

Rstudio TD 1

(Prise en main / Stat descriptives)

Clément Rau

Laboratoire de Mathématiques de Toulouse

Université Paul Sabatier-IUT GEA Ponsan

- 1 **Prise en main**
 - Introduction
 - Opérations élémentaires
 - Affectation
 - Structures de contrôle
 - Autour des séries
- 2 **Quelques fonctions liées aux statistiques descriptives**
 - Quantités cumulées
 - Calcul des paramètres standards
 - Représentations graphique
- 3 **Importer des données depuis un tableur**
- 4 **Exercices**
 - Exo 1
 - Exo 2
- 5 **Pour aller plus (ou moins) loin...**
 - Exo 3

Remarques sur R

- R est un logiciel permettant de faire des analyses statistiques et de produire des graphiques.

Remarques sur R

- R est un logiciel permettant de faire des analyses statistiques et de produire des graphiques.
- R est aussi un langage de programmation complet.

Remarques sur R

- R est un logiciel permettant de faire des analyses statistiques et de produire des graphiques.
- R est aussi un langage de programmation complet.
- ► R est un logiciel gratuit qui se télécharge sur www.r-project.org (Vidéo tuto [ici](#).)

Remarques sur R

- R est un logiciel permettant de faire des analyses statistiques et de produire des graphiques.
- R est aussi un langage de programmation complet.
- ▶ R est un logiciel gratuit qui se télécharge sur www.r-project.org (Vidéo tuto [ici](#).)
 - ▶ Il existe également des versions en ligne [ici](#) , ou [ici](#) ou encore [ici](#) , qui ne nécessitent donc pas d'installation, (parfois juste la création d'un compte).

Remarques sur R

- R est un logiciel permettant de faire des analyses statistiques et de produire des graphiques.
- R est aussi un langage de programmation complet.
- - ▶ R est un logiciel gratuit qui se télécharge sur www.r-project.org (Vidéo tuto [ici](#).)
 - ▶ Il existe également des versions en ligne [ici](#) , ou [ici](#) ou encore [ici](#) , qui ne nécessitent donc pas d'installation, (parfois juste la création d'un compte).
 - ▶ Enfin, pour les "fans d'ios", sachez qu'il existe un appli gratuite sur iphone (nommée *R compiler*) qui fonctionne très bien. (Probablement qu'une appli similaire sur Android existe aussi.)

Remarques sur R

- R est un logiciel permettant de faire des analyses statistiques et de produire des graphiques.
- R est aussi un langage de programmation complet.
- - ▶ R est un logiciel gratuit qui se télécharge sur www.r-project.org (Vidéo tuto [ici](#).)
 - ▶ Il existe également des versions en ligne [ici](#) , ou [ici](#) ou encore [ici](#) , qui ne nécessitent donc pas d'installation, (parfois juste la création d'un compte).
 - ▶ Enfin, pour les "fans d'ios", sachez qu'il existe un appli gratuite sur iphone (nommée *R compiler*) qui fonctionne très bien. (Probablement qu'une appli similaire sur Android existe aussi.)
- De l'aide sur la fonction `bloumbloom` s'obtient par
`>help(bloumbloom)`

Interface, Fenêtres dans R

The screenshot displays the RStudio interface with several windows and annotations:

- Source Editor:** Contains R code for creating a data frame and reading a CSV file. A blue box highlights the text "Vue des données, fonctions et scripts" (View of data, functions and scripts).
- Environment:** Shows the loaded data frame 'ra2010' with 2879 observations and 21 variables. A blue box highlights the text "Vue des éléments en mémoire et de l'historique des commandes" (View of elements in memory and command history).
- Console:** Shows the execution of R commands. A blue box highlights the text "Console".
- Files:** Shows a file browser view of the 'Home' directory. A blue box highlights the text "Vue des répertoires et graphiques, des packages et de l'aide" (View of directories and graphics, packages and help).

Répertoire de travail

R a par défaut un répertoire de travail. Cela signifie que si aucun chemin n'est spécifié, R va lire ou écrire dans ce répertoire.

Répertoire de travail

R a par défaut un répertoire de travail. Cela signifie que si aucun chemin n'est spécifié, R va lire ou écrire dans ce répertoire.

Pour connaître ce répertoire de travail, saisir dans la console de commande :

```
getwd()
```

Répertoire de travail

R a par défaut un répertoire de travail. Cela signifie que si aucun chemin n'est spécifié, R va lire ou écrire dans ce répertoire.

Pour connaître ce répertoire de travail, saisir dans la console de commande :

```
getwd()
```

Pour changer de répertoire de travail, saisir dans la console de commande :

```
setwd("chemin du répertoire souhaité")
```

Répertoire de travail

R a par défaut un répertoire de travail. Cela signifie que si aucun chemin n'est spécifié, R va lire ou écrire dans ce répertoire.

Pour connaître ce répertoire de travail, saisir dans la console de commande :

```
getwd()
```

Pour changer de répertoire de travail, saisir dans la console de commande :

```
setwd("chemin du répertoire souhaité")
```

Attention, dans les chemins Windows, les \ doivent être remplacé par des / ou des \\

- 1 **Prise en main**
 - Introduction
 - Opérations élémentaires
 - Affectation
 - Structures de contrôle
 - Autour des séries
- 2 **Quelques fonctions liées aux statistiques descriptives**
 - Quantités cumulées
 - Calcul des paramètres standards
 - Représentations graphique
- 3 **Importer des données depuis un tableur**
- 4 **Exercices**
 - Exo 1
 - Exo 2
- 5 **Pour aller plus (ou moins) loin...**
 - Exo 3

Essayer qqes opérations élémentaires : addition, multiplication etc...

Essayer qqes opérations élémentaires : addition,
multiplication etc...

$>2 * 5$

Essayer qqes opérations élémentaires : addition, multiplication etc...

>2 * 5

>4 + 2

Essayer qqes opérations élémentaires : addition, multiplication etc...

>2*5

>4+2

>2,5*2

Essayer qqes opérations élémentaires : addition, multiplication etc...

>2*5

>4+2

>2,5*2

>2.5*2

Essayer qqes opérations élémentaires : addition, multiplication etc...

