

傾向性のある場合の手法

白石高章

南山大学情報理工学部

1 モデル

k 標本のモデル

$$X_{ij} = \mu_i + \epsilon_{ij} = \nu + \alpha_i + \epsilon_{ij}, \text{ ただし } \sum_{i=1}^k n_i \alpha_i = 0$$

を考える, ϵ_{ij} は誤差確率変数とよばれ独立で同一の平均 0 分散 σ^2 の分布関数 $F(\frac{x}{\sigma})$ をもつ. 多標本問題では μ_i に制限を置かなかったが, ここでは位置パラメータに傾向性の制約

$$\mu_1 \leq \mu_2 \leq \cdots \leq \mu_k \iff \alpha_1 \leq \alpha_2 \leq \cdots \leq \alpha_k \quad (1.1)$$

がある場合での統計解析法を論じる. 興味ある帰無仮説と対立仮説は,

$$\begin{cases} \text{帰無仮説 } H_0 : \mu_1 = \cdots = \mu_k \\ \text{対立仮説 } H_1^* : \mu_1 \leq \mu_2 \leq \cdots \leq \mu_k \text{ (少なくとも1つの不等号は } < \text{)} \end{cases}$$
$$\iff \begin{cases} \text{帰無仮説 } H_0 : \alpha_1 = \alpha_2 = \cdots = \alpha_k = 0 \\ \text{対立仮説 } H_1^* : \alpha_1 \leq \alpha_2 \leq \cdots \leq \alpha_k \text{ (少なくとも1つの不等号は } < \text{)} \end{cases}$$

であり. 傾向性の制約が正しいかどうかを確かめるためにもう 1 つの帰無仮説と対立仮説は,

$$\begin{cases} \text{帰無仮説 } H_0 \\ \text{対立仮説 } H_1^\# : \text{ある } i \text{ について } \mu_i > \mu_{i+1} \end{cases}$$
$$\iff \begin{cases} H_0 \\ H_1^\# : \text{ある } i \text{ について } \alpha_i > \alpha_{i+1} \end{cases}$$

である. 帰無仮説 H_0 vs. 対立仮説 $H_1^\#$ で H_0 が棄却されれば, 傾向性 (1.1) が成立しないことが解る. さらに傾向性の制約 (1.1) の下でのパラメータ $(\alpha_1, \dots, \alpha_k), \nu, (\mu_1, \dots, \mu_k)$ の点推定に興味がある.

2 解析ソフト

20代, 40代, 60代の健康な成人男性のコレステロール値 ($mg/100ml$) を測定したものが表1で得られた. 年齢はコレステロール値に影響を及ぼすかを調べる.

表1. 年代とコレステロール値

番号	1	2	3	4	5	6	7	8	9	10
20才代	135	222	252	260	269	235	235	386	252	201
40才代	294	311	286	264	277	336	208	346	239	
60才代	370	244	353	420	333	357	317			

この例を基に以下の節で詳しく紹介される統計手法のソフトプログラムの使用方法を紹介する.

< 1 > 正規性の検定法

Nfitk.exe : 正規性の検定法.

[実行例:入力データをテキストファイル 『A:¥input¥ksamp21.txt』 から読み込む]

- (1) 『ワードパッド』を起動し, 標本の個数 k , k 個の標本サイズ (個数), k 組の標本観測値の順で表1のデータを例に次のように入力しそれを 『A:¥input¥ksamp1.txt』として保存する.

```
3
10 9 7
135 222 252 260 269 235 235 386 252 201
294 311 286 264 277 336 208 346 239
370 244 353 420 333 357 317
```

- (2) Nfitk.exe を使って, 実行結果を画面に表示する.

E:¥Esoft > Nfitk < A:¥input¥ksamp1.txt
によって実行結果が画面に表示される.

エンターキー

< 2 > データの従っている分布に近い分布を探す

Seekfk.exe : データの従っている分布に近い分布を探す.

[実行例:入力データをテキストファイル 『A:¥input¥ksamp1.txt』 から読み込む]

E: ¥ Esoft > Seekfk < A: ¥ input ¥ ksamp1.txt

< 3 > 経験分布関数と正規分布関数の重ねがきグラフ

Empk.exe 及び Excel : 関数の値を求めエクセルで重ねがきグラフ.

○ Empk.exe を使って入力データを『A: ¥ input ¥ ksamp2.txt』から読み込み、ファイル『A: ¥ kekka ¥ empk.txt』に実行結果を書き込む。続いて、A: ¥ kekka ¥ empk.txt をエクセルで開き、経験分布関数と正規分布関数のグラフを描く手順を示す。

- (1) 『ワードパッド』を起動し、観測値の小数点以下の桁数 (整数のときは 0)、標本の個数 k , k 個の標本サイズ (個数), k 組の標本観測値の順で表 1 のデータを入力しそれを『A: ¥ input ¥ ksamp2.txt』とする。すなわち、このファイルは『A: ¥ input ¥ ksamp1.txt』の 1 行目に 0 を追加したものである。
- (2) Empk.exe を使って、エクセルに読み込むテキストファイルを作る。

E: ¥ Esoft > Empk < A: ¥ input ¥ ksamp2.txt > A: ¥ kekka ¥ empk.txt

- (3) ¥ empn.txt を ¥ empk.txt に替えた著書『統計科学』の 5.8 節 < 3 > の (3) から (9) と同様の操作を行う。

< 4 > 正規母集団での最良手法

Seir.exe : 正規母集団での最良手法による解析.

