1 質的因子の 1 因子実験

1.3 多重比較

テキスト

芳賀敏郎（2014）医薬品開発のための統計解析

第２部　実験計画法　改訂版、サイエンティスト社、p.294

第２部　実験計画法

１因子実験・・・質的因子　1.1 繰り返し数が等しい場合、1.2 繰り返し数が異なる場合
1.3 多重比較、1.4 ばらつきを特性値とする実験
1.5 ノンパラメトリック検定
量的因子　2.1 直線関係の場合、2.2 非直線関係の場合
2.3 ダミー変数による質的因子の効果の推定

乱塊法・・・・・3.1 質的因子の乱塊法、3.2 量的因子の乱塊法、3.3 欠測値のある場合

共分散分析・・・4.1 共分散分析の目的、4.2 解析手順、4.3 医薬品開発における共分散分析の例

２因子実験・・・5.1 ２因子実験の基礎、5.2 質的因子×質的因子、5.3 質的因子×量的因子
5.4 質的因子×量的因子（変形）、5.5 量的因子×量的因子

多因子実験・・・6.1 多因子実験の基礎、6.2 スクリーニング計画、6.3 応答曲面計画

変量模型ほか・・7.1 １因子実験、7.2 枝分れ実験、7.3 乱塊法の拡張、7.4 経時データ、7.5 交差試験

1.3 多重比較

p.46

テキストの
該当ページ

★プレゼンテーションの
スピーカーノートを、
PDF の注釈に変換してあります

（1）実験単位の α
（2）Bonferroni（ボンフェローニ）の方法
（3）Holm（ホルム）の方法
（4）Tukey（テューキー）の方法
（5）基準となる水準との比較
（6）Dunnett（ダネット）の方法
（7）JMPによる解析
（8）Williams（ウィリアムズ）の方法
（9）方法の比較
（10）多重比較法を実施する上での注意

使用するファイル
　Excelファイル
　　「DE改1-1因子(質).xlsm」
　JMPファイ
　　「1-1因子1.jmp」
　サイエンティスト社のホームページ
　からダウンロード

JMP 10.0.2 の出力を表示

p.46

（1）実験単位の α

多重性の問題

実験単位の α

p.46

●実験単位における検定の回数

実験：４種類の薬剤を動物に投与して、薬効を表す数値の平均を得た

（表示 1.1.1　p.15）

水準	平均
A1	10.30
A2	10.80
A3	11.20
A4	11.30

検定：２水準ごとの t 検定（p.17）

水準番号		差	t 値	p 値
1	2	-0.50	-2.142	*0.048*
1	3	-0.90	-3.855	**0.001**
1	4	-1.00	-4.284	**0.001**
2	3	-0.40	-1.713	0.106
2	4	-0.50	-2.142	*0.048*
3	4	-0.10	-0.428	0.674

検定回数６回

分散分析の F 検定（p.25）

要因	平方和	自由度	平均平方	F 比	p 値
水準間	3.10	3	1.033	7.584	**0.0022**
残差	2.18	16	0.136	1.000	
全体	5.28	19	0.278		

検定回数１回

分散分析の F 検定では、因子の効果を１回検定する

２水準ごとに平均値を比較すると、最大で 水準数×(水準数－１) / 2 の回数の検定をする

水準数 a = 7 の場合、最大で $a\times(a\text{-}1) = 7\times(7\text{-}1)/2 = 21$ 回の検定を行う　→　何が問題？

実験単位の α

p.46

●個別の起こる確率と実験単位の確率

失敗する確率 0.05 の事象を 3 回繰り返したときの全体の確率（事象は独立して起こるとする）

	起こる事象			確率		計
	1 回目	2 回目	3 回目			
1 回失敗	●	○	○	$0.05\times(1\text{-}0.05)^2=$	0.0451	0.1426
	○	●	○		0.0451	
	○	○	●		0.0451	
2 回失敗	●	●	○	$0.05^2\times(1\text{-}0.05)=$	0.0024	
	○	●	●		0.0024	
	●	○	●		0.0024	
3 回失敗	●	●	●	$0.05^3=$	0.0001	
失敗しない	○	○	○	$(1-0.05)^3=$	0.8574	0.8574
				計	1	

実験単位の α

p.46

●個別の起こる確率と実験単位の確率

失敗する確率 0.05 の事象を３回繰り返したときの全体の確率（事象は独立して起こるとする）

	起こる事象 1回目	2回目	3回目	確率		計
１回失敗	●	○	○	$0.05\times(1-0.05)^2=$	0.0451	0.1426
	○	●	○		0.0451	
	○	○	●		0.0451	
２回失敗	●	●	○	$0.05^2\times(1-0.05)=$	0.0024	
	○	●	●		0.0024	
	●	○	●		0.0024	
３回失敗	●	●	●	$0.05^3=$	0.0001	
失敗しない	○	○	○	$(1-0.05)^3=$	0.8574	0.8574
				計	1	

全体で
少なくとも１回
失敗する確率
$1-(1-0.05)^3=0.1426$
↓
これが問題

１回の失敗する確率が 0.05 でも、３回繰り返すと全体で少なくとも１回失敗する確率は 0.14

たとえば、21回繰り返す場合、少なくとも１回失敗する確率は $1-(1-0.05)^{21}=0.659 \fallingdotseq 2/3$

失敗する → 誤る・・・α=0.05 の検定を繰り返すと、全体で誤る確率は0.05より大きくなる

注）実験単位で繰り返す検定はお互いに独立ではないので、このような単純な計算はできない

実験単位の α

p.46

●個別の起こる確率と実験単位の確率

実験：４種類の薬剤を動物に投与して、薬効を表す数値の平均を得た

（表示 1.1.1　p.15）

検定：２水準ごとの t 検定（p.17）

水準番号		差	t 値	p 値
1	2	-0.50	-2.142	*0.048*
1	3	-0.90	-3.855	**0.001**
1	4	-1.00	-4.284	**0.001**
2	3	-0.40	-1.713	0.106
2	4	-0.50	-2.142	*0.048*
3	4	-0.10	-0.428	0.674

個々の検定
$\alpha = 0.05$

実験単位
$\alpha > 0.05$

α ：第１種の誤りの確率
差がないのにも関わらず
差があると誤って判断する確率

１回の比較を α=0.05 で検定

実験単位（6 回の比較）の α は 0.05 よりも高くなる

→ 6 回の検定で誤って差が有意となる確率が 0.05 より高い＝検定の信頼性が保障されない

実験単位の α

α、第１種の誤りの確率、危険率、有意水準
同じ意味

●多重性の問題

実験単位の α ＝ 全体（複数回の検定）で少なくとも１回誤る確率、確からしさの保証
α = 0.05 の検定を繰り返すと、実験単位の α は増大
差が無くても検定を繰り返せば、差が有意になる確率は高くなる　→　検定の信頼性が問題

●実験単位の α を制御

個々の検定の α を制御して、
実験単位の α が 0.05 以下であることを保証する＝ 多重比較法

●事例

水準数 a =7 の比較実験で
個々の検定の α を 0.0024 にして
実験単位の α が 0.05 以下であることを保証（Bonferroniの方法）

水準	平均
A1	4.14
A2	4.17
A3	5.00
A4	5.80
A5	6.17
A6	7.33
A7	8.00

注）繰り返し数 3

	水準		p 値	個々の検定の危険率、α 検定結果 0.05	0.0024
1	A7	A1	0.0006	有意	有意
2	A7	A2	0.0006	有意	有意
3	A6	A1	0.0023	有意	有意
4	A6	A2	0.0028	有意	―
5	A7	A3	0.0041	有意	―
6	A6	A3	0.0185	有意	―
7	A7	A4	0.0249	有意	―
8	A5	A1	0.0385	有意	―
9	A5	A2	0.0385	有意	―
10	A7	A5	0.0550	―	―
			....		
21	A4	A1	0.6818	―	―

注）Bonferroni の方法で調整した危険率 0.05/21 = 0.0024

実験単位の α

p.46

α、第１種の誤りの確率、危険率、有意水準
同じ意味

●多重性の問題

実験単位の α = 全体（複数回の検定）で少なくとも１回誤る確率、確からしさの保証

α = 0.05 の検定を繰り返すと、実験単位の α は増大

差が無くても検定を繰り返せば、差が有意になる確率は高くなる　→　検定の信頼性が問題

●実験単位の α を制御

個々の検定の α を制御して、実験単位の α が 0.05 以下であることを保証する = 多重比較法

●事例

水準数 a =7 の比較実験で個々の検定の α を 0.0024 にして実験単位の α が 0.05 以下であることを保証（Bonferroniの方法）

水準	平均
A1	4.14
A2	4.17
A3	5.00
A4	5.80
A5	6.17
A6	7.33
A7	8.00

注）繰り返し数 3

	水準		p 値	個々の検定の危険率、α 検定結果 0.05	0.0024
1	A7	A1	0.0006	有意	有意
2	A7	A2	0.0006	有意	有意
3	A6	A1	0.0023	有意	有意
4	A6	A2	0.0028	有意	–
5	A7	A3	0.0041	有意	–
6	A6	A3	0.0185	有意	–
7	A7	A4	0.0249	有意	–
8	A5	A1	0.0385	有意	–
9	A5	A2	0.0385	有意	–
10	A7	A5	0.0550	–	–
			……………………		
21	A4	A1	0.6818	–	–

注）Bonferroni の方法で調整した危険率 0.05/21 = 0.0024

実験単位の α

p.46

●実験の目的に応じた多重比較の使い分

（i）すべての水準間の対について同時に比較する場合

Bonferroni の方法、Holm の方法、Tukey の方法

$m = a\,(a\text{-}1)\,/\,2$　水準数 a、比較回数 m

（ii）基準となる水準（コントロール水準）と２つ以上の比較対象群があって、
コントロール水準と比較対象群との対のみを比較する場合

Bonferroni の方法、Holm の方法、

Dunnett の方法（$m = a - 1$）

（iii）水準の間で母平均が単調増加、単調減少
するという事前情報がある場合

Williams の方法

水準
A1
A2
A3
A4

組合せ（i）
A1-A2
A1-A3
A1-A4
A2-A3
A2-A4
A3-A4

組合せ（ii）
A1-A2
A1-A3
A1-A4

$m = 4 - 1$
$=3$

$m = 4(4 - 1)/2$
$=6$

補足

p.46

●LSD 法（最小有意差法）

2 水準ごとの t 検定（p.17）

水準番号		差	t 値	p 値
1	2	-0.50	-2.142	*0.048*
1	3	-0.90	-3.855	**0.001**
1	4	-1.00	-4.284	**0.001**
2	3	-0.40	-1.713	0.106
2	4	-0.50	-2.142	*0.048*
3	4	-0.10	-0.428	0.674

