分散の異なる場合の多群正規標本モデルにおける すべての平均相違の多重比較法

2009SE057 廣田拓也　2009SE204 二階堂健吾

指導教員：白石高章

1　はじめに

分散の異なる 2 群問題の検定として Welch の方法がある.Welch の検定統計量を用いて多群モデルのおける多重比較法を Games-Howell が提案した. 本研究では，Games-Howell 法の分布論について考察する. さらに，Games-Howell 法と漸近理論についてのプログラムを作成した. それを用いてデータ解析を行う. データ解析では，白血球の量が大腸がんに関連があるのか調査したデータを用いる.

2　Welch の検定

2 つの母集団の母平均の差に関する検定を考える. この検定を行うために頻繁に用いられる検定方法が 2 標本 t 検定である. なお，母集団がいずれも正規分布に従いそれぞれ独立であると仮定した上での、平均が等しいかどうかの検定である. その中でも等分散性が仮定できないときに用いるのが Welch の検定である. 以下の内容は，参考文献 [2]，[5] を基に記述した.

2 組のデータを，

$X_{1j} \sim N(\mu_1, \sigma_1^2) \qquad (j = 1, \cdots, n_1),$

$X_{2j} \sim N(\mu_2, \sigma_2^2) \qquad (j = 1, \cdots, n_2).$

それぞれの標本平均，分散の推定量を，

$$\bar{X}_{1\cdot} \equiv \frac{1}{n_1}\sum_{j=1}^{n_1} X_{1j}, \bar{X}_{2\cdot} \equiv \frac{1}{n_2}\sum_{j=1}^{n_2} X_{2j},$$

$$\widehat{\sigma}_1^2 \equiv \frac{\sum_{j=1}^{n_1}(X_{1j} - \bar{X}_{1\cdot})^2}{n_1 - 1}, \widehat{\sigma}_2^2 \equiv \frac{\sum_{j=1}^{n_2}(X_{2j} - \bar{X}_{2\cdot})^2}{n_2 - 1},$$

統計量を，$T \equiv \dfrac{\bar{X}_{1\cdot} - \bar{X}_{2\cdot}}{\sqrt{\dfrac{\widehat{\sigma}_1^2}{n_1} + \dfrac{\widehat{\sigma}_2^2}{n_2}}}$ とする.

$$\begin{cases} \text{帰無仮説 } H_0 : \mu_1 = \mu_2 \\ \text{対立仮説 } H_1 : \mu_1 \neq \mu_2 \end{cases} \tag{1}$$

の検定を考える.

ここで，自由度 ν について考察する.H_0 の下で T は近似的に t_ν に従うと考える. (t_ν は自由度 ν の t 分布)

$E(\bar{X}_{1\cdot} - \bar{X}_{2\cdot}) = 0,\ V(\bar{X}_{1\cdot} - \bar{X}_{2\cdot}) = \dfrac{\sigma_1^2}{n_1} + \dfrac{\sigma_2^2}{n_2}$ より，

$$\frac{\bar{X}_{1\cdot} - \bar{X}_{2\cdot}}{\sqrt{\dfrac{\sigma_1^2}{n_1} + \dfrac{\sigma_2^2}{n_2}}} \sim N(0,\ 1),\ \text{さらに}\frac{N(0,\ 1)}{\sqrt{\frac{\chi_\nu^2}{\nu}}} \sim t_\nu \text{ より，}$$

$$T = \frac{(\bar{X}_{1\cdot} - \bar{X}_{2\cdot}) \Big/ \sqrt{\dfrac{\sigma_1^2}{n_1} + \dfrac{\sigma_2^2}{n_2}}}{\sqrt{\dfrac{\widehat{\sigma}_1^2}{n_1} + \dfrac{\widehat{\sigma}_2^2}{n_2}} \Big/ \sqrt{\dfrac{\sigma_1^2}{n_1} + \dfrac{\sigma_2^2}{n_2}}},$$

近似的に，$\dfrac{\dfrac{\widehat{\sigma}_1^2}{n_1} + \dfrac{\widehat{\sigma}_2^2}{n_2}}{\dfrac{\sigma_1^2}{n_1} + \dfrac{\sigma_2^2}{n_2}} \sim \dfrac{\chi_\nu^2}{\nu}$ となり，

$$Y \equiv \frac{\nu\left(\dfrac{\widehat{\sigma}_1^2}{n_1} + \dfrac{\widehat{\sigma}_2^2}{n_2}\right)}{\dfrac{\sigma_1^2}{n_1} + \dfrac{\sigma_2^2}{n_2}} \text{ が近似的に } \chi_\nu^2 \text{ に従う.}$$

(χ_ν^2 は自由度 ν の χ^2 分布)

$W \sim \chi_\nu^2$ とする.

