Plans Factoriels

A/ Blocs Complets

On considère 2 facteurs A et B dont les nombres de niveaux respectifs sont K_A et K_B. En fonction du nombre d'observations que l'on possède pour chaque couple de niveaux des facteurs A et B, on utilise différentes méthodes. Ce type d'analyse permet de tester l'égalité des moyennes théoriques d'une variable quantitative pour chaque niveau du facteur A et du facteur B (et éventuellement de l'interaction entre ces facteurs).

A1/ Blocs Complets avec répétitions (r répétitions)

Les observations sont sous la forme :

			Facteur B		
	niveau	1		K_B	TOTAL
	1	$x_{1,1,1}$ . . $x_{r,1,1}$	.....	$x_{1,K_B,1}$ . . $x_{r,K_B,1}$	$T_{A,1}$
Facteur A	. . . . .	. . . . .	. . . . .	. . . . .	. . . . .
	K_A	$x_{1,1,K_A}$ . . $x_{r,1,K_A}$	.....	x_{1,K_B,K_A} . . x_{r,K_B,K_A}	T_{A,K_A}
	TOTAL	$T_{B,1}$		T_{B,K_B}	T_g

Le modèle de cette analyse est le suivant :

$$x_{ijk} = \mu + \alpha_i + \beta_j + \gamma_{ij} + e_{ijk}$$

Les hypothèses testées :

H_0 : égalité des moyennes des K_A niveaux du facteur A ou
les effets du facteur A sont nuls, tous les α_i prennent pour valeur 0

H_0' : égalité des moyennes des K_B niveaux du facteur A ou
les effets du facteur B sont nuls, tous les β_j prennent pour valeur 0

H_0'' : les effets de l'interactions entre les niveaux du facteur A et les niveaux du facteurs B sont nuls, tous les γ_{ij} prennent pour valeur 0

H_1 : au moins un des α_i est différent de 0

H_1' : au moins un des β_j est différent de 0

H_1'' : au moins un des γ_{ij} est différent de 0

Les conditions d'applications :

Tous les x_{ijk} suivent des lois Normales de même variance σ^2 (estimée par s_e^2), ou, ce qui est identique, les e_{ijk} sont "Normaux", indépendants et de même variance σ^2 (estimée par s_e^2).

Le tableau d'Analyse de Variance est le suivant :

Origine	**Σ des carrés** (a)	**ddl** (b)	**Variance** (a)/(b)	**F**
Totale (A)	$\sum_{i=1}^{K_A}\sum_{j=1}^{K_B}\sum_{l=1}^{r} x_{lji}^2 - \frac{T_g^2}{N}$	N-1		
Σ des Groupes (B)	$\sum_{i=1}^{K_A}\sum_{j=1}^{K_B} \frac{T_{i,j}^2}{r} - \frac{T_g^2}{N}$	$K_A K_B$-1	S^2	$\frac{S^2}{s_e^2}$
Intragroupe (C)	(A) - (B)	N-$K_A K_B$	s_e^2	
InterA (D)	$\sum_{i=1}^{K_A} \frac{T_{A,i}^2}{rK_B} - \frac{T_g^2}{N}$	K_A-1	S_A^2	$\frac{S_A^2}{s_e^2}$
InterB (E)	$\sum_{j=1}^{K_B} \frac{T_{B,j}^2}{rK_A} - \frac{T_g^2}{N}$	K_B-1	S_B^2	$\frac{S_B^2}{s_e^2}$
InterAB (F)	(B) - (D) - (E)	(K_A-1)(K_B-1)	S_{AB}^2	$\frac{S_{AB}^2}{s_e^2}$

Exercice :

```
%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%
%
% ===================
% ANALYSE DE VARIANCE
% ===================
%
%
%
% FONCTIONS UTILISEES :
% =====================
% ANOVA1.m   Bartlett.m   ANOVA2.m
%
% UTILISER "help name_function" pour obtenir des informations
% sur les fonctions utilisees
%
%
%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%
%
% Exercice 1 :
% ------------
% Lors d'une étude portant sur la consommation d'oxygène de 2 types
% d'organisme en fonction de la salinité de l'eau de mer. L'expérimentation
% a été effectuée a 22°C et 8 répétitions ont été effectuée pour chaque
% organisme et chaque salinité :
%
%  salinité 1 (100%) espèce 1  : 10.5 12.9 11.2 10.1  9.0 10.9 13.6 10.3
%  salinité 1 (100%) espèce 2  :  6.6  8.9 11.4 10.4  7.8  8.2  8.1  6.0
%  salinité 2 ( 75%) espèce 1  :  6.9  9.8  9.9  8.4  6.6  8.9  9.1  6.5
%  salinité 2 ( 75%) espèce 2  :  7.3  6.5  5.6  8.9  7.0  7.7  7.4  8.3
%  salinité 3 ( 50%) espèce 1  : 12.3 12.9  9.9 13.5 12.9  9.4 11.9 12.6
%  salinité 3 ( 50%) espèce 2  : 10.3 12.7 10.3 14.7 13.8 12.0 12.1 13.7
%
%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%
```

