

R コマンダー入門*

John Fox

Version 1.9-1 (2012年9月4日)

日本語訳[†]

2012年10月9日

1 R コマンダーを起動する

本稿は、R の Windows バージョンの下での R コマンダー (Rcmdr) の利用法を説明する。Mac OS X と Linux の元での R コマンダーの見かけと利用法には少し異なる点がある。これらの OS における R コマンダーのインストール方法については、R コマンダーのウェブページ <<http://socserv.socsci.mcmaster.ca/jfox/Misc/Rcmdr/index.html>>, または、直接 <tinyurl.com/Rcmdr>にあるインストールノートへのリンクをたどることにより利用可能である。

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 コマンダーのグラフィカル・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 コマンダ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 コマンダ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 コマンダーと 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 コマンダ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 コマンダーの GUI で利用された R のコマンドは、R コマンダーの主ウィンドウの上部にあるテキスト

* 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-2.15.1, Rcmdr.1.9-1 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の方法がある。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。

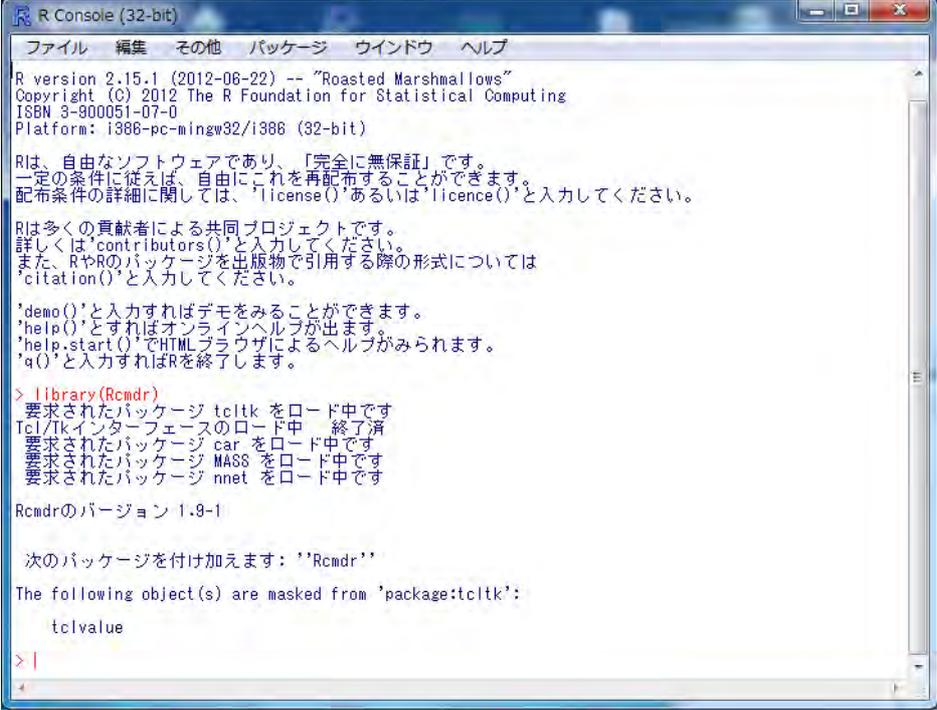
^{*2} Rcmdr は、R とともに配布されている推奨パッケージに加えていくつかのパッケージを必要とし、Rcmdr の起動とともにこれらをロードする。Rcmdr および必要なパッケージ、他の多くの追加パッケージは、Comprehensive R Archive Network (CRAN) <<http://cran.r-project.org/>> よりダウンロードできる。

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。

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** コマンダーの開発目的は、コマンドを入力しないで済むようにすることにある。

- 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他の情報（“メモ”）を表示する。
- グラフを作成すると、それらは別の *Graphics Device* ウィンドウに表示される。



```
R Console (32-bit)
ファイル 編集 その他 パッケージ ウィンドウ ヘルプ

R version 2.15.1 (2012-06-22) -- "Roasted Marshmallows"
Copyright (C) 2012 The R Foundation for Statistical Computing
ISBN 3-900051-07-0
Platform: i386-pc-mingw32/i386 (32-bit)

Rは、自由なソフトウェアであり、「完全に無保証」です。
一定の条件に従えば、自由にこれを再配布することができます。
配布条件の詳細に関しては、'license()'あるいは'licence()'と入力してください。

Rは多くの貢献者による共同プロジェクトです。
詳しくは'contributors()'と入力してください。
また、RやRのパッケージを出版物で引用する際の形式については
'citation()'と入力してください。

'demo()'と入力すればデモをみることができます。
'help()'とすればオンラインヘルプが出ます。
'help.start()'でHTMLブラウザによるヘルプがみられます。
'q()'と入力すればRを終了します。

> library(Rcmdr)
要求されたパッケージ tcltk をロード中です
Tcl/Tkインターフェースのロード中 終了済
要求されたパッケージ car をロード中です
要求されたパッケージ MASS をロード中です
要求されたパッケージ nnet をロード中です

Rcmdrのバージョン 1.9-1

次のパッケージを付け加えます: "Rcmdr"

The following object(s) are masked from 'package:tcltk':

  tclvalue

>|
```

図 1 パッケージ 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（Fox, 2007 および Fox and Carvalho, 2012 参照）、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

