MODULE 3

DATA PRODUCTION

Objectives:

1. Identify major differences between data produced from experiments and observational studies.
2. Understand basic ideas of simple random experiments with one and two factors.
3. Describe the principles of experimental design.
4. Describe the principles of observational studies.
5. Understand basic ideas of designing simple observational studies, and
6. Explain the importance of randomization in both experiments and observational studies.

Contents

Statistical inference is the process of making conclusions about an entire population based on the results from the individuals in a single sample. To make conclusions about the larger population from a sample requires a sample that fairly represents the larger population. In this module, two ways of producing data – (1) Experiments and (2) Observational Studies – are described. The proper collection (or production) of data is critical to statistics (and science in general) so that proper inferences and conclusions can be made.

△ **Inference**: The process of forming conclusions about the unknown parameters of a population by computing statistics from the individuals in a sample.

◇ **If data are not properly collected, then inferences cannot be made.**

3.1 Experiments

An experiment deliberately imposes a condition, or treatment, on individuals to observe their response. In a properly designed experiment all variables that are not of interest are held constant while the variable(s) that are of interest are changed among treatments. As long as the experiment is designed properly (see below), tests for differences in the response variable among treatments can be made. If differences among treatments occur, then those differences are due either to the variable(s) that were deliberately changed or randomness (chance). Thus, strong cause-and-effect statements can be made from data collected with a carefully designed experiment.

◇ **An experiment deliberately imposes a condition, or treatment, on individuals in order to observe their response.**

◇ **Strong cause-and-effect statements can be made from data collected with a carefully designed experiment.**

3.1.1 Single-factor Experiments

A factor is the variable that is deliberately manipulated to determine its effect on the response variable. Sometimes the factor is called an explanatory variable because we are attempting to determine how it affects (or "explains") the response variable. The simplest experiment is a single-factor experiment where the individuals are split into groups defined by the categories of a single factor variable.

For example, suppose that a research group wants to examine the effect of temperature on the total number of bacterial cells after two weeks. They have inoculated 120 agars (petri dishes with a growth medium for the bacteria) with the bacteria and placed them in a chamber where all environmental conditions (e.g., temperature, humidity, light) can be controlled exactly. The researchers will use only two temperatures in this simple experiment – 10^{o}C and 15^{o}C. Thus, temperature is the only factor in this simple experiement because it is the only variable manipulated to different values to determine its impact on the number of bacterial cells.

△ **Factor(s)**: The variable(s) that is (are) deliberately manipulated in the experiment to determine its effect on the response variable. Sometimes called the explanatory variable.

△ **Response**: The variable observed in an experiment to identify the effect of the factors on it.

◇ **In a single-factor experiment only one explanatory variable (i.e., factor) is allowed to vary; all other explanatory variables are held constant.**

Levels are the number of categories of the factor variable. In this example, there are two levels – 10^oC and 15^oC. Treatments are the number of unique conditions that individuals in the experiment are exposed to. In a single-factor experiment, the number of treatments is the same as the number of levels of the factor. Thus, in this simple experiment, there are two treatments – 10^oC and 15^oC. Treatments are discussed more thoroughly in the next section.

The number of replicates in an experiment is the number of individuals that will receive each treatment. In this example, the replicates are the number of inoculated agars that will receive each of the two temperature treatments. The number of replicates is determined by dividing the total number of available individuals (120) by the number of treatments (2). In this case, the number of replicates is 60 inoculated agars.

△ **Levels**: The number of categories or groupings of the factor.

◇ **In single-factor experiments, the number of treatments in the experiment equals the number of levels of the single factor.**

△ **Replicates**: The number of individuals in each treatment group.

◇ **The number of replicates is determined by dividing the total number of available individuals by the number of treatments.**

The agars used in this experiment will be randomly allocated to the two temperature treatments. All other variables – humidity, light, etc. – are kept the same for each treatment. At the end of two weeks, the total number of bacterial cells on each agar (i.e., the response variable) will be recorded and compared between the agars kept at both temperatures.[1] Any difference in mean number of bacterial cells will be due to either different temperature treatments or randomness, because all other variables were the same between the two treatments.

