STATISTIQUE ET LOGICIELS

ANALYSE DE LA VARIANCE A EFFETS MIXTES UTILISATION DE LA PROC MIXED : MAIS QUE RESTE-T-IL A LA PROC GLM ?

Michel Tenenhaus
e-mail : tenenhaus@hec.fr
Groupe HEC (Jouy-en-Josas)

Introduction

L'analyse de la variance joue un rôle tout à fait particulier en Statistique. C'est depuis son origine un univers en perpétuelle expansion. La demande des praticiens a obligé les statisticiens à construire des modèles plus performants, plus souples, s'adaptant mieux à la réalité des données. L'histoire de l'analyse de la variance est bien résumée à travers les procédures SAS d'analyse de la variance : ANOVA, GLM et MIXED. La Proc ANOVA ne traite que les données équilibrées : elle est donc entièrement fondée sur des calculs de sommes de carrés. La Proc GLM embrasse des situations plus complexes : données déséquilibrées, comparaisons multiples, analyse de la variance multivariée, analyse de la variance de mesures répétées, analyse de la variance à effets mixtes. Cependant si la Proc GLM fournit les tests appropriés en analyse de la variance à effets mixtes, elle donne des résultats faux au niveau de l'estimation. Ces limitations de la Proc GLM sont levées dans la Proc MIXED. La Proc MIXED a considérablement simplifié la vie du chercheur en permettant une étude « juste » des modèles à effets mixtes. L'objet de cette conférence est d'identifier avec précision ces limitations de la Proc GLM et de présenter les solutions justes apportées par la Proc MIXED. Il reste cependant deux points pour lesquels la Proc GLM propose des solutions intéressantes : (1) l'approche multivariée dans le traitement des mesures répétées sans données manquantes conduit à une meilleure évaluation des niveaux de signification que le même modèle étudié par la Proc MIXED lorsque la matrice de covariance des résidus est de type « unstructured », (2) la Proc GLM propose des tests basés sur les sommes de carrés de type IV lorsque le plan d'expérience contient des cases vides. Cette possibilité est absente de la Proc MIXED. Chaque thème présenté sera traité autour d'un exemple issu de Milliken & Johnson (1984).

1. L'exemple de Milliken & Johnson : *Effet d'un traitement sur le rythme cardiaque*

On souhaite étudier les effets de trois traitements (AX23, BWW9 et Contrôle) sur le rythme cardiaque. Après que le médicament ait été administré, le rythme cardiaque est mesuré quatre fois aux instants 5 mn, 10 mn, 15 mn, 20 mn. Il y a huit personne par traitement. Les données figurent dans le Tableau 1.

Tableau 1 : Effet d'un traitement sur le rythme cardiaque

	Traitement											
Sujet dans	**AX23**				**BWW9**				**Contrôle**			
traitement	**T1**	**T2**	**T3**	**T4**	**T1**	**T2**	**T3**	**T4**	**T1**	**T2**	**T3**	**T4**

1	72	86	81	77	85	86	83	80	69	73	72	74
2	78	83	88	81	82	86	80	84	66	62	67	73
3	71	82	81	75	71	78	70	75	84	90	88	87
4	72	83	83	69	83	88	79	81	80	81	77	72
5	66	79	77	66	86	85	76	76	72	72	69	70
6	74	83	84	77	85	82	83	80	65	62	65	61
7	62	73	78	70	79	83	80	81	75	69	69	68
8	69	75	76	70	83	84	78	81	71	70	65	65
moyenne	70.5	80.5	81.0	73.125	81.75	84.0	78.625	79.75	72.75	72.375	71.5	71.25

2. Les modèles étudiés

On peut décrire les données de Milliken & Johnson avec plusieurs modèles qui vont nous permettre d'illustrer les nombreuses possibilités de la Proc MIXED.

Modèle 1

Le modèle suivi par les données s'écrit :

$$Y_{ijk} = \delta + \alpha_i + \beta_j + \gamma_{ij} + s_{k(i)} + \varepsilon_{ijk}$$

où :

- α_i , i = 1, 2, 3, correspond aux produits (AX23, BWW9, Contrôle),
- β_j , j = 1, …, 4 correspond aux instants de mesure ($T_1 = 5$, $T_2 = 10$, $T_3 = 15$, $T_4 = 20$),
- γ_{ij} , correspond aux termes d'interaction Produit*Temps,
- $s_{k(i)}$, k=1,…, 8 correspond à l'effet aléatoire Sujet(Produit),
- ε_{ijk} , terme résiduel correspondant à la j-ième mesure sur le k-ième sujet du groupe de traitement i.

