

# 統計解析ソフト R のスクリプト集

[Ver. 15.0]

本冊子では、データファイルとして CSV (コンマ区切り) (\*.csv) で保存した Shift JIS コードの csv ファイルを用いることを前提としており、`read.table` や `write.table` において、`fileEncoding = "shift-jis"` オプションを付けています。UTF-8 CSV ファイルを用いる場合は、このオプションを削除してください。

2024年4月

名古屋大学 大学院教育発達科学研究科・教育学部

心理発達科学専攻 計量心理学領域

石井研究室

## はじめに

この冊子は、心理学や教育学の研究でよく用いられる統計手法に関して、統計解析ソフト「R」のスク립ト（プログラム）をまとめたもので、授業の補助教材として作成されたものです。

Rは、スク립トと言われるプログラムを書いてそれを実行する、という使い方をします。例えば、クロス表を表示するとしたら、スク립ト画面に `table(d1$x1, d1$x2)` と書き、この部分を選択して実行します。すると、出力画面に、実行したスク립トとその結果が表示されます。

ここで、`table` がゴシック体になっていますが、これは、`table` がRの関数であることを表しています。スク립ト画面上でこのような表示がされるわけではありません。この冊子においては、Rの関数とそうでないもの（分析者が指定する変数名など）を区別するために、このような表示をしています。

本冊子では、重要と思われるコメントもゴシック体にしてあります。これは基本的に和文で、「#」を前に書いてコメントアウトしてあるので（Rは認識しない）、それがRの関数でないことはすぐわかると思います。

正しいスク립トでないとソフトは動かず、エラーとなります。正しいスク립トを書いて結果が表示されると、とても嬉しい気持ちになります。また、スク립トを書くのに慣れてくるにつれ、分析法を考える力が増し、統計分析力が向上します。

本冊子はRの実行に主眼を置いているため、分析結果をどのように解釈するかについてはほとんど触れていません。統計分析の手法や結果の解釈については、別の成書にあたって下さい。参考までに、本冊子と関連の強い書籍を挙げておきます。1冊目には、本スク립ト集と同じデータを用いた分析例が出てきます。

人間科学のための統計分析—ここに關心があるすべての人のために— 医歯薬出版 2014  
<http://www.amazon.co.jp/dp/4263731611/>

統計分析のここが知りたい—保健・看護・心理・教育系研究のまとめ方— 文光堂 2005  
<http://www.amazon.co.jp/dp/4830644605/>

Essentials of Statistical Analysis and Psychometrics in Human Sciences.  
 See Sein Publishing:Myanmar, 2019 「人間科学のための統計分析」の英語版)

公認心理師カリキュラム準拠 臨床統計学 [心理学統計法・心理学研究法] 医歯薬出版 2020  
<https://www.amazon.co.jp/dp/4263266358/>



Rは開発環境が公開されているため、多くの研究者がボランティアで開発に関わっている、世界的に定評のあるフリーソフトです。新しい統計手法がすぐに取り入れられるという利点がありますが、同じ分析をするのでも開発者が複数いて、多数の関数が存在することがあったり、古い関数が使えなくなったりするという欠点もあります。ユーザー自身が使いやすい関数を選び、時折、知識を更新する必要があります。

Rは日々開発されていますので、この冊子の内容もすぐに古いものになってしまいます。ですので、こんな関数があるとか、ここの記述は間違ってる、これはもう使えないなど、気づいたことがありましたら、教えて頂ければ幸いです。