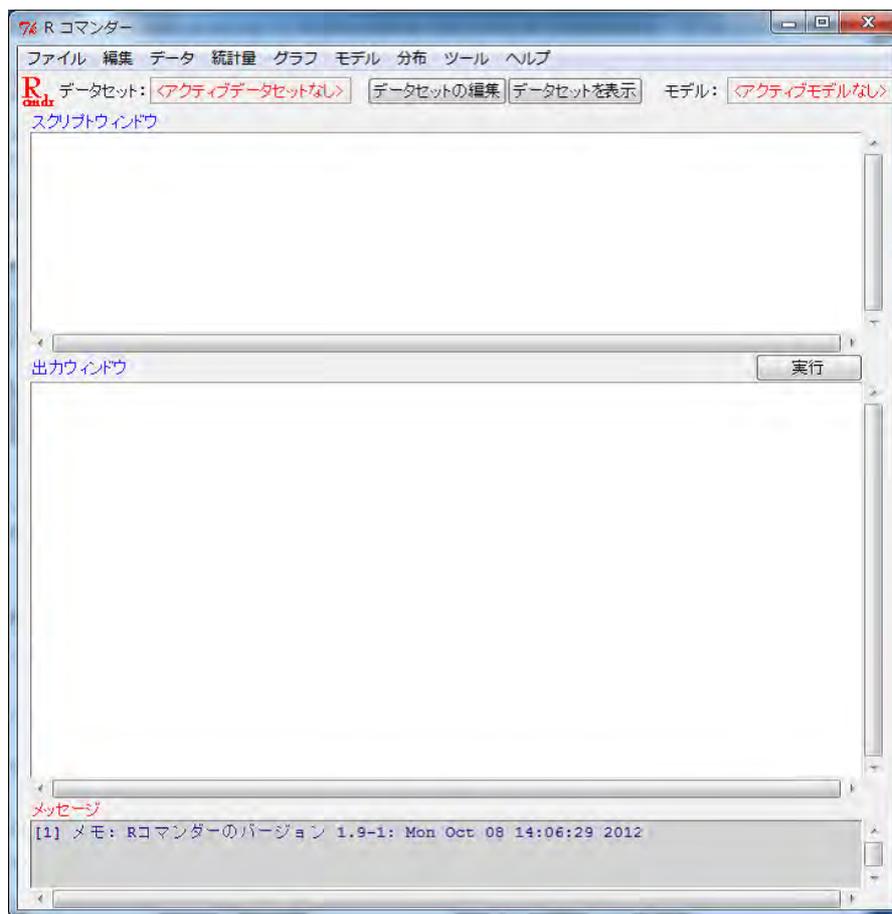
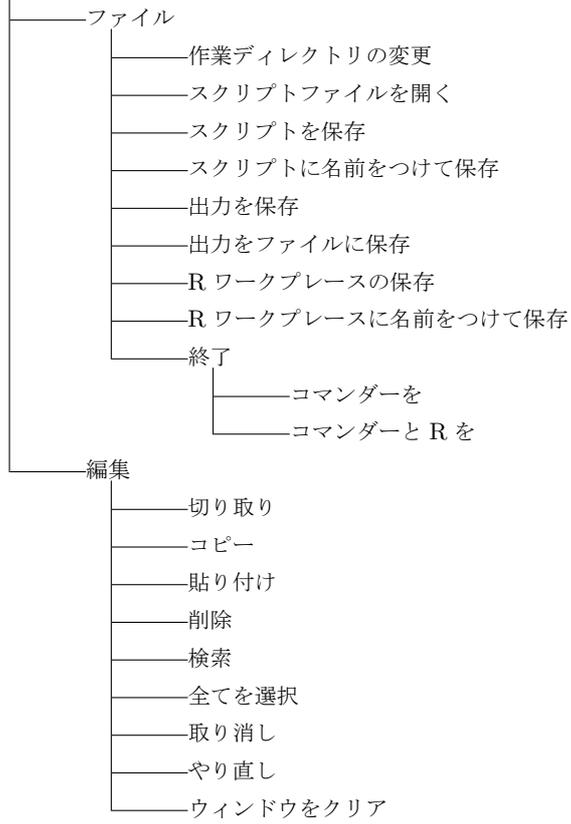


図2 起動直後の R コマンダーのウィンドウ

ログボックスはヘルプボタンを持つ（下記参照）。

R コマンダー（バージョン 1.9-0）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が表示さる。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。

Rcmdr



Rcmdr

データ

- 新しいデータセット
- データセットのロード
- データセットの統合
- 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
 - SPSS データセットから
 - SAS のエクスポートファイルから
 - Minitab データセットから
 - STATA データセットから
 - Excel または Access, dBase のデータセットから [32-bit Windows のみ]
 - Excel ファイルから [現在は, 64-bit Windows のみ]
- 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
 - 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
- アクティブデータセット
 - アクティブデータセットの選択
 - アクティブデータセットの更新
 - アクティブデータセットのヘルプ (可能なら)
 - アクティブデータセット内の変数
 - ケース名の設定
 - アクティブデータセットの部分集合を抽出
 - アクティブデータセット内の変数の集計
 - アクティブデータセットから行を削除
 - アクティブデータセット内の変数を積み重ねて結合
 - 欠測値のあるケースを削除
 - アクティブデータセットの保存
 - アクティブデータセットのエクスポート
- アクティブデータセット内の変数の管理
 - 変数の再コード化
 - 新しい変数の計算
 - データセットに観測値番号を追加
 - 変数の標準化
 - 数値変数を因子に変換
 - 数値変数を区間で区分
 - 因子水準の再順序化
 - 因子に対する対比を定義
 - 変数名の変更
 - データセットから変数を削除

Rcmdr

統計量

要約

- アクティブデータセット
- 数値による要約
- 頻度分布
- 欠測値を数える
- 層別の統計量
- 相関行列
- 相関の検定
- シャピローウィルクの正規性の検定

