

Conditional on the Observations Permutation Tests for Two-way Anova with Application on Industrial Engineering

*Test di Permutazione Condizionati alle Osservazioni per l'Anova a Due
Vie con Applicazione all'Ingegneria Industriale*

Rosa Arboretti Giancristofaro

Dipartimento di Economia, istituzioni e territorio - Università di Ferrara
Corso Ercole I d'Este, 44, 44100 Ferrara; r.arboretti@economia.unife.it

Livio Corain

Dipartimento di Scienze Statistiche - Università di Padova
Via C. Battisti 241/3, 35121 Padova; livio.corain@unipd.it

Riassunto: La metodica dell'Anova a due vie è sicuramente uno dei modelli più importanti nell'ambito della teoria degli esperimenti programmati, ne è riprova il gran numero di proposte di soluzione fornite in letteratura. Tra queste, alcune sono di tipo non parametrico ed in particolare quelle basate sui test di permutazione condizionati alle osservazioni hanno riscontrato particolare interesse. L'obiettivo di questo lavoro è di illustrare queste proposte e proporre un'applicazione nell'ambito dell'ingegneria industriale.

Keywords: Anova, conditional test, permutation test.

1. Introduction

In experimental planning, factorial experiments are of particular interest as they allow us to separately examine the effects on the response variable of interest of two or more factors in all possible level combinations. Consequently, analysis of variance (Anova) is a statistical tool used very extensively in almost all research fields, from life and human sciences to engineering. In the usual linear model for the analysis of variance, if the error components are not normally distributed, parametric analyses may not be appropriate. It has been shown that the parametric F test is not robust when applied to heavy-tailed or asymmetric distributions (Mansouri and Chang, 1995). In these situations, nonparametric solutions are to be preferred.

Since Friedman (1937), there has been a great deal of work on nonparametric testing in two-way Anova, most of which has been published in the last two decades. In the literature, a lot of nonparametric two-way Anova solutions have been proposed and in the last years, thanks to the availability of even more powerful computing equipments, the research on permutation procedure has gained vigour and strength.

Provided that the null hypothesis implying the exchangeability of data with respect to the factors' levels holds, permutations tests not only are independent of underlying likelihood model related to population distribution but they are provided with good general properties: permutations tests are always exact procedures, their conditional

rejection regions are similar or α -invariant, and permutations test are both conditionally unbiased and consistent procedures (Pesarin, 2001). The key point is on the fact that permutation tests are conditional inferential procedures in which conditioning is on observed data set which is always a set of sufficient statistics in null hypothesis for whatever underlying distribution.

Let us suppose that we are carrying out a two-factor factorial design where factor A has I levels and factor B has J levels. If we perform all possible factor level combinations we get a full factorial design with $I \times J$ treatments. For two-way Anova, the usual linear model with additive fixed effects is defined by

$$Y_{ijk} = \mu + a_i + b_j + (ab)_{ij} + \varepsilon_{ijk}, i=1, \dots, I, j=1, \dots, J, k=1, \dots, n,$$

where Y_{ijk} are the experimental responses, μ is the general mean, a_i and b_j are the main effects, $(ab)_{ij}$ are interaction effects and ε_{ijk} are exchangeable experimental error with zero mean and unknown continuous distribution P .

For the sake of identification, effects are required to satisfy the usual side-conditions:

$$\sum_i a_i = \sum_j b_j = 0; \sum_i (ab)_{ij} = 0 \forall j; \sum_j (ab)_{ij} = 0 \forall i.$$

Usually, the experimenter's major interest is testing separately for two main effects and interaction. Hence, there are three sub-null hypotheses of interest:

1. H_{0A} : $\{a_i = 0, \forall i\}$, against H_{1A} : $\{\exists i: a_i \neq 0\}$;
2. H_{0B} : $\{b_j = 0, \forall j\}$, against H_{1B} : $\{\exists j: b_j \neq 0\}$;
3. H_{0AB} : $\{(ab)_{ij} = 0, \forall i, j\}$, against H_{1AB} : $\{\exists i, j: (ab)_{ij} \neq 0\}$,

where the emphasis is on finding three separate and possibly uncorrelated tests for each one of the three sub-null hypotheses of interest. In order to attain this goal within a permutation framework, we must find the proper set of jointly sufficient statistics for all three testing sub-problem: H_{0A} irrespective of whether H_{0B} or H_{0AB} are true or not, H_{0B} irrespective of $H_{0A} \cup H_{0AB}$, H_{0AB} irrespective of $H_{0B} \cup H_{0A}$. In this framework this set is $\mathbf{Y} = [\mathbf{Y}_{11}, \mathbf{Y}_{12}, \dots, \mathbf{Y}_{IJ}]$, i.e. the vector of the observed responses partitioned into $I \times J$ blocks.

