

Une introduction à SciViews R

(Une interface graphique pour R)

Faouzi LYAZRHI

UP Biostatistique

Ecole Nationale Vétérinaire, 23, chemin des Capelles, BP 87614, F-31076 Toulouse cédex

email : f.lyazrhi@envt.fr

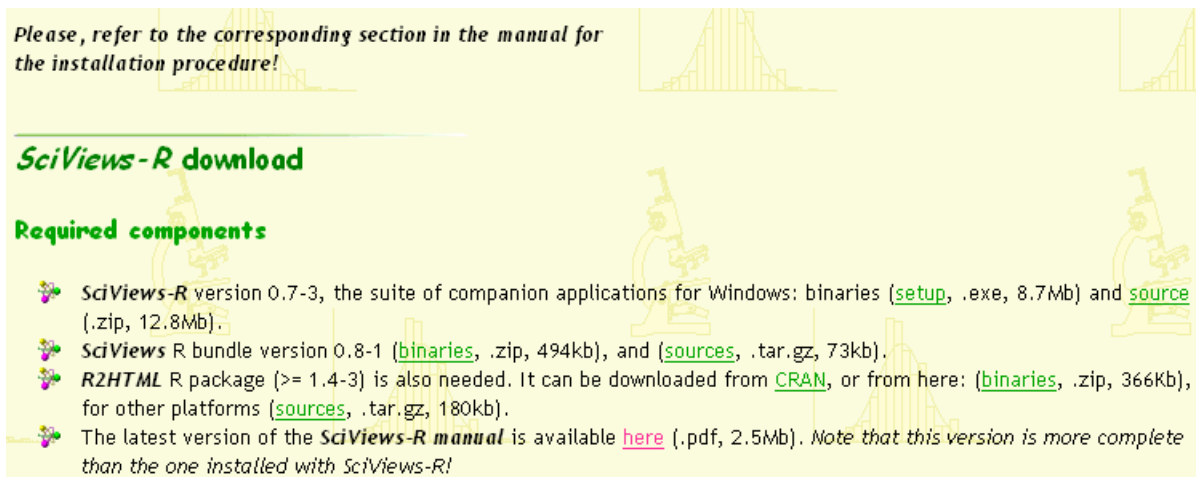
1. Installer SciViews R	3
1.2. Installer R.....	4
2. Importer des données.....	9
3. Manipuler des données	10
3.1 Transformer une variable.....	10
3.2 Sélectionner un sous-ensemble de variables	11
4. Quelques distributions statistiques.....	12
5. Quelques outils statistiques.....	13
5.1 Statistiques descriptives élémentaires.....	14
5.2 Tests Statistiques usuels.....	15
5.2.1 Tests sur les moyennes	15
5.2.2 Tests sur les variances.....	16
5.2.3 Tests sur les pourcentages.....	17
5.2.4 Tests du Khi-deux.....	18
5.3 Tests non paramétriques	20
5.4 ANOVA à un facteur	21
5.4 Régression multiple	24
5.5 Modèle linéaire généralisé	25

SciViews est une interface graphique pour R très conviviale. C'est un logiciel libre comme R et qui possède les outils statistiques et graphiques les plus utilisés. SciViews peut être téléchargé sur le site <http://www.sciviews.org/SciViews-R/index.html>. Vous pouvez également accéder sur ce site à la documentation de SciViews (<http://www.sciviews.org/SciViews-R/Manual.pdf>) vous y trouverez en particulier le détail concernant l'installation de SciViews.

 **Remarque** : Le logiciel R et SciViews sont régulièrement actualisés, vous choisirez de télécharger les dernières versions disponibles.

1. Installer SciViews R

En cliquant sur le lien du site <http://www.sciviews.org/SciViews-R/index.html>, sélectionner dans la page d'accueil (affichée comme ci-dessous) sur setup.exe.



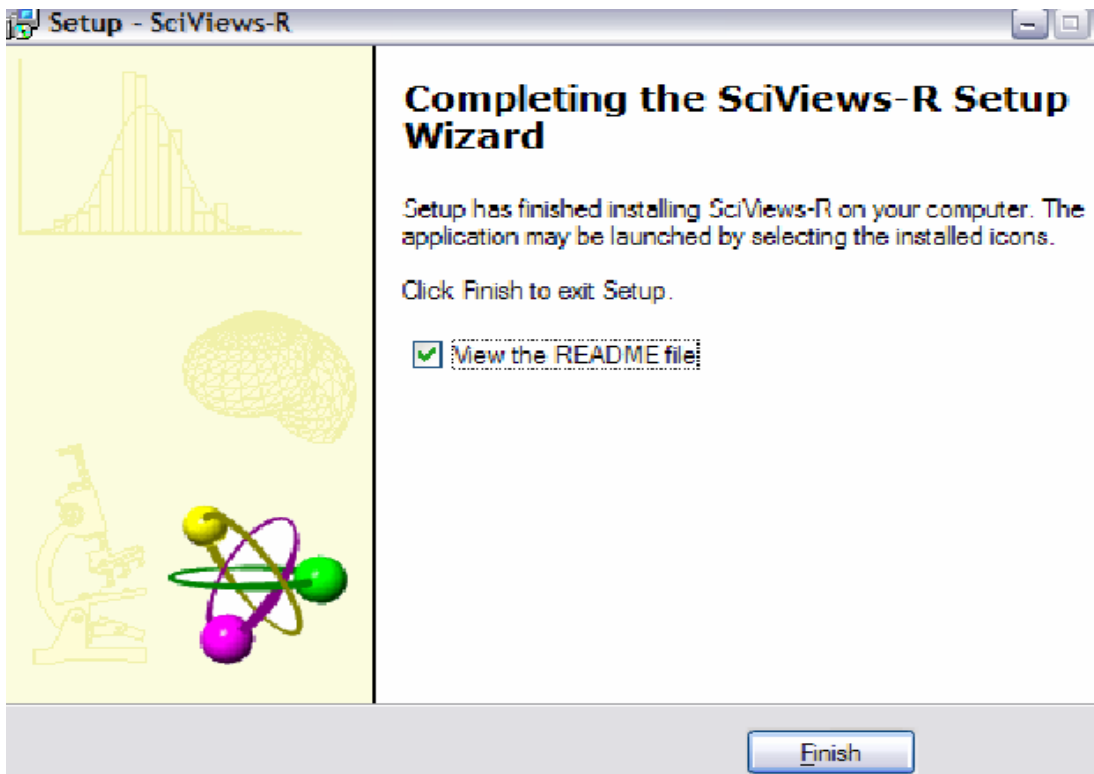
Please, refer to the corresponding section in the manual for the installation procedure!

SciViews-R download

Required components

- SciViews-R version 0.7-3, the suite of companion applications for Windows: binaries ([setup](#), .exe, 8.7Mb) and [source](#) (.zip, 12.8Mb).
- SciViews R bundle version 0.8-1 ([binaries](#), .zip, 494kb), and ([sources](#), .tar.gz, 73kb).
- R2HTML R package (>= 1.4-3) is also needed. It can be downloaded from [CRAN](#), or from here: ([binaries](#), .zip, 366Kb), for other platforms ([sources](#), .tar.gz, 180kb).
- The latest version of the *SciViews-R manual* is available [here](#) (.pdf, 2.5Mb). Note that this version is more complete than the one installed with SciViews-R!

Double cliquer sur setup.exe et cliquer sur *next* en acceptant les options par défaut. A la fin des différentes étapes vous verrez s'inscrire à l'écran la boîte de dialogue suivante :



En cliquant sur Finish, l'installation de SciViews est terminée. Pour compléter l'installation de SciViews, il faut installer R. Il est indispensable de choisir la version 2001 ou plus récent.

1.2. Installer R

R peut-être installé à partir su site <http://cran.cict.fr/index.html>, en cliquant sur ce lien puis sur Windows(95 and later) comme ci-dessous :

Precompiled Binary Distributions

Base system and contributed packages. **Windows and Mac** users most likely want these versions of R.

[Linux](#)

[MacOS X \(10.2.x and above\)](#) This version of R for the Mac is actively maintained.

[MacOS \(System 8.6 to 9.1 and MacOS X up to 10.1.x\)](#) Last supported version of R is 1.7.1, there will be no more updates.

[Windows \(95 and later\)](#)

Cliquer sur base :

Note: CRAN does not have Windows systems and cannot check these binaries for viruses. Use the normal precautions with downloaded executables.

Subdirectories:

[base](#) Binaries for base distribution (managed by Duncan Murdoch)

[contrib](#) Binaries of contributed packages (managed by Uwe Ligges)

Puis télécharger le fichier `rw2001.exe` :

[README.rw2001](#) Installation and other instructions.

[CHANGES](#) New features of this Windows version.

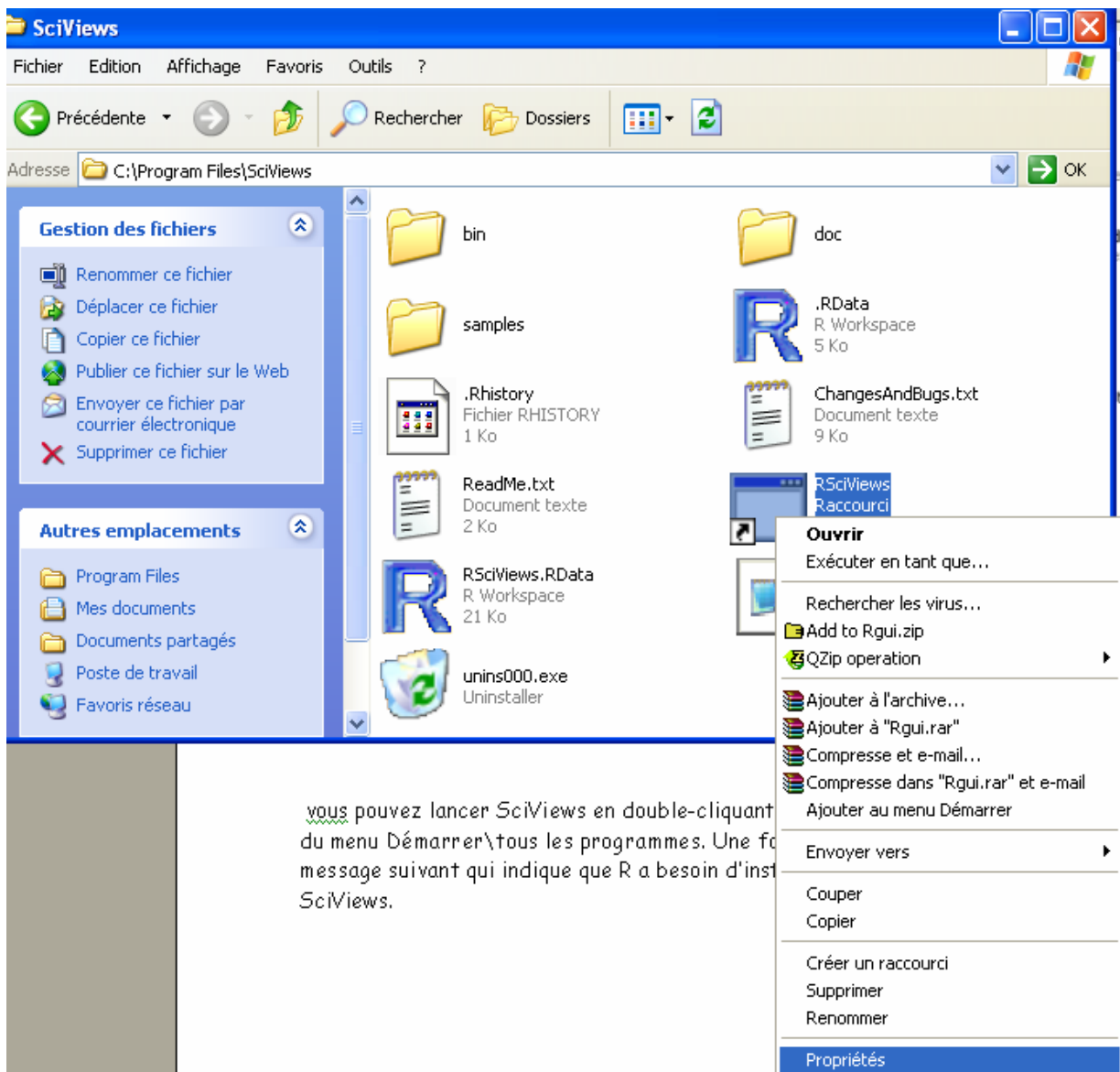
[NEWS](#) New features of all versions.

