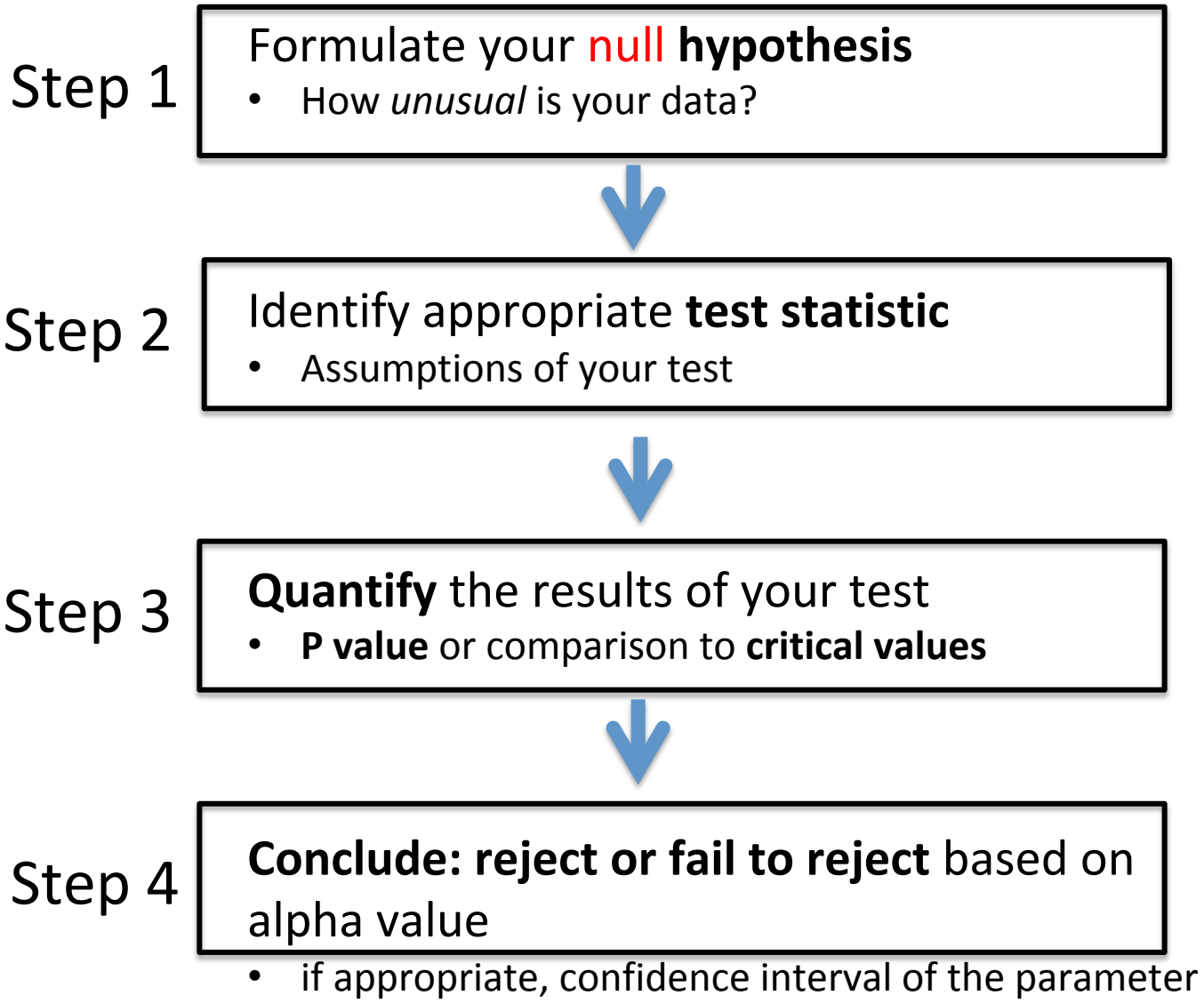
Hypothesis Testing

Your pipeline for hypothesis testing in statistics

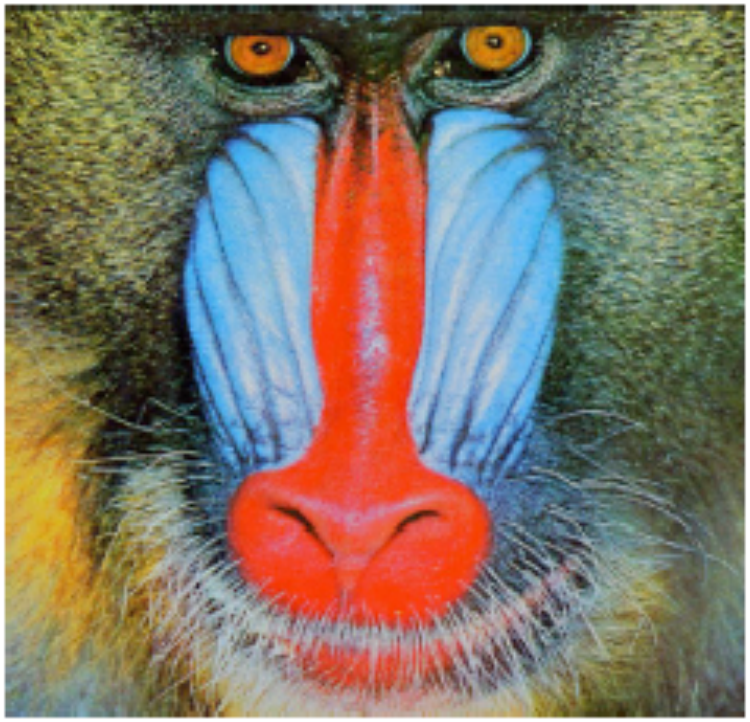


Hypothesis testing automates binary decision making:

- If p-value < alpha (also called significance level)
  - Reject null hypothesis
- If p-value > alpha
  - Fail to reject null hypothesis
- We can outline steps that help us make decisions
