Plan d'expérience

Pierre Druilhet

Université Blaise Pascal, Clermont Ferrand

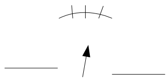
Do Well B. 2013

Première partie I

Pourquoi faire un plan d'expérience ?

Exemple introductif : le plan de pesées

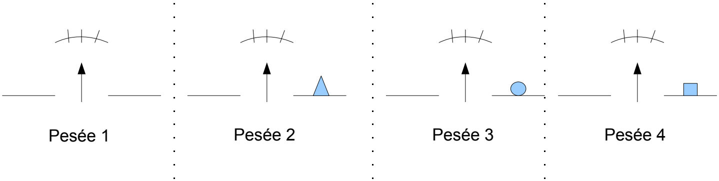
- On a trois objets à peser à l'aide d'une balance de Roberval.
- On n'a droit qu'à 4 pesées. Au départ, la balance est déséquilibrée (défaut de calibrage)



Balance à vide

- La précision de la balance est donnée par la variance des **erreurs de mesures** σ^2.
- Objectif : estimer avec le plus de précision possible le poids des trois objets.

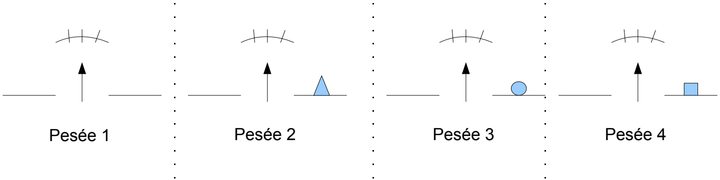
Premier plan de pesées



Modèle statistique associé

$$
\begin{cases}
y_1 = p_o + \varepsilon_1 \\
y_2 = p_o + p_1 + \varepsilon_2 \\
y_3 = p_o + p_2 + \varepsilon_3 \\
y_4 = p_o + p_3 + \varepsilon_4
\end{cases}
$$

Premier plan de pesées



Modèle statistique associé

$$\begin{cases} y_1 = p_o + \varepsilon_1 \\ y_2 = p_o + p_1 + \varepsilon_2 \\ y_3 = p_o + p_2 + \varepsilon_3 \\ y_4 = p_o + p_3 + \varepsilon_4 \end{cases}$$

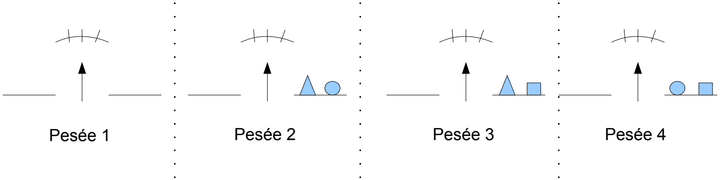
Modèle statistique associé

$$\begin{pmatrix} y_1 \\ y_2 \\ y_3 \\ y_4 \end{pmatrix} = \begin{pmatrix} 1 & 0 & 0 & 0 \\ 1 & 1 & 0 & 0 \\ 1 & 0 & 1 & 0 \\ 1 & 0 & 0 & 1 \end{pmatrix} \begin{pmatrix} p_0 \\ p_1 \\ p_2 \\ p_3 \end{pmatrix} + \begin{pmatrix} \varepsilon_1 \\ \varepsilon_2 \\ \varepsilon_3 \\ \varepsilon_4 \end{pmatrix} = X\,p + \varepsilon$$

Estimation par MCO

$$\begin{array}{lcl} \hat{p}_1 = y_2 - y_1 & \longrightarrow & \mathrm{var}(\hat{p}_1) = 2\sigma^2 \\ \hat{p}_2 = y_3 - y_1 & \longrightarrow & \mathrm{var}(\hat{p}_2) = 2\sigma^2 \\ \hat{p}_3 = y_4 - y1 & \longrightarrow & \mathrm{var}(\hat{p}_3) = 2\sigma^2 \end{array}$$

Deuxième plan de pesées



Modèle statistique associé

$$\begin{pmatrix} y_1 \\ y_2 \\ y_3 \\ y_4 \end{pmatrix} = \begin{pmatrix} 1 & 0 & 0 & 0 \\ 1 & 1 & 1 & 0 \\ 1 & 1 & 0 & 1 \\ 1 & 0 & 1 & 1 \end{pmatrix} \begin{pmatrix} p_0 \\ p_1 \\ p_2 \\ p_3 \end{pmatrix} + \begin{pmatrix} \varepsilon_1 \\ \varepsilon_2 \\ \varepsilon_3 \\ \varepsilon_4 \end{pmatrix} = X\,p + \varepsilon$$

