

DATA SCIENCE ET APPLICATIONS

DM Analyse de survie

Documents :

Diapositives sur l'analyse de survie

Librairie Python : lifelines

Travail à faire

Etudier les fonctionnalités de la librairie « lifelines » qui permettent de faire de l'analyse de survie.

Ecrire un programme Python permettant de réaliser les fonctionnalités suivantes :

- Lecture du fichier de données (fichier fourni)
- Calcul des statistiques descriptives sur la variable « time »
- Afficher un histogramme de la durée (variable « time ») pour l'ensemble de la population ou par sexe (F/H)
- Estimer la probabilité de survie et l'intervalle de confiance en utilisant la fonction Kaplan-Meier. Afficher le tableau des proportions de survivants à l'instant t ($t=0, \dots, n$)
- Représenter graphiquement la courbe de survie avec l'intervalle de confiance
- Représenter la courbe de Kaplan-Meier pour chacun des deux groupes (H/F).

Facultatif

Pour plus de convivialité, vous pouvez utiliser la librairie « Streamlit » qui permet de faciliter l'affichage sur une page Web et ajouter des boutons, des listes déroulantes,

Une autre librairie peut être utilisée : « plotly » couplée à la librairie « Dash ».

Travail à faire à 3 personnes maximum