STAT 110: Homework 6 (20 pts) **Name:** _____
Fall 2017

1. A study was conducted to determine whether the proportion of inmates living with depression differed depending on whether the inmate was serving a life sentence or not. The contingency table below summarizes the data that was collected on 649 randomly selected inmates of prisons in Northwest Amhara, Ethiopia.

	Depression	No Depression	Total
Life sentence	72	66	138
Not a life sentence	212	299	511
Total	284	365	649

Source: Beyen et al. 2017. More than eight in every nineteen inmates were living with depression at prisons of Northwest Amhara Regional State, Ethiopia, a cross sectional study design. *BMC Psychiatry*; 17:31.

a. Estimate the difference in the proportion living with depression between those serving a life sentence and those not serving a life sentence. (1 pt)

$\hat{\pi}_{\text{Depression} \mid \text{Life sentence}} - \hat{\pi}_{\text{Depression} \mid \text{Not a life sentence}}$ **= 72/138 – 212/511 = 0.5217 – 0.4149 = 0.1068**

b. Find the margin of error associated with this difference in proportions. (1 pt)

$$1.96\sqrt{\frac{.5217(1-.5217)}{138}+\frac{.4149(1-.4149)}{511}} = 0.0937$$

c. Find the lower and upper endpoints of the confidence interval for the difference in proportions. (1 pt)

Lower endpoint = **0.1068 – 0.0937 = 0.0131**

Upper endpoint = **0.1068 + 0.0937 = 0.2005**

d. Based on this confidence interval alone (i.e., you do not need to find a p-value), does the study provide evidence that the proportion living with depression differs significantly between those serving a life sentence and those not serving a life sentence? Explain your reasoning. (2 pts)

Since the confidence interval for the difference in proportions does not include zero, we are 95% certain that the two proportions are different. So, the study does provide evidence that the proportion living with depression differs significantly between those serving a life sentence and those not serving a life sentence.

Additional information: Note that the confidence interval tells us we can be 95% certain that the true proportion with depression for those serving a life sentence is anywhere from 0.01 to 0.20 higher than that of those not serving a life sentence.

2. A study was conducted to investigate the relationship between obesity and cardiovascular disease (CVD). The data are summarized in the table below.

	CVD	No CVD	Total
Obese	46	254	300
Not Obese	60	640	700
Total	106	894	1,000

a. Enter the data into JMP and find a p-value to test whether the proportion with CVD differs significantly between the obese and not obese groups. Write a conclusion in the context of the problem. (2 pts)

Tests

N	DF	-LogLike	RSquare (U)
1000	1	4.7795097	0.0141

Test	ChiSquare	Prob>ChiSq
Likelihood Ratio	9.559	0.0020*
Pearson	10.132	0.0015*

Fisher's Exact Test	Prob	Alternative Hypothesis
Left	0.9993	Prob(CVD=yes) is greater for Group=Not obese than Obese
Right	0.0014*	Prob(CVD=yes) is greater for Group=Obese than Not obese
2-Tail	0.0023*	Prob(CVD=yes) is different across Group

p-value: **0.0023 using Fisher's exact test; 0.0015 using the chi-square test**

Conclusion: **There is evidence that the proportion with CVD differs significantly between the obese and not obese groups.**

b. Find the risk of CVD for the obese subjects. You must show your work. (1 pt)

46/300 = 0.1533

c. Find the risk of CVD for the subjects who were not obese. You must show your work. (1 pt)

60 / 700 = 0.0857

d. Find and interpret the relative risk ratio. You must show your work. (1.5 pts)

RR = 0.1533 / 0.0857 = 1.79. This means that the risk of CVD was 1.79 times as large for the obese subjects in this study as for the subjects who were not obese (i.e., the obese subjects in this study were 1.79 times as likely to have CVD as were the subjects who were not obese).

e. Find the odds of CVD for the obese subjects. You must show your work. (1 pt)

46/254

f. Find the odds of CVD for the subjects who were not obese. You must show your work. (1 pt)

60/640

g. Find and interpret the odds ratio. You must show your work (1.5 pts)

OR = (46/254) / (60/640) = 1.93. This means that in this study, the odds of CVD were 1.93 times as large for the obese subjects as for the subjects who were not obese.

3. Note that the relationship between obesity and cardiovascular disease (CVD) in Problem 2 did not adjust for any confounding variables. For example, consider age as a potential confounding variable. Obesity tends to be associated with older age, and older age is also a risk factor for CVD. To account for this, the researchers decided to investigate the relationship between obesity and CVD after controlling for age. They stratified the analysis by creating two age groups: Age < 50 and Age 50+. The data for each group can then be analyzed separately.

Age < 50			
	CVD	No CVD	Total
Obese	10	90	100
Not Obese	35	465	500
Total	45	555	600

Age 50+			
	CVD	No CVD	Total
Obese	36	164	200
Not Obese	25	175	200
Total	61	339	400

a. For the younger subjects (Age < 50), find the risk of CVD for both the obese and not obese subjects, and then find the relative risk ratio. <u>You must show your work</u>. (2 pts)

risk of CVD for obese = 10/100 = 0.10
risk of CVD for not obese = 35/500 = 0.07
RR = 0.10 / 0.07 = 1.43

b. For the older subjects (Age 50+), find the risk of CVD for both the obese and not obese subjects, and then find the risk ratio. <u>You must show your work</u>. (2 pts)

risk of CVD for obese = 36/200 = 0.18
risk of CVD for not obese = 25/200 = 0.125
RR = 0.18/0.125 = 1.44

c. After controlling for age, is there a significant association between obesity and cardiovascular disease (CVD)? Explain your reasoning (note that you should refer to either a confidence interval or a p-value to determine whether the association is significant). (2 pts)

Age <50

Tests

N	DF	-LogLike	RSquare (U)
600	1	0.50290716	0.0031

Test	ChiSquare	Prob>ChiSq
Likelihood Ratio	1.006	0.3159
Pearson	1.081	0.2985

Fisher's Exact Test	Prob	Alternative Hypothesis
Left	0.8910	Prob(CVD=yes) is greater for Group=Not obese than Obese
Right	0.1991	Prob(CVD=yes) is greater for Group=Obese than Not obese
2-Tail	0.2999	Prob(CVD=yes) is different across Group

Age 50+

Tests

N	DF	-LogLike	RSquare (U)
400	1	1.1757471	0.0069

Test	ChiSquare	Prob>ChiSq
Likelihood Ratio	2.351	0.1252
Pearson	2.341	0.1260

Fisher's Exact Test	Prob	Alternative Hypothesis
Left	0.9527	Prob(cvd=yes) is greater for group=not than obese
Right	0.0820	Prob(cvd=yes) is greater for group=obese than not
2-Tail	0.1639	Prob(cvd=yes) is different across group

No. After controlling for age, it was still the case that obese subjects in this study were more likely to have CVD than the subjects who were not obese. However, the results are no longer statistically significant in either age group (p-value = 0.2999 for Age<50 and p-value = 0.1639 for Age 50+).