第 4 章のつづき

浅川 伸一

2007 年 01 月 18 日

5 多重比較の実際

データ例として向後先生のハンバーガーショップを用いる
http://kogolab.jp/elearn/hamburger/chap6/sec3.html

	わくわく	もぐもぐ	ぱくぱく	
	80	75	80	
	75	70	80	
	80	80	80	
	90	85	90	
	95	90	95	
	80	75	85	
	80	85	95	
	85	80	90	
	85	80	85	
	80	75	90	
	90	80	95	
	80	75	85	
	75	70	98	
	90	85	95	
	85	80	85	
	85	75	85	
	90	80	90	
	90	80	90	
	85	90	85	
	80	80	85	
データ数	20	20	20	60
群平均	84.00	79.50	88.15	83.88
分散	29.0	29.75	28.53	41.57

分散分析表は であるから、ハンバーガーショップによる違いは有意である。

変動要因	変動	自由度	分散	F	p
要因	748.63	2	374.37	12.22	0.0001
誤差	1745.55	57	30.62		
計	2494.19	59			

このデータによれば、（一袋あたりのポテトの本数）はもぐもぐ店が最も少く、次にわくわく店である。そこで帰無仮説 H0: $\bar{x}_1 = \bar{x}_2$ を scheffe の方法によって検討してみる。$c_1 = 1,\ c_2 = -1,\ c_3 = 0$ とおくと scheffe の式、左辺は

$$\frac{|c_1\bar{x}_1 + c_2\bar{x}_2 + c_3\bar{x}_3|}{\sqrt{\left(\frac{c_1^2}{n_1} + \frac{c_2^2}{n_2} + \frac{c_3^2}{n_3}\right)\hat{\sigma}_e^2}} = \frac{|\bar{x}_1 - \bar{x}_2|}{\sqrt{\left(\frac{1^2}{20} + \frac{(-1)^2}{20} + \frac{0^2}{20}\right)30.62}} \tag{21}$$

$$= \frac{|84 - 79.5|}{1.75} \tag{22}$$

$$= 2.571 \tag{23}$$

である。一方右辺は

$$\sqrt{(k-1)F_{k-1,m(k-1)}(\alpha)} = \sqrt{(3-1)F_{2,3(20-1)}(0.05)} \tag{24}$$

$$= \sqrt{2 \times \mathrm{Finv}(0.05,2,57)} \tag{25}$$

$$= 2.51 \tag{26}$$

であるから、右辺の 5 % の有意水準の値よりも左辺の方が大きい。したがって、もぐもぐ店とわくわく店には違いがあると言える。

同様の計算をモグモグ店とわくわく店で比較するには $c_1 = 1,\ c_2 = 0,\ c_3 = -1$ として計算すれば良い。右辺の値は変らず、左辺の分母の値も変らないので左辺分子の値のみ入れ替えて $|84.0 - 88.15| = 4.15$ であるから右辺は $4.15/2.571 = 2.37$ となって 5 % の有意水準を示す右辺の値 2.51 よりも小さくなる。このことから、モグモグ店とわくわく店の間には有意な差が認められるとは言えないことになる。

6 一因子の実験計画の復習

線形モデルによる実験計画とは以下のようにまとめられた。

なるデータに対して、

条件 1	条件 2	$\cdots$	条件 k	
x_{11}	x_{12}	$\cdots$	x_{1k}	
x_{21}	x_{22}	$\cdots$	x_{2k}	
$\vdots$	$\vdots$	$\cdots$	$\vdots$	
$x_{n_1 1}$	$x_{n_2 2}$	$\cdots$	$x_{n_k k}$	
$\bar{x}_{\cdot 1}$	$\bar{x}_{\cdot 2}$	$\cdots$	$\bar{x}_{\cdot k}$	$\bar{x}_{\cdot\cdot}$

表 1: 1 要因 k 水準の実験データ

$$\begin{pmatrix} x_{11} \\ x_{21} \\ \vdots \\ x_{n_1 1} \\ x_{12} \\ x_{22} \\ \vdots \\ x_{n_2 2} \\ \vdots \\ x_{1k} \\ x_{2k} \\ \vdots \\ x_{n_k k} \end{pmatrix} = \begin{pmatrix} 1 & 1 & 0 & \cdots & 0 \\ 1 & 1 & 0 & \cdots & 0 \\ \vdots & \vdots & \cdots & \cdots & \vdots \\ 1 & 1 & 0 & & \vdots \\ 1 & 0 & 1 & & \vdots \\ 1 & 0 & 1 & & \vdots \\ \vdots & \vdots & \cdots & \cdots & \vdots \\ 1 & 0 & 1 & & \vdots \\ \vdots & \vdots & \cdots & \cdots & \vdots \\ 1 & 0 & 0 & \vdots & 1 \\ 1 & 0 & 0 & \vdots & 1 \\ \vdots & \vdots & \cdots & \cdots & \vdots \\ 1 & 0 & 0 & \cdots & 1 \end{pmatrix} \begin{pmatrix} \theta_0 \\ \theta_1 \\ \theta_2 \\ \vdots \\ \theta_k \end{pmatrix} + \begin{pmatrix} \epsilon_{11} \\ \epsilon_{21} \\ \vdots \\ \epsilon_{n_1 1} \\ \epsilon_{12} \\ \epsilon_{22} \\ \vdots \\ \epsilon_{n_2 2} \\ \vdots \\ \epsilon_{1k} \\ \epsilon_{2k} \\ \vdots \\ \epsilon_{n_k k} \end{pmatrix} \tag{27}$$

