処理効果に順序制約のある乱塊法モデルにおける 対照群との比較を行う閉検定手順

M2017SS004：大畑航平

指導教員：白石高章

1　はじめに

k 標本正規分布モデルにおいて母平均に順序制約がある場合の対照群との多重比較法を Williams(1971,1972) は提案した．また、二元配置モデルにおいて処理効果に順序制約がある場合のノンパラメトリック検定法として Page 検定がある．本研究では処理効果に順序制約がある場合の正規分布を仮定した乱塊法モデルについて考察する．この乱塊法モデルにおいて線形型統計量を基にした多重比較検定法を提案する．さらに，Williams 型多重比較検定法を提案する．これらの多重比較検定法の検出力をシミュレーションにより比較をする．

2　Page 検定

2.1　二元配置モデルの設定

第 i ブロック $(i = 1, 2, \cdots, n)$，第 j 処理 $(j = 1, 2, \cdots, k)$ の確率変数 X_{ij} を $P(X_{ij} \leqq x) = F(x - \beta_i - \tau_j)$ とする．ただし，$F(x)$ は連続型分布関数とする．β_i はブロック効果を表す定数，τ_j は処理効果を表す定数である．すべての X_{ij} は互いに独立とする．R_{ij} は $X_{i1}, \cdots, X_{ik}$ の中での X_{ij} の順位とする．

表 1　二元配置モデル

	第 1 処理	第 2 処理	$\cdots$	第 k 処理
第 1 ブロック	R_{11}	R_{12}	$\cdots$	R_{1k}
第 2 ブロック	R_{21}	R_{22}	$\cdots$	R_{2k}
$\vdots$	$\vdots$	$\vdots$	$\cdots$	$\vdots$
第 n ブロック	R_{n1}	R_{n2}	$\cdots$	R_{nk}

2.2　検定統計量

R_{ij} の平均と分散はそれぞれ

$$E(R_{ij}) = \frac{k+1}{2} \tag{1}$$

$$V(R_{ij}) = \frac{k^2-1}{12} \tag{2}$$

となる．$R_{\cdot j} = \sum_{i=1}^{n} R_{ij}$ とすると (1)，(2) より

$$E(R_{\cdot j}) = \frac{n(k+1)}{2} \tag{3}$$

$$V(R_{\cdot j}) = \frac{n(k^2-1)}{12} \tag{4}$$

を得る．帰無仮説 $H_0 : \tau_1 = \cdots = \tau_k$ vs. 対立仮説 $H_A : \tau_1 \leqq \cdots \leqq \tau_k$ (少なくとも 1 つの $\leqq$ は $\lneqq$ である) の Page 検定統計量 (文献 [1]) は

$$Q \equiv \frac{1}{\sqrt{n}} \sum_{j=1}^{k} \left(j - \frac{k+1}{2} \right) \left(R_{\cdot j} - \frac{n(k+1)}{2} \right) \tag{5}$$

で与えられる．Q の平均は帰無仮説 H_0 の下で $E_0(Q) = 0$ となる．Q の分散は

$$V_0(Q) = \frac{k^2(k^2-1)(k+1)}{144} \tag{6}$$

である．文献 [1] より帰無仮説 H_0 の下で

$$\frac{Q - E_0(Q)}{\sqrt{V_0(Q)}} \xrightarrow{L} N(0,1) \tag{7}$$

が成り立つ．

2.3　検定方式

標準正規分布の上側 $100\alpha\%$ 点を $z(\alpha)$ とすると，$1 - \Phi(z(\alpha)) = \alpha$ が成り立つ．ただし，$\Phi(\cdot)$ は標準正規分布の分布関数である．水準 α の検定は，

$$\begin{cases} Q \geqq z(\alpha)\sqrt{V(Q)} \text{のとき帰無仮説 } H_0 \text{を棄却する} \\ Q < z(\alpha)\sqrt{V(Q)} \text{のとき帰無仮説 } H_0 \text{を棄却しない} \end{cases}$$

