

1.2. サンプルデータの準備

看護統計支援研究所をインターネット検索すると、
 そのホームページから、サンプルデータ、
 Sample_Data_for_EZR_Handbook.xls
 がダウンロードできます。
 (本書の最後に、データのインターネットアドレスの
 QRコードも示しました。)
 (Google Chrome が不調の場合は、
 MS-Edge など、試してみてください。)

このエクセルファイルを、C:\EZADATA に入れます。

エクセルファイルを、ダブルクリックすると、
 次のスライドのような、エクセルシートが立ち上がります。
 このシートを用いて、本章の演習を進めます。

	A	B	C	D	E	F	G	H	I	J	K	L	M
1													
2		例題1	性別	血色素量		例題2	空腹時血糖	HbA1c		例題3	ジム前	3月後	
3		血液	女	13		空腹時	100	3.9		ジム	85	82	
4		色素量	女	12.9		血糖値	132	4.6		体重	73	64	
5		の	女	10.8		と	157	5.9		減少	59	57	
6		性差	女	14		HbA1c	177	6.5		?	80	61	
7		?	女	14.6		の	193	6.9			58	53	
8			女	13.4		関係	159	7.4			59	55	
9			女	13.9		?	162	7.7			82	62	
10			女	14.2			190	9.1			64	66	
11			女	15.2			266	10			59	53	
12			女	13.9			295	11.1			62	57	
13			男	15.3									
14			男	14									
15			男	14.5									
16			男	17									
17			男	15.2									
18			男	15.3									
19			男	14.6									
20			男	16.2									
21			男	17									
22			男	15.1									
23													
24													

これらのデータは、

① データ入力を行ってみたいという要望があることに配慮し、また、

② 対応のない t 検定のためのデータ構造を明示することを目的としています。

サンプル数 10 は、小さ過ぎます。

これについては、6章3節を見て下さい。

< Google Chrome による
サンプルデータの
ダウンロード(1/2)>

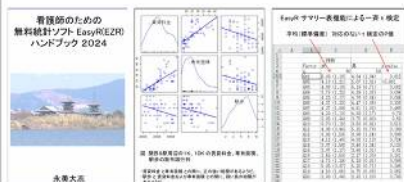
EasyR(EZR)ハンドブック
のホームページで、

EasyR(EZR)エクセル
のサンプルデータ(エクセル)
を押す

画面の右上に、
右のような表示が出る。
速やかに、保存を押す。

次ページへ、

看護師のための無料統計ソフト
EasyR(EZR)ハンドブック 2024
～インストールから、因子分析まで～



EasyR(EZR)ハンドブック カラーサンプル (インストール、と 保存、期間分付)

EasyR(EZR)ハンドブック の サンプルデータ (エクセル)

看護統計支援研究所へのお問い合わせ



< Google Chrom による
サンプルデータの
ダウンロード (2/2)>

右の図のように、
「Sample_Data_for_EZR_Handbook.xls 完了」となり、

エクスプローラで、ダウンロード・フォルダーを見ると、
この エクセル・ファイル があるので、

C:\EZADATA に移す。



1.3. 対応のない t 検定

2群の平均の間に差があるかどうかを検定する手法で、最も使用頻度の高い検定である。

EasyRでは、この検定結果を表示するのに最も妥当である、棒グラフ(標準偏差付き)が自動的に出力される。

グラフは、その中で、右ボタンを押し、ビットマップにコピーを行うと、パワーポイントなどに貼り付けることができる。

初めて見ると、変なデータ構造と感じられるかも知れないが、アンケート調査のデータシートを構築し、解析を始めてみると、妥当な構造であることが、お解りいただけると思う。

多くの統計ソフトが、この構造を採用している。

<例題.1.

