General Linear Models Procedure

Class Level Information

Class	Levels	Values
TESTEUR	2	1 2
ORIGINE	3	1 2 3

Number of observations in data set = 12

General Linear Models Procedure

Dependent Variable : HUILE

Source	DF	Sum of Squares	Mean Square	F Value	Pr > F
Model	5	47.18144167	9.43628833	6.18	0.0233
Error	6	9.16665000	1.52777500		
Corrected Total	11	56.34809167			

R-Square	C.V.	Root MSE	HUILE Mean
0.837321	2.671010	1.236032	46.27583

$$R^2 = \frac{SCE_{M_0} - SCE_{M_3}}{SCE_{M_0}}$$

Source			
Model	IJ-1	$SCE_{M_0} - SCE_{M_3}$	$\frac{(SCE_{M_0} - SCE_{M_3}) / IJ-1}{SCE_{M_3} / N-IJ}$
Error	N-IJ	SCE_{M_3}	
Corrected Total	N-1	SCE_{M_0}	

Tableau "Facteurs"

Dependent Variable : HUILE

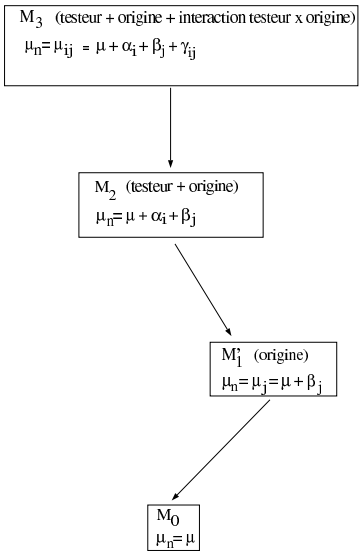
Source	DF	Type I SS	Mean Square	F Value	Pr > F
TESTEUR	1	9.48740833	9.48740833	6.21	0.0470
ORIGINE	2	33.74581667	16.87290833	11.04	0.0097
TESTEUR*ORIGINE	2	3.94821667	1.97410833	1.29	0.3415

Dependent Variable : HUILE

Source	DF	Type III SS	Mean Square	F Value	Pr > F
TESTEUR	1	9.48740833	9.48740833	6.21	0.0470
ORIGINE	2	33.74581667	16.87290833	11.04	0.0097
TESTEUR*ORIGINE	2	3.94821667	1.97410833	1.29	0.3415

TESTEUR	I-1	$SCE_{M_0} - SCE_{M_1}$	$\frac{(SCE_{M_0} - SCE_{M_1}) / I-1}{SCE_{M_3} / N-IJ}$
ORIGINE	J-1	$SCE_{M_1} - SCE_{M_2}$	$\frac{(SCE_{M_1} - SCE_{M_2}) / J-1}{SCE_{M_3} / N-IJ}$
TESTEUR*ORIGINE	(I-1)(J-1)	$SCE_{M_2} - SCE_{M_3}$	$\frac{(SCE_{M_2} - SCE_{M_3}) / (I-1)(J-1)}{SCE_{M_3} / N-IJ}$

```
1 1 43.54          data trv ;
1 1 45.30          infile 'tourn' ;
1 2 44.25          input testeur origine huile ;
1 2 42.55          run ;
1 3 47.28          proc glm ;
1 3 49.40          class testeur origine ;
2 1 47.21          model huile = origine testeur testeur*origine ;
2 1 47.73          run ;
2 2 44.34
2 2 46.49
2 3 47.75
2 3 49.47
```



Sortie SAS

General Linear Models Procedure

Class Level Information

Class	Levels	Values
Testeur	2	1 2
Origine	3	1 2 3

Number of observations in data set = 12

General Linear Models Procedure

Dependent Variable : HUILE

Source	DF	Sum of Squares	Mean Square	F Value	Pr > F
Model	5	47.18144167	9.43628833	6.18	0.0233
Error	6	9.16665000	1.52777500		
Corrected Total	11	56.34809167			

R-Square	C.V.	Root MSE	HUILE Mean
0.837321	2.671010	1.236032	46.27583

Dependent Variable : HUILE

Source	DF	Type I SS	Mean Square	F Value	Pr > F
ORIGINE	2	33.74581667	16.87290833	11.04	0.0097
TESTEUR	1	9.48740833	9.48740833	6.21	0.0470
TESTEUR*ORIGINE	2	3.94821667	1.97410833	1.29	0.3415

