

R コマンドー入門*

John Fox

Version 1.7-0 (2011 年 7 月 25 日)

日本語訳†

2011 年 11 月 28 日

1 R コマンドーを起動する

本稿は、R の Windows バージョンの下での R コマンドー (Rcmdr) の利用法を説明する。R が起動しているとき、*R Console* にコマンド `library(Rcmdr)` を入力して Rcmdr をロードすることにより、R コマンドーのグラフィカル・ユーザー・インターフェース (GUI) を起動することができる。Windows のもとで R コマンドーが適切に機能するには、R がシングル・ドキュメント・インターフェース (SDI) の形で設定されている必要がある。^{*1} パッケージをロードした後、*R Console* と R コマンドーのウィンドウは図 1 と図 2 のようになる。本稿におけるスクリーンイメージは、Windows 7 の元で作成している。Windows の他のバージョン (当然、他のプラットフォームも) を利用している場合は、スクリーンの様子は異なる可能性がある。^{*2}

R コマンドーと *R Console* のウィンドウは、デスクトップ上で自由に移動させることができる。R コマンドーのメニューとダイアログボックスを利用して、データを読み込み、処理、分析を行うことができる。

- R コマンドーの GUI で利用された R のコマンドは、R コマンドーの主ウィンドウの上部にあるテキストウィンドウ (スクリプトウィンドウという) に表示される。スクリプトウィンドウまたは *R Console* のプロンプト (`>`) の位置に直接キー入力してもよい。しかし、R コマンドーの開発目的は、コマンドを入力しないで済むようにすることにある。

* 本マニュアルは、Fox(2005) の改訂版である。連絡は jfox@mcmaster.ca まで。

† 翻訳版における画像等は、R-2.13.1, Rcmdr-1.7-0 に基づく。本翻訳に関する問い合わせ等は、荒木孝治 (arakit@kansai-u.ac.jp) まで。

^{*1} R の Windows バージョンは、通常、マルチ・ドキュメント・インターフェース (MDI) で動く。それは、*R Console* ウィンドウ、セッション中に作られるグラフィカル・デバイス・ウィンドウ、他の R のプロセスに関連したウィンドウを 1 つの主ウィンドウに含む。これに対してシングル・ドキュメント・インターフェース (SDI) では、*R Console* とグラフィカル・デバイスは主ウィンドウとは独立して表示される。R を SDI モードで動かすにはいくつか方法がある。例えば、R の `etc` サブディレクトリにある `Rconsole` ファイルを編集するか、R のデスクトップアイコンのプロパティのショートカットタブのリンク先に「`--sdi`」を追加すればよい。Rcmdr パッケージの限界は、`tcltk` パッケージの限界による。

^{*2} Rcmdr は、R とともに配布されている推奨パッケージに加えていくつかのパッケージを必要とし、Rcmdr の起動とともにこれらをロードする。Rcmdr および必要なパッケージ、他の多くの追加パッケージは、Comprehensive R Archive Network (CRAN) (<http://cran.r-project.org/>) よりダウンロードできる。

Windows の「R GUI」から Rcmdr をインストールすると、Rcmdr が依存するパッケージ全てがインストールされるとは限らない。これらがインストールされていないと、Rcmdr はインターネットまたはローカルファイル (例えば、CD-ROM) からインストールするかどうか聞いてくる。これが Rcmdr をインストールする推奨方法である。他方、関数 `install.packages` を用いて Rcmdr をインストールするときに、引数を `dependencies = TRUE` と設定しておくことにより、依存するパッケージ全てをインストールすることもできる。しかし、この方法は、必要以上のパッケージをインストールする可能性がある。

Dirk Eddelbuettel のおかげで、Debian Linux のユーザーはコマンド `$ apt-get install r-cran-rcmdr` を用いるだけで、Rcmdr と他の必要なパッケージ全てをインストールすることができる。いずれにしても、Linux システム上でパッケージ Rcmdr をビルドし、インストールする方が一般に簡単である。Mac OS X での作業はより面倒である。それは、Rcmdr が依存するパッケージ `tcltk` が X-Windows のために `Tcl/Tk` のインストールを要求し、R が X-Window とともに作動するからである。

詳細は、<http://socserv.socsi.mcmaster.ca/jfox/Misc/Rcmdr/index.html> にある R コマンドーのインストールに関するメモを参照のこと。

- 計算等の結果は，出力ウィンドウという 2 番目のテキストウィンドウに表示される．
- 下部にあるグレーのウィンドウ（メッセージというラベルがついているウィンドウ）は，エラーメッセージや警告，あるいは，図 2 に示すスタートアップメッセージといった情報を表示する．
- グラフを作成すると，別のグラフィックスデバイスのウィンドウに表示される．

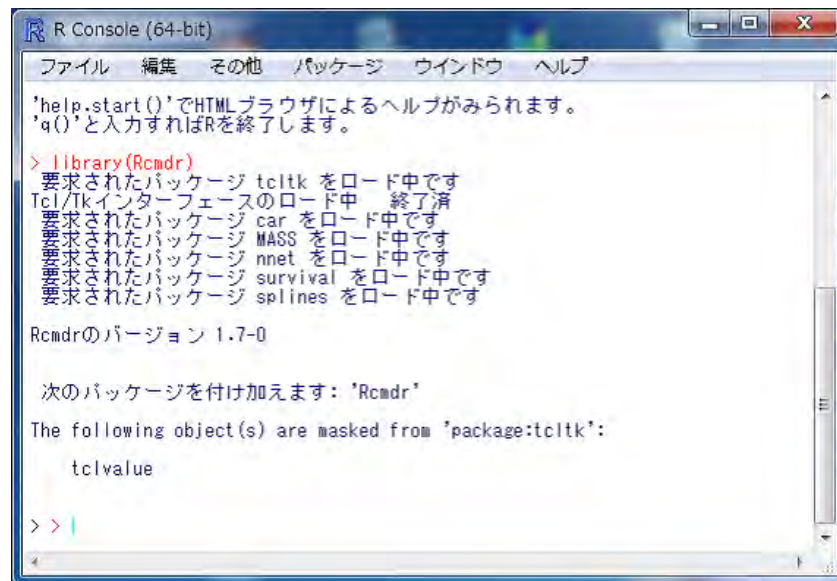


図 1 Rcmdr ロード後の R Console ウィンドウ

R コマンダーの上部には次のようなメニューがある．

- ファイル スクリプトファイルを読み込んだり保存したりする．出力や R ワークスペースの保存，終了の機能を持つ．
- 編集 スクリプトウィンドウと出力ウィンドウの内容を編集（切り取り，コピー，ペースト他）するためのメニュー．スクリプトウィンドウまたは出力ウィンドウで右クリックすると，編集のコンテキストメニューが表示される．
- データ データの読み込み，データ処理のためのメニュー．
- 統計量 基本的な統計分析を行うためのメニュー．
- グラフ 簡単な統計グラフを作成するためのメニュー．
- モデル 統計モデルに対する数値による要約，信頼区間，仮説検定，診断，グラフのためのメニュー．残差といった診断の統計量をデータセットに追加する機能を持つ．
- 分布 標準的な分布の累積確率，確率密度，分位点を求め（数値表の代わりに用いることができる），グラフを作成する．分布からのサンプルを得ることも可能．
- ツール Rcmdr とは関係のないパッケージをロードしたり（例えば，他のパッケージに付属するデータセットにアクセスする），Rcmdr のプラグインパッケージを起動したり，オプションを設定したりするためのメニュー．
- ヘルプ R コマンダーの情報（このマニュアルを含む）を得るためのメニュー．なお，R コマンダーの各ダイアログボックスは *Help* ボタンを持つ（下記参照）．

R コマンダー（バージョン 1.7-0）の完全なメニューツリーを以下に示す．ほとんどのメニュー項目では，後で例示するように，ダイアログボックスを表示する．表示した状況において利用できないメニューはグレイ表示され，選択できないようになっている．

