

# Fonctions de base pour l'utilisation du logiciel R

## Remarques :

Un **argument** mis en **rouge** indique qu'il est obligatoire.

Un **argument** mis en **vert** indique la valeur qui est prise par défaut.

## Aide et bases

?**"sujet"** : retourne le fichier d'aide correspondant au **sujet** demandé (une fonction, un nom de fichier de données, ...).

help(**"sujet"**) : retourne le fichier d'aide correspondant au **sujet** demandé (une fonction, un nom de fichier de données, ...).

help.search(**"sujet"**) : cherche dans **R** tout les fichiers d'aide comportant le **sujet** en question.

apropos(**"sujet"**) : retourne le nom des objets comportant le **sujet** en question.

attach(**"NomFichier"**, **pos=2**, ...) : attache le fichier se nommant **NomFichier** dans un répertoire interne à **R** à la 2<sup>ème</sup> position de son arborescence.

detach(**"NomFichier"**, **pos=2**, ...) : détache le fichier se nommant **NomFichier** du répertoire interne à **R** qui se trouve à la 2<sup>ème</sup> position de son arborescence.

attributes(**"MonObjet"**) : retourne les attributs de **MonObjet** ainsi que leurs valeurs.

## Utilisation des fonctions, dataset d'un package

library(**"Nom\_de\_la\_librairie"**) # exemple : library(survival) pour charger la librairie survival

## Importation de données

read.csv2(**"AdresseFichier"**, **sep=";"**, **header=TRUE**,...) : pour importer un fichier enregistré au format \*.csv, séparateur : point-virgule ;

L'argument **AdresseFichier** correspond au nom du fichier dont les données doivent être lues ;

L'adresse doit être écrite de la manière suivante : C:\\**AdresseFichier**.csv. **Attention, il faut utiliser \\ ou / à la place du classique \** ;

L'argument **sep=";"** indique que par défaut le séparateur utilisé est le point-virgule. La valeur de cet argument peut être modifiée en fonction du séparateur utilisé (exemple : **sep=","** si le séparateur virgule est utilisé ;

L'option **header=TRUE** doit être présente lorsque la première ligne du fichier de données comporte le nom des variables.

read.table(**"AdresseFichier"**, **header=TRUE**, **sep=" "**,...) : pour importer un fichier enregistré au format \*.txt avec un espace comme séparateur ;

L'argument **AdresseFichier** correspond au nom du fichier dont les données doivent être lues ;

L'adresse doit être écrite de la manière suivante : C:\\**AdresseFichier**.csv. **Attention, il faut utiliser \\ ou / à la place du classique \** ;

L'argument **sep=" "** indique que par défaut le séparateur utilisé est l'espace. ;

L'option **header=TRUE** doit être présente lorsque la première ligne du fichier de données comporte le nom des variables.

## Statistiques descriptives

length(**x**) : retourne le nombre d'éléments du vecteur **x**.

max(**x**) : retourne la valeur maximale du vecteur **x**.

min(**x**) : retourne la valeur minimale du vecteur **x**.

range(**x**) : retourne la valeur maximale et la valeur minimale du vecteur **x**.

median(**x**) : retourne la valeur médiane du vecteur **x**.

mean(**x**) : retourne la valeur moyenne du vecteur **x**.

quantile(**x**, **probs=0:4/4**) : retourne les valeurs extrêmes et les quartiles du vecteur **x**.

sum(**x**) : retourne la somme des éléments du vecteur **x**.

summary(**x**) : retourne la valeur moyenne, les valeurs extrêmes et les quartiles du (des) vecteur(s) **x**.

cor(**x**, **y**, **method=c("pearson")**) : retourne la valeur du coefficient de corrélation de Pearson entre le vecteur **x** et le vecteur **y** ;

Si **x** est une matrice, retourne une matrice contenant les coefficients de corrélation de Pearson entre les différents éléments de **x** ;

# Fonctions de base pour l'utilisation du logiciel R

L'argument **method** permet de choisir la méthode d'estimation des coefficients de corrélation : "**pearson**", "**kendall**" ou "**spearman**".