>2*5

>4+2

>2,5*2

>2.5*2

>2^3

Essayer qqes opérations élémentaires : addition, multiplication etc...

```
>2*5
```

```
>4+2
```

```
>2,5*2
```

```
>2.5*2
```

```
>2^3
```

```
>sqrt(16)
```

Essayer qqes opérations élémentaires : addition, multiplication etc...

```
>2*5
```

```
>4+2
```

```
>2,5*2
```

```
>2.5*2
```

```
>2^3
```

```
>sqrt(16)
```

```
>ln(2)
```

Essayer qqes opérations élémentaires : addition, multiplication etc...

```
>2*5
```

```
>4+2
```

```
>2,5*2
```

```
>2.5*2
```

```
>2^3
```

```
>sqrt(16)
```

```
>ln(2)
```

```
>log(2)
```

Essayer qqes opérations élémentaires : addition, multiplication etc...

```
>2*5
```

```
>4+2
```

```
>2,5*2
```

```
>2.5*2
```

```
>2^3
```

```
>sqrt(16)
```

```
>ln(2)
```

```
>log(2)
```

```
>log(100,10)
```


Essayer qqes opérations élémentaires : addition, multiplication etc...

```
>2*5
```

```
>4+2
```

```
>2,5*2
```

```
>2.5*2
```

```
>2^3
```

```
>sqrt(16)
```

```
>ln(2)
```

```
>log(2)
```

```
>log(100,10)
```

```
>exp(1)
```

Essayer qqes opérations élémentaires : addition, multiplication etc...

```
>2*5
```

```
>4+2
```

```
>2,5*2
```

```
>2.5*2
```

```
>2^3
```

```
>sqrt(16)
```

```
>ln(2)
```

```
>log(2)
```

```
>log(100,10)
```

```
>exp(1)
```

```
>log(exp(3))
```

Nombre de décimales souhaitées

On utilise la fonction :

$$\text{round}(x, n)$$

où x désigne un réel et n un entier.

Nombre de décimales souhaitées

On utilise la fonction :

$$\text{round}(x, n)$$

où x désigne un réel et n un entier.

La fonction renvoie une valeur approchée du réel x avec uniquement les n premières décimales.

- 1 **Prise en main**
 - Introduction
 - Opérations élémentaires
 - Affectation
 - Structures de contrôle
 - Autour des séries
- 2 **Quelques fonctions liées aux statistiques descriptives**
 - Quantités cumulées
 - Calcul des paramètres standards
 - Représentations graphique
- 3 **Importer des données depuis un tableur**
- 4 **Exercices**
 - Exo 1
 - Exo 2
- 5 **Pour aller plus (ou moins) loin...**
 - Exo 3

Affectation

Si l'on veut affecter à x la valeur 3, on peut taper :

`>x=3` ou bien `>x<-3`

Affectation

Si l'on veut affecter à x la valeur 3, on peut taper :

`>x=3` ou bien `>x<-3`

On peut alors demander que vaut x :

`>x`

Affectation

Si l'on veut affecter à x la valeur 3, on peut taper :

`>x=3` ou bien `>x<-3`

On peut alors demander que vaut x :

`>x`

Donner à x et y deux valeurs et faire $x+y$, $3x-y$. Contrôler.

Affectation

Si l'on veut affecter à x la valeur 3, on peut taper :

`>x=3` ou bien `>x<-3`

On peut alors demander que vaut x :

`>x`

Donner à x et y deux valeurs et faire $x+y$, $3x-y$. Contrôler.

!!! Attention R distingue majuscule et minuscule !!!

Affectation

Si l'on veut affecter à x la valeur 3, on peut taper :

`>x=3` ou bien `>x<-3`

On peut alors demander que vaut x :

`>x`

Donner à x et y deux valeurs et faire $x+y$, $3x-y$. Contrôler.

!!! Attention R distingue majuscule et minuscule !!!

Pour affecter plusieurs variables, on les sépare par un ;

Affectation

Si l'on veut affecter à x la valeur 3, on peut taper :

`>x=3` ou bien `>x<-3`

On peut alors demander que vaut x :

`>x`

Donner à x et y deux valeurs et faire $x+y$, $3x-y$. Contrôler.

!!! Attention R distingue majuscule et minuscule !!!

Pour affecter plusieurs variables, on les sépare par un ;

Exemple : `a=2 ; b=3; c=0`

- Si l'on veut affecter à S1 la série de nombres 3,5,1,12,20.
On tape :
`>S1=c (3, 5, 1, 12, 20)`

- Si l'on veut affecter à S1 la série de nombres 3,5,1,12,20.
On tape :
`>S1=c (3, 5, 1, 12, 20)`
Vérifier en tapant `>S1`

- Si l'on veut affecter à S1 la série de nombres 3,5,1,12,20.
On tape :
`>S1=c(3, 5, 1, 12, 20)`
Vérifier en tapant `>S1`
- Si l'on veut affecter à S2 la série de caractères : rouge, bleu, vert. On tape :
`>S2=c("rouge", "bleu", "vert")`

- Si l'on veut affecter à S1 la série de nombres 3,5,1,12,20.
On tape :
`>S1=c(3, 5, 1, 12, 20)`
Vérifier en tapant `>S1`
- Si l'on veut affecter à S2 la série de caractères : rouge, bleu, vert. On tape :
`>S2=c("rouge", "bleu", "vert")`
Vérifier en tapant `> S2`

- Si l'on veut affecter à S1 la série de nombres 3,5,1,12,20.

On tape :

```
>S1=c(3, 5, 1, 12, 20)
```

Vérifier en tapant >S1

- Si l'on veut affecter à S2 la série de caractères : rouge, bleu, vert. On tape :

```
>S2=c("rouge", "bleu", "vert")
```

Vérifier en tapant > S2

En tapant >mode(S2), R nous indique la nature de la série S2.

En tapant >length(S1), R nous indique la taille de la série S1.

Rentrer une série rapidement

En utilisant SCAN la saisie d'une série est moins fastidieuse.

taper `>S=scan()`

On rentre successivement les valeurs de la série. Essayer

Pense bête sur les affectations

- la fonction `>ls.str()` permet d'afficher tous les détails sur les affectations en cours.

Pense bête sur les affectations

- la fonction `>ls.str()` permet d'afficher tous les détails sur les affectations en cours.
- la fonction `>rm(list=ls())` permet d'effacer toutes les affectations.

