Tadeusz P r z y b y s z (Lublin)

PORÓWNANIE DWÓCH PŁODOZMIANÓW O TEJ SAMEJ DŁUGOŚCI ROTACJI

1.1.Wstęp

Celem niniejszej pracy jest przedstawienie metody porównania za pomocą analizy wariancji dwóch płodozmianów o tej samej długości rotacji.

Cechą charakterystyczną płodozmianów jest z góry ustalone następstwo występowania różnych roślin w kolejnych latach na tym samym polu płodozmianowym. Za podstawę do porównania płodozmianów bierze się plony otrzymane w obu płodozmianach z tak zwanej rośliny testowej. Wybór rośliny testowej zależy od charakteru płodozmianu. Płodozmiany można porównywać wówczas, gdy w każdym z nich występuje roślina testowa.

Doświadczenie płodozmianowe rozpoczyna się równocześnie wszystkimi fazami, a występowanie rośliny testowej na poszczególnych polach określa się za pomocą macierzy incydencji G. Macierz G składa się z zer i jedynek; jedynki wskazują na występowanie rośliny testowej. Przez b oznaczono liczbę wierszy odpowiadającą liczbie lat prowadzenia doświadczenia, natomiast liczba a wskazuje liczbę pól płodozmianowych. Przykład macierzy incydencji do porównania dwóch płodozmianów czteropolowych.

$$(1.1.1) \qquad \underset{4\times 8}{G} = \left[\begin{array}{cccc:cccc} 1 & 0 & 0 & 1 & 1 & 0 & 1 & 0 \\ 1 & 1 & 0 & 0 & 0 & 1 & 0 & 1 \\ 0 & 1 & 1 & 0 & 1 & 0 & 1 & 0 \\ 0 & 0 & 1 & 1 & 0 & 1 & 0 & 1 \end{array}\right] = [G_1 \vdots G_2]$$

W macierzy (1.1.1) $b = 4$ lata, $a = 8$ pól płodozmianowych; $c = 2$ oznacza liczbę wystąpienia rośliny testowej na tym samym polu płodozmianowym w ciągu b (czterech) lat oraz $d = 4$ wskazuje liczbę pól płodozmianowych, na których występuje roślina testowa w jednym roku.

Problemem porównywania płodozmianów przy pomocy metod analizy wariancji zajmowali się Yates [1954] i Patterson [1964]. Patterson zajmował się głównie metodami oceny komponentów wariancyjnych. Yates w swojej pracy ograniczył się do porównania płodozmianu trójpolowego z czteropolowym; uzyskane przez niego wyniki odnosiły się do rozpatrywanego przykładu bez wyprowadzenia ogólnych wzorów. Nie sprecyzował on wyraźnie modelu matematycznego ani założeń; zredukowane równania normalne przedstawił bez wyprowadzenia. Uogólnienie wyników uzyskanych przez F.Yatesa w odniesieniu do porównania płodozmianów o różnej długości rotacji znajdzie Czytelnik w pracy autora [1976].

1.2. Notacja.

n - liczba obserwacji w pojedynczej replikacji $n=a\cdot c=bd$

r - liczba replikacji

N - liczba obserwacji w całym doświadczeniu; $N = nr$

Przez X oznaczamy macierz układu, która może być podzielona na macierze blokowe związane z poszczególnymi parametrami modelu

$$\underset{N \times q}{X} = [J_N \vdots X_R \vdots X_A \vdots X_B \vdots X_{BR} \vdots X_{AB}]$$

X' - macierz transponowana

E_n - macierz kwadratowa $n \times n$ złożona z samych jedynek

I_r - macierz jednostkowa $r \times r$

$\underset{n\times a}{X_a}, \underset{n\times b}{X_b}, \underset{n\times a}{M_a}$ - macierze odnoszące się do pojedynczych replikacji

$I_r \otimes X_a$ - iloczyn kroneckerowski macierzy

M_a^- - uogólniona macierz odwrotna macierzy M_a

$\underset{N\times N}{V}$ - macierz kowariancji wektora Y

V_n - macierz kowariancji wektora Y_i (wektora obserwacji i-tej replikacji)

P_1, P_2, P_A, P_B i.t.d. - operatory rzutowe o wymiarach $N\times N$

P_a, P_b - operatory rzutowe o wymiarach $n\times n$ odnoszące się do pojedynczej replikacji

$\underset{r\times 1}{\rho}, \underset{a\times 1}{\alpha}, \underset{b\times 1}{\beta}, \underset{br\times 1}{\gamma}, \underset{\eta\times 1}{\varphi}$ - wektory kolumnowe nieznanych parametrów

$\underset{N\times 1}{e_1}, \underset{N\times 1}{e_2}$ - losowe wektory błędów eksperymentalnych

J_N - wektor kolumnowy złożony z samych jedynek

$\underset{N\times 1}{Y}$ - wektor obserwacji w całym doświadczeniu

$\underset{n\times 1}{Y_i}$ - wektor i-tej obserwacji

$\underset{a\times 1}{Y_j}$ - wektor sum obiektowych

$\underset{b\times 1}{Y_k}$ - wektor sum dla lat

SS - suma kwadratów odchyleń

MS - średni kwadrat odchyleń

Oznaczenia innych symboli podajemy w tekście.

