

### 1. Premières notions.

- L'environnement R est un *interpréteur* qui lit et évalue des expressions.
- Création et affectation d'un objet : <- ou =
- Séparateur d'instructions : ; Inutile en fin de ligne s'il n'y a qu'une instruction.
- R distingue minuscules et majuscules.
- `ls()` affiche la liste des objets créés et installés dans l'environnement courant.

### 2. L'aide.

Le logiciel possède un système d'aide en ligne. On peut y accéder de multiples manières.

- L'instruction `?identificateur` ; affiche la documentation associée à l'objet `identificateur`. Par exemple la commande  
`> ?mean`

affiche la documentation associée à la fonction `mean` qui calcule la moyenne (éventuellement censurée) d'un échantillon.

- `help(identificateur)` fait exactement la même chose.
- `help.start()` permet de naviguer sur l'aide en ligne officielle.
- `apropos("mot")` affiche la liste des fonctions et objets qui contiennent le mot-clé `mot` dans leur nom.

### 3. Premiers essais.

Entrer successivement les commandes

```
> ? mean
> help(mean)
> help.start()
> apropos("mean")
```

Sachant qu'en anglais écart-type se dit *standard deviation* trouver grâce à l'aide la commande qui fournit l'écart-type.

### 4. Création de vecteurs.

Entrer successivement les commandes

```
> a = c(1,2,4,8,16,32)
> a
> b = c(3.14, 2.71, 1.414, 1.732)
> b
> c = c(2,4, 3.14)
> c
> ls()
```

On peut penser que `c` est l'initiale du mot *colonne*.

On peut aussi créer des vecteurs progressions arithmétiques :

```
> vpa = seq(5,15,0.3)
> vpa
> length(vpa)
> ls()
```

### 5. Concaténation de vecteurs.

Entrer successivement les commandes

```
> d = c(a,b,c)
> d
> length(a); length(b); length(c); length(d)
> ls()
```

### 6. Paramètres d'une série statistique.

Entrer successivement les commandes

```
> x = d
> mean(x)
> median(x)
> var(x)
> "écart-type"(x)
> summary(x)
> fivenum(x)
```

Comment est définie la médiane ?

Vérifier dans le gestionnaire de packages que le Package "datasets" est chargé. Entrer successivement les commandes

```
> airmiles
> ?airmiles
> lynx
> ?lynx
> LakeHuron
> ?LakeHuron
> min(LakeHuron)
> max(LakeHuron)
```

Se renseigner sur le fonctionnement de la fonction quantile  $Q$  et trouver  $Q(0.9)$  pour la série Lake Huron. *Interprétation.* C'est la hauteur qui n'est dépassée qu'une fois tous les dix ans en moyenne.

### 7. Histogrammes.

Entrer successivement les commandes

```
> hist(LakeHuron)
> hist(LakeHuron,60)
> hist(LakeHuron,30)
> hist(LakeHuron,3)
> coupures = c(min(LakeHuron),578,580,max(LakeHuron))
> hist(LakeHuron,coupures)
```

Quelle est la méthode par défaut utilisée par la commande `hist` pour choisir le nombre de cellules de l'histogramme ?

Essayer la méthode de Freedman-Diaconis. Rappeler le principe des deux méthodes.

Trouver la commande qui affiche le graphe de la fonction de répartition  $F$  d'une série et celle qui affiche le graphe de la fonction quantile  $Q$ .

Entrer successivement les commandes

```
> plot(ecdf(LakeHuron))
> ecdf(LakeHuron)
```

Quelle est la signification de `ecdf` ?

### 8. Boîtes à moustaches.

Entrer successivement les commandes `boxplot(LakeHuron)`, `boxplot(airlines)`, `boxplot(lynx)`.

Il y a plusieurs façons de définir la boîte à moustaches. Quelle est la méthode utilisée par R ?

### 9. Représentation d'une suite.

Examiner la suite statistique (= le vecteur) `lynx`. Il y a de très fortes oscillations d'une période de dix ans environ.

Trouver le moyen de représenter graphiquement la suite comme une chaîne de points ( $i$ , l'année, en abscisse, `lynx[i]`, le nombre de prises de l'année  $i$ , en ordonnée), puis comme la ligne brisée qu'on obtient en joignant ces points par des segments.

