Introduction to Statistics (MAT/SST 115.03 2008S)

Class 04: An Introduction to R, Continued

Held: Wednesday, 30 January 2008

Summary: We continue our explorations of the R environment for statistics.

Notes:

- Can we design a study that will help us predict which stairwell will be open?
- I may ask questions as a way of taking attendance.
- Cassie is setting up office hours Sunday, Tuesday, and Thursday evenings. We're still working on where.

- Handouts: Getting Started with R.
- Due: Topic 3.

Overview:

- Detour: Comments on Topic 3.
- Short review.
- Lab, continued.
- Reflection.

Quick Review of Topic 3

Some questions for you:

- What is a *population*?
- What is a *sample*?
- If we care about populations, why do we often look at samples, rather than complete populations?
- What is a *parameter*?
- What is a *statistic*?
- What makes a sample *representative*?
- What makes a process for generating samples *biased*?
- What is an *observational study*?
- What are some difficulties of observational studies?
- What are the alternatives to observational studies?

Some questions to ask when you read studies (cribbed from p. 43)

- To what population can you reasonably generalize the results of the study?
- Can you reasonably draw a cause-and-effect connection between the explanatory variable and the response variable?

Observation:

- Most statisticians believe you cannot draw cause-and-effect connections in observational studies.
- A surprising number of economists believe that you *can* design observational studies that produce cause-and-effect conclusions.

R, Continued

- Quick review
  - Four key data types: Numbers, strings, vectors, frames
  - Can do computations
  - Can build graphs
  - Can build misleading graphs (example of barplot from data that needs to be tabulated)
- Lab!
- Reflection
