

TP Analyse de données

Les TP utilisent le logiciel gratuit R. Voir

<http://www.r-project.org/>

<http://cran.r-project.org/>

Séance 1 – Introduction à R

R est en logiciel basé sur des commandes en ligne. Voici quelques commandes utiles: `q()`, `help(nom de commande)` `demo()`, `data()`

1. Exécuter et comprendre le sens des commandes suivantes:

```
x<-c(3.7, 2.9, 10.8, 3.1, 7.5, 9.2, 1.8)
x
x[2:4]
y<-sort(x)
y
sort(x)->z
z
2*x+3
x^2
hist(x)
mean(x)
median(x)
var(x)
summary(x)
cov(x,y)
list()
```

2. Créer deux vecteurs de données v et w , et vérifier la relation $\text{var}(v + w) = \text{var}(v) + \text{var}(w) + 2 \text{cov}(v, w)$.
3. La commande suivante génère une simulation de 100 variables aléatoires binomiales, correspondant au nombre de pile parmi 10 jets de pièce: `b<-binom(100,10,0.5)`
Créer plusieurs échantillons de 100, 1000 et 10000 jets de pièce, et déterminer leurs moyennes, variances et histogrammes.