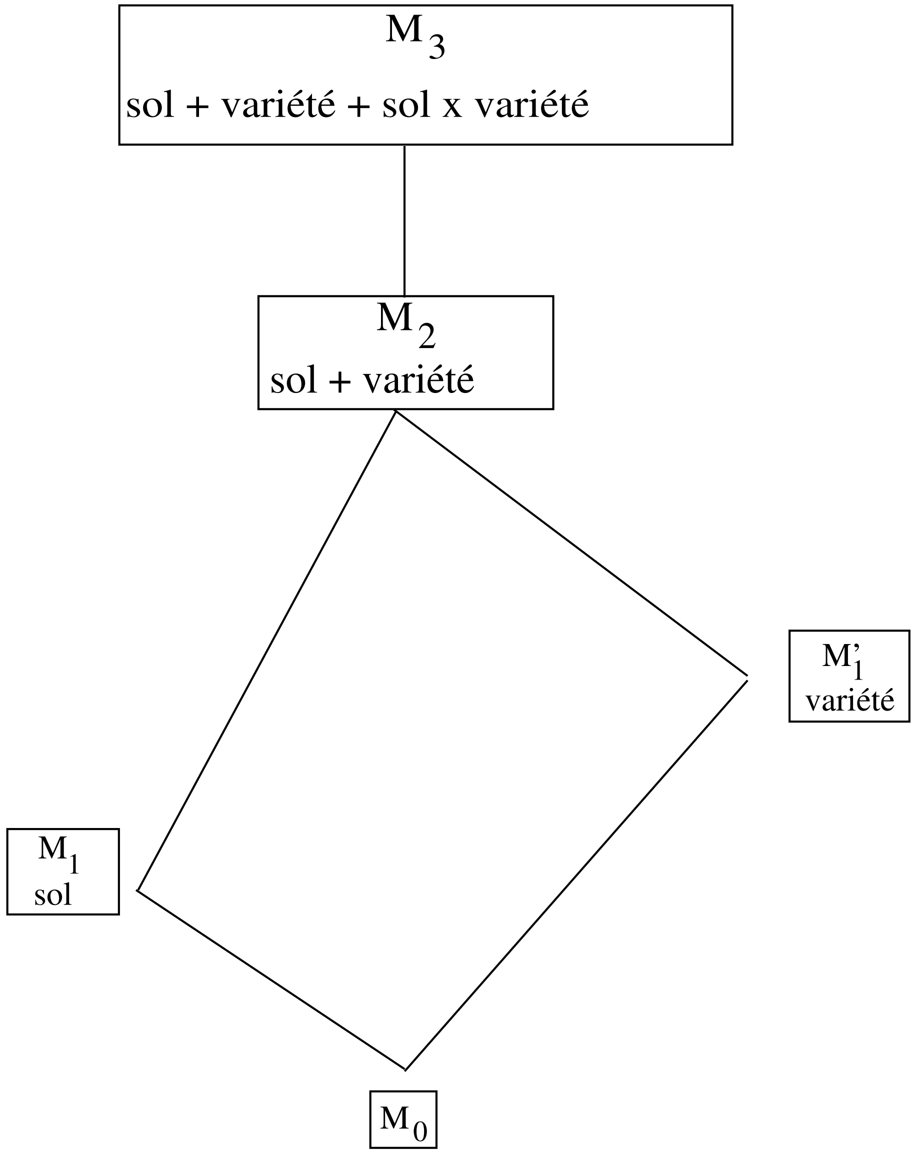
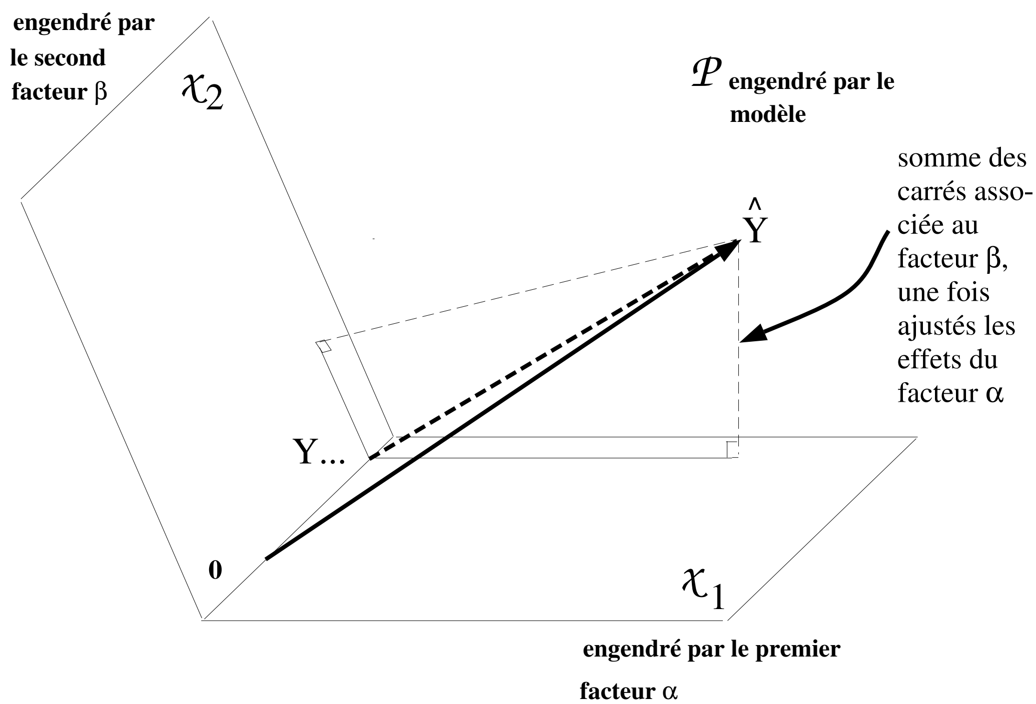
SCE type I :



$$\mathrm{SCE}_{\mathrm{M}_0} - \mathrm{SCE}_{\mathrm{M}'_{1_{\text{variété}}}} \neq \mathrm{SCE}_{\mathrm{M}_{1_{\text{sol}}}} - \mathrm{SCE}_{\mathrm{M}_{2_{\text{sol+variété}}}}$$
$$\mathrm{SCE}_{\mathrm{M}_0} - \mathrm{SCE}_{\mathrm{M}_{1_{\text{sol}}}} \neq \mathrm{SCE}_{\mathrm{M}'_{1_{\text{variété}}}} - \mathrm{SCE}_{\mathrm{M}_{2_{\text{sol+variété}}}}$$

