

## Modele ANOVA i ANCOVA jako szczególne modele liniowe

Problem porównywania średnich w  $k$  zależnych lub niezależnych grupach można potraktować jako szczególny problem testowania hipotezy liniowej.

**Model:**  $Y = \varphi + \varepsilon$

gdzie  $Y \in R^n$  jest wektorem obserwacji  
 $\varphi \in \Omega \subset R^n$  jest wektorem średnich, o którym wiadomo, że należy do pewnej **właściwej podprzestrzeni liniowej  $\Omega$  przestrzeni  $R^n$** , tzn.  $\Omega \subset R^n$  i  $\dim(\Omega) = p < n$   
 $\varepsilon \in R^n$  jest losowym wektorem błędów o rozkładzie  $\varepsilon \sim N_n(\mathbf{0}, \sigma^2 \mathbf{I})$

**Uwaga:**  $Y \sim N_n(\varphi, \sigma^2 \mathbf{I})$

Interesuje nas problem testowania hipotezy

$$H_0: \varphi \in \omega \subset \Omega \quad \text{przeciwko} \quad H_1: \varphi \in \Omega - \omega,$$

gdzie  $\omega$  jest pewną **właściwą podprzestrzenią liniową** przestrzeni  $\Omega$ ,

czyli  $\omega \subset \Omega$  i  $\dim(\omega) = p - r < \dim(\Omega)$ .

**Uwaga.** Zbiory  $\omega$  i  $\Omega$  są **podprzestrzeniami liniowymi**, czyli nie mogą to być zbiory ograniczone np. kule w  $R^n$ .

Po wyborze parametryzacji (bazy w przestrzeni  $R^n$ ) rozważany model liniowy można zapisać w postaci gaussowskiego modelu liniowego

$$Y = X\beta + \varepsilon$$

$$\varepsilon \sim N(\mathbf{0}, \sigma^2 \mathbf{I})$$

$X$  jest tu nielosową macierzą wymiaru  $(n, p)$  zwaną **macierzą planu eksperymentu** o kolumnach  $x_i$ ,  $i=0, \dots, p-1$  przy czym  $x_0 = \mathbf{1}$ , liniowo niezależnych tzn.  $\text{rz}(X) = p$ .

Wektor  $\varphi = E(Y) \in \Omega \subset R^n$  przy czym  $\Omega = \text{span}\{x_i; i=0, \dots, p-1\} = \text{Im}(X)$

Testować będziemy hipotezę liniową

$$H_0: G\beta = \mathbf{0} \quad \text{przeciwko} \quad H_1: G\beta \neq \mathbf{0},$$

gdzie  $G$  jest macierzą  $(r, p)$   $r < p$  pełnego rzędu tzn.  $\text{rz}(G) = r$ , która specyfikuje podprzestrzeń

$\omega = \text{Im}(X|_{\text{Ker}G})$  ( $\dim(\omega) = p - r$ ) będącą obrazem obcięcia przekształcenia  $X$  do jądra przekształcenia  $G$ .

**Przykład.** Porównywanie  $k$  średnich

Rozważmy  $k$  niezależnych prób prostych

$$\begin{array}{c} Y_{11}, \dots, Y_{1n_1} \\ \vdots \\ Y_{k1}, \dots, Y_{kn_k} \end{array} \quad \text{pochodzących odpowiednio z rozkładów}$$

$N(m_1, \sigma^2), \dots, N(m_k, \sigma^2)$ . Można więc napisać

$$Y_{ij} = m_i + \varepsilon_{ij}, \quad i=1, \dots, k, \quad j=1, \dots, n_i,$$

przy czym  $\varepsilon_{ij}$  są niezależnymi zmiennymi losowymi o identycznych rozkładach normalnych  $N(0, \sigma^2)$ .

Testowana jest hipoteza  $H_0: m_1 = \dots = m_k$  wobec alternatywy  $H_1: \sim H_0$

W zapisie macierzowym

$$\begin{bmatrix} Y_{11} \\ \vdots \\ Y_{1n_1} \\ Y_{21} \\ \vdots \\ Y_{2k_2} \\ \vdots \\ Y_{k1} \\ \vdots \\ Y_{kn_k} \end{bmatrix} = m_1 \begin{bmatrix} 1 \\ \vdots \\ 1 \\ 0 \\ \vdots \\ 0 \\ \vdots \\ 0 \\ \vdots \\ 0 \end{bmatrix} + m_2 \begin{bmatrix} 0 \\ \vdots \\ 0 \\ 1 \\ \vdots \\ 1 \\ \vdots \\ 0 \\ \vdots \\ 0 \end{bmatrix} + \dots + m_k \begin{bmatrix} 0 \\ \vdots \\ 0 \\ 0 \\ \vdots \\ 0 \\ \vdots \\ 1 \\ \vdots \\ 1 \end{bmatrix} + \begin{bmatrix} \varepsilon_{11} \\ \vdots \\ \varepsilon_{1n_1} \\ \varepsilon_{21} \\ \vdots \\ \varepsilon_{2k_2} \\ \vdots \\ \varepsilon_{k1} \\ \vdots \\ \varepsilon_{kn_k} \end{bmatrix}$$

