

Questions and Exercises - Tutorial #1

Ettore Gallo

April 17, 2018

Review of Statistics

1. What is an Expected value? Define the expected value $[E(x)]$ in the cases:
 - (a) x is a binomial random variable;
 - (b) x is a continuous random variable.Briefly recall the properties of expected values.
2. Define Variance $[Var(x) = \sigma_x^2]$ and standard deviation $[\sigma_x]$. Briefly recall the properties of variance.
3. Define the covariance between two random variables $cov(x, y)$. Briefly recall the properties of covariance.
4. Define the t-statistic and briefly recall what is hypothesis testing (H_0 vs. H_a).
5. Recall the reason why \bar{y} is used to estimate μ_y (*Hint*: Central Limit Theorem / Law of large numbers). What does it mean that the estimator is unbiased, that is consistent and that is the "least square" estimator?

Regression Analysis

1. Recall the Gauss-Markow assumptions (MLR.1-MLR.5). What does it mean that under these assumptions the OLS estimator is BLUE?
2. Let *kids* denote the number of children ever born to a woman, and let *educ* denote years of education for the woman. A simple model relating fertility to years of education is:

$$kids = \beta_0 + \beta_1 educ + u \quad (1)$$

where u is the unobserved error.

- (i) What kinds of factors are contained in u ? Are these likely to be correlated with level of education?
 - (ii) Will a simple regression analysis uncover the *ceteris paribus* effect of education on fertility? Explain.
3. In the simple linear regression model $y = \beta_0 + \beta_1 x + u$, suppose that $E(u) \neq 0$. Letting $\alpha_0 = E(u)$, show that the model can always be rewritten with the same slope, but a new intercept and error, where the new error has a zero expected value.
 4. Suppose that a researcher, using data on class size (*CS*) and average test scores from 100 third-grade classes, estimates the OLS regression,

$$TestScore = 520.4 - 5.82 \times CS, R^2 = 0.08, SER = 11.5 \quad (2)$$

- a. A classroom has 22 students. What is the regression's prediction for that classroom's average test score?
- b. Last year a classroom had 19 students, and this year it has 23 students. What is the regression's prediction for the change in the classroom average test score?
- c. The sample average class size across the 100 classrooms is 21.4. What is the sample average of the test scores across the 100 classrooms? (*Hint*: Review the formulas for the OLS estimators.)
- d. What is the sample standard deviation of test scores across the 100 classrooms? (*Hint*: Review the formulas for the R^2 and SER)