分散分析の F 検定（p.25）

要因	平方和	自由度	平均平方	F 比	p 値
水準間	3.10	3	1.033	7.584	**0.0022**
残差	2.18	16	0.136	1.000	
全体	5.28	19	0.278		

2 水準ごとに t 検定を繰り返す手法（誤差を比較する 2 水準から推定、p.18　脚注）

2 水準ごとに t 検定を繰り返す手法（誤差を全水準から推定）→　LSD法と呼ぶ場合がある

補足

p.46

●LSD 法（最小有意差、最小有意差法）

2 水準ごとの t 検定（p.17）

水準番号		差	t 値	p 値
1	2	-0.50	-2.142	*0.048*
1	3	-0.90	-3.855	**0.001**
1	4	-1.00	-4.284	**0.001**
2	3	-0.40	-1.713	0.106
2	4	-0.50	-2.142	*0.048*
3	4	-0.10	-0.428	0.674

分散分析の F 検定（p.25）

要因	平方和	自由度	平均平方	F 比	p 値
水準間	3.10	3	1.033	7.584	**0.0022**
残差	2.18	16	0.136	1.000	
全体	5.28	19	0.278		

2 水準ごとに t 検定を繰り返す手法（誤差を比較する 2 水準から推定、p.18　脚注）

2 水準ごとに t 検定を繰り返す手法（誤差を全水準から推定）→　LSD法と呼ぶ場合がある

（多重比較法としては問題あり）

制約付 LSD 法（Fisher の LSD 法、PLSD法）

：分散分析で有意になった場合、 2 水準ごとの t 検定を行う

無制約 LSD 法：分散分析の結果とは関係なく、 2 水準ごとの t 検定を行う

本テキストおよび JMP では、LSD は単に有意となる平均の差（最小有意差）を指している

（2）Bonferroni（ボンフェローニ）の方法

多重比較法の中で最もシンプルな方法

Bonferroni の方法

p.46

●Bonferroni の方法の考え方

危険率 α^* で個々の検定を m 回実行したとき、

実験単位の危険率 α（m 回の内、少なくとも１回誤る確率）は $m\,\alpha^*$ を超えることはない

$$\alpha \leqq m\,\alpha^* \qquad (1.3.1)$$

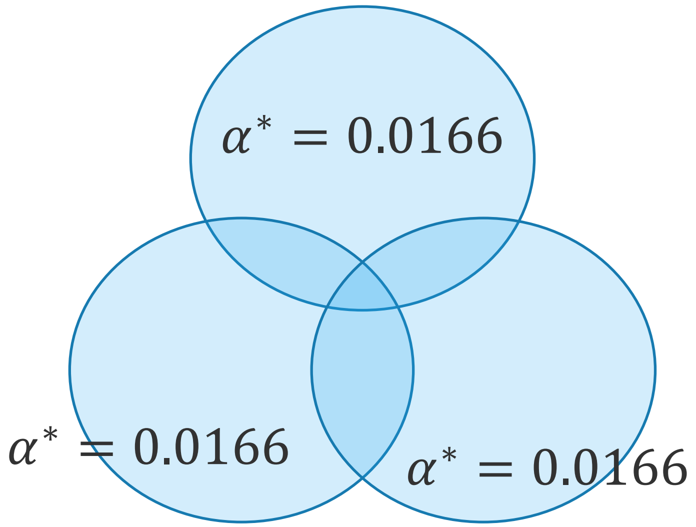
例　$\alpha^* = 0.0166$ の検定を３回（m）行うと、 α は 0.0498 以下

$$\alpha \leqq m\,\alpha^* = 3 \times 0.0166 = 0.0498$$

個々の検定の α^* を α/m にすると実験単位の危険率は α 以下になる

$$\alpha^* = \alpha / m \qquad (1.3.2)$$

$$\alpha^* = \alpha / m = 0.05 / 3 = 0.01666$$



Bonferroniの方法

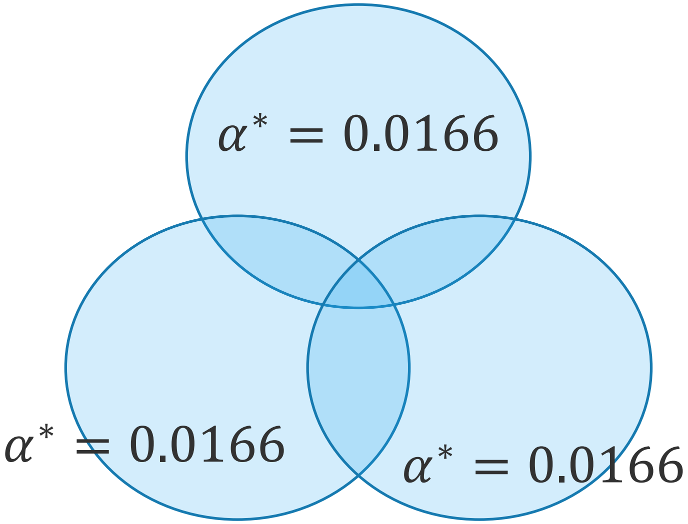
●Bonferroniの方法の計算方法

危険率 α^* で個々の検定を m 回実行したとき、
実験単位の危険率 α（m 回の内、少なくとも1回誤る確率）は $m\,\alpha^*$ を超えることはない

$$\alpha \leqq m\,\alpha^* \qquad (1.3.1)$$

例　$\alpha^* = 0.0166$ の検定を3回（m）行うと、 α は 0.0498 以下

$$\alpha \leqq m\,\alpha^* = 3 \times 0.0166 = 0.0498$$



個々の検定の α^* を α/m にすると実験単位の危険率は α 以下になる

$$\alpha^* = \alpha / m \qquad (1.3.2)$$

$$\alpha^* = \alpha / m = 0.05 / 3 = 0.01666$$

検定での計算方法

（1）危険率 α を修正：p 値が危険率 $\alpha^* = \alpha/m$ よりも小さいときに有意とする

（2）p 値を修正：p 値を m 倍した値 mp が α よりも小さいと有意にする
（mp が1を超えるときは1とする）

どちらも同じ結果になる

Bonferroni の方法

p.46

●具体例

４水準を２水準ごとに t 検定し（表示1.1.3 p.17）、水準の組合せを p 値の昇順で並べ替え

Bonferroniの方法では、α= 0.05 を検定回数 m=6 で除す、または p 値に m=6 を乗じる

t 検定では４組合せで有意、Bonferroniの方法では２組合せで有意

全てのデータを使って誤差分散を推定 p.18 脚注参照

表示 1.3.1 全ての比較の検定（一部改変）

水準番号		$\|t$ 値$\|$	t 検定		Bonferroni の方法				
			p 値	$\alpha = 0.05$	αを修正 $\alpha^* = 0.05/m$		p 値を修正 pm		$\alpha = 0.05$
1	4	4.284	0.0006	*	0.0083 = 0.05/6	*	0.0034	= 0.0006×6	*
1	3	3.855	0.0014	*	0.0083 = 0.05/6	*	0.0084	= 0.0014×6	*
2	4	2.142	0.0479	*	0.0083 = 0.05/6		0.2877	= 0.0479×6	
1	2	2.142	0.0479	*	0.0083 = 0.05/6		0.2877	= 0.0479×6	
2	3	1.713	0.1059		0.0083 = 0.05/6		0.6357	= 0.1059×6	
3	4	0.428	0.6741		0.0083 = 0.05/6		1.0000	← 0.6741×6	

p 値の昇順に並べる

絶対値

Bonferroni の方法

p.46

●具体例

４水準を２水準ごとに t 検定し（表示1.1.3 p.17）、水準の組合せを p 値の昇順で並べ替え

Bonferroniの方法では、α= 0.05 を検定回数 m=6 で除す、または p 値に m=6 を乗じる

t 検定では４組合せで有意、Bonferroniの方法では２組合せで有意

表示 1.3.1 全ての比較の検定（一部改変）

p 値の昇順に並べる

水準番号		$\|t$ 値$\|$	t 検定		Bonferroni の方法			
			p 値	α = 0.05	αを修正 α^*= 0.05/m		p 値を修正 pm	α = 0.05
1	4	4.284	0.0006	*	0.0083 = 0.05/6	*	0.0034 = 0.0006×6	*
1	3	3.855	0.0014	*	0.0083 = 0.05/6	*	0.0084 = 0.0014×6	*
2	4	2.142	0.0479	*	0.0083 = 0.05/6		0.2877 = 0.0479×6	
1	2	2.142	0.0479	*	0.0083 = 0.05/6		0.2877 = 0.0479×6	
2	3	1.713	0.1059		0.0083 = 0.05/6		0.6357 = 0.1059×6	
3	4	0.428	0.6741		0.0083 = 0.05/6		1.0000 ← 0.6741×6	