[rw2001.exe](#) Setup program (about 23 megabytes). Please download this from a [mirror near you](#). This corresponds to the file named **SetupR.exe** in pre-1.6.0 releases.

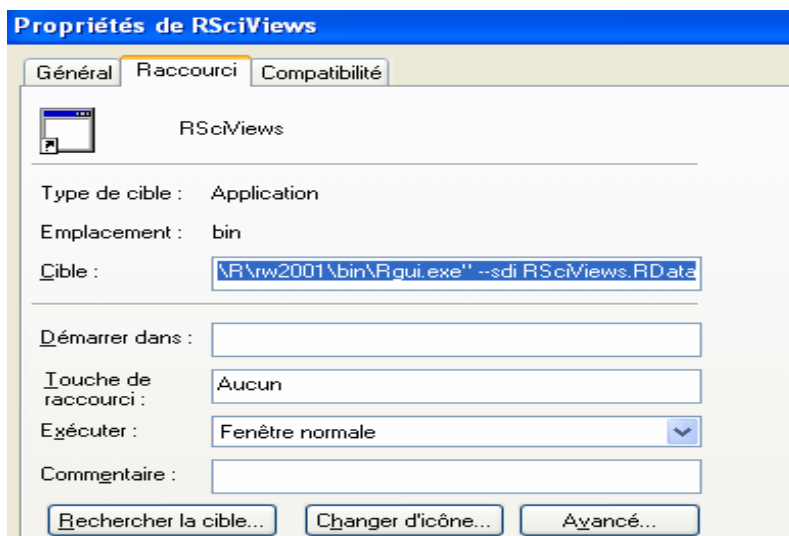
Le fichier rw2001.exe est un exécutable, une fois sauvegardé sur votre disque dur, vous double-cliquez dessus. Il est conseillé de garder les options par défaut, il ne vous reste plus qu'à cliquer à chaque fois sur next, à la fin de l'installation la boîte de dialogue suivante s'affichera à l'écran.



Pour finaliser l'installation de SciViews et R, il faut vérifier le répertoire de travail du raccourci de SciViews. Pour cela aller dans c:\Program Files\SciViews (le répertoire par défaut ou est installé SciViews). Cliquez avec le bouton droit sur le raccourci de SciViews puis propriétés.

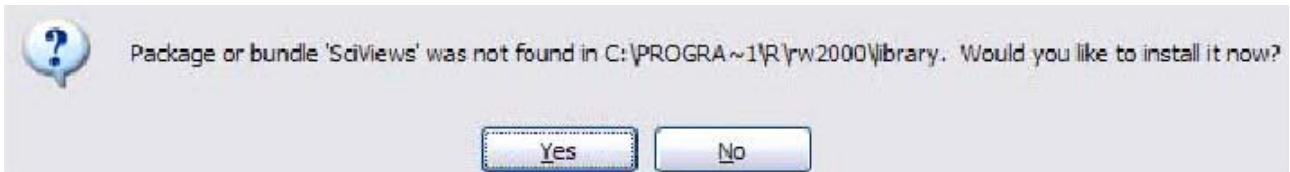


Vérifier comme ci-dessous que dans la case Cible, il y a bien la répertoire où est installé R suivi de -sdi RSciViews.RData

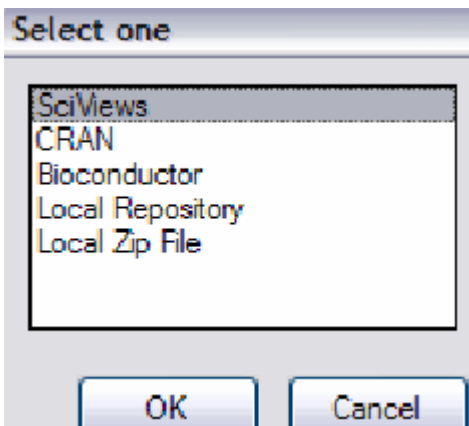


Maintenant l'installation est complètement terminée vous pouvez lancer SciViews en double-cliquant sur l'icône SciViews créée sur votre bureau ou à l'aide du menu Démarrer\tous les programmes.

Une fois la console R lancée, vous verrez apparaître le message suivant qui indique que R a besoin d'installer d'autres modules nécessaires pour l'interface SciViews.



En cliquant sur yes, vous verrez apparaître une autre boîte de dialogue :



En sélectionnant SciViews et une fois l'installation terminée, vous verrez apparaître dans la console SciViews R le message ci-dessous, en tapant y les fichiers temporaires font être détruits.

```
bundle 'SciViews' successfully unpacked and MD5 sums checked
```

```
Delete downloaded files (y/N)? █
```

D'autres modules vont être installés de la même manière en particulier le module R2HTML.

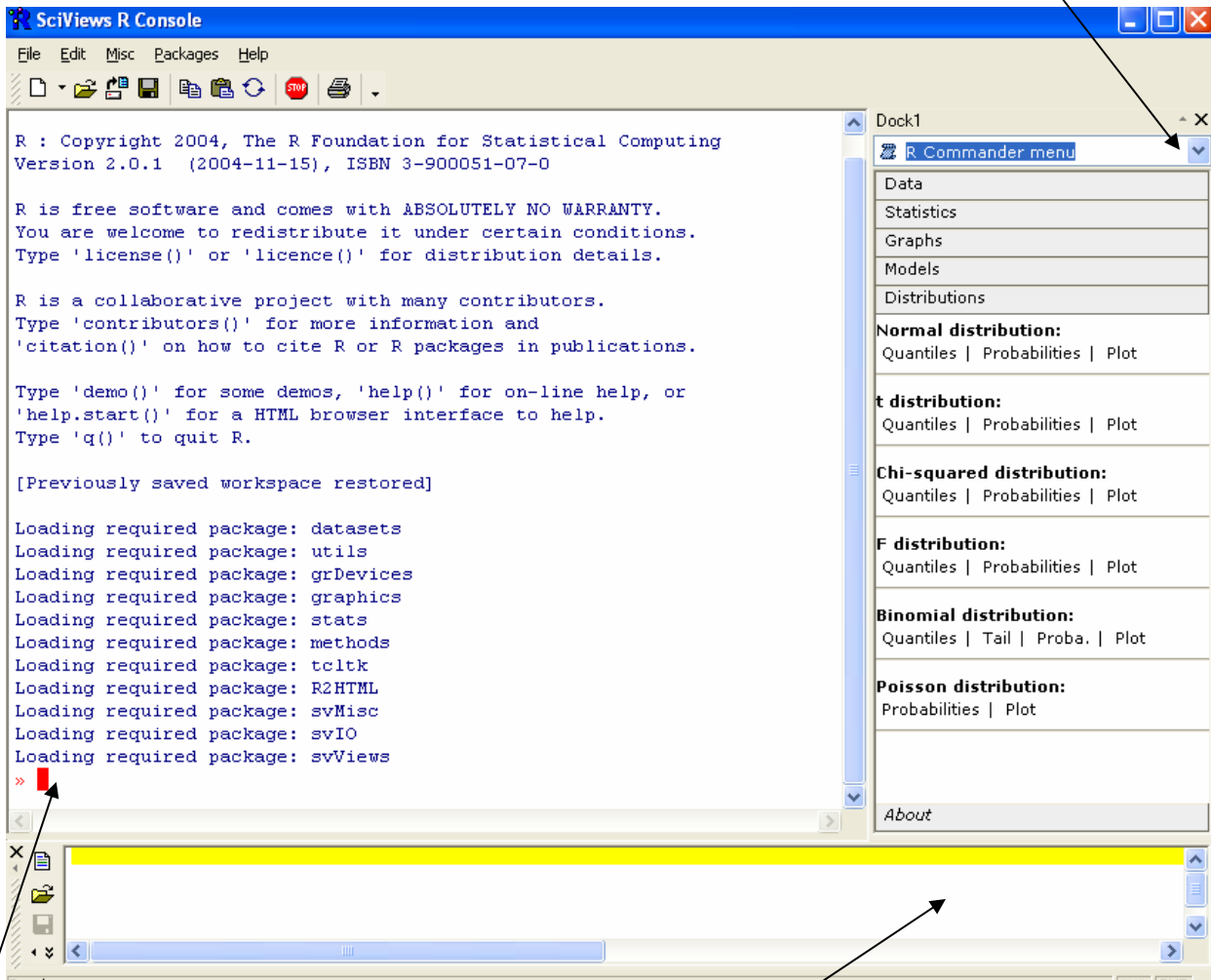
Une fois les modules nécessaires au lancement de la console R de SciViews, vous verrez apparaître la console R de SciViews ci-dessous.



Remarque très importante : Si vous ne disposez pas d'une connexion à Internet en continu, vous devez installer tous les packages nécessaires à R et SciViews dès la première installation.

Le menu déroulant (Dock) :

Graphiques, statistiques, manuels, liste des fonctions R, explorateur de fichiers, aide en ligne,etc.



