



UNIMORE
UNIVERSITÀ DEGLI STUDI DI
MODENA E REGGIO EMILIA



CLADAG 2013

9th SCIENTIFIC MEETING OF THE CLASSIFICATION
AND DATA ANALYSIS GROUP OF THE
ITALIAN STATISTICAL SOCIETY

September 18 - 20, 2013

University of Modena and Reggio Emilia
San Geminiano Complex - Modena, Italy

Book of Abstracts

Editors: Tommaso Minerva, Isabella Morlini, Francesco Palumbo

CLEUP

ISBN: 9788867871179

Patronage



Dipartimento di
Economia Marco Biagi



Dipartimento di
Comunicazione ed Economia

Dipartimento di
Scienze Fisiche, Informatiche e Matematiche

Table of Contents

Tommaso Agasisti, Patrizia Falzetti <i>Socioeconomic sorting and test scores: an empirical analysis in the Italian junior secondary schools</i>	pag. 2
Dario Albarello, Vera D'Amico <i>Empirical testing of probabilistic seismic hazard models</i>	pag. 9
Federico Andreis, Pier Alda Ferrari <i>A proposal for the multidimensional extension of CUB models</i>	pag. 15
Morten Arendt Rasmussen, Evrim Acar <i>Data fusion in the framework of coupled matrix tensor factorization with common, partially common and unique factors</i>	pag. 19
Luigi Augugliaro, Angelo M. Mineo <i>Estimation of Sparse Generalized Linear Models: the dglars package</i>	pag. 20
Antonio Balzanella, Lidia Rivoli, Elvira Romano <i>A comparison between two tools for data stream summarization</i>	pag. 24
Lucio Barabesi, Giancarlo Diana, Pier Francesco Perri <i>Gini Index Estimation in Randomized Response Surveys</i>	pag. 28
Francesco Bartolucci, Federico Belotti, Franco Peracchi <i>A test for time-invariant individual effects in generalized linear models for panel data</i>	pag. 32
Erich Battistin, Carlos Lamarche, Enrico Rettore <i>Identification of the distribution of the causal effect of an intervention using a generalised factor model</i>	pag. 36
Matilde Bini, Lucio Masserini <i>Internal effectiveness of educational offer and students' satisfaction: a SEM approach</i>	pag. 37
Matilde Bini, Leopoldo Nascia, Alessandro Zeli <i>Groups heterogeneity and sectors concentration: a structural equation modeling for micro level analysis of firms</i>	pag. 41
Giuseppe Boari, Marta Nai Ruscone <i>Use of Relevant Principal Components to Define a Simplified Multivariate Test Procedure of Optimal Clustering</i>	pag. 45
Giuseppe Boari, Gabriele Cantaluppi, Angelo Zanella <i>Some Distance Proposals for Cluster Analysis in Presence of Ordinal Variables</i>	pag. 49

