

4.2 ノンパラメトリック法の実際

4.2.1 ウィルコクスの順位和検定

例題1：2つの中心位置の比較

次のデータは、生活習慣の違いによりA群とB群に分けて総コレステロールの値を記録したデータである。前章(3.2節)では、このデータに t 検定を適用した。ここでは、ウィルコクスン(Wilcoxon)の順位和検定を適用する。

総コレステロールの値の記録

| A | B |
|-----|-----|
| 182 | 178 |
| 188 | 184 |
| 186 | 182 |
| 198 | 194 |
| 205 | 201 |
| 206 | 202 |
| 210 | 206 |
| 191 | 187 |
| 200 | 196 |
| 186 | 182 |
| 203 | 199 |
| 168 | 164 |
| 202 | 198 |
| 196 | 192 |
| 218 | 214 |
| 184 | 180 |
| 219 | 215 |
| 204 | 200 |
| 228 | 224 |
| 211 | 207 |

帰無仮説 H_0 ：2つの母集団分布の中心位置に差がない

対立仮説 H_1 ：2つの母集団分布の中心位置に差がある

データの入力

```
> A <- c(182,188,186,198,205,206,210,191,200,186,203,168,202,196,218,184,219,
+ 204,228,211)
> B <- c(178,184,182,194,201,202,206,187,196,182,199,164,198,192,214,180,215,
+ 200,224,207)
```

前章ではExcelを用いたデータの入力方法を紹介したが、ここでは、単純にRに直接入力する方法を紹介している。

ウィルコクソンの順位和検定のための関数

```
> wilcox.test(A, B, paired=F)
```

ウィルコクソンの順位和検定には、`wilcox.test()` という関数を利用する。書式は次のようになる。

```
wilcox.test(第1グループのデータ, 第2グループのデータ, paired=F,
alternative="two.sided")
```

「`paired=F`」は、データが対応づいているかどうかの設定である。対応がある場合は「`paired=F`」（FはFALSEの略、FALSEと入力しても可）と入力し、対応がない場合は「`paired=T`」（TはTRUEの略）と入力する。

「`alternative="two.sided"`」は、両側検定での p 値の計算を意味する。省略すると、自動的に両側検定となる。

検定の結果

```
Wilcoxon rank sum test with continuity correction

data: A and B
W = 236, p-value = 0.3367
alternative hypothesis: true location shift is not equal to 0

警告メッセージ:
In wilcox.test.default(A, B, paired = F) :
  タイがあるため、正確な p 値を計算することができません
```

$p\text{-value} = 0.3367 > \text{有意水準} = 0.05$

であり、有意ではない。したがって、2つの母集団分布の中心位置に差があるとはいえない。

「タイがあるため」というのは、同じ値のデータがあるため、同順位になるデータが存在するということを意味している。次のデータ例のように、タイがないときには、このメッセージは表記されない。

タイのないデータの例

| A | B |
|-----|-----|
| 168 | 164 |
| 179 | 181 |
| 182 | 178 |
| 184 | 180 |
| 186 | 183 |
| 188 | 185 |
| 191 | 187 |
| 196 | 192 |
| 198 | 194 |
| 200 | 197 |
| 202 | 199 |
| 203 | 201 |
| 204 | 207 |
| 205 | 208 |
| 206 | 209 |
| 210 | 212 |
| 211 | 213 |
| 218 | 214 |
| 219 | 215 |
| 228 | 224 |

```
> A <- c(168,179,182,184,186,188,191,196,198,200,202,203,204,205,206,210,211,
+ 218,219,228)
> B <- c(164,181,178,180,183,185,187,192,194,197,199,201,207,208,209,212,213,
+ 214,215,224)
> wilcox.test(A, B, paired=F)
```

```
Wilcoxon rank sum test
```

```
data: A and B
W = 210, p-value = 0.7994
alternative hypothesis: true location shift is not equal to 0
```

ウィルコクソンの順位和検定は、原データを、AグループとBグループを区別せずに、データの順位をつけ、次に、グループごとの順位の和（合計）を求めて、その値の差を検討している。

