SSE-DP-2026-2

Lecture on Basic Concepts of Multiple Comparison Procedures

by

Tetsuhisa Miwa

The Institute of Statistical Mathematics

March 2026

(Summary)

This lecture was delivered through the course of the project for training experts in statistical sciences. The topic concerns basic concepts of multiple comparison procedures.

In statistical analyses we often perform several tests for one set of observed data, Then the problem of multiplicity arises. This lecture first explains the basic principles to cope with this problem of multiplicity. Then some multiple comparison procedures are described for practical problems.

講義資料: 多重比較法の基礎

三輪哲久
統計数理研究所

2026 年 **3** 月

（要約）

本資料は，「統計エキスパート人材育成プロジェクト」における「大学統計教員育成研修」の一環として実施された講義「多重比較法の基礎」の講義資料である。読者の教育・研究に利用していただければ幸いである。

統計的データ解析においては，一組のデータセットに対し複数回の検定を行なうことは多い。そのとき，検定の多重性の問題が生じる。本講義では先ず検定の多重性に対処するための基礎的な概念を説明する。引き続き，具体的な問題に対して個別の多重比較手法を解説する。

大学統計教員育成研修: 先端的データサイエンス特論

多重比較法の基礎

三輪哲久

統計数理研究所　大学統計教員育成センター

2025 年 12 月，2026 年 1 月

目 次

1　多重比較の問題

1.1　はじめに

本講義では多重比較法について，その基礎的な考え方を解説する。最初に分散分析モデル（正規線形モデル）を中心として，理論と適用法を説明する。

統計的データ解析において検定を複数回行なう場合，**多重比較法** (multiple comparison procedures) を使うべきであるとよく言われる。実際，論文投稿に際してレフリーから多重検定に関して指摘されることも多い。また各種の統計パッケージにおいても，数多くの多重比較手法が提供されている。図 1.1 に SPSS の多重比較実行画面を示す。

図 1.1　SPSS の多重比較画面

SAS の GLM プロシージャにおける MEANS ステートメントでは，多重比較法を実行するために，次のようなオプションが利用可能である。

BON, DUNCAN, DUNNETT, DUNNETTL, DUNNETTU,
GABRIEL, GT2, LSD, REGWQ, SCHEFFE, SIDAK,
SMM, SNK, T, TUKEY, WALLER

一方で，多重比較手法の使い方に関しては，その誤用や乱用について様々な意見がある。

- In my view, multiple comparison methods have no place at all in the interpretation of data (Nelder, 1971).
- SNK (Student-Newman-Keuls) 法や Duncan 法は，多重比較法としては正しい方法ではないので，使用してはならない。

著名な統計学者である Nelder の上記のコメントは，多重比較手法の使用に否定的な人々から，よく引用されている。

このように，多重比較手法には多くの手法が存在するとともに，様々な意見があるため，その使用法に関して混乱をまねきやすい。本講義では，多重比較法の本質的な考え方を理解することが期待されている。

多重比較法の標準的なテキストとしては，Miller (1981), Hochberg & Tamhane (1987), Hsu (1996) などがある。また和書では，永田・吉田 (1997)，坂巻ほか (2019) が多重比較法を取り扱っている。その他にも，書籍の一部として多重比較法が解説されている（丹後・松井, 2018）。

1.2　多重比較の例 – 対比較

例 1　大麦収量比較実験（一元配置乱塊法）

多重比較の例として，有名な **Duncan (1955)** の大麦収量比較実験のデータを表 **1.1** に示す。この実験では，$A_1, \ldots, A_7$ の **7** つの品種が，**6** 反復の乱塊法 **(randomised block design)** で配置されている。実は **Duncan (1955)** の論文では，処理平均値（各品種の平均収量）と分散分析表のみが与えられている。表 **1.1** のデータは，同じ処理平均値・分散分析表を持つように，乱数を発生させて作成したものである（**bushel** は体積の単位で，**1 bushel $\approx$ 35.24 liters; 1 bushel = 48 pounds (barley) $\approx$ 21.77kg; 1 acre $\approx$ 0.4047 ha; 1 bushel/acre $\approx$ 53.80 kg/ha**）。

表 **1.1**　**Barley yields (bushels/acre)**

	Varieties							
Block	A_1	A_2	A_3	A_4	A_5	A_6	A_7	Mean
R_1	39	63	66	47	80	55	47	56.7
R_2	64	63	69	58	73	61	79	66.7
R_3	57	82	59	62	79	63	56	65.4
R_4	42	67	56	64	55	62	57	57.6
R_5	49	65	85	75	73	56	58	65.9
R_6	47	87	71	63	68	51	69	65.1
Mean	49.7	71.2	67.7	61.5	71.3	58.0	61.0	62.9

その分散分析表を表 **1.2** に示す。品種間の効果は高度に有意 ($p = 0.002$) である。**R** では，出力 **1.3** のような分散分析表が得られる。

表 **1.2**　**Analysis of Variance**

変動因（要因）	自由度	平方和	平均平方（分散）	F 比	p 値
品種 A	6	2201.29	366.88	4.60**	0.0020
ブロック R	5	709.62	141.92	1.78	0.1470
誤差 E	30	2390.71	79.69		
全体	41	5301.62			

出力 **1.3**　**R** による分散分析の計算

```
> ### Barley data in Randomised Blocks from Duncan (1955).
> ## yield
> y <- c(39, 63, 66, 47, 80, 55, 47,
+        64, 63, 69, 58, 73, 61, 79,
+        57, 82, 59, 62, 79, 63, 56,
+        42, 67, 56, 64, 55, 62, 57,
+        49, 65, 85, 75, 73, 56, 58,
+        47, 87, 71, 63, 68, 51, 69)
## generate levels (gl() function generates factor levels)
> block <- gl(n=6, k=7, labels=paste("R", 1:6, sep=""))
> var <- gl(n=7, k=1, length=42, labels=paste("A", 1:7, sep=""))
> ## data frame
> barley.dat <- data.frame(y, block=block, var=var)
> ## ANOVA
> barley.aov <- aov(y ~ block+var, data=barley.dat)
> anova(barley.aov)
Analysis of Variance Table
```

```
Response: y
          Df  Sum Sq Mean Sq F value   Pr(>F)
block      5  709.62  141.92  1.7809 0.147021
var        6 2201.29  366.88  4.6038 0.002002 **
Residuals 30 2390.71   79.69
---
Signif. codes:  0 ’***’ 0.001 ’**’ 0.01 ’*’ 0.05 ’.’ 0.1 ’ ’ 1
```

分散分析の F 検定では，全ての処理が等しいかどうか（帰無仮説 $H^0_{1\cdots7}\colon \mu_1 = \cdots = \mu_7$）が検定される。それに対して，処理の効果の違いを詳細に検討するときによく使われる手法が対比較（ついひかく，**pairwise comparisons**）である。対比較とは，処理の全てのペアに対して，処理効果が等しいかどうか比較を行なう手法である。たとえば a とおりの処理があるときには，${}_aC_2 = a(a-1)/2$ とおりのペアを考えることになる。表 **1.1** の例では，$a = 7$ であるから

$$H^0_{12}\colon \mu_1 = \mu_2, H^0_{13}\colon \mu_1 = \mu_3, \cdots, H^0_{67}\colon \mu_6 = \mu_7$$

の ${}_7C_2 = 21$ 回の比較を行なうことになる。なお，対比較のための手法については，後で詳しく説明する。

注. この実験計画は，**乱塊法** (**randomised block design**) とよばれ，**Fisher** の **3** 原則（反復，無作為化，局所管理）を満たす方法である。農業実験で広く使われている。まず圃場（ほじょう，実験農地）全体を，いくつかのブロックに分け，各ブロック内で比較したい処理をランダムに配置する。ブロック間に大きな差があっても，処理の比較には影響しない。

A_5	A_2	A_3	A_7	A_6	A_1	A_4

ブロック **1** (R_1)

A_4	A_6	A_3	A_2	A_7	A_1	A_5

ブロック **2** (R_2)

A_5	A_3	A_7	A_4	A_6	A_2	A_1

ブロック **3** (R_3)

A_3	A_1	A_4	A_6	A_7	A_5	A_2

ブロック **4** (R_4)

A_4	A_1	A_7	A_2	A_6	A_3	A_5

ブロック **5** (R_5)

A_6	A_2	A_7	A_4	A_1	A_3	A_5

ブロック **6** (R_6)

図 **1.2**　一元配置乱塊法実験の配置

注. Duncan の **1955** 年の多重比較に関する論文

Duncan, D. B. (1955): Multiple range and multiple F tests, *Biometrics*, 11, 1–42.
の **Google Scholar** での被引用数は **46,574**（**2025** 年 **12** 月 **4** 日）である。**2020** 年以降も数多く引用されている。

1.3　多重比較とは

多重比較 (**multiple comparisons**) とは, その名のとおり一組の実験データに対して複数回の比較を行なうことである（図 **1.3**)。あるいは検定を複数回行なうので**多重検定** (**multiple tests**) とよばれることもある。多重比較を行なうための統計的手法が**多重比較法** (**multiple comparison procedures**) である。

注. 文献によっては，多水準間の対比較 (**pairwise comparisons**) を指して多重比較とよんだり，ファミリー単位過誤率を制御する手法のみを多重比較法とよぶことがあるので注意が必要である。本講義では，図 **1.3** のように，一組の実験データに対して複数の比較（検定）を行なう手法を多重比較法とよぶ。

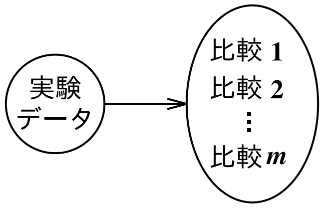


図 1.3　多重比較とは

検定を複数回行なう場面は多い。たとえば，次のような問題がある。

- 多水準間（多群）の比較（歴史的には，この問題が取り上げられた）
- 複数の評価項目（複数のエンドポイントを評価）
- 経時測定データ（複数の時点での検定）
- グループ逐次デザイン（段階的に複数の検定を行なう）

本講義では，最初に多水準間の多重比較に関して解説する。

多水準間の比較に関しても，最初に統計パッケージの例で示したように，多くの多重比較法が提案されている。これらの手法は行なうべき比較のタイプによって表 1.4 のように整理される（各問題のタイプと手法は以下の節で解説する）。問題のタイプのよって使用する手法が異なっており，さらに問題のタイプは実験の目的によって決まるので，たとえば，Tukey 法・Dunnett 法・Scheffé 法のどれを使うかを迷うことはない。実験の目的が対照との比較であれば Dunnett 法を使うか，t 検定を行なうかのいずれかである。問題のタイプの中での各手法は，過誤率の制御の方法が異なっている。

表 1.4　多重比較の問題のタイプと手法

対比較	対照との比較	対比	要因効果の分散分析	回帰分析
最小有意差法	最小有意差法	t 検定	主効果	多項式回帰
Duncan 法	Dunnett 法（両側）	Scheffé 法	交互作用	多次元回帰
Newman-Keuls 法	逐次 Dunnett 法		直交表	
R-E-G-W 法	...	...	...	...
Tukey 法				
Tukey-Kramer 法				
Hayter 法	Dunnett 法（片側）	Bartholomew 法		
	（修正）Williams 法			
（Bartholomew 法）	（Bartholomew 法）			

注. テキストを作成するときの注意
テキストを作るとき，例えば表 1.1 の大麦収量データを一つだけ示して，次のように各種の多重比較手法を実行させると，初心者はどの手法を使えばよいのか混乱してしまう。

(1) 表 1.1 のデータに対して，Tukey 法を使って全てのペアの対比較を行なえ。
(2) 表 1.1 において処理 A_1 が対照品種であるとして Dunnett の検定を行なえ。
(3) 品種群 $\{A_1, A_2, A_3\}$ と $\{A_4, A_5, A_6, A_7\}$ とは親の系統が異なっている。対比

$$\frac{\mu_1 + \mu_2 + \mu_3}{3} - \frac{\mu_4 + \mu_5 + \mu_6 + \mu_7}{4}$$

に対して Scheffé の検定を実行せよ。

このような問題の与え方ではなく，各手法（Tukey 法，Dunnett 法，Scheffé 法）ごとに適切なデータセットを用意すべきである。

1.4　処理平均の分散分析モデル

因子 A は a 個の水準 $A_1, \ldots, A_a$ をもつとする（なお一元配置実験においては水準と処理とは同じ意味になる）。各処理平均を

$$\bar{y}_{i\cdot} \sim N\Big(\mu_i,\ \frac{\sigma^2}{n_i}\Big) \quad (i = 1, \ldots, a) \tag{1.1}$$

とする。n_i は処理 A_i の繰返し数である。反復数 r の乱塊法においては繰返し数は等しく $n_i \equiv n = r$ となる。$\bar{y}_{i\cdot}$ は n_i 個のデータの平均なので，分散は σ^2/n_i となる。誤差分散 σ^2 は，分散分析表において誤差の平均平方 V_e の欄から

$$\hat{\sigma}^2 = V_e = S_e/\nu_e \sim \sigma^2\chi^2(\nu_e)/\nu_e \tag{1.2}$$

と推定される（ν_e は誤差分散の自由度である）。完全無作為化法か乱塊法かに応じて，それぞれの分散分析表の誤差の行の平均平方から得られる。なお，本講義では，確率変数と実現値を同じ大文字または小文字で表わす。

表 1.1 の乱塊法による大麦品種比較実験の例（分散分析表は表 1.2）では，

$$a = 7,\ n_i \equiv n = 6,$$
$$\bar{y}_{1\cdot} = 49.7,\ \bar{y}_{2\cdot} = 71.2,\ \bar{y}_{3\cdot} = 67.7,\ \bar{y}_{4\cdot} = 61.5,\ \bar{y}_{5\cdot} = 71.3,\ \bar{y}_{6\cdot} = 58.0,\ \bar{y}_{7\cdot} = 61.0,$$
$$\hat{\sigma}^2 = V_e = 79.69,\ \nu_e = 30$$

である。

2　仮説のファミリーと過誤率

2.1　第 I 種の過誤と第 II 種の過誤

まず，$a = 2$ の場合（比較が 1 つの場合）について，

$$\begin{cases} \text{帰無仮説 } H^0:\ \ \mu_1 = \mu_2 \\ \text{対立仮説 } H^A:\ \ \mu_1 \neq \mu_2 \end{cases}$$

に対する統計的検定における過誤を考える。統計的仮説検定においては，次の 2 とおりの過誤が考えられる。

- **第 I 種の過誤 (Type I error):** 実際には差がないのに（帰無仮説 H^0 が成り立つときに），有意差有りと判定する。
- **第 II 種の過誤 (Type II error):** 実際には差があるのに（対立仮説 H^A が成り立つときに），有意差無しと判定する。

この両方の過誤（の確率）を同時に小さくすることは不可能で，片方の過誤の確率を小さくすれば，他方の過誤の確率は高くなる。統計的仮説検定においては，第 I 種の過誤率を α（例えば 5%）以下に保ったうえで，第 II 種の過誤率が小さくなる方法を求める（第 II 種の過誤率を β と表わすと，$1 - \beta$ は検出力 (power) とよばれ，第 II 種の過誤率 β が小さい手法は検出力が高いと表現される）。

処理平均と分散の推定値が分散分析により (1.1) 式と (1.2) 式で与えられている場合は

$$|t| = \frac{|\bar{y}_{1\cdot} - \bar{y}_{2\cdot}|}{\hat{\sigma}\sqrt{1/n_1 + 1/n_2}} > t(\nu_e; \alpha/2) \tag{2.1}$$

のときに帰無仮説 H^0: $\mu_1 = \mu_2$ を棄却する方式が最良な検定方式（第 I 種の過誤を α 以下に保ち，あらゆる対立仮説のもとで検出力が最大になる検定方式）であることが数理統計学の理論から知られている。ここで，$t(\nu_e; \alpha/2)$ は自由度 ν_e の t 分布の片側 $\alpha/2$ 点（両側 α 点）である。

なお，第 III 種の過誤として

- **第 III 種の過誤 (Type III error):** 実際の差の符号を反対に判定してしまう。

を考えることもある。(2.1) 式の検定方式の場合，

$$t = \frac{\bar{y}_{1\cdot} - \bar{y}_{2\cdot}}{\hat{\sigma}\sqrt{1/n_1 + 1/n_2}}$$

に対して，

$$\begin{array}{lcl} t > t(\nu_e; \alpha/2) & \Rightarrow & \mu_1 - \mu_2 > 0 \text{ と判定} \\ |t| \leq t(\nu_e; \alpha/2) & \Rightarrow & H^0 \text{ を受容} \\ t < -t(\nu_e; \alpha/2) & \Rightarrow & \mu_1 - \mu_2 < 0 \text{ と判定} \end{array} \tag{2.2}$$

のような判定方式を採用すると，第 III 種の過誤が生じる可能性がある。つまり，対立仮説として H^{A+}: $\mu_1 - \mu_2 > 0$ が成り立っているにもかかわらず，$\mu_1 - \mu_2 < 0$ と判定してしまう確率はゼロではない。

演習問題 1　(2.2) 式の判定方式において，第 III 種の過誤の確率は $\alpha/2$ 以下になることを確認せよ。 □

2.2　仮説と検定のファミリー

一般に m 個の未知パラメータ（母数）$\theta_1, \ldots, \theta_m$ に興味があり，m 個の帰無仮説と対立仮説

$$\left\{ \begin{array}{l} \text{帰無仮説 } H_i^0\text{:}\ \ \theta_i = \theta_i^0 \\ \text{対立仮説 } H_i^A\text{:}\ \ \theta_i \neq \theta_i^0 \end{array} \right. \quad (i = 1, \ldots, m)$$

を考える。θ_i^0 $(i = 1, \ldots, m)$ は既知の定数である。ここで m 個の帰無仮説の集合

$$\mathcal{H} = \{H_1^0, \ldots, H_m^0\}$$

を**帰無仮説のファミリー** (family) とよぶ。

たとえば a 水準の一元配置分散分析において，a 個の母平均 $\mu_1, \ldots, \mu_a$ の全てのペア μ_i, μ_j $(i \neq j)$ に対して値が等しいかどうかを検定する場合（対比較，第 4 節参照），検定の数は $m = {}_aC_2 = a(a-1)/2$ であり，m 個の帰無仮説は

$$H_1^0\text{: } \theta_1 = \mu_1 - \mu_2 = 0, \ldots, H_m^0\text{: } \theta_m = \mu_{a-1} - \mu_a = 0$$

と表わされる $(\theta_1^0 = \cdots = \theta_m^0 = 0)$。

各帰無仮説 H_i^0: $\theta_i = \theta_i^0$ に対して，観測データ $\mathbb{y}$ に基づく検定結果を

$$\phi_i(\mathbb{y}) = \left\{ \begin{array}{ll} 1, & \text{帰無仮説 } H_i^0 \text{ を棄却} \\ 0, & \text{帰無仮説 } H_i^0 \text{ を受容} \end{array} \right.$$

と表わし，その全体

$$\phi_{\mathcal{H}}(\mathbb{y}) = \{\phi_1(\mathbb{y}), \ldots, \phi_m(\mathbb{y})\}$$

を**検定のファミリー**とよぶ。

注.　本講義資料では，スライドでの可視性を考慮して，アルファベットを太字（ボールド）で表示している。ベクトルを表わすためには，$\mathbb{y}$ や $\mathbb{c}$ などの黒板太字 (blackboard bold) を用いる。

2.3 部分帰無仮説と完全帰無仮説

多重比較では複数の帰無仮説 $H_1^0, \ldots, H_m^0$ が対象となる。どの帰無仮説が真に成り立っているかは未知なので，過誤率の評価においては，帰無仮説および対立仮説の様々な場合を想定しておく必要がある。実際に成り立っているのは m 個の帰無仮説のうち一部分だけの可能性がある。そこで，添え字の部分集合 $W \subset \{1, \ldots, m\}$ を考え，W に含まれる添え字をもつ帰無仮説のみが成り立っているという状況を

$$H_W^0 = \bigcap_{i \in W} H_i^0 \tag{2.3}$$

で表わし，**部分帰無仮説** (partial null hypothesis) とよぶ。特に $W = \{1, \ldots, m\}$ の場合，すなわちファミリー $\mathcal{H}$ の全ての帰無仮説が成り立っている場合を**完全帰無仮説** (complete null hypothesis) とよぶ。実際にはパラメータ θ_i $(i = 1, \ldots, m)$ の値は未知なので，どの部分帰無仮説が成り立っているのかは未知である。

2.4 多重比較における過誤率

特定の帰無仮説 H_i^0 に対しては，通常の仮説検定と同様に，H_i^0 が成り立っているときに H_i^0 を棄却すること $(\phi_i(\mathbb{y}) = 1)$ を第 I 種の過誤 (Type I error) といい，H_i^0 が成り立っていないときに H_i^0 を受容すること $(\phi_i(\mathbb{y}) = 0)$ を第 II 種の過誤 (Type II error) という。多重比較においても第 I 種の過誤を中心に考える場合が多い。

いま真に成り立っている帰無仮説を $H_W^0 = \bigcap_{i \in W} H_i^0$ とする。すなわち，集合 W に含まれる添え字をもつ帰無仮説 H_i^0 $(i \in W)$ が成り立っているとする。この成り立っている帰無仮説に対して，仮説を棄却する（$\phi_i(\mathbb{y}) = 1$ と判定する）と第 I 種の過誤をおかすことになる。m 回の検定のうち第 I 種の過誤の数を $V(\mathbb{y} \mid H_W^0)$ とすると，「$\phi_i(\mathbb{y}) = 1 \Leftrightarrow$ 帰無仮説 H_i^0 を棄却」に注意して，

$$V(\mathbb{y} \mid H_W^0) = \sum_{i \in W} \phi_i(\mathbb{y})$$

と表わされる。$V(\mathbb{y} \mid H_W^0)$ は観測データ $\mathbb{y}$ に依存するので確率変数である。このとき，

$$FWER = \Pr\{V(\mathbb{y} \mid H_W^0) > 0 \mid H_W^0\} \tag{2.4}$$

を**ファミリー単位過誤率** (family-wise error rate) という。すなわち $FWER$ は，真に成り立っている帰無仮説のうち，少なくとも 1 つ以上で第 I 種の過誤をおかす確率である。文献によっては，$FWER$ を**実験単位過誤率** (experiment-wise error rate) とよぶこともある（たとえば SAS）。

一方，個別の帰無仮説ごとに第 I 種の過誤の確率を考えた

$$CWER = \max_{i \in W} \Pr\{\phi_i(\mathbb{y}) = 1 \text{（仮説 } H_i^0 \text{ を棄却）} \mid H_i^0\} \tag{2.5}$$

を**比較単位過誤率** (comparison-wise error rate) という。$CWER$ では他の帰無仮説に関係なく，特定の帰無仮説 H_i^0 に着目して過誤の確率を考えている。

一般に (2.4) 式の確率は，成立している帰無仮説 H_i^0 $(i \in W)$ だけでなく，対立仮説 $H_j^A: \theta_j \neq \theta_j^0$ $(j \notin W)$ のパラメータの値 θ_j にも影響される。前述のように，実際に成立している帰無仮説，および対立仮説のパラメータの値は未知であるから，特定の多重比較法についての $FWER$ は，想定される全ての場合に対して (2.4) 式の上限として定義される。

