

# Lab 15 - Multivariate Regression & Interpretation

*Ramin Jabbarialghanab*

*November 30, 2017*

Complete the following exercises below and include all code used to find the answers. Knit together the PDF document and commit both the Lab 15 RMD file and the PDF document to Git. Push the changes to GitHub so both documents are visible in your public GitHub repository.

**1. Select a second explanatory variable from your dataset that you think has implications for the theoretical association of your focal relationship.**

```
library(readxl)
joint_unemployment_total <- read_excel("~/Desktop/Autumn 2017/Statistics 321/unemployment rate/joint_unemployment_total.xlsx")
model3 <- lm(value ~ gender + year, data = joint_unemployment_total)

summary(model3)
```

```
##
## Call:
## lm(formula = value ~ gender + year, data = joint_unemployment_total)
##
## Residuals:
##      Min       1Q   Median       3Q      Max
## -2.9121 -1.3316  0.3235  1.1000  2.3310
##
## Coefficients:
##              Estimate Std. Error t value Pr(>|t|)
## (Intercept)  131.50226   213.02026    0.617   0.546
## genderwomen    8.54444    0.79962   10.686 2.07e-08 ***
## year        -0.06029    0.10602   -0.569   0.578
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Residual standard error: 1.696 on 15 degrees of freedom
## Multiple R-squared:  0.8842, Adjusted R-squared:  0.8687
## F-statistic: 57.25 on 2 and 15 DF,  p-value: 9.516e-08
```

a. Describe the theoretical reasoning for selecting this variable.

I would examine the explanatory variable of gender over unemployment rate (value). Since, the rate can be change by year I decided to include explanatory variable of year to control the variable of year.

b. What type of relationship do you think this variable has with your focal variables? Given that, what do you expect to happen to your focal relationship when it is added to the model?

Acuually, I do not think that “year” has any strong relationship with my focal relationships, however, I thoguht that the effect of economic sunctions against Iran’s nuclear program could hve an intervening relationship with my focal relationship. So, I decided to control year to whether know I am correct.

c. Is it a continuous or categorical variable? What implications does this have for a multivariate regression equation?

It is a categorical variable but numeric, I used factor function and added year as a explnatory variable to the relationship. Controlling the variable only changed the relationship a little bit.

- d. Conduct a multivariate linear regression with this additional explanatory variable and save the model object. Print out the full results by calling `summary()` on your model object.
- e. Describe the results of the multivariate analysis, highlighting:
  - the apparent association between the control variable and the focal response variable
  - how the focal association changed when you incorporated the control variable
  - the implications of these results for your focal association

The relationship between the focal variables is changed from Multiple R-squared of 0.8817 in the bivariate regression model to Adjusted R-squared of 0.8687 in the multivariate analysis. So, the fits in the both bivariate and multivariate analysis. However, controlling year variable does not improve the relationship. For all the given years the association of the years with the focal relationship is negative and the negativity is so small for all years. With this result we can conclude that our focal variables have direct relationship.

- f. How well does this model fit the data? Is it an improvement over the bivariate model? Why or why not?

I think this model fits because Adjusted R-squared value equals 0.8687. Thus, the model explains 87% of the variation. It is changed little. So, the relationship did not change significantly. Bivariate explains more than multivariate.

**2. Select any additional variables you want to incorporate into your final model. For each additional variable added to the model answer the following questions:**

This part does not apply to me because I only three variables. There is no additional variable to use for here.

- a. Describe the theoretical reasoning for selecting this variable.
- b. What type of relationship do you think this variable has with your focal variables? Given that, what do you expect to happen to your focal relationship when it is added to the model?
- c. Is it a continuous or categorical variable? What implications does this have for a multivariate regression equation?
- d. Conduct a multivariate linear regression by adding one explanatory variable at a time and save the model objects. Print out the full results by calling `summary()` on each model object.
- e. Describe the results of the multivariate analysis, highlighting:
  - the apparent association between each additional control variable and the focal response variable
  - how the focal association changed when you incorporated each control variable
  - the implications of these results for your focal association
- f. How well does the full (all explanatory variables included) model fit? Are any of the other models you ran a better fit? Explain how you came to the conclusion you did.
- g. Select the model that you think best fits the data. Provide a brief synopsis of the analysis of your data using this model and describe the implications for the theoretical arguments you set out to test.