Some solutions to the Behrens-Fisher problem for multivariate ecological data

M. J. Anderson [a]

[a] Department of Statistics,
University of Auckland,
Tamaki campus, Auckland,
New Zealand
mja@stat.auckland.ac.nz

Abstract: The Behrens-Fisher problem (BFP) is one of the oldest puzzles in statistics. The essence of this problem is how validly to compare means (or multivariate centroids) between two or more populations when their variances (or multivariate dispersions) differ. This is especially irksome in ecology whenever variables consist of counts of abundances of species, because any differences in means will be accompanied by differences in variances as well. Some solutions to the BFP do exist for the univariate case, but they assume variables are normally distributed, whereas species' counts tend to be long-tailed and overdispersed, with many zeros. This issue is exacerbated for multivariate ecological data (counts of many species in a community), which often also have more variables (species) than samples. This means none of the existing attempts to solve the multivariate BFP (all of which also assume normality) can be used in practice.

In this talk, I will outline, compare and contrast some potential solutions to the multivariate BFP that rely on some rather clever permutation, bootstrap or Monte Carlo (re-)sampling methods. While the permutation approach tends to be mildly liberal, the bootstrap approach, even with reasonable empirical bias-corrections, tends to be overly conservative. But will the Monte Carlo approach come to the rescue with a more exact test? And at what cost in terms of additional underlying assumptions?