2.1. Model matematyczny.

Doświadczenia płodozmianowe można opisać przy użyciu zmodyfikowanego modelu matematycznego split block (por. Yates [1954], Harter [1961], Przybysz [1964, 1976]). Mamy tu do czynienia z klasyfikacją podwójną $A{\times}B$; rolę czynnika A spełniają pola płodozmianowe, drugi kierunek klasyfikacji (czynnik B) wyznaczają kolejne lata. Modyfikacja modelu polega na tym, że obserwacje nie występują w każdej podklasie $A{\times}B$.

Rozważmy model

$$y_{ijk} = \mu + \rho_i + \alpha_j + e_{ij} + \beta_k + \gamma_{ik} + \varphi_{jk} + e_{ijk} \tag{2.1.1}$$

μ - wartość średnia populacji

ρ_i - efekt i-tej replikacji $i = 1,2,\ldots,r$

α_j - efekt j-tego obiektu $j = 1,2,\ldots,a$

β_k - efekt k-tego roku $k = 1,2,\ldots,b$

γ_{ik} oraz φ_{jk} - efekty interakcyjne

e_{ij} - błąd eksperymentalny związany z poletkami płodozmianowymi

e_{ijk} - błąd eksperymentalny odnoszący się do całego doświadczenia

y_{ijk} - pojedyncze obserwacje

Przyjmujemy, że w modelu (2.1.1) składniki e_{ij} oraz e_{ijk} są nieskorelowanymi zmiennymi losowymi o rozkładzie normalnym

$$e_{ij} \sim N(0,\sigma_1^2); \quad e_{ijk} \sim N(0,\sigma_2^2).$$

Wszystkie pozostałe składniki są stałymi parametrami. Model (2.1.1) w zapisie macierzowym przyjmuje postać

(2.1.2) $$Y=J_N\mu+X_R\rho+X_A\alpha+X_B\beta+X_{BR}\gamma+X_{AB}\varphi+X_1e_1+X_2e_2$$

Postać macierzy X zależy od uporządkowania wektora obserwacji $\underset{N\times 1}{Y}$. Dogodnie jest przyjąć uporządkowanie wektora Y według kolejności replikacji (i), obiektów (j) oraz lat (k). Wtedy macierze blokowe można wyrazić w postaci iloczynów kroneckerowskich.

(2.1.3) $$\underset{N\times r}{X_R} = I_r \otimes J_n$$

(2.1.4) $$\underset{N\times a}{X_a} = J_r \otimes X_a = J_r \otimes I_a \otimes J_c$$

(2.1.5) $$\underset{N\times b}{X_b} = J_r \otimes X_b$$

Macierz X_b określa się za pomocą transponowanej macierzy incydencji G′ w następujący sposób:

1. Macierz X_b ma n wierszy i b kolumn
2. W każdym wierszu macierzy X_b występuje tylko raz jedynka, pozostałe elementy są zerami
3. Miejsce jedynki w pierwszym wierszu macierzy X_b pokrywa się z pierwszą jedynką macierzy G′
4. Jedynka w drugim wierszu macierzy X_b występuje w tej kolumnie, w której występuje druga jedynka w pierwszym wierszu macierzy G′
5. Po wyczerpaniu jedynek z pierwszego wiersza macierzy G′ przechodzimy do wiersza drugiego; postępujemy tak aż do wyczerpania wszystkich wierszy macierzy G′.

