Lab 2

Ben Chu

January 30, 2018

### Loading in packages

library(tidyverse)  
library(car)  
library(QuantPsyc)  
library(stats)  
library(lmSupport)  
library(papaja)

#### Creating the dataset

student <- seq(1:12)  
exam <- c(100,72,84,41,69,74,95,94,81,83,65,61)  
attend <- c(13,15,10,5,9,9,12,9,10,11,2,8)  
gpa <- c(3.4,3.9,3.4,2.3,3.0,2.6,4.0,3.9,2.9,3.4,2.2,3.8)  
class <- data.frame (student,exam,attend,gpa)

## Q1a

attendgpa <-lm(exam~attend, data = class)  
summary(attendgpa)

##   
## Call:  
## lm(formula = exam ~ attend, data = class)  
##   
## Residuals:  
## Min 1Q Median 3Q Max   
## -22.592 -7.622 2.230 10.379 18.642   
##   
## Coefficients:  
## Estimate Std. Error t value Pr(>|t|)   
## (Intercept) 48.885 12.029 4.064 0.00227 \*\*  
## attend 2.941 1.205 2.440 0.03484 \*   
## ---  
## Signif. codes: 0 '\*\*\*' 0.001 '\*\*' 0.01 '\*' 0.05 '.' 0.1 ' ' 1  
##   
## Residual standard error: 13.79 on 10 degrees of freedom  
## Multiple R-squared: 0.3732, Adjusted R-squared: 0.3105   
## F-statistic: 5.955 on 1 and 10 DF, p-value: 0.03484

I would tell the professor that attendence significantly predicts exam scores *R*2 = 0.3732, *F*(1,10) = 5.95, *p* <.05

## Q1b

examgpa <- lm(exam~gpa, data = class)  
summary(examgpa)

##   
## Call:  
## lm(formula = exam ~ gpa, data = class)  
##   
## Residuals:  
## Min 1Q Median 3Q Max   
## -24.260 -6.706 4.552 7.138 20.865   
##   
## Coefficients:  
## Estimate Std. Error t value Pr(>|t|)   
## (Intercept) 27.076 22.247 1.217 0.2515   
## gpa 15.312 6.763 2.264 0.0471 \*  
## ---  
## Signif. codes: 0 '\*\*\*' 0.001 '\*\*' 0.01 '\*' 0.05 '.' 0.1 ' ' 1  
##   
## Residual standard error: 14.17 on 10 degrees of freedom  
## Multiple R-squared: 0.3389, Adjusted R-squared: 0.2728   
## F-statistic: 5.125 on 1 and 10 DF, p-value: 0.04705

I would tell the professor that GPA scores significantly predicts exam scores *R*2 = 0.3389, *F*(1,10) = 5.125, *p* <.05

## Q1c

examscoresattend <- lm(exam~gpa+attend, data = class)  
summary(examscoresattend)

##   
## Call:  
## lm(formula = exam ~ gpa + attend, data = class)  
##   
## Residuals:  
## Min 1Q Median 3Q Max   
## -20.386 -8.050 4.042 8.335 15.273   
##   
## Coefficients:  
## Estimate Std. Error t value Pr(>|t|)  
## (Intercept) 33.798 22.993 1.470 0.176  
## gpa 7.650 9.858 0.776 0.458  
## attend 1.917 1.805 1.062 0.316  
##   
## Residual standard error: 14.07 on 9 degrees of freedom  
## Multiple R-squared: 0.4125, Adjusted R-squared: 0.282   
## F-statistic: 3.16 on 2 and 9 DF, p-value: 0.0913

I would tell the professors that GPA scores with attendence do not significantly predict exam scores *R*2 = 0.4126, *F*(2,9) = 3.16, *p* = 0.09

## Q1d

cor.test(class$gpa,class$attend)

##   
## Pearson's product-moment correlation  
##   
## data: class$gpa and class$attend  
## t = 3.394, df = 10, p-value = 0.006839  
## alternative hypothesis: true correlation is not equal to 0  
## 95 percent confidence interval:  
## 0.2719185 0.9194684  
## sample estimates:  
## cor   
## 0.7316438

* The instructor is referring to a multiple regression.
* The contradiction exists because the model now consists of multiple predictor values, which if correlated may deflate the *r*2 value

## Q1e

new.dat = data.frame(attend = 0, gpa = 2.0)  
badgrade <- predict(examscoresattend,new.dat)

