

**III-1. Plan  $S<A \times B>$  avec  $A$  et  $B$  deux facteurs à effets fixes**

**III-1-1 Généralités**

Un tel plan quasi-complet défini par le croisement de deux facteurs  $A$  et  $B$ , est aussi appelé plan factoriel complet d'ordre 2. Chaque condition est répliquée sur des sujets différents. Le facteur sujet est emboîté dans le croisement  $A \times B$ . L'analyse diffère selon que l'emboîtement est équilibré ou non. Dans toute la suite on ne considère que le cas équilibré.

On se propose d'étudier l'effet de la présence simultanée d'un facteur  $A$  à  $r$  modalités et d'un facteur  $B$  à  $c$  modalités sur une variable réponse  $Y$ . Le plan étant équilibré, on dispose pour chaque croisement des modalités  $i$  du facteur  $A$  et  $j$  du facteur  $B$  de  $n$  observations. On note  $y_{s(i,j)}$  l'observation du sujet  $s$  dans le croisement  $(i,j)$ .

|   |       | Facteur $B$ |       |     |  |     |       |
|---|-------|-------------|-------|-----|--|-----|-------|
|   |       | $b_1$       | $b_2$ | ... | $b_j$                                      | ... | $B_c$ |
| A | $a_1$ |             |       |     |  |     |       |
|   | $a_2$ |             |       |     |  |     |       |
|   | ...   |             |       |     |  |     |       |
|   | $a_i$ |             |       |     | $y_{1(i,j), y_{2(i,j), \dots, y_{n(i,j)}}$ |     |       |
|   | ...   |             |       |     |  |     |       |
|   | $a_r$ |             |       |     |  |     |       |

**Remarque :** dans le cas où  $n=1$ , l'interaction des facteurs  $A$  et  $B$  est confondue avec le facteur sujet, on ne peut donc séparer les effets d'interaction des effets résiduels et on ne peut tester leur existence. Un tel plan n'est justifié que lorsque l'on peut modifier a priori la forme de l'interaction entre les facteurs. Dans toute la suite nous développons le cas  $n > 1$ .

**III-1-2 Analyse descriptive des effets (principaux et interaction) :**

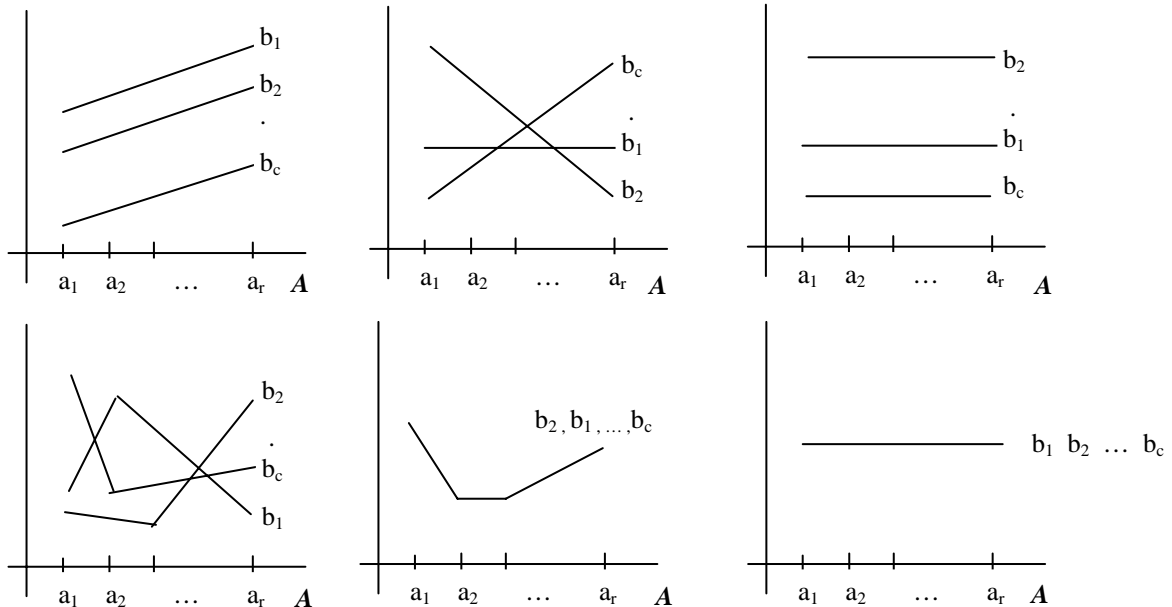
Considérons, pour chaque croisement  $(i,j)$  de la modalité  $i$  de  $A$  et de la modalité  $j$  de  $B$  la moyenne obtenue :

