

Fiche de modélisations n°1

Sélection automatique de modèles prédictifs sur la pauvreté subjective

Kim Antunez

1 Objectif

L'objectif de cette première série de modèles est de détecter les variables de l'enquête du Baromètre d'opinion de la DREES qui permettent le mieux de prédire la **pauvreté subjective**.

Cette partie n'a **aucune visée explicative**. L'objectif est ici **purement prédictif**, dans un but de statistiques descriptives / défrichage de terrain, pour permettre de déceler une première série de variables particulièrement liées à la pauvreté subjective. Les modèles qui suivront permettront certainement d'aller beaucoup plus en profondeur dans l'analyse des véritables **déterminants de la pauvreté subjective**.

2 Analyses

Le principal facteur qui explique la pauvreté subjective est lui-même subjectif, il s'agit de trouver que sa situation actuelle est assez mauvaise. Vient ensuite le facteur institutionnel de disposer d'une allocation logement puis la pauvreté monétaire. Puis, le fait de ne pas vivre en couple (situation d'isolement social), d'être locataire (facteur important de précarité sociale), d'être au RSA (autre facteur institutionnel en plus des APL). Ce n'est qu'ensuite que les facteurs de précarité sur le marché du travail entrent en compte : celui de ne pas exercer d'activité professionnelle, d'occuper certaines PCS particulières (ouvrier) ou qu'elles soient occupées par la personne référente du ménage.

Dans une moindre mesure, d'autres facteurs subjectifs entrent ensuite en compte et sont des indices d'une forte insécurité sociale : une forte préoccupation concernant la thématique du logement, de la pauvreté, du pouvoir d'achat et du niveau des retraites (mais pas des crises financières et de la dette, sujets de préoccupations attribués aux riches), le fait de ne pas être couvert par une mutuelle (indicateur cette fois-ci objectif) ou de ne pas se sentir prêts à épargner en vu d'une situation de dépendance ou de la retraite (pour soi ou ses proches). Les personnes qui se sentent pauvres adhèrent aussi plus généralement au stéréotype de genre de la femme au foyer, votent plus souvent à l'extrême-gauche, se sentent insatisfaits du système de protection sociale, en particulier le système de soin, et ont plus souvent un entourage socialement précaire (SMIC, SDF...). Bien sûr, ils reçoivent également moins souvent des revenus complémentaires à leurs salaires (patrimoine financier et immobilier).

Résultats détaillés :

- Quand $\lambda = 0,1$, deux modalités permettent à elles seules de prédire le sentiment de pauvreté avec 85 % de chance d’avoir bon (*accuracy*). Il s’agit de la modalité “assez mauvaise” de la question “OG1” qui interroge les Français sur le fait de qualifier leur situation actuelle (+¹), ainsi que le fait d’avoir perçu une allocation logement au cours des 12 derniers mois au sein de son ménage (+).
- En pénalisant moins le modèle $\lambda = 0,05$, viennent s’ajouter des variables de revenus (-) et de niveau de vie (-), en d’autres termes, la pauvreté monétaire telle que mesurée dans le Baromètre.
- Quand $\lambda = 0,04$, s’ajoute le fait de vivre en couple (-), d’être locataire (+), d’être au RSA (+).
- Quand $\lambda = 0,03$, s’ajoute le fait de ne pas exercer d’activité professionnelle (+), le pessimisme envers l’avenir (+).
- Quand $\lambda = 0,02$, s’ajoute l’année 2019 (+), le fait que la personne de référence soit sans emploi ou ouvrier (+), le fait d’être préoccupé par la thématique du logement (+) ou le fait de juger que le niveau général des retraites est insuffisant (+), le fait d’avoir un diplôme du supérieur (-).
- Puis, quand $\lambda = 0,01$, 60 variables sont retenues et l’*accuracy* vaut 86 %, soit seulement un point de plus qu’avec le premier modèle à 2 variables ! S’ajoute l’année 2017 (+ mais faible), de nombreuses autres PCS (que ce soit de l’individu ou de la personne de référence), de nombreux autres sujets de préoccupations sociales : logement, pauvreté, pouvoir d’achat (+) contrairement à dette et crises financières (-). D’autres facteurs entrent également en compte : l’adhésion au stéréotype de genre de la femme au foyer (+), la prise en charge des parents en situation de dépendance en les accueillant chez eux (+) mais sans être prêts à épargner pour soi ni pour les autres en prévision de la vieillesse (retraite ou dépendance), le mauvais état de santé et l’insatisfaction envers le système de soin et le niveau de protection sociale (+) ne souhaitant pas limiter les remboursements de la sécurité sociale (+), avoir un entourage qui lui-même connaît des situations de précarités (SMIC, SDF, handicap, +), recevoir des revenus d’actifs financiers ou de locations (-), et voter extrême-gauche (+).

