

Analyse statistique de données qualitatives et quantitatives en sciences sociales : TP

Modèle log-linéaire (chapitre 2)

Interprétation des sorties

PLAN

- Table 2x2 (modèle saturé 4 paramètres)
 - Evaldemo2 (Y) et blgetmg (X)
- Table 2x3 (modèle saturé 6 paramètres)
 - Evaldemo2 (Y) et reg3 (X)
- Table 2x2x3 (modèle saturé 12 paramètres)
 - Evaldemo2 (X) , blgetmg (Y), reg3(Z)
 - Evaldemo2 (X) , polintr2 (Y), reg3(Z)
- Table 4x2x3 (modèle saturé 24 paramètres)
 - Evaldemo4 (X) , polintr2 (Y), reg3(Z)

TABLE 2x2

L'évaluation de la démocratie est-elle liée au fait d'appartenir à une minorité ethnique?

Point de départ : table de contingence
 \Leftrightarrow fréquences observées

Fréquences attendues sous l'hypothèse
 d'indépendance

Minorité Ethnique	Evaluation de la démocratie		
	Positive	Négative	Total
Non	$n_{11} = 1034$	$n_{12} = 715$	$n_{1+} = 1749$
Oui	$n_{21} = 74$	$n_{22} = 30$	$n_{2+} = 104$
Total	$n_{+1} = 1108$	$n_{+2} = 745$	$n_{++} = 1853$

Minorité Ethnique	Evaluation de la démocratie		
	Positive	Négative	Total
Non	1045,8	703,2	1749
Oui	62,2	41,8	104
Total	1108	745	1853

$$\frac{n_{i+} \times n_{+j}}{n_{++}}$$

Il semblerait que le nombre de personnes qui évaluent négativement la démocratie dans le groupe « minorité ethnique » est surestimé sous l'hypothèse d'indépendance.

Le modèle va permettre de déterminer si cette hypothèse est trop réductrice (\Leftrightarrow différences observées trop grandes que pour négliger le lien entre X et Y) via un test de rapport de vraisemblance sur le paramètre λ_{ij}^{XY} .

TABLE 2x2

L'hypothèse d'indépendance est-elle plausible?

- Tests de type 1 et de type 3 (modèle saturé)

LR Statistics For Type 1 Analysis				
Source	Deviance	DF	Khi-2	Pr > Khi-2
Intercept	1845.4015			
EVALDEMO2	1773.8284	1	71.57	<.0001
blgetmg	6.1510	1	1767.68	<.0001
EVALDEMO2*blgetmg	0.0000	1	6.15	0.0131

LR Statistics For Type 3 Analysis			
Source	DF	Khi-2	Pr > Khi-2
EVALDEMO2	1	58.51	<.0001
blgetmg	1	992.55	<.0001
EVALDEMO2*blgetmg	1	6.15	0.0131

La statistique $G^2 = 6,15 > X_1^2(0,95) = 3,84 \Leftrightarrow$ la p-valeur associée au terme d'interaction $< 0,05$

➤ Conclusion : il existe un lien significatif entre 'blgetmg' et 'evaldemo2' : **on sélectionne le modèle saturé** :

$$\log \mu_{ij} = \gamma + \lambda_i^X + \lambda_j^Y + \lambda_{ij}^{XY} \quad (\lambda_1^X = \lambda_1^Y = \lambda_{1j}^{XY} = \lambda_{i1}^{XY} = 0)$$

NB : dans ce cas-ci, on utilise indifféremment les tests de type 1 et de type 3 car on s'intéresse à un seul terme (le dernier) pour sélectionner le modèle (indépendance Vs saturé).

Les 2 types livrent la même information **uniquement** pour le dernier paramètre inclus dans l'équation du modèle.

TABLE 2x2

Modèle saturé : interprétation des coefficients estimés

Analysis Of Maximum Likelihood Parameter Estimates									
Parameter			DDL	Valeur estimée	Erreur type	Wald 95% Confidence Limits		Wald Chi-Square	Pr > Khi-2
Intercept			1	6.9412	0.0311	6.8802	7.0021	49818.2	<.0001
EVALDEMO2	Low		1	-0.3689	0.0486	-0.4642	-0.2736	57.53	<.0001
blgetmg	Yes		1	-2.6371	0.1203	-2.8730	-2.4013	480.26	<.0001
EVALDEMO2*blgetmg	Low	Yes	1	-0.5340	0.2218	-0.9688	-0.0992	5.79	0.0161
Scale			0	1.0000	0.0000	1.0000	1.0000		

■ EVALDEMO2 : $\hat{\lambda}_2^Y = -0.3689 \Leftrightarrow$
 Estimation du logarithme de la cote de succès (évaluation négative) **chez les personnes qui n'appartiennent pas à une minorité ethnique** : \rightarrow cote de « succès » = $e^{-0.3689} = 0,6915$: entre 0 et 1, pas facile à interpréter \rightarrow on inverse : $1/0,6915 = 1,446 \Leftrightarrow$ la probabilité d'évaluer positivement la démocratie est supérieure de 45% environ par rapport à celle de l'évaluer négativement dans ce groupe (IC 95% = **[1,31 ; 1,59]**).

■ Estimation du logarithme de la cote de succès (évaluation négative) **chez les personnes appartiennent à une minorité ethnique** : $\hat{\lambda}_2^Y + \hat{\lambda}_{22}^{XY} = -0.3689 - 0.5340 = -0.9029$
 \rightarrow cote de « succès » = $e^{-0.9029} = 0,405$: entre 0 et 1, pas facile à interpréter \rightarrow on inverse : $1/0,405 = 2,47 \Leftrightarrow$ la probabilité d'évaluer positivement la démocratie est environ 2,5 fois supérieure à celle de l'évaluer négativement parmi les répondants de ce groupe.

