

論文作成に役立つ 統計解析のいろは

矢原耕史 (久留米大学バイオ統計センター)

鈴木里和 (感染研・細菌第二部)

宮本真理 (キアゲンバイオインフォマティクス)

配付資料の一部
御机上

WS企画の概要

- 細菌学会からの依頼による開催
- 目標
 - ➡ 細菌学の先生方の論文作成に直接役立つよう
 - ➡ 細菌学の先生方から事前に頂いた「生の声」に答えるよう
- 発表者の特徴
 - ➡ 矢原：生物医学統計の学位と教育研究経験、
疫学からゲノムまで幅広いデータの統計解析の経験
 - ➡ 鈴木：感染症専門医、UC Berkley公衆衛生大学院に留学
 - ➡ 宮本：統計科学の学位、インフォマティクス

細菌学の先生方から事前に頂いた「生の声」：例1

- 今まで自分流で統計解析をしてきたけれど、それが正しいかどうか分からない
 - ➡ 2群比較の際、順位和検定(U検定)ばかりを使っているが？
 - なぜt検定ではいけない？
 - ➡ 比較する群の数が増えた場合、どうすべき？
 - ➡ ...
- 論文に「有意」と書いてあるが、疑わしい場合がある
 - ➡ 正しいかどうか、どうすれば判断できる？

細菌学の先生方から事前に頂いた「生の声」: 例2

- ありがちな「やってはいけない」ラボデータの統計処理なんかを説明すると有意義かも
 - ➡ 3群間の比較を普通にt検定繰り返し
 - ➡ 3倍希釈で測定している抗体価の平均値を出してt検定
 - ➡ 標準誤差と標準偏差、95%信頼区間の記述が混同
 - ➡ ...

細菌学の先生方から事前に頂いた「生の声」：例3

● 細菌学分野に特有のデータをどう統計解析すべき？

➡ CFU

- 生の値は非常に大きい

➡ 薬剤感受性試験によるMIC

- 濃度1, 2, 4, 8と倍々の濃度、もしくは0.5, 0.25といった1/2ずつに希釈して測定

➡ 同様に、抗体価

- 10倍希釈や2倍希釈などの段階的な希釈系で測定

● 遺伝子発現の群間比較で統計が必須なので、詳しく知りたい

発表タイトル・概要

- 統計学への準備体操 (15分)

鈴木

- 統計解析の基礎と実践的入門(50分)

-マウス操作型ソフトを使いながら-

- 生物医学統計
- 「わかった」「これは使える」と思ってもらえるように

矢原

- QA、討論1 (10分)

- 遺伝子発現の群間比較の統計解析 (20-30分)

- RNA-Seqデータを念頭に

宮本

- 総合QA、討論2 (25-15分)

統計解析の基礎と実践的入門

-マウス操作型ソフトを使いながら-

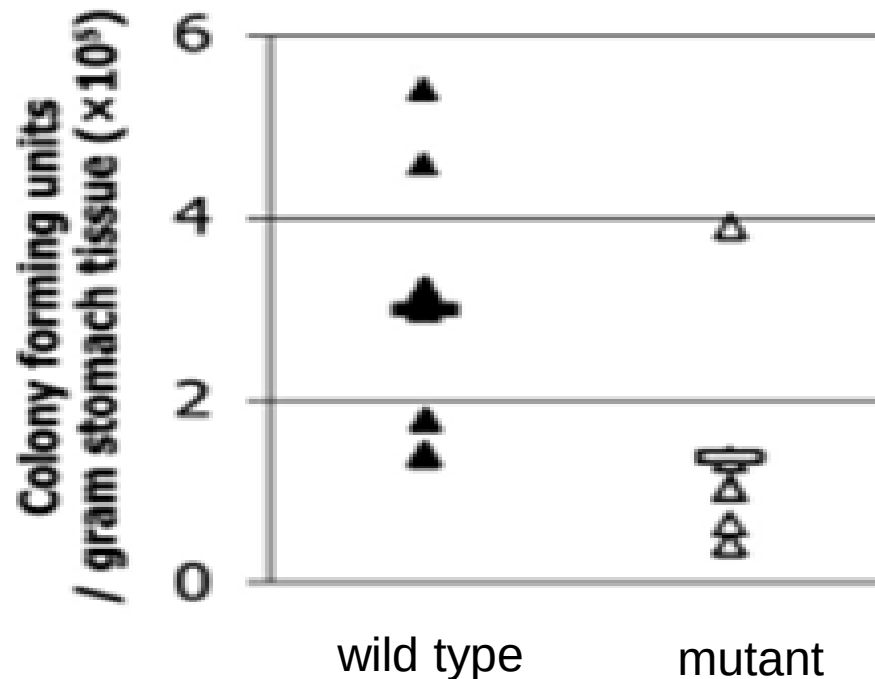
矢原耕史

(久留米大学バイオ統計センター)

koji_yahara@med.kurume-u.ac.jp

ラボデータの統計解析に焦点を当てて

- 多くの場合、やりたいことは「群間の比較」
 - ➡ 各群の測定値の平均に、有意な差があるかどうか？
 - ○○検定によるp値で、判断

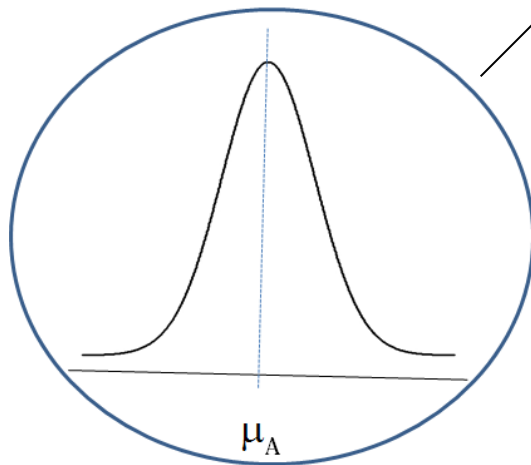


日本細菌学会誌の
論文より改変

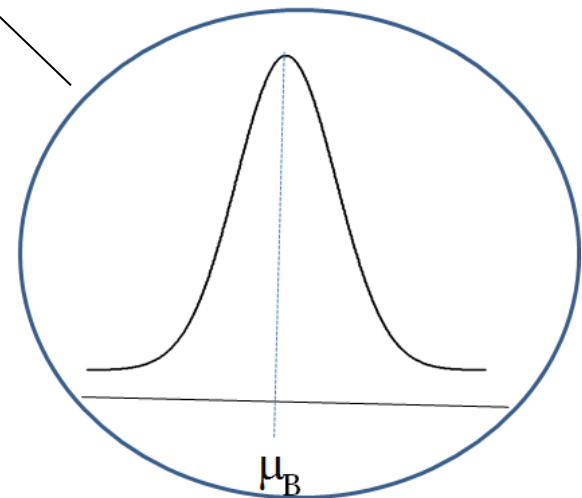
- ➡ 2群の比較なら、t検定か、ウィルコクソンの順位和検定(U検定)

検定の考え方

- 測定したデータは、各群からサンプリングされたもの(「標本」)
 - ➡ 各群の「母集団」の一部



母集団:A群



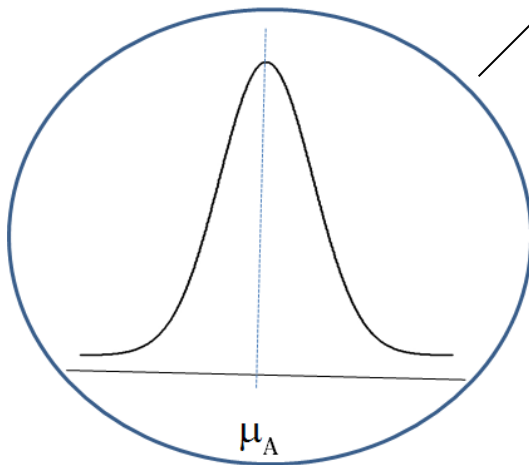
母集団:B群

検定の考え方

- 知りたいのは、測定したデータに基づいて、母集団の平均に差がある、と言えるかどうか



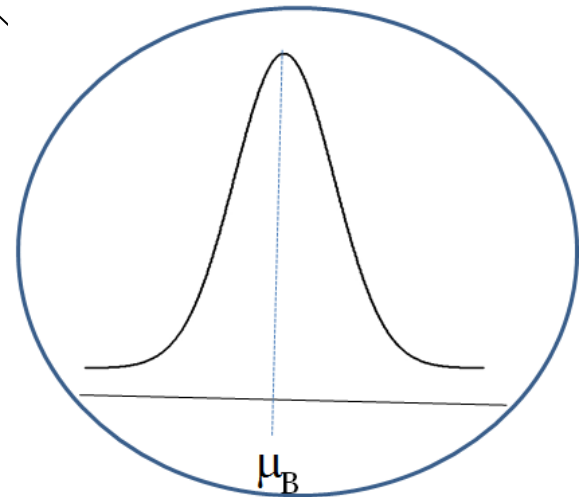
↑ ↓ 測定したデータから
計算した平均の差
(測定する毎にバラつき、変わる)



母集団:A群



母集団の平均に
差があるか？



母集団:B群

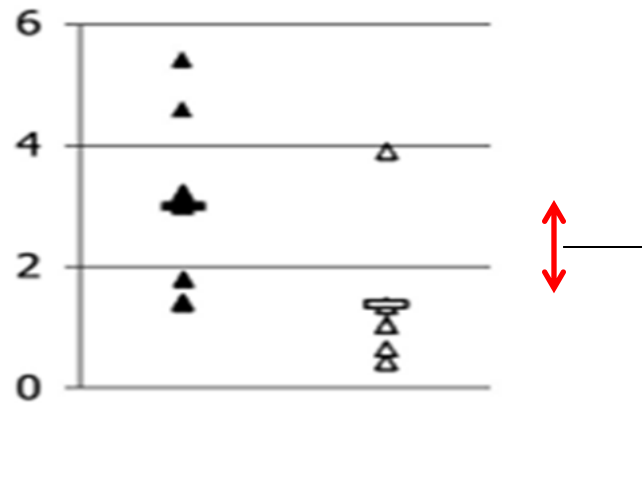
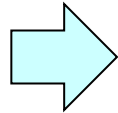
p値の考え方

- 測定したデータから計算した平均の差が、
「本当は母集団の平均には差がない」と考えた場合に
どれくらいの確率で生じるか(どれだけ起こりにくいか)

もし

$$\mu_A = \mu_B$$

とすると



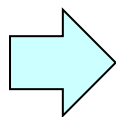
- ・測定したデータからこの差が生じる確率は？
- ・この差はどれだけ生じにくい？
- ・ $\mu_A = \mu_B$ と考えること(帰無仮説)にどれだけ無理がある？

p値の考え方

帰無仮説

$$\mu_A = \mu_B$$

が正しいとすると



測定したデータから
この差が生じる
確率は？

● p値が小さい

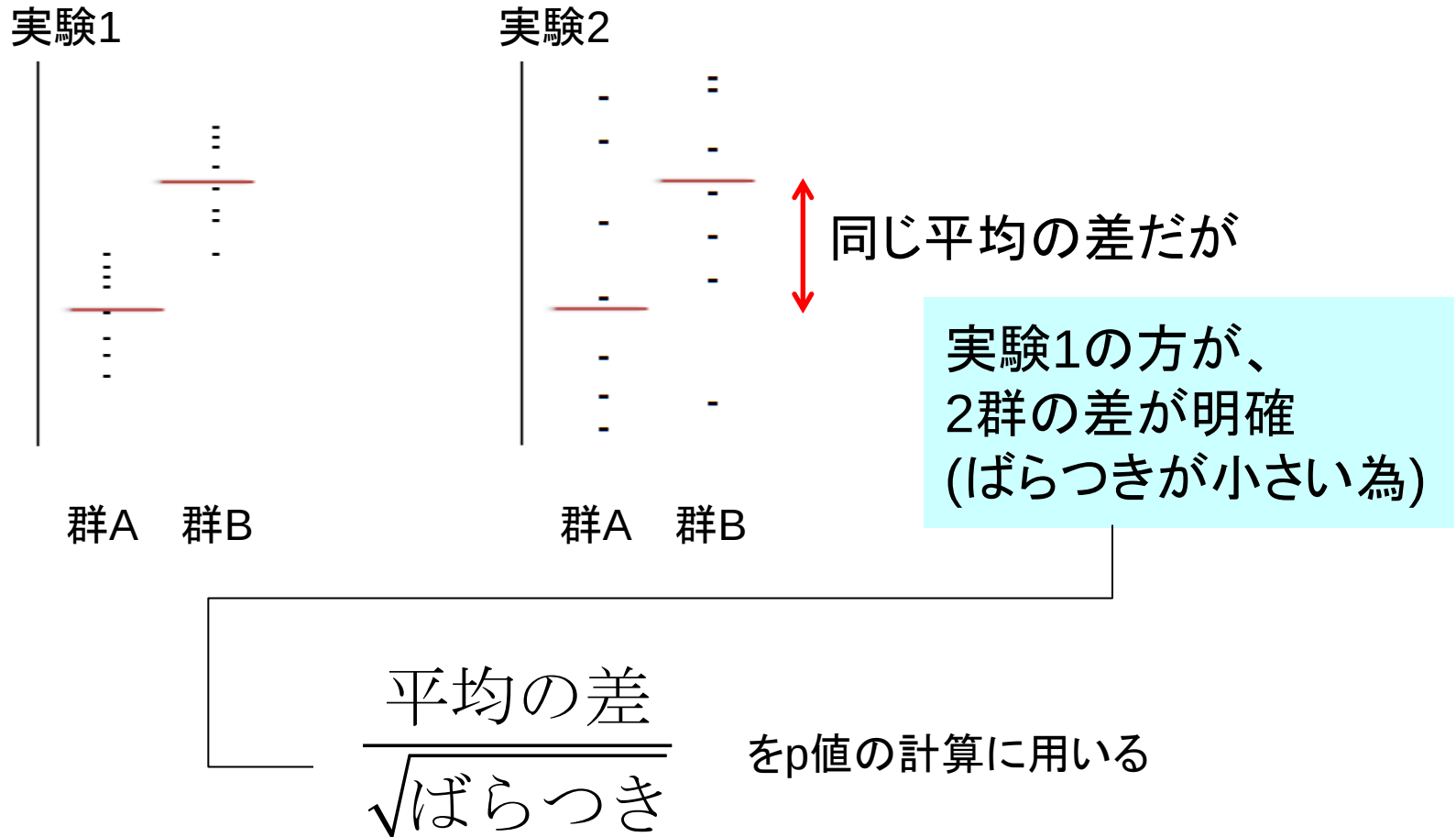
- ➡ 帰無仮説が正しいとすると、こんな大きな差が生じる確率は低い
- ➡ よって「母集団の平均には差がある」(対立仮説)の方が妥当
 - $p\text{値} < 0.05 \rightarrow$ 「有意水準5%で有意差がある」

