

A K A D E M I A E K O N O M I C Z N A

im. Oskara Langego we Wrocławiu

Jerzy Wawrzynek

OPTYMALNE PLANY EKSPERYMENTÓW
DLA ESTYMACJI LINIOWEJ FUNKCJI REGRESJI
NA KOSTCE WIELOWYMIAROWEJ

Praca doktorska

Promotor: Prof.dr hab. Zdzisław Hellwig

W R O C Ł A W 1 9 7 4

S p i s t r e ś c i

Wstęp	1
I Planowanie eksperymentów dla zagadnień regresji	10
II Regresja liniowa na odcinku prostej	30
III Regresja liniowa na kostce dwuwymiarowej (kwadracie)	48
IV Regresja liniowa na kostce m-wymiarowej	72
V Regresja liniowa na prostopadłościu m-wymia- rowym	91
VI Komputerowe badanie efektywności planowania eksperymentów	97
Uwagi końcowe	114
Tabulogram programu obliczeń	117
Tabulogram wyników	123
Literatura cytowana	151

W S T Ę P

W wielu gałęziach nauki, w tym także i w naukach ekonomicznych, występują problemy, których zbadanie wymaga przeprowadzenia obserwacji lub bardziej złożonych eksperymentów. Pomiar dowolnej wielkości eksperymentalnej odbywa się zawsze w obecności pewnych czynników ubocznych, które - niezależnie od wysiłków badacza mających na celu minimizację ich wpływu - nigdy nie mogą być całkowicie wyeliminowane. Oznacza to, że badacz ma do czynienia z wielkościami losowymi, a nie z wielkościami deterministycznymi. Stąd też dla analizy wyników eksperymentu należy zastosować metody statystyki matematycznej i rzeczywiście odpowiednie metody statystyczne są od dawna wykorzystywane w praktyce opracowywania wyników doświadczeń.

Przez długi okres czasu w statystyce matematycznej zwracano uwagę na usprawnianie metod opracowania wyników eksperymentu, przeprowadzonego według pewnego określonego planu. Wybór planu eksperymentu, tzn. określenie jak, gdzie i kiedy przeprowadzić pomiary lub obserwacje, pozostawiano zasadniczo intuicji badacza. Jeśli problem badawczy jest stosunkowo prosty zarówno z punktu widzenia teoretycznego, jak i z punktu widzenia realizacji eksperymentu i przy tym eksperyment nie wymaga znacznych nakładów (środków pieniężnych, czasu, określonych zasobów materiałów), to dokonanie wyboru planu eksperymentu w oparciu o intuicję badacza nie prowadzi na ogół do istotnych strat.

Rozwój nauki i techniki doprowadził do znacznej złożoności

w teoretycznej interpretacji obserwowanych wyników i w realizacji badań eksperymentalnych. Złożoność w realizacji badań eksperymentalnych wyraziła się w poważnym wzroście kosztów eksperymentu (np. konieczność zastosowania niezwykle drogich i skomplikowanych instalacji doświadczalnych, użycia kosztownych substancji chemicznych, wydatkowania wielkich ilości energii), czasu trwania eksperymentu oraz zużycia materiałów, surowców i energii. W tej sytuacji niezwykle aktualne staje się zagadnienie wyciągnięcia możliwie najtrafniejszych wniosków o badanym procesie przy ograniczonych nakładach (środkach pieniężnych, czasie, zasobach materiałów i energii) przeznaczonych na jego badanie. Próby rozwiązania tego zagadnienia w oparciu o intuicję eksperymentatora budzą coraz mniej nadziei. W związku z tym absolutną koniecznością staje się zastosowanie metod, które podają nie tylko sposób opracowania wyników eksperymentu, ale pozwalają na optymalne zorganizowanie, planowanie samego eksperymentu.

Prace w dziedzinie planowania eksperymentów zaczęły się rozwijać w latach trzydziestych w odniesieniu do doświadczeń, których wyniki opracowywano metodami analizy wariancji. Zadanie dokładnego określenia efektów poszczególnych czynników i ich interakcji powoduje tam konieczność zaplanowania doświadczenia odpowiednio zrandomizowanego (por. np. [11], [12], [37]). W latach pięćdziesiątych prace w dziedzinie planowania eksperymentów objęły zagadnienia regresji, tzn. zagadnienia, w których wynik eksperymentu traktuje się jako funkcję m zmiennych rzeczywistych nazywanych zmiennymi kontrolowanymi.

Metody matematyczne, wykorzystywane w teorii planowania eksperymentu, stanowią kompozycję metod statystyki matematycz-

nej i metod rozwiązywania zagadnień ekstremalnych. Przy tym metody statystyczne służą dla sformułowania kryteriów optymalności eksperymentu oraz interpretacji zasadniczych własności tych kryteriów, natomiast w wyniku sformułowania odpowiedniego kryterium optymalności zagadnienie optymalnej organizacji eksperymentu sprowadza się do rozwiązania pewnego zadania ekstremalnego.

Optymalna organizacja planu eksperymentu dla zagadnień regresji może być dokonana zasadniczo na dwa sposoby:

1^o Jeśli ustalona jest wielkość każdego rodzaju nakładów (np. środków pieniężnych, czasu, zasobów materiałów i energii) przeznaczonych na realizację eksperymentu i znane są funkcje zużycia każdego z tych rodzajów nakładów w zależności od liczby obserwacji (pomiarów, badań) to można wyznaczyć liczbę N taką, że realizacja N obserwacji jest możliwa przy zadanych nakładach, natomiast zwiększenie liczby obserwacji ponad wyznaczoną liczbę prowadziłyby do przekroczenia zadanej wielkości co najmniej jednego rodzaju nakładów. Po wyznaczeniu w ten sposób liczby N obserwacji następuje ostateczne ustalenie planu eksperymentu poprzez określenie - w wyniku rozwiązania pewnego zadania ekstremalnego związanego z odpowiednim kryterium optymalności eksperymentu - wartości zmiennych kontrolowanych dla każdej obserwacji. Postępowanie takie prowadzi do wyznaczenia najlepszego - z punktu widzenia danego kryterium optymalności - planu eksperymentu, który można zrealizować przy nakładach ustalonych z góry.

2^o Jeśli z określonym kryterium optymalności eksperymentu związane pewien dodatkowy warunek (np. dotyczący liczbowej

wartości tego kryterium), to spośród planów eksperymentów spełniających ten warunek wybrać można plan o minimalnej liczbie obserwacji. Przy znajomości funkcji zużycia każdego z rodzajów nakładów w zależności od liczby obserwacji można wyznaczyć minimalną wielkość każdego rodzaju nakładów, potrzebnych do realizacji obserwacji. Postępowanie takie prowadzi do określenia niezbędnych wielkości każdego rodzaju nakładów potrzebnych dla realizacji eksperymentu, który ma posiadać pewne z góry określone własności. Pozwala to uniknąć zgromadzenia niewłaściwych (zbyt małych lub zbyt wysokich) nakładów w trakcie przygotowywania eksperymentu i zabezpiecza zarówno przed przerwaniem eksperymentu z powodu braku środków jak i przed zamrożeniem lub marnotrawstwem nadmiaru środków.

W praktyce łączy się obydwie opisane sposoby i dokonuje się bilansowania wielkości nakładów przeznaczonych na eksperyment z własnościami tegoż eksperymentu. W dalszym ciągu pracy zajmować będziemy się pewnymi aspektami przedstawionych sposobów planowania eksperymentu dla zagadnień regresji, a mianowicie zajmujemy się wyznaczaniem optymalnych wartości zmiennych kontrolowanych przy ustalonej liczbie obserwacji oraz wyznaczaniem minimalnej liczby obserwacji potrzebnych do otrzymania planu eksperymentalnego o określonych własnościach.

Charakterystyka treści rozdziałów I - VI.

Rozdział I przedstawia te pojęcia i metody z dziedziny planowania eksperymentów dla zagadnień regresji, które będą wykorzystywane w dalszych rozdziałach pracy. W szczególności zdefiniowano pojęcie planu eksperymentalnego P_N dla N obserwacji oraz scharakteryzowano zagadnienie estymacji współ-

czynników funkcji regresji i estymacji całej funkcji regresji metodą najmniejszych kwadratów. Ponieważ rozkład estymatorów otrzymanych metodą najmniejszych kwadratów zależy m.in. od planu eksperymentalnego, zgodnie z którym wykonywano obserwacje funkcji regresji, więc powstaje problem wyboru optymalnego planu eksperymentalnego. Sformułowano cztery kryteria optymalności planów eksperymentalnych, a mianowicie kryteria A - optymalności (kryterium 1), D - optymalności (kryterium 2), G - optymalności (kryterium 3) oraz I - optymalności (kryterium 4). Z kolei naszkicowano teorię planów przybliżonych i zdefiniowano pojęcie planu przybliżonego, które stanowi uogólnienie pojęcia planu eksperymentalnego. W ramach teorii planów przybliżonych sformułowano cztery zadania będące uogólnieniem odpowiednich kryteriów A -, D -, G - oraz I - optymalności. Związki między tymi zadaniami przedstawiono w postaci twierdzeń 2 i 3. Rozdział kończy się krótkim przeglądem wyników teorii planów przybliżonych.

W rozdziałach II - V niniejszej pracy konstruuje się pewne plany eksperymentalne dla estymacji liniowej funkcji regresji jednej lub wielu zmiennych. Dla najważniejszych z tych planów dowodzi się twierdzeń orzekających, że są to plany spełniające kryteria 2 i 3, a często również kryterium 1, a więc plany równocześnie D - i G - optymalne, a często A - optymalne. Teorię planów przybliżonych wykorzystuje się przy dowodzeniu D - i G - optymalności w ten sposób, że pokazuje się, iż dany plan eksperymentalny dla N obserwacji spełnia warunek (c) twierdzenia 2 (Kiefera i Wolfowitza). Oznacza to, że dany plan eksperymentalny dla N obserwacji spełnia warunek dostateczny D - i G - optymalności w sensie zadań 2 i 3 w szerszej klasie

planów przybliżonych, a więc w sensie kryteriów 2 i 3 jest to plan D - i G - optymalny również w klasie planów eksperymentalnych dla N obserwacji. Analogicznie, przez sprawdzenie równości (1.10) z twierdzenia 3, wykorzystuje się teorię planów przybliżonych dla dowodu A - optymalności planu eksperymentalnego.

W rozdziale II dla liniowej funkcji regresji jednej zmiennej obserwowanej i estymowanej na odcinku $[a, b]$ prostej rzeczywistej przedstawia się G - optymalne plany eksperymentalne dla dowolnej liczby N obserwacji ($N \geq 2$). Jeśli liczba N jest parzysta, to plany te są równocześnie planami D - optymalnymi. Pokazuje się (patrz twierdzenie 5), że przy użyciu planów G - optymalnych: 1^o suma wariancji estymatorów współczynników funkcji regresji maleje ze wzrostem N ; 2^o pole elipsy koncentracji rozkładu estymatorów współczynników funkcji regresji maleje ze wzrostem N ; 3^o maksymalna wartość wariancji estymatora funkcji regresji liniowej maleje ze wzrostem N ; 4^o średnia wartość wariancji estymatora funkcji regresji liniowej maleje ze wzrostem N . Rozdział kończy się porównaniem efektywności rozmieszczania obserwacji według planu G - optymalnego z efektywnością rozmieszczania obserwacji w sposób losowy.

W rozdziale III dla liniowej funkcji regresji dwu zmiennych, obserwowanej i estymowanej na kwadracie $X = \{x = (x_1, x_2), -1 \leq x_i \leq 1, i = 1, 2\}$, przedstawiono 4 rodziny planów eksperymentalnych dla $N = 4k, 4k+1, 4k+2, 4k+3$ obserwacji ($k = 1, 2, \dots$). Zgodnie z tymi planami obserwacje wykonuje się jedynie w punktach wierzchołkowych kwadratu X . Ro-

dzina planów dla $N = 4k$ obserwacji jest rodziną planów A -, D - i G - optymalnych. Poza tym, po wykonaniu N obserwacji według planu należącego do jednej z przedstawionych czterech rodzin, istnieje możliwość wykonania dodatkowych d obserwacji w taki sposób, jak gdyby $N+d$ obserwacji wykonano znów według planu należącego do jednej z tych czterech rodzin. Pokazuje się (patrz twierdzenie 6), że przy wykorzystaniu tych planów : 1^o suma wariancji estymatorów współczynników funkcji regresji maleje ze wzrostem N ; 2^o objętość elipsoidy koncentracji rozkładu estymatorów współczynników funkcji regresji maleje ze wzrostem N ; 3^o maksymalna wartość wariancji estymatora funkcji regresji maleje ze wzrostem N ; 4^o średnia wartość wariancji estymatora funkcji regresji maleje ze wzrostem N . Omówione plany porównuje się - ze względu na kryteria A -, D -, G- oraz I optymalności - z odpowiednimi spośród pewnych planów eksperymentalnych, zgodnie z którymi obserwacje wykonuje się w punktach wierzchołkowych oraz w punkcie środkowym kwadratu X . Porównanie to wypada generalnie na niekorzyść ostatnich planów. Ze względu na kryterium G - optymalności porównano również efektywność rozmieszczania obserwacji zgodnie z planami wprowadzonymi na początku rozdziału z efektywnością rozmieszczania obserwacji w sposób losowy.

W rozdziale IV konstruuje się plany A -, D - i G - optymalne dla liniowej funkcji regresji m zmiennych, obserwowanej i estymowanej na kostce m -wymiarowej. Przedstawia się optymalne plany eksperymentalne kolejno dla liczby N obserwacji równej:

1° $n \cdot 2^m$ gdzie $n = 1, 2, \dots$ (twierdzenie 7)

2° $n \cdot (2^{m-p_1} + 2^{m-p_2} + \dots + 2^{m-p_r})$ gdzie $n = 1, 2, \dots$;

$2^{m-p_i} > m$; $i = 1, 2, \dots, r$ (twierdzenie 8)

3° $n \cdot (N_m + 4t)$ gdzie $n = 1, 2, \dots$; $N_m = 4 + m - m(\text{mod}4)$;

$t = 0, 1, \dots$; $2 \leq m \leq 91$ (twierdzenie 10).

Twierdzenie 8 jest rozszerzeniem twierdzenia 7, a twierdzenie 10 rozszerzeniem twierdzeń 7 i 8 polegającym na tym, że każde następne z nich dołącza nowe elementy do zbioru liczb obserwacji, dla których twierdzenie poprzednie podaje sposób optymalnego zaplanowania eksperymentu. Np. gdy funkcja regresji zależy od $m = 5$ zmiennych kontrolowanych, to twierdzenie 7 określa plany eksperymentalne dla $N = 32, 64, 96, 128, \dots$ obserwacji, twierdzenie 8 określa optymalne plany eksperymentalne dla $N = 8, 16, 24, \underline{32}, 40, 48, 56, \underline{64}, 72, \dots$ obserwacji, natomiast twierdzenie 10 określa optymalne plany eksperymentalne dla $N = \underline{8}, 12, \underline{16}, 20, \underline{24}, 28, \underline{32}, 36, \underline{40}, 44, \dots$ obserwacji.

Rozdział V przedstawia konstrukcję D - i G - optymalnych planów eksperymentalnych dla estymacji liniowej funkcji regresji m zmiennych na zbiorze $X = \{x = (x_1, \dots, x_m) : a_i \leq x_i \leq b_i, i = 1, \dots, m\}$. Dla tego ważnego w praktyce przypadku dowodzi się, że optymalne plany eksperymentalne można uzyskać przez przekształcenie D - i G - optymalnych planów eksperymentalnych na kostce m -wymiarowej.

Rozdział VI przedstawia wyniki komputerowego badania efektywności eksperymentów ze względu na kryterium G - optymalności. Otrzymane wyniki pozwalają na ilościową ocenę niektórych aspektów strat wynikających z mniejszej efektywności eksperymentów realizowanych w sposób nieoptymalny.

Rozdział I

PLANOWANIE EKSPERYMENTÓW DLA ZAGADNIEŃ REGRESJI

Niech w m -wymiarowej przestrzeni euklidesowej \mathcal{E}^m dany będzie podzbiór domknięty X . Dla każdego punktu $x = (x_1, x_2, \dots, x_m)$ należącego do zbioru X niech będzie określona zmienna losowa Y_x o wartości oczekiwanej

$$(1.1) \quad EY_x = \eta(x, \theta),$$

gdzie $\theta = \begin{bmatrix} \theta_1 \\ \cdot \\ \cdot \\ \cdot \\ \theta_k \end{bmatrix}$ jest wektorem nieznanym parametrów.

Równanie (1.1) nazywa się funkcją regresji zmiennych losowych Y_x względem X , często też równanie to bywa nazywane powierzchnią odpowiedzi.

Niech podzbiorem (właściwym lub niewłaściwym) zbioru X będzie zbiór domknięty X_1 . Zakłada się, że dokonywanie obserwacji zmiennych losowych Y_x jest możliwe wyłącznie na podzbiorku X_1 . Podzbiór X_1 nazywa się obszarem obserwacji albo obszarem działania. Określenie obszaru działania odgrywa istotną rolę w planowaniu eksperymentu. W pewnych przypad-

kach obszar ten jest określony jednoznacznie przez charakter zmiennych kontrolowanych x_1, \dots, x_m . Np. jeśli zmienne kontrolowane x_1, \dots, x_m oznaczają składniki mieszaniny chemicznej wyrażone w procentach, to procent określonego składnika x_1 nie może być mniejszy niż 0 i nie może być większy niż 100, a obszar X_1 staje się w tym przypadku $(m-1)$ -wymiarowym simpleksem położonym w \mathcal{E}^m . Podobnie zmienna kontrolowana oznaczająca średnicę kółka kulkowego nie może mieć wartości ujemnej, etc. Z kolei w innych przypadkach obszar działania wyznaczony jest przez charakterystyki aparatury eksperymentalnej albo przez charakter badanego procesu. Np. górna granica temperatury może zależeć od wydajności źródła ciepła albo od własności materiałów izolacyjnych; podobnie górna granica prędkości cząstek elementarnych zależy od parametrów akcelatora.

Eksperymentator obserwujący w zbiorze X_1 wartości y_x zmiennych losowych Y_x ma zwykle do czynienia z jedną z trzech sytuacji, które opiszemy poniżej.

Pierwsza z możliwych sytuacji polega na tym, że postać analityczna funkcji $\eta(x, \theta)$ określającej wartość oczekiwaną zmiennej losowej Y_x w zależności od wartości zmiennych kontrolowanych jest nieznaną. W tej sytuacji zadanie eksperymentatora polega na dostatecznie dobrej aproksymacji funkcji $\eta(x, \theta)$ przez funkcje należące do pewnej klasy danych funkcji oraz na estymacji wektora parametrów θ .

Druga z możliwych sytuacji polega na tym, że postać analityczna funkcji $\eta(x, \theta)$ jest identyczna z jedną

spośród funkcji $\eta_1(x, \theta^{(1)})$, $\eta_2(x, \theta^{(2)})$, ... ,
 $\eta_s(x, \theta^{(s)})$, których postać analityczna jest znana. Inny-
mi słowy, opisywana sytuacja polega na tym, że danych jest
s modeli powierzchni odpowiedzi. Zadanie eksperymentatora
polega tu na wyborze funkcji identycznej z funkcją $\eta(x, \theta)$
oraz na wyestymowaniu właściwego dla tej funkcji wektora θ .
Informacje na temat planowania doświadczeń w opisanej sytua-
cji można znaleźć w pracach Bandemera i współautorów [4],
Boxa [5], [6], Fiedorowa [36], str. 209-263, [9], rozdz.
5 i 6, [8], [10], Hilla i Huntera [14], [15], [21] oraz
innych autorów.

Trzecia z możliwych sytuacji polega na tym, że postać
analityczna funkcji $\eta(x, \theta)$ jest znana, zaś zadanie eks-
perymentatora polega na tym, aby na podstawie wykonanych ob-
serwacji wyestymować nieznaną wektor θ , albo pewną funkcję
parametrów $\theta_1, \dots, \theta_k$. W pracy niniejszej omawia się pla-
nowanie eksperymentów przy założeniach, które wyznaczają
przypadek szczególny takiej właśnie sytuacji.

W dalszym ciągu będziemy zatem zakładać, że dla każdego
 $x \in X$ określona jest zmienna losowa Y_x o nieznannej, jed-
nakowej dla każdego $x \in X$, wariancji σ^2 , natomiast wartość
oczekiwana zmiennej losowej Y_x dana jest równaniem

$$(1.2) \quad EY_x = \theta_1 f_1(x) + \dots + \theta_k f_k(x),$$

gdzie $\theta = \begin{bmatrix} \theta_1 \\ \cdot \\ \cdot \\ \cdot \\ \theta_k \end{bmatrix}$ jest wektorem nieznanym parametrów, nato-

miast

$$f(x) = \begin{bmatrix} f_1(x) \\ \vdots \\ f_k(x) \end{bmatrix} \quad \text{oznacza wektor, którego składowymi jest } k$$

znanych, liniowo niezależnych ciągłych funkcji rzeczywistych f_1, \dots, f_k określonych na zbiorze domkniętym $X \subset \mathcal{E}^m$, tzn. każda z nich jest funkcją m zmiennych niezależnych x_1, \dots, x_m .

Jednym z możliwych zadań jest estymacja wektora nieznanych współczynników $\theta_1, \dots, \theta_k$ oraz całej funkcji regresji (1.2). Zadanie to można zrealizować obserwując wartości zmiennych losowych Y_x w wybranych punktach $x^{(1)}, x^{(2)}, \dots, x^{(p)}$ zbioru X_1 . Zakładamy, że łącznie wykonuje się N obserwacji i że w punktach $x^{(1)}, x^{(2)}, \dots, x^{(p)}$ wykonuje się odpowiednio n_1, n_2, \dots, n_p obserwacji ($\sum_{i=1}^p n_i = N$), przy czym obserwacje te są wzajemnie nieskorelowane (dotyczy to również każdych dwu obserwacji wykonanych w tym samym punkcie x zbioru X_1). Ustalenie łącznej liczby obserwacji, wybór punktów zbioru X_1 , w których obserwacje te będzie się wykonywać oraz przydzielenie odpowiedniej liczby obserwacji wybranym punktom zbioru X_1 oznacza zaplanowanie eksperymentu.

DEFINICJA. Planem eksperymentalnym P_N dla N obserwacji nazywamy zbiór wartości liczbowych

$$\left\{ \begin{array}{l} x^{(1)}, x^{(2)}, \dots, x^{(p)} \\ \lambda_1, \lambda_2, \dots, \lambda_p \end{array} \right\}$$

gdzie $x^{(l)} \in X_1, l = 1, 2, \dots, p$,

$$\lambda_1 = n_1/N, \quad N = \sum_{l=1}^p n_l,$$

zaś n_1 jest liczbą całkowitą określającą liczbę obserwacji, które zostaną wykonane w punkcie $x^{(1)}$ ($l = 1, 2, \dots, p$).

Dla planu eksperymentalnego P_N niech $M(P_N)$ oznacza macierz kwadratową o elementach

$$(1.3) \quad m_{ij} = \sum_{l=1}^p \lambda_l f_i(x^{(l)}) f_j(x^{(l)}) \quad (i, j = 1, 2, \dots, k)$$

skąd

$$(1.4) \quad M(P_N) = \sum_{l=1}^p \lambda_l f(x^{(l)}) f^T(x^{(l)}) .$$

W dalszym ciągu rozważamy tylko takie plany eksperymentalne, dla których $\det M(P_N) \neq 0$. Możemy wtedy zdefiniować funkcję

$$d(x, P_N) = f^T(x) M^{-1}(P_N) f(x) .$$

Na podstawie N obserwacji $y_{11}, \dots, y_{1n_1}, y_{21}, \dots, y_{pn_p}$ zmiennych losowych $Y_x(1), Y_x(2), \dots, Y_x(p)$, wykonywanych według planu eksperymentalnego P_N można wyestymować wektor θ .

Do estymacji wektora $\theta = \begin{bmatrix} \theta_1 \\ \vdots \\ \theta_k \end{bmatrix}$ będziemy stosowali metodę naj-

mniej kwadratów. Estymator uzyskany tą metodą oznaczmy

przez
$$\hat{\theta} = \begin{bmatrix} \hat{\theta}_1 \\ \vdots \\ \hat{\theta}_k \end{bmatrix} .$$

Natychmiast otrzymuje się stąd estymator $\hat{E}Y_x$ wartości oczekiwanej zmiennej losowej Y_x określony następująco:

$$\hat{E}Y_x = \hat{\theta}_1 f_1(x) + \dots + \hat{\theta}_k f_k(x) .$$

Z teorii estymacji metodą najmniejszych kwadratów wiado-

mo, że :

1° estymator $\hat{\theta}_1$ współczynnika θ_1 ($i = 1, \dots, k$) ma minimalną wariancję w klasie estymatorów liniowych i nieobciążonych;

2° dla każdego ustalonego $x \in X$ estymator $\hat{E}Y_x$ funkcji regresji (1.2) ma minimalną wariancję w klasie estymatorów liniowych i nieobciążonych;

3° macierz wariancji i kowariancji estymatorów $\hat{\theta}_1, \dots, \hat{\theta}_k$

$$(1.5) \quad E(\hat{\theta} - \theta)(\hat{\theta} - \theta)^T = \frac{\sigma^2}{N} M^{-1}(P_N) ;$$

4° objętość k -wymiarowej elipsoidy koncentracji rozkładu estymatorów $\hat{\theta}_1, \dots, \hat{\theta}_k$ wyraża się wzorem

$$(1.6) \quad \frac{\sigma^k}{\Gamma(\frac{1}{2}k + 1)} \sqrt{\frac{(k + 2)^k \cdot \pi^k}{N^k \cdot \det M(P_N)}} ;$$

5° wariancja estymatora $\hat{E}Y_x$

$$(1.7) \quad \text{Var}(\hat{E}Y_x) = \frac{\sigma^2}{N} f^T(x) M^{-1}(P_N) f(x) = \frac{\sigma^2}{N} d(x, P_N).$$

Dla danego zbioru X_1 istnieje na ogół cała rodzina \mathcal{D}_N planów eksperymentalnych dla N obserwacji. Powstaje więc tu zasadniczy problem wyboru najlepszego lub przynajmniej dobrego planu eksperymentalnego. Wybór ten zależy bezpośrednio od przyjętego kryterium optymalności planu eksperymentalnego. W literaturze na temat planowania eksperymentów dla zagadnień regresji operuje się najczęściej następującymi czterema kryteriami:

KRYTERIUM 1. Plan eksperymentalny $P_{N,A} \in \mathcal{D}_N$ nazywamy planem A - optymalnym jeśli

$$\sum_{i=1}^k m^{ii}(P_{N,A}) = \min_{P_N} \sum_{i=1}^k m^{ii}(P_N),$$

gdzie $m^{ii}(P_N)$ ($i = 1, \dots, k$) oznacza i -ty element głównej przekątnej macierzy $M^{-1}(P_N)$.

KRYTERIUM 2. Plan eksperymentalny $P_{N,D} \in \mathcal{D}_N$ nazywamy planem D- optymalnym jeśli

$$\det M(P_{N,D}) = \max_{P_N \in \mathcal{D}_N} \det M(P_N).$$

KRYTERIUM 3. Plan eksperymentalny $P_{N,G} \in \mathcal{D}_N$ nazywamy planem G - optymalnym jeśli

$$\max_{x \in X_2} d(x, P_{N,G}) = \min_{P_N \in \mathcal{D}_N} \max_{x \in X_2} d(x, P_N),$$

gdzie X_2 jest zbiorem domkniętym zawartym w zbiorze X .

KRYTERIUM 4. Plan eksperymentalny $P_{N,I} \in \mathcal{D}_N$ nazywamy planem I - optymalnym jeśli

$$\int_{X_2} d(x, P_{N,I}) dx = \min_{P_N \in \mathcal{D}_N} \int_{X_2} d(x, P_N) dx$$

Z porównania kryterium 1 ze wzorem (1.5), kryterium 2 ze wzorem (1.6), a kryteriów 3 i 4 ze wzorem (1.7) wynika, że w klasie planów eksperymentalnych dla N obserwacji

- 1° plan A - optymalny minimalizuje sumę (lub średnią arytmetyczną) $\text{Var } \hat{\theta}_1 + \text{Var } \hat{\theta}_2 + \text{Var } \hat{\theta}_3 + \dots + \text{Var } \hat{\theta}_k$ wariancji estymatorów $\hat{\theta}_1, \hat{\theta}_2, \dots, \hat{\theta}_k$;
- 2° plan D - optymalny minimalizuje objętość elipsoidy koncentracji rozkładu estymatorów $\hat{\theta}_1, \hat{\theta}_2, \dots, \hat{\theta}_k$;
- 3° plan G - optymalny minimalizuje maksymalną na zbiorze X_2 wartość wariancji estymatora $\hat{E}Y_x$ funkcji regresji (1.2);

4^o plan I - optymalny minimalizuje średnią na zbiorze X_2 wartość wariancji estymatora $\hat{E}Y_x$ funkcji regresji (1.2).

Zbiór X_2 występujący w kryteriach 3 i 4 może być podzbiorem zbioru X_1 , odwrotnie - X_1 może być podzbiorem X_2 , wreszcie iloczyn zbiorów X_1 i X_2 może być podzbiorem właściwym każdego ze zbiorów X_1 i X_2 lub zbiorem pustym, w końcu może zachodzić relacja $X_1 = X_2$. Np. niech X oznacza odcinek czasu $[0, x_t]$, zbiór X_1 , na którym można wykonywać obserwacje, niech będzie odcinkiem czasu $[0, x_1]$, gdzie x_1 oznacza moment, w którym należy zakończyć obserwacje. Szczególnie ważna może być estymacja (ekstrapolacja) funkcji regresji na zbiorze $X_2 = [x_2, x_t]$, gdzie $x_1 < x_2 < x_t$. Tutaj $X_1 \cap X_2 = \emptyset$.

Przykład. Niech zbiór X oznacza prostą rzeczywistą ξ^1 i niech funkcja regresji będzie wielomianem stopnia trzeciego

$$EY_x = \theta_0 + \theta_1 x + \theta_2 x^2 + \theta_3 x^3 .$$

Przypuśćmy, że dla wyestymowania wartości funkcji EY_x w punkcie $x = 2$ można wykonać $N = 52$ obserwacje na odcinku $X_1 = [-1, 1]$. Wiadomo również, że $\sigma = 1$. Zadanie to można zrealizować tradycyjnie w ten sposób, że wykonuje się po 13 obserwacji w każdym z czterech następujących punktów : $x^{(1)} = -1$, $x^{(2)} = -1/3$, $x^{(3)} = 1/3$, $x^{(4)} = 1$, które rozłożone są równomiernie na odcinku $[-1, 1]$. Przy takim wyborze planu eksperymentu okazuje się, że wariancja estymatora funkcji regresji w punkcie $x = 2$ jest w przybliżeniu równa 20. Można jednak wybrać taki plan eksperymentu, przy którym wariancja estymatora funkcji regresji w punkcie $x = 2$ będzie minimalna. Innymi słowy, należy wybrać plan G - optymalny

dla przypadku jednopunktowego zbioru $X_2 = \{2\}$. Można pokazać (por. [20]), że plan G - optymalny polega tu na rozmieszczeniu 52 obserwacji w ten sposób, że w punkcie $x^{(1)} = -1$ wykonuje się 5 obserwacji, w punkcie $x^{(2)} = -1/2$ wykonuje się 12 obserwacji, w punkcie $x^{(3)} = 1/2$ wykonuje się 20 obserwacji i w punkcie $x^{(4)} = 1$ wykonuje się 15 obserwacji. Dla tego planu G - optymalnego wariancja estymatora funkcji regresji w punkcie $x = 2$ jest w przybliżeniu równa 13.

Teoria planów przybliżonych. Zdefiniowaliśmy uprzednio plan eksperymentalny dla N obserwacji jako zbiór wartości liczbowych

$$\left\{ \begin{array}{l} x^{(1)}, x^{(2)}, \dots, x^{(p)} \\ \lambda_1, \lambda_2, \dots, \lambda_p \end{array} \right\},$$

gdzie $x^{(1)}$ ($1 = 1, 2, \dots, p$) są wybranymi punktami zbioru X_1 , λ_1 są nieujemnymi ułamkami wymiernymi o wspólnym mianowniku równym N, zaś $\sum_{l=1}^p \lambda_l = 1$. Pod koniec lat pięćdziesiątych w teorii planowania eksperymentów pojawiło się, wprowadzone przez Kiefera i Wolfowitza ([30], [27]), pojęcie planu przybliżonego. Pod pojęciem planu przybliżonego rozumie się dowolną miarę probabilistyczną $\bar{\xi}$ określoną na polu borelowskim \mathcal{B} generowanym przez zbiory otwarte przestrzeni X_1 . Miara taka spełnia warunek $\int_{X_1} \bar{\xi}(dx) = 1$.

Pojęcie planu przybliżonego jest uogólnieniem pojęcia planu eksperymentalnego w tym sensie, że plan eksperymentalny

$$\left\{ \begin{array}{l} x^{(1)}, x^{(2)}, \dots, x^{(p)} \\ \lambda_1, \lambda_2, \dots, \lambda_p \end{array} \right\}$$

można potraktować jako empiry-

czny rozkład prawdopodobieństwa na zbiorze X_1 , który wyzna-

cza na tym zbiorze specjalną dyskretną miarę probabilistyczną.

Niech \mathfrak{D} oznacza rodzinę wszystkich miar probabilistycznych na X_1 . Uogólniając wzór (1.3) można dla każdej miary probabilistycznej $\xi \in \mathfrak{D}$ określić macierz kwadratową $M(\xi)$ o elementach

$$(1.8) \quad m_{ij} = \int_{X_1} f_i(x) f_j(x) \xi(dx) \quad i, j = 1, \dots, k$$

Własności macierzy $M(\xi)$ podaje następujące

TWIERDZENIE 1.

(a) Dla każdego $\xi \in \mathfrak{D}$ macierz $M(\xi)$ jest określona nieujemnie.

(b) Jeśli miara ξ jest skoncentrowana dokładnie w p punktach oraz $p < k$, to $\det M(\xi) = 0$.

(c) Rodzina macierzy $M(\xi)$, $\xi \in \mathfrak{D}$, tworzy zbiór wypukły i zwarty.

(d) Dla każdego $\xi \in \mathfrak{D}$ macierz $M(\xi)$ można napisać w postaci (1.4), gdzie $p \leq \frac{k(k+1)}{2} + 1$.

Dowód twierdzenia 1 zob. np. [25], str. 324.

Podobnie jak dla planów eksperymentalnych P_N zdefiniujemy dla każdego planu przybliżonego ξ takiego, że $\det M(\xi) \neq 0$, funkcję

$$(1.9) \quad d(x, \xi) = f^T(x) M^{-1}(\xi) f(x).$$

Sformułujemy teraz cztery zadania ekstremalne:

ZADANIE 1. Wyznaczyć te miary $\xi \in \mathfrak{D}$, które minimalizują

$\sum_{i=1}^k m^{ii}(\xi)$, gdzie $m^{ii}(\xi)$ oznaczają i -ty ($i = 1, \dots, k$) element głównej przekątnej macierzy $M^{-1}(\xi)$.

ZADANIE 2 . Wyznaczyć te miary $\xi \in \mathcal{D}$, które maksymalizują $\det M(\xi)$.

ZADANIE 3 . Wyznaczyć te miary $\xi \in \mathcal{D}$, dla których $\sup_{x \in X_2} d(x, \xi) = \min$.

ZADANIE 4 . Wyznaczyć te miary $\xi \in \mathcal{D}$, dla których $\int_{X_2} d(x, \xi) dx = \min$.

W ważnym przypadku $X_1 = X_2$ rozwiązania zadań 2 i 3 pokrywają się. Zachodzi bowiem

TWIERDZENIE 2 (Kiefera i Wolfowitza) . W przypadku $X = X_1 = X_2$ warunki

(a) ξ^* maksymalizuje $\det M(\xi)$,

(b) ξ^* minimalizuje $\sup_{x \in X_2} d(x, \xi)$

(c) $\sup_{x \in X_2} d(x, \xi) = k$

są równoważne . Zbiór B złożony ze wszystkich miar ξ spełniających warunki (a) , (b) lub (c) jest wypukły i domknięty, zaś macierz $M(\xi)$ jest jednakowa dla wszystkich $\xi \in B$.

Dowód twierdzenia 2 zob. [31] .

Kolejne twierdzenie dotyczy związków zadania 1 z zadaniem 2.

TWIERDZENIE 3. W klasie \mathcal{D} wszystkich miar probabilistycznych na zbiorze X_1 miara ξ^* maksymalizuje $\det M(\xi)$ oraz równocześnie minimalizuje $\sum_{i=1}^k m^{ii}(\xi)$, jeśli

$$(1.10) \quad cM^{-1}(\xi^*) = [M^{-1}(\xi^*)]^2 ,$$

gdzie c jest pewną stałą. Przy tym $\sum_{i=1}^k m^{ii}(\xi^*) = ck$.

