CHAPTER 3

Parametric Models

As an aid to understanding the role of nonparametric maximum likelihood, this chapter is a review of some of the basic features of the standard parametric models used in mixture modeling, together with basic features of maximum likelihood estimation in these models. We will consider the two different schools of modeling. In one, the latent distributions Q are assumed discrete, with a fixed number of components m. In the other, the latent distributions is assumed to come from some parametric family of continuous distributions.

3.1. Discrete versus continuous. There are certainly instances in which the latent distribution is logically and naturally modeled in either the discrete or continuous form due to the nature of the application. For example, it might be known that there are a finite and known number m of physical components in the population or, alternatively, that there is an inherent continuity expected in the latent variate Φ, such as when it represents a continuous covariate that was not measured.

More typically, however, we are on uncertain ground in specifying the number of components in a discrete latent distribution. Alternatively, in the case of a continuous variate Φ, we have no reason to believe it follows any particular parametric form. Although it is standard practice to assume normality for the latent variate, there is an element of arbitrariness in choosing the appropriate function of the latent variate to be normally distributed.

Moreover, as we have indicated already, we generally obtain very little goodness-of-fit information about the latent distribution from the data, and so there is little hope of having a large enough sample to investigate the true distributional form.

My own preference is for the use of discrete distributions in cases of doubt. The argument is that it makes maximum likelihood numerically simpler, and so is to be preferred if the form of the latent distribution specification makes little difference to the desired statistical inference. A number of investigations have found that misspecifying the latent distribution has very little effect on bias, and minimal effect on standard errors beyond the necessary correction