◇ **Differences among treatments are either caused by randomness (chance) or the factor.**

The single factor is not restricted to just two levels. For example, more than two temperatures, say 10^oC, 12.5^oC, 15^oC, and 17.5^oC, could have been tested. With this modification, there is still only one factor – temperature – but there are now four levels (and only four treatments).

[1] Methods for making this comparison are in Module 17.

3.1.2 Multi-factor Experiments

More than one factor can be tested in one experiment. In fact, it is more efficient to have a properly designed experiment where more than one factor is varied at a time than it is to use separate experiments in which only one factor is varied in each. However, before showing this benefit, let's examine the definitions from the previous section in a multi-factor experiment.

Suppose that the previous experiment was modified to also examine the effect of relative humidity on the number of bacteria cells. This modified experiment has two factors – temperature (with two levels of 10^oC or 15^oC) and relative humidity (with four levels of 20%, 40%, 60%, and 80%). The number of treatments, or combinations of all factors, in this experiment is found by multiplying the levels of all factors (i.e., $2\times4=8$ in this case). The number of replicates in this experiment is now 15 (i.e., total number of available agars divided by eight; 120/8).

A quick drawing of the experimental design can be instructive (below). The drawing is a grid where the levels of one factor are the rows and the levels of the other factor are the columns. The number of rows and columns correspond to the levels of the two factors, respectively, whereas the number of cells in the grid is the number of treatments (numbered in this table to show eight treatments).

	Relative Humidity			
	20%	40%	60%	80%
10^oC	1	2	3	4
15^oC	5	6	7	8

△ **Treatments**: The number of combinations of all factors in the experiment.

⋄ **The number of treatments equals the product of the levels for each factor.**

⋄ **The number of treatments is determined for the overall experiment, whereas the number of levels is determined for each factor.**

The analysis of a multi-factor experimental design is more involved than what will be shown in this course. However, multi-factor experiments have many benefits, which can be illustrated by comparing a multi-factor experiment to separate single-factor experiments. Let's continue with the experiment to identify the effect of both temperature and relative humidity on the number of bacterial cells. However, consider for the moment that (1) separate single-factor experiments will also be conducted to determine the effect of each factor and (2) we cannot use any of the individuals (i.e., agars) in more than one experiment.

To conduct the two separate experiments, randomly split the 120 available agars into two equally-sized groups of 60. The first 60 will be split into two groups of 30 for the first experiment with temperature. The second 60 will be split into four groups of 15 for the second experiment with relative humidity. These separate single-factor experiments are summarized in the following tables (where the numbers in the cells represent the number of replicates in each treatment).

Temperature	
10^oC	15^oC
30	30

Relative Humidity			
20%	40%	60%	80%
15	15	15	15

Now reconsider the design where both factors were varied at once (the table below was modified to include the number of replicates in each treatment).

	Relative Humidity			
	20%	40%	60%	80%
10^oC	15	15	15	15
15^oC	15	15	15	15

The key to examining the benefits of the multi-factor experiment is to determine the number of individuals that give "information" about (i.e., are exposed to) each factor. From the last table it is seen that all 120 individuals are exposed to one of the temperature levels with 60 individuals exposed to each level. In contrast, only 30 individuals were exposed to these levels in the single-factor experiment. In addition, all 120 individuals are exposed to one of the relative humidity levels with 30 individuals exposed to each level. Again, this is in contrast to the single-factor experiment where only 15 individuals were exposed to these levels. Thus, the first advantage of multi-factor experiments is that the available individuals are used more efficiently. In other words, more "information" (i.e., the responses of more individuals) is obtained from a multi-factor experiment than from combinations of single-factor experiments.[2]

> ◇ **Multi-factor experiments use individuals more efficiently; i.e., more "information" about the effect of the factors on the response is gained from the same number of individuals.**

A properly designed multi-factor experiment also allows researchers to determine if multiple factors interact to impact an individual's response. For example, consider the hypothetical results from this experiment in Figure 3.1.[3] The effect of relative humidity is to increase the growth rate for those individuals at 10^oC (black line) but to decrease the growth rate for those individuals at 15^oC (blue line). That is, the effect of relative humidity differs depending on the level of temperature. When the effect of one factor differs depending on the level of the other factor, then the two factors are said to *interact*. Interactions cannot be determined from the two single-factor experiments because the same individuals are not exposed to levels of the two factors at the same time.

