Assign9

Natalie Schmer

28 points total, 2 points per problem part unless otherwise noted.

Q1 Toxic Tomatoes

```
#A
#New
new_1a <- 5/50

#Old
old_1a <- 9/50

#B
prop.test(c(5, 9), c(50, 50), correct = T)

##
##  2-sample test for equality of proportions with continuity correction
##
## data:  c(5, 9) out of c(50, 50)
## X-squared = 0.74751, df = 1, p-value = 0.3873
## alternative hypothesis: two.sided
## 95 percent confidence interval:
##  -0.23510963  0.07510963
## sample estimates:
## prop 1 prop 2
##   0.10   0.18

#X-squared = 0.74751, p-value = 0.3873

#C
#set up matrix
tomatoes <- matrix(c(5, 45, 9, 41), byrow = T, nrow = 2)
colnames(tomatoes) <- c("Toxic", "Non-toxic")
rownames(tomatoes) <- c("New", "Old")
tomatoes

##     Toxic Non-toxic
## New     5        45
## Old     9        41

#test
(tomato_test <- chisq.test(tomatoes, correct = T))

##
##  Pearson's Chi-squared test with Yates' continuity correction
##
```

```
## data:  tomatoes
## X-squared = 0.74751, df = 1, p-value = 0.3873

#Z-squared = 0.748, p-value = 0.3873

#D
tomato_test$expected

##     Toxic Non-toxic
## New     7        43
## Old     7        43

#E
fisher.test(tomatoes)

##
##  Fisher's Exact Test for Count Data
##
## data:  tomatoes
## p-value = 0.3881
## alternative hypothesis: true odds ratio is not equal to 1
## 95 percent confidence interval:
##  0.1235385 1.8596694
## sample estimates:
## odds ratio
##  0.5095856

#p-value = 0.3881
```

F. **Based on part D, fisher's test is preferred because the cell counts are very small as compared to the overall sample size.**

```
#G
prop.test(c(5, 9), c(50, 50), correct = T, alternative = "less")

##
##  2-sample test for equality of proportions with continuity correction
##
## data:  c(5, 9) out of c(50, 50)
## X-squared = 0.74751, df = 1, p-value = 0.1936
## alternative hypothesis: less
## 95 percent confidence interval:
##  -1.00000000  0.05338758
## sample estimates:
## prop 1 prop 2
##   0.10   0.18
```

Q2 Anesthesia

```
#A
#DrugA
a_2 <- 22/47
```

```r
#DrugB
b_2 <- 13/47

#B (4 pts)
#Since the sample sizes are small, fischer test
drugs <- matrix(c(4, 18, 9, 16), byrow=T, nrow =2)
rownames(drugs) <- c("Drug A Yes", "Drug A No")
colnames(drugs) <- c("Drug B Yes", "Drug B No")
drugs

##             Drug B Yes Drug B No
## Drug A Yes          4        18
## Drug A No           9        16

fisher.test(drugs)

##
##  Fisher's Exact Test for Count Data
##
## data:  drugs
## p-value = 0.2071
## alternative hypothesis: true odds ratio is not equal to 1
## 95 percent confidence interval:
##  0.07541563 1.79917977
## sample estimates:
## odds ratio
##  0.4029197

#odds ratio = 0.4029197, p-value= 0.2071
```

Q3 Case Control Study

3A.

```r
bird <- (98/199)/(1-(98/199))
nobird <- (141/469)/(1-(141/469))
odds_ratio <- bird/nobird
```

The bird group has higher odds of getting lung cancer by about 2.25:1

3B.

```r
birds <- matrix(c(328, 141, 101, 98), nrow = 2, byrow = T)
colnames(birds) <- c("Healthy control", "Cancer patients")
rownames(birds) <- c("No Bird", "Bird")
birds

##         Healthy control Cancer patients
## No Bird             328             141
## Bird                101              98
```

```r
epitools::oddsratio(birds, method = "wald")
```

```
## $data
##          Healthy control Cancer patients Total
## No Bird              328             141   469
## Bird                 101              98   199
## Total                429             239   668
## 
## $measure
##                          NA
## odds ratio with 95% C.I. estimate   lower    upper
##                  No Bird 1.000000      NA       NA
##                  Bird    2.257145 1.60518 3.173915
## 
## $p.value
##          NA
## two-sided   midp.exact fisher.exact   chi.square
##   No Bird           NA           NA           NA
##   Bird    3.052348e-06 3.938413e-06 2.243712e-06
## 
## $correction
## [1] FALSE
## 
## attr(,"method")
## [1] "Unconditional MLE & normal approximation (Wald) CI"
```

Based on the interval, it does appear that there is a relationship between bird ownership and lung cancer, where bird owners are more likely to have lung cancer. These estimates are significantly different than 1, since 1 is not in the CI.

3C.

```r
(bird_test <- chisq.test(birds, correct = T))
```

```
## 
##  Pearson's Chi-squared test with Yates' continuity correction
## 
## data:  birds
## X-squared = 21.547, df = 1, p-value = 3.452e-06
```

Since the p value is much less than 0.05, we reject the null hypothesis that the proportion of lung cancer patients is the same between bird owners and those that didn't own birds, and conclude that there is a greater rate of lung cancer with bird owners.