Statistics
an Overview

What We Will Cover in This Section

- What statistics are.
- Descriptive Statistics
  - Frequency distributions
  - Graphs
  - Mean
  - Standard deviation
- Inferential Statistics
  - Z-scores
- Hypothesis Testing

Basic Terms and Concepts

Basic Terminology

STATISTICS

Numerical techniques for describing groups of people or events.

Fundamental Uses

DESCRIPTIVE STATISTICS

Techniques used to organize, summarize, and describe sets of numbers.

INFERENTIAL STATISTICS

Techniques that allow us to make estimates about POPULATIONS based on SAMPLE data.

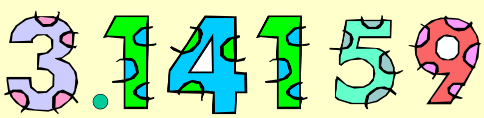
Population vs. Sample

- **Population**
  - ALL members of a group that are alike on some characteristic.
  - Infinite in size.
  - Parameters are indicated by Greek letters: μ σ
- **Sample**
  - A subset of a population.
  - Finite in size.
  - Numerical estimators are called *statistics*.
  - Statistics are indicated by Roman letters: M, S.

Why Used?

- We cannot collect data from populations.
- We collect data from samples.
- On the basis of the numerical characteristics of samples we try to make conclusions about populations.

Using Numbers



Levels of Measurement

NOMINAL SCALE

Numbers are used as labels.

ORDINAL SCALE

Numbers are used to indicate rank order.

Levels of Measurement

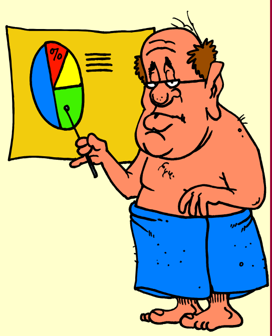
INTERVAL SCALE

Numbers are used to indicate an actual amount and there is an equal unit of measurement between adjacent numbers.

RATIO SCALE

Numbers indicate an actual amount and there is a true zero.

Frequency Tables and Graphs



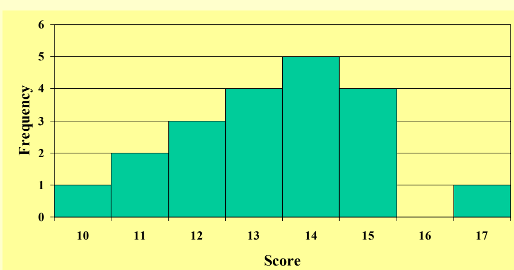
Common Statistics

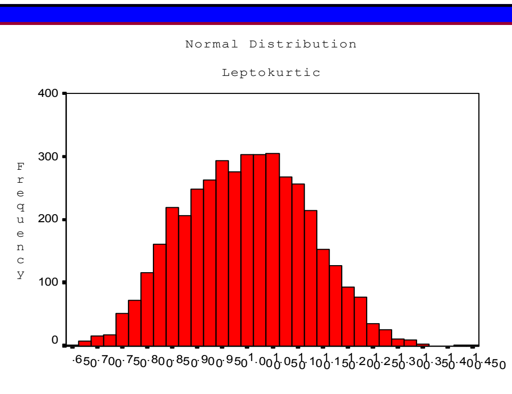
- **Frequency**
  - The number of people who got a certain score.
  - Symbolized with f.
- **Number**
  - The total count of observations in a sample.
  - Symbolized with N.
- **Percent**
  - The number in a group (f) divided by the total number (N).
- **Percentile**
  - The percent of people who got a score <u>and</u> lower.
  - Symbolized with P_n.

