

# STAT 477/577 - Technology Guide

## Module 2 - Section 6

### Matched Pair Designs

Below is an explanation of the R commands and functions needed to calculate McNemar's test for 2 by 2 tables and an extension of McNemar's test to  $J$  by  $J$  tables.

- **McNemar's Test**

The base package in R has a function called `mcnemar.test` which will calculate the test statistic and  $p$ -value for McNemar's test of marginal homogeneity. To obtain the test statistic and  $p$ -value from the lecture notes for the equality of the proportion of correct responses from the two histogram questions, we will first read in the data from the file `histogramq.csv`.

```
hist.data<- read.csv(file.choose(), header = T)
```

Then we will calculate the contingency table for the responses from the two histogram questions.

```
hist.table<- table(hist.data$Question1, hist.data$Question2)
```

Finally, we will use the `mcnemar.test` function to obtain the test statistic and  $p$ -value for the test. Similar to earlier in this module, we will not use the continuity correction to calculate the test statistic.

```
mcnemar.test(hist.table, correct = F)

##
##  McNemar's Chi-squared test
##
## data:  hist.table
## McNemar's chi-squared = 7.7143, df = 1, p-value = 0.005479
```

If the test determines evidence of a difference between the two marginal probabilities, a confidence interval can be calculated using the function `mcnemar.ci`. For the histogram example, a 95% confidence interval for the difference in the proportion of correct responses for the two histogram questions is given by:

```
mcnemar.ci(hist.table, conf.level = 0.95)

## Confidence Interval =  -0.05198398 -0.009136565
```

- **Extension of McNemar's Test**

The base package in R does not include a function for the extension of McNemar's test of marginal homogeneity to  $J$  by  $J$  tables. A function for calculating the test statistic and  $p$ -value called `stuart.maxwell.mh` is included in the R package `irr`. To obtain the test statistic and  $p$ -value from the lecture notes for the equality of the distribution of the responses from the two baby questions, we will first read in the data from the file `babyhat.csv`.

```
baby.data<- read.csv(file.choose(), header = T)
```

Then we will calculate the contingency table for the responses from the two baby questions.

```
baby.table<- table(baby.data$Question1, baby.data$Question2)
```

Finally, we will use the `stuart.maxwell.mh` function to obtain the test statistic and  $p$ -value for the test.

```
stuart.maxwell.mh(baby.table)

##  Stuart-Maxwell marginal homogeneity
##
##  Subjects = 195
##    Raters = 2
##    Chisq = 24.2
##
##  Chisq(3) = 24.2
##    p-value = 2.23e-05
```

The required information for the test from lecture is given in the last two lines of the output. The first line is the test statistic  $W$  (with the degrees of freedom) and the second line is the  $p$ -value of the test statistic.

Note: this function also works to calculate the test statistic and  $p$ -value for McNemar's test.