Permutation Testing for Alternative Nonlinear Models with Application to Aging Curves of Refrigerated Vehicles

Rosa Arboretti Giancristofaro[1], Livio Corain[2], Samuela Franceschini[2], Andrey Pepelyshev[3], and Stefano Rossi[4]

[1] Department of Mathematics, University of Ferrara; Via Machiavelli, 35 - 44100 Ferrara, Italy, *rbrrso@unife.it*
[2] Department of Management and Engineering, University of Padova; Stradella S. Nicola, 3 - 3600 Vicenza, Italy, *livio.corain@unipd.it*
[3] Department of Stochastic Simulation, St. Petersburg State University; Bibliotechnaya sq.2, St.Petersburg - 198904, Russia
[4] Italian National Research Council, Construction Technologies Institute; Corso Stati Uniti, 4 - 35127 Padova, Italy

Abstract. Testing of model fitting for alternative nonlinear model comparisons within the parametric approach is traditionally a difficult topic due to the complexity of studying the null distribution of test statistics. The nonparametric permutation approach is a flexible method, suitable to be implemented for nonlinear models. In this paper, we introduce a novel algorithm within the nonparametric permutation framework able to perform proper inference on parameters of any specified nonlinear model. The algorithm, although general in its kind, offers a well-rounded approach to make inference via permutation test. Finally, we show the usefulness of the proposed method by applying it for making inference on parameters of a nonlinear aging curve for refrigerated vehicles.

Keywords: aging curve, nonlinear models, permutation tests

1 Nonparametric permutation inference on linear models

Permutation tests are conditional inferential procedures where conditioning is performed with respect to the sub-space associated with the set of sufficient statistics under the null hypothesis for all nuisance entities, including the underlying, known or unknown, distribution. For details, see Edgington (1995) and Pesarin (2001). The observed dataset is always a set of sufficient statistics under the null hypothesis for whatever underlying distribution. Therefore, permutation tests can be viewed as nonparametric inferential procedures, conditioned to the space generated by all possible data assignments. Provided that the null hypothesis implies the exchangeability of data, in the framework of permutation tests, the reference distribution of a relevant test statistic is then constructed by calculating its value for all possible permutations (re-orderings) of the observations. Thus a p-value can be computed as the proportion of the permutation values of the statistic that are equal to or greater than the observed value. For a more detailed introduction on permutation tests, we refer the reader to Pesarin (2001).

With regard to inference on linear models, permutation tests applied to multiple regression analysis have been proposed in the literature by Cade and Richards (1995) and by Kennedy and Cade (1996). These authors suggested to permute the residuals which are calculated with respect to estimated regression models. The model parameters can be estimated using the least squares method (ter Braak, 1992; Kennedy and Cade,1996) or the least absolute deviations method (Cade and Richards, 1995, Mielke and Berry, 2001). Cade and Richards (1996) proposed a permutation test for hypothesis testing on LAD (Least Absolute Deviation) regression models, based on permutation of the observed data. The test statistic was drawn from the F test, used in the least squares regression to evaluate the goodness of the estimated models. Kennedy and Cade (1995) employed the permutation test in the comparison of nested models evaluated using the least squares method. They showed that this permutation strategy is valid only when the effects which are not under testing are null. Stapel and ter Braak (1994), however, showed that the method is valid when estimating the largest possible effect, since the other effects influence only marginally such estimation.

2 A general algorithm for permutation inference in nonlinear models

In this section, we present a novel general residual-based algorithm for inference on a single parameter or on a subset of parameters within a nonparametric permutation approach. The proposed technique is suitable for both linear and nonlinear models.

Let us consider any specified nonlinear model

$$Y_i = f(X_i; \beta) + \varepsilon_i, \quad i = 1, \dots, n$$

where Y_i is the response variable, $f(\bullet; \beta)$ is the nonlinear link function, X_i is the vector of explicative variables and ε_i are exchangeable random errors with zero mean and unknown continuous distribution P, $i = 1, \dots, n$. Let us suppose an appropriate method is available to calculate $\hat{\beta}$, i.e. the estimate of the parameter vector β. The null hypothesis of interest is

$$\boldsymbol{H_0} : \tilde{\boldsymbol{\beta}} = \mathbf{0} \ \text{ vs. } \ \boldsymbol{H_1} : \tilde{\boldsymbol{\beta}} \neq \mathbf{0}$$

where $\tilde{\beta}$ is a single parameter or a subset of parameters from β.