3　ファミリー単位過誤率の制御

3.1　検定の多重性

比較単位過誤率 $CWER$ を一定値 α 以下にするためには，個々の検定 $\phi_i(\mathrm{y})$ を第 I 種の過誤率が α 以下になるように設計すればよい。しかし，個々の検定の過誤率が α 以下であっても，多数の検定を行なうと，そのどれかで過誤をおかしてしまう確率（ファミリー単位過誤率 $FWER$）は α よりも大きくなってしまう。このことを**検定の多重性** (multiplicity) という。ファミリー単位過誤率を制御する方法とは，この検定の多重性を考慮した方法である。$FWER$ を α 以下に保つことを，$FWER$ を保障する (guarantee, maintain)，あるいは制御する (control) と表現する。

(2.4) 式の $FWER$ の定義は仮定している帰無仮説に依存する。そして実際にどの帰無仮説が成立しているかは未知である。そこで，完全帰無仮説を含めて全ての部分帰無仮説のもとで $FWER$ を α 以下に保障することを，**強い意味で**制御する (strongly control) という。一方，完全帰無仮説のもとでのみ $FWER$ を α 以下に保障することを，**弱い意味で**制御する (weekly control) という。以下に，強い意味で $FWER$ を制御する方法を説明する。

事後的な仮説

データを見た後で事後的に仮説を設定して検定を行なうことには注意が必要である。たとえば表 1.1 の大麦品種比較実験の例では，データを見ると品種 A_5 の収量が最大で，品種 A_1 の収量が最小である。そこで，品種 A_1 と A_5 の比較を行なうことにすれば，仮に比較の回数は 1 回であるとしても検定結果は有意になりやすい。この場合，事前に想定される全てのペアの対比較を考慮して（全てのペアの対比較をファミリーと考えて），ファミリー単位過誤率の制御を考えればよい。

以下，強い意味で $FWER$ を制御するための方法として次の 3 つを考える。

- 同時信頼区間に基づく方法
- Bonferroni の方法
- 閉検定手順

3.2　同時信頼区間に基づく方法

m 個の母数 $\theta_1, \ldots, \theta_m$ に対する**同時信頼区間** (simultaneous confidence intervals) を $\theta_i \in [\hat{L}_i(\mathrm{y}),\ \hat{U}_i(\mathrm{y})]$ $(i = 1, \ldots, m)$ とする。すなわち，

$$\Pr\{\theta_i \in [\hat{L}_i(\mathrm{y}),\ \hat{U}_i(\mathrm{y})],\ i = 1, \ldots, m\} \geq 1 - \alpha \tag{3.1}$$

が成り立つように信頼区間 $[\hat{L}_i(\mathrm{y}),\ \hat{U}_i(\mathrm{y})]$ を構成する。ここで，下側の限界 $\hat{L}_i(\mathrm{y})$ と上側の限界 $\hat{U}_i(\mathrm{y})$ は観測データ y から計算される。

各帰無仮説 $H_i^0\colon \theta_i = \theta_i^0$ に対して，

$$\theta_i^0 \notin [\hat{L}_i(\mathrm{y}),\ \hat{U}_i(\mathrm{y})] \implies H_i^0 \text{ を棄却}$$

とすれば，$FWER$ は α 以下に保障される。なぜなら，いま成立している部分帰無仮説を $H_W^0 = \bigcap_{i \in W} H_i^0$ とする。同時信頼区間 (3.1) は全ての θ_i に対するものであるから，成り立っている帰無仮説について

$$\Pr\{\theta_i^0 \in [\hat{L}_i(\mathrm{y}),\ \hat{U}_i(\mathrm{y})],\quad i \in W\} \geq 1 - \alpha$$

である。したがって，どれかの $i \in W$ で $\theta_i^0 \notin [\hat{L}_i(\mathrm{y}),\ \hat{U}_i(\mathrm{y})]$ となる確率（すなわち間違って H_i^0 を棄却する確率）は α 以下となる。このことはどの部分帰無仮説においても成り立つので，この同時信頼区間に基づく検定方式は強い意味でファミリー単位過誤率 $FWER$ を制御する。

演習問題 2　上記の言葉による説明を，数式で表わせ。 □

同時信頼区間に基づく方法は，帰無仮説が棄却されたとき，次のようにパラメータの方向に関する判定も行なうことができる。

$$\begin{cases} \theta_i^0 < [\hat{L}_i(\mathrm{y}),\ \hat{U}_i(\mathrm{y})] & \Longrightarrow \quad \theta_i^0 < \theta_i \text{ と判定} \\ \theta_i^0 \in [\hat{L}_i(\mathrm{y}),\ \hat{U}_i(\mathrm{y})] & \Longrightarrow \quad H_i^0 \text{を受容} \\ [\hat{L}_i(\mathrm{y}),\ \hat{U}_i(\mathrm{y})] < \theta_i^0 & \Longrightarrow \quad \theta_i < \theta_i^0 \text{ と判定} \end{cases} \tag{3.2}$$

ここで $\theta_i^0 < [\hat{L}_i(\mathrm{y}),\ \hat{U}_i(\mathrm{y})]$ は，$[\hat{L}_i(\mathrm{y}),\ \hat{U}_i(\mathrm{y})]$ の要素がすべて θ_i^0 より大きいことを意味する。

演習問題 3　(3.2) 式によりパラメータの方向に関する判定を行なった場合，第 III 種の過誤に関してファミリー単位の過誤率が α 以下になることを確認せよ。 □

対比較における Tukey 法，対照との比較における Dunnett 法，対比に対する Scheffé 法は同時信頼区間に基づく方法である。

例 2　$x_1 \sim N(\theta_1,\ 1)$，$x_2 \sim N(\theta_2,\ 1)$，x_1 と x_2 とは独立とする。θ_1 と θ_2 について，個別の信頼区間は

$$\Pr\{\theta_1 \in [x_1 - 1.96,\ x_1 + 1.96]\} = 0.95$$
$$\Pr\{\theta_2 \in [x_2 - 1.96,\ x_2 + 1.96]\} = 0.95$$

である。

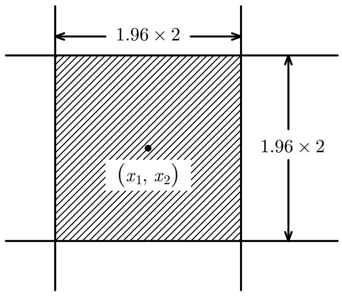


図 3.1　個別の信頼区間

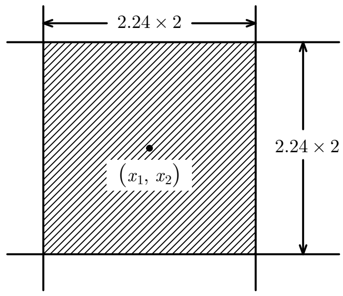


図 3.2　同時信頼区間

しかし，θ_1 と θ_2 を同時に含む確率は

$$\Pr\{\theta_1 \in [x_1 - 1.96,\ x_1 + 1.96] \quad \text{かつ}$$
$$\theta_2 \in [x_2 - 1.96,\ x_2 + 1.96]\} = 0.95 \times 0.95 = 0.9025$$

となり，0.95 より小さい。同時に含む確率を 0.95 とするにはどうすればよいか。

$0.975^2 = 0.95\ (0.975 = \sqrt{0.95})$

`qnorm(0.5*(1.0-sqrt(0.95)), lower.tail=FALSE)=2.236477`

より，同時信頼区間は

$$\Pr\{\theta_1 \in [x_1 - 2.24,\ x_1 + 2.24] \quad \text{かつ}$$
$$\theta_2 \in [x_2 - 2.24,\ x_2 + 2.24]\} = 0.975 \times 0.975 = 0.95$$

で与えられる。この同時信頼区間を利用して，帰無仮説

$$H_1^0\colon \theta_1 = 0$$

に対して，$x_1 > 2.24\ (x_1 - 2.24 > 0)$ となったときには，$\theta_1 > 0$ と判定してよい。θ_2 についても同様である。

[例 2 終]

3.3 Bonferroni の方法

m 個の任意の確率事象 $E_1, \ldots, E_m$ に対して，

$$\Pr\left(\bigcup_{i=1}^{m} E_i\right) \le \sum_{i=1}^{m} \Pr(E_i) \tag{3.3}$$

が成り立つ。この不等式を Bonferroni **の不等式**という（Bool の不等式とよばれることもある）。検定のファミリー $\phi_{\mathcal{H}}(\mathbb{y}) = \{\phi_1(\mathbb{y}), \ldots, \phi_m(\mathbb{y})\}$ の各検定において，第 I 種の過誤をおかすという事象を E_i とすると，(3.3) 式の左辺 $\Pr\left(\bigcup_{i=1}^{m} E_i\right)$ は，どれかの検定で第 I 種の過誤をおかす確率，すなわち，ファミリー単位過誤率 $FWER$ を表わす。ここで個々の検定に対して，第 I 種の過誤率を

$$\Pr(E_i \mid H_i^0) = \Pr\{\phi_i(\mathbb{y}) = 1 \mid H_i^0\} \le \alpha/m$$

となるように設計すれば，任意の部分帰無仮説 $H_W^0 = \bigcap_{i \in W} H_i^0$ に対して

$$FWER = \Pr\left(\bigcup_{i \in W} E_i \mid H_W^0\right) \le \sum_{i \in W} \Pr(E_i \mid H_i^0) \le \#\{W\} \times \frac{\alpha}{m} \le \alpha$$

が成り立ち $FWER$ が α 以下に保障される（$\#\{W\}$ は集合 W の要素数）。この方法を，検定の多重性に対する Bonferroni **の調整法**という。

演習問題 4　Bonferroni の不等式 (3.3) を証明せよ。

注.　ヒント: $m = 2, m = 3$ の場合。

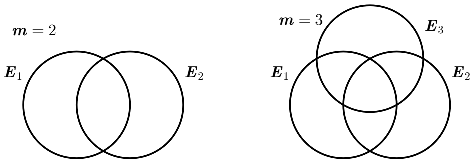


図 3.3　Bonferroni の不等式

$m = 2$ の場合を証明し，数学的帰納法を使えばよい。

□

各検定が独立な場合

事象 $E_1, \ldots, E_m$ が互いに独立な場合は，$\Pr\{E\} = 1 - \Pr\{E^c\}$ の関係を使って，

$$\begin{aligned}\Pr\left(\bigcup_{i=1}^{m} E_i\right) &= \Pr(E_1 \cup \cdots \cup E_m) = 1 - \Pr\{(E_1 \cup \cdots \cup E_m)^c\} \\ &= 1 - \Pr\{E_1^c \cap \cdots \cap E_m^c\} = 1 - \Pr\{E_1^c\} \times \cdots \times \Pr\{E_m^c\} \\ &= 1 - (1 - \Pr\{E_1\}) \times \cdots \times (1 - \Pr\{E_m\})\end{aligned} \tag{3.4}$$

が成り立つ。

各検定で同じ水準を用いて $\Pr\{E_i\} \equiv \alpha'$ とすると

$$\Pr(E_1 \cup \cdots \cup E_m) = 1 - (1 - \Pr\{E_i\})^m = 1 - (1 - \alpha')^m$$

であるから，

$$1-(1-\alpha')^m=\alpha$$

を逆に解いて

$$\alpha'=1-(1-\alpha)^{1/m}$$

とすれば，

$$FWER=\Pr(E_1\cup\cdots\cup E_m)=\alpha$$

となる（この方法は，Šidák **の方法**とよばれることもある)。

ここで，

$$\alpha/m\le 1-(1-\alpha)^{1/m} \tag{3.5}$$

が成り立つので，Bonferroni の方法より検出力が高くなる。

演習問題 5　(3.5) 式を確認せよ。 □

各事象が独立でない場合には，

$$\Pr(E_1\cup\cdots\cup E_m)<1-(1-\Pr\{E_1\})\times\cdots\times(1-\Pr\{E_m\})$$

となる場合や，

$$\Pr(E_1\cup\cdots\cup E_m)>1-(1-\Pr\{E_1\})\times\cdots\times(1-\Pr\{E_m\})$$

となる場合がある。

表 3.1　(1) 独立な事象

	E_2	E_2^c	
E_1	0.16	0.24	0.4
E_1^c	0.24	0.36	0.6
	0.4	0.6	1.0

表 3.2　(2) 同傾向の事象

	E_2	E_2^c	
E_1	0.3	0.1	0.4
E_1^c	0.1	0.5	0.6
	0.4	0.6	1.0

表 3.3　(3) 反対傾向の事象

	E_2	E_2^c	
E_1	0.1	0.3	0.4
E_1^c	0.3	0.3	0.6
	0.4	0.6	1.0

(1) 独立な事象

$$\Pr\{E_1\cup E_2\}=0.16+0.24+0.24=0.64=1-0.6^2$$

(2) 同傾向の事象

$$\Pr\{E_1\cup E_2\}=0.3+0.1+0.1=0.5<1-0.6^2=0.64$$

薬効試験

$$\begin{cases} E_1: & \text{解熱作用の検定} \\ E_2: & \text{頭痛軽減効果の検定} \end{cases}$$

(3) 反対傾向の事象

$$\Pr\{E_1\cup E_2\}=0.1+0.3+0.3=0.7>1-0.6^2=0.64$$

稲の品種改良試験

$$\begin{cases} E_1: & \text{収穫量の検定} \\ E_2: & \text{食味の検定} \end{cases}$$

したがって Šidák の方法は，各検定が独立であるかどうかに依存する。一方，(1), (2), (3) いずれの場合も Bonferroni の不等式は成り立つ。

$$\Pr\{E_1 \cup E_2\} \le \Pr\{E_1\} + \Pr\{E_2\} = 0.4 + 0.4 = 0.8$$

Bonferroni の不等式 (3.3) 式は，どのような事象 $E_1, \ldots, E_m$ に対しても成立する。したがって，どのような検定の問題にも適用できる（各検定が独立でなくてもよいし，異なる種類の検定手法が含まれていてもよい）。ただし検定の数 m が多くなると，α/m の値は極端に小さくなるので，個々の検定は非常に厳しい（有意差を検出しにくい保守的な）ものになる。

3.4　閉検定手順

帰無仮説のファミリー

$$\mathcal{H} = \{H_1^0, \ldots, H_m^0\}$$

に対して，全ての部分帰無仮説（および完全帰無仮説）$H_W^0 = \bigcap_{i \in W} H_i^0$ からなる集合

$$\bar{\mathcal{H}} = \{H_W^0 \mid W \subset \{1, \ldots, m\}\}$$

を $\mathcal{H}$ の**閉包** (closure) とよぶ。$\bar{\mathcal{H}}$ は共通部分の演算に関して閉じている。すなわち

$$H_W^0 \in \bar{\mathcal{H}},\ H_U^0 \in \bar{\mathcal{H}} \implies H_W^0 \cap H_U^0 \in \bar{\mathcal{H}}$$

が成り立つ。

次に各部分帰無仮説 H_W^0 に対し

$$\Pr\{\phi_W(\mathrm{y}) = 1 \mid H_W^0\} \le \alpha$$

となるように**水準 α の**検定方式 $\phi_W(\mathrm{y})$ を定める。特定の帰無仮説 $H_i^0 \in \mathcal{H}$ は，$i \in W$ なる全ての W に対して部分帰無仮説 H_W^0 が棄却された場合に棄却する。この方法を**閉検定手順** (closed testing procedure) という (Marcus, 1976)。

いま，真に成り立っている帰無仮説を $H_W^0 = \bigcap_{i \in W} H_i^0$ とする。H_i^0 $(i \in W)$ のうち，どれか 1 つでも間違って棄却すれば，ファミリー単位の第 I 種の過誤をおかしたことになる。ところが，H_i^0 $(i \in W)$ が棄却される場合には，必ず H_W^0 が棄却されていなければならない。閉検定手順の定義により，仮説 H_W^0 が成り立っているときに H_W^0 を棄却する確率は α 以下であるから，$FWER$ も α 以下になる。

演習問題 6　上記の言葉による説明を，数式で表わせ。　□

例 3　例 2 と同じ問題を考える。$x_1 \sim N(\theta_1, 1)$，$x_2 \sim N(\theta_2, 1)$，x_1 と x_2 とは独立である。興味のある帰無仮説を

$$H_1^0\colon \theta_1 = 0$$
$$H_2^0\colon \theta_2 = 0$$

とする。すなわち，仮説のファミリーは

$$\mathcal{H} = \{H_1^0,\ H_2^0\}$$

であり，その閉包は

$$\bar{\mathcal{H}} = \{H_1^0,\ H_2^0,\ H_1^0 \cap H_2^0\}$$

である。ここで $H^0_{\{1,2\}} = H^0_1 \cap H^0_2$ は H^0_1 と H^0_2 とが同時に成り立つという仮説（この例の場合は完全帰無仮説）であり，

$$H^0_{\{1,2\}} = H^0_1 \cap H^0_2\colon \theta_1 = 0,\ \theta_2 = 0$$

を表わしている。閉包 $\bar{\mathcal{H}}$ に含まれる 3 つの帰無仮説に対して，個別に次のような検定方式を考える。

$$\begin{aligned}
&H^0_1 \cap H^0_2\colon\ \max\{|x_1|, |x_2|\} > 2.24 \Longrightarrow H^0_1 \cap H^0_2 \text{ を棄却} \\
&H^0_1\colon \qquad |x_1| > 1.96 \qquad\qquad \Longrightarrow H^0_1 \text{ を棄却} \\
&H^0_2\colon \qquad |x_2| > 1.96 \qquad\qquad \Longrightarrow H^0_2 \text{ を棄却}
\end{aligned} \tag{3.6}$$

例 2 の計算から，帰無仮説 $H^0_1 \cap H^0_2$ に対して (3.6) 式が有意水準 5%の検定方式を与えることが分かる（$\Pr\{\max\{|x_1|, |x_2|\} > 2.24 \mid H^0_1 \cap H^0_2\} = 0.05$）。

閉検定手順では，

$$\begin{aligned}
&H^0_1 \cap H^0_2 \text{ を棄却},\ H^0_1 \text{ を棄却} \implies H^0_1 \text{ を棄却} \\
&H^0_1 \cap H^0_2 \text{ を棄却},\ H^0_2 \text{ を棄却} \implies H^0_2 \text{ を棄却}
\end{aligned}$$

により判定する。たとえば，

$$x_1 = 2.5,\ x_2 = 2.0$$

が観測されたとき，$H^0_1 \cap H^0_2$，H^0_1，H^0_2 の全てが棄却されるので，閉検定手順として H^0_1 と H^0_2 がともに棄却される。

同時信頼区間に基づく方法（例 2）では，

$$\begin{cases} \theta_1 \in [x_1 - 2.24,\ x_1 + 2.24] = [0.26,\ 4.74] \\ \theta_2 \in [x_2 - 2.24,\ x_2 + 2.24] = [-0.24,\ 4.24] \end{cases}$$

であるから，仮説 $H^0_2\colon \theta_2 = 0$ は棄却されない。

［例 3 終］

H^0_W に対する検定方式 $\phi_W(\mathrm{y})$ は，確率分布が仮定できる場合は，その確率分布に基づいて構成すればよい。例としては対比較における REGW 法がある。一方 W の要素の数を $w = \#\{W\}$ とすれば，H^0_W は w 個の帰無仮説が同時に成り立つという仮説であるから，Bonferroni の不等式により，各 H^0_i $(i \in W)$ を α/w の水準で検定することができる。この方法を Holm の方法という (Holm, 1979)。すべての検定を α/m 水準で行なう Bonferroni 法よりも，Holm の方法の方が検出力が高くなる（Holm の方法については，後節で説明する）。

閉検定手順に基づく方法は，信頼区間に基づく方法や Bonferroni の方法よりも，高い検出力をもつ。なおかつ $FWER$ を α 以下に制御することができる。ただし，パラメータの方向に関する判定を行なった場合に第 III 種の過誤率が制御されるかどうかは証明されていない（一部の問題に対しては第 III 種の過誤率が制御されることが証明されている）。

4　対比較

本節では，分散分析モデルにおける多水準間（多群間）の比較の問題として対比較を取り上げる。処理 A_i と A_j の全ての対（つい，ペア）を考えて，その母平均が等しいという帰無仮説

$$H^0_{ij}\colon \mu_i = \mu_j \quad (\mu_i - \mu_j = 0), \quad 1 \le i, j \le a$$

の検定（比較）を行なう方法を**対比較**（ついひかく，**pairwise comparisons**）という。ここで，記号 H_{ij}^0 の上付きの "**0**" は帰無仮説を表わし，下付きの "ij" は，i 番目の処理と j 番目の処理の比較を表わしている。帰無仮説のファミリーは

$$\mathcal{H} = \{H_{ij}^0 \colon \mu_i = \mu_j \ (\mu_i - \mu_j = 0), \quad 1 \le i, j \le a\} \tag{4.1}$$

であり，対象となる比較の数は $m = {}_aC_2 = a(a-1)/2$ である。

最初に繰返し数の等しい場合 $n_i \equiv n$ について説明する。例として，**Duncan** の大麦収量データ（表 **1.1**）：

$$a = 7,\ n = 6,$$
$$\bar{y}_{1\cdot} = 49.7,\ \bar{y}_{2\cdot} = 71.2,\ \bar{y}_{3\cdot} = 67.7,\ \bar{y}_{4\cdot} = 61.5,\ \bar{y}_{5\cdot} = 71.3,\ \bar{y}_{6\cdot} = 58.0,\ \bar{y}_{7\cdot} = 61.0,$$
$$\hat{\sigma}^2 = V_e = 79.69,\ \nu_e = 30,\ (\hat{\sigma} = \sqrt{V_e} = \sqrt{79.69} = 8.93)$$

を解析する。

4.1 最小有意差法（LSD 法）

計算手順

判定基準値

$$LSD(\alpha) = \hat{\sigma}\sqrt{2/n} \cdot t(\nu_e; \alpha/2) \tag{4.2}$$

を計算する。$t(\nu_e; \alpha/2)$ は自由度 ν_e の t 分布の片側 $\alpha/2$ 点（両側 α 点）である。`R` では `qt()` 関数を使って計算できる。大麦収量実験の例では，

$$a = 7,\ n = 6,\ \hat{\sigma} = 8.93,\ \nu = 30,\ t(30; 0.025) = 2.042$$

を使って，

$$LSD(0.05) = \hat{\sigma}\sqrt{2/n} \cdot t(\nu_e; \alpha/2) = 8.93 \times \sqrt{2/6} \times 2.042 = 10.53$$

となる。$LSD(\alpha)$ は**最小有意差** (**Least Significant Difference, LSD**) とよばれる。

ふたつの処理平均の差の絶対値が

$$|\bar{y}_{i\cdot} - \bar{y}_{j\cdot}| > LSD(\alpha) \tag{4.3}$$

のときに，処理 A_i と A_j に有意差有りと判定する。

多くの統計パッケージでこの計算手順を実行するプログラムが用意されている。統計パッケージを利用しない場合は次のように実行すればよい。処理平均を昇順に並べたものを，

$$\bar{y}_{(1)\cdot} \le \bar{y}_{(2)\cdot} \le \cdots \le \bar{y}_{(a)\cdot}$$

とする。外側の $\bar{y}_{(a)\cdot} - \bar{y}_{(1)\cdot}$ から始め，内側に向かって比較をすすめていく。途中で $\bar{y}_{(i)\cdot}$ と $\bar{y}_{(j)\cdot}$ $(i < j)$ が有意差無しと判定されれば，$i \le k \le j$ なる k については比較を継続する必要はない（その理由を説明せよ）。

表 **1.1** のデータに対する最小有意差法による判定結果は以下のとおりである。この表記法は，下線で結ばれた処理のあいだには，有意差がないことを示している。

A_1	A_6	A_7	A_4	A_3	A_2	A_5
49.7	58.0	61.0	61.5	67.7	71.2	71.3

LSD 法の特徴と考え方

(4.3) 式の判定方式は，

$$|\bar{y}_{i\cdot} - \bar{y}_{j\cdot}| > LSD(\alpha) \iff |t| = \frac{|\bar{y}_{i\cdot} - \bar{y}_{j\cdot}|}{\hat{\sigma}\sqrt{2/n}} > t(\nu_e; \alpha/2)$$

