ノンパラメトリック検定

はじめに

はじめに

これまでの章では，母集団分布が正規分布や特定の分布に従う検定問題を扱ってきた.

しかしながら，母集団分布に特定の分布を仮定する根拠が見いだせない事は多く存在する.

そのような状況の下での仮説検定はどのように扱えばよいだろうか?

この章では，正規分布や特定な分布を仮定できない場合の検定法について述べる.

はじめに

本節で扱うノンパラメトリック検定は，基本的な検定統計量に焦点を当てている.

しかし，長年の研究によって様々な検定統計量やノンパラメトリック法に関する理論が確立されている.

ノンパラメトリック法について，更に詳しく知りたい読者は，例えば 「村上秀俊著 (2015) ノンパラメトリック法，朝倉書店」を参照されたい.

はじめに：歴史

ノンパラメトリック法といっても様々な手法があるが，大きく「統計的仮説検定」と「密度推定」に分類することができる.

特に，ノンパラメトリック法の中でも順位に基づく手法がノンパラメトリック法の大部分を占めている.

ノンパラメトリック法は，1930 年代初め頃から医学，自然科学，心理学，経済学，工学といった様々な分野で広く利用，そして応用されてきた.

はじめに：歴史

そのことから，ノンパラメトリック法の始まりは1936年としている文献もある.

しかし，実際は18世紀初頭に2項確率計算に基づく方法，すなわち符号検定が使用されていた.

1930年代終わりには，パラメータ（母数）を利用しない母集団分布の比較について研究がされ始めた.

1940年代に入ると，順位に基づく手法の出発点となる論文が発表され，1960年代初めまでは，順位に基づく手法の研究が最新の研究課題となっていた.

はじめに：歴史

1940年代後半から1950年代にかけて，母集団分布を仮定する従来のパラメトリック法と，ノンパラメトリック法との関連性について明らかになってきた.

1960年代，ノンパラメトリック法は最も重要な研究課題となり，順位検定からどのように推定量が導かれるか証明され，そして，その推定量が非常に価値のある性質であることが示された.

そして，1960 年代後半には，ノンパラメトリック法の土台となる理論が数多く構築された.

はじめに：歴史

1970年代初期には，線形モデルへノンパラメトリック法が展開され，生存解析に対するノンパラメトリックモデルの研究の火付け役となる論文が発表された.

また，この頃からノンパラメトリックベイズ法に関する論文も発表されてきて，生存曲線のノンパラメトリックベイズ推定量も提案され始めた.

1980 年代初期には，電子データ処理の急速な発達が，ノンパラメトリック法の更なる発展のためには必要不可欠なことであることが予見されていた.

はじめに：歴史

実際に，コンピュータの発達に伴って，ノンパラメトリック法は目覚ましい発展がなされてきた.

情報技術の発展の結果，1980年代後半から1990年代にかけて，マルコフ連鎖モンテカルロ法が要因となって，ベイズ法への関心が一気に高まってきた.

21世紀に入っても，ノンパラメトリック法の提案が活発になされており，ビッグデータ時代と言われる今日において，データ分析における密度推定とノンパラメトリック法の重要性が確認されている.

はじめに：歴史

現在でも，ノンパラメトリック法は統計学の中で最も重要な手法の一つとして知られ，様々な統計ソフトにも標準的な機能として組み込まれている.

では，なぜノンパラメトリック法が広く用いられるようになったかというと

主な要因として，母集団分布に関する仮定が少なく，応用しやすい

からである.

はじめに：歴史

これまでの歴史を見て分かるように，ノンパラメトリック法は大きく統計的仮説検定と密度（分布）推定の 2 つに分類することができる.

ノンパラメトリック法における仮説検定は，順位もしくは経験分布関数に基づく手法であるが，

順位: Wilcoxon (1945)
経験分布関数: Kolmogorov (1933) と Smirnov (1936)

が有名な論文である.

はじめに：歴史

ノンパラメトリック法における密度推定は，カーネル関数に基づく推定方法が有名であるが，

Rosenblatt (1956) と Parzen (1962)

が有名な論文である.

ノンパラメトリック法における分布推定は，経験分布関数による推定が有名であり，基本的な統計学の参考書には記載されている方法であることを補足する.

はじめに：注意点

しかし，どんな場合にもノンパラメトリック法を用いれば良いという訳ではない.

例えば，母集団分布に特定の分布が仮定できることが分かっている場合は，ノンパラメトリック法よりパラメトリック法が良い結果を得ることが出来る.

線形順位和検定

線形順位和検定

ノンパラメトリック法において，統計的仮説検定では順位に基づく方法が大部分を占めているので，まず初めに順位変換について紹介する.

例えば，次のようなデータが与えられているとする.

$\boldsymbol{X}_1$	0.64	-2.54	-1.60	1.28	1.57
$\boldsymbol{X}_2$	1.04	0.77	-0.82		

このとき，第 1 標本の合計および平均は，それぞれ -0.65 と -0.13 である．同様に，第2標本の合計および平均は，それぞれ 0.99 と 0.33 である.

線形順位和検定

ここで第 1 標本と第 2 標本のデータを結合し，昇順する.

データが第 1 標本 に属するものであれば 1，第 2 標本に属するものであれば 0 とすると,

$\boldsymbol{X}_1$	0.64	-2.54	-1.60	1.28	1.57
$\boldsymbol{X}_2$	1.04	0.77	-0.82		

は

データ	-2.54	-1.60	-0.82	$\mathbf{0.64}$	0.77	1.04	$\mathbf{1.28}$	$\mathbf{1.57}$
順位	1	2	3	4	5	6	7	8
属性	1	1	0	1	0	0	1	1

となる.

線形順位和検定

これを元々の標本

$\boldsymbol{X}_1$	0.64	-2.54	-1.60	1.28	1.57
$\boldsymbol{X}_2$	1.04	0.77	-0.82		

に割り当てみると

$\boldsymbol{X}_1$	4	1	2	7	8
$\boldsymbol{X}_2$	6	5	3		

と順位変換を施すこととなる.

線形順位和検定

ところで，データの収集や入力をする際，何らかの間違いが起こる可能性もある.

例えば，第 1 標本のあるデータは，本来は 1.57 であるが，入力ミスによって 157 入力されたとすると，第 1 標本の合計は，154.78 となるが，順位を用いた場合は 4, 1, 2, 7, 8 と変わらない.

ノンパラメトリック法が外れ値に影響されにくい，すなわち頑健である所以のひとつである.

線形順位和検定

もうひとつ具体例を挙げて，順位について紹介をする．
例えば，次のようなデータが与えられているとする．

$\boldsymbol{X}_1$	-0.64	-2.54	-1.60	-1.28	-1.57
$\boldsymbol{X}_2$	1.04	0.77	0.82		

第 1 標本と第 2 標本の合計は，それぞれ -7.63 と 2.63 である．前の例と同様の方法で順位をつけると

$\boldsymbol{X}_1$	5	1	2	4	3
$\boldsymbol{X}_2$	8	6	7		

となる．

線形順位和検定

第 1 標本の順位の合計は 15 となる.

第 2 標本のデータが，第 1 標本のデータよりも大きい傾向にあるならば，第 2 標本のデータの順位は第 1 標本のデータの順位よりも大きくなる傾向がある.

順位に変換しても本来のデータの特性を反映している！

線形順位和検定

$\boldsymbol{X}_1 = (X_{11}, X_{12}, \ldots, X_{1n_1})$ を分布関数 $F_1(x)$ を持つ母集団からの確率標本

$\boldsymbol{X}_2 = (X_{21}, X_{22}, \ldots, X_{2n_2})$ を分布関数 $F_2(x)$ を持つ母集団からの確率標本

このとき, 帰無仮説 $H_0 : F_1(x) = F_2(x)$ について考える.

ただし, $F_1(x)$と $F_2(x)$については連続性だけが仮定され, それ以外については分からないものとする.

線形順位和検定

検定統計量の構成手順

1. $\boldsymbol{X}_1 = (X_{11}, X_{12}, \ldots, X_{1n_1})$と$\boldsymbol{X}_2 = (X_{21}, X_{22}, \ldots, X_{2n_2})$を統合する.
2. 統合した確率標本を昇順に並べて新たな確率標本$\boldsymbol{Y} = (Y_{(1)}, Y_{(2)}, \ldots, Y_{(N)})$を構成する.
3. 次のような指示関数を定義する.

$$Z_i = \begin{cases} 1; & Y_{(i)} \in \boldsymbol{X}_1 \\ 0; & Y_{(i)} \in \boldsymbol{X}_2 \end{cases}$$

ただし, $N = n_1 + n_2$ であり, $Y_{(i)}$は順序統計量と呼ばれている.