The proposed algorithm is defined by the following steps:

(i) estimate the parameter vector β from two models: the first estimate $\hat{\beta}_0$ related to the first model M_0, i.e. under H_0 (with the constraint $\tilde{\beta} = 0$), and the second estimate $\hat{\beta}_1$ related to the second model M_1, i.e. under H_1 (without the constraint $\tilde{\beta} = 0$);

(ii) calculate two vectors of estimated response values: the first $\widehat{Y}_0 = f(X, \hat{\beta}_0)$ (from the model M_0) and the second $\widehat{Y}_1 = f(X, \hat{\beta}_1)$ (from the model M_1);

(iii) calculate two vectors of residuals: the first $\boldsymbol{R_0} = \boldsymbol{Y} - \widehat{\boldsymbol{Y}}_{\boldsymbol{0}}$ (under H_0) and the second $\boldsymbol{R_1} = \boldsymbol{Y} - \widehat{\boldsymbol{Y}}_{\boldsymbol{1}}$ (under H_1);

(iv) calculate S_0, that is the observed value of an appropriate statistic $S(R_0, R_1)$, based on R_0 and R_1. As test statistic we propose for example

$$S = (\mathrm{SSE}_0 - \mathrm{SSE}_1)/\mathrm{SSE}_1$$

where SSE_0 and SSE_1 are respectively the sum of square of residuals under H_0 and under H_1. Other residual-based statistics related to alternative model comparison (Burnham and Anderson, 2002) may be suitable;

(v) randomly permute the paired elements of R_0 and R_1 to obtain R_0^* and R_1^*. Note that the null hypothesis H_0 implies the exchangeability of random errors ε_i, $i = 1, \ldots, n$, with respect to the models M_0 and M_1. Thus, if H_0 holds, we can randomly permute residuals;

(vi) calculate $\boldsymbol{Y_0^*} = \widehat{\boldsymbol{Y}}_{\boldsymbol{0}} + \boldsymbol{R_0^*}$ and $\boldsymbol{Y_1^*} = \widehat{\boldsymbol{Y}}_{\boldsymbol{1}} + \boldsymbol{R_1^*}$;

(vii) from Y_0^* and Y_1^*, re-estimate the parameter vector $\hat{\beta}_0$ and $\hat{\beta}_1$ the two model M_0 and M_1 and the corresponding residuals;

(viii) re-calculate the value of S so that we have $S^* = S(R_0^*, R_1^*)$;

(ix) carry out $B - 1$ independent repetitions of steps (5)-(8), so that we have S_j^*, $j = 1, \ldots, B$, (i.e. a random sampling from the permutation distribution of S);

(x) the permutation estimated p-value $\hat{p}$ for H_0 vs. H_1 is given by

$$\hat{p} = \#(S_j^* \geq S_0)/B;$$

(xi) if $\hat{p} < \alpha$, the null hypothesis H_0 is rejected at the significance level α.

3 Simulation study

In this section, we evaluate the appropriateness of the proposed method through the use of a Monte Carlo simulation study. Let us to consider the well-known three-parameter logistic model:

$$Y_i = \frac{\theta}{1 + \beta_2 e^{-\beta_3 X_{1i}}} + \varepsilon_i, \quad i = 1, \ldots, n.$$

Moreover, we assume that

$$\theta = \beta_1 + \beta_4 X_{2i} + \beta_5 X_{3i}$$

where X_2 is a dummy variable representing some sort of possible fixed effects in Y while X_3 is a numerical covariate. We set the value of parameters as follows $\beta_1 = \beta_2 = \beta_3 = 1$, $\beta_4 = 0.1$, $\beta_5 = 0.07$ and we generate the value of numerical covariates and random errors as follows: X_{1i} are i.i.d from Uniform$[0, 4]$, X_{2i} are i.i.d from Bernoulli$[1/2]$, X_{3i} are i.i.d from a $N(0, 1)$ and ε_i are i.i.d from a $N(0, 0.1)$ since such distributions seem appropriate to represent real data configurations. Hence, possible hypotheses of interest are

- $H_{01} : \beta_4 = \beta_5 = 0$ vs. H_{11} : at least one β_i, $i = 4, 5$ is different from 0;
- $H_{02} : \beta_4 = 0$ vs. $H_{12} : \beta_4 \neq 0$;
- $H_{03} : \beta_5 = 0$ vs. $H_{13} : \beta_5 \neq 0$.

Note that the hypotheses of interest are related to any possible nonlinear model which is alternative to the more simple three-parameter logistic model. Hence, parameters β_4 and β_5 represent a possible different model which we would like to identify using the proposed permutation testing procedure.

Suitable MatLab routines were implemented in order to numerically estimate the parameters of the nonlinear model using the Nelder-Mead algorithm (Lagarias et.al., 1999) and to execute the proposed permutation test. These programs are available upon request by authors.

