

Déterminants de l'accès au crédit au Pays-Bas

Charbel AHOUANDOKOUN, Nazifou AFOLABI

22 juillet, 2023

1- Analyse descriptive

A- Variables quantitatives

Table 1: Statistiques descriptive variable Age

Statistique	Valeur
Min	15.00
1er Quartile	38.75
Médiane	56.00
Moyenne	53.55
3è Quartile	68.00
Max	92.00

Commentaire :

La moitié des individus de notre étude ont plus de 56 ans, on a donc un échantillon plutôt vieux (75% au ont plus de 38 ans) ce qui est normal pour ce genre d'étude. L'âge minimum est de 15 ans, le maximum est de 92 années. On décompte quand même 22 mineurs dans notre base.

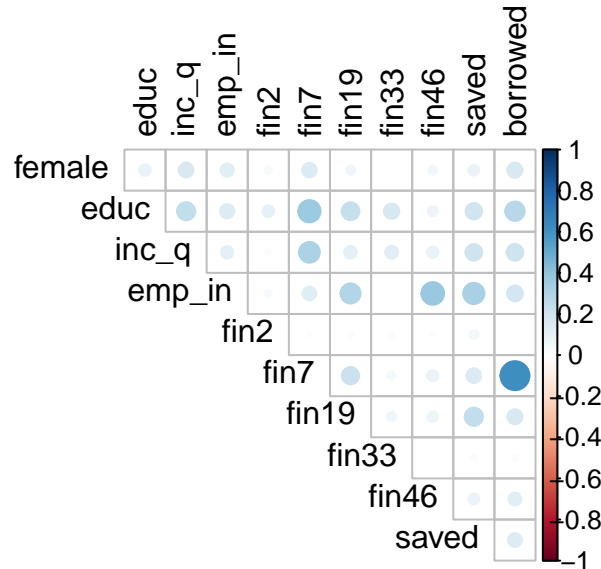
B- Variables qualitatives

Characteristic	**N = 1,000**
female	
male	531 (53%)
female	469 (47%)
educ	
Primary	43 (4.3%)
Secondary	577 (58%)
Tertiary	374 (38%)
Unknown	6
inc_q	
first_q	151 (15%)
second_q	173 (17%)
middle_q	208 (21%)
fourth_q	214 (21%)
fifth_q	254 (25%)
emp_in	
out_of_wf	390 (39%)
wf	610 (61%)

Commentaire

Notre base contient 53% d'individus de sexe masculin. Le niveau d'éducation le plus représenté est le secondaire (58% des individus) et le moins représenté est le primaire (4,3%). 25% des personnes enquêtées font partie des 20% ayant le revenu le plus élevé dans le pays.

