Producing Data

There are basically two ways to gather data (which we will then run through our battery of statistical analyses)—we can gather data without any attempt to influence outcomes, or we can gather data while trying to influence some (or all) of the outcomes.

When there is no deliberate attempt to influence any observations, we are conducting an **observational study** (we're just observing). In this case, a primary concern is obtaining a *sample* from the population.

When we attempt to influence some (or all) observations, then we are conducting an **experiment**. In this case, we often do not have a sample—the primary concern will be the *design* of the experiment.

Designing Samples

Sampling Techniques

Voluntary Response

When the subjects select themselves for the sample, then you have a **Voluntary Response Sample**. For example, call-in polls use voluntary response (since the caller decides to call in, and become part of the sample).

Voluntary Response is bad—there is often some reason why the subjects join the sample, and this reason may have some effect on what you are trying to measure.

Convenience

When the subjects are chosen based on ease of access, then you have a **Convenience Sample**. For example, if I want to know about the opinions of High School students, and I decide to sample from IHS students only, then I've chosen my sample from those that are easiest to reach—that's a convenience sample.

This is bad.

Simple Random

A **Simple Random Sample** (SRS; of size n) is chosen in such a way that every possible group (of size n) from the population has an equal chance of being selected. The best example to get the idea of an SRS is drawing names out of a hat.

Since every group has an equal chance of being selected, every individual also has an equal chance. *This is not reversible*! If every *individual* has an equal chance, there is no guarantee that every *group* has an equal chance.

Probability Sample

A **Probability Sample** assigns each individual some chance of being selected. An SRS is a Probability Sample where each individual gets the same chance. An example where everyone does not get an equal chance would be a lottery or raffle—your chance of being selected (winning) depends on how many tickets you have. Every person with at least one ticket has some chance, but not every individual has the same chance.

Stratified

To take a **Stratified Sample**, first divide the population into homogenous groups (strata). Now take an SRS from each group. Combine these individual SRS's into a single sample from the population. For example, I could divide the student body into Freshmen, Sophomores, Juniors and Seniors; then take an SRS from each group; and finally combine these to make a single sample.

Multistage

"Layers of randomization." Or, "strata within strata." Rather than try to explain it here, see the fine example on page 279 of the textbook.

Cluster

A technique where all individuals within a certain physical area are selected—for example, selecting all houses on a certain (randomly selected) street. This is often used as the last step in a multistage design.

Using the Table of Random Digits

You may be asked to use a table of random digits to select samples. Here are the steps:

[1] Number the individuals. Every individual needs a number with the same number of digits as every other individual—this is often handled by "padding" a number with leading zeros. Note that it is OK (in some designs) for individuals to have more than one number (especially the probability sample). In some cases, this may already have been done for you!

[2] Read a group of digits from the table. Every individual has a number with the same number of digits, so read that many digits from the table! There is always some question about where to start—for the AP, we'll always start at the upper left corner, and read from left to right. When you reach the end of a line, continue onto the next line without any pause or break. If you reach the end of the table, continue again at the beginning.

[3] Decide if the selected number gives a new member to the sample. It is possible that the selected number is not assigned to an individual. It's also possible that the number is for an individual that has already been selected for the sample—you can't select an individual twice for the same sample!

[4] Repeat [2] and [3] until the sample is complete. When you repeat step [2], pick up in the table of random digits right where you left off.

Bias

Bias is something that creates results that are different from what they *ought* to be—in other words, something that systematically favors certain outcomes/measurements over others.

There are three basic types of bias: selection bias, measurement/response bias, nonresponse bias.

Selection Bias

Selection bias is introduced during the selection process—certain individuals are given greater (than intended) probabilities of being selected, or are excluded from the selection process. Failing to include all individuals in the selection process is often called *undercoverage*.

Notice that a voluntary response sample (or even a convenience sample) automatically suffers from selection bias!

Measurement/Response Bias

Measurement bias is introduced when the measurement process tends to give results that differ (systematically) from the population. For example, if a light meter is not properly calibrated, then the measurements it gives will not be correct! A common source of measurement bias is wording bias—the way in which a question is worded can often have an effect on the responses. See the example on page 282 in the text.

Some people use the terms measurement bias and response bias synonymously, but they don't quite mean the same thing. Measurement bias refers exclusively to problems with the measurement device, where response bias refers to anything that might affect the measured results—for example, you are likely to get very different answers to a survey if the person conducting the survey is wearing a Darth Vader costume!

Nonrepsonse Bias

Nonresponse bias is introduced when individuals (people, mostly!) refuse to be measured/refuse to answer questions. Telephone surveys suffer from this! This type of bias is almost unavoidable, so minimizing its effects is important. There are a host of methods (beyond the scope of this course) to address this.

Designing Experiments

Vocabulary

Variables

What are you measuring? Weight? Time required to fall asleep? Dosage (mg) of a certain drug? These are your **variables**. You could have any number of variables, but they fall into three types: explanatory, response, and extraneous. The explanatory variable (or *factor*) is the one that we will change, in order to observe the resulting effect on the response variable. Extraneous variables are just that: extraneous. We don't want/care about them, and must attempt to filter out their effects.