[実行例: キーボードで入力し画面に実行結果を表示する]

E: ¥ Esoft > Seik

標本の個数を入力して下さい

3 入力後エンターキー

3 つの標本サイズを入力して下さい

10 9 7 入力後エンターキー

第 1 標本の観測値を 10
個入力してください。

135 222 252 260 269 235 235 386 252 201 入力後エンターキー

第 2 標本の観測値を 9
個入力してください。

294 311 286 264 277 336 208 346 239 入力後エンターキー

第 3 標本の観測値を 7

個入力してください.

370 244 353 420 333 357 317

入力後エンターキー

有意水準を入力して下さい.

0.01 入力後エンターキー

『実行結果が画面に表示される』

[実行例 2:入力データをテキストファイル 『A:¥input ¥ksamp3.txt』 から読み込む]

著書 『統計科学』の 1.7 節 < 3 > 実行例 2 の手順 (1) から (5) のようにして, A:¥input ¥ksamp3.txt に入力データを入力し保存する.

E:¥Esoft > Seir < A:¥input ¥ksamp3.txt

によって実行結果が画面に表示される .

< 5 > ノンパラメトリック法

Nonpr.exe : 順位手法による解析.

[実行例:入力データをテキストファイル 『A:¥input ¥ksamp3.txt』 から読み込む]

E:¥Esoft > Nonpr < A:¥input ¥ksamp3.txt

によって実行結果が画面に表示される .

< 6 > M 統計量に基づくセミパラメトリック法.

Robr.exe : セミパラメトリック法による解析.

[実行例:入力データをテキストファイル 『A:¥input ¥ksamp3.txt』 から読み込む]

E:¥Esoft > Robr < A:¥input ¥ksamp3.txt

によって実行結果が画面に表示される .

◇ プログラム集

番号と内容	プログラム
< 1 > 正規性の検定	Nfitk.exe
< 2 > 分布の探索	Seekfk.exe
< 3 > 経験分布と正規分布関数	Empk.exe
< 4 > 正規母集団での最良手法	Seir.exe
< 5 > ノンパラメトリック法	Nonpr.exe
< 6 > セミパラメトリック法	Robr.exe

解析チャート

< 1 > 正規性の検定, < 2 > 分布の探索, < 3 > 経験分布関数と正規分布関数の重ねかきグラフ



< 1 > で正規性が棄却されず, < 2 > で正規分布が選択され,
< 3 > により正規性が妥当と認められれば < 4 > を選択
< 2 > で両側指数分布またはワイブル分布, 対数正規分布,
指数分布が選択されれば < 5 > を選択
これら以外であれば < 6 > を選択

3 傾向性制約での極値

手法を論じる前に, $\mathbf{X} \equiv (X_{11}, \dots, X_{1n_1}, \dots, X_{k1}, \dots, X_{kn_k})$ の k 個の関数値 $g_1(\mathbf{X}), \dots, g_k(\mathbf{X})$ に対して

$$u_1 \leq \dots \leq u_k \text{ の下で } \sum_{i=1}^k n_i \{u_i - g_i(\mathbf{X})\}^2 \text{ を最小}$$

にする $u_1 = g_1^*, \dots, u_k = g_k^*$, すなわち

$$\sum_{i=1}^k n_i \{g_i^* - g_i(\mathbf{X})\}^2 = \min_{u_1 \leq \dots \leq u_k} \sum_{i=1}^k n_i \{u_i - g_i(\mathbf{X})\}^2$$

なる g_1^*, \dots, g_k^* を求める必要がある. そのアルゴリズムを紹介する. 関数値 $g_1(\mathbf{X}), \dots, g_k(\mathbf{X})$ と標本サイズ n_1, \dots, n_k に対して p 番目の値 $g_p(\mathbf{X})$ から q 番目の値 $g_q(\mathbf{X})$ までの n_p から n_q を重みとする重みつき平均を

$$g_{[p,q]} \equiv \frac{\sum_{i=p}^q n_i g_i(\mathbf{X})}{\sum_{i=p}^q n_i} \quad (p \leq q)$$

とおく. このとき,

$$h_1^* \equiv \min_{1 \leq m \leq k} g_{[1,m]}, \quad m_1 \equiv \max\{m : g_{[1,m]} = h_1^*, 1 \leq m \leq k\}$$

とおけば, $g_1^* = \dots = g_{m_1}^* \equiv h_1^*$ となり, m_1 個の $g_1^*, \dots, g_{m_1}^*$ が定まる. もし, $m_1 < k$ ならば,

$$h_2^* \equiv \min_{m_1+1 \leq m \leq k} g_{[m_1+1,m]}, \quad m_2 \equiv \max\{m : g_{[m_1+1,m]} = h_2^*, m_1+1 \leq m \leq k\}$$

