

# 統計・多変量解析と ソフトコンピューティング

## 第8章 多重比較法—閾値の調整—

### 本稿掲載の Web ページ

[http://mybook-pub-site.sakura.ne.jp/Statistics\\_Multivariate/index.html](http://mybook-pub-site.sakura.ne.jp/Statistics_Multivariate/index.html)

古橋 武

## 第8章

### 多重比較法—閾値の調整—

前章のように互いに独立な事象の検定を繰り返す場合には、シダックの方法により名義水準を設定すれば良い。しかし、本章で取り扱う各群間の平均値の差の検定においては、この平均値の差は互いに独立ではない。このような場合、シダックの方法では第1種の過誤の確率を公称の有意水準より低くしてしまう。本章では第1種の過誤の確率を公称の有意水準と同じにするよう検定統計量の閾値を調整する多重比較法を紹介する。

#### 8.1 母平均の差の検定（データ数が等しい場合）

##### 8.1.1 テューキーの方法による検定の例

ある日、ある工場のある製品のラインにおいて、ラインのスピードを3通りに変えながらあるパラメータの測定がなされたとする。スピード設定1, 2, 3からそれぞれ9個の標本を取り出すことができ、以下の値が得られたとする。

設定1: 3.3, 3.2, 3.4, 3.3, 3.3, 3.2, 3.5, 3.2, 3.5

設定2: 3.2, 3.1, 3.3, 3.2, 3.2, 3.1, 3.4, 3.2, 3.0

設定3: 3.1, 3.0, 3.3, 3.1, 3.2, 3.3, 3.0, 3.1, 3.1

ライン1, 2, 3の平均値  $\bar{x}_1, \bar{x}_2, \bar{x}_3$  はそれぞれ3.32, 3.19, 3.13であった。スピード設定を変えたことで母平均に差が出たと言えるか？ただし、いずれの設定の母分散も分かっていないが、等しいとみなせるとする。

検定の仮説は、設定 1, 2, 3 の母平均を  $\mu_1, \mu_2, \mu_3$  とすると

$$\begin{aligned} \text{帰無仮説} &: \mu_i = \mu_j \\ \text{対立仮説} &: \mu_i \neq \mu_j \\ &\text{ただし, } i, j = 1, 2, 3, i \neq j \end{aligned} \quad (8.1)$$

であり、公称の有意水準  $\alpha = 0.05$  とする。各データ群は互いに独立であるが、群間の平均値の差同士は独立ではない。このため次項に示すように多重比較による第 1 種の過誤の確率は約 12% であり、前章のボンフェローニの方法やシダックの方法では名義水準を厳しくしすぎてしまう。群の全ての対について平均値の差を検定する多重比較法にテューキー (Tukey) の方法がある。設定  $i$  と設定  $j$  の検定統計量  $q_{ij}$  は次式により与えられる。

$$q_{ij} = \frac{|\bar{x}_i - \bar{x}_j|}{\sqrt{v_E^2 \left( \frac{1}{n_i} + \frac{1}{n_j} \right)}} \quad (8.2)$$

ただし、

$$v_E^2 = \frac{\sum_{k=1}^a \sum_{l=1}^{n_k} (x_{kl} - \bar{x}_k)^2}{\sum_{k=1}^a (n_k - 1)} \quad (8.3)$$

である。これは平均平方と呼ばれる。  $x_{kl}$  は設定  $k$  の第  $l$  番目の測定値、  $n_k (k = 1, 2, 3)$  は各設定のデータ数、  $a$  は設定の数である。この  $q_{ij}$  値は自由度  $\nu$  が

$$\nu = \sum_{k=1}^a (n_k - 1) \quad (8.4)$$

である  $t$  分布に従う。テューキーの方法は各群のデータ数が同じであることを前提としている。そこで、各群の不偏分散を  $v_{ek}^2 (k = 1, 2, \dots, a)$  とすると、上式は

$$q_{ij} = \frac{|\bar{x}_i - \bar{x}_j|}{\sqrt{\frac{2v_E^2}{n}}} \quad (8.5)$$

$$v_E^2 = \frac{\sum_{k=1}^a v_{ek}^2}{a} \quad (8.6)$$

となる。式 (8.5) の  $q_{ij}$  値は自由度  $\nu$  が

$$\nu = a(n - 1) \quad (8.7)$$

である  $t$  分布に従う。テューキーの方法は第 1 種の過誤の確率が公称の有意水準  $\alpha$  となるように、 $q$  値の分布を考慮して検定統計量  $q_{ij}$  の閾値  $q_0$  を調整する検定法である。

|    | A                  | B                   | C        | D         | E            |
|----|--------------------|---------------------|----------|-----------|--------------|
| 1  | 母平均の差の検定(テューキーの方法) |                     |          |           |              |
| 2  |                    |                     |          |           |              |
| 3  | 入力                 | ステューデント化された範囲の上側5%点 |          |           |              |
| 4  |                    | 3.532               |          |           |              |
| 5  |                    |                     |          |           |              |
| 6  |                    | 設定1                 | 設定2      | 設定3       |              |
| 7  |                    | 3.3                 | 3.2      | 3.1       |              |
| 8  |                    | 3.2                 | 3.1      | 3.0       |              |
| 9  |                    | 3.4                 | 3.3      | 3.3       |              |
| 10 |                    | 3.3                 | 3.2      | 3.1       |              |
| 11 |                    | 3.3                 | 3.2      | 3.2       |              |
| 12 |                    | 3.2                 | 3.1      | 3.3       |              |
| 13 |                    | 3.5                 | 3.4      | 3.0       |              |
| 14 |                    | 3.2                 | 3.2      | 3.1       |              |
| 15 |                    | 3.5                 | 3.0      | 3.1       |              |
| 16 |                    |                     |          |           |              |
| 17 | 計算値                | 設定数 a               | データ数 n   | 自由度 $\nu$ |              |
| 18 |                    | 3                   | 9        | 24        |              |
| 19 |                    | 平均x1                | x2       | x3        |              |
| 20 |                    | 3.32                | 3.19     | 3.13      |              |
| 21 |                    | 不偏分散 $v_1^2$        | $v_2^2$  | $v_3^2$   | 平均平方 $v_E^2$ |
| 22 |                    | 0.0144              | 0.0136   | 0.0125    | 0.0135       |
| 23 |                    | 閾値 $q_0$            |          |           |              |
| 24 |                    | 2.498               |          |           |              |
| 25 | 出力                 | 検定統計量 $q_{12}$      | $q_{13}$ | $q_{23}$  |              |
| 26 |                    | 2.43                | 3.45     | 1.01      |              |

図 8.1: 母平均の差の検定の例(テューキーの方法) (テューキーによる検定.xlsx)

図 8.1 はテューキーの方法による母平均の差の検定の実行画面である。セル E22 に式 (8.6) の平均平方  $v_E^2$  を求め、セル B26, C26, D26 にそれぞれ設定 1, 2 間, 1, 3 間, 2, 3 間の平均値の差の  $q$  値  $q_{12}, q_{13}, q_{23}$  を求めている。セル B24 は検定統計量  $q$  値の調整された閾値  $q_0$  である。この値は第 1 種の過誤が公称の有意水準  $\alpha = 0.05$  となるように設定されている。これはステューデント化された範囲の 5% 点と呼ばれる。数表を本書の末尾の表 A.1 に付してある。データ群の数  $a = 3$ , 自由度  $\nu = 24$  であるので、この値は数表から 3.532 と読める。セル B4 にこの値が入力され、セル B24 でこの値を  $\sqrt{2}$  で割ることで閾値  $q_0$  を得ている。閾値を上回っている設定 1, 3 間の平均値に有意な差があったと結論づけられる。

ところで、何故わざわざ数表の値を  $\sqrt{2}$  で割っているのでしょうか？ステューデント化された範囲は

$$q_{ij} = \frac{|\bar{x}_i - \bar{x}_j|}{\sqrt{\frac{v_E^2}{n}}} \quad (8.8)$$

を検定統計量として求められたものである。式 (8.5) の分母から  $\sqrt{2}$  が取り除かれた式である。ステューデント化された範囲の数表をテューキーの方法にのみ適用する場合には式 (8.8) の修正された検定統計量が望ましい。数表の値を  $\sqrt{2}$  で割る必要がない。しかし、これを認めると、後述のテューキー・クレーマーの方法 (8.2 節), ゲームズ・ハウウェ

ルの方法 (8.3 節) など、テューキーの方法の改良手法に数表を適用する場合にも、事前に検定統計量を  $\sqrt{2}$  倍しておかないといけない。後の改良手法においては不自然に見えてしまう  $\sqrt{2}$  の項を必要とする。そこで、本書ではテューキーの方法の検定統計量に式 (8.5) の本来の式を用い、数表の値を  $\sqrt{2}$  で割って用いることとする。こうすれば、どの方法を適用する場合にも、ステューデント化された範囲の数表を常に  $\sqrt{2}$  で割って用いるということで一貫する。もちろん、ステューデント化された範囲の数表を式 (8.5) の検定統計量をもとに新たに作り直す誘惑に駆られるが、歴史ある数表に異なるバージョンを付け加えることはいたずらに混乱を招くだけと思われる。

### 8.1.2 多重性の問題とテューキーの方法のシミュレーション

シミュレーションにより母平均の差の検定における多重性の問題とテューキーの方法の効果を確認する。図 8.2 は、 $q$  値の閾値の調整をしていない場合である。セル B10~B18, C10~C18, D10~D18 にそれぞれ平均  $\mu_1, \mu_2, \mu_3$ , 分散  $\sigma_1^2, \sigma_2^2, \sigma_3^2$  の正規分布に従う乱数 9 個ずつ 3 群を生成している。セル B27, C27, D27 ではそれぞれ 1, 2 群間, 1, 3 群間, 2, 3 群間の平均値の差の  $q$  値  $q_{12}, q_{13}, q_{23}$  を式 (8.5) により求めている。そして、セル B29, C29, D29 にて  $q_{ij}$  値が自由度  $a(n-1)$  の  $t$  分布に従うとして  $p_{ij}$  値を求めている。31 行目では名義水準  $\alpha' = 0.05$  のもとで有意差判定を行い、セル B33 で各対のいずれかで有意差ありと判定された場合に 1 を出力している。以上の検定を 1000 組について行った結果、各対のいずれかで有意差ありと判定された頻度をセル B35 に出力している。図 8.3 は以上の 1000 組のシミュレーションを 100 回繰り返して、第 1 種の過誤の確率の 95% 信頼区間を求めた画面である。各群間の平均値の差は互いに独立ではないので、第 1 種の過誤の確率  $p^*$  はおよそ 12% となっている。一方、図 7.1 の互いに独立な事象の検定の繰り返しにおける第 1 種の過誤の確率は約 14% であった。シダックの方法はこの確率を 5% に抑えるよう名義水準を設定していた。本節のような互いに独立でない事象の検定の繰り返しにシダックの方法を適用すれば、第 1 種の過誤の確率を 5% より小さくしてしまう。

図 8.4 は、 $q$  値の閾値を調整した場合である。図 8.2 との違いは、セル C6 にて調整された  $q$  値である  $q_0$  値が数表から手作業で入力され、29 行目のセルで 27 行目の  $q_{ij}$  値の絶対値が  $q_0$  値を超えた場合を有意差ありとして 1 を出力している点である。図 8.5 は、1000 組のシミュレーションを 100 回繰り返して得られた 95% 信頼区間を示す。第 1 種の過誤の確率  $p^*$  が公称の有意水準に抑えられていることが分かる。

図 8.6 は  $q$  値と 3 群の中で最も絶対値の大きかった  $q$  値 ( $q_{max}$  値とする) の頻度を求めているシミュレーション画面である。図 8.2, 図 8.4 と同様にデータ群数  $a = 3$ , 各群のデータ数  $n = 9$  としている。検定統計量の自由度  $\nu = 24$  である。異なるのは、27 行目

|    | A                                | B                                   | C                  | D                  | E                  | F                  | G                  |
|----|----------------------------------|-------------------------------------|--------------------|--------------------|--------------------|--------------------|--------------------|
| 1  | 母平均の差の検定のシミュレーション(分散未知, q値の調整なし) |                                     |                    |                    |                    |                    |                    |
| 2  |                                  |                                     |                    |                    |                    |                    |                    |
| 3  | 入力                               | $\mu 1$                             | $\mu 2$            | $\mu 3$            | $\sigma 1$         | $\sigma 2$         | $\sigma 3$         |
| 4  |                                  | 1                                   | 1                  | 1                  | 1                  | 1                  | 1                  |
| 5  |                                  | 公称の有意水準 $\alpha$                    | 名義水準 $\alpha'$     |                    |                    |                    |                    |
| 6  |                                  | 0.05                                | 0.05               |                    |                    |                    |                    |
| 7  |                                  |                                     |                    |                    |                    |                    |                    |
| 8  |                                  | 1組目                                 |                    |                    | 2組目                |                    |                    |
| 9  | 計算値                              | x1                                  | x2                 | x3                 | x1                 | x2                 | x3                 |
| 10 |                                  | 1.982                               | 0.431              | 2.422              | 1.267              | 1.911              | 1.314              |
| 11 |                                  | 0.129                               | 0.127              | -0.374             | 2.195              | 1.286              | 0.303              |
| 12 |                                  | -0.132                              | 0.715              | 0.338              | 0.967              | -0.082             | 0.571              |
| 20 |                                  | データ群数 a                             | データ数 n             | 組数 G               |                    |                    |                    |
| 21 |                                  | 3                                   | 9                  | 1000               |                    |                    |                    |
| 26 |                                  | 検定統計量 q12                           | q13                | q23                | q12                | q13                | q23                |
| 27 |                                  | 0.229                               | -0.091             | 0.320              | -0.295             | -0.016             | -0.279             |
| 28 |                                  | p12                                 | p13                | p23                | p12                | p13                | p23                |
| 29 |                                  | 0.821                               | 0.928              | 0.752              | 0.770              | 0.987              | 0.783              |
| 30 |                                  | if p12 < $\alpha'$<br>then 1 else 0 | if p13 < $\alpha'$ | if p23 < $\alpha'$ | if p12 < $\alpha'$ | if p13 < $\alpha'$ | if p23 < $\alpha'$ |
| 31 |                                  | 0                                   | 0                  | 0                  | 0                  | 0                  | 0                  |
| 32 |                                  | 各対のいずれかで有意差あり                       |                    |                    | 各対のいずれかで有意差あり      |                    |                    |
| 33 |                                  | 0                                   |                    |                    | 0                  |                    |                    |
| 34 | 出力                               | 有意頻度 fr                             |                    |                    |                    |                    |                    |
| 35 |                                  | 0.122                               |                    |                    |                    |                    |                    |

|    |    |                                     |                    |                    |                    |                    |                    |
|----|----|-------------------------------------|--------------------|--------------------|--------------------|--------------------|--------------------|
| 26 |    | 検定統計量 q12                           | q13                | q23                | q12                | q13                | q23                |
| 27 |    | 0.229                               | -0.091             | 0.320              | -0.295             | -0.016             | -0.279             |
| 28 |    | p12                                 | p13                | p23                | p12                | p13                | p23                |
| 29 |    | 0.821                               | 0.928              | 0.752              | 0.770              | 0.987              | 0.783              |
| 30 |    | if p12 < $\alpha'$<br>then 1 else 0 | if p13 < $\alpha'$ | if p23 < $\alpha'$ | if p12 < $\alpha'$ | if p13 < $\alpha'$ | if p23 < $\alpha'$ |
| 31 |    | 0                                   | 0                  | 0                  | 0                  | 0                  | 0                  |
| 32 |    | 各対のいずれかで有意差あり                       |                    |                    | 各対のいずれかで有意差あり      |                    |                    |
| 33 |    | 0                                   |                    |                    | 0                  |                    |                    |
| 34 | 出力 | 有意頻度 fr                             |                    |                    |                    |                    |                    |
| 35 |    | 0.122                               |                    |                    |                    |                    |                    |

図 8.2: 母平均の差の検定における第 1 種の過誤のシミュレーション (q 値の調整なし) (q 値調整なしのシミュレーション.xlsm)

|    |     |          |         |                 |        |
|----|-----|----------|---------|-----------------|--------|
| 53 |     | 有意頻度 frk | 100回の平均 |                 |        |
| 54 | fr1 | 0.119    | 0.1188  |                 |        |
| 55 | fr2 | 0.119    | 95%信頼区間 |                 |        |
| 56 | fr3 | 0.107    | 0.1170  | $\leq p^* \leq$ | 0.1207 |

図 8.3: 母平均の差の検定における第 1 種の過誤の確率の 95% 信頼区間 (q 値の調整なし)