$$\mathbf{Y} = m_1 \mathbf{x}_1 + m_2 \mathbf{x}_2 + \dots + m_k \mathbf{x}_k + \boldsymbol{\varepsilon}.$$

$$\mathbf{Y} = \mathbf{X} \mathbf{m} + \boldsymbol{\varepsilon}.$$

W tym problemie

$$\Omega = \text{span} \left\{ \begin{bmatrix} 1 \\ \vdots \\ 1 \\ 0 \\ \vdots \\ 0 \\ \vdots \\ 0 \\ \vdots \\ 0 \end{bmatrix}, \begin{bmatrix} 0 \\ \vdots \\ 0 \\ 1 \\ \vdots \\ 1 \\ \vdots \\ 0 \\ \vdots \\ 0 \end{bmatrix}, \dots, \begin{bmatrix} 0 \\ \vdots \\ 0 \\ 0 \\ \vdots \\ 0 \\ \vdots \\ 1 \\ \vdots \\ 1 \end{bmatrix} \right\}, \quad \dim \Omega = k, \quad \omega = \text{span} \left\{ \begin{bmatrix} 1 \\ \vdots \\ 1 \\ 1 \\ \vdots \\ 1 \\ \vdots \\ 1 \\ \vdots \\ 1 \end{bmatrix} \right\}, \quad \dim \omega = 1$$

### Inna parametryzacja w problemie porównywania $k$ średnich

Niech  $m = \frac{1}{k} \sum_{i=1}^k m_i$  będzie średnią arytmetyczną grupowych wartości oczekiwanych. Oznaczając

$\alpha_i = m_i - m$ ,  $i=1, \dots, k$  odchylenie wartości oczekiwanej w  $i$ -tej grupie od  $m$  otrzymujemy ograniczenie

$$\sum_{i=1}^k \alpha_i = 0.$$

Przy tej parametryzacji model można zapisać w postaci

$$Y_{ij} = m + \alpha_i + \varepsilon_{ij}, \quad i=1, \dots, k, \quad j=1, \dots, n_i, \quad \text{przy czym} \quad \sum_{i=1}^k \alpha_i = 0.$$

Jest to tzw. parametryzacja z sigma ograniczeniami

Parametr  $m$  reprezentuje średni poziom (poziom odniesienia) dla wszystkich obserwowanych zmiennych. Parametr  $\alpha_i$  jest odchyleniem wartości oczekiwanej w  $i$  tej grupie od poziomu odniesienia. Problem testowania równości  $k$  średnich sprowadza się w tym przypadku do testowania hipotezy

$H_0: \alpha_1 = \dots = \alpha_k$  wobec alternatywy  $H_1: \sim H_0$  przy czym  $\sum_{i=1}^k \alpha_i = 0$  (parametr  $m$  nas nie interesuje)

Przy tej parametryzacji przestrzenie  $\Omega$  i  $\omega$  wyglądają następująco

$$\Omega = \text{span} \left\{ \begin{bmatrix} 1 \\ \vdots \\ 1 \\ 1 \\ \vdots \\ 1 \\ \vdots \\ 1 \\ \vdots \\ 1 \end{bmatrix}, \begin{bmatrix} 1 \\ \vdots \\ 0 \\ \vdots \\ 0 \\ \vdots \\ 0 \\ \vdots \\ 0 \end{bmatrix}, \begin{bmatrix} 0 \\ \vdots \\ 1 \\ \vdots \\ 1 \\ \vdots \\ 0 \\ \vdots \\ 0 \end{bmatrix}, \dots, \begin{bmatrix} 0 \\ \vdots \\ 0 \\ \vdots \\ 0 \\ \vdots \\ 1 \\ \vdots \\ 1 \end{bmatrix} \right\}, \dim \Omega = k, \omega = \text{span} \left\{ \begin{bmatrix} 1 \\ \vdots \\ 1 \\ 1 \\ \vdots \\ 1 \\ \vdots \\ 1 \\ \vdots \\ 1 \end{bmatrix} \right\}, \dim \omega = 1$$

