Jonah Meherg

Oktaviano Harsono (Vino)

Stat 469

Professor Matthew Heaton

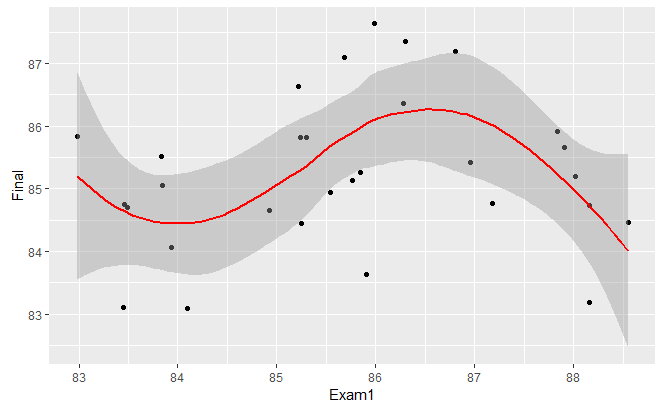
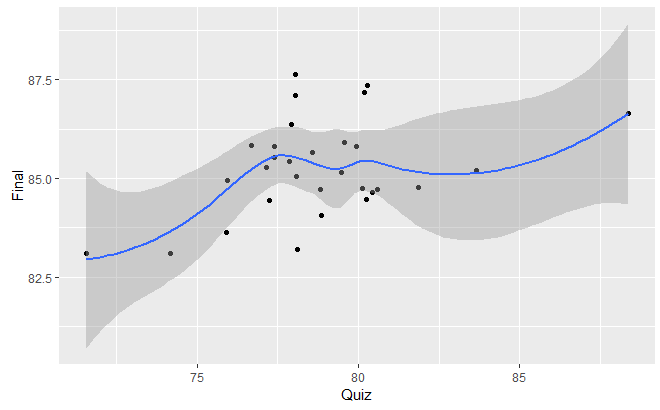
Teaching Pedagogy Analysis

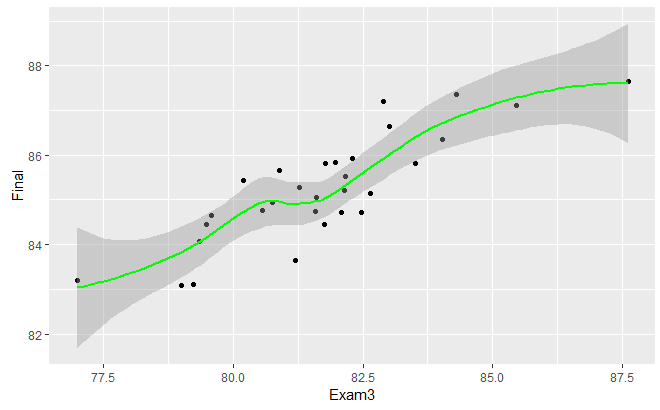
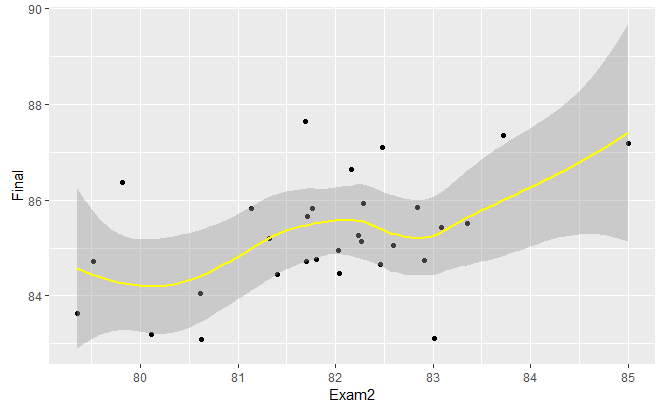
Section 0: Executive Summary

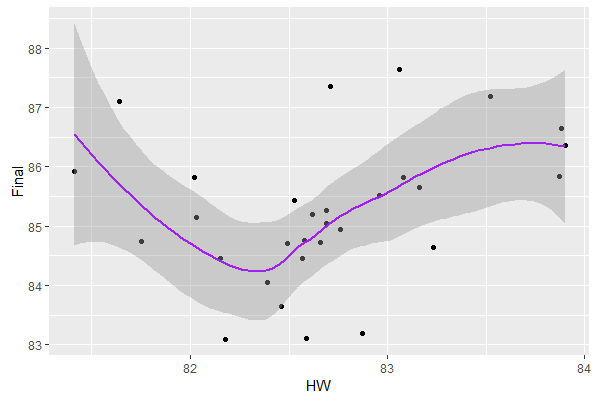
           After a statistical analysis of the pedagogy data provided to us, we have determined that there are some actions that need to take place in our student learning activities, in particular towards our use of quizzes. Exams and homework have a positive effect on student learning and should be emphasized more in a student’s learning experience. Quizzes however have not shown to have an impact on a student’s final exam score.