Multi-factor experiments are used to detect the presence or absence of interaction, not just the presence of it. The hypothetical results in Figure 3.2 show that the growth rate increases with increasing relative humidity at about the same rate for both temperatures. Thus, because the effect of relative humidity is the same for each temperature (and vice versa), there does not appear to be an interaction between the two factors. Again, this could not be determined from the separate single-factor experiments.

> ◇ **Multi-factor experiments can be used to detect interactions between multiple factors.**

3.1.3 Allocating Individuals

In the previous examples, each individual[4] was allocated to (i.e., placed into) treatments. Individuals should be randomly allocated to treatments. Randomization will tend to even out differences among groups for variables not considered in the experiment. In other words, randomization should help assure that all groups

[2]The real importance of this advantage will become apparent when statistical power is introduced in Module 13.
[3]The means of each treatment are plotted and connected with lines in this plot.
[4]When discussing experiments, an "individual" is often referred to as a "replicate" or an "experimental unit."

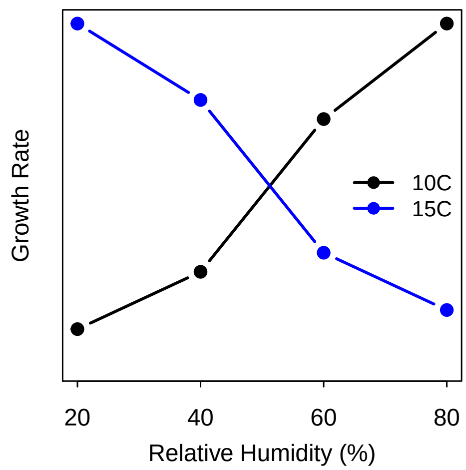

Figure 3.1. Mean growth rates in a two-factor experiment that depict an interaction effect.

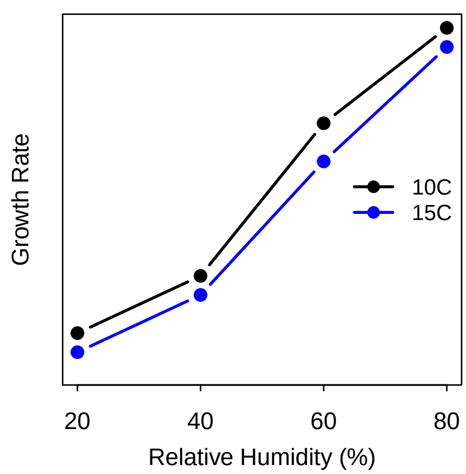

Figure 3.2. Mean growth rates in a two-factor experiment that depict no interaction effect.

are similar before the treatments are imposed. Thus, randomly allocating individuals to treatments removes any bias (foreseen or unforeseen) from entering the experiment.

In the single-factor experiment above – two treatments of temperature – there were 120 agars. To randomly allocate these individuals to the treatments, 60 pieces of paper marked with "10" and 60 marked with "15" could be placed into a hat. One piece of paper would be drawn for each agar and the agar would receive the temperature found on the piece of paper. Alternatively, each agar could be assigned a unique number between 1 and 120 and pieces of paper with these numbers could be placed into the hat. The agars corresponding to the first 60 numbers drawn from the hat could then be placed into the first treatment with the agars for the next (or remaining) 60 numbers in the second treatment. This process is essentially the same as randomly ordering the 120 numbers. A random order of numbers is obtained with R by including the count of numbers as the only argument to `sample()`. For example, randomly ordering the numbers 1 through 120 is accomplished with

```
> sample(120)
```

```
  [1]  80  30 100  90  21  68 104  79  64 106  98  16  73  91 107   1  60  54  26  99
 [21] 108 111  31  47  57  92   5  58  37  50  34  88  41  66  65  29 110 113   4  75
 [41]  93  23  49  97  35  84  74   7  15  39  70  94 114  14  71  20  33  67  86   8
 [61]   6  28  52  48  13  18  63  72  69 120  55  83  42   3  77  82  38  22  96  43
 [81]  56  89  78  17 112  44 103  46  59  85 109 115 118  87  32  62  51  95  24  40
[101] 119 102  19  27 116  36   2  12  45  53  11  76 117  61 105   9 101  25  81  10
```

Thus, the first five agars in the 10^{o}C treatment are 80, 30, 100, 90, and 21. The first five agars in the 15^{o}C treatment are 6, 28, 52, 48, and 13.