- **Remember: What is statistically significant is somewhat arbitrary:**

  p-value of 0.04999 is not so different from 0.050001

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



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

Hypothesis Testing

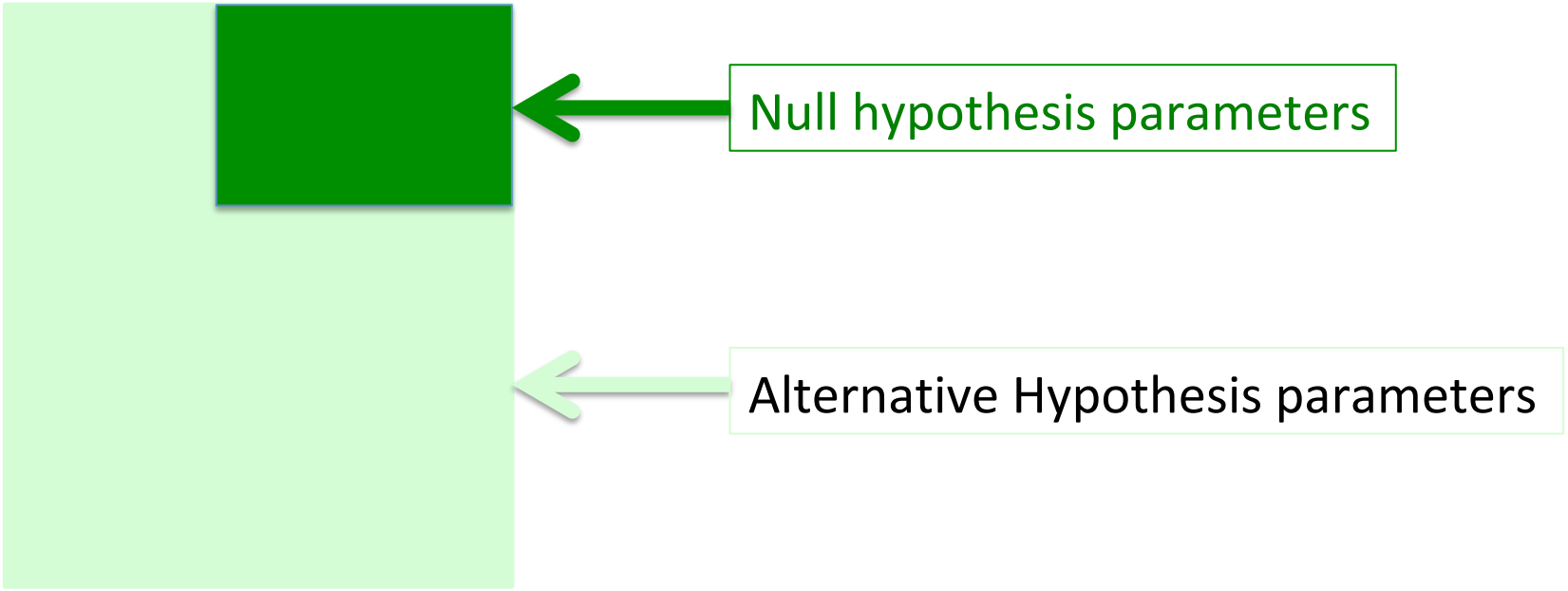
Does wearing a red shirt help win in combat sports?