線形順位和検定

検定統計量は

$$\mathrm{T} = \sum_{i=1}^{N} a_i Z_i \tag{1}$$

で定義され，a_i をスコア関数と呼ぶ.

また，帰無仮説 の下では以下を得る.

$$\begin{aligned} \mathrm{E}[Z_i] &= \frac{n_1}{N}, \\ \mathrm{V}[Z_i] &= \frac{n_1 n_2}{N^2}, \\ \mathrm{cov}(Z_i, Z_j) &= \frac{-n_1 n_2}{N^2(N-1)}, \ (i \neq j). \end{aligned}$$

線形順位和検定

期待値，分散，共分散の証明:

Z_i は 1 もしくは 0 の値をとる確率変数なので 2 項分布に従うことから, 帰無仮説の下では

$$\Pr(Z_i = 1) = \frac{n_1}{N}, \quad \Pr(Z_i = 0) = \frac{n_2}{N}$$

となる. よって, 平均と分散は

$$\mathrm{E}[Z_i] = \frac{n_1}{N}, \quad \mathrm{V}[Z_i] = \frac{n_1 n_2}{N^2}$$

で与えられる.

線形順位和検定

また, $i \neq j$ のとき, Z_i と Z_j の同時積率が

$$\mathrm{E}[Z_i Z_j] = \Pr(Z_i = 1 \cap Z_j = 1) = \frac{{}_{n_1}\mathrm{C}_2}{{}_{N}\mathrm{C}_2} = \frac{n_1(n_1-1)}{N(N-1)}$$

となることを用いることで,

$$\mathrm{cov}(Z_i, Z_j) = \mathrm{E}[Z_i Z_j] - \mathrm{E}[Z_i]\mathrm{E}[Z_j] = \frac{-n_1 n_2}{N^2(N-1)}$$

を得る.

線形順位和検定：はじめに

帰無仮説の下，検定統計量 T の期待値，分散は

$$\mathrm{E[T]} = \frac{n_1}{N}\sum_{i=1}^{N} a_i, \tag{2}$$

$$\mathrm{V[T]} = \frac{n_1 n_2}{N^2(N-1)}\left[N\sum_{i=1}^{N} a_i^2 - \left(\sum_{i=1}^{N} a_i\right)^2\right] \tag{3}$$

で与えられる.

線形順位和検定

2 つの線形順位和統計量を

$$\mathrm{T} = \sum_{i=1}^{N} a_i Z_i, \qquad \mathrm{T}^* = \sum_{i=1}^{N} a_i^* Z_i$$

とする.

帰無仮説 の下, 検定統計量 T および T^* の共分散は

$$\mathrm{cov}(\mathrm{T}, \mathrm{T}^*) = \frac{n_1 n_2}{N^2(N-1)} \left(N \sum_{i=1}^{N} a_i a_i^* - \sum_{i=1}^{N} a_i \sum_{i=1}^{N} a_i^* \right)$$

で与えられる.

位置の違いの検定方法

ウィルコクソン順位和検定

未知母数 Δ に対して

$$F_2(x) = \Pr(X_2 \leq x) = \Pr(X_1 \leq x - \Delta) = F_1(x - \Delta)$$

とする．ただし，$\Delta \neq 0$ である．

X_2 と $X_1 + \Delta$ は同じ分布，もしくは，X_1 と $X_2 - \Delta$ は同じ分布となる．

$\Delta > 0$ ならば，分布関数 F_1 は右方向へ平行移動する．
$\Delta < 0$ ならば，分布関数 F_1 は左方向へ平行移動する．
$\Delta = 0$ の場合，分布関数 F_1 と F_2 は等しくなる．

ウィルコクソン順位和検定

位置母数に対する検定問題

帰無仮説 $H_0 : F_1(x) = F_2(x)$

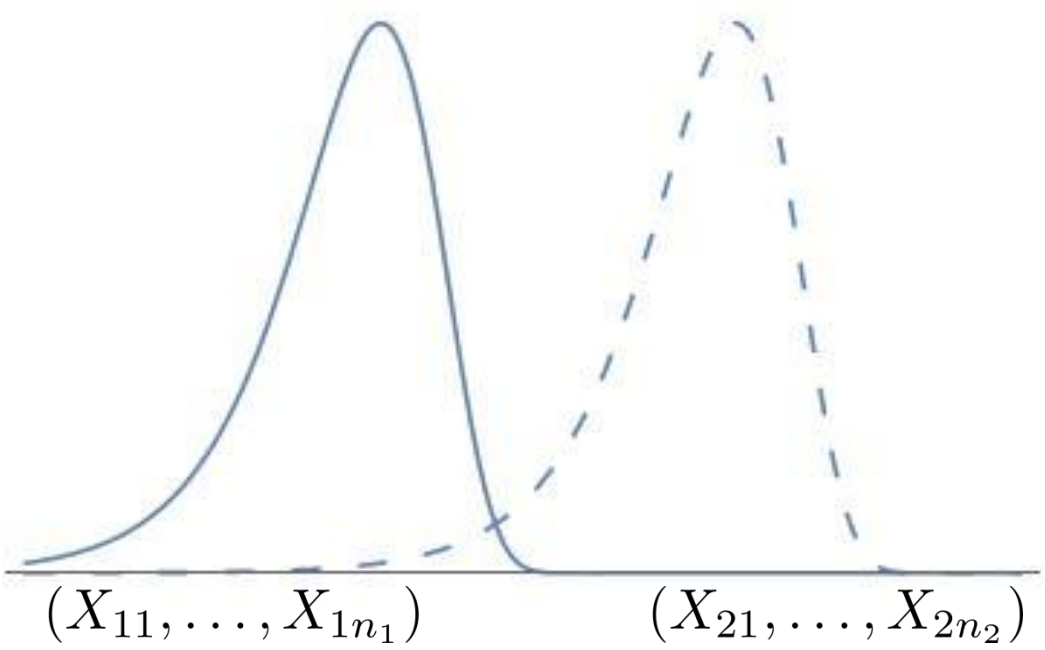
に対して,

両側対立仮説 $H_{11} : F_2(x) = F_1(x - \Delta), \quad \Delta \neq 0$

となる.

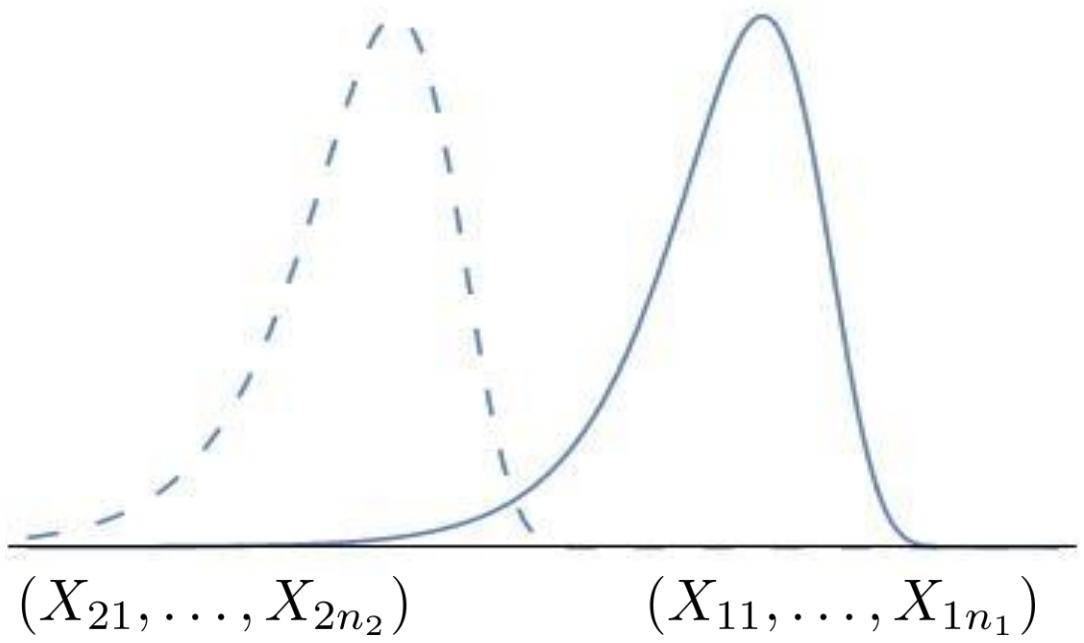
ウィルコクソン順位和検定

$\boldsymbol{X}_2$ の中央値が $\boldsymbol{X}_1$よりも大きいならば, $\boldsymbol{X}_2$ のデータの順位は $\boldsymbol{X}_1$ のデータの順位よりも大きくなる傾向がある.