Bonferroni の方法

●具体例

４水準を２水準ごとに t 検定し（表示1.1.3 p.17）、水準の組合せを p 値の昇順で並べ替え

Bonferroniの方法では、α= 0.05 を検定回数 m=6 で除す、または p 値に m=6 を乗じる

t 検定では４組合せで有意、Bonferroniの方法では２組合せで有意

表示 1.3.1 全ての比較の検定（一部改変）

水準番号		$\|t$ 値$\|$	t 検定		Bonferroni の方法			
			p 値	α = 0.05	αを修正 α^*= 0.05/m		p 値を修正 pm	α = 0.05
1	4	4.284	0.0006	*	0.0083 = 0.05/6	*	0.0034 = 0.0006×6	*
1	3	3.855	0.0014	*	0.0083 = 0.05/6	*	0.0084 = 0.0014×6	*
2	4	2.142	0.0479	*	0.0083 = 0.05/6		0.2877 = 0.0479×6	
1	2	2.142	0.0479	*	0.0083 = 0.05/6		0.2877 = 0.0479×6	
2	3	1.713	0.1059		0.0083 = 0.05/6		0.6357 = 0.1059×6	
3	4	0.428	0.6741		0.0083 = 0.05/6		1.0000 ← 0.6741×6	

p 値の昇順に並べる

m 倍した値が１を超えるときは１とする

Bonferroni の方法

- Bonferroni の方法の特徴

  多重比較の中でも保守的な方法

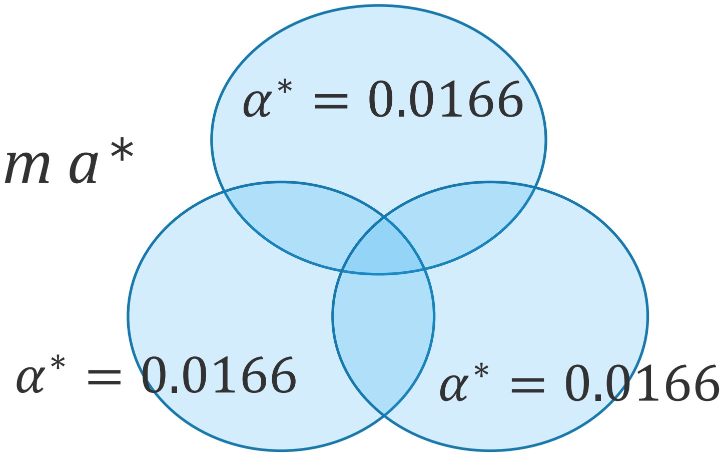
  実際の実験単位の α は、期待する値よりも低くなる

  （第２種の誤りの確率 β が大きくなり、検出力が低下する）

  すべての検定が独立していることを考えているため、

  どのような相関がある場合にも使用できる

$$\alpha \leqq m\,\alpha^*$$



- 多重度

  実験単位の α と、個々の比較に使う α^* の比を多重度 f とする

$$f = \frac{\alpha}{\alpha^*} \qquad (1.3.3)$$

$$\alpha^* = \frac{\alpha}{m} \quad \rightarrow$$

$$f = \frac{\alpha}{\alpha^*} = m$$

Bonferroni の方法

  多重比較は、個々の検定の α^* を下げて、実験単位の α を保証する

  多重度が１より大きいほど検出力は低下（多重度が１の場合、検出力の低下が最も少ない）

  Bonferroni の方法は多重度 f を m とした方法（多重比較の中でも、検出力が低い手法）

  実験単位の α を保証しつつ、多重度 f を m から１に近づけるように多重比較が開発されてきた

Bonferroni の方法

p.47

●Bonferroni の方法の特徴

多重比較の中でも保守的な方法

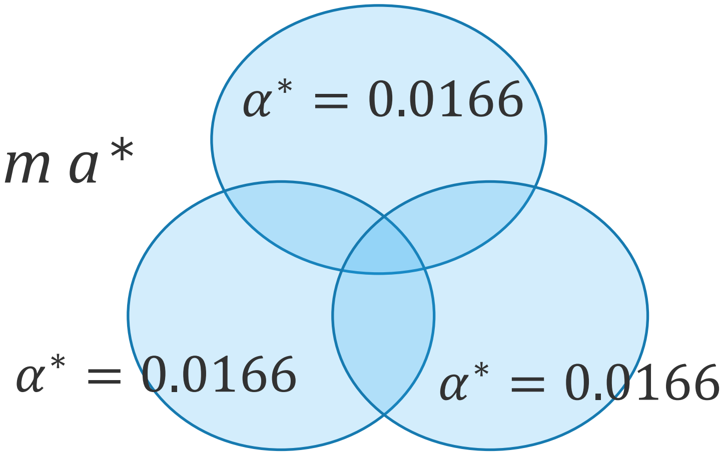
実際の実験単位の α は、期待する値よりも低くなる

（第２種の誤りの確率 β が大きくなり、検出力が低下する）

すべての検定が独立していることを考えているため、

どのような相関がある場合にも使用できる

$$\alpha \leqq m\,\alpha^*$$



●多重度

実験単位の α と、個々の比較に使う α^* の比を多重度 f とする

$$f = \frac{\alpha}{\alpha^*} \qquad (1.3.3)$$

$$\alpha^* = \frac{\alpha}{m} \quad \rightarrow$$

$$f = \frac{\alpha}{\alpha^*} = m$$

Bonferroni の方法

多重比較は、個々の検定の α^* を下げて、実験単位の α を保証する

多重度が１より大きいほど検出力は低下（多重度が１の場合、検出力の低下が最も少ない）

Bonferroni の方法は多重度 f を m とした方法（多重比較の中でも、検出力が低い手法）

実験単位の α を保証しつつ、多重度 f を m から１に近づけるように多重比較が開発されてきた

（3）Holm（ホルム）の方法

Bonferroni の方法の改良版

Holm の方法

p.47

●Holm の方法の考え方

Bonferroni の方法よりも有意差が出やすいように閉手順を利用して改良

４水準を２水準ごとにt検定し（表示1.1.3 p.17）、 p 値の昇順で並べ替え

Holm の方法では、α= 0.05 を k で除す、または p 値に k を乗じる

p 値の低い組合せから開始、k は m から１ずつ減、有意差がない組合せで検定終了

$k = m$
$k = m - 1$
$k = m - 2$
・・・
$k = 1$

p 値の昇順に並べる

表示 1.3.1 全ての比較の検定（一部改変）

水準番号		|t 値|	t 検定		Holm の方法					
			p 値	α = 0.05	αを修正 α^*= 0.05/k			p 値を修正 pk		α = 0.05
1	4	4.284	0.0006	*	0.0083	= 0.05/6	*	0.0034	= 0.0006×6	*
1	3	3.855	0.0014	*	0.0100	=0.05/5	*	0.0070	= 0.0014×5	*
2	4	2.142	0.0479	*	0.0125	=0.05/4		0.1918	= 0.0479×4	
1	2	2.142	0.0479	*						
2	3	1.713	0.1059							
3	4	0.428	0.6741							

Holm の方法

p.47

●Holm の方法の考え方

Bonferroni の方法よりも有意差が出やすいように閉手順を利用して改良

４水準を２水準ごとにt検定し（表示1.1.3 p.17）、 p 値の昇順で並べ替え

Holm の方法では、α= 0.05 を k で除す、または p 値に k を乗じる

p 値の低い組合せから開始、k は m から１ずつ減、有意差がない組合せで検定終了

$k = m$
$k = m - 1$
$k = m - 2$
・・・
$k = 1$

表示 1.3.1 全ての比較の検定（一部改変）

p 値の昇順に並べる

水準番号		$\|t$ 値$\|$	t 検定		Holm の方法				
			p 値	α = 0.05	αを修正 $\alpha^* = 0.05/k$			p 値を修正 pk	α = 0.05
1	4	4.284	0.0006	*	0.0083 = 0.05/6	*	0.0034	= 0.0006×6	*
1	3	3.855	0.0014	*	0.0100 =0.05/5	*	0.0070	= 0.0014×5	*
2	4	2.142	0.0479	*	0.0125 =0.05/4		0.1918	= 0.0479×4	
1	2	2.142	0.0479	*					
2	3	1.713	0.1059						
3	4	0.428	0.6741						

有意差がない組合せで検定終了

Holm の方法

●Holm の方法の計算方法

Bonferroni の方法よりも有意差が出やすいように閉手順を利用して改良

４水準を２水準ごとにt検定し（表示1.1.3 p.17）、 p 値の昇順で並べ替え

Holm の方法では、α= 0.05 を k で除す、または p 値に k を乗じる

p 値の低い組合せから開始、k は m から１ずつ減、有意差がない組合せで検定終了

$k = m$
$k = m - 1$
$k = m - 2$
・・・
$k = 1$

表示 1.3.1 全ての比較の検定（一部改変）

p 値の昇順に並べる

水準番号		$\|t$ 値$\|$	t 検定		Holm の方法			
			p 値	α = 0.05	αを修正 α^* = 0.05/k		p 値を修正 pk	α = 0.05
1	4	4.284	0.0006	*	0.0083 = 0.05/6	*	0.0034 = 0.0006×6	*
1	3	3.855	0.0014	*	0.0100 =0.05/5	*	0.0070 = 0.0014×5	*
2	4	2.142	0.0479	*	0.0125 =0.05/4		0.1918 = 0.0479×4	
1	2	2.142	0.0479	*				
2	3	1.713	0.1059					
3	4	0.428	0.6741					

有意差がない組合せで検定終了

Holm の方法

p.47

●Holm の方法とBonferroni の方法の比較

Bonferroni の方法よりも有意差が出やすいように改良（検出力が高まっている）

$k = m$
$k = m - 1$
$k = m - 2$
・・・
$k = 1$

表示 1.3.1 全ての比較の検定（一部改変）

水準番号		$\|t$ 値$\|$	t 検定		Bonferroni の方法		Holm の方法	
			p 値	$\alpha = 0.05$	p 値を修正 pm	$\alpha = 0.05$	p 値を修正 pk	$\alpha = 0.05$
1	4	4.284	0.0006	*	0.0034 = 0.0006×6	*	0.0034 = 0.0006×6	*
1	3	3.855	0.0014	*	0.0084 = 0.0014×6	*	0.0070 = 0.0014×5	*
2	4	2.142	0.0479	*	0.2877 = 0.0479×6		0.1918 = 0.0479×4	
1	2	2.142	0.0479	*	0.2877 = 0.0479×6			
2	3	1.713	0.1059		0.6357 = 0.1059×6			
3	4	0.428	0.6741		1.0000 ← 0.6741×6			