で各群の  $q$  値の小数点以下を四捨五入し, 29 行目で 3 群の中の  $q_{max}$  値を求め, 31 行目で  $q_{max}$  値の小数点以下を四捨五入している点である. 図 8.7 は得られた頻度分布を示す. 横軸は  $q$  値であり, 縦軸は頻度および確率である.  $q$  値の頻度は 1000 組のシミュレーションにおいて各群の  $q_{ij}$  値が整数値  $\pm 0.5$  の範囲内の値をとった回数 (最大 3000 回) をデータ群数  $\times$  組数 ( $3 \times 1000 = 3000$ ) で割って求めている. 隣の白抜きの棒グラフは  $t_{24}$  の確率であり, 自由度 24 の  $t$  分布の分布関数 (TDIST 関数) を利用して, 整数値  $\pm 0.5$  の範囲の確率を求めたものである. 一方,  $q_{max}$  値の頻度は各組の  $\max|q_{ij}|$  値が整数値  $\pm 0.5$  の範囲内の値をとった回数 (最大 1000 回) を組数 (1000) で割って求めている.  $q_{max}$  の確率は 8.1.3 節の式 (8.41) を用いて求めた理論値である. Excel シートの再計算を実行してシミュレーションを繰り返すことで, 頻度分布が理論値の周りにばらつくことを確認できる. 図 8.8 は TDIST() 関数を用いて描いた自由度 24 の  $t$  分布と式 (8.41) を用いて描いた

|    | A  | B                              | C             | D             | E                              | F             | G             |
|----|--|--------------------------------|---------------|---------------|--------------------------------|---------------|---------------|
| 1  | 母平均の差の検定のシミュレーション(分散未知, <b>α値の調整あり(テューキーの方法)</b> ) |                                |               |               |                                |               |               |
| 2  |  |                                |               |               |                                |               |               |
| 3  | 入力   | μ1                             | μ2            | μ3            | σ1                             | σ2            | σ3            |
| 4  |  | 1                              | 1             | 1             | 1                              | 1             | 1             |
| 5  |  | 公称の有意水準 α                      | 調整 q0         |               |                                |               |               |
| 6  |  | 0.05                           | 2.4975        | = 3.532/√2    |                                |               |               |
| 7  |  |                                |               |               |                                |               |               |
| 8  |  | 1組目                            |               |               | 2組目                            |               |               |
| 9  | 計算値  | x1                             | x2            | x3            | x1                             | x2            | x3            |
| 10 |  | -0.628                         | 1.715         | 3.520         | -2.002                         | -1.563        | 1.284         |
| 11 |  | 0.368                          | -0.173        | 2.816         | 2.285                          | 3.262         | 1.439         |
| 26 |  | 検定統計量 q12                      | q13           | q23           | q12                            | q13           | q23           |
| 27 |  | 0.279                          | -1.958        | 2.237         | 0.331                          | 1.383         | -1.053        |
| 28 |  | if  q12  > q0<br>then 1 else 0 | if  q13  > q0 | if  q23  > q0 | if  q12  > q0<br>then 1 else 0 | if  q13  > q0 | if  q23  > q0 |
| 29 |  | 0                              | 0             | 0             | 0                              | 0             | 0             |
| 30 |  | 各対のいずれかで有意差あり                  |               |               | 各対のいずれかで有意差あり                  |               |               |
| 31 |  | 0                              |               |               | 0                              |               |               |
| 32 | 出力   | 有意頻度 fr                        |               |               |                                |               |               |
| 33 |  | 0.061                          |               |               |                                |               |               |

図 8.4: 母平均の差の検定における第1種の過誤のシミュレーション (テューキーの方法)  
(テューキーによるシミュレーション.xlsx)

|    |     |          |         |        |        |
|----|-----|----------|---------|--------|--------|
| 53 |     | 有意頻度 frk | 100回の平均 |        |        |
| 54 | fr1 | 0.052    | 0.0497  |        |        |
| 55 | fr2 | 0.062    | 95%信頼区間 |        |        |
| 56 | fr3 | 0.057    | 0.0483  | < p* ≤ | 0.0511 |

図 8.5: 母平均の差の検定における第1種の過誤の確率の95%信頼区間 (テューキーの方法)

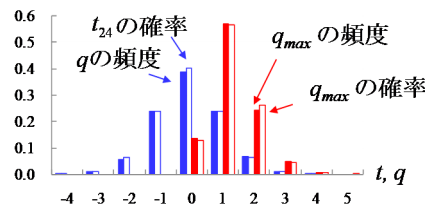
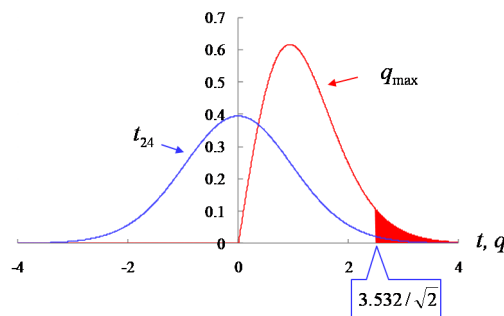
$q_{max}$  値の確率密度関数の形状である。  $q_{max}$  値の分布は  $t$  分布よりも  $t$  値の大きな値にまで延びており、  $q_{max}$  値が  $3.532/\sqrt{2}$  より大きくなる確率が5%である。ここで  $3.532/\sqrt{2}$  は5%点の値と呼ばれる。

### 8.1.3 テューキーの方法の理論

#### 母分散既知の場合

8.1.2 項のシミュレーションにより、各群間の母平均の差の検定を繰り返すことで第1種の過誤の確率が公称の有意水準  $\alpha$  より大きくなってしまいう例を見た。本項ではこのことを理論的に考察する。まず、簡単のために母分散を既知で等しいとする。事象  $X_1, X_2, X_3$  がそれぞれ平均  $\mu_1, \mu_2, \mu_3$ 、分散  $\sigma_1^2, \sigma_2^2, \sigma_3^2$  の正規分布に従うとする。すなわち、  $X_1 \sim N(\mu_1, \sigma_1^2), X_2 \sim N(\mu_2, \sigma_2^2), X_3 \sim N(\mu_3, \sigma_3^2)$  である。分散は皆等しく、しかも簡

|    | A                               | B             | C       | D       | E             | F          | G          |
|----|---------------------------------|---------------|---------|---------|---------------|------------|------------|
| 1  | 平均値の差のシミュレーション (q値とmax q値の分布形状) |               |         |         |               |            |            |
| 2  |                                 |               |         |         |               |            |            |
| 3  | 入力                              | $\mu 1$       | $\mu 2$ | $\mu 3$ | $\sigma 1$    | $\sigma 2$ | $\sigma 3$ |
| 4  |                                 | 1             | 1       | 1       | 1             | 1          | 1          |
| 5  |                                 |               |         |         |               |            |            |
| 6  |                                 | 1組目           |         |         | 2組目           |            |            |
| 7  | 計算値                             | x1            | x2      | x3      | x1            | x2         | x3         |
| 8  |                                 | 1.917         | 1.916   | 2.736   | 1.323         | -0.093     | 1.858      |
| 9  |                                 | 1.183         | 0.631   | 1.041   | 1.679         | -0.605     | 1.339      |
| 24 | 出力                              | 検定統計量 q12     | q13     | q23     | q12           | q13        | q23        |
| 25 |                                 | 0.605         | 0.442   | 0.164   | -0.195        | -0.410     | 0.215      |
| 26 |                                 | qijを四捨五入      |         |         | qijを四捨五入      |            |            |
| 27 |                                 | 1             | 0       | 0       | 0             | 0          | 0          |
| 28 |                                 | max  qij      |         |         | max  qij      |            |            |
| 29 |                                 | 0.605         |         |         | 0.410         |            |            |
| 30 |                                 | max qij を四捨五入 |         |         | max qij を四捨五入 |            |            |
| 31 |                                 | 1             |         |         | 0             |            |            |

図 8.6:  $q$  値と  $q_{max}$  値を求めるシミュレーション ( $q$  値と  $q_{max}$  値のシミュレーション.xlsxm)図 8.7:  $q$  値と  $q_{max}$  値の頻度分布図 8.8:  $q$  値と  $q_{max}$  値の確率密度関数の形状

単のために 1 であるとする ( $\sigma_1^2 = \sigma_2^2 = \sigma_3^2 = 1$ ). さらに, 各群のデータ数は同じとする. 検定統計量の閾値を  $u_0$  とする. 事象  $X_1, X_2, X_3$  のデータ群の平均値をそれぞれ  $\bar{x}_1, \bar{x}_2, \bar{x}_3$  とすると, データ群  $i, j (i, j = 1, 2, 3, i \neq j)$  の平均値間の検定統計量  $u_{ij}$  は式 (6.1) に,  $n_i = n_j = n, \sigma_i = \sigma_j = 1$  を代入することで

$$u_{ij} = \frac{\bar{x}_i - \bar{x}_j}{\sqrt{\frac{2}{n}}} \quad (8.9)$$



と得られる。  $u_{ij}$  は標準正規分布に従う。検定の仮説は、

$$\begin{aligned} \text{帰無仮説} & : \mu_i = \mu_j \\ \text{対立仮説} & : \mu_i \neq \mu_j \\ & \text{ただし, } i, j = 1, 2, 3, i \neq j \end{aligned} \quad (8.10)$$

である。この帰無仮説は

$$\max_{i,j=1,2,3,i \neq j} |u_{ij}| \leq u_0 \quad (8.11)$$

もしくは、

$$|u_{12}| \leq u_0 \quad \text{and} \quad |u_{13}| \leq u_0 \quad \text{and} \quad |u_{23}| \leq u_0 \quad (8.12)$$

と表現できる。いま、  $\bar{x}_1 \geq \bar{x}_2, \bar{x}_3$  の場合を考える。このとき上式を満たす  $\bar{x}_1, \bar{x}_2, \bar{x}_3$  は

$$\begin{aligned} 0 & \leq \frac{\bar{x}_1 - \bar{x}_2}{\sqrt{\frac{2}{n}}} \leq u_0 \\ 0 & \leq \frac{\bar{x}_1 - \bar{x}_3}{\sqrt{\frac{2}{n}}} \leq u_0 \end{aligned} \quad (8.13)$$

の範囲にある。なお、この式により  $|u_{23}| \leq u_0$  は満たされる。仮定の  $\mu_1 = \mu_2 = \mu_3$  を挿入すると

$$\begin{aligned} 0 & \leq \frac{\bar{x}_1 - \mu_1 - \bar{x}_2 + \mu_2}{\sqrt{\frac{2}{n}}} \leq u_0 \\ 0 & \leq \frac{\bar{x}_1 - \mu_1 - \bar{x}_3 + \mu_3}{\sqrt{\frac{2}{n}}} \leq u_0 \end{aligned} \quad (8.14)$$

であり、変形すると

$$\begin{aligned} \sqrt{n}(\bar{x}_1 - \mu_1) - \sqrt{2}u_0 & \leq \sqrt{n}(\bar{x}_2 - \mu_2) \leq \sqrt{n}(\bar{x}_1 - \mu_1) \\ \sqrt{n}(\bar{x}_1 - \mu_1) - \sqrt{2}u_0 & \leq \sqrt{n}(\bar{x}_3 - \mu_3) \leq \sqrt{n}(\bar{x}_1 - \mu_1) \end{aligned} \quad (8.15)$$

となる。式(3.67)より、

$$\sqrt{n}(\bar{X}_i - \mu_i) \sim N(0, 1) \quad (i = 1, 2, 3) \quad (8.16)$$

であることに注意すれば、  $\bar{x}_1 = \bar{x}_{10}$  のとき、  $\bar{x}_2, \bar{x}_3$  が式(8.15)を満たす確率は、

$$\left\{ \Phi(w_{10}) - \Phi(w_{10} - \sqrt{2}u_0) \right\}^2 \quad (8.17)$$

と求められる。ただし、 $w_{10} = \sqrt{n}(\bar{x}_{10} - \mu_1)$  である。また、 $\Phi(\cdot)$  は標準正規分布の確率分布関数であり、次式で与えられる。

$$\Phi(x) = \int_{-\infty}^x \frac{1}{\sqrt{2\pi}} \exp\left(-\frac{t^2}{2}\right) dt \quad (8.18)$$

これは、Excel では NORMSDIST(x) 関数の出力値である。 $f(x)$  を標準正規分布の確率密度関数

$$f(x) = \frac{1}{\sqrt{2\pi}} \exp\left(-\frac{x^2}{2}\right) \quad (8.19)$$

とすると、 $w_{10}$  が標準正規分布に従うので、 $\bar{x}_1 \geq \bar{x}_2, \bar{x}_3$  の場合に式 (8.11) を満たす確率  $P_1$  は

$$P_1 = \int_{-\infty}^{\infty} \left\{ \Phi(w_{10}) - \Phi(w_{10} - \sqrt{2}u_0) \right\}^2 f(w_{10}) dw_{10} \quad (8.20)$$

と求められる。

次に  $\bar{x}_2 \geq \bar{x}_1, \bar{x}_3$  と  $\bar{x}_3 \geq \bar{x}_1, \bar{x}_2$  の各場合に式 (8.11) を満たす確率  $P_2, P_3$  を考える。 $P_1$  の場合と同様の式展開により

$$P_1 = P_2 = P_3 \quad (8.21)$$

であることが分かる。また、 $P_1, P_2, P_3$  をもたらす各事象がそれぞれ同時には起こりえない排反事象であることに注意すれば、式 (8.11) を満たす確率  $P(\max |u_{i,j}| \leq u_0)$  は

$$\begin{aligned} P(\max |u_{i,j}| \leq u_0) &= P_1 + P_2 + P_3 \\ &= 3 \int_{-\infty}^{\infty} \left\{ \Phi(x) - \Phi(x - \sqrt{2}u_0) \right\}^2 f(x) dx \end{aligned} \quad (8.22)$$

である。第 1 種の過誤の確率  $P(\max |u_{i,j}| > u_0)$  は

$$P(\max |u_{i,j}| > u_0) = 1 - P(\max |u_{i,j}| \leq u_0) \quad (8.23)$$

と求められる。

図 8.9 は、式 (8.23) を数値計算により求めている Excel の画面である。セル D4 に公称の有意水準  $\alpha = 0.05$  が与えられ、ここでは、この  $\alpha$  からセル B6 にて両側検定の閾値  $u_0$  を決めている。検定統計量  $u_{ij}$  が標準正規分布に従うので標準正規分布の分布関数の逆関数を用いている。数値積分を実行するにあたって  $5\sigma$  より外の区間の確率は低いとして無視し、 $-5 \leq x \leq 5$  の区間を 1000 等分している。セル C12 以下では  $x_k = -5 + 0.01 \times k$  ( $k = 0, 1, \dots, 999$ ) の座標毎に、微小区間ごとの積分の近似値

$$3 \left\{ \Phi(x_k) - \Phi(x_k - \sqrt{2}u_0) \right\}^2 f(x_k) \Delta x \quad (8.24)$$

|    | A                           | B              | C           | D                |
|----|-----------------------------|----------------|-------------|------------------|
| 1  | 第1種の過誤の数値計算(母平均の差の検定, 分散既知) |                |             |                  |
| 2  |                             |                |             |                  |
| 3  | 入力                          | データ群数 a        | 群内データ数 n    | 公称の有意水準 $\alpha$ |
| 4  |                             | 3              | 9           | 0.05             |
| 5  | 計算値                         | 検定統計量の閾値 $u_0$ |             |                  |
| 6  |                             | 1.9600         |             |                  |
| 7  | 出力                          | 第1種の過誤 $p^*$   |             |                  |
| 8  |                             | 0.12227        |             |                  |
| 9  |                             |                |             |                  |
| 10 |                             |                |             |                  |
| 11 | 数値計算値                       | $x_k$          | $\Delta p$  |                  |
| 12 |                             | -5             | 3.66487E-21 |                  |
| 13 |                             | -4.99          | 4.27327E-21 |                  |
| 14 |                             | -4.98          | 4.9812E-21  |                  |
| 15 |                             | -4.97          | 5.80467E-21 |                  |

$=\text{NORMSINV}(1-\alpha/2)$

$=1-\text{SUM}(C12:C1011)$

$3\{\Phi(x_k) - \Phi(x_k - \sqrt{2}u_0)\}^2 f(x_k)\Delta x$

図 8.9: 多重性の問題における第1種の過誤の確率の数値計算(母平均の差の検定, 母分散既知)(母分散既知での数値計算.xlsx)

を求めている。  $\Delta x = 1/100$  である。得られた第1種の過誤の確率  $p^* = 0.12227$  であった。

図 8.10 は3個のデータ群においてそれぞれ標準正規分布から9個ずつの乱数を生成して、式(6.1)の検定統計量  $u$  による検定を1000組について行っているシミュレーションの画面である。24行目にて、 $u_{ij}$  が標準正規分布に従うので、標準正規分布の分布関数を用いて両側検定の  $p_{ij}$  値を求めている。図 8.11 は1000組のシミュレーションを100回繰り返して、第1種の過誤の確率の95%信頼区間を求めた結果である。シミュレーションの結果は数値計算の結果と一致している。

