

拡張二変量順位和統計量を用いたクロスオーバー 臨床試験における薬効比較法のSASマクロ

川口 淳
久留米大学バイオ統計センター

SAS Macro for Application of Extensions of Bivariate Rank Sum Statistics to the Crossover Design to Compare Two Treatments

Atsushi Kawaguchi
Biostatistics Center, Kurume University

発表内容

- はじめに(動機, 目的)
- 方法の説明
 - » 数理的な説明, シミュレーション結果
- SASマクロの説明
 - » 入手方法, 使用例, 出力
- まとめ

動機となった臨床試験データ

- 眼科領域における双方性クロスオーバー臨床試験
- 試験薬 (T) とプラセボ (P) の薬効比較
 - 2 処置期 (第 1 期 = 左側, 第 2 期 = 右側)
 - 4 処置系列群 T:T, T:P, P:T, P:P (Balaam デザイン)
 - 被験者は無作為に 4 群に割り付けられる.
- 主要評価項目: 重症度 (順序反応変数)
 - (よい) 0, 0.5, 1.0, 1.5, 2.0, 2.5, 3.0 (悪い)

目的

- データの特徴を活かした検出力のより高い検定方法 (Kawaguchi et al., 2009 Biometrics) に対する SAS マクロの開発
- データの特徴 (提案法での工夫)
 - Balaam デザイン (2x4 セルから自由度縮小)
 - 順序反応変数 (順位に基づくノンパラメトリック検定法)
 - 左右の相関 (順位ベクトルの分散共分散行列, 重み付き最小二乗推定量)
- SAS マクロの特徴
 - 単純な入力 (4 つの変数を指定)
 - 出力表 (SAS PROC 風. 検定統計量, p 値, 適合度指標)

記号の定義

- $i=1, 2, 3, 4$: 処置系列群 T:T, T:P, P:T, P:P を表すインデックス
- $j=1, 2$: 左側, 右側を表すインデックス.
- $k=1, 2, \dots, n_i$: i 群に n_i 人いる被験者のインデックス.
- $N = n_1 + n_2 + n_3 + n_4$ (試験全体での被験者数)
- y_{ijk} : i 群 j 側における k 番目の被験者の反応変数.
- R_{ijk} : 全被験者の j 側においてとられた反応変数の順位のうち, i 群の k 番目のもの.

コンセプトモデル

	$j=1$ (左側)	$j=2$ (右側)
$i=1$ (T:T)	$\mu_1 + \tau$	$\mu_2 + \tau$
$i=2$ (T:P)	$\mu_1 + \tau$	μ_2
$i=3$ (P:T)	μ_1	$\mu_2 + \tau$
$i=4$ (P:P)	μ_1	μ_2

- μ_j は ($j=1, 2$), 左右それぞれにおける P の位置パラメータ
- τ は T と P の効果の差を表すパラメータ

帰無仮説 $H_0: \tau = 0$

標準化順位平均

$$\bar{\mathbf{R}} = (\bar{R}_{11}, \bar{R}_{12}, \bar{R}_{21}, \bar{R}_{22}, \bar{R}_{31}, \bar{R}_{32}, \bar{R}_{41}, \bar{R}_{42})'$$

$$\text{ただし, } \bar{R}_{ij} = \sum_{k=1}^{n_i} R_{ijk} / n_i (N+1)$$

- \mathbf{R} の線形変換による検定統計量を考える
 - \mathbf{R} の H_0 の下での分散共分散行列より, 各統計量の分散を算出できる(左右の相関が考慮された).
- 順位付けは左右別々に行う
 - パラメータ μ_1 と μ_2 に関しては任意に扱える