Illustration géométrique



Hypothèses testées : sommes de type I

germination = sol variété sol $*$ variété

sol $\qquad\qquad\qquad\qquad\qquad\qquad\qquad\qquad\qquad\qquad$ H_0

$$F = \frac{\left(SCE_{M_0} - SCE_{M_1}\right)/\ I-1}{SCE_{M_3}/\ N-IJ} \qquad \begin{cases} \forall\ i \\ \alpha_i + \dfrac{\sum_j n_{ij}\beta_j}{n_{i+}} + \dfrac{\sum_j n_{ij}\gamma_{ij}}{n_{i+}} = 0 \end{cases}$$

variété

$$F = \frac{\left(SCE_{M_1} - SCE_{M_2}\right)/\ J-1}{SCE_{M_3}/\ N-IJ} \qquad \begin{cases} \forall\ j \\ \beta_j + \quad f\left(n_{ij}, \gamma_{ij}\right) = 0 \end{cases}$$

sol $*$ variété

$$F = \frac{\left(SCE_{M_2} - SCE_{M_3}\right)/\ IJ-1}{SCE_{M_3}/\ N-IJ} \qquad \begin{cases} \forall\ (i,j) \\ \gamma_{ij} = \quad 0 \end{cases}$$

germination = variété sol sol $*$ variété

variété

$$F = \frac{\left(SCE_{M_0} - SCE_{M'_1}\right)/\ J-1}{SCE_{M_3}/\ N-IJ} \qquad \begin{cases} \forall\ j \\ \beta_j + \dfrac{\sum_i n_{ij}\ \alpha_i}{n_{+j}} + \dfrac{\sum_i n_{ij}\ \gamma_{ij}}{n_{+j}} = 0 \end{cases}$$

sol

$$F = \frac{\left(SCE_{M'_1} - SCE_{M_2}\right)/\ I-1}{SCE_{M_3}/\ N-IJ} \qquad \begin{cases} \forall\ i \\ \alpha_i + \quad g\left(n_{ij}, \gamma_{ij}\right) = 0 \end{cases}$$

sol $*$ variété

$$F = \frac{\left(SCE_{M_2} - SCE_{M_3}\right)/\ IJ-1}{SCE_{M_3}/\ N-IJ} \qquad \begin{cases} \forall\ (i,j) \\ \gamma_{ij} = \quad 0 \end{cases}$$

Hypothèses testées : sommes de type III

	H_0
sol	$\begin{cases} \forall\, i \\ \alpha_i + \dfrac{\sum_j \gamma_{ij}}{J} = 0 \end{cases}$
variété	$\begin{cases} \forall\, j \\ \beta_j + \dfrac{\sum_i \gamma_{ij}}{I} = 0 \end{cases}$
interaction	$\begin{cases} \forall\, (i, j) \\ \gamma_{ij} = 0 \end{cases}$

Toujours utiliser les <u>sommes du type III</u>

On ne teste jamais un <u>effet additif seul</u>

On teste un effet <u>principal</u>

<u>Seule l'interaction</u> est testée <u>indépendamment</u> des <u>autres effets</u>

Comment étudier les effets des facteurs et non les effets principaux ?

1) Ne pas déclarer l'interaction et tester le modèle additif

2) Tester chaque effet seul

Commandes SAS

Exemple des tournesols

```
data trv ;
infile 'tourn' ;
input test orig huile ;
run ;
proc glm ;
class test orig ;
model huile = orig test / ss3 ;
run ;
```

① structuration en A et B

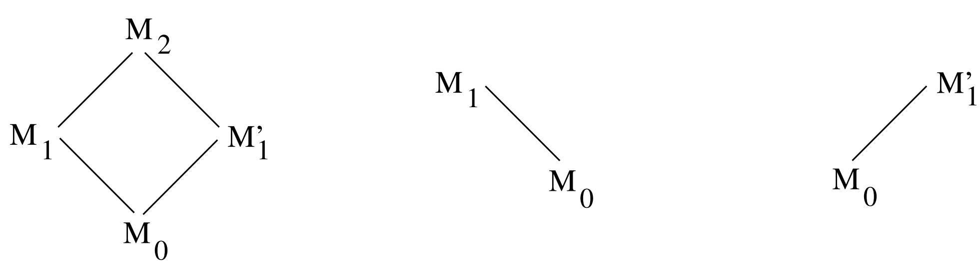
```
proc glm ;
class orig ;
model huile = orig / ss3 ;
```

② structuration en A

```
proc glm ;
class test ;
model huile= test / ss3 ;
```

③ structuration en B

① ② ③



①	②	③
modèle : $M_2 - M_0$ effet A : $M_2 - M'_1$ effet B : $M_2 - M_1$	modèle = effet A $M_1 - M_0$	modèle = effet B $M'_1 - M_0$
3 tests de F intrinsèques	1 test de F intrinsèque	1 test de F intrinsèque

Sortie SAS

Exemple des tournesols

General Linear Models Procedure
Class Level Information

Class	Levels	Values
TEST	2	1 2
ORIG	3	1 2 3

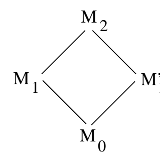
Number of observations in data set = 12

Dependent Variable : HUILE

Source	DF	Sum of Squares	Mean Square	F Value	Pr > F
Model	3	43.23322500	14.41107500	8.79	0.0065
Error	8	13.11486667	1.63935833		
Corrected Total	11	56.34809167			

R-Square	C.V.	Root MSE	DONNEE Mean
0.767253	2.766831	1.280374	46.27583

1



Dependent Variable : HUILE

Source	DF	Type III SS	Mean Square	F Value	Pr > F
ORIG	2	33.74581667	16.87290833	10.29	0.0061
TEST	1	9.48740833	9.48740833	5.79	0.0428

Class Level Information

Class	Levels	Values
ORIG	3	1 2 3

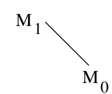
Number of observations in data set = 12

Dependent Variable : HUILE

Source	DF	Sum of Squares	Mean Square	F Value	Pr > F
Model	2	33.74581667	16.87290833	6.72	0.0164
Error	9	22.60227500	2.51136389		
Corrected Total	11	56.34809167			

R-Square	C.V.	Root MSE	DONNEE Mean
0.598881	3.424527	1.584728	46.27583

2



Dependent Variable : HUILE

Source	DF	Type III SS	Mean Square	F Value	Pr > F
ORIG	2	33.74581667	16.87290833	6.72	0.0164

Class Level Information

Class	Levels	Values
TEST	2	1 2

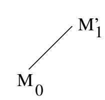
Number of observations in data set = 12

Dependent Variable : HUILE

Source	DF	Sum of Squares	Mean Square	F Value	Pr > F
Model	1	9.48740833	16.87290833	2.02	0.1852
Error	10	46.86068333	4.68606833		
Corrected Total	11	56.34809167			

R-Square	C.V.	Root MSE	DONNEE Mean
0.168371	4.677891	2.164733	46.27583

3



Dependent Variable : HUILE

Source	DF	Type III SS	Mean Square	F Value	Pr > F
TEST	1	9.48740833	9.48740833	2.02	0.1852

