Chapter 1

The Completely Randomized Design

1.1 Study Suggestions

A statistic is a number. Statistics, the discipline, is a collection of principles and methods that help us learn about our world. Statistics is concerned with efficient learning (learning as quickly as possible), and valid learning (learning things that, in fact, are true).

The text will adopt, as much as possible, the following approach. A scientific question will be introduced and during its investigation the subject of Statistics will be shown to be useful. The presentation will not be a whitewash; learning the limitations of Statistics is arguably as important as learning its successful applications.

There are several important ideas in Chapter 1, but my students agree it is the easiest chapter in the text. In Chapter 1, I first want to convince you that comparative studies are *interesting*. To that end, the chapter presents data from 32 projects performed by my students, and several studies performed by professional researchers. In addition, two of the exercises invite you to perform a study of your own.

Chapter 1 introduces a special case of the completely randomized design (CRD). The CRD is a very good way to perform a comparative study. By *special case* I mean that the CRD presented in Chapter 1 is restricted to studies with two treatments and a dichotomous response. In general, a CRD can be used to compare any number of treatments, or it can be used with a more complicated response. Extending the CRD to more than two treatments is a very specialized topic for a first course in statistics, and I discuss this topic only briefly in Chapter 11. Extending the CRD to a more complicated response is very important, and this topic is considered in Chapters 11 and 16.

Chapter 1 introduces you to a large number of studies that can be analyzed in "the same way." I put *the same way* in quotations because there is a common collection of statistical principles and methods that can be applied to each of these studies; but at another level every study is *unique*, and the uniqueness of it is reflected in our interpretation of the answer(s) we get from Statistics.

This seems like a good time to tell you: words matter a great deal in your course! There are two primary ways in which words matter. First, certain words and expressions have a precise technical meaning. When used in your course, a technical word or expression should be given its technical meaning, and not its everyday meaning. For example, **random** has a precise technical meaning that might be different from your everyday use of it.

Second, after performing a statistical analysis of data, or simply completing a homework question, it is important that you be able to explain what you have learned in nontechnical language. But remember, being nontechnical is not an excuse for sloppiness or faulty reasoning.

Chapter 1 introduces the notion that there are two types of subjects—distinct individuals and trials. This dichotomy of subject-types appears throughout the book, and the material will be easier to understand if you take some time to appreciate the distinction now.

Please refer to the *Infidelity study*, introduced on page 11 of the text. Therese's interest motivated her choice of treatments (husband cheating or wife cheating) and response (tell or not tell). Therese chose her subjects—20 friends. After Therese decides to show each friend exactly one treatment, Statistics enters the scene by suggesting that Therese assign subjects to treatments by randomization. At this point, there is no claim that randomization is a good thing. It is simply something that statisticians do. Later in Chapter 1 (see especially Section 1.2.1, *Medical Studies*), two desirable features of randomization are revealed. First, it is fair to the treatments. Second, it is fair to the subjects. (Granted, if the subjects are inanimate, then this second feature is of dubious value.) In Chapter 2, we will learn that randomization also can be used as a basis for **statistical inference** (a technical expression).

It is very important that you learn the mechanics of randomization. Take care always to randomize properly, and do not succumb to the temptation of some 'easier' way to assign subjects to treatments. In particular, after you have studied Chapter 5 you will be able to show that assigning subjects to treatments by tossing a coin will not yield a CRD; thus, *do not* assign subjects to treatments by tossing a coin! Chapter 1 attempts to justify randomization as a reasonable and fair way to assign subjects to treatments. In Chapter 2 you will learn that having a properly randomized study is a critical component of the analysis of a comparative study.

Subtraction. As a child I learned of subtraction as a confrontational topic. I was given the question:

> Bob has nine cookies and his older sister takes away seven cookies. How many cookies are left for Bob?

The answer, of course, is two; nine *take away* (or subtract) seven equals two. In this course, it usually is better to think of subtraction as a way to *compare* two numbers. For example, my dog Casey is ten years old and my son's dog Bailey is two years old; comparing their ages we find that Casey is $10-2=8$ years older than Bailey.