The considered simulation setting consists of 1000 Monte Carlo simulations for the generation of 100 observations ($n = 100$), where the true values are added to standard normally distributed random errors. For each one of the 1000 simulated data we separately estimated the permutation p-values (with 1000 random permutations) following the proposed algorithm for each one of the hypotheses of interest.

Simulations under H_0, which are reported here for hypothesis H_{01} in the last row of Table 1, show that the test distribution follows the achievable nominal levels. The rejection rates under each specific alternative are displayed in Table 1.

Table 1. Permutation test rejection rates under H_{11}, H_{12} and H_{13}.

Hypothesis	nominal level α				
	0.01	0.05	0.1	0.2	0.3
H_{11}(all param.)	0.642	0.945	1.000	1.000	1.000
$H_{12}(\beta_4)$	0.079	0.424	0.782	0.970	1.000
$H_{13}(\beta_5)$	0.242	0.703	0.939	1.000	1.000
H_{01} (all param.)	0.007	0.027	0.082	0.178	0.260

Note that the proposed permutation tests show in general a very good power. since they allow us to identify the true alternative for each of the three hypotheses of interest. For example, when setting the significance level α at 0.05, we reject the false null hypothesis H_{11} 94.5% times, H_{12} 42.4% times and H_{13} 70.3% times. Note that, the The procedure is more powerful in presence of numerical covariate (H_{12}) than for categorical variable (H_{13}).

4 Application to aging curve of refrigerated vehicles

The Accord Transport Perishable (ATP), established in 1970 among some European states and ratified in Italy in 1977, defines the precise structural characteristics of isothermal units at controlled temperature to be used in the transport of perishable products. Over time the insulated capacity of refrigerated transportation vehicles tends to diminish, thus allowing for an increase in the so-called overall coefficient of heat transfer K which represents the insulating capacity of the equipment and is defined as $K = W/(S \triangle T)$ where W is the thermal capacity required in a body of mean surface area S

to maintain the absolute difference $\triangle T$ between the mean inside temperature T_i and the mean outside temperature T_e, during continuous operation, when the mean outside temperature T_e is constant. The mean surface area S of the body is the geometric mean of the inside surface area S_i and the outside surface area S_e (United Nations Economic Commission for Europe, 1970).

Several factors contribute to the increase in K: some maybe regarded as structural deformation due to wear and tear while others are related to an increase of water in the polyurethane slab. A mathematical formulation of the aging curve was derived to study the effects that structural characteristics, as well as operation and maintenance practices, have on the life of refrigerated transport units. For details see Sicuro (2006). This theoretical model was derived as a combination of the physical processes involved in the heat transfer within the insulating panel and was calibrated with respect to the data available through the ATP database. This model formulation allows comparing the effective age of different type of refrigeration units independently from their manufacturing, structure or use.

Based on this model, the theoretical aging, Y_t of a refrigeration unit at its time-life t can be computed according to (1):

$$Y_t = 100\left(\frac{TC_t}{TC_0} - 1\right)\theta \tag{1}$$

where TC_0 and TC_t are the thermal conductivities respectively computed at time $t = 0$ (i.e. when the refrigeration unit is new) and at time t during the life of the unit (i.e. while the refrigeration unit is in use). These thermal conductivities are obtained using (2) and (3).

$$TC_0 = \frac{2}{3}(1 - K_1)\left(1 - \frac{K_2}{2}\right)K_3 + K_1\frac{K_4K_5 + K_4K_5}{K_4 + K_5} + \frac{16}{3}\frac{K_8K_9K_{10}}{\sqrt{3.68\frac{K_{11}}{K_{12}}}} \tag{2}$$

$$\begin{gathered} TC_t = \tfrac{2}{3}(1 - K_1)\left(1 - \tfrac{K_2}{2}\right)K_3 + K_1(K_{13}\cdot t)K_{14}+ \\ K_1(1 - K_{13}\cdot t)\left[\left(\tfrac{P_{air,t}}{P_{tot,t}}\right)K_4 + \left(\tfrac{P_{R,t}}{P_{tot,t}}\right)K_6\right] + \tfrac{16}{3}\tfrac{K_8\cdot K_9\cdot K_{10}}{\sqrt{3.68\frac{K_{11}}{K_{12}}}} \end{gathered} \tag{3}$$