TABLE 2x2

Modèle saturé : interprétation du terme d'interaction

Analysis Of Maximum Likelihood Parameter Estimates									
Parameter			DDL	Valeur estimée	Erreur type	Wald 95% Confidence Limits		Wald Chi-Square	Pr > Khi-2
Intercept			1	6.9412	0.0311	6.8802	7.0021	49818.2	<.0001
EVALDEMO2	Low		1	-0.3689	0.0486	-0.4642	-0.2736	57.53	<.0001
blgetmg	Yes		1	-2.6371	0.1203	-2.8730	-2.4013	480.26	<.0001
EVALDEMO2*blgetmg	Low	Yes	1	-0.5340	0.2218	-0.9688	-0.0992	5.79	0.0161
Scale			0	1.0000	0.0000	1.0000	1.0000		

- IC 95% pour le rapport des cotes d'évaluation **positive** : (1.104, 2.635)
 - 1 ne fait pas partie des valeurs plausibles (confirme la non plausibilité de l'hypothèse d'indépendance)
 - Toutes les valeurs sont supérieures à 1 : les personnes minoritaires d'un point de vue ethnique évaluent plus souvent la démocratie en Belgique de manière positive par rapport aux « non minoritaires ».
 - Largeur de l'intervalle \Leftrightarrow grande incertitude (rappel : faible effectif pour blgetmg = « Yes »)

▪ Le logarithme du rapport de cotes d'évaluation négative pour les « minority » Vs « no minority » est estimé par : $\hat{\lambda}_{22}^{XY} = -0,534$

▪ Sur base de l'approximation normale pour la distribution *a posteriori* de ce paramètre ($\lambda_{22}^{XY} | données$) $\approx N(-0.534, 0.2218^2)$ on peut construire un intervalle de crédibilité 95% approximatif pour ce log(RC) : (-0.9688, -0.0992)

▪ IC 95% pour le rapport des cotes d'évaluation **négative** : (0.3795, 0.90556)

NB : à partir des **estimations du modèle d'indépendance**, on retrouve les **fréquences attendues** calculées à partir de la table de contingence.

Analysis Of Maximum Likelihood Parameter Estimates								
Parameter		DDL	Valeur estimée	Erreur type	Wald 95% Confidence Limits		Wald Chi-Square	Pr > Khi-2
Intercept		1	6.9526	0.0306	6.8926	7.0125	51719.5	<.0001
EVALDEMO2	Low	1	-0.3969	0.0474	-0.4898	-0.3041	70.18	<.0001
blgetmg	Yes	1	-2.8224	0.1009	-3.0202	-2.6246	781.97	<.0001
Scale		0	1.0000	0.0000	1.0000	1.0000		

Fréquences attendues sous l'hypothèse d'indépendance

- $\hat{\mu}_{11} = e^{6,9526} = 1045,8$
- $\hat{\mu}_{12} = e^{6,9526-0,3969} = 703,2$
- $\hat{\mu}_{21} = e^{6,9526-2,8224} = 62,2$
- $\hat{\mu}_{22} = e^{6,9526-2,8224-0,3969} = 41,8$

Minorité Ethnique	Evaluation de la démocratie		
	Positive	Négative	Total
Non	1045,8	703,2	1749
Oui	62,2	41,8	104
Total	1108	745	1853

NB : à partir des **estimations du modèle saturé**, on retrouve les **fréquences observées** de la table de contingence.

Analysis Of Maximum Likelihood Parameter Estimates									
Parameter			DDL	Valeur estimée	Erreur type	Wald 95% Confidence Limits		Wald Chi-Square	Pr > Khi-2
Intercept			1	6.9412	0.0311	6.8802	7.0021	49818.2	<.0001
EVALDEMO2	Low		1	-0.3689	0.0486	-0.4642	-0.2736	57.53	<.0001
blgetmg	Yes		1	-2.6371	0.1203	-2.8730	-2.4013	480.26	<.0001
EVALDEMO2*blgetmg	Low	Yes	1	-0.5340	0.2218	-0.9688	-0.0992	5.79	0.0161
Scale			0	1.0000	0.0000	1.0000	1.0000		

Point de départ : table de contingence
 ⇔ fréquences observées

- $\hat{\mu}_{11} = e^{6,9412} = 1034$
- $\hat{\mu}_{12} = e^{6,9412-0,3689} = 715$
- $\hat{\mu}_{21} = e^{6,9412-2,6371} = 74$
- $\hat{\mu}_{22} = e^{6,9412-0,3689-2,6371-0,5340} = 30$

Minorité Ethnique	Evaluation de la démocratie		
	Positive	Négative	Total
Non	$n_{11} = 1034$	$n_{12} = 715$	$n_{1+} = 1749$
Oui	$n_{21} = 74$	$n_{22} = 30$	$n_{2+} = 104$
Total	$n_{+1} = 1108$	$n_{+2} = 745$	$n_{++} = 1853$

TABLE 2x3

L'évaluation de la démocratie en Belgique est-elle liée à la région du pays (3 niveaux)?

L'hypothèse d'indépendance est-elle plausible?

LR Statistics For Type 1 Analysis				
Source	Deviance	DF	Khi-2	Pr > Khi-2
Intercept	883.0848			
EVALDEMO2	811.8234	1	71.26	<.0001
reg3	40.1542	2	771.67	<.0001
EVALDEMO2*reg3	0.0000	2	40.15	<.0001

La statistique $G^2 = 40,15 > X_2^2(0,95) = 5,99 \Leftrightarrow$ la p-valeur associée au terme d'interaction $< 0,05$

- Conclusion : il existe un lien significatif entre l'évaluation de la démocratie (2 niveaux) et la région (3 niveaux): **on sélectionne le modèle saturé**

LR Statistics For Type 3 Analysis			
Source	DF	Khi-2	Pr > Khi-2
EVALDEMO2	1	12.48	0.0004
reg3	2	331.86	<.0001
EVALDEMO2*reg3	2	40.15	<.0001

$$\log \mu_{ij} = \gamma + \lambda_i^X + \lambda_j^Y + \lambda_{ij}^{XY} \quad (\lambda_1^X = \lambda_1^Y = \lambda_{1j}^{XY} = \lambda_{i1}^{XY} = 0)$$

