**Recursive Partitioning**

Once a best split is found, CART repeats the search process for each node below (child nodes), until either further splitting is stopped by a criterion, or splitting is impossible.

Common stopping conditions include:

- Minimum number of cases has been reached.
- A certain fraction of the total number of cases is in the node.
- A maximum number of levels of splitting has been achieved.
- The maximum number of nodes has been reached.

Conditions under which further splitting is impossible include when:

- Only one case is left in a node.
- All cases are duplicates of each other.
- The node is pure (all target values agree).

**Pruning Trees**

Rather than focusing on when to stop pruning, CART trees are grown larger than they need to be and then pruned back to find the best tree. CART determines the best tree by using the testing data set or by using the process of $V$-fold cross-validation. The testing validation is performed by scoring the tree with the data set not used for training the model. Cross-validation is a form of resampling, which draws a number of samples from the entire distribution and trains models on all samples. The $V$-fold cross-validation is performed by:

1. Partitioning the entire data set into a number ($V$) of parts (folds);
2. Training $V$ models on different combinations of $V - 1$ folds, with the error estimated each time using the $V$th fold;
3. Using the mean (and sigma) of the $V$ error measurements to estimate tree accuracy on new data;
4. Choosing the design parameters (e.g., complexity penalty) that minimize the error in step 3;
5. Refitting the tree, using all the data, using the parameters of step 4.

Figure 7.23 shows a 3-fold cross-validation operation.

The cross-validation process provides a number of independent estimates of the error associated with the algorithm itself rather than due to the randomness in the data. A model created with a CART algorithm (or any other algorithm, for that matter) should not be accepted until the prediction error is partitioned in this manner.

**General Comments about CART for Statisticians**

1. CART is nonparametric and does not require specification of a data distribution.
2. The final modeling variables are not selected beforehand but selected automatically by the algorithm.