



# 1

## Introduction

### 1.1 Repeated Measurements

This book describes, discusses, and demonstrates a variety of statistical methods for the analysis of repeated measurements. The term “repeated measurements” refers broadly to data in which the response of each experimental unit or subject is observed on multiple occasions or under multiple conditions. Although the response variable could itself be either univariate or multivariate, we restrict consideration to univariate response variables measured at multiple occasions for each subject. The term “multiple” will usually mean “more than two,” since the topic of paired measurements is addressed in many other books.

The term “longitudinal data” is also often used to describe repeated measurements data. Some authors use this term when referring to data in which the repeated measurements factor is time. In this usage, longitudinal data could be viewed as a special case of repeated measurements data. Other authors make an alternative distinction and use the term “longitudinal data” to refer to data collected over an extended period of time, often under uncontrolled conditions. The term “repeated measurements” is then used to describe data collected over a relatively short time period, frequently under experimental conditions. Using this definition, repeated measurements data can be regarded as a special case of longitudinal data. In this book, we will use the term “repeated measurements” in the broad sense to refer to the situation in which multiple measurements of the response variable are obtained from each experimental unit.

Research in many areas of application frequently involves study designs in which repeated measurements are obtained. Studies in which the response variable is measured at multiple points in time from each subject are one important and commonly used application. In other applications, the response from each experimental unit is measured under multiple conditions rather than at multiple time points.

In some settings in which repeated measurements data are obtained, the independent experimental units are not individual subjects. For example, in a toxicological study, the experimental units might be litters; responses are then obtained from the multiple newborns in each litter. In a genetic study, experimental units might be defined by families; responses are then obtained from the members of each family.

## 1.2 Advantages and Disadvantages of Repeated Measurements Designs

A key strength of studies in which repeated measurements are obtained from each subject is that this is the only type of design in which it is possible to obtain information concerning individual patterns of change. This type of design also economizes on subjects. For example, when studying the effects of a treatment over time, it is usually desirable to observe the same subjects repeatedly rather than to observe different subjects at each specified time point. Another advantage is that subjects can serve as their own controls in that the outcome variable can be measured under both control and experimental conditions for each subject. Because between-subjects sources of variability can be excluded from the experimental error, repeated measurements designs often provide more efficient estimators of relevant parameters than cross-sectional designs with the same number and pattern of measurements. A final consideration is that data can often be collected more reliably in a study in which the same subjects are followed repeatedly than in a cross-sectional study.

There are two main difficulties in the analysis of data from repeated measures studies. First, the analysis is complicated by the dependence among repeated observations made on the same experimental unit. Second, the investigator often cannot control the circumstances for obtaining measurements, so that the data may be unbalanced or partially incomplete. For example, in a longitudinal study, the response from a subject may be missing at one or more of the time points due to factors that are unrelated to the outcome of interest. In toxicology or genetic studies, litter or family sizes are variable rather than fixed; hence, the number of repeated measures is not constant across experimental units.

Although many approaches to the analysis of repeated measures data have been studied, most are restricted to the setting in which the response

TABLE 1.1. General layout for repeated measurements

Subject	Time Point	Missing Indicator	Response	Covariates		
1	1	$\delta_{11}$	$y_{11}$	$x_{111}$	$\cdots$	$x_{11p}$
	$\vdots$	$\vdots$	$\vdots$	$\vdots$	$\ddots$	$\vdots$
	$j$	$\delta_{1j}$	$y_{1j}$	$x_{1j1}$	$\cdots$	$x_{1jp}$
	$\vdots$	$\vdots$	$\vdots$	$\vdots$	$\ddots$	$\vdots$
	$t_1$	$\delta_{1t_1}$	$y_{1t_1}$	$x_{1t_11}$	$\cdots$	$x_{1t_1p}$
.....						
$i$	1	$\delta_{i1}$	$y_{i1}$	$x_{i11}$	$\cdots$	$x_{i1p}$
	$\vdots$	$\vdots$	$\vdots$	$\vdots$	$\ddots$	$\vdots$
	$j$	$\delta_{ij}$	$y_{ij}$	$x_{ij1}$	$\cdots$	$x_{ijp}$
	$\vdots$	$\vdots$	$\vdots$	$\vdots$	$\ddots$	$\vdots$
	$t_i$	$\delta_{it_i}$	$y_{it_i}$	$x_{it_i1}$	$\cdots$	$x_{it_ip}$
.....						
$n$	1	$\delta_{n1}$	$y_{n1}$	$x_{n11}$	$\cdots$	$x_{n1p}$
	$\vdots$	$\vdots$	$\vdots$	$\vdots$	$\ddots$	$\vdots$
	$j$	$\delta_{nj}$	$y_{nj}$	$x_{nj1}$	$\cdots$	$x_{njp}$
	$\vdots$	$\vdots$	$\vdots$	$\vdots$	$\ddots$	$\vdots$
	$t_n$	$\delta_{nt_n}$	$y_{nt_n}$	$x_{nt_n1}$	$\cdots$	$x_{nt_np}$