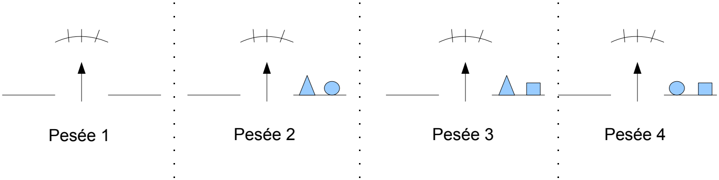
Estimation par MCO

$$\hat{p}_1 = \frac{1}{2}(y_2 + y_3 - y_1 - y_2) \quad \longrightarrow \quad \mathrm{var}(\hat{p}_1) = \sigma^2$$

$$\hat{p}_2 = \frac{1}{2}(y_2 + y_4 - y_1 - y_3) \quad \longrightarrow \quad \mathrm{var}(\hat{p}_2) = \sigma^2$$

$$\hat{p}_3 = \frac{1}{2}(y_3 + y_4 - y_1 - y_2) \quad \longrightarrow \quad \mathrm{var}(\hat{p}_3) = \sigma^2$$

Deuxième plan de pesées



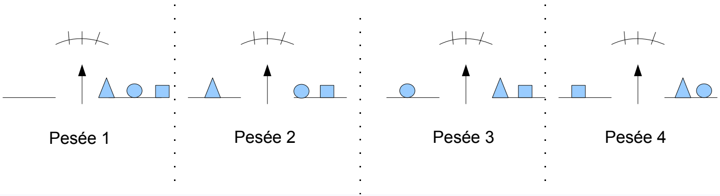
Modèle statistique associé

$$\begin{pmatrix} y_1 \\ y_2 \\ y_3 \\ y_4 \end{pmatrix} = \begin{pmatrix} 1 & 0 & 0 & 0 \\ 1 & 1 & 1 & 0 \\ 1 & 1 & 0 & 1 \\ 1 & 0 & 1 & 1 \end{pmatrix} \begin{pmatrix} p_0 \\ p_1 \\ p_2 \\ p_3 \end{pmatrix} + \begin{pmatrix} \varepsilon_1 \\ \varepsilon_2 \\ \varepsilon_3 \\ \varepsilon_4 \end{pmatrix} = X\,p + \varepsilon$$

Estimation par MCO

$$\begin{aligned} \hat{p}_1 &= \tfrac{1}{2}(y_2 + y_3 - y_1 - y_2) & \longrightarrow \quad & \mathrm{var}(\hat{p}_1) = \sigma^2 \\ \hat{p}_2 &= \tfrac{1}{2}(y_2 + y_4 - y_1 - y_3) & \longrightarrow \quad & \mathrm{var}(\hat{p}_2) = \sigma^2 \\ \hat{p}_3 &= \tfrac{1}{2}(y_3 + y_4 - y_1 - y_2) & \longrightarrow \quad & \mathrm{var}(\hat{p}_3) = \sigma^2 \end{aligned}$$

Troisième plan de pesées



Modèle statistique associé

$$\begin{pmatrix} y_1 \\ y_2 \\ y_3 \\ y_4 \end{pmatrix} = \begin{pmatrix} 1 & 1 & 1 & 1 \\ 1 & -1 & 1 & 1 \\ 1 & 1 & -1 & 1 \\ 1 & 1 & 1 & -1 \end{pmatrix} \begin{pmatrix} p_0 \\ p_1 \\ p_2 \\ p_3 \end{pmatrix} + \begin{pmatrix} \varepsilon_1 \\ \varepsilon_2 \\ \varepsilon_3 \\ \varepsilon_4 \end{pmatrix} = X\,p + \varepsilon$$

