乱塊法における順序制約がある場合の Page 型検定法

2013SE155：大畑航平

指導教員：白石高章

1 はじめに

乱塊法モデルにおいて処理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．本論文では，正規分布を仮定した乱塊法モデルにおける Page 型検定法について考察する．

2 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

$$Y_{ij} \equiv \mu + B_i + \alpha_j + e_{ij} \tag{1}$$

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 $B_i \sim N(0, \{\sigma_B\}^2)$，$\alpha_j$ は処理効果を表すパラメータで $\sum_{j=1}^{k} \alpha_j = 0$，$e_{ij}$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．

表 1　二元配置乱塊法

	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}

各標本平均は

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

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

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

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

3 検定統計量

帰無仮説 $H_0 : \alpha_1 = \cdots = \alpha_k = 0$, 対立仮説 $H_A : \alpha_1 \leqq \cdots \leqq \alpha_k$ (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 (1) 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

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

となる．T の平均は

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

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

$$V_0(T) = \frac{k(k-1)(k+1)\sigma^2}{12}$$

である．よって，帰無仮説 H_0 の下で

$$T \sim N\left(0, \frac{k(k-1)(k+1)\sigma^2}{12}\right) \tag{3}$$

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

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

とおいたとき，次の命題 1 を得る．

命題 1　T_N と $\tilde{\sigma}^2$ が独立である

証明　文献 [3] の定理 2.29 より，S_e と T が独立であることを示せばよい．(2) の式で

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

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

$$\mathrm{Cov}(\bar{e}_{\cdot j} - \bar{e}_{\cdot\cdot}, e_{ij} - \bar{e}_{i\cdot} - \bar{e}_{\cdot j} + \bar{e}_{\cdot\cdot}) = 0 \tag{4}$$

である．同様に，$j \neq j'$ のとき，

$$\mathrm{Cov}(\bar{e}_{\cdot j'} - \bar{e}_{\cdot\cdot}, e_{ij} - \bar{e}_{i\cdot} - \bar{e}_{\cdot j} + \bar{e}_{\cdot\cdot}) = 0 \tag{5}$$

である．(4)，(5) より

$$T \text{ と } S_e \text{は独立である} \tag{6}$$

ことが示された．文献 [3] の定理 2.29 より結論は導かれる．

4 検定方式

(3) の正規分布において，H_0 の下で文献 [3] の系 3.6 を適用すると

$$\sqrt{\frac{12}{k(k-1)(k+1)\sigma^2}}T \sim N(0,1) \tag{7}$$

を得る．さらに文献 [5] の定理 7.44 より

$$\frac{S_e}{\sigma^2} \sim \chi^2_{(k-1)(n-1)} \tag{8}$$

を得る．(6)，(7)，(8) と文献 [3] の定理 3.21 より，

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

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．

5 C 言語プログラムの解説

Page 型検定法による検定結果を出力するプログラムを C 言語により作成した．ただし，上側確率を求めるために，文献 [4] を参考にした．本論文では Page 型検定法の main 関数プログラムのみを以下に記載する．Page 型検定法の詳細なプログラムについては，本稿に記載した．

```
int main(void){
    int n,k;
    float Y[100][100];
    printf("ブロック数を入力:");
    scanf("%d",&n);
    printf("処理数を入力:");
    scanf("%d",&k);
    load(Y,n,k);
    output(Y,n,k);
    return(0);
    }
```

1. ブロック数と処理数を入力する．
2. 関数 load により，データをテキストファイルから読み込んで，データのブロック，処理を表示する．
3. 関数 Page により，T の値を計算する．
4. 関数 SE により，S_e の値を計算する．
5. 関数 sigma により，$\tilde{\sigma}^2$ を計算する．
6. 関数 Tn により，T_N を計算する．
7. 関数 output により，検定結果を出力する．

5.1 アメリカの月平均気温データ

アメリカの月平均気温が上昇しているかどうかを調べるにあたり，1997 年から 2016 年の 20 年間の 1 月と 8 月のデータ (文献 [2]) を使用した．観測地点をニューヨーク，セントルイス，ミネアポリス，ポートランド，サンフランシスコ，アトランタ，マイアミ，ニューオーリンズの 8 地点とし月平均気温データを集めた．

5.2 解析結果

1 月は $T_N = 0.59$, p 値 $= 0.275386$ となり，$\alpha = 0.05, 0.01$ のどちらの場合も帰無仮説 H_0 を棄却できなかった．8 月は $T_N = 3.19$, p 値 $= 0.000876$ となり，$\alpha = 0.01$ の場合に帰無仮説 H_0 は棄却された．この検定により 20 年間でアメリカの 8 月の月平均気温だけが上昇していることが確認できた．また，中国や日本でも同様の解析を行った．その結果，中国は 15 年間で 8 月の月平均気温だけが上昇していることが確認できた．アメリカと中国では 1 月の月平均気温が上昇していることは確認できなかった．一方，日本は 20 年間で 1 月と 8 月の両方の月平均気温が上昇していてアメリカと中国とは異なる結果になった．

6 おわりに

本論文では乱塊法モデルにおいての検定方式を提案した．検定を行うための C 言語プログラムを作成し，実際のデータを用いることによって乱塊法における Page 型検定法をより理解を深めることができた．

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] 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.

[5] 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.