|          |       | Facteur $B$    |                |     |                |     |                | moyennes                 |
|----------|-------|----------------|----------------|-----|----------------|-----|----------------|--------------------------|
|          |       | $b_1$          | $b_2$          | ... | $b_j$          | ... | $B_c$          |                          |
| A        | $a_1$ | $\bar{y}_{11}$ | $\bar{y}_{12}$ | ... | $\bar{y}_{1j}$ | ... | $\bar{y}_{1c}$ | $\bar{y}_{10}$           |
|          | $a_2$ | $\bar{y}_{21}$ | $\bar{y}_{22}$ | ... | $\bar{y}_{2j}$ | ... | $\bar{y}_{2c}$ | $\bar{y}_{20}$           |
|          | ...   | ...            | ...            | ... | ...            | ... | ...            | ...                      |
|          | $a_i$ | $\bar{y}_{i1}$ | $\bar{y}_{i2}$ | ... | $\bar{y}_{ij}$ | ... | $\bar{y}_{ic}$ | $\bar{y}_{i0}$           |
|          | ...   | ...            | ...            | ... | ...            | ... | ...            | ...                      |
|          | $a_r$ | $\bar{y}_{r1}$ | $\bar{y}_{r2}$ | ... | $\bar{y}_{rj}$ | ... | $\bar{y}_{rc}$ | $\bar{y}_{r0}$           |
| moyennes |       | $\bar{y}_{01}$ | $\bar{y}_{02}$ | ... | $\bar{y}_{0j}$ | ... | $\bar{y}_{0c}$ | $\bar{y}_{00} = \bar{y}$ |

La colonne marginale des moyennes par modalité de  $A$  (indépendamment de  $B$ ) « traduit » l'effet principal du facteur  $A$ . De même la ligne marginale des moyennes par modalité de  $B$  (indépendamment de  $A$ ) « traduit » l'effet principal du facteur  $B$ . Les différentes moyennes correspondant au croisement des modalités « traduisent » l'interaction des deux facteurs  $A$  et  $B$ .

### III-1-3 Représentation géométrique des différents effets :

Les différentes combinaisons des effets principaux et d'interaction peuvent être illustrés par des graphiques dans lesquels on représente les moyennes en fonction des modalités de l'un des facteurs. Il y a deux graphiques possibles selon le facteur porté en abscisse.



Le parallélisme des lignes traduit le fait que l'effet du facteur A (ou B) sur la variable réponse Y est le même quelque soit la valeur du facteur B (ou A). Les effets des facteurs A et B sont additifs et il n'y a pas d'effet de l'interaction entre les deux facteurs A et B sur Y.

Le non parallélisme des lignes traduit le fait que l'effet du facteur A (ou du facteur B) n'est pas le même en fonction de la modalité du facteur B (ou du facteur A). Les effets des facteurs A et B sur la variable Y ne sont pas additifs. Aux effets simples des facteurs s'ajoutent l'effet de l'interaction entre les deux facteurs.

Le parallélisme avec l'axe des abscisses traduit l'absence d'effet du facteur A (ou B). De plus si les différentes lignes correspondant aux modalités de B (ou A) sont confondues c'est l'absence totale d'effets.

