

1

Linear Mixed-Effects Models: Basic Concepts and Examples

Many common statistical models can be expressed as linear models that incorporate both *fixed effects*, which are parameters associated with an entire population or with certain repeatable levels of experimental factors, and *random effects*, which are associated with individual experimental units drawn at random from a population. A model with both fixed effects and random effects is called a *mixed-effects* model.

Mixed-effects models are primarily used to describe relationships between a response variable and some covariates in data that are grouped according to one or more classification factors. Examples of such *grouped data* include *longitudinal data*, *repeated measures data*, *multilevel data*, and *block designs*. By associating common random effects to observations sharing the same level of a classification factor, mixed-effects models flexibly represent the covariance structure induced by the grouping of the data.

In this chapter we present an overview of linear mixed-effects (LME) models, introducing their basic concepts through the analysis of several real-data examples, starting from simple models and gradually moving to more complex models. Although the S code to fit these models is shown, the purpose here is to present the motivation for using LME models to analyze grouped data and not to concentrate on the software for fitting and displaying the models. This chapter serves as an appetizer for the material covered in later chapters: the theoretical and computational methods for LME models described in Chapter 2 and the linear mixed-effects modeling facilities available in the `nlme` library, covered in detail in Chapter 4.

The examples described in this chapter also serve to illustrate the breadth of applications of linear mixed-effects models.

1.1 A Simple Example of Random Effects

The data shown in [Figure 1.1](#) are from an experiment in nondestructive testing for longitudinal stress in railway rails cited in Devore (2000, Example 10.10, p. 427). Six rails were chosen at random and tested three times each by measuring the time it took for a certain type of ultrasonic wave to travel the length of the rail. The only experimental setting that changes between the observations is the rail. We say these observations are arranged in a *one-way classification* because they are classified according to a single characteristic—the rail on which the observation was made. These data are described in greater detail in Appendix A.26.

The quantities the engineers were interested in estimating from this experiment are the average travel time for a “typical” rail (the *expected travel time*), the variation in average travel times among rails (the *between-rail variability*), and the variation in the observed travel times for a single rail (the *within-rail variability*). We can see from [Figure 1.1](#) that there is considerable variability in the mean travel time for the different rails. Overall the between-rail variability is much greater than the within-rail variability.

The data on the rails experiment are given in an object called `Rail` that is available with the `nlme` library. Giving the name `Rail` by itself to the `S` interpreter will result in the data being displayed.

```
> Rail
Grouped Data: travel ~ 1 | Rail
  Rail travel
    1      55
    2      53
    3      54
  . . .
   17      85
   18      83
```

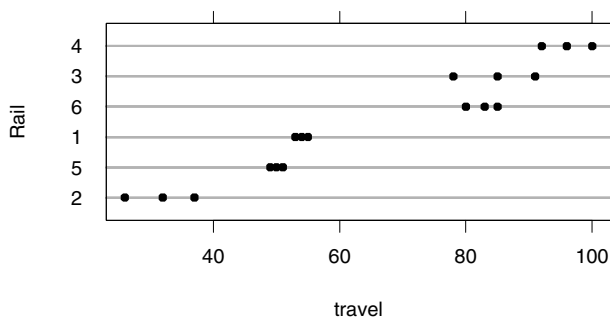


FIGURE 1.1. Travel time in nanoseconds for ultrasonic head-waves in a sample of six railroad rails. The times shown are the result of subtracting 36,100 nanoseconds from the original observation.