### 母分散未知の場合

次に、母分散未知で等しいとみなせる場合について第1種の過誤の確率を求める。事象  $X_1, X_2, X_3$  の分散  $\sigma_1^2 = \sigma_2^2 = \sigma_3^2 = \sigma^2$  とする。両側検定の検定統計量の閾値を  $q_0$  とする。データ群  $i, j (i, j = 1, 2, 3, i \neq j)$  の平均値間の検定統計量  $q_{ij}$  は式(8.5)により

$$q_{ij} = \frac{|\bar{x}_i - \bar{x}_j|}{\sqrt{\frac{2v_E^2}{n}}} \quad (8.25)$$

である。ただし、

$$v_E^2 = \frac{v_{e1}^2 + v_{e2}^2 + v_{e3}^2}{3} \quad (8.26)$$

|    | A                          | B                                  | C                 | D                 | E                                | F                 | G                 |
|----|----------------------------|------------------------------------|-------------------|-------------------|----------------------------------|-------------------|-------------------|
| 1  | 母平均の差の検定のシミュレーション(分散既知の場合) |                                    |                   |                   |                                  |                   |                   |
| 2  |                            |                                    |                   |                   |                                  |                   |                   |
| 3  | 入力                         | $\mu$                              | $\sigma$          | 公称の有意水準 $\alpha$  |                                  |                   |                   |
| 4  |                            | 0                                  | 1                 |                   | 0.05                             |                   |                   |
| 5  | 計算値                        | 第1組                                |                   |                   | 第2組                              |                   |                   |
| 6  |                            | x1                                 | x2                | x3                | x1                               | x2                | x3                |
| 7  |                            | 1.0321                             | 1.7386            | 0.0416            | 0.5920                           | 1.0066            | -0.7591           |
| 8  |                            | 0.6710                             | -0.7537           | -1.6745           | -0.6028                          | -4.3083           | -0.3604           |
| 21 |                            | 検定統計量 u12                          | u13               | u23               | u12                              | u13               | u23               |
| 22 |                            | 1.067                              | 0.294             | 1.361             | 0.406                            | 0.079             | 0.486             |
| 23 |                            | p12                                | p13               | p23               | p12                              | p13               | p23               |
| 24 |                            | 0.2861                             | 0.7687            | 0.1735            | 0.6845                           | 0.9367            | 0.6271            |
| 25 |                            | if p12 < $\alpha$<br>then 1 else 0 | if p13 < $\alpha$ | if p23 < $\alpha$ | if p12 < $\alpha$<br>then 1 else | if p13 < $\alpha$ | if p23 < $\alpha$ |
| 26 |                            | 0                                  | 0                 | 0                 | 0                                | 0                 | 0                 |
| 27 |                            | 各対のいずれか<br>で有意差あり                  |                   |                   | 各対のいずれ<br>れかで有意                  |                   |                   |
| 28 |                            | 0                                  |                   |                   | 0                                |                   |                   |
| 29 | 出力                         | 有意頻度 fr                            |                   |                   |                                  |                   |                   |
| 30 |                            | 0.123                              |                   |                   | =2*(1-NORMSDIST(B22))            |                   |                   |

図 8.10: 母平均の差の検定における多重性の問題のシミュレーション（3群，母分散既知）（母分散既知でのシミュレーション.xlsxm）

|    |     |          |                      |  |  |
|----|-----|----------|----------------------|--|--|
| 51 |     | 有意頻度 frk | 100回の平均              |  |  |
| 52 | fr1 | 0.119    | 0.1234               |  |  |
| 53 | fr2 | 0.12     | 95%信頼区間              |  |  |
| 54 | fr3 | 0.128    | 0.1212 < p* < 0.1255 |  |  |

図 8.11: 母平均の差の検定における第1種の過誤の確率の95%信頼区間（3群，分散既知）

である。  $q_{ij}$  は自由度  $\nu = 3(n - 1)$  の  $t$  分布に従う。検定の仮説は，データ群  $i$  の母平均を  $\mu_i$  とすると，前項と同様に

$$\begin{aligned}
 \text{帰無仮説} &: \mu_i = \mu_j \\
 \text{対立仮説} &: \mu_i \neq \mu_j \\
 &\text{ただし, } i, j = 1, 2, 3, i \neq j
 \end{aligned} \tag{8.27}$$

である。この帰無仮説は

$$\max_{i,j=1,2,3,i \neq j} |q_{ij}| \leq q_0 \tag{8.28}$$

もしくは，

$$|q_{12}| \leq q_0 \quad \text{and} \quad |q_{13}| \leq q_0 \quad \text{and} \quad |q_{23}| \leq q_0 \tag{8.29}$$

と表現できる．いま， $\bar{x}_1 \geq \bar{x}_2, \bar{x}_3$  の場合を考える．このとき上式を満たす  $\bar{x}_1, \bar{x}_2, \bar{x}_3$  は

$$\begin{aligned} 0 &\leq \frac{\bar{x}_1 - \bar{x}_2}{\sqrt{\frac{2v_E^2}{n}}} \leq q_0 \\ 0 &\leq \frac{\bar{x}_1 - \bar{x}_3}{\sqrt{\frac{2v_E^2}{n}}} \leq q_0 \end{aligned} \quad (8.30)$$

の範囲にある．仮定の  $\mu_1 = \mu_2 = \mu_3, \sigma_1^2 = \sigma_2^2 = \sigma_3^2 = \sigma^2$  を挿入して，変形すると

$$\begin{aligned} \frac{\bar{x}_1 - \mu_1}{\sqrt{\frac{\sigma^2}{n}}} - \sqrt{2} \frac{v_E}{\sigma} q_0 &\leq \frac{\bar{x}_2 - \mu_2}{\sqrt{\frac{\sigma^2}{n}}} \leq \frac{\bar{x}_1 - \mu_1}{\sqrt{\frac{\sigma^2}{n}}} \\ \frac{\bar{x}_1 - \mu_1}{\sqrt{\frac{\sigma^2}{n}}} - \sqrt{2} \frac{v_E}{\sigma} q_0 &\leq \frac{\bar{x}_3 - \mu_3}{\sqrt{\frac{\sigma^2}{n}}} \leq \frac{\bar{x}_1 - \mu_1}{\sqrt{\frac{\sigma^2}{n}}} \end{aligned} \quad (8.31)$$

となる．式(3.67)より，

$$\frac{\bar{X}_i - \mu_i}{\sqrt{\frac{\sigma^2}{n}}} \sim N(0, 1) \quad (i = 1, 2, 3) \quad (8.32)$$

であり， $\sqrt{n/\sigma^2}(\bar{X}_i - \mu_i)$  は互いに独立に標準正規分布に従う．

では， $v_E/\sigma$  はどのような分布に従うであろうか？

$$\sqrt{\frac{V_E^2}{\sigma^2}} = \sqrt{\frac{1}{3(n-1)} \frac{\sum_{i=1}^3 \sum_{j=1}^{n_i} (X_{ij} - \bar{X}_i)^2}{\sigma^2}} \quad (8.33)$$

において，式(6.78)より，

$$\frac{\sum_{i=1}^3 \sum_{j=1}^{n_i} (X_{ij} - \bar{X}_i)^2}{\sigma^2} \sim \chi_{3(n-1)}^2 \quad (8.34)$$

である． $x$  が自由度  $\nu$  の  $\chi^2$  分布に従うとすると， $x$  の確率密度関数  $g(x)$  は，式(4.71)より

$$g(x) = \frac{1}{\Gamma\left(\frac{\nu}{2}\right) 2^{\frac{\nu}{2}}} x^{\frac{\nu}{2}-1} \exp\left(-\frac{x}{2}\right) \quad (8.35)$$

と表される．式(8.33)を考慮して

$$s = \sqrt{\frac{x}{\nu}} \quad (8.36)$$

により変数変換を行うと

$$dx = 2\nu s ds \quad (8.37)$$

より,

$$\begin{aligned}
 h_\nu(s) &= \frac{\nu^{\frac{\nu}{2}-1}}{\Gamma\left(\frac{\nu}{2}\right) 2^{\frac{\nu}{2}}} s^{\nu-2} \exp\left(-\frac{\nu s^2}{2}\right) \frac{dx}{ds} \\
 &= \frac{\nu^{\frac{\nu}{2}}}{\Gamma\left(\frac{\nu}{2}\right) 2^{\frac{\nu}{2}-1}} s^{\nu-1} \exp\left(-\frac{\nu s^2}{2}\right)
 \end{aligned}
 \tag{8.38}$$

となる. よって,  $h_{3(n-1)}(s)$  は式 (8.33) の事象  $S = V_E/\sigma$  の確率密度関数を与える.

4.8.2項の結果より, 事象  $\sqrt{n/\sigma^2}(\bar{X}_i - \mu_i)$  と  $\sum_{j=1}^{n_i} (X_{ij} - \bar{X}_i)^2$  は独立である. また, もともと事象  $\sqrt{n/\sigma^2}(\bar{X}_i - \mu_i)$  ( $i = 1, \dots, a$ ) は互いに独立である. よって, 事象  $\sqrt{n/\sigma^2}(\bar{X}_i - \mu_i)$  ( $i = 1, \dots, a$ ) と事象  $V_E^2 = \sum_i^a \sum_{j=1}^{n_i} (X_{ij} - \bar{X}_i)^2$  は互いに独立である. このことから, 事象  $\sqrt{n/\sigma^2}(\bar{X}_i - \mu_i)$  ( $i = 1, \dots, a$ ) と事象  $S = V_E/\sigma$  が互いに独立であるので,  $\bar{x}_1 \geq \bar{x}_2, \bar{x}_3$  の場合に式 (8.28) を満たす確率  $P_1$  は,  $\bar{x}_1 = x_{10}$  のとき,  $w_{10} = \sqrt{n/\sigma^2}(x_{10} - \mu_1)$  とし, また,  $s = v_E/\sigma$  とすると

$$P_1 = \int_0^\infty \int_{-\infty}^\infty \left\{ \Phi(w_{10}) - \Phi(w_{10} - \sqrt{2}sq_0) \right\}^2 f(w_{10}) h_{3(n-1)}(s) dw_{10} ds
 \tag{8.39}$$

と求められる.

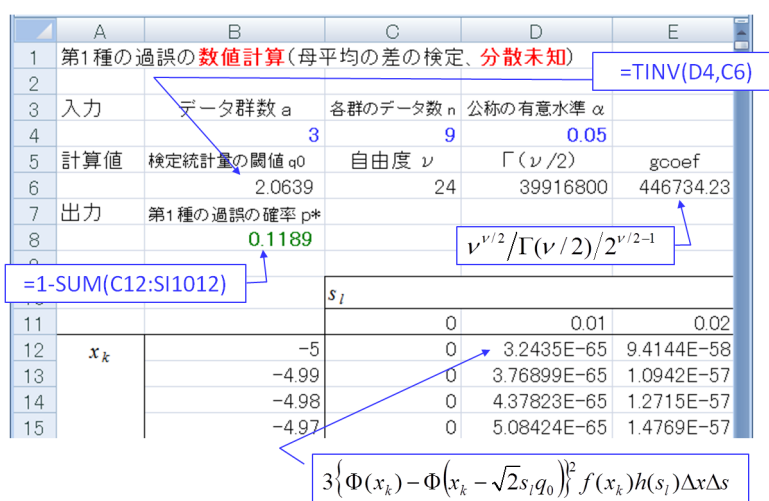


図 8.12: 多重性の問題における第 1 種の過誤の確率の数値計算（平均値の差の検定, 母分散未知）(母分散未知での数値計算.xlsm)

以降, 母分散既知の場合と同様の展開により, 式 (8.28) を満たす確率  $P(\max |q_{ij}| \leq q_0)$  は

$$P(\max |q_{ij}| \leq q_0) = 3 \int_0^\infty \int_{-\infty}^\infty \left\{ \Phi(x) - \Phi(x - \sqrt{2}sq_0) \right\}^2 f(x) h_{3(n-1)}(s) dx ds
 \tag{8.40}$$

```

Sub Γの計算()
'Γ(ν/2)の計算
Application.Calculation = xlCalculationManual 'シートの計算をマニュアル実行とする.
nu = Cells(6, 3) '自由度νの読み込み
Fact = 1
If nu / 2 - Int(nu / 2) > 0 Then '自由度が奇数の時のΓの計算
For i = nu - 2 To 1 Step -2
Fact = Fact * i / 2
Next i
Fact = Fact * Sqr(3.141592654)
Else '自由度が偶数の時のΓの計算
b = nu / 2
For i = b - 1 To 1 Step -1
Fact = Fact * i
Next i
End If
Cells(6, 4).Value = Fact 'Γ(ν/2)をセルE6に出力
Calculate 'シートの計算の実行
Application.Calculation = xlCalculationAutomatic 'シートの計算のマニュアル実行の解除
End Sub

```

図 8.13: Γ関数計算のマクロ

と得られ、第1種の過誤の確率  $P(\max |q_{i,j}| > q_0)$  は

$$P(\max |q_{ij}| > q_0) = 1 - P(\max |q_{ij}| \leq q_0) \quad (8.41)$$

となる。

図 8.12 は式 (8.41) を数値計算により求めている Excel の画面である。入力項目は、データ群数  $a$  (セル B4)、各群のデータ数  $n$  (セル C4)、公称の有意水準  $\alpha$  (セル D4) である。Excel はこれらの項目が手作業で入力される度にシートの再計算を自動的に実行する。ただし、セル D6 の  $\Gamma(\nu/2)$  の値は次の段落で述べるマクロを手作業で実行させる必要がある。マクロの起動は 6.4.1 を参照されたい。マクロはセル C6 の自由度  $\nu$  を読み込んで、セル D6 に  $\Gamma(\nu/2)$  の値を書き込む。この書き込みによりシートの再計算をさせる。セル D4 に公称の有意水準  $\alpha = 0.05$  が与えられると、式 (8.25) の検定統計量が自由度  $\nu = 3(n-1)$  の  $t$  分布に従うので、セル B6 では  $\alpha$  をもとにセル C6 の自由度  $\nu$  を用いて  $t$  分布の分布関数の逆関数により閾値  $q_0$  を決めている。数値積分にあたっては  $|x| \geq 5, s \geq 5$  の範囲の確率は低いとして無視し、 $-5 \leq x \leq 5$  の区間を 1000 等分、 $0 \leq s \leq 5$  の区間を 500 等分している。12 行目以下では  $x_k = -5 + 0.01 \times k$  ( $k = 0, 1, \dots, 999$ )、C 列から右の列では  $s_l = 0 + 0.01 \times l$  ( $l = 0, 1, \dots, 499$ ) の微小区間に分け、微小面積  $\Delta x \Delta s$  ごとの積分の近似値

$$3 \left\{ \Phi(x_k) - \Phi(x_k - \sqrt{2}s_l q_0) \right\}^2 f(x_k) h(s_l) \Delta x \Delta s \quad (8.42)$$

を求めている。  $\Delta x \Delta s = 1/10000$  である。セル B8 にて、積分の近似値の総和を 1 から引

くことで第1種の過誤の確率  $p^*$  を求めている。結果は  $p^* = 0.1189$  であった。8.1.2 項の図 8.3 のシミュレーション結果は数値計算の結果と一致している。

なお、式 (8.38) の確率密度関数の値を求めるには、4.6.2 項に述べた  $\Gamma$  関数の計算を必要とする。この計算は自由度の偶数、奇数に応じて場合分けをしなければならないので、図 8.13 に示すマクロを作成した。このマクロはセル C6 (マクロでは行と列の順番が反転して、(6, 3) と表現している) の自由度を読み込んで、奇数の場合には式 (4.51), (4.50) により、偶数の場合には式 (4.51), (4.38) により  $\Gamma$  関数値を計算し、セル D6 (同 (6, 4)) に結果を出力する。なお、セル E6 では式 (8.38) の中で変数  $s$  に依存しない部分の値を求め、数値積分の全てのセルでこの値を利用している。

#### 閾値 $q_0$ の調整—チューキーの方法—

図 8.14 はチューキーの方法を用いた場合の第1種の過誤の確率の数値計算結果である。図 8.12 との違いは検定統計量の閾値  $q_0$  の与え方だけである。チューキーの方法は第1種の過誤の確率が 0.05 となるように、検定統計量  $q_0$  の値を調整する。付録の表 A.1 より、データ群数  $a = 3$ 、自由度  $\nu = 24$  のとき、 $\sqrt{2}q_0 = 3.532$  と得られる。セル B6 では INDIRECT 関数により閾値の  $q_0$  値を自動で得ている。

$$= \text{INDIRECT}(\text{ADDRESS}(\text{C6}, \text{B4}, , , \text{"Sheet2"})) / \text{SQRT}(2) \quad (8.43)$$

