

Exercice du thème 6

Nous traitons un exemple de score de crédit. Nous cherchons à déterminer les facteurs explicatifs de l'accord ou du refus d'une demande de crédit de clients auprès d'un établissement bancaire.

Les données sont dans le fichier « AcceptationCredit.csv ».

Plusieurs variables ont été mesurées sur un échantillon de 100 clients:

Acceptation crédit	Variable binaire (1 si Oui, 0 sinon)
Âge	Âge du client
Revenu	Revenu par tête dans le ménage (milliers de \$ par mois)
Infraction bancaire	« Oui » : si au moins un problème avec l'établissement bancaire a été rapporté, « Non » sinon

1. Effectuez une régression logistique de la variable réponse « Acceptation crédit » sur toutes les variables explicatives décrites ci-dessus. Quelle est l'estimation de l'équation de la régression logistique obtenue? Interprétez les estimations obtenues en termes de « Odds ratio ».
2. En utilisant un niveau de 5%, est-ce que la régression est globalement significative?
3. En utilisant un niveau de 5%, identifiez les variables significatives dans le modèle.
4. Quelle proportion de la variation de la variable réponse est expliquée par le modèle?
5. Effectuez un test d'hypothèse au seuil de 5%, pour juger de la qualité d'ajustement du modèle.
6. Discutez le pouvoir discriminant du modèle estimé à l'aide de la table de classification.
7. Pour un individu âgé de 38 ans ayant un revenu par tête dans le ménage de 1700\$ et n'ayant jamais eu de problème avec l'établissement bancaire, la demande de crédit sera-t-elle acceptée selon le modèle étudié.