$E(W) = \nu,\ V(W) = 2\nu$ より，(2) を得る.

$$\frac{V(W)}{\{E(W)\}^2} = \frac{2\nu}{\nu^2} = \frac{2}{\nu}. \tag{2}$$

$$E(Y) = \nu,\ V(Y) = \frac{\nu\left(\dfrac{2\sigma_1^4}{n_1^2(n_1-1)} + \dfrac{2\sigma_2^4}{n_2^2(n_2-1)}\right)}{\left(\dfrac{\sigma_1^2}{n_1} + \dfrac{\sigma_2^2}{n_2}\right)^2}.$$

W を Y に変えて (2) に代入すると，

$$\frac{1}{\nu} = \frac{\dfrac{\sigma_1^4}{n_1^2(n_1-1)} + \dfrac{\sigma_2^4}{n_2^2(n_2-1)}}{\left(\dfrac{\sigma_1^2}{n_1} + \dfrac{\sigma_2^2}{n_2}\right)^2} \text{ を得る.}$$

ここで σ_1^2, σ_2^2 は未知であるので，代わりに推定量 $\widehat{\sigma}_1^2, \widehat{\sigma}_2^2$ 入れると，

$$\nu \equiv \frac{\left(\dfrac{\widehat{\sigma}_1^2}{n_1} + \dfrac{\widehat{\sigma}_2^2}{n_2}\right)^2}{\dfrac{\widehat{\sigma}_1^4}{n_1^2(n_1-1)} + \dfrac{\widehat{\sigma}_2^4}{n_2^2(n_2-1)}}$$

となる. しかし，実際の計算では自由度が整数にならないので，四捨五入して扱う.

(1) に対する水準 α の Welch の検定は，

$|T| > t(\nu; \alpha/2)$ のとき，帰無仮説 H_0 を棄却し，対立仮説 H_1 を受け入れ，$\mu_1 \neq \mu_2$ と判定することである.

ただし，$t(\nu;\alpha/2)$ は自由度 ν の t 分布の上側 $100(\alpha/2)$ % 点とする.

$\mu_1-\mu_2$ についての信頼係数 $1-\alpha$ の同時信頼区間は,

$$\bar{X}_{1\cdot}-\bar{X}_{2\cdot}-t\left(\nu;\frac{\alpha}{2}\right)\cdot\sqrt{\frac{\widehat{\sigma}_1^2}{n_1}+\frac{\widehat{\sigma}_2^2}{n_2}}$$

$$<\mu_1-\mu_2<\bar{X}_{1\cdot}-\bar{X}_{2\cdot}+t\left(\nu;\frac{\alpha}{2}\right)\cdot\sqrt{\frac{\widehat{\sigma}_1^2}{n_1}+\frac{\widehat{\sigma}_2^2}{n_2}}$$

で与えられる.

3 Games-Howell 法

Games-Howell 法は Tukey 型の方法で，Welch の統計量を基礎とした多重比較法である. 以下の内容は，参考文献 [4] を基に記述した.

k 群のデータを次の表 1 のようにまとめた.

表 1 k 群モデル

群	サイズ	データ	平均	分布関数
第 1 群	n_1	$X_{11},\cdots,X_{1n_1}$	μ_1	$N(\mu_1,\sigma_1^2)$
$\vdots$	$\vdots$	$\vdots\ \vdots\ \vdots$	$\vdots$	$\vdots$
第 k 群	n_k	$X_{k1},\cdots,X_{kn_k}$	μ_k	$N(\mu_k,\sigma_k^2)$

分散の推定量を,

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

統計量を，$T_{ii'}\equiv\dfrac{\bar{X}_{i\cdot}-\bar{X}_{i'\cdot}}{\sqrt{\dfrac{\widehat{\sigma}_i^2}{n_i}+\dfrac{\widehat{\sigma}_{i'}^2}{n_{i'}}}}$,

