統計・多変量解析と
ソフトコンピューティング

第7章　多重比較法–名義水準の調整–

本稿掲載のWebページ

http://mybook-pub-site.sakura.ne.jp/Statistics_Multivariate/index.html

古橋　武

目次

第7章

多重比較法–名義水準の調整–

7.1 母平均の検定の例–シダックの方法–

ある日，ある工場の3種類の製品のラインにおいて，あるパラメータの測定がなされたとする．製品1, 2, 3からそれぞれ9, 8, 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．

製品1：3.3, 3.2, 3.4, 3.3, 3.3, 3.2, 3.5, 3.2, 3.5
製品2：5.5, 5.1, 5.3, 5.4, 4.9, 5.2, 5.3, 5.4
製品3：8.3, 8.0, 8.1, 8.2, 8.3, 8.2, 8.1

この日の平均値 $\overline{x}_1 = 3.32, \overline{x}_2 = 5.26, \overline{x}_3 = 8.17$ であった．製品1, 2, 3の通常の平均値はそれぞれ $\mu_{10} = 3.25, \mu_{20} = 5.1, \mu_{30} = 8.3$ であったとする．この日の各製品の母平均 μ_1, μ_2, μ_3 は通常の平均値と比べて，異なっているものがあると言えるか？ ただし，これらの製品の母分散 $\sigma_1^2, \sigma_2^2, \sigma_3^2$ は分かっていない．

検定の仮説は

$$
\begin{aligned}
\text{帰無仮説} &\quad : \quad \mu_i = \mu_{i0} \\
\text{対立仮説} &\quad : \quad \mu_i \neq \mu_{i0} \\
&\quad \quad \text{ただし}, i = 1, 2, 3
\end{aligned} \tag{7.1}
$$

であり，検定の有意水準 $\alpha = 0.05$（公称の有意水準という）とする．4.1節の母平均との差の検定を3群のデータに対して適用する．ただし，平均 $\overline{x}_i$ と母平均 $\mu_i (i = 1, 2, 3)$ の差

の検定の個別の有意水準 α'（名義水準という）は次式により決定する．

$$\begin{aligned}\alpha' &= 1-(1-\alpha)^{\frac{1}{3}} \\ &= 1-(1-0.05)^{\frac{1}{3}} \\ &= 0.0170 \end{aligned} \tag{7.2}$$

検定統計量 $t_i(i=1,2,3)$ は式 (4.1) より

$$t_i = \frac{\overline{x}_i - \mu_{i0}}{\frac{v_{ei}}{\sqrt{n_i}}} \tag{7.3}$$

となる．ここで，n_i は製品 i の標本の大きさであり，v_{ei}^2 は製品 i の不偏分散である．

	A	B	C	D	E
1	母平均の検定(Sidak)				
2					
3	入力	通常の平均値μ1	μ2	μ3	公称の有意水準α
4		3.25	5.1	8.3	0.05
5		x1	x2	x3	
6		3.3	5.5	8.3	
7		3.2	5.1	8	
8		3.4	5.3	8.1	
9		3.3	5.4	8.2	
10		3.3	4.9	8.3	
11		3.2	5.2	8.2	
12		3.5	5.3	8.1	
13		3.2	5.4		
14		3.5			
15	計算値	データ数 n1	n2	n3	名義水準α′
16		9	8	7	0.0170
17		平均 x1	x2	x3	
18		3.32	5.26	8.17	
19		不偏分散v1^2	v2^2	v3^2	
20		0.0144	0.0370	0.0124	
21		検定統計量 t1	t2	t3	
22		1.8028	2.3906	-3.0571	
23	出力	p値 p1	p2	p3	
24		0.1091	0.0481	0.0223	

$1-(1-\alpha)^{1/3}$

=TDIST(B22,B16-1,2)

図 7.1: 母平均の検定 (シダックの方法)(シダックの検定.xlsx)

図 7.1 は検定を実行している Excel の画面である．図説の「シダックの検定.xlsx」は対応する Excel ファイルのファイル名である．これら Excel ファイルは以下の URL