**Uwaga:** pierwszy z  $k+1$  wektorów jest sumą pozostałych.

**I jeszcze inna parametryzacja.** Przypuśćmy, że jedna z grup (dla ustalenia uwagi pierwsza) jest wyróżniona np. jest to **grupa kontrolna**. Przyjmując parametry

$$\beta_1 = m_1$$

$$\beta_i = m_i - m_1, \quad i=2, \dots, k$$

$$\begin{bmatrix} X_{11} \\ \vdots \\ X_{1m_1} \\ X_{21} \\ \vdots \\ X_{2k_2} \\ \vdots \\ X_{k1} \\ \vdots \\ X_{kn_k} \end{bmatrix} = \beta_1 \begin{bmatrix} 1 \\ \vdots \\ 1 \\ \vdots \\ 1 \\ \vdots \\ 1 \\ \vdots \\ 1 \\ \vdots \\ 1 \end{bmatrix} + \beta_2 \begin{bmatrix} 0 \\ \vdots \\ 0 \\ \vdots \\ 1 \\ \vdots \\ 1 \\ \vdots \\ 0 \\ \vdots \\ 0 \end{bmatrix} + \dots + \beta_k \begin{bmatrix} 0 \\ \vdots \\ 0 \\ \vdots \\ 0 \\ \vdots \\ 0 \\ \vdots \\ 1 \\ \vdots \\ 1 \end{bmatrix} + \begin{bmatrix} \varepsilon_{11} \\ \vdots \\ \varepsilon_{1m_1} \\ \varepsilon_{21} \\ \vdots \\ \varepsilon_{2k_2} \\ \vdots \\ \varepsilon_{k1} \\ \vdots \\ \varepsilon_{kn_k} \end{bmatrix}$$

Testowanie hipotezy o równości średnich sprowadza się do testowania hipotezy  $H_0: \beta_i = 0, \quad i=2, \dots, k$ .

Parametr  $\beta_1 = m_1$  wyznaczający „poziom odniesienia” nas nie interesuje.

**Parametryzacja z grupą kontrolną jest przykładem parametryzacji oszczędnej. Taka parametryzacja jest stosowana w R**

**Zadanie 1.** Linie lotnicze zainteresowane są porównaniem przydatności 3 wyświetlaczy dla kontrolerów ruchu lotniczego. Każdy wyświetlacz był sprawdzany w 5 symulowanych warunkach zagrożenia. Trzydziestu wysoko wykwalifikowanych kontrolerów z podobnym doświadczeniem zawodowym zostało wybranych do badań. Kontrolerzy zostali losowo przypisani do klatek (wyświetlacz, warunki zagrożenia) po dwóch do każdej klatki. Mierzono czas w sekundach potrzebny do opanowania sytuacji awaryjnej. Wyniki przedstawia tabela

Wyświetlacz	Warunki zagrożenia				
	1	2	3	4	5
1	18	31	22	39	15
	16	35	27	36	12
2	13	33	24	35	10
	15	30	21	38	16
3	24	42	40	52	28
	28	46	37	57	24

Chcemy odpowiedzieć na pytania

- czy wyświetlacze różnią się między sobą?
- czy warunki zagrożenia wpływają na przydatność danego wyświetlacza?

#### Model strukturalny efektów głównych (model addytywny)

$$Y_{ijk} = m + \alpha_i + \beta_j + \varepsilon_{ijk} \quad i=1,2,3; j=1,2,3,4,5; k=1,2$$

$$\sum_i \alpha_i = 0, \quad \sum_j \beta_j = 0, \quad \varepsilon_{ijk} \sim iid(N0, \sigma^2).$$

#### Funkcje aov() i lm() w analizie wariancji

Parametryzacja R

```

y x1 w2 x3 z2 z3 z4 z5
1 18 1 0 0 0 0 0 0
2 16 1 0 0 0 0 0 0
3 13 1 1 0 0 0 0 0
4 15 1 1 0 0 0 0 0
5 24 1 0 1 0 0 0 0
6 28 1 0 1 0 0 0 0
7 31 1 0 0 1 0 0 0
8 35 1 0 0 1 0 0 0
9 33 1 1 0 1 0 0 0
10 30 1 1 0 1 0 0 0
11 42 1 0 1 1 0 0 0
12 46 1 0 1 1 0 0 0
13 22 1 0 0 0 1 0 0
14 27 1 0 0 0 1 0 0
15 24 1 1 0 0 1 0 0
16 21 1 1 0 0 1 0 0
17 40 1 0 1 0 1 0 0
18 37 1 0 1 0 1 0 0
19 39 1 0 0 0 0 1 0
20 36 1 0 0 0 0 1 0
21 35 1 1 0 0 0 1 0
22 38 1 1 0 0 0 1 0
23 52 1 0 1 0 0 1 0
24 57 1 0 1 0 0 1 0
25 15 1 0 0 0 0 0 1
26 12 1 0 0 0 0 0 1
27 10 1 1 0 0 0 0 1
28 16 1 1 0 0 0 0 1
29 28 1 0 1 0 0 0 1
30 24 1 0 1 0 0 0 1