- 1 **Prise en main**
 - Introduction
 - Opérations élémentaires
 - Affectation
 - Structures de contrôle
 - Autour des séries
- 2 **Quelques fonctions liées aux statistiques descriptives**
 - Quantités cumulées
 - Calcul des paramètres standards
 - Représentations graphique
- 3 **Importer des données depuis un tableur**
- 4 **Exercices**
 - Exo 1
 - Exo 2
- 5 **Pour aller plus (ou moins) loin...**
 - Exo 3

Maintenant que nous avons fait un tour d'ensemble des fonctionnalités de R, revenons à quelques rudiments de programmation.

Structure conditionnelle, Boucle if

Structure conditionnelle, Boucle if

Syntaxe :

```
if (condition) A else B
```

calcule les instructions A si la condition est vraie, et les instructions B sinon.

Structure conditionnelle, Boucle if

Structure conditionnelle, Boucle if

Exemple 1 :

```
>a=3
```

```
>if (a>0) a=a+1 else a=a-1
```

```
>a
```

```
[1] 4
```

Structure conditionnelle, Boucle if

Exemple 1 :

```
>a=3
```

```
>if (a>0) a=a+1 else a=a-1
```

```
>a
```

```
[1] 4
```

Essayer avec $a = -5$

Structure conditionnelle, Boucle if

Structure conditionnelle, Boucle if

Exemple 2 :

```
>a=2
```

```
>if (a>0) b=log(a) else b=0
```

```
>b
```

```
[1] 0.6931472
```

Structure conditionnelle, Boucle if

Exemple 2 :

```
>a=2
```

```
>if (a>0) b=log(a) else b=0
```

```
>b
```

```
[1] 0.6931472
```

Essayer avec $a = -3$

Structure itérative, Boucle for

Structure itérative, Boucle for

Syntaxe :

```
for (var in seq) commandes
```

Structure itérative, Boucle for

Structure itérative, Boucle for

Exemple 1 :

```
> a=4
```

```
> for (i in 1:2) a=a+i
```

```
> a
```

```
[1] 7
```

Structure itérative, Boucle for

Structure itérative, Boucle for

Exemple 2 :

```
> for (i in 1:5) print(i)
```

```
[1] 1
```

```
[1] 2
```

```
[1] 3
```

```
[1] 4
```

```
[1] 5
```

Structure itérative, Boucle for

Structure itérative, Boucle for

Exemple 3 :

```
> a=100
```

```
> for (i in 1:5) (a=a+1) & print(a) & print(i)
```

```
[1] 101
```

```
[1] 1
```

```
[1] 102
```

```
[1] 2
```

```
[1] 103
```

```
[1] 3
```

```
[1] 104
```

```
[1] 4
```

```
[1] 105
```

```
[1] 5
```

Structure itérative, Boucle for

Structure itérative, Boucle for

Exemple 4 :

```
> b=numeric(length(10))
```

```
> for (i in 1:10) b[i]=i^2
```

```
> b
```

```
[1] 1 4 9 16 25 36 49 64 81 100
```

```
>
```

Structure itérative, Boucle while

Structure itérative, Boucle while

Syntaxe :

```
while (condition) commandes
```

Structure itérative, while

Structure itérative, while

Exemple 1 :

```
> n=1
> i<-1
>
> while (i<6) {
+ print(i)
+ n<-i*n
+ i<-i+1
+ print(n) }
```

Structure itérative, while

Exemple 1 :

```
> n=1
> i<-1
>
> while (i<6) {
+ print(i)
+ n<-i*n
+ i<-i+1
+ print(n) }
```

Remarque : shift + entrée, vous permet d'aller à la ligne dans les instructions sous R.

Structure itérative, while

Exemple 1 :

```
> n=1
> i<-1
>
> while (i<6) {
+ print(i)
+ n<-i*n
+ i<-i+1
+ print(n) }
```

Remarque : shift + entrée, vous permet d'aller à la ligne dans les instructions sous R.

Vous venez de fabriquer factoriel de 5, noté 5!

- 1 **Prise en main**
 - Introduction
 - Opérations élémentaires
 - Affectation
 - Structures de contrôle
 - Autour des séries
- 2 **Quelques fonctions liées aux statistiques descriptives**
 - Quantités cumulées
 - Calcul des paramètres standards
 - Représentations graphique
- 3 **Importer des données depuis un tableur**
- 4 **Exercices**
 - Exo 1
 - Exo 2
- 5 **Pour aller plus (ou moins) loin...**
 - Exo 3

Accès à une ou des valeurs de la série

Soit S une série, par exemple $S = (12, 10, 11, 19, 9, 10)$. On souhaite parfois avoir accès à la i ème valeur, on utilise alors :

```
>S[3]
```

Accès à une ou des valeurs de la série

Soit S une série, par exemple $S = (12, 10, 11, 19, 9, 10)$. On souhaite parfois avoir accès à la i ème valeur, on utilise alors :

```
>S[3]
```

Essayer !

Accès à une ou des valeurs de la série

Soit S une série, par exemple $S = (12, 10, 11, 19, 9, 10)$. On souhaite parfois avoir accès à la i ème valeur, on utilise alors :

```
>S[3]
```

Essayer !

Si l'on souhaite avoir uniquement les 3, 4 et 5ème valeurs de la série, on tape :

```
>S[3:5]
```

Concaténation de 2 séries

Soit S1, et S2 deux séries. Si l'on veut "rassembler" les données en une série S3, (on dit *concaténer*) on tape :

```
>S3=c (S1, S2)
```

Concaténation de 2 séries

Soit S_1 , et S_2 deux séries. Si l'on veut "rassembler" les données en une série S_3 , (on dit *concaténer*) on tape :

```
>S3=c (S1, S2)
```

Essayer !

Opération sur les séries

Soit S une série, par exemple $S = (12, 10, 11, 19, 9, 10)$. Si l'on souhaite multiplier par 2 chaque valeurs puis addtionner 1 à chaque valeur, on tape alors :

$$>S = 2 * S + 1$$

Opération sur les séries

Soit S une série, par exemple $S = (12, 10, 11, 19, 9, 10)$. Si l'on souhaite multiplier par 2 chaque valeurs puis addtionner 1 à chaque valeur, on tape alors :

$$>S = 2 * S + 1$$

Ici on a écrasé les anciennes valeurs de S . Essayer

Opération sur les séries

Soit S une série, par exemple $S = (12, 10, 11, 19, 9, 10)$. Si l'on souhaite multiplier par 2 chaque valeurs puis addtionner 1 à chaque valeur, on tape alors :

$$>S = 2 * S + 1$$

Ici on a écrasé les anciennes valeurs de S . Essayer

Remplacement d'une valeur d'une série

Soit S une série, par exemple $S = (12, 10, 11, 19, 9, 10)$.

