

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

Régression Logistique (modèles chapitre 1)

Interprétation des sorties

PLAN

- Régression logistique avec une explicative catégorielle
 - Evaldemo2 en fonction de educ3
- Régression logistique avec deux explicatives catégorielles
 - Evaldemo2 en fonction de educ3 + hincfel2
 - Evaldemo2 en fonction de hincfel2 + plinsoc3
- Régression logistique avec une explicative continue
 - Evaldemo2 en fonction de agea
- Régression logistique avec explicatives mixtes
 - Evaldemo2 en fonction de agea + polintr2

RÉGRESSION LOGISTIQUE AVEC UNE EXPLICATIVE CATÉGORIELLE

Évaluation de la démocratie en fonction du niveau d'éducation

On repart de la table de contingence

educ3(Level of education - 3 Levels)	EVALDEMO2(How democratic Belgium is overall - 2 Levels)		
	High	Low	Total
High	347 72.90	129 27.10	476
Low	286 50.09	285 49.91	571
Middle	470 58.68	331 41.32	801
Total	1103	745	1848
Fréquence manquante = 21			

→ Hypothèse à tester : la proportion de répondants qui évaluent négativement la démocratie en Belgique est plus faible quand le niveau d'éducation est plus élevé (différence plus marquée entre les extrêmes);

En bref, questions à se poser :

1. Impact significatif de la variable explicative testée sur la réponse?
2. Si oui, quels groupes diffèrent significativement?
3. Quelles sont les valeurs plausibles pour les rapports de cotes des groupes significativement différents?

1. La variable explicative testée a-t-elle un effet *significatif* sur la réponse?

- Question de recherche : le niveau d'éducation (3 niveaux) a-t-il un impact sur la tendance à évaluer négativement la démocratie en Belgique (2 niveaux)?

Testing Global Null Hypothesis: BETA=0			
Test	Chi-Square	DDL	Pr > Khi-2
Rapport de vrais	58.1135	2	<.0001
Score	56.7397	2	<.0001
Wald	55.3283	2	<.0001

Type 3 Analysis of Effects			
Effect	DDL	Wald Chi-Square	Pr > Khi-2
educ3	2	55.3283	<.0001

- Ici, une seule variable explicative testée → les deux sorties livrent la même information.
- Hypothèse nulle : les cotes de succès (↔ évaluation négative de la démocratie) sont les mêmes dans les 3 groupes.
- Pr > Khi-2 = probabilité d'observer une valeur au moins aussi grande pour la statistique du Chi-Carré si H_0 est vraie ; p-valeur très faible (< 0.05) → Rejet de H_0 ↔ **Le niveau d'éducation a un effet significatif sur la variable réponse ↔ Au moins 2 groupes parmi les 3 sont statistiquement significativement différents.**
- NB : Si les conclusions diffèrent entre la méthode de WALD et celle du test de rapport de vraisemblance, il est préférable de privilégier la seconde quand l'approximation normale est peu fiable (fréquences petites dans la table de contingence).

2. Quels sont les contrastes significatifs?

- Sortie 1 (proc logistic par défaut) :

Analysis of Maximum Likelihood Estimates							
Parameter		DDL	Estimate	Standard Error	Wald Chi-Square	Pr > Khi-2	Exp(Est)
Intercept		1	-0.3506	0.0718	23.8755	<.0001	0.704
educ3	High	1	-0.6389	0.1256	25.8634	<.0001	0.528
educ3	Low	1	0.3471	0.1102	9.9133	0.0016	1.415

- Niveau de référence défini au niveau de l'instruction « class » : Middle → Par défaut, analyse des contrastes « High Vs Middle » et « Low Vs Middle »
 - High Vs Middle : estimation du log de RC < 0 et Pr >Khi-2 < 0.05 → conclusion : les personnes ayant un niveau d'éducation « élevé » ont une cote de succès significativement plus faible que les personnes avec un niveau d'éducation « moyen ».
 - Low Vs Middle : estimation du log de RC > 0 et Pr >Khi-2 < 0.05 → conclusion : les personnes ayant un niveau d'éducation « faible » ont une cote de succès significativement plus grande que les personnes avec un niveau d'éducation « moyen ».
- On ne peut rien dire sur base de cette sortie à propos du contraste « High Vs Low »