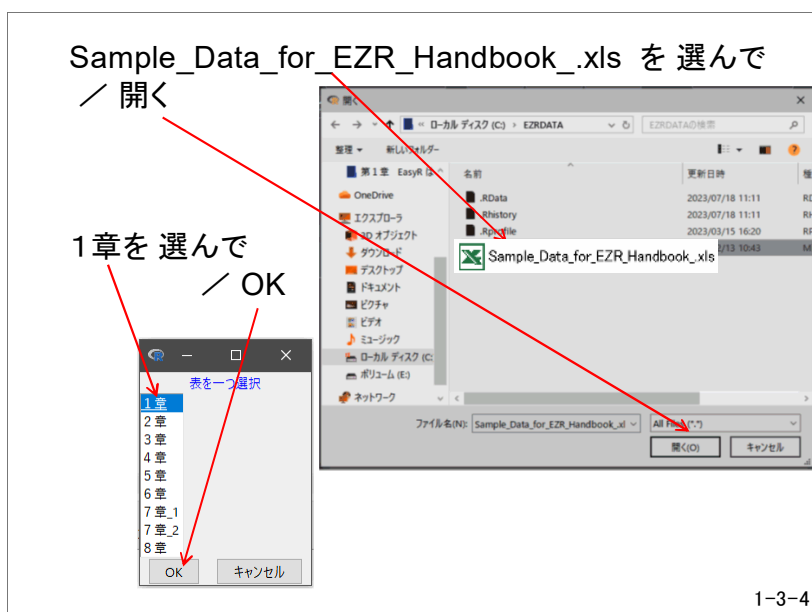
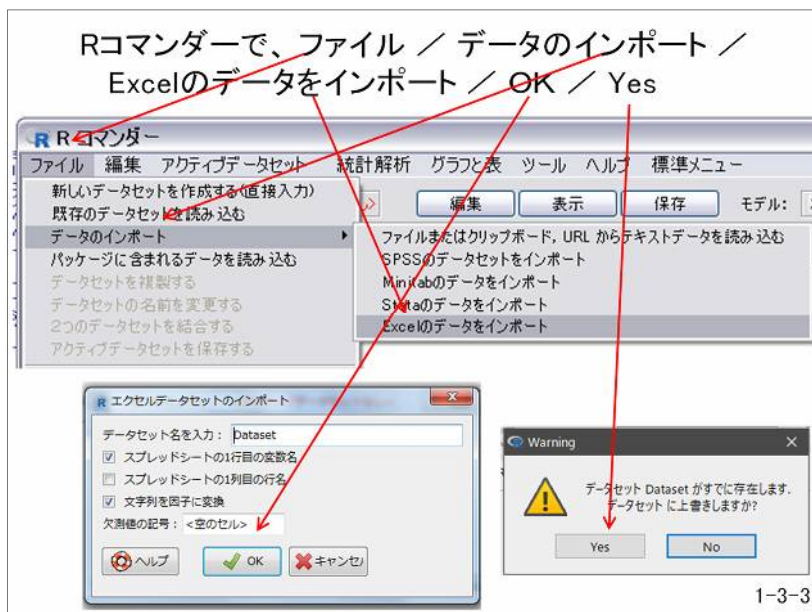
対応のない t 検定>

ある事業所の健康診査における、血液データから、男女、各10人の血色素量の数値を抽出した。

このデータでは、性差はあったといえるか、検定しよう。

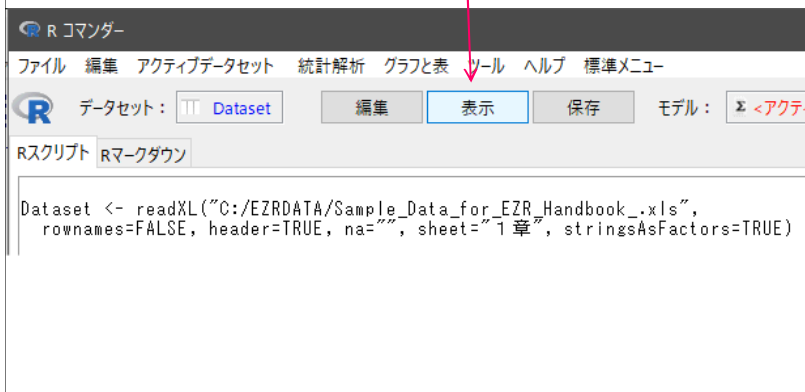
(データ入力する場合は、構造に注意。
前のページの下の、
サンプル・エクセル・ファイル、1章を
見てください。)