● p値が小さくない

- ➡ 帰無仮説が正しいとしても、この程度の差は生じる (p値の確率で)

p値を計算するためには

- 平均の差だけでなく、ばらつきも考慮する必要がある



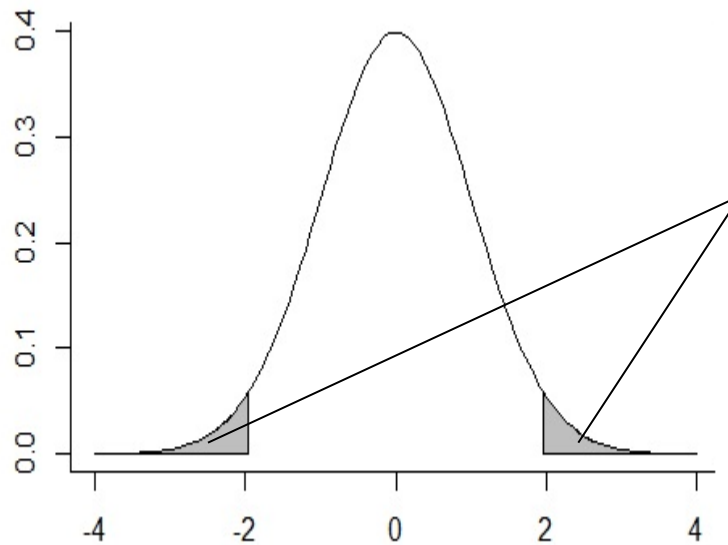
これを「**検定統計量**」: その値が大きいほど、p値は小さくなる

p値の計算法

- データから計算された検定統計量 $\frac{\text{平均の差}}{\sqrt{\text{ばらつき}}}$ の値が、

帰無仮説が正しいと仮定した場合に生じる確率 = p値

- 帰無仮説が正しいと仮定した場合の検定統計量の分布がわかれば、計算できる

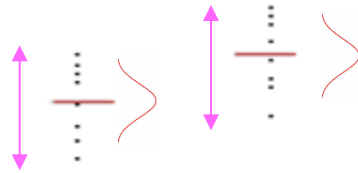


例)
検定統計量の値が1.96
→ それが標準正規分布
するなら、 $p=0.05$

検定統計量の値

t検定 (2群の分散分析)

- 仮定: 各群のデータが独立に正規分布しており、分散が等しい



- この仮定のもとで、帰無仮説が正しい場合に、検定統計量

$\frac{\text{平均の差}}{\sqrt{\text{ばらつき}}}$ の分布はt分布であることが、理論的に証明済

- t分布を使って、p値を計算する
 - ➡ t分布は正規分布と似ているが、上の仮定が成り立つなら、
 - 正規分布よりも正確にp値を計算できる
 - 測定したデータ数が少なくてもOK

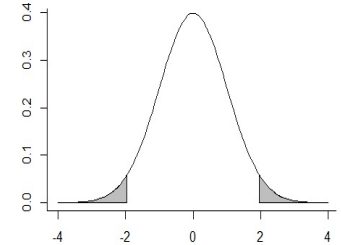
検定の仮定が成り立たない場合は？

- 仮定を無視して検定を行うと、p値がおかしくなる

- p値とは、

- ➡ 検定の仮定が成り立ち、
- ➡ 帰無仮説が正しいとした場合の検定統計量

$\frac{\text{平均の差}}{\sqrt{\text{ばらつき}}}$ の理論的な分布をもとに、計算される



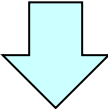
- t検定の仮定が成り立たないデータに対し、t分布を使ってp値を計算するのは間違い。

- ➡ 小さなp値 → 帰無仮説に無理がある、とは言い切れない

t検定が使えない場合、何を使えばよい？

● ウィルコクソンの順位和検定 (U検定)

- ➡ 測定した1つ1つの値を直接使うのではなく、それを全体での順位に変換した上で、群間の比較を行う

測定値	A群	1.3	3.5	6.2	8.3	13.2	15.6		
	B群		4.1	7.1	10.5		16.0	18.2	40.5
									
順位	A群	1	2	4	6	8	9		
	B群		3	5	7		10	11	12

A群での順位和と、B群での順位和に、差はある？

t検定が使えない場合、何を使えばよい？

● ウィルコクソンの順位和検定 (U検定)

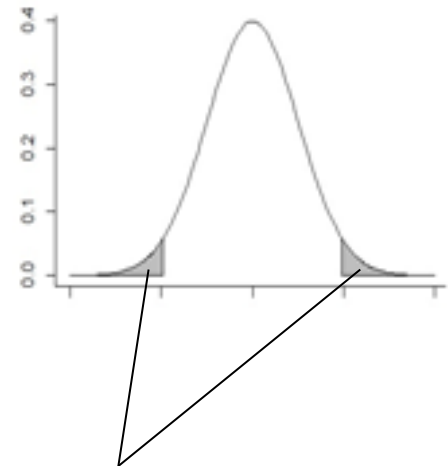
➡ データの分布に特定の形 (例えば正規分布) を仮定しない

- 「ノンパラメトリック」法

- t検定は、データが正規分布すると仮定した「パラメトリック法」

➡ 帰無仮説: どの群でも、個々のデータは、1番目から最後 (N番目) までの順位を、等確率でランダムにとる

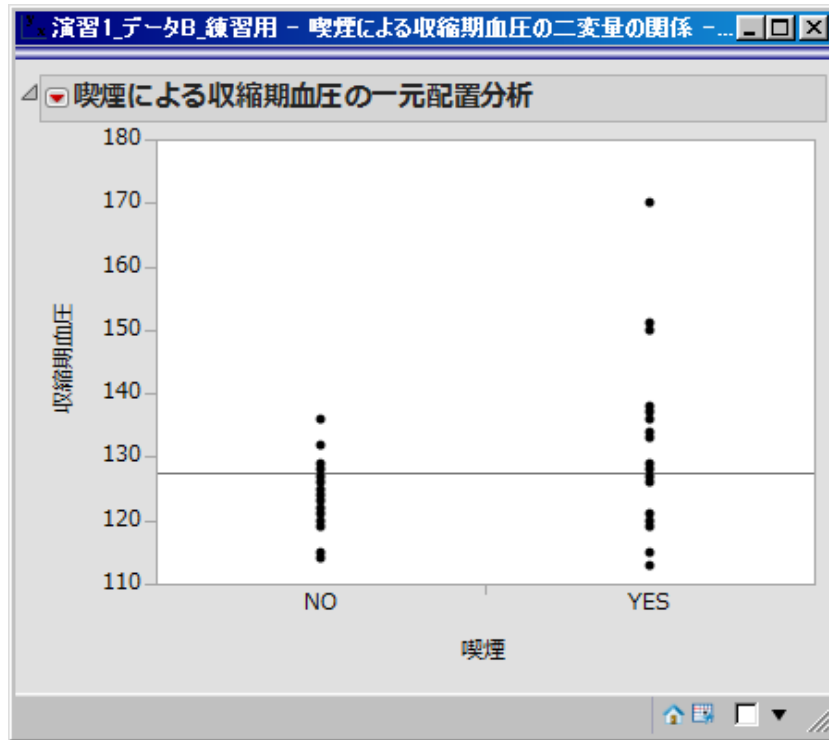
- 帰無仮説が正しいとした場合に、
A群の順位和がとり得る値の分布は
簡単なシミュレーションで求められる



- その分布を使って、実際のデータから計算されたA群の順位和が得られる確率 = p値

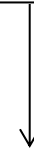
間違いの例

- 喫煙者・非喫煙者25名ずつの収縮期血圧を測定したデータ



平均値の差は
 $130.92 - 123.92 = 7$

この差は有意だと言えるか？

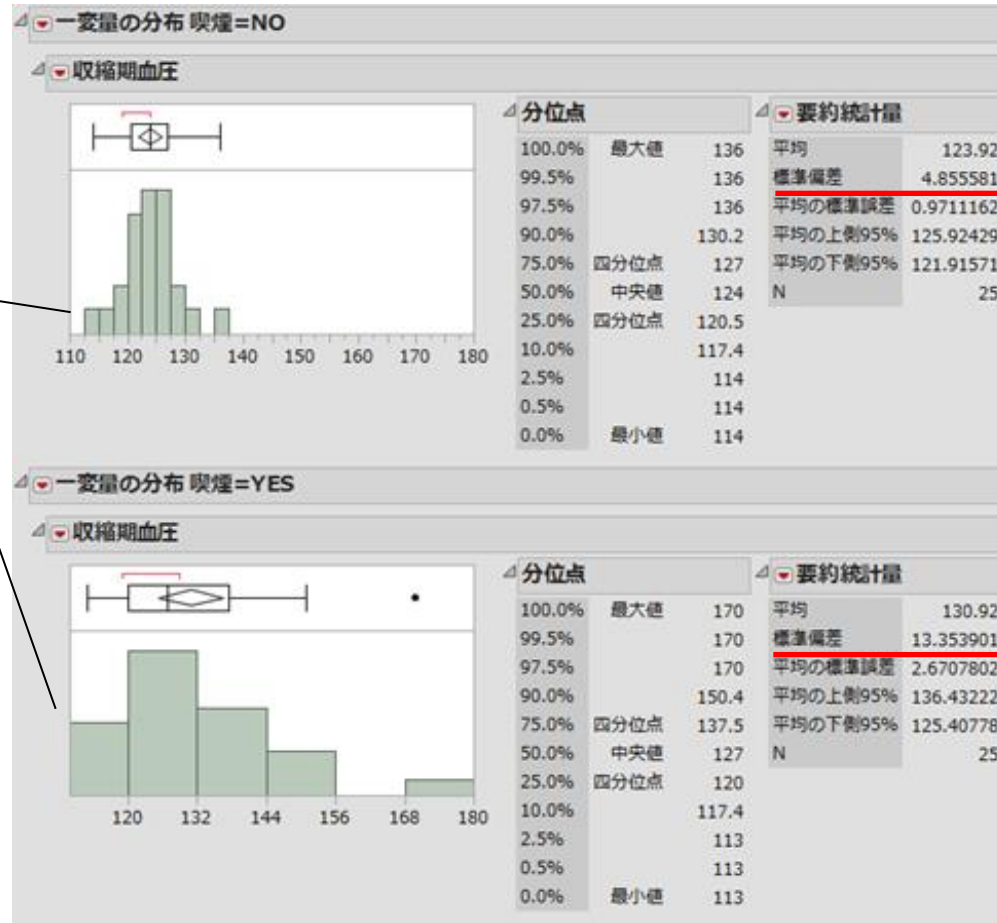
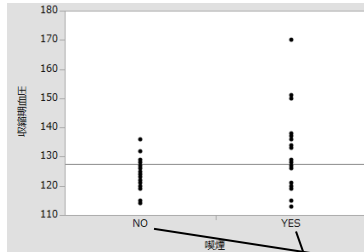