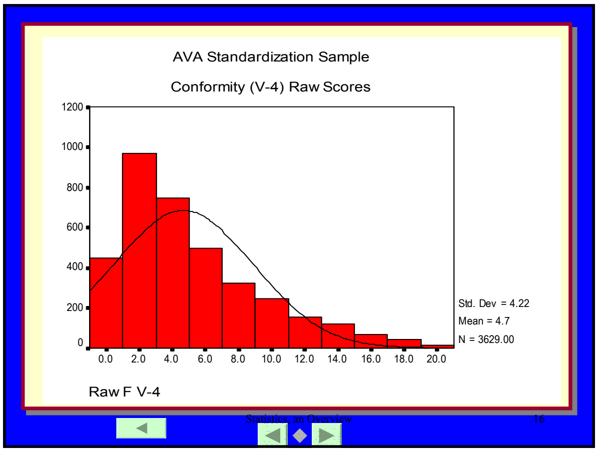
Simple frequency distribution (N=20)

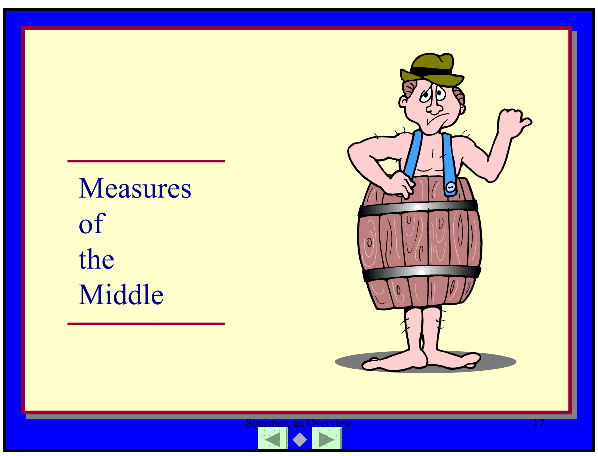
Score	Frequency (f)	%	Cum % **(Percentile)**
17	1	5	100
16	0	0	95
15	4	20	95
14	5	25	75
13	4	20	50
12	3	15	30
11	2	10	15
10	1	5	5

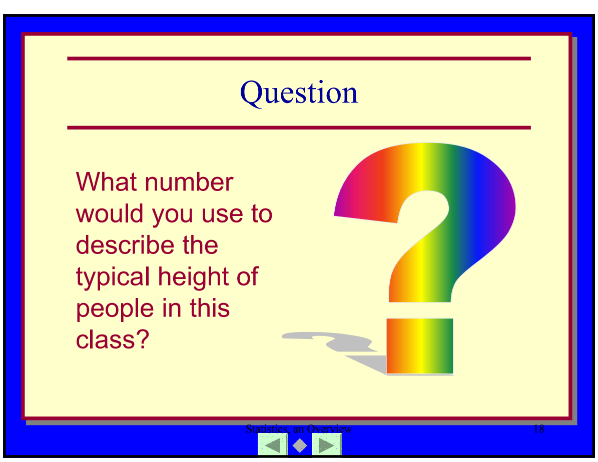
Histogram






AVA Standardization Sample
Conformity (V-4) Raw Scores
1200
1000
800
600
400
200
0
0.0
2.0
4.0
6.0
8.0
10.0
12.0
14.0
16.0
18.0
20.0
Std. Dev = 4.22
Mean = 4.7
N = 3629.00
Raw F V-4
Statistics, an Overview
16


Measures
of
the
Middle
Statistics, an Overview
17


Question
What number would you use to describe the typical height of people in this class?
Statistics, an Overview
18

Mean

- Sum the scores and divide by the number of scores.
- Symbols
  - Sample: M or $\overline{X}$

Median

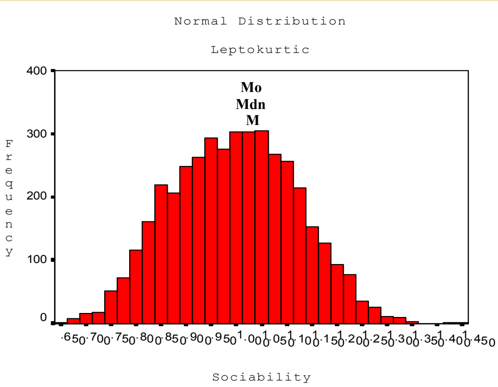
- The score below which 50% of the scores fall.
- Represents P_{50}.
- Divides the distribution in half.
- Symbol: Mdn

