Alternative Hypothesis

Statistics and Big Data

Niccolò Salvini, PhD

UCSC

Academic Year 2025-2026

Course: Statistics and Big Data

Overview

The Three Pillars of a Statistical Test

To evaluate our hypotheses, we employ statistical tests. A statistical test requires three essential components:

1. **Data**: The observations collected from the experiment.
2. **Null Hypothesis (H_0)**: The hypothesis stating no effect or no difference. It always contains a statement of equality (e.g., $\mu = 100$, or $\mu_1 = \mu_2$).
3. **Alternative Hypothesis (H_1)**: The hypothesis that contradicts the null hypothesis. This is what we are trying to find evidence for.

The Three Pillars of a Statistical Test

To evaluate our hypotheses, we employ statistical tests. A statistical test requires three essential components:

1. **Data**: The observations collected from the experiment.
2. **Null Hypothesis (H_0)**: The hypothesis stating no effect or no difference. It always contains a statement of equality (e.g., $\mu = 100$, or $\mu_1 = \mu_2$).
3. **Alternative Hypothesis (H_1)**: The hypothesis that contradicts the null hypothesis. This is what we are trying to find evidence for.

The *form* of the alternative hypothesis is crucial because it defines what counts as "evidence against H_0".

Types of Alternative Hypotheses

The alternative hypothesis can be framed in different ways depending on the research question. Let's assume we are testing a population mean μ and our null hypothesis is $H_0 : \mu = \mu_0$.

Bilateral (Two-Tailed) Test

We are interested if the mean is **different** from μ_0, in either direction.

- $H_1 : \mu \neq \mu_0$

Example: "Is the average response time different from 12 minutes?"

Unilateral (One-Tailed) Test

We are interested if the mean is specifically **greater than** or **less than** μ_0.

- $H_1 : \mu > \mu_0$ (Right-tailed)
- $H_1 : \mu < \mu_0$ (Left-tailed)

Example: "Is the new drug **better** (i.e., recovery time is **less** than) the old one?"

Bilateral (Two-Tailed) Test: $H_1 : \mu \neq \mu_0$

In a two-tailed test, we reject H_0 if our sample statistic is unusually large **or** unusually small. The significance level α is split between the two tails.

Decision Rule

We reject H_0 if the evidence is strongly in favor of the mean being either significantly larger or significantly smaller than what is stated in H_0.

Unilateral (Upper-Tailed) Test: $H_1 : \mu > \mu_0$

In an upper-tailed test, we only reject H_0 if our sample statistic is unusually large. The entire significance level α is in the right tail.

Decision Rule

We reject H_0 only if the evidence is strongly in favor of the mean being significantly larger than what is stated in H_0. (A similar logic applies to a lower-tailed test).

Expanding to Multiple Groups

What happens when we introduce a third drug, E, and want to compare C, D, and E?

Expanding to Multiple Groups

What happens when we introduce a third drug, E, and want to compare C, D, and E? Now, our null hypothesis is that there is no difference among any of the drugs.

Null Hypothesis for k Groups

$H_0 : \mu_C = \mu_D = \mu_E$

Expanding to Multiple Groups

What happens when we introduce a third drug, E, and want to compare C, D, and E? Now, our null hypothesis is that there is no difference among any of the drugs.

Null Hypothesis for k Groups

$H_0 : \mu_C = \mu_D = \mu_E$

The alternative hypothesis becomes more complex.

Alternative Hypothesis for k Groups

H_1 : Not all population means are equal.

Expanding to Multiple Groups

What happens when we introduce a third drug, E, and want to compare C, D, and E? Now, our null hypothesis is that there is no difference among any of the drugs.

Null Hypothesis for k Groups

$H_0 : \mu_C = \mu_D = \mu_E$

The alternative hypothesis becomes more complex.

Alternative Hypothesis for k Groups

H_1 : Not all population means are equal.

This could mean many things: $\mu_C \neq \mu_D = \mu_E$, or $\mu_C = \mu_D \neq \mu_E$, or all three are different.

The Tool for Multiple Groups: ANOVA

Testing multiple pairs of means one-by-one with standard tests (like a t-test) is incorrect and inflates the Type I error rate.

The Tool for Multiple Groups: ANOVA

Testing multiple pairs of means one-by-one with standard tests (like a t-test) is incorrect and inflates the Type I error rate.

Correct Procedure

When comparing the means of three or more groups, we must use a dedicated statistical method called **Analysis of Variance (ANOVA)**.

The Tool for Multiple Groups: ANOVA

Testing multiple pairs of means one-by-one with standard tests (like a t-test) is incorrect and inflates the Type I error rate.

Correct Procedure

When comparing the means of three or more groups, we must use a dedicated statistical method called **Analysis of Variance (ANOVA)**.

- ANOVA tests the null hypothesis $H_0 : \mu_1 = \mu_2 = \cdots = \mu_k$.
- It does this by comparing the variation *between* the groups to the variation *within* the groups.
- If the variation between groups is significantly larger than the variation within them, we reject H_0.

This provides a single test for the overall hypothesis, controlling our α level.

Summary of Key Concepts

1. The **Alternative Hypothesis (H_1)** is driven by the research question and determines the structure of the test.
2. It can be **bilateral (two-tailed)**, testing for any difference ($\neq$).
3. Or it can be **unilateral (one-tailed)**, testing for a specific direction of difference ($>$ or $<$).
4. When comparing more than two groups, the alternative is "not all means are equal", and the correct tool is **ANOVA**.

Exercises

For each scenario, define the null (H_0) and alternative (H_1) hypotheses. State whether the test should be one-tailed or two-tailed.

Exercise 1

A factory manager wants to know if a new production process has changed the average number of defects per day, which was historically 15.

Exercises

For each scenario, define the null (H_0) and alternative (H_1) hypotheses. State whether the test should be one-tailed or two-tailed.

Exercise 1

A factory manager wants to know if a new production process has changed the average number of defects per day, which was historically 15.

Exercise 2

A researcher believes a new teaching method will **Increase** student test scores compared to the old average of 75.

Exercises

For each scenario, define the null (H_0) and alternative (H_1) hypotheses. State whether the test should be one-tailed or two-tailed.

Exercise 1

A factory manager wants to know if a new production process has changed the average number of defects per day, which was historically 15.

Exercise 2

A researcher believes a new teaching method will **Increase** student test scores compared to the old average of 75.

Exercise 3

An environmental agency is testing if the concentration of a pollutant in a river is **below** the safe limit of 50 ppm.