M A T H E M A T I C A L  E C O N O M I C S

No. 13(20) 2017

# ON THE USE OF THE DESIGN OF EXPERIMENTS THEORY IN INDIVIDUAL DATA ANALYSIS

**Małgorzata Złotoś**

**Abstract.** The methods of the design of experiments are one of tools which are used in statistical quality control. The application of the design of experiments in the projection or modification of the manufacturing process leads to the improvement of the technological parameters and to reducing the overall costs of the process. Among the designs of experiments it is possible to indicate the classic and factorial designs of experiments. Nowadays the theory of the design of experiments applies in natural science and in the analysis of the spatial data. Microeconometrics is the part of econometrics which was developed in the 1950s. Microeconometric research deals with the analysis of individual data called micro-data, which are increasingly used in research in the fields of economics, management and finance. The aim of this paper is to present the possibility of using the theory of experimental design in the analysis of the individual data. The proposed method will be presented for a selected set of individual data.

**Keywords:** design of experiments, individual data, response surf.

**JEL Classification:** C19, C99.

**DOI:** 10.15611/me.2017.13.07.

## 1. Introduction

The design of experiments is the one of the most important tools of statistical quality control, but the first use of the methods of experiments' design took place at the beginning of the 20th century in agricultural experiments [Elandt 1964; Kończak 2007]. Currently, the theory of experimental design is being developed in many areas of science: biology, medicine, chemistry and engineering, as well as in spatial data analysis.

The methods of the design of experiments lead to a very deep analysis of the studied phenomenon and especially the production process. The introduction of the algorithms of creating the designs of experiments, as well as the analysis of its results may be essential for using them in other areas of study.

**Małgorzata Złotoś**
University of Economics in Katowice
e-mail: malgorzata.zlotos@ue.katowice.pl
ORCID: 0000-0002-860-4848

## 2. Factorial designs of experiments

In statistical quality control the methods of the design of experiments lead to the determination of the factors which significantly affect the variable characterization of the investigated process, as well as allow to specify the values of the factors for which the result variable reaches the desired properties (e.g. proper value or the smallest variability).

The valid application of the design of experiments needs the proper preparation which should consist of the following points [Montgomery 1997]:

- recognition and definition of the problem by determining all the aspects, circumstances and potential objectives of the experiment;
- appropriate selection of the factors, their levels and ranges, and exploration of the possibility of considering them in the experiment;
- defining the response variable;
- choosing a proper design of experiment, i.e. determine the number of experimental trials and the possible randomization restrictions;
- performing the experiment;
- analysing the results using statistical methods;
- formulating conclusions and recommendations resulting from the analysis of the results.

The experiment is a sequence of $n$ successive experimental trials which are a single result of the value of response variable $Y$, with the fixed values of $X_1, X_2,\ldots, X_m$ specified on sets $\mathrm{X}_1, \mathrm{X}_2,\ldots, \mathrm{X}_\mathrm{m}$ respectively. The experimental area is a set of points $\mathrm{x} = (x_1, x_2,\ldots, x_m)$, where $x_i \in \mathrm{X}_i, i = 1,2,\ldots,m$. Then the set of pairs $P_n = \{\mathrm{x}_j, p_j\}_{j=1}^{n}$ is a design of experiment with $n$ experimental trials, where $\mathrm{x}_j = (x_{1j}, x_{2j},\ldots, x_{mj})$ and $p_j = \frac{n_j}{n}$, where $n_j$ is the number of experimental trials in point $\mathrm{x}_j$ of the experimental area, and also $\sum_{j=1}^{n} n_j = n$, $\sum_{j=1}^{n} p_j = 1$ for $j = 1,2,\ldots,n$. The relationship between the set of factors and the response variable characterizing the process in the best way is presented in the form of the following statistical model [Wawrzynek 2009]