ウィルコクソン順位和検定

$\boldsymbol{X}_2$ の中央値が $\boldsymbol{X}_1$よりも小さいならば, $\boldsymbol{X}_2$ のデータの順位は $\boldsymbol{X}_1$ のデータの順位よりも小さくなる傾向がある.



ウィルコクソン順位和検定

Wilcoxon (1945) は, $\boldsymbol{X}_1$ の順位和を用いた検定統計量

$$\mathrm{W} := \mathrm{T} = \sum_{i=1}^{N} iZ_i = \sum_{i=1}^{n_1} R_i$$

を提案した.

R_i は確率標本 $\boldsymbol{Y}$ における $\boldsymbol{X}_1$ の順位であり，これは (1) 式において $a_i = i$ とした統計量である.

ウィルコクソン順位和検定

帰無仮説の下での検定統計量 W の期待値および分散は，同順位がない場合，(2) 式および (3) 式より

$$\mathrm{E[W]} = \frac{n_1(N+1)}{2}, \qquad \mathrm{V[W]} = \frac{n_1 n_2 (N+1)}{12}$$

で与えられる．

標本サイズが十分大きい場合には，検定統計量 $\mathrm{T_W}$

$$\mathrm{T_W} = \frac{\mathrm{W} - \mathrm{E[W]}}{\sqrt{\mathrm{V[W]}}}$$

は漸近的に標準正規分布に従う．

ウィルコクソン順位和検定

このことは，チェルノフ・サーベージ（Chernoff and Savage, 1958）が線形順位和検定統計量が漸近的に標準正規分布に従うことを示したが，ウィルコクソン順位和検定がその条件を満たすことに由来する.

検定統計量 $\mathrm{T_W}$ が標準正規分布に従うとすると，

$|\mathrm{T_W}| \geq 1.96$ であるならば，有意水準 5% の検定で帰無仮説は棄却される.

標本サイズが小さい場合には，検定統計量 W の棄却点が正確に計算されている (付録参照).

ウィルコクソン順位和検定

例えば，標本サイズが $n_1 = n_2 = 7$ の場合，

W	16	17	18	19	20	21
正確	.0006	.0017	.0052	.0122	.0256	.0466
近似	.0016	.0035	.0073	.0141	.0232	.0478
W	22	23	24	25	26	27
正確	.0804	.1270	.1894	.2652	.3537	.4493
近似	.0794	.1233	.1742	.2604	.3502	.4487

となる.

ウィルコクソン順位和検定

現実問題として，$n_1 \geq 10$，$n_2 \geq 10$ ならばウィルコクソン順位和検定には正規近似を用いてもよいが，この近似の良さは稀である．

正規分布を仮定した下で，t 検定は位置の違いに関して一様最強力検定である．その t 検定に対するウィルコクソン順位和検定の漸近相対効率は $3/\pi \approx 0.955$ となる．

ロジスティック分布を仮定した下で，ウィルコクソン順位和検定は局所最強力検定となる．

ウィルコクソン順位和検定：例題

20匹のマウスをランダムに 2 群に分け，一方は対称群 Π_1 とし，他方の実験群 Π_2 にはある薬品を一定量投与した後，それぞれのリンパ球を数えた.

2 群の母集団分布の形はほぼ同等ということはわかっている．我々は薬品の投与によってリンパ球数は影響を受けるのかを調べたい．すなわち

$$
\begin{aligned}
H_0 &: F_1(x) = F_2(x) \\
H_{11} &: F_2(x) = F_1(x-\Delta), \quad \Delta \neq 0
\end{aligned}
$$

を検定する.

ウィルコクソン順位和検定：例題

この実験から次の表のような結果を得た.

マウスのリンパ球数

Π_1	59	91	64	53	75	69	36	43	54	99
Π_2	67	93	72	55	109	86	50	80	95	103

表の数字は, 5×10^2 個を単位としている.

この表のデータを大きさの順に並べ, Π_1 の観測値を太字にしたものが次表となる.

ウィルコクソン順位和検定：例題

順位付けされたデータ

$Y_{(i)}$	**36**	**43**	50	**53**	**54**	55	**59**	**64**	67	**69**
順位	1	2	3	4	5	6	7	8	9	10
Z_i	1	1	0	1	1	0	1	1	0	1
$Y_{(i)}$	72	**75**	80	86	**91**	93	95	**99**	103	109
順位	11	12	13	14	15	16	17	18	19	20
Z_i	0	1	0	0	1	0	0	1	0	0

この表から

$$\mathrm{W} = 1 + 2 + 4 + 5 + 7 + 8 + 10 + 12 + 15 + 18 = 82$$

$$\mathrm{E}[\mathrm{W}] = 105, \qquad \mathrm{V}[\mathrm{W}] = 175$$

となる.

ウィルコクソン順位和検定：例題

これを検定統計量 $\mathrm{T_W}$ に適用してみると

$$\mathrm{T_W} = \frac{|82 - 105|}{\sqrt{175}} = 1.73$$

となり，標準正規分布の両側 5%点の 1.96 より小さいので，有意水準 5%で帰無仮説を棄却できない.

正確な W の有意水準 5% 確率は

$$\Pr(79 \leq \mathrm{W} \leq 131) = 0.9566$$

で与えられる．W は 82 であるので，いずれにしても帰無仮説を棄却することはできない.

ウィルコクソン順位和検定：例題

このように，位置については不明であるが母集団分布の形が等しいことがわかっているときに，ウィルコクソン順位和検定は有効であると考えられる.

ウィルコクソン順位和検定を $\mathrm{W}^* = |\mathrm{W} - \mathrm{E}[\mathrm{W}]|$ として扱う場合が非常に多い.

その理由のひとつとして，検定統計量 W が平均に対して対称となるからである.

マン・ウィットニーの U 検定

確率標本 $\boldsymbol{X}_1$ が確率標本 $\boldsymbol{X}_2$ より大きくなる個数を, 指示関数

$$U(i,j) = \begin{cases} 0; & X_{1i} < X_{2j} \\ 1; & X_{1i} \geq X_{2j} \end{cases}$$

を用いて定義する.

Mann and Whitney (1947) によって検定統計量

$$\mathrm{MW} = \sum_{i=1}^{n_1} \sum_{j=1}^{n_2} U(i,j)$$

が提案されている.

マン・ウィットニーの U 検定

確率標本 $\boldsymbol{Y}$ における $\boldsymbol{X}_1$ の順位 R_i は

$$R_i = \sum_{j=1}^{n_2} U(i,j) + i$$

と表わすことができる．ただし，一般性を失うことなく $X_{11} < X_{12} < \cdots < X_{1n_1}$ が仮定されるものとする．

したがって，検定統計量 W は

$$\mathrm{W} \quad = \quad \sum_{i=1}^{N} iZ_i = \sum_{i=1}^{n_1} R_i = \sum_{i=1}^{n_1}\sum_{j=1}^{n_2} U(i,j) + \frac{n_1(n_1+1)}{2}$$

となる．

マン・ウィットニーの U 検定

右辺の第 2 項はデータ数のみに依存するので必ず定数となることから，本質的にウィルコクソン順位和検定とマン・ウィットニーの U 検定は等しいことが分かる.

帰無仮説の下，検定統計量 MW の期待値，分散は

$$\mathrm{E}[\mathrm{MW}] = \frac{n_1 n_2}{2}, \quad \mathrm{V}[\mathrm{MW}] = \frac{n_1 n_2 (N+1)}{12}$$

で与えられる.

例えば，確率母関数や理論的に考えるときなど，マン・ウィットニーの U 検定で考えたほうが便利な場合がある.

マン・ウィットニーの U 検定：R

統計ソフトRにおけるウィルコクソン順位和検定は，マン・ウィットニーの U 検定が用いられている.

先程の例では，下記のプログラムを実行することにより，ウィルコクソン順位和検定（マン・ウィットニーの U 検定）の p 値を求めることが出来る.

```
data1 <- c(59,91,64,53,75,69,36,43,54,99)
data2 <- c(67,93,72,55,109,86,50,80,95,103)
wilcox.test(data1, data2, paired=FALSE,
alternative=c("two.sided"), conf.int=TRUE,
conf.level=0.95, correct=FALSE)
```

マン・ウィットニーの U 検定：R

ちなみに,

- 片側検定の場合, alternative のオプションを ``less'' もしくは ``greater'' とする
- 信頼係数を変更する場合は, conf.level の数値を変更する
- 信頼区間を構成しない場合は, conf.int のオプションを FALSE とする
- 連続性の補正を行なう場合は, correct のオプションを TRUE とする

ことに注意されたい.

