In statistical pattern recognition, the multivariate normal distribution is commonly assumed as the distribution describing the features. One reason for its use is that it is well studied and the properties related to marginal, conditional, and component distributions are simple. While it is possible to use the methods described below on features that follow other distributions, most computer packages provide programs only for the normal case. Hence it may be difficult to implement such methods when data are non-normal.

In practice, non-normality is not all that uncommon. For example, if we want to develop a classifier for deciding if an individual should get a bank loan, there are certainly going to be questions leading to binary features. For example:

"Have you ever had a previous loan?"

"Have you ever defaulted on a loan?"