Les variables aléatoires $s_{k(i)}$ suivent indépendamment une loi normale $N(0,\sigma_s^2)$, où σ_s^2 est la variance inter-sujets. Les variables aléatoires ε_{ijk} suivent indépendamment une loi normale $N(0,\sigma_\varepsilon^2)$, où σ_ε^2 est la variance intra-sujets. Les variables aléatoires $s_{k(i)}$ et ε_{ijk} sont indépendantes. On déduit de ce modèle que $\mathrm{Corr}(Y_{ijk} , Y_{i'j'k'}) = \dfrac{\sigma_s^2}{\sigma_s^2 + \sigma_\varepsilon^2}$ pour des mesures sur le même sujet, = 0 sinon. La structure des corrélations entre les mesures associées à un même sujet est donc de type CS (*Compound Symmetry*). Ce modèle est accessible par la Proc GLM.

Modèle 2

On reprend le modèle 1 en supposant que la variance intra-sujet dépend du traitement : le terme résiduel ε_{ijk} suit maintenant une loi normale $N(0,\sigma_{i\varepsilon}^2)$. On en déduit : $\mathrm{Corr}(Y_{ijk} , Y_{i'j'k'}) = \dfrac{\sigma_s^2}{\sigma_s^2 + \sigma_{i\varepsilon}^2}$

pour des mesures sur le même sujet suivant le traitement i, = 0 sinon.

Modèles 3

On modifie le modèle 1 en supprimant l'effet sujet, mais en autorisant les résidus ε_{ijk} associés à un même sujet à être corrélés entre eux selon une structure spécifiée par l'utilisateur.

$$Y_{ijk} = \delta + \alpha_i + \beta_j + \gamma_{ij} + \varepsilon_{ijk}$$

où :

$$\varepsilon_{i.k} = [\varepsilon_{i1k}, \varepsilon_{i2k}, \varepsilon_{i3k}, \varepsilon_{i4k}] \sim N(0,\Sigma)$$

La nature de la matrice Σ doit être définie par l'utilisateur. On trouve dans la Proc MIXED de nombreux types de matrice. Dans l'exemple nous utiliserons les types suivants : UN (Unstuctured), CS (Compound Symmetry), AR(1) (Autorégressif d'ordre 1). Le type CS correspond au modèle 1.
Le type UN est accessible par la Proc GLM, options MANOVA ou REPEATED.

Modèles 4

On reprend les modèles 3 en autorisant la matrice Σ à varier d'un groupe de traitement à l'autre.

3. Les résultats

Nous donnons dans l'annexe 2 le programme SAS permettant d'estimer les différents modèles proposés dans la section 2. La comparaison de ces modèles est réalisée dans le tableau 2.

Tableau 2 : Comparaison des modèles

Modèle	SBC	-2 Res Log Likelihood	$F_{Produit}$	F_{Temps}	$F_{Produit*Temps}$
1	-248.019	487.1769	5.95	12.95	12.16
2	-251.197	484.6700	6.08	12.81	13.97
3 UN	-260.545	476.7820	5.95	16.42	22.36
CS	-248.019	487.1769	5.95	12.95	12.16
AR(1)	-246.392	483.9218	6.46	16.04	13.54
4 UN	-292.281	451.6381	6.71	16.42	22.61
CS	-253.102	479.6197	6.71	12.95	13.94
AR(1)	-252.544	478.5030	7.42	15.69	13.73

Le critère du SBC maximum conduit au choix du modèle 3 avec une matrice de covariance **Σ** de type AR(1) : la corrélation entre les mesures réalisées sur un même sujet décroît avec l'intervalle de temps entre les mesures. De plus on peut admettre que la matrice **Σ** est la même pour les trois traitements. Cependant, la différence entre les SBC des modèles (1) et (3-AR(1)) étant faible, nous préférons poursuivre l'exposé avec le modèle (1) plus simple et en partie accessible par la Proc GLM.

4. Etude comparée du modèle (1) avec les procédures GLM et MIXED

4.1 Utilisation de la Proc GLM

Le tableau d'analyse de la variance associé au modèle 1, et fourni par la Proc GLM, est donné dans le tableau 3. On retrouve exactement les résultats des tests sur les effets fixes de la Proc MIXED.

Tableau 3 : Analyse de la variance du modèle (1)

Source de variation	Degrés de liberté	Carré moyen	E(Carré moyen)	F
Produit	2	657.54167	$\sigma_\varepsilon^2 + 4\sigma_s^2$+Q(Produit, Produit*Temps)	5.95
Temps	3	94.20486	σ_ε^2 + Q(Temps, Produit*Temps)	12.95
Produit*Temps	6	88.52778	σ_ε^2 + Q(Produit*Temps)	12.16
Sujet(Produit)	21	110.48363	$\sigma_\varepsilon^2 + 4\sigma_s^2$	15.182
Résidu	63	7.27728	σ_ε^2	

La présence de l'interaction Produit*Temps nous conduit à comparer les produits entre eux non pas globalement, mais à chaque instant. Notons μ_{ij} la réponse moyenne pour le traitement i à l'instant j. Supposons que le chercheur soit intéressé par la différence entre le produit AX23 et le Contrôle à l'instant T1. Il souhaite obtenir un intervalle de confiance à 95% de $\mu_{11} - \mu_{31}$.

