



HAL
open science

Plans en blocs incomplets : application à la santé

Mamadou Koné, Annick Valibouze

► To cite this version:

Mamadou Koné, Annick Valibouze. Plans en blocs incomplets : application à la santé. Journées Internationales Analyse Statistique : Théorie et Applications, JIATA2012, Jun 2012, Oujda, Maroc. To appear. hal-00694912v1

HAL Id: hal-00694912

<https://hal.science/hal-00694912v1>

Submitted on 7 May 2012 (v1), last revised 30 May 2012 (v2)

HAL is a multi-disciplinary open access archive for the deposit and dissemination of scientific research documents, whether they are published or not. The documents may come from teaching and research institutions in France or abroad, or from public or private research centers.

L'archive ouverte pluridisciplinaire **HAL**, est destinée au dépôt et à la diffusion de documents scientifiques de niveau recherche, publiés ou non, émanant des établissements d'enseignement et de recherche français ou étrangers, des laboratoires publics ou privés.

Plans en blocs incomplets : application à la santé

¹ Mamadou KONÉ et ² Annick VALIBOUZE

L.S.T.A., Université Pierre et Marie Curie-Paris 6.

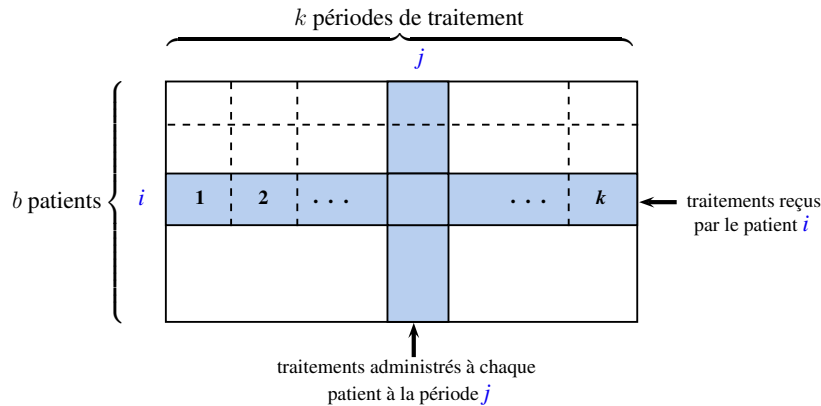
Tour 15-25, 4 place Jussieu, 75252 Paris Cedex 05, France.

Résumé. Le problème de recherche de plans optimaux pour une structure de corrélation générale a été évoqué par Kushner (1997). Le présent article donne des conditions d'optimalité universelle pour des plans en blocs incomplets lorsque les observations sur le même patient sont modélisées par un processus autorégressif $AR(m)$ d'ordre m arbitraire. Il étend et généralise les résultats existants pour les modèles $AR(1)$ et $AR(2)$ comme par exemple ceux de Grondona et Cressie (1993) pour les $AR(2)$ et ceux de Gill et Shukla (1985a) et Kunert (1997) pour les $AR(1)$.

Mots clés : Structures de corrélation, plans en blocs incomplets équilibrés, optimalité universelle, critère d'optimalité, tableaux semi-équilibrés.

1 Introduction

Nous considérons des situations expérimentales dans lesquelles $v \geq 1$ traitements sont administrés à $b \geq 1$ patients au cours de $k \geq 1$ périodes distinctes, le terme générique de période désigne, par convention, les temps, ou bien les instants, distincts où les traitements sont administrés. La figure ci-dessous illustre ce plan expérimental.



Le problème est de construire la meilleure structure expérimentale possible, l'efficacité étant mesurée par des critères d'optimalité évaluant la précision avec laquelle on estime la valeur moyenne des effets. La précision est mesurée par la matrice d'information des estimations. Nous notons $\Omega_{v,b,k}$ l'ensemble de toutes les structures possibles de ce type. Plus précisément, nous raisonnerons dans le cadre d'une structure expérimentale en *blocs incomplets* où un patient donné reçoit exactement k traitements distincts répartis dans les différentes périodes. Dans chacune de ces périodes, le patient reçoit un seul traitement parmi les v possibles, ce qui donne lieu à une mesure expérimentale scalaire unique.

