

7.1 On considère le nombre de semaines pendant lequel des mères allaitent leur nouveaux nés. La base de donnée allaitement contient les variables suivantes :

- *duree* : durée de l'allaitement (en semaines)
- *censure* : variable indicatrice pour la censure, 1 si le temps correspond à la fin de la période d'allaitement, 0 si le temps est censuré à droite
- *race* : race de la mère, soit blanc (1), noir (2) ou autre (3)
- *pauvrete* : est-ce que le revenu de la famille est inférieur au seuil de pauvreté? soit oui (1) ou non (0)
- *fumeur* : est-ce que la mère fumait à la naissance de l'enfant? soit oui (1) ou non (0)
- *alcool* : est-ce que la mère consommait de l'alcool à la naissance de l'enfant? soit oui (1) ou non (0)
- *agemere* : âge de la mère à la naissance de l'enfant
- *annee* : année de la naissance de l'enfant
- *scolarite* : nombre d'années de scolarité de la mère
- *prenat3mois* : absence de soins prénataux : 0 si la mère a requis des soins prénataux dans les trois premiers mois de la grossesse, 1 sinon.

- (a) Estimez et tracez les courbes de survie pour les deux niveaux de la variable du statut de fumeur en utilisant l'estimateur de Kaplan–Meier. Commentez sur les fonctions de survie estimées.
- (b) Selon les résultats de la partie (a), quelle est la probabilité estimée qu'une mère non-fumeuse allaite pendant plus de 36 semaines? Qu'en est-il pour une mère fumeuse?
- (c) Quelle est la médiane et la moyenne du temps d'allaitement pour une mère non-fumeuse? Qu'en est-il pour une mère fumeuse?
- (d) Testez si les courbes de survie des mères non-fumeuses et fumeuses sont les mêmes.
- (e) Ajustez un modèle de Cox à risques proportionnels pour évaluer l'impact du statut de pauvreté, du statut de fumeur, de l'âge de la mère et du nombre d'années de scolarité sur la fonction de risque. Écrivez la fonction de risque en fonction des estimés  $\hat{\beta}$  et interprétez les estimés de ces paramètres.

7.2 Un commerce de chaussures de Montréal veut optimiser son inventaire afin de maximiser ses profits et fait appel à votre société de conseil. Les données chaussures contiennent les variables suivantes :

- *statut* : variable catégorielle, 0 s'il est vendu, 1 si l'article est toujours en stock, 2 s'il est déstocké.
- *temps* : temps de stockage de l'article (en mois).
- *prix* : prix de vente de l'article, arrondi à l'unité près.
- *sexe* : variable catégorielle, 0 pour modèle pour homme, 1 pour femme.

Notre objectif premier est d'estimer la fonction de survie du temps de l'article en stock.

- (a) Que représente la censure dans cet exemple?
- (b) Estimez nonparamétriquement la fonction de survie du temps de stockage à l'aide d'un modèle de Kaplan–Meier et rapportez les estimés des quartiles.
- (c) Ajustez un modèle à risque proportionnel de Cox pour la durée de stockage en fonction du sexe et du prix de vente.
  - i. Rapportez et interprétez les coefficients des variables. Est-ce que les effets estimés sont significatifs?