Voici la demande Proc GLM :

```
proc glm data=coeur;
class sujet produit temps;
model rythme = produit temps produit*temps sujet(produit);
random sujet(produit)/test;
estimate 'mu11 vs mu31/GLM' produit 1 0 -1 produit*temps  1 0 0 0
                                                           0 0 0 0
                                                          -1 0 0 0;
run;
```

Et voici la réponse SAS :

```
Dependent Variable: RYTHME

                                    T for H0:    Pr > |T|   Std Error of
Parameter              Estimate    Parameter=0                Estimate

mu11 vs mu31/GLM    -2.25000000          -1.67     0.1003     1.34882187
```

La Proc GLM ne permet pas d'obtenir directement le calcul de l'intervalle de confiance, mais il sera possible de l'obtenir avec la Proc MIXED.

4.2 Utilisation de la Proc MIXED

Nous demandons maintenant à la Proc MIXED de calculer un intervalle de confiance de $\mu_{11} - \mu_{31}$.

Voici le code SAS :

```
proc mixed data=coeur;
class sujet produit temps;
model rythme = produit temps produit*temps;
random sujet(produit);
estimate 'mu11 vs mu31' produit 1 0 -1 produit*temps  1 0 0 0
                                                      0 0 0 0
                                                     -1 0 0 0 / cl;
run;
```

Et voici les résultats :

```
                              ESTIMATE Statement Results

 Parameter                 Estimate     Std Error    DF       t  Pr > |t|

 mu11 vs mu31           -2.25000000    2.87571161    63   -0.78    0.4369

                        Alpha     Lower     Upper

                         0.05   -7.9967    3.4967
```

4.3 Comparaison des résultats

Les données étant équilibrées on trouve la même estimation « naturelle » $\bar{y}_{11} - \bar{y}_{31} = 70.5–72.75 = -2.25$ de $\mu_{11} - \mu_{31}$ avec les deux procédures. Ce ne serait plus le cas si les données avaient été déséquilibrées. Par contre, on constate une différence importante au niveau de l'écart-type de la

différence : 1.3488 avec la proc GLM contre 2.8757 avec la Proc MIXED. Cette différence de résultats illustre parfaitement les limites de la Proc GLM. En effet, dans la Proc GLM, les effets aléatoires sont considérés comme fixes au niveau de l'estimation. Ceci entraîne que la fonction $\mu_{11} - \mu_{31} = \alpha_1 - \alpha_3 + \gamma_{11} - \gamma_{13}$ n'est pas estimable pour la Proc GLM. Par conséquent la Proc GLM complète l'instruction

```
estimate 'mu11 vs mu31/GLM' produit 1 0 -1 produit*temps  1 0 0 0
                                                           0 0 0 0
                                                          -1 0 0 0;
```

afin d'obtenir une fonction estimable.

L'instruction précédente crée en fait la fonction $\alpha_1 - \alpha_3 + \gamma_{11} - \gamma_{13} + \frac{1}{8}\sum_{k=1}^{8} s_{k(1)} - \frac{1}{8}\sum_{k=1}^{8} s_{k(3)}$. On vérifie facilement ce résultat en utilisant l'option E de ESTIMATE. Et si l'on essaye l'instruction

```
estimate 'mu11 vs mu31' produit 1 0 -1 produit*temps 1 0 0 0 0 0 0 0 -1 0 0 0
sujet(produit) 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0;
```

la réponse de SAS est :

```
mu11 vs mu31 is not estimable.
```

On peut maintenant demander à la Proc MIXED d'estimer la fonction estimable créée par la Proc GLM. On retrouve les résultats de la Proc GLM *plus* l'intervalle de confiance souhaité.

Voici l'instruction SAS :

```
proc mixed data=coeur;
class sujet produit temps;
model rythme = produit temps produit*temps;
random sujet(produit);
estimate 'mu11 vs mu31/narrow ' produit 8 0 -8 produit*temps  8 0 0 0 0 0 0 0 -8 0 0 0 |
sujet(produit) 1 1 1 1 1 1 1 1 0 0 0 0 0 0 0 0 -1 -1 -1 -1 -1 -1 -1 -1 / cl divisor=8;
run;
```

Voici les résultats :

```
                      ESTIMATE Statement Results

 Parameter                 Estimate     Std Error    DF       t  Pr > |t|

 mu11 vs mu31/narrow    -2.25000000    1.34882187    63   -1.67    0.1003

                         Alpha     Lower     Upper

                          0.05   -4.9454    0.4454
```

En conclusion, la Proc GLM fournit un intervalle de confiance de $\mu_{11} - \mu_{31}$ correspondant à une estimation « étroite » de $\mu_{11} - \mu_{31}$ (*narrow inference*), c'est à dire limitée aux sujets participant à l'expérimentation. On obtient l'intervalle [-4.9454 ; 0.4454]. Par contraste la Proc MIXED fournit un intervalle de confiance de $\mu_{11} - \mu_{31}$ correspondant à une estimation « large » de $\mu_{11} - \mu_{31}$ (*broad inference*), c'est à dire généralisable à toute la population étudiée. L'intervalle obtenu alors est évidemment plus large : [-7.9967 ; 3.467], mais il correspond en général à l'objectif réel du chercheur.

