原 著

母分散が一樣でない多群モデルにおける
すべての母平均相違の閉検定手順

Closed Testing Procedures
for Pairwise Multiple Comparison
in Multi-Sample Models with Unequal Variances

白石高章・早川由宏

Taka-aki Shiraishi and Yoshihiro Hayakawa

南山大学理工学部

Faculty of Science and Engineering, Nanzan University

e-mail:marble@nanzan-u.ac.jp

We consider pairwise multiple comparison in k normal populations with unequal variances. Games and Howell (1976) proposed a single-step multiple comparison test based on two-sample test statistics of Welch (1949). We propose a closed testing procedure for pairwise multiple comparison. By computational simulations of the power, it is shown that the closed testing procedure is superior to the single-step multiple comparison test procedure of Games and Howell (1976).

Key words: Single-step procedures, Closed testing procedures, Exact theory, Asymptotic theory, Simulation.

1. はじめに

分散の等しい正規分布を仮定した多群モデルにおけるすべての平均相違の多重比較検定法が, Tukey (1953) と Kramer (1956) によって提案され, 現在ではテューキー・クレマー (Tukey-Kramer) 法とよばれている. 2 群間の t 検定統計量を $t_{ii'}$ とするとき, $\max_{i < i'} |t_{ii'}|$ の分布が学生化された範囲の $\sqrt{2}$ 倍の統計量の分布によって下から抑えられることを Hayter (1984) は示した. 長年正しいとされていた $\max_{i < i'} |t_{ii'}|$ の正確な分布は, 実際は $\max_{i < i'} |t_{ii'}|$ の分布を上から抑える分布であることを白石 (2006) は証明した. 白石 (2006) は, これらの不等式と数値積分を利用して, テューキー・クレマー法の保守度を調べることできた. 更に白石 (2011b) は Tukey-Kramer 法を優越する閉検定手順を提案した.

本論文では, 分散の異なる正規分布を仮定した多群モデルにおけるすべての平均相違の多重比較検定法について論じる. Welch (1949) の検定統計量を基にして, Games and Howell (1976) はシングルステップの多重比較検定法を提案した. 現在ではゲイムス・ハウエル (Games-Howell) 法とよばれている. マルチステップ法として閉検定手順を提案する. この閉検定手順がゲイムス・

ハウエル法を優越することをシミュレーションによって論述する.

2. モデルの設定と統計量の基本的性質

ある要因 A があり, k 個の水準 $A_1, \dots, A_k$ を考える. 水準は群ともよばれる. 水準 A_i における標本の観測値 $(X_{i1}, X_{i2}, \dots, X_{in_i})$ は第 i 標本または第 i 群とよばれ, μ_i, σ_i^2 は未知とし, X_{ij} が正規分布 $N(\mu_i, \sigma_i^2)$ に従うとする. すなわち,

$$P(X_{ij} \leq x) = \Phi\left(\frac{x - \mu_i}{\sigma_i}\right), \quad E(X_{ij}) = \mu_i, \quad V(X_{ij}) = \sigma_i^2 \quad (1)$$

である. さらにすべての X_{ij} は互いに独立であると仮定する. 総標本サイズを $n \equiv n_1 + \dots + n_k$ とおく. k 個の水準の平均母数のすべての比較を考える.

$$\text{帰無仮説 } H_{(i,i')}: \mu_i = \mu_{i'} \text{ vs. 対立仮説 } H_{(i,i')}^A: \mu_i \neq \mu_{i'} \quad (2)$$

に対して帰無仮説のファミリーを,

$$\mathcal{H} \equiv \{H_v \mid v \in \mathcal{U}\}$$

とおく. ただし,

$$\mathcal{U} \equiv \{(i, i') \mid 1 \leq i < i' \leq k\} \quad (3)$$

とする. 便宜上, 一様性の帰無仮説を

$$H_0: \mu_1 = \dots = \mu_k \quad (4)$$

とおく.

定数 α ($0 < \alpha < 1$) をはじめに決める.

$\mathbf{X} \equiv (X_{11}, \dots, X_{1n_1}, \dots, X_{k1}, \dots, X_{kn_k})$ の実現値 $\mathbf{x}$ によって, 任意の $H_v \in \mathcal{H}$ に対して H_v を棄却するかしないかを定める検定方式を $\phi_v(\mathbf{x})$ とする.

$\boldsymbol{\mu} \equiv (\mu_1, \dots, \mu_k)$ とおく. すべての $1 \leq i < i' \leq k$ に対して, $\mu_i \neq \mu_{i'}$ のときは, 有意水準は関係しないので,

$$\begin{aligned} \Theta &\equiv \{\boldsymbol{\mu} \mid 1 \text{ つ以上の帰無仮説 } H_{(i,i')} \text{ が真}\} \\ &= \{\boldsymbol{\mu} \mid \text{ある } i < i' \text{ が存在して, } \mu_i = \mu_{i'}\} \end{aligned} \quad (5)$$

とおき, $\boldsymbol{\mu} \in \Theta$ とする. このとき, 正しい帰無仮説 $H_{(i,i')}$ は 1 つ以上ある. また, 確率は $\boldsymbol{\mu}$ に依存するので, 確率測度を $P_{\boldsymbol{\mu}}(\cdot)$ で表わす.

このとき, 任意の $\boldsymbol{\mu} \in \Theta$ に対して

$$P_{\boldsymbol{\mu}}(\text{正しい帰無仮説のうち少なくとも 1 つが棄却される}) \leq \alpha \quad (6)$$

を満たす検定方式 $\{\phi_v(\mathbf{x}) \mid v \in \mathcal{U}\}$ を, $\mathcal{H}$ に対する水準 α の多重比較検定法とよんでいる. (6) の左辺を, ($\boldsymbol{\mu}$ を固定したときの) 第 1 種の過誤の確率またはタイプ I FWER (type I familywise error rate) とよぶ. また, (6) の右辺の α は全体としての有意水準である. 任意の $\boldsymbol{\mu} \in \Theta$ に対して (6) が成り立つことは,

$$\sup_{\mu \in \Theta} ((6) \text{ の左辺}) \leq \alpha$$

と同値である. すなわち, タイプ I FWER の上限が α 以下である必要がある.

