

# 統計のいろは

大阪大学大学院 工学研究科

ビジネスエンジニアリング専攻 田中 勇気

# 研究とは

## ➤ 勉強と研究の違い

### ● 勉強

先人が発見し、定説となっている知識を自分の中に新しく仕入れる**受動的**な営み（**先生から答えを教えてください**）

### ● 研究

何かの現象の原因や関連要因を自分で新しく説明したり、新しい現象の存在を自分が発見する**能動的**な営み（**自ら答えを見つける**）

文献調査など

研究を進めていくうえでは、研究目的に沿った勉強が必要となる

# 研究の概念

## ➤ 勉強と研究の違い(イメージ)



この証明には  
こういう実験設定  
を考えれば...

こんな結果に  
なった原因は  
恐らく...



理論値通り  
になったぞ!



勉強(講義・実験)

研究

# 研究のスタイル

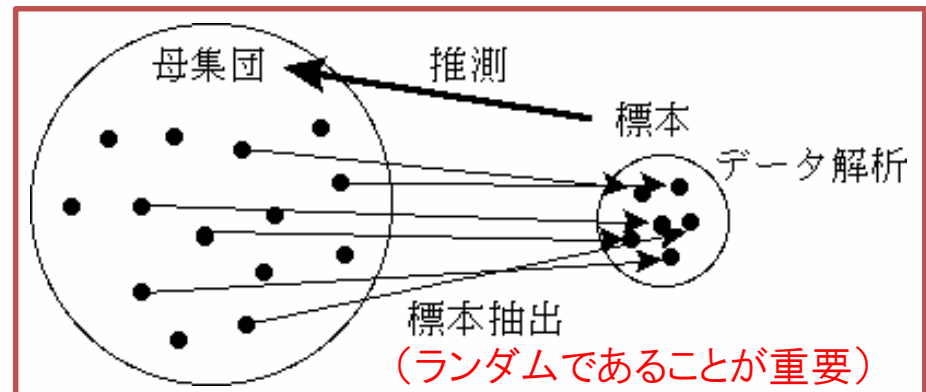
## ➤ 研究は基本的に帰納法のスタイルをとる

### ● 帰納法

経験的事実から共通性を見つけ、普遍的な法則を求める  
実験(具体) 結論(抽象)

- ◆ 考えうる全事例を網羅しない限り、結論は真理とならない

現実的に無理



母集団から抽出した標本を対象として実験することがほとんど

# 研究価値

➤ 研究の価値は新規性と有用性で決まる

- 新規性

過去に誰もやっていない、新しい研究であるか

- 有用性

研究成果が世の中の役に立つか

素晴らしい結果  
が出ましたよ！

新規性

それ僕の  
研究結果と  
同じ

プログラミングが  
出来る人は  
身長が高い！

有用性


それで  
どうすると？

# 研究の客観性

➤ 研究自体は多くの人々が納得し、同じ結論へと到達できる客観的な方法で実施する必要がある


✓ なぜそう言えるのか？

✓ だから何が言えるのか？ を明確にする



僕が思うに結論はこうなります。

なぜ？



結果としてはA群が高いです。

だから？

で？

# 統計的仮説検定の基礎

統計的仮説検定とは  
母集団について立てた仮説が正しいかどうかを標本から推測すること

# 仮説検定の手順

実際はあってほしい

1. まず「A群とB群の間に有意差はない」という帰無仮説を立てる
  - 有意差: ある事柄の起こる確率が有意水準未満であること  
つまり偶然であるとは考えにくいこと
  - 帰無仮説: 「無に帰する」ことを目指す仮説  
つまり「行った実験は何の意味もない」という仮説
2. データに対して適切な有意差検定を行う
3. 有意差があれば帰無仮説が棄却され、A群とB群の間の差は偶然ではなく、明確に異なる差であると結論付けられる



# 仮説検定の手順(補足)

## ➤ 「有意差がある」とはどういうことか

有意水準5%で検定したときに検定統計量が0.05未満であれば有意差あり

- 有意差: ある事柄の起こる確率が有意水準未満であること  
つまり偶然であるとは考えにくいこと
- 帰無仮説: 「無に帰する」ことを目指す仮説  
つまり「行った実験は何の意味もない」という仮説
- 有意水準: 検定統計量と自由度に基づいて算出される確率  
一般には5%(たまに1%)が使われる
- 検定統計量: 平均値の差などの数値が偶然得られたのか、  
理由があって得られたのか判断する材料となる数値
- 自由度: 自由に変動できる値の数  
一般にはデータ数もしくはデータ数-1の値が使われる