マン・ウィットニーの U 検定：R

同順位がある場合，上記のプログラムでは正確な p 値を計算することが出来ないため，

```
wilcox.exact(data1, data2, paired=FALSE,
alternative=c("two.sided"), conf.int=TRUE,
conf.level=0.95, correct=FALSE)
```

を実行する.

ただし，パッケージ ``exactRankTests'' をインストールする必要がある.

尺度の違いの検定方法

ムード検定

未知母数 Δ に対して

$$F_2(x) = \Pr(X_2 \leq x) = \Pr\left(\frac{X_1}{\Delta} \leq x\right) = F_1(\Delta x), \quad \Delta > 0$$

とする.

X_2 と X_1/Δ は同じ分布，もしくは，X_1 と ΔX_2 は同じ分布となる.

$\Delta > 1$ なら，分布関数 F_1 は左右に広がる.
$\Delta < 1$ なら，分布関数 F_1 は左右にあまり広がらない.
$\Delta = 1$ なら，分布関数 F_1 と F_2 は等しくなる.

ムード検定

$\boldsymbol{X}_1 = (X_{11}, X_{12}, \ldots, X_{1n_1})$ を分布関数 $F_1(x)$ を持つ母集団からの確率標本

$\boldsymbol{X}_2 = (X_{21}, X_{22}, \ldots, X_{2n_2})$ を分布関数 $F_2(x)$ を持つ母集団からの確率標本

帰無仮説 $H_0 : F_1(x) = F_2(x)$

に対して

両側対立仮説 $H_{21} : F_2(x) = F_1(\Delta x), \quad \Delta \neq 1, \quad \Delta > 0$

について考える.

ムード検定

同順位がない N 個の順位付けされた確率標本において，結合順位の平均は $(N+1)/2$ となる.

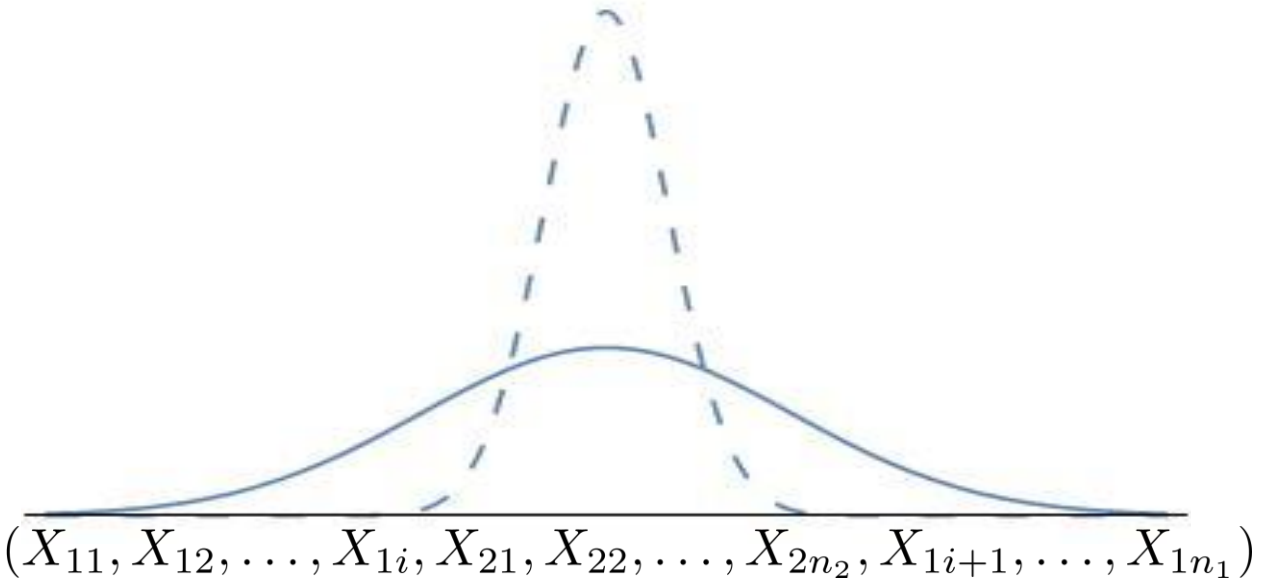
尺度母数の違いを検定するために，$\boldsymbol{X}_1$ と $\boldsymbol{X}_2$ の中央値が既知であるか，もしくは未知であっても等しいことが仮定される必要がある.

中央値が未知であり，等しいことが言えない場合，その確率標本の推定値を利用して中央値を合わせる必要がある．この場合，ノンパラメトリック検定ではないという意見もあるが，本書は基本を学ぶことを目的としているため，深くは追求しないこととする.

ムード検定

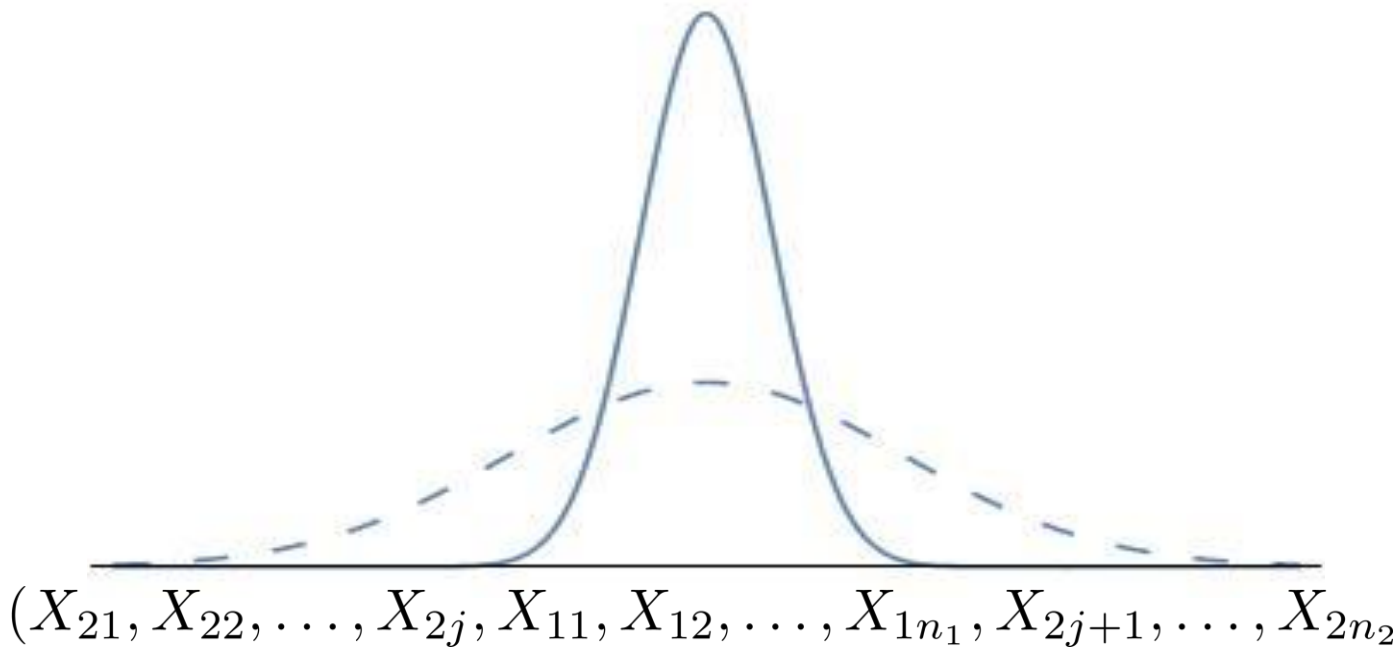
$f_1(x)$と $f_2(x)$ をそれぞれ分布関数 $F_1(x)$ と $F_2(x)$ の密度関数とする．そのとき $f_2(x) = \Delta f_1(\Delta x)$ のような関係であるとする．

Δの値が大きいと確率標本 $\boldsymbol{X}_2$ の順位は中央に位置づけられ，下図のようになる．



ムード検定

$f_2(x) = \Delta f_1(\Delta x)$ のような状況において $0 < \Delta < 1$ のとき，確率標本 $\boldsymbol{X}_2$の順位は両側に位置づけられ，下図のようになる.



ムード検定

Mood (1954) は X_1 の順位に基づく検定統計量

$$\mathrm{MO} := \mathrm{T} = \sum_{i=1}^{N} \left(i - \frac{N+1}{2} \right)^2 Z_i = \sum_{i=1}^{n_1} \left(R_i - \frac{N+1}{2} \right)^2$$

を提案した.