t検定: $p=0.017$

ウィルコクソン順位和検定: $p=0.056$

t検定の仮定に反している

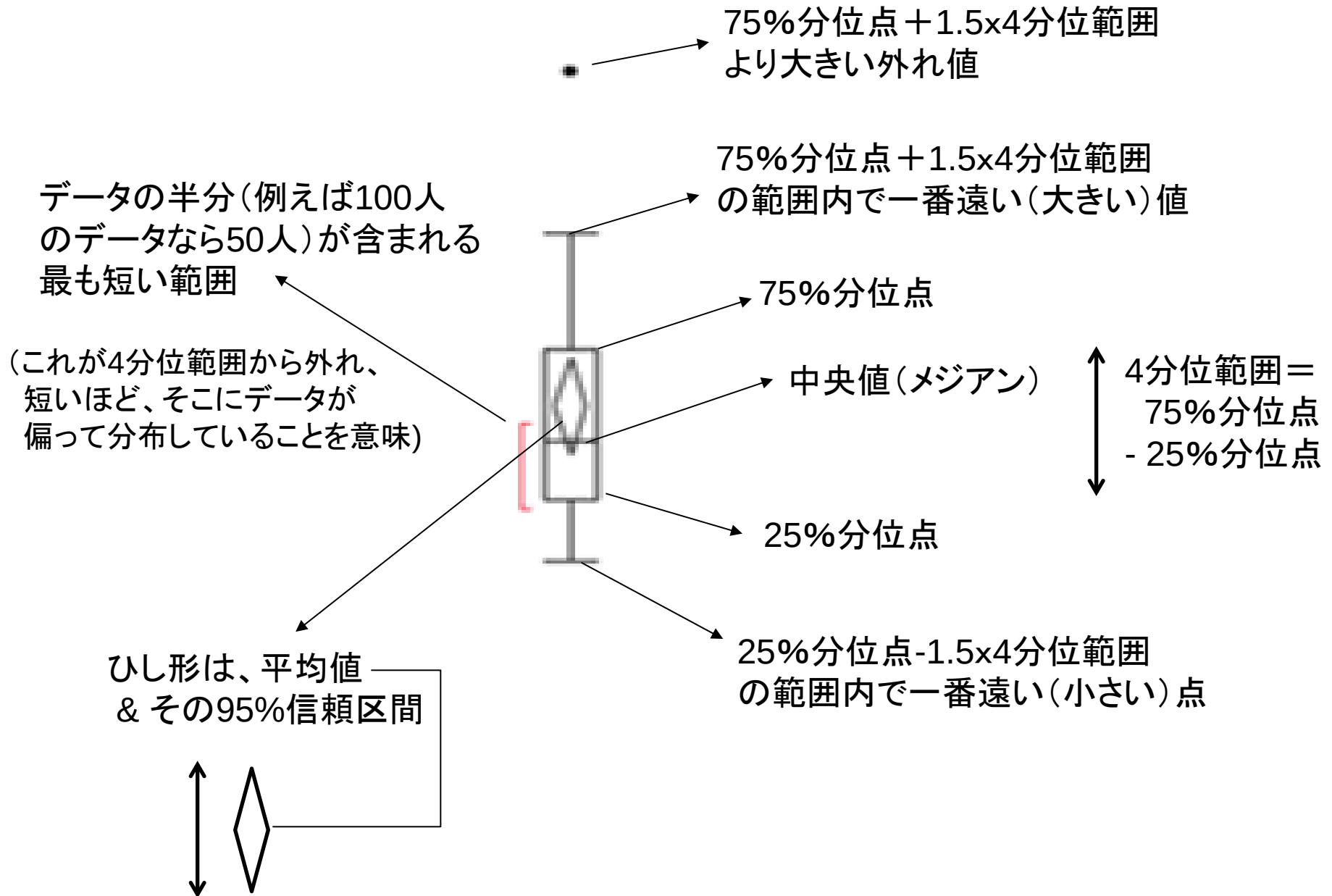
- 各群でデータは正規分布？→No(下の喫煙群の分布)
- 各群の分散は等しい？→No



マウス操作型ソフト
JMP(後述)による。
箱ひげ図の見方は
次スライドに。

この場合、t検定を使って有意差を主張するのは、間違い

(補足) データの分布を吟味するには箱ひげ図が便利



それでは、いつも順位和検定を使えばよい？

- しかし、t検定の仮定を満たすデータなら、t検定を使った方が有意差を示すのに必要なサンプル数は少なくて済む
- 順位和検定の場合
 - ➡ $p < 0.05$ を示すための最小サンプル数は、各群で4
 - 4回ともA群の方が大きい(小さい)なら、 $p = 0.029$
 - ➡ サンプル数が各群で3だと
 - 3回ともA群の方が大きく(小さく)ても、 $p = 0.1$

データの分布が分からない場合、ノンパラメトリック法で検定
→ただし、有意差を示すために、より多くのサンプルが必要

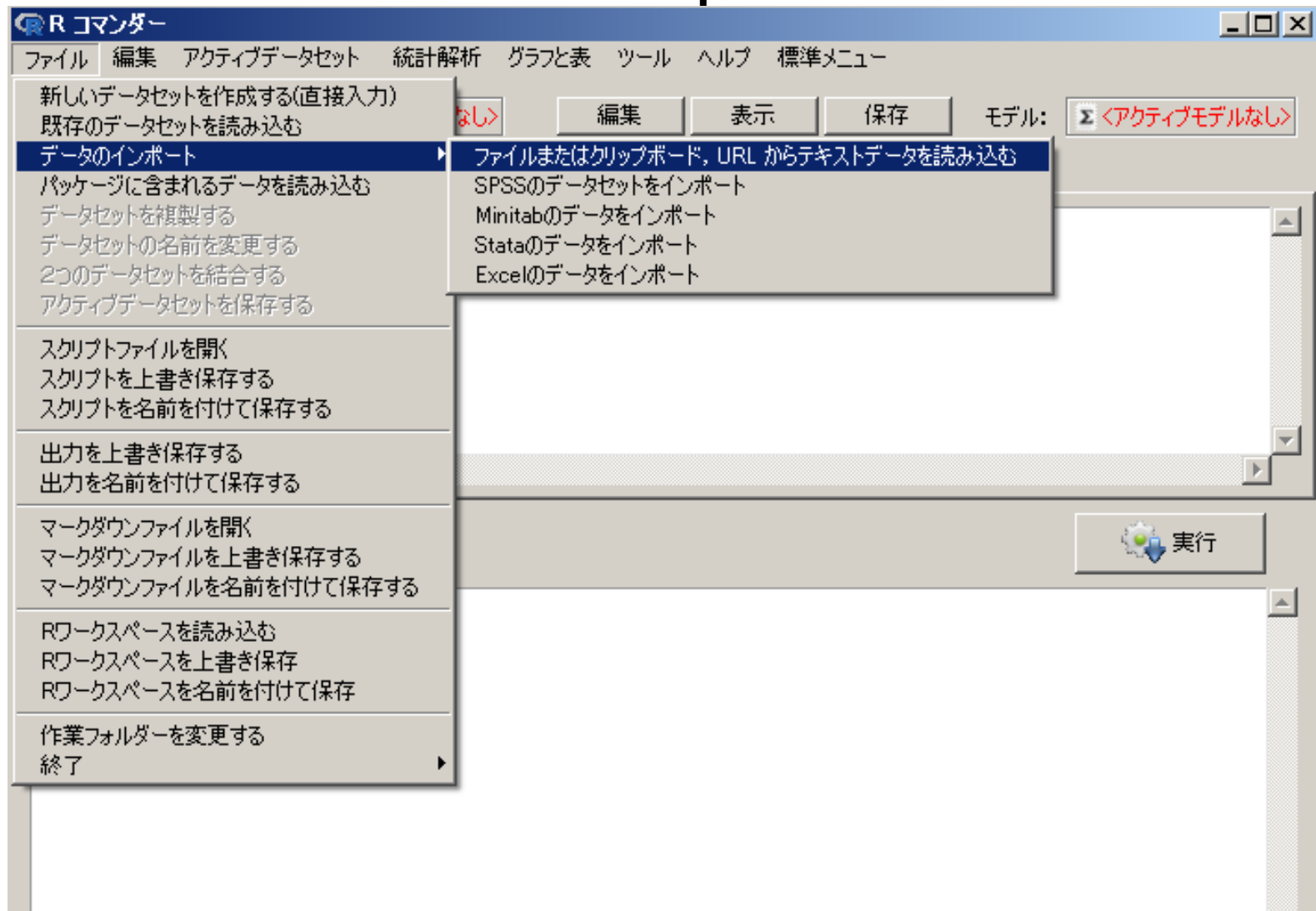
無償マウス操作型ソフトで、これまでの解析を実行

● EZR (Easy R)

- ➡ フリーの統計解析プログラミングソフトRをマウス操作で使えるようにしたもの



EZRを起動→「ファイル」メニュー→「データのインポート」



※Excelファイルを開こうとすると上手くいかないことがあるので
csv形式で別名保存して読み込むのがよい。

データファイルの形式の指定

● 区切り記号

➡ デフォルトは「カンマ」

● csvファイルを読み込むなら、この画面では何も変更不要

ファイルまたはクリップボード、URL からテキストデータを読み込む

データセット名を入力:

ファイル内に変数名あり: ☒

列数があわない場合に調整する: ☒

文字列の場合にも空欄は欠損値(NA)として読み込む: ☒

空欄以外に欠損値として読み込むべき記号:

データファイルの場所

☒ ローカルファイルシステム

☐ クリップボード

☐ インターネットの URL

フィールドの区切り記号

☐ 空白

☒ カンマ

☐ タブ

その他 ☐ 指定:

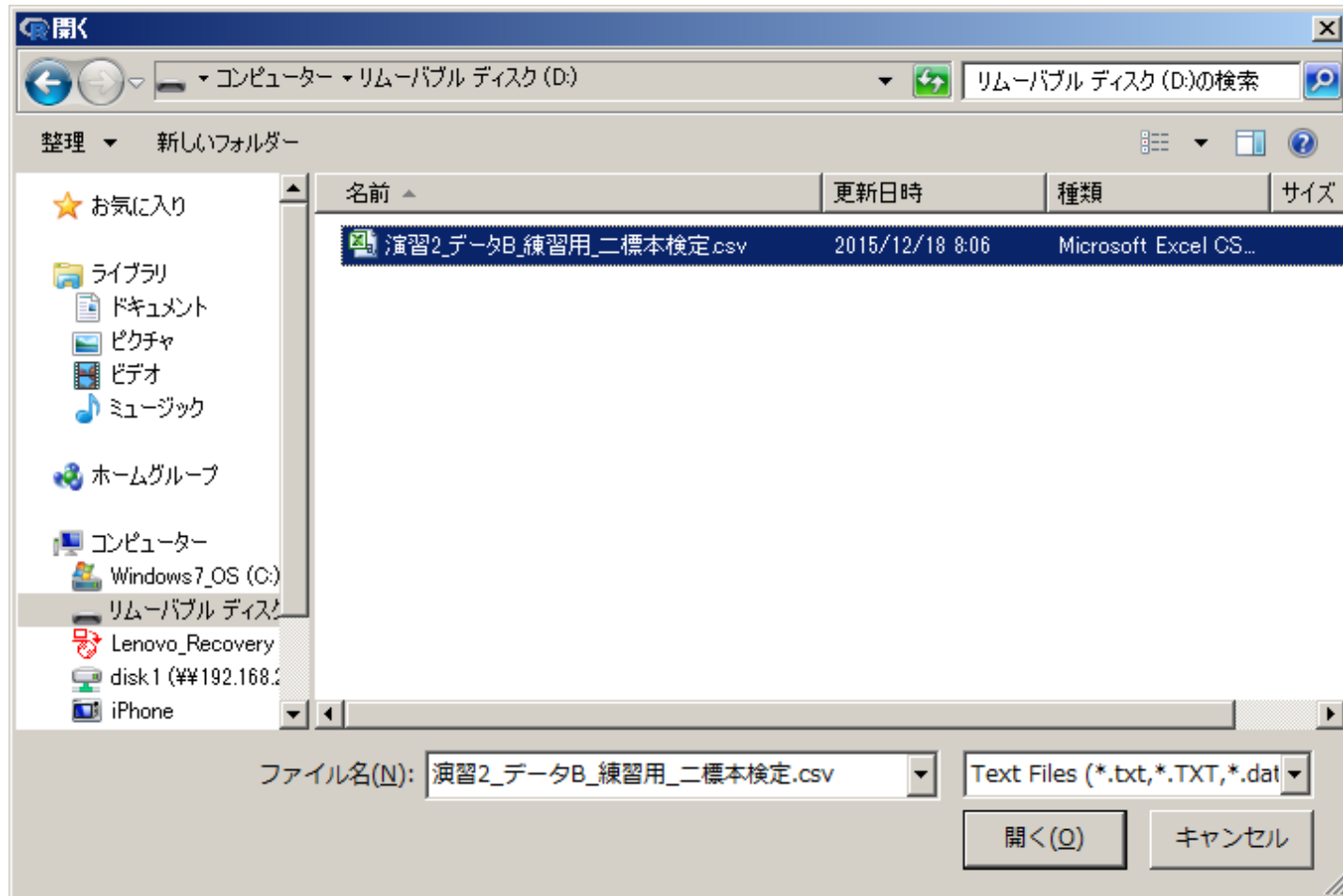
小数点の記号

☒ ピリオド[]

☐ カンマ[]

