

Contributor name: Sakshee Phade

Book Proposed: Biostatistics: A foundation for analysis in the health sciences – Wayne Daniel, Chad Cross

Total Chapters: 14

Total Examples (Chapter-wise): 3, 14, 10, 25, 9, 15, 18, 8, 6, 7, 11, 13, 14, 3

Total Examples: 156

Codable Examples (Chapter-wise): 2, 14, 10, 25, 9, 15, 18, 7, 6, 7, 11, 13, 14, 3

Codable Examples: 154

Chapter 1: Introduction to biostatistics

Example 1.4.1 – Codable

Example 1.4.2 – Codable

Example 1.4.3 – Non-Codable (Reason: Theoretical example which discusses the best way of sampling)

12 CHAPTER 1 INTRODUCTION TO BIOSTATISTICS

Stratified Random Sampling A common situation that may be encountered in a population under study is one in which the sample units occur together in a grouped fashion. On occasion, when the sample units are not inherently grouped, it may be possible and desirable to group them for sampling purposes. In other words, it may be desirable to partition a population of interest into groups, or *strata*, in which the sample units within a particular stratum are more similar to each other than they are to the sample units that compose the other strata. After the population is stratified, it is customary to take a random sample independently from each stratum. This technique is called *stratified random sampling*. The resulting sample is called a *stratified random sample*. Although the benefits of stratified random sampling may not be readily observable, it is most often the case that random samples taken within a stratum will have much less variability than a random sample taken across all strata. This is true because sample units within each stratum tend to have characteristics that are similar.

EXAMPLE 1.4.3

Hospital trauma centers are given ratings depending on their capabilities to treat various traumas. In this system, a level 1 trauma center is the highest level of available trauma care and a level 4 trauma center is the lowest level of available trauma care. Imagine that we are interested in estimating the survival rate of trauma victims treated at hospitals within a large metropolitan area. Suppose that the metropolitan area has a level 1, a level 2, and a level 3 trauma center. We wish to take samples of patients from these trauma centers in such a way that the total sample size is 30.

Solution: We assume that the survival rates of patients may depend quite significantly on the trauma that they experienced and therefore on the level of care that they receive. As a result, a simple random sample of all trauma patients, without regard to the center at which they were treated, may not represent true survival rates, since patients receive different care at the various trauma centers. One way to better estimate the survival rate is to treat each trauma center as a stratum and then randomly select 10 patient files from each of the three centers. This procedure is based on the fact that we suspect that the survival rates within the trauma centers are less variable than the survival rates across trauma centers. Therefore, we believe that the stratified random sample provides a better representation of survival than would a sample taken without regard to differences within strata. ■

It should be noted that two slight modifications of the stratified sampling technique are frequently employed. To illustrate, consider again the trauma center example. In the first place, a systematic sample of patient files could have been selected from each trauma center (stratum). Such a sample is called a *stratified systematic sample*.

The second modification of stratified sampling involves selecting the sample from a given stratum in such a way that the number of sample units selected from that stratum is proportional to the size of the population of that stratum. Suppose, in our trauma center example that the level 1 trauma center treated 100 patients and the level 2 and level 3 trauma centers treated only 10 each. In that case, selecting a random sample of 10 from

