

Overview and Main Advances in Permutation Tests for Linear Regression Models

Massimiliano Giacalone¹, Angela Alibrandi²

1. School of Economics, Management and Statistics, Bologna University, Piazza Scaravilli 2, Bologna, Italy.

2. Department of Economical, Business and Environmental Sciences and Quantitative Methods, Messina University, Via dei Verdi 75, Messina, Italy.

Received: December 12, 2014 / Accepted: January 06, 2015 / Published: February 25, 2015.

Abstract: When the population, from which the samples are extracted, is not normally distributed, or if the sample size is particularly reduced, become preferable the use of not parametric statistic test. An alternative to the normal model is the permutation or randomization model. The permutation model is nonparametric because no formal assumptions are made about the population parameters of the reference distribution, i.e., the distribution to which an obtained result is compared to determine its probability when the null hypothesis is true. Typically the reference distribution is a sampling distribution for parametric tests and a permutation distribution for many nonparametric tests. Within the regression models, it is possible to use the permutation tests, considering their ownerships of optimality, especially in the multivariate context and the normal distribution of the response variables is not guaranteed. In the literature there are numerous permutation tests applicable to the estimation of the regression models. The purpose of this study is to examine different kinds of permutation tests applied to linear models, focused our attention on the specific test statistic on which they are based. In this paper we focused our attention on permutation test of the independent variables, proposed by Oja, and other methods to effect the inference in non parametric way, in a regression model. Moreover, we show the recent advances in this context and try to compare them.

Key words: Permutation Tests, Linear Regression Models, Non Parametric Approach.

1. Permutation Tests in Regression Models

In many cases, when within the regression models the classical condition are not respected, it's possible to use the permutation tests, considering their ownerships of optimality, especially in the multivariate context (Shadrokh and d'Aubigny, 2010; Shadrokh, 2011). The evaluation of the parameters significance is an inferential procedure, based on randomization tests (if the same experimental plan justifies them) or permutation tests (if the observed samples are random, so that the analyzed samples justify the calculations) (Kempthorne and Doerfler, 1969). In permutation tests, the reference distribution against which the statistic is tested is obtained by

randomly permuting the data under study, without reference to any statistical population. The test is valid as long as the reference distribution has been generated by a procedure related to a null hypothesis that makes sense for the problem at hand, irrespective of whether or not the data set is representative of a larger statistical population. This is the reason why the data do not have to be a random sample from some larger statistical population. The only information the permutation test provides is whether the pattern observed in the data is likely, or not, to have arisen by chance. For this reason, one may think that permutation tests are not as good or interesting as classical tests of significance because they might not allow one to infer conclusions that apply to a statistical population.

Permutation tests provide a promising approach to

Corresponding author: Massimiliano Giacalone, School of Economics, Management and Statistics, Bologna University, Italy. E-mail: massimilia.giacalone@unibo.it.

testing hypotheses in a variety of data structures. According to Good (1994), permutation tests are among the most powerful of statistical procedures available, offering robust alternatives in the face of violations of the assumptions of traditional parametric tests. Through the use of the permutation tests, we assess the null hypothesis of casualness: in fact, it suggests that, if the examined phenomenon has a certain tendency, confirmed by a model that appears as gives, it is a purely accidental effect of the observations in casual order (Hogarty and Kromrey, 2003). The alternative hypothesis of the test, however, states that the inferential model, at the base of the studied phenomenon, it is not the result of chance. Performing the randomization test is a way to determine if the null hypothesis is reasonable in this situation. We proceed choosing an useful S statistic test to measure the entity of the phenomenon of interest in relationship to the observed data and we compare the observed s statistic test value of S and the distribution of S , obtained casually rearranging the data. The test founds on the following principle: if the null hypothesis were true, then all the possible arrangements of the observations would have equal probability to verify, that is the order of the observed data is one of the possible equally probable arrangements and s appears as one of the possible values of the randomization distribution of S . If s is significant value, that is the null hypothesis has not confirmed by test, then, for implication, the alternative hypothesis is considered more reasonable. The significance level of s is, therefore, the percentage of the values that are great or equal to s in the randomization distribution. It represents a measure of evidence strength against the null hypothesis.