```

**Model strukturalny czynnikowy pełny (model pełny z interakcją Wyświetlacz × Zagrożenie)**

$$Y_{ijk} = m + \alpha_i + \beta_j + \gamma_{ij} + \varepsilon_{ijk} \quad i=1,2,3; j=1,2,3,4,5, \text{ przy czym}$$

$$\sum_i \alpha_i = 0, \sum_j \beta_j = 0, \sum_i \gamma_{ij} = 0 \text{ dla każdego } j, \sum_j \gamma_{ij} = 0 \text{ dla każdego } i.$$

	y	x1	x2	x3	x4	x5	x6	x7	x8	x9	x10	x11	x12	x13	x14	x15
1	18	1	0	0	0	0	0	0	0	0	0	0	0	0	0	0
2	16	1	0	0	0	0	0	0	0	0	0	0	0	0	0	0
3	13	1	1	0	0	0	0	0	0	0	0	0	0	0	0	0
4	15	1	1	0	0	0	0	0	0	0	0	0	0	0	0	0
5	24	1	0	1	0	0	0	0	0	0	0	0	0	0	0	0
6	28	1	0	1	0	0	0	0	0	0	0	0	0	0	0	0
7	31	1	0	0	1	0	0	0	0	0	0	0	0	0	0	0
8	35	1	0	0	1	0	0	0	0	0	0	0	0	0	0	0
9	33	1	1	0	1	0	0	0	1	0	0	0	0	0	0	0
10	30	1	1	0	1	0	0	0	1	0	0	0	0	0	0	0
11	42	1	0	1	1	0	0	0	1	0	0	0	0	0	0	0
12	46	1	0	1	1	0	0	0	0	1	0	0	0	0	0	0
13	22	1	0	0	0	1	0	0	0	0	0	0	0	0	0	0
14	27	1	0	0	0	1	0	0	0	0	0	0	0	0	0	0
15	24	1	1	0	0	1	0	0	0	0	1	0	0	0	0	0
16	21	1	1	0	0	1	0	0	0	0	1	0	0	0	0	0
17	40	1	0	1	0	1	0	0	0	0	0	1	0	0	0	0
18	37	1	0	1	0	1	0	0	0	0	0	1	0	0	0	0
19	39	1	0	0	0	0	1	0	0	0	0	0	0	0	0	0
20	36	1	0	0	0	0	1	0	0	0	0	0	0	0	0	0
21	35	1	1	0	0	0	1	0	0	0	0	0	1	0	0	0
22	38	1	1	0	0	0	1	0	0	0	0	0	1	0	0	0
23	52	1	0	1	0	0	1	0	0	0	0	0	0	1	0	0
24	57	1	0	1	0	0	1	0	0	0	0	0	0	1	0	0
25	15	1	0	0	0	0	0	1	0	0	0	0	0	0	0	0
26	12	1	0	0	0	0	0	1	0	0	0	0	0	0	0	0
27	10	1	1	0	0	0	0	1	0	0	0	0	0	0	1	0
28	16	1	1	0	0	0	0	1	0	0	0	0	0	0	1	0
29	28	1	0	1	0	0	0	1	0	0	0	0	0	0	0	1
30	24	1	0	1	0	0	0	1	0	0	0	0	0	0	0	1

## Uwagi o diagnostyce modeli ANOVA

### Testy jednorodności wariancji Bartletta i Levene,a

#### Test Levene'a jednorodności wariancji

Problem: na podstawie  $k$  **niezależnych** prób prostych z odpowiednich rozkładów normalnych

$$X_{11}, \dots, X_{1n_1} \quad \text{iid } N(m_1, \sigma_1^2)$$

$$X_{21}, \dots, X_{2n_2} \quad \text{iid } N(m_2, \sigma_2^2)$$

$$\dots \dots \dots$$

$$X_{k1}, \dots, X_{kn_k} \quad \text{iid } N(m_k, \sigma_k^2)$$

zweryfikować hipotezę

$$H_0: \sigma_1^2 = \sigma_2^2 = \dots = \sigma_k^2$$

wobec alternatywy

$$H_1: \sim H_0 \text{ (inaczej } \exists(i,j), \sigma_i^2 \neq \sigma_j^2).$$

Test Levene'a.

Definiujemy absolutne odchylenia  $D_{ij} = |X_{ij} - \bar{X}_{i\cdot}|$ , gdzie  $\bar{X}_{i\cdot} = \frac{1}{n_i} \sum_{j=1}^{n_i} X_{ij}$  i obliczamy wartość

statystyki testowej

$$L = \frac{n-k}{n-1} \frac{\sum_{i=1}^k (\bar{D}_{i\cdot} - \bar{D}_{\cdot\cdot})^2}{\sum_{i=1}^k \sum_{j=1}^{n_i} (D_{ij} - \bar{D}_{i\cdot})^2},$$

która jest statystyką ANOVA dla absolutnych odchylen, mającą przy prawdziwości  $H_0$  rozkład Snedecora Fishera  $F_{n-k, n-1}$

Test Levene'a jest mniej czuły na odstające obserwacje i na odchylenia od normalności niż test Bartletta. Jest to spowodowane użyciem absolutnych odchylen zamiast kwadratów odchylen.