Bonferroni の方法より
p 値が小さい

p.48

（4）Tukey（テューキー）の方法

Tukey の方法

p.48

●Tukey の方法の計算

2 水準ごとの t 検定における棄却限界値との比較

$$|t_{ii'}| \geq t(\nu_e; 0.05)$$

$t(\quad)$：t 分布

ν_e ：誤差の自由度

$$|t_{ii'}| = \frac{|d_{ii'}|}{s.e.[d_{ii'}]}$$

(p.18)

$$= |d_{ii'}| \Big/ \sqrt{\left(\frac{1}{n_i} + \frac{1}{n_{i'}}\right) V_e}$$

2 水準ごとの Tukey の方法における棄却限界値との比較

$$|t_{ii'}| \geq q(a, \nu_e; 0.05)/\sqrt{2}$$

$q(\quad)$：ステュデント化された範囲の分布の上側 5 %点（両側検定に対応）

a ：比較する水準数

ν_e ：誤差の自由度

Tukey の方法

●Tukey の方法の計算

2 水準ごとの t 検定における棄却限界値との比較

$$|t_{ii'}| \geq t(\nu_e; 0.05)$$

$t(\quad)$：t 分布

ν_e ：誤差の自由度

2 水準ごとの Tukey の方法における棄却限界値との比較

$$|t_{ii'}| \geq q(a, \nu_e; 0.05)/\sqrt{2}$$

$$|t_{ii'}| \geq q(a, \nu_e; 0.05)/\sqrt{2}$$

$$\frac{|d_{ii'}|}{s.e.[d_{ii'}]} \geq q(a, \nu_e; 0.05)/\sqrt{2}$$

$$|d_{ii'}| \geq s.e.[d_{ii'}] \times q(a, \nu_e; 0.05)/\sqrt{2}$$

$$= \mathrm{HSD}$$

（Honestly Significant Difference）

$q(\quad)$：ステュデント化された範囲の分布の上側 5 %点（両側検定に対応）

a ：比較する水準数

ν_e ：誤差の自由度

Tukey(1953)は、各水準の繰返し数が全て同じ場合の多重比較法として開発

Kramer(1956)が異なる繰返し数の多重比較に拡張、Hayter(1984)が拡張法の正しいことを証明

現状で、Tukey の方法、Tukey-Kramer のHSD検定は同じ手法を指している

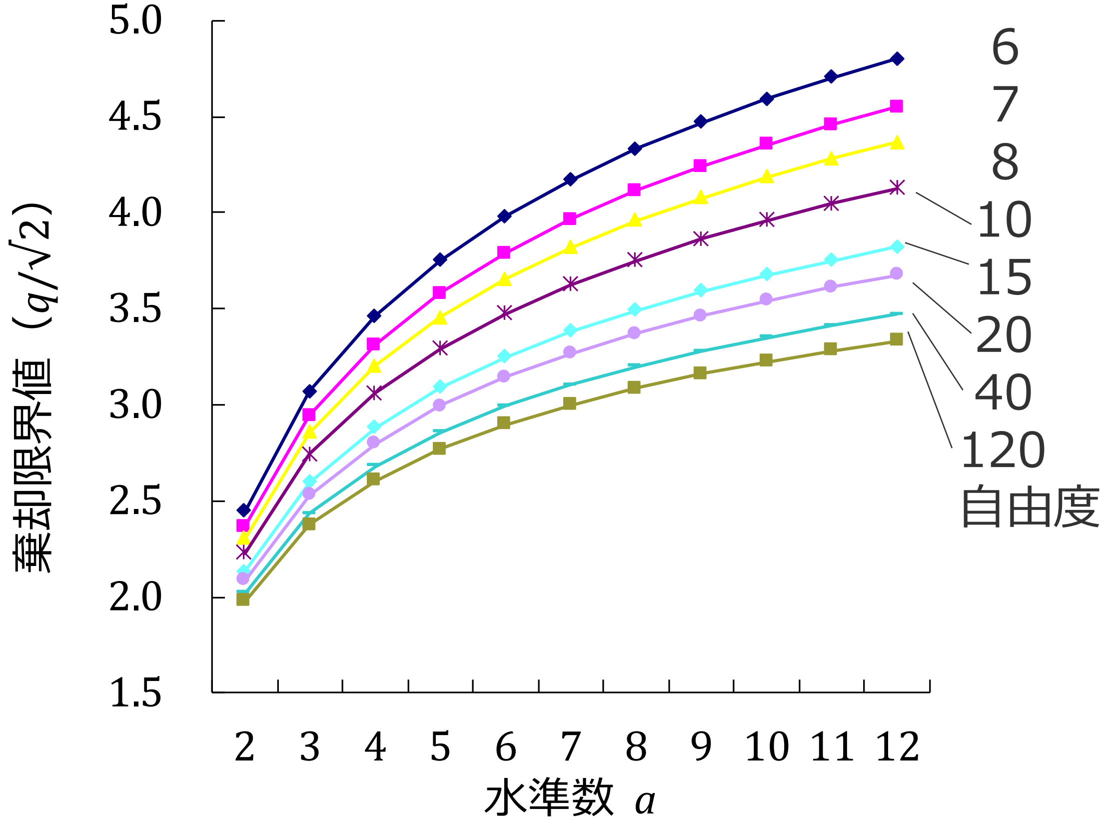
JMPは、異なる繰り返し数に対応し、多重性を調整した正確な p 値を出力

Tukey の方法

●数値表「ステュデント化された範囲の分布の上側５%点」

比較する水準数 a 、誤差の自由度 ν_e の組み合わせ、数値を $\sqrt{2}$ で除して使用

比較する水準数が多いほど、自由度が小さいほど棄却限界値は大きくなる → 有意差が出にくい



奥野・芳賀(1969)　シート「§1.3 Tukeyの数値表」q/√2

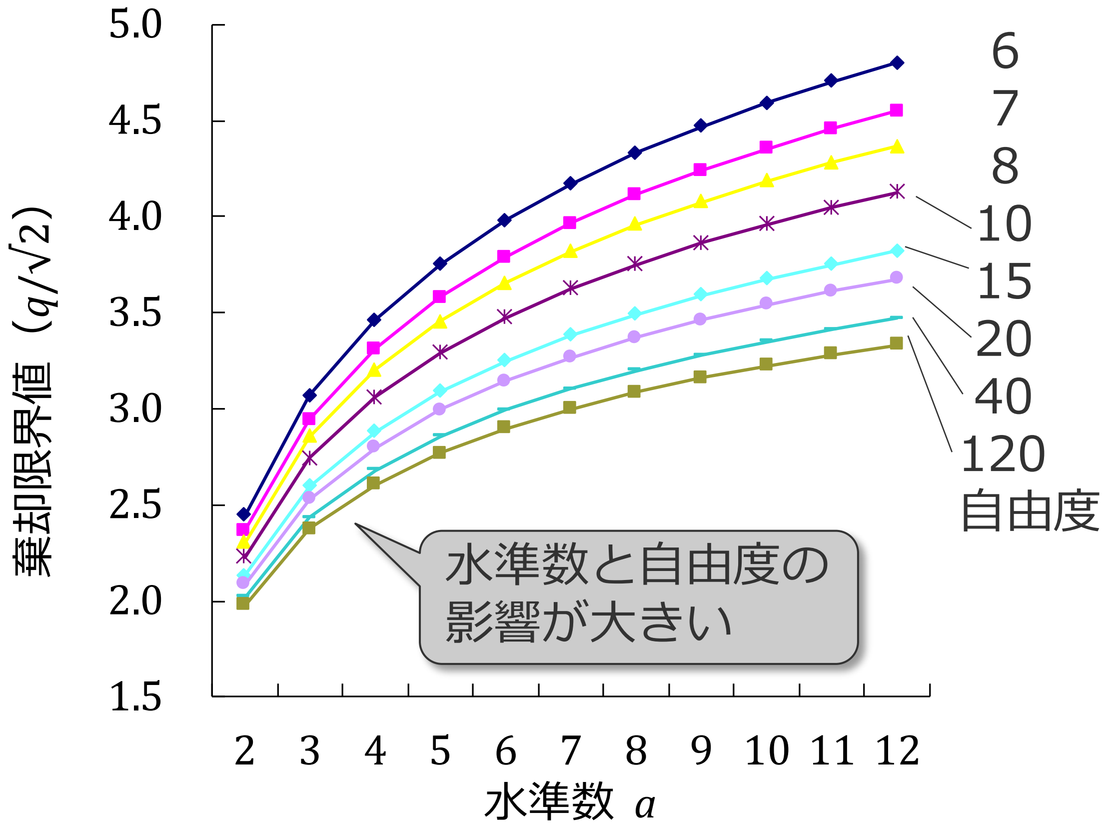
ν_e ＼ a	2	3	4	5	・・・	11	12
・							
6	2.447	3.068	3.462	3.751		4.702	4.801
7	2.365	2.945	3.310	3.578		4.456	4.547
8	2.306	2.857	3.202	3.455		4.281	4.366
・							
10	2.228	2.741	3.060	3.291		4.046	4.125
・							
15	2.131	2.598	2.882	3.088		3.752	3.821
16	2.120	2.580	2.861	3.064		3.717	3.784
・							
40	2.021	2.434	2.681	2.856		3.411	3.468
・							
120	1.980	2.373	2.606	2.770		3.282	3.333

Tukey の方法

●数値表「ステュデント化された範囲の分布の上側５％点」

比較する水準数 a 、誤差の自由度 ν_e の組み合わせ、数値を $\sqrt{2}$ で除して使用

比較する水準数が多いほど、自由度が小さいほど棄却限界値は大きくなる → 有意差が出にくい



奥野・芳賀(1969)　シート「§1.3 Tukeyの数値表」q/√2

ν_e \ a	2	3	4	5	・・・	11	12
・							
6	2.447	3.068	3.462	3.751		4.702	4.801
7	2.365	2.945	3.310	3.578		4.456	4.547
8	2.306	2.857	3.202	3.455		4.281	4.366
・							
10	2.228	2.741	3.060	3.291		4.046	4.125
・							
15	2.131	2.598	2.882	3.088		3.752	3.821
16	2.120	2.580	2.861	3.064		3.717	3.784
・							
40	2.021	2.434	2.681	2.856		3.411	3.468
・							
120	1.980	2.373	2.606	2.770		3.282	3.333

Tukeyの方法

●具体例

４水準の実験結果（表示1.1.3 p.17）

比較する水準数 a = 4, 誤差の自由度 ν_e = 16、危険率 α = 0.05

$$q(a, \nu_e; 0.05)/\sqrt{2} = q(4, 16; 0.05)/\sqrt{2} = 4.046/\sqrt{2} = 2.861$$

Tukeyの方法の棄却限界値は 2.861、t 検定の棄却限界値 2.120 よりもやや大きい

水準番号		$\lvert t$ 値$\rvert$	t 検定 棄却限界値 t (16, 0.05) = 2.120	Tukeyの方法 棄却限界値 q (4, 16; 0.05)/$\sqrt{2}$ = 2.861
1	4	4.284	*	*
1	3	3.855	*	*
2	4	2.142	*	
1	2	2.142	*	
2	3	1.713		
3	4	0.428		