とおけば, $g_{m_1+1}^* = \dots = g_{m_2}^* \equiv h_2^*$ となり, $m_2 - m_1$ 個の $g_{m_1+1}^*, \dots, g_{m_2}^*$ が定まる. $m_2 < k$ ならば, 以下 $3 \leq r$ に対して順次下の枠内を $m_r = k$ となるまで続ける.

$$\begin{aligned} h_r^* &\equiv \min_{m_{r-1}+1 \leq m \leq k} g_{[m_{r-1}+1,m]} \\ m_r &\equiv \max\{m : g_{[m_{r-1}+1,m]} = h_r^*, m_{r-1}+1 \leq m \leq k\} \\ g_{m_{r-1}+1}^* &= \dots = g_{m_r}^* \equiv h_r^* \end{aligned}$$

$m_\ell = k$ ならば

$$g_1^* = \dots = g_{m_1}^* < g_{m_1+1}^* = \dots = g_{m_2}^* < \dots < g_{m_{\ell-1}+1}^* = \dots = g_{m_\ell}^*$$

となり g_i^* の ℓ 個が異なる. この場合の興味あるグラフは,

$$W_i \equiv \sum_{j=1}^i n_j, \quad G_i \equiv \sum_{j=1}^i n_j g_j \quad (i = 1, \dots, k),$$

$$Q_0 \equiv (0, 0), \quad Q_i \equiv (W_i, G_i) \quad (i = 1, \dots, k),$$

$$W_i^* \equiv \sum_{j=1}^{m_i} n_j, \quad G_i^* \equiv \sum_{j=1}^{m_i} n_j g_j^* \quad (i = 1, \dots, \ell),$$

$Q_0^* \equiv (0, 0)$, $Q_i^* \equiv (W_i^*, G_i^*)$ ($i = 1, \dots, \ell$) とおくと, Q_i^* ($i = 0, 1, \dots, \ell$) を結んでできた折れ線は, Q_i ($i = 0, 1, \dots, k$) を結んでできた折れ線以下にある下に凸な曲線で最大のものである. 例として, (n_i, g_i) ($i = 1, \dots, 6$) が, $(3, \frac{8}{3})$, $(2, 1.0)$, $(2, 3.0)$, $(3, 6.0)$, $(2, 1.0)$, $(3, 5.0)$ のときの $Q_0 = (0, 0)$, $Q_1 = (3, 8)$, $Q_2 = (5, 10)$, $Q_3 = (7, 16)$, $Q_4 = (10, 34)$, $Q_5 = (12, 36)$, $Q_6 = (15, 51)$ を結んでできた折れ線と $Q_0^* = (0, 0)$, $Q_1^* = (5, 10)$, $Q_2^* = (7, 16)$, $Q_3^* = (12, 36)$, $Q_4^* = (15, 51)$ を結んでできた折れ線のグラフを図 1 に重ね描きしている. $g_1^* = g_2^* = 2.0 < g_3^* = 3.0 < g_4^* = g_5^* = 4.0 < g_6^* = 5.0$ となる.

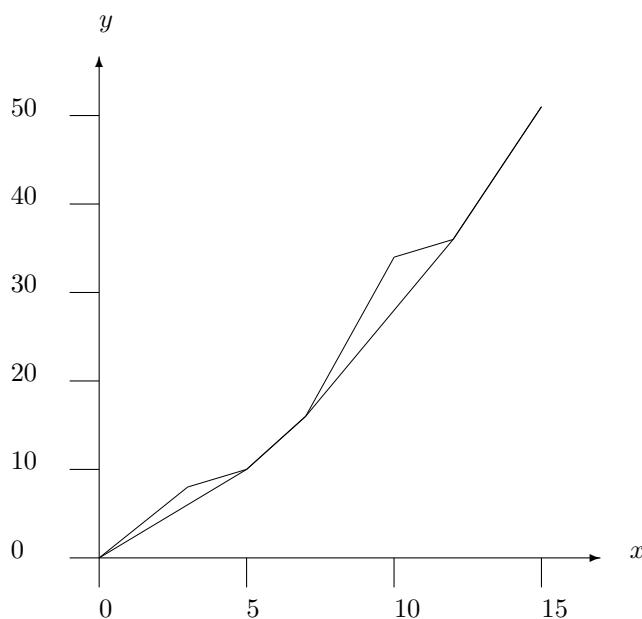


図 1 Q_i と Q_i^* の折れ線グラフ

4 正規母集団での最良手法

$F(\frac{x-\mu_i}{\sigma})$ を正規分布 $N(\mu_i, \sigma^2)$ の分布関数とする.

