

PLAN EN CARRÉ LATIN (L2, L3, L6)

(08 / 01 / 2020, © Monfort, Dicostat2005, 2005-2020)

Un **plan en carré latin** (R.C. BOSE) est un **plan à deux dimensions** dont les blocs sont disposés selon deux dimensions :

(a) soit pour étudier l'effet d'un **facteur expérimental** F donné ;

(b) soit pour étudier l'effet de trois **facteurs** (un facteur F tq le précédent, un facteur « ligne » et un facteur « colonne »).

Ce **plan d'expérience** a notamment pour but l'élimination de l'influence due à l'**hétérogénéité** entre **unités expérimentales**.

(i) Dans un **plan en carré latin** (pcl), on dispose B blocs dans deux dimensions selon l lignes et l colonnes (avec $B = l^2$), on suppose que :

(a) chaque **bloc** comporte une seule **unité expérimentale** ($N_b = 1, \forall b \in N_B^*$) ;

(b) chaque **traitement** (ie niveau du facteur F) i figure une seule fois en ligne et une seule fois en colonne ;

(c) chacun de ces traitements est affecté au **hasard** entre les lignes et entre les colonnes (par application du **principe de randomisation**).

Par suite :

(1) $\text{card } \mathcal{T}_e = l$, en notant $\mathcal{T}_e = N_l^* = \{1, \dots, l\}$

(ensemble des niveaux du facteur F, ie ensemble des traitements).

La **disposition** précédente définit un **carré latin** (L. EULER) dont la forme générale (forme « standard ») est la suivante :

$$(2) \quad C = \begin{bmatrix} 1 & \dots & \dots & \dots & l \\ i_{\sigma_2}(1) & \dots & \dots & \dots & i_{\sigma_2}(l) \\ \dots & \dots & \dots & \dots & \dots \\ i_{\sigma_l}(1) & \dots & \dots & \dots & i_{\sigma_l}(l) \end{bmatrix},$$

où $\sigma_2, \dots, \sigma_l$ sont des éléments de σ_l (**groupe** des **permutations** de N_l^*) qui satisfont les conditions précédentes (avec $\sigma_1 = \text{id} (\sigma_1)$) et où $i_{\sigma_l(c)}$ indique qu'un traitement i est affecté à une unité dans la ligne l et la colonne c (en notant $\sigma_l(c)$ pour désigner $\sigma_l(c)$).

Un pcl est un **plan incomplet** car il ne met en oeuvre que l^2 combinaisons de niveaux (celles du facteur « ligne » et celles du facteur « colonne ») sur les l^3 possibles (en combinant avec le facteur F).

(ii) On peut lui associer un **modèle d'analyse de la variance** de forme additive suivant :

$$(3) \quad y_{lci} = \beta_0 + \beta_l^1 + \beta_c^2 + \beta_i^3 + u_{lci}, \quad \forall (l, c, i) \in A_l,$$

où $A_l \subset (N_l^*)^3$ est tq $\text{card } A_l = l^2$ (ce qu'exprime la **matrice du plan C**).

Le modèle (3) devient un **modèle identifiable** si l'on impose des contraintes d'**identification** tq les suivantes sur les paramètres (cf aussi **conditions d'identification**) :

$$(4) \quad \sum_{l=1}^l \beta_l^1 = \sum_{c=1}^l \beta_c^2 = \sum_{i=1}^l \beta_i^3 = 0,$$

l'indice l étant relatif aux lignes du carré latin C, l'indice c à ses colonnes et l'indice i aux niveaux (ou traitements) du facteur F.

(iii) Les **procédures statistiques** usuelles du **modèle linéaire** (avec **contrainte sur les paramètres**) s'appliquent alors (cf **singularité**).

En pratique, on peut constituer un carré latin par **permutation** cyclique successive sur les lignes (resp colonnes) de C, à partir d'une première ligne (resp colonne) définie par permutation au hasard sur l'ensemble N_l^* .

On emploie souvent, au lieu des notations de (2), l lettres « latines » A, B,..., L pour constituer C comme indiqué (d'où la dénomination).

(iv) On utilise généralement un carré latin lorsqu'on dispose de trois facteurs (notés F, $\mathcal{L}$, $\mathcal{C}$) ayant le même nombre de niveaux et susceptibles d'influer sur une grandeur η (**variable endogène** de (3) dont les **observations** sont les y_{lci}).

Pour éviter d'avoir à observer l^3 unités expérimentales (ce qui n'est pas toujours possible), la construction d'un carré latin permet, à l'aide de l^2 unités seulement, d'appliquer toutes les l^3 combinaisons possibles $(l, c, i) \in A_l$. C'est un plan incomplet, car il manque des unités pour certains triplets (l, c, i) , et équilibré, car l est le même pour chaque facteur.

On note parfois $y_{lc,n(l,c)}$ la valeur de la variable η correspondant à l'observation n_{lc} (ici notée par commodité $n(l,c)$) associée à la « case » (l, c) de C, l'indice n_{lc} dépendant, par construction, du couple (l, c) .

On écrit aussi parfois la **matrice des observations** Y des $(y_{lc,n(l,c)})$ selon :

$$(5) \quad Y = \begin{bmatrix} y_{111} & y_{122} & \cdots & y_{1l1} \\ y_{212} & y_{223} & \cdots & y_{2l1} \\ \cdots & \cdots & \cdots & \cdots \\ y_{l1l} & y_{l2l} & \cdots & y_{ll,l-1} \end{bmatrix}.$$

(v) L'usage des pcl est limité par les deux conditions suivantes :

- (a) chacun des « facteurs » possède le même nombre l de niveaux ;
- (b) le modèle associé (3) est additif (ie sans interaction).

Il existe des plans voisins ou dérivés des pcl, notamment le **plan en carré gréco-latin** et le **plan de YOUNG**.