Now consider the modified experiment with two factors – temperature and relative humidity – with eight treatments containing 15 agars each. Here, it is more efficient to save the random numbers into an object and then select the numbers in the first 15 positions, then the second 15 positions, etc.

```
> ragars2 <- sample(120)
> ragars2[1:15]     # "grab" the first 15 numbers
 [1]  61  82 103  31  66  81 105  40 104 106   5   9  71  36   8
> ragars2[16:30]    # "grab" the second 15 numbers, and so on
 [1] 120   6  26  41  62 111  83  20  57   1  63  86  70  85  73
```

This design might be shown with the following table, where the numbers in each cell represent the first two agars selected to receive that treatment.[5]

	Relative Humidity			
	20%	40%	60%	80%
10^{o}C	61,82,$\cdots$	120,6,$\cdots$	60,72,$\cdots$	89,49,$\cdots$
15^{o}C	78,10,$\cdots$	109,101,$\cdots$	22,2,$\cdots$	114,77,$\cdots$

◇ **Individuals should be randomly allocated to treatments to remove bias.**

3.1.4 Design Principles

There are many other methods of designing experiments and allocating individuals, including blocked designs, nested designs, etc., that are beyond the scope of this book. However, all experimental designs contain these three basic principles.

1. **Control** the effect of variables on the response variable by deliberately manipulating factors to certain levels and maintaining constancy among other variables.
2. **Randomize** the allocation of individuals to treatments to eliminate bias.
3. **Replicate individuals** (use many individuals) in the experiment to reduce chance variation in the results.

[5] Only the first two numbers are shown because of space constraints.

Proper control in an experiment allows for strong cause-and-effect statements; i.e., to state that an observed difference in the response variable was due to the levels of the factor or chance variation rather than some other variable (foreseen or unforeseen). Randomly allocating individuals to treatments removes any bias that may be included in the experiment. For example, if we do not randomly allocate the agars to the treatments, then it is possible that a set of all "poor" agars may end up in one treatment. In this case, any observed differences in the response may not be due to the levels of the factor but to the prior quality of the agars. Replication means that there should be more than one or a few individuals in each treatment. This reduces the effect of each individual on the overall results. For example, if there was one agar in each treatment, then, even with random allocation, the effect of that treatment may be due to some inherent properties of that agar rather than the levels of the factors. Replication, along with randomization, helps assure that the groups of individuals in each treatment are as alike as possible at the start of the experiment.

◇ **Control, Randomization, and Replication are the three major principles of experimental design.**

Review Exercises

3.1 While studying the foraging ecology of northern elephant seals, marine biologists from California observed the health of wild seals in fenced enclosures of two different water temperatures ($< 47^oF$ and $> 47^oF$) and compared these results to the health of domestic seals in two pools, with water temperatures analogous to the wild seals. The wild seals were allowed to eat what they wanted, but the domestic seals were fed a known diet. There were 20 wild seals and 20 domestic seals, each of which was randomly allocated to the two water temperatures (enclosures for the wild seals). Use this information to answer the questions below. *Answer*

(a) Construct a simple diagram to represent this experiment.
(b) What is the response variable?
(c) What are the factors (list all of them)?
(d) How many levels are there (list in same order as factors in answer c)?
(e) How many treatments are there?
(f) How many replicates are there?

3.2 An agronomist is interested in the effect of plowing depth (10 cm, 17 cm, and 25 cm) and amount of applied fertilizer (none or 3 kg per acre) on the harvest of sugar beets. There are 36 nearly identical plots (fields) available for research. The agronomist has asked you to help design an experiment. Specifically, you are asked the questions below. *Answer*

(a) What are the factors?
(b) List the levels for each factor.
(c) How many treatments?
(d) How many replicates for each treatment?
(e) Physically, what is a replicate in this case?
(f) Describe how you would allocate individuals to treatments. Show your R work.