- Si l'on souhaite remplacer la 2ème valeur par un 5, on tape : `>S[2]=5`

Remplacement d'une valeur d'une série

Soit S une série, par exemple $S = (12, 10, 11, 19, 9, 10)$.

- Si l'on souhaite remplacer la 2ème valeur par un 5, on

tape : `>S[2]=5`

Essayer

Remplacement d'une valeur d'une série

Soit S une série, par exemple $S = (12, 10, 11, 19, 9, 10)$.

- Si l'on souhaite remplacer la 2ème valeur par un 5, on tape : `>S[2]=5`

Essayer

- Si l'on souhaite remplacer tous les 10 par 100, on tape : `>S[S==10]=100`

Remplacement d'une valeur d'une série

Soit S une série, par exemple $S = (12, 10, 11, 19, 9, 10)$.

- Si l'on souhaite remplacer la 2ème valeur par un 5, on tape : `>S[2]=5`

Essayer

- Si l'on souhaite remplacer tous les 10 par 100, on tape : `>S[S==10]=100`

Essayer

Remplacement d'une valeur d'une série

Soit S une série, par exemple $S = (12, 10, 11, 19, 9, 10)$.

- Si l'on souhaite remplacer la 2ème valeur par un 5, on tape : `>S[2]=5`

Essayer

- Si l'on souhaite remplacer tous les 10 par 100, on tape : `>S[S==10]=100`

Essayer

- Si l'on souhaite remplacer toutes les valeurs supérieures à 10 par 15, on tape : `>S[S>10]=15`

Remplacement d'une valeur d'une série

Soit S une série, par exemple $S = (12, 10, 11, 19, 9, 10)$.

- Si l'on souhaite remplacer la 2ème valeur par un 5, on tape : `>S[2]=5`

Essayer

- Si l'on souhaite remplacer tous les 10 par 100, on tape : `>S[S==10]=100`

Essayer

- Si l'on souhaite remplacer toutes les valeurs supérieures à 10 par 15, on tape : `>S[S>10]=15`

Essayer

Remplacement d'une valeur d'une série

Soit S une série, par exemple $S = (12, 10, 11, 19, 9, 10)$.

- Si l'on souhaite remplacer la 2ème valeur par un 5, on tape : `>S[2]=5`

Essayer

- Si l'on souhaite remplacer tous les 10 par 100, on tape : `>S[S==10]=100`

Essayer

- Si l'on souhaite remplacer toutes les valeurs supérieures à 10 par 15, on tape : `>S[S>10]=15`

Essayer

Tri

Soit S une série, par exemple $S = (12, 10, 11, 19, 9, 10)$.

- Si l'on souhaite trier les valeurs de la série par ordre croissant, on tape :

```
>sort(S)
```

Tri

Soit S une série, par exemple $S = (12, 10, 11, 19, 9, 10)$.

- Si l'on souhaite trier les valeurs de la série par ordre croissant, on tape :

```
>sort(S)
```

Essayer !

Tri

Soit S une série, par exemple $S = (12, 10, 11, 19, 9, 10)$.

- Si l'on souhaite trier les valeurs de la série par ordre croissant, on tape :

```
>sort(S)
```

Essayer !

- Essayer ensuite

```
>rev(sort(S))
```


Regroupement par caractère d'une série

La fonction `table` donne les effectifs de chaque modalité de la variable. Rentrer une série qualitative ou quantitative avec quelques redondances, notée `S`. Puis taper :

```
>table(S)
```

- 1 **Prise en main**
 - Introduction
 - Opérations élémentaires
 - Affectation
 - Structures de contrôle
 - **Autour des séries**
- 2 **Quelques fonctions liées aux statistiques descriptives**
 - Quantités cumulées
 - Calcul des paramètres standards
 - Représentations graphique
- 3 **Importer des données depuis un tableur**
- 4 **Exercices**
 - Exo 1
 - Exo 2
- 5 **Pour aller plus (ou moins) loin...**

- Exo 3
- Exo 4

Effectifs cumulés

Soit S une série. Pour obtenir la série des valeurs cumulées croissantes, on tape :

```
>cumsum(S)
```

Effectifs cumulés

Soit S une série. Pour obtenir la série des valeurs cumulées croissantes, on tape :

```
>cumsum(S)
```

Rentrer une série S , puis taper `>cumsum(S)`.

Effectifs cumulés

Soit S une série. Pour obtenir la série des valeurs cumulées croissantes, on tape :

```
>cumsum(S)
```

Rentrer une série S , puis taper `>cumsum(S)`.

Question : Soit S une série. Comment obtenir les effectifs cumulés croissants ?

Effectifs cumulés

Question : Soit S une série. Comment obtenir les effectifs cumulés croissants ?

Effectifs cumulés

Question : Soit S une série. Comment obtenir les effectifs cumulés croissants ?

Prenons par exemple $S = (12, 10, 10, 12, 11, 19, 9, 10)$.

Effectifs cumulés

Question : Soit S une série. Comment obtenir les effectifs cumulés croissants ?

Prenons par exemple $S = (12, 10, 10, 12, 11, 19, 9, 10)$. Taper
>EC=cumsum (table (S))

Effectifs cumulés

Question : Soit S une série. Comment obtenir les effectifs cumulés croissants ?

Prenons par exemple $S = (12, 10, 10, 12, 11, 19, 9, 10)$. Taper
`>EC=cumsum(table(S))` Demander `>EC`, vous obtenez :

```
> EC
9  10  11  12  19
1  4  5  7  8
```

Effectifs cumulés

Question : Soit S une série. Comment obtenir les effectifs cumulés croissants ?

Prenons par exemple $S = (12, 10, 10, 12, 11, 19, 9, 10)$. Taper
`>EC=cumsum(table(S))` Demander `>EC`, vous obtenez :

```
> EC
9  10  11  12  19
1  4   5   7   8
```

Ce qui correspond bien aux effectifs cumulés croissants de la série S regroupée par "valeur" :

```
> table(S)
9  10  11  12  19
1  3   1   2   1
```

Fréquences cumulées

Question : Soit S une série. Fabriquez les fréquences cumulées croissantes ?