### III-1-4 Le modèle :

Rappelons que l'on veut étudier l'effet de la présence simultanée d'un facteur A à r modalités et d'un facteur B à c modalités sur une variable réponse Y. Le plan étant équilibré, on dispose pour chaque modalité i du facteur A et pour chaque modalité j du facteur B de n observations ( $n > 1$ ). On note  $y_{s(i,j)}$  l'observation du sujet s dans le croisement (i,j).  $y_{s(i,j)}$  est la réalisation de la variable aléatoire  $Y_{s(i,j)}$  décrite par le modèle suivant :

$$Y_{s(i,j)} = \mu_{ij} + \varepsilon_{s(i,j)}$$

où les variables aléatoires  $\varepsilon_{s(i,j)}$  sont indépendantes et de même loi  $\mathcal{N}(0, \sigma^2)$  et

$$\mu_{ij} = \mu + \alpha_i + \beta_j + (\alpha\beta)_{ij} \quad \text{où}$$

- $\mu$  mesure la moyenne de la population
- les  $\alpha_i$  mesurent les effets principaux du facteur A
- les  $\beta_j$  mesurent les effets principaux du facteur B
- les  $(\alpha\beta)_{ij}$  mesurent les effets d'interaction.

On rajoute (comme hypothèses) les contraintes d'identifiabilité suivantes :

$$\sum_i \alpha_i = 0 \quad \sum_j \beta_j = 0 \quad \sum_j (\alpha\beta)_{ij} = 0 \quad \sum_i (\alpha\beta)_{ij} = 0$$

### III-1-5 Décomposition de la variation :

A partir du tableau des données on calcule les Sommes des Carrés associées aux différentes sources de variation. Le croisement des facteurs A et B permet de décomposer la variation totale en somme de deux variations :

$$\mathbf{SCT}_{\text{obs}} = \mathbf{SCI}_{\text{intercase}} + \mathbf{SCI}_{\text{intracase}} = \mathbf{SC}(A \times B)_{\text{obs}} + \mathbf{SCR}_{\text{obs}}$$

avec :

$$\mathbf{SCT}_{\text{obs}} = \sum_i \sum_j \sum_s (y_{s(i,i)} - \bar{y})^2 \quad \mathbf{SC}(A \times B)_{\text{obs}} = n \sum_i \sum_j (\bar{y}_{ij} - \bar{y})^2 \quad \mathbf{SCR}_{\text{obs}} = \sum_i \sum_j \sum_s (y_{s(i,i)} - \bar{y}_{ij})^2$$

D'autre part le facteur élémentaire A (resp B) définit une partition de l'ensemble des observations en classes à laquelle correspond une somme de carrés observée mesurant la variation associée à l'effet principal de A (resp de B) :

$$\mathbf{SCA}_{\text{obs}} = nc \sum_i (\bar{y}_{i0} - \bar{y})^2 \quad \mathbf{SCB}_{\text{obs}} = nr \sum_j (\bar{y}_{0j} - \bar{y})^2$$

**Définition :** On appelle somme des carrés associée à l'interaction la quantité :

$$\mathbf{SCAB}_{\text{obs}} = \mathbf{SC}(A \times B)_{\text{obs}} - \mathbf{SCA}_{\text{obs}} - \mathbf{SCB}_{\text{obs}}$$

$$\text{On montre que : } \mathbf{SCAB}_{\text{obs}} = \sum_i \sum_j n(\bar{y}_{ij} - \bar{y}_{i0} - \bar{y}_{0j} + \bar{y})^2$$

Dans le cadre du modèle statistique ces sommes de carrés sont des réalisations de variables aléatoires dont on calcule les espérances et plus généralement les distributions des probabilités.

