

Exemple 6.10 : Illustration des étapes de spécification d'un modèle multiniveau complet (reprise de l'exemple de modélisation du jugement des enseignants)

resm= jugement que portent les enseignants sur la valeur scolaire de leurs élèves en mathématiques

scom= scores obtenus par les élèves aux épreuves nationales de CE2 en mathématiques (performances scolaires effectives de début d'année)

cscom= variable scom centrée sur la moyenne des scores obtenus par les élèves aux épreuves nationales de CE2 en mathématiques

mscom= score moyen de la classe aux évaluations nationales de début de CE2 en mathématiques

parti= enfants dont le père est artisan

pinter= enfants dont le père exerce une profession intermédiaire

pempl= enfants dont le père est employé

pouvr= enfants dont le père est ouvrier

pautr= enfants qui ne peuvent être intégrés dans les modalités précédentes

retard= élèves en retard scolaire

tott1bis= score d'internalité des élèves

Syntaxe du modèle 1 (modèle vide):

```
proc mixed data=work.jugement noclprint covtest;  
class classe;  
model resm=/solution ddfm=bw chisq;  
random intercept /sub=classe;  
run;
```

Syntaxe détaillée:

Ligne 20: L'option « noclprint » indique que l'on ne souhaite pas que les identifiants des différents groupes apparaissent dans la fenêtre des sorties SAS, autrement dit, nous n'aurons pas dans les sorties SAS les différents identifiants attribués à chaque classe.

Ligne 20 : L'option « covtest » permet d'estimer les composants de covariance (lorsque les variances sont significativement différentes de 0).

31 *Ligne 21* : L'énoncé « class » indique que la variable à traiter est catégorielle. Dans cet
32 exemple, il s'agit de la variable « classe » qui représente l'identifiant de chacune des classes
33 de l'échantillon (rappelons ici que l'échantillon est composé de 17 classes).

34 *Ligne 22* : L'énoncé « model » indique les effets fixes à estimer. Dans le cas d'un modèle vide,
35 il n'y a pas d'autres effets fixes à estimer que celui de la constante. Nous verrons dans les
36 modèles suivants, qui ont d'autres effets fixes à estimer, que les noms des variables seront
37 derrière le signe « = ».

38 *Ligne 22* : L'option « /solution » précise que l'on souhaite qu'une estimation des effets fixes
39 figure dans les sorties SAS.

40 *Ligne 22* : L'option « ddfm=bw » spécifie la manière avec laquelle les DDL seront calculés
41 (dans cet exemple « bw » correspond à « between and within method »).

42 *Ligne 22* : L'option « chisq » permet de spécifier que l'on souhaite que les valeurs du χ^2
43 apparaissent dans la fenêtre des sorties SAS qui présente les « tests des effets fixes ».

44 *Ligne 23* : L'énoncé « random » indique les effets aléatoires à estimer. Cet énoncé indique
45 que l'on autorise la constante à varier d'une classe à l'autre.

46 *Ligne 23* : L'option « /sub=classe » indique quelles sont les unités statistiques de niveau 2 (les
47 classes) dans lesquelles sont regroupées les unités statistiques de niveau 1 (les élèves).

48 Sorties SAS :

49

51
52
53
54
55
56
57
58
59
60
61
62
63
64
65
66
67
68
69
70
71
72
73
74
75
76
77
78
79
80
81
82
83
84
85
86
87
88
89
90
91
92
93
94
95
96
97
98
99
100
101
102
103
104
105
106
107
108
109
110
111
112
113
114
115
116

