# Math 4311 Contingency Tables Lab

1. A randomized double-blind experiment compared the effectiveness of several drugs in ameliorating postoperative nausea. All patients were anesthetized with nitrous oxide and ether. The following table shows the incidence of nausea during the first four postoperative hours for each of several treatments and a placebo. In this problem, we will test the null hypothesis that all of the treatments and the placebo have the same probability of alleviating nausea at the 5% significance level.

| | Patients without Nausea | Patients with Nausea |
|---|---|---|
| Placebo | 70 | 95 |
| Chlorpromazine | 100 | 52 |
| Dimenhydrinate | 33 | 52 |
| Pentobarbital (100 mg) | 32 | 35 |
| Pentobarbital (150 mg) | 48 | 37 |

   (a) Create a $5 \times 2$ matrix `A` containing the above data. Recall that `c(1,2,3)` yields the vector $(1, 2, 3)$. Given two vectors `u` and `v`, `cbind(u,v)` will bind them together into a matrix with columns `u` and `v`.

   (b) Perform the chi-square test using the command `chisq.test(A)`. What is the reported chi-square statistic?

   (c) How many degrees of freedom are reported by this chi-square test? Does this match $df = (r-1)(c-1)$, as discussed in class?

   (d) What is the $p$-value for the test?

   (e) Do you reject the null hypothesis that all of the treatments and the placebo have the same probability of alleviating nausea?

2. In problem 1, we found a statistically significant difference between the nausea treatments. Now, we will perform pairwise post-hoc tests to compare each pair of treatments.

   (a) Let's start by comparing the placebo to chlorpromazine. We can perform a chi-square test that only uses the first two rows of data using the command `chisq.test(A[c(1,2),])`. What is the resulting $p$-value?

   (b) These post-hoc tests will involve $\binom{5}{2} = 10$ pairwise comparisons, so we need to compute adjusted $p$-values,

$$\text{APV} = \min\left\{\binom{5}{2} \cdot p\text{-value}, 1\right\}.$$

   (In other words, the APV is ten times the $p$-value, but if that result is greater than 1, the APV is 1).

   What is the APV for comparing the placebo to chlorpromazine? Is there a statistically significant difference between the placebo and chlorpromazine at the $\alpha = 0.05$ significance level?

(c) Now let's repeat parts (a) and (b) for the other nine pairwise comparisons. This is pretty tedious to do manually, so let's program some for-loops. As an example, the for-loop below prints every pair of numbers $(i, j)$, such that $1 \leq i < j \leq 5$, along with their sums and products.

```
for(i in 1:4){
  for(j in (i+1):5){
    sum=i+j
    product=i*j

    print(paste("For the pair",i,j))
    print(paste("the sum is",sum))
    print(paste("the product is",product))
  }
}
```

You can modify this code to print APV's for all ten pairwise comparisons. Printing results is reasonable when you have a small number of tasks to complete. For larger scale problems, it's better to store the results in a data frame so they can be viewed later.

```
#First, we create a matrix for storing i, j, sum, and product.
#This matrix starts with zero rows, because it has no data.

results.matrix=matrix(nrow=0,ncol=4)

for(i in 1:4){
  for(j in (i+1):5){
    sum=i+j
    product=i*j

    #Creating a new row
    new.row=c(i,j,sum,product)

    #Attaching this new row to our results matrix
    results.matrix=rbind(results.matrix,
                         new.row)
  }
}

#Storing results in a data frame
results=data.frame(results.matrix)

#Renaming columns and deleting row names
colnames(results)=c("i","j","sum","product")
rownames(results)=NULL

#Viewing results
results
```

(d) Which pairs of treatments, including the placebo, were statistically significantly different?