

EZR を用いた演習の手引き

作成日： 2025 年 3 月 1 日

作成者： 原田 亜紀子

● 演習用データの内容

500 人（男性 160 人，女性 340 人）の健康診査の結果および生活習慣についての仮想データ（Microsoft Excel ファイル）です。

● EZR とは

EZR は，統計解析に用いられるプログラミング言語である R の GUI 環境の 1 つである R コマンドーを，自治医科大学附属さいたま医療センターの神田善伸教授が医学統計用にカスタマイズしたものです。

本教材は，R Version 4.2.2，R コマンドー Version 2.8-0，EZR Version 1.6.1（Windows 版）で動作確認済みです。

A EZR の導入と基本操作

① EZR のインストール

①自治医大付属埼玉医療センターのホームページにアクセスする。

<https://www.jichi.ac.jp/saitama-sct/SaitamaHP.files/statmed.html>

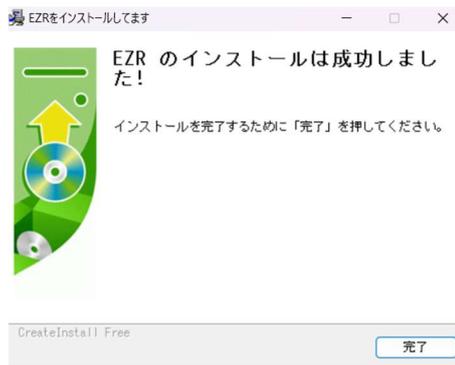
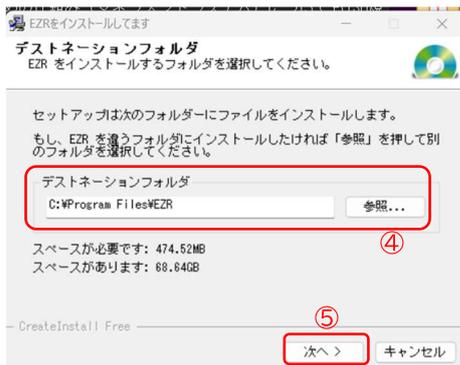
②使用している OS に合わせて，セットアップファイルをダウンロードする。

The image shows two screenshots from the EZR website. The left screenshot is the homepage, with a red box highlighting the download links for Windows, Mac OS X, and Linux. The right screenshot is the 'Windows版のインストール' (Windows Version Installation) page, with a red box highlighting the instruction to click on the download link for Windows. Both screenshots have a circled '2' next to them.

③ダウンロードした「EZRsetup」をクリックし，インストールを開始する。

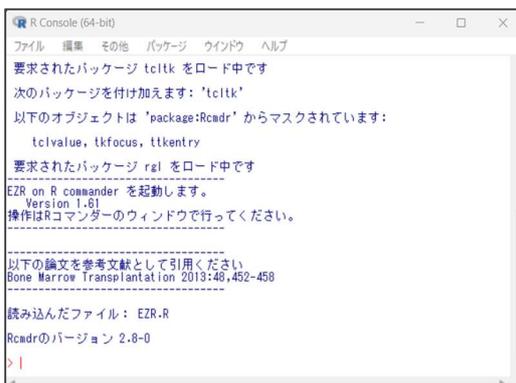
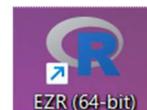
④インストール先を指定する。

⑤「次へ」で作業を進める。

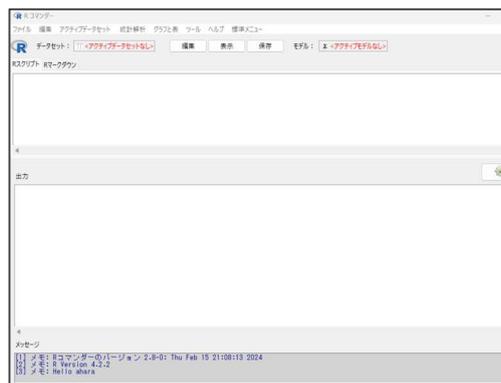


② EZR の起動とデータの読み込み

- ①ショートカット（右図）をクリックするなどして、EZR を起動する。
- ②EZR を起動すると、2 つの画面が表示される。下図左の R コンソール画面は、基本的に操作する必要はなく、下図右の R コマンド画面を操作する。



