Graduate Statistics

Hypothesis Testing

What We Will Cover in This Section

- Introduction.
- Logic of hypothesis testing.
- Errors.
- Types of hypothesis tests.
- Tails of Tales.

What is hypothesis testing?

A set of logical and statistical guidelines used to make inferential decisions from sample statistics to population characteristics.

Types of Hypotheses

- Research hypothesis.
- Logical hypotheses.
  - Null hypothesis (H_o).
  - Alternative hypothesis (H_a).
- Statistical hypothesis.

Research Hypothesis

Statement in words as to what the investigator expects to find.

Example.

Students who drink caffeine will be able to memorize information faster than students who do not drink caffeine.

Logical Hypotheses

Null Hypothesis (H_o).
Statement that the treatment does not have the expected effect.

Alternative Hypothesis (H_a).
Statement that the treatment had the expected effect.

Characteristics of the Logical Hypotheses

1. They are mutually exclusive.
2. They are mutually exhaustive.

Caffeine Example

- H_o
  Students who drink caffeine will not be able to memorize information faster than people who do not drink caffeine.
  - Non-caffeine and caffeine drinkers are the same.
  - Non-caffeine drinkers are faster.
- H_a
  Students who drink caffeine will memorize information faster than those who do not drink caffeine.

Statistical Hypotheses

Statement in statistical (mathematical) terms as to what would be found if the alternative hypothesis is true.

Example.

$M_a > M_b$

$M_a > M_b$ or $M_a < M_b$

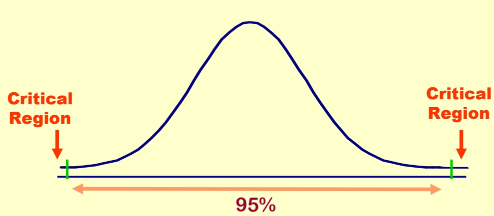
Caffeine Example, AGAIN!

H_a: $M_{caffeine} < M_{no\ caffeine}$

Decision Making Criteria

1. Since we are dealing with sampling error there is always a possibility that differences we find could have happened by chance.
2. We make statistical inferences based on the probability that the results <u>may or may not</u> have <u>happened by chance</u>.
3. Our model for making this decision is founded on the normal distribution.

How the Decision Works



Decision Steps

1. We start by assuming that the Null Hypothesis is true.
2. When a statistical result is rare we conclude that it probably did not happen by chance.
3. If we conclude that a result did not happen by chance (e.g. it is rare), we reject H_o.
4. The only option is to conclude that the true state of affairs is represented by H_a.

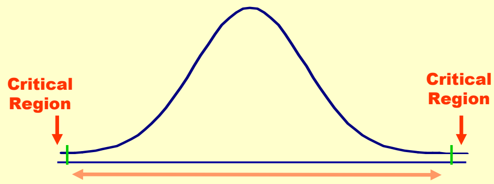
ERRORS

Outcomes of the Statistical Decision

		Actual Situation	
		No Predicted Effect (H_o True)	Predicted Effect (H_o False)
Experimenter's Decision	Reject H_o Research hypothesis is correct	Type I Error	Correct Decision
	Retain H_o Research hypothesis is not correct.	Correct Decision	Type II Error

Alpha Level

The probability that a statistical test will lead to a Type I error.



A Tale of Tails, Review

One-tailed test.

Researcher makes a prediction that one group will be bigger than the other.

Two-tailed test.

Researcher makes a prediction that the groups will be different, but does not state which will be bigger.

A Tale of Tails, Review

Above what z-score do 5% of the cases fall? (α = .05)	1.64
Below what z-score do 1% of the scores fall? (α = .01)	-2.33
Between which two z-scores do 95% of the cases fall? (α = .05)	±1.96
Between which two z-scores do 99% of the cases fall? (α = .01)	±2.58

Key Learning Points #1

1. Science is conservative.
2. We assume that the research hypothesis is invalid until the evidence is so strong that we must conclude that it is true.
3. We statistically ‘test’ the assumption that the research hypothesis is not true.
4. If the data are so strong that we believe that they could not have happened by chance, then we reject H_o.

Key Learning Points #2

5. Since our decisions are based on probability theory not absolute surety, we can make mistakes.
6. The probability of concluding that the research hypothesis is correct when it isn’t (rejecting Ho when it is true) is represented by alpha (α).
7. The probability of failing to find a result when there is one is represented by beta (β).

The End