



## B-3 統計方法論

(2015年4月)

公益社団法人 日本理学療法士協会  
生涯学習課

1

## 学習目標

- 理学療法における実験研究・調査研究のデータ解析方法について学ぶ
- 具体的な統計的解析手法の手順を理解する
  - ▶ 研究計画やデータの取り方に関する知識は省略し、統計的解析について述べる

2

## データ解析のための基礎

3

## 重要となる基礎事項

- ① データ尺度
  - ▶ 名義尺度, 順序尺度, 間隔尺度, 比率尺度
- ② 分布
  - ▶ 正規分布か, それ以外か
- ③ **特性値の決定**
  - ▶ 代表値 (中心的な値)
  - ▶ 散布度 (ばらつき)



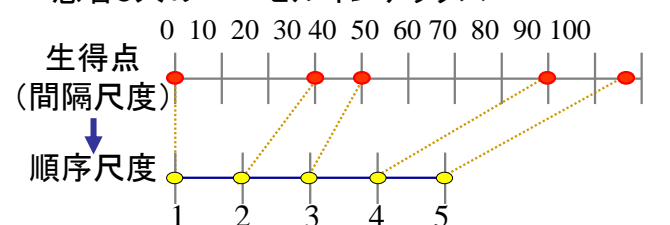
## ①データの尺度

- 比率尺度 (比尺度ともいう)
  - 筋力, 身長, 体重など
- 間隔尺度
  - BI, HDS-R, VASなど
- 順序尺度
  - MMTなど
- 名義尺度
  - 性別や疾患分類など



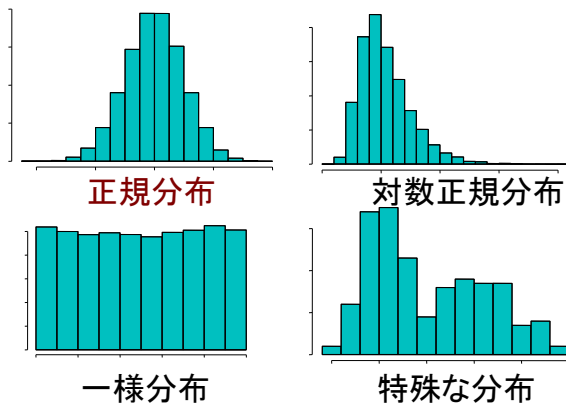
5

## 順序尺度？間隔尺度？

- 患者5人のバーセルインデックス
- 
- BIの生得点を順序尺度のデータとして考えるときは, 1から順に順位を付ける
  - 生得点の間隔の差を考慮したいときは間隔尺度のデータとなる

6

## ②分布の判断: ヒストグラム



7

## 正規分布を判断する方法

- ヒストグラムの観察
  - 主観的な判断
  - $n < 30$ ぐらいだと判断は不可能に近い
- シャピロ・ウィルク (Shapiro-Wilk) 検定
  - $p \geq 0.05$ で「正規分布に従わないとはいえない」=「正規分布する」と判断する
  - 客観的な判断となる

8

## 重要となる基礎事項

### ①データ尺度

- 名義尺度, 順序尺度, 間隔尺度, 比率尺度

### ②分布の判断

- 正規分布か, それ以外か

### ③特性値の決定

- 代表値(平均, 中央値)
- 散布度(標準偏差)