2. Short description of conditional on the observations permutations tests

Let us now shortly illustrate the two-way Anova permutation tests conditional on the observations proposed in literature, among which we can identify a subclass based on the strategy of restricted permutations from those based on more general permutation strategies. Among the restricted permutation procedures, we can distinguish the Pesarin's Constrained Synchronized Permutation (CSP) test (2001), which compare pairs of levels of the factor under hypothesis, from those inspired by the two-way Anova F test, such as Sprent (1998) and Edgington (1995). The remaining permutation tests conditional on the observations are based on more general strategies where permutations are performed without any restrictions among factors' levels, such as the F-inspired permutation test for the interaction proposed by Manly (1997). It is worth highlighting that only tests using restricted permutations are actually exact, whereas those which permute data without any particular restrictions are approximated.

Some other unconditional permutation procedures, belonging mainly to the field of multiple regression analysis, have also been proposed in the literature. All these unconditional strategies permute the residuals that are calculated by means of an estimate of regression models. The model parameters can be estimated using the least squares method (Kennedy and Cade, 1996) or the least absolute deviations method (Cade and Richards, 1995). A test based on the distance function approach is also proposed (Mielke and Berry, 2001). Since this kind of proposals can not be referred to the framework of conditional procedures, they are not coherent with the aim of our paper and clearly they do not ensure the good general properties of conditional permutation tests. For these reasons we have decided to discard them from this work. From results of a recent simulation study (Corain et al., 2005), where all proposed in the literature conditional permutation tests for two-way Anova have been compared, it has been possible to deduce that in case of normality the power of permutation tests are almost the same as the parametric F test, while in case of heavy tail errors the permutations tests are more powerful. Particularly, CSP test is the more robust in case of failure of normality assumption.

3. Application on industrial engineering

With purpose of illustration and validation of conditional permutation tests for two-way Anova let us consider their application on an example proposed by Montgomery (2004). When designing a new battery it is possible to choose the covering material among the three types A, B and C, but on the other hand it is known that temperature will also affect the battery time life. Hence, a balanced complete 3^2 design was carried out on a set of batteries, by establishing the temperature level on 15°, 70° e 125° Fahrenheit which are compatible values with the environment where battery will be employed. For each combination of material and temperature a set of 4 batteries have been tried and the 36 experimental units have been randomly assigned to treatments. The goal of the experiment is understand whether material, temperature and their interaction do affect the battery time life. The observed values of battery time life, for each of the 9 treatments, are reported on Table 1.

Table 1: *observed battery time life, for type of material and temperature level*

Material	Temperature											
	15°				70°				125°			
A	130	155	74	180	34	40	80	75	20	70	82	58
B	150	188	159	126	136	122	106	115	25	70	58	45
C	138	110	168	160	174	120	150	139	96	104	82	60

Results on effect significance for main factors and interaction from parametric F test, from exact CSP permutation tests and from approximated Manly, Edginton and Sprent permutation tests are shown in Table 2.

We observe that all considered tests agree one each other, also with the traditional parametric F test, suggesting the significant active presence both of main factors and their interaction.

Table 2: *p-values of parametric F test and permutation tests for the experiment of battery time life*

Test	Effect		
	A	B	AB
parametric F	0.0020	0.0000	0.0186
CSP	0.0286	0.0286	0.0286
Manly	0.0030	0.0010	0.0220
Edgington	0.0010	0.0010	-
Sprenst	0.0050	0.0000	-

4. Conclusions

In this work we have shortly presented the different conditional on observations permutation tests for two-way Anova proposed in the literature and, as suggested from the developed example, we can confirm that they are very good alternatives respect to the traditional parametric F test. Hence, in each experimental situation where the normality assumption is hard to justify, these nonparametric procedures can be considered a valid solution for the problems under study.

References

- Cade B. S., Richards J. D. (1996) Permutation tests for least absolute deviation regression, *Biometrics*, 52, 886-902.
- Corain L., Salmaso L., Vedovetto V. (2005) Studio comparativo sui test di permutazione condizionati alle osservazioni per l'ANOVA a due vie, accepted for the publication in *I quaderni di Statistica*.
- Edgington E. S. (1995) *Randomization tests* (3rd edn), Marcel Dekker, New York.
- Friedman M. (1937) The use of ranks to avoid the assumption of normality implicit in the ANOVA, *Journal of American Statistical Association*, 32, 675-701.
- Kennedy P. E., Cade B. S. (1996) Randomization tests for multiple regression, *Communications in Statistics. Simulation and Computation*, 25, 923-936.
- Manly B. F. J. (1997) *Randomization, Bootstrap and Monte Carlo Methods in Biology* (2nd edn), Chapman & Hall, London.
- Mielke P. W. jr., Berry K. J. (2001) *Permutation Methods: A distance Function Approach*, Springer Series in Statistics.
- Mansouri H., Chang G. H. (1995) A comparative study of some rank tests for interaction, *Computational Statistics and Data Analysis*, 9, 85-96.
- Montgomery D.C. (2004) *Design and Analysis of Experiments* (6th edn), Wiley, Chichester.
- Pesarin F. (2001) *Multivariate Permutation Tests: With Applications in Biostatistics*, Wiley, Chichester.
- Sprenst P. (1998) *Data Driven Statistical Methods*, Chapman & Hall, London.