R コンソール画面



R コマンド画面

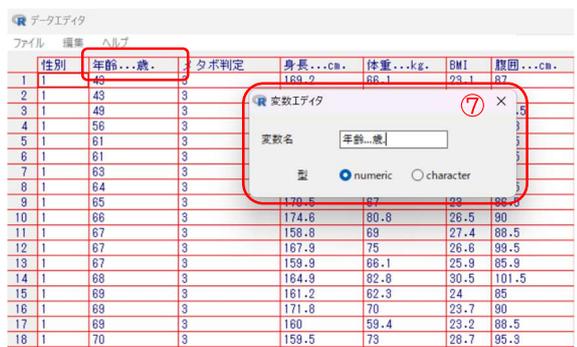
- ③データは直接入力することも、外部ファイルを読み込む（インポートする）こともできる。読み込み可能なデータ形式には、SPSS ファイル、Excel スプレッドシート、CSV ファイル、SAS データファイル、Stata データファイルなどがある。今回は、事前にダウンロードした「dataset_2.xlsx」を用いる。
- ④[ファイル]-[データのインポート]-[Excel のデータをインポート]を選択する。
- ⑤データが読み込まれる。



| | 性別 | 年齢...歳 | メタボ判定 | 身長...cm | 体重...kg | BMI | 腹囲...cm | SBP | DBP | Hb...g/dL | Ht...cm |
|----|----|--------|-------|---------|---------|------|---------|-----|-----|-----------|---------|
| 1 | 1 | 43 | 3 | 169.2 | 66.1 | 23.1 | 87.0 | 161 | 109 | 16.5 | 49.2 |
| 2 | 1 | 43 | 3 | 169.2 | 81.0 | 28.3 | 97.0 | 131 | 90 | 17.1 | 52.0 |
| 3 | 1 | 49 | 3 | 174.5 | 99.0 | 32.5 | 101.5 | 138 | 85 | 15.3 | 47.9 |
| 4 | 1 | 56 | 3 | 166.6 | 73.7 | 26.6 | 96.3 | 136 | 79 | 15.4 | 48.2 |
| 5 | 1 | 61 | 3 | 166.7 | 65.9 | 23.7 | 89.5 | 142 | 60 | 14.5 | 45.7 |
| 6 | 1 | 61 | 3 | 166.0 | 77.6 | 28.2 | 91.5 | 147 | 82 | 15.4 | 47.5 |
| 7 | 1 | 63 | 3 | 170.7 | 68.5 | 23.5 | 87.0 | 119 | 73 | 14.4 | 45.4 |
| 8 | 1 | 64 | 3 | 167.6 | 77.4 | 27.6 | 97.5 | 133 | 73 | 14.4 | 44.1 |
| 9 | 1 | 65 | 3 | 170.5 | 67.0 | 23.0 | 86.5 | 135 | 83 | 15.4 | 49.3 |
| 10 | 1 | 66 | 3 | 174.6 | 80.8 | 26.5 | 90.0 | 156 | 88 | 15.6 | 49.4 |
| 11 | 1 | 67 | 3 | 158.8 | 69.0 | 27.4 | 88.5 | 146 | 84 | 13.3 | 41.3 |
| 12 | 1 | 67 | 3 | 167.9 | 75.0 | 26.6 | 99.5 | 132 | 70 | 15.3 | 48.4 |
| 13 | 1 | 67 | 3 | 159.9 | 66.1 | 25.9 | 85.9 | 133 | 94 | 14.0 | 44.1 |
| 14 | 1 | 68 | 3 | 164.9 | 82.6 | 30.5 | 101.5 | 155 | 79 | 17.0 | 52.4 |
| 15 | 1 | 69 | 3 | 161.2 | 62.3 | 24.0 | 85.0 | 146 | 76 | 12.4 | 38.7 |
| 16 | 1 | 69 | 3 | 171.8 | 70.0 | 23.7 | 90.0 | 124 | 83 | 13.7 | 41.9 |
| 17 | 1 | 69 | 3 | 160.0 | 59.4 | 23.2 | 88.5 | 126 | 71 | 14.8 | 47.7 |
| 18 | 1 | 70 | 3 | 159.5 | 73.0 | 28.7 | 95.3 | 160 | 85 | 15.3 | 47.3 |
| 19 | 1 | 72 | 3 | 155.1 | 57.0 | 23.7 | 86.0 | 136 | 88 | 14.3 | 45.6 |
| 20 | 1 | 72 | 3 | 163.4 | 64.1 | 24.0 | 90.5 | 141 | 76 | 14.4 | 44.3 |
| 21 | 1 | 74 | 3 | 164.5 | 65.0 | 24.0 | 89.5 | 131 | 77 | 13.0 | 40.3 |
| 22 | 1 | 74 | 3 | 165.4 | 81.2 | 29.7 | 106.0 | 124 | 80 | 15.9 | 45.7 |
| 23 | 1 | 25 | 2 | 164.9 | 61.8 | 22.7 | 85.0 | 123 | 77 | 15.3 | NA |
| 24 | 1 | 31 | 2 | 178.5 | 75.2 | 23.6 | 87.0 | 157 | 109 | 14.9 | 47.0 |
| 25 | 1 | 47 | 2 | 171.8 | 72.9 | 24.7 | 89.5 | 111 | 85 | 14.0 | 42.8 |
| 26 | 1 | 48 | 2 | 159.1 | 64.8 | 25.6 | 89.0 | 122 | 79 | 16.3 | NA |

