ECCOMAS Congress 2016
VII European Congress on Computational Methods in Applied Sciences and Engineering
M. Papadrakakis, V. Papadopoulos, G. Stefanou, V. Plevris (eds.)
Crete Island, Greece, 5–10 June 2016

THE BOOTSTRAP APPROACH TO THE STATISTICAL SIGNIFICANCE OF PARAMETERS IN RSM MODEL

Jacek Pietraszek[1], Leszek Wojnar[2]

[1] Department of Software Engineering and Applied Statistics
Faculty of Mechanical Engineering, Cracow University of Technology
37 Jana Pawla II, 31-864 Krakow, Poland
e-mail: pmpietra@mech.pk.edu.pl

[2] Chair of Computer Science and Biocybernetics
Faculty of Mechanical Engineering, Cracow University of Technology
37 Jana Pawla II, 31-864 Krakow, Poland
e-mail: leszek.wojnar@gmail.com

Keywords: Bootstrap, Statistical Significance, Response Surface Model, ANOVA, Materials Science, Biomechanics.

Abstract. *Since G.E.P. Box introduced central composite designs in early fifties of 20th century, the classic design of experiments (DoE) utilizes response surface models (RSM), however usually only in the simple form of low-degree polynomials. The typical procedure assumes the normal distribution of the noise and uses the least square method (LSQ) to identify parameters of the model with a priori assumed structure. The terms of model are repeatedly eliminated in the specific backward stepwise regression, while three indicators (the least significance of parameters, the significance of the lack of fit and the conformity of residuals with the normal distribution) are simultaneously observed to make decision to stop or to continue elimination procedure.*

Practically, in the case of small size datasets, the conformity with the normal distribution has very weak reliability and it leads to very uncertain assessment of parameters statistical significance. The bootstrap approach with simulation-based identification of parameters confidence intervals (CI) appears to be better solution than theoretically proved but only asymptotically equal t-distribution based evaluations.

The case study presented in this paper utilizes data obtained during investigation on compression vertebral fractures prediction based on computer tomography (CT) and micro-tomography (μCT) images. The significant difference in a resolution between these two class of images leaded to different prediction models. The small sample size (23 compressed and scanned vertebraes) and the high dimensionality of detected properties imposed the necessity of an alternative approach to the analysis, other than classic one derived with a requirement of the normality.

1 INTRODUCTION

The research investigation in a materials science always leads to the datasets. The large (or even huge) datasets should be analyzed to obtain summary results expressed in qualitative or quantitative forms. Such conclusions should be semi-proved by a statistical analysis.

Methods of approximation and prediction rapidly evolved in recent years. New approaches came from the artificial intelligence area: non-parametric, data-driven, stochastic. Apart from this, the classic approach of the response surface methodology (RSM) is still very useful and popular. The significance analytical improvements have appeared in the background of RSM in recent decades: statistics based on widespread strict requirements of the normal distribution have been replaced by so-called robust statistics based on the weak requirement of the continuous distribution or even on the statistics based on numerical simulation like e.g. bootstrap approach.

The recognized phenomenon may be studied in two possible ways:

- *a priori* selected treatments are imposed to the phenomenon (through a set of controlled factors) to observe changes in phenomenon behavior – such kind of study is especially specific to engineering research and usually coordinated by the design of experiments (DoE) methodology,
- the phenomenon (one or more objects) is observed passively and its properties are collected to dataset to study regressions between them however potential cause-and-effect relationship may be unknown – such kind of study is especially specific for research on live organisms, which individual properties may vary widely.

The models typically used in the first kind of research approaches are fixed-effects or response surface [1]. The second approaches usually used random-effects or response surface [1]. Additionally, Runkler [2] distinguished three aspects of models:

(a) parametric, which structure is *a priori* known and invariant through identification process,
(b) nonparametric, which structure is a priori unknown and it is constructed during identification process (other name: *data-driven*),
(c) semiparametric [3], which is sequential combination of parametric and nonparametric models.

