

ELEMENTY TEORII RANDOMIZACJI

III. RANDOMIZACJA W DOŚWIADCZENIACH BLOKOWYCH

RADOŚLAW KALA

Katedra Metod Matematycznych i Statystycznych
Akademii Rolniczej w Poznaniu

Streszczenie

Celem pracy jest pokazanie wpływu randomizacji na strukturę modelu doświadczeń blokowych. Rozpatrywane są tutaj dwa rodzaje randomizacji, randomizacja jednostek wewnątrz bloków oraz randomizacja bloków. W obu przypadkach rozważa się zagadnienie istnienia najlepszego estymatora parametrów obiektowych. Rozważania te obejmują także modele analizy wewnątrz i międzyblokowej.

1. WSTĘP

Dążenie do zwiększenia precyzji eksperymentu leży u podstaw teorii układów blokowych. W tej pracy prześledzimy wpływ procesów randomizacyjnych na postacie modeli takich doświadczeń. Omówiona tu będzie zarówno randomizacja wewnątrz bloków, jak i randomizacja bloków. Wyprowadzone modele, podobnie jak te rozważane w pracy Kali (1990), będą rozważone z punktu widzenia możliwości uzyskania najlepszych liniowych estymatorów nieobciążonych funkcji parametrów stałych. W tym kontekście przedyskutowana będzie praktyczna przydatność modeli analizy wewnątrz i międzyblokowej.

W pracy obecnej, będącej kontynuacją poprzednich dwóch prac Kali (1989, 1990), korzysta się z pojęć i wyników tam wprowadzonych. Odwołania do wzorów i twierdzeń podanych w tamtych pracach poprzedzone są odpowiednio oznaczeniami I i II.

Słowa kluczowe: addytywność, jednorodność, doświadczenia ślepe, modele stałe, modele losowe, modele mieszane

2. RANDOMIZACJA JEDNOSTEK W POPULACJACH WARSTWOWYCH

Zastosowanie w eksperymencie randomizacji jednostek jest podyktowane przede wszystkim potrzebą ich modelowej unifikacji. Randomizacja jest przy tym postępowaniem konkurencyjnym w stosunku do selektywnego wyboru jednostek identycznych, który jednakże nie zawsze może być akceptowany, gdyż nieraz zbyt radykalnie modyfikuje populację podstawową na tle której formułowane mają być wnioski końcowe. W dalszym ciągu zajmiemy się postępowaniem stanowiącym kompromis pomiędzy tymi dwoma sposobami wyrównania jednostek.

Przypuśćmy, że populacja podstawowa jednostek eksperymentalnych jest skończona oraz że jednostki tego zbioru dają się pogrupować w tzw. warstwy złożone z jednostek charakteryzujących się identycznością wartości ustalonego podzbioru cech wyróżnionych. Grupowanie takie jest szczególnie łatwe, gdy wśród cech wyróżnionych są cechy dyskretne, bądź, gdy bez straty na dokładności opisu jednostek, niektóre spośród wyróżnionych na nich cech ciągłych można skategoryzować.

Formalizując ten proces przyjmijmy, że cechy stanowiące podstawę grupowania jednostek tworzą podwektor $x^2 \in R^{q_2}$ wektora cech wyróżnionych x ,

$$x = \begin{pmatrix} x^1 \\ x^2 \end{pmatrix} \in R^q, \quad (2.1)$$

przy czym $q = q_1 + q_2$. Populację \mathcal{P}_N jednostek eksperymentalnych dzielimy teraz na warstwy dokonując selekcji jednostek identycznych względem cech ujętych w wektorze x^2 . W rezultacie k -ta warstwa jest zbiorem jednostek $\{u_{1k}, u_{2k}, \dots, u_{N_k k}\}$, gdzie jednostka u_{jk} scharakteryzowana jest parą wektorów

$$\begin{aligned} x^1(u_{jk}) &= x^1_{jk}, \\ x^2(u_{jk}) &= x^2_k, \end{aligned} \quad j = 1, 2, \dots, N_k, \quad (2.2)$$

przy czym drugi z wymienionych tu wektorów, x^2_k , jest taki sam dla wszystkich jednostek tej warstwy.

W każdej wyselekcjonowanej warstwie prowadzimy teraz randomizację jednostek. W rezultacie jednostki k -tej warstwy można przedstawić za pomocą wektorów losowych $x_{(j)k}$, $j = 1, 2, \dots, N_k$, każdy o wartości oczekiwanej

$$E(x_{(j)k}) = E \begin{pmatrix} x^1_{(j)k} \\ x^2_k \end{pmatrix} = \begin{pmatrix} x^1_k \\ x^2_k \end{pmatrix}, \quad (2.3)$$

gdzie

$$\bar{x}_k^{-1} = \frac{1}{N_k} \sum_{j=1}^{N_k} x^1_{jk}, \quad (2.4)$$

i o macierzy dyspersji

$$D(\mathbf{x}_{(j),k}) = \begin{bmatrix} \Sigma_k^{11} & \mathbf{0} \\ \mathbf{0} & \mathbf{0} \end{bmatrix}, \quad (2.5)$$

gdzie

$$N_k \Sigma_k^{11} = \sum_{j=1}^k \mathbf{x}_{jk}^1 (\mathbf{x}_{jk}^1)' - N_k \bar{\mathbf{x}}_k^1 (\bar{\mathbf{x}}_k^1)', \quad (2.6)$$

Macierz kowariancji dwóch zrandomizowanych jednostek pochodzących z tej samej warstwy określa natomiast wzór

$$C(\mathbf{x}_{(j),k}, \mathbf{x}_{(j'),k}) = \frac{-1}{N_k - 1} \begin{bmatrix} \Sigma_k^{11} & \mathbf{0} \\ \mathbf{0} & \mathbf{0} \end{bmatrix}, \quad j \neq j', \quad (2.7)$$

przy czym kowariancja pomiędzy dowolnymi zrandomizowanymi jednostkami z dwóch różnych warstw jest równa zeru. W rezultacie jednostki wewnątrz każdej warstwy charakteryzuje symetria rozkładu łącznego, przy czym własność ta w odniesieniu do cech zawartych w wektorze \mathbf{x}^2 została uzyskana poprzez selekcję, a w odniesieniu do cech zawartych w wektorze \mathbf{x}^1 , poprzez randomizację.

Przedstawione rozważania pozwalają również łatwo skonstruować model jednostek wybranych losowo z populacji, w której warstwy są nieskończone. W tym przypadku jednostki ustalonej, powiedzmy k -tej, warstwy będą opisane zmiennymi losowymi \mathbf{x}_{jk} , $j = 1, 2, \dots$, każda o wartości oczekiwanej

$$E(\mathbf{x}_{jk}) = E \begin{bmatrix} \mathbf{x}_{jk}^1 \\ \mathbf{x}_{jk}^2 \end{bmatrix} = \begin{bmatrix} \boldsymbol{\mu}_k^1 \\ \mathbf{x}_k^2 \end{bmatrix}, \quad (2.8)$$

gdzie $\boldsymbol{\mu}_k^1$ jest wartością oczekiwaną warunkową wektora losowego \mathbf{x}^1 przy warunku $\mathbf{x}^2 = \mathbf{x}_k^2$, i macierzy dyspersji

$$D(\mathbf{x}_{jk}) = \begin{bmatrix} \Sigma_k^{11} & \mathbf{0} \\ \mathbf{0} & \mathbf{0} \end{bmatrix}, \quad (2.9)$$

gdzie Σ_k^{11} jest warunkową macierzą dyspersji wektora \mathbf{x}^1 , gdy $\mathbf{x}^2 = \mathbf{x}_k^2$. W tym przedstawieniu zmienne losowe dotyczące dwóch różnych jednostek będą traktowane jako niezależne.

3. MODELE BLOKOWE

Zajmiemy się teraz konstrukcją modeli dla eksperymentów, w których zbiór γ_v ustalonych obiektów jest porównywany na jednostkach pochodzących z populacji o strukturze warstwowej. Uwzględnimy tutaj populacje skończone i nieskończone, przy czym liczba warstw będzie w obu przypadkach skończona, a warstwy będą bądź wszystkie zbiorami skończonymi, bądź wszystkie zbiorami nieskończonymi. Ponadto założymy, że w eksperymencie

wszystkie warstwy mają być reprezentowane.