(2.1.6) $$\underset{N\times br}{X_{BR}} = I_r \otimes X_b$$

(2.1.7) $$\underset{N\times n}{X_{AB}} = J_r \otimes I_n$$

(2.1.8) $$\underset{N\times ar}{X_1} = I_r \otimes X_a$$

(2.1.9) $$X_2 = I_N$$

Macierz kowariancji V wektora Y wyrazi się wzorem

(2.1.10) $$\underset{N\times N}{V} = \sigma_1^2 X_1 X_1' + \sigma_2^2 X_2 X_2' = I_r \otimes V_n$$

(2.1.11) $$V_n = \sigma_1^2 I_a \otimes E_c + \sigma_2^2 I_n$$

Po kolejnych przekształceniach otrzymujemy

$$V_n = I_{ra} \otimes \begin{bmatrix} \sigma_1^2 + \sigma_2^2 & \sigma_1^2 \\ \sigma_1^2 & \sigma_1^2 + \sigma_2^2 \end{bmatrix}$$

Macierz kowariancji V jest pełnego rzędu (por.Seber [1966]), można zatem wyznaczyć macierz odwrotną V^{-1}

$$V^{-1} = I_r \otimes V_n^{-1}$$

(2.1.12) $$V_n^{-1} = \frac{1}{\sigma_2^2(c\sigma_1^2+\sigma_2^2)} I_{ra} \otimes \begin{bmatrix} \sigma_1^2 + \sigma_2^2, & -\sigma_1^2 \\ -\sigma_1^2, & \sigma_1^2 + \sigma_2^2 \end{bmatrix}$$

2.2. Równania normalne.

Celem oszacowania nieznanych parametrów posłużymy się metodą najmniejszych kwadratów. Równania normalne dla modelu (2.1.2) przyjmują postać (por.Searle [1971]).

$$X'V^{-1}X\Theta = X'V^{-1}Y \tag{2.2.1}$$

Przez Θ oznaczyliśmy wektor nieznanych parametrów; przy czym

$$\Theta' = [\mu' \,\vdots\, \rho' \,\vdots\, \alpha' \,\vdots\, \beta' \,\vdots\, \gamma' \,\vdots\, \rho']$$

Głównym przedmiotem analizy wariancji jest wyznaczenie ocen efektów płodozmianowych δ_1 i δ_2, które wyrażają się poprzez oceny $\hat{\alpha}$ wektora α.

Z tego względu zajmiemy się głównie wyznaczeniem oceny wektora $\hat{\alpha}$. Spośród równań (2.2.1) rozpatrzmy szczegółowo tylko te, które odnoszą się bezpośrednio do wektora α.

Równania te przyjmują postać

$$\begin{aligned} &X_A'V^{-1}J_N\hat{\mu} + X_A'V^{-1}X_R\hat{\rho} + X_A'V^{-1}X_A\hat{\alpha} + X_A'V^{-1}X_B\hat{\beta} + \\ &+ X_A'V^{-1}X_{BR}\hat{\gamma} + X_A'V^{-1}X_{AB}\hat{\varphi} = X_A'V^{-1}Y \end{aligned} \tag{2.2.2}$$

$$\begin{aligned} &X_B'V^{-1}J_N\hat{\mu} + X_B'V^{-1}X_R\hat{\rho} + X_B'V^{-1}X_A\hat{\alpha} + X_B'V^{-1}X_B\hat{\beta} + X_B'V^{-1}X_{BR}\hat{\gamma} + \\ &+ X_B'V^{-1}X_{AB}\hat{\varphi} = X_B'V^{-1}Y \end{aligned}$$

Po wykonaniu zaznaczonych działań otrzymamy dla pierwszego równania (2.2.2)

$$X_A' V^{-1} J_N = \frac{rc}{c\,\sigma_1^2 + \sigma_2^2} J_a$$

$$X_A' V^{-1} X_R = \frac{c}{c\,\sigma_1^2 + \sigma_2^2} J_a \otimes J_r'$$

$$X_A' V^{-1} X_A = \frac{rc}{c\,\sigma_1^2 + \sigma_2^2} I_a$$

$$X_A' V^{-1} X_B = \frac{r}{c\,\sigma_1^2 + \sigma_2^2} G'$$

$$X_A' V^{-1} X_{BR} = \frac{1}{c\,\sigma_1^2 + \sigma_2^2} J_r' \otimes G'$$

$$X_A' V^{-1} X_{AB} = \frac{1}{c\,\sigma_1^2 + \sigma_2^2} I_a \otimes J_c'$$

$$X_A' V^{-1} Y = \frac{1}{c\,\sigma_1^2 + \sigma_2^2} Y_j$$

Analogiczne działania wykonujemy dla drugiego równania (2.2.2)

$$X_B' V^{-1} Y_N = \frac{dr}{c\,\sigma_1^2 + \sigma_2^2} J_b$$

$$X_B' V^{-1} X_R = \frac{d}{c\,\sigma_1^2 + \sigma_2^2} J_b \otimes J_r'$$

$$X_B' V^{-1} X_A = \frac{r}{c\,\sigma_1^2 + \sigma_2^2} G$$

$$X_B' V^{-1} X_B = r X_b' V^{-1} X_b$$

$$X_B'V^{-1}X_{BR} = Y_R' \otimes X_b'V_n^{-1}X_b$$

$$X_B'V^{-1}X_{AB} = rX_b'V_n^{-1}$$

Z modelem (2.1.2) związane są następujące restrykcje, które nakładamy na parametry

$$J_r'\rho = 0, \quad J_a'\hat{\alpha} = 0, \quad J_b'\beta = 0$$

$\sum_i \gamma_{ik} = 0$ dla każdego k oraz $\sum_k \gamma_{ik} = 0$ dla każdego i.