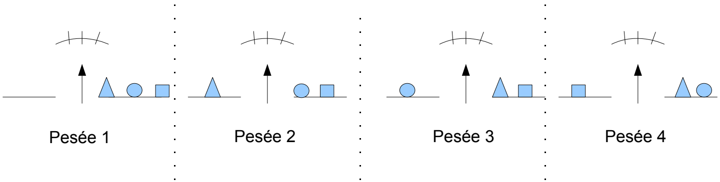
Estimation par MCO

$$\hat{p}_1 = \frac{1}{2}(y_1 - y_2) \quad \longrightarrow \quad \mathrm{var}(\hat{p}_1) = \frac{1}{2}\sigma^2$$

$$\hat{p}_2 = \frac{1}{2}(y_1 - y_3) \quad \longrightarrow \quad \mathrm{var}(\hat{p}_2) = \frac{1}{2}\sigma^2$$

$$\hat{p}_3 = \frac{1}{2}(y_1 - y_4) \quad \longrightarrow \quad \mathrm{var}(\hat{p}_3) = \frac{1}{2}\sigma^2$$

Troisième plan de pesées



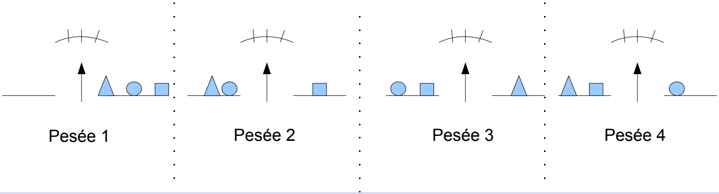
Modèle statistique associé

$$\begin{pmatrix} y_1 \\ y_2 \\ y_3 \\ y_4 \end{pmatrix} = \begin{pmatrix} 1 & 1 & 1 & 1 \\ 1 & -1 & 1 & 1 \\ 1 & 1 & -1 & 1 \\ 1 & 1 & 1 & -1 \end{pmatrix} \begin{pmatrix} p_0 \\ p_1 \\ p_2 \\ p_3 \end{pmatrix} + \begin{pmatrix} \varepsilon_1 \\ \varepsilon_2 \\ \varepsilon_3 \\ \varepsilon_4 \end{pmatrix} = X\,p + \varepsilon$$

Estimation par MCO

$$\begin{aligned} \hat{p}_1 &= \tfrac{1}{2}(y_1 - y_2) & \longrightarrow \quad \mathrm{var}(\hat{p}_1) &= \tfrac{1}{2}\sigma^2 \\ \hat{p}_2 &= \tfrac{1}{2}(y_1 - y_3) & \longrightarrow \quad \mathrm{var}(\hat{p}_2) &= \tfrac{1}{2}\sigma^2 \\ \hat{p}_3 &= \tfrac{1}{2}(y_1 - y_4) & \longrightarrow \quad \mathrm{var}(\hat{p}_3) &= \tfrac{1}{2}\sigma^2 \end{aligned}$$

Quatrième plan de pesées



Modèle statistique associé

$$\begin{pmatrix} y_1 \\ y_2 \\ y_3 \\ y_4 \end{pmatrix} = \begin{pmatrix} 1 & 1 & 1 & 1 \\ 1 & -1 & -1 & 1 \\ 1 & 1 & -1 & -1 \\ 1 & -1 & 1 & -1 \end{pmatrix} \begin{pmatrix} p_0 \\ p_1 \\ p_2 \\ p_3 \end{pmatrix} + \begin{pmatrix} \varepsilon_1 \\ \varepsilon_2 \\ \varepsilon_3 \\ \varepsilon_4 \end{pmatrix} = X\,p + \varepsilon$$

Estimation par MCO

$$\begin{aligned} \hat{p}_1 &= \tfrac{1}{4}(y_1 - y_2 + y_3 - y_4) & \longrightarrow \quad \mathrm{var}(\hat{p}_1) &= \tfrac{1}{4}\sigma^2 \\ \hat{p}_2 &= \tfrac{1}{4}(y_1 - y_2 - y_3 + y_4) & \longrightarrow \quad \mathrm{var}(\hat{p}_2) &= \tfrac{1}{4}\sigma^2 \\ \hat{p}_3 &= \tfrac{1}{4}(y_1 + y_2 - y_3 - y_4) & \longrightarrow \quad \mathrm{var}(\hat{p}_3) &= \tfrac{1}{4}\sigma^2 \end{aligned}$$

Conclusion de l'exemple

Que retenir de l'exemple :

- La stratégie de planification va influencer directement la précision statistique des résultats.
- On peut montrer que le dernière plan de pesées est optimal : la précision des estimateurs est optimale.
- Dans le dernier plan de pesées, le plan est orthogonal. Tout se passe comme si l'on avait pesé 4 fois chaque objet avec une balance équilibrée.

