AP Statistics – 9.1Power 9.1B	Name: 2020 KEY
Goal: Understand Power and Review Type I and Type II Errors	Date:

Section A - REVIEW Type I and Type II Errors

Test Yourself:
Fill in the table: Reject Ho, FTR Ho, Ho true, Ho False, Type I error, Type II error, Power, α, β

Truth about the population

Conclusion based on sample		H_0 TRUE	H_0 FALSE
	Reject H_0	TYPE I ERROR (α)	CORRECT (power = $1-\beta$)
	FAIL TO REJECT H_0	CORRECT	TYPE II ERROR (β)

Example "*Faster fast food?*" The manager of a fast-food restaurant wants to reduce the proportion of drive-through customers who have to wait to receive their food once their order was placed. Based on store records, 63% of customers had to wait when they got to the cashier's window. To reduce this wait time proportion, the manager assigns an additional employee to assist with drive-through orders. During the next month the manager will collect a random sample of drive-through times.

Ho: p = 0.63 Ha: p < 0.63 IMPROVED CUSTOMER SATISFACTION

where ***p = the true proportion of drive-through customers who have to wait more than 2 minutes after their order is placed to receive their food.***

1) Describe **Type I Error** in this setting. Explain the financial consequence(include the impact with new employee)

THE MANAGER FINDS CONVINCING EVIDENCE CUSTOMER SATISFACTION HAS IMPROVED WITH THE EXTRA EMPLOYEE, WHEN IN FACT 63% OF CUSTOMERS WAIT LONGER THAN 2 MINUTES.
CONSEQUENCE: MGR IS SPENDING MORE $'S FOR AN ADDITIONAL EMPLOYEE AND C.S. NOT IMPROVED.

2) Describe **Type II Error** in this setting. Explain the financial consequence(include the impact with new employee)

THE MANAGER DOES NOT FIND CONVINCING EVIDENCE CUSTOMER SERVICE HAS IMPROVED WITH EXTRA EMPLOYEE, WHEN IN FACT 63% OF CUSTOMERS WAIT LESS THAN 2 MINUTES.
CONSEQUENCE: MGR FIRES ADDITIONAL EMPLOYEE AND UPSETS CUSTOMERS WITH POOR SERVICE

WHICH IS WORSE?

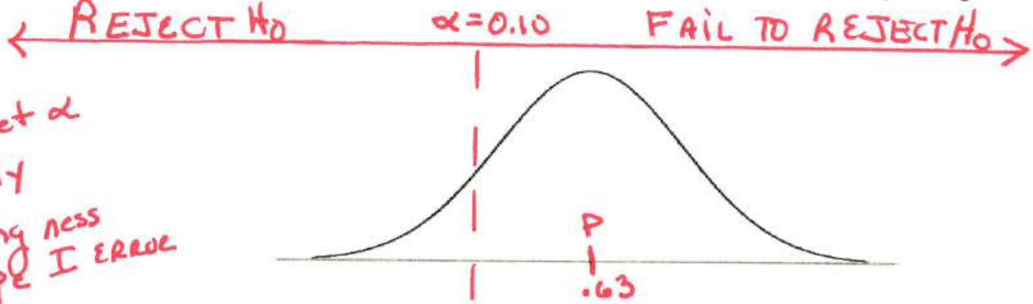
TYPE I ERROR → LOWER α

TYPE II ERROR → INCREASE α

3) Suppose that the manager decided to carry out this test using a significance level **of $\alpha = 0.10$.** He takes a random sample of **250** orders, the manager found **135** customers (about 53%) who had to wait when they got to the cashier's window to receive their food.

- Make a graph labeling the population parameter, α, the rejection region and fail to reject region.

NOTES
- Research set α Before Study
- α is willingness to make Type I ERROR



- What is the probability of a making a Type I error? P(TYPE I ERROR) = α = 0.10
- What is the p-value for this test?