Fenêtre principale :

Affichage des résultats, saisir et exécuter des commandes en interactif

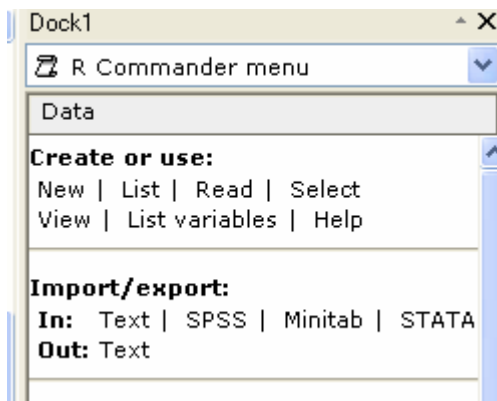
Barre de commandes :

Exécuter des commandes, créer, ouvrir ou sauvegarder un fichier de commandes

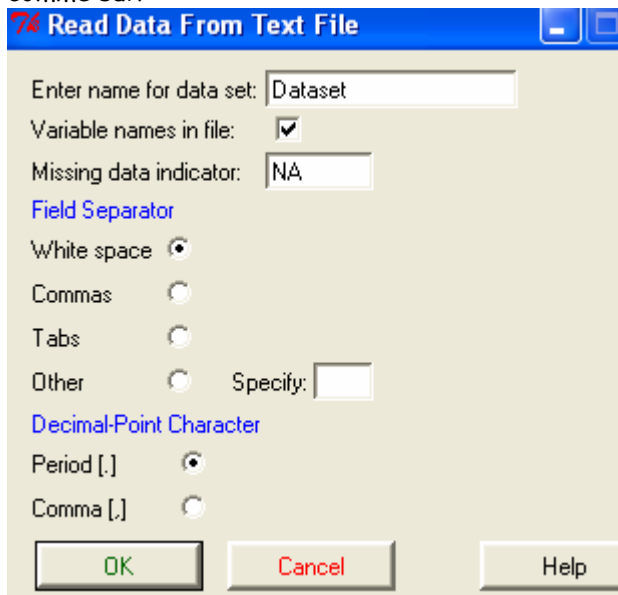
2. Importer des données

Comme pour le langage R, la meilleure façon d'importer un fichier de données avec SciViews reste le format texte (.txt)

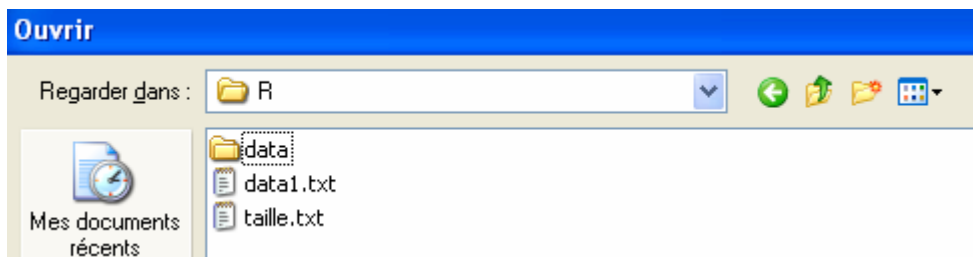
Pour importer un fichier de données sauvegardé sous format '.txt', il suffit de cliquer sur le menu data dans le " Dock " R Commander menu :



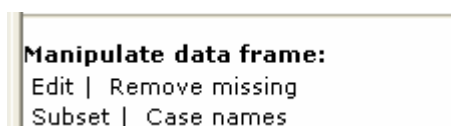
Puis en cliquant sur *in : Text* du sous menu Import/export, une boîte de dialogue s'affichera à l'écran comme suit :



En cliquant sur ok, choisir le répertoire où se trouve le fichier de données que vous voulez ouvrir.



R ne fait que lire le contenu du fichier à ce stade pour l'afficher à l'écran clique sur *edit* du sous menu *manipulate data frame* :



Une feuille de calcul s'affiche à l'écran comme suit :

R Data Editor						
File Edit Help						
	poids	taille	fumeur	sexe	sport	etat
1	85	184	oui	homme	1	malade
2	65	175	oui	homme	1	malade
3	74	180	oui	homme	2	gueri
4	79	175	oui	homme	2	malade
5	71	165	non	homme	3	gueri

3. Manipuler des données

A l'aide de l'éditeur vous pouvez modifier le contenu d'une variable ou saisir de nouvelles variables, en cliquant sur le nom de la variable vous pouvez spécifier sa nature (numérique ou caractère).

3.1 Transformer une variable

Toutes les fonctions arithmétiques usuelles (`sqrt()`, `log()`, `exp()`,...) peuvent être exécutées de façon interactive dans la fenêtre principale. Pour affecter le contenu de la variable transformée dans une nouvelle variable est symbolisé par `<-` (le symbole inférieur suivi du signe -) :

```
>> lpoids <- log(poids)
>> █
```

R Data Editor								
File Edit Help								
	poids	taille	fumeur	sexe	sport	etat	lpoids	vs
1	85	184	oui	homme	1	malade	117	
2	65	175	oui	homme	1	malade	97	
3	74	180	oui	homme	2	gueri	106	

Une deuxième façon de faire consiste à utiliser le menu data, pour cela il suffit de cliquer sur le sous-menu *compute* :

Manipulate variables:

- Compute | Standardize
- Recode | Bin (make classes)
- Factor | Reorder | Contrasts

Dans la boîte de dialogue qui s'affiche comme suit, saisir le nom de la variable à transformer et l'expression de la transformation dans les cases correspondantes

Compute New Variable

Current variables (list only)

etat	▲
fumeur	■
lpoids	▼
poids	▼

New variable name:

Expression to compute:

3.2 Sélectionner un sous-ensemble de variables

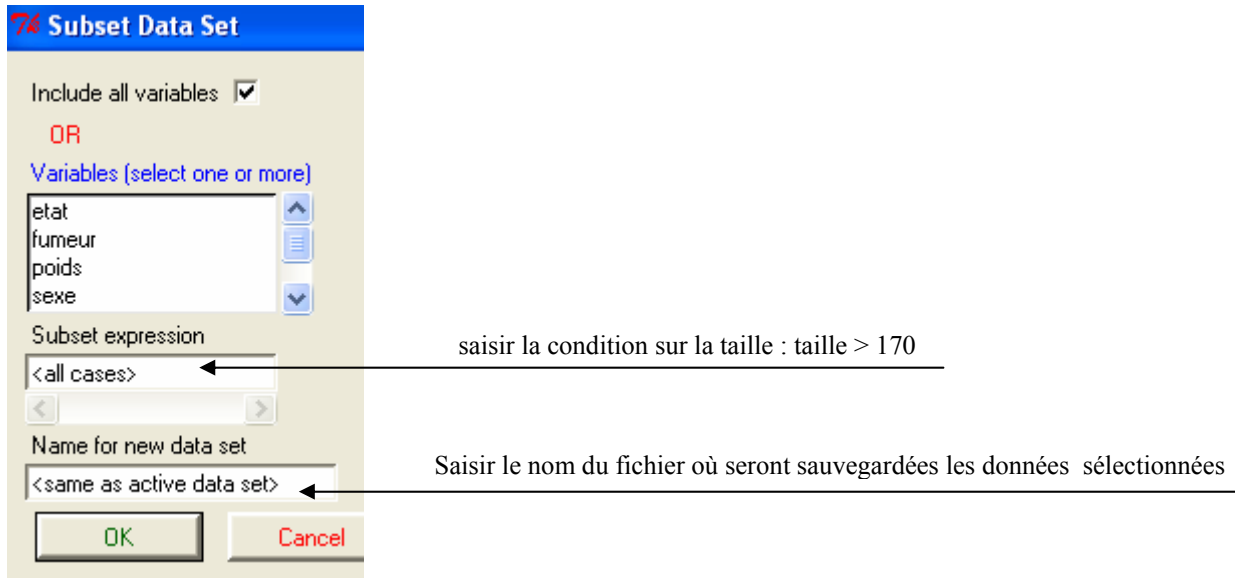
La sélection de variables se fait à l'aide de la commande *subset* du sous menu data :

Manipulate data frame:

Edit | Remove missing
Subset | Case names

Exemple : On voudrait sélectionner dans le fichier data1.txt les individus qui correspondent à une taille supérieure à 170cm.

Pou cela cliquer sur subset, une boîte de dialogue s'ouvre comme suit :



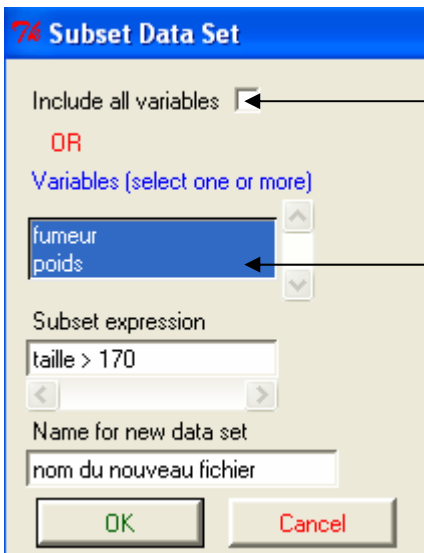
Ensuite vous cliquez sur edit :

	row.names	poids	taille	fumeur	sexe	sport	etat	va
1	1	85	184	oui	homme	1	malade	
2	2	65	175	oui	homme	1	malade	
3	3	74	180	oui	homme	2	gueri	
4	4	79	175	oui	homme	2	malade	
5	6	80	185	non	homme	3	gueri	
6	7	75	180	non	homme	4	malade	
7	10	64	171	oui	femme	1	gueri	
8	16	70	175	non	femme	4	gueri	

Remarque :

Vous pouvez sélectionner les individus suivant un critère et ne garder que quelques variables.

Exemple : on voudrait sélectionner les variables poids et fumeur correspond aux individus de taille supérieure à 170 et ne garder que les variables poids et fumeur dans le nouveau fichier :



Décocher cette case

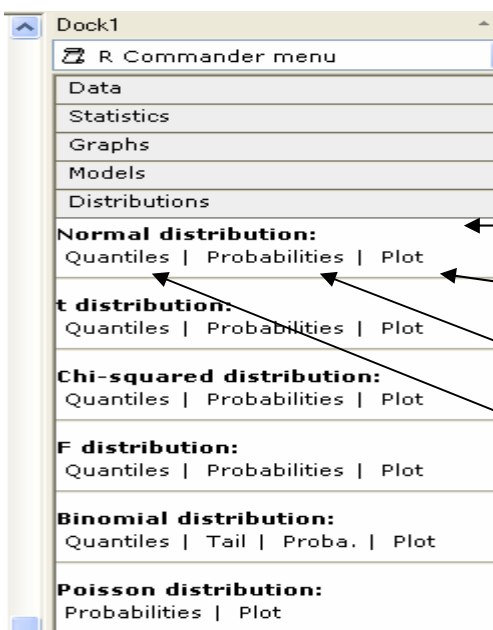
Sélectionner avec la souris les variables que vous voulez sauvegarder dans le nouveau fichier

Le résultat obtenu est le suivant :

	row.names	fumeur	poids	va
1	1	oui	85	
2	2	oui	65	
3	3	oui	74	
4	4	oui	79	
5	6	non	80	
6	7	non	75	
7	10	oui	64	
8	16	non	70	
9				

4. Quelques distributions statistiques

En cliquant sur le sous-menu *distributions* du menu R Commander, la liste des distributions statistiques usuelles s'affichera comme suit. Pour chaque distribution, on peut calculer les quantiles, probabilités ainsi que leurs représentations graphiques.



