新統計学シリーズ=2

実験計画法

東京大学教授 理学博士　山陽国策パルプ
奥野忠一　芳賀敏郎　共著

培風館

新統計学シリーズ＝2

実験計画法

東京大学教授 理学博士　山陽国策パルプ
奥野忠一　芳賀敏郎
共著

培風館

著者略歴

奥野忠一（おくの ただかず）

1944年　東京大学理学部数学科卒業
現　在　東京大学工学部教授（理博）

主要著書

標本調査法入門（共著，小石川書房）
スネデカー：統計的方法（共訳，岩波書店）
バートレット：確率過程入門
（共訳，東大出版会）
多変量解析（共著，日科技連出版）
続多変量解析（共著，日科技連出版）
実験計画法—その発展と最近の話題
（共著，東大出版会）

芳賀敏郎（はが としろう）

1951年　東京大学理学部化学科卒業
1965年　千葉大学工学部非常勤講師
現　在　山陽国策パルプ株式会社勤務

主要著書

官能検査ハンドブック（共編，JUSE出版）
統計計算ポケットブック
（共著，日本規格協会）
統計実務講座
（文部省認定通信教育）（共著，実務教育研究所）
多変量解析（共著，日科技連出版）
続多変量解析（共著，日科技連出版）

昭和44年9月5日　初　版　発　行
昭和53年10月30日　初版第12刷発行

新統計学シリーズ 2
実　験　計　画　法

著　者　奥野忠一
　　　　芳賀敏郎
発行者　山本健二

発行所　株式会社　培　風　館

東京都千代田区九段南4-3-12・郵便番号102
電話(03)262-5256(代表)・振替東京4-44725

定価　¥1600.

日東紙工組版・一葉印刷・三水舎製本

3333-0852-6955

はしがき

本書は，数理統計学の一応用分野である実験計画法を，現実の工場実験や農業試験への適用を常に念頭において，その考え方と手法を中心にまとめたものである．結果的には，かなりの部分を実験から得られたデータの解析法に割いたのであるが，著者らの意図は，統計解析法よりは実験の計画の仕方に重点をおくことにあった．そのため，次の諸点で類書とは異なる取扱い方をした．

（1） 一元配置法・二元配置法というような，データ解析の手順を示すには便利であるが，実験の割付け方を示すにはあいまいな用語を避けた（それとの関連は第3章・第4章に述べてある）．

（2） いわゆる変量模型は，数学的にもあいまいで，実際の解釈も一義的には決まらない場合が多いにもかかわらず，その取扱い方の数式的展開には多くの紙数を浪費するので，**12・2** に，その常識的な説明を与えるにとどめた．

（3） 直交表の使い方として，わが国では，必要な要因効果（主効果と一部の2因子交互作用）の割付けを容易にするための「線点図」が多く用いられているが，本書では，「どの2因子交互作用も主効果とは別名にしない」という条件の下での割付け表および線点図を工夫した（第8～10章および付表[12]）．

（4） 直交表を用いて多くの要因効果をいちどに評価するような実験のデータ解析は，コンピュータによるべきであるとの前提に立って，そのアウトプット形式を例示した（**12・5**，**12・6**）．

著者ら2人は，それぞれ主として農業および工業の場で働いているが，相協力して，実験計画法の適用とその理論の開発に努めてきた．本書は，そのような活動の1つとして行なった，日本規格協会主催の「実験計画法入門講習会」のテキストを骨組としている．そのテキストは，主として工場技術者を対象にして執筆しているが，このシリーズの成書として出版するに当たっては，これを全面的に書きかえ，かつ演習問題などにも工夫を加えて，数理統計学を学ぶ大学の学生諸君にも使用してもらえるように配慮した．

この本の草稿を通読し，種々の有益な助言を与えて下さった，慶応義塾大学工学部助教授　鷲尾泰俊氏に，ここでお礼を申し上げたい．

本書の企画をされてから脱稿まで実に4年を費やした．これはまさに著者らの怠慢と遅筆のためで，その間忍耐強く待ち，また執筆の便を計って下さった

培風館の森平勇三・小関　清氏には，ここにお詫びと感謝の意を表わしたい．また，最後の段階で校正をすべて担当され，面倒な注文にも快く応じて下さった培風館の牧野末喜氏に謝意を表したい．

