ZESZYTY PROBLEMOWE POSTĘPÓW NAUK ROLNICZYCH 1989, z. 382

ANALIZA ZDOLNOŚCI KOMBINACYJNEJ I EFEKTÓW KRZYŻOWAŃ PRZECIWNYCH W NIEKOMPLETNYM UKŁADZIE CZYNNIKOWYM

Wiesław Mądry

Katedra Statystyki Matematycznej i Doświadczalnictwa SGGW-AR w Warszawie

W pracy Mądrego i Piotrowskiego [7] przedstawiono zasady analizy statystycznej pewnego zmodyfikowanego układu diallelicznego, w którym ocenia się mieszańce pomiędzy dwiema grupami form rodzicielskich oraz mieszańce wsobne dla form rodzicielskich z obu grup. Przedstawiony tam algorytm dotyczy przypadku niekompletnego, tzn. zakłada się, że występują braki niektórych kombinacji mieszańcowych w stosunku do potencjalnej /docelowej/ formuły tego układu kojarzenia wg Chuang-Sheng Lin [2].

Można zauważyć, że kiedy pominiemy mieszańce wsobne dla form rodzicielskich z obu grup uzyskujemy czynnikowy układ krzyżowania, zawierający obok krzyżowań prostych przeciwne. Taki wzbogacony o mieszańce przeciwne układ czynnikowy daje dodatkowo możliwości oceny efektów dziedzicznych przekazywanych przez cytoplazmę. W analizie statystycznej można oprzeć się w dalszym ciągu na zmodyfikowanym modelu Griffinga [5] podobnie jak w typowym układzie Chuang-Sheng Lin [2].

Celem niniejszej pracy jest przedstawienie opracowanego algorytmu, dostosowanego do analizy niekompletnego układu czynnikowego, zawierającego mieszańce przeciwne. Jako punkt wyjścia traktuje się algorytm i jego uzasadnienie teoretyczne zawarte w pracy Mądrego i Piotrowskiego [7].

MODEL STATYSTYCZNY I JEGO ZAŁOŻENIA

Rozpatrzmy niekompletny układ czynnikowy z krzyżowaniami przeciwnymi, w którym formy rodzicielskie, pochodzące z pierwszej grupy oznaczamy numerami i ze zbioru $A = \{1,2,\ldots,a\}$ zaś formy rodzicielskie z drugiej grupy numerami ze zbioru $A = \{a+1, a+2, \ldots,a+b\}$. Załóżmy, że jest to układ spójny [4]. Pewnym przykładem mogą być krzyżowania dwóch grup odmian europejskich i amerykańskich pszenicy ozimej /tab. 1/.

Tabela 1

Plan kojarzeń dwóch grup odmian pszenic ozimych amerykańskich i europejskich w niekompletnym układzie czynnikowym z mieszańcami przeciwnymi

Matki	Ojcowie					
	pierwsza grupa			druga grupa		
	Scaut	Sturdy	Omacha	WS 1003	WWRN	Norin
Scaut				X	X	
Sturdy					X	X
Omacha				X	X	X
WS 1003	X	X				
WWRN		X				
Norin	X	X	X			

Dla średnich arytmetycznych otrzymanych w eksperymentalnym układzie całkowicie losowym lub w układzie z m powtórzeniami mieszańca pomiędzy i'-tą formą rodzicielską mateczną oraz j'-tą formą rodzicielską ojcowską przyjmujemy następujący model analizy wariancji:

$$y_{i'j'} = m + a_i + b_j + s_{ij} + r_{ij} + \bar{e}_{ij} \qquad /1/$$

$$i'j' = (ij) \in A\times B \quad \text{lub} \quad i'j': (ji) \in B\times A$$

Definicje efektów w modelu (1) oraz warunki ważone, które nałożono na nie są takie same jak w pracy Mądrego i Piotrowskiego [7].

Model (1) może być przedstawiony w zapisie macierzowym jako:

$$\underset{n\times 1}{y} = \underset{n\times 1}{Nm} + \underset{n\times a}{A_1}\underset{a\times 1}{a} + \underset{n\times b}{A_2}\underset{b\times 1}{b} + \underset{n\times M}{B}\underset{M\times 1}{s} + \underset{n\times n}{C}\underset{n\times 1}{r} + \underset{n\times 1}{\bar{e}}, \quad \bar{e} \sim N_n(0, \sigma_e^2 I), \qquad /2/$$

gdzie

n - liczba wszystkich mieszańców,

M - liczba par rodzicielskich.