自由度を，$\nu_{ii'}=\left[\dfrac{\left(\dfrac{\widehat{\sigma}_i^2}{n_i}+\dfrac{\widehat{\sigma}_{i'}^2}{n_{i'}}\right)^2}{\dfrac{\widehat{\sigma}_i^4}{n_i^2(n_i-1)}+\dfrac{\widehat{\sigma}_{i'}^4}{n_{i'}^2(n_{i'}-1)}}+0.5\right]$

とする. [ ] はガウス記号.

$$\begin{cases}\text{帰無仮説 } H_{(i,\ i')}:\mu_i=\mu_{i'}\\ \text{対立仮説 } H_{(i,\ i')}^A:\mu_i\neq\mu_{i'}\end{cases}$$

に対する水準 α の多重比較検定は,

1. $|T_{ii'}|>ta(k,\nu_{ii'};\alpha)$ 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.

2. $|T_{ii'}|<ta(k,\nu_{ii'};\alpha)$ となる $i,\ i'$ に対して帰無仮説 $H_{(i,\ i')}$ を棄却しない.

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

とする. ただし,

$$\Phi(x)=\int_{-\infty}^x\varphi(z)dz,\varphi(x)=\frac{1}{\sqrt{2\pi}}e^{-\frac{x^2}{2}},$$

$$g(s)\equiv\frac{m^{m/2}}{\Gamma(m/2)2^{m/2-1}}s^{m-1}e^{-ms^2/2},\ m\equiv n-k,$$

$$\Gamma(a)\equiv\int_0^\infty x^{a-1}e^{-x}dx$$

である. $TA(t)=1-\alpha$ を満たす t の解を $ta(k,\ m;\alpha)$ とする.

$\mu_i-\mu_{i'}$ $(1\le i<i'\le k)$ についての信頼係数 $1-\alpha$ の同時信頼区間は,

$$\bar{X}_{i\cdot}-\bar{X}_{i'\cdot}-ta(k,\nu_{ii'};\alpha)\cdot\sqrt{\frac{\widehat{\sigma}_i^2}{n_i}+\frac{\widehat{\sigma}_{i'}^2}{n_{i'}}}$$

$$<\mu_i-\mu_{i'}<\bar{X}_{i\cdot}-\bar{X}_{i'\cdot}+ta(k,\nu_{ii'};\alpha)\cdot\sqrt{\frac{\widehat{\sigma}_i^2}{n_i}+\frac{\widehat{\sigma}_{i'}^2}{n_{i'}}}$$

$(1\le i<i'\le k)$ で与えられる.

4 漸近理論

統計量 $T_{ii'}$ について次の補題 1 が得られる.

補題 1 $\{X_{ij}\mid\ j=1,\cdots,n_i,\ i=1,\cdots,k\}$ は互いに独立であると仮定し,

$$X_{ij}\sim N(\mu_i,\sigma_i^2),\ \lim_{n\to\infty}\frac{n_i}{n}=\lambda_i>0\quad(i=1,\cdots,k)$$

$$T_{ii'}(\boldsymbol{\mu})\equiv\frac{(\bar{X}_{i\cdot}-\mu_i)-(\bar{X}_{i'\cdot}-\mu_{i'})}{\sqrt{\dfrac{\widehat{\sigma}_i^2}{n_i}+\dfrac{\widehat{\sigma}_{i'}^2}{n_{i'}}}}\ \text{とおく.}$$

ただし，$\widehat{\sigma}_i^2\equiv\dfrac{1}{n_i-1}\displaystyle\sum_{j=1}^{n_i}(X_{ij}-\bar{X}_{i\cdot})^2$ とする.

このとき，$T_{ii'}(\boldsymbol{\mu})\xrightarrow{\mathcal{L}}N(0,\ 1)$,

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

が成り立つ.

ただし，$Y_i\sim N(0,\sigma_i^2/\lambda_i)$ とする.

