Hypothesis Testing

Introduction

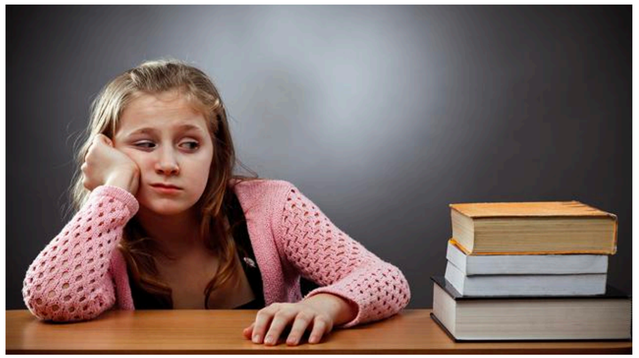
- Hypotheses
- Assumptions about parameters
- Simple & composite
- Null & Alternative
- Test
- Accept or reject null

Preview

← Parameter estimation

Point estimates

Confidence intervals



→ Become the…

Grade credit

“Hypothesis Testing”

Hypotheses

Assumptions (statements) about parameters

Distribution

Population

A coin is biased

Average student GPA is < 3.0

Amazon average delivery time > 2 days

People tweet more on weekend

Men play more video games than women on average

Hypotheses Types

Simple — Parameter takes a single specific value

$\mu = 4.5$ $\sigma = 2.4$

Composite — Parameter takes one of several values

$\mu \in \{4.5, 6.3\}$ $\mu > \sigma$ $\sigma \in [4.5, 6.3)$

One-sided $\mu \leq 2.3$ $\mu > 4.5$

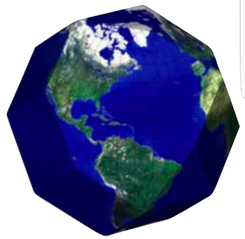
Two-sided $\mu \leq 2.3$ or > 4.5 $\mu < 2.3$ or > 2.3 $\mu \neq 2.3$

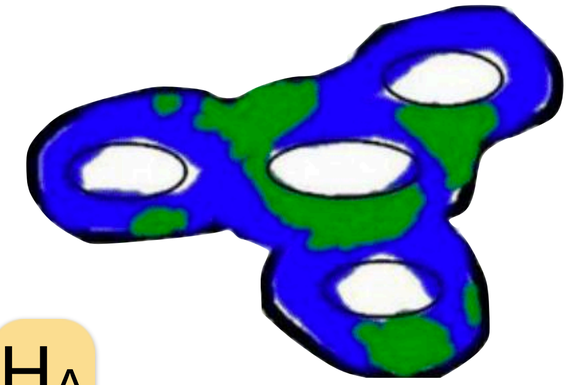
Null and Alternative Hypotheses

Often Assumption believed to be true

Null Status quo hypothesis H_0

Complementary view





Alternative Research hypothesis H_1 H_A

H_A often Complement or “one-side complement” of H_0

Simple H_0

Null

Alternative

Unbiased

$p_h = 0.5$

Biased $p_h \neq 0.5$ 2-sided

Heads more likely $p_h > 0.5$ 1-sided

Gender equality average GPA

Same average GPA

Not exactly simple: {(x,x)}

Different average GPA 2-sided

Men’s average GPA is higher 1-sided

One-Sided H_0

Smartphones
iOS x Android

Null

≥ 60% use iOS

Alternative

< 60% of phones use iOS

Checkout
Self x Cashier

Not exactly one sided: {(x,y): x<y}

Self checkout faster

Self checkout slower

How to Test

Design experiment

Gather data

Data consistent with null hypothesis?

❌ Reject null in favor of alternative

✅ Do not reject null

Equivalently

Strong evidence for alternative hypothesis?

✅

❌

Conservative

Reject null (status quo)

Only if strong evidence against it

Two analogies

Test v. Trial

Hypothesis test

Strong evidence for alternative hypothesis?

✅ Reject null in favor of alternative

❌ Do not reject null

Court trial

Strong incriminating evidence?

