SPRING 1997

Statistics 200B
Lab 3: Further Adventures in Fort Evans

This lab assignment consists of two parts. In the first part you will use stratified random sampling to try to get a more accurate idea of the average family income. Suppose you find squirreled away in the filing cabinet of a fellow city bureaucrat the information that the population sizes of the six regions are

Region 1 36,096

Region 2 24,670

Region 3 37,051

Region 4 21,430

Region 5 14,382

Region 6 21,568

Total 155,197

How do these numbers square with the confidence intervals you produced in the last lab?

Take another sample of total size 500, but allocated proportionally among the six regions (strata). From it estimate the proportion of unemployed adults seeking work, give an estimate of the standard error of your estimate, and give a 95% confidence interval for the population mean. Do the same for the total number. Compare these to the corresponding results you found in Lab 2. Explain the difference in the estimated standard errors.

For the second part of this assignment, use sampling to investigate one of the questions below, or pose a similar question and investigate it.

- What is the relation of a person's age and his or her family income?
- Is the age distribution of the population similar across the different regions, or races, or ethnicities?

- How does the marital status of adults differ across regions?

- How do family incomes differ across regions, races, or ethnicities?

Summarize your sampling analyses and observations. Do your results correspond with your intuition about these social and economic relationships? Why or why not?

NOTE: The function `regionsamp`, like `sampler`, is a homemade function that you can run from your account. It does not appear in the general S documentation.