

Z Katedry Statystyki Matematycznej Wydziału Rolniczego WSR w Lublinie
Kierownik: doc. dr Wiktor Oktaba

Tadeusz PRZYBYSZ

O niejednorodnym błędzie w analizie wariancji dla kwadratu
łacińskiego w modelu z pojedynczo rozszczepionymi poletkami

Об неоднородной ошибке в дисперсионном анализе латинского
квадрата с распределенными деланками

On the Heterogeneous Error of Split Plot Design in Latin Square

§ 1. PROBLEM I STRESZCZENIE

Jednym ze schematów eksperymentalnych stosowanych powszechnie jest układ z pojedynczo rozszczepionymi poletkami (*split plot*). Jak wiadomo, zasada rozszczepionych poletek może być wykorzystana w blokach zrandomizowanych i w kwadracie łacińskim (por. F. Yates [7]).

Odpowiednie testy istotności F i przedziały ufności opierają się na dwu rodzajach błędów E_a i E_b , zależnie od tego, czy weryfikujemy hipotezę dla efektów czynnika A (na dużych poletkach — inaczej poletkach pierwszego rzędu), czy też hipotezę dla czynnika B (na małych poletkach — inaczej poletkach drugiego rzędu). Według H. L. Hartera [1] błąd E_b należy traktować jako sumę dwóch ortogonalnych i niejednorodnych składników: interakcji czynnika B z replikacjami R oraz reszty. Przez niejednorodność rozumiemy różne wariancje składników stanowiących błąd. Niektórzy autorzy nie uwzględniają takiego podziału. Harter, rozpatrując układ z rozszczepionymi poletkami w blokach zrandomizowanych, warunkuje dokonanie takiego podziału odrzuceniem hipotezy, że wariancja interakcji RB jest zerem.

W niniejszej pracy przedstawiono sposób rozbicia błędu E_b na składniki dla układu z pojedynczo rozszczepionymi poletkami, w którym poziomy czynnik A zostały rozlokowane według kwadratu łacińskiego. Przedstawione wzory na funkcje testowe dla weryfikowania hipotezy

braku zróżnicowania w poziomach czynnika A oraz interakcji AB są takie same, jak u Hartera, natomiast wzory na funkcje testowe dla weryfikowania hipotezy o braku różnic w poziomach czynnika B różnią się od analogicznych wzorów podanych przez Hartera. Omówiono sposób przeprowadzenia analizy wariancji oraz obliczono błędy standardowe, które mogą być wykorzystane przy konstruowaniu przedziałów ufności dla różnic średnich arytmetycznych. W § 5 pokazano na przykładzie liczbowym praktyczny sposób zastosowania uzyskanych wyników.

§ 2. MODEL MATEMATYCZNY I TESTY ISTOTNOŚCI

Model układu z rozszczepionymi poletkami oparty na kwadracie łacińskim, którym interesujemy się w niniejszej pracy ma postać następującą:

$$(1) y_{ijhk} = \mu + r_i + q_j + a_h + d_{ijh} + b_k + (rb)_{ik} + (qb)_{jk} + (ab)_{hk} + e_{ijhk}$$

gdzie: $i, j, h = 1, 2, \dots, a$; $k = 1, 2, \dots, b$

Przyjmujemy następujące założenia.

Obserwacje y_{ijhk} mają rozkład normalny;

1° r_i, q_j są to efekty odpowiednio i -tego wiersza i j -tej kolumny

d_{ijh}, e_{ijhk} — błędy E_a i E_b , przy czym e_{ijhk} może być dodatnio skorelowane dla różnych wartości k w obrębie tych samych ijh , o współczynniku korelacji równym ρ .

$(rb)_{ik}, (qb)_{jk}$ — efekty interakcyjne odpowiednio i -tego wiersza i j -tej kolumny z k -tym poziomem czynnika B .

Składniki 1° są zmiennymi losowymi o rozkładzie normalnym ze średnią zero i odpowiednimi wariancjami:

$$\sigma_r^2, \sigma_q^2, \sigma_{Ea}^2, \sigma_{Eb}^2, \sigma_{rb}^2, \sigma_{qb}^2.$$

2° $a_h, b_k, (ab)_{hk}$ są efektami odpowiednio h -tego poziomu czynnika A , k -tego poziomu czynnika B oraz interakcji h -tego poziomu czynnika A z k -tym poziomem czynnika B .