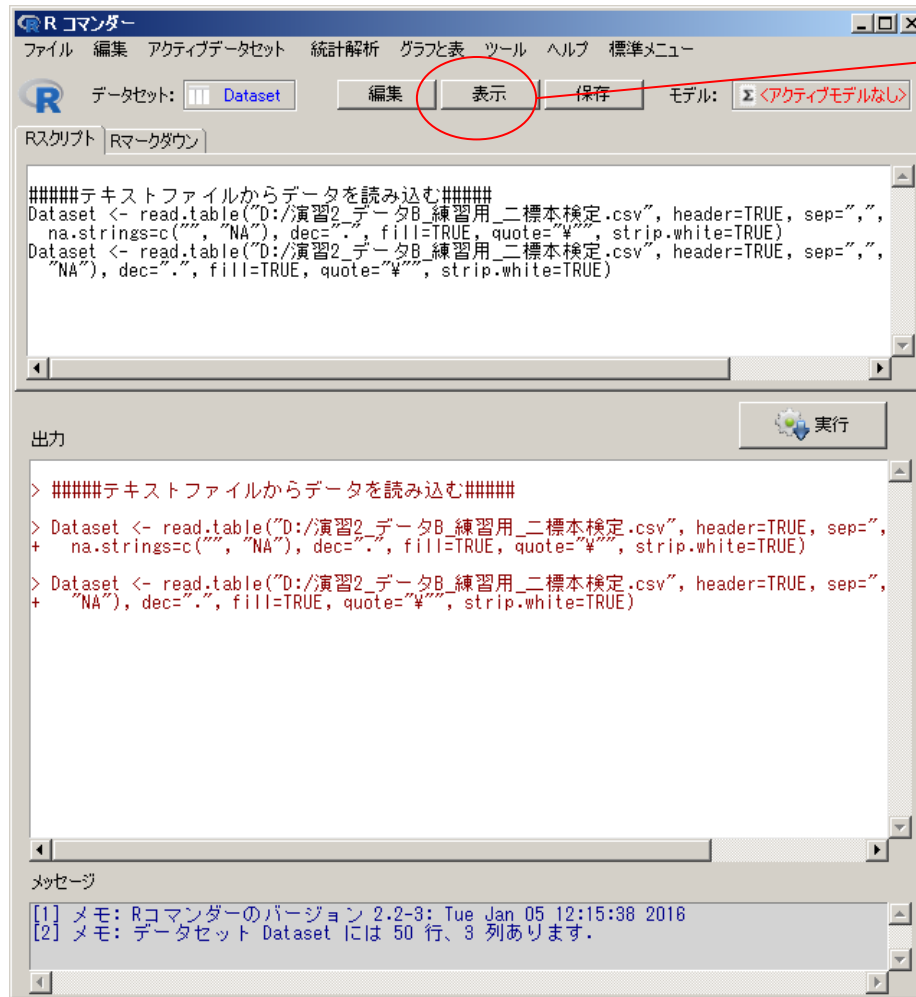
ヘルプ OK キャンセル

読み込むデータファイルの選択



データファイルを読み込めたら

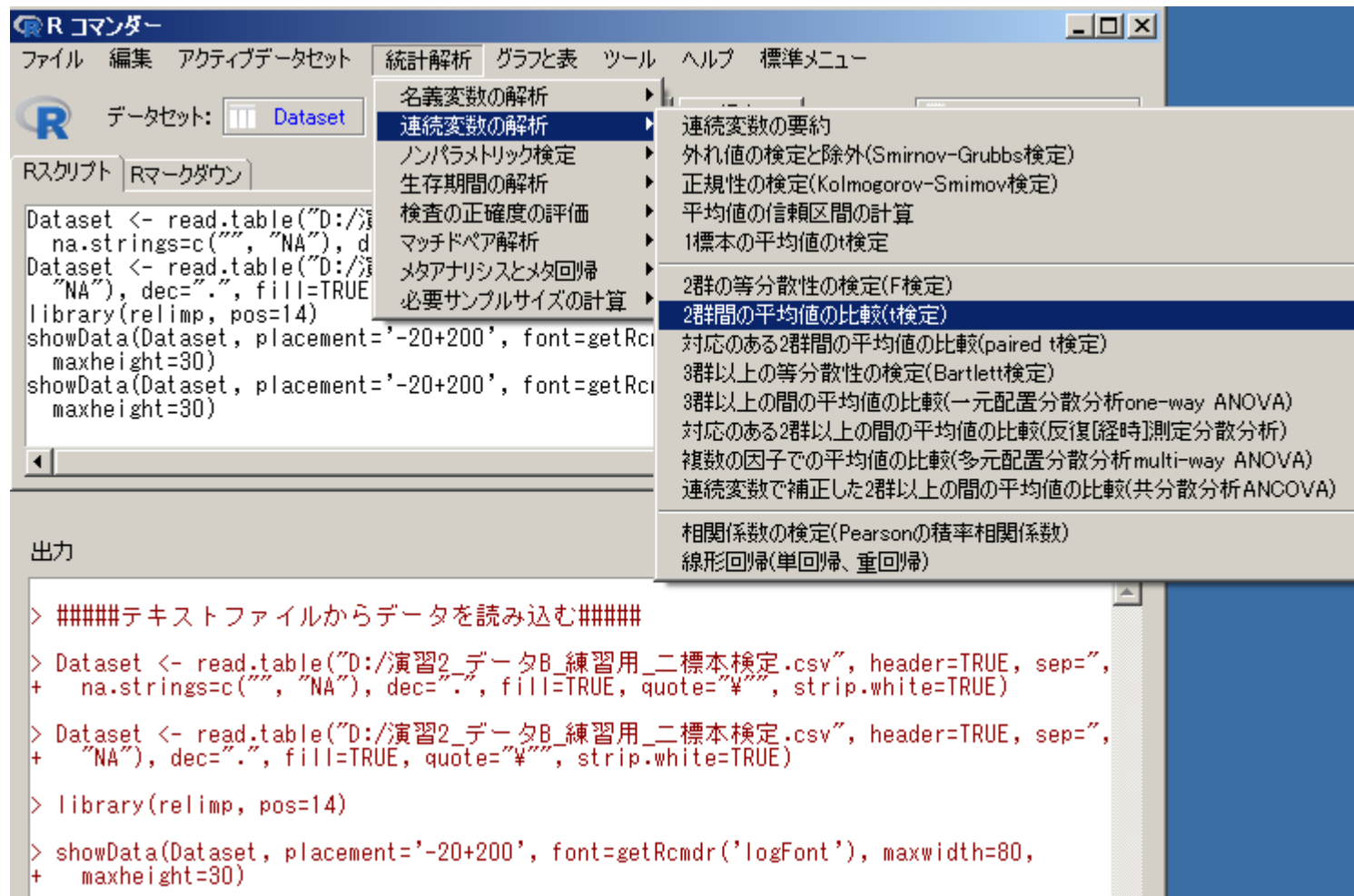
● 「表示」ボタンで内容の確認



	id	収縮期血圧	喫煙
1	1	150	YES
2	2	151	YES
3	3	121	YES
4	4	120	YES
5	5	120	YES
6	6	119	YES
7	7	119	YES
8	8	119	YES
9	9	138	YES
10	10	150	YES
11	11	138	YES
12	12	137	YES
13	13	136	YES
14	14	134	YES
15	15	133	YES
16	16	115	YES
17	17	113	YES
18	18	170	YES
19	19	129	YES
20	20	128	YES
21	21	127	YES
22	22	127	YES
23	23	127	YES
24	24	126	YES
25	25	126	YES
26	26	136	NO
27	27	132	NO
28	28	129	NO
29	29	128	NO
30	30	127	NO

「統計解析」メニューから

● 連続変数の解析→2群間の平均値の比較(t検定)



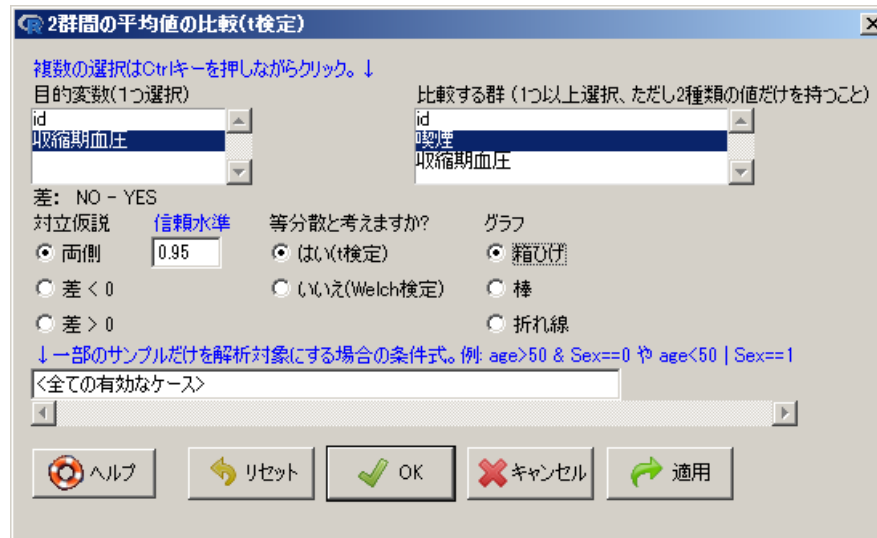
The screenshot shows the R Commander window with the 'Statistical Analysis' menu open. The path 'Continuous Variable Analysis' > 't-test for two groups' is highlighted. The console shows the following R code:

```
Dataset <- read.table("D:/演習2_データB_練習用_二標本検定.csv", header=TRUE, sep=";", na.strings=c("", "NA"), dec=".", fill=TRUE, quote="\"", strip.white=TRUE)
Dataset <- read.table("D:/演習2_データB_練習用_二標本検定.csv", header=TRUE, sep=";", na.strings=c("", "NA"), dec=".", fill=TRUE, quote="\"", strip.white=TRUE)
library(relimp, pos=14)
showData(Dataset, placement='-20+200', font=getRcmdrFont('logFont'), maxwidth=80, maxheight=30)
showData(Dataset, placement='-20+200', font=getRcmdrFont('logFont'), maxwidth=80, maxheight=30)
```

The output window displays the following text:

```
> #####テキストファイルからデータを読み込む#####
> Dataset <- read.table("D:/演習2_データB_練習用_二標本検定.csv", header=TRUE, sep=";",
+   na.strings=c("", "NA"), dec=".", fill=TRUE, quote="\"", strip.white=TRUE)
> Dataset <- read.table("D:/演習2_データB_練習用_二標本検定.csv", header=TRUE, sep=";",
+   "NA"), dec=".", fill=TRUE, quote="\"", strip.white=TRUE)
> library(relimp, pos=14)
> showData(Dataset, placement='-20+200', font=getRcmdrFont('logFont'), maxwidth=80,
+   maxheight=30)
```

目的変数、説明変数の指定



- 変数: Excelデータファイルの各列(統計解析の単位)
- 統計解析の多くは、目的変数Yと説明変数Xの関係の分析
 - ➡ この例の場合、
 - 目的変数Y: 収縮期血圧 (連続)
 - 説明変数X: 喫煙有無 (名義、離散、カテゴリー)

t検定の結果の確認

```
R コマンダー
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>

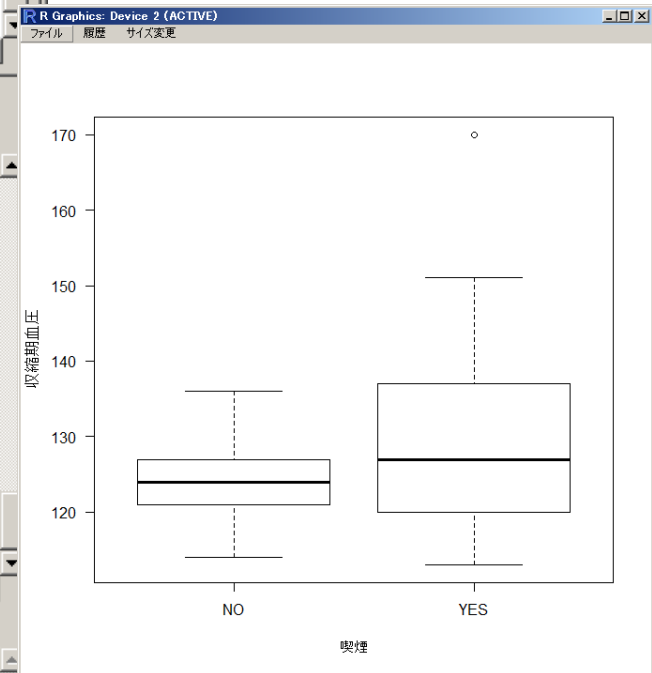
Rスクリプト Rマークダウン

group.names <- c(group.names, "喫煙=YES")
group.means <- c(group.means, bar.means[2])
group.sds <- c(group.sds, bar.sds[2])
group.p <- c(group.p, "")
remove(res)
summary.ttest <- data.frame(mean=group.means, sd=group.sds, p.value=group.p)
rownames(summary.ttest) <- group.names
colnames(summary.ttest) <- gettextRcmdr(colnames(summary.ttest))
summary.ttest

出力

> group.means <- c(group.means, bar.means[2])
> group.sds <- c(group.sds, bar.sds[2])
> group.p <- c(group.p, "")
> remove(res)
> summary.ttest <- data.frame(mean=group.means, sd=group.sds, p.value=group.p)
> rownames(summary.ttest) <- group.names
> colnames(summary.ttest) <- gettextRcmdr(colnames(summary.ttest))
> summary.ttest
      平均 標準偏差  P値
喫煙=NO 123.92  4.855581 0.0174
喫煙=YES 130.92 13.353901

メッセージ
[1] メモ: Rコマンダーのバージョン 2.2-3: Tue Jan 05 12:15:38 2016
[2] メモ: データセット Dataset には 50 行、3 列あります。
```