4.2.2 ウィルコクソンの符号つき順位検定

例題2：対応があるときの2つの中心位置の比較

摂取することにより、体脂肪を減らす効果があるといわれている飲料水の効果を測定したいと考えている。従業員から無作為に18人を選択し、当該飲料水を摂取する前の腹囲を測定した結果（変数A）と、同じ18人に対し、当該飲料水を毎日摂取してもらい、4週間後の腹囲を測定した結果（変数B）を記録した。このデータから、飲料水を摂取することにより、体脂肪を減らす効果があると判断してよいだろうか。

次のデータは、生活習慣の違いによりA群とB群に分けて総コレステロールの値を記録したデータである。前章（3.4節）では、このデータに対応のある t 検定を適用した。ここでは、ウィルコクソン（Wilcoxon）の符号つき順位検定を適用する。

摂取前（A）と摂取後（B）の腹囲

| A | B |
|------|------|
| 97.4 | 98.8 |
| 90.7 | 88.4 |
| 68.1 | 66.7 |
| 94.4 | 90.2 |
| 84.5 | 84.4 |
| 89 | 90.5 |
| 92.5 | 93.8 |
| 74.5 | 73.9 |
| 91.3 | 90.9 |
| 76.3 | 77.4 |
| 65.9 | 65.7 |
| 89.6 | 88.3 |

（次ページに続く）

摂取前 (A) と摂取後 (B) の腹囲 (前ページからの続き)

| A | B |
|-------|------|
| 84.1 | 78.9 |
| 72.4 | 71.5 |
| 96 | 96.8 |
| 63.3 | 62.4 |
| 105.3 | 99.3 |
| 64 | 64.2 |

帰無仮説 H_0 : 差 (A - B) の母集団分布の中心位置に差がない

対立仮説 H_1 : 差 (A - B) の母集団分布の中心位置は正である

データの入力

```
> A <- c(97.4, 90.7, 68.1, 94.4, 84.5, 89, 92.5, 74.5, 91.3, 76.3, 65.9, 89.6, 84.1, 72.4, 96,
+ 63.3, 105.3, 64)
> B <- c(98.8, 88.4, 66.7, 90.2, 84.4, 90.5, 93.8, 73.9, 90.9, 77.4, 65.7, 88.3, 78.9, 71.5,
+ 96.8, 62.4, 99.3, 64.2)
```

ウィルコクソンの符号つき順位検定のための関数

```
> wilcox.test(A, B, paired=T, alternative="greater")
```

「paired=T」は、データに対応があることを意味している。このときにはウィルコクソンの符号つき順位検定が実行され、「paired=F」とすると、前述のウィルコクソンの順位和検定が実行される。「alternative="greater"」は、片側検定を実施する。

検定の結果

```
Wilcoxon signed rank test with continuity correction

data: A and B
V = 116.5, p-value = 0.09196
alternative hypothesis: true location shift is greater than 0

警告メッセージ:
In wilcox.test.default(A, B, paired = T, alternative = "greater") :
```

タイがあるため、正確な p 値を計算することができません

$p\text{-value} = 0.09196 > \text{有意水準} = 0.05$

であり、有意ではない。したがって、差 $(A - B)$ の母集団分布の中心位置は正であるとはいえない。

4.2.3 ムッド検定

例題3：ばらつきの比較

Aクラス10人とBクラス12人の学生に、数学のテストを受けさせた。以下は、その結果を記録しているデータである。これより、2つのグループの分散は等しいといえるだろうか。前章(3.5節)では、このデータにF検定を適用した。ここでは、ムッド(Mood)検定を適用する。

AクラスとBクラスの数学のテスト結果

| A | B |
|----|----|
| 82 | 72 |
| 57 | 63 |
| 66 | 64 |
| 75 | 74 |
| 89 | 66 |
| 72 | 52 |
| 67 | 55 |
| 62 | 40 |
| 49 | 49 |
| 74 | 74 |
| | 62 |
| | 74 |

帰無仮説 H_0 : 2つの母集団分布のばらつきに差がない
対立仮説 H_1 : 2つの母集団分布のばらつきに差がある

May not be copied, displayed, distributed, modified, published, reproduced, stored, transmitted all or any part of the content from this site in any medium to anyone except for personal and non-commercial use permitted under the copyright law of Japan.