これをデータベクトル $\boldsymbol{x}$, 計画行列 $\boldsymbol{A} = [\boldsymbol{A}_0, \boldsymbol{A}_1]$, 平均ベクトル $\boldsymbol{\theta}$, および誤差ベクトル $\boldsymbol{\epsilon}$ を用いて

$$\boldsymbol{x} = \boldsymbol{A}\boldsymbol{\theta} + \boldsymbol{\epsilon} \tag{28}$$

と書くことができた。このとき計画行列のランクは $\mathrm{rann}\boldsymbol{A} = k$ なので

$$\boldsymbol{A}^T\boldsymbol{A} = \begin{pmatrix} N & n_1 & n_2 & \cdots & n_k \\ n_1 & n_1 & 0 & \cdots & 0 \\ n_2 & 0 & n_2 & & 0 \\ \vdots & \vdots & 0 & \cdots & \vdots \\ n_k & 0 & 0 & \cdots & n_k \end{pmatrix} \tag{29}$$

逆行列は存在しない。一つの解法としては $\boldsymbol{A}_0$ で張られる空間の補空間を考えてその上で直交補空間への分割を考えれば良い。

すなわち、データ行列を一つの列ベクトル $\boldsymbol{x} = (\ x_{11}\ ,x_{21}, \cdots ,x_{n_1 1}, x_{12}, x_{22}, \cdots, x_{n_2 2}, \cdots, x_{1k}, x_{2k}, \cdots, x_{n_k k}\ )^T$ で表現し、全ての要素が 1 からなる $(1, 1, \cdots, 1)^T \theta_0$ 列ベクトル ($\theta_0 = \bar{x}_{..}$) と右辺第二項 0 と 1 とでできた行列を $\boldsymbol{A}$ と表すことにする。さらに右辺第三項のベクトルを $\boldsymbol{\epsilon}$ とすると上式は

$$\boldsymbol{x} = \boldsymbol{A\theta} + \boldsymbol{\epsilon} \tag{30}$$

と表現された。

7　2 要因の場合

測定すべき実験変数が複数の場合を考える。

7.1　繰り返しのない場合

各条件で一度しか測定が行われない場合である。表 2 では列方向に条件 A

	条件 A1	条件 A2	条件 A3	
条件 B1	x_{11}	x_{12}	x_{13}	$\bar{x}_{1\cdot}$
条件 B2	x_{21}	x_{22}	x_{23}	$\bar{x}_{2\cdot}$
条件 B3	x_{31}	x_{32}	x_{33}	$\bar{x}_{3\cdot}$
条件 B4	x_{41}	x_{42}	x_{43}	$\bar{x}_{4\cdot}$
	$\bar{x}_{\cdot 1}$	$\bar{x}_{\cdot 2}$	$\bar{x}_{\cdot 3}$	$\bar{x}_{..}$

表 2: 2 元配置の分散分析に用いられるデータ

(3 水準)、行方向に 条件 B (4 水準) が表現されている。実際には各条件毎に一つしかデータが得られないことは稀である。しかし、例えば、条件 B を繰り返しの順序と考えれば実験の最初に得られた条件 B1 と最後に得られた B4 とを比較することで、練習効果、疲労効果などを考慮にいれた分析をすることになり、しばしば用いられる手法である。