$\sum_j \varphi_{jk} = 0$ dla każdego k oraz $\sum_k \varphi_{jk} = 0$ dla każdego j.

Przy sformułowaniu restrykcji dla efektów φ_{jk} trzeba zauważyć, że w danym roku k występują tylko niektóre składniki φ_{jk}, nie obejmujące wszystkich wartości j; to samo można powiedzieć o efektach φ_{jk} dla ustalonego poziomu j. Jeśli zatem piszemy $\sum_j \varphi_{jk} = 0$ to rozumiemy, że suma ta rozciąga się tylko na te wskaźniki j, które występują dla ustalonego k, analogicznie $\sum_k \varphi_{jk} = 0$ dotyczy tych k, w których występuje ustalony obiekt "j". Taki wybór restrykcji wynika z charakteru modelu, pozwala on uzyskać jednoznaczne rozwiązania układu równań normalnych. Możliwe jest również podejście ogólniejsze - poprzez reparametryzacje modelu (2.1.1).

Podstawiając do układu (2.2.2) w uzyskane wyniki przeprowadzonych działań na macierzach oraz uwzględniając restrykcje otrzymamy układ

$$(2.2.3)\qquad \begin{aligned} &\frac{rc}{c\,\sigma_1^2+\sigma_2^2} J_a\hat{\mu} + \frac{rc}{c\,\sigma_1^2+\sigma_2^2} X_A\hat{\alpha} + \frac{r}{c\sigma_1^2+\sigma_2^2} G'\hat{\beta} = \frac{1}{c\sigma_1^2+\sigma_2^2} Y_j \\ &\frac{dr}{c\,\sigma_1^2+\sigma_2^2} J_b\hat{\mu} + \frac{r}{c\,\sigma_1^2+\sigma_2^2} G\hat{\alpha} + rX_b'V_n^{-1}X_b\hat{\beta} = X_B'V^{-1}Y \end{aligned}$$

Z uwagi na to, że macierz X_b nie wyraża się w postaci iloczynu kroneckerowskiego, działania obejmujące tę macierz muszą być wykonywane oddzielnie dla każdego schematu eksperymentalnego. Eliminując z układu (2.2.3) nieznany parametr $\hat{\mu}$ oraz $\hat{\beta}$ otrzymuje się tak zwany zredukowany układ równań normalnych dla wektora α w postaci

$$(2.2.4)\qquad M\hat{\alpha} = Q$$

Jest godne podkreślenia, że układ (2.2.4) nie obejmuje nieznanych wariancji σ_1^2 i σ_2^2, tak więc ocena wektora $\hat{\alpha}$ nie zależy od σ_1^2 i σ_2^2.

Macierz $\underset{a\times a}{M}$ wyznaczamy z następującej relacji

$$M = r\,M_a$$

gdzie $\underset{a\times a}{M_a}$ daje się przedstawić jako

$$(2.2.5)\qquad M_a = X_a'X_a - \frac{1}{d}G'G = c\,I_a - \frac{1}{d}G'G$$

Natomiast wektor prawych stron $\underset{a\times 1}{Q}$ wyraża się wzorem

$$(2.2.6)\qquad Q = X_A'Y - \frac{1}{d}G'X_B'Y = Y_j - \frac{1}{d}G'Y_k$$

Macierz M_a jest macierzą osobliwą, dlatego do jednoznacznego rozwiązania układu (2.2.4) należy dołączyć restrykcje $J_a'\hat{\alpha} = 0$.

Do ogólnych rozważań wygodnie jest przedstawić rozwiązanie układu równań (2.2.4) przy użyciu uogólnionej macierzy odwrotnej M_a^-. Wektor ocen $\hat{\alpha}$ wyrazi się wtedy wzorem

$$\hat{\alpha} = \frac{1}{r} M_a^- Q \tag{2.2.7}$$

Wektor $\hat{\alpha}$ daje oceny wektorów obiektowych dla całego eksperymentu z uwzględnieniem wszystkich replikacji.

