Unit3 HW Solutions

Jacob Turner

January 9, 2019

## 2-Way ANOVA Conceptual Questions

1. The assumptions for two way ANOVA are identical to that of multiple linear regression since it is a special case of multiple linear regression where there are ust two categorical predictor variables used to model the response. Therefore, the assumptions are that the errors are indepdendent, normally distributed, and have constant variance.
2. An interaction effect between the two explanatory variables of the two way anova model indicates that the changes in the mean of response variable due to one of the explanatory variable depends on the level of the other factor. For example. in the ACT data set, if an interaction existed, then the changes that exist between males and females depend on what level of background they have. So there may be no difference between males and females at background a, while there is a very large difference between the two if they have background b.
3. The family wise error rate is the probability of rejecting at least one null hypothesis from a collection of hypothesis tests of which all of the null hypothesis defined by the tests are true. Multiple testing is the act of conducting numerous hypothesis test and comes up quite often in anova models because often the interest is determing what factors (and it’s levels) have different mean values. By conducting multiple hypothesis tests using a significance level at 0.05, the family wise error rate however is not controled at the 0.05 level. It is much higher which and increases as the number of tests increases.
4. FALSE. The type=III F-tests only tell us that there is at least one difference in the means for the given factor of interest that the F test is being reported for. To determine specific differences, additional t-tests must be applied along with multiple test corrections.

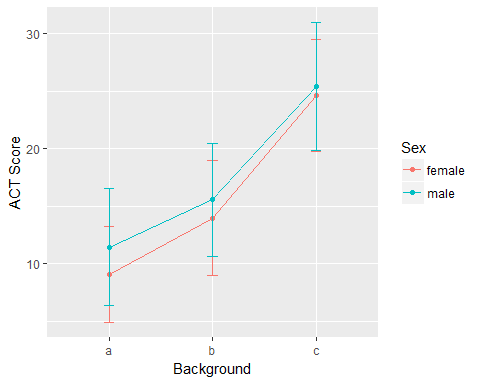
## HW Exercise 1 Q1

setwd("D:/MSDS6372/HWMark")  
  
ACT<-read.csv("MathACT.csv")  
#Attaching the data set, creating a function, and creating a summary stats table. Note: In line 44 below, you can add other statistics like median, IQR,etc.  
  
attach(ACT)  
mysummary<-function(x){  
 result<-c(length(x),mean(x),sd(x),sd(x)/length(x),min(x),max(x),IQR(x))  
 names(result)<-c("N","Mean","SD","SE","Min","Max","IQR")  
 return(result)  
}  
sumstats<-aggregate(Score~Background\*Sex,data=ACT,mysummary)  
sumstats<-cbind(sumstats[,1:2],sumstats[,-(1:2)])  
sumstats

## Background Sex N Mean SD SE Min Max IQR  
## 1 a female 82 9.073171 4.186340 0.05105293 0 18 6.0  
## 2 b female 387 13.963824 5.000905 0.01292224 0 28 6.0  
## 3 c female 54 24.629630 4.849806 0.08981122 15 34 7.0  
## 4 a male 48 11.458333 5.086312 0.10596483 2 25 8.0  
## 5 b male 223 15.565022 4.888305 0.02192065 2 29 7.0  
## 6 c male 67 25.432836 5.554752 0.08290675 13 36 8.5

## HW Exercise 1 Q2.

The means plot using Standard Deviations instead of Standard Errors is provided below. The plot below is much better into gaining insight into the assumption of equal variance. The reason for this is that the SE is not a measure of variability of the data (about its specific mean depending on what treatment combination you are looking at). The SE is the variability of the mean estimate and thus is a function of n and decreases as n gets larger.

For this data set, the sample sizes are quite unbalanced and thus can vary quite a bit. Some SE’s are 5-10 times bigger than others. If the data set was more balaned with equal number of samples in each treatment group, then either graphic would be helpful for assessing equal variances. 

