


九州看護福祉大学  
Kyushu University of Nursing and Social Welfare

1

## 第1回 医療データサイエンスプログラム 事前説明会

2024-12-27  
理学療法士 吉田 勇一

- 19時から開始します。しばらくお待ちください。
- カメラ、マイクはOFFにしてください。
- 参加者名簿と資料のダウンロード  
[y2pt.com/sign/](https://y2pt.com/sign/)



九州看護福祉大学  
Kyushu University of Nursing and Social Welfare

2

九州看護福祉大学 リカレント教育（専門分野）  
スタートアップ事業

### 医療データサイエンスプログラム —超入門 Rとリハビリテーション統計学—

講師 吉田勇一先生

本プログラムでは、臨床データの効果判定やリハビリテーション研究で不可欠な統計学、Rの使用法を習得することを旨とします。

- 本回は統計学の初学者向けの内容となります
- 参加者は第1回～第3回までの全ての受講が可能な方に限ります
- ノートパソコンを持参いただきます（Windows推奨）

対象者 熊本県内で勤務されているPT,OT,ST  
会場 九州看護福祉大学 2号館小講義室1

講座日時と内容

- 第1回： 令和 7年 2月 8日（土）  
13:00～16:00 ・ Rの使い方  
・ 1 標本問題
- 第2回： 令和 7年 2月 15日（土）  
13:00～16:00 ・ 2 群の比較  
・ 分散分析
- 第3回： 令和 7年 2月 22日（土）  
13:00～16:00 ・ 線形回帰モデル  
・ 線形混合効果モデル

申込方法 **※定員10名**  
以下のQRコードからお申込みください  
最終締切：令和 7年 1月 17日（金）正午  
締切後に参加可否をEメールにてご連絡致します

参加費  
スタートアップのため  
全講座無料

お問い合わせ  
九州看護福祉大学 生涯教育研究センター  
TEL 0968-75-1821

参加者名簿と資料ダウンロード [y2pt.com/sign/](https://y2pt.com/sign/)

九州看護福祉大学  
Kyushu University of Nursing and Social Welfare

3

## 統計学に関する実績

2016年 統計検定2級 合格証番号16-2-1426

2023年 医学バイオ統計学博士（テーマ：理学療法法の因果推論）

2023年 統計学備忘録の開設

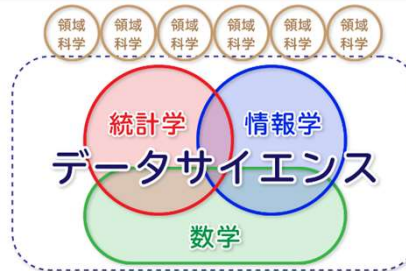
理学療法法の因果推論に関する論文  
EFFICACY OF CATEGORIES IN PHYSICAL THERAPY FOR IMPROVING  
MOTOR FUNCTION OF PATIENTS WITH STROKE. Bulletin of informatics  
and cybernetics, 55(2), 1-10, 2023

参加者名簿と資料ダウンロード [y2pt.com/sign/](https://y2pt.com/sign/)

九州看護福祉大学  
Kyushu University of Nursing and Social Welfare

4

## データサイエンスとは？



滋賀大学データサイエンス学部教授 杉本知之  
日本で最初にデータサイエンス学部を創設したのが滋賀大学（2017年）  
<https://www.stat.go.jp/dstart/point/lecture/07.html>

参加者名簿と資料ダウンロード [y2pt.com/sign/](https://y2pt.com/sign/)

九州看護福祉大学  
Kyushu University of Nursing and Social Welfare

5


### 医療データサイエンスプログラムの目的

- Rを使用してデータサイエンスとデータマイニングの基本を学ぶ
- リハビリテーション統計学の学習方法を習得する

データサイエンス  
データから有用な情報・知識を引き出すための基本原理

データマイニング  
基本原理を組み込んだ技法を活用して、データから有用な情報・知識を引き出す行為

統計学



Foster Provost (著), Tom Fawcett (著), 竹田 正和(監訳); 戦略的データサイエンス入門 ―ビジネスに活かすコンセプトとテクニック 単行本 (ソフトカバー), オライリージャパン, 2014, p2

参加者名簿と資料ダウンロード [y2pt.com/sing/](http://y2pt.com/sing/)

九州看護福祉大学  
Kyushu University of Nursing and Social Welfare

6

### 統計学備忘録

リハビリテーション統計学

統計学  
備忘録  
y2pt.com



リハビリ 統計 吉田

参加者名簿と資料ダウンロード [y2pt.com/sing/](http://y2pt.com/sing/)

九州看護福祉大学  
Kyushu University of Nursing and Social Welfare

7

### 統計学備忘録の実績

アナリティクス

すべてのアカウント > yoshida931  
統計学備忘録

ホーム

アクティブユーザー 2.1万  
↑ 193.8%

イベント数 12万  
↑ 199.9%

表示回数 3.5万  
↑ 170.0%



イベント数  
ページビューやクリック、フォームの送信、動画の再生開始、ダウンロードなど、


参加者名簿と資料ダウンロード [y2pt.com/sing/](http://y2pt.com/sing/)

九州看護福祉大学  
Kyushu University of Nursing and Social Welfare

8

### 旧・統計学備忘録にこれまで寄せられたコメント

hatena blog



6〜7年前は、数学警ら隊の方々から厳しいコメントもいただいたが、現在では、皆無となった

参加者名簿と資料ダウンロード [y2pt.com/sing/](http://y2pt.com/sing/)