3. ゲイムス・ハウエル (Games-Howell) の方法

ゲイムス・ハウエル法は Tukey 型の方法で, ウェルチ (Welch) の統計量を基礎とした多重比較法である. k 個の水準の平均母数のすべての比較を考える. 分散の一樣最小分散不偏推定量は,

$$\tilde{\sigma}_i^2 \equiv \frac{1}{n_i - 1} \sum_{j=1}^{n_i} (X_{ij} - \bar{X}_{i\cdot})^2 \quad (i = 1, \dots, k)$$

で与えられる. ただし, $\bar{X}_{i\cdot} \equiv \sum_{j=1}^{n_i} X_{ij} / n_i$ である. 統計量を,

$$T_{ii'}^G \equiv \frac{\bar{X}_{i\cdot} - \bar{X}_{i'\cdot}}{\sqrt{\frac{\tilde{\sigma}_i^2}{n_i} + \frac{\tilde{\sigma}_{i'}^2}{n_{i'}}}}$$

とする. $\mu \equiv (\mu_1, \dots, \mu_k)$ に対して

$$T_{ii'}^G(\mu) \equiv \frac{\bar{X}_{i\cdot} - \bar{X}_{i'\cdot} - (\mu_i - \mu_{i'})}{\sqrt{\frac{\tilde{\sigma}_i^2}{n_i} + \frac{\tilde{\sigma}_{i'}^2}{n_{i'}}}}$$

とおく. このとき, $t \geq 0$ に対して

$$P\left(\max_{1 \leq i < i' \leq k} |T_{ii'}^G(\mu)| \leq t\right) = P_0\left(\max_{1 \leq i < i' \leq k} |T_{ii'}^G| \leq t\right) \quad (7)$$

が成り立つ. ただし, $P_0(\cdot)$ は H_0 の下での確率測度とする. $[y]$ は y を超えない最大の整数とし

$$\hat{m}_{ii'} \equiv \left\lceil \frac{\left(\frac{\tilde{\sigma}_i^2}{n_i} + \frac{\tilde{\sigma}_{i'}^2}{n_{i'}}\right)^2}{\frac{\tilde{\sigma}_i^4}{n_i^2(n_i - 1)} + \frac{\tilde{\sigma}_{i'}^4}{n_{i'}^2(n_{i'} - 1)}} + \frac{1}{2} \right\rceil \quad (8)$$

とおくとき, (2) の帰無仮説 $H_{(i,i')}$ vs. 対立仮説 $H_{(i,i')}^A$ に対する分散の異なる場合の水準 α の 2 標本検定として, ウェルチは $T_{ii'}^G \geq t(\hat{m}_{ii'}; \alpha)$ のとき $H_{(i,i')}$ を棄却する方法を提案した. ただし, $t(\hat{m}_{ii'}; \alpha)$ は自由度 $\hat{m}_{ii'}$ の t 分布の上側 $100\alpha\%$ 点である. この手法はウェルチの検定とよばれしばしばデータ解析に用いられている. $2 \leq \ell \leq k$ となる整数 ℓ と 3 以上の整数 m_0 に対して

$$TA(t | \ell, m_0) \equiv \ell \int_0^\infty \left[\int_{-\infty}^\infty \{\Phi(x) - \Phi(x - \sqrt{2} \cdot ts)\}^{\ell-1} d\Phi(x) \right] g(s | m_0) ds$$

とする. ただし, $g(s | m_0)$ は, 確率変数 $\sqrt{\chi_{m_0}^2 / m_0}$ の従う分布の密度関数である. すなわち,

$$g(s | m_0) \equiv \frac{(m_0)^{m_0/2}}{\Gamma(m_0/2) 2^{m_0/2-1}} s^{m_0-1} \exp(-m_0 s^2 / 2)$$

である.

$$\text{方程式 } TA(t | \ell, m_0) = 1 - \alpha \text{ を満たす } t \text{ の解を } ta(\ell, m_0; \alpha) \quad (9)$$

とする. このとき, 次の多重比較法を紹介する.

[1] ゲイムス・ハウエル多重比較検定と同時信頼区間

{帰無仮説 $H_{(i,i')}: \mu_i = \mu_{i'}$ vs. 対立仮説 $H_{(i,i')}^A: \mu_i \neq \mu_{i'} \mid (i,i') \in \mathcal{U}$ } に対する水準 α のシングルステップ多重比較検定は, 次の (i), (ii) で与えられる.

- (i) $|T_{ii'}^G| \geq ta(k, \hat{m}_{ii'}; \alpha)$ となる i, i' に対して帰無仮説 $H_{(i,i')}$ を棄却し, 対立仮説 $H_{(i,i')}^A$ を受け入れ, $\mu_i \neq \mu_{i'}$ と判定する.
- (ii) $|T_{ii'}^G| < ta(k, \hat{m}_{ii'}; \alpha)$ となる i, i' に対して帰無仮説 $H_{(i,i')}$ を棄却しない.
- $\mu_i - \mu_{i'} \ (1 \leq i < i' \leq k)$ についての信頼係数 $1 - \alpha$ の同時信頼区間は,

$$\begin{aligned} & \bar{X}_i - \bar{X}_{i'} - ta(k, \hat{m}_{ii'}; \alpha) \cdot \sqrt{\frac{\tilde{\sigma}_i^2}{n_i} + \frac{\tilde{\sigma}_{i'}^2}{n_{i'}}} \\ & < \mu_i - \mu_{i'} < \bar{X}_i - \bar{X}_{i'} + ta(k, \hat{m}_{ii'}; \alpha) \cdot \sqrt{\frac{\tilde{\sigma}_i^2}{n_i} + \frac{\tilde{\sigma}_{i'}^2}{n_{i'}}} \quad (1 \leq i < i' \leq k) \end{aligned} \quad (10)$$

で与えられる.