Fréquences cumulées

Question : Soit S une série. Fabriquez les fréquences cumulées croissantes ?

Prenons la même série $S = (12, 10, 10, 12, 11, 19, 9, 10)$.

Fréquences cumulées

Question : Soit S une série. Fabriquez les fréquences cumulées croissantes ?

Prenons la même série $S = (12, 10, 10, 12, 11, 19, 9, 10)$. Taper
>`FC=cumsum (table (S)) / 8`

Fréquences cumulées

Question : Soit S une série. Fabriquez les fréquences cumulées croissantes ?

Prenons la même série $S = (12, 10, 10, 12, 11, 19, 9, 10)$. Taper `>FC=cumsum(table(S))/8` Demander `>FC`, vous obtenez :

```
> FC
  9      10      11      12      19
0.125  0.500  0.625  0.875  1.000
```

Max, Min

Rentrer une série, que l'on nommera "notes". Par exemple vos notes en maths !

Max, Min

Rentrer une série, que l'on nommera "notes". Par exemple vos notes en maths !

Taper `>max (notes)` puis `>min (notes)`

Max, Min

Rentrer une série, que l'on nommera "notes". Par exemple vos notes en maths !

Taper `>max(notes)` puis `>min(notes)`

Fabriquer alors l'étendue.

Moyenne

- Calculer à la main la moyenne de votre série notes

Moyenne

- Calculer à la main la moyenne de votre série notes
- Taper `>sum (notes)`, puis `>length (notes)` et enfin `>sum (notes) /length (notes)`

Moyenne

- Calculer à la main la moyenne de votre série notes
- Taper `>sum (notes)` , puis `>length (notes)` et enfin `>sum (notes) /length (notes)`
- Taper `>mean (notes)`

Moyenne

- Calculer à la main la moyenne de votre série notes
- Taper `>sum (notes)` , puis `>length (notes)` et enfin `>sum (notes) /length (notes)`
- Taper `>mean (notes)`
- Conclusion ?

Moyenne avec données manquantes

Moyenne avec données manquantes

Parfois, certaines données manquantes d'une série sont rentrées comme "NA". Pour calculer la moyenne en occultant ces données, on tape :

```
>mean(notes, na.rm = TRUE)
```


Moyenne avec données manquantes

Parfois, certaines données manquantes d'une série sont rentrées comme "NA". Pour calculer la moyenne en occultant ces données, on tape :

```
>mean(notes, na.rm = TRUE)
```

Essayer en rentrant une série contenant des NA.

Moyenne de zone de série

- Si l'on ne veut calculer la moyenne que des 3 premières notes, on tapera :

```
>mean (notes [1:3] )
```

- Exercices : Calculer la moyenne des 5 notes les plus faibles.

Moyenne de zone de série

- Si l'on ne veut calculer la moyenne que des 3 premières notes, on tapera :

```
>mean (notes [1 : 3] )
```

- Exercices : Calculer la moyenne des 5 notes les plus faibles.

Indication : utiliser `mean(sort(notes)[1 :5])`

Moyenne pondérée

Supposons que l'on veuille calculer la moyenne de la série suivante :

Caractère	3	4	9	15
Effectif	1	2	3	2

Moyenne pondérée

Supposons que l'on veuille calculer la moyenne de la série suivante :

Caractère	3	4	9	15
Effectif	1	2	3	2

- On rentre d'abord la série des caractères distincts et la série des effectifs.

```
>S=c(3, 4, 9, 15)
```

```
>Eff=c(1, 2, 3, 2)
```

Moyenne pondérée

Supposons que l'on veuille calculer la moyenne de la série suivante :

Caractère	3	4	9	15
Effectif	1	2	3	2

- On rentre d'abord la série des caractères distincts et la série des effectifs.

```
>S=c(3, 4, 9, 15)
```

```
>Eff=c(1, 2, 3, 2)
```

- Puis, on tape

```
>weighted.mean(S, Eff)
```

Médiane

Taper

```
>median(notes)
```

Quantiles

Taper

```
>quantile(notes)
```


Quantiles

Taper

```
>quantile(notes)
```

- Pour obtenir uniquement le 1er quantile, on tape :

```
>quantile(notes, probs = .25)
```

Quantiles

Taper

```
>quantile(notes)
```

- Pour obtenir uniquement le 1er quantile, on tape :

```
>quantile(notes, probs = .25)
```

- Et si l'on veut uniquement la valeur, on tape :

```
> quantile(notes, probs = .25, names =  
FALSE)
```

Quantiles

Taper

```
>quantile(notes)
```

- Pour obtenir uniquement le 1er quantile, on tape :

```
>quantile(notes, probs = .25)
```

- Et si l'on veut uniquement la valeur, on tape :

```
> quantile(notes, probs = .25, names =  
FALSE)
```

Exercice : Fabriquer l'écart inter quantile

Déciles

Taper

```
> quantile(notes, probs = .1)
```

Déciles

Taper

```
> quantile(notes, probs = .1)
```

Pour obtenir tous les déciles, on tape :

```
>quantile(notes, probs = 0:10/10)
```

Déciles

Taper

```
> quantile(notes, probs = .1)
```

Pour obtenir tous les déciles, on tape :

```
>quantile(notes, probs = 0:10/10)
```

Taper `>0:10/10` pour comprendre...

Variance

- Taper

```
>var (notes)
```

Variance

- Taper

```
>var (notes)
```

- Taper `>mean ((notes-mean (notes)) ^2)`

Variance

- Taper

```
>var (notes)
```

- Taper `>mean((notes-mean(notes))^2)`

Comparer vos résultats ? Explication ?