Hypothesis	***Sample Results*** USE [1-PROPZTEST]		
$H_0: p = 0.63$	n=250	x=133	$\hat{p} = \frac{133}{250} \approx .53$
$H_a: p < 0.63$	**Find the P-value =** P(Z < -3.21) = 0.001		

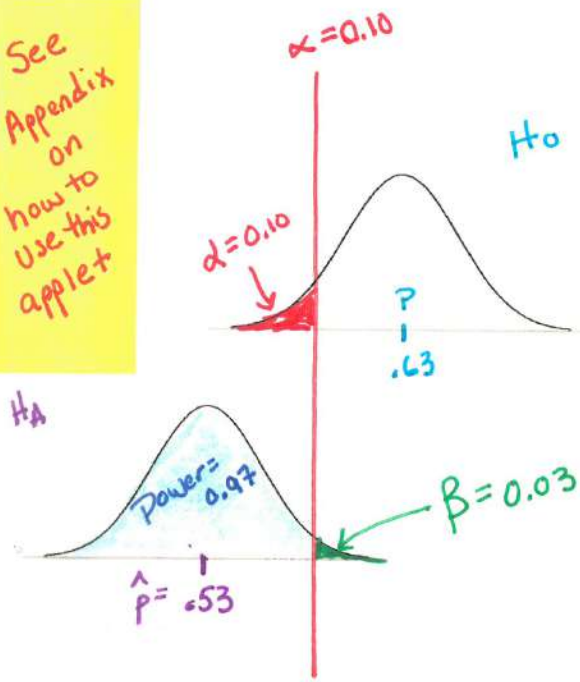
- What decision would the manager make)? Since the pvalue (0.001) IS LESS THAN α=0.10, WE REJECT H_0. WE HAVE CONVINCING EVIDENCE THE CUSTOMER WAIT TIME

Section B - NEW CONCEPTS - Understanding Relationship between Type I and Type 2 Errors AND POWER

DEMONSTRATION: Launch applet "*Improved Batting Averages (Power)*" www.rossmanchance.com/applets

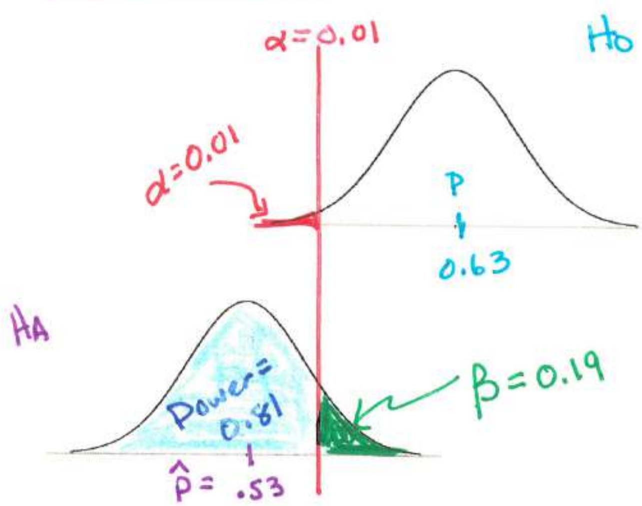
(a) Sketch the Ho and Ha graphs for **α= 0.10** :

See Appendix on how to use this applet



→ n= 250 α= 0.10 Power= .97 β= 0.03

(b) Significance level: To reduce the possibility of a Type I error and avoid the possibility of unnecessarily paying an extra employee, the manager **reduces the significance level from 0.10 to** 0.01

