

Modèles à effets aléatoires et modèles mixtes

Résumé

Un modèle mixte est un modèle comportant à la fois des facteurs à effets fixes, tels qu'ils ont été introduits au chapitre 3, et des facteurs à effets aléatoires, notion nouvelle, un peu particulière, introduite au début de ce chapitre. Les méthodes usuelles dans le modèle linéaire standard, estimations, tests et prévisions, deviennent assez délicates dans le cadre d'un modèle mixte. Certaines sont détaillées dans ce chapitre, d'autres seront simplement évoquées.

Retour au [plan du cours](#)

1 Introduction

Les références bibliographiques les plus importantes sur les modèles linéaires mixtes sont les ouvrages de Miller (1997), Searle *et al.* (1992) et Verbeke & Molenberghs (2000).

Une nouvelle notion est introduite dans ce chapitre : celle de facteur à effets aléatoires. Jusqu'à présent, les facteurs considérés dans les chapitres 3, 4 et 5 étaient des facteurs à effets fixes : les différents niveaux en étaient fixés une fois pour toutes et les effets associés étaient des paramètres à estimer, ces paramètres intervenant dans la moyenne du modèle. Les facteurs à effets aléatoires vont avoir, a priori, une grande quantité de niveaux, les observations réalisées correspondant à un nombre restreint de ces niveaux, pris aléatoirement. On va ainsi modéliser ces niveaux en tant qu'observations d'une variable aléatoire normale, de moyenne nulle (la moyenne du modèle sera définie par les effets fixes) et de variance inconnue, à estimer. Chaque facteur à effets aléatoires sera donc caractérisé par un paramètre de variance qu'il faudra estimer en plus de la variance des erreurs du modèle. D'où le nom de **composantes de la variance** qu'on rencontre également pour de tels modèles.

On appelle modèles mixtes des modèles comportant à la fois des facteurs à effets fixes (ces effets entrant dans la définition de la moyenne du modèle) et

des facteurs à effets aléatoires (ces effets entrant, quant à eux, dans la définition de la variance du modèle). La nécessité d'estimer simultanément plusieurs paramètres de moyenne et plusieurs paramètres de variances dans les modèles mixtes va compliquer la procédure d'estimation. Ainsi, la méthode du maximum de vraisemblance, qui entraîne un biais systématique dans l'estimation de la variance, n'est pas la plus appropriée dans ce cas : on lui préfère, en général, la méthode dite du **maximum de vraisemblance restreint**.

Les tests de significativité des effets aléatoires sont encore des tests de Fisher dans le cas de plans d'expériences équilibrés. Mais, les statistiques de khi-deux intervenant au dénominateur de la statistique de Fisher diffèrent parfois de ce qu'on a vu au chapitre 3, afin de tenir compte de la particularité des modèles mixtes. D'autre part, dans le cas déséquilibré, les tests de Fisher sont en fait des tests approchés et sont d'un usage plus délicat.

2 Modèle à un facteur à effets aléatoires

Les deux premiers paragraphes de ce chapitre sont consacrés à des modèles comportant uniquement des facteurs à effets aléatoires. Ces modèles ont peu d'intérêt dans la pratique ; nous les traitons essentiellement dans un but pédagogique, afin d'introduire de manière progressive les divers constituants des modèles mixtes. Notons que le seul effet fixe d'un modèle à effets aléatoires est l'effet (moyen) général.

2.1 Écriture du modèle pour une observation

Le modèle standard d'analyse de variance (ANOVA) à un seul facteur (à effets fixes) a été défini, au chapitre 3, sous l'une des formes suivantes :

$$Y_{ij} = \beta_j + U_{ij} = \mu + \alpha_j + U_{ij} = m + a_j + U_{ij}.$$

L'écriture de gauche est l'écriture initiale, la suivante correspond au paramétrage centré, la dernière correspond au paramétrage SAS. Dans ces écritures, les différents paramètres intervenant ($\beta_j, \mu, \alpha_j, m, a_j$) sont des réels non aléatoires qu'il convient d'estimer par une méthode appropriée, en général le maximum de vraisemblance. On les appelle parfois des **effets fixes** (c'est-à-dire non aléatoires) et il arrive que l'on parle de **modèle à effets fixes** pour désigner l'ANOVA. Dans un tel modèle, les seules variations aléatoires que l'on envisage sont celles liées à la variable aléatoire réelle (v.a.r.) U_{ij} qui représente

l'erreur du modèle, autrement dit les variations aléatoires non identifiées, non expliquées.

Dans la pratique, il peut se faire que l'on souhaite intégrer dans le modèle des variations aléatoires liées à un phénomène connu ; par exemple, les réactions aléatoires de différents individus à l'absorption d'un médicament donné. Pour ce faire, on intègre une v.a.r. autre que U_{ij} dans le modèle qui devient ainsi un modèle à effets aléatoires. On ne mélangera pas les effets de variance, que l'on va chercher à prendre en compte à travers la v.a.r. en question, avec les effets de moyenne, que l'on continuera à attribuer aux facteurs à effets fixes mis dans le modèle. Ainsi, on modélisera ces effets aléatoires au moyen d'une v.a.r. de moyenne nulle et de variance inconnue, à estimer. Pour des raisons de cohérence, dans le cadre d'un modèle gaussien, on ne considèrera que des v.a.r. normales (gaussiennes), de sorte qu'un modèle à un seul facteur à effets aléatoires (et sans effet fixe autre que l'effet général) s'écrira sous la forme suivante :

$$Y_{ij} = \mu + A_j + U_{ij}$$

(on notera que, compte tenu de la spécificité d'un tel modèle, cette écriture constituera le seul paramétrage considéré ici).

Dans l'écriture ci dessus :

- Y_{ij} est la v.a.r. réponse ;
- μ est un effet fixe, non aléatoire, à estimer (c'est l'effet général, unique effet fixe entrant dans ce modèle) ;
- A_j ($j = 1, \dots, J$) est une v.a.r. $\mathcal{N}(0, \sigma_a^2)$; les différentes v.a.r. A_j sont supposées indépendantes et de même loi ; elles sont donc i.i.d., gaussiennes, centrées, de même variance inconnue ; ainsi, les J niveaux du facteur à effets aléatoires considéré sont J observations indépendantes de cette loi $\mathcal{N}(0, \sigma_a^2)$;
- $U_{ij} \sim \mathcal{N}(0, \sigma^2)$, les U_{ij} étant également i.i.d. ; pour un niveau j fixé du facteur aléatoire, on réalise n_j observations indicées par i , de sorte que l'indice i varie de 1 à n_j et que le nombre total d'observations est $n = \sum_{j=1}^J n_j$;
- on suppose de plus que chaque v.a.r. A_j est indépendante de chaque v.a.r. $U_{ij'}$, $\forall (i, j, j')$.

On déduit ainsi de ce modèle

$$\mathbb{E}(Y_{ij}) = \mu, \text{Var}(Y_{ij}) = \sigma_a^2 + \sigma^2,$$

les deux termes de variance σ_a^2 et σ^2 étant inconnus et devant être estimés, de même que μ .

Les problèmes que nous allons aborder dans ce paragraphe sont l'estimation des paramètres μ , σ_a^2 et σ^2 , le test de significativité du modèle, autrement dit du facteur à effets aléatoires, ainsi que la prédiction des effets $\mu + A_j$.

Remarque. — On notera que les deux termes de variance définis ci-dessus ne dépendent pas des indices i et j : le modèle reste donc homoscédastique.

Remarque. — Pour un même indice j et deux indices i et i' différents, on a :

$$\text{Cov}(Y_{ij}, Y_{i'j}) = \text{Cov}(A_j + U_{ij}, A_j + U_{i'j}) = \sigma_a^2;$$

il n'y a donc pas indépendance entre les deux v.a.r. Y_{ij} et $Y_{i'j}$: ceci est un élément nouveau et très important dans les modèles à effets aléatoires, comme dans les modèles mixtes.

Remarque. — De façon concrète, on décide de considérer un facteur comme étant à effets aléatoires lorsque les J niveaux de ce facteur ne sont pas les seuls qui intéressent l'expérimentateur, mais sont J niveaux pris au hasard dans une population de niveaux quasiment infinie ou, en tout cas, très importante. Le choix de ces niveaux est donc lui-même aléatoire et doit se traduire par l'introduction d'un facteur à effets aléatoires dans le modèle, afin de pouvoir appliquer ce dernier à tout niveau du facteur.

Remarque. — En pratique, comme facteurs à effets aléatoires, on trouve des individus (les "sujets" d'une étude médicale, biologique, génétique...), des animaux (même chose, avec une étude pharmacologique, vétérinaire...), des variétés végétales (étude agronomique ou autre), voire des blocs dans une étude agronomique.

2.2 Écriture matricielle du modèle

Comme indiqué plus haut, on suppose que l'on fait n_j observations (non indépendantes) de la v.a.r. réponse Y au niveau j du facteur aléatoire ; on supposera $n_j \geq 1$, de sorte que l'on ne considèrera ici que des plans complets ; le nombre total d'observations réalisées est noté n ($n = \sum_{j=1}^J n_j$).

Sous forme matricielle, le modèle s'écrit :

$$Y = \mu \mathbf{1}_n + \mathbf{Z}A + U.$$

Dans l'écriture ci-dessus :

- Y et U sont des vecteurs aléatoires de \mathbb{R}^n (que l'on supposera muni de la base canonique) dont les composantes sont, respectivement, les v.a.r. Y_{ij} et U_{ij} ;
- le vecteur $\mathbb{1}_n$ est le vecteur de \mathbb{R}^n dont toutes les composantes sur la base canonique valent 1 ;
- la matrice \mathbf{Z} , de dimension $n \times J$, comporte dans ses colonnes les indicatrices des niveaux du facteur considéré : elle ne contient donc que des 0 et des 1 ;
- enfin, le vecteur

$$A = \begin{pmatrix} A_1 \\ \vdots \\ A_J \end{pmatrix}$$

est un vecteur aléatoire gaussien de \mathbb{R}^J (ici, $p = J$, comme en ANOVA à un seul facteur) : $A \simeq \mathcal{N}_J(0, \sigma_a^2 \mathbf{I}_J)$, où \mathbf{I}_J est la matrice identité d'ordre J .

On a : $\text{Var}(\mathbf{Z}A) = \mathbf{Z}\text{Var}(A)\mathbf{Z}' = \sigma_a^2 \mathbf{Z}\mathbf{Z}'$; on en déduit : $\text{Var}(Y) = \sigma_a^2 \mathbf{Z}\mathbf{Z}' + \sigma^2 \mathbf{I}_n = \mathbf{V}$. On obtient ainsi : $Y \sim \mathcal{N}_n(\mu \mathbb{1}_n, \mathbf{V})$. On dit que σ_a^2 et σ^2 sont les composantes de la variance \mathbf{V} de Y .

EXEMPLE 1 *Considérons le cas très simple dans lequel $J = 2$, $n_1 = 2$ et $n_2 = 3$ (on a donc $n = 5$). Il vient :*

$$\mathbf{Z} = \begin{pmatrix} 1 & 0 \\ 1 & 0 \\ 0 & 1 \\ 0 & 1 \\ 0 & 1 \end{pmatrix}; \quad A = \begin{pmatrix} A_1 \\ A_2 \end{pmatrix};$$

$$\mathbf{V} = \left(\begin{array}{cc|ccc} \sigma_a^2 + \sigma^2 & \sigma_a^2 & 0 & 0 & 0 \\ \sigma_a^2 & \sigma_a^2 + \sigma^2 & 0 & 0 & 0 \\ \hline 0 & 0 & \sigma_a^2 + \sigma^2 & \sigma_a^2 & \sigma_a^2 \\ 0 & 0 & \sigma_a^2 & \sigma_a^2 + \sigma^2 & \sigma_a^2 \\ 0 & 0 & \sigma_a^2 & \sigma_a^2 & \sigma_a^2 + \sigma^2 \end{array} \right).$$

Remarque. — Comme dans le modèle à un seul facteur à effets fixes, on notera que, au sein d'un même niveau du facteur à effets aléatoires, les observations

de ce facteur sont constantes. Seules les observations des v.a.r. U_{ij} changent : $y_{ij} = \mu + a_j + u_{ij}$; $y_{i'j} = \mu + a_j + u_{i'j}$ (a_j est l'observation de la v.a.r. A_j).

2.3 Estimation de la moyenne

2.3.1 Cas général

Le cas général correspond au cas où le plan considéré peut être déséquilibré, autrement dit au cas où les effectifs n_j ne sont pas nécessairement tous égaux. Alors, dans le modèle écrit ci-dessus, la matrice \mathbf{V} des variances-covariances du vecteur aléatoire associé aux observations a une structure irrégulière (ses blocs diagonaux n'ont pas tous la même dimension ; ainsi, dans l'exemple ci-dessus, il y a un bloc carré d'ordre 2 et un autre carré d'ordre 3). Cela entraîne une modification de l'expression usuelle de $\hat{\mu}$, estimation de μ . On notera, d'ailleurs, que cette expression (que nous donnons plus bas) est la même dans le cas gaussien, avec estimation par maximum de vraisemblance, et dans le cas sans hypothèse gaussienne, avec estimation par moindres carrés (au demeurant, dans ce dernier cas, la technique appropriée est la méthode des moindres carrés généralisés, associée à la métrique de Mahalanobis, définie par la matrice \mathbf{V}^{-1}). On se reportera au point 6.3.3 pour plus de détails.

