There is another way in which we can represent the structure given by (1.8) that allows a flexible generalisation. Consider a model whose random part is given by

$$u_{0j} + u_{1j}x_{1ij} + e_{0ij} + e_{1ij}x_{3ij},$$

where now we assume that e_{0ij} has variance σ^2_{e0}, e_{0ij} and e_{1ij} have covariance σ_{e01}, and (somewhat bizarrely) e_{1ij} has variance equal to zero. Then the level-1 contribution to the total variance is given by

$$\sigma^2_{e0} + 2x_{3ij}\sigma_{e01},$$

which amounts to a variance of σ^2_{e0} for boys and one of $\sigma^2_{e0} + 2\sigma_{e01}$ for girls (remember that the covariance may be positive or negative, so that either one of the variances may be the larger of the two). The constraining of one of the variances to equal zero while permitting a non-zero covariance is a device to introduce model complexities into the structure of the variability that can be extended in many directions. In particular, it allows us to model the level-1 variation as a linear function of several explanatory variables (for a more detailed discussion, see Goldstein, 1995, Chapter 3). The terms σ^2_{e0} and σ_{e01} are best thought of as parameters in such a linear function rather than as variances and covariances in the usual sense.

In this chapter, the basic multilevel models have been presented; in subsequent chapters, there will be further elaborations with applications to substantive areas.

246.

CHAPTER 2

Modelling repeated measurements

Harvey Goldstein and Geoff Woodhouse
Mathematical Sciences, Institute of Education, University of London, UK

2.1 INTRODUCTION

When measurements are repeated on the same subjects, for example students or animals, a two-level hierarchy is established with measurement repetitions or occasions as level-1 units and subjects as level-2 units. Such data are often referred to as 'longitudinal' as opposed to 'cross-sectional' where each subject is measured only once. Thus, we may have repeated measures of body weight on growing animals or children, repeated test scores on students or repeated interviews with survey respondents. Figure 2.1 is a plot of height measurements on each of four boys (Goldstein, 1989a; see below) between the ages of 10.5 and 16.5 years. Several things are worth noting. First, for each boy, the

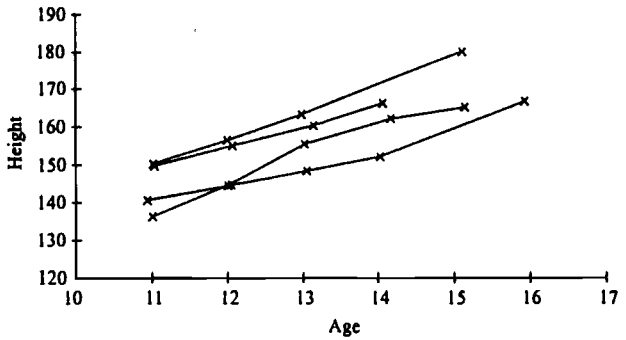

Figure 2.1 Repeated height measurements for four boys.

Multilevel Modelling of Health Statistics Edited by A.H. Leyland and H. Goldstein

growth curve is very approximately a straight line. Secondly, if we drew a simple regression line for each boy, the variation about this line would be small relative to the variation between the lines. In other words the level-1 (within-individual-between-occasion) variation is smaller than the level-2 (between-individual) variation. In contrast with many other two-level data sets, where most of the variation is at level 1, we have a strong hierarchical structure where any failure to model it will result in serious model misspecification. Finally, we note that the lines for each boy have varying slopes; they grow at different rates and we will need to fit models of greater complexity than simple variance component models. Also, as we shall see below, we can fit more complex functions than straight lines to these data.

It is important to distinguish two types of model for repeated measurement data. In one, earlier measurements are treated as covariates rather than responses. In the other, as in the growth example, all the measurements are considered as responses and are related to time or age. The first case will often arise when there are a small number of distinct occasions and where *different* measures are used at each one. In this situation, it will often make little sense to study how the measures are related to age or time: to do so would require us to standardise each measurement to a common metric, but this would still leave problems of interpretation. Plewis (1993) discusses a standardisation where the coefficient of variation at each age is fixed to have a constant value. In general, however, different standardisations may be expected to lead to different inferences. The choice of standardisation is in effect a choice about the appropriate scale along which measurements can be 'equated', so any interpretation needs to recognise this.