である．

3　Page 型検定法

3.1　乱塊法モデルの設定

第 i ブロック $(i = 1, 2, \cdots, n)$，第 j 処理 $(j = 1, 2, \cdots, k)$ の確率変数 Y_{ij} を

$$Y_{ij} \equiv \mu + B_i + \tau_j + e_{ij} \tag{8}$$

で表現される乱塊法モデルを考える．ただし，μ は総平均，B_i はブロック効果を表す変量で $B_i \sim N(0, \{\sigma_B\}^2)$，$\tau_j$ は処理効果を表すパラメータで $\sum_{j=1}^{k} \tau_j = 0$，$e_{ij}$は誤差を表す変量で $e_{ij} \sim N(0, \sigma^2)$ である．$\{e_{ij} | i = 1, \cdots, n; j = 1, \cdots, k\}$ は互いに独立，$B_1, \cdots, B_n$は互いに独立，e_{ij} と B_iは互いに独立とする．このとき，Y_{ij} は各ブロックごとに独立となる．

各標本平均は

$$\bar{Y}_{i\cdot} = \frac{1}{k} \sum_{j=1}^{k} Y_{ij} = \mu + B_i + \bar{e}_{i\cdot}$$

$$\bar{Y}_{\cdot j} = \frac{1}{n} \sum_{i=1}^{n} Y_{ij} = \mu + \bar{B}_{\cdot} + \tau_j + \bar{e}_{\cdot j}$$

$$\bar{Y}_{\cdot\cdot} = \frac{1}{nk} \sum_{i=1}^{n} \sum_{j=1}^{k} Y_{ij} = \mu + \bar{B}_{\cdot} + \bar{e}_{\cdot\cdot}$$

表 2　乱塊法モデル

	第 1 処理	第 2 処理	$\cdots$	第 k 処理
第 1 ブロック	Y_{11}	Y_{12}	$\cdots$	Y_{1k}
第 2 ブロック	Y_{21}	Y_{22}	$\cdots$	Y_{2k}
$\vdots$	$\vdots$	$\vdots$	$\cdots$	$\vdots$
第 n ブロック	Y_{n1}	Y_{n2}	$\cdots$	Y_{nk}

で表現できる．ただし，$\bar{e}_{i\cdot} = \dfrac{1}{k}\displaystyle\sum_{j=1}^{k} e_{ij}$, $\bar{e}_{\cdot j} = \dfrac{1}{n}\displaystyle\sum_{i=1}^{n} e_{ij}$, $\bar{e}_{\cdot\cdot} = \dfrac{1}{nk}\displaystyle\sum_{i=1}^{n}\sum_{j=1}^{k} e_{ij}$, $\bar{B}_{\cdot} = \dfrac{1}{n}\displaystyle\sum_{i=1}^{n} B_i$ とする．

3.2　検定統計量

帰無仮説:$H_0 : \tau_1 = \cdots = \tau_k = 0$ vs. 対立仮説:$H_A : \tau_1 \leqq \cdots \leqq \tau_k$ (少なくとも 1 つの $\leqq$ は $\lneqq$ である) を考える．e_{ij} を連続分布としてモデル (8) で帰無仮説 H_0vs. 対立仮説 H_A に対するノンパラメトリック検定として，Page(1963) は線形順位検定を提案した．Page 検定統計量で順位 R_{ij} を $Y_{ij} - \bar{Y}_{i\cdot}$ に替えた統計量は

$$
\begin{aligned}
T &\equiv \frac{1}{\sqrt{n}}\sum_{j=1}^{k}\left(j - \frac{k+1}{2}\right)\sum_{i=1}^{n}\left(Y_{ij} - \bar{Y}_{i\cdot}\right) \\
&= \frac{1}{\sqrt{n}}\sum_{j=1}^{k}\left(j - \frac{k+1}{2}\right)\sum_{i=1}^{n}\left(\tau_j + e_{ij} - \bar{e}_{i\cdot}\right) \qquad (9)
\end{aligned}
$$

となる．T の平均は

$$
E(T) = \sqrt{n}\sum_{j=1}^{k} j\tau_j \qquad (10)
$$

より，帰無仮説 H_0 の下で $E_0(T) = 0$ である．T の分散は

$$
\begin{aligned}
V_0(T) &= E_0\left[\left\{\frac{1}{\sqrt{n}}\sum_{j=1}^{k} j \sum_{i=1}^{n}(e_{ij} - \bar{e}_{i\cdot})\right\}^2\right] \\
&= \frac{k(k-1)(k+1)\sigma^2}{12} \qquad (11)
\end{aligned}
$$