5. Etude comparée du modèle [(3), Σ de type UN] avec les Proc GLM et MIXED

5.1 Utilisation de la Proc GLM

Lorsqu'il n'y a pas de données manquantes, l'analyse de la variance multivariée disponible dans la Proc GLM permet d'étudier le modèle (3) en supposant la matrice Σ de type UN. Les tests des effets fixes réalisés en utilisant la statistique de Wilks sont plus précis que les tests issus de la Proc MIXED dans le sens suivant : les niveaux de signification calculés en utilisant la transformation de Rao sont exacts ou mieux approchés qu'en utilisant les tests F des effets fixes de la Proc MIXED (Roger & Kenward, 1993).

Rappelons brièvement le test de Wilks.

Notons μ_i le vecteur-ligne $(\mu_{i1}, \ldots, \mu_{i4})$, $Y_{k(i)} = (Y_{i1k}, \ldots, Y_{i4k})$ le vecteur-ligne des mesures réalisées sur le k-ième sujet du groupe i, et Y la matrice 24×4 formée des 24 $Y_{k(i)}$. On suppose ici que les $Y_{k(i)}$ sont indépendants entre eux et suivent chacun une loi multinormale $N(\mu_i , \Sigma)$. Notons enfin μ la matrice 3×4 formée des moyennes μ_{ij}.

La vraisemblance de l'ensemble des données s'écrit :

$$L(Y/\mu,\Sigma) = \prod_{i=1}^{m}\prod_{k=1}^{n_i} \frac{1}{(2\pi)^{d/2}|\Sigma|^{1/2}} \exp[-\frac{1}{2}(Y_{k(i)} - \mu_i)\Sigma^{-1}(Y_{k(i)} - \mu_i)']$$

où m est le nombre de groupes, n_i la taille du groupe i, et d le nombre de répétitions. Ici m = 3, $n_1 = n_2 = n_3 = 8$ et d = 4. On note N la somme de tous les n_i : N = 24.

Tous les tests disponibles dans la Proc GLM à travers les options REPEATED ou MANOVA sont de la forme H_0 : LμM = 0. La matrice L concerne les comparaisons inter-sujets, alors que la matrice M permet de construire les comparaisons intra-sujets. La statistique utilisée pour ce test est le Λ de Wilks définie par

$$\Lambda = \left[\frac{\underset{\mu:L\mu M=0}{\text{Maximum}}\ L(YM/\mu M, M'\Sigma M)}{\underset{\mu,\Sigma}{\text{Maximum}}\ L(YM/\mu M, M'\Sigma M)} \right]^{2/N}$$

où la vraisemblance L(YM / μM, M′ΣM) se déduit immédiatement de L(Y / μ, Σ).

On rejette H_0 pour Λ petit. La statistique F de Rao associée permet de calculer le niveau de signification du test :

$$F = \frac{1-\Lambda^{1/t}}{\Lambda^{1/t}} . \frac{rt-2u}{pq}$$

avec : p = rang (M)
q = rang (L)
v = N – m
r = v – (p – q + 1)/2
u = (pq –2)/4
$t = \sqrt{\frac{p^2q^2-4}{p^2+q^2-5}}$ si $p^2 + q^2 - 5 > 0$, = 1 sinon.

Lorsque l'hypothèse H_0 est vraie, la statistique F suit approximativement une loi de Fisher-Snedecor à pq et rt – 2u degrés de liberté. La loi est exacte si le minimum de (p, q) est inférieure ou égale à 2.

Applications

Test de l'effet Produit

Le test usuel consiste à comparer les moyennes par produit $\mu_{i.} = (\mu_{i1} + \ldots + \mu_{i4})/4$: $H_0 : \mu_{1.} = \mu_{2.} = \mu_{3.}$.