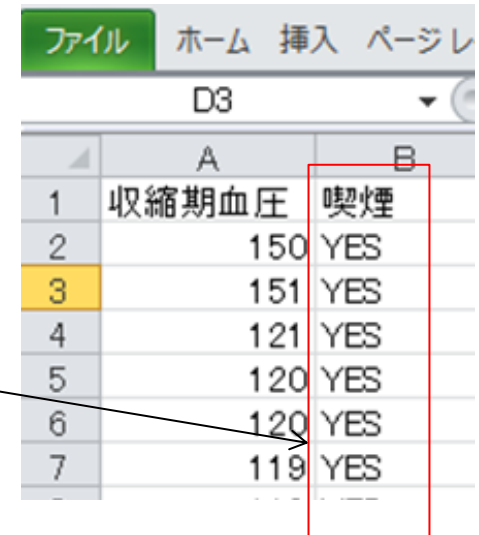
標準偏差が群間で大きく異なるので、t検定を使ったのは間違い

「変数」の種類について

- 連続変数と名義変数に大別

- 名義変数とは

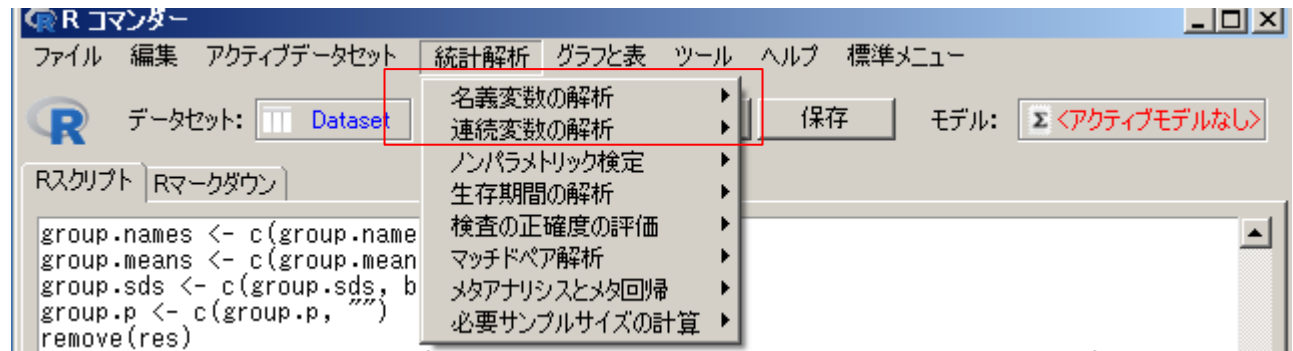
- ➡ 値が連続ではなく、むしろカテゴリーを意味
 - 任意の文字、カテゴリーに対応する数字
- ➡ 離散型変数と呼ぶことも



	A	B
1	収縮期血圧	喫煙
2	150	YES
3	151	YES
4	121	YES
5	120	YES
6	120	YES
7	119	YES

- 「変数」を「尺度」と呼ぶことも

- EZRは、目的変数が連続変数か名義変数かで、基本メニューを分けている



群間で分散は異なるが、データが正規分布なら？

● Welchのt検定

➡ 目的変数、説明変数を指定する画面にオプションあり

2群間の平均値の比較(t検定)

複数の選択はCtrlキーを押しながらクリック。↓

目的変数(1つ選択)

id
収縮期血圧

比較する群(1つ以上選択、ただし2種類の値だけを持つこと)

id
喫煙
収縮期血圧

差: NO - YES

対立仮説 信頼水準

☒ 両側 0.95

☐ 差 < 0

☐ 差 > 0

等分散と考えますか?

☐ はい(t検定)

☒ いいえ(Welch検定)

グラフ

☐ 箱ひげ

☒ 棒

☐ 折れ線

↓一部のサンプルだけを解析対象にする場合の条件式。例: age>50 & Sex==0 や age<50 | Sex==1

<全ての有効なケース>

ヘルプ リセット OK キャンセル 適用

ただし、この例の場合、喫煙あり群のデータは正規分布ではないのでWelchのt検定も使えない

t検定が使えないので、ノンパラメトリック検定

R コマンダー

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

```
group.names <- c(group.name
group.means <- c(group.mean
group.sds <- c(group.sds, b
group.p <- c(group.p, "")
remove(res)
summary.ttest <- data.frame(mean=group.means, sd=group.sds, p.value=group.p)
rownames(summary.ttest) <- group.names
colnames(summary.ttest) <- gettextRcmd
summary.ttest
```

統計解析

- 名義変数の解析
- 連続変数の解析
- ノンパラメトリック検定
- 生存期間の解析
- 検査の正確度の評価
- マッチドペア解析
- メタアナリシスとメタ回帰
- 必要サンプルサイズの計算

保存 モデル: <アクティブモデルなし>

2群間の比較(Mann-Whitney U検定)

対応のある2群間の比較(Wilcoxon符号付順位和検定)

3群以上の間の比較(Kruskal-Wallis検定)

対応のある3群以上の間の比較(Friedman検定)

連続変数の傾向の検定(Jonckheere-Terpstra検定)

相関係数の検定(Spearmanの順位相関係数)

2群間の比較(Mann-Whitney U検定)

目的変数(1つ選択)

id

収縮期血圧

比較する群(1つ以上選択、ただし2種類の値だけを持つこと)

id

喫煙

収縮期血圧

対立仮説

検定のタイプ

☒ 両側 ☒ デフォルト

☐ 差 < 0 ☐ 正確

☐ 差 > 0 ☐ 正規近似

☐ 連続修正を用いた
正規近似

↓一部のサンプルだけを解析対象にする場合の条件式。例: age>50 & Sex==0 や age<50 | Sex==1

<全ての有効なケース>

ヘルプ リセット OK キャンセル 適用

出力

```
> group.means <- c(group.means, bar.m
> group.sds <- c(group.sds, bar.sds[2
> group.p <- c(group.p, "")
> remove(res)
> summary.ttest <- data.frame(mean=gr
> rownames(summary.ttest) <- group.na
> colnames(summary.ttest) <- gettextR
> summary.ttest
```

	平均	標準偏差	P値
喫煙=NO	123.92	4.855581	0.0174
喫煙=YES	130.92	13.353901	

ウィルコクソン順位和検定の結果の確認

R コマンダー

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

```
group.max <- c(group.max, with(Dataset, max(収縮期血圧[喫煙=='YES'], na.rm=TRUE)))
group.p <- c(group.p, "")
remove(res)
mannwhitney.table <- data.frame(Minimum=group.min, Q1=group.Q1,
  Median=group.median, Q3=group.Q3, Maximum=group.max, p.value=group.p)
rownames(mannwhitney.table) <- group.names
colnames(mannwhitney.table)[c(2,4)] <- c("25%", "75%")
colnames(mannwhitney.table) <- gettextRcmdr(colnames(mannwhitney.table))
mannwhitney.table
```

出力

```
> group.max <- c(group.max, with(Dataset, max(収縮期血圧[喫煙=='YES'], na.rm=TRUE)))
> group.p <- c(group.p, "")
> remove(res)
> mannwhitney.table <- data.frame(Minimum=group.min, Q1=group.Q1,
+   Median=group.median, Q3=group.Q3, Maximum=group.max, p.value=group.p)
> rownames(mannwhitney.table) <- group.names
> colnames(mannwhitney.table)[c(2,4)] <- c("25%", "75%")
> colnames(mannwhitney.table) <- gettextRcmdr(colnames(mannwhitney.table))
> mannwhitney.table
```

	最小	25%	メディアン	75%	最大	P値
喫煙=NO	114	121		124	127	136 0.0566
喫煙=YES	113	120		127	137	170

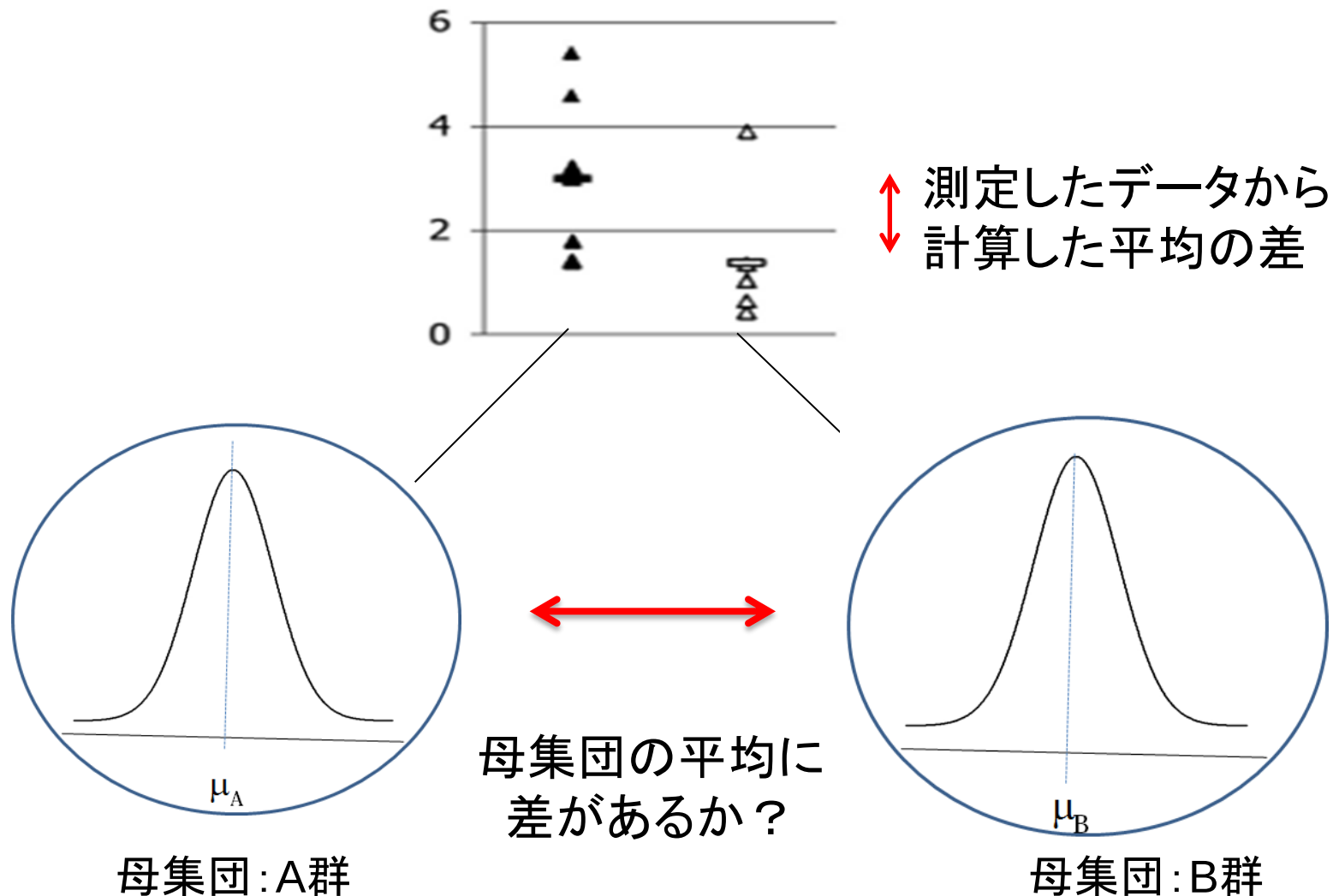
メッセージ

で警告がありました:
cannot compute exact p-value with ties

有意水準5%で
有意とは言えない

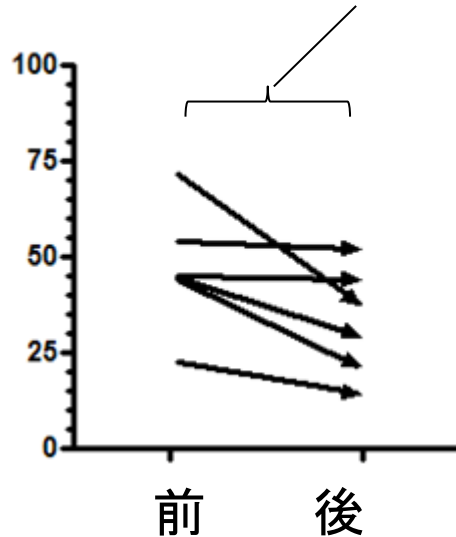
ここまでは「二標本」検定

- 2つの母集団からサンプルされた、2群のデータ(二標本)から、母集団の平均に差がある、と言えるかどうかの検定



もし同じ個体から2回測定したデータを比べる場合

- 例: 6株についてそれぞれ、試薬Aの投与前後で、ある遺伝子の発現量を測定。発現量が有意に変化したかどうかを調べたい。



対応のある (paired) t検定

or

ウィルコクソン符号付順位和検定

➡ 以下の点を考慮して検定

- 同じ個体からの2回の測定値の間には、相関がある
 - ・ 投与前が高い値なら、投与後も高い値を示しやすい
- 投与前と後の値は、2つの母集団から独立にサンプルされたものではない

EZRで、対応のある (paired) t検定

● 右のサンプルデータを読み込んで

	A	B
1	薬投与前	薬投与後
2	75	30
3	55	53
4	45	44
5	45	30
6	45	25
7	25	20

● データが正規分布すると考えられるなら、
「連続変数の解析」から

The screenshot shows the R Commander interface. The 'Stat' menu is open, and the path 'Continuous Data Summary' > 'Comparison of Means' > 'Comparison of Means of Two Groups' > 'Comparison of Means of Two Groups (paired t-test)' is highlighted. The 'paired t-test' dialog box is also shown, with 'Drug before' and 'Drug after' selected for the two groups. The confidence level is set to 0.95, and the alternative hypothesis is 'two-sided'.