Dowód twierdzenia 3 zob. [9], § 2.11 .

Wprowadzenie pojęcia planu przybliżonego i sformułowanie wyżej wymienionych zadań ekstremalnych zaowocowało w postaci szeregu ważnych wyników teoretycznych osiągniętych przy rozwiązywaniu tych zadań (a zwłaszcza zadań 2 i 3) dla różnych specjalnych postaci wektora funkcji $f(x)$ oraz zbiorów X_1 i X_2 . Przedstawimy teraz niektóre z tych rezultatów dla funkcji regresji, która jest wielomianem jednej lub wielu zmiennych.

Pierwsze, historycznie rzecz biorąc, wyniki osiągnięto dla przypadku, gdy funkcja regresji (1.2) przyjmuje postać

$$(1.11) \quad EY_x = \theta_1 + \theta_2 x + \theta_3 x^2 + \dots + \theta_k x^{k-1}$$

tnz. gdy jest ona wielomianem jednej zmiennej stopnia $k-1$. Dla takiej funkcji regresji podano w pracach Guesta [13] i Hoela [16] rozwiązanie zadań 2 i 3, gdy $X_1 = X_2 = [-1, 1]$. Rozwiązaniem jest tu miara probabilistyczna przyporządkowująca jednakową miarę $1/k$ każdemu z k punktów $x^{(1)}, x^{(2)}, \dots, x^{(k)} \in [-1, 1]$, przy czym punkty te są miejscami zerowymi funkcji $(1-x^2)p'_{k-1}(x)$, gdzie $p'_{k-1}(x)$ jest pochodną wielomianu Legendre'a stopnia $k-1$:

$$p_{k-1}(x) = \frac{1}{2^{k-1} \cdot (k-1)!} \cdot \frac{d^{(k-1)}(x^2 - 1)^{k-1}}{dx^{(k-1)}} .$$

Dla omawianego przypadku $X_1 = X_2 = [-1, 1]$ podano również (zob. Maljutow i Fiedorow [34]) rozwiązanie zadania 4. Rozwiązująca to zadanie miara probabilistyczna jest skoncentrowana w tych samych k punktach odcinka $[-1, 1]$, co miara rozwiązująca zadania 2 i 3, natomiast przyporządkowuje tym punktom odpowiednio miary

$$\lambda_1 = \frac{|p_{k-1}^{-1}(x^{(1)})|}{\sum_{j=1}^k |p_{k-1}^{-1}(x^{(j)})|} \quad \text{gdzie } l = 1, 2, \dots, k.$$

Hoel i Levine [20] podali rozwiązania zadania 3 dla omawianej tu funkcji regresji (1.11) w przypadku, gdy $X_1 = [-1, 1]$, natomiast $X_1 \neq X_2$. W szczególności rozważano jednopunktowy zbiór $X_2 = \{b\}$, gdzie $b \notin X_1$; następnie odcinek $X_2 = [a, b]$, gdzie $1 \leq a < b$ oraz odcinek $X_2 = [a, b]$, gdzie $-1 \leq a < 1$, $b > b_1 > 1$. Przy każdym z wymienionych tu zbiorów X_2 , rozwiązaniem zadania 3 jest miara probabilistyczna, która przyporządkowuje k punktom $x^{(l)} = -\cos[(l-1)\pi/(k-1)]$ zbioru X_1 odpowiednio miary

$$\lambda_1 = \frac{|L_1(b)|}{\sum_{j=1}^k |L_j(b)|} \quad l = 1, 2, \dots, k$$

gdzie

$$L_j(x) = \frac{(x - x^{(1)}) \dots (x - x^{(j-1)}) (x - x^{(j+1)}) \dots (x - x^{(k)})}{(x^{(j)} - x^{(1)}) \dots (x^{(j)} - x^{(j-1)}) (x^{(j)} - x^{(j+1)}) \dots (x^{(j)} - x^{(k)})},$$

natomiast liczbę b_1 wyznacza się jednoznacznie z równania

$$\max_{a \leq x \leq 1} \sum_{j=1}^k (L_j^2(x) / |L_j(b_1)|) = \sum_{j=1}^k |L_j(b_1)|.$$

Z innych rezultatów dotyczących funkcji regresji postaci (1.11) należy wspomnieć rozwiązanie zadania 3 dla przypadku $X_1 = [-1, 1]$ i $X_2 = [-a, a]$ przy $a \rightarrow 0$ (patrz: Kiefer i Wolfowitz [32]) oraz dla przypadku $X_1 = [-1, a] \cup [c, 1]$, gdzie $-1 < a < 0 < 1$, zaś $X_2 = \{b\}$, gdzie $a < b < 0$ (patrz: Hoel [18]).

Znacznie bardziej skomplikowane i mniej systematyczne wyniki znane są dla rozległej klasy funkcji regresji, będących wielomianami wielu zmiennych. Tylko w najprostszych przypadkach, gdzie funkcja regresji wielu zmiennych ma postać produktu kartezjańskiego pewnej liczby wielomianów jednej zmiennej, można bezpośrednio uogólnić niektóre rozwiązania jednowymiarowe na większą liczbę zmiennych (por. np. Hoel [17], [18]). Natomiast dla funkcji regresji, która jest wielomianem m zmiennych nie mającym postaci produktu kartezjańskiego, najczęściej spotkać można w literaturze wyniki dotyczące rozwiązania zadań 2 i 3, gdy zbiory X_1 i X_2 określone są następująco :

1° $X_1 = X_2 = X = \left\{ x = (x_1, \dots, x_m) : \sum_{i=1}^m x_i^2 \leq 1 \right\}$, tzn. zbiór X tworzy kulę jednostkową w przestrzeni \mathcal{E}^m ,

2° $X_1 = X_2 = X = \left\{ x = (x_1, \dots, x_m) : -1 \leq x_i \leq 1 \right.$
 $\left. (i = 1, \dots, m) \right\}$ tzn. zbiór X tworzy kostkę w przestrzeni \mathcal{E}^m ,

3° $X_1 = X_2 = X = \left\{ x = (x_1, \dots, x_m) : \sum_{i=1}^m x_i = 1, x_i \geq 0 \right.$
 $\left. (i = 1, \dots, m) \right\}$, tzn. zbiór X tworzy simpleks $m-1$ wymiarowy w przestrzeni \mathcal{E}^m .

Kiefer [29] podał ogólne rozwiązanie zadań 2 i 3 dla przypadku, gdy $X_1 = X_2 = X$, gdzie zbiór X tworzy kulę jednostkową w przestrzeni \mathcal{E}^m , a funkcja regresji jest wielomianem k -tego stopnia m zmiennych. Rozwiązaniem jest miara probabilistyczna skoncentrowana na $(k+1)/2$ sferach $S_1, S_2, \dots, S_{\frac{k+1}{2}}$ o środku w punkcie $(0, \dots, 0)$, przy czym jedna z tych sfer ma zawsze promień równy 1. Jeśli k jest

liczbą parzystą, to środek kuli liczy się jako $1/2$ sfery. Miary $\lambda_1, \lambda_2, \dots, \lambda_{\frac{k+1}{2}}$ przyporządkowane odpowiednim sferom mają na każdej z tych sfer rozkład jednostajny. Liczba sfer jest funkcją niemalejącą stopnia wielomianu będącego funkcją regresji, natomiast dla ustalonego stopnia k funkcji regresji miary $\lambda_1, \lambda_2, \dots, \lambda_{\frac{k+1}{2}}$ oraz promienie $r_1, r_2, \dots, r_{\frac{k+1}{2}}$ sfer $S_1, S_2, \dots, S_{\frac{k+1}{2}}$ są funkcją liczby zmiennych. Gdy $k = 1$, to oczywiście cała miara 1 jest przyporządkowana powierzchni kuli. Gdy $k = 2$, to środkowi kuli jest przyporządkowana miara $2/[(m+1)(m+2)]$, a powierzchni kuli - miara $\frac{m(m+3)}{(m+1)(m+2)}$, (por. [29]). W tabeli 1.1 podano dla przypadku $k = 3$ miary λ_2 i promienie r_2 (oczywiście $r_1 = 1, \lambda_1 = 1 - \lambda_2$), gdy liczba zmiennych m nie jest większa od 10. (por. [7]).

T a b e l a 1.1

m	r_2	λ_2
2	0,447	0,3077
3	0,516	0,2455
4	0,545	0,1695
5	0,560	0,1241
6	0,570	0,0948
7	0,576	0,0749
8	0,581	0,0607
9	0,584	0,0502
10	0,587	0,0422

W sytuacji, gdy $X_1 = X_2 = X$, przy czym zbiór X tworzy

kostkę w przestrzeni ξ^m , natomiast funkcja regresji jest wielomianem m zmiennych, rozwiązania zadań 2 i 3 znane są przy założeniu, że wielomian ten jest co najwyżej stopnia drugiego. Poza tym dla wielomianów niektórych wyższych stopni rozwiązania są znane dla określonej liczby zmiennych, np. dla wielomianu stopnia trzeciego dwu zmiennych (por. [7]). Dla wielomianu stopnia pierwszego rozwiązania zadań 2 i 3 wynikają bezpośrednio z wyników przedstawionych w rozdziale IV, po uniezależnieniu ich od liczby obserwacji. Pewne rozwiązanie zadań 2 i 3 dla kwadratowej funkcji regresji co najwyżej 5 zmiennych podał Kiefer w pracy [28]. Rozwiązaniem jest miara probabilistyczna przyporządkowująca miarę λ_w każdemu z 2^m wierzchołków kostki, miarę λ_k każdemu z $m2^{m-1}$ środków krawędzi łączących wierzchołki kostki oraz miarę λ_s każdemu z $m(m-1)2^{m-3}$ środków ścian dwuwymiarowych kostki m -wymiarowej. Numeryczne wartości miar λ_w , λ_k i λ_s w zależności od wymiaru m podane są w tabeli 1.2. Rozwiązanie zadań 2 i 3 dla kwadratowej funkcji regresji dowolnej liczby

T a b e l a 1.2

m	λ_w	λ_k	λ_s
2	0,1458	0,08015	0,0962
3	0,071975	0,01895	0,0328
4	0,03705	0,0038375	0,01185
5	0,01928	0,0003125	0,004475

zmiennych uzyskali Farrell, Kiefer i Walbran [7]. Wyniki tych autorów są następujące: Niech J^j oznacza podzbiór m -wymiarowej kostki $X = \{x = (x_1, \dots, x_m), -1 \leq x_i \leq 1$

$(i = 1, \dots, m)$ } złożony z tych i tylko tych $2^{m-j} \binom{m}{j}$ punktów kostki, których dokładnie j współrzędnych równa się zeru, natomiast pozostałe współrzędne są równe $+1$ lub -1 . Np. J^0 jest 2^m -elementowym zbiorem wierzchołków kostki, J^1 jest $m \cdot 2^{m-1}$ -elementowym zbiorem środków krawędzi kostki, a J^m jest zbiorem, którego jedynym elementem jest środek kostki. Niech $J(j_1, j_2, j_3) = \bigcup_{i=1}^3 J^{j_i}$, gdzie $0 \leq j_1 < j_2 < j_3 \leq m$, tzn. zbiór $J(j_1, j_2, j_3)$ jest sumą trzech różnych zbiorów spośród zbiorów $J^0, J^1, J^2, \dots, J^m$. Miara probabilistyczna ξ^* spełniająca dla nieujemnych całkowitych liczb a_i , gdzie $\sum_{i=1}^m a_i \leq 4$, warunki

$$1^{\circ} \int_X x_1^{a_1} x_2^{a_2} \dots x_m^{a_m} \xi^*(dx) = \begin{cases} 0 & \text{gdy co najmniej jedna z} \\ & \text{liczb } a_i \text{ jest nieparzysta} \\ \lambda(a_1, \dots, a_m) & \text{gdy wszystkie} \\ & \text{liczby } a_i \text{ są parzyste} \end{cases}$$

$2^{\circ} \lambda(a_1, \dots, a_m) = \text{const}$ dla dowolnej permutacji ustalonego ciągu a_1, \dots, a_m ,

$$3^{\circ} \int_{J(j_1, j_2, j_3)} \xi^*(dx) = 1,$$

jest rozwiązaniem zadań 2 i 3 wtedy i tylko wtedy, jeśli

$$\begin{cases} j_1 = 0, j_2 = 1, 2 \leq j_3 \leq m, \text{ gdy } 2 \leq m \leq 5 \\ j_1 = 0, j_2 = 1 \text{ lub } 2, 3 \leq j_3 \leq m, \text{ gdy } m \geq 6, \end{cases}$$

$$\text{oraz } \lambda(2, 0, 0, \dots, 0) = \lambda(4, 0, 0, \dots, 0) =$$

$$= \frac{(m+3)}{4(m+1)(m+2)^2} \left[(2m^2 + 3m + 7) + (m-1) \sqrt{(4m^2 + 12m + 17)} \right],$$

$$\lambda(2, 2, 0, \dots, 0, \dots, 0) =$$

$$= \frac{(m+3)}{8(m+1)(m+2)^3} \left[(4m^3 + 8m^2 + 11m - 5) + (2m^2 + m + 3) \sqrt{(4m^2 + 12m + 17)} \right].$$

Łatwo zauważyć, że wcześniejszy wynik Kiefera [28] dla $m \leq 5$ jest szczególnym przypadkiem rozwiązań Farrella, Kiefera i

Walbrana, a mianowicie cała miara jest tam skoncentrowana na zbiorze $J(0, 1, 2)$.

Przypadek $X_1 = X_2 = X$, gdzie zbiór $X = \{x = (x_1, \dots, x_m) : \sum_{i=1}^m x_i = 1, x_i \geq 0 \text{ (} i = 1, \dots, m)\}$ tworzy simpleks $m-1$ wymiarowy w przestrzeni ξ^m , interesuje tych eksperymentatorów, którzy zajmują się planowaniem doświadczeń z mieszaninami, a więc np. chemików i technologów. Rozwiązania zadań 2 i 3 przy funkcjach regresji, które są rozmaicie określonymi wielomianami m zmiennych, można dla tego przypadku znaleźć np. w pracach Kiefera [28], Farrella, Kiefera i Walbrana [7], Atwooda [1] oraz Golikowej i Mikeszinej [36].

Spośród wyników teorii planów przybliżonych należy wymienić również grupę prac Kiefera i Wolfowitza [33], Hoëla [19], Studdena [38], [39] i innych autorów, które zawierają uogólnienie cytowanych wyżej wyników Hoëla i Levine'a [20] na przypadek, gdy funkcje $f_1(x), \dots, f_k(x)$ w równaniu $EY_x = \theta_1 f_1(x) + \dots + \theta_k f_k(x)$ tworzą system Czebyszewa, tzn. gdy są one takimi funkcjami ciągłymi określonymi na odcinku $[a, b]$ prostej rzeczywistej, że każda kombinacja liniowa $a_1 f_1(x) + a_2 f_2(x) + \dots + a_k f_k(x)$, gdzie $a_1^2 + a_2^2 + \dots + a_k^2 > 0$, ma na odcinku $[a, b]$ co najwyżej $k-1$ różnych miejsc zerowych. Z punktu widzenia zastosowań interesująca jest seria prac na temat sekwencyjnego generowania rozwiązania zadania 2. O metodach sekwencyjnego osiągnięcia takiej optymalnej miary probabilistycznej traktują np. prace Wynna [40], Fiedorowa [9] i najnowsza praca Atwooda [2]. Bezpośrednie znaczenie dla zastosowań ekonomicznych mogą mieć prace Bandemera [3] i Junga [23], w których dla liniowej funkcji regresji jednej zmiennej

explicito rozważa się zagadnienie kosztów eksperymentu.

Przegląd wyników teorii planów przybliżonych osiągniętych do roku 1965 zawiera rozdział 10 monografii Karlina i Studena [25], przy czym oryginalne wyniki autorów tej monografii zebrane są w [26]. Zbiór [36] prac autorów związanych z Laboratorium Metod Statystycznych Uniwersytetu Moskiewskiego stanowi w części próbę wykorzystania teorii planów przybliżonych w praktycznym planowaniu eksperymentów, a mutacje niektórych artykułów (por. np. [35]) z tego zbioru opublikowano w czasopiśmie zachodnich. Rangę tej grupy uczonych radzieckich określa fakt, że monografia Fiedorowa [9] poświęcona teorii planowania eksperymentów została przetłumaczona na język angielski i wydana w Stanach Zjednoczonych. Zainteresowanie problematyką planowania eksperymentów dla zagadnień regresji jest również bardzo żywe w NRD, a ostatnio i w Polsce, czego dowodem mogą być np. niedawno wydane książki Bandemera i współautorów [4] oraz Kacprzyńskiego [24].

Należy jeszcze zwrócić uwagę na związki między sformułowanymi uprzednio kryteriami 1 - 4 optymalności planów eksperymentalnych i postawionymi w teorii planów przybliżonych zadaniami 1 - 4. Używając terminologii kryteriów 1 - 4 można powiedzieć, że zadania 1 - 4 polegają na wyznaczeniu odpowiednio miar A , D , G lub I - optymalnych w szerszej klasie \mathcal{A} miar probabilistycznych. Jest oczywiste, że plan optymalny dla rodziny \mathcal{A}_N może nie być rozwiązaniem żadnego z zadań 1 - 4. Z drugiej strony, z uwagi na punkt (d) twierdzenia 1, można sprowadzić każde rozwiązanie zadań 1 - 4 do miary dyskretnej skoncentrowanej na skończonej liczbie

punktów. Jeśli w każdym z tych punktów skoncentrowane są miary będące liczbami wymiernymi, to takie rozwiązanie jednego z zadań 1 - 4 jest dla pewnego N planem optymalnym w sensie odpowiadającego temu zadaniu kryterium. Dowlaną miarę rozwiązującą jedno z zadań 1 - 4 można też - dla dostatecznie dużego N - zmodyfikować tak, żeby otrzymać plan eksperymentalny dla N obserwacji a p r o k s y m u j ą c y optymalny plan eksperymentalny w sensie odpowiedniego z kryteriów 1 - 4 . Dla niezbyt wielkiej liczby N bezpośrednia modyfikacja rozwiązań zadań 1 - 4 może jednak dać plan eksperymentalny daleki od wymagań optymalności stawianych przez odpowiednie z kryteriów 1 - 4 . Powoduje to istotne ograniczenia w zastosowaniu teorii planów przybliżonych do konstrukcji optymalnych planów eksperymentalnych. Podkreślić trzeba, że planowanie eksperymentów z małą liczbą obserwacji jest bardzo ważnym zagadnieniem w praktyce doświadczalnictwa.

Rozdział II

REGRESJA LINIOWA NA ODCINKU PROSTEJ

Założmy, że $X = X_1 = X_2 = [a, b]$, $b > a$, $k = 2$,
 $f_1(x) = 1$, $f_2(x) = x$. Wówczas funkcja (1.2) ma postać

$$(2.1) \quad EY_x = \theta_1 + \theta_2 x .$$

Przedstawimy teraz plany eksperymentalne dla estymacji tej funkcji regresji liniowej, które spełniają niektóre z podanych w rozdziale I kryteriów optymalności.

Jeżeli liczba obserwacji jest nieparzysta, tzn. $N = 2n + 1$, gdzie $n = 1, 2, \dots$, to plan eksperymentalny $P_{N,1} \in \mathcal{D}_N$ określony następująco

$$\left\{ \begin{array}{ccc} a & \frac{a+b}{2} & b \\ \frac{n}{2n+1} & \frac{1}{2n+1} & \frac{n}{2n+1} \end{array} \right\}$$

jest planem G - optymalnym ¹⁾.

1) Ze względów identyfikacji kolejne plany eksperymentalne P_N rozważane w tej pracy będą oznaczane dodatkowymi wskaźnikami, np. w rozdziale bieżącym rozważa się plany $P_{N,1}$ oraz $P_{N,2}$, a rozdział III zaczyna się od prezentacji planu $P_{N,10}$.

Z określenia planu eksperymentalnego $P_{N,1}$ oraz ze wzoru (1.3) wynika, że

$$(2.2) \quad M(P_{N,1}) = \begin{bmatrix} 1 & \frac{a+b}{2} \\ \frac{a+b}{2} & \frac{4n(a^2+b^2)+(a+b)^2}{4n(2n+1)} \end{bmatrix}$$

$$(2.3) \quad M^{-1}(P_{N,1}) = \begin{bmatrix} \frac{4n(a^2+b^2)+(a+b)^2}{2n(a-b)^2} & \frac{-(2n+1)(a+b)}{n(a-b)^2} \\ \frac{-(2n+1)(a+b)}{n(a-b)^2} & \frac{2(2n+1)}{n(a-b)^2} \end{bmatrix}$$

Wobec tego¹⁾

$$(2.4) \quad d(x, P_{N,1}) = \begin{bmatrix} 1 & x \end{bmatrix} M^{-1}(P_{N,1}) \begin{bmatrix} 1 \\ x \end{bmatrix} =$$

$$= \frac{2(2n+1)}{n(a-b)^2} x^2 - \frac{2(2n+1)(a+b)}{n(a-b)^2} x + \frac{4n(a^2+b^2)+(a+b)^2}{2n(a-b)^2}$$

Prawa strona równania (2.4) jest funkcją kwadratową zmiennej x o dodatnim współczynnikiem przy x^2 . Oznacza to, że funkcja $d(x, P_{N,1})$ nie osiąga maksimum wewnątrz odcinka $[a, b]$. Stąd łatwo widać, że

$$(2.5) \quad \max_{x \in [a, b]} d(x, P_{N,1}) = d(a, P_{N,1}) = d(b, P_{N,1}) =$$

$$= \frac{4n+1}{2n} = 2 + \frac{1}{2n} = 2 + \frac{1}{N-1}.$$

1) Zapis $\begin{bmatrix} 1 & x \end{bmatrix}$ oznacza jednowierszową macierz o dwu elementach 1 oraz x ; zapis $\begin{bmatrix} 1 \\ x \end{bmatrix}$ oznacza wektor o dwu składowych równych odpowiednio 1 i x .

W świetle punktu c twierdzenia 2 otrzymany wynik nie rozstrzyga o G - optymalności planu $P_{N,1}$. W rzeczywistości plan $P_{N,1}$ tę własność posiada. Dowód G - optymalności planu eksperymentalnego $P_{N,1}$ dla $N = 2n + 1$ ($n = 1, 2, \dots$) obserwacji jest głównym wynikiem pracy Junga [22].

Przytoczmy z kolei znane

TWIERDZENIE 4. Niech $X = X_1 = X_2 = [a, b]$, $k = 2$, $f_1(x) = 1$, $f_2(x) = x$. Wówczas dla każdej parzystej liczby obserwacji N , tzn. $N = 2n$, gdzie $n = 1, 2, \dots$, plan eksperymentalny $P_{N,2} \in \mathfrak{D}_N$ określony następująco

$$\left\{ \begin{array}{cc} a & b \\ \frac{n}{N} & \frac{n}{N} \end{array} \right\} \quad \text{gdzie} \quad \frac{n}{N} = \frac{1}{2}$$

jest planem D - optymalnym i równocześnie planem G - optymalnym.

Dowód. Ze wzoru (1.3) wynika, że dla planu eksperymentalnego $P_{N,2}$ mamy

$$(2.6) \quad M(P_{N,2}) = \begin{bmatrix} 1 & \frac{a+b}{2} \\ \frac{a+b}{2} & \frac{a^2+b^2}{2} \end{bmatrix}$$

$$(2.7) \quad M^{-1}(P_{N,2}) = \begin{bmatrix} \frac{2(a^2+b^2)}{(a-b)^2} & \frac{-2(a+b)}{(a-b)^2} \\ \frac{-2(a+b)}{(a-b)^2} & \frac{4}{(a-b)^2} \end{bmatrix}$$

Wobec tego

$$(2.8) \quad d(x, P_{N,2}) = \begin{bmatrix} 1 & x \end{bmatrix} M^{-1}(P_{N,2}) \begin{bmatrix} 1 \\ x \end{bmatrix} =$$

$$= \frac{4}{(a-b)^2} x^2 - \frac{4(a+b)}{(a-b)^2} x + \frac{2(a^2+b^2)}{(a-b)^2} .$$

Otrzymane wyrażenie jest kwadratową funkcją zmiennej x o współczynniku dodatnim przy x^2 . Oznacza to, że funkcja $d(x, P_{N,2})$ nie osiąga maksimum wewnątrz odcinka $[a, b]$. Stąd łatwo widać, że

$$(2.9) \quad \max_{x \in [a, b]} d(x, P_{N,2}) = d(a, P_{N,2}) = d(b, P_{N,2}) = 2 .$$

Przy poczynionych założeniach z twierdzenia 2 wynika, że dla dowolnego planu eksperymentalnego zachodzi nierówność

$\max_{x \in [a, b]} d(x, P_N) \geq 2$. Z równości (2.9) i z tegoż twierdzenia 2 wnioskujemy zatem, że plan $P_{N,2}$ jest D - i G - optymalny.

c.b.d.o.

Twierdzenie 4 dla przypadku $a = -1, b = 1$ pozostaje w bezpośrednim związku z omawianymi w rozdziale I wynikami Guesta [13] i Hoela [16] dla funkcji regresji będącej wielomianem dowolnego stopnia, gdy stopień wielomianu jest równy 1. Ponadto przy $a = -1$ i $b = 1$ plan $P_{N,2}$ jest nie tylko D - i G - optymalny, ale także A - oraz I - optymalny. A - optymalność wynika łatwo z twierdzenia 3, gdyż dla $a = -1$ i $b = 1$ macierz $M(P_{N,2}) = \begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix}$ i równanie (1.10) z twierdzenia 3 jest spełnione dla $c = 1$. Podobnie I - optymalność otrzymujemy bezpośrednio z cytowanego na str. 21-22 (rozdz.I) wyniku Maljutowa i Fiedorowa [34].

Warto jeszcze zauważyć, że zgodnie z oczekiwaniem prawdziwe jest następujące

TWIERDZENIE 5. Niech N_1 i N_2 oznaczają dwie liczby naturalne takie, że $N_2 > N_1 \geq 2$. Jeżeli w celu estymacji

funkcji regresji (2.1) na odcinku $[a, b]$ wykonano N_1 obserwacji według planu eksperymentalnego $P_{N_1,1}$, gdy N_1 jest liczbą nieparzystą, zaś według planu $P_{N_1,2}$, gdy N_1 jest liczbą parzystą, a ponadto jeśli wykonano N_2 obserwacji według planu $P_{N_2,1}$, gdy N_2 jest liczbą nieparzystą, zaś według planu $P_{N_2,2}$, gdy N_2 jest liczbą parzystą, to:

$$1^{\circ} \quad \left(\sum_{i=1}^2 \text{Var } \hat{\theta}_1 \right)_1 > \left(\sum_{i=1}^2 \text{Var } \hat{\theta}_1 \right)_2,$$

gdzie $\left(\sum_{i=1}^2 \text{Var } \hat{\theta}_1 \right)_j$ oznacza sumę wariancji estymatorów $\hat{\theta}_1$ i $\hat{\theta}_2$, gdy wykonano N_j ($j = 1, 2$) obserwacji;

$$2^{\circ} \quad \text{Vol} (\hat{\theta}_1, \hat{\theta}_2)_1 > \text{Vol} (\hat{\theta}_1, \hat{\theta}_2)_2,$$

gdzie $\text{Vol} (\hat{\theta}_1, \hat{\theta}_2)_j$ oznacza pole elipsy koncentracji rozkładu estymatorów $\hat{\theta}_1, \hat{\theta}_2$, gdy wykonano N_j ($j = 1, 2$) obserwacji;

$$3^{\circ} \quad \sigma_1^2 > \sigma_2^2,$$

gdzie $\sigma_j^2 = \max_{x \in [a, b]} \text{Var} (\hat{E}Y_x) = \frac{\sigma^2}{N_j} \max_{x \in [a, b]} d(x, P_{N_j, i})$ oznacza maksymalną wartość wariancji estymatora funkcji regresji (2.1) na odcinku $[a, b]$, gdy wykonano N_j obserwacji według planu $P_{N_j, i}$ ($i, j = 1, 2$);

$$4^{\circ} \quad \overline{\sigma}_1^2 > \overline{\sigma}_2^2,$$

gdzie $\overline{\sigma}_j^2 = \frac{1}{b-a} \int_a^b \text{Var} (\hat{E}Y_x) dx = \frac{\sigma^2}{b-a} \int_a^b \frac{1}{N_j} d(x, P_{N_j, i}) dx$ oznacza wartość średnią wariancji estymatora funkcji regresji (2.1) na odcinku $[a, b]$, gdy wykonano N_j obserwacji według planu $P_{N_j, i}$ ($i, j = 1, 2$).

Dowód każdej z nierówności od 1^o do 4^o podzielimy na 3 przypadki:

Przypadek A, gdy N_1 i N_2 są dwiema liczbami nieparzystymi.

Przypadek B, gdy N_1 i N_2 są dwiema liczbami parzystymi.

Przypadek C, gdy jedna z liczb N_1, N_2 jest liczbą parzystą, a druga - nieparzystą.

Udowodnimy najpierw nierówność 1^o.

Ze wzoru (1.5) wynika, że

$$\left(\sum_{l=1}^2 \text{var } \hat{\theta}_l\right)_j = \frac{\sigma^2}{N_j} m^{11}(P_{N_j,1}) + m^{22}(P_{N_j,1}),$$

gdzie $m^{ll}(P_{N_j,1})$ oznacza l -ty ($l = 1, 2$) element głównej przekątnej macierzy $M^{-1}(P_{N_j,1})$.

Jeśli N_j jest liczbą nieparzystą, to ze wzoru (2.3) otrzymujemy

$$m^{11}(P_{N_j,1}) = \frac{4n_j(a^2 + b^2) + (a+b)^2}{2n_j(a-b)^2}, \quad m^{22}(P_{N_j,1}) = \frac{2(2n_j+1)}{n_j(a-b)^2},$$

skąd

$$\begin{aligned} (2.10) \quad \left(\sum_{l=1}^2 \text{var } \hat{\theta}_l\right)_j &= \frac{\sigma^2}{N_j} \cdot \frac{4n_j(a^2+b^2) + (a+b)^2 + 4(2n_j+1)}{2n_j(a-b)^2} = \\ &= \frac{\sigma^2}{N_j} \cdot \frac{(2N_j-1)(a^2+b^2) + 2ab + 4N_j}{(N_j-1)(a-b)^2}, \end{aligned}$$

gdzież $N_j = 2n_j + 1$.

Jeśli N_j jest liczbą parzystą, to ze wzoru (2.7) otrzymujemy

$$m^{11}(P_{N_j,2}) = \frac{2(a^2 + b^2)}{(a-b)^2}, \quad m^{22}(P_{N_j,2}) = \frac{4}{(a-b)^2},$$

skąd

$$(2.11) \quad \left(\sum_{j=1}^2 \text{Var } \hat{\theta}_1 \right)_j = \frac{\sigma^2}{N_j} \cdot \frac{2(a^2 + b^2 + 2)}{(a-b)^2} .$$

Przypadek A. Należy udowodnić, że (por. wzór (2.10)), że

$$(2.12) \quad \frac{\sigma^2}{N_1} \cdot \frac{(2N_1-1)(a^2+b^2) + 2ab + 4N_1}{(N_1-1)(a-b)^2} - \\ - \frac{\sigma^2}{N_2} \cdot \frac{(2N_2-1)(a^2+b^2) + 2ab + 4N_2}{(N_2-1)(a-b)^2} > 0 .$$

W tym celu wystarczy pokazać, że powyższa nierówność jest spełniona dla $N_2 = N_1 + 2$, gdzie $N_1 = 3, 5, 7, \dots$. Otrzymujemy wtedy

$$\frac{\sigma^2}{N_1} \cdot \frac{(2N_1-1)(a^2+b^2)+2ab+4N_1}{(N_1-1)(a-b)^2} - \frac{\sigma^2}{(N_1+2)} \cdot \frac{(2N_1+3)(a^2+b^2)+2ab+4(N_1+2)}{(N_1+1)(a-b)^2} = \\ = \frac{\sigma^2}{(a-b)^2} \cdot \frac{4N_1^2 (a^2+b^2+2) + 4(2N_1+1)ab + 2(2N_1-1)(a^2+b^2)}{(N_1-1) N_1 (N_1+1)(N_1+2)} .$$

Ponieważ $N_1 \geq 3$, więc otrzymane wyrażenie jest dodatnie, tzn. spełniona jest nierówność (2.12).

Przypadek B. Ponieważ $N_2 > N_1$, więc nierówność 1^o wynika natychmiast ze wzoru (2.11).

Przypadek C. Z udowodnionych już przypadków A i B wynika, że nierówność 1^o wystarczy dowieść tu dla $N_2 = N_1 + 1$.

Jeżeli N_1 jest liczbą nieparzystą ($N_1 = 3, 5, 7, \dots$), to ze wzorów (2.10) i (2.11) mamy

$$\left(\sum_{i=1}^2 \text{Var } \hat{\theta}_1 \right)_1 - \left(\sum_{i=1}^2 \text{Var } \hat{\theta}_1 \right)_2 =$$

$$\begin{aligned}
 &= \frac{\sigma^2}{N_1} \cdot \frac{(2N_1-1)(a^2+b^2)+2ab+4N_1}{(N_1-1)(a+b)^2} - \frac{\sigma^2}{N_1+1} \cdot \frac{2(a^2+b^2+2)}{(a-b)^2} = \\
 &= \frac{\sigma^2}{(a-b)^2} \cdot \frac{(3N_1-1)(a^2+b^2) + 2(N_1+1)ab + 8N_1}{(N_1-1) N_1 (N_1+1)} > 0.
 \end{aligned}$$

Jeżeli N_1 jest liczbą parzystą ($N_1 = 2, 4, 6, \dots$),
to ze wzorów (2.11) i (2.10) mamy

$$\begin{aligned}
 &\left(\sum_{i=1}^2 \text{Var } \hat{\theta}_1\right)_1 - \left(\sum_{i=1}^2 \text{Var } \hat{\theta}_1\right)_2 = \\
 &= \frac{\sigma^2}{N_1} \cdot \frac{2(a^2+b^2+2)}{(a-b)^2} - \frac{\sigma^2}{N_1+1} \cdot \frac{(2N_1+1)(a^2+b^2) + 2ab + 4(N_1+1)}{N_1(a-b)^2} = \\
 &= \frac{\sigma^2}{N_1(N_1+1)} > 0,
 \end{aligned}$$

co kończy dowód nierówności 1°.

Udowodnimy teraz nierówność 2°.

Ze wzoru (1.6) wynika, że nierówność 2° jest równoważna nierówności

$$(2.13) \quad N_1^2 \cdot \det M(P_{N_1,1}) < N_2^2 \cdot \det M(P_{N_2,1}),$$

gdzie $1,1 = 1, 2$.

Przypadek A. Ze wzoru (2.2) mamy

$$\begin{aligned}
 (2.14) \quad \det M(P_{N_j,1}) &= \frac{4n_j(a^2+b^2) + (a+b)^2}{4(2n_j+1)} - \left(\frac{a+b}{2}\right)^2 = \\
 &= \frac{n_j(a-b)^2}{2(2n_j+1)} = \frac{(N_j-1)(a-b)^2}{4 N_j},
 \end{aligned}$$

gdyż $N_j = 2n_j + 1$. Stąd

$$N_1^2 \det M(P_{N_1,1}) = \frac{N_1(N_1-1)(a-b)^2}{4}$$

oraz

$$N_2^2 \det M(P_{N_2,1}) = \frac{N_2(N_2-1)(a-b)^2}{4} .$$

Porównując prawe strony obu równań otrzymujemy - wobec faktu, że $N_2 > N_1$ - relację

$$N_1^2 \det M(P_{N_1,1}) < N_2^2 \det M(P_{N_2,1}) ,$$

która oznacza, że w rozważanym przypadku nierówność (2.13) jest prawdziwa.

Przypadek B. Ze wzoru (2.6) mamy

$$(2.15) \quad \det M(P_{N_1,2}) = \det M(P_{N_2,2}) = \frac{a^2+b^2}{2} - \left(\frac{a+b}{2}\right)^2 = \left(\frac{a-b}{2}\right)^2$$

Ponieważ $N_1 < N_2$, więc nierówność (2.13) jest w tym przypadku oczywista.

Przypadek C. Z udowodnionych już przypadków A i B wynika, że prawdziwość nierówności (2.13) wystarczy pokazać tu dla $N_2 = N_1 + 1$.

