hal-00589585, version 1 - 29 Apr 2011

Plans en blocs pour la structure de corrélation NNm

Mamadou KONE et 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. Les plans $NN1$ et $NN2$-optimaux ont été pleinement caractérisés par Kiefer et Wynn (1981), et Morgan et Chakravarti (1988). Nous étendons ces résultats pour des plans ayant la structure de corrélation NNm, correspondant à des valeurs de m supérieures ou égales à 3. Nous donnons des conditions d'optimalités pour le modèle NNm. Pour la construction des plans NNm-optimaux, nous utilisons les plans à voisinage équidistant, et les tableaux semi-équilibrés.

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

1 Introduction

Nous considérons des situations expérimentales dans lesquelles v traitements sont administrés à b patients dans k périodes. Le problème est de donner la meilleure structure d'expérience, vis-à-vis de critères d'optimalité pour la mesure des effets. Nous notons $\Omega_{v,b,k}$ la classe de toutes les structures possibles de ce type.

Soit $Y_{is\ell}$ la mesure expérimentale obtenue lorsque le $i^{\text{ième}}$ $(1 \leq i \leq v)$ traitement est appliqué au $s^{\text{ième}}$ $(1 \leq s \leq b)$ patient à la $\ell^{\text{ième}}$ $(1 \leq \ell \leq k)$ période. La structure de corrélation du voisin le plus proche du $m^{\text{ième}}$ ordre, NNm (NN pour "*nearest-neighbor*") considérée est, pour $1 \leq m \leq k-1$,

$$\mathbb{C}\text{ov}(Y_{is\ell}, Y_{i's'\ell'}) = \begin{cases} \sigma^2 \rho_{|\ell-\ell'|} & \text{si } s = s', |\ell - \ell'| \leq m \\ 0 & \text{sinon ,} \end{cases} \tag{1}$$

i.e., toutes les observations ont la même variance σ^2; la corrélation entre les observations effectuées sur différents patients est nulle ; deux observations réalisées sur le même patient sont d'autant plus corrélées qu'elles sont voisines plus voisine. Dans le modèle (1), ρ_1 désigne le coefficient de corrélation entre observations immédiatement successives, ρ_2 le coefficient de corrélation entre observations séparées par deux intervalles de temps, ..., et ρ_{k-1} le coefficient de corrélation entre observations extrêmes. les valeurs de ces coefficients sont supposées positives, et décroissantes de sorte que $\rho_0 = 1 \geq \rho_1 \geq \rho_2 \geq \rho_3 \geq \ldots \geq \rho_m > 0$, $\rho_{m+1} = \ldots = \rho_{k-1} = 0$. Pour déterminer les plans optimaux pour la structure de corrélation (1), deux approches sont proposées par Kiefer et Wynn (1981):

1. Considérez la classe Ω (définie ci-dessus) de tous les plans et identifions la classe $\Omega' \subset \Omega$ qui est optimale pour des erreurs non corrélées.

2. En utilisant l'estimateur des moindres carrés, déterminez la sous classe $\Omega^* \subset \Omega'$ qui est optimale pour la structure de corrélation (1).

Cette approche a été utilisée par Chêng (1983), Ipinyomi (1986), Kunert (1987), Russell et Eccleston (1987a), Morgan et Chakravarti (1988), Jacroux (1998).

Le travail fondamental de Kiefer et Wynn (1981) a eu pour objet l'élaboration d'un nouveau modèle, le modèle $NN1$. Ils ont défini des critères d'optimalité sur la matrice de variances-covariances aboutissant à la notion d'*optimalité universelle faible*, en utilisant l'estimateur des *moindres carrés ordinaire* (MCO). Or lorsque la structure de covariance est précisément connue, les plans réellement optimaux sont définis en utilisant l'estimateur des *moindres carrés généralisés* (MCG). Kiefer et Wynn ont justifié le choix de MCO en

E-mail address : mamadou.kone@upmc.fr, annick.valibouze@upmc.fr

hal-00589585, version 1 - 29 Apr 2011

montrant que la perte de précision résultant de l'utilisation de cet estimateur est assez faible dans le modèle de corrélation NN, Ils ont abouti à des conditions d'optimalité faible dans ce modèle, et ont construit des plans en *blocs incomplets équilibrés à voisinage équidistant* (voir le paragraphe 4), désignés par EBIBD (pour "*equineighboured balanced incomplete block design*") $NN1$-optimaux en utilisant des *carrés latins* et des *ensembles de différences*. Chêng (1983) a utilisé le même modèle que Kiefer et Wynn, il a introduit des méthodes de construction des plans en *blocs incomplets équilibrés*, désignés par BIBD (pour "*balanced incomplete block design*") (voir le paragraphe 2) basées sur la théorie des graphes ayant des blocs de taille 3 et $v-1$. D'autres méthodes de construction plus générales des EBIBD pour le cas $k=3$, et basées essentiellement sur la technique des *ensembles de différences*, sont données par Jacroux (1998). Dans le même contexte que Kiefer et Wynn, Morgan et Chakravarti (1988), ont défini la structure de covariance du voisin le plus proche du $m^{\text{ième}}$ ordre, et ont établi des conditions d'*optimalité universelle faible* des plans en blocs pour les modèles $NN1$ et $NN2$. On pourra se reporter à Cutler (1993), Uddin et Morgan (1997), Benchekroun et Chakravarti (1999), et Uddin (2008) pour une approche alternative des plans en blocs pour des observations corrélées, utilisant l'estimateur des MCG et une structure de corrélation différente.

Suivant la nomenclature de Morgan et Chakravarti (1988), nous qualifions de NNm-équilibrés les plans d'expérience en blocs incomplets, lorsque ceux-ci sont équilibrés pour les périodes distantes dans le temps de m unités, ou moins. La procédure d'équilibrage temporelle permet d'éliminer des résultats d'expérience les biais résultant d'effets d'interaction dus à la proximité de l'administration de certains traitements au même patient. Une condition minimale pour qu'un plan en blocs incomplets soit utilisable en présence d'effets, dus à l'ordre de l'administration des traitements, est que chaque traitement soit appliqué le même nombre de fois dans chaque période. Cette condition impose que le nombre v de traitements divise le nombre b de patients. Agrawal (1966b,a) a énoncé le résultat suivant : pour $d \in \Omega_{v,b,k}$ avec $v|b$, les traitements peuvent être réarrangés pour chaque patient pour obtenir un plan équilibré dans les périodes. Pour $2 \leq v \leq 7$ et $2 \leq k \leq 5$, Deheuvels et Dzerko (1991) ont donné une liste de plans $d \in \Omega^*_{v,b,k}$ qui sont $NN1$, $NN2$-équilibrés.

Le paragraphe 2 de cet article est consacrée à l'introduction du modèle des plans en blocs pour décrire les résultats d'expériences associés à un plan. Dans le paragraphe 3 nous présentons d'abord un résultat fondamental connu relatif aux plans *universellement faiblement optimaux*, et ensuite nous étendons les résultats relatifs au plans $NN1$, $NN2$-optimaux pour des plans NNm correspondant à des valeurs de m supérieures ou égales à 3. Les paragraphes 4 et 5 présentent respectivement quelques exemples de constructions des plans NNm-optimaux et les démonstrations des résultats énoncés dans cet article.

2 Description du modèle

Supposons que la mesure expérimentale obtenue lorsque le $i^{\text{ième}}$ traitement est appliqué au $s^{\text{ième}}$ patient à la $\ell^{\text{ième}}$ période soit de la forme

$$Y_{is\ell} = \mu + \eta_i + \beta_s + \varepsilon_{is\ell}, \tag{2}$$

où μ désigne l'effet moyen des traitements sur l'ensemble des patients, β_s est l'effet relatif particulier au $s^{\text{ième}}$ patient, η_i est l'effet relatif particulier au traitement i, et où les résidus $\{\varepsilon_{is\ell}\}$ vérifient la structure de corrélation définie dans (1) ci-dessus. On impose à ce modèle les contraintes classiques d'identification,

$$\sum_{i=1}^{v} \eta_i = \sum_{s=1}^{b} \beta_s = 0. \tag{3}$$

Un plan $d \in \Omega_{v,b,k}$ sera défini comme une fonction

$$\begin{aligned} d : [\![1,b]\!] \times [\![1,k]\!] &\to [\![1,v]\!] \\ (s,\ell) &\mapsto d(s,\ell), \end{aligned}$$

hal-00589585, version 1 - 29 Apr 2011

où $d(s,\ell)$ désigne le traitement appliqué au $s^{\text{ième}}$ $(1 \leq s \leq b)$ patient dans la $\ell^{\text{ième}}$ $(1 \leq \ell \leq k)$ période. Ce plan d sera représenté par un tableau $b \times k$ d'éléments pris parmi l'ensemble des v traitements, où les lignes représentent les patients, et les colonnes les périodes. Pour un plan $d \in \Omega_{v,b,k}$, soit $n_{d,i,s}$ le nombre de fois que le traitement i est appliqué au $s^{\text{ième}}$ patient, $r_{d,i}$ le nombre de fois que le traitement i est répété dans la totalité de l'expérience, et $k_{d,s}$ le nombre total de traitements reçu par le $s^{\text{ième}}$ patient. Si $n_{d,i,s} = 0$ ou 1, $r_{d,i} = r_{d,i'} = r$, $k_{d,s} = k_{d,s'} = k$, le plan d est dit respectivement *binaire*, *équirépliqué*, et *propre*. Nous considérons seulement les plans en *blocs binaires propres équirépliqués*, désignés par PBERD (pour "*proper binary equireplicated block designs*"), où le $s^{\text{ième}}$ patient reçoit k traitements distincts $d(s,1),\ldots,d(s,k)$ sur les périodes $1,\ldots,k$, chacun repliqué r fois. En notation matricielle, le modèle (2) peut être écrite et précisée sous la forme,

$$\mathbf{Y}_d = \mu \mathbf{1}_{bk} + A_d \eta + (\mathbf{I}_b \otimes \mathbf{1}_k)\beta + \varepsilon, \text{ avec } \mathbb{E}(\varepsilon) = 0 \text{ et } \mathrm{Var}(\varepsilon) = \mathbf{I}_b \otimes \Delta, \tag{4}$$

où $\mathbf{Y}_d = (Y_{d(1,1)11},\ldots,Y_{d(1,k)11},\ldots,Y_{d(b,1)b1},\ldots,Y_{d(b,k)bk})'$ est le vecteur des observations, $\mathbf{1}_{bk}$ est le vecteur $bk \times 1$ colonne dont toutes les composantes sont égales à 1, $\mathbf{I}_b$ est la matrice $b \times b$ identité, $\otimes$ dénote le produit de Kronecker, $\eta = (\eta_{d(1,1)},\ldots,\eta_{d(1,k)},\ldots,\eta_{d(b,1)},\ldots,\eta_{d(b,k)})'$ est le vecteur des effets traitements, la matrice $A_d = [T_1',\ldots,T_b']'$ est déterminée par le plan d lorsque le $(\ell,i)^{\text{ième}}$ élément de $T_s = (\mathbf{t}_{\ell,i}(s))_{(1\leq\ell\leq k,\ 1\leq i\leq v)}$,

$$\mathbf{t}_{\ell,i}(s) = \begin{cases} 1 & \text{si } d(s,\ell) = i, \\ 0 & \text{sinon,} \end{cases} \tag{5}$$

et, représente la matrice $bk \times v$ d'incidence période-traitement, $\beta = (\beta_1,\ldots,\beta_b)'$ est le vecteur des effets patients, $(\mathbf{I}_b \otimes \mathbf{1}_k)$ est la matrice $bk \times b$ d'incidence période-patient; son $(s,m)^{\text{ième}}$ élément

$$\mathbf{p}_{s,m} = \begin{cases} 1 & \text{si le } m^{\text{ième}} \text{ coordonnée de } \mathbf{Y}_d \text{ est une observation sur le } s^{\text{ième}} \text{ patient,} \\ 0 & \text{sinon,} \end{cases}$$

ε est le vecteur $bk \times 1$ des erreurs aléatoires correlées avec des espérances nulles et matrice de covariances $\mathbf{I}_b \times \Delta$, $\Delta = (\sigma^2 \rho_{\ell,\ell'})$ avec $\rho_{\ell,\ell} = 1$ est la matrice $k \times k$ de variances-covariances des observations sur le même patient (voir l'identité (1)). Nous dénotons par $r(i,s)$ la temporalité à laquelle le traitement i a été administré au patient s, $r(i,s) = \ell$ si et seulement si $d(s,\ell) = i$