Section 1: Introduction and Problem Background

    We are trying to determine which class activities, if any, are effective for content mastery. If there are some activities that are not effective, we should discard or modify them accordingly. Likewise, if there are some activities that are effective, we should implement more time in those activities to improved learning. We are looking at data from a statistics department and past semesters over 10 academic years of an introductory statistics class.







According to the data and using scatterplots, we see linear relationships between the response variable (final scores) and the explanatory variables. The final exam can be weighted towards the newer material, which would make the results of exam 3 have more of an impact than it normally would with an equally weighted cumulative final. There could be a couple possible lurking variables such as the way a professor teaches or their teaching style, the type of exam or quizzes given (take home, in class, testing center, frequent pop quizzes), the types of students in the class (motivation, race, culture, prior knowledge). If we ignore these issues, we expect the students that devote their time into this class and with prior knowledge have a better grade than those who do not. If we account for the same teacher teaching the same way, it would be helpful to know that it may be the professor and not the material or curriculum itself affecting student’s grade on the final exam.

We noticed that some of the scatter plots had heteroskedasticity, this could be a result due to the sample size of students across all semesters. So we decided to use a heteroskedastic approach for our model to alleviate some of the issues stated earlier. We will use a heteroskedastic multiple linear regression using fixed weights model since we know sample size. We will then fit the model to our data, make statistical inferences, and use cross validation to check how well our model fits predictions fit.

Section 2: Statistical Model

    The model used was a heteroskedastic multiple linear regression model with a fixed variance, weighted by number of students per semester. The following model was used:

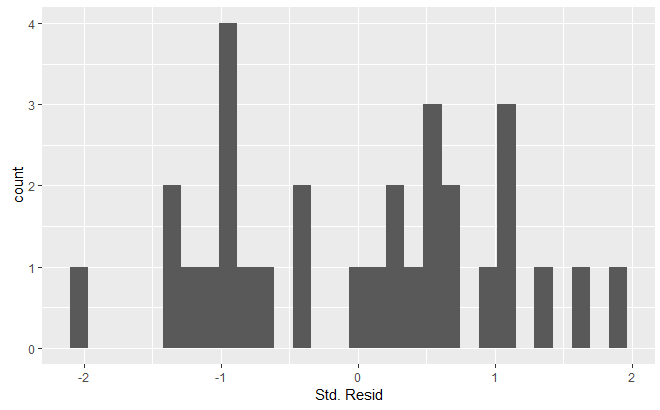
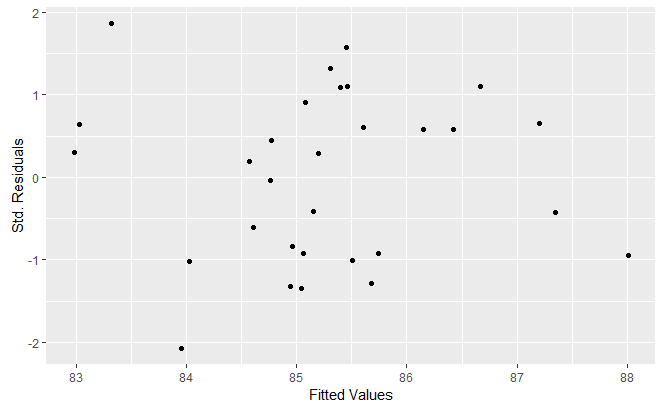
y~*N*(Xß, σ2D),

Where, y is an n **x** 1 vector giving the response variable of final score that is following a normal distribution with mean Xß and variance σ2D. X is a n **x** (p + 1) matrix of explanatory data, and ß is a (p + 1) **x** 1 vector of the coefficients of the variables. σ2 is variance of the final score about the fitted regression line, and D is a n **x** n matrix with dii on the diagonal where dii=|vi| where vi is our inverse sample size of number of students.

    Our model is assuming linearity on account of the model being a linear model. Independence is assumed based on each student group not retaking the course as there is nothing to assume that any two data points are correlated. Normality is assumed because of the normally distributed model. Variance is assumed to be equal among each group, and, after taking a heteroskedastic model weighted by the inverse of number of students, it is assumed to be equal.

Section 3: Model Validation

We assumed linearity in our model by looking at the scatterplot of our data. We assumed independence since each student should not affect another student. We assume normality by looking at the histogram of the standard residuals and they fall between 3 standard deviations. It does not look too normal, but looks decently normal. We also have 30 observations which is a small sample size, so since it looks decently normal it is safe for us to proceed. We assume equal variance because looking at the fitted values vs standard residual scatterplot, we see a normal spread.