C- Matrice de Crammer

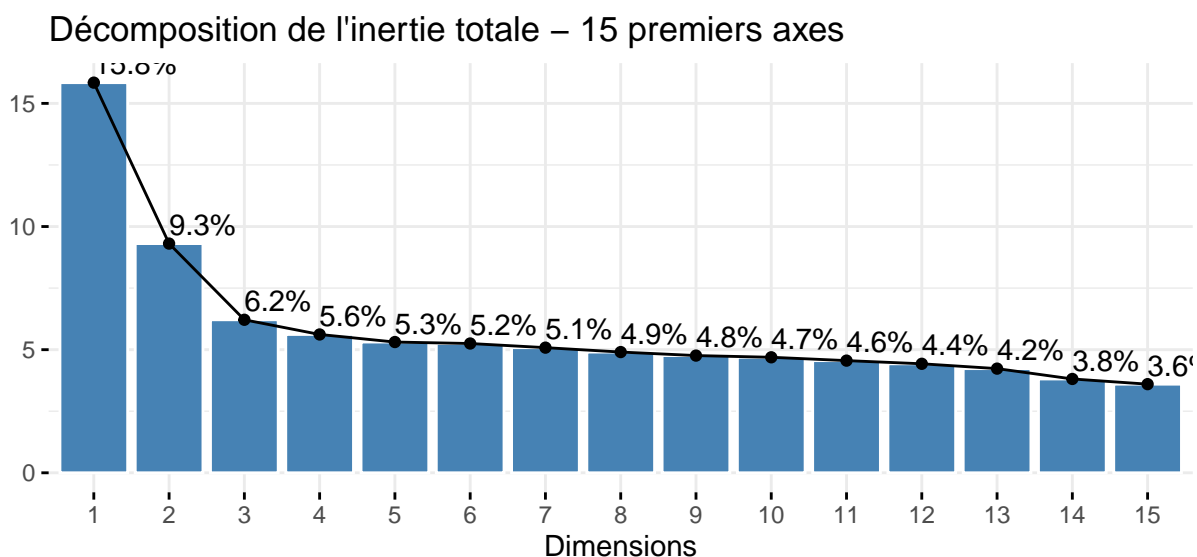


Commentaire

La matrice de Cramer présente les liens entre les variables qualitatives de notre base. On y voit notamment une très forte 'corrélation' entre *fin7* et *borrowed*. Aussi, les variables les moins liées à *fin19*, notre variable d'intérêt sont : *fin2*, *fin33* et *fin46*.

