

PERMUTATION TEST

It is useful to assess the strength of evidence of an association between the response variable and the prognostic variable (categorized into a binary variable using a cutoff value obtained from the results of the binary partitioning algorithm).

Partitioning methods in which splits are determined by the use of p-values are afflicted by the problem of multiple testing which arises due to the repeated ‘looks’ at the data that are involved in obtaining the optimal split [1]. Approaching binary partitioning from a hypothesis testing framework necessitates that the multiple testing issue be addressed; typically some penalty is applied, similar to a Bonferroni adjustment in the case of multiple hypotheses being tested.

Since we take the approach of obtaining an optimal split of the data based on a goodness-of-split measure, we are not in the same hypothesis testing scenario as that described above. Therefore we must approach this issue from a different angle.

A distribution-free permutation test [2] is used to assess the strength of evidence of an association between the prognostic variable which has been dichotomized into ‘high’ and ‘low’ levels via the binary partitioning method proposed in this paper and the response variable. Within the same framework as presented by Kim et al. [3] we implement the permutation test in the following several steps:

1. Calculate $\Delta D_{observed}(s^*, \mathcal{N})$, the optimal goodness-of-fit measure obtained from the observed data.
2. Permute the observed data to obtain J permutation data sets \mathcal{N}_j , where $j = 1, \dots, J$. This is simply achieved by randomly permuting the covariate values while holding the response variable fixed.
3. For each of these data sets, compute the optimal goodness-of-fit measure $\Delta D_j(s^*, \mathcal{N}_j)$, where $j = 1, \dots, J$.
4. The p-value is obtained from the permutational distribution of the goodness-of-fit statistics. In fact, the p-value can be simply calculated as

$$p = \frac{\sum_{j=1}^J [I(\Delta D_j(s^*, \mathcal{N}_j) \geq \Delta D_{observed}(s^*, \mathcal{N}))]}{J}.$$

We sample from the permutation distribution because it is typically computationally prohibitive to obtain $n!$ permutations of the observed data and to compute an optimal goodness-of-fit measure for each of those data sets. Therefore, as an approximation, one can chose J sufficiently large in order to achieve as many significant digits as desired for the p-value.

References

- [1] Mazumdar M, Glassman JR. Tutorial in Biostatistics – Categorizing a prognostic variable: review of methods, code for easy implementation and applications to decision-making about cancer treatments. *Statistics in Medicine* 2000; **19**:113–132.
- [2] Cox DR, Hinkley DV. *Theoretical Statistics* Chapman and Hall: London, 1974.
- [3] Kim H-J, Fay MP, Feuer EJ, Midthune DN. Permutation tests for joinpoint regression with applications to cancer rates. *Statistics in Medicine* 2000; **19**:335-351.