As a simple example, consider Fig. 7.11, where the distributions of a single feature are represented. Note that the populations overlap. What this means is that some of the data from population c_1 will fall close to those of population c_2, and vice versa. One approach to developing a rule is based on the overlap between the distributions. Suppose, for example, that the rule we decide to use selects population c_1 if $x \leq A$ and population c_2 if $x > A$. Then sometimes the decision will be incorrect. Let $P(c_1|c_2)$ be the probability that population c_1 is selected when in fact the true population is c_2. This represents the probability of an incorrect decision, the probability of misclassification for group c_2. There is a similar probability for group c_1, $P(c_2|c_1)$.