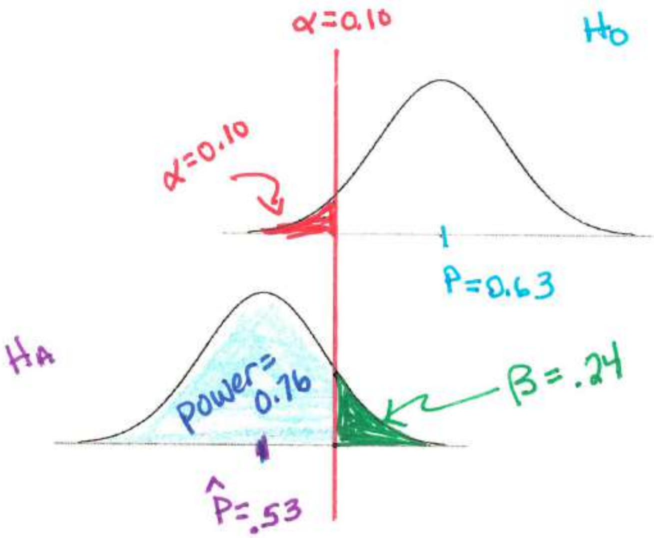


→ n= 250 α= 0.01 Power= 0.81 β= 0.19

Conclusion: reduce α → α↓ β↑ Power↓

Notes/
- α + β have inverse relationships
- α + Power go in same direction

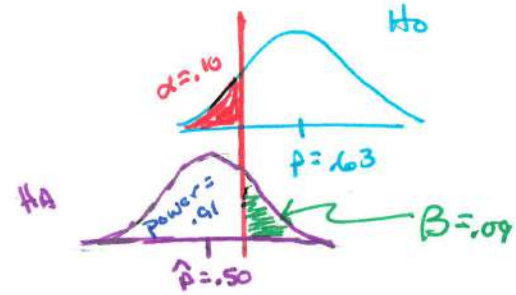
(c) To get faster results, the manager reduces the sample size from 250 **to 100**. **Use α=.10**



NOTICE:
n=250 ⟶ POWER=0.97
n=100↓ ⟶ Power = 0.76 ↓

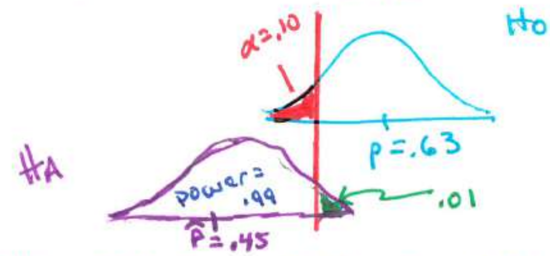
→ n= 100 α=0.10 Power= 0.76 β= 0.24

(d) The manager takes 2 random sample of **100** customers. His first sample, he found **50** customers **(50%)** who had to wait. **Use α=.10**



→ **n=100 α=0.10** Power=0.91 β= 0.09 .

(e) His next sample, He found **45** out **100** customers **(45%)** who had to wait. **Use α=.10**



→ **n=100 α=0.10** Power= 0.99 β= 0.01 .

CONCLUSION: NOTICE, when $\hat{p}$ is further away from p=.63, the power gets bigger

- **Summarize what you have learned about the factors that affect the power of a test.**

What happens when you **increase α?**   α↑ THEN β↓ AND Power↑
What happens when you increase **sample size (n)?**   n↑ THEN power↑
What can researcher control to reduce errors? (3) Good Experimental DESIGN PRACTICES   (1) Sample Size   (2) Predetermine α   * REDUCE EXTRANEOUS FACTORS   * USE OF CONTROL/BLOCKING/STRATIFYING   * REDUCE BIAS ADMINISTER SURVEYS

- What do you do to maximize the **POWER** of a test?

1) Sample Size: ↑ POWER THEN INCREASE "n" (decreases variability)

2) Alpha Level: To ↑ POWER THEN INCREASE α (makes rejecting Ho easier)

3) Alternative P-value: Power is large as the Ha value is more extreme + further away from Ha (pop parameter)

Part D – Good AP Test definitions to know

- Definitions for **POWER** of a test
  - **Power is a correct decision**
  - **Power interpretation**: assuming the Pop. Parameter is _____, there is a ___ probability of finding convincing evidence of the alternative _____
  - **Power Probability Statement** → P(Reject Ho | Ha is true)

- **Probability Statements**
  - **P(Type I Error) = α** → P(Reject Ho | Ho is true)
  - **P(Type II Error) = β** → P(Fail to Reject Ho | Ho is false)
  - **β= 1- Power**

Part E – Power of a Test – Past Multiple Choice Questions

2002 Question #35

In a test of the hypothesis H_0: $\mu = 100$ *versus* H_a: $\mu > 100$, the power of the test when $\mu = 101$ would be greatest for which of the following choices of sample size *n* and significance level α?