九州看護福祉大学  
Kyushu University of Nursing and Social Welfare

9

## 統計学備忘録のユーザー

アクティブユーザー (国 ID)



国	アクティブ ユー
Japan	2万 ↑200...
United States	233 ↑46.5%
China	201 ↑1,91...
France	57 ↓6.6%
United Kingdom	30 ↑36.4%
Canada	23 ↑109...
Germany	22 ↑29.4%

過去 12 か月間

国を表示 →

参加者名簿と資料ダウンロード [y2pt.com/sing/](https://y2pt.com/sing/)

九州看護福祉大学  
Kyushu University of Nursing and Social Welfare

10

## リハビリテーション統計学 統計学備忘録

This site is powered by  
 エックスサーバー <https://www.xserver.ne.jp/>  
 WordPress <https://ja.wordpress.org/>  
 Cocoon <https://wp-cocoon.com/>



y2pt.com



サイト運営による利益相反はありません

参加者名簿と資料ダウンロード [y2pt.com/sing/](https://y2pt.com/sing/)

九州看護福祉大学  
Kyushu University of Nursing and Social Welfare

11

## 医療データサイエンスプログラムの内容

参加者名簿と資料ダウンロード [y2pt.com/sing/](https://y2pt.com/sing/)

九州看護福祉大学  
Kyushu University of Nursing and Social Welfare

12

## 医療データサイエンスプログラムの内容

- Rの使い方
- データの取り込み
- 記述統計  
(エクセルのような使い方を練習)
- グラフを作成してデータの可視化
- 1標本問題
  - 連続変数
  - 比率の検定
- 2群の比較
  - 連続変数
  - 比率
- 分散分析
  - 一元配置分散分析  
二元配置分散分析
  - 級内相関係数
- 線形回帰モデル
  - 一般線形  
(単回帰分析、重回帰分析)
  - 一般化線形  
(ロジスティック回帰分析)
- 線形混合効果モデル
  - クロスオーバー試験
  - 経時データの解析  
(反復測定分散分析)

参加者名簿と資料ダウンロード [y2pt.com/sing/](https://y2pt.com/sing/)

九州看護福祉大学  
Kyushu University of Nursing and Social Welfare

13

## プログラム内容は統計学備忘録に収録

統計学 備忘録 リハビリテーション統計学

HOME PROFILE

**01**  
確率変数と確率分布 / 二項分布とポアソン分布 / 正規分布 / 超幾何分布 / 一様分布 / 正規分布 / 正規性の検定 / Q-Qプロット / 裾野分布 / カイニ乗分布 / t分布 / F分布 / 平均の95%信頼区間 (母分散が未知) / 平均の95%信頼区間 (母分散が既知)

**02**  
1標本の比率の検定 / 分類表について / 比率の差の検定 / Fisherの正確検定 / カイニ乗検定 / ROC曲線

**03**  
対応のある検定 / 対応のない検定 (等分散) / フェルスの検定 / マルコフ連鎖の確率分布 / 線形スコアマッチング

**04**  
二元配置分散分析 / 二元配置分散分析 (繰り返し測定なし) / 二元配置分散分析 (繰り返し測定) / ICC(1,1) / ICC(2,1) / ICC(3,1)

**05**  
回帰分析の基本 (単回帰分析) / 重回帰分析 / 重回帰分析の検定 / 主成分分析 / ロジスティック回帰 (1) / ロジスティック回帰 (2) / ロジスティック回帰の標準変異 / 回帰診断 (VIF) / 交互作用

参加者名簿と資料ダウンロード [y2pt.com/sing/](http://y2pt.com/sing/)

九州看護福祉大学  
Kyushu University of Nursing and Social Welfare

14

## 医療データサイエンスプログラムの内容

$$f(x) = \frac{1}{\sqrt{2\pi\sigma^2}} \exp\left\{-\frac{(x-\mu)^2}{2\sigma^2}\right\}$$

$$\log\left(\frac{rate}{1-rate}\right) = \beta_0 + \beta_1 x_1 + \beta_2 x_2$$

医療データサイエンスプログラムでは、数理は必要最小限とします

参加者名簿と資料ダウンロード [y2pt.com/sing/](http://y2pt.com/sing/)

九州看護福祉大学  
Kyushu University of Nursing and Social Welfare

15

## RとAIを使いながらデータを可視化します

	A	B	C	D
1	ID	歩行練習	練習前	短縮時間
2	1	B	22.5	5.6
3	2	A	25.5	7.8
4	3	B	19.3	3.3
5	4	A	25.6	9.6
6	5	A	31.2	10.8
7	6	B	16.5	4
8	7	B	19.6	4.5
9	8	A	26.3	8.6
10	9	A	22.2	6.3

$f(x)$

rate

歩行速度

参加者名簿と資料ダウンロード [y2pt.com/sing/](http://y2pt.com/sing/)

九州看護福祉大学  
Kyushu University of Nursing and Social Welfare

16

## RとAIを使いながらデータを可視化します

短縮時間

歩行練習

短縮時間

練習前の10m歩行時間

参加者名簿と資料ダウンロード [y2pt.com/sing/](http://y2pt.com/sing/)

九州看護福祉大学  
Kyushu University of Nursing and Social Welfare

17

RとAIを使いながら必要に応じて、解析します

```
R
t.test(短縮時間 ~ 歩行練習, data=dat)
```

参加者名簿と資料ダウンロード [y2pt.com/sing/](https://y2pt.com/sing/)