Jeżeli N_1 jest liczbą nieparzystą ($N_1 = 3, 5, 7, \dots$), to ze wzoru (2.14) otrzymujemy, że

$$N_1^2 \det M(P_{N_1,1}) = N_1(N_1-1) \frac{(a-b)^2}{4} ,$$

zaś ze wzoru (2.15) otrzymujemy dla $N_2 = N_1 + 1$

$$(N_1+1)^2 \det M(P_{N_1+1,2}) = (N_1+1)^2 \frac{(a-b)^2}{4} .$$

Porównując prawe strony obu równań otrzymujemy relację

$$(2.16) \quad N_1^2 \det M(P_{N_1,1}) < (N_1+1)^2 \det M(P_{N_1+1,2}) .$$

Jeżeli N_1 jest liczbą parzystą ($N_1 = 2, 4, 6, \dots$),
to ze wzoru (2.15) otrzymujemy, że

$$N_1^2 \det M(P_{N_1,2}) = N_1^2 \frac{(a-b)^2}{4} ,$$

zaś ze wzoru (2.14) otrzymujemy dla $N_2 = N_1+1$

$$(N_1+1)^2 \det M(P_{N_1+1,1}) = (N_1+1) N_1 \frac{(a-b)^2}{4}$$

Porównując prawe strony obu równań otrzymujemy relację

$$(2.17) \quad N_1^2 \det M(P_{N_1,2}) < (N_1+1)^2 \det M(P_{N_1+1,1}) .$$

Z nierówności (2.16) i (2.17) wynika, że w omawianym
przypadku spełniona jest nierówność (2.13), co równocześnie
kończy dowód nierówności 2^o.

Z kolei udowodnimy nierówność 3^o.

Przypadek A. Wystarczy tu pokazać, że różnica

$$\sigma_1^2 - \sigma_2^2 = \frac{\sigma^2}{N_1} \max_{x \in [a,b]} d(x, P_{N_1,1}) - \frac{\sigma^2}{N_2} \max_{x \in [a,b]} d(x, P_{N_2,1})$$

jest dodatnia dla $N_2 = N_1+2$ ($N_1 = 3, 5, 7, \dots$). Ze
wzoru (2.5) mamy wtedy

$$\begin{aligned} \sigma_1^2 - \sigma_2^2 &= \frac{\sigma^2}{N_1} \left(2 + \frac{1}{N_1-1}\right) - \frac{\sigma^2}{N_1+2} \left(2 + \frac{1}{N_1+1}\right) = \\ &= \sigma^2 \cdot \frac{2(2N_1^2 + 2N_1 - 1)}{(N_1-1) N_1 (N_1+1) (N_1+2)} > 0 . \end{aligned}$$

Przypadek B. Dla planu $P_{N,2}$ mamy ze wzorów (1.7) i (2.9),

że

$$\max_{x \in [a, b]} \text{Var} (\hat{E}Y_x) = \frac{\sigma^2}{N} \max_{x \in [a, b]} d(x, P_{N,2}) = \sigma^2 \cdot \frac{2}{N}.$$

Stąd wynika, że maksymalna wariancja estymatora funkcji regresji (2.1) otrzymanego po wykonaniu N obserwacji zgodnie z planem $P_{N,2}$ jest malejącą funkcją N . Wobec tego w przypadku tym spełniona jest nierówność 3^0 .

Przypadek C, Z udowodnionych już przypadków A i B wynika, że nierówność 3^0 wystarczy dowieść tu dla $N_2 = N_1 + 1$.

Jeżeli N_1 jest liczbą nieparzystą ($N_1 = 3, 5, 7, \dots$), to

$$\sigma_1^2 - \sigma_2^2 = \frac{\sigma^2}{N_1} \max_{x \in [a, b]} d(x, P_{N_1,1}) - \frac{\sigma^2}{N_1+1} \max_{x \in [a, b]} d(x, P_{N_1+1,2}).$$

Stąd oraz ze wzorów (2.5) i (2.9) otrzymujemy

$$\sigma_1^2 - \sigma_2^2 = \frac{\sigma^2}{N_1} \left(2 + \frac{1}{N_1-1}\right) - \frac{\sigma^2}{N_1+1} \cdot 2 = \sigma^2 \cdot \frac{3N_1 - 1}{(N_1-1)N_1(N_1+1)} > 0.$$

Jeżeli N_1 jest liczbą parzystą ($N_1 = 2, 4, 6, \dots$), to

$$\sigma_1^2 - \sigma_2^2 = \frac{\sigma^2}{N_1} \max_{x \in [a, b]} d(x, P_{N_1,2}) - \frac{\sigma^2}{N_1+1} \max_{x \in [a, b]} d(x, P_{N_1+1,1}).$$

Stąd oraz ze wzorów (2.9) i (2.5) otrzymujemy

$$\sigma_1^2 - \sigma_2^2 = \frac{\sigma^2}{N_1} \cdot 2 - \frac{\sigma^2}{N_1+1} \left(2 + \frac{1}{N_1}\right) = \sigma^2 \frac{1}{N_1(N_1+1)} > 0.$$

Łącznie z obydwu nierówności wynika, że spełniona jest nierówność 3^0 .

Pozostaje jeszcze pokazać, że prawdziwa jest nierówność 4^0 .

Przypadek A. Dla nieparzystej liczby $N = 2n + 1$ ($n = 1, 2, \dots$) obserwacji niech (por. wzór (2.4))

$$g(x; n, a, b) = \frac{1}{N} d(x, P_{N,1}) =$$

$$= \frac{1}{2n+1} \left[\frac{2(2n+1)}{n(a-b)^2} x^2 - \frac{2(2n+1)(a+b)}{n(a-b)^2} x + \frac{4n(a^2+b^2) + (a+b)^2}{2n(a-b)^2} \right] .$$

Nierówność 4^o zostanie dla tego przypadku udowodniona jeśli pokażemy, że w wyrażeniu

$$\frac{\sigma^2}{b-a} \int_a^b \frac{1}{N} d(x, P_{N,1}) dx = \frac{\sigma^2}{b-a} \int_a^b g(x; n, a, b) dx$$

dodatnia funkcja podcałkowa $g(x; n, a, b)$ jest dla każdego ustalonego $x \in [a, b]$ funkcją malejącą parametru n . W tym celu wystarczy obliczyć różnicę $g(x; n+1, a, b) - g(x; n, a, b)$ gdzie $n = 1, 2, \dots$. Po pewnych przekształceniach mamy

$$g(x; n+1, a, b) - g(x; n, a, b) =$$

$$= \frac{-2(2x^2 - 2ax - 2bx + a^2 + b^2)(4n^2 + 8n + 3) + (a-b)^2(4n+3)}{2n(n+1)(2n+1)(2n+3)(a-b)^2}$$

Dla wartości x zmieniających się w przedziale $[a, b]$ różnica ta osiąga maksimum w punkcie $\frac{a+b}{2}$. Maksimum to jest równe

$$\frac{-2}{(2n+1)(2n+3)} < 0 .$$

Zatem $g(x; n+1, a, b) - g(x; n, a, b) < 0$ dla $x \in [a, b]$, tzn. $g(x; n, a, b)$ jest malejącą funkcją n ($n = 1, 2, \dots$), co w danym przypadku dowodzi słuszności nierówności 4^o.

Przypadek B. W dowodzie twierdzenia 4 (patrz wzór (2.8)) pokazano, że wyrażenie $d(x, P_{N,2})$ nie zależy od liczby N wykonanych obserwacji. Wobec tego w omawianym przypadku

$$\overline{\sigma}_1^2 = \frac{\sigma^2}{b-a} \int_a^b \frac{1}{N_1} d(x, P_{N_1,2}) dx = \frac{\sigma^2}{b-a} \cdot \frac{1}{N_1} \cdot G(a,b)$$

oraz

$$\overline{\sigma}_2^2 = \frac{\sigma^2}{b-a} \int_a^b \frac{1}{N_2} d(x, P_{N_2,2}) dx = \frac{\sigma^2}{b-a} \cdot \frac{1}{N_2} \cdot G(a,b) ,$$

gdzie

$$G(a,b) = \int_a^b d(x, P_{N_j,2}) dx \quad (j = 1, 2) .$$

Ponieważ $N_2 > N_1$, więc z powyższych równości otrzymujemy, że

$$\overline{\sigma}_1^2 > \overline{\sigma}_2^2 .$$

Przypadek C. Z udowodnionych już przypadków A i B wynika, że nierówność 4^o wystarczy dowieść tu dla $N_2 = N_1 + 1$.

Jeżeli N_1 jest liczbą nieparzystą, to można napisać $N_1 = 2n + 1$, $N_2 = 2n + 2$, gdzie $n = 1, 2, \dots$. Wówczas ze wzorów (2.4) i (2.8) otrzymujemy, że

$$\frac{1}{2n+1} d(x, P_{2n+1,1}) = \frac{4(2n+1)x^2 - 4(2n+1)(a+b)x + 4n(a^2+b^2) + (a+b)^2}{2n(2n+1)(a-b)^2}$$

oraz

$$\frac{1}{2n+2} d(x, P_{2n+2,2}) = \frac{2x^2 - 2(a+b)x + (a^2+b^2)}{(n+1)(a-b)^2} .$$

Łącząc prawe strony powyższych równań znakiem większości otrzymujemy nierówność

$$\frac{4(2n+1)x^2 - 4(2n+1)(a+b)x + 4n(a^2+b^2) + (a+b)^2}{2n(2n+1)(a-b)^2} > \frac{2x^2 - 2(a+b)x + (a^2+b^2)}{(n+1)(a-b)^2}$$

którą można sprowadzić do postaci

$$4(2n+1)(x^2 - ax - bx) + (n+1)(a+b)^2 + 2n(a^2+b^2) > 0.$$

Lewa strona tej nierówności osiąga minimum dla $x = \frac{a+b}{2}$.

Przy tym minimalną wartością lewej strony jest $(a-b)^2$, co oznacza, że

$$(2.18) \quad \frac{1}{2n+1} d(x, P_{2n+1,1}) > \frac{1}{2n+2} d(x, P_{2n+2,2}) \quad \text{dla każdego} \\ x \in (-\infty, \infty).$$

Jeżeli N_1 jest liczbą parzystą, to można napisać $N_1 = 2n$, $N_2 = 2n + 1$, gdzie $n = 1, 2, \dots$. Wówczas ze wzorów (2.8) i (2.4) otrzymujemy, że

$$\frac{1}{2n} d(x, P_{2n,2}) = \frac{2x^2 - 2(a+b)x + (a^2+b^2)}{n(a-b)^2}$$

oraz

$$\frac{1}{2n+1} d(x, P_{2n+1,1}) = \frac{4(2n+1)x^2 - 4(2n+1)(a+b)x + 4n(a^2+b^2) + (a+b)^2}{2n(2n+1)(a-b)^2}$$

złącząc prawe strony powyższych równań znakiem większości otrzymujemy nierówność

$$\frac{2x^2 - 2(a+b)x + (a^2+b^2)}{n(a-b)^2} > \frac{4(2n+1)x^2 - 4(2n+1)(a+b)x + 4n(a^2+b^2) + (a+b)^2}{2n(2n+1)(a-b)^2},$$

którą można przekształcić do postaci

$$(a - b)^2 > 0.$$

Oznacza to, że

$$(2.19) \quad \frac{1}{2n} d(x, P_{2n,2}) > \frac{1}{2n+1} d(x, P_{2n+1,1}) \quad \text{dla każdego} \\ x \in (-\infty, \infty)$$

Ponieważ $\overline{\sigma_j^2} = \frac{\sigma^2}{b-a} \int_a^b \frac{1}{N_j} d(x, P_{N_j, i}) dx$ oraz

$\frac{1}{N_j} d(x, P_{N_j, i}) > 0$ dla każdego $x \in (-\infty, \infty)$, gdzie $i, j = 1, 2$, więc w rozważanym przypadku nierówność 4^o wynika natychmiast ze wzorów (2.18) i (2.19).

o.b.d.o.

Dla różnych liczb obserwacji wykonywanych według planu eksperymentalnego $P_{N,1}$ lub $P_{N,2}$ w zależności od tego, czy liczba obserwacji jest nieparzysta czy parzysta, podano w przedostatniej kolumnie tabeli 2.1 maksymalne wartości wariancji estymatora funkcji regresji (2.1) podzielone przez współczynnik σ^2 . W ostatniej kolumnie tej tabeli zamieszczono maksymalne wartości wariancji estymatora funkcji regresji (2.1) podzielone przez σ^2 , gdy N obserwacji wykonuje się w punktach $x^{(1)}, x^{(2)}, \dots, x^{(N)}$ otrzymanych z generatora liczb pseudolosowych o rozkładzie określonym gęstością

$$g_1(x) = \begin{cases} \frac{1}{2} & \text{dla } -1 \leq x \leq 1 \\ 0 & \text{poza tym} \end{cases}$$

tzn. o rozkładzie jednostajnym na odcinku $[-1, 1]$. Generowanie punktów o rozkładzie $g_1(x)$ nazwiemy tutaj sposobem losowym rozmieszczenia obserwacji. Dla ustalonej liczby N wartości maksymalne $\max_{x \in [-1, 1]} \text{Var}(\hat{E}Y_x) / \sigma^2$ podane w ostatniej kolumnie tabeli 2.1 są średnimi otrzymanymi w serii 30 eksperymentów komputerowych (por. rozdział VI).

Jeżeli kryterium dobroci eksperymentu mającego na celu estymację funkcji regresji $EY_x = \theta_1 + \theta_2 x$ na odcinku $[-1, 1]$ jest kryterium G - optymalności, to wyniki zamieszczone w tabeli 2.1 pozwalają ocenić korzyści stosowania

T A B E L A 2.1

Maksymalne wartości wariancji

estymatora funkcji regresji $EY_x = \theta_1 + \theta_2 x$

Ogólna liczba obser- wacji N	Sposób rozmieszczenia obserwacji				
	według planu $P_{N,1}$ albo $P_{N,2}$			losowy	
	Liczba obserwacji w punkcie			$\max_{x \in [a,b]} \text{Var}(\hat{EY}_x) / \sigma^2$	$\max_{x \in [-1,1]} \text{Var}(\hat{EY}_x) / \sigma^2$
	a	$\frac{a+b}{2}$	b		
2	1	-	1	1	58.777
3	1	1	1	.833	25.269
4	2	-	2	.500	3.803
5	2	1	2	.450	1.679
6	3	-	3	.333	1.589
7	3	1	3	.310	1.069
8	4	-	4	.250	.754
9	4	1	4	.236	.597
10	5	-	5	.200	.560
11	5	1	5	.191	.490
12	6	-	6	.167	.477
13	6	1	6	.160	.412
14	7	-	7	.143	.373
15	7	1	7	.138	.370
16	8	-	8	.125	.330
17	8	1	8	.121	.316
18	9	-	9	.111	.279
19	9	1	9	.108	.294
20	10	-	10	.100	.248
21	10	1	10	.098	.238
22	11	-	11	.091	.226
23	11	1	11	.089	.205
24	12	-	12	.083	.198
25	12	1	12	.082	.197
30	15	-	15	.067	.160
35	17	1	17	.058	.133
40	20	-	20	.050	.119
45	22	1	22	.045	.098
50	25	-	25	.040	.095
60	30	-	30	.033	.075
70	35	-	35	.029	.064
80	40	-	40	.025	.054
90	45	-	45	.022	.048
100	50	-	50	.020	.044
200	100	-	100	.010	.021
500	250	-	250	.004	.008

G - optymalnych planów eksperymentalnych $P_{N,1}$ lub $P_{N,2}$ w porównaniu ze sposobem losowym rozmieszczenia obserwacji. Np., gdy $N = 4$, to wykonanie obserwacji według planu $P_{4,2}$ daje wynik około 7,6 razy dokładniejszy, gdy $N = 15$, to wykonanie obserwacji według planu $P_{15,1}$ daje wynik około 2,7 razy dokładniejszy, gdy $N = 500$, to wykonanie obserwacji według planu $P_{500,2}$ daje wynik 2 razy dokładniejszy w porównaniu ze sposobem losowym rozmieszczenia czterech, piętnastu lub pięciuset obserwacji. Podobnie, jeśli np. żąda się, żeby $\max_{x \in [-1,1]} \text{Var}(\hat{E}Y_x) / \sigma^2 \leq \frac{1}{3}$, to dla spełnienia tego warunku potrzeba wykonać albo 6 obserwacji według G - optymalnego planu

$$\left\{ \begin{array}{cc} -1 & -1 \\ \frac{2}{6} & \frac{2}{6} \end{array} \right\},$$

albo też wykonać średnio 16 obserwacji w punktach rozmieszczonych losowo na odcinku $[-1, 1]$. Analogicznie, jeśli np. eksperymentator dysponuje środkami na wykonanie $N = 30$ obserwacji zmiennych losowych Y_x o wartości oczekiwanej $EY_x = \theta_1 + \theta_2 x$, to przy losowym sposobie rozmieszczenia 30 obserwacji na odcinku $[-1, 1]$ osiągnie on dokładność estymacji daną następującą przybliżoną równością

$$\max_{x \in [-1,1]} \text{Var}(\hat{E}Y_x) / \sigma^2 \approx 0,16 .$$

Tą samą dokładność można jednak osiągnąć wykonując o wiele mniej, bo tylko 13 obserwacji według planu

$$\left\{ \begin{array}{ccc} -1 & 0 & 1 \\ \frac{6}{13} & \frac{1}{13} & \frac{6}{13} \end{array} \right\},$$

co pozwala eksperymentatorowi zaoszczędzić znaczną część środków. Jeśli potrzeba oszczędzania jest zbyt duża, to w miarę możliwości eksperymentator powinien wykonać 30 obserwacji według G - optymalnego planu

$$\left\{ \begin{array}{cc} -1 & 1 \\ \frac{15}{30} & \frac{15}{30} \end{array} \right\}$$

co daje 2,4 razy większą dokładność estymacji.

Rozdział III

REGRESJA LINIOWA NA KOSTCE DWUWYMIAROWEJ (KWADRACIE)

Kostką dwuwymiarową nazywamy (por. określenie podane na początku rozdziału IV) zbiór $X = \{x = (x_1, x_2), -1 \leq x_1 \leq 1, i = 1, 2\}$. Liniowa funkcja regresji na tak określonym zbiorze X ma postać

$$(3.1) \quad EY_x = \theta_0 + \theta_1 x_1 + \theta_2 x_2 .$$

Dla tej funkcji regresji będziemy rozważali cztery następujące rodziny planów eksperymentalnych na kwadracie:

I. Jeżeli liczba obserwacji $N = 4k$, gdzie $k = 1, 2, \dots$, to bierzemy pod uwagę plan eksperymentalny $P_{N,10}$ określony następująco

$$(3.2) \quad \left\{ \begin{array}{cccc} (1, 1) & (1, -1) & (-1, 1) & (-1, -1) \\ \frac{k}{4k} & \frac{k}{4k} & \frac{-k}{4k} & \frac{k}{4k} \end{array} \right\},$$

II. Jeżeli liczba obserwacji $N = 4k + 1$, gdzie $k = 1, 2, \dots$, to rozważamy plan eksperymentalny $P_{N,11}$ określony następująco

$$(3.3) \quad \left\{ \begin{array}{cccc} (1, 1) & (1, -1) & (-1, 1) & (-1, -1) \\ \frac{k+1}{4k+1} & \frac{k}{4k+1} & \frac{k}{4k+1} & \frac{k}{4k+1} \end{array} \right\}.$$

III. Jeżeli $N = 4k + 2$, gdzie $k = 1, 2, \dots$, to definiujemy plan $P_{N,12}$ w sposób następujący

$$(3.4) \quad \left\{ \begin{array}{cccc} (1, 1) & (1, -1) & (-1, 1) & (-1, -1) \\ \frac{k+1}{4k+2} & \frac{k}{4k+2} & \frac{k}{4k+2} & \frac{k+1}{4k+2} \end{array} \right\} .$$

IV. Jeżeli $N = 4k+3$, gdzie $k = 0, 1, 2, \dots$, to plan $P_{N,13}$ określamy następująco

$$(3.5) \quad \left\{ \begin{array}{cccc} (1, 1) & (1, -1) & (-1, 1) & (-1, -1) \\ \frac{k+1}{4k+3} & \frac{k+1}{4k+3} & \frac{k}{4k+3} & \frac{k+1}{4k+3} \end{array} \right\} .$$

Plan eksperymentalny $P_{N,10}$ jest tu rozważany dlatego, że na mocy twierdzenia 7 (rozdział IV) jest to plan A -, D - i G - optymalny dla funkcji regresji (3.1) przy liczbie obserwacji $N = 4k$ ($k = 1, 2, \dots$). Ograniczenie się do planów $P_{N,10} - P_{N,13}$ ma tę własność, że po wykonaniu N obserwacji według jednego z nich istnieje możliwość wykonania dodatkowych d obserwacji w taki sposób, jak gdyby łączna liczba $N + d$ obserwacji była wykonana znów według jednego z planów eksperymentalnych $P_{N,10} - P_{N,13}$.

Podamy teraz niektóre charakterystyki planów eksperymentalnych $P_{N,10} - P_{N,13}$.

I. Plan eksperymentalny $P_{N,10}$:

$$(3.6) \quad M(P_{N,10}) = M^{-1}(P_{N,10}) = \begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{bmatrix}$$

$$\sum_{i=1}^3 m^{ii}(P_{N,10}) = 3$$

$$(3.7) \quad \det M(P_{N,10}) = 1$$

$$d(x, P_{N,10}) = 1 + x_1^2 + x_2^2$$

$$(3.8) \quad \max_{x \in X} d(x, P_{N,10}) = 3$$

$$(3.9) \quad \frac{1}{4} \int_{-1}^1 \int_{-1}^1 \text{Var} (\hat{E}Y_x) dx_1 dx_2 = \\ = \frac{\sigma^2}{4} \int_{-1}^1 \int_{-1}^1 \frac{1}{4k} d(x, P_{N,10}) dx_1 dx_2 = \sigma^2 \frac{5}{12k} .$$

II. Plan eksperymentalny $P_{N,11}$:

$$M(P_{N,11}) = \begin{bmatrix} 1 & \frac{1}{4k+1} & 1 \\ \frac{1}{4k+1} & 1 & \frac{1}{4k+1} \\ \frac{1}{4k+1} & \frac{1}{4k+1} & 1 \end{bmatrix}$$

$$M^{-1}(P_{N,11}) = \begin{bmatrix} \frac{8k^2+6k+1}{8k^2+6k} & \frac{-(4k+1)}{2(8k^2+6k)} & \frac{-(4k+1)}{2(8k^2+6k)} \\ \frac{-(4k+1)}{2(8k^2+6k)} & \frac{8k^2+6k+1}{8k^2+6k} & \frac{-(4k+1)}{2(8k^2+6k)} \\ \frac{-(4k+1)}{2(8k^2+6k)} & \frac{-(4k+1)}{2(8k^2+6k)} & \frac{8k^2+6k+1}{8k^2+6k} \end{bmatrix}$$

$$(3.10) \quad \sum_{i=1}^3 m^{ii}(P_{N,11}) = 3 \cdot \frac{8k^2+6k+1}{8k^2+6k}$$

$$(3.11) \quad \det M(P_{N,11}) = 1 + \frac{8k-1}{(4k+1)^3}$$

$$d(x, P_{N,11}) = \frac{(8k^2+6k+1)(x_1^2+x_2^2+1) - (4k+1)(x_1+x_1x_2+x_2)}{8k^2+6k}$$

$$(3.12) \quad \max_{x \in X} d(x, P_{N,11}) = 3 + \frac{2k+2}{4k^2+3k}$$

$$(3.13) \quad \frac{1}{4} \int_{-1}^1 \int_{-1}^1 \text{Var} (\hat{E}Y_x) dx_1 dx_2 =$$

$$= \frac{\sigma^2}{4} \int_{-1}^1 \int_{-1}^1 \frac{1}{4k+1} d(x, P_{N,11}) dx_1 dx_2 = \sigma^2 \frac{5(2k+1)}{6(4k^2+3k)}$$

III. Plan eksperymentalny $P_{N,12}$:

$$M(P_{N,12}) = \begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & \frac{1}{2k+1} \\ 0 & \frac{1}{2k+1} & 1 \end{bmatrix}$$

$$M^{-1}(P_{N,12}) = \begin{bmatrix} 1 & 0 & 0 \\ 0 & \frac{(2k+1)^2}{(2k+1)^2-1} & \frac{-(2k+1)}{(2k+1)^2-1} \\ 0 & \frac{-(2k+1)}{(2k+1)^2-1} & \frac{(2k+1)^2}{(2k+1)^2-1} \end{bmatrix}$$

$$(3.14) \quad \sum_{i=1}^3 m^{11}(P_{N,12}) = 1 + \frac{(2k+1)^2}{2k(k+1)}$$

$$(3.15) \quad \det M(P_{N,12}) = 1 - \frac{1}{(2k+1)^2}$$

$$d(x, P_{N,12}) = \frac{(2k+1)^2(x_1^2+x_2^2) - 2(2k+1)x_1x_2}{4k(k+1)} + 1$$

$$(3.16) \quad \max_{x \in X} d(x, P_{N,12}) = 3 + \frac{1}{k}$$

$$(3.17) \quad \frac{1}{4} \int_{-1}^1 \int_{-1}^1 \text{Var} (\hat{E}Y_x) dx_1 dx_2 =$$

$$= \frac{6^2}{4} \int_{-1}^1 \int_{-1}^1 \frac{1}{4k+2} d(x, P_{N,12}) dx_1 dx_2 = 6^2 \cdot \frac{10k^2+10k+1}{12k(k+1)(2k+1)}$$

IV. Plan eksperymentalny $P_{N,13}$:

$$M(P_{N,13}) = \begin{bmatrix} 1 & \frac{1}{4k+3} & \frac{-1}{4k+3} \\ \frac{1}{4k+3} & 1 & \frac{1}{4k+3} \\ \frac{-1}{4k+3} & \frac{1}{4k+3} & 1 \end{bmatrix}$$

$$M^{-1}(P_{N,13}) = \begin{bmatrix} \frac{(4k+3)(2k+1)}{2(4k+1)(k+1)} & \frac{-(4k+3)}{4(4k+1)(k+1)} & \frac{4k+3}{4(4k+1)(k+1)} \\ \frac{-(4k+3)}{4(4k+1)(k+1)} & \frac{(4k+3)(2k+1)}{2(4k+1)(k+1)} & \frac{-(4k+3)}{4(4k+1)(k+1)} \\ \frac{4k+3}{4(4k+1)(k+1)} & \frac{-(4k+3)}{4(4k+1)(k+1)} & \frac{(4k+3)(2k+1)}{2(4k+1)(k+1)} \end{bmatrix}$$

$$(3.18) \quad \sum_{i=1}^3 m^{ii}(P_{N,13}) = \frac{3(4k+3)(2k+1)}{2(4k+1)(k+1)}$$

$$(3.19) \quad \det M(P_{N,13}) = 1 - \frac{12k+11}{(4k+3)^3}$$

$$d(x, P_{N,13}) = \frac{(4k+3)(2k+1)(x_1^2+x_2^2+1) + (4k+3)(x_2-x_1)(x_2-x_1)}{2(4k+1)(k+1)}$$

$$(3.20) \quad \max_{x \in X} d(x, P_{N,13}) = 3 + \frac{6}{4k+1}$$

$$(3.21) \quad \frac{1}{4} \int_{-1}^1 \int_{-1}^1 \text{Var}(\hat{E}Y_x) dx_1 dx_2 =$$

$$= \frac{\sigma^2}{4} \int_{-1}^1 \int_{-1}^1 \frac{1}{4k+3} d(x, P_{N,13}) dx_1 dx_2 = \sigma^2 \cdot \frac{5(2k+1)}{6(4k+1)(k+1)}$$

Dla planów eksperymentalnych $P_{N,10} - P_{N,13}$ zachodzi poniższe

TWIERDZENIE 6. Niech N_1 i N_2 oznaczają dwie liczby naturalne takie, że $N_2 > N_1 \geq 3$. Jeżeli w celu estymacji funkcji regresji (3.1) na kwadracie $X = \{x = (x_1, x_2), -1 \leq x_i \leq 1, i = 1, 2\}$, wykonano N_1 obserwacji według planu eksperymentalnego $P_{N_1,i}$, gdzie $i = 10 + N_1 \pmod{4}$, tzn. N_1 obserwacji wykonano według jednego z planów $P_{N_1,10} - P_{N_1,13}$, a ponadto jeśli wykonano N_2 obserwacji według planu eksperymentalnego $P_{N_2,i}$, gdzie $i = 10 + N_2 \pmod{4}$, tzn. N_2 obserwacji wykonano według jednego z planów $P_{N_2,10} - P_{N_2,13}$, to :

$$1^{\circ} \quad \left(\sum_{l=0}^2 \text{Var } \hat{\theta}_l \right)_1 > \left(\sum_{l=0}^2 \text{Var } \hat{\theta}_l \right)_2,$$

gdzie $\left(\sum_{l=0}^2 \text{Var } \hat{\theta}_l \right)_j$ oznacza sumę wariancji estymatorów $\hat{\theta}_0, \hat{\theta}_1$ i $\hat{\theta}_2$, gdy wykonano N_j ($j = 1, 2$) obserwacji;

$$2^{\circ} \quad \text{Vol} (\hat{\theta}_0, \hat{\theta}_1, \hat{\theta}_2)_1 > \text{Vol} (\hat{\theta}_0, \hat{\theta}_1, \hat{\theta}_2)_2,$$

gdzie $\text{Vol} (\hat{\theta}_0, \hat{\theta}_1, \hat{\theta}_2)_j$ oznacza objętość elipsoidy koncentracji rozkładu estymatorów $\hat{\theta}_0, \hat{\theta}_1, \hat{\theta}_2$, gdy wykonano N_j ($j = 1, 2$) obserwacji;

$$3^{\circ} \quad \sigma_1^2 > \sigma_2^2,$$

gdzie $\sigma_j^2 = \max_{x \in X} \text{Var} (\hat{E}Y_x) = \frac{\sigma^2}{N_j} \max_{x \in X} d(x, P_{N_j,i})$ oznacza

maksymalną wartość wariancji estymatora funkcji regresji

(3.1) na kwadracie X , gdy N_j obserwacji wykonano według planu $P_{N_j, i}$ ($i = 10, 11, 12, 13$; $j = 1, 2$);

$$4^0 \quad \overline{\sigma}_1^2 > \overline{\sigma}_2^2,$$

$$\begin{aligned} \text{gdzie } \overline{\sigma}_j^2 &= \frac{1}{4} \int_{-1}^1 \int_{-1}^1 \text{Var}(\hat{E}Y_x) dx_1 dx_2 = \\ &= \frac{\sigma^2}{4} \int_{-1}^1 \int_{-1}^1 \frac{1}{N_j} d(x, P_{N_j, i}) dx_1 dx_2 \end{aligned}$$

oznacza wartość średnią wariancji estymatora funkcji regresji (3.1) na kwadracie X , gdy N_j obserwacji wykonano według planu $P_{N_j, i}$ ($i = 10, 11, 12, 13$; $j = 1, 2$).

Dowód.

Udowodnimy najpierw nierówność 1^0 .

Ze wzoru (1.5) wynika, że

$$\left(\sum_{i=0}^2 \text{Var } \hat{\theta}_1 \right)_j = \frac{\sigma^2}{N_j} \left[m^{11}(P_{N_j, i}) + m^{22}(P_{N_j, i}) + m^{33}(P_{N_j, i}) \right].$$

Wobec tego prawdziwość nierówności 1^0 zostanie udowodniona, jeżeli dla dowolnie ustalonej liczby naturalnej k pokażemy, że

$$(3.22) \quad \frac{\sigma^2}{4k} \sum_{i=1}^3 m^{ii}(P_{4k, 10}) > \frac{\sigma^2}{4k+1} \sum_{i=1}^3 m^{ii}(P_{4k+1, 11}) >$$

$$\frac{\sigma^2}{4k+2} \sum_{i=1}^3 m^{ii}(P_{4k+2, 12}) > \frac{\sigma^2}{4k+3} \sum_{i=1}^3 m^{ii}(P_{4k+3, 13}) >$$

$$\frac{\sigma^2}{4k+4} \sum_{i=1}^3 m^{ii}(P_{4k+4, 10})$$

oraz, że

$$(3.23) \quad \frac{\sigma^2}{3} \sum_{i=1}^3 m^{ii}(P_{3,13}) > \frac{\sigma^2}{4} \sum_{i=1}^3 m^{ii}(P_{4,10}) .$$

Obliczając $\sum_{i=1}^3 m^{ii}(P_{4k,10})$, $\sum_{i=1}^3 m^{ii}(P_{4k+4,10})$ i $\sum_{i=1}^3 m^{ii}(P_{4,10})$ ze wzoru (3.6), $\sum_{i=1}^3 m^{ii}(P_{4k+1,11})$ ze wzoru (3.10), $\sum_{i=1}^3 m^{ii}(P_{4k+2,12})$ ze wzoru (3.14), natomiast

$\sum_{i=1}^3 m^{ii}(P_{4k+3,13})$ i $\sum_{i=1}^3 m^{ii}(P_{3,13})$ ze wzoru (3.18), otrzymujemy równoważną nierówności (3.22) nierówność

$$(3.24) \quad \sigma^2 \frac{3}{4k} > \sigma^2 \cdot \frac{3(2k+1)}{2k(4k+3)} > \sigma^2 \cdot \frac{6k^2+6k+1}{4k(2k+1)(k+1)} > \\ \sigma^2 \cdot \frac{3(2k+1)}{2(4k+1)(k+1)} > \sigma^2 \cdot \frac{3}{4(k+1)} ,$$

oraz równoważną nierówności (3.23) nierówność

$$\frac{3}{2} \sigma^2 > \frac{3}{4} \sigma^2 .$$

Ostatnia nierówność jest oczywista, natomiast prawdziwość nierówności poczwórnej (3.24) wynika łatwo po przekształceniu każdej z czterech tworzących ją nierówności pojedynczych do postaci typu $\sigma^2 h(k) > 0$, co kończy dowód nierówności 1^o.

Udowodnimy teraz nierówność 2^o.

Ze wzoru (1.6) wynika, że nierówność 2^o jest równoważna nierówności

$$(3.25) \quad N_1^3 \cdot \det M(P_{N_1,i}) < N_2^3 \cdot \det M(P_{N_2,l})$$

gdzie $i, l = 10, 11, 12, 13$. Prawdziwość nierówności

(3.25) zostanie udowodniona, jeżeli dla dowolnie ustalonej liczby naturalnej k pokażemy, że

$$(3.26) \quad (4k)^3 \det M(P_{4k,10}) < (4k+1)^3 \det M(P_{4k+1,11}) < \\ (4k+2)^3 \det M(P_{4k+2,12}) < (4k+3)^3 \det M(P_{4k+3,13}) < \\ (4k+4)^3 \det M(P_{4k+4,10})$$

oraz, że

$$(3.27) \quad 3^3 \cdot \det M(P_{3,13}) < 4^3 \det M(P_{4,10}) .$$

Obliczając $\det M(P_{4k,10})$, $\det M(P_{4k+4,10})$ i $\det M(P_{4,10})$ ze wzoru (3.7), $\det M(P_{4k+1,11})$ ze wzoru (3.11), $\det M(P_{4k+2,12})$ ze wzoru (3.15), natomiast $\det M(P_{4k+3,13})$ i $\det M(P_{3,13})$ ze wzoru (3.19), otrzymujemy równoważną nierówności (3.26) nierówność

$$64k^3 < 64k^3 + 48k^2 + 20k < 64k^3 + 96k^2 + 32k < \\ 64k^3 + 144k^2 + 96k + 16 < 64k^3 + 192k^2 + 192k + 64$$

oraz równoważną nierówności (3.27) nierówność

$$16 < 64 .$$

Prawdziwość obu otrzymanych nierówności jest oczywista, co kończy dowód nierówności (3.25), a tym samym dowód nierówności 2^o.

Z kolei udowodnimy nierówność 3^o.

Prawdziwość nierówności 3^o zostanie udowodniona, jeżeli dla dowolnie ustalonej liczby naturalnej k pokażemy, że

$$(3.28) \quad \frac{\sigma^2}{4k} \max_{x \in X} d(x, P_{4k,10}) > \frac{\sigma^2}{4k+1} \max_{x \in X} d(x, P_{4k+1,11}) >$$

$$\frac{\sigma^2}{4k+2} \max_{x \in X} d(x, P_{4k+2, 12}) > \frac{\sigma^2}{4k+3} \max_{x \in X} d(x, P_{4k+3, 13}) > \frac{\sigma^2}{4k+4} \max_{x \in X} d(x, P_{4k+4, 10})$$

oraz, że

$$(3.29) \quad \frac{\sigma^2}{3} \max_{x \in X} d(x, P_{3, 13}) > \frac{\sigma^2}{4} \max_{x \in X} d(x, P_{4, 10})$$

Obliczając $\max_{x \in X} d(x, P_{4k, 10})$, $\max_{x \in X} d(x, P_{4k+4, 10})$ i $\max_{x \in X} d(x, P_{4, 10})$ ze wzoru (3.8), $\max_{x \in X} d(x, P_{4k+1, 11})$ ze wzoru (3.12), $\max_{x \in X} d(x, P_{4k+2, 12})$ ze wzoru (3.16), natomiast $\max_{x \in X} d(x, P_{4k+3, 13})$ i $\max_{x \in X} d(x, P_{3, 13})$ ze wzoru (3.20), otrzymujemy równoważną nierówności (3.28) nierówność

$$(3.30) \quad \sigma^2 \frac{3}{4k} > \sigma^2 \frac{3k+2}{k(4k+3)} > \sigma^2 \frac{3k+1}{2k(2k+1)} > \sigma^2 \cdot \frac{3}{4k+1} > \sigma^2 \frac{3}{4(k+1)}$$

oraz równoważną nierówności (3.29) nierówność

$$3 \sigma^2 > \frac{3}{4} \sigma^2$$

Ostatnia nierówność jest oczywista, natomiast prawdziwość nierówności poczwórnej (3.30) wynika łatwo po przekształceniu każdej z czterech tworzących ją nierówności pojedynczych do postaci typu $\sigma^2 h(k) > 0$, co kończy dowód nierówności 3^o.