$$Y(X_1, X_2,\ldots, X_m) = y(X_1, X_2,\ldots, X_m) + \varepsilon, \tag{1}$$

where $EY(X_1, X_2, \ldots, X_m) = y(X_1, X_2, \ldots, X_m)$, $E\varepsilon = 0$ and $V\varepsilon = \sigma^2$ where $\sigma^2$ is a constant value. It is assumed that the errors are uncorrelated with each other. Model (1) can be presented as a general linear model [Wawrzynek 2009] as $\mathrm{Y} = \mathrm{F}\beta + \varepsilon$ where

$$\mathrm{Y}^{\mathrm{T}} = \left(Y_1 Y_2 \ldots Y_n\right) \tag{2}$$

$$\varepsilon^{\mathrm{T}} = \left(\varepsilon_1 \varepsilon_2 \ldots \varepsilon_n\right) \tag{3}$$

$$\beta^{\mathrm{T}} = \left\lfloor \beta_1 \beta_2 \ldots \beta_k \right\rfloor \tag{4}$$

$$\mathrm{f}^{\mathrm{T}}(\mathrm{x}) = \left(f_1(\mathrm{x})\, f_2(\mathrm{x}) \ldots f_k(\mathrm{x})\right) \tag{5}$$

$$\mathrm{F} = \begin{bmatrix} f_1(\mathrm{x}_1) & \ldots & f_k(\mathrm{x}_1) \\ \vdots & \ddots & \vdots \\ f_1(\mathrm{x}_n) & \ldots & f_k(\mathrm{x}_n) \end{bmatrix} = \left[\mathrm{f}(\mathrm{x}_1)\ \mathrm{f}(\mathrm{x}_2) \ldots \mathrm{f}(\mathrm{x}_n)\right]^{\mathrm{T}} \tag{6}$$

and $f_i(\mathrm{x}_j) \equiv x_{ij}$, for $i = 1,2,\ldots k,\ j = 1,2,\ldots,n$.

The response surface function is defined using the formula y = Fβ. The realization of the factorial design of experiments usually leads to the estimation of the parameters of response surface function which is the first or second degree polynomial defined for variables $X_1, X_2, \ldots, X_m$ [Wawrzynek 2009]. In order to estimate of the parameters of a response surface function the least squares method is usually used [Wawrzynek 1993, 2009]. The significance of each parameter is usually verified using the *t* – test, whereas to determine whether the estimated model adequately fits to the data it is possible to use the $R^2$ value, adjusted $R^2$ value and the lack of fit test [Walesiak, Gatnar (eds.) 2009].

In practice, the production companies particularly used the factorial designs of experiments in which the factors occur on two or three levels.

The factorial design of experiment $2^m$ involves *m* factors $X_1, X_2, \ldots, X_m$ on two levels: upper – denoted "+" or "1" and lower – denoted "-" or " –1". Then the $2^m$ factorial design of the experiment needs $n = 2^m$ experimental trials and leads to the estimation of the parameters of response surface function expressed with the following equation [Wawrzynek 2009]

$$\mathrm{y}(\mathrm{x}) = \beta_0 + \beta_1 x_1 + \beta_2 x_2 + \ldots + \beta_m x_m + \beta_{12} x_1 x_2 + \ldots + \beta_{12\ldots m} x_1 x_2 \ldots x_m. \tag{7}$$

The $2^2$ factorial design of the experiment which involves *k* replications is presented in Table 1. This design allows for the estimation of the response surface function as follows
[Wawrzynek 2009]:

$$y(\mathrm{x}) = \beta_0 + \beta_1 x_1 + \beta_2 x_2 + \beta_{12} x_1 x_2 . \tag{8}$$

The $3^m$ factorial design of the experiment involves *m* factors $X_1, X_2, \ldots, X_m$ on three levels: upper – denoted "+" or "1", middle – denoted "0" and lower – denoted " – " or " –1". In particular the $3^m$ design of the experiment is a result of $n = 3^m$ experimental trials and allows for the estimation of the response surface function [Wawrzynek 2009]

$$\begin{aligned} y(\mathrm{x}) = {} & \beta_0 + \beta_1 x_1 + \beta_2 x_2 + \ldots + \beta_m x_m + \beta_{12} x_1 x_2 + \ldots + \beta_{m-1\,m} x_{m-1} x_m + \\ & + \beta_{123} x_1 x_2 x_3 + \ldots + \beta_{12\ldots m} x_1 x_2 \ldots x_m + \beta_{11} x_1^2 + \beta_{22} x_2^2 + \ldots + \beta_{m\,m} x_m^2 \end{aligned} . \tag{9}$$

The $3^2$ full factorial design of the experiment is presented in Table 2.