分割表

- 2 元表
- 多元表
- 2 元表の入力と分析

平均

- 1 標本の t 検定
- 独立標本 t 検定
- 対応のある t 検定
- 1 元配置分散分析
- 多元配置分散分析

比率

- 1 標本比率の検定
- 2 標本比率の検定

分散

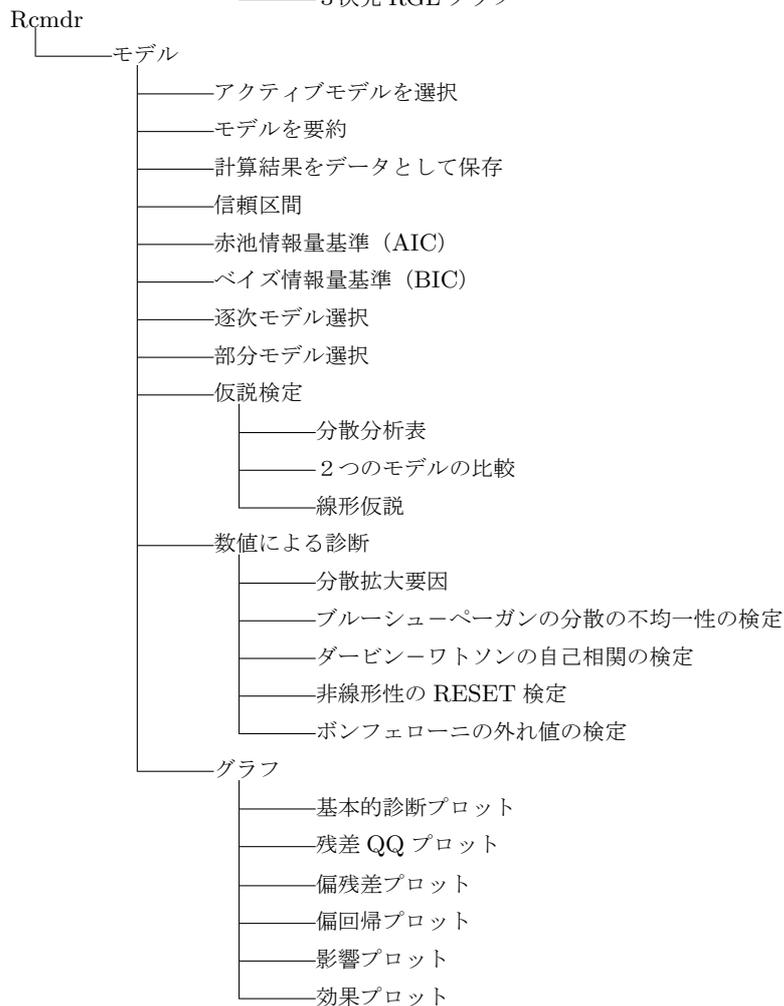
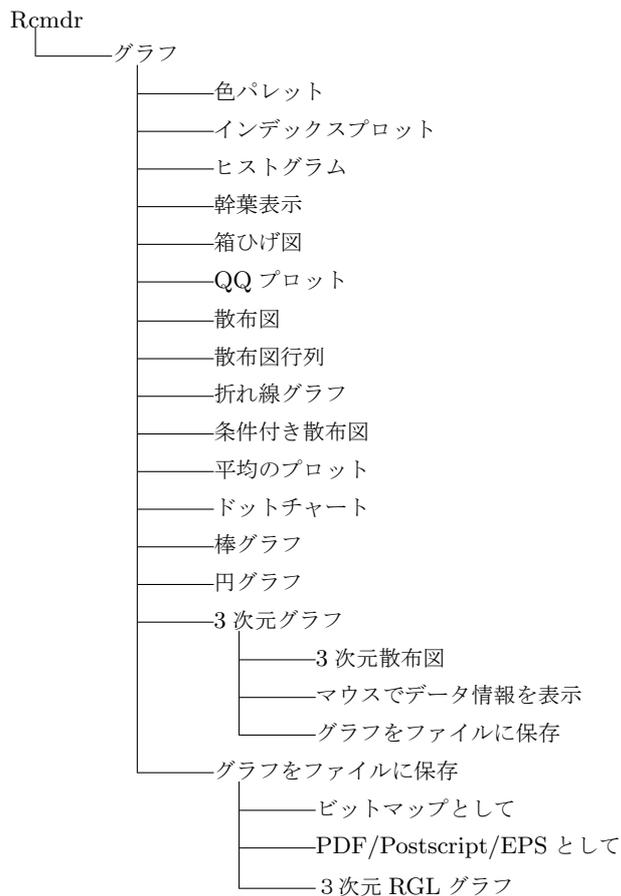
- 分散の比の F 検定
- バートレットの検定
- ルビーンの検定