Sortie SAS

Exemple des carottes

General Linear Models Procedure

Class Level Information

Class	Levels	Values
SOL	2	1 2
VAR	3	1 2 3

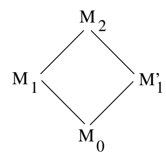
Number of observations in data set = 15

Dependent Variable : GERMINA

Source	DF	Sum of Squares	Mean Square	F Value	Pr > F
Model	3	177.2340426	59.0780142	1.90	0.1888
Error	11	342.7659574	31.1605416		
Corrected Total	14	520.0000000			

R-Square	C.V.	Root MSE	DONNEE Mean
0.340835	37.21442	5.582163	15.00000

①



Dependent Variable : GERMINA

Source	DF	Type III SS	Mean Square	F Value	Pr > F
SOL	1	83.9007092	83.9007092	2.69	0.1291
VAR	2	124.7340426	62.3670213	2.00	0.1814

Class Level Information

Class	Levels	Values
SOL	2	1 2

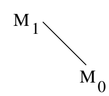
Number of observations in data set = 15

Dependent Variable : GERMINA

Source	DF	Sum of Squares	Mean Square	F Value	Pr > F
Model	1	52.50000000	52.50000000	1.46	0.2485
Error	13	467.50000000	39.96153846		
Corrected Total	14	520.00000000			

R-Square	C.V.	Root MSE	DONNEE Mean
0.100962	39.97863	5.996794	15.00000

②



Dependent Variable : GERMINA

Source	DF	Type III SS	Mean Square	F Value	Pr > F
SOL	1	52.50000000	52.50000000	1.46	0.2485

Class Level Information

Class	Levels	Values
VAR	3	1 2 3

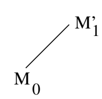
Number of observations in data set = 15

Dependent Variable : GERMINA

Source	DF	Sum of Squares	Mean Square	F Value	Pr > F
Model	2	93.33333333	46.66666667	1.31	0.3051
Error	12	426.66666667	35.55555556		
Corrected Total	14	520.00000000			

R-Square	C.V.	Root MSE	DONNEE Mean
0.179487	39.75232	5.962848	15.00000

③



Dependent Variable : GERMINA

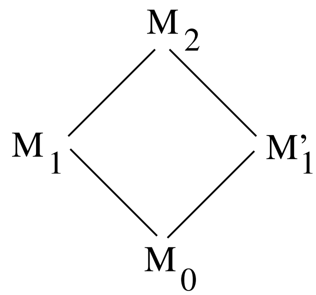
Source	DF	Type III SS	Mean Square	F Value	Pr > F
VAR	2	93.33333333	46.66666667	1.31	0.3051

Exemple des tournesols

① modèle significatif

effet A significatif (origine)

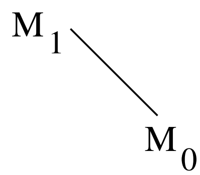
effet B significatif (testeur)



② modèle significatif

⇕

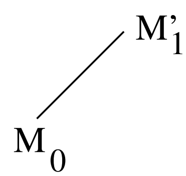
effet A significatif (origine)



③ modèle non significatif

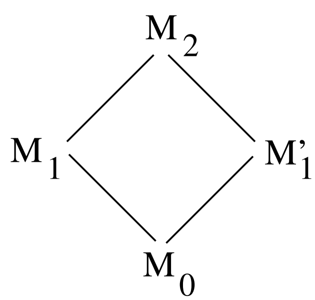
⇕

effet B non significatif (testeur)



Exemple des carottes

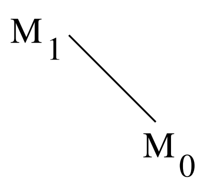
① modèle non significatif



② modèle non significatif

⇕

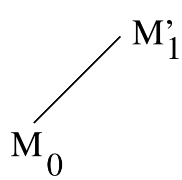
effet A non significatif (sol)



③ modèle non significatif

⇕

effet B non significatif (variété)



Réduire la complexité du modèle pour
tester les facteurs

$$\Downarrow$$

tests de F intrinsèques des facteurs

mais

perte de puissance

- $\nearrow$ variance d'erreur
- $\searrow$ R^2
- $\nearrow$ C.V

Analyse de Variance à deux facteurs croisés
Synthèse

A	B	A x B
(I)	(J)	(IJ)

Test du modèle

- N.S. : on s'arrête
- S. : il existe une structuration en AB (A ? , B ? , A x B ?)
  - Test de l'interaction
    - N.S.

      ⇓

      le test de A
      le test de B
      sont sans ambiguité

      ⇓

      si A S. : comparaison multiple des μ_i

      si B S. : comparaison multiple des μ_j
    - S.