Do wyznaczenia w analizie wariancji sumy kwadratów odchyleń dla tak zwanego pierwszego błędu potrzebne są również oceny efektów obiektowych wyznaczonych dla każdej replikacji oddzielnie. Jeżeli przez $\underset{a\times 1}{\alpha_i}$ oznaczymy wektor nieznanych efektów obiektowych dla i-tej replikacji, wtedy ocenę $\hat{\alpha}_i$ otrzymamy z następującej relacji

$$\hat{\alpha}_i = M_a^- Q_i \tag{2.2.8}$$

Wektor $\underset{a\times 1}{Q_i}$ wyznacza się oddzielnie dla każdej replikacji ze wzoru

$$Q_i = X_a' Y_i - \frac{1}{d} G' X_b' Y_i$$

Macierz M_a ma tę samą postać i nie zależy od replikacji.

3.1. Sumy kwadratów odchyleń.

Podstawą do przeprowadzenia analizy wariancji dla modelu (2.1.2) jest podział ogólnej sumy kwadratów odchyleń SS_y na składniki zgodnie z tożsamością (por. Yates [1954], Patterson [1964]).

$$SS_y = SS_R + SS_A + SS_B + SS_{BR} + SS_{AB} + SS_1 + SS_2 \tag{3.1.1}$$

Wzory rachunkowe na poszczególne sumy kwadratów przyjmują postać

(3.1.2) $$SS_A = \hat{\alpha}Q = \frac{1}{r} Q'M_a^- Q$$

(3.1.3) $$SS_1 = \sum_i^r \hat{\alpha}_i' Q_i - \hat{\alpha}' Q = \sum_i^r Q_i'M_a^- Q_i - \frac{1}{r} Q'M_a^- Q$$

Pozostałe sumy kwadratów odchyleń oblicza się w zwykły sposób

(3.1.4)
$$SS_R = \frac{\sum_i Y_{i...}^2}{n} - \frac{Y_{...}^2}{nr}$$

$$SS_B = \frac{\sum_k Y_{..k}^2}{dr} - \frac{Y_{...}^2}{nr}$$

$$SS_{BR} = \frac{\sum_i \sum_k Y_{i.k}^2}{d} - \frac{Y_{i..}^2}{n} - \frac{Y_{i..}^2}{dr} + \frac{Y_{...}^2}{nr}$$

$$SS_{AB} = \frac{\sum_j \sum_k Y_{..k}^2}{r} - \frac{Y_{..k}^2}{dr} - SS_A$$

Sumę kwadratów SS_2 oblicza się przez dopełnienie do całości.

Tożsamość (3.1.1) daje się wyrazić za pomocą operatorów rzutowych P jako

(3.1.4) $$Y'I_NY - Y'P_NY = Y'P_RY + Y'P_AY + Y'P_BY + Y'P_{BR} + \\ + Y'P_{AB}Y + Y'P_1Y + Y'P_2Y$$

Operatory P związane są z odpowiednimi macierzami blokowymi X. Korzystamy tu z następującej relacji

(3.1.5) $$P[X] = X[X'X]^{-1} X'$$

gdzie X jest dowolną macierzą blokową.

Zgodnie z przyjętymi oznaczeniami

$$P_N = P[J_N] = \frac{1}{N} E_N$$

(3.1.6) $$P_R = P[X_R] \; P_N = X_R[X_R' X_R]X_R - \frac{1}{N} E_N =$$

$$= \frac{1}{n} I_r \otimes E_n - \frac{1}{N} E_N = \frac{1}{N}[r\, I_r - E_r] \otimes E_n$$

Operator P_A wyznaczamy wychodząc ze wzoru (3.1.2)

$$SS_A = \frac{1}{r} Q' M_a^- Q = \frac{1}{r} Y' [X_A \frac{1}{d} X_B\, G] M_a^- [X_A' - \frac{1}{d} G' X_B'] Y$$

Stąd operator P_A przyjmuje postać

$$P_A = \frac{1}{r} [X_A - \frac{1}{d} X_B\, G] M_a^- [X_A' - \frac{1}{d} G' X_B']$$

Po przeprowadzeniu odpowiednich przekształceń otrzymamy

(3.1.7) $$P_A = \frac{1}{r} E_r \otimes [X_a - \frac{1}{d} X_b\, G] M_a^- [X_a' - \frac{1}{d} G' X_b'] = \frac{1}{r} E_r \otimes P_a$$

Przez P_a oznaczyliśmy operator odnoszący się do pojedynczej replikacji.