2.1 Plans en blocs incomplets équilibrés (BIBD)

Les BIBD sont introduits et analysés par Yates dans les années (1936-1940), ils jouent un rôle spécial parmi la classe des PBERD. Soit $\lambda_{d,i,i'}$ le nombre (cumulé) de fois où les traitements i et i' sont appliqués au même patient. Les BIBD sont tels que $\lambda_{d,i,i'} = \lambda$, ils sont notés par BIBD(v,b,r,k,λ). Des conditions nécessaires pour l'existence d'un tel plan dans $\Omega_{v,b,k}$ sont,

$$\begin{aligned} vr &= kb \\ \lambda(v-1) &= r(k-1) \end{aligned} \tag{6}$$

Les plans en *blocs complets*, désignés par CBD (pour "*complete block design*") sont les plans tels que $k = v$ et donc $r = b = \lambda$. Les conditions nécessaires (6) ne sont suffisantes seulement que dans certains cas, elles le sont trivialement pour $k = 2$. Bose (1939) a montré qu'elles ſaient suffisantes pour $k = 3$ et $\lambda = 2$. Hanani (1961) a amélioré ces résultats en établissant que ces conditions étaient également suffisantes pour $k = 3$ ou $k = 4$ pour tout λ, et lorsque $k = 5$ avec $\lambda = 1, 4, 20$. Une tabulation quasiment exhaustive des BIBD connus pour $r \leq 41$ et $k \leq v/2$ donnant l'existence, la non existence ou connaissance d'un BIBD est donnée dans Mathon et Rosa (1996). Une liste de BIBD pour $v \leq 25$ et $k \leq 11$ peut être trouvée dans l'ouvrage de Hinkelmann et Kempthorne (2005), pp. 115-118. Nous renvoyons à l'ouvrage de Giesbrecht et Gumpertz (2004), pp. 235-240, pour un catalogue de BIBD avec $4 \leq v \leq 100$ et $r \leq 15$. Parmi les

hal-00589585, version 1 - 29 Apr 2011

combinaisons pour lesquelles aucune solution n'existe (du moins, au mieux des connaissances actuelles), on a par exemple, $v = 15, b = 21, k = 5, r = 7$ et $\lambda = 2$. Soit $C_d = r\mathbf{I}_v - k^{-1}\mathbf{\Lambda}$, où $\mathbf{I}_v$ est la matrice identité, et $\mathbf{\Lambda} = (\lambda_{d,i,i'})$, la matrice d'information d'un BIBD d. Un *contraste* (traitement) est une combinaison linéaire $c'\eta = \sum_{i=1}^{v} c_i\eta_i$ des effets traitements telle que $\sum_{i=1}^{v} c_i = 0$. Un *contrast élémentaire* est tout contraste de la forme $\eta_i - \eta_j = c'\eta$. Un *contraste* est dit *estimable* s'il possède un estimateur linéaire sans biais. Un PBERD dans lequel tout les contrastes traitements sont *estimables* est appelé *connexe*. Nous supposons que tout les plans compétiteurs vérifient cette propriété de connexité, c'est à dire le rang $\mathrm{rang}(\mathbf{C}_d) = v - 1$ (Chakrabarti (1962), John (1980), pp. 9-13, Rasch et Herrendörfer (1986), pp. 39-40, et Dey (1986)). Nous restreignons $\Omega_{v,b,k}$ à la classe de tout les PBERD connectés.

2.2 Estimation des effets traitements

Dans le modèle (4), si $\widehat{\eta}$ est l'estimateur des MCO du vecteur traitement η, alors les équations normales réduites de $\widehat{\eta}$ sont données par,

$$\left(A_d'A_d - \frac{1}{k}A_d'(\mathbf{I}_b \otimes \mathbf{1}_k)(\mathbf{I}_b \otimes \mathbf{1}_k)'A_d\right)\widehat{\eta} = \left(A_d' - \frac{1}{k}A_d'(\mathbf{I}_b \otimes \mathbf{1}_k)(\mathbf{I}_b \otimes \mathbf{1}_k)'\right)Y_d. \tag{7}$$

Pour tout les détails nous renvoyons au chapitre 1 de l'ouvrage de Hinkelmann et Kempthorne (2005). Le terme à droite de l'égalité dans l'équation (7) est la somme des traitements ajustés, il est noté Q_d dans la littérature. L'autre terme (sans $\widehat{\eta}$) représente la matrice d'information (noté C_d) de η lorsqu'on est dans le cas d'observations non corrélées : $\Delta = \sigma^2\mathbf{I}_{bk}$. Dans ce cas, il est bien connu que cette matrice est définie positive et que la somme de ses éléments diagonaux et extra-diagonaux est nulle pour tout $d \in \Omega_{v,b,k}$.
Le système (7) n'est pas résoluble directement en $\widehat{\eta}$, du fait que la matrice C_d n'est pas inversible. Le lemme 2.1 de Shah (1960), permet de pallier cette difficulté, nous ne rentrons pas ici dans les détails.

Soit $\hat{\gamma}$ le meilleur estimateur (MCO) linéaire sans biais du vecteur $\gamma' = (\gamma_1, \ldots, \gamma_v)$ des effets traitements corrigés, où $\gamma_i = \eta_i - \frac{1}{v}\sum_{j=1}^{v}\eta_j$, pour $i = 1, \ldots, v$. Les résultats d'optimalité d'un plan $d \in \Omega_{v,b,k}$ seront basés sur l'ensemble des contrastes $\gamma = (\mathbf{I}_v - \frac{\mathbf{E}_v}{v})\eta$, ce choix est raisonnable puisque $c'\gamma = c'\eta$ pour tout vecteur contraste c. Pour les deux type de plans considérés, les estimateurs sont donnés par (voir par exemple Tinsson (2010)),

(a) si le plan $d \in \Omega_{v,b,k}$ est un $\mathrm{BIBD}(v, b, r, k, \lambda)$,

$$\widehat{\gamma} \quad = \quad k(\lambda v)^{-1}\left(A_d' - \frac{1}{k}A_d'(\mathbf{I}_b \otimes \mathbf{1}_k)(\mathbf{I}_b \otimes \mathbf{1}_k)'\right)Y_d. \tag{8}$$

(b) si le plan $d \in \Omega_{v,b,k}$ est un $\mathrm{CBD}(v, b, r, k, \lambda)$,

$$\widehat{\gamma} \quad = \quad b\left(A_d' - \frac{1}{k}A_d'(\mathbf{I}_b \otimes \mathbf{1}_k)(\mathbf{I}_b \otimes \mathbf{1}_k)'\right)Y_d. \tag{9}$$

3 Optimalité universelle faible des plans NNm

L'étude de la matrice d'information fournit des critères d'optimalité forte, mais lorsque la forme de la matrice d'information est trop complexe pour être étudiée, on est amené à définir des critères sur la matrice de variance aboutissant à la notion d'*optimalité universelle faible* de Kiefer et Wynn (1981).

3.1 Résultats préliminaires

Notons $\mathcal{V} = \{D_d : d \in \Omega_{v,b,k}\}$ l'ensemble des matrices de variances-covariance pour les estimations d'un ensemble de contrastes. Une telle matrice de variances-covariance est sémi-définie positive de dimension $v \times v$.

hal-00589585, version 1 - 29 Apr 2011

Définition 3.1 Une fonction $\Phi : \mathcal{V} \mapsto]-\infty, +\infty]$ est appelé critère. Nous nous intéressons aux critères Φ satisfaisant les conditions (voir Kiefer et Wynn (1981)) suivantes.

(i) Pour tout $D \in \mathcal{V}$, $\Phi(D)$ est invariant par toutes permutations appliquées aux lignes et aux colonnes de D;

(ii) Φ est convexe, i.e., $\Phi\{aD_1+(1-a)D_2\} \leq a\Phi(D_1)+(1-a)\Phi(D_2)$ pour tout $D_1, D_2 \in \mathcal{V}$ et $0 \leq a \leq 1$;

(iii) $\Phi(aD) \leq \Phi(D)\ \forall\ D \in \mathcal{V}$ dès que $0 < a < 1$.

Un plan d appartenant à $\Omega_{v,b,k}$ est dit *faiblement universellement optimal* si sa matrice de variance D_d minimise simultanément tous les critères Φ satisfaisant les conditions ((i)-(iii)) ci-dessus.

La proposition suivante fournit une approche simple pour examiner l'*optimalité universelle faible* d'un plan parmi une classe donnée.

Proposition 3.2 *(Kiefer et Wynn (1981)) Supposons que $\boldsymbol{I}'_v D_d = 0'_v$ pour $d \in \Omega_{v,b,k}$, et qu'il existe $d \in \Omega^*_{v,b,k}$ tel que : (i) Sa matrice de variance D_d est complètement symétrique, i.e, $D_d = \alpha \boldsymbol{I}_v + \beta \boldsymbol{E}_v$, où α et β sont des scalaires, $\boldsymbol{I}_v$ est la matrice $v \times v$ identité et $\boldsymbol{E}_v$ est la matrice $v \times v$ composée de 1 partout, (ii) La trace de D_d est minimale dans l'ensemble $\{D_d, d \in \Omega^*_{v,b,k}\}$, alors d est universellement faiblement optimal dans $\Omega^*_{v,b,k}$.*

Preuve Voir Kiefer (1975), et Kiefer et Wynn (1981).

Dans ce travail, les plans optimaux dans la classe Ω' considérés sont les CBD et les BIBD (voir Kiefer et Wynn (1981)). Le plan $d \in \Omega_{v,b,k}$ sera selectionné dans la sous classe $\Omega^*_{v,b,k}$ des BIBD ou des CBD.

3.2 Plans NNm-optimaux

Notations générales Nous généralisons les notations de Morgan et Chakravarti (1988) appliquées à $m = 1, 2$. Nous fixons m, l'ordre de voisinage, qui est aussi un intervalle de temps entre les observations.

Pour un plan $d \in \Omega_{v,b,k,}$, soit $A_i = \#\{s : n_{d,i,s} = 1\}$, le nombre de patients recevant le traitement i. Posons $\forall\ \ell \in [\![1, m]\!]$,

$\phi^{\ell}_{d,i} = \#\{s : r(i,s) \in \{\ell, k-\ell+1\}\}$ le nombre de patients pour lesquels le traitement i est appliqué à la $\ell^{\text{ième}}$ ou $(k-\ell+1)^{\text{ième}}$ période;

$\phi^{\ell}_{d,i,i'} = \#\{s : s \in A_i \cap A_{i'}, r(i,s) \in \{\ell, k-\ell+1\}\} + \#\{s : s \in A_i \cap A_{i'}, r(i',s) \in \{\ell, k-\ell+1\}\}$, le nombre de patients recevant les traitements i et i' pour lesquels i ou i' est appliqué à $\ell^{\text{ième}}$ ou $(k-\ell+1)^{\text{ième}}$ période, où un patient est compté deux fois si i et i' sont à la $\ell^{\text{ième}}$ et $(k-\ell+1)^{\text{ième}}$ période;

$N^{\ell}_{d,i,i'} = \#\{s \in A_i \cap A'_i : |r(i,s) - r(i',s)| \leq \ell\}$, le nombre de patients recevant les traitements i et i' pour lesquels i et i' sont appliqués à ℓ intervalles de temps l'un de l'autre avec $N^{\ell}_{d,i,i} = 0$.

Définissons,

$$(\phi^{\ell}_d)^T = (\phi^{\ell}_{d,1}, \phi^{\ell}_{d,2}, \ldots, \phi^{\ell}_{d,v});\ \text{et, pour } i \neq i', \Phi^{\ell} = (\phi^{\ell}_{d,i,i'});\ \ N^{\ell}_d = (N^{\ell}_{d,i,i'}).$$

Les correspondances avec les notations e_i, f_i, $e_{ii'}$ et $f_{ii'}$ de Morgan et Chakravarti (1988) sont les suivantes :

$$\phi^1_{d,i} = e_i,\ \ \phi^2_{d,i} = f_i,\ \ \phi^1_{d,i,i'} = e_{ii'} \text{ et } \phi^2_{d,i,i'} = f_{ii'}.$$

D'autres notations sont nécessaires afin de donner la matrice de variances-covariances de l'estimateur $\hat{\gamma}$ des effets traitements corrigés défini dans (8) et (9). Posons $\Delta^* = \mathbf{I}_b \otimes \Delta$, où Δ est défini dans le paragraphe 2. Notons $D_d(\Delta^*)$, la matrice de variances-covariances de $\hat{\gamma}$. Et enfin posons $\mathbf{E}_{p\times q} = \mathbf{1}_p \mathbf{1}'_q$ la matrice $p \times q$ composée de 1 partout avec, en particulier $\mathbf{E}_p = \mathbf{E}_{p\times p}$ et $\mathbf{I}_p = \mathbf{I}_{p\times p}$.