2. Quels sont les contrastes significatifs?

- Sortie 2 (via l'instruction supplémentaire contrast) :

Contrast Test Results			
Contrast	DDL	Wald Chi-Square	Pr > Khi-2
Middle Vs High	1	25.8634	<.0001
Low Vs Middle	1	9.9133	0.0016
Low Vs High	1	55.1172	<.0001

- On retrouve les mêmes valeurs pour la statistique du chi-carré de Wald et les p-valeurs pour les 2 comparaisons déjà évoquées supra (même si inversion des catégories pour le premier contraste).
 - De plus, on peut à présent conclure que le contraste entre les deux catégories extrêmes est également significatif!
- Puisqu'un effet significatif a pu être mis en évidence pour toutes les comparaisons 2 à 2, on peut maintenant s'intéresser à l'ensemble des valeurs plausibles pour les 3 (log de) RC.

3. Quelles sont les valeurs plausibles pour les rapports de cotes des groupes significativement différents?

Contrast Estimation and Testing Results by Row									
Contrast	Type	Row	Estimate	Standard Error	Alpha	Confidence Limits		Wald Chi-Square	Pr > Khi-2
Middle Vs High	PARM	1	0.6389	0.1256	0.05	0.3927	0.8851	25.8634	<.0001
Middle Vs High	EXP	1	1.8944	0.2380	0.05	1.4809	2.4233	25.8634	<.0001
Low Vs Middle	PARM	1	0.3471	0.1102	0.05	0.1310	0.5632	9.9133	0.0016
Low Vs Middle	EXP	1	1.4150	0.1560	0.05	1.1400	1.7563	9.9133	0.0016
Low Vs High	PARM	1	0.9860	0.1328	0.05	0.7257	1.2463	55.1172	<.0001
Low Vs High	EXP	1	2.6805	0.3560	0.05	2.0662	3.4775	55.1172	<.0001

Dans la sortie, deux lignes pour chaque contraste :

- PARM : log du RC (IC ne contient pas 0 si $Pr > Khi-2 < 0.05$)
- EXP : RC (IC ne contient pas 1 si $Pr > Khi-2 < 0.05$)

- Choix dans la définition des contrastes : numérateur = niveau d'éducation le plus faible des 2
- Tous les intervalles pour les RC ne contenant pas 1 et étant > 1 , on en déduit que la cote de succès est d'autant plus grande que le niveau d'éducation est faible (rappel : cote de succès = probabilité d'évaluer négativement plutôt que positivement la démocratie en Belgique).
- Quid en termes de « quantité »? Par exemple, on peut dire (avec 95% de chance de ne pas se tromper) que les individus avec un niveau d'éducation « faible » ont une cote de succès entre 2 et 3,5 fois plus grande que les individus avec un niveau d'éducation « élevé » dans la population belge.

Remarque : à partir des estimations issues de la régression logistique, on retrouve les fréquences relatives en ligne de la table de contingence (probabilités estimées pour la variable réponse conditionnellement à la catégorie k de la variable explicative : $\hat{\pi}_k$)

Analysis of Maximum Likelihood Estimates							
Parameter		DDL	Estimate	Standard Error	Wald Chi-Square	Pr > Khi-2	Exp(Est)
Intercept		1	-0.3506	0.0718	23.8755	<.0001	0.704
educ3	High	1	-0.6389	0.1256	25.8634	<.0001	0.528
educ3	Low	1	0.3471	0.1102	9.9133	0.0016	1.415

Table de educ3 par EVALDEMO2			
educ3(Level of education - 3 Levels)	EVALDEMO2(How democratic Belgium is overall - 2 Levels)		
	High	Low	Total
High	347 72.90	129 27.10	476
Low	286 50.09	285 49.91	571
Middle	470 58.68	331 41.32	801
Total	1103	745	1848

