

Psychological 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 decisions from sample statistics to population characteristics.



Types of Hypotheses

- Research hypothesis.
- Logical hypotheses.
 - Null hypothesis (H_0).
 - Alternative hypothesis (H_a).
- Statistical hypothesis.

9/29/2005

P331 Hypothesis Testing

4



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.

9/29/2005

P331 Hypothesis Testing

5



Logical Hypotheses

Null Hypothesis (H_0).

Statement that the treatment does not have the expected effect.

Alternative Hypothesis (H_a).

Statement that the treatment had the expected effect.

9/29/2005

P331 Hypothesis Testing

6



Characteristics of the Logical Hypotheses

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

9/29/2005

P331 Hypothesis Testing

7



Caffeine Example

- H_0
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.

9/29/2005

P331 Hypothesis Testing

8



Statistical Hypotheses

Statement in statistical terms as to what would be found if the research hypothesis is true.

Example.

$$M_a > M_b \text{ (one tail)}$$

$$M_a > M_b \text{ or } M_a < M_b \text{ (two tail)}$$

9/29/2005

P331 Hypothesis Testing

9



Caffeine Example, AGAIN!

$$H_a: M_{\text{caffeine}} > M_{\text{no caffeine}}$$

$$H_o: M_{\text{caffeine}} = M_{\text{no caffeine}}$$

or

$$M_{\text{caffeine}} < M_{\text{no caffeine}}$$

9/29/2005

P331 Hypothesis Testing

10



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 may or may not have happened by chance.
3. Our model for making this decision is founded on the normal distribution.

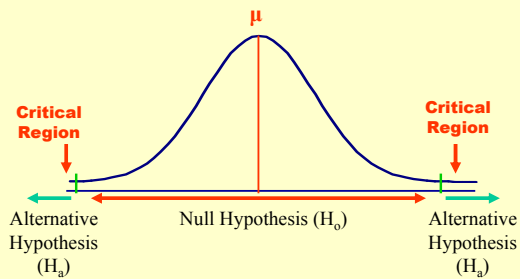
9/29/2005

P331 Hypothesis Testing

11



How the Decision Works



9/29/2005

P331 Hypothesis Testing

12



Decision Steps

1. We start by **assuming that the Null Hypothesis is true**.
2. When a statistical **result is rare** (less than 5% or 1% of the time) 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_0** .
4. The only option is to conclude that the true state of affairs is represented by H_a .

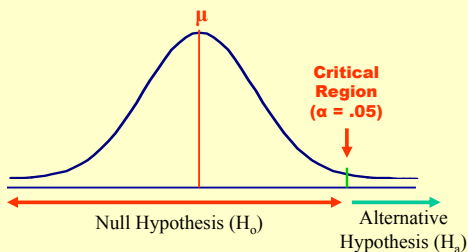
9/29/2005

P331 Hypothesis Testing

13



One-tail Example



9/29/2005

P331 Hypothesis Testing

14



FOOF

ERRORS





9/29/2005

P331 Hypothesis Testing

15



Outcomes of the Statistical Decision

		Actual Situation	
		No Predicted Effect (H_0 True)	Predicted Effect (H_0 False)
Experimenter's Decision	Reject H_0 Research hypothesis is correct	Type I Error 	Correct Decision 
	Retain H_0 Research hypothesis is not correct.	Correct Decision 	Type II Error 

9/29/2005

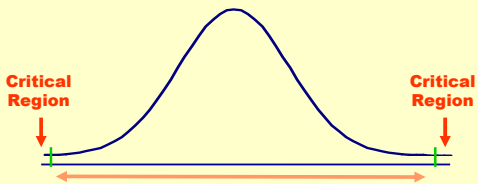
P331 Hypothesis Testing

16



A Two-tailed Test

The probability that a two-tailed test will lead to a Type I error.



9/29/2005

P331 Hypothesis Testing

17



Practice



9/29/2005

P331 Hypothesis Testing

18



Z-Test Again

Rudy Beggah thought that a diet that consisted of root vegetables would help high school football players gain weight. Beggah knew that the national average weight for a football player was $\mu = 167$ and $\sigma = 12$.

1. What is the research hypothesis?
2. What is the null hypothesis?
3. What is the alternative hypothesis?
4. What is the statistical hypothesis?
5. Is this a one-tail or two-tail test?
6. What alpha level should Rudy set?

9/29/2005

P331 Hypothesis Testing

19



The Results

Rudy got a sample of 25 football players and supplemented their diet with 2 pounds of root vegetables a day. At the end of six weeks this group's weight was $M=175$.

$$\sigma_{\bar{x}} = \frac{12}{\sqrt{25}} = \frac{12}{5} = 2.4$$

$$Z = \frac{175 - 167}{2.4} = 3.33$$

Did the diet help?

9/29/2005

P331 Hypothesis Testing

20



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_0 .

9/29/2005

P331 Hypothesis Testing

21



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 H_0 when it is true) is represented by alpha (α).
7. The probability of failing to find a result when there is one is a Type II error (β).

9/29/2005

P331 Hypothesis Testing

22



The End

9/29/2005

P331 Hypothesis Testing

23

t-tests