### The student who does not go to class and has a gpa of 2.0 would receive a 49

## Q2

### Loading data

load("C:/Users/Branly Mclanbry/Downloads/lab2AA.RData")

renaming and cleaning data

clean\_dat <- lab2AA %>%  
 mutate(  
 p\_agree = q1agree,  
 p\_fair = q1fair,  
 p\_eff = q1eff,  
 education = q4,  
 employment = q5,  
 happiness = q7,  
 job\_choice = q8,  
 job\_satis = q9,  
 ethnicity = qa,  
 aa\_support = (p\_agree + p\_fair + p\_eff))%>%  
 na.omit()

### General linear model with all variables

every.mod <- lm(aa\_support ~ education + employment + happiness + job\_choice + job\_satis, dat = clean\_dat)  
summary(every.mod)

##   
## Call:  
## lm(formula = aa\_support ~ education + employment + happiness +   
## job\_choice + job\_satis, data = clean\_dat)  
##   
## Residuals:  
## Min 1Q Median 3Q Max   
## -12.2728 -2.6192 0.0128 2.6491 12.5421   
##   
## Coefficients:  
## Estimate Std. Error t value Pr(>|t|)   
## (Intercept) 4.9150 0.4605 10.674 < 2e-16 \*\*\*  
## education 0.2213 0.1896 1.167 0.244005   
## employment 0.5061 0.2035 2.488 0.013349 \*   
## happiness 0.1991 0.1987 1.002 0.317049   
## job\_choice -0.1227 0.2071 -0.592 0.554076   
## job\_satis 0.7391 0.2101 3.518 0.000494 \*\*\*  
## ---  
## Signif. codes: 0 '\*\*\*' 0.001 '\*\*' 0.01 '\*' 0.05 '.' 0.1 ' ' 1  
##   
## Residual standard error: 4.041 on 334 degrees of freedom  
## Multiple R-squared: 0.3187, Adjusted R-squared: 0.3085   
## F-statistic: 31.25 on 5 and 334 DF, p-value: < 2.2e-16

lm.beta(every.mod)

## education employment happiness job\_choice job\_satis   
## 0.09871666 0.21827856 0.07550335 -0.05187809 0.29905318

### Cleaning across racial lines

white <- clean\_dat %>%   
 filter(ethnicity == "White") %>%  
 na.omit()  
   
minority <- clean\_dat %>%  
 filter(ethnicity != "White")

## Running series of linear models

white.1 <- lm(aa\_support ~ education + employment, dat = white)  
white.2 <- lm(aa\_support ~ education + employment + happiness + job\_choice + job\_satis, dat = white)  
minority.1 <- lm(aa\_support ~ education + employment, dat = minority)  
minority.2 <- lm(aa\_support ~ education + employment + happiness + job\_choice + job\_satis, dat = minority)

### Let’s take a look at all the models and some standardized units *b*\*

summary(white.1)

##   
## Call:  
## lm(formula = aa\_support ~ education + employment, data = white)  
##   
## Residuals:  
## Min 1Q Median 3Q Max   
## -8.2340 -2.7881 -0.6037 2.0100 11.8866   
##   
## Coefficients:  
## Estimate Std. Error t value Pr(>|t|)   
## (Intercept) 6.2893 0.5106 12.319 < 2e-16 \*\*\*  
## education 1.0700 0.3324 3.219 0.00153 \*\*   
## employment -0.2459 0.3671 -0.670 0.50383   
## ---  
## Signif. codes: 0 '\*\*\*' 0.001 '\*\*' 0.01 '\*' 0.05 '.' 0.1 ' ' 1  
##   
## Residual standard error: 3.867 on 173 degrees of freedom  
## Multiple R-squared: 0.1417, Adjusted R-squared: 0.1318   
## F-statistic: 14.29 on 2 and 173 DF, p-value: 1.81e-06

lm.beta(white.1)

## education employment   
## 0.45580267 -0.09484471

summary(white.2)

##   
## Call:  
## lm(formula = aa\_support ~ education + employment + happiness +   
## job\_choice + job\_satis, data = white)  
##   
## Residuals:  
## Min 1Q Median 3Q Max   
## -8.9124 -2.5246 -0.6429 1.9482 12.2180   
##   
## Coefficients:  
## Estimate Std. Error t value Pr(>|t|)   
## (Intercept) 5.688558 0.539302 10.548 < 2e-16 \*\*\*  
## education 0.947837 0.344958 2.748 0.00665 \*\*   
## employment -0.191080 0.357908 -0.534 0.59412   
## happiness -0.007591 0.315762 -0.024 0.98085   
## job\_choice -0.730417 0.290372 -2.515 0.01282 \*   
## job\_satis 1.074669 0.322146 3.336 0.00104 \*\*   
## ---  
## Signif. codes: 0 '\*\*\*' 0.001 '\*\*' 0.01 '\*' 0.05 '.' 0.1 ' ' 1  
##   
## Residual standard error: 3.696 on 170 degrees of freedom  
## Multiple R-squared: 0.2295, Adjusted R-squared: 0.2068   
## F-statistic: 10.13 on 5 and 170 DF, p-value: 1.669e-08

lm.beta(white.2)