この時得られたデータは次のモデルに従うと考える

$$x_{ij} = \mu + \alpha_i + \beta_j + e_{ij} \tag{31}$$

すなわち、データは全平均の推定値と条件 A の効果と条件 B の効果と誤差

とで表される。このときの線形モデルは

$$\begin{pmatrix} x_{11} \\ x_{21} \\ x_{31} \\ x_{41} \\ x_{12} \\ x_{22} \\ x_{32} \\ x_{42} \\ x_{13} \\ x_{23} \\ x_{33} \\ x_{43} \end{pmatrix} = \begin{pmatrix} 1 & 1 & & & 1 & & & \\ 1 & 1 & & & & 1 & & \\ 1 & 1 & & & & & 1 & \\ 1 & 1 & & & & & & 1 \\ 1 & & 1 & & 1 & & & \\ 1 & & 1 & & & 1 & & \\ 1 & & 1 & & & & 1 & \\ 1 & & 1 & & & & & 1 \\ 1 & & & 1 & 1 & & & \\ 1 & & & 1 & & 1 & & \\ 1 & & & 1 & & & 1 & \\ 1 & & & 1 & & & & 1 \end{pmatrix} \begin{pmatrix} \alpha_1 \\ \alpha_2 \\ \alpha_3 \\ \beta_1 \\ \beta_2 \\ \beta_3 \\ \beta_4 \end{pmatrix} + \begin{pmatrix} \epsilon_{11} \\ \epsilon_{21} \\ \epsilon_{31} \\ \epsilon_{41} \\ \epsilon_{12} \\ \epsilon_{22} \\ \epsilon_{32} \\ \epsilon_{42} \\ \epsilon_{13} \\ \epsilon_{23} \\ \epsilon_{33} \\ \epsilon_{43} \end{pmatrix} \tag{32}$$

式中、要素が書かれていない個所は 0 である。これは

$$\boldsymbol{x} = \boldsymbol{A\theta} + \boldsymbol{\epsilon} \tag{33}$$

と書くことができて、形式的には一要因（一元配置）の分散分析と全く同じ式である。

$$\boldsymbol{A}^T \boldsymbol{x} = \begin{pmatrix} \bar{x}_{\cdot\cdot} \\ \bar{x}_{\cdot 1} \\ \bar{x}_{\cdot 2} \\ \bar{x}_{\cdot 3} \\ \bar{x}_{1\cdot} \\ \bar{x}_{2\cdot} \\ \bar{x}_{3\cdot} \\ \bar{x}_{4\cdot} \end{pmatrix} \tag{34}$$

である。また、

$$\boldsymbol{A}^T \boldsymbol{A} = \begin{pmatrix} 12 & 4 & 4 & 4 & 3 & 3 & 3 & 3 \\ 4 & 4 & & & 1 & 1 & 1 & 1 \\ 4 & & 4 & & 1 & 1 & 1 & 1 \\ 4 & & & 4 & 1 & 1 & 1 & 1 \\ 3 & 1 & 1 & 1 & 3 & & & \\ 3 & 1 & 1 & 1 & & 3 & & \\ 3 & 1 & 1 & 1 & & & 3 & \\ 3 & 1 & 1 & 1 & & & & 3 \end{pmatrix} \tag{35}$$

統計的検討をする場合には帰無仮説が二つあって $H1 : \alpha_1 = \alpha_2 = \alpha_3 = 0$, $H2 : \beta_1 = \beta_2 = \beta_3 = \beta_4 = 0$ である。

式 (32), (33) の計画行列 $\boldsymbol{A}$ の中で第一列を $\boldsymbol{A}_0$, 第 2 列から 4 列までを $\boldsymbol{A}_1$, 第 5 列から第 8 列までを $\boldsymbol{A}_2$ とする。

$$\boldsymbol{A}_0 = \begin{pmatrix} 1 \\ 1 \\ 1 \\ 1 \\ 1 \\ 1 \\ 1 \\ 1 \\ 1 \\ 1 \\ 1 \\ 1 \end{pmatrix} \tag{36}$$

$$\boldsymbol{A}_1 = \begin{pmatrix} 1 & & \\ 1 & & \\ 1 & & \\ 1 & & \\ & 1 & \\ & 1 & \\ & 1 & \\ & 1 & \\ & & 1 \\ & & 1 \\ & & 1 \\ & & 1 \end{pmatrix} \tag{37}$$

$$\boldsymbol{A}_2 = \begin{pmatrix} 1 & & & \\ & 1 & & \\ & & 1 & \\ & & & 1 \\ 1 & & & \\ & 1 & & \\ & & 1 & \\ & & & 1 \\ 1 & & & \\ & 1 & & \\ & & 1 & \\ & & & 1 \end{pmatrix} \tag{38}$$

さらに $\boldsymbol{\alpha} = (\theta_1, \theta_2, \theta_3)^T$, $\boldsymbol{\beta} = (\beta_1, \beta_2, \beta_3, \beta_4)^T$, とおくとデータは

$$\boldsymbol{x} = \boldsymbol{A}_0\mu + \boldsymbol{A}_1\boldsymbol{\alpha} + \boldsymbol{A}_2\boldsymbol{\beta} + \boldsymbol{\epsilon} \tag{39}$$