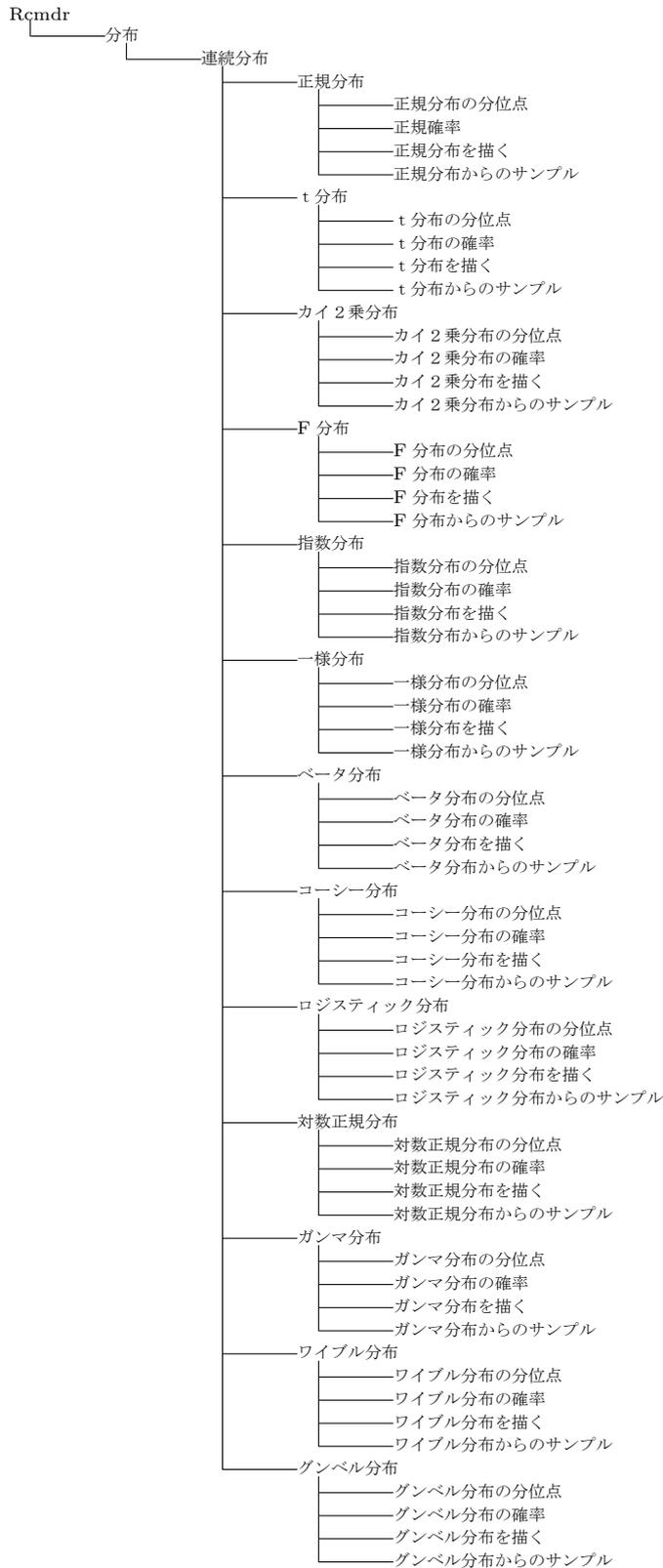
ノンパラメトリック検定

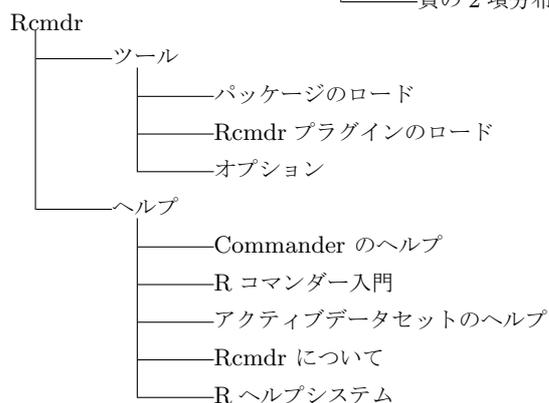
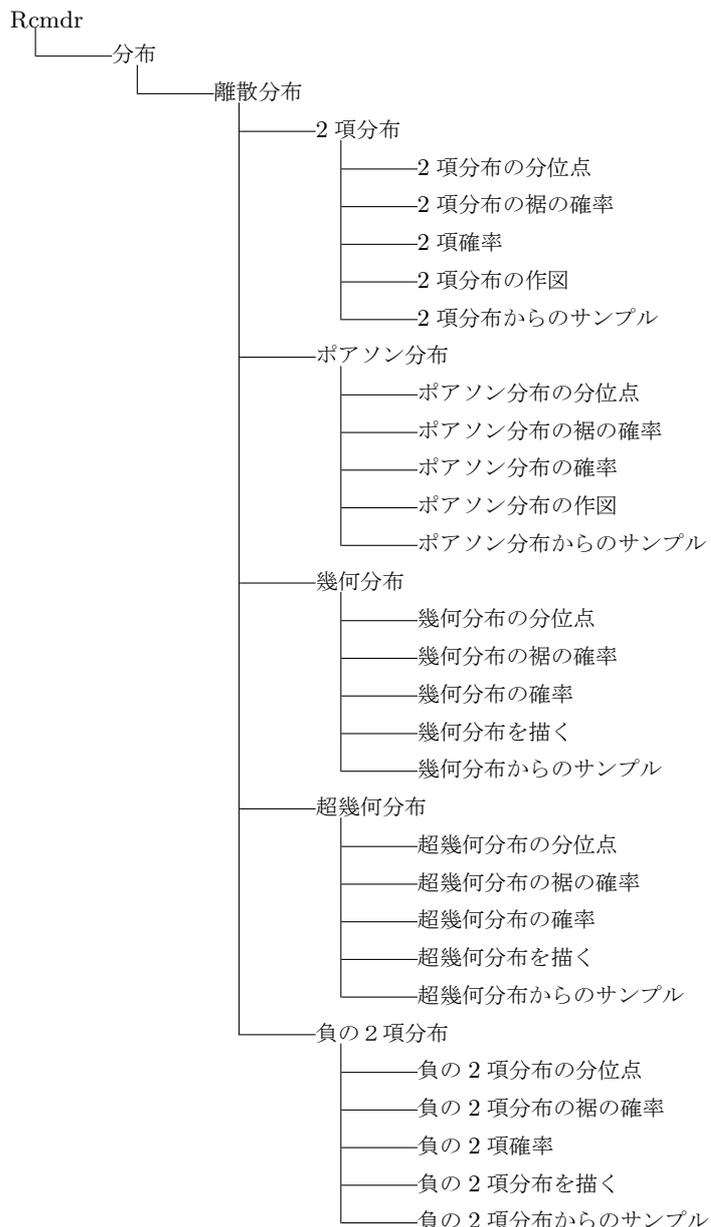
- 2 標本ウィルコクソン検定
- 対応のあるウィルコクソン検定
- クラスカル-ウォリスの検定
- フリードマンの順位和検定

次元解析

- スケールの信頼性
- 主成分分析
- 因子分析
- 検証的因子分析
- クラスタ分析
 - k-平均クラスタ分析
 - 階層的クラスタ分析
 - 階層的クラスタリングの結果の要約
 - 階層的クラスタリングの結果をデータセットに保存
- モデルへの適合
 - 線形回帰
 - 線形モデル
 - 一般化線形モデル
 - 多項ロジットモデル
 - 順序回帰モデル







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ア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コマンドを複数行入力してもよいが、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しておく必要がある。キーの組合せ入力 *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 コマンド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 コマンド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 コマンド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 のオプションで設定する。^{*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 コマンド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 コマンドを起動した直後には、アクティブデータセットはない。

R コマンドによるデータの入力方法にはいくつかある。^{*7}

- **Mac OS X** 以外のプラットフォームでは、**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。データセット数が非常に少ないときは、これでよい。
- プレーンテキスト（“ascii”）ファイルまたはクリップボード、URL によるインターネット、他の統計

^{*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アは、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かわりに すこし不細工であるが、R のデータビューア（*View*）を利用する。変数の数に関係なく R のデータビューアを利用するには、限界値を 0 に設定する。詳細については、R コマンドのヘルプファイルを参照のこと。

^{*4} R コマンドのプラグインパッケージ（Fox, 2007; Fox and Carvalho, 2012）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 GUI の標準的な仕様である。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。

^{*7} これらのデータソースは、全てのプラットフォームで利用可能ではない。

パッケージ (Minitab や SPSS, Stata) から,あるいは (Windows) では, 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.*⁸ 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.

*⁸ 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.

