Let’s go back to Reilly’s claim that 75% of prom shoppers prefer to get their prom dress (shoes and all) on Revolve. Annabelle is convinced that the percentage is way less (she prefers Lulus). Annabelle does a google form survey and is able to get a random sample of 156 students shopping for a prom dress to give their first choice. Her results are that of the 156 responses, 107 of them swear by Revolve.

How would we know if Annabelle makes a faulty inference from her data?

First let’s review the process

Steps in Hypothesis Testing

1. Define the population characteristic (i.e. parameter) about which hypotheses are to be tested.

Proportion of Revolve users

2. State the null hypothesis H_0 Reilly's claim $\rightarrow p = 0.75$

3. State the alternative hypothesis H_a Annabelle's claim $\rightarrow p < 0.75$

4. State the significance level for the test α What value of α are we going with?

5. Check all assumptions. We've done this before

6. State the name of the test. Sample proportion *z*?

7. State *df* (degrees of freedom) if applicable (not applicable in proportion land).

8. Display the test statistic to be used without any computation at this point. Which formula?

9. Compute the value of the test statistic, showing specific numbers used. Show work on AP Exam

10. Calculate the *P* – value. Is *p* - value greater than or less than significance level? This determines the outcome, reject or fail to reject.

11. Sketch a picture of the situation.

12. State the conclusion in two sentences -
    1. Summarize in theory discussing H_0
    2. Summarize in context discussing H_a

Single Sample Hypothesis Tests for Proportions

Steps in Proportion Hypothesis Testing

1. $p =$ ……. Proportion of Revolve users
2. $H_0 : p = \#$ Reilly's claim → $p = 0.75$
3. $H_a : p$ $\neq$ / $<$ / $>$ Annabelle's claim → $p < 0.75$
4. State the significance level α for the test

5. <u>Assumptions:</u>
   1. Random Sample
   2. $np \geq 10$, $n(1-p) \geq 10$
   3. SSSRTP

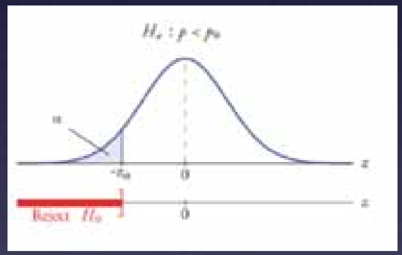
6. 1 Sample Proportion z Test

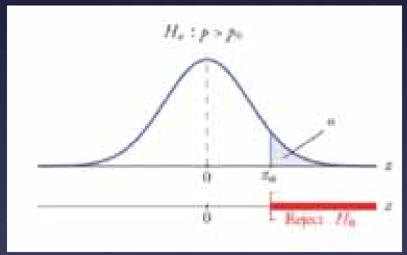
7. $df = N/A$

8/9. $$z = \frac{\hat{p} - p}{\sqrt{\frac{p(1-p)}{n}}} = \# \longleftarrow z = \frac{x - \mu}{\sigma}$$

Note the formula for z score using proportions

11.

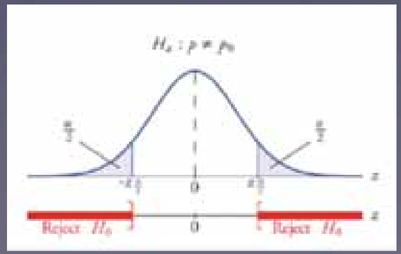




10. $P - value =$

$$P(z > \#) = normalcdf(\#, 1E99, 0, 1)$$
$$P(z < \#) = normalcdf(-1E99, \#, 0, 1)$$
} one-sided tests

$$2P(z > \#) = 2 * normalcdf(\#, 1E99, 0, 1)$$
$$2P(z < \#) = 2 * normalcdf(-1E99, \#, 0, 1)$$
} two-sided tests



Is p - value greater than or less than significance level? This determines the outcome, reject or fail to reject.

12. State the conclusion in two sentences -
    1. Summarize in theory discussing H_0.
    2. Summarize in context discussing H_a.

Errors - We make them, even though we're awesome

INCREDIBLY IMPORTANT: We do not "accept" the null hypothesis here. We "fail to reject" it which is not the same thing.