九州看護福祉大学  
Kyushu University of Nursing and Social Welfare

18

RとAIを使いながら必要に応じて、解析します

```
R
fit <- lm(短縮時間 ~ 歩行練習 + 練習前, data=dat)
summary(fit)
confint(fit)
```

参加者名簿と資料ダウンロード [y2pt.com/sing/](https://y2pt.com/sing/)

九州看護福祉大学  
Kyushu University of Nursing and Social Welfare

19

医療データサイエンスプログラムではRを使用します

<https://www.r-project.org/>

参加者名簿と資料ダウンロード [y2pt.com/sing/](https://y2pt.com/sing/)

九州看護福祉大学  
Kyushu University of Nursing and Social Welfare

20

統計ソフトについて

- SPSS** (GUI: グラフィカルユーザーインターフェース)
  - 簡単に操作可能で、多くのユーザーがいる
  - 書籍も多く出版されている
  - ただし有料 サブスク月額99米ドル (14,000円) ~
  - 永久ライセンス 3,610米ドル (530,000) ~
- SAS** 治験など生物統計学者が使用している高額ソフト。プログラムの習得必要。
- STATA** GUIで操作が簡単。世界のトップジャーナルで散見される。サブスク80,000/年~。
- R** 無料統計ソフト、オープンソースなので世界中の統計家により日々更新されている
  - EZRやRコマンドー は、Rを使いやすくしたもの (SPSSのようなGUI)

どの統計ソフトを使えば良いのか？  
無料かつ世界のトップジャーナルに掲載されている論文で使用されているソフトが良い

参加者名簿と資料ダウンロード [y2pt.com/sing/](https://y2pt.com/sing/)

九州看護福祉大学  
21

### 世界五大医学雑誌

- The Lancet - エルゼビア発行 (IF=202.731)
- New England Journal of Medicine - 米国マサチューセッツ内科外科学会が発行(IF=176.079)
- JAMA (Journal of the American Medical Association) - 米国医師会が発行(IF=157.335)
- BMJ (British Medical Journal) - 英国医師会が発行(IF=93.333)
- Annals of Internal Medicine - 米国内科学会が発行(IF=51.598)

参加者名簿と資料ダウンロード [y2pt.com/sing/](http://y2pt.com/sing/)

九州看護福祉大学  
22

### Lancet 2024; 404: 134-44

 Effectiveness and cost-effectiveness of an individualised, progressive walking and education intervention for the prevention of low back pain recurrence in Australia (WalkBack): a randomised controlled trial

 Natasha C Pocovi, Chung-Wei Christine Lin, Simon D French, Petra L Graham, Johanna M van Dongen, Jane Latimer, Dafna Merom, Anne Tiedemann, Christopher G Maher, Ornella Clavisi, Shuk Yin Kate Tong, Mark J Hancock

オーストラリアの理学療法士らの研究グループは、個別に段階的に歩行練習と理学療法士が教育介入することが、腰痛の再発予防に効果的であり、腰痛予防に関わる費用対効果がよいことを報告した。

independent data monitoring committee was not used due to the low-risk nature of the intervention. All analyses (clinical effectiveness and cost-effectiveness) were performed using R (version 4.2.2).

参加者名簿と資料ダウンロード [y2pt.com/sing/](http://y2pt.com/sing/)

九州看護福祉大学  
23

The NEW ENGLAND JOURNAL of MEDICINE

ORIGINAL ARTICLE

### Dapagliflozin in Patients with Heart Failure and Reduced Ejection Fraction

John J.V. McMurray, M.D., Scott D. Solomon, M.D., Silvio E. Inzucchi, M.D., Lars Køber, M.D., D.M.Sc., et al., for the DAPA-HF Trial Committees and Investigators\*

We analyzed the total symptom score on the Kansas City Cardiomyopathy Questionnaire as a composite, rank-based outcome, incorporating patient vital status at 8 months along with a change in score from baseline to 8 months in surviving patients, using the rank analysis of covariance method, with a corresponding win ratio used to estimate the magnitude of treatment effect.<sup>17</sup> We assessed the consistency of the treatment effect among 14 prespecified subgroups. The safety analyses were performed in patients who had undergone randomization and received at least one dose of dapagliflozin or placebo. We used Fisher's exact test to compare the incidence of adverse events. All the analyses were performed with the use of Stata software, version 15 (StataCorp) and [R, version 3.5.1 \(R Foundation for Statistical Computing\)](https://www.foundry.com/).

参加者名簿と資料ダウンロード [y2pt.com/sing/](http://y2pt.com/sing/)

九州看護福祉大学  
24

PTJ Physical Therapy & Rehabilitation Journal IF: 3.5

PTJ Physical Therapy & Rehabilitation Journal | Physical Therapy, 2023,103:1-10  
<https://doi.org/10.1080/00140139.2023.2241102>  
Advance access publication date October 12, 2022  
Original Research

### The Role of Exercise Self-Efficacy in Exercise Participation Among Women With Persistent Fatigue After Breast Cancer: A Mixed-Methods Study

Stephen Wechsler, PT, DPT<sup>1,2</sup>, Mei R. Fu, PhD, RN, FAAN<sup>3</sup>, Kathleen Lyons, ScD, OTR/L<sup>1</sup>, Kelley C. Wood, PhD, ACSM-CEP, CET<sup>3</sup>, Lisa J. Wood Magee, PhD, RN, FAAN<sup>4</sup>

