STAT 110 **Name:** _____
Practice Exam 2 Solutions
Spring 2017

True/False Questions: Circle the correct answer.

1. A *risk difference* (i.e., difference in proportions) of 0 indicates no significant difference in the outcome of interest between the two groups being compared.	(TRUE) FALSE
2. An *odds ratio* of 1 indicates no significant difference in the outcome of interest between the two groups being compared.	(TRUE) FALSE
3. Observational studies allow researchers to establish a cause and effect relationship between the explanatory variable and the response variable, as long as the results are statistically significant.	TRUE (FALSE)
4. Suppose a chi-square test indicates that there is a significant difference between the proportion of whites who are registered democrats and the proportion of African Americans who are registered democrats. This means that there is an association between *Race* and *Political Affiliation*.	(TRUE) FALSE
5. The following describes an observational study, not a randomized controlled experiment: *Researchers at the University of Massachusetts investigated the effects of two different weight training regimes in children by recruiting 43 young volunteers from the YMCA after-school program. The children were randomly assigned to one of three different weight training regimes. Improvements in muscular strength were measured for all groups.*	TRUE (FALSE)
6. The following describes an observational study, not a randomized controlled experiment: *To assess the association between mercury dental fillings (known as amalgams) and Multiple Sclerosis (MS), investigators select 200 subjects with MS and 200 subjects without MS and interview each about amalgams received since childhood. The proportion with amalgams is compared between the MS and No MS groups.*	(TRUE) FALSE

Questions 7 - 9 refer to the following scenario. A random sample of U.S. adults were interviewed and asked whether they had ever seen a ghost. The results are summarized for two age groups below.

	Has reportedly seen a ghost	Has not seen a ghost	*Total*
18-29 years old	212	1,313	1,525
30 years old or older	465	3,912	4,377
Total	677	5,225	5,902

7. To investigate whether the proportion that reported having seen a ghost differs across age group, it is most appropriate to compare the proportions 212/677 and 465/677.	TRUE (FALSE)
8. The p-value from the chi-square test of independence was found to be 0.0005. This indicates that there is a statistically significant difference in the proportion that reported having seen a ghost between the two age groups.	(TRUE) FALSE
9. The p-value from the chi square test of independence (0.0005) can be used to test whether the proportion who have reportedly seen a ghost, overall, is significantly lower than 50%.	TRUE (FALSE)

Multiple Choice Questions: Circle the correct answer.

Questions 10 - 11 refer to the following scenario. Suppose a study is conducted to investigate whether heart attack patients who own a pet tend to recover more often than patients who do not own a pet. They randomly select a sample of heart attack patients from a large hospital and follow them for one year. They then compare the sample proportions that have survived and find that 92% of those with pets are still alive while only 64% of those without pets have survived.

10. Which of the following statements most accurately identifies the explanatory and response variables?

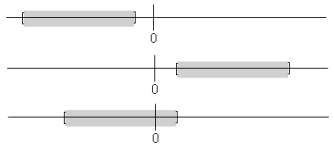
a. The explanatory variable is having a pet and the response variable is not having a pet.
b. The explanatory variable is having a heart attack and the response variable is survival status.
c. The explanatory variable is survival status and the response variable is whether or not the patient has a pet.
d. The explanatory variable is pet ownership and the response variable is survival status.

11. Assume the p-value for testing for a difference in these proportions is less than 0.05. Which of the following statements is most correct?

a. There is evidence that pets increase survival rates among heart attack victims.
b. There is evidence that pets do not increase survival rates among heart attack victims.
c. There is not enough evidence to make any conclusions regarding the association between pet ownership and survival rates after a heart attack.
d. There is evidence that pet ownership is associated with increased survival rates among heart attack victims, but a cause/effect relationship can't be established because of the way the study was designed.

12. Suppose that a study was conducted to test whether there is a difference in the proportion of parents who celebrate Halloween and the proportion of non-parents who celebrate Halloween. Of those surveyed, 80% of parents celebrated Halloween, while only 68% of non-parents did. The difference was found to be statistically significant (p-value < 0.05). Identify which of the following 95% confidence intervals is most likely correct for $\pi_{\text{celebrate | parent}} - \pi_{\text{celebrate | non-parent}}$.

a.
b.
c.



