PS231 B PS 2

Johanna Reyes

3/7/2022

# 1

## Exercise 2

The correlation between mass and elite tolerance scores is 0.52; between mass tolerance scores and repression scores, −0.26; between elite tolerance scores and repression scores,−0.42.

The equation for which we are computing the coefficients is Repression = Mass tolerance + Elite tolerance + .

For computing purposes, let’s define repression as U, mass tolerance as V and elite tolerance as X. Therefore, the converted equation is . In matrix form, , where M is the set of matrices, M = (V X).

Due to standardization, .

M’M= , therefore

M’M= n , and

M’U= = n= n

To compute the coefficients, we solve for

MM\_dt <- c(1,0.52,0.52, 1)  
M\_prime\_M <- matrix(MM\_dt,nrow=2)  
M\_prime\_M

## [,1] [,2]  
## [1,] 1.00 0.52  
## [2,] 0.52 1.00

M\_prime\_U <- matrix(c(-0.26,-0.42),ncol=1)  
M\_prime\_U

## [,1]  
## [1,] -0.26  
## [2,] -0.42

coef <- (solve(M\_prime\_M)) %\*% M\_prime\_U  
coef

## [,1]  
## [1,] -0.05701754  
## [2,] -0.39035088

The standardization of the variables allows us to use the methods in lecture 5; namely, we can apply the OLS assumptions and use matrix algebra to compute the coefficients and their variance.

## Exercise 3

=residual variance of , and using equation 8 in Freedman, p. 85, . Then, we multiple . Since there are two variables on the right hand side of equation (10):, and since the sample size is small, p=3.

a\_hat <- -0.06  
b\_hat <- -0.39  
n <- 36   
p <- 3   
sigma2\_hat <- 1 - (a\_hat)^2 - (b\_hat)^2 - 2\*(a\_hat\*b\_hat\*0.52)  
sd <- sqrt(sigma2\_hat\*(36/33))  
sd

## [1] 0.9457834

## Exercise 4

The formula for the standard errors of the coefficients is .

var <- as.matrix(sigma2\_hat \* solve(M\_prime\_M))  
se <- sqrt(var[1,1])  
se

## [1] 1.06012

var\_diff <- var[1,1] + var[2,2] - (2\*var[1,2])   
var\_diff

## [1] 3.416517

se\_diff <- sqrt(var\_diff)  
se\_diff

## [1] 1.848382

t\_a <- a\_hat/se  
t\_a

## [1] -0.05659737

t\_b <- b\_hat/se  
t\_b

## [1] -0.3678829

t\_diff <- (a\_hat-b\_hat)/se\_diff  
t\_diff

## [1] 0.1785345

Unlike Gibson, none of the t-ratios calculated imply significance for either coefficient OR their difference.

# 2

Given that the regression is standardized, the variance of the coefficients equals 1, so the sample size cancels out when we compute each matrix, as shown below.

= n