

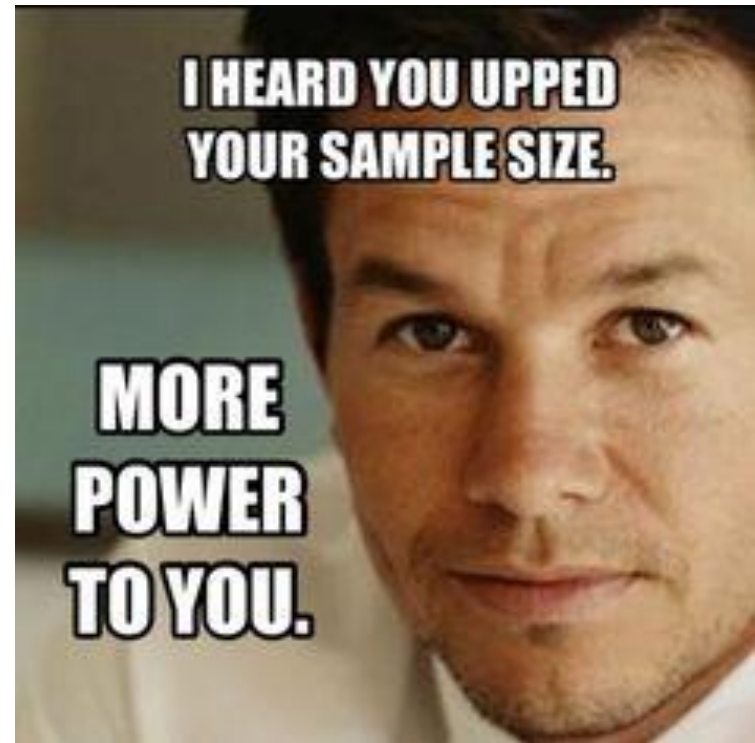
# Study Session Week of 4/23

## Objectives:

- I will discuss the differences between Type I and Type II errors
- I will make connections between errors, alpha, beta and power

## Agenda:

- Review errors
- Review alpha, beta & power
- Practice with MC
- Practice with FR



# Type I Errors vs. Type II Errors

- **Type I Errors** are choosing to **reject the null** based on evidence (p-value is lower than alpha), but in reality **the population null is true**

ART is my BFF  
 $\alpha$   $\beta$

- **Type II errors** are choosing to **fail to reject the null**, but in reality **the population null is false**.

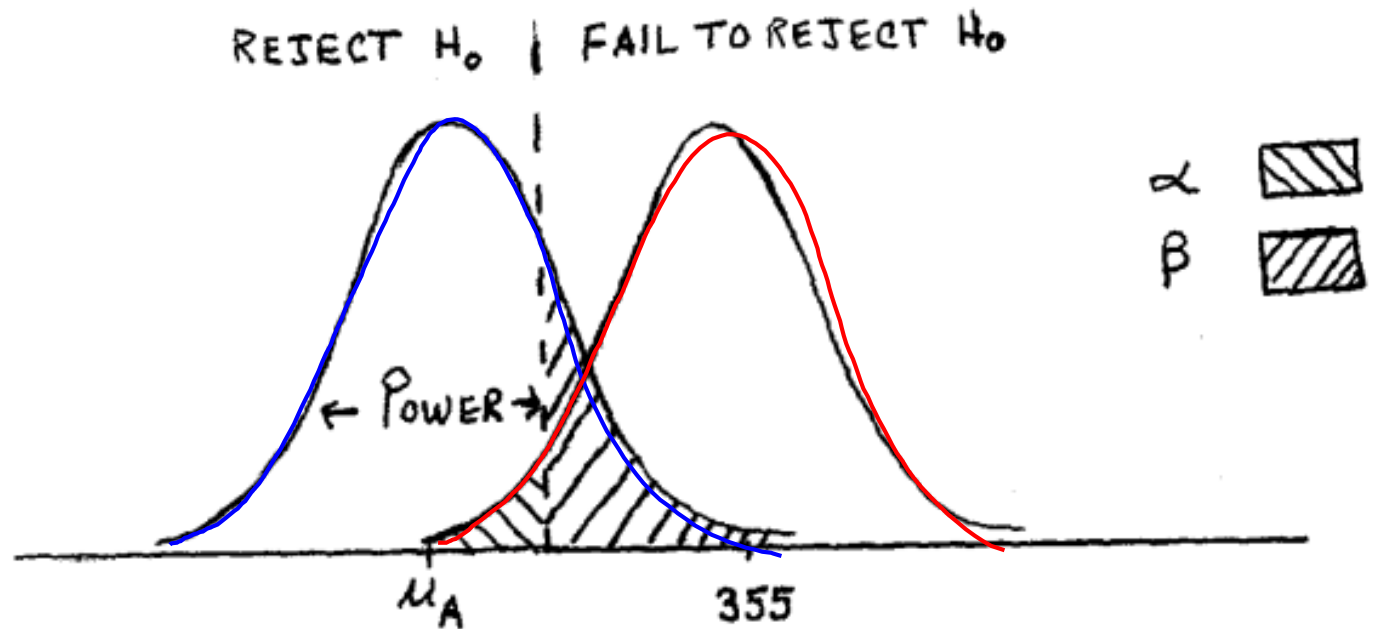
- When addressing these, you will also have to put it in context – PRO