Chapter 2: Descriptive statistics

Example 2.2.1 – Codable

Example 2.3.1 – Codable

Example 2.3.2 – Codable

Example 2.4.1 – Codable

Example 2.4.2 – Codable

Example 2.4.3 – Codable

Example 2.4.4 – Codable

Example 2.4.5 – Codable

Example 2.4.6 – Codable

Example 2.5.1 – Codable

Example 2.5.2 – Codable

Example 2.5.3 – Codable

Example 2.5.4 – Codable

Example 2.5.5 – Codable

Chapter 3: Some basic probability concepts

Example 3.4.1 – Codable

Example 3.4.2 – Codable

Example 3.4.3 – Codable

Example 3.4.4 – Codable

Example 3.4.5 – Codable

Example 3.4.6 – Codable

Example 3.4.7 – Codable

Example 3.4.8 – Codable

Example 3.4.9 – Codable

Example 3.5.1 – Codable

Chapter 3: Probability distributions

Example 4.2.1 – Codable

Example 4.2.2 – Codable

Example 4.2.3 – Codable

Example 4.2.4 – Codable

Example 4.2.5 – Codable

Example 4.2.6 – Codable

Example 4.2.7 – Codable

Example 4.2.8 – Codable

Example 4.3.1 – Codable

Example 4.3.2 – Codable

Example 4.3.3 – Codable

Example 4.3.4 – Codable

Example 4.4.1 – Codable

Example 4.4.2 – Codable

Example 4.4.3 – Codable

Example 4.4.4 – Codable

Example 4.4.5 – Codable

Example 4.6.1 – Codable

Example 4.6.2 – Codable

Example 4.6.3 – Codable

Example 4.6.4 – Codable

Example 4.6.5 – Codable

Example 4.7.1 – Codable

Example 4.7.2 – Codable

Example 4.7.3 – Codable

Chapter 5: Some important sampling distributions

Example 5.3.1 – Codable

Example 5.3.2 – Codable

Example 5.3.3 – Codable

Example 5.4.1 – Codable

Example 5.4.2 – Codable

Example 5.5.1 – Codable

Example 5.5.2 – Codable

Example 5.6.1 – Codable

Example 5.6.2 – Codable

Chapter 6: Estimation

Example 6.2.1 – Codable

Example 6.2.2 – Codable

Example 6.2.3 – Codable

Example 6.2.4 – Codable

Example 6.3.1 – Codable

Example 6.4.1 – Codable

Example 6.4.2 – Codable

Example 6.4.3 – Codable

Example 6.4.4 – Codable

Example 6.5.1 – Codable

Example 6.6.1 – Codable

Example 6.7.1 – Codable

Example 6.8.1 – Codable

Example 6.9.1 – Codable

Example 6.10.1 – Codable

Chapter 7: Hypothesis testing

Example 7.2.1 – Codable

Example 7.2.2 – Codable

Example 7.2.3 – Codable

Example 7.2.4 – Codable

Example 7.2.5 – Codable

Example 7.3.1 – Codable

Example 7.3.2 – Codable

Example 7.3.3 – Codable

Example 7.3.4 – Codable

Example 7.4.1 – Codable

Example 7.5.1 – Codable

Example 7.6.1 – Codable

Example 7.7.1 – Codable

Example 7.8.1 – Codable

Example 7.8.2 – Codable

Example 7.9.1 – Codable

Example 7.9.2 – Codable

Example 7.10.1 – Codable

Chapter 8: Analysis of variance

Example 8.1.1 – Non-codable (Reason: Theoretical example explaining the procedure.)

the structure of the models used to represent them will prove to be valuable for understanding some of the most commonly used inferential statistics.

Analysis of Variance This chapter is concerned with *analysis of variance*, which may be defined as a technique whereby the total variation present in a set of data is partitioned into two or more components. Associated with each of these components is a specific source of variation, so that in the analysis it is possible to ascertain the magnitude of the contributions of each of these sources to the total variation.

The development of analysis of variance (ANOVA) is due mainly to the work of R. A. Fisher (1), whose contributions to statistics, spanning the years 1912 to 1962, have had a tremendous influence on modern statistical thought (2,3).

Applications Analysis of variance finds its widest application in the analysis of data derived from experiments. The principles of the design of experiments are well covered in many books, including those by Hinkelmann and Kempthorne (4), Montgomery (5), and Myers and Well (6). We do not study this topic in detail, since to do it justice would require a minimum of an additional chapter. Some of the important concepts in experimental design, however, will become apparent as we discuss analysis of variance.

Analysis of variance is used for two different purposes: (1) to estimate and test hypotheses about population variances, and (2) to estimate and test hypotheses about population means. We are concerned here with the latter use. However, as we will see, our conclusions regarding the means will depend on the magnitudes of the observed variances.

The concepts and techniques that we cover under the heading of analysis of variance are extensions of the concepts and techniques covered in Chapter 7. In Chapter 7 we learned to test the null hypothesis that two means are equal. In this chapter we learn to test the null hypothesis that three or more means are equal. Whereas, for example, what we learned in Chapter 7 enables us to determine if we can conclude that two treatments differ in effectiveness, what we learn in this chapter enables us to determine if we can conclude that three or more treatments differ in effectiveness. The following example illustrates some basic ideas involved in the application of analysis of variance. These will be extended and elaborated on later in this chapter.