Operator P_1 wyznaczymy biorąc za punkt wyjścia wzór (3.1.3). We wzorze tym sumowania po (Σ_i) można zapisać mnożeniem kroneckerowskim operatora P_a przez macierz jednostkową I_r, tak więc otrzymamy

(3.1.8) $$P_1 = I_r \otimes P_a - P_A = \frac{1}{r}[r\, I_r - E_r] \otimes P_a$$

Pozostałe operatory wyznaczamy analogicznie jak operator P_R.

(3.1.9) $$P_B = \frac{1}{rd} E_r \otimes X_b X_b' - \frac{1}{N} E_N = \frac{1}{r} E_r \otimes P_b$$

gdzie $$P_b = [\frac{1}{d} X_b X_b - \frac{1}{bd} E_n]$$

(3.1.10) $$P_{BR} = \frac{1}{r}[rI_r - E_r] \otimes P_b$$

(3.1.11) $$P_{AB} = \frac{1}{r} E_r \otimes [I_n - \frac{1}{d} X_b X_b' - P_a]$$

Do wyznaczenia operatora P_2 posłużymy się podziałem liczby stopni swobody dla całości na składniki (por. Oktaba [1958])

$$\nu_e = rn - rb - n + b - (a-1)(r-1)$$

Zgodnie z tym podziałem

$$P_2 = I_N - \frac{1}{d} I_r \otimes X_b X_b' - \frac{1}{r} E_r \otimes I_n + \frac{1}{rd} E_r \otimes X_b X_b' - P_1$$

Po dokonaniu przekształceń P_2 przyjmuje postać

(3.1.12) $$P_2 = \frac{1}{r}[r\, I_r - E_r] \otimes [I_n - \frac{1}{d} X_b X_b' - P_a]$$

Łatwo udowodnić, że formy kwadratowe SS_A i SS_1 są niezależne. W tym celu wystarczy wykazać, że $P_A V P_1 = 0$ (por. Graybill [1961], tw. 4.2.1)

$$P_A = \frac{1}{r} E_r \otimes P_a$$

$$P_1 = \frac{1}{r} [r\, I_r - E_r] \otimes P_a$$

$$V = I_r \otimes V_n$$

$$P_A V P_1 = (\frac{1}{r} E_r \otimes P_a)(I_r \otimes V_n)(\frac{1}{r}[r\, I_r - E_r] \otimes P_a) =$$

$$= \frac{1}{r^2} E_r [r\, I_r - E_r] P_a V_n P_a = \frac{1}{r^2}[r\, E_r - E_r^2] \otimes P_a V_n P_a = 0$$

ponieważ $rE_r = E_r^2$. W analogiczny sposób dowodzi się niezależności form kwadratowych; SS_{AB} i SS_B i SS_E, SS_B i SS_{BR} oraz SS_A i SS_R.

W analizie wariancji składnik SS_A podlega dalszemu podziałowi

$$SS_A = SS_{Pł} + SS_{Re}$$

Przez $SS_{Pł}$ oznaczamy sumę kwadratów odchyleń płodozmianów, obliczamy ją według wzoru

$$SS_{Pł} = \frac{Y_I^2}{cpr} + \frac{Y_{II}^2}{cpr} - \frac{Y_{...}^2}{nr} \tag{3.1.13}$$

Przez Y_I i Y_{II} oznaczono odpowiednio sumy wszystkich wyników pierwszego i drugiego płodozmianu; p jest liczbą pól porównywanych płodozmianów. W przykładzie macierzy incydencji (1.1.1) p=4 oznacza liczbę kolumn macierzy blokowej G_1 i G_2, macierze te łączą się z oboma porównywanymi płodozmianami.

$$Y_I = \sum_i^r \sum_{j=1}^4 \sum_k^b y_{ijk}; \quad Y_{II} = \sum_i^r \sum_{j=5}^8 \sum_k^b y_{ijk}$$

Operator rzutowy przy pomocy którego można wyznaczyć $SS_{Pł}$ przyjmuje postać

$$P_{Pł} = \frac{1}{r}E_r \otimes \left\{ \frac{1}{cp} \left[\begin{array}{c|c} E_{cp} & O \\ \hline O & E_{cp} \end{array} \right] - \frac{1}{n}E_n \right\} \tag{3.1.14}$$

Sumy kwadratów $SS_{Pł}$ i SS_1 są niezależne, wynika to stąd, że $P_{Pł}VP_1 = 0$; dowód tej równości jest natychmiastowy.

