

Modèles pour données ordinales

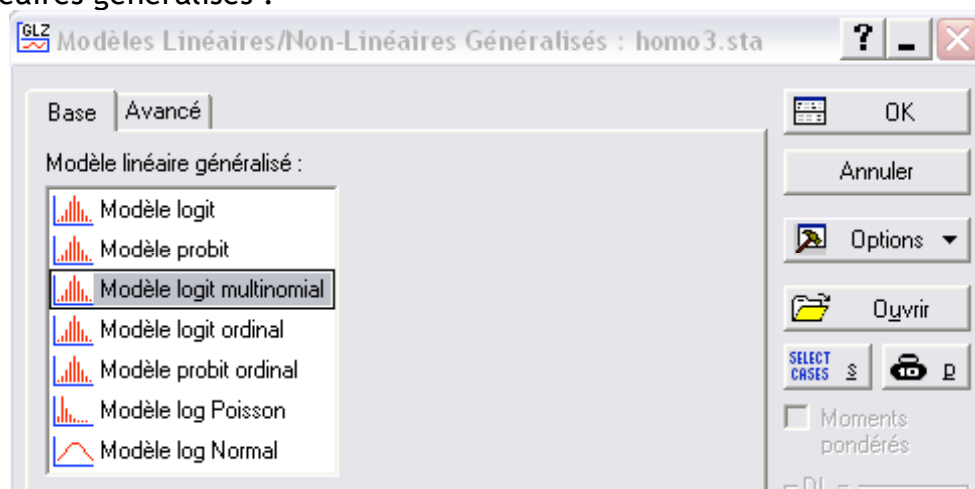
1. Régression logistique polytomiale

Avant d'étudier les modèles pour données ordinales, remarquons qu'il est également possible d'utiliser la régression logistique lorsque la réponse est polytomiale : elle consiste à confronter séparément chaque réponse possible avec une réponse privilégiée préalablement choisie. On obtient exactement les mêmes résultats qu'avec le modèle log-linéaire.

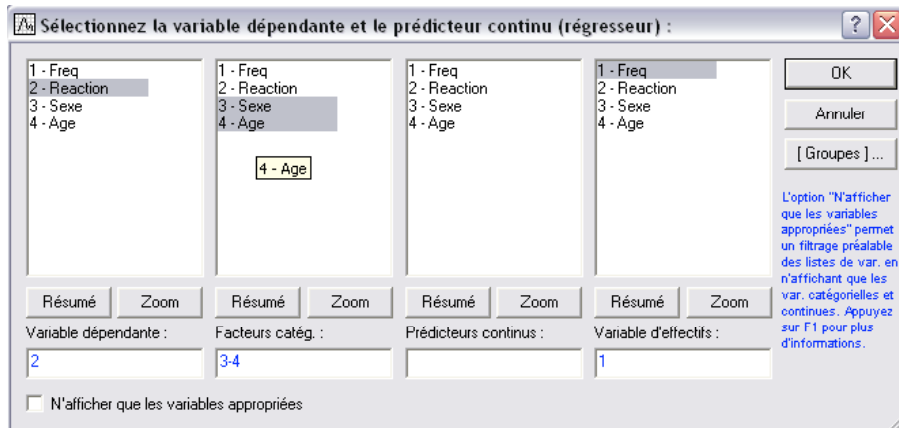
- Le fichier de départ est homo3.sta :

Données : homo3.sta (4 var. et 50 obs.)				
	1 Freq	2 Reaction	3 Sexe	4 Age
1	129	1	Femme	<30
2	85	2	Femme	<30
3	11	3	Femme	<30
4	7	4	Femme	<30
5	5	5	Femme	<30
6	70	1	Femme	30-39
7	55	2	Femme	30-39
8	7	3	Femme	30-39
9	6	4	Femme	30-39
10	4	5	Femme	30-39
11	82	1	Femme	40-49
12	59	2	Femme	40-49
13	9	3	Femme	40-49
14	12	4	Femme	40-49
15	8	5	Femme	40-49
16	94	1	Femme	50-64
17	72	2	Femme	50-64
18	18	3	Femme	50-64
19	7	4	Femme	50-64

- Choisir « Modèle logit multinomial » dans Statistiques > Modèles Linéaires/Non linéaires généralisés :



- Renseigner les variables en présence et en particulier la colonne des fréquences :



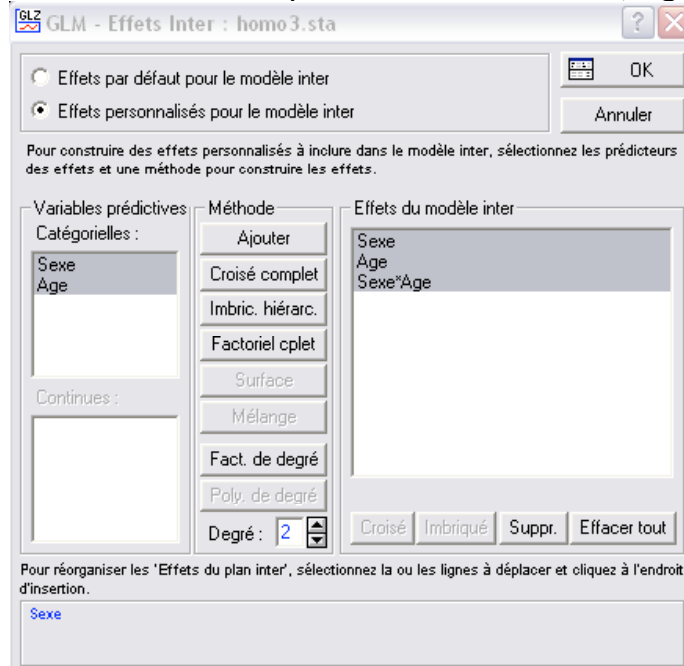
- C'est la dernière catégorie renseignée pour la réponse qui est la référence : forçons donc le choix pour la première :