3 Code et résultats

```
#chargement des packages
library(dplyr) #manipuler les bases de données
library(glmnet) #sélection automatique de modèles (lasso, ridge, elasticnet)

# alea fixé
set.seed(1)
```

Création de la base de données permettant de réaliser les différents modèles économétriques. Différents traitements sont effectués :

1. On crée des indicatrices en remplacement des variables catégorielles (par défaut on supprime la première modalité de la variable pour que les modèles économétriques soient identifiables). On renomme ces variables selon la nomenclature suivante : `LASSO_nomvar_nummodalite`.
2. On crée la variable `y_LASSO` que l’on cherche à prédire. C’est une indicatrice qui vaut 0 quand l’individu répond “Je me considère déjà comme pauvre” à la question “PE3. Vous personnellement, pensez-vous qu’il y a un risque que vous deveniez pauvre dans les cinq prochaines années ?”.

1. Un signe + indique que la variable influe positivement sur le sentiment de pauvreté, inversement pour un signe -.

3. On supprime les colonnes qui contiennent des valeurs manquantes (qui ne sont pas disponibles pour tous les millésimes ou pour tous les enquêtés) : `sdstatemp`, `sdnbenf`, `og2`²...

```
bdd_lasso <- bdd
for(col in setdiff(colnames(bdd_lasso),c("ident"))){
  if(is.character(bdd_lasso[[col]])){
    modalites <- unique(bdd_lasso[[col]])[-1]
    for(mod in modalites){
      nom_var = paste0("LASSO_",col,"-",mod)
      bdd_lasso[[nom_var]] = 1*c(bdd_lasso[[col]]==mod)
    }
  } else{
    nom_var = paste0("LASSO_",col)
    bdd_lasso[[nom_var]] = bdd_lasso[[col]]
  }
}
bdd_lasso <- bdd_lasso %>%
  mutate( LASSO_y = factor(ifelse(
    subj_risque_pauvrete=="Je me considère déjà\ncomme pauvre",1,0))
  ) %>%
  select(-LASSO_subj_risque_pauvrete) %>%
  select(starts_with("LASSO"))

#On supprime les colonnes qui contiennent des valeurs manquantes
for(col in colnames(bdd_lasso)){
  if(any(is.na(bdd_lasso[[col]]))){
    bdd_lasso[[col]] <- NULL
  }
}
```

On sélectionne aléatoirement :

1. Un échantillon d'entraînement pour entraîner le modèle [70 % des données]
2. Un échantillon test pour calculer les performances du modèles en l'occurrence l'*accuracy* (le pourcentage de bien prédits) [30 % des données]

```
select_train = sort(sample(nrow(bdd_lasso), nrow(bdd_lasso)*.7))
train<-bdd_lasso[select_train,]
test<-bdd_lasso[-select_train,]
```

On essaie différents modèles de type Lasso avec plusieurs niveau de pénalisation λ . Le principe est que plus λ est élevé, plus on pénalise le fait de choisir un grand nombre de variables et moins le nombre de variables sélectionnées est important. Inversement, un modèle avec un faible λ est peu sélectif.

On affiche pour chaque valeur de λ le nombre de variables retenues³ (dont le coefficient de la régression est non nul) et on les liste.

2. Une autre option aurait été de les recoder en une modalité à part, nous avons choisi l'option de les supprimer par simplicité.

3. modalités en cas de variables catégorielles et nom de la variable en cas de variables continues

```

# Note alpha=1 for lasso only and alpha=0 ridge only.
for(lambda in c(0.9,0.5,0.1,0.05,0.04,0.03,0.02, 0.01)){
  cat(paste0("Valeur de la pénalisation : ", lambda,"\n"))
  lasso.model <- glmnet(x=as.matrix(train[-ncol(train)]),
                        y=train$LASSO_y, alpha=1,family="binomial",
                        lambda = lambda) #cv.lasso$lambda.min

  myCoefs <- coef(lasso.model, s=0)
  myCoefs_nonnuls <- myCoefs@Dimnames[[1]][which(myCoefs != 0 ) ][-1]
  cat(paste0("Nombre de variables retenues : ",
            length(myCoefs_nonnuls),"\n"))
  if(length(myCoefs_nonnuls)!=0){
    cat(paste0(
      do.call(c,lapply(myCoefs_nonnuls,
                      function(x){
                        paste0("- ", couper_string(colnames_to_nomexplicite(x),90))
                      })),collapse = "\n"
    ))
  }

  probabilities <- lasso.model %>%
    predict(newx = as.matrix(test[, -ncol(test)]))
  cat(paste0("\n>>> Pourcentages de bien prédits : ",
            round(100*mean(
              ifelse(probabilities > 0.5, 1, 0) == test$LASSO_y
            )," %"))
  }
  cat("\n\n")
}