Principe d'une démarche expérimentale

1. Définition des objectifs de l'expérience par l'expérimentateur :
   - estimation
   - test d'hypothèse
   - prévision
   - discrimination
   - ...
2. Quelle sont les contraintes expérimentales :
   - contraintes de temps,
   - contrainte d'espace,
   - contraintes de coût,
   - contraintes éthiques
   - ...

3. Définition des facteurs pouvant influencer l'expérience :
   - température pour une expérience biologique.
   - âge, état de santé, sexe d'un patient
   - effets périodes pour des mesures répétées dans le temps.
   - ...
4. Détermination des sources de variabilité possibles :
   - hétérogénéité du terrain,
   - différents fournisseurs d'animaux,
   - expérience sur plusieurs terrains exposés différemment,
   - effet sujets, expérimentateur,...
   - effet du centre d'expérimentation (essais multicentrique).
   - effet de l'étude (méta-analyse)
   - ...
5. Définition des unités expérimentales :
   - individu,
   - individu x période,
   - parcelle de terrain,
   - ...

6. Élaboration de l'expérience = **planification** : à partir des données précédentes, élaborer un plan qui
   - permette la meilleure précision statistique des résultats,
   - garantisse l'impartialité de l'étude (randomisation + simple ou double aveugle si possible).
7. Réalisation des expériences,
8. Analyse des résultats.

Comment gagner de la précision statistique ?

- Augmenter le nombre de répétitions :
  - la variance des estimateurs diminue,
  - mais la variance résiduelle peut augmenter si on est obligé de prendre une population plus hétérogène.
    Exemple : tests sur des rats élevés dans un même laboratoire vs rats obtenus par différents fournisseurs. Ou bien nécessité de mener l'expérience sur plusieurs champs, plusieurs centres.
- Contrôle local (ou blocking) : on regroupe les unités expérimentales en fonctions de critères ou facteur :
  - diminue la variance et augmente la puissance des tests en enlevant des sources de variabilité.
  - peut augmenter la variance si le plan n'est pas orthogonal et l'homogénéisation insuffisante.
  - accroît la complexité de l'analyse.

Deuxième partie II

Techniques de blocage (blocking)

Exemple 2.1Choix du nombre de block

Problème

On dispose de p lots d'animaux de taille n. On teste un médicament et on observe la réponse x. A taille de l'expérience np fixée comment choisir n et p pour avoir la meilleure réponse ?

Réponse

La réponse s'écrit $x_{ij} = \mu + a_i + \varepsilon_{ij}$. On note $\sigma_A^2 = \mathbb{V}\mathrm{ar}(a_i)$ la variabilité inter lots et $\sigma^2 = \mathbb{V}\mathrm{ar}(\varepsilon_{ij})$ la variabilité intra lot, on a :

$$\mathbb{V}\mathrm{ar}(\bar{X}) = \frac{\sigma_A^2}{p} + \frac{\sigma^2}{np}$$

est minimum pour $n = 1$, i.e. p maximal. On diminue ainsi l'impact de la variabilité inter lot.

Exemple 2.2 : blocage spatial

On souhaite tester 5 variétés de blés. On découpe un champ en parcelles. Pour contrôler les variations de fertilité, on regroupe les parcelles en blocs homogènes.

Bloc 1	1	2	3	4	5
Bloc 2	3	1	5	2	4
			⋮		
Bloc b	4	3	1	5	2

Définition

Plan complet équilibré : Chaque traitement apparaît le même nombre de fois dans chaque bloc (équilibré).

- C'est le plan qui est universellement optimal parmi tous les plans possibles.
- Le caractère aléatoire ou fixe des blocs ne va pas influencer l'analyse sur les traitements
- L'estimation des effets traitements n'est pas modifiée si on enlève l'effet bloc du modèle (car non significatif).
- L'impact du "blocage" se fait essentiellement sur l'estimation de la variance résiduelle (dénominateur du test de student ou de Fisher).