Kolumny macierzy układu $X = [N\ A_1\ A_2\ B\ C]$ rozpinają n-wymiarową przestrzeń estymacji. Podprzestrzenie N, A_1, A_2, B i C spełniają warunki zawierania [7]:

$$N \subset A_1 \subset B \subset C$$
$$N \subset A_2 \subset B \subset C$$

Estymacja parametrów modelu (1)

Oceny parametrów modelu (1) z warunkami ważonymi uzyskane metodą najmniejszych kwadratów muszą być wyznaczane z powodu nieortogalności podprzestrzeni $A_1 - N$ i $A_2 - N$ dla układów niekompletnych w następujących podprzestrzeniach wzajemnie ortogonalnych:

$$N, \langle A_1\ A_2\rangle - N,\ B - \langle A_1\ A_2\rangle \text{ i } C - B.$$

Estymowane parametry tworzą następujące rzuty ortogonalne na wymienione podprzestrzenie /dopełnienia ortogonalne/ wg Keuls i Garretsen, Nawrockiego, Calińskiego i Mikosa [1, 3, 6, 8-10]:

$$y_N = N\hat{m},$$

$$y_{\langle A_1A_2\rangle -N} = y_{\langle A_1A_2\rangle} - y_N = (A_1A_2)\begin{pmatrix}\hat{a}^*\\ \hat{b}^*\end{pmatrix} - N\hat{m} = A_1\hat{a} + A_2\hat{b} \qquad /3/$$

$$y_{B-\langle A_1A_2\rangle} = y_B - y_{\langle A_1A_2\rangle} = B\hat{s}^* - (A_1\hat{a}^* + A_2\hat{b}^*) = B\hat{s},$$

$$y_{C-B} = y_C - y_B = C\hat{r}^* - B\hat{s}^* = C\hat{r}.$$

Wszystkie parametry z gwiazdką wyznaczamy z następujących równań:

$$(N^T N)\hat{m} = N^T y,$$
$$(A_1\ A_2)^T (A_1\ A_2)\begin{pmatrix}\hat{a}^*\\ \hat{b}^*\end{pmatrix} = (A_1\ A_2)^T y,$$
$$(C^T C)\hat{r}^* = C^T y.$$

Spośród macierzy układów równań (4) jedna tylko macierz $(A_1A_2)^T(A_1A_2)$ jest osobliwa. Jej rząd wynosi a + b - 1; jest więc o jeden mniejszy niż wymiar tej macierzy. W kategoriach przestrzeni wektorowych oznacza to, że suma wektorów podprzestrzeni A_1 jest równa sumie wektorów podprzestrzeni A_2. Możemy to wyrazić formułą $(A_1A_2)\,t = 0$, gdzie $t' = (\underbrace{1\,1\,\ldots\,1}_{a},\underbrace{-1\,-1\,\ldots\,-1}_{b})$. Ponieważ rząd macierzy układu $(A_1A_2)^T(A_1A_2)$ jest o jeden mniejszy niż liczba elementów wektora parametrów $\binom{\hat{a}^*}{\hat{a}^*}$, zatem w celu rozwiązania omawianego układu równań należy uwzględnić jedną restrykcję nałożoną na wektor parametrów w postaci $\sum_i k_i\,\hat{b}_i^* = 0$. Wobec tego $\hat{b}^*$ jest natychmiast oceną $\hat{b}$. Przyjętą restrykcję wyrażamy w postaci $d'\binom{\hat{a}^*}{\hat{b}^*} = 0$, gdzie $d' = (0,\ldots,0,k_1,k_2,\ldots,k_{a+b})$ zaś współczynniki k_j określają liczbę mieszańców z j-tą formą rodzicielską.

W konsekwencji otrzymujemy:

$$\binom{\hat{a}^*}{\hat{b}^*} = [(A_1A_2)^T\,(A_1A_2) + dd']^{-1}\,(A_1A_2)^T y,$$ gdzie macierz $[(A_1A_2)^T(A_1A_2) + dd']^{-1}$ /oznaczamy ją symbolem W/ jest uogólnioną macierzą odwrotną do macierzy $(A_1A_2)^T(A_1A_2)$ wg Scheffego [12] i Rao [11].