```

Valeur de la pénalisation : 0.9
 Nombre de variables retenues : 0

Valeur de la pénalisation : 0.5
 Nombre de variables retenues : 0

Valeur de la pénalisation : 0.1
 Nombre de variables retenues : 2
 - OG1. Vous personnellement, comment qualifieriez-vous votre situation actuelle ? Assez mauvaise
 - SDRES_9. Au cours des douze derniers mois, votre ménage a-t-il perçu des ressources provenant des allocations de logement ? Oui
 >>> Pourcentages de bien prédits : 85 %

Valeur de la pénalisation : 0.05
 Nombre de variables retenues : 5
 - OG1. Vous personnellement, comment qualifieriez-vous votre situation actuelle ? Assez mauvaise
 - OG1. Vous personnellement, comment qualifieriez-vous votre situation actuelle ? Très

mauvaise

- SDRES_9. Au cours des douze derniers mois, votre ménage a-t-il perçu des ressources provenant des allocations de logement ? Oui
 - SDREVTR. Revenus mensuels nets avant impôts du foyer (en tranches) De 1000 à moins de 1400 euros par mois
 - SDNIVIE. Quintiles de niveau de vie mensuel imputé 1er quintile
- >>> Pourcentages de bien prédits : 85 %

Valeur de la pénalisation : 0.04

Nombre de variables retenues : 10

- SDCOUPLE. Vivez-vous en couple ? Oui
 - OG1. Vous personnellement, comment qualifieriez-vous votre situation actuelle ? Assez mauvaise
 - OG1. Vous personnellement, comment qualifieriez-vous votre situation actuelle ? Très mauvaise
 - LO1. Concernant votre résidence principale, êtes-vous... Locataire d'un logement social (y compris HLM)
 - SDPROXIM_7. Dans votre famille ou en dehors de votre famille, connaissez-vous quelqu'un au RSA ? Oui, vous-même
 - SDRES_3. Au cours des douze derniers mois, votre ménage a-t-il perçu des ressources provenant du RSA (Revenu de solidarité active) ? Oui
 - SDRES_9. Au cours des douze derniers mois, votre ménage a-t-il perçu des ressources provenant des allocations de logement ? Oui
 - SDREVTR. Revenus mensuels nets avant impôts du foyer (en tranches) De 1000 à moins de 1400 euros par mois
 - SDREVTR. Revenus mensuels nets avant impôts du foyer (en tranches) Moins de 1000 euros par mois
 - SDNIVIE. Quintiles de niveau de vie mensuel imputé 1er quintile
- >>> Pourcentages de bien prédits : 85 %

Valeur de la pénalisation : 0.03

Nombre de variables retenues : 14

- SDPRSITUA. Situation de la personne de référence Il.elle n'exerce aucune activité professionnelle
- SDCOUPLE. Vivez-vous en couple ? Oui
- OG1. Vous personnellement, comment qualifieriez-vous votre situation actuelle ? Assez mauvaise
- OG1. Vous personnellement, comment qualifieriez-vous votre situation actuelle ? Très mauvaise
- OG3_1. Quand vous pensez à l'avenir, êtes-vous plutôt optimiste ou plutôt pessimiste... pour vous-même ? Plutôt optimiste
- OG3_1. Quand vous pensez à l'avenir, êtes-vous plutôt optimiste ou plutôt pessimiste... pour vous-même ? Très pessimiste
- LO1. Concernant votre résidence principale, êtes-vous... Locataire d'un logement social (y compris HLM)
- SDPROXIM_7. Dans votre famille ou en dehors de votre famille, connaissez-vous quelqu'un au RSA ? Oui, vous-même
- SDRES_3. Au cours des douze derniers mois, votre ménage a-t-il perçu des ressources provenant du RSA (Revenu de solidarité active) ? Oui

- SDRES_9. Au cours des douze derniers mois, votre ménage a-t-il perçu des ressources provenant des allocations de logement ? Oui
 - SDREVTR. Revenus mensuels nets avant impôts du foyer (en tranches) De 1000 à moins de 1400 euros par mois
 - SDREVTR. Revenus mensuels nets avant impôts du foyer (en tranches) Moins de 1000 euros par mois
 - SDNIVIE. Quintiles de niveau de vie mensuel imputé 2e quintile
 - SDNIVIE. Quintiles de niveau de vie mensuel imputé 1er quintile
- >>> Pourcentages de bien prédits : 85 %

