Brandt C3.4 HW

Kate Brandt

January 18, 2019

This includes my code script (I completed the assignment in R) and my answers

**Setting up**

# set workspace for SOCI709  
setwd("~/R/SOCI709/analysis")  
  
# import necessary packages to import .dta file  
library(foreign)  
library(tidyverse)  
  
# import data and create variables  
attend <-read.dta("data/attend.dta")  
atndrte <- attend$atndrte  
priGPA <- attend$priGPA  
ACT <- attend$ACT

**(i) Min, max, mean for selected variables**

atndrte

* Min: 6.25
* Max: 100
* Mean: 81.70956

priGPA

* Min: 0.857
* Max: 3.93
* Mean: 2.58775

ACT

* Min: 133
* Max: 32
* Mean: 22.51029

# Code for (i) to find Min, Max, Mean  
i\_funs <- c("min", "max", "mean")  
attend %>% select(atndrte, priGPA, ACT) %>% summarise\_all(funs\_(i\_funs))

**(ii) Estimate model**

Estimated Model:

atndrte = 75.7 + 17.261(priGPA) - 1.717(ACT) + u

Interpretation:

# Code for (ii) to estimate model  
## model: atndrte = B0 + b1(priGPA) + b2(ACT) + u  
  
## set variables   
head(attend)  
y<- (atndrte)  
x1<- (priGPA)  
x2<- (ACT)  
  
## Create equation and summarize  
attend <-  
 cbind.data.frame(attend,y,x1,x2)  
head(attend)  
mod\_ii <- lm(y~x1+x2, data=attend)  
summary(mod\_ii)

**(iii)**

**(iv) Using estimated model to predict attendance rate**

Predicted attendance rate: 104.3627%

Interpretation: This is an unrealistic result- nobody can have over a 100% attendance rate.

# Code for (iv) to calculate priGPA = 3.65 and ACT = 20  
atndrte\_iv <- 75.7 + 17.261\*3.65 - 1.717\*20  
atndrte\_iv

**(v) Calculating differences between predicted values**

Predicted attendance rate for student A: 93.1521%

Predicted attendance rate for student B: 67.3061%

Difference between rates: 25.846%

Mathematically, this difference makes sense because of the negative coefficient associated with ACT score and higher emphasis on GPA value in this model.

# Code for (v) to predict values and take the difference  
#predict and print values  
studentA <- 75.7 + 17.261\*3.1 - 1.717\*21  
studentA  
studentB <- 75.7 + 17.261\*2.1 - 1.717\*26  
studentB  
#calculate difference  
diffAB <- studentA - studentB  
diffAB