R_i は確率標本 $\boldsymbol{Y}$ における $\boldsymbol{X}_1$ の順位であり，これは (1) において $a_i = \{i - (N+1)/2\}^2$ とした検定統計量である.

ムード検定

帰無仮説の下でのムード検定の期待値および分散は，同順位がない場合，(2) 式および (3) 式より

$$\mathrm{E}[\mathrm{MO}] = \frac{n_1(N^2-1)}{12}, \quad \mathrm{V}[\mathrm{MO}] = \frac{n_1 n_2 (N+1)(N^2-4)}{180}$$

で与えられる.

標本サイズが大きい場合には，検定統計量 T_{MO}

$$\mathrm{T}_{\mathrm{MO}} = \frac{\mathrm{MO} - \mathrm{E}[\mathrm{MO}]}{\sqrt{\mathrm{V}[\mathrm{MO}]}}$$

は，帰無仮説のもとで漸近的に標準正規分布に従う.

ムード検定

チェルノフ・サーベージ (Chernoff and Savage, 1958) によって，線形順位和検定統計量が漸近的に標準正規分布に従うことが示されたが，ムード検定がその条件を満たすことに由来する.

つまり，$|\mathrm{T_{MO}}| \geq 1.96$ であるならば，有意水準 5% の検定で帰無仮説は棄却される.

標本サイズが小さい場合には，検定統計量 MO の正確な棄却点が計算されている（付録参照).

ムード検定

例えば，標本サイズが $n_1 = n_2 = 8$ の場合，

MO	172	182	194	206	222	236	248	262
正確	.4844	.3934	.2856	.1936	.0984	.0496	.0239	.0078
近似	.4796	.3792	.2693	.1782	.0914	.0454	.0228	.0092

となる．現実問題として，正規近似を用いるためには $n_1 \geq 20$，$n_2 \geq 20$は欲しいところである．

正規分布を仮定した下，F 検定に対するムード検定の漸近相対効率は $15/2\pi^2 = 0.76$ であり，自由度 2 の t 分布を仮定した下では，ムード検定が局所最強力検定となる．

ムード検定：例

2 種類の睡眠薬 Π_1，Π_2 の効能の比較のため，24 匹のマウスをランダムに 2 群に分けて実験を行なった．ただし，それぞれの効能の強さはほぼ同じである．

睡眠薬 Π_1 に対する反応は個体差が大きく関係し，睡眠薬 Π_2 は個体差の影響は少ないことが予想されている．

睡眠薬による個体差にばらつきがあるか調べるため，

$$\begin{aligned} &\text{帰無仮説 } H_0 \quad : \quad F_1(x) = F_2(x) \\ &\text{対立仮説 } H_{21} \quad : \quad F_2(x) = F_1(\Delta x), \quad \Delta \neq 1, \quad \Delta > 0 \end{aligned}$$

を検定する．

ムード検定：例

それぞれの睡眠薬を同じ濃度に希釈し，モルモットが眠るまでの時間を観測したところ下表のような結果を得た.

モルモットが眠るまでの時間（秒）

Π_1	45	60	68	47	66	40	41	55	30	58	51	31
Π_2	50	56	48	57	53	54	61	42	46	39	49	52

このデータに対して，大きさの順に並べたものが次表である.

ただし，睡眠薬 Π_1 の観測値は太字で表すことにする.

ムード検定：例

順位付けられたデータ

$Y_{(i)}$	**30**	**31**	39	**40**	**41**	42	**45**	46	**47**	48	49	50
順位	1	2	3	4	5	6	7	8	9	10	11	12
Z_i	1	1	0	1	1	0	1	0	1	0	0	0
$Y_{(i)}$	**51**	52	53	54	**55**	56	57	**58**	**60**	61	**66**	**68**
順位	13	14	15	16	17	18	19	20	21	22	23	24
Z_i	1	0	0	0	1	0	0	1	1	0	1	1

この表から，検定統計量 MO は

$$\begin{aligned} \mathrm{MO} &= 132.25 + 110.25 + 72.25 + 56.25 + 30.25 + 12.25 \\ &\quad +0.25 + 20.25 + 56.25 + 72.25 + 110.25 \\ &= 805. \end{aligned}$$

ムード検定：例

検定統計量 MO の期待値と分散は

$$\mathrm{E}[\mathrm{MO}] = 575, \qquad \mathrm{V}[\mathrm{MO}] = 11440$$

で与えられるので

$$\mathrm{T}_{\mathrm{MO}} = \frac{805 - 575}{\sqrt{11440}} = 2.1504$$

となる. 検定統計量 T_{MO} の値は標準正規分布の両側 5% 点の 1.96 より大きいので帰無仮説は棄却される.

ムード検定：R

統計ソフトRにおけるムード統計量は，基準化されていること，すなわち T_{MO} であることに注意されたい.

例では，下記のプログラムを実行することにより，検定統計量 T_{MO} の p 値を求めることができる.

```
data1 <- c(45,60,68,47,66,40,41,55,30,58,51,31)
data2 <- c(50,56,48,57,53,54,61,42,46,39,49,52)
mood.test(data1,data2,paired=FALSE,
alternative = c("two.sided"), conf.int=TRUE,
conf.level=0.95, correct=FALSE)
```

位置と尺度の違いを検定する方法

ラページ型統計量

- 位置母数を検定するのに有効な検定統計量は，尺度母数の違いにはあまり有用ではない.
- 尺度母数を検定するのに有効な検定統計量は，2 つの分布の中央値が等しいという仮定が必要であった.

$X_2 = (X_1 + \Delta_1)/\Delta_2$ という変数変換が行えるとき，

$$\begin{aligned} P(X_2 \leq x) &= P\left(\frac{X_1 + \Delta_1}{\Delta_2} \leq x\right) = P(X_1 \leq \Delta_2 x - \Delta_1) \\ &= F_1(\Delta_2 x - \Delta_1) \end{aligned} \tag{4}$$

となる．ただし，$\Delta_1 \neq 0$，$\Delta_2 > 0, \Delta_2 \neq 1$である.

ラページ型統計量

母集団分布について前提となる知識がない場合には，どのような検定が適切だろうか?

帰無仮説 $H_0 : F_1(x) = F_2(x)$

に対して，

両側対立仮説 $H_{31} : F_1(x) \neq F_2(x)$

とする.

片側対立仮説は，「すべての x に対して」

$H_{32} : F_1(x) \geqslant F_2(x)$, もしくは $H_{33} : F_1(x) \leqslant F_2(x)$

となる.

ラページ型統計量

上記の片側対立仮説の場合，「すべての x に対して」としたが，「ある x に対して」の場合，対立仮説は

$H'_{32} : F_1(x) > F_2(x)$ もしくは $H'_{33} : F_1(x) < F_2(x)$

となる.

対立仮説 H_{31} に対して，Lepage (1971) は，

ウィルコクソン順位和検定

および

アンサリー・ブラッドレー検定 (Ansari and Bradley, 1960)

に基づく平方和検定統計量を提案した.

ラページ型統計量

ラページ統計量が提案されて以降,

- Pettitt (1976)
- Büning and Thadewald (2000)
- Büning (2002)
- Neuhäuser (2000)
- Murakami (2007)

など，様々なラページ型検定統計量が提案されている.

ラページ型統計量

最も有名なラページ型統計量は，Pettitt (1976) によって提案された

$$\mathrm{LP} = \left(\frac{\mathrm{W} - \mathrm{E[W]}}{\sqrt{\mathrm{V[W]}}} \right)^2 + \left(\frac{\mathrm{MO} - \mathrm{E[MO]}}{\sqrt{\mathrm{V[MO]}}} \right)^2$$

である.

まず，ウィルコクソン順位和検定とムード検定が無相関となることを示す.

(1)式より，2 つの線形順位和統計量のスコア関数を a_i および b_i と定義する.

ラページ型統計量

Randles and Hogg 条件 (1971)

2 つの線形順位和統計量のスコア関数がそれぞれ

$$a_i + a_{N+1-i} = k, \quad i = 1, 2, \ldots, N$$

および

$$b_i = b_{N+1-i}, \quad i = 1, 2, \ldots, N$$

を満たすとき，線形順位統計量 T および T^* は，帰無仮説の下で無相関になる．ただし，k は定数とする．

ラページ型統計量

ウィルコクソン順位和検定とムード検定は，それぞれ (6) 式と (7) 式を満たすことから無相関となる.

ウィルコクソン順位和検定とムード検定はチェルノフ・サーベージの漸近正規性 (Chernoff and Savage, 1958) を満たすことから，ラページ型検定統計量 LP の極限分布は自由度 2 の χ^2 分布となる.