Laura Bocci, Donatella Vicari <i>A general model for INDCLUS with external information</i>	pag. 53
Paola Bongini, Paolo Trivellato, Mariangela Zenga <i>The financial literacy and the undergraduates</i>	pag. 57
Riccardo Bramante, Marta Nai Ruscone, Pasquale Spani <i>Credit risk measurement and ethical issue: some evidences from the italian banks</i>	pag. 61
Pierpaolo Brutti, Lucio Ceccarelli, Fulvio De Santis, Stefania Gubbiotti <i>On the Stylometric Authorship of Ovid's Double Heroides: An Ensemble Clustering Approach</i>	pag. 65
Silvia Caligaris, Fulvia Mecatti and Patrizia Farina <i>Causal Inference in Gender Discrimination in China: Nutrition, Health, Care</i>	pag. 69
Giorgio Calzolari, Antonino Di Pino <i>Self-Selection and Direct Estimation of Across-Regime Correlation Parameter</i>	pag. 73
Maria Gabriella Campolo, Antonino Di Pino, Ester Lucia Rizzi <i>Modern Vs. Traditional: A cluster-based specification of gender and familistic attitudes and their influence on the division of labour of Italian couples</i>	pag. 77
Gabriele Cantaluppi, Marco Passarotti <i>Clustering the Four Gospels in the Greek, Latin, Gothic and Old Church Slavonic Translations</i>	pag. 81
Carmela Cappelli, Francesca Di Iorio <i>Regression Trees for change point analysis: methods, applications and recent developments</i>	pag. 85
Roberto Casarin and Marco Tronzano and Domenico Sartore <i>Bayesian Stochastic Correlation Models</i>	pag. 89
Rosalia Castellano, Gennaro Punzo, Antonella Rocca <i>Evaluating the selection effect in labour markets with a low female participation</i>	pag. 93
Paola Cerchiello, Paolo Giudici <i>A statistical based H index for the evaluation of e-markets</i>	pag. 97
Annalisa Cerquetti <i>Bayesian nonparametric estimation of global disclosure risk</i>	pag. 101
Enrico Ciavolino, Roberto Savona <i>The Forecasting side of Sovereign Risk: a Generalized Cross Entropy Approach</i>	pag. 105

Nicoletta Cibella, Tiziana Tuoto, Luca Valentino <i>What data tell you that models can't say</i>	pag. 109
Roberto Colombi, Sabrina Giordano <i>Multiple Hidden Markov Models for Categorical Time Series</i>	pag. 114
Pier Luigi Conti, Daniela Marella <i>Asymptotics in survey sampling for high entropy sampling designs</i>	pag. 118
Claudio Conversano, Massimo Cannas, Francesco Mola <i>On the Use of Recursive Partitioning in Causal Inference: A Proposal</i>	pag. 122
Franca Crippa, Marcella Mazzoleni, Mariangela Zenga <i>Keeping the pace with higher education. A fuzzy states gender study</i>	pag. 128
F. Cugnata, C. Guglielmetti and S. Salini <i>CUB model to validate FACIT TS-PS measurement instrument</i>	pag. 133
Rosario D'Agata, Venera Tomaselli <i>Multilevel Approach in Meta-Analysis of Pre-Election Poll Accuracy</i>	pag. 137
Alfonso Iodice D'Enza and Angelos Markos <i>Low-dimensional tracking of association structures in categorical data</i>	pag. 141
Giulio D'Epifani <i>Self-censored Categorical Responses A device for recovering latent behaviors</i>	pag. 145
Pierpaolo D'Urso, Marta Disegna, Riccardo Massari <i>Tourism Market Segmentation with Imprecise Information</i>	pag. 150
Utkarsh J. Dang, Salvatore Ingrassia, Paul D. McNicholas and Ryan Browne <i>Cluster-weighted models for multivariate response and extensions</i>	pag. 154
Cristina Davino, Domenico Vistocco <i>Unsupervised Classification through Quantile Regression</i>	pag. 158
F. Marta L. Di Lascio, Simone Giannerini <i>A copula-based approach to discover inter-cluster dependence relationships</i>	pag. 162
José G. Dias, Sofia B. Ramos <i>Hierarchical market structure of Euro area regime dynamics</i>	pag. 166
Drago Carlo, Balzanella Antonio <i>Consensus Community Detection: a Nonmetric MDS Approach</i>	pag. 170
Fabrizio Durante, Roberta Pappadà and Nicola Torelli <i>Clustering financial time series by measures of tail dependence</i>	pag. 174