[ゲイムス・ハウエル法多重比較検定法の正当性]

$\Theta_G \equiv \{\mu \mid 1 \text{ つ以上の帰無仮説 } H_{(i,i')} \text{ が真}\}$, $\mu_0 \in \Theta_G$ とする. この $\mu_0 \equiv (\mu_{01}, \dots, \mu_{0k})$ を与えたとき, 任意の $(i, i') \in \mathcal{A} \subset \mathcal{U}$ に対して帰無仮説 $H_{(i,i')}$ が真 ($\mu_{0i} = \mu_{0i'}$) で, 任意の $(i, i') \in \mathcal{A}^c \cap \mathcal{U}$ に対して帰無仮説 $H_{(i,i')}$ が偽 ($\mu_{0i} \neq \mu_{0i'}$) となるように $\mathcal{A} (\neq \emptyset)$ を決める. ただし, $\mathcal{U}$ は (3) で定義したものとする. $\mu \equiv (\mu_1, \dots, \mu_k)$ に対して

$$B_{ii'}(\mu) \equiv \{|T_{ii'}^G(\mu)| \geq ta(k, \hat{m}_{ii'}; \alpha)\}, \quad B_{ii'} \equiv \{|T_{ii'}^G| \geq ta(k, \hat{m}_{ii'}; \alpha)\}$$

とすると, $(i, i') \in \mathcal{A}$ と μ_0 に対して

$$\begin{aligned} B_{ii'} & \subset \bigcup_{(i,i') \in \mathcal{A}} B_{ii'} \\ & = \bigcup_{(i,i') \in \mathcal{A}} B_{ii'}(\mu_0) \\ & \subset \bigcup_{(i,i') \in \mathcal{U}} B_{ii'}(\mu_0) = \bigcup_{(i,i') \in \mathcal{U}} \{|T_{ii'}^G(\mu_0)| \geq ta(k, \hat{m}_{ii'}; \alpha)\} \end{aligned} \quad (11)$$

が成り立つ.

$$\Theta_1 \equiv \{\mu \mid (i, i') \in \mathcal{A} \text{ に対して } \mu_i = \mu_{i'}, (i, i') \in \mathcal{A}^c \cap \mathcal{U} \text{ に対して } \mu_i \neq \mu_{i'}\}$$

とおく. $\mu_0 \in \Theta_1$ に対して

$$\begin{aligned} P_{\mu_0} \left(\bigcup_{(i,i') \in \mathcal{A}} B_{ii'}(\mu_0) \right) & = P_{\mu_0} \left(\bigcup_{(i,i') \in \mathcal{A}} B_{ii'} \right) \\ & \leq P_0 \left(\bigcup_{(i,i') \in \mathcal{U}} \{|T_{ii'}^G| \geq ta(k, \hat{m}_{ii'}; \alpha)\} \right) \end{aligned} \quad (12)$$

が示される. ただし, $P_0(\cdot)$ は (4) の一様性の帰無仮説 H_0 の下での確率測度とする. (11) と (12) より, ゲイムス・ハウエル法では

$$P_0 \left(\bigcup_{(i,i') \in \mathcal{U}} \{|T_{ii'}^G| \geq ta(k, \hat{m}_{ii'}; \alpha)\} \right) \lesssim \alpha$$

が成り立つ必要がある。これは

$$P_0 \left(\bigcap_{(i,i') \in \mathcal{U}} \{|T_{ii'}^G| < ta(k, \hat{m}_{ii'}; \alpha)\} \right) \gtrsim 1 - \alpha \quad (13)$$

と同値である。さらに任意の μ に対して

$$P_\mu \left(\bigcup_{(i,i') \in \mathcal{U}} B_{ii'}(\mu) \right) = P_0 \left(\bigcup_{(i,i') \in \mathcal{U}} B_{ii'} \right) \lesssim \alpha$$

である。

4. 閉検定手順

$\emptyset \subsetneq V \subset \mathcal{U}$ を満たす V に対して, $\bigwedge_{v \in V} H_v$ は k 個の母平均に関していくつかが等しいという仮説となる。 $I_1, \dots, I_J$ ($I_j \neq \emptyset, j = 1, \dots, J$) を $1, \dots, k$ の互いに素な部分集合の組と同じ I_j ($j = 1, \dots, J$) に含まれる添字をもつ母平均は等しいという帰無仮説を $H(I_1, \dots, I_J)$ で表す。このとき, $\emptyset \subsetneq V \subset \mathcal{U}$ を満たす任意の V に対して, ある自然数 J と上記のある $I_1, \dots, I_J$ が存在して

$$\bigwedge_{v \in V} H_v = H(I_1, \dots, I_J) \quad (14)$$

が成り立つ。 $M \equiv M(I_1, \dots, I_J) \equiv \sum_{j=1}^J \ell_j$, $\ell_j \equiv \#(I_j)$ とする。

[2] 閉検定手順

(a) $J \geq 2$ のとき, $\ell = \ell_1, \dots, \ell_J$ に対して

$$\alpha(M, \ell) \equiv 1 - (1 - \alpha)^{\frac{\ell}{M}} \quad (15)$$

で $\alpha(M, \ell)$ を定義する。 $1 \leq j \leq J$ となる整数 j と $i, i' \in I_j$ となる $i < i'$ が存在して $ta(\ell_j, \hat{m}_{ii'}; \alpha(M, \ell_j)) \leq |T_{ii'}^G|$ ならば (14) の帰無仮説 $\bigwedge_{v \in V} H_v$ を棄却する。

(b) $J = 1$ ($M = \ell_1$) のとき, $i, i' \in I_1$ となる $i < i'$ が存在して $ta(M, \hat{m}_{ii'}; \alpha) \leq |T_{ii'}^G|$ ならば帰無仮説 $\bigwedge_{v \in V} H_v$ を棄却する。

(a), (b) の方法で, $(i, i') \in V \subset \mathcal{U}$ を満たす任意の V に対して, $\bigwedge_{v \in V} H_v$ が棄却されるとき, 多重比較検定として, $H_{(i,i')}$ を棄却する。

定理 3.1 [2] の閉検定手順は近似的に水準 α の多重比較検定である。

証明 (a) の検定の近似的な有意水準は α であることを示せばよい。

$$P_0(\text{ある } j \text{ と } i, i' \in I_j \text{ が存在して, } |T_{ii'}^G| \geq ta(\ell_j, \hat{m}_{ii'}; \alpha(M, \ell_j)))$$

$$= 1 - P_0(\text{任意の } j \text{ と } i, i' \in I_j \text{ となる任意の } i, i' \text{ に対して } |T_{ii'}^G| < ta(\ell_j, \hat{m}_{ii'}; \alpha(M, \ell_j))) \quad (16)$$

である.