variable is normally distributed and the data are balanced and complete. Although the development of methods for the analysis of repeated measures categorical data has received substantially less attention in the past, this has more recently become an important and active area of research. Still, the methodology is not nearly as well-developed as for continuous, normally distributed outcomes. The practical application of methods for repeated categorical outcomes also lags behind that for normal-theory methods due to the lack of readily accessible software.

### 1.3 Notation for Repeated Measurements

The notation used to describe methods for the analysis of repeated measurements varies considerably in the statistical literature. Table 1.1 shows the general layout for repeated measurements that will be used in this book. Let  $n$  denote the number of independent experimental units (subjects) from which repeated measurements are obtained, let  $t_i$  denote the number of

measurements from subject  $i$ , and let  $y_{ij}$  be the response from subject  $i$  at time point (or occasion)  $j$  for  $j = 1, \dots, t_i$  and  $i = 1, \dots, n$ . In addition, let  $p$  denote the number of covariates, and let  $\mathbf{x}_{ij} = (x_{ij1}, \dots, x_{ijp})'$  denote the vector of covariates associated with  $y_{ij}$ . In general, the values of the covariates may vary across the repeated measurements from a subject; such occasion-specific variables are called time-dependent or within-subject covariates. Because there may be missing values of  $y_{ij}$  and/or missing components in the vector  $\mathbf{x}_{ij}$ , it is convenient to define indicator variables

$$\delta_{ij} = \begin{cases} 1 & \text{if } y_{ij} \text{ and } \mathbf{x}_{ij} \text{ are observed,} \\ 0 & \text{otherwise.} \end{cases}$$

One special case of the general layout shown in Table 1.1 is when repeated measurements are obtained (or scheduled to be obtained) at a common set of  $t$  measurement occasions for all subjects. In this case,  $t_1 = \dots = t_n = t$ .

An important and commonly occurring situation is when repeated measurements are obtained from  $s$  subpopulations (groups) of subjects at a common set of  $t$  time points (or measurement occasions). In this case, let  $n_h$  be the number of subjects in group  $h$  for  $h = 1, \dots, s$ . In terms of the general notation,  $n = \sum_{h=1}^s n_h$ . The  $s$  groups may be defined by the  $s$  levels of a single covariate. In other situations, the groups may be defined by the cross-classification of the levels of several categorical covariates. In terms of the general layout shown in Table 1.1, the  $s$  groups can be described in terms of  $p = s - 1$  time-independent (or between-subject) categorical covariates. Although data of this type can be displayed using the general layout of Table 1.1, it may be more convenient to present the data as shown in Table 1.2. In this case, instead of letting  $y_{ij}$  denote the response at time  $j$  from subject  $i$ , we let  $y_{hij}$  denote the response at time  $j$  from subject  $i$  in group  $h$  for  $j = 1, \dots, t$ ,  $i = 1, \dots, n_h$ , and  $h = 1, \dots, s$ .

The final special case we will consider is the situation where repeated measurements are obtained (or scheduled to be obtained) at  $t$  time points from  $n$  subjects from a single population. In this case, the data can be displayed in an  $n \times t$  matrix, as shown in Table 1.3. Here,  $y_{ij}$  denotes the  $j$ th measurement from the  $i$ th subject for  $j = 1, \dots, t$ ,  $i = 1, \dots, n$ . The corresponding missing value indicators are defined by

$$\delta_{ij} = \begin{cases} 1 & \text{if } y_{ij} \text{ is observed,} \\ 0 & \text{otherwise.} \end{cases}$$

