

The Heuristic Approach to the Selection of Experimental Design, Model and Valid Pre-Processing Transformation of DoE Outcome

PIETRASZEK Jacek^{1,a}, GOROSHKO Andrii^{2,b}

¹ Institute of Applied Informatics, Cracow University of Technology
Al. Jana Pawła II 37, 31-864 Kraków, Poland

² Khmelnytsky National University, 11 Institutaska Str, 29016 Khmelnytsky, Ukraine

^amp Pietra@mech.pk.edu.pl, ^biftomm@ukr.net

Keywords: heuristic decision, design of experiments, experimental design, DoE, fractional factorials, response surface, logistics regression.

Abstract. The typical problem in the design of experiments methodology is to avoid the risk of obtaining predictions outside of ranges with a technological or physical sense. Such a situation occurs very often if a simple linear regression or another unbounded one is involved as a predictive model. The possible solution is to provide mappings of the outcome from the range of interest into the full range from negative to plus infinity before the linear regression is applied. The intermediate values obtained from the unbounded predictive model are re-transformed into the physical domain by the inverse transformation. The key issue is to decide if the mapping is required and – subsequently – what mapping is necessary. The author proposed a simple heuristic solution supporting this decision and tested such a solution in some examples described in this paper.

Introduction

The design of experiments methodology (DoE) dates back to the 1920s, when Ronald Fisher, employed in the Rothamsted Experimental Station at Harpenden (Hertfordshire, England), developed an analysis of variance (ANOVA) and introduced a Latin square experimental design as the optimal sampling scheme for ANOVA [1, 2]. His experiences and ideas have been summarized in the monograph “Design of Experiments” [3] published in 1935. In the same year 1935, F. Yates proposed [4] an idea of a split-plot experimental design. In 1946, R.L. Plackett and J.P. Burman proposed [5] an idea of special experimental designs targeted at ordering particular input factors from the most significant to the least ones. Such designs, named screening designs, differ from typical two-level factorial designs proposed by R.A. Fisher and they allow for obtaining information at greatly reduced cost. G.E.P. Box and K.B Wilson augmented [6] DoE methodology and included response surface methodology (RSM), which extended original Fisher concept onto quantitative continuous factors. Next, in 1958, H. Scheffé proposed such modification of RSM which allows for the analysis of problems where allowable factors values are restricted by a linear condition, usually in the form of a limited sum. His solution is known as a mixture design.

Simultaneously, in 1952, G. Elfving in his article [7] defined a problem describing the construction of an optimal experimental design related to a particular regression model. It is commonly recognized as the beginning of research on optimal designs. His theoretical approach was extended and redefined by J. Kiefer and J. Wolfowitz [8] in 1959 by a concept of continuous design, which was the generalization of an experimental design concept. They formally defined the term of “optimal design”. Subsequently, by using these ideas, they analyzed in detail a problem of the so-called D-optimality. In 1972 V.V. Fedorov [9] and H. Chernoff [10] published independently two monographs, summarizing the ‘state-of-the-art’ of the theoretical works on optimal designs. Simultaneously, they proposed effective methods to construct such designs. The year 1984 approached [11] an idea originated from Japan: a robust design involved and developed by G. Taguchi. A comprehensive monograph describing the robust design was published in 1989 by

M.S.Phadke [12]. In 1994, K. Hinkelmann and O. Kempthorne published two volumes [13, 14], in which they summarized the theoretical and practical ‘state-of-the-art’ related to the design of experiments. These volumes were re-edited and re-published: in 2005 the second volume, in 2008 – the first one. In 2012, K. Hinkelman published – as an editor – the third volume [15] addressing the specific issues of DoE applications in such domains which still did not exist in 1994 or had not been explored experimentally. These are research into genetic structures, spatial analysis in agronomic research, adaptable randomized clinical research, logistic regression for qualitative factors, directional statistics and designing of computer simulation research.

Outline of the Concept

It is necessary to set treatments for all input factors while implementing the selected experimental design on a particular research object. One set of factors settings (one value per factor) is named a treatment [16]. A set of all potentially available treatments constitutes a design space [17]. An experimental design is a scheme for providing systematic observations of an object response. Typically, in a classical DoE, an object response is a quantitative factor, which means that its observed value is a real number. In 1946, S.S. Stevens proposed a taxonomy [18] classifying possible observations into two groups: qualitative and quantitative. Furthermore, he divided these groups into six classes (scales/levels of measurement):

- qualitative observations: binary, nominal, ordinal,
- quantitative observations: interval, rational, absolute.