と書き直すことができる。すなわち LSD 法は，各比較に有意水準 α の t 検定を実行していることと同値である。したがって，比較単位過誤率 $CWER$ は α 以下に保障される。しかしペアの総数は $a(a-1)/2 = 7(7-1)/2 = 21$ であり，これら 21 回の検定のうち，どこかで間違って有意差有りと判定する確率，すなわちファミリー単位過誤率 $FWER$ は α より高くなる。

もし各検定が独立であるとすれば，(3.4) 式により，完全帰無仮説 $H^0_{1\cdots7}\colon \mu_1 = \cdots = \mu_7$ のもとで，ファミリー単位の過誤率は

$$FWER' = 1 - (1-0.05)^{21} = 1 - 0.95^{21} = 0.66$$

となる。実際には，(4.3) 式において各検定は独立とはならないので，正確なファミリー単位の過誤率はこの値とは異なる（各検定が独立でないことを確認せよ）。

演習問題 7　表 1.1 の大麦収量実験の例 ($a = 7, \nu_e = 30$) において，完全帰無仮説 $H^0_{1\cdots7}\colon \mu_1 = \cdots = \mu_7$ のもとでの正確なファミリー単位の過誤率 $FWER$ を計算せよ。 □

(4.2) 式において，$LSD(\alpha)$ ではなく，$LSD(\alpha/m)$ を用いれば，Bonferroni 法により，ファミリー単位過誤率 $FWER$ は α 以下に保障される（ただし，$m = a(a-1)/2$）。

演習問題 8　大麦収量実験の例 ($a = 7, \nu_e = 30$) において，Bonferroni 法で調整した水準による LSD 法の実行手順を示せ。また，完全帰無仮説 $H^0_{1\cdots7}\colon \mu_1 = \cdots = \mu_7$ のもとでのファミリー単位過誤率 $FWER$ を計算せよ。 □

保護付き LSD 法

まず最初に分散分析表の F 検定で，完全帰無仮説

$$H^0_{1\cdots a}\colon \mu_1 = \cdots = \mu_a$$

に対する検定を行ない，この完全帰無仮説が棄却された場合のみ，LSD 法を実行する方法を**保護つき LSD 法** (protected LSD, PLSD, Fisher's PLSD) という。一方，完全帰無仮説の検定を実行することなく，LSD 法による検定を行なう方法を**保護なし LSD 法** (unprotected LSD) という。

保護つき LSD 法は，完全帰無仮説のもとでは，ファミリー単位過誤率を α 以下に保つ（その理由を説明せよ）。しかし，その他の部分帰無仮説のもとでは必ずしもファミリー単位過誤率を保障しない。

演習問題 9　大麦収量実験の例 ($a = 7, \nu_e = 30$) において，品種 A_1 , ..., A_6 には全く差が無く，品種 A_7 のみ収量が大きく異なる，すなわち

$$\mu_1 = \cdots = \mu_6 \ll \mu_7$$

と仮定する。このとき，完全帰無仮説 $H^0_{1\cdots7}\colon \mu_1 = \cdots = \mu_7$ は，ほとんど確実に棄却されるものとする。この場合，保護つき LSD 法を 5%水準で実施した場合の $FWER$ を求めよ。 □

注.　**演習問題** 7 - 9 は次節で説明するスチューデント化した範囲の確率分布を利用する。

4.2 Tukey 法

計算手順

Tukey（テューキー）法では，LSD 法よりも厳しい判定基準値

$$HSD_a(\alpha) = (\hat{\sigma}/\sqrt{n}) \cdot q(a, \nu_e; \alpha) \tag{4.4}$$

を計算する。$q(a, \nu_e; \alpha)$ はスチューデント化した範囲の上側 α 点である。R では，`qtukey()` 関数で計算できる。$HSD_a(\alpha)$ は Honestly Significant Difference とよばれる。$HSD_a(\alpha)$ の値は処理の数 a に依存し，a が増えるほど値が大きくなる（判定が厳しくなる）。

判定手順は LSD 法と同じで

$$|\bar{y}_i. - \bar{y}_j.| > HSD_a(\alpha) \tag{4.5}$$

のときに，処理 A_i と A_j とに有意差有りと判定する。

表 1.1 の大麦品種比較試験の例では，判定基準値は

$$a = 7,\ n = 6,\ \hat{\sigma} = 8.92,\ \nu_e = 30,\ q(7, 30; 0.05) = 4.464$$
$$HSD_7(0.05) = 8.93/\sqrt{6} \times 4.464 = 16.27$$

となる。判定結果は以下のとおりである。

A_1	A_6	A_7	A_4	A_3	A_2	A_5
49.7	58.0	61.0	61.5	67.7	71.2	71.3

Tukey 法の特徴と考え方

Tukey 法は同時信頼区間に基づく方法（第 3.2 節）である。処理の全てのペアに対して母平均の差 $\mu_i - \mu_j$ の同時信頼区間

$$\begin{aligned}\mu_i - \mu_j &\in [\hat{L}_{ij}(\mathrm{y}),\ \hat{U}_{ij}(\mathrm{y})] \\ &= \left[\bar{y}_i. - \bar{y}_j. - HSD_a(\alpha),\ \bar{y}_i. - \bar{y}_j. + HSD_a(\alpha)\right] \quad (1 \le i, j \le a)\end{aligned} \tag{4.6}$$

を考える。この同時信頼区間に関して

$$\Pr\{\mu_i - \mu_j \in [\hat{L}_{ij}(\mathrm{y}),\ \hat{U}_{ij}(\mathrm{y})], \quad 1 \le i, j \le a\} = 1 - \alpha \tag{4.7}$$

が成り立つ。スチューデント化した範囲の上側 α 点 $q(a, \nu_e; \alpha)$ は，(4.7) 式が成り立つように定められている。

(4.5) 式の判定方式は，信頼区間がゼロを含まないことと同値，すなわち

$$|\bar{y}_i. - \bar{y}_j.| > HSD_a(\alpha) \quad \Longleftrightarrow \quad 0 \notin [\hat{L}_{ij}(\mathrm{y}),\ \hat{U}_{ij}(\mathrm{y})]$$

である（確認せよ）。したがって，Tukey 法は強い意味でファミリー単位過誤率 $FWER$ を α 以下に保障する。Tukey 法は，パラメータの同時信頼区間を与えるという点で有効な方法である。

スチューデント化した範囲

$u_1, \ldots, u_a, \chi^2$ が互いに独立に

$$u_i \sim N(0, 1), \quad 1 \le i \le a,$$
$$\chi^2 \sim \chi^2(\nu_e)$$

に従うとき，

$$Q = \max_{i,j} \frac{|u_i - u_j|}{\sqrt{\chi^2/\nu_e}} = \frac{\max_i u_i - \min_i u_i}{\sqrt{\chi^2/\nu_e}} \tag{4.8}$$

を**スチューデント化した範囲** (Studentized range) という。その上側 α 点 $q(a, \nu_e; \alpha)$ は，

$$\Pr\{Q > q(a, \nu_e; \alpha)\} = \alpha$$

で与えられる。

分散分析により，処理平均と分散の推定値が (1.1) 式と (1.2) 式

$$\bar{y}_{i\cdot} \sim N\Big(\mu_i,\ \frac{\sigma^2}{n}\Big) \quad (i = 1, \ldots, a)$$
$$\hat{\sigma}^2 = V_e = S_e/\nu_e \sim \sigma^2\chi^2(\nu_e)/\nu_e$$

で与えられている場合，

$$Q = \max_{i,j} \frac{|(\bar{y}_{i\cdot} - \mu_i) - (\bar{y}_{j\cdot} - \mu_j)|}{\hat{\sigma}/\sqrt{n}} \tag{4.9}$$

は，スチューデント化した範囲の分布に従う（確認せよ）。

演習問題 10　スチューデント化した範囲の定義 (4.8) 式と (4.9) 式から，全てのペアの同時信頼区間が (4.6) 式で与えられることを示せ。 □

注.　ヒント。

$$\begin{aligned}
1 - \alpha &= \Pr\Big\{ \max_{i,j} \frac{|(\bar{y}_{i\cdot} - \mu_i) - (\bar{y}_{j\cdot} - \mu_j)|}{\hat{\sigma}/\sqrt{n}} \le q(a, \nu_e; \alpha)\Big\} \\
&= \Pr\Big\{ \text{すべての } i,\ j \text{ で } \frac{|(\bar{y}_{i\cdot} - \mu_i) - (\bar{y}_{j\cdot} - \mu_j)|}{\hat{\sigma}/\sqrt{n}} \le q(a, \nu_e; \alpha)\Big\} \\
&= \Pr\Big\{ \text{すべての } i,\ j \text{ で } |(\bar{y}_{i\cdot} - \mu_i) - (\bar{y}_{j\cdot} - \mu_j)| \le HSD_a(\alpha)\Big\} \\
&= \Pr\Big\{ \text{すべての } i,\ j \text{ で } |(\bar{y}_{i\cdot} - \bar{y}_{j\cdot}) - (\mu_i - \mu_j)| \le HSD_a(\alpha)\Big\}
\end{aligned}$$

また，完全帰無仮説 $H^0_{1\cdots a}$： $\mu_1 = \cdots = \mu_a$ のもとでは，

$$Q = \max_{i,j} \frac{|\bar{y}_{i\cdot} - \bar{y}_{j\cdot}|}{\hat{\sigma}/\sqrt{n}} = \frac{\max_i \bar{y}_{i\cdot} - \min_i \bar{y}_{i\cdot}}{\hat{\sigma}/\sqrt{n}}$$

がスチューデント化した範囲の分布に従うことから，

$$\max_{i,j} |\bar{y}_{i\cdot} - \bar{y}_{j\cdot}| > HSD_a(\alpha) \implies H^0_{1\cdots a}\text{: } \mu_1 = \cdots = \mu_a \text{ を棄却}$$

とすれば，帰無仮説 $H^0_{1\cdots a}$： $\mu_1 = \cdots = \mu_a$ に対する水準 α の検定方式が与えられる。

注.　**演習問題** 7 – 9 のヒント。

完全帰無仮説 $H^0_{1\cdots a}$： $\mu_1 = \cdots = \mu_a$ のもとで，Tukey 法によるファミリー単位過誤率 $FWER$ は

$$FWER = \Pr\{ \text{どれかの } i,\ j \text{ で } |\bar{y}_{i\cdot} - \bar{y}_{j\cdot}| > HSD_a(\alpha)\}$$

$$
\begin{aligned}
&= \Pr\left\{ \max_{i,j} |\bar{y}_{i\cdot} - \bar{y}_{j\cdot}| > HSD_a(\alpha) \right\} \\
&= \Pr\left\{ \max_{i,j} \frac{|\bar{y}_{i\cdot} - \bar{y}_{j\cdot}|}{\hat{\sigma}/\sqrt{n}} > q(a, \nu_e; \alpha) \right\} \\
&= \Pr\{Q > q(a, \nu_e; \alpha)\} = \alpha
\end{aligned}
$$

である。しかし，判定基準値として $HSD_a(\alpha)$ ではなく，$LSD(\alpha)$ や，Bonferroni 法の $LSD(\alpha/m)$ を使えば，$FWER$ は異なる値になる。たとえば，LSD 法では

$$
\begin{aligned}
FWER &= \Pr\{\text{ どれかの } i,\ j \text{ で } |\bar{y}_{i\cdot} - \bar{y}_{j\cdot}| > LSD(\alpha)\} \\
&= \Pr\left\{ \max_{i,j} |\bar{y}_{i\cdot} - \bar{y}_{j\cdot}| > LSD(\alpha) \right\} \\
&= \Pr\left\{ \max_{i,j} \frac{|\bar{y}_{i\cdot} - \bar{y}_{j\cdot}|}{\hat{\sigma}/\sqrt{n}} > \boxed{\quad\text{ア}\quad} \right\} \\
&= \Pr\{Q > \boxed{\quad\text{ア}\quad}\} = \boxed{\quad\text{イ}\quad}
\end{aligned}
$$

となる。スチューデント化した範囲の上側確率は，たとえば R では

ptukey(q, nmeans=a, df=ν_e, lower.tail=FALSE)

により計算できる。

4.3 REGWQ 法

対比較においてファミリー単位過誤率を制御する方法として，第 **3.4** 節の閉検定手順を利用した **REGWQ 法**について説明する。

計算手順

○ 判定基準値の計算

$a - 1$ 個の判定基準値 R_p $(p = 2, \ldots a)$ を計算する。

$$
R_p = \frac{\hat{\sigma}}{\sqrt{n}} \cdot q(p, \nu_e; \alpha_p), \quad p = 2, \ldots, a \tag{4.10}
$$

$$
\alpha_p = \begin{cases} 1 - (1 - \alpha)^{p/a} & p < a - 1 \\ \alpha & p = a - 1, a \end{cases}
$$

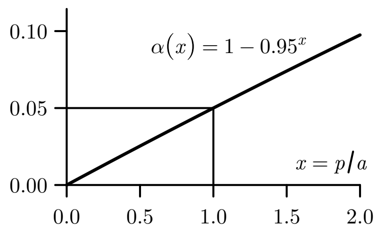


図 **4.1**　$1 - 0.95^x$ のグラフ

$q(p, \nu_e; \alpha_p)$ は p 個の処理平均に対するスチューデント化した範囲の上側 α_p 点である。$p < a-1$ に対しては，$\alpha_p < \alpha$ となる。ここで

$$
R_2 \leq R_3 \leq \cdots \leq R_a \tag{4.11}
$$

となるように調整する。すなわち，(**4.10**) 式の計算により，$R_p > R_{p+1}$ となった場合は，R_{p+1} を R_p で置き換える $(p = 2, \ldots, a-1)$。付表 **9.1** に，$\alpha = 0.05$ に対する $q(p, \nu_e; \alpha_p)$ の値を与える。ただし表 **9.1** では，$q(2, \nu_e; \alpha_2) \leq q(3, \nu_e; \alpha_3) \leq \cdots \leq q(k, \nu_e; \alpha_a)$ となるように調整されている。また，$p = a$ のときは $R_p = R_a = HSD_a(\alpha)$ である。

○ 判定手順 1

母平均 μ_i $(i = 1, \ldots, a)$ の添え字の部分集合 $P = \{i_1, i_2, \ldots, i_p\} \subset \{1, \ldots, a\}$ を考え，P に含まれる添え字をもつ母平均が全て等しいという仮説を

$$
H^0(P)\text{：}\ \mu_i = \mu_j, \quad i, j \in P \tag{4.12}
$$

と表わす。たとえば $P = \{1, 2, 5, 7\}$ の場合，$H^0(P)$：$\mu_1 = \mu_2 = \mu_5 = \mu_7$ である。

注.　この帰無仮説は，2 重添え字が許されるならば

$$H^0(P)\colon\ \mu_{i_1} = \mu_{i_2} = \cdots = \mu_{i_p}$$

と書いてもよい。

仮説 $H^0(P)$ を (4.10) 式の判定基準 R_p を用いて，Tukey 法と同じように検定する。すなわち，

$$\max_{i,\,j \in P} |\bar{y}_{i\cdot} - \bar{y}_{j\cdot}| > R_p \tag{4.13}$$

のときに，仮説 $H^0(P)$ を棄却する。

そして，特定のペアの仮説 $H^0_{ij}\colon \mu_i = \mu_j$ については，$\{i, j\} \subset P \subset \{1, \ldots, a\}$ となる全ての部分集合 P に対して，仮説 $H^0(P)$ が (4.13) 式の検定で棄却されたときに，対象となる仮説 $H^0_{ij}\colon \mu_i = \mu_j$ を棄却する。

○ **判定手順 2**

繰返し数が等しい場合 ($n_i \equiv n$) は，次のように計算すればよい。LSD 法と同様に，昇順に処理平均を並べたものを

$$\bar{y}_{(1)\cdot} \le \bar{y}_{(2)\cdot} \le \cdots \le \bar{y}_{(a)\cdot}$$

とし，外側から順次，内側へ向かって検定する。

$$\begin{array}{lll}
\bar{y}_{(a)\cdot} - \bar{y}_{(1)\cdot} & > R_a & \Longrightarrow \text{有意差有り} \\
\bar{y}_{(a)\cdot} - \bar{y}_{(2)\cdot} & > R_{a-1} & \Longrightarrow \text{有意差有り} \\
\vdots & \vdots & \vdots \\
\bar{y}_{(a)\cdot} - \bar{y}_{(a-1)\cdot} & > R_2 & \Longrightarrow \text{有意差有り} \\
\bar{y}_{(a-1)\cdot} - \bar{y}_{(1)\cdot} & > R_{a-1} & \Longrightarrow \text{有意差有り} \\
\vdots & \vdots & \vdots \\
\bar{y}_{(2)\cdot} - \bar{y}_{(1)\cdot} & > R_2 & \Longrightarrow \text{有意差有り}
\end{array}$$

一般的に表わせば，$\bar{y}_{(j)\cdot} - \bar{y}_{(i)\cdot} > R_{j-i+1}$ $(i < j)$ のとき，処理 $A_{(i)}$ と $A_{(j)}$ に有意差有りと判定する。ただし，$\bar{y}_{(i)\cdot}$ と $\bar{y}_{(j)\cdot}$ $(i < j)$ が有意差無しと判定されれば，$i \le k \le j$ なる $\bar{y}_{(k)\cdot}$ については，いずれも有意差無しと判定する（$A_{(i)}$ は $\bar{y}_{(i)\cdot}$ に対応する処理を表わす)。

異なる R_p の値を用いて判定するので，$i \le k < \ell \le j$ なる k, ℓ に対して，

$$\bar{y}_{(j)\cdot} - \bar{y}_{(i)\cdot} \le R_{j-i+1}$$
$$\bar{y}_{(\ell)\cdot} - \bar{y}_{(k)\cdot} > R_{\ell-k+1}$$

となることがある。しかし，外側の $\bar{y}_{(i)\cdot}$ と $\bar{y}_{(j)\cdot}$ との間に有意差無しと判定するので，その間にある処理間には有意差無しと判定する（すなわち，$\bar{y}_{(k)\cdot}$ と $\bar{y}_{(\ell)\cdot}$ は有意差有りとは判定しない)。

注.　（**演習問題**）判定手順 1 と判定手順 2 は同じ結果を与えることを示せ。

処理平均を昇順に並べたときの $\bar{y}_{(i)\cdot}$ と $\bar{y}_{(j)\cdot}$ に着目する:

$$\bar{y}_{(1)\cdot} \le \cdots \le \bar{y}_{(k)\cdot} \le \cdots \le \bar{y}_{(i)\cdot} \le \cdots \le \bar{y}_{(j)\cdot} \le \cdots \le \bar{y}_{(\ell)\cdot} \le \cdots \le \bar{y}_{(a)\cdot}$$

判定手順 2 で $\bar{y}_{(i)\cdot}$ と $\bar{y}_{(j)\cdot}$ とに有意差有りと判定されるのは

$$k \le i < j \le \ell$$

なる全ての k と ℓ に関して

$$\bar{y}_{(\ell)\cdot} - \bar{y}_{(k)\cdot} > R_{\ell-k+1} \tag{4.14}$$

となる場合である。

$\bar{y}_{(i)\cdot}$ と $\bar{y}_{(j)\cdot}$ を含む任意の集合

$$\{\bar{y}_{(i)\cdot},\ \bar{y}_{(j)\cdot},\ \ldots\} \subset \{\bar{y}_{1\cdot}, \ldots, \bar{y}_{a\cdot}\}$$

に対して，判定手順 1 により (4.13) が成り立つときは，明らかに (4.14) 式が成り立つ（判定手順 2 が成り立つ）。逆に，判定手順 2 により (4.14) 式が成り立っている場合には，判定基準値 R_p の単調性（(4.11) 式）により，$\bar{y}_{(i)\cdot}$ と $\bar{y}_{(j)\cdot}$ を含む任意の集合 $\{\bar{y}_{(i)\cdot},\ \bar{y}_{(j)\cdot},\ \ldots\}$ に対して，(4.13) が成り立つ。 □

大麦収量実験の例（表 1.1）では，

$$a = 7,\ n_i \equiv n = 6,\ \hat{\sigma} = 8.93,\ \nu_e = 30,$$

および，付表 9.1 を使って，$a - 1 = 6$ 個の判定基準値は

p	2	3	4	5	6	7
$q(p, \nu_e; \alpha_p)$	3.668	4.003	4.184	4.305	4.305	4.464
R_p	13.37	14.59	15.25	15.69	15.69	16.27

である。判定結果は，

A_1	A_6	A_7	A_4	A_3	A_2	A_5
49.7	58.0	61.0	61.5	67.7	71.2	71.3

となり，Tukey 法と同じである。

REGWQ 法の特徴と考え方

REGWQ 法は閉検定手順（第 3.4 節）に基づく方法である。対比較の帰無仮説のファミリー

$$\mathcal{H} = \{H_{ij}^0\colon \mu_i = \mu_j\ (1 \leq i, j \leq a)\}$$

に対して，その要素の共通部分全体からなる閉包集合 $\bar{\mathcal{H}}$ を考える。$\bar{\mathcal{H}}$ の要素は a 個の母平均に関して，いくつかが等しいという仮説になる。たとえば，$a = 7$ の場合，

$$\mu_1 = \mu_2 = \mu_5 = \mu_7, \quad \mu_3 = \mu_4 = \mu_6$$

のような部分帰無仮説が $\bar{\mathcal{H}}$ の要素となる。一般的に $P_1, \ldots, P_g$ を添え字 $\{1, \ldots, a\}$ の互いに素な部分集合の組とし，同じ P_k $(k = 1, \ldots, g)$ に含まれる添え字をもつ母平均は等しいという仮説を $H^0(P_1, \ldots, P_g)$ と表わす。上記の例の場合，$P_1 = \{1, 2, 5, 7\}$，$P_2 = \{3, 4, 6\}$ とすると，

$$H^0(P_1, P_2)\colon\ \mu_1 = \mu_2 = \mu_5 = \mu_7, \quad \mu_3 = \mu_4 = \mu_6$$

である。

$\bar{\mathcal{H}}$ の要素 $H^0(P_1, \ldots, P_g)$ に関して，個々の部分集合に対する仮説 $H^0(P_k)$ を，(4.10) 式の有意水準で検定すれば，$H^0(P_1, \ldots, P_g)$ に対する検定の有意水準は α 以下に保たれる。たとえば，$P_1 = \{1, 2, 5, 7\}$，$P_2 = \{3, 4, 6\}$ の場合，$H^0(P_1)$ と $H^0(P_2)$ に関する検定を，それぞれ

$$\alpha_4 = 1 - (1 - \alpha)^{4/7}$$
$$\alpha_3 = 1 - (1 - \alpha)^{3/7}$$

の水準で検定すれば，$H^0(P_1, P_2)$ の検定に対する水準は

$$1-(1-\alpha_4)(1-\alpha_3)=1-(1-\alpha)^{4/7}\cdot(1-\alpha)^{3/7}=1-(1-\alpha)=\alpha$$

となる。ここで，$H^0(P_1)$ に対する検定と $H^0(P_2)$ に対する検定は正の相関をもつので，Šidák の方法が適用できる (Einot & Gabriel, 1975)。