Source	DF	Type III SS	Mean Square	F Value	Pr > F
ORIGINE	2	33.74581667	16.87290833	11.04	0.0097
TESTEUR	1	9.48740833	9.48740833	6.21	0.0470
TESTEUR*ORIGINE	2	3.94821667	1.97410833	1.29	0.3415

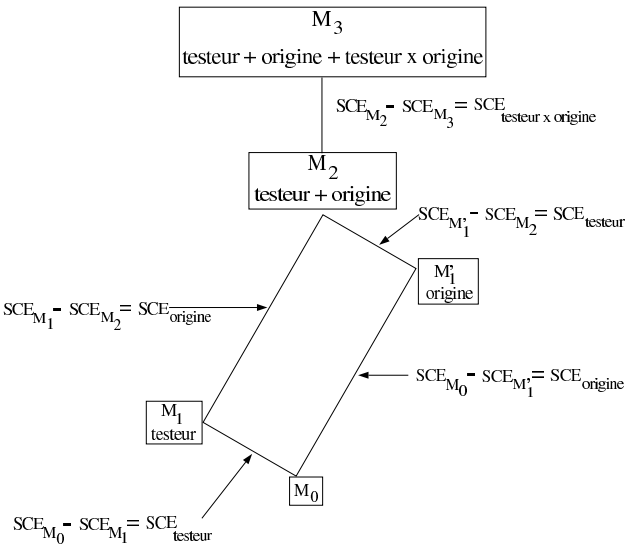
Dependent Variable : HUILE

Source	DF	Type I SS	Mean Square	F Value	Pr > F
ORIGINE	2	33.74581667	16.87290833	11.04	0.0097
TESTEUR	1	9.48740833	9.48740833	6.21	0.0470
TESTEUR*ORIGINE	2	3.94821667	1.97410833	1.29	0.3415

Dependent Variable : HUILE

Source	DF	Type III SS	Mean Square	F Value	Pr > F
ORIGINE	2	33.74581667	16.87290833	11.04	0.0097
TESTEUR	1	9.48740833	9.48740833	6.21	0.0470
TESTEUR*ORIGINE	2	3.94821667	1.97410833	1.29	0.3415

ORIGINE	J-1	$SCE_{M_0} - SCE_{M'_1}$	$\frac{\left(SCE_{M_0} - SCE_{M'_1}\right) / J-1}{SCE_{M_3} / N-IJ}$
TESTEUR	I-1	$SCE_{M'_1} - SCE_{M_2}$	$\frac{\left(SCE_{M'_1} - SCE_{M_2}\right) / I-1}{SCE_{M_3} / N-IJ}$
TESTEUR*ORIGINE	(I-1)(J-1)	$SCE_{M_2} - SCE_{M_3}$	$\frac{\left(SCE_{M_2} - SCE_{M_3}\right) / (I-1)(J-1)}{SCE_{M_3} / N-IJ}$



$$\text{SCE type I} = \text{SCE type III}$$

Une unique décomposition de la somme des carrés des écarts totale :

$$SCE_{totale} = SCE_{testeur} + SCE_{origine} + SCE_{testeur\ x\ origine} + SCE_{résiduelle}$$

Plan orthogonal : que teste-t-on avec les tests F ?

Plan orthogonal $\Leftrightarrow$ SCE type I = SCE type III

	F	H_0
origine	$\left\{ \begin{array}{l} \frac{\left(SCE_{M_0} - SCE_{M_1'}\right) / \ J-1}{SCE_{M_3} / \ N-IJ} \\ = \\ \frac{\left(SCE_{M_1} - SCE_{M_2}\right) / \ J-1}{SCE_{M_3} / \ N-IJ} \end{array} \right.$	$\left\{ \begin{array}{l} \forall j \\ \beta_j + \frac{\sum_i \gamma_{ij}}{I} = 0 \end{array} \right.$
testeur	$\left\{ \begin{array}{l} \frac{\left(SCE_{M_0} - SCE_{M_1}\right) / \ I-1}{SCE_{M_3} / \ N-IJ} \\ = \\ \frac{\left(SCE_{M_1'} - SCE_{M_2}\right) / \ I-1}{SCE_{M_3} / \ N-IJ} \end{array} \right.$	$\left\{ \begin{array}{l} \forall i \\ \alpha_i + \frac{\sum_j \gamma_{ij}}{J} = 0 \end{array} \right.$
testeur x origine	$\frac{\left(SCE_{M_2} - SCE_{M_3}\right) / \ IJ-1}{SCE_{M_3} / \ N-IJ}$	$\left\{ \begin{array}{l} \forall (i,j) \\ \gamma_{ij} = 0 \end{array} \right.$