The identification of a model is only the first step of analysis. The second one is a statistical estimation of uncertainty to determine the sensitivity of the model and statistical significance of its terms. However this definition is intuitively clear, it requires some quantitative measure and its evaluation. Typical procedure requires some important assumptions about mutual independence of observations, a normality of residuals and so on. If these conditions are met then algorithm leads simultaneously to the analysis of variance (ANOVA) [1] and tests of significance [1] – both result in the same conclusions.

The typical test of significance checks if the tested parameter may be treated as zero. The null hypothesis H_0 states that the parameter is zero against the alternative hypothesis H_1 that the parameter is not zero. If previously mentioned assumptions are met and the least squares method was used to identification then the ratio of the parameter divided by its standard deviation is tested against t-Student distribution [1]. But the problem arises if the assumptions are not met or the reliability of tests is very weak due to small size of a sample.

The proposed solution is the replacement of the classic test of parameters significance with the equivalent bootstrap-based checking of zero existence inside the parameter confidence intervals. This solution is more practical for datasets with the unknown distribution because it does not require to meet the mentioned assumptions.

Authors have been tried [4] previously to combine a bootstrap approach with an artificial neural network approximator to analyze materials science data processed by image analysis however they have with many numerical artifacts. Such approach was more successfully for processing non-parametric statistical analysis for surface layer [5] and analysis of human vertebraes scannings [6].

The bootstrap approach appears to be a promising solution for some aspects of fuzzy statistics which is developed mainly by Buckley [7, 8]. The general workflow for such implementations has been proposed by Grzegorzewski [9]. Some preliminary investigations in the field of a design of experiments (DoE) have been made by Pietraszek [10, 11].

2 COMPUTATIONAL METHODS

Two main methods were used: the response surface model (RSM) [12] and the bootstrap method [13]. The RSM model is used to construct prediction model to fit observed strength of vertebrae with the crushing force. The bootstrap method is used to evaluate confidence intervals for effects. Finally, the existence of zero inside the intervals is inspected. If any of intervals contains zero, the null hypothesis of is rejected.

2.1 RSM model

The RSM model consist of three terms [14]:

$$R_c = b_{const} + b_{BV/TV} \cdot BVTV + b_{branches} \cdot Branches + b_{Junctions} \cdot Junctions \tag{1}$$

where:
R_c – observed strength of a vertebrae
b_{const} – constant term,
$b_{BV/TV}$ – coefficient of relative density of a trabecular bone,
$b_{branches}$– coefficient of a average number of branches in trabecular bone,
$b_{junctions}$– coefficient of a average number of junctions on the branches in trabecular bone.
All coefficients were evaluated using the ordinary least squares (OLS) [1].

2.2 Bootstrap

The bootstrap approach [13] to RSM model (Eq.1) is described by the following workflow. The original dataset **D** is a source for iteratively made drawing with replacement which results in the bootstrapped dataset $\mathbf{D}_{Boot}$. Then, the parameters $\hat{b}_i$ of the model are identified from the dataset $\mathbf{D}_{Boot}$ using the least squares criterion L

$$\min L(b_i) = L(\hat{b}_i) \tag{2}$$

and obtained parameters are collected until large number of iterations will be reached. Finally, the quantiles are evaluated (typically 2.5% and 97.5%) for datasets of parameters as a range of confidence intervals and at last zero existence inside confidence intervals is checked.

3 MATERIALS AND DATASET

The data (Table 1) were obtained during investigation on the development and the optimization of diagnosis methods for the estimation risk of fractures in osteoporosis based on a three-dimensional images of trabecular bones obtained *in vivo* [14].