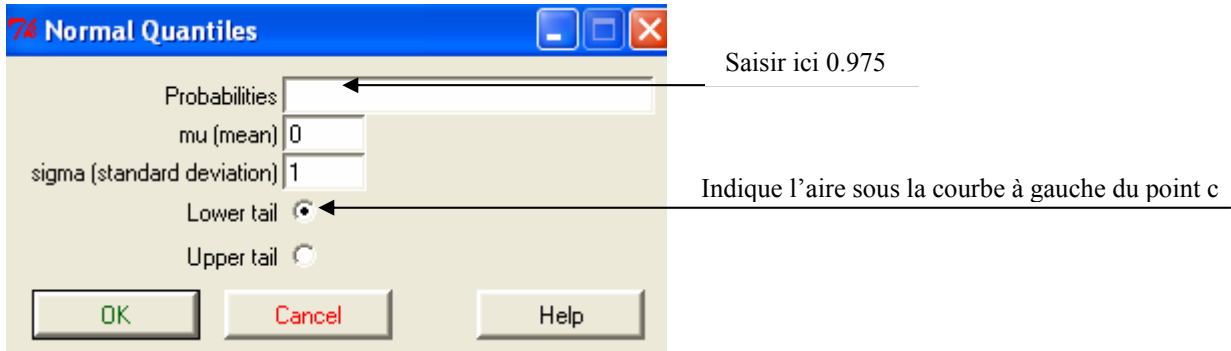
Loi normale

représentation graphique de la distribution

Calcul des probabilités (aire sous la courbe)

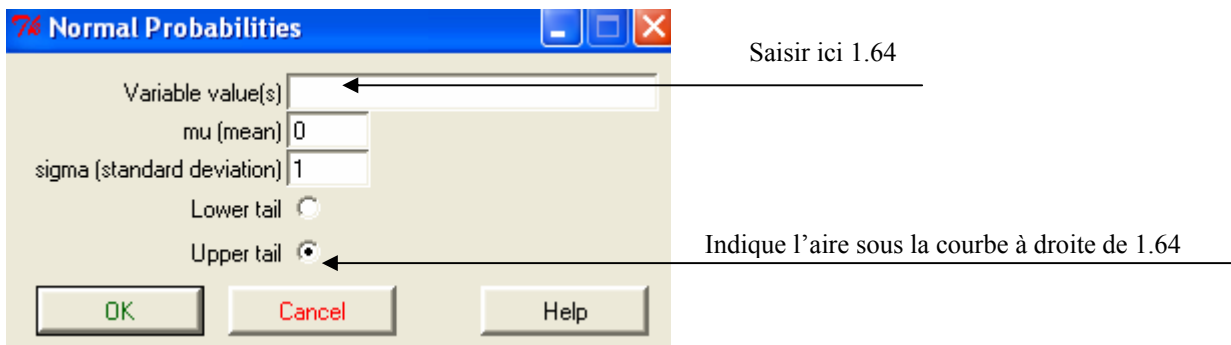
Calcul des quantiles

Exemple 1 : calculer $P(X < c) = 0.975$, où X suit une loi normale centrée réduite
 Pour cela cliquer sur *Quantiles* du sous menu Normal distributions, la boîte de dialogue ci-dessous s'affichera :



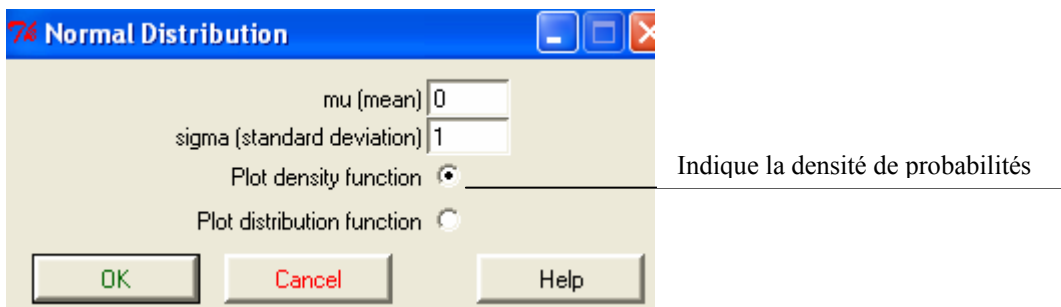
Exemple 2 : calculer $P(X > 1.64)$, où X suit une loi normale centrée réduite

Pour cela cliquer sur *Probabilities* du sous menu Normal distribution :



Exemple 3 : Représentation graphique de la densité de la loi normale centrée réduite

Cliquer sur plot du sous menu Normal distribution :

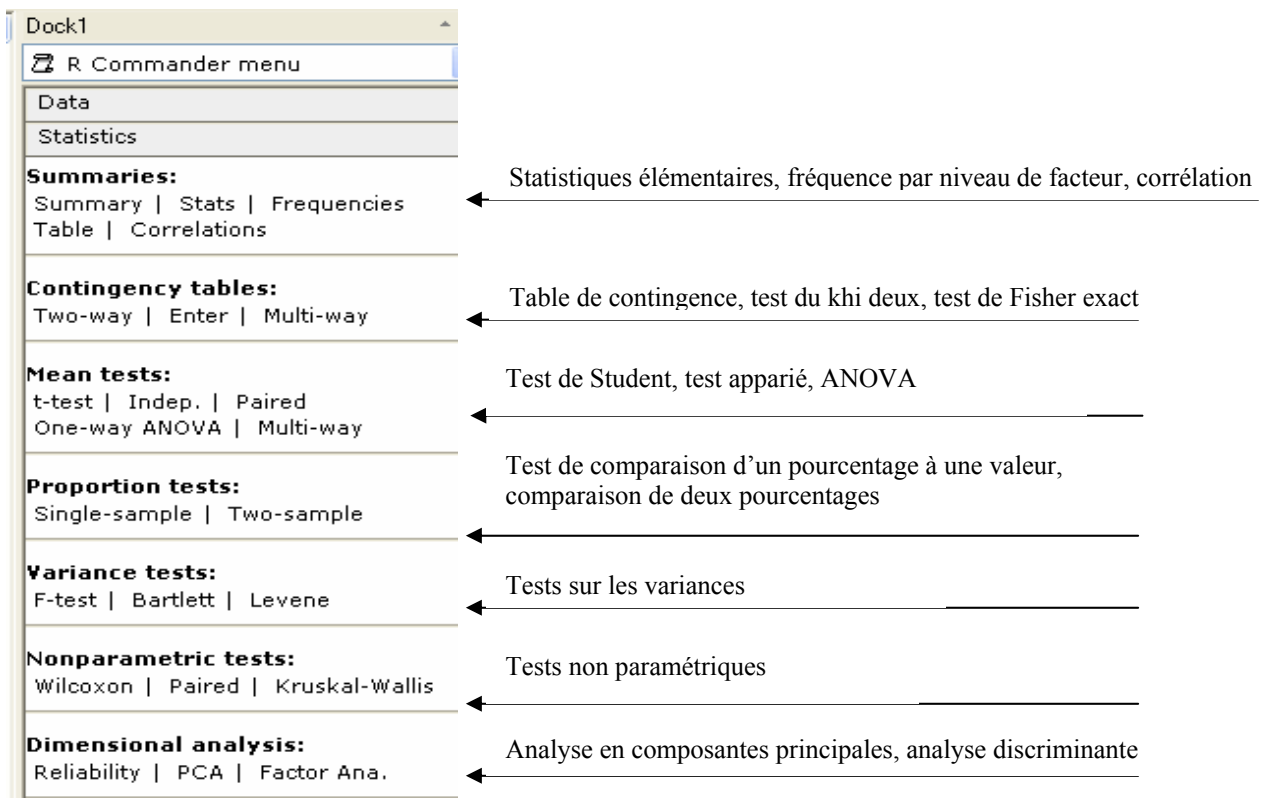


En cliquant sur OK, une fenêtre graphique s'affichera à l'écran, avec le bouton vous pouvez sauvegarder le graphique sous différents formats ainsi que l'imprimer.

Remarque : Vous pouvez représenter la fonction de répartition au lieu de la densité pour cela il suffit de cocher la case *Plot distribution function* dans la boîte de dialogue ci-dessus.

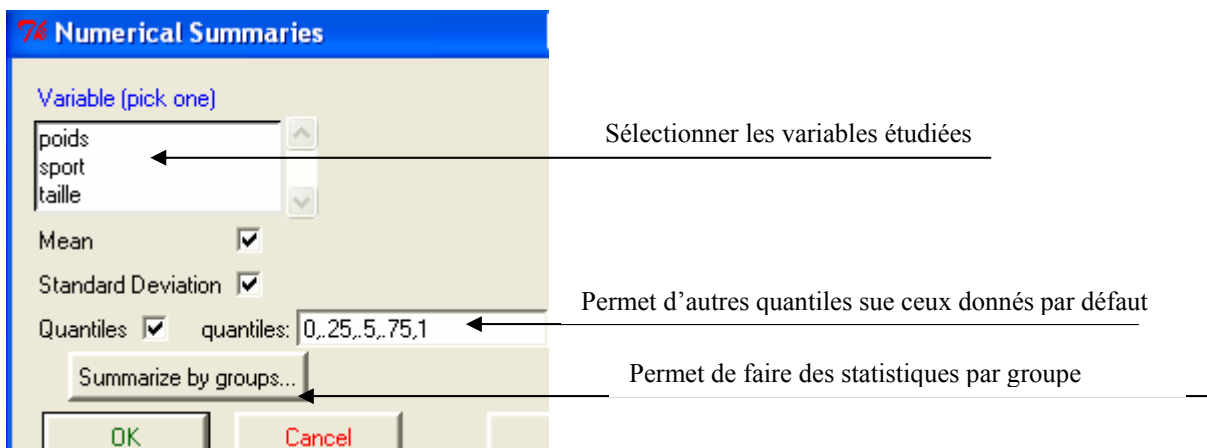
5. Quelques outils statistiques

Dans ce paragraphe, nous présentons le menu des outils statistiques les plus utilisés. Pour cela, il suffit de cliquer sur le sous menu *statistics* du menu R Commander.



5.1 Statistiques descriptives élémentaires

Les statistiques élémentaires moyenne (mean), écart-type (sd) et les 3 quartiles (quantiles) s'obtiennent à l'aide de la commande Stats. Comme le montre l'exemple suivant appliqué au fichier data1.txt :



Dans notre exemple les statistiques sur le poids :

```
» Rcmdr » mean(Dataset$poids, na.rm=TRUE)
[1] 70.375

» Rcmdr » sd(Dataset$poids, na.rm=TRUE)
[1] 7.446476

» Rcmdr » quantile(Dataset$poids, c( 0,.25,.5,.75,1 ), na.rm=TRUE)
  0%   25%   50%   75%  100%
55.00 65.00 70.00 74.25 85.00
```

Remarque : La commande `summary` permet d'obtenir les statistiques élémentaires sur toutes les variables contenues dans le fichier, en cliquant sur cette commande on obtient les résultats suivants concernant le fichier data1.txt :

```

» Rcmdr » summary(Dataset)
      poids      taille      fumeur      sexe      sport      etat
Min.   :55.00   Min.   :155.0   non:8    femme:8   Min.   :1.00   gueri :8
1st Qu.:65.00   1st Qu.:165.0   oui:8    homme:8   1st Qu.:1.75   malade:8
Median :70.00   Median :170.0
Mean   :70.38   Mean   :170.8
3rd Qu.:74.25   3rd Qu.:176.3
Max.   :85.00   Max.   :185.0
Max.   :4.00
» ■

```

5.2 Tests Statistiques usuels

Dans ce paragraphe nous allons présenter les différents tests statistiques usuels. Pour cela nous allons dans un premier temps présenter les tests sur les moyennes puis les variances, les tests sur les pourcentages et enfin les tests non paramétriques.

