

IST772 Problem Set 9

Shivani Sanjay Mahaddalkar

The homework for week 10 is based on exercises 2-4, 6, 8 and 9 on pages 209-210 but with changes as noted in this notebook (i.e., follow the problems as given in this document and not the textbook).

Attribution statement: (choose only one) 1. I did this homework by myself, with help from the book and the professor 2. I did this homework with help from the book and the professor and these Internet sources: 3. I did this homework with help from but did not cut and paste any code

Chapter 9, Exercise 2

```
str(CO2)

## Classes 'nfnGroupedData', 'nfGroupedData', 'groupedData' and 'data.frame':
## 84 obs. of 5 variables:
## $ Plant      : Ord.factor w/ 12 levels "Qn1"<"Qn2"<"Qn3"<...: 1 1 1 1 1 1 1
## 2 2 2 ...
## $ Type       : Factor w/ 2 levels "Quebec","Mississippi": 1 1 1 1 1 1 1 1 1
## 1 ...
## $ Treatment: Factor w/ 2 levels "nonchilled","chilled": 1 1 1 1 1 1 1 1 1
## 1 ...
## $ conc       : num  95 175 250 350 500 675 1000 95 175 250 ...
## $ uptake     : num  16 30.4 34.8 37.2 35.3 39.2 39.7 13.6 27.3 37.1 ...
## - attr(*, "formula")=Class 'formula' language uptake ~ conc | Plant
## .. ..- attr(*, ".Environment")=<environment: R_EmptyEnv>
## - attr(*, "outer")=Class 'formula' language ~Treatment * Type
## .. ..- attr(*, ".Environment")=<environment: R_EmptyEnv>
## - attr(*, "labels")=List of 2
## ..$ x: chr "Ambient carbon dioxide concentration"
## ..$ y: chr "CO2 uptake rate"
## - attr(*, "units")=List of 2
## ..$ x: chr "(uL/L)"
## ..$ y: chr "(umol/m^2 s)"

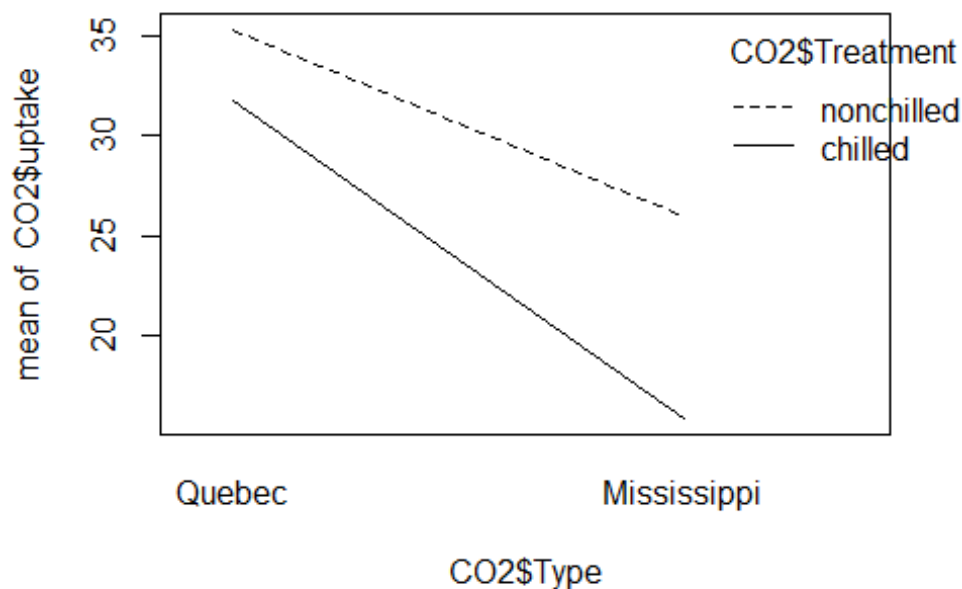
is.na(CO2$uptake)

## [1] FALSE FALSE FALSE FALSE FALSE FALSE FALSE FALSE FALSE FALSE FALSE
## [13] FALSE FALSE FALSE FALSE FALSE FALSE FALSE FALSE FALSE FALSE FALSE
## [25] FALSE FALSE FALSE FALSE FALSE FALSE FALSE FALSE FALSE FALSE FALSE
## [37] FALSE FALSE FALSE FALSE FALSE FALSE FALSE FALSE FALSE FALSE FALSE
```

```
FALSE
## [49] FALSE FALSE FALSE FALSE FALSE FALSE FALSE FALSE FALSE FALSE FALSE
FALSE
## [61] FALSE FALSE FALSE FALSE FALSE FALSE FALSE FALSE FALSE FALSE FALSE
FALSE
## [73] FALSE FALSE FALSE FALSE FALSE FALSE FALSE FALSE FALSE FALSE FALSE
FALSE
```

Using the built in dataset *CO2*, use the `interaction.plot()` command to display a means plot of the “uptake” variable, using “Type” and “Treatment” as the factors. Interpret the results: Without knowing any of the statistics, would you guess that there is or is not an interaction between Type and Treatment? (1 pt)

```
interaction.plot(x.factor = CO2$Type, trace.factor = CO2$Treatment,
response =CO2$uptake)
```



From the plot we can say, that the mean of uptake for different treatments is different. The means for nonchilled treatment is higher than the mean for chilled treatment. The plant in Quebec shows higher uptake rates than the plant at Mississippi. From the plot, we could say that there is a possibility of the interaction between treatment and type having an effect on the uptake rates.

