1.5 The pseudo-F statistic

Once the partitioning has been done we are ready to calculate a test statistic associated with the general multivariate null hypothesis of no differences among the groups. For this, following R. A. Fisher's lead, a pseudo-*F* ratio is defined as:

$$ F = \frac{ SS _ A / \left(a - 1 \right)} { SS _ {Res} / \left(N - a \right) } \tag{1.3} $$

where (*a* – 1) are the degrees of freedom associated with the factor and (*N* – *a*) are the residual degrees of freedom. It is clear here that, as the pseudo-*F* statistic in (1.3) gets *larger*, the likelihood of the null hypothesis being true *diminishes*. Interestingly, if there is only one variable in the analysis and one has chosen to use Euclidean distance, then the resulting PERMANOVA *F* ratio is exactly the same as the original *F* statistic in traditional ANOVA[10] (Fisher (1924)). In general, however, the PERMANOVA *F* ratio should be thought of as a "pseudo" *F* statistic, because it does *not* have a known distribution under a true null hypothesis. There is only one situation for which this distribution is known and corresponds to Fisher's traditional *F* distribution, namely: (i) if the analysis is being done on a single response variable *and* (ii) the distance measure used was Euclidean distance *and* (iii) the single response variable is normally distributed. In all other cases (multiple variables, non-normal variables and/or non-Euclidean dissimilarities), all bets are off! Therefore, in general, we cannot rely on traditional tables of the *F* distribution to obtain a *P*-value for a given multivariate data set.

Some other test statistics based on resemblance measures (and using randomization or permutation methods to obtain *P*-values, see the next section) have been suggested for analysing one-way ANOVA designs (e.g., such as the average between-group similarity divided by the average within-group similarity as outlined by Good (1982) and Smith, Pontasch & Cairns (1990) , see also all of the good ideas in the book by Mielke & Berry (2001) and references therein). Unlike pseudo-*F*, however, these can be limited in that they may not necessarily yield straightforward extensions to multi-way designs.

---

[10] In fact, a nice way to familiarise oneself with the routine is to do a traditional univariate ANOVA using some other package and compare this with the outcome from the analysis of that same variable based on Euclidean distances using PERMANOVA.

Revision #5
Created 2 August 2022 11:31:41 by Arden
Updated 8 April 2024 01:59:35 by Marti