HW2_1830

Yue Zhang

2024-09-29

##2.15

(a) Answer:Response variables: lower risk of overall cancer and lower risk of colorectal cancers. Explanatory variable: regualr aspirin use. For colorectal cancers, RR = 0.81, which means that the probability for people who use aspirin regularly had colorectal cancers was 19% (1-0.81=0.19) lower than those who didn't take aspirin regularly. Also, 95% CI = (0.75,0.88) which means that we're 95% confident taht the relative risk will lie between 0.75 and 0.88. SInce our CI doesn't include 1, which means that taking aspirin regularly has a significant effect on colorectal cancers.

(b)(i) Since 95% CI = (0.94,0.99) which doesn't include 1, that means taking aspirin regularly has a significant association with overall cancer risk. (ii) Sicne RR = 0.97 which is close to 1, this indicates that it is associated with 3% (1-0.97=0.03) reduction in the risk of overall cancer. So the effect of taking aspirin regularly is relatively weak.

##2.16 (a)

```
# DF
I = 3
J = 3
df = (I - 1) * (J - 1)
df
```

```
## [1] 4
```

```
# Create a matrix of observed counts
Table1 = matrix(c(21, 159, 110, 53, 372, 221, 94, 249, 83), nrow = 3,
    byrow = TRUE)

# Add row and column names for clarity
rownames(Table1) = c("Above Average", "Average", "Below Average")
colnames(Table1) = c("Not Too Happy", "Pretty Happy", "Very Happy")

# View the observed counts matrix
Table1
```

```
##               Not Too Happy Pretty Happy Very Happy
## Above Average            21          159        110
## Average                  53          372        221
## Below Average            94          249         83
```

```r
# Perform the chi-square test
chi_square_test = chisq.test(Table1)

# View the test results
chi_square_test
```

```
##
##  Pearson's Chi-squared test
##
## data:  Table1
## X-squared = 73.352, df = 4, p-value = 4.444e-15
```

Answer: P-value = 4.444e-15 < 0.05. It shows statistically significant that we can reject the null hypothesis. And there's a strong association between a person's received happiness and their family income.

(b) In cell n11 (21), the standardized residual is -2.973. This means that fewer people that have above average income fell not too happy than expected. In cell n13 (110), the standardized residual is 3.144, which means that more people whose income is above average feel very happy than expected. In cell n31 (94), the standardized residual is 7.368, which indicates that many more people whose income is below average feel not too happy than expected. In cell n33 (83), the standardized residual is -5.907. This indicates that far fewer people whose income is below average feel very happy than expected. The most significant standardized deviations are in n31 and n33. These residuals show that happiness and income are not independent.

##2.24

```r
# Create a matrix of observed counts
Table2 = matrix(c(21, 2, 15, 3), nrow = 2, byrow = TRUE)

# Add row and column names for clarity
rownames(Table2) = c("Surgery", "Radiation therapy")
colnames(Table2) = c("Cancer Controlled", "Cancer Not Controlled")

# Perform the Fisher test
fisher.test(Table2, alternative = "greater")
```

```
##
##  Fisher's Exact Test for Count Data
##
## data:  Table2
## p-value = 0.3808
## alternative hypothesis: true odds ratio is greater than 1
## 95 percent confidence interval:
##  0.2864828       Inf
## sample estimates:
## odds ratio
##   2.061731
```

```r
fisher.test(Table2, alternative = "two.sided")
```

```
##
```

```
##  Fisher's Exact Test for Count Data
##
## data:  Table2
## p-value = 0.6384
## alternative hypothesis: true odds ratio is not equal to 1
## 95 percent confidence interval:
##   0.2089115 27.5538747
## sample estimates:
## odds ratio
##   2.061731
```

Answer: Both p-values (0.38 and 0.64) are higher than 0.05, which means that we don't have enough evidence to reject the null hypothesis. There is no significant evidence to suggest that surgery is more effective than radiation therapy for cancer control.

##2.30 (a)True (b)False As when we change the response and explanatory variables, the nominator and denominator remain the same (c)True (d)True