TABLE 2x3

Modèle saturé : interprétation des coefficients estimés

Analysis Of Maximum Likelihood Parameter Estimates									
Parameter			DDL	Valeur estimée	Erreur type	Wald 95% Confidence Limits		Wald Chi-Square	Pr > Khi-2
Intercept			1	5.9189	0.0518	5.8173	6.0205	13032.4	<.0001
EVALDEMO2	Low		1	-0.5072	0.0846	-0.6730	-0.3415	35.97	<.0001
reg3	Bxl		1	-1.0361	0.1013	-1.2347	-0.8375	104.59	<.0001
reg3	Fl		1	0.4913	0.0658	0.3623	0.6203	55.70	<.0001
EVALDEMO2*reg3	Low	Bxl	1	-0.8492	0.2101	-1.2610	-0.4374	16.34	<.0001
EVALDEMO2*reg3	Low	Fl	1	0.2935	0.1041	0.0895	0.4975	7.95	0.0048
Scale			0	1.0000	0.0000	1.0000	1.0000		

- Dans les trois régions du pays, les répondants évaluent plus souvent la démocratie positivement (plutôt que nég.)
- La cote de la région bruxelloise se distingue plus fortement
- Reste à savoir si ces différences sont significatives (et si oui, dans quelle mesure...)

- Log de la cote d'évaluation nég. Vs pos. chez les **Wallons** :
 $\hat{\lambda}_2^Y = -0,5072$
 → Cote d'éval. pos. Vs nég. : $\frac{1}{e^{-0,5072}} = \mathbf{1,66}$

- Log de la cote d'évaluation nég. Vs pos. chez les **Bruxellois** :
 $\hat{\lambda}_2^Y + \hat{\lambda}_{22}^{XY} = -0,5072 - 0,8492 = -1,3564$
 → Cote d'éval. pos. Vs nég. : $\frac{1}{e^{-1,3564}} = \mathbf{3,88}$

- Log de la cote d'évaluation nég. Vs pos. chez les **Flamands** :
 $\hat{\lambda}_2^Y + \hat{\lambda}_{32}^{XY} = -0,5072 + 0,2935 = -0,2137$
 → Cote d'éval. pos. Vs nég. : $\frac{1}{e^{-0,2137}} = \mathbf{1,24}$

TABLE 2x3

Modèle saturé : interprétation du terme d'interaction

Analysis Of Maximum Likelihood Parameter Estimates									
Parameter			DDL	Valeur estimée	Erreur type	Wald 95% Confidence Limits		Wald Chi-Square	Pr > Khi-2
EVALDEMO2*reg3	Low	Bxl	1	-0.8492	0.2101	-1.2610	-0.4374	16.34	<.0001
EVALDEMO2*reg3	Low	Fl	1	0.2935	0.1041	0.0895	0.4975	7.95	0.0048

Analysis Of Maximum Likelihood Parameter Estimates									
Parameter			DDL	Valeur estimée	Erreur type	Wald 95% Confidence Limits		Wald Chi-Square	Pr > Khi-2
EVALDEMO2*reg3	Low	Bxl	1	-1.1427	0.2017	-1.5380	-0.7475	32.11	<.0001
EVALDEMO2*reg3	Low	Wal	1	-0.2935	0.1041	-0.4975	-0.0895	7.95	0.0048

	Log RC EVAL négative	RC EVAL négative	RC EVAL positive	IC 95% pour RC EVAL positive	
Bxl Vs Fl	-1,142	$e^{-1,142} = 0,3192$	$\frac{1}{0,3192} = 3,13$	$\frac{1}{e^{-0,7475}} = 2,11$	$\frac{1}{e^{-1,5380}} = 4,66$
Bxl Vs Wal	-0,8492	$e^{-0,8492} = 0,4278$	$\frac{1}{0,4278} = 2,34$	$\frac{1}{e^{-0,4374}} = 1,55$	$\frac{1}{e^{-1,2610}} = 3,53$
Wal Vs Fl	-0,2935	$e^{-0,2935} = 0,7456$	$\frac{1}{0,7456} = 1,34$	$\frac{1}{e^{-0,0895}} = 1,09$	$\frac{1}{e^{-0,4975}} = 1,64$

- Aucun des IC pour les RC ne contient 1 \Leftrightarrow Tous les contrastes sont significatifs (p-valeurs < 0,05)
- Clairement, Bruxelles se distingue assez fortement : la probabilité d'évaluer la démocratie positivement y est significativement plus élevée qu'en Wallonie et en Flandre.
- Le contraste Wallonie/Flandre est statistiquement significatif mais moins remarquable quantitativement.

TABLE 2x2x3 – exemple a

Évaluation de la démocratie en fonction de la région du pays ET du fait d'appartenir ou non à une minorité ethnique

- Les deux modèles précédents nous ont permis de déterminer que blgetmg et reg3 sont toutes deux liées significativement avec evaldemo2 de manière marginale (↔ quand les effets sont testés séparément).
- Nous allons à présent inclure ces deux variables explicatives simultanément dans un même modèle afin d'examiner la structure de dépendance entre les 3 variables. Nous allons découvrir que l'effet de l'appartenance à une minorité ethnique n'est plus significatif une fois que l'on contrôle la région du pays.