13. A study was conducted to investigate whether or not white women were either more (or less) at risk for having a low birth weight baby than non-white women. The 95% confidence interval obtained was $0.02 \le \pi_{\text{low birth weight | non-white}} - \pi_{\text{low birth weight | white}} \le 0.28$. Which of the following interpretations is most correct?

a. We can be 95% certain the proportion of low birth weight babies born to both non-white and white women is somewhere between 0.02 and 0.28.
b. We can be 95% certain the proportion of white women that have low birth weight babies is anywhere from 2 to 28 percentage points higher than that of non-white women.
c. We can be 95% certain the proportion of non-white women that have low birth weight babies is anywhere from 2 to 28 percentage points higher than that of white women.
d. We can be 95% certain that non-white women are anywhere from 0.02 to 0.28 times as likely to have a low birth weight baby as are white women.

Questions 14 – 16 refer to the following scenario. A survey study was used to classify subjects based on both the type of *Community* in which they lived (rural vs. suburban) and their *Ideology* (conservative vs. liberal). Both the observed counts and the expected counts needed for a chi-square test are shown below.

Observed Counts

	Conservative	Liberal	Moderate
Rural	122	37	125
Suburban	290	112	290

Expected Counts

	Conservative	Liberal	Moderate
Rural	119.9	43.3	120.8
Suburban	292.1	105.6	294.2

14. Which of the following shows the correct calculations for the chi-square test statistic?

a. $\frac{(122-119.9)^2}{119.9}+\frac{(37-43.3)^2}{43.3}+\frac{(125-120.8)^2}{120.8}$

b. $\frac{(290-292.1)^2}{292.1}+\frac{(112-105.6)^2}{105.6}+\frac{(290-294.2)^2}{294.2}$

c. $\sqrt{\frac{122}{119.9}+\frac{37}{43.3}+\frac{125}{120.8}+\frac{290}{292.1}+\frac{112}{105.6}+\frac{290}{294.2}}$

(d.) $\frac{(122-119.9)^2}{119.9}+\frac{(37-43.3)^2}{43.3}+\frac{(125-120.8)^2}{120.8}+\frac{(290-292.1)^2}{292.1}+\frac{(112-105.6)^2}{105.6}+\frac{(290-294.2)^2}{294.2}$

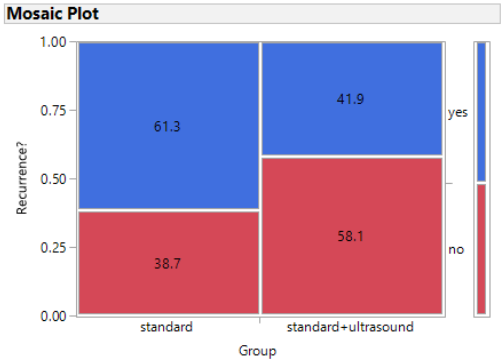
15. If the p-value associated with the chi-square test statistic for this problem is greater than 0.05, then which of the following conclusions is most correct?

(a.) There is not enough evidence to conclude that an association exists between *Community* and *Ideology.*
b. There is a significant difference between the sample and the population.
c. There is no significant association between the sample and the population.
d. There is a significant association between *Community* and *Ideology.*

16. If the p-value associated with the chi-square test statistic for this problem is greater than 0.05, then which of the following statements is most correct?

a. It would be very surprising to obtain the observed sample results if there is really <u>no association</u> between *Community* and *Ideology*.
b. It would be very surprising to obtain the observed sample results if there is really an association between *Community* and *Ideology*.
(c.) It would <u>not</u> be surprising to obtain the observed sample results if there is really <u>no association</u> between *Community* and *Ideology*.
d. It would <u>not</u> be surprising to obtain the observed sample results if there is really an association between *Community* and *Ideology*.

