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 = β_1Mass tolerance + β_2Elite tolerance + δ.

For computing purposes, let's define repression as U, mass tolerance as V and elite tolerance as X. Therefore, the converted equation is $U = aV + bX + \delta$. In matrix form, $U = M\binom{a}{b} + \delta$, where M is the set of matrices, M = (V X).

Due to standardization, $r_{vx} = \frac{1}{n}\sum_{i=1}^{n} V_i X_i$.

M'M= $\begin{pmatrix} \sum_{i=1}^{n} V_i^2 & \sum_{i=1}^{n} V_i X_i \\ \sum_{i=1}^{n} V_i X_i & \sum_{i=1}^{n} X_i^2 \end{pmatrix}$, therefore

M'M= n $\begin{pmatrix} 1 & 0.52 \\ 0.52 & 1 \end{pmatrix}$, and

M'U= $\begin{pmatrix} \sum_{i=1}^{n} V_i U_i \\ \sum_{i=1}^{n} X_i U_i \end{pmatrix}$ = n$\begin{pmatrix} r_{VU} \\ r_{XU} \end{pmatrix}$= n$\begin{pmatrix} -0.26 \\ -0.42 \end{pmatrix}$

To compute the coefficients, we solve for $(M'M)^{-1}M'U = \begin{pmatrix} \hat{a} \\ \hat{b} \end{pmatrix}$

```
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

$\hat{\sigma}^2$=residual variance of $\hat{\delta}$, and using equation 8 in Freedman, p. 85, $\hat{\sigma}^2 = 1 - \hat{a}^2 - \hat{b}^2 - \hat{a}\hat{b}r_{VX}$. Then, we multiple $\sigma^2(\frac{n}{n-p})$. Since there are two variables on the right hand side of equation (10):$U = aV + bX + \delta$, 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 $SE_{\hat{\beta}} = \hat{\sigma}^2[X'X]^{-1}$.

```
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
```

```r
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.

$\hat{\beta} = (M'M)^{-1}M'U$

$M'M = n\begin{pmatrix} \mathbf{E}(V_i^2) & \mathbf{E}(V_iX_i) \\ \mathbf{E}(V_iX_i) & \mathbf{E}(X^2) \end{pmatrix}$

$M'M = n\begin{pmatrix} 1 & r_{VX} \\ r_{VX} & 1 \end{pmatrix}$

$\hat{\beta} = (M'M)^{-1}M'U = \mathrm{n}\begin{pmatrix} 1 & r_{VU} \\ r_{XU} & 1 \end{pmatrix} * \frac{1}{n\begin{pmatrix} 1 & r_{VX} \\ r_{VX} & 1 \end{pmatrix}}$