日本大学医学部総合医学研究所紀要
Vol.4（2016）pp.61-65

医学研究支援部門報告

実験計画法の概要

間﨑武郎[1)]，渡部和浩[2)]，嶋村政男[2)]，石井敬基[1)]

Overview of experimental design

Takero MAZAKI[1)], Kazuhiro WATANABE[2)], Masao SHIMAMURA[2)], Yukimoto ISHII[1)]

1. はじめに

実験（研究）において実験計画ほど無視されるものはない。統計学的手法を用いてP値を計算する方法を学ぶことに熱心であっても，実験を適切に計画する方法を学ぶことはほとんどない。入念な実験計画は研究の質を担保するうえで必須である。実験計画法はイギリスの統計学者R. A. Fisher（1890-1962）が1920年代に農業実験に導入した方法であり，自然科学分野だけでなくすべての研究に共通する技術の総称である。今回この実験計画法を概説する。

2. Fisherの3原則

実験の場を適切に管理し，効率的に情報を得るための指針が実験計画法である。実験計画は以下に述べるFisherの3原則に基づいて構築される（表1）。

① 繰り返し（repetition, replication）

実験を繰り返すことで処理間の比較を十分に正確にすることを保証する；すなわち，1回の測定の場合それが水準の違いによるものか，誤差によるものか判別できない。2回以上繰り返すことで，より多くの情報を得ることができて実験で生じる誤差の大きさを評価することができる。

② 無作為化（randomization）

誤差には偶然誤差と系統誤差がある。系統誤差は，たとえば農業では地域・圃場環境，医療では年齢・性別・治療法のアンバランスなどのように，一方的な偏りを生じさせるような誤差である。このような系統誤差は，処理を実験単位に無作為に割り当てることによって偶然誤差に転化させることができる。

③ 局所管理（local control）

処理や反復が多い場合大きな系統誤差が生じる可能性がある。このとき実験全体を複数のブロックに分割し，系統誤差を取り除く方法を局所管理あるいはブロッキング（blocking）という。ブロック内では実験場を均一，無作為化する必要がある。

表1　Fisherの3原則

繰り返し (repetition, replication)
実験を繰り返すことで処理間の比較を十分に正確にすることを保証する
無作為化 (randomization)
処理を実験単位に無作為に割り当てること
局所管理 (local control)
実験全体を複数のブロックに分割し系統誤差を取り除く方法

1）日本大学医学部医学研究企画・推進室
2）日本大学医学部総合医学研究所 医学研究支援部門化学分析室
間﨑武郎：mazaki.takero@nihon-u.ac.jp

3. 実験の処理

Fisherの3原則は実験の配置に関する指針と考えることができる。一方，実験の処理の選定も必要であり，以下に処理について概説する。

① 因子，要因と水準

因子は，実験系で調査研究したい形質に影響を与えると想定している原因のことである。因子の効果を要因という。要因には，各因子の水準が変わることで生じる主効果と複数の因子の組み合わせで生じる交互作用がある。水準とは因子を質的，量的に変える場合の各段階のことである。例えば，ある薬剤の濃度がマウスの活動に与える影響に興味がある場合，薬剤が因子になりその濃度が水準になる。

② 交互作用（interaction）

2因子を考えたとき，一方の因子の効果が他方の因子の水準ごとに異なるとき，これら2つの因子間に交互作用が存在するという（図1）。交互作用が存在する場合，結果の解釈には注意が必要である。

③ 因子数による実験の分類

取り上げた因子の水準のすべての組み合わせを実施するものを完全実施要因計画（full factorial designs）といい，一部の組み合わせしか実施しないものを一部実施要因計画（fractional factorial design）という。1つの因子を研究する実験を一元配置実験とよぶ（表2a）。この場合交互作用は評価できない。取り上げる因子数により二元配置，三元配置などとよび，2因子以上を総称して多元配置とよぶ（表2b）。多元配置実験では複数の因子を同時に調べられるので，一元配置実験に比較して観測数を節約できる利点がある。多元配置において，完全実施要因計画ではなく一部実施要因計画を行う場合は，各水準の組み合わせを決定する際に直交表を利用するので直交配列表実験という。一元配置あるいは多元配置実験にしろ，因子の効果は平均値で表される；すなわち，因子の効果の差は平均値の差で評価される。

