

Name: **KEY** Official Class: Date:
Teacher: Period: Class:

Statistics Review Sheet

Vocabulary

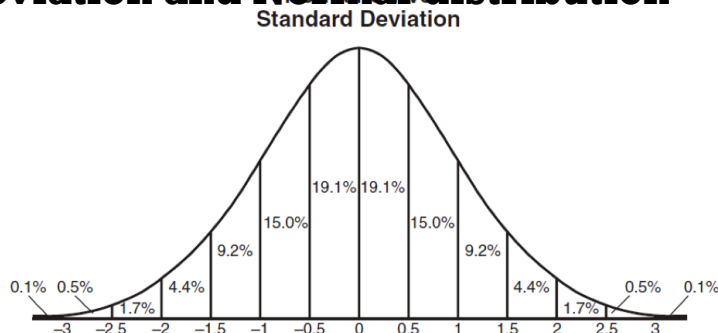
Directions: Using the word box, define all of the words

Word Bank:

Normal Distribution Variable	Experimental Condition Independent Variable Sampling Error	Population Statistics Control Condition	Data Dependent Variable
---------------------------------	--	---	----------------------------

- Data:** measurements or observations (aka scores)
- Variable:** A characteristic or condition which has different values for different individuals (ex. height, test scores, gender)
- Independent variable:** The variable that is controlled by an experimenter
- Dependent variable:** A variable that is allowed to vary and is observed in relation to the IV (dependent on the independent variable)
- Statistics:** A set of calculations used to organize, summarize and interpret info
- Population:** ALL of the individual you wish to study (ex. all students in the US)
- Sample:** ONLY SOME of the individuals/objects you wish to study from a population (ex. 1000 students from New York)
- Sampling Error:** A discrepancy which occurs between a sample and its population
- Control condition:** Individuals in this type of experimentation are given no experimental treatment or are given a type of placebo (This condition is used to have a base of reference for the experimental group)
- Experimental condition:** Individuals in this type of experimentation do receive the treatment being tested
- Normal Distribution:** This type of distribution is seen when the variables are clustered together with gradual decrease on either side of the distribution. This type of distribution is used often in calculations assuming a normal population distribution.

Standard Deviation and Normal distribution



Directions: Use the diagram below to answer the following questions.

- In the United States, the mean height of adult males is 69.2 inches. Translate this height to feet and inches.

5 feet and 9.2 inches

Name: **KEY** Official Class: Date:

Teacher: Period: Class:

2. The heights of men in the United States are normally distributed with a standard deviation of 2.8 inches. What is the range of heights within 1σ of the mean?

$-3\sigma = 60.8$
 $-2.5\sigma = 62.2$
 $-2\sigma = 63.6$
 $-1.5\sigma = 65$
 $-1\sigma = 66.4$

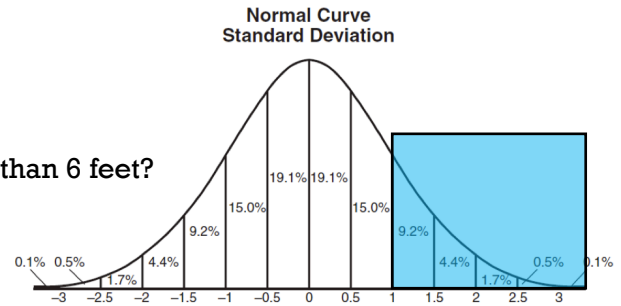
$-0.5\sigma = 67.8$
 $0\sigma = 69.2$ inches
 $0.5\sigma = 70.6$
 $1\sigma = 72$
 $1.5\sigma = 73.4$

$2\sigma = 74.8$
 $2.5\sigma = 76.2$
 $3\sigma = 77.6$

66.4 inches to 72 inches

3. Approximately what percent of adult U.S. males are taller than 6 feet?

Add percentages in highlighted section on curve
 $9.2 + 4.4 + 1.7 + 0.5 + 0.1 = 15.7\%$

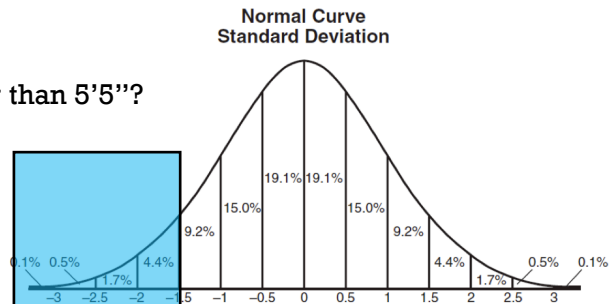


4. In a sample of 50,000 men, approximately how many would be taller than 6 feet?

15.7 % of 50,000 so $(0.157 \times 50000) = 7850$ men

5. Approximately what percent of U.S. males are shorter than 5'5"?

Add percentages in highlighted section on curve
 $4.4 + 1.7 + 0.5 + 0.1 = 6.5\%$



6. Out of every million men, approximately how many are shorter than 5'5"?

6.5% of 1 million so $(1,000,000 \times 0.065) = 65,000$ men

Type I and Type II Error

1. A machine fills milk bottles, the mean amount of milk in each bottle is supposed to be 32 Oz with a standard deviation of 0.06 Oz. Suppose the mean amount of milk is approximately normally distributed. To check if the machine is operating properly, 36 filled bottles will be chosen at random and the mean amount will be determined.

