CHAPITRE 3

Réponses aux exercices du Chapitre 3

MTH3302 - Méthodes probabilistes et statistiques pour l'I.A.

Jonathan Jalbert – Automne 2021

1. a) Oui, la classe semble influencer la probabilité de survie.

   b) On modélise la probabilité de survie $\mu_{\boldsymbol{x}}$ en fonction des variables explicatives $\boldsymbol{x}_i$ avec le modèle suivant :
   $$\ln\left(\frac{\theta_i}{1-\theta_i}\right) = \boldsymbol{x}_i^\top \boldsymbol{\beta}$$

   c) Oui, les coefficients de régression sont significativement différents de 0 (leurs valeurs-p sont plus petites que 5%).

   d) Pour un passager de première classe, on a que $\boldsymbol{x}_0^\top = \begin{bmatrix}1 & 1 & 0\end{bmatrix}$. La probabilité de survie associée à ce $\boldsymbol{x}_0$ est donnée par
   $$\frac{\exp(\hat\beta_0+\hat\beta_1)}{1+\exp(\hat\beta_0+\hat\beta_1)} = \frac{\exp(-1.14+1.67)}{1+\exp(-1.14+1.67)}.$$

   e) L'ordonnée à l'origine β_0 correspond au logarithme de la cote de la probabilité de survie d'un passager de troisième classe.

2. a)
   $$\mathbb{P}(Y=1|X=100) = \frac{e^{-0,6}}{1+e^{-0,6}} \approx 0,35.$$

b) Le logarithme de la cote qu'une personne habitant à 0m d'une station de métro emprunte le transport en commun.

c) On a besoin de 3 variables indicatrices :

x_2	x_3	x_4	Catégorie
0	0	0	Étudiant
1	0	0	Travailleur
0	1	0	Retraité
0	0	1	Autre

Le modèle devient $\mathbb{E}(Y|X=\boldsymbol{x}) = \frac{\exp(\boldsymbol{x}^\top\boldsymbol{\beta})}{1+\exp(\boldsymbol{x}^\top\boldsymbol{\beta})}$ avec $X^\top = \begin{bmatrix}1 & X_1 & X_2 & X_3 & X_4\end{bmatrix}$ et $\boldsymbol{\beta}^\top = \begin{bmatrix}\beta_0 & \beta_1 & \beta_2 & \beta_3 & \beta_4\end{bmatrix}$.

d) Selon le modèle proposé en (c), l'ordonnée à l'origine correspond au logarithme de la cote qu'un étudiant habitant à 0m d'une station de métro emprunte le transport en commun.

3. a) $\hat{\theta} = \bar{Y}$.

b) Le nombre moyens d'accidents par année.

c) L'année pourrait être incluse comme variable explicative.

d) $\mathbb{E}(Y|X=\boldsymbol{x}) = \exp(\boldsymbol{x}^\top\boldsymbol{\beta})$, avec $X^\top = \begin{bmatrix}1 & X_1\end{bmatrix}$ où X_1 est l'année et $\boldsymbol{\beta}^\top = \begin{bmatrix}\beta_0 & \beta_1\end{bmatrix}$.