Na zakończenie dowodu twierdzenia 6 udowodnimy nierówność 4^o. Prawdziwość nierówności 4^o zostanie udowodniona, jeśli dla dowolnie ustalonej liczby naturalnej k pokażemy, że

$$\begin{aligned}
 (3.31) \quad & \int_{-1}^1 \int_{-1}^1 \frac{1}{4k} d(x, P_{4k, 10}) dx_1 dx_2 > \\
 & \int_{-1}^1 \int_{-1}^1 \frac{1}{4k+1} d(x, P_{4k+1, 11}) dx_1 dx_2 > \\
 & \int_{-1}^1 \int_{-1}^1 \frac{1}{4k+2} d(x, P_{4k+2, 12}) dx_1 dx_2 > \\
 & \int_{-1}^1 \int_{-1}^1 \frac{1}{4k+3} d(x, P_{4k+3, 13}) dx_1 dx_2 > \\
 & \int_{-1}^1 \int_{-1}^1 \frac{1}{4k+4} d(x, P_{4k+4, 10}) dx_1 dx_2
 \end{aligned}$$

oraz, że

$$\begin{aligned}
 (3.32) \quad & \int_{-1}^1 \int_{-1}^1 \frac{1}{3} d(x, P_{3, 13}) dx_1 dx_2 > \\
 & \int_{-1}^1 \int_{-1}^1 \frac{1}{4} d(x, P_{4, 10}) dx_1 dx_2 .
 \end{aligned}$$

Wykorzystując w celu obliczenia całek podwójnych kolejno wzory (3.9), (3.13), (3.17), (3.21) i (3.9) oraz (3.21) i (3.9), otrzymujemy równoważną nierówności (3.31) nierówność

$$\begin{aligned}
 (3.33) \quad & \frac{5}{3k} > \frac{10(2k+1)}{3k(4k+3)} > \frac{10k^2+10k+1}{3k(k+1)(2k+1)} > \\
 & \frac{10(2k+1)}{3(4k+1)(k+1)} > \frac{5}{3(k+1)}
 \end{aligned}$$

oraz równoważną nierówności (3.32) nierówność

$$\frac{10}{3} > \frac{5}{3} .$$

Ostatnia nierówność jest oczywista, natomiast prawdziwość poczwórnej nierówności (3.33) wynika łatwo po przekształ-

ceniu każdej z czterech tworzących ją nierówności pojedynczych do postaci typu $\sigma^2 h(k) > 0$, co kończy dowód nierówności 4^o.

c.b.d.o.

Zanotujmy, że dla planu eksperymentalnego $P_{N,12a}$ określonego następująco

$$\left\{ \begin{array}{cccc} (1, 1) & (1, -1) & (-1, 1) & (-1, -1) \\ \frac{k+1}{4k+2} & \frac{k+1}{4k+2} & \frac{k}{4k+2} & \frac{k}{4k+2} \end{array} \right\},$$

gdzie $N = 4k + 2$, $k = 1, 2, \dots$, mamy

$$(3.34) \quad \sum_{i=1}^3 m^{ii}(P_{N,12a}) = 1 + \frac{(2k+1)^2}{2k(k+1)},$$

$$(3.35) \quad \det M(P_{N,12a}) = 1 - \frac{1}{(2k+1)^2},$$

$$(3.36) \quad \max_{x \in X} d(x, P_{N,12a}) = 3 + \frac{1}{k},$$

$$(3.37) \quad \frac{1}{4} \int_{-1}^1 \int_{-1}^1 \text{Var}(\hat{E}Y_x) dx_1 dx_2 =$$

$$\frac{\sigma^2}{4} \int_{-1}^1 \int_{-1}^1 \frac{1}{4k+2} d(x, P_{N,12a}) dx_1 dx_2 =$$

$$= \sigma^2 \frac{5k^2 + 5k + 1}{6k(k+1)(2k+1)}.$$

Porównując prawa strony wzorów (3.14) - (3.17) odpowiednio z prawymi stronami wzorów (3.34) - (3.37) otrzymujemy, że

$$\sum_{i=1}^3 m^{ii}(P_{N,12}) = \sum_{i=1}^3 m^{ii}(P_{N,12a})$$

$$\det M(P_{N,12}) = \det M(P_{N,12a})$$

$$\max_{x \in X} d(x, P_{N,12}) = \max_{x \in X} d(x, P_{N,12a})$$

$$\frac{\sigma^2}{4} \int_{-1}^1 \int_{-1}^1 \frac{1}{4k+2} d(x, P_{N,12}) dx_1 dx_2 >$$

$$\frac{\sigma^2}{4} \int_{-1}^1 \int_{-1}^1 \frac{1}{4k+2} d(x, P_{N,12a}) dx_1 dx_2 .$$

Oznacza to, że plany eksperymentalne $P_{N,12}$ i $P_{N,12a}$ dla $N = 4k + 2$ ($k = 1, 2, \dots$) obserwacji są równoważne z punktu widzenia kryteriów A -, D - i G - optymalności, natomiast z punktu widzenia kryterium I - optymalności nieznacznie¹⁾ lepszy jest plan $P_{N,12}$.

Dodać trzeba, że na drodze analogicznych porównań dochodzi się do wniosku, iż z punktu widzenia kryteriów A -, D -, G - oraz I - optymalności planowi eksperymentalnemu $P_{N,11}$ (por. wzór (3.3)) równoważne są 3 następujące plany:

$$\left\{ \begin{array}{cccc} (1, 1) & (1, -1) & (-1, 1) & (-1, -1) \\ \frac{k}{4k+1} & \frac{k+1}{4k+1} & \frac{k}{4k+1} & \frac{k}{4k+1} \end{array} \right\},$$

$$\left\{ \begin{array}{cccc} (1, 1) & (1, -1) & (-1, 1) & (-1, -1) \\ \frac{k}{4k+1} & \frac{k}{4k+1} & \frac{k+1}{4k+1} & \frac{k}{4k+1} \end{array} \right\},$$

¹⁾Wartość średnia wariancji estymatora funkcji regresji (3.1) na kostce dwuwymiarowej, gdy $N = 4k+2$ obserwacji wykonano według planu $P_{N,12a}$, jest większa co najwyżej o $\frac{1}{21}$ analogicznej wartości dla planu $P_{N,12}$.

$$\left\{ \begin{array}{cccc} (1, 1) & (1, -1) & (-1, 1) & (-1, -1) \\ \frac{k}{4k+1} & \frac{k}{4k+1} & \frac{k}{4k+1} & \frac{k+1}{4k+1} \end{array} \right\};$$

planowi eksperymentalnemu $P_{N,12}$ (por. wzór (3.4)) równoważny jest plan określony następująco:

$$\left\{ \begin{array}{cccc} (1, 1) & (1, -1) & (-1, 1) & (-1, -1) \\ \frac{k}{4k+2} & \frac{k+1}{4k+2} & \frac{k+1}{4k+2} & \frac{k}{4k+2} \end{array} \right\};$$

planowi eksperymentalnemu $P_{N,12a}$ równoważne są 3 następujące plany:

$$\left\{ \begin{array}{cccc} (1, 1) & (1, -1) & (-1, 1) & (-1, -1) \\ \frac{k+1}{4k+2} & \frac{k}{4k+2} & \frac{k+1}{4k+2} & \frac{k}{4k+2} \end{array} \right\}$$

$$\left\{ \begin{array}{cccc} (1, 1) & (1, -1) & (-1, 1) & (-1, -1) \\ \frac{k}{4k+2} & \frac{k+1}{4k+2} & \frac{k}{4k+2} & \frac{k+1}{4k+2} \end{array} \right\};$$

$$\left\{ \begin{array}{cccc} (1, 1) & (1, -1) & (-1, 1) & (-1, -1) \\ \frac{k}{4k+2} & \frac{k}{4k+2} & \frac{k+1}{4k+2} & \frac{k+1}{4k+2} \end{array} \right\};$$

planowi eksperymentalnemu $P_{N,13}$ (por. wzór (3.5)) równoważne są 3 następujące plany:

$$\left\{ \begin{array}{cccc} (1, 1) & (1, -1) & (-1, 1) & (-1, -1) \\ \frac{k}{4k+3} & \frac{k+1}{4k+3} & \frac{k+1}{4k+3} & \frac{k+1}{4k+3} \end{array} \right\};$$

$$\left\{ \begin{array}{cccc} (1, 1) & (1, -1) & (-1, 1) & (-1, -1) \\ \frac{k+1}{4k+3} & \frac{k}{4k+3} & \frac{k+1}{4k+3} & \frac{k+1}{4k+3} \end{array} \right\};$$

$$\left\{ \begin{array}{cccc} (1, 1) & (1, -1) & (-1, 1) & (-1, -1) \\ \frac{k+1}{4k+3} & \frac{k+1}{4k+3} & \frac{k+1}{4k+3} & \frac{k}{4k+3} \end{array} \right\}.$$

Plan eksperymentalny $P_{N,12a}$ oraz plany eksperymentalne równoważne z punktu widzenia kryteriów A -, D -, G - oraz I - optymalności odpowiednio planom $P_{N,10}$, $P_{N,11}$, $P_{N,12}$, $P_{N,12a}$ i $P_{N,13}$ zostały tu przedstawione oczywiście dlatego, że po wykonaniu N obserwacji według jednego z tych planów istnieje możliwość wykonania dodatkowych d obserwacji w taki sposób, jak gdyby łączna liczba $N+d$ obserwacji była wykonana według jednego z planów eksperymentalnych przedstawionych dotychczas w rozdziale III.

Plany eksperymentalne $P_{N,11}$, $P_{N,12}$ / $P_{N,12a}$ i $P_{N,13}$ będziemy teraz porównywać z odpowiednimi elementami pewnej rodziny symetrycznych planów eksperymentalnych dla wykonania $N = 4k + 1$ ($k = 1, 2, \dots$; $i = 1, 2, 3$) obserwacji na kostce dwuwymiarowej. Elementy $P_{N,0i}$ tej rodziny definiujemy następująco:

$$\left\{ \begin{array}{ccccc} (1, 1) & (1, -1) & (-1, 1) & (-1, -1) & (0, 0) \\ \frac{k}{4k+1} & \frac{k}{4k+1} & \frac{k}{4k+1} & \frac{k}{4k+1} & \frac{i}{4k+1} \end{array} \right\}.$$

Rodzina $P_{N,0i}$ dzieli się na 3 podrodziny $P_{N,01}$, $P_{N,02}$, $P_{N,03}$ planów eksperymentalnych. Plany $P_{N,01}$ pozwalają wykonać $N = 4k + 1$ ($k = 1, 2, \dots$) obserwacji, plany $P_{N,02}$ umożliwiają wykonanie $N = 4k + 2$ obserwacji, zaś plany $P_{N,03}$ umożliwiają wykonanie $N = 4k + 3$ obserwacji na kostce dwuwymiarowej.

Oto niektóre charakterystyki planów P_{N, O_1} :

$$M(P_{N, O_1}) = \begin{bmatrix} 1 & 0 & 0 \\ 0 & \frac{4k}{4k+1} & 0 \\ 0 & 0 & \frac{4k}{4k+1} \end{bmatrix}$$

$$M^{-1}(P_{N, O_1}) = \begin{bmatrix} 1 & 0 & 0 \\ 0 & \frac{4k+1}{4k} & 0 \\ 0 & 0 & \frac{4k+1}{4k} \end{bmatrix}$$

$$(3.38) \quad \sum_{j=1}^3 m^{jj}(P_{N, O_1}) = 1 + \frac{4k+1}{2k}$$

$$(3.39) \quad \det M(P_{N, O_1}) = \frac{16k^2}{(4k+1)^2}$$

$$d(x, P_{N, O_1}) = \frac{4k+1}{4k} (x_1^2 + x_2^2) + 1$$

$$(3.40) \quad \max_{x \in X} d(x, P_{N, O_1}) = 3 + \frac{1}{2k}$$

$$(3.41) \quad \frac{1}{4} \int_{-1}^1 \int_{-1}^1 \text{Var}(\hat{E}Y_x) dx_1 dx_2 =$$

$$= \frac{\sigma^2}{4} \int_{-1}^1 \int_{-1}^1 \frac{1}{4k+1} d(x, P_{N, O_1}) dx_1 dx_2 =$$

$$= \sigma^2 \left(\frac{1}{6k} + \frac{1}{4k+1} \right)$$

Zauważmy, że dla wykonania $N = 4k + 3$ ($k = 1, 2, \dots$) obserwacji w pięciu punktach $(1, 1)$, $(1, -1)$, $(-1, 1)$, $(-1, -1)$, $(0, 0)$ kostki dwuwymiarowej można, oprócz pla-

nu $P_{N,03}$, skonstruować jeszcze inne plany symetryczne względem początku układu współrzędnych, np. plan eksperymentalny $P_{N,03a}$ określony następująco

$$\left\{ \begin{array}{ccccc} (1, 1) & (1, -1) & (-1, 1) & (-1, -1) & (0, 0) \\ \frac{k+1}{4k+3} & \frac{k}{4k+3} & \frac{k}{4k+3} & \frac{k+1}{4k+3} & \frac{1}{4k+3} \end{array} \right\}$$

Dla planu $P_{N,03a}$ mamy:

$$M(P_{N,03a}) = \begin{bmatrix} 1 & 0 & 0 \\ 0 & \frac{4k+2}{4k+3} & \frac{2}{4k+3} \\ 0 & \frac{2}{4k+3} & \frac{4k+2}{4k+3} \end{bmatrix}$$

$$M^{-1}(P_{N,03a}) = \begin{bmatrix} 1 & 0 & 0 \\ 0 & \frac{(4k+3)(2k+1)}{8k(k+1)} & -\frac{4k+3}{8k(k+1)} \\ 0 & -\frac{4k+3}{8k(k+1)} & \frac{(4k+3)(2k+1)}{8k(k+1)} \end{bmatrix}$$

$$(3.42) \quad \sum_{i=1}^3 m^{ii}(P_{N,03a}) = 1 + \frac{(4k+3)(2k+1)}{4k(k+1)}$$

$$(3.43) \quad \det M(P_{N,03a}) = \frac{16k(k+1)}{(4k+3)^2}$$

$$d(x, P_{N,03a}) = 1 + \frac{(4k+3)(2k+1)}{8k(k+1)}(x_1^2 + x_2^2) - \frac{4k+3}{4k(k+1)}x_1x_2$$

$$(3.44) \quad \max_{x \in X} d(x, P_{N,03a}) = 1 + \frac{4k+3}{2k}$$

$$(3.45) \quad \frac{1}{4} \int_{-1}^1 \int_{-1}^1 \text{var}(\hat{E}Y_x) dx_1 dx_2 =$$

$$= \frac{\sigma^2}{4} \int_{-1}^1 \int_{-1}^1 \frac{1}{4k+3} d(x, P_{N, O3a}) dx_1 dx_2 = \sigma^2 \frac{20k^2+22k+3}{12k(k+1)(4k+3)}$$

Porównując wzory (3.38) - (3.41), w których $i = 3$, odpowiednio ze wzorami (3.42) - (3.45) otrzymujemy, że

$$\sum_{i=1}^3 m^{ii}(P_{N, O3}) > \sum_{i=1}^3 m^{ii}(P_{N, O3a})$$

$$\det M(P_{N, O3}) < \det M(P_{N, O3a})$$

$$\max_{x \in X} d(x, P_{N, O3}) = \max_{x \in X} d(x, P_{N, O3a})$$

$$\frac{\sigma^2}{4} \int_{-1}^1 \int_{-1}^1 \frac{1}{4k+3} d(x, P_{N, O3}) dx_1 dx_2 >$$

$$\frac{\sigma^2}{4} \int_{-1}^1 \int_{-1}^1 \frac{1}{4k+3} d(x, P_{N, O3a}) dx_1 dx_2 .$$

Powyższe równania i nierówności oznaczają, że w sensie kryteriów A -, D - oraz I - optymalności plan eksperymentalny $P_{N, O3a}$ jest lepszy od planu eksperymentalnego $P_{N, O3}$, natomiast w sensie kryterium G - optymalności plany $P_{N, O3}$ i $P_{N, O3a}$ są równoważne.

Dodać trzeba, że plan eksperymentalny $P_{N, O3b}$ określony następująco

$$\left\{ \begin{array}{ccccc} (1, 1) & (1, -1) & (-1, 1) & (-1, -1) & (0, 0) \\ \frac{k}{4k+3} & \frac{k+1}{4k+3} & \frac{k+1}{4k+3} & \frac{k}{4k+3} & \frac{1}{4k+3} \end{array} \right\} .$$

jest w sensie kryteriów A -, D -, G - oraz I - optymalności równoważny planowi eksperymentalnemu $P_{N, O3a}$.

Porównywanie planów eksperymentalnych $P_{N,11}$,
 $P_{N,12} / P_{N,12a}$, $P_{N,13}$ z planami $P_{N,01}$, $P_{N,02}$,
 $P_{N,03} / P_{N,03a}$ ze względu na A - optymalność.

Dla każdej ustalonej liczby naturalnej k porównanie odpowiedniego planu $P_{N,11}$ dla $N = 4k + 1$ obserwacji z planem $P_{N,01}$ dla tej samej liczby obserwacji polega na porównaniu wzoru (3.10) ze wzorem (3.38), gdzie $i = 1$. Podobnie porównanie planu $P_{N,12}$ dla $N = 4k + 2$ obserwacji z planem $P_{N,02}$ dla tejże liczby obserwacji sprowadza się do porównania wzoru (3.14) ze wzorem (3.38), gdzie $i = 2$. Wreszcie plan $P_{N,13}$ dla $N = 4k + 3$ obserwacji porównuje się z planem $P_{N,03a}$ przez porównanie wzoru (3.18) ze wzorem (3.42). Łatwo dochodzi się wtedy do wniosku, że dla każdej ustalonej liczby k ($k=1,2,\dots$) ze względu na minimalizację sumy (lub średniej) wariancji estymatorów $\hat{\theta}_0, \hat{\theta}_1, \hat{\theta}_2$ współczynników funkcji regresji (3.1) plan eksperymentalny $P_{N,11}$ jest lepszy od planu $P_{N,01}$, plany $P_{N,12}$ i $P_{N,12a}$ są lepsze od planu $P_{N,02}$, plan $P_{N,13}$ jest lepszy od planu $P_{N,03a}$ oraz tym bardziej od planu $P_{N,03}$.

Porównywanie planów eksperymentalnych $P_{N,11}$,
 $P_{N,12} / P_{N,12a}$, $P_{N,13}$ z planami $P_{N,01}$, $P_{N,02}$,
 $P_{N,03} / P_{N,03a}$ ze względu na D - optymalność.

Dla każdej ustalonej liczby naturalnej k porównanie odpowiedniego planu $P_{N,11}$ dla $N = 4k + 1$ obserwacji z planem $P_{N,01}$ dla tej samej liczby obserwacji polega na zbadaniu relacji pomiędzy wzorem (3.11), a wzorem

(3.39) , w którym $i = 1$. Podobnie porównanie planu $P_{N,12}$ dla $N = 4k + 2$ obserwacji z planem $P_{N,02}$ dla tejże liczby obserwacji sprowadza się do zbadania relacji zachodzącej między wzorami (3.15) i (3.39), w którym $i = 2$. Wreszcie plan $P_{N,13}$ dla $N = 4k + 3$ obserwacji porównuje się z planem $P_{N,03a}$ przez zbadanie relacji między wzorami (3.19) i (3.43). Zbadanie tych relacji prowadzi łatwo do stwierdzenia, że dla każdej ustalonej liczby k ($k = 1, 2, \dots$) ze względu na minimalizację objętości elipsoidy koncentracji rozkładu estymatorów $\hat{\theta}_0, \hat{\theta}_1, \hat{\theta}_2$ plan eksperymentalny $P_{N,11}$ jest lepszy od planu $P_{N,01}$, plany $P_{N,12}$ i $P_{N,12a}$ są lepsze od planu $P_{N,02}$, plan $P_{N,13}$ jest lepszy od planu $P_{N,03a}$ oraz tym bardziej od planu $P_{N,03}$.

Porównywanie planów eksperymentalnych $P_{N,11}$,

$P_{N,12} / P_{N,12a}$, $P_{N,13}$ z planami $P_{N,01}$, $P_{N,02}$,

$P_{N,03} / P_{N,03a}$ ze względu na G - optymalność.

Dla każdej ustalonej liczby naturalnej k porównanie odpowiedniego planu $P_{N,11}$ dla $N = 4k + 1$ obserwacji z planem $P_{N,01}$ dla tej samej liczby obserwacji polega na porównaniu wzoru (3.12) ze wzorem (3.40), gdzie $i = 1$. Podobnie porównanie planu $P_{N,12}$ dla $N = 4k + 2$ obserwacji z planem $P_{N,02}$ dla tejże liczby obserwacji sprowadza się do porównania wzoru (3.16) ze wzorem (3.40), gdzie $i = 2$. Wreszcie plan $P_{N,13}$ dla $N = 4k + 3$ obserwacji porównuje się z planem $P_{N,03}$ przez porównanie wzoru (3.20) ze wzorem (3.40), w którym $i = 3$. Ła-

two dochodzi się wtedy do wniosku, że ze względu na minimalizację maksymalnej wariancji estymatora funkcji regresji plan $P_{N,01}$ jest lepszy od planu $P_{N,11}$, plan $P_{N,02}$ jest tak samo dobry jak plany $P_{N,12}$ i $P_{N,12a}$, plany $P_{N,03}$ i $P_{N,03a}$ są gorsze od planu $P_{N,13}$ przy każdej ustalonej liczbie k ($k = 1, 2, \dots$).

Porównywanie planów eksperymentalnych $P_{N,11}$, $P_{N,12}$ / $P_{N,12a}$, $P_{N,13}$ z planami eksperymentalnymi $P_{N,01}$, $P_{N,02}$, $P_{N,03}$ / $P_{N,03a}$ ze względu na I - optymalność.

Dla każdej ustalonej liczby naturalnej k porównanie odpowiedniego planu $P_{N,11}$ dla $N = 4k + 1$ obserwacji z planem $P_{N,01}$ dla tej samej liczby obserwacji polega na zbadaniu relacji pomiędzy wzorem (3.13), a wzorem (3.41), w którym $i = 1$. Podobnie porównanie planu $P_{N,12a}$ dla $N = 4k + 2$ obserwacji z planem $P_{N,02}$ dla tejże liczby obserwacji sprowadza się do zbadania relacji zachodzącej między wzorami (3.37) i (3.41), w którym $i = 2$. Wreszcie plan $P_{N,13}$ dla $N = 4k + 3$ obserwacji porównuje się z planem $P_{N,03a}$ przez zbadanie relacji między wzorami (3.21) i (3.45). Zbadanie tych relacji prowadzi łatwo do stwierdzenia, że dla każdej ustalonej liczby k ($k = 1, 2, \dots$) ze względu na minimalizację średniej wartości wariancji estymatora funkcji regresji plan eksperymentalny $P_{N,11}$ jest lepszy od planu $P_{N,01}$, plan $P_{N,12}$, a nawet plan $P_{N,12a}$ jest lepszy od planu $P_{N,02}$, plan $P_{N,13}$ jest lepszy od planu $P_{N,03a}$ oraz tym bardziej od planu $P_{N,03}$.

Porównywanie planów eksperymentalnych $P_{N,10} - P_{N,13}$
ze sposobem losowym rozmieszczenia obserwacji w kwadracie X .

W przedostatniej kolumnie tabeli 3.1 zamieszczono, podzielone przez σ^2 , maksymalne wartości wariancji estymatora funkcji regresji (3.1) obliczone na kwadracie $X = \{x = (x_1, x_2), -1 \leq x_i \leq 1, i = 1, 2\}$, gdy obserwacje wykonuje się według jednego z planów eksperymentalnych $P_{N,10}, P_{N,11}, P_{N,12}, P_{N,13}$ w zależności od tego, czy liczba obserwacji N podzielona przez 4 daje odpowiednio resztę 0, 1, 2 lub 3. W ostatniej kolumnie tej tabeli podano maksymalne wartości wariancji estymatora funkcji regresji (3.1) podzielone przez σ^2 , gdy N obserwacji wykonuje się w punktach $(x_1^{(1)}, x_2^{(1)}), (x_1^{(2)}, x_2^{(2)}), \dots, (x_1^{(N)}, x_2^{(N)})$ otrzymanych z generatora liczb pseudolosowych o rozkładzie określonym gęstością

$$g_2(x_1, x_2) = \begin{cases} \frac{1}{4} & \text{dla } -1 \leq x_i \leq 1, \quad i = 1, 2 \\ 0 & \text{poza tym} \end{cases}$$

tzn. o rozkładzie jednostajnym na kwadracie X . Dla ustalonej liczby N wartości maksymalne $\max_{x \in X} \text{Var}(\hat{E}Y_x) / \sigma^2$ podane w ostatniej kolumnie są średnimi otrzymanymi w serii 30 eksperymentów komputerowych (por. rozdział VI).

Jeżeli zastosujemy kryterium G - optymalności jako kryterium dobroci eksperymentu mającego na celu estymację funkcji regresji $EY_x = \theta_0 + \theta_1 x_1 + \theta_2 x_2$ na kwadracie $X = \{(x_1, x_2) : -1 \leq x_i \leq 1, i = 1, 2\}$, to wyniki zamieszczone w tabeli 3.1 pozwalają - analogicznie jak

T A B E L A 3.1

Maksymalne wartości wariancji estymatora funkcji regresji

$$EY_x = \theta_0 + \theta_1 x_1 + \theta_2 x_2 \quad \text{na kwadracie } X = \{-1 \leq x_1 \leq 1, -1 \leq x_2 \leq 1\}$$

Ogólna liczba obser- wacji N	Sposób rozmieszczenia obserwacji					
	według planów $P_{N,10} - P_{N,13}$				losowy	
	Liczba obserwacji w punkcie (1,1) (1,-1) (-1,-1) (-1,1)				$\max_{x \in X} \text{Var}(\hat{EY}_x) / \sigma^2$	$\max_{x \in X} \text{Var}(\hat{EY}_x) / \sigma^2$
3	1	1	1	-	3	1256.494
4	1	1	1	1	.750	20.396
5	2	1	1	1	.714	6.987
6	2	1	2	1	.667	4.920
7	2	2	2	1	.600	3.106
8	2	2	2	2	.375	2.312
9	3	2	2	2	.364	2.321
10	3	2	3	2	.350	1.475
11	3	3	3	2	.333	1.215
12	3	3	3	3	.250	1.222
13	4	3	3	3	.244	1.002
14	4	3	4	3	.238	.855
15	4	4	4	3	.231	.705
16	4	4	4	4	.188	.758
17	5	4	4	4	.184	.688
18	5	4	5	4	.181	.645
19	5	5	5	4	.176	.565
20	5	5	5	5	.150	.515
21	6	5	5	5	.148	.522
22	6	5	6	5	.145	.458
23	6	6	6	5	.143	.479
24	6	6	6	6	.125	.429
25	7	6	6	6	.123	.385
30	8	7	8	7	.105	.326
35	9	9	9	8	.091	.265
40	10	10	10	10	.075	.223
45	12	11	11	11	.068	.193
50	13	12	13	12	.062	.182
60	15	15	15	15	.050	.147
70	18	17	18	17	.044	.124
80	20	20	20	20	.038	.104
90	23	22	23	22	.034	.094
100	25	25	25	25	.030	.082
200	50	50	50	50	.015	.039
500	125	125	125	125	.006	.015

wyniki zamieszczone w tabeli 2.1, rozdz. II - ocenić korzyści stosowania G - optymalnego planu eksperymentalnego $P_{N,10}$ oraz planów $P_{N,11}$, $P_{N,12}$ i $P_{N,13}$ w porównaniu ze sposobem losowym wykonywania obserwacji w punktach otrzymanych z rozkładu $g_2(x_1, x_2)$.

Rozdział IV

REGRESJA LINIOWA NA KOSTCE m - WYMIAROWEJ

Przez kostkę m -wymiarową będziemy w dalszym ciągu rozumieć zbiór $X = \{x = (x_1, x_2, \dots, x_m), -1 \leq x_i \leq 1, i = 1, 2, \dots, m\}$. Na tak określonym zbiorze X będziemy zajmowali się konstrukcją optymalnych planów eksperymentalnych dla estymacji liniowej funkcji regresji

$$EY_x = \theta_0 + \theta_1 x_1 + \dots + \theta_m x_m .$$

Zacytujemy najpierw

TWIERDZENIE 7 . Niech $X = X_1 = X_2 = \{x = (x_1, \dots, x_m), x_i: -1 \leq x_i \leq 1, i = 1, 2, \dots, m\}$, $k = m + 1$, $f_1(x) = 1$, $f_2(x) = x_1, \dots, f_k(x) = x_m$. Rozważmy plan eksperymentalny $P_{N,4} \in \mathfrak{D}_N$, gdzie $N = n \cdot 2^m$, określony następująco

$$\left\{ \begin{array}{lll} (1, 1, \dots, 1) & (-1, 1, \dots, 1) & (1, -1, \dots, 1) \dots \\ 2^{-m} & 2^{-m} & 2^{-m} \\ \\ (1, \dots, 1, -1) & (-1, -1, \dots, 1) & (-1, 1, -1, \dots, 1) \dots \\ 2^{-m} & 2^{-m} & 2^{-m} \\ \\ (1, \dots, 1, -1, -1) \dots & (-1, -1, \dots, -1) & \\ 2^{-m} & 2^{-m} & \end{array} \right\}$$

tzn. skoncentrowany z jednakowymi miarami $2^{-m} = \frac{n}{N}$ dokładnie we wszystkich wierzchołkach zbioru X . Plan $P_{N,4}$ jest

planem A -, D - oraz G - optymalnym dla każdej liczby $N = n \cdot 2^m$ obserwacji.

Dowód: Dla planu $P_{N,4}$ macierz $M(P_{N,4})$ jest macierzą jednostkową stopnia $m + 1$. Wobec tego $M(P_{N,4}) = M^{-1}(P_{N,4}) = I$, skąd

$$d(x, P_{N,4}) = \begin{bmatrix} 1 & x_1 & \dots & x_m \end{bmatrix} M^{-1}(P_{N,4}) \begin{bmatrix} 1 \\ x_1 \\ \vdots \\ x_m \end{bmatrix} = 1 + \sum_{i=1}^m x_i^2 .$$

Zatem

$$\max_{x \in X} d(x, P_{N,4}) = \max_{x \in X} (1 + \sum_{i=1}^m x_i^2) = m + 1 = k ,$$

co zgodnie z twierdzeniem 2 oznacza, że plan $P_{N,4}$ jest D - i G - optymalny. Ponadto, ponieważ $M^{-1}(P_{N,4}) = I$, spełnione jest równanie $M^{-1}(P_{N,4}) = [M^{-1}(P_{N,4})]^2$. Stąd i z twierdzenia 3 wynika, że plan $P_{N,4}$ jest również planem A - optymalnym.

c.b.d.o.

Podamy teraz kilka uwag dotyczących interpretacji planu eksperymentalnego $P_{N,4}$ z twierdzenia 7 w terminach doświadczenia czynnikowego. Wyobraźmy sobie, że należy wykonać pełne doświadczenie czynnikowe typu 2^m i że dla każdego z m czynników tzw. poziom dolny = - 1, zaś poziom górny = + 1. Oznacza to, że należy wykonać po jednej obserwacji dla każdej z 2^m możliwych kombinacji poziomów czynników uwzględnianych w doświadczeniu, które to kombinacje wyrażają się przez ciągi m -elementowe o elementach równych + 1 lub - 1. Oczywiście zbiór tych ciągów można traktować jako zbiór współrzędnych wierzchołków kostki m -wymia-

rowej. Możemy powiedzieć, że zbiór 2^m -elementowy wierzchołków kostki wyznacza pełne doświadczenie czynnikowe typu 2^m . Tak więc pełne doświadczenie czynnikowe pokrywa się z planem eksperymentalnym $P_{N,4}$ z twierdzenia 7 dla $N = 2^m$ obserwacji.

Niech $F_{m,0}$ oznacza macierz o $N = 2^m$ wierszach oraz $m + 1$ kolumnach taką, że wszystkie elementy pierwszej kolumny są równe $+1$, zaś pozostałe elementy kolejnego wiersza - oprócz określonych już elementów pierwszej kolumny - dane są przez współrzędne kolejnego wierzchołka kostki m -wymiarowej. Jeśli przez x_{ui} oznaczymy element znajdujący się w wierszu u ($u = 1, \dots, N$) oraz kolumnie $i + 1$ ($i=1, \dots, m$) macierzy $F_{m,0}$, to x_{ui} jest i -tą współrzędną wierzchołka kostki, któremu przyporządkowano numer u . Macierz $F_{m,0}$ nazywamy macierzą planowania pełnego doświadczenia czynnikowego. Łatwo widać, że macierz $M(P_{N,4}) = 2^{-m} F_{m,0}^T F_{m,0} = I$. W tabelach 4.1 i 4.2 przedstawiamy przykładowo macierze planowania dla doświadczenia czynnikowego 2^2 oraz 2^3 , przy czym w macierzach tych zastąpiono liczby $+1$ i -1 odpowiednio znakami $+1$ i -1 .

Macierz $F_{m,0}$ generuje macierz $V_{m,0}$ o $N = 2^m$ wierszach oraz $2^m - (m+1) = \binom{m}{2} + \binom{m}{3} + \dots + \binom{m}{m}$ kolumnach w sposób następujący: Elementami każdego ustalonego wiersza u ($u = 1, \dots, N$) macierzy $F_{m,0}$ są $1, x_{u1}, x_{u2}, \dots, x_{um}$. Otrzymuje się stąd najpierw $\binom{m}{2}$ kolumn macierzy $V_{m,0}$ tak, że element wiersza u ($u = 1, \dots, N$) wybranej kolumny jest równy iloczynowi $x_{ui}x_{uj}$ ($i < j$, $i = 1, \dots, m-1$; $j = 2, \dots, m$), gdzie para i, j jest ustalona przez wybór kolumny w sposób wzajemnie jednoznaczny. Następnie otrzymu-

Tabela 4.1

Macierz planowania pełnego
doświadczenia czynnikowego
typu 2^2

1	x_1	x_2
+	-	-
+	+	-
+	-	+
+	+	+

Tabela 4.2

Macierz planowania pełnego
doświadczenia czynnikowego
typu 2^3

1	x_1	x_2	x_3
+	-	-	-
+	+	-	-
+	-	+	-
+	+	+	-
+	-	-	+
+	+	-	+
+	-	+	+
+	+	+	+

je się $\binom{m}{3}$ dalszych kolumn macierzy $V_{m,0}$, gdzie element wiersza u ($u = 1, \dots, N$) ustalonej (przez wybór i, j, k) kolumny spośród tych $\binom{m}{3}$ kolumn jest równy iloczynowi $x_{ui}x_{uj}x_{uk}$ ($i < j < k$, $i = 1, \dots, m-2$; $j = 2, \dots, m-1$; $k = 3, \dots, m$). Dalej otrzymuje się analogicznie $\binom{m}{4}$, $\binom{m}{5}$ itd. kolejnych kolumn macierzy $V_{m,0}$. Elementy ostatniej kolumny macierzy $V_{m,0}$ są równe iloczynowi $x_{u1}x_{u2} \dots x_{um}$. Niech F_m oznacza macierz kwadratową o $N = 2^m$ wierszach i kolumnach powstałą przez dołączenie macierzy $V_{m,0}$ do macierzy $F_{m,0}$ (Tabela 4.3 przedstawia macierz F_3). Zachodzi następujący

LEMAT 1. Kolumny macierzy F_m są ortogonalne.

