CSPC 440 : Lab Handout 11

Kyle Payne

November 11, 2015

# 1.

* : The response from the kth experimental unit receiving the ith level of factor A and the jth level of factor B.
* : (i,j) treatment mean
* : Overall mean, an unknown constant.
* : An effect due to the ith level of factor A, an unknown constant.
* : An effect due to the jth level of factor B, an unknown constant. tbij: An interaction effect of the ith level of factor A with the jth level of factor B, an unknown constant.
* : A random error associated with the response from the kth experimental unit receiving the ith level of factor A combined with the jth level of factor B. We require that the have a normal distribution with mean 0 and common variance . In addition, the errors must be independent.

Suppose we have a completely randomized experiment with r replications of the treatments constructed by crossing factor A having 3 levels and factor B having 3 levels. The following model

was fit to the data. Answer the following questions:

* 1. After imposing the necessary constraints on the parameters, interpret these parameters in terms of the treatment means, .
  2. State the null and alternative hypotheses for testing for an interaction in terms of the parameters.
  3. State the null and alternative hypotheses for testing for an interaction in terms of the treatment means.
  4. Provide two profile plots, one in which there is an interaction between factors A and B and one where there is not an interaction. (Just write these on the back of the page)

# 2.

In testing food products for palatability, General Foods employed a 7-point scale from -3 (terrible to +3 (excellent) with 0 representing "average". Their standard method for testing palatability was to conduct a taste test with 50 persons - 25 men and 25 women. The experiment reported here involved the effects on palatability of a course versus fine screen and of a low versus high concentration of a liquid compo- nent. Four groups of 50 consumers each were recruited from local churches and club groups. Persons were assigned randomly to the four treatment groups as they were recruited. The experiment was replicated four times, so that there were 16 groups of 50 consumers each in the entire experiment.

Taken from (E. Street and M.G. Carroll, "Preliminary Evaluation of a Food Product," in Statistics: A Guide to the Unknown (Judith M. Tanur, Ed.) Holden-Day, San Francisco, 1972, 220-238)

library(knitr) ## don't run this line  
  
  
## just use the import dataset button  
taste\_test <- read.csv("~/Desktop/taste\_test.txt", sep="")  
  
## fit the ANOVA model  
  
  
model\_1 <- lm(SCORE ~ SCR\*LIQ, data=taste\_test)  
  
## ------   
## kable just prints it out in a nicer format for   
## this sheet. just run anova(model\_1)  
kable(anova(model\_1))

|  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- |
|  | Df | Sum Sq | Mean Sq | F value | Pr(>F) |
| SCR | 1 | 10609.00 | 10609.0000 | 27.269573 | 0.0002140 |
| LIQ | 1 | 1024.00 | 1024.0000 | 2.632109 | 0.1306838 |
| SCR:LIQ | 1 | 420.25 | 420.2500 | 1.080218 | 0.3191377 |
| Residuals | 12 | 4668.50 | 389.0417 | NA | NA |

## plot(model\_1) ## this might be useful for part a.

## a. Test the assumptions of ANOVA (i.e. Normality, Constant Variance, etc. You can use either statistical tests or Diagnostic Plots for these).

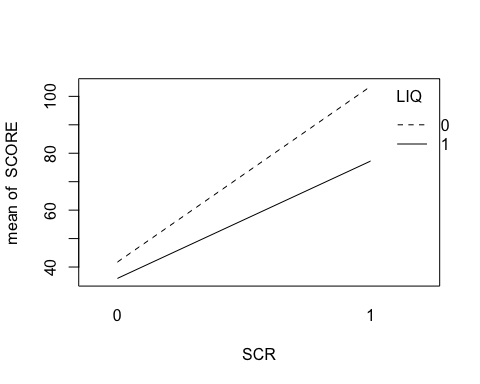
## 

## b. Is the interaction term between SCR and LIQ statistically significant at the level?

## 

## c. Produce an interaction plot

## ------  
## This is how you can produce an interaction plot between  
## two factors  
with(taste\_test,   
interaction.plot( x.factor = SCR, trace.factor= LIQ,  
 response= SCORE)  
)



The with function is a convenient way to use a function that does not have a data = argument, by basically stating, "use this function with this data frame".

## d. Is there an interaction between the factors SCR and LIQ?

## e. What are the main features you should look for in a profile plot to determine whether there is a significant interaction or a lack of interaction between two factors?

# 3.

Six magazine advertisements were randomly selected from each of three magazines selected from magazines grouped into three classes according to educational level of their readers. For each advertisement selected the number of words of advertising copy, the number of sentences in the advertising copy, and the number of three plus syllable words in the advertising copy were observed. Questions arise as to whether significant differences exist in the characteristics of advertising copy among the magazines or the groups of magazines. Also relevant to readability are the number of words per sentence and the proportion of three syllable words in the copy.

magazine\_data <- read.csv("~/Desktop/magazine\_data.txt", sep="")  
## again, just use the import dataset function.  
  
magazine\_data$MAG <- factor(magazine\_data$MAG)  
magazine\_data$GROUP <- factor(magazine\_data$GROUP)

## a. Fit a two-way ANOVA interaction model with MAG and GROUP as factor variables.

## b. Check the assumptions of ANOVA (Can be done using diagnostic plots or Tests). Are these assumptions violated?

c. Is the interaction effect of MAG and GROUP Signficant?

d. Produce an Interaction Plot as in Question 1.