D- Analyse des Correspondances Multiples

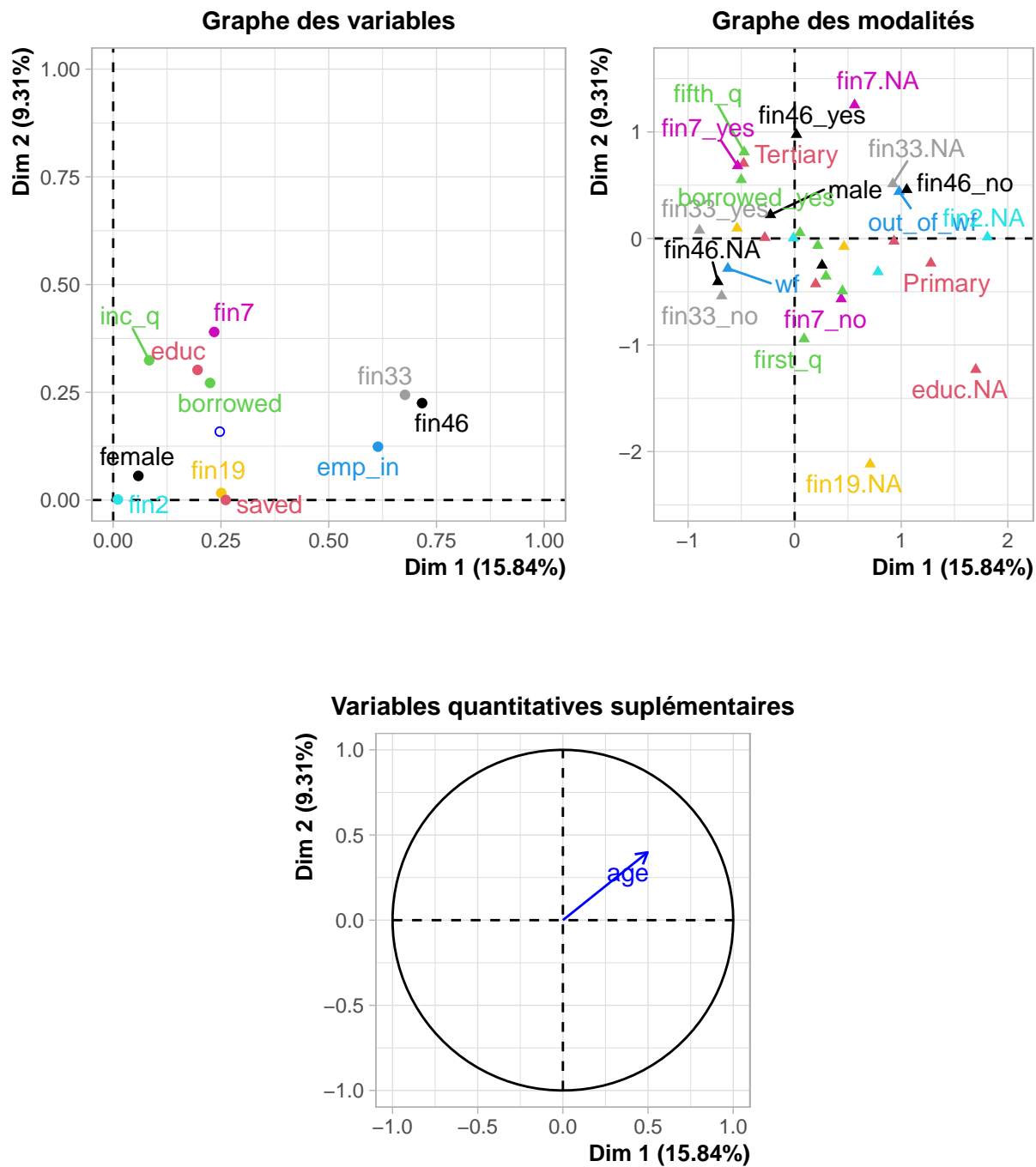
a- Choix du nombre d'axes factorielles



Commentaire On retient les 2 premiers axes factoriels car il n'y a plus de variation significative de la part d'inertie portée par les axes après le 2^e (règle du coude). Ce plan résume 25.1% de la variabilité totale de nos

données

b- Graphiques



Commentaire

- Le premier graphique présente les variables dans le plan formé par les deux premiers axes factoriels. On peut relever notamment que les variables *fin33* et *fin46* sont très proches. Les variables *borrowed*, *educ*, *fin7*, *inc_q* sont également très proches; de même pour *fin19* et *saved*. Plus une variable est proche de l'origine, moins il est lié aux axes du plan.
- Le deuxième graphique fait la même analyse mais cette fois-ci par rapport aux modalités des variables.
- Le troisième représente la variable quantitative *age* dans la plan. La longueur de la flèche et son sens (pointe dans la même direction que nos deux axes factoriels) montre que cette variable en plus d'être fortement corrélée aux axes contribue fortement à la variation des ces derniers.

Table 2: Eta 2- Lien entre les axes et les variables

	Lien avec l'axe 1	Lien avec l'axe 2
female	0.0583	0.0560
educ	0.1957	0.3019
inc_q	0.0834	0.3244
emp_in	0.6143	0.1238
fin2	0.0107	0.0012
fin7	0.2343	0.3900
fin19	0.2507	0.0163
fin33	0.6769	0.2441
fin46	0.7166	0.2249
saved	0.2611	0.0002
borrowed	0.2247	0.2716

Lien avec les axes et choix des variables

Le tableau ci-dessus présente le lien entre nos variables et les deux axes factoriels retenus. Nous avons choisi comme mesure le Eta2, et retiendrons comme critère de sélection un Eta2 minimum de 0.1 sur chacun des deux axes. Ainsi, les variables qui satisfont cette condition sont: *fin7, educ, borrowed, fin33, fin46, emp_in*. Cependant, les variables *fin33* et *fin46* ont des corrélations (mesurées par la matrice de Cramer) quasi nulles, respectivement 0.06 et 0.08 avec la variable explicative *fin19*, on ne les utilisera pas.

2- Modèle à probabilité linéaire:

Table 3: Modèle à probabilité linéaire

	Dependent variable:
	fin19
age	0.004*** (0.001)
as.factor(educ)2	0.191*** (0.074)
as.factor(educ)3	0.313*** (0.077)
as.factor(fin7)1	0.086** (0.040)
as.factor(emp_in)1	0.360*** (0.039)
as.factor(borrowed)1	0.030 (0.038)
Constant	-0.267*** (0.100)
Observations	990
R ²	0.153
Adjusted R ²	0.148
Residual Std. Error	0.461 (df = 983)
F Statistic	29.622*** (df = 6; 983)

Note: * p<0.1; ** p<0.05; *** p<0.01

- Test de Breusch-Pagan

Table 4: Résultats du test de Pagan

	Stat	p.value
BP	28.01	0.00009

Commentaire

La p-value du test de Breusch-Pagan est inférieure à 5% nous pouvons donc rejeter l'hypothèse nulle selon laquelle la variance des résidus est constante et en déduire que l'hétéroscédasticité est effectivement présente. On va donc la corriger.

- Modèle avec correction d'hétéroscédasticité

Après correction de l'hétéroscédasticité, les coefficients associés à nos variables ainsi que la valeur des écarts-types liées à ces derniers ont variés. La part expliquée de la variabilité de l'accès au crédit (*fin19*) par notre modèle a augmenté, on est passé de 15,3% à 75,6% . Etre plus âgé a un effet positif sur les chances d'obtenir un crédit, être un salarié, avoir un crédit ont aussi un effet positif. Avoir fait des études supérieures augmente la probabilité d'avoir crédit par rapport au fait d'avoir complété les études primaires. Tandis que par rapport à la même situation, le gain d'avoir fait des études secondaires semble presque nul. Avoir emprunté de l'argent l'année d'avant l'étude ne semble pas avoir un effet statistiquement significatif sur l'accès au crédit (voir table ci-dessous).

