

## 5.6 Cross-validation

The procedure of pulling out one sample at a time and checking the ability of the model to correctly classify that sample into its appropriate group is also called *cross-validation*. An important part of the CAP output from a discriminant type of analysis is the table showing the specific cross-validation results obtained for a chosen value of  $m$ . This gives specific information about *how distinct* the groups are and *how well* the PCO axes discriminate among the groups. No matter what patterns seem to be apparent from the CAP plot, nor how small the  $P$ -value from the permutation test (see the following section), this table of cross-validation results is actually the best way to assess the validity and utility of the CAP model. Indeed, we suggest that when using CAP for discrimination, no CAP plot should be presented without also providing cross-validation results, or at least providing the figure for overall misclassification error (or, equivalently, allocation success). This is because the CAP plot will look better and better (i.e., it will look more and more in tune with the hypothesis) the more PCO axes we choose to use. This does *not* mean, however, that the predictive capability of the underlying CAP model is improved! Indeed, we have just seen in the previous example how increases in the number of PCO axes (beyond  $m = 7$ ) actually *reduces* the allocation success of the model. So, the cross-validation provides a necessary check on the potential arbitrariness of the results.

Furthermore, the more detailed cross-validation results provided in the CAP output provide information about which groups are more distinct than others. Although, in this case, the groups had roughly comparable mis-classification errors (~70-76%, see Fig. 5.7), these errors can sometimes vary quite widely among the groups. The output file also indicates in which direction mistakes are made and for which individual samples this occurred. For example, looking at the cross-validation table for the Poor Knights fish data, 4 of the 15 samples from September 1998 were incorrectly classified as belonging to the group sampled in September 1999, while none were incorrectly classified as belonging to the group sampled in March 1999. Furthermore, the individual samples that were mis-classified (and the group into which they were erroneously allocated) are shown directly under the summary table. For example, the samples numbered 1, 2, 4 and 15 were the particular ones from September 1998 that were mis-classified (Fig. 5.7).

As a rule of thumb, bear in mind that, with three groups, one would expect an allocation success of around 33.3% simply by chance alone. Similarly, one would expect an allocation success of around 50% by chance in the case of two groups, or 25% in the case of 4 groups, etc. If the allocation success is substantially greater than would be expected by chance (as is the case for the Poor Knights data), then the CAP model obtained is a potentially useful one for making future predictions and allocations. Thus, the results of the cross-validation give a direct measure of the relative distinctiveness of the groups and also the potential utility of the model for future classification or prediction.