# Practical 9: Log-linear models for contingency tables

## Question 1

Refer to Rodriguez Table 5.1 (page 2 of Rodriguez Chapter 5).

TABLE 5.1: Serum Cholesterol and Heart Disease

| Serum Cholesterol | Heart Disease | | Total |
|---|---|---|---|
| | Present | Absent | |
| < 260 | 51 | 992 | 1043 |
| 260+ | 41 | 245 | 286 |
| Total | 92 | 1237 | 1329 |

**Create a single vector for the four counts.**

```
Count<-c(51,992,41,245)
```

**Create a single vector to represent the corresponding Serum Cholesterol categories, using appropriate numeric codes.**

```
SC <-c(1, 1, 2, 2)
```

**Create a single vector to represent the corresponding Heart Disease categories, using appropriate numeric codes.**

```
HD <-c(1, 2, 1, 2)
```

**Use R to implement an appropriate Poisson regression model to test the hypothesis of independence of Serum Cholesterol and Heart Disease.**

1. What is the deviance of the fitted model?

2. What is the p-value associated with the deviance?

3. What is the critical value associated with the deviance?

4. What is your conclusion about the fitted model?

5. What is your conclusion about independence of Serum Cholesterol and Heart Disease?

**Question 2**

The traditional approach to testing this hypothesis calculates expected counts under independence and compares observed and expected counts using Pearson's chi-squared test.
The raw data (before cross-tabulation) is stored as '**Table_5_1.txt**'.

**Download this file from Canvas, read it into an R dataframe and attach the dataframe.**

Pearson's chi-squared test can be performed in R using `chisq.test()`. Use the `help(chisq.test)` to learn more about this function.

**Apply the test to the data, without the continuity correction and compare the value of the test statistic to that reported on page 5 of Rodriguez Chapter 5.**

**Question 3: Summer 2017 Q4 (b)**

Consider the following 3x3 contingency table:

| | Variable B | | |
|---|---|---|---|
| Variable A | Level 1 | Level 2 | Level 3 |
| Level 1 | 72 | 88 | 60 |
| Level 2 | 66 | 76 | 56 |
| Level 3 | 86 | 92 | 63 |

It is of interest to determine whether variables A and B are independent.

**Create a single vector for the nine counts.**

```
Count<-c(72,66,86,88,76,92,60,56,63)
```

**Create a single vector to represent the corresponding Variable A categories, using appropriate numeric codes.**

```
A<-c(1,2,3,1,2,3,1,2,3)
```

**Create a single vector to represent the corresponding Variable B categories, using appropriate numeric codes.**

```
B<-c(1,1,1,2,2,2,3,3,3)
```

**Use R to implement an appropriate Poisson regression model to test the hypothesis of independence of Variable A and Variable A.**

6. What is the deviance of the fitted model?

7. What is the p-value associated with the deviance?

8. What is the critical value associated with the deviance?

9. What is your conclusion about the fitted model?

10. What is your conclusion about independence variables A and B?