## education employment happiness job\_choice job\_satis   
## 0.403758044 -0.073700034 -0.002896568 -0.303625714 0.442818796

summary(minority.1)

##   
## Call:  
## lm(formula = aa\_support ~ education + employment, data = minority)  
##   
## Residuals:  
## Min 1Q Median 3Q Max   
## -11.6510 -2.7331 0.4535 3.1570 10.7669   
##   
## Coefficients:  
## Estimate Std. Error t value Pr(>|t|)   
## (Intercept) 5.3107 1.0151 5.232 5.16e-07 \*\*\*  
## education 0.2647 0.2486 1.065 0.289   
## employment 1.0697 0.2479 4.314 2.79e-05 \*\*\*  
## ---  
## Signif. codes: 0 '\*\*\*' 0.001 '\*\*' 0.01 '\*' 0.05 '.' 0.1 ' ' 1  
##   
## Residual standard error: 4.455 on 161 degrees of freedom  
## Multiple R-squared: 0.2222, Adjusted R-squared: 0.2126   
## F-statistic: 23 on 2 and 161 DF, p-value: 1.638e-09

lm.beta(minority.1)

## education employment   
## 0.0987673 0.4001477

summary(minority.2)

##   
## Call:  
## lm(formula = aa\_support ~ education + employment + happiness +   
## job\_choice + job\_satis, data = minority)  
##   
## Residuals:  
## Min 1Q Median 3Q Max   
## -13.2071 -2.6187 0.6848 2.9531 9.5060   
##   
## Coefficients:  
## Estimate Std. Error t value Pr(>|t|)   
## (Intercept) 4.43767 1.01223 4.384 2.12e-05 \*\*\*  
## education 0.04644 0.24532 0.189 0.8501   
## employment 0.63099 0.26690 2.364 0.0193 \*   
## happiness 0.26387 0.26218 1.006 0.3157   
## job\_choice 0.25861 0.29922 0.864 0.3887   
## job\_satis 0.49471 0.28797 1.718 0.0878 .   
## ---  
## Signif. codes: 0 '\*\*\*' 0.001 '\*\*' 0.01 '\*' 0.05 '.' 0.1 ' ' 1  
##   
## Residual standard error: 4.284 on 158 degrees of freedom  
## Multiple R-squared: 0.2942, Adjusted R-squared: 0.2719   
## F-statistic: 13.17 on 5 and 158 DF, p-value: 1.003e-10

lm.beta(minority.2)

## education employment happiness job\_choice job\_satis   
## 0.01732864 0.23604759 0.09719706 0.09817709 0.19297729

The writes up suggest that for white participants, affirmitave action is supported multiple variables *R*2 = 0.14, *F*(2,173) = 14.29, *p* <.05. Although, not all variables contributed to the prediction. Greater perceived benefit of education (*b*\* = .45, *p* < .05) related to more support for affirmative action. Greater perceived benefit of employment did not relate to affirmative action (*b*\* = -.10, *p* = .50)

For minority participants, affirmitave action is supported multiple variables *R*2 = 0.22, *F*(2,161) = 23, *p* <.05. Although, not all variables contributed to the prediction. Greater perceived benefit of employment (*b*\* = .40, *p* < .05) is related to more support for affirmative action. Greater perceived benefit of education did not relate to affirmative action (*b*\* = .10, *p* = .29)

## Lastly, comparison of models against each other.

anova(minority.1,minority.2)  
anova(white.1,white.2)

I like to use modelCompare

modelCompare(minority.1,minority.2)

## SSE (Compact) = 3195.92   
## SSE (Augmented) = 2900.129   
## Delta R-Squared = 0.07198628   
## Partial Eta-Squared (PRE) = 0.09255275   
## F(3,158) = 5.371601, p = 0.001509836

modelCompare(white.1,white.2)

## SSE (Compact) = 2587.155   
## SSE (Augmented) = 2322.668   
## Delta R-Squared = 0.08773977   
## Partial Eta-Squared (PRE) = 0.1022308   
## F(3,170) = 6.452751, p = 0.0003660598

Looking at the hierarchical multiple regression, we find that the *R*2 change = 0.09, *p* < .05 for white participants. For minorities, the *R*2 change = 0.07,*p* < .05