Valeur de la pénalisation : 0.02

Nombre de variables retenues : 21

- ANNEE. Année 2019
 - SDACT. Activité principale de l'interviewé.e (ou ancienne activité) Autre inactif.ive
 - SDPRSITUA. Situation de la personne de référence Il.elle n'exerce aucune activité professionnelle
 - SDPRPCS7. Profession de la personne de référence (ou ancienne profession) en 7 catégories Ouvrier.ère
 - SDCOUPLE. Vivez-vous en couple ? Oui
 - SDMATRI. Statut marital Célibataire
 - OG1. Vous personnellement, comment qualifieriez-vous votre situation actuelle ? Assez mauvaise
 - OG1. Vous personnellement, comment qualifieriez-vous votre situation actuelle ? Très mauvaise
 - OG3_1. Quand vous pensez à l'avenir, êtes-vous plutôt optimiste ou plutôt pessimiste... pour vous-même ? Plutôt optimiste
 - OG3_1. Quand vous pensez à l'avenir, êtes-vous plutôt optimiste ou plutôt pessimiste... pour vous-même ? Très pessimiste
 - OG4_1. Êtes-vous personnellement préoccupé(e) par... le logement ? Beaucoup
 - LO1. Concernant votre résidence principale, êtes-vous... Locataire d'un logement social (y compris HLM)
 - RE1. À l'heure actuelle, comment évaluez-vous le niveau de vie moyen des retraités par rapport à celui de l'ensemble de la population ? Bien moins bon
 - SDPROXIM_7. Dans votre famille ou en dehors de votre famille, connaissez-vous quelqu'un au RSA ? Oui, vous-même
 - SDRES_3. Au cours des douze derniers mois, votre ménage a-t-il perçu des ressources provenant du RSA (Revenu de solidarité active) ? Oui
 - SDRES_9. Au cours des douze derniers mois, votre ménage a-t-il perçu des ressources provenant des allocations de logement ? Oui
 - SDREVTR. Revenus mensuels nets avant impôts du foyer (en tranches) De 1000 à moins de 1400 euros par mois
 - SDREVTR. Revenus mensuels nets avant impôts du foyer (en tranches) Moins de 1000 euros par mois
 - SDDIPL. Niveau de diplôme Vous avez un diplôme supérieur (2ème, 3ème cycle, grande école, L3, M1, M2)
 - SDNIVIE. Quintiles de niveau de vie mensuel imputé 2e quintile
 - SDNIVIE. Quintiles de niveau de vie mensuel imputé 1er quintile
- >>> Pourcentages de bien prédits : 86 %

Valeur de la pénalisation : 0.01

Nombre de variables retenues : 66

- ANNEE. Année 2017
- ANNEE. Année 2018
- ANNEE. Année 2019
- SDACT. Activité principale de l'interviewé.e (ou ancienne activité) Autre inactif.ive
- SDPCS7. Profession de l'interviewé.e (ou ancienne profession) en 7 catégories Profession libérale, cadre supérieur.e
- SDPRSITUA. Situation de la personne de référence Il.elle n'exerce aucune activité professionnelle
- SDPRSITUA. Situation de la personne de référence Il.elle travaille de façon intermittente
- SDPRPCS7. Profession de la personne de référence (ou ancienne profession) en 7 catégories Profession libérale, cadre supérieur.e
- SDPRPCS7. Profession de la personne de référence (ou ancienne profession) en 7 catégories Ouvrier.ère
- SDPRPCS7. Profession de la personne de référence (ou ancienne profession) en 7 catégories Profession intermédiaire
- SDPRPCS10. Profession actuelle de la personne de référence en 10 catégories Employé.e
- SDPRPCS10. Profession actuelle de la personne de référence en 10 catégories Artisan.e ou commerçant.e
- SDCOUPLE. Vivez-vous en couple ? Oui
- SDMATRI. Statut marital Célibataire
- SDMUT. Êtes-vous couvert.e par une mutuelle ou une assurance complémentaire maladie ? Non
- OG1. Vous personnellement, comment qualifieriez-vous votre situation actuelle ? Assez mauvaise
- OG1. Vous personnellement, comment qualifieriez-vous votre situation actuelle ? Très mauvaise
- OG3_1. Quand vous pensez à l'avenir, êtes-vous plutôt optimiste ou plutôt pessimiste... pour vous-même ? Plutôt optimiste
- OG3_1. Quand vous pensez à l'avenir, êtes-vous plutôt optimiste ou plutôt pessimiste... pour vous-même ? Très pessimiste
- OG4_1. Êtes-vous personnellement préoccupé(e) par... le logement ? Beaucoup
- OG4_3. Êtes-vous personnellement préoccupé(e) par... la pauvreté ? Beaucoup
- OG4_3. Êtes-vous personnellement préoccupé(e) par... la pauvreté ? Peu
- OG4_5. Êtes-vous personnellement préoccupé(e) par... la dette de la France ? Beaucoup
- OG4_5. Êtes-vous personnellement préoccupé(e) par... la dette de la France ? Pas du tout
- OG4_7. Êtes-vous personnellement préoccupé(e) par... les crises financières internationales ? Pas du tout
- OG4_9. Êtes-vous personnellement préoccupé(e) par... le niveau des salaires et du pouvoir d'achat ? Beaucoup
- OG8_9. « Dans l'idéal, les femmes devraient rester à la maison pour élever leurs enfants » Plutôt d'accord
- LO1. Concernant votre résidence principale, êtes-vous... Locataire hors logement social, c'est-à-dire dans le parc privé
- LO1. Concernant votre résidence principale, êtes-vous... Locataire d'un logement social (y compris HLM)
- RE1. À l'heure actuelle, comment évaluez-vous le niveau de vie moyen des retraités par rapport à celui de l'ensemble de la population ? Bien moins bon
- RE10. À l'avenir, que faudrait-il faire en priorité en matière de retraite ? Il faut