Fréquence manquante = 21

Régression logistique : Log de la cote de succès pour la catégorie k	Table de contingence : probabilité de succès pour la catégorie k
$\log\left(\frac{\hat{\pi}_1}{1-\hat{\pi}_1}\right) = \hat{\theta}_1 = \hat{\mu} = -0,3506$	$\hat{\pi}_1 = \frac{e^{-0,3506}}{1+e^{-0,3506}} = 0,413$
$\log\left(\frac{\hat{\pi}_2}{1-\hat{\pi}_2}\right) = \hat{\theta}_2 = \hat{\mu} + \hat{\alpha}_2$ $= -0,3506 - 0,6389 = -0,9895$	$\hat{\pi}_2 = \frac{e^{-0,9895}}{1+e^{-0,9895}} = 0,271$
$\log\left(\frac{\hat{\pi}_3}{1-\hat{\pi}_3}\right) = \hat{\theta}_3 = \hat{\mu} + \hat{\alpha}_3$ $= -0,3506 + 0,3471 = -0,0035$	$\hat{\pi}_3 = \frac{e^{-0,0035}}{1+e^{-0,0035}} = 0,499$

RÉGRESSION LOGISTIQUE AVEC DEUX EXPLICATIVES CATÉGORIELLES

1. Évaluation de la démocratie en fonction du niveau d'éducation et du niveau d'aisance (revenu)

Table 1 de educ3 par EVALDEMO2			
Contrôle pour hincfel2=Confortable			
educ3(Level of education - 3 Levels)	EVALDEMO2(How democratic Belgium is overall - 2 Levels)		
	High	Low	Total
High	323 73.91	114 26.09	437
Low	199 53.49	173 46.51	372
Middle	367 59.97	245 40.03	612
Total	889	532	1421
Fréquence manquante = 14			

Table 2 de educ3 par EVALDEMO2			
Contrôle pour hincfel2=Difficult			
educ3(Level of education - 3 Levels)	EVALDEMO2(How democratic Belgium is overall - 2 Levels)		
	High	Low	Total
High	24 61.54	15 38.46	39
Low	87 44.16	110 55.84	197
Middle	102 54.55	85 45.45	187
Total	213	210	423
Fréquence manquante = 6			

% de Succès observés		
	Hincfel2 « Confortable »	Hincfel2 « Difficult »
Educ3 = « High »	26,09	38,46
Educ3 = « Middle »	40,03	45,45
Educ3 = « Low »	46,51	55,84

Sélection du modèle

Type 3 Analysis of Effects			
Effect	DDL	Wald Chi-Square	Pr > Khi-2
educ3	2	38.0373	<.0001
hincfel2	1	1.7335	0.1880
educ3*hincfel2	2	0.9591	0.6191

Type 3 Analysis of Effects			
Effect	DDL	Wald Chi-Square	Pr > Khi-2
educ3	2	43.4255	<.0001
hincfel2	1	7.8968	0.0050

- Pour un niveau d'éducation et un niveau d'aisance donnés, l'interaction entre ces deux explicatives ***dans leurs effets sur l'évaluation de la démocratie*** n'est pas significative.
 - Il faut supprimer le terme d'interaction de l'équation et refaire tourner le modèle.
- On constate alors que l'effet de hincfel2, qui n'était pas significatif sous contrôle du terme d'interaction, devient significatif une fois qu'on ne contrôle plus que educ3 (s'interroger sur la raison – suggestion : examen de la table de contingence).
 - On conserve le modèle avec les deux explicatives (sans le terme d'interaction) pour la suite des analyses.

Attention TYPE 3 ⇔ évalue le caractère significatif de l'effet d'un terme en contrôlant tous les autres simultanément (que ceux-ci soient inclus avant ou après dans l'équation) → le retrait d'un terme affecte les statistiques du Chi² et les p-valeurs calculées pour tous les termes restants !!!

2. Quels sont les contrastes significatifs pour le modèle retenu?

Log (cote) _{jk}	k = 1 : Aisance = 'Confortable'	k = 2 : Aisance = 'Difficult'
j = 1 : Educ = 'Middle'	μ	$\mu + \beta_2$
j = 2 : Educ = 'High'	$\mu + \alpha_2$	$\mu + \alpha_2 + \beta_2$
j = 3 : Educ = 'Low'	$\mu + \alpha_3$	$\mu + \alpha_3 + \beta_2$

Class Level Information			
Class	Value	Design Variables	
educ3	High	1	0
	Low	0	1
	Middle	0	0
hincfel2	Confortable	0	
	Difficult	1	

Pour un niveau donné de l'autre explicative :

- Contrast Educ « Middle Vs High » : $-\alpha_2$
- Contrast Educ « Low Vs Middle » : α_3
- Contrast Educ « Low Vs High » : $\alpha_3 - \alpha_2$
- Constrat Hincfel « Difficult Vs confortable » : β_2

```
educ3 -1 0
educ3  0 1
educ3 -1 1
hincfel2 1
```