この Excel ブックには Sheet2 に表 A.1 と同じものをコピーしてある (図 8.15)。表では行がセル C6 の自由度  $\nu$  に対応づけられ、列がセル B4 のデータ群数  $a$  に対応づけられている。 $\nu, a$  を行、列のパラメータとして ADDRESS 関数が Sheet2 の対応セルのアドレスを読み出し、INDIRECT 関数がそのアドレスの指示するセル内の値を読み出している。セル B6 では読み出した値に  $1/\sqrt{2}$  をかけて  $q_0$  値を求めている。第1種の過誤の確率の計算結果は  $p^* = 0.04998$  となった。8.1.2 項の図 8.5 のシミュレーション結果は数値計算の結果と一致している。

式 (8.40) は、データ群数  $a$ 、各群のデータ数  $n$  の場合に容易に拡張できる。

$$P(\max |q_{ij}| \leq q_0) = a \int_0^\infty \int_{-\infty}^\infty \left\{ \Phi(x) - \Phi(x - \sqrt{2}sq_0) \right\}^{a-1} f(x) h_{a(n-1)}(s) dx ds \quad (8.44)$$

実は、図 8.12 は上式の数値計算を実行できる設定となっている。この Excel のシートを利用して、表 A.1 のステューデント化された範囲の 5% 点の値を求めることができる。図 8.16 はその画面である。ユーザは公称の有意水準  $\alpha$ 、データ群数（水準数とも呼ばれる） $a$ 、自由度  $\nu$  を入力し、図 8.17 のマクロを実行することで、セル B9 にステューデント化



|    | A                                   | B               | C         | D                | E          | F          |
|----|-------------------------------------|-----------------|-----------|------------------|------------|------------|
| 1  | 第1種の過誤の数値計算(母平均の差の検定、分散未知、テューキーの方法) |                 |           |                  |            |            |
| 2  |                                     |                 |           |                  |            |            |
| 3  | 入力                                  | データ群数 a         | 各群のデータ数 n | 公称の有意水準 $\alpha$ |            |            |
| 4  |                                     | 3               | 9         | 0.05             |            |            |
| 5  | 計算値                                 | 検定統計量の閾値 $q_0$  | 自由度 $\nu$ | $\Gamma(\nu/2)$  | gcoef      |            |
| 6  |                                     | 2.4975          | 24        | 39916800         | 446734.23  |            |
| 7  | 出力                                  | 第1種の過誤の確率 $p^*$ |           |                  |            |            |
| 8  |                                     | 0.04998         |           |                  |            |            |
| 9  |                                     |                 |           |                  |            |            |
| 10 | 数値計算値                               |                 | $s_i$     |                  |            |            |
| 11 |                                     |                 |           | 0                | 0.01       | 0.02       |
| 11 |                                     |                 |           | 0                | 0.01       | 0.02       |
| 12 | $x_k$                               | -5              | 0         | 4.60932E-65      | 1.3002E-57 | 2.7661E-53 |
| 13 |                                     | -4.99           | 0         | 5.35639E-65      | 1.5115E-57 | 3.2164E-53 |
| 14 |                                     | -4.98           | 0         | 6.22258E-65      | 1.7564E-57 | 3.7389E-53 |
| 15 |                                     | -4.97           | 0         | 7.2265E-65       | 2.0405E-57 | 4.3449E-53 |

図 8.14: テューキーの方法による第1種の過誤の確率の数値計算 (母平均の差の検定)  
(テューキーによる数値計算.xlsm)

|    | A         | B     | C     | D     | E      | F      | G      |
|----|-----------|-------|-------|-------|--------|--------|--------|
| 1  | $\nu$ \ a | 2     | 3     | 4     | 5      | 6      | 7      |
| 2  | 2         | 6.085 | 8.331 | 9.798 | 10.881 | 11.734 | 12.435 |
| 3  | 3         | 4.501 | 5.91  | 6.825 | 7.502  | 8.037  | 8.478  |
| 4  | 4         | 3.926 | 5.04  | 5.757 | 6.287  | 6.706  | 7.053  |
| 5  | 5         | 3.635 | 4.602 | 5.218 | 5.673  | 6.033  | 6.33   |
| 6  | 6         | 3.46  | 4.339 | 4.896 | 5.305  | 5.628  | 5.895  |
| 7  | 7         | 3.344 | 4.165 | 4.681 | 5.06   | 5.359  | 5.606  |
| 8  | 8         | 3.261 | 4.041 | 4.529 | 4.886  | 5.167  | 5.399  |
| 9  | 9         | 3.199 | 3.948 | 4.415 | 4.755  | 5.024  | 5.244  |
| 10 | 10        | 3.151 | 3.877 | 4.327 | 4.654  | 4.912  | 5.124  |
| 11 | 11        | 3.113 | 3.82  | 4.256 | 4.574  | 4.823  | 5.028  |

図 8.15: Sheet2にコピーされたステューデント化された範囲の表

された範囲の5%点 ( $\sqrt{2}q_0$ )を得ることができる。マクロは、まず  $\Gamma(\nu/2)$  の計算を行い、結果をセル B5 に書き込む。次に  $q_0$  の候補値をセル C7 に書き込んだ後にシートの計算を実行させる。この計算によりセル D7 には第1種の過誤の確率  $p^*$  が出力される。マクロはこの値を読み込んで、これがセル B3 で指定された有意水準  $\alpha \pm 0.0000001$  以内となるか、もしくはシートの計算実行回数が30回 (探索回数  $n_{search}$  とよぶ) になるまで探索を繰り返す。探索には  $(q_{min}, q_{max}) = (2.7, 32)$  の区間の2分探索法を採用している。探索終了時には第1種の過誤の確率を  $\alpha$  の値とする  $q_0$  値がセル C7 に、そしてその  $\sqrt{2}$  倍の値であるステューデント化された範囲の  $\alpha\%$  点 がセル B9 に得られる。  $2 \leq a \leq 9, 2 \leq \nu \leq 240$  の範囲で数表の値と一致することを確認してある。

|    | A                | B                    | C                            | D                             | E                     |
|----|------------------|----------------------|------------------------------|-------------------------------|-----------------------|
| 1  | ステューデント化された範囲の探索 |                      |                              |                               |                       |
| 2  | 入力               | 公称の有意水準<br>$\alpha$  | データ群数<br>$a = 2 \sim 10$ の整数 | 自由度<br>$\nu = 2 \sim 250$ の整数 |                       |
| 3  |                  | 0.05                 | 3                            | 24                            |                       |
| 4  | 計算値              | $\Gamma(\nu/2)$      | gの係数                         |                               |                       |
| 5  |                  | 39916800             | 446734.2296                  |                               |                       |
| 6  |                  | 探索回数<br>$n_{search}$ | $q_0$                        | 第1種の過誤の<br>確率 $p^*$           |                       |
| 7  |                  | 24                   | 2.4973                       | 0.0500000                     |                       |
| 8  | 出力               | ステューデント化され<br>た範囲の%点 |                              |                               | $=1-SUM(C13:AAA1213)$ |
| 9  |                  | 3.532                | $=\sqrt{2}q_0$               |                               |                       |
| 10 |                  |                      |                              |                               |                       |
| 11 |                  |                      | $s_i$                        |                               |                       |
| 12 |                  |                      | 0                            | 0.01                          | 0.02                  |
| 13 | $x_k$            | -6                   | 0                            | 3.03998E-72                   | 8.3065E-65            |
| 14 |                  | -5.99                | 0                            | 3.64021E-72                   | 9.9497E-65            |
| 15 |                  | -5.98                | 0                            | 4.35764E-72                   | 1.1914E-64            |
| 16 |                  | -5.97                | 0                            | 5.21491E-72                   | 1.4263E-64            |

図 8.16: ステューデント化された範囲の探索（ステューデント化された範囲の探索.xlsxm）

```

Sub q0値の探索マクロ() '(qmin qmax)の区間の二分探索法
Application.Calculation = xlCalculationManual 'シートの計算をマニュアル実行とする。
Cells(7, 2).Value = "探索中"
alpha = Cells(3, 2).Value '有意水準の読み込み
NsearchMax = 30 '探索回数の上限設定
nsearch = 1 '探索回数の格納用変数
Call Γの計算サブ 'Γ(ν/2)の計算
qmax = 32 / Sqr(2) 'q値の候補の上限值設定
qmin = 2.7 / Sqr(2) 'q値の候補の下限値設定
qmean = (qmin + qmax) / 2 '暫定q値の計算
Cells(7, 3).Value = qmean
Calculate '暫定q値をセル(7,3)に書きだした直後にp値の計算を実行
p_value = Cells(7, 4).Value 'p値の読み込み
While Abs(p_value - alpha) > 0.0000001 And nsearch < NsearchMax
    'p値が有意水準±0.0000001の範囲内に入るまで
    'もしくは探索回数qNsearchMaxとなるまでq値の探索を実行
    If p_value - alpha > 0 Then
        qmin = qmean
    Else
        qmax = qmean
    End If
    qmean = (qmin + qmax) / 2 '暫定q値の計算
    Cells(7, 3).Value = qmean
    Calculate '暫定q値をセル(7,3)に書き出した直後にp値の計算を実行
    p_value = Cells(7, 4).Value 'p値の読み込み
    nsearch = nsearch + 1 '探索回数のカウントアップ
Wend
If nsearch = NsearchMax Then
    Cells(7, 2).Value = "探索回数が上限(N=30)に達しましたの強制終了しました。"
Else
    Cells(7, 2).Value = nsearch '探索回数の表示
End If
Application.Calculation = xlCalculationAutomatic 'シートの計算のマニュアル実行の解除。
End Sub

```

図 8.17: ステューデント化された範囲の探索マクロ

## 8.2 母平均の差の検定（データ数が異なる場合）

### 8.2.1 テューキー・クレーマーの方法による検定の例

ある日、ある工場のある製品のラインにおいて、ラインのスピードを3通りに変えながらあるパラメータの測定がなされたとする。スピード設定1, 2, 3からそれぞれ11, 9, 7個の標本を取り出すことができ、以下の値が得られたとする。

設定1: 3.3, 3.2, 3.4, 3.3, 3.3, 3.2, 3.5, 3.2, 3.2, 3.4, 3.5

設定2: 3.2, 3.1, 3.3, 3.2, 3.2, 3.1, 3.4, 3.0, 3.2

設定3: 3.1, 3.0, 3.3, 3.1, 3.2, 3.0, 3.2

ライン1, 2, 3の平均値  $\bar{x}_1, \bar{x}_2, \bar{x}_3$  はそれぞれ 3.32, 3.19, 3.13 であった。スピード設定を変えたことで母平均に差が出たと言えるか？ただし、いずれの設定の母分散にも差がないとする。

テューキーの方法は各群のデータ数が同じ場合の多重比較法であった。各群でデータ数が異なる場合にはテューキーの方法による判定結果は理論的に保証されたものではない。各群のデータ数が異なる場合の母平均の差の検定の多重比較法にテューキー・クレーマー (Tukey Kramer) の方法 [16] がある。設定  $i, j (i, j = 1, 2, 3, i \neq j)$  の平均値間の検定統計量  $q_{ij}$  は、式 (8.2) に立ち返って、

$$q_{ij} = \frac{|\bar{x}_i - \bar{x}_j|}{\sqrt{v_E^2 \left( \frac{1}{n_i} + \frac{1}{n_j} \right)}} \quad (8.45)$$

で与えられる。ただし、

$$v_E^2 = \frac{\sum_{k=1}^a \sum_{l=1}^{n_k} (x_{kl} - \bar{x}_k)^2}{\sum_{k=1}^a (n_k - 1)} \quad (8.46)$$

である。  $n_k$  は設定  $k$  のデータ数であり、  $a$  は設定の数である。この  $q_{ij}$  値は自由度

$$\nu = \sum_{k=1}^a (n_k - 1) \quad (8.47)$$

の  $t$  分布に従う。検定の仮説は、データ群  $i$  の母平均を  $\mu_i$  とすると、前項と同様に

$$\text{帰無仮説} : \mu_i = \mu_j$$

$$\text{対立仮説} : \mu_i \neq \mu_j$$

$$\text{ただし, } i, j = 1, 2, 3, i \neq j \quad (8.48)$$

である。図 8.18 は検定の実行画面である。26 行目のセルにて式 (8.45) による  $q_{12}, q_{13}, q_{23}$  の値を求めている。得られた  $q_{ij}$  値はそれぞれ 2.49, 3.40, 1.04 であった。設定数  $a = 3$ , 自由度  $\nu = 24$  であるので、表 A.1 よりステューデント化された範囲の 5% 点の値は 3.532 と読める。よって、閾値  $q_0 = 3.532/\sqrt{2} = 2.498$  であり、設定 1, 3 間に有意な差があったと結論づけられる。

|    | A                       | B                   | C           | D           | E            | F         |
|----|-------------------------|---------------------|-------------|-------------|--------------|-----------|
| 1  | 平均値の差の検定(テューキー・クレマーの方法) |                     |             |             |              |           |
| 2  |                         |                     |             |             |              |           |
| 3  | 入力                      | ステューデント化された範囲の上側5%点 |             |             |              |           |
| 4  |                         | 3.532               |             |             |              |           |
| 5  |                         |                     |             |             |              |           |
| 6  |                         | 設定1                 | 設定2         | 設定3         |              |           |
| 7  |                         | 3.3                 | 3.2         | 3.1         |              |           |
| 8  |                         | 3.2                 | 3.1         | 3.0         |              |           |
| 9  |                         | 3.4                 | 3.3         | 3.3         |              |           |
| 10 |                         | 3.3                 | 3.2         | 3.1         |              |           |
| 11 |                         | 3.3                 | 3.2         | 3.2         |              |           |
| 12 |                         | 3.2                 | 3.1         | 3.0         |              |           |
| 13 |                         | 3.5                 | 3.4         | 3.2         |              |           |
| 14 |                         | 3.2                 | 3.0         |             |              |           |
| 15 |                         | 3.2                 | 3.2         |             |              |           |
| 16 |                         | 3.4                 |             |             |              |           |
| 17 |                         | 3.5                 |             |             |              |           |
| 18 |                         |                     |             |             |              |           |
| 19 | 計算値                     | 設定数 a               | データ数 n1     | n2          | n3           | 自由度 $\nu$ |
| 20 |                         | 3                   | 11          | 9           | 7            | 24        |
| 21 |                         | 平均 $\bar{x}_1$      | $\bar{x}_2$ | $\bar{x}_3$ |              |           |
| 22 |                         | 3.32                | 3.19        | 3.13        |              |           |
| 23 |                         | 不偏分散 $v_1^2$        | $v_2^2$     | $v_3^2$     | 平均平方 $v_E^2$ | 閾値 $q_0$  |
| 24 |                         | 0.01364             | 0.01361     | 0.01238     | 0.0133       | 2.498     |
| 25 | 出力                      | 検定統計量 $q_{12}$      | $q_{13}$    | $q_{23}$    |              |           |
| 26 |                         | 2.49                | 3.40        | 1.04        |              |           |

図 8.18: 母平均の差の検定 (データ数が異なる場合, テューキー・クレマーの方法)  
(テューキー・クレマーによる検定.xlsx)

### 8.2.2 テューキー・クレマーの方法のシミュレーション

テューキー・クレマーの方法により、各群のデータ数が異なる場合の母平均の差の検定のシミュレーションを行う。図 8.19 は、データ数が 11 個, 9 個, 7 個の場合におい

て、それぞれ、平均  $\mu_1 = \mu_2 = \mu_3 = 1$ 、分散  $\sigma_1^2 = \sigma_2^2 = \sigma_3^2 = 1$  の正規乱数を生成させ、検定統計量の  $q$  値を式 (8.45) により求めて、式 (8.48) の仮説の下での検定を 1000 組について繰り返したシミュレーションの画面の一部である。セル C6 にて調整  $q_0$  値が手作業により入力されている。3.532 ( $= \sqrt{2}q_0$ ) の値は表 A.1 のステューデント化された範囲においてデータ群数  $a = 3$ 、自由度  $\nu = 24$  に対応する値である。図 8.20 は以上の 1000 組のシミュレーションを 100 回繰り返したときの、第 1 種の過誤の確率の 95% 信頼区間を示す。各群のデータ数が異なる場合にも第 1 種の過誤の真の確率  $p^*$  をほぼ 5% とする結果が得られている。

|    | A   | B                | C        | D                  | E          | F          | G          |
|----|---|------------------|----------|--------------------|------------|------------|------------|
| 1  | 平均値の差の検定のシミュレーション(各群のデータ数が異なる場合(テューキー・クレマーの方法)) |                  |          |                    |            |            |            |
| 2  |   |                  |          |                    |            |            |            |
| 3  | 入力  | $\mu_1$          | $\mu_2$  | $\mu_3$            | $\sigma_1$ | $\sigma_2$ | $\sigma_3$ |
| 4  |   | 1                | 1        | 1                  | 1          | 1          | 1          |
| 5  |   | 公称の有意水準 $\alpha$ | 調整 $q_0$ |                    |            |            |            |
| 6  |   | 0.05             | 2.4975   | $= 3.532/\sqrt{2}$ |            |            |            |
| 7  |   |                  |          |                    |            |            |            |
| 8  | 計算値   | 1組目              |          |                    | 2組目        |            |            |
| 9  |   | X1の乱数            | X2の乱数    | X3の乱数              | X1の乱数      | X2の乱数      | X3の乱数      |
| 10 |   | 0.492            | -0.179   | 1.715              | 1.870      | 1.064      | 1.758      |
| 11 |   | 1.417            | 1.901    | 1.147              | 1.435      | 1.274      | 2.213      |

|    |    |                                      |                     |                     |                                      |                     |                     |
|----|----|--------------------------------------|---------------------|---------------------|--------------------------------------|---------------------|---------------------|
| 28 |    | 検定統計量 $q_{12}$                       | $q_{13}$            | $q_{23}$            | $q_{12}$                             | $q_{13}$            | $q_{23}$            |
| 29 |    | 0.548                                | 0.094               | 0.399               | -0.270                               | 0.340               | -0.567              |
| 30 |    | if $ q_{12}  > q_0$<br>then 1 else 0 | if $ q_{13}  > q_0$ | if $ q_{23}  > q_0$ | if $ q_{12}  > q_0$<br>then 1 else 0 | if $ q_{13}  > q_0$ | if $ q_{23}  > q_0$ |
| 31 |    | 0                                    | 0                   | 0                   | 0                                    | 0                   | 0                   |
| 32 |    | 各対のいずれかで有意差あり                        |                     |                     | 各対のいずれかで有意差あり                        |                     |                     |
| 33 |    | 0                                    |                     |                     | 0                                    |                     |                     |
| 34 | 出力 | 有意頻度 $fr$                            |                     |                     |                                      |                     |                     |
| 35 |    | 0.047                                |                     |                     |                                      |                     |                     |