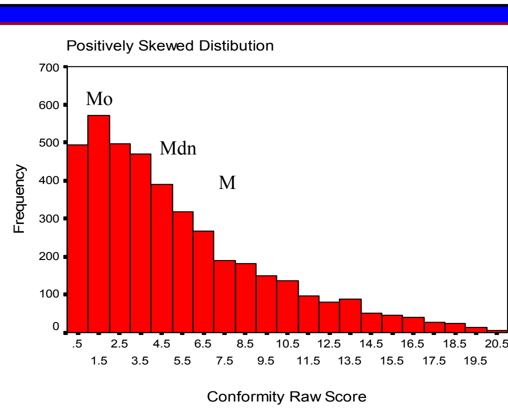
Example

8	**9**	**10**	**11**	**12**	**13**	**14**	**15**	**16**
8	**9**	**10**	**11**	**12**	**13**	**16**	**16**	**46**

Mode

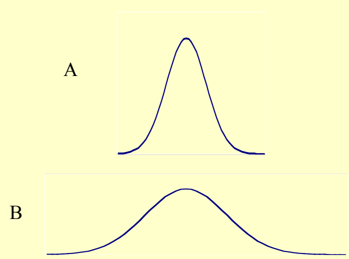
- The score that occurs most frequently in a distribution.
- Used for nominal scales or higher.
- Symbol: Mo





Measures of Variability

Two Normal Distributions with the Same Mean



Overview

The Mean describes the ‘typical score’; measures of variability give an index of how much the rest of the scores in the sample are spread out around the mean.

Range

- The distance between the lowest and highest score.
- Formula

  Range = Highest Score – Lowest Score
- Example

1 3 4 6 8 12 15 16 18 19

1 3 4 6 8 12 15 16 18 79

Deviation Score

	Score	X - M_x	$(X - Mx)^2$
	5	-2.5	6.25
	6	-1.5	2.25
	7	-.5	.25
	8	.5	.25
	9	1.5	2.25
	10	2.5	6.25
Sum	45	0	17.50
Mean	7.5	0	2.92

Variance

Variance (S^2 or σ^2)

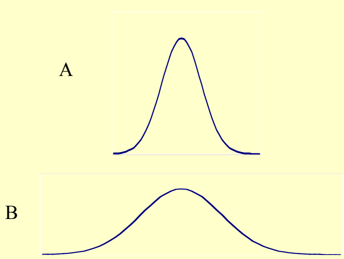
Mean <u>squared</u> deviation score around the mean.

Standard Deviation (S or σ)

Square root of
the variance.

$$\sqrt{2.92} = 1.71$$

Large and Small Standard Deviations



Key Learning Points

- Most behavioral characteristics are normally distributed.
- The ***Mean*** represents the 'typical' score for a sample.
- The ***Variance*** and ***Standard Deviation*** measure the variability of scores in a sample.

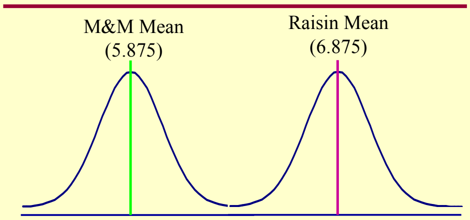
Inferential Statistics

Issue

- In research we collect data from SAMPLES and try to generalize those results to POPULATIONS.
- To what degree does a sample mean *represent* the population to which we want to make inferences?
- Does this sample represent population A or does it represent population B?

	Raisin	M&M
	6	5
	7	8
	5	9
	8	6
	8	6
	3	4
	11	4
	7	5
Number	8	8
Sum	55	47
Mean	6.875	5.875
Variance	5.554	3.268
Standard Deviation	2.357	1.808

Underlying Concept

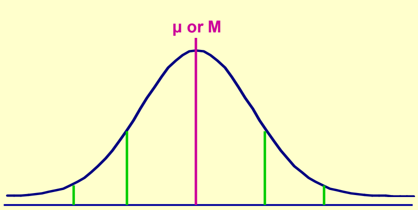


Question: Do these two samples represent the same population or do they represent two different populations?

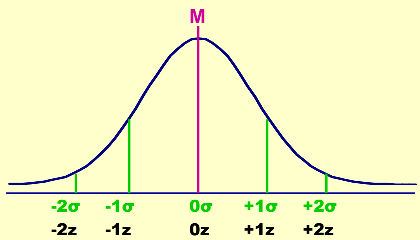
What is Rare?

Means that are so far apart that we conclude that the distance between them could not have happened by chance.