Back to the Infidelity study. The number 0.70 equals the proportion of those subjects who read about a cheating husband who responded that they would tell. Similarly, the number 0.40 equals the proportion of those subjects who read about a cheating wife who responded that they would tell. We will want to compare these numbers. First, qualitatively, we see that the first proportion is larger than the second. Quantitatively, the first proportion is 0.30 larger than the second proportion. (Because, of course, $0.70 - 0.40 = 0.30$.)

It is accurate to say that, "The proportion who would tell on a husband is 0.30 larger than the proportion who would tell on a wife." Alternatively, it is acceptable to say that, "The proportion who would tell on a husband is 30 **percentage points** larger than the proportion who would tell on a wife." But it is misleading to say that, "The proportion who would tell on a husband is 30 percent larger than the proportion who would tell on a wife." (In fact, 0.70 is 75 percent larger than 0.40, because $1.75(0.40) = 0.70$. Especially if you enjoy mathematics you might remember that computing percent increase or decrease is related to comparing two numbers by examining their ratio. A ratio can be a good way to compare two numbers, or it can be grossly misleading—two degrees Fahrenheit is *not* twice as warm as one degree Fahrenheit. Only on rare occasions will we compare two numbers by computing their ratio; differences usually work better in Statistics.)

The remainder of this section consists of a list of errors commonly made by my students. These are errors of enthusiasm; that is, these errors are caused by trying to overuse the CRD. A CRD with two treatments and a dichotomous response is a very powerful tool, but, of course, it is not applicable to, or appropriate for, all scientific problems.

1. A researcher should not change a response that is naturally numerical to a dichotomy just for statistical expediency. For example, suppose a person wants to compare hitting a golf ball with a 3-iron to hitting a golf ball with a 3-wood. A natural response is the distance the ball travels when hit, which is a number, not a dichotomy. I recommend against changing the response to a dichotomy, such as, 'farther than 150 yards' or not.

2. A researcher should take care to distinguish between studies for which randomization is possible and those for which it is not (for example, comparing females and males). Observational studies are very important and interesting, but

they are not CRD's. You will learn about observational comparative studies in Chapters 7, 11, and 16.

3. Sometimes it makes good scientific sense to have each subject consist of several dichotomous trials. In such cases, the total number of successes per subject is a natural response, and it is a number, not a dichotomy. For example, suppose you want to investigate the effect of caffeine versus no caffeine on a person's ability to make a 10 foot putt in golf. It would take a very long time to perform a CRD if each putt was a subject. (Why?) If, however, a subject consists of 25 putts, one could perform a reasonable study in a few days.

4. There is little or no reason to use a CRD to compare responses to two questions, *unless* exposure to both questions would likely influence the subjects' responses.

5. Unfortunately, in the book I have given the impression that a CRD for a sequence of trials must be performed by a single metasubject performing trials two different ways. The CRD also can be used to compare two metasubjects who are performing the trials the same way. For example, a student of mine, Yolanda, used a CRD to perform a study of her ability to shoot free throws with her left (dominant) hand versus her right hand. Alternatively, Yolanda could have used a CRD to compare her ability to shoot free throws with her sister Cary's ability.

1.2 Solutions to Odd-Numbered Exercises

Solutions for Section 1.1

9. The researcher did not assign rats to diets by randomization. The failure to randomize could have introduced any number of possible distortions. For example, perhaps the 10 rats easiest to catch were assigned to the first treatment. Perhaps again, ease of catching is related to a factor (weight?) that influences weight gain in the study. (Other explanations are possible.)

11. If students are able to choose their professor, then the choice made might be related to factors that influence the score on the final exam. For example, perhaps Professor X's students perform worse on the final because he/she has a disproportionate share of students with a weaker math background.

Solutions for Section 1.2

1. Either bar chart reveals the 30-percentage-point difference in the proportions between the two versions of Therese's question.

3. For either version of Becca's question, a majority of subjects responded that they would tell. More said they would tell if the cheater was male, with a difference of 20 percentage points.

5. There is a huge difference, 60 percentage points, between the proportions of successes for the two versions of Julie's question. When pretending to be 65, the subjects showed a great preference for the tranquil end, but when pretending to be 35, the subjects showed a great preference for the aggressive treatment. Of course, one should remember that the number of subjects in this study was very small.