図 8.19: 母平均の差の検定のシミュレーション (各群のデータ数相違, テューキー・クレマーの方法)

(テューキー・クレマーによるシミュレーション.xlsxm)

|    |        |             |         |                 |        |
|----|--------|-------------|---------|-----------------|--------|
| 53 |        | 有意頻度 $fr_k$ | 100回の平均 |                 |        |
| 54 | $fr_1$ | 0.046       | 0.0494  |                 |        |
| 55 | $fr_2$ | 0.049       | 95%信頼区間 |                 |        |
| 56 | $fr_3$ | 0.051       | 0.0480  | $\leq p^* \leq$ | 0.0508 |

図 8.20: 母平均の差の検定, テューキー・クレマーの方法における第 1 種の過誤の確率の 95% 信頼区間

## 8.2.3 テューキー・クレーマーの方法の理論 [23] [24]

式 (8.48) の帰無仮説は

$$\max_{i,j=1,2,3,i \neq j} |q_{ij}| \leq q_0 \quad (8.49)$$

もしくは,

$$|q_{12}| \leq q_0 \quad \text{and} \quad |q_{13}| \leq q_0 \quad \text{and} \quad |q_{23}| \leq q_0 \quad (8.50)$$

と表現できる. いま,  $\bar{x}_1 \geq \bar{x}_2, \bar{x}_3$  の場合を考える. このとき

$$\begin{aligned} 0 &\leq \frac{\bar{x}_1 - \bar{x}_2}{\sqrt{v_E^2 \left( \frac{1}{n_1} + \frac{1}{n_2} \right)}} \leq q_0 \\ 0 &\leq \frac{\bar{x}_1 - \bar{x}_3}{\sqrt{v_E^2 \left( \frac{1}{n_1} + \frac{1}{n_3} \right)}} \leq q_0 \end{aligned} \quad (8.51)$$

の範囲を考える. 文献 [17] には, この範囲にある  $\bar{x}_1, \bar{x}_2, \bar{x}_3$  が (8.50) 式を満たすと記されている. これに対して文献 [23] [24] にて, **上式の範囲にある  $\bar{x}_1, \bar{x}_2, \bar{x}_3$  が (8.50) 式を満たさない場合がある**ことが指摘されている. 文献 [23] [24] にならい, 具体例を紹介する.

(例 8.1) データ群数  $a = 3$ , データ数  $n_1 = 10, n_2 = 20, n_3 = 30$  とする. また,  $\bar{x}_1 = 1.2, \bar{x}_2 = -1.7, \bar{x}_3 = 0.8, v_E^2 = 60$  であったとする. 簡単のため  $q_0 = 1$  としても一般性を失わない. このとき,

$$\begin{aligned} 0 &\leq \frac{\bar{x}_1 - \bar{x}_2}{\sqrt{v_E^2 \left( \frac{1}{n_1} + \frac{1}{n_2} \right)}} = 0.97 \leq q_0 \\ 0 &\leq \frac{\bar{x}_1 - \bar{x}_3}{\sqrt{v_E^2 \left( \frac{1}{n_1} + \frac{1}{n_3} \right)}} = 0.14 \leq q_0 \end{aligned} \quad (8.52)$$

であり,  $\bar{x}_1, \bar{x}_2, \bar{x}_3$  は (8.51) 式を満たしている. しかし,

$$\frac{\bar{x}_2 - \bar{x}_3}{\sqrt{v_E^2 \left( \frac{1}{n_2} + \frac{1}{n_3} \right)}} = 1.12 > q_0 \quad (8.53)$$

であり, (8.50) 式を満たしていない.

この例では  $n_1 < n_2, n_3$  のときに, (8.51) 式では考慮していない  $\bar{x}_2 - \bar{x}_3$  の項が  $q_0$  を超えてしまっている.  $\bar{x}_1, \bar{x}_2, \bar{x}_3$  が (8.51) 式を満たしていても, 第1種の過誤が起きる場合があることを示している.

(8.51) 式において、仮定の  $\mu_1 = \mu_2 = \mu_3$  を挿入して、変形すると

$$\begin{aligned} \sqrt{\frac{n_2 \bar{x}_1 - \mu_1}{n_1} \frac{\sigma^2}{n_1}} - \sqrt{n_2} \sqrt{\frac{1}{n_1} + \frac{1}{n_2} \frac{v_E}{\sigma}} q_0 &\leq \frac{\bar{x}_2 - \mu_2}{\sqrt{\frac{\sigma^2}{n_2}}} \leq \sqrt{\frac{n_2 \bar{x}_1 - \mu_1}{n_1} \frac{\sigma^2}{n_1}} \\ \sqrt{\frac{n_3 \bar{x}_1 - \mu_1}{n_1} \frac{\sigma^2}{n_1}} - \sqrt{n_3} \sqrt{\frac{1}{n_1} + \frac{1}{n_3} \frac{v_E}{\sigma}} q_0 &\leq \frac{\bar{x}_3 - \mu_3}{\sqrt{\frac{\sigma^2}{n_3}}} \leq \sqrt{\frac{n_3 \bar{x}_1 - \mu_1}{n_1} \frac{\sigma^2}{n_1}} \end{aligned} \quad (8.54)$$

となる。式 (3.67) より、

$$\frac{\bar{X}_i - \mu_i}{\sqrt{\frac{\sigma^2}{n_i}}} \sim N(0, 1) \quad (i = 1, 2, 3) \quad (8.55)$$

であり、事象  $\sqrt{n_i/\sigma^2}(\bar{X}_i - \mu_i)$  は互いに独立に標準正規分布に従う。また、事象  $S = V_E/\sigma$  は式 (8.38) において、 $\nu = n_1 + n_2 + n_3 - 3$  としたときの確率密度関数  $h_{n_1+n_2+n_3-3}(s)$  の分布に従う。これらの事象が互いに独立であるとして、 $\bar{x}_1 \geq \bar{x}_2, \bar{x}_3$  の場合に (8.51) 式を満たす確率  $P_1$  は、 $\bar{x}_1 = x_{10}$  のとき、 $w_{10} = \sqrt{n/\sigma^2}(x_{10} - \mu_1)$  とし、また、 $s = v_E/\sigma$  とすると

$$\begin{aligned} P_1 = \int_0^\infty \int_{-\infty}^\infty \prod_{j=2,3} \left\{ \Phi \left( \sqrt{\frac{n_j}{n_1}} w_{10} \right) - \Phi \left( \sqrt{\frac{n_j}{n_1}} w_{10} - \sqrt{\frac{n_j}{n_1} + 1} s q_0 \right) \right\} \\ \times f(w_{10}) h_{n_1+n_2+n_3-3}(s) dw_{10} ds \end{aligned} \quad (8.56)$$

と求められる。 $\bar{x}_2 \geq \bar{x}_1, \bar{x}_3$ ,  $\bar{x}_3 \geq \bar{x}_1, \bar{x}_2$  の各場合の確率  $P_2, P_3$  を同様に求めることができる。

$P_1, P_2, P_3$  には例 8.1 のように第 1 種の過誤が起きる確率が含まれていることがある。したがって、 $1 - P_1 - P_2 - P_3$  は第 1 種の過誤の確率  $P(\max|q_{ij}| > q_0)$  以下となる。

$$P(\max|q_{ij}| > q_0) \geq 1 - P_1 - P_2 - P_3. \quad (8.57)$$

等号が成立するのは  $n_1 = n_2 = n_3$  のときである [23] [24].

データ群が  $a$  個の場合に拡張する。検定の仮説は、データ群  $i$  の母平均を  $\mu_i$  とすると

$$\begin{aligned} \text{帰無仮説} &: \mu_i = \mu_j \\ \text{対立仮説} &: \mu_i \neq \mu_j \\ &\text{ただし, } i, j = 1, \dots, a, i \neq j \end{aligned} \quad (8.58)$$

である。ここで、データ群  $i$  のデータ数を  $n_i$ 、自由度を  $\nu = \sum_k (n_k - 1)$  とし、定数  $\lambda_{ij}$  を

$$\lambda_{ij} = \frac{n_j}{n_i + n_j} \quad (8.59)$$

とする.

$$P_a = 1 - \sum_{i=1}^a \int_0^\infty \int_{-\infty}^\infty \prod_{j(\neq i)}^a \left\{ \Phi \left( \frac{\sqrt{\lambda_{ij}}}{\sqrt{1-\lambda_{ij}}} w \right) - \Phi \left( \frac{\sqrt{\lambda_{ij}} w - s q_0}{\sqrt{1-\lambda_{ij}}} \right) \right\} \times f(w) h_\nu(s) dw ds \quad (8.60)$$

とすると, 第1種の過誤の起きる確率  $P(\max|q_{ij}| > q_0)$  との関係は

$$P(\max|q_{ij}| > q_0) \geq P_a \quad (8.61)$$

となる.

テューキー・クレーマーの方法では,  $q_0$  として表 A.1 のステューデント化された範囲 ( $q_{0st'ed}$  と記す) を利用している. ステューデント化された範囲により, 第1種の過誤の確率  $P(\max|q_{ij}| > q_0)$  を公称の有意水準  $\alpha$  以下とできることがヘイター (Hayter) [18] により示されている. すなわち,

$$\alpha = P(\max|q_{ij}| > q_{0st'ed}) \geq P(\max|q_{ij}| > q_0) \geq P_a \quad (8.62)$$

である. 等号が成り立つのは  $n_i = n_j = n$  のときである [23] [24].

図 8.21 は, テューキー・クレーマーの方法による検定において (8.60) 式の  $P_a$  が 0.05 となる  $q_0$  を閾値とした場合のシミュレーション画面の一部である. 3群で, 各群のデータ数を  $n_1 = 10, n_2 = 20, n_3 = 30$  とした場合である. (8.60) 式による閾値は緩く, 1000 組の中で第1種の過誤が起きる頻度は公称の有意水準の 5% より高い値が出る傾向にある. 図 8.22 は 1000 組のシミュレーションを 100 回繰り返して, 第1種の過誤の確率の 95% 信頼区間を求めた結果である. (8.62) 式の通り,  $P_a$  が 5% となるように閾値を設定すると, 第1種の過誤の確率  $P(\max|q_{ij}| > q_0)$  は 5% よりも大きな値となる. 図 8.23 は, 図 8.21 と同じ設定において, 閾値  $q_0$  のみをステューデント化された範囲に変更した場合について第1種の過誤の確率の 95% 信頼区間を求めた結果である. 各群のデータ数が  $n_1 = 10, n_2 = 20, n_3 = 30$  のときでも, ステューデント化された範囲を閾値とすることで第1種の過誤の確率を 5% とする結果が得られている. この 100 回のシミュレーションを何度も繰り返すと, 第1種の過誤の確率が 5% よりも低く抑えられていることを確認できる.



|    | A  | B                              | C             | D             | E             | F             | G             |
|----|--|--------------------------------|---------------|---------------|---------------|---------------|---------------|
| 1  | 平均値の差の検定のシミュレーション(各群のデータ数が異なる場合(テューキー・クレーマーの方法)) |                                |               |               |               |               |               |
| 2  |  |                                |               |               |               |               |               |
| 3  | 入力   | $\mu 1$                        | $\mu 2$       | $\mu 3$       | $\sigma 1$    | $\sigma 2$    | $\sigma 3$    |
| 4  |  | 1                              | 1             | 1             | 1             | 1             | 1             |
| 5  |  | 公称の有意水準 $\alpha$               | 調整 $q_0$      |               |               |               |               |
| 6  |  | 0.05                           | 2.387         |               |               |               |               |
| 7  |  |                                |               |               |               |               |               |
| 8  | 計算値  | 1組目                            |               |               | 2組目           |               |               |
| 9  |  | X1の乱数                          | X2の乱数         | X3の乱数         | X1の乱数         | X2の乱数         | X3の乱数         |
| 10 |  | -0.393                         | 0.240         | 1.359         | -0.066        | -0.705        | 0.537         |
| 11 |  | 2.695                          | 0.133         | 1.925         | -0.347        | 0.036         | 0.309         |
| 41 |  | データ群数 a                        | データ数 n1       | n2            | n3            | 組数 G          |               |
| 42 |  | 3                              | 10            | 20            | 30            | 1000          |               |
| 43 |  |                                |               |               |               |               |               |
| 44 |  | 平均X1                           | X2            | X3            | X1            | X2            | X3            |
| 45 |  | 0.931                          | 1.165         | 1.040         | 0.967         | 0.986         | 1.185         |
| 46 |  | 平均平方 $v_g^2$                   |               |               | 平均平方 $v_g^2$  |               |               |
| 47 |  | 1.210                          |               |               | 1.176         |               |               |
| 48 |  | 検定統計量 q12                      | q13           | q23           | q12           | q13           | q23           |
| 49 |  | -0.551                         | -0.271        | -0.396        | -0.046        | -0.551        | 0.636         |
| 50 |  | if  q12  > q0<br>then 1 else 0 | if  q13  > q0 | if  q23  > q0 | if  q12  > q0 | if  q13  > q0 | if  q23  > q0 |
| 51 |  | 0                              | 0             | 0             | 0             | 0             | 0             |
| 52 |  | 各対のいずれかで有意差あり                  |               |               | 各対のいずれかで有意差あり |               |               |
| 53 |  | 0                              |               |               | 0             |               |               |
| 54 | 出力   | 有意頻度 fr                        |               |               |               |               |               |
| 55 |  | 0.055                          |               |               |               |               |               |

図 8.21: 母平均の差の検定のシミュレーション ( $n_1 = 10, n_2 = 20, n_3 = 30$ , テューキー・クレーマーの方法,  $P_a = 0.05$  とする閾値の場合) (テューキー・クレーマーによるシミュレーション\_n.10.20.30.xlsm)

|    |     |          |         |                 |        |
|----|-----|----------|---------|-----------------|--------|
| 61 |     | 有意頻度 frk | 100回の平均 |                 |        |
| 62 | fr1 | 0.05     | 0.0517  |                 |        |
| 63 | fr2 | 0.051    | 95%信頼区間 |                 |        |
| 64 | fr3 | 0.044    | 0.0503  | $\leq p^* \leq$ | 0.0532 |

図 8.22: 母平均の差の検定, テューキー・クレーマーの方法 ( $n_1 = 10, n_2 = 20, n_3 = 30$ ,  $P_a = 0.05$  とする閾値の場合) における第1種の過誤の確率の95%信頼区間

|    | A   | B                | C        | D                 | E               | F          | G          |
|----|---|------------------|----------|-------------------|-----------------|------------|------------|
| 1  | 平均値の差の検定のシミュレーション (各群のデータ数が異なる場合 (テューキー・クレマーの方法)) |                  |          |                   |                 |            |            |
| 2  |   |                  |          |                   |                 |            |            |
| 3  | 入力  | $\mu 1$          | $\mu 2$  | $\mu 3$           | $\sigma 1$      | $\sigma 2$ | $\sigma 3$ |
| 4  |   | 1                | 1        | 1                 | 1               | 1          | 1          |
| 5  |   | 公称の有意水準 $\alpha$ | 調整 $q_0$ |                   |                 |            |            |
| 6  |   | 0.05             | 2.406    | 3.403/ $\sqrt{2}$ |                 |            |            |
| 61 |   | 有意頻度 frk         |          | 100回の平均           |                 |            |            |
| 62 | fr1   | 0.047            |          | 0.0492            |                 |            |            |
| 63 | fr2   | 0.033            |          | 95%信頼区間           |                 |            |            |
| 64 | fr3   | 0.058            |          | 0.0479            | $\leq p^* \leq$ | 0.0506     |            |

図 8.23: 母平均の差の検定, テューキー・クレマーの方法 ( $n_1 = 10, n_2 = 20, n_3 = 30$ , 閾値をステューデント化された範囲とした場合) における第1種の過誤の確率の95%信頼区間

### 8.3 母平均の差の検定 (母分散が異なる場合)

#### 8.3.1 ゲイムズ・ハウウェルの方法による検定の例

ある日, ある工場のある製品のラインにおいて, ラインのスピードを3通りに変えながらあるパラメータの測定がなされたとする. スピード設定1, 2, 3からそれぞれ11, 9, 7個の標本を取り出すことができ, 以下の値が得られたとする.