注. $\#\{P\}=p=a-1$ のときは，$H^0(P)$ の検定は水準 α で検定してよい（(4.10) 式)。なぜなら，$\#\{P\}=p=a-1$ とすると，残りの部分は単一の要素 $\{1,\ldots,a\}\setminus P=\{i\}$ となり，第 I 種の過誤は生じないからである。

演習問題 11　$a=3$ のとき，完全帰無仮説 $H^0_{123}\colon \mu_1=\mu_2=\mu_3$ が棄却された後は，各対比較 $H^0_{ij}\colon \mu_i=\mu_j$ を有意水準 α で実行してよいことを説明せよ。すなわち $a=3$ の場合は，保護付き LSD 法はファミリー単位過誤率 $FWER$ を α 以下に保障する。
$a=4$ の場合はどうか。 □

注. Ryan (1960) は，α_p として，Bonferroni の不等式を利用した

$$\alpha_p=\alpha\cdot(p/a)$$

を提案した。その後 Einot & Gabriel (1975) は，少し検出力が高くなる

$$\alpha_p=1-(1-\alpha)^{p/a}$$

を提案した。そして，Welsch (1977) は，$p=a-1$ の場合は，$\alpha_{a-1}=\alpha$ でよいことを示した。これらの歴史により，この方法は Ryan-Einot-Gabriel-Welsch 法とよばれている（たとえば，Hsu (1996) や，SAS, SPSS などの統計パッケージ)。

ところが，(4.10) 式の α_p は Tukey (1953) の有名な未公刊の文献ですでにに紹介されていた。したがって，REGW 法は Tukey-Welsch 法とよばれることもある（例えば，Hochberg & Tamhane (1987), 永田ら (1997) など)。

さらに帰無仮説 $H^0(P)$ の検定にスチューデント化した範囲を使う方法を REGWQ 法という。帰無仮説 $H^0(P)$ の検定に F 検定を使うこともできる。その方法は REGWF 法とよばれる。

4.4　その他の方法（SNK 法，Duncan 法）

統計パッケージで利用できる対比較手法として，Student-Newman-Keuls 法や Duncan 法などがある。

Student-Newman-Keuls 法の計算手順

Student-Newman-Keuls 法（スチューデント-ニューマン-クールズ法，SNK 法）の計算手順は REGWQ 法と同様である。ただし $a-1$ 個の判定基準値として

$$R_p^{SNK}=\frac{\hat{\sigma}}{\sqrt{n}}\cdot q(p,\nu_e;\alpha),\quad p=2,\ldots,a \tag{4.15}$$

を用いる。REGWQ 法に比較して $q(p,\nu_e;\alpha)\le q(p,\nu_e;\alpha_p)$ であるから，SNK 法の方が REGWQ 法よりも判定基準値が小さくなる。すなわち有意差有りと判定する確率が高くなる。

大麦品種比較実験（表 1.1）のデータについての判定結果は REGWQ 法，Tukey 法と同じである。

Student-Newman-Keuls 法の特徴と考え方

SNK 法では，想定される部分帰無仮説として

$$H^0(P):\ \mu_1 = \mu_2 = \mu_5 = \mu_7$$

のタイプだけを考えている。たとえば処理平均が複数のグループに分かれているような部分帰無仮説

$$H^0(P_1, P_2):\ \mu_1 = \mu_2 = \mu_5 = \mu_7, \quad \mu_3 = \mu_4 = \mu_6$$

は考えていない。したがって，このようなタイプの帰無仮説のもとでは，ファミリー単位過誤率 $FWER$ は α 以下に保障されない。

Duncan 法

Duncan 法（ダンカン法）も計算手順は REGWQ 法と同様である。ただし $a-1$ 個の判定基準値として

$$R_p^D = \frac{\hat{\sigma}}{\sqrt{n}} \cdot q(p, \nu_e; \alpha_p^D), \quad p = 2, \ldots, a \tag{4.16}$$
$$\alpha_p^D = 1 - (1-\alpha)^{p-1}$$

を用いる。

大麦品種比較実験（表 1.1）のデータについての判定結果は最小有意差法と同じである。

Duncan 法の特徴と考え方

Duncan 法における確率 α_p^D の値は，$p = 2$ 以外では $\alpha_p^D > \alpha$ $(p > 2)$ である。たとえば $p = 7, \alpha = 0.05$ とすると，$\alpha_7^D = 1 - 0.95^{7-1} = 0.265$ となる。したがって Duncan 法はファミリー単位過誤率 $FWER$ を α 以下に保障することはない。実際，大麦品種比較試験の例 $(a = 7)$ で完全帰無仮説が成り立っている場合（全ての処理の母平均が等しい場合），5% の名目で Duncan 法を実施すると，どれかの比較で間違って有意差有りと判定する確率は 27% 近くになる。

Duncan 法では，処理の数 a が増えたときに，ファミリー単位過誤率は大きくなってもよいと考えられている。その対価として，検出力を高くしようというものである。最小有意差法と同じくらい差が有意であると判定される場合が多い。Duncan 法についての議論は三輪 (1997) を参照されたい。

注　多重範囲検定

REGWQ 法・SNK 法・Duncan 法では，判定基準値 R_p $(p = 2, \ldots, a)$ が変化するので，多重範囲検定 (multiple range tests) とよばれることもある。たとえば Duncan 法は DMRT (Duncan's multiple range tests) などとよばれる。これに対し，LSD 法・Tukey 法では一定の判定基準値を使用する。

注　各手法の判定基準値

処理平均値を昇順に並べて判定する場合の判定基準値:

p	2	3	4	5	6	7
LSD 法	10.53	10.53	10.53	10.53	10.53	10.53
Duncan 法	10.53	11.06	11.41	11.66	11.84	11.99
SNK 法	10.53	12.71	14.01	14.95	15.68	16.27
REGWQ 法	13.37	14.59	15.25	15.69	15.69	16.27
Tukey 法	16.27	16.27	16.27	16.27	16.27	16.27

4.5　アンバランストモデルでの対比較

一元配置完全無作為化法実験では，処理の繰返し数 n_i がそろっていない場合がある。そのときは，処理 A_i, A_j の繰返し数を n_i, n_j とすると，繰返し数が揃っている場合の判定基準値計算の $\hat{\sigma}/\sqrt{n}$ の部分を全て

$$\hat{\sigma}\sqrt{\frac{1}{2}\left(\frac{1}{n_i}+\frac{1}{n_j}\right)}$$

に置き換えて判定基準値を計算すればよい。たとえば，最小有意差法・Tukey 法・REGWQ 法では

LSD 法: $$LSD(\alpha)(i,j)=\hat{\sigma}\sqrt{\frac{1}{n_i}+\frac{1}{n_j}}\cdot t(\nu_e;\alpha/2)$$

Tukey 法: $$HSD_a(\alpha)(i,j)=\hat{\sigma}\sqrt{\frac{1}{2}\left(\frac{1}{n_i}+\frac{1}{n_j}\right)}\cdot q(k,\nu_e;\alpha)$$

REGWQ 法: $$R_p(i,j)=\hat{\sigma}\sqrt{\frac{1}{2}\left(\frac{1}{n_i}+\frac{1}{n_j}\right)}\cdot q(p,\nu_e;\alpha_p)$$

となる。判定基準値は n_i と n_j に依存するので，比較するペアごとに計算しなければならない。

Tukey-Kramer 法

とくに Tukey 法で，アンバランストな場合に，

$$|\bar{y}_{i\cdot}-\bar{y}_{j\cdot}|>HSD_a(\alpha)(i,j)\quad\Longrightarrow\quad \text{有意差有り}$$

によって検定する方式を Tukey-Kramer （クレイマー）法という。これは，アンバランストな場合でも同時信頼区間

$$[\hat{L}_{ij}(\mathbb{y}),\ \hat{U}_{ij}(\mathbb{y})]=\left[\bar{y}_{i\cdot}-\bar{y}_{j\cdot}-HSD_a(\alpha)(i,j),\ \bar{y}_{i\cdot}-\bar{y}_{j\cdot}+HSD_a(\alpha)(i,j)\right]$$

に対して，

$$\Pr\{\mu_i-\mu_j\in[\hat{L}_{ij}(\mathbb{y}),\ \hat{U}_{ij}(\mathbb{y})],\quad 1\le i,j\le a\}\ge 1-\alpha \tag{4.17}$$

が成り立つことに基づく。(4.17) 式が成り立つことは，多くの人によって予想されていた（Tukey-Kramer 予想とよばれる）。この予想は Hayter (1984) によって正しいことが証明された。

4.6　SAS による対比較の実行例

出力 4.1 に，SAS による一元配置乱塊法実験データ（大麦品種比較実験，表 1.1）の解析プログラムを示す。SAS で多重比較を実行するには，`GLM` プロシージャの `Means` ステートメントにおいて，`LSD`，`Tukey` などのオプションを指定する。なお出力 4.1 では例示のために，`LSD` から `Tukey` までを指定している。通常の解析では，実験の目的に応じてどれか 1 つの多重比較オプションを指定する。出力 4.2 に出力結果を示す。ここでは，SAS の前バージョンによるテキスト形式での出力を示している。

現在，SAS では，**OnDemand for Academics:**

`https://www.sas.com/ja_jp/software/on-demand-for-academics.html`

により無償で SAS にアクセスできる。

出力 4.1　SAS による一元配置乱塊法実験の解析プログラム

```
Title 'Pairwise Comparisons';
Data duncan;
  Do R = 1 to 6;
    Do A = 1 to 7;
      Input y @@;
      Output;
    End;
  End;
DataLines;
  39 63 66 47 80 55 47
  64 63 69 58 73 61 79
  57 82 59 62 79 63 56
  42 67 56 64 55 62 57
  49 65 85 75 73 56 58
  47 87 71 63 68 51 69
;
Run;

Proc GLM Data = duncan;
  Class R A;
  Model y = R A / SS3;
  Means A / LSD Duncan SNK REGWQ Tukey;
Run;
```

出力 **4.2　SAS** による対比較の出力結果

```
                          t Tests (LSD) for y

  NOTE: This test controls the Type I comparisonwise error rate, not the
                         experimentwise error rate.

                     Alpha                              0.05
                     Error Degrees of Freedom             30
                     Error Mean Square              79.69048
                     Critical Value of t             2.04227
                     Least Significant Difference     10.526

        Means with the same letter are not significantly different.

                t Grouping          Mean       N    A

                         A        71.333       6    5
                         A        71.167       6    2
                    B    A        67.667       6    3
                    B    A        61.500       6    4
                    B    A        60.000       6    7
                    B    C        58.000       6    6
                         C        49.667       6    1

                    Duncan's Multiple Range Test for y

  NOTE: This test controls the Type I comparisonwise error rate, not the
                         experimentwise error rate.

                       Alpha                          0.05
                       Error Degrees of Freedom         30
                       Error Mean Square          79.69048

Number of Means          2          3          4          5          6          7
Critical Range       10.53      11.06      11.41      11.66      11.84      11.99
```

Means with the same letter are not significantly different.

Duncan Grouping		Mean	N	A
	A	71.333	6	5
	A	71.167	6	2
B	A	67.667	6	3
B	A	61.500	6	4
B	A	60.000	6	7
B	C	58.000	6	6
	C	49.667	6	1

Student-Newman-Keuls Test for y

NOTE: This test controls the Type I experimentwise error rate under the complete null hypothesis but not under partial null hypotheses.

Alpha	0.05
Error Degrees of Freedom	30
Error Mean Square	79.69048

Number of Means	2	3	4	5	6	7
Critical Range	10.525728	12.705867	14.014232	14.949541	15.676311	16.269306

Means with the same letter are not significantly different.

SNK Grouping		Mean	N	A
	A	71.333	6	5
	A	71.167	6	2
	A	67.667	6	3
B	A	61.500	6	4
B	A	60.000	6	7
B	A	58.000	6	6
B		49.667	6	1

Ryan-Einot-Gabriel-Welsch Multiple Range Test for y

NOTE: This test controls the Type I experimentwise error rate.

Alpha	0.05
Error Degrees of Freedom	30
Error Mean Square	79.69048

Number of Means	2	3	4	5	6	7
Critical Range	13.365951	14.589316	15.248963	15.690638	15.690638	16.269306

Means with the same letter are not significantly different.

REGWQ Grouping		Mean	N	A
	A	71.333	6	5
	A	71.167	6	2
	A	67.667	6	3
B	A	61.500	6	4
B	A	60.000	6	7
B	A	58.000	6	6
B		49.667	6	1

```
                 Tukey's Studentized Range (HSD) Test for y

   NOTE: This test controls the Type I experimentwise error rate, but it
           generally has a higher Type II error rate than REGWQ.

                 Alpha                                       0.05
                 Error Degrees of Freedom                      30
                 Error Mean Square                       79.69048
                 Critical Value of Studentized Range      4.46418
                 Minimum Significant Difference            16.269

        Means with the same letter are not significantly different.

             Tukey Grouping          Mean       N    A

                          A        71.333       6    5
                          A        71.167       6    2
                          A        67.667       6    3
                     B    A        61.500       6    4
                     B    A        60.000       6    7
                     B    A        58.000       6    6
                     B             49.667       6    1
```

4.7 R による対比較の計算

1) TukeyHSD() 関数

- **Tukey** 法の実行。

2) multcomp パッケージ

- **Tukey** 法，**Dunnett** 法，**Scheffé** 法。
- 一般化線形モデルのパラメータに対する多重比較も実行できる。

3) agricolae パッケージ

- **LSD** 法，**Duncan** 法，**SNK** 法，**REGW** 法，**HSD** 法（**Tukey** 法）。

出力 **4.3** **R** による対比較の実行

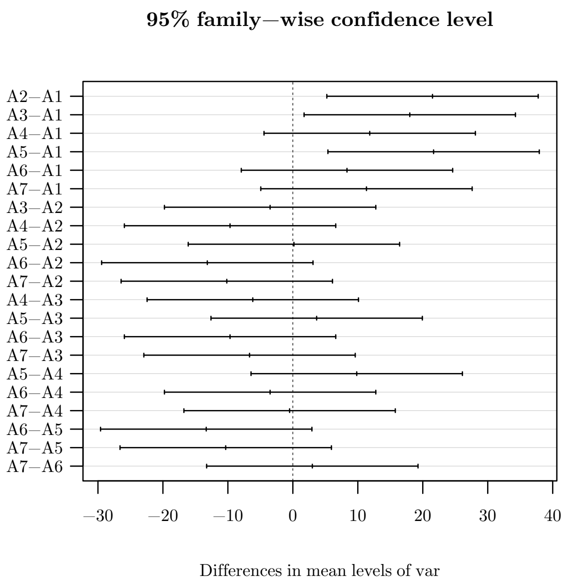
```
>### Pairwise comparisons
> ### Barley experiment in randomised blocks from Duncan (1955)
> ## data frame
> y <- c(39, 63, 66, 47, 80, 55, 47,
+        64, 63, 69, 58, 73, 61, 79,
+        57, 82, 59, 62, 79, 63, 56,
+        42, 67, 56, 64, 55, 62, 57,
+        49, 65, 85, 75, 73, 56, 58,
+        47, 87, 71, 63, 68, 51, 69)
> ## generate levels
> block <- gl(n=6, k=7, labels=paste("R", 1:6, sep=""))
> var <- gl(n=7, k=1, length=42, labels=paste("A", 1:7, sep=""))
> barley.dat <- data.frame(y, block=as.factor(block), var=as.factor(var))
>
> ## ANOVA
> barley.aov <- aov(y ~ block+var, data=barley.dat)
> anova(barley.aov)
Analysis of Variance Table

Response: y
          Df  Sum Sq Mean Sq F value   Pr(>F)
block      5  709.62  141.92  1.7809 0.147021
```

```
var        6 2201.29  366.88  4.6038 0.002002 **
Residuals 30 2390.71   79.69
---
Signif. codes:  0 ’***’ 0.001 ’**’ 0.01 ’*’ 0.05 ’.’ 0.1 ’ ’ 1
>
> ### pairsiwe comparisons
> ## Tukey procedure by TukeyHSD() function
> alpha <- 0.05
> barley.hsd <- TukeyHSD(barley.aov, "var", conf.level=1.0-alpha)
> barley.hsd
  Tukey multiple comparisons of means
    95% family-wise confidence level

Fit: aov(formula = y ~ block + var, data = barley.dat)

$var
             diff        lwr       upr     p adj
A2-A1  21.5000000   5.230695 37.769305 0.0039984
A3-A1  18.0000000   1.730695 34.269305 0.0226197
A4-A1  11.8333333  -4.435972 28.102638 0.2787277
A5-A1  21.6666667   5.397362 37.935972 0.0036702
A6-A1   8.3333333  -7.935972 24.602638 0.6729635
A7-A1  11.3333333  -4.935972 27.602638 0.3257583
A3-A2  -3.5000000 -19.769305 12.769305 0.9928393
A4-A2  -9.6666667 -25.935972  6.602638 0.5108892
A5-A2   0.1666667 -16.102638 16.435972 1.0000000
A6-A2 -13.1666667 -29.435972  3.102638 0.1763459
A7-A2 -10.1666667 -26.435972  6.102638 0.4516801
A4-A3  -6.1666667 -22.435972 10.102638 0.8900025
A5-A3   3.6666667 -12.602638 19.935972 0.9908378
A6-A3  -9.6666667 -25.935972  6.602638 0.5108892
A7-A3  -6.6666667 -22.935972  9.602638 0.8495169
A5-A4   9.8333333  -6.435972 26.102638 0.4909109
A6-A4  -3.5000000 -19.769305 12.769305 0.9928393
A7-A4  -0.5000000 -16.769305 15.769305 0.9999999
A6-A5 -13.3333333 -29.602638  2.935972 0.1658760
A7-A5 -10.3333333 -26.602638  5.935972 0.4325091
A7-A6   3.0000000 -13.269305 19.269305 0.9968936
>
> plot(barley.hsd, las=1)
```



```
> ## Tukey procedure by multcomp package
> library(multcomp)
> barley.tukey <- glht(barley.aov, linfct=mcp(var="Tukey"))
> summary(barley.tukey)

     Simultaneous Tests for General Linear Hypotheses

Multiple Comparisons of Means: Tukey Contrasts


Fit: aov(formula = y ~ block + var, data = barley.dat)

Linear Hypotheses:
             Estimate Std. Error t value Pr(>|t|)
A2 - A1 == 0  21.5000     5.1540   4.172  0.00393 **
A3 - A1 == 0  18.0000     5.1540   3.492  0.02297 *
A4 - A1 == 0  11.8333     5.1540   2.296  0.27874
A5 - A1 == 0  21.6667     5.1540   4.204  0.00368 **
A6 - A1 == 0   8.3333     5.1540   1.617  0.67305
A7 - A1 == 0  11.3333     5.1540   2.199  0.32582
A3 - A2 == 0  -3.5000     5.1540  -0.679  0.99284
A4 - A2 == 0  -9.6667     5.1540  -1.876  0.51093
A5 - A2 == 0   0.1667     5.1540   0.032  1.00000
A6 - A2 == 0 -13.1667     5.1540  -2.555  0.17624
A7 - A2 == 0 -10.1667     5.1540  -1.973  0.45149
A4 - A3 == 0  -6.1667     5.1540  -1.196  0.88999
A5 - A3 == 0   3.6667     5.1540   0.711  0.99083
A6 - A3 == 0  -9.6667     5.1540  -1.876  0.51087
A7 - A3 == 0  -6.6667     5.1540  -1.293  0.84950
A5 - A4 == 0   9.8333     5.1540   1.908  0.49075
A6 - A4 == 0  -3.5000     5.1540  -0.679  0.99283
A7 - A4 == 0  -0.5000     5.1540  -0.097  1.00000
A6 - A5 == 0 -13.3333     5.1540  -2.587  0.16587
A7 - A5 == 0 -10.3333     5.1540  -2.005  0.43220
A7 - A6 == 0   3.0000     5.1540   0.582  0.99690
---
Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
(Adjusted p values reported -- single-step method)

Warning message:
In RET$pfunction("adjusted", ...) : Completion with error > abseps
>
> confint(barley.tukey)

     Simultaneous Confidence Intervals

Multiple Comparisons of Means: Tukey Contrasts

Fit: aov(formula = y ~ block + var, data = barley.dat)

Quantile = 3.1607
95% family-wise confidence level

Linear Hypotheses:
             Estimate lwr      upr
A2 - A1 == 0  21.5000   5.2096  37.7904
A3 - A1 == 0  18.0000   1.7096  34.2904
A4 - A1 == 0  11.8333  -4.4571  28.1237
A5 - A1 == 0  21.6667   5.3763  37.9571
A6 - A1 == 0   8.3333  -7.9571  24.6237
```

```
A7 - A1 == 0  11.3333  -4.9571  27.6237
A3 - A2 == 0  -3.5000 -19.7904  12.7904
A4 - A2 == 0  -9.6667 -25.9571   6.6237
A5 - A2 == 0   0.1667 -16.1237  16.4571
A6 - A2 == 0 -13.1667 -29.4571   3.1237
A7 - A2 == 0 -10.1667 -26.4571   6.1237
A4 - A3 == 0  -6.1667 -22.4571  10.1237
A5 - A3 == 0   3.6667 -12.6237  19.9571
A6 - A3 == 0  -9.6667 -25.9571   6.6237
A7 - A3 == 0  -6.6667 -22.9571   9.6237
A5 - A4 == 0   9.8333  -6.4571  26.1237
A6 - A4 == 0  -3.5000 -19.7904  12.7904
A7 - A4 == 0  -0.5000 -16.7904  15.7904
A6 - A5 == 0 -13.3333 -29.6237   2.9571
A7 - A5 == 0 -10.3333 -26.6237   5.9571
A7 - A6 == 0   3.0000 -13.2904  19.2904
>
> ## Tukey procedure by agricolae package
> library(agricolae)
> HSD.test(barley.aov,  "var", console=TRUE)

Study: barley.aov ~ "var"

HSD Test for y

Mean Square Error:  79.69048

var,  means

          y       std r       se Min Max   Q25  Q50   Q75
A1 49.66667  9.373722 6 3.644413  39  64 43.25 48.0 55.00
A2 71.16667 10.553041 6 3.644413  63  87 63.50 66.0 78.25
A3 67.66667 10.269697 6 3.644413  56  85 60.75 67.5 70.50
A4 61.50000  9.093954 6 3.644413  47  75 59.00 62.5 63.75
A5 71.33333  9.136009 6 3.644413  55  80 69.25 73.0 77.50
A6 58.00000  4.732864 6 3.644413  51  63 55.25 58.5 61.75
A7 61.00000 11.260551 6 3.644413  47  79 56.25 57.5 66.25

Alpha: 0.05 ; DF Error: 30
Critical Value of Studentized Range: 4.464177

Minimun Significant Difference: 16.26931

Treatments with the same letter are not significantly different.

          y groups
A5 71.33333      a
A2 71.16667      a
A3 67.66667      a
A4 61.50000     ab
A7 61.00000     ab
A6 58.00000     ab
A1 49.66667      b
>
> ## REGWQ procedure by agricolae package
> REGW.test(barley.aov, "var", console=TRUE)

Study: barley.aov ~ "var"

Ryan, Einot and Gabriel and Welsch multiple range test
```

```
for y

Mean Square Error:  79.69048

var,  means

          y       std r       se Min Max   Q25  Q50   Q75
A1 49.66667  9.373722 6 3.644413  39  64 43.25 48.0 55.00
A2 71.16667 10.553041 6 3.644413  63  87 63.50 66.0 78.25
A3 67.66667 10.269697 6 3.644413  56  85 60.75 67.5 70.50
A4 61.50000  9.093954 6 3.644413  47  75 59.00 62.5 63.75
A5 71.33333  9.136009 6 3.644413  55  80 69.25 73.0 77.50
A6 58.00000  4.732864 6 3.644413  51  63 55.25 58.5 61.75
A7 61.00000 11.260551 6 3.644413  47  79 56.25 57.5 66.25

Alpha: 0.05 ; DF Error: 30

Critical Range
       2        3        4        5        6        7
13.36699 14.58932 15.24967 15.69070 15.69070 16.26931

Means with the same letter are not significantly different.

          y groups
A5 71.33333      a
A2 71.16667      a
A3 67.66667      a
A4 61.50000     ab
A7 61.00000     ab
A6 58.00000     ab
A1 49.66667      b
>
> ## SNK procedure by agricolae package
> SNK.test(barley.aov, "var", console=TRUE)

Study: barley.aov ~ "var"

Student Newman Keuls Test
for y

Mean Square Error:  79.69048

var,  means

          y       std r       se Min Max   Q25  Q50   Q75
A1 49.66667  9.373722 6 3.644413  39  64 43.25 48.0 55.00
A2 71.16667 10.553041 6 3.644413  63  87 63.50 66.0 78.25
A3 67.66667 10.269697 6 3.644413  56  85 60.75 67.5 70.50
A4 61.50000  9.093954 6 3.644413  47  75 59.00 62.5 63.75
A5 71.33333  9.136009 6 3.644413  55  80 69.25 73.0 77.50
A6 58.00000  4.732864 6 3.644413  51  63 55.25 58.5 61.75
A7 61.00000 11.260551 6 3.644413  47  79 56.25 57.5 66.25

Alpha: 0.05 ; DF Error: 30

Critical Range
       2        3        4        5        6        7
10.52583 12.70595 14.01423 14.94967 15.67631 16.26931

Means with the same letter are not significantly different.
```

```
          y groups
A5 71.33333      a
A2 71.16667      a
A3 67.66667      a
A4 61.50000     ab
A7 61.00000     ab
A6 58.00000     ab
A1 49.66667      b
>
> ## Duncan procedure by agricolae package
> duncan.test(barley.aov, "var", console=TRUE)

Study: barley.aov ~ "var"

Duncan's new multiple range test
for y

Mean Square Error:  79.69048

var,  means

          y       std r       se Min Max   Q25  Q50   Q75
A1 49.66667  9.373722 6 3.644413  39  64 43.25 48.0 55.00
A2 71.16667 10.553041 6 3.644413  63  87 63.50 66.0 78.25
A3 67.66667 10.269697 6 3.644413  56  85 60.75 67.5 70.50
A4 61.50000  9.093954 6 3.644413  47  75 59.00 62.5 63.75
A5 71.33333  9.136009 6 3.644413  55  80 69.25 73.0 77.50
A6 58.00000  4.732864 6 3.644413  51  63 55.25 58.5 61.75
A7 61.00000 11.260551 6 3.644413  47  79 56.25 57.5 66.25

Alpha: 0.05 ; DF Error: 30

Critical Range
       2        3        4        5        6        7
10.52583 11.06157 11.40886 11.65674 11.84390 11.99047

Means with the same letter are not significantly different.

          y groups
A5 71.33333      a
A2 71.16667      a
A3 67.66667     ab
A4 61.50000     ab
A7 61.00000     ab
A6 58.00000     bc
A1 49.66667      c
>
> ## LSD procedure by agricolae package
> LSD.test(barley.aov, "var", console=TRUE)

Study: barley.aov ~ "var"

LSD t Test for y

Mean Square Error:  79.69048

var,  means and individual ( 95 %) CI

          y       std r       se      LCL      UCL Min Max   Q25  Q50   Q75
```

```
A1 49.66667  9.373722 6 3.644413 42.22378 57.10955  39  64 43.25 48.0 55.00
A2 71.16667 10.553041 6 3.644413 63.72378 78.60955  63  87 63.50 66.0 78.25
A3 67.66667 10.269697 6 3.644413 60.22378 75.10955  56  85 60.75 67.5 70.50
A4 61.50000  9.093954 6 3.644413 54.05712 68.94288  47  75 59.00 62.5 63.75
A5 71.33333  9.136009 6 3.644413 63.89045 78.77622  55  80 69.25 73.0 77.50
A6 58.00000  4.732864 6 3.644413 50.55712 65.44288  51  63 55.25 58.5 61.75
A7 61.00000 11.260551 6 3.644413 53.55712 68.44288  47  79 56.25 57.5 66.25

Alpha: 0.05 ; DF Error: 30
Critical Value of t: 2.042272

least Significant Difference: 10.52583

Treatments with the same letter are not significantly different.

          y groups
A5 71.33333      a
A2 71.16667      a
A3 67.66667     ab
A4 61.50000     ab
A7 61.00000     ab
A6 58.00000     bc
A1 49.66667      c
>
```