## 1.4 Missing Data

As was mentioned in Section 1.2, the occurrence of missing data is common in studies where repeated measurements are obtained. Although this book does not focus specifically on the analysis of incomplete repeated measurements, many of the methods described in subsequent chapters can

TABLE 1.2. Layout for the special case of multiple samples

Group	Subject	Time Point				
		1	...	$j$	...	$t$
1	1	$y_{111}$	...	$y_{11j}$	...	$y_{11t}$
	$\vdots$	$\vdots$	$\ddots$	$\vdots$	$\ddots$	$\vdots$
	$i$	$y_{1i1}$	...	$y_{1ij}$	...	$y_{1it}$
	$\vdots$	$\vdots$	$\ddots$	$\vdots$	$\ddots$	$\vdots$
	$n_1$	$y_{1n_1 1}$	...	$y_{1n_1 j}$	...	$y_{1n_1 t}$
.....						
$h$	1	$y_{h11}$	...	$y_{h1j}$	...	$y_{h1t}$
	$\vdots$	$\vdots$	$\ddots$	$\vdots$	$\ddots$	$\vdots$
	$i$	$y_{hi1}$	...	$y_{hij}$	...	$y_{hit}$
	$\vdots$	$\vdots$	$\ddots$	$\vdots$	$\ddots$	$\vdots$
	$n_h$	$y_{hn_h 1}$	...	$y_{hn_h j}$	...	$y_{hn_h t}$
.....						
$s$	1	$y_{s11}$	...	$y_{s1j}$	...	$y_{s1t}$
	$\vdots$	$\vdots$	$\ddots$	$\vdots$	$\ddots$	$\vdots$
	$i$	$y_{si1}$	...	$y_{sij}$	...	$y_{sit}$
	$\vdots$	$\vdots$	$\ddots$	$\vdots$	$\ddots$	$\vdots$
	$n_s$	$y_{sn_s 1}$	...	$y_{sn_s j}$	...	$y_{sn_s t}$

TABLE 1.3. Layout for the one-sample case

Subject	Time Point				
	1	...	$j$	...	$t$
1	$y_{11}$	...	$y_{1j}$	...	$y_{1t}$
$\vdots$	$\vdots$	$\ddots$	$\vdots$	$\ddots$	$\vdots$
$i$	$y_{i1}$	...	$y_{ij}$	...	$y_{it}$
$\vdots$	$\vdots$	$\ddots$	$\vdots$	$\ddots$	$\vdots$
$n$	$y_{n1}$	...	$y_{nj}$	...	$y_{nt}$

be used when the data are incomplete. The mechanism that results in missing data must, however, be considered when selecting an appropriate method of analysis. Little and Rubin (1987) and Schafer (1997) provide comprehensive treatments of the analysis of incomplete data. Laird (1988), Gornbein et al. (1992), Heyting et al. (1992), Little (1995), and Kenward and Molenberghs (1999) provide reviews focused specifically on repeated measurements.

In particular, Little and Rubin (1987) have described missing-data mechanisms as follows:

1. Missing completely at random (MCAR): if the probability of observing the response is independent of both the observed and unobserved outcome values;
2. Missing at random (MAR): if the probability of observing the response depends on the observed outcome values but is independent of the unobserved outcome values;
3. Nonignorable: if the probability of observing the response depends on the unobserved outcome values.

The nonignorable missing-data mechanism is also called informative or non-random.

With specific reference to repeated measurements, consider a study in which the outcome variable of interest is scheduled to be measured at a fixed number of occasions (visits) for each subject. The missing-data mechanism is MCAR if subjects miss their visits totally at random. A MAR missing-data mechanism would result if the probability of missing a visit is directly related to prior observed responses. An example of a nonrandom (nonignorable) missing-data mechanism would be if, in addition to prior observed responses affecting whether the response at a specific subsequent visit is missing, subjects would be more or less likely to miss a visit based on the unobserved value of their response at that specific visit.

In their discussion of missing data in repeated measurements, Diggle and Kenward (1994) refer to MCAR as the completely random dropout (CRD) mechanism. They propose the term “random dropout” (RD) for the MAR mechanism. The situation in which the missing-data mechanism is nonignorable is called the informative dropout (ID) mechanism.