Zależnie od charakteru doświadczenia czynniki A i B można traktować jako stałe lub losowe. W związku z tym dla efektów 2° otrzymamy cztery warianty:

- I. A i B stałe, $\Sigma a_h = 0$; $\Sigma b_k = 0$; $\Sigma_h (ab)_{hk} = \Sigma_k (ab)_{hk} = 0$.
- II. A i B losowe, $a_h \sim N(0, \sigma_a^2)$; $b_k \sim N(0, \sigma_b^2)$; $(ab)_{hk} \sim N(0, \sigma_{ab}^2)$.
- III. A stałe, B losowe, $(ab)_{hk} \sim N\left(0, \frac{a-1}{a} \sigma_{ab}^2\right)$; $\Sigma_h (ab)_{hk} = 0$.
- IV. A losowe, B stałe, $(ab)_{hk} \sim N\left(0, \frac{b-1}{b} \sigma_{ab}^2\right)$; $\Sigma_k (ab)_{hk} = 0$.

Wartości oczekiwane średnich kwadratów zestawione w tab. 1 obliczono posługując się regułą E. F. Schultza [6]. Jeżeli w tab. 1 czynnik A uważamy za stały, należy pominąć współczynnik a^{**} , w przypadku zaś stałego czynnika B pomijamy współczynnik a^* . Dla stałych czynników A i B umownie przyjmujemy

$$\sigma_a^2 = \frac{\Sigma a_h^2}{a-1}, \quad \sigma_b^2 = \frac{\Sigma b_k^2}{b-1}, \quad \sigma_{ab}^2 = \frac{\Sigma \Sigma (ab)_{hk}^2}{(a-1)(b-1)}$$

Tab. 1. Wartości oczekiwane średnich kwadratów
Expected values of mean squares

Średni kwadrat	Liczba stopni swobody	σ_r^2	σ_q^2	σ_a^2	σ_{Ea}^2	σ_b^2	σ_{rb}^2	σ_{qb}^2	σ_{ab}^2	σ_{Eb}^2
V_r	$a-1$	ab			b		a^*			$1 + (b-1)\rho$
V_q	$a-1$		ab		b			a^*		$1 + (b-1)\rho$
V_a	$a-1$			ab	b				a^*	$1 + (b-1)\rho$
V_{Ea}	$(a-1)(a-2)$				b					$1 + (b-1)\rho$
V_b	$b-1$					a^2	a	a	a^{**}	$1-\rho$
V_{rb}	$(a-1)(b-1)$						a			$1-\rho$
V_{qb}	$(a-1)(b-1)$							a		$1-\rho$
V_{ab}	$(a-1)(b-1)$								a	$1-\rho$
V_{Eb}	$(a-1)(a-2)(b-1)$									$1-\rho$

Znajomość wartości oczekiwanych pozwala skonstruować testy istotności F (por. H. Scheffe [5]). Odpowiadające im funkcje testowe zestawiono w tab. 2. Wzory te nie ulegną zmianie, jeśli współczynnik korelacji $\rho = 0$.

Tab. 2. Hipotezy i funkcje testowe
Tests of hypotheses

Hipoteza	1	2	3	4	5
	$\sigma_a^2 = 0$	$\sigma_b^2 = 0$	$\sigma_{ab}^2 = 0$	$\sigma_{rb}^2 = 0$	$\sigma_{qb}^2 = 0$
Funkcja testowa dla warian-tów	$\frac{V_a}{V_{Ea}}$ [I, IV]	$\frac{V_b}{V_{rb} + V_{qb} - V_{Eb}}$ [I, III]	$\frac{V_{ab}}{V_{Eb}}$	$\frac{V_{rb}}{V_{Eb}}$	$\frac{V_{qb}}{V_{Eb}}$
	$\frac{V_a}{V_{Ea} + V_{ab} - V_{Eb}}$ [II, III]	$\frac{V_b}{V_{rb} + V_{qb} + V_{ab} - 2V_{Eb}}$ [II, IV]			

Wzory na funkcje testowe 1 i 2 mogą być uproszczone, jeżeli zostaną przyjęte tzw. wstępne hipotezy 3, 4 i 5. Obszar krytyczny α dla testów weryfikujących te hipotezy może dochodzić wg Hartera do 0,50 (por. także A. E. Paul [2]).