5　対照処理との比較

処理のうちの **1** つが**対照**（標準，**control**）処理であり，この対照処理と他の試験処理との比較のみに興味がある場合を考える。

例 4　子豚の体重増への飼料の影響（対照処理との比較）

例として，タンパク質を強化した飼料が現行飼料に対して子豚の体重増に効果があるかどうかを調べるために行なった実験を考える。取り上げた処理は次の **6** つである。

A_1:	現行飼料
A_2:	植物性タンパク質を標準量添加したもの
A_3:	植物性タンパク質を標準量の **2** 倍添加したもの
A_4:	植物性タンパク質を標準量の **3** 倍添加したもの
A_5:	動物性タンパク質を標準量添加したもの
A_6:	動物性タンパク質を標準量の **2** 倍添加したもの

完全無作為化法 (**completely randomised design**) により各処理をランダムに **3** 頭の子豚に与えた。一定期間（**6** 週間）後の体重増のデータを表 **5.1** に，その分散分析表を表 **5.2** に示す。分散分析の結果は 1% 水準で有意である ($p = 0.0059$)。

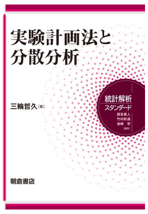
この実験の目的は，現行飼料（対照処理）A_1 に対して効果のある飼料（試験処理）を選ぶことであり，試験処理 $A_2, \ldots, A_5$ のあいだの差には，とくに興味はない。

表 **5.1**　子豚の体重増 (**kg**)

飼料	A_1	A_2	A_3	A_4	A_5	A_6
	27	24	25	31	33	36
	25	23	26	29	30	31
	23	27	31	34	27	34
平均	25.0	24.7	27.3	31.3	30.0	33.7

表 **5.2**　子豚の体重増の分散分析表（一元配置完全無作為化法）

変動因	自由度	平方和	平均平方	F 比	p-値
飼料 A	5	195.33	39.07	5.81**	0.0059
誤差 E	12	80.67	6.72		
全体 T	17	276.00			



【参考】
三輪 哲久 (著) 『実験計画法と分散分析』
朝倉書店 (2015/9/25), 216 ページ

5.1　帰無仮説のファミリーと対立仮説

処理のうちの 1 つが対照（control）処理であり，この対照処理と他の試験処理との比較のみに興味がある。ここでは A_1 を対照処理とする。帰無仮説のファミリーは

$$\mathcal{H} = \{H_{i1}^{0}:\ \mu_i = \mu_1 \quad (\mu_i - \mu_1 = 0), \quad i = 2, \ldots, a\}$$

であり，興味の対象となる比較の数は $m = a - 1$ となる。

対立仮説

対照処理との比較の問題においては，「他の処理が対照処理と異なるかどうか」を検討したい場合と，「他の処理は対照処理よりも値が大きいか（あるいは小さいか)」を検討したい場合がある。この検出したい事柄を対立仮説として

1)　$H_{i1}^{A\pm} : \mu_i \neq \mu_1$　（両側対立仮説）
2)　$H_{i1}^{A+} : \mu_i > \mu_1$　（上片側対立仮説）
3)　$H_{i1}^{A-} : \mu_i < \mu_1$　（下片側対立仮説）

のように表わす。1) に対応する検定を両側検定，2) と 3) に対応する検定を片側検定という。

注．両側検定を行なうか片側検定を行なうか（さらに上側か下側か）は，データを見た後で決めてはいけない。たとえば，対照処理よりも大きな値が観測されたので上片側対立仮説を設定するという方法では，ファミリー単位過誤率は保障されない（その理由を考えよ)。どの対立仮説を設定するかは研究（実験）の目的に依存し，実験計画の段階で決めておくべきである。

例 4 では，実験の目的が現行飼料（対照処理）A_1 に対して効果のある飼料を選ぶことなので，上片側対立仮説

$$H_{i1}^{A+} : \mu_i > \mu_1 \quad (i = 2, \ldots, 5)$$

を考える。また，対照処理以外の処理 $A_2, \ldots, A_5$ のあいだの差の比較には興味はない。

5.2　Dunnett 法

対照との比較において，ファミリー単位過誤率を制御するための代表的な方法が **Dunnett 法**（ダネット法）である。

- 両側 Dunnett 法の計算手順
  両側対立仮説 $H_{i1}^{A\pm}$： $\mu_i \neq \mu_1$ に対しては，

$$|\bar{y}_{i\cdot} - \bar{y}_{1\cdot}| > D''(\alpha) = \hat{\sigma}\sqrt{2/n} \cdot d''(a-1, \nu_e; \alpha) \tag{5.1}$$

  のときに，μ_i は μ_1 とは有意に異なると判定する。
- 片側 Dunnett 法の計算手順
  上片側対立仮説 H_{i1}^{A+}： $\mu_i > \mu_1$ に対しては，

$$\bar{y}_{i\cdot} - \bar{y}_{1\cdot} > D'(\alpha) = \hat{\sigma}\sqrt{2/n} \cdot d'(a-1, \nu_e; \alpha) \tag{5.2}$$

  のときに，μ_i は μ_1 より有意に大きいと判定する。一方，下片側対立仮説 H_{i1}^{A-}： $\mu_i < \mu_1$ に対しては，

$$\bar{y}_{i\cdot} - \bar{y}_{1\cdot} < -D'(\alpha) = -\hat{\sigma}\sqrt{2/n} \cdot d'(a-1, \nu_e; \alpha) \tag{5.3}$$

  のときに，μ_i は μ_1 より有意に小さいと判定する。

判定基準値計算のための $d''(a-1, \nu_e; \alpha)$ と $d'(a-1, \nu_e; \alpha)$ については，Hochberg & Tamhane (1987)，Hsu (1996) などの多重比較法の標準的なテキストに与えられている。

表 5.1 の飼料比較実験の例では，

$$a = 6,\ n = 3,$$
$$\bar{y}_{1\cdot} = 25.0,\ \bar{y}_{2\cdot} = 24.7,\ \bar{y}_{3\cdot} = 27.3,\ \bar{y}_{4\cdot} = 31.3,\ \bar{y}_{5\cdot} = 30.0,\ \bar{y}_{6\cdot} = 33.7$$
$$\hat{\sigma} = \sqrt{V_e} = \sqrt{6.72} = 2.59,\ \nu_e = 12$$
$$d'(5, 12; 0.05) = 2.502$$

を使って，判定基準値は

$$D'(0.05) = 2.59 \times \sqrt{2/3} \times 2.502 = 5.30$$

となる。各処理と対照処理との差は

$$\bar{y}_{2\cdot} - \bar{y}_{1\cdot} = -0.3,\ \bar{y}_{3\cdot} - \bar{y}_{1\cdot} = 2.3,\ \bar{y}_{4\cdot} - \bar{y}_{1\cdot} = 6.3,$$
$$\bar{y}_{5\cdot} - \bar{y}_{1\cdot} = 5.0,\ \bar{y}_{6\cdot} - \bar{y}_{1\cdot} = 8.7$$

であるから，A_1 に比べて有意に効果のあるのは A_4 と A_6 である。

Dunnett 法の特徴と考え方

Dunnett 法は Tukey 法と同様に，同時信頼区間を構成することによってファミリー単位過誤率 $FWER$ を制御する方法である。

両側対立仮説 $H_{i1}^{A\pm}$： $\mu_i \neq \mu_1$ については，両側信頼区間

$$\begin{aligned}\mu_i - \mu_1 &\in [\hat{L}_{i1}(\mathbb{y}),\ \hat{U}_{i1}(\mathbb{y})] \\ &= \left[\bar{y}_{i\cdot} - \bar{y}_{1\cdot} - D''(\alpha),\ \bar{y}_{i\cdot} - \bar{y}_{1\cdot} + D''(\alpha)\right] \quad (i = 2, \ldots, a)\end{aligned} \tag{5.4}$$

を考える。このとき

$$\Pr\{\mu_i - \mu_1 \in [\hat{L}_{i1}(\mathbb{y}),\ \hat{U}_{i1}(\mathbb{y})],\quad i = 2, \ldots, a\} = 1 - \alpha$$

が成り立つ（この式が成り立つように $d''(a-1,\nu_e;\alpha)$ の値が定められている)。(5.1) 式の判定方式は，この信頼区間がゼロを含まないことと同値であり，

$$|\bar{y}_{i\cdot}-\bar{y}_{1\cdot}|>D''(\alpha) \iff 0\notin[\hat{L}_{i1}(\mathrm{y}),\ \hat{U}_{i1}(\mathrm{y})]$$

が成り立つ。したがって，同時信頼区間に関する議論（第 2.4，3.2 節）よりファミリー単位過誤率 $FWER$ が保障される。

上片側対立仮説 H_{i1}^{A+}： $\mu_i>\mu_1$ に対しては，上片側信頼区間

$$\mu_i-\mu_1\in[\hat{L}_{i1}(\mathrm{y}),\ \infty)=\big[\bar{y}_{i\cdot}-\bar{y}_{1\cdot}-D'_a(\alpha),\ \infty\big)\quad(i=2,\dots,a) \tag{5.5}$$

を考える。$d'(a-1,\nu_e;\alpha)$ の値は

$$\Pr\{\mu_i-\mu_1\in[\hat{L}_{i1}(\mathrm{y}),\ \infty),\quad i=2,\dots,a\}=1-\alpha$$

が成り立つように決められている。両側検定の場合と同様に

$$\bar{y}_{i\cdot}-\bar{y}_{1\cdot}>D'_a(\alpha)\quad\iff\quad 0\notin[\hat{L}_{i1}(\mathrm{y}),\ \infty)$$

が成り立ち，(5.2) 式の判定方式は信頼区間がゼロを含まないことと同値である。下側同時信頼区間についても同様に構成することができる。

注. $d''(a-1,\nu_e;\alpha)$ と $d'(a-1,\nu_e;\alpha)$ についての数理的説明
$u_1,\dots,u_a,\chi^2$ が互いに独立に

$$u_i\sim N(0,1),\quad 1\le i\le a,$$
$$\chi^2\sim\chi^2(\nu_e)$$

に従うとき，Dunnett 法の基準値 $d'(a-1,\nu_e;\alpha)$, $d''(a-1,\nu_e;\alpha)$ は

$$\Pr\left\{\max_{2\le i\le a}\frac{u_i-u_1}{\sqrt{2\cdot\chi^2/\nu_e}}>d'(a-1,\nu_e;\alpha)\right\}=\alpha,$$
$$\Pr\left\{\max_{2\le i\le a}\frac{|u_i-u_1|}{\sqrt{2\cdot\chi^2/\nu_e}}>d''(a-1,\nu_e;\alpha)\right\}=\alpha$$

を満たすように決められている。

この定義から，$a-1$ 個の母平均の差 $\mu_i-\mu_1$ $(2\le i\le a)$ の同時信頼区間が (5.4) 式，(5.5) 式で与えられる。その証明は Tukey 法の場合と同様である。

演習問題 12　Tukey 法の場合（演習問題 10）を参考に，対照との比較における $\mu_i-\mu_1$ $(2\le i\le a)$ の同時信頼区間が (5.4) 式，(5.5) 式で与えられることを確認せよ。 □

アンバランストなモデルでの Dunnett 法

一元配置完全無作為化法で繰返し数 n_i が不揃いの場合（アンバランストなモデルの場合)，繰返し数が等しいことを仮定した $d''(a-1,\nu_e;\alpha)$，$d'(a-1,\nu_e;\alpha)$ を使うとファミリー単位過誤率 $FWER$ が保障されない場合がある。この点は対比較の Tukey-Kramer 法とは異なる。しかし，$d''(a-1,\nu_e;\alpha)$，$d'(a-1,\nu_e;\alpha)$ の計算は数値計算的に困難ではないので，統計パッケージを用いてアンバランストな場合の Dunnett 法を実行することができる。

5.3　t 検定

対照処理との比較において比較単位過誤率 $CWER$ を制御するには，各比較の過誤率が α 以下になるように個々の検定を設計する。したがって，各比較において通常の t 検定を実行すればよい。

1) 両側対立仮説 $H_{i1}^{A\pm}$：$\mu_i \neq \mu_1$

$$|\bar{y}_{i\cdot} - \bar{y}_{1\cdot}| > \hat{\sigma}\sqrt{2/n} \cdot t(\nu_e; \alpha/2) \implies \text{帰無仮説 } H_{i1}^{0} : \mu_i = \mu_1 \text{ を棄却}$$

2) 上片側対立仮説 H_{i1}^{A+}：$\mu_i > \mu_1$

$$\bar{y}_{i\cdot} - \bar{y}_{1\cdot} > \hat{\sigma}\sqrt{2/n} \cdot t(\nu_e; \alpha) \implies \text{帰無仮説 } H_{i1}^{0} : \mu_i = \mu_1 \text{ を棄却}$$

3) 下片側対立仮説 H_{i1}^{A-}：$\mu_i < \mu_1$

$$\bar{y}_{i\cdot} - \bar{y}_{1\cdot} < -\hat{\sigma}\sqrt{2/n} \cdot t(\nu_e; \alpha) \implies \text{帰無仮説 } H_{i1}^{0} : \mu_i = \mu_1 \text{ を棄却}$$

ここで，$t(\nu_e; \alpha)$ は自由度 ν_e の t 分布の片側 α 点である。実は **Dunnett** 法の (**5.1**) 式 – (**5.3**) 式は，t 検定における $t(\nu_e; \alpha/2)$ を $d''(a-1, \nu_e; \alpha)$ で置き換え，$t(\nu_e; \alpha)$ を $d'(a-1, \nu_e; \alpha)$ で置き換えたものである。

注　対照処理との比較では，例 4 のように，探索的に多数の処理を実験に取り入れて対照処理との比較を行なう場合が多い。このとき比較単位の過誤のみを考慮していると間違って多くの処理を有意差ありと判定してしまう可能性があるので注意が必要である。

比較単位過誤率の制御を考えるのは，各比較のそれぞれが個別の実験目的となっている場合である。

5.4　SAS による Dunnett 法の計算

SAS による **Dunnett** 法の解析プログラム（出力 **5.3**）と実行例（**5.4**）を示す。**Dunnett** 法の上片側検定を実行するには，`Means` ステートメントにおいて，`DunnettU` を指定する。

出力 **5.3**　**SAS** による **Dunnett** 法の解析プログラム

```
Title 'Comparisons with a Control';
Data swinedata;
  Do R = 1 to 3;
    Do A = 1 to 6;
      Input w @@;
      Output;
    End;
  End;
DataLines;
  27  24  25  31  33  36
  25  23  26  29  30  31
  23  27  31  34  27  34
  ;
Run;

Proc GLM Data=swinedata;
  Class R A;
  Model w = A / SS3;
  Means A / DunnettU;
Run;
```

出力 **5.4**　**SAS** による **Dunnett** 法の出力結果

```
                         GLM プロシジャ
                      分類変数の水準の情報
                分類            水準     値
                R                  3    1 2 3
                A                  6    1 2 3 4 5 6

               読み込んだオブザベーション数          18
               使用されたオブザベーション数          18

従属変数 : w
要因                     自由度          平方和         平均平方      F 値   Pr > F
Model                         5    195.3333333     39.0666667      5.81   0.0059
Error                        12     80.6666667      6.7222222
Corrected Total              17    234.5000000

               R2 乗         変動係数        Root MSE        w の平均
            0.707729       9.044389        2.592725        28.66667

                                         Type III
要因                     自由度          平方和         平均平方      F 値   Pr > F
A                             5    195.3333333     39.0666667      5.81   0.0059

                         GLM プロシジャ
                  w に対する Dunnett の片側 t 検定
NOTE: この検定は全処理群とコントロールの間の比較に対する第 1 種の過誤の確率を
                            制御します。

                     アルファ                        0.05
                     誤差の自由度                      12
                     誤差の平均平方            6.722222
                     Dunnett の t の棄却値    2.50225
                     最小な有意差                5.2971

        有意水準 0.05 で有意に差があることを *** で示しています。

                  A                    同時 95% 信頼限
                比較       平均の差           界

                6 - 1       8.667      3.370 Infinity   ***
                4 - 1       6.333      1.036 Infinity   ***
                5 - 1       5.000     -0.297 Infinity
                3 - 1       2.333     -2.964 Infinity
                2 - 1      -0.333     -5.630 Infinity
```

5.5 R による Dunnett 法の計算

出力 **5.5** **R** による対照との比較の計算

```
> ### Dunnett test for swine weight-gain data
> ## data frame
> y <- c(27, 24, 25, 31, 33, 36,
+        25, 23, 26, 29, 30, 31,
+        23, 27, 31, 34, 27, 34)
> ## Generate levels
> diet <- gl(n=6, k=1, length=length(y), labels=paste("A", 1:6, sep=""))
> swine.dat <- data.frame(y, diet)
>
> ## ANOVA
> swine.aov <- aov(y ~ diet, data=swine.dat)
```

```
> summary(swine.aov)
            Df Sum Sq Mean Sq F value  Pr(>F)
diet         5 195.33   39.07   5.812 0.00594 **
Residuals   12  80.67    6.72
---
Signif. codes:  0 ’***’ 0.001 ’**’ 0.01 ’*’ 0.05 ’.’ 0.1 ’ ’ 1
>
> ## Dunnett test
> library(multcomp)
> swine.dnt <- glht(swine.aov, linfct=mcp(diet="Dunnett"),
+                    alternative="greater")
> summary(swine.dnt)
     Simultaneous Tests for General Linear Hypotheses
Multiple Comparisons of Means: Dunnett Contrasts
Fit: aov(formula = y ~ diet, data = swine.dat)

Linear Hypotheses:
             Estimate Std. Error t value  Pr(>t)
A2 - A1 <= 0  -0.3333     2.1170  -0.157 0.87515
A3 - A1 <= 0   2.3333     2.1170   1.102 0.38229
A4 - A1 <= 0   6.3333     2.1170   2.992 0.02136 *
A5 - A1 <= 0   5.0000     2.1170   2.362 0.06328 .
A6 - A1 <= 0   8.6667     2.1170   4.094 0.00303 **
---
Signif. codes:  0 ’***’ 0.001 ’**’ 0.01 ’*’ 0.05 ’.’ 0.1 ’ ’ 1
(Adjusted p values reported -- single-step method)

> confint(swine.dnt)
     Simultaneous Confidence Intervals
Multiple Comparisons of Means: Dunnett Contrasts
Fit: aov(formula = y ~ diet, data = swine.dat)

Quantile = -2.5018
95% family-wise confidence level

Linear Hypotheses:
             Estimate lwr     upr
A2 - A1 <= 0 -0.3333  -5.6296     Inf
A3 - A1 <= 0  2.3333  -2.9630     Inf
A4 - A1 <= 0  6.3333   1.0370     Inf
A5 - A1 <= 0  5.0000  -0.2963     Inf
A6 - A1 <= 0  8.6667   3.3704     Inf
```