NB : effectifs faibles pour blgetmg = Yes par région

Table 1 de reg3 par EVALDEMO2			
Contrôle pour blgetmg=Yes			
reg3(Region - 3 Levels)	EVALDEMO2(How democratic Belgium is overall - 2 Levels)		
	High	Low	Total
Bxl	31	5	36
Fl	27	14	41
Wal	16	11	27
Total	74	30	104
Fréquence manquante = 1			

Table 2 de reg3 par EVALDEMO2			
Contrôle pour blgetmg=No			
reg3(Region - 3 Levels)	EVALDEMO2(How democratic Belgium is overall - 2 Levels)		
	High	Low	Total
Bxl	98	27	125
Fl	580	476	1056
Wal	356	212	568
Total	1034	715	1749
Fréquence manquante = 7			

TABLE 2x2x3 – exemple a

Sélection du modèle

- On part du modèle saturé : $\log \mu_{ijk} = \gamma + \lambda_i^X + (\lambda_j^Y + \lambda_k^Z + \lambda_{jk}^{YZ}) + \lambda_{ij}^{XY} + \lambda_{ik}^{XZ} + \lambda_{ijk}^{XYZ}$
- On teste le retrait du terme d'interaction de 3^{ème} ordre λ_{ijk}^{XYZ}

LR Statistics For Type 3 Analysis			
Source	DF	Khi-2	Pr > Khi-2
IVALDEMO2	1	36.91	<.0001
blgetmg	1	383.72	<.0001
reg3	2	380.27	<.0001
blgetmg*reg3	2	48.35	<.0001
IVALDEMO2*blgetmg	1	0.13	0.7214
IVALDEMO2*reg3	2	31.50	<.0001
IVALDEM*blgetmg*reg3	2	1.62	0.4458

LR Statistics For Type 1 Analysis				
Source	Deviance	DF	Khi-2	Pr > Khi-2
Intercept	2726.7495			
IVALDEMO2	2655.1764	1	71.57	<.0001
blgetmg	887.4990	1	1767.68	<.0001
reg3	104.7317	2	782.77	<.0001
blgetmg*reg3	44.4517	2	60.28	<.0001
IVALDEMO2*blgetmg	38.3007	1	6.15	0.0131
IVALDEMO2*reg3	1.6158	2	36.68	<.0001
IVALDEM*blgetmg*reg3	0.0000	2	1.62	0.4458

À partir des tests de **Type 1** et de **Type 3**, on conclut que l'hypothèse selon laquelle $\lambda_{ijk}^{XYZ} = 0$ est plausible $G^2 = 1,62 < X_2^2(0,95) = 5,99$ ($p = 0,4458 > 0,05$)
 → 1^{ère} simplification du modèle : retrait de λ_{ijk}^{XYZ} .

À partir des tests de **Type 1**, on peut aussi dire que l'indépendance entre l'évaluation de la démocratie et la région *conditionnellement* à *blgetmg* n'est pas une hypothèse plausible! $\Delta G^2 = 36,86 > X_2^2(0,95) = 5,99$ ($p < 0,0001$) → on ne pourra pas retirer le terme λ_{ik}^{XZ} du modèle déjà amputé de λ_{ijk}^{XYZ}

Reste à tester si l'indépendance entre l'évaluation de la démocratie et l'appartenance à une minorité ethnique conditionnellement à la région du pays est une hypothèse plausible...

TABLE 2x2x3 – exemple a Sélection du modèle

Pour tester X (EVALDEMO2) $\perp\!\!\!\perp$ Y (blgetmg) $|Z$ (reg3), il faut veiller à l'ordre dans lequel on inclut les termes dans l'instruction `model` de la proc `genmod` :

- `model count = X Y Z Y*Z X*Y X*Z / . . . ;` → on pourra utiliser uniquement le **Type 3** pour se prononcer sur l'hypothèse d'indépendance conditionnelle de X et Y sachant Z
 - `model count = X Y Z Y*Z X*Z X*Y / . . . ;` → on pourra utiliser indifféremment les tests de **Type 1 et de Type 3** pour se prononcer sur cette hypothèse
- La statistique calculée permettant de tester cette hypothèse est $\Delta G^2 = 1,57 < X_1^2(0,95) = 3,84$ avec une p -valeur = 0,2096 > 0,05 (cf. slides suivant) → on peut simplifier le modèle en retirant le terme $XY|Z$.

→ Modèle retenu : Indépendance de `evaldemo2` et `blgetmg` conditionnelle à `reg3`

$$X \perp\!\!\!\perp Y|Z : \log \mu_{ijk} = \gamma + \lambda_i^X + (\lambda_j^Y + \lambda_k^Z + \lambda_{jk}^{YZ}) + \lambda_{ik}^{XZ}$$

→ On peut à présent examiner les estimations dans ce modèle final

`model count = X Y Z Y*Z X*Y X*Z/...;`

`model count = X Y Z Y*Z X*Z X*Y/...;`

LR Statistics For Type 1 Analysis				
Source	Deviance	DF	Khi-2	Pr > Khi-2
Intercept	2726.7495			
EVALDEMO2	2655.1764	1	71.57	<.0001
blgetmg	887.4990	1	1767.68	<.0001
reg3	104.7317	2	782.77	<.0001
blgetmg*reg3	44.4517	2	60.28	<.0001
EVALDEMO2*blgetmg	38.3007	1	6.15	0.0131
EVALDEMO2*reg3	1.6158	2	36.68	<.0001

LR Statistics For Type 1 Analysis				
Source	Deviance	DF	Khi-2	Pr > Khi-2
Intercept	2726.7495			
EVALDEMO2	2655.1764	1	71.57	<.0001
blgetmg	887.4990	1	1767.68	<.0001
reg3	104.7317	2	782.77	<.0001
blgetmg*reg3	44.4517	2	60.28	<.0001
EVALDEMO2*reg3	3.1899	2	41.26	<.0001
EVALDEMO2*blgetmg	1.6158	1	1.57	0.2096

LR Statistics For Type 3 Analysis			
Source	DF	Khi-2	Pr > Khi-2
EVALDEMO2	1	35.36	<.0001
blgetmg	1	417.94	<.0001
reg3	2	387.21	<.0001
blgetmg*reg3	2	55.70	<.0001
EVALDEMO2*blgetmg	1	1.57	0.2096
EVALDEMO2*reg3	2	36.68	<.0001

LR Statistics For Type 3 Analysis			
Source	DF	Khi-2	Pr > Khi-2
EVALDEMO2	1	35.36	<.0001
blgetmg	1	417.94	<.0001
reg3	2	387.21	<.0001
blgetmg*reg3	2	55.70	<.0001
EVALDEMO2*reg3	2	36.68	<.0001
EVALDEMO2*blgetmg	1	1.57	0.2096