Procédure Mixed					
Informations sur le modèle					
Table	WORK.JUGEMENT				
Variable dépendante	RESM				
Structure de covariance	Variance Components				
Effet du sujet	CLASSE				
Méthode d'estimation	REML				
Méthode de variance résiduelle	Profil				
Méthode SE des effets fixes	Basé(e) sur le modèle				
Méthode des degrés de liberté	Between-Within				
Dimensions					
Paramètres de covariance	2				
Colonnes dans X	1				
Colonnes dans Z par sujet	1				
Sujets	17				
Max. obs. par sujet	27				
Nombre d'observations					
Nombre d'observations lues	342				
Nombre d'observations utilisées	342				
Nombre d'observations non utilisées	0				
Historique des itérations					
Itération	Evaluations	-2 Log-vrais. restreinte	Critère		
0	1	1456.87405564			
1	3	1434.89969271	0.00006175		
2	1	1434.87310607	0.00000070		
3	1	1434.87281973	0.00000000		
Correspond aux critères de convergence.					
Valeurs estimées des paramètres de covariance					
Param de cov	Sujet	Valeur estimée	Erreur type	Valeur Z	Pr > Z
Intercept	CLASSE	0.4730	0.2278	2.08	0.0189
Residual		3.6481	0.2856	12.78	<.0001
Statistiques d'ajustement					
		-2 log-vraisemblance rest	1434.9		
		AIC (préférer les petites	1438.9		
		AICC (préférer les petite	1438.9		
		BIC (préférer les petites	1440.5		
Le Système SAS					
Procédure Mixed					
Solution pour effets fixes					
Effet	Valeur estimée	Erreur type	DDL	Valeur du test t	Pr > t
Intercept	6.6113	0.1990	16	33.22	<.0001

117 Lignes 51 à 88 : Résumés d'informations concernant le modèle.

118 Ligne 90 : Indication concernant la convergence du modèle. Dans cet exemple, nous pouvons
119 constater que le modèle correspond aux critères de convergence.

120 Ligne 97 : Dans la colonne « valeur estimée » nous avons l'estimation de la variance
121 interclasse au niveau de la constante (niveau 2). Cette estimation indique qu'il existe des

122 différences significatives du niveau moyen du jugement des enseignants entre les différences
123 classes.

124 *Ligne 98 : Dans la colonne « valeur estimée » nous avons l'estimation de la variance inter-*
125 *élèves. Autrement dit, nous constatons des différences de jugement des enseignants à*
126 *l'intérieur des classes.*

127 *Ligne 116 : Dans la colonne « valeur estimée », nous avons l'estimation moyenne de la*
128 *constante. Cette estimation correspond aux jugements que les enseignants portent sur la valeur*
129 *scolaire de leurs élèves en mathématiques, toutes classes confondues.*

130 *Avant de présenter la syntaxe du modèle 2, une étape doit être conduite, celle de la centration*
131 *de la variable « scom ». En effet, dans cet exemple, les scores obtenus par les élèves aux*
132 *épreuves nationales de CE2 en mathématiques ont été centrés autour de la moyenne de*
133 *l'échantillon. Cela revient à soustraire à chaque observation de « scom » la moyenne obtenue*
134 *par l'ensemble de l'échantillon (41.19). Le programme SAS suivant permet d'effectuer cette*
135 *centration.*

136 Syntaxe pour la centration de « scom » autour de la moyenne de l'échantillon :

```
137 data work.jugement2;set work.jugement;  
138 cscom = scom-41.19;  
139 inter_niv=cscom*mscom;  
140 run;
```

cette syntaxe permet de créer un fichier de données temporaire dans lequel la variable « scom » est centrée autour de la moyenne de l'échantillon. La variable « scom » centrée se nomme « cscom ».

142 Syntaxe du modèle 2 :

```
143 proc mixed data=work.jugement2 noclprint covtest;  
144 class classe;  
145 model resm= cscom mscom/solution ddfm=bw chisq;  
146 random intercept /sub=classe;  
147 run;
```

nous travaillons à présent à partir du fichier créé précédemment « work.jugement2 »

introduction de deux variables explicatives dans l'énoncé « model » : cscom et mscom

151 Sorties SAS :

152

153
154
155
156
157
158
159
160
161
162
163
164
165
166
167
168
169
170
171
172
173
174
175
176
177
178
179
180
181
182
183
184
185
186
187
188
189
190
191
192
193
194
195
196
197
198
199
200
201
202
203
204
205
206
207
208
209
210
211
212
213
214