Les lemmes suivants seront utilisés pour la démonstration des théorèmes 3.8 et 3.9.

hal-00589585, version 1 - 29 Apr 2011

Lemme 3.3 *D'après les notations précédentes, nous avons pour m fixé*

$$\begin{aligned}
(i)\ \sum_{i'\neq i} N^{\ell}_{d,i,i'} &= 2r - \sum_{p=1}^{\ell} \phi^{p}_{d,i}\ ,\ \forall\ \ell \in [\![1,m]\!] \\
(ii)\ \sum_{i'\neq i} \phi^{p}_{d,i,i'} &= (k-2)\phi^{p}_{d,i} + 2r,\ \forall\ p \in [\![1,m]\!] \\
(iii)\ \sum_{i=1}^{v} \phi^{1}_{d,i} &= 2b\ \textit{et}\ ,\ \textit{pour}\ \ k>3,\ \sum_{i=1}^{v} \phi^{p}_{d,i} = 2b,\ \forall\ p \in [\![2,m]\!].
\end{aligned}$$

Pour $k=3$, il est facile de vérifier le lemme suivant.

Lemme 3.4 *Pour $k=3$, nous avons les identités suivantes,*

$$\begin{aligned}
(i)\ N^{\ell}_{d,i,i'} &= 0\ \forall\ \ell \in [\![3,m]\!], \\
(ii)\ N^{1}_{d,i,i'} + N^{2}_{d,i,i'} &= \lambda, \\
(iii)\ \phi^{2}_{d,i,i'} &= N^{1}_{d,i,i'}, \\
(iv)\ \phi^{\ell}_{d,i,i'} &= 0\ \forall\ \ell \in [\![3,m]\!], \\
(v)\ N^{1}_{d,i,i'} + \phi^{1}_{d,i,i'} &= 2\lambda.
\end{aligned}$$

Lemme 3.5 *Dans la structure de covariance (1),*

$$D_d(\Delta^*) = (\lambda v)^{-2} k^2 \sum_{s=1}^{b} T_s' \mathbb{W}(\Delta) T_s, \tag{10}$$

où $\mathbb{W}(\Delta) = (\boldsymbol{I}_k - k^{-1}\boldsymbol{E}_k)\Delta(\boldsymbol{I}_k - k^{-1}\boldsymbol{E}_k)$. En particulier les éléments extra-diagonaux de $D_d(\Delta^*)$ sont

$$D_{d,i,i'}(\Delta^*) = (\lambda v)^{-2} k^2 \sum_{s\in A_i \cap A_{i'}} \mathbb{W}_{r(i,s),r(i',s)}, \tag{11}$$

où $\mathbb{W}_{r(i,s),r(i',s)}$ est le $(r(i,s),r(i',s))^{\text{ième}}$ élément de $\mathbb{W}(\Delta)$.

Lemme 3.6 *Pour $k \geq 2m$, si $d \in \Omega_{v,b,k}$ est un $BIBD$ pour le modèle NNm, alors, nous avons*

$$\begin{aligned}
\sigma^{-2} D_{d,i,i}(\Delta^*) &= (\lambda v)^{-2}\Big\{ r[k(k-1) - 2(k+1)\rho_1 - 2(k+2)\rho_2 - \cdots - 2(k+m)\rho_m] \\
&\quad + 2k \sum_{\ell=1}^{m} \rho_\ell \sum_{p=1}^{\ell} \phi^{p}_{d,i} \Big\},
\end{aligned}$$

$$\begin{aligned}
\sigma^{-2} D_{d,i,i'}(\Delta^*) &= (\lambda v)^{-2}\Big\{ -\lambda[k + 2(k+1)\rho_1 + 2(k+2)\rho_2 + \cdots + 2(k+m)\rho_m] \\
&\quad + k \sum_{\ell=1}^{m} \rho_\ell \left(\sum_{p=1}^{\ell} \phi^{p}_{d,i,i'} + k N^{\ell}_{d,i,i'} \right) \Big\}.
\end{aligned}$$

Lemme 3.7 *Pour $k=3$, si $d \in \Omega_{v,b,k}$ est un BIBD pour le modèle NNm, alors nous avons*

$$\sigma^{-2} D_{d,i,i}(\Delta^*) = 2(\lambda v)^{-2}\{r(3 - 4\rho_1 + \rho_2) + 3e_{d,i}(\rho_1 - \rho_2)\}$$

$$\sigma^{-2} D_{d,i,i'}(\Delta^*) = (\lambda v)^{-2}\{6(\rho_1 - \rho_2) N^{1}_{d,i,i'} - \lambda(3 + 2\rho_1 - 5\rho_2)\}.$$

Nous pouvons maintenant énoncer le résultat d'optimalité suivant. Soit

$$F(i,i') = \sum_{\ell=1}^{m} \rho_\ell \left(\sum_{p=1}^{\ell} \phi^{p}_{d,i,i'} + k N^{\ell}_{d,i,i'} \right) \tag{12}$$

hal-00589585, version 1 - 29 Apr 2011

Théorème 3.8 *Pour $k \geq 2m$, un $BIBD(v,b,r,k,\lambda)$ est universellement faiblement optimal dans $\Omega^*_{v,b,k}$ pour le modèle NNm, si et seulement si les quantités respectives facteurs des ρ_ℓ dans $F(i,i')$, sont, chacune, indépendantes de i, i' $(1 \leq i \neq i' \leq v)$. Pour $k = 3$, cette condition est équivalente à l'egalité de tout les $N^1_{d,i,i'} (1 \leq i \neq i' \leq v)$.*

Remarque 1 *Les théorèmes 5.1 et 2.1 de Morgan et Chakravarti (1988), et Kiefer et Wynn (1981) respectivement sont des corollaires de ce théorème lorsqu'on les applique au cas $m = 1, 2$.*

Dans le cas des plans en blocs complets ($k = v$ et $\lambda = b$), nous obtenons le résultat suivant.

Théorème 3.9 *Soit un $CBD(v,b,r,k,\lambda)$ pour le modèle NNm, sa matrice de variances-covariances $D_d(\Delta^*)$ est complètement symétrique si et seulement si les quantités $N^\ell_{d,i,i'}$ sont chacune indépendantes de i, i', $1 \leq i \neq i' \leq v$.*

Remarque 2 *Le théorème 2.11 de Morgan (1983) est un corollaire de ce théorème lorsqu'on l'applique au cas $m = 2$.*

Corollaire 3.10 *Pour $k \geq 2m$, un $CBD(v,b,r,k,\lambda)$ est universellement faiblement optimal dans $\Omega^*_{v,b,k}$ pour le modèle NNm si les quantités $N^\ell_{d,i,i'}$ sont chacune indépendantes de i, i', $1 \leq i \neq i' \leq v$.*

Remarque 3 *Pour les cas $m = 1, 2$ nous retrouvons respectivement les théorèmes 4.1 et 2.4 de Kiefer et Wynn (1981), et Morgan et Chakravarti (1988) .*

Pour montrer les théorèmes 3.8 et 3.9, nous utilisons la proposition 3.2 de Kiefer et Wynn (1981).

Conséquence des lemmes 3.6 et 3.7. Soit $d \in \Omega^*_{v,b,k}$. D'après les lemmes 3.6, la trace $tr(D_d)$ est indépendante du choix d'un BIBD. En effet, d'après le lemme 3.3(iii), $\sum_{i=1}^v \phi^1_{d,i} = 2b$ pour $k > 3$ et $\ell \in [\![2, m]\!]$. Nous avons

$$\begin{aligned} tr(D_d) = \sum_{i=1}^{v} D_{d,i,i}(\Delta) \quad &= \quad \sigma^2(\lambda v)^{-2}\{kb[k(k-1) - 2(k+1)\rho_1 - 2(k+2)\rho_2 - 2(k+3)\rho_3 \\ &\quad - \cdots - 2(k+m)\rho_m] + 4kb(\rho_1 + \rho_2 + \rho_3 + \cdots + \rho_m) \\ &\quad + \ 4kb(\rho_2 + \rho_3 + \cdots + \rho_m) + 4kb(\rho_3 + \cdots + \rho_m) \\ &\quad + \ \cdots + \ 4kb\rho_m\}, \end{aligned} \tag{13}$$

et d'après le lemme 3.7,

$$tr(D_d) = 2(\lambda v)^{-2}\{r(3 - 4\rho_1 + \rho_2)v + 6b(\rho_1 - \rho_2)\} \text{ puisque } \sum_{i=1}^{v} \phi^\ell_{d,i} = 2b.$$

Ainsi tous les plans compétiteurs possèdent la même trace. Donc, nous allons identifier les plans dans $\Omega^*_{v,b,k}$ qui ont une matrice de variances-covariances complètement symétrique, i.e, les plans qui sont universellement faiblement optimaux dans $\Omega^*_{v,b,k}$.

3.3 Existence des plans NNm-optimaux

Dans ce paragraphe, nous cherchons à généraliser l'étude des plans minimaux $NN1$-optimaux de Kiefer et Wynn (1981), et $NN2$-optimaux de Morgan et Chakravarti (1988) aux plans NNm-optimaux. Ici, plans minimaux signifient qu'ils correspondent à la valeur minimale de b pour k et v donnés. Nous supposerons dans la suite que m est un entier fixé supérieur ou égale à 1.

On cherche une condition nécessaire pour que $\forall\ m, \forall\ \ell \in [\![1, m]\!]$, chacune des quantités facteur de ρ_ℓ des $F(i,i')$ soit une constante indépendante de i, i' $(i \neq i')$.

hal-00589585, version 1 - 29 Apr 2011

Proposition 3.11 *Soit* $m > 1$. *Un* $BIBD(v, b, r, k, \lambda)$ *universellement faiblement optimal dans* $\Omega^*_{v,b,k}$ *pour le modèle* NNm *satisfait*

$$k(k-1)|4\lambda. \text{ Si, de plus, } k \not\equiv 0 \ (\mod 4) \text{ alors}$$

$$k(k-1)|2\lambda. \tag{14}$$

Remarque 4 *Le théorème 2.1 de Morgan et Chakravarti (1988) est un corollaire de cette proposition lorsqu'on l'applique au cas* $m = 2$.

Corollaire 3.12 *Soit* $m > 1$. *Le nombre minimum de patients* b *pour lequel il existe un* $CBD(v, b, r, k, \lambda)$ *universellement faiblement optimal dans* $\Omega^*_{v,b,k}$ *pour le modèle* NNm *est*

$$b = \frac{v(v-1)}{2}. \tag{15}$$

Remarque 5 *Au départ de cette étude nous cherchions à généraliser les formules (14) et (15) établies dans le cas* $m = 2$ *sous les formes*

$$\prod_{\ell=1}^{m}(k-\ell+1)|m!\lambda, \quad et \quad b = \frac{1}{m!}\prod_{\ell=1}^{m}(v-\ell+1).$$

La proposition 3.11 et son corollaire nous renseignent de l'invariance de ces formules $\forall\ m > 1$.

4 Construction de plans NNm-optimaux

Dans ce paragraphe nous donnons quelques exemples de constructions des plans NNm-équilibrés.