Marco Enea, Antonella Plaia <i>Influence diagnostics for generalized linear mixed models: a gradient-like statistic</i>	pag. 178
Enrico Fabrizi, Maria R. Ferrante, Carlo Trivisano <i>Joint estimation of poverty and inequality parameters in small areas</i>	pag. 182
Giorgio Fagiolo, Andrea Roventini <i>Macroeconomic Policy in DSGE and Agent-Based Models</i>	pag. 187
Salvatore Fasola, Mariangela Sciandra <i>New Flexible Probability Distributions for Ranking Data</i>	pag. 191
Maria Brigida Ferraro, Paolo Giordani <i>A new fuzzy clustering algorithm with entropy regularization</i>	pag. 195
Camilla Ferretti, Piero Ganugi, Renato Pieri <i>Mobility measures for the dairy farms in Lombardy</i>	pag. 199
Silvia Figini, Marika Vezzoli <i>Model averaging and ensemble methods for risk corporate estimation</i>	pag. 203
Luis Angel García-Escudero, Alfonso Gordaliza, Carlos Matrán, Agustín Mayo-Isar <i>New proposals for clustering based on trimming and restrictions</i>	pag. 207
Andreas Geyer-Schulz, Fabian Ball <i>Formal Diagnostics for Graph Clustering: The Role of Graph Automorphisms</i>	pag. 211
Massimiliano Giacalone, Angela Alibrandi <i>An overview on multiple regression models based on permutation tests</i>	pag. 215
Francesca Giambona, Mariano Porcu <i>The determinants of Italian students' reading scores: a Quantile Regression analysis</i>	pag. 219
Paolo Giordani, Henk A.L. Kiers, Maria Antonietta Del Ferraro <i>The R Package ThreeWay</i>	pag. 223
Giuseppe Giordano, Ilaria Primerano <i>Co-occurrence Network from Semantic Differential Data</i>	pag. 227
Paolo Giudici <i>Financial risk data analysis</i>	pag. 231
Silvia Golia, Anna Simonetto <i>A Comparison between SEM and Rasch model: the polytomous case</i>	pag. 237
Anna Gottard <i>Some considerations on VCUB models</i>	pag. 241

Francesca Greselin, Salvatore Ingrassia <i>Data driven EM constraints for mixtures of factor analyzers</i>	pag. 245
Leonardo Grilli, Carla Rampichini, Roberta Varriale <i>Predicting students' academic performance: a challenging issue in statistical modelling</i>	pag. 249
Luigi Grossi, Fany Nan <i>Robust estimation of regime switching models</i>	pag. 255
Kristian Hovde Liland <i>Variable selection in sequential multi-block analysis</i>	pag. 259
Maria Iannario <i>Robustness issues for a class of models for ordinal data</i>	pag. 260
Maria Iannario, Domenico Piccolo <i>A class of ordinal data models in R</i>	pag. 264
Salvatore Ingrassia, Antonio Punzo <i>Parsimony in Mixtures with Random Covariates</i>	pag. 268
Hiroshi Inoue <i>International Relations Based on the Voting Behavior in General Assembly</i>	pag. 272
Carmela Iorio, Massimo Aria, Antonio D'Ambrosio <i>Visual model representation and selection for classification and regression trees</i>	pag. 276
Monia Lupparelli, Luca La Rocca, Alberto Roverato <i>Log-Mean Linear Parameterizations for Smooth Independence Models</i>	pag. 284
Marica Manisera, Paola Zuccolotto <i>Nonlinear CUB models</i>	pag. 288
Marica Manisera, Marika Vezzoli <i>Finding number of groups using a penalized internal cluster quality index</i>	pag. 292
Daniela Marella, Paola Vicard <i>Object-Oriented Bayesian Network to deal with measurement error in household surveys</i>	pag. 296
Angelos Markos, Alfonso Iodice D'Enza, Michel Van de Velden <i>Beyond tandem analysis: joint dimension reduction and clustering in R</i>	pag. 300
F. Martella and M. Alfò <i>A biclustering approach for discrete outcomes</i>	pag. 304

