Flip a coin 100 times.
H_0: p = 0.5 (coin in fair)
H_1: p ≠ 0.5 (coin is unfair)

Sample #1: You get 51 heads. Conclusion: retain the null.
You do not need MATH 241 to come to this conclusion.
This is just common sense!
Sample #2: You get 98 heads. Conclusion: reject the null.
You do not need MATH 241 to come to this conclusion.
This is just common sense!
Sample #3: you get 61 heads. Conclusion: ????
I need MATH 241.

Let's run the test for sample #3. Let alpha = 5%. Then we compute the retainment region:

$$RET = \left[p_0 \pm z_{\frac{\alpha}{2}} \sqrt{\frac{p_0(1-p_0)}{n}}\right]$$

$$= \left[0.5 \pm 2 \sqrt{\frac{0.5(1-0.5)}{100}}\right]$$

$$= [.371, .629]$$

$$\hat{p} = \frac{61}{100} = .61 \in RET \Rightarrow \text{Retain Null.}$$

This test is called a 2-sided 1-proportion z test.

2-sided: reject on left and right,
1-proportion: your sample stat is a proportion
z-test: use the std normal distr. from the CLT

The choice of alpha really does matter. Slightly larger and the result of this test would've been "reject".

There are two types of errors in hypothesis testing:

		decision: Retain H_0	decision: Reject H_0
truth	H_0	✓	Type I error
	H_a	Type II error	✓

If you reject the null, it is possible you made a Type I error. If you retain the null, it is possible you made a Type II error. You have no way of knowing if you made an error or not!!!

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

$P(\text{Type II error})$ is not covered in 241 but is covered in 369

$\text{Power} := 1 - P(\text{Type II error})$ is the probability of finding your effect (proving your theory)

$$P(\text{Type I error}) \uparrow \Rightarrow P(\text{Type II err}) \downarrow$$

$$P(\text{Type I err}) \downarrow \Rightarrow P(\text{Type II err}) \uparrow$$

There are different costs to Type I and Type II errors. For ex:

Clinical trial example
Null: drug doesn't cure cancer
Alt: drug cures cancer

Type I err: conclude drug cures cancer but in reality drug does not
Type II err: conclude drug does not cure cancer but in reality it does

Cost of Type I error: ???
Cost of Type II error: ???

Fire alarm example
Null: no fire
Alt: fire

Type I err: conclude fire but in reality no fire (false alarm)
Type II err: conclude no fire but in reality fire (alarm never sounded)

Cost of Type I error: low
Cost of Type II error: high
Thus, alpha should be HIGH

Let's run the gender birth experiment again this time with a whole lot more data (higher n, the sample's size). In 2000, there were n = 4,247,000 babies born in America and 2,173,000 were male. Let's see what happens:

$$H_a: p \neq 0.5, \quad H_0: p = 0.5, \quad \alpha = 5\%$$

$$RET = \left[0.5 \pm 2\sqrt{\frac{0.5(1-0.5)}{4247000}}\right]$$

$$= [.49516, .50484]$$

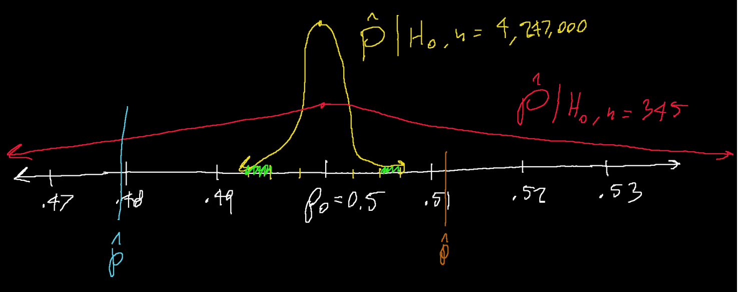
this RET is much tighter than last class with n=345

$$\hat{p} = \frac{2,173,000}{4,247,000} = .51165 \notin RET \Rightarrow \text{Reject } H_0$$

We conclude that P(male child) is not = 50%.
Last class we retained the null! However, that was a Type II error. This is correct. We have a much lower P(Type II error) here since n is much higher.

$$n \uparrow \Rightarrow P(\text{Type II err}) \downarrow$$

$$n \uparrow \Rightarrow P(\text{Type I err})$$ doesn't change



Higher n gives you a lower standard error of the sampling distribution (P-hat) and thus you have a small retainment region allowing for detection of small deviations from the null.

I
H_0: Aliens and UFO's don't exist
H_a: Aliens and UFO's do exist
α: Low

II
H_0: Aliens and UFO's don't exist
H_a: Aliens and UFO's do exist
α: HIGH

III
H_0: Aliens and UFO's do exist
H_a: Aliens and UFO's don't exist
α: Low

IV
H_0: Aliens and UFO's do exist
H_a: Aliens and UFO's don't exist
α: HIGH