where (in brackets we report the value of constants K_i, $i = 1, \ldots, 15$) K_1 (0.97) is the porosity of the polyurethane slab, which is given by the ratio of the volume occupied by the air and gas and the total volume of the slab; K_2 (0.85) the fraction of volume occupied by the solid in the selected geometrical representation of the polyurethane structure. For this particular derivation, the insulating material is represented as in line cubic cells of equal dimensions surrounded by a layer of gas. The constants K_3 (0.29 W/mK), K_4 (0.026 W/mK), K_6 (0.0078 W/mK) and K_{14} (0.6163 W/mK) are the thermal conductivity coefficients of the solid, the air, the compressed gas (R) and the water in the polyurethane slab, respectively, K_5 (1000 Pa) and K_7 (90000 Pa) are the initial partial pressures of the air and of the compressed gas, respectively, $P_{tot,t}$ is the total pressure in the polyurethane slab at time t [Pa], and is given by the sum of $P_{air,t}$ and $P_{R,t}$, the partial pressures of the air and gas at time t, respectively, $P_{air,t}$ and $P_{R,t}$ are computed according to (4) and (5),

$$P_{air,t} = K_5 + (K_{15} - K_5)(1 - P(t, A_a)), \tag{4}$$

$$P_{R,t} = K_7 + P(t, A_R), \tag{5}$$

K_{13} (3.0e-12 mc/mc-s) is the flux of condensed water in the slab; K_8 (5.6704E-8 W/mK3) is the Stephan-Boltzmann constant; K_9 (298 K) is the mean temperature of two faces of polyurethane slab; K_{10} (0.0005 m) is the mean equivalent diameter of the cells and K_{11} (35 Kg/mc) and K_{12} (1200 Kg/mc) are the densities of the foam and the solid, respectively.

In (4), K_{15} (101325 Pa) is the partial pressure of the air outside the refrigeration unit and $P(t, A)$ is a function of (6) that computes the partial pressure of a gas at time t based on the parameter A, given by (7).

$$P(t, \beta_i) = \frac{8\sqrt{2}}{1.01895\pi^2}\left(e^{-\beta_i t} - \frac{e^{-9\beta_i t}}{9}\right) \tag{6}$$

$$\beta_i = \left(\frac{\pi}{2K_{16}}\right)^2 \tilde{\beta}_i, \quad i = 2, 3 \tag{7}$$

where K_{16} (0.1 m) is the mean thickness of the polyurethane layer and β_i is a calibrated unknown parameter corresponding to the coefficient of diffusivity of the air or of the compressed gas respectively for $\tilde{\beta}_2$ and $\tilde{\beta}_3$. Equation (6) represents the mean pressure over time and across the polyurethane layer and was derived as a simplification of the diffusivity processes of the gases present in the polyurethane layer, based on Fick's Law (Smith, 2004).

From an engineering point of view, the parameter θ in (1) can be interpreted as the aging velocity of the refrigeration unit. To account for structural characteristics and specifications that might contribute to the overall aging of the isothermal unit, we can represent θ as an additional linear model such as in (8). This allows evaluating the aging results of refrigerated transportation systems that might differ by structural factors or by method of employment.

$$\begin{aligned}\theta = \beta_1 + \beta_4 X_1 + \textstyle\sum \beta_{5j} X_{2j} + \beta_6 X_3 + \beta_7 X_4 + \beta_8 X_5 + \beta_9 X_6 \\ + \beta_{10} X_7 + \beta_{11} X_8 + \beta_{12} X_9 + \beta_{13} X_{10}\end{aligned} \tag{8}$$

where β_1 is a constant, β_i, $i = 4, 6 \ldots, 13$, and β_{5j}, $j = 1, \ldots, 10$, are parameters related to several possible relevant factors potentially affecting the aging velocity. The variables under study are: X_1 : type of use, X_{2j}, $j = 1, \ldots, 10$: type of transported perishables (j = catering, dairy, deep frozen foods, dry, fish, fruit and vegetables, general perishable, ice cream, meat, poultry), X_3 : number of leafs for the second door, X_4 : number of leafs for the first door, X_5 : total perimeter doors, X_6 : presence of refrigerating unit in the vehicle, X_7 : presence of meat rails in the roof of the vehicle, X_8 : average thermal thickness, X_9 : average geometrical thickness, X_{10} : full working status.

In Table 2 the permutation p-values (with 1000 independent random permutations) are obtained for each factor, from a database of nearly 4,000 records of measurements and real aging data, available from 1998 to 2007 at the *Laboratories of Chill Techniques* (LCT) within the Italian National Research Council, Construction Technologies Institute, Padova, Italy. The LCT is one of the centers certified to measure the overall coefficient of heat transfer in transported refrigeration systems.