血色素量	
女	男
13	15.3
12.9	14
10.8	14.5
14	17
14.6	15.2
13.4	15.3
13.9	14.6
14.2	16.2
15.2	17
13.9	15.1



< 読み込んだデータの確認 (1/2) >

8ページに従って読み込んだ後、表示



< 読み込んだデータの確認(1/2) >

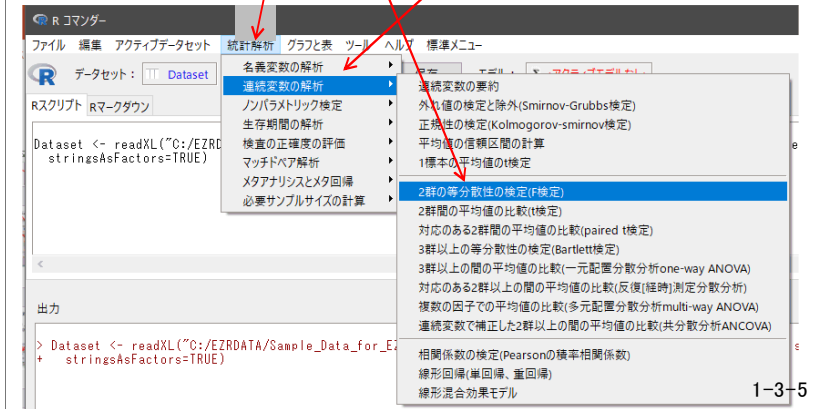
Dataset ウィンドウ が立ち上がる。

NAは、欠損値と認識され、EasyRは、無視して解析を進めてくれる。

Dataset									
例題 1	性別	血色素量	...	例題 2	空腹時血糖	HbA1c	...	例題 3	...
1	血液	女	13.0	NA	空腹時	100	3.9	NA	ジム
2	色	女	12.9	NA	血糖値	132	4.6	NA	体重
3	素量	女	10.8	NA	と	157	5.9	NA	減少
4	の	女	14.0	NA	HbA1c	177	6.5	NA	?
5	性差	女	14.6	NA	の	193	6.9	NA	<NA>
6	?	女	13.4	NA	関係	159	7.4	NA	<NA>
7	<NA>	女	13.9	NA	?	162	7.7	NA	<NA>
8	<NA>	女	14.2	NA	<NA>	190	9.1	NA	<NA>
9	<NA>	女	15.2	NA	<NA>	266	10.0	NA	<NA>
10	<NA>	女	13.9	NA	<NA>	295	11.1	NA	<NA>
11	<NA>	男	15.3	NA	<NA>	NA	NA	NA	NA
12	<NA>	男	14.0	NA	<NA>	NA	NA	NA	NA
13	<NA>	男	14.5	NA	<NA>	NA	NA	NA	NA
14	<NA>	男	17.0	NA	<NA>	NA	NA	NA	NA
15	<NA>	男	15.2	NA	<NA>	NA	NA	NA	NA
16	<NA>	男	15.3	NA	<NA>	NA	NA	NA	NA
17	<NA>	男	14.6	NA	<NA>	NA	NA	NA	NA
18	<NA>	男	16.2	NA	<NA>	NA	NA	NA	NA
19	<NA>	男	17.0	NA	<NA>	NA	NA	NA	NA
20	<NA>	男	15.1	NA	<NA>	NA	NA	NA	NA

<等分散性のF検定> (対応のない t 検定の準備)