#### Data Analysis

Quantitative analyses were conducted in R.<sup>63</sup> Frequency distributions, and descriptive statistics were examined for all study variables. All data met assumptions and therefore did not require transformation. To improve interpretability of self-reported CRE, SF-36-VS scores were centered and reversed so that higher values represent higher fatigue levels. To improve interpretability and ability to compare ESE scores, percentages were calculated by dividing participants' self-efficacy for exercise scale scores by the max score of 90. Correlation coefficients were examined for all independent

参加者名簿と資料ダウンロード [y2pt.com/sing/](http://y2pt.com/sing/)

九州看護福祉大学 25

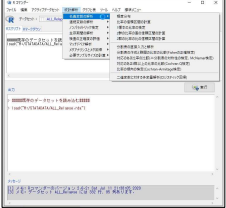
## Rを使うなら、コーディングとGUIのどちらがよいか？

コーディングによるR操作  
プログラミング

```
# データの準備
xa <- c(75, 87, 89, 88, 84, 81, 88, 83, 88, 88, 82,
72, 74, 93, 77, 67, 88, 84, 68, 84, 80, 78, 75, 71,
82, 74, 84, 77, 79, 76, 83, 75, 86, 76, 80, 76, 68,
72, 75, 85)
xb <- c(64, 77, 79, 73, 89, 82, 59, 85, 80, 75, 65,
79, 65, 74, 73, 72, 69, 83, 98, 73, 88, 59, 62, 88,
64, 74, 81, 79, 69, 87, 83, 87, 72, 71, 72, 78, 78,
82, 72, 71)
# ヒストグラムのプロット
par(mfrow=c(1, 2)) # グラフを横につづる設定
hist(xa, main="Histogram of xa", xlab="Values",
col="blue", border="black")
hist(xb, main="Histogram of xb", xlab="Values",
col="red", border="black")
```

入力が見えにくい

GUIによるR操作  
EZR, Rコマンダー



操作が簡単

- グラフィカルユーザーインターフェイス (GUI)

参加者名簿と資料ダウンロード [y2pt.com/sing/](https://y2pt.com/sing/)

九州看護福祉大学 26

## 医療データサイエンスプログラムではコーディングを使用

### コーディングの利点

- コードのコピー&ペーストが可能
  - webサイトのコードを利用
  - AIが生成したコードを利用
- 解析経過を一括保存
  - 解析の再現性
- コードの共有可能
  - 解析経過を共有できる

```
# データの準備
xa <- c(75, 87, 89, 88, 84, 81, 88, 83, 88, 88, 82,
72, 74, 93, 77, 67, 88, 84, 68, 84, 80, 78, 75, 71,
82, 74, 84, 77, 79, 76, 83, 75, 86, 76, 80, 76, 68,
72, 75, 85)
xb <- c(64, 77, 79, 73, 89, 82, 59, 85, 80, 75, 65,
79, 65, 74, 73, 72, 69, 83, 98, 73, 88, 59, 62, 88,
64, 74, 81, 79, 69, 87, 83, 87, 72, 71, 72, 78, 78,
82, 72, 71)
# ヒストグラムのプロット
par(mfrow=c(1, 2)) # グラフを横につづる設定
hist(xa, main="Histogram of xa", xlab="Values",
col="blue", border="black")
hist(xb, main="Histogram of xb", xlab="Values",
col="red", border="black")
```

参加者名簿と資料ダウンロード [y2pt.com/sing/](https://y2pt.com/sing/)

九州看護福祉大学 27

## Rのプログラム

プログラム作成は、AIが助けてくれるようになった

### 生成AIはプログラミングに強い！

【2024年最新】生成AIサービス10選を徹底比較				
種類	サービス名	特徴	料金	日本語対応
文章	Chat GPT	・ OpenAIが開発 ・ 世界中で最も普及	無料～月額20ドル	○
	Copilot	・ Microsoft+OpenAIが ・ 実装	無料～月額3750円	○
	Gemini	・ Googleが開発 ・ マルチモーダル機能	無料～月額2800円	○

AI総研 <https://metaversesouken.com/ai/>

参加者名簿と資料ダウンロード [y2pt.com/sing/](https://y2pt.com/sing/)

九州看護福祉大学 28

## AIを使用した統計解析

例) Copilotへのプロンプト

1, 2, 3, 5, 7 と 3, 3, 4, 9, 12を比較する図を作成して、2群の差を検定するためのRのプログラムを教えてください。

参加者名簿と資料ダウンロード [y2pt.com/sing/](https://y2pt.com/sing/)



九州看護福祉大学  
29

### Copilotの回答

次に、2群の差を検定するためのRのプログラムを紹介し、このプログラムでは、t検定を使用して2つのグループの平均値の差を検定します。

```

# データの設定
group1 <- c(1, 2, 3, 5, 7)
group2 <- c(3, 3, 4, 9, 12)

# 箱ひげ図の作成
boxplot(group1, group2, names = c("Group 1", "Group 2"), main = "Comparison of Two Groups",
        ylab = "Values")

# t検定の実行
t_test_result <- t.test(group1, group2)

# 結果の表示
print(t_test_result)
    
```

このプログラムを実行すると、箱ひげ図が表示され、t検定の結果が出力されます。t検定の結果には、t値とp値が含まれます。これにより、2つのグループの平均値に有意な差があるかどうかを判断できます。

何か他にお手伝いできることがあれば教えてください。

参加者名簿と資料ダウンロード [y2pt.com/sing/](https://y2pt.com/sing/)