Na początek zauważmy, że posługując się randomizacją, w wypadku warstw skończonych, lub konstruując próby proste, w wypadku warstw nieskończonych, uzyskujemy własność symetrii (wymienialności) jednostek w ramach poszczególnych warstw populacji podstawowej. W związku z tym jednostki tworzące wspólną warstwę stanowią dogodną bazę dla porównania badanych obiektów. Fakt ten winien być uwzględniony w planie eksperymentu, który określa tym razem liczby replikacji dla poszczególnych obiektów w ramach każdej warstwy. Jeżeli zatem plan eksperymentu przewiduje n_{sk} replikacji dla s -tego obiektu w k -tej warstwie, to z warstwy tej należy pobrać próbę prostą (w wypadku warstwy nieskończonej \mathcal{P}_∞) lub próbę zrandomizowaną (w wypadku warstwy skończonej \mathcal{P}_{N_k}), o licznosci $n_k = \sum_{s=1}^w n_{sk}$. Przy B warstwach, pobranych prób będzie B . Oznaczmy je odpowiednio przez $\mathcal{P}_{n_1}, \mathcal{P}_{n_2}, \dots, \mathcal{P}_{n_B}$ i będziemy nazywać blokami.

Z uwagi na własność symetrii (wymienialności) jednostek w blokach, przydzielanie obiektów do jednostek zrandomizowanych wewnątrz bloków, a dokładniej, do numerów tych jednostek ustalonych w procesie randomizacji, jest zupełnie dowolne. Oczywiście musi być ono zgodne z planem doświadczenia, który winien gwarantować pożądany stopień spójności i efektywności układu, a także uwzględniać koszt eksperymentu. W tym sensie proces randomizacji jednostek nie stoi w sprzeczności z teorią optymalności układów doświadczalnych (porównaj Harville, 1975, s.27).

Kolejnym elementem wymagającym sprecyzowania jest cecha badana. Tak jak w przypadku modeli już rozważanych, jest ona kombinacją liniową cech wyróżnionych na jednostkach, którą obecnie można zapisać w postaci

$$\mathbf{x}'\mathbf{a} = (\mathbf{x}^1)'\mathbf{a}^1 + (\mathbf{x}^2)'\mathbf{a}^2, \quad (3.1)$$

gdzie \mathbf{a}^1 i \mathbf{a}^2 są podwektorami wektora \mathbf{a} odpowiedniego rozmiaru. Zauważmy, że w tym przedstawieniu faktycznie tylko pierwszy składnik reprezentuje zmienną losową, natomiast składnik drugi jest stałą charakterystyczną dla każdego bloku.

Korzystając ze wzorów (2.3), (2.4), (2.8) i (2.9) można łatwo ustalić parametry cechy badanej. Wartość oczekiwaną i wariancję dla jednostek k -tego bloku wyrazimy w dalszym ciągu używając oznaczeń

$$m_{u,k} = \begin{cases} (\mu_k^1)'\mathbf{a}^1 + (\mathbf{x}_k^2)'\mathbf{a}^2, & \mathcal{P}_{n_k} \subset \mathcal{P}_\infty, \\ (\bar{\mathbf{x}}_k^1)'\mathbf{a}^1 + (\mathbf{x}_k^2)'\mathbf{a}^2, & \mathcal{P}_{n_k} \subset \mathcal{P}_{N_k}, \end{cases} \quad (3.2)$$

oraz

$$\sigma_{u,k}^2 = \begin{cases} (\mathbf{a}^1)'\Sigma_k^{11}\mathbf{a}^1, & \mathcal{P}_{n_k} \subset \mathcal{P}_\infty, \\ \{N_k/(N_k-1)\}(\mathbf{a}^1)'\mathbf{S}_k^{11}\mathbf{a}^1, & \mathcal{P}_{n_k} \subset \mathcal{P}_{N_k}. \end{cases} \quad (3.3)$$

Chcąc opisać potencjalną obserwację cechy badanej uzyskanej na j -tej jednostce w k -tym bloku poddanej działaniu s -tego obiektu zauważmy, że jeśli tylko w odniesieniu do wszystkich obiektów i jednostek biorących

udział w eksperymencie spełniona jest zasada addytywności, to każdy blok stanowi w istocie proste doświadczenie, w którym cecha badana może być wyrażona zgodnie z równością (II.4.1). Stąd obserwację tę możemy zapisać w postaci

$$y_{jk}(s) = y_{jk} + t_s, \quad (3.4)$$

gdzie y_{jk} reprezentuje wartość cechy badanej na j -tej jednostce w k -tym bloku "wolnej" od działania obiektu s , natomiast t_s jest efektem obiektu s jednakowym dla wszystkich jednostek potraktowanych tym obiektem.

Zwróćmy tutaj uwagę, że przyjęcie zasady addytywności oznacza w istocie zignorowanie tzw. efektu interakcji blokowo-objektowej. Gdyby takie współdziałanie miało być w modelu uwzględnione, należałoby w równości (3.4) efekt t_s zastąpić efektem t_{ks} uzależniając go tym samym od bloku. Taka modyfikacja równości (3.4) daje się łatwo uzasadnić, jeśli proces modelowania oprzeć o słabszą zasadę zachowania symetrii. Mimo tej możliwości, prowadzącej do jeszcze bardziej rozbudowanych modeli, w dalszym ciągu postąpimy tradycyjnie, zgodnie z zasadą addytywności.

Biorąc teraz pod uwagę równość (3.4), oznaczenia (3.2) i (3.3) oraz fakt skorelowania jednostek zrandomizowanych, wyrażony relacją (2.7), możemy zapisać dwa modele dla wszystkich obserwacji zebranych w wyniku przeprowadzenia eksperymentu blokowego. Dla populacji nieskończonej mamy mianowicie model

$$\{y, D'm + \Delta't, \text{diag}_{k=1}^B ((\sigma_{u,k}^2 + \sigma_{\epsilon}^2) I_{n_k})\}, \quad \mathcal{P}_{n_k} \subset \mathcal{P}_{\infty}, \quad (3.5)$$

a dla populacji skończonej

$$\{y, D'm + \Delta't, \text{diag}_{k=1}^B (\sigma_{u,k}^2 (I_{n_k} - \frac{1}{N_k} J_{n_k})) + \sigma_{\epsilon}^2 I_{n_k}\}, \quad \mathcal{P}_{n_k} \subset \mathcal{P}_{N_k}, \quad (3.6)$$

gdzie

$$y = (y_{11}(1), y_{21}(1), \dots, y_{n_B}(v))' \quad (3.7)$$

jest wektorem obserwowanych zmiennych losowych, wektor t określony jest w (II.4.5),

$$m = (m_{u,1}, m_{u,2}, \dots, m_{u,B})' \quad (3.8)$$

jest wektorem nieznanych efektów blokowych,

$$D' = \text{diag}_{k=1}^B (1_{n_k}) \quad (3.9)$$

jest $n \times B$ -wymiarową macierzą dla bloków, a

$$\Delta' = (\text{diag}_{i=1}^v (1'_{n_{1i}}) : \dots : \text{diag}_{i=1}^v (1'_{n_{1i}}))' \quad (3.10)$$

jest $n \times v$ -wymiarową macierzą określającą przyporządkowanie obiektów do jednostek w kolejnych blokach.

Jak łatwo zauważyć, w obydwu wyprowadzonych modelach wartości oczekiwane są takie same, natomiast macierze dyspersji są różne. W modelu (3.5) obserwacje nie są skorelowane, a łączna liczba komponentów wariancji jest równa B. Są nimi wielkości

$$\sigma_k^2 = \sigma_{u,k}^2 + \sigma_e^2, \quad k = 1, 2, \dots, B. \quad (3.11)$$

W modelu (3.6) obserwacje wewnątrz bloków są skorelowane, a łączna liczba komponentów wariancji jest o jeden większa. Bliższa analiza wymienionych parametrów pozwala stwierdzić, że zmniejszenie wielkości komponentów, co w przypadku rozważanych modeli wpływa na poprawienie precyzji eksperymentu, jest związane z taką selekcją jednostek, która dostarczy warstw o możliwie najmniejszym rozproszeniu jednostek, przy czym jest to ważne tylko dla tych cech wyróżnionych, które w kombinacji $(x^1)'a^1$ występują z niezerowymi współczynnikami. Gdyby rozproszenie to można było zaniedbać, wtedy macierze dyspersji obu modeli zredukowałyby się do struktury prostej. Analogiczną redukcję możnaby uzyskać, gdyby cecha badana była kombinacją liniową tylko tych cech, które stanowią podstawę tworzenia warstw, a w konsekwencji bloków. Kładąc mianowicie $a^1 = 0$ mamy $\sigma_{u,k}^2 = 0$ dla $k = 1, 2, \dots, B$.