Soit l'hypothèse H_0 :

$$\begin{bmatrix} -1 & 1 & 0 \\ -1 & 0 & 1 \end{bmatrix} \begin{bmatrix} \mu_{11} & \mu_{12} & \mu_{13} & \mu_{14} \\ \mu_{21} & \mu_{22} & \mu_{23} & \mu_{24} \\ \mu_{31} & \mu_{32} & \mu_{33} & \mu_{34} \end{bmatrix} \begin{bmatrix} 0.25 \\ 0.25 \\ 0.25 \\ 0.25 \end{bmatrix} = 0$$

Voici les instructions SAS permettant de réaliser ce test :

```
data coeur2;
input  rythme1 rythme2 rythme3 rythme4 produit $;
cards;
72          86          81          77      AX23
78          83          88          81      AX23
71          82          81          75      AX23
72          83          83          69      AX23
66          79          77          66      AX23
74          83          84          77      AX23
62          73          78          70      AX23
69          75          76          70      AX23
85          86          83          80      BWW9
82          86          80          84      BWW9
71          78          70          75      BWW9
83          88          79          81      BWW9
86          85          76          76      BWW9
85          82          83          80      BWW9
79          83          80          81      BWW9
83          84          78          81      BWW9
69          73          72          74      Contrôle
66          62          67          73      Contrôle
84          90          88          87      Contrôle
80          81          77          72      Contrôle
72          72          69          70      Contrôle
65          62          65          61      Contrôle
75          69          69          68      Contrôle
71          70          65          65      Contrôle
;

proc glm data=coeur2;
class produit;
model rythme1-rythme4 = produit/nouni noint;
contrast 'mu1 = mu2 = mu3' produit -1 1 0, produit -1 0 1;
manova m=(.25 .25 .25 .25);
run;
```

L'instruction CONTRAST crée la matrice L et l'instruction MANOVA M= crée la matrice M.
Voici les résultats :

```
Statistic                    Value          F      Num DF    Den DF  Pr > F

Wilks' Lambda             0.63824026     5.9515          2        21  0.0090
```

Le niveau de signification du test est exact puisque q = 1.

Test de l'effet Temps

Le test usuel consiste à comparer les moyennes à chaque instant $\mu_{.j} = (\mu_{1j} + \mu_{2j} + \mu_{3j})/3$:
$H_0 : \mu_{.1} = \mu_{.2} = \mu_{.3} = \mu_{.4}$.

Soit l'hypothèse H_0 :

$$\begin{bmatrix}1 & 1 & 1\end{bmatrix}\begin{bmatrix}\mu_{11} & \mu_{12} & \mu_{13} & \mu_{14}\\ \mu_{21} & \mu_{22} & \mu_{23} & \mu_{24}\\ \mu_{31} & \mu_{32} & \mu_{33} & \mu_{34}\end{bmatrix}\begin{bmatrix}-1 & -1 & -1\\ 1 & 0 & 0\\ 0 & 1 & 0\\ 0 & 0 & 1\end{bmatrix} = 0$$

Voici les instructions SAS permettant de réaliser ce test :

```
proc glm data=coeur2;
class produit;
model rythme1-rythme4 = produit/nouni noint;
contrast 'mu.1 = mu.2 = mu.3 = mu.4' produit 1 1 1;
manova m=(-1 1 0 0,-1 0 1 0,-1 0 0 1);
run;
```

Voici les résultats :

```
Statistic                    Value          F      Num DF    Den DF  Pr > F

Wilks' Lambda            0.29886466    14.8580         3        19  0.0001
```

Le test est exact puisque p = 1.

*Test de l'interaction Produit*Temps*

Il s'agit de tester si les évolutions dans les différents groupes sont parallèles .
Soit : H_0 : $\mu_{21} - \mu_{11} = \mu_{22} - \mu_{12} = \mu_{23} - \mu_{13} = \mu_{24} - \mu_{14}$,
$\mu_{31} - \mu_{11} = \mu_{32} - \mu_{12} = \mu_{33} - \mu_{13} = \mu_{34} - \mu_{14}$
Soit encore H_0 :

$$\begin{bmatrix}-1 & 1 & 0\\ -1 & 0 & 1\end{bmatrix}\begin{bmatrix}\mu_{11} & \mu_{12} & \mu_{13} & \mu_{14}\\ \mu_{21} & \mu_{22} & \mu_{23} & \mu_{24}\\ \mu_{31} & \mu_{32} & \mu_{33} & \mu_{34}\end{bmatrix}\begin{bmatrix}-1 & -1 & -1\\ 1 & 0 & 0\\ 0 & 1 & 0\\ 0 & 0 & 1\end{bmatrix} = 0$$

Voici les instructions SAS :

```
proc glm data=coeur2;
class produit;
model rythme1-rythme4 = produit/nouni noint;
contrast 'Produit*Temps' produit -1 1 0, produit -1 0 1;
manova m=(-1 1 0 0,-1 0 1 0,-1 0 0 1);
run;
```

Voici les résultats :

```
Statistic                    Value          F      Num DF    Den DF  Pr > F

Wilks' Lambda            0.10831123    12.9107         6        38  0.0001
```

Le niveau de signification est exact puisque p = 2.