3.3 Translocation is an important tool in modern wildlife management. Current techniques, however, result in the death of many translocated individuals shortly after release in their new homes. Researchers in

France (Letty *et al.* 2000) simultaneously examined the use of tranquilization (tranquilized or not) and acclimatization pens (pens where an individual can "get used to" the new environment; used acclimatization pen or not) on the survival rate (survived or not) of translocated rabbits. Their experiment used a total of 64 European wild rabbits. Use this information to answer the questions below. Answer

(a) Construct a diagram to represent this experiment.
(b) What is the response variable?
(c) What are the factors (list all of them)?
(d) How many levels are there (list in same order as factors in answer c)?
(e) How many treatments are there?
(f) How many replicates are there?
(g) What is an individual in this experiment?

3.4 In 1994, biologists studied the health of whitetail deer as it relates to eating habits. Sixty-four deer were randomly allocated into four groups. One group was to be kept on a deer farm and fed a strict diet. The other two groups would be sent to Channel Island off the coast of Alaska. One of the Channel Island groups would be restricted to browsing in prairies to simulate farm fields. The second was to be restricted to browsing in hardwood forests. The third Channel Island group would be fed a strict diet on the island. The researchers literally followed these deer around for 9 months, recording what the deer ate as they moved. Urine was also collected to assess the health of the deer. Use this information to answer the questions below. Answer

(a) What is the response variable?
(b) What are the factors (list all of them)?
(c) How many levels are there (list in same order as factors in answer b)?
(d) How many treatments are there?
(e) How many replicates are there?
(f) What is an individual in this experiment?

3.5 A chemical engineer is designing the production process for a new product. The chemical reaction that produces the product may have higher or lower yield, depending on the temperature and stirring rate in the vessel in which the reaction takes place. The engineer decides to investigate the effect on yield of two temperatures (50C and 60C) and three stirring rates (60, 90, and 120 rpm). A new vessel should be used for each production and only 30 vessels exist. Help the engineer set up this experiment by answering the questions below. Answer

(a) What are the factors (list all of them)?
(b) How many levels are there (list in same order as factors in answer a)?
(c) How many treatments are there?
(d) What is the response variable?
(e) How many replicates are there?
(f) Physically, what is a replicate (i.e., not a number)?
(g) Identify the individuals for each treatment. Show your R work.
(h) Use a simple table to diagram the experimental setup.

3.6 A student is designing an experiment to determine the simultaneous effects of calcium in the diet and regular exercise on blood pressure. In this experiment, some subjects will be given a calcium supplement pill and some will be given a placebo sugar pill. In addition, some subjects will be required to perform aerobic exercises once a day, whereas others will not. The researcher has 32 male subjects available that are as similar as possible (similar ages, weights, initial blood pressures, etc.). Help the student design this experiment by answering the questions below. Answer

(a) What are the factors (list all of them)?
(b) How many levels are there (list in same order as factors in answer a)?
(c) How many treatments are there?
(d) What is the response variable?
(e) How many replicates are there?
(f) Physically, what is a replicate (i.e., not a number)?
(g) Identify the individuals for each treatment. Show your R work.
(h) Use a simple table to diagram the experimental setup.

3.2 Observational Studies – Sampling

In observational studies the researcher has no control over any of the variables observed for an individual. The researcher simply observes individuals, disturbing them as little as possible, trying to get a "picture" of the population. Observational studies cannot be used to make cause-and-effect statements because all variables that may affect the outcome may not have been measured or specifically controlled. Thus, any observed difference among groups may be caused by the variables measured, some other unmeasured variables, or chance (randomness).

Consider the following as an example of the problems that can occur when all variables are not measured. For many years scientists thought that the brains of females weighed less than the brains of males. They used this finding to support all kinds of ideas about sex-based differences in learning ability. However, these earlier researchers failed to measure body weight, which has since been found to be strongly related to brain weight in both males and females. After controlling for the effect of differences in body weights, there was no difference in brain weights between the sexes. Thus, many sexist ideas persisted for years because cause-and-effect statements were inferred from data where all variables were not recorded.

◇ Strong cause-and-effect statements CANNOT be made from observational studies.

In observational studies, it is important to understand to what population inferences will be made.[6] To make useful inferences from a sample, the sample must be an unbiased representation of the population. In other words, it must not systematically favor certain individuals or outcomes.

