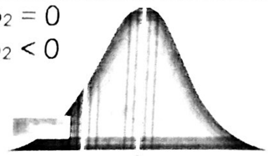
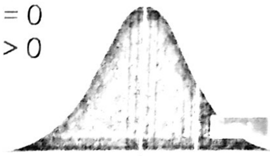
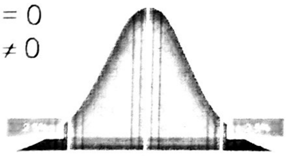
Statistical Reasoning
Hypothesis Tests key

Name: ______ Date: ______ Class: ______

Hypothesis Testing the Difference of Two Proportions

***State:* State null and alternative hypotheses and identify statistical data.**

- Hypothesis: Null hypothesis (H_0) and Alternative hypothesis (H_a)
- Statistical Data
  - Parameter: Population mean (μ) or proportion (p)
  - Standard deviation: Population std. dev. (σ) or sample std. dev. (s)
  - Significance level (α): the critical value where the null hypothesis may be rejected.
    - 10% (0.1), 5% (0.05), or 1% (0.01)
  - p_1 ***and*** p_1 — *proportions 1 and 2*
  - x_1 ***and*** x_1 — *number of successes from proportions 1 and 2*
  - n_1 ***and*** n_2 — *number of trials from proportions 1 and 2*

Decreased: Lower-tailed test	*Increased: Upper-tailed test*	*Changed: Two-tailed test*
H_0: $p_1 - p_2 = 0$ H_a: $p_1 - p_2 < 0$ 	H_0: $p_1 - p_2 = 0$ H_a: $p_1 - p_2 > 0$ 	H_0: $p_1 - p_2 = 0$ H_a: $p_1 - p_2 \neq 0$ 

***Plan*: Select the appropriate test and consider conditions.**

- Test statistic: formula and substitution—identify the appropriate test and substitute

Z- Test	*T- Test*	*1 Proportion*	*2 Proportions*
If population mean and population standard deviation is known (σ)	*If population mean and sample standard deviation is known* (s)	*If population proportion sample proportion is known*	*If 2 sample proportions are known*
$z = \frac{\bar{x} - \mu}{\frac{\sigma}{\sqrt{n}}}$	$z = \frac{\bar{x} - \mu}{\frac{s}{\sqrt{n}}}$	$z = \frac{\hat{p} - p}{\sqrt{\frac{p(1-p)}{n}}}$	$= \sqrt{\frac{p_1(1-p_1)}{n_1} + \frac{p_2(1-p_2)}{n_2}}$

Do: *Compute the p-value, show all calculations and calculator inputs*

Conclude: *Explain your findings in context to the problem.*

- Reject H_0: (is p-value < alpha?)—a small p-value provides evidence against the null hypothesis, because data has been observed that would be unlikely if the null hypothesis were correct.

"With a *p* value of ______ at the ______ % significance level we have sufficient evidence to reject the null hypothesis and can conclude (*insert alternative hypothesis in context here*)"

- Failed to re reject H_0: (is p-value > alpha?)— a large p-value does not provide enough evidence to reject the null in favor of the alternative.

"With a *p* value of ______ at the ______ % significance level we don't have sufficient evidence to reject the null hypothesis and cannot conclude (*insert alternative hypothesis in context here*)"

Example 1:

A group of college stat students was interested in how many students were happy with their choice of university. The students went out and obtained a random sample and asked the question, "Do you plan on returning next year?" The responses along with the gender of the person responding are summarized in the following table.

Test to see if the proportion of students planning on returning is the same for both genders at the 0.05 level of significance?

		Response			
		Yes	No	Maybe	
Gender	Male	211	45	19	275
	Female	141	32	9	182

- **State:** Set up hypotheses and identify statistical data

  Ho $P_{male} = P_{female}$ $\alpha = 0.05$

  Ha $P_{male} \neq P_{female}$

  male $X = 211$ $n = 275$ female $X = 141$ $n = 182$

- **Plan:** Select the appropriate and consider conditions. **Test type:** 2 prop

- **Do:** Compute the statistic.

  P –value = 0.85 **z-score =** −0.19

- **Conclude:** **Reject the null / Fail to Reject the null**

With a p value of 0.85 at the 0.05 % significance level we **(don't have/ have)** sufficient evidence to reject the null hypothesis and **(can/cannot)** conclude that the proportion of males returning is different from females.

(insert alternative hypothesis in context here)

Example 2:

A consumer agency spokesman stated that he thought that the proportion of households having a washing machine was higher for suburban households then for urban households. To test to see if that statement was correct at the 0.05 level of significance, a reporter randomly selected a number of households in both suburban and urban environments and obtained the following data.

	Number surveyed	Number having washing machines	Proportion having washing machines
Suburban	300	243	0.810
Urban	250	181	0.724

- **State:** Set up hypotheses and identify statistical data

  Ho $P_{sub} = P_{urban}$ $\alpha = 0.05$

  Ha $P_{sub} > P_{urban}$

  suburban $X = 243$ $n = 300$ urban $X = 181$ $n = 250$

- **Plan:** Select the appropriate and consider conditions. **Test type:** 2 sample prop

- **Do:** Compute the statistic.

  P –value = 0.008 **z-score =** 2.39

- **Conclude:** **Reject the null / Fail to Reject the null**

With a p value of 0.008 at the 5 % significance level we **(don't have/ have)** sufficient evidence to reject the null hypothesis and **(can/cannot)** conclude that the % of households with washing machines is higher in the suburbs.

Statistical Reasoning
Hypothesis Tests key

Name: Date: Class:

Hypothesis Testing the Difference of Two Proportions Practice

Speech Delay

One of the most striking human accomplishments is the achievement of intelligible speech by age five. Deficits in speech acquisition often occur when a child has hearing difficulty or is slow to develop the motor functions associated with speech. Other deficits in speech acquisition are of unknown origin. Males generally develop physiologically at a slower rate than females, and in a recent large study of speech development in young children investigators hypothesized that a greater proportion of boys would have speech deficits for this reason. A random sample of healthy two-year-olds was followed over the course of a year. At age 3, each child was classified as having a speech delay or not. The data by gender is given in the table below:

Speech delay and gender at 3 years old

Gender	Speech delay	No Speech Delay	Total
Male	22	280	302
Female	78	259	337
Total	100	539	639

At the .05 significance level, do the data provide evidence of the hypothesized developmental differences? Provide statistical evidence for your conclusion.

- **State:** Set up hypotheses and identify statistical data

H_o $p_{males} = p_{females}$ $\alpha = .05$

H_a $p_{males} > p_{females}$

male $x = 22$ $n = 302$

female $x = 78$ $n = 337$

- **Plan:** **Test type:** 2 sample prop
- **Do:** **P –value =** 0.99
- **Conclude:** **Reject the null / Fail to Reject the null**

With a p value of 0.99 at the 5 % significance level we ***(don't have/ have)*** sufficient evidence to reject the null hypothesis and ***(can/cannot)*** conclude that the proportion of boys with delayed speech is greater than girls.
(insert alternative hypothesis in context here)