Comparaison entre μ_{11} et μ_{31}

Le test s'écrit :

$$\begin{bmatrix}1 & 0 & -1\end{bmatrix}\begin{bmatrix}\mu_{11} & \mu_{12} & \mu_{13} & \mu_{14}\\ \mu_{21} & \mu_{22} & \mu_{23} & \mu_{24}\\ \mu_{31} & \mu_{32} & \mu_{33} & \mu_{34}\end{bmatrix}\begin{bmatrix}1\\0\\0\\0\end{bmatrix} = 0$$

Voici le programme SAS :

```
proc glm data=coeur2;
class produit;
model rythme1-rythme4 = produit/nouni noint;
contrast 'mu11 - mu31' produit 1 0 -1;
manova m=(1 0 0 0);
run;
```

Voici les résultats :

```
Statistic                   Value          F      Num DF    Den DF  Pr > F

Wilks' Lambda            0.96937618     0.6634         1        21  0.4245
```

Le test est exact puisque p = q = 1.

5.2 Comparaison des résultats entre les Proc GLM multivariée et Proc MIXED

Les résultats de l'analyse de la variance multivariée sont ici exacts. Le tableau suivant montre que les résultats donnés par la Proc MIXED sont équivalents pour des comparaisons inter-sujets, mais nettement différents pour les comparaison intra-sujets.

Test	F GLM multivariée	Niveau de signification	F MIXED	Niveau de signification
Produit	5.95	0.0090	5.95	0.0090
Temps	14.85	0.0001	16.42	0.0001
Produit*Temps	12.91	0.0001	22.36	0.0001
$\mu_{11} - \mu_{31} = 0$	0.6634	0.4245	0.66	0.4245

6. Conclusion

Nous avons essentiellement considéré dans cet exposé le cas des mesures répétées. Quelle stratégie adopter pour choisir la bonne procédure SAS ? Les avantages de la Proc MIXED sont nombreux. La Proc MIXED permet de retrouver tous les résultats justes de l'approche univariée de la Proc GLM. Elle donne accès à des résultats généralisables à toute la population étudiée (*broad inference*) alors que la Proc GLM ne fournit, au niveau de l'estimation, que des résultats limités à l'échantillon étudié (*narrow inference*). La Proc MIXED propose une très grande variété de modèle au niveau du choix de la matrice de covariance Σ des mesures intra-individus. Mais il est sans doute préférable de s'appuyer sur une expérience large pour choisir le type de la matrice de covariance, plutôt que de chercher une structure s'adaptant bien aux données disponibles souvent limitées et volatiles. Mentionnons une propriété très importante de la Proc MIXED : sa capacité à prendre en compte des variances hétérogènes aussi bien pour les modèles à effets fixes que les modèles à effets mixtes.

La Proc GLM garde cependant un intérêt dans deux cas extrêmes. Lorsque les données sont complètes sur tous les individus, nous avons vu que l'approche multivariée donnait des tests plus exacts que la Proc MIXED pour les facteurs intra-sujets, lorsque la matrice de covariance est de type « Unstructured ». Lorsqu'il y a des cases vides (par exemple aucun sujet n'a essayé le traitement 1 au temps 1) la Proc GLM propose des tests particuliers accessibles par les sommes de carrés de type 4. Ces tests sont absents de la Proc MIXED, mais on peut les reconstituer à l'aide de contrastes.

Finalement les avantages de la Proc MIXED sur la Proc GLM me semblent considérables. Dans la plupart des applications s'appuyant sur des modèles mixtes, le chercheur devrait plus facilement construire les modèles adéquats en utilisant la Proc MIXED plutôt que la Proc GLM. Ajoutons enfin qu'elle est plus simple à utiliser et beaucoup plus complète que la Proc GLM pour l'analyse de modèles mixtes.

Références

1. Cook, N.R. (1998) : Restricted Maximum Likelihood. In *Encyclopedia of Biostatistics*, P. Armitage & T. Colton (Eds). John Wiley & Sons, vol. 5, 3827-3830.
2. Fai, A.H. & Cornelius, P.L. (1996) : Approximate F-tests of multiple degree of freedom hypotheses in generalized least squares analyses of unbalanced split-plot experiments. *Journal of Statistical Computing and Simulation*, 54, 363-378.
3. Littel, R.C., Milliken, G.A., Stroup, W.W. & Wolfinger, R.D. (1996) : *SAS System for Mixed Models*. Cary ; NC :SAS Institute Inc.
4. Milliken, G.A. & Johnson, D.E. (1984) : *Analysis of messy data*.. New York :Van Nostrand Reinhold.
5. Roger, J.H. & Kenward, M.G. (1993) : Repeated measures using proc mixed instead of proc glm. In *Proceedings of the First Annual South-East SAS Users Group conference*, Cary, NC, U.S.A. : SAS Institute, 199-208.
6. Verbeke, G. & Molenberghs, G. (1997) :*Linear Mixed Models in Practice. A SAS-Oriented Approach*. New York : Springer Verlag.