3.2. Wartości oczekiwane form kwadratowych i analiza wariancji.

Do przeprowadzenia analizy wariancji niezbędne jest uprzednie wyznaczenie wartości oczekiwanych średnich kwadratów. Korzystamy tu z twierdzenia, które orzeka, że jeśli wektor obserwacji ma rozkład z wartością średnią i macierzą kowariancji V to wartość oczekiwana formy kwadratowej $Y'PY$ przyjmuje postać

(3.2.1) $$E(Y'PY) = tr(PV) + \mu' P \mu$$

We wzorze (3.2.1) przez $tr(PV)$ oznaczono ślad macierzy, natomiast składnik $\mu' P \mu$ jest parametrem niecentralności.

Do stosowania wzoru 3.2.1 wymagana jest znajomość jawnej postaci operatora P, a niestety operatory P_A, P_B, P_{BR} oraz P_1 i P_2 zależą bądź od uogólnionej macierzy odwrotnej M_a^- lub macierzy blokowej X_b. Stronę rachunkową związaną z wyznaczeniem wartości oczekiwanych sum kwadratów pomijamy. Uzyskane wyniki zestawiono w tablicy 1.

Parametry δ_1 i δ_2 są związane z efektami α_j następującymi wzorami

$$\delta_1 = \frac{1}{p}\sum_{j=1}^{p}\alpha_j \ , \qquad \delta_2 = \frac{1}{p}\sum_{j=p+1}^{a}\alpha_j$$

Parametry niecentralności λ_a i λ_{ab} przyjmują postać

$$\lambda_a = \frac{r\,c(d-1)}{(a-1)d}\sum_{j=1}^{a}\alpha_j^2 = \frac{rc(d-1)}{(a-1)d}\,\alpha' I_a \alpha$$

$$\lambda_{ab} = \frac{r}{n-a-b+1}\sum_{j}^{a}\sum_{k}^{b}\varphi^2_{jk} = \frac{r}{n-a-b+1}\,\varphi' I_n \varphi .$$

Głównym celem analizy wariancji jest ocena istotności kontrastu

Tablica 1. Wartości oczekiwane średnich kwadratów

Źródło zmienności	Liczba stopni swobody	Średni kwadrat	Wartości oczekiwane E(MS)
Replikacje	r - 1	MS_R	$\sigma_2^2 + c\sigma_1^2 + n\,\rho' I_r \rho$
Obiekty - pola płodozmianowe	a - 1	MS_A	$\sigma_2^2 + \frac{n-b}{a-1}\sigma_1^2 + \lambda_a$
Płodozmiany	1	$MS_{Pł}$	$\sigma_2^2 + c\,\sigma_1^2 + pr(\delta_1 - \delta_2)^2$
Reszta	a - 2	MS_{Re}	$\sigma_2^2 + \frac{n-b-c}{a-2}\sigma_1^2 + \lambda_{re}$
Błąd I	(a-1)(r-1)	MS_1	$\sigma_2^2 + \frac{n-b}{a-1}\sigma_1^2$
Lata	b - 1	MS_B	$\sigma_2^2 + \frac{b-c}{b-1}\sigma_1^2 + \frac{dr}{b-1}\beta' I_b \beta$
Interakcja Repl. x Lata	(b-1)(r-1)	MS_{BR}	$\sigma_2^2 + \frac{b-c}{b-1}\sigma_1^2 + \frac{d}{(b-1)(r-1)}\gamma' I_{br}\gamma$
Interakcja Obiekty x Lata	n-a-b+1	MS_{AB}	$\sigma_2^2 + \lambda_{ab}$
Błąd II	(r-1)(n-a-b+1)	MS_2	σ_2^2

$$\delta_1 - \delta_2 = \frac{1}{p}\sum_{j=1}^{p}\alpha_j - \frac{1}{p}\sum_{j=p+1}^{a}\alpha_j$$

Ocenę tego kontrastu traktujemy jako ocenę różnicy efektów porównywanych płodozmianów.

Zgodnie z tym hipotezę zerową zapiszemy jako

$$H_o : \delta_1 = \delta_2 \tag{3.2.2}$$

wobec hipotezy alternatywnej $H_1 : \delta_1 \neq \delta_2$.

W swoich rozważaniach skupiamy się tylko na hipotezie (3.2.2). Cała analiza wariancji podporządkowana jest temu celowi. Oczywiście formalnie można podstawić i sprawdzać hipotezy dotyczące innych parametrów modelu (2.1.1). Uzyskane wnioski wynikające ze sprawdzenia tych hipotez nie miałyby praktycznego znaczenia lub byłyby z punktu widzenia praktyki rolniczej oczywiste.

Takimi oczywistymi wnioskami są z reguły istotne różnice w planowaniu w różnych latach - wynika to ze zmienności warunków meteorologicznych; podobnie formalny charakter miałyby wnioski wynikające ze stwierdzenia istotnych różnic pomiędzy efektami α_j, które związane są z poszczególnymi polami płodozmianowymi.

