

## STAT 110 – PRACTICE PROBLEM #1 FOR FINAL

The data in the file BreastDiag.JMP come from a study of breast tumors conducted at the University of Wisconsin-Madison. The goal was determine if malignancy of a tumor could be established by using shape characteristics of cells obtained via fine needle aspiration (FNA) and digitized scanning of the cells. The sample of tumor cells were examined under an electron microscope and a variety of cell shape characteristics were measured.

Research Question: Is there evidence that the average cell radius is greater for malignant tumors?

1. Describe the population(s) of interest in this study.

Breast tumor cells (malignant & benign)

2. One variable in this study is DIAGNOSIS (malignant or benign). Is this variable categorical or numerical?

3. Another variable in this study is cell RADIUS. Is this variable categorical or numerical?

4. Which variable is the predictor – the one in question 2 or question 3?

Diagnosis

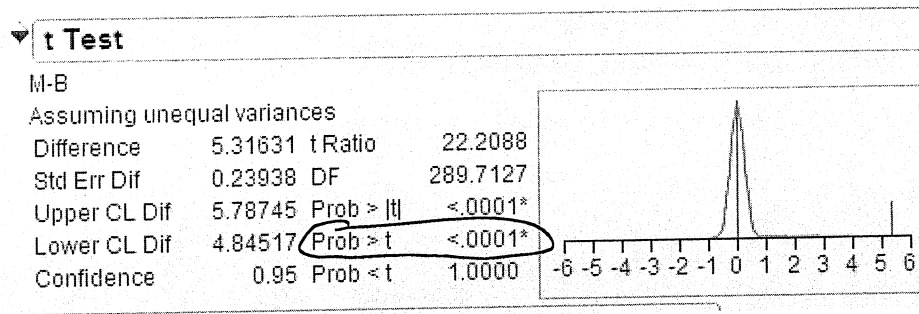
5. Complete the following sentence to explain why this is an observational study and not a designed experiment. "The subjects were..."

not randomly assigned to be malignant or benign

6. Given the predictor and response for this study, what is the name of a test we could use to investigate this research question?

two-sample t-test

7. Carry out this hypothesis test using JMP (be sure to clearly identify the p-value and write a conclusion in the context of the problem).



We have evidence the average <sup>breast</sup> tumor cell radius is greater for M than B tumors (p-value <.0001).

## STAT 110 – PRACTICE PROBLEM #2 FOR FINAL

A designed experiment was carried out in which 160 subjects were randomly assigned to one of four popular diet plans: Atkins, Ornish, Weight Watchers, and Zone (40 subjects per diet). All subjects were overweight or obese with body mass index values between 27 and 42.

Research Question: Is there a significant difference in the drop-out rates across the four diet plans?

	Atkins	Ornish	Weight Watchers	Zone
Completed	21	20	26	26
Dropped out	19	20	14	14

1. Describe the population(s) of interest in this study.

*all overweight/obese people who go on one of the four diets*

2. One variable in this study is DIET PLAN. Is this variable categorical or numerical?

3. Another variable in this study is cell WHETHER THE SUBJECT COMPLETED THE STUDY OR DROPPED OUT. Is this variable categorical or numerical?

4. Which variable is the response – the one in question 2 or question 3?

5. Complete the following sentence to explain why this is a designed experiment. "The subjects were..."

Randomly assigned to one of the four diet plans

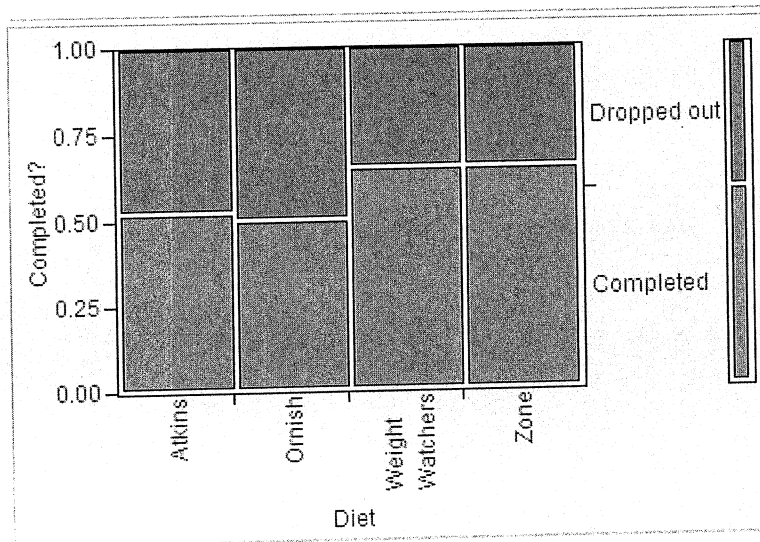


Note: Random selection alone doesn't make designed expts.