Table 2. Permutation p-values associated with the analyzed factors affecting the aging curve of refrigerated vehicles.

Factor	Permutation p-value
All factors	0.001
Type of use	0.098
Type of transported perishables	0.040
Number of leafs for the second door	0.010
Number of leafs for the first door	0.001
Total perimeter doors	0.050
Refrigerating unit	0.159
Meat rails in the roof	0.003
Average thermal thickness	0.038
Average geometrical thickness	0.088
Full working	0.001

Results in Table 2 may suggest several practical conclusions. In fact, the more relevant factors affecting the aging curve of refrigerated vehicles are those with a smaller permutation p-value: Number of leafs for the first door and Full working, followed by Meat rails in the roof. When these structural characteristics are introduced in the refrigerated vehicle we can expect a great changing in the related aging curve.

As illustration of partial results, presented in Figure 1 we have four estimated aging curves considering the contribution of two significant detected factors. More precisely:

- A: Meat rails in the roof = NO, Number of leafs for the first door = 0
- B: Meat rails in the roof = YES, Number of leafs for the first door = 0
- C: Meat rails in the roof = NO, Number of leafs for the first door = 2
- D: Meat rails in the roof = YES, Number of leafs for the first door = 2

5 Conclusions

In this work we have introduced a novel algorithm within the nonparametric permutation framework able to perform proper inference on parameters of any specified nonlinear model. As suggested by the simulation study and by

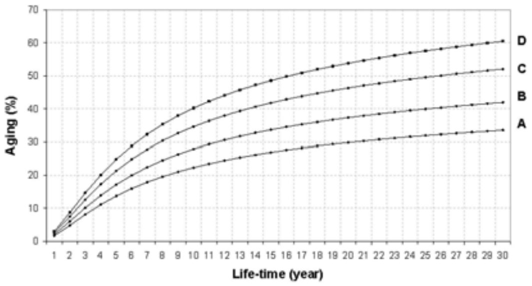

Fig. 1. Four estimated aging curves by considering the contribution of few significant detected factor.

the application to a real case study, we can state that the proposed methods offers a well-rounded approach to make inference on nonlinear models. Therefore, in each situation where the normality assumption is hard to justify or where the null distribution of test statistics is too hard to cope with, the proposed nonparametric procedure can be considered as a valid solution. We believe that in many experimental and observational studies this permutation approach may provide a significant contribution to successful research related to nonlinear and also linear models.

References

BURNHAM, K.P. and ANDERSON, D.R. (2002): *Model selection and multimodel inference: a practical information-theoretic approach (2nd edn).* Springer-Verlag, New York.

CADE, B.S. and RICHARDS, J.D. (1996): Permutation tests for least absolute deviation regression. *Biometrics, 52, 886–902.*

EDGINGTON, E.S. (1995): *Randomization tests (3rd edn).* Marcel Dekker.

LAGARIAS, J.C.; REEDS, J.A.; WRIGHT, M.H. and WRIGHT, P.E. (1999): Convergence properties of the Nelder-Mead simplex method in low dimensions. *SIAM J. Optim. 9 , no. 1, 112–147.*

KENNEDY, P.E. and CADE, B.S. (1996): Randomization tests for multiple regression. *Communications in Statistics. Simulation and Computation, 25, 923-936.*

MIELKE, P.W.JR. and BERRY, K.J. (2001): *Permutation Methods: A distance Function Approach.* Springer Series in Statistics.

PESARIN, F. (2001): *Multivariate Permutation Tests: With Applications in Biostatistics.* Wiley, Chichester.

SICURO, A. (2006): *Modello analitico dell'invecchiamento dei pannelli isolanti, mezzi isotermi e confronto con dati sperimentali.* Degree Thesis, supervisor Prof. C. Bonacina. Dipartimento di Fisica Tecnica, Università di Padova.

SMITH, W.F. (2004): *Foundations of Materials Science and Engineering* 3rd ed., McGraw-Hill.

STAPEL, M. and TER BRAAK, C.F.J. (1994): Randomization and boostrap test in factorial experiments: Does analysis follow from design? *Munster: Dutch-German Biometrics Meeting, 15-18 May, 1994.*

TER BRAAK, C.F.J. (1992): *Permutation versus bootstrap significance tests in multiple regression and ANOVA. In: K.H. Jöckel, G. Rothe and W. Sendler (Eds.): Bootstrapping and related resampling techniques*. Springer, 79–86.

UNITED NATIONS ECONOMIC COMMISSION FOR EUROPE (1970): *Agreement on the International Carriage of Perishable Foodstuffs and on the Special Equipment to be used for such Carriage (ATP)*. Transportation Division.