In the second case, which is usually referred to as a 'repeated measures' model, it is more natural to ask questions about how the relationship between a common measure such as height or weight changes with age, and it is this class of models that we shall discuss here. A detailed description of the distinction between the former 'conditional' models and the latter 'unconditional' models can be found in Goldstein (1979) and Plewis (1985).

We may also have repetition at higher levels of a data hierarchy. For example, we may have annual data about smoking habits on successive cohorts of 16-year-old students in a sample of schools. In this case, the school is the level-3 unit, year is the level-2 unit and student the level-1 unit. We may even have a combination of repetitions at different levels: in the previous example, with the students themselves being questioned on successive occasions. We shall also look at an example where there are responses at both level 1 and level 2, that is specific to the occasion and to the subject.

The link with multivariate data models (see Chapter 5) is also apparent where the occasions are fixed. This can be seen in Table 2.1 where we have four measurements on each individual; the first subscript refers to occasion and the second to individual.

We can regard this as a multivariate response vector with four responses for each child, and specify a model, for example relating the measurements to a polynomial function of age. This multivariate approach has traditionally been

Table 2.1 Measurements at four occasions for three individuals.

Individual	Occasion 1	Occasion 2	Occasion 3	Occasion 4
1	y_{11}	y_{21}	y_{31}	y_{41}
2	y_{12}	y_{22}	y_{32}	y_{42}
3	y_{13}	y_{23}	y_{33}	y_{43}

used with repeated measures data (Grizzle and Allen, 1969). It cannot, however, deal with data with an arbitrary spacing of time points or number of occasions, and we shall not consider it further.

In all the models considered so far, we have assumed that the level-1 residuals are uncorrelated. For some kinds of repeated measures data, however, this assumption will not be reasonable, and we shall also investigate models that allow a serial correlation structure for these residuals.

Our examples deal only with continuous response variables, but a discussion of how to apply these procedures where responses are discrete will be given at the end of the chapter.

2.2 A TWO-LEVEL REPEATED MEASURES MODEL

Consider a data set consisting of repeated measurements of the heights of a random sample of children. Thus, for the data in Figure 2.1, we can write a simple model with linear growth as

$$y_{ij} = \beta_{0j} + \beta_{1j}x_{ij} + e_{ij}. \tag{2.1}$$

This model assumes that height Y is linearly related to age X, with each subject having their own intercept and slope, so that

$$E(\beta_{0j}) = \beta_0, \quad E(\beta_{1j}) = \beta_1,$$
$$\text{var}(\beta_{0j}) = \sigma^2_{u0}, \quad \text{var}(\beta_{1j}) = \sigma^2_{u1}, \quad \text{cov}(\beta_{0j}, \beta_{1j}) = \sigma_{u01}, \quad \text{var}(e_{ij}) = \sigma^2_e.$$

There is no restriction on the number or spacing of ages, so that we can fit a single model to subjects who may have one or several measurements. We can clearly extend (2.1) to include further explanatory variables, measured either at the occasion level, such as time of year or state of health, or at the subject level such as birthweight or gender. We can also extend the basic linear function in (2.1) to include higher-order terms and we can further model the level-1 residual so that the level-1 variance is a function of age (see Chapter 1).

Table 2.2 presents the results of an analysis fitting (2.1) and also a model that includes further polynomial growth terms. The data consist of 436 measurements of the heights of 108 boys between the ages of 11 and 16 years (Goldstein, 1989a). For convenience, age is now measured about the (approximate) mean age of 13.0 years. When we calculate polynomial terms and fit random coefficients this 'centring' will avoid numerical problems arising from approximate collinearities.