である．(10)，(11) より，帰無仮説 H_0 の下で

$$
T \sim N\left(0, \frac{k(k-1)(k+1)\sigma^2}{12}\right) \qquad (12)
$$

である．統計量 T_N，誤差分散 $\tilde{\sigma}^2$，誤差平方和 S_e を

$$
\begin{aligned}
T_N &\equiv \sqrt{\frac{12}{k(k-1)(k+1)\tilde{\sigma}^2}}\,T \qquad (13) \\
\tilde{\sigma}^2 &\equiv \frac{S_e}{(k-1)(n-1)} \\
S_e &\equiv \sum_{i=1}^{n}\sum_{j=1}^{k}(Y_{ij} - \bar{Y}_{i\cdot} - \bar{Y}_{\cdot j} + \bar{Y}_{\cdot\cdot})^2 \\
&= \sum_{i=1}^{n}\sum_{j=1}^{k}(e_{ij} - \bar{e}_{i\cdot} - \bar{e}_{\cdot j} + \bar{e}_{\cdot\cdot})^2
\end{aligned}
$$

とおいたとき，次の命題 1 を得る．

命題 1　T と $\tilde{\sigma}^2$ が独立である

証明　文献 [3] の定理 2.29 より，T と S_e が独立であることを示せばよい．(9) の式で

$$
\sum_{i=1}^{n}(\tau_j + e_{ij} - \bar{e}_{i\cdot}) = n(\bar{e}_{\cdot j'} - \bar{e}_{\cdot\cdot})
$$

と計算できる．よって，文献 [3] の定理 2.22 の (3) より $\mathrm{Cov}(\bar{e}_{\cdot j'} - \bar{e}_{\cdot\cdot}, e_{ij} - \bar{e}_{i\cdot} - \bar{e}_{\cdot j} + \bar{e}_{\cdot\cdot}) = 0$ を求めればよい．

$j = j'$ のとき

$$
\mathrm{Cov}(\bar{e}_{\cdot j} - \bar{e}_{\cdot\cdot}, e_{ij} - \bar{e}_{i\cdot} - \bar{e}_{\cdot j} + \bar{e}_{\cdot\cdot}) = 0 \qquad (14)
$$

である．同様に，$j \neq j'$ のとき，

$$
\mathrm{Cov}(\bar{e}_{\cdot j'} - \bar{e}_{\cdot\cdot}, e_{ij} - \bar{e}_{i\cdot} - \bar{e}_{\cdot j} + \bar{e}_{\cdot\cdot}) = 0 \qquad (15)
$$

である．(14)，(15) より

$$
T \text{ と } S_e \text{は独立である} \qquad (16)
$$

ことが示された．文献 [3] の定理 2.29 より結論は導かれる．

3.3　検定方式

(12) より，H_0 の下で文献 [3] の系 3.6 を適用すると

$$
\sqrt{\frac{12}{k(k-1)(k+1)\sigma^2}}\,T \sim N(0,1) \qquad (17)
$$

を得る．さらに文献 [6] の定理 7.44 より

$$
\frac{S_e}{\sigma^2} \sim \chi^2_{(k-1)(n-1)} \qquad (18)
$$

を得る．(17)，(18)，命題 1 と文献 [3] の定理 3.21 より，H_0 の下で，

$$
T_N \sim t_{(k-1)(n-1)}
$$

が成り立つ．T_N を検定統計量として，自由度 $(k-1)(n-1)$ の t 分布の上側 $100\alpha\%$ 点を $t((k-1)(n-1);\alpha)$ とすると，水準 α の検定関数 は検定関数$\phi(\cdot)$ を用いて，

$$
\phi(\mathbf{Y}) = \begin{cases} 1 & (T_N > t((k-1)(n-1);\alpha) \text{ のとき}) \\ 0 & (T_N < t((k-1)(n-1);\alpha) \text{ のとき}) \end{cases}
$$

と表現される．すなわち，$T_N > t((k-1)(n-1);\alpha)$ のとき H_0 を棄却し，$T_N < t((k-1)(n-1);\alpha)$ のとき H_0 を棄却しない．