- ajouter au système actuel un complément d'assurance ou d'épargne individuelle
- DE4. Seriez-vous prêt.e à épargner davantage en prévision d'une éventuelle situation de dépendance vous concernant ? Plutôt prêt.e
 - DE4. Seriez-vous prêt.e à épargner davantage en prévision d'une éventuelle situation de dépendance vous concernant ? Tout à fait prêt.e
 - DE5. Si l'un de vos parents proches devenait dépendant, que feriez-vous ? Vous l'accueilleriez chez vous
 - DE5. Si l'un de vos parents proches devenait dépendant, que feriez-vous ? Vous consacreriez une partie de votre revenu à lui payer des aides de manière à ce qu'il reste à son domicile
 - PS15_1. « Notre système de sécurité sociale... fournit un niveau de protection suffisant » Pas d'accord du tout
 - SA1. Comment est votre état de santé général ? Très bon
 - SA1. Comment est votre état de santé général ? Moyen
 - SA5. Selon vous, au cours des dernières années, l'état de santé des Français s'est... (NSP)
 - SA7_3. En général, êtes-vous satisfait de la qualité des soins offerts par les ... cliniques privées ? Plutôt satisfait.e
 - SA7_3. En général, êtes-vous satisfait de la qualité des soins offerts par les ... cliniques privées ? Ni satisfait.e ni insatisfait.e
 - SA7_3. En général, êtes-vous satisfait de la qualité des soins offerts par les ... cliniques privées ? (Je ne suis jamais allé.e/ je ne connais pas)
 - SA7_7. En général, êtes-vous satisfait de la qualité des soins offerts par les ... infirmiers/infirmières (hors hôpital) ? (NSP)
 - SA8_2. Pour réduire le déficit de la branche maladie de la Sécurité sociale, seriez-vous plutôt favorable ou opposé.e à... limiter le remboursement de la Sécurité sociale pour certaines prestations ? Plutôt opposé.e
 - SA22_4. Avez-vous, au cours de ces dernières années, fait appel à un dentiste / orthodontiste pour vous ou pour vos proches ? Non
 - CS19_2. « Beaucoup de personnes perçoivent des allocations alors qu'elles n'y ont pas droit » Pas du tout d'accord
 - SDPROXIM_3. Dans votre famille ou en dehors de votre famille, connaissez-vous quelqu'un qui est SDF ? Oui, hors famille
 - SDPROXIM_4. Dans votre famille ou en dehors de votre famille, connaissez-vous quelqu'un élevant seul.e ses enfants avec un revenu inférieur au SMIC ? Oui, vous-même et hors famille
 - SDPROXIM_7. Dans votre famille ou en dehors de votre famille, connaissez-vous quelqu'un au RSA ? Oui, vous-même
 - SDPROXIM_7. Dans votre famille ou en dehors de votre famille, connaissez-vous quelqu'un au RSA ? Oui, vous-même, dans votre famille et hors famille
 - SDPROXIM_7. Dans votre famille ou en dehors de votre famille, connaissez-vous quelqu'un au RSA ? Oui, vous-même et hors famille
 - SDPROXIM_8. Dans votre famille ou en dehors de votre famille, connaissez-vous une personne handicapée de moins de 60 ans ? Oui, vous-même, dans votre famille et hors famille
 - SDPROXIM_9. Dans votre famille ou en dehors de votre famille, connaissez-vous une personne âgée dépendante ? Oui, vous-même
 - SDRES_3. Au cours des douze derniers mois, votre ménage a-t-il perçu des ressources provenant du RSA (Revenu de solidarité active) ? Oui
 - SDRES_6. Au cours des douze derniers mois, votre ménage a-t-il perçu des ressources provenant de revenus d'actifs financiers ? Oui