Ipinyomi (1986) a introduit le concept des plans à voisinage équidistant, ED, (ED pour "*equi-neighboured design*"). Il a considéré la matrice de variances-covariances, $\Delta = \sum_{\ell=0}^{k-1} \rho_\ell U_{k,\ell}$, où $(U_{k,\ell})_{i,j} = 1$ si $|i-j| = \ell$, et 0 sinon, avec $U_{k,0} = I_k$ et $\rho_0 = 1$. Cela équivaut à l'hypothèse de la structure de corrélation du voisin le plus proche du $(k-1)^{\text{ième}}$ ordre $NN(k-1)$ dans la structure de corrélation donnée dans l'identité (1). Les ED sont tels que $N^{\ell}_{d,i,i'}$ $(1 \leq \ell \leq k-1)$ est indépendant de $1 \leq i \neq i' \leq v$. Par le théorème 3.9, un ED est un CBD NNm-optimal. Ipinyomi a construit les ED avec le nombre minimum de patients $b = mv$ pour $v = 2m+1$ et premier. Il a aussi construit les ED dans lesquels $b = v(v-1)$ en utilisant des carrés latins mutuellement orthogonaux $v \times v$ et, a suggéré que si un ensemble de carré latin n'existe pas, le concept d'un *ensemble de différence* peut être utilisé en général pour la construction des ED. Notons que ces ED sont des BIBD. Ainsi, les ED sont plus restrictifs que les EBIBD de Kiefer et Wynn (1981), Chêng (1983) et Jacroux (1998), qui sont des plans tels que $N^1_{d,i,i'}$ est indépendant de $1 \leq i \neq i' \leq v$.

Rao (1946, 1947) a introduit le concept des tableaux orthogonaux. Nous abordons dans ce paragraphe la construction des BIBD NNm-optimaux en utilisant les tableaux orthogonaux comme outils de base. Le lecteur pourra consulter, parmi les références traitant du sujet, Raghavarao (1971), Rao (1961, 1973), Morgan et Chakravarti (1988) et Hedayat *et al.* (1999). Considérons un ensemble Σ à v éléments et un tableau T $(b \times k)$ d'éléments pris dans Σ. Soit t un entier tel que $0 \leq t \leq k$. Un *tableau orthogonal* est souvent noté par $\text{OA}(b, k, v, t)$, ce sigle étant d'origine englo-saxon, OA mis pour "orthogonal array". Le nombre b désigne le nombre de lignes (ou de patients), la taille du OA, le nombre k est le nombre de colonnes (ou le nombre de traitements reçus par patient), v est le nombre de niveaux (ou le nombre totale de traitements dans l'expérience). Les paramètres d'un OA satisfont, $l = b/v^t$, où l est appelé index du tableau. Le cas particulier $l = 1$ est un cas important. Nous disons qu'on est en présence d'un OA d'index unité. Nous disons que T est un OA de niveau v de force t et d'index l, si dans tout bloc formé de t colonnes de T, les v^t t-uplets possibles apparaissent le même nombre de fois l chacun. L'existence et la construction des OA sont étudiés par Bush (1952), Bose et Bush (1952), Shrikhande (1964) et Yamamoto *et al.* (1984), parmi d'autres.

hal-00589585, version 1 - 29 Apr 2011

Deux autres arrangements de tableaux sont définis par Rao (1961) comme un OA de type I et II, renommés plus tard par un tableau *semi-équilibré* et un tableau *transitive*, notés respectivement par $\mathrm{SB}(b,k,v,t)$, SB mis pour "*semi-balanced*" et $\mathrm{TA}(b,k,v,t)$, TA mis pour "*transitive array*". Nous disons que T est un OA de type I ou un $\mathrm{TA}(b,k,v,t)$ de niveau s de force t et d'index l, si dans tout bloc formé de t colonnes de T, les $v!/(v-t)!$ t-uplets ordonnés d'éléments distincts apparaissent le même nombre de fois l chacun. Le tableau T est un OA de type II ou un $\mathrm{SBA}(b,k,v,t)$ de niveau s de force t et d'index l, si dans tout bloc formé de t colonnes de T, les $v!/t!(v-t)!$ t-uplets non ordonnés d'éléments distincts apparaissent le même nombre de fois l chacun. En particulier, un $\mathrm{SBA}(b,k,v,2)$ est un tableau $b \times k$ dans lequel chaque paire d'éléments non ordonnés dans chaque paire de colonnes apparaît le même nombre de fois l. Il est claire qu'un $\mathrm{TA}(b,k,v,t)$ d'index l est un $\mathrm{SBA}(b,k,v,t)$ d'index $l(t!)$.

Martin et Eccleston (1991) ont introduit le concept des plans fortement à voisinage équidistant, SDEN (SDEN pour "*strongly equineighboured design*"). Ils ont montré qu'un SDEN est *universellement faiblement optimal* pour $\hat{\gamma}$ sur $\Omega^*_{v,b,k}$ pour toute matrice de variances-covariances Δ. Dans ce type de plan, chaque traitement apparaît le même nombre de fois dans chaque période, et le nombre de fois qu'une paire (i,i') de traitements non ordonnés appliqués au même patient sur les périodes ℓ et ℓ' est indépendant de $1 \leq i \neq i' \leq v$, et $1 \leq \ell \neq \ell' \leq k$. les SDEN sont liés au plans NNm-équilibrés. Plusieurs auteurs ont remarqué qu'un SDEN pour lequel $k \geq 3$, est équivalent à un SB de force 2. Deheuvels et Dzerko (1991) ont utilisé les termes *totalement équilibrés* pour les SDEN et SBA, et *universellement équilibrés* pour les TA. Martin et Eccleston (1991) ont étudié en détail les propriétés d'optimalité des SDEN.

Les SBA et TA peuvent être vu comme des BIBD composés de v traitements et de b patients recevant chacun k traitements dans chaque k périodes. Le recours au SBA pour la construction de plans optimaux est examiné par Majumdar et Martin (2004).

Théorème 4.1 *Dans le modèle NNm, l'existence d'un $SBA\left(\frac{lv(v-1)}{2},k,v,2\right)$ implique l'existence d'un $BIBD\left(\frac{lv(v-1)}{2},v,k,\frac{lk(k-1)}{2},\frac{lk(v-1)}{2}\right)$ universellement faiblement optimale.*

Remarque 6 *Le théorème 4.1 de Morgan et Chakravarti (1988) est un corollaire de ce théorème lorsqu'on l'applique au cas $m = 2$.*

Corollaire 4.2 *Si $k \not\equiv 0 \mod (4)$, alors le BIBD obtenu à partir du $SB\left(\frac{v(v-1)}{2},k,v,2\right)$ est minimum NNm-optimal.*

Remarque 7 *Le corollaire 4.1(i) de Morgan et Chakravarti (1988) est similaire à ce corollaire lorsqu'on l'applique au cas $m = 2$.*

Remarque 8 *Pour réaliser un SBA il faut que b soit un entier multiple de $v!/(t!(v-t)!)$ ($b = lv!/(t!(v-t)!)$). Une condition nécessaire d'existence d'un TA est $b = lv!/(v-t)!$). Rao (1961, 1973) a montré que le nombre minimum de lignes pour réaliser un SBA de force 2 et d'index 1 est $b = v(v-1)/2$ si v est impair, $v(v-1)$ si v est pair, et que si v est une puissance d'un nombre premier impair, un SBA peut être construit à partir d'un corps de Galois à v éléments. Mukhopadhyay (1972) a construit un $\mathrm{SBA}(lv(v-1)/2,k,v,2)$ à partir d'un SBA et un OA, et a montré l'existence d'un SBA$(v(v-1)/2,k,v,2)$ pour $1 \leq k \leq 5$ et $v = 20s+1$ ou $20s+5$, $s \geq 1$. Lindner et al. (1987) ont construit des SBA pour $k = 4,5$ de force 2 et d'indexe 1 pour tout $v \geq 5$ impaire à l'exception de $v = 15, 39$. Morgan et Chakravarti (1988) ont construit des SBA de force 2 et d'indexe 2 pour $k = 3$ et v paire, et des SBA de k colonnes, de force 2 et d'indexe 1. Ils ont montré qu'un BIBD universellement faiblement optimal avec $k = 3$ pour les modèles NN1 et NN2 est équivalent à un SBA.*

5 Preuves

Nous rappelons que $\ell \in [\![1,m]\!]$, pour m fixé, A_i est le nombre de patients recevant le traitement i, $\#(A_i \cap A_{i'}) = \lambda$ est le nombre (cumulé) de fois où la paire (i,i') de traitements est appliquée au même patient, et

hal-00589585, version 1 - 29 Apr 2011

$\#A_i = r$ est le nombre de fois que le traitement i est répété dans la totalité de l'expérience. Nous inspirons des preuves de Morgan (1983); Morgan et Chakravarti (1988), et Kiefer et Wynn (1981) pour les modèles $NN1$ et $NN2$ en les généralisant.

5.1 Preuve du lemme 3.3

(i) Pour i fixé, $\sum_{i' \neq i} N^{\ell}_{d,i,i'}$ représente le nombre de patients recevant les traitements i et i' pour lesquels i et i' sont appliqués à ℓ intervalles de temps l'un de l'autre, sommés sur i', $\sum_{p=1}^{\ell} \phi^p_{d,i}$ est le nombre de patients pour lesquels le traitement i est appliqué à la $p^{\text{ième}}$ ou $(k-p+1)^{\text{ième}}$ période, sommés sur p. Il s'agit d'énumérer ces éléments. Soit λ_i^p le nombre de voisins d'ordre p du traitement i. Pour chaque patient recevant le traitement i la somme $\sum_{i' \neq i} N^{\ell}_{d,i,i'}$ est incrementée de $2 - \lambda_i^p$ où $\lambda_i^p = 0$ si i est appliqué à la $(p+1)^{\text{ième}}$ ou $(k-p)^{\text{ième}}$ période et $\lambda_i^p = 1$ sinon. Donc, comme il existe r patients recevant le traitement i, $\sum_{i' \neq i} N^{\ell}_{d,i,i'} = r(2 - \lambda_i^p) = 2r - r\lambda_i^p = 2r - \sum_{p=1}^{\ell} \phi^p_{d,i}$. ce qui achève la preuve.

□

(ii) $\phi^p_{d,i,i'}$ est le nombre de patients recevant les traitements i et i' pour lesquels i ou i' est appliqué à la $p^{\text{ième}}$ ou $(k-p+1)^{\text{ième}}$ période, où un patient est compté deux fois si i et i' sont à la $\ell^{\text{ième}}$ et $(k-p+1)^{\text{ième}}$ période. Le nombre $\phi^{\ell}_{d,i}$ de patients recevant le traitement i à la $p^{\text{ième}}$ ou $(k-p+1)^{\text{ième}}$ période contribue de k à la somme $\sum_{i' \neq i} \phi^p_{d,i,i'}$. Les $r - \phi^p_{d,i}$ patients pour lesquels le traitement i n'apparaît pas à la $p^{\text{ième}}$ ou à la $(k-p+1)^{\text{ième}}$ période contribue de 2 à la somme $\sum_{i' \neq i} \phi^p_{d,i,i'}$. Nous avons ainsi,

$$\sum_{i' \neq i} \phi^{\ell}_{d,i,i'} = k\phi^{\ell}_{d,i} + 2(r - \phi^p_{d,i}) = (k-2)\phi^p_{d,i} + 2r.$$

□

(iii) En énumérant, nous obtenons $2b$ traitements appliqués aux b patients à la $p^{\text{ième}}$ ou $(k-p+1)^{\text{ième}}$ période.