A2/ Blocs Complets sans répétitions

Les observations sont sous la forme :

			Facteur B		
	niveau	1		K_B	TOTAL
	1	$x_{1,1}$	.....	$x_{K_B,1}$	$T_{A,1}$
Facteur A	. . . . .	. . . . .	. . . . .	. . . . .	. . . . .
	K_A	x_{1,K_A}	.....	x_{K_B,K_A}	T_{A,K_A}
	TOTAL	$T_{B,1}$		T_{B,K_B}	T_g

Le modèle de cette analyse est le suivant :

$$x_{ij} = \mu + \alpha_i + \beta_j + e_{ij}$$

Les hypothèses testées :

H_0 : égalité des moyennes des K_A niveaux du facteur A ou
les effets du facteur A sont nuls, tous les α_i prennent pour valeur 0

H_0' : égalité des moyennes des K_B niveaux du facteur A ou
les effets du facteur B sont nuls, tous les β_j prennent pour valeur 0

H_1 : au moins un des α_i est différent de 0

H_1' : au moins un des β_j est différent de 0

Les conditions d'applications :

Tous les x_{ij} suivent des lois Normales de même variance σ^2 (estimée par s_e^2), ou, ce qui est identique, les e_{ij} sont "Normaux", indépendants et de même variance σ^2 (estimée par s_e^2).

Le tableau d'Analyse de Variance est le suivant :

Origine	**Σ des carrés** (a)	**ddl** (b)	**Variance** (a)/(b	**F**
Totale (A)	$\sum_{i=1}^{K_A}\sum_{j=1}^{K_B} x_{ij}^2 - \frac{T_g^2}{N}$	N-1		
InterA (B)	$\sum_{i=1}^{K_A}\frac{T_{A,i}^2}{K_B} - \frac{T_g^2}{N}$	K_A-1	S_A^2	$\frac{S_A^2}{s_e^2}$
InterB (C)	$\sum_{j=1}^{K_B}\frac{T_{B,j}^2}{K_A} - \frac{T_g^2}{N}$	K_B-1	S_B^2	$\frac{S_B^2}{s_e^2}$
Intragroupe (D)	(A) - (B) - (C)	N-K_A-K_B+1	s_e^2	

Exercice :

```
%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%
%
% Exercice 2 :
% ------------
% On souhaite analyser la qualité des eaux de 20 petits bassins versants en
% fonction d'un indice de qualité (qui intègre différentes mesures) et ce, en
% fonction de 4 périodes de l'année. Peut-on considérer qu'il existe des
% différences entre bassins versants ? Existe-t-il un effet de la période à
% laquelle est effectuée la séries de mesures ?
% Les observations sont les suivantes :
% data = [46,38,42,47,48,40,43,43,42,49,45,51,47,52,50,48,47,47,47,45
%         46,42,44,39,49,38,47,45,47,55,50,42,47,51,45,47,50,47,44,48
%         47,42,42,51,54,40,49,46,45,47,39,51,45,49,45,49,49,44,45,51
%         48,47,44,45,51,44,47,48,47,57,49,55,48,48,46,48,54,54,44,48];
% chaque ligne correspond à une période avec les indices des 20 bassins versants.
% Comparer les résultats obtenus avec un ANOVA à un facteur (bassins versants
% et/ou période) à ceux d'une ANOVA à deux facteurs.
%
%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%
```

B/ Blocs Incomplets

Les observations ont sous la forme :