と表現できる。ここで μ は全平均を表す。

一要因の分散分析で行ったように、全分散 (の総データ数倍) は

$$12S_T^2 = \sum\sum(x_{ij} - \bar{x}_{..})^2 \tag{40}$$

で表されるが、これを変形し

$$\begin{aligned}
12S_T^2 &= \sum\sum(x_{ij} - \bar{x}_{..})^2 && (41)\\
&= \sum\sum(x_{ij} - \bar{x}_{.j} - \bar{x}_{i.} + \bar{x}_{..} + \bar{x}_{.j} - \bar{x}_{..} + \bar{x}_{i.} + \bar{x}_{..})^2 && (42)\\
&= \sum\sum\{(\bar{x}_{.j} - \bar{x}_{..}) + (\bar{x}_{i.} - \bar{x}_{..}) && (43)\\
&\quad + (\bar{x}_{ij} - \bar{x}_{.j} - \bar{x}_{i.} + \bar{x}_{..})\}^2\\
&= 4\sum(\bar{x}_{.j} - \bar{x}_{..})^2 + 3\sum(\bar{x}_{i.} - \bar{x}_{..})^2 && (44)\\
&\quad + \sum\sum(x_{ij} - \bar{x}_{.j} - \bar{x}_{i.} + \bar{x}_{..})^2\\
&= S_0^2 + S_1^2 + S_2^2 + S_e^2 && (45)
\end{aligned}$$

従って上記のごとき分散の分解によって次の分散分析表を得る。

要因	平方和	自由度	平均平方	F
A	S_1	m-1	$S_1/(m-1)$	$\dfrac{S_1/(m-1)}{S_e/(m-1)(k-1)}$
B	S_2	k-1	$S_2/(k-1)$	$\dfrac{S_2/(m-1)}{S_e/(m-1)(k-1)}$
誤差	S_e	$(m-1)(s-1)$	$S_e/(m-1)(k-1)$	
計	$\boldsymbol{x}^T\boldsymbol{x}$	mk - 1		

表 3: 繰り返しのない 2 要因の分散分析表

線形モデルに即して考えれば、計画行列 $\boldsymbol{A}$ の部分行列によって構成される空間 $\boldsymbol{A}_0$, $\boldsymbol{A}_1^\perp$, $\boldsymbol{A}_2^\perp$, および $\boldsymbol{A}_e^\perp$

によって直和分解され $L(A) = A_0 \oplus A_1^\perp \oplus A_2^\perp$ と表される。$\boldsymbol{A}^\perp$ は $\boldsymbol{A}_0$ で張られる空間とは直交補空間の関係にある空間を表す。このとき各部分空間への射影行列を $\boldsymbol{Q}_0$, $\boldsymbol{Q}_1$, $\boldsymbol{Q}_2$, $\boldsymbol{Q}_e$ とすると $\boldsymbol{I} = \boldsymbol{Q}_0 + \boldsymbol{Q}_1 + \boldsymbol{Q}_2 + \boldsymbol{Q}_e$ である。

従って、これらの射影行列を用いれば、一要因の分散分析の場合と全く同じようにしてベクトルの長さ $\boldsymbol{x}^T\boldsymbol{x}$ を分解することができる。

$$\begin{aligned}
\boldsymbol{x}^T\boldsymbol{x} &= \boldsymbol{x}^T\boldsymbol{Q}_0\boldsymbol{x} + \boldsymbol{x}^T\boldsymbol{Q}_1\boldsymbol{x} + \boldsymbol{x}^T\boldsymbol{Q}_2\boldsymbol{x} + \boldsymbol{x}^T\boldsymbol{Q}_e\boldsymbol{x} && (46)\\
&= \boldsymbol{x}^T\boldsymbol{Q}_0\boldsymbol{x} + \boldsymbol{x}^T\boldsymbol{Q}_1\boldsymbol{x} + \boldsymbol{x}^T\boldsymbol{Q}_2\boldsymbol{x} + \boldsymbol{x}^T(\boldsymbol{I} - \boldsymbol{Q})\boldsymbol{x} && (47)\\
&= S_0 + S_1 + S_2 + S_e && (48)
\end{aligned}$$

7.2 繰り返しのある 2 要因の分散分析

表 2 のように、各条件あたりのデータが 1 つしかない場合を複数個のデータを持つことができるように拡張する。分析に用いられるデータは次の表 4 のようになる。

	条件 A1	条件 A2	条件 A3	
条件 B1	x_{111}	x_{121}	x_{131}	$\bar{x}_{1..}$
	x_{112}	x_{122}	x_{132}	
条件 B2	x_{211}	x_{221}	x_{231}	$\bar{x}_{2..}$
	x_{212}	x_{222}	x_{232}	
	$\bar{x}_{.1.}$	$\bar{x}_{.2.}$	$\bar{x}_{.3.}$	$\bar{x}_{...}$