⑥データセットを編集する場合は[編集]をクリックする。

⑦編集画面で、変数名やデータ型、入力値の修正などが行える。たとえば変数名の変更は、変数名をクリックすると「変数エディタ」が表示され、可能になる。



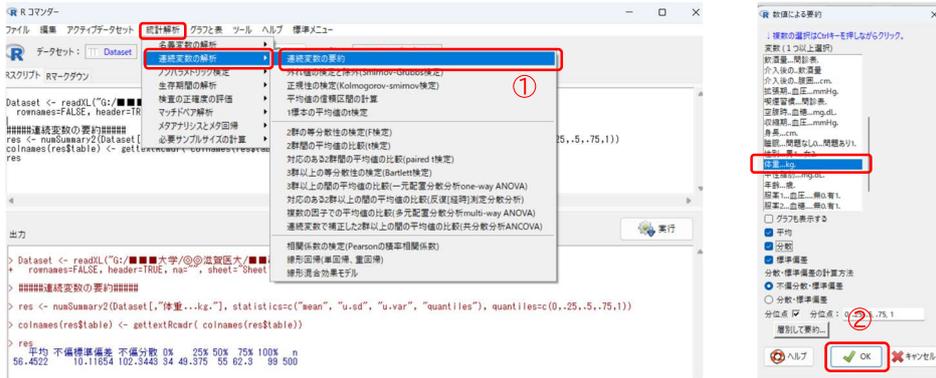
B データの種類と分布

① 度数分布表とヒストグラムの作成

a データの値の範囲を調べる（体重の集計の例）

①[統計解析]-[連続変数の解析]-[連続変数の要約]を選択する。

②数値の要約を行う変数（演習例では「体重」）を選択し、[OK]を選択する。



③出力画面に要約統計が表示される。演習例では、対象数 (n) は 500 例、平均 56.45、最小値 (0%) 34、最大値 (100%) 99 となる。

```
> res
平均 不偏標準偏差 不偏分散 0% 25% 50% 75% 100% n
56.4522 10.11654 34 49.375 55 62.3 99 500
```

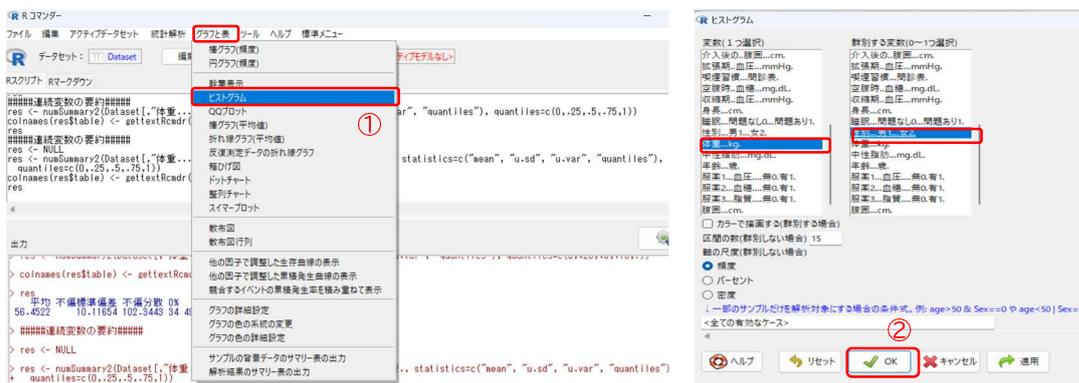
④②において、[層別して要約]→[性別...男 1...女 2]を選択したのち、[OK]を選択すると、男女別に集計される。

```
> res
平均 不偏標準偏差 不偏分散 0% 25% 50% 75% 100% data:n
1 63.28312 9.505597 42.9 57.550 62.4 67.525 99.0 180
2 53.23765 8.711024 34.0 47.775 52.3 58.025 86.2 340
```