設定1: 3.3, 3.2, 3.4, 3.3, 3.3, 3.2, 3.5, 3.2, 3.2, 3.4, 3.5

設定2: 3.2, 3.1, 3.3, 3.2, 3.2, 3.1, 3.4, 3.0, 3.2

設定3: 3.1, 3.0, 3.3, 3.1, 3.1, 2.9, 3.4

ライン1, 2, 3の平均値  $\bar{x}_1, \bar{x}_2, \bar{x}_3$  はそれぞれ3.32, 3.19, 3.13であった. スピード設定を変えたことで母平均に差が出たと言えるか? ただし, いずれの設定の母標準偏差も分かっていないとする.

検定の仮説は, データ群  $i$  の母平均を  $\mu_i$  とすると

帰無仮説 :  $\mu_i = \mu_j$

対立仮説 :  $\mu_i \neq \mu_j$

ただし,  $i, j = 1, 2, 3, i \neq j$

(8.63)

である。

テューキーの方法およびテューキー・クレーマーの方法は各群の母分散が等しいことを前提とした多重比較法である。本節の例では設定  $i$  の不偏分散を  $v_{ei}^2$  とすると、 $v_{e1}^2 = 0.0136$ ,  $v_{e2}^2 = 0.0136$ ,  $v_{e3}^2 = 0.029$  であり異なっている。母分散が等しいことを前提とした多重比較法では第1種の過誤の確率を5%以下に抑える保証はない。なお、母分散を等しいとみなせるかどうかについては次章の多群の場合の母分散の検定を行わなければならない。

母分散が等しいとみなせない場合には、ウェルチ (Welch) の検定統計量を用いた **ゲームズ・ハウエル (Games Howell [19]) の方法** を適用する。データ群  $i$  の平均値を  $\bar{x}_i$  とすると、データ群  $i, j$  間の検定統計量

$$q_{ij} = \frac{\bar{x}_i - \bar{x}_j}{\sqrt{\frac{v_{ei}^2}{n_i} + \frac{v_{ej}^2}{n_j}}} \quad (8.64)$$

が近似的に等価自由度  $\nu^*$  の  $t$  分布に従うことを利用する。等価自由度  $\nu^*$  は

$$\nu^* = \frac{\left(\frac{v_{ei}^2}{n_i} + \frac{v_{ej}^2}{n_j}\right)^2}{\frac{1}{n_i-1} \left(\frac{v_{ei}^2}{n_i}\right)^2 + \frac{1}{n_j-1} \left(\frac{v_{ej}^2}{n_j}\right)^2} \quad (8.65)$$

により与えられる。等価自由度は小数点以下を四捨五入して用いる。ステューデント化された範囲の数表において、データ群数  $a$ 、自由度  $\nu^*$  の値を読みこれを閾値とする。

図 8.24 はゲームズ・ハウエルの方法により平均値の差の検定を実行している Excel の画面である。セル B25 では設定 1, 2 間の等価自由度  $\nu_{12}^*$  を式 (8.65) により求めている。ROUND( $x, y$ ) 関数において  $y = 0$  とすることで、 $x$  の小数点以下を四捨五入している。セル B27 では設定 1, 2 間の検定統計量  $q_{12}$  を式 (8.64) により求めている。セル B5 では、表 A.1 のステューデント化された範囲より設定数  $a = 3$  と等価自由度  $\nu^* = 17$  に対応する値 3.628 が手作業により入力されている。C 列では設定 1, 3 間の検定統計量  $q_{13} = 2.58$  に対して、等価自由度  $\nu^* = 10$  をもとに数表から手入力された値より閾値  $q_{0.13} = 2.741$  が得られている。D 列は設定 2, 3 間の比較結果である。いずれの設定間にも有意差はなく、帰無仮説は棄却されない。

図 8.25 は母分散に差がないとはみなせないにもかかわらずテューキー・クレーマーの方法を適用した場合の Excel の画面である。図 8.18 の Excel の画面において、データのみを本項のものに入れ替えただけである。設定 1, 3 間に有意差のある結果が得られている。母分散に差がある場合にテューキー・クレーマーの方法を適用したことで第1種の過誤の確率が5%より大きくなってしまったと推測される。次項のシミュレーションにてより詳細な比較を行う。

|    | A                                 | B                   | C       | D     | E         | F      | G      |
|----|-----------------------------------|---------------------|---------|-------|-----------|--------|--------|
| 1  | 母平均の差の検定 (母分散に差あり, ゲイムズ・ハウウェルの方法) |                     |         |       |           |        |        |
| 2  |                                   |                     |         |       |           |        |        |
| 3  | 入力                                | ステューデント化された範囲の上側5%点 |         |       |           |        |        |
| 4  |                                   | q0.12               | q0.13   | q0.23 |           |        |        |
| 5  |                                   | 2.565               | 2.741   | 2.741 |           |        |        |
| 6  |                                   | 設定1                 | 設定2     | 設定3   |           |        |        |
| 8  |                                   | 3.3                 | 3.2     | 3.1   |           |        |        |
| 9  |                                   | 3.2                 | 3.1     | 3.0   |           |        |        |
| 10 |                                   | 3.4                 | 3.3     | 3.3   |           |        |        |
| 11 |                                   | 3.3                 | 3.2     | 3.1   |           |        |        |
| 12 |                                   | 3.3                 | 3.2     | 3.1   |           |        |        |
| 13 |                                   | 3.2                 | 3.1     | 2.9   |           |        |        |
| 14 |                                   | 3.5                 | 3.4     | 3.4   |           |        |        |
| 15 |                                   | 3.2                 | 3.0     |       |           |        |        |
| 16 |                                   | 3.2                 | 3.2     |       |           |        |        |
| 17 |                                   | 3.4                 |         |       |           |        |        |
| 18 |                                   | 3.5                 |         |       |           |        |        |
| 19 |                                   |                     |         |       |           |        |        |
| 20 | 計算値                               | 設定数 a               | データ数 n1 | n2    | n3        |        |        |
| 21 |                                   | 3                   | 11      | 9     | 7         |        |        |
| 22 |                                   | 平均x1                | x2      | x3    | 不偏分散ve1^2 | ve2^2  | ve3^2  |
| 23 |                                   | 3.32                | 3.19    | 3.13  | 0.0136    | 0.0136 | 0.0290 |
| 24 |                                   | 等価自由度ν*12           | ν*13    | ν*23  |           |        |        |
| 25 |                                   | 17                  | 10      | 10    |           |        |        |
| 26 | 出力                                | 検定統計量 q12           | q13     | q23   |           |        |        |
| 27 |                                   | 2.46                | 2.58    | 0.80  |           |        |        |

$= 3.628/\sqrt{2}$  (設定1)       $= 3.877/\sqrt{2}$  (設定3)

$$\text{ROUND}\left(\frac{(v_a^2/n_1 + v_b^2/n_2)^2}{(v_a^2/n_1)/(n_1-1) + (v_b^2/n_2)/(n_2-1)}, 0\right)$$

$$= |x_1 - x_2| / \sqrt{v_{e1}^2/n_1 + v_{e2}^2/n_2}$$

図 8.24: 母平均の差の検定 (母分散に差あり, ゲイムズ・ハウウェルの方法) (ゲイムズ・ハウウェルによる検定.xlsx)

|    | A                                     | B                   | C       | D       | E        | F     |
|----|---------------------------------------|---------------------|---------|---------|----------|-------|
| 1  | 母平均の差の検定 (母分散に差があるのにテューキー・クレマーの方法を適用) |                     |         |         |          |       |
| 2  |                                       |                     |         |         |          |       |
| 3  | 入力                                    | ステューデント化された範囲の上側5%点 |         |         |          |       |
| 4  |                                       | 3.532               |         |         |          |       |
| 5  |                                       |                     |         |         |          |       |
| 6  |                                       | 設定1                 | 設定2     | 設定3     |          |       |
| 7  |                                       | 3.3                 | 3.2     | 3.1     |          |       |
| 8  |                                       | 3.2                 | 3.1     | 3.0     |          |       |
| 19 | 計算値                                   | 設定数 a               | データ数 n1 | n2      | n3       | 自由度ν  |
| 20 |                                       | 3                   | 11      | 9       | 7        | 24    |
| 21 |                                       | 平均x1                | x2      | x3      |          |       |
| 22 |                                       | 3.31                | 3.19    | 3.13    |          |       |
| 23 |                                       | 不偏分散ve1^2           | ve2^2   | ve3^2   | 平均平方vE^2 | 閾値 q0 |
| 24 |                                       | 0.01091             | 0.01361 | 0.02905 | 0.0163   | 2.498 |
| 25 | 出力                                    | 検定統計量 q12           | q13     | q23     |          |       |
| 26 |                                       | 2.09                | 2.92    | 0.94    |          |       |

図 8.25: 母平均の差の検定 (母分散に差があるのにテューキー・クレマーの方法を適用) (テューキー・クレマーによる検定 (母分散に差あり) .xlsx)

### 8.3.2 ゲイムズ・ハウウェルの方法のシミュレーション

ゲイムズ・ハウウェルの方法により, 母分散に差がある場合の平均値の差の検定のシミュレーションを行う. 図 8.26 は, 図 8.19 の設定においてデータ群3の母標準偏差  $\sigma_3$  の

みを1から3へと変更して、1000組のシミュレーションを100回繰り返して、第1種の過誤の確率の95%信頼区間を求めた結果である。前項の図8.25の結果と同様に、母分散に差がある場合の平均値の差の検定にテューキー・クレーマーの方法を適用することで第1種の過誤の確率を大きくしてしまう結果となった。

|   | A  | B                | C        | D       | E          | F          | G          |
|---|--|------------------|----------|---------|------------|------------|------------|
| 1 | 母平均の差の検定のシミュレーション(母分散に差があるのにテューキー・クレーマーの方法を適用) |                  |          |         |            |            |            |
| 2 |  |                  |          |         |            |            |            |
| 3 | 入力   | $\mu 1$          | $\mu 2$  | $\mu 3$ | $\sigma 1$ | $\sigma 2$ | $\sigma 3$ |
| 4 |  | 1                | 1        | 1       | 1          | 1          | 3          |
| 5 |  | 公称の有意水準 $\alpha$ | 調整 $q_0$ |         |            |            |            |
| 6 |  | 0.05             | 2.4975   |         |            |            |            |

|    |     |          |         |                 |        |  |
|----|-----|----------|---------|-----------------|--------|--|
| 53 |     | 有意頻度 frk | 100回の平均 |                 |        |  |
| 54 | fr1 | 0.129    | 0.1254  |                 |        |  |
| 55 | fr2 | 0.157    | 95%信頼区間 |                 |        |  |
| 56 | fr3 | 0.125    | 0.1234  | $\leq p^* \leq$ | 0.1273 |  |

図 8.26: 母平均の差の検定のシミュレーションと第1種の過誤の確率の95%信頼区間(母分散に差があるのにテューキー・クレーマーの方法を適用)(テューキー・クレーマーによるシミュレーション(母分散に差あり).xlsm)

図8.27は、ゲイズ・ハウウェルの方法によるシミュレーションの画面である。等価自由度 $\nu^*$ および検定統計量 $q$ をそれぞれ式(8.65)、(8.64)により求めている。セルB34では次の関数により閾値 $q_0$ 値を自動で得ている。

$$= \text{INDIRECT}(\text{ADDRESS}(\nu_{12}^*, a, , , \text{"Sheet2"})) / \sqrt{2} \quad (8.66)$$

図8.28は以上の1000組のシミュレーションを100回繰り返して得られた第1種の過誤の確率の95%信頼区間である。ゲイズ・ハウウェルの方法により第1種の過誤の確率を5%辺りとする結果が得られている。

|    | A   | B                | C       | D       | E          | F          | G          |
|----|---|------------------|---------|---------|------------|------------|------------|
| 1  | 母平均の差の検定のシミュレーション(母分散に差あり, ゲイムズ・ハウウェルの方法) |                  |         |         |            |            |            |
| 2  |   |                  |         |         |            |            |            |
| 3  | 入力  | $\mu 1$          | $\mu 2$ | $\mu 3$ | $\sigma 1$ | $\sigma 2$ | $\sigma 3$ |
| 4  |   | 1                | 1       | 1       | 1          | 1          | 3          |
| 5  |   | 公称の有意水準 $\alpha$ |         |         |            |            |            |
| 6  |   | 0.05             |         |         |            |            |            |
| 7  |   |                  |         |         |            |            |            |
| 8  |   | 1組目              |         |         | 2組目        |            |            |
| 9  |   | x1               | x2      | x3      | x1         | x2         | x3         |
| 10 |   | 0.622            | 0.450   | -0.638  | 1.227      | 1.643      | 0.702      |
| 11 |   | 1.511            | 0.119   | 1.548   | 0.284      | 2.117      | -1.726     |
| 12 |   | 0.556            | 1.102   | -0.725  | -0.261     | 0.053      | -3.933     |
| 22 |   | データ群数 a          | データ数 n1 | n2      | n3         | 組数 G       |            |
| 23 |   | 3                | 11      | 9       | 7          | 1000       |            |
| 29 |   | 等価自由度 v*_12      | v*_13   | v*_23   | v*_12      | v*_13      | v*_23      |
| 30 |   | 17               | 8       | 9       | 13         | 6          | 7          |
| 31 |   | 検定統計量 q12        | q13     | q23     | q12        | q13        | q23        |
| 32 |   | 0.01             | 2.06    | 2.01    | 1.12       | 2.07       | 1.74       |
| 33 |   | 閾値 q0_12         | q0_13   | q0_23   | q0_12      | q0_13      | q0_23      |
| 34 |   | 2.565            | 2.857   | 2.792   | 2.640      | 3.068      | 2.945      |

$= \text{INDIRECT}(\text{ADDRESS}(v_{12}^*, a_{,,}, \text{"Sheet2"})) / \sqrt{2}$

図 8.27: 母平均の差の検定のシミュレーション (母分散に差あり, ゲイムズ・ハウウェルの方法) (ゲイムズ・ハウウェルによるシミュレーション.xlsm)

|    |     |          |         |                 |        |
|----|-----|----------|---------|-----------------|--------|
| 53 |     | 有意頻度 frk | 100回の平均 |                 |        |
| 54 | fr1 | 0.052    | 0.0498  |                 |        |
| 55 | fr2 | 0.051    | 95%信頼区間 |                 |        |
| 56 | fr3 | 0.059    | 0.0485  | $\leq p^* \leq$ | 0.0511 |

図 8.28: 母平均の差の検定, ゲイムズ・ハウウェルの方法における第1種の過誤の確率の95%信頼区間

## 8.4 母平均の差の検定（対照群と処理群の間の検定）

検定の課題には、全ての群間の平均値の差だけでなく、例えば、従来薬により得られたデータの一群（対照群）と各種の新薬により得られた複数のデータ群（処理群）との差を検定したい場合もある。もしくは、改良案により得られたデータの一群を種々の従来法から得られた複数のデータ群と比較して、改良案の有効性を示したい場合もある。そのような場合には、これまでの複数群間の全組み合わせを調べる多重比較法では調整  $q_0$  値を厳しくしすぎてしまうことがある。本節ではダネット (Dunnett) の方法を紹介する。なお、この方法は前提条件として母分散が等しいとしている。

### 8.4.1 ダネットの方法による検定の例

ある日、ある工場のある製品のラインにおいて、ラインのスピードを新たに2通りに変えながらあるパラメータの測定をした。従来設定（設定1とする）および新しい設定（設定2, 3とする）からそれぞれ9個ずつの標本を取り出すことができ、以下の値が得られたとする。

設定1: 3.5, 3.3, 3.3, 3.4, 3.3, 3.2, 3.5, 3.4, 3.2

設定2: 3.2, 3.2, 3.4, 3.2, 3.2, 3.1, 3.4, 3.1, 3.2

設定3: 3.3, 3.2, 3.4, 3.3, 3.2, 3.4, 3.3, 3.2, 3.1

ライン1, 2, 3の平均値  $\bar{x}_1, \bar{x}_2, \bar{x}_3$  はそれぞれ 3.34, 3.22, 3.27 であった。新しいスピード設定はこのパラメータの値を変えることに効果があったと言えるか？ただし、いずれの設定の母分散も分かっていないとする。