前、後の変数を
第1、第2の変数に指定

EZRで、ウィルコクソン符号付順位和検定

- データが正規分布すると考えられない場合
「ノンパラメトリック検定」から

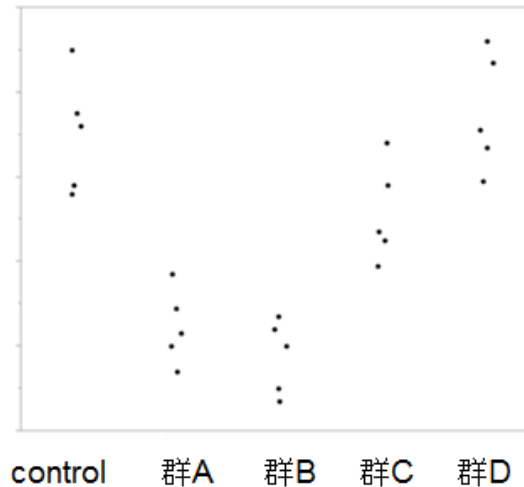
The screenshot shows the EZR (Easy R) software interface. The main window is titled "R コマンドー" and contains a menu bar with options like "ファイル", "編集", "アクティブデータセット", "統計解析", "グラフと表", "ツール", "ヘルプ", and "標準メニュー". Below the menu bar, there's a "データセット:" field with a dropdown menu showing "Dataset". To the right, there's a "保存" button and a "モデル:" field with a dropdown menu showing "<アクティブモデルなし>".

The "統計解析" menu is open, showing a list of non-parametric tests. The "ノンパラメトリック検定" option is selected, which has opened a sub-menu. In this sub-menu, the "対応のある2群間の比較(Wilcoxon符号付順位和検定)" option is highlighted.

A dialog box titled "対応のある2群間の比較(Wilcoxon符号付順位和検定)" is open in the foreground. It contains two dropdown menus for selecting variables: "第1の変数(1つ選択)" and "第2の変数(1つ選択)". Both are currently set to "薬投与後" and "薬投与前" respectively. Below these, there are radio buttons for "対立仮説" (Hypothesis) and "検定のタイプ" (Type of test). The "対立仮説" section has options for "両側" (Two-sided), "差 < 0" (Difference < 0), and "差 > 0" (Difference > 0). The "検定のタイプ" section has options for "デフォルト" (Default), "正確" (Exact), "正規近似" (Normal approximation), and "連続修正を用いた正規近似" (Normal approximation with continuity correction). At the bottom, there's a text field for "一部のサンプルだけを解析対象にする場合の条件式。例: age>50 & Sex==0 や age<50 | Sex==1" (Condition for analyzing only some samples. Example: age>50 & Sex==0 or age<50 | Sex==1) and a button for "全ての有効なケース" (All valid cases). The dialog box also has buttons for "ヘルプ", "リセット", "OK", "キャンセル", and "適用".

前、後の変数を
第1、第2の変数に指定

比較する群の数が増えた場合は？



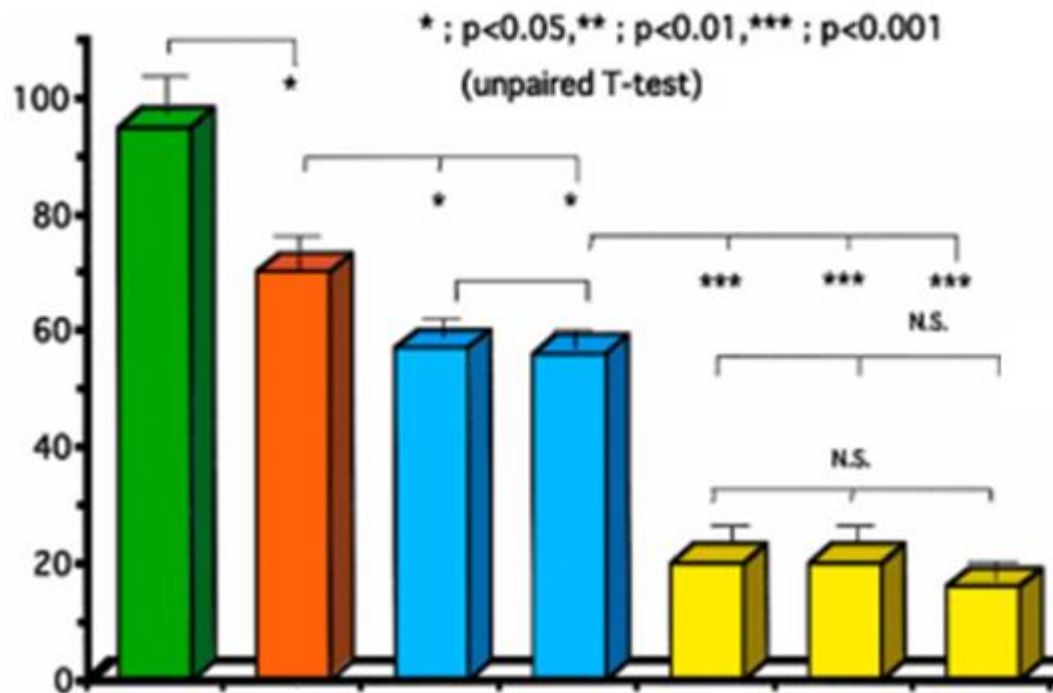
データが正規分布に従う仮定	3群以上	2群の場合
できる	分散分析	t検定に一致
できない (ノンパラメトリック法が必要)	Kruskal-Wallis検定	ウィルコクソン 順位和検定に一致

● 検定の帰無仮説:「どの群間にも、平均値に差がない」

➡ P値が小さい→ どこかの群間に、平均値の差がある

さらに、どの群間に差があるかを突き止める(「対比較」)

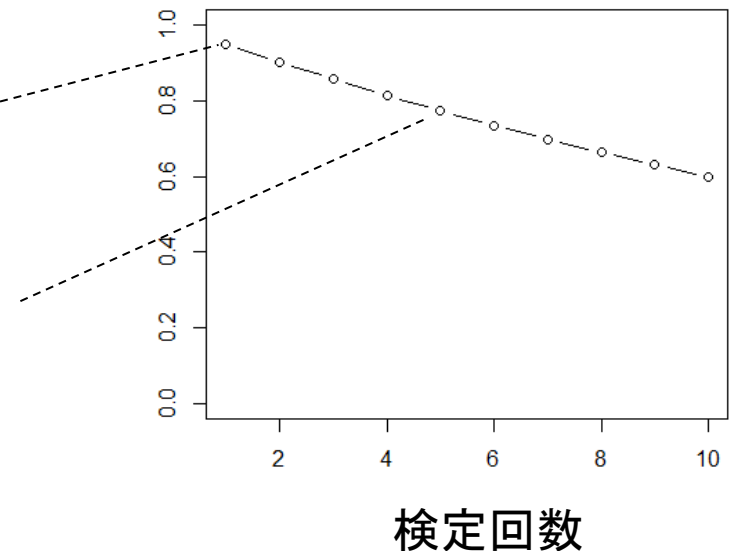
- よくある間違い:t検定・ウィルコクソンの順位和検定をすべての群間で繰り返す



日本細菌学会誌の
論文より改変

検定を繰り返すことの問題点

- 誤って有意と判断してしまう確率が高くなる。なぜなら、
 - ➡ $p < 0.05$ を有意と判断する(有意水準5%で検定を行う)場合、帰無仮説が正しくても、これぐらいの差は5%の確率で生じる
 - ➡ 1回検定を行うと、誤って有意と判断する確率は5%
 - false positive
 - α エラー(第一種の誤り)
 - ➡ **その誤りの生じない確率**
 - 1回の検定につき95%
 - 2回検定を行うと $0.95^2 = 90\%$
 - 5回検定を行うと $0.95^5 = 77\%$



対比較の方法

- 検定を繰り返しても、第1種の誤り(false positive)の生じる確率が大きくならないよう、有意水準かp値を調整
- 一番基本的なBonferroni法
 - ➡ 例えば検定を5回行う場合
 - 1回あたり $p < 0.05/5 = 0.01$ なら、有意と判断する
 - 各検定での有意水準を厳しくする
 - 5回を通じて第1種の誤りの生じない確率は
 - $(1 - 0.05/5)^5 = 95\%$ に収まる
 - ➡ シンプルで理解しやすいが、これでは厳しすぎて、有意差があると考えられるのに、有意と言えない場合が多い

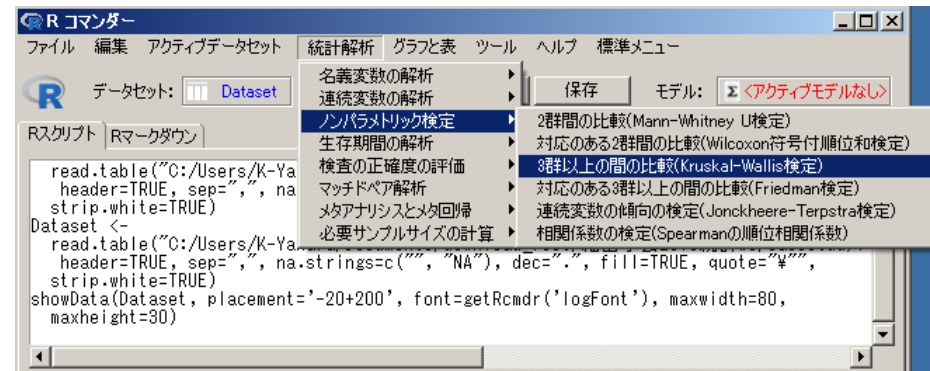
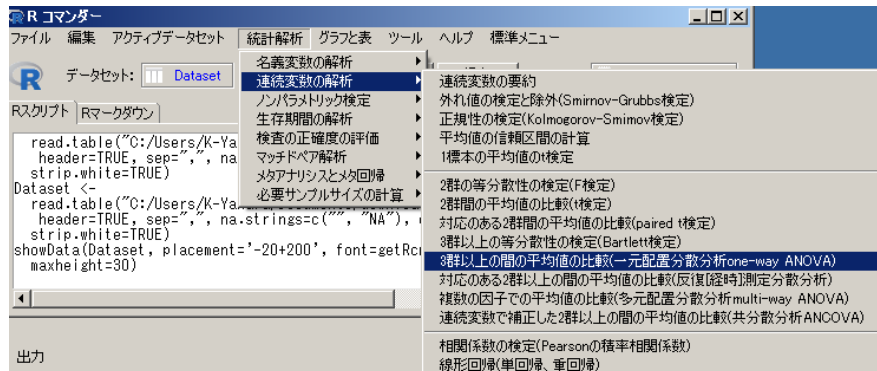
対比較の方法 (Bonferroniより緩い)

● 最もよく使われるのは以下の4つ

データが正規分布に従う仮定 できる	全ての群を ペアワイズに比較 Tukey法	Control群 vs その他の群を1つずつ比較 Dunnett法
できない (ノンパラメトリック法が必要)	Steel-Dwass法	Steel法

● EZRでのマウス操作を通じて、この場合分けを理解しましょう

➡ 分散分析またはKruskal-Wallis検定を選んで



分散分析 → 対比較 の場合

3群以上の間の平均値の比較(一元配置分散分析 one-way ANOVA)

モデル名を入力: AnovaModel2

複数の選択はCtrlキーを押しながらクリック。↓

目的変数(1つ選択)

t
value

比較する群(1つ以上選択)

t
treat
value

グラフ

等分散と考えますか?

☒ 箱ひげ ☒ はい(一元配置分散分析)

☐ 棒 ☐ いいえ(Welch検定)

☐ 折れ線

↑ Welch検定の場合は多重比較やモデルの保存は行われません。
↓ 2組ずつの比較(post-hoc検定)は比較する群が1つの場合のみ実施される。

☐ 2組ずつの比較(Bonferroniの多重比較)

☐ 2組ずつの比較(Holmの多重比較)

☐ 2組ずつの比較(Tukeyの多重比較)

☒ 2組ずつの比較(Dunnettの多重比較)

↑ アルファベット順で一番若い群が対照群となる。変更したい場合は因子水準を再順序化する。

☐ モデル解析用に解析結果をアクティブモデルとして残す

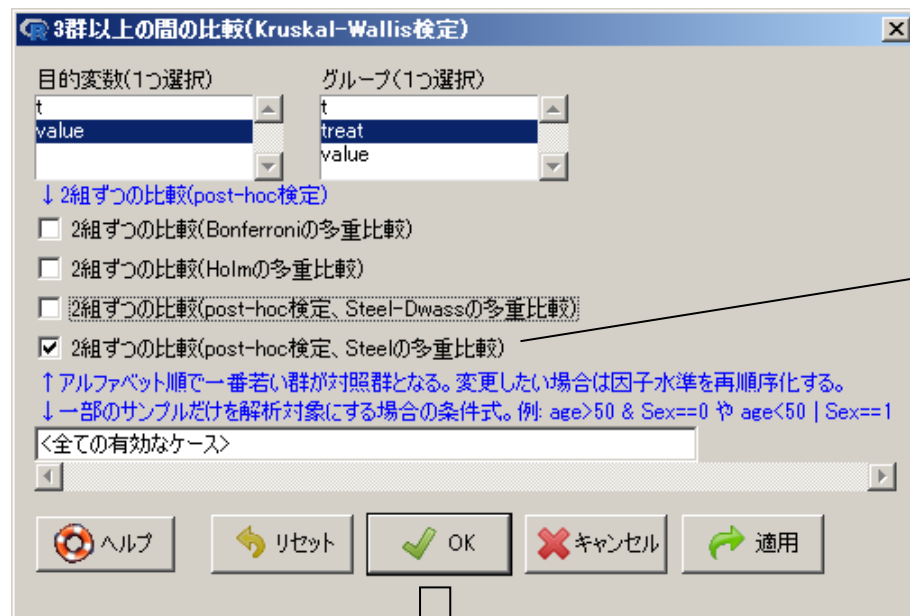
↓ 一部のサンプルだけを解析対象にする場合の条件式。例: age>50 & Sex==0 や age<50 | Sex==1

<全ての有効なケース>

ヘルプ リセット OK キャンセル 適用

Control群
vs その他の群を1つ1つ比較
(Dunnett法)