Table 2.2 Height (cm) for adolescent growth, bone age, and adult height for a sample of boys. Age measured about 13.0 years. Level-2 variances and correlations are shown. All random parameters are significant at the 5% level.

Parameter				Model A estimate (SE)	Model B estimate (SE)
Fixed					
Intercept				153.2	153.1
Age				7.10 (0.14)	7.06 (0.17)
Age^2				0.25 (0.06)	0.32 (0.06)
Age^3				−0.21 (0.02)	−0.21 (0.03)
Random					
Level 2:					
	Intercept	Age	Age^2		
Intercept	59.3				52.2
Age	0.39	0.79			
Age^2	−0.49	−0.35	0.19		
Level 1:					
σ_e^2				1.32	4.49
−2 log-like.				2182.6	2300.6

The simple 'variance components' model, which fits only an intercept at level 2, is a poor fit, as shown by the deviance statistic of 118.0 with five degrees of freedom. A variance components model, sometimes known as a 'compound symmetry' model, is anyway implausible since it assumes that the correlation between two measurements is $\sigma_{u0}^2(\sigma_{u0}^2 + \sigma_e^2)^{-1}$, the intra-unit correlation, and hence does not depend on the age difference. In fact, for these data, we can go on to fit a quartic term in the fixed part and make the coefficient of the cubic term random at the individual level; we have omitted this extended model for simplicity.

For each individual, we can estimate the posterior level-2 residuals for the intercept, linear and quadratic coefficients (see Chapter 1). Using these, we may therefore construct the predicted growth curves for each individual. Figure 2.2 shows these for the same four individuals as in Figure 2.1, and we can see the very different growth patterns.

We could go on to further elaborate this model in a straightforward way by adding covariates, for example social class, allowing us to investigate how the growth patterns vary by type of child. It is also possible to elaborate the model in an interesting way by including further response variables so defining a multivariate repeated measures model. This has a number of useful properties, as we shall explain below.

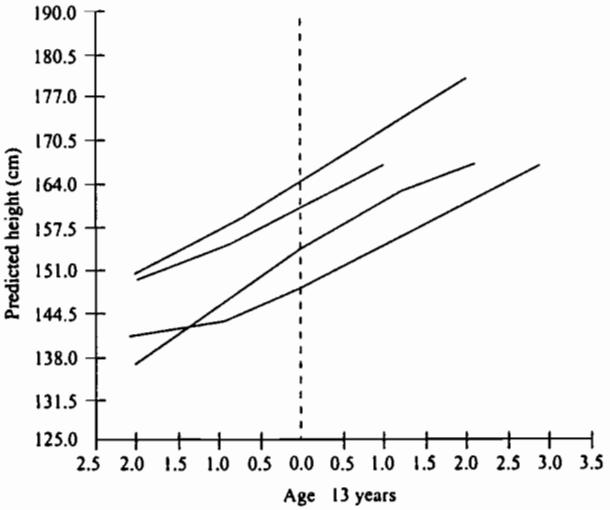

Figure 2.2 Predicted growth patterns for four individuals; fitting the model in Table 2.2.

2.3 A POLYNOMIAL MODEL EXAMPLE FOR ADOLESCENT GROWTH IN HEIGHT AND BONE AGE TOGETHER WITH ADULT HEIGHT

Our next example combines the basic two-level repeated measures model with a multivariate model to show how a general growth prediction model can be constructed. The data are as before, together with measurements of their height as adults and estimates of their bone ages at each height measurement based upon wrist radiographs. We first write down the three basic components of the model, starting with a simple repeated measures model for height using a third-degree polynomial with coefficients up to the quadratic random at the individual level:

$$y_{ij}^{(1)} = \sum_{h=0}^{3} \beta_h^{(1)} x_{ij}^h + \sum_{h=0}^{2} u_{hj}^{(1)} x_{ij}^h + e_{ij}^{(1)}, \qquad (2.2)$$

where the level-1 term e_{ij} is now allowed to have a complex structure, for example a decreasing variance with increasing age, and the x_{ij}^h represent powers of the child's age.