Procédure Mixed					
Informations sur le modèle					
Table	WORK.JUGEMENT2				
Variable dépendante	RESM				
Structure de covariance	Variance Components				
Effet du sujet	CLASSE				
Méthode d'estimation	REML				
Méthode de variance résiduelle	Profil				
Méthode SE des effets fixes	Basé(e) sur le modèle				
Méthode des degrés de liberté	Between-Within				
Dimensions					
Paramètres de covariance	2				
Colonnes dans X	3				
Colonnes dans Z par sujet	1				
Sujets	17				
Max. obs. par sujet	27				
Nombre d'observations					
Nombre d'observations lues	342				
Nombre d'observations utilisées	342				
Nombre d'observations non utilisées	0				
Historique des itérations					
Itération	Evaluations	-2 Log-vrais. restreinte			Critère
0	1	1263.29563329			
1	3	1224.33177272			0.00009570
2	1	1224.30093857			0.00000133
3	1	1224.30053519			0.00000000
Correspond aux critères de convergence.					
Valeurs estimées des paramètres de covariance					
Param de cov	Sujet	Valeur estimée	Erreur type	Valeur Z	Pr > Z
Intercept	CLASSE	0.4045	0.1818	2.22	0.0130
Residual		1.8692	0.1467	12.74	<.0001
Statistiques d'ajustement					
		-2 log-vraisemblance rest	1224.3		
		AIC (préférer les petites	1228.3		
		AICC (préférer les petite	1228.3		
		BIC (préférer les petites	1230.0		

Procédure Mixed						
Solution pour effets fixes						
Effet	Valeur estimée	Erreur type	DDL	Valeur du test t	Pr > t	
Intercept	9.3635	1.2127	15	7.72	<.0001	
cscm	0.1274	0.007208	324	17.67	<.0001	
mscm	-0.06607	0.02931	15	-2.25	0.0396	
Type 3 Tests des effets fixes						
Effet	DDL Num.	DDL Res.	Khi-2	Valeur F	Pr > Khi-2	Pr > F
cscm	1	324	312.24	312.24	<.0001	<.0001
mscm	1	15	5.08	5.08	0.0242	0.0396

Ligne 225: Les estimations indiquent une relation positive et significative entre le jugement des enseignants et les scores obtenus par les élèves aux épreuves nationales de CE2 en mathématiques. Autrement dit, plus l'élève est fort au sein de sa classe, mieux il est jugé.

Ligne 226 : Les estimations indiquent une relation négative et significative entre le jugement des enseignants et le niveau moyen de la classe aux évaluations nationales de CE2 en mathématiques (effet de contexte significatif et négatif). Autrement dit, à score individuel donné, les élèves sont d'autant moins bien jugés que leur classe est plus forte.

Syntaxe du modèle 3:

```
proc mixed data=work.jugement2 noclprint covtest;
class classe;
model resm= cscm/mscm/solution ddfm=bw chisq;
random intercept cscm/sub=classe type=un;
run;
```

indique que l'on autorise une variation aléatoire des pentes de régression du jugement sur les scores des élèves

L'option « type=un » indique que l'on souhaite spécifier dans notre modélisation une matrice de covariances constantes-pentes non structurée (*unstructured*)

Sorties SAS :

254
255
256
257
258
259
260
261
262
263
264
265
266
267
268
269
270
271
272
273
274
275
276
277
278
279
280
281
282
283
284
285
286
287
288
289
290
291
292
293
294
295
296
297
298
299
300
301
302
303
304
305
306
307
308
309
310
311
312

Procédure Mixed					
Informations sur le modèle					
Table	WORK.JUGEMENT2				
Variable dépendante	RESM				
Structure de covariance	Unstructured				
Effet du sujet	CLASSE				
Méthode d'estimation	REML				
Méthode de variance	Profil				
résiduelle					
Méthode SE des	Basé(e) sur le modèle				
effets fixes					
Méthode des degrés	Between-Within				
de liberté					
Dimensions					
Paramètres de covariance	4				
Colonnes dans X	3				
Colonnes dans Z par sujet	2				
Sujets	17				
Max. obs. par sujet	27				
Nombre d'observations					
Nombre d'observations lues	342				
Nombre d'observations utilisées	342				
Nombre d'observations non utilis	0				
Historique des itérations					
Itération	Evaluations	-2 Log-vrais. restreinte		Critère	
0	1	1263.29563329			
1	2	1206.22524297		0.00047903	
2	1	1206.06698445		0.00002724	
3	1	1206.05872359		0.00000011	
4	1	1206.05869167		0.00000000	
Correspond aux critères de convergence.					
Valeurs estimées des paramètres de covariance					
Param de cov	Sujet	Valeur estimée	Erreur type	Valeur Z	Pr Z
UN(1,1)	CLASSE	0.3419	0.1589	2.15	0.0157
UN(2,1)	CLASSE	-0.01156	0.009376	-1.23	0.2175
UN(2,2)	CLASSE	0.001672	0.000886	1.89	0.0296
Residual		1.6944	0.1362	12.44	<.0001

