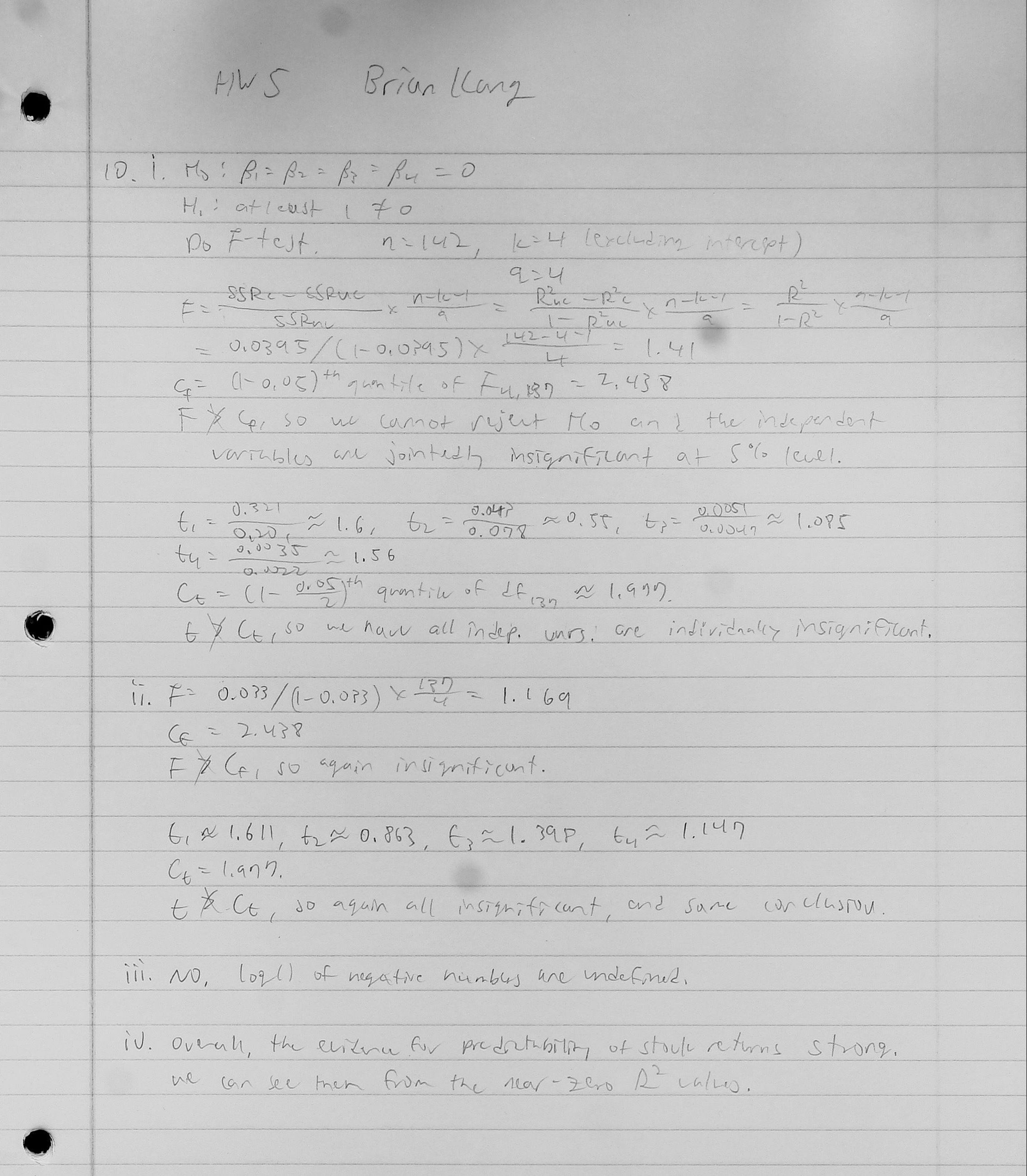
hw05

Brian Kang

Fri Feb 15 23:22:36 2019



library("wooldridge")

## Warning: package 'wooldridge' was built under R version 3.5.2

data("lawsch85")  
  
# C2)  
# i)  
# Null Hypotheses H0: Beta5=0  
# Alt. Hypotheses H1: Beta5=/=0  
lm.1 <- lm(log(salary)~LSAT+GPA+log(libvol)+log(cost)+rank,   
 data = lawsch85, na.action = na.omit)  
summary(lm.1)

##   
## Call:  
## lm(formula = log(salary) ~ LSAT + GPA + log(libvol) + log(cost) +   
## rank, data = lawsch85, na.action = na.omit)  
##   
## Residuals:  
## Min 1Q Median 3Q Max   
## -0.301356 -0.084982 -0.004359 0.077935 0.288614   
##   
## Coefficients:  
## Estimate Std. Error t value Pr(>|t|)   
## (Intercept) 8.3432262 0.5325192 15.667 < 2e-16 \*\*\*  
## LSAT 0.0046965 0.0040105 1.171 0.24372   
## GPA 0.2475238 0.0900371 2.749 0.00683 \*\*   
## log(libvol) 0.0949932 0.0332544 2.857 0.00499 \*\*   
## log(cost) 0.0375539 0.0321061 1.170 0.24427   
## rank -0.0033246 0.0003485 -9.541 < 2e-16 \*\*\*  
## ---  
## Signif. codes: 0 '\*\*\*' 0.001 '\*\*' 0.01 '\*' 0.05 '.' 0.1 ' ' 1  
##   
## Residual standard error: 0.1124 on 130 degrees of freedom  
## (20 observations deleted due to missingness)  
## Multiple R-squared: 0.8417, Adjusted R-squared: 0.8356   
## F-statistic: 138.2 on 5 and 130 DF, p-value: < 2.2e-16