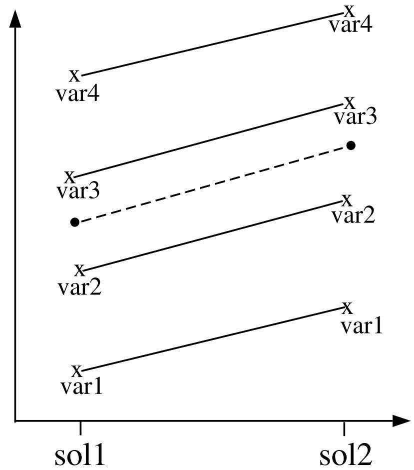
      ⇓

      le test de A
      le test de B
      sont ambigus

      ⇓

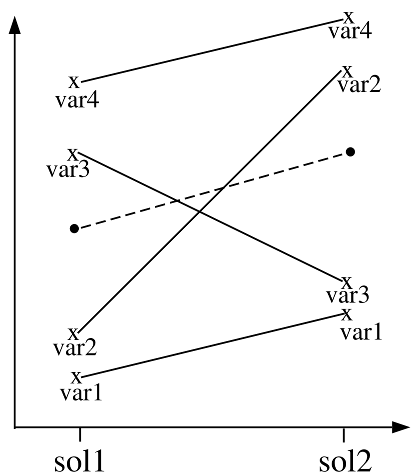
      comparaison multiple des μ_{ij}

On s'intéresse à un des facteurs croisés



F de l'interaction Non Significatif

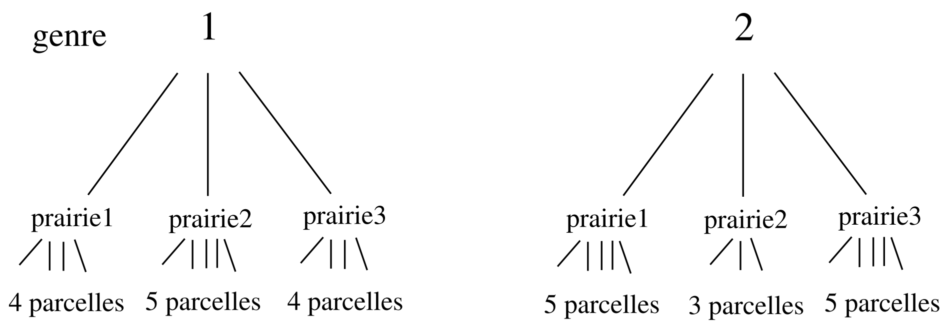
la comparaison sol1/sol2 est correcte



F de l'interaction Significatif

comparaison sol1/sol2 fausse ; la réponse dépend de la variété

Analyse de Variance à deux facteurs facteurs hiérarchisés



Les données

Genre	Prairie	Rendement
1	1	29.9
1	1	19.8
1	1	29.5
1	1	27.0
1	2	15.9
1	2	26.3
1	2	19.8
1	2	22.5
1	2	20.9
1	3	19.2
1	3	21.4
1	3	13.3
1	3	18.3
2	1	29.1
2	1	32.7
2	1	34.5
2	1	39.2
2	1	43.4
2	2	26.9
2	2	32.5
2	2	31.1
2	3	24.3
2	3	21.7
2	3	23.7
2	3	28.9
2	3	22.4

M_2 genre + prairie(genre)

$$\mu_n = \mu + \alpha_i + \beta_{j/i}$$

M_1 genre

$$\mu_n = \mu + \alpha_i$$

M_0

$$\mu_n = \mu$$

Commandes et sortie SAS

```
data trv ;
infile 'prairie' ;
input genre prairie rdt ;
run ;
proc glm ;
class genre prairie ;
model rdt = genre prairie (genre) ;
run ;
```

General Linear Models Procedure

Class Level Information

Class	Levels	Values
GENRE	2	1 2
PRAIRIE	3	1 2 3

Number of observations in data set = 26

General Linear Models Procedure

Dependent Variable : RDT

Source	DF	Sum of Squares	Mean Square	F Value	Pr > F
Model	5	921.4527179	184.2905436	11.07	0.0001
Error	20	332.9626667	16.6481333		
Corrected Total	25	1254.4153846			

R-Square	C.V.	Root MSE	RDT Mean
0.734567	15.73502	4.080212	25.93077

General Linear Models Procedure

Dependent Variable : RDT

Source	DF	Type I SS	Mean Square	F Value	Pr > F
GENRE	1	437.0600000	437.0600000	26.25	0.0001
PRAIRIE (GENRE)	4	484.3927179	121.0981795	7.27	0.0009

Source	DF	TYpe III SS	Mean Square	F Value	Pr > F
GENRE	1	417.6403101	417.6403101	25.09	0.0001
PRAIRIE (GENRE)	4	484.3927179	121.0981795	7.27	0.0009

Tableau “Modèle”

General Linear Models Procedure

Class Level Information

Class	Levels	Values
GENRE	2	1 2
PRAIRIE	3	1 2 3

Number of observations in data set = 26

General Linear Models Procedure

Dependent Variable : RDT

Source	DF	Sum of Squares	Mean Square	F Value	Pr > F
Model	5	921.4527179	184.2905436	11.07	0.0001
Error	20	332.9626667	16.6481333		
Corrected Total	25	1254.4153846			

R-Square	C.V.	Root MSE	RDT Mean
0.734567	15.73502	4.080212	25.93077

Source			
Model	IJ-1	$SCE_{M_0} - SCE_{M_2}$	$\dfrac{\left(SCE_{M_0} - SCE_{M_2}\right) / IJ-1}{SCE_{M_2} / N-IJ}$
Error	N-IJ	SCE_{M_2}	
Corrected Total	N-1	SCE_{M_0}	

modèle hiérarchisé : SCE type I

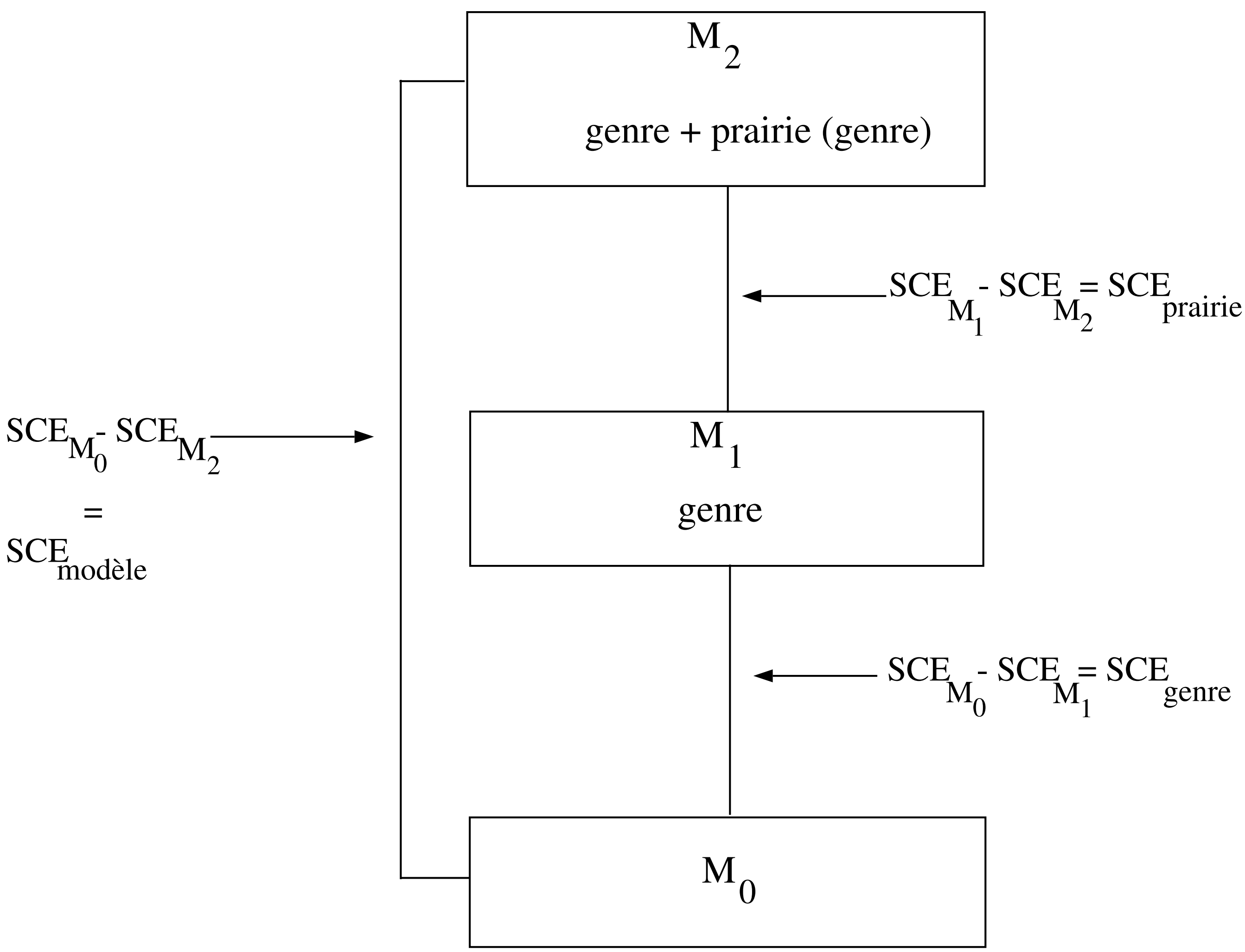


Tableau "Facteurs"