Dowód: Wystarczy pokazać, że

T A B E L A 4.3

Macierz F_3

1	x_1	x_2	x_3	x_1x_2	x_1x_3	x_2x_3	$x_1x_2x_3$
+	-	-	-	+	+	+	-
+	+	-	-	-	-	+	+
+	-	+	-	-	+	-	+
+	+	+	-	+	-	-	-
+	-	-	+	+	-	-	+
+	+	-	+	-	+	-	-
+	-	+	+	-	-	+	-
+	+	+	+	+	+	+	+

macierz $F_{3,0}$
macierz $V_{3,0}$

$$(4.1) \quad \sum_{u=1}^N x_{ui_1} x_{ui_2} \dots x_{ui_k} = 0 \quad \text{dla } 1 \leq i_1 < i_2 < \dots < i_k \leq m.$$

Wynika to z tego, że iloczyn skalarny dowolnych dwu kolumn macierzy F_m jest postaci

$$\sum_{u=1}^N x_{ui_1}^{p_1} x_{ui_2}^{p_2} \dots x_{ui_k}^{p_k}, \quad \text{gdzie } p_j \ (j = 1, \dots, k)$$

jest równe 1 lub 2. W przypadku $p_j = 2$ mamy $x_{ui_j}^2 = 1$, skąd redukcja do (4.1). Ponumerujmy tak wierzchołki kostki m -wymiarowej, żeby dla $u = 1, \dots, N/2$ było $x_{ui_k} = 1$, zaś dla $u = N/2+1, \dots, N$ było $x_{ui_k} = -1$. Stąd zamiast równania (4.1) otrzymujemy równoważne równanie

$$(4.2) \quad \sum_{u=1}^{N/2} x_{ui_1} x_{ui_2} \dots x_{ui_{k-1}} = \sum_{u=N/2+1}^N x_{ui_1} x_{ui_2} \dots x_{ui_{k-1}}$$

Wartość dowolnego składnika $x_{ui_1} x_{ui_2} \dots x_{ui_{k-1}}$ każdej z dwu sum w równaniu (4.2) jest wyznaczona przez współrzędne $(x_{ui_1}, \dots, x_{ui_{k-1}})$. Ponieważ ustalone współrzędne $(x_{ui_1}^0, \dots, x_{ui_{k-1}}^0)$ posiada 2^{m-k} składników o indeksie $u \leq N/2$ oraz 2^{m-k} składników o indeksie $u \geq N/2+1$, więc równanie (4.2), a tym samym równanie (4.1), jest spełnione.

cbdo

Z dowodu tego lematu wynikają natychmiast poniższe wnioski.

WNIOSEK 1. Zachodzi równanie $2^{-m} F_m^T F_m = I$.

WNIOSEK 2. Macierz S_m powstała przez opuszczenie dowolnej liczby kolumn macierzy F_m zachowuje ortogonalność kolumn i dla macierzy tej $2^{-m} S^T S_m = I$.

WNIOSEK 3. Macierz, powstała przez dołączenie do macierzy F_m dowolnej liczby wierszy o elementach 1, 0, 0, ..., 0, zachowuje ortogonalność kolumn.

Oprócz pełnego doświadczenia czynnikowego można też rozważać tzw. ułamkowe doświadczenie czynnikowe typu 2^{m-p} . Jest to doświadczenie, w którym uwzględnia się m czynników na dwu poziomach każdy, ale ogólna liczba obserwacji $N = 2^{m-p}$. Sformułujemy algorytm na wyznaczanie macierzy planowania $F_{m,p}$ ułamkowego doświadczenia czynnikowego typu 2^{m-p} z macierzy F_{m-p} .

ALGORYTM. Macierz planowania $F_{m,p}$ ułamkowego doświadczenia czynnikowego typu 2^{m-p} , gdzie $2^{m-p} > m$, można otrzymać z macierzy kwadratowej F_{m-p} o $N = 2^{m-p}$ wierszach i kolumnach w sposób następujący: Pierwszą kolumną macierzy

$F_{m,p}$ jest pierwsza kolumna macierzy F_{m-p} . Każdą z pozostałych m kolumn macierzy $F_{m,p}$ otrzymuje się tak, że ustalona kolumna macierzy $F_{m,p}$ jest równa z dokładnością do znaku jednej z tych kolumn macierzy F_{m-p} , które nie zostały jeszcze przyporządkowane innym kolumnom macierzy $F_{m,p}$. Dla utworzenia macierzy $F_{m,p}$ nie wykorzystuje się zatem $2^{m-p} - (m+1)$ kolumn macierzy F_{m-p} . Elementy wiersza u ($u=1, \dots, 2^{m-p}$) macierzy $F_{m,p}$ są postaci $1, x_{u1}, x_{u2}, \dots, x_{um}$, gdzie $x_{ui} = \pm 1$ ($i = 1, \dots, m$).

Ciąg $(x_{u1}, x_{u2}, \dots, x_{um})$ można interpretować jako współrzędne pewnego wierzchołka kostki m -wymiarowej.

Przykłady.

1. Potraktujmy macierz F_3 przedstawioną w tabeli 4.3 jako macierz F_{4-1} . Zgodnie z podanym algorytmem możemy z tej macierzy otrzymać macierz $F_{4,1}$ ułamkowego doświadczenia czynnikowego typu 2^{4-1} ($m = 4$ czynniki, $N = 2^{4-1} = 8$ obserwacji) na $2^4 \binom{7}{4} = 16 \cdot 35 = 560$ różnych sposobów. Tabele 4.4 i 4.5 przedstawiają dwie spośród tak otrzymanych macierzy $F_{4,1}$.
2. Potraktujmy macierz F_3 z tabeli 4.3 jako macierz F_{5-2} . Macierz $F_{5,2}$ doświadczenia typu 2^{5-2} ($m = 5$ czynników, $N = 2^{5-2} = 8$ obserwacji) można otrzymać z macierzy F_3 - według algorytmu - na jeden z $2^5 \binom{7}{5} = 672$ możliwych sposobów. Tabele 4.6 i 4.7 przedstawiają dwie spośród tak otrzymanych macierzy $F_{5,2}$.
3. Potraktujmy macierz F_3 z tabeli 4.3 jako macierz F_{0-5} . Nie jest wtedy spełniony wymieniony w algorytmie

Tabela 4.4

Macierz $F_{4,1}$ otrzymana z macierzy F_{4-1} przez przyporządkowanie $x_1=x_1, x_2=x_2, x_3=x_3, x_4 = -x_1x_2$

1	x_1	x_2	x_3	x_4
+	-	-	-	-
+	+	-	-	+
+	-	+	-	+
+	+	+	-	-
+	-	-	+	-
+	+	-	+	+
+	-	+	+	+
+	+	+	+	-

Tabela 4.5

Macierz $F_{4,1}$ otrzymana z macierzy $F_{4-1} = F_3$ przez przyporządkowanie $x_1=x_1, x_2=x_2, x_3=x_3, x_4 = x_1x_2x_3$

1	x_1	x_2	x_3	x_4
+	-	-	-	-
+	+	-	-	+
+	-	+	-	+
+	+	+	-	-
+	-	-	+	+
+	+	-	+	-
+	-	+	+	-
+	+	+	+	+

Tabela 4.6

Macierz $F_{5,2}$ otrzymana z macierzy $F_{5-2} = F_3$ przez przyporządkowanie $x_1=x_1, x_2=x_2, x_3=x_3, x_4 = x_1x_2, x_5 = x_1x_2x_3$

1	x_1	x_2	x_3	x_4	x_5
+	-	-	-	-	-
+	+	-	-	+	+
+	-	+	-	+	+
+	+	+	-	-	-
+	-	-	+	-	+
+	+	-	+	+	-
+	-	+	+	+	-
+	+	+	+	-	+

Tabela 4.7

Macierz $F_{5,2}$ otrzymana z macierzy $F_{5-2} = F_3$ przez przyporządkowanie $x_1=x_1, x_2=x_2, x_3=x_3, x_4 = x_2x_3, x_5 = x_1x_3$

1	x_1	x_2	x_3	x_4	x_5
+	-	-	-	+	+
+	+	-	-	+	-
+	-	+	-	-	+
+	+	+	-	-	-
+	-	-	+	-	-
+	+	-	+	-	+
+	-	+	+	+	-
+	+	+	+	+	+

warunek $2^{m-p} > m$. Tym samym nie działa algorytmiczna reguła konstrukcji macierzy $F_{8,5}$ na podstawie macierzy F_3 .

LEMAT 2. Określona wyżej macierz $F_{m,p}$ ułamkowego doświadczenia czynnikowego 2^{m-p} ($2^{m-p} > m$) ma kolumny ortogonalne.

Dowód: W klasie macierzy S_{m-p} powstałych przez opuszczenie dowolnej liczby kolumn macierzy F_{m-p} istnieje macierz S_{m-p}^* o liczbie kolumn równej liczbie kolumn macierzy $F_{m,p}$ i taka, że dla każdej kolumny macierzy S_{m-p}^* istnieje równa jej z dokładnością do znaku kolumna macierzy $F_{m,p}$. Zmiana znaku kolumny może zmienić wartość iloczynu skalarnego dwu kolumn na wartość przeciwną, natomiast nie zmienia modułu tej wartości. Zgodnie z wnioskiem 2 macierzy S_{m-p}^* ma kolumny ortogonalne. Tę samą własność ma zatem również macierz $F_{m,p}$.

c.b.d.o.

Natychmiast mamy

WNIOSEK 4. Zachodzi równanie $2^{p-m} P_{m,p}^T P_{m,p} = I$.

Wprowadzimy teraz wygodną dla naszych celów definicję:

DEFINICJA. Nieregularnym doświadczeniem czynnikowym typu $2^{m-p_1, m-p_2, \dots, m-p_k}$, gdzie $2^{m-p_i} > m$ ($i = 1, \dots, k$) nazywamy dowolne doświadczenie o macierzy planowania

$$F_{m, (p_1, p_2, \dots, p_k)} = \begin{bmatrix} F_{m, p_1} \\ F_{m, p_2} \\ \vdots \\ F_{m, p_k} \end{bmatrix}$$

Rozwijając tę definicję można powiedzieć, że nieregularne doświadczenie czynnikowe dla m czynników na dwu poziomach każdy jest sumą k ułamkowych ($p_i > 0$) lub

pełnych ($p_1 = 0$) doświadczeń czynnikowych odpowiednio typu $2^{m-p_1}, 2^{m-p_2}, \dots, 2^{m-p_k}$ o łącznej liczbie obserwacji $N = 2^{m-p_1} + 2^{m-p_2} + \dots + 2^{m-p_k}$.

Przykłady:

4. Określony w twierdzeniu 7 plan eksperymentalny $P_{N,4}$ dla $N = n \cdot 2^m$ obserwacji jest dany macierzą planowania

$$F_{m,(\underbrace{0,0,\dots,0}_n \text{ zer})} = \left[\begin{array}{c} F_{m,0} \\ F_{m,0} \\ \vdots \\ F_{m,0} \end{array} \right] \left. \vphantom{\begin{array}{c} F_{m,0} \\ F_{m,0} \\ \vdots \\ F_{m,0} \end{array}} \right\} n \text{ macierzy } F_{m,0}$$

5. Nieregularne doświadczenie czynnikowe typu $2^{5-2,5-2}$ ma macierz planowania $F_{5,(2,2)} = \left[\begin{array}{c} F_{5,2} \\ F_{5,2} \end{array} \right]$. Wstawmy za jedną z macierzy $F_{5,2}$ macierz zamieszczoną w tabeli 4.6 a za drugą z nich - macierz zamieszczoną w tabeli 4.7. Wtedy $N = 2^{5-2} + 2^{5-2} = 16$ obserwacji wykonuje się w 14 wierzchołkach kostki pięciowymiarowej, przy tym w wierzchołkach o współrzędnych $(1, 1, -1, -1, -1)$ i $(-1, 1, 1, 1, -1)$ wykonuje się po dwie obserwacje.

LEMAT 3. Kolumny macierzy $F_{m,(p_1,p_2,\dots,p_k)}$ są ortogonalne.

Dowód: Z definicji macierzy $F_{m,(p_1,p_2,\dots,p_k)}$ widać, że iloczyn skalarny dwu kolumn o numerach i oraz j ($i, j = 1, \dots, m+1; i \neq j$) tejże macierzy jest równy sumie k iloczynów skalarnych kolumn o tych samych numerach macierzy $F_{m,p_1}, F_{m,p_2}, \dots, F_{m,p_k}$. Zgodnie z lema-

tem 2 każda z macierzy F_{m,p_i} ($i = 1, \dots, k$) ma kolumny ortogonalne. Wobec tego również macierz $F_{m,(p_1,p_2,\dots,p_k)}$ ma kolumny ortogonalne.

c.b.d.o.

Oczywisty jest

WNIOSEK 5. Zachodzi równanie

$$(2^{m-p_1} + 2^{m-p_2} + \dots + 2^{m-p_k})^{-1} F_{m,(p_1,\dots,p_k)}^T F_{m,(p_1,\dots,p_k)} = I$$

Macierzy planowania $F_{m,(p_1,p_2,\dots,p_k)}$ nieregularnego doświadczenia czynnikowego przyporządkujemy następujący plan eksperymentalny $P_{N,5}$ dla $N = n(2^{m-p_1} + 2^{m-p_2} + \dots + 2^{m-p_k})$ obserwacji wykonywanych w wierzchołkach kostki m -wymiarowej:

$$(4.3) \quad \left\{ \begin{array}{cccc} x^{(1)} & x^{(2)} & \dots & x^{(2^m)} \\ \lambda_1 & \lambda_2 & \dots & \lambda_{2^m} \end{array} \right\},$$

gdzie $x^{(u)}$ oznacza u -ty ($u = 1, \dots, 2^m$) wierzchołek dany przez współrzędne $x_{u1}, x_{u2}, \dots, x_{um}$, przy czym $x_{ui} = \pm 1$ ($i = 1, \dots, m$), natomiast

$$\lambda_u = \lambda(x_{u1}, x_{u2}, \dots, x_{um}) = \frac{n \cdot n_u}{N} = \frac{n_u}{2^{m-p_1} + 2^{m-p_2} + \dots + 2^{m-p_k}}$$

$u = 1, \dots, 2^m$, zaś n_u ($n_u \geq 0$) oznacza liczbę wierszy macierzy $F_{m,(p_1,p_2,\dots,p_k)}$ o elementach $1, x_{u1}, x_{u2}, \dots, x_{um}$.

Przykład.

6. Rozważanej w przykładzie 5 macierzy planowania $F_{5,(2,2)}$

odpowiada plan eksperymentalny $P_{N,5}$ dla każdej wielokrotności 16 obserwacji na wierzchołkach kostki pięciowymiarowej, dla którego (por. tabele 4.6 i 4.7):

$$\begin{aligned} \lambda(-1,-1,-1,-1,-1) &= \lambda(1,-1,-1,1,1) = \lambda(-1,1,-1,1,1) = \\ \lambda(-1,-1,1,-1,1) &= \lambda(1,-1,1,1,-1) = \lambda(1,1,1,-1,1) = \\ \lambda(-1,-1,-1,1,1) &= \lambda(1,-1,-1,1,-1) = \lambda(-1,1,-1,-1,1) = \\ \lambda(-1,-1,1,-1,-1) &= \lambda(1,-1,1,-1,1) = \lambda(1,1,1,1,1) = \\ &= 1/16 = 0,0625 ; \end{aligned}$$

$$\lambda(1,1,-1,-1,-1) = \lambda(-1,1,1,1,-1) = 1/8 = 0,125 ;$$

Potrzebne dalej własności planu $P_{N,5}$ przedstawia następujące

TWIERDZENIE 8. Niech $X = X_1 = X_2 = \{x = (x_1, \dots, x_m), x_1: -1 \leq x_1 \leq 1, i = 1, \dots, m\}$, $k = m + 1$, $f_1(x) = 1$, $f_2(x) = x_1, \dots, f_k(x) = x_m$. Rozważmy plan eksperymentalny $P_{N,5} \in \mathfrak{D}_N$ dla $N = n(2^{m-p_1} + 2^{m-p_2} + \dots + 2^{m-p_k})$ obserwacji określony wzorem (4.3). Plan $P_{N,5}$ jest planem A -, D - i G - optymalnym. Przy tym macierz

$$(4.4) \quad M(P_{N,5}) = (2^{m-p_1} + \dots + 2^{m-p_k})^{-1} F_{m, (p_1, \dots, p_k)}^T F_{m, (p_1, \dots, p_k)}$$

Dowód: Równość (4.4) wynika bezpośrednio z określenia planu $P_{N,5}$. Dalej, z wniosku 5 mamy, że prawa strona tej równości jest macierzą jednostkową, więc oczywiście $M(P_{N,5}) = I$. Stąd, tak jak w dowodzie twierdzenia 7, wynika A -, D - i G - optymalność planu $P_{N,5}$.

c.b.d.o.

DEFINICJA. Macierz kwadratową o n wierszach i kolumnach nazywamy macierzą Hadamarda H_n , jeśli wszystkie jej elementy są równe $+1$ lub -1 i jeśli $n^{-1}H_n^T H_n = I$.

Przykłady.

7. Łatwo skonstruować macierze Hadamarda dla $n = 1$ i 2 :

$$\begin{bmatrix} 1 \end{bmatrix}, \quad \begin{bmatrix} 1 & 1 \\ 1 & -1 \end{bmatrix}$$

8. Skonstruowane wcześniej macierze kwadratowe F_m o 2^m wierszach i kolumnach są macierzami Hadamarda.

Dowolną macierz Hadamarda można przekształcić przy pomocy mnożenia odpowiednich wierszy przez -1 tak, że pierwsza kolumna składa się tylko z elementów $+1$ (takie przekształcenie daje znów macierz Hadamarda). Odtąd będziemy zakładali, że macierze Hadamarda są doprowadzone do tej postaci. Dalej można też udowodnić, że stopień n macierzy Hadamarda wynosi $1, 2$ lub $4t$ ($t = 1, 2, \dots$). Pokazano (por. np. [36], str. 71 - 82) istnienie macierzy Hadamarda stopnia $4t$ dla kolejnych wartości t od 1 do 46 włącznie, tzn. od $n = 4 \cdot 1 = 4$ do $n = 4 \cdot 46 = 184$, oraz dla wielu większych wartości t począwszy od $t = 48$, tzn. $n = 192$.

WNIOSEK 6. Niech $H_{n,m}$ oznacza macierz o n wierszach i $m+1 \leq n$ kolumnach otrzymaną przez skreślenie dowolnych $n - m - 1$ kolumn macierzy H_n różnych od kolumny pierwszej. Zachodzi równanie

$$n^{-1}H_{n,m}^T H_{n,m} = I.$$

Dowód wniosku wynika natychmiast z definicji macierzy Hadamarda.

Uwaga. W przypadku gdy $n = m + 1$, mamy $H_{n,m} = H_n$.

Założmy, że liczby naturalne m, n_1, n_2, \dots, n_k są takie, że macierze $H_{n_1,m}, H_{n_2,m}, \dots, H_{n_k,m}$ istnieją. Rozważmy teraz doświadczenie dla m czynników o łącznej liczbie obserwacji $N = n_1 + n_2 + \dots + n_k$, które jest sumą k doświadczeń czynnikowych odpowiednio o n_1, n_2, \dots, n_k obserwacjach, przy czym plany doświadczeń składowych dane są odpowiednio macierzami $H_{n_1,m}, H_{n_2,m}, \dots, H_{n_k,m}$. Łączne doświadczenie ma zatem macierz planowania

$$H_{(n_1, n_2, \dots, n_k), m} = \begin{bmatrix} H_{n_1, m} \\ H_{n_2, m} \\ \vdots \\ H_{n_k, m} \end{bmatrix}$$

LEMAT 4. Kolumny macierzy $H_{(n_1, n_2, \dots, n_k), m}$ są ortogonalne.

Dowód: Z definicji macierzy $H_{(n_1, n_2, \dots, n_k), m}$ widać, że iloczyn skalarny kolumn o numerach i oraz j ($i, j = 1, \dots, m+1, i \neq j$) jest równy sumie k iloczynów skalarnych kolumn i oraz j macierzy $H_{n_1, m}, H_{n_2, m}, \dots, H_{n_k, m}$. Z wniosku 6 wynika ortogonalność kolumn każdej z macierzy $H_{n_i, m}$ ($i = 1, \dots, k$). Wobec tego macierz $H_{(n_1, n_2, \dots, n_k), m}$ ma również kolumny ortogonalne.

c.b.d.o.

Oczywisty jest też

WNIOSEK 7. Zachodzi równanie

$$(n_1+n_2+\dots+n_k)^{-1} H_{(n_1, n_2, \dots, n_k), m}^T H_{(n_1, n_2, \dots, n_k), m} = I.$$

Macierzy planowania $H_{(n_1, n_2, \dots, n_k), m}$ przyporządkujemy następujący plan eksperymentalny $P_{N,6}$ dla $N = n(n_1 + n_2 + \dots + n_k)$ obserwacji wykonywanych w wierzchołkach kostki m -wymiarowej:

$$(4.5) \quad \left\{ \begin{array}{cccc} x^{(1)} & x^{(2)} & \dots & x^{(2^m)} \\ \lambda'_1 & \lambda'_2 & \dots & \lambda'_{2^m} \end{array} \right\},$$

gdzie $x^{(u)}$ oznacza u -ty ($u = 1, 2, \dots, 2^m$) wierzchołek dany przez współrzędne $x_{u1}, x_{u2}, \dots, x_{um}$, przy czym $x_{ui} = \pm 1$ ($i = 1, \dots, m$), natomiast $\lambda'_u =$

$$= \lambda'(x_{u1}, x_{u2}, \dots, x_{um}) = \frac{n \cdot n'_u}{N} = \frac{n'_u}{n_1+n_2+\dots+n_k},$$

$u = 1, \dots, 2^m$, zaś n'_u ($n'_u \geq 0$) oznacza liczbę wierszy macierzy $H_{(n_1, \dots, n_k), m}$ o elementach $1, x_{u1}, x_{u2}, \dots, x_{um}$.

TWIERDZENIE 9. Załóżmy, że liczby naturalne m, n_1, n_2, \dots, n_k są takie, że macierz $H_{(n_1, n_2, \dots, n_k), m}$ istnieje i niech $X = X_1 = X_2 = \{x = (x_1, \dots, x_m), x_i: -1 \leq x_i \leq 1, i = 1, \dots, m\}$, $k = m+1$, $f_1(x) = 1, f_2(x) = x_1, \dots, f_k(x) = x_m$. Rozważmy plan eksperymentalny $P_{N,6} \in \mathfrak{D}_N$ dla $N = n \cdot (n_1 + n_2 + \dots + n_k)$ obserwacji dany wzorem (4.5). Plan $P_{N,6}$ jest planem A-, D- i G- optymalnym.

Przy tym macierz

$$(4.6) \quad M(P_{N,6}) = \frac{1}{n_1+n_2+\dots+n_k} H^T(n_1, \dots, n_k), m \quad H(n_1, \dots, n_k), m$$

Dowód: Równość (4.6) wynika natychmiast z określenia planu $P_{N,6}$. Ponadto z wniosku 7 mamy, że prawa strona wzoru (4.6) jest macierzą jednostkową, skąd oczywiście $M(P_{N,6}) = I$. Z tego, tak jak w dowodzie twierdzenia 7, wynika A -, D - i G - optymalność planu $P_{N,6}$.

c.b.d.o.

TWIERDZENIE 10. Jeśli liczba czynników m spełnia nierówność $2 \leq m \leq 91$, to A -, D - i G - optymalny plan eksperymentalny $P_{N,6}$, tzn. plan o macierzy planowania postaci $H(n_1, \dots, n_k), m$, można skonstruować dla każdej liczby obserwacji $N = N_m + 4t$, gdzie $N_m = 4 + m - m(\text{mod } 4)$, zaś $t = 0, 1, 2, \dots$.

Dowód: Przy założeniu, że $2 \leq m \leq 91$, z istnienia macierzy Hadamarda stopnia $4t$ dla kolejnych wartości t od 1 do 46, tzn. macierzy stopnia 4, 8, 12, ..., 184 wynika istnienie planów eksperymentalnych $P_{N,6}$ określonych z pomocą macierzy planowania

$$H(n_1), m = \left[H_{n_1, m} \right]$$

dla każdej liczby obserwacji $n_1 = N_m, N_m+4, N_m+8, \dots, 180, 184$. Udowodnimy teraz istnienie A -, D - i G - optymalnych planów $P_{N,6}$ dla każdej liczby obserwacji $N = N_m + 4t \geq 188$. Ponieważ $\max_{2 \leq m \leq 91} N_m = 92$, więc ciąg

$N_m, N_{m+4}, \dots, 184$ zawiera również liczby 92 i 96.

Plan $P_{N,6}$ określony z pomocą macierzy planowania

$$H_{(92,96),m} = \begin{bmatrix} H_{92,m} \\ H_{96,m} \end{bmatrix}, \quad 2 \leq m \leq 91,$$

jest zatem, zgodnie z twierdzeniem 9, A -, D - i G - optymalnym planem dla $N = 188$ obserwacji. Uczyńmy teraz indukcyjne założenie, że taki plan istnieje dla każdej liczby $N \leq N_0$, gdzie $N(\text{mod } 4) = N_0(\text{mod } 4) = 0$ i $N_0 \geq 188$, tzn. plan A -, D - i G - optymalny istnieje dla każdej liczby obserwacji z ciągu $N_m, N_{m+4}, \dots, 184, \dots, N_0$. Plan $P_{N,6}$ dla $N_0 \geq 188$ obserwacji może być określony z pomocą macierzy planowania

$$H_{(92, N_0 - 92), m} = \begin{bmatrix} H_{92, m} \\ H_{N_0 - 92, m} \end{bmatrix}, \quad 2 \leq m \leq 91.$$

W tym celu wystarczy zauważyć, że $N_m \leq 92$, $N_0 - 92 < N_0$ oraz, że $(N_0 - 92)(\text{mod } 4) = 0$, tzn. liczby 92 oraz $N_0 - 92$ są elementami ciągu N_m, N_{m+4}, \dots, N_0 i wobec tego macierz planowania

$$H_{(92, N_0 - 92 + 4), m} = \begin{bmatrix} H_{92, m} \\ H_{N_0 - 88, m} \end{bmatrix}, \quad 2 \leq m \leq 91,$$

określa A -, D - i G - optymalny plan $P_{N,6}$ dla $N_0 + 4$ obserwacji ¹⁾.

c.b.d.o.

¹⁾ "Biometrische Zeitschrift", t.15, zesz. 4(1973), str.287
przynosi sprawozdanie z trzeciej NRD-owskiej sesji robo-

Jeżeli w konkretnej sytuacji eksperymentalnej istnieje możliwość wykonania N obserwacji funkcji regresji na kostce m -wymiarowej, to dla zagwarantowania A -, D - i G - optymalności planu dla regresji liniowej można ograniczyć się - zgodnie z twierdzeniem 10 - do wykonania $N - N(\text{mod } 4)$ obserwacji według planu $P_{N-N(\text{mod } 4),6}$. Przy tym musi oczywiście zachodzić nierówność $N - N(\text{mod } 4) \geq N_m = 4 + m - m(\text{mod } 4)$.

czej poświęconej optymalnemu planowaniu eksperymentów, która odbyła się w dniach 2 - 7.10.1972 w Heringsdorf. Sprawozdanie zawiera m.in. streszczenie referatu "G-optimale Versuchsplanung auf dem Würfel und auf der Kugel", który wygłosił W. Jung. W przekładzie polskim streszczenie to brzmi następująco: Dla modelu regresji $y(\underline{x}) = \theta_0 + \theta_1 x_1 + \dots + \theta_k x_k$, gdzie $\underline{x} = (x_1, \dots, x_k)'$, konstruuje się dla $\underline{x} \in \{ \underline{x} : -1 \leq x_i \leq 1, i = 1, \dots, k \}$ względnie dla $\underline{x} \in \{ \underline{x} : \underline{x}'\underline{x} \leq k \}$ konkretne G - optymalne plany eksperymentalne V_n o liczebności n . Wychodzą z tego, że ich macierze informacyjne $M(V_n)$ są równe macierzy jednostkowej (co stanowi warunek dostateczny G - i D - optymalności tych planów), podaje się sposób otrzymywania takich planów z macierzy Hadamarda dla dowolnych wartości k i wartości n podzielnych przez 4. Obok tych wspólnych dla kostki i kuli planów bada się specjalnie dla kuli pewną regułę konstrukcyjną, dostarczającą plany G - optymalne dla dalszych wartości n .

Zobacz też [4], str. 53 - 55.

Liczba czynników (wymiarów kostki) jest wprowadzicie w twierdzeniu 10 ograniczona z góry przez liczbę 91, ale ograniczenie to wydaje się nie mieć konsekwencji praktycznych. Z drugiej strony łatwo podać przykłady planów eksperymentalnych dla N obserwacji, które ze względu na kryteria A -, D -, G - oraz I - optymalności są lepsze od planu eksperymentalnego $P_{N-N(\bmod 4),6}$ dla $N-N(\bmod 4)$ obserwacji. Przykładem takiego planu eksperymentalnego dla N obserwacji na kostce m -wymiarowej jest plan polegający na rozmieszczeniu $N-N(\bmod 4)$ obserwacji według planu $P_{N-N(\bmod 4),6}$ i umieszczeniu pozostałych $N(\bmod 4)$ obserwacji w środku kostki, tzn. w punkcie o współrzędnych $(0, 0, \dots, 0)$.

Rozdział V

REGRESJA LINIOWA NA PROSTOPADŁOŚCIANIE m - WYMIAROWYM

Bardzo często w praktyce w zagadnieniach regresji wielu zmiennych mamy sytuację, w której poszczególne zmienne kontrolowane mogą się zmieniać w określonych przedziałach $[a_i, b_i]$ ($i = 1, 2, \dots, m$). Powstaje wtedy m -wymiarowy prostopadłościan

$$Z = \{z_i: a_i \leq z_i \leq b_i, \quad i = 1, \dots, m\},$$

na którym określona jest liniowa funkcja regresji

$$(5.1) \quad EY_Z = \mathcal{V}_0 + \mathcal{V}_1 z_1 + \mathcal{V}_2 z_2 + \dots + \mathcal{V}_m z_m.$$

Pokażemy, że ważne zadanie wyznaczenia optymalnych planów eksperymentalnych dla estymacji współczynników $\mathcal{V}_0, \mathcal{V}_1, \dots, \mathcal{V}_m$ i całej funkcji regresji (5.1) - przy założeniu, że obserwacje można wykonywać w dowolnym punkcie z prostopadłościanu Z oraz że estymacja funkcji (5.1) jest interesująca w całym zbiorze Z - można rozwiązać w oparciu o skonstruowane w rozdziale poprzednim D - optymalne i G - optymalne plany eksperymentalne na m -wymiarowej kostce X .

Zachodzi mianowicie

TWIERDZENIE 11. Niech na m -wymiarowej kostce $X = \{x_i: -1 \leq x_i \leq 1, \quad i = 1, \dots, m\}$ dany będzie następujący plan

eksperymentalny P_N dla N obserwacji

$$\left\{ \begin{array}{cccc} x^{(1)} & x^{(2)} & \dots & x^{(p)} \\ \lambda_1 & \lambda_2 & \dots & \lambda_p \end{array} \right\},$$

który jest D - optymalnym i G - optymalnym planem eksperymentalnym dla funkcji regresji $EY_x = \Theta_0 + \Theta_1 x_1 + \dots + \Theta_m x_m$.

Niech dana będzie macierz

$$A = \begin{bmatrix} 1 & 0 & 0 & \dots & 0 & 0 \\ \frac{a_1+b_1}{2} & \frac{b_1-a_1}{2} & 0 & \dots & 0 & 0 \\ \frac{a_2+b_2}{2} & 0 & \frac{b_2-a_2}{2} & \dots & 0 & 0 \\ \dots & \dots & \dots & \dots & \dots & \dots \\ \frac{a_m+b_m}{2} & 0 & 0 & \dots & 0 & \frac{b_m-a_m}{2} \end{bmatrix}$$

przekształcająca w $(m+1)$ -wymiarowej przestrzeni euklidesowej m -wymiarową kostkę $X = \{x = (1, x_1, \dots, x_m), -1 \leq x_i \leq 1, i = 1, \dots, m\}$ na m -wymiarowy prostopadłoscian $Z = \{z = (1, z_1, \dots, z_m), a_i \leq z_i \leq b_i, b_i > a_i, i = 1, \dots, m\}$ w ten sposób, że $z = Ax$, tzn.

$$z_i = \frac{(a_i + b_i) + (b_i - a_i)x_i}{2}.$$

Wówczas następujący plan eksperymentalny \tilde{P}_N dla N obserwacji na prostopadłoscianie Z

$$\left\{ \begin{array}{cccc} z^{(1)} & z^{(2)} & \dots & z^{(p)} \\ \tilde{\lambda}_1 & \tilde{\lambda}_2 & \dots & \tilde{\lambda}_p \end{array} \right\},$$

otrzymany z planu P_N w ten sposób, że $\tilde{\lambda}_1 = \lambda_1$

($l = 1, 2, \dots, p$) oraz

$$(5.2) \quad z^{(l)} = \begin{bmatrix} z_1^{(l)} \\ z_2^{(l)} \\ \cdot \\ \cdot \\ \cdot \\ z_m^{(l)} \end{bmatrix} = \begin{bmatrix} \frac{(a_1 + b_1) + (b_1 - a_1)x_1^{(l)}}{2} \\ \frac{(a_2 + b_2) + (b_2 - a_2)x_2^{(l)}}{2} \\ \cdot \\ \cdot \\ \cdot \\ \frac{(a_m + b_m) + (b_m - a_m)x_m^{(l)}}{2} \end{bmatrix}$$

$l = 1, 2, \dots, p$, jest planem D - optymalnym i G - optymalnym dla funkcji regresji

$$EY_z = \vartheta_0 + \vartheta_1 z_1 + \dots + \vartheta_m z_m,$$

którą można obserwować i należy estymować na prostopadłościennym $Z = \{z_i : a_i \leq z_i \leq b_i, i = 1, \dots, m\}$.

Dowód: Nieosobliwa macierz A pozwala na dokonanie odwrotnego przekształcenia prostopadłościennemu Z na kostkę X.

Zgodnie z tym przekształceniem $x = A^{-1}z$, tzn.

$$x_i = \frac{2z_i - a_i - b_i}{b_i - a_i} \quad (i = 1, \dots, m).$$

Zatem funkcję regresji $EY_x = \theta_0 + \theta_1 x_1 + \dots + \theta_m x_m$ można przedstawić w postaci

$$(5.3) \quad EY_{x(z)} = \theta_0 + \theta_1 \left(\frac{2z_1 - a_1 - b_1}{b_1 - a_1} \right) + \dots + \theta_m \left(\frac{2z_m - a_m - b_m}{b_m - a_m} \right) = \theta^T \varphi(z),$$

gdzie $\varphi(z) = \left[1 \quad \frac{2z_1 - a_1 - b_1}{b_1 - a_1} \quad \dots \quad \frac{2z_m - a_m - b_m}{b_m - a_m} \right]^T$.

Estymację współczynników $\theta_0, \theta_1, \dots, \theta_m$ oraz całej funkcji regresji (5.3) można wykonać w oparciu o plan eksperymentalny \tilde{P}_N . Wówczas (por. wzory (1.9) i (1.4))

$$\begin{aligned} d(z, \tilde{P}_N) &= \varphi^T(z) M^{-1}(\tilde{P}_N) \varphi(z) = \\ &= \varphi^T(z) \left[\sum_{l=1}^p \lambda_l \varphi(z^{(l)}) \varphi^T(z^{(l)}) \right]^{-1} \varphi(z) = \\ &= \varphi^T(Az) \left[\sum_{l=1}^p \lambda_l \varphi(Ax^{(l)}) \varphi^T(Ax^{(l)}) \right]^{-1} \varphi(Ax). \end{aligned}$$

Stąd, ponieważ $\varphi(z) = \varphi(Ax) = f(x)$ gdzie $f(x) = [1 \ x_1 \ \dots \ x_m]^T$, otrzymujemy

$$\begin{aligned} d(z, \tilde{P}_N) &= f^T(x) \left[\sum_{l=1}^p \lambda_l f(x) f^T(x) \right]^{-1} f(x) = \\ &= f^T(x) M^{-1}(P_N) f(x) = d(x, P_N). \end{aligned}$$

Równość $d(z, \tilde{P}_N) = d(x, P_N)$ zachodzi dla każdej pary punktów z i x związanej przekształceniem A , wobec czego

$$(5.4) \quad \max_{z \in Z} d(z, \tilde{P}_N) = \max_{x \in X} d(x, P_N).$$

Plan P_N jest planem D - optymalnym i G - optymalnym, więc $\max_{x \in X} d(x, P_N) = m + 1$. Wobec tego z równości (5.4) i twierdzenia 2 wynika D - optymalność i G - optymalność planu P_N dla estymacji funkcji regresji (5.3) na prostopadłościźnie Z .