EXAMPLE 8.1.1

Suppose we wish to know if three drugs differ in their effectiveness in lowering serum cholesterol in human subjects. Some subjects receive drug A, some drug B, and some drug C. After a specified period of time, measurements are taken to determine the extent to which serum cholesterol was reduced in each subject. We find that the amount by which serum cholesterol was lowered is not the same in all subjects. In other words, there is *variability among* the measurements. Why, we ask ourselves, are the measurements not all the same? Presumably, one reason they are not the same is that the subjects received different drugs. We now look at the measurements of those subjects who received drug A. We find that the amount by which serum cholesterol was lowered is not the same among these subjects. We find this to be the case when we look at the measurements for subjects

8.1 INTRODUCTION 307

who received drug B and those subjects who received drug C. We see that there is *variability among* the measurements *within* the treatment groups. Why, we ask ourselves again, are these measurements not the same? Among the reasons that come to mind are differences in the genetic makeup of the subjects and differences in their diets. Through an analysis of the *variability* that we have observed, we will be able to reach a conclusion regarding the equivalence of the effectiveness of the three drugs. To do this we employ the techniques and concepts of analysis of variance. ■

Variables In our example we allude to three kinds of variables. We find these variables to be present in all situations in which the use of analysis of variance is appropriate. First, we have the *treatment variable*, which in our example was “drug.” We had three “values” of this variable, drug A, drug B, and drug C. The second kind of variable we refer to is the *response variable*. In the example it is change in serum cholesterol. The response variable is the variable that we expect to exhibit different values when different “values” of the treatment variable are employed. Finally, we have the other variables that we mention—genetic composition and diet. These are called *extraneous variables*. These variables may have an effect on the response variable, but they are not the focus of our attention in the experiment. The treatment variable is the variable of primary concern, and the question to be answered is: Do the different “values” of the treatment variable result in differences, on the average, in the response variable?

Assumptions Underlying the valid use of analysis of variance as a tool of statistical inference is a set of fundamental assumptions. Although an experimenter must not expect to find all the assumptions met to perfection, it is important that the user of analysis of variance techniques be aware of the underlying assumptions and be able to recognize when they are substantially unsatisfied. Because experiments in which all the assumptions are perfectly met are rare, analysis of variance results should be considered as approximate rather than exact. These assumptions are pointed out at appropriate points in the following sections.

We discuss analysis of variance as it is used to analyze the results of two different experimental designs, the completely randomized and the randomized complete block designs. In addition to these, the concept of a factorial experiment is given through its use in a completely randomized design. These do not exhaust the possibilities. A discussion of additional designs may be found in the references (4–6).

Example 8.2.1 – Codable

Example 8.2.2 – Codable

Example 8.3.1 – Codable

Example 8.4.1 – Codable

Example 8.4.2 – Codable

Example 8.5.1 – Codable

Example 8.5.2 – Codable

Chapter 9: Simple Linear Regression and Correlation

Example 9.3.1 – Codable

Example 9.4.1 – Codable

Example 9.4.2 – Codable

Example 9.4.3 – Codable

Example 9.7.1 – Codable

Example 9.7.2 – Codable

Chapter 10: Multiple Regression and Correlation

Example 10.3.1 – Codable

Example 10.4.1 – Codable

Example 10.4.2 – Codable

Example 10.4.3 – Codable

Example 10.5.1 – Codable

Example 10.6.1 – Codable

Example 10.6.2 – Codable

Chapter 11: Regression Analysis: Some Additional Techniques

Example 11.1.1 – Codable

Example 11.1.2 – Codable

Example 11.2.1 – Codable

Example 11.2.2 – Codable

Example 11.2.3 – Codable

Example 11.3.1 – Codable

Example 11.4.1 – Codable

Example 11.4.2 – Codable

Example 11.4.3 – Codable

Example 11.4.4 – Codable

Example 11.4.5 – Codable

Chapter 12: The chi-square distribution and the analysis of frequencies

Example 12.3.1 – Codable

Example 12.3.2 – Codable

Example 12.3.3 – Codable

Example 12.3.4 – Codable

Example 12.3.5 – Codable

Example 12.4.1 – Codable

Example 12.4.2 – Codable

Example 12.5.1 – Codable

Example 12.6.1 – Codable

Example 12.7.1 – Codable

Example 12.7.2 – Codable

Example 12.7.3 – Codable

Example 12.7.4 – Codable

Chapter 13: Nonparametric and distribution-free statistics

Example 13.3.1 – Codable

Example 13.3.2 – Codable

Example 13.4.1 – Codable

Example 13.5.1 – Codable

Example 13.6.1 – Codable

Example 13.7.1 – Codable

Example 13.8.1 – Codable

Example 13.8.2 – Codable

Example 13.9.1 – Codable

Example 13.9.2 – Codable

Example 13.10.1 – Codable

Example 13.10.2 – Codable

Example 13.11.1 – Codable

Example 13.11.2 – Codable

Chapter 14: Survival analysis

Example 14.3.1 – Codable

Example 14.4.1 – Codable

Example 14.5.1 – Codable