The measure of bone age is already standardised, since the average bone age for boys of a given chronological age is equal to this age for the population. Thus we model bone age using an overall constant to detect any average departure for this group together with between-individual and within-individual variation ($u_{0j}^{(2)}$ and $u_{1j}^{(2)}$ respectively):

$$y_{ij}^{(2)} = \beta_0^{(2)} + \sum_{h=0}^{1} u_{hj}^{(2)} x_{ij}^h + e_{ij}^{(2)}. \tag{2.3}$$

For adult height, we have a simple model with an overall mean and level-2 variation given by

$$y_j^{(3)} = \beta_0^{(3)} + u_{0j}^{(3)}. \tag{2.4}$$

If we had more than one adult measurement on individuals, we would be able to estimate also the level-1 (within-individual) variation among adult height measurements; in effect measurement errors. We now combine (2.2)–(2.4) into a single model using the following indicators.

$\delta_{ij}^{(1)} = 1$ if growth period measurement, 0 otherwise;

$\delta_{ij}^{(2)} = 1$ if bone age measurement, 0 otherwise;

$\delta_j^{(3)} = 1$ if adult height measurement, 0 otherwise;

$$\begin{aligned} y_{ij} = {} & \delta_{ij}^{(1)} \left(\sum_{h=0}^{3} \beta_h^{(1)} + \sum_{h=0}^{2} u_{hj}^{(1)} x_{ij}^h + e_{ij}^{(1)} \right) \\ & + \delta_{ij}^{(2)} \left(\beta_0^{(2)} + \sum_{h=0}^{1} u_{hj}^{(2)} x_{ij}^h + e_{ij}^{(2)} \right) \\ & + \delta_j^{(3)} (\beta_0^{(3)} + u_{0j}^{(3)}). \end{aligned} \tag{2.5}$$

This is now a multivariate model, with the multivariate structure specified at level 1 using the three dummy variables above. Since adult height is defined at the individual level, its residual can only co-vary with random coefficients at that level and not at level 1. The variances and a covariance between bone age and height are specified at level 2, and the between-individual variation involving height, bone age and adult height at level 3. In fact, for simplicity, we shall assume that the residuals for bone age and height at level 2 are largely 'measurement errors', and hence it is reasonable to assume they will be independent, although dependences might arise, for example if the model were incorrectly specified at level 2. Table 2.3 shows the fixed and random parameters for this model, omitting the estimates for the between-individual variation in the quadratic and cubic coefficients of the polynomial growth curve.

From the positive value of the bone age intercept we infer that this sample is slightly advanced compared with the general population, but with a large between-individual variance of 0.70.

We see that there are non-zero correlations between adult height and both the height intercept and growth coefficients, but a smaller correlation between adult height and the bone age intercept. This suggests that the growth measurements can be used to make predictions of adult height, but that little is gained by including the bone age. To predict adult height, we require the estimated residuals for adult height from the model. For a new individual, with information

Table 2.3 Height (cm) for adolescent growth, bone age and adult height for a sample of boys. Age measured about 13.0 years. Level-2 variances and correlations shown.

Parameter	Estimate (SE)
Fixed	
Adult height:	
Intercept	174.6
Height:	
Intercept	153.1
Age	7.08 (0.16)
Age^2	0.30 (0.05)
Age^3	−0.20 (0.03)
Bone age:	
Intercept	0.21 (0.09)
Age	0.04 (0.02)

Random

Level 2:

	Adult height	Height intercept	Age	Age^2	Bone Intercept	Age
Adult height	63.4					
Height intercept	0.76	58.6				
Age	0.22	0.50	0.70			
Age^2	0.19	−0.50	−0.48	0.17		
Bone age Intercent.	0.06	0.58	0.34	−0.86	0.70	

Level 1 variances:	
Height	1.64
Bone age	0.36

available at one or more ages on height or bone age, we would estimate the adult height residual using the model parameters. This therefore provides a quite general method for predicting adult height using any collection of height and bone age measurements at a set of ages within the range fitted by the model. Table 2.4 shows the estimated standard errors associated with predictions made on the basis of varying amounts of information. It is clear that the main gain in efficiency comes with the use of height with a smaller gain from the addition of bone age.

