1.

The study described in the lecture about “Obesity & Heart Disease” also included men. The data for men are shown in the accompanying table.

(a) Compute (i) the sample proportions of CVD deaths for the obese and nonobese groups; (ii) the standard error for the difference in sample proportions; (iii) a 95% confidence interval for the difference in populations.

(i)$\widehat{\pi_1} = \frac{22}{1201}$ =0.01832

$\widehat{\pi_2} = \frac{22}{1431}$=0.01537

$\widehat{\pi_1 - \pi_2}$=0.00295

(ii)

SE=$\sqrt{\frac{0.01832\times(1-0.01832)}{1201} - \frac{0.01537\times(1-0.01537)}{1431}} = 0.00505$

(iii) 95% CI= $(\widehat{\pi_1} - \widehat{\pi_2} \pm 1.96SE) = 0.00295 \pm 1.96 * 0.00505 = (-0.00695, 0.012857)$

(b) Find a one-sided p-value for the test of equal population proportions (using the standard error for testing equality).

Z=$\frac{\widehat{\pi_1} - \widehat{\pi_2}}{SE} = \frac{0.00295}{0.00505} = 0.58415$

p-value→p(z>=0.58415)=0.2796

(c) Compute (i) the sample odds of CVD death for the obese and nonobese groups; (ii) the estimated odds ratio; (iii) the standard error of the estimated log odds ratio; (iv) a 95% confidence interval for the odds ratio.

(i)

$$odds_1 = \frac{22}{1179} = 0.01866\,, odds_2 = \frac{22}{1409} = 0.01561$$

(ii)

$$OR = \frac{0.01866}{0.01561} = 1.1952$$

(iii)

ASE{ln(OR)}=$\sqrt{\frac{1}{22} + \frac{1}{22} + \frac{1}{1179} + \frac{1}{1409}} = 0.30408$

Log(OR)=0.1783

(iv)

95% CI for the log odds ratio is → $1.1952 \pm 1.96 * 0.30408$

95%CI for the odds ratio→$e^{\log(OR) \pm 1.96SE} = (0.659, 2.169)$

(d) Write a concluding sentence that incorporates the confidence interval from part (c)

There is no evidence that the odds ration differs from 1 and no sufficient evidence that odds ratio is higher in obese because 1 is contained inside the 95% CI.

2. For the data in “Smoking & Lung Cancer” case, test whether the odds of lung cancer for smokers are equal to the odds of lung cancer for nonsmokers, using Fisher’s Exact Test.

Set null hypothesis:

$H_0: \varphi = 1$

$H_A: \varphi \neq 1$

Note the p-value is significant and the confidence interval doesn't contain 1. Therefore, we reject the null hypothesis of the odds ratio being 1.

<Code input>

```
> fisher.test(case1803)


    Fisher's Exact Test for Count Data


data:  case1803

p-value = 0.008823

alternative hypothesis: true odds ratio is not equal to 1

95 percent confidence interval:

  1.409691 30.094245

sample estimates:

odds ratio

  5.333256
```

3. A management behavior analyst has been studying the relationship between male/female supervisory structures in the workplace and the level of employee’s job satisfaction. The results of a recent survey are shown in the table below. Is there sufficient evidence to infer that the level of hob satisfaction depends on the boss/employee gender relationship?

To test the hypothesis is that the level of job satisfaction depends on the boss/employee gender relationship at 5% significance level.

The null and alternative hypothesis is:

$$H_0: The\ level\ of\ job\ satisfaction\ independent\ on\ the\ boss - employeegender\ relationship.$$

$$H_A: The\ level\ of\ job\ satisfaction\ depends\ on\ the\ boss - employee\ gender\ relationship.$$

The chi-squared test:

<code input>

```
> job<-matrix(c(21,25,54,71,39,49,50,38,31,48,10,11),nrow=3,byrow=TRUE)

> colnames(job)<-c("female/male","female/female","male/female","female/female")

> rownames(job)<-c("satisfied","neutral","dissatisfied")

> job=as.table(job)

> chisq.test(job,correct=F)


        Pearson's Chi-squared test


data:  job

X-squared = 74.478, df = 6, p-value = 4.915e-14
```

Conclusion:

Based on the chi-squared test, $X^2 = 74.478,$ and p-value=0.0000. Since the p-value is less than $\alpha = 0.05$, the null hypothesis is rejected at 5% level of significance. There is sufficient evidence to infer that the level of job satisfaction depends on the boss/employee gender relationship. The result is statistically significant.