The purpose of this study is to examine different kinds of permutation tests applied to linear models, focused our attention on the specific test statistic on which they are based. In particular we focused our attention on permutation test of the independent variables, proposed by Oja, and other methods to

effect the inference in non parametric way. Moreover, we show the main advances in this context and try to compare them. The paper is so structured:

- in section 2 we describe the Oja permutation test and its reformulations proposed by other authors;
- in section 3 the Ter Braak residual permutation test of the complete model is presented with its appreciable properties;
- in section 4 three kinds of residual permutation tests of the reduced models are introduced (respectively Friedman and Lane test, Kennedy test and Manly test);
- in section 5 the permutation test of the dependent variable approach is examined, including three possible reason to justify this type of permutation approach (Manly);
- in section 6 the exact restricted permutation tests for partial regression models, proposed by Brown and Maritz, is shortly presented.
- in section 7 we discuss the main characteristics of the examined methods and we underline some remarks related to the permutation and randomization aspects of this paper.

2. The Oja Permutation Test of the Independent Variables

In the literature, the permutation tests applicable to the estimation of regression models are numerous. Firstly, we focus our attention on Oja test. The experimental plan presented by Oja (1987) considers n subjects to which a treatment variable x is assigned in order to study their effects on a response variable Y . In addition, for each k subject, further Z explanatory variables (covariates) are considered. The non-parametric permutation tests proposed by Oja are relative to a completely permuted plane: in fact, they based on the assumption that the treatment values are randomly assigned to subjects. Therefore, the permutation distribution used to verify the significance of a relationship between X and Y , taking into account the effects of the Z covariates, is obtained

by permuting the X values to the n statistical units. In formal terms, this is a regression plan model where the results can be generalized to multiple regression. The model can be expressed as:

$$Y_i = \alpha + \beta X_i + \gamma Z_i + \varepsilon_i$$

with $i=1, \dots, n$, where α , β and γ are unknown parameters, X is the explicative variable of the plane such that $\sum_{i=1}^n X_i = 0$, Z is the explanatory covariable and $\varepsilon_1 \dots \varepsilon_n$ are independent and identically distributed random errors with null mean. The attention is focused on the β parameter; therefore the null hypothesis is expressed by $H_0: \beta = 0$ and α and γ are nuisance parameters. Let's suppose that Y , X and Z are given for all i : the X variable is considered as a realization of the random permutation x^* of x . Then, the corresponding y values, which have not been realized, are $y^* = y + \beta(x^* - x)$, from which we can easily obtain $y^* - \beta x^* = y - \beta x$. The test statistic proposed by OJA to assess the null hypothesis is:

$$T = \sum_{i < j < k} \Delta_{ijk}^y \Delta_{ijk}^{x^*} \quad (1)$$

where

$$\Delta_{ijk}^y = \begin{vmatrix} 1 & 1 & 1 \\ y_i & y_j & y_k \\ z_i & z_j & z_k \end{vmatrix} \quad (2)$$

with $i < j < k$ and similarly for $\Delta_{ijk}^{x^*}$. This statistic is not easy from the computational point of view; so, Oja proposed an alternative form of this test in order to facilitate the calculations:

$$T = \sum_i \hat{y}_i x_i^* \quad (3)$$

where $\hat{y}_i = \sum_{j < k} \Delta_{ijk}^y \delta_{jk}^z$ with

$$\Delta_{jk}^z = \begin{vmatrix} 1 & 1 \\ z_j & z_k \end{vmatrix} \quad (4)$$

with $j < k$. Before Oja and then Collins (1987) tried to

approximate the permutation distributions of test statistics proposed by them, with other distributions. In particular Oja has suggested a standardized normal approximation or, equivalently, to square the test statistic proposal and compare the result with the critical value of a χ_1^2 . Of course there is no certainty that these distributions provide adequate approximation to the corresponding distributions of permutation. Collins (1987) has proposed, in his work, a reformulation of the Oja statistics, to use easier methods of calculation, to obtain the explicit formulas of the moments of permutation and especially to have the advantage of being able to recognize a beta distribution as an approximation of the exact distribution of null permutation. These procedures presented by Oja and Collins have not had much success and development because, by permuting the independent variables, they violate the principle of ancillarity according to which the Plan should be subject to maintain the collinearity between the explanatory variables (Kennedy, 1995). To overcome this drawback Kennedy and Cade (1996) argue that the use of an asymptotically pivotal statistic (such as t or F statistics) should prevent the violation of this principle and thus the degree of collinearity of the independent variables should not affect. However, neither Collins nor Oja using pivotal test statistics. Circumventing the problem, they are limited to randomly assign the values of the explanatory variables x and z , and this leads to the variables to be "close" to the orthogonality. Moreover also Manly (1991) critiques the permutation procedure of the observations, arguing that it is not appropriate for testing the effect of one or more independent variables, given that the other explanatory variables may or may not have an effect on the dependent variable.