The method can be used for any measurements, either to be predicted or as predictors. In particular, covariates such as family size or social background can be included to improve the prediction. We can also predict other events of interest, such as the estimated age at maximum growth velocity. Pan and Goldstein (1998), for example, provide estimates of growth rates and accelerations

Table 2.4 Standard errors for height predictions for specified combinations of height and bone age measurements.

Bone age measures		Height measures (age)		
		None	11.0	11.0 12.0
None			4.3	4.2
11.0		7.9	3.9	3.8
11.0	12.0	7.9	3.7	3.7

for individuals from any set of serial measurements taken during growth. They model height and weight in a bivariate response model and also provide 'conditional' predictions and norms for current weight or height given any set of previous weights and heights.

2.4 MODELLING AN AUTOCORRELATION STRUCTURE AT LEVEL-1

So far we have assumed that the level-1 residuals are independent. In many situations, however, such an assumption would be false. For growth measurements the specification of level-2 variation serves to model a separate curve for each individual, but the between-individual variation will typically involve only a few parameters, as in the previous example. We can think of each curve as a smooth summary of growth with small random departures at each measurement occasion. If, however, measurements on an individual are obtained very close together in time, they will have a similar departure from that individual's underlying growth curve. This implies that the level-1 residuals will be positively correlated; there will be 'autocorrelation' between them. Examples occur in other areas, such as economics, where measurements on each unit, for example an enterprise or economic system, exhibit an autocorrelation structure and where the parameters of the separate time series will vary across units at level-2.

A detailed discussion of multilevel time series models is given by Goldstein *et al.* (1994). They discuss both the discrete-time case, where the measurements are made at the same set of equal intervals for all level-2 units, and the continuous-time case, where the time intervals can vary. We shall develop the continuous-time model here, since it is both more general and flexible.

To simplify the presentation, we shall drop the level-1 and -2 subscripts and write a general model for the level-1 residuals as follows:

$$\mathrm{cov}(e_t e_{t-s}) = \sigma_e^2 f(s). \tag{2.6}$$

This states that the covariance between two measurements s units in time apart, depends on the level-1 variance (σ_e^2, which may be a function of age) and

Table 2.5 Some choices for the covariance function g for level-1 residuals.

$g = \beta_0 s$	For equal intervals, this is a first-order autoregressive series
$g = \beta_0 s + \beta_1(t_1 + t_2) + \beta_2(t_1^2 + t_2^2)$	For time points t_1 and t_2, this implies that the variance is a quadratic function of time
$g = \begin{cases} \beta_0 s & \text{if no replicate} \\ \beta_1 & \text{if replicate} \end{cases}$	For replicated measurements this gives an estimate of measurement reliability $\exp(-\beta_1)$
$g = (\beta_0 + \beta_1 z_{1j} + \beta_2 z_{2ij})s$	The covariance is allowed to depend on an individual level characteristic (e.g. gender) and a time-varying characteristic (e.g. season of the year or age)
$g = \begin{cases} \beta_0 s + \beta_1 s^{-1} & (s > 0) \\ 0 & (s = 0) \end{cases}$	Allows a flexible functional form, when the time intervals are not close to zero

a function involving the time difference. The latter function is conveniently described by a negative exponential reflecting the common assumption that with increasing time difference the covariance will tend to a fixed value, $\alpha\sigma_e^2$ (in the following example, we shall assume that this is zero, but in other cases this may not be reasonable):

$$f(s) = \alpha + \exp[-g(\beta, z, s)], \tag{2.7}$$

where β is a vector of parameters for further explanatory variables z. Some choices for g are given in Table 2.5.

