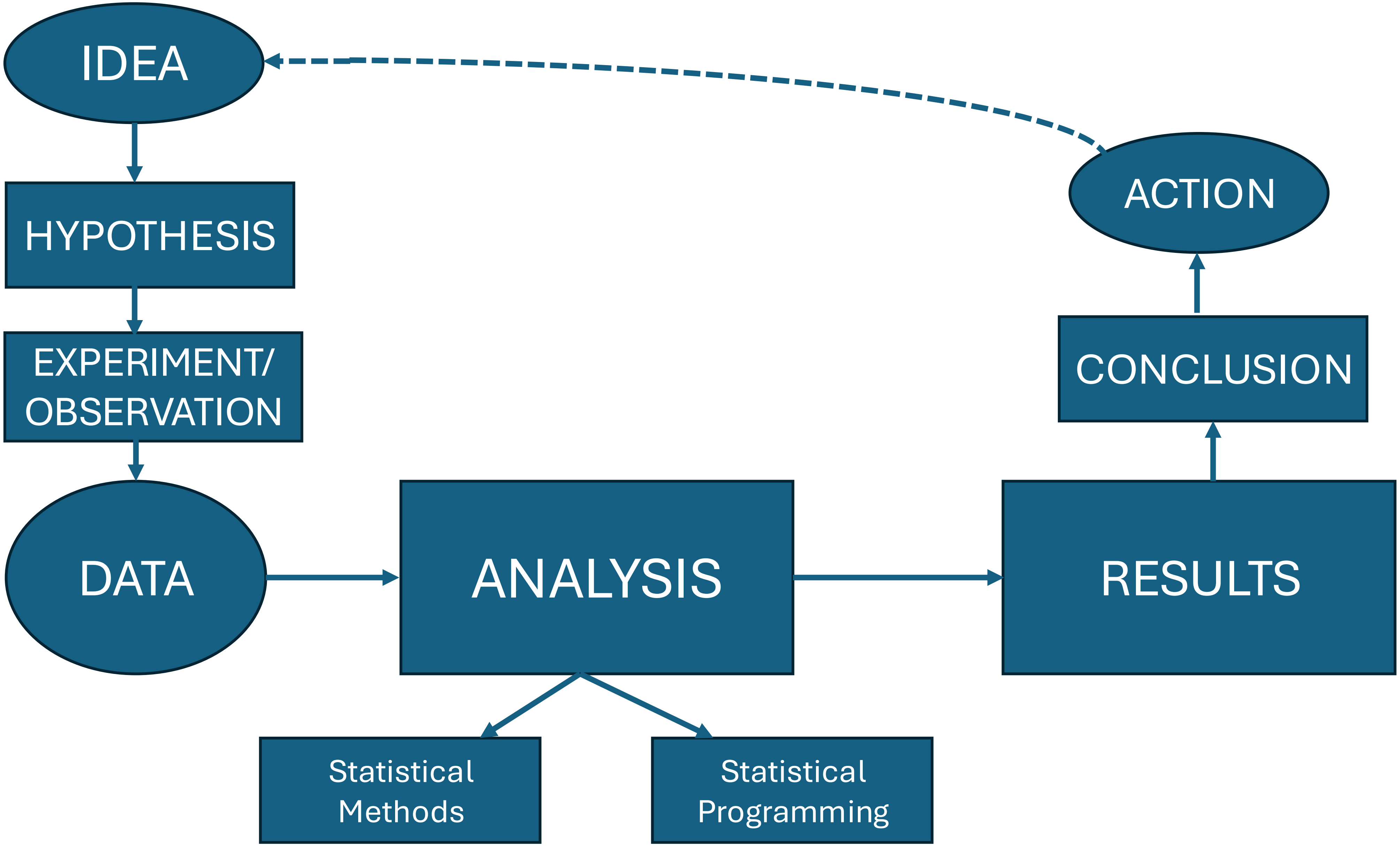
Module 4A: Hypothesis Testing

Applied Epistemology: A Framework for how we know things scientifically

Agenda:

- Working through examples of hypothesis testing
  - **Binomial Example**
  - χ^2 Goodness of fit tests


IDEA
HYPOTHESIS
EXPERIMENT/
OBSERVATION
DATA
ANALYSIS
RESULTS
CONCLUSION
ACTION
Statistical
Methods
Statistical
Programming

Your pipeline for hypothesis testing in statistics

Step 1

Formulate your null **hypothesis**

- Null hypothesis is ***only*** *hypothesis that is tested*
- Falsification: *want* to reject your null

Step 2

Identify appropriate **test statistic**

- Assumptions of your test

Step 3

Quantify the results of your test

- **P value** or comparison to **critical values**
- How *unusual* is your data?

Step 4

Conclude: reject or fail to reject

- based on alpha value
- if appropriate, confidence interval of the parameter

Does wearing a red shirt help win during a wrestling match?

Data from the 2004 Olympics in combat sports: wrestling, taekwondo and boxing.

<u>Does wearing a red shirt help win during a wrestling match?</u>

16 out of **20 rounds** had more red-shirted than blue-shirted winners in the 2004 Olympics in wrestling, taekwondo and boxing.

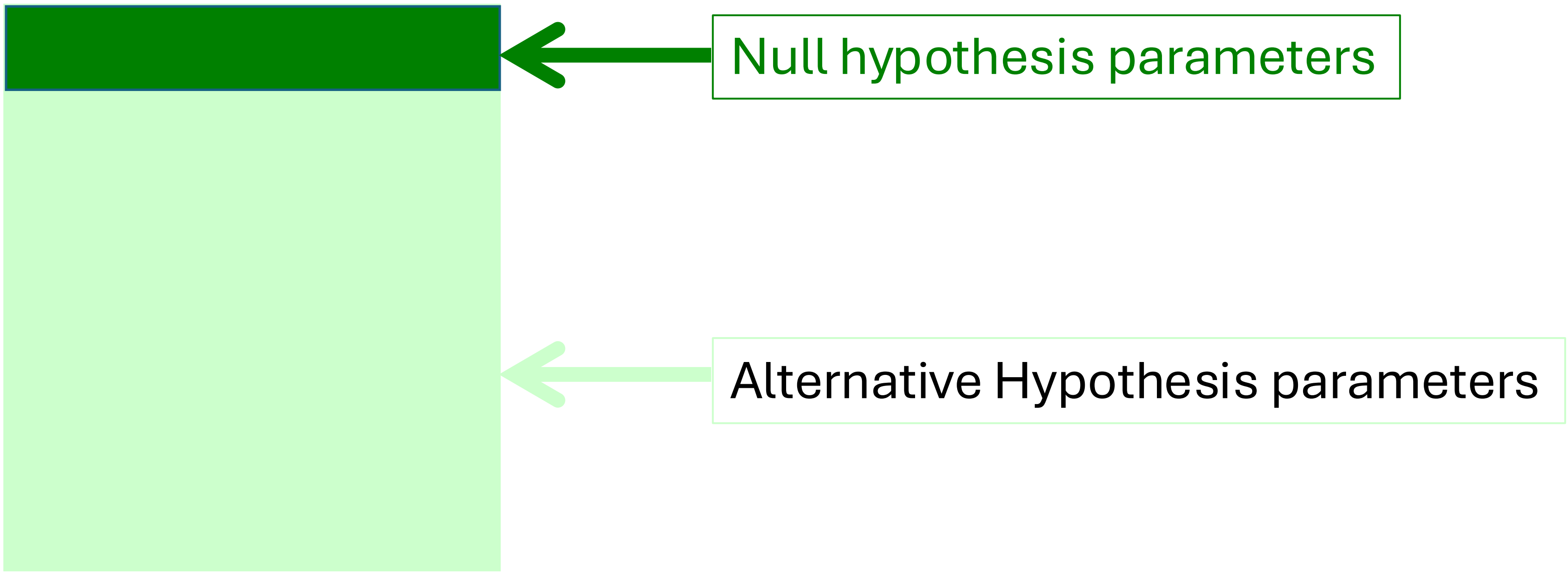
Does wearing a red shirt help win in combat sports?