このとき検定の仮説は、設定  $i$  の母平均を  $\mu_i$  とすると

$$\begin{aligned} \text{帰無仮説} &: \mu_1 = \mu_j \\ \text{対立仮説} &: \mu_1 \neq \mu_j \\ &\text{ただし, } j = 2, 3 \end{aligned} \tag{8.67}$$

である。

設定1のデータ群を対照群、設定2, 3のデータ群を処理群とよぶ。設定1,  $j(j = 2, 3)$



の平均値間の検定統計量  $q_{1j}$  は、式 (8.45) と同じであり、

$$q_{1j} = \frac{\bar{x}_1 - \bar{x}_j}{\sqrt{v_E^2 \left( \frac{1}{n_1} + \frac{1}{n_j} \right)}} \quad (8.68)$$

で与えられる。ただし、

$$v_E^2 = \frac{\sum_{k=1}^a \sum_{l=1}^{n_k} (x_{kl} - \bar{x}_k)^2}{\sum_{k=1}^a (n_k - 1)} \quad (8.69)$$

である。  $x_{kl}$  は設定  $k$  の第  $l$  番目の測定値、  $n_k$  は設定  $k$  のデータ数、  $a$  は設定の数である。この  $q_{1j}$  値は自由度

$$\sum_{k=1}^a (n_k - 1) \quad (8.70)$$

の  $t$  分布に従う。図 8.29 はダネットの方法により検定を実行している Excel の画面である。図 8.1 および図 8.18 との違いは 24 行目の検定統計量を設定 1, 2 間, 1, 3 間に限定して求めている点にある。検定の繰り返し回数が減ることで、 $q$  値の閾値の調整を緩めることができる。本項の検定は仮説より両側検定であるので、表 A.4 のダネット方法の 5% 点（両側検定）においてデータ群数  $a = 3$ 、自由度  $\nu = n_1 + n_2 + n_3 - a = 24$ 、データ数比  $\lambda_{1j} = 0.5$  より、閾値  $q_0 = 2.349$  と読める。ただし、データ数比  $\lambda_{1j}$  は

$$\lambda_{1j} = \frac{n_j}{n_1 + n_j} \quad (8.71)$$

により与えられる。ここで、 $n_j$  は新設定（設定 2, 3）のデータ数であり、設定 2, 3 間でデータ数は等しいことを前提としている。この閾値  $q_0$  がセル C4 に手作業で入力されている。24 行目の  $q_{12}, q_{13}$  値と比較すると、 $q_{12} > q_0, q_{13} < q_0$  であるので、設定 1, 2 間には有意な差があったと結論づけられる。なお、表 A.2～表 A.11 のダネットの方法の 5% 点は式 (8.68) をもとに作られているので、表の値に  $1/\sqrt{2}$  をかける必要はない。

設定数  $a = 3$ 、自由度  $\nu = 24$  のとき、テューキーの方法における閾値  $q_0 = 3.532/\sqrt{2} = 2.4975$  である。この閾値は 24 行目のいずれの  $q$  値よりも大きい。テューキーの方法ではいずれの設定値間にも有意差はないという結果となってしまう。これはテューキーの方法における閾値の調整が式 (8.67) の仮説に対しては厳しすぎることによる。

#### 8.4.2 ダネットの方法のシミュレーション

シミュレーションによりダネットの方法の有効性を確認する。図 8.30 はダネットの方法によるシミュレーションを行っている Excel の画面である。事象  $X_1$  のデータを対照群



|    | A                          | B                | C                | D        | E                    | F         |
|----|----------------------------|------------------|------------------|----------|----------------------|-----------|
| 1  | 母平均の差の検定(対照群との比較, ダネットの方法) |                  |                  |          |                      |           |
| 2  |                            |                  |                  |          |                      |           |
| 3  | 入力                         | 公称の有意水準 $\alpha$ | ダネットの方法のための両側5%点 |          |                      |           |
| 4  |                            | 0.05             | 2.349            |          |                      |           |
| 5  |                            |                  |                  |          |                      |           |
| 6  |                            | 設定1              | 設定2              | 設定3      |                      |           |
| 7  |                            | 3.5              | 3.2              | 3.3      |                      |           |
| 8  |                            | 3.3              | 3.2              | 3.2      |                      |           |
| 9  |                            | 3.3              | 3.4              | 3.4      |                      |           |
| 10 |                            | 3.4              | 3.2              | 3.3      |                      |           |
| 11 |                            | 3.3              | 3.2              | 3.2      |                      |           |
| 12 |                            | 3.2              | 3.1              | 3.4      |                      |           |
| 13 |                            | 3.5              | 3.4              | 3.3      |                      |           |
| 14 |                            | 3.4              | 3.1              | 3.2      |                      |           |
| 15 |                            | 3.2              | 3.2              | 3.1      |                      |           |
| 16 |                            |                  |                  |          |                      |           |
| 17 | 計算値                        | 設定数 a            | データ数 n1          | n2       | n3                   | 自由度 $\nu$ |
| 18 |                            | 3                | 9                | 9        | 9                    | 24        |
| 19 |                            | 平均 $x_1$         | $x_2$            | $x_3$    | データ数比 $\lambda_{1j}$ |           |
| 20 |                            | 3.34             | 3.22             | 3.27     | 0.5                  |           |
| 21 |                            | 不偏分散 $ve_1^2$    | $ve_2^2$         | $ve_3^2$ | 平均平方 $\nu_E^2$       |           |
| 22 |                            | 0.0128           | 0.0119           | 0.0100   | 0.0116               |           |
| 23 | 出力                         | 検定統計量 $q_{12}$   | $q_{13}$         |          |                      |           |
| 24 |                            | 2.41             | 1.53             |          |                      |           |

図 8.29: 母平均の差の検定(対照群と処理群, ダネットの方法) (ダネットによる検定.xlsx)

とし, 事象  $X_2, X_3$  のそれらを処理群としている. 各群それぞれに母平均  $\mu_1, \mu_2, \mu_3$ , 母分散  $\sigma_1^2, \sigma_2^2, \sigma_3^2$  の手入力を受けて正規乱数を 9 個ずつ生成させ, 式 (8.68) によりセル B30 でデータ群 1, 2 間の検定統計量  $q_{12}$ , セル C30 で  $q_{13}$  を求めている. 32 行目ではこれらの  $q_{ij}$  値とセル C6 に手入力されたダネットの方法の  $q_0$  値とを比較して,  $q_{ij}$  値の方が大きい場合に 1 を出力し, 34 行目では以上の比較でいずれかの対で 1 の場合に 1 を出力している. 以上の検定を 1000 組について繰り返している. セル B36 にはデータ群 1, 2 間, 1, 3 間の少なくともどちらかで有意差ありと判定された頻度 (第 1 種の過誤の頻度) を出力している. このシミュレーション例では 1000 組中 56 組にて第 1 種の過誤が起きていた. 図 8.31 は以上の 1000 組のシミュレーションを 100 回繰り返して得られた第 1 種の過誤の確率の 95% 信頼区間を示す. データ群 1, 2 間, 1, 3 間の母平均に差がないのに差があると誤判定された確率は約 5% に抑えられたことが分かる.

図 8.32 にテューキーの方法の  $q_{max}$  値の確率密度関数とダネットの方法の  $q_{max}$  値の確率密度関数を並べて示す. データ群数  $a = 3$ , 自由度  $\nu = 24$  の場合である. ダネットの方法ではデータ数比  $\lambda_{1j} = 0.5$  としている. テューキーの方法の  $q_{max}$  値の分布は式 (8.41) により求めた. また, ダネットの方法では後述の式 (8.77) を用いた. ダネットの方法では確率密度関数は  $q$  の大きな領域において裾が伸びない形状となり, 閾値  $q_0$  は小さな値となるように調整されている様子を見ることができる.

|    |                                     |                     |                     |         |                     |                     |            |
|----|-------------------------------------|---------------------|---------------------|---------|---------------------|---------------------|------------|
| 1  | 母平均の差の検定のシミュレーション(対照群との比較, ダネットの方法) |                     |                     |         |                     |                     |            |
| 2  |                                     |                     |                     |         |                     |                     |            |
| 3  | 入力                                  | $\mu 1$             | $\mu 2$             | $\mu 3$ | $\sigma 1$          | $\sigma 2$          | $\sigma 3$ |
| 4  |                                     | 1                   | 1                   | 1       | 1                   | 1                   | 1          |
| 5  | 公称の有意水準 $\alpha$                    | ダネットの $q_0$         |                     |         |                     |                     |            |
| 6  |                                     | 0.05                | 2.349               |         |                     |                     |            |
| 7  |                                     |                     |                     |         |                     |                     |            |
| 8  |                                     | 1組目                 |                     |         | 2組目                 |                     |            |
| 9  | 計算値                                 | x1                  | x2                  | x3      | x1                  | x2                  | x3         |
| 10 |                                     | 1.352               | -0.235              | 0.723   | -0.130              | 0.824               | 2.876      |
| 11 |                                     | 0.862               | 3.314               | 0.859   | 0.435               | 1.457               | -1.755     |
| 29 |                                     | 検定統計量 $q_{12}$      | $q_{13}$            |         | $q_{12}$            | $q_{13}$            |            |
| 30 |                                     | 0.300               | 0.237               |         | 0.474               | 1.996               |            |
| 31 |                                     | if $ q_{12}  > q_0$ | if $ q_{13}  > q_0$ |         | if $ q_{12}  > q_0$ | if $ q_{13}  > q_0$ |            |
| 32 |                                     | then 1 else 0       |                     |         |                     |                     |            |
| 33 |                                     | 0                   | 0                   |         | 0                   | 0                   |            |
| 34 |                                     | 各対のいずれかで有意差あり       |                     |         |                     |                     |            |
| 35 | 出力                                  | 有意頻度 fr             |                     |         |                     |                     |            |
| 36 |                                     | 0.056               |                     |         |                     |                     |            |

図 8.30: 母平均の差の検定のシミュレーション（対照群と処理群，ダネットの方法）（ダネットによるシミュレーション.xlsxm）

|    |     |          |         |                 |        |
|----|-----|----------|---------|-----------------|--------|
| 53 |     | 有意頻度 frk | 100回の平均 |                 |        |
| 54 | fr1 | 0.057    | 0.0502  |                 |        |
| 55 | fr2 | 0.045    | 95%信頼区間 |                 |        |
| 56 | fr3 | 0.049    | 0.0487  | $\leq p^* \leq$ | 0.0517 |

図 8.31: 母平均の差の検定，ダネットの方法における第1種の過誤の確率の95%信頼区間

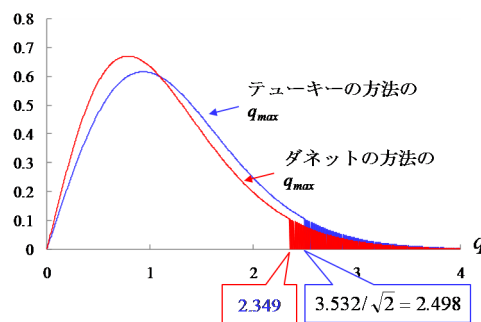


図 8.32:  $q_{max}$  の確率密度関数の形状（チューキーの方法とダネットの方法の比較）

8.4.3 ダネットの方法の理論（両側検定）

式 (8.67) の帰無仮説は，両側検定であるので

$$\max_{j=2,3} |q_{1j}| \leq q_0 \tag{8.72}$$

もしくは,

$$|q_{12}| \leq q_0 \quad \text{and} \quad |q_{13}| \leq q_0 \quad (8.73)$$

と表現できる. このとき, 上式を満たす  $\bar{x}_1, \bar{x}_2, \bar{x}_3$  は

$$\begin{aligned} -q_0 &\leq \frac{\bar{x}_1 - \bar{x}_2}{\sqrt{v_E^2 \left( \frac{1}{n_1} + \frac{1}{n_2} \right)}} \leq q_0 \\ -q_0 &\leq \frac{\bar{x}_1 - \bar{x}_3}{\sqrt{v_E^2 \left( \frac{1}{n_1} + \frac{1}{n_3} \right)}} \leq q_0 \end{aligned} \quad (8.74)$$

の範囲にある. 8.2.3項にて, テューキー・クレーマーの方法では,  $\bar{x}_1, \bar{x}_2, \bar{x}_3$  が上式を満たしても,  $|\bar{x}_2 - \bar{x}_3|$  が第1種の過誤を起こす場合があることを指摘した. **ダネットの方法では  $|\bar{x}_2 - \bar{x}_3|$  は検定の対象外**であることに注意されたい. 仮定の  $\mu_1 = \mu_2 = \mu_3$  を挿入して, 変形すると

$$\begin{aligned} \sqrt{\frac{n_2}{n_1}} \frac{\bar{x}_1 - \mu_1}{\sqrt{\frac{\sigma^2}{n_1}}} - \sqrt{n_2} \sqrt{\frac{1}{n_1} + \frac{1}{n_2}} \frac{v_E}{\sigma} q_0 &\leq \frac{\bar{x}_2 - \mu_2}{\sqrt{\frac{\sigma^2}{n_2}}} \leq \sqrt{\frac{n_2}{n_1}} \frac{\bar{x}_1 - \mu_1}{\sqrt{\frac{\sigma^2}{n_1}}} + \sqrt{n_2} \sqrt{\frac{1}{n_1} + \frac{1}{n_2}} \frac{v_E}{\sigma} q_0 \\ \sqrt{\frac{n_3}{n_1}} \frac{\bar{x}_1 - \mu_1}{\sqrt{\frac{\sigma^2}{n_1}}} - \sqrt{n_3} \sqrt{\frac{1}{n_1} + \frac{1}{n_3}} \frac{v_E}{\sigma} q_0 &\leq \frac{\bar{x}_3 - \mu_3}{\sqrt{\frac{\sigma^2}{n_3}}} \leq \sqrt{\frac{n_3}{n_1}} \frac{\bar{x}_1 - \mu_1}{\sqrt{\frac{\sigma^2}{n_1}}} + \sqrt{n_3} \sqrt{\frac{1}{n_1} + \frac{1}{n_3}} \frac{v_E}{\sigma} q_0 \end{aligned} \quad (8.75)$$

となる. 8.1.3項の考察により, 事象  $\sqrt{n_i/\sigma^2}(\bar{X}_i - \mu_i)$  は互いに独立に標準正規分布に従い, 事象  $S = V_E/\sigma$  は式(8.38)の確率密度関数  $h_\nu(s)$  の分布に従う. ただし,  $\nu = n_1 + n_2 + n_3 - 3$  である. また,  $n_2 = n_3$  を前提としているので,  $n_2 = n_3 = n_j$  とおく. これらの事象が互いに独立であるとして, 式(8.72)を満たす確率  $P_1$  は

$$\begin{aligned} P_1 = \int_0^\infty \int_{-\infty}^\infty \left\{ \Phi \left( \sqrt{\frac{n_j}{n_1}} w_{10} + \sqrt{\frac{n_j}{n_1} + 1s} q_0 \right) - \Phi \left( \sqrt{\frac{n_j}{n_1}} w_{10} - \sqrt{\frac{n_j}{n_1} + 1s} q_0 \right) \right\}^2 \\ \times f(w_{10}) h_\nu(s) dw_{10} ds \end{aligned} \quad (8.76)$$

と求められる.

第1種の過誤の確率  $P(\max|q_{ij}| > q_0)$  は

$$P(\max|q_{1j}| > q_0) = 1 - P_1 \quad (8.77)$$

により求められる.