□

5.2 Preuve du lemme 3.5

Montrons l'identité (10). En utilisant l'estimateur linéaire sans biais des effets traitements corrigés $\hat{\gamma}$ défini dans (8), nous avons,

$$\begin{aligned}
D_d(\Delta^*) &= \mathbb{V}\text{ar}(\widehat{\gamma}) = k^2(\lambda v)^{-2} \left(A'_d - \frac{1}{k} A'_d (\mathbf{I}_b \otimes \mathbf{1}_k \mathbf{1}'_k) \right) (\mathbf{I}_b \otimes \Delta) \\
&\quad \times \left(A_d - \frac{1}{k} (\mathbf{I}_b \otimes \mathbf{1}_k \mathbf{1}'_k) A_d \right) \\
&= k^2(\lambda v)^{-2} \left(A'_d \left\{ (\mathbf{I}_b \otimes \Delta) - \frac{1}{k} (\mathbf{I}_b \otimes \mathbf{E}_k \Delta) - \frac{1}{k} (\mathbf{I}_b \otimes \Delta \mathbf{E}_k) + k^{-2} (\mathbf{I}_b \otimes \mathbf{E}_k \Delta \mathbf{E}_k) \right\} A_d \right) \\
&= k^2(\lambda v)^{-2} A'_d \left\{ \mathbf{I}_b \otimes \left(\mathbf{I}_k - \frac{1}{k} \mathbf{E}_k \right) \Delta \left(\mathbf{I}_k - \frac{1}{k} \mathbf{E}_k \right) \right\} A_d \\
&= k^2(\lambda v)^{-2} \sum_{s=1}^{b} T'_s \mathbb{W}(\Delta) T_s \quad \text{puisque} \quad A'_d = [T'_1, \ldots, T'_b]. \qquad (16)
\end{aligned}$$

hal-00589585, version 1 - 29 Apr 2011

Montrons maintenant l'identité (11). Nous cherchons à calculer $\mathbb{C}\mathrm{ov}(\hat{\gamma}_i, \hat{\gamma}_{i'})$. D'après l'identité (16), nous avons

$$\begin{aligned}(\lambda v)^2 k^{-2} D_{d,i,i'}(\Delta^*) &= \mathbb{C}\mathrm{ov}(\hat{\gamma}_i, \hat{\gamma}_{i'}) \\ &= \left(\sum_{s=1}^{b} T_s' \mathbb{W}(\Delta) T_s\right)_{i,i'} \\ &= \sum_{s=1}^{b}\sum_{\ell=1}^{v}\sum_{\ell'=1}^{v} \mathbf{t}_{\ell,i}(s)\mathbf{t}_{\ell',i'}(s)\mathbb{W}_{\ell,\ell'}.\end{aligned}$$

Puisque le patient s reçoit les traitements i et i' si et seulement si $\mathbf{t}_{\ell,i}(s)$ et $\mathbf{t}_{\ell',i'}(s)$ sont non nuls, voir le paragraphe (2), alors nous avons,

$$(\lambda v)^2 k^{-2} D_{d,i,i'}(\Delta^*) = \sum_{s\in A_i\cap A_{i'}}\sum_{\ell=1}^{v}\sum_{\ell'=1}^{v} \mathbf{t}_{\ell,i}(s)\mathbf{t}_{\ell',i'}(s)\mathbb{W}_{\ell,\ell'}.$$

Et, par définition de r (temporalité), on a,

$$(\lambda v)^2 k^{-2} D_{d,i,i'}(\Delta^*) = \sum_{s\in A_i\cap A_{i'}} \mathbb{W}_{r(i,s),r(i',s)}.$$

□

5.3 Preuve du lemme 3.6

Nous cherchons à calculer $\mathbb{C}\mathrm{ov}(\hat{\gamma}_i, \hat{\gamma}_{i'})$ avec $i \neq i'$ en explicitant l'expression de $\mathbb{W}_{\ell,\ell'}$. Nous avons,

$$\begin{aligned}\mathbb{W}_{\ell,\ell'} &= ((\mathbf{I}_k - k^{-1}\mathbf{E}_k)\Delta(\mathbf{I}_k - k^{-1}\mathbf{E}_k))_{\ell,\ell'} \\ &= \sigma^2\left(\rho_{\ell,\ell'} - k^{-1}\sum_{\ell=1}^{k}\rho_{\ell,\ell'} - k^{-1}\sum_{\ell'=1}^{k}\rho_{\ell,\ell'} + k^{-2}\sum_{\ell=1}^{k}\sum_{\ell'=1}^{k}\rho_{\ell,\ell'}\right). \qquad (17)\end{aligned}$$

Calculons chacun des termes du membre droit de l'identité (17) que nous sommons au besoin sur les patients s appartenant à $A_i \cap A_{i'}$. Posons $\ell = r(i,s)$ et $\ell' = r(i',s)$ les temporalités respectives auxquelles les traitements i et i' sont appliqués au patient s. Nous avons

$$\begin{aligned}\rho_{\ell,\ell'} &= \rho_{r(i,s),r(i',s)} \\ &= \rho_{|r(i,s)-r(i',s)|}\mathbb{1}_{\{s\,:\,|r(i,s)-r(i',s)|\leq m\}}, \text{ d'après (1).}\end{aligned}$$

Or $\{s : |r(i,s) - r(i',s)| \leq m\}$ est l'union $\bigcup_{p=1}^{m}\{|s : r(i,s) - r(i',s)| = p\}$. Cette union est disjointe car chaque patient reçoit un même traitement au plus une fois. Ainsi

$$\begin{aligned}&\mathbb{1}_{\{\{s\,:\,|r(i,s)-r(i',s)|=1\}\cup\{s\,:\,|r(i,s)-r(i',s)|=2\}\cup\cdots\cup\{s\,:\,|r(i,s)-r(i',s)|=m\}\}} \\ &\quad = \mathbb{1}_{\{s\,:\,|r(i,s)-r(i',s)|=1\}} + \mathbb{1}_{\{s\,:\,|r(i,s)-r(i',s)|=2\}} + \cdots + \mathbb{1}_{\{s\,:\,|r(i,s)-r(i',s)|=m\}}.\end{aligned}$$

Par conséquent

$$\begin{aligned}\sum_{s\in A_i\cap A_{i'}}\rho_{\ell,\ell'} &= \sum_{s\in A_i\cap A_{i'}}\rho_{|r(i,s)-r(i',s)|}\left[\mathbb{1}_{\{|r(i,s)-r(i',s)|=1\}} + \mathbb{1}_{\{|r(i,s)-r(i',s)|=2\}} + \cdots + \mathbb{1}_{\{|r(i,s)-r(i',s)|=m\}}\right] \\ &= \rho_1\sum_{s\in A_i\cap A_{i'}}\{s : |r(i,s) - r(i',s)| = 1\} + \rho_2\sum_{s\in A_i\cap A_{i'}}\{s : |r(i,s) - r(i',s)| = 2\} \\ &\quad + \cdots + \rho_m\sum_{s\in A_i\cap A_{i'}}\{s : |r(i,s) - r(i',s)| = m\} \\ &= \rho_1\#\{s\in A_i\cap A_{i'} : |r(i,s) - r(i',s)| = 1\} + \rho_2\#\{s\in A_i\cap A_{i'} : |r(i,s) - r(i',s)| = 2\} \\ &\quad + \cdots + \rho_m\#\{s\in A_i\cap A_{i'} : |r(i,s) - r(i',s)| = m\} \\ &= \sum_{p=1}^{m}\rho_p N^p_{d,i,i'}. \qquad (18)\end{aligned}$$

hal-00589585, version 1 - 29 Apr 2011

Remarque 9 Comme $i \neq i'$, $r(i,s) \neq r(i',s)$ et l'ensemble $\{s \ : \ |r(i,s) - r(i',s)| = 0\} = \phi$. Ainsi $\mathbb{1}_{\{s \ : \ |r(i,s)-r(i',s)|=0\}} = 0$.

Toujours dans l'objectif d'expliciter $\mathbb{W}_{\ell,\ell'}$, calculons $\sum_{\ell'=1}^{k} \rho_{\ell,\ell'}$ apparaissant dans le troisième terme de (17). On a

$$\begin{aligned}
\sum_{\ell'=1}^{k} \rho_{\ell,\ell'} &= \sum_{r(i',s)=1}^{k} \rho_{r(i,s),r(i',s)} \\
&= 1 \times \mathbb{1}_{\{r(i',s) \ : \ |r(i,s)-r(i',s)|=0\}} + \sum_{r(i',s)=1}^{k} \rho_{|r(i,s)-r(i',s)|} \mathbb{1}_{\{r(i',s) \ : \ 0<|r(i,s)-r(i',s)|\leq m\}} \\
&= 1 + \sum_{p=1}^{m} \rho_p \#\{b \in [\![1,k]\!] \ : \ |r(i,s) - b| = p\} \text{ en posant } b = r(i',s). \qquad (19)
\end{aligned}$$

Posons $D = \sum_{p=1}^{m} \rho_p \#\{b \in [\![1,k]\!] \ : \ |r(i,s) - b| = p\}$. Si $\exists \ p_0 \in [\![1,m]\!]$ tel que $r(i,s) \in \{p_0, k - p_0 + 1\}$ (i.e., le traitement i est administré à la $p_0^{\text{ième}}$ ou $(k-p_0+1)^{\text{ième}}$ période au patient $s^{\text{ième}}$) alors

$$\#\{b \in [\![1,k]\!] \ : \ |r(i,s) - b| = p\} = \begin{cases} 2 & \text{pour } p \in [\![1,p_0]\!] \text{ lorsque } p_0 > 1 \\ 1 & \text{pour } p \in [\![p_0+1,m]\!] \ (= [\![1,m]\!] \text{ lorsque } p_0 = 1) \end{cases}$$

Dans le cas contraire, si $r(i,s) \in [\![m+1, k-m]\!]$ alors $\forall \ p \in [\![1,m]\!]$

$$\#\{b \in [\![1,k]\!] \ : \ |r(i,s) - b| = p\} = 2.$$

Par conséquent, $\forall \ s \in A_{i'}$, si $s \in A_i$ alors

$$D = \begin{cases} 2\rho_1 + 2\rho_2 + 2\rho_3 + \cdots + 2\rho_{p_0-1} + \rho_{p_0} + \cdots + \rho_m & \text{si } \exists \ p_0 \in [\![1,m]\!] \ : \ r(i,s) \in \{p_0, k-p_0+1\} \\ 2\rho_1 + 2\rho_2 + 2\rho_3 + \cdots + 2\rho_m & \text{si } r(i,s) \in [\![m+1, k-m]\!]. \end{cases}$$

Si $s \notin A_i$ alors $D = 0$. Ainsi,

$$\begin{aligned}
\sum_{s \in A_{i'} \cap A_i} (1 + D) &= \sum_{s \in A_{i'} \cap A_i} (1 + \rho_1 + \rho_2 + \rho_3 + \cdots + \rho_m \\
&\quad + \ \rho_1 \mathbb{I}^1_{r(i,s)} + \rho_2 \mathbb{I}^2_{r(i,s)} + \rho_3 \mathbb{I}^3_{r(i,s)} +, \cdots, +\rho_m \mathbb{I}^m_{r(i,s)}), \qquad (20)
\end{aligned}$$

où, pour tout $p \in [\![1,m]\!]$

$$\mathbb{I}^p_{r(i,s)} = \begin{cases} 1 & \text{si } r(i,s) \in [\![p+1, k-p]\!] \\ 0 & \text{sinon .} \end{cases}$$

De même, pour le terme $\sum_{\ell=1}^{k} \rho_{\ell,\ell'}$, on a $\forall \ s \in A_i$, si $s \in A_{i'}$ alors

$$\begin{aligned}
\sum_{s \in A_i \cap A_{i'}} (1 + C) &= \sum_{s \in A_i \cap A_{i'}} (1 + \rho_1 + \rho_2 + \rho_3 + \cdots + \rho_m + \rho_1 \mathbb{I}^1_{r(i',s)} \\
&\quad + \ \rho_2 \mathbb{I}^2_{r(i',s)} + \rho_3 \mathbb{I}^3_{r(i',s)} + \cdots + \rho_m \mathbb{I}^m_{r(i',s)}), \qquad (21)
\end{aligned}$$

où,

$$C = \sum_{p=1}^{m} \rho_p \#\{a \in [\![1,k]\!] \ : \ |a - r(i',s)| = p\}. \qquad (22)$$

Terminons par le dernier membre de (17), c'est-à- dire $\sum_{\ell=1}^{k} \sum_{\ell'=1}^{k} \rho_{\ell,\ell'}$. Nous avons,

hal-00589585, version 1 - 29 Apr 2011

$$\begin{aligned}
\sum_{\ell=1}^{k}\sum_{\ell'=1}^{k}\rho_{\ell,\ell'} &= \sum_{r(i,s)=1}^{k}\sum_{r(i',s)=1}^{k}\rho_{r(i,s),r(i',s)} \\
&= k\times \mathbb{1}_{\{(r(i,s),r(i',s))\,:\,|r(i,s)-r(i',s)|=0\}} \\
&\quad + \sum_{r(i,s)=1}^{k}\sum_{\substack{r(i',s)\in[\![1,k]\!]:\\ i'\neq i}}\rho_{|r(i,s)-r(i',s)|}\mathbb{1}_{\{(r(i,s),r(i',s))\,:\,0<|r(i,s)-r(i',s)|\leq m\}} \\
&= k+\sum_{p=1}^{m}\rho_p \#\{(a,b)\ \in\ [\![1,k]\!]^2\ :\ a\neq b\ \text{ et } |a-b|=p\} \qquad (23)\\
&\quad (\text{ en posant } a=r(i,s) \text{ et } b=r(i',s)).
\end{aligned}$$