2. Quels sont les contrastes significatifs pour le modèle retenu?

Contrast Test Results			
Contrast	DDL	Wald Chi-Square	Pr > Khi-2
Educ Middle Vs High	1	21.5019	<.0001
Educ Low Vs Middle	1	7.5577	0.0060
Educ Low Vs High	1	43.1456	<.0001
HINCFEL Difficult Vs Comfortable	1	7.8968	0.0050

- Les 3 contrastes relatifs au niveau d'éducation restent significatifs ***pour un niveau d'aisance donné;***
- Le contraste relatif au niveau d'aisance est significatif, ***pour un niveau d'éducation donné*** (logique : la variable « hincfel2 » s'est révélée significative et elle ne possède que 2 niveaux dans cette analyse → 1 seul contraste, forcément significatif).

3. Quelles sont les valeurs plausibles pour les rapports de cotes des groupes significativement différents?

Contrast Estimation and Testing Results by Row									
Contrast	Type	Row	Estimate	Standard Error	Alpha	Confidence Limits		Wald Chi-Square	Pr > Khi-2
Educ Middle Vs High	PARM	1	0.5890	0.1270	0.05	0.3401	0.8380	21.5019	<.0001
Educ Middle Vs High	EXP	1	1.8023	0.2289	0.05	1.4050	2.3118	21.5019	<.0001
Educ Low Vs Middle	PARM	1	0.3061	0.1114	0.05	0.0879	0.5244	7.5577	0.0060
Educ Low Vs Middle	EXP	1	1.3582	0.1512	0.05	1.0919	1.6894	7.5577	0.0060
Educ Low Vs High	PARM	1	0.8952	0.1363	0.05	0.6281	1.1623	43.1456	<.0001
Educ Low Vs High	EXP	1	2.4477	0.3336	0.05	1.8740	3.1972	43.1456	<.0001
HINCFEL Difficult Vs Confortable	PARM	1	0.3241	0.1153	0.05	0.0981	0.5502	7.8968	0.0050
HINCFEL Difficult Vs Confortable	EXP	1	1.3828	0.1595	0.05	1.1030	1.7336	7.8968	0.0050

- Pour un niveau d'aisance donné, la probabilité d'évaluer négativement la démocratie en Belgique (plutôt que positivement) diminue significativement avec le niveau d'éducation.

NB : on constate cependant que les estimations sont légèrement affectées par le contrôle du niveau d'aisance (ex. « Low Vs Middle » : IC modèle1 : RC \in [2.07 ; 3.48] Vs modèle2 [1.87 ; 3.20])

- Pour un niveau d'éducation donné, les personnes qui ont des difficultés à vivre avec leur revenu actuel ont une cote de succès (entre 10% et 73%) supérieure à celles vivent plus confortablement.

RÉGRESSION LOGISTIQUE AVEC DEUX EXPLICATIVES CATÉGORIELLES

2. Évaluation de la démocratie en fonction du niveau d'aisance et de la place que l'individu estime occuper dans la société

Table 1 de plinsoc3 par EVALDEMO2			
Contrôle pour hincfel2=Confortable			
plinsoc3(Your place in society - 3 Levels)	EVALDEMO2(How democratic Belgium is overall - 2 Levels)		
	High	Low	Total
Bottom	39 43.82	50 56.18	89
Middle	704 62.69	419 37.31	1123
Top	149 69.63	65 30.37	214
Total	892	534	1426
Fréquence manquante = 9			

Table 2 de plinsoc3 par EVALDEMO2			
Contrôle pour hincfel2=Difficult			
plinsoc3(Your place in society - 3 Levels)	EVALDEMO2(How democratic Belgium is overall - 2 Levels)		
	High	Low	Total
Bottom	60 47.24	67 52.76	127
Middle	148 53.62	128 46.38	276
Top	6 27.27	16 72.73	22
Total	214	211	425
Fréquence manquante = 4			

% de Succès observés		
Plinsoc3\Hincfel2	« Confortable »	« Difficult »
« Top »	30,37	72,73
« Middle »	37,31	46,38
« Bottom »	56,18	52,76