[1] 検定; $\bar{X}_{i.} \equiv \frac{1}{n_i} \sum_{j=1}^{n_i} X_{ij}$, $\bar{X}_{..} \equiv \frac{1}{n} \sum_{i=1}^k \sum_{j=1}^{n_i} X_{ij}$ とおき, $\tilde{\mu}_1^*, \dots, \tilde{\mu}_k^*$ を

$$\sum_{i=1}^k n_i \{\tilde{\mu}_i^* - \bar{X}_{i.}\}^2 = \min_{u_1 \leq \dots \leq u_k} \sum_{i=1}^k n_i \{u_i - \bar{X}_{i.}\}^2$$

を満たすものとする. このとき, 帰無仮説 H_0 vs. 対立仮説 H_1^* の検定統計量と帰無仮説 H_0 vs. 対立仮説 $H_1^\#$ の検定統計量はそれぞれ

$$T_S^* \equiv \frac{\sum_{i=1}^k n_i (\tilde{\mu}_i^* - \bar{X}_{..})^2}{\sum_{i=1}^k \sum_{j=1}^{n_i} (X_{ij} - \bar{X}_{i.})^2}, \quad T_S^\# \equiv \frac{\sum_{i=1}^k n_i (\bar{X}_{i.} - \tilde{\mu}_i^*)^2}{\sum_{i=1}^k \sum_{j=1}^{n_i} (X_{ij} - \bar{X}_{i.})^2},$$

と表され、 T_S^* が大きいとき H_0 を棄却し H_1^* を受け入れる。また、 $T_S^\#$ が大きいとき H_0 を棄却し $H_1^\#$ を受け入れる。

まず、 $n - k$ が 50 未満のときの手順を述べる。

- 手順 1. $0 < \alpha < 1$ となるように有意水準 α を決め、観測値

$$\mathbf{X} \equiv (X_{11}, \dots, X_{kn_k})$$

から、 T_S^* と $T_S^\#$ を計算する。

- 手順 2. シミュレーションの繰り返し数 M として 1000 を採り、

$i = 1, 2, \dots, M - 1$ について以下の手順 2.1 を繰り返す。

- 手順 2.1. 乗算合同法の標準正規乱数を n 個生成させ、それを $X_{11}(i), \dots, X_{kn_k}(i)$

とし、

$$T_S^*(i) \equiv \frac{\sum_{i=1}^k n_i \{\tilde{\mu}_i^*(i) - \bar{X}_{..}(i)\}^2}{\sum_{i=1}^k \sum_{j=1}^{n_i} \{\bar{X}_{ij}(i) - \bar{X}_{..}(i)\}^2},$$

$$T_S^\#(i) \equiv \frac{\sum_{i=1}^k n_i \{\bar{X}_{i.}(i) - \tilde{\mu}_i^*(i)\}^2}{\sum_{i=1}^k \sum_{j=1}^{n_i} \{X_{ij}(i) - \bar{X}_{i.}(i)\}^2},$$

を計算する、ただし、 $\bar{X}_{i.}(i)$, $\bar{X}_{..}(i)$ は $\bar{X}_{i.}$, $\bar{X}_{..}$ と類似の定義。

$\tilde{\mu}_1^*(i), \dots, \tilde{\mu}_k^*(i)$ を

$$\sum_{i=1}^k n_i \{\tilde{\mu}_i^*(i) - \bar{X}_{i.}(i)\}^2 = \min_{u_1 \leq \dots \leq u_k} \sum_{i=1}^k n_i \{u_i - \bar{X}_{i.}(i)\}^2$$

を満たすものとする。

- 手順 3. $N_1 \equiv \sum_{i=1}^{M-1} I(T_S^* < T_S^*(i))$, $N_2 \equiv \sum_{i=1}^{M-1} I(T_S^\# < T_S^\#(i))$ とおく。

$\frac{N_1}{M} \leq \alpha \implies H_0$ を棄却し、 H_1^* を受け入れる。

$\frac{N_2}{M} \leq \alpha \implies H_0$ を棄却し、 $H_1^\#$ を受け入れる。

次に $n - k \geq 50$ のときの手順を述べる。

- 手順 1. $0 < \alpha < 1$ となるように有意水準 α を決め、観測値

$$\mathbf{X} \equiv (X_{11}, \dots, X_{kn_k})$$

から、 $Z_S^* \equiv (n - 1)T_S^*$ と $Z_S^\# \equiv (n - k)T_S^\#$ を計算する。

- 手順 2. $i = 1, \dots, k$ について $\lambda_i \equiv \frac{n_i}{n}$ とおく。シミュレーションの繰り返し数 M として 1000 を採り、

$i = 1, 2, \dots, M - 1$ について以下の手順 2.1 を繰り返す。

- 手順 2.1. それぞれ $N(0, \frac{1}{\lambda_1}), \dots, N(0, \frac{1}{\lambda_k})$ に従う正規乱数 k 個

$y_1(i), \dots, y_k(i)$ 生成する.

$$Z_S^*(i) \equiv \sum_{i=1}^k \lambda_i \{y_i^*(i) - \sum_{j=1}^k \lambda_j y_j(i)\}^2 \quad (4.1)$$

$$Z_S^\#(i) \equiv \sum_{i=1}^k \lambda_i \{y_i(i) - y_i^*(i)\}^2 \quad (4.2)$$

とおく, ただし, $y_1^*(i), \dots, y_k^*(i)$ を

$$\sum_{i=1}^k \lambda_i \{y_i^*(i) - y_i(i)\}^2 = \min_{u_1 \leq \dots \leq u_k} \sum_{i=1}^k \lambda_i \{u_i - y_i(i)\}^2$$

を満たすものとする.

