rate of 0.97. What is the probability of making a type II error? What is the power of the test?

$p_{actual} = 0.97$

① Calculate p^* for $\alpha = .05$

invNorm(.95, .86, .0491)

$= .9408$



② Calculate β = normalcdf(−.999, .9408, .97, .0491)

$= .2760$

③ Calculate power $= 1 - \beta$

$= 1 - .2760$

$= \boxed{.724}$

g) The effect size is the difference between p_0 and the TRUE, ACTUAL population proportion (which we will most likely never know). What happens to the power of the test as the effect size increases?

As the effect size increases, it becomes easier and more likely to detect that effect as significant.
Because power is the probability of detecting an effect if the effect is present, the power will increase if effect size increases.

h) We have no control over the effect size. What could we control to decrease P(type II error) and hence increase the power of the test?

1) We could increase the sample size;

$n \uparrow \quad SE \downarrow \quad \alpha \downarrow \quad \beta \downarrow$ and power $= 1 - \beta \uparrow$

2) We could also increase α;

for a fixed n, $\alpha \uparrow \quad \beta \downarrow \quad$ power $= 1 - \beta \uparrow$

(but this would increase the probability of a type I error)

13. Testing cars. A clean air standard requires that vehicle exhaust emissions not exceed specified limits for various pollutants. Many states require that cars be tested annually to be sure they meet these standards. Suppose state regulators double check a random sample of cars that a suspect repair shop has certified as okay. They will revoke the shop's license if they find significant evidence that the shop is certifying vehicles that do not meet standards.

a) In this context, what is a Type I error?
b) In this context, what is a Type II error?
c) Which type of error would the shop's owner consider more serious?
d) Which type of error might environmentalists consider more serious?

H_0: Emissions within limits.
H_A: Emissions not within limits

	Ho T	Ho F
R	I	
NR		II

a) Emissions are actually within limits, but we happen to sample enough outside limits to not certify the shop.
b) Emissions are actually outside limits but we fail to detect this, and certify the shop.
c) Owner would consider Type I worse.
d) Environmentalists would consider Type II worse.

15. Cars again. As in Exercise 13, state regulators are checking up on repair shops to see if they are certifying vehicles that do not meet pollution standards.

a) In this context, what is meant by the power of the test the regulators are conducting?
b) Will the power be greater if they test 20 or 40 cars? Why?
c) Will the power be greater if they use a 5% or a 10% level of significance? Why?
d) Will the power be greater if the repair shop's inspectors are only a little out of compliance or a lot? Why?

a)

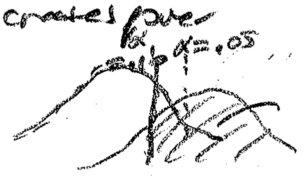
	T	F
R	I	$1-\beta$ power
N		II β

Power is the probability that, if the shop is not within limits (H_0 False) that this study will correctly detect this and keep the shop from getting certified.

b) Power will increase with 40 cars $n\uparrow$ $\alpha\downarrow$ and $\beta\downarrow$
$1-\beta=$ power $\nearrow$

power $\uparrow$ because $\sigma\downarrow$

c) $\alpha\uparrow$ $\beta\downarrow$ $1-\beta\uparrow$ higher α increases power
so higher with $\boxed{\alpha=.10}$



p^* moves left with higher α

d) Power will be greater if shop is "a lot" out of compliance because the probability of detecting this increases
(it is easier to detect a larger effect size)

20. **Ads.** A company is willing to renew its advertising contract with a local radio station only if the station can prove that more than 20% of the residents of the city have heard the ad and recognize the company's product. The radio station conducts a random phone survey of 400 people.

a) What are the hypotheses?

b) The station plans to conduct this test using a 10% level of significance, but the company wants the significance level lowered to 5%. Why?

c) What is meant by the power of this test?

d) For which level of significance will the power of this test be higher? Why?

e) They finally agree to use $\alpha = 0.05$, but the company proposes that the station call 600 people instead of the 400 initially proposed. Will that make the risk of Type II error higher or lower? Explain.

a) $H_0: p = .20$ 20% of residents heard ads
$H_A: p > .20$ more than 20% of residents heard ads

b) lowering $\alpha \downarrow$ decreases probability of type I error of reporting that more than 20% have heard ads when only 20% actually have.