4. 乱塊法を用いた基本的な考え方

完全無作為化法（completely randomized design）とは実験単位で処理を完全にランダムに配置する方法をいう。Fisherの3原則のうち，無作為化と繰り返しの2項目を満たす。この実験は完全無作為化一

表 2

a

$A1$	$A2$	$A3$
A_{11}	A_{21}	A_{31}
A_{12}	A_{22}	A_{32}
A_{13}	A_{23}	A_{33}
A_{14}	A_{24}	A_{34}

b

	$A1$	$A2$	$A3$
$B1$	AB_{11}	AB_{21}	AB_{31}
$B2$	AB_{12}	AB_{22}	AB_{32}
$B3$	AB_{13}	AB_{23}	AB_{33}
$B4$	AB_{14}	AB_{24}	AB_{34}

a. 一元配置実験。因子Aを3水準に設定し各水準で3回の繰り返し実験を行うとすると計12回の実験を要する。

b. 繰り返しのない二元配置実験。因子Aを3水準，因子Bを4水準に設定し各組み合わせで1回ずつの実験を行うとすると計12回の実験を要する。なお，aでは因子$A1$の1番目のデータをA_{11}，bでは因子$A1$の1番目と因子$B1$の1番目のデータをAB_{11}と記している。

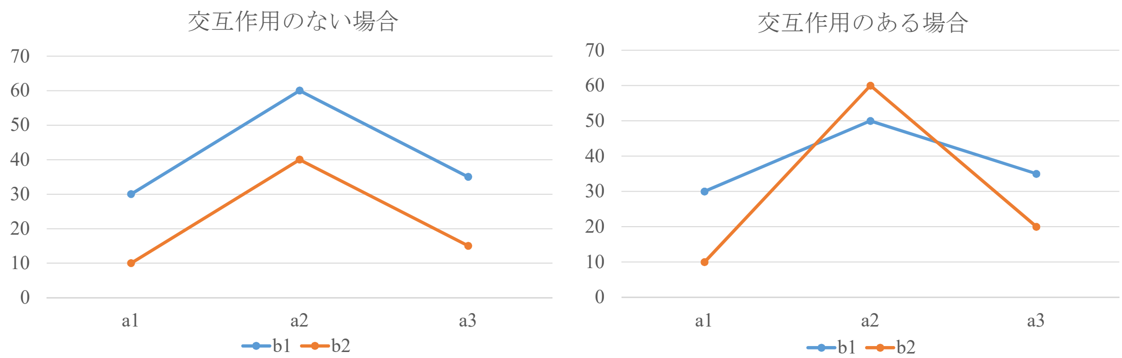

図 1 因子A（3水準）と因子B（2水準）の二元配置実験における交互作用の有無。交互作用がない場合は，各水準の効果は他の因子の水準間で同じであるため各折れ線は平行である。一方，交互作用がない場合，各直線は交差する。

元配置であり，統計学的解析方法は一元配置分散分析（one-way analysis of variance, one-way ANOVA）を用いる。一方，乱塊法（randomized block design）とは実験条件が同じブロックを作り，ブロック内で比較したい水準一揃いの実験をランダムに行うことをいい，Fisherの3原則をすべて満たす。たとえば，因子Aの3個の水準を3日間で3回行う実験を考える。しかし1日3水準の実験しか行えないとする。完全無作為化法では1日目にA_3が2回行われている（表3）。もし，“日”の違いが実験結果に影響を及ぼす場合，その効果が因子Aによって及ぼされたかのような誤解を生ずる。一方，乱塊法では“日”の違いも因子として取り上げ，1日の中で各水準をそれぞれ1回ずつ無作為に行う。これを3日間にわたり繰り返す。この場合，“日”がブロック因子であり統計学的解析方法は二元配置分散分析（two-way ANOVA）となる（表3）。

5. 直交表実験

上の例では最も簡単な二元配置要因計画を考えたが，実際はさらに因子数と水準数の多い実験を計画する場合もある。ここで，2水準を因子A, B, C, Dと4つある場合を考える。水準すべての組み合わせは$2^4 = 16$通りである。すべての組み合わせを1回以上行う完全実施要因計画を行うことが望ましいが，因子数や水準数の多い場合は不可能である。このような場合は一部実施要因計画を実施し，その際に直交表を利用してどの水準の組み合わせで実験するか決定する。

