Małgorzata Szerszunowicz
Uniwersytet Ekonomiczny w Katowicach

O PEWNYM PLANIE EKSPERYMENTU CZYNNIKOWEGO WYKORZYSTUJĄCYM MODELE PÓL LOSOWYCH*

Wprowadzenie

Planowanie eksperymentów, będące narzędziem statystycznej kontroli jakości, pozwala na określenie zależności pomiędzy czynnikami uwzględnianymi w realizacji procesu produkcyjnego a rezultatami tego procesu. Klasyczne plany eksperymentów wymagają odpowiedniego zdefiniowania poziomów poszczególnych czynników oraz wyznaczenia kolejnych punktów obszaru eksperymentowania, co zasadniczo wiąże się z kosztami realizacji eksperymentu.

Przedmiotem artykułu będzie przedstawienie częściowego planu eksperymentu czynnikowego wykorzystującego modele charakteryzujące badaną właściwość, zdefiniowane za pomocą pól losowych. Plan ten umożliwi wyznaczenie kolejnych punktów obszaru eksperymentowania wtedy, gdy zastosowanie klasycznych planów doświadczeń nie jest możliwe bądź wiąże się ze zwiększeniem kosztów realizacji eksperymentu.

1. Planowanie eksperymentów w statystycznej kontroli jakości

Planowanie eksperymentów jest istotnym etapem poprzedzającym proces produkcyjny, dlatego jego stosowanie powinno przebiegać zgodnie z określonymi kryteriami, które D.C. Montgomery formułuje następująco (Montgomery, 1997):

– identyfikacja i sformułowanie problemu, polegające na określeniu wszystkich aspektów, okoliczności i potencjalnych celów eksperymentu,

* Projekt został sfinansowany ze środków Narodowego Centrum Nauki przyznanych na podstawie decyzji numer DEC-2011/03/B/HS4/05630.

– odpowiedni dobór czynników, ich poziomów, zakresów zmienności, jak również ustalenie możliwości uwzględnienia ich w planowanym eksperymencie,
– określenie zmiennej objaśnianej na wyjściu,
– wybór odpowiedniego planu eksperymentu polegający na określeniu liczby doświadczeń oraz ewentualnych reguł randomizacji,
– realizacja eksperymentu,
– analiza otrzymanych wyników mająca charakter statystyczny,
– wnioski i zalecenia wynikające dla badanego sformułowane na podstawie analizy wyników eksperymentu.

Pod pojęciem eksperymentu należy rozumieć ciąg n kolejnych doświadczeń, gdzie pojedyncze doświadczenie jest jednorazowym uzyskaniem wartości zmiennej objaśnianej Y, przy ustalonych wartościach poszczególnych czynników $X_1, X_2, \ldots, X_m$. Niech $\mathbf{X_1}, \mathbf{X_2}, \ldots, \mathbf{X_m}$ oznaczają zbiory wszystkich możliwych wartości czynników $X_1, X_2, \ldots, X_m$, wówczas obszar eksperymentowania jest zbiorem punktów $\mathbf{x} = (x_1, x_2, \ldots, x_m)$, gdzie $x_i \in \mathrm{X}_i$, $i = 1, 2, \ldots, m$. Zbiór par w postaci $P_n = \{\mathrm{x}_j, p_j\}_{j=1}^{n}$ określa plan eksperymentu obejmujący n doświadczeń, gdzie $\mathbf{x}_j = (x_{1j}, x_{2j}, \ldots, x_{mj})$ oraz $p_j = \frac{n_j}{n}$, przy czym n_j oznacza liczbę doświadczeń w punkcie $\mathbf{x}_j$ obszaru eksperymentowania, ponadto $\sum_{j=1}^{n} n_j = n$ oraz $\sum_{j=1}^{n} p_j = 1$ dla $j = 1, 2, \ldots, n$.

