18

Pattern-Mixture Models

18.1 Introduction

The high sensitivity of selection modeling results to the correct specification of the measurement model as well as the dropout model, about which little is often known, has been extensively documented. See also Sections 15.3, 15.4, 17.1, 17.2.2, and 17.5. This has lead to growing interest in pattern-mixture modeling, based on the factorization (15.2) (Little 1993, Glynn, Laird and Rubin 1986, Hogan and Laird 1997). After initial mention of pattern-mixture models (Glynn, Laird, and Rubin 1986, Little and Rubin 1987), they are receiving more attention lately (Little 1993, 1994a, 1995, Hogan and Laird 1997, Ekholm and Skinner 1998, Molenberghs, Michiels, Kenward, and Diggle 1998, Molenberghs, Michiels, and Kenward 1998).

18.1.1 A Simple Illustration

We will first illustrate the idea of pattern-mixture modeling using a simple setting. Let us adopt pattern-mixture decomposition (15.2) and suppress dependence on covariates:

$$f(\boldsymbol{y}_i, \boldsymbol{r}_i|\boldsymbol{\theta}, \boldsymbol{\psi}) = f(\boldsymbol{y}_i|\boldsymbol{r}_i, \boldsymbol{\theta})f(\boldsymbol{r}_i, \boldsymbol{\psi}),$$

with notation as laid out in Chapter 15. Restricting attention to dropout (Section 15.9), we obtain, using (15.7),

$$f(\boldsymbol{y}_i, d_i|\boldsymbol{\theta}, \boldsymbol{\psi}) = f(\boldsymbol{y}_i|d_i, \boldsymbol{\theta})f(d_i|\boldsymbol{\psi}). \tag{18.1}$$

Equivalently, using (15.8),

$$f(\boldsymbol{y}_i, t_i|\boldsymbol{\theta}, \boldsymbol{\psi}) = f(\boldsymbol{y}_i|t_i, \boldsymbol{\theta})f(t_i|\boldsymbol{\psi}). \tag{18.2}$$

Consider a continuous response at three times of measurement which will be modeled using a trivariate Gaussian distribution. Assume that there may be dropout at time 2 or 3, and let the dropout indicator T_i take the values 1 and 2 to indicate that the last observation occurred at these times and 3 to indicate no dropout. Then, in the first instance, the model implies a different distribution for each time of dropout. We can write

$$\boldsymbol{y}_i \mid t_i \sim N(\boldsymbol{\mu}(t_i), \Sigma(t_i)), \tag{18.3}$$

where

$$\boldsymbol{\mu}(t) = \begin{pmatrix} \mu_1(t) \\ \mu_2(t) \\ \mu_3(t) \end{pmatrix} \quad \text{and} \quad \Sigma(t) = \begin{pmatrix} \sigma_{11}(t) & \sigma_{21}(t) & \sigma_{31}(t) \\ \sigma_{21}(t) & \sigma_{22}(t) & \sigma_{32}(t) \\ \sigma_{31}(t) & \sigma_{32}(t) & \sigma_{33}(t) \end{pmatrix},$$

for $t = 1, 2, 3$. Recall that t indicates length of sequences, as defined in Section 15.9, rather than time points of measurements actually taken. Let $P(t) = \pi_t = f(t_i|\boldsymbol{\psi})$, then the marginal distribution of the response is a mixture of normals with, for example, mean

$$\boldsymbol{\mu} = \sum_{t=1}^{3} \pi_t \boldsymbol{\mu}(t).$$

Its variance can be derived by application of the delta method (see Sections 18.3, 18.4, 20.6.2, and 24.4.2).

However, although the π_t can be simply estimated from the observed proportions in each dropout group, only 16 of the 27 response parameters can be identified from the data without making further assumptions. These 16 comprise all the parameters from the completers plus those from the following two submodels. For $t = 2$

$$N\left(\begin{pmatrix} \mu_1(2) \\ \mu_2(2) \end{pmatrix}; \begin{pmatrix} \sigma_{11}(2) & \sigma_{21}(2) \\ \sigma_{31}(2) & \sigma_{32}(2) \end{pmatrix}\right),$$

and for $t = 1$

$$N\left(\mu_1(1); \sigma_{11}(1)\right).$$