Chapter 11: Experiments and Observational Studies p 318

Observation vs Experiment

An **observational study** observes individuals and measures variables of interest but does not attempt to influence the response.

An **experiment** imposes a treatment on individuals in order to observe their response.

- Observational studies are valuable for discovering trends and possible relationships.
- However, it is not possible for observational studies to demonstrate a causal relationship
- When our goal is to understand the cause and effect, experiments are the only source of fully convincing data.

Example

How do you find out if exercise helps insomnia?

- look at a bunch of people, find out if they exercise and how much, ask them to rate their insomnia.
- Suppose the people who exercise more suffer less from insomnia. Can you conclude that people who suffer from insomnia should be recommended to exercise?
  - *maybe: but this is only an association, not cause and effect.*

- This is an example of an observational study.
- Assesses association but not cause and effect (like correlation).

Observational studies p318

– Commonly used

– May help identify variables that have an effect
  – but does not prove cause and effect

– ***Retrospective* study** "looking back", like one above:
  – measure exercise and insomnia from historical records.

– ***Prospective* study** "looking forward":
  – identify subjects in advance, collect data as events happen.

– Are data from the past even accurate?
– Which is better, retrospective or prospective?

- prospective: usually less confounding and bias. Outcome needs to be a common one. Takes a long time.

- retrospective: easier to obtain enough data for rare events. Takes less time to do. More concerns about bias/confounding

Experiments p319

How do we establish cause and effect?

- need to randomly choose some subjects and instruct them to exercise
- the other subjects are instructed *not* to exercise

- assess insomnia for all subjects.

Why is this better? How does it level out effects of other variables?

- *choosing two groups at random means that the groups should start out relatively equal in terms of anything that might matter*

- *if the groups end up unequal in terms of insomnia, then evidence that exercise made a difference.*

Terminology

- People/animals/whatever participating in experiment called *experimental units / subjects.*

- Experimenter has at least one explanatory variable, a ***factor***, to manipulate.

- At least one *response variable* measured.
- Specific values chosen for factor called ***levels****.*

- Combination of manipulated levels of factors called ***treatment***.

Example:

Variation of exercise/insomnia experiment: add diet

- three kinds of exercise: none, moderate, strenuous
- two different diets: fruit/veg, “normal”
- factors are:
  - exercise, with 3 levels
  - diet, with 2 levels
- 3 x 2 = 6 treatments (6 combinations of 2 factors)
- divide subjects into 6 groups at random.

Principles of experimental design p321

1. Control
- We control sources of variation other than the factors we are testing by making conditions as similar as possible for all treatment groups.

2. Randomize
- allows us to equalize the effects of unknown or uncontrollable sources of variation.
- It does not eliminate the effects of these sources, but it spreads them out across the treatment levels so that we can see past them.

3. Replicate

- get many measurements of response for each treatment.

4. Blocking

- divide experimental units into groups of similar ones and sample appropriately

If we group similar individuals together and then randomize within each of these blocks, we can remove much of the variability due to the difference among the blocks.

- **Statistical Significance p325**
  An observed effect so large that it would rarely occur by chance is called **statistically significant**.

Placebos p328

- A placebo is a "fake" treatment designed to look like a real one.

- Why is that important?
  - Known that receiving *any* treatment will cause a subject to improve.

  - Want to show that the "real" treatment is not just effective, but better than a placebo. Then have evidence that the treatment is worth knowing about.

- Can also use current standard treatment to compare with.
- Subjects getting placebo/standard treatment called *control group.*

Blinding p327

- When we know what treatment was assigned, it’s difficult not to let that knowledge influence our assessment of the response, even when we try to be careful.

- In order to avoid the bias that might result from knowing what treatment was assigned, we use **blinding**.

- There are two main classes of individuals who can affect the outcome of the experiment:

  - those who could influence the results (subjects, treatment administrators, technicians)

  - those who evaluate the results (judges, treating physicians, etc.)

- When every individual in *either one* of these classes is blinded, an experiment is said to be **single-blind**.

When everyone in *both* classes is blinded, the experiment is called **double-blind**

The best experiments

are:

- randomized
- comparative
- double-blind
- placebo-controlled.

- **Completely randomized design** (CRD), p330
-

When all experimental units are allocated at random among all treatments, the experimental design is **completely randomized**

Example Want to compare three teaching methods A , B and C.

- Select a group of students at random e.g. 30 students.
- Divide them into here groups at random. One group studying from method A, one from method B and the other from method C

- Assess performance of all students and compare the results.

- **Block design** p330

A **block** is a group of experimental units or subjects that are known before the experiment to be similar in some way that is expected to affect the response to the treatments. In a **randomized block design (RBD)**, the random assignment of units to treatments is carried out separately within each block.

Example Want to compare three teaching methods A , B and C. The strength of the students can affect results (i.e. strength is a blocking variable).

- Select a group of students at random e.g. 30 students.
- Divide the students into blocks, say 10 blocks: the best three students, the next best three and so on.
- Form each block pick one of them at random to study from method A, one from method B and the other from method C. i.e. the subjects in each block are randomized into the three teaching methods.

- Assess performance of all students and compare the results.

Confounding p332 (We discussed this in correlation)

Ethical experiments

Idea of *imposing* treatments on subjects might be questionable:

- what if study effects of smoking on lung disease?
- would have to prevent some subjects from smoking, and *make some subjects smoke* for duration of study (!!!)

There are some known unhealthy/dangerous things you cannot ask subjects to do. Also,

- giving a placebo when a best proven treatment is available *is not ethical.*
- subjects who receive placebo *must not be subject to serious harm* by so doing.

See **Declaration of Helsinki**, which governs experiments on human subjects:

www.wma.net/en/30publications/10policies/b3/index.html