Generalized least squares fit by maximum likelihood

 Model: Final ~ .-Quiz

 Data: Pedagogy

|  |  |  |
| --- | --- | --- |
| AIC | BIC | logLik |
| 34.77784 | 45.98742 | -9.388922 |

Variance function:

Structure: fixed weights

Formula: ~(1/NStudents)

Coefficients:

|  |  |  |  |  |
| --- | --- | --- | --- | --- |
|  | Value | Std. Error | t-value | p-value |
| (Intercept) | -30.151271 | 12.594564 | -2.393991 | 0.0252 |
| Semester | -0.001329 | 0.024366 | -0.054542 | 0.9570 |
| NStudents | 0.000095 | 0.000274 | 0.345869 | 0.7326 |
| Exam1 | 0.181473 | 0.044782 | 4.052364 | 0.0005 |
| Exam2 | 0.345743 | 0.058083 | 5.952518 | 0.0000 |
| Exam3 | 0.454548 | 0.034720 | 13.091800 | 0.0000 |
| HW | 0.415269 | 0.128062 | 3.242716 | 0.0036 |

Correlation:

|  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- |
|  | (Intercept) | Semester | NStudents | Exam1 | Exam2 | Exam3 |
| Semester | -0.219 |  |  |  |  |  |
| NStudents | 0.073 | -0.014 |  |  |  |  |
| Exam1 | -0.434 | -0.069 | -0.329 |  |  |  |
| Exam2 | -0.401 | 0.158 | -0.096 | 0.299 |  |  |
| Exam3 | -0.112 | -0.203 | 0.059 | 0.118 | -0.240 |  |
| HW | -0.820 | 0.256 | 0.044 | -0.008 | -0.017 | -0.067 |

Standardized residuals:

|  |  |  |  |  |
| --- | --- | --- | --- | --- |
| Min | Q1 | Med | Q3 | Max |
| -2.2134409 | -0.9307151 | 0.1679617 | 0.7803931 | 1.9904860 |

Residual standard error: 6.967232

Degrees of freedom: 30 total; 23 residual

When we first used the generalized least squares fit by maximum likelihood we saw that quiz was not significant (p-value greater than 0.05), so we took it out of our model and used the rest of the explanatory variables. We first used all of the explanatory variables and noticed that quizzes were not significant, and then upon testing without quizzes we achieved a better AIC and stronger model. We kept semesters and nstudents because intuitively it would make sense to keep those variables and it did not affect our coverage.

|  |  |  |  |
| --- | --- | --- | --- |
| RPMSE | CVG | WID | BIAS |
| 0.4577995 | 0.8666667 | 1.393601 | -0.009493274 |

Our root prediction mean square error is 0.4577995 which means that our prediction can be off by that much, which is good. Our coverage is 0.8666667 meaning 86.66% of our prediction intervals contains the true value, which is pretty high. Our width of our prediction intervals is 1.393601 and our bias is -0.00949 which is really good because the width is small and we want our bias as close to zero as we can.

Section 4: Analysis Results

       The class activities that are associated with improved final scores were exam 1, exam 2, exam 3, and homework. Quizzes however were shown to not be associated with improved learning. For exam 1, for every percent increase we should expect an increased final score percent of 0.18% holding all other variables constant. For exam 2, for every percent increase we should expect an increased final score percent of 0.34% holding all other variables constant. For exam 3, for every percent increase we should expect an increased final score percent of 0.45% holding all other variables constant. For homework, for every percent increase we should expect an increased final score percent of 0.41% holding all other variables constant.

    Our statistical model resulted in an AIC of 34.77784 and an RPMSE of 0.4577995 which suggests that our model shows a strong effect of the explanatory variables on the final score and that they explain learning very well.

            The 3 semesters that did better in terms of student learning were spring semester of year 3, fall semester of year 4, and spring semester of year 2. These did about 2 points higher than the predicted 85% average final score. The 3 semesters that did worse in terms of student learning were spring semester of year 5, winter semester of year 8, and fall semester of year 3. These did about 2 points lower than the predicted 85% average final score. We chose 3 highest and 3 lowest because it would be the top and bottom 10% of our data.