On montre le théorème fondamental suivant :

**Théorème :** Sous les hypothèses du modèle, les statistiques **SCA**, **SCB**, **SCAB** et **SCR** sont indépendantes et d'espérances respectives :

$$E(\mathbf{SCA}) = (r-1)\sigma^2 + \sum_i nc\alpha_i^2$$

$$E(\mathbf{SCB}) = (c-1)\sigma^2 + \sum_j nr\beta_j^2$$

$$E(\mathbf{SCAB}) = (r-1)(c-1)\sigma^2 + \sum_i \sum_j n(\alpha\beta)_{ij}^2$$

$$E(\mathbf{SCR}) = (N - rc)\sigma^2 = rc(n-1)\sigma^2$$

On ramène toutes les sommes de carrés à des moyennes de carrés en divisant par les degrés de liberté correspondant.  $\mathbf{MCA} = \frac{\mathbf{SCA}}{r-1}$   $\mathbf{MCB} = \frac{\mathbf{SCB}}{c-1}$   $\mathbf{MCAB} = \frac{\mathbf{SCAB}}{(r-1)(c-1)}$   $\mathbf{MCR} = \frac{\mathbf{SCR}}{rc(n-1)}$

Ceci permet de réécrire le théorème précédent pour les moyennes des carrés :

**Théorème :** Sous les hypothèses du modèle, les statistiques **MCA**, **MCB**, **MCAB** et **MCR** sont indépendantes et d'espérances respectives :

$$E(\mathbf{MCA}) = \sigma^2 + \frac{nc}{(r-1)} \sum_i \alpha_i^2$$

$$E(\mathbf{MCB}) = \sigma^2 + \frac{nr}{(c-1)} \sum_j \beta_j^2$$

$$E(\mathbf{MCAB}) = \sigma^2 + \frac{n}{(r-1)(c-1)} \sum_i \sum_j (\alpha\beta)_{ij}^2$$

$$E(\mathbf{MCR}) = \sigma^2$$

### III-1-6 Tests et décisions statistiques :

Le théorème précédent nous permet de tester l'existence des différents effets des facteurs. On peut construire des tests indépendants sur chacune des sources de variations :

#### Test 1 :

hypothèse nulle  $H_{0A}$  : pas d'effet principal du facteur A  $\Leftrightarrow (\forall i) \alpha_i = 0$  contre  
l'hypothèse alternative :  $H_{1A}$  : il existe un effet principal du facteur A  $\Leftrightarrow (\exists i) \alpha_i \neq 0$

#### Test 2 :

hypothèse nulle  $H_{0B}$  : pas d'effet principal du facteur B  $\Leftrightarrow (\forall j) \beta_j = 0$  contre  
l'hypothèse alternative :  $H_{1B}$  : il existe un effet principal du facteur B  $\Leftrightarrow (\exists j) \beta_j \neq 0$

#### Test 3 :

hypothèse nulle  $H_{0AB}$  : pas d'effet d'interaction  $\Leftrightarrow (\forall i, j) (\alpha\beta)_{ij} = 0$  contre  
l'hypothèse alternative :  $H_{1AB}$  : il existe une interaction  $\Leftrightarrow (\exists i, j) (\alpha\beta)_{ij} \neq 0$

La construction de ces tests résulte des trois corollaires du théorème suivant :

#### Théorème :

Dans le cadre du modèle statistique, la statistique  $SCR/\sigma^2$  suit une loi de  $\chi^2$  à  $rc(n-1)$  degrés de liberté, sous l'hypothèse nulle  $H_{0A}$  la statistique  $SCA/\sigma^2$  suit une loi de  $\chi^2$  à  $r-1$  degrés de liberté, sous l'hypothèse nulle  $H_{0B}$  la statistique  $SCB/\sigma^2$  suit une loi de  $\chi^2$  à  $c-1$  degrés de liberté, sous l'hypothèse nulle  $H_{0AB}$  la statistique  $SCAB/\sigma^2$  suit une loi de  $\chi^2$  à  $(r-1)(c-1)$  degrés de liberté et ces statistiques sont indépendantes.