Nous supposons que la corrélation entre des observations effectuées sur des patients distincts est nulle. Désignons par $\varepsilon_{i,j}$ le processus d'erreur obtenue à la $j^{\text{ième}}$ ($1 \leq j \leq k$) période sur le $i^{\text{ième}}$ ($1 \leq i \leq b$)

1. mamadou.kone@upmc.fr

2. annick.valibouze@upmc.fr

patient. Nous supposons que $\varepsilon_{i,j}$ est une réalisation partielle d'un processus autorégressif de $m^{\text{ième}}$ ordre désigné dans la suite par $\text{AR}(m)$ (AR pour "autoregressif"). Il est classiquement caractérisé par les relations

$$\varepsilon_{i,j} - \sum_{\ell=1}^m \theta_{\ell} \varepsilon_{i,j-\ell} = w_{i,j} ; \quad j = 0, \pm 1, \pm 2, \dots, \quad (1)$$

où les θ_{ℓ} désignent les paramètres du modèle et les $w_{i,j}$ sont des variables aléatoires indépendantes et identiquement distribuées, d'espérance nulle et de variance σ^2 . La fonction de covariance γ d'un processus $\text{AR}(m)$ satisfait l'équation de différence ci-dessous (voir par exemple Wei (1990))

$$\gamma(s) - \sum_{\ell=1}^m \theta_{\ell} \gamma(s-\ell) = \begin{cases} 0 & \text{pour } s > 0 \\ \sigma^2 & \text{pour } s = 0. \end{cases} \quad (2)$$

où, pour tout $i \in \{1, \dots, b\}$ fixé et pour tout $j \in \{1, \dots, k\}$

$$\gamma(s) = \text{Cov}(\varepsilon_{i,j}, \varepsilon_{i,j+s}) \quad (3)$$

Lorsque la structure de covariance est connue de manière exacte, il est naturel de construire les plans d'expériences optimaux en faisant usage de l'estimateur des Moindres Carrés Généralisé (MCG). Parmi d'autres auteurs, Kunert (1985, 1987), Azzalini et Giovagnoli (1987), Gill et Shukla (1985a,b), Martin et Eccleston (1991), Grondona et Cressie (1993), Satpati *et al.* (2007) ont utilisé cette approche, pour les structures de corrélation $\text{AR}(1)$ et $\text{AR}(2)$.

Dans le cas des modèles $\text{NN}m$ (NN pour "nearest-neighbor"), la structure de covariance offre la possibilité d'utiliser l'estimateur des Moindres Carrés Ordinaires (MCO) (voir Kiefer et Wynn (1981) pour le modèle $\text{NN}1$, Morgan et Chakravarti (1988) pour le modèle $\text{NN}2$ et Koné et Valibouze (2011) pour le modèle $\text{NN}m$ pour m arbitraire).

Dans les deux modèles $\text{NN}m$ et $\text{AR}(m)$, les plans considérés sont $\text{NN}m$ -équilibrés au sens de la définition suivante :

Définition 1.1 Pour m ($m \geq 1$) fixé, on appellera plan $\text{NN}m$ -équilibré (ou plan équilibré pour les voisinages à distance m) si pour tout entier $\ell \in \{1, \dots, m\}$, le nombre de fois où deux traitements distincts administrés au même patient sont voisins à distance ℓ est indépendant du choix de ces deux traitements ; c'est-à-dire c'est une valeur qui ne dépend que de ℓ et du plan.

Exemple 1 Le plan de la table 1 est $\text{NN}2$ -équilibré. En effet, ici $m = 2$ et chaque paire de traitement distincts apparaît exactement 3 fois à distance $\ell = 1$ et 2 fois à distance $\ell = 2$.

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

TABLE 1 – Plan $\text{NN}2$ -équilibré

Nous donnons des conditions d'optimalité universelle de plans pour le modèle $\text{AR}(m)$, $m \geq 1$, pour l'estimation des contrastes de traitements via l'estimateur MCG. Le paragraphe 2 de cet article est consacré à l'introduction du modèle décrivant la structure de corrélation. Dans le paragraphe 3, nous présentons un résultat connu relatif aux plans universellement optimaux. Pour terminer, dans le paragraphe 4 nous présentons des conditions d'optimalité universelle ainsi que la construction des plans satisfaisant ces conditions.