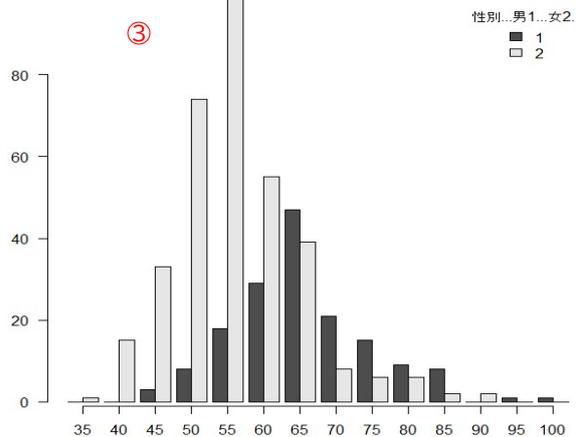
b 階級ごとに集計する

①[グラフと表]-[ヒストグラム]を選択する。

②表示する変数と群別する変数を選択し、[OK]を選択する。演習例では、それぞれ[体重]、[性別]である。



③ヒストグラムが表示される。



② 基本統計量の算出

①[標準メニュー]-[統計量]-[要約]-[数値による要約]を選択する。



②出力する変数と統計量を選択する。統計量については、平均、標準偏差、標準誤差、四分位範囲、変動係数、分位数などが選択できる。



③結果が出力される。下図の mean は平均，sd は標準偏差，IQR は四分位範囲である。

```
> numSummary(exercise[, "体重...kg.", drop=FALSE], stati
  mean      sd      IQR 0%    25% 50%  75% 100%  n
56.4522 10.11654 12.925 34 49.375 55 62.3  99 500
```

C 関連の指標

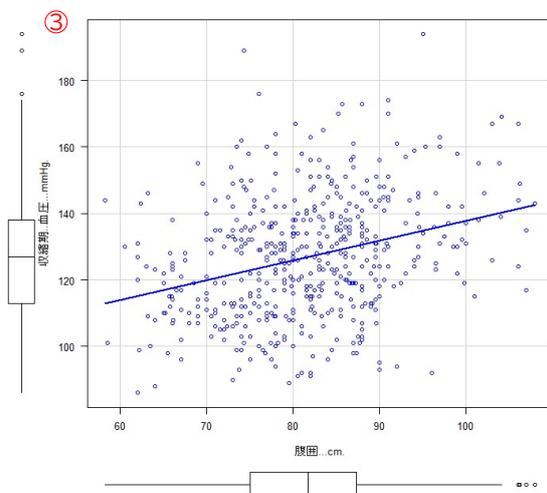
① 散布図の描画

①[グラフと表]-[散布図]を選択する。

②X と Y の変数を指定し、[OK]をクリックする。今回は、「腹囲」と「収縮期血圧」を指定する。



③散布図が出力される。



② 相関の分析

- ①[統計解析]-[連続変数の解析]-[相関係数の検定 (Pearson の積率相関係数)]を選択する。
- ②相関を分析したい 2 変数を選択し, [OK]をクリックする。今回は, 「収縮期血圧」と「腹囲」を指定する。

The screenshot shows the R Commander interface. On the left, the 'Statistical Analysis' menu is open, and 'Correlation Test (Pearson's Product-Moment Correlation)' is selected. On the right, a dialog box titled 'Correlation Test (Pearson's Product-Moment Correlation)' is open, showing a list of variables. 'exercise\$SBP' and 'exercise\$腹囲...cm.' are selected. The 'OK' button is highlighted with a red circle.