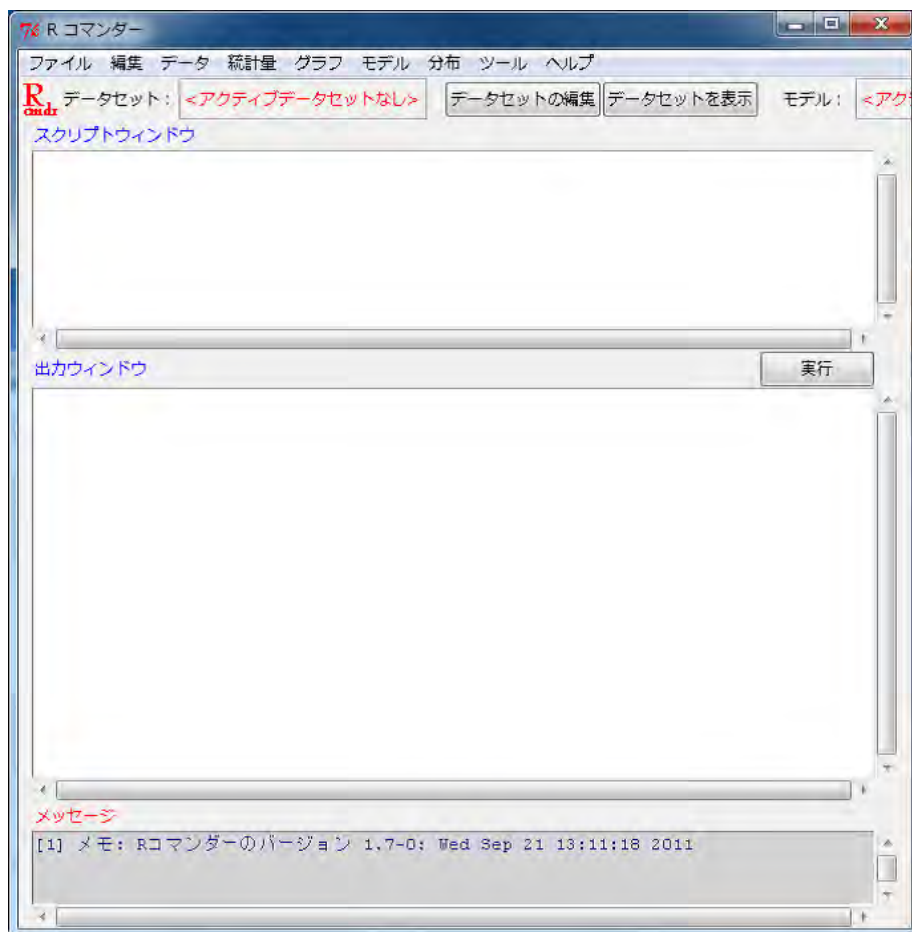
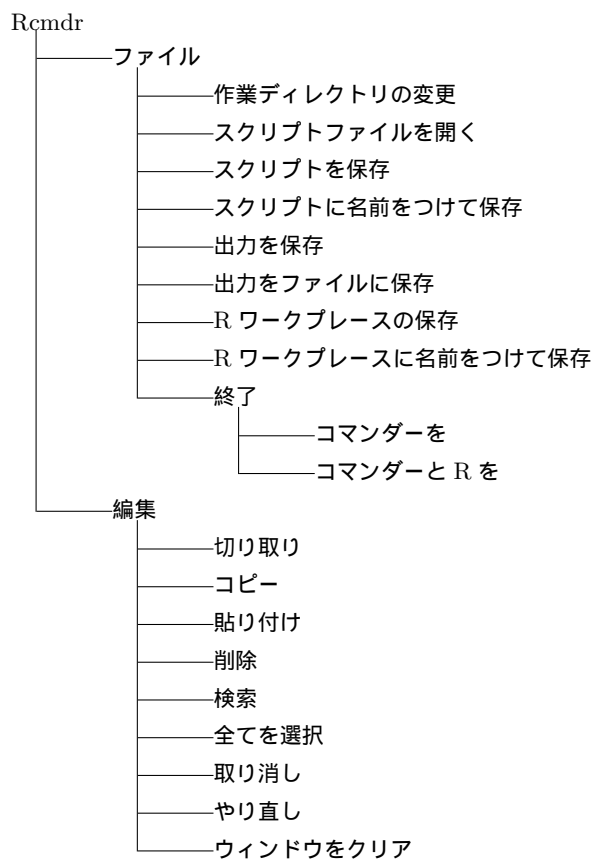
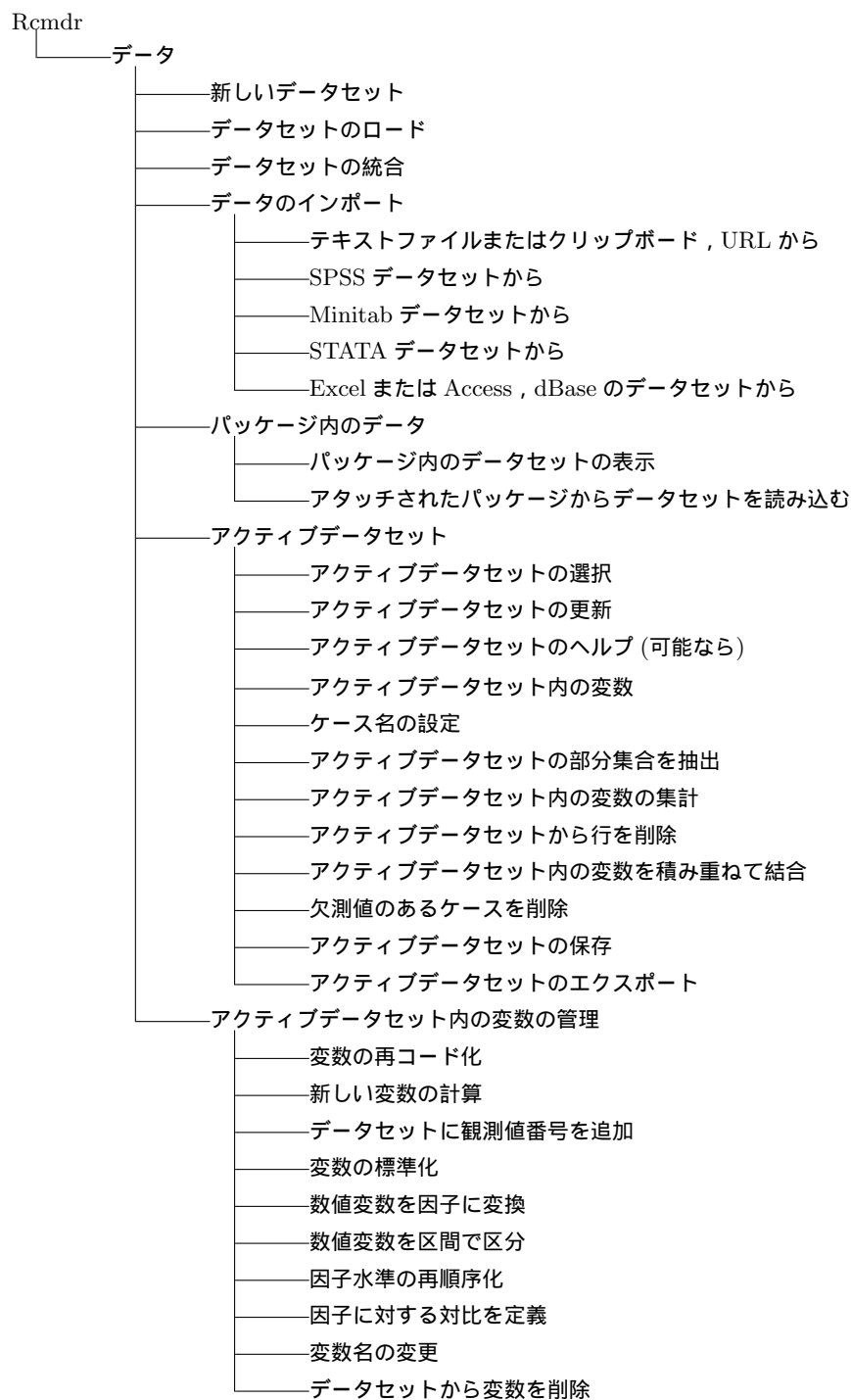
