

Exercice du thème 7

Ce contexte concerne les efforts déployés par une entreprise de télécommunication pour réduire le taux de désabonnement de sa clientèle. Chaque ligne correspond à un client distinct et enregistre diverses informations démographiques et d'utilisation des services. Les données sont dans le fichier « telco.csv ».

La variable « tenure » correspond au nombre de mois d'abonnement. La variable « churn » est une variable binaire indiquant si le client a rompu le contrat ou pas encore (1 = Churn, 0 = Not churn).

1. Construisez et commentez la table de survie et la courbe de survie pour le temps d'abonnement.
2. On souhaite examiner le temps d'abonnement selon la catégorie du client (variable « custcat »). Construisez les courbes de survie des différentes catégories et testez l'égalité des fonctions de survie entre elles. Que peut-on conclure?
3. Le temps d'abonnement diffère-t-il selon le genre du client? Selon la région?
4. La compagnie souhaite identifier les facteurs démographiques qui expliquent le temps d'abonnement. Quelle méthode préconisez-vous? Pouvez-vous identifier ces facteurs?