Step 1: Formulate Hypothesis

Four steps in hypothesis testing:

1. Formulate Hypothesis

- Most of the mental effort
- Quantifies how unusual data is *if you assume that the null hypothesis is true*
- H_0 and H_A - mutually exclusive



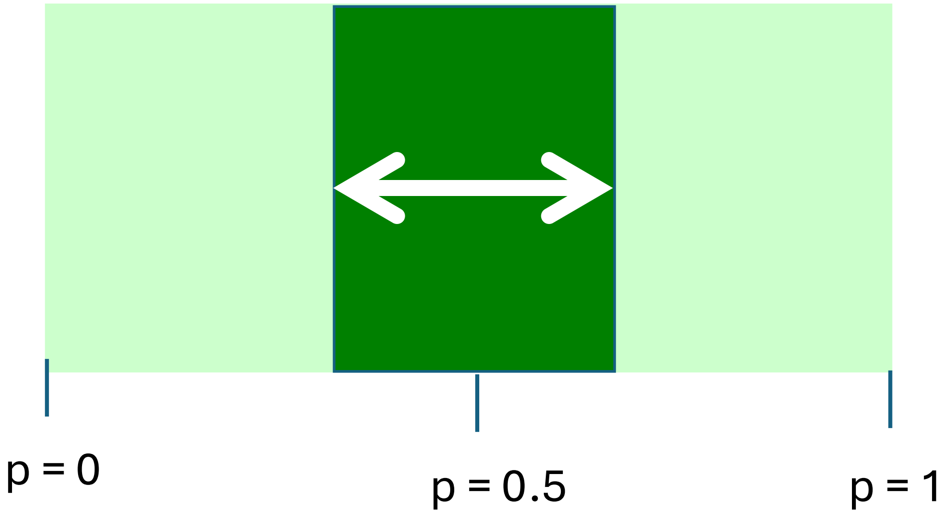
Hypothesis Testing

Does wearing a red shirt help win in combat sports?

Step 1: Formulate Hypothesis

H_0: Red and blue shirted athletes are equally likely to win (proportion = 0.5)

H_A: Red and blue shirted athletes are not equally likely to win (proportion ≠0.5)



Does wearing a red shirt help win in combat sports?

Step 1: Formulate Hypothesis

H_0: Red and blue shirted athletes are equally likely to win (proportion = 0.5)

H_A: Red and blue shirted athletes are not equally likely to win (proportion ≠0.5)

Step 2: Identify test statistic

Does wearing a red shirt help win in combat sports?

Step 1: Formulate Hypothesis

H_0: Red and blue shirted athletes are equally likely to win (proportion = 0.5)

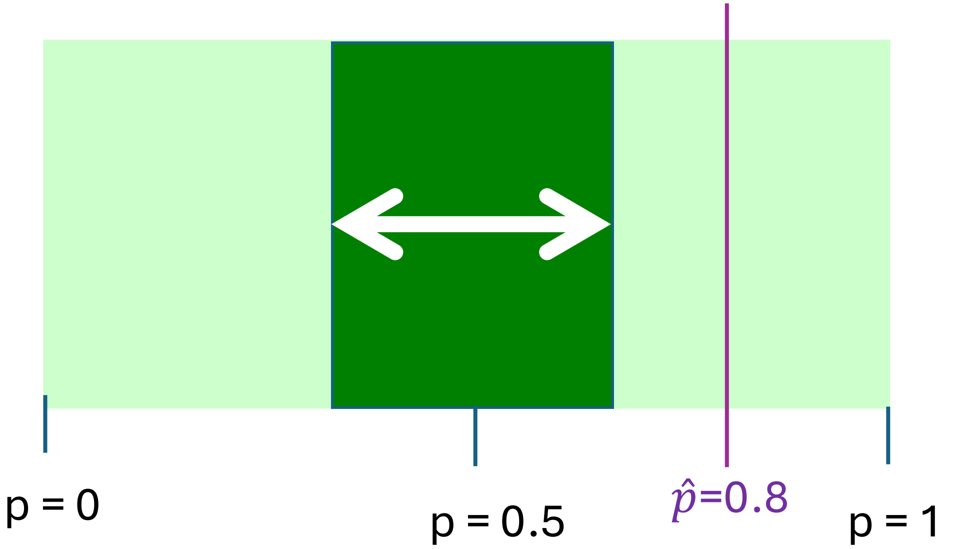
H_A: Red and blue shirted athletes are not equally likely to win (proportion ≠0.5)