Posons $\bar{Y}_{\bullet j} = \frac{1}{n_j} \sum_{i=1}^{n_j} Y_{ij}$, moyenne empirique des composantes du vecteur aléatoire Y correspondant au niveau j du facteur. On peut vérifier que l'on a :

$$\begin{aligned} \mathbb{E}(\bar{Y}_{\bullet j}) &= \mu; \\ \text{Var}(\bar{Y}_{\bullet j}) &= \frac{1}{n_j^2} \text{Var}\left(\sum_{i=1}^{n_j} Y_{ij}\right) = \frac{\sigma_a^2 + \sigma^2}{n_j} + \frac{n_j(n_j - 1)\sigma_a^2}{n_j^2} \\ &= \sigma_a^2 + \frac{\sigma^2}{n_j} = \tau_j^2. \end{aligned}$$

L'estimation (maximum de vraisemblance ou moindres carrés généralisés) du paramètre de moyen-ne μ est alors :

$$\hat{\mu} = \sum_{j=1}^J w_j \bar{y}_{\bullet j}, \quad \text{avec } w_j = \frac{\frac{1}{\tau_j^2}}{\sum_{j=1}^J \frac{1}{\tau_j^2}}.$$

Remarque. — On notera que le calcul de $\hat{\mu}$ donné ci-dessus nécessite la connaissance des poids w_j , donc des éléments τ_j^2 , autrement dit des composantes de la variance σ_a^2 et σ^2 . Comme ces composantes sont inconnues, elles doivent être estimées avant de pouvoir calculer l'estimation de la moyenne.

Remarque. — La justification de l'expression de $\hat{\mu}$ est donnée en 6.3.3, dans la remarque 71.

2.3.2 Cas équilibré

Dans ce cas, on a $n_j = n_0, \forall j$, et $n = n_0 J$. Les éléments τ_j ne dépendent plus de j et les poids w_j sont tous égaux à $\frac{1}{J}$. On en déduit immédiatement :

$$\hat{\mu} = \frac{1}{n} \sum_{j=1}^J \sum_{i=1}^{n_0} y_{ij} = \bar{y}_{\bullet\bullet} \quad (\text{moyenne générale des observations } y_{ij}).$$

On voit que, dans ce cas, le calcul effectif de $\hat{\mu}$ peut se faire avant d'estimer les composantes de la variance.

2.4 Estimation des composantes de la variance

Dans les modèles à effets aléatoires, il existe différentes méthodes d'estimation des composantes de la variance (les paramètres σ_a^2 et σ^2 dans le cas d'un seul facteur). Dans les cas simples (en particulier celui d'un plan équilibré), les différentes méthodes peuvent être équivalentes, mais ce n'est plus vrai dans le cas général. De plus, il n'y a pas de méthode qui soit uniformément meilleure que les autres, d'où les difficultés.

2.4.1 Estimation par ANOVA

Indiquons tout de suite que cette appellation n'est pas vraiment appropriée, même si c'est la plus couramment utilisée dans la littérature statistique (raison pour laquelle nous l'utilisons). On trouve également l'appellation de "méthode de type I" dans le logiciel SAS, ce qui est tout aussi peu approprié. Il s'agit en fait d'une méthode d'estimation de type moments : on écrit un système d'équations en égalant moments empiriques et moments théoriques (d'ordre 1) de certaines sommes de carrés ; la solution de ce système fournit des valeurs pour les composantes de la variance et ces valeurs sont prises comme estimations

de ces composantes. On pourrait donc encore l'appeler "méthode par sommes de carrés", ou "méthode par formes quadratiques", mais nous utiliserons plutôt l'appellation usuelle.

Considérons donc les deux sommes de carrés définies ci-dessous.

– SSA est la somme des carrés des écarts pris en compte par le modèle, c'est-à-dire par le facteur aléatoire A . On a :

$$SSA = \sum_{j=1}^J n_j (\bar{Y}_{\bullet j} - \bar{Y}_{\bullet\bullet})^2.$$

Le calcul de l'espérance mathématique de SSA , un peu lourd, conduit au résultat suivant :

$$\mathbb{E}(SSA) = (J-1)\sigma^2 + \frac{1}{n}(n^2 - \sum_{j=1}^J n_j^2)\sigma_a^2.$$

On notera que, dans le cas équilibré ($n_j = n_0, \forall j$), cette espérance se simplifie et devient :

$$\mathbb{E}(SSA) = (J-1)(\sigma^2 + n_0\sigma_a^2).$$

Par ailleurs, les solutions du système que l'on va écrire feront intervenir le carré moyen relatif au facteur A , qui sera noté MSA : $MSA = \frac{1}{J-1}SSA$ (SSA est une quantité à $J-1$ degrés de liberté). On obtient :

$$\mathbb{E}(MSA) = \sigma^2 + \frac{1}{n(J-1)}(n^2 - \sum_{j=1}^J n_j^2)\sigma_a^2,$$

qui devient, dans le cas équilibré :

$$\mathbb{E}(MSA) = \sigma^2 + n_0\sigma_a^2.$$

– SSE est la somme des carrés des écarts résiduels (ceux associés aux erreurs du modèle, c'est-à-dire non pris en compte par le modèle) :

$$SSE = \sum_{j=1}^J \sum_{i=1}^{n_j} (Y_{ij} - \bar{Y}_{\bullet j})^2.$$

L'espérance de SSE s'écrit :

$$\mathbb{E}(SSE) = (n - J)\sigma^2.$$

On utilisera encore le carré moyen relatif aux erreurs, noté MSE :

$$MSE = \frac{SSE}{n - J} \quad (SSE \text{ est à } n - J \text{ degrés de liberté}). \text{ Son espérance vaut } \mathbb{E}(MSE) = \sigma^2.$$

On définit ensuite un système de 2 équations à 2 inconnues en égalant les espérances mathématiques de MSA et MSE avec les valeurs observées de ces 2 carrés moyens qui seront respectivement notées $MSA(y)$ et $MSE(y)$. Ce système est le suivant :

$$\begin{aligned} MSA(y) &= \sigma^2 + \frac{1}{n(J-1)} \left(n^2 - \sum_{j=1}^J n_j^2 \right) \sigma_a^2; \\ MSE(y) &= \sigma^2. \end{aligned}$$

Sa résolution (immédiate) fournit les estimations ANOVA des composantes de la variance. L'estimation de la composante relative au facteur aléatoire A sera indiquée par 1, afin de ne pas la confondre avec les autres estimations obtenues plus loin.

$$\begin{aligned} \hat{\sigma}^2 &= MSE(y); \\ \hat{\sigma}_{a1}^2 &= \frac{n(J-1)[MSA(y) - MSE(y)]}{n^2 - \sum_{j=1}^J n_j^2}. \end{aligned}$$

Par la suite, nous noterons $\hat{\Sigma}^2$ la statistique (la v.a.r.) dont $\hat{\sigma}^2$ est l'observation ($\hat{\Sigma}^2 = MSE$) et $\hat{\Sigma}_{a1}^2$ la statistique dont $\hat{\sigma}_{a1}^2$ est l'observation :

$$\hat{\Sigma}_{a1}^2 = \frac{n(J-1)[MSA - MSE]}{n^2 - \sum_{j=1}^J n_j^2}$$

(dans les expressions ci-dessus, $\hat{\Sigma}^2$ et $\hat{\Sigma}_{a1}^2$ sont les estimateurs, tandis que $\hat{\sigma}^2$ et $\hat{\sigma}_{a1}^2$ sont les estimations).

Propriétés des estimateurs ANOVA

- On a nécessairement : $\hat{\sigma}^2 \geq 0$. Par contre, le signe de $\hat{\sigma}_{a1}^2$ est celui de $MSA(y) - MSE(y)$ et peut être négatif. Dans ce cas, par convention, on posera $\hat{\sigma}_{a1}^2 = 0$, ce qui revient à considérer que chaque niveau A_j est presque sûrement nul (comme observation d'une loi normale de moyenne et de variance nulles). Le facteur aléatoire est donc sans effet et on doit, dans ce cas, le retirer du modèle.
- On peut montrer la relation suivante :

$$\frac{(n - J)\hat{\Sigma}^2}{\sigma^2} = \frac{SSE}{\sigma^2} \sim \chi_{n-J}^2.$$

On en déduit immédiatement que $\hat{\Sigma}^2$ est un estimateur sans biais et convergent de σ^2 .

- Par ailleurs, on vérifie sans difficulté que $\hat{\Sigma}_{a1}^2$ est un estimateur sans biais de σ_a^2 .
- Lorsque le plan considéré est équilibré, l'expression de $\hat{\Sigma}_{a1}^2$ se simplifie et devient :

$$\hat{\Sigma}_{a1}^2 = \frac{MSA - MSE}{n_0}.$$

- Toujours dans le cas d'un plan équilibré, on peut aussi montrer la relation suivante :

$$\frac{SSA}{n_0\sigma_a^2 + \sigma^2} \sim \chi_{J-1}^2$$

(voir, par exemple, Searle *et al.*, 1992). Par contre, il n'y a pas de résultat de ce type dans le cas déséquilibré.

- Les statistiques SSA et SSE sont indépendantes (que le plan soit équilibré ou non ; on pourra encore se reporter à Searle *et al.*, 1992).
- Par contre, $\hat{\Sigma}_{a1}^2$ et $\hat{\Sigma}^2$ ne sont pas indépendantes (c'est immédiat, compte tenu de l'expression de $\hat{\Sigma}_{a1}^2$).
- On n'a pas de résultat de convergence concernant $\hat{\Sigma}_{a1}^2$, qu'on soit dans le cas équilibré ou déséquilibré.
- Enfin, on sait écrire $\text{Var}(\hat{\Sigma}_{a1}^2)$ et $\text{Cov}(\hat{\Sigma}_{a1}^2, \hat{\Sigma}^2)$ dans tous les cas, mais les expressions sont compliquées, surtout dans le cas déséquilibré.

2.4.2 Estimation par maximum de vraisemblance

La vraisemblance de l'échantillon s'écrit :

$$L(y, \mu, \sigma_a^2, \sigma^2) = \frac{1}{(2\pi)^{n/2} (\det \mathbf{V})^{1/2}} \exp\left[-\frac{1}{2}(y - \mu \mathbb{1}_n)' \mathbf{V}^{-1} (y - \mu \mathbb{1}_n)\right],$$

où \mathbf{V} a été définie en 6.1.2 et vaut : $\mathbf{V} = \sigma_a^2 \mathbf{Z}\mathbf{Z}' + \sigma^2 \mathbf{I}_n$.

La log-vraisemblance s'écrit donc :

$$\begin{aligned} l(y, \mu, \sigma_a^2, \sigma^2) &= \log[L(y, \mu, \sigma_a^2, \sigma^2)] \\ &= \text{constante} - \frac{1}{2} \log(\det \mathbf{V}) - \frac{1}{2} (y - \mu \mathbf{1}_n)' \mathbf{V}^{-1} (y - \mu \mathbf{1}_n). \end{aligned}$$

Il est possible d'expliciter $\det \mathbf{V}$, ainsi que \mathbf{V}^{-1} , en fonction de σ_a^2 et σ^2 . On peut alors dériver $l(y, \mu, \sigma_a^2, \sigma^2)$ selon ces 2 variables et obtenir les équations de vraisemblance. Toutefois, les solutions de ces équations ne sont explicites que dans le cas équilibré.

Cas équilibré

La résolution des équations de vraisemblance conduit aux expressions suivantes :

$$\hat{\Sigma}^2 = MSE; \quad \hat{\Sigma}_{a2}^2 = \frac{1}{n} [(1 - \frac{1}{J})MSA - MSE].$$

L'expression de $\hat{\Sigma}^2$ est la même que celle obtenue par ANOVA ; il en va donc de même pour ses propriétés. Par contre, ce n'est pas le cas pour l'expression ci-dessus de $\hat{\Sigma}_{a2}^2$ qui est différente de celle obtenue par ANOVA dans le cas équilibré ($\hat{\Sigma}_{a1}^2$). Il est encore possible que sa valeur calculée $\hat{\sigma}_{a2}^2$ soit négative ; comme précédemment, elle est alors ramenée à 0. Le problème majeur de cet estimateur $\hat{\Sigma}_{a2}^2$ est qu'il est biaisé pour σ_a^2 , sans qu'on puisse en corriger le biais. En effet, on peut vérifier sans difficulté : $\mathbb{E}(\hat{\Sigma}_{a2}^2) = \frac{1}{n_0 J^2} [n_0 (J-1) \sigma_a^2 - \sigma^2]$. Enfin, signalons qu'on sait encore expliciter dans ce cas la matrice des variances-covariances du couple $(\hat{\Sigma}^2, \hat{\Sigma}_{a2}^2)$.