			Facteur B		
	niveau	1		K_B	TOTAL
	1	$x_{1,1}$	.....	$x_{K_B,1}$	$T_{A,1}$
Facteur A	. . . . .	. . . . .	. . . . .	. . . . .	. . . . .
	K_A	x_{1,K_A}	.....	x_{K_B,K_A}	T_{A,K_A}
	TOTAL	$T_{B,1}$		T_{B,K_B}	T_g

mais toutes les cases ne seront pas remplis, par exemple :

niveau	1	2	3	4	5	Total
1	x_{11}				x_{12}	T_{A1}
2		x_{21}	x_{22}			T_{A2}
3			x_{31}	x_{32}		T_{A3}
4	x_{41}	x_{42}				T_{A4}
5				x_{51}	x_{52}	T_{A5}
Total	T_{B1}	T_{B2}	T_{B3}	T_{B4}	T_{B5}	Tg

Le modèle de cette analyse est le suivant :

$$x_{ij} = \mu + \alpha_i + \beta_j + e_{ij}$$

Les hypothèses testées :

H_0 : égalité des moyennes des K_A niveaux du facteur A ou
les effets du facteur A sont nuls, tous les α_i prennent pour valeur 0

H_0' : égalité des moyennes des K_B niveaux du facteur A ou
les effets du facteur B sont nuls, tous les β_j prennent pour valeur 0

H_1 : au moins un des α_i est différent de 0

H_1' : au moins un des β_j est différent de 0

Les conditions d'applications :

Tous les x_{ij} suivent des lois Normales de même variance σ^2 (estimée par s_e^2), ou, ce qui est identique, les e_{ij} sont "Normaux", indépendants et de même variance σ^2 (estimée par s_e^2).

Le tableau d'Analyse de Variance est le suivant :

Origine	**Σ des carrés** (a)	**ddl** (b)	**Variance** (a)/(b	**F**
Totale (A)	$\sum_{i=1}^{K_A}\sum_{j=1}^{K_B} x_{ij}^2 - \frac{T_g^2}{N}$	N-1		
InterA (B)	$\sum_{i=1}^{K_A} \frac{T_{A,i}^2}{K_B} - \frac{T_g^2}{N}$	K_A-1	S_A^2	$\frac{S_A^2}{s_e^2}$
InterB (C)	$\sum_{j=1}^{K_B} \frac{\left(T_{B,j} - ST_j/k\right)^2}{Er}$	K_B-1	S_B^2	$\frac{S_B^2}{s_e^2}$
Intragroupe (D)	(A) - (B) - (C)	N-K_A-K_B+1	s_e^2	

avec :
ST_i est la somme des T_A intervenant dans les $T_{B,i}$ (e.g. $ST_1 = T_{A1} + T_{A4}$ pour T_{B1})
r = nombre de niveau du facteur A utilisé pour chaque niveau du facteur B
k = nombre de niveau du facteur B utilisé pour chaque niveau du facteur A
E est le facteur d'efficacité définie comme : $E = \frac{(k-1)K_B}{(K_B-1)k}$

```
%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%
%
% Exercice 3 :
% ------------
% On souhaite comparer l'efficacité de 6 méthodes de mesures pour mesurer la DCO
% des eaux de boisson. On teste 10 eaux "types", mais le volume des échantillons
% ne permettent que d'utiliser 3 des 6 méthodes. Par tirage au sort, on affect 3
% méthodes par échantillon d'eau. Les résultats sont les suivants :
% data = [156 160   0 129   0   0
%           0 199   0 180   0 180
%           0 158 157   0 144   0
%           0 109 122   0   0 115
%         212 223   0   0 150   0
%           0   0 149 183 131   0
%         203   0 189 196   0   0
%           0   0   0 201 176 196
%         349   0 335   0   0 334
%         197   0   0   0 137 174];
% chaque ligne correspond à une eau "type" et les "zéros" correspondent a une
% absence de mesure.
% Analyser ces résultats.
%
%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%
```