Z tej ostatniej obserwacji wypływa praktyczny wniosek mówiący o konieczności tworzenia bloków w oparciu o te ich cechy, które mogą mieć istotny wpływ na zachowanie cechy badanej.

Zauważmy także, że spełnienie równości

$$\sigma_{u,1}^2 = \sigma_{u,2}^2 = \dots = \sigma_{u,B}^2 = \sigma_*^2, \quad (3.12)$$

którą nazwiemy warunkiem jednorodności bloków, powoduje uproszczenie struktury macierzy dyspersji w obu modelach. Model (3.5) staje się wtedy modelem stałym układu blokowego. Charakteryzuje go prosta postać macierzy dyspersji z jednym komponentem wariancji $\sigma^2 = \sigma_*^2 + \sigma_e^2$. Teoria tego modelu jest dobrze opracowana (patrz np. Pearce i in., 1974).

Model (3.6), przy spełnieniu warunku jednorodności, pozostaje nadal modelem mieszanym, lecz liczba komponentów wariancji redukuje się do dwóch. W obu tak uproszczonych modelach estymacja parametrów stałych oparta na metodzie najmniejszych kwadratów, dostarcza estymatorów najlepszych w klasie estymatorów liniowych i nieobciążonych. Spostrzeżenie to w odniesieniu do modelu (3.5) z warunkiem (3.12) jest oczywiste. W odniesieniu do modelu (3.6) wynika z następującego twierdzenia.

Twierdzenie 1. Jeżeli w modelu (3.6) spełniony jest warunek (3.12), to najlepszym liniowym estymatorem nieobciążonym dowolnej estymowalnej funkcji parametrycznej jest estymator otrzymany metodą najmniejszych kwadratów.

Dowód. W świetle lematu II.B wystarczy sprawdzić prawdziwość następującej relacji

$$R(V(D':\Delta')) \subset R(D':\Delta'),$$

gdzie

$$V = \text{diag}_{k=1}^B (I_{n_k} - \frac{1}{N_k} J_{n_k}) = I_n - D' \text{diag}_{k=1}^B (\frac{1}{N_k} \mathbf{1}'_{n_k}),$$

a macierze D' i Δ' określone są odpowiednio w (3.9) i (3.10). W tym celu zauważmy, że

$$VD' = D' \{ I_n - \text{diag}_{k=1}^B (\frac{1}{N_k} \mathbf{1}'_{n_k}) D' \},$$

co implikuje $R(VD') \subset R(D') \subset R(D':\Delta')$. Z kolei $V\Delta' = \Delta' - D'Z$, gdzie

$$Z = \begin{bmatrix} \frac{1}{N_1} (n_{11} & n_{21} & \dots & n_{v1}) \\ \dots & \dots & \dots & \dots \\ \frac{1}{N_B} (n_{1B} & n_{2B} & \dots & n_{vB}) \end{bmatrix}.$$

Stąd $R(V\Delta') \subset R(D':\Delta')$. \square

Ważność warunku jednorodności bloków dla problemu estymacji w modelach (3.5) i (3.6) podkreśla jeszcze bardziej kolejne twierdzenie. Zanim jednak je sformułujemy, wprowadzimy pojęcie bloku zdegenerowanego. Określenie to będziemy wiązać z blokiem, którego jednostki nie są poddane działaniu żadnego obiektu, bądź z blokiem, w którym wszystkie jednostki poddano działaniu tego samego obiektu.

Twierdzenie 2. Jeżeli w modelach (3.5) i (3.6) warunek jednorodności bloków nie jest spełniony, a bloki nie są zdegenerowane, to nie istnieje najlepszy liniowy nieobciążony estymator wartości oczekiwanej wektora y .

Dowód. Ponieważ macierz jednostkowa należy do klasy macierzy dyspersji modelu (3.5), więc w świetle lematu II.B najlepszy liniowy nieobciążony estymator wartości oczekiwanej w tym modelu istnieje wtedy i tylko wtedy, gdy

$$R\{V_k(D':\Delta')\} \subset R(D':\Delta') \quad \text{dla } k = 1, 2, \dots, B, \quad (3.13)$$

gdzie

$$V_k = \text{diag}(0, \dots, I_{n_k}, \dots, 0). \quad (3.14)$$

Resztę dowodu poprowadzimy nie wprost pokazując, że warunek (3.13) jest sprzeczny z założeniem o braku bloków zdegenerowanych. W tym celu zauważmy, że dla ustalonego numeru bloku k relacja (3.13) implikuje istnienie takich macierzy A i B , że

$$V_k \Delta' = D'A + \Delta'B.$$

Porównując wybrane wiersze macierzy $V_k \Delta'$ oraz $D'A + \Delta'B$, a dokładniej wiersze odpowiadające pierwszemu i k -temu blokowi, mamy w szczególności

$$0 = (1_{n_1} : 0 : \dots : 0)A + \text{diag}_{i=1}^v (1_{n_{i1}})B \quad (3.15)$$

$$\text{diag}_{i=1}^v (1_{n_{ik}}) = (0 : \dots : 1_{n_k} : \dots : 0)A + \text{diag}_{i=1}^v (1_{n_{ik}})B. \quad (3.16)$$

Mnożąc równość (3.15) lewostronnie przez macierz $\text{diag}_{i=1}^v (n_{11}^{-1} \mathbf{1}_{n_{11}}')$ mamy

$$\mathbf{B} = -(\mathbf{1}_v : \dots : \mathbf{0})\mathbf{A}, \quad (3.17)$$

skąd, po podstawieniu do (3.16), wynika równość

$$\text{diag}_{i=1}^v (\mathbf{1}_{n_{1k}}) = [(\mathbf{0} : \dots : \mathbf{1}_{n_k} : \dots : \mathbf{0}) - (\mathbf{1}_{n_k} : \mathbf{0} : \dots : \mathbf{0})]\mathbf{A}. \quad (3.18)$$

W rezultacie warunek (3.13) implikuje relacje

$$\mathbf{R}\{\text{diag}_{i=1}^v (\mathbf{1}_{n_{1k}})\} \subset \mathbf{R}(\mathbf{1}_{n_k}), \quad (3.19)$$

która jest możliwa tylko wtedy, gdy k -ty blok jest zdegenerowany.

Analogicznie przebiega dowód w odniesieniu do modelu (3.6), w którym warunkiem koniecznym i dostatecznym istnienia estymatora najlepszego jest również układ relacji (3.13), lecz z macierzami \mathbf{V}_k postaci

$$\mathbf{V}_k = \text{diag}(\mathbf{0}, \dots, \mathbf{I}_{n_k} - \frac{1}{N_k} \mathbf{J}_{n_k}, \dots, \mathbf{0}).$$

Powtarzając przytoczone wyżej argumenty łatwo uzyskać równość

$$(\mathbf{I}_{n_k} - \frac{1}{N_k} \mathbf{J}_{n_k}) \text{diag}_{i=1}^v (\mathbf{1}_{n_{1k}}) = (-\mathbf{1}_{n_k} : \mathbf{0} : \dots : \mathbf{1}_{n_k} : \dots : \mathbf{0})\mathbf{A},$$

która ponownie implikuje relację (3.19). \square

Wykorzystując przedstawione tu rezultaty, a warunek (3.12) jednorodności bloków w szczególności, można stwierdzić, że eksperymenty ze stałymi blokami mogą mieć zastosowanie, jeśli grupowanie jednostek w bloki tak eliminuje ich zmienność, że w przyzmacie cechy badanej jednostki w blokach można uznać za identyczne, tj., gdy $\sigma_{u,k}^2 = 0$ dla $k = 1, 2, \dots, B$, lub jeśli grupowanie w bloki zapewnia co najmniej jednakową wariancję cechy badanej w blokach, tj., gdy $\sigma_{u,k}^2 = \sigma_*^2$ dla $k = 1, 2, \dots, B$ (porównaj Kempthorne, 1952, s.166). W tym drugim przypadku precyzja eksperymentu będzie oczywiście mniejsza, a sensowność bloków uzasadniona tylko wtedy, gdy oczekiwane stałe efekty blokowe będą się faktycznie różniły. Warto w tym miejscu zauważyć, że połączenie twierdzeń 1 oraz 2 prowadzi do wniosku, iż dla modeli doświadczzeń z blokami niezdegenerowanymi równość (3.12) jest warunkiem koniecznym i dostatecznym istnienia najlepszego liniowego estymatora wartości oczekiwanej wektora \mathbf{y} . Oznacza to, że jeżeli nie możemy zapewnić jednakowej wariancji w każdym bloku, to należy z układu o blokach stałych zrezygnować. W przeciwnym bowiem razie nie tylko komplikuje się analiza wyników eksperymentalnych, która uwzględniać musi model mieszany z wieloma komponentami wariancji, ale także, jeśli nie przede wszystkim, staje się niemożliwe wyznaczenie dla dowolnych estymowalnych funkcji efektów stałych liniowych estymatorów nieobciążonych, które miałyby minimalną wariancję.