If we assume multivariate normality for the response variable, maximum-likelihood estimates are available (details are given by Goldstein *et al.*, 1994).

We now have a model that consists of two distinct covariance structures: the between-individual and the within-individual. From the interpretational point of view it is convenient to have parameters that summarise individual characteristics such as average growth and rate of growth. From this point of view, the within-individual structure exists only to provide a full description of the covariance structure in order to obtain a properly specified model. In some situations, however, where data may not be very extensive, we may be able to describe the overall structure *either* by fitting a small number of higher-level random coefficients together with an elaborated serial correlation structure at level 1, *or* by an elaborate higher level structure and simple, independent, variation at level 1. Diggle (1988), for example, fitted a variance components model together with a level-1 serial correlation structure with $g = \beta_0 s$ to repeated measurements. Sometimes it may be possible to make a choice in terms of goodness of fit, for example using the AIC criterion based upon comparing deviances (Lindsey, 1999), but more generally the aim should be to parameterise the model so that a useful interpretation can be placed upon the parameters.

2.5 A GROWTH MODEL WITH AUTOCORRELATED RESIDUALS

The data for this example consist of a sample of 26 boys, each measured on nine occasions between the ages of 11 and 14 years (Harrison and Brush, 1990). The measurements were taken approximately three months apart. Table 2.6 shows the estimates from a model that assumes independent level-1 residuals with a constant variance. The model also includes a cosine term to model the seasonal variation in growth with time measured from the beginning of the year. If the seasonal component has amplitude α and phase γ, we can write

$$\alpha \cos(t + \gamma) = \alpha_1 \cos t - \alpha_2 \sin t.$$

In the present case, the second coefficient is estimated to be very close to zero, and is set to zero in the following model. This component results in an average growth difference between summer and winter estimated to be about 0.5 cm.

We now fit in Table 2.7 the model with $g = \beta_0 s$, which is the continuous-time version ofthe first-order autoregressive model.

The fixed part and level-2 estimates are little changed. The autocorrelation parameter implies that the correlation between residuals three months (0.25 years) apart is 0.18: $\exp(-\beta s) = \exp(-1.725) = 0.18$. For measurements six months apart, this drops to 0.03. This suggests that once measurements are taken less than three months apart, it will become important to fit a serial correlation model in order to specify the data structure correctly. Failure to do this will still provide consistent estimates for the fixed parameters, but will tend

Table 2.6 Height as a fourth-degree polynomial on age, measured about 13.0 years. Standard errors in parentheses; correlations in parentheses for covariance terms.

Parameter	Estimate (SE)		
Fixed			
Intercept	148.9		
Age	6.19 (0.35)		
Age^2	2.17 (0.46)		
Age^3	0.39 (0.16)		
Age^4	−1.55 (0.44)		
cos (time)	−0.24 (0.07)		
Random			
Level 2:			
	Intercept	Age	Age^2
Intercept	61.6 (17.1)		
Age	8.0 (0.61)	2.8 (0.7)	
Age^2	1.4 (0.22)	0.9 (0.67)	0.7 (0.2)
Level 1:			
σ_e^2	0.20 (0.02)		

Table 2.7 Height as a fourth-degree polynomial on age, measured about 13.0 years. Standard errors in parentheses; correlations in parentheses for covariance terms. Autocorrelation structure fitted for level-1 residuals.

Parameter	Estimate (SE)		
Fixed			
Intercept	148.9		
Age	6.19 (0.35)		
Age^2	2.16 (0.45)		
Age^3	0.39 (0.17)		
Age^4	−1.55 (0.43)		
cos (time)	−0.24 (0.07)		
Random			
Level 2:			
	Intercept	Age	Age^2
Intercept	61.5 (17.1)		
Age	7.9 (0.61)	2.7 (0.7)	
Age^2	1.5 (0.25)	0.9 (0.68)	0.6 (0.2)
Level 1:			
α_e^2	0.23 (0.04)		
β	6.90 (2.07)		

to underestimate standard errors and also not provide consistent estimates for the random parameters.