TABLE 2x2x3 – exemple a

Estimations dans le modèle final

- Ce qui nous intéresse ici, c'est de comprendre la nature des liens entre région et évaluation.
- On a déjà examiné les RC d'évaluation de la démocratie pour les 3 contrastes Bxl Vs Fl, Bxl Vs Wal, et Wal Vs Fl à partir de la table 2x3.
- On va maintenant examiner dans quelle mesure les valeurs des RC (et des bornes des IC) sont influencées par le contrôle de la variable blgetmg

```
title2 "Modèle final : Indépendance de evaldemo2 et blgetmg conditionnelle à reg3";
title3 "Région de référence : Wallonie";
proc genmod data = ex3a;
  class evaldemo2 blgetmg reg3 (ref='Wal')/param = ref ref = first;
  model count = evaldemo2 blgetmg reg3 blgetmg*reg3 evaldemo2*reg3 /dist=poisson link = log type1 type3;
run;
title3 "Région de référence : Flandre";
proc genmod data = ex3a;
  class evaldemo2 blgetmg reg3 (ref='Fl')/param = ref ref = first;
  model count = evaldemo2 blgetmg reg3 blgetmg*reg3 evaldemo2*reg3 /dist=poisson link = log type1 type3;
run;
```


TABLE 2x2x3 – exemple a Estimations dans le modèle final

Analysis Of Maximum Likelihood Parameter Estimates										Analysis Of Maximum Likelihood Parameter Estimates									
Parameter			DDL	Valeur estimée	Erreur type	Wald 95% Confidence Limits		Wald Chi-Square	Pr > Khi-2	Parameter			DDL	Valeur estimée	Erreur type	Wald 95% Confidence Limits		Wald Chi-Square	Pr > Khi-2
EVALDEMO2*reg3	Low	Bxl	1	-1.1800	0.2066	-1.5849	-0.7750	32.61	<.0001	EVALDEMO2*reg3	Low	Bxl	1	-0.8824	0.2149	-1.3035	-0.4612	16.86	<.0001
EVALDEMO2*reg3	Low	Wal	1	-0.2976	0.1042	-0.5019	-0.0933	8.15	0.0043	EVALDEMO2*reg3	Low	Fl	1	0.2976	0.1042	0.0933	0.5019	8.15	0.0043

	Log RC EVAL négative	RC EVAL négative	RC EVAL positive	IC 95% pour RC EVAL positive	
Bxl Vs Fl	-1,18	$e^{-1,18} = 0,3073$	$\frac{1}{0,3073} = 3,25$	$\frac{1}{e^{-0,775}} = 2,17$	$\frac{1}{e^{-1,5849}} = 4,88$
Bxl Vs Wal	-0,8824	$e^{-0,8824} = 0,4138$	$\frac{1}{0,4138} = 2,42$	$\frac{1}{e^{-0,4612}} = 1,59$	$\frac{1}{e^{-1,3035}} = 3,68$
Wal Vs Fl	-0,2976	$e^{-0,2976} = 0,7426$	$\frac{1}{0,7426} = 1,35$	$\frac{1}{e^{-0,0933}} = 1,1$	$\frac{1}{e^{-0,5019}} = 1,65$

Conclusion : les estimations et les ensembles de valeurs plausibles pour les rapports de cotes sont très peu affectées par le contrôle de blgetmg → les conclusions restent les mêmes qu’au slide 11 mais on peut à présent ajouter que celles-ci restent valables quand on contrôle l’appartenance à une minorité ethnique.

TABLE 2x2x3 – exemple b

Évaluation de la démocratie en fonction de la région du pays et l'intérêt pour la politique

- Rappel : grâce au modèle de **régression logistique** (cf. slides corrigé séance 2), on avait déjà examiné l'effet de **l'intérêt pour la politique pour un âge donné** et on avait conclut que les personnes qui se déclarent pas ou peu intéressées par la politique ont une cote de succès (évaluation négative Vs positive) significativement supérieure à celles qui se déclarent plutôt ou très intéressées par la politique.
- Grâce au modèle log-linéaire, nous allons à présent pouvoir évaluer (notamment) si cet effet reste similaire une fois que l'on contrôle la région du pays plutôt que l'âge

Table 1 de reg3 par EVALDEMO2			
Contrôle pour polintr2=Interested			
reg3(Region - 3 Levels)	EVALDEMO2(How democratic Belgium is overall - 2 Levels)		
	High	Low	Total
Bxl	52	12	64
Fl	354	189	543
Wal	165	69	234
Total	571	270	841
Fréquence manquante = 2			

Table 2 de reg3 par EVALDEMO2			
Contrôle pour polintr2=Not Interested			
reg3(Region - 3 Levels)	EVALDEMO2(How democratic Belgium is overall - 2 Levels)		
	High	Low	Total
Bxl	79	22	101
Fl	254	302	556
Wal	207	155	362
Total	540	479	1019
Fréquence manquante = 6			

TABLE 2x2x3 – exemple b

Sélection du modèle

- On part du modèle saturé
- On teste le retrait du terme d'interaction de 3^{ème} ordre λ_{ijk}^{XYZ}
- À partir des tests de **Type 1 et de Type 3**, on conclut que l'hypothèse selon laquelle $\lambda_{ijk}^{XYZ} = 0$ est plausible : $G^2 = 2,67 < X_2^2(0,95) = 5,99$ ($p = 0,2637 > 0,05$) → 1^{ère} simplification du modèle : retrait de λ_{ijk}^{XYZ} .
- À partir des tests de **Type 1**, on peut aussi dire que l'indépendance entre l'évaluation de la démocratie et la région *conditionnellement* à *polintr2* n'est pas une hypothèse plausible: $\Delta G^2 = 47,77 > X_2^2(0,95) = 5,99$ ($p < 0,0001$) → on ne peut pas retirer le terme λ_{ik}^{XZ} du modèle déjà amputé de λ_{ijk}^{XYZ}
- Reste à tester si l'indépendance entre l'évaluation de la démocratie et l'intérêt pour la politique conditionnellement à la région du pays est une hypothèse plausible...