表 4: 繰り返しのある 2 元配置の分散分析に用いられるデータ

対応する線形モデルは

$$x_ijk = \mu + \alpha_i + \beta_j + (\gamma)_{ij} + \epsilon_{ijk} \tag{49}$$

これを書き下すと次のようになる。

$$\begin{pmatrix} x_{111} \\ x_{112} \\ x_{121} \\ x_{122} \\ x_{131} \\ x_{132} \\ x_{211} \\ x_{212} \\ x_{221} \\ x_{222} \\ x_{231} \\ x_{232} \end{pmatrix} = \begin{pmatrix} 1 & 1 & & 1 & & & 1 & & & & & \\ 1 & 1 & & 1 & & & 1 & & & & & \\ 1 & 1 & & & 1 & & & 1 & & & & \\ 1 & 1 & & & 1 & & & 1 & & & & \\ 1 & 1 & & & & 1 & & & 1 & & & \\ 1 & 1 & & & & 1 & & & 1 & & & \\ 1 & & 1 & 1 & & & & & & 1 & & \\ 1 & & 1 & 1 & & & & & & 1 & & \\ 1 & & 1 & & 1 & & & & & & 1 & \\ 1 & & 1 & & 1 & & & & & & 1 & \\ 1 & & 1 & & & 1 & & & & & & 1 \\ 1 & & 1 & & & 1 & & & & & & 1 \end{pmatrix} \begin{pmatrix} \theta_0 \\ \alpha_1 \\ \alpha_2 \\ \alpha_3 \\ \beta_1 \\ \beta_2 \\ \gamma_{11} \\ \gamma_{12} \\ \gamma_{13} \\ \gamma_{21} \\ \gamma_{22} \\ \gamma_{23} \end{pmatrix} + \begin{pmatrix} \epsilon_{111} \\ \epsilon_{112} \\ \epsilon_{121} \\ \epsilon_{122} \\ \epsilon_{131} \\ \epsilon_{132} \\ \epsilon_{211} \\ \epsilon_{212} \\ \epsilon_{221} \\ \epsilon_{222} \\ \epsilon_{231} \\ \epsilon_{232} \end{pmatrix} \tag{50}$$

これまでと全く同様にして

$$\boldsymbol{x} = \boldsymbol{A\theta} + \boldsymbol{\epsilon} \tag{51}$$

と書くことができて、計画行列が $\boldsymbol{A} : [\boldsymbol{A}_0, \boldsymbol{A}_1, \boldsymbol{A}_2, \boldsymbol{A}_3]$ となっただけである。$\boldsymbol{\theta} = (\mu, \boldsymbol{\alpha}, \boldsymbol{\beta}, \boldsymbol{\gamma})^T$ であり、$\boldsymbol{\alpha} = (\alpha_1, \alpha_2, \alpha_3)^T$, $\boldsymbol{\beta} = (\beta_1, \beta_2)^T$, $\boldsymbol{\gamma} = (\gamma_1 1, \gamma_1 2, \gamma_1 3, \gamma_2 1, \gamma_2 2, \gamma_2 3)^T$ である。

空間の直和分解については

$$R = L(A_0) \oplus L^{\perp}(A_1) \oplus L^{\perp}(A_2) \oplus L^{\perp}(A_3) \oplus L^{\perp}(A_e) \tag{52}$$

帰無仮説は 3 つあって、それぞれ対応する空間へと射影したベクトルの長さが誤差ベクトルと比べて十分に長いと言えるか否かを検討することになる。

分散分析表は表 5 のようになる。

要因	平方和	自由度	平均平方	F
A	S_1	m-1	$S_1/(m-1)$	$\frac{S_1/(m-1)}{S_e/(m-1)(k-1)}$
B	S_2	k-1	$S_2/(k-1)$	$\frac{S_2/(m-1)}{S_e/(m-1)(k-1)}$
A×B	(m-1)(k-1)	S_3	$S_3/(m-1)(k-1)$	$\frac{S_3/(m-1)(k-1)}{S_e/mk(t-1)}$
誤差	S_e	$mk(t-1)$	$S_e/mk(t-1)$	
計	$\boldsymbol{x}^T\boldsymbol{x}$	mk - 1		

表 5: 繰り返しのある 2 要因の分散分析表

8 交互作用

2 要因以上の実験計画で繰り返しが存在する場合には、各条件毎に分散を求めることができる。このとき、各個の条件毎に差が存在する場合がある。

むしろ、積極的に交互作用の効果を検討するような実験が立案されることも多い。

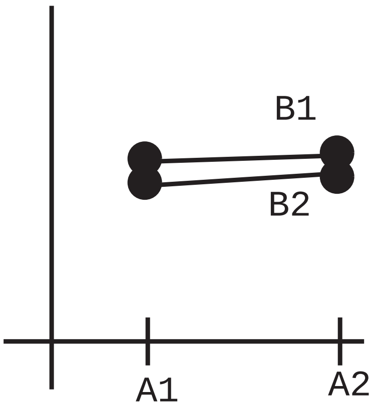

図 2: いずれの効果も認められない例

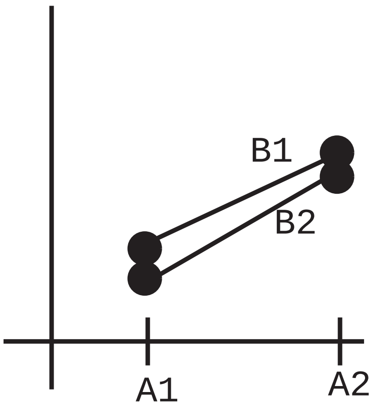

図 3: 要因 A の効果が認められる例

図 2 は横軸に条件 (要因) A(2 水準) をとり、条件 (要因) B (同じく 2 水準) を異なる線として表現したものである。この図 2 の場合は条件 A にも条件 B にも差異が認められず、すべてのデータが全平均の近くに集まっている。