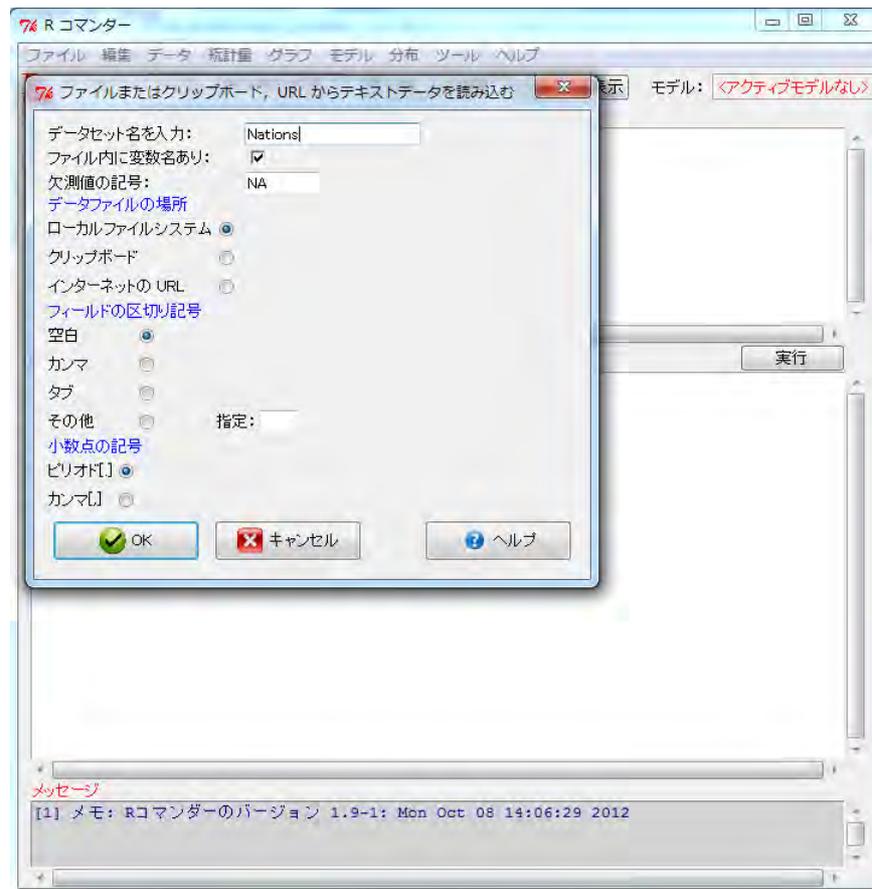


図 3 テキストファイルからのデータの読み込み

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** コマンダ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** コマンダ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メモが表示される。**R** コマンダーはまた、**relimp** パッケージをロードするために **library** コマンドを実行するので、データセットが表示される。このように、一般に、パッケージは必要なときに自動的にロードされる。

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** コマンダ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^{*9}。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。

^{*9} **R** データエディタの制約により、**Mac OS X** のもとで直接データセットを入力することはできないので、対応するメニュー項目を以下表示しない。

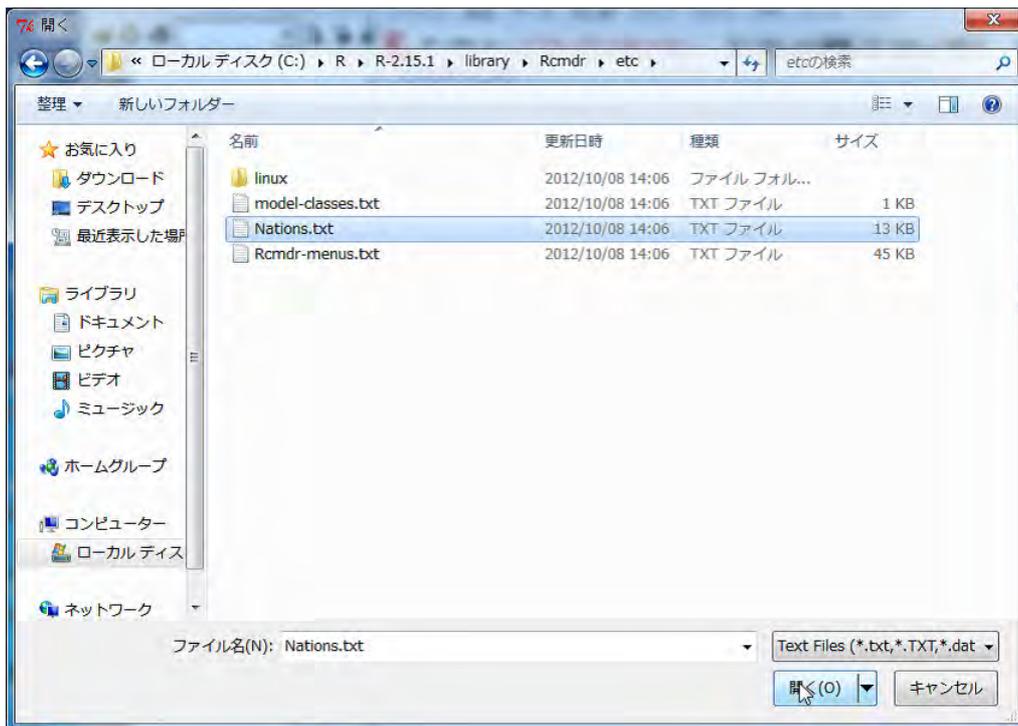


図4 テキスト形式データファイルを読み込むためのファイルを開くダイアログボックス

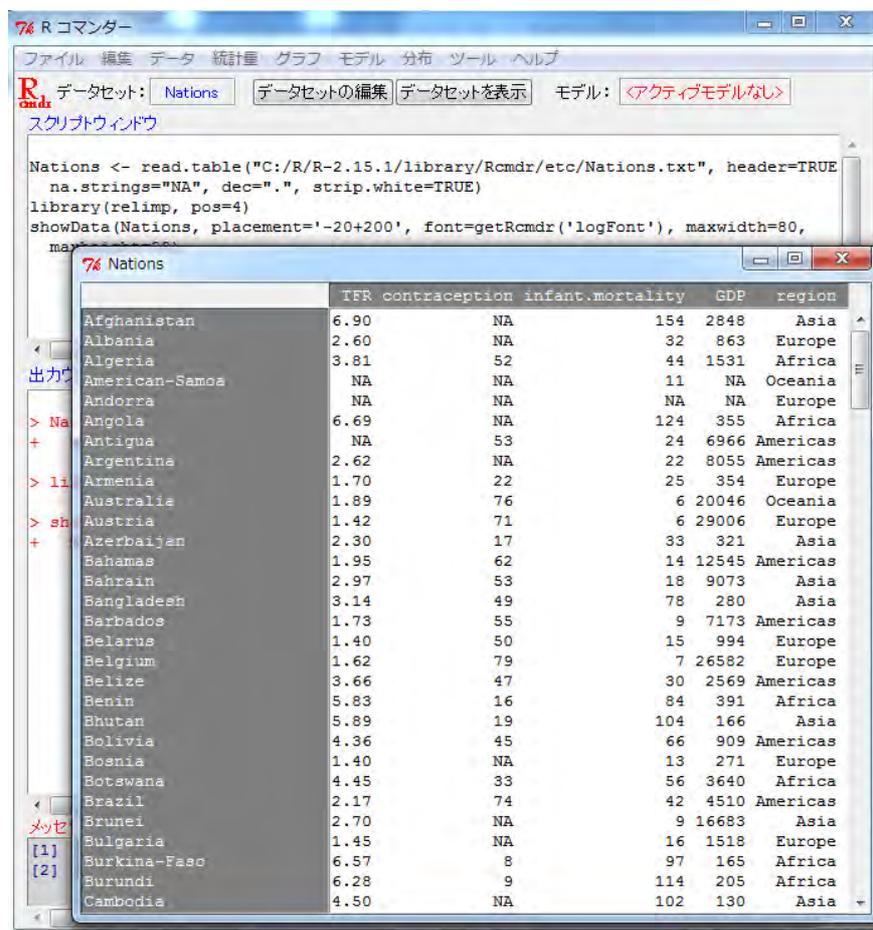


図5 アクティブデータセットの表示

- **R** コマンダーのメニューから、**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。
- 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