Variance, explication

- La fonction `>var (notes)` de R, renvoie l'expression suivante :

$$\frac{1}{n-1} \sum_i n_i [x_i - \mathbb{E}(X)]^2$$

Variance, explication

- La fonction `>var (notes)` de R, renvoie l'expression suivante :

$$\frac{1}{n-1} \sum_i n_i [x_i - \mathbb{E}(X)]^2$$

- alors que la "vraie" variance correspond à :

$$\frac{1}{n} \sum_i n_i [x_i - \mathbb{E}(X)]^2$$

Variance, explication

- La fonction `>var (notes)` de R, renvoie l'expression suivante :

$$\frac{1}{n-1} \sum_i n_i [x_i - \mathbb{E}(X)]^2$$

- alors que la "vraie" variance correspond à :

$$\frac{1}{n} \sum_i n_i [x_i - \mathbb{E}(X)]^2$$

⇒ Ainsi pour obtenir la "vraie" variance, vous avez 2 options :

-Taper `>mean ((notes-mean (note)) ^2)`

-Taper

`> ((length (notes) - 1) / length (notes)) * var (notes)`

Variance, explication

Remarques :

- la variance "modifiée" de R, est sans biais et sera utile dans les estimations et tests statistiques.

Variance, explication

Remarques :

- la variance "modifiée" de R, est sans biais et sera utile dans les estimations et tests statistiques.
- Lorsqu'on veut calculer la variance d'une série où les caractères sont pondérées par un effectif, on peut effectuer la même adaptation avec la commande

```
weighted.mean()
```

Variance, explication

Remarques :

- la variance "modifiée" de R, est sans biais et sera utile dans les estimations et tests statistiques.
- Lorsqu'on veut calculer la variance d'une série où les caractères sont pondérées par un effectif, on peut effectuer la même adaptation avec la commande

`weighted.mean()` Exemple : pour la série

Caractère	3	4	9	15
Effectif	1	2	3	2

Variance, explication

Remarques :

- la variance "modifiée" de R, est sans biais et sera utile dans les estimations et tests statistiques.
- Lorsqu'on veut calculer la variance d'une série où les caractères sont pondérées par un effectif, on peut effectuer la même adaptation avec la commande

`weighted.mean()` Exemple : pour la série

Caractère	3	4	9	15
Effectif	1	2	3	2

- On rentre d'abord la série des caractères distincts et la série des effectifs.

```
>S=c(3,4,9,15)
```

```
>Eff=c(1,2,3,2)
```


Variance, explication

Remarques :

- la variance "modifiée" de R, est sans biais et sera utile dans les estimations et tests statistiques.
- Lorsqu'on veut calculer la variance d'une série où les caractères sont pondérées par un effectif, on peut effectuer la même adaptation avec la commande

`weighted.mean()` Exemple : pour la série

Caractère	3	4	9	15
Effectif	1	2	3	2

- ▶ On rentre d'abord la série des caractères distincts et la série des effectifs.

```
>S=c(3,4,9,15)
```

```
>Eff=c(1,2,3,2)
```

- ▶ Puis, on calcule la moyenne notée m par :

```
>m=weighted.mean(S, Eff)
```

Variance, explication

Remarques :

- la variance "modifiée" de R, est sans biais et sera utile dans les estimations et tests statistiques.
- Lorsqu'on veut calculer la variance d'une série où les caractères sont pondérées par un effectif, on peut effectuer la même adaptation avec la commande

`weighted.mean()` Exemple : pour la série

Caractère	3	4	9	15
Effectif	1	2	3	2

- On rentre d'abord la série des caractères distincts et la série des effectifs.

```
>S=c(3,4,9,15)
```

```
>Eff=c(1,2,3,2)
```

- Puis, on calcule la moyenne notée m par :

```
>m=weighted.mean(S, Eff)
```

- Et enfin, on tape :

```
>weighted.mean( (S-m)^2, Eff)
```

Ecart type

```
Taper >sd (notes)
```

Ecart type

Taper `>sd (notes)`

Là encore, pour obtenir le "vrai" écart-type, vous pouvez :

-soit le calculer manuellement en prenant la racine carré de la "vraie" variance

-soit taper

```
>sqrt ( (length (notes) -1) /length (notes) )  
*sd (notes)
```

Paramètres en un seul clic

Taper

```
>summary (notes)
```

- 1 **Prise en main**
 - Introduction
 - Opérations élémentaires
 - Affectation
 - Structures de contrôle
 - Autour des séries
- 2 **Quelques fonctions liées aux statistiques descriptives**
 - Quantités cumulées
 - Calcul des paramètres standards
 - **Représentations graphique**
- 3 **Importer des données depuis un tableur**
- 4 **Exercices**
 - Exo 1
 - Exo 2
- 5 **Pour aller plus (ou moins) loin...**

- Exo 3
- Exo 4

Graphe basique

Rentrer une série S de quelques valeurs, puis Taper

```
>plot(S)
```


Diagramme en bâtons

Rentrer une série S de quelques valeurs (avec des redondances), puis Taper

```
>plot(table(S))
```

Diagramme en bâtons

Rentrer une série S de quelques valeurs (avec des redondances), puis Taper

```
>plot(table(S))
```

On peut améliorer la présentation en tapant :

```
>plot(table(S), main="titre", ylab="axe des  
y")
```

histogramme

- Avec la même série S de quelques valeurs, taper :

```
>hist(S)
```

histogramme

- Avec la même série S de quelques valeurs, taper :

```
>hist(S)
```

On constate que les classes sont imposées...

histogramme

- Avec la même série S de quelques valeurs, taper :

```
>hist(S)
```

On constate que les classes sont imposées...

- Pour choisir les classes, on tape :

```
>hist(S,breaks=c(0,4,15,20))
```

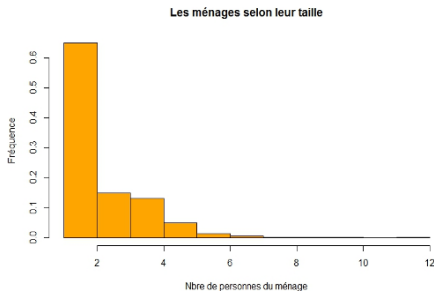
si l'on souhaite que les bornes des classes soient 0, 4, 15 et 20.