自由度や検定統計量の算出方法など  
詳しく知りたい場合は参考文献の統計専門書などを参照

# 検定のいろいろ

- 2群の平均値の差をみたい場合
  - 対応のあるt検定
  - 対応のないt検定
- 3群以上の平均値の差をみたい場合
  - 分散分析
    - ◆ 多重比較
- 2群間の関係をみたい場合
  - 相関分析
  - 回帰分析

ここではよく使われる「2群の平均値の差の検定」について説明

# なぜ検定を行う？

- 2群の平均値の差が偶然によるものでないことを確かめるため

被験者	中間試験[点]	期末試験[点]
A君	52	87
B君	44	98
C君	48	11
D君	56	4
平均値	50	50

データの分散が見えないため、  
平均値だけで結論付けるのは危険



難易度が同じなら  
次はもっと難しくしよう♪

# なぜ検定を行う？

- 2群の平均値の差が偶然によるものでないことを確かめるため

この勉強法でやったら  
英語の点が上がった！

ほんと！？  
真似しよう！

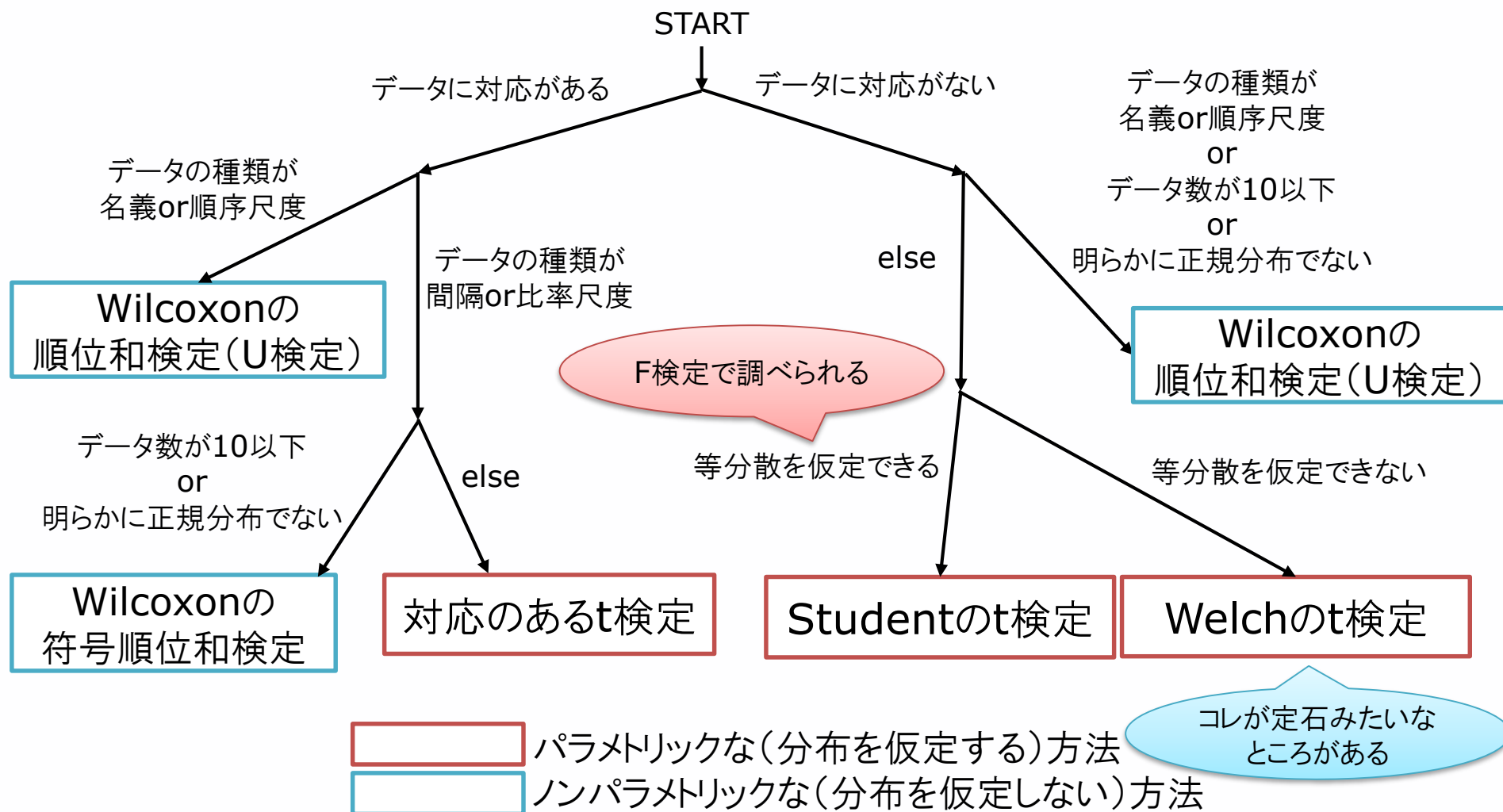
点数	中間試験[点]	期末試験[点]
英語	64	65

点数が上がったのは事実だが、  
差の度合いによっては捉え方が全く異なる

正しい検定手法を選んで使おう！



# 検定手法の選び方



# 尺度の水準

## ➤ 名義尺度

数値の大きさが意味をもたない尺度(数値を入れ替えてもOK)

ex. リレーにおけるレーン 1:赤組 2:白組 3:青組

## ➤ 順序尺度

数値の大きさが意味をもつ尺度

ex. リレーにおける順位 1位:白組 2位:赤組 3位:青組