W przypadku przyjęcia hipotez 4 i 5 sumy kwadratów odchyłeń odpowiadające składnikom $(rb)_{ik}$ i $(qb)_{jk}$ włączamy do błędu E_b , dodając również odpowiednie liczby stopni swobody. Wstępną hipotezę 3 weryfikujemy tylko dla losowych składników a_h lub b_k . Może się zdarzyć, że tylko niektóre hipotezy wstępne zostaną przyjęte. Procedura łączenia dotyczy wtedy tylko tych sum kwadratów, dla których przyjęto hipotezy. Jeżeli hipotezy wstępne zostaną odrzucone, pozostają wówczas w mocy wzory na funkcje testowe zestawione w tab. 2.

W przypadku stosowania testów przybliżonych 1 i 2 zachodzi potrzeba określenia liczby stopni swobody, dla których odczytuje się wartości testowe z tablic F Snedecora. Odnosi się to jedynie do mianowników funkcji testowych, ponieważ w liczniku występuje zawsze pojedynczy średni kwadrat ze znaną liczbą stopni swobody. Szukaną liczbę stopni swobody szacujemy posługując się wzorem F. E. Satterthwaite'a [4].

§ 3. PRZEDZIAŁY UFNOŚCI

W tab. 3 przedstawiono kwadraty błędów standardowych, z których korzystamy przy konstruowaniu przedziałów ufności dla różnic średnich arytmetycznych. Przedziały ufności konstruuje się tylko dla stałych czynników A i B .

Tab. 3. Kwadraty błędów standardowych
Squares of standard erros

Porównywane średnie	dla wariantu I	dla wariantów III lub IV
$\bar{y} \dots h - \bar{y} \dots h'$	$\frac{2}{ab} Ea$	$\frac{2}{ab} [V_{Ea} + ab - V_{Ea}]$ III
$\bar{y} \dots k - \bar{y} \dots k'$	$\frac{2}{a^2} [V_{rb} + V_{qb} - E_{Eb}]$	$\frac{2}{a^2} [V_{rb} + V_{qb} + V_{ab} - 2V_{Eb}]$ IV
$\bar{y} \dots hk - \bar{y} \dots h'k$	$\frac{2}{a^2b} [V_{Ea} + (b-1)V_{Eb}]^*$	$\frac{2}{ab} [V_{Ea} + bV_{ab} - V_{Eb}]^*$ III
$\bar{y} \dots hk - \bar{y} \dots h'k'$	$\frac{2}{a^2} [V_{rb} + V_{qb} + (a-2)V_{Eb}]$	$\frac{2}{a^2} [V_{rb} + V_{qb} + aV_{ab} - 1V_{Eb}]$ IV
$\bar{y} \dots hk - \bar{y} \dots h'k'$	$\frac{2}{a^2b} \{aV_{Ea} + bV_{rb} + bV_{qb} + [(a-2)(b-1)-2]V_{Eb}\}^*$ Tylko dla wariantu I	

* Wzory te wyprowadzono zakładając, że współczynnik korelacji $\rho = 0$

These formulas are obtained under condition that coefficient of correlation $\rho = 0$

§ 4. WNIOSKI

Przedstawiona zasada podziału błędu E_b na składniki niejednorodne może być wykorzystana przez tych doświadczalników, którzy w praktyce posługują się omawianym modelem. Nieuwzględnienie takiego podziału w przypadku wystąpienia interakcji RB i QB prowadzi do zakłóceń w prawidłowym wnioskowaniu. Wyraża się to w zbyt częstym odrzuceniu hipotez $\sigma_a^2 = 0$ i $\sigma_b^2 = 0$ oraz przyjmowaniu hipotezy $\sigma_{ab}^2 = 0$. Tak więc podział błędu E_b wpływa na zmianę mocy testów stosowanych w analizie wariancji. Podział ten nie wpływa tylko na zmianę funkcji testowych weryfikujących hipotezę $\sigma_a^2 = 0$ dla wariantów I i IV.