統計解析 / 連続変数の解析 /
2群の等分散性の検定(F検定) /



血色素量 / 性別 / OK



R コマンドー の出力に

F検定 P値 = 0.628

0.628 は 0.05 より大きく
等分散仮説が棄却されなかったので、
t 検定に進む。

(P値 < 0.05 の場合は、
Welchの検定を行う。)

```
R コマンドー
ファイル 編集 アクティブデータセット 統計解析 グラフと表 ツール ヘルプ 標準メニュー
データセット: Dataset
Rスクリプト Rマークダウン

Dataset <- readXL("C:/EZDATA/Sample_Data_for_EZR_Handbook_.xls")
stringsAsFactors=TRUE)
#####2群の等分散性の検定(F検定)#####
tapply(Dataset$血色素量, Dataset$性別, var, na.rm=TRUE)
res <- NULL
(res <- var.test(血色素量 ~ 性別, alternative='two.sided', conf.
cat(gettextRcmdr( "F test"), " ", gettextRcmdr( "p.value"), "
, sep=" ")

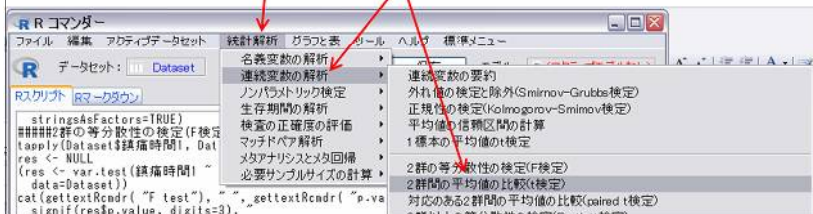
出力
F test ~ NULL
> (res <- var.test(血色素量 ~ 性別, alternative='two.sided', con
F test to compare two variances
data: 血色素量 by 性別
F = 1.3943, num df = 9, denom df = 9, p-value = 0.6285
alternative hypothesis: true ratio of variances is not equal to
95 percent confidence interval:
0.3463164 5.6133145
sample estimates:
ratio of variances
1.394268

> cat(gettextRcmdr( "F test"), " ", gettextRcmdr( "p.value"), "
, sep=" ")
F検定 P値 = 0.628
```