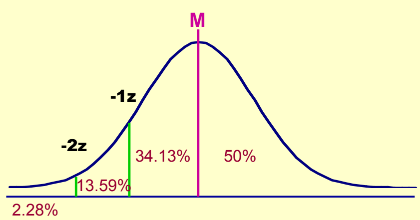
Units of Measurement



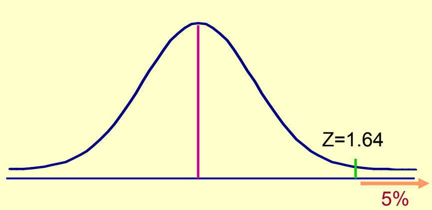
Units of Measurement



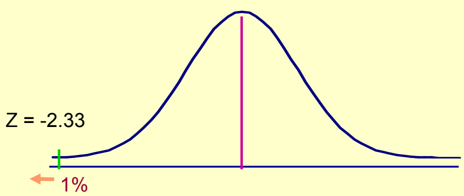
Slicing the Normal Curve



Key Unit #1



Key Unit #2



What Is Rare?

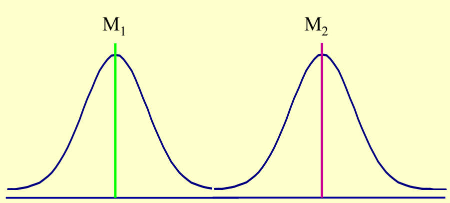
- Some event that has a low probability of happening.
- In research we choose this ‘rare’ value.
- Typically it is set at 5% (.05) or less.
- Any event that occurs 5% of the time or less is considered to be rare.
- Indicated by: $p < .05$

Applications

How Inferential Statistics are Used

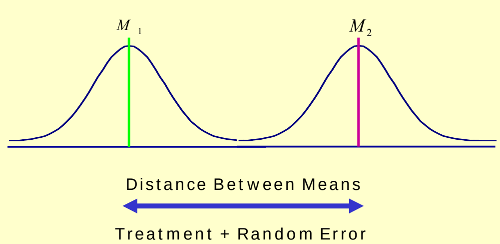
1. When we want to know if the scores for two groups are different.
   - t-test
   - Analysis of Variance (ANOVA)
2. When we want to see if there is a relationship between scores.
   - Correlation coefficient

The t-test

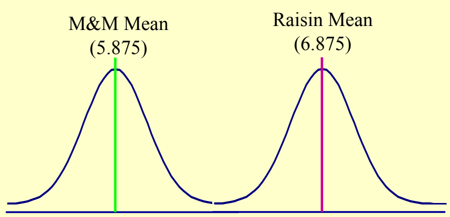


Question: Do these two samples represent the same population or do they represent two different populations?

t-test Logic



Remember the Memory Study?



t-Test

$t_{(14)} = .95,\ p < .36$

t	The name of the statistic
(14)	Degrees of freedom (df). Two less than the number of people in the study.
.95	The calculated value for t. It ranges from 0 to large. It is possible to have negative values for t.
p<.36	An indicator as to how rare this value is. It indicates the number of times out of 100 you would get this difference between means based on the sample size.

Another Situation

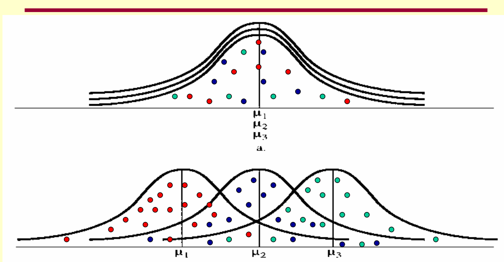
The management of Sal T. Dogg's restaurant wanted to see if the saltiness of appetizers would influence the number of drinks people purchased. Three sections of the club are targeted to receive appetizers that have either low, medium, or high saltiness. The dependent variable is the number of drinks ordered.

Appetizer saltiness and number or drinks ordered.		
Group 1 Low Salt	Group 2 Medium Salt	Group 3 High Salt
2	3	3
3	4	1
1	4	2
1	5	2
2	6	3
2	4	1
1	3	1
2	2	2
2	4	1
4	4	2
M = 2.00	M = 3.90	M = 1.80

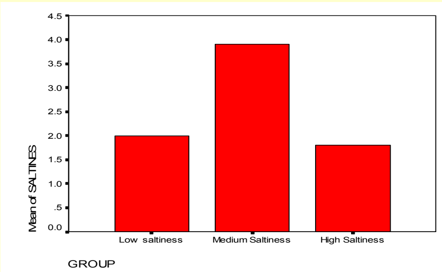
Issue

How to determine if one mean is significantly different from the other means while minimizing the probability of committing a Type I error.