17. A study was conducted to assess the clinical effectiveness of ultrasound therapy to heal leg ulcers. Some of the patients were assigned to the standard care only group, while others were assigned to standard care + ultrasound therapy. Researchers measured whether or not the ulcers recurred after they had been healed. The data were analyzed using JMP to compare the recurrence rates, and the results are shown below.



Fisher's Exact Test	Prob	Alternative Hypothesis
Left	0.1018	Prob(Recurrence?=yes) is greater for Group=standard than standard+ultrasound
Right	0.9628	Prob(Recurrence?=yes) is greater for Group=standard+ultrasound than standard
2-Tail	0.2035	Prob(Recurrence?=yes) is different across Group

Which of the following conclusions is most correct?

a. The study does not provide enough evidence that the recurrence rate is significantly greater for those getting standard care than for those getting standard care + ultrasound therapy.
b. The study does provide evidence that the recurrence rate is significantly different between those getting standard care and those getting standard care + ultrasound therapy.
c. The study does provide evidence that the recurrence rate is significantly greater for the standard care group than for the standard care + ultrasound therapy group.
d. The study does provide evidence that the recurrence rate is significantly greater for the standard care + ultrasound therapy group than for the standard care group.

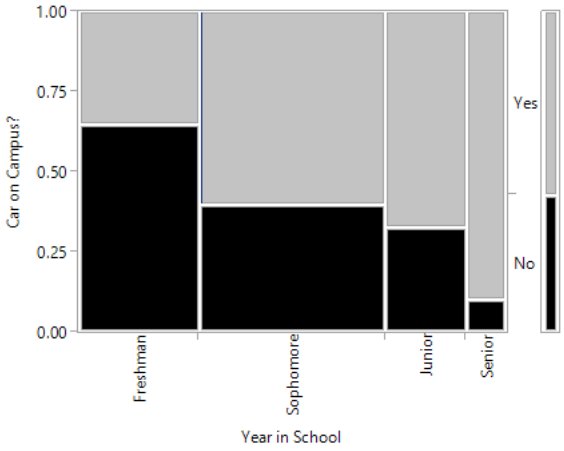
18. In a study examining coffee drinking as a risk factor for bladder cancer, there were 500 cases identified from the local cancer registry and 500 controls identified from the community. All participants were then asked about their coffee drinking habits. Coffee drinkers were defined as people who drank at least one cup per day. Suppose the odds ratio was computed as follows:

$$\frac{\text{odds of cancer for non-coffee drinkers}}{\text{odds of cancer for coffee drinkers}} = 0.286$$

Which of the following is the most correct interpretation of this odds ratio?

a. The odds of bladder cancer are higher for non-coffee drinkers than for coffee drinkers; specifically, the odds of bladder cancer are 28.6% higher for non-coffee drinkers as for coffee drinkers.
b. The odds of bladder cancer are higher for coffee drinkers than for non-coffee drinkers; specifically, the coffee drinkers are 28.6% more likely to have bladder cancer as are the non-coffee drinkers.
c. The odds of bladder cancer are higher for coffee drinkers than for non-coffee drinkers; specifically, the odds of bladder cancer are 0.286 times as large for the non-coffee drinkers as for the coffee drinkers.
d. The odds of bladder cancer are higher for coffee drinkers than for non-coffee drinkers; specifically, the odds of bladder cancer are 28.6% higher for the coffee drinkers as for the non-coffee drinkers.

19. Suppose a random sample of Winona State students was surveyed. The following mosaic plot describes the relationship between *Year in School* and whether or not a student had a *Car on Campus*.



Which of the following tables shows the data that produced this mosaic plot?

a.

	No	Yes	Totals
Freshman	59	32	91
Sophomore	20	41	61
Junior	56	85	141
Senior	3	27	30

b.

	No	Yes	Totals
Freshman	32	59	91
Sophomore	85	56	141
Junior	41	20	61
Senior	27	3	30

c.

	No	Yes	Totals
Freshman	59	32	91
Sophomore	56	85	141
Junior	20	41	61
Senior	3	27	30

d.

