

Case-based Social Statistics II

CSSS 322

Professor: Mark S. Handcock

Homework 2

Due Thursday, April 18, 2002

Problems to be handed in:

- 1) Based on responses of 1467 subjects in the General Social Surveys in the mid-1980s, a 95% confidence interval for the mean number of close friends equals (6.8, 8.0). Which of the following interpretations of this interval is (are) correct?
 - a) We can be 95% confident that \bar{X} is between 6.8 and 8.0.
 - b) We can be 95% confident that μ is between 6.8 and 8.0.
 - c) Ninety-five percent of the values of $X =$ the number of close friends (for this sample) are between 6.8 and 8.0.
 - d) If random samples of size 1467 were repeatedly selected, then 95% of the time \bar{X} would be between 6.8 and 8.0.
 - e) If random samples of size 1467 were repeatedly selected, then in the long run 95% of the confidence intervals formed would contain the true value of μ .
- 2) A survey is taken to estimate the mean annual family income for families living in public housing in Chicago. For a random sample of 30 families, the annual incomes (in hundreds of dollars) are as follows:

83	90	77	100	83	64	78	92	73	122
96	60	85	86	108	70	139	56	94	84
111	93	120	70	92	100	124	59	112	79

The data is entered into Datatools on the course website.

- a) Using Datatools, construct a stem-and-leaf plot of the incomes. What do you predict about the shape of the population distribution?
 - b) Construct and interpret estimates of μ and σ , the mean and standard deviation of the family incomes of all families living in public housing in Chicago.
 - c) Construct and interpret a 95% confidence interval for μ .
 - d) Construct a 95% confidence interval for μ . Interpret the interval and compare it to the one in part (c).
- 3) A study is conducted of a geographic distribution of the residences of the employees at a large factory, in order to determine the suitability of initiating busing to that factory. One variable considered is the distance the employee lives from the factory. For a random

sample of 100 employees, the mean distance is 6.3 miles and the standard deviation is 4.0 miles.

- a) Find and interpret a 90% confidence interval for the mean residential distance from the factory for all employees.
 - b) Find and interpret a 90% *prediction* interval for the residential distance from the factory for a randomly chosen employee.
 - c) About how large a sample would have been adequate if we merely needed to estimate the mean to within 1.0 miles, with 95% confidence?
- 4) In the 1991 General Social Survey, respondents were asked whether people convicted of murder should receive the death penalty. 1078 responded yes and 336 responded no. Construct a 99% confidence interval for the proportion of American adults who would answer yes. Interpret this interval. Can you conclude that more than half of all American adults would answer yes? Explain briefly.

Extra Credit Problem:

- 5) The General Social Survey is a means of collecting information about the U.S. population. It is an annual survey that is based on randomly sampling people and interviewing them about the economic and social aspects of their lives. Suppose that you took a random sample of 100 families to find out about the weekly income of families in the U.S. Suppose you find that the mean income of families in the sample was \$874 and the sample standard deviation of the incomes was \$186.
- a) Find the 95% confidence interval for the population mean.
 - b) Suppose a local urban planning organization is interested in the total income for all families in Western Washington. If there are 543,000 families in Western Washington, find a 95% confidence interval for the *total* weekly income for all these people. Briefly explain to the meaning of your answer.
 - c) Suppose that the urban planning organization undertook a complete accounting of the entire population of families, and found a total of \$450,714,000. Your job is on the line. What do you say to him?