Analyse de variance à deux facteurs
cas non orthogonal

(pas de case vide)

3 variétés de carottes cultivées dans 2 sols

Y : nombre de jours avant germination

Données expérimentales

Y_{ijr}	Variété 1	Variété 2	Variété 3
Sol 1	6 10 11	13 15	14 22
Sol 2	12 19 15 18	31	18 9 12

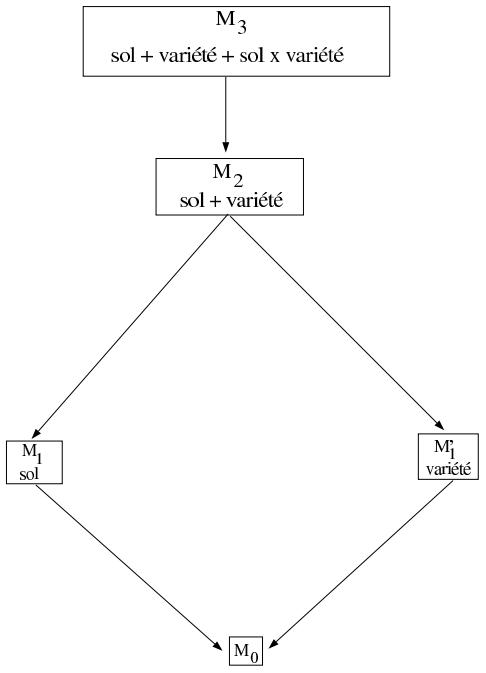
ex : dispositif déséquilibré

Facteur A = sol I = 2 niveaux

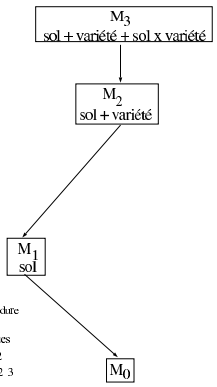
Facteur B = variété J = 3 niveaux

Fichier des données

```
1 1 16
1 1 10
1 1 11
1 2 13
1 2 15
1 3 14
1 3 22
2 1 12
2 1 19
2 1 15
2 1 18
2 2 31
2 3 18
2 3 9
2 3 12
```

```
data trv ;
infile 'solvar' ;
input sol variete germina ;
run ;
proc glm ;
class sol variete ;
model germina = sol variete  sol * variete ;
run ;
```



General Linear Models Procedure
Class Level Information

Class	Levels	Values
SOL	2	1 2
VARIETE	3	1 2 3

Number of observations in data set = 15

General Linear Models Procedure

Dependent Variable : GERMINA

Source	DF	Sum of Squares	Mean Square	F Value	Pr > F
Model	5	400.0000000	80.0000000	6.00	0.0103
Error	9	120.0000000	13.3333333		
Corrected Total	14	520.0000000			

R-Square	C.V.	Root MSE	GERMINA Mean
0.769231	24.34322	3.651484	15.00000

General Linear Models Procedure

Dependent Variable : GERMINA

Source	DF	Type I SS	Mean Square	F Value	Pr > F
SOL	1	52.5000000	52.5000000	3.94	0.0785
VARIETE	2	124.7340426	62.3670213	4.68	0.0405
SOL*VARIETE	2	222.7659574	111.3829787	8.35	0.0089

Source	DF	Type III SS	Mean Square	F Value	Pr > F
SOL	1	123.7714286	123.7714286	9.28	0.0139
VARIETE	2	192.1276596	96.0638298	7.20	0.0135
SOL*VARIETE	2	222.7659574	111.3829787	8.35	0.0089

Tableau "Modèle"

