
TP : Introduction   R

Objectif :

- prendre en main les concepts de base du logiciel R
- r aliser des ANOVA

1 Introduction

R¹ est un logiciel de statistiques gratuit et open-source, disponible sous Linux, Windows et OS X.

Pour lancer R, tapez **R** en ligne de commande. Il est  galement possible d' crire des scripts et de les interpr ter avec R. Pour obtenir de l'aide en ligne sur une commande, il suffit de la faire pr c der du symbole `?`. Les fl ches haut et bas permettent de naviguer dans l'historique de commandes. Les commentaires sont pr c d s par le symbole `#`. La touche `tab` peut  tre utilis e pour la compl tion des commandes et des noms de variables. Il est recommand  de travail avec l'IDE RStudio et d'utiliser la syntaxe Markdown (fichiers avec l'extension `Rmd`).

2 Premiers pas

Le moyen le plus rapide d'entrer un faible nombre de donn es dans R est la fonction `c` (*combine operator*). Ainsi la commande suivante permet d'affecter un ensemble de r els   la variable `v`. Pour visualiser le contenu d'une variable, il suffit de taper son nom. Dans le cas pr sent le `[1]` qui est affich  indique que la variable est de type vecteur. Calculer la moyenne, m diane, variance ... de ces valeurs s'obtient simplement en tapant le nom de la fonction correspondante en passant le nom de la variable en param tre. Notez que le symbole `=` est utilis  pour l'affectation mais il est  galement possible d'utiliser `<-`.

Modifier le contenu d'un vecteur peut se faire en indiquant l'indice de l' l ment   modifier entre crochets. De nombreuses op rations sont possibles sur les vecteurs comme d terminer la somme des  l ments dont la valeur est sup rieure   10. Les autres types de donn es fr quemment manipul s sont les matrices, les listes et les data frames.

```
> v = c(12, .4, 5, 2, 50, 8, 3, 1, 4, .25) # data
> v
[1] 12.00  0.40  5.00  2.00 50.00  8.00  3.00  1.00  4.00  0.25
> mean(v)
[1] 8.565
> v[2]=7
> v
[1] 12.00  7.00  5.00  2.00 50.00  8.00  3.00  1.00  4.00  0.25
> sum(v[v>10])
[1] 62
```

Q 1. Cherchez comment calculer le ni me percentile d'un vecteur. Calculez le 90e percentile de `v`.

Toutes ces commandes peuvent  tre enregistr es dans un fichier qui peut  tre interpr t  en sp cifiant son chemin avec la commande `setwd` et lanc  avec la commande `source` (dans le cas de l'utilisation en ligne de commande).

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

3 Graphiques

Q 2. Testez les commandes **boxplot** et **barplot** pour représenter graphiquement les données d'un vecteur. Il est possible de mettre plusieurs vecteurs en paramètre. Dans ce cas, chacun des vecteurs sera représenté.

4 Analyses statistiques sur l'article ForceEdge

L'objectif est de refaire une partie des analyses statistiques détaillées dans l'article suivant :

Axel Antoine, Sylvain Malacria, and G ry Casiez. 2017. ForceEdge: Controlling Autoscroll on Both Desktop and Mobile Computers Using the Force². In Proceedings of the 2017 CHI Conference on Human Factors in Computing Systems (CHI '17). ACM, New York, NY, USA, 3281-3292.

Nous allons nous concentrer uniquement sur la premi re exp rience. Commencez par lire le protocole exp rimental.

Q 3. Commencez par faire fonctionner les exemples du cours³ et lisez attentivement le code.

Q 4. Reproduisez les graphiques de la figure 4 (le point important est d'afficher les intervalles de confiance   95%).

Q 5. Refaites les analyses statistiques sur le temps.

Q 6. Refaites les analyses sur les donn es NASA-TLX.

5 Analyses de vos propres donn es

Si vous avez d j  r alis  des exp riences contr l es et que vous avez les donn es   disposition l'objectif est de faire / refaire des analyses statistiques avec R.

2. <http://cristal.univ-lille.fr/~casiez/acm.php?id=N37457>

3. TPR.zip