Mariagiulia Matteucci, Stefania Mignani, Roberto Ricci <i>A Multidimensional IRT approach to analyze learning achievement of Italian students</i>	pag. 309
Sabina Mazza <i>Extending the Forward Search to the Combination of Multiple Classifiers: A Proposal</i>	pag. 314
Fulvia Mecatti, M. Giovanna Ranalli <i>Plug-in Bootstrap for Sample Survey Data</i>	pag. 318
Alessandra Menafoglio, Matilde Dalla Rosa and Piercesare Secchi <i>A BLU Predictor for Spatially Dependent Functional Data of a Hilbert Space</i>	pag. 322
Maria Adele Milioli, Lara Berzieri, Sergio Zani <i>Comparing fuzzy and multidimensional methods to evaluate well-being at regional level</i>	pag. 326
Michelangelo Misuraca, Maria Spano <i>Comparing text clustering algorithms from a multivariate perspective</i>	pag. 331
Cristina Mollica, Luca Tardella <i>Mixture models for ranked data classification</i>	pag. 335
Isabella Morlini, Stefano Orlandini <i>Cluster analysis of three-way atmospheric data</i>	pag. 339
Roberto Nardecchia, Roberto Sanzo, Margherita Velucchi, Alessandro Zeli <i>Productivity transition probabilities: A microlevel data analysis for Italian manufacturing sectors (1998-2007)</i>	pag. 345
Andrea Neri, Giuseppe Ilardi <i>Interviewers, co-operation and data accuracy: is there a link?</i>	pag. 349
Akinori Okada, Satoru Yokoyama <i>Nonhierarchical Asymmetric Cluster Analysis</i>	pag. 353
Marco Perone Pacifico <i>SuRF: Subspace Ridge Finder</i>	pag. 357
Andrea Pagano, Francesca Torti, Jessica Cariboni, Domenico Perrotta <i>Robust clustering of EU banking data</i>	pag. 361
Giuseppe Pandolfo, Giovanni C. Porzio <i>On depth functions for directional data</i>	pag. 365
Andrea Pastore, Stefano F. Tonellato <i>A generalised Silhouette-width measure</i>	pag. 369

Fulvia Pennoni, Giorgio Vittadini <i>Hospital efficiency under two competing panel data models</i>	pag. 373
Alessia Pini, Simone Vantini <i>The Interval-Wise Control of the Family-Wise Error Rate for Testing Functional Data</i>	pag. 377
Mariano Porcu, Isabella Sulis <i>Detecting differences between primary schools in mathematics and reading achievement by using schools added-value measures of performance</i>	pag. 381
Antonio Punzo, Paul D. McNicholas, Katherine Morris, Ryan P. Browne <i>Outlier Detection via Contaminated Mixture Distributions</i>	pag. 387
Emanuela Raffinetti, Pier Alda Ferrari <i>New perspectives for the RDI index in social research fields</i>	pag. 392
Monia Ranalli, Roberto Rocci <i>Mixture models for ordinal data: a pairwise likelihood approach</i>	pag. 396
Marco Riani, Andrea Cerioli, Gianluca Morelli <i>Issues in robust clustering</i>	pag. 400
Stéphane Robin <i>Deciphering and modeling heterogeneity in interaction networks</i>	pag. 404
Rosaria Romano, Francesco Palumbo <i>Partial Possibilistic Regression Path Modeling</i>	pag. 409
Renata Rotondi <i>Classification of composite seismogenic sources through probabilistic score indices</i>	pag. 413
Gabriella Schoier <i>On Wild Bootstrap and M Unit Root Test</i>	pag. 417
Luca Scrucca <i>On the implementation of a parallel algorithm for variable selection in model-based clustering</i>	pag. 421
Paolo Sestito <i>The Role of Learning Measurement in the Governance of an Education System: an Overview of the Issues</i>	pag. 425
John Shawe-Taylor, Blaz Zlicar <i>Novelty Detection with Support Vector Machines</i>	pag. 430
Nadia Solaro <i>Multidimensional scaling with incomplete distance matrices: an insight into the problem</i>	pag. 431