4　対照群との比較を行う多重比較検定法

4.1　Page 型多重比較検定法

表 2 のモデルにおいて処理効果に傾向性の制約

$$
\tau_1 \leqq \cdots \leqq \tau_k \qquad (19)
$$

は成り立っているものとする．第 1 群を対照群，第 2 群から第 k 群を処理群とする．j を $2 \leqq j \leqq k$ とする．1 つの比較のための検定は

$$\text{帰無仮説 } H_j : \tau_1 = \tau_j \text{ vs. 対立仮説 } H_j^A : \tau_1 < \tau_j$$

となる．$\ell = 2, \cdots, k$ に対して

$$T_N(\ell) \equiv \sqrt{\frac{12}{\ell(\ell-1)(\ell+1)\tilde{\sigma}^2(\ell)}} T(\ell) \tag{20}$$

とする．ただし，

$$T(\ell) \equiv \sum_{j=1}^{\ell} \left(j - \frac{\ell+1}{2} \right) \sum_{i=1}^{n} (Y_{ij} - Y_{i\cdot}(\ell))$$

$$\tilde{\sigma}^2(\ell) \equiv \frac{\sum_{i=1}^{n} \sum_{j=1}^{\ell} (Y_{ij} - \bar{Y}_{i\cdot}(\ell) - \bar{Y}_{\cdot j} + \bar{Y}_{\cdot\cdot}(\ell))^2}{(\ell-1)(n-1)}$$

$$\bar{Y}_{i\cdot}(\ell) \equiv \frac{1}{\ell} \sum_{j=1}^{\ell} Y_{ij}$$

$$\bar{Y}_{\cdot\cdot}(\ell) \equiv \frac{1}{n\ell} \sum_{i=1}^{n} \sum_{j=1}^{\ell} Y_{ij}$$

とする．ただし，$Y_{\cdot j}$ は第 3 章で表記した値と同じものとする．水準 α の多重比較検定として $T_N(\ell)$ に基づいた手法 [1] が提案できる．

[1] $T_N(\ell)$ **に基づく多重比較検定**

$j \leqq \ell \leqq k$ となる任意の ℓ に対して $T_N(\ell) > t((\ell-1)(n-1); \alpha)$ ならば，{ 帰無仮説 H_j vs. 対立仮説 $H_j^A \,|\, j = 2, \cdots, k\}$ に対する水準 α の多重比較検定として帰無仮説 H_j を棄却し，対立仮説 H_j^A を受け入れ $\tau_1 < \tau_j$ と判定する．

定理 1 [1] の検定方式は水準 α の多重比較検定である．

証明　帰無仮説のファミリーを，

$$\mathcal{H} \equiv \{H_j \,|\, j \in \mathcal{J}\}$$

とおく．ただし，$\mathcal{J} \equiv \{j \,|\, 2 \leqq j \leqq k\}$ とする．

$\boldsymbol{\tau} \equiv (\tau_1, \cdots, \tau_k)$ とおく．$\tau_1 < \tau_2$ のときは，有意水準は関係しないので，

$$\begin{aligned} \boldsymbol{\Theta}_0 &\equiv \{\boldsymbol{\tau} \,|\, 1 \text{ つ以上の帰無仮説 } H_j \text{が真 }\} \\ &= \{\boldsymbol{\tau} \,|\, \text{ある } j \in J \text{ が存在して}, \tau_1 = \tau_j\} \end{aligned} \tag{21}$$

とおき，$\boldsymbol{\tau} \in \Theta_0$ とする．このとき，正しい帰無仮説 H_j は 1 つ以上ある．また，確率は $\boldsymbol{\tau}$ に依存するので，確率測度を $P_{\boldsymbol{\tau}}(\cdot)$ で表す．$\tau_j = \tau_1$ を満たす最大の自然数 j を j_0 とする．事象 D_ℓ を

$$D_\ell \equiv \{T_N(\ell) \geqq t((\ell-1)(n-1); \alpha)\}$$

とおく．$2 \leqq j \leqq j_0$ を満たす整数 j に対して [1] の方法で正しい帰無仮説 H_j を棄却する事象は，$\bigcap_{\ell=j}^{k} D_\ell$ であるので，[1] の方法で 1 つ以上の正しい帰無仮説 H_j を棄却する確率は，

$$P_{\boldsymbol{\tau}} \left(\bigcup_{j=2}^{j_0} \left\{ \bigcap_{\ell=j}^{k} D_\ell \right\} \right) \leqq P_{\boldsymbol{\tau}}(D_{j_0}) = P_0(D_{j_0}) \leqq \alpha$$