4. RANDOMIZACJA WARSTW

Unifikacja jednostek eksperymentalnych w populacji rozwarstwionej uzyskiwana w wyniku randomizacji jednostek w warstwach prowadziła w konsekwencji do modeli, w których stałe efekty bloków pojawiają się w wartości oczekiwanej obserwowanego wektora y , a różne komponenty wariancji występują w kolejnych diagonalnych podblokach macierzy dyspersji wektora y . Z punktu widzenia estymacji najefektywniejszej doświadczenia odpowiadające tym modelom są przydatne tylko wtedy, gdy uda się zapewnić spełnienie warunku jednorodności bloków, tzn., gdy w strukturze macierzy dyspersji wektora y kolejne podbloki diagonalne faktycznie nie zależą od różnych komponentów wariancyjnych. Może się jednak okazać, że wspomniana własność jednorodności bloków nie jest możliwa do osiągnięcia. Można wtedy dążyć do uproszczenia struktury macierzy dyspersji wprowadzając dodatkową randomizację warstw.

Niech jednostki populacji podstawowej będą pogrupowane w B warstw $\{u_{1k}, u_{2k}, \dots, u_{nk}\}$, $k = 1, 2, \dots, B$, względem cech stanowiących podwektor x^2 wektora $x \in \mathbb{R}^q$ cech wyróżnionych. Wprowadzając teraz, podobnie jak w paragrafie I.5, zmienne losowe δ_{rk} określone na zbiorze wszystkich permutacji indeksów warstw, wektor reprezentujący (j) -tą zrandomizowaną jednostkę w (k) -tej zrandomizowanej warstwie możemy zapisać w postaci

$$x_{(j)k} = \sum_{r=1}^B \delta_{rk} x_{(j)r}, \quad (4.1)$$

gdzie $x_{(j)r}$ jest (j) -tą zrandomizowaną jednostką r -tej warstwy. Ponieważ zmienne losowe δ_{rk} i $x_{(j)r}$ są niezależne, więc korzystając z z (2.3) można łatwo ustalić, że

$$E(x_{(j)k}) = \frac{1}{B} \sum_{r=1}^B E(x_{(j)r}) = \begin{bmatrix} \bar{x}^1 \\ \bar{x}^2 \end{bmatrix}, \quad (4.2)$$

gdzie

$$\bar{x}^1 = \frac{1}{B} \sum_{r=1}^B \bar{x}_r^1 = \frac{1}{B} \sum_{r=1}^B \frac{1}{N_r} \sum_{j=1}^N x_{rj}^1, \quad (4.3)$$

$$\bar{x}^2 = \frac{1}{B} \sum_{r=1}^B x_r^2. \quad (4.4)$$

Podobnie, korzystając z własności zmiennych δ_{rk} , mamy

$$E(x_{(j)k} x'_{(j)k}) = \sum_{r=1}^B E(\delta_{rk} \delta_{rk}) E(x_{(j)r} x'_{(j)r}) + \quad (4.5)$$

$$+ \sum_{r=1}^B \sum_{r'=1}^B (\delta_{rk} \delta_{r'k}) E(x_{(j)r} x'_{(j)r'})$$

$$= \begin{cases} \frac{1}{B} \sum_{r=1}^B E(x_{(j)r} x'_{(j)r}), & k = k' \\ \frac{1}{B(B-1)} \sum_{r=1}^B \sum_{r'=1}^B E(x_{(j)r} x'_{(j)r'}), & k \neq k' \end{cases}$$

Na mocy (2.3), (2.5), (4.5) oraz (4.2) macierz dyspersji zmiennej $x_{(kj)}$ przyjmuje postać

$$D(x_{(kj)}) = \frac{1}{B} \sum_{r=1}^B D(x_{(j),k}) + \frac{1}{B} \sum_{r=1}^B E(x_{(j),r}) E(x'_{(j),r}) - E(x_{(j,k)}) E(x'_{(j,k)}) \quad (4.6)$$

$$= \begin{bmatrix} \frac{1}{B} \sum_{r=1}^B S_r^{11} & \mathbf{O} \\ \mathbf{O} & \mathbf{O} \end{bmatrix} + \begin{bmatrix} F^{11} & F^{12} \\ F^{21} & F^{22} \end{bmatrix},$$

gdzie

$$F^{11} = \frac{1}{B} \left\{ \sum_{r=1}^B \bar{x}_r^1 (\bar{x}_r^1)' - B \bar{x}^1 (\bar{x}^1)' \right\},$$

$$F^{12} = \frac{1}{B} \left\{ \sum_{r=1}^B \bar{x}_r^1 (\bar{x}_r^2)' - B \bar{x}^1 (\bar{x}^2)' \right\} = (F^{21})', \quad (4.7)$$

$$F^{22} = \frac{1}{B} \left\{ \sum_{r=1}^B \bar{x}_r^2 (\bar{x}_r^2)' - B \bar{x}^2 (\bar{x}^2)' \right\}.$$

Z kolei macierze kowariancji dwóch jednostek z tej samej zrandomizowanej warstwy oraz z dwóch różnych zrandomizowanych warstw są odpowiednio postaci

$$C(x_{(jk)}, x_{(j \cdot k)}) = \frac{1}{B} \sum_{r=1}^B C(x_{(j),r}, x_{(j \cdot),r}) + \quad (4.8)$$

$$+ \frac{1}{B} \sum_{r=1}^B E(x_{(j),r}) E(x'_{(j \cdot),r}) - E(x_{(j,k)}) E(x'_{(j \cdot k)}) =$$

$$= \begin{bmatrix} -\frac{1}{B} \sum_{r=1}^B \frac{1}{N_r - 1} S_r^{11} & \mathbf{O} \\ \mathbf{O} & \mathbf{O} \end{bmatrix} + F, \quad j \neq j \cdot,$$

oraz

$$C(x_{(jk)}, x_{(j \cdot k \cdot)}) = \frac{1}{B(B-1)} \sum_{r=1}^B \sum_{r \neq r \cdot} E(x_{(j),r}) E(x'_{(j \cdot),r \cdot}) + \quad (4.9)$$

$$- \frac{1}{B^2} \sum_{r=1}^B E(x_{(j),r}) \sum_{r \cdot=1}^B E(x'_{(j \cdot),r \cdot}) = \frac{-1}{B-1}, \quad k \neq k \cdot.$$

gdzie

$$= \frac{-1}{B-1} \begin{bmatrix} F^{11} & F^{12} \\ F^{21} & F^{22} \end{bmatrix},$$

przy czym podbloki macierzy określone są w (4.7).

Wyprowadzone parametry pozwalają stwierdzić, że jednostki eksperymentalne w populacji warstwowej ze zrandomizowanymi jednostkami i ze zrandomizowanymi warstwami mają własność symetrii (wymienialności) wewnątrz warstw. Modelowo wszystkie jednostki opisane są zmiennymi losowymi o wspólnej wartości oczekiwanej. Ponadto jednostki wewnątrz warstw mają jednakowe macierze dyspersji oraz jednakowe macierze kowariancji. Macierze kowariancji jednostek z różnych warstw, choć równe pomiędzy sobą, różnią się jednak od macierzy kowariancji jednostek

wewnątrz warstw. Tak więc własność symetrii przysługuje jedynie jednostkom wewnątrz warstw.