Si la taille des blocs est $< t$, on ne peut pas construire un plan équilibré $\rightsquigarrow$

Définition

Plans en blocs incomplets équilibrés si $k < t$

- chaque traitement apparaît le même nombre de fois, noté r dans le plan (plan équirépété).
- chaque paire de traitements distincts apparaît λ fois ensemble (i.e. dans un même bloc) :

$$\lambda = \frac{r(k-1)}{I-1}$$

où $I = bk/r$

Exemple 2.3 : PBIE ($b = 10$, $k = 3$, $t = 5$, $r = 6$, $\lambda = 3$).

bloc 1	1	2	3
bloc 2	2	3	5
bloc 3	1	3	4
bloc 4	2	4	5
bloc 5	1	3	5
bloc 6	1	2	5
bloc 7	2	3	4
bloc 8	1	2	4
bloc 9	3	4	5
bloc 10	1	4	5

Propriétés des PBIE

- Les PBIE sont universellement optimaux parmi tous les plans de même taille.
- Les PBIE ne sont pas des plans orthogonaux en général (sauf si complet).
- Il faudra tenir compte du caractère aléatoire ou non des blocs.

Faut-il utiliser des plans en blocs ?

Les plans en blocs permettent de regrouper des unités expérimentales qui se ressemblent.

- Si les unités expérimentales ne sont pas raisonnablement similaires, on a toujours intérêt à les regrouper en blocs,
- mais utiliser des blocs de manière injustifiée peut diminuer la performance d'un plan d'expérience,
- ...sauf si on utilise un plan en blocs complets équilibrés.

Troisième partie III

Randomisation

Introduction

Les expériences randomisées ont été introduite dans les années 1920 par Fisher en agronomie.

Exemple

On veut comparer t engrais ou variétés de plantes. On découpe le champs en parcelles (sans structure de blocs). Sur chaque parcelle, on applique un traitement.

Plan systématique

Un premier plan qui vient à l'esprit est un **plan systématique.**

1	2	3	1	2	3	1	2	3

Inconvénient : il peut y avoir une structure cachée de fertilité (due par exemple à un essai antérieur) qui corresponde à la structure du plan systématique. Cela peut favoriser une traitement.

Plan avec parcelles témoins

- Une autre possibilité est d'utiliser des **parcelles témoins**. Par exemple

1	T	2	3	T	1	2	T	3

- On peut montrer que le nombre n_o optimal de témoins pour une expérience avec t traitements et r répétitions est :

$$n_o = r\sqrt{t}$$

Par exemple, si t=4, il faudra mettre deux fois plus de répétitions pour les témoins que pour les autres traitements.

- Inconvénient des témoins : les parcelles utilisées pour les témoins sont perdues pour les traitements. Cette méthode ancienne est maintenant peu utilisée.

Définition

La randomisation est une affection aléatoire des traitements.

- La randomisation a transformé le biais en variance
- la randomisation évite le **biais de sélection** par l'expérimentateur :
- **La randomisation garantit l'objectivité de l'étude !**

Randomisation restreinte

- La randomisation totale peut aboutir à une disposition indésirable. Par exemple, on peut avoir :

1	1	1	2	2	2	3	3	3

  même si c'est avec une probabilité faible.
- Dans le même ordre d'idée, si on a une structure en bloc, on veut que la randomisation respecte cette structure.
- Il faut s'assurer que la randomisation valide le modèle.

Randomisation d'un plan en bloc

On considère un plan en b blocs de taille k. On peut conduire la randomisation suivante :

1. On construit un plan en bloc virtuel (non réalisé), par exemple un PBCE ou un PBIE.
2. On crée une permutation aléatoire des blocs.
3. A l'intérieur de chaque bloc, on crée une permutation aléatoire des parcelles

Remarque

- *Si le plan initial est un PBCE (resp. un PBIE), il le restera après randomisation.*
- *Les blocs devenant aléatoires, ils devront être traités comme des effets aléatoires dans l'analyse.*

Exemple : randomisation d'un PBIE

1. Construction d'un PBIE virtuel.

bloc 1	1	2	3
bloc 2	1	2	4
bloc 3	1	3	4
bloc 4	2	3	4

2. Permutation aléatoire des blocs : $1 \to 3$, $2 \to 2$, $3 \to 4$ et $4 \to 1$.

bloc 4	2	3	4
bloc 2	1	2	4
bloc 1	1	2	3
bloc 3	1	3	4

3. On permute aléatoirement les parcelles dans les blocs. On obtient le plan final :

2	4	3
4	1	2
4	3	1
1	2	3

Notion générale de randomisation

On peut distinguer trois types de randomisations restreintes

- **Type I** : randomisation des parcelles. À partir d'un plan virtuel donné, on choisit aléatoirement une permutation des parcelles parmi un ensemble fixé de permutations possibles.
- **Type II** : randomisation des traitements. À partir d'un plan virtuel donné, on choisit aléatoirement une permutation des traitements parmi un ensemble fixé de permutations possibles.
- **Type III** : sélection aléatoire d'un plan. On choisit aléatoirement un plan parmi un ensemble fixé de plans.

