7.1 The Logic of Hypothesis Testing

Dr. Lauren Perry

Goals

1. Understand the logic of hypothesis testing.
2. Identify Type I and Type II error from a set of hypotheses.

Broadly, our goal is to make decisions about the value of a parameter.

We have confidence intervals, but we might also want to ask questions like

- Do cans of soda actually contain 12 oz?
- Is Medicine A better than Medicine B?

Hypotheses

A **hypothesis** is a statement that something is true.

A hypothesis test involves two (competing) hypotheses:

1. The **null hypothesis**, denoted H_0, is the hypothesis to be tested. This is the "default" assumption.
2. The **alternative hypothesis**, denoted H_A is the alternative to the null.

A **hypothesis test** helps us decide whether the null hypothesis should be rejected in favor of the alternative.

Example

Cans of soda are labeled with "12 FL OZ". Is this accurate?

The default, or uninteresting, assumption is that cans of soda contain 12 oz.

- H_0: the mean volume of soda in a can is 12 oz.
- H_A: the mean volume of soda in a can is NOT 12 oz.

We can write these hypotheses in words or in statistical notation.

Statistical Notation

The null specifies a single value of μ

- H_0: $\mu = \mu_0$

We call μ_0 the **null value**. When we run a hypothesis test, μ_0 will be replaced by some number.

The alternative specifies a *range* of possible values for μ:

- H_A: $\mu \neq \mu_0$. "The true mean is different from the null value."

Example

- H_0: the mean volume of soda in a can is 12 oz.
  - The null value is 12. In statistical notation, $H_0 : \mu = 12$.
- H_A: the mean volume of soda in a can is NOT 12 oz.
  - In statistical notation, $H_A : \mu \neq 12$.

The Logic of Hypothesis Testing

- Take a random sample from the population.
- If the data area consistent with the null hypothesis, do not reject the null hypothesis.
- If the data are inconsistent with the null hypothesis *and* supportive of the alternative hypothesis, reject the null in favor of the alternative.

Example: Jury Trials

In the US court system, jurors are told to assume the defendant is "innocent until proven guilty".

Innocence is the default assumption, so

- H_0: the defendant is innocent.
- H_A: the defendant is guilty.

Example: Jury Trials

- It is not the jury's job to decide if the defendant is innocent!
  - That should be their default assumption.
- They are *only* there to decide if the defendant is guilty or if there is not enough evidence to override that default assumption.

The *burden of proof* lies on the alternative hypothesis.

Notice the careful language in the logic of hypothesis testing: we either reject, or fail to reject, the null hypothesis.

We never "accept" a null hypothesis.

Decision Errors

- A **Type I Error** is rejecting the null when it is true. (Null is true, but we conclude null is false.)
- A **Type II Error** is not rejecting the null when it is false. (Null is false, but we do not conclude it is false.)

H_0 is

True

False

Decision

Do not reject H_0

Correct decision

Type II Error

Reject H_0

Type I Error

Correct decision

Example

In our jury trial,

- H_0: the defendant is innocent.
- H_A: the defendant is guilty.

A Type I error is concluding guilt when the defendant is innocent.

A Type II error is failing to convict when the person is guilty.

How likely are we to make errors?

$P(\text{Type I Error}) = \alpha$, the **significance level**.

- This is the same α we saw in confidence intervals!

$P(\text{Type II Error}) = \beta$.

- This is something we don't have time to cover in detail.

We would like both α and β to be small but,

- If we decrease α, then β will increase.
- If we increase α, then β will decrease.

In practice, we set α (as we did in confidence intervals).

We can improve β by increasing sample size.

Example

Consider two possible criminal charges:

1. Defendant is accused of stealing a loaf of bread. If found guilty, they may face some jail time and will have a criminal record.
2. Defendant is accused of murder. If found guilty, they will have a felony and may spend decades in prison.

Since these are moral questions, I will let you consider the consequences of each type of error.

However, keep in mind that we do make scientific decisions that have lasting impacts on people's lives.

Hypothesis Test Conclusions

- If the null hypothesis is rejected, we say the result is **statistically significant**. We can interpret this result with:
  - At the α level of significance, the data provide sufficient evidence to support the alternative hypothesis.
- If the null hypothesis is *not* rejected, we say the result is **not statistically significant**. We can interpret this result with:
  - At the α level of significance, the data do *not* provide sufficient evidence to support the alternative hypothesis.

When we write these types of conclusions, we will write them in the context of the problem.