変数名:

型: numeric character

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

パッケージの多くはデータを含んでいる。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リストとして表示でき、**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。^{*10} 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名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右のリストボックスにデータセット名のリストが表示される。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。^{*11} データセットのスクロールボックス内の記号キーをクリックすると、その記号で始まる次のデータセットに移動する。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

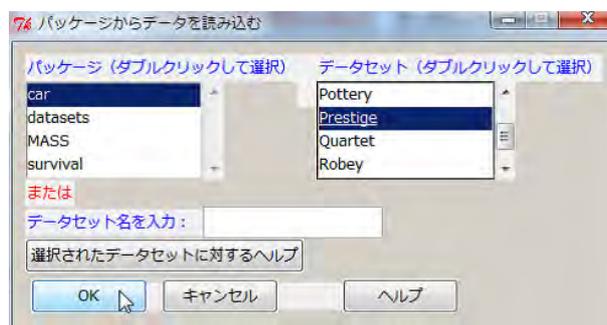


図9 アタッチされたパッケージからのデータセットの読み込み—今の場合、car パッケージの Prestige データセット

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 コマンド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。ここでは、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 コマンドーがどのように機能するかを一度見ると、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ことを期待する。

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している。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

^{*10} 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く、データフレームのみが R コマンド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。

^{*11} R コマンドーでは一般に、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。

る。R コマンドのウィンドウの左上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。

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ことができるように設定されている。

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。^{*12}

```
> numSummary(Nations[, "infant.mortality"], statistics=c("mean", "sd", "IQR",
+ "quantiles"), quantiles=c(0,.25,.5,.75,1))
      mean      sd IQR 0% 25% 50% 75% 100%   n NA
43.47761 38.75604  54  2  12  30  66  169 201  6
```

デフォルトでは、平均と標準偏差 (sd)、四分位範囲 (IQR)、および、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 は欠測値の数である。

R コマンドのダイアログは通常そうであるが、図 11 に示す**数値による要約**ダイアログボックスのように、**OK**、**キャンセル**、**ヘルプ**という 3 つのボタンを持つ。ヘルプボタンにより、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 (ブラウザに表示) を参照することができる。リセットボタンは、ほとんどの R コマンドのダイアログにあるが、ダイアログをそのオリジナルな状態に戻す。さもなければ、以前の選択からの起動状態を保持する。アクティブデータセットが切り替えられたとき、ダイアログの状態もリセットされる。

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", "IQR", "quantiles"), quantiles=c(0,.25,.5,.75,1))
```

