Economics 468

November 14, 2016 **R. Davidson**

Assignment 3

1. Consider the linear regression model with n observations,

$$\boldsymbol{y} = \delta_1 \boldsymbol{d}_1 + \delta_2 \boldsymbol{d}_2 + \boldsymbol{u}, \;\; \boldsymbol{u} \sim \mathrm{IID}(\mathbf{0}, \sigma^2 \mathbf{I}). \tag{1}$$

The two regressors are dummy variables, with every element of $\boldsymbol{d}_2$ equal to 1 minus the corresponding element of $\boldsymbol{d}_1$. The vector $\boldsymbol{d}_1$ has n_1 elements equal to 1, and the vector $\boldsymbol{d}_2$ has $n_2 = n - n_1$ elements equal to 1.

The parameter of interest is $\gamma \equiv \delta_2 - \delta_1$. Find the standard deviation of $\hat{\gamma}$ (that is, the square root of its true variance) and write it as a function of σ, n, and either n_1 or n_2.

Suppose the data for regression (1) come from an experiment that you design and administer. If you can only afford to collect 800 observations, how should you choose n_1 and n_2 in order to estimate γ as efficiently as possible?

2. Consider the linear regression model with n observations,

$$\boldsymbol{y} = \beta_1 + \beta_2 \boldsymbol{d} + \boldsymbol{u}, \;\; \boldsymbol{u} \sim \mathrm{NID}(\mathbf{0}, \sigma^2 \mathbf{I}). \tag{2}$$

The only regressor here is a dummy variable, with each element equal to 1 for n_1 observations and equal to 0 for the remaining $n - n_1$ observations.

First, find the standard error of $\hat{\beta}_2$ as a function of n, n_1, and σ. Student's t distribution with n degrees of freedom can be defined as the distribution of the ratio of a standard normal variable to the square root of an independent χ^2 variable with n degrees of freedom divided by n; schematically,

$$t_n = \frac{\mathrm{N}(0,1)}{\sqrt{\chi^2_n / n}}.$$

The non-central t distribution, with non-centrality parameter (NCP) λ, can be defined similarly as $\mathrm{N}(\lambda, 1)/\sqrt{\chi^2_n/n}$. Suppose that the null hypothesis that $\beta_2 = 0$ is tested by a t test. If we denote by β_2 the true value of the parameter, find the distribution of the t statistic as a function of n, n_1, β_2, and σ^2.

3. Consider the linear regression model

$$\boldsymbol{y} = \boldsymbol{X\beta} + \boldsymbol{u}, \quad \boldsymbol{u} \sim \mathrm{N}(\mathbf{0}, \sigma^2 \mathbf{I}), \quad \mathrm{E}(\boldsymbol{u} \,|\, \boldsymbol{X}) = \mathbf{0},$$

where $\boldsymbol{X}$ is an $n \times k$ matrix. If σ_0 denotes the true value of σ, how is the quantity $\boldsymbol{y}^\top \boldsymbol{M_X y}/\sigma_0^2$ distributed? Use this result to derive a test of the null hypothesis that $\sigma = \sigma_0$. Is this a one-tailed test or a two-tailed test?

Suppose that $k = 3$. Plot the power function for this test at the .05 level for the null hypothesis that $\sigma = 1$ over the interval $0.5 < \sigma < 2.0$ for three values of the sample size, namely, $n = 13$, $n = 23$, and $n = 43$. **Hint:** This exercise does not require any simulations, but it does require you to calculate the cumulative χ^2 distribution function many times and its inverse a few times.

4. Consider again the data on consumption (or saving) and income used for Assignment 1. Now consider the following dynamic model

$$c_t = \alpha + \beta c_{t-1} + \gamma_0 y_t + \gamma_1 y_{t-1} + u_t,$$

where $c_t = \log C_t$ and $y_t = \log Y_t$. Compute a t statistic for the hypothesis that $\gamma_0 + \gamma_1 = 0$. On the basis of this test statistic, perform an asymptotic test, a parametric bootstrap test, and a semiparametric bootstrap test using rescaled residuals.