### Krótki przegląd konserwatywnych testów post hoc.

Wszystkie rozważane testy konserwatywne prowadzą do następującej reguły

**Średnie  $m_i$  i  $m_j$  są istotnie różne jeżeli  $|\bar{Y}_i - \bar{Y}_j| >$  wartość progowa (właściwa dla danego testu)**

Rozważane testy kontrolują prawdopodobieństwa różnych błędów

**Test Fishera NIR** (najmniejsza istotna różnica) (inaczej *LSD* least significant difference).

#### Algorytm

1. Przeprowadzić ANOVA w celu przetestowania  $H_0: m_1 = \dots = m_k$  przeciwko alternatywie  $H_1$ : co najmniej dwie średnie różnią się między sobą
2. Jeżeli nie ma podstaw do odrzucenia  $H_0$  kończymy analizę.
3. Jeżeli  $H_0$  została odrzucona, to definiujemy najmniejszą istotną różnicę *NIR* (*LSD*) pomiędzy średnimi próbkowymi, którą należy zaobserwować, aby uznać odpowiadające im średnie w odpowiednich populacjach za istotnie różne.
4. Dla wyspecyfikowanej wartości  $\alpha$  w celu porównania  $m_i$  z  $m_j$  obliczamy *NIR* wg wzoru

$$NIR = t_{1-\frac{\alpha}{2}, \sum_{i=1}^k n_i} \sqrt{S_w^2 \left( \frac{1}{n_i} + \frac{1}{n_j} \right)}, \text{ gdzie } t_{1-\frac{\alpha}{2}, \sum_{i=1}^k n_i} \text{ jest kwantylem rzędu } 1-\frac{\alpha}{2} \text{ rozkładu } t\text{-Studenta o}$$

$\sum_{i=1}^k n_i$  stopniach swobody a  $S_w^2$  jest sumą kwadratów wewnątrz grup (z ANOVA)

5. Porównujemy wszystkie pary średnich próbkowych. Jeżeli  $|\bar{Y}_i - \bar{Y}_j| > NIR$ , to uznajemy, że  $m_i$  i  $m_j$  są istotnie różne.
6. Dla wszystkich porównań par prawdopodobieństwo błędu I rodzaju jest ustalone na poziomie  $\alpha$

Komputerowe symulacje wykazały, że jeśli stosujemy test Fishera w połączeniu z ANOVA (tak jak opisano powyżej), to poziom istotności złożonego testu porównań wielokrotnych jest w przybliżeniu równy poziomowi istotności testu  $F$  (procedura Fisher's protected LSD).

Jeżeli stosujemy test *NIR* samodzielnie (bez ANOVA), to kontrolujemy jedynie błąd przy pojedynczych porównaniach (**per comparison**). Odpowiada to wielokrotnemu stosowaniu testu istotności różnicy dwóch średnich opartego na statystyce  $t$ -Studenta, przy czym estymator wariancji opieramy na całej próbie (z powodu jednorodności wariancji w grupach) a nie tylko na obserwacjach z aktualnie porównywanych grup.

**Test W Tukey'a** oparty jest na studentyzowanym rozstępie  $\sqrt{n} \frac{\bar{Y}_{\max} - \bar{Y}_{\min}}{S}$  pomiędzy średnimi próbkowymi.

#### Algorytm (dla jednakowo licznych grup)

1. Uporządkować średnie próbkowe  $\bar{Y}_i$ ,  $i=1, \dots, k$
2.  $m_i$  i  $m_j$  są istotnie różne jeżeli  $|\bar{Y}_i - \bar{Y}_j| \geq W$ , gdzie  $W = q_\alpha(k, df) \sqrt{\frac{S_w^2}{n}}$ ,  $df$  jest liczbą stopni swobody w  $S_w^2$ ,  $n$  ilość obserwacji w każdej grupie,  $q_\alpha(k, df)$  prawostronna wartość krytyczna studentyzowanego rozstępu (tablice).
3. Prawdopodobieństwo zaobserwowania fałszywie istotnej różnicy dla porównań parami jest równe  $\alpha$ .

Jeżeli grupy nie są równoliczne, to zamiast  $W$  należy użyć  $W_{ij} = q_\alpha(k, df) \sqrt{\frac{S_w^2}{2} \left(\frac{1}{n_i} + \frac{1}{n_j}\right)}$

**Test Tukey'a kontroluje błąd I rodzaju dla wszystkich porównań parami**, tzn. prawdopodobieństwo (przy  $H_0: m_1 = \dots = m_k$ ) zaobserwowania takiego układu średnich próbkowych  $\bar{Y}_i$ ,  $i=1, \dots, k$ , dla którego przynajmniej jedna różnica pomiędzy średnimi  $m_i$  jest fałszywie uznana za istotną jest równe  $\alpha$ . Jeżeli jest  $k$  grup, to test Tukey'a kontroluje łączny błąd I rodzaju dla  $\binom{k}{2}$  porównań jednocześnie (**per experiment**)

**Test Newman-Keulsa** jest modyfikacją testu Tukey'a wykorzystującą informację o ilości miejsc pomiędzy badanymi średnimi w uporządkowanym ciągu średnich.