Cas déséquilibré

Dans ce cas, il n'y a pas de solution analytique des équations de vraisemblance et l'on a recours à une méthode numérique de résolution d'un système non linéaire (méthode itérative, de type Fisher scoring). La solution $\hat{\Sigma}_{a2}^2$ obtenue dans ce cas est, en général, encore biaisée, sans que l'on puisse écrire explicitement le biais, et encore moins le corriger.

Remarque. — Le caractère biaisé de l'estimateur maximum de vraisemblance $\hat{\Sigma}_{a2}^2$ vient de ce que la log-vraisemblance donnée plus haut ne sépare pas le

paramètre de moyenne (μ) des paramètres de variance (σ_a^2 et σ^2). Toutefois, il est possible de faire une décomposition appropriée de la vraisemblance : c'est le principe de la méthode du maximum de vraisemblance restreint.

2.4.3 Estimation par maximum de vraisemblance restreint

Il est donc possible de factoriser la vraisemblance $L(y, \mu, \sigma_a^2, \sigma^2)$ selon deux termes : le premier contient μ , et sa maximisation conduit à l'expression de $\hat{\mu}$ donnée en 6.1.3 ; le second ne dépend que de σ_a^2 et de σ^2 . La méthode du maximum de vraisemblance restreint consiste à maximiser le logarithme de ce seul second terme pour obtenir une estimation des composantes de la variance. Cette méthode, introduite par Patterson & Thomson (1971), est maintenant connue sous l'acronyme de REML (pour *REstricted -ou RESidual- Maximum Likelihood*, appellation introduite par Harville (1977) ; on trouve aussi les termes de vraisemblance résiduelle et de vraisemblance marginale). Elle fournit, en général, des estimateurs des composantes de la variance ayant de meilleures propriétés que les estimateurs maximum de vraisemblance. Toutefois, comme pour ces derniers, il n'y a de solution explicite que dans le cas équilibré. On trouvera plus de détails sur cette méthode en 6.3.3.

Cas équilibré

La résolution des équations, obtenues par dérivation de la partie de la log-vraisemblance ne dépendant que des termes de variance, conduit aux expressions suivantes :

$$\hat{\Sigma}^2 = MSE; \quad \hat{\Sigma}_{a3}^2 = \frac{MSA - MSE}{n_0}.$$

Notons tout d'abord que l'expression de $\hat{\Sigma}^2$ est la même que dans les deux cas précédents. Par ailleurs, dans le cas équilibré, on notera que : $\hat{\Sigma}_{a3}^2 = \hat{\Sigma}_{a1}^2 \neq \hat{\Sigma}_{a2}^2$. Les propriétés de $\hat{\Sigma}_{a3}^2$ sont donc les mêmes que celles de $\hat{\Sigma}_{a1}^2$; en particulier, il s'agit d'un estimateur sans biais. Bien sûr, sa valeur calculée $\hat{\sigma}_{a3}^2$ peut être négative et est alors ramenée à 0.

Cas déséquilibré

Qu'on utilise la vraisemblance complète ou la vraisemblance restreinte, les équations obtenues n'ont pas de solution analytique dans le cas déséquilibré. On est donc conduit, comme précédemment, à utiliser un algorithme de résolution numérique. Bien sûr, on n'a pas de résultat sur le biais des estimateurs, mais des simulations ont mis en évidence que les estimateurs obtenus par maximum de vraisemblance restreint sont, en général, "meilleurs" que ceux obtenus par maximum de vraisemblance.

2.4.4 Estimations MINQUE et MIVQUE

Avant que les ordinateurs ne fournissent des solutions commodes au problème de la résolution numérique d'un système d'équations non linéaires sans solution analytique, un certain nombre de statisticiens ont cherché à obtenir, de façon relativement simple, des solutions "raisonnables", à défaut d'être optimales, pour les estimateurs des composantes de la variance dans les modèles à effets aléatoires et les modèles mixtes. Parmi eux, C.R. Rao a publié, dans les années 1970–72, quatre articles sur le sujet (voir la bibliographie) proposant deux classes d'estimateurs : les estimateurs MINQUE et MIVQUE. Nous donnons, dans ce paragraphe, quelques éléments succincts sur ces estimateurs aujourd'hui peu utilisés. Notons toutefois que le logiciel SAS utilise des estimations MIVQUE pour initialiser l'algorithme REML.

L'idée générale est d'estimer toute combinaison linéaire des composantes de la variance, du type $c_1\sigma_a^2 + c_2\sigma^2$, par une forme quadratique des observations y_{ij} (en remarquant que tout estimateur d'une variance fait intervenir, d'une façon ou d'une autre, des sommes de carrés des observations, ce qui est aussi l'idée de l'estimation par ANOVA). Si \mathbf{Q} désigne une matrice réelle, carrée d'ordre n et symétrique, on prend donc pour estimation de $c_1\sigma_a^2 + c_2\sigma^2$ la forme quadratique $y'\mathbf{Q}y$. On impose alors à la matrice \mathbf{Q} diverses propriétés. Tout d'abord, elle doit être telle que $y'\mathbf{Q}y$ soit invariant par translation sur μ (autrement dit, l'estimation de la variance ne doit pas dépendre de la moyenne) ; ensuite, l'estimation fournie doit être sans biais (pour $c_1\sigma_a^2 + c_2\sigma^2$). Ceci ne suffisant pas pour obtenir un estimateur unique, on doit imposer une autre condition d'optimalité. Selon la condition choisie, on obtient l'une des deux classes précisées ci-dessous.

Si l'on impose, en plus, à $y'\mathbf{Q}y$ de minimiser une certaine norme, on obtient

un estimateur appelé *Minimum Norm Quadratic Unbiased Estimator*, d'où l'acronyme MINQUE (ce fut la première méthode proposée ; en général, on lui préfère la suivante).

Si l'on impose maintenant à $y'\mathbf{Q}y$ d'être de variance minimum, on obtient alors un estimateur appelé *Minimum Variance Quadratic Unbiased Estimator*, autrement dit MIVQUE (cette seconde propriété est plus naturelle pour un estimateur).

Concrètement, ces deux procédures fonctionnent en une seule itération : on choisit tout d'abord une valeur initiale pour chaque composante de la variance ; on met ensuite à jour ces composantes par un calcul unique faisant intervenir \mathbf{Q} et y . Par construction, les estimateurs ainsi définis sont sans biais, mais ils peuvent encore conduire à des solutions négatives. D'autre part, il est clair qu'ils dépendent du choix de la solution initiale.

Remarque. — Dans le cas équilibré, on notera tout d'abord que toutes les méthodes exposées ci-dessus conduisent à la même estimation de la composante résiduelle σ^2 de la variance :

$$\hat{\Sigma}^2 = MSE.$$

D'autre part, en ce qui concerne l'autre composante σ_a^2 , celle liée au facteur aléatoire A , les estimations par ANOVA, par maximum de vraisemblance restreint, ainsi que les estimations MINQUE et MIVQUE, conduisent au même résultat (toujours dans le cas équilibré) :

$$\hat{\Sigma}_{a1}^2 = \frac{MSA - MSE}{n_0}.$$

Par contre, la solution fournie par le maximum de vraisemblance est différente :

$$\hat{\Sigma}_{a2}^2 = \frac{1}{n_0 J} \left[\left(1 - \frac{1}{J}\right) MSA - MSE \right].$$

Cela met bien en évidence la faiblesse du maximum de vraisemblance dans ce cas : c'est la seule méthode conduisant à une estimation biaisée.

Remarque. — Dans le cas déséquilibré, les méthodes du maximum de vraisemblance et du maximum de vraisemblance restreint nécessitent l'usage d'un algorithme itératif, donc le choix de solutions initiales dont dépend, bien sûr, les résultats obtenus. Il est souvent conseillé, et c'est le choix que nous

préconisons, de prendre dans ce cas l'estimation par ANOVA comme solution initiale.

Il faut toutefois noter que ce n'est pas le choix du logiciel statistique SAS qui, dans les procédures VARCOMP et MIXED, utilise comme solution initiale une solution particulière de la méthode MIVQUE, appelée MIVQUE(0). La méthode MIVQUE nécessitant elle-même le choix de valeurs initiales (et opérant ensuite en une seule étape), on appelle MIVQUE(0) la méthode obtenue en prenant 1 pour valeur initiale de σ^2 et 0 pour valeur initiale de σ_a^2 .

2.5 Intervalles de confiance

2.5.1 Intervalle de confiance pour μ

Dans le cas équilibré, on peut construire un intervalle de confiance pour μ de la même façon que dans une ANOVA à un facteur (à partir de la loi de Student). En effet, on a dans ce cas

$$\hat{\mu} = \bar{Y}_{\bullet\bullet} \sim \mathcal{N}\left(\mu, \frac{n_0\sigma_a^2 + \sigma^2}{n}\right)$$

et, en utilisant le fait que $\frac{SSA}{n_0\sigma_a^2 + \sigma^2} \sim \chi_{J-1}^2$ (indépendant de $\hat{\mu}$), on obtient un intervalle de confiance, de type Student, de la forme :

$$\bar{Y}_{\bullet\bullet} \pm t \sqrt{\frac{MSA}{n}}.$$

Dans l'expression ci-dessus, t est le quantile d'ordre $1 - \frac{\alpha}{2}$ d'une loi de Student

à $J - 1$ degrés de liberté, $MSA = \frac{1}{J-1} \sum_{j=1}^J n_j (\bar{Y}_{\bullet j} - \bar{Y}_{\bullet\bullet})^2$ et $n = n_0 J$.

Dans le cas déséquilibré, les choses sont plus délicates. En effet, dans ce cas, il vient :

$$\hat{\mu} = \sum_{j=1}^J w_j \bar{Y}_{\bullet j} \sim \mathcal{N}(\mu, v^2), \text{ avec : } v^2 = \frac{1}{\sum_{j=1}^J \frac{n_j}{n_j \sigma_a^2 + \sigma^2}}.$$

En utilisant les estimations des paramètres de variances σ_a^2 et σ^2 , on peut construire un intervalle approximatif à partir de la loi normale. Toutefois, on

n'est même pas sûr que le risque asymptotique de cet intervalle soit α , compte tenu qu'on n'est pas sûr de la convergence de l'estimateur $\hat{\Sigma}_a^2$.

2.5.2 Intervalle de confiance pour σ^2

Qu'on soit dans le cas équilibré ou déséquilibré, on a vu plus haut que : $\frac{SSE}{\sigma^2} \sim \chi_{n-J}^2$. En utilisant les quantiles appropriés de la loi de khi-deux, on en déduit donc un intervalle de confiance exact (non asymptotique) pour le paramètre σ^2 .

2.5.3 Intervalle de confiance pour une fonction des deux paramètres de variance

Dans le cas équilibré, on peut construire des intervalles de confiance exacts pour les rapports $\frac{\sigma_a^2}{\sigma_a^2 + \sigma^2}$, $\frac{\sigma^2}{\sigma_a^2 + \sigma^2}$ et $\frac{\sigma_a^2}{\sigma^2}$, à partir des lois de Fisher appropriées. Toutefois, en pratique, ces intervalles ne sont pas d'un grand intérêt. Par ailleurs, ils ne sont valables que dans le cas équilibré.

2.5.4 Intervalle de confiance pour σ_a^2

Il n'existe pas d'intervalle de confiance exact pour le paramètre σ_a^2 . On peut juste obtenir un intervalle approché dans le cas équilibré, mais nous ne le détaillons pas ici. Nous renvoyons encore une fois à Searle *et al.* (1992) pour plus de détails.

2.6 Test de l'effet du facteur

Tester la significativité du facteur aléatoire A , supposé de loi $\mathcal{N}(0, \sigma_a^2)$, revient à tester $\{H_0 : \sigma_a^2 = 0\}$ contre $\{H_1 : \sigma_a^2 > 0\}$, avec un niveau α fixé (cela revient encore, dans le cas d'un seul facteur, à tester la significativité du modèle).

Sous l'hypothèse nulle H_0 , qu'on soit dans le cas équilibré ou déséquilibré, c'est le modèle constant qui est considéré et la statistique $F = \frac{MSA}{MSE}$ est distribuée selon une loi de Fisher à $(J-1)$ et $(n-J)$ degrés de liberté, comme dans le cas d'un facteur à effets fixes. On peut donc utiliser le test de Fisher standard pour tester H_0 .

Par contre, sous H_1 , la distribution de F est différente selon que l'on a

affaire à un facteur à effets fixes, à un facteur à effets aléatoires avec un plan équilibré, ou encore à un facteur à effets aléatoires avec un plan déséquilibré. C'est donc la détermination de la puissance de ce test qui est délicate (nous ne l'aborderons pas dans ce cours).