6. Given the predictor and response for this study, what is the name of a test we could use to investigate this research question?

*Chi-square test*

7. Carry out this hypothesis test using JMP (be sure to clearly identify the p-value and write a conclusion in the context of the problem).



▼ **Tests**

	N	DF	-LogLike	RSquare (U)
	160	3	1.5841674	0.0146
Test		ChiSquare	Prob>ChiSq	
Likelihood Ratio		3.168	0.3664	
Pearson		3.158	0.3678	<i>- p-value</i>

*No evidence drop-out rate differs across diet.*

## STAT 110 – PRACTICE PROBLEM #3 FOR FINAL

Researchers used magnetic resonance imaging to measure the volumes of various regions of the brain for a sample of 15 identical twins, where one twin was affected by schizophrenia and the other not (“unaffected”). The twins were found in a search through the U.S. and Canada, and ages ranges from 25 to 44 years, with 8 male and 7 female pairs. The data (in cubic centimeters) for the left hippocampus region of the brain appear here and in the file Hippocampus.JMP.

pair	Schizophrenia	
	affected	unaffected
1	1.27	1.94
2	1.63	1.44
3	1.47	1.56
4	1.39	1.58
5	1.93	2.06
6	1.26	1.66
7	1.71	1.75
8	1.67	1.77
9	1.28	1.78
10	1.85	1.92
11	1.02	1.25
12	1.34	1.93
13	2.02	2.04
14	1.59	1.62
15	1.97	2.08

Research Question: Is there evidence of a difference in hippocampus volumes between those affected by schizophrenia and those unaffected?

1. Identify the response variable in this study.

*hippocampus volume*

2. Identify the explanatory variable in this study.

*Group (affected vs. unaffected)*

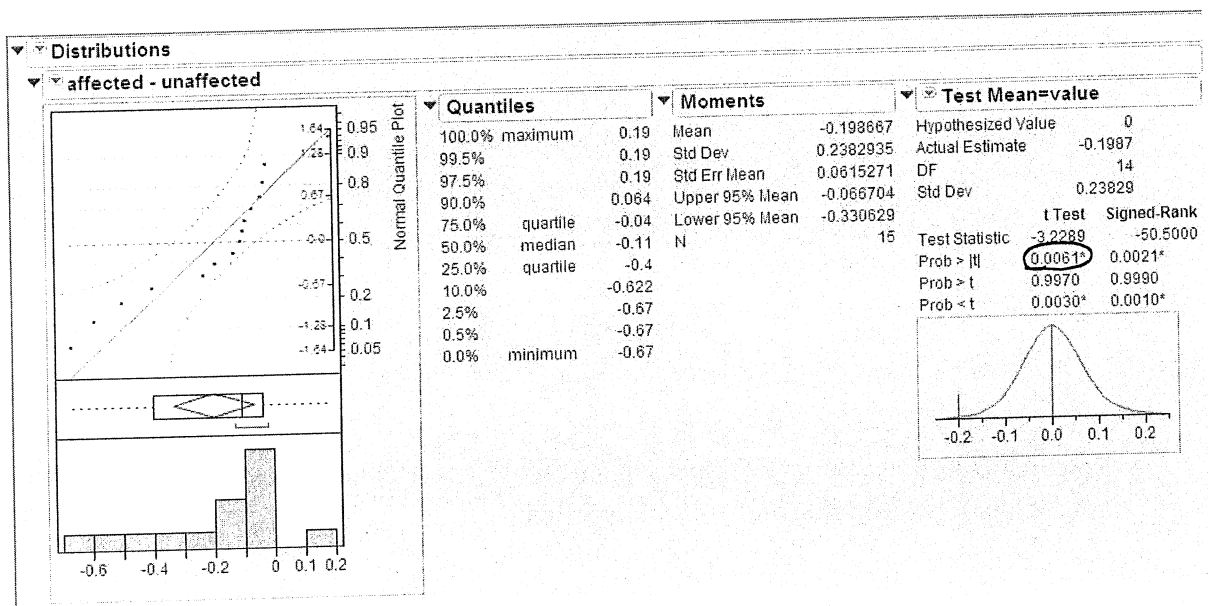
3. Complete the following sentence to explain why this is an observational study and not a designed experiment. "The subjects were..."

not randomly assigned to have Schizophrenia or not.

