

ノンパラメトリック検定

資料: Rによる統計的検定と推定

範囲: 第4章 (P.97～P.116)

著者: 内田 治

西澤 英子

60191008

大石 聖也

正規分布

- ・平均的な値が最も多く、その値から離れるほど少なくなる

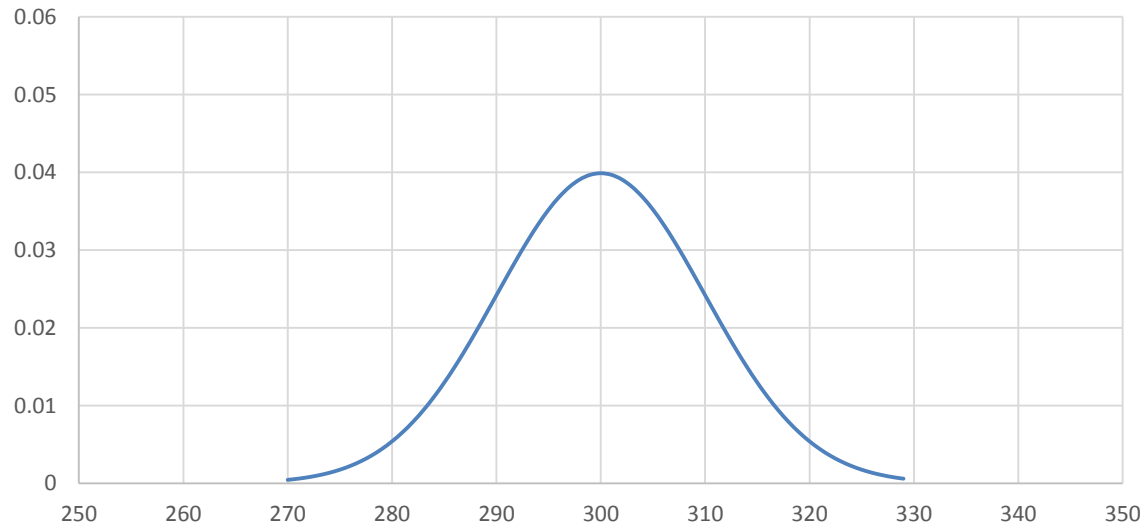


図1. 平均300、標準偏差10の正規分布

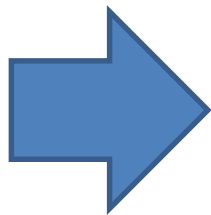
尺度

名称	説明	例
名義尺度	他と区別し分類するための名称のようなもの	血液型
順序尺度	順序・大小に意味があり、間隔には意味がないもの	1位・2位・3位
間隔尺度	メモリが等間隔になっているもので、その間隔に意味があるもの	西暦
比例尺度	0が原点であり、間隔と比例に意味があるもの	体重

多くの検定手法について

- ・母集団分布が正規分布であることを仮定している

→正規分布に従わないデータに適用しても信用できない



ノンパラメトリック法

ノンパラメトリック法とは

- データに特定の分布を仮定しない解析方法の総称
- **順位値**を解析の対象とする
→ もとのデータの分布を問題としない解析

ノンパラメトリック法の適用場面

1. 母集団のデータが正規分布か疑わしい
2. 外れ値を含めた解析を行いたい
3. データが順位値である
4. データが順序尺度のデータである

ノンパラメトリック法の注意点

- 正規分布を仮定した手法に比べ、検出力が低下する

- 検出力

本当は差がある状態を正しく判定する確率

ノンパラメトリック法の種類

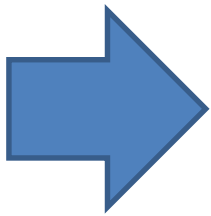
ノンパラメトリック法（手法の総称）

- ウィルコクスの順位和検定
- ウィルコクスの符号つき順位検定
- ムッド検定
- クラスカル-ウォリス検定
- フリードマン検定
- 順位相関係数

平均値に関するt検定の代用

- ・t検定

2つの独立した標本における平均値の差が統計学的に意味があるかを判定



ウィルコクソンの順位和検定

ウィルコクソンの順位和検定

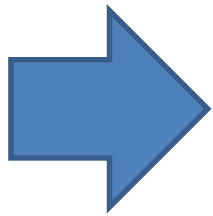
- ・グループを区別せずにデータの順位をつける
- ・グループごとの順位合計を求め、その値の差を検討