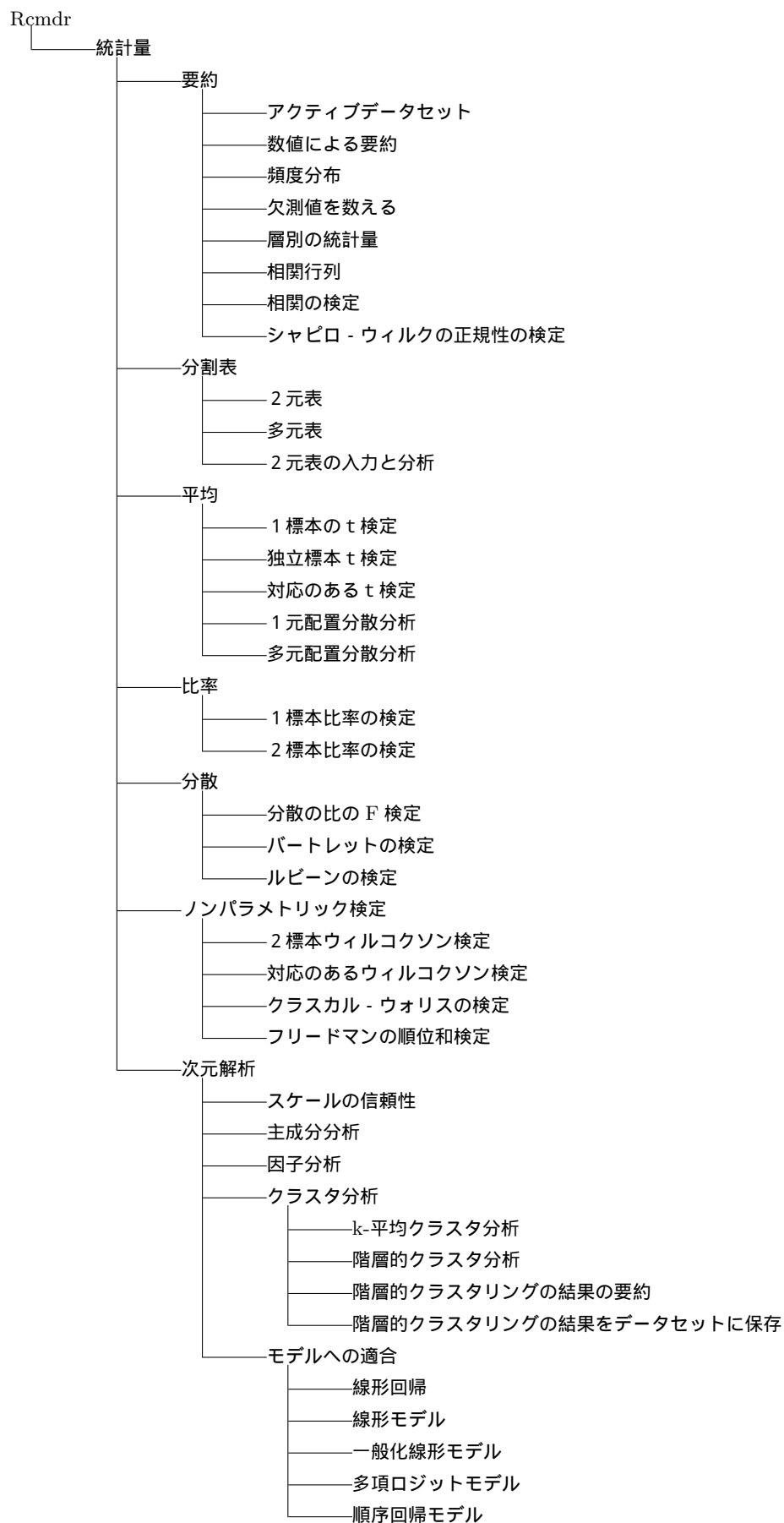
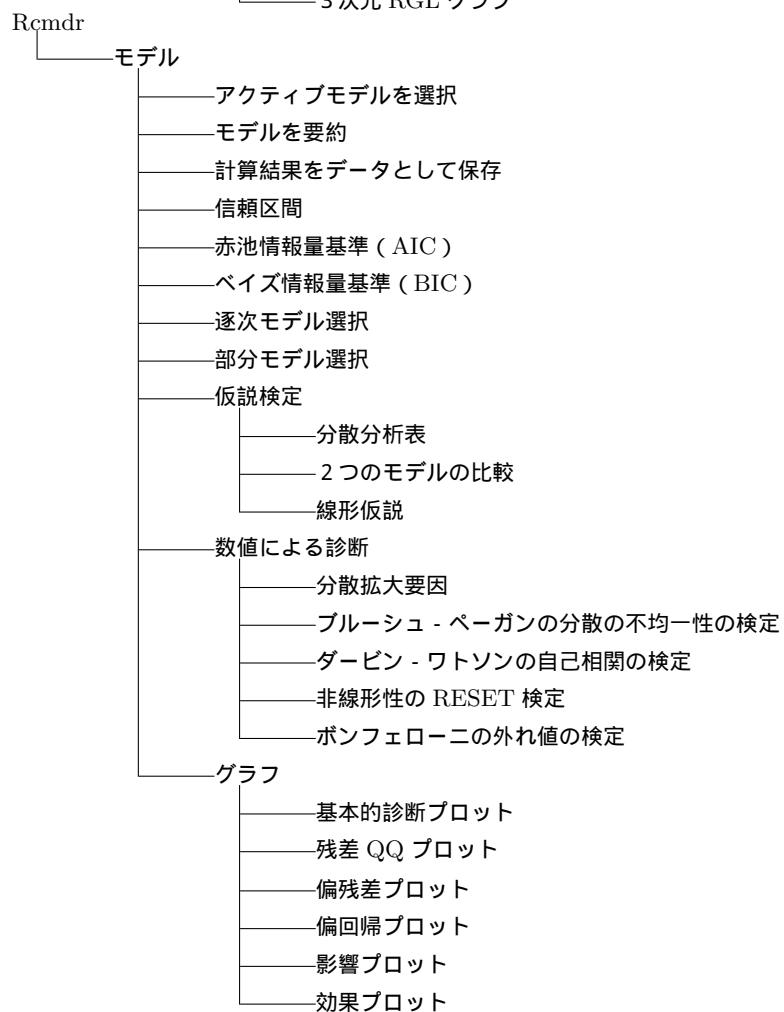
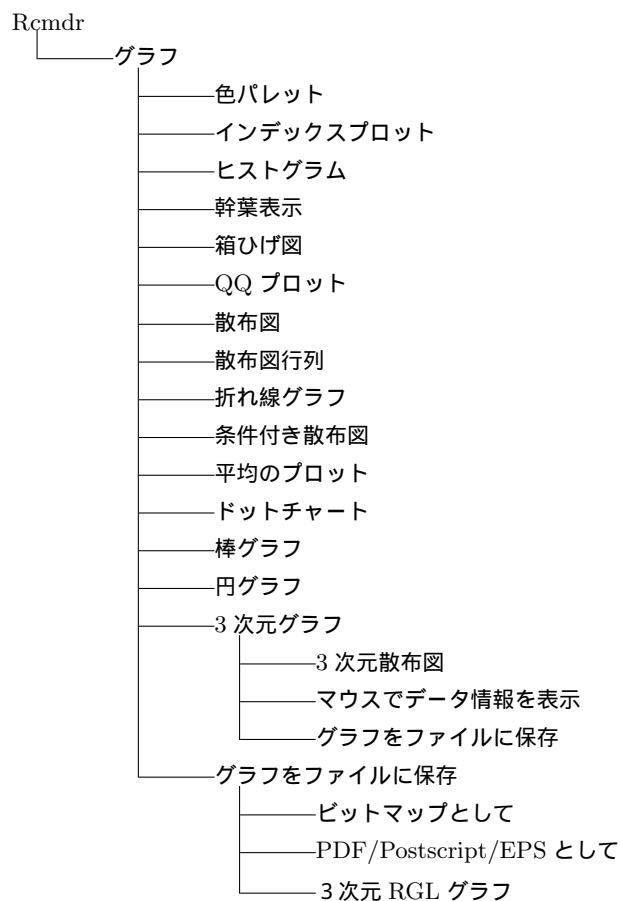