- ③結果が出力される。

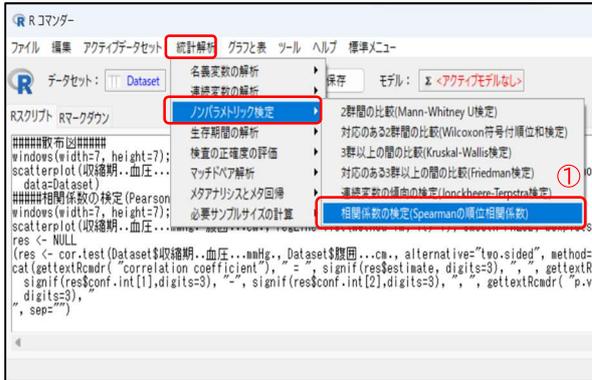
```
> (res <- cor.test(exercise$SBP, exercise$腹囲...cm., alterna
Pearson's product-moment correlation
data: exercise$SBP and exercise$腹囲...cm.
t = 7.1454, df = 498, p-value = 3.208e-12
alternative hypothesis: true correlation is not equal to 0
95 percent confidence interval:
 0.2232210 0.3824073
sample estimates:
cor
0.3049425

> cat(gettextRcmdr("correlation coefficient"), " = ", signif
+ signif(res$conf.int[1], digits=3), "-", signif(res$conf.in
+ digits=3), "
+ ", sep="")
相関係数 = 0.305, 95%信頼区間 0.223-0.382, P値 = 3.21e-12
```

■ 参考 スピアマンの順位相関係数

対象とする変数が正規分布していない場合は, ノンパラメトリックな手法であるスピアマンの相関係数による分析を行う。

- ①[統計解析]-[ノンパラメトリック検定]-[相関係数の検定 (Spearman の順位相関係数)]を選択する。
- ②相関を分析したい 2 変数を選択し, [OK]をクリックすると, 結果が出力される。



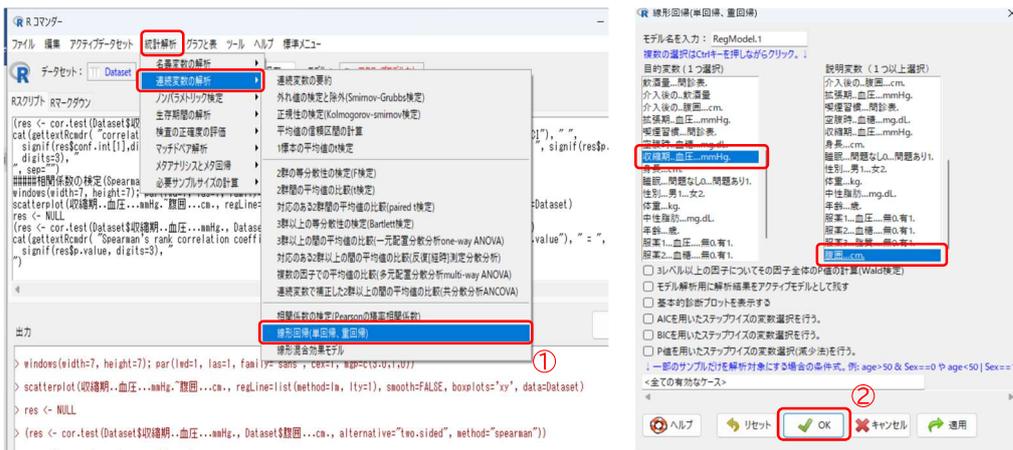
```

②
> (res <- cor.test(exercise$SBP, exercise$腹囲...cm
                Spearman's rank correlation rho
data: exercise$SBP and exercise$腹囲...cm.
S = 14587835, p-value = 7.858e-12
alternative hypothesis: true rho is not equal to 0
sample estimates:
rho
0.2997907 |
> cat(gettextRcmdr("Spearman's rank correlation coefficient",
+ "signif(res$p.value, digits=3),",
+ "Spearmanの順位相関係数 0.3 P値 = 7.86e-12

```

③ 回帰分析

- ① [統計解析]-[連続変数の解析]-[線形回帰 (単回帰、重回帰)]を選択する。
- ② 目的変数と説明変数を選択し、[OK]をクリックする。今回は「収縮期血圧」と「腹囲」を選択する。