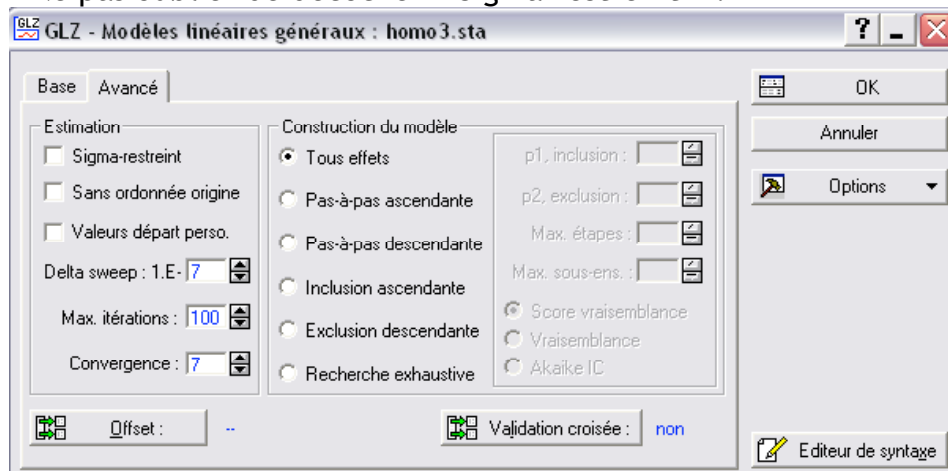
- De même, forçons nos choix pour les références pour Age et Sexe :



- Le modèle saturé prévoit des effets Sexe, Age et Sexe*Age :



- Ne pas oublier de décocher « Sigma-restreint » :



- On retrouve bien les mêmes résultats pour les tests qu'en log-linéaire :

Reaction - Test Vraisemblance Type 1 (hor
 Distribution : MULTINOMIALE
 Fonction de Liaison : LOGIT

Effet	Degré de Liberté	Log-Vraisbnc	Chi²	p
Ord.Orig.	4	-2246,44		
Sexe	4	-2238,73	15,41537	0,003913
Age	16	-2211,28	54,90658	0,000004
Sexe*Age	16	-2200,57	21,42196	0,162855

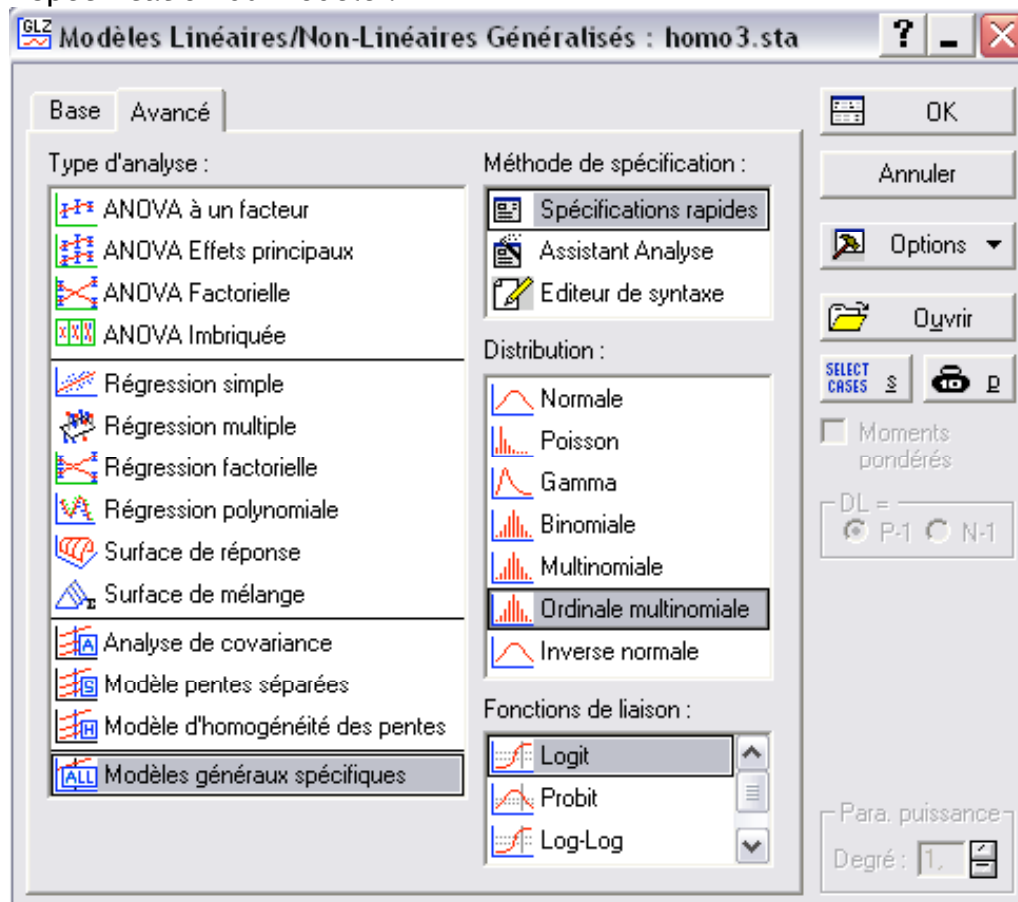
Reaction - Test Vraisemblance Type 1 (hon Distribution : MULTINOMIALE Fonction de Liaison : LOGIT				
Effet	Degré de Liberté	Log-Vraisbnc	Chi²	p
Ord.Orig.	4	-2246,44		
Age	16	-2219,89	53,09457	0,000007
Sexe	4	-2211,28	17,22738	0,001746
Sexe*Age	16	-2200,57	21,42196	0,162855