General Linear Models Procedure

Class Level Information

Class	Levels	Values
SOL	2	1 2
VARIETE	3	1 2 3

Number of observations in data set = 15

General Linear Models Procedure

Dependent Variable : GERMINA

Source	DF	Sum of Squares	Mean Square	F Value	Pr > F
Model	5	400.0000000	80.0000000	6.00	0.0103
Error	9	120.0000000	13.3333333		
Corrected Total	14	520.0000000			

R-Square	C.V.	Root MSE	GERMINA Mean
0.769231	24.34322	3.651484	15.00000

$$r^2 = \frac{SCE_{M_0} - SCE_{M_3}}{SCE_{M_0}}$$

Source			
Model	IJ-1	$SCE_{M_0} - SCE_{M_3}$	$\frac{(SCE_{M_0} - SCE_{M_3}) / IJ-1}{SCE_{M_3} / N-IJ}$
Error	N-IJ	SCE_{M_3}	
Corrected Total	N-1	SCE_{M_0}	

Dependent Variable : GERMINA

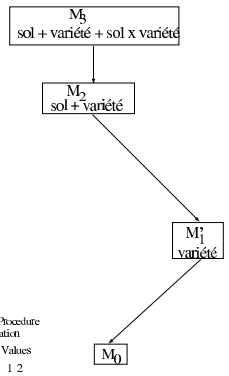
Source	DF	Type I SS	Mean Square	F Value	Pr > F
SOL	1	52.5000000	52.5000000	3.94	0.0785
VARIETE	2	124.7340426	62.3670213	4.68	0.0405
SOL*VARIETE	2	222.7659574	111.3829787	8.35	0.0089

Source	DF	Type III SS	Mean Square	F Value	Pr > F
SOL	1	123.7714286	123.7714286	9.28	0.0139
VARIETE	2	192.1276596	96.0638298	7.20	0.0135
SOL*VARIETE	2	222.7659574	111.3829787	8.35	0.0089

Type I

SOL	I-1	$SCE_{M_0} - SCE_{M_1}$	$\frac{(SCE_{M_0} - SCE_{M_1}) / I-1}{SCE_{M_3} / N-IJ}$
VARIETE	J-1	$SCE_{M_1} - SCE_{M_2}$	$\frac{(SCE_{M_1} - SCE_{M_2}) / J-1}{SCE_{M_3} / N-IJ}$
SOL*VARIETE	IJ-1	$SCE_{M_2} - SCE_{M_3}$	$\frac{(SCE_{M_2} - SCE_{M_3}) / IJ-1}{SCE_{M_3} / N-IJ}$


```
data trv ;
infile 'solvar' ;
input sol variete germina ;
run ;
proc glm ;
class sol variete ;
model germina = variete sol  variete*sol ;
run ;
```



General Linear Models Procedure
Class Level Information

Class	Levels	Values
SOL	2	1 2
VARIETE	3	1 2 3

Number of observations in data set = 15

General Linear Models Procedure

Dependent Variable : GERMINA

Source	DF	Sum of Squares	Mean Square	F Value	Pr > F
Model	5	400.0000000	80.0000000	6.00	0.0103
Error	9	120.0000000	13.3333333		
Corrected Total	14	520.0000000			

R-Square	C.V.	Root MSE	GERMINA Mean
0.769231	24.34322	3.651484	15.000000

General Linear Models Procedure

Dependent Variable : GERMINA

Source	DF	Type I SS	Mean Square	F Value	Pr > F
VARIETE	2	93.3333333	46.6666667	3.50	0.0751
SOL	1	83.9007092	83.9007092	6.29	0.0334
SOL*VARIETE	2	222.7659574	111.3829787	8.35	0.0089

Source	DF	Type III SS	Mean Square	F Value	Pr > F
VARIETE	2	192.1276596	96.0638298	7.20	0.0135
SOL	1	123.7714286	123.7714286	9.28	0.0139
SOL*VARIETE	2	222.7659574	111.3829787	8.35	0.0089

Dependent Variable : GERMINA

Source	DF	Type I SS	Mean Square	F Value	Pr > F
VARIETE	2	93.3333333	46.6666667	3.50	0.0751
SOL	1	83.9007092	83.9007092	6.29	0.0334
SOL*VARIETE	2	222.7659574	111.3829787	8.35	0.0089