- 手順 3. $N_1' \equiv \sum_{i=1}^{M-1} I(Z_S^* < Z_S^*(i))$, $N_2' \equiv \sum_{i=1}^{M-1} I(Z_S^\# < Z_S^\#(i))$ とおく.

$$\frac{N_1'}{M} \leq \alpha \implies H_0 \text{ を棄却し, } H_1^* \text{ を受け入れる.}$$

$$\frac{N_2'}{M} \leq \alpha \implies H_0 \text{ を棄却し, } H_1^\# \text{ を受け入れる.}$$

[2] 点推定量; 傾向性の制約 $\mu_1 \leq \mu_2 \leq \dots \leq \mu_k$ の下でのそれぞれのパラメータの推定量は次のとおりである.

$$\hat{\alpha}_i^* = \tilde{\mu}_i^* - \bar{X}_.. \quad (i = 1, \dots, k), \quad \tilde{\nu}^* = \bar{X}_.., \quad \tilde{\mu}_i^* \quad (i = 1, \dots, k).$$

5 ノンパラメトリック法

$F(x)$ を連続型未知の分布関数とする.

[3] 順位検定; 『多標本問題』の5節 [5] で定義したように n 個すべての観測値 X_{11}, \dots, X_{kn_k} を小さい方から並べたときの X_{ij} の順位を R_{ij} とする. $\bar{R}_i. \equiv \frac{1}{n_i} \sum_{j=1}^{n_i} R_{ij}$ とし, $\bar{R}_1^*, \dots, \bar{R}_k^*$ を

$$\sum_{i=1}^k n_i (\bar{R}_i^* - \bar{R}_i.)^2 = \min_{u_1 \leq \dots \leq u_k} \sum_{i=1}^k n_i (u_i - \bar{R}_i.)^2$$

を満たすものとする. このとき, 帰無仮説 H_0 vs. 対立仮説 H_1^* の検定統計量と帰無仮説 H_0 vs. 対立仮説 $H_1^\#$ の検定統計量はそれぞれ

$$T_R^* \equiv \sum_{i=1}^k n_i \left(\bar{R}_i^* - \frac{n+1}{2} \right)^2, \quad T_R^\# \equiv \sum_{i=1}^k n_i (\bar{R}_i. - \bar{R}_i^*)^2,$$

と表され, T_R^* が大きいとき H_0 を棄却し H_1^* を受け入れる. また, $T_R^\#$ が大きいとき H_0 を棄却し, $H_1^\#$ を受け入れる. まず, ある i について n_i が 30 未満のときの手順を述べる.

- 手順 1. $0 < \alpha < 1$ となるように有意水準 α を決め,

$$\mathbf{R}(0) \equiv (R_1(0), \dots, R_n(0)) \equiv (R_{11}, \dots, R_{1n_1}, R_{21}, \dots, R_{kn_k})$$

とおき, 統計量

$$T_R^* \equiv \sum_{i=1}^k n_i \left\{ \bar{R}_i^*(0) - \frac{n+1}{2} \right\}^2,$$

$$T_R^\# \equiv \sum_{i=1}^k n_i \{ \bar{R}_i(0) - \bar{R}_i^*(0) \}^2,$$

を計算する, ただし, $L_0 \equiv 0$, $L_i \equiv \sum_{j=1}^i n_j$,

$\bar{R}_i(0) \equiv \sum_{j=L_{i-1}+1}^{L_i} R_j(0)/n_i$, $\bar{R}_i^*(0), \dots, \bar{R}_i^*(0)$ を

$$\sum_{i=1}^k n_i \{ \bar{R}_i^*(0) - \bar{R}_i(0) \}^2 = \min_{u_1 \leq \dots \leq u_k} \sum_{i=1}^k n_i \{ u_i - \bar{R}_i(0) \}^2$$

を満たすものとする.

- 手順 2. シミュレーションの繰り返し数 M として 1000 を採り,
 - $i = 1, 2, \dots, M-1$ について以下の手順 2.1 から手順 2.4 までを繰り返す.
 - 手順 2.1. 乗算合同法の疑似一様乱数を $n - n_1$ 個生成させ, それを $u_1(i), \dots, u_{n-n_1}(i)$ とする.
 - 手順 2.2. $j = 1, \dots, n - n_1$ について以下の手順 2.2.1 と手順 2.2.2 を繰り返す.
 - 手順 2.2.1. $\ell \equiv [(n - j + 1) \cdot u_j(i)] + 1$ とおき, $\ell = n - j + 2$ ならば新たに $\ell \equiv 1$ とおく, ただし, $[x]$ は x を越えない最大の整数とする.
 - 手順 2.2.2. n 次元ベクトル $\mathbf{R}(\ell - 1) \equiv (R_1(\ell - 1), \dots, R_n(\ell - 1))$ の $n - j + 1$ 成分 $R_{n-j+1}(\ell - 1)$ と ℓ 成分 $R_\ell(\ell - 1)$ とを入れ替えたものを新たに, $\mathbf{R}(\ell - 1)$ とおく.
 - 手順 2.3. 手順 2.2 で成分入れ替えによって最終的に決定された $\mathbf{R}(\ell - 1)$ を $\mathbf{R}(i)$ とおく.
 - 手順 2.4. $\mathbf{R}(i) \equiv (R_1(i), \dots, R_n(i))$ より

$$T_R^*(i) \equiv \sum_{i=1}^k n_i \left\{ \bar{R}_i^*(i) - \frac{n+1}{2} \right\}^2,$$

$$T_R^\#(i) \equiv \sum_{i=1}^k n_i \{ \bar{R}_i(i) - \bar{R}_i^*(0) \}^2,$$

を計算する, ただし, $\bar{R}_i(i), \bar{R}_i^*(i)$ は $\bar{R}_i(0), \bar{R}_i^*(0)$ と類似の定義.