証明 $T_{ii'}(\boldsymbol{\mu})=\dfrac{\sqrt{n}(\bar{X}_{i\cdot}-\mu_i)-\sqrt{n}(\bar{X}_{i'\cdot}-\mu_{i'})}{\sqrt{\dfrac{n}{n_i}\widehat{\sigma}_i^2+\dfrac{n}{n_{i'}}\widehat{\sigma}_{i'}^2}}$

と式を変形できる. 中心極限定理より,

$\sqrt{n_i}(\bar{X}_{i\cdot}-\mu_i)\xrightarrow{\mathcal{L}}\sigma_iZ_i\sim N(0,\sigma_i^2)$ が成り立つ. 故に,

$$\sqrt{n}(\bar{X}_{i\cdot}-\mu_i)=\sqrt{n_i}\frac{\sqrt{n}}{\sqrt{n_i}}(\bar{X}_{i\cdot}-\mu_i)$$
$$\xrightarrow{\mathcal{L}}\frac{\sigma_i}{\sqrt{\lambda_i}}Z_i\sim N\left(0,\frac{\sigma_i^2}{\lambda_i}\right)$$
が言える．このとき，$Z_i\sim N(0,\,1)$，$Z_1,\cdots,\,Z_k$ は互いに独立である．大数の法則より，
$$\widehat{\sigma}_i^2=\frac{n_i}{n_i-1}\left\{\frac{1}{n_i}\sum_{j=1}^{n_i}X_{ij}^2-(\bar{X}_{i\cdot})^2\right\}$$
$$\xrightarrow{P}E(X_{i1}^2)-\{E(X_{i1})\}^2=\sigma_i^2$$
が成り立つ．よって，スラツキーの定理より，
$$T_{ii'}(\boldsymbol{\mu})\xrightarrow{\mathcal{L}}S_{ii'}\equiv\frac{\frac{\sigma_i}{\sqrt{\lambda_i}}Z_i-\frac{\sigma_{i'}}{\sqrt{\lambda_{i'}}}Z_{i'}}{\sqrt{\frac{\sigma_i^2}{\lambda_i}+\frac{\sigma_{i'}^2}{\lambda_{i'}}}}$$
である．このとき $S_{ii'}$ の平均，分散を求めると，$E(S_{ii'})=0$，$V(S_{ii'})=1$ である．
よって，$T_{ii'}(\boldsymbol{\mu})\xrightarrow{\mathcal{L}}N(0,\,1)$ となる．また，
$$\max_{1\le i<i'\le k}|T_{ii'}(\boldsymbol{\mu})|\xrightarrow{\mathcal{L}}\max_{1\le i<i'\le k}\frac{\left|\frac{\sigma_i}{\sqrt{\lambda_i}}Z_i-\frac{\sigma_{i'}}{\sqrt{\lambda_{i'}}}Z_{i'}\right|}{\sqrt{\frac{\sigma_i^2}{\lambda_i}+\frac{\sigma_{i'}^2}{\lambda_{i'}}}}$$
$$\Leftrightarrow\max_{1\le i<i'\le k}|T_{ii'}(\boldsymbol{\mu})|\xrightarrow{\mathcal{L}}\max_{1\le i<i'\le k}\frac{|Y_i-Y_{i'}|}{\sqrt{\frac{\sigma_i^2}{\lambda_i}+\frac{\sigma_{i'}^2}{\lambda_{i'}}}}$$
が成り立つ．ただし，$Y_i\equiv\sigma_iZ_i/\sqrt{\lambda_i}\sim N(0,\sigma_i^2/\lambda_i)$ とする． □

補題 1 の条件 $X_{ij}\sim N(\mu_i,\sigma_i^2)$ の替わりに，次の一般的な分布の

$$(\textbf{条件 1})\quad\begin{cases}P(X_{ij}\le x)=F\left(\dfrac{x-\mu_i}{\sigma_i}\right),\\ f(x)\equiv F'(x),\ \displaystyle\int_{-\infty}^{\infty}xf(x)dx=0,\\ \displaystyle\int_{-\infty}^{\infty}x^2f(x)dx=1.\end{cases}$$