Analysis of Variance (ANOVA)



Graph of Saltiness Ratings



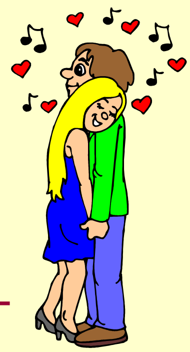
ANOVA Summary Table

Source	SS	df	MS	$F_{(crit=3.35)}$
Between Groups	26.87	2	13.435	14.77
Within Groups	24.50	27	.91	
Total	51.37	29		

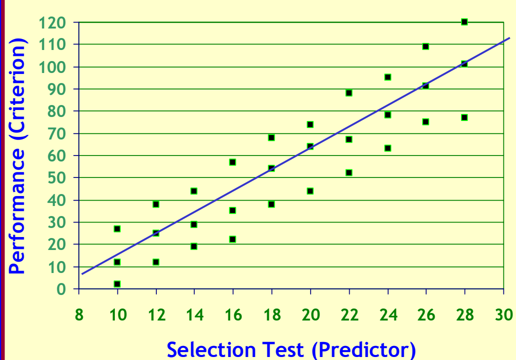
How to Express F

$F_{(2,27)} = 14.77, \; p<.05$

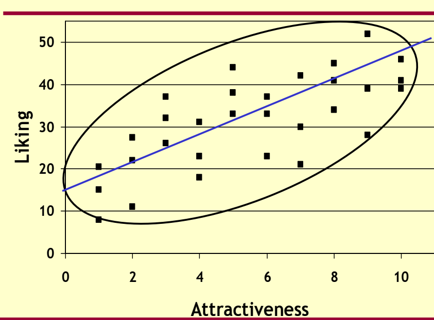
- Degrees of freedom (between, within)
- Calculated F
- Rare?