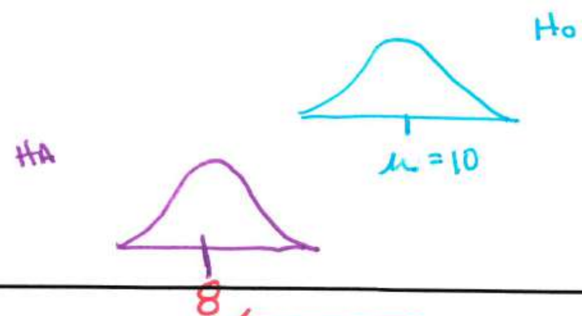
A) $n = 10$, $\alpha = 0.05$
B) $n = 10$, $\alpha = 0.01$
C) $n = 20$, $\alpha = 0.05$
D) $n = 20$, $\alpha = 0.01$
E) It cannot be determined from the information given.

ANSWER C — select the largest sample size (n)
— Select the largest α

2012 Admin version

20. Suppose that on a hypothesis test for a single population mean, H_a: $\mu < 10$. Assume that H_a is true. For a fixed sample size and significance level α, the power of the test will be greatest if the actual mean is which of the following?

(A) 8
(B) 9
(C) 10
(D) 11
(E) 13



Select the mean the furthest away for Ho (μ=10)

ANSWER A

Appendix: <u>Power Demonstration</u>: *How would the following changes affect the power of the test?*

Launch Applet (*Improved Batting Averages*)- http://www.rossmanchance.com/applets/power.html

<u>Example</u> "***Faster fast food?***" $H_0: p = 0.63$ versus $H_a: p < 0.63$

He found from the random sample of 250 orders, that 53% of customers waited more than 2 minutes to receive their food once their order is placed.

(a) **Reduces the significance level: α=0.10 → α=0.01.**

Launch applet

- ***Improved Batting Averages (Power)***
- **www.rossmanchance.com/applets**

For our test of:
Ho: p = 0.63
Ha: p < 0.63

We assume:
phat = 0.53, n = 250, α = 0.10

<u>Step 1 - Enter</u>

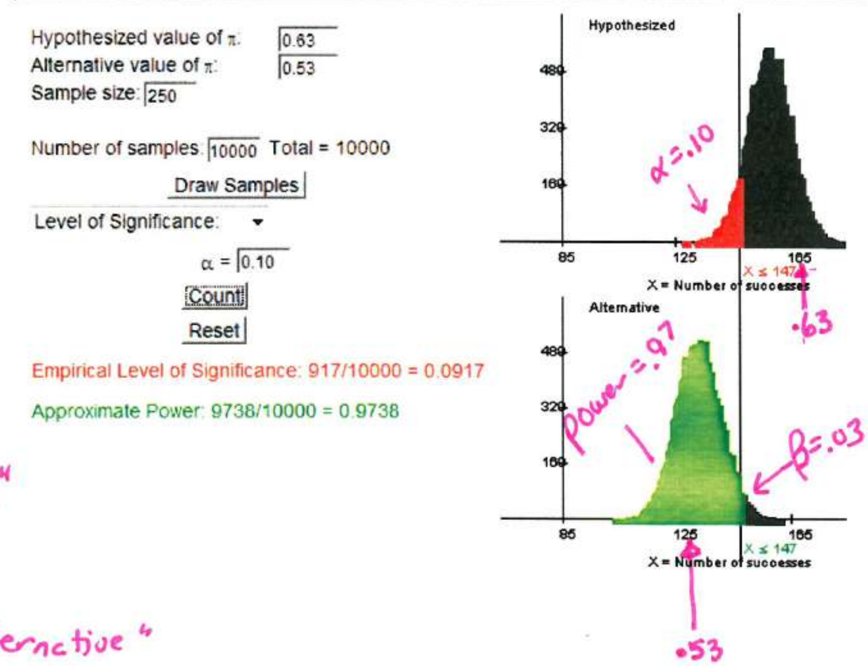
- 0.63 for the hypothesized value of p or π
- 0.53 for the alternative hypothesis
- 250 for the sample size, and
- 10,000 for the number of samples.
- **Press Draw Samples.**