二変量クラスカル-ワリス統計量

	j=1 (左側)	j=2 (右側)
i=1 (T:T)	$\mu_1 + \tau$	$\mu_2 + \tau$
i=2 (T:P)	$\mu_1 + \tau$	μ_2
i=3 (P:T)	μ_1	$\mu_2 + \tau$
i=4 (P:P)	μ_1	μ_2

$$\mathbf{d} = (\bar{R}_{11} - \bar{R}_{41}, \bar{R}_{12} - \bar{R}_{42}, \bar{R}_{21} - \bar{R}_{41}, \bar{R}_{22} - \bar{R}_{42}, \bar{R}_{31} - \bar{R}_{41}, \bar{R}_{32} - \bar{R}_{42})' = \mathbf{C}_0 \bar{\mathbf{R}}$$

$$Q_0 = \mathbf{d}' \mathbf{V}_{\mathbf{d},0}^{-1} \mathbf{d} \sim \chi_6^2$$

自由度6の統計量 \Rightarrow パラメータ1つの H_0 に対しては不適切

左右別ウィルコクソン順位和統計量

	j=1 (左側)	j=2 (右側)
i=1 (T:T)	$\mu_1 + \tau$	$\mu_2 + \tau$
i=2 (T:P)	$\mu_1 + \tau$	μ_2
i=3 (P:T)	μ_1	$\mu_2 + \tau$
i=4 (P:P)	μ_1	μ_2

$$\text{左側 } F_1 = \left\{ \frac{n_1 \bar{R}_{11} + n_2 \bar{R}_{21}}{n_1 + n_2} \right\} - \left\{ \frac{n_3 \bar{R}_{31} + n_4 \bar{R}_{41}}{n_3 + n_4} \right\}$$

$$\text{右側 } F_2 = \left\{ \frac{n_1 \bar{R}_{12} + n_3 \bar{R}_{32}}{n_1 + n_3} \right\} - \left\{ \frac{n_2 \bar{R}_{22} + n_4 \bar{R}_{42}}{n_2 + n_4} \right\}$$

$$f_1 = (F_1 + F_2)/2 \quad \Rightarrow \quad Q_{ave} = f_1^2 / v_{f,1,0} \sim \chi_1^2$$

左右の相関が考慮されていない

提案する検定統計量構成の流れ

- 左右別ウィルコクソン順位和統計量の平均 f_1 に加え, 0と期待できる比較(例えば, $f_2 = F_1 - F_2$)統計量 $\mathbf{g}=(f_2, f_3, f_4, f_5, f_6)$ を導入する
 - $\mathbf{f} = (f_1, f_2, f_3, f_4, f_5, f_6)'$ は \mathbf{R} の線形変換($\mathbf{f} = \mathbf{A}\mathbf{C}_0\mathbf{R}$)
 - \mathbf{f} の分散共分散行列が \mathbf{R} のから求める事ができる
- 0制約も組み込まれたモデルを立て, τ の推定量を重み付き最小二乗推定法により求める
 - \mathbf{f} の分散共分散行列の逆行列を利用(左右の相関を考慮)

検定統計量の構成

の統計量
0と期待される

線形モデル

$$\begin{bmatrix} f_1 \\ f_2 \\ f_3 \\ f_4 \\ f_5 \\ f_6 \end{bmatrix} \hat{=} \begin{bmatrix} 1 \\ 0 \\ 0 \\ 0 \\ 0 \\ 0 \end{bmatrix} b$$

重み付き最小二乗推定量

$$\mathbf{x} = (1 \ 0 \ 0 \ 0 \ 0 \ 0)'$$

$$b = (\mathbf{x}' \mathbf{V}_{f,0}^{-1} \mathbf{x})^{-1} \mathbf{x}' \mathbf{V}_{f,0}^{-1} \mathbf{f}$$

$$v_{b,0} = (\mathbf{x}' \mathbf{V}_{f,0}^{-1} \mathbf{x})^{-1}$$

検定統計量

$$Q_b = b^2 / v_{b,0} \sim \chi_1^2$$

検定統計量の解釈

- 二変量クラスカル-ワリス検定統計量の分解

$$Q_0 = Q_b + Q_1$$

d.f. = 6 検定統計量(d.f.=1) 適合度統計量(d.f.=5)