Remarque

Notons que le type III est la forme la plus générale d'envisager la randomisation.

Propriétés d'une randomisation

Les trois principales propriétés recherchées pour une randomisation sont :

- **le non biais** : si l'on s'intéresse à l'estimation des contrastes des traitements, on désire que l'espérance (à travers la randomisation) des estimateurs des contrastes soit non biaisée.
- **la validité faible** : on cherche à "valider" le test de nullité des effets traitements, en imposant au numérateur et au dénominateur du test de Fisher d'être en moyenne égaux sous l'hypothèse nulle : on désire que, en l'absence d'effets traitements, l'espérance du carré moyen associé aux traitements soit égale à l'espérance du carré moyen associé à l'erreur.
- **la validité forte** : on cherche à "valider" le test correspondant à une sous-hypothèse du test précédent : on désire que, en l'absence d'effets traitements, l'espérance du carré moyen associé à tout sous-espace de l'espace des effets traitements soit égal à l'espérance du carré moyen associé à l'erreur.

Proposition

Bailey & Rowley, 1987
(i) La randomisation d'un plan en blocs complets à t traitements est faiblement valide si, pour chaque couple de parcelles de blocs différents, la probabilité que ces deux parcelles reçoivent le même traitement est de $1/t$.
(ii) La randomisation d'un plan en blocs complets à t traitements est fortement valide si, pour chaque couple de parcelles distinctes d'un même bloc, la probabilité que ces deux parcelles reçoivent un couple donné de traitements distincts est de $1/t(t-1)$; alors que, pour chaque couple de parcelles de blocs différents, la probabilité que ces deux parcelles reçoivent un couple donné de traitements est de $1/t^2$
(iii) Une randomisation faiblement valide d'un plan en blocs complets à t traitements peut être rendue fortement valide en randomisant totalement les traitements.

Quatrième partie IV

Plans factoriels

Plans factoriels complet

On considère maintenant des expériences avec plus de deux facteurs. Les facteurs peuvent être

- contrôlés (âge, sexe,traitement,...)
- d'homogénéisation (lots, centres, blocs...).

Définition

Un plan d'expérience factoriel est **complet** si on a au moins une observation pour chaque combinaison de niveaux des facteurs.

- Par exemple, un plan complet à 3 facteurs avec 3,4 et 5 niveaux aura au minimum $3 \times 4 \times 5 = 120$ observations.
- Dans un plan complet, **toutes les interactions sont estimables**.

Définition

Un plan d'expérience factoriel complet est **équilibré** si on a le même nombre d'observations pour chaque combinaison de niveaux des facteurs.

- Dans un plan complet équilibré, tous les effets (principaux et interactions) sont orthogonaux.
- Un PBCE est une plan complet équilibré pour les effets blocs et traitements.

Exemple de plan incomplets : les carrés latins

Problèmes :

- On a 3 facteurs ayant le même nombre de niveaux m.
- Un plan complet nécessiterait au moins m^3 expériences.
- On ne dispose que de m^2 expériences.
- On est obligé de sacrifier l'estimabilité de certains effets.
- On fait donc l'hypothèse qu'**il n'y a pas d'interaction** entre les effets et on veut que les effets principaux soient estimables entièrement.
- Sous cette hypothèse, on veut que les **effets principaux soient orthogonaux** (l'estimation des effets d'un facteur se font indépendamment de la présence ou non des autres facteurs : type 1=type 3).

Le **carré latin** répond à toutes ces contraintes.