Tukey の方法

p.49

●多重度

Tukey の方法の棄却限界値は 2.861 → t 分布では、危険率 α= 0.0113 の両側検定に相当

=TDIST(2.861, 16, 2) = 0.0113

したがって、式(1.3.3.) から多重度は 4.42 になる

$$f = \frac{\alpha}{\alpha^*} = \frac{0.05}{0.0113} = 4.42 \qquad (1.3.4)$$

Bonferroni の方法の多重度

$$f = \frac{\alpha}{\alpha^*} = m = 6.00 \qquad (1.3.3)$$

この値は、Bonferroni の方法の 6 から25%以上小さくなった（α^* が α に近づいた）

●水準間の全ての対を比較する多重比較

実験単位の α=0.05 が保証される三者の中で、この事例ではTukeyの方法の検出力が高い

Holmの方法は検定の度に棄却限界値（比較する α）が変わるが、その他の 2 つの方法は同じ棄却限界値（比較する α ）を用いる

表示 1.3.1　全ての比較の検定（一部改変）

水準番号		$\|t$ 値$\|$	p 値			2.861
			t 検定	Bonf.	Holm	Tukey
1	4	4.284	**0.0006**	**0.0034**	**0.0034**	**0.0029**
1	3	3.855	**0.0014**	**0.0084**	**0.0070**	**0.0069**
2	4	2.142	**0.0479**	0.2877	0.1918	0.1822
1	2	2.142	**0.0479**	0.2877		0.1822
2	3	1.713	0.1059	0.6357		0.3492
3	4	0.428	0.6741	1.0000		0.9728

Tukey の方法

p.49

● 多重度

Tukey の方法の棄却限界値は 2.861 → t 分布では、危険率 α= 0.0113 の両側検定に相当

=TDIST(2.861, 16, 2) = 0.0113

したがって、式(1.3.3.) から多重度は 4.42 になる

$$f = \frac{\alpha}{\alpha^*} = \frac{0.05}{0.0113} = 4.42 \qquad (1.3.4)$$

Bonferroni の方法の多重度

$$f = \frac{\alpha}{\alpha^*} = m = 6 \qquad (1.3.3)$$

この値は、Bonferroni の方法の 6 から25%以上小さくなった（α^* が α に近づいた）

● 水準間の全ての対を比較する多重比較

実験単位の α=0.05 が保証される三者の中で、この事例ではTukeyの方法の検出力が高い

Holmの方法は検定の度に棄却限界値（比較する α）が変わるが、その他の 2 つの方法は同じ棄却限界値（比較する α）を用いる

表示 1.3.1　全ての比較の検定（一部改変）

水準番号		| t 値|	p 値			2.861
			t 検定	Bonf.	Holm	Tukey
1	4	4.284	**0.0006**	**0.0034**	**0.0034**	**0.0029**
1	3	3.855	**0.0014**	**0.0084**	**0.0070**	**0.0069**
2	4	2.142	**0.0479**	0.2877	0.1918	0.1822
1	2	2.142	**0.0479**	0.2877		0.1822
2	3	1.713	0.1059	0.6357		0.3492
3	4	0.428	0.6741	1.0000		0.9728

p.49

（5）基準となる水準との比較

コントロール水準と比較対象との多重比較

基準となる水準との比較

p.49

- Bonferroni の方法、Holm の方法によるコントロール水準との比較
  - コントロール水準（A1）と他の３水準を２水準ごとに t 検定（表示1.1.3 p.17）、$m = 4 - 1 = 3$
  - p 値の昇順で並べ替え、方法は全ての対の比較と同じ、比較回数 m=3,　k= 3, 2, 1

表示 1.3.2 基準となる水準との比較の検定（一部改変）

水準番号		$\|t$ 値$\|$	t 検定		Bonferroni の方法				
			p 値	α = 0.05	αを修正 α^*= 0.05/m		p 値を修正 pm		α=0.05
1	4	4.284	0.0006	*	0.0167 = 0.05/3	*	0.0017	= 0.0006×3	*
1	3	3.855	0.0014	*	0.0167 = 0.05/3	*	0.0042	= 0.0014×3	*
1	2	2.142	0.0479	*	0.0167 = 0.05/3		0.1438	= 0.0479×3	

水準番号		$\|t$ 値$\|$	t 検定		Holm の方法				
			p 値	α = 0.05	αを修正 α^*= 0.05/k		p 値を修正 pk		α=0.05
1	4	4.284	0.0006	*	0.0167 = 0.05/3	*	0.0017	= 0.0006×3	*
1	3	3.855	0.0014	*	0.0250 =0.05/2	*	0.0028	= 0.0014×2	*
1	2	2.142	0.0479	*	0.0500 =0.05/1	*	0.0479	= 0.0479×1	*

基準となる水準との比較

p.49

●Bonferroni の方法、Holm の方法によるコントロール水準との比較

コントロール水準（A1）と他の３水準を２水準ごとに t 検定（表示1.1.3 p.17）、$m = a - 1$

p 値の昇順で並べ替え、方法は全ての対の比較と同じ、比較回数 m=3,　k= 3, 2, 1

表示 1.3.2 基準となる水準との比較の検定（一部改変）

水準番号		$\|t$ 値$\|$	t 検定		Bonferroni の方法					
			p 値	$\alpha = 0.05$	αを修正 $\alpha^* = 0.05/m$			p 値を修正 pm		α=0.05
1	4	4.284	0.0006	*	0.0167	= 0.05/3	*	0.0017	= 0.0006×3	*
1	3	3.855	0.0014	*	0.0167	= 0.05/3	*	0.0042	= 0.0014×3	*
1	2	2.142	0.0479	*	0.0167	= 0.05/3		0.1438	= 0.0479×3	

水準番号		$\|t$ 値$\|$	t 検定		Holm の方法					
			p 値	$\alpha = 0.05$	αを修正 $\alpha^* = 0.05/k$			p 値を修正 pk		α=0.05
1	4	4.284	0.0006	*	0.0167	= 0.05/3	*	0.0017	= 0.0006×3	*
1	3	3.855	0.0014	*	0.0250	=0.05/2	*	0.0028	= 0.0014×2	*
1	2	2.142	0.0479	*	0.0500	=0.05/1	*	0.0479	= 0.0479×1	*

p 値の昇順

（6） Dunnett（ダネット）の方法

コントロール水準と比較対象との多重比較

Dunnett の方法

p.49

●Dunnett の方法の計算

２水準ごとの t 検定における棄却限界値（第１部）

$$|t_{ii'}| \geq t(\nu_e, 0.05)$$

$t(\quad)$：t 分布

ν_e：誤差の自由度

$$|t_{ii'}| = \frac{|d_{ii'}|}{s.e.[d_{ii'}]} = |d_{ii'}| \Big/ \sqrt{\left(\frac{1}{n_i} + \frac{1}{n_{i'}}\right) V_e}$$

（p.18）

２水準ごとの Dunnett の方法における棄却限界値

$$|t_{ii'}| \geq d(a, \nu_e, \rho; 0.05)$$

$d(\quad)$：ダネットの方法のための両側５%点（ρ）

a：比較する水準数

ν_e：誤差の自由度

ρ：相関係数、対となる水準の繰り返し数で選択、同数の場合は 0.5

（永田・吉田（1997）には、ρ = 0.1, 0.3, 0.5, 0.7, 0.9 の表がある）

JMPは、異なる繰り返し数に対応し、多重性を調整した正確な p 値を出力

Dunnett の方法

●具体例

４水準の実験結果（表示1.1.3 p.17）、水準１をコントロール水準とする

比較する水準数 a = 4, 誤差の自由度 ν_e = 16、危険率 α = 0.05　（ρ=0.5）

比較する２水準の繰り返し数が同じ

$$d(a,\nu_e,\rho;0.05) = d(4,16,0.5;0.05) = 2.592$$

Dunnett の方法の棄却限界値は 2592、t 検定の棄却限界値 2.120 よりもやや大きい

コントロール水準

水準番号		$\lvert t\ 値\rvert$	t 検定 棄却限界値 $t\,(16, 0.05) = 2.120$	Dunnett の方法 棄却限界値 $d\,(4, 16, 0.5\,;0.05) = 2.592$
1	4	4.284	*	*
1	3	3.855	*	*
1	2	2.142		

Dunnett の方法

p.49

●多重度

Dunnett の方法の棄却限界値は 2.592 → これは危険率 α= 0.01966 の両側検定に相当

=TDIST(2.592, 16, 2) = 0.01966

したがって、式(1.3.3.) から多重度は 2.54 になる

$$f = \frac{\alpha}{\alpha^*} = \frac{0.05}{0.01966} = 2.54$$

Bonferroni の方法の多重度

$$f = \frac{\alpha}{\alpha^*} = m = 3 \qquad (1.3.3)$$

この値は、Bonferroni の方法の３から15%以上小さくなった（α^* が α に近づいた）

●コントロール水準と比較する多重比較

実験単位の α=0.05 が保証される三者の中で、この事例では Holm の方法の検出力が高い

表示 1.3.2 基準となる水準との比較の検定（一部改変）

水準番号		|t 値|	p 値			2.592
			t 検定	Bonf.	Holm	Dunnett
1	4	4.284	**0.0006**	**0.0017**	**0.0017**	**0.0016**
1	3	3.855	**0.0014**	**0.0042**	**0.0028**	**0.0038**
1	2	2.142	**0.0479**	0.1438	**0.0479**	0.1166

Dunnett の方法

p.49

●多重度

Dunnett の方法の棄却限界値は 2.592 → これは危険率 α= 0.01966 の両側検定に相当

=TDIST(2.592, 16, 2) = 0.01966

したがって、式(1.3.3.) から多重度は 2.54 になる

$$f = \frac{\alpha}{\alpha^*} = \frac{0.05}{0.01966} = 2.54$$

Bonferroni の方法の多重度

$$f = \frac{\alpha}{\alpha^*} = m = 3 \qquad (1.3.3)$$

この値は、Bonferroni の方法の３から15%以上小さくなった（α* が α に近づいた）

●コントロール水準と比較する多重比較

実験単位の α=0.05 が保証される三者の中で、この事例では Holm の方法の検出力が高い

表示 1.3.2 基準となる水準との比較の検定（一部改変）

水準番号		|t 値|	p 値			2.592
			t 検定	Bonf.	Holm	Dunnett
1	4	4.284	**0.0006**	**0.0017**	**0.0017**	**0.0016**
1	3	3.855	**0.0014**	**0.0042**	**0.0028**	**0.0038**
1	2	2.142	**0.0479**	0.1438	**0.0479**	0.1166

補足

p.49

●コントロール水準がある実験の解析

コントロール水準と比較対象が設定されている実験であっても、
比較対象同士の比較も行うのであれば、
対象とする対の比較をする多重比較（たとえば Tukey の方法）を利用する