Funkcję regresji (5.3) przekształćmy następująco:

$$EY_Z = (\theta_0 - \theta_1 - \dots - \theta_m) + \frac{2}{b_1 - a_1} \theta_1 z_1 + \dots + \frac{2}{b_m - a_m} \theta_m z_m$$

i wprowadźmy oznaczenia: $\psi_0 = \theta_0 - \theta_1 - \dots - \theta_m$

$$\varphi_i = \frac{2}{b_i - a_i} \mathbf{0}_i \quad (i = 1, \dots, m)$$

$$\varphi = \begin{bmatrix} \varphi_0 & \varphi_1 & \dots & \varphi_m \end{bmatrix}^T$$

$$f(z) = \begin{bmatrix} 1 & z_1 & \dots & z_m \end{bmatrix}^T.$$

Zatem dla każdego $z \in Z$ mamy $EY_z = \theta^T \varphi(z) = \varphi^T f(z)$ przy czym $\varphi = B\theta$ oraz $f(z) = (B^T)^{-1} \varphi(z)$, gdzie

$$B = \begin{bmatrix} 1 & -1 & -1 & \dots & -1 & -1 \\ 0 & \frac{2}{b_1 - a_1} & 0 & \dots & 0 & 0 \\ 0 & 0 & \frac{2}{b_2 - a_2} & \dots & 0 & 0 \\ \dots & \dots & \dots & \dots & \dots & \dots \\ 0 & 0 & 0 & \dots & 0 & \frac{2}{b_m - a_m} \end{bmatrix}$$

Zastosujemy plan eksperymentalny \tilde{P}_N do estymacji współczynników $\varphi_0, \varphi_1, \dots, \varphi_m$ oraz całej funkcji regresji $EY_z = \varphi^T f(z) = \varphi_0 + \varphi_1 z_1 + \dots + \varphi_m z_m$. Macierz $M^*(\tilde{P}_N)$ (gwiazdką zaznaczono macierz M planu \tilde{P}_N dla funkcji $EY_z = \varphi^T f(z)$ w odróżnieniu od macierzy M tego samego planu dla funkcji $EY_z = \theta^T \varphi(z)$) ma tutaj postać

$$M^*(\tilde{P}_N) = \sum_{l=1}^p \lambda_l f(z^{(l)}) f^T(z^{(l)}),$$

$$\text{a funkcja } d^*(z, \tilde{P}_N) = f^T(z) [M^*(\tilde{P}_N)]^{-1} f(z) =$$

$$= f^T(z) \left[\sum_{l=1}^p \lambda_l f(z^{(l)}) f^T(z^{(l)}) \right]^{-1} f(z) =$$

$$= \varphi^T(z) B^{-1} \left[\sum_{l=1}^p \lambda_l (B^T)^{-1} \varphi(z^{(l)}) \varphi^T(z^{(l)}) B^{-1} \right]^{-1} (B^T)^{-1} \varphi(z) =$$

$$\begin{aligned}
 &= \varphi^T(z) B^{-1} \left\{ (B^T)^{-1} \left[\sum_{l=1}^p \lambda_l \varphi(z^{(l)}) \varphi^T(z^{(l)}) \right] B^{-1} \right\}^{-1} (B^T)^{-1} \varphi(z) \\
 &= \varphi^T(z) B^{-1} B \left[\sum_{l=1}^p \lambda_l \varphi(z^{(l)}) \varphi^T(z^{(l)}) \right]^{-1} B^T (B^T)^{-1} \varphi(z) = \\
 &= \varphi^T(z) \left[\sum_{l=1}^p \lambda_l \varphi(z^{(l)}) \varphi^T(z^{(l)}) \right]^{-1} \varphi(z) = \\
 &= \varphi^T(z) M^{-1}(\tilde{P}_N) \varphi(z) = d(z, \tilde{P}_N).
 \end{aligned}$$

Otrzymaliśmy więc, że $d^*(z, \tilde{P}_N) = d(z, \tilde{P}_N)$. Wcześniej pokazaliśmy, że $\max_{z \in Z} d(z, \tilde{P}_N) = m + 1$, wobec czego również $\max_{z \in Z} d^*(z, \tilde{P}_N) = m + 1$, co zgodnie z twierdzeniem 2 oznacza, że plan eksperymentalny P_N jest D - optymalny i G - optymalny dla estymacji funkcji regresji

$$EY_Z = \varphi_0 + \varphi_1 z_1 + \dots + \varphi_m z_m$$

na prostopadłościannie $Z = \{z_i: a_i \leq z_i \leq b_i, i = 1, \dots, m\}$

c.b.d.o.

Twierdzenie 11 stanowi specjalny przypadek ogólniejszych twierdzeń podanych w [9], § 2.2 .

Rozdział VI

KOMPUTEROWE BADANIE EFEKTYWNOŚCI PLANOWANIA EKSPERYMENTÓW

W bardzo wielu zagadnieniach eksperymentalnych spotykamy się z sytuacją, w której realizacja eksperymentu w sposób optymalny jest niemożliwa, albo poważnie utrudniona. Nie można mówić o optymalnym planowaniu eksperymentu np. wtedy, gdy eksperymentator nie ma wpływu na wartości zmiennych x_1, x_2, \dots, x_m , a planowanie ogranicza się co najwyżej do wyznaczenia liczby obserwacji. Poważnie utrudniona jest realizacja eksperymentu zgodnie z planem optymalnym wtedy, gdy badacz jest w stanie utrzymywać wartości zmiennych x_1, x_2, \dots, x_m jedynie w pobliżu wartości optymalnych. Taki przypadek występuje dość często w praktyce ze względu na to, że na ogół optymalne wartości x_1, x_2, \dots, x_m są zarazem wartościami ekstremalnymi obszaru obserwacji X_1 . Zdarzają się również sytuacje, w których - dla uniknięcia związanego z tym wysiłku organizacyjnego - świadomie rezygnuje się z planu optymalnego decydując się na eksperyment bierny, w którym wartości zmiennych x_1, x_2, \dots, x_m nie są ustalone z góry. Z takim postępowaniem spotykamy się często wtedy, gdy wykonanie obserwacji nie wymaga wielkich nakładów czasu, środków i materiałów, wobec czego rezygnację z planu optymalnego można

zrekompensować przez wykonanie odpowiednio większej liczby obserwacji. Jeśli zatem eksperyment nie jest realizowany zgodnie z planem, który jest optymalny ze względu na określone kryterium optymalności, to ocenę niektórych aspektów strat wynikających z odstępstwa od optymalności można otrzymać w odpowiedzi na dwa następujące pytania:

- 1^o Jaka jest wartość liczbowej charakterystyki dobroci eksperymentu polegającego na wykonaniu N obserwacji według pewnego planu, o którym wiadomo, że nie jest planem optymalnym, w porównaniu z wartością analogicznej liczbowej charakterystyki dobroci eksperymentu polegającego na wykonaniu tej samej liczby N obserwacji, ale zgodnie z planem spełniającym określone kryterium optymalności ?
- 2^o Ile obserwacji trzeba wykonać w ramach eksperymentu realizowanego przez badacza w pewien przyjęty przez niego sposób, aby tak przeprowadzony eksperyment był co najmniej tak dobry, jak eksperyment polegający na wykonaniu N obserwacji zgodnie z planem spełniającym określone kryterium optymalności ?

Oceny strat wynikłych z odstępstwa od planów optymalnych badane będą dla eksperymentów mających na celu estymację liniowej funkcji regresji

$$EY_x = \theta_0 + \theta_1 x_1 + \dots + \theta_m x_m$$

na zbiorze X , który jest kostką m -wymiarową (por. określenie podane w rozdz. IV), przy czym za kryterium optymalności przyjmuje się kryterium G - optymalności. Zatem liczbową charakterystyką dobroci eksperymentu polegającego na wykonaniu n obserwacji według pewnego planu P_n jest wyrażenie (por.

wzór (1.7) oraz kryterium 3, rozdz. I)

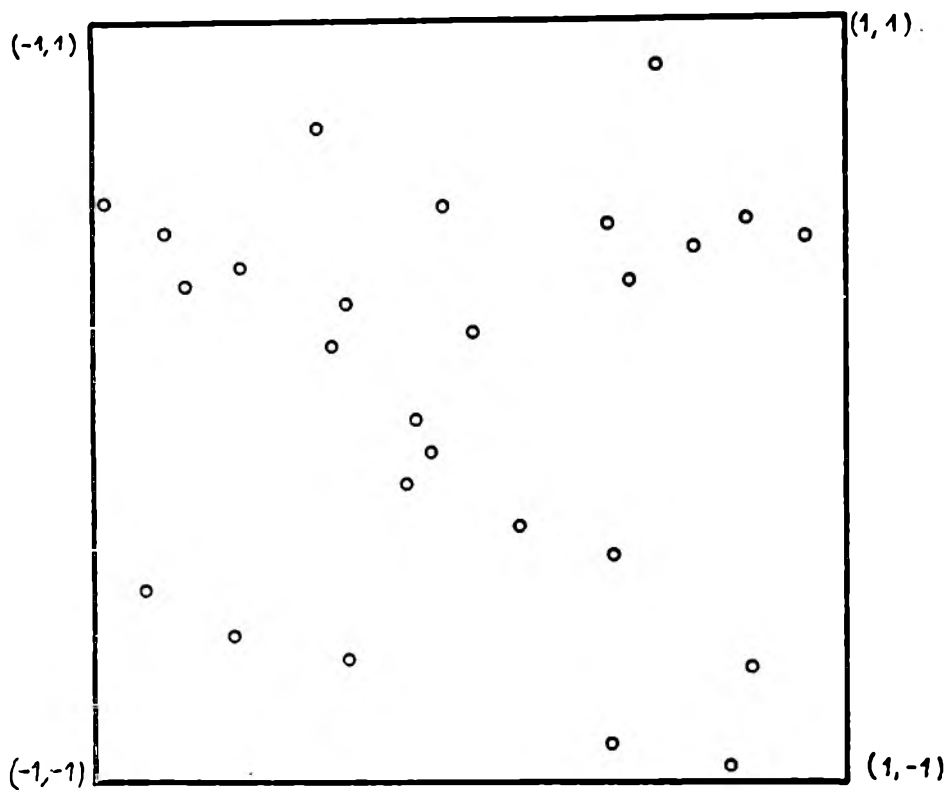
$$(6.1) \quad \max_{x \in X} \frac{1}{n} d(x, P_n) ,$$

którego wartość będziemy dalej nazywać *dokładnością* eksperymentu przeprowadzonego zgodnie z planem P_n . Wobec tego odpowiedzi na pytania 1^o i 2^o będą polegały na dokonaniu odpowiednich porównań pomiędzy dokładnością eksperymentu przeprowadzonego zgodnie z planem G - optymalnym dla N obserwacji z dokładnością eksperymentów przeprowadzanych przy innych - akceptowanych przez badacza - sposobach rozmieszczania obserwacji. W dalszym ciągu uwzględnione zostaną dwa sposoby rozmieszczania obserwacji, które zrealizowano metodą symulacji komputerowej.

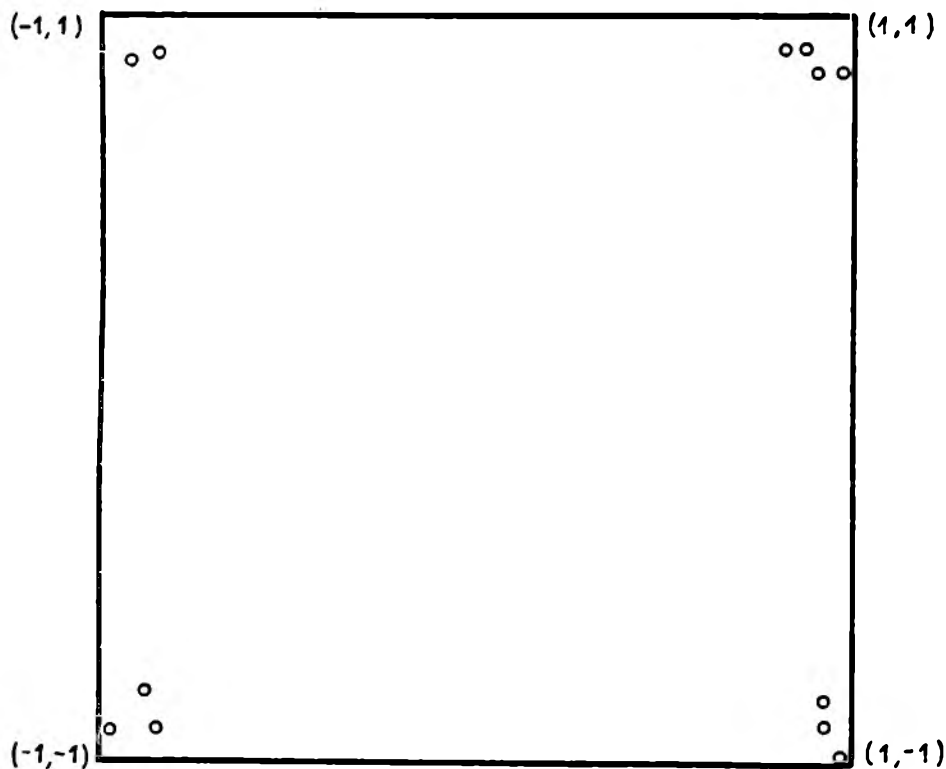
Pierwszy sposób rozmieszczania obserwacji polega na tym, że obserwacje wykonuje się w punktach $(x_1^{(1)}, x_2^{(1)}, \dots, x_m^{(1)})$, $(x_1^{(2)}, x_2^{(2)}, \dots, x_m^{(2)})$, ... otrzymanych z generatora liczb pseudolosowych o rozkładzie określonym gęstością

$$g_m(x) = g_m(x_1, x_2, \dots, x_m) = \begin{cases} \frac{1}{2^m} & \text{dla } -1 \leq x_i < 1, \quad i = 1, \dots, m \\ 0 & \text{poza tym,} \end{cases}$$

tzn. o rozkładzie jednostajnym na m -wymiarowej kostce X . Pewne skończone ciągi tak wygenerowanych punktów kostki X będą dalej nazywane planami losowymi. Rys. 6.1 przedstawia rozmieszczenie 26 punktów wygenerowanych według rozkładu $g_2(x_1, x_2)$ na kwadracie X ($m = 2$), które wyznaczają plan losowy dla wykonania 26 obserwacji. Plany losowe są modelem sytuacji, w której eksperymentator nie ma wpływu na wartość zmiennych x_1, x_2, \dots, x_m .



Rys. 6.1.
Rozmieszczenie ciągu 26 punktów wygenerowanych zgodnie z rozkładem $g_2(x_1, x_2)$ na kwadracie $X = \{x = (x_1, x_2), -1 \leq x_i \leq 1, i = 1, 2\}$



Rys. 6.2.
Rozmieszczenie ciągu 12 punktów wygenerowanych zgodnie z rozkładem $\tilde{g}_2(x_1, x_2)$ na kwadracie $X = \{x = (x_1, x_2), -1 \leq x_i \leq 1, i = 1, 2\}$

Drugi sposób rozmieszczania obserwacji polega na tym, że obserwacje wykonuje się w punktach $(x_1^{(1)}, x_2^{(1)}, \dots, x_m^{(1)})$, $(x_1^{(2)}, x_2^{(2)}, \dots, x_m^{(2)})$, ... otrzymanych z generatora liczb pseudolosowych o rozkładzie określonym gęstością

$$\tilde{g}_m(x) = \tilde{g}_m(x_1, x_2, \dots, x_m) = \begin{cases} \left(\frac{5}{2}\right)^m & \text{jeśli } \bigcap_{i=1}^m (0,8 \leq |x_i| \leq 1) \\ 0 & \text{poza tym .} \end{cases}$$

Pewne skończone ciągi tak wygenerowanych punktów kostki X będą dalej nazywane planami quasioptymalnymi. Rys. 6.2 przedstawia rozmieszczenie 12 punktów wygenerowanych według rozkładu $\tilde{g}_2(x_1, x_2)$ na kwadracie X , które wyznaczają plan quasioptymalny dla wykonania 12 obserwacji. Plany quasioptymalne są modelem sytuacji, w których eksperymentator jest w stanie utrzymywać wartości zmiennych x_1, x_2, \dots, x_m jedynie w pobliżu wartości optymalnych. (Przypominamy, że współrzędne punktów planów G - optymalnych dla estymacji liniowej funkcji regresji na kostce X są równe $+1$ lub -1 , Zob. też rozdz. IV).

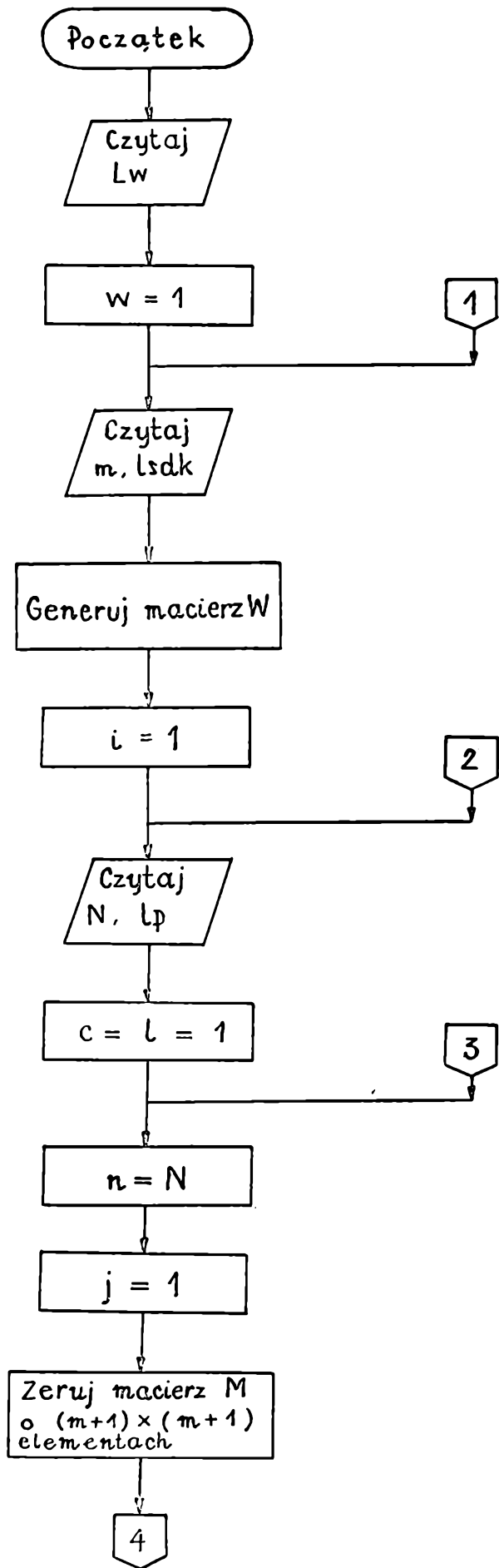
Zakres i algorytm obliczeń. Porównywanie - ze względu na wartość wyrażenia (6.1) - planów losowych i quasioptymalnych z planami G - optymalnymi realizowane było na komputerze ODRA 1204 w ośrodku obliczeniowym Instytutu Cybernetyki Ekonomicznej Akademii Ekonomicznej we Wrocławiu. Obliczenia wykonywano dla kostki X o wymiarach 2, 3, 4 i 5. Dla ustalonego wymiaru m ($m = 2, 3, 4, 5$) kostki X bierze się pod uwagę plany G - optymalne dla wykonania 8, 12, 16, 20, 24, 28, 32 obserwacji. Dla ustalonej liczby N ($N = 8, 12, 16, 20, 24, 28, 32$) obserwacji oblicza się (ze wzoru

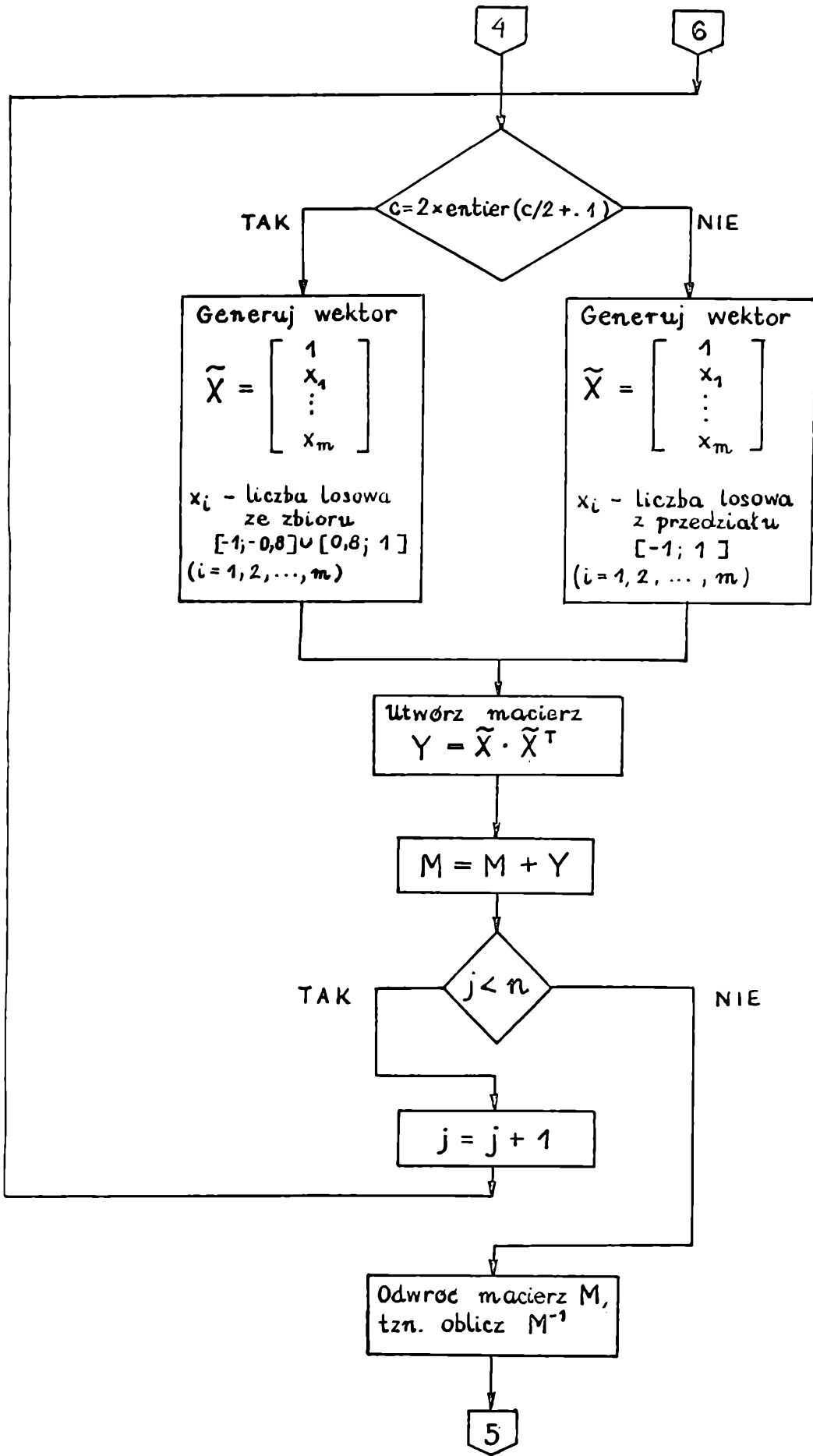
(6.1) i twierdzenia 2) dokładność eksperymentu zrealizowanego według planu G - optymalnego. Dokładność ta jest równa $\frac{m+1}{N}$. Z kolei dla ustalonych m i N powtarza się 30 razy następujące postępowanie: Generuje się N punktów według rozkładu $g_m(x)$, oblicza się dokładność eksperymentu wyznaczonego przez tak otrzymany plan losowy dla N obserwacji, a następnie powiększa się N -elementowy zbiór wygenerowanych punktów o dalsze punkty generowane według rozkładu $g_m(x)$ aż do chwili otrzymania n -elementowego ($n > N$) zbioru punktów wyznaczających plan P_n , dla którego

$$(6.2) \quad \max_{x \in X} \frac{1}{n} d(x, P_n) \leq \frac{m+1}{N}.$$

Na podstawie wyników 30-krotnego powtórzenia opisanego postępowania oblicza się średnią dokładność eksperymentu przeprowadzonego zgodnie z planem losowym o liczebności N oraz wariancję i odchylenie standardowe tej dokładności, a także średnią liczebność planów losowych spełniających warunek (6.2) oraz wariancję i odchylenie standardowe tej liczebności. Analogiczną procedurę powtarza się 30 razy dla planów quasioptymalnych, których punkty generowane są według rozkładu $\tilde{g}_m(x)$, a w końcu oblicza się wartość średnią, wariancję i odchylenie standardowe dokładności eksperymentu przeprowadzonego zgodnie z planem quasioptymalnym o liczebności N oraz wartość średnią, wariancję i odchylenie standardowe liczebności planów quasioptymalnych spełniających warunek (6.2).

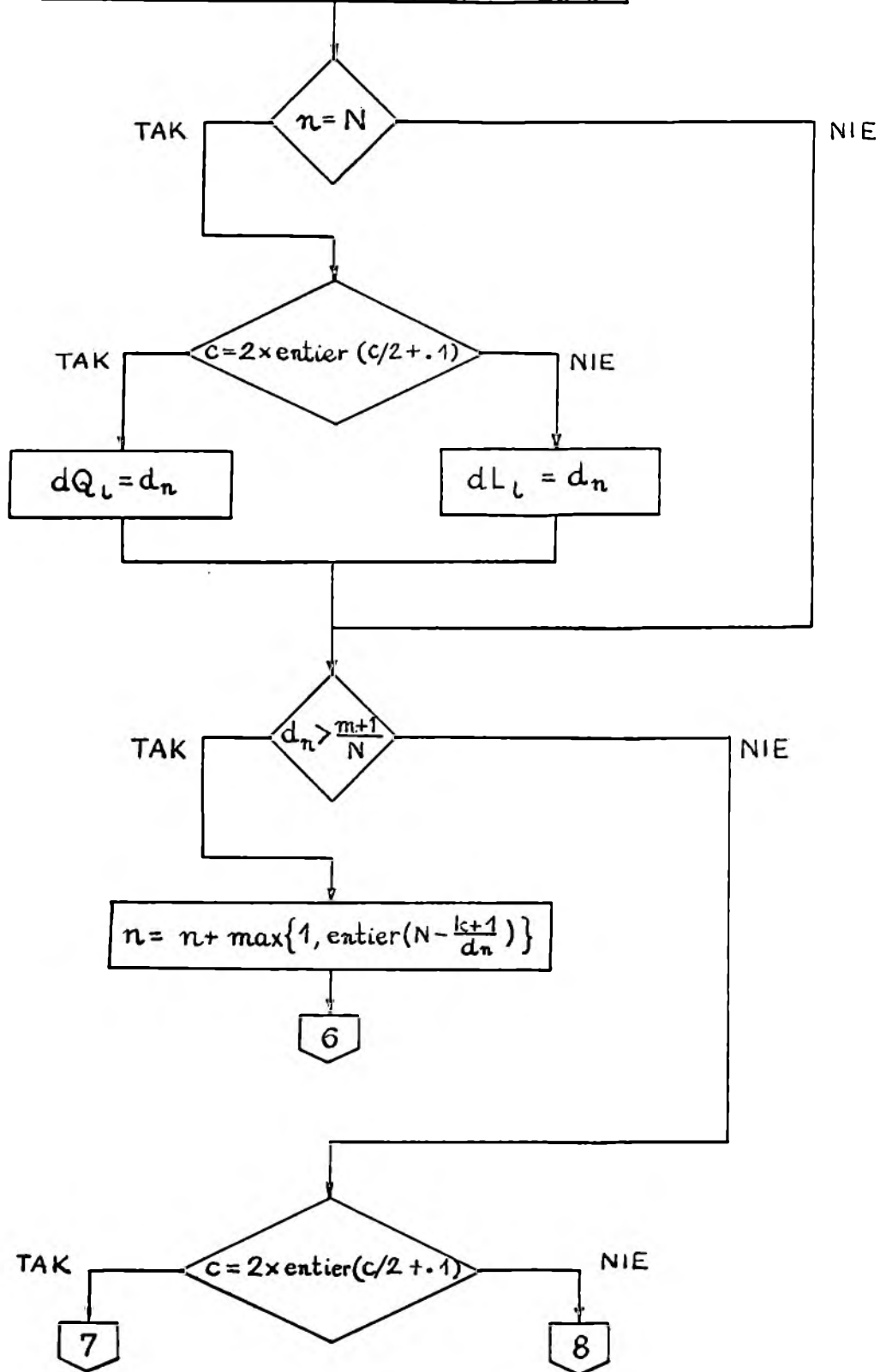
Algorytm obliczeń przedstawiono w postaci schematu blokowego (patrz str. 103 - 106) oraz programu napisanego w języku ALGOL 1204 (patrz str. 117 - 122).

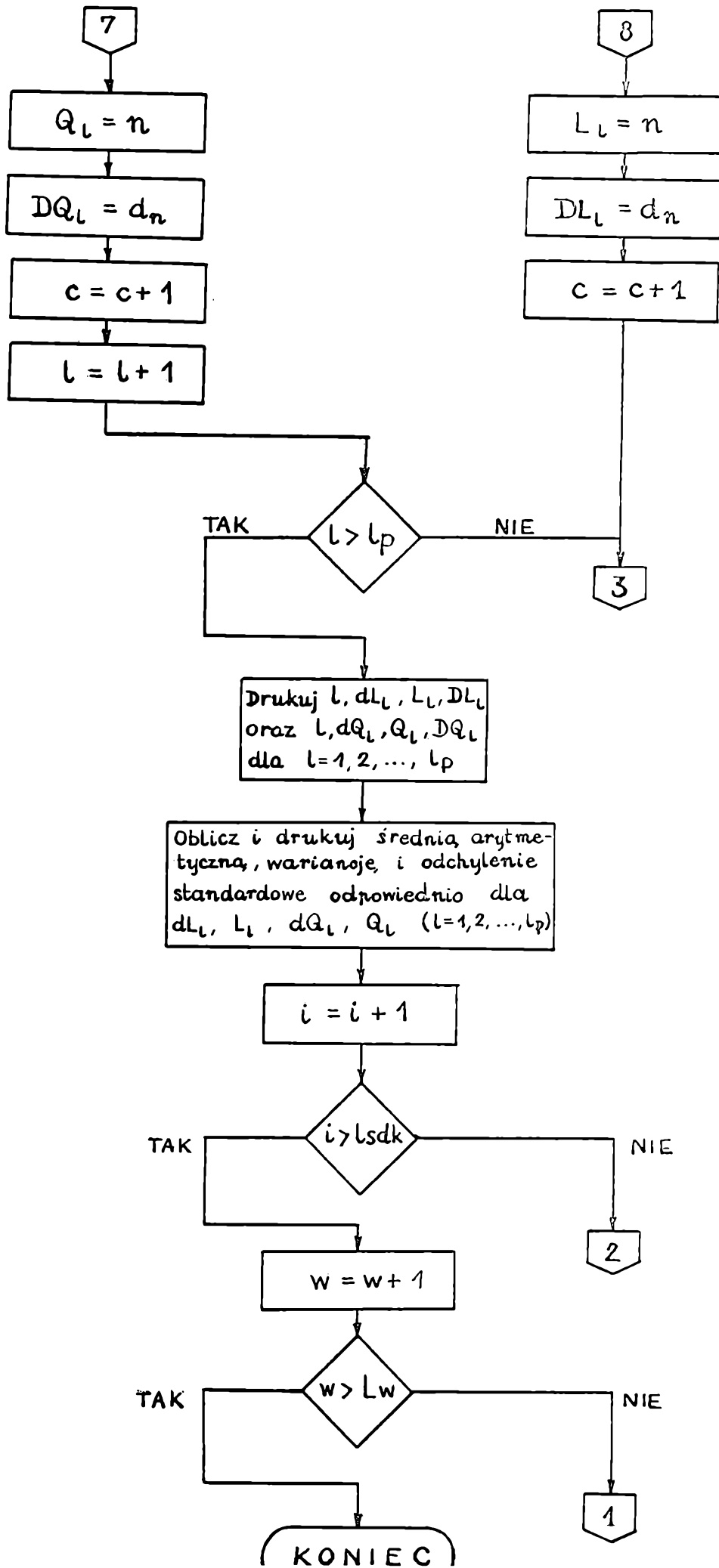




5

Oblicz $d_n = \max_{k=1, \dots, 2^m} \tilde{X}_k^T M^{-1} \tilde{X}_k$
gdzie: $\tilde{X}_k^T = [1 \ x_{k,1} \ x_{k,2} \ \dots \ x_{k,m}]$
 $x_{k,i}$ - element macierzy W na przecięciu wiersza k ($k=1, 2, \dots, 2^m$) oraz kolumny i ($i=1, 2, \dots, m$)





Użyte w schemacie blokowym symbole wczytywanych danych mają następujące znaczenie:

- Lw - liczba kostek o różnych wymiarach,
m - wymiar kostki,
lsdk - liczba różnych wartości dokładności eksperymentów na kostce m-wymiarowej,
N - liczba obserwacji w planie G - optymalnym, określająca wartość dokładności eksperymentu jako równą $(m+1)/N$,
lp - liczba powtórzeń podprogramu, który generuje plany losowe lub quasiptymalne o dokładności równej $(m+1)/N$.

Przy opisanym wyżej zakresie obliczeń mamy:

- Lw = 4 (generowanie punktów odbywa się dla kostek o wymiarach 2, 3, 4, 5)
m = 2, 3, 4, 5
lsdk = 7 (dla kostki o ustalonym wymiarze m określa się 7 wartości dokładności eksperymentu równych $\frac{m+1}{N}$, gdzie N przyjmuje 7 różnych wartości)
N = 8, 12, 16, 20, 24, 28, 32
lp = 30

Generowanie macierzy W w współrzędnych wierzchołków kostki m-wymiarowej odbywa się w programie w ten sposób, że w k-tej ($k = 1, 2, \dots, m$) kolumnie macierzy W występuje na przemian 2^{m-k} elementów równych $+1$ i 2^{m-k} elementów równych -1 . Np. macierz W w współrzędnych wierzchołków kostki trójwymiarowej ma postać:

$$W = \begin{bmatrix} 1 & 1 & 1 \\ 1 & 1 & -1 \\ 1 & -1 & 1 \\ 1 & -1 & -1 \\ -1 & 1 & 1 \\ -1 & 1 & -1 \\ -1 & -1 & 1 \\ -1 & -1 & -1 \end{bmatrix}$$

Dla generowania punktów zgodnie z rozkładem $g_m(x)$ lub $\tilde{g}_m(x)$ wykorzystuje się w programie funkcję "unif", znajdującą się w bibliotece programów języka ALGOL 1204. Kolejnymi wartościami funkcji "unif" są liczby pseudolosowe o rozkładzie jednostajnym w przedziale $(0, 1)$. Współrzędne punktu $x = (x_1, x_2, \dots, x_m)$ generowanego zgodnie z rozkładem $g_m(x)$ otrzymuje się wtedy z równania

$$x_k = 2 \cdot \text{unif} - 1 \quad (k = 1, 2, \dots, m),$$

natomiast k -tą ($k = 1, 2, \dots, m$) współrzędną punktu generowanego zgodnie z rozkładem $\tilde{g}_m(x)$ otrzymuje się według wzoru

$$x_k = \begin{cases} 0,4 \cdot \text{unif} - 1 & \text{jeśli } \text{unif} < 0,5 \\ 0,4 \cdot \text{unif} + 0,6 & \text{jeśli } \text{unif} \geq 0,5 \end{cases}$$

Odwracanie macierzy realizowane jest w programie przy użyciu procedury bibliotecznej "inver 1". Przy pomocy procedury "inver 1" można odwracać dodatnio określone macierze symetryczne.

Dokładność d_n eksperymentu zrealizowanego według planu P_n dla n obserwacji wyznacza się w algorytmie według wzoru

$$d_n = \max_{k=1, \dots, 2^m} \tilde{X}_k^T M^{-1} \tilde{X}_k .$$

Postępowanie to jest konsekwencją tego, że funkcja $\tilde{X}^T M^{-1} \tilde{X}$ jest funkcją wypukłą zmiennych x_1, x_2, \dots, x_m (por. np. [4], str. 62). Macierz M oraz wektory \tilde{X} i \tilde{X}_k określone są tak, jak w załączonym schemacie blokowym.

Załączony tekst programu zawiera dwie instrukcje warunkowe umożliwiające drukowanie współrzędnych punktów wybranych planów losowych lub quasioptymalnych. Należy zaznaczyć, że zapisane są one niezgodnie ze składnią języka ALGOL. W konkretnej sytuacji należy albo wyeliminować obie te instrukcje, albo zmodyfikować je w sposób określony w komentarzach poprzedzających te instrukcje.

Wymiar kostki jest oznaczony w tekście programu symbolem k .