- 手順 3. $N_{1u} \equiv \sum_{i=1}^{M-1} I(|T_R^*| \leq |T_R^*(i)|)$, $N_{1\ell} \equiv \sum_{i=1}^{M-1} I(|T_R^*| < |T_R^*(i)|)$,
 $N_{2u} \equiv \sum_{i=1}^{M-1} I(|T_R^\#| \leq |T_R^\#(i)|)$, $N_{2\ell} \equiv \sum_{i=1}^{M-1} I(|T_R^\#| < |T_R^\#(i)|)$ とおく.
 一様乱数 1 つを生成させ, それを v とする.

$$\frac{N_{1u}}{M} \leq \alpha \text{ または } \left(\frac{N_{1u}}{M} > \alpha \text{ かつ } \frac{N_{1\ell}}{M} < \alpha \text{ かつ } v < \frac{M \cdot \alpha - N_{1\ell}}{N_{1u} - N_{1\ell}} \right)$$

$$\begin{aligned} &\implies H_0 \text{ を棄却する.} \\ \frac{N_{2u}}{M} \leq \alpha &\text{ または } \left(\frac{N_{2u}}{M} > \alpha \text{ かつ } \frac{N_{2\ell}}{M} < \alpha \text{ かつ } v < \frac{M \cdot \alpha - N_{2\ell}}{N_{2u} - N_{2\ell}}\right) \\ &\implies H_0 \text{ を棄却する.} \end{aligned}$$

次に $n_1, \dots, n_k \geq 30$ のときの手順を述べる.

- 手順 1. $0 < \alpha < 1$ となるように有意水準 α を決め, 観測値

$$\mathbf{X} \equiv (X_{11}, \dots, X_{kn_k}) \text{ から,}$$

$$Z_R^* \equiv \frac{(n-1)T_R^*}{\sum_{j=1}^n (R_j - \frac{n+1}{2})^2}, \quad Z_R^\# \equiv \frac{(n-1)T_R^\#}{\sum_{j=1}^n (R_j - \frac{n+1}{2})^2}$$

を計算する.

- 手順 2. $i = 1, \dots, k$ について $\lambda_i \equiv \frac{n_i}{n}$ とおく. シミュレーションの繰り返し数 M として 1000 を採り,

$i = 1, 2, \dots, M-1$ について以下の手順 2.1 を繰り返す.

- 手順 2.1. それぞれ $N(0, \frac{1}{\lambda_1}), \dots, N(0, \frac{1}{\lambda_k})$ に従う正規乱数 k 個 $y_1(i), \dots, y_k(i)$

生成すし, (4.1), (4.2) で定義した $Z_S^*(i)$ と $Z_S^\#(i)$ を計算する.

- 手順 3. $N'_1 \equiv \sum_{i=1}^{M-1} I(Z_R^* < Z_S^*(i))$, $N'_2 \equiv \sum_{i=1}^{M-1} I(Z_R^\# < Z_S^\#(i))$ とおく.

$$\frac{N'_1}{M} \leq \alpha \implies H_0 \text{ を棄却し, } H_1^* \text{ を受け入れる.}$$

$$\frac{N'_2}{M} \leq \alpha \implies H_0 \text{ を棄却し, } H_1^\# \text{ を受け入れる.}$$

- [4] 順位推定量; $\hat{\alpha}_i$ ($i = 1, \dots, k$) と $\hat{\nu}$ をそれぞれ『多標本問題』の 5 節 [6] で紹介した α_i ($i = 1, \dots, k$) と ν の順位推定とする. $\hat{\alpha}_1^*, \dots, \hat{\alpha}_k^*$ を

$$\sum_{i=1}^k n_i \{\hat{\alpha}_i^* - \hat{\alpha}_i\}^2 = \min_{u_1 \leq \dots \leq u_k} \sum_{i=1}^k n_i \{u_i - \hat{\alpha}_i\}^2$$

を満たすものとする. このとき, 傾向性の制約 $\mu_1 \leq \mu_2 \leq \dots \leq \mu_k$ の下でのそれぞれのパラメータの順位推定量は次のとおりである.

$\hat{\alpha}_i^*$ ($i = 1, \dots, k$), $\hat{\nu}$, $\hat{\mu}_i^* = \hat{\nu} + \hat{\alpha}_i^*$ ($i = 1, \dots, k$). $\hat{\alpha}_i^*$ は密度関数 $f(x) \equiv F'(x)$ に対称性を必要としないが, $\hat{\nu}$, $\hat{\mu}_i^*$ には $f(x)$ の対称性を必要とする.

6 頑健な手法

$F(x)$ を連続型未知の分布関数とする.