実験

水準	
A0	コントロール
A1	比較対象
A2	比較対象
A3	比較対象

解析

比較する組合せ		
1	A0	A1
2	A0	A2
3	A0	A3

比較する組合せ		
1	A0	A1
2	A0	A2
3	A0	A3
4	A1	A2
5	A1	A3
6	A2	A3

コントロール水準があれば、
コントロール水準との比較だけ
というわけではない
両方の解析がありうる
実験前に、どちらかを決めておく
（結果を見てから選択するのではない）

（7）JMPによる解析

多重比較による解析

t 検定の繰り返し

Tukey の方法

Dunnett の方法

JMPによる解析

p.50

●JMPファイルの読み込みと表示

JMP ファイル「1-1因子1.jmp」を読み込み

●データ

表示 1.1.1、表示 1.1.10 と同じデータ、１因子、４水準

水準の列名は「群」、観測値の列名は「y」

●解析

［二変量の関係］

▼＞［平均／ANOVA］

▼＞［平均の比較］

＞［各ペア、Student の t 検定］

＞［すべてのペア Tukey の HSD 検定］

＞［コントロール群との比較（Dunnett）］＞［コントロール群を選択］

→「A1」を選択

表示 1.1.1

水準	1	2	3	4	5
A1	10.8	9.9	9.7	10.4	10.7
A2	10.7	10.6	11.0	10.8	10.9
A3	11.4	10.7	10.9	11.3	11.7
A4	11.9	11.2	11.0	11.1	11.3

縦に並べる

表示 1.1.10

列名

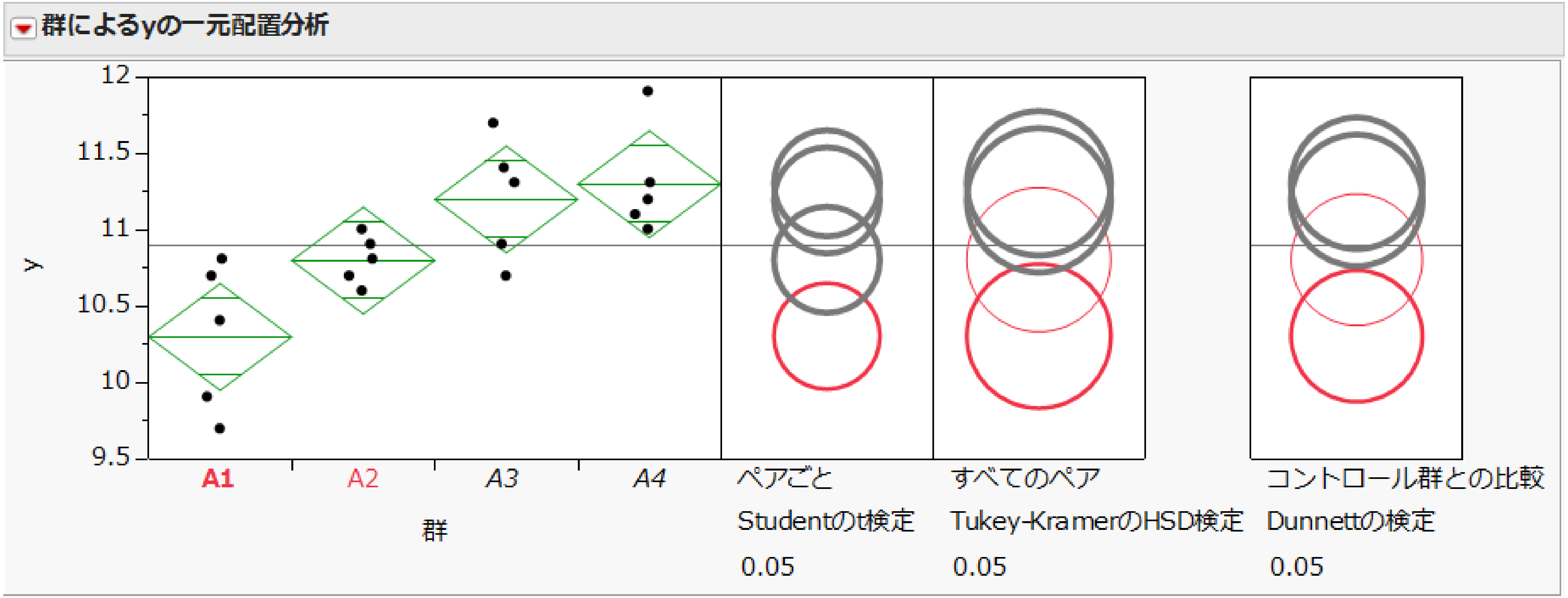
	群	y
1	A1	10.8
2	A1	9.9
3	A1	9.7
4	A1	10.4
5	A1	10.7
6	A2	10.7
7	A2	10.6
8	A2	11
9	A2	10.8
10	A2	10.9
11	A3	11.4
12	A3	10.7
13	A3	10.9
14	A3	11.3
15	A3	11.7
16	A4	11.9
17	A4	11.2
18	A4	11
19	A4	11.1
20	A4	11.3

JMPによる解析［二変量の関係］

●グラフ化

表示 1.3.3 JMP による多重比較（1）

多重度 f = 1.00　4.42　2.54



JMPによる解析［二変量の関係］

p.50

- t 検定

  §1で説明済み

  誤差は全ての水準から推定

  実験単位の α が保証されていない

  （無制約 LSD 法）

  ただし、利用場面がある（後述）

Studentのt検定を使ったペアごとの比較

棄却限界値

t	Alpha
2.11991	0.05

LSD閾値行列

LSD
最小有意差

Abs(Dif)-LSD

	A4	A3	A2	A1
A4	-0.49490	-0.39490	0.00510	0.50510
A3	-0.39490	-0.49490	-0.09490	0.40510
A2	0.00510	-0.09490	-0.49490	0.00510
A1	0.50510	0.40510	0.00510	-0.49490

値が正の場合、ペアになっている平均の間に有意差があることを示します。

JMPによる解析［二変量の関係］

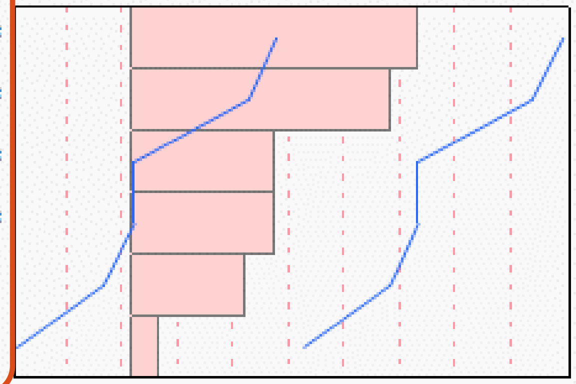
p.50

- Bonferroni 法、Holm 法

水準	t 検定		Bonferroni 法			Holm 法		
	p 値		k	修正 p 値		k	修正 p 値	
A4 - A1	0.0006	*	6	0.0036	*	6	0.0036	*
A3 - A1	0.0014	*	6	0.0084	*	5	0.0070	*
A4 - A2	0.0479	*	6	0.2874		4	0.1916	
A2 - A1	0.0479	*	6	0.2874		3		
A3 - A2	0.1059		6	0.6354		2		
A4 - A3	0.6741		6	1		1		

差の順位レポート

水準	- 水準	差	差の標準誤差	下側信頼限界	上側信頼限界	p値
A4	A1	1.000000	0.2334524	0.505103	1.494897	0.0006*
A3	A1	0.900000	0.2334524	0.405103	1.394897	0.0014*
A4	A2	0.500000	0.2334524	0.005103	0.994897	0.0479*
A2	A1	0.500000	0.2334524	0.005103	0.994897	0.0479*
A3	A2	0.400000	0.2334524	-0.094897	0.894897	0.1059
A4	A3	0.100000	0.2334524	-0.394897	0.594897	0.6741



JMPによる解析［二変量の関係］

●Tukey の方法

「Tukey-KramerのHSD検定」と表記

Tukey-KramerのHSD検定を使ったすべてのペアの比較

棄却限界値

q*	Alpha
2.86102	0.05

LSD閾値行列

HSD・・・LSDと同じ意味

Abs(Dif)-HSD

	A4	A3	A2	A1
A4	-0.66791	-0.56791	-0.16791	0.33209
A3	-0.56791	-0.66791	-0.26791	0.23209
A2	-0.16791	-0.26791	-0.66791	-0.16791
A1	0.33209	0.23209	-0.16791	-0.66791

値が正の場合、ペアになっている平均の間に有意差があることを示します。

JMPによる解析［二変量の関係］

●Tukey の方法

文字の接続レポート

水準	~文字A列	~文字B列	~文字C列	字E列	~文字F列	平均
A4	A					11.300000
A3	A					11.200000
A2	A	B				10.800000
A1		B				10.300000

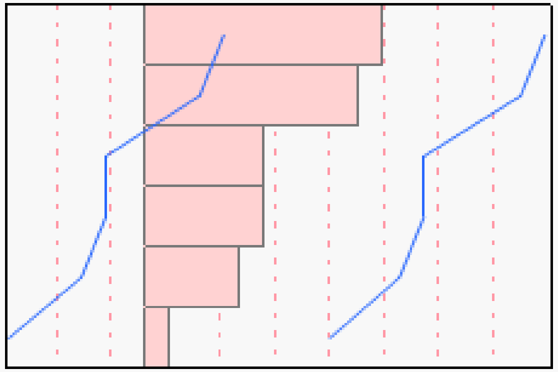
同じ文字でつながっていない水準は有意に異なります。

水準	平均	
A4	11.3	a
A3	11.2	a
A2	10.8	ab
A1	10.3	b

差の順位レポート

多重性が調整された*p* 値

水準	- 水準	差	差の標準誤差	下側信頼限界	上側信頼限界	p値
A4	A1	1.000000	0.2334524	0.332088	1.667912	0.0029*
A3	A1	0.900000	0.2334524	0.232088	1.567912	0.0069*
A4	A2	0.500000	0.2334524	-0.167912	1.167912	0.1822
A2	A1	0.500000	0.2334524	-0.167912	1.167912	0.1822
A3	A2	0.400000	0.2334524	-0.267912	1.067912	0.3492
A4	A3	0.100000	0.2334524	-0.567912	0.767912	0.9728



JMPによる解析［二変量の関係］

p.50

●Dunnett の方法

Dunnettの検定を使ったコントロール群との比較

コントロール群 = A1

棄却限界値

|d|	Alpha
2.59240	0.05

LSD閾値行列

LSD

多重性が調整されたp 値

水準	Abs(Dif)-LSD	p値
A4	0.395	0.0016*
A3	0.295	0.0038*
A2	-0.11	0.1166
A1	-0.61	1.0000

