Review - Week 5

Read: Chapters 13-15

Review: Experiments

Observational studies – investigations in which one simply observes the state of some population without trying to influence the responses.

Experimental studies – active imposition of treatments on the subjects of an experiment.

In an experiment, one or more treatments are imposed on the subjects. Each treatment is a combination of levels of the explanatory variable, which we call factors.

The design of an experiment is a specification of the treatments to be used and the manner in which subjects are assigned to these treatments. The basic principles of experimental design are control, randomization and replication.

1. Control for the effects of lurking variables.

2. Randomization is the use of chance to assign subjects into groups. It produces groups of subjects that are more likely to be similar in all respects before the treatments are applied, than when using non-random methods.

3. Repeating an experiment on many subjects is called replication. Experiments consisting of more subjects are more likely to detect differences than those with fewer subjects.

Additional control can be achieved by splitting subjects into blocks. A block is a group of subjects that are similar in ways that are expected to affect the response. Randomization is thereafter carried out separately in each block.

In a matched pairs design we choose blocks of two subjects that are as closely matched as possible. Each member of the pair then receives different treatments.

Experiments should compare two or more treatments in order to avoid confounding the effect of a treatment with other influences. Often one of these treatments is a placebo, or fake treatment, and those subjects receiving the placebo are referred to as the control group.

In a double blind experiment neither the subjects nor the people who have contact with them know which treatment a particular subject received.

If the observed differences among the treatment groups are big enough, we attribute the differences to the treatment. The difference is statistically significant if the observed effect is so large that it would rarely occur by chance.
Exercises:

**Exercise 1:** A teacher wants to compare the effectiveness of a new reading curriculum with the standard one. She tests the reading ability of each student in the class, and then randomly divides them into two groups. One group uses the new curriculum, while the other group uses the standard curriculum. After three months she retests the students and compares the increase in reading ability in the two groups.

(a) Is this an observational study or an experiment? Why?
(b) What are the explanatory and response variables?

**Exercise 2:** Eight students have volunteered to be in a sleep deprivation study. They are Bob, Bill, John, Stan, Sue, Mary, Betty and Marge. We want to assign 4 students into each of two groups that will be assigned different treatments.

(a) Use the random number table to assign the 8 students into 2 groups.
(b) Suppose we want to ensure that each of the two groups consist of two women and two men. Randomly assign the students in each block to the two groups.

**Exercise 3:** A researcher wants to determine whether two new pesticides can effectively limit crop damage that is due to a certain species of beetle. She plans on applying the pesticides to a variety of plots of corn, wait a week and thereafter check the number of beetles found on each plot. The researcher wants to know whether or not either of the new pesticides works and whether there is a significant difference in effectiveness between them. Design an appropriate experiment.

Review: Probability

The set, S, of all possible outcomes of a random phenomena is called the sample space. An event is any collection of possible outcomes of random phenomena, that is, any subset of S.

Given any two events A and B, we have the following elementary set operations:

- The union of A and B, written $A \cup B$, is the set of elements that belong to either A or B or both.
- The intersection of A and B, written $A \cap B$, is the set of elements that belong to both A and B.
- The complement of A, written $A^c$, is the set of all elements that are not in A.

If two events A and B have no outcomes in common, they are said to be disjoint.
A probability model consists of a sample space and an assignment of probabilities $P$, that assign a number $P(A)$ to each event $A$ as its probability.

The probability of any outcome of a random phenomenon is the proportion of times the outcome would occur in a very long series of repetitions. The law of large numbers says that this proportion gets closer and closer to the true relative frequency as the number of trials increases.

**Rules of Probability** - Consider an experiment whose sample space is $S$. For each event $A$ of the sample space we assume that a number $P(A)$ is defined and satisfies the following rules:

1. $0 \leq P(A) \leq 1$
2. $P(S) = 1$
3. $P(A^c) = 1 - P(A)$
4. If $A$ and $B$ are disjoint, then $P(A \cup B) = P(A) + P(B)$
5. If $A$ and $B$ are independent then $P(A \cap B) = P(A)P(B)$.

**Exercises:**

**Exercise 1:** Roll a six-sided die. Let $A$ be the event that you roll an even number, $B$ be the event that you roll an odd number and $C$ be the event that you roll a three or less. What outcomes are contained in the event:

(a) $A \cup B$  (b) $A \cap B$  (c) $A^c \cup B$  (d) $A \cap C$  (e) $A \cap C^c$

**Exercise 2:** Suppose that you roll two six-sided dice.

(a) What outcomes are contained in the sample space?  
(b) Which outcomes are contained in the event that you roll a seven?  
(c) What is the probability that you roll a seven?  
(d) What is the probability that you roll something other than a seven?  
(e) Which outcomes are contained in the event that you roll a five?  
(f) What is the probability that you roll a five?  
(g) What is the probability that you roll a seven or a five?

**Exercise 3:** A basketball player makes 70% of his free throws. Suppose he shoots 10 in a row. What is the probability that he makes all of them? What is the probability that he misses at least one? You may assume that each shot is independent of the others.