Uzyskane rezultaty dają się bezpośrednio przenieść na przypadek populacji z warstwami nieskończonymi. Jednostki uzyskane w procesie losowania wewnątrz warstw, które poddano randomizacji, można przedstawić za pomocą wektorów losowych. Oznaczmy je symbolem $x_{j(k)}$, gdzie (k) jest indeksem warstwy po randomizacji, a j jest indeksem jednostki wewnątrz warstwy. Parametry tych zmiennych łatwo wyznaczyć uwzględniając we wzorach już wyprowadzonych parametry określone w (2.8) i (2.9), a także biorąc pod uwagę fakt niezależności jednostek w warstwach. I tak, z równości (4.2), po uwzględnieniu (2.8), mamy wartość oczekiwaną

$$E(x_{j(k)}) = \begin{bmatrix} \bar{\mu}^1 \\ \bar{x}^2 \end{bmatrix}, \quad (4.10)$$

gdzie wektor \bar{x}^2 określony jest w (4.4), a

$$\bar{\mu}^1 = \frac{1}{B} \sum_{r=1}^B \mu_r^1. \quad (4.11)$$

Ze wzoru (4.6), po uwzględnieniu (2.8) i (2.9), mamy

$$D(x_{j(k)}) = \begin{bmatrix} \frac{1}{B} \sum_{r=1}^B \Sigma_r^{11} & \mathbf{0} \\ \mathbf{0} & \mathbf{0} \end{bmatrix} + \begin{bmatrix} \Phi^{11} & \Phi^{12} \\ \Phi^{21} & \Phi^{22} \end{bmatrix}, \quad (4.12)$$

gdzie

$$\Phi^{11} = \frac{1}{B} \{ \sum_{r=1}^B \mu_r^1 (\mu_r^1)' - B \bar{\mu}^1 (\bar{\mu}^1)' \},$$

$$\Phi^{12} = \frac{1}{B} \{ \sum_{r=1}^B \mu_r^1 (x_r^2)' - B \bar{\mu}^1 (\bar{x}^2)' \} = (\Phi^{21})', \quad (4.13)$$

$$\Phi^{22} = F^{22}.$$

I wreszcie ze wzorów (4.8) i (4.9) oraz równości $C(x_{j_r}, x_{j_r}) = \mathbf{0}$ wynika, że macierz kowariancji pomiędzy dowolnymi jednostkami tej samej wybranej losowo warstwy jest jednakowa i ma postać

$$C(x_{j(k)}, x_{j'(k)}) = \Phi, \quad j \neq j', \quad (4.14)$$

podczas, gdy macierz kowariancji pomiędzy jednostkami z dowolnych dwóch zrandomizowanych warstw ma postać

$$C(x_{j(k)}, x_{j'(k')}) = \frac{-1}{B-1} \Phi, \quad k \neq k', \quad (4.15)$$

gdzie podbloki macierzy Φ określone są w (4.13).

W rezultacie można stwierdzić, że w przypadku populacji z warstwami nieskończonymi, randomizacja warstw prowadzi do własności symetrii (wymienialności) jednostek podobnej do tej scharakteryzowanej w odniesieniu do populacji z warstwami skończonymi.

5. MODELE ZE ZRANDOMIZOWANYMI BLOKAMI

Przez modele ze zrandomizowanymi blokami będziemy rozumieć modele dostosowane do opisu doświadczeń, w których ustalone obiekty tworzące zbiór J_v są porównywane w oparciu o jednostki pochodzące z populacji warstwowej, przy czym uzyskane są one w wyniku losowania, lub randomizacji, wewnątrz warstw oraz randomizacji samych warstw.

Na początek rozważymy populację skończoną z B warstwami o liczebnościach odpowiednio N_1, N_2, \dots, N_B . Utworzenie próby jednostek eksperymentalnych stanowiących podmiot działania obiektów polegać będzie na zrandomizowanym wyborze b warstw, a następnie na pobraniu b prób zrandomizowanych $\mathcal{P}_{n_1}, \mathcal{P}_{n_2}, \dots, \mathcal{P}_{n_b}$ z każdej spośród warstw wybranych.

Z uwagi na wyprowadzoną w poprzednim paragrafie własność symetrii jednostek tworzących warstwę, obiekty będą mogły być skutecznie porównywane na jednostkach pochodzących z tej samej warstwy. Fakt ten winien być uwzględniony w planie eksperymentu, który dla zachowania możliwości nieskrępowanego wyboru warstw musi ponadto spełniać warunek

$$\max_k \{n_k\} \leq \min_k \{N_k\}. \quad (5.1)$$

Poza wymienionymi ograniczeniami, przyporządkowanie obiektów do poszczególnych jednostek, a dokładniej, do numerów jednostek w blokach, może przebiegać zupełnie dowolnie byle niezależnie od wyniku randomizacji.

Cechę badaną określimy standardowo za pomocą równości (3.1). Tym razem jednakże obydwie składniki reprezentują zmienne losowe z parametrami określonymi we wzorach (4.2), (4.6), (4.8) i (4.9). W rezultacie potencjalną obserwację uzyskaną z (jk) -tej zrandomizowanej jednostki eksperymentalnej, na której zastosowano s -ty obiekt, wyrazimy równością (3.4). Model wektorowy wszystkich obserwacji przyjmuje ostatecznie postać

$$\{y, m\}_n + \Delta' t, \frac{1}{B} \sum_{r=1}^B \sigma_{u,r}^2 \left(I_n - \frac{1}{N} D'D \right) + \sigma_f^2 \left(D'D - \frac{1}{B} J_n \right) + \sigma_e^2 I_n \}, \quad (5.2)$$

$$\mathcal{P}_{n_k} \subset \mathcal{P}_{N_k}, \quad k = 1, 2, \dots, b, \quad b \leq B,$$

gdzie

$$y = (y_{11}(1), y_{21}(1), \dots, y_{nb}(v))',$$

$$\Delta' = (\text{diag}_{i=1}^v (1'_{n_{1i}}) : \dots : \text{diag}_{i=1}^v (1'_{n_{bi}}))', \quad (5.3)$$

$$D' = \text{diag}_{k=1}^b (1_{n_k}), \quad (5.4)$$

$$t = (t_1, t_2, \dots, t_v)',$$

$$m = (\bar{x}^1)' a^1 + (\bar{x}^2)' a^2, \quad (5.5)$$

$$\sigma_f^2 = \frac{B}{B-1} a' F a, \quad (5.6)$$

przy czym podbloki macierzy F określone są w (4.7), a komponenty $\sigma_{u,r}^2$, $r = 1, 2, \dots, B$, w (3.3).

W przypadku populacji z warstwami nieskończonymi budowa modelu potencjalnych obserwacji przebiega podobnie, z tym, że jego struktura jest ustalona z wykorzystaniem wzorów (4.10), (4.12), (4.14) i (4.15). Wektorowa postać modelu wygląda ostatecznie następująco:

$$\{y, m1_n + \Delta't, (\sigma_u^2 + \sigma_r^2)I_n + \sigma_r^2(D'D - \frac{1}{B}J_n)\}, \quad (5.7)$$

$$P_{n_k} \subset P_{\infty}, \quad k = 1, 2, \dots, b, \quad b \leq B,$$

gdzie tym razem parametr wspólny ma postać

$$m = (\bar{m}^1)'a^1 + (\bar{x}^2)'a^2,$$

natomiasz komponenty σ_r^2 i σ_u^2 określone są wzorami

$$\sigma_r^2 = a' \bar{a}_r, \quad (5.8)$$

$$\sigma_u^2 = \frac{1}{B} \sum_{r=1}^B \sigma_{u,r}^2. \quad (5.9)$$

Porównując przedstawione tu modele z modelami (3.6) i odpowiednio (3.5), wyprowadzonymi w paragrafie 3, łatwo zauważyć, że mają one prostrzą strukturę wartości oczekiwanej oraz bardziej wyrównaną strukturę macierzy dyspersji, w której poszczególne podbloki diagonalne zależą od tych samych komponentów wariancji, mimo że komponentów tych jest więcej. W modelu (5.7) pojawił się komponent σ_r^2 , określony w (5.8) i wyrażający międzywarstwową wariancję cechy badanej, oraz komponent σ_u^2 będący średnią arytmetyczną wariancji określających zmienność wewnątrz poszczególnych warstw. Należy tu podkreślić fakt zależności parametru σ_u^2 od wszystkich warstwowych komponentów wariancji $\sigma_{u,r}^2$, $r = 1, 2, \dots, B$, a nie tylko od tych odpowiadających warstwom reprezentowanym w doświadczeniu.

W strukturze macierzy dyspersji modelu (5.2) poza komponentami σ_r^2 i σ_u^2 występują jeszcze wszystkie komponenty warstwowe $\sigma_{u,r}^2$, $r = 1, 2, \dots, B$. Każdy z nich pojawia się w modelu jako współczynnik poprzedzający macierz

$$I_n - \frac{1}{N} D'D \equiv L_r. \quad (5.10)$$

Wszystkie macierze L_r , $r = 1, 2, \dots, B$, są nieujemnie określone, przy czym własność ta jest w istocie konsekwencją warunku (5.1). Warto tu odnotować, że unifikacja macierzy L_r , prowadząca do redukcji liczby komponentów wariancyjnych w modelu (5.2), a tym samym do uproszczenia modelu, ma miejsce, gdy wszystkie warstwy są jednakowo liczne, czego jednakże nie należy utożsamiać z jednakową wielkością bloków.