# HW Exercise 2 Q1.

The estimated difference for mean ACT scores (Males vs Females with background A) are the same for the two outputs, 2.385. The p-values are different, Tukey being more conservative because these procedures are trying to control the family wise error rate. Under the Tukey correction, the number of tests in the family is 15, while the bonferroni table is only considering a family of 3 tests.

Since more tests yield higher rates for the family wise error rate. The Tukey adjustment is more conservative since the chances of finding at least one false positive test (aka Reject at least one ) are higher.

# HW Exercise 3 1.

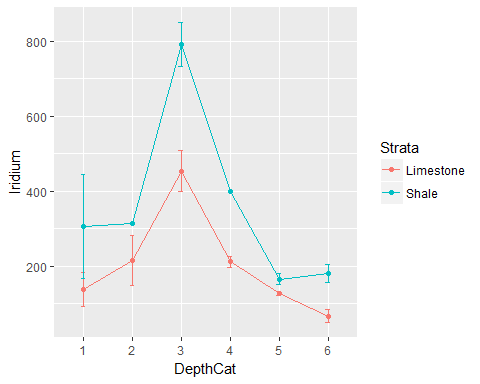
Below is a means plot of the data with the error bars using SE’s. The sample sizes (some have one so there is no variablity at all!) are quite low here so it is very hard to get sense of the equal variance assumption using this plot. A qq-plot of the residuals will better in this case. In terms of an addative versus non addative model, it appears at face value that changes in mean Iridium values between limestone and shale depend on the depth categoriy. Depth Cat 3 looks particulary interesting. (Again, in small samples we have to lean heavy on the ANOVA F-tests)

library(Sleuth3)

## Warning: package 'Sleuth3' was built under R version 3.5.1

mysummary<-function(x){  
 result<-c(length(x),mean(x),sd(x),sd(x)/length(x),min(x),max(x),IQR(x))  
 names(result)<-c("N","Mean","SD","SE","Min","Max","IQR")  
 return(result)  
}  
sumstats<-aggregate(Iridium~Strata\*DepthCat,data=ex1317,mysummary)  
sumstats<-cbind(sumstats[,1:2],sumstats[,-(1:2)])  
ggplot(sumstats,aes(x=DepthCat,y=Mean,group=Strata,colour=Strata))+  
 ylab("Iridium")+  
 geom\_line()+  
 geom\_point()+  
 geom\_errorbar(aes(ymin=Mean-SE,ymax=Mean+SE),width=.1)

## Warning: Removed 2 rows containing missing values (geom\_errorbar).



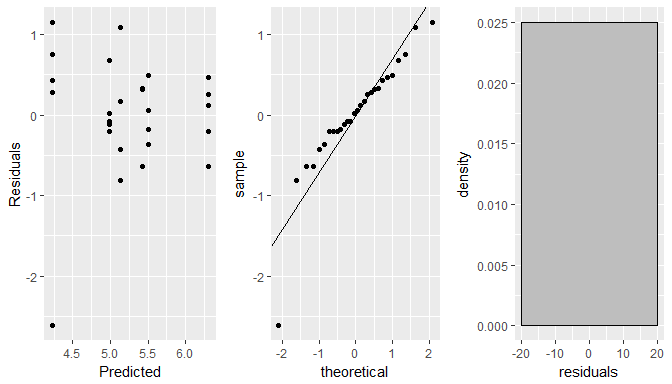
# HW Exercise 3 Q2.

Below are the residual diagnostics of a nonaddative, two way analysis of variance. Overall the normality of the residuals is reasonable with. No exterem outliers were present in previous graphs as well as the qq-plot (middle and right). There is a little concern with the constant variance assumption. The residual versus fitted graph (left) suggests that the variance may tend to increase as the mean of Iridium increases. This is a tough call on whether to transform or not. I would examine an analysis on the transformed scale and then make a call.

model.fit<-aov(log(Iridium)~DepthCat,data=ex1317)  
library(gridExtra)

## Warning: package 'gridExtra' was built under R version 3.5.1

myfits<-data.frame(fitted.values=model.fit$fitted.values,residuals=model.fit$residuals)  
  