For example, consider that you want to determine the mean length of all fish in a particular lake (e.g., Square Lake from Section 2.1). Using a net with large mesh, such that only large fish are caught, would produce a biased sample because interest is in all fish not just the large fish in Square Lake. Setting the nets near spawning beds (i.e., only adult fish) would also produce a biased sample. In both instances, a sample would be collected from a population other than the population of interest. Thus it is important to select a sample from the specified population.

◇ It is important to understand what the population is before considering how to take a sample.

[6]Thus, it is very important to first perform an IVPPS as discussed in Section 2.1.

3.2.1 Types of Sampling Designs

Three common types of sampling designs – voluntary response, convenience, and probability-based samples – are considered in this section. Voluntary response and convenience samples tend to produce biased samples, whereas proper probability-based samples will produce an unbiased sample.

A voluntary response sample consists of individuals that have chosen themselves for the sample by responding to a general appeal. An example of a voluntary response sample would be the group of people that respond to a general appeal placed in the school newspaper. If the population of interest in this sample was all students at the school, then this type of general appeal would likely produce a biased sample of students that (i) read the school newspaper, (ii) feel strongly about the topic, or (iii) both.

A convenience sample consists of individuals who are easiest to reach for the researcher. An example of a convenience sample is when a researcher queries only those students in a particular class. This sample is “convenient” because the individuals are easy to gather. However, if the population of interest was all students at the school, then this type of sample would likely produce a biased sample of students that is likely (i) of one major or another, (ii) in one or two-years (e.g., Freshman or Sophomores), or (iii) both.

△ **Voluntary Response Sample**: A sample of individuals that choose themselves for the sample by responding to a general appeal.

△ **Convenience Sample**: A sample of individuals who are easiest to reach for the researcher.

◇ **Voluntary response and convenience samples often produces a biased sample.**

In probability-based sampling, each individual of the population has a known chance of being selected for the sample. The simplest probability-based sample is the **Simple Random Sample** (SRS) where each individual has the same chance of being selected. Proper selection of an SRS requires each individual to be assigned a unique number. The SRS is then formed by choosing random numbers and collecting the individuals that correspond to those numbers.

For example, an auditor may need to select a sample of 30 financial transactions from all transactions of a particular bank during the previous month. Because each transaction is numbered, the auditor may know that there were 1112 transactions during the previous month (i.e., the population). The auditor would then number each transaction from 1 to 1112 (likely already done in this case), randomly select 30 numbers (with no repeats) from between 1 and 1112, and then physically locate the 30 transactions that correspond to the 30 selected numbers. Those 30 transactions are the SRS.

Random numbers are selected in R by including the population size as the first and sample size as the second argument to `sample()`. For example, 30 numbers from between 1 and 1112 is selected with

```
> sample(1112,30)
 [1]   75  320  874  104  128  870  607 1091 1030 1053 1031  518  433  893  816  903
[17]  342 1016  136  580  670  376  576 1076 1034  365  492  189  409   66
```

Thus, accounts 75, 320, 874, 104, and 128 would be the first five selected.

There are other more complex types of probability-based samples – e.g., stratified samples and nested or multistage samples – that are beyond the scope of this course. However, the goal of these more complex types of samples is generally to impart more control into the sampling design.

△ **Probability-based Sample**: A sample where each individual of the population has a known chance of being selected for the sample.

△ **Simple Random Sample**: A probability-based sample where each individual of the population has the same chance of being selected for the sample. Usually abbreviated as SRS.

◇ **To conduct a proper SRS each individual of the population must be able to be assigned a unique number.**

If the population is such that a numerical label cannot be assigned to each individual, then the researcher must try to use a method of selection for which they feel each individual has an equal chance of being selected. Usually this means randomizing the technique rather than the individuals. In the fish example discussed on the previous page, the researcher may consider choosing random mesh sizes, random locations for placing the net, or random times for placing the net. Thus, in many real-life instances the researcher simply tries to use a method that is likely to produce an SRS or something very close to it.

◇ **If a number cannot be assigned to each individual in the population, then the researcher should randomize the "technique" to assure as close to a random sample as possible.**

Polls, campaign or otherwise, are examples of observational studies that you are probably familiar with. The following are links to sites that discuss various aspects of polling.

- How Polls are Conducted by Frank Newport, Lydia Saad, and David Moore, The Gallup Organization.
- Why Do Campaign Polls Zigzag So Much? by G.S. Wasserman, Purdue U.