である．ゆえに，定理の主張は証明された．

4.2　Williams 型多重比較検定法

$2 \leqq \ell \leqq k$ となる任意の ℓ に対して，統計量 T_ℓ^o と $\tilde{\mu}_\ell^o$ を，

$$T_\ell^o \equiv \frac{\sqrt{n}(\tilde{\mu}_\ell^o - \bar{Y}_{\cdot 1})}{\sqrt{2\tilde{\sigma}^2}}$$

$$\tilde{\mu}_\ell^o \equiv \max_{2 \leqq s \leqq \ell} \frac{\sum_{j=s}^{\ell} \bar{Y}_{\cdot j}}{\ell - s + 1}$$

で定義する．このとき，$\tilde{\mu}_\ell^o$ は，$\mu_2 \leqq \mu_3 \leqq \cdots \leqq \mu_\ell$ の下での，μ_ℓ の最尤推定量である．自由度 $m \equiv (\ell-1)(n-1)$ とし，Williams 型統計量 T_ℓ^o の上側 $100\alpha\%$ 点 $td(\ell, m; \alpha)$ を文献 [4] の付表 24 の数値とするとき，水準 α の多重比較検定として T_ℓ^o に基づいた手法 [2] が提案できる．

[2] T_ℓ^o **に基づく多重比較検定**

$j \leqq \ell \leqq k$ となる任意の ℓ に対して，$T_\ell^o > td(\ell, m; \alpha)$ ならば，{ 帰無仮説 H_j vs. 対立仮説 $H_j^A \,|\, j = 2, \cdots, k\}$ に対する水準 α の多重比較検定として帰無仮説 H_j を棄却し，対立仮説 H_j^A を受け入れ $\tau_1 < \tau_j$ と判定する．

定理 2 [2] の検定方式は水準 α の多重比較検定である．

定理 2 の証明は本稿に記載した．これ以降，4.1 節の検定法を「Page 型」，4.2 節の検定法を「Williams 型」と表記する．

5　処理群ごとの検出力の比較

繰り返し回数を 100,000 回のシミュレーションにより，各検定法の処理群ごとの検出力の比較を行う．$j = 2, \cdots, k$ のとき，$NT_j(re)$ を

$$NT_j(re) = \begin{cases} 1 & (\text{帰無仮説 } H_j \text{ が棄却される場合}) \\ 0 & (\text{帰無仮説 } H_j \text{ が棄却されない場合}) \end{cases}$$

として，対立仮説 H_j^A の検出力 $PT_j(re)$ を

$$PT_j(re) \equiv \frac{1}{rep} \sum_{re=1}^{rep} NT_j(re)$$

とおいた．ただし，rep はシミュレーションの繰り返し回数である．$n = 11, 21,\ k = 4, 5, 6,\ \alpha = 0.05, rep = 100,000$

とし未知母数 τ_j を

$$\tau_j(1) = \Delta\left(j - \frac{1}{k}\sum_{\ell=1}^{k}\ell\right)$$

$$\tau_j(2) = \Delta\left(j^2 - \frac{1}{k}\sum_{\ell=1}^{k}\ell^2\right)$$

$$\tau_j(3) = \Delta\left(\sqrt{j} - \frac{1}{k}\sum_{\ell=1}^{k}\sqrt{\ell}\right)$$

$$\tau_j(4) = \Delta\left(\log j - \frac{1}{k}\sum_{\ell=1}^{k}\log \ell\right)$$

と設定しシミュレーションを行った．このとき，Δ の値は Page 型における対立仮説 H^A_{k-1} の検出力が 0.7 に近くなるように定めた．表 3,4,5,6 は $n = 21, k = 5, \tau_j = \tau_j(1), \cdots, \tau_j(4)$ の場合の結果である． H^A_3, H^A_4, H^A_5 は