313
314
315
316
317
318
319
320
321
322
323
324
325
326
327
328
329
330
331
332
333
334
335
336
337
338
339
340
341
342
343
344
345
346
347
348

Statistiques d'ajustement						
-2 log-vraisemblance rest	1206.1					
AIC (préférer les petites	1214.1					
AICC (préférer les petite	1214.2					
BIC (préférer les petites	1217.4					
Test du rapport de vraisemblance du modèle nul						
DDL	Khi-2	Pr > Khi-2				
3	57.24	<.0001				
Solution pour effets fixes						
Effet	Valeur estimée	Erreur type	DDL	Valeur du test t	Pr > t	
Intercept	10.9249	1.3530	15	8.07	<.0001	
cscm	0.1302	0.01246	324	10.45	<.0001	
mscm	-0.1024	0.03225	15	-3.17	0.0063	
Type 3 Tests des effets fixes						
Effet	DDL Num.	DDL Res.	Khi-2	Valeur F	Pr > Khi-2	Pr > F
cscm	1	324	109.10	109.10	<.0001	<.0001
mscm	1	15	10.07	10.07	0.0015	0.0063

349

- 350 *Ligne 306* : Indique l'estimation de la variance inter-classes au niveau de la constante.
- 351 *Ligne 307* : Indique l'estimation de la covariance constantes-pentes. La colonne Pr Z indique
- 352 que ce paramètre n'est pas significatif. Il sera par conséquent retiré du modèle suivant.
- 353 *Ligne 308* : Indique l'estimation de la variance inter-classes au niveau de la pente.

354 Syntaxe du modèle 4:

```
355 proc mixed data=work.jugement2 noclprint covtest;
356 class classe;
357 model resm= cscm/mscm/solution ddfm=bw chisq;
358 random intercept cscm/sub=classe;
359 run;
```

cette syntaxe reprend celle du modèle 3
mais on a retiré l'option « type=un » car
le paramètre estimant la covariance
pentes-constantes n'était pas significatif

360 Sorties SAS :

361

362
363
364
365
366
367
368
369
370
371
372
373
374
375
376
377
378
379
380
381
382
383
384
385
386
387
388
389
390
391
392
393
394
395
396
397
398
399
400
401
402
403
404
405
406
407
408
409
410
411
412
413
414
415
416
417
418
419
420
421
422
423
424
425
426

Procédure Mixed					
Informations sur le modèle					
Table	WORK.JUGEMENT2				
Variable dépendante	RESM				
Structure de covariance	Variance Components				
Effet du sujet	CLASSE				
Méthode d'estimation	REML				
Méthode de variance	Profil				
résiduelle					
Méthode SE des	Basé(e) sur le modèle				
effets fixes					
Méthode des degrés	Between-Within				
de liberté					
Dimensions					
Paramètres de covariance	3				
Colonnes dans X	3				
Colonnes dans Z par sujet	2				
Sujets	17				
Max. obs. par sujet	27				
Nombre d'observations					
Nombre d'observations lues	342				
Nombre d'observations utilisées	342				
Nombre d'observations non utilisées	0				
Historique des itérations					
Itération	Evaluations	-2 Log-vrais. restreinte			Critère
0	1	1263.29563329			
1	3	1207.91984259			0.00053995
2	1	1207.74096233			0.00003155
3	1	1207.73134215			0.00000015
4	1	1207.73129933			0.00000000
Correspond aux critères de convergence.					
Valeurs estimées des paramètres de covariance					
Param de cov	Sujet	Valeur estimée	Erreur type	Valeur Z	Pr > Z
Intercept	CLASSE	0.3388	0.1572	2.16	0.0156
cscm	CLASSE	0.001744	0.000911	1.91	0.0278
Residual		1.6922	0.1358	12.46	<.0001
Statistiques d'ajustement					
-2 log-vraisemblance rest				1207.7	
AIC (préférer les petites				1213.7	
AICC (préférer les petite				1213.8	
BIC (préférer les petites				1216.2	

