KM 8(1982)
pag 47-49

A non-parametric test for two way ANOVA when the number of treatments is a power of 2.

N.J.D. Nagelkerke, Laboratory of Medical Physics, University of Amsterdam, Herengracht 196, 1016 BS Amsterdam, The Netherlands.

Abstract

A simple distribution-free method for a two-way mixed model analysis of variance is presented. The proposed method generalizes Wilcoxon's signed rank test and has better power than Friedman's test. The test can be applied whenever the number of levels of the fixed factor is a power of 2.

1. Introduction.

A frequently occurring problem in statistical practice is the analysis of data obtained from a design with one fixed factor at k levels, e.g. treatments, which is the factor of interest and one "nuisance" factor, usually a random factor, e.g. patients.
When the standard assumptions are satisfied, the appropriate method of analysis is the two way mixed model analysis of variance. When these assumptions are violated, the usual test procedure is the one proposed by Friedman (1937), which makes use of the intrablock rankings of the outcome of the experiment. For $k = 2$ this is equivalent to applying the sign test, which is known to have poor efficiency for most underlying distributions. For this case, the Wilcoxon signed rank test is felt to be more satisfactory.
For $k > 2$, procedures more efficient than Friedman's have been proposed by Hodges and Lehmann (1962). Their method consists of aligning the values in each block before ranking to make use of interblock information. For $k = 2$ this procedure is asymptotically equivalent to applying the signed rank test. In this paper an alternative to the aligned rank test is proposed, that consists of a straightforward combination of k-1 Wilcoxon signed rank statistics, which is applicable whenever $k = 2^r$.

2. The proposed test.

The model considered for a two-factor factorial design without interaction is:

$$y_{ij} = \mu + \alpha_j + p_i + e_{ij} \qquad j = 1,2,\dots,k; \qquad i = 1,\dots,n; \qquad k = 2^r,$$

where the α_j denote the fixed effects and the p_i the random effects. The e_{ij} are independent random error variables, for each index i having the same distribution function F_i.
The hypothesis to be tested is H_0: $\alpha_1 = \alpha_2 = \dots = \alpha_k$ against H_1 that at least one of the equalities is violated.
The k-1 dimensional space of hypotheses is spanned by k-1 parameters

$$c_i : \sum_{j=1}^{k} \gamma_{ij}\alpha_j, \quad i = 1,2,\dots,k-1$$

$$\sum_{j}^{k} \gamma_{ij} = 0; \quad \sum_{j}^{k} \gamma_{ij}\gamma_{lj} = 0; \quad i \neq l.$$

H_0 corresponds to the origin of the space:

$$c_1 = c_2 = \dots = c_{k-1} = 0.$$

Under H_0, for a special choice of $\{\gamma_{ij}\}$, to be discussed later on, the signed rank statistic V_i, defined as the sum of the products of the signs and ranks of the absolute values, obtained from n values of the i-th contrast $\sum_j \gamma_{ij}\gamma_{lj}$ is

asymptotically normally distributed with $E(V_i) = 0$ and $Var(V_i) = n(n+1)(2n+1)/6$. Besides, the statistics V_i and V_j are independent, which is argued at the end of section 4.
As a result we can use

$$T = \sum_{j}^{k-1} V_j^{*2}, \qquad (1)$$

where $V_j^* = \frac{|V_j|(-1)}{n(n+1)(2n+1)/6}$ as an asymptotically χ^2_{k-1} variate and reject the null hypothesis for large values of T. The -1 in the expression for V_j^* is a continuity correction which presumably improves the χ^2 approximation. For samples not too small (say, $n > 15$) the χ_1^2 is known to provide a good approximation to the distribution of V_j^{*2}. This result readily extends to a sum of such variates. When there are ties, the usual modifications to V_j^*, i.e. to work conditional upon the observed ties, can be applied.