- [5] M 検定; $\bar{\psi}_{i.}(\mathbf{X}) \equiv \frac{1}{n_i} \sum_{j=1}^{n_i} \psi(\frac{X_{ij} - \hat{\mu}}{\hat{s}_n})$ と $\bar{\psi}_{..}(\mathbf{X})$ を『多標本問題』の 6 節 [7] で定義したものとする. $\bar{\psi}_{1.}^*(\mathbf{X}), \dots, \bar{\psi}_{k.}^*(\mathbf{X})$ を

$$\sum_{i=1}^k n_i \{\bar{\psi}_{i.}^*(\mathbf{X}) - \bar{\psi}_{i.}(\mathbf{X})\}^2 = \min_{u_1 \leq \dots \leq u_k} \sum_{i=1}^k n_i \{u_i - \bar{\psi}_{i.}(\mathbf{X})\}^2$$

を満たすものとする. このとき, 帰無仮説 H_0 vs. 対立仮説 H_1^* の検定統計量と帰無仮説 H_0 vs. 対立仮説 $H_1^\#$ の検定統計量はそれぞれ

$$T_M^* \equiv \sum_{i=1}^k n_i \{ \bar{\psi}_i^*(\mathbf{X}) - \bar{\psi}_{..}(\mathbf{X}) \}^2,$$

$$T_M^\# \equiv \sum_{i=1}^k n_i \{ \bar{\psi}_i(\mathbf{X}) - \bar{\psi}_i^*(\mathbf{X}) \}^2,$$

と表され, T_M^* が大きいとき H_0 を棄却し H_1^* を受け入れる. また, $T_M^\#$ が大きいとき H_0 を棄却し, $H_1^\#$ を受け入れる. まず, ある i について n_i が 30 未満のときの手順を述べる.

- 手順 1. $0 < \alpha < 1$ となるように有意水準 α を決め,

$$\begin{aligned} \mathbf{w}(0) &\equiv (w_1(0), \dots, w_n(0)) \\ &\equiv \left(\psi\left(\frac{X_{11} - \tilde{\nu}}{\hat{s}_n}\right), \dots, \psi\left(\frac{X_{1n_1} - \tilde{\nu}}{\hat{s}_n}\right), \psi\left(\frac{X_{21} - \tilde{\nu}}{\hat{s}_n}\right), \dots, \psi\left(\frac{X_{kn_k} - \tilde{\nu}}{\hat{s}_n}\right) \right) \end{aligned}$$

とおき, 統計量

$$T_M^* \equiv \sum_{i=1}^k n_i \{ \bar{w}_i^*(0) - \bar{w}_{..}(0) \}^2$$

$$T_M^\# \equiv \sum_{i=1}^k n_i \{ \bar{w}_i(0) - \bar{w}_i^*(0) \}^2$$

を計算する, ただし, $L_0 \equiv 0$, $L_i \equiv \sum_{j=1}^i n_j$,

$\bar{w}_i(0) \equiv \sum_{j=L_{i-1}+1}^{L_i} w_j(0)/n_i$, $\bar{w}_{..}(0) \equiv \sum_{j=1}^n w_j(0)/n$, $\bar{w}_1^*(0), \dots, \bar{w}_k^*(0)$ を

$$\sum_{i=1}^k n_i \{ \bar{w}_i^*(0) - \bar{w}_i(0) \}^2 = \min_{u_1 \leq \dots \leq u_k} \sum_{i=1}^k n_i \{ u_i - \bar{w}_i(0) \}^2$$

を満たすものとする.

- 手順 2. シミュレーションの繰り返し数 M として 1000 を採り,
 - $i = 1, 2, \dots, M-1$ について以下の手順 2.1 から手順 2.4 までを繰り返す.
 - 手順 2.1. 乗算合同法の疑似一様乱数を $n - n_1$ 個生成させ, それを $u_1(i), \dots, u_{n-n_1}(i)$ とする.
 - 手順 2.2. $j = 1, \dots, n - n_1$ について以下の手順 2.2.1 と手順 2.2.2 を繰り返す.
 - 手順 2.2.1. $\ell \equiv [(n - j + 1) \cdot u_j(i)] + 1$ とおき, $\ell = n - j + 2$ ならば新たに $\ell \equiv 1$ とおく, ただし, $[x]$ は x を越えない最大の整数とする.
 - 手順 2.2.2. n 次元ベクトル $\mathbf{w}(i-1) \equiv (w_1(i-1), \dots, w_n(i-1))$ の $n - j + 1$ 成分 $w_{n-j+1}(i-1)$ と ℓ 成分 $w_\ell(i-1)$ とを入れ替えたものを新たに, $\mathbf{w}(i-1)$ とおく.

- ・手順 2.3. 手順 2.2 で成分入れ替えによって最終的に決定された $w(i-1)$ を $w(i)$ とおく.
- ・手順 2.4. $w(i) \equiv (w_1(i), \dots, w_n(i))$ より

$$T_M^* \equiv \sum_{i=1}^k n_i \{\bar{w}_i^*(i) - \bar{w}_{..}(i)\}^2,$$

$$T_M^\# \equiv \sum_{i=1}^k n_i \{\bar{w}_i(i) - \bar{w}_i^*(i)\}^2,$$

を計算する, ただし, $\bar{w}_i(i)$, $\bar{w}_{..}(i)$, $\bar{w}_i^*(i)$ は $\bar{w}_i(0)$, $\bar{w}_{..}(0)$, $\bar{w}_i^*(0)$ と類似の定義.