7. When the word *homemade* was included in the description, the proportion of subjects selecting the waffle cone increased by 36 percentage points.

9. A very high proportion of college students would recommend that a friend use the counseling service, but a small proportion would use it themselves. The difference between these proportions is very large—40 percentage points.

11. The proportion of subjects who supported Afrocentric schools was much lower when the exclusion of females was mentioned. The difference is huge—67 percentage points.

13. The inclusion of statistical tables in the history text increased the proportion of subjects who found the author's argument convincing when compared to no tables in the text. The difference is quite large, 28 percentage points.

15. None of the women in the study knew whether she was receiving aspirin or the placebo. The physician(s) who determined the response also did not know which subjects received aspirin and which received the placebo.

    On either treatment, a majority of women did not develop pregnancy-induced hypertension, but the women on aspirin had better results. The difference was 23 percentage points. Alternatively stated, women on the placebo were approximately three times as likely to develop pregnancy-induced hypertension than women on aspirin.

17. Knowledge of one's treatment could conceivably influence a subject's response. In addition, in borderline cases, knowledge of treatment could subconsciously influence a physician's classification of response.

    The proportion of relapses is approximately three times as high in the low-dose group. Alternatively, the proportion of relapses is 25 percentage points higher in the low-dose group. This suggests continuing the standard dose.

Solutions for Section 1.3

5. Pam's son's proportion of successes was higher with his right hand than with his left, but the difference is only 8 percentage points.

7. Loretta's husband shot much better from the left of the basket than he did from in front of the basket. The difference between the proportions is 48 percentage points.

9. With either treatment, Raul's roommate achieved a success on a minority of his trials. He was, however, much more proficient with his eyes open—his proportion was eight times, or 40 percentage points, larger with his eyes open than blindfolded.

11. Kimberly had a low proportion of successes from each treatment, but the proportion was six times, or 20 percentage points, larger from the shorter distance.

13. Franchell shot free throws somewhat better with the larger ball. Franchell's proportion of successes with the larger ball was 16 percentage points higher than the proportion of successes with the smaller ball.

15. Kimberly threw darts much better with her right hand than with her left hand. The difference between the proportion of successes is 48 percentage points.

17. With either treatment, failures by Pearl were rare. She performed slightly better with a left lead, but the difference between the proportion of successes is only 8 percentage points.

19. Comparing the two types of wet vanes, the proportion of failures was seven times—or 24 percentage points—larger with the feathers.

1.3 Exam Questions

1. Numbers are assigned to eight subjects as listed below:

    Al–1; Bev–2; Cy–3; Di–4; Eve–5;
    Fred–6; Gene–7; and Hal–8.

    A box with eight cards numbered one through eight is obtained. Five cards are selected at random from this box yielding the cards numbered 2, 5, 3, 7, and 8. If this drawing is the basis for a randomization, what are the names of the five persons assigned to the first treatment?

2. A CRD yields $\hat{p}_1 = 0.40$ and $\hat{p}_2 = 0.60$. Which of the following statements is true?

    (a) The proportion of successes on the first treatment is twenty percent larger than the proportion of successes on the second treatment.

    (b) The proportion of successes on the first treatment is twenty percentage points larger than the proportion of successes on the second treatment.

    (c) The proportion of failures on the first treatment is twenty percent larger than the proportion of failures on the second treatment.

(d) The proportion of failures on the first treatment is twenty percentage points larger than the proportion of failures on the second treatment.

3. Refer to the alternate randomization scheme presented in Figure 1.1 on page 11 of the text. Subjects numbered 1, 2, ..., 10 select cards in order (that is, subject number 1 selects first, then subject number 2, and so on) from the top of a deck of cards, as illustrated in Figure 1.1. The colors on the cards in the deck are, from top to bottom (B is for black, R is for red),

   B, R, R, R, B, B, B, R, B, B.

   Which four subjects are assigned to version 2?

4. Al performs a CRD with 50 subjects on the first treatment, 25 subjects on the second treatment, and obtains a total of 12 successes—10 on the first treatment, and two on the second treatment.