No	BV/TV	Branches [mm^{-3}]	Junctions [mm^{-3}]	Rc [MPa]
1	0.1166	5.8627	3.4724	8.08
2	0.2341	10.2597	5.9767	18.81
3	0.1603	11.1857	6.5363	7.31
4	0.2495	9.4670	5.5925	18.61
5	0.1704	9.3460	5.5847	12.27
6	0.1959	7.8127	4.5305	16.21
7	0.0727	3.1667	1.8487	6.50
8	0.2181	8.0648	4.7760	14.76
9	0.1156	2.4981	1.4564	10.87
10	0.1459	6.0010	3.5627	8.91
11	0.1392	8.3790	4.8488	10.52
12	0.0956	3.2275	1.9075	7.13
13	0.2139	9.8098	5.8136	14.10
14	0.1061	5.2762	3.1694	6.01
15	0.1022	4.0528	2.4041	6.36
16	0.1065	7.8878	4.5475	7.11
17	0.0837	6.7055	3.8857	6.26
18	0.1546	4.7831	2.8448	10.92
19	0.1957	9.6873	5.6839	16.32
20	0.0938	5.8117	3.4954	6.48
21	0.1410	9.2607	5.4136	9.55
22	0.1407	6.8481	4.0794	9.21
23	0.1220	6.0651	3.6250	8.57

Table 1: Raw data (source [14])

4 ANALYSIS

The model was identified from the equation (1) and the dataset (Table 1). The obtained parameters and their statistics evaluated classically are presented in Table 2.

Term	Value	SE	t	p	CI	
					-95%	+95%
b_{const}	0.34	0.91	0.39	0.701	-1.55	2.26
$b_{BV/TV}$	92.61	8.08	11.45	0.000	75.69	109.53
$b_{Branches}$	7.26	3.27	2.22	0.039	0.42	14.10
$b_{Junctions}$	-13.16	5.64	-2.33	0.031	-24.98	-1.35

Table 2: Descriptive statistics of transformed data (source [14])

The ANOVA protocol leads to the classic ANOVA table which decomposes the total variation into part assigned to grouping factor and the remain assigned to all other factors grouped under name ‘error’ which should be treated rather as ‘unexplained’ than only ‘random error’. For the mentioned data, the ANOVA table (Table 3) revealed that critical p-Value is less than significance level $\alpha = 0.05$ for all terms, what means that all terms (except constant) are statistically significant.

Effect	SS	df	MS	F	p
$b_{BV/TV}$	222.04	1	222.04	131.2	0.000
$b_{Branches}$	8.34	1	8.34	4.93	0.039
$b_{Junctions}$	9.21	1	9.21	5.44	0.031
Error	32.15	19	1.69	–	–
Total	271.74	22		–	–

Table 3: ANOVA table for transformed data (source [14])

Now, the approach was switched to the bootstrap. The number of draw iterations was set to 10.000 to easy selection of quantiles from the bootstrapped dataset. After the full bootstrap procedure, the descriptive statistics were evaluated for model parameters (Table 4).

The bounds of the confidence intervals were easy identified due to the selected number of bootstrap iterations. They were values found at positions 250 and 9750 in the sorted bootstrapped results. Similarly, the bootstrapped p-Value was evaluated as relative position of sign switching inside the sorted bootstrapped results.

Effect	Parameter mean	SE	$\text{p-Value}_{\text{bootstrapped}}$ [1] / $\text{p-Value}_{\text{theoretical}}$ [2]	-95 CI	+95 CI
b_{const}	0.29	0.82	0.682 / 0.729	-1.42	1.84
$b_{BV/TV}$	91.98	8.93	0.000 / 0.000	74.46	109.42
$b_{Branches}$	7.60	3.27	0.016 / 0.031	1.57	14.57
$b_{Junctions}$	-13.72	5.66	0.010 / 0.026	-25.79	-3.39

Table 4: Parameters of fixed-effects model for transformed data

1) bootstrapped p-Value was evaluated from relative position of sign switching inside a bootstrap table

2) theoretical p-Value was evaluated from ratio (mean/SE) and t-Student distribution at d.o.f. = 19