Kruskal-Wallis検定 → 対比較 の場合



Control群
vs その他の群を1つ1つ比較
(Steel法)

3群以上の間の比較(Kruskal-Wallis検定)

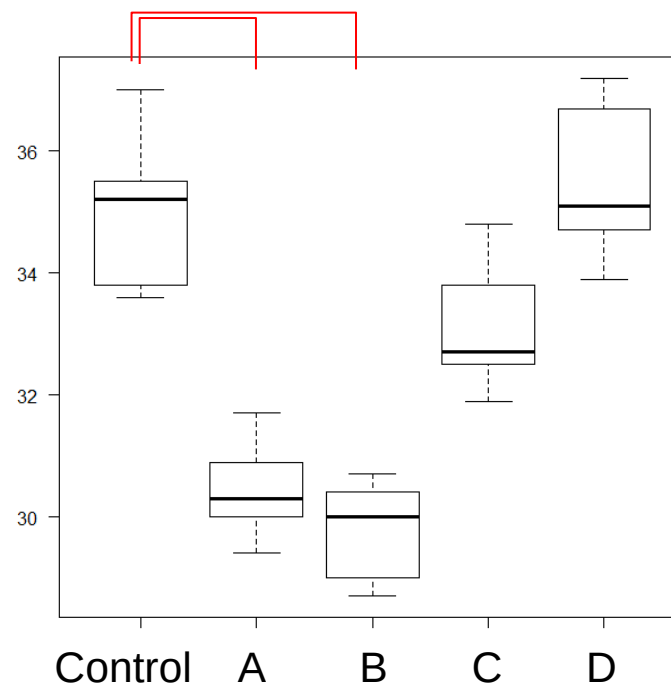
P値 = 0.000575

```
> remove(res)
```

```
> Steel(Dataset$value, Dataset$treat)
```

	t	rho	p.value
control:drugA	2.611165	0.5	<u>0.03167478</u>
control:drugB	2.611165	0.5	<u>0.03167351</u>
control:drugC	1.885762	0.5	0.18168440
control:drugD	0.522233	0.5	0.95919900

調整済p値



エラーバーについて

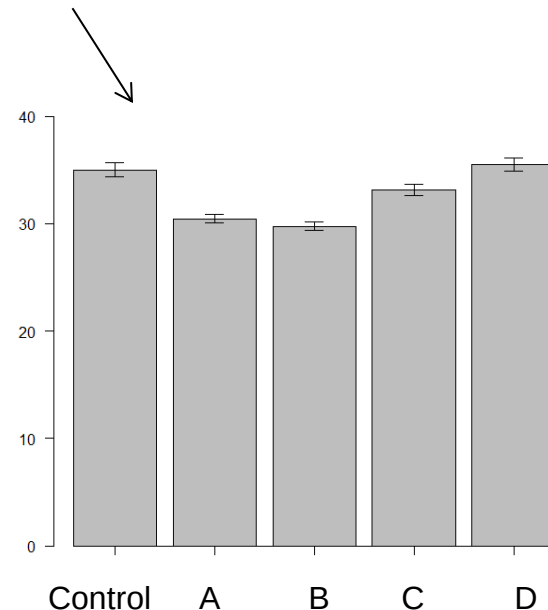
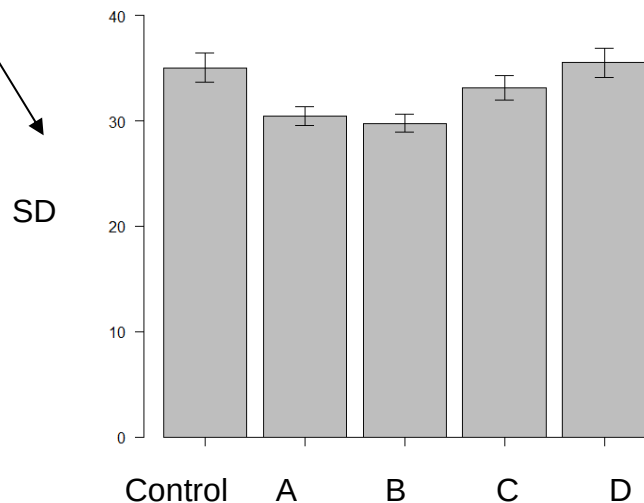
● 3つの可能性

① データの標準偏差 SD

② データの平均の標準偏差 (= 標準誤差, Standard Error of Mean)

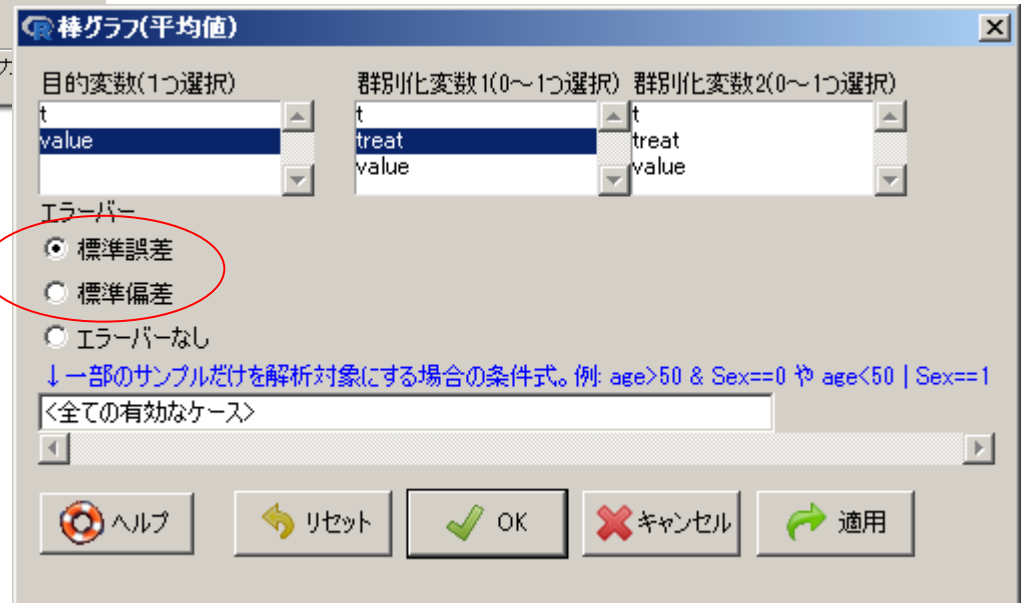
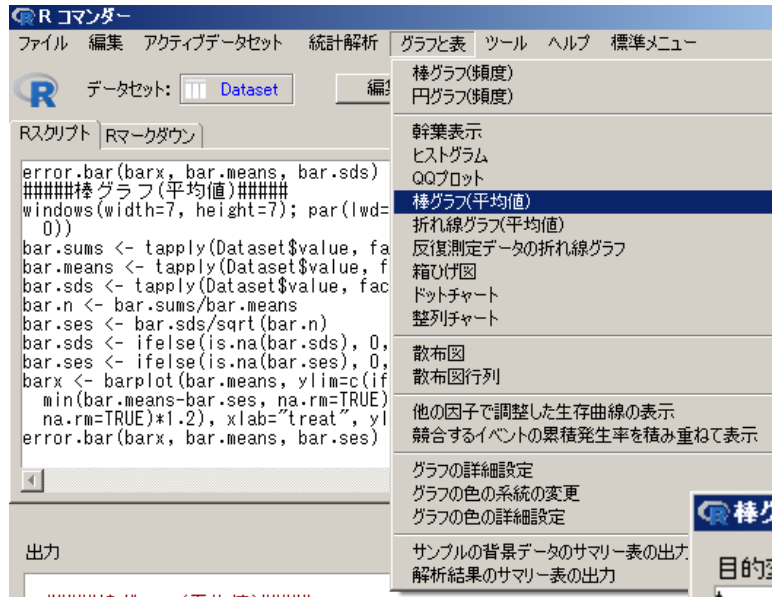
- 同じ実験を繰り返しても、得られる平均値は一致しない。それがどれだけバラつくのか、の指標。

- 当然、データの標準偏差より小さい $\frac{SD}{\sqrt{\text{サンプル数}}}$



標準偏差、標準誤差の使い分けの指定

● EZRで、前ページの棒グラフを作るには

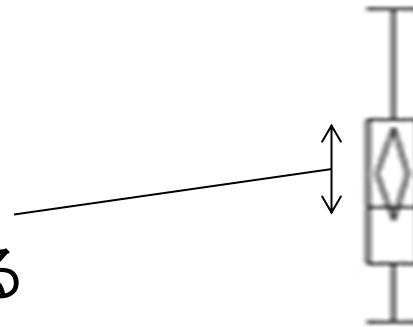


誤差バーについて

● 3つの可能性

③データの平均の95%信頼区間

- 95%の確率で、平均値はこの範囲にある
- サンプル数が大きければ、近似的に
 - 平均値 $\pm 1.96 \times$ 標準誤差
- 有償ソフトJMPでは
 - 箱ひげ図(前述)にデフォルトで表示される
 - 棒グラフでの表示も「グラフビルダー」で可能



誤差バーがどれを意味しているのかを明記する必要あり

細菌学分野に特有のデータについて

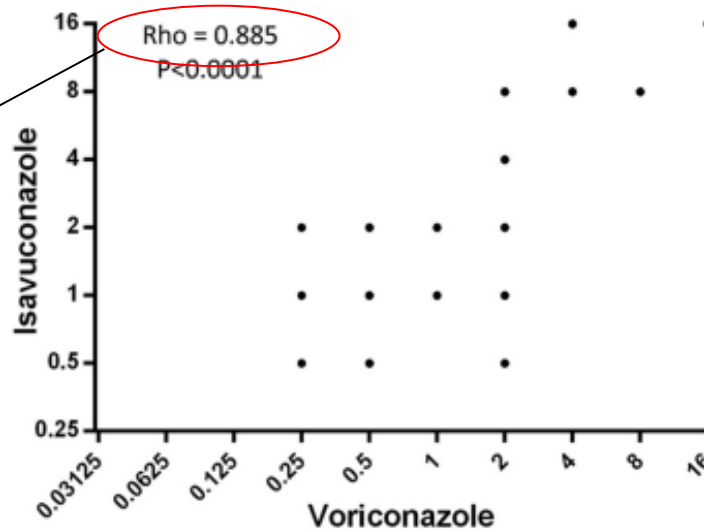
● CFU

- ➡ 2群比較には、ウィルコクソン順位和検定を使うのが無難
 - t検定の前提(データが正規分布する)が満たされる保証なし(もし正規分布することが事前に分かっているなら、t検定でよい)

● 薬剤感受性試験のMICや抗体価

- ➡ 測定値は、段階的で飛び飛びの値
 - 例) 1, 2, 4, 8、もしくは0.5, 0.25 ...
- ➡ 連続変数ではなく、「順序変数」(名義変数の一部)
 - データが連続型分布に従うことを仮定するパラメトリック法(例えばt検定)を使うのは不適切
 - ノンパラメトリック法、または名義変数(カテゴリーデータ)の解析法を使う必要あり

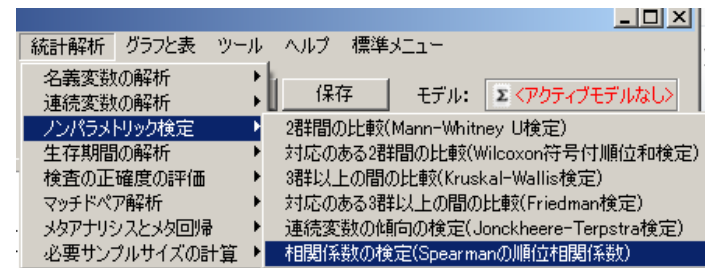
例: 薬A,BのMICの相関は？



Gregson (2013),
Anti. Agents. Chem.