(http://www.kyoritsu-pub.co.jp/bookdetail/9784320122680) からダウンロードできる．各群の平均値に対してそれぞれ??節の検定を適用している．各検定の p 値（両側検定）はそれぞれ 0.1091, 0.0481, 0.0223 でありいずれも名義水準 $\alpha' = 0.0170$ より大きい．したがって，帰無仮説を棄却できずこの日の母平均が通常の母平均と異なる製品があるとは言えないという結果が得られた．

7.2　多重性の問題

3 群のデータに対して平均値の検定を繰り返すシミュレーションを行う．公称の有意水準 $\alpha = 0.05$，名義水準 $\alpha' = 0.05$ とする. 事象 X_1, X_2, X_3 は互いに独立であり，それぞれ平均 μ_1, μ_2, μ_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． 図 7.2 は事象毎に検定を行うことを 1000 回繰り返すシミュレーションの Excel の画面である．事象 X_1, X_2, X_3 につきそれぞれ 9 個，8 個，7 個の正規乱数を生成し，それぞれを 1 群として，3 群のデータを 1 組としている．図 7.3 は図 7.2 のつづきの画面である．群毎に??節の母平均との差の検定を行っている．事象 X_i の平均値 $\overline{X}_i$ と母平均 μ_i との差の p 値が名義水準 $\alpha' = 0.05$（両側検定）を下まわったとき，33 行目のセルに 1 を出力する．35 行目のセルには 3 群のいずれかで有意水準を下まわった場合に 1 を出力する．このセルの値が 1 であるとき第 1 種の過誤が起きていることを意味する．

以上のシミュレーションを 1000 組について行い，35 行目の 1 の数を数えて 1000 で割った結果をセル B37 に出力している．この例では 1000 組中 147 組において第 1 種の過誤が起きていたことを示している．

以上の 1000 組のシミュレーションを 100 回行い，第 1 種の過誤の母平均 p^* の 95% 信頼区間の出力画面を図 7.4 に示す．$0.1399 \le p^* \le 0.1443$ であった．第 1 種の過誤の確率は公称の有意水準 $\alpha = 0.05$ を大きく超えてしまった．このように「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を超えてしまうこと」を多重性の問題とよぶ．

	A	B	C	D	E
1	t検定の繰り返しのシミュレーション（名義水準の調整なし）				
2					
3	入力	μ1	μ2	μ3	
4		1	1	1	
5		σ1	σ2	σ3	
6		1	1	1	
7		公称の有意水準α	名義水準α'		
8		0.05	0.05		
9					
10	計算値	1組目			2組目
11		x1	x2	x3	x1
12		2.223	2.342	2.195	1.908
13		0.626	0.614	1.228	0.716
14		2.482	2.305	2.039	1.139
15		2.203	1.695	1.607	0.824
16		1.367	0.873	−0.157	0.740
17		−0.470	2.025	2.216	2.803
18		1.517	0.692	0.016	0.517
19		1.521	2.100		1.446
20		−0.536			0.541

図 7.2: 母平均の検定における第 1 種の過誤のシミュレーション（名義水準の調整なし）

名義水準調整なしのシミュレーション.xlsm

	A	B	C	D	E
22		データ群数 a	組数 G		
23		3	1000		
24		データ数 n1	n2	n3	
25		9	8	7	
26		平均値X1	X2	X3	X1
27		1.215	1.581	1.306	1.182
28		検定統計量 t1	t2	t3	t1
29		0.575	2.226	0.807	0.717
30		p値 p1	p2	p3	p1
31		0.581	0.061	0.451	0.494
32		if p1 < α′ then 1 else 0	if p2 < α′	if p3 < α′	if p1 < α′
33		0	0	0	0
34		3群のいずれかで有意差あり			3群のいずれかで有意差あり
35		0			0
36	出力	有意頻度fr			
37		0.147			

図 7.3: 母平均の検定における第 1 種の過誤のシミュレーション（名義水準の調整なし）（つづき）

53		有意頻度 frk	100回の平均		
54	fr1	0.134	0.1421		
55	fr2	0.152	95%信頼区間		
56	fr3	0.138	0.1399	≦ p* ≦	0.1443