問い合わせ先 ((a)を@に変えて下さい)  
 ishii.lab.nuedu.psychometrics(a)gmail.com

石井研究室Webサイト  
<http://www.educa.nagoya-u.ac.jp/~ishii-h/>

石井 秀宗

## 目次

はじめに .....	2
<b>1 Rを使うために</b>	
Rの概要 .....	9
Rを使うための準備 .....	10
管理者権限について .....	11
Rのインストール .....	13
Rの起動と終了 .....	17
パッケージのインストール .....	18
パッケージの読み込み・取り外し .....	19
スクリプトの新規作成・保存・読み込み .....	20
スクリプトの実行 .....	22
オブジェクトの確認・消去 .....	23
ヘルプの使い方 .....	24
スクリプトの外部ファイルへのコピー .....	25
出力結果の外部保存 - Wordファイルなどに保存 .....	26
出力結果の外部保存 - sink関数 .....	27
図の外部保存 - Wordファイルなどに貼りつけ .....	29
図の外部保存 - PDFファイルに保存 .....	30
出力結果をExcelで整形保存 .....	31
Rの情報源 .....	34
Mac版のRで作業ディレクトリを設定 .....	35
Mac版のRでread.table関数でCSVファイルを読み込み .....	37
Mac版のRで図を描いたときに日本語を正しく表示させる .....	38
R Studioについて .....	39
Google Colaboratory で R を使うための準備 .....	41
Google Colaboratory で R を使う .....	44
Google Colaboratory にCSVデータをアップロードする .....	45
Google Colaboratory に日本語フォントをインストールする .....	46
Google Colaboratory のノートブックを保存する .....	47
Google Colaboratory のノートブックを共有する .....	48
<b>2 Rのスクリプト作成時の注意点とエラーへの対処</b>	
スクリプト作成にあたっての一般的注意 .....	49
スクリプトの冒頭に作業メモリを空にするコマンドを書く	
バックスラッシュと円マーク	
コメントを書くときは # を使う	
半角カナ文字を使わない	
全角文字もなるべく使わない	
スクリプトの各行の先頭に「>」「+」「1:」などを書かない	
データファイルの変数名とRスクリプトの変数名を一致させる	
Rの予約語をオブジェクト名に用いない	
pi, T, Fなどをオブジェクト名に用いない	
代入結果を表示するときは当該のスクリプト部分を () で囲う	
改行の位置に気をつける	
欠測のあるデータの分析について	
四捨五入について	
作成したスクリプトを少しずつ実行して確認する	
よくある質問・エラー .....	55
データが読み込めない (rm, setwd, read.table)	
添え字が許される範囲外です と表示される	
作成した覚えのない変数が存在する	
incomplete final line found by readTableHeader と表示される	
結果の数値に e-01 のような表記が付く	
パッケージが開けない	
パッケージの関数が使えない	
エラーが発生したときの対処 .....	58
半角カナ文字を使っていませんか?	
お使いのPCのOSは何ですか?	
Excelでデータを保存したCSVデータと、Rでの読み込み形式は合っていますか?	
作業メモリ上に余計なデータが保存されていませんか?	

実行するスクリプトの範囲選択は正しいですか？  
 データのあるフォルダを正しく設定していますか？  
 データファイル名は正しいですか？  
 変数名を間違えていませんか？  
 データファイルの変数名にスペースが入っていませんか？  
 全角記号が混ざっていませんか？  
 スペルミスをしていませんか？  
 スクリプトを書くときに「>」「+」「1:」などを左端に書いていませんか？  
 大文字と小文字を間違えていませんか？  
 記号を間違えていませんか？  
 カンマを忘れていませんか？  
 カンマを多く入れていませんか？  
 カンマとピリオドを間違えていませんか？  
 文字列をダブルクォーテーションで囲うのを忘れていませんか？  
 カッコをつけ忘れていませんか？  
 カッコの種類を間違えていませんか？  
 カッコで括らなければならないのを忘れていませんか？  
 カッコをつける位置を間違えていませんか？  
 オプションは正しく書かれていますか？  
 if - else文を書くとき、if 文の終わり と else 文の始まりの間で改行してませんか？  
 1行に複数の命令文を書いていませんか？  
 1つの数式を複数行に分けて書くとき、次の行の先頭が+ - \* / などの記号になっていませんか？  
 パッケージは正しくインストールされていますか？  
 R Studioを使っていますか？  
 Rが複数個、起動していませんか？  
 コンピュータが再起動待ちの状態になっていませんか？  
 全角文字を使っていることが原因のこともあります  
 パソコンのユーザー名が全角文字になっていることが原因のこともあります  
 Rやコンピュータを再起動することも有効です  
 Rを再インストールすることも有効です

### 3 データの作成・読み込み・保存

Excelにおけるデータファイルの作成	64
Excelを使わずにCSVファイルを作成	65
ExcelにおけるデータファイルのCSV保存	66
作業ディレクトリの指定	67
データの読み込み	68
読み込んだデータの確認	70
R上のデータのCSV保存	71
UTF-8 CSVファイルをExcelで読み込む方法	73
Excelデータファイルの読み込み・保存	74

### 4 データ操作

欠測値の取り扱い	77
データ値に基づいたカテゴリ変数の生成 — 離散データ	81
データ値に基づいたカテゴリ変数の生成 — 連続データ	82
カテゴリ変数の再カテゴリ化 — 水準の合併	83
不要な水準の削除・必要な水準の追加	84
回答データの採点	85
逆転項目のデータ処理	87
尺度得点の計算	88
データの標準化・中心化・偏差値化	89
行名・列名（変数名）・要素名の指定	91
行や列の抽出・削除	93
データの並べ替え	96
データの結合	97
対応のあるデータ（縦断的データ）のフォーマット変換 — Long Format and Wide Format	100
データ型の確認・変換	104
データ構造の確認・変換	105
文字型の数字を数値型の数値に変換	106
factor型のデータ値の扱い	108