- a. What are the null and alternative hypotheses?

Null hypothesis: mean is equal to 32 ounces

Alternate hypothesis: mean is NOT equal to 32 ounces

- b. Make the table like the one above to show what and where your type I and type II errors are

Name: _____ KEY _____ Official Class: _____ Date: _____
Teacher: _____ Period: _____ Class: _____

		The Fact is...	
		H_0 True	H_A True
We say that	Reject H_0	Type I Error – come to the conclusion that the defendant is guilty when his is not	Correct Decision
	Fail to Reject H_0	Correct Decision	Type II Error – come to the conclusion that the defendant is not guilty when he really is guilty

c. Describe your type I error in context.

The fact is that the mean is 32 ounces but we say that it isn't.

d. Describe your type II error in context.

The fact is that the mean is not 32 ounces, but we say that it is.

2. A certain type of seed has always grown to a mean height of 8.5 inches, with a standard deviation of 1 inch. Based on past experiment, the mean height of a seed is known to be distributed approximately normal. A researcher wishes to find out whether some new enriched conditions would improve the mean height. He wants to use $\alpha = .01$ test and would like to have a 96% chance of rejecting the null hypothesis if the mean height is 9.5 inches.

a. What are the null and alternative hypotheses?

Null hypothesis: Mean height is 8.5/no effect on growth

Alternate hypothesis: Mean height is greater than 8.5/there is an effect on growth

b. Make the table like the one above to show what and where your type I and type II errors are

		The Fact is...	
		H_0 True	H_A True
We say that	Reject H_0	Type I Error – come to the conclusion that the defendant is guilty when his is not	Correct Decision
	Fail to Reject H_0	Correct Decision	Type II Error – come to the conclusion that the defendant is not guilty when he really is guilty

c. Describe your type I error in context.

The fact is that there is no effect on plant growth but we say that there isn't.

d. Describe your type II error in context.

The fact is that there is an effect on plant growth but we say that there isn't.

Correlation vs. Causation

1. If a data set has a correlation coefficient of 0.01, what does that mean?

No correlation

2. If a data set has a correlation coefficient of -1.0, what does that mean?

Perfect negative correlation

3. If a data set has a correlation coefficient of 0.20, what does that mean?

No correlation to weak positive correlation

Name: _____ **KEY** _____ Official Class: _____ Date: _____
Teacher: _____ Period: _____ Class: _____

4. If a data set has a correlation coefficient of -0.8 , what does that mean?

Strong negative correlation

5. If a data set has a correlation coefficient of 0.5 , what does that mean?

Weak positive correlation.

Null/Alternate Hypothesis and P-Values

Directions: Write the null and alternate hypothesis. Then determine if the data is statistically significant and how it relates to the null hypothesis of the specific experiment.

1. A study wants to determine the effects of caffeine on reaction time. The p-value is 0.03 .
Null Hypothesis: There is no effect of caffeine on reaction time
Alternate Hypothesis: There is an effect of caffeine on reaction time
The data is **statistically significant** and you **reject the null hypothesis**
2. A study wants to determine the effects of a drug on heart rate. The p-value is 0.00005 .
Null Hypothesis: There is no effect of the drug on heart rate
Alternate Hypothesis: There is an effect the drug on heart rate
The data is **statistically significant** and you **reject the null hypothesis**
3. A study wants to determine whether or not Blue Island is the best place to fish. The p-value is 0.50 .
Null Hypothesis: Blue Island is not the best place to fish
Alternate Hypothesis: Blue Island is the best place to fish
The data is **not statistically significant** and you **fail to reject the null hypothesis**
4. A study wants to compare the average number of free throws made by the varsity basketball team, and players in the NBA. The p-value is 0.973 .
Null Hypothesis: don't worry about this one
Alternate Hypothesis: don't worry about this one
The data is **not statistically significant** and you **fail to reject the null hypothesis**
5. A study want to determine if there is a link between the number of tennis balls a person can hit and the number of baseballs they can hit. The p-value is 0.05 .
Null Hypothesis: There is no link between tennis and baseball
Alternate Hypothesis: There is a link between tennis and baseball
Since the p-value is equal to 0.05 , it is up to the research to decide what to do with the dat

T-Tests

Directions: Determine the number type of t-test that should be used. Explain your reasoning.

1. A study wants to determine if the number of glasses of water a person drinks, is related to the number of hours that specific person sleeps.

A paired t-test because you are looking at how water effects a SPECIFIC person's sleeping

Name: _____ **KEY** _____ Official Class: _____ Date: _____
Teacher: _____ Period: _____ Class: _____

2. A study wants to determine what type of tires a driver should use on their own car to best handle the slippery wet conditions of rain and snow.

A paired t-test because a driver wants to compare his own specific car with different tires.

3. A study wants to determine if they have discovered a new melting point of calcium.

Fixed-value t-test because you are comparing to a known fact; the melting point of calcium.

4. A study wants to determine if which bag of M&Ms has more candy, the peanuts or the milk chocolate.

Unpaired t-test because you are comparing the two general groups of M&Ms

5. A study wants to determine if a person who likes salmon also likes cod.

A paired t-test because you want to look at if a specific person like one fish that they also like the other.