	Fail to reject	Reject
true	Hooray!	Type I error
true	Type II error	Hooray!

Analogous to a false positive test

Analogous to finding an innocent person guilty

Type I error - reject H_0 when H_0 is true

Type II error - fail to reject H_0 when H_0 is false

Analogous to a false negative test

Analogous to acquitting a guilty person

OR

Type I error - 1st equation correct and you pick the 2nd "equation"

Type II error - 2nd "equation" correct and you pick the 1st equation

α VS β

$$P(\text{Type I error}) = \alpha$$

$$P(\text{Type II error}) = \beta$$

So our level of significance α is also our probability of a Type I error.

If α goes up, then β goes down.

If α goes down, then β goes up.

Game plan - determine which error is worse, then choose the appropriate α and β.

Errors - We make them, even though we're awesome

	Fail to reject	Reject
True	Reilly is right and Annabelle's evidence says so!	Reilly is right but Annabelle's evidence says otherwise
False	Reilly only appears to be right according to Annabelle's evidence	Reilly is wrong and Annabelle's evidence says so!

What probability of being wrong (α) are we going with?

Consequences of each type of error

Type I error - Revolve doesn't gain as many customers because people believe Annabelle's data

Type II error - Revolve gains customers they shouldn't get because Annabelle's error inflates the perception of Revolve

Now let's go back to Raven and her claim of Mahomes being the GOAT. She claims that after some sampling that the proportion of Mahomes believers at SI is 0.92 (that's 92%!). Kevin is having none of this and decides to do his own sampling. Because he has March Madness on his mind he only takes the time for one sample of 140 SI students and finds that 125 are believers.

What are the consequences of each error?

Errors - We make them, even though we’re awesome

	Fail to reject	Reject
True	Raven is right and Kevin’s evidence says so	Kevin claims to be right but his evidence says otherwise
False	Raven only appears to be right according to Kevin’s evidence	Kevin is right and his own evidence says so

Type I error with probability = α

Type II error with probability = β

Type I error - Kevin gloats and Raven simmers but Kevin was wrong the whole time

Type II error - Kevin is sad but shouldn’t be because he’s right despite the evidence

Game plan - determine which error is worse, then choose the appropriate α and β.

So which would we rather have: Angry Kevin or angry Raven?

Errors - We make them, even though we're awesome

When we correctly reject a particular null hypothesis, we get to something called…

	Fail to reject H_0	Reject H_0
H_0 **true**	Hooray!	Type I error α
H_a **true**	Type II error β	POWER!!

POWER is the probability that we correctly reject the null hypothesis

In other words, POWER is $1-\beta$

Type I error - reject H_0 when H_0 is true

Type II error - fail to reject H_0 when H_0 is false

And in case you're wondering, Power is a good thing:)

OR

Type I error - 1st equation correct and you pick the 2nd equation

Type II error - 2nd equation correct and you pick the 1st equation

Also called 'level of significance' or 'significance level'.

α vs β

$P(\text{Type I error}) = \alpha$

$P(\text{Type II error}) = \beta$

If α goes up, then β goes down.

If α goes down, then β goes up.

You won't be asked to calculate power from scratch but you will be expected to understand what it represents and calculate it using β

$$\text{Power} = P(\text{rejecting a false } H_0) = 1 - \beta$$

Power is a good thing:)

4 ways for Power to increase

1. Increase α because β will go down
2. Increase Sample Size
3. Decrease the Standard Error size
4. The larger the discrepancy (distance) between the hypothesized parameter value and the true parameter value, the larger the power

Example 3 The U.S. Department of Transportation reported that during a recent period, 77% of all domestic passenger flights arrived on time (meaning within 15 minutes of the scheduled arrival). Suppose that an airline with a poor on-time record decides to offer its employees a bonus if, in an upcoming month, the airline's proportion of on-time flights exceeds the overall industry rate of 0.77. Let p be the true proportion of the airline's flights that are on time during the month of interest.

(a) What are the null and alternate hypotheses?

(b) What are the Type I and Type II errors in this context?

(c) What are the consequences of these errors?

If $\beta = 0.14$ What is the power of this test? Power $= 1 - \beta = 0.86$