値が正の場合、ペアになっている平均の間に有意差があることを示します。

（8）Williams（ウィリアムズ）の方法

単調増加、単調減少の事前情報がある場合

Williams の方法

●概要

水準に無投与、低用量、中用量、高用量のような順序関係があり、
特性値 y が単調に増加または減少するという事前情報がある場合
事前情報を基にあらかじめ決めた順に検定を進め、有意でなくなった組合せで検定を終了
（結果を見てから順番を決めるのではない）

	無投与		低用量		中用量		高用量
平均値	100	＜	120	＜	140	＜	210

(3)　(2)　(1)
比較

単調増加が期待されたが，高用量の平均値が，中用量の平均値よりも小さかった場合、
高用量と中用量の差は誤差によるものと考え，両者の平均値を取る（調整平均値）
効果の減少を含まない折れ線で表わされるような前処理を施してから順序に検定
通常、増加か増減か予見できるので片側検定
JMP ではサポートされていない

表示 1.3.5 Williams の方法

	無投与		低用量		中用量		高用量
平均値	100	＜	120	＜	140	＞	134
調整平均値	100	＜	120	＜	137	＝	137

Williams の方法

p.51

●概要

水準に無投与、低用量、中用量、高用量のような順序関係があり、
特性値 y が単調に増加または減少するという事前情報がある場合
事前情報を基にあらかじめ決めた順に検定を進め、有意でなくなった組合せで検定を終了
（結果を見てから順番を決めるのではない）

	無投与		低用量		中用量		高用量
平均値	100	＜	120	＜	140	＜	210

(3)　(2)　(1)
比較

単調増加が期待されたが，高用量の平均値が，中用量の平均値よりも小さかった場合、
高用量と中用量の差は誤差によるものと考え，両者の平均値を取る（調整平均値）
効果の減少を含まない折れ線で表わされるような前処理を施してから順序に検定
通常、増加か増減か予見できるので片側検定
JMP ではサポートされていない

表示 1.3.5 Williams の方法

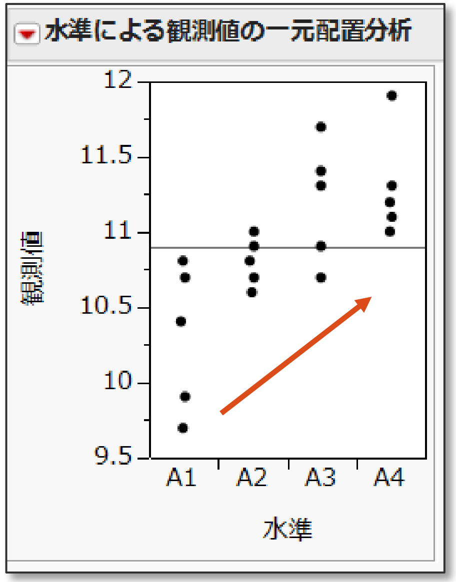
	無投与		低用量		中用量		高用量
平均値	100	＜	120	＜	140	＞	134
調整平均値	100	＜	120	＜	137	＝	137

Williams の方法

●具体例

４水準の実験結果（表示1.1.3 p.17）

水準A1を基準として単調増加するという事前情報あり（A1 ≦ A2 ≦ A3 ≦ A4）



p 値とは違う

Williams の数値表 片側5%

Williams の数値表 両側5%

Ve = 0.13625　　φ_E = 16

比較の順番	水準番号		p	Mp	tp	$W(p, \varphi_E; 0.05)$		$W(p, \varphi_E; 0.025)$	
1	A1	A4	4	11.3	4.284	1.860	*	2.216	*
2	A1	A3	3	11.2	3.855	1.831	*	2.193	*
3	A1	A2	2	10.8	2.142	1.746	*	2.120	*

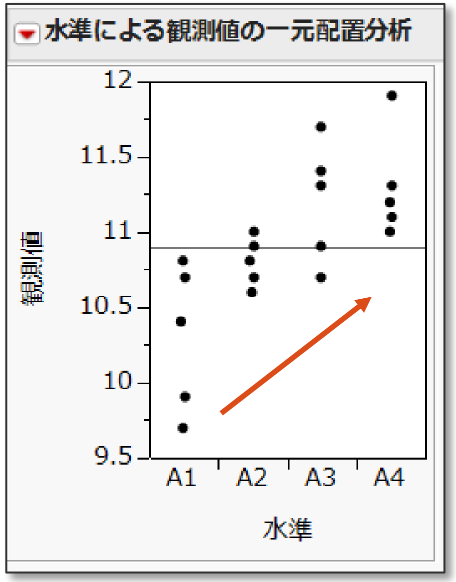
注）記号の表記は永田・吉田（1979）に準じる

Williams の方法

p.53

●具体例

４水準の実験結果（表示1.1.3 p.17）
水準 A1 を基準として単調増加するという事前情報あり（A1 ≦ A2 ≦ A3 ≦ A4）
下表の p の値（4, 3, 2）に対応した t 値（tp）と数値表の棄却限界値を比較
A1 と A4 の比較で有意 → A1 と A3 の比較で有意 → A1 と A2 の比較で有意
（この事例では、途中で検定を終了することなく、最後まで比較）
具体的な計算方法は永田・吉田（1979）を参照



p 値とは違う

Williams の数値表 片側5%

Williams の数値表 両側5%

Ve = 0.13625　　φ_E = 16

比較の順番	水準番号		p	Mp	tp	$W(p, \varphi_E; 0.05)$		$W(p, \varphi_E; 0.025)$	
1	A1	A4	4	11.3	4.284	1.860	*	2.216	*
2	A1	A3	3	11.2	3.855	1.831	*	2.193	*
3	A1	A2	2	10.8	2.142	1.746	*	2.120	*

注）記号の表記は永田・吉田（1979）に準じる

Williams の方法

p.53

● Dunnett の方法とWilliams の方法の比較

４水準の実験結果（表示1.1.3 p.17）両側５%の検定

Dunnett の方法の棄却限界値は 2.592

Williams の方法の棄却限界値は 2.216→2.193→2.120

比較する水準数が多くなると両者とも棄却限界値か高くなる、Dunnett の方法の増加程度は大

Williamsの方法の多重度は極めて１に近く、検出力の低下はわずかである

Bonferroni の方法の多重度

$$f = \frac{\alpha}{\alpha^*} = m = 3 \qquad (1.3.3)$$

表示 1.3.6 Dunnettの方法と Williams の方法の比較（ν_e = 16、両側５%）

	a	2	3	4	5	6	7	8	9
棄却限界値	Dunnett	2.120	2.424	2.592	2.708	2.796	2.866	2.924	2.974
	Williams	2.120	2.193	2.216	2.227	2.234	2.238	2.241	2.243
α^*	Dunnett	0.050	0.028	0.020	0.016	0.013	0.011	0.010	0.009
	Williams	0.050	0.043	0.042	0.041	0.040	0.040	0.040	0.039
多重度	Dunnett	1.000	1.814	2.544	3.223	3.862	4.463	5.034	5.585
	Williams	1.000	1.151	1.204	1.230	1.247	1.256	1.264	1.269

p.53

（9）方法の比較

多重比較法の多重度の比較

方法の比較

p.53

●多重度による多重比較法の比較

全ての対の比較：比較回数 m は水準数 a が多くなると急激に増加

Bonferroii の方法は f=m、検出力が著しく低下、Tukey の方法も多重度はかなり増加

基準との比較：比較回数 m は、「全ての対の比較」ほどは増加しない

多重度の増加程度は大きくない

特に Williams の方法の多重度は低いまま維持されている

多重度 f：1から増加すると検出力は低下

$$f = \frac{\alpha}{\alpha^*} \qquad (1.3.3)$$

表示 1.3.7
各種多重比較法の比較（1）
（一部改変）
両側 $\alpha = 0.05$
自由度 $\nu = \infty$

		水準数 a	2	3	4	5	6	7	8	9
全ての対の比較	比較数 m		1	3	6	10	15	21	28	36
	多重度	Bonferroni	1	3	6	10	15	21	28	36
		Tukey	1.00	2.62	4.90	7.84	11.43	15.65	20.50	25.99
基準との比較	比較数 m		1	2	3	4	5	6	7	8
	多重度	Bonferroni	1	2	3	4	5	6	7	8
		Dunnett	1.00	1.85	2.66	3.42	4.16	4.87	5.57	6.25
		Williams	1.00	1.14	1.19	1.21	1.22	1.23	1.24	1.24

方法の比較

p.53

●多重度による多重比較法の比較

(1.3.3)

全ての対の比較：比較回数 m は水準数 a が多くなると急激に増加

Bonferroii の方法は f=m、検出力が著しく低下、Tukey の方法も多重度はかなり増加

基準との比較：比較回数 m は、「全ての対の比較」ほどは増加しない

多重度の増加程度は大きくない

特に Williams の方法の多重度は低いまま維持されている

多重度 f：1から増加すると検出力は低下

$$f = \frac{\alpha}{\alpha^*} \qquad (1.3.3)$$

		水準数 a	2	3	4	5	6	7	8	9
全ての対の比較	比較数 m		1	3	6	10	15	21	28	36
	多重度	Bonferroni	1	3	6	10	15	21	28	36
		Tukey	1.00	2.62	4.90	7.84	11.43	15.65	20.50	25.99
基準との比較	比較数 m		1	2	3	4	5	6	7	8
	多重度	Bonferroni	1	2	3	4	5	6	7	8
		Dunnett	1.00	1.85	2.66	3.42	4.16	4.87	5.57	6.25
		Williams	1.00	1.14	1.19	1.21	1.22	1.23	1.24	1.24