### 5 統計図表

度数分布表	110
クロス集計表	111

割合なども入ったクロス集計表	113
群別のクロス集計表	116
フラットなクロス集計表	1117
フラットなクロス集計表から通常のクロス集計表を作成	120
円グラフ	122
帯グラフ	123
棒グラフ	126
ヒストグラム	129
2群の積み上げヒストグラム	131
箱ひげ図	132
折れ線グラフ	133
散布図	134
ひまわり図	135
バブルプロット	136
ラベルつき散布図	137
グラフオプション	138
ggplotを使った作図	141
度数を比較する棒グラフ — ggplot	143
平均値を比較する棒グラフ — ggplot	144
帯グラフ — ggplot	145
ヒストグラム — ggplot	146
箱ひげ図 — ggplot	147
折れ線グラフ — ggplot	148
散布図 — ggplot	149

## 6 量的変数の記述統計量

平均, 標準偏差, 最大値, 中央値, 最小値, 分位数	150
群別の平均, 標準偏差, 最大値, 中央値, 最小値, 分位数	151
要約統計量 — describe, describeBy	153
2要因以上の群別要約統計量	155
共分散・相関係数	156
群別の共分散・相関係数	157
偏相関係数	159
四分相関係数・多分相関係数	161
アルファ係数	163
オメガ係数	165
級内相関係数	166

## 7 1群または2群の平均値に関する推測 — t検定

1群の平均値の推測	167
対応のある2群の平均値の比較	168
対応のある2群の平均値差の効果量の推定	169
対応のない2群の平均値の比較 — 素データ	170
対応のない2群の平均値の比較 — 要約統計量	171
対応のない2群の平均値差の効果量の推定	172
平均値の非劣性・同等性の検証 — 素データ	175
平均値の非劣性・同等性の検証 — 要約統計量	177

## 8 分散に関する推測

2群の分散の比較 — 素データ	178
2群の分散の比較 — 要約統計量	179
多群の分散の比較 — 素データ	180
多群の分散の比較 — 要約統計量	181

## 9 多群の平均値に関する推測 — 分散分析・多重比較

1つの被験者間要因(1 B)	182
1つの被験者内要因(1 W)	189
2つの被験者間要因(2 B)	193
1つの被験者間要因と1つの被験者内要因(1 B 1 W)	1202
2つの被験者内要因(2 W)	208
3つの被験者間要因(3 B)	213
2つの被験者間要因と1つの被験者内要因(2 B 1 W)	224
1つの被験者間要因と2つの被験者内要因(1 B 2 W)	230

3つの被験者内要因(3W) .....	237
<b>10 分布の位置に関する推測 — ノンパラメトリック法</b>	
対応のある2群の分布の位置の比較 — ウィルコクソンの符号順位検定 .....	244
対応のある多群の分布の位置の比較 — フリードマンの検定 .....	246
対応のない2群の分布の位置の比較 — ウィルコクソンの順位和検定(マン・ホイットニーの検定) .....	248
対応のない多群の分布の位置の比較 — クラスカル・ウォリスの検定 .....	250
<b>11 相関係数に関する推測</b>	
相関係数の推測 .....	252
相関係数行列の各要素の検定 .....	253
2群の相関係数の差の検定 .....	254
2群の相関係数の差の信頼区間 .....	255
多群の相関係数の等質性の推測 .....	257
<b>12 分割表に関する推測 — <math>\chi^2</math>検定・フィッシャーの正確検定</b>	
2×2表 — ファイ係数, カイ2乗検定, 尤度比検定, フィッシャーの正確検定 .....	260
r×c表 — クラメルの変数係数, カイ2乗検定, 尤度比検定 .....	263
ピアソン残差・デビアンズ残差 .....	266
2名の評定者の評定の一致度 — コーエンのカップ係数 .....	269
複数名の評定者の評定の一致度 — Fleissのカップ係数 .....	271
<b>13 比率に関する推測</b>	
1群の比率の推測 — 二項検定 .....	272
対応のある2群の比率の比較 — マクネマーの検定 .....	273
対応のある多群の比率の比較 — コ克蘭のQ検定 .....	275
リスク差・リスク比・オッズ比の推測 .....	277
対応のない多群の比率の等質性の推測 .....	282
<b>14 線形回帰分析</b>	
単回帰分析 — lm関数 .....	284
重回帰分析 — lm関数 .....	286
重回帰分析 — glm関数 .....	289
重回帰分析 — sem関数 .....	291
主成分得点を用いた重回帰分析 .....	296
<b>15 パス解析</b>	
パス解析 — specifyEquations関数 .....	299
パス解析 — specifyModel関数 .....	303
多母集団のパス解析 — specifyEquations関数 .....	307
多母集団のパス解析 — specifyModel関数 .....	311
多母集団パス解析におけるパス係数の検定 .....	315
<b>16 因子分析</b>	
確認的因子分析 .....	321
探索的因子分析 .....	326
2次因子分析 .....	329
階層因子分析 .....	335
<b>17 潜在変数のある構造方程式モデリング(共分散構造分析)</b>	
多重指標モデル .....	342
潜在変数のある構造方程式モデリング — semパッケージ .....	345
潜在変数のある構造方程式モデリング — lavaanパッケージ .....	349
成長曲線モデル — lavaan パッケージ .....	353
多母集団分析 — semパッケージ .....	356