Remarque. — Il est important de bien voir que le test de significativité du facteur, dans le cas d'un seul facteur, est le même que ce facteur soit à effets fixes ou à effets aléatoires. On ne peut donc pas espérer s'en servir pour choisir entre les deux types d'effets. Redisons ici que ce choix dépend des conditions expérimentales et absolument pas d'une quelconque technique statistique.

2.7 Prédiction d'un effet aléatoire

Il convient de faire attention au fait que la prédiction de la v.a.r. Y , lorsqu'elle n'est pas observée mais qu'on connaît le niveau du facteur A auquel est réalisée son observation, est assez délicate. Soit j le niveau en question et $Y_{\ell j}$ la v.a.r. correspondante à prédire. On dispose de $\mathbb{E}(Y_{\ell j}) = \mu$, mais une prédiction par $\hat{\mu}$ ne tient pas compte du niveau auquel on se trouve et ne convient donc pas. Par ailleurs, $\bar{y}_{\bullet j}$ correspond à la prédiction qu'on ferait dans le cas d'un facteur à effets fixes et ne convient pas davantage, pas plus que $\mathbb{E}(A_j) = 0$.

La solution consiste en fait à prévoir $Y_{\ell j}$ par $\hat{\mu} + \mathbb{E}(A_j / \bar{Y}_{\bullet j} = \bar{y}_{\bullet j})$. Autrement dit, on fait intervenir l'espérance conditionnelle de la v.a.r. A_j , sachant que la v.a.r. $\bar{Y}_{\bullet j}$ prend la valeur $\bar{y}_{\bullet j}$, moyenne observée, au sein de l'échantillon, à ce niveau du facteur.

On sait que $A_j \sim \mathcal{N}(0, \sigma_a^2)$ et que $Y_{ij} \sim \mathcal{N}(\mu, \sigma_a^2 + \sigma^2)$. On a vu en 6.1.3 que $\bar{Y}_{\bullet j} \sim \mathcal{N}(\mu, \sigma_a^2 + \frac{\sigma^2}{n_j})$ et on vérifie sans difficulté que $\text{Cov}(A_j, \bar{Y}_{\bullet j}) = \sigma_a^2$ (dans les calculs, on doit faire attention au fait que Y_{ij} et $Y_{i'j}$ ne sont pas indépendantes). Cela permet d'écrire la loi conjointe des v.a.r. A_j et $\bar{Y}_{\bullet j}$:

$$\mathcal{N}_2 \left(\left(\begin{array}{c} 0 \\ \mu \end{array} \right); \left(\begin{array}{cc} \sigma_a^2 & \sigma_a^2 \\ \sigma_a^2 & \sigma_a^2 + \frac{\sigma^2}{n_j} \end{array} \right) \right).$$

Les règles usuelles du calcul de l'espérance conditionnelle d'une composante de loi normale bidimensionnelle sachant la seconde composante permettent

d'écrire

$$\mathbb{E}(A_j / \bar{Y}_{\bullet j} = \bar{y}_{\bullet j}) = 0 + (\bar{y}_{\bullet j} - \mu) \frac{\sigma_a^2}{\sigma_a^2 + \frac{\sigma^2}{n_j}} = (\bar{y}_{\bullet j} - \mu) \frac{\sigma_a^2}{\tau_j^2},$$

qui sera "estimée" par $(\bar{y}_{\bullet j} - \hat{\mu}) \frac{\hat{\sigma}_a^2}{\hat{\tau}_j^2}$. Finalement, la prédiction de $Y_{\ell j}$ sera donnée par :

$$\hat{Y}_{\ell j} = \hat{\mu} + (\bar{y}_{\bullet j} - \hat{\mu}) \frac{\hat{\sigma}_a^2}{\hat{\tau}_j^2}.$$

2.8 Illustration

Il s'agit d'un exemple fictif à un seul facteur aléatoire, ce facteur comportant 4 niveaux. Les niveaux, notés 1, 2, 3 et 4, figurent en première colonne du fichier des données. La variable réponse figure dans la colonne suivante. Le plan est complet, avec répétitions, déséquilibré. Il y a 13 individus observés, donc 13 lignes dans le fichier des données reproduit ci-dessous.

Les données

1	9
1	10
1	11
1	12
2	15
2	16
2	17
3	13
3	13
3	14
3	15
4	25
4	28

Le programme SAS

Le programme SAS ci-dessous permet de traiter ces données en tenant compte du caractère aléatoire du facteur. Outre la procédure GLM, il est ici

nécessaire d'utiliser la procédure VARCOMP pour estimer les composantes de la variance (σ_a^2 et σ^2). On notera que la méthode d'estimation doit être spécifiée, la méthode par défaut étant MIVQUE (0).

```
* ----- ;
* options facultatives pour la mise en page des sorties ;
* ----- ;
options linesize=76 pagesize=64 nodate;
title;
footnote 'Effets aleatoires - Exemple fictif';
* ----- ;
*          lecture des donnees ;
* (le fichier "fic.don" contient les donnees ;
* et se trouve dans le repertoire de travail) ;
* ----- ;
data fic;
infile 'fic.don';
input effet $ reponse;
run;
* ----- ;
*          procedure GLM ;
* ----- ;
proc glm data=fic;
class effet;
model reponse = effet / ss3;
random effet;
run;
quit;
* ----- ;
*          procedure VARCOMP ;
* (chacune des 4 options est utilisee successivement) ;
* ----- ;
proc varcomp data=fic method=typel;
class effet;
model reponse = effet;
run;
* ----- ;
proc varcomp data=fic method=mivque0;
class effet;
model reponse = effet;
```

```
run;
* ----- ;

proc varcomp data=fic method=ml;
class effet;
model reponse = effet;
run;
* ----- ;
proc varcomp data=fic method=reml;
class effet;
model reponse = effet;
run;
quit;
```

Les sorties de la procédure GLM

```
PAGE 1
-----
                        The GLM Procedure

                        Class Level Information

Class          Levels   Values
effet          4        1 2 3 4

Number of observations      13

PAGE 2
-----
                        The GLM Procedure

Dependent Variable: reponse

Source          DF          Sum of Squares      Mean Square      F Value      Pr > F
Model           3          354.0576923        118.0192308       74.54      <.0001
Error           9           14.2500000          1.5833333
Corrected Total 12          368.3076923

R-Square      Coeff Var      Root MSE      reponse Mean
0.961310      8.261603       1.258306       15.23077

PAGE 3
-----
                        The GLM Procedure
```

```
Source          Type III Expected Mean Square      effet          4      1 2 3 4
effet          Var(Error) + 3.1795 Var(effet)      Number of observations      13
```

Les sorties de la procédure VARCOMP

PAGE 1 Variance Components Estimation Procedure

```
Class Level Information
Class      Levels  Values
effet      4      1 2 3 4

Number of observations      13
Dependent Variable:      reponse
```

Type 1 Analysis of Variance

Source	DF	Sum of Squares	Mean Square
effet	3	354.057692	118.019231
Error	9	14.250000	1.583333
Corrected Total	12	368.307692	.

Type 1 Analysis of Variance

```
Source          Expected Mean Square
effet          Var(Error) + 3.1795 Var(effet)
Error          Var(Error)
Corrected Total .
```

Type 1 Estimates

```
Variance Component      Estimate
Var(effet)              36.62097
Var(Error)              1.58333
```

PAGE 2 Variance Components Estimation Procedure

```
Class Level Information
Class      Levels  Values
```

MIVQUE(0) SSQ Matrix

Source	effet	Error	reponse
effet	31.90533	9.53846	906.47337
Error	9.53846	12.00000	368.30769

MIVQUE(0) Estimates

```
Variance Component      reponse
Var(effet)              25.23143
Var(Error)              10.63656
```

PAGE 3

Variance Components Estimation Procedure

Class Level Information

```
Class      Levels  Values
effet      4      1 2 3 4

Number of observations      13
Dependent Variable:      reponse
```

Maximum Likelihood Iterations

Iteration	Objective	Var(effet)	Var(Error)
0	29.6188217655	12.0015563076	5.0593729821
1	23.0175866155	33.3583321167	1.6232181526
2	23.0101144107	35.1435488604	1.5855418629
3	23.0101132719	35.1211357555	1.5859817377
4	23.0101132718	35.1208834116	1.5859866946

Convergence criteria met.

Maximum Likelihood Estimates

```
Variance Component      Estimate
```

```

Var(effet)      35.12088
Var(Error)     1.58599
    
```

```

Var(effet)      1520.6      -0.28093
Var(Error)     -0.28093      0.55920
    
```

Asymptotic Covariance Matrix of Estimates

```

                Var(effet)      Var(Error)
Var(effet)     640.34038      -0.28152
Var(Error)     -0.28152      0.56084
    
```

PAGE 4

Variance Components Estimation Procedure

```

Class Level Information
Class      Levels  Values
effet      4      1 2 3 4

Number of observations      13
Dependent Variable:      reponse
    
```

REML Iterations

Iteration	Objective	Var(effet)	Var(Error)
0	26.7855408891	13.0016859999	5.4809873972
1	19.3392040165	34.5534023567	1.7968792220
2	19.1250010581	49.6309606128	1.5586243312
3	19.1198401813	47.1506894694	1.5841973236
4	19.1198374354	47.0942665176	1.5848199498
5	19.1198374354	47.0942665176	1.5848199498

Convergence criteria met.

REML Estimates

```

Variance Component      Estimate
Var(effet)              47.09427
Var(Error)              1.58482
    
```

Asymptotic Covariance Matrix of Estimates

```

                Var(effet)      Var(Error)
    
```

3 Modèle à deux facteurs croisés à effets aléatoires

On suppose maintenant que la v.a.r. réponse Y dépend de deux facteurs à effets aléatoires notés A et B et, éventuellement, de leur interaction qui sera notée C . On note encore J le nombre de niveaux observés de A ($J \geq 2$), ces niveaux étant indicés par j , et K le nombre de niveaux observés de B ($K \geq 2$), ces niveaux étant maintenant indicés par k . Les deux facteurs A et B sont croisés et on note n_{jk} ($n_{jk} \geq 1$) le nombre d'observations réalisées dans la cellule (j, k) du plan obtenu par ce croisement. Enfin, on pose $n = \sum_{j=1}^J \sum_{k=1}^K n_{jk}$: n est le nombre total d'observations réalisées.

3.1 Écritures du modèle

Pour une observation Y_{ijk} de la v.a.r. réponse Y , le modèle s'écrit :

$$Y_{ijk} = \mu + A_j + B_k + C_{jk} + U_{ijk}.$$

- Comme précédemment, μ est l'effet général ; c'est l'unique effet fixe de ce modèle.
- A_j est l'effet aléatoire du niveau j du facteur A ($j = 1, \dots, J$) ; on suppose : $A_j \sim \mathcal{N}(0, \sigma_a^2)$.
- B_k est l'effet aléatoire du niveau k du facteur B ($k = 1, \dots, K$) ; on suppose de même : $B_k \sim \mathcal{N}(0, \sigma_b^2)$.
- C_{jk} est l'effet aléatoire de l'interaction de A et de B dans la cellule (j, k) (l'interaction entre deux facteurs à effets aléatoires est nécessairement elle-même à effets aléatoires) ; on suppose maintenant : $C_{jk} \sim \mathcal{N}(0, \sigma_c^2)$.
- U_{ijk} est la v.a.r. erreur du modèle et l'on suppose : $U_{ijk} \sim \mathcal{N}(0, \sigma^2)$.

Enfin, les v.a.r. A_j , B_k , C_{jk} et U_{ijk} sont supposées mutuellement indépendantes.

On peut réécrire le modèle sous la forme matricielle suivante :

$$Y = \mu\kappa_n + \mathbf{Z}_1 \begin{pmatrix} A_1 \\ \vdots \\ A_J \end{pmatrix} + \mathbf{Z}_2 \begin{pmatrix} B_1 \\ \vdots \\ B_K \end{pmatrix} + \mathbf{Z}_3 \begin{pmatrix} C_1 \\ \vdots \\ C_{JK} \end{pmatrix} + U$$

$$= \mu\kappa_n + \mathbf{Z}_1 A + \mathbf{Z}_2 B + \mathbf{Z}_3 C + U.$$

Dans l'écriture ci-dessus, Y et U sont des vecteurs aléatoires de \mathbb{R}^n et κ_n est le vecteur dont toutes les coordonnées (sur la base canonique de \mathbb{R}^n) sont égales à 1. D'autre part, \mathbf{Z}_1 (respectivement \mathbf{Z}_2 , resp. \mathbf{Z}_3) est la matrice des indicatrices des niveaux de A (resp. de B , resp. des cellules), de dimension $n \times J$ (resp. $n \times K$, resp. $n \times JK$).

