

GV903: Advanced Research Methods

Class 12 Introduction to R

In this session we will have a gentle introduction in [R](#). To make things much easier we will use the interface [RStudio](#). We will go over the R script from Moodle together.

Exercises:

1. Create a vector `x1` that takes 1000 values from a Normal distribution with mean 1 and variance 2, a vector `x2` that takes 1000 values from a Bernoulli distribution with probability 0.5.

2. Generate a vector `y` as

$$y_i = 10 + 2x_{1i} + 5x_{2i} + \epsilon_i$$

where $\epsilon_i \sim N(0, 3^2)$.

3. Put all the vectors together into a data frame and name it `lab12`.
4. Using the variables in this dataset fit a regression of y on x_1 and x_2 , plot a scatter plot of y and x_1 and the fitted lines from the model.
5. Test whether the model is correctly specified. What is the specification of the regression equation?
6. Include an interaction of x_1 and x_2 in the regression model and decide whether you should include this term in the model.
7. Include a quadratic term of x_1 in the regression model and decide whether you should include this term in the model.