```
Variable: GDP
      mean      sd      IQR 0%   25%   50%   75% 100%  n NA
Africa  1196.000 2089.614  795.50 36  209.00 389.5 1004.50 11854 54  1
Americas 5398.000 6083.311 5268.50 386 1749.25 2765.5 7017.75 26037 40  1
Asia    4505.051 6277.738 6062.50 122  345.00 1079.0 6407.50 22898 39  2
Europe  13698.909 13165.412 24582.25 271 1643.75 9222.5 26226.00 42416 44  1
Oceania  8732.600 11328.708 16409.25 654 1102.75 2348.5 17512.00 41718 20  5
```

^{*12} 変数リストボックスで 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、追加選択することができる。

Variable: infant.mortality

	mean	sd	IQR	0%	25%	50%	75%	100%	n	NA
Africa	85.27273	35.188095	50.0	7	61.00	85.0	111.00	169	55	0
Americas	25.60000	17.439713	24.0	6	12.00	21.5	36.00	82	40	1
Asia	45.65854	32.980001	50.0	5	22.00	37.0	72.00	154	41	0
Europe	11.85366	7.122363	10.0	5	6.00	8.0	16.00	32	41	4
Oceania	27.79167	29.622229	26.5	2	9.25	20.0	35.75	135	24	1

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。

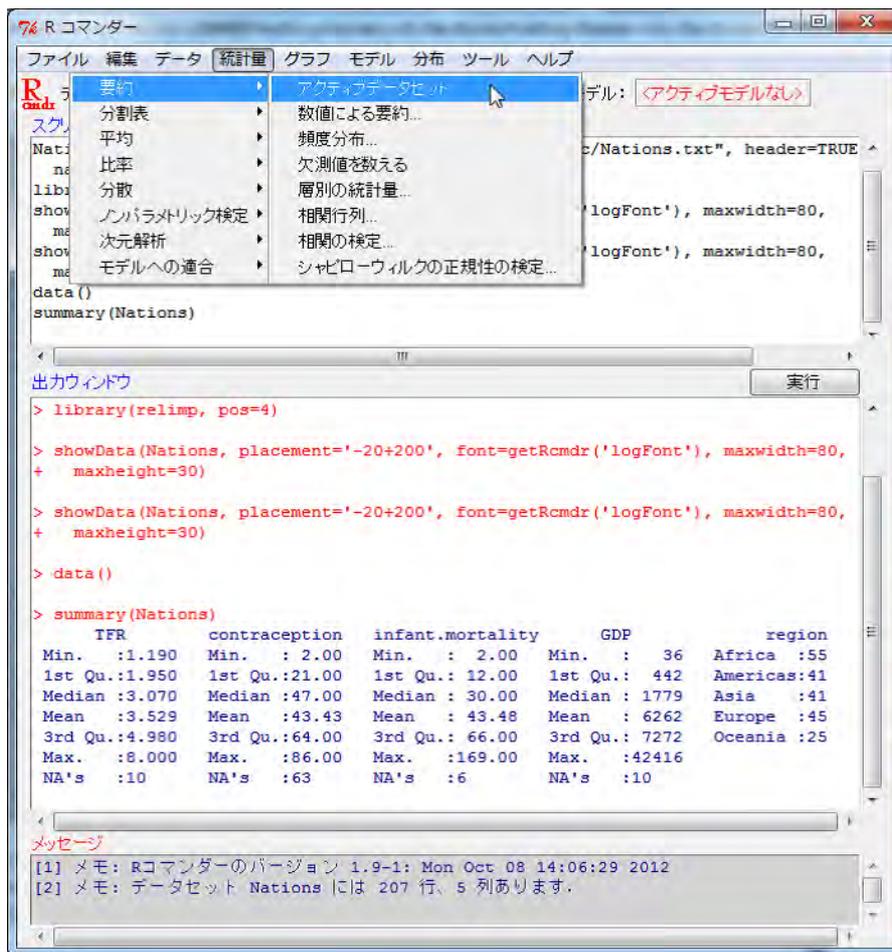


図 10 アクティブデータセットに対する変数の要約情報の取得

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** が表示された **グラフィックスウィンドウ** が表れる。

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** には通常、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、前のグラフを呼び出すことができる。^{*13}

^{*13} 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**

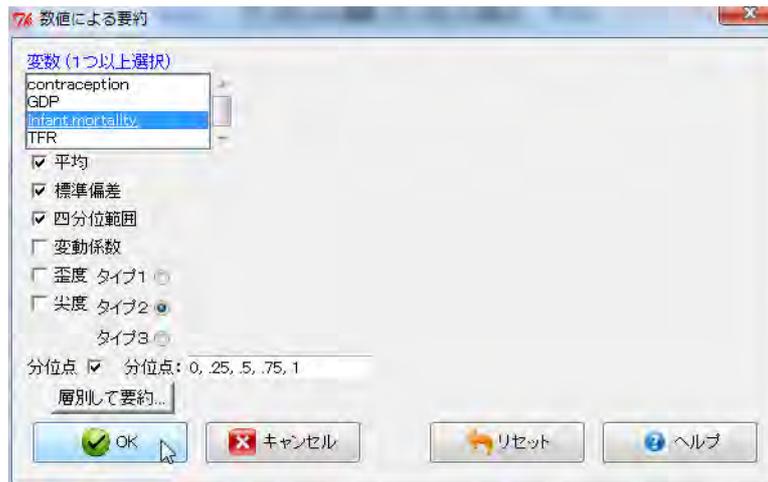


図 11 数値による要約のダイアログボックス

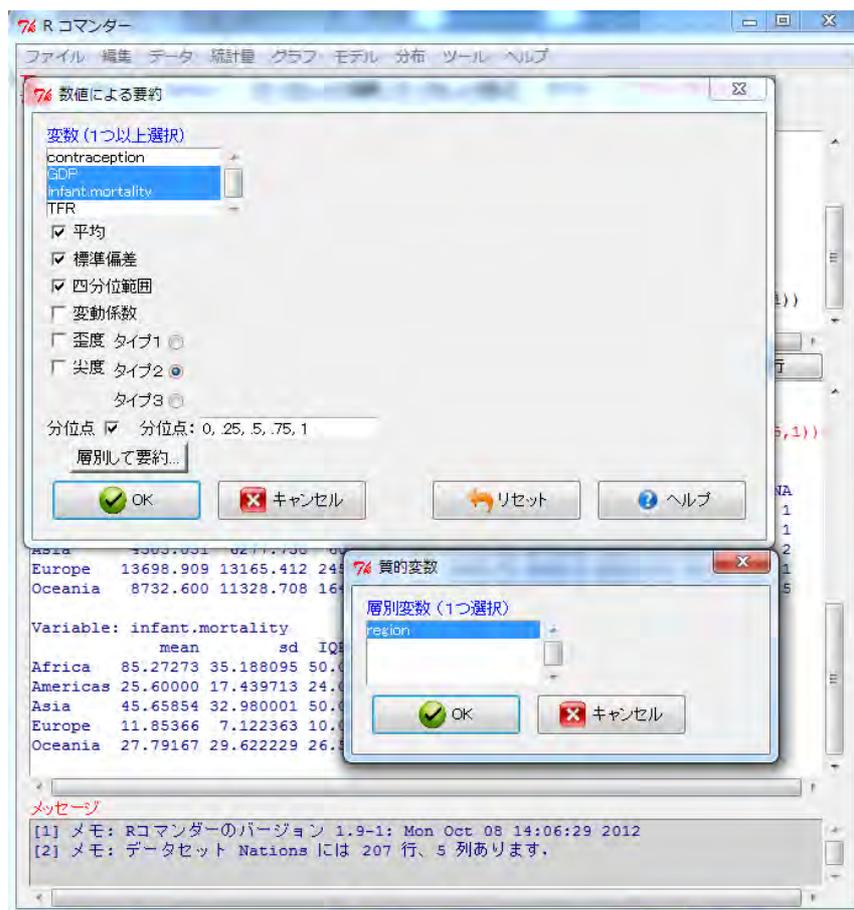


図 12 質的変数ダイアログボックスで層別変数を選択する

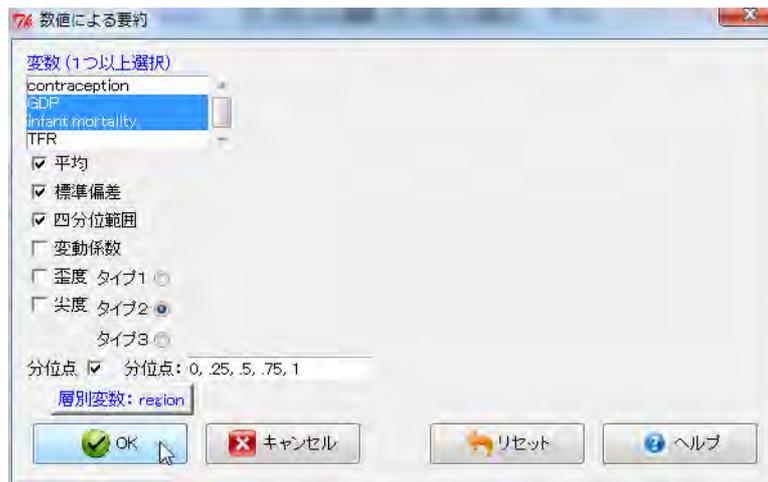


図 13 層別変数を指定した後の数値による要約ダイアログボックス



図 14 ヒストグラムのダイアログボックス

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 コマンダー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... メニューにより）、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、

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。

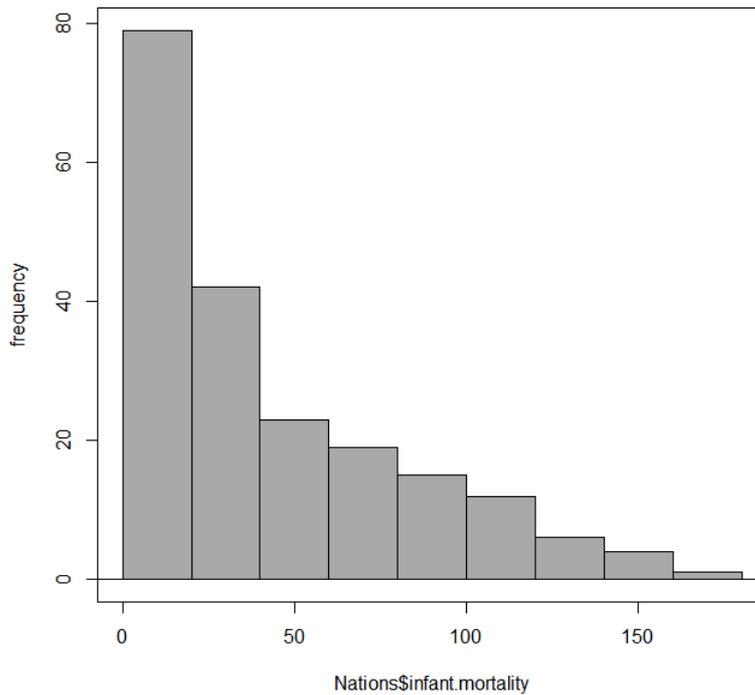


図 15 Nations データセットの `infant.mortality` のヒストグラムを表示するグラフィックスウィンドウ

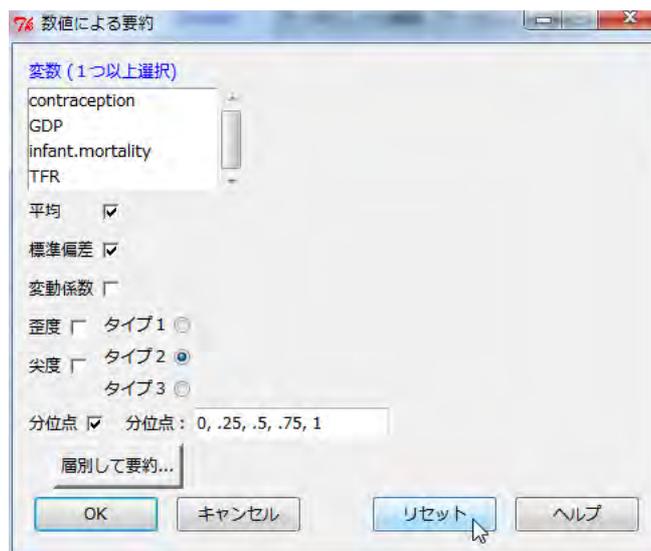


図 16 数値による要約ダイアログボックス — `dialog.memory` オプションを **TRUE** にし、リセットボタンを表示している。

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 16 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。^{*14}

- 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。