に変えても補題 1 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 1 は成り立つ．

補題 1 から次の命題 2 を得る．

命題 2　補題 1 の条件の下で，
$$A(t)\le\lim_{n\to\infty}P\left(\max_{1\le i<i'\le k}|T_{ii'}(\boldsymbol{\mu})|\le t\right)\le A^*(t)$$
が成り立つ．ただし，
$$A(t)\equiv k\int_{-\infty}^{\infty}\{\Phi(x)-\Phi(x-\sqrt{2}\cdot t)\}^{k-1}d\Phi(x),$$
$$A^*(t)\equiv\sum_{i'=1}^{k}\int_{-\infty}^{\infty}\prod_{\substack{i=1\\ i\neq i'}}^{k}\left\{\Phi\left(\sqrt{\frac{\lambda_i\sigma_{i'}^2}{\lambda_{i'}\sigma_i^2}}\cdot x\right)\right.$$
$$\left.-\Phi\left(\sqrt{\frac{\lambda_i\sigma_{i'}^2}{\lambda_{i'}\sigma_i^2}}\cdot x-\sqrt{\frac{\lambda_i\sigma_{i'}^2+\lambda_{i'}\sigma_i^2}{\lambda_{i'}\sigma_i^2}}\cdot t\right)\right\}d\Phi(x)$$
とする．$\sigma_1^2/\lambda_1=\cdots=\sigma_k^2/\lambda_k$ が満たされるとき，上の不等式で等号が成り立つ．

証明　任意の $t>0$ に対して，
$$\lim_{n\to\infty}P\left(\max_{1\le i<i'\le k}|T_{ii'}(\boldsymbol{\mu})|\le t\right)$$
$$=P\left(\max_{1\le i<i'\le k}\frac{|Y_i-Y_{i'}|}{\sqrt{\frac{\sigma_i^2}{\lambda_i}+\frac{\sigma_{i'}^2}{\lambda_{i'}}}}\le t\right)$$
が補題 1 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，参考文献 [1] の付録の定理 A.5, A.6 を適用すると，
$$A(t)\le\lim_{n\to\infty}P\left(\max_{1\le i<i'\le k}|T_{ii'}(\boldsymbol{\mu})|\le t\right)\le A^*(t)$$
が導かれる． □

(条件 1) のような一般的な条件に変えても命題 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命題 2 は成り立つ．

$n_1,\cdots,\,n_k$ が大きいとき命題 2 より，
$\{$ 帰無仮説 $H_{(i,\ i')}$ vs. 対立仮説 $H_{(i,\ i')}^A|1\le i<i'\le k\}$ に対する水準 α の多重比較検定は，

1. $|T_{ii'}|>a(k;\alpha)$ 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．
2. $|T_{ii'}|<a(k;\alpha)$ となる $i,\,i'$ に対して帰無仮説 $H_{(i,\ i')}$ を棄却しない．

$A(t)=1-\alpha$ を満たす t の解を $a(k;\alpha)$ とする．

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

$$\bar{X}_{i\cdot}-\bar{X}_{i'\cdot}-a(k;\alpha)\cdot\sqrt{\frac{\widehat{\sigma}_i^2}{n_i}+\frac{\widehat{\sigma}_{i'}^2}{n_{i'}}}$$
$$<\mu_i-\mu_{i'}<\bar{X}_{i\cdot}-\bar{X}_{i'\cdot}+a(k;\alpha)\cdot\sqrt{\frac{\widehat{\sigma}_i^2}{n_i}+\frac{\widehat{\sigma}_{i'}^2}{n_{i'}}}$$
$(1\le i<i'\le k)$ で与えられる．

5　Games-Howell 法の解析プログラム

Games-Howell 法で解析を行うプログラムを作成した．自由度が 200 以上になると漸近理論を用いて応用解析す

る．$\alpha = 0.05,\ 0.01,\ 3 \leq k \leq 10$ に対する $ta(k,\ m;\alpha)$ の数表が参考文献 [6] の表 6.1，表 6.2 に，$a(k;\alpha)$ の数表が参考文献 [1] の付表 12 に掲載されている. これらの数値を活用した．ページ数の都合上，C 言語によるデータ解析プログラムは，main プログラムだけを以下に示し，その説明を行う. なお，全体のプログラムは卒業論文に掲載されている．

```
int main(void){
    taa();
    input();
    keisan();
    taakakuno();
    sinrai();
    kentei();
    return 0;
}
```

taa(); は $ta(k,\ m;\alpha)$，$a(k;\alpha)$ の値をファイルから読み込む.

input(); は検定したい群の数，それぞれの標本サイズ，有意水準，観測値を入力する．

keisan(); は標準平均，分散の推定量などを求めそれらを用いて統計量，自由度を計算する．

taakakuno(); は入力された群の数，有意水準，計算によって求められた自由度から $ta(k,\ m;\alpha)$，$a(k;\alpha)$ を求める．

sinrai(); は同時信頼区間を求める．

kentei(); 自由度の大きさによって Games-Howell 法の検定か漸近理論の検定のどちらかの結果を表示する．

6　データ解析

がん患者の白血球量に違いがあるのか調べるためにそれぞれの群について Games-Howell 法で実際に解析してみた. なお，このデータは参考文献 [3] に掲載されている．

