

1 - Installation

J. Thioulouse & D. Chessel

Résumé

Cette fiche est une introduction à l'utilisation du logiciel R pour les trois environnements Unix, Windows et MacOS.

Plan

1.	INTRODUCTION	2
2.	INSTALLATION.....	2
3.	UTILISATION.....	4
4.	PREMIER EXEMPLE (WINDOWS)	6

1. Introduction

R est un logiciel de calcul statistique disponible pour les systèmes d'exploitation Windows et Unix. Il existe une version Macintosh, mais dont la mise à jour a été arrêtée à la version 0.64.0. R est un logiciel libre, qui est distribué selon la licence GNU General Public License. Il est très largement compatible avec le logiciel Splus

La version officielle actuelle est la version 1.1.1. Elle est disponible dans les archives du réseau CRAN (Comprehensive R Archive Network) dont le site de base est situé en Autriche : <http://www.ci.tuwien.ac.at/R/>. Il existe plusieurs sites miroirs, et une version locale se trouve sur le serveur FTP du laboratoire de Biométrie : <ftp://pbil.univ-lyon1.fr/pub/Splus/R/>

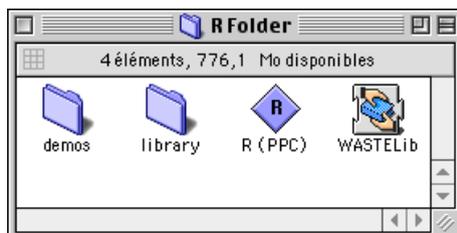
2. Installation

2.1. Version Unix

Les fichiers sources sont distribués sous la forme d'une archive tar compressée avec gzip. L'installation s'effectue de façon standard, sans difficultés particulières (lire le fichier INSTALL). A noter qu'il faut disposer d'un compilateur C (gcc) et d'un compilateur FORTRAN (f77). La version 1.0.1 est installée sur les serveurs du laboratoire de Biométrie (pbil.univ-lyon1.fr) dans le répertoire /appli.

2.2. Version Macintosh

La version Macintosh est la version 0.64.0. Le développement de cette version a en effet été arrêté et la mise à jour ne se fera probablement pas avant la sortie de Mac OS X. Elle est disponible sous forme d'une archive CompactPro auto-extractible et codée en BinHex (MacR-0.64a-BigDist.sea.Hqx). Cette archive contient une application directement exécutable. Le décompactage fournit le dossier suivant :

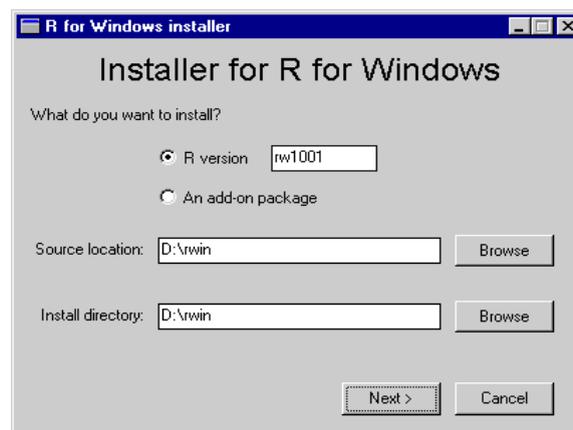


2.3. Version Windows

La distribution se présente sous la forme de dix archives zip, de deux fichiers ReadMe, et d'un programme d'installation exécutable (rwinst.exe).



Lancer le programme d'installation, qui demande de préciser le dossier où doit avoir lieu l'installation :



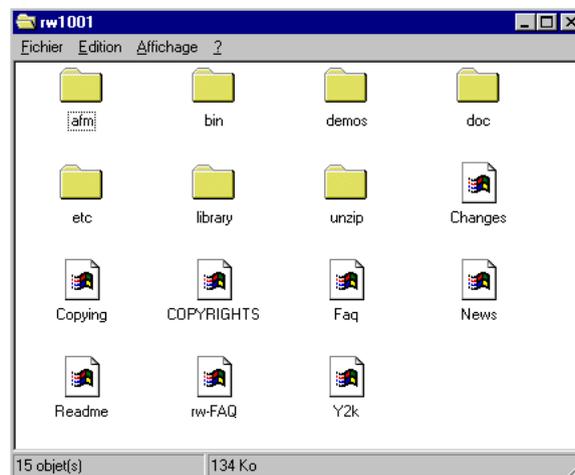
Sélectionner les composants dont on désire réaliser l'installation (conserver les options par défaut en cas de doute) :



Cliquer sur le bouton Next pour terminer l'installation :



Cliquer sur le bouton Finish pour terminer l'installation. On obtient un dossier intitulé rw1001 :



Ce dossier contient en particulier un dossier intitulé bin, qui contient lui-même l'application Rgui.exe, constituant l'interface utilisateur graphique de R.