   Present Al's data in a 2×2 table and compute the table of row proportions.

5. Peg decided to perform a balanced CRD with ten trials. The result of her randomization was that the first five trials were assigned to the first treatment and the last five trials were assigned to the second treatment.

   Al looked at the assignment and said, "That does not look random to me." Comment.

6. A CRD yields $\hat{p}_1 = 0.6$ and $\hat{p}_2 = 0.7$. Construct a bar chart of the proportions of failures on the two treatments.

7. True or false? If the total number of successes equals the total number of failures for a balanced CRD, then $\hat{p}_1 = 0.5$.

8. True or false? If the total number of successes equals the total number of failures for a balanced CRD, then $\hat{p}_1 + \hat{p}_2 = 1$.

9. True or false? The purpose of follow-up in a medical study is to prevent each subject from discovering his or her own treatment.

10. True or false? In a double-blind medical study, each treatment receives the same number of subjects.

11. A basketball player wants to compare her ability to shoot free throws with her left and right hands. She decides to perform a study consisting of five shots with each hand. She begins her study by tossing a coin. If the coin lands heads, then her first shot will be with her right hand and she will alternate hands on every shot until 10 shots are completed. If, however, the coin lands tails, then her first shot will be with her left hand and she will alternate hands on every shot until 10 shots are completed. Is this a CRD? Explain your answer.

12. Statistics is concerned with learning that results from

   (a) Randomized experiments
   (b) Memorization and repetition
   (c) The collection and interpretation of information
   (d) The media

13. A completely randomized design is a study where

   (a) Subjects are assigned to treatments by randomization.
   (b) Subjects give a dichotomous response.
   (c) Treatments are assigned to responses by randomization.
   (d) Treatments yield dichotomous responses.

14. In the Three-Point Basket study Clyde Gaines attempted 100 shots from behind the three-point line, 50 from in front of the basket and 50 from the left corner. Suppose that Clyde's performance gradually got worse because of fatigue. This performance pattern is an example of a

   (a) Sequence of trials
   (b) Time trend
   (c) Poorly designed study
   (d) Placebo effect

15. A researcher performs a balanced controlled comparative study with 10 subjects. The subjects are assigned numbers according to the following scheme: Alice (1), Barbara (2), Carla (3), Diana (4), Elena (5), Felicia (6), Georgia (7), Helena (8), Juanita (9), and Kinsey (10).

    The process of randomization assigns subjects 3, 10, 1, 5, and 2 to the first treatment.

    The responses of Alice, Carla, Elena, and Kinsey are successes, and the the responses of the remaining subjects are failures.

    Present these data in a contingency table.

1.4 Solutions to Exam Questions

1. Bev, Cy, Eve, Gene, and Hal.

2. (d).

3. Subjects numbered 2, 3, 4, and 8 are assigned to version 2.

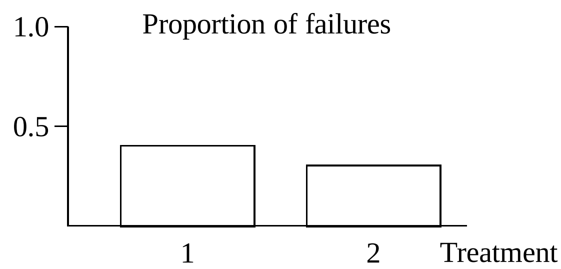
4. Al's tables are below.

Treatment	S	F	Total
1	10	40	50
2	2	23	25
Total	12	63	75

Treatment	S	F	Total
1	0.20	0.80	1.00
2	0.08	0.92	1.00

5. Randomization does not guarantee that the assignment will *look random* to everybody. (In fact, for a small number of subjects it is quite easy to detect a "pattern" in any assignment.)

6. The bar chart is below.



7. False.

8. True.

9. False.

10. False.

11. No. With her method many assignments, including LLLRRRLLRR for example, are impossible. But with a CRD all assignments are possible (additionally, looking ahead to Chapter 2, all are equally likely).

12. (c).

13. (a).

14. (b).

15. The data are presented below.

Treatment	S	F	Total
1	4	1	5
2	0	5	5
Total	4	6	10