Zależność będącą przedmiotem badania eksperymentalnego, która określa wpływ losowych i nielosowych czynników $X_1, X_2, \ldots, X_m$ na zmienną wynikową Y, można przedstawić w postaci modelu statystycznego określonego równaniem (Wawrzynek, 2009):

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

gdzie $EY(X_1, X_2, \ldots, X_m) = y(X_1, X_2, \ldots, X_m)$, $E\varepsilon = 0$ oraz $V\varepsilon = \sigma^2$, przy czym σ^2 jest wielkością ustaloną, niezależną od wartości poszczególnych czynników, natomiast funkcja $y(x_1, x_2, \ldots, x_m)$ jest nazywana powierzchnią odpowiedzi. Argumentami funkcji powierzchni odpowiedzi są realizacje m nielosowych zmiennych $X_1, X_2, \ldots, X_m$. Model (1) przedstawia się w postaci ogólnego modelu liniowego $\mathbf{Y} = \mathbf{F}\boldsymbol{\beta} + \boldsymbol{\varepsilon}$, gdzie:

$$\mathbf{Y}^{\mathbf{T}} = \left(Y_1 Y_2 \ldots Y_n\right),$$
$$\boldsymbol{\varepsilon}^{\mathbf{T}} = \left(\varepsilon_1 \varepsilon_2 \ldots \varepsilon_n\right),$$
$$\boldsymbol{\beta}^{\mathbf{T}} = \left[\beta_1 \beta_2 \ldots \beta_k\right],$$
$$\mathbf{f}^{\mathbf{T}}(\mathbf{x}) = \left(f_1(\mathbf{x}) f_2(\mathbf{x}) \ldots f_k(\mathbf{x})\right),$$
$$\mathbf{F} = \begin{bmatrix} f_1(\mathbf{x}_1) & \ldots & f_k(\mathbf{x}_1) \\ \vdots & \ddots & \vdots \\ f_1(\mathbf{x}_n) & \ldots & f_k(\mathbf{x}_n) \end{bmatrix} = \left[\mathbf{f}(\mathbf{x}_1)\ \mathbf{f}(\mathbf{x}_2) \ldots \mathbf{f}(\mathbf{x}_n)\right]^{\mathbf{T}},$$

przy czym $f_i(\mathbf{x}_j) \equiv x_{ij}$, dla $i = 1,2,\ldots,k, j = 1, 2,\ldots, n$. Estymacji parametrów funkcji powierzchni odpowiedzi w postaci $\mathbf{y} = \mathbf{F}\boldsymbol{\beta}$ dokonuje się metodą najmniejszych kwadratów (Wawrzynek, 1993). Wówczas wariancja estymatora funkcji powierzchni odpowiedzi wyraża się wzorem:

$$\mathbf{V}\tilde{\mathbf{y}}(\mathbf{x}) = \sigma^2 \mathbf{f}^{\mathbf{T}}(\mathbf{x})(\mathbf{F}^{\mathbf{T}}\mathbf{F})^{-1}\mathbf{f}(\mathbf{x}), \tag{2}$$

ponadto jej wartość zależy wyłącznie od wartości elementów macierzy kwadratowej $\mathbf{F}^{\mathbf{T}}\mathbf{F}$, czyli od wyboru odpowiedniego planu eksperymentu.

Zazwyczaj w literaturze dla klasycznych planów eksperymentów rozważa się funkcje powierzchni odpowiedzi w postaci:

$$y(\mathrm{x}) = \beta_0 + \beta_1 x_1 + \beta_2 x_2 + \ldots + \beta_m x_m + \beta_{12} x_1 x_2 + \ldots + \beta_{12\ldots m} x_1 x_2 \ldots x_m, \tag{3}$$

dla której estymacja parametrów funkcji powierzchni odpowiedzi polega na realizacji eksperymentu uwzględniającego m czynników występujących na n_i poziomach każdy, obejmującego $n = \prod_{i=1}^{m} n_i$ losowo przeprowadzonych doświadczeń.

2. Pola losowe a wytrzymałość materiałów

Teoria pól losowych jest wykorzystywana w metodach statystyki przestrzennej, w szczególności w takich dziedzinach, jak meteorologia czy genealogia. W niniejszym artykule pola losowe będą służyć jako narzędzie wspomagające planowanie eksperymentów. Niech zatem dana będzie przestrzeń probabilistyczna (Ω, A, P). Funkcja $U: \mathbf{R}^n \times \Omega \to \mathbf{R}^l$ taka, że dla każdego $\boldsymbol{t} \in \mathbf{R}^n$ odwzorowanie $\omega \mapsto U(\boldsymbol{t}, \omega)$ jest zmienną losową w Ω, jest n-wymiarowym polem losowym. Pole losowe jest wyznaczane za pomocą łącznej dystrybuanty określonej dla zmiennych losowych $U(\boldsymbol{t}_1), U(\boldsymbol{t}_2), \ldots, U(\boldsymbol{t}_m)$ w postaci:

$$F_{\boldsymbol{t}_1,\boldsymbol{t}_2,\ldots,\boldsymbol{t}_m}(u_1,u_2,\ldots,u_m) = P(U(\boldsymbol{t}_1) < u_1, U(\boldsymbol{t}_2) < u_2, \ldots, U(\boldsymbol{t}_m) < u_m),$$

gdzie $U(\boldsymbol{t})$ jest n-wymiarowym polem losowym, natomiast $\boldsymbol{t}_1, \boldsymbol{t}_2, \ldots, \boldsymbol{t}_m$ są wartościami parametru $\boldsymbol{t}$. Jeżeli $U(\boldsymbol{t})$ jest ciągłym polem losowym, czyli funkcja $U : \mathbf{R}^n \times \Omega \to \mathbf{R}^1$ jest ciągła w zbiorze $\mathbf{R}^n \times \Omega$, to wyrażenie:

$$\begin{aligned} m_{j_1,j_2,\ldots,j_s}(\boldsymbol{t}_1,\boldsymbol{t}_2,\ldots,\boldsymbol{t}_s) &= E((U(\boldsymbol{t}_1))^{j_1}(U(\boldsymbol{t}_2))^{j_2}\ldots(U(\boldsymbol{t}_s))^{j_s}) = \\ &= \int_\Omega (u(\boldsymbol{t}_1))^{j_1}(u(\boldsymbol{t}_2))^{j_2}\ldots(u(\boldsymbol{t}_s))^{j_s}\,dP(u) \end{aligned}$$

jest zwykłym momentem rzędu $j_1 + \ldots + j_s$ w punktach $\boldsymbol{t}_1, \ldots, \boldsymbol{t}_s$, dla $j_k > 0,\ k = 1,2,\ldots,s,\ s \geq 1$. Analogicznie określa się momenty centralne rzędu $j_1 + \ldots + j_s$ w punktach $\boldsymbol{t}_1, \ldots, \boldsymbol{t}_s$ w postaci:

$$\begin{aligned} \mu_{j_1,j_2,\ldots,j_s}(\boldsymbol{t}_1,\boldsymbol{t}_2,\ldots,\boldsymbol{t}_s) &= E([U(\boldsymbol{t}_1) - E(U(\boldsymbol{t}_1))]^{j_1}\ldots[U(\boldsymbol{t}_s) - E(U(\boldsymbol{t}_s))]^{j_s}) = \\ &= \int_\Omega [u(\boldsymbol{t}_1) - E(U(\boldsymbol{t}_1))]^{j_1}\ldots[u(\boldsymbol{t}_s) - E(U(\boldsymbol{t}_s))]^{j_s}\,dP(u), \end{aligned}$$

gdzie $j_k > 0,\ k = 1,2,\ldots,s,\ s \geq 1$. Ponadto dla ustalonych wartości $\boldsymbol{t}_1, \boldsymbol{t}_2$ parametru $\boldsymbol{t} \in \mathbf{R}^n$ można określić funkcję korelacji pola losowego w zbiorze $I \subset \mathbf{R}^n$ dla $\boldsymbol{t}_1, \boldsymbol{t}_2 \in I$ za pomocą wyrażenia:

$$R(U(\boldsymbol{t}_1), U(\boldsymbol{t}_2)) = R(\boldsymbol{t}_1, \boldsymbol{t}_2) = \frac{\operatorname{cov}(U(\boldsymbol{t}_1), U(\boldsymbol{t}_2))}{\sigma(\boldsymbol{t}_1)\sigma(\boldsymbol{t}_2)},$$

gdzie $\operatorname{cov}(U(\boldsymbol{t}_1), U(\boldsymbol{t}_2))$ oznacza kowariancję zmiennych losowych $U(\boldsymbol{t}_1)$ i $U(\boldsymbol{t}_2)$, natomiast $\sigma^2(\boldsymbol{t}_1)$ i $\sigma^2(\boldsymbol{t}_2)$ odpowiednie wariancje tych zmiennych. Dalsze rozważania zostaną ograniczone jedynie do modeli pól losowych stacjonarnych w szerszym sensie, czyli takich, że dla każdego $\boldsymbol{t} \in \mathbf{R}^n$ funkcja $U(\boldsymbol{t})$ ma skończoną wartość oczekiwaną i skończoną wariancję oraz dla każdych $\boldsymbol{t}_1, \boldsymbol{t}_2 \in \mathbf{R}^n$ jest spełniona równość:

$$R(U(\boldsymbol{t}_1), U(\boldsymbol{t}_2)) = R(\boldsymbol{t}_1, \boldsymbol{t}_2) = R(\overrightarrow{\boldsymbol{t}_1\boldsymbol{t}_2}),$$

gdzie $\overrightarrow{\boldsymbol{t}_1\boldsymbol{t}_2}$ jest wektorem łączącym punkty $\boldsymbol{t}_1, \boldsymbol{t}_2$. Stacjonarne w szerszym sensie pole losowe nazywa się izotropowym, gdy zachodzi warunek:

$$R(\boldsymbol{t}_1, \boldsymbol{s}_1) = R(\boldsymbol{t}_2, \boldsymbol{s}_2),$$

zakładając równość odległości pomiędzy punktami $\boldsymbol{t}_1$ i $\boldsymbol{s}_1$ oraz $\boldsymbol{t}_2$ i $\boldsymbol{s}_2$.

Ze względu na losowy charakter struktury materiałów wykorzystywanych w procesach produkcyjnych, cechy wytrzymałościowe materiałów mogą być opisane z wykorzystaniem teorii pól losowych.

Niech $D \subset \mathbf{R}^2$ będzie częścią materiału, którego wytrzymałość należy zbadać. Wówczas pole losowe $w(\boldsymbol{p})$ określone w zbiorze D opisuje jedną z cech wytrzymałościowych badanego materiału w punkcie $\boldsymbol{p} \in D$ (Szczepankiewicz, 1985). W celu badania materiałów, w których nie można wyróżnić kierunków zmienności cechy wytrzymałości, wykorzystuje się pola losowe stacjonarne w szerszym sensie, izotropowe (Szczepankiewicz, 1985).

3. Częściowy plan eksperymentu czynnikowego

Niech dany będzie eksperyment polegający na weryfikacji wytrzymałości materiału, uwzględniający k czynników występujących na pewnej liczbie poziomów każdy. Celem eksperymentu czynnikowego będzie estymacja funkcji powierzchni odpowiedzi opisującej cechę wytrzymałości materiału w postaci (3), ponadto niech koszty eksperymentu będą ograniczone tak, że realizacja doświadczeń we wszystkich punktach obszaru eksperymentowania nie jest możliwa.

Zgodnie z regułami planowania eksperymentów w przypadku całkowitych i ułamkowych eksperymentów czynnikowych, kolejne doświadczenia powinny być zrealizowane w losowo wybranych punktach obszaru eksperymentowania. Ze względu na losowy charakter struktury materiałów podejście to może prowadzić do niewiarygodnych wyników. Wówczas można się posłużyć planem eksperymentu wykorzystującym informacje dotyczące cechy wytrzymałości materiału określone za pomocą pól losowych stacjonarnych w szerszym sensie, izotropowych. Zatem niech na podstawie fachowej wiedzy eksperymentatora i wyników pochodzących z wcześniejszych pomiarów, ustalona cecha wytrzymałościowa wykorzystywanego materiału będzie określona za pomocą stacjonarnego w szerszym sensie, izotropowego pola losowego $w(\boldsymbol{p})$. Wtedy rozważany eksperyment można zrealizować następująco:

1. Zdefiniowanie wszystkich punktów obszaru eksperymentowania.

2. Wyznaczenie wartości pola losowego $w(\boldsymbol{p})$ dla każdego punktu obszaru eksperymentowania.
3. Wybór punktów obszaru eksperymentowania (na podstawie fachowej wiedzy eksperymentatora) z uwzględnieniem realizacji pola losowego $w(\boldsymbol{p})$.
4. Losowa realizacja doświadczeń we wskazanych punktach obszaru eksperymentowania.

