Permutation Methods: A Distance Function Approach, by Paul W. Mielke, Jr, Kenneth J. Berry, New York: Springer 2001, ISBN 0-387-98882-3, xv+352pp, $74.95.

Permutation methods as the authors, Mielke,Jr. and Berry, suggested is a generalization of Fisher-Pitman permutation test, which is nicely outlined in the chapter 2; "Description of MRPP".

MRPP (Multiresponse Permutation Procedures) are a class of permutation methods which is based on the following distance function:

$\delta = \sum_{i=1}^{g} C_i \xi_i$; Where C_i are group weights which added up to 1 and $\xi_i = \left(\frac{n_i (n_i - 1)}{2} \right)^{-1} \sum_{I<J} \Delta_{I,J}$ is the average distance function value for all distinct pairs of objects in the group S_i. Under null hypothesis H_0, equal probabilities are assigned to each of the $M = \frac{N!}{\prod_{i=1}^{g+1} n_i!}$ possible allocations of the N objects in Ω to the g+1 groups $(S_1, \ldots, S_{g+1})$. The δ statistic compares the within group clumping of response measurements against the model specified by random allocation under null hypotheses. This procedure extends the permutation test to wide array of multiple response cases.

The merit of this method is critically depending on how is the distance function $\Delta_{I,J}$ between two distinct objects in the groups being defined. The authors pointed out that the Euclidean distance being used in the least square sometimes is not as robust as the distance being defined as the absolute deviation, since the former is average based and the later is the median based .

Three approached are used to compute p-values for the test. These are, exact, resampling and moment approximations. In exact test, the distance is computed based on the observed data, then the data are permuted over all possible allocations of the objects in different groups, and the distance being computed. The p-value is the proportion of the arrangements which have the distance is as extreme or more extreme than the value of the test statistics. When the total objects N in the test is large, the exact approach becomes computationally prohibitive. Resampling and moment approximation approaches make computation manageable, however, the results are not always comparable.

MRPP are applied in the LAD (Least Absolute Distance) regression, randomized block experimental design, Goodness-of-fit tests, the traditional and generalized contingency tables. MRPP serves as a none parametric and versatile alternate tools to a lot traditional statistical approaches.

MRPP is a computation intensive approach and it only makes possible and practical as the computation cost is steadily trending downward.

As a whole, this is a well written book with an intent to generalize this seemly limited tool to a lot applications. For this book to become more valuable, it should also provide more in depth coverage of the comparisons of the exact test, resampling and moment approximation test to other than just to display the simulated or computed p-values.

Shin Ta Liu
Lynx Systems