直交表には2水準系，3水準系，多水準系などがある。最も簡単な2水準系の直交表である$L_8(2^7)$を例にとって考える（図2）。$L_8(2^7)$直交表の一番上の列番1, 2 ⋯ , 7は単なる列番号である；一方，各列が1つの実験に対応しており，左の列の型番1, 2 ⋯ , 8は単なる実験番号を表している（表4）。実験番号順に実験を行う必要はなく，無作為に順番を決める。いま，この直交表の第2列と第5列を選ぶ（表5a）。第2列の1と2を因子Aの水準に，第5列の1と2を因子Bの水準として1回目から8回目までの実験を表の組み合わせで行う。これを水準の組み合わせに整理すると表5bのようになり，これは二元配置実験である（図3）。さらに，第6列の1と2を因子Cの水準として加えると実験の組み合わせは表6aのようになり，水準ごとに整理すれば表の組み合わせになり，これは3元配置実験である（表6b）。このように8回の実験で効率的に実験が可能である。ここで注意すべきことは，$A \times B$の交互作用は上の直交表では第7列に表れるので，因子Cを第6列では

表3　完全無作為化法と乱塊法実験

	1日目	2日目	3日目
完全無作為化法	$A_3\ A_3\ A_1$	$A_1\ A_3\ A_2$	$A_1\ A_2\ A_2$
乱塊法	$A_1\ A_2\ A_3$	$A_1\ A_3\ A_2$	$A_3\ A_1\ A_2$

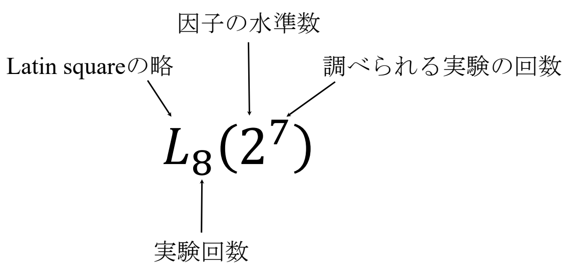

図2　直交表$L_8(2^7)$の記号の意味

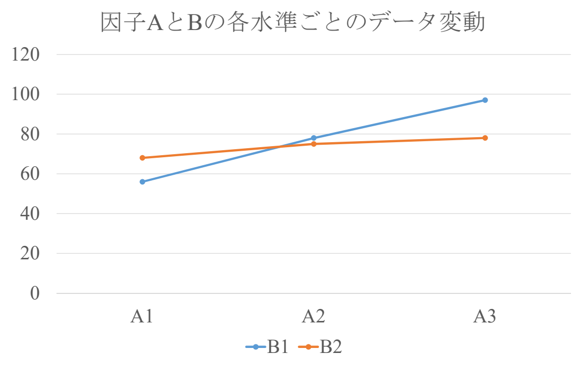

図 3　水準B_1の場合，因子Aの効果は増加傾向にある。一方，水準B_2の場合，因子Aの効果はなだらかな増加である。両直線は交差しており，交互作用が認められる。

表4　$L_8(2^7)$ 直交表

		列番						
		1	2	3	4	5	6	7
型番	1	1	1	1	1	1	1	1
	2	1	1	1	2	2	2	2
	3	1	2	2	1	1	2	2
	4	1	2	2	2	2	1	1
	5	2	1	2	1	2	1	2
	6	2	1	2	2	1	2	1
	7	2	2	1	1	2	2	1
	8	2	2	1	2	1	1	2