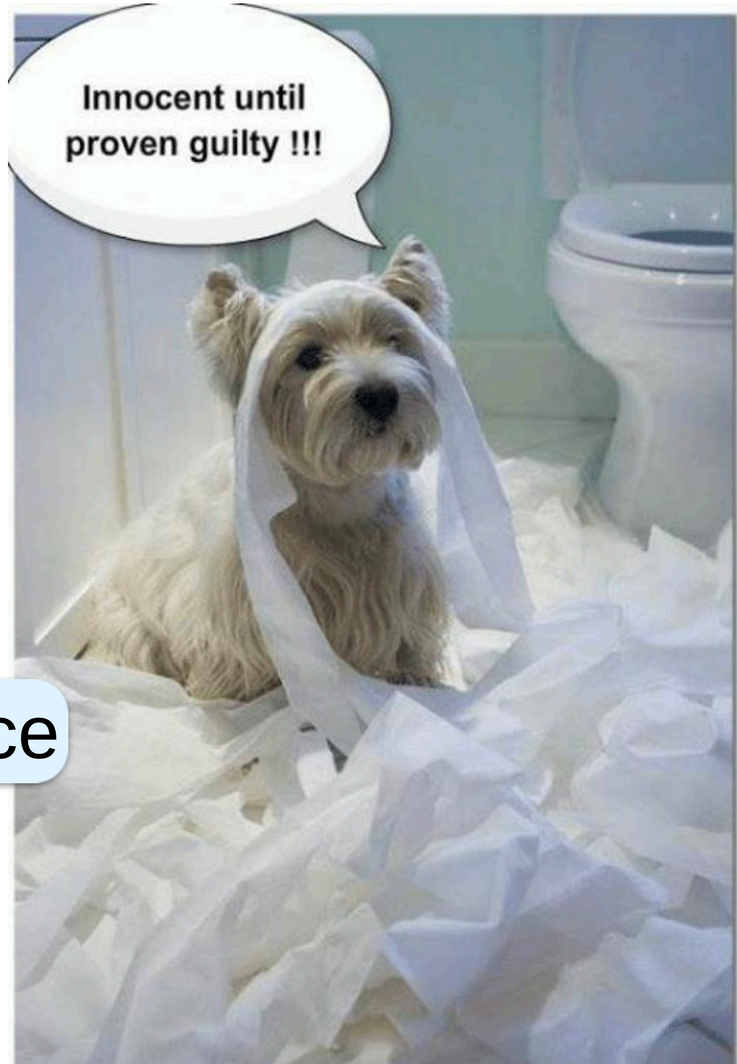
✅ Guilty

❌ Not guilty

Innocence = Null

Presumed innocence

Rejected only by strong evidence



Test vs. Myth

Hypothesis test

Strong evidence for alternative hypothesis?

✅ Reject null in favor of alternative

❌ Do not reject null

Design a test

Strong evidence for myth?