^{*14} 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。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 と一緒にインストールされる *Introduction to R* マニュアルを参照。これには、R Console の Help メニューからアクセスすることができる。

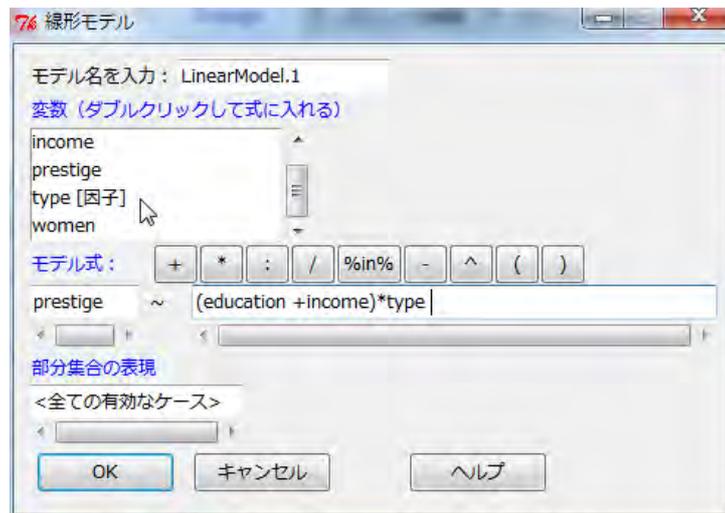


図 17 線形モデルダイアログボックス

- モデル式の上にならんでいるボタンは、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、 $\log(\text{income})$ 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 `verb=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。

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、デフォルトの“Type II”を選択すると、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 コマンドーの**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 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、表示され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。^{*15} **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、それぞれ、**R** コマンダーの**ファイルメニュー**、**グラフ** → **グラフをファイルで保存**メニューを利用することもできる。

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が、通常は *No* を選択するとよい。

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・実行するための機能を簡単に提供する。**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、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。**R** で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、プログラミングエディタやインタラクティブな開発環境、例えば、RStudio <www.rstudio.org>を用いる方がよい。^{*16}

参考文献

- [1] Fox, J. (2003). Effect displays in R for generalised linear models. *Journal of Statistical Software*, 8(15):1-27.
- [2] Fox, J. (2005). The R Commander: A basic-statistics graphical user interface to R. *Journal of Statistical Software*, 19(9):1-42.
- [3] Fox, J. (2007). Extending the Rcmdr by “Plug-in” Packages. *R News*, 7(3):46-52.
- [4] Fox, J. and Carvalho, Marilia S. (2012). The RcmdrPlugin.survival package: Extending the R Commander interface to survival analysis. *Journal of Statistical Software*, 49(7):1-32.
- [5] 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.
- [6] Moore, D. S. (2000). *The Basic Practice of Statistics, Second Edition*. Freeman, New York.
- [7] Venables, W. N. and Ripley, B. D. (2002). *Modern Applied Statistics with S, Fourth Edition*. Springer, New York.

^{*15} 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り、一般に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。

^{*16} **R** コマンダーは、**RStudio** の元で機能する。しかし、デフォルトでは、**R** コマンダーの出力とメッセージは **RStudio** の中で **R** コンソールに送られるが、不安定性や **RStudio** のグラフィックスデバイスとの不適合性といった問題がある。