The Basic Practice of Statistics
Ninth Edition

David S. Moore William I. Notz

Chapter 9
Producing Data:
Experiments

Lecture Slides

In Chapter 9, we cover …

- Observation versus experiment
- Subjects, factors, and treatments
- How to experiment badly
- Randomized comparative experiments
- The logic of randomized comparative experiments
- Cautions about experimentation
- Matched pairs and other block designs

Recalling: explanatory variable versus response variable

In a study to relate two conditions, researchers often define one as the ***explanatory variable*** and other as the outcome or ***response variable***.

In a study to determine whether surgery or chemotherapy results in higher survival rates for a certain type of cancer, whether the patient survived is one variable, and whether he or she received surgery or chemotherapy is the other.

Which is the explanatory variable and which is the response variable?

- **Explanatory variable** is one that may explain or may cause differences in a **response variable** (or outcome variable).

Randomized Experiment versus Observational Studies

- **Randomized experiment**: create differences in the explanatory variable and examine results (response variable).
- **Observational study**: observe differences in the explanatory variable and notice whether these are related to differences in the response variable.

Two reasons why we must sometimes use an observational study instead of an experiment …

1. It is **unethical** or **impossible to assign people** to receive a specific treatment.(e.g, smoking)
2. Certain explanatory variables are **inherent traits** and cannot be randomly assigned. (e.g., handedness)

Observational study versus experiment

- In contrast to observational studies, experiments don't just observe individuals or ask them questions. They actively impose some treatment in order to measure the response.

OBSERVATIONAL STUDY VERSUS EXPERIMENT

- An observational study observes individuals and measures variables of interest, but it does not attempt to influence the responses. The purpose of an observational study is to describe some group or situation.

- An experiment, on the other hand, deliberately imposes some treatment on individuals to observe their responses. The purpose of an experiment is to study whether the treatment causes a change in the response.

- When our goal is to understand cause and effect, experiments are the preferred source for fully convincing data.

- The distinction between observational studies and experiments is one of the most important in statistics.

Confounding

- A lurking (or confounding) variable is a variable that is not among the explanatory or response variables in a study but that may influence the response variable.

- Two variables (an explanatory variable and a lurking variable) are confounded when their effects on a response variable cannot be separated from each other.

- Confounding is a **bigger problem in observational studies**. Observational studies of the effect of one variable on another often fail because of confounding between the explanatory variable and one or more lurking variables.
- Well-designed experiments take steps to avoid confounding.

Confounding - example

Study of the relationship between smoking during pregnancy and child's subsequent IQ a few years after birth.

- **Explanatory variable**: whether or not the mother smoked during pregnancy
- **Response variable**: subsequent IQ of the child
- Women who smoke may also have poor nutrition, lower levels of education, or lower income.
- **Possible Confounding Variables**: Mother's nutrition, education, and income.

Subjects, factors, and treatments

- An experiment is a statistical study in which we actually do something (apply a **treatment**) to people, animals, or objects to observe the **response**. Here is the basic vocabulary of experiments.

- The individuals studied in an experiment are often called subjects, particularly when they are people.
- The explanatory variables in an experiment are often called factors.
- A treatment is any specific experimental condition applied to the subjects. If an experiment has more than one factor, a treatment is a combination of specific values of each factor.

How to experiment badly (1of 3)

- *Experiments don't guarantee good data.*
- Good designs are essential for effective experiments, just as they are for sampling.

Example: A college regularly offers a review course to prepare students for the GMAT. This year, it will offer only an online version of the course.

Students → Online Course → GMAT Scores

- Suppose the results were 10% higher than the long-term average for those who took the *classroom* review course.
- **Can we conclude that the online course is more effective?**

How to experiment badly (2 of 3)

- No, we cannot conclude that the online course is more effective.
- A closer look at the course reveals that students in the current year are older and more mature than in previous years. This may have contributed to higher GMAT scores.
- The effect of online versus in-person class is confounded with the effect of lurking variables.
- This is an example of **an uncontrolled experiment**.
- **Solution:** randomized students in the same year to online and in-person class groups.

How to experiment badly (3 of 3)

- Many laboratory experiments use a design like the one in the online GMAT course example:

Subjects → **Treatment** → **Measure Response**

- In the laboratory environment, simple designs often work well.
- Field experiments and experiments with animals or people deal with more variable conditions.
- *Outside the laboratory, badly designed experiments often yield worthless results because of confounding with lurking variables.*

Randomized comparative experiments (1 of 2)

- The remedy for confounding is to perform a comparative experiment in which some individuals receive one treatment and similar units receive another. Most well-designed experiments compare two or more treatments.
- Some experiments may include a control group that receives an inactive treatment (*placebo*) or an existing baseline treatment.