The preceding characterizations of missing-data mechanisms refer only to the response variable and do not address the effect of covariates on the missing-data mechanism. For example, it may be important to consider the influence of a fully observed covariate on the probability of response. Little and Rubin (1987) have classified the mechanisms that govern missing data when the influence of a covariate is taken into account. If the probability of response is independent of the covariate and of the observed and unobserved responses, then the missing-data mechanism is said to be MCAR. If the

probability of response depends on the covariate but is independent of the unobserved responses, then the missing-data mechanism is said to be MAR provided that we have conditioned on the value of the covariate. If the probability of response depends on the unobserved responses with a possible (but not necessary) dependence on the covariate, then the missing-data mechanism is said to be nonignorable.

Suppose that the probability of observing a response depends on the value of the covariate but not on the observed and unobserved responses. For example, suppose that the probability of dropping out of a study varies according to the value of a covariate. Little and Rubin (1987) classify this mechanism as MAR due to the dependence on the covariate. There are, however, differing opinions on the classification of the missing-data mechanism in this situation. Diggle and Kenward (1994), among others, have classified this mechanism as MCAR provided that one conditions on the covariate in the analysis. Little (1995) suggests using the term covariate dependent dropout to describe this situation (provided that one conditions on all of the necessary covariates) and reserves the term MCAR only for a dropout that is independent of the covariate and observed and unobserved responses.

If the missing-data mechanism is MCAR, most standard approaches to analysis will be valid, and the issue of interest is simply the difficulty in implementing an analysis when the data are incomplete. In particular, analyses that omit experimental units with missing data (“complete case” analyses) are valid, although they may be inefficient. If the missing-data mechanism is MAR, then the nonresponse mechanism is said to be ignorable. In this case, likelihood-based inferences are still valid. Moment-based analysis methods, however, are biased when the missing-data mechanism is MAR. Although MAR is a weaker assumption than MCAR, nonignorable missing-data mechanisms are certainly much more common than either MCAR or MAR mechanisms.

If the missing-data mechanism is nonignorable, both likelihood-based and moment-based methods of analysis are biased. The development of methods for the analysis of repeated measurements that are valid in the case of nonignorable missingness is a difficult task.

Wu and Carroll (1988) discuss a special type of nonignorable missingness that they call “informative dropout;” this special case has been studied by several authors. In particular, Wu and Carroll (1988), Wu and Bailey (1989), and Mori et al. (1992, 1994) propose methodology for estimating the rate of change of a continuous repeated outcome when the dropout mechanism is informative. This approach has been extended to generalized linear mixed models (Follmann and Wu, 1995) and to repeated count data (Albert and Follmann, 2000).

Other authors have considered other types of models that adjust for nonignorable missingness. These include the approaches of Stasny (1987), Conaway (1992, 1993, 1994), Dawson and Lagakos (1993), Diggle and Ken-



ward (1994), Follmann et al. (1994), Cook and Lawless (1997), Molenberghs et al. (1997), and Albert (2000). Such methods for the analysis of repeated measurements when the missing-data mechanism is nonignorable are not yet available in standard statistical software packages.

As an alternative to parametrically modeling the dropout process, Verbeke et al. (2001) recommend the use of a sensitivity analysis based on local influence (Cook, 1986) to examine the potential effects of nonrandom dropout. Rotnitzky et al. (1998) also propose a procedure for carrying out a sensitivity analysis that examines how inferences concerning regression parameters change depending on assumptions about the nonresponse mechanism. Kenward (1998) provides an example illustrating the use of sensitivity analyses for repeated measurements. For normally distributed endpoints, Brown (1990) proposes a “protective” estimator that also does not require one to address the missingness model explicitly. Michiels and Molenberghs (1997) extend Brown’s approach to repeated categorical outcomes with nonrandom dropout.

## 1.5 Sample Size Estimation

This book describes methods for the analysis of data when the response variable is measured repeatedly for each independent experimental unit. Although the design of repeated measurements studies is equally important, this is not, however, a focus of the following chapters.

One important issue in study design is estimating the sample size required to detect an effect of a given magnitude with specified power or to estimate the power with which an effect of a given magnitude can be detected using a specified sample size. When the outcome variable is measured once for each experimental unit, procedures for estimating sample size and power are well-known and widely applied. The corresponding situation for repeated measurements data, however, is less well-developed. The complexity is due both to the fact that repeated observations from the same experimental unit are correlated and also that the repeated measurements situation requires more assumptions and parameters to be specified.