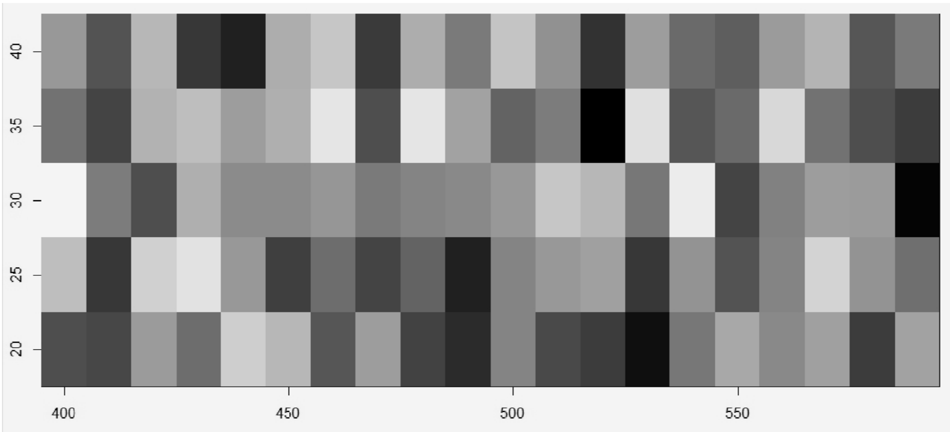
Eksperyment przeprowadzony zgodnie z podanym schematem pozwala na wyznaczenie planu doświadczeń z uwzględnieniem cech wytrzymałościowych wykorzystywanego materiału, ponadto może być zrealizowany przy ograniczonych kosztach eksperymentu czy też przy ograniczonej liczbie doświadczeń. W szczególności przedstawiony algorytm realizacji eksperymentu może być wykorzystywany wtedy, gdy obszar eksperymentowania składa się z wielu punktów.

4. Zastosowanie częściowego planu eksperymentu czynnikowego

D.C. Montgomery rozważa eksperyment polegający na badaniu wytrzymałości arkusza papieru w zależności od dwóch czynników: stężenia masy celulozowej (X_1) oraz poziomu ciśnienia (X_2) (Montgomery, 2001). Niech zatem podobnie stężenie masy celulozowej mieści się w przedziale od 20% do 40%, natomiast ciśnienie waha się od 400 Pa do 590 Pa.

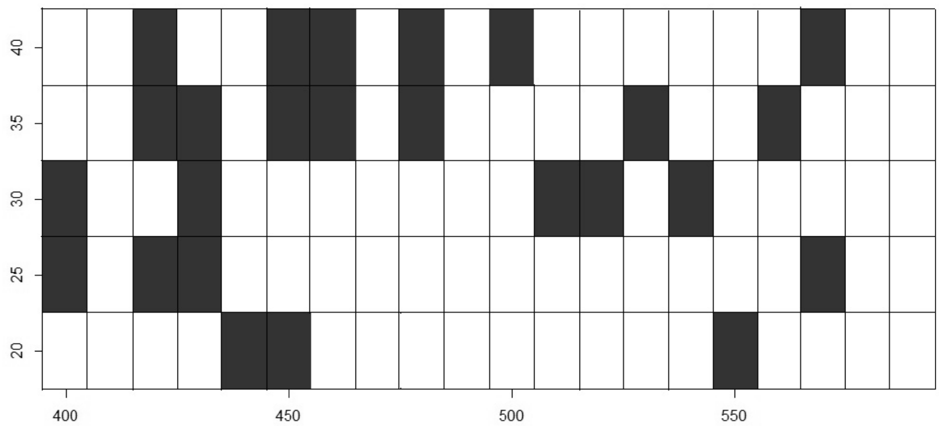
Na podstawie fachowej wiedzy eksperymentatora i wcześniejszych pomiarów w badaniu założono, że wytrzymałość papieru kształtuje się zgodnie z realizacjami stacjonarnego w szerszym sensie, izotropowego pola losowego, dla którego zmienne losowe mają rozkład normalny o ustalonych parametrach μ_w i σ_w. Ponadto przyjęto, że czynnik X_1 występuje na 5 poziomach, natomiast czynnik X_2 na 20 poziomach. Wówczas obszar eksperymentowania składa się ze 100 punktów, zatem realizacja całkowitego eksperymentu czynnikowego wymagałaby przeprowadzenia 100 doświadczeń, co wiąże się z wydłużeniem i zwiększonymi kosztami planowanego eksperymentu. Zatem niech eksperyment zostanie zrealizowany zgodnie z proponowanym częściowym planem wykorzystującym teorię pól losowych. Obszar eksperymentowania składa się z punktów o współrzędnych odpowiadających poszczególnym poziomom czynników X_1 i X_2. Na podstawie danych historycznych wyznaczono wartości pola losowego w poszczególnych punktach zadanego obszaru eksperymentowania (rysunek 1).