Sélection du modèle

Type 3 Analysis of Effects			
Effect	DDL	Wald Chi-Square	Pr > Khi-2
hincfel2	1	7.6006	0.0058
plinsoc3	2	17.2427	0.0002
hincfel2*plinsoc3	2	11.5924	0.0030

- Les deux explicatives testées ainsi que leur interaction sur la réponse sont significatives, et ce, en contrôlant respectivement tous les autres termes du modèle.
 - On conserve donc ce modèle « saturé » pour la suite des analyses

Estimation des paramètres et p-valeurs

Analysis of Maximum Likelihood Estimates								
Parameter			DDL	Estimate	Standard Error	Wald Chi-Square	Pr > Khi-2	Exp(Est)
Intercept			1	-0.5189	0.0617	70.7272	<.0001	0.595
hincfel2	Difficult		1	0.3737	0.1356	7.6006	0.0058	1.453
plinsoc3	Bottom		1	0.7673	0.2224	11.9078	0.0006	2.154
plinsoc3	Top		1	-0.3105	0.1609	3.7231	0.0537	0.733
hincfel2*plinsoc3	Difficult	Bottom	1	-0.5118	0.3092	2.7399	0.0979	0.599
hincfel2*plinsoc3	Difficult	Top	1	1.4364	0.5193	7.6524	0.0057	4.206

Catégories de référence :

- niveau d'aisance : « revenu confortable »
- place dans la société : « middle »

Log (cote)jk

j=1 : hincfel2 = 'Confortable'

j=2 : hincfel2 = 'Difficult'

k = 1 :
plinsoc3 = 'Middle'

μ

$\mu + \alpha_2$

k = 2 :
plinsoc3 = 'Bottom'

$\mu + \beta_2$

$\mu + \alpha_2 + \beta_2 + \gamma_{22}$

k = 3 :
plinsoc3 = 'Top'

$\mu + \beta_3$

$\mu + \alpha_2 + \beta_3 + \gamma_{23}$

Interprétation des paramètres

Pour les individus qui se voient plutôt au milieu de l'échelle sociale et qui vivent 'confortablement' :

- $\mu < 0$ avec $p < 0,05 \rightarrow$ le succès (évaluation négative de la démocratie) est moins probable que l'échec (évaluation positive de la démocratie)

Pour les individus qui se voient plutôt au milieu de l'échelle sociale :

- $\alpha_2 > 0$ avec $p < 0,05 \rightarrow$ la cote de succès des individus qui ont des difficultés financières est plus grande que celle de ceux qui vivent confortablement (idem supra).

Parmi les individus qui vivent 'confortablement' :

- $\beta_2 > 0$ avec $p < 0,05 \rightarrow$ les individus qui se voient en bas de l'échelle sociale ont une cote de succès supérieure par rapport à ceux qui se voient plutôt au milieu.
- $\beta_3 < 0$ avec $p \approx 0,05 \rightarrow$ les individus qui se voient en haut de l'échelle sociale ont une cote de succès inférieure par rapport à ceux qui se voient plutôt au milieu mais effet moins net (proximité seuil 0,05 : il est plus difficile de trancher).

Interprétation des termes d'interaction :

- $\gamma_{22} < 0$ mais $p > 0,05 \rightarrow$ on ne peut pas dire que l'effet de la variable hincfel2 est modulé par le passage de 'Middle' à 'Bottom'
- $\gamma_{23} > 0$ avec $p < 0,05 \rightarrow$ par contre, quand on passe de 'Middle' à 'Top' dans la hiérarchie sociale, le fait d'avoir des difficultés financières va augmenter la plausibilité d'évaluer négativement (plutôt que positivement) la démocratie.

NB : On peut également présenter les résultats sous la forme des cotes de succès estimées pour chaque profil

Analysis of Maximum Likelihood Estimates								
Parameter			DDL	Estimate	Standard Error	Wald Chi-Square	Pr > Khi-2	Exp(Est)
Intercept			1	-0.5189	0.0617	70.7272	<.0001	0.595
hincfel2	Difficult		1	0.3737	0.1356	7.6006	0.0058	1.453
plinsoc3	Bottom		1	0.7673	0.2224	11.9078	0.0006	2.154
plinsoc3	Top		1	-0.3105	0.1609	3.7231	0.0537	0.733
hincfel2*plinsoc3	Difficult	Bottom	1	-0.5118	0.3092	2.7399	0.0979	0.599
hincfel2*plinsoc3	Difficult	Top	1	1.4364	0.5193	7.6524	0.0057	4.206