Step 2: Identify test statistic

16 out of 20 red shirted winners

--> proportion = **0.8**

This is a discrepancy of **0.3** from H_0. Can it be due to chance alone?



Does wearing a red shirt help win in combat sports?

Step 1: Formulate Hypothesis

H_0: Red and blue shirted athletes are equally likely to win (proportion = 0.5)

H_A: Red and blue shirted athletes are not equally likely to win (proportion ≠0.5)

Step 2: Identify test statistic

16 out of 20 red shirted winners --> proportion = 0.8

Step 3: Calculate the P-Value/Compare to critical values or fixed Significance

Does wearing a red shirt help win in combat sports?

Step 1: Formulate Hypothesis

H_0: Red and blue shirted athletes are equally likely to win (proportion = 0.5)

H_A: Red and blue shirted athletes are not equally likely to win (proportion ≠0.5)

Step 2: Identify test statistic

16 out of 20 red shirted winners --> proportion = 0.8

Step 3: Calculate the P-Value/Com16pare to critical values or fixed significance

If H_0 is true, what is the chance of observing a test statistic with a value at least as extreme as the one we have observed? ←p-value

https://www.wolframalpha.com/input?i=binomial+calculator

Does wearing a red shirt help win in combat sports?

Step 1: Formulate Hypothesis

H_0: Red and blue shirted athletes are equally likely to win (proportion = 0.5)

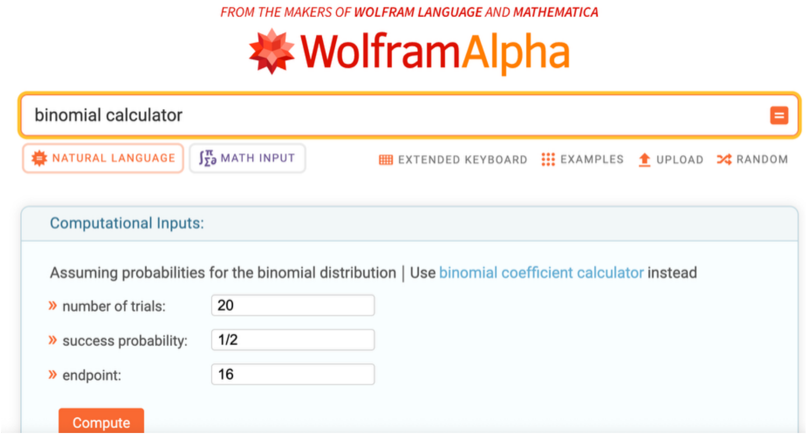
H_A: Red and blue shirted athletes are not equally likely to win (proportion ≠0.5)

