CHAPTER 8

The Semiparametric MLE

Although there are certainly examples where interest in the mixture NPMLE method has focused on the estimated latent distribution itself or functionals of it, the largest class of investigations and applications has occurred in the arena of semiparametric estimation, in which the latent distribution is included in the model to allow extra heterogeneity, but the focus is on a set of auxiliary parameters, generally of the regression type.

The earliest extensive applied investigation of semiparametric mixture maximum likelihood was by Heckman and Singer (1984), who were investigating the effects of latent distribution misspecification and were comparing, therefore, the nonparametric and parametric approaches to modeling the latent distributions. Among the many further investigations of this type we might count Follman and Lambert (1989), Brännäs and Rosenqvist (1994), Butler and Louis (1992), and Davies (1993). See Lindsay and Lesperance (1995) for a survey of the results in this area.

Since there exists a substantial literature on the implementation of this methodology, we will focus herein instead on special simplifying structures that can exist when we have a likelihood with both latent and auxiliary parameters. We will consider three different semiparametric models. One is an exponential family random effects model, of which the Rasch model is an illustration. The second is a measurement error problem, with the additional complication of case-control sampling. The final problem is an outlier distribution model for contingency tables that leads to a new method for assessing the fit of a parametric model.

We start with a general optimization result that is important because it leads to a simplification and clarification of what would otherwise be numerically more difficult problems. It will be used in the examples.

8.1. An equivalence theorem. In this section, we will use the notation β and G to designate two arbitrary "parameters" upon which our model depends, but in our applications, β will be real-valued parameters of interest and G will be the latent distribution. Our interest is in whether the β parameters