表 2　$\beta-$ 白血球の量 $(*10^9/L)$

第 1 群 健常者	6.0	6.3	5.1	6.2	10.4	4.4	7.4	7.0
	5.6	5.3	2.6	6.3	6.1	5.3	5.4	5.2
	4.3	4.9	7.3	4.9	6.9	4.3	5.6	5.1
第 2 群 A	7.7	7.8	6.1	9.6	5.5	5.8	4.0	5.4
第 3 群 B	10.4	5.6	7.0	8.2	9.0	8.4	8.1	8.0
	6.5	9.1	11.0	10.9	10.6	5.2	7.9	7.6
	5.8	7.0						
第 4 群 C	8.0	6.7	9.3	6.6	9.3	7.2	5.2	9.8
	6.2	10.1	9.3	9.4	6.5	5.4	7.6	9.2
第 5 群 D	9.5	7.8	5.7	8.0	9.6	13.7	6.3	7.3
	6.2							

表 2 のデータは，大腸癌患者の β -白血球について検討した. 数値は β -白血球 $(*10^9/L)$ を表している．病状 A, B，C，D を以下のように分類する．

A (95 %)：がんが大腸壁内にとどまるもの
B (80 %)：がんが大腸壁を貫くがリンパ節転移のないもの
C (70 %)：リンパ節転移のあるもの
D (25 %)：腹膜、肝、肺などへの遠隔転移のあるもの

検定結果は以下のようになった.

$|T_{12}| = 1.070 < ta(5,\ 10; 0.01) = 4.339$ より $H_{(1,\ 2)}$を棄却しない. $\mu_1 - \mu_2$の信頼係数 0.99 の同時信頼区間は, $-3.749 < \mu_1 - \mu_2 < 2.265.$

$|T_{13}| = 4.588 > ta(5,\ 32; 0.01) = 3.548$ より $H_{(1,\ 3)}$を棄却する. $-4.224 < \mu_1 - \mu_3 < -0.540.$

$|T_{14}| = 4.165 > ta(5,\ 30; 0.01) = 3.569$ より $H_{(1,\ 4)}$を棄却する. $-3.930 < \mu_1 - \mu_4 < -0.303.$

$|T_{15}| = 2.840 < ta(5,\ 10; 0.01) = 4.339$ より $H_{(1,\ 5)}$を棄却しない. $-6.288 < \mu_1 - \mu_5 < 1.313.$

残りの帰無仮説に対しても同様に検定をした結果，棄却されなかった.

健康な人と比べて病状 B，C の人は白血球が多いことが分かった．信頼区間に 0 を含まないことと帰無仮説を棄却することが一致していることを確認できた．また，ボンフェローニ法やホルムの方法を用いても同様の結果が得られた．作成したプログラムで解析を行ったところ，同様の結果を得られた.

7　おわりに

Welch の検定や Games-Howell 法について述べてきた. 多重比較法について研究してきたが，Games-Howell 法のプログラムを作ることで，今まで研究してきたことの理解がより深まった．2 群問題の Welch の検定を基礎として多重比較法の Games-Howell 法を研究していくことで，2 群問題では考えなかった条件などを論じた.

参考文献

[1] 白石高章：『多群連続モデルにおける多重比較法』共立出版，東京，2011.

[2] 矢島美寛　等：『自然科学の統計学』東京大学出版会，東京，1992.

[3] Daniel，W.W.：Biostatistics：basic concepts and methodology for the health sciences，Wiley，2010.

[4] Games，P.A. and Howell，J.F.： Pairwise Multiple Comparison Procedures with Unequal N's and/or Variances：A Monte Carlo Study，Journal of Educational and Behavioral Statistics，1976.

[5] Welch，B.L.：The significance of the difference between two means when the population variances are unequal．Biometrika，25，350-362，1938.

[6] 早川由宏: 『Mathematica と C 言語による統計プログラミングの基礎』. 2012 年度南山大学情報理工学部情報システム数理学科卒業論文，2013 年 1 月.