Table 1. The $2^2$factorial design of the experiment

| Sign of experimental trial | 1 | $X_1$ | $X_2$ | $X_1X_2$ | Values of the response variable | | | |
|---|---|---|---|---|---|---|---|---|
| | | | | | 1 | 2 | ... | k |
| *(1)* | + | - | - | + | $y_{11}$ | $y_{12}$ | ... | $y_{1k}$ |
| *a* | + | + | - | - | $y_{21}$ | $y_{22}$ | ... | $y_{2k}$ |
| *b* | + | - | + | - | $y_{31}$ | $y_{32}$ | ... | $y_{3k}$ |
| *ab* | + | + | + | + | $y_{41}$ | $y_{42}$ | ... | $y_{4k}$ |

Source: own elaboration.

Table 2. The $3^2$ factorial design of the experiment

| Sign of experimental trial | 1 | $X_1$ | $X_2$ | $X_1X_2$ | $X_1^2$ | $X_2^2$ |
|---|---|---|---|---|---|---|
| *(1)* | + | - | - | + | + | + |
| *a* | + | 0 | - | 0 | 0 | + |
| $a^2$ | + | + | - | - | + | + |
| *b* | + | - | 0 | 0 | + | 0 |
| *ab* | + | 0 | 0 | 0 | 0 | 0 |
| $a^2b$ | + | + | 0 | 0 | + | 0 |
| $b^2$ | + | - | + | - | + | + |
| $ab^2$ | + | 0 | + | 0 | 0 | + |
| $a^2\,b^2$ | + | + | + | + | + | + |

Source: own elaboration.

The design of the experiment which is presented in Table 2 allows for the estimation of the following response surface function [Wawrzynek 2009]

$$y(\mathrm{x}) = \beta_0 + \beta_1 x_1 + \beta_2 x_2 + \beta_{12} x_1 x_2 + \beta_{11} x_1^2 + \beta_{22} x_2^2 . \tag{10}$$

In the literature [Lawson 2015; Montgomery 2001; Ryan 2007; Wawrzynek 2009], the designs of experiments are also defined for factors with different number of levels. Then one may involve factors $X_1$, $X_2$, …, $X_m$, as follows

- $m_1$ factors on $p_1$ levels
- $m_2$ factors on $p_2$ levels
- …
- $m_k$ factors on $p_k$ levels,

where $m_1 + m_2 + \ldots + m_k = m$, so the defined design of the experiment includes $p_1^{m_1} \cdot p_2^{m_2} \cdot \ldots \cdot p_k^{m_k}$ experimental trials and leads to the estimation of response surface function (7) or (9). In Table 3 the design of the experiment for two factors with two levels and one factor with three levels is presented.

Table 3. The design of the experiment defined for factors with different levels of factors

| No. | 1 | $X_1$ | $X_2$ | $X_3$ |
|---|---|---|---|---|
| 1 | + | + | + | + |
| 2 | + | + | - | 0 |
| 3 | + | + | + | - |
| 4 | + | + | - | + |
| 5 | + | + | + | 0 |
| 6 | + | + | - | - |
| 7 | + | - | + | + |
| 8 | + | - | - | 0 |
| 9 | + | - | + | - |
| 10 | + | - | - | + |
| 11 | + | - | + | 0 |
| 12 | + | - | - | - |

Source: own elaboration.

In practice, the design of the factorial experiment with a different number of factor levels is defined with proper tables denoted as orthogonal arrays [Wawrzynek 2009].