**Corollaire :** sous l'hypothèse  $H_{0A}$ , la statistique  $F_A = \frac{MCA}{MCR}$  suit une loi de Fischer à  $r-1$  et  $rc(n-1)$  degrés de liberté.

Le test1 est alors défini au seuil de signification  $\alpha$  par la règle de décision suivante :

si  $F_{A \text{ obs}} \geq \lambda_\alpha$  alors on rejette l'hypothèse nulle  
où  $\lambda_\alpha$  est donné par l'équation :  $\alpha = P(F_A \geq \lambda_\alpha)$ .

**Corollaire :** sous l'hypothèse  $H_{0B}$ , la statistique  $F_B = \frac{MCB}{MCR}$  suit une loi de Fischer à  $c-1$  et  $rc(n-1)$  degrés de liberté.

Le test2 est alors défini au seuil de signification  $\alpha$  par la règle de décision suivante :

si  $F_{B \text{ obs}} \geq \lambda_\alpha$  alors on rejette l'hypothèse nulle  
où  $\lambda_\alpha$  est donné par l'équation :  $\alpha = P(F_B \geq \lambda_\alpha)$ .

**Corollaire :** sous l'hypothèse  $H_{0AB}$ , la statistique  $F_{AB} = \frac{MCAB}{MCR}$  suit une loi de Fischer à  $(r-1)(c-1)$  et  $rc(n-1)$  degrés de liberté.

Le test3 est alors défini au seuil de signification  $\alpha$  par la règle de décision suivante :

si  $F_{AB \text{ obs}} \geq \lambda_\alpha$  alors on rejette l'hypothèse nulle  
où  $\lambda_\alpha$  est donné par l'équation :  $\alpha = P(F_{AB} \geq \lambda_\alpha)$ .

On présente l'étude dans le :

#### tableau d'analyse de la variance :

| Source | $SC_{obs}$   | ddl              | $MC_{obs}$   | $F_{obs}$     |
|--------|--------------|------------------|--------------|---------------|
| A      | $SCA_{obs}$  | $r - 1$          | $MCA_{obs}$  | $F_{A\ obs}$  |
| B      | $SCB_{obs}$  | $c - 1$          | $MCB_{obs}$  | $F_{B\ obs}$  |
| AB     | $SCAB_{obs}$ | $(r - 1)(c - 1)$ | $MCAB_{obs}$ | $F_{AB\ obs}$ |
| R      | $SCR_{obs}$  | $rc(n - 1)$      | $MCR_{obs}$  |               |
| Total  | $SCT_{obs}$  | $rcn - 1$        |              |               |

**Pratique des calculs :**

Calcul de SCA :

$$SCA_{obs} = nc \sum_i (\bar{y}_{i0} - \bar{y})^2 = ncr \times \text{variance des moyennes de A}$$

Calcul de SCB :

$$SCB_{obs} = nr \sum_{j=1}^c (\bar{y}_{0j} - \bar{y})^2 = ncr \times \text{variance des moyennes de B}$$

Calcul de SC(A×B) :

$$SC(A \times B)_{obs} = n \sum_{i=1}^r \sum_{j=1}^c (\bar{y}_{ij} - \bar{y})^2 = ncr \times \text{variance des moyennes des carrés croisés}$$

Calcul de SC(AB) :

$$SC(AB)_{obs} = SC(A \times B)_{obs} - SCA_{obs} - SCB_{obs}$$

Calcul de SCT :

$$SCT_{obs} = ncr \times \text{variance de toutes les données}$$

Calcul de SCR :

$$SCR_{obs} = SCT_{obs} - SC(A \times B)_{obs} = SCT_{obs} - SC(AB)_{obs} - SCA_{obs} - SCB_{obs}$$