- SDRES_7. Au cours des douze derniers mois, votre ménage a-t-il perçu des ressources provenant de revenus de locations ? Oui
 - SDRES_9. Au cours des douze derniers mois, votre ménage a-t-il perçu des ressources provenant des allocations de logement ? Oui
 - SDREVTR. Revenus mensuels nets avant impôts du foyer (en tranches) De 1000 à moins de 1400 euros par mois
 - SDREVTR. Revenus mensuels nets avant impôts du foyer (en tranches) Moins de 1000 euros par mois
 - SDDIPL. Niveau de diplôme Vous avez un certificat d'aptitude professionnelle (CAP), un brevet d'enseignement professionnel (BEP)
 - SDDIPL. Niveau de diplôme Vous n'avez pas de diplôme
 - SDDIPL. Niveau de diplôme Vous avez un diplôme supérieur (2ème, 3ème cycle, grande école, L3, M1, M2)
 - SDPOLTR. Sur une échelle de 0 à 10, où vous classeriez-vous politiquement ? (en 5 tranches) 0-2 extrême gauche
 - SDNIVIE. Quintiles de niveau de vie mensuel imputé 2e quintile
 - SDNIVIE. Quintiles de niveau de vie mensuel imputé 4e quintile
 - SDNIVIE. Quintiles de niveau de vie mensuel imputé 1er quintile
- >>> Pourcentages de bien prédits : 86 %

Pour le dernier modèle (le moins pénalisé), on regarde la valeur des coefficients associés aux variables, en les triant par ordre d'importance en valeur absolue.

```
myResults <- data.frame(
  features = myCoefs@Dimnames[[1]][ which(myCoefs != 0 ) ],
  coefs    = round(myCoefs          [ which(myCoefs != 0 ) ],2)
) %>%
  mutate(features_nom =
lapply(features, colnames_to_nomexplicite)) %>%
  mutate(absv=abs(coefs)) %>%
  arrange(desc(absv)) %>%
  select(-absv)

myResults %>%
  select(features,features_nom,coefs) %>%
  kable(format = 'latex', booktabs = TRUE, longtable = TRUE) %>%
  kableExtra::column_spec(1, width = "4cm") %>%
  kableExtra::column_spec(2, width = "9cm") %>%
  kableExtra::column_spec(3, width = "2cm") %>%
  kableExtra::kable_styling(latex_options = c("hold_position", "repeat_header"))
```

| features | features_nom | coefs |
|-------------|--|-------|
| (Intercept) | NULL | -2.84 |
| LASSO_og1-4 | OG1. Vous personnellement, comment qualifieriez-vous votre situation actuelle ? Très mauvaise | 1.60 |
| LASSO_og1-3 | OG1. Vous personnellement, comment qualifieriez-vous votre situation actuelle ? Assez mauvaise | 1.16 |

(continued)

| features | features_nom | coefs |
|--------------------|--|-------|
| LASSO_sdproxim_7-4 | SDPROXIM_7. Dans votre famille ou en dehors de votre famille, connaissez-vous quelqu'un au RSA ? Oui, vous-même | 0.69 |
| LASSO_sdnivie-1 | SDNIVIE. Quintiles de niveau de vie mensuel imputé 1er quintile | 0.64 |
| LASSO_sdproxim_4-7 | SDPROXIM_4. Dans votre famille ou en dehors de votre famille, connaissez-vous quelqu'un élevant seul.e ses enfants avec un revenu inférieur au SMIC ? Oui, vous-même et hors famille | 0.54 |
| LASSO_sdres_9-1 | SDRES_9. Au cours des douze derniers mois, votre ménage a-t-il perçu des ressources provenant des allocations de logement ? Oui | 0.52 |
| LASSO_sdnivie-2 | SDNIVIE. Quintiles de niveau de vie mensuel imputé 2e quintile | 0.52 |
| LASSO_sdproxim_7-8 | SDPROXIM_7. Dans votre famille ou en dehors de votre famille, connaissez-vous quelqu'un au RSA ? Oui, vous-même, dans votre famille et hors famille | 0.44 |
| LASSO_lo1-2 | LO1. Concernant votre résidence principale, êtes-vous... Locataire d'un logement social (y compris HLM) | 0.32 |
| LASSO_sdproxim_8-8 | SDPROXIM_8. Dans votre famille ou en dehors de votre famille, connaissez-vous une personne handicapée de moins de 60 ans ? Oui, vous-même, dans votre famille et hors famille | 0.30 |
| LASSO_sdcouple-1 | SDCOUPLE. Vivez-vous en couple ? Oui | -0.29 |
| LASSO_og3_1-2 | OG3_1. Quand vous pensez à l'avenir, êtes-vous plutôt optimiste ou plutôt pessimiste... pour vous-même ? Plutôt optimiste | -0.27 |
| LASSO_annee-2019 | ANNEE. Année 2019 | 0.26 |
| LASSO_lo1-3 | LO1. Concernant votre résidence principale, êtes-vous... Locataire hors logement social, c'est-à-dire dans le parc privé | 0.22 |
| LASSO_re1-5 | RE1. À l'heure actuelle, comment évaluez-vous le niveau de vie moyen des retraités par rapport à celui de l'ensemble de la population ? Bien moins bon | 0.21 |
| LASSO_sdrevtr-2 | SDREVTR. Revenus mensuels nets avant impôts du foyer (en tranches) De 1000 à moins de 1400 euros par mois | 0.21 |
| LASSO_sdprpcs10-2 | SDPRPCS10. Profession actuelle de la personne de référence en 10 catégories Artisan.e ou commerçant.e | -0.20 |
| LASSO_de4-1 | DE4. Seriez-vous prêt.e à épargner davantage en prévision d'une éventuelle situation de dépendance vous concernant ? Tout à fait prêt.e | -0.20 |
| LASSO_sdrevtr-1 | SDREVTR. Revenus mensuels nets avant impôts du foyer (en tranches) Moins de 1000 euros par mois | 0.20 |