- Après de l'interaction non-significative Age*Sexe, on obtient les mêmes estimations qu'en log-linéaire avec une série de paramètres pour chacune des 4 confrontations possibles de la réponse avec sa catégorie de référence :

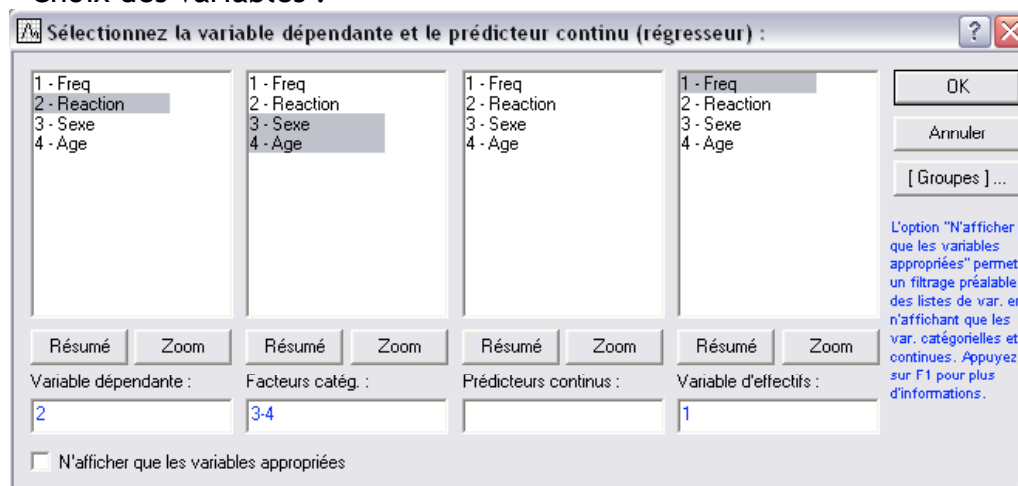
Reaction - Paramètres estimés (homo3.sta) Distribution : MULTINOMIALE Fonction de Liaison : LOGIT							
Effet	Niveau Effet	Niveau Réponse	Colonne	Estimat.	Standard Erreur	Wald Stat.	p
Ord.Orig 1		2	1	-0,35777	0,117067	9,33966	0,002242
Age	30-39	2	2	0,15488	0,161817	0,91615	0,338489
Age	40-49	2	3	0,15352	0,162392	0,89372	0,344472
Age	50-64	2	4	0,35670	0,153967	5,36707	0,020520
Age	>65	2	5	0,49513	0,183339	7,29331	0,006921
Age	<30	2	6	0,00000			
Sexe	Femme	2	7	-0,13883	0,108403	1,64009	0,200313
Sexe	Homme	2	8	0,00000			
Ord.Orig 2		3	9	-1,36041	0,176875	59,15675	0,000000
Age	30-39	3	10	-0,07225	0,266104	0,07372	0,785993
Age	40-49	3	11	-0,01925	0,266136	0,00523	0,942347
Age	50-64	3	12	0,22983	0,245576	0,87584	0,349343
Age	>65	3	13	0,82331	0,260063	10,02233	0,001547
Age	<30	3	14	0,00000			
Sexe	Femme	3	15	-0,68699	0,174281	15,53809	0,000081
Sexe	Homme	3	16	0,00000			
Ord.Orig 3		4	17	-2,09344	0,232019	81,40865	0,000000
Age	30-39	4	18	-0,28022	0,366714	0,58389	0,444793
Age	40-49	4	19	0,48305	0,301877	2,56054	0,109561
Age	50-64	4	20	-0,09999	0,342529	0,08522	0,770348
Age	>65	4	21	0,96786	0,315964	9,38314	0,002190
Age	<30	4	22	0,00000			
Sexe	Femme	4	23	-0,32749	0,214365	2,33389	0,126585
Sexe	Homme	4	24	0,00000			
Ord.Orig 4		5	25	-2,62692	0,285591	84,60694	0,000000
Age	30-39	5	26	0,23281	0,401382	0,33641	0,561908
Age	40-49	5	27	0,54638	0,374509	2,12843	0,144589
Age	50-64	5	28	0,69642	0,356797	3,80982	0,050953
Age	>65	5	29	1,77387	0,339506	27,29903	0,000000
Age	<30	5	30	0,00000			
Sexe	Femme	5	31	-0,27756	0,221464	1,57070	0,210105
Sexe	Homme	5	32	0,00000			
Echelle				1,00000	0,000000		