補事象の確率は

$$\begin{aligned} & P_0(\text{任意の } j \text{ と } i, i' \in I_j \text{ となる任意の } i, i' \text{ に対して } |T_{ii'}^G| < ta(\ell_j, \hat{m}_{ii'}; \alpha(M, \ell_j))) \\ &= P_0 \left(\bigcap_{j=1}^J \left[\bigcap_{i, i' \in I_j} \{|T_{ii'}^G| < ta(\ell_j, \hat{m}_{ii'}; \alpha(M, \ell_j))\} \right] \right) \end{aligned}$$

ここで $\bigcap_{i, i' \in I_j} \{|T_{ii'}^G| < ta(\ell_j, \hat{m}_{ii'}; \alpha(M, \ell_j))\}$ は $C_j \equiv \{X_{im} \mid m = 1, \dots, n_i, i \in I_j\}$ の関数で $I_1 \cup \dots \cup I_J \subset \{1, \dots, k\}$, $I_j \cap I_{j'} = \emptyset$ ($j \neq j'$) であるから $C_1, \dots, C_J$ は互いに独立である. このことと (13) により

$$\begin{aligned} & P_0(\text{任意の } j \text{ と } i, i' \in I_j \text{ となる任意の } i, i' \text{ に対して } |T_{ii'}^G| < ta(k, \hat{m}_{ii'}; \alpha)) \\ &= \prod_{j=1}^J P_0 \left(\bigcap_{i, i' \in I_j} \{|T_{ii'}^G| < ta(\ell_j, \hat{m}_{ii'}; \alpha(M, \ell_j))\} \right) \\ &\gtrsim \prod_{j=1}^J (1 - \alpha(M, \ell_j)) \\ &= \prod_{j=1}^J (1 - \alpha)^{\frac{\ell_j}{M}} = 1 - \alpha \end{aligned} \quad (17)$$

となる.

(16), (17) 式より

$$P_0(\text{ある } j \text{ と } i, i' \in I_j \text{ が存在して, } |T_{ii'}^G| \geq ta(\ell_j, \hat{m}_{ii'}; \alpha(M, \ell_j))) \lesssim \alpha$$

が成り立ち, 結論を得る. □

5. 漸近理論

ゲイムス・ハウエル法の標本サイズが大きい場合を考える.

(条件 1) $\lim_{n \rightarrow \infty} (n_i/n) = \lambda_i > 0 \quad (1 \leq i \leq k)$

を仮定する. 統計量 $T_{ii'}^G$ について次の補題 3.2 が得られる.

補題 3.2 (条件 1) の下で

$$\max_{1 \leq i < i' \leq k} |T_{ii'}^G(\boldsymbol{\mu})| \xrightarrow{\mathcal{L}} \max_{1 \leq i < i' \leq k} \frac{|Y_{i'} - Y_i|}{\sqrt{\frac{\sigma_i^2}{\lambda_i} + \frac{\sigma_{i'}^2}{\lambda_{i'}}}}$$

が成り立つ. ただし, $Y_1, \dots, Y_k$ は互いに独立で $Y_i \sim N(0, \sigma_i^2/\lambda_i)$ とする.

証明 中心極限定理により,

$$\sqrt{n_i}(\bar{X}_{i\cdot} - \mu_i) \xrightarrow{\mathcal{L}} \sigma_i Z_i \sim N(0, \sigma_i^2)$$

が成り立つ. ここで, $Z_i \sim N(0, 1)$, $Z_1, \dots, Z_k$ は互いに独立である. これにより,

$$\sqrt{n}(\bar{X}_{i\cdot} - \mu_i) = \left(\frac{\sqrt{n}}{\sqrt{n_i}} \right) \sqrt{n_i}(\bar{X}_{i\cdot} - \mu_i) \xrightarrow{\mathcal{L}} \frac{\sigma_i}{\sqrt{\lambda_i}} Z_i \sim N\left(0, \frac{\sigma_i^2}{\lambda_i}\right)$$

である. 大数の法則より, $\hat{\sigma}_i^2 \xrightarrow{P} \sigma_i^2$ が成り立つ. ここで, スラツキーの定理より,

$$T_{ii'}^G(\mu) = \frac{\sqrt{n}(\bar{X}_{i'\cdot} - \mu_{i'}) - \sqrt{n}(\bar{X}_{i\cdot} - \mu_i)}{\sqrt{\frac{n}{n_i}\hat{\sigma}_i^2 + \frac{n}{n_{i'}}\hat{\sigma}_{i'}^2}} \xrightarrow{\mathcal{L}} \frac{\frac{\sigma_{i'}}{\sqrt{\lambda_{i'}}}Z_{i'} - \frac{\sigma_i}{\sqrt{\lambda_i}}Z_i}{\sqrt{\frac{\sigma_i^2}{\lambda_i} + \frac{\sigma_{i'}^2}{\lambda_{i'}}}} \sim N(0, 1)$$

を得る. ゆえに補題の主張

$$\begin{aligned} \max_{1 \leq i < i' \leq k} |T_{ii'}^G(\mu)| &\xrightarrow{\mathcal{L}} \max_{1 \leq i < i' \leq k} \frac{\left| \frac{\sigma_{i'}}{\sqrt{\lambda_{i'}}}Z_{i'} - \frac{\sigma_i}{\sqrt{\lambda_i}}Z_i \right|}{\sqrt{\frac{\sigma_i^2}{\lambda_i} + \frac{\sigma_{i'}^2}{\lambda_{i'}}}} \\ &\iff \max_{1 \leq i < i' \leq k} |T_{ii'}^G(\mu)| \xrightarrow{\mathcal{L}} \max_{1 \leq i < i' \leq k} \frac{|Y_{i'} - Y_i|}{\sqrt{\frac{\sigma_i^2}{\lambda_i} + \frac{\sigma_{i'}^2}{\lambda_{i'}}}} \end{aligned}$$

が示される. $\square$

補題 3.2 の条件 $X_{ij} \sim N(\mu_i, \sigma_i^2)$ の替わりに, $F(x)$ を連続型の分布の分布関数とし, 一般的な

$$(\text{条件 2}) \quad P(X_{ij} \leq x) = F\left(\frac{x - \mu_i}{\sigma_i}\right), \quad \int_{-\infty}^{\infty} x dF(x) dx = 0, \quad \int_{-\infty}^{\infty} x^2 dF(x) dx = 1$$

に変えても補題 3.2 は成り立つ. すなわち, 連続型の分布ならば正規分布でなくても補題 3.2 は成り立つ.