Calculons le coefficient de ρ_p. Pour les $2p$ valeurs a appartenant à $[\![1,p]\!]\bigcup[\![k-p+1,k]\!]$, il n'éxiste qu'une valeur b dans $[\![1,k]\!]$ telle que $|a-b|=p$. Soit $2p$ valeurs possibles pour a et b. Pour chacune des $(k-2p)$ valeurs a appartenant à $[\![p+1,k-p]\!]$, il existe 2 valeurs b dans $[\![1,k]\!]$ telles que $|a-b|=p$. Soit $2(k-2p)$ valeurs au total. Donc le coefficient de ρ_p est

$$\#\{(a,b)\ \in\ [\![1,k]\!]^2\ :\ a\neq b\ \text{ et } |a-b|=p\}=2p+2(k-2p)=2(k-p).$$

En conséquence, on a

$$\begin{aligned}
\sum_{s\in A_i\cap A_{i'}}\sum_{\ell=1}^{k}\sum_{\ell'=1}^{k}\rho_{\ell,\ell'} &= \lambda[k+(2(k-2)+2)\rho_1+(2(k-4)+4)\rho_2+(2(k-6)+6)\rho_3 \\
&\quad +\cdots+(2(k-2p)+2p)\rho_p+\cdots+(2(k-2m)+2m)\rho_m] \\
&= \lambda[k+2(k-1)\rho_1+2(k-2)\rho_2+2(k-3)\rho_3+(2(k-2p)+2p)\rho_p \\
&\quad +\cdots+2(k-m)\rho_m]. \qquad (24)
\end{aligned}$$

Ainsi, en combinant les équations (18), (20), (21) et (24), on obtient que,

$$\begin{aligned}
(\lambda v)^2\sigma^{-2}\mathbb{C}\mathrm{ov}(\hat{\gamma}_i,\hat{\gamma}_{i'}) &= k^2\sum_{\ell=1}^{m}\rho_\ell N^\ell_{d,i,i'}-k\rho_1\sum_{s\in A_i\cap A_{i'}}\left(\mathbb{I}^1_{r(i,s)}+\mathbb{I}^1_{r(i',s)}\right) \\
&\quad -\ k\rho_2\sum_{s\in A_i\cap A_{i'}}\left(\mathbb{I}^2_{r(i,s)}+\mathbb{I}^2_{r(i',s)}\right) \\
&\quad -\ \cdots-k\rho_m\sum_{s\in A_i\cap A_{i'}}\left(\mathbb{I}^m_{r(i,s)}+\mathbb{I}^m_{r(i',s)}\right) \\
&\quad -\ \lambda[k+2(\rho_1+2\rho_2+3\rho_3+\cdots+m\rho_m)] \\
&= k^2\sum_{\ell=1}^{m}\rho_\ell N^\ell_{d,i,i'}-k\rho_1(2\lambda-\phi^1_{d,i,i'})-k\rho_2(2\lambda-\phi^1_{d,i,i'}-\phi^2_{d,i,i'}) \\
&\quad -\ \cdots-k\rho_m(2\lambda-\phi^1_{d,i,i'}-\phi^2_{d,i,i'}-\cdots-\phi^m_{d,i,i'}) \\
&\quad -\ \lambda[k+2(\rho_1+2\rho_2+3\rho_3+\cdots+m\rho_m)] \\
&= k^2\sum_{\ell=1}^{m}\rho_\ell N^\ell_{d,i,i'}-k\sum_{\ell=1}^{m}\rho_\ell\left(2\lambda-\sum_{p=1}^{\ell}\phi^p_{d,i,i'}\right) \\
&\quad -\ \lambda[k+2(\rho_1+2\rho_2+3\rho_3+\cdots+m\rho_m)] \\
&= -\lambda[k+2(k+1)\rho_1+2(k+2)\rho_2+2(k+3)\rho_3 \\
&\quad +\cdots+2(k+m)\rho_m]+k\sum_{\ell=1}^{m}\rho_\ell\left(\sum_{p=1}^{\ell}\phi^p_{d,i,i'}+kN^\ell_{d,i,i'}\right),
\end{aligned}$$

ce qu'il fallait démontrer.

□

Le calcul de la variance est rigoureusement identique à celui de la covariance en posant $A_i = A_{i'}$ dans les équations (18), (20), (21) et (24).

$$\begin{aligned}
(\lambda v)^2\sigma^{-2}\mathbb{V}\mathrm{ar}(\widehat{\gamma}_i) &= \sum_{s\in A_i}[k^2-k-2(\rho_1+2\rho_2+3\rho_3+\cdots+m\rho_m)\\
&\quad -2k\rho_1\mathbb{I}^1_{r(i,s)}-2k\rho_2\mathbb{I}^2_{r(i,s)}-\cdots-2k\rho_m\mathbb{I}^m_{r(i,s)}]\\
&= r[k^2-k-2(\rho_1+2\rho_2+3\rho_3+\cdots+m\rho_m)]\\
&\quad -2k\rho_1\sum_{s\in A_i}\mathbb{I}^1_{r(i,s)}-2k\rho_2\sum_{s\in A_i}\mathbb{I}^2_{r(i,s)}\\
&\quad -\cdots-2k\rho_m\sum_{s\in A_i}\mathbb{I}^m_{r(i,s)}\\
&= r[k(k-1)-2(\rho_1+2\rho_2+3\rho_3+\cdots+m\rho_m)]\\
&\quad -2k\rho_1(r-\phi^1_{d,i})-2k\rho_2(r-\phi^1_{d,i}-\phi^2_{d,i})\\
&\quad -\cdots-2k\rho_m(r-\phi^1_{d,i}-\phi^2_{d,i}-\cdots-\phi^m_{d,i})\\
&= r[k(k-1)-2(\rho_1+2\rho_2+3\rho_3+\cdots+m\rho_m)]\\
&\quad -2rk\rho_1-2rk\rho_2-\cdots-2rk\rho_m\\
&\quad +2k\rho_1\phi^1_{d,i}+2k\rho_2(\phi^1_{d,i}+\phi^2_{d,i})\\
&\quad +\cdots+2k\rho_m(\phi^1_{d,i}+\phi^2_{d,i}+\cdots+\phi^m_{d,i})\\
&= r[k(k-1)-2(k+1)\rho_1-2(k+2)\rho_2-2(k+3)\rho_3\\
&\quad -\cdots-2(k+m)\rho_m]+2k\sum_{\ell=1}^{m}\rho_\ell\sum_{p=1}^{\ell}\phi^p_{d,i},
\end{aligned}$$

ce qu'il fallait démontrer.

□

5.4 Preuve du lemme 3.7

Si $k=3, \forall\, \ell \in [\![3,m]\!]$ d'après les (i) et (iv) du lemme 3.4 on a $N^\ell_{d,i,i'}=0$ et $\phi^\ell_{d,i,i'}=0$. Ainsi, en procédant comme dans la preuve du lemme 3.6, nous obtenons,

$$\sum_{s\in A_i\cap A_{i'}}\rho_{\ell,\ell'}=\rho_1 N^1_{d,i,i'}+\rho_2 N^2_{d,i,i'}. \tag{25}$$

Si $s\notin A_i$ alors $D=0$.

$$\sum_{s\in A_{i'}\cap A_i}\sum_{\ell'=1}^{3}\rho_{\ell,\ell'}=\sum_{s\in A_{i'}\cap A_i}(1+D)=\sum_{s\in A_{i'}\cap A_i}(1+\rho_1+\rho_2+(\rho_1-\rho_2)\mathbb{I}_{r(i,s)=2}), \tag{26}$$

où,

$$D=\sum_{p=1}^{2}\rho_p\#\{b\in[\![1,3]\!]\,:\,|r(i,s)-b|=p\}.$$

De même, pour le terme

$$\sum_{s\in A_i\cap A_{i'}}\sum_{\ell=1}^{3}\rho_{\ell,\ell'}=\sum_{s\in A_i\cap A_{i'}}(1+C)=\sum_{s\in A_i\cap A_{i'}}(1+\rho_1+\rho_2+(\rho_1-\rho_2)\mathbb{I}_{r(i',s)=2}), \tag{27}$$

hal-00589585, version 1 - 29 Apr 2011

hal-00589585, version 1 - 29 Apr 2011

où,

$$C = \sum_{p=1}^{2} \rho_p \#\{a \in [\![1,3]\!] : |a - r(i',s)| = p\}. \tag{28}$$

Nous obtenons, pour le dernier membre de (17),

$$\sum_{s \in A_i \cap A_{i'}} \sum_{\ell=1}^{3} \sum_{\ell'=1}^{3} \rho_{\ell,\ell'} = \lambda(3 + 4\rho_1 + 2\rho_2). \tag{29}$$

En combinant les identités (25), (26), (27) et (29), on obtient que

$$\begin{aligned}
(\lambda v)^2 \sigma^{-2} \mathbb{C}\mathrm{ov}(\hat{\gamma}_i, \hat{\gamma}_{i'}) &= 9(\rho_1 N^1_{d,i,i'} + \rho_2 N^2_{d,i,i'}) - 3(\rho_1 - \rho_2) \sum_{s \in A_i \cap A_{i'}} (\mathbb{I}_{r(i,s)=2} + \mathbb{I}_{r(i',s)=2}) \\
&\quad - \lambda(3 + 2\rho_1 + 4\rho_2) \\
&= 9(\rho_1 N^1_{d,i,i'} + \rho_2 N^2_{d,i,i'}) - 3(\rho_1 - \rho_2)(2\lambda - \phi^1_{d,i,i'}) \\
&\quad - \lambda(3 + 2\rho_1 + 4\rho_2)
\end{aligned}$$

Pour achever la preuve on utilise les (ii) et (v) du lemme 3.4, en posant $N^2_{d,i,i'} = \lambda - N^1_{d,i,i'}$ et $2\lambda - \phi^1_{d,i,i'} = N^1_{d,i,i'}$ dans la dernière égalité. Ainsi,

$$\begin{aligned}
(\lambda v)^2 \sigma^{-2} \mathbb{C}\mathrm{ov}(\hat{\gamma}_i, \hat{\gamma}_{i'}) &= 9(\rho_1 N^1_{d,i,i'} + \rho_2(\lambda - N^1_{d,i,i'})) \\
&\quad - 3(\rho_1 - \rho_2) N^1_{d,i,i'} - \lambda(3 + 2\rho_1 + 4\rho_2) \\
&= 9\rho_1 N^1_{d,i,i'} - 9\rho_2 N^1_{d,i,i'} - 3\rho_1 N^1_{d,i,i'} + 3\rho_2 N^1_{d,i,i'} \\
&\quad - \lambda(3 + 2\rho_1 + 4\rho_2 - 9\rho_2),
\end{aligned}$$

ce qu'il fallait démontrer.

□

Comme dans le calcul précédent de la variance, remplaçons $A_1 = A_2$ dans les identités (25), (26), (27) et (29), on obtient, alors, en les combinant,

$$\begin{aligned}
(\lambda v)^2 \sigma^{-2} \mathrm{Var}(\hat{\gamma}_i) &= 9r - r(3 + 2\rho_1 + 4\rho_2) - 6(\rho_1 - \rho_2) \sum_{s \in A_i} \mathbb{I}_{r(i,s)=2} \\
&= r(6 - 2\rho_1 - 4\rho_2) - 6(\rho_1 - \rho_2)(r - \phi^1_{d,i}) \\
&= 2r[(3 - 4\rho_1 + \rho_2) + 3(\rho_1 - \rho_2)\phi^1_{d,i}],
\end{aligned}$$

ce qu'il fallait démontrer.

□

5.5 Preuve du théorème 3.9

Nous rappelons qu'une matrice complètement symétrique a ses éléments diagonaux et extra-diagonaux égaux.