表5　因子AとBの水準の組み合わせ

a

実験番号	2	5
1回目の実験	A_1	B_1
2回目の実験	A_1	B_2
3回目の実験	A_2	B_1
4回目の実験	A_2	B_2
5回目の実験	A_1	B_2
6回目の実験	A_1	B_1
7回目の実験	A_2	B_2
8回目の実験	A_2	B_1

b

	B_1	B_2
A_1	1回目 6回目	2回目 5回目
A_2	3回目 8回目	4回目 7回目

表6　因子A，BとCの水準の組み合わせ

a

実験番号	2	5	6
1回目の実験	A_1	B_1	C_1
2回目の実験	A_1	B_2	C_2
3回目の実験	A_2	B_1	C_2
4回目の実験	A_2	B_2	C_1
5回目の実験	A_1	B_2	C_1
6回目の実験	A_1	B_1	C_2
7回目の実験	A_2	B_2	C_2
8回目の実験	A_2	B_1	C_1

b

		B_1	B_2
A_1	C_1 C_2	1回目 6回目	2回目 5回目
A_2	C_1 C_2	3回目 8回目	4回目 7回目

なく第7列に割り付けてしまうと因子Cと交互作用$A \times B$は交絡(confounding)し区別がつかなくなる。

6. ANOVA

One-way ANOVAであれtwo-way ANOVAであれ，分散分析の原理はデータのばらつき（総平方和）を要因によるばらつき（水準間平方和）と誤差によるばらつき（誤差平方和）に分解し，水準間平方和が誤差平方和に対して有意かどうかをF検定で判断する。Two-way ANOVAで，因子Aをl水準(i=1, ⋯ , l)，因子Bをm水準(j=1, ⋯ , m)，繰り返し回数をr回(k=1, ⋯ , r)，観測値x_{ijk}をとすると，総平方和

$$\sum_{i=1}^{l}\sum_{j=1}^{m}\sum_{k=1}^{r}\left(x_{ijk} - \bar{x}_{\dots}\right)^2$$

$$= mr\sum_{i=1}^{l}(\bar{x}_{i\cdot\cdot} - \bar{x}_{\dots})^2 + lr\sum_{j=1}^{m}\left(\bar{x}_{\cdot j\cdot} - \bar{x}_{\dots}\right)^2$$

$$+ r\sum_{i=1}^{l}\sum_{j=1}^{m}\left(\bar{x}_{ij\cdot} - \bar{x}_{i\cdot\cdot} - \bar{x}_{\cdot j\cdot} + \bar{x}_{\dots}\right)^2$$

$$+ \sum_{i=1}^{l}\sum_{j=1}^{m}\sum_{k=1}^{r}\left(x_{ijk} - \bar{x}_{ij\cdot}\right)^2$$

と，分解できる。なお，右辺第一項は因子Aの主効果，第二項は因子Bの主効果，第三項は因子AとBの交互作用項，第四項は誤差平方和である。

2つの因子A（3水準），B（2水準）と3ブロックからなる乱塊法を考える。観測値x_{ijk}とするとこのモデルは次の式

$$x_{ijk} = \mu + A_{ik} + B_{jk} + (AB)_{ijk} + \varepsilon_{ijk}$$

で表される。なお，uは全平均，A_i(i=1,2,3)は因子Aの主効果，B_j(j=1, 2)は因子Bの主効果，$(AB)_{ijk}$（ブロックk=1, 2, 3）は交互作用，ε_{ijk}は誤差である。誤差は平均0，分散σ^2の正規分布に従うと仮定する。ここで重要なことは，ANOVAでは誤差の(1)正規性，(2)等分散性，(3)独立性，(4)線形性を仮定している点である。ANOVAは非正規性にも頑強であるが，解析前にそれらを検討することは重要である。注意すべき点は，交互作用が存在するときは主効果についてコメントはできないということである。

7. まとめ

簡単に実験計画の概要を述べた。実験計画はFisherの3原則に則って立案すべきである。完全実施要因計画には完全無作為化法と乱塊法がある。一部実施要因計画を実施するときは直交表を用いて効率的に実験を計画する。いずれの方法も統計学的解析はANOVAで行う。なお，ANOVAで仮定している4つの条件をpost-hoc analysisで検定することや，正規性が認められない場合に行うFriedman検定（Friedman rank test）は割愛したので，成書を参照されたい。

文　献

1）丹後敏郎ほか：医学統計学の辞典．朝倉書店，2012
2）中村義作：よくわかる実験計画法．近代科学社，2013
3）日本統計学会編：統計学．東京図書，2013
4）森田浩：よくわかる最新実験計画法の基本と仕組み．秀和システム，2014