補題 3.2 から次の命題 3.3 を得る.

命題 3.3 補題 3.2 の条件の下で,

$$A(t) \leq \lim_{n \rightarrow \infty} P\left(\max_{1 \leq i < i' \leq k} |T_{ii'}^G(\mu)| \leq t\right) \leq A^*(t) \quad (18)$$

が成り立つ. ただし,

$$\begin{aligned} A(t) &\equiv k \int_{-\infty}^{\infty} \{\Phi(x) - \Phi(x - \sqrt{2} \cdot t)\}^{k-1} d\Phi(x), \\ A^*(t) &\equiv \int_{-\infty}^{\infty} \sum_{i'=1}^k \prod_{\substack{i=1 \\ i \neq i'}}^k \left\{ \Phi\left(\sqrt{\frac{\lambda_i \sigma_{i'}^2}{\lambda_{i'} \sigma_i^2}} \cdot x\right) - \Phi\left(\sqrt{\frac{\lambda_i \sigma_{i'}^2}{\lambda_{i'} \sigma_i^2}} \cdot x - \sqrt{\frac{\lambda_i \sigma_{i'}^2 + \lambda_{i'} \sigma_i^2}{\lambda_{i'} \sigma_i^2}} \cdot t\right) \right\} d\Phi(x) \end{aligned} \quad (19)$$

とする. $\sigma_1^2/\lambda_1 = \dots = \sigma_k^2/\lambda_k$ が満たされるとき, 上の両方の不等式で等号が成り立つ.

証明 任意の $t \geq 0$ に対して,

$$\lim_{n \rightarrow \infty} P \left(\max_{1 \leq i < i' \leq k} |T_{ii'}^G(\boldsymbol{\mu})| \leq t \right) = P \left(\max_{1 \leq i < i' \leq k} \frac{|Y_i - Y_{i'}|}{\sqrt{\frac{\sigma_i^2}{\lambda_i} + \frac{\sigma_{i'}^2}{\lambda_{i'}}}} \leq t \right)$$

が補題 3.2 より成り立つ. ただし, $Y_i \sim N(0, \sigma_i^2/\lambda_i)$ とする.

σ_i^2 を σ_i^2/λ_i として, 白石 (2011a) の付録の定理 A.5 を適用すると,

$$A(t) \leq \lim_{n \rightarrow \infty} P \left(\max_{1 \leq i < i' \leq k} |T_{ii'}^G(\boldsymbol{\mu})| \leq t \right) \leq A^*(t)$$

が導かれる. □

(20) の $A(t)$ に対応して, $2 \leq \ell \leq k$ となる整数 ℓ に対して

$$A(t | \ell) \equiv \ell \int_{-\infty}^{\infty} \{\Phi(x) - \Phi(x - \sqrt{2} \cdot ts)\}^{\ell-1} d\Phi(x)$$

とし,

$$\text{方程式 } A(t | \ell) = 1 - \alpha \text{ を満たす } t \text{ の解を } a(\ell; \alpha) \quad (20)$$

とする. $n_1, \dots, n_k$ が大きいとき, 命題 3.3 より, $\{\text{帰無仮説 } H_{(i,i')} \text{ vs. 対立仮説 } H_{(i,i')}^A \mid 1 \leq i < i' \leq k\}$ に対する水準 α の多重比較検定は, 次で与えられる.

- (i) $|T_{ii'}^G| \geq a(k; \alpha)$ となる i, i' に対して帰無仮説 $H_{(i,i')}$ を棄却し, 対立仮説 $H_{(i,i')}^A$ を受け入れ, $\mu_i \neq \mu_{i'}$ と判定する.
- (ii) $|T_{ii'}^G| < a(k; \alpha)$ となる i, i' に対して帰無仮説 $H_{(i,i')}$ を棄却しない.

$\mu_i - \mu_{i'} \ (1 \leq i < i' \leq k)$ についての信頼係数 $1 - \alpha$ の漸近的な同時信頼区間は,

$$\bar{X}_i - \bar{X}_{i'} - a(k; \alpha) \cdot \sqrt{\frac{\tilde{\sigma}_i^2}{n_i} + \frac{\tilde{\sigma}_{i'}^2}{n_{i'}}} < \mu_i - \mu_{i'} < \bar{X}_i - \bar{X}_{i'} + a(k; \alpha) \cdot \sqrt{\frac{\tilde{\sigma}_i^2}{n_i} + \frac{\tilde{\sigma}_{i'}^2}{n_{i'}}} \quad (1 \leq i < i' \leq k)$$

で与えられる.

[3] 漸近的な閉検定手順

$T^G(I_j) \equiv \max_{i < i', i, i' \in I_j} |T_{ii'}^G|$ とする.

(A) $J \geq 2$ のとき, (15) 式で $\alpha(M, \ell)$ を定義する. $1 \leq j \leq J$ となる整数 j と $i, i' \in I_j$ となる $i < i'$ が存在して $a(\ell_j; \alpha(M, \ell_j)) \leq T^G(I_j)$ ならば (14) の帰無仮説 $\bigwedge_{v \in V} H_v$ を棄却する.

(B) $J = 1 \ (M = \ell_1)$ のとき, $i, i' \in I_1$ となる $i < i'$ が存在して $a(M; \alpha) \leq T^G(I_j)$ ならば帰無仮説 $\bigwedge_{v \in V} H_v$ を棄却する.

(A), (B) の方法で, $(i, i') \in V \subset \mathcal{U}$ を満たす任意の V に対して, (14) の $\bigwedge_{v \in V} H_v$ が棄却されるとき, 漸近的な多重比較検定として, $H_{(i,i')}$ を棄却する.

定理 3.4 (条件 1), (条件 2) の下で, [3] の閉検定手順は水準 α の漸近的な多重比較検定である.

証明 (B) の検定の漸近的な有意水準が α であることは自明であるので, (A) の検定の漸近的な

有意水準が α であることを示せばよい.