Levels

The explanatory variable/s is/are set by the experimenter: how much of this drug will be used? How much weight will be placed on the girder? How long will we boil the solution? When we decide this, we are setting the **levels** of the explanatory variable/s. For example, if we are studying the effects of a certain drug, then we might try several dosages—say, 50mg, 100mg and 150mg. This is one variable (Dosage) with three levels (50, 100, 150).

Note that we will NOT decide on levels! That is far beyond our ability. The levels should either be given, or obvious—the most obvious levels being presence/absence…

Treatments

A **treatment** consists of one level from each explanatory variable. For example, let's say that you are conducting an experiment on the effects of stirring rate (rpm) and duration (seconds) on the incorporation of an ingredient in a solution/mixture. Furthermore, let's say that you've decided to test three stirring rates (60, 120 and 240rpm) and two durations (1 and 2 minutes). There would be six treatments—60rpm and 1 minute; 60rpm and 2 minutes; 120rpm and 1 minute; 120rpm and 2 minutes; 240rpm and 1 minute; 240rpm and 2 minutes. See the example on page 290 of the text.

Experimental Units / Subjects

The things on which we experiment are called the **experimental units**—perhaps they are batches of paint, or lots of bearings. When the experimental units are people, we call them *subjects*.

Comparative Experiments

It is typically the case that an experiment is designed to determine if changes in A cause (yes, cause) changes in B. The only way that we can be sure that this is the case is to control the effects of any extraneous variables. Sometimes this is possible—often, it is not. Thus, we need to conduct **comparative experiments**, where we compare results of several treatments that are given under similar conditions.

The Placebo Effect

Many experiments compare a treatment against a placebo. A **placebo** is anything which has a known (or no) effect on the response variable. The word *placebo* refers to a sugar pill, but a placebo does not need to be a pill! For example—does advertising with subliminal messages cause people to have a higher opinion of certain products? When conducting this experiment, we might show people advertising with subliminal messages, and then show some other people advertising without the subliminal messages. The ads with subliminals are an actual treatment; the ads without them make a placebo.

Historical note: it was once common to include placebos in a doctor's bag. When patients came and described symptoms which didn't point to an obvious cause, a doctor would give the patient a placebo. If the patient did not return (if the condition improved), then all was well—either the condition improved on its own, or it never really existed (psychosomatic). If the patient did not get better, *then* the case warranted further attention…

The Control Group

Those individuals that actually receive a treatment are in the **experimental group**. Those that receive a placebo are in the **control group**. It is important that similar individuals are in each group. The best way to do that is…

Randomization

…to randomly assign individuals to these groups! This will "level out" the effects of any unknown, extraneous variables in the experiment.

The Three Principles

The Three Principles of Experimental Design are:

[1] Randomization—individuals must be assigned to treatment groups at random!

[2] Control—the effects of extraneous variables must be taken into account (i.e., controlled). Unknown variables are controlled through randomization. Known variables are handled through direct manipulation (e.g., I can directly control the temperature in the room), or blocking (see below).

[3] Replication—larger samples give better results; more individuals in the experiment will provide more reliable results!

Really, these can be reduced to a single phrase: **reduce variation!**

Additional Precautions

Blinding

Placebos are used as the basis of comparison (and one form of control). Unfortunately, knowing that you might have been given a placebo could introduce bias into the experiment. Thus, it is important that the subjects do not know what kind of treatment has been received. When this is the case, the experiment is **single blind**.

It is also important for the person taking the measurements not to know what kind of treatment the subjects received—otherwise, there may be some hidden bias. When neither the subject, nor those who take the measurements know which treatment was applied, then the experiment is **double-blind**.

Realism

In an experiment, we change something. Unfortunately, we can't always do that in an environment similar to that in which the change might naturally occur…for example: does noise level affect performance on a test? Noise level is explanatory; that's what we'll change. It is not easy to make noise, though, that is similar to that which would normally be experienced during a test. (Personally, I participated in exactly this type of experiment in college. The noise was simply a TV broadcast, with the volume turned up very loud.) In this case, there is a lack of realism that will have some effect on the results.

Matched Pairs

One way to deal with unknown extraneous variables is randomization, but that isn't always the best answer. Often, a better choice is to use a **matched pairs design**. This involves matching one experimental unit in the experimental group with one experimental unit in the control group—these two units are either selected because they are very similar, or they are made to be very similar. We've already looked at this in a previous chapter…

Block Designs

Extraneous variables that are known need to be controlled. Sometimes, this is easy—I can adjust the temperature in the room, or I can cover up the label on the can. There are other extraneous variables that I can't easily control or affect—I can't change your gender, nor can I change whether or not you've ever taken the SAT. When we encounter extraneous variables that can't be changed, then we must block. This involves splitting the experimental units into homogenous groups (blocks), and then performing the experiment within each block.

Simulating Experiments

There is a nice summary of this on page 310 of the text. Note how they mark below the random digits—this is precisely what you should do on the AP Exam!