③ 結果が出力される。

```

> multireg.table <- NULL
> multireg.table <- cbind(res$coefficients[,1], confint(RegModel.1),res$coefficients[,2:4])
> colnames(multireg.table)[1] <- "Estimate"
> colnames(multireg.table) <- gettextRcmdr( colnames(multireg.table))
> multireg.table

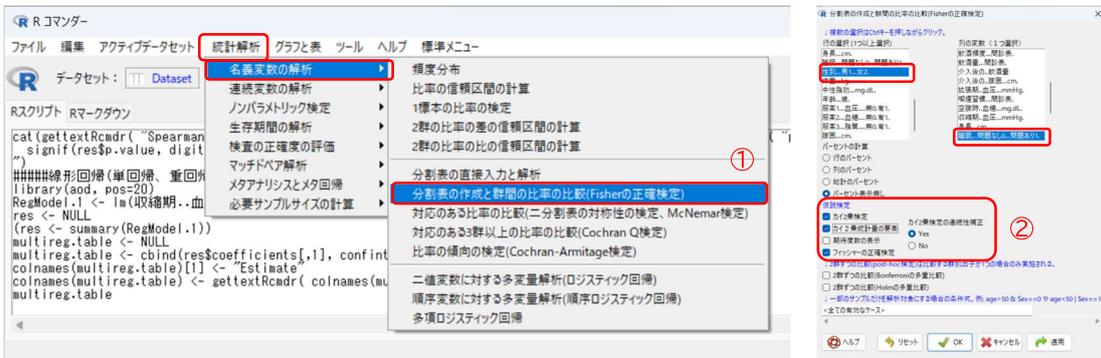
```

| | 回帰係数推定値 | 95%信頼区間下限 | 95%信頼区間上限 | 標準誤差 | t統計量 | P値 |
|-------------|------------|------------|------------|------------|----------|--------------|
| (Intercept) | 78.1970375 | 64.7263390 | 91.6677360 | 6.85622805 | 11.40528 | 6.253862e-27 |
| 腹囲...cm. | 0.5968214 | 0.4327161 | 0.7609267 | 0.08352526 | 7.14540 | 3.207820e-12 |

D 統計分析

① 独立性の検定 (χ^2 検定)

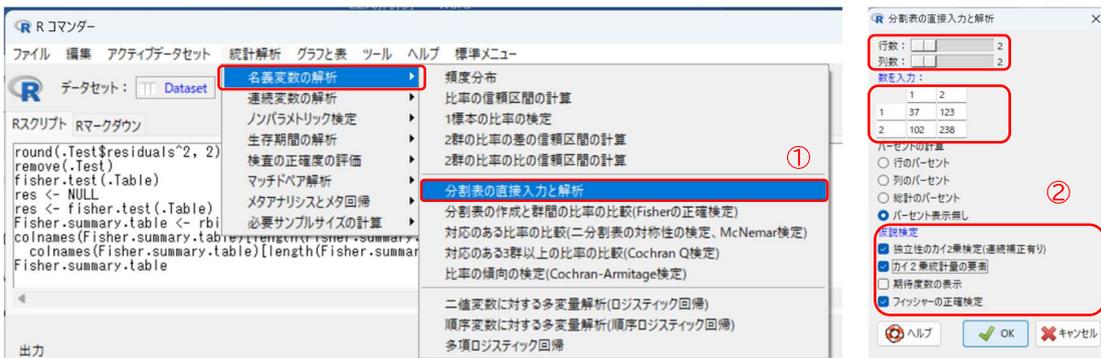
- ① [統計解析]-[名義変数の解析]-[分割表の作成と群間の比率の比較]を選択する。
- ② 行・列の変数と検定の種類を選択する。今回は「性別」と「睡眠問題」について選択し、カイ2乗検定を行う。
- ③ 結果が出力される。



■ 参考

クロス集計表が手元にある場合は、以下の方法で独立性の検定が可能である。

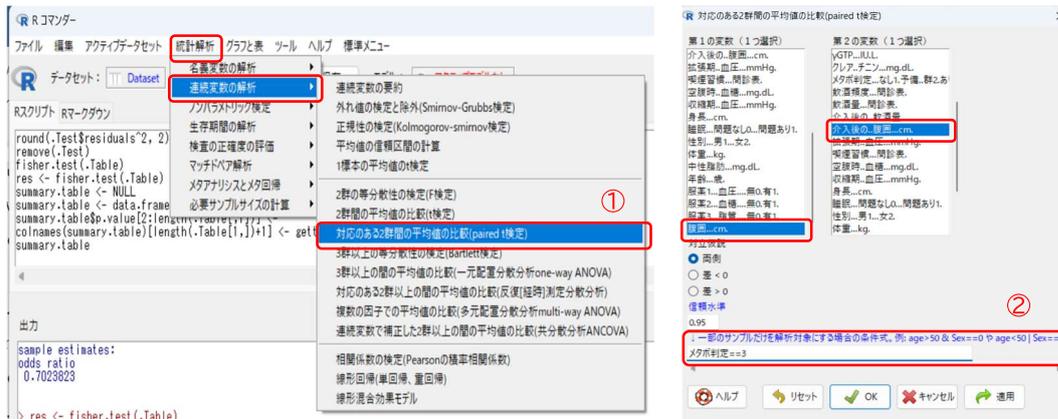
- ① [統計解析]-[名義変数の解析]-[分割表の直接入力と解析]を選択する。
- ② 行数と列数を選択し、それぞれのセルに値を入力し、検定の種類を選択する。
- ③ 結果が出力される。