For selected cases, D.C. Montgomery also presents the construction method of the design of an experiment taking into account the factors on a various number of levels. This method, defined for quantitative variables, consists of the appropriate determination of factor levels using the additional factors [Montgomery 2001]. When the factors are qualitative variables on more than two levels, then the additional variables determine the proper levels of factors.

## 3. The proposed method of the use of factorial design of an experiment in individual data analysis

In this paper the data about a single unit with objective or subjective information from surveys or from available administrative databases, is considered. In the literature this type of data is called microdata, and their analysis is the subject of microeconometrics [Gruszczyński (ed.) 2010; Bąk 2013].

Therefore a set of data is determined by the obtained survey, the purpose of which was the Y estimated (explanatory variable) of a certain attribute based on explanatory variables $X_1$ and $X_2$. It was assumed that each of the variables $X_1$ and $X_2$ receives two values of the characteristic $A_i$ and $B_i$ for $i = 1,2$. The results of such a specific survey, for $k$ obtained complete answers, are presented in Table 4.

Table 4. Results of the survey with $k$ complete answers

| No. | Explanatory variables | | Values of response variable | | | |
|---|---|---|---|---|---|---|
| | $X_1$ | $X_2$ | 1 | 2 | ... | $k$ |
| 1 | $A_1$ | $A_2$ | $y_{11}$ | $y_{12}$ | ... | $y_{1k}$ |
| 2 | $B_1$ | $A_2$ | $y_{21}$ | $y_{22}$ | ... | $y_{2k}$ |
| 3 | $A_1$ | $B_2$ | $y_{31}$ | $y_{32}$ | ... | $y_{3k}$ |
| 4 | $B_1$ | $B_2$ | $y_{41}$ | $y_{42}$ | ... | $y_{4k}$ |

Source: own elaboration.

When the independent variable $X_1$ and $X_2$ are denoted as factors with two levels $A_i$ and $B_i$ for $i=1,2$, and the values of the dependent variable Y are denoted as empirical values of response surface function for $k$ replication, then the results of the surveys presented in Table 4 can be recognized as the realizations of the experimental trials of the $2^2$ factorial design of the experiment. Then the dependence of response variable Y and explanatory variables $X_1$ and $X_2$ can be defined as response surface function (8). The usage of the least squares method leads to the estimation of the coefficients of response surface function and allows for the estimation of the impact of factors on the response variable.

## 4. The use of the proposed method for empirical data

The set of survey data about the evaluation of CDs is considered. The survey includes the following explanatory variables (factors): $X_1$ – the volume of the CD, $X_2$ – the package of the CD and $X_3$ – colour of the CD, and the response variable $Y$ is the rate of individual CD variants denoted in the metric scale. Factor $X_1$ takes two levels: lower ”650MB” and upper – ”700MB”, factor $X_2$ also takes two levels: lower – ”box”, upper – ”envelope”, and factor $X_3$ takes three levels: lower – ”blue”, middle – ”silver” and upper – ”gold”. Then the complete design of experiments for two factors on two levels and one factor on three levels is presented in Table 5.

The purpose of the experiment is to estimate the values of the parameters of the response surface function as follows

$$y(x) = \beta_0 + \beta_1 x_1 + \beta_2 x_2 + \beta_3 x_3. \tag{11}$$

The values of response variable obtained six times for eight CD variants presented in Table 6.

Factor $X_3$ occurs on three levels, therefore in Table 7 the additional variables $X_3^{(1)}$ and $X_3^{(2)}$ are determined. Variables $X_3^{(1)}$ and $X_3^{(2)}$ define the values of variable $X_3$.

Table 5. The complete design of the experiment for empirical data

| No. | $X_1$ – Volume | $X_2$ – Package | $X_3$ – Colour |
|---|---|---|---|
| 1 | 650MB | box | blue |
| 2 | 700MB | box | blue |
| 3 | 650MB | envelope | blue |
| 4 | 700MB | envelope | blue |
| 5 | 650MB | box | silver |
| 6 | 700MB | box | silver |
| 7 | 650MB | envelope | silver |
| 8 | 700MB | envelope | silver |
| 9 | 650MB | box | gold |
| 10 | 700MB | box | gold |
| 11 | 650MB | envelope | gold |
| 12 | 700MB | envelope | gold |

Source: own elaboration.