**var(x)** : retourne la variance du vecteur **x**.

Si **x** est une matrice, retourne la matrice de variance-covariance de **x** ;

**sd(x)** : retourne l'écart-type du vecteur **x**.

Si **x** est une matrice, retourne l'écart-type de chaque colonnes de **x** ;

**table(...)** : donne la fréquence absolue des éléments **x** quand il s'agit d'un vecteur, **table(x)**, ou un (des) tableau(x) de contingence(s) des fréquences absolues quand il y a plusieurs vecteurs, **table(x, y)**.

**prop.table(x, margin=NULL)** : donne les fréquences des cellules contenues dans le tableaux **x**. L'argument **margin** détermine par rapport à quelle marge du tableau les fréquences seront calculées : **prop.table(x, margin=1)** pour le total ligne et **prop.table(x, margin=2)** pour le total colonne.

**margin.table(x, margin=NULL)** : donne la somme des entrées d'un tableau **x**. L'argument **margin** détermine pour quelle marge du tableau les sommes seront calculées : **margin.table(x, margin=1)** pour le total ligne et **margin.table(x, margin=2)** pour le total colonne.

## Graphiques

**plot(x, y, ...)** : fait le graphique de **x** (vecteur d'abscisse) en fonction de **y** (vecteur des ordonnées).

**x** peut être aussi un objet **R** pour lequel il existe une fonction graphique prédéfinie (par exemple, un objet de classe "survfit" retourné par la fonction `survfit(...)`) ;

**...** : arguments permettant d'apporter des spécifications au graphique. Par exemple, `type="l"`, pour représenter une ligne, `type="p"`, pour représenter des points, `xlab="MonLabel"`, pour personnaliser le label mis sur l'axe des abscisses, ...

**hist(x, probability=F, ...)** : retourne l'histogramme des fréquences des valeurs de **x** (où **x** est numérique). Retourne l'histogramme des probabilité lorsque **probability=T**.

**barplot(x, probability=F, ...)** : retourne l'histogramme des valeurs de **x** (où **x** peut être un facteur).

**boxplot(x, ...)** : retourne un diagramme en boîte à moustache des éléments de **x**.

**pie(x, ...)** : retourne un graphique en camembert des éléments de **x**.

**jpeg("Adresse Graphique")** : enregistre un graphique au format \*.jpeg à l'adresse spécifiée.

**bmp("Adresse Graphique")** : enregistre un graphique au format \*.bmp à l'adresse spécifiée.

## Tests statistiques

**chisq.test(x, y, correct="T", ...)** : réalise le test du Chi-deux de comparaison de **x** et **y**. Utilisable également lorsque **x** est sous la forme d'un tableau de contingence.

L'argument **correct** peut prendre les valeurs logiques "**T**" ou "**F**" selon que l'on utilise la correction de Yates ou non.

**t.test(x, y, alternative="two.sided", paired="F", var.equal="T", ...)** : réalise le test t de Student de comparaison de **x** et **y**.

L'argument **alternative** peut prendre les valeurs "**two.sided**", "**greater**" ou "**less**" selon que le test est réalisé en situation bilatérale, unilatéral de supériorité ou unilatéral d'infériorité ;

L'argument **paired** peut prendre les valeurs logiques "**F**" ou "**T**" selon que le test est non apparié ou apparié ;

L'argument **var.equal** peut prendre les valeurs logiques "**F**" ou "**T**" selon que l'hypothèse d'égalité des variances n'est pas vérifiée ou qu'elle est vérifiée.

**cor.test(x, y, alternative="two.sided", paired="F", method=c("pearson"), ...)** : réalise le test de corrélation entre **x** et **y**.

L'argument **alternative** peut prendre les valeurs "**two.sided**", "**greater**" ou "**less**" selon que le test est réalisé en situation bilatérale, unilatéral de supériorité ou unilatéral d'infériorité ;

L'argument **method** peut prendre les valeurs "**pearson**", "**kendall**" ou "**spearman**" selon la méthode choisie pour estimer les coefficients de corrélation.

Exécutez la commande **help.search("test")** pour afficher les noms des commandes utilisant le mot test.