3.2.2 Of What Value are Observational Studies?

In this module it became apparent that properly designed experiments can lead to "cause-and-effect" statements, whereas observational studies (even properly designed) are unlikely to lead to such statements. Furthermore, in the last section, it was suggested that it is very difficult to take a proper probability-based sample because it is hard to assign a number to each individual in the population (precisely because entire populations are very difficult to "see"). So, do observational studies have any value? There are at least three reasons why observational studies are useful.

The scientific method begins with a scientist making an observation about a natural phenomenon. Observational studies may serve to provide such an observation. Alternatively, observational studies may be deployed after an observation has been made to see if that observation is "prevalent" and worthy of further investigation. Thus, observational studies may lead directly to hypotheses that form the basis of experiments.

Experiments are often conducted under very confined and controlled conditions so that the effect of one or more factors on the response variable can be identified. However, at the conclusion of an experiment it is often questioned whether a similar response would be observed "in nature" under much less controlled conditions. For example, one might determine that a certain fertilizer increases growth of a certain plant in the greenhouse, with consistent soil characteristics, temperatures, lighting, etc. However, it is a much different, and, perhaps, more interesting, question to determine if that fertilizer elicits the same response when applied to an actual field.

Finally, there are situations where conducting an experiment simply cannot be done, either for ethical, financial, size, or other constraints. For example, it is generally accepted that smoking causes cancer in humans even though an experiment where one group of people was forced to smoke while another was not allowed to smoke has not been conducted. Similarly, it is also very difficult to perform valid experiments on "ecosystems." In these situations, an observational study is simply the best study allowable. Cause-and-effect statements are arrived at in these situations because observational studies can be conducted with some, though not absolute, control and control can be imparted mathematically into some analyses.[7] In addition, a "preponderance of evidence" may be arrived at if enough observational studies point to the same conclusion.

Review Exercises

3.7 The National Institutes of Health (NIH) established the Women's Health Initiative (WHI) in 1991 to address the most common causes of death, disability and impaired quality of life in postmenopausal women. The WHI addressed cardiovascular disease, cancer, and osteoporosis. The WHI was a 15 year multi-million dollar endeavor, and one of the largest U.S. prevention studies of its kind. One aspect of the WHI enlisted 93,676 postmenopausal women between the ages of 50-79 from 40 Clinical Centers from throughout the United States (see this map). The women were not asked to take any medication or change their health habits. The health of participants was tracked over an average of eight years by periodically asking the women to complete health forms. What type of study is this? *Answer*

3.8 The U.S. Department of Transportation sponsored a study to determine the transportation patterns and motivations for driving among offenders before, during, and after suspension of their driver's license for an alcohol-related offense (more information here.). The travel patterns for each subject were examined for two four-hour periods during the last month of the suspension period (one observation Monday-Thursday 6 - 10 a.m. and the other observation Friday or Saturday evening 6 - 10 p.m.; actual days were randomly selected). These observation periods were selected to include a time period when the subject would likely be traveling to work and a time period when the subject would likely be traveling for personal, recreational, or social reasons. Similar examinations were conducted at least one month after drivers had had their license reinstated. These post-suspension observations were conducted for each subject at the same times of day and days of the week as the during-suspension observations. What type of study is this? Why? *Answer*

3.9 I have noticed that the needles of white pine trees near major highways are brown. I hypothesized that this may be caused by increased levels of carbon monoxide (CO; ppm) and salt (ppt) near the roads. I am considering two studies to test this hypothesis. First, at two types of sites – near highways and far from highways – I will count the number of trees that are mostly brown and measure levels of CO and salt. Second, I will determine the effect of CO and salt levels by growing 20 randomly-selected nearly-identical seedlings in pots that only differ in the levels of CO and salt – 0 and 5 ppm CO and 0 and 4 ppt salt (NOTE: the 0 levels correspond to normal background levels). *Answer*

(a) Use a diagram to clearly depict the experimental situation described above.
(b) Write the numeric label for each individual in the appropriate place on your diagram.
(c) In the experiment, which treatment is considered a control? Why?
(d) Which study will provide a definitive answer to the stated hypothesis? Explain why!

---

[7] These analyses are beyond the scope of this book, though.