TIP: mention **sample evidence** vs. **population reality**

	Reject H0	Fail to Reject H0
Reality: H0 is True	Type I error (probability = $\alpha$ )	Probability = $1-\alpha$
Reality: H0 is False	Power ( $1-\beta$ )	Type II error (probability = $\beta$ )

# Alpha & Beta

The diagram illustrates the relationship between the chances of making a Type I error, making a Type II error, and the power of the hypothesis test. If the sample size is increased the two curves will overlap less because the standard error will be smaller, thus increasing the power. If  $\alpha$  is increased, then the decision line for the test will move right in this case and increase the power. Increasing the distance between 355 (the sample mean) and the true alternative mean moves the left curve farther left, increasing the power.



# Power – how to increase

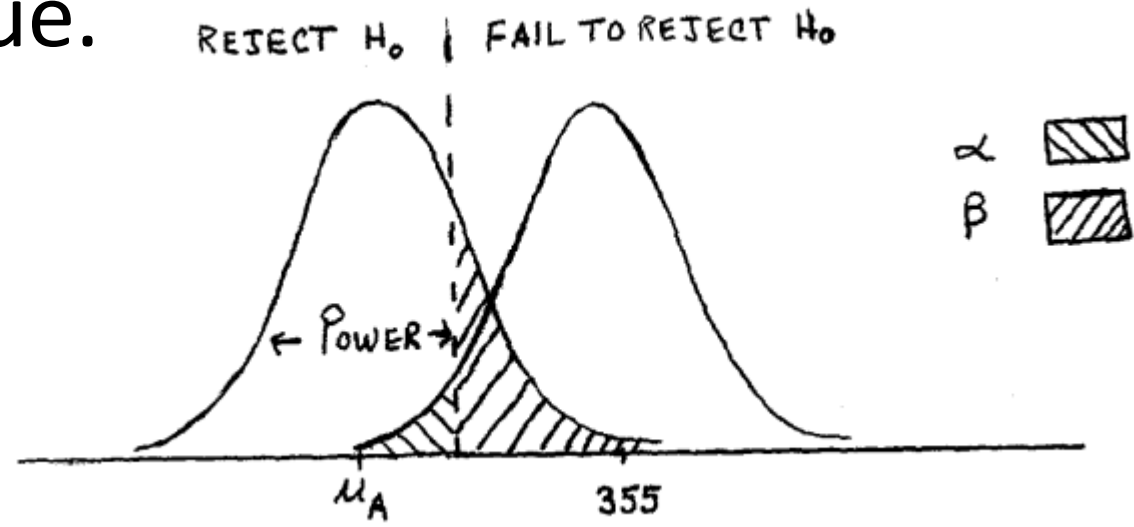
- Increase the sample size  $n$  ↑
- Increase the probability of making a Type I error  $\alpha$  ↑
- Increase the distance between the hypothesized parameter and the true alternative value.

power =  $1 - \beta$

correctly reject a false null

type II error  
failing to reject FALSE null

effect size



<https://bit.ly/2EUt09A>



# Multiple Choice Practice #1

Suppose a machine that makes pegs to be used in holes to hold furniture parts together is malfunctioning, but the manufacturer doesn't know it. A quality control test is conducted bimonthly with the null hypothesis stating that the machine works properly. The p-value of the most recent test was 0.185. What probably happens as a result of this test?

$H_0$ : machine works — assume true  
 $H_A$ : machine doesn't work

Fail to reject +

- (A) The test correctly fails to reject  $H_0$ .
- ~~(B)~~ The test correctly rejects  $H_0$ .
- ~~(C)~~  $H_0$  is rejected, resulting in a Type I error.
- ~~(D)~~  $H_0$  is not rejected, resulting in a Type I error.
- (E)  $H_0$  is not rejected, resulting in a Type II error.

not type I  
BFF

# Multiple Choice Practice #2

Which sample size and significance level will give a test of highest power?