表示 1.3.7
各種多重比較法の比較（1）
（一部改変）
両側 α = 0.05
自由度 ν = ∞

方法の比較

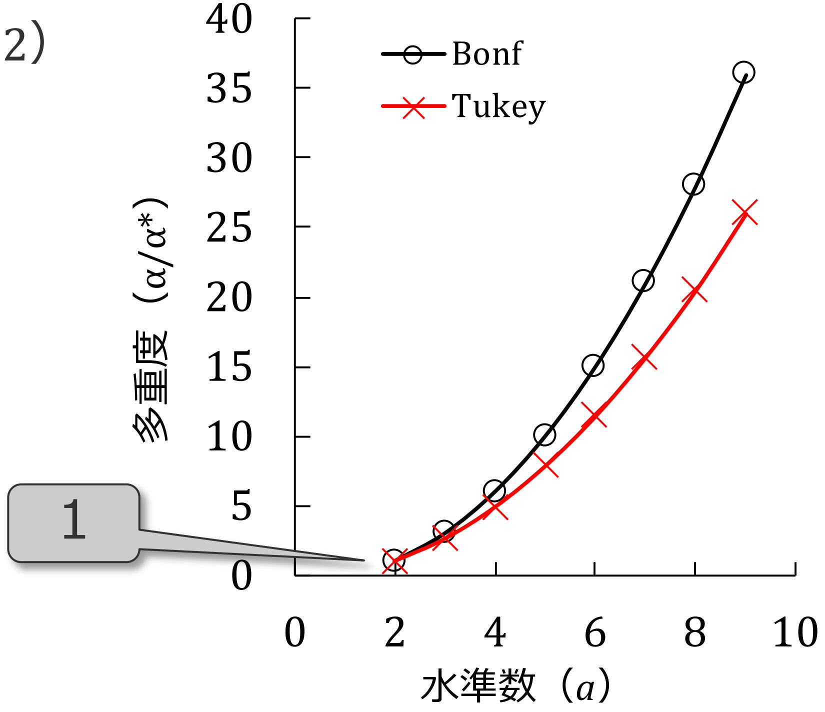
●多重度による多重比較法の比較

t 検定の繰り返しは、常に多重度が 1

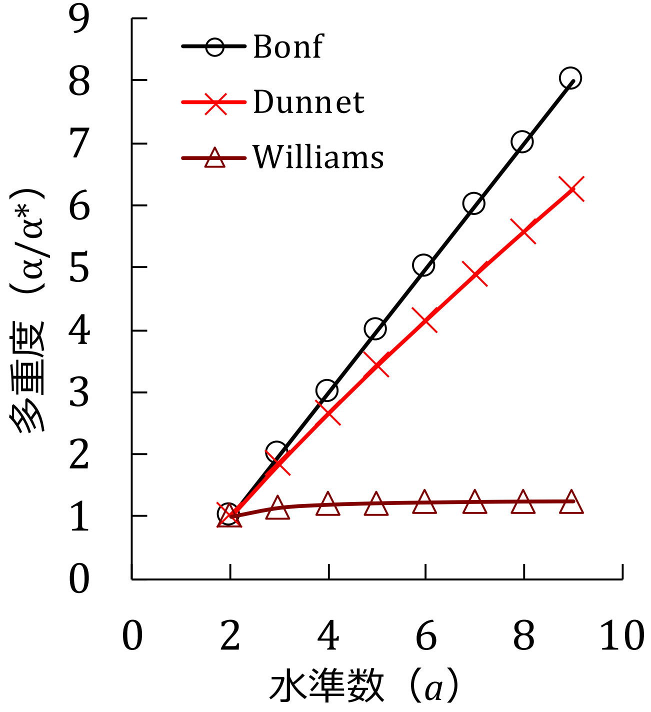
比較する水準数が多いほど多重度は増加、検出力が低下（第２種の誤りの確率が高まる）

表示 1.3.8
各種多重比較法の比較（2）

全ての対の比較



基準との比較



p.54

（10）多重比較法を実施する上での注意

多重比較の適用場面

多重比較法を実施する上での注意

p.54

●多重比較

実験単位の α（第１種の誤りの確率、危険率、有意水準）を保証

（検定を繰り返しても、差がないのに有意差があると誤る確率を抑えることはできる）

その反面、第２種の誤りの確率 β は大きくなる

（個々の比較の α を下げるために、検出力（$1-\beta$）は低下する）

●多重比較を薬剤開発の実験に使った場合

新薬が従来の薬よりも有効である（差がある）ことを保証する（常識的に多重比較を利用）

しかし、開発初期の段階では、有効な新薬を見逃す確率が高まる（重大な損失）

新薬の毒性、副作用を調べる場合、見逃す確率が高まる（見逃す危険は重大）

●多重比較を使用する指針（本テキストの著者の考え方）

第１種の誤りを厳格に保証しなければならないときに使用（例えば検証的試験）

それ以外の場合は、t 検定の繰り返しを使用（例えば探索的試験、JMP でも実施可能）

多重比較法を実施する上での注意

p.54

●多重比較

実験単位の α（第１種の誤りの確率、危険率、有意水準）を保証
（検定を繰り返しても、差がないのに有意差があると誤る確率を抑えることはできる）
その反面、第２種の誤りの確率 β は大きくなる
（個々の比較の α を下げるために、検出力（$1-\beta$）は低下する）

●多重比較を薬剤開発の実験に使った場合

新薬が従来の薬よりも有効である（差がある）ことを保証する（常識的に多重比較を利用）
しかし、開発初期の段階では、有効な新薬を見逃す確率が高まる（重大な損失）
新薬の毒性、副作用を調べる場合、見逃す確率が高まる（見逃す危険は重大）

●多重比較を使用する指針（本テキストの著者の考え方）

第１種の誤りを厳格に保証しなければならないときに使用（例えば検証的試験）
それ以外の場合は、t 検定の繰り返しを使用（例えば探索的試験、JMP でも実施可能）

補足

p.54

●多重比較の結果の解釈

多重比較の結果、有意であれば、積極的に結果を述べることができる
有意でない場合、結果（有意差がないこと）を積極的に支持できない
（個々の比較の検出力を落としいてるので、差を見逃している可能性がある）

水準	平均	
A1	32.2	a
A2	34.3	a
A3	32.1	a
A4	35.6	a

●多重比較を行う実験の設計

検証的試験では、必要な水準に絞り、適切な繰り返し数（サンプルサイズ）を用意する
適切な多重比較法を採用する
（無用な水準を加えて水準数を増やし、各水準の繰り返し数を少なくして、
Bonferroniの方法で全ての対を比較すると、恣意的に有意差を出さないようにできる
→　これは、誤用、悪用になる）
実験前に、比較方法と適用する多重比較法を選択（結果を見てから判断するのではない）
分散分析と多重比較法とは別の統計手法、分散分析→多重比較法という順番ではない

補足

p.54

●多重比較の結果の解釈

多重比較の結果、有意であれば、積極的に結果を述べることができる

有意でない場合、結果（有意差がないこと）を積極的に支持できない

（個々の比較の検出力を落としいてるので、差を見逃している可能性がある）

水準	平均	
A1	32.2	a
A2	34.3	a
A3	32.1	a
A4	35.6	a

●多重比較を行う実験の設計

検証的試験では、必要な水準に絞り、適切な繰り返し数（サンプルサイズ）を用意する

適切な多重比較法を採用する

（無用な水準を加えて水準数を増やし、各水準の繰り返し数を少なくして、

Bonferroni の方法で全ての対を比較すると、恣意的に有意差を出さないようにできる

→　これは、誤用、悪用になる）

実験前に、比較方法と適用する多重比較法を選択（結果を見てか選択するのではない）

分散分析と多重比較法とは別の統計手法、分散分析→多重比較法という順番ではない

補足

p.54

●多重比較として利用できない方法

多重比較法にはここに紹介した以上に多種多様な方法がある

書籍や統計ソフトに含まれている手法であっても、実験単位の α が保証されない手法がある

（永田・吉田, 1997 参照）

実験単位の α が保証されない手法（誤った手法）

2標本の t 検定の繰り返し（誤差を比較する2水準から推定）

2標本の t 検定の繰り返し（誤差を全水準から推定）の内、無制約 LSD 法

Duncan の方法

４水準以上の場合に実験単位の α が保証されない手法（３水準の多重比較法としては正しい）

2標本の t 検定の繰り返し（誤差を全水準から推定）の内、制約付 LSD法（Fisher の LSD法）

Newman-Keuls の方法

このうち、２標本の t 検定の繰り返し（誤差を全水準から推定）を使用する考え方について、
本テキストの著者の考え方は本節（10）で説明済

補足

p.54

●多重比較として利用できない方法

多重比較法にはここに紹介した以上に多種多様な方法がある

書籍や統計ソフトに含まれている手法であっても、実験単位の α が保証されない手法がある

（永田・吉田, 1997 参照）

実験単位の α が保証されない手法（誤った手法）

2 標本の t 検定の繰り返し（誤差を比較する 2 水準から推定）

2 標本の t 検定の繰り返し（誤差を全水準から推定）の内、無制約 LSD 法

Duncan の方法

4 水準以上の場合に実験単位の α が保証されない手法（3 水準の多重比較法としては正しい）

2 標本の t 検定の繰り返し（誤差を全水準から推定）の内、制約付 LSD法（Fisher の LSD法）

Newman-Keuls の方法

このうち、2 標本の t 検定の繰り返し（誤差を全水準から推定）を使用する考え方について、本テキストの著者の考え方は本節（10）で説明済

補足

p.54

●パラメトリック検定とノンパラメトリック検定における多重比較

バラメトリック検定の多重比較

Bonferroni の方法（t 検定）、Holm の方法（t 検定）、Tukey の方法、Dunnett の方法

Williams の方法・・・いずれも正規性、等分散性が前提

ノンパラメトリック検定での多重比較

Bonferroniの方法、Holmの方法はノンパラメトリック検定でも利用可

t 検定の代わりにWilcoxon検定、χ^2検定等を使う（近似計算による p 値が必要）

ノンパラメトリック検定の多重比較法

近似計算に基づいているので、十分な繰り返し数（サンプルサイズ）が必要

Steel-Dwass の方法、Steel の方法はJMPでサポートされている

永田・吉田（1997）参照

	パラメトリック検定	ノンパラメトリック検定
全ての対の比較	Tukey の方法	Steel-Dwass の方法
基準との比較	Dunnett の方法	Steel の方法
単調増加、単調減少	Williams の方法	Shirley-Williams の方法

まとめ

p.55

●多重比較の理解

なぜ多重比較が必要なのか

ここでは基本的な事項を解説

●多重比較法の選択

多重比較法には多くの種類がある

現実の問題にぶつかったときに復習

永田・吉田（1979）を参照

●多重比較を利用する場面

多重比較法を実施する上での注意を理解

● 引用

永田　靖・吉田道弘（1997）統計的多重比較法の基礎、サイエンティスト社

三輪哲久（2015）実験計画法と分散分析、朝倉書店

奥野忠一・芳賀敏郎（1969）実験計画法、培風館

● 作成　片瀬雅彦

● 監修　松本一彦、長谷文雄

● 作成時期　2019年6月19日

● 改訂　2019年 9月10 日、2023年9月5日