

Paired Permutation Test for Effects in Unreplicated Factorials

*Test con Permutazioni Appaiate per la Verifica di Ipotesi
sugli Effetti in Piani Fattoriali Non Replicati*

Fortunato Pesarin, Luigi Salmaso, Dario Basso, Livio Finos

Department of Statistics, University of Padova, Italy, via C. Battisti 241 Padova,
pesarin@stat.unipd.it - salmaso@stat.unipd.it - rio_basso@yahoo.it - lfinos@stat.unipd.it

Riassunto: Il test di permutazione proposto per valutare gli effetti nei piani fattoriali non replicati a due livelli è basato sulla nozione di matrice di riallineamento e sulle permutazioni appaiate. Vengono discussi vantaggi e limiti della procedura di permutazione esatta per piani fattoriali non replicati e viene presentata una nuova soluzione basata sui residui idonea per sottoporre a verifica di ipotesi i rimanenti $2^k - k - 1$ effetti.

Keywords: Factorial design, Paired permutations, Permutation test, Realignment.

1. Introduction

In this paper we present a new permutation test based on paired permutations and on residuals for unreplicated factorials. In the literature there are some tentative solutions, and for a comprehensive review we refer the reader to *Hamada & Balakrishnan* (1998). The most efficient nonparametric solution is that of *Loughin & Noble* (1997), which is exact only for the largest absolute effect and, based on residuals, approximate for other effects. Our proposal allows us to obtain an exact test for k largest effects by using the realigning procedure in a 2^k unreplicated factorial. The realigning technique and paired permutations applied to residuals allow us also to test for the remaining $2^k - k - 1$ effects. The latter test is approximate but has a good behavior in terms of power. Due to lack of space in the 5th section (Results) we shall only present a short simulation study on power behavior of exact tests for the 4 largest effects in a 2^4 factorial.

2. Lenth test

Lenth (1989) considers m contrasts k_i of a factorial design and supposes that their estimates c_i are independent realizations of $N(k_i, \tau^2)$, $i=1, \dots, m$.

He proposes a robust estimator of the contrast standard error τ , which he calls *pseudo-standard error* :

$$\text{PSE} = 1.5 \times \text{Median} \{ |c_i| < 2.5 s_0 \} |c_i| \quad \text{where} \quad s_0 = 1.5 \times \text{Median} \{ |c_i| \}.$$

$c_i < 2.5 s_0$ i

He also defines a *margin of error*:

$$\mathbf{ME} = t_{.975; d} \times \mathbf{PSE}$$

$\mathbf{ME}$ is useful to construct confidence intervals for k_i using $c_i \pm \mathbf{ME}$. With a large number of contrasts one can expect one or two estimates of inactive contrasts to exceed the ME leading to false conclusions. To account for this possibility he also defines a *simultaneous margin of error*:

$$\mathbf{SME} = t_{\gamma; d} \times \mathbf{PSE},$$

where $t_{\gamma; d}$ is the γ quantile of a t distribution with $m/3$ d.f. and $\gamma = (1 + .95^{1/m})/2$.

A contrast whose estimate extends beyond the SME is clearly active, one which does not extend beyond the ME cannot be deemed active, and one in between leads to an undecidable conclusion.

3. Loughin & Noble test

Define the *design matrix* $\mathbf{X}_N$, $N=2^k$, of a 2^k complete factorial design as:

$$\mathbf{X}_N = [\mathbf{1} \ \mathbf{D} \ \mathbf{P}_D]$$

where $\mathbf{1}$ is the column vector of the general mean μ which is composed of all +1 elements, $\mathbf{D}$ is a $N \times k$ matrix whose rows contain all combinations of two distinct levels (± 1) of the k main factors, and $\mathbf{P}_D$ is a $N \times [N - k - 1]$ matrix whose columns correspond to the t factor interactions and these columns are Kronecker products of columns corresponding to t main factors, $2 \leq t \leq k$. The usual linear model for the analysis of variance of a 2^k complete factorial design with fixed effects is $\mathbf{y} = \mathbf{X}_N \mathbf{b} + \mathbf{e}$, where $\mathbf{X}_N$ is the square design matrix, $\mathbf{y}$ is an $N \times 1$ vector of responses, $\mathbf{b}$ is an $N \times 1$ vector of parameters and $\mathbf{e}$ is an $N \times 1$ vector of i.i.d. random errors. If the information matrix $\mathbf{X}_N^T \mathbf{X}_N$ is of full rank, the unique OLS (Ordinary Least Squares) estimate of parameters $\mathbf{b}$ is given by:

$$\mathbf{b} = (\mathbf{X}_N^T \mathbf{X}_N)^{-1} \mathbf{X}_N^T \mathbf{y} = \frac{1}{2^k} \mathbf{X}_N^T \mathbf{y},$$

since all the columns of $\mathbf{X}_N$ are orthogonal; b_0 is the mean of observed responses and $b_1, \dots, b_{N-1}$ are the contrasts estimating the effects.