Step 2: Identify test statistic

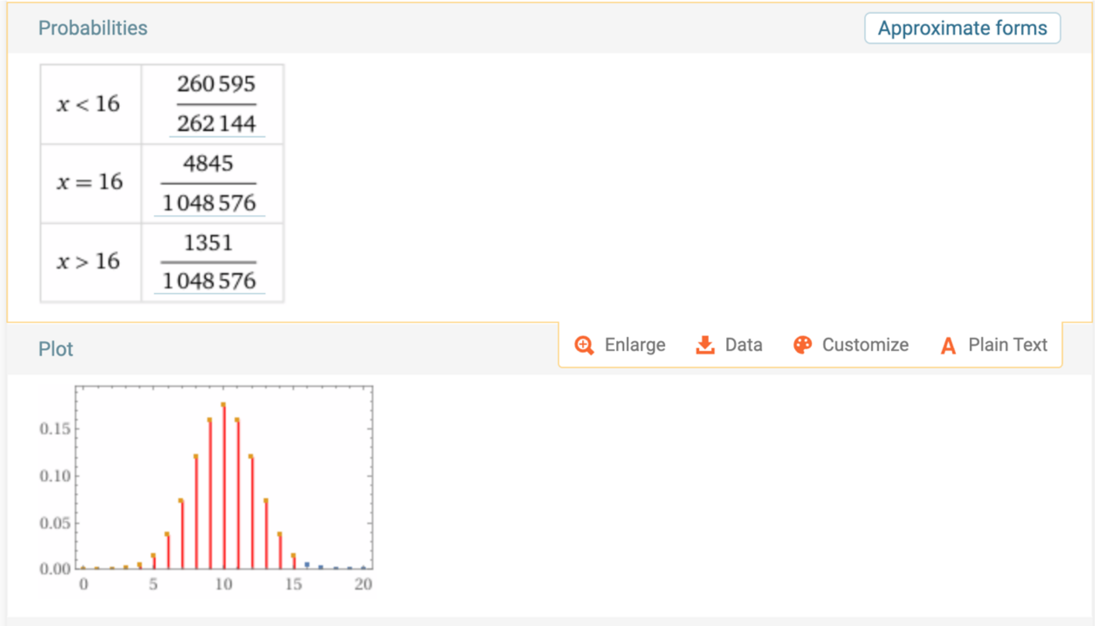
16 out of 20 red shirted winners --> proportion = 0.8

Step 3: Calculate the P-Value/Com16pare to critical values or fixed significance

If H_0 is true, what is the chance of observing a test statistic with a value at least as extreme as the one we have observed? ←p-value



https://www.wolframalpha.com/input?i=binomial+calculator



Does wearing a red shirt help win in combat sports?

Step 1: Formulate Hypothesis

H_0: Red and blue shirted athletes are equally likely to win (proportion = 0.5)

H_A: Red and blue shirted athletes are not equally likely to win (proportion ≠0.5)

Step 2: Identify test statistic

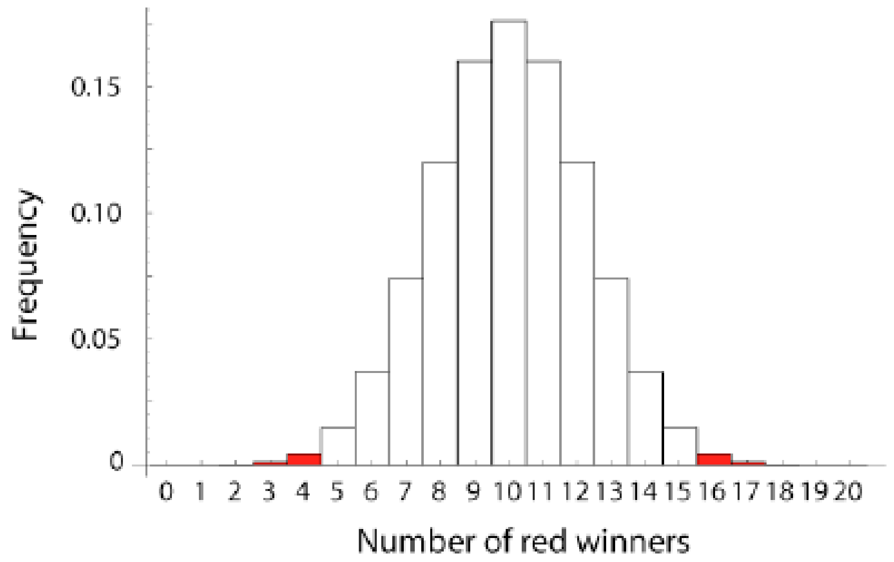
16 out of 20 red shirted winners --> proportion = 0.8

Step 3: Calculate the P-Value/Compare to critical values or fixed significance

Null Distribution of the sample proportion

The Binomial Distribution explains this type of proportion data

If H_0 is true, what is the chance of observing a test statistic value ***at least as extreme*** as the one we have observed?



Does wearing a red shirt help win in combat sports?