ラページ型統計量：例

昆虫の生息地域によって成長に違いがあるか検定する.

北の方に生息する昆虫 7 匹（グループ Π_1）と南の方に生息する昆虫 7 匹（グループ Π_2）をある方法で抽出して，昆虫の成長の違いを調査した.

検定する仮説として，

帰無仮説: 北方に生息する昆虫と南方に生息する昆虫の成長の分布は等しい

対立仮説: 北方に生息する昆虫と南方に生息する昆虫の成長の分布は等しくない

について考える.

ラページ型統計量：例

下表のようなデータを得た.

昆虫の大きさ (mm)

Π_1	191	173	188	163	184	200	174
Π_2	211	185	201	195	189	199	180

このデータに対してラページ型検定を行う.

ラページ型統計量の計算のため，2 つの母集団からの標本を結合して昇順に並べ，Π_1 の観測値を太字で表したものが次表である.

ラページ型統計量：例

順位付けられたデータ

$Y_{(i)}$	**163**	**173**	**174**	180	**184**	185	**188**
順位	1	2	3	4	5	6	7
Z_i	1	1	1	0	1	0	1
$Y_{(i)}$	189	**191**	195	199	**200**	201	211
順位	8	9	10	11	12	13	14
Z_i	0	1	0	0	1	0	0

この表より，ラページ検定統計量 LP は

$$\mathrm{LP} = 2.97551 + 0.08163 = 3.0571$$

となり，有意水準 5% の値 5.991 より小さいので帰無仮説を棄却しない.

ラページ型統計量：例

ちなみに，$n_1 = n_2 \leq 13$ の場合には正確な棄却点が求められており，$n_1 = n_2 = 7$ の棄却点は 5.560 である (付録参照).

よって，いずれにしても帰無仮説を棄却することはできない.

バウムガートナー型統計量

対立仮説 (5) に対して，最も知られている検定統計量はコルモゴロフ・スミルノフ検定 (Kolmogorov, 1933; Smirnov, 1939) であろう.

しかし，近年 Neuhäuser et al. (2017) によって `Instead we recommend the modified Baumgartner test and Zhang's test. These tests can be used routinely instead of the widely used Smirnov test, offering generally considerably enhanced power while still controlling type I error rate.'と述べれていることから，本節で取り扱う検定統計量に限定する.

バウムガートナー型統計量

連続な分布関数 $F_1(x)$ から得られる大きさ n_1 の順序統計量を

$$\boldsymbol{X}_1 = (X_{1(1)}, X_{1(2)}, \ldots, X_{1(n_1)})$$

連続な分布関数 $F_2(x)$ から得られる大きさ n_2 の順序統計量を

$$\boldsymbol{X}_2 = (X_{2(1)}, X_{2(2)}, \ldots, X_{2(n_2)})$$

とする.

さらに, $R_1 < \cdots < R_{n_1}$ と $H_1 < \cdots < H_{n_2}$ は大きさの順に並べた $\boldsymbol{X}_1$ と $\boldsymbol{X}_2$ の順位を表わす.

バウムガートナー型統計量

Baumgartner et al. (1998) によって，検定統計量

$$\mathrm{B} = \frac{1}{2}(B_{X_1} + B_{X_2})$$

が提案されている．ただし，

$$B_{X_1} = \frac{1}{n_2 N} \sum_{i=1}^{n_1} \frac{(R_i - \frac{N}{n_1} i)^2}{\frac{i}{n_1+1}(1 - \frac{i}{n_1+1})},$$

$$B_{X_2} = \frac{1}{n_1 N} \sum_{j=1}^{n_2} \frac{(H_j - \frac{N}{n_2} j)^2}{\frac{j}{n_2+1}(1 - \frac{j}{n_2+1})}$$

である．

バウムガートナー型統計量

クラーメル・フォン ミーゼス検定（Cramér, 1928; von Mises, 1931; Anderson, 1962）と比較すると，バウムガートナー統計量はクラーメル・フォン ミーゼス検定よりも分布の裾を強調する検定統計量となっている.

この検定統計量の検出力は，位置母数の違いに対してウィルコクソン順位和検定の検出力と近く，尺度母数などの違いに対してコルモゴロフ・スミルノフ検定やクラーメル・フォン ミーゼス検定よりも検出力が高くなることがシミュレーション実験により示されている (Baumgartner et al., 1998; Murakami, 2006).

バウムガートナー型統計量

R_i と H_j の期待値と分散はそれぞれ

$$
\begin{aligned}
\mathrm{E}[R_i] &= \frac{N+1}{n_1+1} i, \\
\mathrm{V}[R_i] &= \frac{i}{n_1+1}\left(1-\frac{i}{n_1+1}\right)\frac{n_2(N+1)}{n_1+2} \\
\mathrm{E}[H_j] &= \frac{N+1}{n_2+1} j, \\
\mathrm{V}[H_j] &= \frac{j}{n_2+1}\left(1-\frac{j}{n_2+1}\right)\frac{n_1(N+1)}{n_2+2}
\end{aligned}
$$

となる.

バウムガートナー型統計量

Murakami (2006) によって，検定統計量

$$\mathrm{B}^* = \frac{1}{2}(B^*_{X_1} + B^*_{X_2})$$

が提案されている．ただし，

$$B^*_{X_1} = \frac{1}{n_1}\sum_{i=1}^{n_1} \frac{(R_i - \frac{N+1}{n_1+1}i)^2}{\frac{i}{n_1+1}(1-\frac{i}{n_1+1})\frac{n_2(N+1)}{n_1+2}},$$

$$B^*_{X_2} = \frac{1}{n_2}\sum_{j=1}^{n_2} \frac{(H_j - \frac{N+1}{n_2+1}j)^2}{\frac{j}{n_2+1}(1-\frac{j}{n_2+1})\frac{n_1(N+1)}{n_2+2}}$$

である．

バウムガートナー型統計量

バウムガートナー型統計量を用いる利点のひとつとして，$\Delta \neq 0$ の条件の下で，

$$F_2(x) = F_1(x - \Delta) \text{ と } F_2(x) = F_1(x + \Delta)$$

に対するバウムガートナー型統計量の検出力が等しくなる (Murakami, 2008;Neuhäuser et al., 2017).

n_1 と n_2 が小さい場合には，検定統計量 B^* の棄却点が正確に計算されている (付録参照).

バウムガートナー型統計量

ラページ型検定統計量を説明したデータを用いて，バウムガートナー型統計量を説明する.

知りたい仮説は

帰無仮説: 北方に生息する昆虫と南方に生息する昆虫の成長の分布は等しい

対立仮説: 北方に生息する昆虫と南方に生息する昆虫の成長の分布は等しくない

であった.

バウムガートナー型統計量

そのデータを昇順に並べ，Π_1 の観測値を太字で表した表より，バウムガートナー型統計量 B^* は

$$\mathrm{B}^* = \frac{1}{2}\left(\frac{14739}{8575} + \frac{3191}{1715}\right) \approx 1.78974$$

となり，有意水準 5% の棄却点 2.493 より小さいので棄却することは出来ない.

$n_1 = n_2 = 7$ のときの正確な棄却点は 2.804 であり，いずれにしても仮説を棄却することはできない.

バウムガートナー型統計量：R

統計ソフトRにおけるバウムガートナー型統計量は，下記のプログラムを実行することにより，バウムガートナー型統計量を求めることができる.

```
data1 <- c(191,173,188,163,184,200,174)
data2 <- c(211,185,201,195,189,199,180)
murakami_stat(data1, data2, flavor = 1L)
```

ただし，パッケージ ``BWStest'' をインストール必要がある.

バウムガートナー型統計量：R

また，flavor を以下のように変更することによって，
様々な検定統計量を実行することが可能である.

- 0L: バウムガートナー統計量 (Baumgartner et al., 1998)
- 1L: バウムガートナー型統計量 (Murakami, 2006)
- 2L: バウムガートナー型統計量 $(\mathrm{Neuhäuser},\ 2001)$
- 3L: バウムガートナー型統計量 (Murakami, 2012 の B_3 統計量)
- 4L: バウムガートナー型統計量 (Murakami, 2012 の B_4 統計量)
- 5L:バウムガートナー型統計量 (Murakami, 2012 の B_5 統計量)

バウムガートナー型統計量：R

バウムガートナー統計量 B は片側対立仮説にはあまり有用ではないことから，Neuhäuser (2001) によって片側バウムガートナー型統計量が提案されている.

2 項分布に対して，バウムガートナー型統計量に基づく検定統計量が Shan et al. (2013) によって提案されている.