九州看護福祉大学  
30

### Copilotの回答をRにペースト

```

> # 箱ひげ図の作成
> boxplot(group1, group2, names = c("Group 1", "Group 2"), mai
> # t検定の実行
> t_test_result <- t.test(group1, group2)
> # 結果の表示
> print(t_test_result)

Welch Two Sample t-test

data:  group1 and group2
t = -1.2257, df = 6.4794, p-value = 0.263
alternative hypothesis: true difference in means is not equal to 0
95 percent confidence interval:
 -7.698999  2.498999
sample estimates:
mean of x mean of y
      3.6      6.2
    
```

参加者名簿と資料ダウンロード [y2pt.com/sing/](https://y2pt.com/sing/)

九州看護福祉大学  
31

### 同じ回答が出て意味が分からない → Copilotに質問します

Rの解析結果です。結果について詳しく教えてください。統計学は不得意です。

Welch Two Sample t-test

data: value by group  
t = -1.2257, df = 6.4794, p-value = 0.263  
alternative hypothesis: true difference in means between group  
group1 and group group2 is not equal to 0  
95 percent confidence interval:  
-7.698999 2.498999  
sample estimates:  
mean in group group1 mean in group group2  
3.6 6.2

参加者名簿と資料ダウンロード [y2pt.com/sing/](https://y2pt.com/sing/)

九州看護福祉大学  
32

### Copilotの解説

**結果の全体像**  
この結果は、2つのグループ (group1とgroup2) の平均値に差があるかどうかを調べるために、Welchのt検定を行ったものです。

**各項目の意味**  

- t = -1.2257: t統計量と呼ばれる値です。この値が大きいほど、2つのグループの平均値に差がある可能性が高いことを示します。
- df = 6.4794: 自由度です。統計学的計算で用いられる値で、サンプルサイズやグループ数に関連します。
- p-value = 0.263: p値です。この値が0.05より大きい場合、通常は2つのグループの平均値に有意な差がないと判断されます。
- alternative hypothesis: 対立仮説です。「2つのグループの平均値に差がある」という仮説を指します。
- 95 percent confidence interval: 95%信頼区間です。真の平均値の差が、この区間内に含まれる確率が95%であることを示します。
- sample estimates: 各グループの平均値です。

**結果の解釈**  
今回の結果では、p値が0.263と、有意水準である0.05よりも大きくなっています。そのため、**2つのグループの平均値に有意な差はない**という結論になります。

**まとめ**  
今回の検定の結果から、group1とgroup2のデータに有意な差は見られませんでした。つまり、両グループの平均値はほぼ同じであると考えられます。

**次に詳しく知りたい方へ**  

- Welchの検定: 2つの標本が異なる分散を持つ場合に用いられる検定です。
- t分布: t統計量が従う確率分布です。
- 有意水準: どの程度の確率で誤った判断を許容するかを決める基準です。
- 仮説: この結果についてさらに詳しい説明が必要な場合は、お気軽にご質問ください。
- 可視化: データをグラフなどで可視化することで、より直感的に結果を理解することができます。例えば、箱ひげ図やヒストグラムなどを用いて、2つのグループの分布を比較してみましょう。
- 他の統計手法: データの特性や分析目的によっては、t検定以外の統計手法が適している場合もあります。
- データの前提条件: データが正規分布を仮定している場合、分布が歪んでいる場合には、適切な前処理を行う必要があります。


これらの点についても、ご希望があれば詳しく説明します。

参加者名簿と資料ダウンロード [y2pt.com/sing/](https://y2pt.com/sing/)




九州看護福祉大学  
33

### (余談) RとPythonの比較



Rは統計学が得意

**目的 = 『説明』**  
原因と結果の因果関係についてじっくり考える。予測というより、入力と出力の関係性を考える場合。何が結果に影響を与えているか？



Pythonは機械学習や開発が得意

**目的 = 『予測』**  
結果ベースで考えていく（とにかくこの画像を判明すればOK）。機械がパターンやルールを見つけ出す。ブラックボックスになっていても予測精度がよければOK。

\*どちらかが使えれば、他方も使えるようになる！

参加者名簿と資料ダウンロード [y2pt.com/sing/](https://y2pt.com/sing/)

九州看護福祉大学  
34

### (余談) RとPythonの比較

**R**

```

# データの準備
xa <- c(75, 87, 89, 80, 84, 81, 88, 83, 88, 88, 82, 72, 74, 93, 77, 67, 88, 84, 68, 86, 80, 78, 75, 71, 82, 74, 84, 77, 72, 76, 83, 75, 80, 76, 80, 68, 72, 75, 85)
xb <- c(64, 77, 79, 73, 89, 82, 59, 85, 80, 75, 65, 79, 65, 74, 73, 72, 69, 83, 90, 73, 88, 59, 62, 80, 63, 74, 83, 70, 69, 67, 81, 67, 72, 71, 72, 78, 79, 82, 72, 71)

# ヒストグラムのプロット
par(mfrow=c(2,1)) # グラフを横並びで表示
hist(xa, main="Histogram of xa", xlab="Values", col="blue", border="black")
hist(xb, main="Histogram of xb", xlab="Values", col="red", border="black")
        
```