Histogramme

Fonction **hist** pour les histogrammes :

```
hist(RCity$TAILLEMENAGE, main="Les ménages selon leur taille",  
     xlab="Nbre de personnes du ménage", ylab="Fréquence",  
     probability=TRUE, col="orange")
```

probability=TRUE permet de
représenter les fréquences



Camembert

Soit S une série qualitative, on tape :

```
>pie(table(S))
```

Boxplot-Boîte à moustache

- Avec la même série S de quelques valeurs, taper :
`>boxplot (S)`

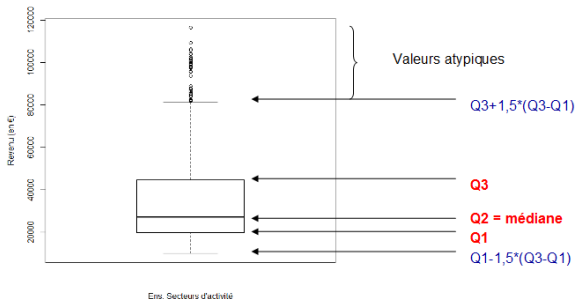
Boxplot-Boîte à moustache

- Avec la même série S de quelques valeurs, taper :
`>boxplot(S)`
- Pour obtenir la boxplot horizontalement, taper :
`>boxplot(S, horizontal=TRUE)`

Boxplot-Boîte à moustache

Fonction **boxplot** (indicateurs de distribution) :

`boxplot (RCity$SALAIRE, xlab="Ens. Secteurs d'activité",
ylab="Revenu(en €)")`



- 1 **Prise en main**
 - Introduction
 - Opérations élémentaires
 - Affectation
 - Structures de contrôle
 - Autour des séries
- 2 **Quelques fonctions liées aux statistiques descriptives**
 - Quantités cumulées
 - Calcul des paramètres standards
 - Représentations graphique
- 3 **Importer des données depuis un tableur**
- 4 **Exercices**
 - Exo 1
 - Exo 2
- 5 **Pour aller plus (ou moins) loin...**

- Exo 3
- Exo 4

Il est assez courant de vouloir importer des données saisies avec un tableur du type Open Office ou Excel. Rstudio permet d'importer (entre autres) :

Il est assez courant de vouloir importer des données saisies avec un tableur du type Open Office ou Excel. Rstudio permet d'importer (entre autres) :

- des fichiers .csv
- des fichiers .xlsx

Il est assez courant de vouloir importer des données saisies avec un tableur du type Open Office ou Excel. Rstudio permet d'importer (entre autres) :

- des fichiers .csv
- des fichiers .xlsx

Il suffit d'aller dans la fenêtre en bas à droite de Rstudio, de sélectionner l'emplacement du fichier de données. Puis par un clic gauche sur la souris, sélectionner sur "import dataset".

Il est assez courant de vouloir importer des données saisies avec un tableur du type Open Office ou Excel. Rstudio permet d'importer (entre autres) :

- des fichiers .csv
- des fichiers .xlsx

Il suffit d'aller dans la fenêtre en bas à droite de Rstudio, de sélectionner l'emplacement du fichier de données. Puis par un clic gauche sur la souris, sélectionner sur "import dataset". Le tableau de données apparaît alors dans la fenêtre en haut à gauche de Rstudio.

- On peut visualiser ce tableau de données en tapant

```
>View(nom du fichier)
```

- On peut visualiser ce tableau de données en tapant

```
>View(nom du fichier)
```

- Si l'on souhaite accéder seulement aux effectifs associé à un caractère X (une colonne), on tapera

```
>nom du fichier$X
```

- 1 **Prise en main**
 - Introduction
 - Opérations élémentaires
 - Affectation
 - Structures de contrôle
 - Autour des séries
- 2 **Quelques fonctions liées aux statistiques descriptives**
 - Quantités cumulées
 - Calcul des paramètres standards
 - Représentations graphique
- 3 **Importer des données depuis un tableur**
- 4 **Exercices**
 - Exo 1
 - Exo 2
- 5 **Pour aller plus (ou moins) loin...**

● Exo 3

● Exo 4

Exo 1

A l'aide du fichier `formulaire.xlsx` ou `csv` (formulaire sur les moyens de transport) à télécharger sur la page web, répondre aux questions suivantes.

- 1 Faire apparaître ce tableau de données dans Rstudio.
- 2 Combien de personnes utilisent le bus comme moyen de transport ?
- 3 Quel est le budget moyen par mois pour le transport ?
Calculer l'écart type. Tracer la boxplot relative au budget.
- 4 Après avoir déterminer des classes raisonnables concernant le budget, tracer l' histogramme des effectifs associés à vos classes.
- 5 Tracer le diagramme en bâton relatif aux marques de véhicules, puis tracer le camembert associé. (commencer par faire un regroupement par marque.)

Exo 1 suite

- 6 Calculer les pourcentages de chaque marque de véhicules.
- 7 Combien de personnes ont une voiture personnelle ?
- 8 Calculer les quantiles des distances parcourues par les personnes de cette étude. Tracer la boxplot.
- 9 Combien de personnes utilisent uniquement le bus ?

Exo 1 indication pour question 7

On pourra taper :

Exo 1 indication pour question 7

On pourra taper :

```
>b=0  
> for (i in 1:length(Formulaire$VOITUREPERSO))  
  (if (Formulaire$VOITUREPERSO[i]=="Oui") b=b+1)  
> b  
[1] 27
```


Exo 2

On a relevé les temps de retard (en min) des employés dans une entreprise :

5, 7, 12, 1, 5, 9, 4, 0, 0, 0, 14, 4, 22, 2, 6, 10, 5

Exo 2

On a relevé les temps de retard (en min) des employés dans une entreprise :

5, 7, 12, 1, 5, 9, 4, 0, 0, 0, 14, 4, 22, 2, 6, 10, 5

- 1 Calculer la moyenne, l'étendue, la variance, l'écart type
- 2 Calculer les quantiles, ainsi que l'écart inter quantile.
- 3 Donner la valeur du 3ième décile.
- 4 Représenter la boxplot
- 5 Tracer un diagramme en batons.
- 6 Y a t'il une (ou des) valeur(s) aberrante(s) ?
(Rappel : une valeur est dite aberrante si elle est supérieure à $Q_3 + 1,5(Q_3 - Q_1)$ ou inférieure à $Q_1 - 1,5(Q_3 - Q_1)$)

- 1 **Prise en main**
 - Introduction
 - Opérations élémentaires
 - Affectation
 - Structures de contrôle
 - Autour des séries
- 2 **Quelques fonctions liées aux statistiques descriptives**
 - Quantités cumulées
 - Calcul des paramètres standards
 - Représentations graphique
- 3 **Importer des données depuis un tableur**
- 4 **Exercices**
 - Exo 1
 - Exo 2
- 5 **Pour aller plus (ou moins) loin...**
 - Exo 3

Pour définir une fonction en R

Supposons que l'on veuille définir la fonction f définie par

$$f(x) = x^2 - 5x + 1.$$

On peut taper en R :

```
f<-function(x) { x^2-5*x +1 }
```

Pour définir une fonction en R

Supposons que l'on veuille définir la fonction f définie par

$$f(x) = x^2 - 5x + 1.$$

On peut taper en R :

```
f<-function(x) { x^2-5*x +1}
```

Si l'on veut avoir la valeur de $f(1)$. Il suffit de taper :

```
f(1)
```

```
[1] -3
```

Exo 3

- 1 Programmer la fonction factorielle avec une boucle "while".
On la nommera *fact*
- 2 Programmer la fonction factorielle de manière récursive.
On l'appelera *factrec*.
- 3 A l'aide d'une des 2 fonctions, en déduire un moyen
d'afficher $\binom{n}{k}$. Tester en calculant par ex $\binom{5}{2}$.