2. Régression ordinale polytomiale : modèle à cotes proportionnelles

- Spécification du modèle :



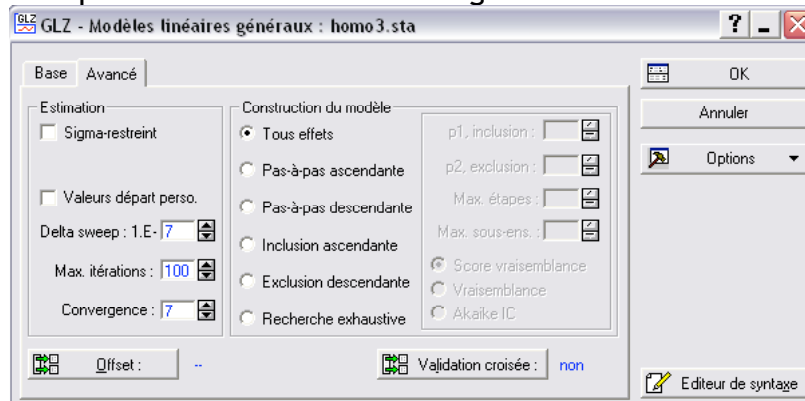
- Choix des variables :



- Catégories de référence : l'ordre de 1 à 5 (de tolérant à intolérant) pour la réponse convient ; on choisit '<30 ans' pour Age et 'Homme' pour Sexe :



- Ne pas oublier de décocher 'Sigma-restreint' :



- Les tests associés au modèle saturé (Sexe+Age+Sexe*Age) sont :

Effet	Degré de Liberté	Log-Vraisbnc	Chi ²	p
Ord.Orig.	4	-2246,44		
Sexe	1	-2241,97	8,95080	0,002773
Age	4	-2221,98	39,97516	0,000000
Sexe*Age	4	-2219,30	5,34956	0,253273

Effet	Degré de Liberté	Log-Vraisbnc	Chi ²	p
Ord.Orig.	4	-2246,44		
Age	4	-2227,05	38,77267	0,000000
Sexe	1	-2221,98	10,15330	0,001440
Sexe*Age	4	-2219,30	5,34956	0,253273

L'interaction Sexe*Age est non-significative.

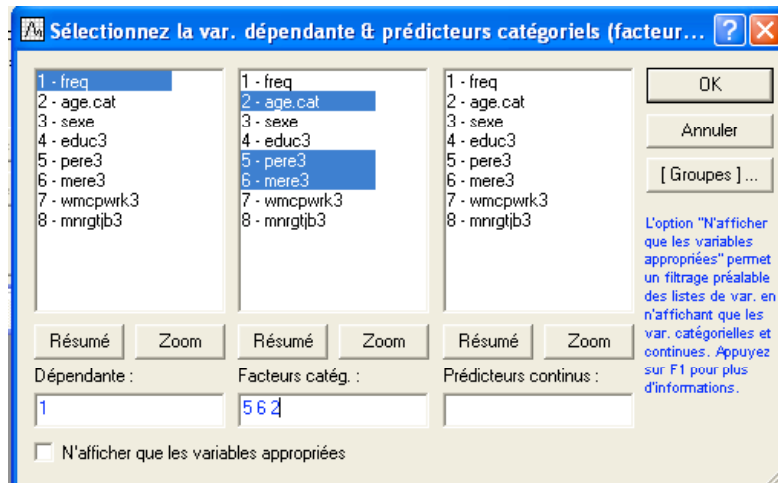
- Voici les estimations obtenues pour le modèle retenu Sexe+Age :

Effet	Niveau Effet	Colonne	Estimat.	Standard Erreur	Wald Stat.	p
Ord.Orig 1		1	-0,163902	0,098214	2,7850	0,095153
Ord.Orig 2		2	1,445892	0,104884	190,0427	0,000000
Ord.Orig 3		3	2,205581	0,115811	362,6987	0,000000
Ord.Orig 4		4	3,000426	0,137474	476,3489	0,000000
Age	30-39	5	-0,028886	0,137141	0,0444	0,833176
Age	40-49	6	-0,208703	0,134445	2,4097	0,120583
Age	50-64	7	-0,248055	0,128850	3,7062	0,054210
Age	>65	8	-0,854650	0,142257	36,0936	0,000000
Age	<30	9	0,000000			
Sexe	Femme	10	0,283774	0,089162	10,1294	0,001459
Sexe	Homme	11	0,000000			

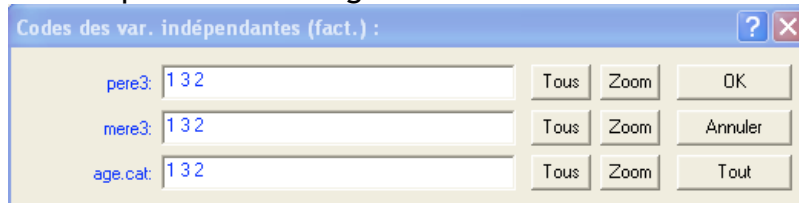
3. Variable réponse multivariée (fichier femme_et_travail.sta)