一方、図 3 では要因 B の効果が認められず (B1, B2 間で差がない) 要因 A の効果のみが有意であるような場合である。

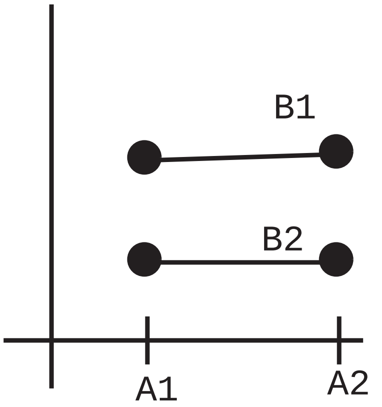

図 4: 要因 B の効果が認められる例

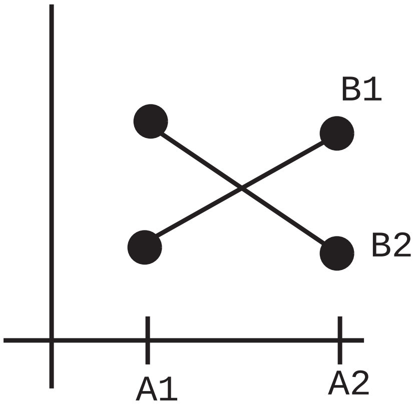

図 5: A と B の交互作用効果が認められる例

図 4 は反対に要因 B の効果は有意であるが、要因 A では A1 と A2 との間に差が認められない場合である。

図 5 は交互作用が有意となった場合である。条件毎に要因 A の効果と要因 B の効果の現れ方が異なる。

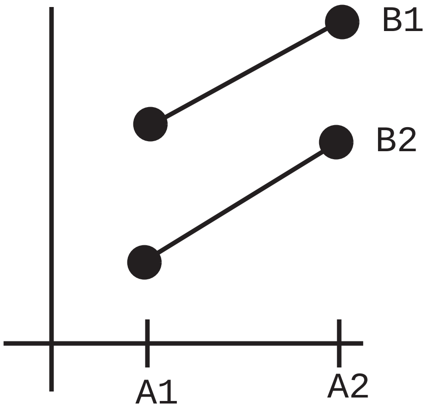

図 6: 2 つの要因の主効果が共に有意であると考えられる例

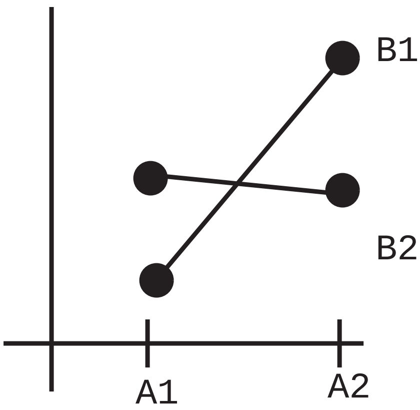

図 7: 要因 A の主効果と交互作用が認められる例

図 6 は A1 と A2 との間にも、B1 と B2 との間にも共に有意な差が認められる場合の例である。

図 7 は A1 と A2 との間に有意な差が認められ、かつ交互作用も存在する

が要因 B の効果は認められない例である。あるいは A1 と A2 との差が有意となったのは、B2 条件のみであったということもできる。

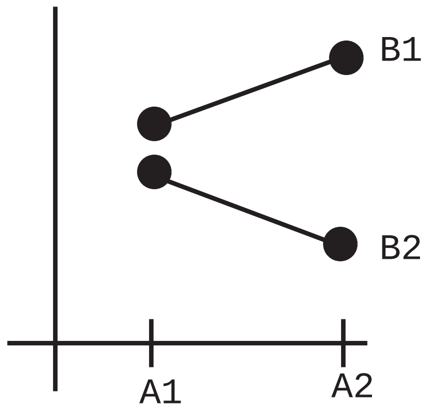

図 8: B の主効果と交互作用が認められる例

図 8 は B1 と B2 との間に有意な差が認められ、かつ交互作用も存在するが要因 A の効果は認められない例である。これは B1 と B2 との差が認められたのは A2 条件においてのみであったと解釈することもできる。