多母集団分析 — lavaanパッケージ	363
等値制約 — semパッケージ	368
等値制約 — lavaanパッケージ	374

## 1 8 さまざまな多変量データ解析

クラスター分析	378
対応分析	382
多重対応分析	385
主成分分析	389
主成分回帰分析・PLS回帰分析	391
ロジスティック回帰分析 — 素データ	394
ロジスティック回帰分析 — 集計データ	397
多項ロジスティック回帰分析	400
対数線形モデル — glm関数	403
対数線形モデル — loglm関数	410
カウントデータの分析	417

## 1 9 階層線形モデル（マルチレベルモデル）

階層線形モデルの例	419
Rで階層線形モデル分析をするためのデータ構造	420
階層線形モデル — lmer関数	421
階層線形モデル — lme関数	428

## 2 0 検定力分析

2群の平均値	434
相関係数	436
クロス集計表	438
2群の比率	440
1つの被験者間要因(1 B)	443
1つの被験者内要因(1 W)	446
2つの被験者間要因(2 B)	449
1つの被験者間要因と1つの被験者内要因(1 B 1 W)	453
2つの被験者内要因(2 W)	458

## 2 1 信頼区間に基づく標本サイズの推定

1群の平均値	463
2群の平均値	464
相関係数	465
比率	466

## 2 2 項目分析

項目分析 — 自作関数	467
項目分析 — itemanalysisパッケージ	477

## 2 3 項目応答理論

項目パラメタの推定 — 1PL, 2PL, 3PLモデル	481
特性関数・情報関数の図示	485
特性値の推定	487
局所独立性の確認	489
等化 — 共通項目法	492
D I F分析	498

## 2 4 付録

文字列関数 — 文字列の切り出し・結合・検索・置換	504
演算記号・算術関数	507
切り上げ・切り下げ・四捨五入	509
集合関数	511

比較演算子 .....	512
論理演算子 .....	513
ベクトルの生成 .....	514
ベクトルの要素の評価 .....	515
行列演算 .....	517
制御コマンド .....	521
コマンドを生成して実行 .....	524
一括分析・総当たりの分析 .....	527
関数の作成 .....	530
スクリプトの呼び出し .....	531
多変量正規乱数の発生 .....	532
確率関数 .....	534
いくつかの図をまとめてPDFファイルに出力 .....	535
正規分布と標準得点の対応関係の図を描く .....	537
統計的有意性検定の概念図を描く .....	541
同なるべくグループメンバーが重ならないグループ分けを複数回行う .....	543



## Rの概要

## Rの画面

The screenshot shows the R RStudio environment. On the left, the console window displays statistical results for a principal component analysis, including the total number of observations, the number of variables, and the eigenvalues of the components. The script editor in the center shows R code for performing PCA using the `prcomp` function. On the right, a plot window titled "scree plot" shows the eigenvalues of the components, with the first two components having significantly higher eigenvalues than the others.

結果表示画面

スクリプト編集画面

図出力画面

## どういふものか

- ・統計分析や統計図表の作成ができる
- ・SPSS や EXCELより, SAS や FORTRAN に近い
- ・スクリプト (プログラム) を書く
- ・本体とパッケージからなる
  - 本体 データ管理, 編集, 領域普遍的な分析
  - パッケージ 領域固有的な分析, 便利機能

## デメリット

- ・取っ付きが悪い
- ・関数名や文法を知らないと使えない
- ・正しくスクリプト (プログラム) を書かないと動かない
- ・パッケージが山のようにあって, どれを使えば良いのか調べるのが大変
- ・同じ分析を実行するパッケージや関数が複数あったりする
  - 分散分析 aov, anova, Anova
  - 共分散構造分析 sem, lavaan
- ・異なるパッケージの中に, 同じ名前の関数があったりする
- ・パッケージ, 関数によって, 仮定や数値計算法の違いなどから, 結果が多少異なることがある

## メリット

- ・フリーソフト
- ・世界標準
- ・世界中のボランティア開発者が, よってたかってパッケージを開発
- ・新しい分析機能がどんどん追加されていく
- ・信頼性が高い
- ・分析のスクリプト (プログラム) を保存
  - 同じ分析を簡単に再現できる
  - SPSSは分析結果を保存し, 過程を保存しないので, 過去の分析を再現するのが難しいことがある
  - Rは, 分析結果ではなく, スクリプト (分析方法, 過程) を保存するので, 結果を容易に再現できる