Table 6. The results of the surveys

| No. | $X_1$ | $X_2$ | $X_3$ | $y_1$ | $y_2$ | $y_3$ | $y_4$ | $y_5$ | $y_6$ |
|---|---|---|---|---|---|---|---|---|---|
| 1 | 1 | 1 | 1 | 70 | 90 | 80 | 80 | 60 | 90 |
| 2 | 1 | –1 | –1 | 30 | 40 | 30 | 80 | 90 | 50 |
| 3 | –1 | 1 | –1 | 20 | 30 | 20 | 15 | 40 | 35 |
| 4 | –1 | –1 | 1 | 50 | 60 | 50 | 40 | 50 | 70 |
| 5 | –1 | –1 | 0 | 50 | 50 | 40 | 40 | 50 | 60 |
| 6 | 1 | 1 | 0 | 40 | 80 | 70 | 70 | 80 | 90 |
| 7 | –1 | 1 | 0 | 60 | 70 | 60 | 30 | 20 | 80 |
| 8 | –1 | 1 | 1 | 90 | 80 | 70 | 80 | 90 | 90 |

Source: own elaboration.

Table 7. The definition of variable $X_3$ with the additional variables $X_3^{(1)}$ and $X_3^{(2)}$

| $X_3$ | $X_3^{(1)}$ | $X_3^{(2)}$ |
|---|---|---|
| 1 | 1 | 0 |
| 0 | –1 | –1 |
| –1 | 0 | 1 |

Source: own elaboration.

Then the considered factorial design of the experiment can be presented as in Table 8.

Table 8. The factorial design of the experiment with additional variables

| No. | $X_1$ | $X_2$ | $X_3$ | | $y_1$ | $y_2$ | $y_3$ | $y_4$ | $y_5$ | $y_6$ |
|---|---|---|---|---|---|---|---|---|---|---|
| | | | $X_3^{(1)}$ | $X_3^{(2)}$ | | | | | | |
| 1 | 1 | 1 | 1 | 0 | 70 | 90 | 80 | 80 | 60 | 90 |
| 2 | 1 | –1 | –1 | –1 | 30 | 40 | 30 | 80 | 90 | 50 |
| 3 | –1 | 1 | –1 | –1 | 20 | 30 | 20 | 15 | 40 | 35 |
| 4 | –1 | –1 | 1 | 0 | 50 | 60 | 50 | 40 | 50 | 70 |
| 5 | –1 | –1 | 0 | 1 | 50 | 50 | 40 | 40 | 50 | 60 |
| 6 | 1 | 1 | 0 | 1 | 40 | 80 | 70 | 70 | 80 | 90 |
| 7 | –1 | 1 | 1 | 0 | 60 | 70 | 60 | 30 | 20 | 80 |
| 8 | 1 | –1 | 1 | 0 | 90 | 80 | 70 | 80 | 90 | 90 |

Source: own elaboration.

Moreover the response surface function (11) is given in the form

$$y(\mathrm{x}) = \beta_0 + \beta_1 x_1 + \beta_2 x_2 + \beta_3^{(1)} x_3^{(1)} + \beta_3^{(2)} x_3^{(2)}. \tag{12}$$

On the basis of the obtained results, the parameters of response surface function (12) were estimated for *k* replications. The results of this estimation are presented in Table 9.