§ 5. PRZYKŁAD LICZBOWY *

Założono doświadczenie w układzie rozszczepionych poletek (*split plot*), w którym poziomy czynnika A (cztery odmiany wyki), występujące na poletkach pierwszego rzędu rozlokowano według kwadratu łacińskiego; poziomy czynnika B (cztery stosunki wysiewu wyki z owsem) rozlosowano na poletkach drugiego rzędu. Doświadczenie założono w czterech replikacjach. Cechą badaną był plon nasion mieszanki.

O czynnikach A i B zakładamy, że są stałe. Uzyskane wnioski odnosimy tylko do tych czterech odmian i stosunków wysiewu. Analizę wariancji uwzględniającą podział błędu E_b na składniki niejednorodne przedstawiono w tab. 4.

Tab. 4. Analiza wariancji
Analysis of variance

Rodzaj zmienności		Liczba stopni swobody	Suma kwadratów nS^2	Średni kwadrat V	Wartość funkcji testowej F_0	Wartość graniczna	
						$F_{0,05}$	$F_{0,25}$
Wiersze	R	3	2,57	0,857	4,74	4,76	
Kolumny	Q	3	2,08	0,693			
Odmiany	A	3	4,80	1,600			
Błąd	E_a	6	2,02	0,337			
Stosunki wysiewu	B	3	4,84	1,613	5,68	3,86	1,423
Interakcje							
	RB	9	2,01	0,223	1,335		
	QB	9	2,56	0,284	1,701		
	AB	9	7,99	0,888	4,77	2,25	
Błąd	E_b	18	3,00	0,167			

* Dane eksperymentalne uzyskano z Katedry Hodowli Roślin WSR w Lublinie.

Na podstawie wstępnej analizy wariancji, przeprowadzonej przy poziomie istotności $\alpha=0,25$, podejmujemy decyzję o wydzieleniu z błędu E_b składnika $(qb)_{jk}$ ($F^\circ = 1,701$; $F_{0,25} = 1,423$) oraz o włączeniu do błędu drugiego rzędu składnika $(rb)_{ik}$ ($F^\circ = 1,335$; $F_{0,25} = 1,423$). W ten sposób otrzymujemy nowy błąd eksperymentalny E_b , którego suma kwadratów równa się: $nS_{Eb}^2 = 3,00 + 2,01 = 5,01$ i jest oparta na 27 stopniach swobody.

Średni kwadrat dla nowego błędu $V_{Eb} = 0,186$. We właściwej analizie wariancji należy sprawdzić hipotezy $\sigma_b^2 = 0$ i $\sigma_{ab}^2 = 0$. Włączając składnik $(rb)_{ik}$ do błędu drugiego rzędu zmieniamy kształt modelu matematycznego [1]; otrzymujemy nowy model, który nie zawiera tego składnika.

Funkcja testowa F dla sprawdzenia hipotezy $\sigma_b^2 = 0$ przyjmuje wówczas postać $F = V_b : V_{qb}$, a jej wartość liczbową $F^\circ = 1,613 : 0,284 = 5,68$ wobec $F_{0,05} = 3,86$ dla 3 i 9 stopni swobody. Hipotezę $\sigma_{ab}^2 = 0$ sprawdzamy obliczając wartość testową $F^\circ = 0,888 : 0,186 = 4,77$, przy $F_{0,05} = 2,25$ dla 9 i 27 stopni swobody.

Na podstawie przeprowadzonej analizy wariancji formułujemy następujące wnioski:

1. Nie stwierdzono istotnych różnic pomiędzy odmianami. Należy zauważyć, że porównanie poziomów czynnika A w układzie z rozszczepionymi poletkami jest iluzoryczne ze względu na małą liczbę stopni swobody, odpowiadającą błędowi E_a .