General Linear Models Procedure

Dependent Variable : RDT

Source	DF	Type I SS	Mean Square	F Value	Pr > F
GENRE	1	437.0600000	437.0600000	26.25	0.0001
PRAIRIE (GENRE)	4	484.3927179	121.0981795	7.27	0.0009

Source	DF	Type III SS	Mean Square	F Value	Pr > F
GENRE	1	417.6403101	417.6403101	25.09	0.0001
PRAIRIE (GENRE)	4	484.3927179	121.0981795	7.27	0.0009

SCE type I

GENRE	I-1	$SCE_{M_0} - SCE_{M_1}$	$\dfrac{\left(SCE_{M_0} - SCE_{M_1}\right) / I-1}{SCE_{M_2} / N-IJ}$
PRAIRIE (GENRE)	I(J-1)	$SCE_{M_1} - SCE_{M_2}$	$\dfrac{\left(SCE_{M_1} - SCE_{M_2}\right) / I(J-1)}{SCE_{M_2} / N-IJ}$

Hypothèses testées : SCE type I

$$\mathrm{H_0}$$

prairie (genre) $$\begin{cases} \forall\ (\mathrm{i},\mathrm{j}) \\ \beta_{\mathrm{j/i}} = 0 \end{cases}$$

genre $$\begin{cases} \forall\ \mathrm{i} \\ \alpha_{\mathrm{i}} + \dfrac{\sum\limits_{\mathrm{j}} \mathrm{n_{ij}}\beta_{\mathrm{j/i}}}{\mathrm{n_{i+}}} = 0 \end{cases}$$

Hypothèses testées : SCE type III

$$\mathrm{H_0}$$

prairie (genre) $\begin{cases} \forall\,(\mathrm{i},\mathrm{j}) \\ \beta_{\mathrm{j/i}} = 0 \end{cases}$

genre $\begin{cases} \forall\,\mathrm{i} \\ \alpha_{\mathrm{i}} + \dfrac{\sum_{\mathrm{j}} \beta_{\mathrm{j/i}}}{\mathrm{J}} = 0 \end{cases}$

les hypothèses sont indépendantes des effectifs

$\Rightarrow$ utiliser les sommes de type III

Analyse de Variance à deux facteurs hiérarchisés
Synthèse

A B(A)

Test du modèle

N.S. : on s'arrête

S. : il existe une structuration
en AB (A ? , B (A) ?)

B (A) ?

N.S.
⇓
le test de A
est sans ambiguité

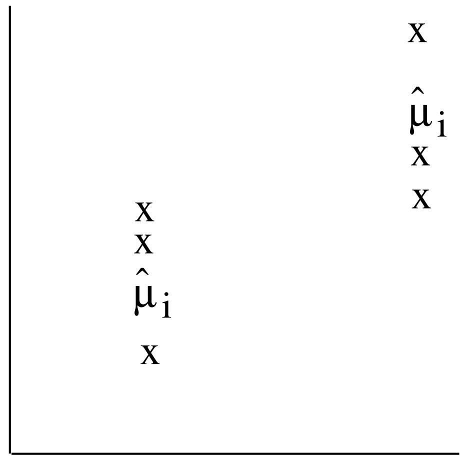
S.
⇓
le test de A
est ambigu

A ?

N.S. :

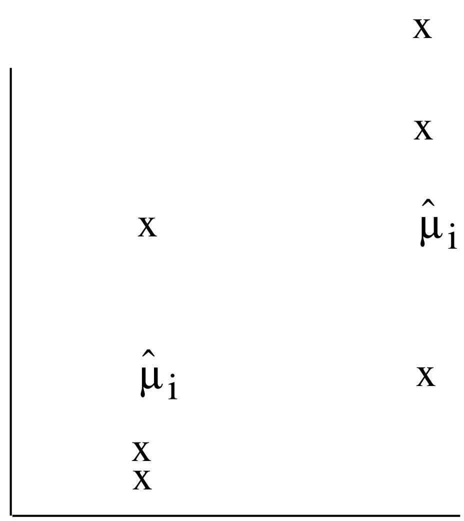
S.
⇓
comparaisons multiples
des moyennes de A