$T^G(I_1), \dots, T^G(I_J)$ は互いに独立より,

$$\lim_{n \rightarrow \infty} P_0(T^G(I_j) < a(\ell_j; \alpha(M, \ell_j)), j = 1, \dots, J) = \prod_{j=1}^J \left\{ \lim_{n \rightarrow \infty} P_0(T^G(I_j) < a(\ell_j; \alpha(M, \ell_j))) \right\} \quad (21)$$

が成り立つ.

命題 3.3 より

$$\begin{aligned} (21) &\geq \prod_{j=1}^J A(a(\ell_j; \alpha(M, \ell_j)) \mid \ell_j) \\ &= \prod_{j=1}^J (1 - \alpha(M, \ell_j)) \\ &= \prod_{j=1}^J \left\{ (1 - \alpha)^{\frac{\ell_j}{M}} \right\} = 1 - \alpha \end{aligned} \quad (22)$$

を得る.

(22) 式より

$$\begin{aligned} &\lim_{n \rightarrow \infty} P_0(\text{ある } j \text{ が存在して, } T^G(I_j) \geq a(\ell_j; \alpha(M, \ell_j))) \\ &= 1 - \lim_{n \rightarrow \infty} P_0(T^G(I_j) < a(\ell_j; \alpha(M, \ell_j)), j = 1, \dots, J) \\ &\leq \alpha \end{aligned}$$

が示される. ここで, (A) の検定は, 漸近的に有意水準 α である. $\square$

6. シミュレーションによる検出力の優越性

ゲイムス・ハウエルの多重比較検定法を (GH), [2] の閉検定手順を (SH) とする. これらの多重比較検定法がタイプ I FWER の水準が近似的にたもたれていることをシミュレーションにより検証する. 群の数 $k = 4$, 有意水準 $\alpha = 0.05, 0.01$, 標本サイズ $n_1 = n_2 = n_3 = n_4 = 5, 10, 20$ とする. (12) より, タイプ I FWER は, 一樣性の帰無仮説 H_0 の下で最大値をとるので, $X_{i1} \sim N(0, 1 + (\ell - 1)(i - 1)/3)$ ($i = 1, \dots, 4; \ell = 1, \dots, 5, 10$) とした正規分布に従う標本を生成した. ゲイムス・ハウエルの多重比較検定法と提案した閉検定手順のタイプ I FWER の最大値をそれぞれ GHF , SHF とし, SHF と GHF の推定値を表 1 に載せている. シミュレーションの繰り返し数を 100,000 回とした. この表から, サイズの小さい $n_1 = 5$ とした場合を除いて, 水準はおおよそ保たれている. 標本サイズを不揃いの $n_i = 5i, 10i, 20i$ ($i = 1, 2, 3, 4$) とした場合の SHF と GHF の推定値を表 2 に載せている.

ゲイムス・ハウエル法 (GH) と提案した閉検定手順 (SH) による総対検出力をシミュレーションを用いて比較する. 総対検出力は, Ramsey (1979) によって導入され, 松田・永田 (1990) によって日本語に訳された. 群の数 $k = 4$, 有意水準 $\alpha = 0.05$, 標本サイズ $n_i = 20$ とし表 3-5 のように平均が等間隔の場合 (EQ1-3), 最小範囲の場合 (MIN1-3), 最大範囲の場合 (MAX1-3) に, それぞれ 100,000 回シミュレーションを行った. その結果を表 6-8 に掲載している. これらの表から, Jpn J Biomet Vol.35, No.2, 2014

表 1. $k = 4$, $n_1 = \dots = n_4 = 5, 10, 20$, $X_{i1} \sim N(0, 1 + (\ell - 1)(i - 1)/3)$ ($i = 1, \dots, 4$; $\ell = 1, \dots, 5, 10$) とした場合の提案した閉検定手順のタイプ I FWER の最大値 SHF とゲイムス・ハウエルの多重比較検定法のタイプ I FWER の最大値 GHF の値

(i) 水準 0.05, $n_1 = n_2 = n_3 = n_4 = 5$ のとき						
ℓ の値	1	2	3	4	5	10
SHF	0.0420	0.0445	0.0466	0.0490	0.0492	0.0505
GHF	0.0462	0.0479	0.0510	0.0547	0.0549	0.0569
(ii) 水準 0.05, $n_1 = 10$ のとき						
ℓ の値	1	2	3	4	5	10
SHF	0.0492	0.0496	0.0500	0.0499	0.0493	0.0489
GHF	0.0509	0.0514	0.0510	0.0502	0.0495	0.0486
(iii) 水準 0.05, $n_1 = 20$ のとき						
ℓ の値	1	2	3	4	5	10
SHF	0.0501	0.0504	0.0487	0.0486	0.0480	0.0475
GHF	0.0511	0.0508	0.0482	0.0470	0.0464	0.0451
(iv) 水準 0.01, $n_1 = 5$ のとき						
ℓ の値	1	2	3	4	5	10
SHF	0.0066	0.0072	0.0088	0.0097	0.0103	0.0109
GHF	0.0070	0.0078	0.0100	0.0109	0.0120	0.0129
(v) 水準 0.01, $n_1 = 10$ のとき						
ℓ の値	1	2	3	4	5	10
SHF	0.0093	0.0097	0.0102	0.0102	0.0102	0.0099
GHF	0.0095	0.0099	0.0105	0.0107	0.0109	0.0100
(vi) 水準 0.01, $n_1 = 20$ のとき						
ℓ の値	1	2	3	4	5	10
SHF	0.0097	0.0103	0.0097	0.0096	0.0096	0.0093
GHF	0.0099	0.0103	0.0095	0.0094	0.0094	0.0087

(SH) による総対検出力, (GH) による総対検出力をそれぞれ (SHP), (GHP) とすると,

① 等間隔 EQ1, EQ2, EQ3 のとき

$$0.2705 \leq (\text{SHP}) - (\text{GHP}) \leq 0.3825$$

② 最小範囲 MIN1, MIN2, MIN3 のとき

$$0.1741 \leq (\text{SHP}) - (\text{GHP}) \leq 0.2468$$

③ 最大範囲 MAX1, MAX2, MAX3 のとき

$$0.1723 \leq (\text{SHP}) - (\text{GHP}) \leq 0.3011$$

を得, 提案した閉検定手順による総対検出力が Games-Howell 法による総対検出力よりも非常に高い. 特に, 等間隔のときが顕著である.