Table 9. The estimated coefficients of the response surface function for *k* replications

| $y_k(x)$ | Values of the coefficients of the response surface function | | | | |
|---|---|---|---|---|---|
| | $\beta_0$ | $\beta_1$ | $\beta_2$ | $\beta_3^{(1)}$ | $\beta_3^{(2)}$ |
| $y_1(x)$ | 45.83 | 6.25 | –3.75 | 21.67 | –0.83 |
| $y_2(x)$ | 58.33 | 10.00 | 5.00 | 16.67 | 6.67 |
| $y_3(x)$ | 48.33 | 10.00 | 5.00 | 16.67 | 6.67 |
| $y_4(x)$ | 53.33 | 23.13 | –5.63 | 4.17 | 1.67 |
| $y_5(x)$ | 61.67 | 20.00 | –10.00 | –6.67 | 3.33 |
| $y_6(x)$ | 66.67 | 9.38 | 3.13 | 15.83 | 8.33 |

Source: own elaboration.

In order to assess the importance and influence of factors on the response variable for each of the parameters, the *p*-values of t – test were estimated.

The results are presented in Table 10. Table 11 presents the outcome of the assessment of fit of the response surface function to the empirical data.

Table 10. The *p*-values of the estimated coefficients

| $y_k(x)$ | *p*-values of the coefficients | | | | |
|---|---|---|---|---|---|
| | $\beta_0$ | $\beta_1$ | $\beta_2$ | $\beta_3^{(1)}$ | $\beta_3^{(2)}$ |
| $y_1(x)$ | **0.0028** | 0.2828 | 0.4906 | **0.0426** | 0.9184 |
| $y_2(x)$ | **0.0000** | **0.0000** | **0.0000** | **0.0000** | **0.0000** |
| $y_3(x)$ | **0.0000** | **0.0000** | **0.0000** | **0.0000** | **0.0000** |
| $y_4(x)$ | **0.0002** | **0.0016** | 0.0754 | 0.2341 | 0.6473 |
| $y_5(x)$ | **0.0003** | **0.0062** | **0.0405** | 0.1817 | 0.5137 |
| $y_6(x)$ | **0.0000** | **0.0088** | 0.1339 | **0.0045** | **0.0400** |

Source: own elaboration.

Table 11. The $R^2$ values, adjusted $R^2$ values, values of $F$ statistic with the corresponding *p*-values, of the estimated response surface functions

| $y_k(x)$ | $R^2$ | $\overline{R}^2$ | $F$ | *p*-value |
|---|---|---|---|---|
| $y_1(x)$ | 0.8423 | 0.632 | 4.006 | 0.1418 |
| $y_2(x)$ | 1 | 1 | $9.243 \cdot 10^{30}$ | 0.0000 |
| $y_3(x)$ | 1 | 1 | $2.908 \cdot 10^{31}$ | 0.0000 |
| $y_4(x)$ | 0.9778 | 0.948 | 32.93 | 0.0082 |
| $y_5(x)$ | 0.9545 | 0.8939 | 15.75 | 0.0236 |
| $y_6(x)$ | 0.9814 | 0.9566 | 3954 | 0.0063 |

Source: own elaboration.

Based on the obtained results, it can be assumed that the estimated response surface functions $y_2(x)$, $y_3(x)$, $y_4(x)$, $y_5(x)$, to $y_6(x)$ are significant.

The design of the experiment which allows for the estimation of the response surface function parameters for the mean of value of the response variable is presented in Table 12.

As a result the following estimator of the response surface function was obtained

$$y_{sr}(\mathrm{x}) = 55.69 + 13.125x_1 - 1.04x_2 + 11.39x_3^{(1)} + 4.31x_3^{(2)}. \qquad (13)$$

The *p*-values of the parameters of response surface function (13) are presented in Table 13.

Table 12. The design of the experiment for the mean value of the response variable

| No. | $X_1$ | $X_2$ | $X_3$ | | $y_{sr}$ |
|---|---|---|---|---|---|
| | | | $X_3^{(1)}$ | $X_3^{(2)}$ | |
| 1 | 1 | 1 | 1 | 0 | 78.33 |
| 2 | 1 | –1 | –1 | –1 | 53.33 |
| 3 | –1 | 1 | –1 | –1 | 26.67 |
| 4 | –1 | –1 | 1 | 0 | 53.33 |
| 5 | –1 | –1 | 0 | 1 | 48.33 |
| 6 | 1 | 1 | 0 | 1 | 71.67 |
| 7 | –1 | 1 | 0 | 1 | 53.33 |
| 8 | –1 | 1 | 1 | 0 | 83.33 |

Source: own elaboration.