データの入力

```
> A <- c(82, 57, 66, 75, 89, 72, 67, 62, 49, 74)
> B <- c(72, 63, 64, 74, 66, 52, 55, 40, 49, 74, 62, 74)
```

ムッド検定のための関数

```
> mood.test(A, B)
```

検定の結果

```
Mood two-sample test of scale

data:  A and B
Z = 0.5729, p-value = 0.5667
alternative hypothesis: two.sided
```

$p\text{-value} = 0.5667 > \text{有意水準} = 0.05$

であり、有意ではない。したがって、AとBの母集団分布のばらつきに差があるとはいえない。

4.2.4 クラスカル-ウォリス検定

例題4：3つ以上の中心位置の比較

次のデータは3箇所の異なる水源（変数 x ）から収集した上水道水に含まれる塩素の量（変数 y ）を記録したものである。水源地の違いによって含まれる塩素の量に違いがあるといえるだろうか。前章（3.7節）では、このデータに一元配置分散分析を適用した。ここでは、クラスカル-ウォリス（Kruskal-Wallis）検定を適用する。

3つの水源 (x) に含まれる塩素の量 (y) の記録

| x | y |
|-----|------|
| 1 | 1.75 |
| 1 | 0.63 |
| 1 | 0.14 |
| 1 | 1.74 |
| 1 | 1.27 |
| 1 | 1.2 |
| 1 | 1.05 |
| 2 | 1.12 |
| 2 | 0.77 |
| 2 | 0.3 |
| 2 | 2.16 |
| 2 | 0.29 |
| 2 | 0.11 |
| 2 | 0.06 |
| 3 | 0.14 |
| 3 | 0.07 |
| 3 | 0.42 |
| 3 | 0.74 |
| 3 | 0.56 |
| 3 | 1.03 |
| 3 | 0.21 |

帰無仮説 H_0 : 3つの母集団分布の中心位置に差がない

対立仮説 H_1 : 3つの母集団分布の中心位置に差がある

データの入力

```
> A1 <- c(1.75, 0.63, 0.44, 1.74, 1.27, 1.2, 1.05)
> A2 <- c(1.12, 0.77, 0.3, 1.16, 0.29, 0.11, 0.06)
> A3 <- c(0.64, 0.47, 0.42, 0.74, 0.56, 1.03, 0.21)
```

クラスカル-ウォリス検定のための関数

```
> kruskal.test(list(A1, A2, A3))
```

クラスカル-ウォリス検定を行うには、`kruskal.test()` という関数を用いることになる。引数として、ベクトル入力した A1、A2、A3 の個別のデータを、

`list()` という関数を使ってデータフレーム化している。

検定の結果

```
Kruskal-Wallis rank sum test

data: list(A1, A2, A3)
Kruskal-Wallis chi-squared = 6.1076, df = 2, p-value = 0.04718
```

$p\text{-value} = 0.04718 < \text{有意水準} = 0.05$

であり、有意である。したがって、3つの母集団分布の中心位置に差があるといえる。

4.2.5 フリードマン検定

例題5：順位の一貫性に関する検定

5つの食品A1、A2、A3、A4、A5を6人の評価者が試食した。美味しいと感じる順に順位をつけた結果、次のようになった。

5つの食品を6人の評価者が試食した順位

| | A1 | A2 | A3 | A4 | A5 |
|------|----|----|----|----|----|
| 評価者1 | 4 | 1 | 2 | 3 | 5 |
| 評価者2 | 3 | 2 | 1 | 4 | 5 |
| 評価者3 | 5 | 1 | 3 | 2 | 4 |
| 評価者4 | 3 | 2 | 1 | 4 | 3 |
| 評価者5 | 4 | 1 | 2 | 3 | 5 |
| 評価者6 | 2 | 1 | 3 | 4 | 5 |

評価者間に順位の一貫性があるかどうかを検定する。

帰無仮説 H_0 : 6人の評価者に順位の一貫性がない

対立仮説 H_1 : 6人の評価者に順位の一貫性がある

このデータは因子として、食品と評価者の2つの因子を取り上げているので、

二元配置実験のデータとなる。測定値は順位値であるので、各評価者の平均値は全員が3となり、評価者間の平均値の差を議論することには意味がない。この例における評価者のような因子をブロック因子と呼び、因子としてブロック因子を含んだ実験を乱塊法と呼んでいる。通常の測定値であれば、乱塊法による実験データは、二元配置分散分析により解析することが可能となるが、このような順位データのときにはフリードマン (Friedman) 検定が適用できる。