<u>Step 2 - Enter</u> Select "PROPORTION OF SUCCESSES"

- In the drop down menu that says "Choose option," choose Level of Significance and enter 0.10 for α.
- Press "count" ☐ ✓ "show alternative"
- **Result: Power of the test is ~97% and β=.03**

α=0.10 [part a]

Power Simulation $n = 250$



<u>Step 3 -</u>

- Change the value of α = 0.01 and
- Press Count.

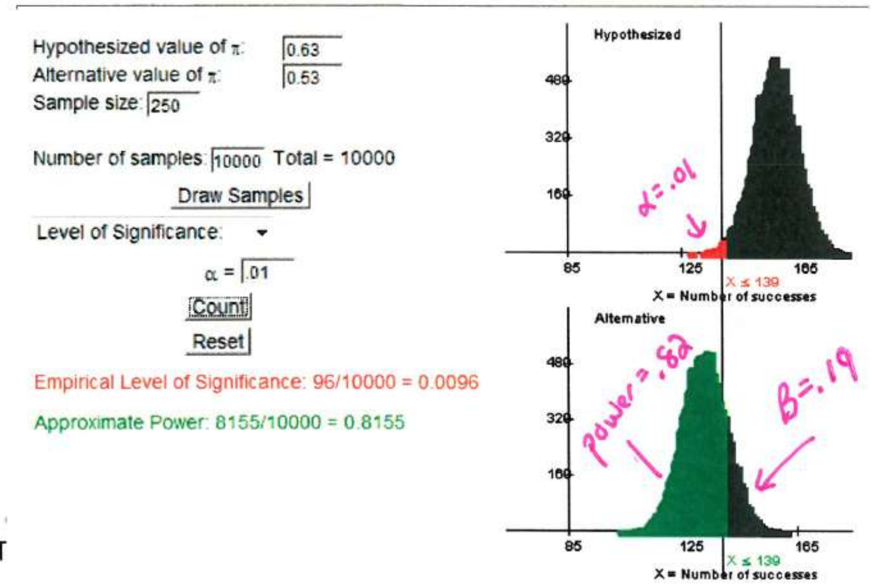
How does the power change?