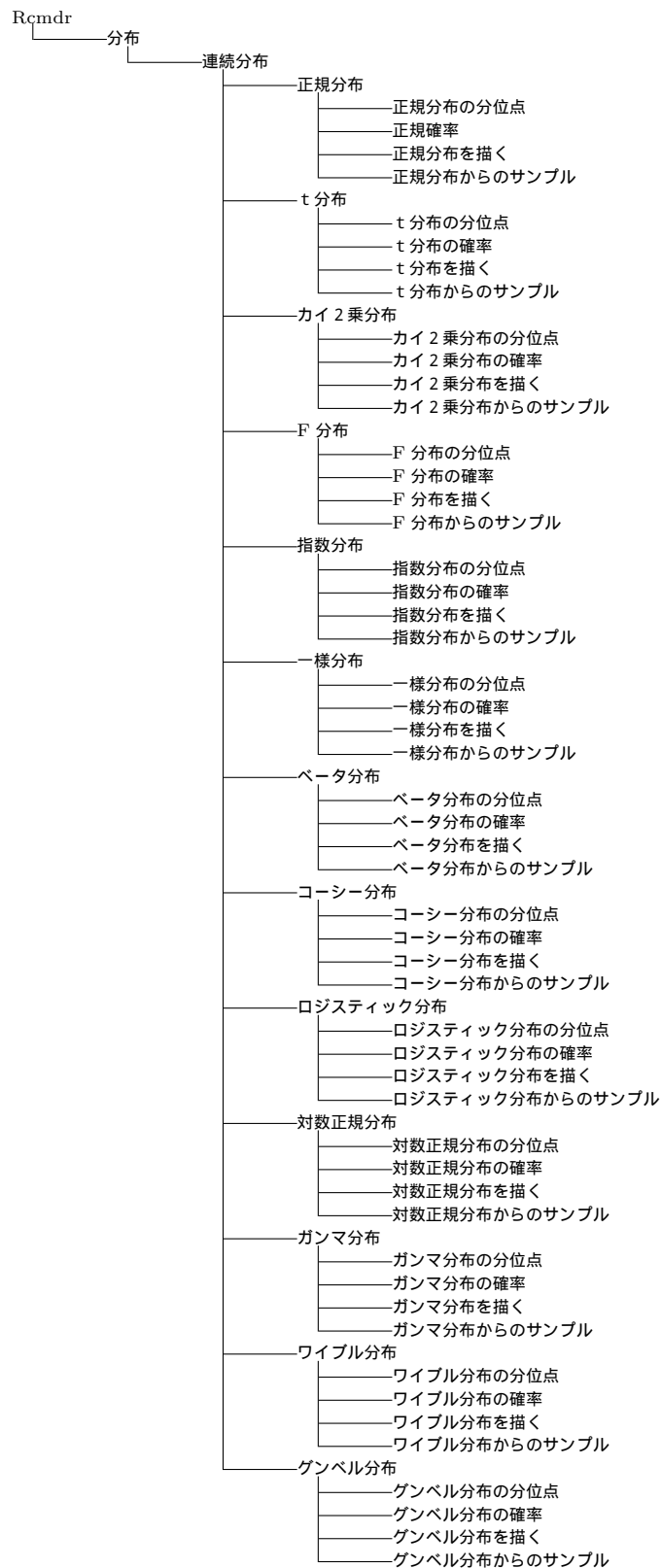
図 2 起動直後の R コマンダーのウィンドウ

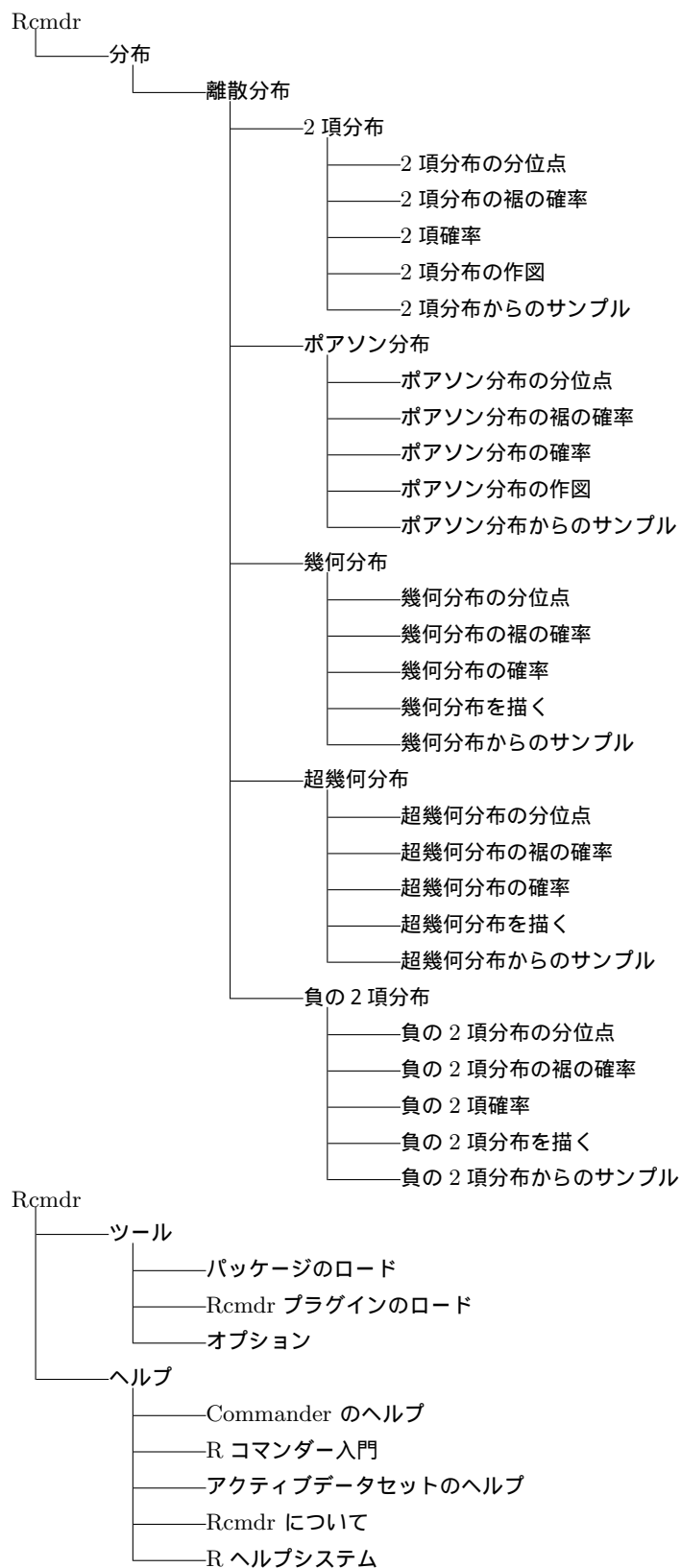












R コマンダーのインターフェースは、メニューとダイアログに加えて、他にいくつかの要素を持つ。

- メニューの下にボタンが並んだツールバーがある。
 - － 最も左にある（フラット）ボタンはアクティブデータセットの名前を表示する。最初、アクティブデータセットはない。このボタンを押すことにより、現在メモリにあるデータセットを（2 つ以上あれば）選択することができる。R コマンダーのメニューとダイアログのほとんどは、アクティブデー

タセットを対象とする（ファイル，編集，分布メニューを除く）。

- アクティブデータセットを編集したり表示したりするためのボタンが2つある。他のオペレーションが実行されているとき，データセットビューを開いたままにしておくことができる。^{*3}
- フラットボタンには，アクティブな統計モデル — 線形モデル（線形回帰モデル），一般化線形モデル，多項ロジットモデル，比例オッズモデル — の名前が表示される。^{*4} 最初，アクティブモデルはない。メモリにモデルが2つ以上あるとき，このボタンをクリックすることによりモデルを選択することができる。
- ツールバーの直下に，スクロール可能な大きなテキストウィンドであるスクリプトウィンドウ（この名前が表示されている）がある。GUIで生成されたコマンドは，このウィンドウに表示される。スクリプトウィンドウのテキストを編集したり，Rのコマンドをキー入力したりすることができる。スクリプトウィンドウの右下にある実行ボタンをクリックする（*Ctrl-r* でもよい）ことにより，カーソルがある行のスクリプトを実行することができる。ドラッグして複数行を選択した場合，実行ボタンにより全てを実行することができる。スクリプトウィンドウに入力したコマンドは，複数行に渡ってもよいが，そのときは，2行目以降を1つ以上の空白またはタブにより字下げする必要がある。キー入力の組合せ *Ctrl-a* により，スクリプトウィンドウ内の全てのテキストを選択することができ，*Ctrl-s* によりウィンドウの内容を保存するためのダイアログボックスを表示することができる。
- スクリプトウィンドウの下に，スクロール可能で編集可能な出力のためのテキストウィンドウがある。このウィンドウに送られたコマンドは赤で，出力結果はダークブルーで表示される（*R Console* と同じ）。
- スクリプトウィンドウの下には，メッセージを表示するための小さなグレーのテキストウィンドウがある。エラーメッセージは赤で，警告は緑で，他のメッセージはダークブルーで表示される。エラーと警告はベル音によっても知らされる。

パッケージ *Rcmdr* を起動した後，*R Console* を最小化してもよい。R コマンダーのウィンドウの大きさは，通常の方法で変更したり最小化したりできる。R コマンダーの大きさを調整すると，以降の出力は出力ウィンドウの大きさに自動的に調整される。

R コマンダーでは設定を柔軟にできる。ここで示したのは標準設定である。設定の変更は，ツール → オプションメニューを用いるか，より徹底的に行うには，R のオプションで行う。^{*5} 詳細については，*Rcmdr* のヘルプファイルを参照のこと。

2 データの入力

R コマンダーのほとんどの手続きは，アクティブデータセットがあるという前提で実行される。^{*6} メモリに複数のデータセットがある場合，1つのみを選択することができる。R コマンダーを起動した直後には，アクティブデータセットはない。

R コマンダーによるデータの入力方法にはいくつかある。

- データ → 新しいデータセット... より，直接入力することができる。データセット数が少ないときは，これでよい。

^{*3} David Firth のパッケージ *relimp* にある関数 *showData* によるデータビューは，多くの変数を持つデータセットを表示することができる。変数の数の限界（初期設定は100）を越えた場合，データセットを表示する代わりに R のデータビュー（View）を利用する。変数の数に関係なく R のデータビューを利用するには，限界値を0に設定する。詳細については，R コマンダーのヘルプファイルを参照のこと。データセットを表示するのに R のデータビューを利用することの欠点は，他のオペレーションが実行されている間，エディタウィンドウが表示されないことである。

^{*4} 必要なダイアログボックスとメニューアイテムを準備し，R の *etc* ディレクトリにある *model-classes.txt* を編集することにより，別の統計モデルを付加することができる。

^{*5} 省略記号 (...) が付いているメニュー項目を選択すると，ダイアログボックスが表示される。これは標準的な GUI の仕様である。本稿では，→ はメニュー項目やサブメニューを選択することを意味する。

^{*6} 分布メニューで選択された手続きはそうではない。例えば，メニュー統計量 → 分割表の2元表を入力して分析... の場合である。

- テキスト（“ascii”）ファイルまたはクリップボード，他の統計パッケージ（Minitab や SPSS，Stata），Excel または Access，dBade のデータをインポートすることができる．
- R のパッケージに含まれるデータセットを読み込むことができる．その名前を知っている場合はそれをキー入力するか，ダイアログボックスから選択する．

2.1 テキストファイルからデータを読み込む

例として，データファイル Nations.txt を取り上げる．*7 このデータの最初の数行は次のようになっている．

```
TFR contraception infant.mortality GDP region
Afghanistan      6.90    NA   154   2848   Asia
Albania           2.60    NA    32    863   Europe
Algeria           3.81    52    44   1531   Africa
American-Samoa    NA      NA    11    NA   Oceania
Andorra           NA      NA    NA    NA   Europe
Angola            6.69    NA   124    355   Africa
Antigua           NA      53    24   6966   Americas
Argentina         2.62    NA    22   8055   Americas
Armenia           1.70    22    25    354   Europe
Australia         1.89    76     6  20046   Oceania
. . .
```

- ファイルの第 1 行に変数名がある．これらは，TFR（出生率で，女性 1 人当たりの子供の数），contraception（既婚女性当たりの避妊具利用率（％）），infant.mortality（出生児 1000 人当たりの乳児死亡率），GDP（1 人当たり国民総生産，単位は US ドル），region である．
- 2 行目以下には，国単位でデータ値が入力されている．データ値は余白（1 つ以上の空白またはタブ）で区切られている．データ値は縦に並んでいると見やすいが，そうである必要はない．データ行が国名で始まっていることに注意．これをデータセットの行名としたいので，国名に対応する変数名を入れていない．すなわち，変数名は 5 つだが，データ値は 6 つある．このような場合，R は各行の最初の値を行名として取り扱う．
- データ値には欠測値がある．R では，欠測値のコードとして NA（not available の意味）を用いるのがよい．
- TFR，contraception，infant.mortality，GDP は数値（量的）変数である．これに対して，region には地域名が入力されている．これが読み込まれると R は region を因子，つまり質的変数として取り扱う．R コマンドは，数値変数と因子とを区別する．

データファイルを R に読み込むには，R コマンドのメニューから，データ → データのインポート → テキストファイルまたはクリップボード，URL から... を選択する．この操作により，図 3 に示すテキストファイルまたはクリップボード，URL からデータを読み込むというダイアログボックスが表示される．データセットのデフォルト名は Dataset であるが，Nations に変更している．

R において，データセット名は，大文字または小文字のアルファベット（または，ピリオド“.”で始まり，以降，全アルファベット，アンダースコア（_），数字（0-9）で構成される必要がある．空白を用いることはできない．また，R では大文字と小文字を区別する．そのため，nations，Nations，NATIONS 等は区別され，異なるデータセット名となる．