評価者間に順位の一致性があるかどうかを見る検定であるが、これは食品間に差があるかどうかを見ていることにもなる。したがって、次のような仮説を検定していることにもなる。

帰無仮説 H_0 : 5つの食品間の順位に差がない

対立仮説 H_1 : 5つの食品間の順位に差がある

データの入力 (形式1)

```
> C1 <- c(4,1,2,3,5)
> C2 <- c(3,2,1,4,5)
> C3 <- c(5,1,3,2,4)
> C4 <- c(3,2,1,4,3)
> C5 <- c(4,1,2,3,5)
> C6 <- c(2,1,3,4,5)
> X <- rbind(C1, C2, C3, C4, C5, C6)
```

フリードマンの検定で、因子が2つ登場する列方向にベクトル入力したときには、`rbind()` という関数を使って各列を結合し、二元表の形にしている。

データの入力 (形式2)

```
> A1 <- c(4,3,5,3,4,2)
> A2 <- c(1,2,1,2,1,1)
> A3 <- c(2,1,3,1,2,3)
> A4 <- c(3,4,2,4,3,4)
> A5 <- c(5,5,4,3,5,5)
> X <- cbind(A1, A2, A3, A4, A5)
```

行方向にベクトル入力したときには、`cbind()` という関数を使って各行を結合し、二元表の形にしている。

フリードマン検定のための関数

```
> friedman.test(X)
```

フリードマンの検定を行うには、`friedman.test()` という関数を用いる。引数は二元表のデータセット名となる。

検定の結果

```
Friedman rank sum test

data: X
Friedman chi-squared = 16.6387, df = 4, p-value = 0.002272
```

$p\text{-value} = 0.002272 < \text{有意水準} = 0.05$

であり、有意である。したがって、6人の評価者に順位の一貫性があるといえる。すなわち、5つの食品間の順位に差があるといえる。

4.2.6 マクネマー検定

マクネマー (McNemar) 検定は、比べたいグループが2つあり、またカテゴリが2つ (たとえば、好きか嫌いか) で、データに対応がある場合に、割合の違いを検定するための方法である。

例題6：データに対応があるときの2×2クロス集計表

次のような2×2クロス集計表 (分割表) がある。これは数学と英語の好き嫌いを学生100人に調査した結果である。

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 |

英語が好きな学生の割合と、数学が好きな学生の割合に差があるといえるかどうかを検定する。

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

このデータから、英語の好き嫌いと数学の好き嫌いは関係があるかどうかという検定を行うのであれば、3.13節で説明した独立性のカイ2乗検定を適用すればよい。これは次のように実施する。

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

Pearson's Chi-squared test with Yates' continuity correction

data: X
X-squared = 7.5208, df = 1, p-value = 0.006099
```

$p\text{-value} = 0.006099 < \text{有意水準} = 0.05$

であり、有意である。したがって、英語の好き嫌いと数学の好き嫌いは関係があるといえるという結論になる。

また、正確確率検定を行うならば、次のように実施される。正確確率検定は、データの数が少ないときに実施される精密な検定法で、正規近似やカイ2乗近似を用いる方法とは異なるものである。クロス集計表に関する検定では `fisher.test()` という関数を用いる。

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

Fisher's Exact Test for Count Data

```

data: X
p-value = 0.003843
alternative hypothesis: true odds ratio is not equal to 1
95 percent confidence interval:
  1.398364  10.151515
sample estimates:
odds ratio
  3.695297

```

$p\text{-value} = 0.003843 < \text{有意水準} = 0.05$

であり、有意である。したがって、英語の好き嫌いと数学の好き嫌いは関係があるといえるという結論になる。

さて、この例題では、行と列の関係を検定するのではなく、英語の好きな割合と数学の好きな割合を検定するのが目的である。カイ2乗検定にせよ、正確確率検定にせよ、独立性の検定は行と列の関係を検定しているので、この例題には不適切である。このようなときには、マクネマー検定が適用できる。

データの入力

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

マクネマー検定のための関数

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

マクネマー検定を行うには、`mcnemar.test()` という関数を用いる。引数は 2×2 クロス集計表のデータセット名となる。

検定の結果

McNemar's Chi-squared test with continuity correction

