

7.7. BAYES' RULE WITH EMPIRICAL P.D.F. AND UNKNOWN DISTRIBUTIONS

The methods developed in the previous sections all involved the assumption that the distribution of the data is known. Indeed, we assumed the distributions to be multivariate normal. In practice this assumption may not always be valid. For example, rather than having a unimodal, symmetric distribution, it may be bimodal or even multimodal or may be skewed. Also, some of the data may be discrete while other data are continuous. We saw at the end of the last chapter how great could be the divergence between an assumed bivariate normal and the actual distribution of the observed data points. In this section we describe two alternative approaches to the estimation of the required p.d.f., namely, the kernel method and the nearest-neighbor method. These techniques are nonparametric; that is, we make no assumptions about the distributions belonging to a family of functions, described by distinct parameter values.