## Rを使うための準備

### Rを使うにあたって必要なもの

- ・インターネットにつながるパソコン  
Rをダウンロード・実行するのに必要  
Windows, Mac, Linux いずれでもよい  
(本スクリプト集は基本的にWindows環境で作成しています)
- ・エクセルなどデータ入力できるソフト  
データをCSV形式で保存するのに使う  
他の形式 (エクセル形式など) でも構わないのだが, 汎用性の高いCSV形式を推奨  
表入力・CSV保存できるソフトなら何でも構わない
- \* Google Colaboratory を使って, インターネット上で R を使うこともできます。その場合は, Google アカウントが必要です。  
Google Colaboratory でRを実行する場合は, OSの違いはあまり意識しないで済みますが, インターネットへの常時接続が必要である, CSVデータを毎回アップロードしなければならない, 利用時間に制約があるなど, いくつか不便な点もあります。

### Rを使うにあたって必要なこと

- ・所定のWebサイトからRファイルをダウンロードしてパソコンにRをインストールする  
フリーソフトなので料金はかからない  
安全なソフトとされている
- ・エクセル等でデータを入力しCSV形式で保存する  
データのある場所が, Rの作業領域になるので, 処理の速さが気にならないければ  
USBなどの外部メモリに保存しておくで良い。(バックアップは別にとっておく)
- ・Rで分析を行うための, スクリプト (プログラム) の書き方, 関数名を覚える  
データの読み込み, 書き出しのためのスクリプト  
分析のためのスクリプト  
データ加工のためのスクリプト など  
おそらくこの部分が一番のハードルになる

### ファイルの拡張子を表示させる

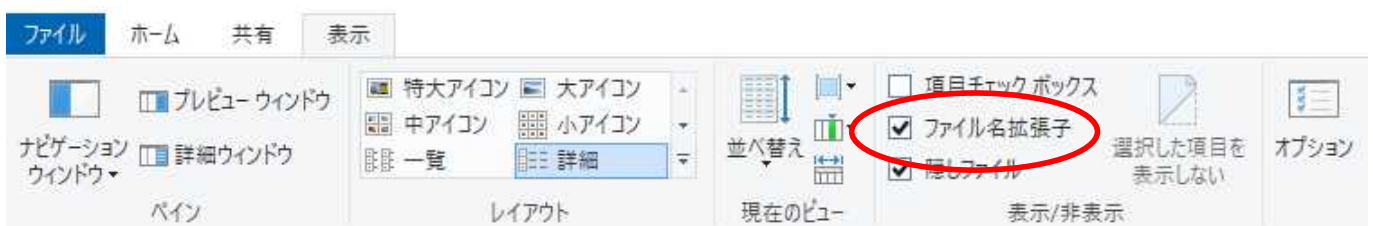
たいていの場合, コンピュータ上のファイル名には, 拡張子というものがついています。～. pdf, ～. docx の. pdfや. docxなどです。ファイルの属性を表すのに用いられています。

Windowsのエクスプローラは, 初期設定では拡張子を表示しません。この状態だと, ～の部分と同じファイル名のファイルが, PDFファイルなのかWordファイルなのか, 表示されるファイル名だけでは分からなくなります。

そこで, エクスプローラにおいて, 拡張子も表示させるように設定を変更しておきます。

エクスプローラ → 表示 → ファイル名拡張子をクリック としてください。そうすると, 拡張子がついたファイル名が表示されます。

ファイルの種類を確認するためにも, 拡張子を常に表示させることをお勧めします。



## 管理者権限について

Windowsマシンの場合、管理者としてパソコンを実行していないとRをインストールできないことがある。このような場合は、パソコンの管理者にRをインストールしてもらうか、自分にも管理者権限を持たせてもらう必要がある。Macの場合は、管理者権限がないとRを実行できないので、管理者権限を意識することはまずない。

管理者権限を持っていなくてもRを使うことはできる。しかし、Rを管理者権限として実行できていないと、パッケージをインストールした際、Rのライブラリフォルダに保存されず、別のフォルダにインストールされる場合がある。例えば、パッケージのインストールを行った際、以下のような警告メッセージが出ることもある。

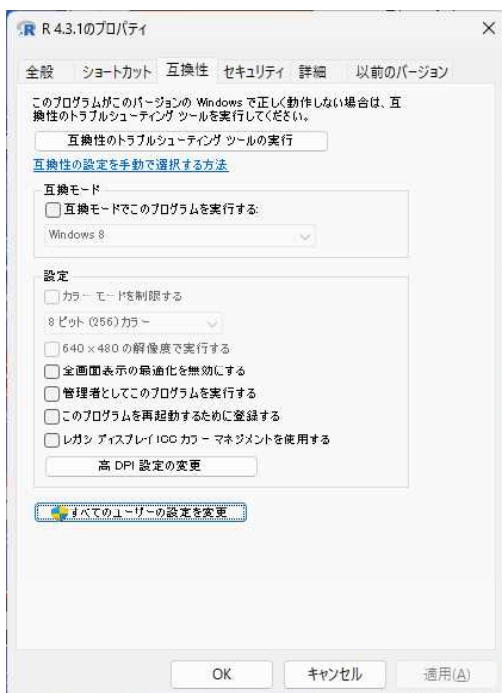
```
Warning in install.packages(NULL, .libPaths()[1L], dependencies = NA, type = type) :
'lib = "C:/Program Files/R/R-3.2.0/library"' は書き込み可能ではありません
ダウンロードされたパッケージは、以下にあります
C:\Users\%○○○\AppData\Local\Temp\RtmpWi0uAi\downloaded_packages
```