COTE_{jk}

j=1 : hincfel2 = 'Comfortable'

j=2 : hincfel2 = 'Difficult'

k = 1 :
plinsoc3 = 'Middle'

$$e^{-0,5189}$$

$$= 0,595$$

$$e^{-0,5189+0,3737}$$

$$= 0,865$$

k = 2 :
plinsoc3 = 'Bottom'

$$e^{-0,5189+0,7673}$$

$$= 1,282$$

$$e^{-0,5189+0,3737+0,7673-0,5118}$$

$$= 1,117$$

k = 3 :
plinsoc3 = 'Top'

$$e^{-0,5189-0,3105}$$

$$= 0,436$$

$$e^{-0,5189+0,3737-0,3105+1,4364}$$

$$= 2,666$$

NB : si on change la catégorie de référence, cela va avoir un impact sur les valeurs estimées des paramètres mais on retrouve toujours les mêmes estimations des cotes de succès pour chaque profil!!!

Analysis of Maximum Likelihood Estimates								
Parameter			DDL	Estimate	Standard Error	Wald Chi-Square	Pr > Khi-2	Exp(Est)
Intercept			1	-0.8294	0.1486	31.1376	<.0001	0.436
hincfel2	Difficult		1	1.8101	0.5012	13.0413	0.0003	6.111
plinsoc3	Bottom		1	1.0779	0.2603	17.1525	<.0001	2.938
plinsoc3	Middle		1	0.3105	0.1609	3.7231	0.0537	1.364
hincfel2*plinsoc3	Difficult	Bottom	1	-1.9482	0.5731	11.5551	0.0007	0.143
hincfel2*plinsoc3	Difficult	Middle	1	-1.4364	0.5193	7.6524	0.0057	0.238

(cote)jk

j=1 : hincfel2 = 'Confortable'

j=2 : hincfel2 = 'Difficult'

k = 1 :
plinsoc3 = 'Top'

$$e^{-0,8294} = 0,436$$

$$e^{-0,8294+1,8101} = 2,666$$

k = 2 :
plinsoc3 = 'Bottom'

$$e^{-0,8294+1,0779} = 1,282$$

$$e^{-0,8294+1,8101+1,0779-1,9482} = 1,117$$

k = 3 :
plinsoc3 = 'Middle'