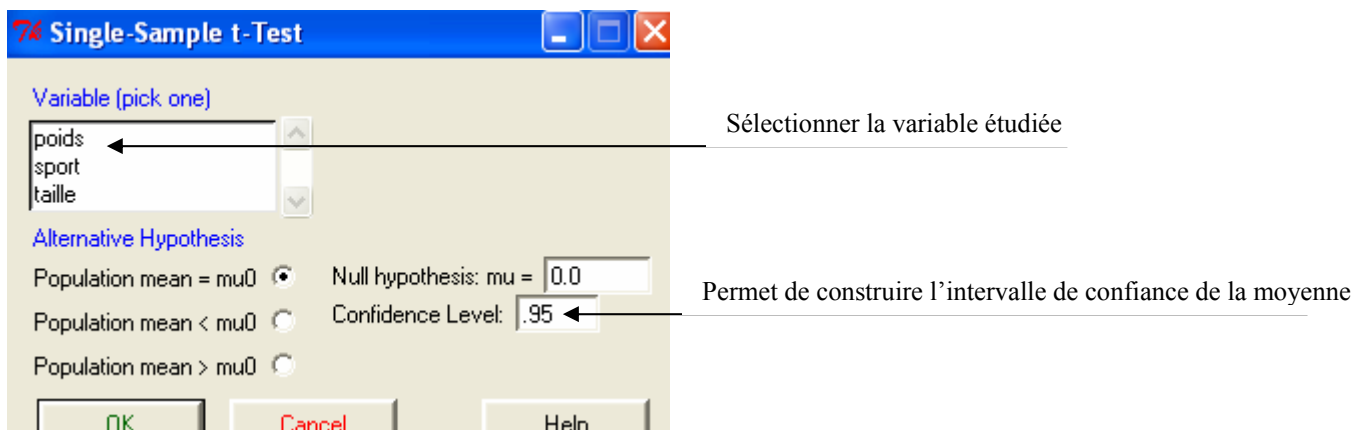
5.2.1 Tests sur les moyennes

Les tests statistiques usuels portant sur la moyenne sont donnés par le sous menu *Mean tests* du menu *Statistics*.

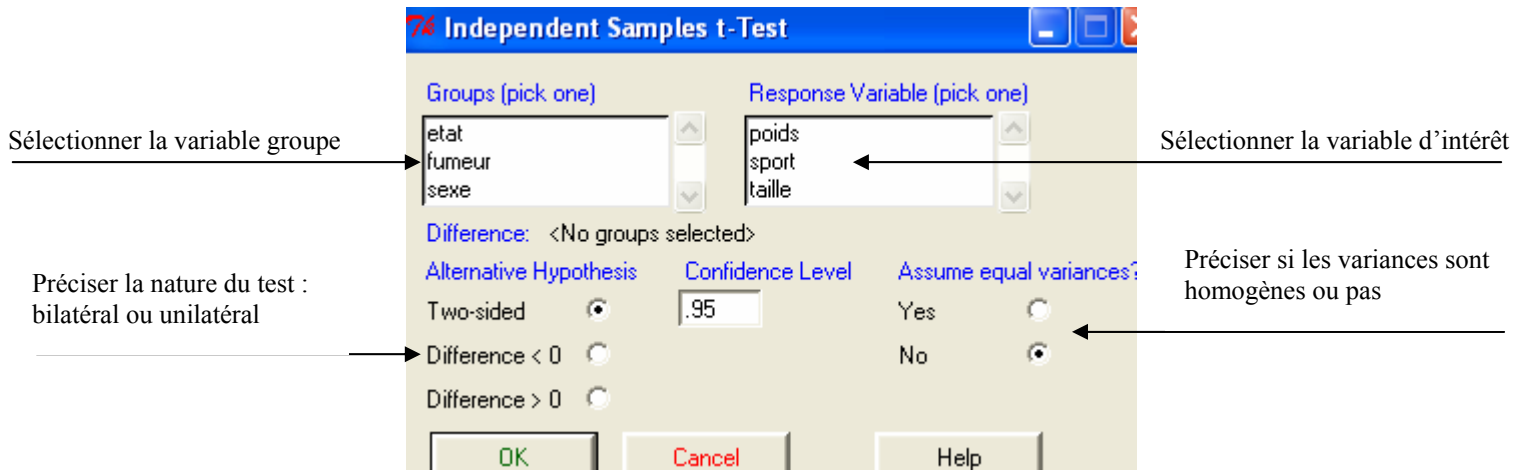
Mean tests:
t-test | Indep. | Paired
One-way ANOVA | Multi-way

➤ Comparaison d'une moyenne à une valeur donnée (t-Test)

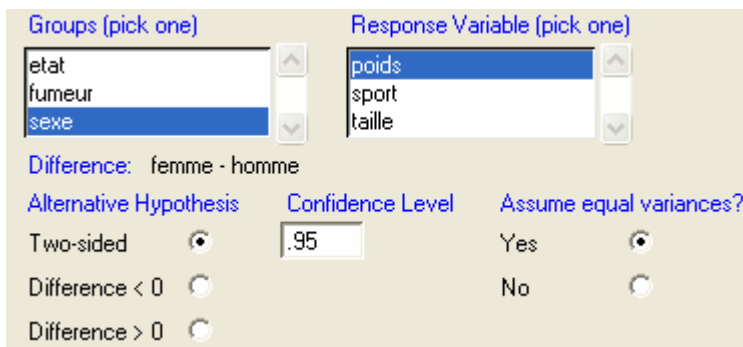
En cliquant sur le sous menu t-Test, la boîte de dialogue suivante s'affiche à l'écran, test bilatéral, test unilatéral ainsi que l'intervalle de confiance d'une moyenne.



➤ Test de comparaison de deux moyennes (échantillons indépendants)



Exemple : On voudrait tester les poids moyens suivant le sexe, on suppose que les variances sont homogènes :



Les résultats se présentent comme suit, noter que l'intervalle de confiance de la différence des moyennes peut-être obtenu à l'aide de cette même commande.

```
» Rcmdr » t.test(poids~sexe, alternative='two.sided', conf.level=.95, var.equal=TRUE, data=Dataset)
```

```
Two Sample t-test

data: poids by sexe
t = -2.8564, df = 14, p-value = 0.01269
alternative hypothesis: true difference in means is not equal to 0
95 percent confidence interval:
 -15.320168 -2.179832
sample estimates:
mean in group femme mean in group homme
      66.00           74.75
```

Remarque : le test de comparaison de deux moyennes dans le cas d'échantillons appariés s'obtient à l'aide de la commande *Paired* du sous menu *Mean tests* :

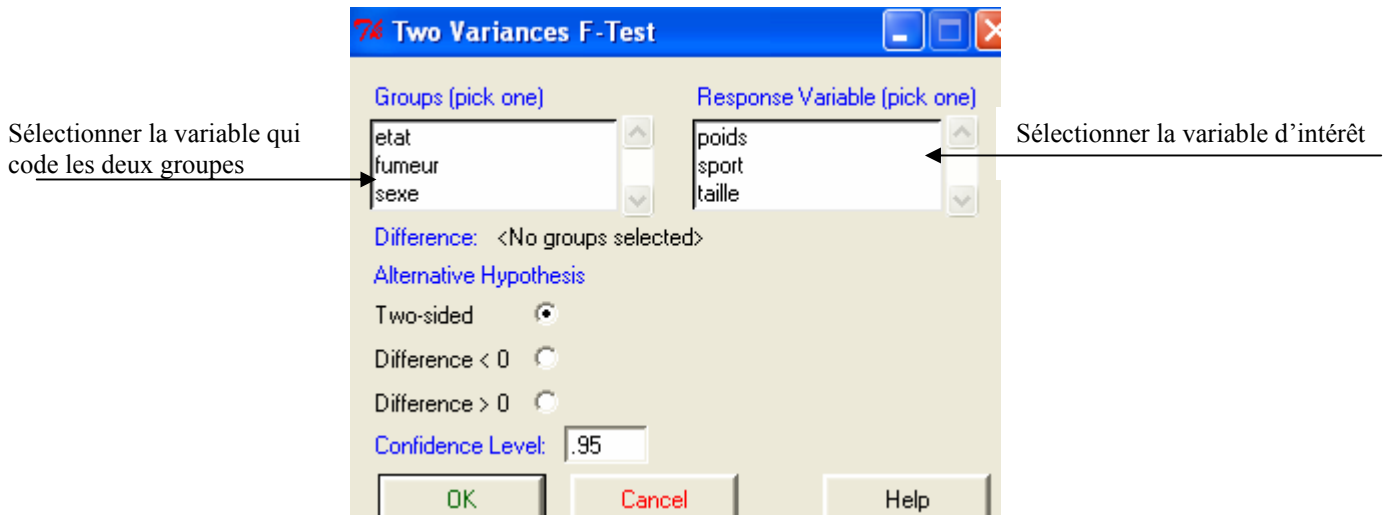
Mean tests:
t-test | Indep. | Paired

5.2.2 Tests sur les variances

Deux tests usuels sont donnés dans le sous menu *Variance tests*. Le test de Fisher dans le cas de deux variances, le test de Bartlett étant une généralisation à plus de deux variances.

Variance tests:
F-test | Bartlett | Levene

En cliquant sur *F-test*, la boîte de dialogue ci-dessous s'affiche à l'écran.



Remarque : Le test de Fisher est un test bilatéral. Le test de Bartlett est un test unilatéral et se déroule de la même manière que celui de Fisher il suffit de cliquer sur Bartlett du sous menu Variance tests

Exemple : On voudrait comparer les variances du poids des individus suivant l'activité sportive (données du fichier data1.txt).