Table 5: Modèles à probabilité linéaire

	<i>Dependent variable:</i>	
	fin19	
	Modèle	Modèle corrigé
	(1)	(2)
age	0.004*** (0.001)	0.001*** (0.0003)
as.factor(educ)2	0.191*** (0.074)	0.053*** (0.014)
as.factor(educ)3	0.313*** (0.077)	0.269*** (0.026)
as.factor(fin7)1	0.086** (0.040)	0.206*** (0.025)
as.factor(emp_in)1	0.360*** (0.039)	0.466*** (0.020)
as.factor(borrowed)1	0.030 (0.038)	-0.013 (0.014)
Constant	-0.267*** (0.100)	-0.078*** (0.019)
Observations	990	990
R ²	0.153	0.756
Adjusted R ²	0.148	0.754
Residual Std. Error (df = 983)	0.461	0.902
F Statistic (df = 6; 983)	29.622***	506.908***
<i>Note:</i>	* p<0.1; ** p<0.05; *** p<0.01	

3- Modèle logit

Table 6: Modèle Logit

	<i>Dependent variable:</i>
	fin19
age	0.021*** (0.005)
as.factor(educ)2	1.345*** (0.503)
as.factor(educ)3	1.894*** (0.512)
as.factor(fin7)1	0.365** (0.184)
as.factor(emp_in)1	1.671*** (0.193)
as.factor(borrowed)1	0.152 (0.178)
Constant	-4.076*** (0.627)
Observations	990
Log Likelihood	-601.631
Akaike Inf. Crit.	1,217.262
<i>Note:</i>	* p<0.1; ** p<0.05; *** p<0.01

Commentaire

Toutes nos variables sont significatifs à l'exception de la variable *borrowed*.

- par rapport au fait d'avoir un niveau d'étude primaire, avoir complété des études secondaire et supérieures augmente les chances d'avoir accès au crédit.
- Etre plus âgé augmente la probabilité d'avoir un crédit.
- Posséder une carte de crédit, être salarié augmente les chances d'avoir accès au crédit.