#### Algorytm (dla jednakowo licznych grup)

1. Uporządkować średnie próbkowe  $\bar{Y}_i$ ,  $i=1, \dots, k$
2. Dla dwóch średnich  $\bar{Y}_i$  i  $\bar{Y}_j$  odległych o  $r$  miejsc odpowiadające im średnie  $m_i$  i  $m_j$  są istotnie różne jeżeli  $|\bar{Y}_i - \bar{Y}_j| \geq W_r$ , gdzie  $W_r = q_\alpha(r, df) \sqrt{\frac{S_w^2}{n}}$ ,  $df$  jest liczbą stopni swobody w  $S_w^2$ ,  $n$  ilość obserwacji w każdej grupie,  $q_\alpha(r, df)$  prawostronna wartość krytyczna studentyzowanego rozstępu (tablice). **Uwaga.** Przyjmujemy, że sąsiednie średnie odległe są o 2 a skrajne o  $k$  czyli  $r_{ij} = |\text{ranga}(\bar{Y}_i) - \text{ranga}(\bar{Y}_j)| + 1$

Jeżeli grupy nie są równoliczne, to zamiast  $W_r$  należy użyć  $W_{rij} = q_\alpha(r, df) \sqrt{\frac{S_w^2}{2} \left(\frac{1}{n_i} + \frac{1}{n_j}\right)}$ .

**Test wielokrotnych rozstępów Duncana** jest podobny do dwóch poprzednich, gdyż jest oparty na studentyzowanym rozstępie, lecz różni się od poprzednich tym, że zmienny jest poziom istotności

przy poszczególnych porównaniach. Gdy średnie próbkowe zostaną uporządkowane i są odległe o  $r$  miejsc, to istotność różnicy odpowiadających im średnich w populacjach jest testowana na poziomie  $1-(1-\alpha)^{r-1}$

### Algorytm (dla jednakowo liczych grup)

1. Uporządkować średnie próbkowe  $\bar{Y}_i$ ,  $i=1, \dots, k$
2. Średnie  $m_i$  i  $m_j$  są istotnie różne jeżeli odpowiadające im średnie próbkowe odległe o  $r$  miejsc spełniają warunek  $|\bar{Y}_i - \bar{Y}_j| \geq W_r'$ , gdzie  $W_r' = q'_\alpha(r, df) \sqrt{\frac{S_w^2}{n}}$ ,  $df$  jest liczbą stopni swobody w  $S_w^2$ ,  $n$  ilość obserwacji w każdej grupie,  $q'_\alpha(r, df)$  prawostronna wartość krytyczna testu Duncana (tablice).

Jeżeli grupy są w przybliżeniu równoliczne, to zamiast  $n$  należy użyć  $\tilde{n} = \frac{k}{\frac{1}{n_1} + \frac{1}{n_2} + \dots + \frac{1}{n_k}}$ .

### Test Scheffe'go

1. Rozważmy dowolny kontrast  $I_a = \sum_{i=1}^k a_i m_i$ ,  $\sum_{i=1}^k a_i = 0$ . Chcemy zweryfikować hipotezę

$$H_0: \forall \mathbf{a} \quad I_a = \sum_{i=1}^k a_i m_i = 0 \text{ wobec alternatywy } H_1: \exists \mathbf{a} \quad I_a \neq 0.$$

2. Rozważmy dowolny kontrast próbkowy  $\hat{I}_a = \sum_{i=1}^k a_i \bar{X}_i$  dla którego nieobciążonym

$$\text{estymatorem wariancji jest } \hat{V}(\hat{I}) = \frac{1}{n-k} S_w^2 \sum_{i=1}^k \frac{1}{n_i} a_i^2$$

3. Kontrast  $I_a$  uznamy za istotny (istotnie różny od 0), gdy

$$\left| \frac{(n-k)\hat{I}_a^2}{(k-1)S_w^2 \sum_{i=1}^k \frac{a_i^2}{n_i}} \right| > F_{k-1, n-k, 1-\alpha},$$

gdzie  $F_{k-1, n-k, 1-\alpha}$  jest kwantylem rzędu  $1-\alpha$  rozkładu centralnego Snedecora Fishera  $F$ .

4. Prawdopodobieństwo zaobserwowania fałszywie istotnego kontrastu jest równe  $\alpha$ .

Test Scheffe'go kontroluje łączny błąd większej liczby porównań niż test Tukey'a, więc jest bardziej konserwatywny (trudniej odrzucić  $H_0$ ). W teście Scheffego na poziomie  $\alpha$  prawdopodobieństwo zaobserwowania fałszywie istotnego kontrastu nie przekracza  $\alpha$ .

**Porównania testów.** W poniższej tabeli zestawiono wartości progowe po przekroczeniu których różnice uznajemy za istotne. Rozważono porównanie 6 grup, przy czym próbka dla każdej grupy liczy  $n=5$  elementów a  $S_w^2=2451$ .