2.6 MULTIVARIATE REPEATED MEASURES MODELS

We have already discussed the bivariate repeated measures model where the level-1 residuals for the two responses are independent. In the general multivariate case where correlations at level 1 are allowed, we can fit a full multivariate model by adding a further lowest level as described in Chapter 5. For the autocorrelation model, this will involve extending the models to include cross-correlations. For example, for two response variables with the model of Table 2.7 we would write the cross-correlation as

$$g = \sigma_{e1}\sigma_{e2} \exp(-\beta_{12} s).$$

The special case of a repeated measures model where some or all occasions are fixed is of interest. We have already dealt with one example of this where adult height is treated separately from the other growth measurements. The same approach could be used with, for example, birthweight or length at birth. In some studies, all individuals may be measured at the same initial occasion, and we can choose to treat this as a covariate rather than as a response. This might be appropriate where individuals were divided into groups for different treatments following initial measurements.

2.7 CROSSOVER DESIGNS

A common procedure for comparing the effects of two different treatments, A and B, is to divide the sample of subjects randomly into two groups and then to assign A to one group followed by B, and B to the other group followed by A. The potential advantage of such a design is that the between-individual variation can be removed from the treatment comparison. A basic model for such a design with two treatments, repeated measurements on individuals and a single group effect can be written as follows:

$$y_{ij} = \beta_0 + \beta_1 x_{1ij} + \beta_2 x_{2ij} + u_{0j} + u_{2ij} x_{2ij} + e_{ij}, \tag{2.8}$$

where X_1 is a dummy variable for time period and X_2 is a dummy variable for treatment. In this model we have not modelled the responses as a function of time within treatment, but this can be added in the standard fashion described in previous sections. In the random part at level-2 we allow between-individual variation for the treatment difference, and we can also structure the level-1 variance to include autocorrelation or different variances for each treatment or time period.

One of the problems with such designs is so called 'carry-over' effects whereby exposure to an initial treatment leaves some individuals more or less likely to respond positively to the second treatment. In other words, the u_{2j} may depend on the order in which the treatments were applied. To model this, we can add an additional term to the random part of the model, say $u_{3j}\delta_{3ij}$, where δ_{3ij} is a dummy variable that is 1 when A precedes B and the second treatment is being applied, and zero otherwise. This will also have the effect of allowing level-2 variances to depend on the ordering of treatments. The extension to more than two treatment periods and more than two treatments is straightforward.

2.8 DISCRETE RESPONSE DATA

The methods we have described for continuous data can be used for discrete responses, with suitable modifications to the model and estimation procedures. We shall not go into detail here, but will sketch out a simple model for binary responses. Suppose we have data on whether or not teenage children smoke, measured at successive occasions approximately six months apart. The response is yes (1) or no (0), and we have explanatory variables such as age, gender, and social class. We may write a standard model as follows.

$$\left.\begin{aligned} \text{logit}(\pi_{ij}) &= (X\beta)_{ij} + u_j, \\ y_{ij} &\sim \text{Binomial}(1, \pi_{ij}), \end{aligned}\right\} \tag{2.9}$$

where y_{ij} is the observed response for the jth child at the ith occasion, $(X\beta)_{ij}$ is the fixed part linear predictor containing explanatory variables, and u_j is the random effect, assumed to have a normal distribution, for the jth child

measuring propensity to smoke in comparison with the population mean. The use of the logit link function is a standard procedure, and we assume that, given the fixed predictors and the individual propensity, we have independent binomial variation with probability π_{ij}, whether we observe smoking or not. Goldstein and Rasbash (1996) discuss the estimation issues for such models.

A major difficulty with (2.9) is that there will typically be many individuals who always smoke or never smoke, giving probabilities π_{ij} of 1 or 0. This implies that they have values at $\pm\infty$ for u_j. In practice, what happens if we attempt to fit such a model is that we encounter a great deal of 'underdispersion' because the level-1 variation is less than that required by the binomial assumption. One approach to this problem using a multivariate binary model is given by Yang *et al.* (2000) for the case of a small number of discrete occasions, and this approach is currently being extended to general repeated measures structures using a formulation similar to the time series model described above.