2. Stwierdzono istotne różnice pomiędzy stosunkami wysiewu.

3. Stwierdzono istotną interakcję pomiędzy odmianami i stosunkami wysiewu.

Bez podziału błędu E_b przebieg analizy byłby inny. Dotyczy to sprawdzania hipotez $\sigma_b^2 = 0$ i $\sigma_{ab}^2 = 0$. Funkcje testowe opierałyby się na tym samym błędzie $nS_{Eb}^2 = 2,01 + 2,56 + 3,00 = 7,57$, przy 36 stopniach swobody ze średnim kwadratem $V_{Eb} = 0,210$. Wartość testową dla B obliczylibyśmy $F^\circ = 1,613 : 0,210 = 7,68$ a dla AB $F^\circ = 0,888 : 0,210 = 4,23$.

Istotną cechą sugerowanej metody postępowania jest przyjęcie założenia, że w modelu matematycznym [1] mogą wystąpić losowe efekty interakcyjne pomiędzy poziomami czynnika B oraz wierszami R i kolumnami Q . Założenie to znajduje plastyczną ilustrację w doświadczeniach polowych. Istotne efekty interakcyjne $(rb)_{ik}$ i $(qb)_{jk}$ mogą wystąpić tylko na skutek niejednakowego reagowania tego samego poziomu czynnika B na warunki glebowe w obrębie poletek pierwszego rzędu. Przyczyną tego może być duża zmienność glebowa między poletkami drugiego rzędu w obrębie tego samego dużego poletka. Jak wiadomo, zmienność glebowa między poletkami drugiego rzędu powinna mieć charakter fluk-

tuacyjny. Tak więc istotność interakcji RB i QB może wskazywać na niejednorodność warunków glebowych na małych poletkach. W takim przypadku efekty główne czynnika B mogą być uwikłane ze zmiennością glebową występującą na małych poletkach.

Stosując przedstawioną zasadę podziału błędu E_b zmniejszamy średni kwadrat V_{Eb} występujący w mianowniku funkcji testowej $F = V_{ab} : V_{Eb}$. Dzięki temu łatwiej wykryć istotność interakcji AB , co, jak wiadomo, jest jednym z głównych celów w układach z rozszczepionymi poletkami.

Zagadnienie podziału błędu drugiego rzędu w innych modelach z rozszczepionymi poletkami oraz związane z tym problemy zostały szczegółowo omówione [3].

PIŚMIENNICTWO

1. Harter H. L.: On the Analysis of Split Plot Experiments. *Biometrics*, vol. 17, 1961.
2. Paul A. E.: On a Preliminary Test for Pooling Mean Squares in the Analysis of Variance. *The Annals of Mathematical Statistics*, vol. 21, 1950.
3. Przybysz T.: Pojedyncze i wielokrotne doświadczenia oparte na zasadzie rozszczepionych poletek. *Ann. Univ. Mariae Curie-Skłodowska, sectio E*, vol. XIX (1964), 16, Lublin 1965.
4. Satterthwaite F. E.: An Approximate Distribution of Estimates of Variance Components. *Biom. Bull.*, vol. 2, 1946.
5. Scheffe H.: *The Analysis of Variance*. New York 1959.
6. Schultz E. F.: Rules of Thumb for Determining Expectations of Mean Squares in Analysis of Variance. *Biometrics*, vol. 11, 1955.
7. Yates F.: Complex Experiments. Supplement to the *Journal of the Royal Statistical Society*, vol. II, 1935.

РЕЗЮМЕ

Рассматривается схема с распределенными делянками, в которой уровни главного фактора A расположены по латинскому квадрату. Уровни фактора B размещены внутри уровней фактора A . Факторы A и B могут быть постоянны или случайны, таким образом получается четыре модели. Для этих моделей представлены критерии значимости главных эффектов A и B а также их взаимодействия. Критерии значимости получаются при условии неоднородной ошибки E_b , которая может быть записана в виде нескольких слагаемых. На числовом примере представлено применение полученных результатов.

SUMMARY

The error E_b of split plot design in Latin square is partitioned into three components.

Four mathematical models with fixed or random effects of factors A and B are considered. The exact and approximate tests F and standard errors of differences between treatment means are presented. Numerical example is given.