Ostatecznie otrzymujemy następujące formuły na oceny nie obciążone efektów modelu (1) z warunkami ważonymi:

$$\begin{aligned}
\hat{m} &= \Sigma y_{ij}/n \\
\hat{a}_i &= \hat{a}_i^* - \hat{m}, \\
\hat{b}_j &= \hat{b}_j^*, \\
\hat{s}_{ij} &= \hat{s}_{ij}^* - \hat{a}_i^* - \hat{b}_j^* \\
\hat{r}_{ij} &= y_{ij} - \hat{s}_{ij}^*
\end{aligned} \qquad /5/$$

Błędy standardowe uzyskanych estymatorów modelu (5) są następujące:

$$\sigma_{\hat{m}} = \frac{\sigma_e^2}{n}\,; \quad \sigma_{(\hat{a}_i - \hat{a}_{i'})} = \sqrt{\mathrm{var}\,(\hat{a}_i^* - \hat{a}_{i'}^*)}\,;$$

$$\sigma\hat{a}_i = \sqrt{\mathrm{var}\,\hat{a}_i^* - \mathrm{var}\,\hat{m}}\,; \quad \sigma_{(\hat{b}_j - \hat{b}_{j'})} = \sqrt{\mathrm{var}\,(\hat{b}_j^* - \hat{b}_{j'}^*)}\,;$$

$$\sigma_{\hat{b}_j} = \sqrt{\mathrm{var}\ \hat{b}_j^*}\ ; \quad \sigma_{\hat{r}_{ij}} = \frac{\sigma_e^2}{2}\ , \quad \sigma_{\hat{s}_{ij}} = \sqrt{\mathrm{var}\ \bar{y}_{ij} - \mathrm{var}(\hat{a}_i^* + \hat{b}_j^*)}\ ,$$

gdzie:

$$\mathrm{var}\ (\hat{a}_i^* - \hat{a}_{i'}^*) = (w_{ii} - 2w_{ii'} + w_{i'i'})\ \sigma_e^2\ ;$$

$$\mathrm{var}\ \hat{a}_i^* = \left[w - \left(\frac{1}{d't}\right)^2 t\ t'_{ii}\right] \sigma_e^2\ ;$$

$$\mathrm{var}\ (\hat{b}_j^* - \hat{b}_{j'}^*) = (w_{jj} - 2w_{jj'} + w_{j'j'})\ \sigma_e^2\ ; \qquad /6/$$

$$\mathrm{var}\ \hat{b}_j^* = \left[w - \left(\frac{1}{d't}\right)^2 tt'_{jj}\right] \sigma_e^2\ ;$$

$$\mathrm{var}\ (\hat{a}_i^* + \hat{b}_j^*) = (w_{ii} + 2w_{ij} + w_{jj})\, \sigma_e^2\ ;$$

$\mathrm{var}\ \hat{m} = (1/n)\ \sigma_e^2$ (σ_e^2 jest wariancją błędu losowego średniej arytmetycznej z m obserwacji)

$$\mathrm{var}\ \bar{y}_{ij} = \frac{1}{2}\ \sigma_e^2\ .$$

Tabela 2

Analiza wariancji według modelu (1) dla niekompletnego układu czynnikowego z mieszańcami przeciwnymi

Źródło zmienności	Stopień swobody	Suma kwadratów	Średni kwadrat
GCA rodziców pierwszej grupy	$a - 1$	$y^2_{\langle A,B\rangle} - y^2_B$	s^2_A
GCA rodziców drugiej grupy	$b - 1$	$y^2_{\langle A,B\rangle} - y^2_A$	s^2_B
SCA	$M - a - b + 1$	$y^2_C - y^2_{\langle A,B\rangle}$	s^2_S
Efekty przeciwne	$n - M$	$y^2 - y^2_C$	s^2_r
Błąd losowy	$v_{\bar{e}}$	-	s^2_e

$s^2_{\bar{e}} = \dfrac{s^2_e}{m}$; s^2_e średni kwadrat dla błędu we wstępnej analizie wariancji wykonanej na danych pochodzących z pojedynczych poletek.