Éléments de réponse (exo 3)

En utilisant "while"

```
fact<-function(x) { res=1 ;  
while(x>1) { res=res * x; x= x-1 }  
res }
```

Eléments de réponse (exo 3)

En utilisant "while"

```
fact<-function(x) { res=1 ;  
while(x>1) { res=res * x; x= x-1 }  
res }
```

Vérifier la valeur de *fact*(5)

Éléments de réponse (exo 3)

Forme récursive

```
factrec<-function(x){tmp<-1  
if(x>0){ tmp<-factrec(x-1)*x }  
return(tmp)}
```

Eléments de réponse (exo 3)

Forme récursive

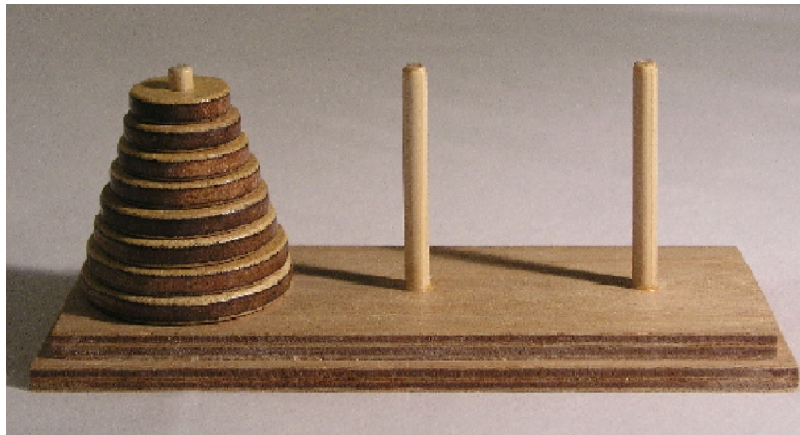
```
factrec<-function(x){tmp<-1  
if(x>0){ tmp<-factrec(x-1)*x }  
return(tmp)}
```

Vérifier la valeur de *factrec*(5)

Calcul des coefficients binomiaux

```
n=6; k=3;  
print(paste(k, "parmi", n, "vaut"));  
fact(n) / (fact(k) * fact(n-k))
```

Tours de Hanoi



Tours de Hanoi



Le jeu consiste à déplacer les n palets de la tour 1 sur la tour 3. Le jeu n'aurait évidemment guère d'intérêt si on ne rajoutait pas la contrainte suivante : à chaque étape, aucun palet ne doit être posé sur un palet de diamètre inférieur.

Tours de Hanoi



Le jeu consiste à déplacer les n palets de la tour 1 sur la tour 3. Le jeu n'aurait évidemment guère d'intérêt si on ne rajoutait pas la contrainte suivante : à chaque étape, aucun palet ne doit être posé sur un palet de diamètre inférieur.

Ecrire une fonction `hanoi(n, a, b, c)` qui prend comme arguments : n (nombre de palets sur la tour 1), et a, b, c qui sont les noms de tours, et qui renvoie la suite des instructions pour résoudre le problème.

Eléments de réponse : Tours de Hanoi

```
hanoi<-function(n,a,b,c){ if (n==3)
{print(paste("mettre palet de la tour", a, "a la tour",c))
print(paste("mettre palet de la tour", a, "à la tour",b))
print(paste("mettre palet de la tour", c, "à la tour",b))
print(paste("mettre palet de la tour", a, "à la tour",c))
print(paste("mettre palet de la tour", b, "à la tour",a))
print(paste("mettre palet de la tour", b, "à la tour",c))
print(paste("mettre palet de la tour", a, "à la tour",c))
} else hanoi(n-1,a,c,b)
print(paste("mettre palet de la tour", a, "à la tour",c))
hanoi(n-1,b,a,c)
}
```

Eléments de réponse : Tours de Hanoi

```

hanoi<-function(n,a,b,c){ if (n==3)
{print(paste("mettre palet de la tour", a, "a la tour",c))
print(paste("mettre palet de la tour", a, "à la tour",b))
print(paste("mettre palet de la tour", c, "à la tour",b))
print(paste("mettre palet de la tour", a, "à la tour",c))
print(paste("mettre palet de la tour", b, "à la tour",a))
print(paste("mettre palet de la tour", b, "à la tour",c))
print(paste("mettre palet de la tour", a, "à la tour",c))
} else hanoi(n-1,a,c,b)
print(paste("mettre palet de la tour", a, "à la tour",c))
hanoi(n-1,b,a,c)
}
Tester avec hanoi(3,1,2,3)

```



```
>hanoi(4,1,2,3)
[1] "mettre palet de la tour 1 à la tour 2"
[1] "mettre palet de la tour 1 à la tour 3"
[1] "mettre palet de la tour 2 à la tour 3"
[1] "mettre palet de la tour 1 à la tour 2"
[1] "mettre palet de la tour 3 à la tour 1"
[1] "mettre palet de la tour 3 à la tour 2"
[1] "mettre palet de la tour 1 à la tour 2"
[1] "mettre palet de la tour 1 à la tour 3"
[1] "mettre palet de la tour 2 à la tour 3"
[1] "mettre palet de la tour 2 à la tour 1"
[1] "mettre palet de la tour 3 à la tour 1"
[1] "mettre palet de la tour 2 à la tour 3"
[1] "mettre palet de la tour 1 à la tour 2"
[1] "mettre palet de la tour 1 à la tour 3"
[1] "mettre palet de la tour 2 à la tour 3"
```