表 2. $k = 4$, $n_i = 5i, 10i, 20i$, $X_{i1} \sim N(0, 1 + (\ell - 1)(i - 1)/3)$ ($i = 1, \dots, 4$; $\ell = 1, \dots, 5, 10$) とした場合の提案した閉検定手順のタイプ I FWER の最大値 SHF とゲイムス・ハウエルの多重比較検定法のタイプ I FWER の最大値 GHF の値

(i) 水準 0.05, $n_i = 5i$ ($i = 1, 2, 3, 4$) のとき						
ℓ の値	1	2	3	4	5	10
SHF	0.0532	0.0509	0.0491	0.0474	0.0473	0.0487
GHF	0.0572	0.0536	0.0503	0.0480	0.0471	0.0483
(ii) 水準 0.05, $n_i = 10i$ のとき						
ℓ の値	1	2	3	4	5	10
SHF	0.0512	0.0502	0.0496	0.0500	0.0498	0.0488
GHF	0.0516	0.0515	0.0503	0.0501	0.0487	0.0468
(iii) 水準 0.05, $n_i = 20i$ のとき						
ℓ の値	1	2	3	4	5	10
SHF	0.0503	0.0505	0.0502	0.0492	0.0493	0.0491
GHF	0.0499	0.0509	0.0507	0.0494	0.0486	0.0471
(iv) 水準 0.01, $n_i = 5i$ のとき						
ℓ の値	1	2	3	4	5	10
SHF	0.0123	0.0106	0.0094	0.0089	0.0090	0.0095
GHF	0.0137	0.0114	0.0094	0.0090	0.0090	0.0094
(v) 水準 0.01, $n_i = 10i$ のとき						
ℓ の値	1	2	3	4	5	10
SHF	0.0104	0.0100	0.0100	0.0099	0.0097	0.0095
GHF	0.0105	0.0100	0.0102	0.0098	0.0098	0.0092
(vi) 水準 0.01, $n_i = 20i$ のとき						
ℓ の値	1	2	3	4	5	10
SHF	0.0103	0.0100	0.0098	0.0096	0.0101	0.0096
GHF	0.0106	0.0102	0.0098	0.0092	0.0100	0.0091

7. 適用例

Bönsch et al. (2005) は, アルコール依存患者において対立遺伝子の長さが発現した α シヌクレインの mRNA のレベルと何らかの関連を示すことを主張している. そのデータが Everitt and Hothorn (2010) の表 14.1 に載せられている. Everitt and Hothorn (2010) の中で, 3 つの群: 対立遺伝子の長さが短い ($n_1 = 24$), 普通 ($n_2 = 58$), 長い ($n_3 = 15$) にカテゴリーされ, 3 つの群において, α シヌクレインの mRNA の発現レベルの分布の平均相違の比較が水準 0.05 の Tukey-Kramer の方法 (TK) により多重比較検定されたがいずれの帰無仮説も棄却されなかった. さらに, Everitt and Hothorn (2010) は分散の不均一性を指摘している. 実際, 統計解析システム「R」を使ってバートレット検定を実行すると分散の一樣性の検定の p 値は 0.0482 となり水準 0.05 で帰無仮説は棄却される. Games-Howell の方法 (GH), 閉検定手順 (SH) により, 有意水準 5% の多重比較検定を行う. $T_{ii'}^G$ の値は

表 3. 等間隔の 4 群モデル設定

	EQ1		EQ2		EQ3	
	平均	分散	平均	分散	平均	分散
第 1 群	Δ	1	Δ	1	Δ	1
第 2 群	2Δ	1	2Δ	1.33	2Δ	2
第 3 群	3Δ	1	3Δ	1.66	3Δ	3
第 4 群	4Δ	1	4Δ	2	4Δ	4

表 4. 最小範囲の 4 群モデル設定

	MIN1		MIN2		MIN3	
	平均	分散	平均	分散	平均	分散
第 1 群	$-\Delta$	1	$-\Delta$	1	$-\Delta$	1
第 2 群	$-\Delta$	1	$-\Delta$	1.33	$-\Delta$	2
第 3 群	Δ	1	Δ	1.66	Δ	3
第 4 群	Δ	1	Δ	2	Δ	4

表 5. 最大範囲の 4 群モデル設定

	MAX1		MAX2		MAX3	
	平均	分散	平均	分散	平均	分散
第 1 群	$-\sqrt{2}\Delta$	1	$-\sqrt{2}\Delta$	1	$-\sqrt{2}\Delta$	1
第 2 群	0	1	0	1.33	0	2
第 3 群	0	1	0	1.66	0	3
第 4 群	$\sqrt{2}\Delta$	1	$\sqrt{2}\Delta$	2	$\sqrt{2}\Delta$	4

表 6. 表 3 の等間隔とした場合の提案した閉検定手順(SH)の総対検出力とゲイムス・ハウエル(GH)の多重比較検定法の総対検出力の値

EQ1			EQ2			EQ3		
Δ	(SH)	(GH)	Δ	(SH)	(GH)	Δ	(SH)	(GH)
0.884	0.4003	0.0866	1.354	0.3992	0.0978	2.394	0.3996	0.1192
0.988	0.6019	0.2194	1.539	0.6002	0.2455	2.781	0.6000	0.2719
1.122	0.8014	0.4588	1.791	0.7998	0.5031	3.317	0.7999	0.5294

表 7. 表 4 の最小範囲での (SH) と (GH) の総対検出力の値

MIN1			MIN2			MIN3		
Δ	(SH)	(GH)	Δ	(SH)	(GH)	Δ	(SH)	(GH)
0.435	0.3997	0.1899	0.674	0.4005	0.1953	1.197	0.4004	0.2048
0.503	0.6004	0.3536	0.783	0.6005	0.3635	1.404	0.6010	0.3811
0.588	0.8002	0.5956	0.899	0.8003	0.6069	1.670	0.8005	0.6264