Source	DF	Type III SS	Mean Square	F Value	Pr > F
VARIETE	2	192.1276596	96.0638298	7.20	0.0135
SOL	1	123.7714286	123.7714286	9.28	0.0139
SOL*VARIETE	2	222.7659574	111.3829787	8.35	0.0089

Type I

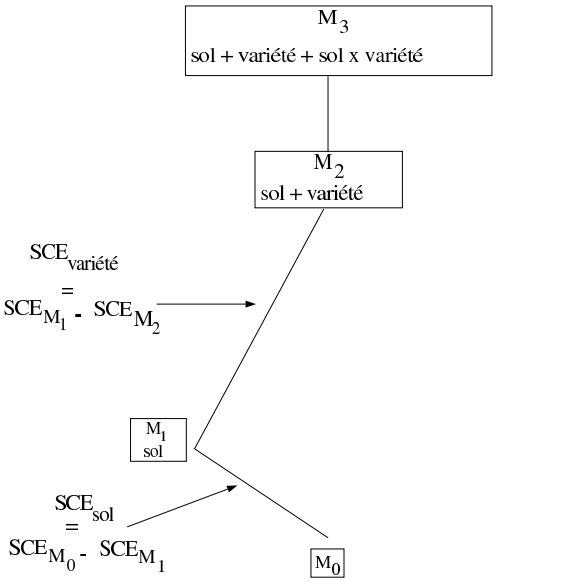
VARIETE	J-1	$SCE_{M'_0} - SCE_{M'_1}$	$\frac{(SCE_{M'_0} - SCE_{M'_1}) / J-1}{SCE_{M_3} / N-IJ}$
SOL	I-1	$SCE_{M'_1} - SCE_{M_2}$	$\frac{(SCE_{M'_1} - SCE_{M_2}) / I-1}{SCE_{M_3} / N-IJ}$
SOL*VARIETE	IJ-1	$SCE_{M_2} - SCE_{M_3}$	$\frac{(SCE_{M_2} - SCE_{M_3}) / IJ-1}{SCE_{M_3} / N-IJ}$

Plusieurs sommes de carrés sont disponibles.

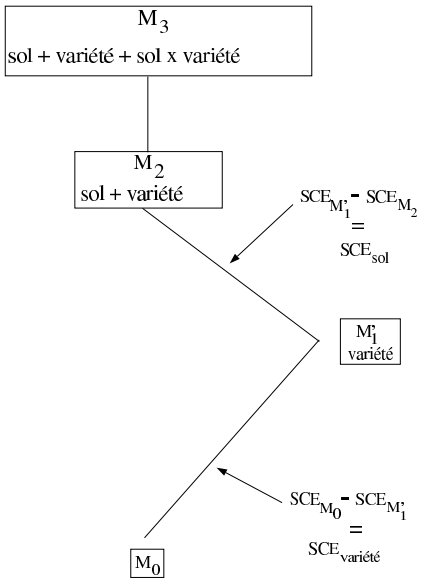
Lesquelles choisir pour tester des modèles ?

Pourquoi ?

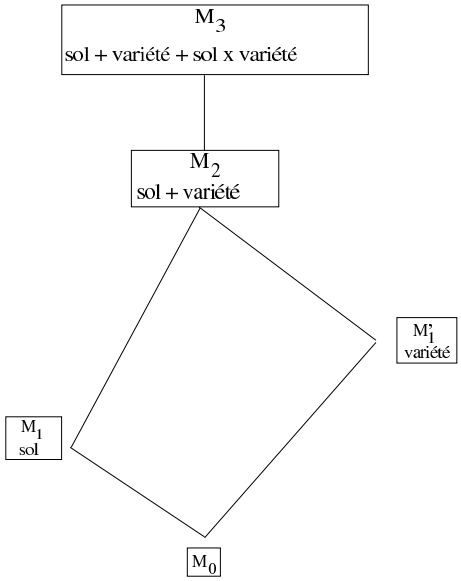
SCE type I :



SCE type I :



SCE type I :



$$SCE_{M_0} - SCE_{M'_{1_{variété}}} \neq SCE_{M_{1_{sol}}} - SCE_{M_{2_{sol+variété}}}$$
$$SCE_{M_0} - SCE_{M_{1_{sol}}} \neq SCE_{M'_{1_{variété}}} - SCE_{M_{2_{sol+variété}}}$$

Illustration géométrique