**Python**

```

# python -
1 import pandas as pd
2 import matplotlib.pyplot as plt
3
4 xa = pd.Series([75, 87, 89, 80, 84, 81, 88, 83, 88, 88, 82, 72, 74, 93, 77, 67, 88, 84, 68, 84, 80, 78, 75, 71, 82, 74, 84, 77, 72, 76, 83, 75, 80, 76, 80, 76, 68, 72, 75, 85])
5 xb = pd.Series([64, 77, 79, 73, 89, 82, 59, 85, 80, 75, 65, 79, 65, 74, 73, 72, 69, 83, 90, 73, 88, 59, 62, 80, 63, 74, 83, 70, 69, 67, 81, 67, 72, 71, 72, 78, 79, 82, 72, 71])
6
7 fig = plt.figure(dpi=96)
8 ax1 = fig.add_subplot(211)
9 ax1.set_xlim(55, 100)
10 ax1.set_ylim(0, 15)
11 ax1.hist(xa, bins=9, range=(55,100), color='r')
12
13 ax2 = fig.add_subplot(212)
14 ax2.set_xlim(55, 100)
15 ax2.set_ylim(0, 15)
16 ax2.hist(xb, bins=9, range=(55,100), color='b')
17
18 plt.show()
19
        
```

参加者名簿と資料ダウンロード [y2pt.com/sing/](https://y2pt.com/sing/)

九州看護福祉大学  
35


### なぜリハビリテーション統計学が必要なのか？


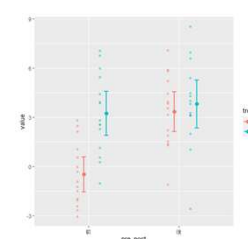
参加者名簿と資料ダウンロード [y2pt.com/sing/](https://y2pt.com/sing/)

九州看護福祉大学  
36

### リハビリテーション統計学の定義

リハビリテーションの効果を説明するための統計学



参加者名簿と資料ダウンロード [y2pt.com/sing/](https://y2pt.com/sing/)

九州看護福祉大学  
37

### 研究デザイン

新治療群 vs 従来治療群

RCT      観察研究

観察研究 { コホート研究  
                 ケース・コントロール研究  
                 横断研究

参加者は無作為に新治療群または従来治療群に割り当てられる。新治療法が一方の群に適用され、もう一方の群は従来の治療法またはプラセボを受けます。介入後のアウトカム（効果、副作用など）を両群で比較し、新治療法の効果を評価。

参加者が自然に受けている治療を観察。新治療群と従来治療群を時間をかけて追跡し、アウトカムのデータを収集。ランダム化がないため、潜在的な偏りや交絡因子の影響を完全に排除することはできません。

できるだけ交絡因子を取り除いた解析が望まれる

参加者名簿と資料ダウンロード [y2pt.com/sing/](https://y2pt.com/sing/)

九州看護福祉大学  
38

### RCTの場合

Association is causation.

新治療群 vs. 従来治療群

RCTであれば t 検定でOK  
クロスオーバー試験もあり

参加者名簿と資料ダウンロード [y2pt.com/sing/](https://y2pt.com/sing/)

九州看護福祉大学  
39

### コホート研究の場合

Association is not causation.

新治療群 vs. 従来治療群

因果 vs. 関連

交絡因子を調整した解析

Causation vs. Association

因果 vs. 関連

Miguel A. Hernán, James M. Robins ; Causal Inference: What If. April 26, 2024, <https://www.hsph.harvard.edu/miguelhernan/causal-inference-book/>

参加者名簿と資料ダウンロード [y2pt.com/sing/](https://y2pt.com/sing/)

九州看護福祉大学  
40

### シンプソンのパラドクス

コホート研究の場合

新治療群 vs. 従来治療群

差がなかった ( $p > 0.05$ )

ベースラインの10m歩行時 (秒)

- 15秒未満 差がなかった ( $p > 0.05$ )
- 15秒以上 差が認められた ( $p < 0.05$ )

アウトカム = 10m歩行時間 (秒)

参加者名簿と資料ダウンロード [y2pt.com/sing/](https://y2pt.com/sing/)

九州看護福祉大学  
41

### コホート研究の因果推論

できるだけ交絡因子を取り除いた解析が望まれる

新治療群 vs. 従来治療群

因果 vs. 関連

ベースラインの10m歩行時間 15秒未満  
有意差なし ( $p > 0.05$ )

ベースラインの10m歩行時間 15秒以上  
有意差あり ( $p < 0.05$ )

関連から「新治療群と従来治療群の効果に差がなかった」とは言えない!

参加者名簿と資料ダウンロード [y2pt.com/sing/](http://y2pt.com/sing/)

九州看護福祉大学  
42

### リハビリテーション統計学の理解

RCTの場合

コホート研究の場合

● 新治療群  
○ 従来治療群

短縮時間

ベースラインの10m歩行時間 (秒)

参加者名簿と資料ダウンロード [y2pt.com/sing/](http://y2pt.com/sing/)

九州看護福祉大学  
43

### リハビリテーション統計学の理解

コホート研究の場合

● 新治療群  
○ 従来治療群

10m歩行時間の短縮時間

ベースラインの10m歩行時間 (秒)

参加者名簿と資料ダウンロード [y2pt.com/sing/](http://y2pt.com/sing/)

九州看護福祉大学  
44

### 因果ダイアグラム (DAG : Directed Acyclic Graph)

DAG : 日本語の正式名称は「有向非巡回グラフ」

交絡因子

ベースラインの10m歩行時間

原因 (治療内容)  
新治療群  
従来治療群

結果 (アウトカム)  
10m歩行時間の短縮

観察研究では、交絡因子が影響している可能性が高い

交絡因子の影響を取り除いて (調整して) 解析しなければならない

岩波データサイエンス Vol.3, [特集]因果推論

参加者名簿と資料ダウンロード [y2pt.com/sing/](http://y2pt.com/sing/)

