



Initiation à l'analyse statistique des données expérimentales

Christophe Pallier et Christophe Lalanne

`christophe.pallier@m4x.org / christophe.lalanne@gmx.net`

Mastère de Sciences Cognitives, EHESS – Paris 5 – ENS



But du cours et Méthode

- ▶ Amener les étudiants au point d'être capable de réaliser des analyses de données simples (comparaisons de moyennes, détection de corrélation,...) en comprenant ce qu'ils font.
- ▶ Enseigner en parallèle les concepts théoriques et la pratique d'un logiciel de statistique ("R")
- ▶ Transmettre de "bonnes habitudes": exploration graphique et descriptive des données.
- ▶ Travailler avec des jeux de données simulés et pas seulement des données réelles



Ressources

- ▶ Documents de tp: [tp1.pdf](#) et [tp2.pdf](#)
- ▶ Livres conseillés:
 - ▶ Howell, D.C. (1998). *Méthodes Statistiques en Sciences Humaines*. De Boeck Université.
 - ▶ Crawley, M.J. (2005). *Statistics – An Introduction using R*. Wiley.
- ▶ Sur le site de R (www.r-project.org) voir la rubrique 'Documentation/Other/Contributed'. A signaler:
 - ▶ *R pour les débutants* ([pdf](#))
 - ▶ *Notes on the use of R for psychology experiments and questionnaires* ([pdf](#))



Choix du logiciel de statistique "R"

- ▶ Avantages:
 - ▶ logiciel très flexible et très puissant. Permettant de réaliser de une analyse de données de A à Z:
 1. manipulations et transformation des données brutes
 2. analyses statistique
 3. génération de graphiques prêts à publiés.
 - ▶ logiciel libre (et donc gratuit...)
 - ▶ multiplateformes (Linux, MacOSX, Windows,...)
 - ▶ de nombreuses documentations existent
- ▶ Particularité: R est un *langage*, ce qui rend son apprentissage plus difficile qu'un logiciel "cliquodrôme".



Démarrage de R, Accéder à l'aide

- ▶ Ouvrir un terminal et entrer “R”.
L'interaction avec R se fait en tapant des *commandes*.
- ▶ Entrer les commandes suivantes:

```
demo(graphics)
```

```
demo(persp)
```

```
help.start()
```

```
help.search('Student')
```

```
?t.test
```

```
q()
```



Utiliser R comme une calculatrice

▶ opérations arithmétiques

```
3.14*5
```

```
1.05^6
```

```
(1+1i) * (5+6i)
```

▶ fonctions

```
sin(pi/4)
```

```
log10(20)
```

```
lfactorial(100)
```

▶ variables

```
a <- 34
```

```
a
```

```
a + 1/a
```



Le vecteur, objet de base de R

```
b <- c(1,10,20,50) # assignment
b
1+b               # opérations
2*b
a <- c(0,1,0,1)
a+b
a %*% b           # produit scalaire

sqrt(b)          # fonctions
log(b)
mean(b)
plot(b)
plot(b,type='l')
```



Créer des vecteurs

```
dat <- scan('')
```

```
dat
```

```
(dat<-c(alpha=1,beta=3,delta=8))
```

```
dat <- runif(100,min=0,max=10)
```

```
seq(1,10)      # séries
```

```
1:10
```

```
seq(1,10,by=2)
```

```
c(b,b)         # concaténation
```

```
c(b,0,b,a)
```

```
rep(b,10)      # répétitions
```

```
rep(a,c(5,4,3,5))
```

```
gl(3,10)
```



Exercices

- ▶ Calculer la somme des nombres de 1 à 100
- ▶ Calculer le produit des nombres de 1 à 100
- ▶ Créer un vecteur contenant 10 '1' suivi de 10 '2'
- ▶ Créer un vecteur contenant 10 fois la série de nombre 1 à 10
- ▶ Créer un vecteur contenant 1 '1', puis 2 '2', puis 3 '3', jusqu'à 20.
- ▶ Afficher la fonction log de factoriel sur l'intervalle [1,100]
- ▶ Afficher la fonction $\sin(x)/x$ sur l'intervalle $[-6\pi, 6\pi]$



Quitter R, Workspace

- ▶ Quitter R: `q()`
- ▶ Redémarrer R, puis entrer:

```
ls()
```

```
history()
```

- ▶ Utiliser un éditeur de texte pour sauver les commandes suivantes dans un fichier 'monscript.R'

```
# monscript.R
```

```
a<-scan('http://www.pallier.org/misc/a.dat')
```

```
max(a)
```

Puis exécuter dans R:

```
source('monscript.R',echo=T)
```

Et, en ligne de commande unix:

```
R -q --vanilla <monscript.R
```