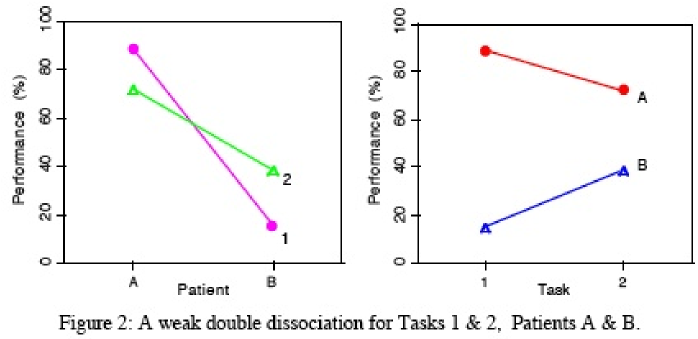

図 9: Bullinaria(1999), Figure 2

図 9 左は横軸に二人の意味記憶に障害を持つ患者の語彙判断課題の成績を示している (Bullinaria,1999)。患者 A は課題 1, 課題 2 の成績とも良好であり、患者 B は相対的に成績が落ちている。

同じデータを書き直したものが図 9 右である。横軸を課題とし患者の違いを異なる線として表現している。

このように同じデータであっても作図の仕方によって印象が異なる。

9　さらに複雑な実験計画

3 要因以上の実験計画であっても基本的な考え方は変らない。どれか一つの要因について繰り返しがなければ、繰り返しのない 2 要因の分散分析と同じく交互作用の検討をすることができない。最も一般的な形で 3 要因の実験計画を表現すると

$$x_{ijkl} = \mu + \alpha_i + \beta_j + \gamma_k + (\alpha\beta)_{ij} + (\beta\gamma)_{jk} + (\gamma\alpha)_{ki} + (\alpha\beta\gamma)_{ijk} + \epsilon_{ijkl} \quad (53)$$

となる。ここで $(\alpha\beta)_{ij}, (\beta\gamma)_{jk}, (\gamma\alpha)_{ki}$ は 2 次の交互作用を表し、$(\alpha\beta\gamma)_{ijk}$ は 3 次の交合作用を表す。データの性質や実験の目的によってはいくつかの交互作用を無視して考えることも行われる。

エクセルの統計ツールで用意されているのは 2 元配置 (2 要因) の分散分析までであるが、どんな複雑な実験計画でも計画行列を作り、射影行列による直和分解を用いれば原理的には計算可能である。

つまり、すべての実験計画の基本的なアイデアはすべて同じであり、計画行列を如何に作るかによってデータの解釈が異なってくる。計画行列を元にデータを直和分解して、その部分空間内での線分の長さに対応する量を考えれば良い。

また、一般に検討すべき要因が増えれば増えるほど手続きは煩雑になり、結果の解釈も難しくなる。高次交互作用の解釈も難しい。

そこで、主効果 (2 次の交互作用も加えることがある) のみを検討する方法が開発されている。

9.1　一部実施と直交計画表

例えば A, B, C, D という 4 つの要因があり、すべて 2 水準であるとする。交互作用については AB についてだけ考慮したいが、それ以外の交互作用については理論的に無視できるものと考えよう。すべての実験条件について $2^4 = 16$ 条件あるが、この一部だけを実施するものとし、以下の表 9.1 のうち x_i の条件だけ実施する。

A の主効果を α_1, α_2, B, C, D の主効果をそれぞれ $\beta, \gamma, delta$ で表すものとする。A と B との交互作用については τ とする。

	A	B	C	D	
1	1	1	1	1	x_1
2	1	1	1	2	
3	1	1	2	1	
4	1	1	2	2	x_2
5	1	2	1	1	
6	1	2	1	2	x_3
7	1	2	2	1	x_4
8	1	2	2	2	
9	2	1	1	1	
10	2	1	1	2	x_5
11	2	1	2	1	x_6
12	2	1	2	2	
13	2	2	1	1	x_7
14	2	2	1	2	
15	2	2	2	1	x_8
16	2	2	2	2	

線形モデルを用いて

$$\begin{pmatrix} x_1 \\ x_2 \\ x_3 \\ x_4 \\ x_5 \\ x_6 \\ x_7 \\ x_8 \end{pmatrix} = \begin{pmatrix} 1 & 1 & 1 & 1 & 1 & 1 \\ 1 & 1 & 1 & -1 & -1 & 1 \\ 1 & 1 & -1 & 1 & -1 & -1 \\ 1 & 1 & -1 & -1 & 1 & -1 \\ 1 & -1 & 1 & 1 & -1 & -1 \\ 1 & -1 & 1 & -1 & 1 & -1 \\ 1 & -1 & -1 & 1 & 1 & 1 \\ 1 & -1 & -1 & -1 & -1 & 1 \end{pmatrix} \begin{pmatrix} \mu \\ \alpha \\ \beta \\ \gamma \\ \delta \\ \tau \end{pmatrix} + \begin{pmatrix} \epsilon_1 \\ \epsilon_2 \\ \epsilon_3 \\ \epsilon_4 \\ \epsilon_5 \\ \epsilon_6 \\ \epsilon_7 \\ \epsilon_8 \end{pmatrix} \tag{54}$$