Analiza wariancji według modelu (1) jest przedstawiona w tabeli 2. Odpowiednie kwadraty rzutów ortogonalnych znajdziemy według wzorów:

$$y^2 = \sum_{ij} y_{ij}^2, \quad y_A^2 = \sum_i \frac{Y_{i.}^2}{k_i}, \quad y_B^2 = \sum_j \frac{Y_{.j}^2}{k_j} \; Y_{i.}, \quad Y_{.j},$$

gdzie:

sumy danych w tabeli diallelicznej dla i-tych form rodzicielskich z pierwszej grupy lub j-tych form rodzicielskich z drugiej grupy, k_i oraz k_j liczb mieszańców z i-tą lub j-tą formą rodzicielską

$$y^2_{<A,B>} = \hat{a}^{*\prime}A'y + \hat{b}^{*\prime}B'y, \quad y_C^2 = \hat{s}^{*\prime}X'y, \quad y_N^2 = \frac{Y_{..}^2}{n},$$

gdzie $Y_{..}$ jest sumą ogólną danych w tabeli diallelicznej.

LITERATURA

1. Caliński T.: On non-orthogonal experiments, Biom. Zeit., 14, 73-84, 1972.
2. Chuang-Sheng Lin: Analysis of diallel crosses between two groups where parental lines are included. Biometrics, 28, 612--618, 1972.
3. Garretsen F., Keuls M.: A general method for the analysis of genetic variation in complete and incomplete diallels and North Carolina II designs. Part II. Procedures and general formulas for the fixed model. Euphytica, 27, 49-68, 1978.
4. Gilbert N.: Additive combining abilities fitted to plant breeding data. Biometrics, 23, 45-49, 1967.
5. Griffing B.: Concept of general and specific combining ability in relation to diallel crossing systems. Aust. J. Biol. Sci., 9, 463-493, 1956.
6. Keuls M., Garretsen F.: A general method for the analysis of genetic variation in complete and incomplete diallels and North Carolina II designs. P. II. Procedures and general formulas for the random model. Euphytica, 26, 537-551, 1977.
7. Mądry W., Piotrowski W.: Analysis of the combining ability of two parental groups in incomplete diallel crosses. J. Appl. Stat., 13, 39-48, 1986.

8. Mikos H.: Operatory rzutowe w analizie wariancji, III Coll. Metodol. Agrobiom., 78-142, 1973.
9. Nawrocki Z.: Teoria i praktyka doświadczenia rolniczego.PWRiL Warszawa, 1967.
10. Nawrocki Z.: Zastosowanie operatorów rzutowych do doświadczalnictwa rolniczego. SGGW-AR, Warszawa 1972.
11. Rao C.R.: Modele liniowe statystyki matematycznej. PWN, Warszawa 1980.
12. Scheffe H.: The analysis of variance. Willy and Sons,New York 1959.

W. Mądry

ANALYSIS OF THE COMBINING ABILITY AND RECIPROCAL CROSSING RESULTS PRESENTED IN AN INCOMPLETE FACTORIAL MATING DESIGN

S u m m a r y

The method of univeriate analysis of variance for an incomplete factorial mating /North Caroline II/ design with reciprocal crossings is described in the paper. The mating design provides testing offspring in an experimental design of random blocks or complete random design with the constant number or replications. The linear model joining in itself properties of the Griffing's model /1956/ with the model for classical cross classification has been assumed in the analysis.

Formulae for estimation of the model parameters and their standard errors as well as formulae for sums of deviation squares and degrees of freedom in the table of analysis of variance have been derived. Some properties and theorems of vector spaces have been made use of.

В. Мондры

АНАЛИЗ КОМБИНАЦИОННОЙ СПОСОБНОСТИ И ЭФФЕКТОВ ОБРАТНОГО СКРЕЩИВАНИЯ В НЕКОМПЛЕТНОЙ ФАКТОРНОЙ СХЕМЕ

Резюме

В статье представлен однопризнаковый дисперсионный анализ для некомплектной факторной схемы с обратными скрещиваниями. Рассматриваемая схема спаривания предусметривает испытание потомства в экспериментальной схеме случайных блоков, или в полностью случайной схеме с постоянным количеством повторений. В анализе принята линейная модель соединяющая в себе признаки модели Гриффинга /1956/ и модели для классической перекрестной классификации.

Были определены формулы как определители параметров модели и их стандартных погрешностей, а также формулы на суммы квадратов отклонений и степени свободы в таблице дисперсионного анализа. Использовали некоторые свойства и теоремы векторных пространств.