Chapter 9, Exercise 3

Use `aov()` to produce significance tests of the main effects of Type and Treatment and the interaction between Type and Treatment using uptake as the dependent variable. (1 pt) Make

sure to state each null hypothesis and then use the correct language to describe the outcomes of the significance tests. Make sure to report “omnibus” statistics as well. (1 pt)

```
aovOut1 <- aov(uptake~Type*Treatment, data= C02)
summary(aovOut1)

##              Df Sum Sq Mean Sq F value    Pr(>F)
## Type           1   3366      3366  52.509 2.38e-10 ***
## Treatment       1    988       988  15.416 0.000182 ***
## Type:Treatment   1    226       226   3.522 0.064213 .
## Residuals      80   5128         64
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
```

Null Hypothesis: Absence of effect of the interaction between Type and Treatment on the Uptake rates. The p-value of 0.06421(>0.05) means that the effect of the interaction between Treatment and Type on the uptake rates is not statistically significant and we fail to reject the null hypothesis. On inspecting the individual effects, we can see that Type and Treatment are statistically significant slopes

Chapter 9, Exercise 4

Use `anovaBF()` to examine the main effects and interaction as described in Exercise 3. Interpret the results in your own words. Contrast the results from the traditional ANOVA analysis with the results of the Bayes Factor ANOVA. (1 pt) Use the output object from the `anovaBF()` procedure to create an odds ratio that compares the complete model (with the interaction) against a main effects-only model. (1 pt)

Important note: The `anovaBF()` function is picky about how the data sets it analyzes are stored. Use `newC02 <- data.frame(C02)` to convert the C02 data set to a data frame. Then conduct your analysis on `newC02`.

```
newC02 <- data.frame(C02)
library(BayesFactor)

## Warning: package 'BayesFactor' was built under R version 4.0.4

## Loading required package: coda

## Loading required package: Matrix

## *****
## Welcome to BayesFactor 0.9.12-4.2. If you have questions, please contact
## Richard Morey (richarddmorey@gmail.com).
##
## Type BFManual() to open the manual.
## *****

aovOut<- anovaBF(uptake~Type*Treatment, data= newC02)
aovOut
```

```
## Bayes factor analysis
## -----
## [1] Type : 2366422 ±0%
## [2] Treatment : 11.62652 ±0%
## [3] Type + Treatment : 243258644 ±2.6%
## [4] Type + Treatment + Type:Treatment : 297715778 ±1.06%
##
## Against denominator:
## Intercept only
## ---
## Bayes factor type: BFlinearModel, JZS
```

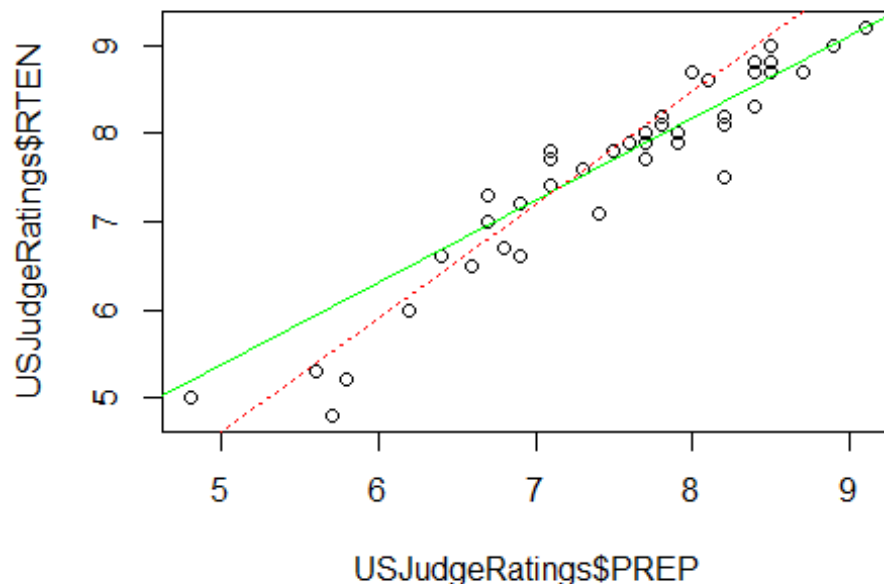
From the Bayesian analysis, we can see that the full model containing the interaction between type and treatment is overwhelmingly in favour of the interaction effect. The individual effects like type and treatment show strong odds to reject the null hypothesis. The odds in favor of the interaction effect does not align with the traditional analysis done before.

```
aovOut[3]/aovOut[4]
## Bayes factor analysis
## -----
## [1] Type + Treatment : 0.8170835 ±2.8%
##
## Against denominator:
## uptake ~ Type + Treatment + Type:Treatment
## ---
## Bayes factor type: BFlinearModel, JZS
```