Lefante (1990), Kirby et al. (1994), and Overall and Doyle (1994) consider sample size estimation when the focus is on hypotheses characterized in terms of a univariate summary statistic across the repeated measurements. These approaches are relevant to the methods of analysis presented in Chapter 2 of this book. Overall et al. (1998) compare the Kirby et al. (1994) and Overall and Doyle (1994) approaches, and Ahn et al. (2001) provide a computer program for sample size estimation.

Several sample size estimation methodologies are available when the response at each time point is normally distributed. These approaches are relevant to the methods of analysis discussed in Chapters 3–6 of this book.

Bloch (1986) and Lui and Cumberland (1992) describe methods for sample size estimation based on the univariate split-plot analysis-of-variance model. Vonesh and Schork (1986) and Rochon (1991) provide sample size estimation procedures based on Hotelling's  $T^2$  statistic. Muller and Barton (1989) and Muller et al. (1992) consider sample size and power for the full multivariate analysis-of-variance model. Diggle et al. (1994, pp. 29–31) and Lindsey (2001) also discuss sample size estimation for normally distributed outcome variables.

Sample size estimation when the response variable at each time point is binary has also been studied; these approaches can be used in the situations discussed in Chapters 7–9 of this book. Lui (1991) and Shoukri and Martin (1992) extend the univariate split-plot model to the binary case. Lee and Dubin (1994) base their approach on the concept of the design effect from sample survey methodology. Rochon (1989) and Lipsitz and Fitzmaurice (1994) use weighted least squares procedures for sample size estimation with binary repeated measurements.

Approaches for estimation of sample size and power based on extensions of generalized linear model methodology to the repeated measurements situation are also available. These methods are useful in conjunction with the analysis approaches described in Chapter 9 of this book. Section 9.5.6 provides references and basic descriptions of the sample size estimation methods proposed by Liu and Liang (1997), Shih (1997), Rochon (1998), and Pan (2001b).

## 1.6 Outline of Topics

Many approaches to the analysis of repeated measurements have been proposed and studied. In addition, numerous books have been published dealing wholly or predominantly with the analysis of repeated measurements. Table 1.4 provides a listing of books that I am aware of that have their focus on statistical methodology for repeated measurements. Useful tutorials and articles reviewing methods for the analysis of repeated measurements include the papers by Everitt (1995), Cnaan et al. (1997), Albert (1999), and Omar et al. (1999). Diggle and Donnelly (1989) provide a selected bibliography on general methods for the analysis of repeated measurements.

Although I have found many of these other references to be quite useful, this book has a somewhat different purpose. Because it is often difficult to select, implement, and apply appropriate statistical methodology, I have sought to provide a broad survey of traditional and modern methods for the analysis of repeated measurements. Whereas some of the existing books are reasonably comprehensive in their coverage, others are more narrowly focused on specialized topics. This book is more comprehensive than many, and is targeted at a lower mathematical level and focused more on ap-