On obtient les résultats suivants, noter la première ligne contient la commande à utiliser en interactif.

```
» Rcmdr » tapply(Dataset$poids, Dataset$sport, var, na.rm=TRUE)
      1      2      3      4
95.33333 45.66667 110.25000 11.58333

» Rcmdr » bartlett.test(poids ~ sport, data=Dataset)

      Bartlett test for homogeneity of variances

data:  poids by sport
Bartlett's K-squared = 3.1869, df = 3, p-value = 0.3637
```

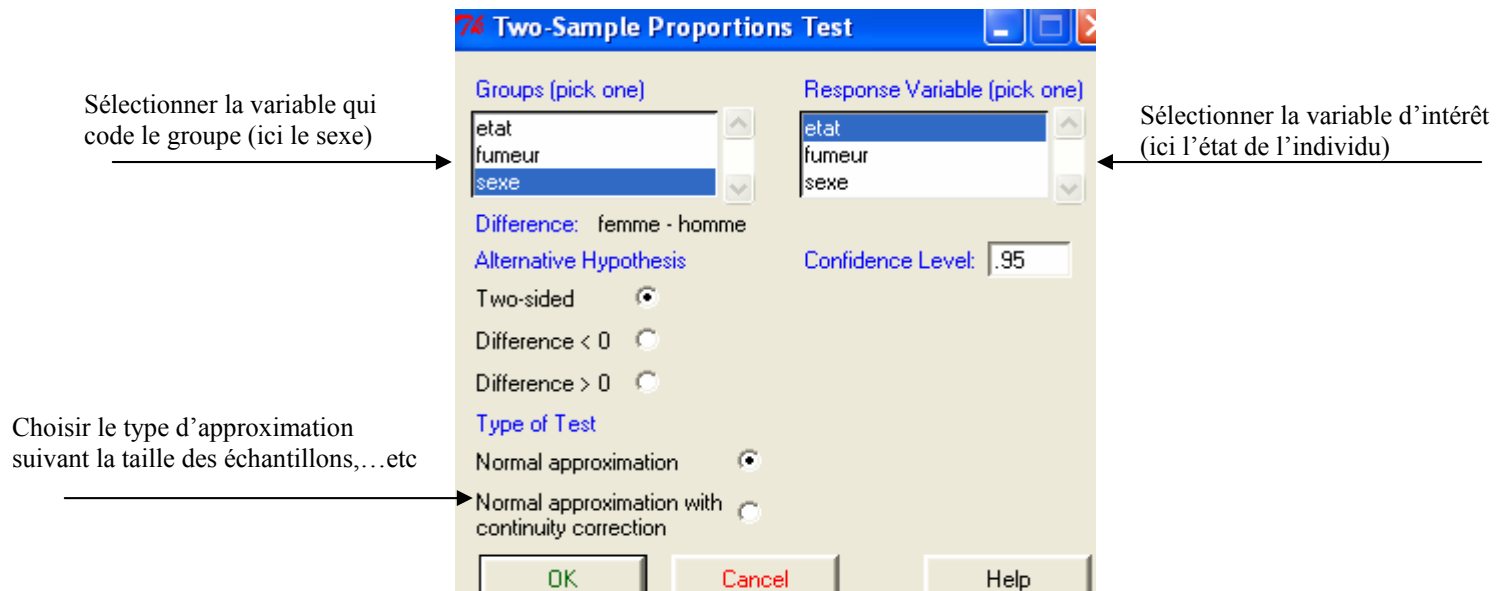
5.2.3 Tests sur les pourcentages

La comparaison d'un pourcentage à une valeur donnée se fait à l'aide de la commande Single-sample du sous menu Proportion tests. La comparaison de deux pourcentages (échantillons indépendants) se fait à l'aide de la commande Two-sample du même sous menu.

Proportion tests:

Single-sample | Two-sample

Exemple : on veut comparer le pourcentage de malades suivant le sexe (données du fichier data1.txt)



Les résultats se présentent comme suit, noter que le calcul de l'intervalle de confiance de la différence des deux pourcentages :

```
» Rcmdr » prop.test(.Table, alternative='two.sided', conf.level=.95, correct=FALSE)
```

```
2-sample test for equality of proportions without continuity
correction
```

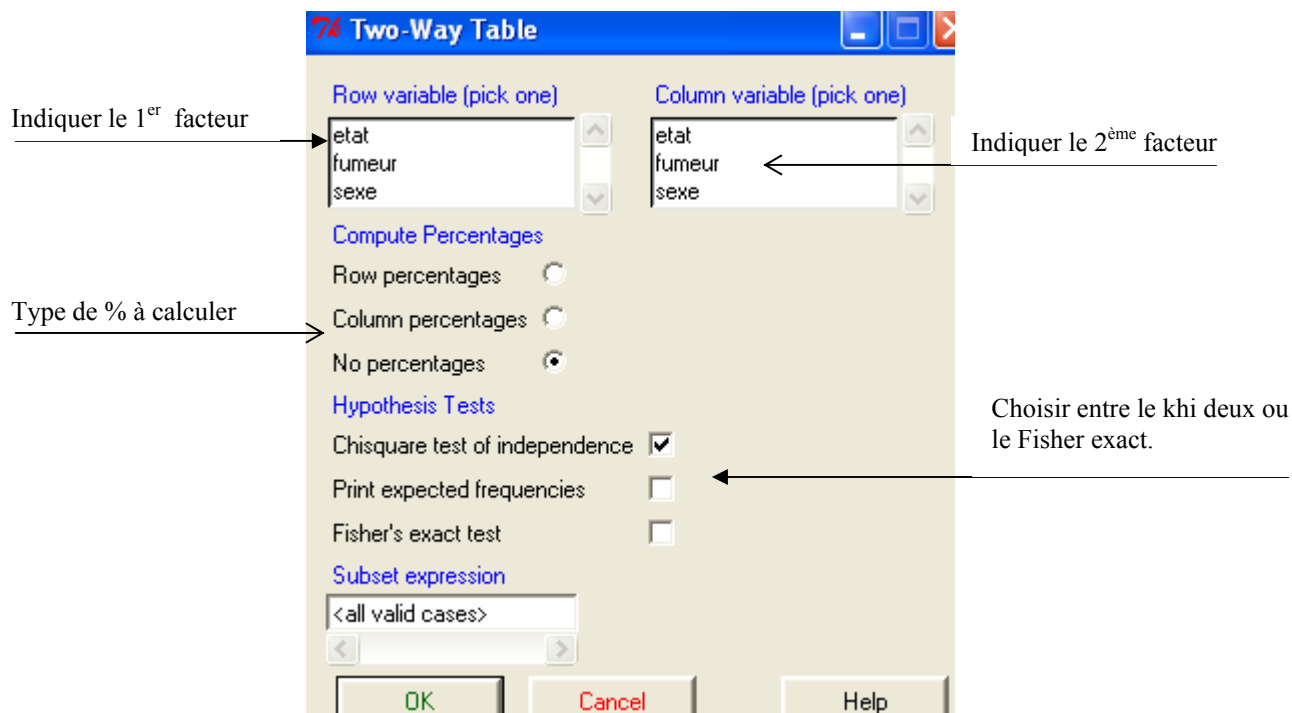
```
data: .Table
X-squared = 1, df = 1, p-value = 0.3173
alternative hypothesis: two.sided
95 percent confidence interval:
 -0.2244317  0.7244317
sample estimates:
prop 1 prop 2
 0.625  0.375
```

5.2.4 Tests du Khi-deux

Dans ce paragraphe nous présentons l'application la plus fréquente du test du khi-deux et qui consiste à tester l'indépendance de deux facteurs. Pour cela il suffit de cliquer sur Two-way du sous menu *Contingency-table* :

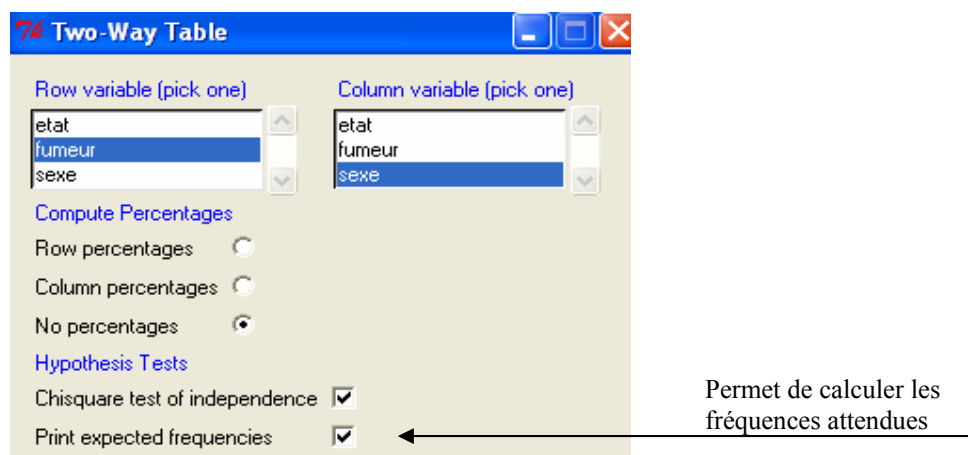
Contingency tables:

Two-way | Enter | Multi-way



Exemple :

On voudrait comparer le pourcentage de fumeur et de non fumer suivant le sexe (données data1.txt) :



Les résultats sont comme suit :

```

» Rcmdr » .Table <- xtabs(~fumeur+sexe, data=Dataset)

» Rcmdr » .Table
      sexe
fumeur femme homme
  non 4      4
  oui 4      4

» Rcmdr » .Test <- chisq.test(.Table, correct=FALSE)

» Rcmdr » .Test

      Pearson's Chi-squared test

data:  .Table
X-squared = 0, df = 1, p-value = 1

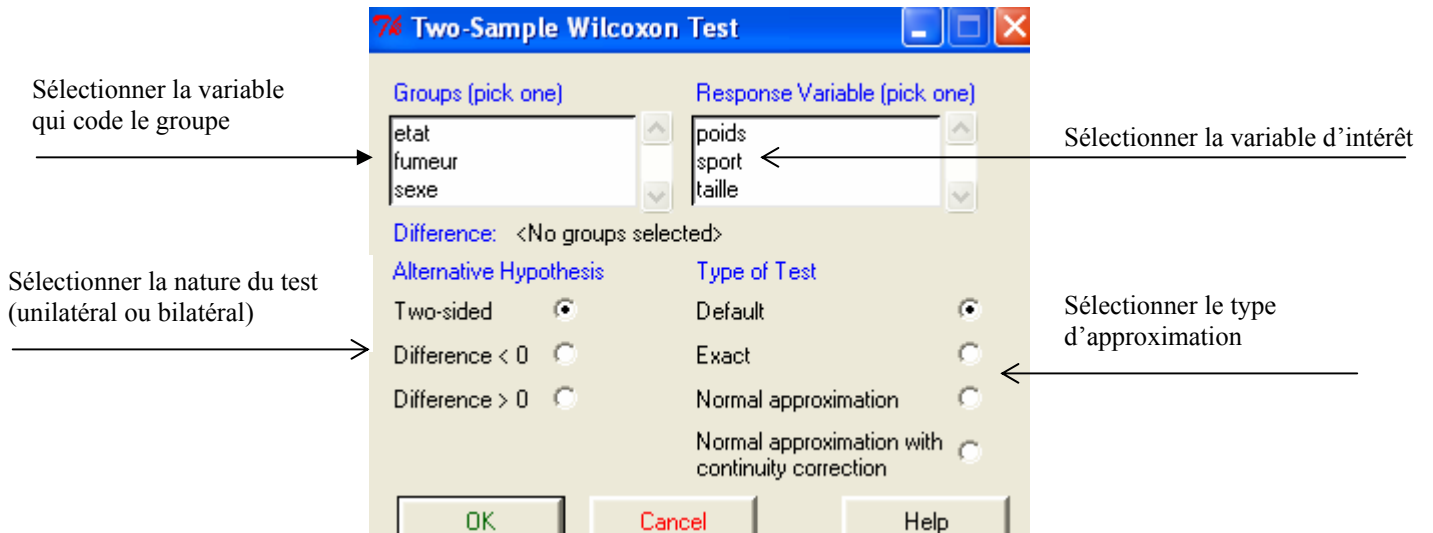
```

Remarque 1 :

Il est possible de saisir la table en interactif, il suffit de cliquer sur la commande Enter du sous menu Contingency tables

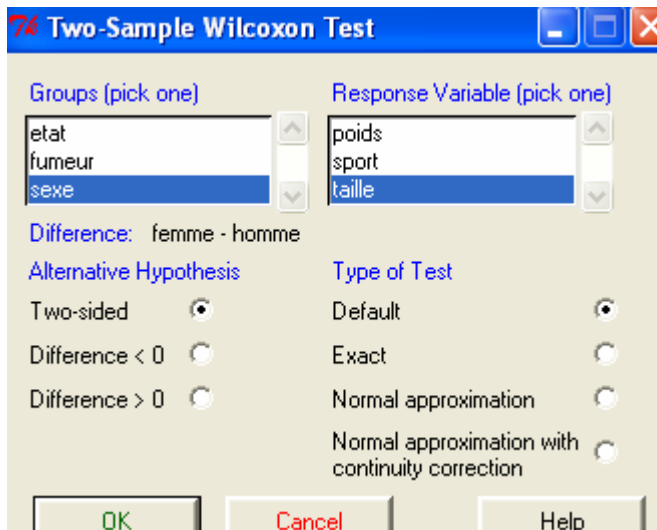
5.3 Tests non paramétriques

Dans ce paragraphe nous présentons les trois tests non paramétriques les plus utilisés, à savoir le test de Wilcoxon pour comparer deux échantillons indépendants, le test de Wilcoxon pour échantillons appariés et enfin le test de Kruskal-Wallis pour plus de deux échantillons indépendants.