Omówienie wyników. Kompletne tabulogramy wyników obliczeń zamieszczono na str. 123-150. Jedna strona tabulogramu zawiera wyniki dotyczące ustalonego wymiaru m kostki X i ustalonej liczebności N planu G - optymalnego dla estymacji liniowej funkcji regresji na kostce X . Dla ustalonych wartości m i N na tabulogramie podano:

- dokładność eksperymentu zrealizowanego według planu G - optymalnego, równą $(m+1)/N$;
- dla każdego z 30 generowanych planów losowych oraz dla każdego z 30 generowanych planów quasioptymalnych:
 - dokładność eksperymentu polegającego na wykonaniu N obserwacji w N pierwszych punktach otrzymanych w trakcie generowania odpowiedniego planu (patrz kolumna

- tabulogramu oznaczona symbolem dN) ;
- liczebność wygenerowanego planu o dokładności $d \leq (m+1)/N$ (patrz kolumna tabulogramu oznaczona symbolem n) ;
 - kontrolny wydruk dokładności wygenerowanego planu (patrz kolumna tabulogramu oznaczona symbolem d) ;
 - wartość średnią, wariancję i odchylenie standardowe dokładności eksperymentu przeprowadzonego według planu losowego o liczebności N oraz wartość średnią, wariancję i odchylenie standardowe dokładności eksperymentu przeprowadzonego według planu quasioptymalnego o liczebności N (patrz poniżej kolumn oznaczonych symbolem dN) ;
 - wartość średnią, wariancję i odchylenie standardowe liczebności planów losowych spełniających warunek (6.2) oraz wartość średnią, wariancję i odchylenie standardowe liczebności planów quasioptymalnych spełniających warunek (6.2) (patrz poniżej kolumn oznaczonych symbolem n).

Na podstawie wyników zawartych w tabulogramach zbudowano dwie tablice, które dostarczają - mającej charakter statystyczny - odpowiedzi na pytania 1^o i 2^o postawione na początku rozdziału.

Tablica 6.1 zawiera odpowiedź na pytanie 1^o. Odpowiedź ta dostarcza liczbowej oceny średniej dokładności eksperymentów wykonanych w sposób nieoptymalny na tle znanej dokładności eksperymentów wykonanych według planu G -optymalnego. W jednym wierszu tej tablicy podaje się - dla kostki X o ustalonym wymiarze - dokładność eksperymentu polegają-

T A B L I C A 6.1

Dokładność eksperymentów przeprowadzonych zgodnie z G- optymalnymi, quasioptymalnymi i losowymi planami eksperymentalnymi

Wymiar kostki	Liczba obserwacji	Dokładność eksperymentu		
		Plan G- opty- malny	Plan quasioptymalny (wartość średnia)	Plan losowy (wartość średnia)
2	8	.375	15.159	2.312
	12	.250	.499	1.222
	16	.188	.397	.758
	20	.150	.271	.515
	24	.125	.213	.429
	28	.107	.164	.366
	32	.094	.154	.308
3	8	.500	9.278	5.052
	12	.333	1.149	2.122
	16	.250	.621	1.427
	20	.200	.415	.955
	24	.167	.333	.689
	28	.143	.263	.599
	32	.125	.231	.475
4	8	.625	8.914	17.944
	12	.417	1.875	4.369
	16	.313	1.066	2.290
	20	.250	.664	1.386
	24	.208	.512	1.068
	28	.179	.374	.850
	32	.156	.325	.657
5	8	.750	323.037	25.226
	12	.500	6.661	6.446
	16	.375	1.507	3.539
	20	.300	.941	2.176
	24	.250	.664	1.450
	28	.214	.527	1.259
	32	.188	.459	.940

T A B L I C A 6.2

Liczba obserwacji wykonywanych zgodnie z planami G- optymalnymi, quasioptymalnymi i losowymi, które są niezbędne do osiągnięcia zadanej dokładności eksperymentu.

Wymiar kostki	Dokładność eksperymentu	Liczba obserwacji		
		Plan G- optymalny	Plan quasioptymalny (wartość średnia)	Plan losowy (wartość średnia)
2	.375	8	16	26
	.250	12	20	37
	.188	16	28	48
	.150	20	32	59
	.125	24	38	68
	.107	28	42	79
	.094	32	49	90
3	.500	8	18	31
	.333	12	26	43
	.250	16	30	55
	.200	20	35	66
	.167	24	42	81
	.143	28	47	90
	.125	32	53	99
4	.625	8	20	34
	.417	12	28	47
	.313	16	33	61
	.250	20	39	70
	.208	24	45	86
	.179	28	49	97
	.156	32	55	106
5	.750	8	24	38
	.500	12	30	54
	.375	16	35	64
	.300	20	42	78
	.250	24	47	89
	.214	28	54	103
	.188	32	61	114

cego na wykonaniu ustalonej liczby obserwacji odpowiednio według planu G - optymalnego, planu quasioptymalnego i planu losowego. Uderzająca niedokładność (duża wartość średnia wyrażenia (6.1)) eksperymentów polegających na wykonaniu 8 obserwacji według planu quasioptymalnego wynika np. z tego, że w niektórych planach quasioptymalnych rozkład 8 wygenerowanych punktów okazuje się wyraźnie niesymetryczny względem środka kostki X.

Tablica 6.2 zawiera odpowiedź na pytanie 2^o. Odpowiedź ta określa średnią liczebność obserwacji w eksperymentach przeprowadzonych w sposób nieoptymalny, których dokładność ma być niegorsza od dokładności eksperymentów wykonanych według planu G - optymalnego dla zadanej liczby obserwacji. W jednym wierszu tej tablicy podaje się - dla kostki X o ustalonym wymiarze - liczebności eksperymentów wykonywanych według planu G - optymalnego, planu quasioptymalnego i planu losowego, które osiągają jednakową zadaną dokładność.

U W A G I K O Ń C O W E

1. Przedstawione w rozdziale I kryteria optymalności planu eksperymentu, mającego na celu estymację współczynników $\theta_1, \theta_2, \dots, \theta_k$ lub estymację całej funkcji regresji (1.2), określają pewien wachlarz pożądanych własności estymatorów tych współczynników lub estymatora funkcji regresji. Eksperyment mający na celu estymację współczynników $\theta_1, \theta_2, \dots, \theta_k$ funkcji (1.2) zaleca się realizować zgodnie z planem spełniającym kryterium A - optymalności lub kryterium D - optymalności. Eksperyment mający na celu estymację całej funkcji regresji (1.2) zaleca się realizować według planu spełniającego kryterium G - optymalności lub kryterium I - optymalności. W różnych przypadkach szczególnych plan eksperymentalny spełniający określone kryterium optymalności spełnia również inne kryteria optymalności.
2. Dla przypadku, gdy funkcja regresji jest liniową funkcją jednej, dwu lub - ogólnie - m zmiennych kontrolowanych, natomiast obszar obserwacji jest odpowiednio odcinkiem prostej rzeczywistej, kwadratem na płaszczyźnie, kostką lub prostopadłością m -wymiarową, zaprezentowano w rozdziałach II - V plany eksperymentalne, które spełniają określone kryteria optymalności.
3. Wykonane porównania - ze względu na kryterium G - optymalności - eksperymentów realizowanych według planów optymalnych z eksperymentami biernymi, w których obserwacje rozmieszcza się w sposób losowy, dowodzą wysokiej

efektywności planów optymalnych (por. tablice 2.1, 3.1, 6.1 i 6.2). Przy zadanej liczebności próby wykonanie obserwacji zgodnie z planem optymalnym pozwala znacznie (w zbadanych przypadkach co najmniej dwukrotnie) poprawić dokładność estymacji (por. tablice 2.1, 3.1 i 6.1). Podobnie, przy ustalonej dokładności estymacji, zastosowanie planu G - optymalnego pozwala w zbadanych przypadkach na zmniejszenie liczebności próby co najmniej o około 60% w porównaniu z liczebnością próby konieczną w eksperymencie biernym (por. tablica 6.2). Wyniki badań zrealizowanych przy pomocy komputera wskazują również (por. tablice 6.1 i 6.2) na wielkie znaczenie eksperymentu quasioptymalnego, który stanowi praktycznie możliwą aproksymację eksperymentu przeprowadzonego ściśle według planu optymalnego. Planuje się przeprowadzenie analogicznych badań efektywności eksperymentu również ze względu na inne kryteria optymalności.

4. W kręgu zainteresowań autora znajdują się problemy konstrukcji planów eksperymentalnych dla innych postaci funkcji regresji (np. funkcji kwadratowej wielu zmiennych) oraz innych obszarów obserwacji. Ponadto dostrzega się potrzebę rozwiązania tych zagadnień przy założeniu, że obserwacje są skorelowane.

TABULOGRAM PROGRAMU OBLICZEŃ

I

TABULOGRAMY WYNIKÓW

```
begin  
  comment program generowania losowych i quasioptymalnych planow  
    eksperymentalnych;  
  integer c,e,i,j,k,ko,l,lp,lSDK,Lw,N,n,w,wi,wie;  
  real d,dN,dr,glvarx,glvaralfa,sl,sq,slN,sqN,val,vaq,vlN,vqN;  
  real procedure unif;  
    begin  
      glvarx=glvaralfa*glvarx;  
      unif=glvarx=glvarx-entier(glvarx)  
    end unif;  
  procedure inver1(n,a, fail);  
    value n;  
    integer n;  
    array a;  
    label fail;  
    begin  
      integer i,ik,i1,j,j1,k;  
      real x,y;  
      for i:=1 step 1 until n do  
        begin  
          i1:=i+1;  
          for j:=i step 1 until n do  
            begin  
              j1:=j+1;  
              x=a[i,j];  
              ik=i-1;  
              for k:=1 step 1 until ik do x=x-a[j1,k]*a[i1,k];  
              if j=i then begin  
                if x<0 ∨ x=0 then go to fail;  
                a[i1,i]:=y=1/sqrt(x);  
              end j=i  
              else a[j1,i]:=x*y  
            end j  
          end i;  
        end i;  
      end i;  
    end i;  
  end i;  
end i;
```

```
for i:=1 step 1 until n do
  for j:=i+1 step 1 until n do
    begin
      x:=0;
      j1:=j+1;
      for k=j-1 step -1 until i do x=x-a[j1,k]*a[k+1,i];
      a[j1,i]:=x*a[j1,j]
    end j,i;
  for i:=1 step 1 until n do
    for j:=i step 1 until n do
      begin
        x:=0;
        j1:=n+1;
        for k=j+1 step 1 until j1 do x=x+a[k,j]*a[k,i];
        a[j+1,i]:=x
      end j,i
    end inver1;
  read(glvaralfa,Lw);
  glvarx=.12345678901;
  w:=1;
  wymiar: read(k,l sdk);
  begin
    integer array W[1:2↑k,1:k];
    array DL,DQ,d1,dQ,L,Q[1:30],M[0:k+1,0:k],MO[0:k,0:k],X[0:k],
      a[1:k+2,1:k+1];
    X[0]:=1;
    for ko:=1 step 1 until k do
      begin
        e:=2↑(k-ko);
        c:=wie:=0;
        for:for wi:=wie+1 step 1 until e+wie do
          if c=2*entier(c/2+.1) then W[wi,ko]:=1 else W[wi,ko]:=-1;
          c:=c+1;
          wie:=wie+e;
          if wie<2↑k then go to for
        end generowanie W;
      i:=1;
```



```
efekt:read(N,lp);
  c:=l:=1;
liczba:n:=N;
  j:=1;
  for wi:=0 step 1 until k do
    for ko:=wi step 1 until k do M[wi,ko]:=0;
comment Ponizsza instrukcje warunkowa nalezy pominac
        lub zastapic <WARUNEK> wyrazeniem logicznym
        zawierajacym zmienne k,l,M;
if <WARUNEK> then
  begin
    format('12');
    print('?????DOKLADNOSC',k+1,'/',N,'PLAN');
    if c=2 × entier(c/2+.1) then print('QUASIOPTYMALNY');
    else print('LOSOWY');

    format('XXXXX1');
    for ko:=1 step 1 until k do print(ko);
    line(2)
  end;
wersja:if c=2 × entier(c/2+.1) then go to wersja2;
  for e:=1 step 1 until k do X[e]:=2 × unif - 1;
  go to x;
wersja2: for e:=1 step 1 until k do
  begin
    d:=unif;
    if d<.5 then X[e]:= .4×d-1
    else X[e]:= .4×d+.6
  end;
comment Decyzja o pominięciu lub odpowiednim zastosowaniu
ponizszej instrukcji warunkowej musi byc identyczna
z decyzja podjeta w wyniku komentarza poprzedniego;
x:if <WARUNEK> then
  begin
    format('123');
    print(j);
    format('1.234');
    for ko:=1 step 1 until k do print(X[k]);
    line(1)
  end;
```

```
for wi:=0 step 1 until k do
  for ko:=wi step 1 until k do
    M[wi,ko]:=M[wi,ko]+ X[wi] × X[ko];
if j<n then begin
  j:= j+1;
  go to wersja
end;
for wi:=0 step 1 until k do
  for ko:=wi step 1 until k do a[wi+1,ko+1]:=M[wi,ko];
inver1(k+1,a, fail);
for wi:=0 step 1 until k do
  for ko:=0 step 1 until wi do
  .begin
    MO[wi,ko]:=a[wi+2,ko+1];
    MO[ko,wi]:=MO[wi,ko]
  end;
d:=0;
e:=2↑k;
for wie:=1 step 1 until e do
  begin
    for ko:=0 step 1 until k do
      begin
        X[ko]:=MO[0,ko];
        for wi:=1 step 1 until k do
          X[ko]:=X[ko]+W[wie,wi] ×MO[wi,ko]
        end;
      dr:=X[0];
      for wi:=1 step 1 until k do dr:=dr+X[wi]×W[wie,wi];
      if dr>d then d:=dr
    end maksimum;
  X[0]:=1;
if n=N then begin
  if c=2×entier(c/2+.1) then dQ[1]:=d else dI[1]:=d
  end;
```

```
if d>(k+1)/N then begin
    e:=entier(N-(k+1)/d);
    if e=0 then n:=n+1 else n:=n+e;
    j:=j+1;
    go to wersja
end;
if c=2*entier(c/2+.1) then go to q;
L[1]:=j;
DL[1]:=d;
c:=c+1;
go to liczba;
q: Q[1]:=j;
DQ[1]:=d;
c:=c+1;
l:=l+1;
if ¬(l>lp) then go to liczba;
line(20);
format('WYMIAR=12 N=12 DOKLADNOSC=1.234');
print(k,N,(k+1)/N,'???');
print('PLAN LOSOWY PLAN QUASIOPTYMALNY');
print('???Lp dN n d');
print('Lp dN n d??');
format('12 11.234 123 1.234');
for l:=1 step 1 until lp do
    print(l,dl[l],L[l],DL[l],l,dQ[l],Q[l],DQ[l],'?');
sl:=sq:=val:=vaq:=slN:=sqN:=vlN:=vqN:=0;
for l:=1 step 1 until lp do
    begin
        sl:=sl+L[l];
        val:=val+L[l]*L[l];
        sq:=sq+Q[l];
        vaq:=vaq+Q[l]*Q[l];
        slN:=slN+dl[l];
        vlN:=vlN+dl[l]*dl[l];
        sqN:=sqN+dQ[l];
        vqN:=vqN+dQ[l]*dQ[l]
    end;
```

```
sl:=sl/lp;
sq:=sq/lp;
slN:=slN/lp;
sqN:=sqN/lp;
val:=val/lp-sl×sl;
vaq:=vaq/lp-sq×sq;
vlN:=vlN/lp-slN×slN;
vqN:=vqN/lp-sqN×sqN;
format('111.234123.4');
print('??Srednia',slN,sl,sqN,sq);
print('?Wariancja',vlN,val,vqN,vaq);
print('?Odchust',sqrt(vlN),sqrt(val),sqrt(vqN),sqrt(vaq));
i:=i+1;
if i>lsdk then w=w+1 else go to efekt;
if w<Lw ∨ w=Lw then go to wymiar;
fail:end
end;
```

WYMIAR = 2 N = 8 DOKLADNOSC = .375

PLAN LOSOWY				PLAN QUASIOPTYMALNY			
Lp	dN	n	d	Lp	dN	n	d
1	1.461	28	.371	1	.765	17	.345
2	1.278	27	.374	2	.678	14	.368
3	1.245	26	.323	3	1.377	16	.335
4	1.237	24	.356	4	1.365	15	.336
5	4.665	24	.362	5	.749	15	.363
6	1.445	24	.370	6	.635	13	.343
7	3.361	25	.355	7	135.432	15	.358
8	2.057	24	.342	8	289.785	16	.373
9	5.777	25	.359	9	1.364	20	.291
10	2.090	27	.347	10	.673	14	.368
11	1.406	19	.368	11	1.774	19	.346
12	1.635	26	.348	12	1.705	20	.311
13	3.168	31	.369	13	.740	13	.368
14	1.600	28	.363	14	.765	13	.341
15	4.213	37	.374	15	.660	11	.375
16	1.451	26	.365	16	1.244	18	.294
17	1.982	27	.370	17	.772	19	.339
18	.738	17	.374	18	.719	17	.316
19	3.450	21	.366	19	.684	16	.321
20	3.969	24	.374	20	1.596	14	.336
21	2.284	25	.370	21	.597	14	.359
22	1.689	24	.374	22	.737	13	.299
23	1.451	27	.357	23	.670	13	.358
24	1.297	20	.361	24	.662	15	.356
25	1.318	23	.336	25	2.929	17	.328
26	1.507	20	.364	26	1.792	14	.365
27	3.673	26	.353	27	.620	12	.372
28	2.798	26	.369	28	1.888	15	.340
29	1.755	37	.334	29	.760	21	.360
30	3.354	29	.370	30	.640	14	.343
Srednia	2.312	25.6			15.159	15.4	
Wariancja	1.479	18.2			3182.099	6.2	
Odch st	1.216	4.3			56.410	2.5	

WYMIAR = 2 N = 12 DOKLADNOSC = .250

PLAN LOSOWY

PLAN QUASIOPTYMALNY

Lp	dN	n	d	Lp	dN	n	d
1	1.986	33	.238	1	.407	17	.220
2	.838	35	.243	2	.356	19	.248
3	2.990	40	.240	3	.563	20	.238
4	2.125	29	.243	4	.564	18	.244
5	.703	31	.232	5	.378	20	.244
6	.786	30	.239	6	1.016	26	.221
7	1.429	32	.237	7	.398	20	.207
8	1.156	44	.239	8	.386	20	.232
9	1.079	37	.245	9	.352	17	.248
10	.873	39	.244	10	.381	16	.223
11	.971	38	.245	11	.872	22	.226
12	1.556	40	.231	12	.378	18	.243
13	1.278	39	.248	13	.947	20	.244
14	1.040	39	.233	14	.401	20	.245
15	.993	37	.236	15	.358	17	.249
16	1.270	36	.245	16	.356	20	.242
17	2.582	39	.247	17	.398	21	.241
18	.910	40	.249	18	.376	19	.208
19	.985	30	.233	19	.533	20	.249
20	.930	31	.245	20	.391	17	.247
21	.710	37	.249	21	.579	20	.249
22	.868	31	.237	22	.532	20	.237
23	.868	37	.248	23	.292	17	.213
24	.750	38	.246	24	.398	20	.249
25	1.102	31	.246	25	.548	18	.250
26	1.270	46	.210	26	.398	22	.237
27	.969	34	.248	27	.386	18	.209
28	.863	40	.241	28	.547	19	.247
29	.873	36	.232	29	1.074	24	.230
30	1.891	36	.232	30	.399	20	.249
Srednia	1.222	36.2			.499	19.5	
Wariancja	.306	17.5			.042	4.4	
Odch st	.553	4.2			.204	2.1	

WYMIAR = 2 N = 16 DOKŁADNOŚĆ = .188

PLAN LOSOWY				PLAN QUASIOPTYMALNY			
Lp	dN	n	d	Lp	dN	n	d
1	.969	41	.183	1	.261	26	.186
2	.880	44	.187	2	.325	22	.183
3	1.024	43	.187	3	.450	34	.183
4	.621	53	.186	4	.743	27	.154
5	.570	47	.175	5	.443	26	.164
6	.928	41	.187	6	.252	22	.181
7	.482	43	.187	7	.639	35	.158
8	.785	41	.187	8	.351	30	.166
9	.673	43	.182	9	.262	23	.164
10	.512	49	.178	10	.328	23	.168
11	.721	52	.185	11	.433	26	.186
12	1.343	48	.187	12	.326	24	.186
13	.671	40	.166	13	.321	26	.178
14	.430	42	.170	14	.449	29	.169
15	.568	57	.176	15	.717	33	.184
16	.506	45	.183	16	.341	28	.174
17	1.046	59	.187	17	.723	32	.184
18	.979	64	.187	18	.266	28	.175
19	.855	48	.182	19	.428	24	.183
20	.705	55	.180	20	.354	27	.158
21	.881	51	.184	21	.323	29	.179
22	.638	46	.186	22	.252	29	.166
23	.836	44	.183	23	.357	32	.168
24	1.112	52	.181	24	.266	26	.163
25	.710	45	.187	25	.338	27	.184
26	.732	53	.187	26	.671	33	.184
27	.585	41	.187	27	.346	27	.185
28	.727	56	.186	28	.349	29	.176
29	.786	48	.179	29	.322	23	.187
30	.457	40	.180	30	.280	26	.163
Srednia	.758	47.7			.397	27.5	
Wariancja	.045	38.3			.022	12.5	
Odch st	.212	6.2			.147	3.5	

WYMIAR = 2 N = 20 DOKLADNOSC = .150

PLAN LOSOWY

PLAN QUASIOPTYMALNY

Lp	dN	n	d	Lp	dN	n	d
1	.430	60	.150	1	.240	30	.145
2	.503	65	.146	2	.247	39	.134
3	.406	58	.146	3	.194	25	.146
4	.633	63	.140	4	.253	39	.143
5	.424	60	.133	5	.425	34	.138
6	.418	68	.142	6	.230	31	.147
7	.443	56	.145	7	.282	34	.138
8	.486	56	.142	8	.287	34	.141
9	.541	56	.141	9	.202	25	.148
10	.461	47	.148	10	.195	29	.149
11	.318	49	.146	11	.252	29	.140
12	.581	57	.143	12	.302	39	.148
13	.415	62	.148	13	.230	28	.149
14	.599	56	.148	14	.201	27	.144
15	.576	65	.147	15	.197	30	.149
16	.534	57	.145	16	.242	32	.150
17	.549	56	.147	17	.308	30	.143
18	.572	61	.150	18	.542	35	.146
19	.550	59	.145	19	.304	30	.148
20	.526	55	.148	20	.198	26	.141
21	.421	59	.147	21	.305	32	.145
22	.915	64	.146	22	.204	31	.137
23	.711	60	.139	23	.256	37	.142
24	.417	53	.147	24	.434	35	.139
25	.346	65	.149	25	.241	27	.138
26	.593	63	.148	26	.306	30	.149
27	.491	58	.148	27	.231	36	.139
28	.634	59	.149	28	.342	31	.144
29	.542	52	.143	29	.201	31	.150
30	.408	58	.144	30	.288	29	.150
Srednia	.515	58.6			.271	31.5	
Wariancja	.014	21.9			.006	15.4	
Odch st	.117	4.7			.079	3.9	

WYMIAR = 2 N = 24 DOKLADNOSC = .125

PLAN LOSOWY				PLAN QUASIOPTYMALNY			
Lp	dN	n	d	Lp	dN	n	d
1	.360	67	.124	1	.187	37	.124
2	.442	66	.124	2	.216	34	.119
3	.389	68	.123	3	.187	37	.120
4	.543	64	.125	4	.193	36	.124
5	.402	70	.124	5	.190	36	.125
6	.543	70	.123	6	.259	36	.122
7	.572	69	.125	7	.265	38	.122
8	.405	71	.123	8	.348	45	.117
9	.387	69	.117	9	.168	38	.125
10	.377	67	.125	10	.253	37	.124
11	.353	72	.123	11	.289	44	.118
12	.339	64	.125	12	.163	29	.124
13	.509	67	.120	13	.310	44	.119
14	.429	76	.123	14	.199	40	.125
15	.311	60	.119	15	.180	37	.122
16	.433	69	.124	16	.229	41	.112
17	.323	70	.121	17	.161	33	.116
18	.364	68	.119	18	.166	35	.125
19	.371	61	.124	19	.190	37	.124
20	.596	66	.121	20	.218	35	.116
21	.787	67	.118	21	.224	36	.114
22	.426	72	.122	22	.218	44	.103
23	.361	59	.117	23	.187	45	.114
24	.386	65	.125	24	.233	41	.115
25	.475	70	.124	25	.162	33	.115
26	.478	61	.120	26	.195	32	.122
27	.394	67	.125	27	.226	41	.124
28	.377	68	.125	28	.178	31	.115
29	.334	74	.124	29	.188	36	.124
30	.409	62	.125	30	.220	34	.124
Srednia					.213	37.4	
Wariancja					.002	17.2	
Odch st					.045	4.2	

WYMIAR = 2 N = 28 DOKLADNOSC = .107

PLAN LOSOWY

PLAN QUASIOPTYMALNY

Lp	dN	n	d	Lp	dN	n	d
1	.343	74	.107	1	.153	40	.098
2	.391	74	.102	2	.156	40	.102
3	.318	79	.107	3	.155	36	.105
4	.565	83	.106	4	.214	55	.106
5	.359	70	.107	5	.176	41	.102
6	.269	84	.104	6	.167	38	.101
7	.393	77	.107	7	.178	43	.107
8	.313	79	.106	8	.139	38	.102
9	.294	74	.106	9	.212	46	.106
10	.352	77	.106	10	.159	42	.107
11	.415	71	.105	11	.153	38	.106
12	.329	87	.104	12	.189	40	.104
13	.452	70	.105	13	.152	42	.106
14	.385	70	.105	14	.136	37	.103
15	.457	84	.107	15	.157	41	.106
16	.306	81	.104	16	.153	42	.106
17	.465	80	.105	17	.207	41	.101
18	.302	84	.104	18	.135	36	.104
19	.414	90	.107	19	.158	43	.104
20	.273	71	.106	20	.132	39	.106
21	.358	73	.105	21	.239	52	.098
22	.467	89	.106	22	.146	45	.101
23	.280	82	.106	23	.154	40	.107
24	.320	90	.103	24	.156	44	.107
25	.476	68	.106	25	.147	39	.099
26	.351	80	.107	26	.199	44	.105
27	.244	78	.105	27	.150	39	.098
28	.498	98	.105	28	.155	44	.105
29	.312	81	.105	29	.142	39	.102
30	.286	69	.106	30	.149	43	.096
Srednia				Srednia			
.366				.164			
Wariancja				Wariancja			
.006				.001			
Odch st				Odch st			
.078				.026			
78.9				41.6			
53.0				16.8			
7.3				4.1			

WYMIAR = 2 N = 32 DOKLADNOSC = .094

PLAN LOSOWY				PLAN QUASIOPTYMALNY			
Lp	dN	n	d	Lp	dN	n	d
1	.282	104	.092	1	.186	49	.094
2	.385	95	.093	2	.179	58	.090
3	.245	86	.090	3	.130	49	.087
4	.275	82	.091	4	.143	45	.090
5	.350	96	.092	5	.146	45	.088
6	.231	82	.093	6	.119	45	.089
7	.308	81	.088	7	.125	47	.094
8	.299	90	.092	8	.256	52	.083
9	.308	84	.094	9	.136	49	.094
10	.290	81	.093	10	.144	43	.091
11	.269	81	.094	11	.131	45	.090
12	.369	107	.093	12	.130	46	.094
13	.271	91	.093	13	.132	47	.093
14	.414	86	.093	14	.183	47	.086
15	.330	104	.094	15	.134	49	.093
16	.381	83	.093	16	.221	50	.091
17	.327	95	.093	17	.158	51	.091
18	.336	96	.093	18	.147	49	.093
19	.300	89	.094	19	.132	43	.089
20	.299	87	.094	20	.130	48	.088
21	.416	109	.092	21	.249	53	.091
22	.367	87	.092	22	.130	48	.093
23	.245	90	.093	23	.167	53	.087
24	.269	88	.087	24	.169	51	.090
25	.240	76	.094	25	.132	45	.091
26	.326	88	.090	26	.120	41	.094
27	.243	76	.092	27	.168	58	.094
28	.281	82	.092	28	.146	43	.092
29	.339	111	.092	29	.131	51	.091
30	.256	81	.093	30	.133	52	.087
Srednia	.308	89.6			.154	48.4	
Wariancja	.003	88.4			.001	16.3	
Odch st	.051	9.4			.035	4.0	

WYMIAR = 3 N = 8 DOKLADNOSC = .500

PLAN LOSOWY

PLAN QUASIOPTYMALNY

Lp	dN	n	d	Lp	dN	n	d
1	3.012	29	.485	1	.921	21	.460
2	3.081	23	.468	2	1.929	16	.474
3	1.911	27	.494	3	1.045	16	.435
4	4.934	28	.485	4	5.966	22	.475
5	4.080	37	.473	5	2.186	21	.403
6	7.212	31	.484	6	.949	21	.465
7	6.972	28	.499	7	1.045	18	.413
8	2.109	28	.451	8	2.722	17	.464
9	5.215	29	.461	9	2.622	15	.458
10	3.422	33	.475	10	1.655	13	.494
11	1.355	28	.483	11	1.716	17	.482
12	3.533	33	.390	12	1.636	17	.431
13	6.176	44	.447	13	2.151	17	.453
14	4.830	27	.459	14	.960	15	.498
15	4.724	38	.483	15	5.182	17	.471
16	3.190	29	.488	16	.932	14	.466
17	4.536	26	.448	17	2.263	21	.469
18	3.221	31	.497	18	3.023	21	.498
19	2.629	27	.483	19	1.269	16	.462
20	6.645	33	.497	20	1.828	14	.466
21	6.458	31	.495	21	1.566	13	.499
22	2.416	27	.477	22	1.675	21	.476
23	7.490	28	.487	23	.924	14	.468
24	4.748	30	.476	24	.972	17	.459
25	7.795	31	.478	25	1.286	18	.414
26	4.929	34	.468	26	1.604	18	.476
27	13.749	29	.425	27	5.029	22	.474
28	2.801	33	.460	28	2.128	21	.496
29	2.768	36	.475	29	2.545	22	.397
30	15.605	30	.452	30	218.625	19	.487

Srednia 5.052 30.6
 Wariancja 9.709 17.2
 Odch st 3.116 4.1

9.278 17.8
 1512.832 8.2
 38.895 2.9

WYMIAR = 3 N = 12 DOKŁADNOŚĆ = .333

PLAN LOSOWY

PLAN QUASIOPTYMALNY

Lp	dN	n	d	Lp	dN	n	d
1	1.983	41	.330	1	.571	19	.322
2	1.767	41	.320	2	1.278	31	.327
3	2.768	31	.328	3	.595	25	.318
4	2.150	43	.327	4	.828	24	.323
5	2.163	48	.333	5	1.324	24	.311
6	1.738	46	.320	6	.535	19	.324
7	1.270	40	.322	7	.546	21	.323
8	1.692	45	.321	8	1.136	28	.261
9	1.561	39	.328	9	1.392	25	.308
10	1.827	36	.308	10	.798	27	.267
11	1.520	46	.311	11	1.505	31	.273
12	2.042	46	.303	12	.886	20	.325
13	2.049	42	.310	13	2.186	27	.319
14	2.402	46	.333	14	1.122	24	.327
15	5.293	37	.328	15	1.100	26	.317
16	2.238	44	.333	16	1.364	32	.326
17	1.094	42	.309	17	.686	25	.308
18	1.770	49	.322	18	1.657	24	.292
19	1.435	36	.331	19	.621	20	.284
20	2.230	51	.324	20	.742	25	.318
21	2.067	45	.331	21	1.946	26	.283
22	1.362	40	.332	22	1.528	24	.314
23	2.840	44	.330	23	1.680	29	.332
24	2.368	41	.333	24	1.041	30	.235
25	3.071	50	.333	25	.677	24	.324
26	1.877	46	.312	26	2.147	22	.282
27	2.293	41	.331	27	1.266	25	.278
28	1.733	36	.329	28	1.529	28	.303
29	3.604	46	.306	29	.968	25	.333
30	1.454	44	.311	30	.815	23	.296
Srednia	2.122	42.7		1.149	25.1		
Wariancja	.638	20.5		.218	11.6		
Odch st	.799	4.5		.467	3.4		

WYMIAR = 3 N = 16 DOKŁADNOŚĆ = .250

PLAN LOSOWY				PLAN QUASIOPTYMALNY			
Lp	dN	n	d	Lp	dN	n	d
1	1.654	58	.248	1	.354	27	.224
2	1.434	46	.249	2	.548	38	.215
3	.800	47	.249	3	1.059	32	.233
4	1.160	43	.244	4	.937	32	.231
5	1.572	59	.250	5	.459	23	.246
6	.804	51	.248	6	.383	25	.227
7	.813	52	.230	7	.453	31	.207
8	1.076	50	.249	8	.815	30	.224
9	2.060	63	.233	9	.438	31	.250
10	.680	51	.232	10	.632	35	.248
11	2.065	55	.246	11	.879	30	.249
12	2.268	57	.249	12	.737	26	.248
13	1.192	61	.238	13	.383	24	.227
14	2.250	65	.246	14	.838	32	.243
15	1.187	54	.239	15	.404	29	.242
16	.871	51	.230	16	.484	25	.245
17	1.634	59	.247	17	.812	32	.248
18	1.574	53	.247	18	.662	30	.241
19	1.505	48	.245	19	.731	26	.241
20	.846	53	.243	20	.885	40	.245
21	1.848	57	.247	21	.436	28	.244
22	1.545	59	.241	22	.443	33	.222
23	1.796	55	.249	23	.445	26	.234
24	1.264	49	.249	24	.623	25	.242
25	1.240	66	.232	25	.735	26	.248
26	1.209	50	.250	26	.462	38	.222
27	2.277	71	.240	27	.417	24	.224
28	2.242	55	.242	28	.930	32	.214
29	.932	53	.249	29	.654	38	.246
30	1.008	46	.246	30	.593	32	.246
Srednia	1.427	54.6			.621	30.0	
Wariancja	.242	40.4			.040	20.9	
Odch st	.491	6.4			.199	4.6	

WYMIAR = 3 N = 20 DOKŁADNOSC = .200

PLAN LOSOWY				PLAN QUASIOPTYMALNY			
Lp	dN	n	d	Lp	dN	n	d
1	.742	76	.197	1	.375	35	.181
2	1.083	67	.194	2	.306	34	.189
3	1.082	71	.188	3	.407	37	.198
4	.820	69	.198	4	.561	41	.196
5	.781	63	.199	5	.343	33	.194
6	1.170	57	.197	6	.377	45	.192
7	1.060	60	.186	7	.372	29	.192
8	1.189	65	.196	8	.464	34	.182
9	1.124	80	.195	9	.354	31	.195
10	1.196	69	.192	10	.332	35	.195
11	.894	59	.196	11	.442	40	.193
12	.824	64	.197	12	.286	30	.194
13	1.503	62	.199	13	.415	35	.197
14	1.027	63	.199	14	.379	33	.194
15	.713	65	.197	15	.413	34	.179
16	.614	72	.193	16	.479	33	.199
17	.959	69	.191	17	.343	33	.191
18	1.446	78	.188	18	.330	32	.180
19	.744	71	.195	19	.332	36	.194
20	.748	56	.199	20	.376	29	.199
21	.738	65	.191	21	.281	33	.181
22	.580	55	.190	22	.291	27	.197
23	.818	68	.199	23	.385	41	.191
24	1.442	60	.198	24	.359	33	.195
25	.878	61	.200	25	.576	33	.179
26	.964	57	.197	26	.562	37	.194
27	.923	70	.200	27	.607	36	.198
28	.825	63	.199	28	.434	35	.180
29	.710	57	.199	29	.450	42	.199
30	1.037	65	.186	30	.811	35	.191
Srednia				Srednia			
.955				.415			
Wariancja				Wariancja			
.056				.013			
Odch st				Odch st			
.237				.113			
65.2				34.7			
40.6				15.7			
6.4				4.0			