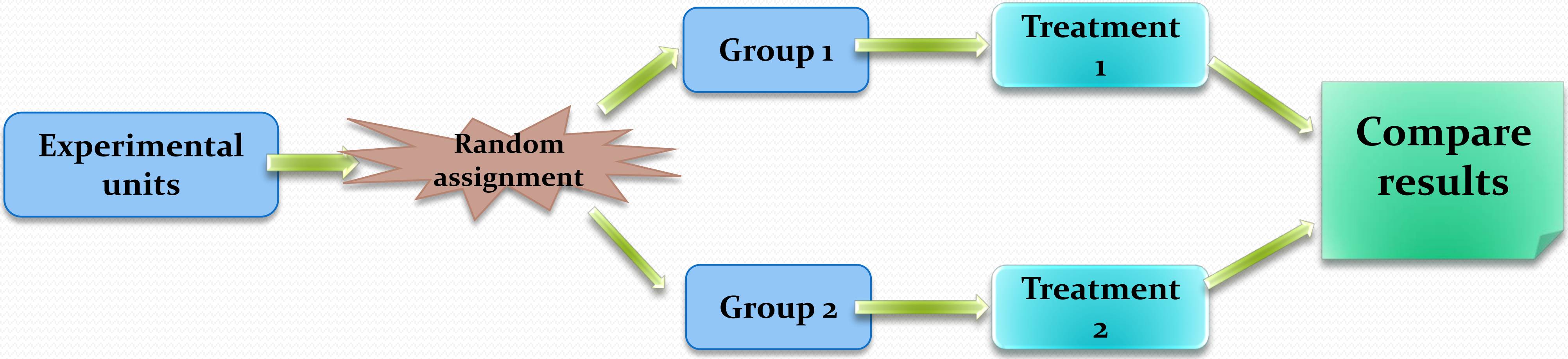
- An experiment that uses both comparison of two or more treatments and random assignment of subjects to treatments is a randomized comparative experiment.

Randomized comparative experiments (2 of 2)

- Comparison alone isn't enough. If the treatments are applied to groups that differ greatly, bias will result. The solution to the problem of bias is random assignment.
- In an experiment, random assignment means that individuals are assigned to treatments at random—that is, using some sort of chance process.

COMPLETELY RANDOMIZED DESIGN

- In a completely randomized experimental design, all the subjects are allocated at random among all the treatments.



The logic of randomized comparative experiments (1 of 2)

- Random assignment of subjects forms groups that should be similar in all respects before the treatments are applied.
- A comparative experiment with randomization ensures that influences other than the experimental treatments operate equally on all groups.
- Differences in average response must be due either to the treatments or to the play of chance in the random assignment of subjects to the treatments.

The logic of randomized comparative experiments (2 of 2)

PRINCIPLES OF EXPERIMENTAL DESIGN

- The basic principles of statistical design of experiments are:

1. Control—restrict the effects of lurking variables on the response, most simply by comparing two or more treatments.
2. Randomization—use chance to assign subjects to treatments.
3. Replication—use enough subjects in each group to reduce chance variation in the results.

- An observed effect so large that it would rarely occur by chance is called statistically significant.
- A statistically significant association in data from a well-designed experiment ***does imply causation***.

Cautions about experimentation

- The logic of a randomized comparative experiment depends on our ability to treat all the subjects the same in every way, except for the actual treatments being compared.
- Many medical experiments are, for example, "placebo controlled."
- A placebo is a dummy treatment that is as similar to the treatment as possible but contains no "active ingredient."

DOUBLE-BLIND EXPERIMENTS

- In a double-blind experiment, neither the subjects nor those who interact with them and measure the response variable know which treatment each subject is receiving.

- Even these have been criticized for **lack of realism**.

Control Groups, Placebos, and Blinding

Control Groups

- Handled identically to the treatment group(s) in all respects, except that they don't receive the actual treatment.

Placebos

- Research shows people also respond to placebos – looks like the real drug but has no active ingredients.
- Randomly assign some patients to receive the drug and others to receive a placebo, without telling which they are receiving.

Blinding

- ***Double-blind***: neither the participant nor the researcher taking measurements know who had which treatment.
- ***Single-blind***: only one of the two (participant or researcher) knows which treatment the participant was assigned.

Example: Blindly Lowering Cholesterol

Which lowers cholesterol more?

Special diet (*portfolio*) versus drug (*lovastatin*)?

Details:

- **Three treatments**: portfolio diet, low-fat diet with lovastatin, low-fat diet with placebo.
- The 46 volunteers were **randomized** by a statistician using a random number generator.
- **Blinding**: researchers and participants both blind as to which drug (lovastatin or placebo) people in those two groups were taking. However, participants and dieticians could not be blind to what the participants were eating. Lab staff evaluating cholesterol measurements were blinded to the treatment.