このような状況では、パッケージはR本体とは別のフォルダにインストールされているため、library()関数で読み込もうとしても、「○○というパッケージはありません」というエラーが出る。

上のような状況を解決するためには、管理者権限のあるユーザーとしてRを実行して、パッケージをインストールする。Rを管理者権限として実行するためには、Rのプロパティを変更する。



デスクトップ上のRのアイコン上で右クリックし、左のようなメニューを表示し、メニューの中から、「プロパティ」を選択する。



「プロパティ」を選択すると、左のようなウィンドウが表示される。

プロパティウィンドウにある複数のタブの中から、「互換性」タブを選択する。

「すべてのユーザーの設定を変更」のボタンをクリックする。

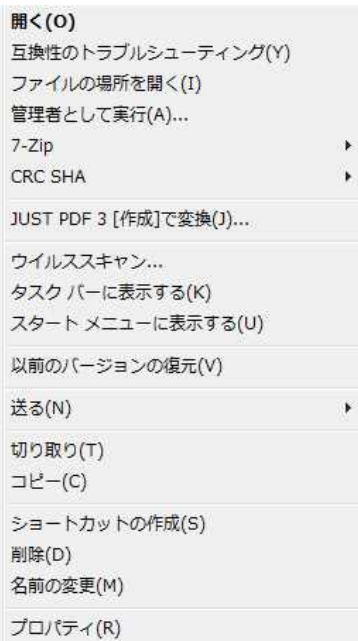


「管理者としてこのプログラムを実行する」のチェックボックスにチェックを入れ、「適用」を選択する。

表示させたすべてのウィンドウについて「OK」を選択して閉じる。

次にRを起動したとき、「コンピュータへの変更を許可しますか？」と聞かれたら「はい」として続行する。

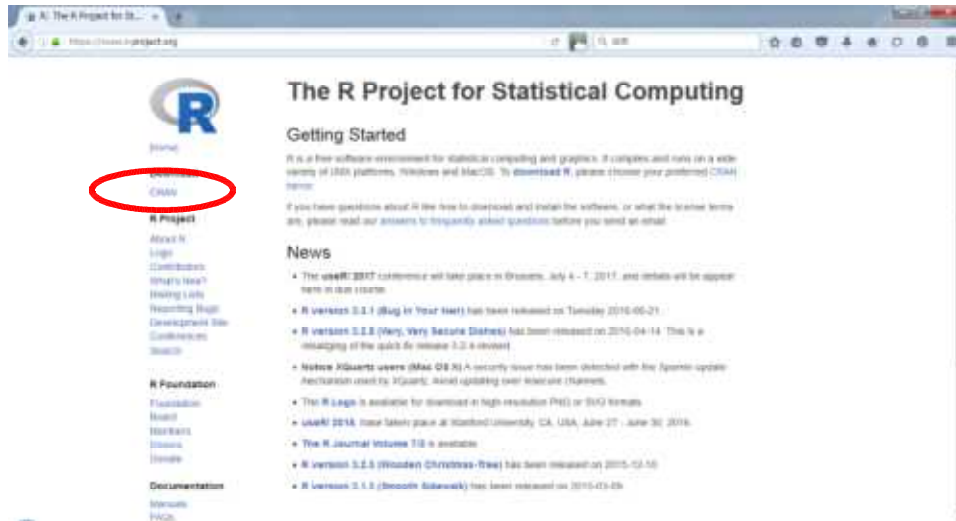
これで、管理者としての実行が可能となるので、この状態で、パッケージのインストールをすればよい。