Test	Liczba miejsc pomiędzy średnimi $r$				
	2	3	4	5	6
Fisher NIR	64.63	64.63	64.63	64.63	64.63
Tukey	96.75	96.75	96.75	96.75	96.75
Newman-Keuls	64.65	78.16	86.35	92.33	96.75
Duncan	64.65	67.97	69.74	71.29	72.62

Scheffe	196.31	196.31	196.31	196.31	196.31
---------	--------	--------	--------	--------	--------

Widać, że najbardziej konserwatywny jest test Scheffe'go a najmniej konserwatywny (czyli najbardziej czuły) test *NIR*. Z uwagi na kontrolę błędu godny polecenia jest test Tukey'a. Symulacje preferują raczej test Newmana-Keulusa. Z uwagi na czułość duże uznanie wśród praktyków wzbudził test Duncana.

Na uwagę zasługuje jeszcze test **Dunetta** wielokrotnych porównań z wyróżnioną grupą kontrolną  
**#Testy post-hoc z pakietu agricolé.**

## ANCOVA

W pewnym eksperymencie rolniczym porównywano plony *Y* trzech różnych odmian kukurydzy. Mierzono również opady deszczu *X*. Zakładając, że odmiany zostały losowo przypisane do poletek zbadać, czy odmiany istotnie różnią się między sobą wysokością plonu.

Odmiana I		Odmiana II		Odmiana III	
X	Y	X	Y	X	Y
14	68,5	30	81,5	15	70,0
36	83,0	33	85,2	11	84,0
31	83,0	39	87,1	42	75,2
27	66,5	26	69,3	14	67,5
15	58,1	43	73,5	15	66,0
17	82,4	31	65,5	42	79,1
		18	73,4	26	75,5
		14	56,1		

### Najpierw zacniemy od modelu jednakowych nachyleń

$$Y_{ij} = m + \alpha_i + \beta X_{ij} + \varepsilon_{ij}$$

$$\sum_{i=1}^3 \alpha_i = 0$$

$$\varepsilon_{ij} \sim N(0, \sigma^2)$$

Parametryzacja R

$$\begin{bmatrix} Y_1 \\ Y_2 \\ Y_3 \end{bmatrix} = \begin{bmatrix} 1 \\ 1 \\ 1 \end{bmatrix} m + \begin{bmatrix} 0 \\ 1 \\ 0 \end{bmatrix} \alpha_2 + \begin{bmatrix} 0 \\ 0 \\ 1 \end{bmatrix} \alpha_3 + \begin{bmatrix} X_1 \\ X_2 \\ X_3 \end{bmatrix} \beta + \begin{bmatrix} \varepsilon_1 \\ \varepsilon_2 \\ \varepsilon_3 \end{bmatrix}$$

### Model niejednakowych nachyleń

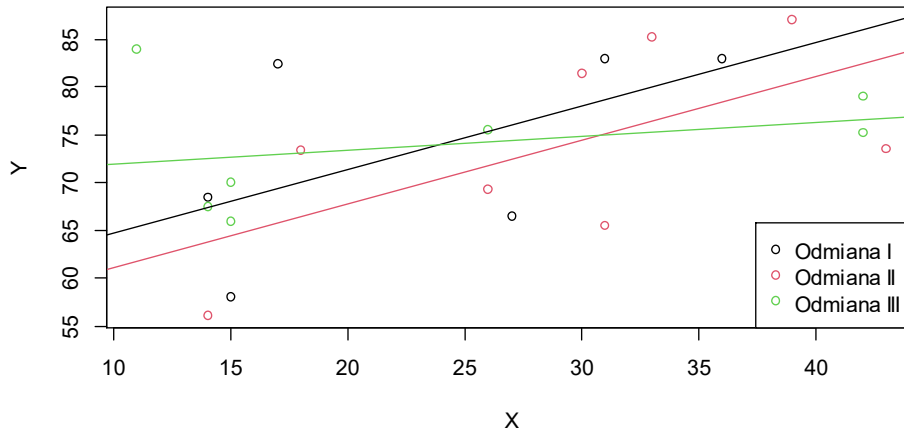
$$Y_{ij} = m + \alpha_i + \beta X_{ij} + \gamma_i X_{ij} + \varepsilon_{ij}$$

$$\sum_{i=1}^3 \alpha_i = 0, \quad \sum_{i=1}^3 \gamma_i = 0$$

$$\varepsilon_{ij} \sim N(0, \sigma^2)$$

Parametryzacja R

$$\begin{bmatrix} Y_1 \\ Y_2 \\ Y_3 \end{bmatrix} = \begin{bmatrix} 1 \\ 1 \\ 1 \end{bmatrix} m + \begin{bmatrix} 0 \\ 1 \\ 0 \end{bmatrix} \alpha_2 + \begin{bmatrix} 0 \\ 0 \\ 1 \end{bmatrix} \alpha_3 + \begin{bmatrix} X_1 \\ X_2 \\ X_3 \end{bmatrix} \beta + \begin{bmatrix} 0 \\ X_1 \\ 0 \end{bmatrix} \gamma_2 + \begin{bmatrix} 0 \\ 0 \\ X_3 \end{bmatrix} \gamma_3 + \begin{bmatrix} \varepsilon_1 \\ \varepsilon_2 \\ \varepsilon_3 \end{bmatrix}$$