Let $|b_{(1)}| \geq |b_{(2)}| \geq \dots \geq |b_{(N-1)}|$ be the ordered absolute estimated effects (OAE) and let us consider the test statistic $W_i = |b_{(i)}|$. Suppose also to perform a test on the largest ordered absolute estimated effect. Under the null hypothesis that the largest effect has mean 0, hence all effects have mean 0, the relationship between elements of $\mathbf{y}$ and columns of $\mathbf{X}_N$ is determined purely by the randomization of factor levels to runs of the experiment. Hence, a reference

distribution for W_1 can be found by computing W_1 for every possible permutation of the elements of $\mathbf{y}$. Let $G^*(w_1^* | \mathbf{y})$, where w_1^* is the test resulting from $\mathbf{y}^*$ - i.e. a random permutation of $\mathbf{y}$ - be the reference distribution, then the null hypothesis will be rejected, at the α level of significance, if $P_1 = 1 - G^*(w_1^* | \mathbf{y}) < \alpha$. The Loughin and Noble's procedure for the remaining $N-2$ effects requires the residualization of $\mathbf{y}$ by the largest effect: $\mathbf{y}' = \mathbf{y} - b_1$. Hence $b_{(2)}$ is now the largest effect, then, in general, it is possible to go on with the permutation procedure to find a reference distribution for W_j and perform the test on effect $\mathbf{b}_j$ ($j = 2, \dots, N-1$), each time residualizing $\mathbf{y}$ with respect to the previous largest effects.

4. Paired permutation test

Loughin and Noble (1997) showed that by permuting the whole vector of observations, the test on the largest absolute effect (identified from the OLS estimates) is the only exact one. Hence, a specific restricted permutation strategy (Pesarin and Salmaso, 2001) is needed to obtain an exact permutation test on a larger number of effects. The hypotheses to test are: $H_{0\beta_1} : \{\beta_1 = 0\}$ against $H_{1\beta_1} : \{\beta_1 \neq 0\}$ irrespective of whether or not $H_{0\beta_2} \cup H_{0\beta_3} \cup \dots \cup H_{0\beta_{N-1}}$ are true, and so forth. Hence, within a permutation framework, we must find a separate test for each single hypothesis.

The realigning procedure allows us to determine different data sub-sets from the original observations in a 2^k factorial design to perform k permutation tests on the k largest effects by using paired permutations. We define the $[N-1] \times k$ realigning matrix $\mathbf{R}_2^k$ composed of all possible combinations of the two distinct levels (± 1) of k factors of a two-level factorial design, excluding the combination with all $+1$ elements. Note that $\mathbf{R}_2^k$ coincides with the matrix $\mathbf{D}$ when the row with all $+1$ s in $\mathbf{D}$ is deleted. The realigning procedure can be summarized by the following steps: 1) Write the design matrix in normal form; 2) Exchange rows in the design matrix such that in the second row columns corresponding to k factors have the same configuration of ± 1 's as the r -th row of the realigning matrix, $k \leq r \leq N-1$; 3) Exchange rows in the design matrix such that each pair of elements e_1, e_2 in two adjacent rows of each of the k -factors with $+1$ in the second row has the same sign, that is either $e_1 = e_2 = +1$ or $e_1 = e_2 = -1$.

If we wish to perform a paired permutation test for an effect, we should find a realigned matrix where such effect is not aligned (i.e. the column relative to this effect has -1 in the 2^{nd} row). Let $|b^r_{(1)}| \geq |b^r_{(2)}| \geq \dots \geq |b^r_{(N-1)}|$ be the OAE for the r -th realignment, $r = 1, \dots, N-1$ of the design matrix $\mathbf{X}_N$ of an unreplicated two-level factorial. For each realignment we can perform a permutation test on the largest effect by considering as a test statistic the absolute value of the corresponding OLS estimate of the effect under testing. The reference distribution for the test statistic on the largest absolute effect $|b^r_{(1)}|$, is obtained by considering the permutation distribution of $|b^{r*}_{(1)}|$, where $|b^{r*}_{(1)}|$ is the permutation value of largest absolute effect obtained from paired permutations applied to realigned observations from the r -th realignment. The paired permutation strategy is based on the following permutation structure: permutations of rows within each pair of adjacent rows in the realigned design matrix (for more details we refer the reader to Pesarin and Salmaso, 2001). By using paired permutations it is also possible to perform exact permutation testing on other different $k-1$ largest effects. Furthermore, if error

components in the linear model for responses are reasonably assumed to be i.i.d., then the step-down approximate permutation approach introduced by *Loughin and Noble (1997)* can be applied for testing the remaining $2^k - k - 1$ effects, together with the paired permutation strategy within the realignments.

5. Some results

The following simulation study applied to the unreplicated 2^4 factorial shows the behaviour of the test by computing the power on the 4 largest effects. T.E. (.) is the true effect of factor (.), α are the nominal achieved significance levels, and for each realignment, the factor whose effect is the largest one in that realignment is indicated. The permutation distribution from paired permutations is computed exactly and the number of Monte Carlo simulations is 12000. Untested effects are supposed to be negligible and we set them equal to zero.

Table 1: Power of the paired permutation test for the 2^4 factorial.

Realignment	T.E. (.)	α	.063	.0125	.188	.250
1(A)	1		.123	.223	.304	.379
2(B)	2		.516	.720	.808	.866
4(C)	3		.872	.973	.988	.995
8(D)	4		.983	.999	1	1

References

- [1] Lenth R.V. (1989) “Quick and easy analysis of unreplicated factorials”. *Technometrics* vol. 31 n.4, 469 – 473.
- [2] Loughin T.M., Noble W. (1997) “A permutation test for effects in an unreplicated factorial design” *Technometrics* 39, 180-190.
- [3] Pesarin F., Salmaso L. (2001) “Exact permutation testing for effects in unreplicated factorial designs” *Proceedings of the International Symposium on Business and Industrial Statistics*, ISI Satellite Meeting, 31-38.