La loi de probabilité du vecteur aléatoire Y a maintenant pour expression

$Y \sim \mathcal{N}_n(\mu\kappa_n; \sigma_a^2 \mathbf{Z}_1 \mathbf{Z}'_1 + \sigma_b^2 \mathbf{Z}_2 \mathbf{Z}'_2 + \sigma_c^2 \mathbf{Z}_3 \mathbf{Z}'_3 + \sigma^2 \mathbf{I}_n)$, soit $\mathcal{N}_n(\mu\kappa_n; \mathbf{V})$, en posant ici :

$$\mathbf{V} = \sigma_a^2 \mathbf{Z}_1 \mathbf{Z}'_1 + \sigma_b^2 \mathbf{Z}_2 \mathbf{Z}'_2 + \sigma_c^2 \mathbf{Z}_3 \mathbf{Z}'_3 + \sigma^2 \mathbf{I}_n.$$

Remarque. — Ici encore, comme dans le cas d'un seul facteur, on notera qu'il y a des corrélations entre observations, en raison de la nature aléatoire des effets considérés. Ainsi, deux observations Y_{ijk} et $Y'_{i'jk}$ de la même cellule (j, k) ont en commun les v.a.r. A_j , B_k et C_{jk} , de sorte que leur covariance vaut : $\sigma_a^2 + \sigma_b^2 + \sigma_c^2$. De même, deux observations Y_{ijk} et $Y'_{ij'k'}$ de la même ligne j ont en commun la v.a.r. A_j , de sorte que leur covariance vaut σ_a^2 . Enfin, on a encore un résultat de même nature pour deux observations d'une même colonne.

3.2 Estimation des composantes de la variance dans le cas équilibré

Dans toute la suite de cette section 6.2, on supposera que l'on a affaire à un plan équilibré, autrement dit on posera : $n_{jk} = n_0 \geq 1, \forall (j, k)$. Il s'ensuit $n = n_0 JK$, et l'estimateur du paramètre de moyenne est $\hat{\mu} = \bar{Y}_{\bullet\bullet\bullet} = \frac{1}{n_0 JK} \sum_{j=1}^J \sum_{k=1}^K \sum_{i=1}^{n_0} Y_{ijk}$. On se reportera au paragraphe 6.3.3 pour le cas déséquilibré (à partir de deux facteurs à effets aléatoires, en présence d'un plan

déséquilibré, on a recours aux procédures générales d'estimation relatives aux modèles mixtes, sur lesquelles nous reviendrons en 6.3.3).

3.2.1 Estimation par ANOVA

Le principe de cette méthode est toujours le même : on définit un système d'équations en égalant les espérances mathématiques de certains carrés moyens avec leur moyenne empirique. Chaque carré moyen est relatif à un terme de variance (A , B , C ou U).

On définit ainsi :

$$SSA = n_0 K \sum_{j=1}^J (\bar{Y}_{\bullet\bullet j} - \bar{Y}_{\bullet\bullet\bullet})^2; \quad MSA = \frac{SSA}{J-1}.$$

$$SSB = n_0 J \sum_{k=1}^K (\bar{Y}_{\bullet\bullet k} - \bar{Y}_{\bullet\bullet\bullet})^2; \quad MSB = \frac{SSB}{K-1}.$$

$$SSC = n_0 \sum_{j=1}^J \sum_{k=1}^K (\bar{Y}_{\bullet jk} - \bar{Y}_{\bullet j\bullet} - \bar{Y}_{\bullet\bullet k} + \bar{Y}_{\bullet\bullet\bullet})^2; \quad MSC = \frac{SSC}{(J-1)(K-1)}.$$

$$SSE = \sum_{j=1}^J \sum_{k=1}^K \sum_{i=1}^{n_0} (Y_{ijk} - \bar{Y}_{\bullet jk})^2; \quad MSE = \frac{SSE}{(n_0 - 1)JK}.$$

On peut ensuite calculer (c'est assez fastidieux, mais pas vraiment difficile) les espérances mathématiques de ces carrés moyens.

$$\mathbb{E}(MSA) = \sigma^2 + n_0 \sigma_c^2 + n_0 K \sigma_a^2; \quad \mathbb{E}(MSB) = \sigma^2 + n_0 \sigma_c^2 + n_0 J \sigma_b^2;$$

$$\mathbb{E}(MSC) = \sigma^2 + n_0 \sigma_c^2; \quad \mathbb{E}(MSE) = \sigma^2.$$

Enfin, en égalant ces espérances avec les observations correspondantes des quantités MSA , MSB , MSC et MSE sur l'échantillon considéré (ces observations seront respectivement notées $MSA(y)$, $MSB(y)$, $MSC(y)$ et $MSE(y)$), on obtient un système linéaire de quatre équations à quatre inconnues, triangulaire, dont on déduit immédiatement les estimations ANOVA des composantes de la variance :

$$\hat{\sigma}^2 = MSE(y); \quad \hat{\sigma}_c^2 = \frac{MSC(y) - MSE(y)}{n_0};$$

$$\hat{\sigma}_b^2 = \frac{MSB(y) - MSC(y)}{n_0 J}; \quad \hat{\sigma}_a^2 = \frac{MSA(y) - MSC(y)}{n_0 K}.$$

En remplaçant ensuite les observations des moyennes de carrés par les statistiques correspondantes, on obtient les expressions analogues pour les estimateurs (il suffit d'enlever les (y)); ces derniers seront respectivement notés $\hat{\Sigma}^2$, $\hat{\Sigma}_c^2$, $\hat{\Sigma}_b^2$ et $\hat{\Sigma}_a^2$.

Propriétés des estimateurs ANOVA

- Il peut arriver que les valeurs calculées des estimateurs $\hat{\Sigma}_a^2$, $\hat{\Sigma}_b^2$ ou $\hat{\Sigma}_c^2$ soient négatives; dans ce cas, elle sont mises à zéro et le facteur correspondant est supprimé du modèle.
- Les quatre estimateurs $\hat{\Sigma}_a^2$, $\hat{\Sigma}_b^2$, $\hat{\Sigma}_c^2$ et $\hat{\Sigma}^2$ sont sans biais (c'est immédiat d'après les formules ci-dessus).
- Parmi les estimateurs sans biais, ils sont de variance minimum (admis).
- On peut encore vérifier : $\frac{SSE}{\sigma^2} \sim \chi_{(n_0-1)JK}^2$.
- On ne sait pas expliciter la loi de probabilité des trois autres estimateurs.

3.2.2 Autres méthodes d'estimation

Comme on ne considère que le cas équilibré dans ce paragraphe, les estimateurs ANOVA, REML, MINQUE et MIVQUE sont identiques. Seuls les estimateurs maximum de vraisemblance sont différents. On obtient encore MSE comme estimateur de σ^2 par maximum de vraisemblance, mais les estimateurs des trois autres composantes de la variance sont en général différents de ceux explicités ci-dessus. De plus, ils sont biaisés.

Remarque. — Dans le cas déséquilibré, les différentes méthodes d'estimation fournissent, en général, des résultats différents (voir le point 6.3.3).

3.3 Tests des effets aléatoires dans le cas équilibré

3.3.1 Propriétés préliminaires

On a déjà signalé :

$$\frac{(n_0 - 1)JK}{\sigma^2} MSE \sim \chi_{(n_0-1)JK}^2.$$

On peut également vérifier :

$$\frac{(J - 1) MSA}{n_0 K \sigma_a^2 + n_0 \sigma_c^2 + \sigma^2} \sim \chi_{J-1}^2; \quad \frac{(K - 1) MSB}{n_0 J \sigma_b^2 + n_0 \sigma_c^2 + \sigma^2} \sim \chi_{K-1}^2;$$

$$\frac{(J - 1)(K - 1) MSC}{n_0 \sigma_c^2 + \sigma^2} \sim \chi_{(J-1)(K-1)}^2.$$

De plus, les quatre statistiques ci-dessus sont indépendantes. Cela permet de définir des statistiques de tests, distribuées selon des lois de Fisher, pour tester les différentes hypothèses nulles relatives au modèle à deux facteurs croisés. Comme déjà vu pour des facteurs à effets fixes, on procède de façon hiérarchique, en commençant par tester les interactions.

3.3.2 Test de $\{H_0^c : \sigma_c^2 = 0\}$ contre $\{H_1^c : \sigma_c^2 > 0\}$

Sous H_0^c , il est clair que $F_c = \frac{MSC}{MSE} \sim F_{(J-1)(K-1); (n_0-1)JK}$. F_c est donc la statistique du test de H_0^c contre H_1^c et c'est encore la même statistique que pour tester l'interaction entre deux facteurs à effets fixes. Si H_0^c est rejetée, on conserve le modèle complet, avec les deux facteurs aléatoires et les interactions. Sinon, on enlève les interactions du modèle et on conserve le modèle additif.

3.3.3 Test de $\{H_0^a : \sigma_a^2 = 0\}$ contre $\{H_1^a : \sigma_a^2 > 0\}$ dans le modèle complet

Sous H_0^a , $F_a = \frac{MSA}{MSC} \sim F_{J-1; (J-1)(K-1)}$. F_a est donc la statistique du test de H_0^a contre H_1^a .

3.3.4 Test de $\{H_0^b : \sigma_b^2 = 0\}$ contre $\{H_1^b : \sigma_b^2 > 0\}$ dans le modèle complet

De façon symétrique, on a maintenant, sous H_0^b : $F_b = \frac{MSB}{MSC} \sim F_{K-1; (J-1)(K-1)}$. F_b est la statistique du test de H_0^b contre H_1^b .

Remarque. — On notera que les dénominateurs des deux dernières statistiques de tests sont MSC et non MSE . Autrement dit, ces tests ne sont pas les mêmes que ceux qu'on ferait dans le cadre d'un modèle à effets fixes.

Remarque. — Toujours en ce qui concerne ces deux derniers tests, on voit qu'ils sont les mêmes, que H_0^c soit vraie ou non. Si on le souhaite, on peut donc les utiliser pour tester les effets principaux au sein du modèle additif. Toutefois, ceci n'est pas l'optique du logiciel SAS.

Remarque. — On ne dispose plus de tels tests dans le cas déséquilibré (voir le point 6.3.5).

Remarque. — Concernant les intervalles de confiance, on dispose toujours du même intervalle que dans le cas d'un seul facteur pour les paramètres μ et σ^2 . Par contre, on n'a pas de résultat précis pour les autres composantes de la variance.

Remarque. — Pour déterminer les valeurs prédites une fois un modèle choisi, nous renvoyons au point 6.3.6.

4 Modèles mixtes

On appelle modèle mixte un modèle statistique dans lequel on considère à la fois des facteurs à effets fixes (qui vont intervenir au niveau de la moyenne du modèle) et des facteurs à effets aléatoires (qui vont intervenir au niveau de la variance du modèle). Un modèle est dit mixte lorsqu'il y a au moins un facteur de chaque nature. Dans le cadre de ce cours, nous ne considérons que des modèles linéaires gaussiens mixtes, mais la notion de modèle mixte se rencontre également dans d'autres contextes, notamment dans le modèle linéaire généralisé. Dans la suite de ce paragraphe, nous ne spécifierons pas le nombre de facteurs à effets fixes, ni celui de facteurs à effets aléatoires : ils seront quelconques.

4.1 Écriture générale d'un modèle linéaire gaussien mixte

Un modèle linéaire gaussien mixte, relatif à n observations, s'écrit sous la forme matricielle suivante :

$$Y = \mathbf{X}\beta + \sum_{k=1}^K \mathbf{Z}_k A_k + U = \mathbf{X}\beta + \mathbf{Z}A + U.$$

Nous précisons ci-dessous les éléments de cette écriture.

- Y est le vecteur aléatoire réponse de \mathbb{R}^n .

- \mathbf{X} est la matrice $n \times p$ relative aux effets fixes du modèle (figurant en colonnes) ; p est donc le nombre total d'effets fixes pris en compte dans le modèle ; la matrice \mathbf{X} est analogue à la matrice d'incidence dans une ANOVA.
- β est le vecteur des p effets fixes β_j ($j = 1, \dots, p$) ; il s'agit de paramètres à estimer.
- \mathbf{Z}_k est la matrice des indicatrices (disposées en colonnes) des niveaux du k -ième facteur à effets aléatoires ($k = 1, \dots, K$) ; on notera q_k le nombre de niveaux de ce facteur ; \mathbf{Z}_k est donc de dimension $n \times q_k$.
- Nous allons noter $A_{k\ell}$ la v.a.r. associée au ℓ -ième niveau du k -ième facteur à effets aléatoires ($\ell = 1, \dots, q_k$) ; pour tout ℓ , on a $A_{k\ell} \sim \mathcal{N}(0, \sigma_k^2)$; pour un indice k donné (autrement dit, pour un facteur déterminé), les v.a.r. $A_{k\ell}$ sont supposées indépendantes (donc i.i.d.) ; bien entendu, deux observations de la v.a.r. réponse Y faites au même niveau ℓ du k -ième facteur sont corrélées, leur covariance comportant le terme σ_k^2 et, éventuellement, d'autres composantes de la variance ; par ailleurs, pour deux indices k et k' distincts, $A_{k\ell}$ et $A_{k'\ell'}$ sont indépendantes, pour tous les niveaux ℓ et ℓ' .