C/ Exercices Supplémentaires

```
%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%
%
% ===================
% ANALYSE DE VARIANCE
% ===================
%
%
%
% FONCTIONS UTILISEES :
% =====================
% ANOVA1.m   Bartlett.m   ANOVA2.m
%
% UTILISER "help name_function" pour obtenir des informations
% sur les fonctions utilisees
%
%
%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%
%
% Exercice 4 :
% ------------
% On souhaite comparer l'efficacité de 2 méthodes de mesures pour mesurer la DCO.
% Pour cela on a appliqué les 2 méthodes sur 5 échantillons d'eau tirés au
% hasard.
% Analyser les résultats obtenus. Peut-on conclure que les 2 méthodes donnent des
% résultats identiques ?
%
% méthode 1 eau 1 : 28.5  21.2  21.3  22.7
% méthode 2 eau 1 : 25.8  24.3  18.7  24.4
% méthode 1 eau 2 : 21.3  21.9  22.6  20.1
% méthode 2 eau 2 : 19.1  24.1  20.9  24.4
% méthode 1 eau 3 : 25.1  25.1  24.4  23.9
% méthode 2 eau 3 : 26.0  29.8  27.6  26.5
% méthode 1 eau 4 : 19.4  27.5  20.4  28.8
% méthode 2 eau 4 : 23.0  25.9  20.5  26.6
% méthode 1 eau 5 : 30.0  23.7  25.5  26.8
% méthode 2 eau 5 : 31.7  31.6  26.9  36.6
%
%
%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%
```

```
%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%
% Exercice 5 :
% -----------
% Des substances A et B sont supposées avoir un effet mutagène. Compte tenu de
% la durée importante des expériences, ces 2 substances ont été testées sur un
% modèle animal dans 4 laboratoires. Les substances A et B sont employées dans
% les deux premiers groupes, tandis que le 3ème reçoit une substance placebo
% (groupe témoin C). Les réactions (en unités arbitraires) sont les suivantes :
%                      Groupe
% Laboratoire      A      B     C
% 1               75     74    60
% 2               70     78    64
% 3               66     72    65
% 4               69     68    55
% 1. Sans tenir compte du laboratoire, une différence entre les trois groupes
% est-elle discernable ? Les substances A et B ont-elles un effet significatif ?
% 2. Pour affiner l'interprétation de ces résultats, on souhaite regarder
% l'influence du laboratoire qui a effectué les mesures. La différence entre
% les trois groupes est-elle encore discernable ? L'effet "laboratoire" est-il
% significatif ?
%
%
%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%
% Exercice 6 :
% -----------
% On mesure les rendements de procédés d'épuration par biomasse fixée (de type
% BIOCARBONE) en fonction du type de biomasse microbienne utilisée lors du
% démarrage du procédé, biomasse "naturelle" ou biomasse sélectionnée pour ces
% performances. 2 types d'eau ont été utilisés. Les résultats sont les suivants :
%  biomasse N  eau 1  : 84  66  56  56  81  79  62  89
%  biomasse N  eau 2  : 61  59  84  74  73  74  71  74
%  biomasse S  eau 1  : 87  89  92 101  77  95  88  91
%  biomasse S  eau 2  : 103 90 107 116  95 112  96  92
% Analyser ces résultats.
%
%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%
%
% Exercice 7 :
% -----------
% On souhaite comparer les performances  (nombre de boisseaux récoltes par ha)
% de quatre variétés de mais sur des terrain de qualité comparable. 4 variétés
% de mais (1,2,3,4) sont cultivées sur 6 parcelles de terre (A,B,C,D,E).
% La variable dépendante sera le nombre de boisseaux récoltes par ha) récoltes.
% Les données obtenues sont le suivantes :
%
%                               Parcelle
% Variété         A     B     C     D     E     F
% 1              232   279   251   278   294   284
% 2              193   220   240   249   238   220
% 3              190   208   235   190   224   215
% 4              225   201   249   251   234   198
%
%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%
%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%
%Exercice 8 :
% ------------
% Dans une étude sur l'impact des pluies acides sur l'écologie des lacs des
% Laurentides (Québec). Des chercheurs ont acidifié 4 petits cours d'eau
% expérimentaux coulant l'un à côté de l'autre. Le premier ruisseau a été
% acidifié au H2SO4 et au HNO3 ; le deuxième seulement au H2SO4 ; le troisième
% seulement au HNO3 ; et le dernier a servi de ruisseau témoin et n'a pas été
% traité. La chlorophylle a été mesurée dans les différents cours d'eau après une
% période d'acidification de 48 H. Les résultats figurent dans le tableau
% suivant :
%  H2SO4 + HNO3 +  :  1.54 1.50 0.99 1.52 1.88
%  H2SO4 + HNO3 -  :  1.65 1.50 1.18 2.19 1.11
%  H2SO4 - HNO3 +  :  1.55 1.16 1.99 1.22 1.95
%  H2SO4 - HNO3 -  :  3.28 2.94 3.50 3.07 3.04];
% Quelle sont les effets des acidifications sur la production de chlorophylle ?
%
%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%
```