Tutorial Assignment 1 - SOLUTIONS

Instructor: Kevin McGregor

MATH 4330

Question 1:

```
library(MASS)
data("birthwt")

# Information on the dataset
?birthwt

# We can see from the documentation that
# smoke=1 if smoker, 0 otherwise
# low=1 if low birthweight, 0 otherwise
table(birthwt$smoke, birthwt$low)

##
##      0  1
##   0 86 29
##   1 44 30
```

From the contingency table, we can see that it's more likely that a baby has low birthweight if the mother is a smoker, with 30 out of a total of $44 + 30 = 74$ smokers; for non-smokers, the number of low birthweight babies is only 29 out of $86 + 29 = 115$.

Question 2:

We have two categorial variables and we want to check if they're associated. We can use the **chi-squared test for independence**.

```
chisq.test(birthwt$smoke, birthwt$low)

##
##  Pearson's Chi-squared test with Yates' continuity correction
##
## data:  birthwt$smoke and birthwt$low
## X-squared = 4.2359, df = 1, p-value = 0.03958
```

If we choose level $\alpha = 0.05$ for our test, we have that $p = 0.03958 < 0.05$. This means that we reject the null hypothesis of independence. Therefore, we have evidence that smoking is associated with low birthweight.