*7 このファイルは，パッケージ Rcmdr のサブディレクトリ etc にある．

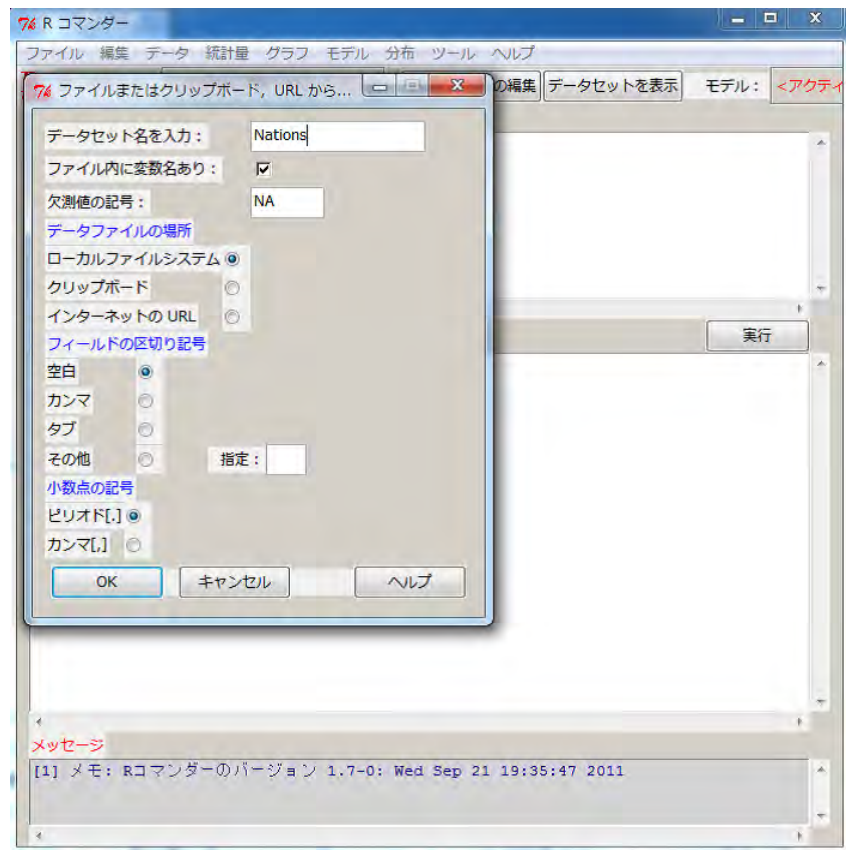


図3 Rcmdrのロード後の *R Console* ウィンドウ

テキストファイルからデータを読み込むダイアログで *OK* ボタンをクリックすると、図4に示すファイルを開くダイアログが表示される。ここでは、*Nations.txt* ファイルを読み込む状況を示している。ダイアログの開くボタンをクリックすると、データファイルが読み込まれる。データファイルが読み込まれると、それは *R* コマンドーのアクティブデータセットとなる。結果として、図5に示すように、読み込まれたデータセット名が *R* コマンドーのウィンドウの左上部にあるデータセットボタンに表示される。

データセットを表示ボタンをクリックすると、図5に示すようなデータビューウィンドウが表示される。データセット *Nations* を読み込み、それを表示するコマンド (*read.table* と *showData*) が、スクリプトウィンドウと出力ウィンドウに表示されていることに注意（データセットの表示により少し隠れてわかりにくい）。データセットが読み込まれてアクティブデータセットになると、メッセージウィンドウにコメントが表示される（さらにコマンド *showData* が実行されると消える）。

コマンド *read.table* は *R* の“データフレーム”を作る、これは、行をケース、列を変数とする表形式のデータセットのオブジェクトである。行は、ケースまたは観測対象を表し、列は変数である。*R* コマンドーのデータセットは、*R* のデータフレームである。

2.2 データを直接入力する

R の表計算に似たデータエディタから直接データを入力するには、次のようにする。例として、Moore (2000) の Problem 2.44 からの非常に小さいデータセットを用いる。

- *R* コマンドーのメニューから、データ → 新しいデータセット... を選択する。オプションとして、データセットの名前を、例えば *Problem2.44* をダイアログボックスに入力し、*OK* をクリックする。（*R* のデータセット名に空白を使用できないことに注意。）これにより、何も入力されていないデータエディタのウィンドウが表示される。

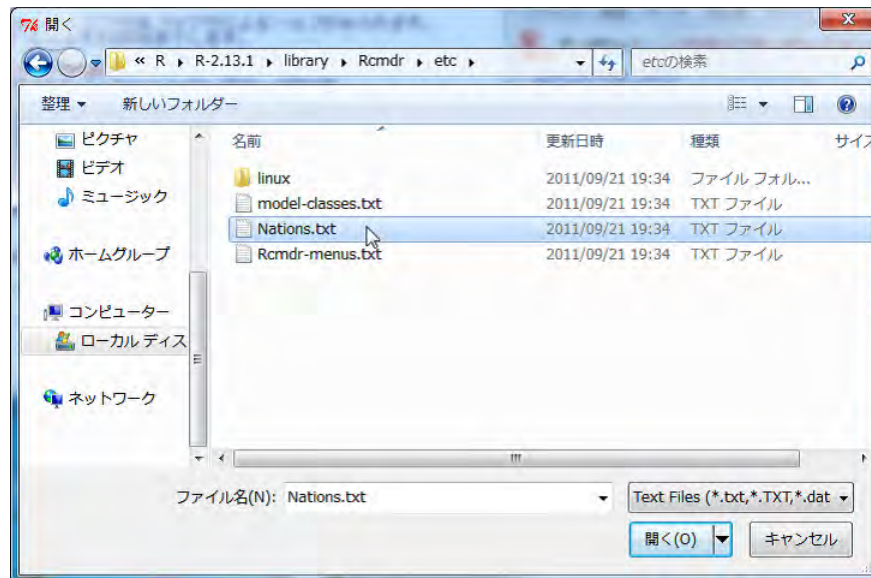


図 4 テキスト形式データファイルを読み込むためのファイルを開くダイアログボックス

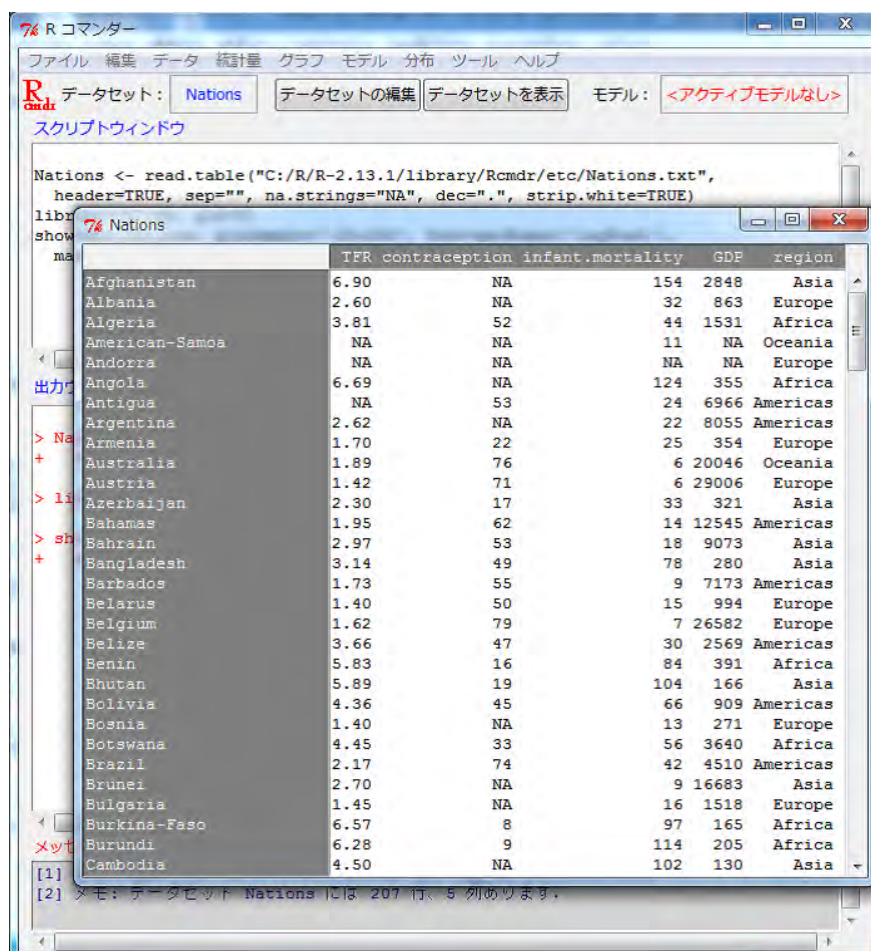


図 5 アクティブデータセットの表示

- データエディタの最初の 2 列にデータを入力する．入力するセルを移動するには，キーボードにある矢印キーやタブキー，Enter キーを用いたり，マウスでポインターを移動して左クリックしたりする．データの入力が終了すると，図 6 のようになる．
- 次に，第 1 列の上部にある名前 var1 をクリックする．これにより，図 7 のような変数エディタのダイアログボックスが表示される．
- 変数名 age を入力し，変数エディタのウィンドウの右上隅にある ×（閉じる）ボタンをクリックするか，Enter キーを押してウィンドウを閉じる．同様にして 2 列目の変数名を height に変更する．データエディタは図 8 のようになる．
- データエディタのメニューよりファイル → 閉じるを選択するか，データエディタの右上にある × ボタンをクリックする．これにより，入力したデータセットは R コマンダーのアクティブデータセットとなる．



	var1	var2	var3	var4	var5	var6
1	36	86				
2	48	90				
3	51	91				
4	54	93				
5	57	94				
6	60	96				
7						
8						
9						
10						
11						
12						
13						
14						
15						
16						
17						
18						
19						