Annexe 1

Le modèle mixte : principaux résultats théoriques

Pour rendre la lecture de cette article plus facile nous avons rassemblé dans cette annexe les principaux résultats théoriques utiles à l'interprétation des sorties de la Proc MIXED. Nous avons essentiellement utilisé ici Cook (1998), l'annexe 1 de Littel, Milliken, Stroup & Wolfinger (1996) et l'annexe A de Verbeke & Molenberghs (1997).

Le modèle mixte s'écrit

$$y = X\beta + Zu + e$$

où X et Z sont des matrices décrivant le plan d'expériences, et où u et e sont respectivement des effets alétoires et les résidus. On suppose que le vecteur u suit une loi multinormale N(0,G), le vecteur e une loi multinormale N(0,R) et que les vecteurs u et e sont indépendants. On note V = ZGZ′ + R la matrice de covariance de y. Il en résulte que y suit une loi N(Xβ,V).

Estimation des paramètres du modèle

Pour estimer β on considère le logarithme de la fonction de vraisemblance :

$$\ell(\beta,V) = -\frac{n}{2}\log(2\pi) - \frac{1}{2}\log|V| - \frac{1}{2}(y - X\beta)'V^{-1}(y - X\beta)$$

L'estimateur du maximum de vraisemblance de β est donc obtenu par moindres carrés généralisés (Generalized Least Squares) :

$$\hat{\beta}_{GLS} = (X'V^{-1}X)^{-1}X'V^{-1}y$$

Il faut ensuite remplacer la matrice V = ZGZ′ + R par son estimation. Deux méthodes permettent d'obtenir des estimations $\hat{G}$ et $\hat{R}$ de G et R : la méthode du maximum de vraisemblance (ML = Maximum Likelihood) et celle du maximum de vraisemblance restreint (REML = Restricted Maximum Likelihood). Les logarithmes des fonctions de vraisemblance maximisés sont respectivement, à une constante près :

$$ML : \ell(G,R) = -\frac{1}{2}\log|V| - \frac{1}{2}r'V^{-1}r$$

$$REML : \ell_R(G,R) = -\frac{1}{2}\log|V| - \frac{1}{2}\log|X'V^{-1}X| - \frac{1}{2}r'V^{-1}r$$

où $r = y - X\hat{\beta}_{GLS}$. L'option REML est habituellement préférable car elle permet de retrouver exactement les résultats de la Proc GLM lorsque les données dont équilibrées. Il y a une exception : dans la construction de tests LRT sur les effets fixes β, il est préférable d'utiliser l'option ML car la fonction maximisée ne dépend pas de X.

On obtient ensuite l'estimation finale de β en remplaçant V par $\hat{V} = Z\hat{G}Z' + \hat{R}$ dans $\hat{\beta}_{GLS}$:

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

On estime ensuite le vecteur des effets aléatoires u en utilisant la formule :

(A.2) $$\hat{u} = \hat{G}Z'\hat{V}^{-1}(y - X\hat{\beta})$$

Le vecteur $\hat{\beta}$ est un estimateur EBLUE (*Empirical Best Linear Unbiased Estimator*) de β, et le vecteur $\hat{u}$ est un prédicteur EBLUP (*Empirical Best Linear Unbiased Predictor*) de u.

Intervalle de confiance et test d'hypothèse

On considère qu'une fonction des paramètres Kβ + Mu est prédictible si la fonction Kβ est estimable. On parle d'inférence large (*broad inference*) lorsque M = 0, d'inférence étroite (*narrow inference*)

lorsque M porte sur l'ensemble des effets aléatoires, et d'inférence intermédiaire (*intermediate inference*) lorsque M ne porte que sur une partie des effets aléatoires. Dans le premier cas l'inférence est réalisée au niveau de la population, dans le deuxième cas elle est restreinte à l'échantillon des effets observés, et dans le troisième cas conditionnellement aux effets sélectionnés par M.

La variance de l'estimateur $K\hat{\beta} + M\hat{u}$ d'une fonction prédictible $K\beta + Mu$ est donnée par la formule $L\hat{C}L'$, où $L = [K,M]$ et

$$\hat{C} = \begin{bmatrix} X'\hat{R}^{-1}X & X'\hat{R}^{-1}Z \\ Z'\hat{R}^{-1}X & Z'\hat{R}^{-1}Z + \hat{G}^{-1} \end{bmatrix}^{-}$$

Lorsque β est estimable on a

(A.3) $$\text{Var}(\hat{\beta}) = (X'\hat{V}^{-1}X)^{-1}$$

C'est la matrice de covariance de $\hat{\beta}$ obtenu par moindres carrés généralisés.