TABLE 1.4. Books focusing on methodology for repeated measurements

- 
- Crowder, M.J. and Hand, D.J. (1990). *Analysis of Repeated Measures*. Chapman and Hall, London.
- Davidian, M. and Giltinan, D.M. (1995). *Nonlinear Models for Repeated Measurement Data*. Chapman and Hall, London.
- Diggle, P.J. (1990). *Time Series: A Biostatistical Introduction*. Oxford University Press, New York.
- Diggle, P.J. et al. (1994, 2002). *Analysis of Longitudinal Data*. Oxford University Press, Oxford.
- Dwyer, J.H. et al. (1992). *Statistical Models for Longitudinal Studies of Health*. Oxford University Press, New York.
- Fahrmeir, L. and Tutz, G. (2001). *Multivariate Statistical Modelling Based on Generalized Linear Models*. Springer-Verlag, New York.
- Girden, E.R. (1992). *ANOVA: Repeated Measures*. Sage Publications, Newbury Park, CA.
- Goldstein, H. (1979). *The Design and Analysis of Longitudinal Studies: Their Role in the Measurement of Change*. Academic Press, New York.
- Hagenaars, J.A. (1990). *Categorical Longitudinal Data: Log-linear Panel, Trend, and Cohort Analysis*. Sage Publications, Newbury Park, CA.
- Hand, D.J. and Crowder, M.J. (1996). *Practical Longitudinal Data Analysis*. Chapman and Hall, London.
- Hand, D.J. and Taylor, C.C. (1987). *Multivariate Analysis of Variance and Repeated Measures*. Chapman and Hall, London.
- Jones, R.H. (1993). *Longitudinal Data with Serial Correlation: A State-Space Approach*. Chapman and Hall, London.
- Kshirsagar, A.M. and Smith, W.B. (1995). *Growth Curves*. Marcel Dekker, New York.
- Lindsey, J.K. (1999). *Models for Repeated Measurements*. Oxford University Press, New York.
- McCulloch, C.E. and Searle, S.R. (2000). *Generalized, Linear, and Mixed Models*. John Wiley and Sons, New York.
- Müller, H.G. (1988). *Nonparametric Regression Analysis of Longitudinal Data*. Springer-Verlag, Berlin.
- Nesselroade, J.R. and Baltes, P.B. (1980). *Longitudinal Methodology in the Study of Behavior and Development*. Academic Press, New York.
- Pan, J.X. and Fang, K.T. (2001). *Growth Curve Models with Statistical Diagnostics*. Springer-Verlag, New York.
- Pickles, A. (1990). *Longitudinal Data and the Analysis of Change*. Oxford University Press, New York.
- Plewis, I. (1985). *Analysing Change: Measurement and Explanation Using Longitudinal Data*. John Wiley and Sons, New York.
- Verbeke, G. and Molenberghs, G. (1997). *Linear Mixed Models in Practice*. Springer-Verlag, New York.
- Verbeke, G. and Molenberghs, G. (2000). *Linear Mixed Models for Longitudinal Data*. Springer-Verlag, New York.
- Vonesh, E.F. and Chinchilli, V.M. (1996). *Linear and Nonlinear Models for the Analysis of Repeated Measurements*. Marcel Dekker, New York.
- von Eye, A. (1990). *Statistical Methods in Longitudinal Research. Volumes I and II*. Academic Press, New York.
-

plications than most. It is designed to be used both as a textbook in a semester-length course and also as a useful reference for statisticians and data analysts.

I have attempted to provide sufficient background material on the methods that are presented to ensure that students and readers will have a good understanding of the methodology. At the same time, the focus is on applying the approaches discussed to real data. Because of this, there are numerous examples in each chapter as well as homework problems at the end of each chapter.

The remaining chapters discuss methods for the analysis of repeated measurements when the response variable is

- continuous and normally distributed;
- categorical;
- continuous and nonnormal.

Note that categorical outcome variables include dichotomous responses, polytomous variables (more than two possible values, not necessarily ordered), ordered categorical responses, and count variables. For each type of outcome variable, methods that can be used in the following settings are discussed:

- one sample ( $p = 0$ );
- multiple samples (one categorical covariate);
- multiple samples ( $p$  categorical covariates);
- regression (quantitative covariates).

Chapter 2 first discusses some simple univariate approaches to the analysis of repeated measurements. These methods involve reducing the multiple measurements obtained from each subject to a single “derived variable” or “summary statistic.” Chapters 3–6 discuss methods for normally distributed response variables. These chapters cover both traditional and modern approaches to the analysis of repeated measurements.

Chapter 7 then describes the weighted least squares approach for the analysis of categorical response variables. Chapter 8 presents the randomization model approach for the analysis of one-sample repeated measurements; this method can be applied both to categorical and continuous outcome variables. Chapter 9 describes extensions of generalized linear model methodology for the analysis of repeated measurements; these methods also can be used for categorical and continuous outcome variables. Finally, Chapter 10 discusses nonparametric methods for the analysis of repeated measurements.

## 1.7 Choosing the “Best” Method of Analysis

This book describes several methods for the analysis of repeated measurements. Although some are old and others are more recent, I have found all (with one exception to be mentioned later in this section) to be useful. Here are some guidelines for selecting an appropriate statistical method for a given application. Additional comments on the advantages and disadvantages of the various methods are provided in each chapter.

Chapter 2 discusses methods that reduce the vector of multiple measurements from each experimental unit to a single measurement. This approach avoids the issue of correlation among the repeated measurements from a subject and is often a useful preliminary or exploratory method of analysis. In situations where the distribution of the outcome variable is unusual, or where the sample size is too small or the number and pattern of repeated measurements are too irregular to permit the use of other methods, the univariate approach to the analysis of repeated measurements may be the only feasible one.