4. What is the name of a test we could use to investigate this research question?

paired t-test

5. Carry out this hypothesis test using JMP (be sure to clearly identify the p-value and write a conclusion in the context of the problem).



Data provides evidence H.V. differs between those w/ and w/o Schizophrenia (p-value = .0061).

## STAT 110 – PRACTICE PROBLEM #4 FOR FINAL

Researchers have conjectured that the use of the words “forbid” and “allow” can affect people’s responses to survey questions. In a study, college students were randomly assigned to answer one of the following questions:

- Should your college allow speeches on campus that might incite violence?
- Should your college forbid speeches on campus that might incite violence?

The results are as follows:

	Given question with “forbid”	Given question with “allow”
In favor of speeches	12	8
Not in favor of speeches	2	3

Research Question: Is there evidence that college students respond more positively toward having such speeches if their question was phrased in terms of “forbid” rather than “allow”?

8. Describe the population(s) of interest in this study.

all college students

9. One variable in this study is WORD CHOICE in the question they responded to (“forbid” vs. “allow”). Is this variable categorical or numerical?

10. Another variable in this study is WHETHER THEIR RESPONSE WAS IN FAVOR OF THE SPEECHES. Is this variable categorical or numerical?

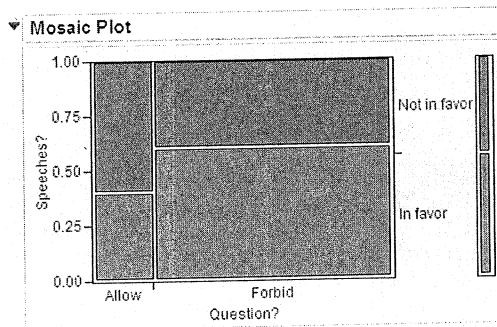
11. Which variable is the response – the one in question 2 or question 3?

12. Complete the following sentence to explain why this is a designed experiment. "The subjects were..."

*Randomly assigned to "Forbid" or "Allow" form*

13. Given the predictor and response for this study, what is the name of a test we could use to investigate this research question? *Fisher's exact test*

14. Carry out this hypothesis test using JMP (be sure to clearly identify the p-value and write a conclusion in the context of the problem).



▼ Tests

	N	DF	-LogLike	RSquare (U)
	25	1	0.32295333	0.0188

Test	ChiSquare	Prob>ChiSq
Likelihood Ratio	0.646	0.4216
Pearson	0.649	0.4203

Fisher's Exact Test	Prob	Alternative Hypothesis
Left	0.9043	Prob(Speeches?=In favor) is greater for Question?=Allow than Forbid
Right	<b>0.3783</b>	Prob(Speeches?=In favor) is greater for Question?=Forbid than Allow
2-Tail	0.6232	Prob(Speeches?=In favor) is different across Question?

*No evidence college students are more likely to be in favor w/ the Forbid form (p-value = 0.3783).*



## STAT 110 – PRACTICE PROBLEM #5 FOR FINAL

Consider data from the Youth Risk Factor Behavior Surveillance System. For each research question, identify which statistical analysis is appropriate.

- A – binomial test for a single proportion
- B – Fisher's Exact Test
- C – Chi-square Test
- D – one-sample t-test
- E – matched pairs t-test (dependent samples)
- F – two-sample t-test (independent samples)

- C 1. Does the proportion of students who have ever smoked a cigarette differ across grade level (freshman, sophomore, junior, senior)?
- A 2. Do the majority of students report never having tried marijuana?
- for C 3. Does the proportion of students who were in a physical fight over the last 12 months differ across gender?
- F 4. Is there a difference in average weight between those who have exercised for the purpose of losing weight in the last 30 days and those who haven't?
- A 5. Do the majority of students report always wearing a seatbelt when riding in a car driven by someone else?
- B or C 6. Does the proportion of students who considered suicide over the past 12 months differ across gender?