Top 10%

|  |  |  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- | --- | --- |
| Observation | Semester | NStudents | Exam1 | Exam2 | Exam3 | HW | Quiz | Final |
| 9 | 3 | 324 | 85.99 | 81.69 | 87.60 | 83.06 | 78.05 | 87.64 |
| 10 | 4 | 873 | 86.30 | 83.72 | 84.29 | 82.71 | 80.28 | 87.35 |
| 6 | 2 | 336 | 86.81 | 85.00 | 82.89 | 83.52 | 80.20 | 87.19 |

Bottom 10%

|  |  |  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- | --- | --- |
| Observation | Semester | NStudents | Exam1 | Exam2 | Exam3 | HW | Quiz | Final |
| 15 | 5 | 362 | 84.10 | 80.63 | 79.00 | 82.18 | 71.57 | 83.09 |
| 23 | 8 | 287 | 83.45 | 83.01 | 79.23 | 82.59 | 74.16 | 83.11 |
| 7 | 3 | 801 | 88.16 | 80.11 | 77.00 | 82.87 | 78.11 | 83.19 |

Section 5: Conclusions

           We have found that quizzes are not an effective student learning activity as they are not statistically significant in relation to the final exam score. Exams and homework, however do have an impact on student learning as they currently are functioning. We suggest that the department should look into a better way to administer quizzes, or to remove them completely from the teaching curriculum if a better way cannot be discovered.

Appendix

library(readr)

library(lubridate)

library(MASS)

library(ggplot2)

library(lmtest)

library(car)

library(tidyverse)

library(magrittr)

library(multcomp)

library(nlme)

source("predictgls.R")

Pedagogy <- read.delim("https://mheaton.byu.edu/Courses/Stat469/Topics/1%20-%20Independence/3%20-%20Project/Data/ClassAssessment.txt", header = TRUE, sep = "")

head(Pedagogy)

tail(Pedagogy)

ggpairs(Pedagogy)

pairs(Pedagogy)

ggplot(data=Pedagogy, mapping=aes(x=Quiz,y=Final)) + geom\_point() + geom\_smooth()

ggplot(data = Pedagogy, aes(x=Exam1, y=Final)) + geom\_point() + geom\_smooth(color = "red")

ggplot(data = Pedagogy, aes(x=Exam2, y=Final)) + geom\_point() + geom\_smooth(color = "yellow")

ggplot(data = Pedagogy, aes(x=Exam3, y=Final)) + geom\_point() + geom\_smooth(color = "green")

ggplot(data = Pedagogy, aes(x=HW, y=Final)) + geom\_point() + geom\_smooth(color = "purple")

scatter.smooth(Pedagogy$Quiz,Pedagogy$Final)

pedagogy.gls <- gls(Final~., data = Pedagogy, weights = varFixed(value = ~(1/NStudents)), method = "ML")

pedagogy.gls

pedagogy.gls <- gls(Final~.-Quiz, data = Pedagogy, weights = varFixed(value = ~(1/NStudents)), method = "ML")

pedagogy.gls$coefficients

coef(pedagogy.gls$modelStruct, unconstrained = FALSE)

pedagogy.gls$sigma

summary(pedagogy.gls)

rpmse <- rep(NA, nrow(Pedagogy))

cvg <- rep(NA, nrow(Pedagogy))

wid <- rep(NA, nrow(Pedagogy))

bias <- rep(NA, nrow(Pedagogy))

ci.level <- 0.95

alpha <- 1-ci.level

for(obs in 1:nrow(Pedagogy)){

 ## Split Test and Training

 test.set <- Pedagogy[obs,]

 train.set <- Pedagogy[-obs,]

 ## Fit gls to training set

 train.gls <- gls(Final~.-Quiz, data=train.set, weights=varFixed(~1/NStudents), method = "ML")

 ## Generate prediction

 pred <- predictgls(train.gls, newdframe=test.set, ci.level)

 ## Get prediction interval

 pred.low <- pred[['Prediction']] - qt(1-alpha/2, df=nrow(Pedagogy)-length(coef(train.gls)))\*pred[['SE.pred']]

 pred.high <- pred[['Prediction']] + qt(1-alpha/2, df=nrow(Pedagogy)-length(coef(train.gls)))\*pred[['SE.pred']]

 ## RPMSE

 rpmse[obs] <- (pred[['Prediction']]-test.set[['Final']])^2

 ## CVG

 cvg[obs] <- (test.set[['Final']] > pred.low) & (test.set[['Final']] < pred.high)

 ## Wid

 wid[obs] <- (pred.high - pred.low) %>% mean()

 ## bias

 bias[obs] <- mean(pred[,'Prediction']-test.set[,'Final'])

 }

data.frame(RPMSE=sqrt(mean(rpmse)), CVG=mean(cvg), WID=mean(wid), BIAS=mean(bias))

ggplot() + geom\_point(aes(x=fitted(pedagogy.gls), y=resid(pedagogy.gls, type="pearson"))) +

 xlab("Fitted Values") + ylab("Std. Residuals")

ggplot() + geom\_histogram(aes(x=resid(pedagogy.gls, type="pearson"))) + xlab("Std. Resid")

Pedagogy[order(Pedagogy$Final - pred.high, decreasing = TRUE)[1:3],]

Pedagogy[order(Pedagogy$Final - pred.low, decreasing = FALSE)[1:3],]