	No	Yes	Totals
Freshman	32	59	91
Sophomore	41	20	61
Junior	85	56	141
Senior	27	3	30

Questions 20 - 23 refer to the following study. You want to investigate a claim that women are more likely than men to dream in color. You take a random sample of men and a random sample of women (in your community) and ask whether they dream in color and compare the proportions of each gender that dream in color.

20. If the difference in the proportions (who dream in color) between the two samples turns out *not* to be statistically significant (the p-value was greater than 0.05), which of the following would be the best conclusion to draw?

a. You have found strong evidence that there is no difference between the proportions of men and women in your community that dream in color.
b. You have not found enough evidence to conclude that there is a difference between the proportions of men and women in your community that dream in color.
c. You have found strong evidence against the claim that there is a difference between the proportions of men and women that dream in color.
d. Because the result is not significant, we can't conclude anything from this study.

21. Suppose that two different studies are conducted on this issue.

- Study A finds that 40 of 100 women sampled dream in color, compared to 20 of 100 men.
- Study B finds that 35 of 100 women dream in color, compared to 25 of 100 men.

Which study (A or B) would provide stronger evidence that there is a genuine difference between men and women on this issue?

a. Study A
b. Study B
c. The strength of evidence would be similar for these two studies

22. Suppose two more studies are conducted on this issue. Both studies find 30% of women sampled dream in color, compared to 20% of men. But Study C consists of 100 people of each sex, whereas Study D consists of 40 people of each sex. Which study would provide stronger evidence that there is a genuine difference between men and women on this issue?

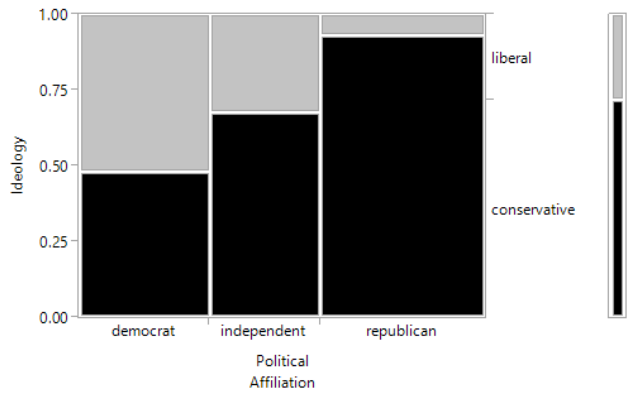
a. Study C
b. Study D
c. The strength of evidence would be similar for these two studies

23. Suppose the difference in the sample proportions who dream in color does turn out to be statistically significant (p-value < 0.05). Below are two possible explanations for this observed difference. Which do you consider to be the more plausible (i.e., believable or reasonable) explanation for the observed result?

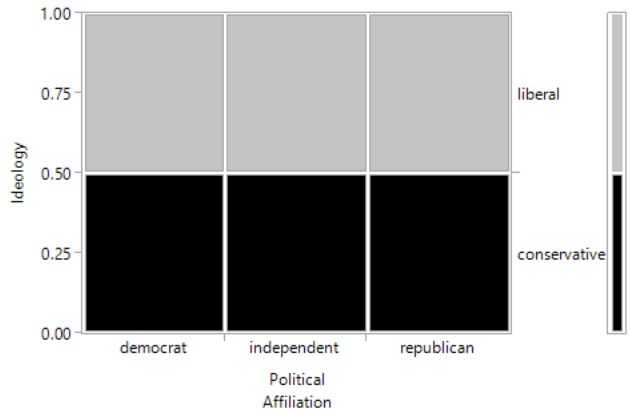
a. Men and women in your community do not differ on this issue but by chance alone the random sampling led to the difference we observed between the two groups.
b. Men and women in your community differ on this issue.
c. (a) and (b) are equally plausible explanations.

Questions 24 and 25 refer to the following mosaic plots. Suppose these plots were used to investigate whether an association exists between *Political Affiliation* and *Ideology*.