- **Result: Power of the test is ~82% and β=.18**

* $\boxed{\text{RESET}}$ AFTER EVERY SIMULATION

α=0.01 [part a]2

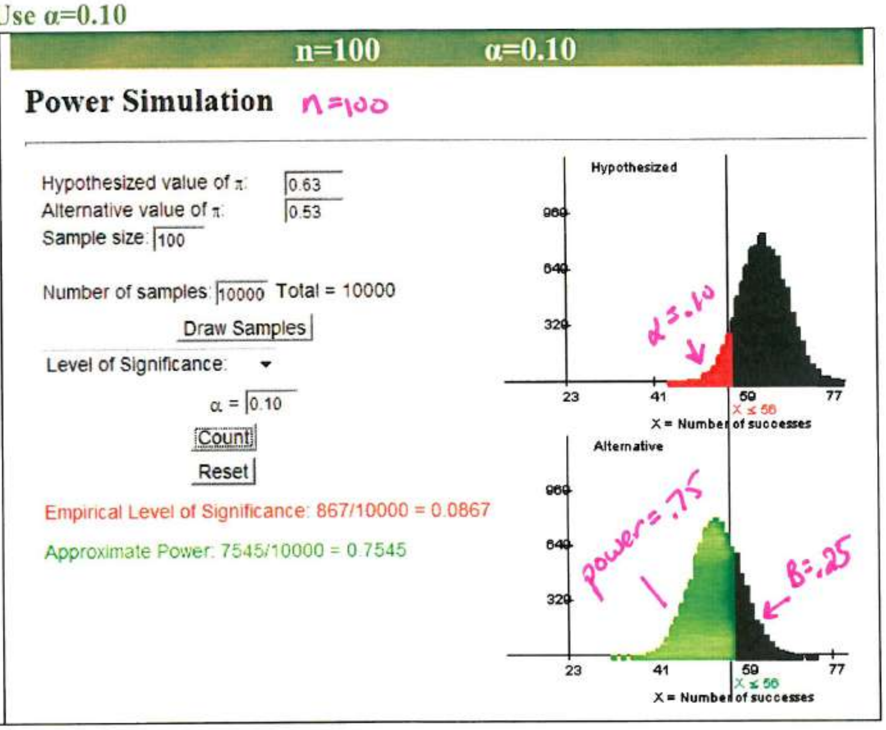
Power Simulation $n = 250$



(b) **Reduces the sample size:** n=250 → n=100. **Use α=0.10**

Step 4 -
- Reset
- Change the sample size to 100
- Draw Samples
- Choose α=.10
- Press "count"

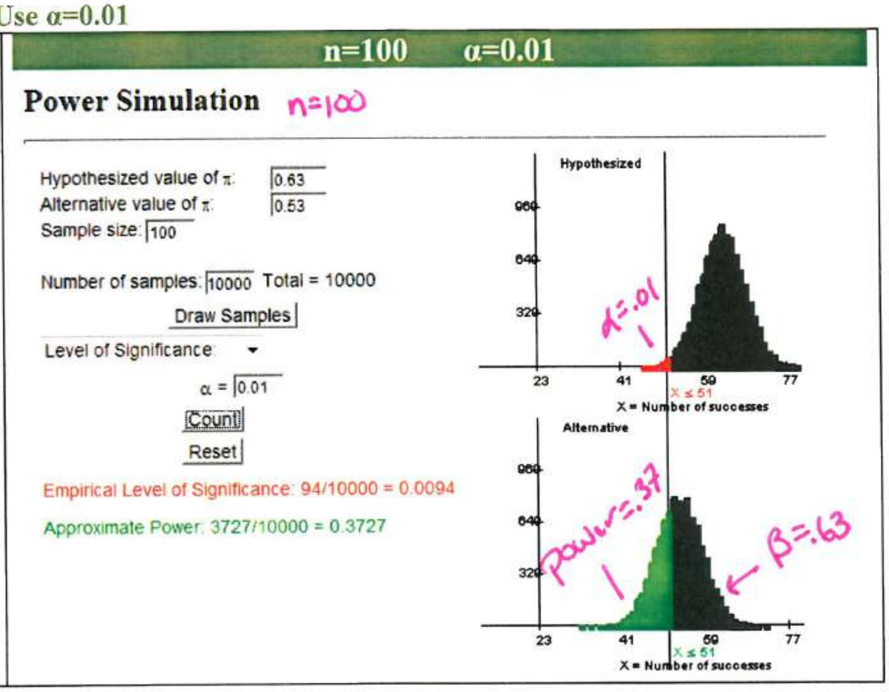
How does the power change?
- **Result: Power of the test is ~75% and β=.25**