3. Utilisation

3.1. Version Unix

Pour utiliser le logiciel, il faut se connecter sur la station Unix, se déplacer dans le dossier de travail choisi (commande "cd") et taper la commande "R" pour lancer l'exécution :

```
thioulou% cd FichSplus
/mnt/users/thioulou/FichSplus
thioulou% R
```

```
R : Copyright 2000, The R Development Core Team
Version 1.0.1 (April 14, 2000)
```

```
R is free software and comes with ABSOLUTELY NO WARRANTY.
```

```

You are welcome to redistribute it under certain conditions.
Type "?license" or "?licence" for distribution details.

R is a collaborative project with many contributors.
Type "?contributors" for a list.

Type "demo()" for some demos, "help()" for on-line help, or
"help.start()" for a HTML browser interface to help.
Type "q()" to quit R.
>

```

Utiliser la commande `getwd()` pour vérifier le dossier de travail courant:

```

>getwd()
[1] "/mnt/users/thioulou/FichSplus"

```

et la commande `q()` pour quitter le logiciel :

```

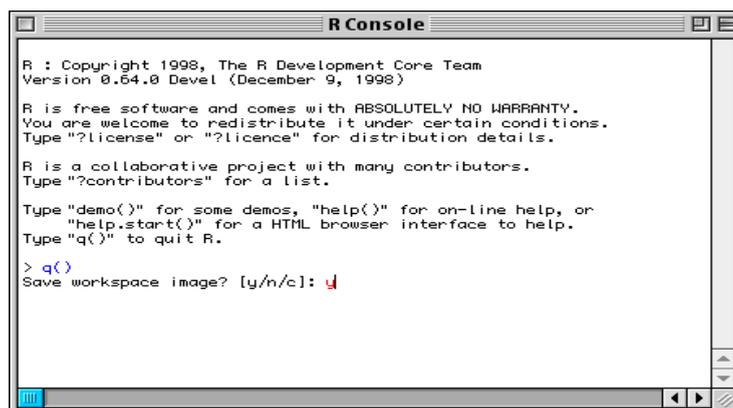
>q()
Save workspace image? [y/n/c]: n
thioulou%

```

Si on répond "yes" à la question "Save workspace image?", les objets créés lors de la session sont enregistrés dans un fichier intitulé ".RData" situé dans le dossier de travail. Les fichiers présents dans le répertoire courant (dossier de travail) sont accessibles directement ("fic.txt"). Les fichiers présents dans des répertoires différents sont accessibles en spécifiant leur chemin complet ("/mnt/users/thioulou/ficbis.txt").

3.2. Version Macintosh

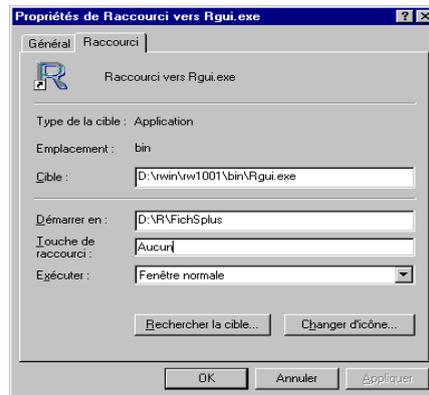
Lancer l'application par un double clic sur "R (PPC)".