- 手順 3. $N_{1u} \equiv \sum_{i=1}^{M-1} I(|T_M^*| \leq |T_M^*(i)|)$, $N_{1\ell} \equiv \sum_{i=1}^{M-1} I(|T_M^*| < |T_M^*(i)|)$,
 $N_{2u} \equiv \sum_{i=1}^{M-1} I(|T_M^\#| \leq |T_M^\#(i)|)$, $N_{2\ell} \equiv \sum_{i=1}^{M-1} I(|T_M^\#| < |T_M^\#(i)|)$ とおく.

一様乱数 1 つを生成させ, それを v とする.

$$\frac{N_{1u}}{M} \leq \alpha \text{ または } \left(\frac{N_{1u}}{M} > \alpha \text{ かつ } \frac{N_{1\ell}}{M} < \alpha \text{ かつ } v < \frac{M \cdot \alpha - N_{1\ell}}{N_{1u} - N_{1\ell}} \right)$$

$\Rightarrow H_0$ を棄却する.

$$\frac{N_{2u}}{M} \leq \alpha \text{ または } \left(\frac{N_{2u}}{M} > \alpha \text{ かつ } \frac{N_{2\ell}}{M} < \alpha \text{ かつ } v < \frac{M \cdot \alpha - N_{2\ell}}{N_{2u} - N_{2\ell}} \right)$$

$\Rightarrow H_0$ を棄却する.

次に $n_1, \dots, n_k \geq 30$ のときの手順を述べる.

- 手順 1. $0 < \alpha < 1$ となるように有意水準 α を決め,
 観測値 $\mathbf{X} \equiv (X_{11}, \dots, X_{kn_k})$ から,

$$Z_M^* \equiv \frac{(n-1)T_M^*}{\sum_{i=1}^k \sum_{j=1}^{n_i} \left\{ \psi\left(\frac{X_{ij} - \bar{\mu}}{\hat{s}_n}\right) - \bar{\psi}_{..}(\mathbf{X}) \right\}^2},$$

$$Z_M^\# \equiv \frac{(n-1)T_M^\#}{\sum_{i=1}^k \sum_{j=1}^{n_i} \left\{ \psi\left(\frac{X_{ij} - \bar{\mu}}{\hat{s}_n}\right) - \bar{\psi}_{..}(\mathbf{X}) \right\}^2}$$

とを計算する.

- 手順 2. $i = 1, \dots, k$ について $\lambda_i \equiv \frac{n_i}{n}$ とおく. シミュレーションの繰り
 返し数 M として 1000 を採り, $i = 1, 2, \dots, M-1$ について以下の手順 2.1 を繰り返す.

- ・ 手順 2.1. それぞれ $N(0, \frac{1}{\lambda_1}), \dots, N(0, \frac{1}{\lambda_k})$ に従う正規乱数 k 個 $y_1(i), \dots, y_k(i)$
 生成する. (2), (3) で定義した $Z_S^*(i)$ と $Z_S^\#(i)$ を計算する.

- 手順 3. $N'_1 \equiv \sum_{i=1}^{M-1} I(Z_M^* < Z_S^*(i))$, $N'_2 \equiv \sum_{i=1}^{M-1} I(Z_M^\# < Z_S^\#(i))$ とおく.

$$\frac{N'_1}{M} \leq \alpha \Rightarrow H_0 \text{ を棄却し, } H_1^* \text{ を受け入れる.}$$

$$\frac{N'_2}{M} \leq \alpha \Rightarrow H_0 \text{ を棄却し, } H_1^\# \text{ を受け入れる.}$$

[6] M 推定量; $\check{\alpha}_i (i = 1, \dots, k)$ と $\check{\nu}$ をそれぞれ『多標本問題』の 6 節 [8] で紹介した

α_i ($i = 1, \dots, k$) と ν の M 推定とする. $\check{\alpha}_1^*, \dots, \check{\alpha}_k^*$ を

$$\sum_{i=1}^k n_i \{\check{\alpha}_i^* - \check{\alpha}_i\}^2 = \min_{u_1 \leq \dots \leq u_k} \sum_{i=1}^k n_i \{u_i - \check{\alpha}_i\}^2$$

を満たすものとする. このとき, 傾向性の制約 $\mu_1 \leq \mu_2 \leq \dots \leq \mu_k$ の下でのそれぞれのパラメータの M 推定量は次のとおりである.

$\check{\alpha}_i^*$ ($i = 1, \dots, k$), $\check{\nu}$, $\check{\mu}_i^* = \check{\nu} + \check{\alpha}_i^*$ ($i = 1, \dots, k$). $\check{\alpha}_i^*$ は密度関数 $f(x) \equiv F'(x)$ に対称性を必要としないが, $\check{\nu}$, $\check{\mu}_i^*$ には $f(x)$ の対称性を必要とする.

密度関数 $f(x)$ に対称性を仮定できない場合は, $\tilde{\nu} = \bar{X}_{..}$ を ν の推定量とする. μ_i の推定量は

$$\check{\mu}_i^* = \tilde{\nu} + \check{\alpha}_i^* \quad (i = 1, \dots, k).$$