F prairie (genre) Non Significatif



La comparaison des genres est correcte

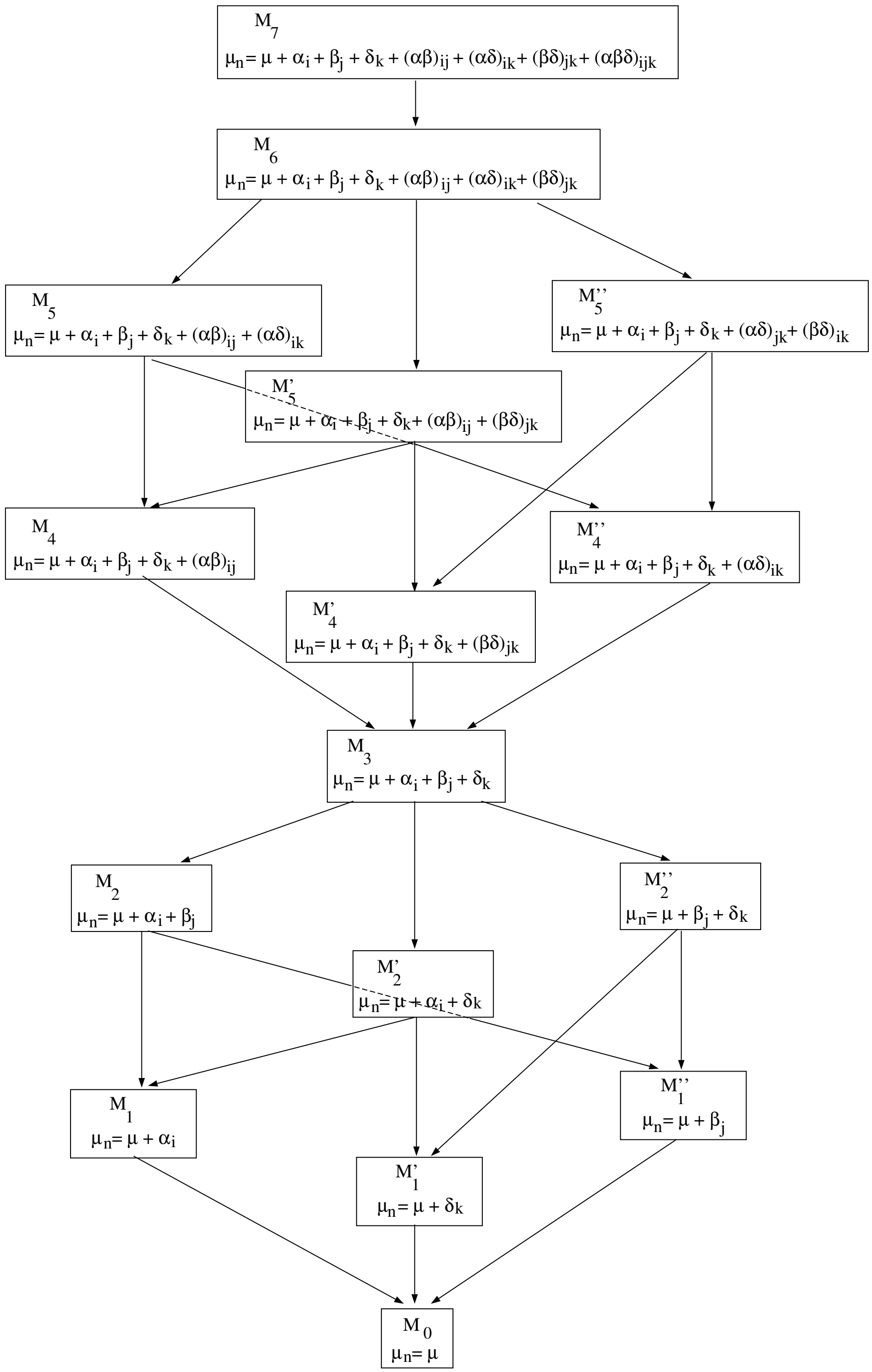
F prairie (genre) Significatif



La comparaison des genres est fausse; la moyenne de chaque genre est fortement influencée par les prairies

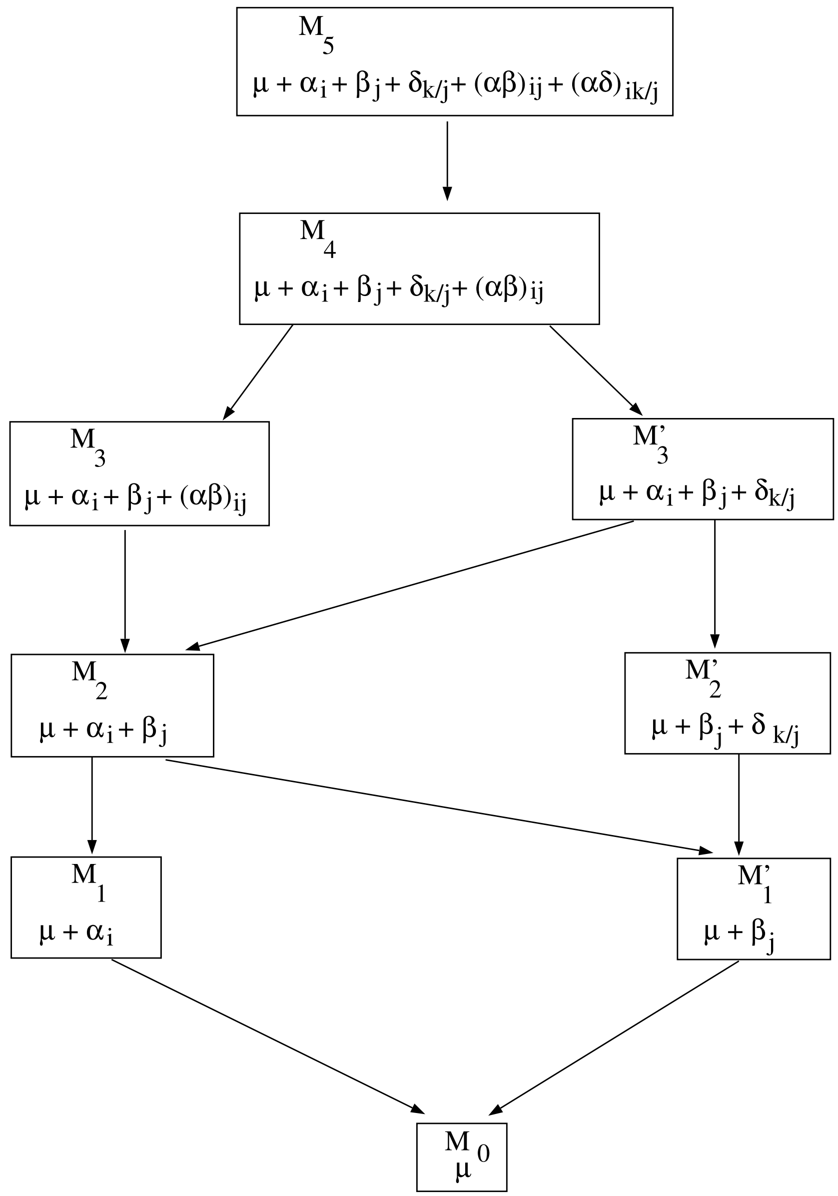
Des exemples d'analyses plus complexes

Trois facteurs croisés $A(\alpha_i)$ $B(\beta_j)$ $C(\delta_k)$



Dispositif blocs complets dans plusieurs lieux pour plusieurs traitements

$$\begin{array}{lcl} \text{Traitement} & \rightarrow & A(\alpha_i) \\ \text{Lieux} & \rightarrow & B(\beta_j) \\ \text{Bloc} & \rightarrow & C\left(\delta_{k/j}\right) \end{array}$$



Comparaisons multiples de moyennes

— Analyse de variance :

test F global significatif

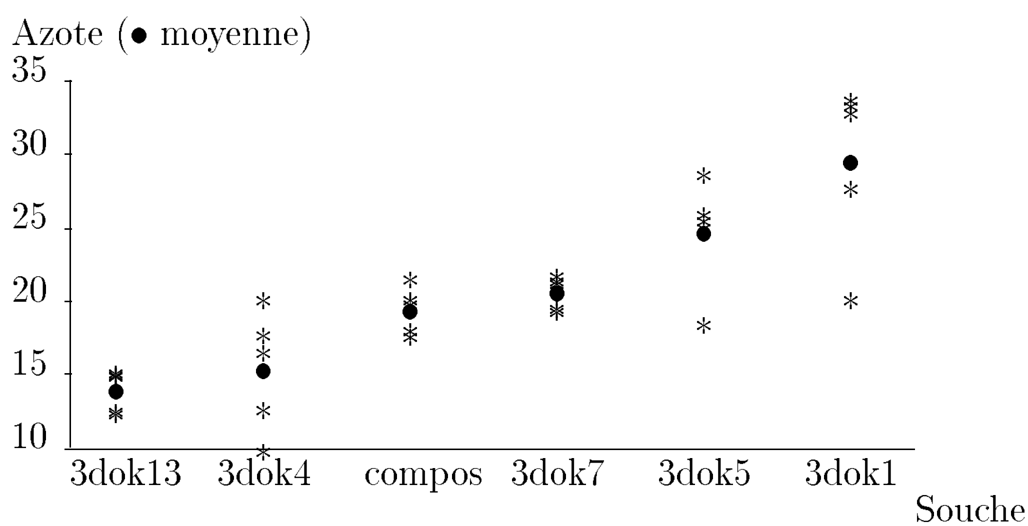
$H_0 : \mu_1 = \mu_2 = \cdots = \mu_I$ rejetée