インストール後、Rを起動するとき、デスクトップ上のRのアイコンをクリックして起動しない場合は、Rのアイコン上で右クリックし、「管理者として実行」を選択する。

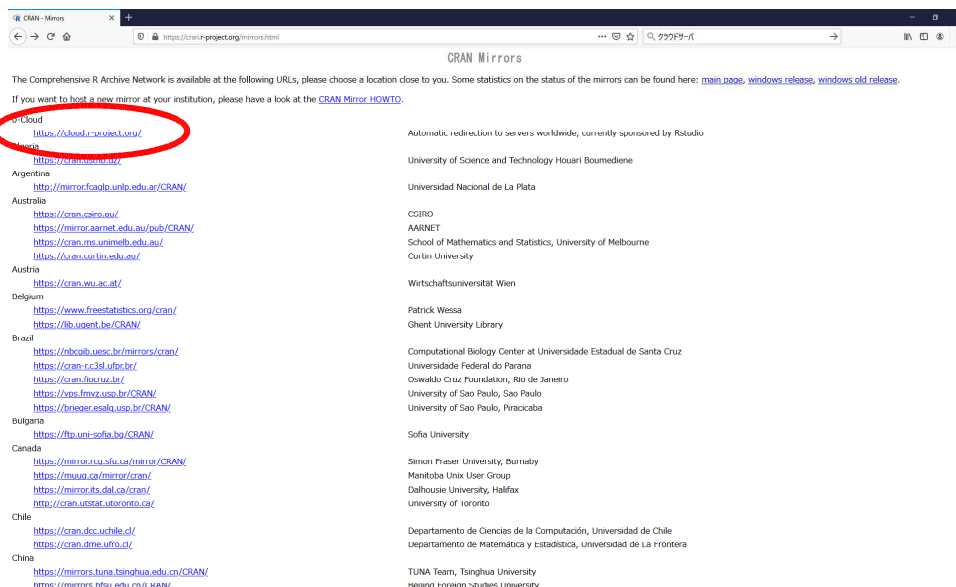
Rのインストール

【注意】 管理者または管理者権限を持ったユーザーとしてパソコンを起動（ログイン）している必要がある。管理者権限については前項参照。

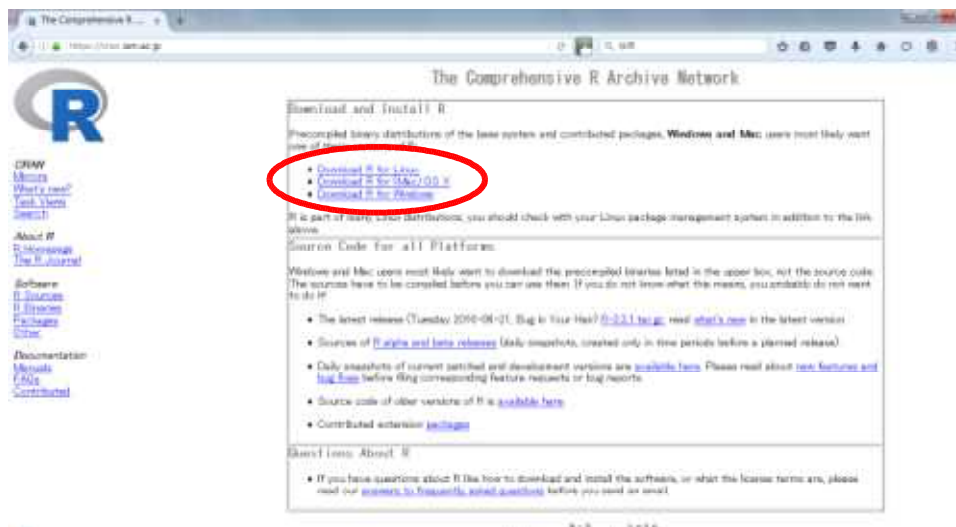


http://www.r-project.org/ を開く。

Downloadの下にある CRAN をクリックして、最寄りのミラーサイトを選ぶ。



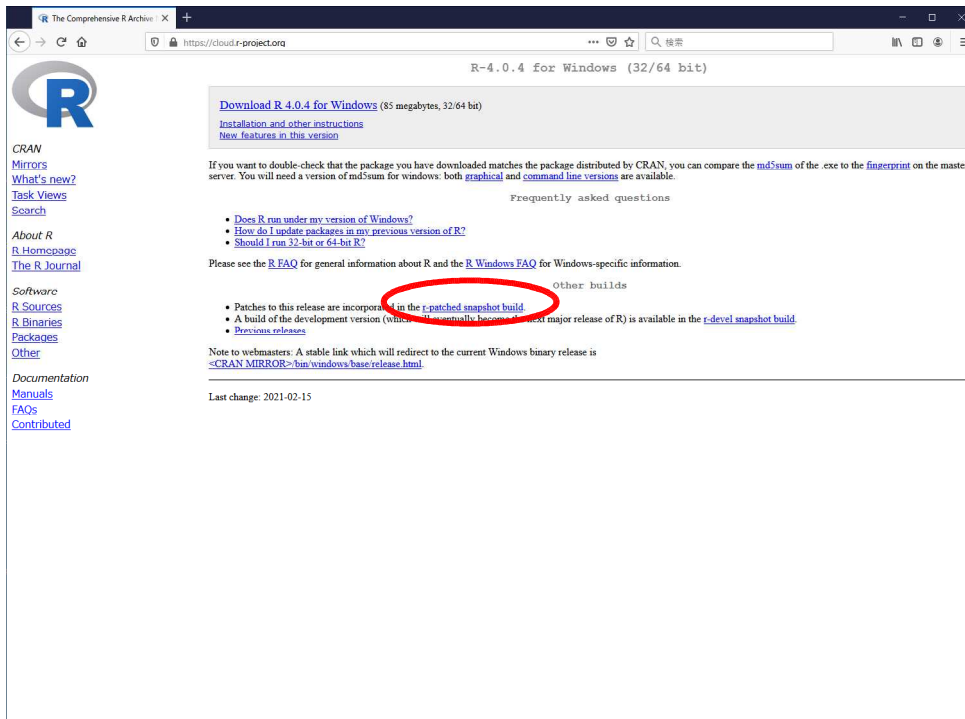
一番上の 0-cloud, もしくは, Japan の The Institute of Statistical Mathematics, Tokyo からダウンロードするのが分かりやすい。