## Semi-partial correlations

### Here is the formula that is used.

values <-summary(white.2)  
r2 <- values$r.squared[1]  
dfr<-values$df[2]  
dfr  
values <-summary(white.2)  
r2 <- values$r.squared[1]  
dfr<-values$df[2]  
t1 <- values$coefficients[2,3]  
t2 <- values$coefficients[3,3]  
t3 <- values$coefficients[4,3]  
t4 <- values$coefficients[5,3]  
t5 <- values$coefficients[6,3]  
sr1<-((t1^2)/dfr)\*(1-r2)  
sr2<-((t2^2)/dfr)\*(1-r2)  
sr3<-((t3^2)/dfr)\*(1-r2)  
sr4<-((t4^2)/dfr)\*(1-r2)  
sr5<-((t5^2)/dfr)\*(1-r2)

Here is the modelEffectSizes.

modelEffectSizes(white.2)

## lm(formula = aa\_support ~ education + employment + happiness +   
## job\_choice + job\_satis, data = white)  
##   
## Coefficients  
## SSR df pEta-sqr dR-sqr  
## (Intercept) 1520.1221 1 0.3956 NA  
## education 103.1508 1 0.0425 0.0342  
## employment 3.8943 1 0.0017 0.0013  
## happiness 0.0079 1 0.0000 0.0000  
## job\_choice 86.4512 1 0.0359 0.0287  
## job\_satis 152.0485 1 0.0614 0.0504  
##   
## Sum of squared errors (SSE): 2322.7  
## Sum of squared total (SST): 3014.4

modelEffectSizes(minority.2)

## lm(formula = aa\_support ~ education + employment + happiness +   
## job\_choice + job\_satis, data = minority)  
##   
## Coefficients  
## SSR df pEta-sqr dR-sqr  
## (Intercept) 352.7864 1 0.1085 NA  
## education 0.6577 1 0.0002 0.0002  
## employment 102.5918 1 0.0342 0.0250  
## happiness 18.5925 1 0.0064 0.0045  
## job\_choice 13.7110 1 0.0047 0.0033  
## job\_satis 54.1720 1 0.0183 0.0132  
##   
## Sum of squared errors (SSE): 2900.1  
## Sum of squared total (SST): 4109.0

varDescribe(white, Detail = 3)

## vars n mean sd median trimmed mad min max range skew  
## q1agree 1 176 2.72 1.62 2 2.55 1.48 1 7 6 0.74  
## q1fair 2 176 2.68 1.56 2 2.47 1.48 1 7 6 1.02  
## q1eff 3 176 2.87 1.42 3 2.78 1.48 1 6 5 0.46  
## q4 4 176 2.37 1.77 1 2.09 0.00 1 7 6 1.03  
## q5 5 176 2.27 1.60 2 2.01 1.48 1 7 6 1.07  
## q7 6 176 2.33 1.58 2 2.09 1.48 1 7 6 1.00  
## q8 7 176 2.33 1.73 2 2.01 1.48 1 7 6 1.25  
## q9 8 176 2.31 1.71 1 2.02 0.00 1 7 6 1.15  
## qa\* 9 176 1.00 0.00 1 1.00 0.00 1 1 0 NaN  
## p\_agree 10 176 2.72 1.62 2 2.55 1.48 1 7 6 0.74  
## p\_fair 11 176 2.68 1.56 2 2.47 1.48 1 7 6 1.02  
## p\_eff 12 176 2.87 1.42 3 2.78 1.48 1 6 5 0.46  
## education 13 176 2.37 1.77 1 2.09 0.00 1 7 6 1.03  
## employment 14 176 2.27 1.60 2 2.01 1.48 1 7 6 1.07  
## happiness 15 176 2.33 1.58 2 2.09 1.48 1 7 6 1.00  
## job\_choice 16 176 2.33 1.73 2 2.01 1.48 1 7 6 1.25  
## job\_satis 17 176 2.31 1.71 1 2.02 0.00 1 7 6 1.15  
## ethnicity\* 18 176 1.00 0.00 1 1.00 0.00 1 1 0 NaN  
## aa\_support 19 176 8.27 4.15 8 7.89 4.45 3 19 16 0.69  
## kurtosis se  
## q1agree -0.35 0.12  
## q1fair 0.46 0.12  
## q1eff -0.68 0.11  
## q4 -0.30 0.13  
## q5 0.04 0.12  
## q7 0.03 0.12  
## q8 0.59 0.13  
## q9 0.30 0.13  
## qa\* NaN 0.00  
## p\_agree -0.35 0.12  
## p\_fair 0.46 0.12  
## p\_eff -0.68 0.11  
## education -0.30 0.13  
## employment 0.04 0.12  
## happiness 0.03 0.12  
## job\_choice 0.59 0.13  
## job\_satis 0.30 0.13  
## ethnicity\* NaN 0.00  
## aa\_support -0.16 0.31

## Making the table

|  |  |
| --- | --- |
| header | yup |
| what | what |

why