Table 13. The *p*-value of estimated coefficients of the response surface function (13)

| $\beta_i$ | $\beta_0$ | $\beta_1$ | $\beta_2$ | $\beta_3^{(1)}$ | $\beta_3^{(2)}$ |
|---|---|---|---|---|---|
| *p*-value | **0.0000** | **0.0002** | 0.2002 | **0.0009** | **0.0228** |

Source: own elaboration.

For response surface function $R^2 = 0.9959$, $\bar{R}^2 = 0.9905$, $F = 182.9$ and *p*-value = 0.0007. Therefore the response surface function (13) is significant.

Table 14. The result of the estimation of the response surface function values for unrealized experimental trials

| No. | $X_1$ | $X_2$ | $X_3$ | | $\tilde{y}$ |
|---|---|---|---|---|---|
| | | | $X_3^{(1)}$ | $X_3^{(2)}$ | |
| 1 | 1 | –1 | 0 | 1 | 74.17 |
| 2 | –1 | 1 | 0 | 1 | 45.83 |
| 3 | –1 | –1 | –1 | –1 | 27.92 |
| 4 | 1 | 1 | –1 | –1 | 52.03 |

Source: own elaboration.

According to the estimation of the response surface function (13), it is possible to estimate the values of the response variable for unrealized experimental trials. The results of this estimation are presented in Table 14.

The implementation of the design of the experiments' methods in individual data analysis allows to determine the proper dependence between the factors and the dependent variable. The presented considerations lead to the assessment of the significance of certain response surface functions and to the estimation of unknown values of the response variable.

## 5. Conclusions

The design of experiments as the statistical method of quality control allows for the achievement of a better production process. In other areas of the study the planned experiments may be used in practice as well as in theory.

This elaboration presents the proposition of using the methodology of planning the experiments in individual data analysis. This method enables the determination of the response surface function which characterizes the studied process or the phenomenon, on the basis of which we can determine the influence of other factors on the response variable. Moreover, estimating the formula of the response surface function for each replicas of the experiment or for the average value, and on these basis the value for missing variants of the studied feature might be set.

## Bibliography

Bąk A. (2013). *Mikroekonometryczne metody badania preferencji konsumentów z wykorzystaniem programu R*. Wydawnictwo C.H. Beck. Warszawa.

Elandt R. (1964). *Statystyka matematyczna w zastosowaniu do doświadczalnictwa rolniczego*. PWN. Warszawa.

Gruszczyński M. (ed.) (2010). *Mikroekonometria. Modele i metody analizy danych indywidualnych*. Wolters Kluwer. Warszawa.

Kończak G. (2007). *Metody statystyczne w sterowaniu jakością produkcji*. Wydawnictwo Akademii Ekonomicznej w Katowicach. Katowice.

Lawson J. (2015). *Design and Analysis of Experiments with R*. CRC Press Taylor & Francis Group. Boca Raton.

Montgomery D.C. (1997). *Introduction to Statistical Quality Control*. John Wiley & Sons, Inc. New York.

Montgomery D.C. (2001). *Design and Analysis of Experiments*. John Wiley & Sons, Inc. New York.

Ryan T.P. (2007). *Modern Experimental Design*. John Wiley & Sons. New Jersey.

Walesiak M., Gatnar E. (eds.) (2009). *Statystyczna analiza danych z wykorzystaniem programu R*. Wydawnictwo Naukowe PWN. Warszawa.

Wawrzynek J. (1993). *Statystyczne planowanie eksperymentów w zagadnieniach regresji w warunkach małej próby*. Wydawnictwo Akademii Ekonomicznej we Wrocławiu. Wrocław.

Wawrzynek J. (2009). *Planowanie eksperymentów zorientowane na doskonalenie jakości produktu*. Wydawnictwo Uniwersytetu Ekonomicznego we Wrocławiu. Wrocław.