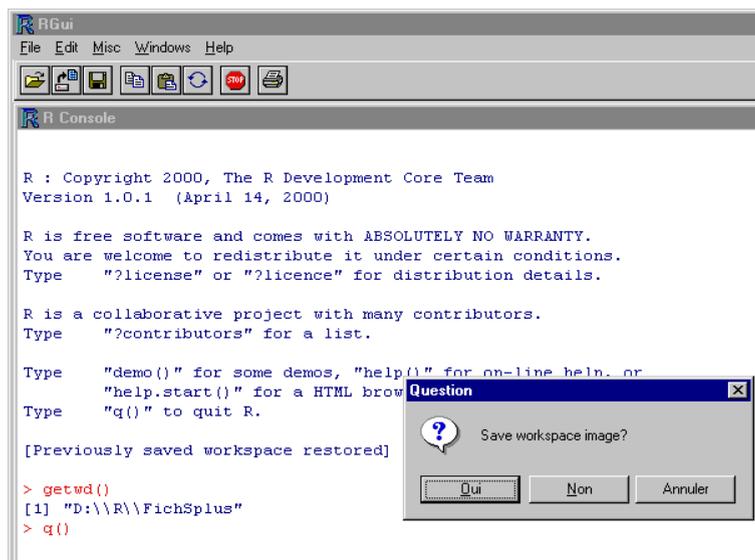
Si on répond yes à la question "Save workspace image?", les objets créés lors de la session sont enregistrés dans un fichier intitulé "R Session" dont on doit indiquer le répertoire. Les fichiers ne sont accessibles qu'en spécifiant leur chemin complet en syntaxe Macintosh (séparateurs = ":") HD4:Data:human.txt

3.3. Version Windows

Créer un raccourci de Rgui.exe et le placer dans un dossier de travail quelconque. Modifier le paramètre "Démarrer en :" de ce raccourci (Clic droit -> Propriétés) pour indiquer le dossier de travail choisi :



Exécuter Rgui.exe en cliquant sur le raccourci et vérifier que le dossier de travail est bien celui désiré avec la fonction `getwd()`. Quitter la session en tapant `q()` :



Si on répond «Oui» à la question "save workspace image?", les objets créés lors de la session sont enregistrés dans un fichier intitulé ".RData" situé dans le dossier de travail. Les fichiers présents dans le répertoire courant (dossier de travail) sont accessibles directement. Les fichiers présents dans des répertoires différents sont accessibles en spécifiant leur chemin complet, en utilisant la syntaxe Windows mais en doublant les \ :

```
D:\\dossier\\fic.txt
```

4. Premier exemple (Windows)

Pour une première prise de contact, éditer le loi binomiale pour $n = 10$ et $p = 1/3$. Après le prompt taper :

```
> dbinom(0:10,10,1/3) loi binomiale
[1] 1.734153e-02 8.670765e-02 1.950922e-01 2.601229e-01 2.276076e-01
[6] 1.365645e-01 5.690190e-02 1.625768e-02 3.048316e-03 3.387018e-04
[11] 1.693509e-05

> options(digits=4) pour avoir 4 chiffres significatifs
```

Rappeler l'ordre avec la touche « flèche en haut »

```
> dbinom(0:10,10,1/3)
[1] 1.734e-02 8.671e-02 1.951e-01 2.601e-01 2.276e-01 1.366e-01 5.690e-02
[8] 1.626e-02 3.048e-03 3.387e-04 1.694e-05
```

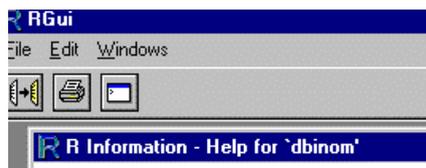
Quelle est la probabilité d'obtenir 1 avec une loi binomiale pour $n = 10$ et $p = 1/3$:

```
> dbinom(1,10,1/3)
[1] 0.0867
```

Quelle est la probabilité d'obtenir plus de 45 et moins de 55 avec une loi binomiale pour $n = 100$ et $p = 1/2$? Taper :

```
> ?dbinom
```

Il vient la fenêtre de documentation de la fonction :



En bas lire :

```
# Compute P(45 < X < 55) for X Binomial(100,0.5)
sum(dbinom(46:54, 100, 0.5))
```

Copier l'ordre et le coller dans la fenêtre de commande :

```
> ?dbinom
> sum(dbinom(46:54, 100, 0.5))
```