Murakami (2006) によって多標本検定統計量へと拡張され，Murakami et al. (2009) や Miyazaki and Murakami (2019) によって近似分布が導出されている.

漸近相対効率

漸近相対効率

ふたつの線形順位検定を T と T^* とし，有意水準は等しいものと仮定する.

また，対立仮説のパラメータ Δ に対する検出力関数を，それぞれ $\psi(\Delta)$ と $\psi^*(\Delta)$ とする．ただし，Δ_0 は帰無仮説の下でのパラメータを表わすものとする.

さらに，

$$\lim_{N\to\infty} \psi(\Delta_N) = \lim_{N\to\infty} \psi^*(\Delta_N)$$

とする.

漸近相対効率

このとき，漸近相対効率は

$$\mathrm{ARE}(\mathrm{T}, \mathrm{T}^*) = \lim_{N \to \infty} \frac{N^*}{N}$$

によって与えられる.

例えば，$\lim_{N,N^* \to \infty} \frac{N^*}{N} = 0.95$ ならば，検定統計量 T^* が 95 個のデータで得ることのできる検出力と同等の検出力を得るためには，検定統計量 T では 100 個のデータが必要であると言える.

漸近相対効率

ここで，検定統計量 に対して次の仮定が成り立つこととする.

1. $\dfrac{d\mathrm{E}(\mathrm{T})}{d\Delta}$ が存在し，かつ $\Delta = 0$ で連続である.

   また，$\dfrac{d^r \mathrm{E}(T_N(Z))}{d\Delta^r}$, $r = 2, 3, \ldots,$ は Δ_0 において

   ゼロとなる.

2. $\lim\limits_{N\to\infty} \dfrac{d\mathrm{E}(\mathrm{T})/d\Delta|_{\Delta=\Delta_0}}{\sqrt{N}\sigma(\mathrm{T})|_{\Delta=\Delta_0}} = C$ となる定数が存在する.

漸近相対効率

3．次の条件を満たす対立仮説の系列 $\{\Delta_N\}$ が存在する．

$$\Delta_N = \Delta_0 + \frac{\delta}{\sqrt{N}}, \quad \delta > 0,$$

$$\lim_{N\to\infty} \frac{d\mathrm{E}(\mathrm{T})/d\Delta|_{\Delta=\Delta_N}}{d\mathrm{E}(\mathrm{T})/d\Delta|_{\Delta=\Delta_0}} = 1,$$

$$\lim_{N\to\infty} \frac{\sigma(\mathrm{T})|_{\Delta=\Delta_N}}{\sigma(\mathrm{T})|_{\Delta=\Delta_0}} = 1.$$

漸近相対効率

4. $$\lim_{N\to\infty} \Pr_{\Delta}\left(\frac{\mathrm{T}-\mathrm{E}(\mathrm{T})|_{\Delta=\Delta_N}}{\sigma(\mathrm{T})|_{\Delta=\Delta_0}} \le x\right) = \Phi(x)$$

が成立し, $\Delta = 0$ の近傍で一様である.

5. $$\lim_{N\to\infty} \Pr(\mathrm{T} \le t_{N,\alpha}; \Delta = \Delta_0) = \alpha, \quad 0 < \alpha < 1.$$

この T が線形順位和検定統計量であるということと, 仮定 1 ～ 5 は無関係である. 同様に, 検定統計量 T^* にも適用可能である.

漸近相対効率

この仮定が満たされていれば，対立仮説の系列に対する線形順位統計量の漸近検出力は

$$\lim_{N\to\infty} \Pr\left(\frac{\mathrm{T} - \mathrm{E}(\mathrm{T})|_{\Delta=\Delta_N}}{\sigma(\mathrm{T})|_{\Delta=\Delta_0}} \geq \frac{t_{N,\alpha} - \mathrm{E}(\mathrm{T})|_{\Delta=\Delta_N}}{\sigma(\mathrm{T})|_{\Delta=\Delta_0}} \right)$$
$$= 1 - \Phi(x_\alpha - \delta C)$$

によって与えられる．ただし，x_α は $1 - \Phi(x_\alpha) = \alpha$ を満たすものとする．

漸近相対効率

2 つの検定統計量は，上記の仮定 1 ～ 5 を満たす検定統計量とする．このとき，漸近相対効率 (Asymptotic Relative Efficiency) は

$$\mathrm{ARE}(\mathrm{T}, \mathrm{T}^*) = \lim_{n\to\infty} \frac{e(\mathrm{T})}{e(\mathrm{T}^*)}$$

で与えられる．ただし

$$e(\mathrm{T}) = \frac{\left[\frac{d\mathrm{E}(\mathrm{T})}{d\Delta}\right]^2\Big|_{\Delta=\Delta_0}}{\sigma^2(\mathrm{T})|_{\Delta=\Delta_0}}, \qquad e(\mathrm{T}^*) = \frac{\left[\frac{d\mathrm{E}(\mathrm{T}^*)}{d\Delta}\right]^2\Big|_{\Delta=\Delta_0}}{\sigma^2(\mathrm{T}^*)|_{\Delta=\Delta_0}}$$

である．

1 元配置分散分析

クラスカル・ウォリス検定

$\{\boldsymbol{X}_{ij}|i=1,\ldots,k,\ j=1,\ldots,n_i\}$ を分布関数 $F_i(x)$ から得られる大きさ $n_1,\ldots,n_k$ の独立な確率標本とする．ただし，$N=n_1+n_2+\cdots+n_k$ である．

帰無仮説　$H_0: F_1(x)=F_2(x)=\cdots=F_k(x)$

に対して，

対立仮説 $H_4: H_0$ が成立しない

とする．

この対立仮説は一般的すぎるため，確率標本は連続で未知な分布関数 $F_i(x-\Delta_i)$ から得られるものとする．

クラスカル・ウォリス検定

このとき，

帰無仮説　$H_0 : \Delta_1 = \Delta_2 = \cdots = \Delta_k$

に対して，

対立仮説　$H_{41} : \Delta_i \neq \Delta_j$

となる．

ただし，少なくともひとつの $i \neq j$ が成立するものとする．

クラスカル・ウォリス検定

帰無仮説が棄却されるということは,

$$\Delta_1 \neq \Delta_2 = \Delta_3 = \cdots = \Delta_k$$

$$\vdots$$

$$\Delta_1 = \Delta_2 = \cdots = \Delta_{k-1} \neq \Delta_k$$

$$\Delta_1 \neq \Delta_2 \neq \Delta_3 = \cdots = \Delta_k$$

$$\vdots$$

$$\Delta_1 \neq \Delta_2 \neq \cdots \neq \Delta_k$$

の全ての場合を含んでいることに注意されたい.

また，帰無仮説が棄却されても，何番目の確率標本が他の確率標本と違うのか等については何も言えない.

クラスカル・ウォリス検定

すべての確率標本を統合し, $N = n_1 + \cdots + n_k$ の観測値を昇順に並べ，それぞれに順位をつける．このとき，R_{ij} を $\boldsymbol{X}_{ij}$ の順位としたのが下表である.

確率標本の順位

標本番号	順位					合計	平均
1	R_{11}	R_{12}	R_{13}	$\cdots$	R_{1n_1}	$R_{1\cdot}$	$\overline{R_1}$
2	R_{21}	R_{22}	R_{23}	$\cdots$	R_{2n_2}	$R_{2\cdot}$	$\overline{R_2}$
$\vdots$	$\vdots$	$\vdots$	$\vdots$		$\vdots$	$\vdots$	$\vdots$
k	R_{k1}	R_{k2}	R_{k3}	$\cdots$	R_{kn_k}	$R_{k\cdot}$	$\overline{R_k}$

クラスカル・ウォリス検定

Kruskal and Wallis (1952) は，検定統計量

$$
\begin{aligned}
\mathrm{H} &= \frac{12}{N(N+1)} \sum_{i=1}^{k} \frac{1}{n_i} \left\{ \sum_{j=1}^{n_i} R_{ij} - \frac{n_i(N+1)}{2} \right\}^2 \\
&= \frac{12}{N(N+1)} \sum_{i=1}^{k} \frac{1}{n_i} \left(\sum_{j=1}^{n_i} R_{ij} \right)^2 - 3(N+1)
\end{aligned}
$$

を提案した.

- $k = 2$ とすると，ウィルコクソン順位和検定と本質的に同じとなることから，ウィルコクソン順位和検定を拡張した検定統計量であり，極限分布は，自由度 $k-1$ の χ^2 分布に従う.

クラスカル・ウォリス検定：例

Cawson et al. (1974) によって与えられたデータを用いて，クラスカル・ウォリス検定を紹介する.