#Residual vs Fitted  
plot1<-ggplot(myfits,aes(x=fitted.values,y=residuals))+ylab("Residuals")+  
 xlab("Predicted")+geom\_point()  
  
#QQ plot of residuals #Note the diagonal abline is only good for qqplots of normal data.  
plot2<-ggplot(myfits,aes(sample=residuals))+  
 stat\_qq()+geom\_abline(intercept=mean(myfits$residuals), slope = sd(myfits$residuals))  
  
#Histogram of residuals  
plot3<-ggplot(myfits, aes(x=residuals)) +   
 geom\_histogram(aes(y=..density..),binwidth=40,color="black", fill="gray")  
 #geom\_density(alpha=.2, fill="red")  
  
grid.arrange(plot1, plot2,plot3, ncol=3)



# HW Exercise 3 Q3.

The question here is asking whether or not an interaction exists between the Depth Category and Strata. The type-3 sums of squares F-test is provided below. The test for an interaction is not significant (F-stat: .4792 p-value: .787). Therefore, we can not conclude that the potential changes in mean Iridium for one of the factors does not depend on the other.

library(car)

## Warning: package 'car' was built under R version 3.5.1

## Loading required package: carData

## Warning: package 'carData' was built under R version 3.5.1

Anova(model.fit,type=3)

## Anova Table (Type III tests)  
##   
## Response: log(Iridium)  
## Sum Sq Df F value Pr(>F)   
## (Intercept) 105.400 1 173.4206 6.538e-12 \*\*\*  
## DepthCat 11.759 5 3.8695 0.01146 \*   
## Residuals 13.371 22   
## ---  
## Signif. codes: 0 '\*\*\*' 0.001 '\*\*' 0.01 '\*' 0.05 '.' 0.1 ' ' 1

# HW Exercise 3 Q4.

Based on the Type-3 F-test, the only statistical differences appear to be between depth categories irregardless of strata. Moving forward we have to ways to move forward, one is to reduce the model to a single one way anova model since the others are not signficant. This would boost the degrees of freedom on the error term giving us more information to estimate the assumed constant variance in the model. The second approach would be to conduct the analysis as is. I will provide the results using the model as is.

The following table provides all pairwise comparisons of the 6 depth categories with adjusted p-values using Tukey’s procedure. Examining the adjusted p-values and confidence intervals, there is one comparison that yields a statisticaly significant result. Category 3 versus 6 has an estimated difference of 2.088 (iridium units).

TukeyHSD(model.fit,"DepthCat",conf.level=.95)

## Tukey multiple comparisons of means  
## 95% family-wise confidence level  
##   
## Fit: aov(formula = log(Iridium) ~ DepthCat, data = ex1317)  
##   
## $DepthCat  
## diff lwr upr p adj  
## 2-1 0.29232257 -1.5625125 2.1471576 0.9960089  
## 3-1 1.17644438 -0.4526759 2.8055647 0.2561207  
## 4-1 0.36951566 -1.3477279 2.0867592 0.9834600  
## 5-1 -0.14271103 -1.6648848 1.3794627 0.9996701  
## 6-1 -0.91110705 -2.5402273 0.7180132 0.5203583  
## 3-2 0.88412181 -0.8894397 2.6576833 0.6357960  
## 4-2 0.07719309 -1.7776420 1.9320281 0.9999940  
## 5-2 -0.43503360 -2.1108917 1.2408245 0.9628608  
## 6-2 -1.20342962 -2.9769911 0.5701319 0.3166065  
## 4-3 -0.80692872 -2.4360490 0.8221916 0.6418297  
## 5-3 -1.31915541 -2.7411682 0.1028574 0.0794772  
## 6-3 -2.08755143 -3.6235008 -0.5516021 0.0040376  
## 5-4 -0.51222669 -2.0344004 1.0099470 0.8961262  
## 6-4 -1.28062271 -2.9097430 0.3484976 0.1829648  
## 6-5 -0.76839602 -2.1904088 0.6536168 0.5560767