### III-1-7 Exemple :

Un expérimentateur veut mettre à l'épreuve l'hypothèse que le niveau d'excitation physiologique d'une personne frustrée dépend du type d'agression autorisée. Il construit alors l'expérience suivante : 30 sujets choisis au hasard sont répartis de façon aléatoire en 6 groupes de 5 sujets, correspondant aux 6 conditions expérimentales définies par le croisement de deux facteurs

A : Facteur frustration :

2 modalités : A<sub>1</sub> : frustré      A<sub>2</sub> : non frustré

B : Facteur agression :

3 modalités : B<sub>1</sub> : non agressif    B<sub>2</sub> : moyennement agressif  
B<sub>3</sub> : très agressif

Les résultats sont donnés par le tableau ci-contre.

|                | A <sub>1</sub> |   |   |   |   | A <sub>2</sub> |   |    |    |    |
|----------------|----------------|---|---|---|---|----------------|---|----|----|----|
| B <sub>1</sub> | 1              | 6 | 4 | 3 | 0 | 8              | 8 | 10 | 8  | 11 |
| B <sub>2</sub> | 3              | 5 | 2 | 4 | 4 | 6              | 8 | 4  | 4  | 3  |
| B <sub>3</sub> | 6              | 7 | 4 | 5 | 3 | 5              | 7 | 3  | -1 | 1  |

#### Calcul de SCA :

$$SCA_{obs} = nc \sum_i (\bar{y}_{i0} - \bar{y})^2 = ncr \times \text{variance}$$

des moyennes de A

$$SCA_{obs} = 30 \times 1,0 = 30,0$$

#### Calcul de SCB :

$$SCB_{obs} = nr \sum_j (\bar{y}_{00j} - \bar{y})^2 = ncr \times \text{variance}$$

des moyennes de B     $SCB_{obs} = 30 \times 0,6067 = 18,2$

#### Calcul de SC(A×B) :

$$SC(A \times B)_{obs} = n \sum_{i=1}^r \sum_j (\bar{y}_{ij} - \bar{y})^2 = ncr \times \text{variance des moyennes des carrés croisés}$$

$$SC(A \times B)_{obs} = 30 \times 4,1867 = 125,6$$

#### Calcul de SC(AB) :

$$SC(AB)_{obs} = SC(A \times B)_{obs} - SCA_{obs} - SCB_{obs} = 125,6 - 30,0 - 18,2 = 77,4$$

#### Calcul de SCT :

$$SCT_{obs} = ncr \times \text{variance de toutes les données} \quad SCT_{obs} = 30 \times 7,16 = 214,8$$

#### Calcul de SCR :

$$SCR_{obs} = SCT_{obs} - SC(A \times B)_{obs} = 214,8 - 125,6 = 89,2$$

On présente l'étude dans le :

**tableau d'analyse de la variance :**

| Source | SC <sub>obs</sub>          | ddl                | MC <sub>obs</sub>           | F <sub>obs</sub>              |
|--------|----------------------------|--------------------|-----------------------------|-------------------------------|
| A      | SCA <sub>obs</sub> = 30,0  | r - 1 = 1          | MCA <sub>obs</sub> = 30,0   | F <sub>A obs</sub> = 8,0717   |
| B      | SCB <sub>obs</sub> = 18,2  | c - 1 = 2          | MCB <sub>obs</sub> = 9,1    | F <sub>B obs</sub> = 2,4484   |
| AB     | SCAB <sub>obs</sub> = 77,4 | (r - 1)(c - 1) = 2 | MCAB <sub>obs</sub> = 38,7  | F <sub>AB obs</sub> = 10,4126 |
| R      | SCR <sub>obs</sub> = 89,2  | rc(n - 1) = 24     | MCR <sub>obs</sub> = 3,7167 |                               |
| Total  | SCT <sub>obs</sub> = 214,8 | rcn - 1 = 29       |                             |                               |