3. Asymptotic efficiency; a heuristic derivation.

A widely employed criterion to compare the performance of tests is the asymptotic relative efficiency (a.r.e.) or Pitman efficiency (Cox and Hinkley, 1974). To establish the asymptotic efficiency of our test relative to the usual F test, when the usual ANOVA assumptions, e.g. normality hold, we consider first the a.r.e. of the signed rank test relative to the t-test. It is well known (Lehmann, 1975) that this is $3/\pi$. Of course this does not change when the statistics are squared, that is - asymptotically - the efficiency of the χ^2 distribution of V_j^{*2} and $F_{1,m}$ with m the degrees of freedom of the errors, are compared.
As the overall $F_{k-1,m}$ distributed test statistic in the usual ANOVA, consists of the sum of the $F_{1,m}$ distributed values of the orthogonal contrasts, which are separately $\pi/3$ times as efficient as the signed rank test, we may conclude that the a.r.e. of the proposed test compared to the parametric ANOVA is also $3/\pi$.
Of course, when the absolute normal score test (Lehmann, 1975) instead of the signed rank test is used, the a.r.e. equals 1. However, as this test is comparatively unknown and computationally not generally available, the slight gain thus obtained is not worth the extra effort.
The a.r.e. in the normal case of the proposed test, relative to Friedman's equals $(k+1)/k$ and relative to the aligned rank test equals unity (cf. Lehmann, 1975).

4. Orthogonal contrasts.

To apply Wilcoxon's signed rank test, the assumption - under H_0 - of a symmetrical distribution about zero has to be satisfied for the contrasts $\sum_j \gamma_{ij} y_{lj}$.

In addition the contrasts should be mutually independent. Both conditions are only met when the contrasts considered are such that $|\gamma_{ij}| = 0$ or a constant, which, without loss of generality, can be taken equal to unity.
It appears that this is only possible when $k = 2^r$ $(r = 1,2,\ldots)$, in which case there is a simple choice of orthogonal contrasts. For example, for $r = 2$, i.e. $K = 4$, this choice is, in obvious notation:

$$\begin{array}{lrrrr} c_1: & 1 & 1 & -1 & -1 \\ c_2: & 1 & -1 & 0 & 0 \\ c_3: & 0 & 0 & 1 & -1 \end{array} \qquad (2)$$

This choice is clearly not unique. A different set of contrast yields a different value of T. Simulations suggest that these differences are not very large. The situation seems to be analogous to the freedom one has in choosing weights for each rank in any rank test, e.g. normal scores, ranks as such, etc.
It remains to be demonstrated that using the contrasts described above, $\mathrm{Cov}(V_i,V_j) = 0$ $(i \neq j)$. Let $z_{i1} = \sum_m \gamma_{im} y_{1m}$, then $Ez_{i1} = 0$. The simultaneous distribution of $y_{11}, \ldots, y_{1k}$ is the same for any permutation of the indices 1 to k. Therefore, z_{i1} is distributed symmetrically about zero; $\mathrm{sign}(z_{i1}).\mathrm{sign}(z_{j1})$ is independent of $\mathrm{abs}(z_{i1})$ and $\mathrm{abs}(z_{j1})$ and $E\mathrm{sign}(z_{i1}).\mathrm{sign}(z_{j1}) = 0$.
So $\mathrm{sign}(z_{i1})$ and $\mathrm{sign}(z_{j1})$ are mutually independent. This implies that V_i and V_j are independent.

5. Example.

44 subjects on a low-salt diet were given (in random order) placebo (P), a diuretic (D), a beta blocker (B) and a combination of the latter two (C). Systolic blood pressures were recorded; yielding average values of 141.0 (P), 135.3 (D), 128.5 (B) and 128.1 (C) mm Hg respectively.
To test the hypothesis that all pills are equally (in)effective: H_0 against the hypothesis H_1 of at least one inequality, Wilcoxon's signed rank test was applied to the following contrasts:

c_1: D - P

c_2: (D+P) - (B+C)

c_3: C - B

The contrasts, c_1 and c_3, can be associated with the effect of the diuretic, whereas the contrast c_2 compares pills without a beta blocker with pills with a beta blocker. The SPSS programm NPAR TEST yielded chi-square values of 8.35, 23.14 and 0.03 for c_1, c_2 and c_3 respectively. The sum, (approximately χ^2_3 under H_0) 31.52, is clearly highly significant, thereby rejecting H_0 in favour of H_1.

References.

Cox, D.R. and Hinkley, D.V. (1974) Theoretical statistics, London: Chapman and Hall.

Friedman, M. (1937) The use of ranks to avoid the assumption of normality implicit in the analysis of variance. J. Am. Stat. Soc. 32, 675-692.

Hodges, J.L. and Lehmann, E.L. (1962) Rank methods for combination of independent experiments in analysis of variance. Ann. Math. Stat. 33, 482-497.

Lehmann, E.L. (1975) Non-parametrics: statistical methods based on ranks. San Francisco: Holden-Day.

Wilcoxon, F. (1945) Individual comparisons by ranking methods. Biometrics Bull. 1, 80-83.

Scheffé, H. (1959) The analysis of variance. New York: Wiley.