On a :

```
> sum(dbinom(46:54, 100, 0.5))
[1] 0.6318
```

Quelle est la probabilité d'obtenir au plus 1 avec une loi binomiale pour $n = 10$ et $p = 1/3$:

```
> pbinom(1,10,1/3)
[1] 0.1040
> 1.734e-02+8.671e-02
[1] 0.1041
> pbinom(0.5,10,1/3)
[1] 0.01734
> pbinom(-0.5,10,1/3)
[1] 0
```

Noter que **d** donne les valeurs $P(X = j)$ (densité, définie pour les valeurs possibles) et que **p** donne les valeurs $P(X \leq x)$ (fonction de répartition, définie pour tout x). Quelle est la probabilité de dépasser strictement 4 pour une loi de Poisson de paramètres 2.7 ?

```
> 1-ppois(4,2.7)
```

```
[1] 0.1371
```

Quelle est la probabilité de dépasser 1.96 pour une loi normale réduite :

```
> 1-pnorm(1.96)
[1] 0.025
```

Quelle est la valeur x telle que $P(X \leq x) = 0.975$ pour une loi normale (quantile) :

```
> qnorm(0.975)
[1] 1.96
```

Quel est le quantile 1% pour une loi T à 5 ddl :

```
> qt(0.01,5)
[1] -3.365
```

Noter que **Q** donne pour y la plus grande valeur de x telle que $P(X \leq x) = y$ (quantile).

Donner un échantillon aléatoire simple de 10 valeurs d'une loi de Poisson de paramètre 2.7 :

```
> rpois(10,2.7)
[1] 2 2 1 0 4 0 7 0 1 4
```

d'une loi normale réduite :

```
> rnorm(10)
[1] -0.2174 -1.3695 -0.3055 -0.3072 -0.3036 1.0070 0.8167 0.4624 -2.3624
[10] 0.6982
```

d'une loi Khi2 à 2 ddl :

```
> rchisq(10,2)
[1] 3.4801 1.1599 0.5118 0.8682 3.2848 0.4763 1.0231 0.9021 1.8163 0.5807
```

d'une loi binomiale $n = 100$ et $p = 1/2$

```
> rbinom(10,100,0.5)
[1] 48 51 49 48 56 51 54 53 51 37
```

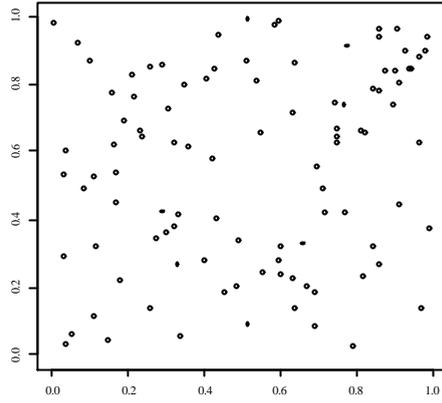
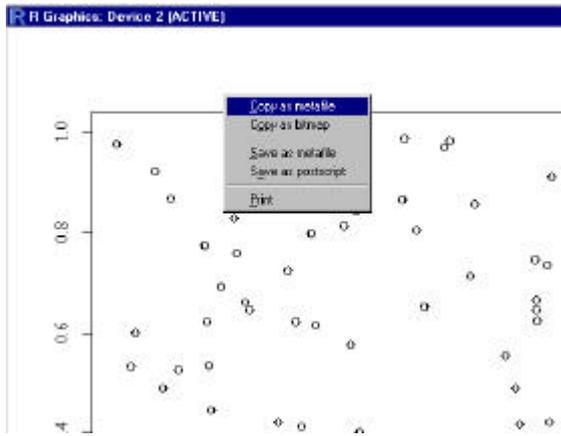
Noter que **r** donne des échantillons. **d**, **p**, **q** et **r** sont disponibles pour :

dbinom	loi binomiale	dnbinom	loi binomiale négative
pois	loi de Poisson	norm	loi normal
lnorm	loi lognormale unif		loi uniforme
t	loi de Student	chisq	loi Chi2
exp	loi exponentielle	f	loi F
gamma, beta, wilcox, ...			

Pour une première figure, placer 100 points au hasard dans un carré :

```
> plot(runif(100),runif(100),xlim=c(0,1),ylim=c(0,1),xlab="",ylab="")
```

Cliquer avec le bouton droit dans la figure et coller dans Word :



à suivre ...