Homework 3

Elisabeth R Silver

2/22/2021

## Multiple Linear Regression with UT Austin Prof. Evaluation Data

library(pacman)  
p\_load(dplyr, stats, psych, apaTables, ggplot2, yhat, AER, pwr, scales, papaja, knitr)  
data("TeachingRatings")  
#Add an indicator for proportion of students who completed a course eval  
TeachingRatings <- TeachingRatings %>%   
 mutate(prop\_rated = students/allstudents)  
  
vars <- c("eval", "beauty", "prop\_rated")  
  
var.names <- c("Student Evaluation", "Physical Attractiveness", "Evaluation Response Rate")

The dataset contains students’ course evaluations for 463 courses during the 2000-2002 academic years at UT Austin. It has (among other variables) data on the course instructor’s student evaluation, physical attractiveness (z-scored and rated by a panel of six students), and the response rate for the course evaluation.

### Hierarchical models

First, I’ll regree evaluation on physical attractiveness:

mod1 <- lm(eval~beauty, data = TeachingRatings)  
mod1.sum <- summary(mod1)  
mod1.sum

##   
## Call:  
## lm(formula = eval ~ beauty, data = TeachingRatings)  
##   
## Residuals:  
## Min 1Q Median 3Q Max   
## -1.80015 -0.36304 0.07254 0.40207 1.10373   
##   
## Coefficients:  
## Estimate Std. Error t value Pr(>|t|)   
## (Intercept) 3.99827 0.02535 157.727 < 2e-16 \*\*\*  
## beauty 0.13300 0.03218 4.133 4.25e-05 \*\*\*  
## ---  
## Signif. codes: 0 '\*\*\*' 0.001 '\*\*' 0.01 '\*' 0.05 '.' 0.1 ' ' 1  
##   
## Residual standard error: 0.5455 on 461 degrees of freedom  
## Multiple R-squared: 0.03574, Adjusted R-squared: 0.03364   
## F-statistic: 17.08 on 1 and 461 DF, p-value: 4.247e-05

Next, I’ll regress evaluation onto the response rate for the course evaluation:

mod2 <- lm(eval~prop\_rated, data = TeachingRatings)  
mod2.sum <- summary(mod2)  
mod2.sum

##   
## Call:  
## lm(formula = eval ~ prop\_rated, data = TeachingRatings)  
##   
## Residuals:  
## Min 1Q Median 3Q Max   
## -2.00326 -0.34653 0.04265 0.41277 1.25611   
##   
## Coefficients:  
## Estimate Std. Error t value Pr(>|t|)   
## (Intercept) 3.4483 0.1147 30.067 < 2e-16 \*\*\*  
## prop\_rated 0.7389 0.1503 4.915 1.24e-06 \*\*\*  
## ---  
## Signif. codes: 0 '\*\*\*' 0.001 '\*\*' 0.01 '\*' 0.05 '.' 0.1 ' ' 1  
##   
## Residual standard error: 0.5415 on 461 degrees of freedom  
## Multiple R-squared: 0.04979, Adjusted R-squared: 0.04773   
## F-statistic: 24.16 on 1 and 461 DF, p-value: 1.237e-06

Finally, I’ll regress course evaluations on physical attractiveness and evaluation response rate:

mod3 <- lm(eval~prop\_rated+beauty, data = TeachingRatings)  
mod3.sum <- summary(mod3)  
mod3.sum

##   
## Call:  
## lm(formula = eval ~ prop\_rated + beauty, data = TeachingRatings)  
##   
## Residuals:  
## Min 1Q Median 3Q Max   
## -1.90926 -0.32381 0.04198 0.39195 1.31608   
##   
## Coefficients:  
## Estimate Std. Error t value Pr(>|t|)   
## (Intercept) 3.50543 0.11443 30.633 < 2e-16 \*\*\*  
## prop\_rated 0.66217 0.15008 4.412 1.28e-05 \*\*\*  
## beauty 0.11264 0.03189 3.532 0.000453 \*\*\*  
## ---  
## Signif. codes: 0 '\*\*\*' 0.001 '\*\*' 0.01 '\*' 0.05 '.' 0.1 ' ' 1  
##   
## Residual standard error: 0.5348 on 460 degrees of freedom  
## Multiple R-squared: 0.07489, Adjusted R-squared: 0.07086   
## F-statistic: 18.62 on 2 and 460 DF, p-value: 1.679e-08

The unique contribution to the variance in evaluations attributable to physical attractiveness is 0.0391499, to response rate is 0.0250945, and their overlap is 0.0106412