(continued)

| features | features_nom | coefs |
|--------------------|--|-------|
| LASSO_sdprsitua-7 | SDPRSITUA. Situation de la personne de référence Il.elle n'exerce aucune activité professionnelle | 0.19 |
| LASSO_sdpcs7-3 | SDPCS7. Profession de l'interviewé.e (ou ancienne profession) en 7 catégories Profession libérale, cadre supérieur.e | -0.17 |
| LASSO_og4_1-1 | OG4_1. Êtes-vous personnellement préoccupé(e) par... le logement ? Beaucoup | 0.17 |
| LASSO_sdprpcs7-4 | SDPRPCS7. Profession de la personne de référence (ou ancienne profession) en 7 catégories Profession intermédiaire | -0.16 |
| LASSO_sdact-9 | SDACT. Activité principale de l'interviewé.e (ou ancienne activité) Autre inactif.ive | 0.15 |
| LASSO_sdproxim_7-7 | SDPROXIM_7. Dans votre famille ou en dehors de votre famille, connaissez-vous quelqu'un au RSA ? Oui, vous-même et hors famille | 0.15 |
| LASSO_de4-2 | DE4. Seriez-vous prêt.e à épargner davantage en prévision d'une éventuelle situation de dépendance vous concernant ? Plutôt prêt.e | -0.14 |
| LASSO_sdres_3-1 | SDRES_3. Au cours des douze derniers mois, votre ménage a-t-il perçu des ressources provenant du RSA (Revenu de solidarité active) ? Oui | 0.14 |
| LASSO_sdprpcs7-6 | SDPRPCS7. Profession de la personne de référence (ou ancienne profession) en 7 catégories Ouvrier.ère | 0.12 |
| LASSO_sdres_6-1 | SDRES_6. Au cours des douze derniers mois, votre ménage a-t-il perçu des ressources provenant de revenus d'actifs financiers ? Oui | -0.12 |
| LASSO_sa5-6 | SA5. Selon vous, au cours des dernières années, l'état de santé des Français s'est... (NSP) | 0.11 |
| LASSO_sddipl-8 | SDDIPL. Niveau de diplôme Vous avez un diplôme supérieur (2ème, 3ème cycle, grande école, L3, M1, M2) | -0.11 |
| LASSO_og4_5-4 | OG4_5. Êtes-vous personnellement préoccupé(e) par... la dette de la France ? Pas du tout | 0.09 |
| LASSO_ps15_1-4 | PS15_1. « Notre système de sécurité sociale... fournit un niveau de protection suffisant » Pas d'accord du tout | 0.09 |
| LASSO_sa7_7-6 | SA7_7. En général, êtes-vous satisfait de la qualité des soins offerts par les ... infirmiers/infirmières (hors hôpital) ? (NSP) | 0.09 |
| LASSO_cs19_2-4 | CS19_2. « Beaucoup de personnes perçoivent des allocations alors qu'elles n'y ont pas droit » Pas du tout d'accord | 0.09 |
| LASSO_sdmatri-3 | SDMATRI. Statut marital Célibataire | 0.08 |

(continued)