Supposons que la matrice D_d définie dans le lemme 3.5 soit complèment symétrique. A partir du lemme

hal-00589585, version 1 - 29 Apr 2011

3.6 et de l'identité (9), pour un CBD dans le modèle NNm, nous avons

$$\begin{aligned}\sigma^{-2}(bv)^2 D_d = & \\ & b[v(v-1) - 2(v+1)\rho_1 - 2(v+2)\rho_2 - 2(v+3)\rho_3 - \cdots - 2(v+m)\rho_m]\mathbf{I}_v \\ & + \ 2v\sum_{\ell=1}^{m}\rho_\ell \sum_{p=1}^{\ell}\mathrm{diag}(\phi_d^p) + v^2\sum_{\ell=1}^{m}\rho_\ell \mathrm{N}_d^\ell - b[v + 2(v+1)\rho_1 + 2(v+2)\rho_2 \\ & + \ 2(v+3)\rho_3 + \cdots + 2(v+m)\rho_m](\mathbf{E}_v - \mathbf{I}_v) + v\sum_{\ell=1}^{m}\rho_\ell\sum_{p=1}^{\ell}\left(\Phi_d^p - 2\mathrm{diag}(\phi_d^p)\right) \\ = \ & bv^2\mathbf{I} - b[v + 2(v+1)\rho_1 + 2(v+2)\rho_2 + 2(v+3)\rho_3 + \cdots + 2(v+m)\rho_m]\mathbf{E} \\ & + \ v^2\sum_{\ell=1}^{m}\rho_\ell N_d^\ell + v\sum_{\ell=1}^{m}\rho_\ell\sum_{p=1}^{\ell}\Phi_d^p. \end{aligned} \tag{30}$$

Soit ε un vecteur orthonormé de $\mathbb{R}^v$ tel que $\mathbf{1}_v'\varepsilon = 0$ (i.e., $\sum_{i=1}^{v}\varepsilon_i = 0$), alors $\mathbf{E}_v\varepsilon = 0_v$. En supposant que

$$\Phi_d^\ell = \phi_d^\ell \mathbf{1}_v' + \mathbf{1}_v\phi_d'^\ell, \ \ \forall \ \ell \in [\![1, m]\!], \tag{31}$$

avec (30) on obtient que

$$\mathbb{V}\mathrm{ar}(\varepsilon'\hat{\gamma}) = \varepsilon' D_d \varepsilon = \sigma^2 b^{-2}\left[b + \sum_{\ell=1}^{m}\rho_\ell\sum_{i=1}^{v}\sum_{i'\neq i}\varepsilon_i\varepsilon_{i'}N_{d,i,i'}^\ell\right]. \tag{32}$$

Le fait que la matrice D_d soit complètement symétrique implique l'égalité de la variance pour tous les contrastes de traitements normalisés. Ainsi pour tout vecteur $\xi \in \mathbb{R}^v$ tel que $\sum_{i=1}^{v}\xi_i = 0$ et $\sum_{i=1}^{v}\xi_i^2 = 1$, nous avons

$$\sum_{\ell=1}^{m}\rho_\ell\left(\sum_{i=1}^{v}\sum_{i'\neq i}\varepsilon_i\varepsilon_{i'}N_{d,i,i'}^\ell - \sum_{i=1}^{v}\sum_{i'\neq i}\xi_i\xi_{i'}N_{d,i,i'}^\ell\right) = 0 \tag{33}$$

Pour les couples (i, i') et (j, j') tels que $i \neq i'$ et $j \neq j'$, avec un choix approprié des vecteurs ε et ξ, l'identité (33) est équivalente à

$$\sum_{\ell=1}^{m}\rho_\ell\left(N_{d,i,i'}^\ell - N_{d,j,j'}^\ell\right) = 0 \ \ \forall \ \ (i,i'), i \neq i' \ \text{ et } \forall \ \ (j,j'), j \neq j'. \tag{34}$$

Pour $m = \ell = 1$, nous avons donc puisque $\rho_1 \neq 0$

$$N_{d,i,i'}^1 = N_{d,j,j'}^1 \ \ \forall \ \ (i,i'), i \neq i' \ \ (j,j'), j \neq j'. \tag{35}$$

Pour $m = 2$, d'après les identités (34) et (35), impliquent

$$\rho_2(N_{d,i,i'}^2 - N_{d,j,j'}^2) = 0, \ \ \text{d'où, comme} \quad \rho_2 \neq 0,$$

$$N_{d,i,i'}^2 = N_{d,j,j'}^2 \ \ \forall \ \ (i \neq i'), \ \ (j \neq j'),$$

et ainsi de suite, $\forall\ \ell \in [\![1, m]\!]$ toutes les quantités $N_{d,i,i'}^\ell$, de l'identité (32) sont identiques, ce qu'il fallait démontrer.

□

Montrons maintenant la symétrie complète de la matrice D_d.
Nous savons que pour un CBD, $\forall\ \ell \in [\![1, m]\!]$,

$$\phi_{d,i,i'}^\ell = \phi_{d,i}^\ell + \phi_{d,i'}^\ell. \tag{36}$$

A partir des quantités $\sum_{i'\neq i}N_{d,i,i'}^\ell = 2r - \sum_{p=1}^{\ell}\phi_{d,i}^p$, définies dans le lemme 3.3(i), nous pouvons voir que l'égalité des termes $N_{d,i,i'}^\ell$ implique l'égalité des termes $\phi_{d,i,i'}^\ell$. Ainsi, la matrice D_d est complètement symétrique, ce qu'il fallait démontrer.

□

hal-00589585, version 1 - 29 Apr 2011

5.6 Preuve du théorème 3.8

D'après la conséquence des lemmes 3.6 et 3.7 (voir paragraphe 2.2) tous les plans compétiteurs possèdent la même trace. Ainsi, par la proposition 3.2, l'*optimalité universelle faible* des BIBD pour le modèle NNm est celle pour laquelle la matrice de variances-covariances D_d est complétement symétrique. Cela signifie que

1. les éléments diagonaux $\sum_{\ell=1}^{m} \rho_\ell \sum_{p=1}^{\ell} \phi_{d,i}^p$ sont indépendantes de i.
2. les éléments extra-diagonaux $\sum_{\ell=1}^{m} \rho_\ell \left(\sum_{p=1}^{\ell} \phi_{d,i,i'}^p + kN_{d,i,i'}^\ell \right)$ sont indépendantes de i, i', $(i \neq i')$.

La condition 2 implique la condition 1 puisque la somme en ligne ou en colonne de D_d est nulle, et elle est équivalente à, $\forall\ (i,i'), i \neq i', (j,j'), j \neq j'$,

$$\sum_{\ell=1}^{m} \rho_\ell \left(\sum_{p=1}^{\ell} \phi_{d,i,i'}^p + kN_{d,i,i'}^\ell \right) - \sum_{\ell=1}^{m} \rho_\ell \left(\sum_{p=1}^{\ell} \phi_{d,j,j'}^p + kN_{d,j,j'}^\ell \right) = 0$$

$\Longleftrightarrow$

$$\sum_{\ell=1}^{m} \rho_\ell \left(\sum_{p=1}^{\ell} \phi_{d,i,i'}^p + kN_{d,j,j'}^\ell - \sum_{p=1}^{\ell} \phi_{d,j,j'}^p - kN_{d,j,j'}^\ell \right) = 0$$

En utilisant le même raisonnement que la preuve du théorème 3.9 dans le paragraphe ci-dessus, nous obtenons le résultat énoncé. Pour $k = 3$, par le lemme 3.4(i), (iv) et (ii), nous avons respectivement, $N_{d,i,i'}^\ell = \phi_{d,i,i'}^\ell = 0\ \ \forall\ \ \ell \in [\![3, m]\!]$, et $\phi_{d,i,i'}^2 = N_{d,i,i'}^1$. Dans ce cas les valeurs des quantités facteurs de ρ_ℓ dans $\sum_{\ell=1}^{m} \rho_\ell \left(\sum_{p=1}^{\ell} \phi_{d,i,i'}^p + kN_{d,i,i'}^\ell \right)$ sont respectivement,

$$\begin{aligned} \phi_{d,i,i'}^1 + kN_{d,i,i'}^1 &= 2\lambda + (k-1)N_{d,i,i'}^1, \\ \phi_{d,i,i'}^1 + \phi_{d,i,i'}^2 + kN_{d,i,i'}^2 &= (k+2)\lambda - kN_{d,i,i'}^1. \end{aligned}$$

Le résultat découle du lemme 3.7. □

5.7 Preuve de la proposition 3.11

Posons

$$F(i,i',\ell,p) = \sum_{p=1}^{\ell} \phi_{d,i,i'}^p + kN_{d,i,i'}^\ell. \tag{37}$$

Nous avons, alors

$$F(i,i') = \rho_1 F(i,i',1,p) + \rho_2 F(i,i',2,p) + \cdots + \rho_m F(i,i',m,p).$$

Pour que chaque $F(i,i',\ell,p)$ soit indépendante de i, i' $(i \neq i')$, il est nécessaire (mais pas suffisant) que la sommation $F = \sum_i \sum_{i' \neq i} F(i,i')$ sur chacune d'elle divisée par le nombre $v(v-1)$ de $\rho_\ell F(i,i',\ell,p)$ sommés dans F soit un entier indépendant de i, i'.

Fixons $\ell \in [\![1, m]\!]$. Le facteur F_ℓ de ρ_ℓ dans F est

$$\begin{aligned} F_\ell = \sum_i \sum_{i' \neq i} F(i,i',\ell,p) &= \sum_i \sum_{i' \neq i} \left[\sum_{p=1}^{\ell} \phi_{d,i,i'}^p + kN_{d,i,i'}^\ell \right] \\ &= \sum_i \sum_{i' \neq i} \sum_{p=1}^{\ell} \phi_{d,i,i'}^p + \sum_i \sum_{i' \neq i} kN_{d,i,i'}^\ell, \\ &= \sum_{p=1}^{\ell} \left[(k-2) \sum_{i=1}^{v} \phi_{d,i}^p + \sum_{i=1}^{v} 2r \right] + k \sum_{i=1}^{v} \left[2r - \sum_{p=1}^{\ell} \phi_{d,i}^\ell \right], \end{aligned}$$

hal-00589585, version 1 - 29 Apr 2011

d'après les (i) et (ii) du lemme 3.3. Et puisque $\sum_i \phi^{\ell}_{d,i} = 2b$, d'après le (iii) du lemme 3.3, nous avons:

$$\begin{aligned} F_\ell = \sum_i \sum_{i' \neq i} F(i,i',\ell,p) &= 2\ell b(k-2) + 2\ell rv + 2rvk - 2\ell bk \\ &= 2vr(k+\ell) - 4b\ell \\ &= 2b[k(k+\ell) - 2\ell] \\ &= 2b[(k-1)(k+2\ell) - k(\ell-1)] \\ &= \frac{2vr[(k-1)(k+2\ell) - k(\ell-1)]}{k} \text{ d'après l'identité } (i) \text{ en (6)}, \\ &= \frac{2\lambda v(v-1)[(k-1)(k+2\ell) - k(\ell-1)]}{k(k-1)}. \end{aligned}$$

Ainsi

$$\begin{aligned} \frac{F_\ell}{v(v-1)} &= \frac{2\lambda(k+2\ell)}{k} - \frac{2\lambda(\ell-1)}{k-1} \\ &= 2\lambda + \frac{4\lambda\ell}{k} - \frac{2\lambda(\ell-1)}{k-1}. \end{aligned} \tag{38}$$

Par ailleurs, si l'hypothèse de la proposition 3.11 est satisfaite alors, d'après le théorème 3.8 les valeurs $F(i,i',\ell,p)$ sont indépendantes de $i, i'(i \neq i')$. Par conséquent

$$\forall\, i, \forall\, i' \neq i, F(i,i',\ell,p) = \frac{F_\ell}{v(v-1)} = 2\lambda + \frac{4\lambda\ell}{k} - \frac{2\lambda(\ell-1)}{k-1}$$

est une quantité entière pour tout $\ell \in [\![1,m]\!]$. En posant $\ell = 1$, pour que

$$F(i,i',1,p) = \frac{F_1}{v(v-1)} = 2\lambda + \frac{4\lambda}{k}$$

soit une valeur entière, il est nécessaire que $k|4\lambda$, et par conséquent

$$2\lambda + \frac{4\lambda\ell}{k} \in \mathbb{N}\ \forall\, \ell \in [\![1,m]\!]. \tag{39}$$

Pour $m > 1$, et $\ell = 2$, pour que

$$F(i,i',2,p) = \frac{F_2}{v(v-1)} = 2\lambda + \frac{8\lambda}{k} - \frac{2\lambda}{k-1}$$

soit une valeur entière, par la condition (39), il est nécessaire que $(k-1)|2\lambda$. Par conséquent

$$2\lambda + \frac{4\lambda\ell}{k} - \frac{2\lambda(\ell-1)}{k-1} \in \mathbb{N}\ \forall\, \ell \in [\![1,m]\!].$$

Puisque que p.g.c.d$(k,k-1) = 1$, les conditions nécessaires $k|4\lambda$ et $(k-1)|2\lambda$ impliquent $k(k-1)|4\lambda$. Si, de plus, $k \not\equiv 0 \pmod 4$, alors

$$k|4\lambda \Rightarrow k|2\lambda \Rightarrow k(k-1)|2\lambda,$$

ce qu'il fallait démontrer.