図 7.4: 母平均の検定における第 1 種の過誤の 95% 信頼区間（名義水準の調整なし）

7.3 ボンフェローニの方法, シダックの方法

前節の例では各群における第 1 種の過誤の確率を α' に設定した．各群において第 1 種の過誤の起きない確率は $1-\alpha'$ である．各群は互いに独立であるので 3 群のいずれにおいても第 1 種の過誤が起きない確率 p_1 は，

$$p_1 = (1-\alpha')^3 \tag{7.4}$$

である．よって，3 群のいずれかにおいて第 1 種の過誤が起きる確率 p^* は以下の式で与えられる．

$$\begin{aligned} p^* &= 1 - p_1 \\ &= 1-(1-\alpha')^3 \end{aligned} \tag{7.5}$$

前節では名義水準 $\alpha' = 0.05$ としていたので $p^* = 0.143$ と求められた．図 7.4 の結果は以上の考察と一致している．

本節で紹介する多重比較法はボンフェローニ (Bonferroni) の方法とシダック (Sidak) の方法 [1] である．ここで，多重比較法とは，多群のデータに対する検定の繰り返しにおい

て第 1 種の過誤の確率が公称の有意水準を超えないようにする検定法をいう．ボンフェローニ, シダックの方法は第 1 種の過誤が公称の有意水準 α を超えないように名義水準 α' を調整する検定法である．考え方は，式 (7.5) の第 1 種の過誤の起きる確率 p^* が公称の有意水準 α となるように名義水準 α' を調整することにある．すなわち

$$\alpha \quad = \quad 1-(1-\alpha')^a \tag{7.6}$$

とすることである．ただし，a は群の数である．これより

$$\alpha' \quad = \quad 1-(1-\alpha)^{\frac{1}{a}} \tag{7.7}$$

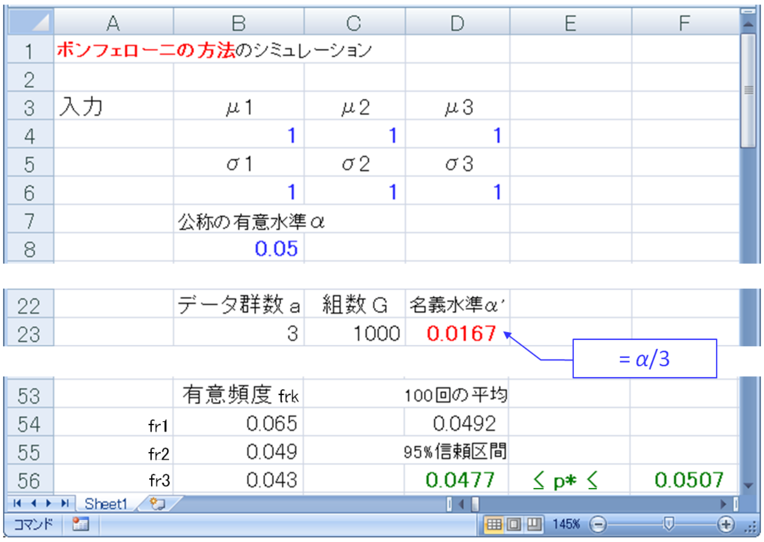
と求められる．この名義水準調整による検定法がシダックの方法である．今日のようにコンピュータが手軽に使えなかった時代には $1/a$ 乗の計算は容易ではなかった．そこで，式 (7.7) を α に関してマクローリン展開して，その第 1 項をとることで次の近似式を得る．

$$\begin{aligned} \alpha' \quad &= \quad \frac{\alpha}{a}+\frac{1}{2!}\frac{1}{a}\left(1-\frac{1}{a}\right)\alpha^2+\frac{1}{3!}\frac{1}{a}\left(1-\frac{1}{a}\right)\left(2-\frac{1}{a}\right)\alpha^3+\cdots \\ &\approx \quad \frac{\alpha}{a} \end{aligned} \tag{7.8}$$