Références

- Azzalini, A. et Giovagnoli, A. (1987). Some optimal designs for repeated measurements with autoregressive errors. *Biometrika*, **74**(4), 725–734.
- Benchekroun, K. (1993). *Association-balanced arrays with applications to experimental design*. Ph.D. thesis, Dept. of Statistics, The University of North Carolina, Chapel Hill.
- Dehevels, P. et Derzko, G. (1991). Block designs for early-stage clinical trials. *Rapport technique LSTA*.
- Gill, P. S. et Shukla, G. K. (1985a). Efficiency of nearest neighbour balanced block designs for correlated observations. *Biometrika*, **72**(3), 539–544.
- Gill, P. S. et Shukla, G. K. (1985b). Experimental designs and their efficiencies for spatially correlated observations in two dimensions. *Comm. Statist. A—Theory Methods*, **14**(9), 2181–2197.
- Grondona, M. O. et Cressie, N. (1993). Efficiency of block designs under stationary second-order autoregressive errors. *Sankhyā Ser. A*, **55**(2), 267–284.
- Kiefer, J. (1975a). Balanced block designs and generalized Youden designs. I. Construction (patchwork). *Ann. Statist.*, **3**, 109–118.
- Kiefer, J. (1975b). Construction and optimality of generalized Youden designs. In *A survey of statistical design and linear models (Proc. Internat. Sympos., Colorado State Univ., Ft. Collins, Colo., 1973)*, pages 333–353. North-Holland, Amsterdam.
- Kiefer, J. et Wynn, H. P. (1981). Optimum balanced block and Latin square designs for correlated observations. *Ann. Statist.*, **9**(4), 737–757.
- Koné, M. et Valibouze, A. (2011). Plans en blocs incomplets pour la structure de corrélation nm . *Pub. Ins. Stat. Univ. Paris*, **55**(55), 65–88.
- Kunert, J. (1987). Neighbour balanced block designs for correlated errors. *Biometrika*, **74**(4), 717–724.
- Kunert, J. (1985). Optimal repeated measurements designs for correlated observations and analysis by weighted least squares. *Biometrika*, **72**(2), 375–389.
- Kushner, H. B. (1997). Optimal repeated measurements designs : the linear optimality equations. *Ann. Statist.*, **25**(6), 2328–2344.
- Martin, R. J. (1998). Optimal designs for small-sized blocks under dependence. *J. Combin. Inform. System Sci.*, **23**(1-4), 95–124. J. N. Srivastava : felicitation volume.
- Martin, R. J. et Eccleston, J. A. (1991). Optimal incomplete block designs for general dependence structures. *J. Statist. Plann. Inference*, **28**(1), 67–81.
- Morgan, J. P. et Chakravarti, I. M. (1988). Block designs for first and second order neighbor correlations. *Ann. Statist.*, **16**(3), 1206–1224.
- Mukhopadhyay, A. C. (1972). Construction of BIBD's from OA's combinatorial arrangements analogous to OA's. *Calcutta Statist. Assoc. Bull.*, **21**, 45–50.
- Passi, R. M. (1976). A weighting scheme for autoregressive time averages. *Journal of Applied Meteorology*, **15**, issue 2, 117–119.
- Rao, C. (1961). Combinatorial arrangements analogous to orthogonal arrays. *Sankhyā Ser A*, pages 283–286.

- Rao, C. R. (1973). Some combinatorial problems of arrays and applications to design of experiments. In *Survey of combinatorial theory (Proc. Internat. Sympos., Colorado State Univ., Ft. Collins, Colo., 1971)*, pages 349–359. North-Holland, Amsterdam.
- Satpati, S. K. et Parsad, R. (2004). Construction and cataloguing of nested partially balanced incomplete block designs. *Ars Combin.*, **73**, 299–309.
- Satpati, S. K., Parsad, R., et Gupta, V. K. (2007). Efficient block designs for dependent observations—a computer-aided search. *Comm. Statist. Theory Methods*, **36**(5-8), 1187–1223.
- Siddiqui, M. M. (1958). On the inversion of the sample covariance matrix in a stationary autoregressive process. *Ann. Math. Statist.*, **29**, 585–588.
- Wei, W. W. S. (1990). *Time series analysis*. Addison-Wesley Publishing Company Advanced Book Program, Redwood City, CA. Univariate and multivariate methods.
- Wise, J. (1955). The autocorrelation function and the spectral density function. *Biometrika*, **42**, 151–159.