427
428
429
430
431
432
433
434
435
436
437
438
439
440
441
442
443
444

Solution pour effets fixes						
Effet	Valeur estimée	Erreur type	DDL	Valeur du test t	Pr > t	
Intercept	10.2249	1.3138	15	7.78	<.0001	
cscm	0.1297	0.01272	324	10.19	<.0001	
mscom	-0.08533	0.03137	15	-2.72	0.0158	

Type 3 Tests des effets fixes						
Effet	DDL Num.	DDL Res.	Khi-2	Valeur F	Pr > Khi-2	Pr > F
cscm	1	324	103.88	103.88	<.0001	<.0001
mscom	1	15	7.40	7.40	0.0065	0.0158

445

446 *Dans la section « Valeurs estimées des paramètres de covariance », l'estimation*
447 *UN(2,1) ne figure plus car nous avons retiré de notre modèle l'option « type=un ».*

448 Syntaxe du modèle 5 (modèle complet):

```
449 proc mixed data=work.jugement2 noclprint covtest;
450 class classe;
451 model resm= cscm mscom parti pinter pempl pouvtr pautr retard
452 tottlbis/solution ddfm=bw chisq;
453 random intercept cscm/sub=classe solution;
454 run;
```

455

456

457 *Lignes 451 et 452: Intégration dans le modèle des variables explicatives*
458 *considérées dans le chapitre 3*

459

cette option permet de spécifier dans le modèle les solutions pour les effets aléatoires.

460
461
462
463
464
465
466
467
468
469
470
471
472
473
474
475
476
477
478
479
480
481
482
483
484
485
486
487
488
489
490
491
492
493
494
495
496
497
498
499
500
501
502
503
504
505
506
507
508
509
510
511
512
513
514
515
516
517
518
519

Procédure Mixed					
Informations sur le modèle					
Table	WORK.JUGEMENT2				
Variable dépendante	RESM				
Structure de covariance	Variance Components				
Effet du sujet	CLASSE				
Méthode d'estimation	REML				
Méthode de variance résiduelle	Profil				
Méthode SE des effets fixes	Basé(e) sur le modèle				
Méthode des degrés de liberté	Between-Within				
Dimensions					
Paramètres de covariance	3				
Colonnes dans X	10				
Colonnes dans Z par sujet	2				
Sujets	17				
Max. obs. par sujet	27				
Nombre d'observations					
Nombre d'observations lues	342				
Nombre d'observations utilisées	342				
Nombre d'observations non utilis	0				
Historique des itérations					
Itération	Evaluations	-2 Log-vrais. restreinte	Critère		
0	1	1237.07452050			
1	3	1190.74190283	0.00056892		
2	1	1190.55573137	0.00003057		
3	1	1190.54649601	0.00000013		
4	1	1190.54645763	0.00000000		
Correspond aux critères de convergence.					
Valeurs estimées des paramètres de covariance					
Param de cov	Sujet	Valeur estimée	Erreur type	Valeur Z	Pr > Z
Intercept	CLASSE	0.2680	0.1293	2.07	0.0191
cscom	CLASSE	0.001807	0.000921	1.96	0.0248
Residual		1.5876	0.1289	12.31	<.0001
Statistiques d'ajustement					
		-2 log-vraisemblance rest	1190.5		
		AIC (préférer les petites	1196.5		
		AICC (préférer les petite	1196.6		
		BIC (préférer les petites	1199.0		

520
521
522
523
524
525
526
527
528
529
530
531
532
533
534
535
536
537
538
539
540
541
542
543
544
545
546
547
548
549
550
551
552
553
554
555
556
557
558
559
560
561
562
563
564
565
566
567
568
569
570
571
572
573
574
575