- 左右の相関(説明のためデータ設定の簡略化)

- $n_1=n_2=n_3=n_4=n$, 反応変数にタイがないと仮定
- H_0 の下での左と右の分散は同一 v_0 , 相関は ρ_0

$$Q_b = \frac{2n(f_1 - \rho_0 f_2)^2}{v_0(1 - \rho_0^2)} \qquad Q_{ave} = \frac{2nf_1^2}{v_0}$$

相関が強まると分散減少 ⇨ 高検出力期待

シミュレーションの目的と設定

【目的】 第1種の過誤と検出力を調べる

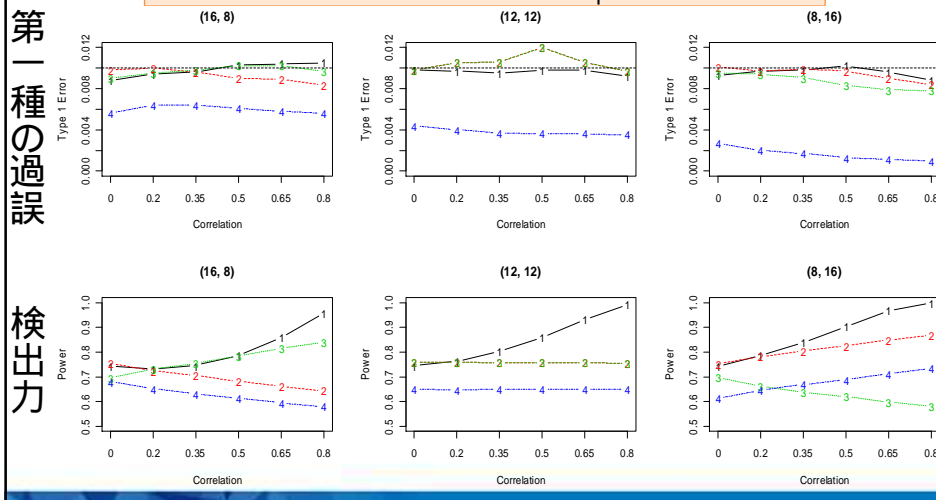
- 2変量の正規乱数 $(y_{11k}, y_{12k}) \sim N(\mu_{1i}, \mu_{2i})$
 - 【分散構造】 4群間では同一の2x2分散共分散行列を仮定する.
 - 分散は1にし, 相関は $\rho = 0, 0.2, 0.35, 0.5, 0.65, 0.8$ (6パターン)を考える.
 - 【平均構造】 コンセプトモデル(1)において, $\mu_1 = 0, \mu_2 = 1$
 - $\alpha = 0$ (第1種の過誤), $\beta = 0.7$ (検出力)

- 【サンプルサイズ】 右表の3パターン(N=48)
- 各 N= 48のデータに対し, 各統計量からp値を算出し, 第一種の過誤または検出力を計算する. 10,000回繰り返す

	サンプルサイズ	
	T:T群とP:P群	T:P群とP:T群
パターン1	16	8
パターン2	12	12
パターン3	8	16

シミュレーション結果

$$1 = Q_b, 2 = Q_{ave}, 3 = Q_{equ}, 4 = Z_c$$



第一種の過誤

検出力

SASマクロの入手方法

- BimetricsのHP (<http://www.biometrics.tibs.org/>) から行う。
 - 左側のメニューから「Paper Information」に進み
 - 「Supplementary Materials」に進む。
 - 「Vol. 65, No. 3 September 2009」に進み、
 - 論文タイトル「Applications of extensions of bivariate rank sum statistics to the crossover design to compare two treatments through four sequence groups」をクリックすれば
 - 「Data and/or Computer Code」というリンクからzipファイルにまとめられたマクロを含むファイル一式を入手できる。
- 著者問い合わせで行う。
 - 久留米大学バイオ統計センター 川口 淳

ファイル一式 (SAS関連のみ)