Strukturę macierzy dyspersji modelu (5.2) można również przekształcić korzystając z następującej tożsamości:

$$\frac{1}{B} \sum_{r=1}^B \sigma_{u,r}^2 (\mathbf{I}_n - \frac{1}{N_r} \mathbf{D}'\mathbf{D}) = (\mathbf{I}_n - \frac{1}{N_H} \mathbf{D}'\mathbf{D}) \sigma_u^2,$$

gdzie komponent σ_u^2 określony jest w (5.9), a

$$N_H = B / \sum_{r=1}^B \frac{1}{N_r} (\sigma_{u,r}^2 / \sigma_u^2)$$

jest ważoną średnią harmoniczną liczebności warstw N_r , $r = 1, 2, \dots, B$. Model o tak przekształconej macierzy dyspersji, pozornie zależnej tylko od trzech komponentów σ_u^2 , σ_f^2 i σ_e^2 , pokrywa się z modelem wyprowadzonym przez Calińskiego i Kageyama (1988).

Nakładając na modele (5.2) i (5.7) dodatkowe warunki, można wyprowadzić z nich jeszcze inne modele znane w literaturze. I tak, przyjmując, że liczba warstw B rośnie nieograniczenie i równocześnie wariancje warstwowe $\sigma_{u,r}^2$, $r = 1, 2, \dots$, są ograniczone (skąd wynika, że $\sigma_u^2 < \infty$), model (5.7) redukuje się do modelu rozważanego przez Pattersona i Thompsona (1971). Ten sam efekt można uzyskać zakładając w modelu (5.2) nieograniczony wzrost liczby warstw B z równoczesnym ograniczeniem liczebności samych warstw N_r , $r = 1, 2, \dots$, co przy skończonych wariancjach $\sigma_{u,r}^2$, $r = 1, 2, \dots$, zapewnia zbieżność średniej harmonicznego N_H .

Warto tu jeszcze zaznaczyć, że przyjęcie założenia o nieograniczonym wzroście liczby warstw B oznacza w istocie, że mamy do czynienia z pewną ich nieskończoną populacją. W procesie zakładania doświadczenia będziemy zatem mieli do czynienia z próbą losową warstw, a to, w świetle twierdzenia 2.I, eliminuje konieczność ich randomizacji.

Nowa struktura macierzy dyspersji modeli (5.2) i (5.7) powoduje, że spełnienie warunku (3.12), jednorodności bloków, zmienia te modele tylko nieznacznie, nie decydując w szczególności o istnieniu najlepszego nieobciążonego estymatora liniowego dla wartości oczekiwanej wektora \mathbf{y} . Okazuje się, że istnienie takiego estymatora w przedstawionych tu modelach nie jest uzależnione od własności bloków, ale zależy od wybranego przez eksperymentatora planu doświadczenia. Plan taki będzie wygodnie scharakteryzować poprzez określenie własności tzw. macierzy incydencji \mathbf{N} . Macierz ta jest $v \times b$ -wymiarowa, a jej sk -ty element, n_{sk} , jest liczbą powtórzeń s -tego obiektu w k -tym bloku. Związek macierzy incydencji z macierzami już wyprowadzonymi jest bezpośredni i ma postać

$$\mathbf{N} = \Delta \mathbf{D}', \quad (5.11)$$

gdzie macierz Δ' określona jest w (5.3), a macierz \mathbf{D}' w (5.4).

Twierdzenie 3. W modelach (5.2) i (5.7) najlepszy nieobciążony estymator wartości oczekiwanej wektora \mathbf{y} istnieje wtedy i tylko wtedy, gdy macierz incydencji planu eksperymentalnego można przedstawić w postaci

$$\mathbf{N} = \text{diag}_{h=1}^g (n_{v_h} \mathbf{1}'_{b_h}), \quad (5.12)$$

gdzie g jest ustaloną liczbą naturalną, $g \geq 1$, $b_1 + b_2 + \dots + b_g = b$, $v_1 + v_2 + \dots + v_g = v$, a n_{v_h} jest v_h -wymiarowym wektorem o składowych naturalnych. Jeżeli warunek ten jest spełniony, to poszukiwanym estymatorem jest estymator otrzymany metodą najmniejszych kwadratów.

Dowód. Wykorzystując lemat II.B łatwo stwierdzić, że warunkiem koniecznym i dostatecznym istnienia estymatora najefektywniejszego w modelu (5.2) jest układ inkluzji

$$R\left\{\left(I_n - \frac{1}{N_r} D' D\right) \left(1_n : \Delta'\right)\right\} \subset R\left(1_n : \Delta'\right), \quad r = 1, 2, \dots, B, \quad (5.13)$$

oraz

$$R\left\{\left(D' D - \frac{1}{B} J_n\right) \left(1_n : \Delta'\right)\right\} \subset R\left(1_n : \Delta'\right). \quad (5.14)$$

W świetle relacji $1_n \in R(\Delta')$ i $R(\alpha A) = R(A)$, gdzie $\alpha \neq 0$, układ warunków (5.13) i (5.14) jest równoważny inkluzji

$$R(D' N') \subset R(\Delta'), \quad (5.15)$$

gdzie N jest macierzą incydencji, określoną w (5.11), a Δ' jest macierzą dla obiektów, zdefiniowaną w (5.3). Warunek (5.15) jest z kolei równoważny równości

$$\left\{I_n - \Delta' (\Delta \Delta')^{-1} \Delta\right\} D' N' = O,$$

która, wobec nieujemnej określoności macierzy ujętej w nawias klamrowy, może być zapisana równoważnie w postaci

$$H N' = O, \quad (5.16)$$

gdzie

$$H = D \left\{I_n - \Delta' (\Delta \Delta')^{-1} \Delta\right\} D' = D D' - N' (\Delta \Delta')^{-1} N.$$

Macierz H związana jest z tzw. dualnym układem blokowym. Jest ona stopnia b i rzędu co najwyżej $b-1$. Wiadomo ponadto, że jeśli rząd jej jest równy $b-g$, to macierz incydencji ma postać

$$N = \text{diag}_{h=1}^g N_h, \quad (5.17)$$

gdzie N_h jest $v_h \times b_h$ -wymiarową podmacierzą macierzy N , której odpowiada macierz

$$H_h = D_h D_h' - N_h' (\Delta_h \Delta_h')^{-1} N_h, \quad (5.18)$$

stopnia b oraz b_h rzędu b_h-1 , a D_h , D' i Δ' są podmacierzami odpowiednio macierzy D' i Δ' przyjmujących wówczas postaci $D' = \text{diag}_{h=1}^g D_h'$ oraz $\Delta' = \text{diag}_{h=1}^g \Delta_h'$, co pozwala przedstawić macierz H jako $H = \text{diag}_{h=1}^g H_h$. Zatem przyjęcie założenia, że rząd macierzy H jest $b-g$ ($g \geq 1$) pozwala w rezultacie warunek (5.16) zapisać jako układ warunków

$$H_h N_h' = O, \quad h = 1, 2, \dots, g, \quad (5.19)$$

gdzie macierze M_h określone są w (5.18).

Z drugiej strony łatwo sprawdzić, że $MI_b = 0$, skąd wynikają równości $M_h^{-1} 1_{b_h} = 0$ dla $h = 1, 2, \dots, g$. Ponieważ równocześnie rząd(M_h) = $b_h - 1$ dla $h = 1, 2, \dots, g$, więc ostatecznie warunki (5.19) zachodzą wtedy i tylko wtedy, gdy każda z macierzy M_h jest rzędu jeden i taka, że $R(N'_h) \subset R(1_{b_h})$, co kończy dowód twierdzenia w odniesieniu do modelu (5.2).

Dla zakończenia dowodu wystarczy zauważyć, że warunkiem koniecznym i dostatecznym istnienia estymatora najlepszego w modelu (5.7) jest inkluzja (5.14), równoważna ze spełnieniem relacji (5.15). \square

W uzupełnieniu udowodnionego twierdzenia wyrazimy jeszcze warunek (5.12) bardziej bezpośrednio. W istocie warunek ten oznacza, że wszystkie bloki w doświadczeniu można podzielić na g ($g \geq 1$) grup, przy czym

(a) poszczególne grupy związane są z rozłącznymi podzbiórami obiektów oraz

(b) w każdym bloku ustalonej grupy obiekty replikowane są tak samo.