表 3 $n = 21, k = 5, \tau_j(1)$

対立仮説	Page 型	Williams 型
H^A_2	0.11370	0.12606
H^A_3	0.37273	0.35994
H^A_4	0.71364	0.66141
H^A_5	0.93695	0.88958

表 4 $n = 21, k = 5, \tau_j(2)$

対立仮説	Page 型	Williams 型
H^A_2	0.08206	0.08704
H^A_3	0.34550	0.32377
H^A_4	0.81229	0.75789
H^A_5	0.99586	0.98507

表 5 $n = 21, k = 5, \tau_j(3)$

対立仮説	Page 型	Williams 型
H^A_2	0.15151	0.16942
H^A_3	0.41956	0.41598
H^A_4	0.71099	0.67415
H^A_5	0.90370	0.86328

表 6 $n = 21, k = 5, \tau_j(4)$

対立仮説	Page 型	Williams 型
H^A_2	0.19321	0.22400
H^A_3	0.46036	0.47629
H^A_4	0.70054	0.68766
H^A_5	0.86599	0.84410

Page 型の検出力が Williams 型を上回ることが多かったが，H^A_2 は Page 型の検出力よりも Williams 型の検出力のほうが高かった．

6 総対検出力の比較

繰り返し回数 100,000 回のシミュレーションにより，各検定法の総対検出力の比較を行う．$j = 2, 3, 4,\ \tau_1 < \tau_2$ のとき $NTT(re)$ を

$$NTT(re) = \begin{cases} 1 & (\text{帰無仮説 } H_j \text{ がすべて棄却される場合}) \\ 0 & (\text{それ以外の場合}) \end{cases}$$

として，総対検出力 APP を

$$APP \equiv \frac{1}{rep}\sum_{re=1}^{rep} NTT(re)$$

とおいた．ただし，rep は繰り返し回数である．$n = 11, 21,\ k = 4,\ \alpha = 0.05,\ rep = 100,000$ として，5 章で設定した $\tau_j(2)$ を除く未知母数についてそれぞれシミュレーションを行った．このとき，Δ の値は Page 型の総対検出力が 0.5 に近くなるように定めた．表 7,8 は，$n = 11, 21\, k = 4,\ \alpha = 0.05$ の場合の結果である．

$n = 11, 21$ のどちらの場合でも Page 型の検出力が

表 7 $n = 11, k = 4, \alpha = 0.05$

母数	Page 型	Williams 型
$\tau_j(1)$	0.5061	0.49488
$\tau_j(3)$	0.50125	0.49074
$\tau_j(4)$	0.50314	0.49102

表 8 $n = 21, k = 4, \alpha = 0.05$

母数	Page 型	Williams 型
$\tau_j(1)$	0.49849	049274
$\tau_j(3)$	0.50476	0.49881
$\tau_j(4)$	0.50271	0.49706

Williams 型よりも高くなっていた．また，$\tau_j(2)$ の総対検出力は 0.5 に近づかなかったため省略した．

7 おわりに

本研究では線形型統計量を基にした多重比較検定法を提案することができた．また，Williams 型多重比較検定法を提案した．シミュレーションにより検出力の比較を行った結果，棄却される帰無仮説の個数が多い場合は Williams 型の検出力のほうが高く，棄却される帰無仮説が少ない場合は Page 型の検出力のほうが高くなることが分かった．総対検出力は Pageg 型が Williams 型よりも一様に高くなることが分かった．

参考文献

[1] Hettmansperger, T. P. *Statistical Inference Based on Ranks(Wiley Series in Probability and Statistics).* Willy, NewYork, 1984.

[2] 国土交通省：気象庁，過去の気象データ検索. http://www.data.jma.go.jp/gmd/cpd/monitor/index.html

[3] 白石高章:『統計科学の基礎』. 日本評論社, 東京. 2012.

[4] 白石高章，杉浦洋:『多重比較法の理論と数値計算』. 共立出版，東京. 2018.

[5] 早川由宏，白石高章：Fortran と C 言語による統計プログラミングの基礎 Mathematica の使い方，研究ノート. 2015.

[6] 宿久洋，村上亨，原恭彦:『確率と統計の基礎 I 増補改訂版』. ミネルヴァ書房，京都. 2009.