| features | features_nom | coefs |
|--------------------|---|-------|
| LASSO_sddipl-1 | SDDIPL. Niveau de diplôme Vous n'avez pas de diplôme | 0.07 |
| LASSO_sdnivie-4 | SDNIVIE. Quintiles de niveau de vie mensuel imputé 4e quintile | -0.07 |
| LASSO_sdprpcs10-5 | SDPRPCS10. Profession actuelle de la personne de référence en 10 catégories Employé.e | 0.06 |
| LASSO_og4_3-1 | OG4_3. Êtes-vous personnellement préoccupé(e) par... la pauvreté ? Beaucoup | 0.06 |
| LASSO_re10-2 | RE10. À l'avenir, que faudrait-il faire en priorité en matière de retraite ? Il faut ajouter au système actuel un complément d'assurance ou d'épargne individuelle | -0.06 |
| LASSO_de5-3 | DE5. Si l'un de vos parents proches devenait dépendant, que feriez-vous ? Vous consacreriez une partie de votre revenu à lui payer des aides de manière à ce qu'il reste à son domicile | -0.06 |
| LASSO_sa1-1 | SA1. Comment est votre état de santé général ? Très bon | -0.06 |
| LASSO_sdproxim_3-3 | SDPROXIM_3. Dans votre famille ou en dehors de votre famille, connaissez-vous quelqu'un qui est SDF ? Oui, hors famille | 0.06 |
| LASSO_sdmut-2 | SDMUT. Êtes-vous couvert.e par une mutuelle ou une assurance complémentaire maladie ? Non | 0.05 |
| LASSO_og4_9-1 | OG4_9. Êtes-vous personnellement préoccupé(e) par... le niveau des salaires et du pouvoir d'achat ? Beaucoup | 0.05 |
| LASSO_og8_9-2 | OG8_9. « Dans l'idéal, les femmes devraient rester à la maison pour élever leurs enfants » Plutôt d'accord | 0.05 |
| LASSO_sdres_7-1 | SDRES_7. Au cours des douze derniers mois, votre ménage a-t-il perçu des ressources provenant de revenus de locations ? Oui | -0.05 |
| LASSO_sddipl-4 | SDDIPL. Niveau de diplôme Vous avez un certificat d'aptitude professionnelle (CAP), un brevet d'enseignement professionnel (BEP) | 0.05 |
| LASSO_sdpoltr-1 | SDPOLTR. Sur une échelle de 0 à 10, où vous classeriez-vous politiquement ? (en 5 tranches) 0-2 extrême gauche | 0.04 |
| LASSO_og4_5-1 | OG4_5. Êtes-vous personnellement préoccupé(e) par... la dette de la France ? Beaucoup | -0.03 |
| LASSO_og4_7-4 | OG4_7. Êtes-vous personnellement préoccupé(e) par... les crises financières internationales ? Pas du tout | 0.03 |
| LASSO_sa1-3 | SA1. Comment est votre état de santé général ? Moyen | 0.03 |

(continued)

| features | features_nom | coefs |
|--------------------|---|-------|
| LASSO_sa7_3-7 | SA7_3. En général, êtes-vous satisfait de la qualité des soins offerts par les ... cliniques privées ? (Je ne suis jamais allé.e/ je ne connais pas) | 0.03 |
| LASSO_sdprpcs7-3 | SDPRPCS7. Profession de la personne de référence (ou ancienne profession) en 7 catégories Profession libérale, cadre supérieur.e | -0.02 |
| LASSO_og3_1-4 | OG3_1. Quand vous pensez à l'avenir, êtes-vous plutôt optimiste ou plutôt pessimiste... pour vous-même ? Très pessimiste | 0.02 |
| LASSO_og4_3-3 | OG4_3. Êtes-vous personnellement préoccupé(e) par... la pauvreté ? Peu | -0.02 |
| LASSO_sa7_3-2 | SA7_3. En général, êtes-vous satisfait de la qualité des soins offerts par les ... cliniques privées ? Plutôt satisfait.e | -0.02 |
| LASSO_sa7_3-3 | SA7_3. En général, êtes-vous satisfait de la qualité des soins offerts par les ... cliniques privées ? Ni satisfait.e ni insatisfait.e | 0.02 |
| LASSO_sa22_4-3 | SA22_4. Avez-vous, au cours de ces dernières années, fait appel à un dentiste / orthodontiste pour vous ou pour vos proches ? Non | 0.02 |
| LASSO_sdprsitua-3 | SDPRSITUA. Situation de la personne de référence Il.elle travaille de façon intermittente | 0.01 |
| LASSO_de5-2 | DE5. Si l'un de vos parents proches devenait dépendant, que feriez-vous ? Vous l'accueilleriez chez vous | 0.01 |
| LASSO_sa8_2-2 | SA8_2. Pour réduire le déficit de la branche maladie de la Sécurité sociale, seriez-vous plutôt favorable ou opposé.e à... limiter le remboursement de la Sécurité sociale pour certaines prestations ? Plutôt opposé.e | 0.01 |
| LASSO_sdproxim_9-4 | SDPROXIM_9. Dans votre famille ou en dehors de votre famille, connaissez-vous une personne âgée dépendante ? Oui, vous-même | 0.01 |
| LASSO_annee-2017 | ANNEE. Année 2017 | 0.00 |
| LASSO_annee-2018 | ANNEE. Année 2018 | 0.00 |

4 Notes méthodologiques

Pour ces modèles quatre vagues du Baromètre ont été empilées : 2016, 2017, 2018 et 2019 (12 114 observations).

Bibliographie

<https://www.math.univ-toulouse.fr/~besse/Wikistat/pdf/st-scenar-reg-penal-prostate.pdf>

<https://glmnet.stanford.edu/articles/glmnet.html>

<http://www.sthda.com/english/articles/36-classification-methods-essentials/149-penalized-logistic-regression-essentials-in-r-ridge-lasso-and-elastic-net/>

Machine Learning Essentials: Practical Guide in R