図 6 データ入力後のデータエディタ

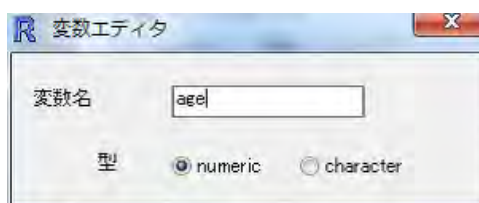


図 7 データエディタの変数名の変更のためのダイアログボックス

2.3 パッケージからデータを読み込む

パッケージの多くにはデータが含まれる．パッケージ内のデータセットは，データ → パッケージ内のデータ → パッケージ内のデータセットの表示によりポップアップウィンドウに表示できる．また，データ → パッケージ内のデータ → アタッチされたパッケージからデータセットを読み込むにより R コマンダーに読み込むことができる．*8 結果として表示されるダイアログボックスを図 9 に示す．パッケージ内のデータセットの名前を知っているときは，それを直接入力してもよい．そうでなければ，パッケージ名をダブルクリックすると

*8 パッケージ内のデータセットは必ずしも全てがデータフレームではない．データフレームのみが R コマンダーに適していることに注意．データフレームでないデータを読み込もうとすると，メッセージウィンドウにエラーメッセージが表示される．

右のリストボックスにデータセット名が表示される．データセット名をダブルクリックすると，その名前がダイアログ中のデータセットを入力欄にコピーされる．*⁹ R の他のパッケージにアクセスするには，ツール → パッケージのロード によって行う．



図 8 2 つの変数名を変更した後のデータエディタウィンドウ

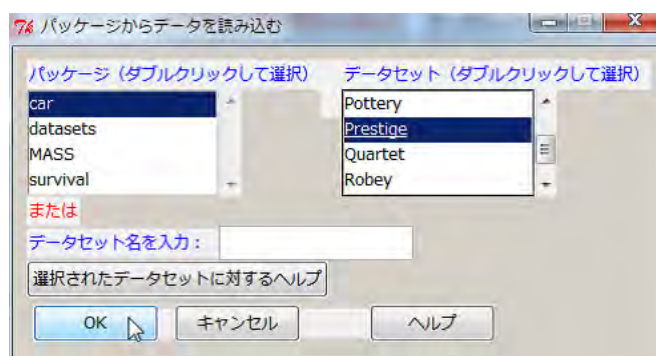


図 9 アタッチされたパッケージからのデータセットの読み込み—今の場合，car パッケージの Prestige データセット

3 数値による要約の実行とグラフの作成

アクティブデータセットがあると，R コマンダーのメニューにより数値による要約やグラフの作成を実行できる．基本的な例をいくつか示す．良い GUI というものはだいたい見れば分かるものである．R コマンダーがどのように機能するかを一度知ると，必要に応じてオンラインヘルプファイルを参照すれば，ほとんどトラブル無く利用できるだろう．

下記の例では，前節でテキストファイルから読み込んだ Nations をアクティブデータセットとする．前節で説明したように，Moore (2000) の 5 つのケースのデータセットをキー入力したり，パッケージ car から Prestige データセットを読み込んだりしているときは，これらのどれかがアクティブデータセットとなっている．R コマンダーのウィンドウの上左部にあるアクティブデータセット名を表示するフラットボタンをクリックし，現在メモリにあるデータセットのリストから選択することにより，アクティブデータセットを切り替えることができる．

*⁹ R コマンダーでは一般に，リストボックス内のアイテムをダイアログの他の箇所にコピーする必要がある場合，ダブルクリックするだけでよい．

統計量 → 要約 → アクティブデータセットにより、図 10 に示す結果を求めることができる。データセット内の各数値変数 (TFR, contraception, infant.mortality, GDP) に対して、最小値と最大値、第 1 四分位数、第 3 四分位数、メディアン、平均、欠測値の数を表示する。質的変数である region に対しては、因子の各水準のデータ数が表示される。データセット内に変数が 10 個以上ある場合は、計算を進めてよいかどうか問うことにより、不要な大量の出力を避けるようになっている。

同様に、統計量 → 要約 → 数値による要約により、図 11 に示すダイアログボックスが表示される。このダイアログには数値変数のみが表示される。因子 region が表示されないのは、因子に対して数値による要約を行っても意味がないからである。infant.mortality をクリックして選択し、OK をクリックすると、次の結果が出力ウィンドウに表示される。^{*10}

```
> numSummary(Nations[, "infant.mortality"], statistics=c("mean", "sd", "quantiles"))
      mean      sd 0% 25% 50% 75% 100%   n NA
43.47761 38.75604  2  12  30  66 169 201  6
```

デフォルトでは、平均と標準偏差 (sd)、最小値 (0 %)、第 1 四分位数、メディアン、第 3 四分位数、最大値 (100 %) に対応する分位点 (パーセント点) を表示する。n は有効な観測数で、NA は欠測値数である。

R コマンドのダイアログは通常そうであるが、数値による要約には OK, Cancel, Help という 3 つのボタンがある。Help ボタンにより、ダイアログボックス自体のヘルプページまたはダイアログが利用する R の関数のヘルプページを参照することができる。

数値による要約のダイアログボックスでは、因子の水準によって定義されるグループ内での要約情報、つまり層別の要約情報を求めることもできる。層別して要約... をクリックすると、図 12 に示す質的変数ダイアログが表示される。データセット Nations には質的変数が 1 つしかないのので、変数のリストには region のみが表示されている。これを選択し、OK ボタンをクリックすると、層別して要約... ボタンが層別変数: region へと変化する (図 13)。OK をクリックすると、次の結果が表示される。

```
> numSummary(Nations[, c("GDP", "infant.mortality")], groups=Nations$region,
+   statistics=c("mean", "sd", "quantiles"), quantiles=c(0,.25,.5,.75,1))
```

Variable: GDP

	mean	sd	0%	25%	50%	75%	100%	n	NA
Africa	1196.000	2089.614	36	209.00	389.5	1004.50	11854	54	1
Americas	5398.000	6083.311	386	1749.25	2765.5	7017.75	26037	40	1
Asia	4505.051	6277.738	122	345.00	1079.0	6407.50	22898	39	2
Europe	13698.909	13165.412	271	1643.75	9222.5	26226.00	42416	44	1
Oceania	8732.600	11328.708	654	1102.75	2348.5	17512.00	41718	20	5

Variable: infant.mortality

	mean	sd	0%	25%	50%	75%	100%	n	NA
Africa	85.27273	35.188095	7	61.00	85.0	111.00	169	55	0
Americas	25.60000	17.439713	6	12.00	21.5	36.00	82	40	1
Asia	45.65854	32.980001	5	22.00	37.0	72.00	154	41	0
Europe	11.85366	7.122363	5	6.00	8.0	16.00	32	41	4
Oceania	27.79167	29.622229	2	9.25	20.0	35.75	135	24	1

^{*10} 変数のリストボックスで 1 つの変数のみを選択するには、その名前を左クリックするだけでよい。2 つ以上の変数を選択したい場合は、通常の Windows での方法を適用する。左クリックで変数を選択することができ、再度左クリックすると、取り消すことができる。Shift キーを押して左クリックすると、選択を拡大することができる。Ctrl キーを押したまま左クリックすると、選択を追加することができる。

R コマンダーの他のダイアログでも、同様にして層別変数を選択することができることがある。

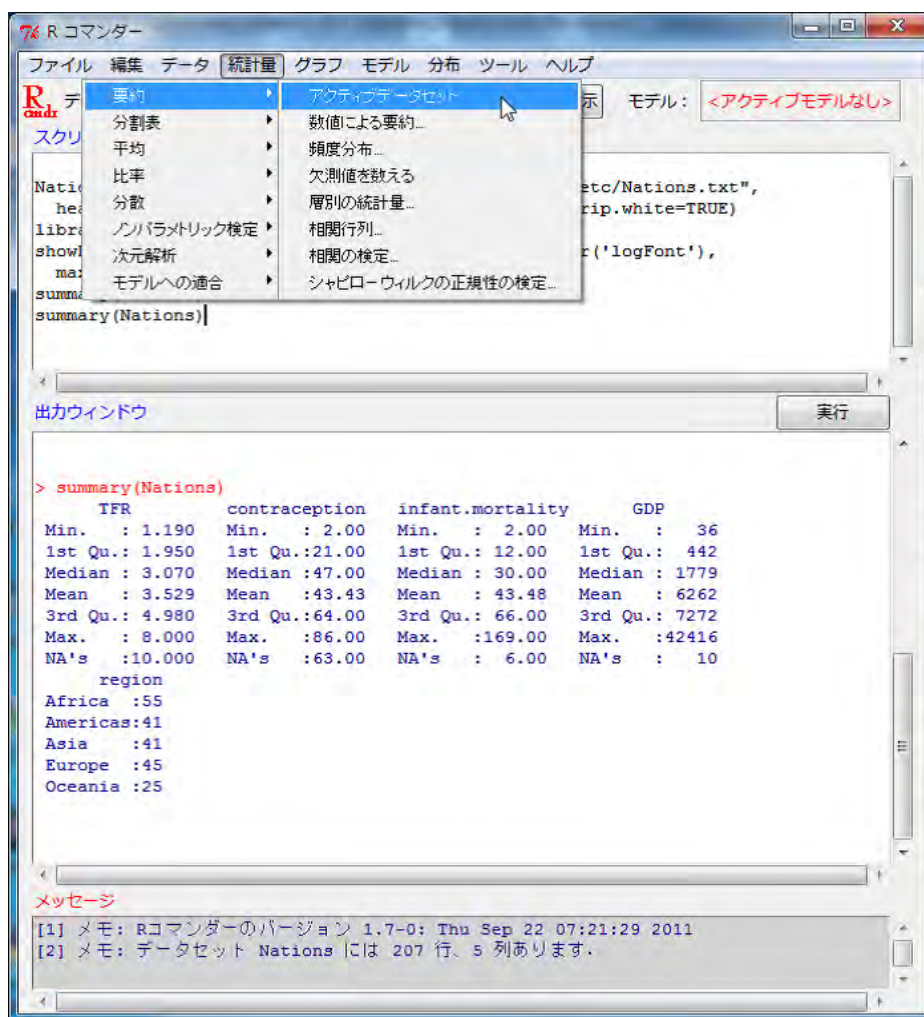


図 10 アクティブデータセットに対する変数の要約情報

R コマンダーでグラフを作成することも簡単である。例えば、R コマンダーのメニューよりグラフ → ヒストグラム... を選択すると、図 14 に示すヒストグラムのダイアログボックスが表示される。infant.mortality を選択し、OK をクリックすると、図 15 に示すヒストグラムを含むグラフィックスウィンドウが表示される。