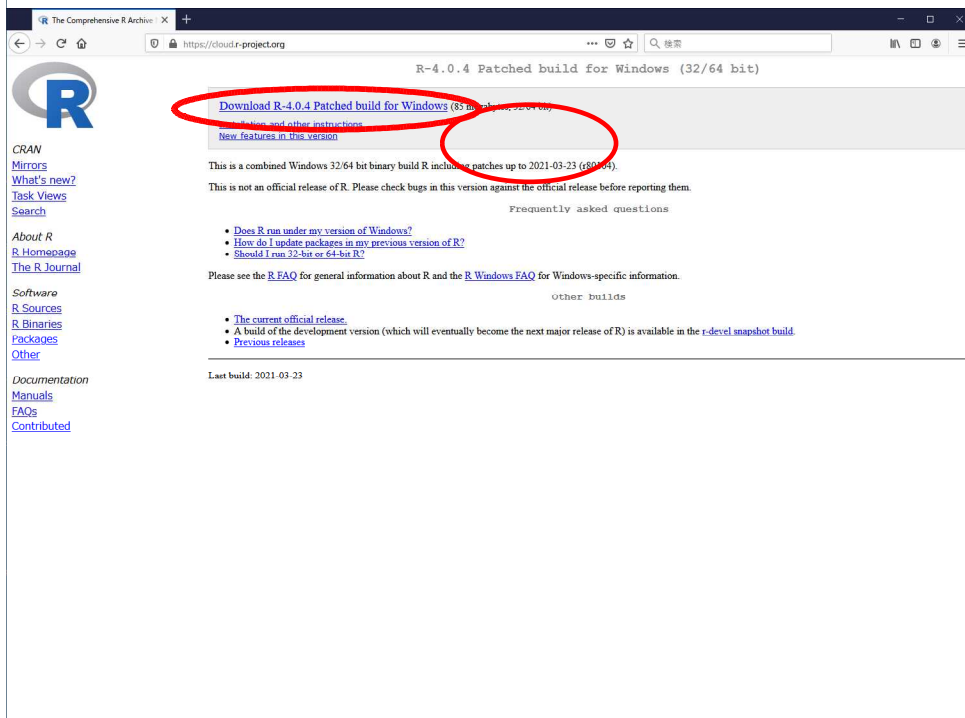
Download and Instal R で, Mac, Windows, Linux のうち, 自分の使っているコンピュータにあうものを選ぶ。



baseを選択する。



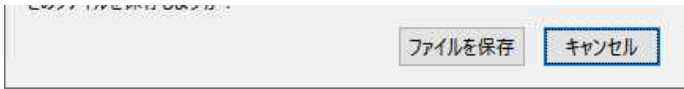
当初リリースの不具合を修正したPatched buildがあればクリックする。



一番上のDownload R... を選ぶ。

ファイルを保存するか聞いてくるので、保存するを選ぶ。

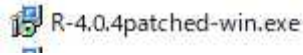
ダウンロードしたファイルが、「ダウンロード」というフォルダか、「デスクトップ」か、どこかにあるはずなので、探し出す。



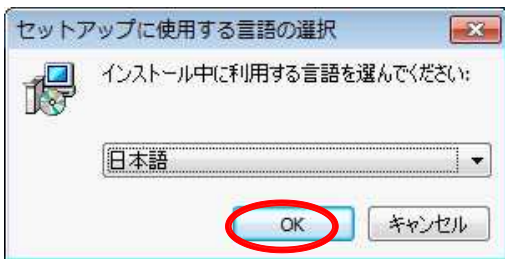
デスクトップに保存した場合



ダウンロードフォルダ等に保存した場合



ダウンロードしたファイル「R-〇〇.exe」をダブルクリックして、インストールを開始する。  
〇〇のところは、RのバージョンやOSなどにより異なる。



利用言語を選びOKとする

特に構わなければ  
次へ 次へ 次へ  
...

32bitマシンか 64bit マシンか選択できるときは、自分のあてはまるほうを選ぶ。(あてはまらないほうのチェックを外す)









































