```
data: X
```

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

$p\text{-value} = 0.2159 > \text{有意水準} = 0.05$

であり、有意でない。したがって、英語の好きな割合と数学の好きな割合に差があるとはいえないという結論になる。

4.2.7 シャピロ-ウィルク検定

データが正規分布に従っているかどうかを検定する方法の1つとして、シャピロ-ウィルク (Shapiro-Wilk) 検定と呼ばれる方法がある。

例題7：正規性の検定

次のような成人20人の体重を測定したデータがある。このデータから、体重は正規分布に従っているかどうかを検定する。

```
51, 69, 59, 64, 65, 49, 31, 65, 61, 85, 62, 63, 63, 65, 64, 72, 63, 70, 58, 60
```

帰無仮説 H_0 : 正規分布に従っている

対立仮説 H_1 : 正規分布に従っていない

母集団分布が正規分かどうかを確かめることは、統計的検定手法を適用するうえで、外れ値の検証とともに重要な検証である。なぜならば、多くの検定手法は正規分布を前提とした理論に基づいて構築されているからである。

正規分布かどうかの検定には、度数分布表に基づいて行う適合度検定と呼ばれる方法と、ここで紹介するシャピロ-ウィルク検定がよく用いられる。

データの入力

```
> X <- c(51, 69, 59, 64, 65, 49, 31, 65, 61, 85, 62, 63, 63, 65, 64, 72, 63, 70, 58, 60)
```

シャピロ-ウィルク検定のための関数

```
> shapiro.test(X)
```

シャピロ-ウィルクの検定を行うには、`shapiro.test()` という関数を用いる。引数はデータセット名となる。

検定の結果

```
Shapiro-Wilk normality test

data:  X
W = 0.8715, p-value = 0.01250
```

$p\text{-value} = 0.01250 < \text{有意水準} = 0.05$

であり、有意である。したがって、正規分布に従っているとはいえないという結論になる。

4.2.8 Rのノンパラメトリック検定関数

Rに装備されているノンパラメトリック検定のための関数をまとめて列挙すると、次のようになる。

- 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()`
- 二標本のばらつきの違いについてのムッド検定 (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 (二元配置分散分析-乱塊法の代用)
`friedman.test()`
- 2×2 クロス集計表におけるマクネマー検定
`mcnemar.test()`
- 正規性に対するシャピロ-ウィルク検定
`shapiro.test()`

以上のほかに次のような検定のための関数もある。

- アンサリー-ブラッドレー (Ansari-Bradley) 検定 (分散の違いに関する検定)
`ansari.test()`
- フリグナー-キリーン (Fligner-Killeen) 検定 (メディアン検定)
`fligner.test()`
- クエード (Quade) 検定 (繰り返しのないブロック化データ)
`quade.test()`

ちなみに、`apropos()` 関数を使って、`apropos(".test")` と入力すると、R に用意されている検定のための関数が、すべてではないが一覧することができる。

```
> apropos(".test")
[1] ".valueClassTest"      "ansari.test"
[3] "bartlett.test"        "binom.test"
[5] "Box.test"             "chisq.test"
[7] "cor.test"             "file_test"
[9] "fisher.test"          "fligner.test"
[11] "friedman.test"        "kruskal.test"
[13] "ks.test"              "mantelhaen.test"
[15] "mauchley.test"        "mauchly.test"
[17] "mcnemar.test"         "mood.test"
[19] "oneway.test"          "pairwise.prop.test"
[21] "pairwise.t.test"      "pairwise.wilcox.test"
[23] "poisson.test"         "power.anova.test"
[25] "power.prop.test"      "power.t.test"
[27] "PP.test"              "prop.test"
[29] "prop.trend.test"      "quade.test"
[31] "shapiro.test"         "t.test"
[33] "var.test"             "wilcox.test"
```