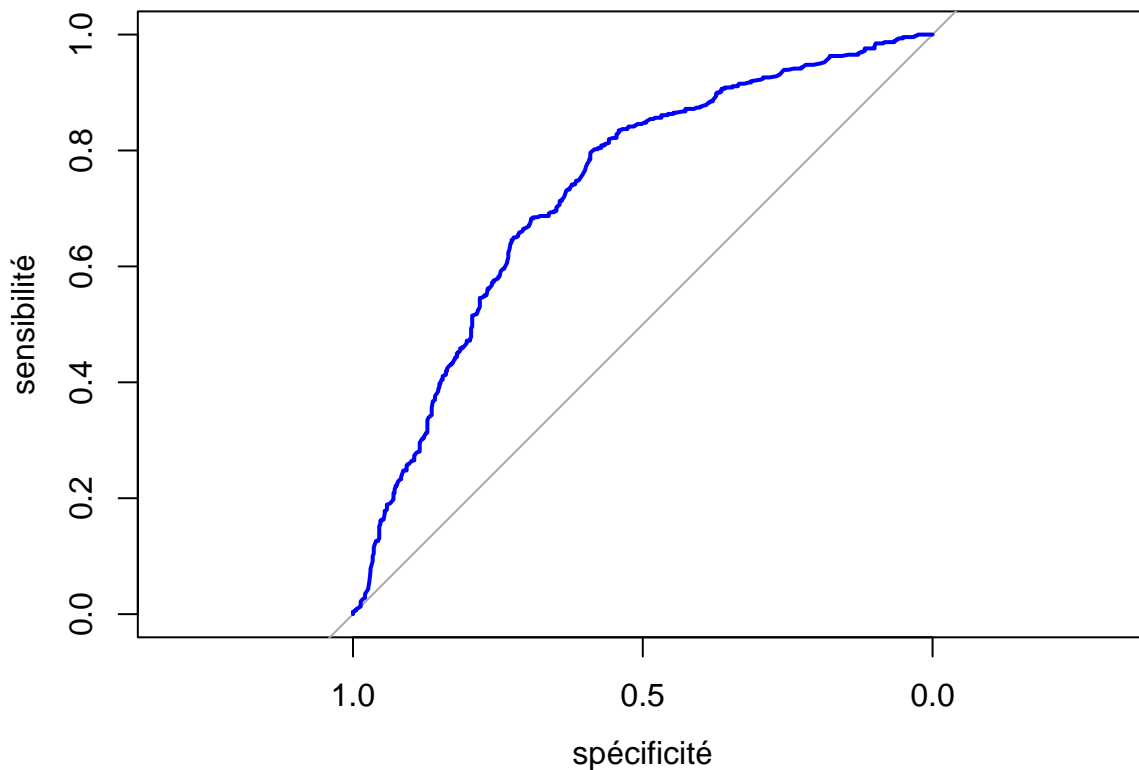
Table 7: Rapports de chance

	Odd_ratio
age	1.021351
educ_2	3.839423
educ_3	6.644675
fin7_1	1.440180
emp_in_1	5.316476
borrowed_1	1.164736

Rapports de chance

- La probabilité d'obtenir un prêt augmente de 2,13% avec l'âge.
- Les personnes qui ont fait des études secondaires et supérieures ont respectivement 3 et 6 fois plus de chance d'accéder au crédit que ceux qui n'ont fait que des études primaires.
- Avoir une carte de crédit, être salarié augmentent les chances d'avoir un crédit respectivement de 44,01% et 5,31.
- Avoir contracté un prêt l'an dernier augmente de 1,16 fois les chances d'accès au crédit à nouveau.

- Courbe ROC



L'aire sous la courbe ROC est de 0.7298195. La capacité de discrimination du modèle est acceptable.

4- Modèle Probit

Table 8:

	Dependent variable:			
	fin19			
	Linéaire	OLS Lin corr	logistic Logit	probit probit
	(1)	(2)	(3)	(4)
age	0.004*** (0.001)	0.001*** (0.0003)	0.021*** (0.005)	0.012*** (0.003)
as.factor(educ)2	0.191*** (0.074)	0.053*** (0.014)	1.345*** (0.503)	0.753*** (0.269)
as.factor(educ)3	0.313*** (0.077)	0.269*** (0.026)	1.894*** (0.512)	1.086*** (0.275)
as.factor(fin7)1	0.086** (0.040)	0.206*** (0.025)	0.365** (0.184)	0.230** (0.112)
as.factor(emp_in)1	0.360*** (0.039)	0.466*** (0.020)	1.671*** (0.193)	1.004*** (0.115)
as.factor(borrowed)1	0.030 (0.038)	-0.013 (0.014)	0.152 (0.178)	0.089 (0.109)
Constant	-0.267*** (0.100)	-0.078*** (0.019)	-4.076*** (0.627)	-2.387*** (0.346)
Observations	990	990	990	990
R ²	0.153	0.756		
Adjusted R ²	0.148	0.754		
Log Likelihood			-601.631	-602.064
Akaike Inf. Crit.			1,217.262	1,218.128
Residual Std. Error (df = 983)	0.461	0.902		
F Statistic (df = 6; 983)	29.622***	506.908***		

Note:

*p<0.1; **p<0.05; ***p<0.01

Commentaire

- les signes et la significativité de nos paramètres sont les mêmes pour toutes les régressions.
- Comme dit plus haut, le modèle linéaire avec correction d'hétéroscédasticité (*Lincorr*) explique plus la variabilité de *fin19* que le modèle à probabilité linéaire simple. Mais ces deux modèles sont inefficaces car elles prédisent des probabilités supérieurs à 1 ou inférieurs à 0 suivant les caractéristiques de certains individus.
- Les modèles *Logit* et *Probit* sont assez proches, avec approximativement les mêmes AIC et Log likelihood. La différence est au niveau de la valeur estimée des paramètres due au méthodes de calculs qui sont différentes.

Conclusion

De notre étude, il ressort que pour évaluer les chances d'accès au crédit il faut estimer un *Modèle à probabilité non linéaire* en tenant en compte notamment de l'âge, le niveau d'éducation, la situation professionnelle, et le fait que la personne détienne ou pas une carte de crédit. Il pourrait être intéressant de regarder en aval la situation bancaire (crédits contractés les années précédentes).