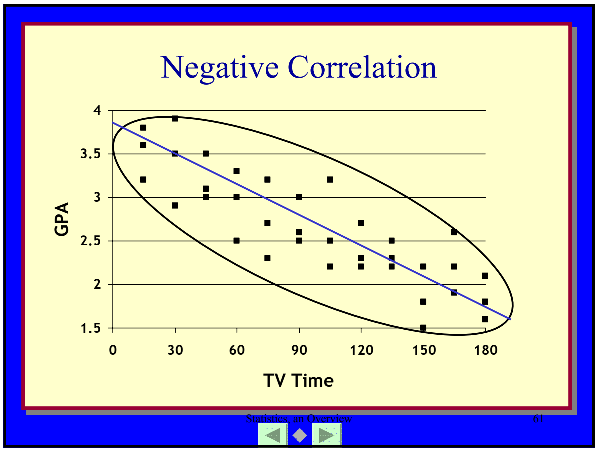


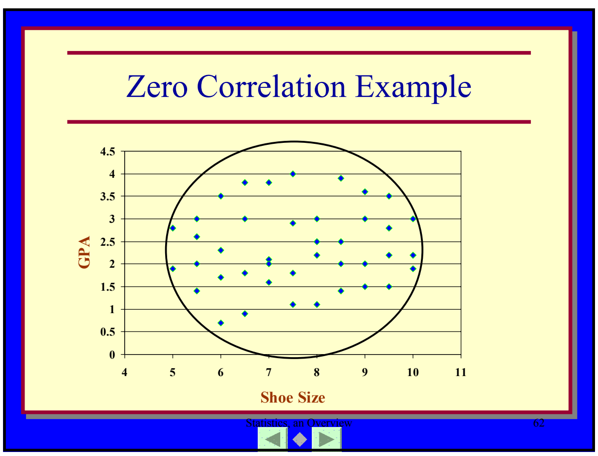
Correlation:
Measuring Relationships

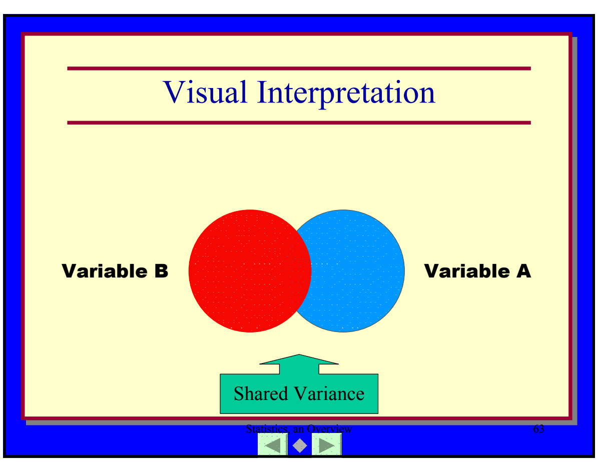


Positive Correlation

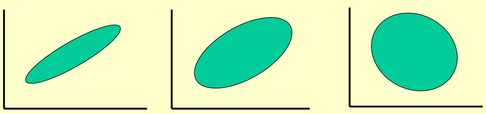



Negative Correlation
GPA
TV Time
4
3.5
3
2.5
2
1.5
0
30
60
90
120
150
180
Statistics: an Overview
61


Zero Correlation Example
GPA
Shoe Size
4.5
4
3.5
3
2.5
2
1.5
1
0.5
0
4
5
6
7
8
9
10
11
Statistics: an Overview
62


Visual Interpretation
Variable B
Variable A
Shared Variance
Statistics: an Overview
63

Relationship Strength



Strong **Moderate** **Weak**

What Correlations Look Like

	Competence (1)	Self Confidence	Recklessness(1)	Competence (2)	Recklessness(2)
Competence(1)	1	-.485(**)	.433(**)	.947(**)	.398(**)
Self Confidence	-.485(**)	1	-.054	-.480(**)	-.013
Recklessness(1)	.433(**)	-.054	1	.415(**)	.986(**)
Competence(2)	.947(**)	-.480(**)	.415(**)	1	.369(**)
Recklessness(2)	.398(**)	-.013	.986(**)	.369(**)	1

** Correlation is significant at the 0.01 level (2-tailed).
* Correlation is significant at the 0.05 level (2-tailed).

Power and Effect Size

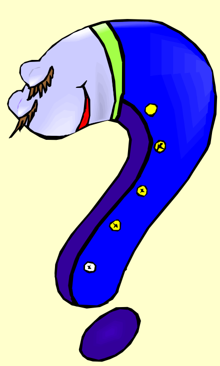
Effect Size

1. How strong was the treatment?
2. How strong is the relationship?

Power

The ability of the statistical procedure to detect the effect being measured.

Hypotheses



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.

How they fit together

Research hypothesis.

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

How They Fit Together #2

- 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.

Caffeine Example, AGAIN!

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

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

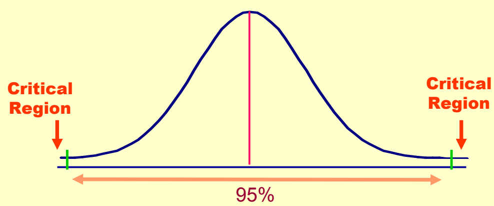
or

$M_{caffeine} > M_{no\ affeine}$

Decision Making Criteria

1. We make statistical inferences based on the probability that the results may or may not have happened by chance.
2. Since we are dealing with sampling error there is always a possibility that data we collect could have happened by chance.
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 (H_o) 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.

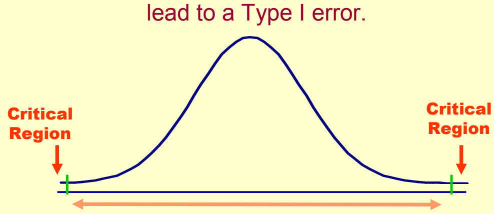
ERRORS

Outcomes of the Statistical Decision

		Actual Situation	
		No Predicted Effect (H_o True)	Predicted Effect (Ho False)
Experimenter's Decision	Reject H_o	**Type I Error**	**Correct Decision**
	Retain H_o	**Correct Decision**	**Type II Error**

Alpha Level

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



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).