計画行列のうち 2 列目が A, 3 列目が B, 4 列目が C, 5 列目が D, 6 列目が A と B との交互作用を表している。これを

$$\boldsymbol{x} = \boldsymbol{A}\boldsymbol{\theta} + \boldsymbol{\epsilon} \tag{55}$$

として解くことを考える。

計画行列 $\boldsymbol{A}$ の列ベクトルは直交していることは

$$\boldsymbol{A}^T\boldsymbol{A} = \begin{pmatrix} 8 & 0 & 0 & 0 & 0 & 0 \\ 0 & 8 & 0 & 0 & 0 & 0 \\ 0 & 0 & 8 & 0 & 0 & 0 \\ 0 & 0 & 0 & 8 & 0 & 0 \\ 0 & 0 & 0 & 0 & 8 & 0 \\ 0 & 0 & 0 & 0 & 0 & 8 \end{pmatrix} \tag{56}$$

となることから確認できる。$\boldsymbol{A}^T\boldsymbol{x}$ を計算すると

$$\begin{pmatrix} x_1+x_2+x_3+x_4+x_5+x_6+x_7+x_8 \\ x_1+x_2+x_3+x_4-x_5-x_6-x_7-x_8 \\ x_1+x_2-x_3-x_4+x_5+x_6-x_7-x_8 \\ x_1-x_2+x_3-x_4+x_5-x_6+x_7-x_8 \\ x_1-x_2-x_3+x_4-x_5+x_6+x_7-x_8 \\ x_1+x_2-x_3-x_4-x_5-x_6+x_7+x_8 \end{pmatrix} \tag{57}$$

各要素がそれぞれ全平均、A の効果、B の効果、C の効果、D の効果、A と B との交互作用の効果に成っている。$(\boldsymbol{A}^T\boldsymbol{A})^{-1}$ は対角要素が 1/8 で非対角要素が全て 0 である（直交しているので）単純な行列になっているので計算が極端に簡単となる。

$$\begin{aligned} \boldsymbol{x}^T\boldsymbol{x} &= \boldsymbol{x}^T\boldsymbol{A}^T(\boldsymbol{A}^T\boldsymbol{A})^{-1}\boldsymbol{x} && (58) \\ &= \boldsymbol{x}^T\boldsymbol{A}_0\boldsymbol{x} + \boldsymbol{x}^T\boldsymbol{A}_1\boldsymbol{x} + \boldsymbol{x}^T\boldsymbol{A}_2\boldsymbol{x} + \boldsymbol{x}^T\boldsymbol{A}_3\boldsymbol{x} + \boldsymbol{x}^T\boldsymbol{A}_4\boldsymbol{x} + \boldsymbol{x}^T\boldsymbol{A}_5\boldsymbol{x} && (59) \\ &= S_0^2 + S_\alpha + S_\beta^2 + S_\gamma^2 + S_\delta^2 + S_\tau^2 + S_e^2 && (60) \end{aligned}$$

従って次の分散分析表 6 を得る。

要因	平方和	自由度	
A	S_α	1	S_α/S_e
B	S_β	1	S_β/S_e
C	S_γ	1	S_γ/S_e
D	S_δ	1	S_δ/S_e
AB	S_τ	1	S_τ/S_e
誤差	S_e	1	

表 6: 分散分析表

この例からわかるとおり計画行列が直交している場合には多数の因子を少い測定で得ることができる。この直交表にそれぞれの因子（主効果）および

交互作用を割り当てればよい。例えば 3 要因の実験計画で全ての交互作用を考慮しなければならないとすれば主効果 3 つ（ A, B, C)、2 次の交互作用 3 つ (AB, BC, CA) 3 次の交互作用 1 つ (ABC) を含む実験に対して次の表**??**を直交計画表という。直交計画表の全ての列は互いに直交することが本質的

データ	μ	A	B	AB	C	AC	BC	ABC
x_1	1	1	1	1	1	1	1	1
x_2	1	1	1	1	-1	-1	-1	-1
x_3	1	1	-1	-1	1	1	-1	-1
x_4	1	1	-1	-1	-1	-1	1	1
x_5	1	-1	1	-1	1	-1	1	-1
x_6	1	-1	1	-1	-1	1	-1	1
x_7	1	-1	-1	1	1	-1	-1	1
x_8	1	-1	-1	1	-1	1	1	-1
(54)A の列	1	2	3	6	4			5
上の例での要因	μ	α	β	τ	δ	e	e	γ

表 7: 直交計画表 $L_u(2^7)$

な点である。交互作用 AB の列の各要素は A と B との列の対応する要素同士の積になっている。

直交行列の 1 列を水準 A と考えることを直交行列に因子（あるいは水準）A を割り付けると言う。直交行列には $L_{16}(2^{15})$ などいくつか考案されている。考え方は全て同じである。