データ群が  $a$  個の場合に拡張する。検定の仮説は、データ群  $i$  の母平均を  $\mu_i$  とすると

$$\begin{aligned} \text{帰無仮説} & : \mu_1 = \mu_j \\ \text{対立仮説} & : \mu_1 \neq \mu_j \\ & \text{ただし, } j = 2, \dots, a \end{aligned} \quad (8.78)$$

である。ここで、データ群  $2, 3, \dots, a$  のデータ数は全て等しいことを前提としている。これらのデータ数を  $n_j$ 、自由度を

$$\nu = a(n_j - 1) \quad (8.79)$$

とし、データ数比  $\lambda_{1j}$  を

$$\lambda_{1j} = \frac{n_j}{n_1 + n_j} \quad (8.80)$$

とする。第1種の過誤の確率  $P(\max|q_{1j}| > q_0)$  は

$$P(\max|q_{1j}| > q_0) = 1 - \int_0^\infty \int_{-\infty}^\infty \left\{ \Phi\left(\frac{\sqrt{\lambda_{1j}}w + sq_0}{\sqrt{1-\lambda_{1j}}}\right) - \Phi\left(\frac{\sqrt{\lambda_{1j}}w - sq_0}{\sqrt{1-\lambda_{1j}}}\right) \right\}^{a-1} \times f(w)h_\nu(s)dw ds \quad (8.81)$$

となる。各群のデータ数が与えられた段階で、上式を用いて公称の有意水準  $\alpha$  を 0.05 もしくは 0.01 とする検定統計量の閾値  $q_0$  を求めることができる。

|    | A   | B                    | C              | D               | E                   | F             |
|----|---|----------------------|----------------|-----------------|---------------------|---------------|
| 1  | 第1種の過誤の数値計算(母平均の差の検定, <b>ダネットの方法の%点, 両側検定</b> ) |                      |                |                 |                     |               |
| 2  |   |                      |                |                 |                     |               |
| 3  | 入力  | データ群数<br>a           | 対照群のデータ数<br>n1 | 処理群のデータ数<br>n2  | 公称の有意水準<br>$\alpha$ | ダネットの閾値<br>q0 |
| 4  |   | 3                    | 9              | 9               | 0.05                | 2.349         |
| 5  | 計算値   | データ数比 $\lambda_{1j}$ | 自由度 $\nu$      | $\Gamma(\nu/2)$ | gcoef               |               |
| 6  |   | 0.5                  | 24             | 39916800        | 446734.2296         |               |
| 7  | 出力  | 第1種の過誤の確率 p*         |                |                 |                     |               |
| 8  |   | 0.05004              |                |                 |                     |               |
| 9  |   |                      |                |                 |                     |               |
| 10 | 数値計算値   | $s_j$                |                |                 |                     |               |
| 11 |   |                      | 0              | 0.01            | 0.02                | 0.03          |
| 12 | $x_k$   | -6                   | 0              | 4.47502E-72     | 1.55464E-64         | 4.157E-60     |
| 13 |   | -5.99                | 0              | 5.35654E-72     | 1.86064E-64         | 4.974E-60     |
| 14 |   | -5.98                | 0              | 6.40977E-72     | 2.2262E-64          | 5.949E-60     |
| 15 |   | -5.97                | 0              | 7.6678E-72      | 2.66279E-64         | 7.115E-60     |

図 8.33: ダネットの方法の5%点の数値計算（ダネットによる数値計算（両側）.xls）

図 8.33 は上式を利用して、第1種の過誤の確率を数値計算により求めている Excel の画面である。図 8.14 との違いは確率の計算式が式 (8.81) によっていることと、検定統計

量の閾値  $q_0$  が表 A.2~A.6 のダネットの方法の5%点（両側検定）により与えられる点だけである。数表には、データ数比  $\lambda_{1j} = 0.1, 0.3, 0.5, 0.7, 0.9$  の各場合が用意されている。この例では  $\lambda_{1j} = 0.5$  であるので、公称の有意水準  $\alpha = 0.05$ 、データ群数  $a = 3$ 、自由度  $\nu = 24$  より、 $q_0 = 2.349$  と読める。数値計算の結果は第1種の過誤の確率  $p^* = 0.05$  となった。図 8.31 のシミュレーション結果はこの数値計算の結果と一致していた。

|    | A                   | B                    | C                            | D                             | E                       |
|----|---------------------|----------------------|------------------------------|-------------------------------|-------------------------|
| 1  | ダネットの方法の%点(両側検定)の探索 |                      |                              |                               |                         |
| 2  |                     |                      |                              |                               |                         |
| 3  | 入力                  | 公称の有意水準<br>$\alpha$  | データ群数<br>$a = 2 \sim 10$ の整数 | 自由度<br>$\nu = 2 \sim 250$ の整数 | データ数比<br>$\lambda_{1j}$ |
| 4  |                     | 0.05                 | 3                            | 24                            | 0.5                     |
| 5  | 計算値                 | $\Gamma(\nu/2)$      | gの係数                         | 探索回数<br>$n_{\text{search}}$   | 第1種の過誤の<br>確率 $p^*$     |
| 6  |                     | 39916800             | 446734.23                    | 24                            | 0.05000001              |
| 7  | 出力                  | ダネットの方法<br>の%点 $q_0$ |                              |                               |                         |
| 8  |                     | 2.3494               |                              |                               |                         |
| 9  |                     |                      |                              |                               |                         |
| 10 | 数値計算値               |                      | $s_j$                        |                               |                         |
| 11 |                     |                      | 0                            | 0.01                          | 0.02                    |
| 12 | $x_k$               | -6                   | 0                            | 4.4766E-72                    | 1.55521E-64             |
| 13 |                     | -5.99                | 0                            | 5.35842E-72                   | 1.86132E-64             |
| 14 |                     | -5.98                | 0                            | 6.41203E-72                   | 2.22702E-64             |
| 15 |                     | -5.97                | 0                            | 7.6705E-72                    | 2.66376E-64             |

図 8.34: ダネットの方法の%点（両側検定）の探索（ダネットの%点（両側）の探索.xlsm）

数表のダネットの方法の5%点（両側検定）の値は式 (8.81) を用いた探索により求めることができる。図 8.34 はダネットの方法の%点（両側検定）の探索を行う Excel の画面である。入力項目は公称の有意水準  $\alpha$ （セル B4）、データ群数  $a$ （セル C4）、自由度  $\nu$ （セル D4）、データ数比  $\lambda_{1j}$ （セル E4）である。出力項目はダネットの方法の%点（セル B8）である。図 8.17 と同様のマクロにより、2分探索法を用いた探索がなされる。得られる  $q_0$  値は  $2 \leq a \leq 9, 2 \leq \nu \leq 240$  の範囲で表 A.2~A.6 の値と一致することを確認してある。

#### 8.4.4 ダネットの方法の理論（片側検定）

ダネットの方法における片側検定の仮説は

$$\begin{aligned}
 \text{帰無仮説} &: \mu_1 = \mu_j \\
 \text{対立仮説} &: \mu_1 > \mu_j \\
 &\text{ただし, } j = 2, 3
 \end{aligned}
 \tag{8.82}$$

である。これより、帰無仮説は

$$\max_{j=2,3} q_{1j} \leq q_0
 \tag{8.83}$$

もしくは,

$$q_{12} \leq q_0 \quad \text{and} \quad q_{13} \leq q_0 \quad (8.84)$$

と表現できる. このとき,  $\bar{x}_1 \geq \bar{x}_2, \bar{x}_3$  の場合のみを考えればよい. 上式を満たす  $\bar{x}_1, \bar{x}_2, \bar{x}_3$  は

$$\begin{aligned} \frac{\bar{x}_1 - \bar{x}_2}{\sqrt{v_E^2 \left( \frac{1}{n_1} + \frac{1}{n_2} \right)}} &\leq q_0 \\ \frac{\bar{x}_1 - \bar{x}_3}{\sqrt{v_E^2 \left( \frac{1}{n_1} + \frac{1}{n_3} \right)}} &\leq q_0 \end{aligned} \quad (8.85)$$

の範囲にある. 仮定の  $\mu_1 = \mu_2 = \mu_3$  を挿入して, 変形すると

$$\begin{aligned} \sqrt{\frac{n_2}{n_1}} \frac{\bar{x}_1 - \mu_1}{\sqrt{\frac{\sigma^2}{n_1}}} - \sqrt{n_2} \sqrt{\frac{1}{n_1} + \frac{1}{n_2}} \frac{v_E}{\sigma} q_0 &\leq \frac{\bar{x}_2 - \mu_2}{\sqrt{\frac{\sigma^2}{n_2}}} \\ \sqrt{\frac{n_3}{n_1}} \frac{\bar{x}_1 - \mu_1}{\sqrt{\frac{\sigma^2}{n_1}}} - \sqrt{n_3} \sqrt{\frac{1}{n_1} + \frac{1}{n_3}} \frac{v_E}{\sigma} q_0 &\leq \frac{\bar{x}_3 - \mu_3}{\sqrt{\frac{\sigma^2}{n_3}}} \end{aligned} \quad (8.86)$$

となる. よって, 式 (8.83) を満たす確率  $P_1$  は

$$P_1 = \int_0^\infty \int_{-\infty}^\infty \left\{ 1 - \Phi \left( \sqrt{\frac{n_j}{n_1}} w_{10} - \sqrt{\frac{n_j}{n_1} + 1} s q_0 \right) \right\}^2 f(w_{10}) h_\nu(s) dw_{10} ds \quad (8.87)$$

と求められる.

第1種の過誤の確率  $P(\max q_{ij} > q_0)$  は

$$P(\max q_{1j} > q_0) = 1 - P_1 \quad (8.88)$$

により求められる.

データ群が  $a$  個の場合に拡張する. 検定の仮説は, データ群  $i$  の母平均を  $\mu_i$  とすると

$$\begin{aligned} \text{帰無仮説} &: \mu_1 = \mu_j \\ \text{対立仮説} &: \mu_1 > \mu_j \\ &\text{ただし, } j = 2, \dots, a \end{aligned} \quad (8.89)$$

である. ここで, データ群  $2, 3, \dots, a$  のデータ数は全て等しいことを前提としている. これらのデータ数を  $n_j$ , 自由度を  $\nu = a(n_j - 1)$  とし, データ数比  $\lambda_{1j}$  を

$$\lambda_{1j} = \frac{n_j}{n_1 + n_j} \quad (8.90)$$

とする. 第1種の過誤の確率  $P(\max q_{1j} > q_0)$  は

$$P(\max q_{1j} > q_0) = 1 - \int_0^\infty \int_{-\infty}^\infty \left\{ 1 - \Phi \left( \frac{\sqrt{\lambda_{1j}} w - s q_0}{\sqrt{1 - \lambda_{1j}}} \right) \right\}^{a-1} f(w) h_\nu(s) dw ds \quad (8.91)$$

となる.

図 8.35 はダネットの方法の%点 (片側検定) の探索を行う Excel の画面である. 入力項目は公称の有意水準  $\alpha$  (セル B4), データ群数  $a$  (セル C4), 自由度  $\nu$  (セル D4), データ数比  $\lambda_{1j}$  (セル E4) である. 出力項目はダネットの方法の%点 (片側検定) (セル B8) である. 得られる  $q_0$  値は  $2 \leq a \leq 9, 2 \leq \nu \leq 240$  の範囲で表 A.7~A.11 の値と一致することを確認してある.

|    | A                     | B                    | C                            | D                             | E                       |
|----|-----------------------|----------------------|------------------------------|-------------------------------|-------------------------|
| 1  | ダネットの方法の%点 (片側検定) の探索 |                      |                              |                               |                         |
| 2  |                       |                      |                              |                               |                         |
| 3  | 入力                    | 公称の有意水準<br>$\alpha$  | データ群数<br>$a = 2 \sim 10$ の整数 | 自由度<br>$\nu = 2 \sim 250$ の整数 | データ数比<br>$\lambda_{1j}$ |
| 4  |                       | 0.05                 | 3                            | 24                            | 0.5                     |
| 5  | 計算値                   | $\Gamma(\nu/2)$      | gの係数                         | 探索回数<br>$n_{search}$          | 第1種の過誤の<br>確率 $p^*$     |
| 6  |                       | 39916800             | 446734.23                    | 22                            | 0.04999994              |
| 7  | 出力                    | ダネットの方法<br>の%点 $q_0$ |                              |                               |                         |
| 8  |                       | 2.0080               |                              |                               |                         |
| 9  |                       |                      |                              |                               |                         |
| 10 | 数値計算値                 |                      | $s_j$                        |                               |                         |
| 11 |                       |                      | 0                            | 0.01                          | 0.02                    |
| 12 | $x_k$                 | -6                   | 0                            | 2.71105E-53                   | 2.26602E-46             |
| 13 |                       | -5.99                | 0                            | 2.87855E-53                   | 2.40602E-46             |
| 14 |                       | -5.98                | 0                            | 3.05609E-53                   | 2.55442E-46             |
| 15 |                       | -5.97                | 0                            | 3.24426E-53                   | 2.7117E-46              |

図 8.35: ダネットの方法の%点 (片側検定) の探索 (ダネットの%点 (片側) の探索.xlsm)

## 8.5 まとめ

母平均の差の検定を繰り返す場合には, 各群の平均値の差が互いに独立ではないために, ボンフェローニの方法やシダックの方法では閾値を厳しくしすぎてしまう. そこで, 本章では第1種の過誤の確率が公称の有意水準を超えないように, 検定統計量の閾値を調整する方法を解説した. 3群以上のデータ群間の母平均の差の検定においては, 以下の場合分けと方法の使い分けがなされる.

データ群間の母平均の対比較の検定法

(i) データ数が等しく, 母分散が等しいと見なせる場合

テューキーの方法

(ii) データ数が異なり，母分散が等しいと見なせる場合

テューキー・クレーマーの方法

(iii) データ数が異なり，母分散が等しいと見なせない場合

ゲイムズ・ハウウェルの方法

対照群と処理群の間の母平均の差の検定法

(iv) 母分散が等しいと見なせる場合

ダネットの方法



## 参考文献

- [1] 金田行雄, 笹井理生監修「超多自由度系の新しい科学」, 計算科学講座 10, 共立出版, 2010.
- [2] L. A. Zadeh, “Foreword to Soft Computing Series,” 日本ファジィ学会編「ソフトコンピューティングシリーズ」朝倉書店, 1995.
- [3] 菅野道夫「ファジィ理論の展開」サイエンス社, 1989.
- [4] 中森義輝「ファジィモデリング」オーム社, 1994.
- [5] 日本ファジィ学会編「ファジィ・ニューラルシステム」講座ファジィ第 12 巻, 日刊工業新聞社, 1995.
- [6] 林 勲, 古橋 武編著「ファジィ・ニューラルネットワーク」日本ファジィ学会編: ソフトコンピューティングシリーズ 6, 朝倉書店, 1996.
- [7] B. A. Wichmann and I. D. Hill, “Algorithm AS 183: An Efficient and Portable Pseudo-Random Number Generator.” *Journal of the Royal Statistical Society. Series C (Applied Statistics)*, Vol. 31, No. 2, pp. 188-190, 1982.
- [8] B. A. Wichmann and I. D. Hill, “Correction: Algorithm AS 183: An Efficient and Portable Pseudo-Random Number Generator.” *Journal of the Royal Statistical Society. Series C (Applied Statistics)*, Vol. 33, No. 1, p. 123, 1984.
- [9] 二見良治, 西敏明「Excel 統計解析講座 2: 必須統計解析の基礎」日科技連, 2001.
- [10] 内田治「すぐわかる EXCEL による統計解析」東京図書, 1996.
- [11] 清水信博「もう悩まない! 論文が書ける統計」オーエムエス出版, 2004.
- [12] G. K. Robinson, “Behrens-Fisher Problem,” *Encyclopedia of Statistical Science*, S. Kots, N. L. Johnson (Eds.), Vol. 1, pp.205-209, 1981.
- [13] 永田靖, 吉田道弘「統計的多重比較法の基礎」サイエンティスト社, 1987.

- [14] 道用大介, 「図解でわかるエクセルのマクロと VBA がみるみるわかる本」 秀和システム, 2010.
- [15] Neil Salkind(Ed.), “Bonferroni Test,” Encyclopedia of Measurement and Statistics, Vol.1, SAGE Publications, pp. 103-107, 2007.
- [16] Neil Salkind(Ed.), “Tukey-Kramer Procedure,” Encyclopedia of Measurement and Statistics, Vol.3, SAGE Publications, pp. 1016-1018, 2007.
- [17] 広津千尋 「実験データの解析—分散分析を超えて—」 共立出版, 1992.
- [18] A. J. Hayter, “A Proof of the Conjecture that the Tukey-Kramer Multiple Comparisons Procedure is Conservative,” The annals of Statistics, Vol. 12, No. 1, pp.61-75, 1984.
- [19] Paul A. Games and John F. Howell, “Pairwise Multiple Comparison Procedures with Unequal N’s and/or Variances: A Monte Carlo Study”, Journal of Educational and Behavioral Statistics, Vol. 1, No. 2, pp.113-125, 1976.
- [20] A. J. Hayter, “The Maximum Familywise Error Rate of Fisher’s Least Significant Difference Test,” Journal of the American Statistical Association, Vol. 81, No. 396, pp. 1000-1004, 1986.
- [21] 稲垣宣生 「数理統計学」 裳華房, 2003.
- [22] 竹村彰通 「現代数理統計学」 創文社, 1991.
- [23] 白石高章 「Tukey-Kramer 法に関連した分布の上界」 計算機統計学, Vol.19, No. 2, pp.77-87, 2006
- [24] 白石高章 「多群連続モデルにおける多重比較法」 共立出版, 2011

## 著者

古橋 武

名古屋大学工学研究科計算理工学専攻

本稿の内容は、

古橋武・宮本定明著

「統計・多変量解析とソフトコンピューティング ―超多自由度系解析を目指して―」

金田・笹井監修，計算科学講座 第3巻，共立出版，2012

<http://www.kyoritsu-pub.co.jp/bookdetail/9784320122680>

から抜粋したものです。共立出版社の許可を得て Web ページに掲載しています。著作権法上で認められている例外を除き，出版社の許可なく複写することはできません。