W szczególności z zasady (b) wynika wniosek, że w każdej grupie bloki obejmują tę samą liczbę jednostek, a w przypadku, gdy układ blokowy jest tzw. układem binarnym, tzn., gdy w każdym bloku każdy obiekt może pojawić się co najwyżej raz, zasada (b) implikuje, że każdy z g podukładów blokowych ma macierz incydencji N_h złożoną z samych jedynek, tj. $N_h = 1_{v_h} 1'_{b_h}$, $h = 1, 2, \dots, g$.

Warto również zauważyć, że macierz incydencji postaci (5.12) spełnia równość $N = N(DD')^{-1}N'(\Delta\Delta')^{-1}N$, która jest warunkiem tzw. ortogonalności układu blokowego (patrz Chakrabarti, 1962, Darroch i Silvey, 1963, a także Caliński i Kageyama, 1988).

Ponadto zwróćmy jeszcze uwagę na to, że przedstawiony w twierdzeniu 3 warunek istnienia estymatora najlepszego w modelach (5.2) i (5.7) jest zbieżny z warunkiem sformułowanym przez Zmysłonego i Kageyamę (1987) w odniesieniu do modelu o uproszczonej macierzy dyspersji, który nietrudno można uzyskać z modeli tu rozważanych przyjmując, że bloki stanowią próbę prostą z nieskończonej populacji bloków ($B = \infty$).

6. MODELE ANALIZY WEWNĄTRZ I MIĘDZYBLOKOWEJ

Spośród modeli przedstawionych do tej pory tylko nieliczne mają szeroko rozwiniętą teorię obejmującą nie tylko estymację nieznanych parametrów, ale także pełną ich analizę. Należą do nich model (II.4.2), ze stałymi parametrami obiektowymi, oraz model (II.6.1), z losowymi parametrami obiektowymi, obydwa odnoszące się do populacji nieskończonej. Wśród modeli blokowych bogatą teorię ma model (3.5) ze spełnionym warunkiem (3.12) jednorodności bloków. W podejściu tu prezentowanym odpowiada on sytuacji, w której populacja podstawowa \mathcal{P} jest traktowana jako zbiór b warstw nieskończonych i jednorodnych względem cechy badanej,

przy czym z każdej warstwy pobrano próbę losową, stanowiącą blok, na którego jednostkach rozmieszczono ustalone obiekty zbioru γ . Uwzględniając warunek jednorodności bloków $\sigma_{u,1}^2 = \sigma_{u,2}^2 = \dots = \sigma_{u,b}^2 = \sigma_*^2$, model ten przyjmuje postać

$$\{y, D'm + A't, \sigma^2 I_n\}, \quad (6.1)'$$

gdzie

$$D' = \text{diag}_{k=1}^b (1_{n_k}), \quad (6.2)$$

$$A' = (\text{diag}_{i=1}^{n_1} (1'_{n_1}) : \dots : \text{diag}_{i=1}^{n_b} (1'_{n_b}))', \quad (6.3)$$

m i t są odpowiednio wektorami efektów blokowych i obiektowych, a

$$\sigma^2 = \sigma_*^2 + \sigma_e^2$$

jest sumą wariancji jednostek i wariancji błędu technicznego.

Model (6.1) odgrywa kluczową rolę w teorii układów blokowych (patrz np. Kempthorne, 1952, s.151, Scheffe, 1959, s.163, Pearce, 1983, s.57) i jako taki przyjmowany jest powszechnie za podstawę analizy danych eksperymentalnych. Jak już powiedzieliśmy, wymaga on jednak spełnienia warunku jednorodności bloków, który daje znaczne uproszczenia teoretyczne, lecz w praktyce eksperymentalnej może być trudny do przyjęcia (patrz np. Kempthorne, 1952, s.166).

Wolnymi od tego założenia są, wyprowadzone w poprzednim paragrafie, modele ze zrandomizowanymi blokami. Wyeliminowanie warunku jednorodności bloków nie usuwa jednakże wszystkich trudności. Ich skalę w zakresie estymacji obrazuje twierdzenie 3, z którego w szczególności wynika, że bez spełnienia, przez plan eksperymentalny, dość restryktywnych warunków nie można się spodziewać istnienia najlepszych liniowych nieobciążonych estymatorów dla dowolnych kontrastów obiektowych. Poszukując w tej sytuacji rozwiązania kompromisowego dokonuje się pewnych transformacji modelu wyjściowego. Stanowią one podstawę dla tzw. analizy wewnątrz blokowej i analizy międzyblokowej. Pojęcia te funkcjonują w teorii eksperymentalnych układów blokowych od bardzo dawna, lecz w odniesieniu do modeli znacznie uproszczonych (porównaj dyskusję w pracy Hautmana i Speeda, 1983, s.1072).

Transformacje, o których wspomniano wyżej polegają na liniowym przekształceniu obserwowanego wektora losowego y tak, aby całkowitą informację w nim zawartą rozdzielić na część mieszczącą się wewnątrz bloków i część mieszczącą się pomiędzy blokami. Algebraicznie proces ten można wyrazić tożsamością

$$y = (I_n - P_D)y + P_D y,$$

gdzie $P_D = D'(DD')^{-1}D$ jest operatorem rzutu ortogonalnego na podprzestrzeń rozpiętą na kolumnach macierzy D' . Ponieważ iloczyn Dy jest wektorem sum blokowych, a DD' jest macierzą diagonalną, której kolejne

niezerowe elementy są wielkościami bloków, tj. $DD' = \text{diag}_{k=1}^b (n_k)$, więc transformacja $P_D y$ polega w istocie na zastąpieniu kolejnych składowych w wektorze y odpowiednimi średnimi obiektowymi, natomiast transformacja $(I_n - P_D) y$ polega na odjęciu od każdej składowej wektora y odpowiedniej średniej blokowej. Pierwsza z tych transformacji, ignorująca zmienność wewnątrz bloków, prowadzi do analizy międzyblokowej, a druga, eliminująca średnie blokowe, prowadzi do analizy wewnątrzblokowej.

W celu uzyskania odpowiednich modeli dla analizy wewnątrz i międzyblokowej wystarczy zastosować opisane transformacje do modeli wcześniej wyprowadzonych pamiętając, że operator P_D , spełnia równości $P_D D' = D'$ oraz $P_D 1_n = 1_n$.

Model analizy wewnątrzblokowej wygodnie będzie zapisać korzystając z następującego oznaczenia

$$\phi = I - P_D, \quad (6.4)$$

które stosowane jest w teorii układów blokowych (patrz np. Pearce, 1983, s.59). Użycie tej transformacji sprowadza model (5.2) do postaci

$$\{\phi y, \phi \Delta' t, \bar{\sigma}^2 \phi\}, \quad (6.5)$$

gdzie wariancja $\bar{\sigma}^2 = \sigma_u^2 + \sigma_e^2$.

Zauważmy, że jeśli przekształcenie (6.4) zastosujemy do modelu (5.7), to również otrzymamy model (6.5). Oznacza to, że postać modelu dla analizy wewnątrzblokowej nie zależy od liczebności warstw w populacji podstawowej.

Wobec lematu II.A oraz idempotentności macierzy ϕ łatwo sprawdzić, że w modelu (6.5) możliwe jest wyznaczenie najlepszego estymatora nieobciążonego funkcji $\phi \Delta' t$ za pośrednictwem metody najmniejszych kwadratów. Odpowiedni układ równań normalnych przyjmuje wtedy postać

$$\Delta \phi \Delta' t = \Delta \phi y, \quad (6.6)$$

gdzie prawa strona stanowi w istocie poszukiwane estymatory dla funkcji $\Delta \phi \Delta' t$. Ponieważ

$$\Delta \phi \Delta' 1_v = \Delta \phi 1_n = 0, \quad (6.7)$$

więc funkcje te są faktycznie kontrastami efektów obiektowych, przy czym liczba liniowo niezależnych kontrastów możliwych tu do wyestymowania jest równa rzędowi macierzy $C = \Delta \phi \Delta'$. Z uwagi na własność (6.7) rząd macierzy C nie przekracza $v-1$, co stanowi maksymalną liczbę liniowo niezależnych kontrastów możliwych do ustalenia pomiędzy efektami v obiektów. Widać zatem, że analiza wewnątrzblokowa może dostarczyć estymatorów dla pewnej liczby kontrastów obiektowych, a niekiedy nawet dla wszystkich porównań między nimi. Należy jednak pamiętać, że estymatory otrzymane z tej analizy mimo, że charakteryzują się własnością nieobciążoności i minimalności wariancji w modelu (6.5), to w ogólności nie posiadają obu tych własności ani w modelu (5.2) ani w modelu (5.7). Wreszcie należy tu zaznaczyć, że

równanie (6.6) pojawia się także w trakcie analizy modelu (6.1). Wtedy jest ono tzw. zredukowanym układem równań normalnych (porównaj np. Houtman i Speed, 1983, s.1072).