1 つのセッションで複数のグラフを作成した場合、グラフィックスデバイスウィンドウには、最新のもののみが表示される。キーボードの Page Up または Page Down キーにより、以前のグラフを呼び出すことができる。^{*11}

3.1 ダイアログボックスにおけるメモリ機能

R コマンダーのバージョン 1.7-0 から、ダイアログでのユーザの前の選択を記憶する機能を導入している。この機能は、R コマンダーを起動する前に、R のコマンド `Rcmdr=list(dialog.memory=TRUE)` を入力しておくことにより有効になる。当面の間、R コマンダーのダイアログは、新しいスキームをに対応させる途中のため、`dialog.memory` オプションの設定を FALSE としている。アクティブデータセットを変更したとき、ダイア

^{*11} R コマンダーは、グラフ履歴メカニズムをスタートアップ時にオンにしている。この機能は Windows のみで利用可能である。グラフ → 3次元グラフ → 3次元散布図... で作られた動的な 3次元散布図は、特別な RGL デバイス中に表示される。同様に、モデル → グラフ → 効果プロットで作られる統計モデルに対する効果プロット (Fox, 2003; Fox and Hong, 2009) は、独立したグラフィックスデバイス・ウィンドウに表示される。

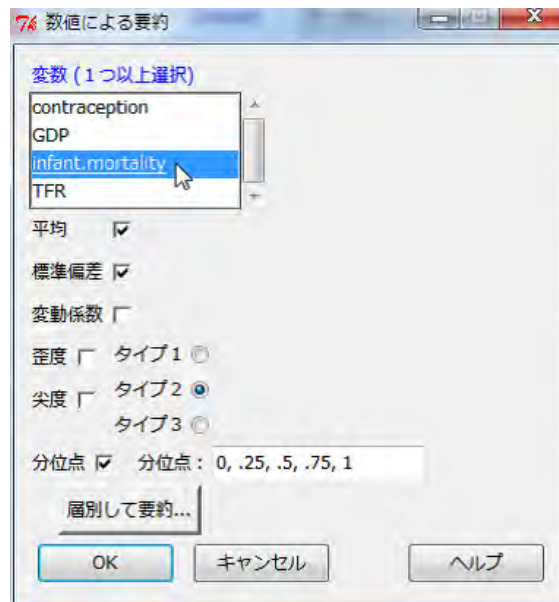


図 11 数値による要約のダイアログボックス

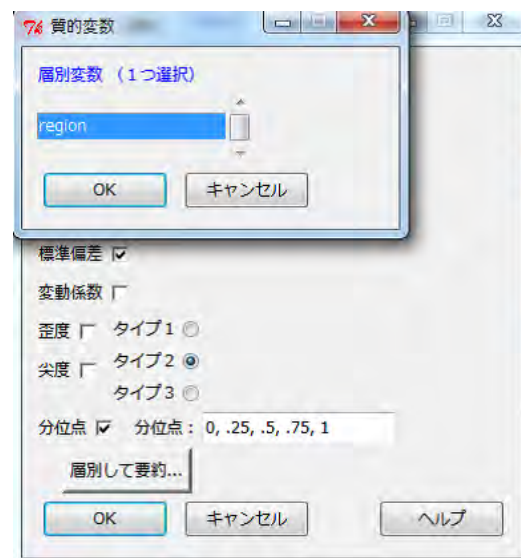


図 12 質的変数ダイアログボックスで層別変数を選択する

ログボックスのメモリは消去される。メモリ機能を持つダイアログボックスには *Reset* ボタンがある。これにより、ダイアログをデフォルトの状態に戻すことができる。数値による要約におけるこのボタンを、図 16 に示す。

4 統計モデル

統計量 → モデルへの適合メニューを用いて、R コマンドでいくつかの統計モデルを作成することができる。線形モデル（線形回帰... および線形モデル... により）、一般化線形モデル、多項ロジットモデル、順序回帰モデル（比例オッズモデル等）である。[最後の 2 つは、それぞれ Venables and Ripley (2002) による 2 つのパッケージ *nnet* と *MASS* による。] ダイアログボックスはモデルによって少し異なるが（例えば、一般化線形モデルのダイアログには、分布族とこれに対応するリンク関数を選択する機能がある）、図 17 に示す線形モデルのダイアログボックスと共通する一般的な構造を持つ。^{*12}

^{*12} 例外は線形回帰ダイアログで、分析対象とするデータセットが持つ数値変数のリストから名前を選択することにより、説明変数と目的変数を指定することができる。以下の説明では、R のモデル式についてよく知っていることを仮定している。詳細については、R

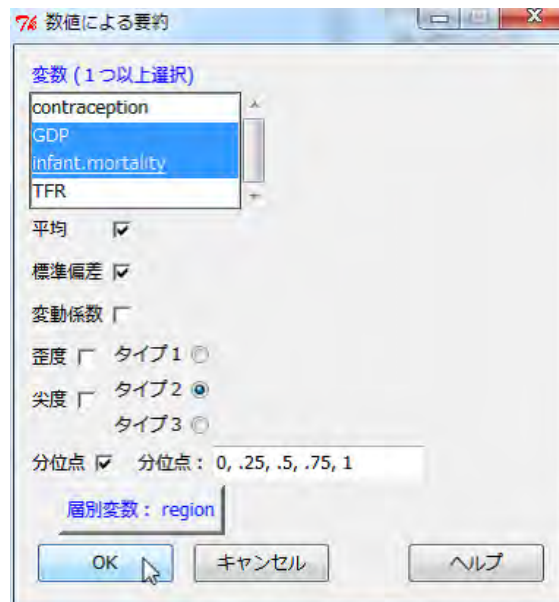


図 13 層別変数を指定した後の数値による要約ダイアログボックス



図 14 ヒストグラムのダイアログボックス

- 変数のリストボックスにある変数をダブルクリックすると、モデル式にそれがコピーされる。式の左辺が空白なら左辺に、そうでないならば右辺に（必要ならば、記号 + が前に追記される）。変数リストの中で、因子については、変数名の後ろに [因子] と記されている。
- モデル式の上に並んでいるボタンは、式の右辺にオペレータや丸括弧を入力するために利用できる。
- モデル式のフィールドに直接キー入力してもよい。例えば、`log(income)` といった項を入力するには、直接入力する必要がある。
- モデル名は、今は `LinearModel.1` となっている。これは自動的に生成されるが、変更することもできる。
- 部分集合の表現というボックスに R の表現を入力することができる。入力があると、これが関数 `lm` の引数 `subset` として送られ、データセット内の観測値の部分集合に対してモデルが適用される。部分集合の表現の 1 つの形として、各データに対して `TRUE` または `FALSE` を評価する論理的な表現がある。例えば、`type != "prof"`（これは、データセット `Prestige` で、非専門的職業全て（`prof` ではない）を指定する）である。

と一緒にインストールされる *Introduction to R* を参照。これは *R Console* の *Help* メニューからアクセスすることができる。