6 対比の検定

本節では，処理効果の線形結合に興味がある場合を考える。

例 5 水稲の葉いもち病斑面積率（アンバランストな一元配置完全無作為化法実験）

表 **6.1** は，水稲 **6** 品種について，温室内のポットを実験単位とし，葉いもち病斑面積率を観測したデータである。各水準の繰返し数 n_i が一定にならなかった（三輪『実験計画法と分散分析』**(2015)**）を参考)。

分散分析表を表 **6.2** に示す。F 比は $F_A = 13.10 > 4.25 = F(5, 18; 0.01)$ であり **1%**水準で有意である。すなわち葉いもち病の発生に関して品種間に効果の違いがある。

この例において **6** つの処理 $A_1, \ldots, A_6$ は対等な立場ではなく，品種群 $\{1, 2, 3\}$ と $\{4, 5, 6\}$ とは，異なる母本（品種の親）からの品種であり，葉いもち病に対する抵抗性が異なっている可能

性がある。そこで，帰無仮説

$$H_{\mathbb{C}}^0\colon\ \frac{\mu_1+\mu_2+\mu_3}{3}-\frac{\mu_4+\mu_5+\mu_6}{3}=0$$

を考える。この線形結合が以下に定義する対比であることは容易に確かめられる。

表 6.1　葉いもち病斑面積率 (%)

（アンバランストな一元配置完全無作為化法）

品種	n_i	繰り返し					平均
A_1	5	36	32	32	34	29	32.6
A_2	3	29	27	25			27.0
A_3	4	29	32	31	34		31.5
A_4	3	28	24	30			27.3
A_5	4	28	26	25	23		25.5
A_6	5	21	23	20	18	25	21.4
計	24					総平均	27.5

表 6.2　水稲 6 品種葉いもち病斑面積率データの分散分析表

変動因	自由度	平方和	平均平方	F 比	p-値
品種 A	5	396.89	79.38	13.10**	1.79e-05
誤差 E	18	109.07	6.06		
全体 T	23	505.96			

$\hat{\sigma}^2 = V_e = 6.06\ (\nu_e = 18)$

6.1　対比のファミリー

処理効果の母平均 $\mu_i\ (i=1,\ldots,a)$ の線形結合

$$\sum_{i=1}^{a} c_i\,\mu_i$$

で，係数 $c_i\ (i=1,\ldots,a)$ の和が 0 になるもの，すなわち，

$$\sum_{i=1}^{a} c_i = 0 \tag{6.1}$$

を満たすものを**対比**（たいひ，contrast）という。この対比が 0 かどうかを実験データから検定するために，帰無仮説

$$H_{\mathbb{C}}^0\colon\ \sum_{i=1}^{a} c_i\,\mu_i = 0 \tag{6.2}$$

を考える。

帰無仮説のファミリー

後述の Scheffé 法では，無限に存在する対比についての帰無仮説全体

$$\mathcal{H}=\left\{H_{\mathbb{C}}^0\colon\ \sum_{i=1}^{a} c_i\,\mu_i=0\quad\Bigl(\sum_{i=1}^{a} c_i=0\Bigr)\right\} \tag{6.3}$$

をファミリーとして考える。

6.2 対比の t 検定

計算手順

対比に対する t **検定**では，

$$\left|\sum_{i=1}^{a} c_i\,\bar{y}_{i\cdot}\right| > \hat{\sigma}\sqrt{\textstyle\sum_{i=1}^{a}(c_i^2/n_i)}\cdot t(\nu_e;\alpha/2) \tag{6.4}$$

のときに，帰無仮説 $H_{\mathbb{C}}^0$： $\sum_{i=1}^{a} c_i\,\mu_i = 0$ を棄却する。ここで，$t(\nu_e;\alpha/2)$ は自由度 ν_e の t 分布の片側 $\alpha/2$ 点（両側 α 点）である。

表 **6.1** の水稲品種比較実験の例では，データから計算される対比の推定値は

$$\sum_{i=1}^{a} c_i\,\bar{y}_{i\cdot} = \frac{\bar{y}_{1\cdot}+\bar{y}_{2\cdot}+\bar{y}_{3\cdot}}{3} - \frac{\bar{y}_{4\cdot}+\bar{y}_{4\cdot}+\bar{y}_{6\cdot}}{3} = 5.6$$

となる。一方，(6.4) 式右辺の判定基準値は，1% 水準の $t(18;0.01/2) = 2.878$ を使うと

$$\sum_{i=1}^{a}\frac{c_i^2}{n_i} = \frac{1}{9}\Big(\frac{1}{5}+\frac{1}{3}+\frac{1}{4}+\frac{1}{3}+\frac{1}{4}+\frac{1}{5}\Big) = 0.174$$

$$\hat{\sigma}\sqrt{\textstyle\sum_{i=1}^{a}(c_i^2/n_i)}\cdot t(\nu_e;\alpha/2) = \sqrt{6.06\times 0.174}\times 2.878 = 2.96$$

である。したがって帰無仮説

$$H_{\mathbb{C}}^0：\ \frac{\mu_1+\mu_2+\mu_3}{3} - \frac{\mu_4+\mu_5+\mu_6}{3} = 0$$

は，1% 水準で棄却される。

t 検定の特徴と考え方

演習問題 13　(6.4) 式の判定方式は，帰無仮説

$$H_{\mathbb{C}}^0：\ \sum_{i=1}^{a} c_i\,\mu_i = 0$$

に対する t 検定であることを説明せよ。 □

したがって，対比に対する t 検定は比較単位過誤率 $CWER$ を α 以下に制御する。しかし複数の対比を考えて，それぞれに t 検定を行なえば，ファミリー単位過誤率 $FWER$ は α よりも大きくなる。ただ 1 つの対比を検定することが実験の目的である場合には t 検定を行なうことができる。

6.3 Scheffé 法

Scheffé 法（シェフェ法）は，全ての対比を考えたときの帰無仮説のファミリー (6.3) 式に対して，ファミリー単位過誤率 $FWER$ を α 以下に保障する方法である。

計算手順

判定基準値

$$S_{\mathbb{C}}(\alpha) = \hat{\sigma}\sqrt{(a-1)\cdot\textstyle\sum_{i=1}^{a}(c_i^2/n_i)\cdot F(a-1,\nu_e;\alpha)} \tag{6.5}$$

を計算し，

$$\left|\sum_{i=1}^{a} c_i\,\bar{y}_{i\cdot}\right| > S_{\mathbb{C}}(\alpha) \tag{6.6}$$

のときに帰無仮説 $H_{\mathbb{c}}^0$： $\sum_{i=1}^{a} c_i\,\mu_i = 0$ を棄却する。$F(a-1,\nu_e;\alpha)$ は自由度 $(a-1,\nu_e)$ の F 分布の上側 α 点である。

表 6.1 の例では，F 分布の 1% 点 $F(5,18;0.01)=4.248$ を用いると

$$\left|\sum_{i=1}^{a} c_i\,\bar{y}_{i\cdot}\right| = \left|\frac{\bar{y}_{1\cdot}+\bar{y}_{2\cdot}+\bar{y}_{3\cdot}}{3} - \frac{\bar{y}_{4\cdot}+\bar{y}_{4\cdot}+\bar{y}_{6\cdot}}{3}\right| = 5.6$$

$$\sum_{i=1}^{a}\frac{c_i^2}{n_i} = \frac{1}{9}\Big(\frac{1}{5}+\frac{1}{3}+\frac{1}{4}+\frac{1}{3}+\frac{1}{4}+\frac{1}{5}\Big) = 0.174$$

$$S_{\mathbb{c}}(0.01) = 2.46\times\sqrt{5\times 0.174\times 4.248} = 4.73$$

であるから 1% 水準で有意である。

Scheffé 法の特徴と考え方

Scheffé 法は，Tukey 法や Dunnett 法と同様に同時信頼区間に基づく方法である。対比についての同時信頼区間

$$\begin{aligned}\sum_{i=1}^{a} c_i\,\mu_i &\in [\hat{L}_{\mathbb{c}}(\mathbb{y}),\ \hat{U}_{\mathbb{c}}(\mathbb{y})]\\ &= \left[\sum_{i=1}^{a} c_i\,\bar{y}_{i\cdot} - S_{\mathbb{c}}(\alpha),\ \sum_{i=1}^{a} c_i\,\bar{y}_{i\cdot} + S_{\mathbb{c}}(\alpha)\right]\end{aligned} \tag{6.7}$$

を考える。このとき

$$\Pr\Big\{\text{全ての対比について}\quad \sum_{i=1}^{a} c_i\,\mu_i \in [\hat{L}_{\mathbb{c}}(\mathbb{y}),\ \hat{U}_{\mathbb{c}}(\mathbb{y})]\Big\} = 1-\alpha$$

が成り立つ。(6.6) 式の判定方式は，(6.7) 式の信頼区間がゼロを含まないこと，すなわち

$$0 \notin \left[\sum_{i=1}^{a} c_i\,\bar{y}_{i\cdot} - S_{\mathbb{c}}(\alpha),\ \sum_{i=1}^{a} c_i\,\bar{y}_{i\cdot} + S_{\mathbb{c}}(\alpha)\right]$$

と同値である。したがって，同時信頼区間に関する議論 (3.2 節) よりファミリー単位過誤率 $FWER$ が保障される。

演習問題 14　全ての対比に対する同時信頼区間が，(6.7) 式で与えられることを証明せよ。 □

ヒント　Cauchy-Schwarz の不等式を使って，任意の $\mathbb{c}=(c_1,\ldots,c_a)^T$ に対し

$$\frac{1}{(a-1)\,\hat{\sigma}^2}\frac{\left|\sum c_i\cdot(\bar{y}_{i\cdot}-\mu_i)\right|^2}{\sum c_i^2/n_i} \le \frac{\sum n_i\cdot\{(\bar{y}_{i\cdot}-\mu_i)-(\bar{y}_{\cdot\cdot}-\mu)\}^2}{(a-1)\,\hat{\sigma}^2}$$

が成り立つことを示し，右辺が自由度 $(a-1,\nu_e)$ の F 分布に従うことを利用すればよい。

Scheffé 法は分散分析の F 検定と密接な関係がある。演習問題 14 の結果を使えば，

$$\frac{1}{(a-1)\,\hat{\sigma}^2}\max_{\mathbb{c}}\frac{\left|\sum c_i\cdot\bar{y}_{i\cdot}\right|^2}{\sum c_i^2/n_i} = \frac{\sum n_i\,(\bar{y}_{i\cdot}-\bar{y}_{\cdot\cdot})^2}{(a-1)\,\hat{\sigma}^2} = F_A$$

は分散分析の F 統計量となる。したがって，次の 2 つの事項

- 分散分析の F 検定で有意となる（すなわち，$F_A > F(a-1,\nu_e;\alpha)$ となる)。
- Scheffé 法で有意となる対比 $\sum_{i=1}^{a} c_i\,\mu_i$ が少なくとも 1 つ存在する。

は同値である。分散分析の F 検定で有意とならないときは，Scheffé 法で有意となる対比は存在しない。

Tukey 法，Dunnett 法との比較

第 4 節の対比較 $H_{ij}^0: \mu_i - \mu_j = 0$，および第 5 節の対照との比較 $H_{i1}^0: \mu_i - \mu_1 = 0$ は，いずれも対比の一種である。したがって，これらの比較にも原理的には Scheffé 法を用いることが可能である。しかし Scheffé 法は無限個の対比に対してファミリー単位過誤率を制御しているので対比較や対照との比較においては検出力が低くなる（第 II 種の過誤率が高くなる）。たとえば第 4 節で検討した大麦品種比較実験の対比較の問題に対して Scheffé 法を用いると，(6.6) 式の判定基準値は

$$a = 7,\ n = 6,\ \hat{\sigma} = 8.93,\ \nu_e = 30,\ F(6, 30; 0.05) = 2.421$$

$$\hat{\sigma}\sqrt{(a-1)\cdot\sum_{i=1}^{a}(c_i^2/n_i)\cdot F(a-1,\nu_e;\alpha)}$$

$$= 8.93 \times \sqrt{6 \times (2/6) \times 2.421} = 19.65$$

である。Tukey 法の判定基準 $HSD_7(0.05) = 16.27$ よりもさらに厳しい判定基準値となる。したがって，対比較や対照との比較のように行なうべき比較が実験の目的から決まっているときには Scheffé 法を使用すべきではない。

6.4　R による対比の検定

出力 6.3　R による対比の検定

```
> ### Tests of contrasts ###
> ## data frame
> y <- c(36, 32, 32, 34, 29,
+        29, 27, 25,
+        29, 32, 31, 34,
+        28, 24, 30,
+        28, 26, 25, 23,
+        21, 23, 20, 18, 25)
> var <- c("A1", "A1", "A1", "A1", "A1",
+          "A2", "A2", "A2",
+          "A3", "A3", "A3", "A3",
+          "A4", "A4", "A4",
+          "A5", "A5", "A5", "A5",
+          "A6", "A6", "A6", "A6", "A6")
> rice.dat <- data.frame(y, var=as.factor(var))
>
> ## ANOVA
> rice.aov <- aov(y ~ var, data=rice.dat)
> anova(rice.aov)
Analysis of Variance Table

Response: y
          Df Sum Sq Mean Sq F value    Pr(>F)
var        5 396.89  79.378    13.1 1.794e-05 ***
Residuals 18 109.07   6.059
---
Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
>
> ## Contrast tests
> library(multcomp)
>
> ## t-test
> rice.glht <- glht(rice.aov,
+                   linfct=mcp(var=c(1/3, 1/3, 1/3, -1/3, -1/3, -1/3)))
> summary(rice.glht)
     Simultaneous Tests for General Linear Hypotheses
Multiple Comparisons of Means: User-defined Contrasts
```

```
Fit: aov(formula = y ~ var, data = rice.dat)

Linear Hypotheses:
       Estimate Std. Error t value Pr(>|t|)
1 == 0    5.622      1.027   5.474 3.37e-05 ***
---
Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
(Adjusted p values reported -- single-step method)
>
> confint(rice.glht)
     Simultaneous Confidence Intervals
Multiple Comparisons of Means: User-defined Contrasts
Fit: aov(formula = y ~ var, data = rice.dat)

Quantile = 2.1009
95% family-wise confidence level

Linear Hypotheses:
       Estimate lwr    upr
1 == 0 5.6222   3.4645 7.7799
>
> ## Scheffe test
> print(t.value <- summary(rice.glht)$test$tstat) # t-statistic
       1
5.474333
> print(df1 <- anova(rice.aov)$Df[1])             # treatment df
[1] 5
> print(df2 <- anova(rice.aov)$Df[2])             # error df
[1] 18
> print(f.value <- t.value^2/df1)                 # F statistic
       1
5.993664
> pf(f.value, df1, df2, lower.tail=FALSE)         # p-value by Scheffe test
          1
0.001963229
```

glht() 関数による対比の検定

glht() 関数では，特定の対比 $\mathbb{c} = (c_1, \ldots, c_a)^T$ に対して t 統計量

$$t_{\mathbb{c}} = \frac{\sum_{i=1}^{a} c_i \, \bar{y}_{i\cdot}}{\hat{\sigma}\sqrt{\sum_{i=1}^{a} c_i^2/n_i}} = 5.474$$

を計算し，次に p-値

$$p = \Pr\{|T| \geq |t_{\mathbb{c}}|\}$$

を出力する。これは対比に対する t 検定となる。

Scheffé 法では，この $t_{\mathbb{c}}$ の値の 2 乗を処理の自由度 $\nu_A = a - 1 = 5$ で割った値

$$F_{\mathbb{c}} = \frac{t_{\mathbb{c}}^2}{a-1} = \frac{(5.474)^2}{5} = 5.994$$

を検定統計量とし，p-値

$$p = \Pr\{F(\nu_A, \nu_e) \geq F_{\mathbb{c}}\} = 0.00196$$

を計算すればよい。

7　ノンパラメトリック法，FDR 法，その他の方法

7.1　ノンパラメトリック法

一般的に，伝統的な実験計画法に基づいた分散分析においては，誤差の確率分布が正規分布から多少はずれていても，結論に大きくは影響しない。しかし，調査データ，環境データ，臨床データなどでは，データの分布が正規分布から大きく外れることがある。本節で順位に基づいた**ノンパラメトリック法**による多重比較手法を説明する。

a 水準 $A_1, \dots, A_a$ をもつ一元配置完全無作為化法実験を考える。水準 A_i の繰返し数を n_i とし，第 k 番目の観測データ y_{ik} $(i = 1, \dots, a;\ k = 1, \dots, n_i)$ は，**確率分布** $\mathcal{F}_i$ に従うものとする。本節では簡単のために $n_1 = \cdots = n_a = n$ とする。サンプルサイズ n_i が異なる場合については，永田・吉田 (1997)，Hochberg & Tamhane (1987) などを参照されたい。

7.1.1　Steel-Dwass 法（対比較）

全ての水準のペア A_i, A_j に対して，両者が同じ分布に従うという帰無仮説のファミリー

$$\mathcal{H} = \{H_{ij}^0:\ \mathcal{F}_i \equiv \mathcal{F}_j, \quad 1 \le i, j \le a\}$$

を考える。

特定の水準組み合わせ (A_i, A_j) について，$2n$ 個のデータ y_{ik}, y_{jk} $(k = 1, \dots, n)$ を込みにして順位 $1, \dots, 2n$ を付ける。ここで，順位付けには (A_i, A_j) についての $2n$ 個のデータのみを用いることに注意が必要である。A_j のデータに付けられた順位を r_{jk} $(k = 1, \dots, n)$ とし，その順位和を

$$W_{ij} = \sum_{k=1}^{n} r_{jk}$$

とする。帰無仮説 $H_{ij}^0: \mathcal{F}_i \equiv \mathcal{F}_j$ のもとで W_{ij} の期待値と分散は

$$\mathrm{E}[W_{ij}] = \frac{n(2n+1)}{2}, \quad \mathrm{V}[W_{ij}] = \frac{n^2(2n+1)}{12}$$

で与えられる（Wilcoxon の順位和検定参照）。帰無仮説 $H_{ij}^0: \mathcal{F}_i \equiv \mathcal{F}_j$ は

$$|t_{ij}| = \frac{|W_{ij} - \frac{1}{2} - \mathrm{E}[W_{ij}]|}{\sqrt{\mathrm{V}[W_{ij}]}} > q(a, \infty; \alpha)/\sqrt{2} \tag{7.1}$$

のときに棄却される。ここで，$q(a, \infty; \alpha)$ は自由度 ∞ に対するスチューデント化した範囲の上側 α 点である。また，(7.1) 式の分子における $1/2$ は連続補正である。この方法を **Steel-Dwass 法**という (Steel (1960), Dwass (1960))。

7.1.2　Steel 法（対照処理との比較）

水準 A_1 が対照処理であり，対照処理との比較のファミリー

$$\mathcal{H} = \{H_{j1}^0:\ \mathcal{F}_j \equiv \mathcal{F}_1, \quad j = 2, \dots, a\}$$

を考える。

対比較の場合と同様に，水準 (A_1, A_j) の $2n$ 個のデータ y_{1k}, y_{jk} $(k = 1, \dots, n)$ を用いて順位付けを行ない，順位和・期待値・分散

$$W_{j1} = \sum_{k=1}^{n} r_{jk}, \quad \mathrm{E}[W_{j1}] = \frac{n(2n+1)}{2}, \quad \mathrm{V}[W_{j1}] = \frac{n^2(2n+1)}{12}$$

を計算する。両側検定では，

$$|t_{j1}| = \frac{|W_{j1} - \frac{1}{2} - \mathrm{E}[W_{j1}]|}{\sqrt{\mathrm{V}[W_{j1}]}} > d''(a-1, \infty; \alpha) \tag{7.2}$$

のときに，帰無仮説 H_{j1}^0: $\mathcal{F}_j \equiv \mathcal{F}_1$ を棄却する。$d''(a-1, \infty; \alpha)$ は両側 Dunnett 検定のための判定基準値である。この方法を **Steel 法**という (Steel, 1959)。

パラメトリック法の場合と同様に，対照処理との比較では片側対立仮説

$$H_{j1}^{A+}\colon\ \mathcal{F}_1 \prec \mathcal{F}_j$$

を考えることもできる。ここで "$\mathcal{F}_1 \prec \mathcal{F}_j$" は，$\mathcal{F}_j$ の方が $\mathcal{F}_1$ よりも確率的に大きな値をとることを意味する。すなわち，それぞれの分布関数を F_1, F_j とすれば，任意の x において $F_1(x) \geq F_j(x)$ が成り立つ。対立仮説 H_{j1}^{A+} に対しては，

$$t_{j1} = \frac{W_{j1} - \frac{1}{2} - \mathrm{E}[W_{j1}]}{\sqrt{\mathrm{V}[W_{j1}]}} > d'(a-1, \infty; \alpha) \tag{7.3}$$

のときに，帰無仮説 H_{j1}^0: $\mathcal{F}_j \equiv \mathcal{F}_1$ を棄却すればよい。

7.2 p 値に基づく方法

m 個の帰無仮説のファミリーを

$$\mathcal{H} = \{H_1^0, H_2^0, \ldots, H_m^0\} \tag{7.4}$$

とする。各帰無仮説 H_i^0 に対する検定方式が存在し，m 個の p 値

$$p_1, p_2, \ldots, p_m$$

が計算されているとする。

特定の帰無仮説 H_i^0 に対しては，

$$p_i \leq \alpha$$

のときに H_i^0 を棄却すれば，比較単位過誤率 $CWER$ は α 以下に制御される。

7.2.1 Bonferroni 法

Bonferroni 法では

$$p_i \leq \alpha/m$$

となる帰無仮説 H_i^0 を棄却する。任意の部分帰無仮説のもとでファミリー単位過誤率 $FWER$ が α 以下に保障される（第 3.3 節）。

7.2.2 Holm 法

Holm 法は，閉検定手順と Bonferroni の方法を組み合わせたものである。

Holm 法の判定手順（手順 1）

m 個の p 値 $p_1, p_2, \ldots, p_m$ を昇順に並べたものを

$$p_{(1)} \leq p_{(2)} \leq \cdots \leq p_{(i)} \leq \cdots \leq p_{(m)}$$

とし，対応する帰無仮説を $H^0_{(1)}, H^0_{(2)}, \ldots, H^0_{(i)}, \ldots, H^0_{(m)}$ とする。Holm 法による判定手順は以下のとおりである:

$$
\begin{array}{lcl}
p_{(1)} \le \alpha/m & \Longrightarrow & H^0_{(1)} \text{ を棄却} \\
p_{(2)} \le \alpha/(m-1) & \Longrightarrow & H^0_{(2)} \text{ を棄却} \\
\quad\vdots & & \\
p_{(i)} \le \alpha/(m-i+1) & \Longrightarrow & H^0_{(i)} \text{ を棄却} \\
\quad\vdots & & \\
p_{(m)} \le \alpha & \Longrightarrow & H^0_{(m)} \text{ を棄却}
\end{array}
$$

ただし，ある i で $p_{(i)} > \alpha/(m-i+1)$ となったときは，全ての $H^0_{(j)}$ $(j > i)$ を受容する。

Holm 法の考え方（手順 2）

帰無仮説のファミリー (7.4) 式に対して，任意の部分帰無仮説を

$$H^0_W = \bigcap_{i \in W} H^0_i, \quad W \subset \{1, \ldots, m\}$$

とする。閉検定手順では，この部分帰無仮説 H^0_W に対する検定方式を構成する必要がある。ここで，$w = \#\{W\}$ を集合 W の要素数として

$$\min_{i \in W} p_i \le \frac{\alpha}{w}$$

のときに H^0_W を棄却することにする。そうすれば，実際に H^0_W が成り立つときに，間違って H^0_W を棄却する確率は α 以下である。

特定の帰無仮説 H^0_i は，

$$i \in W \subset \{1, \ldots, m\}$$

となるような全ての H^0_W が棄却されたときに H^0_i を棄却する。そうすると，閉検定手順の議論（第 3.4 節）により，ファミリー単位過誤率 $FWER$ が α 以下に保障される。