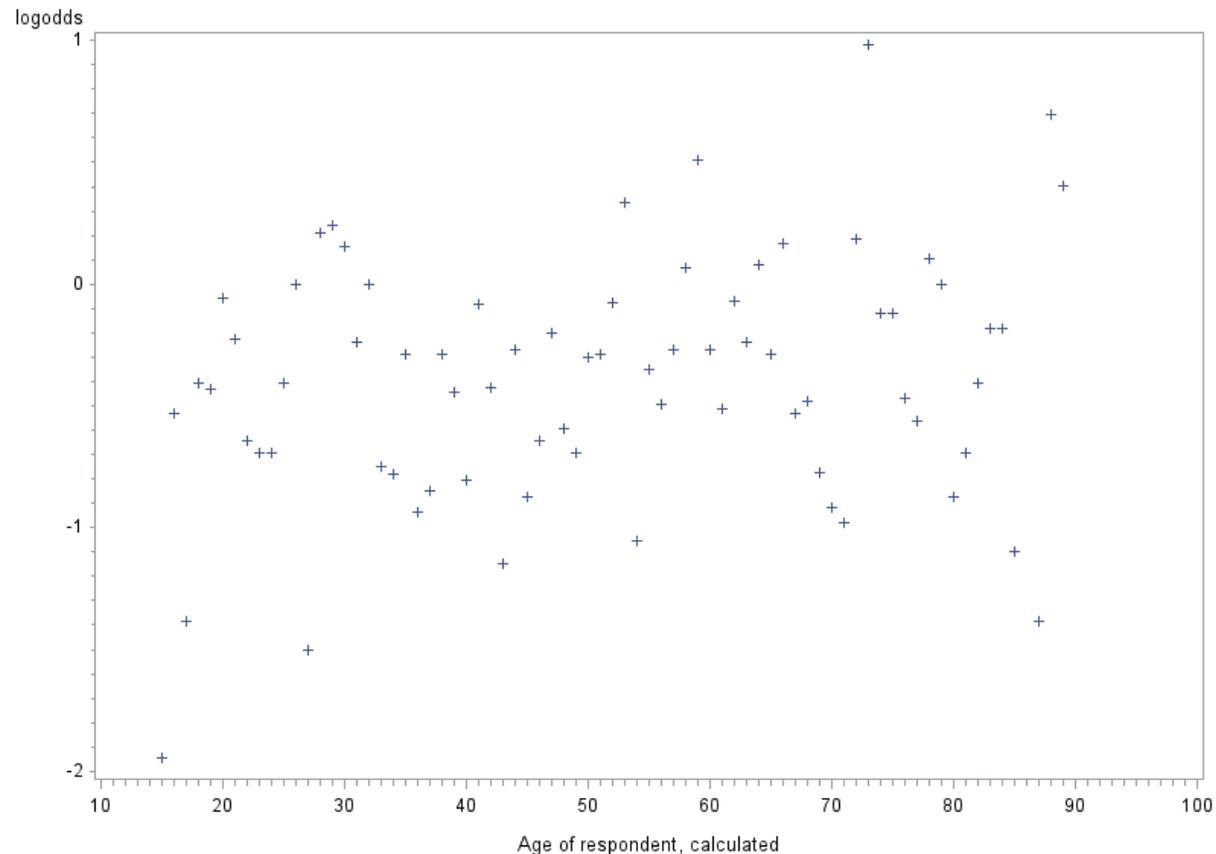
$$e^{-0,8294+0,3105} = 0,595$$

$$e^{-0,8294+1,8101+0,3105-1,4364} = 0,865$$

RÉGRESSION LOGISTIQUE AVEC UNE EXPLICATIVE CONTINUE

Évaluation de la démocratie en fonction de l'âge

- 1^{ère} étape : vérification de l'hypothèse de linéarité (évolution des log(cotes) avec l'âge:



- Pas de contre-indication manifeste à la mise en œuvre du modèle
- Si l'âge a un effet statistiquement significatif, celui-ci doit être faible en pratique (points dispersés, pas de tendance croissante ou décroissante clairement identifiable)

RÉGRESSION LOGISTIQUE AVEC UNE EXPLICATIVE CONTINUE

Évaluation de la démocratie en fonction de l'âge

- Modèle

Testing Global Null Hypothesis: BETA=0			
Test	Chi-Square	DDL	Pr > Khi-2
Rapport de vrais	2.5301	1	0.1117
Score	2.5310	1	0.1116
Wald	2.5285	1	0.1118

- Quelle que soit la méthode utilisée, l'effet de la variable « âge » n'est pas significatif au seuil 0,05 (quand c'est la seule variable insérée dans le modèle).

Analysis of Maximum Likelihood Estimates						
Parameter	DDL	Estimate	Standard Error	Wald Chi-Square	Pr > Khi-2	Exp(Est)
Intercept	1	-0.5824	0.1272	20.9632	<.0001	0.559
agea	1	0.00395	0.00248	2.5285	0.1118	1.004

- Intercept : $\hat{\beta}_0 = -0,5824 \Leftrightarrow \log(\text{cote})$ quand pour un âge = 0 (aucun sens en pratique)
- agea : $\hat{\beta}_1 = 0,00395 \Leftrightarrow \log(\text{rapport des cotes})$ d'un âge x+1 par rapport à un âge x (> 0 échantillon mais Pr > khi-2 > 0,05).

Odds Ratio Estimates			
Effect	Point Estimate	95% Wald Confidence Limits	
agea	1.004	0.999	1.009

- Passage à l'exponentielle : l'augmentation d'une unité de l'âge induit une augmentation de la cote de succès de 0,4% dans l'échantillon mais « 1 » fait partie des valeurs plausibles pour le rapport de cotes au niveau de la population

RÉGRESSION LOGISTIQUE AVEC EXPLICATIVES MIXTES

Évaluation de la démocratie en fonction de l'âge et de l'intérêt pour la politique

- NB : on suppose la linéarité des log(cotes) en fonction de l'âge dans les 2 groupes (interested or not)
- **Sélection du modèle** : examen des effets de type 3