2.9 MISSING DATA

In repeated measures designs data are regarded as missing where one or more of the responses in a complete balanced design such as in Table 2.1 are unavailable. Several broad situations need to be distinguished. In the first, a response may be missing because of the study design or for reasons that are unconnected with the true, but unknown, value of the response. Thus we may deliberately design a study where each individual is measured for only a subset of occasions. Such 'rotation' designs may be practical if time is limited or where a researcher does not wish to impose too great a burden on any respondent. In other cases, the probability of being missing may depend on predictor variables in the model, but otherwise is unrelated to the model parameters, in particular the level-1 and level-2 random effects. For example, if males are more likely to have missing data than females and gender is a covariate in the model, inferences will be consistent. Situations such as these are said to involve 'ignorable' missingness.

The second situation is where the probability of a response being missing depends on the values of other observed responses. In this case, applying maximum-likelihood to the observed data yields estimates with the usual maximum-likelihood properties of consistency etc., so long as the model is properly specified.

The third situation is where the probability of an observation being missing depends on the unknown value of the observation itself. This 'non-ignorable' case is the most difficult case to deal with, and consistent estimates are possible only if one is prepared to make particular assumptions about the nature of the missingness mechanism or the distributions. Such assumptions are generally not robust, although applying a range of such assumptions as part of a general sensitivity analysis may be useful (Kenward, 1998).

In practice, care should be taken to eliminate missing data, or at least to attempt to understand its causes so that variables responsible for it can be

included as covariates in a model. Little (1995) reviews the various procedures for handling missing data.

2.10 CONCLUSIONS

We have shown how very general models for repeated measures data can be constructed, including data with responses at different levels, and models where there are varying numbers of occasions and time points with the addition, where necessary, of a time series structure. We have not discussed nonlinear models such as sometimes occur in growth studies, but see Goldstein (1995) and Palmer *et al.* (1991) for a discussion of these. There are now several computer packages that will fit some or all of the models described (see Chapter 13 for a discussion).

CHAPTER 3

Binomial Regression

Nigel Rice
Centre for Health Economics, University of York, UK

3.1 INTRODUCTION

The majority of health data do not lend themselves to simple model specification allowing a linear link function to relate a set of explanatory variables to a response measured on a continuous scale. Instead, it is quite common to observe outcomes of interest that are qualitative or limited in their range of measurement. Of these, perhaps the most commonly encountered are discrete responses where an outcome may take one of a number of discrete values of either a categorical or non-categorical nature. The simplest of this type of model is one where the dependent variable is binary assuming one of two values, which, without loss of generality, may be denoted by 1 and 0, representing, for example, the presence or absence of an attribute, the success or failure of a trial, or the occurrence or not of an event.

In this chapter, we consider multilevel models in which the dependent variable assumes discrete values. For example, we may be interested in investigating the relationship between lifestyle choices, such as the intake of alcohol, smoking habits and diet, and the incidence of specific diseases such as ischaemic heart disease (IHD). In such a study, we may wish to code an occurrence of IHD as 1 and a non-occurrence as 0, and in so doing, by construction, create a binary response. The coding of such an event in this manner allows the researcher to relate a qualitative response to a set of potential explanatory variables in a regression framework and subject the resulting parameter estimates to standard statistical tests of hypotheses.

In other circumstances, instead of observing responses on individuals, we may observe the outcomes of a group of individuals or repeated experiments on the same individual. Often such data are expressed in terms of proportions, for example the proportion of patients who have shown a favourable response to a particular treatment in a clinical trial. Once again, such responses can be related to a set of explanatory variables and modelled adequately in a regression framework in much the same way as when we observe binary variables.

Multilevel Modelling of Health Statistics Edited by A.H. Leyland and H. Goldstein