

Survie

Feuille de TP 1 : Méthodes nonparamétriques

Exercice 1

Considérons un temps de survie T de loi exponentielle de paramètre $\lambda > 0$.

1. Calculer la fonction de répartition, la fonction de survie, le risque instantané, et le risque cumulé.
2. Calculer le quantile d'ordre α , noté q_α pour $\alpha \in]0, 1[$.
3. Représenter le graphe de la fonction $\alpha \mapsto q_\alpha$.
4. Calculer l'espérance et la variance de T .

Exercice 2

Considérons n individus et leurs temps de survie respectifs $T_1, \dots, T_n$. Considérons que les données sont aléatoirement censurées à droite. On observe donc (X_i, δ_i) , $i = 1, \dots, n$ avec $X_i = \min(T_i, C_i)$, où C_i la censure et $\delta_i = \mathbb{1}_{T_i \leq C_i}$, l'indicateur de censure. On se propose d'estimer S la fonction de survie de T à partir des observations (X_i, δ_i) , $i = 1, \dots, n$. Soient

$$\hat{S}_1(t) = \frac{1}{n} \sum_{i=1}^n \mathbb{1}_{X_i > t} \text{ et } \hat{S}_2(t) = \frac{1}{\sum_{i=1}^n \delta_i} \sum_{i=1}^n \mathbb{1}_{X_i > t, \delta_i = 1}$$

1. Rappeler la définition de $S(t)$.
2. Interpréter $\hat{S}_1(t)$.
3. Vers quelle fonction $\hat{S}_1(t)$ converge-t-elle ? Conclusion ?
4. Interpréter $\hat{S}_2(t)$.
5. Vers quelle fonction $\hat{S}_2(t)$ converge-t-elle ? Conclusion ?

Exercice 3

(Introduction à R- Etude préliminaire)

On s'intéresse au délai d'apparition d'une infection localisée au point d'insertion du cathéter chez des patients souffrant de maladie rénale et ayant un équipement de dialyse portable. Le cathéter peut être enlevé pour des raisons autres que l'infection, ce qui conduit à des observations censurées. Les données se trouvent dans le fichier *kidney.txt*. Le fichier comprend les informations suivantes :

Patient : numéro du patient

Time : délai d'apparition d'une infection

Status : 0 pour observation censurée, 1 sinon

Age : âge

Sex : 1 pour homme, 2 pour femme

Disease : type de la maladie (0=GN, 1=AN, 2=PKD, 3=Other).

1. Compter le nombre de données censurées dans les deux groupes définis par le sexe.
2. Récupérer les âges respectifs des hommes et des femmes dans deux tableaux. Faire une représentation graphique de l'âge en fonction du sexe.
3. Calculer la moyenne et la variance de l'âge dans les deux groupes définis par le sexe.
4. Comparer la moyenne d'âge entre les hommes et les femmes.
5. Comparer la répartition du type d'atteinte rénale en fonction du sexe.
6. L'âge moyen est-il le même quel que soit le type d'atteinte rénale?

Exercice 4

On s'intéresse au temps de survie de patients atteints de leucémie myéloïde aigue. Les données se trouvent dans le fichier *aml.txt*. Le fichier comprend les informations suivantes :

time : temps de survie

status : 0 pour une observation censurée et 1 sinon

x : maintien d'un traitement par chimiothérapie.

1. Charger les librairies **survival**, **ISwR** et **MASS**.
2. Récupérer le jeu de données **aml**, afficher les informations concernant ce jeu de données et les premières observations contenues dans ce jeu de données.
3. Créer un objet de type "survie" sous R à partir de cette base. Pour cela, on utilise la fonction **Surv** avec la syntaxe **base= Surv(times,status==1)**. Les données censurées sont marquées par un "+".
4. Calculer à la main par la méthode de Kaplan-Meier, la fonction de survie globale, puis celles des deux groupes décrits par la variable "maintien du traitement par chimiothérapie".
5. Représenter à la main sur un même graphique ces deux courbes de survie en indiquant par un "+" les dates où ont lieu les censures.
6. Répondre aux deux questions précédentes en utilisant la fonction **survfit** de R.
7. On cherche à mettre en évidence une éventuelle différence de survie entre les deux groupes décrits par la variable "maintien du traitement par chimiothérapie". Calculer les paramètres (moyenne, médiane, quartiles) dans les deux groupes. Comparer avec les résultats sur la population globale.
8. Comparer graphiquement les deux courbes de survie.
9. Mettre en place un test de l'hypothèse "les courbes de survie sont égales dans les deux groupes". On précisera le (les) test(s) utilisé(s), la statistique de test, sa loi sous l'hypothèse nulle, sa zone de rejet. Mettre en oeuvre le(s) test(s) avec R en utilisant la fonction **survdif**. On utilisera l'option **rho** de la fonction **survdif**.

Exercice 5

On s'intéresse à la base `melanom` de la library `ISwR`. On s'intéresse en particulier à la durée de suivi `Days` et au statut `status` (1 : mort liée mélanome, 2 : vivant à la fin de l'étude, 3 : mort d'une autre cause).

1. Pour quelles valeurs de la variable « status », les données sont censurées ? Expliquer.
2. Créer un objet de type "survie" sous R à partir de cette base. Pour cela, on utilise la fonction `Surv`. `base.mel = Surv(days,status==1)`
Les censures sont marquées par un "+".
3. Estimer la courbe de survie par la méthode de Kaplan-Meier, à l'aide de la fonction `survfit`.
4. Estimer les courbes de survie pour chaque sexe (1 : F, 2 : H)
5. On souhaite maintenant tester si le sexe a une influence sur la survie. Quel test proposez vous ? Le mettre en place à l'aide de la fonction `survdif` de R. Quel autre test peut-on réaliser ? Utiliser pour cela l'option `rho` de la fonction `survdif`.