## ➤ 間隔尺度(数値の加減を考えられる)

順序尺度の条件+数値間の間隔が一定の尺度

ex. リレーにおけるタイム 20秒 ←1秒→ 21秒 ←1秒→ 22秒

## ➤ 比率尺度(数値の加減乗除を考えられる)

間隔尺度の条件+倍数を考えることができる尺度

ex. リレーにおける走行距離 100m ←2倍→ 200m

# 対応の有無

## ➤ 「対応がある」とは

データの1行が1人分にあたるような場合

ex. 2回分のテストの得点比較

被験者	中間試験[点]	期末試験[点]
A君	52	87
B君	44	98

52点も87点も  
A君の得点

## ➤ 「対応がない」とは

データの1行に共通性がない場合

ex. 男女間による1回分のテストの得点比較

男性	女性
52	87
44	98

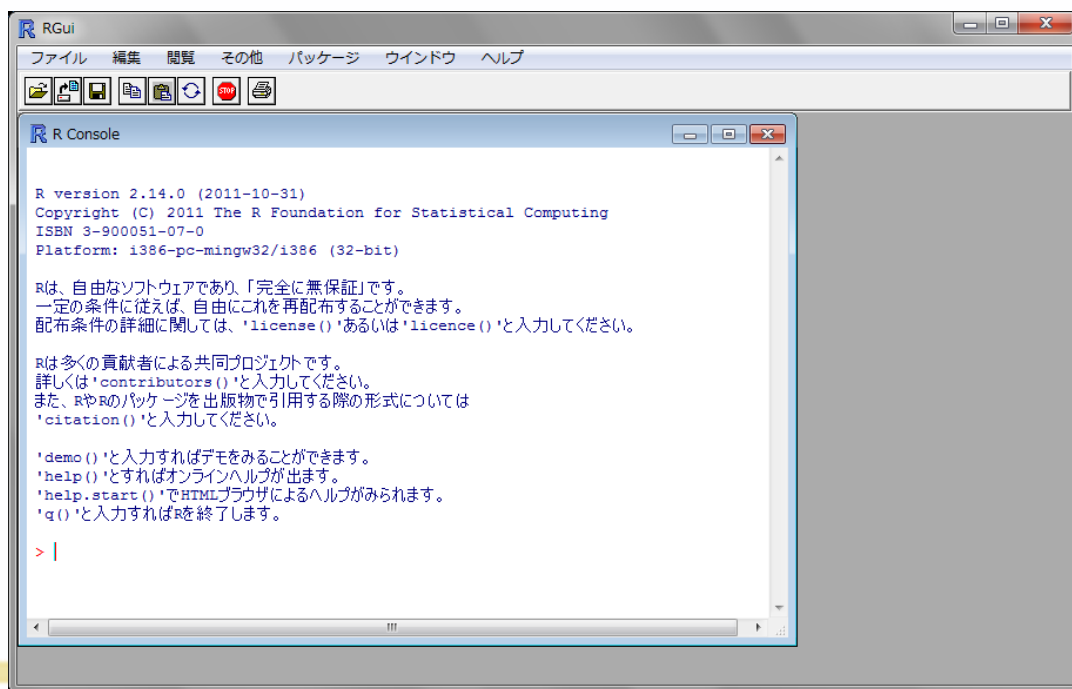
52点と87点は  
異なる学生

# 統計ソフト「R」

## ➤ 統計ソフト「R(R言語)」

ネットから無料でダウンロードできる統計ソフト

多くの統計科学者がRのパッケージの開発に携わっており、最先端の統計手法はまずRで実装されると言っても過言ではない  
有名なExcelとかSPSSとかでもできるけど有料



```
RGui
ファイル 編集 閲覧 その他 パッケージ ウィンドウ ヘルプ
R Console
R version 2.14.0 (2011-10-31)
Copyright (C) 2011 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。

> |
```