## ③特性値: 代表値

データ例: 10, 20, 15, 30, 30

### ● 平均(mean)

$$(10+20+15+30+30) \div 5 = 21$$

### ● 中央値(median)

中央(50%)の値 10, 15, 20, 30, 30

### ● 最頻値(mode)

最も度数の多い値 10, 20, 15, 30, 30

## 散布度(データのばらつき)

データ例: 10, 20, 15, 30, 30

### ● 分散

$$\{(10-21)^2 + (20-21)^2 + \dots + (30-21)^2\} / (5-1) = 80$$

### ● 標準偏差(SD, sd)

$$\sqrt{80} \doteq 8.944$$

論文では平均±標準偏差で表すことが多い

### ● 四分位範囲: 25%の値-75%の値

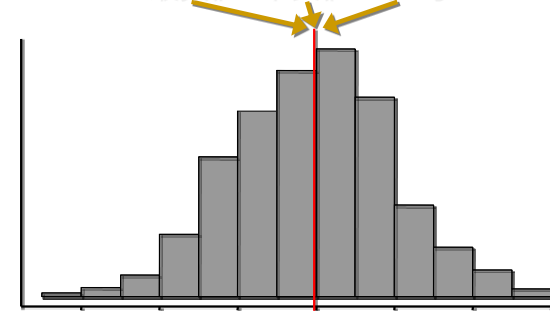
$$30 - 15 = 20$$

11

## 分布と代表値の関係

～正規分布と代表値～

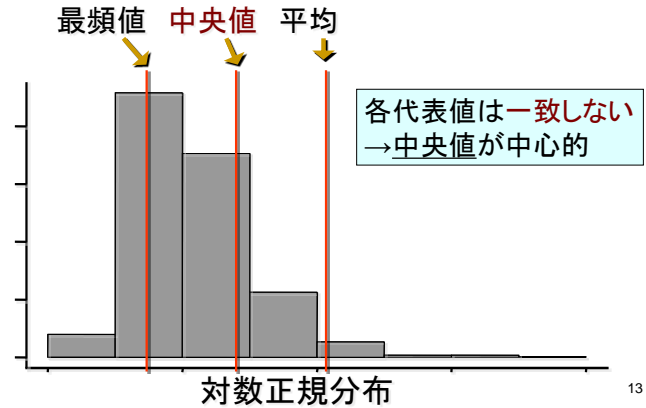
最頻値 中央値 平均



間隔・比率尺度のデータかつ, 正規分布

12

## 正規分布以外での代表値



## 特性値の使い分け

	正規分布	
	する	しない
名義尺度	分布不明: 最頻値	
順序尺度	分布不明: 中央値	
間隔・比尺度	平均 SD	

14

## 基礎事項のまとめ

使える特性値(代表値, 散布度)を決める

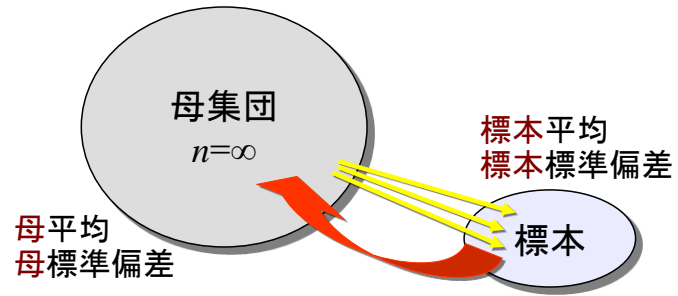
- データの尺度を判断する
  - 比・間隔尺度, 順序尺度, 名義尺度の何れか
- 正規分布するかないかを判断
  - シャピロ・ウィルク検定を利用する

15

## 統計的検定について

16

## 母集団と標本



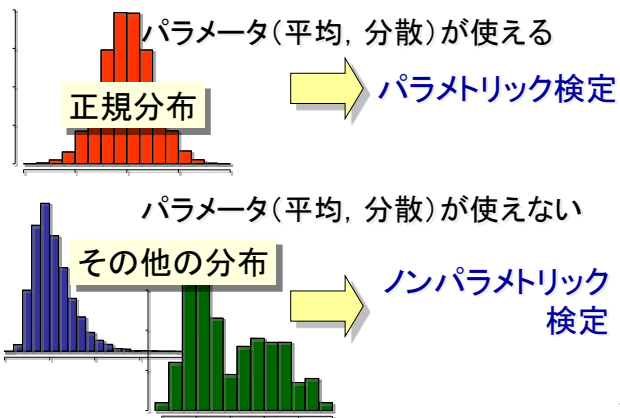
- ↓ 標本平均・標準偏差 → 母平均・標準偏差を推定
- ↓ 統計学的な理論(信頼区間など)を用いて, 推定

## 統計的検定とは

- 標本の特性値(平均・標準偏差)から母集団の特性値を推定する客観的, 数理的な手続き
- 例
  - 差の検定: 母平均・中央値に差があるかを推定
  - 相関・回帰: 2変数の母集団に相関があるか推定
  - 分割表の検定: 母集団の度数に偏りがあるか
- 確率的な表現で判断する

18

## パラメトリックとノンパラメトリック検定



19

## 統計的検定は確率pで判断する

- $p$  = 有意確率という
  - 0~1の範囲を取る. 0.05の場合は5%
- 検定では $p$ が出力される
  - $p$ によって検定の結果を判断する
  - 差の検定では「差のない確率」が $p$
  - 相関の検定では「相関のない確率」が $p$
  - 分割表の検定では「偏りのない確率」が $p$
- $p < 0.05$ のとき「有意に〇〇がある」という
  - 差の検定なら, 差のない確率が5%のように小さいとき, 差があると考えの方が妥当だろう, という意味