1-3-7

<例題.1. 対応のない t 検定>

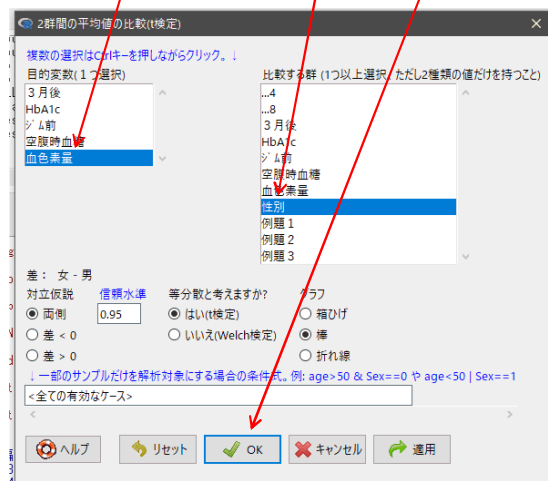
統計解析 / 連続変数の解析 /
2群間の平均値の比較(t 検定) /



1-3-8

第1章 EasyR はじめの一步

血色素量 / 性別 / OK

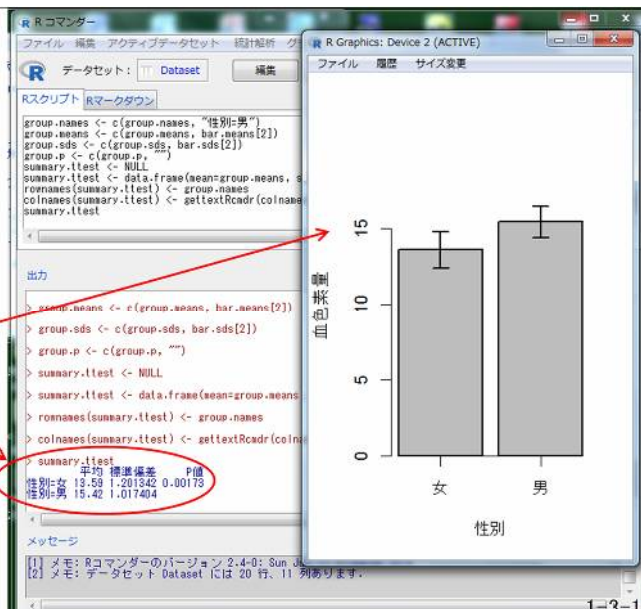


1-3-9

<例題.1.
解析
結果
例>

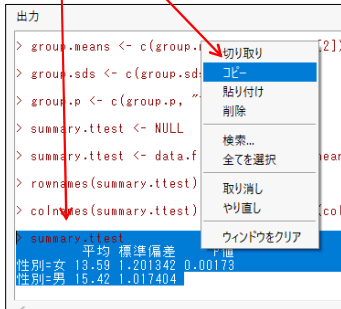
グラフ
ウィンドウ

解析
結果

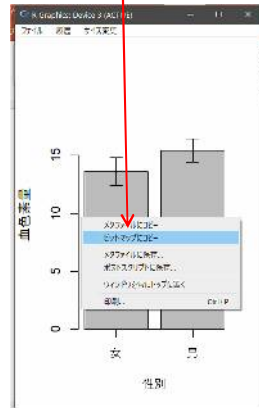


1-3-10

解析結果をドラッグ
(青くなる) / 右ボタン
コピー /
パワーポイントなどに貼付



グラフの上で、右ボタン /
ビットマップにコピー /
パワーポイントなどに貼付



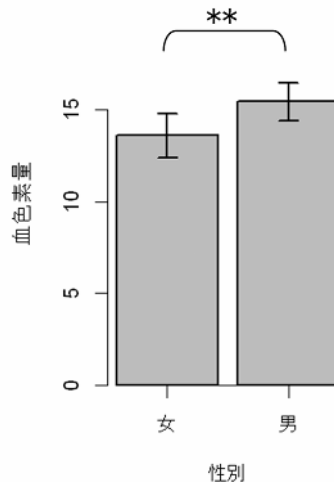
1-3-11

<例題.1. 解析結果例>

等分散の F検定 P値 = 0.628

```
> summary.ttest
      平均 標準偏差      P値
性別=女 13.59 1.201342 0.00173
性別=男 15.42 1.017404
```

血色素量は、男性方が、
1%水準で有意に高かった。



1-3-12

<対応のない t 検定 凡例>

(2群の平均の差の検定)

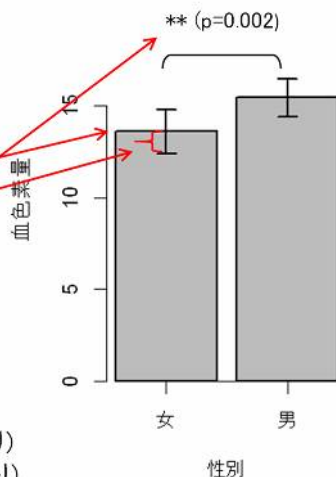
<EZR 出力>

	平均	標準偏差	P値
性別=女	13.59	1.201342	0.00173
性別=男	15.42	1.017404	

平均
標準偏差

P値:
1.0~0.05
0.05~0.01
0.01~0.001
<0.001

判定:
n.s.(not significant)
(有意差なし)
* (5%水準で有意差あり)
** (1%水準で有意差あり)
*** (0.1%水準で有意差あり)



1-3-13

1.4. 相関分析

相関分析は、2つの変量の間に、どのような関係があるか、検討する手法である。

本節の例題では、HbA1cは、空腹時血糖値の代替指標として使用できるかどうかを、検討するが、この方法が、相関分析である。

相関係数(r)が重要な指標であり、代替指標であるためには、概ね 0.9以上である必要があるだろう。

EasyR では、相関分析に最も妥当であると考えられる、散布図が自動的に生成される。

<例題.2. 相関(回帰)分析>

空腹時血糖値(血糖)は、特定健康診査(検診)の重要項目である。10時間絶食が必要であるが、受診者が守れないこともしばしばである。

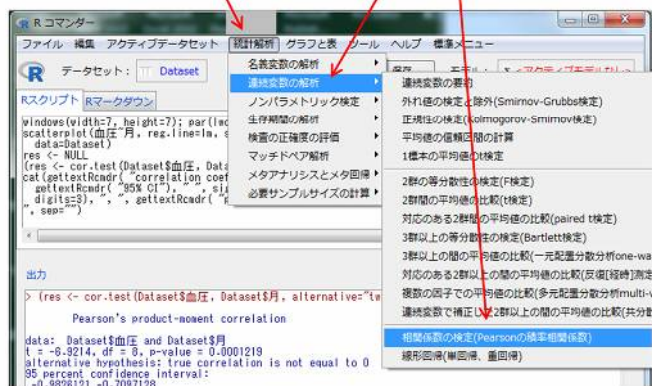
その場合、健診の評価では、ヘモグロビンA1c(HbA1c)を代替指標として使用している。