Step 1: Formulate Hypothesis

H_0: Red and blue shirted athletes are equally likely to win (proportion = 0.5)

H_A: Red and blue shirted athletes are not equally likely to win (proportion ≠0.5)

Step 2: Identify test statistic

16 out of 20 red shirted winners --> proportion = 0.8

Step 3: Calculate the P-Value/Compare to critical values or fixed significance

If H_0 is true, what is the chance of observing a test statistic value at least as extreme as the one we have observed?

The P-value from the null distribution of the proportion is calculated as:

P = [P[0]+P[1]+P[2]+P[3]+P[4]+P[16]+P[17]+P[18]+P[19]+P[20]]

= due to symmetry

=2XP[16]+P[17]+P[18]+P[19]+P[20]]

= 0.012

$$P\binom{20}{16} = \frac{20!}{16!4!} 0.5^{16}(1-0.5)^4$$

https://stattrek.com/online-calculator/binomial

Does wearing a red shirt help win in combat sports?

Step 1: Formulate Hypothesis

H_0: Red and blue shirted athletes are equally likely to win (proportion = 0.5)

H_A: Red and blue shirted athletes are not equally likely to win (proportion ≠0.5)

Step 2: Identify test statistic

16 out of 20 red shirted winners --> proportion = 0.8

Step 3: Calculate the P-Value/Compare to critical values or fixed significance

P = 2x[P[16]+P[17]+P[18]+P[19]+P[20]] = 0.012

Does wearing a red shirt help win in combat sports?

Step 1: Formulate Hypothesis

H_0: Red and blue shirted athletes are equally likely to win (proportion = 0.5)

H_A: Red and blue shirted athletes are not equally likely to win (proportion ≠0.5)

Step 2: Identify test statistic

16 out of 20 red shirted winners --> proportion = 0.8

Step 3: Calculate the P-Value/Compare to critical values or fixed significance

P = 2x[P[16]+P[17]+P[18]+P[19]+P[20]] = 0.012

What is alpha?

$\alpha = 0.05$ and P-value =0.012

P-value $< \alpha$ so we can reject H_0

Does wearing a red shirt help win in combat sports?

Step 1: Formulate Hypothesis

H_0: Red and blue shirted athletes are equally likely to win (proportion = 0.5)

H_A: Red and blue shirted athletes are not equally likely to win (proportion ≠0.5)

Step 2: Identify test statistic

16 out of 20 red shirted winners --> proportion = 0.8

Step 3(a): Calculate the P-Value

P = 2x[P[16]+P[17]+P[18]+P[19]+P[20]] = 0.012

Step 3(b): Compare to a fixed significance

α =0.05 and P-value =0.012

P< α so we can reject H_0

Step 4: ALWAYS CONCLUDE

Athletes in red and blue shirts are **not equally likely** to win

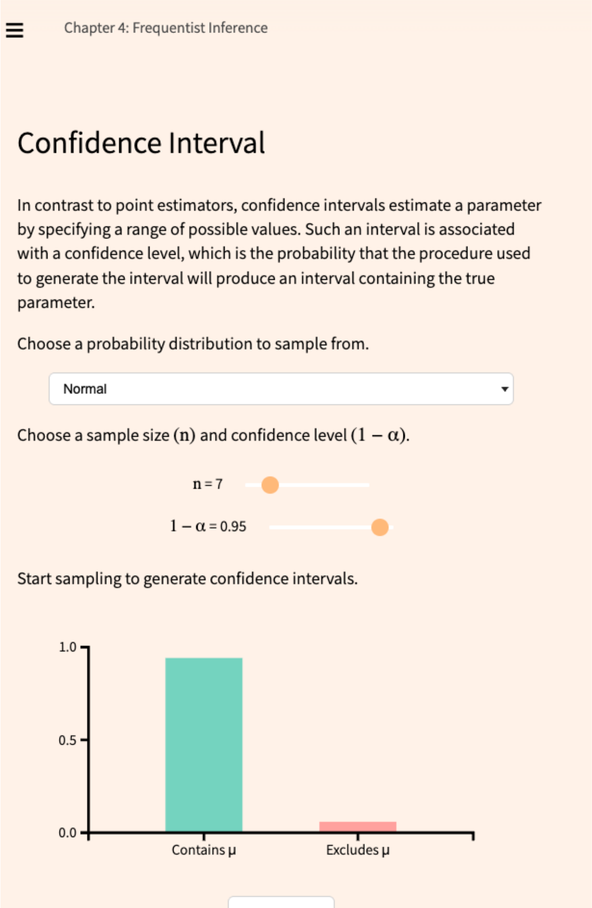
(normally, we also put a confidence interval or any additional information to support our conclusion here, such as confidence interval, effect size calculation or whatever additional evidence is appropriate for your model.)