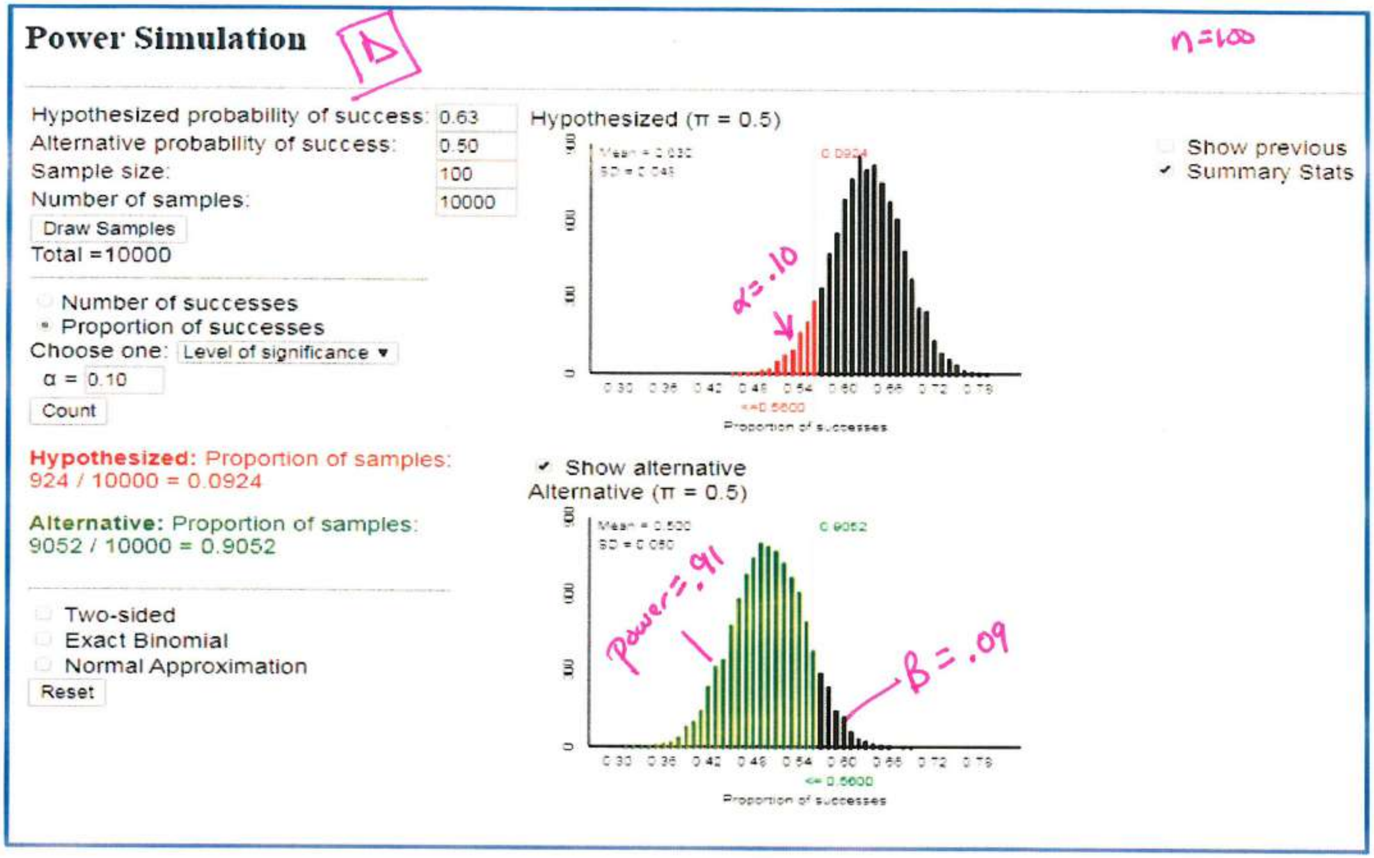


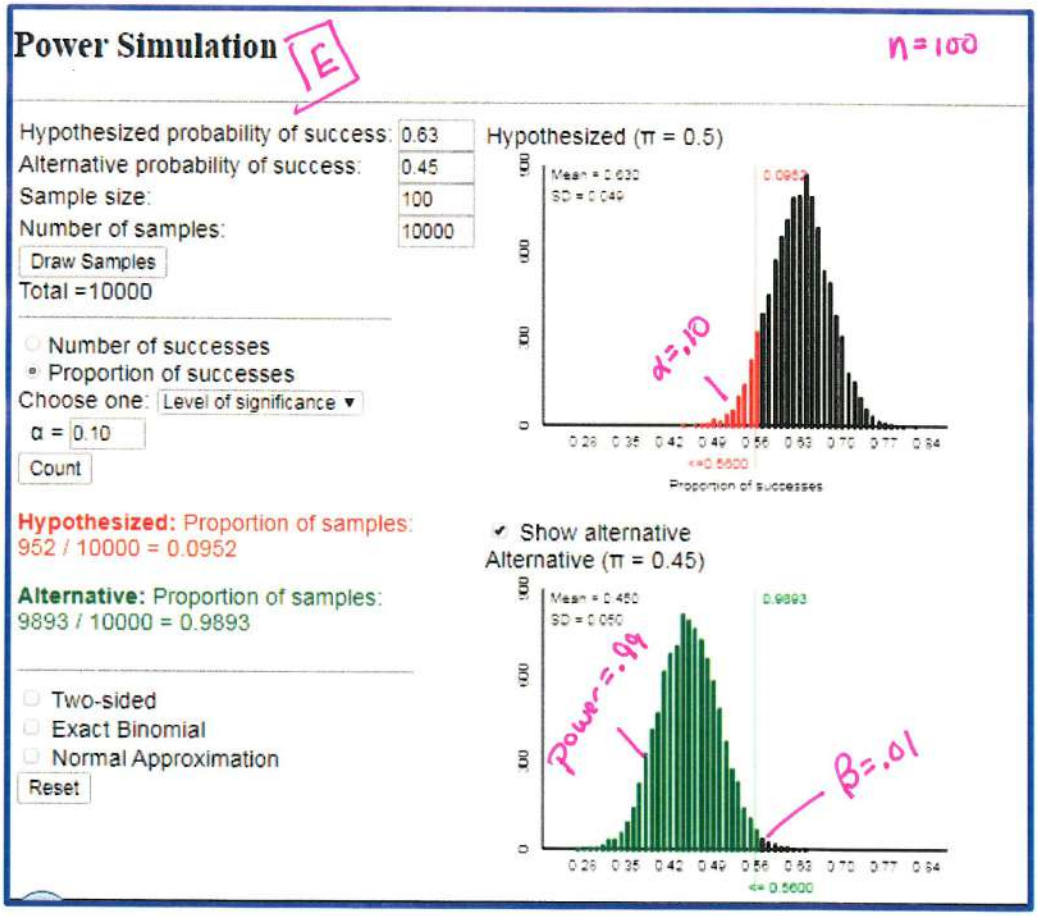
(c) **Reduces the sample size:** n=250 → n=100. **Use α=0.01**

Step 5 -
- Change the value of α = 0.01
- Press "count"

How does the power change?
- **Result: Power of the test is ~37% and β=.63**




Power Simulation
D
n=100
Hypothesized probability of success: 0.63
Alternative probability of success: 0.50
Sample size: 100
Number of samples: 10000
Draw Samples
Total =10000
Number of successes
Proportion of successes
Choose one: Level of significance
α = 0.10
Count
Hypothesized: Proportion of samples:
924 / 10000 = 0.0924
Alternative: Proportion of samples:
9052 / 10000 = 0.9052
Two-sided
Exact Binomial
Normal Approximation
Reset
Hypothesized (π = 0.5)
0.0924
α=.10
0.30 0.36 0.42 0.48 0.54 0.60 0.66 0.72 0.78
<=0.5600
Proportion of successes
Show previous
Summary Stats
Show alternative
Alternative (π = 0.5)
Mean = 0.500
SD = 0.050
0.9052
power=.91
β=.09
0.30 0.36 0.42 0.48 0.54 0.60 0.66 0.72 0.78
<= 0.5600
Proportion of successes


Power Simulation
E
n=100
Hypothesized probability of success: 0.63
Alternative probability of success: 0.45
Sample size: 100
Number of samples: 10000
Draw Samples
Total =10000
Number of successes
Proportion of successes
Choose one: Level of significance
α = 0.10
Count
Hypothesized: Proportion of samples:
952 / 10000 = 0.0952
Alternative: Proportion of samples:
9893 / 10000 = 0.9893
Two-sided
Exact Binomial
Normal Approximation
Reset
Hypothesized (π = 0.5)
Mean = 0.630
SD = 0.049
α=.10
0.28 0.35 0.42 0.49 0.56 0.63 0.70 0.77 0.84
<=0.5600
Proportion of successes
Show alternative
Alternative (π = 0.45)
Mean = 0.450
SD = 0.050
0.9893
power=.99
β=.01
0.28 0.35 0.42 0.49 0.56 0.63 0.70 0.77 0.84
<= 0.5600