3. The Residual Permutation Test of the Complete Model

There are other methods to effect the inference in non parametric way, in a regression model. The residual permutation test of the complete model,

proposed by Ter Braak (1992), is based on permutation of residuals of the full model. The test statistic under permutation for this method is:

$$r_T^2 = \frac{(\sum(R_{Y|XZ}^\pi - k_\pi X)R_{Z|X})^2}{\sum(R_{Y|XZ}^\pi - k_\pi X)^2 \sum R_{Z|X}^2} \quad (1)$$

where $k_\pi = \sum R_{Y|XZ}^\pi X / \sum X^2$ and $R_{Y|XZ}^\pi$ are the permuted least-squares residuals of the full model. Permutation test proposed by Ter Braak is analogous to a bootstrap test and it consists of permuting the residual samples of a multiple regression in order to produce a distribution that can be compared with the value sample of a statistic test. In effects, this test is not a permutation test in traditional sense, because the data are transformed for getting the residues, before their exchange happens. Moreover, it is hybrid between a permutation test and a bootstrap test and its justification can be derive from both value b^* around the true b value in the bootstrap samples. Similarly the variability of F_{obs} to test $\beta = \beta_0$ are similar to the variability of F^* to test $\beta = b$. These appreciable properties are also justified because the F used statistic test is asymptotically pivotal; whatever is the errors distribution, the F asymptotic distribution doesn't depend on the parameters in the model. This test is well applicable with great samples because the variability of b around the truth β is similar to the variability of the resampled parameters that are not tested (Levin and Robbins, 1983; Gail, Tan and Piantadosi, 1988; Kennedy and Cade, 1996). These valuable properties are also justified by the fact that the test statistic F used is asymptotically pivotal, that whatever the distribution of errors, the asymptotic distribution of F does not depend on the parameters of the model.

4. The Residual Permutation Test of the Reduced Model

In the residual permutation test of the reduced model, the effects of one or more explanatory

independent X variables are tested, permuting the residuals of the regression equation of the reduced model, that contains only the other explanatory Z variables that are not tested (Levin e Robbins, 1983; Gail, Tan and Piantadosi, 1988). The permutation test is performed comparing the F sampling test statistic with the distribution that is gotten permuting residuals of Y calculated in precedence. Such procedure is appropriate only when the X variables are incorrelated with the other explanatory variables Z used for calculating the residuals. The first approximate test is that provided by Friedman and Lane (1983); this method is similar to the exact test, except that ε and α are replaced by their least-squares estimates $R_{Y|X}$ and a , respectively.

The statistic under permutation is:

$$r_F^2 = \frac{(\sum(Y_{\pi(F)} - a_{\pi(F)}X)R_{Z|X})^2}{\sum(Y_{\pi(F)} - a_{\pi(F)}X)^2 \sum R_{Z|X}^2} \quad (1)$$

where $a_{\pi(F)} = \sum Y_{\pi(F)} X / \sum X^2$.

Successively, Kennedy (1995) proposed another method of permutation, based on the general idea that the partial regression coefficient is equivalent to the simple regression coefficient of residuals. The test statistic under permutation is the correlation coefficient between $R_{Y|X}^\pi$ and $R_{Z|X}$, expressed as:

$$r_K^2 = \frac{\sum(R_{Y|X}^\pi R_{Z|X})^2}{\sum R_{Y|X}^2 \sum R_{Z|X}^2} \quad (2)$$

The values r_K^2 and r_F^2 are different under permutation although the two examined methods permute residuals $R_{Y|X}$. Also Mainly (1997) proposed a method, based on permuting observed Y values for the test of partial correlation. The test statistic under permutation is the following:

$$r_M^2 = \frac{(\sum(Y^\pi - a_{\pi(M)}X)R_{Z|X})^2}{\sum(Y^\pi - a_{\pi(M)}X)^2 \sum R_{Z|X}^2} \quad (3)$$

where $a_{\pi(M)} = \sum Y^\pi X / \sum X^2$.

