

## Introduction aux logiciels pour les statistiques (IS3) - 2023/2024

### Commandes en R et Python

#### INDICATEURS STATISTIQUES

Commande R	Commande Python	Description
<code>install.packages("dplyr")</code>	<code>pip install pandas</code>	installer un package ou un module
<code>library(dplyr)</code>	<code>import pandas as pd</code>	charger un package ou un module
<code>dim(matable)</code>	<code>matable.shape</code>	nombre de lignes et de colonnes de la table
<code>str(matable)</code>	<code>matable.info()</code>	court descriptif des variables de la table
<code>summary(matable)</code>	<code>matable.describe()</code>	quelques indicateurs statistiques pour les variables de la table
<code>summary(matable\$mavariation)</code>	<code>matable.mavariation.describe()</code>	quelques indicateurs statistiques pour une variable de la table
<code>var(matable\$mavariation)</code>	<code>matable.mavariation.var()</code>	variance corrigée d'une seule variable de la table
<code>sd(matable\$mavariation)</code>	<code>matable.mavariation.std()</code>	écart-type corrigé d'une seule variable de la table

N.B : Pour obtenir la variance non corrigée, on peut utiliser l'astuce vue en cours en R et Python.

## PROBABILITÉS



**lettre****nom\_fonction**(arguments)

### Nom de la fonction

norm, exp, pois...

### Lettres

d : fonction de densité (cas « continue ») ou fonction de probabilité (cas « discret »).

p : fonction de répartition F où  $F(k)=P[X<k]$

q : fonction des quantiles.

r : génère des valeurs aléatoires



```
import scipy.stats as stats
```

**stats****nom\_fonction****mot\_clé**(arguments)

### Nom de la fonction

norm, expon, poisson...

### Mots clés

pdf(x, parametre) : densité en x

pmf(k, parametre) : densité discrète en k

cdf(x,parametre) : fonction de répartition en x

ppf(q, parametre) : q-ième quantile.

rvs : simulation

## MEMENTO R

### Vecteur V

1. On le construit avec la fonction `c()`. Elle permet de combiner les valeurs individuelles dans un vecteur unique.
2. Toutes les valeurs d'un vecteur doivent obligatoirement du même type. Sinon R essaiera de les convertir au mieux.
3. Une valeur manquante *NA* (not available) présente dans un vecteur n'a aucune incidence sur son type.
4. *NULL* signifie objet vide et ne doit pas être confondu avec *NA*.
5. La fonction `class(V)` indique si *V* est un entier (integer), un réel (numeric), une chaîne de caractère (character) ou un booléen (logical).
6. `length(V)` : renvoie le nombre d'éléments de *V*.
7. `rep(k,l)` : créer un vecteur de longueur *l* mais dont toutes les valeurs sont égales à *k*.
8. `seq(i, j, by = d)` : créer un vecteur dont les valeurs vont de *i* à *j* avec un pas de *d*.
9. On indique entre crochets les éléments à conserver de *V*.
  - (a) Indexation par position : `V[c(2,5,6)]` permet de sélectionner les 2<sup>ème</sup>, 5<sup>ème</sup> et 6<sup>ème</sup> éléments de *V*.
  - (b) Indexation par position : `V[c(-2,-5,-6)]` permet de sélectionner tous les éléments de *V* sauf le 2<sup>ème</sup>, 5<sup>ème</sup> et 6<sup>ème</sup>.
  - (c) Indexation par condition : `V[V>8]` permet de sélectionner les éléments de *V* strictement plus grands que 8.
10. `which(V<=2)` : renvoie la position des éléments qui sont inférieurs ou égaux à 2.
11. `V3<-c(V1,V2)` : permet de combiner deux vecteurs *V1* et *V2* en un seul vecteur *V3*.
12. `V[2]<-74` : remplace le 2<sup>ème</sup> élément de *V* par le nombre 46.

### Liste L

1. Contrairement au vecteur, une liste ne contient pas forcément que des éléments de même type .
2. Une liste se crée avec la fonction `list`.
3. On indique entre double crochets les éléments à conserver de *L*.
4. `length(L)` : renvoie le nombre d'éléments de *L*.
5. `L3<-append(L1,L2)` : permet de combiner deux listes *L1* et *L2* en une seule liste *L3*.

### Fonctions

```
somme1<-function(a,b){  
  somme<-a+b  
  return(somme)  
}
```

1. Objet de R qui reçoit des arguments en entrée et renvoie un résultat.
2. Certaines existent déjà en R.
3. Une fonction peut renvoyer un objet de R ou aucun objet de R.
4. Une fonction qui ne renvoie pas d'objet de R est appelée procédure .

```
somme3<-function(a,b){  
  somme<-a+b  
  print(somme)  
}
```

5. Quand on souhaite renvoyer plusieurs objets de R, il faut les stocker dans une liste et renvoyer la liste.

```
somme2<-function(a,b){  
  somme<-a+b  
  resultat<-list(a,b,somme)  
  return(resultat)  
}
```

6. Pour obtenir de l'aide sur une fonction d'un package importé, il faut taper dans la console le nom de ladite fonction précédé d'un point d'interrogation.

## Une matrice M

1. `M<-matrix(c(18,15,12,3,6,14), nrow=2, ncol=3, byrow=TRUE)` : créer une matrice avec 2 lignes et 3 colonnes en remplissant ligne par ligne.
2. Indexation par crochets :
  - (a) `M[i, j]` désigne l'élément de la ligne i et de la colonne j de la matrice M
  - (b) `M[i, ]` désigne le vecteur correspondant à la i<sup>ème</sup> ligne de la matrice M
  - (c) `M[, j]` désigne le vecteur correspondant à la j<sup>ème</sup> colonne de la matrice M

## Une data.frame

1. Objet de R composé de vecteurs nommés et de tailles égales mais pas forcément du même type.
2. On crée une data.frame avec la fonction `data.frame()`. Elle prend en arguments les vecteurs nommés qui la constitueront.
3. Un package de R contient souvent des jeux de données. Il faut importer le package et ensuite taper `data` dans la console. Un menu avec l'ensemble des jeux de données utilisables s'affiche.
4. On peut interroger la table grâce à une approche matricielle, vectorielle ou avec les verbes du package `dplyr` (voir support du cours).
5. On peut renommer une variable avec la fonction `rename`.

### Parcourir la table « participants »

	Prenom	Sexe	Victoires
1	Paul	H	3
2	Julie	F	4
3	Susie	F	2
4	Eric	H	3

participants[,2]  
participants[, "Victoires"]  
participants\$Victoires

participants[3,3]  
participants[3, "Victoires"]  
participants\$Victoires[3]  
participants[Prenom == "Susie", "Victoires"]  
participants\$Victoires[Prenom == "Susie"]

### Remarques :

Contrairement à certains langages de programmation, il ne faut pas définir le type de la variable avant de l'implémenter. En affectant une valeur à une variable, R la crée directement en lui attribuant un type par défaut.

Pour approfondir les notions susmentionnées, le lecteur peut se référer aux sources et à la bibliographie du cours.