Type 3 Analysis of Effects			
Effect	DDL	Wald Chi-Square	Pr > Khi-2
agea	1	7.6764	0.0056
polintr2	1	1.5035	0.2201
agea*polintr2	1	1.6874	0.1939

Type 3 Analysis of Effects			
Effect	DDL	Wald Chi-Square	Pr > Khi-2
agea	1	6.2827	0.0122
polintr2	1	45.5793	<.0001

- Pour un âge et un niveau d'intérêt pour la politique donnés, l'interaction entre ces deux explicatives ***dans leurs effets sur l'évaluation de la démocratie*** n'est pas significative.
 - Il faut supprimer le terme d'interaction de l'équation et refaire tourner le modèle.
- On constate alors que l'effet de polintr2, qui n'était pas significatif sous contrôle du terme d'interaction, devient significatif une fois qu'on ne contrôle plus que agea.
- L'effet de l'âge est aussi significatif dès lors que l'on contrôle le niveau d'intérêt pour la politique.
 - On conserve le modèle avec les deux explicatives (sans le terme d'interaction) pour la suite des analyses.

RÉGRESSION LOGISTIQUE AVEC EXPLICATIVES MIXTES

Évaluation de la démocratie en fonction de l'âge et de l'intérêt pour la politique

- Modèle : $\log\left(\frac{\pi_x}{1-\pi_x}\right) = (\beta_0 + \alpha_k) + \beta_1 x ; (\alpha_1 = 0)$

Analysis of Maximum Likelihood Estimates							
Parameter		DDL	Estimate	Standard Error	Wald Chi-Square	Pr > Khi-2	Exp(Est)
Intercept		1	-1.0686	0.1482	51.9665	<.0001	0.343
agea		1	0.00636	0.00254	6.2827	0.0122	1.006
polintr2	Not Interested	1	0.6622	0.0981	45.5793	<.0001	1.939

Odds Ratio Estimates			
Effect	Point Estimate	95% Wald Confidence Limits	
agea	1.006	1.001	1.011
polintr2 Not Interested vs Interested	1.939	1.600	2.350

- Estimations ponctuelles :
 - $\hat{\alpha}_2 = 0,6622$
 - $\hat{\beta}_1 = 0,00636$
- IC pour les rapports de cotes :
 - Pour un niveau d'intérêt pour la politique donné, l'augmentation d'une unité de l'âge induit une augmentation de la cote de succès comprise entre 0,1% et 1,1% (faible en pratique).
 - Pour un âge donné, les personnes qui ne sont pas intéressées par la politique ont une cote de succès entre 1,6 et 2,35 fois supérieure par rapport à celles qui sont intéressées.

Conclusions

- Parmi les explicatives testées dans ces premiers modèles (educ3, hincfel2, plinsoc3, agea et polintr2), l'effet de l'**âge** sur la probabilité d'évaluer négativement (plutôt que positivement) la démocratie en Belgique est le moins tranché, aussi bien statistiquement (p-valeur < 0,05 ou > 0,05 selon le modèle retenu) que sociologiquement (faible valeur des coefficients estimés).
- L'effet du **niveau d'éducation** va dans le sens de l'hypothèse testée : « la cote de succès est d'autant plus grande que le niveau d'éducation est faible » (attention, pas *continu*). De plus, les tendances et les coefficients restent similaires quand on contrôle le niveau d'aisance.
- L'effet du **niveau d'aisance** va aussi dans le sens de l'analyse descriptive : globalement, le fait d'avoir des difficultés financières augmente significativement la probabilité d'évaluer négativement la démocratie. Ceci est vrai quel que soit le niveau d'éducation de la personne. Cette variable entretient cependant des interactions non négligeable avec la variable plinsoc3 dans ses effets sur la réponse.
- Globalement, plus la personne estime avoir une **position** basse **dans la hiérarchie sociale**, plus sa cote de succès est faible. Néanmoins, le fait d'avoir des difficultés financières peut inverser la tendance pour la catégorie des personnes qui estiment être « au top » de l'échelle sociale.
- Pour terminer, l'effet de l'**intérêt pour la politique** va aussi dans le sens de l'analyse descriptive : les personnes qui se déclarent pas ou peu intéressées par la politique ont une cote de succès significativement supérieure à celles qui se déclarent plutôt ou très intéressées par la politique (pour un âge donné).