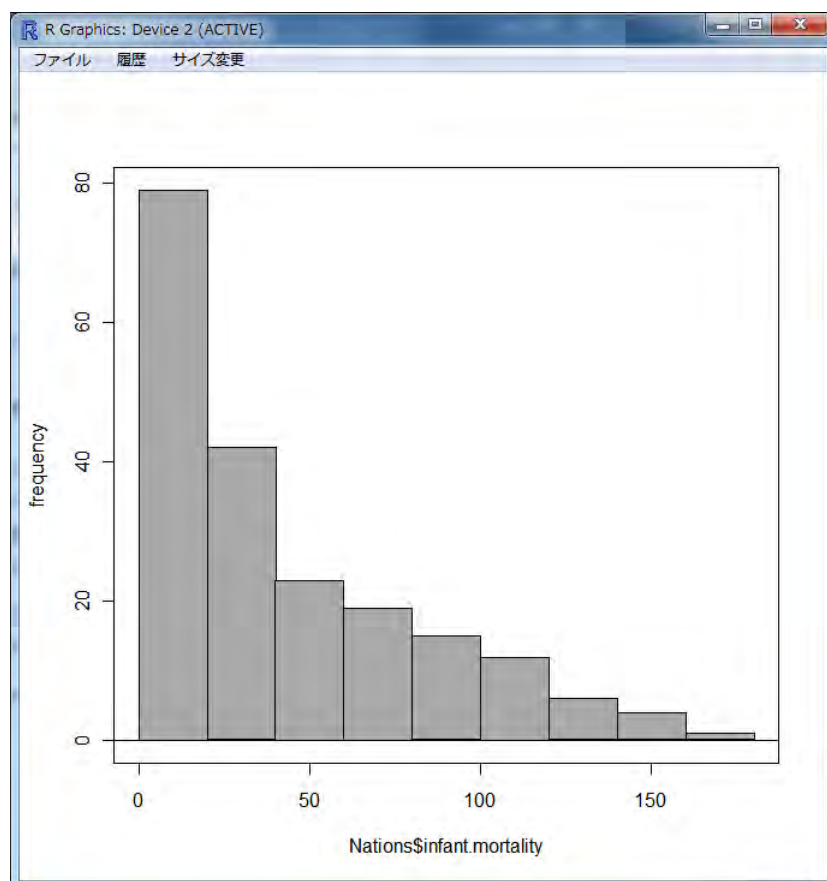


図 15 Nations データセットの infant.mortality (乳児死亡率) のヒストグラムを表示するグラフィカルウィンドウ

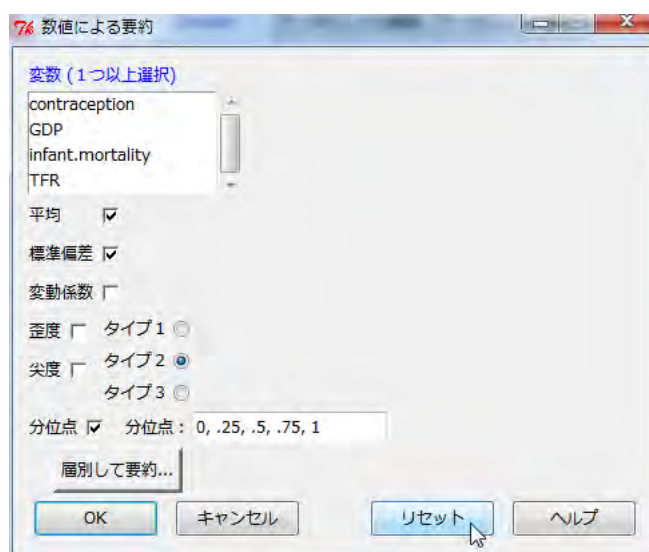


図 16 数値による要約ダイアログボックス — dialog.memory オプションを TRUE にし, リセットボタンを表示している。

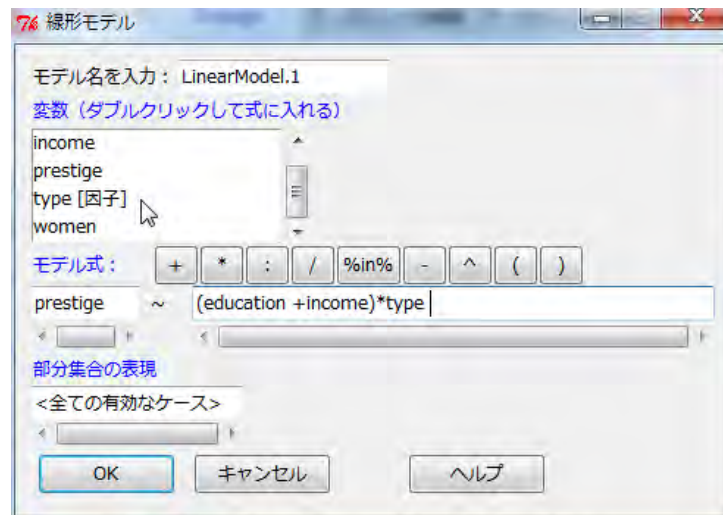


図 17 線形モデルダイアログボックス

OK ボタンをクリックすると、次の結果が出力ウィンドウに表示される。また、`LinearModel.1` がアクティブモデルとなり、それがモデルボタンに表示される。

```
> LinearModel.1 <- lm(prestige ~ (education + income)*type, data=Prestige)
```

```
> summary(LinearModel.1)
```

Call:

```
lm(formula = prestige ~ (education + income) * type, data = Prestige)
```

Residuals:

Min	1Q	Median	3Q	Max
-13.462	-4.225	1.346	3.826	19.631

Coefficients:

	Estimate	Std. Error	t value	Pr(> t)
(Intercept)	2.276e+00	7.057e+00	0.323	0.7478
education	1.713e+00	9.572e-01	1.790	0.0769 .
income	3.522e-03	5.563e-04	6.332	9.62e-09 ***
type[T.prof]	1.535e+01	1.372e+01	1.119	0.2660
type[T.wc]	-3.354e+01	1.765e+01	-1.900	0.0607 .
education:type[T.prof]	1.388e+00	1.289e+00	1.077	0.2844
education:type[T.wc]	4.291e+00	1.757e+00	2.442	0.0166 *
income:type[T.prof]	-2.903e-03	5.989e-04	-4.847	5.28e-06 ***
income:type[T.wc]	-2.072e-03	8.940e-04	-2.318	0.0228 *

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

Residual standard error: 6.318 on 89 degrees of freedom

(4 observations deleted due to missingness)

Multiple R-squared: 0.8747, Adjusted R-squared: 0.8634

F-statistic: 77.64 on 8 and 89 DF, p-value: < 2.2e-16

アクティブモデルに対する操作は、モデルメニューから選択することができる。例えば、モデル → 仮説検定 → 分散分析表により、次の結果が表示される。

```
> Anova(LinearModel.1, type="II")
```

Anova Table (Type II tests)

Response: prestige

	Sum Sq	Df	F value	Pr(>F)
education	1068.0	1	26.7532	1.413e-06 ***
income	1131.9	1	28.3544	7.511e-07 ***
type	591.2	2	7.4044	0.00106 **
education:type	238.4	2	2.9859	0.05557 .
income:type	951.8	2	11.9210	2.588e-05 ***
Residuals	3552.9	89		

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

5 その他

5.1 出力の保存と印刷

R コマンダーのファイルメニューより、テキストの出力を直接保存することができる。これは、R のグラフィックスデバイスのウィンドウにおいて、ファイルメニューによりグラフを保存したり、印刷したりできるのと同じである。しかし、保存したい出力テキストやグラフをワープロ文書として保存しておく方が便利である。このようにすると、R の出力に注記や説明をつけて配布することができる。

Word や OpenOffice の Writer を起動する。Windows のワードパッドでもよい。出力ウィンドウからテキストをコピーするには、テキストの範囲をマウスで指定し、編集メニューからコピーを選択し（あるいは、*Ctrl-c* を押したり、ウィンドウ内で右クリックしてコンテキストメニューからコピーを選択したりする）、編集 → 貼り付け（または、*Ctrl-v*）によりワープロにテキストを貼り付ける。1 つ注意すべきことは、R からのテキストの出力に対しては、Courier New といった等幅フォント（タイプライター体）を使うべきである。でないときれいに整列しない。

同様に、グラフをコピーするには、R のグラフィックスデバイスのメニューからファイル → クリップボードにコピー → メタファイルとしてを選択する。そして、編集 → 貼り付け（または、*Ctrl-v*）によりワープロにグラフを貼り付ける。別の方法として、*Ctrl-w* により R のグラフィックスデバイスからグラフをコピーするか、グラフ上で右クリックして表示したコンテキストメニューよりメタファイルにコピーを選択してもよい。^{*13} R のセッションの最後に、作成したドキュメントを保存または印刷することができるが、これは注釈付きの記録となる。

テキストやグラフを保存するための別の方法として、R コマンダーのファイルメニューと、グラフ → グラフをファイルで保存メニューにより保存することもできる。

^{*13} これらのメニューを調べると分かるように、様々なフォーマットでグラフをファイルのみならずクリップボードに保存することができる。ここで述べた方法は簡単であるが、グラフは高品位である。再度述べるが、ここでの説明は Windows のシステムのみにあてはまるものである。

5.2 R セッションの終了

セッションを終了する方法はいくつかある．例えば，R コマンダーのメニューのファイル → 終了 → コマンダーと R をを選択する．終了してもよいかという確認の後，スクリプトと出力ウィンドウの内容を保存したいかどうかを聞かれる．なお，R Console で，ファイル → 終了を選択してもよい．この場合，R のワークスペース（すなわち，R がメモリに保存しているデータ）を保存するかどうか聞かれる．通常は *No* でよい．

5.3 スクリプトウィンドウにコマンドを入力する

スクリプトウィンドウは，コマンドを編集・入力・実行するための簡単な機能を提供する．R コマンダーが生成したコマンドは，スクリプトウィンドウに表示され，エディタと同じように，コマンドを入力したり編集したりすることができる．しかしながら，R コマンダーは，R に対する真のコンソールではなく，スクリプトウィンドウには限界がある．例えば，複数行にわたるコマンドは，全て同時に実行されなければならない．本気でプログラミングするときは，R の Windows や Mac OS X バージョンが提供するスクリプトエディタを利用するか，より望ましいのは，インタラクティブな開発環境を提供するプログラミングエディタを利用することである．

参考文献

- Fox, J. (2003). Effect displays in R for generalised linear models. *Journal of Statistical Software*, 8(15):1-27.
- Fox, J. (2005). The R Commander: A basic-statistics graphical user interface to R. *Journal of Statistical Software*, 19(9):1-42.
- Fox, J. (2007). Extending the Rcmdr by “Plug-in” Packages. *R News*, 7(3):46-52.
- Fox, J. and Hong, J. (2009). Effect displays in R for multinomial and proportional-odds logit models: Extensions to the effects package. *Journal of Statistical Software*, 32(1):1.24.
- Moore, D. S. (2000). *The Basic Practice of Statistics, Second Edition*. Freeman, New York.
- Venables, W. N. and Ripley, B. D. (2002). *Modern Applied Statistics with S, Fourth Edition*. Springer, New York.