~~(A)~~  $n = 25, \alpha = 0.01$

~~(B)~~  $n = 25, \alpha = 0.05$

**(C)**  $n = 50, \alpha = 0.10$

(D)  $n = 50, \alpha = 0.05$

(E)  $n = 50, \alpha = 0.01$

$n \uparrow$        $\alpha \uparrow$       power  $\uparrow$

# Multiple Choice Practice #3

Which of the following statements are false concerning Type I and Type II errors?

- I. A Type I error is always worse than a Type II error **F**
- II. The higher the probability of a Type I error the lower the probability of a Type II error. **T**
- III. A Type I error incorrectly rejects a true alternative hypothesis. **F**

AR True null  
|  
rejecting

- (A) I only
- (B) II only
- (C) III only
- (D) I and III only**
- (E) II and III only

# Multiple Choice Practice #4

The manufacturer of the latest greatest cell phone uses a survey to determine the opinions of a group of potential consumers hired to test the phone for a period of one month. If the new cell phone is viewed as working better than existing cell phones, the company will launch the product. The hypotheses to a significance test for the survey results are listed below in word form.

$H_0$ : The new phone works the same as existing phones

Think true

actually false (error)

$H_A$ : The new phone works better than existing phones.

Which statement below describes a Type II error and a possible consequence of that error.

Fail to Reject

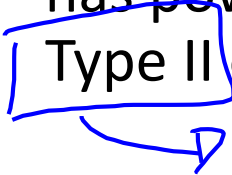
- (A) The new phone does work better than existing phones so the company launches the product.
- (B) The new phone does work better than existing phones but the company does not launch the product.
- (C) The new phone works the same as existing phones but the company launches the product.
- (D) The new phone works the same as existing phones so the company does not launch the product.
- (E) The new phone works the same as existing phones so the company requires a new test group.



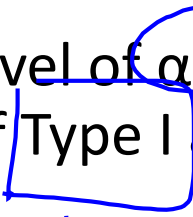
# Free Response Practice #1

$$1 - 0.78 = 0.22$$

You read that a statistical test at a significance level of  $\alpha = .05$  has power of 0.78. What are the probabilities of Type I and Type II errors for this test?



0.22



$\alpha$

0.05



$$\text{power} = 1 - \beta$$

$$\beta = 1 - \text{power}$$

$$\frac{+\text{power}}{-\beta} \quad \frac{+\text{power}}{-\beta}$$

$$\text{power} = 1 - \beta$$

# Free Response Practice #2

You read that a statistical test at a significance level of  $\alpha = .01$  has probability 0.14 of making a Type II error when a specific alternative is true. What is the power of the test against this alternative?

$$1 - 0.14 = 0.86$$

- x 0.01
- x 0.14
- x 0.99
- x 0.15
- x 0.13

wrong

$\times H_0: \mu = 130$  ✖

$H_A: \mu > 130$

# Free Response Practice #3

Your company markets a computerized device for detecting high blood pressure. The device measures an individual's blood pressure once per hour at a randomly selected time throughout a 12-hour period. Then it calculates the mean systolic (top number) pressure for the sample of measurements. Based on the sample results, the device determines whether there is significant evidence that the individual's actual mean systolic pressure is greater than 130. If so, it recommends that the person seek medical attention.

a) State appropriate null and alternative hypotheses in this setting. Be sure to define your parameters.

b) Describe a Type I and a Type II error, and explain the consequences of each.

$\alpha$  RT      OF F

$\uparrow \alpha 0.1$

c) The blood pressure device can be adjusted to decrease one error probability at the cost of an increase in the other error probability. Which error probability would you choose to make smaller. Explain.

# Free Response Practice #4

A drug manufacturer claims that fewer than 10% of patients who take its new drug for treating Alzheimer's disease will experience nausea. To test this claim, a significance test is carried out with

$$H_o: p = 0.10$$

$$H_a: p < 0.10$$

where  $p$  is the proportion of patients on new drug who experience nausea.

You learn that the power of this test at the 5% significance level against the alternative of  $p = 0.08$  is 0.64.

- Explain in simple language what “power = 0.64” means in this setting.
- You could get higher power against the same alternative with the same  $\alpha$  by changing the number of measurements you make. Should you make more measurements or fewer to increase power. Explain.
- If you decide to use  $\alpha = 0.01$  in place of  $\alpha = 0.05$  with no other changes to the test, will power increase or decrease? Justify your answer.
- If you shift your interest to the alternative 0.07 with no other changes, will the power increase, decrease or stay the same? Justify your answer.