Przyjęto, że koszty realizacji eksperymentu pozwalają na realizację co najwyżej 25 doświadczeń. Zatem w kolejnym kroku wybrano 25 punktów obszaru eksperymentowania o najmniejszych wartościach pola losowego.



Rys. 1. Graficzna prezentacja realizacji pola losowego

Wówczas częściowy plan obejmuje 25 punktów obszaru eksperymentowania (rysunek 2), w których należy przeprowadzić doświadczenia w sposób losowy.



Rys. 2. Punkty obszaru eksperymentowania wybrane na podstawie realizacji pola losowego

Dzięki wykorzystaniu informacji o wytrzymałości materiału określonej za pomocą pola losowego częściowy plan eksperymentu pozwolił na ograniczenie liczby doświadczeń przy ograniczonych kosztach eksperymentu.

Tabela 1

Wartości wariancji estymatora funkcji powierzchni odpowiedzi
w zależności od stosowanego planu eksperymentu czynnikowego

Liczba doświadczeń	Plan eksperymentu czynnikowego	Wartość najmniejsza wariancji estymatora funkcji powierzchni odpowiedzi	Wartość największa wariancji estymatora funkcji powierzchni odpowiedzi
100	całkowity	0,0066 σ^2	0,0516 σ^2
25	częściowy	0,0297 σ^2	0,2550 σ^2
	losowy	0,0335 σ^2	0,1944 σ^2
50	częściowy	0,0133 σ^2	0,1019 σ^2
	losowy	0,0140 σ^2	0,1088 σ^2
75	częściowy	0,0087 σ^2	0,0742 σ^2
	losowy	0,0089 σ^2	0,0747 σ^2

Ponadto wyznaczono wartości wariancji funkcji powierzchni odpowiedzi dla punktów obszaru eksperymentowania całkowitego i częściowego planu eksperymentu czynnikowego. Podobne rozważania przeprowadzono dla planu eksperymentu obejmującego 50 i 75 doświadczeń. Wyniki przedstawia tabela 1. Oczywiście najbardziej satysfakcjonujące wartości wariancji estymatora funkcji powierzchni odpowiedzi otrzymano dla całkowitego planu eksperymentu czynnikowego. Natomiast wartości wariancji dla częściowych planów eksperymentu czynnikowego w większości przypadków są mniejsze od wartości wariancji wyznaczonej dla planów doświadczeń skonstruowanych przez losowy wybór punktów obszaru eksperymentowania.

Podsumowanie

W praktyce klasyczne plany eksperymentów czynnikowych często wymagają przeprowadzenia licznych doświadczeń oraz nie uwzględniają właściwości struktury badanego materiału. Realizacja eksperymentu zgodnie z częściowym planem eksperymentu czynnikowego wykorzystującym modele pól losowych pozwala na uwzględnienie losowego charakteru struktury materiałów, co istotnie wpływa na wiarygodność otrzymanych wyników. Proponowany częściowy plan eksperymentu czynnikowego prowadzi również do ograniczenia kosztów realizacji eksperymentu, stanowiąc alternatywę dla klasycznych ułamkowych planów eksperymentów czynnikowych.

Literatura

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

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

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

Szczepankiewicz E. (1985): *Zastosowania pól losowych*. Państwowe Wydawnictwo Naukowe, Warszawa.

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

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

ON THE FACTORIAL DESIGN OF EXPERIMENT USING RANDOM FIELDS MODELS

Summary

A tool of statistical quality control, design of experiments carried out on the basis of established assumptions leads to some improvements in the quality of manufacturing processes and has a significant influence on their economical results. The aim of this paper is to present the design of experiment using models defined by random fields. The presented factorial design of experiment allow one to determine appropriate points of experimental area when the application of classical design of experiment is improper. Moreover, the proposed experimental design leads to limitation of realization costs of the experiment and is an alternative for the fractional design of experiments.