- Dans l'écriture matricielle ci-dessus, on a posé $A_k = \begin{pmatrix} A_{k1} \\ \vdots \\ A_{kq_k} \end{pmatrix}$, de

sorte que l'on a

$A_k \sim \mathcal{N}_{q_k}(0, \sigma_k^2 \mathbf{I}_{q_k})$, les vecteurs aléatoires A_k étant mutuellement indépendants.

- U est le vecteur aléatoire des erreurs du modèle ; il vérifie $U \sim \mathcal{N}_n(0, \sigma^2 \mathbf{I}_n)$; de plus, il est supposé indépendant des A_k .
- Pour obtenir la seconde écriture (simplifiée) du modèle, qui est la forme la plus générale d'un modèle linéaire mixte, on a posé :

$$\mathbf{Z} = (\mathbf{Z}_1 | \dots | \mathbf{Z}_K) ; A = \begin{pmatrix} \overline{A_1} \\ \vdots \\ \overline{A_K} \end{pmatrix}.$$

\mathbf{Z} est une matrice (connue) de dimension $n \times q$, avec $q = \sum_{k=1}^K q_k$ (q est le nombre total d'effets aléatoires considérés) ; A est un vecteur aléatoire gaussien de \mathbb{R}^q .

Remarque. — Il n'est pas très courant (et, dans la mesure du possible, il vaut mieux l'éviter) de considérer des effets d'interactions entre un facteur à effets fixes et un autre facteur à effets aléatoires. Toutefois, dans certains cas pratiques, on peut être amené à le faire. Dans ce cas, les interactions en question doivent nécessairement être considérées comme des effets aléatoires. Dans l'écriture générale des modèles mixtes, elles sont donc intégrées dans la partie A.

4.1.1 Moments de Y

- De façon évidente, il vient : $\mathbb{E}(Y) = \mathbf{X}\beta$.
- D'autre part : $\text{Var}(Y) = \mathbf{V} = \text{Var}(\mathbf{Z}A) + \text{Var}(U) = \sum_{k=1}^K \sigma_k^2 \mathbf{Z}_k \mathbf{Z}_k' + \sigma^2 \mathbf{I}_n$.

Finalement, on obtient $Y \sim \mathcal{N}_n(\mathbf{X}\beta, \mathbf{V})$, les composantes de Y n'étant pas indépendantes au sein d'un même niveau d'un facteur aléatoire donné.

Remarque. — Il est courant de poser : $\mathbf{G} = \text{diag}(\sigma_1^2 \mathbf{I}_{q_1} \cdots \sigma_K^2 \mathbf{I}_{q_K})$. Cela permet d'écrire :

$\sum_{k=1}^K \sigma_k^2 \mathbf{Z}_k \mathbf{Z}_k' = \mathbf{Z}\mathbf{G}\mathbf{Z}'$; d'où : $\mathbf{V} = \mathbf{Z}\mathbf{G}\mathbf{Z}' + \sigma^2 \mathbf{I}_n$ (la partie $\sigma^2 \mathbf{I}_n$ de la variance est parfois notée \mathbf{R}).

Remarque. — Il arrive que l'on intègre le vecteur aléatoire des erreurs U dans la partie aléatoire du modèle mixte ci-dessus, en remarquant que U et chaque

A_k sont de même nature. On peut en effet écrire : $U = \mathbf{I}_n U = \mathbf{I}_n \begin{pmatrix} U_1 \\ \vdots \\ U_n \end{pmatrix}$,

où $U_i \sim \mathcal{N}(0, \sigma^2)$. On peut ainsi considérer que U représente un facteur à n effets aléatoires. Dans ce cas, on le note A_0 , on pose $\mathbf{Z}_0 = \mathbf{I}_n$, $\mathbf{Z}^* = (\mathbf{Z}_0 | \mathbf{Z})$ et $A^* = (A_0 | A)'$, ce qui permet de réécrire : $Y = \mathbf{X}\beta + \mathbf{Z}^* A^*$. Nous utiliserons peu cette écriture simplifiée, pour éviter de mélanger une partie du modèle avec son erreur. Toutefois, elle pourra être implicite, comme par exemple dans l'estimation des composantes de la variance par ANOVA développée en 6.3.3 (cette réécriture du modèle mixte permet surtout d'alléger les notations par la suite).

4.2 Estimation des paramètres dans le cas équilibré

4.2.1 Estimation de β

Le vecteur β des p paramètres β_j correspondant aux effets fixes du modèle a, dans ce cas, toujours la même expression, quelle que soit la méthode utilisée pour estimer les composantes de la variance. Il s'agit de l'expression fournie par la méthode des moindres carrés ordinaires, que nous noterons $\text{OLSE}(\beta)$ (pour *Ordinary Least Squares Estimator*) : $\hat{B} = \text{OLSE}(\beta) = (\mathbf{X}'\mathbf{X})^{-1} \mathbf{X}'Y$ (dans ce cas, il n'est pas nécessaire d'avoir estimé \mathbf{V} pour obtenir l'estimation de β).

4.2.2 Estimation des composantes de la variance

Il s'agit ici d'estimer les $K + 1$ paramètres de variance σ_k^2 ($k = 1, \dots, K$) et σ^2 . Cela revient à estimer la matrice \mathbf{V} des variances-covariances de la variable réponse Y .

Le plan étant équilibré, les méthodes ANOVA, REML, MINQUE et MIVQUE, qui généralisent ce qui a été exposé en 6.1.4, conduisent toutes au même résultat explicite, solution d'un système de $K + 1$ équations linéaires. Les estimateurs obtenus sont sans biais et de variance minimum parmi les estimateurs sans biais et quadratiques en les observations. On peut néanmoins obtenir des valeurs négatives pour certaines composantes de la variance : elles sont alors mises à 0 et le facteur correspondant doit être retiré du modèle.

De son côté, la méthode du maximum de vraisemblance fournit des solutions en général différentes des précédentes et biaisées.

En 6.3.3, nous revenons plus en détails sur les méthodes d'estimation de \mathbf{V} dans le cas général.

4.3 Estimation des paramètres dans le cas déséquilibré

4.3.1 Estimation de β

L'expression que l'on obtient dans ce cas pour \hat{B} fait intervenir l'estimation de la matrice des variances-covariances \mathbf{V} de Y . Si l'expression est unique, la valeur correspondante dépend de la méthode d'estimation de \mathbf{V} . En fait, l'expression obtenue est aussi celle fournie par la méthode des moindres carrés généralisés, pour cette raison notée $\text{GLSE}(\beta)$ (pour *Generalized Least Squares*

Estimator) :

$$\hat{B} = \text{GLSE}(\beta) = (\mathbf{X}'\hat{\mathbf{V}}^{-1}\mathbf{X})^{-1}\mathbf{X}'\hat{\mathbf{V}}^{-1}Y.$$

Dans l'expression ci-dessus, $\hat{\mathbf{V}} = \sum_{k=1}^K \hat{\sigma}_k^2 \mathbf{Z}_k \mathbf{Z}_k' + \hat{\sigma}^2 \mathbf{I}_n$; on voit donc que l'on doit estimer les composantes de la variance avant de pouvoir estimer le vecteur β (la justification de l'expression ci-dessus de \hat{B} est donnée dans le point traitant de l'estimation de \mathbf{V} par maximum de vraisemblance).

Quelle que soit la méthode utilisée pour estimer les composantes de la variance, $\text{GLSE}(\beta)$ est un estimateur sans biais de β .

Remarque. — On peut vérifier que, dans le cas d'un plan équilibré, on obtient $\text{GLSE}(\beta) = \text{OLSE}(\beta)$.

4.3.2 Estimation de \mathbf{V} par ANOVA

Cette méthode consiste à généraliser ici ce qui a été vu en 6.1.4, les sommes de carrés étant maintenant remplacées par des matrices de formes quadratiques sur \mathbb{R}^n . Pour chacun des facteurs à effets aléatoires, ainsi que pour la v.a.r. erreur U , on définit donc une matrice réelle \mathbf{C}_h , carrée d'ordre n , symétrique et strictement définie positive. En affectant l'indice 0 à U , l'indice h va ainsi varier de 0 à K . Ces matrices \mathbf{C}_h , autrement dit ces formes quadratiques, seront choisies de telle sorte que les équations obtenues soient commodes à résoudre (voir plus loin).

Si l'on considère un vecteur aléatoire Y de \mathbb{R}^n , de loi $\mathcal{N}_n(\mu, \mathbf{V})$, on sait que $\mathbb{E}(Y' \mathbf{C}_h Y) = \mu' \mathbf{C}_h \mu + \text{tr}(\mathbf{C}_h \mathbf{V})$. En appliquant cette formule au modèle mixte écrit plus haut, il vient :

$$Y \sim \mathcal{N}_n(\mathbf{X}\beta, \mathbf{V}) \implies \mathbb{E}(Y' \mathbf{C}_h Y) = \beta' \mathbf{X}' \mathbf{C}_h \mathbf{X} \beta + \text{tr}(\mathbf{C}_h \mathbf{V}).$$

Pour que l'estimation de \mathbf{V} , que l'on va faire au moyen de ces formes quadratiques, soit déconnectée de celle de β , on choisit les matrices \mathbf{C}_h telles que $\mathbf{X}' \mathbf{C}_h \mathbf{X} = 0$ (ce qui est équivalent à les choisir telles que $\mathbf{C}_h \mathbf{X} = 0$, dès que \mathbf{C}_h est strictement définie positive). On obtient ainsi :

$$\mathbb{E}(Y' \mathbf{C}_h Y) = \text{tr}(\mathbf{C}_h \mathbf{V}) = \text{tr}(\mathbf{C}_h \sum_{k=0}^K \sigma_k^2 \mathbf{Z}_k \mathbf{Z}_k') = \sum_{k=0}^K \sigma_k^2 \text{tr}(\mathbf{Z}_k' \mathbf{C}_h \mathbf{Z}_k).$$

La méthode d'ANOVA consiste ainsi à suivre les étapes suivantes :

- on choisit $K + 1$ matrices \mathbf{C}_h ($h = 0, 1, \dots, K$) vérifiant les propriétés requises et linéairement indépendantes ;
- on appelle \mathbf{T} la matrice carrée d'ordre $K + 1$ de terme général : $\mathbf{T}_h^k = \text{tr}(\mathbf{Z}_k' \mathbf{C}_h \mathbf{Z}_k)$;
- on pose :

$$\gamma^2 = \begin{pmatrix} \sigma^2 \\ \sigma_1^2 \\ \vdots \\ \sigma_K^2 \end{pmatrix} ; \quad Q(y) = \begin{pmatrix} Q_0(y) \\ Q_1(y) \\ \vdots \\ Q_K(y) \end{pmatrix}, \quad \text{avec } Q_h(y) = y' \mathbf{C}_h y ;$$

- on résout le système $\mathbb{E}(Y' \mathbf{C}_h Y) = y' \mathbf{C}_h y$ ($h = 0, 1, \dots, K$), soit encore : $\mathbf{T} \gamma^2 = Q(y)$;
- ce qui conduit à la solution : $\hat{\gamma}^2 = \mathbf{T}^{-1} Q(y)$.

On voit que toute la méthode repose sur le choix approprié des matrices \mathbf{C}_h . Il est clair que différents choix au niveau de ces matrices conduiront à différentes estimations des composantes de la variance. Il est maintenant courant d'utiliser l'une des méthodes proposées par C.R. Henderson (1953) et améliorées par la suite.

- La méthode dite Henderson I ne s'applique qu'aux modèles à effets aléatoires, autrement dit tels que $\mathbf{X} = 0_n$, ce qui évite d'imposer la propriété $\mathbf{X}' \mathbf{C}_h \mathbf{X} = 0$ aux matrices \mathbf{C}_h . Elle est calquée sur la méthode décrite en 6.2.2, en adaptant les sommes de carrés au cas déséquilibré.
- La méthode Henderson II est une adaptation de la précédente qui définit des formes quadratiques spécifiques permettant d'annuler les effets fixes, donc associées à des matrices \mathbf{C}_h vérifiant $\mathbf{X}' \mathbf{C}_h \mathbf{X} = 0$. Toutefois, cette méthode ne peut s'appliquer s'il y a dans le modèle des interactions entre effets fixes et effets aléatoires.
- La **méthode Henderson III** est la plus générale (elle s'applique dans tous les cas), mais aussi la plus compliquée. C'est la plus utilisée dans la pratique et c'est, en particulier, celle qui est mise en œuvre dans le logiciel SAS.

Toutes ces méthodes fournissent des estimateurs sans biais des composantes de la variance. On trouvera une présentation très détaillée des méthodes de Henderson dans Searle *et al.* (1992).