Luigi Spezia, Cecilia Pinto <i>Markov switching models for high-frequency time series: flapper skate's depth profile as a case study</i>	pag. 435
Ralf Steeking, Klaus B. Schebesch <i>Data Privacy in Credit Scoring: Evaluating SVM Approaches Based on Microaggregated Data</i>	pag. 439
Isabella Sulis, Francesca Giambona, Nicola Tedesco <i>Analyzing university students' careers using Multi-State Models</i>	pag. 443
Luca Tardella, Danilo Alunni Fegatelli <i>BBRecap for Bayesian Behavioural Capture-Recapture Modeling</i>	pag. 447
Cristina Tortora, Paul D. McNicholas, Ryan P. Browne <i>Mixtures of generalized hyperbolic factor analyzers</i>	pag. 451
Giovanni Trovato <i>Testing for endogeneity and country heterogeneity</i>	pag. 455
Joaquin Vanschoren and Mikio L. Braun, Cheng Soon Ong <i>Open science in machine learning</i>	pag. 461
Valerio Veglio <i>Logistic Regression and Decision Tree: Performance Comparisons in Estimating Customers' Risk of Churn</i>	pag. 465
Maurizio Vichi <i>Robust Two-mode clustering</i>	pag. 469
Vincenzina Vitale <i>Hierarchical Graphical Models and Item Response Theory</i>	pag. 470
Sara Viviani <i>Extending the JM library</i>	pag. 474
Adalbert F.X. Wilhelm <i>Visualisations of Classification Tree Models: An Evaluative Comparison</i>	pag. 478

An overview on multiple regression models based on permutation tests

Massimiliano Giacalone¹ and Angela Alibrandi²

Abstract

When the population, from which the samples are extracted, is not normally distributed, or if the sample size is particularly reduced, it becomes preferable to use nonparametric statistical tests. Within the regression models, it is possible to use permutation tests, considering their ownerships of optimality, when the normal distribution of the response variables is not guaranteed especially in the multivariate context. In the literature there are numerous permutation tests applicable to the estimation of regression models. In this paper we focused our attention on the permutation tests of the independent variables, proposed by Oja, and other methods of nonparametric inference, in the regression models context.

1 Permutation test in regression models

In many cases, when the classical conditions of regression models are not respected, it's possible to use the permutation tests, considering their ownerships of optimality, especially in the multivariate context. 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). Through the use of 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. We proceed choosing a useful S statistic test to

¹ Massimiliano Giacalone, lecturer with annual appointment, Bologna University, email: massimilia.giacalone@unibo.it

² Angela Alibrandi, Department of Economical, Business, Environmental Sciences and Quantitative Methods, Messina University, email: aalibrandi@unime.it

measure the entity of the phenomenon of interest in relation to the observed data and we compare the observed s statistic test value of S and the distribution of S , obtained by casually rearranging the data. The test is based 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 a significant value, then the null hypothesis is rejected, then, for implication, the alternative hypothesis is considered more reasonable. The significance level of s is, 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.

2 The Oja permutation test of the independent variables

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 are based on the assumption that the treatment values are randomly assigned to the 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:

$$Y_i = \alpha + \beta X_i + \gamma Z_i + \varepsilon_i \quad (1)$$

with $i=1, \dots, n$, where α , β and γ are unknown parameters, X is the explicative variable of the plane such that $\sum_i^n X_i = 0$, z is the explanatory covariable and $\varepsilon_1, \dots, \varepsilon_n$ are independent and identically distributed random errors with zero mean. The attention is focused on the β parameter; therefore the null hypothesis is expressed by $H_0: \beta=0$; 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 $y^* = \gamma + \beta(x^* - x)$, from which we can easily obtain $y^* - \beta x^* = \gamma - \beta x$. The test statistic proposed by OJA to assess the null hypothesis is:

$$T = \sum_{i < j < k} \Delta_{ijk}^y \Delta_{ijk}^x \quad \text{where} \quad \Delta_{ijk}^y = \begin{vmatrix} 1 & 1 & 1 \\ y_i & y_j & y_k \\ z_i & z_j & z_k \end{vmatrix} \quad (2)$$

with $i < j < k$ and similarly for Δ_{ijk}^x .