A fundamental difference between these groups is a type of observation. It is a typical number for quantitative observations, while for qualitative ones – a non-numeric label.

First heuristic. Such classification leads to a first heuristic selecting appropriate experimental design: qualitative input factors require factorial approach (full factorial, fractional factorial, Plackett-Burman, latin squares, Taguchi) contrary to quantitative ones that require response surface (Box-Wilson, Box-Behnken, mixtures).

A special approach is required for quantitative factors artificially degraded to the ordinal scale. Such a procedure is often observed if a researcher expects a multimodal response. Firstly, the factorial design based on degraded factors gives a dense mesh of predicted values. Secondly, the obtained mesh of values is an input – as substitute of measurements – for non-parametric models such as splines, RBF, neural networks etc.

Second heuristic. If Stevens taxonomy is applied to an outcome of the investigated object, then an appropriate predictive model may be selected. A quantitative outcome leads to a group of typical parametric models predicting a mean (or variance) of an outcome. A qualitative outcome leads to a group of logistic regressions:

- a binary outcome to binary logistic regression, modeling the predicted probability for a dichotomic class of an outcome,
- a nominal outcome to nominal logistic regression, modeling the predicted probability for each of a label included in the factor universum (a set of all labels for a particular factor),
- an ordinal outcome to ordinal logistic regression, modeling the predicted probability for each label included in the factor universum.

In the last case, the additional restriction appears: the classification of hypersurfaces (formally manifolds) should not be intersected, for the fact that a set of events cannot have a smaller probability than its subset.

Third heuristic. If the outcome is quantitative, then its particular Stevens class may suggest a pre-processing transformation of the outcome. This transformation is suggested to fit the unbound characteristic of a linear regression. The Stevens quantitative classes map onto transformation as follows:

- interval – no transformation; measured values may directly be modeled by linear regression,
- rational – logarithmic pre-processing transformation,

- absolute – logit pre-processing transformation; alternatively the probit transformation may be applied, however it is computationally more expensive.

It is obvious that both of the suggested transformations may be replaced by any other which has the same mapping properties if the theoretical background of the problem under consideration provides such formulas.

Case Studies

Quantitative Factors vs. Qualitative (Ordinal) Outcome. In 2009, J. Szczygieł and W. Grzeżożek conducted extensive research [19] on the subjective perception of vibrations in trams. The passengers gave their opinions by means of an electronic panel. They could select one of five levels of comfort. It was a typical qualitative outcome which may be assigned to the ordinal scale. The physically measured accelerations served as input factors. In 2012, J. Pietraszek, J. Szczygieł and W. Grzeżożek proposed [20] a predictive model based on a modified ordinal logistic regression [21-23] by introducing classification surfaces which are not simple hyper-planes. It results in classification surfaces being a patchwork of locally defined manifolds.

The object of the research was defined with 10 quantitative input factors (physical measurements acquired by sensors) and 1 qualitative factor (type of track). The outcome was qualitative (ordinal) with 5 ordered labels. The proposed model was an ordinal logistic regression with the modified, and locally defined, non-planar classification hypersurface. The data set contained 41000 data records. The obtained classifier achieved a satisfactory level of valid predictions.

It should be emphasized that the data set proved to be the 6-th dimensional object immersed in the 11-th dimensional design space. This dimensionality assessment had been achieved through the PCA analysis [24], which is a widely used method [25].

The whole experiment was close to the border between the crisp and fuzzy regression. Formally, the logistic regression predicts the probability assigned to each label of a qualitative outcome and this is the crisp one in such an approach. On the other hand, the personal assessments of each tested passenger are completely subjective and this should be treated as a typical fuzzy outcome. A similar problem appeared in a fuzzy regression analysis of a typical mechanical problem: the assembly of a flywheel made on the multi-spindle machine [26].

Quantitative Factors vs. Quantitative (Rational) Outcome. In 2013, J. Pietraszek and N. Radek [27] analyzed an average frictional coefficient measured for silicon carbides coatings made by electro-spark deposition and textured by a laser beam. The laser beam machining (LBM) was engaged in order to lower the roughness of the coating made by ESD. The elements coated by ESD were tested to determine tribological properties and they were compared before and after LBM. To achieve high reliability of the results, the test was conducted in accordance with the design of experiments (DoE) methodology and the results that were obtained were processed by a statistical analysis.