Rでの検定方法を  
簡単に紹介します



# Rで検定を行う前に

- ディレクトリとファイル名を指定してデータを読み出す

```
>data <- read.csv("file.csv", header=TRUE)
```

変数名

ファイル名

ヘッダ

file.csv

	A	B	C	D
1	s25	s50	s75	s100
2	12.90653	9.786769	13.58686	13.76906
3	10.12766	10.71223	8.86082	12.88216
4	8.573567	12.01133	8.103967	11.79165
5	11.19755	13.08732	13.16595	7.510244
6	9.891765	15.9068	14.11829	10.00944
7	12.23712	12.1806	11.93099	10.58737
8	10.86167	7.852693	10.68842	10.05458
9	12.69706	8.641875	9.918166	11.89757
10	11.79685	13.17343	10.76867	14.2013
11	8.715774	11.24256	11.67802	14.77374
12	9.818111	10.72785	9.781646	13.9257
13	7.750089	11.82814	13.18392	10.59824
14	9.806922	14.95375	11.0036	6.863839
15	10.98579	11.54986	11.37553	11.8797

data\$s50という表記でアクセス可

# 対応のあるt検定

## ➤ 条件

- データに対応がある
- データの種類が間隔尺度あるいは比率尺度である
- データ数が11以上ある
- データの分布が正規分布である、もしくは正規分布が仮定できる

```
>t.test(data$s25, data$s50, paired=TRUE)
```

```
Paired t-test

data:  data$s25 and data$s50
t = -1.9396, df = 22, p-value = 0.06535
alternative hypothesis: true difference in means is not equal to 0
95 percent confidence interval:
 -2.34331764  0.07837075
sample estimates:
mean of the differences
      -1.132473
```

# Studentのt検定

## ➤ 条件

- データに対応がない
- データの種類が間隔尺度あるいは比率尺度である
- データ数が11以上ある
- データの分布が正規分布である、もしくは正規分布が仮定できる
- 等分散を仮定できる

```
>t.test(data$s25, data$s50, var.equal=TRUE)
```

```
Two Sample t-test

data:  data$s25 and data$s50
t = -2.3441, df = 45, p-value = 0.02355
alternative hypothesis: true difference in means is not equal to 0
95 percent confidence interval:
 -2.1151394 -0.1601388
sample estimates:
mean of x mean of y
 10.57169  11.70933
```

# Welchのt検定

## ➤ 条件

- データに対応がない
- データの種類が間隔尺度あるいは比率尺度である
- データ数が11以上ある
- データの分布が正規分布である、もしくは正規分布が仮定できる
- 等分散を仮定できない

```
>t.test(data$s25, data$s50)
```

```
Welch Two Sample t-test

data:  data$s25 and data$s50
t = -2.3566, df = 43.204, p-value = 0.02304
alternative hypothesis: true difference in means is not equal to 0
95 percent confidence interval:
 -2.1110378 -0.1642404
sample estimates:
mean of x mean of y
 10.57169  11.70933
```

# Wilcoxonの検定について

- Rでデフォルトで用意されている関数よりも、より正確な検定統計量を導くことができるwilcox.exact関数が推奨されている
- 下記1行目のコマンドでインストールを行った後、2行目でライブラリを使えるように呼び出す必要がある

```
>install.packages("exactRankTests", repos=http://cran.ism.ac.jp/)  
>library(exactRankTests)
```

- 参考:デフォルトで用意されている関数

```
>wilcox.test(data$s25,data$s50)
```

# Wilcoxonの順位和検定(U検定)

## ➤ 条件

- データに対応がある場合
  - ◆ データの種類が名義尺度あるいは順序尺度である
- データに対応がない場合
  - ◆ データの種類が名義尺度あるいは順序尺度、あるいはデータ数が10以下あるいはデータの分布が明らかに正規分布ではない

```
>wilcox.exact(data$s25,data$s50)
```

```
Exact Wilcoxon rank sum test
```

```
data: data$s25 and data$s50
```

```
W = 175, p-value = 0.03153
```

```
alternative hypothesis: true mu is not equal to 0
```