The comparison of odds with the interaction model vs the main effects only model, shows strong evidence in favour of the main effects only model. The odds in favor of the main effects only model are greater than 3 which means that the traditional analysis is supported. # Chapter 9, Exercise 6

This exercise uses the USJudgeRatings dataset, contains a set of 43 ratings of the competence of state judges in the U.S. Superior Court (from the 1970s). Using similar code as that which appeared in the regression interactions section of this chapter, plot regression lines showing PREP (on the X-axis) and RTEN (on the Y-axis). Plot two regression lines, one for the half of the data set where FAMI is below its median value and one where FAMI is at or above its median value. Without knowing any of the statistics, do you think it likely that there is an interaction between FAMI and PREP? Explain why or why not in your own words. (1 pt)

```
FAMIMedian <- median(USJudgeRatings$FAMI)
plot(USJudgeRatings$PREP, USJudgeRatings$RTEN)
FAMIHi <- subset(USJudgeRatings, USJudgeRatings$FAMI > FAMIMedian)
hilmOut <- lm(RTEN~PREP, data = FAMIHi)
abline(hilmOut, col= "green")
FAMIlo <- subset(USJudgeRatings, USJudgeRatings$FAMI < FAMIMedian)
lolmOut <- lm(RTEN~PREP, data = FAMIlo)
abline(lolmOut, col= "red", lty= 3)
```



The slopes of the two lines is not very different, therefore it seems less likely that there is an interaction between the familiarity with law and preparation for trial.

Chapter 9, Exercise 8

Conduct a regression analysis of the linear and interactive effects of PREP and FAMI on RTEN using the `lm()` function. Be sure to centre the variables first. (1 pt) Interpret the results in your own words, making sure to report the outcomes of the significance tests. (1 pt)

```
RatingsDF <- data.frame(USJudgeRatings)
RatingsDF$scaledPREP <- scale(USJudgeRatings$PREP, center = TRUE, scale=
TRUE)
RatingsDF$scaledFAMI <- scale(USJudgeRatings$FAMI, center = TRUE, scale=
TRUE)
RatingsDF$scaledRTEN <- scale(USJudgeRatings$RTEN, center = TRUE, scale=
TRUE)
lmOut1 <- lm(scaledRTEN~scaledPREP*scaledFAMI, data = RatingsDF)
summary(lmOut1)

##
## Call:
## lm(formula = scaledRTEN ~ scaledPREP * scaledFAMI, data = RatingsDF)
##
## Residuals:
##      Min       1Q   Median       3Q      Max
## -0.83927 -0.17963  0.05695  0.16011  0.58461
```

```
##
## Coefficients:
##              Estimate Std. Error t value Pr(>|t|)
## (Intercept)    0.07497    0.06077   1.234   0.2248
## scaledPREP     0.72424    0.34725   2.086   0.0436 *
## scaledFAMI     0.17915    0.34167   0.524   0.6030
## scaledPREP:scaledFAMI -0.07754    0.03981  -1.948   0.0587 .
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Residual standard error: 0.3084 on 39 degrees of freedom
## Multiple R-squared:  0.9117, Adjusted R-squared:  0.9049
## F-statistic: 134.2 on 3 and 39 DF, p-value: < 2.2e-16
```

The adjusted R squared of the model is statistically significant. Familiarity of law, the preparation for trial and their interaction account for 90.49% of variability in the Ratings. However, the slope for interaction is greater than 0.05, therefore it is statistically insignificant. Only the slope for the preparation of trial is statistically significant.

Chapter 9, Exercise 9

Repeat Exercise 8 using `lmBF()` to conduct one regression analysis that includes the interaction term and one that does not. (1 pt) Interpret the results. Make sure you report an odds ratio for the comparison of a model with linear effects versus a model with linear and interaction effects. Interpret the results in your own words. (1 pt)

```
lmBFInt <- lmBF(scaledRTEN~scaledPREP+scaledFAMI+scaledPREP*scaledFAMI, data
= RatingsDF)
summary(lmBFInt)

## Bayes factor analysis
## -----
## [1] scaledPREP + scaledFAMI + scaledPREP * scaledFAMI : 3.068631e+17 ±0%
##
## Against denominator:
##   Intercept only
## ---
## Bayes factor type: BFlinearModel, JZS

lmBFIntNot <- lmBF(scaledRTEN~scaledPREP+scaledFAMI, data = RatingsDF)
summary(lmBFIntNot)

## Bayes factor analysis
## -----
## [1] scaledPREP + scaledFAMI : 6.298499e+17 ±0%
##
## Against denominator:
##   Intercept only
```

```
## ---
## Bayes factor type: BFlinearModel, JZS

lmBFIntNot/lmBFInt

## Bayes factor analysis
## -----
## [1] scaledPREP + scaledFAMI : 2.052544 ±0%
##
## Against denominator:
##   scaledRTEN ~ scaledPREP + scaledFAMI + scaledPREP * scaledFAMI
## ---
## Bayes factor type: BFlinearModel, JZS
```

The odds are in favour of the main effects only model, however, it only provides weak evidence.