**Problem.** Jak usunąć wpływ zmiennej towarzyszącej na odpowiedź-ćwiczenia

## Testowanie sekwencyjne i brzegowe w modelach liniowych.

Rozważmy model 2 czynnikowej ANOVA z czynnikami **A** i **B**. Używając składni języka R model efektów głównych z odpowiedzią **y** można zapisać w postaci **y~A+B** natomiast pełny model czynnikowy w postaci **y~A\*B** co jest równoważne zapisowi **y~A+B+A:B**. Po przetestowaniu hipotezy  $H_0$  o nieistotnym wpływie na odpowiedź wszystkich czynników polegającym na porównaniu modeli **y~A\*B** lub **y~A+B** z modelem pustym **y~1** i odrzuceniu hipotezy zerowej pojawia się problem testowania istotności poszczególnych czynników i ich interakcji. **Funkcjonują 2 zasadnicze strategie testowania:**

**testowanie typu I – testowanie sekwencyjne (sequential),**

**testowanie typu III –testowanie brzegowe (marginal).**

Testowanie typu I (sekwencyjne) polega na sprawdzeniu istotności kolejnych zmiennych po dołączeniu do coraz większego modelu. W pierwszym kroku testowana jest istotność czynnika A przez porównanie (testem F ANOVA ) modelu **y~A** z modelem pustym **y~1**. W drugim kroku sprawdzana jest istotność drugiego czynnika poprzez porównanie modelu **y~A+B** z modelem **y~A**. Wreszcie w trzecim kroku sprawdzana jest istotność interakcji **A:B** przez porównanie modelu **y~A+B+A:B** z modelem **y~A+B**. Wyniki testowania mogą zależeć (i zwykle zależą) od kolejności zmiennych. Testowanie sekwencyjne dokonywane jest z użyciem funkcji `anova()`

Testowanie typu III (brzegowe) polega na sprawdzeniu istotności zmiennych przez porównanie modelu pełnego z modelem po usunięciu efektu danej zmiennej. Istotność czynnika **A** jest testowana przez porównanie (testem F) modelu  $y \sim A+B+A:B$  z modelem  $y \sim B+A:B$ , podobnie istotność czynnika **B** jest sprawdzana przez porównanie modelu  $y \sim A+B+A:B$  z modelem  $y \sim A+A:B$ . Wreszcie istotność interakcji jest testowana przez porównanie modelu  $y \sim A+B+A:B$  z  $y \sim A+B$ . Wyniki testowania nie zależą od kolejności testowania zmiennych.

Testowanie brzegowe dokonywane jest przy użyciu funkcji `summary()`

**Eksperyment numeryczny**- wygenerujemy 100 obserwacji z 3 wymiarowego rozkładu normalnego

$N(\mathbf{m}, \Sigma)$  z  $\mathbf{m} = \begin{bmatrix} 0 \\ 0 \\ 0 \end{bmatrix}$  i  $\Sigma = \begin{bmatrix} 1 & 0.25 & 0.25 \\ 0.25 & 1 & 0.75 \\ 0.25 & 0.75 & 1 \end{bmatrix}$ . Macierz kowariancyjna jest w tym przypadku równa

macierzy korelacyjnej. Wybraliśmy macierz korelacyjną tak aby 2 i 3 składowa były silnie skorelowane.

Dokonując dekompozycji Choleskiego macierzy  $\Sigma = \mathbf{U}^T \mathbf{U} = \mathbf{L} \mathbf{L}^T$ , gdzie  $\mathbf{U}$  jest macierzą trójkątną górną a

$\mathbf{L}$  trójkątną dolną za pomocą funkcji `chol()` przekształcamy 100 elementową próbkę prostą z rozkładu  $N(\mathbf{0}, \mathbf{I})$

w 100 elementową próbkę prostą z rozkładu  $N(\mathbf{m}, \Sigma)$  za pomocą transformacji  $\mathbf{Y} = \mathbf{U}^T \mathbf{X}$ .

Podsumowując, w rozważanym przykładzie jest zależność pomiędzy  $Y$  a  $X$  oraz  $Y$  i  $Z$ . Zależność ta jest widoczna, gdy porównujemy model pusty z modelem z tylko jedną zmienną objaśnianą. Słabiej ta zależność jest widoczna, gdy model pusty porównujemy z modelem z dwiema zmiennymi objaśniającymi (model pełny) ponieważ każda ze zmiennych objaśniających wyraża tę samą część zmienności zmiennej objaśnianej  $Y$ . Widać to w testach brzegowych, które pokazują że zmienna  $X$  ma nieistotny wpływ na  $Y$  jeżeli jest uwzględniona w modelu zmienna  $Z$  (i symetrycznie). W sytuacji gdy zmienne objaśniane są od siebie zależne należy starannie wybrać model