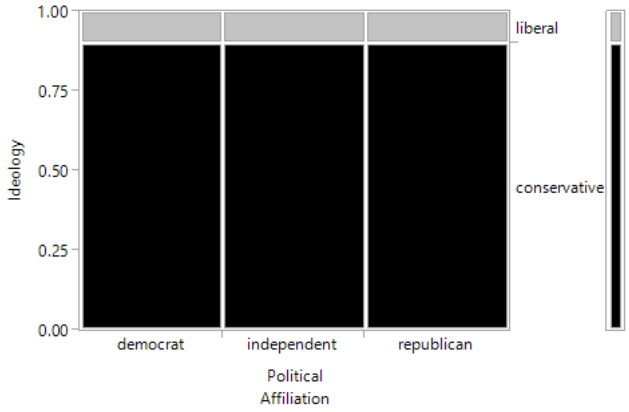
Plot I



Plot II



Plot III



24. Which mosaic plot (or plots) indicate that there is no association between *Political Affiliation* and *Ideology*?

a. Plot I only
b. Plot II only
c. Plot III only
d. Plots II and III
e. Plots I and III

25. The data shown in which plot will yield the smallest p-value?

a. Plot III
b. Plot II
c. Plot I

Short Answer Questions:

26. A study was conducted to examine the association between religion and belief in life after death. The data are summarized in the following table.

	Yes	No	Unsure	Totals
Catholic	118	8	23	149
Protestant	103	7	6	116
Other	22	10	14	46
None	45	49	50	144
Totals	288	74	93	455

a. Use relevant percentages to explain how the proportion that *do* believe in life after death differs across the four levels of religion in this data set.

Catholic: 118/149 = 0.79
Protestant: 103/116 = 0.89
Other: 22/46 = 0.48
None: 45/144 = 0.31

Protestants and Catholics are more likely to believe in life after death than are people of other or no religion.

b. Set up the null and alternative hypothesis to test whether the observed differences are significant (i.e., set up the hypotheses to test for an association between religion and belief in life after death.

H_o: **There is no association between religion and belief in life after death.**

H_a: **There is an association between religion and belief in life after death.**

c. Compute the three expected counts highlighted in the table below (these would be used to compute the chi-square test statistic).

	Yes	No	Unsure	Totals
Catholic	**94.3**		**30.5**	149
Protestant		**18.9**		116
Other				46
None				144
Totals	288	74	93	455

d. Identify the value of the chi-square statistic from the JMP output.

Tests

N	DF	-LogLike	RSquare (U)
455	6	64.012500	0.1547

Test	ChiSquare	Prob>ChiSq
Likelihood Ratio	128.025	<.0001*
Pearson	121.765	<.0001*

Chi-square test statistic: _____**121.765**_____

e. Identify the p-value associated with the chi-square statistic from the JMP output.

p-value: _____**<.0001**_____

f. Write a conclusion in the context of the problem.

The study provides evidence that there is a significant association between religion and belief in life after death.

27. A random sample of 200 subjects who were part of a much larger study on survival of patients following admission to an adult intensive care unit (ICU) was obtained. Below is a contingency table summarizing the relationship between mortality and whether or not the patient was admitted to the ICU from the emergency room.

Admitted to ICU from ER	Died	Survived	Row Totals
Yes	38	109	147
No	2	51	53
Column Totals	40	160	200

a. Find the risk of death for those who were admitted to the ICU from the ER.

38/147 = 0.2585

b. Find the risk of death for those who were *not* admitted to the ICU from the ER.

2/53 = 0.0377

c. Find the risk ratio (also known as the relative risk).

RR = 6.86

d. Interpret the risk ratio from part c.

The risk of death for those who were admitted to the ICU from the ER is 6.86 times as large as the risk of death for those who were admitted to the ICU not directly from the ER (or, those who were admitted to the ICU directly from the ER were 6.86 times as likely to die as were those who were not admitted to the ICU directly from the ER).

e. Find the odds of death for those who were admitted to the ICU from the ER.

38/109

f. Find the odds of death for those who were *not* admitted to the ICU from the ER.

2/51

g. Find the odds ratio.

OR = 8.89

h. Interpret the odds ratio from part g.

The odds of death for those who were admitted to the ICU from the ER is 8.89 times as large as the odds of death for those who were admitted to the ICU not directly from the ER.