Lorsque $K\beta$ est estimable, on peut tester l'hypothèse $H_0 : K\beta + Mu = 0$.

Lorsque L est un vecteur ligne, on utilise la statistique

$$t = \frac{K\hat{\beta} + M\hat{u}}{\sqrt{L\hat{C}L'}}$$

La statistique t suit approximativement une loi de Student avec un nombre de degrés de liberté ν fixé en utilisant la formule de Satterthwaite présentée plus bas. On peut aussi construire un intervalle de confiance au niveau $1-\alpha$ en utilisant la formule

$$K\hat{\beta} + M\hat{u} \pm t_{1-\alpha/2}(\nu)\sqrt{L\hat{C}L'}$$

Lorsque le rang de L est supérieur à 1, on utilise la statistique

$$F = \frac{[\hat{\beta}', \hat{u}']L'(L\hat{C}L')^{-1}L\begin{bmatrix}\hat{\beta} \\ \hat{u}\end{bmatrix}}{\text{rang}(L)}$$

qui suit, approximativement, sous H_0 une loi de Fisher-Snedecor à rang(L) degrés de liberté pour le numérateur et ν degrés de liberté pour le numérateur. On peut remarquer que $F = t^2$ lorsque le rang de L vaut 1.

La formule de Satterthwaite

L'approche moindres carrés de l'analyse de la variance à effets mixtes comprend trois étapes. Tout d'abord on estime le modèle en supposant tous les facteurs fixes et on réalise l'analyse de la variance habituelle à cette situation. Ensuite on calcule les espérances des carrés moyens obtenus en tenant compte de la nature fixe où aléatoire des facteurs. Enfin, on construit les statistiques F adéquates en

examinant les carrés moyens espérés exprimés en fonction des variances des facteurs aléatoires et du résidu. Lorsque le dénominateur du F peut se mettre sous la forme d'une combinaison linéaire de carrés moyens $\sum_{i=1}^{s} a_i CM_i$, alors le nombre de degrés de liberté du dénominateur du F est approché par

$$\nu = \frac{(a_1 CM_1 + ... + a_s CM_s)^2}{\frac{(a_1 CM_1)^2}{dl_1} + ... + \frac{(a_s CM_s)^2}{dl_s}}$$

où dl_i est le nombre de degrés de liberté du carré moyen CM_i.
On trouve dans la Proc Mixed la procédure de calcul de ν proposée par Fai & Cornelius (1996) et généralisant la formule de Satterthwaite à des situations plus générales. Elle est accessible via l'option DDFM = Satterth. Elle permet de retrouver les résultats de la Proc GLM lorsqu'on utilise la Proc MIXED sur des modèles accessibles par la Proc GLM. Elle peut fonctionner mal sur des modèles un peu trop complexes.

Comparaisons multiples

On trouve dans la Proc MIXED des procédures de comparaisons multiples des moyennes ajustées (c'est à dire estimées à l'aide du modèle étudié). On peut citer en particulier les méthodes de Scheffé, Tukey et Dunnett qui ont été adaptées aux modèles à effets mixtes.

Annexe 2

Utilisation de la Proc MIXED : le programme SAS

```
data coeur;
input sujet produit temps rythme;
cards;
1 1 1 72
⋮
8 3 4 65
;

proc mixed data=coeur;
```

```
title 'modèle 1';
class sujet produit temps;
model rythme = produit temps produit*temps;
random sujet(produit);
run;

proc mixed data=coeur;
title 'modèle 2';
class sujet produit temps;
model rythme = produit temps produit*temps;
random sujet(produit);
repeated / group=produit;
run;

proc mixed data=coeur;
title 'modèle 3-UN';
class sujet produit temps;
model rythme = produit temps produit*temps;
repeated temps/ subject=sujet(produit) type=UN;
run;

proc mixed data=coeur;
title 'modèle 4-UN';
class sujet produit temps;
model rythme = produit temps produit*temps;
repeated temps/ subject=sujet(produit) type=UN group=produit;
run;

proc mixed data=coeur;
title 'modèle 3-CS';
class sujet produit temps;
model rythme = produit temps produit*temps;
repeated temps/ subject=sujet(produit) type=CS;
run;

proc mixed data=coeur;
title 'modèle 4-CS';
class sujet produit temps;
model rythme = produit temps produit*temps;
repeated temps/ subject=sujet(produit) type=CS  group=produit;
run;
```

```
proc mixed data=coeur;
title 'modèle 3-AR(1)';
class sujet produit temps;
model rythme = produit temps produit*temps;
repeated temps/ subject=sujet(produit) type=AR(1);
run;

proc mixed data=coeur;
title 'modèle 4-AR(1)';
class sujet produit temps;
model rythme = produit temps produit*temps;
repeated temps/ subject=sujet(produit) type=AR(1) group=produit;
run;
```