右のデータは、10人の検診受診者の、空腹時血糖値と HbA1c の測定結果である。相関(回帰)分析を試みよう。

血糖	HbA1c
100	3.9
132	4.6
157	5.9
177	6.5
193	6.9
159	7.4
162	7.7
190	9.1
266	10
295	11.1

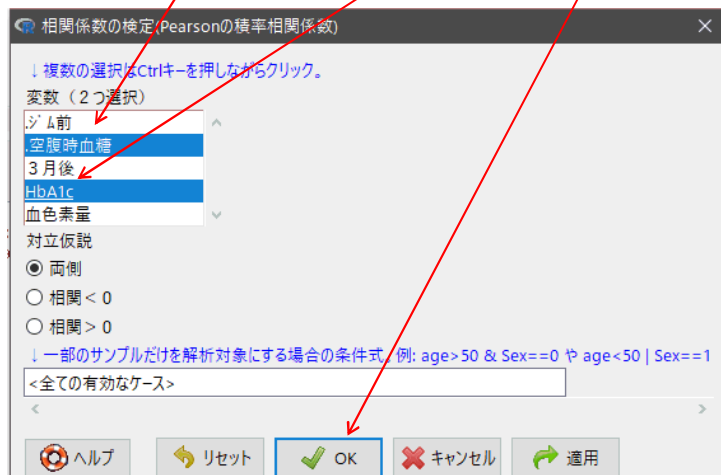
<例題 2. 相関分析>

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) /



1-4-3

空腹時血糖値 / Ctrl+ HbA1c / OK
(コントロールキーを押しながら)

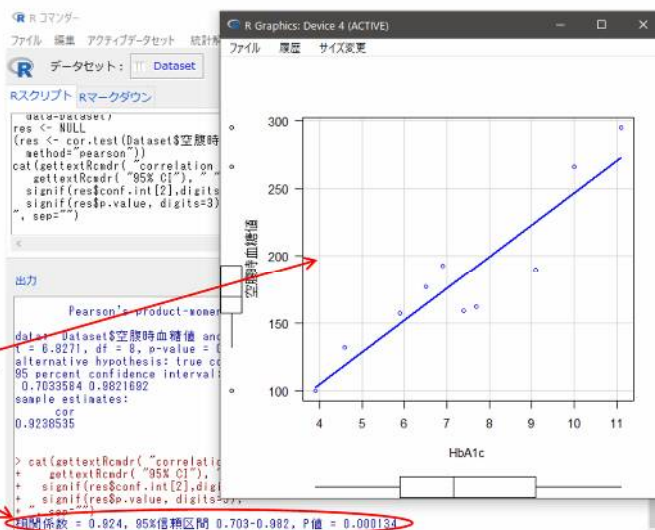


1-4-4

<例題.2. 解析 結果 例>

グラフ
ウィンドウ

解析
結果



1-4-5

<例題.2. 解析結果例>

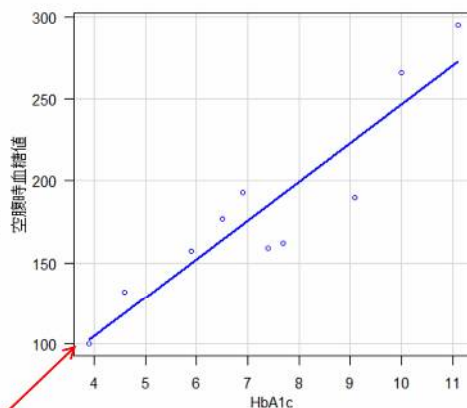
EasyRの出力:

相関係数 = 0.924,
95%信頼区間 0.703-0.982,
P値 = 0.000134

$r = 0.924$
 $p < 0.001$ ***

0.1%水準で有意で、
強い正の相関があった。

図の中の青い実線は、
回帰直線である。



1-4-6

1. 5. 対応のある t 検定

3節では、血色素量に男女差があるかを検討したが、本節では、ジムに通って行う運動により、体重減少したかどうかを、検討する。

同じ人について、一定の期間を置いて、前後の測定を行って比較するので、「繰り返し測定の t 検定」とも言われる。

この解析に適切なグラフは、折れ線グラフであるが、妥当なデータ数(概ね30以上)になると判読が難しくなるので、EZRでは自動出力されない。

<例題.3.

対応のある t 検定>

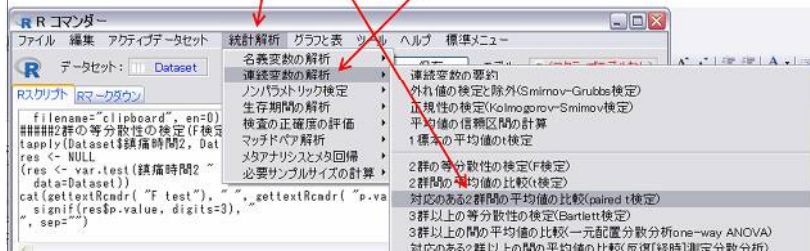
中年男性10人が、一念発起し、ジムのトレーニングにより、減量を試みた。ジムに行く前と、3ヶ月後の体重を測定したところ、右記のようであった。

ジムのトレーニングは有効であったといえるか？検定してみよう。

ジム前	3月後
65	62
73	64
59	57
60	61
58	53
59	55
62	62
64	66
59	53
62	57

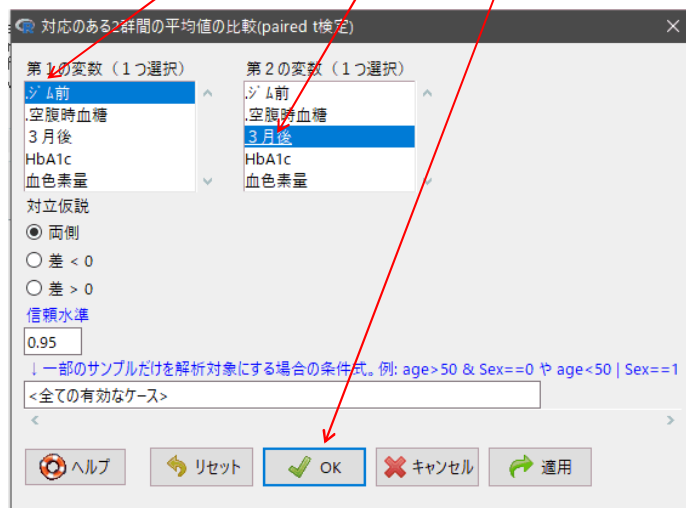
<例題 3. 対応のある t 検定>

統計解析 / 連続変数の解析
対応のある2群間の平均値の比較 /



1-5-3

ジム前 / 3ヶ月後 / OK



1-5-4

<例題.3.
解析
結果
例>

解析
結果

```

R コマンドター
ファイル 編集 アクティブデータセット 統計解析 グラフと表 ツール ヘルプ 標準メニュー
データセット: Dataset 編集 表示 保存 モデル: <アクティブモデルなし>
Rスクリプト Rマークダウン
mean2 <- mean(Dataset$3月後, na.rm=TRUE)
sd1 <- sd(Dataset$ジム前, na.rm=TRUE)
sd2 <- sd(Dataset$3月後, na.rm=TRUE)
summary.ttest <- data.frame(mean=c(mean1, mean2), sd=c(sd1, sd2), p.value=c(signif(res$p.value,
digit=3), ""))
rownames(summary.ttest) <- c("ジム前", "3月後")
colnames(summary.ttest) <- gettextRcmdr(colnames(summary.ttest))
summary.ttest
出力
> mean2 <- mean(Dataset$3月後, na.rm=TRUE)
> sd1 <- sd(Dataset$ジム前, na.rm=TRUE)
> sd2 <- sd(Dataset$3月後, na.rm=TRUE)
> summary.ttest <- NULL
> summary.ttest <- data.frame(mean=c(mean1, mean2), sd=c(sd1, sd2), p.value=c(signif(res$p.value,
digit=3), ""))
> rownames(summary.ttest) <- c("ジム前", "3月後")
> colnames(summary.ttest) <- gettextRcmdr(colnames(summary.ttest))
summary.ttest
平均 標準偏差 P値
ジム前 62.1 4.483302 0.0184
3月後 59.0 4.618802
メッセージ
[1] メモ: Rコマンドターのバージョンは 2.4-0. Fri Jun 19 09:52:10 2016
[2] エラー: 表が選択されていません
[3] メモ: データセット Dataset には 20 行、11 列あります。
1-5-5