Solution pour effets fixes						
Effet		Valeur estimée	Erreur type	DDL	Valeur du test t	Pr > t
Intercept		9.7547	1.2899	15	7.56	<.0001
cscm		0.1208	0.01290	317	9.37	<.0001
mscom		-0.08269	0.02952	15	-2.80	0.0134
parti		-0.1781	0.2994	317	-0.59	0.5525
pinter		-0.2394	0.2476	317	-0.97	0.3343
pempl		-0.2448	0.2334	317	-1.05	0.2951
pouvr		-0.5880	0.2209	317	-2.66	0.0082
pautr		-0.3252	0.2684	317	-1.21	0.2266
retard		-0.7002	0.2073	317	-3.38	0.0008
TOTT1BIS		0.07881	0.03148	317	2.50	0.0128
Solution pour effets aléatoires						
Effet	CLASSE	Valeur estimée	Err type préd	DDL	Valeur du test t	Pr > t
Intercept	2	-0.06394	0.2552	332	-0.25	0.8023
cscm	2	-0.02307	0.02309	332	-1.00	0.3186
Intercept	3	-0.6076	0.2573	332	-2.36	0.0188
cscm	3	-0.02356	0.02226	332	-1.06	0.2907
Intercept	4	0.5189	0.2674	332	1.94	0.0532
cscm	4	-0.00613	0.02357	332	-0.26	0.7948
Intercept	5	-0.07084	0.3416	332	-0.21	0.8359
cscm	5	0.004986	0.02265	332	0.22	0.8259
Intercept	6	-0.2027	0.2912	332	-0.70	0.4868
cscm	6	0.01077	0.02484	332	0.43	0.6649
Intercept	7	-0.1036	0.2753	332	-0.38	0.7068
cscm	7	0.03804	0.02332	332	1.63	0.1038
Intercept	8	-0.4829	0.2683	332	-1.80	0.0728
cscm	8	-0.04579	0.02288	332	-2.00	0.0462
Intercept	9	0.08150	0.3000	332	0.27	0.7860
cscm	9	-0.03522	0.02538	332	-1.39	0.1662
Intercept	10	-0.1304	0.3934	332	-0.33	0.7405
cscm	10	0.08702	0.02296	332	3.79	0.0002
Intercept	21	-0.09209	0.2810	332	-0.33	0.7433
cscm	21	-0.02354	0.02671	332	-0.88	0.3788
Intercept	22	0.9744	0.2812	332	3.47	0.0006
cscm	22	-0.01949	0.02561	332	-0.76	0.4471
Intercept	23	0.4969	0.2963	332	1.68	0.0944
cscm	23	-0.03475	0.02091	332	-1.66	0.0974
Intercept	24	-0.1228	0.4653	332	-0.26	0.7920
cscm	24	0.01199	0.03326	332	0.36	0.7188
Intercept	25	-0.7400	0.2910	332	-2.54	0.0114
cscm	25	0.05412	0.02828	332	1.91	0.0565
Intercept	26	0.1384	0.2828	332	0.49	0.6248
cscm	26	-0.00169	0.02371	332	-0.07	0.9432
Intercept	27	0.08903	0.4375	332	0.20	0.8389

Solution pour effets aléatoires						
Effet	CLASSE	Valeur estimée	Err type préd	DDL	Valeur du test t	Pr > t
cscm	27	-0.01381	0.03716	332	-0.37	0.7104
Intercept	28	0.3177	0.2989	332	1.06	0.2885
cscm	28	0.02014	0.02596	332	0.78	0.4386

Type 3 Tests des effets fixes						
Effet	DDL Num.	DDL Res.	Khi-2	Valeur F	Pr > Khi-2	Pr > F
cscm	1	317	87.74	87.74	<.0001	<.0001
mscm	1	15	7.85	7.85	0.0051	0.0134
parti	1	317	0.35	0.35	0.5520	0.5525
pinter	1	317	0.93	0.93	0.3336	0.3343
pempl	1	317	1.10	1.10	0.2943	0.2951
pouvr	1	317	7.08	7.08	0.0078	0.0082
pautr	1	317	1.47	1.47	0.2257	0.2266
retard	1	317	11.41	11.41	0.0007	0.0008
TOTT1BIS	1	317	6.27	6.27	0.0123	0.0128

Lignes 527 à 535: Cette partie fixe du modèle complet indique que toutes choses égales par ailleurs, les enfants d'ouvriers sont moins bien jugés que les enfants de cadre supérieur (*ligne 532*), les élèves en retard scolaire sont moins bien jugés que ceux qui sont « à l'heure » scolairement » (*ligne 534*) et les élèves qui se montrent plus internes dans leurs explications sont mieux jugés par l'enseignant (*ligne 535*). L'effet positif du score individuel en mathématiques (*ligne 527*) et l'effet négatif du score agrégé (*ligne 528*) restent significatifs malgré l'introduction de nouvelles variables explicatives.