Stats: Modeling the World

Chapter 12: Experimental Design
The Women’s Health Study randomly assigned nearly 40,000 women over the age of 45 to receive either aspirin or a placebo for over 10 years to examine the effect of aspirin on cancer risk to healthy women. This long-term trial was best conducted as?

- A) Census
- B) Observational Study
- C) Randomized comparative experiment
- D) A single-blind randomized comparative experiment
- E) A double-blind randomized comparative experiment
Warm - Up

• The Women’s Health Study randomly assigned nearly 40,000 women over the age of 45 to receive either aspirin or a placebo for over 10 years to examine the effect of aspirin on cancer risk to healthy women. This long-term trial was best conducted as?

• E) A double-blind randomized comparative experiment:

• Because there was a treatment imposed within randomized groups, this was a randomized comparative experiment. It was best conducted as a double-blind, where neither the women nor the evaluators knew which individuals were taking aspirin and which were getting a placebo

Best Experiments:
• Randomized
• Comparative
• Double-blind
• A control group
Matching: Observational Studies

• Matching: In observational study either retrospective or prospective, subjects who are similar in ways not under study may be matched and then compared with each other on the variables of interest.

• E.g. (age, gender, nationality, economic status)

• Matching, Stratified Sampling, and Blocking are all the same concept just applied to different methods.
Grouping individuals of interest that are similar and randomize within those groups as a way to remove unwanted variation. Be careful to keep the terms straight.

Experiment
  Blocking

Sample
  Stratify

Observational Study
  Matching
Confounding

• Confounded: when the levels of one factor are associated with the levels of another factor the

• Our goal is to measure how the factor affects the response variable. But when another variable is intertwined with the factor we can’t tell how much each of these variables is effecting the response variable.
Confounding can also arise from a badly designed multifactor experiment. Here’s a classic. A credit card bank wanted to test the sensitivity of the market to two factors: the annual fee charged for a card and the annual percentage rate charged.

<table>
<thead>
<tr>
<th>Annual Rate</th>
<th>Annual Fee</th>
</tr>
</thead>
<tbody>
<tr>
<td>High</td>
<td>HY, HN</td>
</tr>
<tr>
<td>Low</td>
<td>LY, LN</td>
</tr>
</tbody>
</table>

It sent out 50,000 offers with **low rate and no fee** and 50,000 offers with **higher rate and a fee**. No surprise people preferred the low-rate no fee card. Because the large sample size the bank was able to estimate the difference precisely.

But the question the bank really wanted to answer was “How much change was due to the rate, and how much as due to the fee?” Unfortunately there’s no way to separate out the two effects. If the bank had sent out all four possible different treatments it could have learned about both factors and could have seen what happens when the two factors occur in combination.
Confounding Example:

We wish to see if detergent A cleans better than detergent B.
• Detergent A is used in one washing machine
• Detergent B is used in a different machine.

At the end of this experiment, we have confounding. That is, we cannot tell if a difference in the cleanliness of the clothes is caused by the detergent or by the quality of the machine that was used.
Confounding Example:

After many dogs and cats suffered from upset stomachs caused by contaminated foods, we’re trying to find out whether a newly formulated pet food is safe. Our experiment will feed some animals the new food and others a food known to be safe, and a veterinarian will check the response.

Question: Why would it be a bad design to feed the test food to some dogs and the safe food to some cats?

Answer: This would create confounding. We would not be able to tell whether any differences in animals health were attributable to the food they had eaten or to differences in how the two species responded.

Recall Experimental Principle:

(I) Control: Sources of variation other than the factors by making conditions as similar as possible for all treatment groups.
40. Gas mileage: Do cars get better gas mileage with premium instead of regular unleaded gasoline? It might be possible to test some engines in a laboratory, but we’d rather use real cars and drivers in real day-to-day driving, so we get 20 volunteers. Design the experiment.
This experiment has 1 factor (type of gasoline), at 2 levels (premium and regular), resulting in two treatments. The response variable is gas mileage. An experiment diagram for a simple design appears above. Randomly assign each of the 20 volunteers to the premium or regular groups. Ask them to keep driving logs (the number of miles driven and the gallons of gasoline) for one month. Compare the differences in the fuel economy for the two groups.
What Could Effect Car Mileage?

- Size of the Car
- Car Brand
- # of Passengers

In our case these differences could potentially effect how much mileage we get from the cars. Blocking isolates the variability due to the differences between the blocks.
Car Blocked Design

- This experiment has 1 factor (type of gasoline), at 2 levels (premium and regular), resulting in two treatments. The response variable is gas mileage. An experiment diagram for a blocked design. The car size may have some effect on gas mileage. To reduce differences between car sizes we will divide the volunteers based on car size into 3 blocks (large, medium, small). Randomly within each block assign each of the volunteers to the premium or regular groups. Ask them to keep driving logs (the number of miles driven and the gallons of gasoline) for one month. Compare the differences in the fuel economy for the two groups.
Insomnia (pg. 327): An experiment investigating a dietary approach to treating bipolar disorder. Researchers randomly assigned 30 subjects to two treatment groups, one group taking a high dose of mega-3 fats and the other a placebo.

a) Why was it important to randomize in assigning the subjects to the two groups? Subjects’ responses might be related to other factors, like diet, exercise, or genetics. Randomization should equalize the two groups with respect to unknown factors.

a) What would be the advantage and disadvantages of using 100 subjects instead of 30? More subjects would minimize the impact of individual variability in the responses, but the experiment would become more costly and time-consuming. The size of the sample is important not its proportion to the population of interest.
#17. Torn ACL: Having at least one **15-minute warm-up session per week** resulted in a drastic reduction in tears in the ACL. In a study involving about 45000 **adolescent girls’ soccer players in Sweden**, one group was **randomly assigned** to warm up with neuromuscular exercise session. This group had 64% fewer ACL tears than the control group.

a) Is this an experiment or an observational study? Explain why
   This is an experiment. Subjects were randomly assigned to treatments.

b) Identify the treatments in this study. What is the response variable?
   **Factors:** Warm Up
   **Levels:** Y or N
   **Treatments:** Warm Up Yes and Warm Up No
   (What are all the combinations of my factor levels)
   **Response:** ACL tears
   (What information are we comparing at the end)

c) Give one statistical advantage of using only Swedish girls who played soccer in this study (What is my population of interest? What variation am I controlling?)
   **Population:** Adolescent Girl Soccer Players in Sweden
   Less variation in ACL tears would be expected among players of the same gender in the same sport, and similar age which would allow a difference between the treatments to be detected more easily

d) Give one statistical disadvantage of using only Swedish girls who played soccer in this study. (What is my population of interest? What further generalization can’t I conclude?)
   The conclusions of the study cannot be generalized to males, females in other countries, nor to players of other sports
Assessments:

• Quiz: Chpt.11: Sample Surveys & Chpt. 12: Experiments & Observational Studies: **Friday 9/16/16 & Thursday 9/15/16**

• **Review Day: Monday 9/19/16 & Tuesday 9/20/16**

• Exam 1: Chpt1: Chpt.10: Chpt.11: Chpt.12: **Wednesday 9/21/16 & Thursday 9/22/16**
Homework & Classwork

• Homework: pg. 327 #14, 39, 24

• Classwork: pg. 327 # 18, 23, 35, 38