Dépendance entre Pere3 et Mere3 et leur lien avec Age.cat

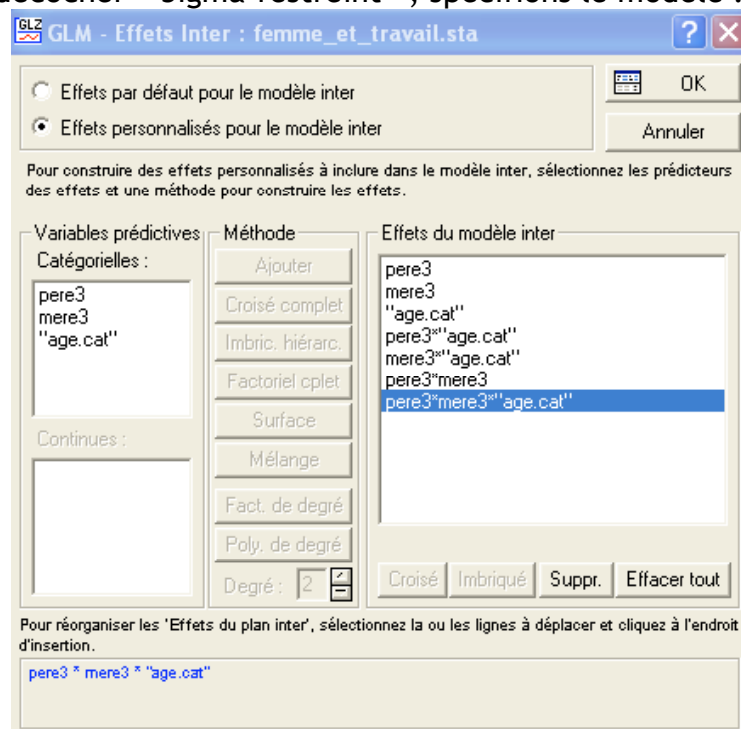
- Spécification des variables :



- Prenons systématiquement la catégorie intermédiaire en référence :



- Après avoir décocher « Sigma-restreint », spécifions le modèle :



- Les carrés de type I sont:

freq - Test Vraisemblance Type 1 (femme_)				
Distribution : POISSON				
Fonction de Liaison : LOG				
Effet	Degré de Liberté	Log-Vraisbnc	Chi²	p
Ord.Orig.	1	-2386,61		
"pere3"	2	-2279,37	214,4780	0,000000
"mere3"	2	-2040,69	477,3691	0,000000
age.cat	2	-1972,19	136,9872	0,000000
"pere3"*age.cat	4	-1930,99	82,4139	0,000000
"mere3"*age.cat	4	-1858,49	144,9896	0,000000
"pere3"*"mere3"	4	-1639,08	438,8150	0,000000
"pere3"*"mere3"*age.cat	8	-1637,64	2,8786	0,941746

L'association Père-Mère ne change pas significativement avec l'époque considérée.

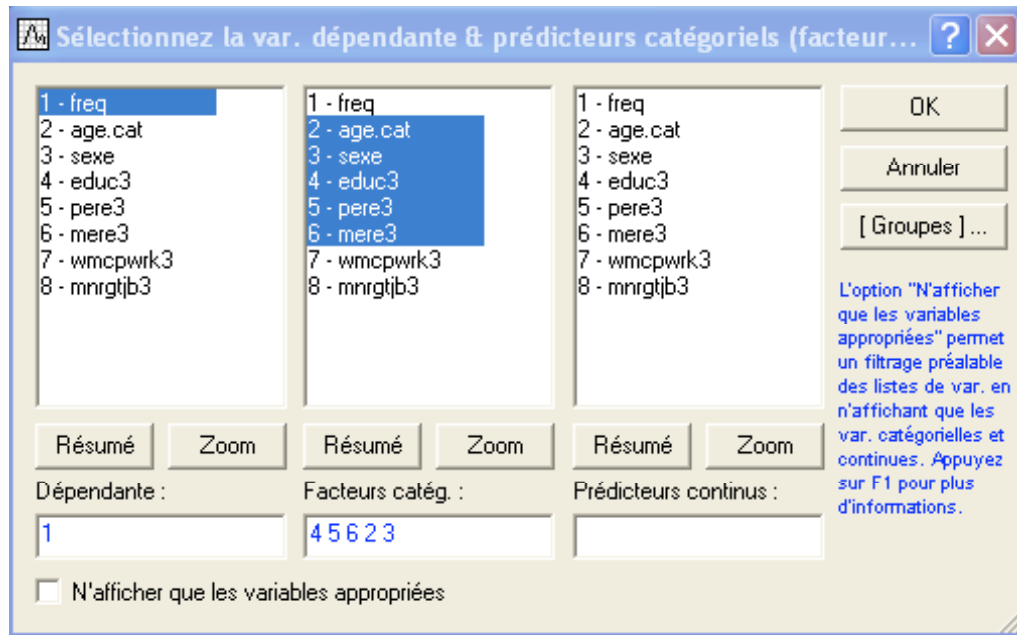
- Si on coche 'Sigma-restreint' avec le modèle sans cette interaction, on peut obtenir les tests de type III qui évaluent séparément le retrait unique de chaque terme du modèle (les autres restant présents) :