九州看護福祉大学  
45

### 交絡調整 ①

層別解析 交絡因子 = ベースラインの10m歩行時間

10m歩行時間短縮

ベースラインの10m歩行時間が15秒未満 新治療群 vs 従来治療群

ベースラインの10m歩行時間が15秒以上 新治療群 vs 従来治療群

参加者名簿と資料ダウンロード [y2pt.com/sing/](https://y2pt.com/sing/)

九州看護福祉大学  
46

### 交絡調整 ②

回帰分析 交絡因子 = ベースラインの10m歩行時間

コホート研究の場合

10m歩行時間の短縮時間

ベースラインの10m歩行時間 (秒)

参加者名簿と資料ダウンロード [y2pt.com/sing/](https://y2pt.com/sing/)

九州看護福祉大学  
47

### 交絡調整 ③

傾向スコアマッチング 交絡因子 = ベースラインの10m歩行時間

傾向スコア：対象者の背景因子に基づいて算出されるスコア（確率）

新治療群と従来治療群から、傾向スコアが類似している対象者を選び出し、両群を比較することで、練習Aと練習Bの効果の差異を検証する方法です。

新治療群 vs 従来治療群

参加者名簿と資料ダウンロード [y2pt.com/sing/](https://y2pt.com/sing/)

九州看護福祉大学  
48

### リハビリテーション統計学の理解

交絡調整の方法

- 層別解析
- 回帰分析
- 傾向スコア
- ・・・その他

これらの手法を実装するために、リハビリテーション統計学を学習しなければならない

参加者名簿と資料ダウンロード [y2pt.com/sing/](https://y2pt.com/sing/)

九州看護福祉大学  
Kyushu University of Nursing and Social Welfare 49

## 観察研究のためのSTROBE声明

参加者名簿と資料ダウンロード [y2pt.com/sing/](http://y2pt.com/sing/)

九州看護福祉大学  
Kyushu University of Nursing and Social Welfare 50

### RCTのガイドライン

- CONSORTガイドライン (Consolidated Standards of Reporting Trials)
- SPIRITガイドライン (Standard Protocol Items: Recommendations for Interventional Trials)
- PRISMAガイドライン (Preferred Reporting Items for Systematic Reviews and Meta-Analyses)
- ICH-GCP (International Conference on Harmonisation - Good Clinical Practice)

参加者名簿と資料ダウンロード [y2pt.com/sing/](http://y2pt.com/sing/)

九州看護福祉大学  
Kyushu University of Nursing and Social Welfare 51

### 観察研究のガイドライン STROBE声明

多くの医学研究は観察研究である。観察研究には報告の質が不十分なものが散見され、研究の強み (strength) や弱み (weakness)、および、一般化可能性 (generalizability) の評価が妨げられている。そこで、方法論の専門家 (methodologist) や雑誌編集者が集い、経験上のエビデンスや理論を考慮した上で、STROBE (Strengthen the Reporting of Observational Studies in Epidemiology) 推奨 (recommendations) を作成した。これは観察研究の報告の質を改善するためのものである。

STROBE は、記述的な観察研究や、曝露と健康に関するアウトカム (outcomes) との関連を検討する研究の報告において、一般的に含めるべき内容について言及している。またSTROBEでは、観察研究における3つの主要な研究デザイン (study design) である、**コホート研究 (cohort study)**、**ケース・コントロール研究 (case-control study)**、**横断研究 (cross-sectional study)** について言及している。

<http://www.strobe-statement.org/>

参加者名簿と資料ダウンロード [y2pt.com/sing/](http://y2pt.com/sing/)

九州看護福祉大学  
Kyushu University of Nursing and Social Welfare 52

### STROBE声明：観察研究の報告に関するガイドライン

記述的データ [descriptive data]	14*	(a) 参加者の特徴 (例：人口統計学的、臨床的、社会的特徴) と曝露や潜在的交絡因子の情報を示す。 (b) それぞれの変数について、データが欠損した参加者数を記載する。 (c) コホート研究：追跡期間を平均および合計で要約する。
統計・分析方法 [statistical method]	12	(a) 交絡因子の調整に用いた方法を含め、すべての統計学的方法を示す。 (b) サブグループと相互作用 [interaction] の検証に用いたすべての方法を示す。 (c) 欠損データ [missing data] をどのように扱ったかを説明する。 (d) ・コホート研究：該当する場合は、脱落例 [loss to follow-up] をどのように扱ったかを説明する。 ・ケース・コントロール研究：該当する場合は、ケースとコントロールのマッチングをどのように行ったかを説明する。 ・横断研究：該当する場合は、サンプリング方式 [sampling strategy] を考慮した分析法について記述する。 (e) あらゆる感度分析 [sensitivity analysis] の方法を示す。

参加者名簿と資料ダウンロード [y2pt.com/sing/](http://y2pt.com/sing/)

九州看護福祉大学  
Kyushu University of Nursing and Social Welfare

53

## STROBE声明：観察研究の報告に関するガイドライン

おもな結果[main result]

16 (a) 調整前[unadjusted]の推定値と、該当する場合は交絡因子での調整後の推定値、そしてそれらの精度(例：95%信頼区間)を記述する。どの交絡因子が、なぜ調整されたかを明確にする。

(b) 連続変数[continuous variable]がカテゴリー化されているときは、カテゴリー境界[category boundary]を報告する。

(c) 意味のある[relevant]場合は、相対リスク[relative risk]を、意味をもつ期間の絶対リスク[absolute risk]に換算することを考慮する。

参加者名簿と資料ダウンロード [y2pt.com/sing/](http://y2pt.com/sing/)