➡ データが正規分布に従うことを仮定
→ピアソンの相関係数

➡ 仮定しない(ノンパラメトリック)
→スピアマンの相関係数



※データを予め順位化した上で、相関を計算

本日紹介した統計解析法の整理

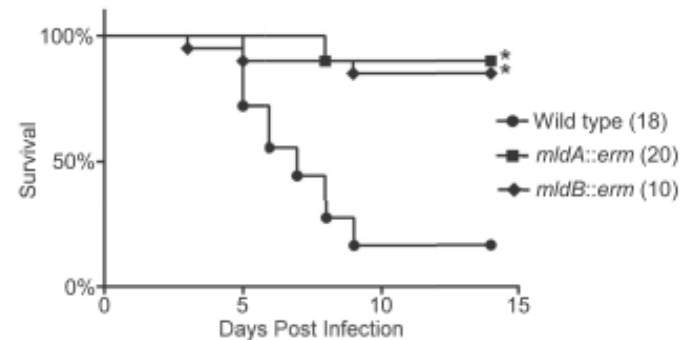
	正規分布を仮定	ノンパラメトリック
2群比較	t検定	ウィルコクソンの順位和検定
同一個体の 2回測定と比較	対応のあるt検定	ウィルコクソンの符号付順位和検定
3群以上の比較	分散分析 → Tukey法 or Dunnett法	Kruskal-Wallis検定 → Steel-Dwass法 or Steel法

相関係数	ピアソン	スピアマン
------	------	-------

- いずれも、フリーのマウス操作型ソフトEZRで使えます
- データが正規分布する保証がないならノンパラメトリック法
 - ➡ CFU, MIC, 抗体価 ...
 - ただし、有意差を示すのに、より多くのサンプルサイズが必要

今回は対象外とさせて頂いた内容

- 目的変数が名義変数(例:ある菌が薬剤耐性かどうか)
 - ➡ 分割表とカイ2乗検定・Fisherの正確検定、ロジスティック回帰
- 細菌を動物に感染させた後の生存時間に関する解析



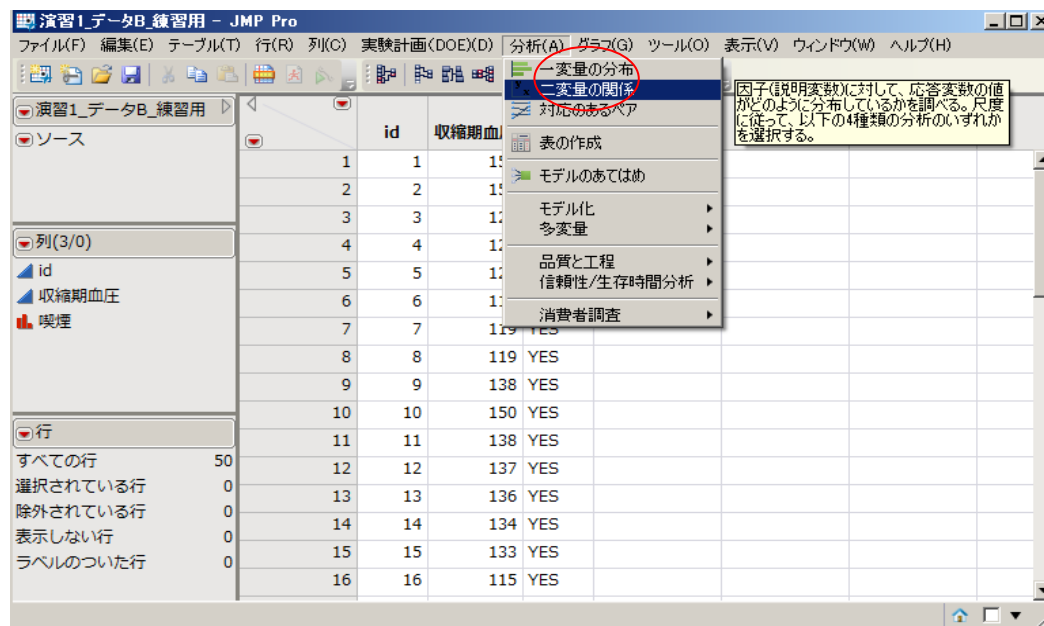
- 疫学研究に必要な
 - ➡ サンプルサイズ設計
 - 有意差によって説得力のある結果を示すために、何株・どれだけのデータを集めればよいのか？
 - ➡ 重回帰分析
 - 複数の要因と目的変数の関係を探る

マウス操作型統計解析ソフトの比較

● JMP (ジャンプ)

➡ 久留米大、九州大、京都大等、全国の医学部の講義・演習で活用

➡ 全てマウス操作。ラボデータ解析の大半は「分析」メニューの「二変量の関係」から統一的に可能。



➡ 有償だが、最も使いやすく定評のあるマウス操作型ソフト

「分析」→「二変量の関係」でt検定を行う例

演習1_データB_練習用 - JMP Pro

ファイル(F) 編集(E) テーブル(T) 行(R) 列(C) 実験計画(DOE)(D) 分析(A) グラフ(G) ツール(O) 表示(V) ウィンドウ(W) ヘルプ(H)

一変量の分布
二変量の関係
 対応のあるペア
 表の作成
 モデルのあてはめ
 モデル化
 多変量
 品質と工程
 信頼性/生存時間分析
 消費者調査

因子(説明変数)に対して、応答変数の値がどのように分布しているかを調べる。尺度に従って、以下の4種類の分析のいずれかを選択する。

	id	収縮期血
1	1	15
2	2	15
3	3	12
4	4	12
5	5	12
6	6	13
7	7	119 YES
8	8	119 YES
9	9	138 YES
10	10	150 YES
11	11	138 YES
12	12	137 YES
13	13	136 YES
14	14	134 YES
15	15	133 YES
16	16	115 YES

演習1_データB_練習用

ソース

列(3/0)

id
 収縮期血圧
 喫煙

行

すべての行 50
 選択されている行 0
 除外されている行 0
 表示しない行 0
 ラベルのついた行 0

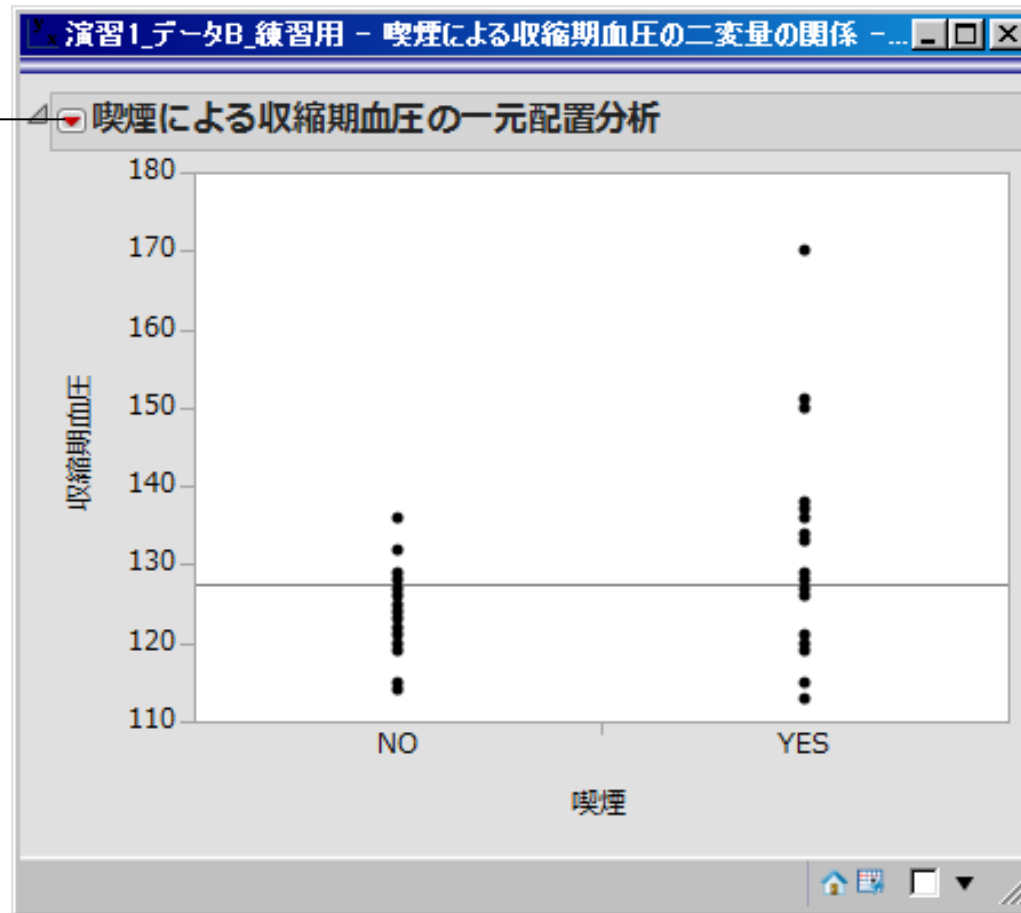
「Y, 目的変数」と「X, 説明変数」を指定

➡ 例えば「CFU」と「実験条件」

➡ 先ほどの例なら「収縮期血圧」と「喫煙有無」



XとYの関係のプロット



赤いボタンを押すと、この先に可能な解析のリストが表示

t検定を実行

演習1_データB_練習用 - 喫煙による収縮期血圧の二変量の関係 -

喫煙による収縮期血圧の一元配置分析

- 分位点
- 平均/ANOVA/プーリングしたt検定
 - 平均と標準偏差
 - 個々の分散を用いたt検定
 - 平均分析法
 - 平均の比較
 - ノンパラメトリック
 - 等分散性の検定
 - 同等性の検定
 - ロバストなあてはめ
 - 検出力...
 - α水準の設定
- 正規分位点プロット
- 累積確率プロット
- 密度
- 対応のある列を設定...
- 保存
- 表示オプション
- スクリプト

t検定、分散分析、および平均のレポートの表示/非表示を切り替える。

収縮期血圧

喫煙

演習1_データB_練習用 - 喫煙による収縮期血圧の二変量の関係 - JMP Pro

喫煙による収縮期血圧の一元配置分析

一元配置の分散分析

あてはめの要約

R2乗	0.112217
自由度調整R2乗	0.093721
誤差の標準偏差(RMSE)	10.04747
応答の平均	127.42
オブザベーション(または重みの合計)	50

t検定

YES-NO

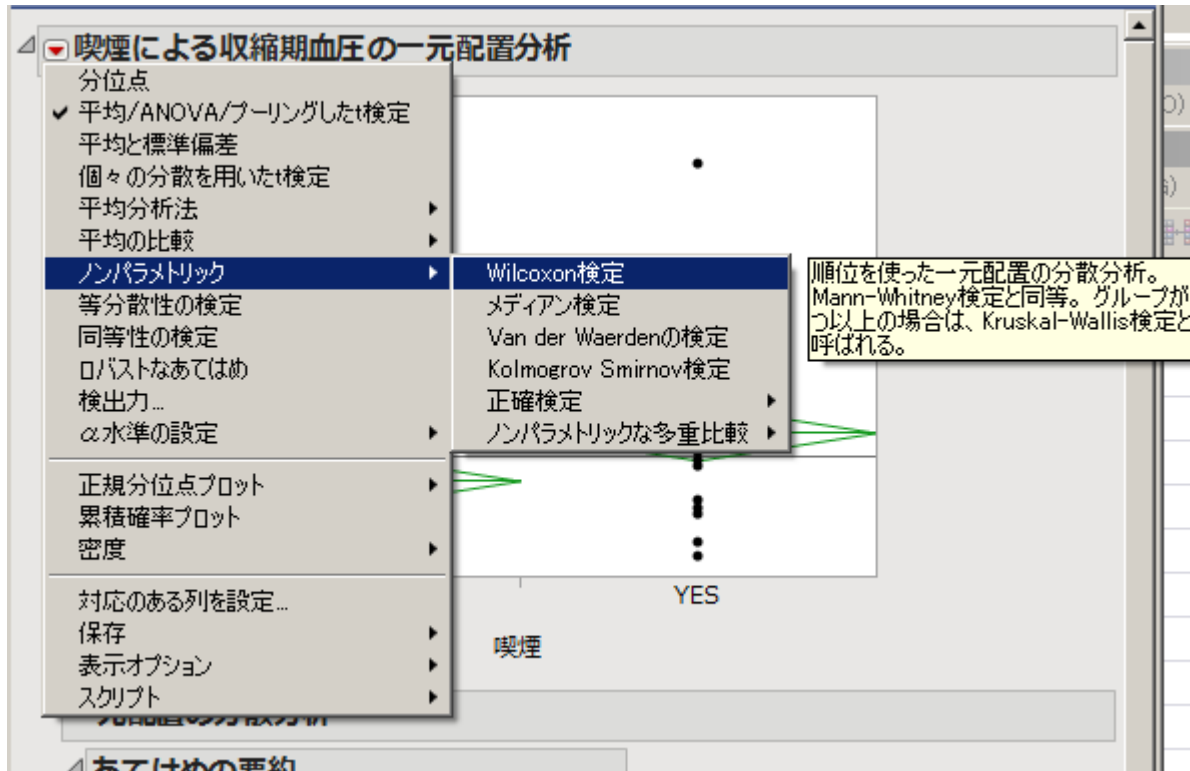
分散が等しいと仮定

差	7.0000	t値	2.463181
差の標準誤差	2.8419	自由度	48
差の上側信頼限界	12.7139	p値(Prob> t)	0.0174*
差の下側信頼限界	1.2861	p値(Prob>t)	0.0087*
信頼率	0.95	p値(Prob<t)	0.9913

分散分析

要因	自由度	平方和	平均平方	F値	p値(Prob>F)
喫煙	1	612.5000	612.500	6.0673	0.0174*
誤差	48	4845.6800	100.952		
全体(修正済み)	49	5458.1800			

ウィルコクソン順位和検定を実行



Wilcoxon/Kruskal-Wallisの検定(順位和)

水準	度数	スコア和	スコアの期待値	スコア平均	(平均-平均0)/標準偏差0
NO	25	539.000	637.500	21.5600	-1.906
YES	25	736.000	637.500	29.4400	1.906

2標本検定(正規近似)

S	Z	p値(Prob> Z)
736	1.90620	0.0566

一元配置検定(カイ2乗近似)

カイ2乗	自由度	p値(Prob>ChiSq)
3.6708	1	0.0554

JMPを買う？ 買わない？

● 買うなら

➡ 全学ライセンス: 年間60万円

- 研究室ライセンス(年間45万円)と大差ない
- 仲間を作って大学で買えば、多くの人にとって有益

● 買わないなら

➡ 30日間、トライアル版が利用可能

- トライアル版で「分析」メニューの「二変量の解析」を使いながら統計学・統計解析の基礎と考え方を学べます

➡ その後は、無償のマウス操作型ソフト

- 本セミナーで使った「EZR」が、現時点では一番おすすめ

謝辞

● 事前にアドバイスを頂いた細菌学者の先生方

➡ 杏林大

- 米澤先生、大崎先生、北条先生、神谷先生

➡ 感染研・細菌第二部

- 鈴木先生、松井先生、筒井先生

➡ 九州大

- 林先生、小椋先生、後藤先生
- 大岡先生(現鹿児島大)

➡ 阪大

- 堀口先生