$\Leftrightarrow$ Il existe i et j tels que $\mu_i \neq \mu_j$

— Erreurs indépendantes, identiquement distribuées $\mathcal{N}(0, \sigma^2)$

Exemple : Influence de 6 souches de rhizobium sur la quantité d'azote de plantes de trèfle rouge

Souche	Azote					Moyenne
3dok1	19.4	32.6	27.0	32.1	33.0	28.82
3dok5	17.7	24.8	27.9	25.2	24.3	23.98
3dok4	17.0	19.4	9.1	11.9	15.8	14.64
3dok7	20.7	21.0	20.05	18.8	18.6	19.92
3dok13	14.3	14.4	11.8	11.6	14.2	13.26
compos	17.3	19.4	19.1	16.9	20.8	18.70



Sortie SAS

General Linear Models Procedure

Dependent Variable : AZOTE

Source	DF	Sum of Squares	Mean Square	F Value	Pr > F
Model	5	847.05	169.41	14.37	0.0001
Error	24	282.93	11.79		
Corrected Total	29	1129.97			

$$\boxed{\widehat{\sigma}^2 = 11.79}$$

Questions :

— Faut-il former des groupes de souches de moyennes homogènes ?

— Si "compos" est le témoin, quelles moyennes diffèrent de la moyenne du témoin ?

Pour comparer 2 moyennes, méthode classique du t de Student

$$\boxed{\begin{array}{l} H_0 : \mu_i = \mu_j \text{ rejetée si} \\[1em] \mid \widehat{\mu}_i - \widehat{\mu}_j \mid \geq t(\alpha, \nu) \times s_d \end{array}}$$

α: risque de 1ère espèce

ν: degrés de liberté des résidus

t: quantile de la loi de Student au seuil α avec ν degrés de liberté

s_d: écart-type de la différence de 2 moyennes $\mu_i - \mu_j$

— en cas d'effectifs égaux $s_d = \widehat{\sigma}\sqrt{2/n}$

— en cas d'effectifs inégaux $s_d = \widehat{\sigma}\sqrt{1/n_i + 1/n_j}$

Application à l'exemple Azote / Rhizobium

Cas équirépété $n = 5$

Ecart-type de la différence $\mu_i - \mu_j$

$$s_d = \sqrt{2\widehat{\sigma}^2/n} = \sqrt{2 \times 11.79/5} = 2.17$$

Degrés de liberté des résidus $\nu = 24$

Valeur de t au seuil $\alpha = 5\%$ avec $\nu = 24$ d.d.l.

$$t(0.05, 24) = 2.064$$

La différence entre 2 moyennes particulières est significative au seuil 5% lorsqu'elle dépasse

$$t(0.05, 24) \times s_d = 2.064 \times 2.17 = 4.48$$

Cette quantité qui sert de critère est la **Plus Petite Différence Significative (P.P.D.S.)**

<u>P.P.D.S. et Risque global</u>

I niveaux du facteur

$\Rightarrow I$ moyennes à comparer soit

$C_I^2 = \frac{I(I-1)}{2}$ comparaisons de 2 moyennes

α risque de 1ère espèce d'une comparaison

α_g risque de 1ère espèce global

(= probabilité de trouver au moins une différence significative alors qu'il n'y en a pas)

Avec la méthode de la P.P.D.S. :

I	C_I^2	α	α_g
2	1	0.05	0.05
5	10	0.05	0.29
10	45	0.05	0.63
15	105	0.05	0.83

$\boxed{\alpha_g \text{ est supérieur à } \alpha}$

Autres méthodes

En contrôlant le risque global α_g,

si $\mid \widehat{\mu}_i - \widehat{\mu}_j \mid \geq$ valeur critique
les moyennes sont déclarées différentes.

Les **valeurs critiques** varient selon les méthodes.

t-corrigé de Bonferonni

on choisit α_g , par exemple 5%

pour 1 comparaison $\alpha = \alpha_{\mathrm{g}} / \frac{I(I-1)}{2}$

$$\boxed{\begin{array}{c} H_0 : \mu_i = \mu_j \text{ rejetée si} \\ \mid \widehat{\mu}_{\mathrm{i}} - \widehat{\mu}_{\mathrm{j}} \mid \geq t\left(\frac{\alpha_{\mathrm{g}}}{\frac{I(I-1)}{2}}, \nu\right) \times \mathrm{s}_d \end{array}}$$

Méthode de Tukey

Si le nombre de répétitions est constant

Intervalles de confiance simultanés
pour les différences $\mu_i - \mu_j$

$$\boxed{\begin{array}{c} H_0 : \mu_i = \mu_j \text{ rejetée si} \\ \mid \widehat{\mu}_i - \widehat{\mu}_j \mid \geq q(\alpha_g, I, \nu) \times s_d/\sqrt{2} \end{array}}$$

Méthode de Newman-Keuls

Si le nombre de répétitions est constant

Variante séquentielle de la méthode
de Tukey, plus puissante.

Dans un ordre déterminé, on
compare la différence entre la plus
grande et la plus petite moyenne
d'un groupe de p moyennes à

$$\boxed{q(\alpha_g, p, \nu) \times s_d/\sqrt{2}}$$

Exemple traité par la méthode de Newman-Keuls

NB de moyennes p	2	3	4	5	6
Etendue critique	4.48	5.42	5.99	6.40	6.71

souche	3dok13	3dok4	compos	3dok7	3dok5	3dok1
moyenne	13.26	14.64	18.70	19.92	23.98	28.82

p=6	3dok1 - 3dok13	>	6.71	significatif
p=5	3dok1 - 3dok4	>	6.40	significatif
	3dok5 - 3dok13	>	6.40	significatif
p=4	3dok1 - compos	>	5.99	significatif
	3dok5 - 3dok4	>	5.99	significatif
	3dok7 - 3dok13	>	5.99	significatif
p=3	3dok1 - 3dok7	>	5.42	significatif
	3dok5 - compos	<	5.42	non significatif
	3dok7 - 3dok4	<	5.42	non significatif
	compos - 3dok13	>	5.42	significatif
p=2	3dok1 - 3dok5	>	4.48	significatif
	3dok4- 3dok13	<	4.48	non significatif

