

Math 218 Mathematical Statistics

Prof. D. Joyce, Clark University

21 Jan 2009

Due Friday. From page 101, exercises 1, 3, 4, 7, 8, 9, 12, 13, 15.

Due Monday. From page 105, exercises 19, 20,

For next meeting. Read Chapter 4 on summarizing and exploring data.

Last meeting. Began a discussion of collecting data. Types of statistical studies. Control groups in a comparative study. Sample surveys. Prospective and retrospective studies. Basic sampling designs. Simple random sampling. Other kinds of random sampling.

Today. Experimental studies. In the terminology of experimental design, purposes of an experiment are to

- evaluate how a set of *factors* (that is, predictor variables) affect one or more response variables
- screen out unimportant factors
- select values for the controllable factors to maximize or minimize a response variable
- fit a model that can be used to make predictions or adjust controllable factors to determine particular values for one or more response variables.

The first purpose, of course, is the primary purpose.

Those factors that can be controlled are also called *treatment factors*. Other factors, some of which may not even be known, are called *noise factors* or *nuisance factors*. The values of a factor are

also called its *levels*. A *treatment* is a particular combination of levels for all the treatment factors.

Treatments are applied to *experimental units* (i.e., subjects), and measuring the responses. A *treatment group* is a collection of experimental units who all get the same treatment, and the result of that treatment on the treatment group is called a *run*.

See example 3.8 which illustrates these concepts.

Strategies to reduce experimental error variation. There are various kinds of experimental errors. One is *systematic error* due to differences between experimental units; various units can be biased to give nonrepresentative responses. Another is *random error* which is the variability of responses that an individual unit will give. A third is *measurement error* which means the same response may be measured differently if a new set of measurements are made.

See example 3.9.

There are some standard strategies that are used to reduce systematic error.