freq - Test Vraisemblance Type 3 (femme_)				
Distribution : POISSON				
Fonction de Liaison : LOG				
Effet	Degré de Liberté	Log-Vraisbnc	Chi²	p
"pere3"	2	-1645,12	12,0649	0,002400
"mere3"	2	-1732,71	187,2566	0,000000
age.cat	2	-1727,89	177,6106	0,000000
"pere3"*age.cat	4	-1645,92	13,6811	0,008386
"mere3"*age.cat	4	-1677,21	76,2567	0,000000
"pere3"*"mere3"	4	-1858,49	438,8150	0,000000

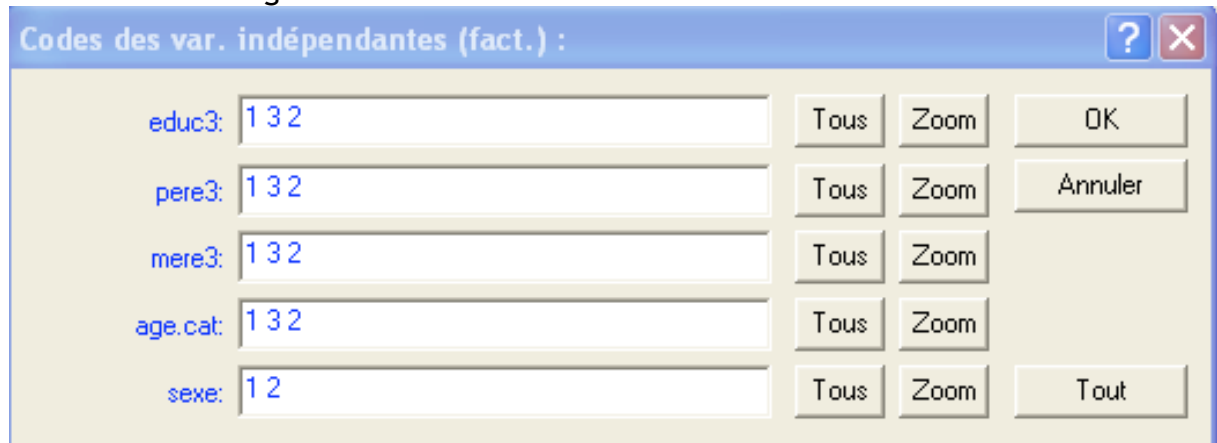
- Tous les termes sont donc indispensables. Il suffit ensuite de réajuster ce modèle après avoir décocher 'Sigma-restreint' pour interpréter les estimations.

Dépendance entre EDUC3 et Pere3, Mere3, Age.cat et Sexe

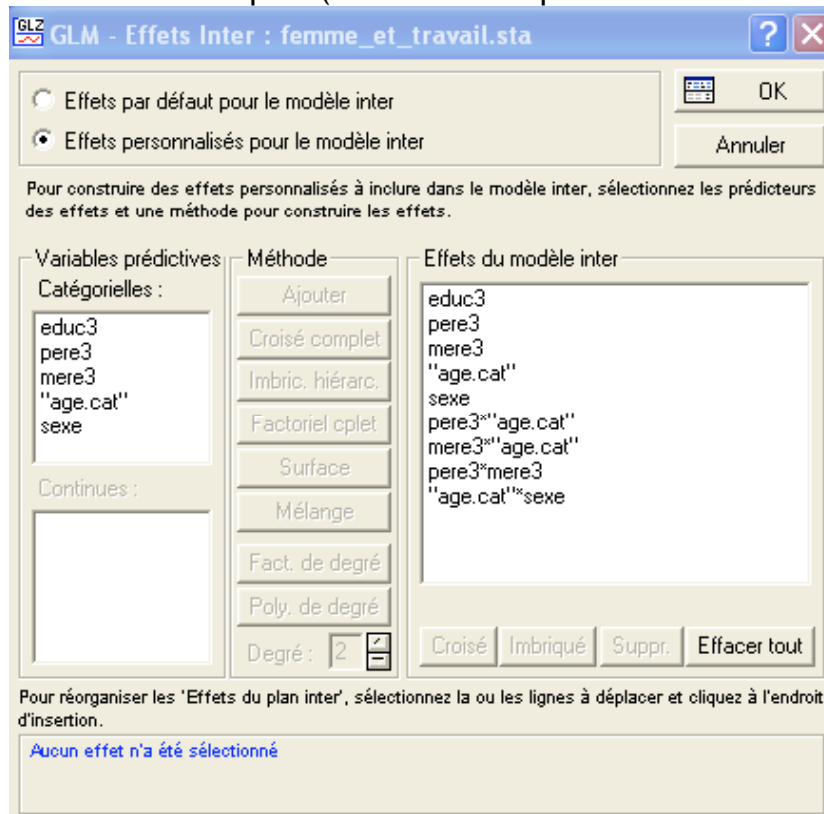
- Choix des variables :



- choix des catégories de référence :