Exemple :

On voudrait comparer la distribution des tailles suivant le sexe (données data1.txt) :



Les résultats se présentent sous la forme suivante :

```
» Rcmdr » wilcox.test(taille ~ sexe, alternative="two.sided", data=Dataset)
```

```
Wilcoxon rank sum test with continuity correction
```

```
data: taille by sexe
```

```
W = 14, p-value = 0.06469
```

```
alternative hypothesis: true mu is not equal to 0
```

5.4 ANOVA à un facteur

✓ Première méthode

L'ANOVA à un facteur est obtenue à l'aide la commande *One-Way ANOVA* du menu Mean test. Pour pouvoir effectuer les comparaisons multiples il suffit de cocher la case *pairwise comparisons of mean*.

Exemple : On voudrait comparer les poids moyens suivant l'activité sportive. La variable sport étant déclarée comme facteur à 4 niveaux.

Sélectionner la variable qui code le groupe (facteur) →

→ Sélectionner la variable réponse

```
» anova(lm(poids ~ sport, data=Dataset))
```

```
Analysis of Variance Table
```

```
Response: poids
```

	Df	Sum Sq	Mean Sq	F value	Pr(>F)
sport	3	965.50	321.83	6.6874	0.00664 **
Residuals	12	577.50	48.13		

```
---
```

```
Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
```



```
Call:
simtest.formula(formula = poids ~ sport, data = Dataset, type = "Tukey")
```

Tukey contrasts for factor sport

Contrast matrix:

	sport1	sport2	sport3	sport4
sport2-sport1	0	-1	1	0
sport3-sport1	0	-1	0	1
sport4-sport1	0	-1	0	1
sport3-sport2	0	0	-1	1
sport4-sport2	0	0	-1	1
sport4-sport3	0	0	0	1

Absolute Error Tolerance: 0.001

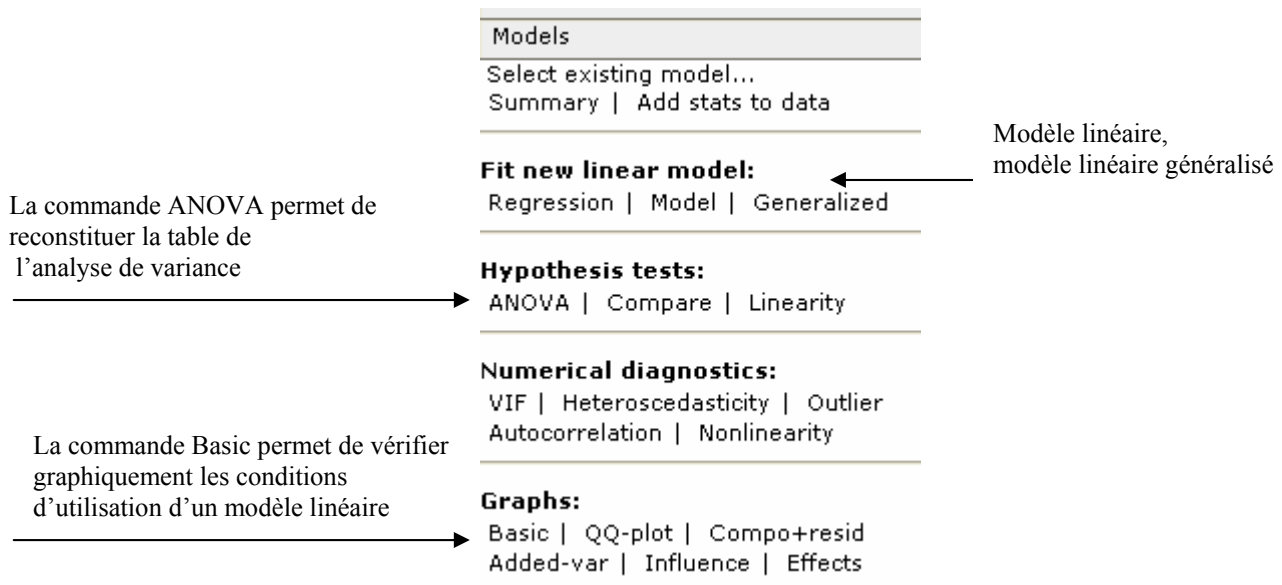
Coefficients:

	Estimate	t value	Std.Err.	p raw	p Bonf	p adj
sport3-sport1	-19.75	-4.026	4.905	0.002	0.010	0.008
sport4-sport1	-17.25	-3.517	4.905	0.004	0.021	0.017
sport2-sport1	-16.00	-3.262	4.905	0.007	0.027	0.022
sport3-sport2	-3.75	-0.764	4.905	0.459	1.000	0.731
sport4-sport3	2.50	-0.510	4.905	0.620	1.000	0.832
sport4-sport2	-1.25	-0.255	4.905	0.803	1.000	0.832

Cette façon de procéder a deux inconvénients, d'une part elle ne permet pas d'obtenir les estimations des paramètres du modèle d'analyse de la variance ($\text{poids} = \mu + \text{sport} + \epsilon$), et d'autre part elle ne fournit pas les résidus dont on a besoin pour vérifier les conditions d'utilisation de l'ANOVA.

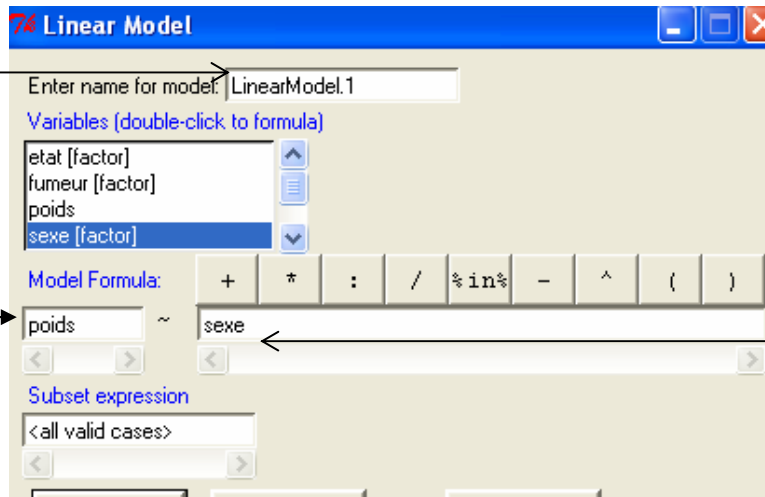
✓ Deuxième méthode

La deuxième façon de faire consiste à utiliser le menu *Models*. Ce menu permet entre autres d'analyser des modèles linéaires, des modèles linéaires généralisés ainsi que les outils graphiques pour vérifier les conditions d'utilisation propres à chaque modèle.



En cliquant sur le sous-menu *model*, la boîte de dialogue ci-dessous s'affiche à l'écran.

Vous pouvez donner un nom au modèle



Sélectionner la variable réponse

Sélectionner le facteur

Les résultats se présentent comme suit, noter qu'en passant par le menu *models*, on dispose de l'estimation des différents paramètres du modèle :

```
lm(formula = poids ~ sexe, data = Dataset)
```

Residuals:

Min	1Q	Median	3Q	Max
-11.000	-3.188	-0.250	4.063	10.250

Coefficients:

	Estimate	Std. Error	t value	Pr(> t)
(Intercept)	66.000	2.166	30.470	3.37e-14 ***
sexe[T.homme]	8.750	3.063	2.856	0.0127 *

Signif. codes: 0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1

Residual standard error: 6.127 on 14 degrees of freedom

Multiple R-Squared: 0.3682, Adjusted R-squared: 0.3231

F-statistic: 8.159 on 1 and 14 DF, p-value: 0.01269

Pour reconstituer la table de l'analyse de la variance il suffit de cliquer sur la commande ANOVA :

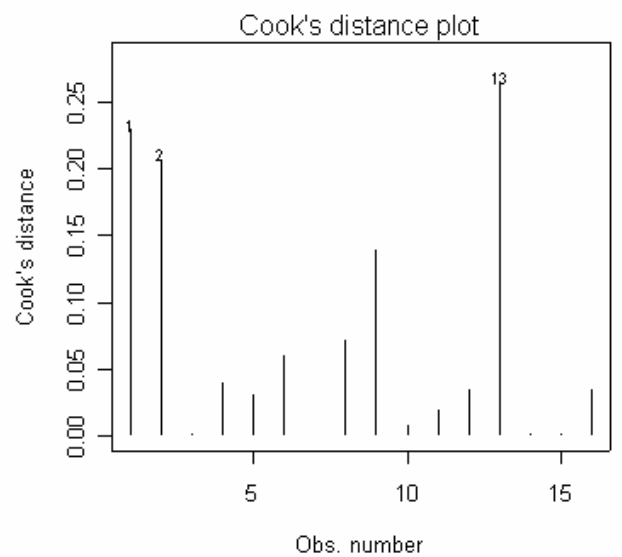
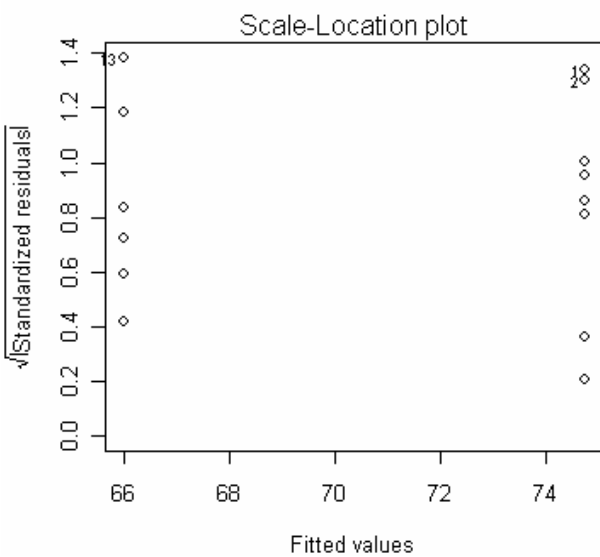
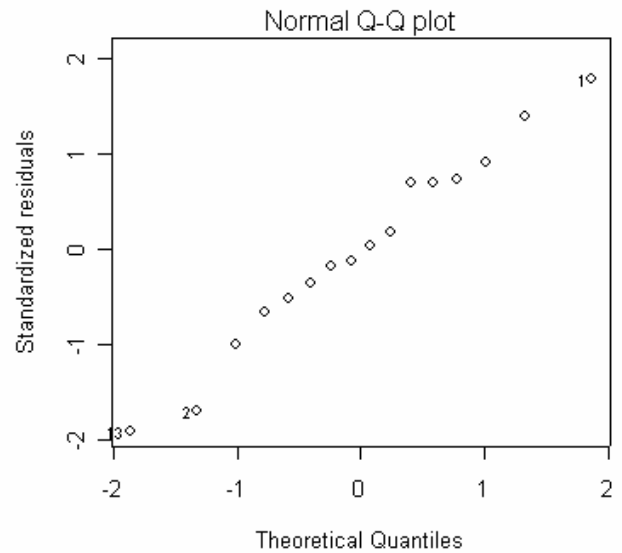
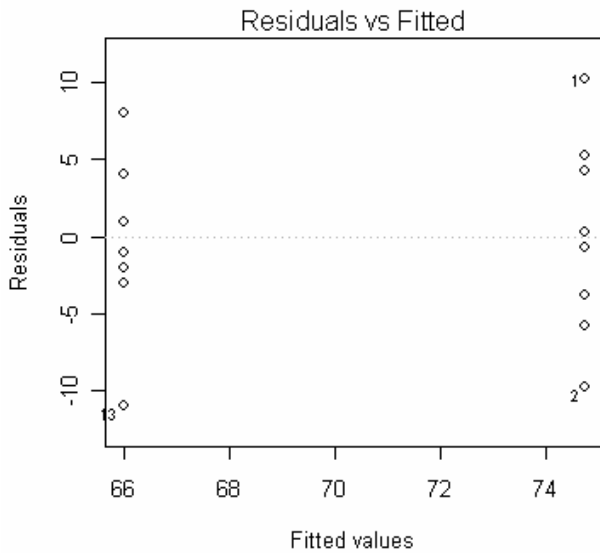
Hypothesis tests:

ANOVA | Compare | Linearity

Pour Vérifier les conditions d'utilisation d'un modèle linéaire il suffit de cliquer sur la commande Basic du menu Graph :

Graphs:

Basic | QQ-plot | Compo+resid
Added-var | Influence | Effects



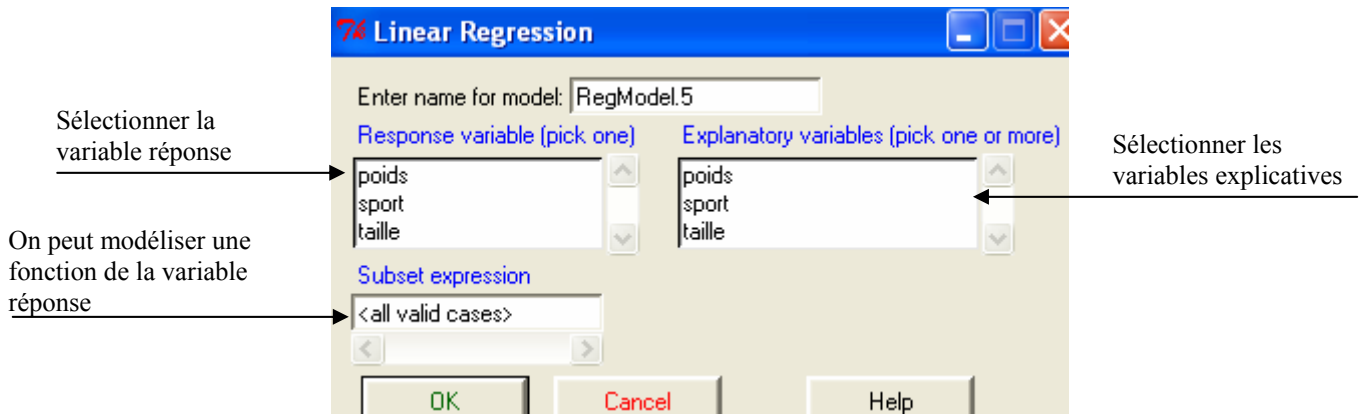
De façon générale, SciViews offre la possibilité de faire des analyses de la variance à plus de deux facteurs avec ou sans interaction (l'interaction est symbolisée par *), des modèles hiérarchiques (niché ou *nested* est symbolisé par %in%)

5.4 Régression multiple

La régression linéaire simple ou multiple se fait à l'aide de la commande *regression* du menu *models* :

Fit new linear model:

Regression | Model | Generalized



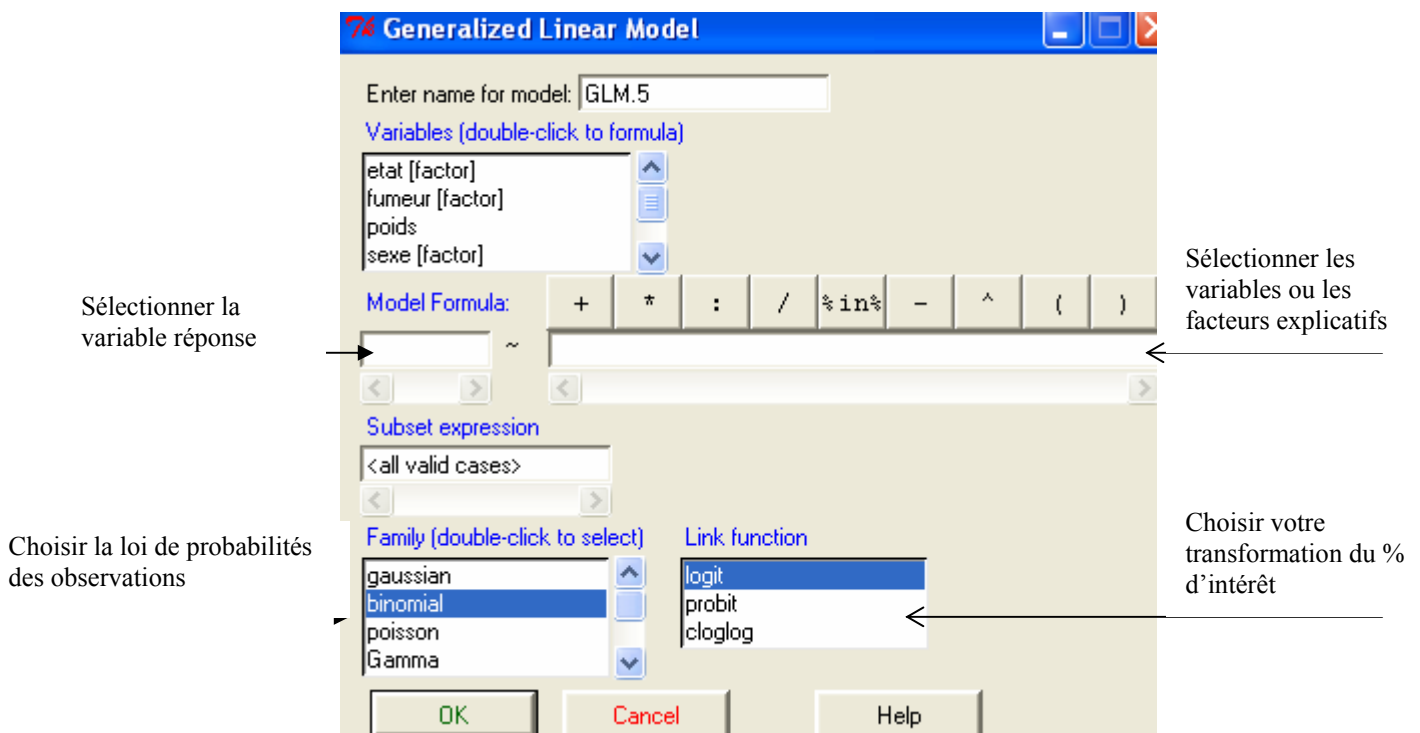
Remarque : La vérification des conditions d'utilisation du modèle linéaire se fait à l'aide de la commande `Graphs/basic`

5.5 Modèle linéaire généralisé

Le modèle linéaire généralisé (logit, probit) se fait à l'aide de la commande `Generalized` :

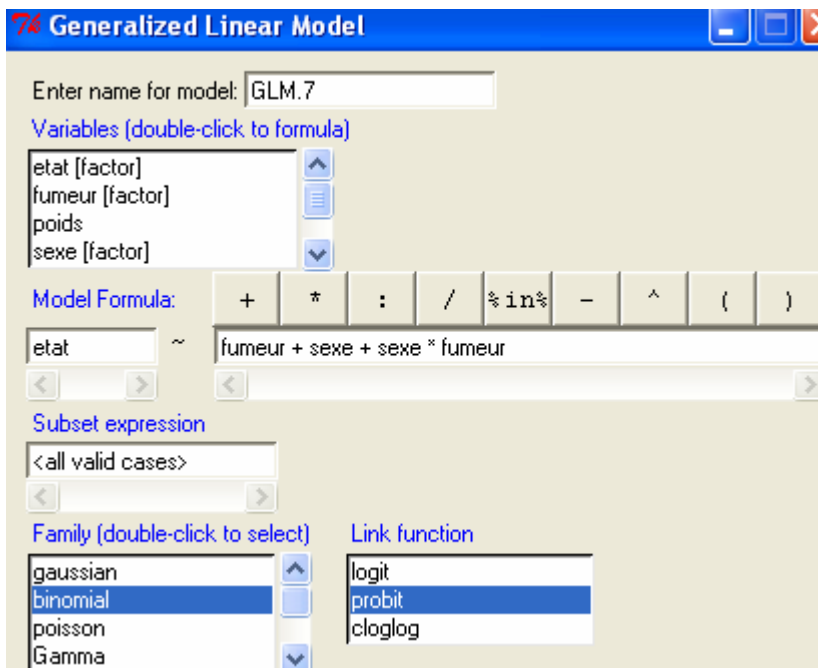
Fit new linear model:

Regression | Model | Generalized



Exemple :

On voudrait expliquer le pourcentage de malade en fonction de l'état et du sexe (données `data1.txt`).



Les résultats sont donnés sous la forme suivante :

```
glm(formula = etat ~ fumeur + sexe + sexe * fumeur, family = binomial(probit),
     data = Dataset)
```

Deviance Residuals:

Min	1Q	Median	3Q	Max
-1.665	-1.177	0.000	1.177	1.665

Coefficients:

	Estimate	Std. Error	z value	Pr(> z)
(Intercept)	8.290e-17	6.267e-01	1.32e-16	1.000
fumeur[T.oui]	-6.745e-01	9.257e-01	-0.729	0.466
sexe[T.homme]	2.256e-16	8.862e-01	2.55e-16	1.000
fumeur[T.oui]:sexe[T.homme]	1.349e+00	1.309e+00	1.030	0.303

(Dispersion parameter for binomial family taken to be 1)

Null deviance: 22.181 on 15 degrees of freedom
 Residual deviance: 20.088 on 12 degrees of freedom
 AIC: 28.088

Number of Fisher Scoring iterations: 4