5. The Permutation Test of the Dependent Variable

This test is used to verify the null hypothesis $H_0 : \beta = 0$. It can be performed comparing the values of the F test statistic with the distribution obtained permuting the Y observations, to casually assign to the sets of the observations of X and Z independent variables. Manly (1991) proposes three possible motives to justify this type of permutation approach: in first place, the n observations can be a casual sample from a population of possible observations, where the Y variable could be independent from the X and Z explanatory variables; in according to place, the values of the experimental variables X and Z can casually be assigned to n statistical units and, therefore, the values of the Y response variable can be observed (Y would not to be influenced by the X and Z variables) besides, if the variable Y and the explanatory variables X and Z are independent, all the possible joining among every value Y and every values X and Z are equally probable in relationship to a potential mechanism that generates the data.

6. The Exact Restricted Permutation Tests for Partial Regression Models

Problems involving complex relationships among variables may require permuting the residuals of some model instead of the raw data. The effect of a nominal covariable may be controlled for by restricted permutations, limited to the objects within the groups defined by the covariable. This method is discussed in detail by Manly (1991). Applications are found in Brown and Maritz (1982), that furnish an exact permutation for test of a partial regression model, within the regression plain. The proposed scheme, united to a suitable experimental plan, is used for the inference on the regression coefficient β of X , when exists another explanatory Z variable that influences the Y response. The X coefficient is

therefore a disturb parameter.

7. Final Remarks

In this paper we revisited the use of permutation test to evaluate the inference in nonparametric way, in a regression model. The permutation strategy involves a comparison of the observed test statistic (e.g., differences in class mean effect sizes or estimated regression weights) with the set of values obtained through rearranging the data. The rearrangements of the data are repeated until a distribution is obtained for all possible permutations (an exact permutation test) or for a large, random sample of the permutations (an approximate permutation test). This distribution of test statistics obtained from the permutations of the observed data provides an empirical sampling distribution within which the observed test statistic may be compared. Comparing the randomization and permutation tests with the conventional test for inference in a regression model, we can underline some aspects. First of all, the randomization and permutation tests have two important advantages: they are also valid and opportunely applicable without casual samples and they allow to select a statistic test appropriated for a particular considered situation. Nevertheless, it's not possible to generalize the conclusions of a randomization test to the whole population of interest. In fact a randomization test identify the probability that a phenomenon of interest is casual. The concept of population from which to extract samples of observations is not fundamental and this is the reason for which for the casual sampling is not required. From other hand, the generalization of results of the conventional tests to the whole population is based on the assumption, not always verifiable, that the observed samples is equivalent to a casual sample or that the data are available for the whole population of interest (but this last condition is practically unrealizable).

Oja's permutation test permutes predictors. The author presents some distribution-free tests applicable

in presence of covariates, when treatment values are randomly assigned. The formulas and calculations are, however, cumbersome, and implementation of the test relies on using a χ^2 approximation to the exact null distribution. Oja discussed the permutation distribution of test statistics when the permute- Z method was used; successively, Collins proposed fitting a Beta distribution to the permutation distribution in order to perform the test. With reference to the alternative approaches of permutation tests in regression, we examine four kind of permutation tests, in which the test statistics is differently obtained.

The permutation distribution suggested by Freedman and Lane is constructed by permuting the residuals and adding them to the predicted values to construct a new set of y variables these y values may be regressed on x to obtain another set of residuals that are unique to this permutation of the data. In this way the residuals and the squared partial correlation coefficient for this permutation are obtained. The residuals involving z and x have not changed in the permutation, because their values are constant across the set of permutations.

The permutation distribution suggested by Kennedy is also constructed by permuting the residuals, but they are not recombined with the original predicted values. Rather, these permuted residuals are entered directly in the calculation of the squared partial correlation. The only value that will change across permutations is the numerator of this formula, because each permutation will result in new pairings of the two residuals, while the sum of the squared residuals remains constant.

Manly suggested that the original observed y values may be permuted, and the regression of these permuted y values on x may be obtained, providing residuals residuals, which will be unique for each permutation of the y vector and are used to compute the partial correlation.

Ter Braak suggested a permutation distribution that

is similar to the Freedman and Lane approach, except that the residuals being permuted are obtained from regressing y on both z and x simultaneously (called the full model residuals). For a given sample of observed values of y , z and x , the observed values of y are regressed on x and z simultaneously to obtain the residuals. The permutation distribution suggested by Ter Braak is constructed by permuting these residuals, then regressing them on x to obtain another set of residuals that are unique to this permutation of the data. In all four methods, the observed value of the squared partial correlation is used as the test statistic, but the methods yield different permutation distributions against which this value is evaluated to obtain a probability statement.