```

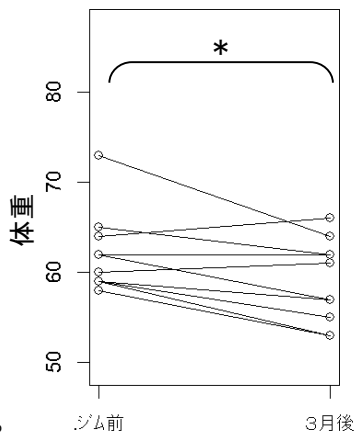
<例題.3. 解析結果例>

```
> summary.ttest
```

	平均	標準偏差	P値
ジム前	62.1	4.483302	0.0184
3月後	59.0	4.618802	

5%水準で有意であった。

3ヶ月間のジムにおける
トレーニングは、有効であった。



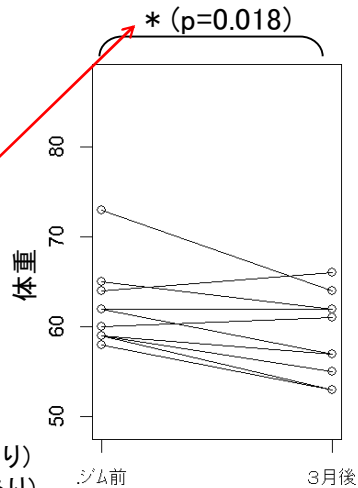
<対応のある t 検定 凡例> (繰返し測定 の 差 の 検定)

<EZR 出力>

	平均	標準偏差	P値
ジム前	62.1	4.483302	0.0184
3月後	59.0	4.618802	

折れ線グラフが妥当であるが、
データ数が多くなると判読不能
になるので、
EZRでは自動出力されない。

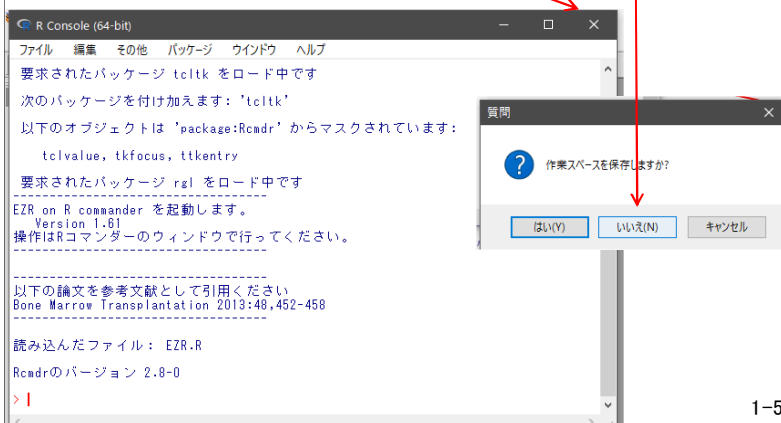
P値:	判定:
1.0~0.05	n.s.(not significant) (有意差なし)
0.05~0.01	* (5%水準で有意差あり)
0.01~0.001	** (1%水準で有意差あり)
<0.001	*** (0.1%水準で有意差あり)



1-5-7

<EasyR の終了方法(最も手順が少ない)>

R Consoleで、× を押す / いいえ



1-5-8