Modele analizy międzyblokowej otrzymamy dokonując transformacji za pomocą operatora P_D , lub równoważnie za pomocą macierzy D , która wektor y redukuje do wektora sum obiektowych Dy . Druga z tych możliwości, z uwagi na rozmiary wektorów i macierzy, jest wygodniejsza i sprowadza modele (5.2) i (5.7) do postaci,

$$\{Dy, N't_*, \frac{1}{B} \sum_{r=1}^B \sigma_{u,r}^2 [DD' - \frac{1}{N} (DD')^2] + \sigma_f^2 [(DD')^2 - \frac{1}{B} DD'] + \sigma_o^2 DD'\}, \quad (6.8)$$

$$P_{n_k} \subset P_{N_k}, \quad k = 1, 2, \dots, b, \quad b \leq B,$$

$$\{Dy, N't_*, \sigma_f^2 [(DD')^2 - \frac{1}{B} (DD')^2] + \bar{\sigma}^2 DD'\}, \quad (6.9)$$

$$P_{n_k} \subset P_{\infty}, \quad k = 1, 2, \dots, b, \quad b \leq B,$$

gdzie N jest macierzą incydencji, a $t_* = mn + t$ jest wektorem nowych parametrów, przy czym n jest wektorem wielkości bloków.

Tym razem uzyskane modele nie są identyczne, a ich przydatność w zakresie estymacji jest uzależniona od możliwości skonstruowania estymatorów kontrastów obiektowych, które miałyby optymalne własności. Niekiedy taka możliwość istnieje stając się źródłem dodatkowej informacji, która może być spożytkowana dla poprawienia estymatorów uzyskanych z analizy wewnątrzblokowej.

Twierdzenie 4. W modelach (6.8) i (6.9) analizy międzyblokowej istnieje najlepszy liniowy nieobciążony estymator dla wartości oczekiwanej wektora Dy wtedy i tylko wtedy, gdy

$$R(DD'N') \subset R(N'). \quad (6.10)$$

Jeśli warunek ten jest spełniony, to poszukiwanym estymatorem jest estymator otrzymany metodą najmniejszych kwadratów.

Dowód. Na początek zauważmy, że macierz DD' jest nieosobliwa. Zatem bez straty możliwości wyznaczenia estymatora najlepszego modele (6.8) i (6.9) można przekształcić za pomocą macierzy $(DD')^{-1/2}$ (patrz Baksalary i Kala, 1980, s.914). W wyniku tego przekształcenia w strukturze macierzy dyspersji pojawi się macierz jednostkowa - w modelu (6.8) przy komponencie σ_o^2 , a w modelu (6.9) przy komponencie $\bar{\sigma}^2$. Stosując teraz do tak przekształconych modeli lemat II.B łatwo otrzymać wniosek, że warunkiem koniecznym i dostatecznym istnienia najlepszego nieobciążonego liniowego estymatora wartości oczekiwanej $E\{(DD')^{-1/2} D'y\} = (DD')^{-1/2} N't_*$ jest

$$R\{(DD')(DD')^{-1/2} N'\} \subset R\{(DD')^{-1/2} N'\},$$

który jest równoważny inkluzji (6.10). Pozostała część twierdzenia wynika z lematu II.B bezpośrednio. \square

Nie trudno zauważyć, że warunek (6.10) jest spełniony np. wtedy, gdy $DD' = n_0 I_b$, tzn. gdy układ blokowy jest właściwy, lub wtedy, gdy układ blokowy jest złożeniem kilku podukładów właściwych, każdy dotyczący innego podzbioru obiektów. W tej drugiej sytuacji układ blokowy jest oczywiście układem niespójnym. Warto równocześnie zaznaczyć, że żaden z warunków tu opisanych nie jest konieczny. Nie mniej, z powyższych spostrzeżeń wynika, że warto w procesie planowania doświadczenia dążyć do jednakowej wielkości bloków.

Jeśli warunek (6.10) jest spełniony, to odpowiedni układ równań normalnych przyjmuje postać

$$N(DD')^{-1}N't_0 = N(DD')^{-1}y, \quad (6.11)$$

dostarczając poszukiwanych estymatorów dla liniowych funkcji $c't_0$. W przypadku, gdy $DD' = n_0 I_b$, równanie (6.11) redukuje się do postaci

$$NN't_0 = Ny. \quad (6.12)$$

Zarówno równanie (6.11) jak i równanie (6.12) pojawiają się w teorii układów blokowych (patrz Pearce i inni, 1974, s.454 lub Pearce, 1981, s.79), lecz ich wyprowadzenie związane jest z modelami uzyskanymi w oparciu o inne założenia.

W zakończeniu pragnę wyrazić podziękowanie prof. drowi hab. Tadeuszowi Calińskiemu za uważne przestudiowanie wcześniejszych wersji pracy oraz wnikliwe uwagi, które przyczyniły się do istotnego poprawienia ostatecznego kształtu przedstawionej teorii randomizacji.

LITERATURA

- Baksalary J.K., Kala R. (1981). Linear transformations preserving best linear unbiased estimators in a general Gauss-Markoff model. *Ann. Statist.* **4**, 913-916.
- Caliński T., Kageyama S. (1988). A randomization theory of intrablock and interblock estimation. Technical report No.230, Statistical Research Group, Hiroshima University, Hiroshima, Japan.
- Chacrabarti M.C. (1962). *Mathematics of Design and Analysis of Experiments*. Asia Publishing House, Bombay.
- Darroch J.N., Silvey S.D. (1963). On testing more than one hypothesis. *Ann. Math. Statist.* **34**, 555-567.
- Harville D.A. (1975). Experimental randomisation: Who needs it? *The Amer. Statist.* **29**, 27-31.
- Houtman A.M., Speed T.P. (1983). Balance in designed experiments with orthogonal block structure. *Ann. Statist.* **11**, 1069-1085.
- Kala R. (1989). Elementy teorii randomizacji. I. Próba zrandomizowana. *Listy Biometryczne - Biometrical Letters* **26**, 41-55.

- Kala R. (1990). Elementy teorii randomizacji. II. Modelowanie doświadczeń prostych. *Listy Biometryczne - Biometrical Letters* 27, 31-45.
- Kempthorne O. (1952). *The Design and Analysis of Experiments*. J. Wiley, New York.
- Patterson H.D., Thompson R. (1971). Recovery of inter-block information when block sizes are unequal. *Biometrika* 58, 545-554.
- Pearce S.C. (1983). *The Agricultural Field Experiment. A Statistical Examination of Theory and Practice*. J. Wiley, Chichester.
- Pearce S.C., Caliński, T., Marshall, T.F.de C. (1974). The basic contrasts of an experimental design with special reference to the analysis of data. *Biometrika* 61, 449-460.
- Scheffe H. (1959). *The Analysis of Variance*. J. Wiley, New York.
- Zmyślony R., Kageyama S. (1987). Estimation of parameters in block designs under mixed models. Preprint 379 Inst. Matem. PAN.

Praca wpłynęła 1 sierpnia 1988;
 w wersji ostatecznej 26 października 1990

ELEMENTS OF THE RANDOMIZATION THEORY.

III. RANDOMIZATION IN BLOCK EXPERIMENTS

Summary

1. INTRODUCTION

The aim of this paper is to show the influence of randomization on the structure of the model of block experiments. Two kinds of randomization are considered, the randomization of units inside blocks, and randomization of blocks. In both models the existence of best linear unbiased estimators of treatment parameters is considered. The discussion comprises also models for the intra-block and the inter-block analysis.

Key words: additivity, homogeneity, null experiments, fixed models, random models, mixed models

Here we consider a situation in which p inbred lines are chosen and certain crosses are made among them. Griffing (1956) discussed four types of diallel crossing systems. In types I and II the p parental lines are also included; in types III and IV they are omitted. In types I and III the reciprocal crosses are included, i.e. both, the cross in which the i -th line is the male line and the j -th line is the female line and the reciprocal cross with the j -th as the male line and the i -th as the female line. In types II and IV the reciprocal crosses are omitted.

Griffing (1956) gave the analysis of the above four types of diallel