The object of the research was defined with 2 quantitative input factors (diameter of cavity and level of blackening). The outcome was quantitative (rational): an average coefficient of friction. The average was calculated over the data records obtained from a tribological tester TM-01. The value of the friction coefficient cannot be lower than 0. The heuristic led to the logarithmic pre-processing transformation. The selected experimental design was a central composite (Box-Wilson) with 9 treatments (4 axial, 4 cube and 1 center) and 5 replications in the center treatment [28].

Qualitative Factors vs. Quantitative (Rational) Outcome. In 2013, J. Pietraszek, M. Kołomycki and E. Kocylowska [29] analyzed the quantitative effects of the introduction of graphical cards with CUDA technology into the computational environment of multi-core CPUs. It is an issue of high importance due to heavy load computations required by typical engineering problems involving a finite element method (FEM) to evaluate stress and strain in constructions [30] or adjustment calculus to improve accuracy of numerical results basing on theoretical assumptions [31].

The object of the research was defined with 2 qualitative factors (existence/non-existence of CUDA card and types of multi-core CPUs). The outcome was quantitative (rational): the duration of the rendering process related to the same 3D scene. The time required for rendering cannot be shorter than 0. The heuristic led to the logarithmic pre-processing transformation. The selected experimental design was full factorial 2^2 [28].

Quantitative Factors vs. Quantitative (Absolute) Outcome. In 2013, J. Pietraszek and A. Gądek-Moszczak [32] investigated the spatial structure of sintered metal powders described by many qualitative and quantitative micro-geometrical properties. The empirical distributions of quantitative properties obtained from the advanced image analysis methods [33-37] were very irregular and, for this reason, inconvenient for further numerical simulations. The smoothing of these distributions was required for data conditioning and preprocessing, however, the use of simple regression techniques is limited due to the strict lower and upper bound of a cumulative distribution function.

A non-parametric regression model based on a multilayer perceptron was introduced because of its well-known smoothing properties. The technical application of such a model required additional provision of the confidence band or any equivalent measure of uncertainty. The highly non-linear structure of the neural network model made it impossible to use typical linear techniques to estimate the variance. Additionally, the simple confidence band estimation led to the non-physical values of the cumulative distribution function: lower than 0 or greater than 1. Due to these facts, the logit pre-processing transformation was used according to the third heuristic.

Summary

The specific difficulties related to the selection of an experimental design and the appropriate predictive model were considered. Additionally, the risk of non-physical predictions obtained from a naive implementation of the linear regression was estimated.

The package of heuristics which help to avoid this risk was proposed and the examples of their use were included.

References

- [1] R.A. Fisher, *Statistical Methods for Research Workers*, Oliver and Boyd Press, Edinburgh 1925.
- [2] R.A. Fisher, *The Arrangement of Field Experiments*, *J. Roy. Statist. Soc.* 33 (1926) 503-513.
- [3] R.A. Fisher, *The Design of Experiments*, Oliver and Boyd Press, Edinburgh 1935.
- [4] F. Yates, *Complex experiments*, *J. Roy. Statist. Soc. Suppl.*2 (1935) 181-247.
- [5] R.L. Plackett, J.P. Burman, *The design of optimum multifactorial experiments*, *Biometrika* 33 (1946) 305-325.
- [6] G.E.P. Box, K.B. Wilson, *On the Experimental Attainment of Optimum Conditions*, *J. Roy. Stat. Soc. B* 13 (1951) (1) 1-45.
- [7] G. Elfving, *Optimum Allocation in Linear Regression Theory*, *Ann. Math. Stat.* 23 (1952) (2) 255-262.
- [8] J. Kiefer, J. Wolfowitz, *Optimum Designs in Regression Problems*, *Ann. Math. Stat.* 30 (1959) (2) 271-294.
- [9] V.V. Fedorov, *Theory of Optimal Experiments*, Elsevier, 1972.
- [10] H. Chernoff, *Sequential Analysis and Optimal Design*, Harvard University, Cambridge, Massachusetts, 1972.
- [11] J. McElroy, Taguchi - Japans Secret Weapon, *Automot. Ind.* 164 (1984) (8) 18-&.
- [12] M.S. Phadke, *Quality Engineering Using Robust Design*, Prentice Hall International, Inc., London 1989.
- [13] O. Kempthorne, K. Hinkelmann, *Design and analysis of experiments. Vol.1. Introduction to experimental design*, John Wiley & Sons, Hoboken, NJ, USA 1994.