4.3.3 Estimation de \mathbf{V} par maximum de vraisemblance

La log-vraisemblance du modèle mixte gaussien s'écrit :

$$l(y, \beta, \mathbf{V}) = c - \frac{1}{2} \log[\det(\mathbf{V})] - \frac{1}{2} (y - \mathbf{X}\beta)' \mathbf{V}^{-1} (y - \mathbf{X}\beta)$$

(généralisation de ce qui a été fait en 6.1.4). On en déduit le système de p équations :

$$\frac{\partial l}{\partial \beta} = \mathbf{X}' \mathbf{V}^{-1} y - \mathbf{X}' \mathbf{V}^{-1} \mathbf{X} \beta$$

dont découlent les équations normales.

On remarque ensuite : $\frac{\partial \mathbf{V}}{\partial \sigma_k^2} = \mathbf{Z}_k \mathbf{Z}_k'$; on en déduit :

$$\frac{\partial l}{\partial \sigma_k^2} = -\frac{1}{2} \text{tr}(\mathbf{V}^{-1} \mathbf{Z}_k \mathbf{Z}_k') + \frac{1}{2} (y - \mathbf{X}\beta)' \mathbf{V}^{-1} \mathbf{Z}_k \mathbf{Z}_k' \mathbf{V}^{-1} (y - \mathbf{X}\beta)$$

(une équation pour chaque σ_k^2 , $k = 0, 1, \dots, K$). Les premières équations de vraisemblance fournissent :

$$\hat{\beta} = (\mathbf{X}' \mathbf{V}^{-1} \mathbf{X})^{-1} \mathbf{X}' \mathbf{V}^{-1} y = \text{GLSE}(\beta).$$

Les suivantes s'écrivent :

$$\text{tr}(\mathbf{V}^{-1} \mathbf{Z}_k \mathbf{Z}_k') = (y - \mathbf{X}\hat{\beta})' \mathbf{V}^{-1} \mathbf{Z}_k \mathbf{Z}_k' \mathbf{V}^{-1} (y - \mathbf{X}\hat{\beta}), \quad k = 0, 1, \dots, K.$$

On obtient ainsi un système de $K + 1 + p$ équations non linéaires à $K + 1 + p$ inconnues que l'on résout par une méthode numérique itérative (de type Fisher scoring). Ces procédures numériques fournissent en plus, à la convergence, la matrice des variances-covariances asymptotiques des estimateurs.

Les estimateurs obtenus par maximum de vraisemblance sont, en général, biaisés : la méthode produit un biais systématique. Ils peuvent être négatifs et sont alors ramenés à 0.

Remarque. — On a vu que les p premières équations ci-dessus fournissent l'estimation maximum de vraisemblance de β , $\hat{\beta} = \text{GLSE}(\beta)$, dans tous les cas de figure. Pour retrouver l'expression donnée pour $\hat{\mu}$ en 6.1.3 dans le cas déséquilibré, il faut remarquer que, dans le cas particulier où μ est le seul effet

fixe, $\mathbf{X} = \mathbb{1}_n$, de sorte que $\mathbf{X}' \mathbf{V}^{-1} \mathbf{X}$ est la somme de tous les termes de \mathbf{V}^{-1} . Comme on a, dans ce cas, $\mathbf{V} = \sigma_a^2 \mathbf{Z} \mathbf{Z}' + \sigma^2 \mathbf{I}_n$ (matrice bloc-diagonale), elle s'inverse par bloc, chaque bloc étant carré d'ordre n_j , de la forme

$$\begin{pmatrix} \sigma_a^2 + \sigma^2 & \sigma_a^2 & \cdots \\ \sigma_a^2 & \sigma_a^2 + \sigma^2 & \cdots \\ \cdots & \cdots & \cdots \end{pmatrix},$$

dont l'inverse s'écrit

$$\frac{1}{(n_j \sigma_a^2 + \sigma^2) \sigma^2} \begin{pmatrix} (n_j - 1) \sigma_a^2 + \sigma^2 & -\sigma_a^2 & \cdots \\ -\sigma_a^2 & (n_j - 1) \sigma_a^2 + \sigma^2 & \cdots \\ \cdots & \cdots & \cdots \end{pmatrix}.$$

Ainsi, la somme de toute ligne (ou de toute colonne) de \mathbf{V}^{-1} vaut $\frac{1}{(n_j \sigma_a^2 + \sigma^2)}$ et le total de tous les termes de cette matrice vaut

$$\sum_{j=1}^J \frac{n_j}{n_j \sigma_a^2 + \sigma^2} = \sum_{j=1}^J \frac{1}{\tau_j^2}.$$

D'autre part, $\mathbb{1}_n' \mathbf{V}^{-1} y$ vaut $\sum_{j=1}^J \frac{\bar{y}_{\bullet j}}{\tau_j^2}$, ce qui permet d'écrire :

$$\mu = \sum_{j=1}^J w_j \bar{y}_{\bullet j} \quad \text{et} \quad \hat{\mu} = \sum_{j=1}^J \hat{w}_j \bar{y}_{\bullet j}.$$

4.3.4 Estimation de \mathbf{V} par maximum de vraisemblance restreint

On a pu constater, dans le point précédent, que les différentes équations obtenues lorsqu'on réalise l'estimation des paramètres par maximum de vraisemblance contiennent à la fois le vecteur β (des paramètres liés à l'espérance de Y) et la matrice \mathbf{V} (des paramètres liés à la variance de Y). Dans un modèle mixte, c'est ce mélange de paramètres de natures différentes dans les mêmes équations qui engendre un biais systématique dans l'estimation par maximum de vraisemblance des composantes de la variance. L'objet de la méthode du

maximum de vraisemblance restreint est précisément de séparer les deux types de paramètres.

L'idée est la suivante : aux colonnes de la matrice \mathbf{X} sont associés des vecteurs de l'espace vectoriel $F = \mathbb{R}^n$ (n est le nombre d'observations réalisées) ; on munit ce dernier de la métrique identité (associée à la matrice \mathbf{I}_n sur la base canonique), ce qui en fait un espace euclidien. En notant F_X le sous-espace vectoriel de F engendré par les colonnes de X (F_X est supposé de dimension p), on sait que le projecteur orthogonal de F dans F_X a pour matrice associée $\mathbf{H} = \mathbf{X}(\mathbf{X}'\mathbf{X})^{-1}\mathbf{X}'$. Par ailleurs, le projecteur sur le s.e.v. F_X^\perp , supplémentaire orthogonal à F_X dans F , est $\mathbf{H}^\perp = \mathbf{I}_n - \mathbf{H}$. Ainsi, en projetant Y sur F_X^\perp et en travaillant avec cette projection (qui est, par définition, orthogonale à toute combinaison linéaire des colonnes de \mathbf{X}), on s'affranchit de β dans l'estimation des composantes de la variance. Toutefois, le s.e.v. F_X^\perp étant de dimension $m = n - p$, le vecteur aléatoire projeté de Y sur F_X^\perp est multinormal d'ordre m . Écrit dans \mathbb{R}^n , c'est un vecteur aléatoire dégénéré (sa matrice des variances-covariances est singulière) qu'on doit transformer pour obtenir un vecteur aléatoire directement écrit dans \mathbb{R}^m . Soit donc \mathbf{M}_0 une matrice $m \times n$, de rang m , réalisant cette transformation et soit $\mathbf{M} = \mathbf{M}_0\mathbf{H}^\perp$. On considère finalement $Y^* = \mathbf{M}Y$; on a ainsi $Y^* \sim \mathcal{N}_m(\mu, \mathbf{V}^*)$, avec :

$$\begin{aligned} \mu &= \mathbf{M}\mathbf{X}\beta = \mathbf{M}_0(\mathbf{H}^\perp\mathbf{X}\beta) = 0 \text{ (par définition de } \mathbf{H}^\perp \text{)} ; \\ \mathbf{V}^* &= \mathbf{M}\mathbf{V}\mathbf{M}' = \mathbf{M}_0\mathbf{H}^\perp\mathbf{V}\mathbf{H}^\perp\mathbf{M}_0'. \end{aligned}$$

En fait, on peut encore écrire :

$$\mathbf{V}^* = \sum_{k=0}^K \sigma_k^2 \mathbf{M}\mathbf{Z}_k\mathbf{Z}_k'\mathbf{M}' = \sum_{k=0}^K \sigma_k^2 \mathbf{Z}_k^*\mathbf{Z}_k^{*'} \text{ (en posant } \mathbf{Z}_k^* = \mathbf{M}\mathbf{Z}_k \text{)}.$$

En réécrivant les $K + 1$ dernières équations de vraisemblance relatives au vecteur aléatoire Y^* , il vient maintenant :

$$\text{tr}(\mathbf{V}^{*-1}\mathbf{Z}_k^*\mathbf{Z}_k^{*'}) = y^{*'}\mathbf{V}^{*-1}\mathbf{Z}_k^*\mathbf{Z}_k^{*'}\mathbf{V}^{*-1}y^*, \quad k = 0, 1, \dots, K \text{ (avec : } y^* = \mathbf{M}y \text{)}.$$

Il s'agit d'un système de $K + 1$ équations non linéaires à $K + 1$ inconnues (les composantes σ_k^2 de la variance ; $k = 0, 1, \dots, K$) dans lequel ne figure plus le vecteur β . En général, il n'admet pas de solution analytique et nécessite une procédure numérique itérative pour sa résolution. Les estimateurs $\hat{\Sigma}_k^2$ ainsi

obtenus ne sont pas nécessairement sans biais, mais ne comportent pas de biais systématique. Là encore, la procédure itérative fournit, à la convergence, la matrice des variances-covariances asymptotiques des estimateurs.

Remarque. — Les équations écrites ci-dessus proviennent de l'annulation des dérivées partielles, selon les σ_k^2 , de la log-vraisemblance du vecteur aléatoire Y^* , autrement dit de la log-vraisemblance de la projection de Y sur le s.e.v. F_X^\perp , ou encore de la restriction de la vraisemblance de Y à ce sous-espace. Les estimateurs obtenus par maximisation de cette restriction de la vraisemblance sont, pour cette raison, appelés estimateurs du maximum de vraisemblance restreint (on devrait dire, de façon plus rigoureuse, du maximum de la vraisemblance restreinte).

4.3.5 Estimation de \mathbf{V} par MINQUE et MIVQUE

Dans un modèle mixte, il est encore possible d'estimer les composantes de la variance en utilisant soit la méthode MINQUE, soit la méthode MIVQUE. Le principe général reste le même que celui exposé en 6.1.4. En fait, ces méthodes sont peu utilisées dans la pratique. Il faut néanmoins rappeler que SAS utilise la méthode dite MIVQUE(0) pour initialiser la procédure itérative utilisée avec la méthode REML (procédures VARCOMP et MIXED).

4.4 Intervalles de confiance

Dans le cas d'un plan équilibré, on peut construire un intervalle de confiance exact, de type Student, pour toute fonction linéaire $c'\beta$ du paramètre β relatif aux effets fixes ($c \in \mathbb{R}^p$). Le principe est le même que celui indiqué au paragraphe 6.1.5. Dans le cas d'un plan déséquilibré, cela n'est plus possible.

Pour la variance des erreurs σ^2 , un intervalle de confiance exact, de type khi-deux, peut être construit dans tous les cas, que ce soit pour un plan équilibré ou pour un plan déséquilibré.

Mais, pour les autres paramètres de variance, on ne dispose pas d'intervalle de confiance précis (pas même asymptotique).

4.5 Tests de significativité des facteurs

Ces tests sont standards dans le cas équilibré, mais deviennent assez problématiques dans le cas déséquilibré.

4.5.1 Cas équilibré

Pour tester la significativité d'un facteur à effets fixes dans un modèle mixte, on utilise le test habituel de Fisher, tel qu'il a été défini en ANOVA : il reste valable dans ce cas (rappelons qu'il s'agit d'un test exact).

Pour tester la significativité d'un facteur à effets aléatoires, c'est-à-dire la nullité d'une composante de la variance, on utilise encore un test de Fisher analogue à ceux définis en 6.2.3.

Tous ces tests sont mis en œuvre par la procédure GLM de SAS, mais seuls ceux relatifs aux effets fixes le sont par la procédure MIXED.

4.5.2 Cas déséquilibré

Il n'y a malheureusement pas de test exact, ni même de test asymptotique, qui permette de tester les effets, que ce soient les effets fixes ou les effets aléatoires, dans un modèle mixte avec un plan déséquilibré. Il existe seulement des tests approchés (dont on ne contrôle pas réellement le niveau, et encore moins la puissance). Nous donnons néanmoins ci-dessous quelques pistes qui permettront d'aider à choisir le modèle le plus approprié relativement à un jeu de données.