WYMIAR = 3 N = 24 DOKŁADNOŚĆ = .167

PLAN LOSOWY

PLAN QUASIOPTYMALNY

Lp	dN	n	d	Lp	dN	n	d
1	.579	76	.163	1	.292	37	.164
2	.507	69	.160	2	.282	40	.164
3	.455	73	.166	3	.371	43	.151
4	.564	76	.167	4	.276	35	.156
5	.710	92	.165	5	.267	39	.153
6	1.017	89	.165	6	.458	49	.163
7	.510	73	.165	7	.229	39	.154
8	.695	81	.166	8	.283	39	.163
9	.579	76	.165	9	.271	41	.163
10	.849	80	.156	10	.351	47	.166
11	.664	93	.159	11	.299	41	.149
12	.746	73	.160	12	.307	46	.163
13	.830	73	.165	13	.305	35	.159
14	.542	76	.163	14	.357	44	.166
15	.944	83	.166	15	.377	43	.142
16	.560	78	.165	16	.537	49	.161
17	1.061	100	.166	17	.297	34	.165
18	.558	77	.163	18	.363	37	.164
19	.622	89	.167	19	.352	44	.164
20	.837	95	.161	20	.370	43	.144
21	.929	94	.166	21	.302	41	.162
22	.750	83	.164	22	.258	42	.162
23	.553	72	.166	23	.305	48	.161
24	.615	81	.165	24	.384	42	.163
25	.835	70	.166	25	.289	43	.162
26	.638	87	.166	26	.328	43	.149
27	.587	73	.164	27	.305	40	.163
28	.588	76	.159	28	.509	51	.158
29	.768	77	.166	29	.310	35	.153
30	.563	74	.165	30	.349	41	.164
Srednia	.689	80.3		.333	41.7		
Wariancja	.025	67.7		.005	18.9		
Odch st	.159	8.2		.069	4.3		

WYMIAR = 3 N = 28 DOKLADNOSC = .143

PLAN LOSOWY

PLAN QUASIOPTYMALNY

Lp	dN	n	d	Lp	dN	n	d
1	.650	83	.142	1	.219	42	.136
2	.456	83	.137	2	.235	57	.142
3	.913	87	.140	3	.234	50	.141
4	.491	79	.142	4	.288	45	.140
5	.889	84	.142	5	.267	45	.129
6	.644	92	.140	6	.319	55	.142
7	.508	83	.141	7	.312	51	.138
8	.556	87	.142	8	.256	48	.131
9	.569	99	.142	9	.285	48	.140
10	.630	107	.143	10	.239	49	.138
11	.496	95	.140	11	.277	47	.134
12	.593	99	.131	12	.269	43	.136
13	.660	91	.142	13	.302	49	.123
14	.731	89	.140	14	.431	48	.143
15	.474	91	.141	15	.258	43	.140
16	.627	95	.134	16	.208	47	.131
17	.580	78	.142	17	.330	45	.142
18	.565	91	.142	18	.246	53	.129
19	.632	92	.134	19	.216	43	.139
20	.455	81	.136	20	.317	46	.143
21	.469	80	.141	21	.215	37	.141
22	.805	96	.139	22	.233	47	.131
23	.486	78	.138	23	.238	45	.138
24	.711	103	.140	24	.220	43	.133
25	.617	85	.142	25	.203	44	.136
26	.538	75	.143	26	.281	46	.133
27	.615	96	.140	27	.272	47	.136
28	.574	93	.142	28	.234	44	.143
29	.537	96	.142	29	.243	45	.139
30	.508	86	.143	30	.231	46	.140
Srednia	.599	89.1			.263	46.6	
Wariancja	.013	61.1			.002	15.4	
Odch st	.16	7.8			.047	3.9	

WYMIAR = 3 N = 32 DOKŁADNOŚĆ = .125

PLAN LOSOWY				PLAN QUASIOPTYMALNY			
Lp	dN	n	d	Lp	dN	n	d
1	.629	113	.120	1	.238	52	.123
2	.383	95	.124	2	.298	57	.116
3	.512	104	.125	3	.183	46	.122
4	.365	84	.125	4	.252	64	.123
5	.471	111	.121	5	.222	45	.125
6	.427	92	.122	6	.212	48	.121
7	.446	107	.124	7	.233	55	.120
8	.464	103	.125	8	.197	51	.115
9	.414	98	.125	9	.204	50	.117
10	.491	104	.124	10	.174	49	.125
11	.485	100	.122	11	.330	55	.120
12	.571	108	.117	12	.213	48	.124
13	.452	98	.124	13	.320	62	.124
14	.493	10	.125	14	.327	54	.123
15	.522	119	.119	15	.187	47	.122
16	.360	92	.124	16	.230	52	.122
17	.372	81	.124	17	.218	46	.124
18	.446	106	.120	18	.224	55	.121
19	.424	88	.124	19	.219	60	.122
20	.629	93	.125	20	.190	50	.122
21	.625	96	.124	21	.276	51	.123
22	.524	101	.124	22	.211	49	.123
23	.546	105	.116	23	.162	45	.116
24	.443	104	.125	24	.202	46	.114
25	.375	86	.122	25	.246	69	.117
26	.371	104	.124	26	.225	50	.125
27	.494	89	.123	27	.186	47	.124
28	.465	95	.125	28	.250	54	.114
29	.559	107	.123	29	.324	58	.124
30	.505	87	.124	30	.192	58	.118
Srednia	.475	99.0		.231	52.4		
Wariancja	.006	80.5		.002	34.8		
Odch st	.077	9.0		.047	5.9		

WYMIAR = 4 N = 8 DOKŁADNOSC = .625

PLAN LOSOWY

PLAN QUASIOPTYMALNY

Lp	dN	n	d	Lp	dN	n	d
1	4.910	35	.606	1	4.847	21	.620
2	41.384	33	.590	2	4.733	18	.535
3	7.486	31	.573	3	2.647	22	.623
4	7.453	31	.542	4	4.288	18	.523
5	8.879	31	.611	5	2.479	20	.587
6	6.181	34	.602	6	4.697	32	.596
7	7.707	35	.577	7	6.755	16	.576
8	5.261	36	.572	8	1.393	19	.552
9	6.823	36	.610	9	3.089	16	.550
10	45.229	35	.603	10	6.361	27	.557
1	10.373	28	.605	11	1.287	21	.495
12	8.770	30	.568	12	13.443	17	.609
13	7.205	39	.571	13	3.041	14	.534
14	17.087	36	.595	14	7.785	20	.605
15	8.030	32	.612	15	9.173	17	.549
16	18.908	32	.602	16	4.402	22	.532
17	3.522	33	.602	17	2.954	15	.597
18	146.322	29	.619	18	4.774	19	.600
19	7.639	44	.572	19	3.328	25	.582
20	5.318	28	.623	20	3.873	23	.621
21	3.541	40	.599	21	8.688	21	.583
22	22.000	38	.538	22	12.737	19	.580
23	3.924	29	.613	23	3.224	21	.624
24	2.554	29	.620	24	3.257	17	.566
25	6.512	29	.621	25	5.627	17	.603
26	65.695	39	.576	26	3.057	18	.600
27	3.839	33	.590	27	6.135	20	.595
28	9.089	31	.576	28	117.839	18	.525
29	41.244	32	.624	29	5.326	18	.605
30	5.424	31	.623	30	6.186	29	.614

Srednia 17.944 33.3
 Wariancja 788.315 15.0
 Odch st 28.077 3.9

8.914 20.0
 417.216 15.7
 20.426 4.0

WYMIAR = 4 N = 12 DOKŁADNOŚĆ = .417

PLAN LOSOWY

PLAN QUASIOPTYMALNY

Lp	dN	n	d	Lp	dN	n	d
1	5.443	59	.354	1	1.101	19	.403
2	5.684	49	.412	2	1.763	21	.402
3	2.250	46	.405	3	3.004	33	.416
4	5.851	45	.409	4	1.492	33	.395
5	1.519	50	.414	5	1.532	25	.410
6	6.684	43	.415	6	1.364	34	.373
7	3.628	42	.410	7	2.265	38	.373
8	2.484	35	.40	8	1.460	27	.390
9	6.389	39	.414	9	2.156	31	.405
10	7.387	48	.413	10	1.775	32	.326
11	2.103	46	.413	11	1.978	28	.369
12	5.584	54	.412	12	.961	23	.415
13	1.607	37	.416	13	1.796	30	.407
14	3.563	63	.380	14	1.786	26	.353
15	3.964	49	.390	15	2.234	28	.350
16	4.793	47	.411	16	2.272	25	.380
17	3.892	55	.414	17	2.284	26	.408
18	10.563	60	.382	18	2.307	26	.416
19	3.096	41	.416	19	1.899	25	.412
20	3.364	46	.410	20	.970	30	.377
21	3.012	46	.400	21	2.417	27	.401
22	3.618	37	.400	22	1.075	22	.388
23	2.514	46	.398	23	1.610	24	.398
24	3.161	48	.414	24	3.323	27	.411
25	9.800	50	.415	25	.915	23	.386
26	3.408	47	.408	26	.993	32	.390
27	6.311	52	.417	27	2.314	22	.416
28	2.819	46	.394	28	3.011	22	.377
29	3.811	43	.410	29	1.600	31	.360
30	2.775	41	.416	30	2.587	26	.312

Srednia 4.369
 Wariancja 4.785
 Odch st 2.187

1.875 27.2
 .400 19.2
 .633 4.4

WYMIAR = 4 N = 16 DOKŁADNOŚĆ = .313

PLAN LOSOWY				PLAN QUASIOPTYMALNY			
Lp	dN	n	d	Lp	dN	n	d
1	1.235	53	.309	1	2.046	35	.251
2	2.580	69	.312	2	1.267	31	.267
3	1.364	62	.290	3	1.487	40	.299
4	3.364	62	.311	4	.651	27	.297
5	1.915	65	.301	5	1.298	34	.280
6	2.189	68	.312	6	.748	28	.298
7	6.929	56	.307	7	.555	34	.296
8	2.279	58	.308	8	1.168	34	.289
9	2.057	61	.300	9	.966	29	.294
10	2.050	61	.312	10	.709	28	.307
11	1.932	58	.294	11	1.489	31	.305
12	2.709	57	.312	12	1.143	30	.289
13	1.867	69	.303	13	.977	34	.273
14	1.797	58	.307	14	.969	27	.300
15	2.399	63	.304	15	.780	32	.284
16	2.458	65	.304	16	.707	27	.309
17	3.062	64	.308	17	1.011	33	.312
18	2.397	56	.307	18	1.002	32	.280
19	2.253	63	.312	19	1.004	36	.287
20	1.801	64	.307	20	1.895	34	.250
21	2.791	67	.312	21	.823	37	.302
22	1.047	47	.310	22	1.360	35	.311
23	1.633	68	.307	23	.941	37	.306
24	1.695	58	.309	24	.820	36	.311
25	1.807	58	.306	25	.909	37	.291
26	2.942	61	.312	26	.562	26	.299
27	1.930	60	.309	27	1.345	32	.233
28	1.310	58	.310	28	1.503	33	.310
29	3.213	61	.309	29	.974	35	.269
30	1.691	57	.299	30	.860	33	.301
Srednia	2.290	60.9		1.066		32.6	
Wariancja	1.072	24.0		.128		12.3	
Odch st	1.035	4.9		.358		3.5	

WYMIAR = 4 N = 20 DOKLADNOSC = .250

PLAN LOSOWY

PLAN QUASIOPTYMALNY

Lp	dN	n	d	Lp	dN	n	d
1	1.223	74	.234	1	.762	45	.248
2	1.518	69	.250	2	.564	36	.220
3	1.345	72	.246	3	.604	37	.246
4	1.124	80	.245	4	.542	45	.249
5	1.790	77	.249	5	.579	45	.237
6	1.065	72	.250	6	.474	35	.236
7	.851	68	.249	7	.972	48	.248
8	2.205	64	.248	8	1.153	41	.220
9	1.051	69	.243	9	.913	41	.244
10	1.498	75	.247	10	.872	48	.232
11	1.542	61	.245	11	.583	40	.249
12	1.107	63	.240	12	.528	36	.229
13	1.066	66	.250	13	.460	34	.236
14	1.685	63	.248	14	.536	36	.233
15	1.293	82	.242	15	.608	39	.247
16	1.146	77	.250	16	.618	33	.239
17	1.280	66	.245	17	.568	37	.249
18	1.416	61	.242	18	.699	39	.210
19	.988	68	.247	19	.420	35	.246
20	1.924	64	.244	20	.662	38	.216
21	1.141	76	.241	21	.510	35	.247
22	1.242	71	.242	22	.661	34	.250
23	2.285	60	.248	23	.774	47	.244
24	1.294	74	.245	24	.587	33	.238
25	1.797	68	.244	25	.898	34	.201
26	.970	64	.247	26	.623	36	.243
27	1.312	74	.242	27	.866	43	.219
28	1.119	79	.249	28	.633	39	.222
29	1.035	69	.250	29	.763	43	.241
30	2.265	66	.245	30	.481	36	.245
Srednia	1.386	69.7			.664	38.9	
Wariancja	.150	35.7			.029	20.8	
Odch st	.387	6.0			.169	4.6	

WYMIAR = 4 N = 24 DOKŁADNOŚĆ = .208

PLAN LOSOWY

PLAN QUASIOPTYMALNY

Lp	dN	n	d	Lp	dN	n	d
1	1.600	87	.206	1	.368	49	.191
2	.853	86	.206	2	.485	51	.208
3	2.217	99	.206	3	.502	45	.208
4	.889	79	.206	4	.436	39	.194
5	.926	77	.208	5	.539	46	.207
6	1.711	100	.207	6	.350	43	.201
7	1.054	97	.206	7	.506	51	.208
8	1.154	77	.205	8	.453	36	.208
9	.635	88	.206	9	1.013	43	.197
10	1.221	73	.207	10	.560	50	.207
11	.914	71	.207	11	.362	40	.196
12	1.363	97	.202	12	.405	50	.203
13	1.510	93	.204	13	.509	42	.207
14	1.181	102	.203	14	.913	53	.190
15	.990	76	.205	15	.548	49	.203
16	1.284	87	.208	16	.407	42	.207
17	1.091	92	.208	17	.651	40	.186
18	.867	82	.208	18	.394	38	.208
19	.890	82	.203	19	.419	45	.203
20	1.315	84	.206	20	.526	45	.206
21	.768	78	.208	21	.648	48	.201
22	.739	82	.199	22	.393	36	.205
23	.904	75	.208	23	.455	42	.205
24	.997	99	.208	24	.413	44	.197
25	.889	85	.200	25	.460	44	.182
26	.673	84	.208	26	.526	45	.205
27	.742	90	.203	27	.552	50	.176
28	.816	82	.206	28	.367	41	.189
29	1.093	89	.206	29	.773	41	.208
30	.758	80	.208	30	.421	36	.205
Srednia	1.068	85.8			.512	44.1	
Wariancja	.117	71.5			.024	22.6	
Odch st	.343	8.5			.154	4.8	

WYMIAR = 4 N = 28 DOKŁADNOŚĆ = .179

PLAN LOSOWY				PLAN QUASIOPTYMALNY			
Lp	dN	n	d	Lp	dN	n	d
1	1.063	87	.175	1	.457	56	.175
2	.757	94	.177	2	.338	45	.165
3	.798	101	.177	3	.331	49	.177
4	.845	107	.178	4	.355	52	.176
5	.574	90	.179	5	.309	54	.177
6	.867	88	.176	6	.397	53	.171
7	1.390	96	.178	7	.323	49	.170
8	1.062	110	.179	8	.526	49	.175
9	.734	98	.177	9	.370	51	.177
10	.998	99	.177	10	.345	47	.168
11	1.100	96	.178	11	.478	45	.170
12	.914	89	.177	12	.427	51	.178
13	.859	93	.176	13	.343	43	.172
14	.650	97	.178	14	.318	47	.173
15	1.054	95	.174	15	.483	49	.178
16	.666	90	.178	16	.362	50	.173
17	1.034	97	.177	17	.351	51	.178
18	.660	103	.178	18	.496	45	.171
19	.685	92	.178	19	.267	45	.178
20	.868	89	.174	20	.344	51	.178
21	1.145	106	.178	21	.323	45	.177
22	.550	88	.178	22	.408	48	.171
23	.741	109	.176	23	.339	44	.176
24	.765	89	.170	24	.456	49	.165
25	.647	89	.178	25	.363	47	.176
26	.871	103	.169	26	.356	43	.177
27	.634	88	.177	27	.328	51	.172
28	.933	108	.178	28	.414	43	.176
29	.867	104	.173	29	.294	52	.178
30	.756	94	.178	30	.328	45	.178
Srednia	.850	96.3			.374	48.3	
Wariancja	.036	50.1			.004	1.9	
Odch st	.190	7.1			.064	3.4	

WYMIAR = 4 N = 32 DOKLADNOSC = .156

PLAN LOSOWY

PLAN QUASIOPTYMALNY

Lp	dN	n	d	Lp	dN	n	d
1	.559	97	.156	1	.275	60	.149
2	.702	97	.149	2	.358	55	.156
3	.611	100	.153	3	.362	52	.147
4	.647	113	.155	4	.369	58	.151
5	.623	109	.155	5	.355	50	.155
6	.783	98	.154	6	.291	52	.156
7	.582	102	.156	7	.363	67	.149
8	.513	110	.151	8	.270	55	.156
9	.665	103	.154	9	.355	49	.150
10	.585	102	.155	10	.291	53	.156
11	.673	99	.155	11	.286	47	.153
12	.467	102	.155	12	.229	59	.156
13	.751	99	.156	13	.329	52	.147
14	.723	120	.154	14	.357	54	.154
15	.747	116	.156	15	.411	67	.148
16	.601	98	.154	16	.422	62	.143
17	.590	99	.153	17	.380	50	.154
18	.729	109	.156	18	.316	55	.154
19	1.030	115	.153	19	.336	49	.154
20	.517	101	.152	20	.279	57	.150
21	.671	120	.154	21	.337	57	.150
22	.680	114	.149	22	.287	55	.156
23	.543	109	.154	23	.394	51	.151
24	.915	96	.156	24	.309	51	.155
25	.715	105	.156	25	.284	53	.146
26	.595	98	.153	26	.383	66	.155
27	.633	98	.156	27	.274	55	.154
28	.641	98	.155	28	.274	56	.154
29	.669	133	.155	29	.309	51	.154
30	.538	101	.145	30	.256	49	.156
Srednia	.657	105.7			.325	54.9	
Wariancja	.013	82.9			.002	27.4	
Odch st	.115	9.1			.049	5.2	

WYMIAR = 5 N = 8 DOKLADNOSC = .750

PLAN LOSOWY				PLAN QUASIOPTYMALNY			
Lp	dN	n	d	Lp	dN	n	d
1	47.648	42	.710	1	7.110	22	.668
2	12.586	29	.710	2	8.078	22	.690
3	9.607	34	.709	3	9.361	30	.748
4	20.376	35	.695	4	4.686	21	.687
5	61.569	36	.697	5	5.050	24	.650
6	10.775	35	.736	6	10.326	20	.632
7	9.589	41	.744	7	505.026	25	.734
8	11.040	58	.710	8	8.473	23	.644
9	50.186	37	.721	9	18.733	20	.626
10	90.743	41	.748	10	389.263	23	.712
11	16.428	36	.727	11	7.645	29	.743
12	20.964	34	.725	12	2.704	27	.607
13	19.839	36	.741	13	3.289	21	.673
14	5.023	34	.745	14	5.075	25	.739
15	5.763	36	.738	15	6.638	22	.725
16	9.864	29	.732	16	6.754	23	.643
17	31.106	36	.704	17	12.132	21	.715
18	12.534	39	.744	18	4.004	19	.666
19	54.067	36	.708	19	6.039	20	.708
20	7.854	32	.750	20	896.327	22	.710
21	120.836	44	.717	21	568.421	19	.640
22	6.361	31	.743	22	7.144	25	.709
23	28.481	43	.740	23	6.924	28	.737
24	8.785	33	.725	24	5.397	25	.676
25	9.758	39	.743	25	5.317	25	.700
26	13.840	39	.712	26	6.453	25	.742
27	31.314	42	.663	27	293.147	26	.624
28	8.408	42	.691	28	5.370	22	.624
29	8.922	35	.743	29	8.009	21	.733
30	12.525	35	.746	30	6868.224	25	.558
Srednia	25.226	37.3			323.037	23.3	
Wariancja	706.389	29.9			1522079.606	8.0	
Odch st	26.578	5.5			1233.725	2.8	

WYMIAR = 5 N = 12 DOKŁADNOŚĆ = .500

PLAN LOSOWY

PLAN QUASIOPTYMALNY

Lp	dN	n	d	Lp	dN	n	d
1	4.150	58	.499	1	1.296	34	.391
2	5.740	49	.496	2	1.511	30	.497
3	4.127	59	.498	3	1.713	25	.476
4	6.069	58	.481	4	3.142	28	.465
5	3.841	48	.500	5	3.503	31	.469
6	8.888	55	.492	6	1.542	33	.462
7	3.163	63	.499	7	2.008	25	.487
8	5.523	51	.475	8	1.128	24	.474
9	13.399	49	.499	9	2.200	36	.437
10	11.398	48	.495	10	1.791	32	.499
11	5.947	51	.496	11	2.691	24	.449
12	8.285	47	.494	12	2.865	27	.479
13	5.828	53	.497	13	4.274	24	.483
14	10.303	59	.479	14	2.510	25	.445
15	6.824	45	.498	15	6.042	30	.430
16	3.097	45	.488	16	2.308	30	.440
17	4.146	51	.485	17	2.216	31	.457
18	4.685	60	.497	18	6.548	35	.50
19	3.613	50	.495	19	2.174	27	.388
20	6.338	55	.489	20	1.705	32	.495
21	4.181	49	.484	21	3.856	33	.479
22	6.418	49	.500	22	5.781	34	.487
23	4.805	56	.496	23	2.982	34	.425
24	14.745	64	.480	24	3.423	27	.497
25	9.068	49	.492	25	118.341	28	.495
26	6.486	61	.434	26	2.035	31	.490
27	8.646	62	.441	27	4.756	28	.456
28	4.911	54	.497	28	1.990	26	.459
29	3.631	55	.441	29	1.929	28	.471
30	5.134	64	.427	30	1.570	34	.481

Srednia 6.446 53.9
 Wariancja 8.521 32.4
 Odch st 2.919 5.7

6.661 29.5
 432.042 13.0
 20.786 3.6

WYMIAR = 5 N = 16 DOKLADNOSC = .375

PLAN LOSOWY

PLAN QUASIOPTYMALNY

Lp	dN	n	d	Lp	dN	n	d
1	2.838	68	.371	1	.993	28	.353
2	2.614	67	.351	2	1.012	37	.353
3	2.368	66	.375	3	1.315	32	.366
4	2.011	53	.373	4	2.255	35	.321
5	3.795	61	.374	5	1.212	38	.357
6	2.270	58	.361	6	1.870	47	.345
7	3.130	67	.373	7	1.471	36	.357
8	3.233	57	.373	8	1.846	36	.340
9	5.531	73	.370	9	1.491	29	.365
10	2.635	65	.369	10	1.123	36	.362
11	2.046	59	.375	11	.929	38	.375
12	2.595	72	.374	12	1.491	37	.339
13	5.575	62	.350	13	1.651	38	.361
14	2.587	68	.333	14	1.436	38	.365
15	1.764	63	.373	15	2.369	36	.369
16	3.875	63	.374	16	1.283	34	.332
17	2.690	54	.375	17	1.431	32	.351
18	3.342	70	.370	18	1.827	35	.374
19	3.201	74	.367	19	1.803	37	.371
20	2.377	59	.372	20	.899	30	.347
21	3.471	70	.370	21	2.264	37	.374
22	4.034	62	.363	22	1.483	36	.352
23	1.293	51	.371	23	.679	26	.360
24	4.009	66	.356	24	1.599	35	.367
25	5.804	60	.360	25	1.786	40	.340
26	2.312	66	.361	26	1.988	36	.361
27	2.182	60	.368	27	.949	30	.366
28	17.629	62	.360	28	2.304	37	.356
29	3.053	60	.355	29	1.150	32	.366
30	1.896	62	.374	30	1.298	27	.349

Srednia 3.539 63.3
 Wariancja 8.029 32.1
 Odch st 2.834 5.7

1.507 34.8
 .200 17.9
 .447 4.2

WYMIAR = 5 N = 20 DOKŁADNOŚĆ = .300

PLAN LOSOWY				PLAN QUASIOPTYMALNY			
Lp	dN	n	d	Lp	dN	n	d
1	1.301	81	.300	1	1.101	46	.274
2	2.364	71	.295	2	.674	36	.299
3	3.208	79	.297	3	.809	46	.284
4	1.568	74	.298	4	.854	48	.251
5	3.461	73	.298	5	1.222	52	.264
6	1.378	76	.293	6	.643	34	.300
7	2.716	82	.299	7	.928	38	.292
8	2.013	83	.294	8	.853	36	.274
9	1.363	78	.278	9	.608	38	.296
10	1.595	84	.293	10	1.102	43	.280
11	2.954	76	.292	11	.604	37	.292
12	2.396	74	.295	12	1.313	46	.294
13	1.262	76	.298	13	1.046	44	.245
14	2.234	78	.290	14	.764	37	.299
15	2.101	71	.294	15	.720	40	.286
16	1.735	69	.295	16	.843	42	.288
17	1.613	79	.297	17	.950	40	.297
18	2.070	87	.282	18	.892	38	.297
19	1.540	74	.299	19	1.079	52	.292
20	2.295	84	.297	20	1.109	42	.299
21	2.996	83	.298	21	1.127	43	.295
22	4.220	80	.280	22	.722	34	.299
23	1.824	76	.275	23	.823	45	.286
24	2.746	86	.298	24	1.061	37	.284
25	1.431	93	.293	25	.954	38	.286
26	1.950	80	.300	26	1.095	47	.299
27	2.708	74	.291	27	1.096	42	.282
28	1.888	71	.285	28	1.068	49	.292
29	1.983	69	.290	29	1.319	44	.294
30	2.365	72	.298	30	.861	42	.275
Srednia	2.176	77.8			.941	41.9	
Wariancja	.488	32.5			.039	24.1	
Odch st	.698	5.7			.197	4.9	

WYMIAR = 5 N = 24 DOKŁADNOŚĆ = .250

PLAN LOSOWY				PLAN QUASIOPTYMALNY			
Lp	dN	n	d	Lp	dN	n	d
1	1.530	92	.238	1	.685	44	.243
2	.965	83	.249	2	.555	42	.245
3	1.312	89	.247	3	.743	45	.249
4	1.143	91	.249	4	.603	38	.247
5	2.000	80	.248	5	.543	41	.247
6	1.502	91	.250	6	.892	46	.242
7	1.197	97	.230	7	.654	53	.245
8	1.830	93	.247	8	.781	50	.234
9	1.052	88	.247	9	1.129	49	.248
10	1.666	85	.248	10	.537	42	.231
11	1.418	97	.238	11	.640	48	.231
12	1.843	83	.248	12	.631	56	.249
13	1.397	98	.235	13	.919	45	.227
14	1.509	97	.248	14	.387	43	.237
15	1.057	91	.243	15	.783	44	.246
16	1.246	78	.248	16	.567	51	.206
17	1.505	91	.242	17	.677	48	.247
18	1.495	92	.249	18	.721	49	.249
19	1.577	82	.248	19	.480	43	.250
20	1.272	94	.248	20	.626	51	.248
21	1.178	99	.243	21	.604	43	.250
22	1.335	85	.247	22	.515	50	.244
23	1.147	85	.250	23	.645	48	.241
24	1.173	93	.250	24	.866	51	.248
25	2.232	83	.245	25	.843	46	.240
26	1.394	93	.249	26	.532	50	.247
27	1.445	96	.245	27	.701	50	.244
28	2.752	78	.249	28	.562	46	.236
29	1.124	81	.248	29	.618	57	.247
30	1.197	85	.250	30	.490	39	.246
Srednia					.664	46.9	
Wariancja					.023	20.7	
Odch st					.153	4.5	

WYMIAR = 5 N = 28 DOKŁADNOŚĆ = .214

PLAN LOSOWY

PLAN QUASIOPTYMALNY

Lp	dN	n	d	Lp	dN	n	d
1	1.108	103	.213	1	.620	49	.211
2	1.002	88	.212	2	.466	46	.205
3	.942	92	.214	3	.433	51	.214
4	1.920	108	.212	4	.590	60	.198
5	.922	96	.211	5	.401	47	.213
6	1.800	120	.212	6	.762	52	.213
7	1.164	105	.211	7	.572	52	.210
8	.881	90	.214	8	.783	69	.210
9	.991	96	.212	9	.526	63	.212
10	1.187	96	.213	10	.495	56	.200
11	1.227	98	.206	11	.444	53	.207
12	1.107	102	.214	12	.680	51	.208
13	1.840	102	.210	13	.494	53	.207
14	1.503	100	.214	14	.363	60	.183
15	1.668	111	.210	15	.491	48	.207
16	1.065	124	.206	16	.516	51	.198
17	1.345	99	.210	17	.450	53	.211
18	1.414	102	.214	18	.956	58	.210
19	.844	110	.206	19	.526	57	.210
20	1.284	114	.202	20	.430	56	.212
21	1.197	97	.210	21	.475	53	.214
22	.995	112	.211	22	.483	48	.211
23	.927	95	.214	23	.412	49	.212
24	1.180	106	.202	24	.534	53	.196
25	1.483	105	.212	25	.570	50	.204
26	2.251	100	.213	26	.433	54	.212
27	1.270	105	.211	27	.417	50	.212
28	.950	110	.206	28	.403	51	.208
29	1.216	95	.211	29	.586	61	.208
30	1.080	95	.209	30	.484	51	.207
Srednia	1.259	102.5			.527	53.5	
Wariancja	.115	69.5			.016	25.9	
Odch st	.339	8.3			.127	5.1	

WYMIAR = 5 N = 32 DOKLADNOSC = .188

PLAN LOSOWY				PLAN QUASIOPTYMALNY			
Lp	dN	n	d	Lp	dN	n	d
1	.912	107	.186	1	.369	55	.183
2	.928	124	.186	2	.451	62	.179
3	.875	111	.181	3	.375	64	.186
4	.630	114	.187	4	.337	57	.177
5	.743	95	.186	5	1.266	68	.186
6	.901	112	.183	6	.420	65	.187
7	.824	102	.183	7	.404	64	.181
8	.982	116	.183	8	.429	54	.182
9	.891	117	.186	9	.424	56	.185
10	.736	134	.175	10	.468	60	.182
11	.736	111	.184	11	.362	59	.178
12	1.341	126	.174	12	.404	54	.179
13	1.270	136	.178	13	.430	62	.182
14	.988	122	.182	14	.516	70	.171
15	.794	110	.185	15	.368	63	.184
16	.942	120	.187	16	.556	57	.179
17	1.115	109	.186	17	.369	56	.182
18	.743	105	.187	18	.541	57	.187
19	.826	101	.186	19	.431	63	.184
20	1.003	109	.185	20	.452	75	.181
21	.767	109	.187	21	.538	67	.182
22	.831	100	.183	22	.419	57	.168
23	.904	113	.184	23	.354	52	.182
24	1.100	122	.184	24	.591	64	.177
25	1.170	106	.187	25	.550	57	.166
26	.724	110	.186	26	.368	54	.186
27	.919	122	.183	27	.388	72	.187
28	.946	118	.182	28	.365	60	.186
29	1.498	114	.186	29	.382	58	.187
30	1.160	119	.181	30	.431	59	.179
Srednia	.940	113.8			.459	60.7	
Wariancja	.038	86.8			.027	31.5	
Odch st	.196	9.3			.164	5.6	

LITERATURA CYTOWANA

- [1] Atwood, C.L., Optimal and efficient designs of experiments, Ann. Math. Statist. 40(1969), str. 1570 - 1602
- [2] Atwood, C.L., Sequences converging to D-optimal designs of experiments, Annals of Statistics 1(1973), str. 342 - 352
- [3] Bandemer, Optimale Wahl der Abszissenwerte für die lineare Regression bei gegebener Kostenfunktion, Math. OF Statist. 1(1970), str. 297 - 308
- [4] Bandemer, H., Bellmann, A., Jung, W., Richter, K., Optimale Versuchsplanung, Berlin 1973
- [5] Box, G.E.P., and Draper, N.R., A basic for selection of response surface design, J. Amer. Statist. Assoc. 54(1959), str. 622 - 654
- [6] Box, G.E.P., and Hill, W.T., Discrimination among mechanistic models, Technometrics 9(1967), str. 57 - 71
- [7] Farrell, R.H., Kiefer, J. and Walbran, A., Optimum multivariate designs, Proc. 5-th Berkeley Symp. of Math. Statist. and Probab. 1(1967), str. 113 - 138
- [8] Fedorov, V.V. and Maljutov, M.B., Optimal designs in regression problems, Math. OF Statist. 3(1972), str. 281 - 308
- [9] Fiedorow, W.W., Teoria optimalnego eksperymentu, Moskwa 1971
- [10] Fiedorow, W.W., Asymptotycznie optymalne plany eksperymentów po dyskryminacji dwóch regresyjnych modeli, Teoria ~~W~~ Prawdopodob. i Statyst. 16(1971), str. 576 - 577
- [11] Finney, D.T., An introduction to the theory of experimental design, Chicago 1967

- [12] Fisher, R.A., The design of experiments, London 1951
- [13] Guest, P.G., The spacing of observations in polynomial regression, Ann. Math. Stat. 29(1958), str. 294 - 299
- [14] Hill, W.J., Hunter W.G. and Wichern, D.W., A joint design criterion for the dual problem of model discrimination and parameter estimation, Technometrics 10(1968), str. 145 - 160
- [15] Hill, W.J. and Hunter, W.G., A note on designs for model discrimination: variance unknown case, Technometrics 11 (1969), str. 396 - 402
- [16] Hoel, P.G., Efficiency problems in polynomial estimation, Ann. Math. Statist. 29(1958), str. 1134 - 1145
- [17] Hoel, P.G., Minimax designs in twodimensional regression, Ann. Math. Statist. 36(1965), str. 1097 - 1106
- [18] Hoel, P.G., Optimum designs for polynomial extrapolation, Ann. Math. Statist. 36(1965), str. 1483 - 1493
- [19] Hoel, P.G., A simple solution for optimal Tchebyshev regression extrapolation, Ann. Math. Statist. 37(1966), str. 720 - 725
- [20] Hoel, P.G. and Levine A., Optimal spacing and weighting in polynomial prediction, Ann. Math. Statist. 35(1964), str. 1553 - 1560
- [21] Hunter, W.G. and Reiner A.M., Design for discriminating between two rival models, Technometrics 7(1965), str. 307 - 324
- [22] Jung, W., G-optimale Versuchsplanung für die Regressionsgerade $y(x) = \theta_1 + \theta_2 x$, Biometrische Zeitschrift 13 (1971), str. 369 - 375

- [23] Jung, W., Kostenoptimale Versuchsplanung im Regressionsmodell, Biometrische Zeitschrift 16(1974), str. 31 - 47
- [24] Kacprzyński, B., Planowanie eksperymentów - podstawy matematyczne, Warszawa 1974
- [25] Karlin, S.J. and Studden, W.J., Tchebycheff systems:with applications in analysis and statistics, New York 1966
- [26] Karlin, S.J. and Studden W.J., Optimal experimental designs, Ann. Math. Statist. 37(1966), str. 783 - 815
- [27] Kiefer, J., Optimum experimental designs, J.Roy.Statist. Soc. Ser. B 21(1959), str. 272 - 319
- [28] Kiefer, J., Optimum designs in regression problems II, Ann. Math. Statist. 32(1961), str. 298 - 325
- [29] Kiefer, J., Optimum experimental designs V, Proc. 4-th Berkeley Symp. of Math. Statist. and Probab. 1(1961), str. 381 - 405
- [30] Kiefer, J. and Wolfowitz J., Optimum designs in regression problems, Ann. Math. Statist. 30(1959), str. 271 - 294
- [31] Kiefer, J. and Wolfowitz, J., The equivalence of two extremum problems, Canadian Journal of Math. 12(1960), str. 363 - 366
- [32] Kiefer, J. and Wolfowitz, J., Optimum extrapolation and interpolation designs, Ann. Inst. of Statistical Math. Tokyo 16(1964), str. 79 - 108, 295 - 303
- [33] Kiefer, J. and Wolfowitz, J., On a theorem of Hoel and Levine on extrapolation designs, Ann. Math. Statist. 36(1965), str. 1627 - 1655
- [34] Maljutow, M.B. i Fiedorow, W.W., O planach wzwieszЕННОj polinomialnoj regressii s minimalnoj sriedniej dispersiej, Teorja Wierojatn. Primien. 16(1971), str. 734 - 738

- [35] Nalimov, V.V., Golikova, T.I. and Mikeschina, N.G., On practical use of the concept of D-optimality, Technometrics 12(1970), str. 799 - 812
- [36] Praca zbiorowa, Nowyje idiei w planirowanii eksperimenta, Moskwa 1969
- [37] Scheffé, H., The analysis of variance, New York 1958
- [38] Studden, W.J., Optimal designs on Tchebycheff points, Ann. Math. Statist. 39(1968), str. 1435 - 1447
- [39] Studden, W.J., Optimal designs for multivariate polynomial extrapolation, Ann. Math. Statist. 42(1971), str. 828 - 832
- [40] Wynn, H.P., The sequential generation of D-optimum experimental designs, Ann. Math. Statist. 41(1970), str. 1655 - 1664