Méthode de Scheffé

Quels que soient les nombres de répétitions

Intervalles de confiance simultanés pour tous les contrastes (combinaisons linéaires de moyennes dont la somme des coefficients est 0), en particulier pour les comparaisons de 2 moyennes

$$\boxed{\begin{array}{c} H_0 : \mu_i = \mu_j \text{ rejetée si} \\ \mid \widehat{\mu}_i - \widehat{\mu}_j \mid \geq \\ \sqrt{(I-1)F(\alpha_g, I-1, \nu)} \times s_d \end{array}}$$

Test conservatif, très général

Méthode de Dunnett

Comparaisons au témoin

On ne fait que (I-1) comparaisons

$$\boxed{\begin{array}{c} H_0 : \mu_i = \mu_0 \text{ rejetée si} \\ \mid \widehat{\mu}_i - \widehat{\mu}_0 \mid \geq d(\alpha_g, I-1, \nu) \times s_d \end{array}}$$

μ_0 moyenne du témoin

Traitements meilleurs que le témoin :

— test unilatéral à droite
$H_0 : \mu_i = \mu_0 \quad , H_1 : \mu_i > \mu_0$
si $\widehat{\mu}_i - \widehat{\mu}_0 > d(\alpha_g\, I-1, \nu) \times s_d$
on rejette H_0.

— test unilatéral à gauche
$H_0 : \mu_i = \mu_0 \quad , H_1 : \mu_i < \mu_0$
si $\widehat{\mu}_i - \widehat{\mu}_0 < -d(\alpha_g\, I-1, \nu) \times s_d$
on rejette H_0.

Tableau récapitulatif

Méthode	Valeur critique
PPDS	$t(\alpha,\nu)\mathrm{s}_d$
Bonferonni	$t\left(\frac{\alpha_{\mathrm{g}}}{I(I-1)/2},\nu\right)\mathrm{s}_d$
Tukey	$q(\alpha_{\mathrm{g}},I,\nu)\mathrm{s}_d/\sqrt{2}$
Newman Keuls	$q(\alpha_{\mathrm{g}},p,\nu)\mathrm{s}_d/\sqrt{2}$ où $p=2,3\cdots I$
Scheffé	$\sqrt{(I-1)F(\alpha_{\mathrm{g}},I-1,\nu)}\times\mathrm{s}_d$
Dunnett	$d(\alpha_{\mathrm{g}},I-1,\nu)\mathrm{s}_d$

I nombre de moyennes à comparer

α erreur 1ère espèce, de chaque comparaison 2 à 2

α_g risque global de toutes les comparaisons 2 à 2

ν degrés de liberté des résidus

$\hat{\sigma}$ écart-type résiduel

s_d écart-type de la différence des moyennes
en cas d'effectifs égaux $s_d=\widehat{\sigma}\sqrt{2/n}$
en cas d'effectifs inégaux $s_d=\widehat{\sigma}\sqrt{1/n_i+1/n_j}$

t quantile de la loi de Student (table)

q quantile “studentized range” (table)

F quantile de la loi de Fisher (table)

d quantile de Dunnett (table)

<u>Caractéristiques</u> <u>des</u> <u>méthodes</u>

	PPDS	Tukey	Newman Keuls
Erreur 1ère espèce contrôlée	pour chaque comparaison ou contraste	pour toutes les comparaisons de 2 moyennes	pour toutes les comparaisons de 2 moyennes
Nombre de répétitions	non constant	constant	constant
Utilisations	quelques comparaisons ou contrastes définis a priori	comparaison de toutes les moyennes	comparaison de toutes les moyennes
Caractéristique		conservatif	

contraste : combinaison linéaire de moyennes dont la somme des coefficients est 0

exemples : $\mu_1 - \mu_2, \mu_1 - \frac{\mu_2+\mu_3}{2}$

test conservatif : test dont le risque de 1ère espèce <u>réel</u> est inférieur au risque choisi (par exemple 5 %).

Caractéristiques des méthodes –suite

	Bonferonni	Scheffé	Dunnett
Erreur 1ère espèce contrôlée	pour toutes les comparaison effectuées	pour tous les contrastes	pour toutes les comparaisons au témoin
Nombre de répétitions	non constant	non constant	constant
Utilisations	comparaisons de moyennes ou contrastes en nombre réduit	contraste (plus de 2 moyennes) comparaison de 2 moyenne en grand nombre	comparaison au témoin recherche des moyennes les plus élevées
Caractéristique	général conservatif	très général conservatif cohérent avec le test F	bilatéral ou unilatéral

contraste : combinaison linéaire de moyennes dont la somme des coefficients est 0

exemples : $\mu_1 - \mu_2, \mu_1 - \frac{\mu_2+\mu_3}{2}$

test conservatif : test dont le risque de 1ère espèce réel est inférieur au risque choisi (par exemple 5 %).

Application à l'exemple Azote/ Rhizobium

Méthode	Risque choisi	Valeur critique pour les différences entre 2 moyennes	Options SAS
P.P.D.S	$\alpha = 5\%$	4.48	LSD ou T
Tukey	$\alpha_{\mathrm{g}} = 5\%$	6.71	TUKEY
Newman Keuls	$\alpha_{\mathrm{g}} = 5\%$	Nb de moyennes / Etendue: 2 : 4.48 3 : 5.42 4 : 5.99 5 : 6.40 6 : 6.71	SNK
Bonferonni	$\alpha_{\mathrm{g}} = 5\%$	7.08	BON
Scheffé	$\alpha_{\mathrm{g}} = 5\%$	7.86	SCHEFFE
Dunnett	$\alpha_{\mathrm{g}} = 5\%$	5.85	DUNNETT ('compos')

Estimation

Souche	Azote moyenne
3dok1	28.82
3dok5	23.98
3dok7	19.92
compos	18.70
3dok4	14.64
3dok13	13.26

Différences entre les moyennes

	3dok1	3dok5	3dok7	compos	3dok4
3dok1	0				
3dok5	4.84	0			
3dok7	8.90	4.06	0		
compos	10.12	5.28	1.22	0	
3dok4	14.18	9.34	5.28	4.06	0
3dok13	15.56	10.72	6.66	5.44	1.38

Classements obtenus

P.P.D.S		Newman Keuls		Tukey		Bonferonni		Scheffé		Moyenne	Souche
A		A			A		A		A	28.82	3dok1
	B		B	B	A	B	A	B	A	23.98	3dok5
C	B	C	B	B	C	B	C	B	C	19.92	3dok7
C	D	C	B	B	C	B	C	B	C	18.70	compos
E	D	C	D		C		C		C	14.64	3dok4
E			D		C		C		C	13.26	3dok13

Les moyennes portant la même lettre ne sont pas significativement différentes.

DUNNETT

SOUCHE

Comparaison

3dok1 - compos ***

3dok5 - compos

3dok7 - compos

3dok4 - compos

3dok13 - compos

Les comparaisons significatives sont repérées par ***