# Wilcoxonの符号付順位和検定

## ➤ 条件

- データに対応がある
- データの種類が間隔尺度あるいは比率尺度である
- データ数が10以下もしくはデータの分布が明らかに正規分布でない

```
>wilcox.exact(data$s25,data$s50,paired=TRUE)
```

```
Exact Wilcoxon signed rank test
```

```
data: data$s25 and data$s50
```

```
V = 79, p-value = 0.07487
```

```
alternative hypothesis: true mu is not equal to 0
```

# 片側検定と両側検定

➤ 第二種の過誤(帰無仮説が誤りなのに採択してしまう間違い)を少なくするために片側or両側検定を選ぶ

- 右側検定・・・平均値の値が「より大きくなる」と予測している場合
- 左側検定・・・平均値の値が「より小さくなる」と予測している場合
- 両側検定・・・平均値の値が「等しくない」と予測している場合

➤ Rでの指定方法(例: Welchのt検定)

- 右側検定( $s_{25} > s_{50}$ と予測している場合)

```
>t.test(data$s25,data$s50,alternative="greater")
```

- 左側検定( $s_{25} < s_{50}$ と予測している場合)

```
>t.test(data$s25,data$s50,alternative="less")
```

alternativeの指定を省略した場合は  
両側検定が行われる



# 片側検定と両側検定(補足)

## ➤ 具体例

総合成績上位の学生の英語の得点(rankH)	総合成績下位の学生の英語の得点(rankL)
72	27
84	58
88	81

普通に考えたら  
総合成績上位の人  
の方が得点も  
高いだろうな.....

有意差が出るかどうか **右側検定**を行う

## ➤ Rでの指定方法(例: Welchのt検定)

```
>t.test(data$rankH,data$rankL,alternative="greater")
```



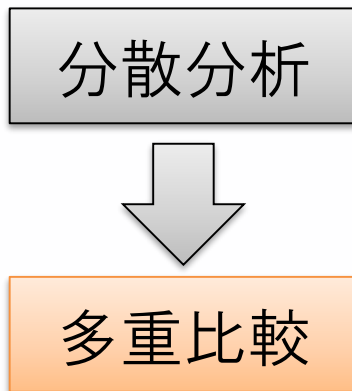
# おまけ

## ➤ 3群以上の平均値の差をみたい場合

- 分散分析(ANOVA)・・・全体として有意差があるかを調べる検定
  - ◆ 多重比較・・・個々の群と群を検定する場合に有意水準を上げずに行う検定

複数回のt検定では  
有意水準が上がってしまう！

## ➤ 検定の手順



通常は多重比較の前に分散分析を行うのが定石であるが、特定の群間に注目するならば分散分析は用いずに直接多重比較を行う

t検定と同様、様々な検定があるので詳細は参考文献のURLを参照

# 参考文献(統計学)

## ➤ 統計学がわかる(ファーストブック)



向後千春・富永敦子 著  
技術評論社

具体例をベースに初心者にもわかりやすく書かれているので統計入門書としてオススメ

## ➤ 統計嫌いのための心理統計の本:統計のキホンと統計手法の選び方

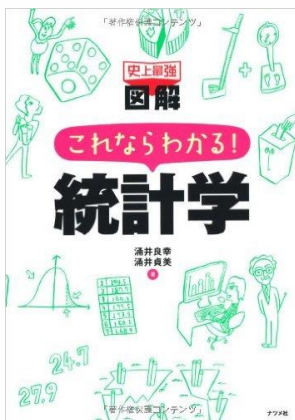


白井祐浩 著  
創元社

実験でアンケート調査したいけど数式は見たくないという人にオススメ

# 参考文献(統計学)

## ➤ 史上最強 図解 これならわかる！ 統計学



涌井良幸・涌井貞美 著  
ナツメ社

難易度は少し高いが  
数学的に理解したい人にオススメ

## ➤ 「統計学が最強の学問である」シリーズ

デジタル社会を生き抜くための武器と読者のための読書術  
Literacy for the Next Generation  
統計学が最強の学問である

10万人の命を贈られる唯一の方法  
[読者体験] 読者・著者・編集者・心理統計学  
[著者体験] エンディング・エンディング(読者体験)  
主要な分野も体系的に解説した画期的な入門書!

西内啓  
著



西内啓  
著



西内啓  
著

西内 啓 著  
ダイヤモンド社

統計学という学問を体系的に  
理解したいという人にオススメ