– *Le test de Fisher, avec degré de liberté calculé selon l'approximation de Satterthwaite.* Nous présentons ce test dans un cadre simple, le principe restant le même dans tout modèle mixte. On se place donc dans le cas d'un modèle à deux facteurs croisés à effets aléatoires. Notons A et B les deux facteurs considérés, J et K leurs nombres de niveaux, C leurs interactions et E les erreurs du modèle considéré. Enfin, on note encore MS les carrés moyens associés à chacun de ces effets. Il est possible d'écrire :

$$\begin{aligned}\mathbb{E}(MSA) &= \sigma^2 + \alpha_1\sigma_c^2 + \alpha_2\sigma_a^2 ; \\ \mathbb{E}(MSB) &= \sigma^2 + \alpha_3\sigma_c^2 + \alpha_4\sigma_b^2 ; \\ \mathbb{E}(MSC) &= \sigma^2 + \alpha_5\sigma_c^2 ; \\ \mathbb{E}(MSE) &= \sigma^2.\end{aligned}$$

On ne dispose pas d'expression explicite pour les coefficients α_i , mais on sait les déterminer numériquement, en général en utilisant la méthode dite de Henderson III (c'est ce qui est fait dans le logiciel SAS).

Si l'on souhaite tester, par exemple, l'hypothèse nulle $\{H_0 : \sigma_a^2 = 0\}$, on peut réécrire, sous H_0 :

$$\begin{aligned}\mathbb{E}(MSA) &= \sigma^2 + \alpha_1\sigma_c^2 \\ &= \sigma^2 + \alpha_1\left[\frac{\mathbb{E}(MSC) - \sigma^2}{\alpha_5}\right] \\ &= \sigma^2\left[1 - \frac{\alpha_1}{\alpha_5}\right] + \frac{\alpha_1}{\alpha_5}\mathbb{E}(MSC) \\ &= \frac{\alpha_1}{\alpha_5}\mathbb{E}(MSC) + \left[1 - \frac{\alpha_1}{\alpha_5}\right]\mathbb{E}(MSE).\end{aligned}$$

En fait, on remplace les espérances (inconnues) des carrés moyens intervenant dans le terme de droite de cette expression par les valeurs empiriques correspondantes (par exemple, $MSC(y)$ remplace $\mathbb{E}(MSC)$). On pose alors :

$$MSA0 = \frac{\alpha_1}{\alpha_5} MSC(y) + \left[1 - \frac{\alpha_1}{\alpha_5}\right] MSE(y).$$

Il a été montré par Satterthwaite (1946) que toute quantité du type $MSA0$ est approximativement distribuée selon une loi de khi-deux dont le degré de liberté q peut être calculé. En fait, si d_1 est le degré de liberté de MSC et d_2 celui de l'erreur du modèle considéré (donc de MSE), la formule donnant q est la suivante :

$$q = \frac{\left[\frac{\alpha_1}{\alpha_5} MSC(y) + \left[1 - \frac{\alpha_1}{\alpha_5}\right] MSE(y)\right]^2}{\frac{\left[\frac{\alpha_1}{\alpha_5} MSC(y)\right]^2}{d_1} + \frac{\left[1 - \frac{\alpha_1}{\alpha_5}\right]^2 MSE(y)^2}{d_2}}.$$

Pour tester H_0 , on fait donc le rapport des deux expressions $MSA(y)$ et $MSA0$ et on le compare à une loi de Fisher à $J - 1$ et q degrés de liberté. Si l'on ne peut pas affirmer que ces deux carrés moyens sont indépendants, le fait qu'ils soient calculés avec des sommes de carrés distinctes permet de penser que le résultat obtenu est une approximation correcte d'une loi de Fisher. En fait, des simulations (ainsi que l'expérience) ont montré que ce test approximatif de Fisher fonctionne plutôt bien. C'est lui qui est mis en œuvre dans la procédure GLM du logiciel SAS, mais pas dans la procédure MIXED. Nous recommandons d'utiliser prioritairement ce test dans les applications.

- *Le test de Wald.* Il permet de tester la significativité des composantes de la variance dans les modèles mixtes. En particulier, c'est le test que l'on trouve dans la procédure MIXED de SAS. Clairement, nous déconseillons ce test. Non seulement il s'agit d'un test asymptotique (qui nécessite donc certaines précautions d'utilisation) mais, surtout, la loi de khi-deux obtenue comme loi limite sous H_0 nécessite certaines conditions techniques qui ne sont manifestement pas vérifiées ici (l'hypothèse nulle d'un tel test est nécessairement du type $\{H_0 : \sigma_a^2 = 0\}$, autrement dit la valeur testée est située sur la frontière de l'espace paramétrique \mathbb{R}_+ , ce qui empêche d'établir le résultat asymptotique qui est donc faux).
- *Une autre solution,* peu courante, consiste à calculer un effectif moyen n^* , en faisant la moyenne harmonique de l'ensemble des effectifs des différentes cellules (les cellules définies par le plan d'expériences considéré). On remplace ensuite chaque effectif par n^* et on opère comme en présence d'un plan équilibré avec n^* répétitions. Cette méthode est celle préconisée dans Miller (1997) lorsque les effectifs sont assez proches les uns des autres ; elle nous semble moins intéressante que la première méthode indiquée ci-dessus.
- Certains praticiens regardent ce que donnent les tests standards de Fisher en considérant tous les effets du modèle comme des effets fixes. Cela peut permettre de se faire une idée de l'importance des différents effets, mais doit seulement être considéré comme un complément aux tests de Fisher approchés exposés plus haut.
- Une comparaison des estimations des différentes composantes de la variance dans le modèle complet (dans lequel on a pris en compte tous les effets considérés au départ) permet aussi de préciser l'importance relative des différents effets. En particulier, la comparaison de chacune de ces composantes avec l'estimation de la variance de l'erreur du modèle est un élément intéressant à prendre en compte dans le choix d'un modèle.
- Lorsque le choix entre plusieurs modèles n'est pas clair (autrement dit, lorsqu'on hésite à prendre en compte certains effets dans le modèle retenu), on peut aussi regarder les critères de choix de modèle tels que *AIC* ou *BIC* (en faisant attention à leur définition dans la procédure MIXED de SAS ; se reporter à l'Annexe D). En particulier, si l'un des modèles envisagés minimise chacun de ces deux critères, c'est le modèle qu'il faudra retenir.
- Enfin, signalons l'existence de tests exacts pour tester certaines hypothèses relatives aux effets d'un modèle mixte avec plan déséquilibré. Ces tests sont assez complexes, difficiles à mettre en œuvre, et leur principe dépasse le cadre de ce cours. De plus, ils ne figurent dans aucun des logiciels statistiques courants. On en trouvera une présentation très détaillée dans l'ouvrage de Khuri *et al.* (1998) auquel nous renvoyons le lecteur intéressé.

4.6 Prévisions dans les modèles mixtes

Sur la base du principe exposé dans le point 6.1.7, la prévision d'une v.a.r. Y non observée, au moyen d'un modèle mixte, fait appel à la notion d'espérance conditionnelle, afin d'obtenir un résultat optimal en un certain sens. C'est ce qu'on appelle le BLUP (*Best Linear Unbiased Predictor*), que nous ne développerons pas ici, mais que l'on trouvera détaillé dans Searle *et al.* (1992). D'autre part, il est possible d'obtenir les prévisions de type BLUP avec la procédure MIXED de SAS.

4.7 Illustration

Il s'agit d'un exemple fictif à deux facteurs : le premier, f_1 , supposé à effets fixes, comporte 3 niveaux (notés 1, 2, 3) ; le second, f_2 , supposé à effets aléatoires, en comporte 4 (notés 1, 2, 3, 4). La réponse, y , est constituée d'entiers compris entre 1 et 30. Le plan est complet, déséquilibré. Il y a 35 individus observés, les données étant reproduites ci-après.

Les données

```

1 1 10
1 1 12
1 1 14
1 2 17
1 2 18
1 3 15
1 3 14
1 3 14
1 4 12
1 4 10
1 4 13
1 4 11
2 1 7
2 1 8
2 2 13
2 2 11
2 2 12
2 3 9
2 3 8
2 3 10
2 3 9
2 4 7
2 4 6
3 1 19
3 1 17
3 1 14
3 2 23
3 2 25
3 2 24
3 2 22
3 3 19
3 3 20
3 4 16
3 4 15
3 4 15
  
```

Le programme SAS

Le programme SAS ci-dessous permet de traiter ces données en tenant compte du caractère aléatoire du second facteur, sans considérer d'interactions. Outre les procédures GLM et VARCOMP, on a utilisé ici la procédure MIXED.

```

* options facultatives pour la mise en page des sorties ;
* ----- ;
options linesize=76 pagesize=64 nodate;
title;
footnote 'Modele mixte - Exemple fictif';
* ----- ;
*           lecture des donnees           ;
*   (le fichier "fic.don" contient les donnees ;
*   et se trouve dans le repertoire de travail) ;
* ----- ;
data fic;
infile 'fic.don';
input f1 f2 y;
run;
* ----- ;
*           procedure GLM :           ;
*           f1 en effets fixes et     ;
*           f2 en effets aleatoires   ;
* ----- ;
proc glm data=fic;
class f1 f2;
model y = f1 f2 / ss3;
random f2 / test;
run;
quit;
* ----- ;
*           procedure VARCOMP         ;
*           (estimation des composantes de la variance ;
*           avec l'option reml)       ;
* ----- ;
proc varcomp data=fic method=reml;
class f1 f2;
model y = f1 f2 / fixed=1;
run;
  
```

```
* ----- ;
*           procedure MIXED ;
* ----- ;
proc mixed data=fic method=reml;
class f1 f2;
model y = f1;
random f2;
run;
quit;
```

```
PAGE 3           The GLM Procedure
-----
Source          Type III Expected Mean Square
f1              Var(Error) + Q(f1)
f2              Var(Error) + 8.5556 Var(f2)
PAGE 4
-----
```

Les sorties de la procédure GLM

```
PAGE 1           The GLM Procedure
-----

Class Level Information

Class          Levels  Values
f1              3      1 2 3
f2              4      1 2 3 4

Number of observations    35
```

```
PAGE 2           The GLM Procedure
-----

Dependent Variable: y

Source          DF          Sum of Squares    Mean Square    F Value    Pr > F
Model           5      811.3221334      162.2644267    103.08    <.0001
Error          29      45.6492952       1.5741136
Corrected Total 34      856.9714286

R-Square      Coeff Var      Root MSE      y Mean
0.946732      8.980018      1.254637      13.97143

Source          DF          Type III SS    Mean Square    F Value    Pr > F
f1              2      571.1145937    285.5572968    181.41    <.0001
f2              3      230.8431290    76.9477097     48.88    <.0001
```

```
Tests of Hypotheses for Mixed Model Analysis of Variance
Dependent Variable: y

Source          DF          Type III SS    Mean Square    F Value    Pr > F
f1              2      571.114594    285.557297    181.41    <.0001
f2              3      230.843129    76.947710     48.88    <.0001
Error: MS(Error) 29      45.649295     1.574114
```

Les sorties de la procédure VARCOMP

```
Variance Components Estimation Procedure

Class Level Information

Class          Levels  Values
f1              3      1 2 3
f2              4      1 2 3 4

Number of observations    35
Dependent Variable: y

REML Iterations

Iteration      Objective      Var(f2)      Var(Error)
0              26.1668391628  9.1679369938  1.5668088318
1              26.1634065807  8.7488110831  1.5740571643
2              26.1634065199  8.7470951138  1.5740883770
3              26.1634065199  8.7470951138  1.5740883770

Convergence criteria met.

REML Estimates
```

Variance Component	Estimate		0	1	167.18609531	
			1	2	124.34322114	0.00000121
			2	1	124.34318122	0.00000000
Var(f2)	8.74710		Convergence criteria met.			
Var(Error)	1.57409					
Asymptotic Covariance Matrix of Estimates			Covariance Parameter Estimates			
	Var(f2)	Var(Error)	Cov Parm		Estimate	
Var(f2)	53.16422	-0.01969	f2	8.7466		
Var(Error)	-0.01969	0.17087	Residual	1.5741		

Les sorties de la procédure MIXED

The Mixed Procedure

Model Information

Data Set	WORK.FIC
Dependent Variable	y
Covariance Structure	Variance Components
Estimation Method	REML
Residual Variance Method	Profile
Fixed Effects SE Method	Model-Based
Degrees of Freedom Method	Containment

Fit Statistics

-2 Res Log Likelihood	124.3
AIC (smaller is better)	128.3
AICC (smaller is better)	128.8
BIC (smaller is better)	127.1

Type 3 Tests of Fixed Effects

Effect	Num DF	Den DF	F Value	Pr > F
f1	2	29	181.42	<.0001

Class Level Information

Class	Levels	Values
f1	3	1 2 3
f2	4	1 2 3 4

Dimensions

Covariance Parameters	2
Columns in X	4
Columns in Z	4
Subjects	1
Max Obs Per Subject	35
Observations Used	35
Observations Not Used	0
Total Observations	35

Iteration History

Iteration	Evaluations	-2 Res Log Like	Criterion
-----------	-------------	-----------------	-----------