c) power is the probability of correctly detecting that more than 20% have heard ads if more than 20% actually have heard ads.

d) power $= 1-\beta \uparrow$ if $\beta \downarrow$ so $\alpha \uparrow$
power will be higher for $\alpha = .10$

e) with increased n $\alpha \downarrow$ $\beta \downarrow$
and power $= 1-\beta \uparrow$

22. **Testing the ads.** The company in Exercise 20 contacts 600 people selected at random, and only 133 remember the ad.

a) Should the company renew the contract? Support your recommendation with an appropriate test.

b) Explain carefully what your P-value means in this context.

a) $\hat{p} = \frac{133}{600} = .222$

conditions
SRS/indep ✓ assumed
$n < 10\%$ pop ✓ 600 < 10% of listeners
$np, nq \geq 10$ ✓
$(600)(.2) = 120$
$(600)(.8) = 480$

$H_0: p = .20$ 20% of residents heard ads
$H_A: p > .20$ more than 20% of residents heard ads

1-PropZTest using: $p_0 = .20$, $x = 133$, $n = 600$, $p > p_0$ p-value $= \boxed{.0922}$

p-value $= .0922 > .05$ is high so we fail to reject H_0.
We do not have sufficient statistical evidence to conclude that more than 20% of residents heard the ads.

b) p-value of .0922 means if 20% of listeners actually heard the ads there is a 9.22% probability of a sample of 600 having 22.2% hearing the ads (like this sample) or higher just due to natural sampling variation.

Chapter 21 Practice Quiz

AP Statistics Quiz C – Chapter 21 **Name**____________________

The owner of a small clothing store is concerned that only 28% of people who enter her store actually buy something. A marketing salesman suggests that she invest in a new line of celebrity mannequins (think Adam Sandler modeling the latest jeans…). He loans her several different "people" to scatter around the store for a two-week trial period. The owner carefully counts how many shoppers enter the store and how many buy something so that at the end of the trial she can decide if she'll purchase the mannequins.

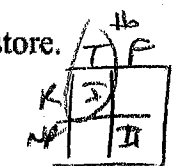
1. Write the owner's null and alternative hypotheses.

   Ho: $p = .28$ mannequins do not change purchases.

   Ha: $p > .28$ mannequins increase purchases.

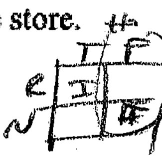
2. In this context describe a Type I error and the impact such an error would have on the store.

   Type I error means mannequins actually don't increase purchases but we think they do. Impact: we waste money on ineffective mannequins.



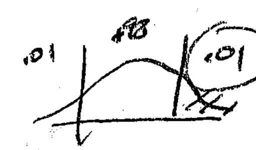
3. In this context describe a Type II error and the impact such an error would have on the store.

   Type II error means mannequins actually do increase sales, but we think they don't. Impact: we don't buy mannequins and miss out on the increased sales.



4. Based on data that she collected during the trial period the store's owner found that a 98% confidence interval for the proportion of all shoppers who might buy something was (27%, 35%). What conclusion should she reach about the mannequins? Explain.

   because 28% is within the confidence interval we don't have evidence that mannequins increase sales.



(HA is one-sided)

5. What alpha level did the store's owner use? .01

6. Describe to the owner an advantage and a disadvantage of using an alpha level of 5% instead.

   Advantage: Increasing α from .01 to .05 would decrease β, decrease the probability of a type II error (less chance of missing out on sales). It also increases the power of the test (more likely to detect mannequins help if they really do).

   Disadvantage: Probability of a type I error (wasting money on ineffective mannequins) is increased.

7. The owner talked the salesman into extending the trial period so that she can base her decision on data for a full month. Will the power of the test increase, decrease, or remain the same?

   This increases sample size:

   $n \uparrow \quad \alpha \downarrow \; \beta \downarrow \quad \text{power} = 1-\beta \uparrow$ power of the test will increase

8. Over the trial month the rate of in-store sales rose to 30% of shoppers. The store's owner decided this increase was statistically significant. Now that she's convinced the mannequins work, why might she still chose not to purchase them?

   28% to 30% is only a 2% increase in sales. This might not be enough of an increase to justify the cost of the mannequins. (difference between statistically significant and practically significant)