LR Statistics For Type 1 Analysis				
Source	Deviance	DF	Khi-2	Pr > Khi-2
Intercept	974.7867			
EVALDEMO2	903.8813	1	70.91	<.0001
polintr2	886.8208	1	17.06	<.0001
reg3	112.5095	2	774.31	<.0001
polintr2*reg3	93.3431	2	19.17	<.0001
EVALDEMO2*polintr2	50.4380	1	42.91	<.0001
EVALDEMO2*reg3	2.6659	2	47.77	<.0001
EVALDEM*polintr*reg3	0.0000	2	2.67	0.2637

LR Statistics For Type 3 Analysis			
Source	DF	Khi-2	Pr > Khi-2
EVALDEMO2	1	40.57	<.0001
polintr2	1	4.75	0.0293
reg3	2	257.25	<.0001
polintr2*reg3	2	26.01	<.0001
EVALDEMO2*polintr2	1	10.93	0.0009
EVALDEMO2*reg3	2	8.31	0.0157
EVALDEM*polintr*reg3	2	2.67	0.2637

TABLE 2x2x3 – exemple b Sélection du modèle

model count = X **Y|Z** X*Z **X*Y**/...;

LR Statistics For Type 1 Analysis				
Source	Deviance	DF	Khi-2	Pr > Khi-2
Intercept	974.7867			
EVALDEMO2	903.8813	1	70.91	<.0001
polintr2	886.8208	1	17.06	<.0001
reg3	112.5095	2	774.31	<.0001
polintr2*reg3	93.3431	2	19.17	<.0001
EVALDEMO2*reg3	53.7592	2	39.58	<.0001
EVALDEMO2*polintr2	2.6659	1	51.09	<.0001

$\Delta G^2 = 51,09 > X_1^2(0,95) = 3,84$ avec une p-valeur <0,0001
 → on rejette l'hypothèse $X \perp\!\!\!\perp Y|Z$.

On retient le modèle sans l'interaction de 3^{ème} ordre :

$$\log \mu_{ijk} = \gamma + \lambda_i^X + (\lambda_j^Y + \lambda_k^Z + \lambda_{jk}^{YZ}) + \lambda_{ij}^{XY} + \lambda_{ik}^{XZ}$$

On peut à présent examiner les estimations dans ce modèle final.

LR Statistics For Type 3 Analysis			
Source	DF	Khi-2	Pr > Khi-2
EVALDEMO2	1	83.20	<.0001
polintr2	1	4.24	0.0395
reg3	2	271.28	<.0001
polintr2*reg3	2	27.35	<.0001
EVALDEMO2*reg3	2	47.77	<.0001
EVALDEMO2*polintr2	1	51.09	<.0001

Analysis Of Maximum Likelihood Parameter Estimates									
Parameter			DDL	Valeur estimée	Erreur type	Wald 95% Confidence Limits		Wald Chi-Square	Pr > Khi-2
Intercept			1	5.1277	0.0719	4.9867	5.2687	5082.11	<.0001
EVALDEMO2	Low		1	-0.9477	0.1074	-1.1582	-0.7371	77.85	<.0001
polintr2	Not Interested		1	0.1873	0.0912	0.0086	0.3661	4.22	0.0400
reg3	Bxl		1	-1.1203	0.1459	-1.4062	-0.8343	58.97	<.0001
reg3	FI		1	0.7227	0.0847	0.5566	0.8888	72.73	<.0001
polintr2*reg3	Not Interested	Bxl	1	0.1359	0.1829	-0.2225	0.4943	0.55	0.4573
polintr2*reg3	Not Interested	FI	1	-0.4746	0.1053	-0.6809	-0.2682	20.32	<.0001
EVALDEMO2*reg3	Low	Bxl	1	-0.8631	0.2122	-1.2791	-0.4472	16.54	<.0001
EVALDEMO2*reg3	Low	FI	1	0.3735	0.1063	0.1652	0.5819	12.35	0.0004
EVALDEMO2*polintr2	Low	Not Interested	1	0.6992	0.0990	0.5052	0.8932	49.88	<.0001
Scale			0	1.0000	0.0000	1.0000	1.0000		

Globalement, les conclusions pour l'effet de la région sur evaldemo2 vont dans le même sens que *supra* (testé marginalement ou avec contrôle de blgtemg) même si les estimations et les bornes des IC sont légèrement modifiées par le contrôle de polintr2 (cf. aussi sortie avec ref = 'FI').

Par ailleurs, ***pour une région donnée***, l'estimation du logarithme du rapport des cotes d'évaluation négative pour les personnes pas intéressées Vs intéressées par la politique est > 0 et l'IC pour le log RC ne contient pas 0. Après passage à l'exponentielle : RC not inter. Vs inter = 2,012 avec IC 95% [1,66; 2,44], ce qui est très proche des valeurs obtenues à partir du modèle de régression logistique où l'on contrôlait l'âge (1,6 ; 2,35).

TABLE 4x2x3 – réponse polytomiale

Évaluation de la démocratie en fonction de la région du pays et l'intérêt pour la politique

Pour terminer, on peut affiner l'étude des facteurs qui influencent la manière dont les citoyens belges évaluent la démocratie dans leur pays en adoptant un autre choix de recatégorisation pour la variable réponse, susceptible de nuancer les conclusions par rapport à l'approche binaire, plus lisible, plus facile à interpréter mais souvent assez réductrice.

dmcntov(How democratic [country] is overall)	EVALDEMO4(How democratic Belgium is overall - 4 Levels)				
	Hardly democratic	Not democratic	Quite democratic	Very democratic	Total
Not at all democratic	0	36	0	0	36
1	0	15	0	0	15
2	0	33	0	0	33
3	0	58	0	0	58
4	0	90	0	0	90
5	257	0	0	0	257
6	260	0	0	0	260
7	0	0	414	0	414
8	0	0	440	0	440
9	0	0	0	176	176
Completely democratic	0	0	0	82	82
Total	517	232	854	258	1861