Définition

Un carré latin est un tableau carré composé de nombre de 1 à m tel que chaque nombre apparaît une et une seule fois dans chaque ligne et chaque colonne. Les lignes correspondent aux niveaux du premier facteurs, les colonnes aux niveaux du deuxième facteur et les nombres aux niveaux du troisième facteur.

	facteur B				
facteur A	1	2	3	4	5
	2	3	4	5	1
	3	4	5	1	2
	4	5	1	2	3
	5	1	2	3	4

Randomisation d'un carré latin

Une randomisation valide d'un carré latin se fait de la façon suivante :

1. On permute aléatoirement les lignes du plan,
2. puis on permute aléatoirement les colonnes du plan.

Par exemple

$$\begin{array}{|ccccc|}\hline 1&2&3&4&5\\ 2&3&4&5&1\\ 3&4&5&1&2\\ 4&5&1&2&3\\ 5&1&2&3&4\\ \hline\end{array} \longrightarrow \begin{array}{|ccccc|}\hline 4&1&3&2&5\\ 1&3&5&4&2\\ 5&2&4&3&1\\ 3&5&2&1&4\\ 2&4&1&5&3\\ \hline\end{array}$$

Cinquième partie V

Plans en cross-over

Définition

On cherche à comparer t traitements. Un plan en cross-over est un dispositif expérimental divisé en p périodes. Pour chaque période, chaque individu reçoit un traitement, en général différent du précédent.

Périodes

`sujet i`	2	1	4	3	5

Pourquoi utiliser un plan en cross-over ?

- En général la principale source de variabilité vient de l'effet sujet. Dans un cross-over, le sujet devient son **propre témoin**. Si le plan est orthogonal, l'effet sujet est éliminé pour l'analyse des effets traitements.
- Mais, on observe souvent un effet de **carry-over** (ou effet **résiduel** ou effet **rémanent**) qui doit être pris en compte dans l'analyse,
- On n'utilisera donc un cross-over que s'il y a une forte variabilité inter-individu.

Un bon plan en cross-over est un plan qui doit :

1. éliminer l'effet période en équilibrant les traitements sur les périodes,
2. éliminer l'impact du carry-over en équilibrant les couples consécutifs (équilibre pour les voisinages),
3. éliminer l'effet sujet en utilisant un PBCE ou en minimiser l'impact en utilisant un PBIE.

Remarque

Les points 2 et 3 sont en fait contradictoires. On privilégiera 2 et on essaiera d'équilibrer pour les voisinages. Les recherches actuelles se portent sur des plans qui réalisent un compromis entre 2 et 3.

Plan en cross-over obtenu par permutations

4 traitements sur 4 périodes et $4! = 24$ sujets

1	2	3	4
1	2	4	3
1	3	2	4
1	3	4	2
1	4	2	3
1	4	3	2
2	1	3	4
2	1	4	3
2	3	1	4
2	3	4	1
2	4	1	3
2	4	3	1
3	1	2	4
3	1	4	2
3	2	1	4
3	2	4	1
3	4	1	2
3	4	2	1
4	1	2	3
4	1	3	2
4	2	1	3
4	2	3	1
4	3	1	2
4	3	2	1

Plan en cross-over obtenu par permutations

4 traitements sur 3 périodes et 4 !/1 !=24 sujets.

1	2	3
1	3	2
2	1	3
2	3	1
3	1	2
3	2	1
1	2	4
1	4	2
2	1	4
2	4	1
4	1	2
4	2	1
1	3	4
1	4	3
3	1	4
3	4	1
4	1	3
4	3	1
2	3	4
2	4	3
3	2	4
3	4	2
4	2	3
4	3	2

C'est un PBIE pour les sujets-traitements avec $r = 18$ et $\lambda = 12$.

Carrés latins complets

Définition

Un **carré latin semi-complet** est un carré latin tel que chaque traitement est précédé une et une seule fois par chaque autre traitement.

Exemple :

périodes

patients

1	2	3	4
3	1	4	2
2	4	1	3
4	3	2	1

- Ce plan est optimal si on incorpore un effet de carry-over dans le modèle.
- Ce plan minimise l'impact du biais du carry-over si celui-ci n'est pas inclus dans l'analyse.

Cas du cross-over AB/BA

Remarque

Les plans en cross-over à 2 traitements $\times$ *2 périodes de type*

$$\begin{matrix} A & B \\ B & A \end{matrix}$$

ne doivent pas être utilisés si on soupçonne un carry-over. Dans ce cas, il y a confusion entre les effets carry-over et traitement.

Bon appétit !