注.　（**演習問題**）判定手順 1 と判定手順 2 は同じ結果を与えることを示せ。

p 値を昇順に並べたときの $p_{(i)}$ に着目する:

$$p_{(1)} \le \cdots \le p_{(j)} \le \cdots \le p_{(i)} \le \cdots \le p_{(m)}$$

手順 1 により $H^0_{(i)}$ が棄却されるとする。$1 \le j \le i$ に対して

$$p_{(j)} \le \frac{\alpha}{m-j+1}$$

である。次に，$p_{(i)}$ を含む任意の集合

$$\{p_{(j)}, \ldots, p_{(i)}, \ldots\}$$

を考える。この集合の最小値を

$$p_{(j)} = \min\{p_{(j)}, \ldots, p_{(i)}\}$$

とすると，この集合の要素数は

$$w = \#\{p_{(j)}, \ldots, p_{(i)}, \ldots\} \le m - j + 1$$

である。したがって，

$$\min\{p_{(j)},\ldots,p_{(i)},\ldots\}=p_{(j)}\le\frac{\alpha}{m-j+1}\le\frac{\alpha}{w}$$

が成り立ち，手順 2 の条件を満足する。

逆に，手順 2 により $H^0_{(i)}$ が棄却されるときは，$1\le j\le i$ に対して

$$\min\{p_{(j)},p_{(j+1)},\ldots,p_{(i)},\ldots,p_{(m)}\}=p_{(j)}\le\frac{\alpha}{m-j+1}$$

が成り立つ（すなわち，手順 1 により $H^0_{(i)}$ が棄却される）。

例 6

$m=5$ つの帰無仮説に対し，観測された p 値を小さいほうから順に

$$p_1=0.006,\ p_2=0.012,\ p_3=0.024,\ p_4=0.048,\ p_5=0.096$$

とする。

$\alpha=0.05$ とすると，$\alpha/m=0.05/5=0.01$ であるから，Bonferroni 法で棄却されるのは H^0_1 のみである。Holm 法の判定基準値は

$$0.05/5=0.01,\ 0.05/4=0.0125,\ 0.05/3=0.0167,\ 0.05/2=0.025,\ 0.05$$

であるから，H^0_1 と H^0_2 が棄却される。 □

7.3 FDR 法

遺伝子研究などにおいては，非常に多数の検定が行なわれることがある。このときファミリー単位過誤率 $FWER$ を制御する手法を適用すると，検定が極めて保守的になる（検出力が弱くなる）可能性がある。この問題に対応するため，Benjamini & Hochberg (1995) は FDR とよばれる過誤率を提案した。ここでは，その方法を紹介する。また，本節では，Benjamini & Hochberg (1995) で使用されている記号を使用する。

m 個の帰無仮説からなるファミリー $\mathcal{H}=\{H^0_1,\ldots,H^0_m\}$ を考える。この m 個の帰無仮説のうち，m_0 個が真に成り立っているとする $(0\le m_0\le m)$。検定結果は表 7.1 のように要約される。

表 7.1　m 個の帰無仮説に対する検定結果（各判定の数）

	帰無仮説を受容	帰無仮説を棄却	計
帰無仮説が真	U	V	m_0
帰無仮説が偽	T	S	$m-m_0$
計	$m-R$	R	m

R は棄却された帰無仮説の数で，観測可能な確率変数である。一方，帰無仮説 $H^0_1,\ldots,H^0_m$ のうち，どの m_0 個が真に成り立っているかは未知（m_0 の値自体も未知）であるから，確率変数 U, V, S, T は観測することができない。

ここで，間違って棄却された帰無仮説の割合

$$Q=\frac{V}{R}=\frac{V}{V+S}$$

を考える。ただし，$R = V + S = 0$ のときは，$Q = 0$ と定義する。この Q も確率変数である（false discovery proportion とよばれることもある）。その期待値

$$FDR = \mathrm{E}[Q] = \mathrm{E}[V/R] = \mathrm{E}[V/(V+S)] \tag{7.5}$$

を FDR (False Discovery Rate) とよぶ。一方ファミリー単位過誤率 $FWER$ は，真の帰無仮説が 1 つ以上間違って棄却される確率であるから

$$FWER = \Pr\{V > 0\} = \Pr\{Q > 0\}$$

である。

FDR は次の 2 つの性質をもっている:

(a) 全ての帰無仮説が真であれば（$m_0 = m$ ならば），$FDR = FWER$ となる。

(b) $m_0 < m$ ならば，$FDR \leq FWER$ となる。

性質 (b) により，$FDR \leq \alpha$ を保障する手法は，必ずしも $FWER \leq \alpha$ を保障するとは限らない。しかし性質 (a) により，すべての帰無仮説が真の場合には（完全帰無仮説のもとでは）$FWER$ は α 以下に保障されている。すなわち，FDR を制御する方法は，弱い意味で $FWER$ を制御する方法である（第 3.1 節）。

演習問題 15　性質 (a) と (b) を確認せよ。
ヒント。(a) $m_0 = m$ ならば，常に $S = 0$ であり，$V > 0$ のとき $Q = V/R = 1$。(b) 一般に $Q = V/(V+S) \leq 1$ に注意。 □

Benjamini & Hochberg (1995) は，FDR を制御するための次のような手順を与えている。帰無仮説 $H_1^0, \ldots, H_m^0$ に対して個別に検定を考え，その p-値を $p_1, \ldots, p_m$ とする。それを昇順に並べたものを

$$p_{(1)} \leq \cdots \leq p_{(m)}$$

とし，$p_{(i)}$ に対応する帰無仮説を $H_{(i)}^0$ で表わす。

- $p_{(i)} \leq (i/m)\,\alpha$ を満たす最大の i を k とする。すなわち次式が成り立つ。

$$p_{(k)} \leq (k/m)\,\alpha$$
$$p_{(i)} > (i/m)\,\alpha \quad (i = k+1, \ldots, m)$$

- このとき，$H_{(i)}^0$ $(i = 1, \ldots, k)$ を全て棄却する。

この方法は，各帰無仮説に対する検定が独立であれば FDR を α 以下に制御する (Benjamini & Hochberg, 1995)。

Holm 法で示した例 6 を考える:

$$p_1 = 0.006,\ p_2 = 0.012,\ p_3 = 0.024,\ p_4 = 0.048,\ p_5 = 0.096$$

$p_3 = 0.024 \leq (3/5) \times 0.05 = 0.03$ であるから，H_1^0, H_2^0, H_3^0 が棄却される。

表 7.2　各手法の判定基準値

観測された p 値	Bonferroni 法	Holm 法	FDR 法
0.006	0.01	0.01	0.01
0.012	0.01	0.0125	0.02
0.024	0.01	0.0167	0.03
0.048	0.01	0.025	0.04
0.096	0.01	0.05	0.05

8　おわりに

8.1　適用上の問題点

実験データに多重比較を適用するためには，

(1) どのタイプの比較を行なうか（表 1.4）
(2) 過誤率をどのように制御するか

を決めなければならない。どちらの問題も，統計学的な考察よりも，研究の具体的な目的によって決まるべきものである。

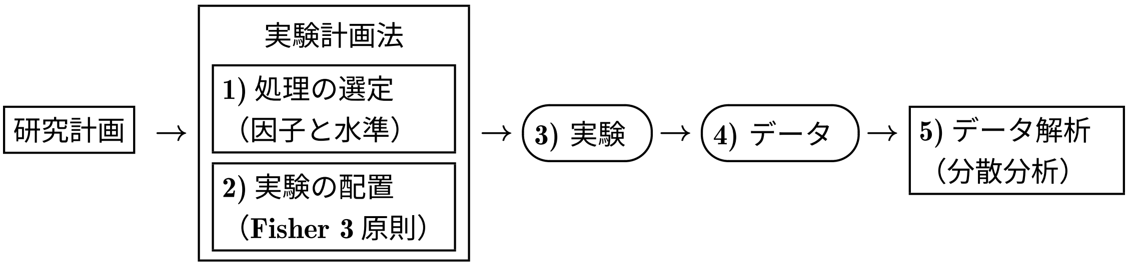


図 8.1　研究計画からデータ解析まで

注.　**R. A. Fisher (1935).** ***The Design of Experiments***
"If the design of an experiment is faulty, any method of interpretation which makes it out to be decisive must be faulty too."（実験の計画が間違っているのに，決定的な解釈を導くような手法があるとすれば，その解釈法もまた間違ったものに違いない。）

また，統計手法を必要とする場面として，

(a) 企業内・研究所内における技術開発の場面
(b) 科学雑誌へ論文として投稿する場面

がある。この両方において，同じように過誤率を制御する必要はない（と筆者は思っている）。

現在は，複数の比較を行なうとき，ファミリー単位過誤率 $FWER$ を制御する手法を用いることが主流である。しかし，一般にファミリー単位過誤率を保障する方法は保守的であることに注意しなければならない（有意な差を検出しにくいことを保守的 (conservative) という）。

技術開発の初期の段階で，保守的な方法を用いると，有効な技術を見逃してしまう可能性がある。また，毒性や副作用の研究では，保守的な手法を用いることは危険ですらある。

一方，科学雑誌へ論文を投稿した場合，論文内で行なわれている比較が，データを取ったあと恣意的に考えられたものかどうか，編集者や査読者には判断できない場合が多い。したがって，論文投稿に際して効果の有効性を主張するためには，ファミリー基準過誤率が保障されている方が，編集者にとっても，読者にとっても安心である。

注.　従来，Duncan の方法はファミリー単位過誤率 $FWER$ を保障しないという理由で，統計理論家や論文のレフリーに不評であった。しかし現在，$FWER$ を保障しない FDR 法は広く使われている。
Google Scholar での被引用数（2026 年 1 月 15 日現在）

Duncan (1955)	46,802
Benjamini & Hochberg (1995)	125,355

8.2　再現性と確認実験

ASA（アメリカ統計学会）の p 値に関する声明

仮説検定や p 値の使用法，結果の再現性に関しては，現在も活発な議論が展開されている。

- ○ **Wasserstein, R. L. and Lazar, N. A. (2016): The ASA's statement on p-values: context, process, and purpose, *The American Statistician*, 70, 129–133.**
- ○ **Wasserstein, R. L., Schirm A. L. and Lazar, N. A. (2019): Moving to a world beyond "$p < 0.05$", *The American Statistician*, 73, 1–19.**
- ○ **Yoav Benjamini, et al. (2021): The ASA president's task force statement on statistical significance and replicability, *The Annals of Applied Statistics*, Vol. 15, 1084–1085.**

確認実験

m 個の帰無仮説のファミリー

$$\mathcal{H} = \{H_1^0, H_2^0, \ldots, H_m^0\}$$

に対して，p 値

$$p_1, p_2, \ldots, p_m$$

が計算されているとする。値の小さい p 値に着目する。

$$p_{(1)} = \min\{p_1, p_2, \cdots, p_m\}$$

ここで，すぐに論文を書くのではなく，もう一度，帰無仮説 $H_{(1)}^0$ に対する実験を行ない有意性を確認すれば，$FWER$ は保障される。

8.3　人の名前の読み方

表 8.1　人の名前の読み方（英語）

かなり確かな読み方	たぶん正しい読み方	聞き伝え
• **Dunnett:** ダネット	• **Duncan:** ダンカン	• **Bonferroni:** ボンフェローニ
• **Hayter:** ヘイター	• **Gabriel:** ガブリエル	• **Einot:** エイノット
• **Hochberg:** ホックバーグ	• **Newman:** ニューマン	• **Holm:** ホルム，ホウム
• **Hommel:** ホメル	• **Tukey:** テューキー	• **Keuls:** クールズ
• **Hsu:** シュー		• **Kramer:** クレイマー
• **Marcus:** マーカス		• **Peritz:** ペリッツ
• **Tamhane:** タムハーニ		• **Scheffé:** シェフェィ，シェフィ，シェフェ
• **Shaffer:** シェィファー		• **Simes:** サイムズ
		• **Welsch:** ウェルシュ

9 数表

表 9.1 REGWQ 法のための数表 ($\alpha = 0.05$)

Critical values q(p, nu; alpha_p) for REGW multiple range tests

	a-> 4	4	4	5	5	5	5	6	6	6	6	6
nu	p-> 2	3	4	2	3	4	5	2	3	4	5	6
1	32.819	32.819	32.819	37.082	37.082	37.082	37.082	40.408	40.408	40.408	40.408	40.408
2	8.718	8.718	9.798	9.772	10.811	10.811	10.881	10.723	11.859	12.036	12.036	12.036
3	5.878	5.910	6.825	6.385	7.138	7.138	7.502	6.826	7.625	7.910	7.910	8.037
4	4.923	5.040	5.757	5.274	5.893	5.893	6.287	5.573	6.220	6.497	6.497	6.706
5	4.458	4.602	5.218	4.739	5.284	5.284	5.673	4.975	5.540	5.801	5.801	6.033
6	4.184	4.339	4.896	4.427	4.926	4.926	5.305	4.630	5.142	5.390	5.390	5.628
7	4.006	4.165	4.681	4.224	4.691	4.691	5.060	4.406	4.883	5.121	5.121	5.359
8	3.880	4.041	4.529	4.082	4.525	4.529	4.886	4.249	4.701	4.930	4.930	5.167
9	3.786	3.948	4.415	3.977	4.403	4.415	4.755	4.134	4.567	4.789	4.789	5.024
10	3.714	3.877	4.327	3.896	4.308	4.327	4.654	4.045	4.463	4.679	4.679	4.912
11	3.657	3.820	4.256	3.832	4.233	4.256	4.574	3.975	4.381	4.593	4.593	4.823
12	3.611	3.773	4.199	3.780	4.172	4.199	4.508	3.918	4.314	4.522	4.522	4.750
13	3.572	3.734	4.151	3.737	4.122	4.151	4.453	3.871	4.259	4.463	4.463	4.690
14	3.540	3.701	4.111	3.700	4.079	4.111	4.407	3.831	4.213	4.414	4.414	4.639
15	3.512	3.673	4.076	3.670	4.043	4.076	4.367	3.798	4.173	4.372	4.372	4.595
16	3.488	3.649	4.046	3.643	4.011	4.046	4.333	3.768	4.139	4.335	4.335	4.557
17	3.467	3.628	4.020	3.620	3.984	4.020	4.303	3.743	4.109	4.304	4.304	4.524
18	3.449	3.609	3.997	3.599	3.960	3.997	4.276	3.721	4.083	4.276	4.276	4.494
19	3.433	3.593	3.977	3.581	3.938	3.977	4.253	3.701	4.060	4.251	4.253	4.468
20	3.418	3.578	3.958	3.565	3.919	3.958	4.232	3.684	4.039	4.229	4.232	4.445
24	3.373	3.532	3.901	3.515	3.860	3.901	4.166	3.629	3.974	4.160	4.166	4.373
30	3.329	3.486	3.845	3.466	3.802	3.845	4.102	3.576	3.912	4.092	4.102	4.301
40	3.286	3.442	3.791	3.418	3.745	3.791	4.039	3.524	3.850	4.027	4.039	4.232
60	3.244	3.399	3.737	3.371	3.690	3.737	3.977	3.474	3.791	3.962	3.977	4.163
120	3.203	3.356	3.685	3.326	3.636	3.685	3.917	3.425	3.733	3.900	3.917	4.096
Inf	3.163	3.314	3.633	3.282	3.584	3.633	3.858	3.377	3.676	3.838	3.858	4.030

Critical values q(p, nu; alpha_p) for REGW multiple range tests

	a-> 7	7	7	7	7	7	8	8	8	8	8	8	8
nu	p-> 2	3	4	5	6	7	2	3	4	5	6	7	8
1	43.119	43.119	43.119	43.119	43.119	43.119	45.397	45.397	45.397	45.397	45.397	45.397	45.397
2	11.597	12.821	13.011	13.011	13.011	13.011	12.409	13.716	13.918	13.918	13.918	13.918	13.918
3	7.218	8.059	8.360	8.471	8.471	8.478	7.573	8.452	8.767	8.884	8.912	8.912	8.912
4	5.834	6.507	6.795	6.940	6.940	7.053	6.068	6.763	7.062	7.213	7.292	7.292	7.347
5	5.180	5.762	6.032	6.183	6.183	6.330	5.362	5.958	6.237	6.393	6.487	6.487	6.582
6	4.804	5.329	5.584	5.735	5.735	5.895	4.957	5.494	5.755	5.910	6.009	6.009	6.122
7	4.561	5.048	5.291	5.439	5.439	5.606	4.697	5.193	5.441	5.592	5.693	5.693	5.815
8	4.392	4.851	5.085	5.230	5.230	5.399	4.516	4.982	5.220	5.368	5.469	5.469	5.596
9	4.267	4.706	4.932	5.074	5.074	5.244	4.383	4.827	5.057	5.202	5.303	5.303	5.432
10	4.171	4.594	4.814	4.954	4.954	5.124	4.281	4.708	4.931	5.074	5.174	5.174	5.304
11	4.096	4.506	4.721	4.858	4.858	5.028	4.200	4.615	4.832	4.972	5.071	5.071	5.202
12	4.034	4.434	4.645	4.781	4.781	4.950	4.135	4.538	4.751	4.889	4.987	4.987	5.119
13	3.984	4.375	4.582	4.716	4.716	4.884	4.082	4.475	4.684	4.820	4.917	4.917	5.049
14	3.941	4.325	4.529	4.661	4.661	4.829	4.037	4.423	4.628	4.762	4.859	4.859	4.990
15	3.905	4.283	4.484	4.615	4.615	4.782	3.998	4.378	4.580	4.713	4.809	4.809	4.940
16	3.874	4.246	4.444	4.575	4.575	4.741	3.965	4.339	4.539	4.670	4.765	4.765	4.896
17	3.847	4.214	4.410	4.539	4.539	4.705	3.936	4.305	4.502	4.633	4.727	4.727	4.858
18	3.823	4.186	4.380	4.509	4.509	4.673	3.911	4.275	4.471	4.600	4.694	4.694	4.824
19	3.802	4.161	4.354	4.481	4.481	4.645	3.889	4.249	4.443	4.571	4.664	4.664	4.794
20	3.783	4.139	4.330	4.457	4.457	4.620	3.869	4.226	4.418	4.545	4.638	4.638	4.768
24	3.725	4.070	4.256	4.380	4.380	4.541	3.807	4.153	4.339	4.464	4.555	4.555	4.684
30	3.668	4.003	4.184	4.305	4.305	4.464	3.747	4.082	4.263	4.385	4.474	4.474	4.601
40	3.612	3.938	4.114	4.232	4.232	4.388	3.688	4.013	4.189	4.308	4.395	4.395	4.521
60	3.559	3.874	4.046	4.161	4.163	4.314	3.632	3.946	4.117	4.232	4.318	4.318	4.441
120	3.506	3.812	3.979	4.091	4.096	4.241	3.576	3.881	4.046	4.159	4.242	4.242	4.363
Inf	3.456	3.752	3.914	4.023	4.030	4.170	3.523	3.817	3.978	4.087	4.168	4.170	4.286

参考文献

[1] Yoav Benjamini, et al. The ASA president's task force statement on statistical significance and replicability. *The Ann. Appl. Statist.*, Vol. 15, No. 3, pp. 1084–1085, 2021.

[2] Yoav Benjamini and Yosef Hochberg. Controlling the false discovery rate: A practical and powerful approach to multiple testing. *J. Roy. Statist. Soc. Ser. B*, Vol. 57, No. 1, pp. 289–300, 1995.

[3] D. B. Duncan. Multiple range and multiple F tests. *Biometrics*, Vol. 11, pp. 1–42, 1955.

[4] C. W. Dunnett. A multiple comparison procedure for comparing several treatments with a control. *J. Amer. Statist. Assoc.*, Vol. 50, pp. 1096–1121, 1955.

[5] M. Dwass. Some k-sample rank-order tests. In I. Olkin, et al., editors, *Contributions to probability and statistics*, pp. 198–202. Stanford University Press, 1960.

[6] I. Einot and K. R. Gabriel. A study of the powers of several methods of multiple comparisons. *J. Amer. Statist. Assoc.*, Vol. 70, pp. 574–583, 1975.

[7] R. A. Fisher. *The Design of Experiments.* Oliver & Boyd, Edinburgh and London, 1935.

[8] A. J. Hayter. A proof of the conjecture that the tukey-kramer multiple comparisons procedure is conservative. *Annals of Statistics*, Vol. 12, pp. 61–75, 1984.

[9] Y. Hochberg and A. C. Tamhane. *Multiple Comparison Procedures.* Wiley, New York, 1987.

[10] S. Holm. A simple sequentially rejective multiple test procedure. *Scand. J. Statist.*, Vol. 6, pp. 65–70, 1979.

[11] Jason C. Hsu. *Multiple Comparisons: Theory and Methods.* Chapman and Hall, London, 1996.

[12] M. Keuls. The use of the 'studentized range' in connection with an analysis of variance. *Euphytica*, Vol. 1, pp. 112–122, 1952.

[13] C. Y. Kramer. Extension of multiple range test to group means with unequal numbers of replications. *Biometrics*, Vol. 12, pp. 307–310, 1956.

[14] W. Liu, Tetsuhisa Miwa, and A. J. Hayter. Simultaneous confidence interval estimation for successive comparisons of ordered treatment effects. *J. Statist. Plann. Inference*, Vol. 88, pp. 75–86, 2000.

[15] R. Marcus, E. Peritz, and K. R. Gabriel. On closed testing procedures with special reference to ordered analysis of variance. *Biometrika*, Vol. 63, pp. 655–660, 1976.

[16] R. G. Miller. *Simultaneous statistical inference.* Springer-Verlag, New York, 2nd edition, 1981.

[17] T. Miwa and A. J. Hayter. Combining the advantages of one-sided and two-sided test procedures for comparing several treatment effects. *J. Amer. Statist. Assoc.*, Vol. 94, No. 445, pp. 302–307, 1999.

[18] 三輪哲久. 農業研究分野における多重比較論争. 応用統計学, Vol. 26, No. 2, pp. 99–109, 1997.

[19] 三輪哲久. 実験計画法と分散分析. 朝倉書店, 東京, 2015.

[20] Tetsuhisa Miwa, A. J. Hayter, and Satoshi Kuriki. The evaluation of general non-centred orthant probabilities. *J. Roy. Statist. Soc. Ser. B*, Vol. 65, No. 1, pp. 223–234, 2003.

[21] Tetsuhisa Miwa, A. J. Hayter, and Wei Liu. Calculations of level probabilities for normal random variables with unequal variances with applications to Bartholomew's test in unbalanced one-way models. *Comput. Statist. Data Anal.*, Vol. 34, pp. 17–32, 2000.

[22] 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, 東京, 1997.

[23] J. A. Nelder. Discussion to the paper by O'Neill and Wetherill. *J. Roy. Statist. Soc. Ser. B*, Vol. 36, pp. 244–246, 1971.

[24] D. Newman. The distribution of the range in samples from a normal population, expressed in terms of an independent estimate of standard deviation. *Biometrika*, Vol. 31, pp. 20–30, 1939.

[25] T. A. Ryan. Significance tests for multiple comparison of proportions, variances, and other statistics. *Psycol. Bull.*, Vol. 57, pp. 318–328, 1960.

[26] 坂巻顕太郎, 寒水孝司, 濱崎俊光. 多重比較法. 朝倉書店, 東京, 2019.