- [14] O. Kempthorne, K. Hinkelmann, Design and analysis of experiments. Vol.2. Advanced experimental design, John Wiley & Sons, Hoboken, NJ, USA 1994.
- [15] K. Hinkelmann, Editor, Design and analysis of experiments. Vol.3. Special Designs and Applications, John Wiley & Sons, Hoboken, NJ, USA (2012).
- [16] ISO 3534-3:2013, "Statistics -- Vocabulary and symbols -- Part 3: Design of experiments", (2013).
- [17] R.L. Mason, R.F. Gunst, J.L. Hess, Statistical Design and Analysis of Experiments, John Wiley & Sons, Hoboken 1989.
- [18] S.S. Stevens, On the theory of scales of measurements, 103 (1946) 677-680.
- [19] W. Grzegożek, J. Szczygieł, M. Król, An Attempt of an Employment of a Continuous Wavelet Transform for Evaluation of Temporary Comfort Disturbance, J. KONES 16 (2009) 165-172.
- [20] J. Pietraszek, W. Grzegożek, J. Szczygieł, Forecasting of Subjective Comfort in Tram Using Ordinal Logistic Regression and Manifold Learning, 19 (2012) (2) 403-410.
- [21] A. Agresti, Categorical Data Analysis, John Wiley & Sons, Hoboken 2002.
- [22] J.M. Hilbe, Logistic Regression Models, CRC Press, Boca Raton 2009.
- [23] D.W. Hosmer, S. Lemeshow, Applied Logistic Regression, John Wiley & Sons, Hoboken 2000.
- [24] I.T. Jolliffe, Principal Component Analysis, Springer, New York 2002.
- [25] E. Skrzypczak-Pietraszek, J. Pietraszek, Chemical profile and seasonal variation of phenolic acid content in bastard balm (*Melittis melissophyllum* L., Lamiaceae), J Pharmaceut Biomed 66 (2012) 154-161.
- [26] J. Pietraszek, Fuzzy Regression Compared to Classical Experimental Design in the Case of Flywheel Assembly, Lect Notes Artif Int 7267 (2012) 310-317.
- [27] J. Pietraszek, N. Radek, DoE-based model for tribology of laser textured surfaces of silicon carbide, in: T. Łodygowski, J. Rakowski, T. Garbowski and W. Sumelka, (Eds.), CMM-2013 - Computer Methods in Mechanics, 27-31.08.2013, Poznań, Poland, Poznań University of Technology (2013), MS14-21-22.
- [28] D.C. Montgomery, Design and Analysis of Experiments, John Wiley & Sons, Inc., Hoboken 2008.
- [29] J. Pietraszek, M. Kołomycki, E. Kocyłowska, The impact of CUDA technology on the efficiency of Blender renderer program, Czasopismo Techniczne 110 (2013) (1-M) 317-324.
- [30] P. Osocha, P. Duda, B. Wegłowski, Determining temperature and stress changes in thick-walled elements of steam lines, Inz Chem Procesowa 25 (2004) (4) 2249-2256.
- [31] T. Styrylska, J. Pietraszek, Numerical Modeling of Non-Steady-State Temperature-Fields with Supplementary Data, Z Angew Math Mech 72 (1992) (6) T537-T539.
- [32] J. Pietraszek, A. Gądek-Moszczak, N. Radek, The Estimation of Accuracy for the Neural Network Approximation in the Case of Sintered Metal Properties, in: A. Badica, (Ed.), Recent Developments in Computational Collective Intelligence., Studies in Computational Intelligence 513, Springer (2014), 125-134.
- [33] A. Gadek, S. Kuciel, L. Wojnar, W. Dziadur, Application of computer-aided analysis of an image for assessment of reinforced polymers structures, Polimery-W 51 (2006) (3) 206-211.
- [34] A. Gadek-Moszczak, L. Wojnar, Objective, Quantitative and Automatic X-Ray Image. Analysis of the Bone Regenerate in the Ilizarov Method, (2009) 453-458.
- [35] A. Gadek-Moszczak, S. Zmudka, Description of 3D microstructure of the composites with polypropylene (PP) matrix and Tuf particles fillers, Solid State Phenom. 197 (2013) 186-191.
- [36] A. Szczotok, R. Przeliorz, Phase transformations in CMSX-4 nickel-base superalloy, IOP Conf. Ser.-Mat. Sci. 35 (2012).
- [37] A. Szczotok, M. Sozanska, A Comparison of Grain Quantitative Evaluation Performed with Standard Method of Imaging with Light Microscopy and EBSD Analysis, Prakt Metallogr-Pr M 46 (2009) (9) 454-468.