This statistic is not easily calculable; so, Oja proposed an alternative form of this test in order to facilitate the calculations:

$$T = \sum_i \bar{y}_i x_i^* \quad \text{where} \quad \bar{y}_i = \sum_{j < k} \Delta_{ijk}^y \delta_{jk}^x \quad \text{with} \quad \Delta_{jk}^x = \begin{vmatrix} 1 & 1 \\ z_j & z_k \end{vmatrix} \quad (3)$$

where $j < k$. Collins (1987) tried to approximate the permutation distributions proposed by Oja, 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^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, using 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).

3 Some alternative approaches of permutation tests in regression

Other methods of nonparametric inference, in a regression model are:

- **The residual permutation test of the complete model**, proposed by Ter Braak (1992), that 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 statistical test. In effects, this test is not a permutation test in the traditional sense, because the data are transformed to get the residues, before their exchange happens. Moreover, it is hybrid between a permutation test and a bootstrap test and its justification can be derived 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 the distribution of errors, 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 that are not tested (Levin and Robbins, 1983; Gail, Tan and Piantadosi, 1988; Kennedy and Cade, 1996).
- **The permutation test of the dependent variable**, used to verify the null hypothesis $H_0: \beta = 0$. It can be performed by comparing the values of the F test statistic with the distribution obtained by 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 relation to a potential mechanism that generates the data;
- **The exact restricted permutation tests for partial regression models**, fundamentally developed by Brown and Maritz (1982), furnishes an exact permutation test for 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.

4 Final remarks

In this paper we revisited the use of permutation tests to evaluate non parametric inference in a regression model. 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 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 identifies 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 the casual sampling is not required. In the 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 are equivalent to a casual sample or that the data are available for the whole population of interest (but this last condition is practically unrealizable). So, randomization and permutation tests represent a methodologically adequate solution in a large number of practical experimental contexts in which the samples are not random. These methods seem to be appropriate for particular conditions, alternatively to conventional tests whose assumptions are too restricted.

References

1. Brown, B. M., Maritz, J. S.: Distribution- free methods in regression. In: Australian Journal of Statistics, 24, pp. 318-333 (1982).
2. Collins, M. F.: A permutation test for planar regression. In: Australian Journal of Statistics, 29, pp. 303-308 (1987).
3. Gail, M. H., Tan, W.Y. , Piantadosi, S.: Tests for no treatment effect in randomized clinical trials. In : Biometrika, 75, pp. 57-64 (1988).
4. Kempthorne, O., Dorfner, T.E.: The behavior of some significance tests under experimental randomization. In: Biometrika, 56, pp. 231-248 (1969).
5. Kennedy, P.E.: Randomization tests in econometrics. In: Journal of Business and Economic Statistics, 13, pp. 85-94 (1995).
6. Kennedy, P.E, Cade, B.S.: Randomization tests for multiple regression. In: Communication in Statistics – Simulation and Computation, 25, pp. 923-936 (1996).
7. Levin, B., Robbins, H.: Urn models for regression analysis, with applications to employment discrimination studies. In: Law and Contemporary Problems, 46, pp. 247-267 (1983).
8. Mainly, B. J. F.: Randomization and Monte Carlo methods in biology. Chapman and Hall. London (1991).
9. Oja, H.: On permutation tests in multiple regression and analysis of covariance problems. In: Australian Journal of Statistics, 29, pp.91-100 (1987).
10. Ter Braak, C. J. F.: Permutation versus bootstrap significance tests in multiple regression and ANOVA. In: Bootstrapping and Related Techniques, (K. H. Jockel, G. Rothe and W. Sendler, Eds.) New York, Springer Verlag, pp. 79-85 (1992).