Statistics with R – Intermediate Level

Section 1

Tests of Association

Lesson 1 - Pearson Correlation

```
hw = read.csv("hw.csv")

View(hw)

#########
### how to perform the Pearson correlation
#########

### we will compute and interpret the correlation
### between subjects' height and weight

###############
### Basic assumptions:

# the variables are normally distributed
# there are no significant outliers
# the relationship between the variables is approximately
linear
###############

### run the correlation test
```

```
cor.test(hw$height, hw$weight, method="pearson",
alternative="two.sided", conf.level=0.95)

### or

cor.test(hw$height, hw$weight)
```

Lesson 2 - Spearman and Kendall Correlation

```
hw = read.csv("hw.csv")

View(hw)

#########
### how to perform the Spearman and Kendall correlation
#########

### we will compute and interpret the correlation
### between subjects' height and weight

############

### Spearman correlation

cor.test(hw$height, hw$weight, method="spearman",
conf.level = 0.95, exact=FALSE)

### we set the option exact to FALSE in order to force the
program
### to use the asymptotic formula

### alternatively, we can use the package pspearman

require(pspearman)

spearman.test(hw$height, hw$weight, approximation = "AS89")

### AS89 is an approximation algorithm used to compute the
asymptotic p-values
### other values for the option approximation are exact and
t-distribution
```

```
############

### Kendall corelation

cor.test(hw$height, hw$weight, method="kendall", conf.level
= 0.95, exact=FALSE)
```

Lesson 3 - Partial Correlation

```
ice = read.csv("icecream.csv")

View(ice)

#########
### how to perform the partial correlation
### using the package ppcor
#########

### we will compute and interpret the partial correlation
### between the icecream sales and the number of heart
attacks
### control variable: temperature at noon

### run the Person correlation between the variables first

cor.test(ice$attacks, ice$icecream)

### load the package

require(ppcor)

### run the partial correlation

pcor.test(ice$attacks, ice$icecream, ice$temp,
method="pearson")
```

Lesson 4 - Chi-Square Test For Independence

```
bf = read.csv("breakfast.csv")

View(bf)

#########
```

```
### how to perform the chi square test for association
### using the package gmodels
#########

### we will compute the chi square for association
### for the variables agecat (age category) and bfast
(breakfast preference)

### load the package

require(gmodels)

### execute the CrossTable function

CrossTable(bf$agecat, bf$bfast, expected=T, prop.r=F,
prop.c=F, prop.t=F, prop.chisq=F)

### if some expected values are lower than 5,
### we can compute the Fisher's exact test

fisher.test(bf$agecat, bf$bfast, simulate.p.value = T)

##########

### how to compute the Cramer's V
### with the lsr package

### load the package

require(lsr)

cramersV(bf$agecat, bf$bfast)
```