Choix pour la version "binaire"

- $0 \leq \text{DMCNTOV} \leq 6 \rightarrow \text{evaldemo2} = \text{'Low'}$
- $7 \leq \text{DMCNTOV} \leq 10 \rightarrow \text{evaldemo2} = \text{'High'}$

Choix pour la version "polytomiale"

- $0 \leq \text{DMCNTOV} \leq 4 \rightarrow \text{evaldemo4} = \text{'Not democratic'}$
- $5 \leq \text{DMCNTOV} \leq 6 \rightarrow \text{evaldemo4} = \text{'Hardly democratic'}$
- $7 \leq \text{DMCNTOV} \leq 8 \rightarrow \text{evaldemo4} = \text{'Quite democratic'}$
- $9 \leq \text{DMCNTOV} \leq 10 \rightarrow \text{evaldemo4} = \text{'Very democratic'}$

TABLE 4x2x3 – réponse polytomiale

Sélection du modèle

LR Statistics For Type 1 Analysis				
Source	Deviance	DF	Khi-2	Pr > Khi-2
Intercept	1483.4633			
EVALDEMO4	965.3691	3	518.09	<.0001
polintr2	948.3086	1	17.06	<.0001
reg3	173.9974	2	774.31	<.0001
polintr2*reg3	154.8309	2	19.17	<.0001
EVALDEMO4*polintr2	103.3377	3	51.49	<.0001
EVALDEMO4*reg3	9.6759	6	93.66	<.0001
EVALDEM*polintr*reg3	0.0000	6	9.68	0.1390

LR Statistics For Type 3 Analysis			
Source	DF	Khi-2	Pr > Khi-2
EVALDEMO4	3	59.72	<.0001
polintr2	1	1.31	0.2521
reg3	2	18.07	0.0001
polintr2*reg3	2	1.33	0.5135
EVALDEMO4*polintr2	3	18.72	0.0003
EVALDEMO4*reg3	6	50.28	<.0001
EVALDEM*polintr*reg3	6	9.68	0.1390

1. Sur base du type 1 et du type 3 (sur la gauche), il apparaît que l'interaction de 3^{ème} ordre n'est pas significative → retrait de λ_{ijk}^{XYZ} .
2. Sur base du type 1, on peut aussi conclure que l'indépendance entre evaldemo4 et reg3 *conditionnellement* à polintr2 n'est pas plausible → on ne peut pas retirer le terme λ_{ik}^{XZ} du modèle déjà amputé de λ_{ijk}^{XYZ}
3. Sur base du type 3 (ci-dessous), l'indépendance entre evaldemo4 et polintr2 *conditionnellement* à reg3 n'est pas plausible → on ne peut pas retirer le terme λ_{ij}^{XY} du modèle déjà amputé de λ_{ijk}^{XYZ}

LR Statistics For Type 3 Analysis			
Source	DF	Khi-2	Pr > Khi-2
EVALDEMO4	3	145.76	<.0001
polintr2	1	0.81	0.3670
reg3	2	30.77	<.0001
polintr2*reg3	2	31.59	<.0001
EVALDEMO4*polintr2	3	63.92	<.0001
EVALDEMO4*reg3	6	93.66	<.0001

On retient le modèle sans l'interaction de 3^{ème} ordre : $\log \mu_{ijk} = \gamma + \lambda_i^X + (\lambda_j^Y + \lambda_k^Z + \lambda_{jk}^{YZ}) + \lambda_{ij}^{XY} + \lambda_{ik}^{XZ}$

TABLE 4x2x3 – réponse polytomiale Interprétation de Evaluation*Intérêt politique

RC du désintérêt Vs intérêt pour la politique *pour une région donnée*

(absence d'interaction sign. du 3^{ème} ordre ⇔ pas nécessaire d'estimer ces RC séparément pour les 3 régions).

Evaluation de la démocratie en Belgique			
Very democratic	Quite democratic	Hardly democratic	Not democratic
1	$e^{0.4287} = 1,535$	$e^{0.9342} = 2,545$	$e^{1.2798} = 3,596$
$e^{-0.4287} = 0,651$	1	$e^{0.5054} = 1,658$	$e^{0.8511} = 2,342$
$e^{-0.9342} = 0,393$	$e^{-0.5054} = 0,603$	1	$e^{0.3457} = 1,413$
$e^{-1.2798} = 0,278$	$e^{-0.8511} = 0,427$	$e^{-0.3457} = 0,708$	1

Conclusion : pour une région du pays donnée, les personnes que ne s'intéressent pas à la politique choisissent plus souvent des modalités d'évaluation plus « négatives » par rapport aux personnes qui s'intéressent à la politique, et ce pour chaque comparaison des modalités d'évaluation de la démocratie.

RC du désintérêt Vs intérêt pour la politique *pour une région donnée*

Analysis Of Maximum Likelihood Parameter Estimates									
Parameter			DDL	Valeur estimée	Erreur type	Wald 95% Confidence Limits		Wald Chi-Square	Pr > Khi-2
EVALDEMO4*polintr2	Hardly democratic	Not Interested	1	0.9342	0.1607	0.6192	1.2492	33.78	<.0001
EVALDEMO4*polintr2	Not democratic	Not Interested	1	1.2798	0.1953	0.8971	1.6626	42.96	<.0001
EVALDEMO4*polintr2	Quite democratic	Not Interested	1	0.4287	0.1470	0.1406	0.7169	8.50	0.0035
EVALDEMO4*polintr2	Hardly democratic	Not Interested	1	0.5054	0.1148	0.2803	0.7305	19.37	<.0001
EVALDEMO4*polintr2	Not democratic	Not Interested	1	0.8511	0.1596	0.5383	1.1639	28.44	<.0001
EVALDEMO4*polintr2	Very democratic	Not Interested	1	-0.4287	0.1470	-0.7169	-0.1406	8.50	0.0035
EVALDEMO4*polintr2	Not democratic	Not Interested	1	0.3457	0.1699	0.0128	0.6786	4.14	0.0418
EVALDEMO4*polintr2	Quite democratic	Not Interested	1	-0.5054	0.1148	-0.7305	-0.2803	19.37	<.0001
EVALDEMO4*polintr2	Very democratic	Not Interested	1	-0.9342	0.1607	-1.2492	-0.6192	33.78	<.0001
EVALDEMO4*polintr2	Hardly democratic	Not Interested	1	-0.3457	0.1699	-0.6786	-0.0128	4.14	0.0418
EVALDEMO4*polintr2	Quite democratic	Not Interested	1	-0.8511	0.1596	-1.1639	-0.5383	28.44	<.0001
EVALDEMO4*polintr2	Very democratic	Not Interested	1	-1.2798	0.1953	-1.6626	-0.8971	42.96	<.0001