The analysis of confidence intervals bounds revealed that an asymmetric exists between left and right side relative to means (Table 5). The coefficients were evaluated as studentized i.e. a quotient of a deviation and standard error, where deviation was difference between confidence interval bound and respective mean:

$$cf = \frac{\pm 95CI - mean}{SE}. \tag{3}$$

Coefficient	Parameters intervals coefficients	
	-95 CI	+95 CI
b_{const}	-2.09	1.89
$b_{BV/TV}$	-1.96	1.95
$b_{Branches}$	-1.84	2.13
$b_{Junctions}$	-2.13	1.83

Table 5: Studentized coefficients of confidence intervals for the bootstrap model

5 CONCLUSIONS

- Bootstrap approach appears to be effective computational method to identify parameters of RSM effects model and their statistical properties.
- Bootstrap approach does not require to make *a priori* inconvenient assumptions.
- Bootstrap approach is very convenient to automatize in computational workflow and further statistical postprocessing.

REFERENCES

[1] O. Kempthorne, K. Hinkelmann, *Design and analysis of experiments. Vol.1. Introduction to experimental design*. John Wiley & Sons, Hoboken, NJ, USA, 2007.

[2] T.A. Runkler, *Data Analytics. Models and Algorithms for Intelligent Data Analysis.* Vieweg+Teubner Verlag | Springer Fachmedien, Wiesbaden, 2012.

[3] J.L. Horowitz, *Semiparametric Models*. W: red. J.E. Gentle, W.K. Hardle, Y. Mori, Handbook of Computational Statistics, Springer, Heidelberg, 2012, 597-617.

[4] J. Pietraszek, A. Gadek-Moszczak, The Smooth Bootstrap Approach to the Distribution of a Shape in the Ferritic Stainless Steel AISI 434L Powders. *Solid State Phenomena* **197**, 162-167, 2013.

[5] J. Pietraszek, N. Radek, K. Bartkowiak, Advanced Statistical Refinement of Surface Layer's Discretization in the Case of Electro-Spark Deposited Carbide-Ceramic Coatings Modified by a Laser Beam. *Solid State Phenomena* **197**, 198-202, 2013.

[6] A. Gadek-Moszczak, J. Pietraszek, B. Jasiewicz, S. Sikorska, L. Wojnar, The Bootstrap Approach to the Comparison of Two Methods Applied to the Evaluation of the Growth Index in the Analysis of the Digital X-ray Image of a Bone Regenerate. *New Trends in Computational Collective Intelligence* **572**, 127-136, 2015.

[7] J.J. Buckley, *Fuzzy probability and statistics*. Springer Verlag, Heidelberg, 2006.

[8] J.J. Buckley, Fuzzy statistics: hypothesis testing. *Soft Computing* **9**, 512-518, 2005.

[9] P. Grzegorzewski, *Decision support under uncertainty. Statistical methods for imprecise data.* EXIT, Warszawa, 2006.

[10] J. Pietraszek, Fuzzy Regression Compared to Classical Experimental Design in the Case of Flywheel Assembly. *Lecture Notes in Artificial Intelligence* **7267**, 310-317, 2012.

[11] J. Pietraszek, A. Gadek-Moszczak, T. Torunski, Modeling of Errors Counting System for PCB Soldered in the Wave Soldering Technology. *Advanced Materials Research* **874**, 139-143, 2014.

[12] S. Heinz, *Mathematical Modeling*. Springer Verlag, Berlin Heidelberg, 2011.

[13] J. Shao, D. Tu, *The Jackknife and Bootstrap*. Springer, New York, 1995.

[14] E. Czerwiński, A. Gądek-Moszczak, T. Konopka, Z. Latała, R. Petryniak, Z. Tabor, L. Wojnar, J. Pietraszek, *Development and optimization of diagnosis methods for estimation risk of fractures in osteoporosis based on a three-dimensional images of trabecular bones obtained in vivo. [in Polish]*. Cracow University of Technology, Kraków, 2012.

[15] J. Pietraszek, A. Goroshko, The Heuristic Approach to the Selection of Experimental Design, Model and Valid Pre-Processing Transformation of DoE Outcome. *Advanced Materials Research* **874**, 145-149, 2014.