九州看護福祉大学  
Kyushu University of Nursing and Social Welfare

54

補足

**JSSC 統計検定**  
Japan Statistical Society Certificate

日本統計学会公式認定

総務省後援  
文部科学省後援  
経済産業省後援  
内閣府後援  
厚生労働省後援

日本統計学会：1948年設立  
日本理学療法士協会：1965年設立

参加者名簿と資料ダウンロード [y2pt.com/sing/](http://y2pt.com/sing/)

九州看護福祉大学  
Kyushu University of Nursing and Social Welfare

55

## 統計検定2級に向けて使用したテキスト

参加者名簿と資料ダウンロード [y2pt.com/sing/](http://y2pt.com/sing/)

九州看護福祉大学  
Kyushu University of Nursing and Social Welfare

56

## 超お勧めのテキスト

- 量的研究におけるリサーチエッセンスの立て方
- 代表的な臨床研究デザインとその特徴
- データの分布と1変数の要約
- 2つの変数の関係と要約
- 推測の基礎
- 検定とp値
- 交絡バイアスに対処するための方法
- 回帰モデル
- イベント発症リスクに対する回帰モデル
- 発症や治癒までの期間を考慮する
- 回帰モデリング
- データ数に関する議論
- 統計解析ソフトRによる図表の作成と統計解析
- 既存データの利用
- メタアナリシス

参加者名簿と資料ダウンロード [y2pt.com/sing/](http://y2pt.com/sing/)

九州看護福祉大学  
Kyushu University of Nursing and Social Welfare

57

## フローレンス・ナイチンゲール

神の考えを理解するには、統計学を学ばなければならない。  
なぜならそれらは神の目的の尺度であるからだ。



「近代看護教育の生みの親」とも呼ばれるイギリスの看護師フローレンス＝ナイチンゲール（1820年-1910年）イギリス政府によって看護師団のリーダーとしてクリミア戦争に派遣されたと野戦病院で骨身を削って看護活動に励んだ。統計に関する知識を存分に使ってイギリス軍の戦死者・傷病者に関する膨大なデータ分析から、彼らの多くが戦闘で受けた傷そのものではなく、傷を負った後の治療や病院の衛生状態が十分でないことが原因で死亡したことを明らかにし、病院内の衛生状況を改善することで傷病兵の死亡率を劇的に引き下げました。

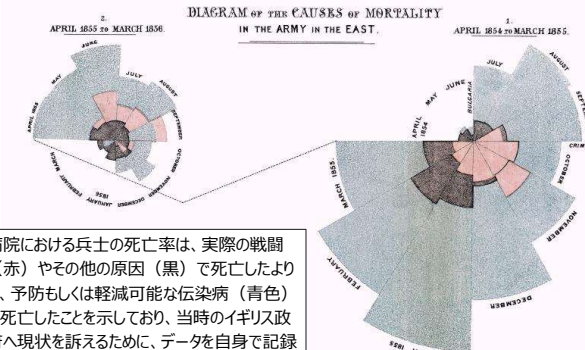
総務省統計局 [https://www.stat.go.jp/naruhodo/15\\_episode/episode/nightingale.html](https://www.stat.go.jp/naruhodo/15_episode/episode/nightingale.html)

参加者名簿と資料ダウンロード [y2pt.com/sing/](https://y2pt.com/sing/)

九州看護福祉大学  
Kyushu University of Nursing and Social Welfare

58

## フローレンス・ナイチンゲールが描いたcoxcomb



病院における兵士の死亡率は、実際の戦闘（赤）やその他の原因（黒）で死亡したよりも、予防もしくは軽減可能な伝染病（青色）で死亡したことを示しており、当時のイギリス政府へ現状を訴えるために、データを自身で記録し、のちに制作したレポートに挿入されました。

<https://visualizing.jp/nightingale-chart/>

参加者名簿と資料ダウンロード [y2pt.com/sing/](https://y2pt.com/sing/)

九州看護福祉大学  
Kyushu University of Nursing and Social Welfare

59

## ご清聴ありがとうございました

九州看護福祉大学 リカレント教育（専門分野）スタートアップ講座

### 医療データサイエンスプログラム

—超入門 Rとリハビリテーション統計学—

講師 吉田勇一先生

本プログラムでは、臨床データの収集・分析やRの活用方法を習得することを目的とします。

- 今回は統計学の初歩者向けの内容となります
- 参加費は無料ですが、教材費や午餐代等は別途お支払いとなります
- ノートパソコンを持参いただきます。（Windows推奨）

対象者 熊本県内で勤務されているPT,OT,ST  
会場 九州看護福祉大学 2号館小講義室1

講座日時と内容

- 第1回：令和7年2月8日（土）  
13:00～16:00 ・Rの使い方  
・1種本問題
- 第2回：令和7年2月15日（土）  
13:00～16:00 ・2群の比較  
・分散分析
- 第3回：令和7年2月22日（土）  
13:00～16:00 ・線形回帰モデル  
・線形混合効果モデル

申込方法 ※定員10名  
以下のQRコードからお申込みください  
最終締切：令和7年1月17日（金）正午  
締切後に参加可否をEメールにてご連絡致します

参加費  
スタートアップのため  
全講座無料

お問い合わせ  
九州看護福祉大学 生涯教育研究センター  
TEL 0968-75-1821

質問  
[yyoshida@kyushu-ns.ac.jp](mailto:yyoshida@kyushu-ns.ac.jp)

本日の参加者は名簿のご記入ください

参加者名簿と資料ダウンロード [y2pt.com/sing/](https://y2pt.com/sing/)