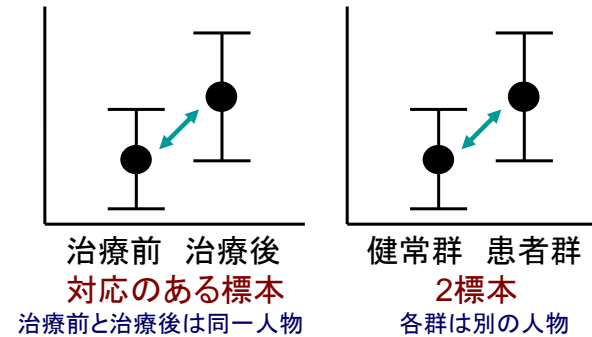
20

## 統計解析について

- 統計解析を手計算するのは時間の無駄
- 統計ソフトを使用すれば良い
  - Excelのような表計算ソフトでも計算可能だが、できる限り専用の統計ソフトを使用すべき
- 推奨される統計ソフト
  - IBM SPSS
    - ・ 有料ソフトであり、簡単には入手できない
  - Rコマンダー
    - ・ 無料統計ソフト。無料でも信頼度の高さは世界的に認められている
    - ・ 参考URL
    - ・ <http://www.hs.hirosaki-u.ac.jp/~pteiki/research/stat/S/>

## 差の検定

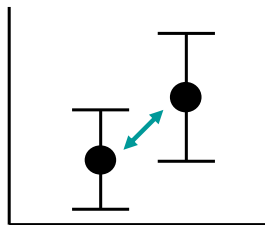
- 対応のある標本と2標本の「差」がある



22

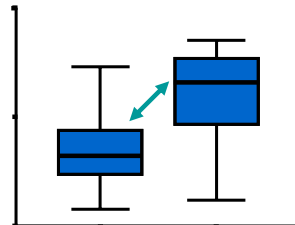
## 対応のある標本の差の検定

平均の差を比較



対応のあるt検定  
(パラメトリック法)

中央値の差を比較

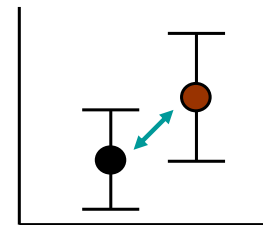


Wilcoxonの検定  
(ノンパラメトリック法)

23

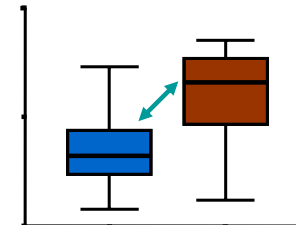
## 2標本の差の検定

平均の差を比較



2標本t検定  
(パラメトリック法)

中央値の差を比較



Mann-Whitneyの検定  
(ノンパラメトリック法)

24

## 対応のあるデータの例

- 被検者5人を対象
- 治療前後のBIIに差はあるか？

### 代表値の判断

- 尺度: 間隔尺度
- 正規分布か  
➤シャピロウィルク検定

	治療前	治療後
a	35	50
b	50	65
c	45	60
d	30	50
e	40	55

25

## 解析の結果 (Rコマンダーを使用)

```
[[3]]
対応のある場合のt検定
                        8.922 × 10-5

データ: a と b
t値 = -16, 自由度 = 4, P値 = 8.922e-05
対立仮説: 母平均の差は, 0ではない
95 パーセント信頼区間: -18.77645 -13.22355
標本推定値:
差の平均値
      -16
```

26

## 2標本データ例

- 歩行可能群と不可能群の年齢に差があるか

### 代表値の判断

- データ尺度: 比率尺度
- 正規分布するか  
➤シャピロウィルク検定

	可能群	不可能群
	74	90
	62	79
	82	71
	75	83
	77	84
	86	

27

## 解析の結果 (Rコマンダーを使用)

```
二標本t検定 (分散が等しいと仮定できるとき)

データ: a を as.numeric(b) で層別
t値 = -1.1564, 自由度 = 9, P値 = 0.2773
対立仮説: 母平均の差は, 0ではない
95 パーセント信頼区間: -15.963195  5.163195
標本推定値:
グループ1の平均値  グループ2の平均値
      76.0              81.4
```