表 8. 表 5 の最大範囲での (SH) と (GH) の総対検出力の値

MAX1			MAX2			MAX3		
Δ	(SH)	(GH)	Δ	(SH)	(GH)	Δ	(SH)	(GH)
0.645	0.3996	0.1431	0.988	0.3994	0.1469	1.755	0.4008	0.1635
0.729	0.6007	0.2996	1.138	0.6000	0.3211	2.077	0.6005	0.3505
0.838	0.7999	0.5513	1.354	0.8002	0.5985	2.546	0.8006	0.6283

$$T_{12}^G = 1.005, T_{13}^G = 2.615, T_{23}^G = 2.314$$

となり, $\hat{m}_{12} = 38, \hat{m}_{13} = 36, \hat{m}_{23} = 36$ である. 白石 (2011a) の付表 10 より, $ta(3, 38; 0.05) = 2.439, ta(3, 36; 0.05) = 2.444$ が得られる. ここで,

$$T_{12}^G = 1.005 < 2.439 = ta(3, \hat{m}_{12}; 0.05), T_{13}^G > ta(3, \hat{m}_{13}; 0.05), T_{23}^G < ta(3, \hat{m}_{23}; 0.05) \quad (23)$$

が導かれ, Games-Howell の方法 (GH) により, 帰無仮説 H_{13} だけが棄却される. さらに, 白石 (2011a) の付表 2 の t 分布表により, $ta(2, 38; 0.05) = 2.024, ta(2, 36; 0.05) = 2.028$ が得られる. ここで,

$$T_{12}^G < 2.024 = ta(2, \hat{m}_{12}; 0.05), T_{13}^G > ta(2, \hat{m}_{13}; 0.05), T_{23}^G > ta(2, \hat{m}_{23}; 0.05) \quad (24)$$

が導かれる. (23), (24) より, 閉検定手順 (SH) を使って検定を行うと, 帰無仮説 H_{13}, H_{23} が棄却され, H_{12} は棄却されない. 以上の水準 0.05 の多重比較検定の結果としてまとめたものを表 9 に示している.

表 9. 対立遺伝子の長さによって分けられた 3 群における発現した α シヌクレインの mRNA のレベルの平均相違の多重比較検定の結果

2 群の組み合わせ	GH	SH	TK
(短い, 普通)	棄却されない	棄却されない	棄却されない
(短い, 長い)	棄却される	棄却される	棄却されない
(普通, 長い)	棄却されない	棄却される	棄却されない

表 9 より, 水準 0.05 の閉検定手順 (SH) により, 2 つの帰無仮説が棄却された. Everitt and Hothorn (2010) に載せられている点推定値の結果では, 対立遺伝子の長さが大きいほど α シヌクレインの mRNA のレベルの平均値は大きくなっている.

8. おしまいに

松田 (1994) は, 群サイズが小さく不揃いで不等分散の場合にはゲイムス・ハウエル法が有意水準を保たれていないことをシミュレーション結果として指摘している. 表 2 の $n_i = 5i$ の場合の (i) と (iv) の場合には有意水準を保たれていない. このため, 松田 (1994) は, 等分散性をパートレット検定統計量を使って予備検定をおこないゲイムス・ハウエル法と Dunnett (1980) の方法を組み合わせた多重比較検定を提唱している. 我々が提案している閉検定手順の SHF は, 表 2 の $n_i = 5i$ の場合の (i) と (iv) の場合に保守的ではないが, ゲイムス・ハウエル法よりも有意水準に近い値となっている. 表 1, 2 より, 群サイズが 10 以上の場合, SHF はおおそ保守的になっている. (条件 1) の下で $n \rightarrow \infty$ のとき, (8) の $\hat{m}_{ij}$ は確率的に大きい値となり, $ta(k, \hat{m}_{ii}; \alpha)$ は $a(k; \alpha)$ に収束する. (7) と (18) により, ゲイムス・ハウエル法と提案している閉検定手順は n が大きいときに保守的な手法となっている. 定理 3.4 より, [3] の漸近的な手法も保守的な手法であることが解る.

謝 辞

本論文の修正にあたり, 貴重なご指摘と有益なご助言をいただいた2人の査読者に変感謝致します。本研究は, 日本学術振興会科学研究費補助金基盤研究(C)(課題番号 24540148) 及び 2014年度南山大学パツへ研究奨励金 I-A-2 の援助を受けて行ったものである。

参考文献

- Bönsch, D., Lederer, T., Reulbach, U., Hothorn, T., Kornhuber, J. and Bleich, S. (2005). Joint analysis of the NACP-REP1 marker within the alpha synuclein gene concludes association with alcohol dependence. *Human Molecular Genetics*, **14**, 967–971.
- Dunnett, C. W. (1980). Pairwise multiple comparison procedures in the unequal variance case. *Journal of the American Statistical Association*, **75**, 796–800.
- Games, P. A. and Howell, J. F. (1976). Pairwise multiple comparison procedures with unequal N's and/or Variances: A Monte Carlo Study. *Journal of Educational Statistics*, **1**, 113–125.
- Hayter, A. J. (1984). A proof of the conjecture that the Tukey-Kramer multiple comparisons procedure is conservative. *Annals of Statistics*, **12**, 61–75.
- Kramer, C. Y. (1956). Extension of multiple range tests to group means with unequal numbers of replications. *Biometrics*, **8**, 75–86.
- Ramsey, P. H. (1978). Power differences between pairwise multiple comparisons. *Journal of the American Statistical Association*, **73**, 479–485.
- Tukey, J. W. (1953). The Problem of Multiple Comparisons. *The Collected Works of John W. Tukey, J. W. (1994), Volume VIII. Multiple Comparisons*. Chapman and Hall.
- Welch, B. L. (1938). The significance of the difference between two means when the population variances are unequal. *Biometrika*, **25**, 350–362.
- 白石高章 (2006). Tukey-Kramer 法に関連した分布の上界. 計算機統計学会 和文誌, **19**, 77–87.
- 白石高章 (2011a). 多群連続モデルにおける多重比較法—パラメトリック, ノンパラメトリックの数理統計. 共立出版.
- 白石高章 (2011b). 多群モデルにおけるすべての平均相違に関する閉検定手順. 計量生物学, **32**, 33–47.
- 松田眞一, 永田靖 (1990). 多重比較における新たな検出力の提案と各手法の特徴比. 応用統計学, **19**, 93–113.
- 松田眞一 (1994). 一元配置多重比較の予備検定の改良. 応用統計学, **23**, 129–145.