Hipotezę (3.2.2) sprawdzamy przy pomocy przybliżonego testu F (Scheffe [1959]).

Funkcja testowa F przyjmuje postać

$$F = \frac{MS_{pł}}{\widehat{E\ MS}_{pł}} \tag{3.2.3}$$

W mianowniku występuje ocena wartości oczekiwanej średniego kwadratu dla płodozmianów.

Do wyznaczenia funkcji testowej F potrzebne są oceny σ_1^2 i σ_2^2. Oceny te uzyskamy przyrównując odpowiednie średnie kwadraty do ich wartości oczekiwanych, to znaczy

$$MS = E(MS)$$

Stąd

$$\hat{\sigma}_2^2 = MS_2$$

oraz

$$MS_1 = \hat{\sigma}_2 + \frac{n-b}{a-1}\hat{\sigma}_1^2$$

Po dokonaniu podstawienia otrzymamy

$$\hat{\sigma}_1^2 = \frac{a-1}{n-b}[MS_1 - MS_2]$$

Zakłada się, że $MS_1 > MS_2$; nierówność ta wynika z charakteru modelu (2.1.1).

Po dokonaniu podstawień otrzymamy

$$E(MS_{pł}) = \hat{\sigma}_2^2 + c\,\hat{\sigma}_1^2 = \frac{c}{n-b}[(a-1)MS_1 - (b-1)MS_2]$$

Funkcja testowa (3.2.3) ma przybliżony rozkład F z 1 stopniem swobody dla licznika i f stopniami swobody dla mianownika. Zgodnie ze wzorem Satterthaite'a [1946]

$$f = \frac{[(a-1)MS_1 - (b-1)MS_2]^2}{\frac{[(a-1)MS_1]^2}{(a-1)(r-1)} + \frac{[(b-1)MS_2]^2}{(n-a-b+1)(r-1)}}$$

Można zauważyć, że dla sprawdzenia hipotezy zerowej dotyczącej istotności efektów α_j otrzymalibyśmy dokładny test postaci

$$F = \frac{MS_A}{MS_1}$$

Wyprowadzone w pracy wzory odnoszą się do pewnej klasy doświadczeń płodozmianowych spełniających określone warunki.

1. Porównujemy dwa płodozmiany o tej samej długości rotacji - - długość rotacji oznaczono literą p.
2. W obu płodozmianach występuje roślina testowa tę samą (c) ilość razy w ciągu b lat.
3. Konstrukcje schematu eksperymentalnego, opisanego za pomocą macierzy incydencji, wymaga, aby wszystkie numery obiektów $j = p + 1, \ p + 2, \ldots, a$ drugiego płodozmianu spotkały się z każdym numerem $j = 1, 2, \ldots p$ obiektów pierwszego płodozmianu.

Przy tych założeniach układ równań (2.2.4) da się rozwiązać bez uciekania się do wyznaczania macierzy M_a^-.

Wszystkie rachunki związane z zastosowaniem przedstawionej metody można przeprowadzić przy użyciu programu na e.m.c. ODRA 1325, którego autorką jest Barbara Jezior; program ten pozostaje w dyspozycji Ośrodka Obliczeniowego Akademii Rolniczej w Lublinie.

Literatura cytowana

Graybill, F.A.; 1961, An introduction to linear statistical models, vol.I, New York.

Harter, H.L.; 1961, On the analysis of split plot experiments, Biometrics 17, str.144-149.

Oktaba, W.; 1958, O wzajemnej i jednoznacznej odpowiedniości między sumami kwadratów i stopniami swobody w analizie wariancji, Annales UMCS, sec.E, 12, str.327-344.

Patterson, H.D.; 1964, Theory of cyclic rotation experiments, J.Roy. Statist.Soc., ser.B, 26, str.1-45.

Przybysz, T.; 1964, Pojedyncze i wielokrotne doświadczenia oparte na zasadzie rozszczepionych poletek, Część I, Analiza wariancji, Annales UMCS, sec.E, 19, str.307-332.

Przybysz, T.; 1976, Statystyczne metody porównywania płodozmianów, AR Lublin, Rozprawy Naukowe, 39.

Satterthwaite, F.E.; 1946, An aproximate distribution of estimates of variance components, Biom.Bull., 2, str.110-114.

Searle, S.R.; 1971, Linear models, Wiley, New York.

Seber, G.A.P.; 1966, The linear hypothesis, Griffin, London.

Scheffe, H.; 1959, The analysis of variance, Wiley, New York.

Yates, F.; 1954, The analysis of experiments containing different crop rotations, Biometrics 10, str.324-346.