# sample size  
n <- length(lm.1$fitted)  
# R squared  
r2 <- summary(lm.1)$r.squared  
# t- stat of rank  
summary(lm.1)$coefficients[5+1,3]

## [1] -9.540787

# critical value  
qt(1-0.05/2, n-1-1)

## [1] 1.977826

# The at the 5% confidence level, the abs value of the t-stat   
# is greater than the critical value; therefore, we reject our  
# null hypothesis and conclude that the rank of law schools has  
# a ceteris paribus effect on median starting salary.  
  
# ii)  
# t-stat of LAST  
summary(lm.1)$coefficients[1+1,3]

## [1] 1.171045

# |1.171045| < 1.977826  
# Therefore we cannot reject the null hypothesis and LSAT is  
# individually insignificant.  
  
# t-stat of GPA  
summary(lm.1)$coefficients[2+1,3]

## [1] 2.749133

# |2.749133| > 1.977826  
# Therefore we reject the null hypothesis in favor of the alt.  
# GPA is individually significant for explaining salary.  
  
# F-stat  
# testing LSAT and GPA jointly  
# H0: Beta1=Beta2=0  
# k=5; q=2;  
reg.r <- lm(log(salary)~log(libvol)+log(cost)+rank,   
 data = lawsch85, na.action = na.omit)  
rsquared.ur <- summary(lm.1)$r.squared  
rsquared.r <- summary(reg.r)$r.squared  
# calculating the F-stat  
((rsquared.ur - rsquared.r)/(1-rsquared.ur))\*(n-5-1)/2

## [1] 8.038666

# critical value  
qf(1-0.05,2,n-5-1)

## [1] 3.065839

# 8.038666 > 3.065839; therefore, LSAT and GPA jointly are  
# significant.  
  
# iii)  
# This time, H0: Beta6,7=0  
# k=7; q=2;  
lm.2 <- lm(log(salary)~LSAT+GPA+log(libvol)+log(cost)+rank  
 +clsize+faculty,   
 data = lawsch85, na.action = na.omit)  
summary(lm.2)

##   
## Call:  
## lm(formula = log(salary) ~ LSAT + GPA + log(libvol) + log(cost) +   
## rank + clsize + faculty, data = lawsch85, na.action = na.omit)  
##   
## Residuals:  
## Min 1Q Median 3Q Max   
## -0.299686 -0.082283 -0.009353 0.078702 0.269288   
##   
## Coefficients:  
## Estimate Std. Error t value Pr(>|t|)   
## (Intercept) 8.416e+00 5.523e-01 15.239 < 2e-16 \*\*\*  
## LSAT 5.582e-03 4.180e-03 1.336 0.18416   
## GPA 2.661e-01 9.325e-02 2.853 0.00508 \*\*   
## log(libvol) 5.516e-02 4.040e-02 1.365 0.17466   
## log(cost) 2.967e-02 3.468e-02 0.856 0.39393   
## rank -3.428e-03 3.573e-04 -9.594 < 2e-16 \*\*\*  
## clsize 1.342e-04 1.535e-04 0.874 0.38379   
## faculty 6.748e-05 3.999e-04 0.169 0.86629   
## ---  
## Signif. codes: 0 '\*\*\*' 0.001 '\*\*' 0.01 '\*' 0.05 '.' 0.1 ' ' 1  
##   
## Residual standard error: 0.1131 on 123 degrees of freedom  
## (25 observations deleted due to missingness)  
## Multiple R-squared: 0.844, Adjusted R-squared: 0.8351   
## F-statistic: 95.05 on 7 and 123 DF, p-value: < 2.2e-16

# sample size  
n2 <- length(lm.2$fitted)  
reg.r <- lm(log(salary)~LSAT+GPA+log(libvol)+log(cost)+rank,   
 data = lawsch85, na.action = na.omit)  
rsquared.ur <- summary(lm.2)$r.squared  
rsquared.r <- summary(reg.r)$r.squared  
# calculating the F-stat  
((rsquared.ur - rsquared.r)/(1-rsquared.ur))\*(n-7-1)/2

## [1] 0.9431595

# critical value  
qf(1-0.05,2,n-7-1)

## [1] 3.066952

# 0.9431595 < 3.066952, so clsize and faculty are not jointly  
# significant at the 5% confidence level. Therefore, they do   
# not need to be added to this equation  
  
# iv)  
cor(lawsch85$rank, lawsch85$lcost, use = "complete.obs")

## [1] -0.4351474

cor(lawsch85$rank, lawsch85$age, use = "complete.obs")

## [1] -0.507981

cor(lawsch85$rank, lawsch85$llibvol, use = "complete.obs")

## [1] -0.6890247

# Looking at the correlation coefficients of the rank and other  
# variables within the data frame, we can see that some outside  
# of our regression equation may influence the rank of the law  
# school. The three I found include lcost,age,and llibevol.