

6.1 On s'intéresse à la satisfaction de clients par rapport à un produit. Cette dernière est mesurée à l'aide d'une échelle de Likert, allant de très insatisfait (1) à très satisfait (5). Le jeu de données `multinom.sas7bdat` contient les variables suivantes :

- `y` : score de satisfaction
- `sexe` : sexe de l'individu, soit homme (0), soit femme (1)
- `educ` : niveau d'éducation le plus élevé complété, soit secondaire (`sec`), soit collégial (`cegep`), soit universitaire (`uni`)
- `revenu` : variable catégorielle indiquant le revenu, soit faible (1), moyen (2) ou élevé (3).
- `age` : âge de l'individu (en années).

Modélisez la satisfaction des clients, `y`, en fonction de l'âge, du niveau d'éducation, du sexe et du niveau de revenu.

- (a) Est-ce que le modèle de régression multinomiale ordinale à cote proportionnelles est une simplification adéquate du modèle de régression multinomiale logistique? Si oui, utilisez ce modèle pour la suite. Si non, ajustez le modèle de régression multinomiale logistique avec 1 comme catégorie de référence pour `y`, 1 pour `revenu` et `sec` pour `education` et utilisez ce dernier pour répondre aux autres questions.
- (b) Interprétez l'effet des variables `education` et `sexe` pour la catégorie 2.
- (c) Est-ce que le modèle avec une probabilité constante pour chaque item est adéquat lorsque comparé au modèle qui inclut toutes les covariables?
- (d) Est-ce que l'effet de la variable `age` est globalement significatif?
- (e) Fournissez un intervalle de confiance à niveau 95% pour l'effet de la variable `age` pour chacune des cote par rapport à très insatisfait (1). Que concluez-vous sur l'effet de `age` pour les réponses 2, ..., 5 par rapport à 1?
- (f) Écrivez l'équation de la cote ajustée pour satisfait (4) par rapport à très insatisfait (1).
- (g) Prédisez la probabilité qu'un homme de 30 ans qui a un diplôme collégial et qui fait partie de la classe moyenne sélectionne une catégorie donnée. Quelle modalité est la plus susceptible?