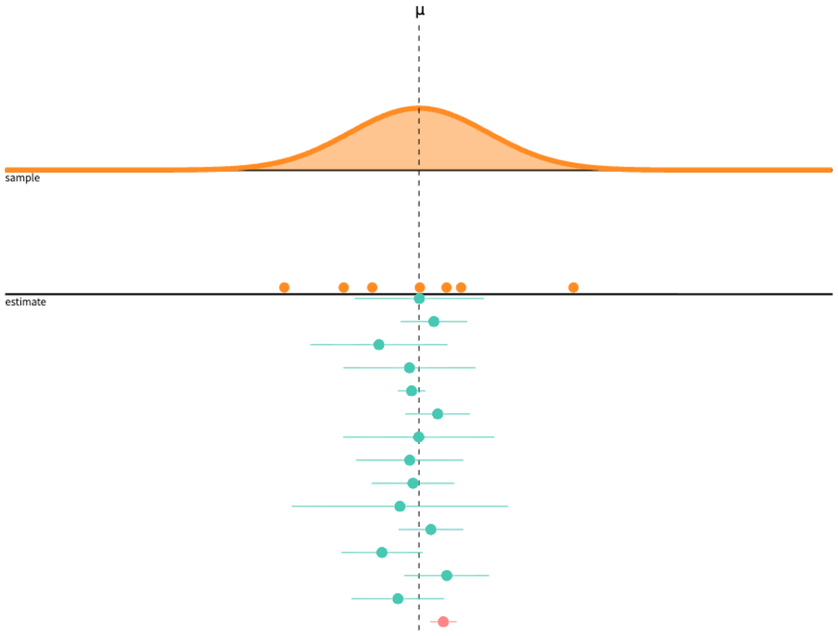
$$\hat{p} - 1.96 * \sqrt{\frac{\hat{p}(1-\hat{p})}{20}} < p < \hat{p} + 1.96 * \sqrt{\frac{\hat{p}(1-\hat{p})}{20}}$$

For Confidence Interval 95%: $0.625 < p < 0.975$

Follow up study → https://www.nature.com/articles/s41598-024-81373-3

Confidence Intervals





https://seeing-theory.brown.edu/frequentist-inference/index.html#section2

95% CI means that if we repeated the study many times, 95% of those intervals would capture the true μ, not that there's a 95% chance *this* one does

Research Claim	H_0 (Null Hypothesis)	H_1 (Alternative Hypothesis)
1. A new cholesterol drug reduces LDL levels compared to placebo.		
2. Sleep duration is associated with fasting glucose levels.		
3. CRISPR editing increases the rate of successful gene knock-in events.		
4. Cancer cells express Gene X more than normal cells.		

Research Claim	H_0 (Null Hypothesis)	H_1 (Alternative Hypothesis)
1. A new cholesterol drug **reduces** LDL levels compared to placebo.	$\mu_1 = \mu_2$	$\mu_1 < \mu_2$
2. Sleep duration is associated with fasting glucose levels.	$\rho = 0$	$\rho \neq 0$
3. CRISPR editing **increases** the rate of successful gene knock-in events.	$p_1 = p_2$	$p_1 > p_2$
4. Cancer cells express Gene X **more than** normal cells.	$\mu_{cancer} = \mu_{normal}$	$\mu_{cancer} > \mu_{normal}$