□

hal-00589585, version 1 - 29 Apr 2011

Preuve du corollaire 3.12 On utilise le théorème 3.9 et la méthode énoncée ci-dessus dans la cas des BIBD.

Soit $G(i,i') = \sum_{\ell=1}^{m} \rho_\ell G(i,i',\ell)$, où $G(i,i',\ell) = N^{\ell}_{d,i,i'}$. Le facteur G_ℓ de ρ_ℓ dans G est

$$G_\ell = \sum_{i=1}^{v} \sum_{i' \neq i} G(i,i',\ell) = 2b(k-\ell).$$

Ainsi

$$\frac{G_\ell}{v(v-1)} = \frac{2b(k-\ell)}{v(v-1)}.$$

Par ailleurs, si l'hypothèse de la proposition 3.2 est satisfaite alors, d'après le théorème 3.9 les valeurs $G(i,i',\ell)$ sont indépendantes de i, i', $(i \neq i')$. Par conséquent

$$\forall\, i, \forall\, i' \neq i, G(i,i',\ell) = \frac{G_\ell}{v(v-1)} = \frac{2b(k-\ell)}{v(v-1)}$$

est une quantité entière pour tout $\ell \in [\![1,m]\!]$. Pour que $G(i,i',\ell)$ soit une valeur entière, il est nécessaire que $v(v-1)|2b$, ce qu'il fallait démontrer.

□

Remarque 10 *Dans la démonstration de la proposition 3.11 et du corollaire 3.12, nous retrouvons les conditions d'existence des plans $NN1, NN2$-optimaux respectivement de Kiefer et Wynn (1981), et Morgan et Chakravarti (1988).*

5.8 Preuve du théorème 4.1

On commence par identifier le nombre b de patients à $\frac{lv(v-1)}{2}$. Nous rappelons que les conditions nécessaires pour l'existence d'un BIBD sont,

$$\begin{aligned} bk &= rv \\ \lambda(v-1) &= r(k-1). \end{aligned} \tag{40}$$

Les équations (40) impliquent,

$$\lambda = \frac{bk(k-1)}{v(v-1)} = \frac{k(k-1)}{2}, \text{ et } r = \frac{bk}{v} = \frac{l(v-1)k}{2}.$$

On a aussi (voir le lemme 3.3(i)),

$$\sum_{i' \neq i} N^{\ell}_{d,i,i'} = 2r - \sum_{p=1}^{\ell} \phi^{p}_{d,i} \ \ ,\forall 1 \leq i \neq i' \leq v, \forall\, \ell \in [\![1,m]\!] \text{ pour } \ m \text{ fixé},$$

par le fait que $\sum_{i=1}^{v} \phi^{1}_{d,i} = 2b$ et pour $k > 3 \ \forall \ \ell \in [\![2,m]\!]$, $\sum_{i=1}^{v} \phi^{\ell}_{d,i} = 2b$, (voir le lemme 3.3$(iii)$), $\Rightarrow$

$$\sum_{i=1}^{v} \sum_{i' \neq i} N^{\ell}_{d,i,i'} = 2rv - \sum_{p=1}^{\ell} \sum_{i=1}^{v} \phi^{p}_{d,i} = \frac{2l(k-\ell)v(v-1)}{2}.$$

D'autre part (voir le lemme 3.3(ii)), on a

$$\sum_{i' \neq i} \phi^{\ell}_{d,i,i'} \ = \ (k-2)\phi^{\ell}_{d,i} + 2r, \ \ \forall \ \ell \in [\![1,m]\!] \text{ pour } m \text{ fixé }, \ \ 1 \leq i \neq i' \leq v.$$

$\Rightarrow \sum_i \sum_{i' \neq i} \phi^{\ell}_{d,i,i'} = (k-2) \sum_i \phi^{\ell}_{d,i} + 2rv = \frac{2(k-2)v(v-1)}{2} + \frac{2lv(v-1)k}{2} = \frac{4l(k-1)v(v-1)}{2}$.
On a donc,

$$N^{\ell}_{d,i,i'} = l(k-\ell), \text{ et } \phi^{\ell}_{d,i,i'} = 2l(k-1), \forall \ \ell \in [\![1,m]\!], \text{ et } (1 \leq i \neq i' \leq v),$$

et comme l'optimalité des BIBD pour le modèle NNm requiert l'égalité de chacune des quantités facteur de ρ_ℓ des $F(i,i')$, alors on a bien le résultats indiqué. □

hal-00589585, version 1 - 29 Apr 2011

Preuve du corollaire 4.2 Il suffit de calculer la valeur $\frac{F_\ell}{v(v-1)}$ et de voir la condition pour qu'elle soit entier.

$$\frac{F_\ell}{v(v-1)} = 2bk(k-\ell) + 4\ell b(k-1) = \frac{2\lambda(k-\ell)}{k-1} + \frac{4\lambda}{k}.$$

□

remerciements

Nous remercions sincèrement le Professeur Paul Deheuvels d'avoir contribué par ses suggestions constructives à la qualité de ce article.

Références

Agrawal, H. (1966a). Some generalizations of distinct representatives with applications to statistical designs. *Ann. Math. Statist.*, **37**, 525–528.

Agrawal, H. (1966b). Some methods of construction of designs for two-way elimination of heterogeneity. I. *J. Amer. Statist. Assoc*, **61**, 1153–1171.

Azzalini, A. et Giovagnoli, A. (1987). Some optimal designs for repeated measurements with autoregressive errors. *Biometrika*, **74**(4), 725–734.

Benchekroun, K. et Chakravarti, I. M. (1999). Correlation uniform incomplete block designs. *J. Statist. Plann. Inference*, **76**(1-2), 263–272.

Bose, R. C. (1939). On the construction of balanced incomplete block designs. *Ann. Eugenics*, **9**, 353–399.

Bose, R. C. et Bush, K. A. (1952). Orthogonal arrays of strength two and three. *Ann. Math. Statistics*, **23**, 508–524.

Bush, K. A. (1952). Orthogonal arrays of index unity. *Ann. Math. Statistics*, **23**, 426–434.

Chakrabarti, M. C. (1962). *Mathematics of design and analysis of experiments*. Asia Publishing House, Bombay-London-New York.

Chêng, C. (1988). A note on the optimality of semi balanced arrays. in: Dodge, federov, wynn (eds),. *Optimal Design and Analysis of Experiment. North-Holland, Amsterdam,*, pages 115–122.

Chêng, C. S. (1983). Construction of optimal balanced incomplete block designs for correlated observations. *Ann. Statist.*, **11**(1), 240–246.

Cutler, D. R. (1993). Efficient block designs for comparing test treatments to a control when the errors are correlated. *J. Statist. Plann. Inference*, **36**(1), 107–125.

Deheuvels, P. et Dzerko, G. (1991). Block designs for early-stage clinical trials. *International Society for Clinical Biostatistic, Brussels*.

Denes, J. et Keedwell, A. D. (1974). *Latin squares and their applications*. Academic Press, New York.

Dey, A. (1986). *Theory of block designs*. A Halsted Press Book. John Wiley & Sons Inc., New York.

Giesbrecht, F. G. et Gumpertz, M. L. (2004). *Planning, construction, and statistical analysis of comparative experiments*. Wiley Series in Probability and Statistics. Wiley-Interscience [John Wiley & Sons], Hoboken, NJ.

hal-00589585, version 1 - 29 Apr 2011

Hanani, H. (1961). The existence and construction of balanced incomplete block designs. *Ann. Math. Statist.*, **32**, 361–386.

Hedayat, A. S., Sloane, N. J. A., et Stufken, J. (1999). *Orthogonal arrays*. Springer Series in Statistics. Springer-Verlag, New York. Theory and applications, With a foreword by C. R. Rao.

Hinkelmann, K. et Kempthorne, O. (2005). *Design and analysis of experiments. Vol. 2*. Wiley Series in Probability and Statistics. Wiley-Interscience [John Wiley & Sons], Hoboken, NJ. Advanced experimental design.

Ipinyomi, R. A. (1986). Equineighboured experimental designs. *Austral. J. Statist.*, **28**(1), 79–88.

Jacroux, M. (1998). On the construction of efficient equineighboured incomplete block designs having block size 3. *Sankhyā Ser. B*, **60**(3), 488–495.

John, P. W. M. (1980). *Incomplete block designs*, volume 1 of *Lecture Notes in Statistics*. Marcel Dekker Inc., New York.

Kiefer, J. (1975). Balanced block designs and generalized Youden designs. I. Construction (patchwork). *Ann. Statist.*, **3**, 109–118.

Kiefer, J. et Wynn, H. P. (1981). Optimum balanced block and Latin square designs for correlated observations. *Ann. Statist.*, **9**(4), 737–757.

Kramer, E. S., Kreher, D. L., Rees, R., et Stinson, D. R. (1989). On perpendicular arrays with $t \geq 3$. *Ars Combin.*, **28**, 215–223.

Kunert, J. (1985). Optimal repeated measurements designs for correlated observations and analysis by weighted least squares. *Biometrika*, **72**(2), 375–389.

Kunert, J. (1987). Neighbour balanced block designs for correlated errors. *Biometrika*, **74**(4), 717–724.

Lindner, C. C., Mullin, R. C., et Van Rees, G. H. J. (1987). Separable orthogonal arrays. *Utilitas Math.*, **31**, 25–32.

Majumdar, D. et Martin, R. J. (2004). Efficient designs based on orthogonal arrays of type I and type II for experiments using units ordered over time or space. *Stat. Methodol.*, **1**(1-2), 19–35.

Martin, R. J. et Eccleston, J. A. (1991). Optimal incomplete block designs for general dependence structures. *J. Statist. Plann. Inference*, **28**(1), 67–81.

Mathon, R. et Rosa, A. (1985). Tables of parameters of BIBDs with $r \leq 41$ including existence, enumeration, and resolvability results. In *Algorithms in combinatorial design theory*, volume 114 of *North-Holland Math. Stud.*, pages 275–307. North-Holland, Amsterdam.

Mathon, R. et Rosa, A. (1985). $2 - (v, k, \lambda)$ designs of small order. In *The CRC Handbook of Combinatorial Designs,*, C. J. Colbourn and J. H. Dinitz (eds.). Boca Raton, FL: CRC, pages 3–41.

Morgan, J. (1983). *Optimum block design for neighbor type covariance structure.* Ph.D. thesis, Mimeo series #1524, Dept. of Statistics, Univ. of North Carolina, Chapel Hill.

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.

hal-00589585, version 1 - 29 Apr 2011

Raghavarao, D. (1971). *Constructions and combinatorial problems in design of experiments*. John Wiley & Sons Inc., New York. Wiley Series in Probability and Mathematical Statistics.

Rao, C. (1946). Hypercubes. of strength 'd' leading to confounded designs in factorial experiments. *Bull. Calcutta Math. Soc.*, **38**, 67–78.

Rao, C. (1947). Factorial experiments derivable from combinatorial arrangements of arrays. *J. R. S. S., Suppl*, **09**, 128–139.

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.

Rasch, D. et Herrendörfer, G. (1986). *Experimental design*. D. Reidel Publishing Co., Dordrecht. Sample size determination and block designs, Translated from the German.

Russell, K. G. et Eccleston, J. A. (1987a). The construction of optimal balanced incomplete block designs when adjacent observations are correlated. *Austral. J. Statist.*, **29**(1), 84–90.

Shah, K. R. (1960). Optimality criteria for incomplete block designs. *Ann. Math. Statist.*, **31**, 791–794.

Shrikhande, S. S. (1964). Generalized Hadamard matrices and orthogonal arrays of strength two. *Canad. J. Math.*, **16**, 736–740.

Tinsson, W. (2010). *Plans d'expérience: constructions et analyses statistiques*, volume 67 of *Mathématiques & Applications (Berlin) [Mathematics & Applications]*. Springer-Verlag, Berlin.

Uddin, N. (2008). MV-optimal block designs for correlated errors. *Statist. Probab. Lett.*, **78**(17), 2926–2931.

Uddin, N. et Morgan, J. P. (1997). Efficient block designs for settings with spatially correlated errors. *Biometrika*, **84**(2), 443–454.

Yamamoto, S., Kuriki, S., et Sato, M. (1984). On existence and construction of some 2-symbol orthogonal arrays. *TRU Math.*, **20**(2), 317–331.