Manly observes that if all slope parameters, in a regression model, are tested simultaneously against a reduced parameter model that includes only the intercept, then the reference permutation distribution for the test statistic is obtained by randomly sampling the $n!$ permutations of the dependent variable to the matrix of independent variables and calculating the test statistic for each permutation.

Concluding our overview we cite Brown and Maritz; they propose a restricted permutation in the context of multiple regression, where predictor variables take several fixed repeated values.

We have also demonstrated that the critical values for all this permutation and randomization tests converge to the same constant under sequences of contiguous alternative hypotheses. So they all have the same asymptotic power. Anderson and Legendre (1999) considering empirical studies show that Friedman and Lane (1983) method generally gives the best results (in terms of type one error or power) and the theoretical results here demonstrate why this is so. So, randomization and permutation tests represent a methodologically adequate solution in a large number of practise experimental contexts in which the samples are not random and it's necessary a statistic test appropriated for a particular condition, alternatively to

conventional tests whose assumptions are too restricted. In addition we underline that these results have special relevance for the of multivariate data (i.e multiple response variables). For univariate analysis the normal theory tests are fairly robust and in situations where this is in doubt, appropriate transformations of the raw data can usually be found. In the opposite, multivariate tests are not so robust to departures from nonnormality, and permutation and randomization tests are very used for nonparametric analysis of multivariate data particularly in the biological and ecological sciences especially for tests based on distances matrices.

References

- [1] Anderson, M. J., Robinson, J.(2001). Permutation tests for linear models. *Australian Journal of Statistics* ,75-88.
- [2] Anderson, M. J. , Legendre, P.(1999). An empirical comparison of permutation methods for test of partial regressions coefficients in a linear model. *Journal of Statistical Computation and Simulation* , 271-303.
- [3] Brown, B. M., Maritz, J. S.(1982). Distribution- free methods in regression. *Australian Journal of Statistics*.,318-333.
- [4] Collins, M. F.(1987). A permutation test for planar regression. *Australian Journal of Statistics*.,303-308.
- [5] Freedman, D., Lane, D.(1983). A nonstochastic interpretation of reported significance levels. *Journal of Business and Economic Statistics*.,292-298.
- [6] Gail, M. H., Tan, W.Y. , Piantadosi,S.(1988). Tests for no treatment effect in randomized clinical trials. *Biometrika* ,57-64.
- [7] Good, P. (1994). *Permutation tests: A practical guide to resampling methods for testing hypotheses*. Springer-Verlag, New York.
- [8] Hogarty, K. Y., Kromrey, J. D.(2003). Permutation tests for linear models in meta-analysis: Robustness and power under non-normality and variance heterogeneity. Paper presented at the annual meeting of the American Educational Research Association, Chicago (2003, April).
- [9] Kempthorne, O., Dorfler, T.E.(1969). The behavior of some significance tests under experimental randomization. *Biometrika* , 231-248.
- [10] Kennedy, P.E.(1995). Randomization tests in econometrics. *Journal of Business and Economic Statistics*., 85-94.
- [11] Kennedy, P.E, Cade, B.S.(1996). Randomization tests for multiple regression. *Communication in Statistics - Simulation and Computation*., 923-936.
- [12] Levin, B., Robbins, H.(1983). Urn models for regression analysis, with applications to employment discrimination studies. *Law and Contemporary Problems* , 247-267.
- [13] Manly, B. J.F.(1991). *Randomization and Monte Carlo methods in biology*. Chapman and Hall, London.
- [14] Oja, H.,(1987). On permutation tests in multiple regression and analysis of covariance problems. *Australian Journal of Statistics* ,91-100.
- [15] Shadrokh, A., d'Aubigny,G.(2010). An analytic comparison of permutation methods for tests of partial regression coefficients in the linear model. *Applied Mathematical Sciences*., 857-878.
- [16] Shadrokh, A.(2011). Comparison of permutation methods tests in linear multiple regression model. *International Journal of Academic Research* ,853-857.
- [17] Ter Braak, C. J. F.(1992). Permutation versus bootstrap significance tests in multiple regression and ANOVA. In *Bootstrapping and Related Techniques* (Edited by K. H. Jockel, G. Rothe and W. Sendler) New York, Springer Verlag,79-85.