ここで，$a \geq 3$ であるので，

$$1-(1-\alpha)^{\frac{1}{a}} \geq \frac{\alpha}{a} \tag{7.9}$$

である．この近似式により名義水準を調整した検定法がボンフェローニの方法である．



	A	B	C	D	E	F
1	ボンフェローニの方法のシミュレーション					
2						
3	入力	μ1	μ2	μ3		
4		1	1	1		
5		σ1	σ2	σ3		
6		1	1	1		
7		公称の有意水準α				
8		0.05				
22		データ群数 a	組数 G	名義水準α'		
23		3	1000	0.0167		
53		有意頻度 frk		100回の平均		
54	fr1	0.065		0.0492		
55	fr2	0.049		95%信頼区間		
56	fr3	0.043		0.0477	≦ p* ≦	0.0507

図 7.5: 多重比較法（ボンフェローニの方法）のシミュレーションと 95% 信頼区間（ボンフェローニによるシミュレーション.xlsm）

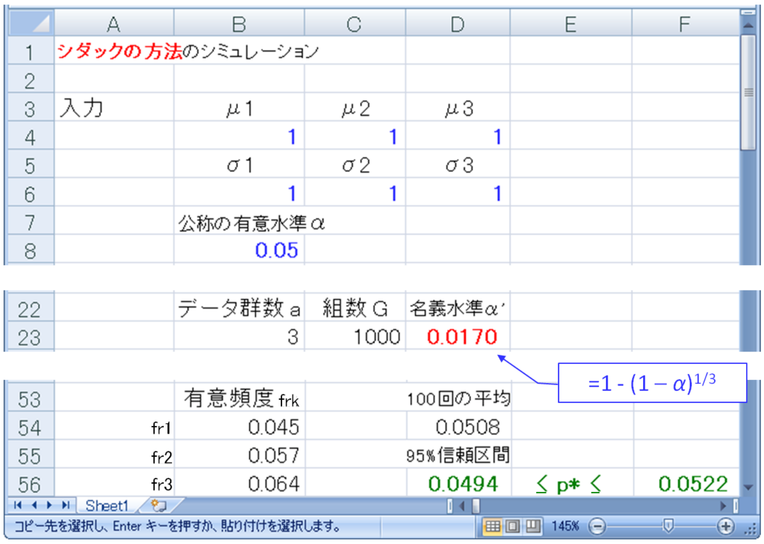


図 7.6: 多重比較法（シダックの方法）のシミュレーションと 95% 信頼区間（シダックによるシミュレーション.xlsm）

図 7.5 はボンフェローニの方法による検定のシミュレーションの Excel の画面の一部と，1000 組のシミュレーションを 100 回行って第 1 種の過誤の確率の 95% 信頼区間を求めた結果である．図 7.2 との設定の違いは，名義水準 α' を式 (7.8) により求めている点だけである．得られた第 1 種の過誤の確率 p^* の 95% 信頼区間は $0.0477 \leq p^* \leq 0.0507$ であり，この例では公称の有意水準 α に合う結果が得られた．しかし，ボンフェローニの方法は名義水準 α' をシダックの方法よりも小さな値に調整する．

図 7.6 はシダックの方法によるシミュレーションの画面の一部と第 1 種の過誤の確率の 95% 信頼区間の結果である．得られた第 1 種の過誤の確率 p^* は $0.0494 \leq p^* \leq 0.0522$ であり，式 (7.5) に一致する結果である．

7.4　まとめ

多重比較法とは，多群のデータに対する検定の繰り返しにおいて第 1 種の過誤の確率が公称の有意水準 α を超えないようにする検定法をいう．ボンフェローニ, シダックの方法は第 1 種の過誤が公称の有意水準 α を超えないように名義水準 α' を調整する検定法である．

a をデータ群の数とすると，シダックの方法は

$$\alpha' \;=\; 1-(1-\alpha)^{\frac{1}{a}} \tag{7.10}$$

とする．ボンフェローニの方法は

$$\alpha' = \frac{\alpha}{a} \tag{7.11}$$

とする．

参考文献

[1] Neil Salkind(Ed.), “Bonferroni Test,” Encyclopedia of Measurement and Statistics, Vol.1, SAGE Publications, pp. 103-107, 2007.

索引

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