Matched pairs

- The idea of this experimental design is to create or use matching pairs of similar experimental units for comparing two treatments.
- A matched pairs design is a randomized experimental design in which—within each matching pair of similar subjects—chance is used to determine which subject gets each treatment.
- Sometimes, a “pair” in a matched pairs design consists of a single unit that receives both treatments. Because the order of the treatments can influence the response, chance is then used to determine which treatment is applied first for each unit.

Block design

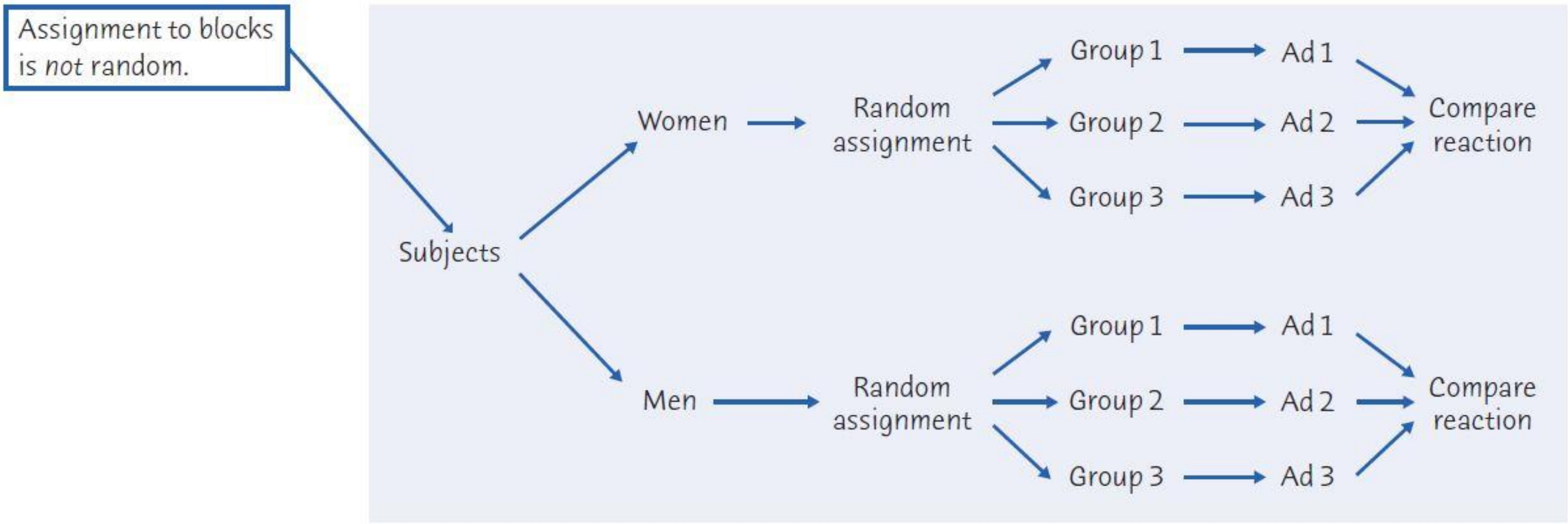
- Matched pairs are one kind of block design, with each pair forming a block.

- A block is a group of individuals that are known, before the experiment, to be similar in some way that is expected to affect their response to the treatments.
- In a block design, the random assignment of experimental units to treatments is carried out separately within each block.

- A wise experimenter will form blocks based on the most important unavoidable sources of variability among the subjects.
- Randomization will then average out the effects of the remaining lurking variables and allow an unbiased comparison of the treatments.

Block design—illustration

Example 9.10 – women and men respond differently to advertising



Block Design – another example

- Suppose there are 3 flower beds (blocks) with each bed containing 20 plants.
- Goal: assess the effect of water temperature and fertilizer type in plant growth.
- Two water temperatures: 20℃ and 30℃
- And two fertilizer types: A and B
- Four treatment combinations → 1) 20℃ with fertilizer A, 2) 20℃ with fertilizer B, 3) 30℃ with fertilizer A and 4) 20℃ with fertilizer B)

Randomization of plants to treatments is performed separately in each block (average of 5 plants per treatment in each block):

Block 1	Block 2	Block 3
Plant 1: Treatment 3)	Plant 1: Treatment 1)	Plant 1: Treatment 2)
Plant 2: Treatment 2)	Plant 2: Treatment 4)	Plant 2: Treatment 3)
Plant 3: Treatment 4)	Plant 3: Treatment 2)	Plant 3: Treatment 4)
Plant 4: Treatment 4)	Plant 4: Treatment 1)	Plant 4: Treatment 4)
And so on	And so on	And so on