Chapter 5

Random Effects and the Linear Mixed Model

Until now, we have considered primarily the estimation of regression parameters from the linear model, i.e., β in the model

$$E\ (Y_i|X_i) = X_i\beta, \tag{5.1}$$

as well as the variance parameters when var(Y_i) has an arbitrary structure. In this chapter we consider the use of random effects in modeling longitudinal data or clustered outcomes. Many researchers view it as more natural to assume that the mean response depends upon a combination of population parameters β and subject-specific effects. In the setting of the linear model where Y_i is linear in the parameters and the error terms, it is natural to also assume Y_i is linear in the subject-specific effects. As we will show, this still leads to the linear model (5.1), but var(Y_i) now has a special random effects structure.

The use of random effects offers several benefits when modeling longitudinal data. First, it provides a way to model correlation in unbalanced designs. Secondly, random effects can be used to estimate subject-specific effects arising in several applications. Finally, it offers an optimal way to combine within- and between-subject data.

Random effects are useful when strict measurement protocols are not followed and we have measurements made at arbitrary, irregularly spaced intervals. It is not desirable to design a study in this way, but such data sets are not uncommon. It can happen that we start with a strict protocol but because of missingness and missed timing, we end up with measurement times that do not conform to a set of protocols-defined occasions. Use of retrospectively collected records for analysis