

Political Science 239

Problem Set 7

Due date: Wednesday, October 25th, 2006

For this problem set you will use simulated data available in the file PSet7.RData, where you will find the data frames "exercise1" and "exercise3". You can find the data on the usual website <http://are.berkeley.edu/~rocio/teaching.html>. You will also find the file "ProblemSet7_data.R", which has the code that was used to generate the data. You should take a look at this code to get an idea of what the true population parameters are.

Exercise 1 *Using the simulated random variables x_1 and x_2 in data frame "exercise1", bootstrap the difference in means (draw 1,000 bootstrap samples). Calculate the bootstrapped 95% confidence interval and the bootstrapped standard error of your estimator.*

Exercise 2 *Is the difference in means between x_1 and x_2 significantly different from zero?*

Exercise 3 *Using the simulated random variables $Y, X_1, X_2,$ and X_3 in data frame "exercise3", calculate bootstrapped 95% confidence intervals and bootstrapped standard errors of the OLS estimators of $\alpha, \beta_1, \beta_2,$ and β_3 in the model $Y = \alpha + \beta_1 X_1 + \beta_2 X_2 + \beta_3 X_3 + \epsilon$. (Draw 1,000 bootstrap samples and use non-parametric bootstrapping). Also, use bootstrapping to calculate the estimated bias of each of the OLS estimators.*

Exercise 4 *Now use bootstrapping to test the null hypothesis that $\beta_1 = \beta_2$. Do you reject this null hypothesis? Plot the bootstrap distribution of your test-statistic and show whether the full sample value of the test-statistic falls in the critical region (use 5% significance level).*