ファイル名	内容
exampledata.sas7bdat	表1のSAS用データ
EBRST2008.sas	マクロ本体
Example.sas	マクロ実行例
a24Rosner.sas	Rosnerの方法マクロ

マクロ内の処理概要

側毎の反応変数の順位の計算 (PROC RANK)

RANKプロシジャにおいてbyステートメントにperiodを指定



群毎に側別の順位の平均を求める (PROC MEANS)

MEANSプロシジャによりclassステートメントにseqとperiodを指定



検定統計量及びp値の算出 (IML)

クロネッカー積演算子 (@)

行列の積(*)及び逆行列演算(inv)

p値の計算 (PROBCHI)

データセットの準備

example.sas (一部改)

```

/** Specify the Location of Data Set */
libname lb 'C:\';

/** Read Data Set */
data tmp; set lb.exempladata; run;

proc print data=tmp; run;
    
```

OBS	seq	res	period
1	1	2.0	2
2	1	1.0	2
3	1	2.0	2
4	1	1.0	2
5	1	2.5	2
6	1	2.5	2
7	2	1.5	2
8	2	2.5	2
9	2	1.0	2
10	2	3.0	2

● 側を表す変数 (1, 2)
 ● 反応変数
 ● 群を表す変数 (1, 2, 3, 4)

マクロ実行方法

本体ファイル「EBRST2008.sas」を読み込み後

```
%EBRST2008(tmp, res, seq, period);
```

マ
ク
ロ
引
数

データセット名(tmp)

反応変数名(res)

群を表す変数名(seq)

側を表す変数名(period)

実行結果

検定統計量

適合度統計量

Final Output			
	DF	Chi-Square	Pr>ChiSq
Qb	1	5.4314815	0.0197769
Q1	5	3.2048734	0.668434
Qave	1	4.8426529	0.0277642
Qequ	1	3.3443932	0.0674343

⇒ 5%有意水準で、試験薬に効果があることが示唆された。

bivariate Kruskal-Wallis statistic Q0		
DF	Chi-Square	Pr>ChiSq
6	8.6363549	0.1950855

$$Q_0 = Q_b + Q_1$$

f
-0.124929
-0.00907
0.1613393
0.0192744
-0.093921
-0.18424

Randomization based covariance matrix of f						
0.0032229	0.0015327	-0.000064	-1.19E-18	-1.08E-19	-1.08E-18	
0.0015327	0.0060391	-1.46E-18	-0.000074	0.0006485	-7.761E-6	
-0.000064	-3.25E-19	0.0173675	-0.008131	1.518E-18	-4.34E-18	
0.0192744	-7.67E-19	-0.000074	-0.008131	0.0149228	-7.761E-6	0.0010998
-0.093921	1.416E-18	0.0006485	8.962E-19	-7.761E-6	0.0246426	-0.000295
-0.18424	-1.24E-18	-7.761E-6	-1.68E-18	0.0010998	-0.000295	0.0417471

まとめと考察

- 本論文では2期2剤4群のクロスオーバーデザインにおいて2期が同一被験者の左右に対応する強い相関を示す場合のノンパラメトリック検定法についてのSASマクロを紹介した。
 - 同一被験者への同時の処置は眼科領域のみならず皮膚科領域などでも考えられる。
 - シミュレーション研究(相関 ρ → 検出力 \uparrow), 適した方法だと考えられる。
 - SASマクロはデータセットを用意すれば, 4つの変数名を指定して簡単に実行ができる。
 - 実行結果は見やすいように, SAS PROCのものに似たようなデザインにしている。
- 方法論はいくつかの拡張を持っている。
 - 処置の反対側への影響 (Kawaguchi, Koch, and Ramaswamy, 2009 Biometrics)
 - 処置期が2からそれ以上への拡張 (Kawaguchi and Koch, 2010 Journal of Biopharmaceutical Statistics)
 - 層別解析 (Kawaguchi, Koch, and Wang, 2010 To appear in Statistics in Biopharmaceutical Research)