Ref = Very democratic

Ref = Quite democratic

Ref = Hardly democratic

Ref = Not democratic

TABLE 4x2x3 – réponse polytomiale Interprétation de Evaluation* Région

RC des Flamands Vs Wallons *pour un niveau d'intérêt pour la politique donné* (absence d'interaction sign. du 3^{ème} ordre ⇔ pas nécessaire d'estimer ces RC séparément pour les 2 niveaux de polintr2).

Evaluation de la démocratie en Belgique			
Very democratic	Quite democratic	Hardly democratic	Not democratic
1	$e^{0.8355} = 2,306$	$e^{1.0495} = 2,856$	$e^{0.9841} = 2,675$
$e^{-0.8355} = 0,434$	1	1	1
$e^{-1.0495} = 0,350$	1	1	1
$e^{-0.9841} = 0,374$	1	1	1

Conclusion : pour un niveau d'intérêt pour la politique donné, les Wallons et les Flamands se distinguent systématiquement par rapport au choix de la catégorie « très démocratique » : les Flamands choisissent plus souvent une catégorie « moins positive » que celle-là par rapport aux Wallons.

TABLE 4x2x3 – réponse polytomiale Interprétation de Evaluation* Région

RC des Flamands Vs Bruxellois pour un niveau d'intérêt pour la politique donné

Evaluation de la démocratie en Belgique			
Very democratic	Quite democratic	Hardly democratic	Not democratic
1	$e^{1,2396} = 3,454$	$e^{2,1390} = 8,491$	$e^{2,2339} = 9,336$
$e^{-1,2396} = 0,289$	1	$e^{0,8994} = 2,458$	$e^{0,9943} = 2,703$
$e^{-2,1390} = 0,118$	$e^{-0,8994} = 0,407$	1	1
$e^{-2,2339} = 0,107$	$e^{-0,9943} = 0,370$	1	1

Conclusion : pour un niveau d'intérêt pour la politique donné, les Flamands et les Bruxellois se distinguent par rapport aux choix de toutes les modalités excepté le contraste « Peu Vs pas démocratique » : les Flamands choisissent plus souvent une catégorie « moins positive » que celle-là par rapport aux Bruxellois.

Les valeurs des RC estimés entre la catégorie « très démocratique » et les 2 plus négatives sont assez remarquables. Il faut toutefois tenir compte de l'incertitude liée au processus inférentiel (ex : estimation **9,336** – IC : [4,47; 19,48] : les valeurs plausibles restent élevées même au voisinage de la borne inférieure).

TABLE 4x2x3 – réponse polytomiale

Interprétation de Evaluation* Région

RC des **Wallons Vs Bruxellois** pour un niveau d'intérêt pour la politique donné

* p=0,0587

Evaluation de la démocratie en Belgique			
Very democratic	Quite democratic	Hardly democratic	Not democratic
1	$e^{0,4042} = 1,498 *$	$e^{1,0896} = 2,973$	$e^{1,2498} = 3,49$
$e^{-0,4042} = 0,668 *$	1	$e^{0,6854} = 1,985$	$e^{0,8457} = 2,33$
$e^{-1,0896} = 0,336$	$e^{-0,6854} = 0,504$	1	1
$e^{-1,2498} = 0,287$	$e^{-0,8457} = 0,429$	1	1

Conclusion :

- Globalement, on observe les mêmes tendances que pour la comparaison Flamands Vs Bruxellois.
- Les valeurs des RC estimés sont toutefois moins remarquables quantitativement.
- Enfin, encore plus que pour la Flandre, on observe une polarisation des positions assez claire entre « Very – Quite » Vs « Hardly – Not » → une réponse binaire est suffisante pour distinguer ces deux groupes.

Conclusions

- Durant cette séance, nous avons étudié l'effet des variables explicatives « région », « intérêt pour la politique » et « appartenance à une minorité ethnique » sur la manière dont les citoyens belges évaluent plus ou moins positivement la démocratie dans leur pays au moyen de modèles log-linéaires.
- L'effet de la **région** s'est avéré statistiquement significatif et important d'un point de vue sociologique, quel que soit le modèle retenu (marginale, avec contrôle de blgetmg ou de polintr2). Globalement, les Bruxellois se distinguent par une évaluation plus positive que les 2 autres régions. Sur base du modèle avec réponse polytomiale, on sait également dire que les Wallons et Flamands ne se distinguent que par rapport au choix (moins fréquent pour les Flamands) des modalités très positives (9 et 10) sur l'échelle d'évaluation.
- L'effet de l'**intérêt pour la politique** a ici été étudié conditionnellement à la région. Les tendances restent similaires à celles qui avaient été mises en évidence grâce au modèle de régression logistique (sous contrôle de l'âge). Le modèle log-linéaire à réponse polytomiale nous apprend de surcroît qu'il existe une gradation : les RC des personnes désintéressées Vs intéressées sont d'autant plus élevés que l'évaluation est négative (au sens ordinal du terme).
- L'effet de l'**appartenance à une minorité ethnique** est statistiquement significatif isolément (dans le sens « les minorités ont une évaluation plus positive ») mais non significatif conditionnellement à la région. La faiblesse des effectifs de la catégorie minoritaire peut toutefois expliquer en partie la « disparition » de cet effet dans les modèles à plusieurs explicatives.