# 4.

The data set we will examine consists of the survival times of animals randomly allocated to three poisons (I, II, III) and four “treatments” (A, B, C, D). The arrangement is called a 3 × 4 factorial treatment structure: each of the 3 poisons appears together with each of the 4 “treatments”. Poisons and “treatments” are the two factors in the treatment structure, each having a certain number of levels: 3 for poisons, 4 for “treatments”. Each combination of a poison and a “treatment” constitutes one treatment in the experimental design, for a total of 3 × 4 = 12 treatments. For example, one treatment would be: administer Poison III and “Treatment” B. The design is replicated 4 times: There are 4 animals in each of the 12 treatment groups. (Since treatment groups have an equal number of animals, the design is said to be balanced.) The design is completely randomized: The 12 × 4 = 48 animals are assigned completely at random to the 12 treatments, subject to having exactly 4 per group. (Notice that the factors have no influence on how the randomization is performed, other than defining the 12 treatment groups.) Suppose both factors (poisons and “treatments”) are of equal interest, and it is possible that they interact.

## install.packages("faraway") ## probably need to run this line too.  
library(faraway)

## Warning: package 'faraway' was built under R version 3.1.2

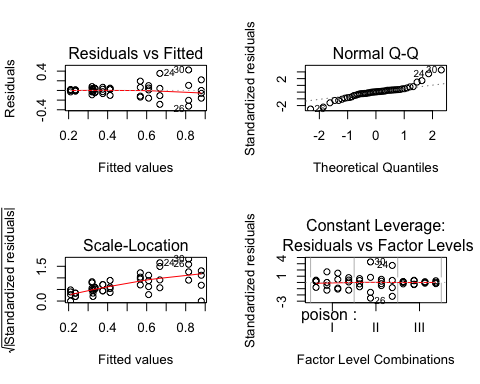
fullmod <- lm(time ~ poison \* treat, data=rats)  
  
# note: poison \* treat = poison + treat + poison:treat  
  
summary(fullmod)

##   
## Call:  
## lm(formula = time ~ poison \* treat, data = rats)  
##   
## Residuals:  
## Min 1Q Median 3Q Max   
## -0.32500 -0.04875 0.00500 0.04312 0.42500   
##   
## Coefficients:  
## Estimate Std. Error t value Pr(>|t|)   
## (Intercept) 0.41250 0.07457 5.532 2.94e-06 \*\*\*  
## poisonII -0.09250 0.10546 -0.877 0.3862   
## poisonIII -0.20250 0.10546 -1.920 0.0628 .   
## treatB 0.46750 0.10546 4.433 8.37e-05 \*\*\*  
## treatC 0.15500 0.10546 1.470 0.1503   
## treatD 0.19750 0.10546 1.873 0.0692 .   
## poisonII:treatB 0.02750 0.14914 0.184 0.8547   
## poisonIII:treatB -0.34250 0.14914 -2.297 0.0276 \*   
## poisonII:treatC -0.10000 0.14914 -0.671 0.5068   
## poisonIII:treatC -0.13000 0.14914 -0.872 0.3892   
## poisonII:treatD 0.15000 0.14914 1.006 0.3212   
## poisonIII:treatD -0.08250 0.14914 -0.553 0.5836   
## ---  
## Signif. codes: 0 '\*\*\*' 0.001 '\*\*' 0.01 '\*' 0.05 '.' 0.1 ' ' 1  
##   
## Residual standard error: 0.1491 on 36 degrees of freedom  
## Multiple R-squared: 0.7335, Adjusted R-squared: 0.6521   
## F-statistic: 9.01 on 11 and 36 DF, p-value: 1.986e-07

anova(fullmod)

## Analysis of Variance Table  
##   
## Response: time  
## Df Sum Sq Mean Sq F value Pr(>F)   
## poison 2 1.03301 0.51651 23.2217 3.331e-07 \*\*\*  
## treat 3 0.92121 0.30707 13.8056 3.777e-06 \*\*\*  
## poison:treat 6 0.25014 0.04169 1.8743 0.1123   
## Residuals 36 0.80073 0.02224   
## ---  
## Signif. codes: 0 '\*\*\*' 0.001 '\*\*' 0.01 '\*' 0.05 '.' 0.1 ' ' 1

par(mfrow=c(2,2))  
plot(fullmod)



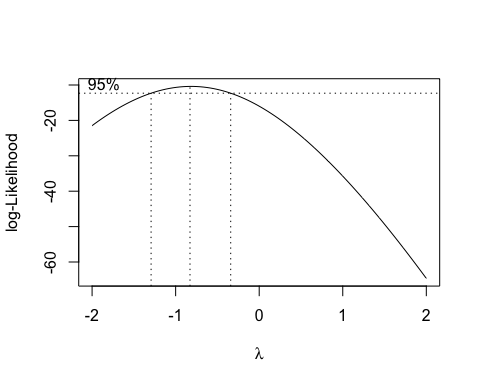
## a. Are the ANOVA assumptions violated?

## b. Perform a box-cox transformation.

library(MASS)

## Warning: package 'MASS' was built under R version 3.1.3

par(mfrow=c(1,1))  
boxcox(fullmod)



## c. Perform the correct transformation, create diagnostic plots, then

produce an interaction plot for the factors Poison and Treatment.