表1. グループに分けたデータ

A	B
6 8	4 27
1 91	3 36
5 15	2 64

対応のあるt検定の代用

- ・ウィルコクソンの順位和検定
2つのグループのデータが独立に収集された
場面で適用できる手法



ウィルコクソンの符号つき順位検定

ウィルコクソンの符号つき順位検定

- ・対応するデータの差(A-B)の絶対値から順位をつける

表2. 1回目(A)と2回目(B)のデータ

- ・順位から検討

	A	B
3	8	27
1	91	36
2	15	64

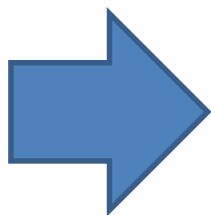
Diagram illustrating the ranking process for the Wilcoxon signed-rank test. The absolute differences are calculated for each pair of data (A-B) and ranked from smallest to largest:

- Rank 3: $|-19|$ (corresponding to A=8, B=27)
- Rank 1: $|55|$ (corresponding to A=91, B=36)
- Rank 2: $|-49|$ (corresponding to A=15, B=64)

等分散性のF検定の代用

- ・F検定

2つの独立した標本において、データが等分散かどうかを検定する手法

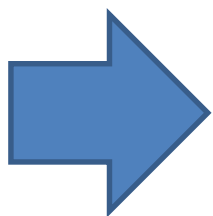


ムツド検定

分散分析の代用

- ・分散分析

データが正規分布に従い、等分散であるときに、
2つ以上の平均値の差を検定する手法



クラスカル-ウォリス検定

クラスカル-ウォリス検定

- ・適用場面

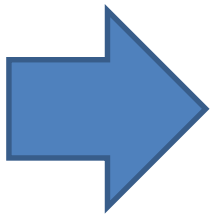
- 正規分布だと仮定できない

- 外れ値があり、等分散が仮定できない

二次元配置分散分析の代用

- ・二次元配置分散分析

分散分析の一種で、比較したい項目の種類が2つあるような場面で用いる手法



フリードマン検定

フリードマン検定

- ・適用場面

- 正規分布、等分散という条件を満たさない

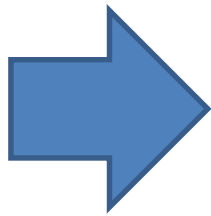
- 測定値が順位値である

表3. 3つの商品を3人の評価者が試食した順位

	A1	A2	A3
評価者1	3	1	2
評価者2	3	2	1
評価者3	3	1	2

相関係数の代用

- ・相関係数（ピアソンの相関係数）
2つの変数間の関係の強さを見る統計量



順位相関係数
(スピアマンの順位相関係数
ケンドールの順位相関係数)

順位相関係数

- 適用場面
 - 正規分布に従わない
 - 順位値、順序尺度のデータである
- スピアマン、ケンドールで計算方法が違い、計算結果も一致しない

Rのノンパラメトリック検定関数1

手法	用途	関数
ウィルコクスの順位和検定	t検定の代用	wilcox.test()
ウィルコクスの符号つき順位検定	対応のあるt検定の代用	wilcox.test()
ムッド検定	F検定の代用	mood.test()
クラスカル-ウォリス検定	一元配置分散分析の代用	kruskal.test()
フリードマン検定	二次配置分散分析の代用	friedman.test()

Rのノンパラメトリック検定関数2

手法	用途	関数
マクネマー検定	割合の違いを検定	<code>mcnemar.test()</code>
シャピロ-ウィルク検定	正規分布かを検定	<code>shapiro.test()</code>
アンサリー-ブラッドレー検定	分散に違いがあるかを検定	<code>ansari.test()</code>
フリグナー-キリン検定	ばらつきの同等性を検定	<code>frigner.test()</code>
クエード検定	繰り返しのない ブロック化データに行う検定	<code>quade.test()</code>

マクネマー検定

- 2つのグループと2つのカテゴリで、データに対応がある場合に、割合の違いを検定する

表4. 数学と英語のクロス集計表

		英語	
		好き	嫌い
数学	好き	18	12
	嫌い	20	50

マクネマー検定

- 帰無仮説 H_0
→ 英語の好きな割合 = 数学の好きな割合
- 対立仮説 H_1
→ 英語の好きな割合 $\neq$ 数学の好きな割合

マクネマー検定

- 入力

```
> X <- matrix(c(18,12,20,50), ncol=2, byrow=T)
```

- 検定

```
> mcnemar.test(X)
```

マクネマー検定

- ・結果

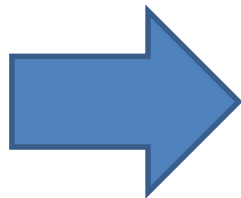
```
> mcnemar.test(X)
```

```
McNemar's Chi-squared test with continuity correction
```

```
data: X
```

```
McNemar's chi-squared = 1.5312, df = 1, p-value = 0.2159
```

- ・p-value > 有意水準 = 0.05



有意でない
差があるとはいえない

まとめ

- 正規分布か疑わしいデータに対しては、ノンパラメトリック検定
- Rには、ノンパラメトリック検定のための関数が用意されている