✅ Accept

❌ Keep default belief

Testing Hypotheses

Design Experiment

Test

Define numerical outcome

T

Test statistic

Related to hypothesis

Determine distribution of T under H_0

$P_{H_0}(T=t)$

Observe data

Calculate value t of the test statistic T

Large

$P_{H_0}(t)$

Small
t towards H_A

H_0 consistent with data

Do not reject H_0

Accept H_0

H_0 inconsistent with data

Reject H_0 in favor of H_A

Intuitive → Formal

Coin Bias, 1-Sided H_A

H_0 Unbiased $p_h = 0.5$ Simple null

H_A Biased towards heads $p_h > 0.5$ 1-sided

Test

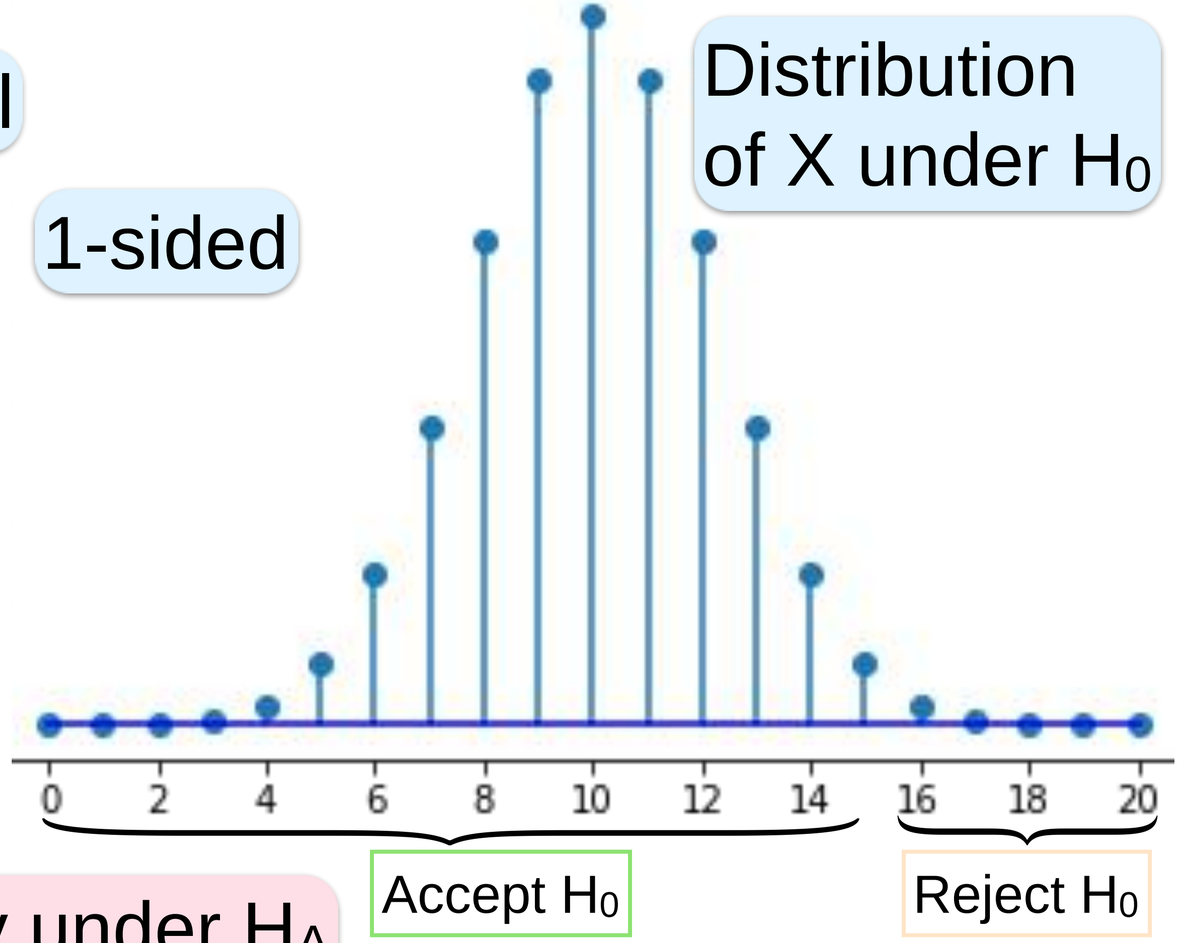
Test statistic Number of heads X

Intuitive