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

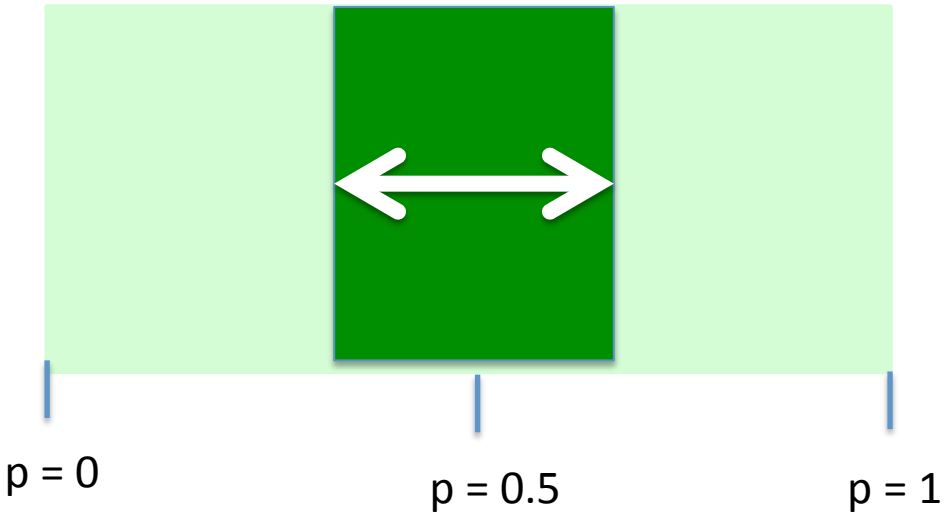


Does wearing a red shirt help win in combat sports?

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/Compare 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?

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

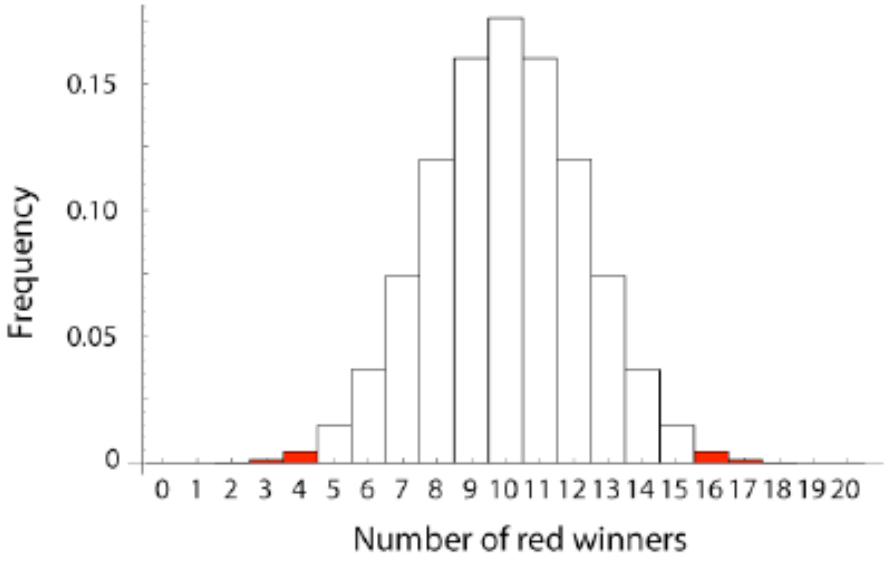
16 out of 20 red shirted winners --> proportion = 0.8

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 (explained in the next chapter)

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 = 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

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)

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?

α =0.05 and P-value =0.012

P < α 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

Someone claims they make 90% of the shots they make on goal in soccer. If this is tested, what would be the null hypothesis?

---

a)You do not have enough information to make a null hypothesis because you don't know how they will do the test

b) $\overline{x}= 0.9$

c)The person does not have a mean of making 90% of their soccer shots

d)The person does have a mean of making 90% of their soccer shots