- Modèle de départ (en tenant compte des conclusions précédentes)

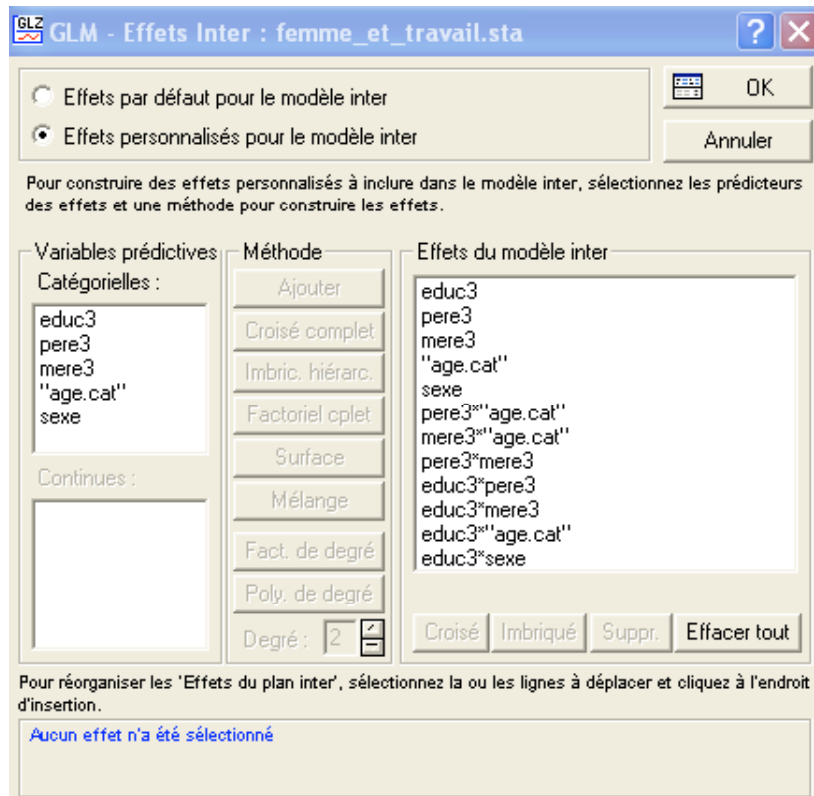


freq - Test Vraisemblance Type 1 (femme_e
Distribution : POISSON
Fonction de Liaison : LOG

Effet	Degré de Liberté	Log-Vraisbnc	Chi²	p
Ord.Orig.	1	-2386,61		
"educ3"	2	-2263,41	246,4058	0,000000
"pere3"	2	-2156,17	214,4780	0,000000
"mere3"	2	-1917,48	477,3691	0,000000
age.cat	2	-1848,99	136,9872	0,000000
sexe	1	-1848,91	0,1575	0,691463
"pere3"*age.cat	4	-1807,70	82,4139	0,000000
"mere3"*age.cat	4	-1735,21	144,9896	0,000000
"pere3"*"mere3"	4	-1515,80	438,8150	0,000000
age.cat*sexe	2	-1514,76	2,0739	0,354529

On peut laisser, comme attendu, age*Sexe

- Introduisons des liens du 1^{er} ordre entre EDUC3 et les autres variables :



Seul le lien entre EDUC3 et Sexe (conditionnellement à...) est non significatif :

freq - Test Vraisemblance Type 3 (femme_ε
Distribution : POISSON
Fonction de Liaison : LOG

Effet	Degré de Liberté	Log-Vraisbnc	Chi²	p
"educ3"	2	-1356,64	100,4632	0,000000
"pere3"	2	-1313,22	13,6334	0,001095
"mere3"	2	-1370,22	127,6237	0,000000
age.cat	2	-1336,41	60,0057	0,000000
sexe	1	-1307,05	1,2862	0,256757
"pere3*"age.cat	4	-1310,73	8,6457	0,070592
"mere3*"age.cat	4	-1335,30	57,7898	0,000000
"pere3**"mere3"	4	-1453,89	294,9754	0,000000
"educ3**"pere3"	4	-1353,45	94,0878	0,000000
"educ3**"mere3"	4	-1327,04	41,2765	0,000000
"educ3**"age.cat	4	-1336,21	59,6082	0,000000
"educ3**"sexe	2	-1307,79	2,7813	0,248917

- La stabilité de ces derniers peut être évaluée en testant la nécessité d'interactions d'un ordre supérieur:

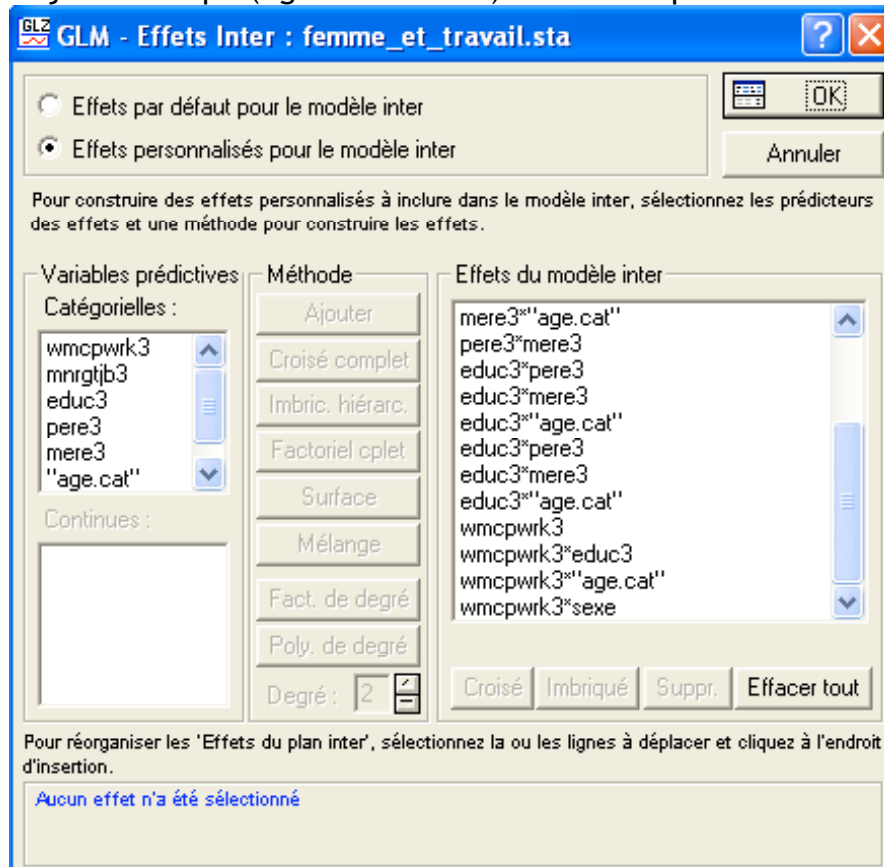
freq - Test Vraisemblance Type 1 (femme_...
Distribution : POISSON
Fonction de Liaison : LOG

Effet	Degré de Liberté	Log-Vraisbnc	Chi²	p
Ord.Orig.	1	-2386,61		
"educ3"	2	-2263,41	246,4058	0,000000
"pere3"	2	-2156,17	214,4780	0,000000
"mere3"	2	-1917,48	477,3691	0,000000
age.cat	2	-1848,99	136,9872	0,000000
sexe	1	-1848,91	0,1575	0,691463
"pere3"*age.cat	4	-1807,70	82,4139	0,000000
"mere3"*age.cat	4	-1735,21	144,9896	0,000000
"pere3"*"mere3"	4	-1515,80	438,8150	0,000000
"educ3"*"pere3"	4	-1367,47	296,6635	0,000000
"educ3"*"mere3"	4	-1337,60	59,7434	0,000000
"educ3"*age.cat	4	-1307,79	59,6082	0,000000
"educ3"*"pere3"*"mere3"	8	-1304,38	6,8190	0,556276
"educ3"*"pere3"*age.cat	8	-1301,78	5,2052	0,735439
"educ3"*"mere3"*age.cat	8	-1296,62	10,3151	0,243603

Cela n'est pas nécessaire !

Dépendance entre Rep1 avec Age, Sexe et EDUC

- Ajoutons Rép1*(Age+Sexe+EDUC) au modèle précédent :



- Les tests de type III donnent comme résultat :

freq - Test Vraisemblance Type 3 (femme_				
Distribution : POISSON				
Fonction de Liaison : LOG				
Effet	Degré de Liberté	Log-Vraisbnc	Chi ²	p
"mere3**age.cat	4	-1221,07	57,7898	0,000000
"pere3**mere3"	4	-1339,66	294,9754	0,000000
"educ3**pere3"	4	-1239,22	94,0878	0,000000
"educ3**mere3"	4	-1212,81	41,2765	0,000000
"educ3**age.cat	4	-1218,10	51,8534	0,000000
"wmcprwk3"	2	-1246,90	109,4434	0,000000
"wmcprwk3**age.cat	4	-1203,96	23,5734	0,000097
"wmcprwk3**sexe	2	-1195,86	7,3637	0,025176
"wmcprwk3**educ3"	4	-1203,84	23,3244	0,000109

- Les termes Rép1*père3 et Rép1*mère3 n'apportent rien :

freq - Test Vraisemblance Type 1 (femme_e
 Distribution : POISSON
 Fonction de Liaison : LOG

Effet	Degré de Liberté	Log-Vraisbnc	Chi²	p
"pere3"*age.cat	4	-1807,70	82,4139	0,000000
"mere3"*age.cat	4	-1735,21	144,9896	0,000000
"pere3"*"mere3"	4	-1515,80	438,8150	0,000000
"educ3"*"pere3"	4	-1367,47	296,6635	0,000000
"educ3"*"mere3"	4	-1337,60	59,7434	0,000000
"educ3"*age.cat	4	-1307,79	59,6082	0,000000
"wmcprk3"	2	-1229,14	157,3134	0,000000
"wmcprk3"*age.cat	4	-1207,52	43,2393	0,000000
"wmcprk3"*sexe	2	-1203,84	7,3637	0,025176
"wmcprk3"*"educ3"	4	-1192,17	23,3244	0,000109
"wmcprk3"*"pere3"	4	-1190,89	2,5678	0,632543
"wmcprk3"*"mere3"	4	-1186,66	8,4571	0,076198

- La même démarche peut être adoptée avec Rép2.