$X \geq 16$ Unlikely under H_0 More likely under H_A

Reject null in favor of H_A

$X < 15$ Do not reject null



Coin Bias, 2-Sided H_A

H_0 Unbiased $p_h = 0.5$ Simple null

H_A Biased $p_h \neq 0.5$ 2-sided

Test

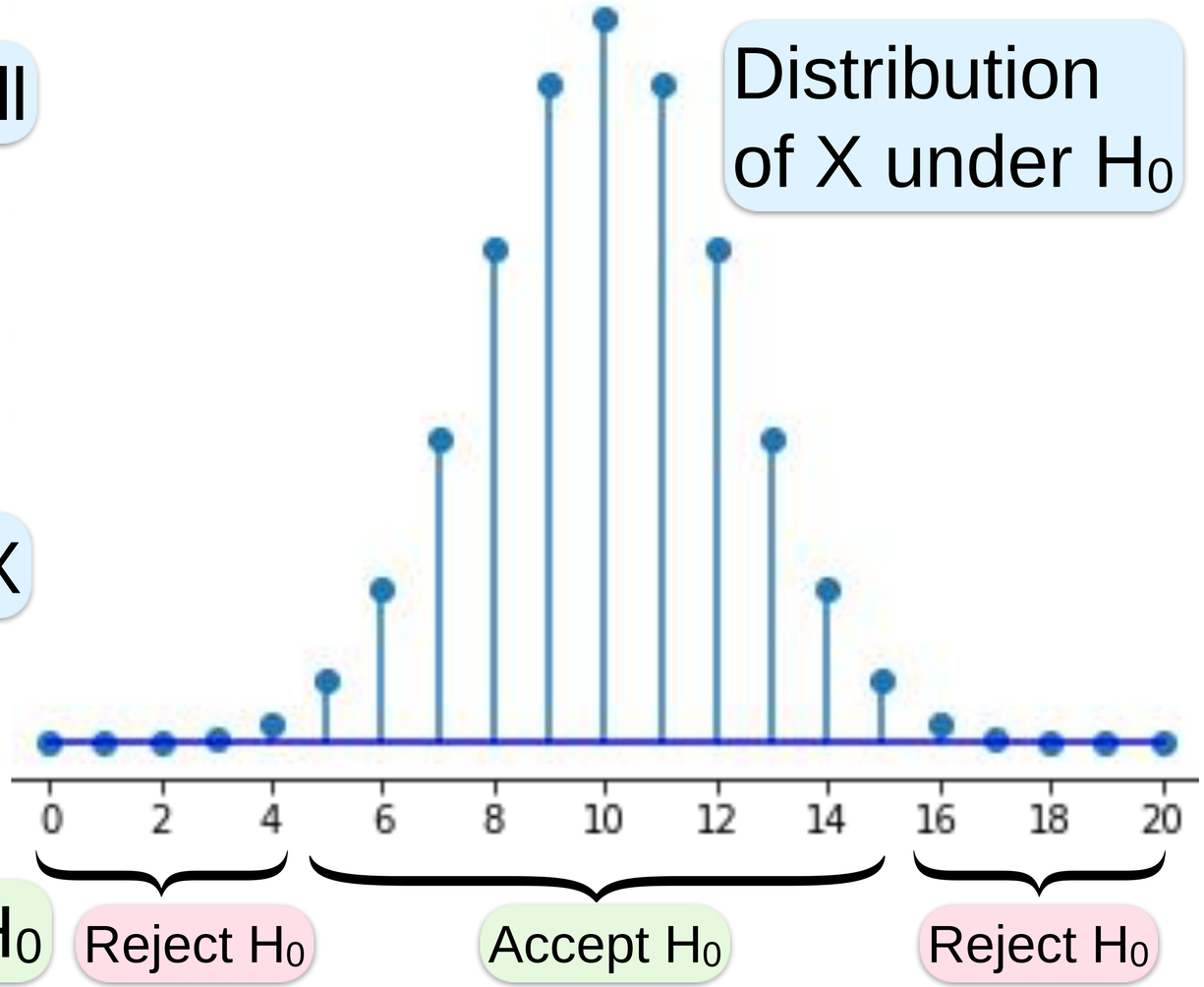
Test statistic Number of heads X

Intuitively

$5 \leq X \leq 15$ Do not reject H_0 Accept H_0

Otherwise Reject H_0 in favor of H_A

Distribution of X under H_0



Hypothesis Testing

Introduction

Assumptions about parameters

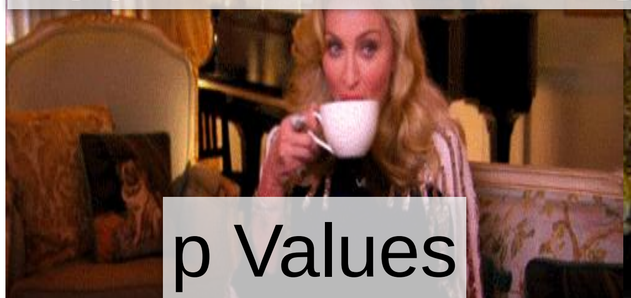
Simple & composite

Null & Alternative

Test

Accept or reject null

Hypothesis Testing



p Values