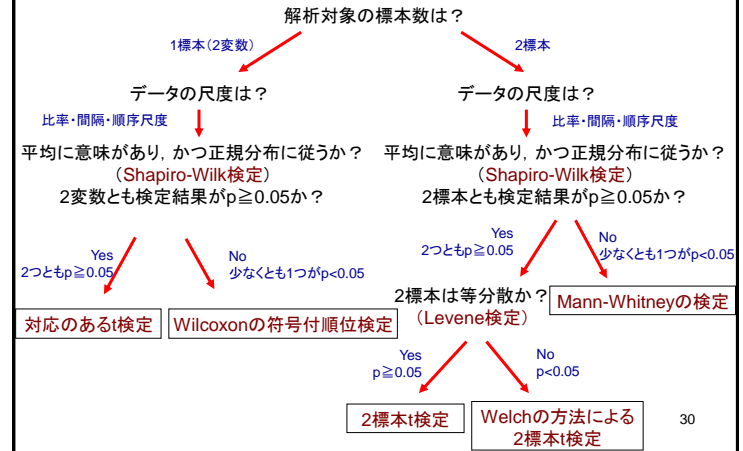
28

## 差の検定のまとめ

- 対応のある標本の差の検定
  - 平均: 対応のあるt検定
  - 中央値: Wilcoxonの符号付順位検定
- 2標本の差の検定
  - 平均: 2標本t検定 (Welchの検定)
  - 中央値: Mann-Whitney検定

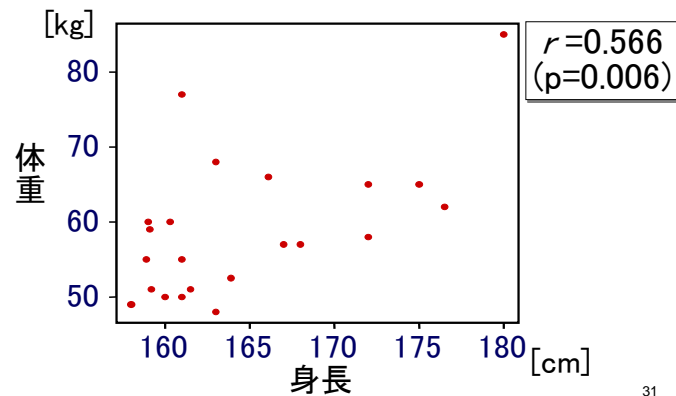
29

## 2変数または2標本の差の検定 検定選択のフローチャート



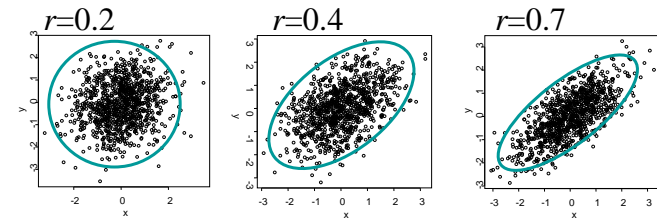
30

## 相関係数の例



31

## 相関係数解釈の目安



なし | 0.2 | やや相関 | 0.4 | かなり相関 | 0.7 | かなり強い相関

根拠のない基準であり、目安として考える

32



## 相関係数の選択方法

2変数とも、間隔・比尺度で  
かつ、正規分布に従うデータか？

Yes  
平均を使う

Pearsonの相関係数

No  
中央値を使う

Spearmanの  
順位相関係数

33

## 分割表の検定 ( $\chi^2$ 検定)

		退院先		合計
		自宅	施設	
歩行	可能	66	38	104
	不可能	40	68	108
合計		106	106	212

単位：人

- 名義尺度のデータに適用される
- 検定結果は $p=0.0001$  ( $p<0.01$ )

34

## p(有意確率)の解釈上の注意

- pが小さい=差や相関の確実性が高い
  - pが小さくても差や相関が大きいのではない
  - pの大きさと差の程度は無関係である
- nが大きいとき、pは小さくなる
  - nが大きいときは有意性に加えて、データの平均差や相関係数の大きさも考慮する

35

## 統計解析を修得するために

- 例題を参考に、統計ソフトを使ってみる
  - 統計ソフトによって操作マニュアルを準備
- 例題に似たデータを、自分でとってみる
  - 対象者5, 6人に同じデータをとってみる
- 論文の統計記載部分を真似てみる
  - 自分のやりたい研究と似通った論文を集めて、同じような手順でやってみる

36