When the outcome variable at each time point is continuous and approximately normally distributed, the methods described in Chapters 3, 4, and 6 should be considered. Although Chapter 5 describes the use of classical repeated measures analysis of variance (ANOVA) for the analysis of continuous, normally distributed repeated measurements, I do not recommend the use of this methodology. I have included a short chapter on repeated measures ANOVA only because this approach is still widely used in some areas of application. Therefore, it is important to describe the restrictive assumptions and shortcomings of this methodology.

Chapter 6 discusses the linear mixed model, the most recent approach to the analysis of normally distributed repeated measurements. A natural question is whether the older multivariate analysis methods described in Chapters 3 and 4 are still necessary. First, the unstructured multivariate analysis approaches based on Hotelling’s  $T^2$  statistic, multivariate analysis of variance, and growth curve analysis are valid methods of analysis when repeated measurements are obtained at a fixed set of time points and there are no missing data. Second, the classical methods described in Chapters 3 and 4 are often based on fewer assumptions than are considered in practical applications of linear mixed model methodology. Third, because the unstructured multivariate analysis approaches are commonly used in some areas of application, familiarity with them is desirable. A final comment is that the simulation studies described in Section 6.5.3 indicate that the unstructured multivariate test statistics may perform better in small and moderate samples than the linear mixed model statistics. Thus, although the methods of Chapter 6 are important, the unstructured multivariate analysis approaches based on Hotelling’s  $T^2$  statistic, multivariate analysis of variance, and growth curve analysis are still often worthy of consideration.

When the outcome variable is categorical, the methods of Chapters 7–9 can be considered. Of these, the weighted least squares (WLS) methodology of Chapter 7 and the methods based on extensions of generalized linear model methodology (Chapter 9) are the most general approaches to the analysis of repeated categorical outcomes. Although some might argue that the methods of Chapter 9 are always to be preferred over the older WLS approach, the methods of Chapter 7 are quite useful when the number of repeated measurements is relatively small and all covariates are categorical. In particular, the WLS approach can be used for analyzing a wide variety of types of linear and nonlinear response functions and also provides a lack-of-fit statistic for assessing the appropriateness of the chosen model.

The methods described in Chapter 9 can also be used to analyze continuous repeated measurements when the marginal distribution at each time point is a member of the exponential family of distributions, such as the normal, gamma, and inverse Gaussian distributions. In particular, when the response is approximately normally distributed, the methods in Chapter 9 provide alternatives to the methods of Chapters 3–6 that may be more robust to departures from assumptions. Wu et al. (2001) discuss the relationships between the methods of Chapters 6 and 9 when the data are normally distributed.

Section 9.8 describes methods appropriate for the analysis of ordered categorical outcomes. These offer the advantage of being able to accommodate continuous covariates but require the restrictive proportional-odds assumption. The WLS approach (Chapter 7) can fit more flexible models to ordered categorical responses and also provides an overall goodness-of-fit test. The disadvantages are that covariates must be categorical and that the sample size must be quite large if any of (a) the number of levels of the response variable, (b) the number of time points, or (c) the number of levels of the cross-classification of the covariates is large.

Chapter 8 discusses the randomization model approach using Cochran–Mantel–Haenszel (CMH) statistics. This methodology requires minimal assumptions concerning the distribution of the response and can be used for both continuous and categorical outcomes. In addition, CMH statistics are applicable in situations where the sample size is too small to justify the use of alternative approaches. The major shortcoming of this method is that it is appropriate only for one-sample problems (i.e., when there are no covariates). In addition, the randomization model approach provides procedures for hypothesis testing only; it is not possible to estimate the parameters of a model.

When the response variable is continuous but nonnormal, nonparametric approaches (Chapter 10) may be the only reasonable option other than the summary-statistic approach. In this case, the Chapter 10 approaches allow one to consider the multivariate nature of the data rather than reducing the multiple responses to a summary measure. The shortcomings include the

lack of estimation procedures and the fact that the repeated measurements nature of the data is not fully taken into account.

Statistical Methods for the Analysis of Repeated  
Measurements

Davis, C.S.

2002, XXIV, 416 p. 72 illus., Hardcover

ISBN: 978-0-387-95370-0