② 対応のある場合の t 検定

- ① [統計解析]-[連続変数の解析]-[対応のある群間の平均値の比較 (paired t 検定)]
- ② 2 変数を選択し、検定を行う。今回は「腹囲」と「介入後の腹囲」を指定する。なお、「介入後の腹囲」は「メタボ判定」が 3 の人のみに存在する変数なので、条件式に「メタ

「ボ判定==3」と入力することで対象を絞り込んでいる。



③結果が出力される。

```
> mean2 <- mean(exercise$介入後の..腹囲...cm., na.rm=TRUE)
> sd1 <- sd(exercise$腹囲...cm., na.rm=TRUE)
> sd2 <- sd(exercise$介入後の..腹囲...cm., na.rm=TRUE)
> summary.ttest <- NULL
> summary.ttest <- data.frame(mean=c(mean1, mean2), sd=c(sd1, sd2), p.value=c(sign
> rownames(summary.ttest) <- c("腹囲...cm.", "介入後の..腹囲...cm.")
> colnames(summary.ttest) <- gettextRcmdr(colnames(summary.ttest))
> summary.ttest
              平均 標準偏差          P値
腹囲...cm.    81.55700  9.310886 0.000000253 ③
介入後の..腹囲...cm. 92.54717  6.169298
```

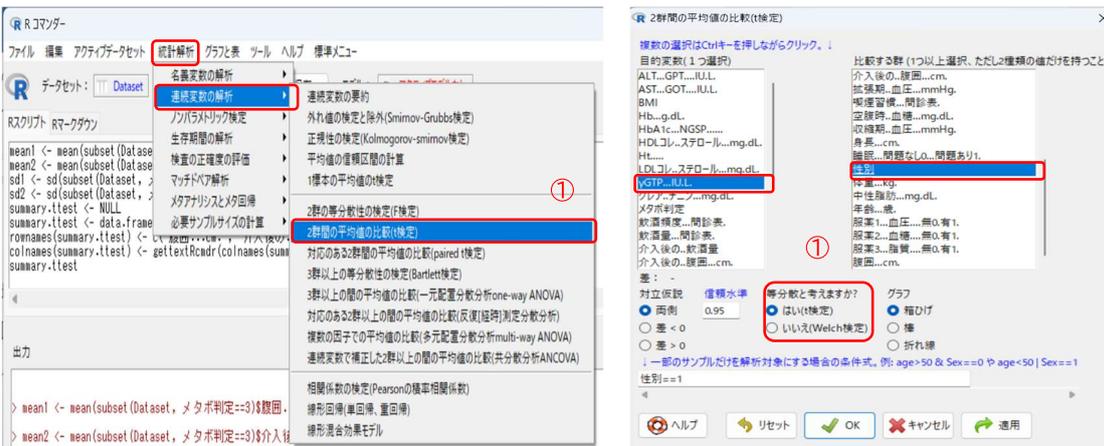
③ 等分散性の検定

- ① [統計解析]-[連続変数の解析]-[2群の等分散性の検定 (F検定)]を選択する。
- ② 目的変数と、グループを分ける変数を指定する。ここでは、「 γ -GTP」と「喫煙習慣」を指定する。
- ③ 結果が出力される。



④ 分散が等しい場合の t 検定

- ①[統計解析]-[2群間の平均値の比較 (t 検定)]を選択する。
- ②目的変数と比較する群を分ける変数を指定し、「等分散と仮定しますか?」で「はい」を選択する。今回は、変数に「 γ -GTP」と「喫煙習慣」を指定する
- ③結果が出力される。



⑤ 分散が等しくない場合の t 検定 (ウェルチの t 検定)

「④分散が等しい場合の t 検定」の、②において「等分散と仮定しますか?」を「いいえ」にすると、ウェルチの t 検定を行うことができる。