

ANOVA: Is 100 years enough?

L. R. LaMotte

LSU Health - New Orleans, United States of America

Abstract

R. A. Fisher introduced analysis of variance for balanced designs in his 1925 book *Statistics for Research Workers*. ANOVA quickly became the most widely used statistical method, particularly in agromonic disciplines.

It partitions total variation into identifiable parts that together account for all possible differences among the cell means. For unbalanced designs, though, no such consensus has been reached in the intervening 100 years, and disputatious colloquy continues to this day. A 2003 paper by Øyvind Langsrud, for example, titled “ANOVA for unbalanced data: Use Type II instead of Type III sums of squares,” has been cited more than 600 times.

The purpose of this presentation is to propose a strategy for unbalanced designs that parallels classical ANOVA for balanced designs. It is based on sums of squares that test exclusively those classical ANOVA effects that are estimable, which in some cases is none, together with other contrasts up to the degrees of freedom available for all differences among the means of the non-empty cells. Techniques are illustrated to express and identify the contrasts that are tested.

Keywords

Unbalanced designs, extended effects, empty cells.