肘前静脈コルチゾール濃度

Π_1	262	307	211	323	454	339	304	154	287	356
Π_2	465	501	455	355	468	362				
Π_3	343	772	207	1048	838	687				

Π_1，Π_2，Π_3 の中央値に違いがあるか，すなわち，

$$H_0 \quad : \quad \Delta_1 = \Delta_2 = \Delta_3$$

$$H_{41} \quad : \quad \Delta_i \neq \Delta_j, \quad 1 \leq i,\ j \leq 3$$

に対して，有意水準 5% の検定を行う.

クラスカル・ウォリス検定：例

データを昇順に並べ，順位をつける．

順位付けられたデータ

	順位										合計
Π_1	4	7	3	8	14	9	6	1	5	12	69
Π_2	16	18	15	11	17	13					90
Π_3	10	20	2	22	21	19					94

したがって，

$$\mathrm{H} = \frac{12}{22(22+1)}\left(\frac{69^2}{10} + \frac{90^2}{6} + \frac{94^2}{6}\right) - 3(22+1) = 9.2316$$

となる．極限分布は χ_2^2 分布に従うので，棄却点は 5.991 となることから，帰無仮説を棄却できる．

クラスカル・ウォリス検定：R

統計ソフトRにおけるクラスカル・ウォリス検定は，先程ののの例では，下記のプログラムを実行することにより，クラスカル・ウォリス検定の p 値を求めることが出来る.

```
data1 <- c(262,307,211,323,454,339,304,154,
           287,356)
data2 <- c(465,501,455,355,468,362)
data3 <- c(343,772,207,1048,838,687)
kruskal.test(list(data1,data2,data3))
```

ヨンキー・タプストラ検定

確率標本 $\boldsymbol{X}_{ij},\ i=1,\ldots,k,\ \ j=1,\ldots,n_i$ は分布関数 $F_i(x-\Delta_i)$ から得られるものとする.

分布関数には連続性だけが仮定され，それ以外については分からないものとする.

例えば，ある薬の投与量を k 段階に分けて効果を調べるとき，投与量を増やすにつれて効果の増加 (もしくは減少) が見込まれる場合，順序制約がある仮説について考える必要がある.

ヨンキー・タプストラ検定

帰無仮説 $H_0 : \Delta_1 = \Delta_2 = \cdots = \Delta_k$

に対して

対立仮説 $H_{42} : \Delta_1 \leq \Delta_2 \leq \cdots \leq \Delta_k,$

とする.

ただし，少なくとも一つの不等号が成立する.

以下の対立仮説についても同様である.

対立仮説 $H_{43} : \Delta_1 \geq \Delta_2 \geq \cdots \geq \Delta_k.$

ヨンキー・タプストラ検定

相異なる 2 つの標本に対して, $\boldsymbol{X}_i = (X_{i1}, X_{i2}, \ldots, X_{in_i})$ と $\boldsymbol{X}_j = (X_{j1}, X_{j2}, \ldots, X_{jn_j})$ の大小比較をし, $X_{is} < X_{jt}$ を満たす組の個数を数える.

すなわち,

$$J_{ij} = \sum_{s=1}^{n_i} \sum_{t=1}^{n_j} U(X_{is}, X_{jt})$$

である. ただし,

$$U(x, y) = \begin{cases} 1: & x < y \\ 0: & x \geq y \end{cases}$$

とする.

ヨンキー・タプストラ検定

Terpstra (1952) と Jonckheere (1954) によって,
$k(k-1)/2$ 個のすべての組み合わせについて J_{ij} を求め,
その和とした検定統計量

$$\mathrm{JT} = \sum_{i=1}^{k} \sum_{j=i+1}^{k} \sum_{s=1}^{n_i} \sum_{t=1}^{n_j} U(X_{is}, X_{jt})$$

が提案された.

ヨンキー・タプストラ検定

帰無仮説の下での検定統計量 JT の期待値と分散は,

$$\mathrm{E}[\mathrm{JT}] = \frac{1}{4}\left(N^2 - \sum_{j=1}^{k} n_j^2\right),$$

$$\mathrm{V}[\mathrm{JT}] = \frac{1}{72}\left\{N^2(2N+3) - \sum_{j=1}^{k} n_j^2(2n_j+3)\right\}$$

で与えられる.

ヨンキー・タプストラ検定：R

クラスカル・ウォリス検定で用いたデータに対して，Π_1，Π_2，Π_3 の平均に傾向性があるか調べるため，有意水準 5% 検定を行う.

ヨンキー・タプストラ検定は, Π_1 と Π_2 , Π_1 と Π_3 , Π_2 と Π_3 で比較を行うため，マン・ウィットニーの U 検定と同様の考え方より

$$\mathrm{JT} = 57 + 49 + 24 = 130$$

となる.

ヨンキー・タプストラ検定：R

E[JT]=78，V[JT]=269 となることから，

$$\frac{130 - 78 - 0.5}{\sqrt{269}} \approx 3.14$$

を得る．よって，帰無仮説を棄却することができる．

ヨンキー・タプストラ検定：R

統計ソフトRにおけるヨンキー・タプストラ検定は，先程の例では，下記のプログラムを実行することにより，ヨンキー・タプストラ検定の p 値を求めることが出来る.

```
data1 <- c(262,307,211,323,454,339,304,154,
           287,356)
data2 <- c(465,501,455,355,468,362)
data3 <- c(343,772,207,1048,838,687)
gdata <- c(rep(1,length(data1)),
rep(2, length(data2)), rep(3, length(data3)))
dataA <- c(data1, data2, data3)
jonckheere.test(dataA, gdata, "increasing")
```

ヨンキー・タプストラ検定：R

- gdata は標本番号を表わすベクトルである.
- パッケージ 「clinfun」をインストールする必要がある.
- 両側検定の場合，SIDE を 「two.sided」とする必要がある.
- 片側検定の場合 SIDE を 「increasing」もしくは「decreasing」とする必要がある.

参考文献

参考文献

Anderson, T.W. (1962). On the distribution of the two-sample Cramér-von Mises Criterion. The Annals of Mathematical Statistics 33, 1148--1159.

Ansari, A.R. and Bradley, R.A. (1960). Rank sum tests for dispersion. The Annals of Mathematical Statistics 31, 1174--1189.

Baumgartner, W., Weiß, P. and Schindler, H. (1998). A nonparametric test for the general two-sample problem. Biometrics 54, 1129--1135.

Chernoff, H. and Savage, I.R. (1958). Asymptotic normality and efficiency of certain nonparametric test statistics. The Annals of Mathematical Statistics 29, 972--994.

Kolmogorov, A. (1933). Sulla determinazione empirica di una legge di distribuzioe. Giornale dell'Instituto Italiano degla Attuari 4, 83--91.

Lepage, Y. (1971). A combination of Wilcoxon's and Ansari-Bradley's statistics, Biometrika 58, 213--217.

参考文献

Mann, H.B. and Whitney, D.R. (1948). On a test of whether one of two random variables is stochastically larger than the other. The Annals of Mathematical Statistics 18, 50--60.

Mood, A.M. (1954). On the Asymptotic Efficiency of Certain Nonparametric Two-Sample Tests. The Annals of Mathematical Statistics 25, 514--522.

Murakami, H. (2006). A k-sample rank test based on the modified Baumgartner statistic and its power comparison. Journal of the Japanese Society of Computational Statistics 19, 1--13.

Murakami, H. (2008). A multisample rank test for location-scale parameters. Communications in Statistics - Simulation and Computation 37, 1347--1355.

Murakami, H. (2012). Modified Baumgartner statistics for the two-sample and multisample problems: A numerical comparison. Journal of Statistical Computation and Simulation 82, 711--728.

村上　秀俊. (2015). ノンパラメトリック法. 朝倉書店.

参考文献

Neuhäuser, M. (2001). One-side two-sample and trend tests based on a modified Baumgartner-Weiss-Schindler statistic. Journal of Nonparametric Statistics 13, 729--739.

Neuhäuser, M., Welz, A. and Ruxton, G.D. (2017). Statistical tests for the comparison of two samples: The general alternative. Communications in Statistics - Simulation and Computation 47, 903--909.

Pettitt, A.N. (1976). A two-sample Anderson-Darling rank statistic, Biometrika 63, 161--168.

Randles, R.H. and Hogg, R.V. (1971). Certain uncorrelated and independent rank statistics, Journal of the American Statistical Association 66, 569--574.

Smirnov, N.V. (1939). Estimate of deviation between empirical distribution functions in two independent samples. Bulletin of Moscow University 2, 3--16.

Wilcoxon, F. (1945). Individual comparisons by ranking methods. Biometrics 1, 80--83.