[27] H. Scheffé. *The Analysis of Variance*. Wiley, New York, 1959.

[28] R. G. D. Steel. A multiple comparison rank sum test: Treatments versus control. *Biometrics*, Vol. 15, pp. 560–572, 1959.

[29] R. G. D. Steel. A rank sum test for comparing all pairs of treatments. *Technometrics*, Vol. 2, pp. 197–207, 1960.

[30] 丹後俊郎, 松井茂之（編）. 医学統計学ハンドブック. 朝倉書店, 東京, 2018.

[31] J. W. Tukey. The problem of multiple comparisons. *Unpublished report*, 1953. Princeton University.

[32] R. L. Wasserstein and N. A. Lazar. The ASA's statement on p-values: context, process, and purpose. *The American Statistician*, Vol. 70, pp. 129–133, 2016.

[33] R. L. Wasserstein, A. L. Schirm, and N. A. Lazar. Moving to a world beyond "$p < 0.05$". *The American Statistician*, Vol. 73, pp. 1–19, 2019.

[34] R. E. Welsch. Stepwise multiple comparison procedures. *J. Amer. Statist. Assoc.*, Vol. 72, pp. 566–575, 1977.

演習問題略解

演習問題 1　対立仮説 $H^{A+}\colon \mu_1-\mu_2>0$ のもとで，第 III 種の過誤の確率は

$$\begin{aligned}&\Pr\{t<-t(\nu_e;\alpha/2)\mid H^{A+}\colon \mu_1-\mu_2>0\}\\&\qquad\le \Pr\{t<-t(\nu_e;\alpha/2)\mid H^{0}\colon \mu_1-\mu_2=0\}=\alpha/2\end{aligned}$$

が成り立つ。対立仮説 $H^{A-}\colon \mu_1-\mu_2<0$ の場合も同様である。

演習問題 2　実際に成立している帰無仮説を H_i^0 $(i\in W\subset\{1,\ldots,m\})$ とする。W は添え字全体の集合 $\{1,\ldots,m\}$ の部分集合である。そうすると

$$\begin{aligned}FWER&=\Pr\{\text{どれかの } i\in W \text{ で } \theta_i^0\notin[\hat{L}_i(\mathrm{y}),\ \hat{U}_i(\mathrm{y})]\}\\&=1-\Pr\{\text{すべての } i\in W \text{ で } \theta_i^0\in[\hat{L}_i(\mathrm{y}),\ \hat{U}_i(\mathrm{y})]\}\\&\le 1-\Pr\{1\le i\le m \text{ で } \theta_i\in[\hat{L}_i(\mathrm{y}),\ \hat{U}_i(\mathrm{y})]\}\\&\le 1-(1-\alpha)=\alpha\end{aligned}$$

ここで，W は $\{1,\ldots,m\}$ の部分集合であるから

$$\begin{aligned}&\Pr\{1\le i\le m \text{ で } \theta_i\in[\hat{L}_i(\mathrm{y}),\ \hat{U}_i(\mathrm{y})]\}\\&\qquad\le\Pr\{\text{すべての } i\in W \text{ で } \theta_i^0\in[\hat{L}_i(\mathrm{y}),\ \hat{U}_i(\mathrm{y})]\}\end{aligned}$$

が成り立つ。

演習問題 3　真の θ_i の値について，$\theta_i>\theta_i^0$ の成り立っている添え字の集合を W^+ とし，$\theta_i<\theta_i^0$ の成り立っている添え字の集合を W^- とする。$\theta_i=\theta_i^0$ の成り立っている添え字の集合は演習問題 2 と同様に W とすると $W^+\cup W\cup W^-=\{1,2,\ldots,m\}$ である。

$$\begin{cases}\theta_i>\theta_i^0 & i\in W^+\\ \theta_i=\theta_i^0 & i\in W\\ \theta_i<\theta_i^0 & i\in W^-\end{cases}$$

第 III 種の過誤が起こるのは

どれかの $i\in W^+$ において $[\hat{L}_i(\mathrm{y}),\ \hat{U}_i(\mathrm{y})]<\theta_i^0$　(したがって $\theta_i<\theta_i^0$ と判定)　(∗)
または
どれかの $i\in W^-$ において $[\hat{L}_i(\mathrm{y}),\ \hat{U}_i(\mathrm{y})]>\theta_i^0$　(したがって $\theta_i>\theta_i^0$ と判定)　(∗∗)

となる場合である。(∗)，(∗∗) いずれの場合も $\theta_i\notin[\hat{L}_i(\mathrm{y}),\ \hat{U}_i(\mathrm{y})]$ である。演習問題 2 と同様にして，第 III 種の過誤に関する $FWER$ は

$$\begin{aligned}FWER&=\Pr\{\text{どれかの } i\in W^+ \text{ で } [\hat{L}_i(\mathrm{y}),\ \hat{U}_i(\mathrm{y})]<\theta_i^0\\&\qquad\qquad\text{または，どれかの } i\in W^- \text{ で } [\hat{L}_i(\mathrm{y}),\ \hat{U}_i(\mathrm{y})]>\theta_i^0\ \}\\&\le\Pr\{\text{どれかの } i\in(W^+\cup W^-) \text{ で } \theta_i\notin[\hat{L}_i(\mathrm{y}),\ \hat{U}_i(\mathrm{y})]\}\\&=1-\Pr\{\text{すべての } i\in(W^+\cup W^-) \text{ で } \theta_i\in[\hat{L}_i(\mathrm{y}),\ \hat{U}_i(\mathrm{y})]\}\\&\le 1-\Pr\{1\le i\le m \text{ で } \theta_i\in[\hat{L}_i(\mathrm{y}),\ \hat{U}_i(\mathrm{y})]\}\\&\le 1-(1-\alpha)=\alpha\end{aligned}$$

演習問題 4　$m=2$ に対して

$$\Pr(E_1 \cup E_2) = \Pr(E_1 \cup (E_2 \setminus E_1)) = \Pr(E_1) + \Pr(E_2 \setminus E_1) \le \Pr(E_1) + \Pr(E_2)$$

が成立する。m に関する数学的帰納法を使う。

演習問題 5

$$\begin{aligned}
&\alpha/m \le 1-(1-\alpha)^{1/m}\\
&\qquad \Updownarrow \quad \text{項を移項して}\\
&(1-\alpha)^{1/m} \le 1-\alpha/m\\
&\qquad \Updownarrow \quad x^m \text{ は単調関数だから両辺を } m \text{ 乗して}\\
&(1-\alpha) \le \bigl(1-\alpha/m\bigr)^m
\end{aligned} \tag{$*$}$$

したがって，$(*)$ 式を数学的帰納法を用いて証明すればよい。

$m=1$ のときは等号が成り立つ。任意の α $(0 \le \alpha \le 1)$ に対して $(*)$ が成り立つとして，$m+1$ の場合を考える。$m+1$ のときの $(*)$ 式の右辺は

$$\begin{aligned}
\Bigl(1-\frac{\alpha}{m+1}\Bigr)^{(m+1)} &= \Bigl(1-\frac{\alpha}{m+1}\Bigr)^m \times \Bigl(1-\frac{\alpha}{m+1}\Bigr)\\
&= \Bigl(1-\frac{1}{m}\frac{m\,\alpha}{m+1}\Bigr)^m \times \Bigl(1-\frac{\alpha}{m+1}\Bigr)\\
&\ge \Bigl(1-\frac{m\,\alpha}{m+1}\Bigr) \times \Bigl(1-\frac{\alpha}{m+1}\Bigr) && (**)\\
&= 1-\frac{m\,\alpha}{m+1}-\frac{\alpha}{m+1}+\frac{m\,\alpha^2}{(m+1)^2}\\
&= 1-\alpha+\frac{m\,\alpha^2}{(m+1)^2} \ge 1-\alpha
\end{aligned}$$

となり，$(*)$ が成り立つ。ここで，$(**)$ の不等式は，$\alpha' = (m\,\alpha)/(m+1)$ とおいて m の場合の $(*)$ を利用している。

- **$(*)$ の別解**

$0 \le \alpha \le 1$ に対して

$$f(\alpha) = \bigl(1-\alpha/m\bigr)^m - (1-\alpha)$$

とおく。

$$\begin{aligned}
&f(0) = 0\\
&f'(\alpha) = -(1-\alpha/m)^{m-1} + 1 \ge 0
\end{aligned}$$

したがって，$f(\alpha) \ge 0$。

演習問題 6　実際に成立している帰無仮説を H_i^0 $(i \in W \subset \{1,\ldots,m\})$ とする。事象として

$$\{i \in W \text{ のどれかの } H_i^0 \text{ を棄却}\} \Longrightarrow \{H_W^0 \text{ が棄却されている}\}$$

であるから

$$\begin{aligned}
FWER &= \Pr\{i \in W \text{ のどれかの } H_i^0 \text{ を棄却} \mid H_W^0\}\\
&\le \Pr\{H_W^0 \text{ を棄却} \mid H_W^0\} \le \alpha
\end{aligned}$$

演習問題 7　4.2 節のスチューデント化した範囲の説明により，完全帰無仮説 $H^0_{1\cdots a}$： $\mu_1 = \cdots = \mu_a$ のもとでの LSD 法の $FWER$ は次式で計算できる。

$$
\begin{aligned}
FWER &= \Pr\{\text{どれかの } i,\ j \text{ で } |\bar{y}_{i\cdot} - \bar{y}_{j\cdot}| > LSD(\alpha)\} \\
&= \Pr\left\{\max_{i,j} |\bar{y}_{i\cdot} - \bar{y}_{j\cdot}| > LSD(\alpha)\right\} \\
&= \Pr\left\{\max_{i,j} |\bar{y}_{i\cdot} - \bar{y}_{j\cdot}| > \hat{\sigma}\sqrt{2/n} \cdot t(\nu_e; \alpha/2)\right\} \\
&= \Pr\left\{\max_{i,j} \frac{|\bar{y}_{i\cdot} - \bar{y}_{j\cdot}|}{\hat{\sigma}/\sqrt{n}} > \sqrt{2} \cdot t(\nu_e; \alpha/2)\right\}\Big\} \\
&= \Pr\{Q > \sqrt{2} \cdot t(\nu_e; \alpha/2)\}
\end{aligned}
$$

LSD 法の $FWER$

```
> ## FWER of Unprotected LSD (Exercise 8)
> print(q.lsd <- sqrt(2.0)*t.value)
[1] 2.888209
> print(FWER.LSD <- ptukey(q.lsd, a, nue, lower.tail=FALSE))
[1] 0.4107989
```

完全帰無仮説 $H^0_{1\cdots 7}$: $\mu_1 = \cdots = \mu_7$ のもとでの LSD 法のファミリー単位過誤率は $FWER = 0.41$ である。

演習問題 8

Bonferroni 調整による LSD 法

```
> ## Bonferroni adjustment for LSD procesure (Exercise 9)
> print(m <- a*(a-1)/2)
[1] 21
> print(t.bon <- qt(0.5*alpha/m, nue, lower.tail=FALSE))
[1] 3.318605
> print(LSD.bon <- sigma.hat*sqrt(2.0/n)*t.bon)
[1] 17.09228
> print(q.bon <- sqrt(2.0)*t.bon)
[1] 4.693216
> print(FWER.Bon <- ptukey(q.bon, a, nue, lower.tail=FALSE))
[1] 0.03433132
```

Bonferroni 法では，名目上の 0.05 に対して $FWER = 0.034$ となっている（有意差を検出しにくくなっている)。

演習問題 9

$H^0_{1\cdots 6}$: $\mu_1 = \cdots = \mu_6$ のもとでの $FWER$

```
> ## FWER of Protected LSD (Exercise 10)
> print(FWER.PLSD <- ptukey(q.lsd, a-1, nue, lower.tail=FALSE))
[1] 0.343549
```

保護付き lSD 法では，

$$\mu_1 = \cdots = \mu_6 \ll \mu_7$$

のときに $FWER = 0.34$ となる。

演習問題 10　ヒントの最後の部分を変形すれば

$$
\begin{aligned}
&|(\bar{y}_{i}.-\bar{y}_{j}.)-(\mu_i-\mu_j)| \le HSD_a(\alpha)\\
&\qquad\qquad \Updownarrow\\
&\bar{y}_{i}.-\bar{y}_{j}.-HSD_a(\alpha) \le \mu_i-\mu_j \le \bar{y}_{i}.-\bar{y}_{j}.+HSD_a(\alpha)
\end{aligned}
$$

が得られる。

演習問題 11　$a=3$ のとき，帰無仮説のファミリーは

$$\mathcal{H}=\{H_{12}^0\colon \mu_1=\mu_2,\ H_{13}^0\colon \mu_1=\mu_3,\ H_{23}^0\colon \mu_2=\mu_3\}$$

である。その共通部分すべてからなる閉包集合は

$$\bar{\mathcal{H}}=\{H_{12}^0\colon \mu_1=\mu_2,\ H_{13}^0\colon \mu_1=\mu_3,\ H_{23}^0\colon \mu_2=\mu_3,\ H_{123}^0\colon \mu_1=\mu_2=\mu_3\}$$

となる。ここで，

$$H_{12}^0\cap H_{13}^0=H_{12}^0\cap H_{23}^0=H_{13}^0\cap H_{23}^0=H_{12}^0\cap H_{13}^0\cap H_{23}^0=H_{123}^0$$

の関係が成り立つので，閉包集合 $\bar{\mathcal{H}}$ の要素は限られている。
　したがって，帰無仮説 $H_{12}^0\colon \mu_1=\mu_2$ を棄却するためには，

$$
\begin{aligned}
&H_{12}^0\colon \mu_1=\mu_2\\
&H_{123}^0\colon \mu_1=\mu_2=\mu_3
\end{aligned}
$$

を有意水準 α で検定すればよい。すなわち，完全帰無仮説 $H_{123}^0\colon \mu_1=\mu_2=\mu_3$ が有意水準 α で棄却されたあとでは，仮説 $H_{12}^0\colon \mu_1=\mu_2$ を有意水準 α で検定してよい。
　$a=4$ のとき，帰無仮説のファミリーは

$$
\begin{aligned}
\mathcal{H}=\{&H_{12}^0\colon \mu_1=\mu_2,\ H_{13}^0\colon \mu_1=\mu_3,\ H_{14}^0\colon \mu_1=\mu_4\\
&H_{23}^0\colon \mu_2=\mu_3,\ H_{24}^0\colon \mu_2=\mu_4,\ H_{34}^0\colon \mu_3=\mu_4\}
\end{aligned}
$$

である。その閉包には，完全帰無仮説 $H_{1234}^0\colon \mu_1=\mu_2=\mu_3=\mu_4$ のほかに，

$$H_{12,34}^0\colon \mu_1=\mu_2,\ \mu_3=\mu_4$$

のように 2 つのグループに分かれた部分帰無仮説も含まれる。したがって，閉検定手順により帰無仮説 $H_{12}^0\colon \mu_1=\mu_2$ を棄却するためには，

$$
\begin{aligned}
&H_{12}^0\colon \mu_1=\mu_2\\
&H_{123}^0\colon \mu_1=\mu_2=\mu_3,\ H_{124}^0\colon \mu_1=\mu_2=\mu_4\\
&H_{12,34}^0\colon \mu_1=\mu_2,\ \mu_3=\mu_4\\
&H_{1234}^0\colon \mu_1=\mu_2=\mu_3=\mu_4
\end{aligned}
$$

を有意水準 α で検定する必要がある。。

演習問題 12　まず，

$$\frac{(\bar{y}_{i}.-\mu_i)-(\bar{y}_{1}.-\mu_1)}{\hat{\sigma}\sqrt{2/n}}=\frac{u_i-u_1}{\sqrt{2\cdot\chi^2/\nu_e}}$$

と表わされることに注意する。片側信頼区間については

$$
\begin{aligned}
1-\alpha &= \Pr\Big\{ \max_{2\le i\le a} \frac{u_i - u_1}{\sqrt{2\cdot\chi^2/\nu_e}} \le d'(a-1,\nu_e;\alpha) \Big\} \\
&= \Pr\Big\{ \max_{2\le i\le a} \frac{(\bar{y}_{i}.-\mu_i)-(\bar{y}_{1}.-\mu_1)}{\hat{\sigma}\sqrt{2/n}} \le d'(a-1,\nu_e;\alpha) \Big\} \\
&= \Pr\Big\{ \text{すべての } i\ (2\le i\le a) \text{ で } \frac{(\bar{y}_{i}.-\mu_i)-(\bar{y}_{1}.-\mu_1)}{\hat{\sigma}\sqrt{2/n}} \le d'(a-1,\nu_e;\alpha) \Big\} \\
&= \Pr\Big\{ \text{すべての } i\ (2\le i\le a) \text{ で } (\bar{y}_{i}.-\mu_i)-(\bar{y}_{1}.-\mu_1) \le D'(\alpha) \Big\} \\
&= \Pr\Big\{ \text{すべての } i\ (2\le i\le a) \text{ で } \bar{y}_{i}.-\bar{y}_{1}.-D'(\alpha) \le \mu_i-\mu_1 \Big\}
\end{aligned}
$$

より，片側同時信頼区間が (5.5) 式で与えられる。両側同時信頼区間も Tukey 法と同様に示すことができる（演習問題 10 参照)。

演習問題 13　データから計算される対比の推定量 $\sum_{i=1}^{a} c_i\,\bar{y}_{i}.$ に関して，その期待値と分散は

$$
\mathrm{E}\Big[\sum_{i=1}^{a} c_i\,\bar{y}_{i}.\Big] = \sum_{i=1}^{a} c_i\,\mu_i, \quad \mathrm{V}\Big[\sum_{i=1}^{a} c_i\,\bar{y}_{i}.\Big] = \sigma^2 \sum_{i=1}^{a} c_i^2/n_i
$$

である。すなわち，

$$
\sum_{i=1}^{a} c_i\,\bar{y}_{i}. \sim N\Big(\sum_{i=1}^{a} c_i\,\mu_i,\ \sigma^2\sum_{i=1}^{a} c_i^2/n_i\Big)
$$

したがって，帰無仮説 $H_{\mathbb{c}}^0$: $\sum_{i=1}^{a} c_i\,\mu_i = 0$ のもとで

$$
u = \frac{\sum_{i=1}^{a} c_i\,\bar{y}_{i}.}{\sigma\sqrt{\sum_{i=1}^{a} c_i^2/n_i}}
$$

は標準正規分布に従う。標準偏差 σ を推定値 $\hat{\sigma}$ で置き換えた

$$
t = \frac{\sum_{i=1}^{a} c_i\,\bar{y}_{i}.}{\hat{\sigma}\sqrt{\sum_{i=1}^{a} c_i^2/n_i}}
$$

は自由度 ν_e の t 分布に従う。(6.4) 式の判定方式は，この t 分布に従う統計量に対して

$$
|t| > t(\nu_e;\alpha/2)
$$

という通常の両側 t 検定を行なっていることと同値である。

演習問題 14　以下，和の記号 $\sum$ は全て $\sum_{i=1}^{a}$ を意味する。対比の係数ベクトルを

$$
\mathbb{c} = (c_1,\ldots,c_a)^T, \quad \sum c_i = 0
$$

と表わす。また，総平均 $\bar{y}..$ と，母平均に関する一般平均 μ を

$$
\begin{aligned}
\bar{y}.. &= \frac{1}{\sum n_i}\sum n_i\,\bar{y}_{i}. \\
\mu &= \frac{1}{\sum n_i}\sum n_i\,\mu_i
\end{aligned}
$$

とおく。$\sum c_i = 0$ であるから

$$\sum c_i\,(\bar{y}_{i\cdot} - \mu_i) = \sum c_i\,\{(\bar{y}_{i\cdot} - \mu_i - (\bar{y}_{\cdot\cdot} - \mu)\} \tag{$*$}$$

である。

Cauchy-Schwarz の不等式により，任意の実数 $A_1, \ldots, A_a, B_1, \ldots, B_a$ に対して

$$\left(\sum A_i\,B_i\right)^2 \le \left(\sum A_i^2\right)\left(\sum B_i^2\right), \quad \frac{\left(\sum A_i\,B_i\right)^2}{\left(\sum A_i^2\right)} \le \left(\sum B_i^2\right)$$

が成り立つ。等号が成り立つのは，定数 λ を用いて $A_i = \lambda B_i$ と表わされるときである。ここで

$$A_i = c_i/\sqrt{n_i}, \quad B_i = \sqrt{n_i}\,\{(\bar{y}_{i\cdot} - \mu_i) - (\bar{y}_{\cdot\cdot} - \mu)\},$$
$$A_i\,B_i = c_i\,\{(\bar{y}_{i\cdot} - \mu_i) - (\bar{y}_{\cdot\cdot} - \mu)\}$$

とおけば，$(*)$ を利用して

$$\frac{1}{(a-1)\,\hat{\sigma}^2}\frac{\{\sum c_i\,(\bar{y}_{i\cdot} - \mu_i)\}^2}{\sum c_i^2/n_i} \le \frac{\sum n_i\,\{(\bar{y}_{i\cdot} - \mu_i) - (\bar{y}_{\cdot\cdot} - \mu)\}^2}{(a-1)\,\hat{\sigma}^2}$$

が任意の $\mathbb{c}$ に対して成り立つ。等号は $c_i = n_i\,\{(\bar{y}_{i\cdot} - \mu_i) - (\bar{y}_{\cdot\cdot} - \mu)\}$ のときに成り立つ。

この式の右辺は自由度 $(a-1, \nu_e)$ の F 分布に従う。したがって，

$$\begin{aligned}
1 - \alpha &= \Pr\left\{\frac{1}{(a-1)\,\hat{\sigma}^2}\max_{\mathbb{c}}\frac{\left|\sum c_i\,(\bar{y}_{i\cdot} - \mu_i)\right|^2}{\sum c_i^2/n_i} \le F(a-1, \nu_e; \alpha)\right\} \\
&= \Pr\left\{\text{すべての } \mathbb{c} \text{ で } \frac{1}{(a-1)\,\hat{\sigma}^2}\frac{\left|\sum c_i\,(\bar{y}_{i\cdot} - \mu_i)\right|^2}{\sum c_i^2/n_i} \le F(a-1, \nu_e; \alpha)\right\} \\
&= \Pr\left\{\text{すべての } \mathbb{c} \text{ で } \left|\sum c_i\,(\bar{y}_{i\cdot} - \mu_i)\right| \le S_{\mathbb{c}}(\alpha)\right\} \\
&= \Pr\left\{\text{すべての } \mathbb{c} \text{ で } \left|\sum c_i\,\bar{y}_{i\cdot} - \sum c_i\,\mu_i\right| \le S_{\mathbb{c}}(\alpha)\right\}
\end{aligned}$$

が成り立つ。ここで (6.5) 式により

$$S_{\mathbb{c}}(\alpha) = \hat{\sigma}\sqrt{(a-1)\cdot\sum(c_i^2/n_i)\cdot F(a-1, \nu_e; \alpha)} \tag{6.5}$$

である。

演習問題 15　性質 (a): $m - m_0 = 0$ であるから常に $S = 0$ であり，$V > 0$ のとき $Q = V/R = 1$，$V = 0$ のとき $Q = 0$ となる。したがって，

$$\mathit{FDR} = \mathrm{E}[Q] = 1 \times \Pr\{Q = 1\} = \Pr\{V > 0\} = \mathit{FWER}$$

が成り立つ。

性質 (b): 一般に $Q = V/(V+S) \le 1$ である。したがって，

$$\mathit{FDR} = \mathrm{E}[Q] \le 1 \times \Pr\{Q > 0\} = \mathit{FWER}$$

が成り立つ。