1969 年 6 月

奥野　忠一
芳賀　敏郎

目　　次

為的に設定し，それらが特性値に与える**効果の比較**を目的とするものである．

（2） その実験結果を示す特性値のとる値（**データ**）は，同じ条件下で繰り返しても一定にはならず，若干の**バラツキ**を伴う．

この2つの性質を備えたあらゆる種類の実験において，「一定の実験費用の下で，得られるべき**情報量**を最大にするための手法」を広い意味で実験計画法と呼ぶことができる．これは，「一定の情報量を獲得するのに，最小の実験費用で行なう手法」といいかえてもよい．ここで，「情報量」という言葉があいまいであるが，これは，その情報の信頼度を表わす場合と，その普遍性を示す場合とがあって，これについては第12章で述べる．いずれにせよ，「情報の質」については，ここでは評価していない．なぜなら，それは，その対象領域における実質科学や固有技術の知識に，より多く依存するものだからである．

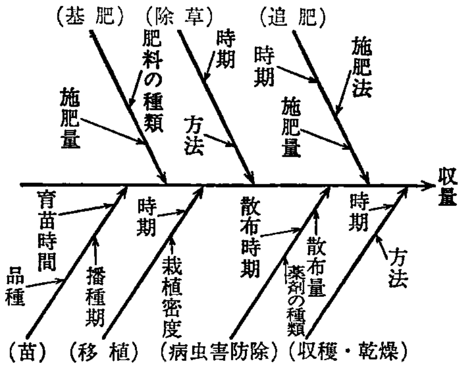

図 1·2　特性要因図（農業試験の例）

上に規定したような実験計画法は，すでに1920年代に，R. A. Fisher（フィッシャー；1890-1962, 英国ロザムステッド農業試験場技師，数学者）によって，その基礎が築かれ，はじめは農業試験に主として用いられてきたが，第2次世界大戦後は工場実験にも広く取り入れられ，最近では官能検査(味見試験)や市場調査研究などの分野でも用いられるようになった．このように適用対象が変わると，それに応じて，用いられる実験計画も少しずつ異なってくる．しかしながら，実験計画法の基本的な考え方とその手順は，共通のものとして論じることができると著者らは考えている．本書では，工業・農業への適用例を中心にして論じてゆく．

実験計画法は，その字義のとおり，実験の計画の仕方に関する手法であるが，そのような計画に従って得られたデータの解析法についても論じなければならないことはいうまでもない．本書では，もちろんこの両者を互いに関連させながら取り扱ってゆく．また，実験の計画の仕方には，次の2つの側面がある：

（a） 処理の選択——因子と水準の選び方

（b）　誤差の制御——実験配置の仕方

この（a）を 1·2 で,（b）を 1·4 で論じる.

1·2　因子と水準の選び方

いかなる実験においても，その実験の成否は，比較のために取り上げた処理条件が適切であったか否かによってほとんど決まる．それゆえ，実験処理の選択は，実験計画において第一義的に重大な課題であって，この課題に立ち向かうためには，実験者は過去の技術情報を可能な限り収集して，評価・検討し，かつ，新しい実験に期待する情報の内容を明確に規定しなければならない．このとき，「情報獲得の効率化」という観点から，処理の選択法に関して，一般的な示唆を与えることができる．それが，本節の主題である，「因子と水準の選び方」である．

因子と水準

まず，因子（factor），水準（level）という言葉の定義を与えよう：

因子——結果の特性値に影響を及ぼすと考えられる種々の原因系のうち，その実験で取り上げて比較されるもの．

水準——因子のとる種々の条件．

因子と水準の例

＜因　子＞	＜水　準＞	
温　度	100°，120°，140°	………… 3水準
触媒量	4%，5%，6%，7%	……… 4水準
原料の種類	甲，乙	……………………… 2水準

たとえば，上の例では，「温度因子」は 100°，120°，140° の3水準に，「触媒量因子」は 4%，5%，6%，7% の4水準に変えられている。また，因子の水準が量的に変えられるか，質的に変えられるかによって，「量的因子」，「質的因子」と区別することがある．「温度」や「触媒量」は前者に属し，「原料の種類」のような因子は，その水準が特定の質（品種・銘柄・産地・購入元など）によって決められるから，後者に属する．

1つの実験で取り上げる因子の数によって，その実験を，「1因子実験」，「2因子実験」，「3因子実験」などと呼ぶ．「1因子実験」では，ただ1つの因子の水準を何段階かにとり，それぞれの水準の下で2回以上繰り返すが，その他の原因系の条件はすべて一定に保つ．「2因子実験」では，2つの因子の水準

のすべての組合せを1回以上実験するが，この際他の原因系の条件はすべて一定に保つ．表 1-1 および表 1-2 に，「温度因子」A と「触媒量因子」B を用いたそれぞれの例を示す．一般に，**多因子実験**とは，2つ以上の因子を同時に取り上げる実験をさすが，特に，**取り上げた因子の水準のあらゆる組合せを各 r 回（$r \geqq 1$）実験する**ものを（多因子）**要因実験**（factorial experiment）という（実際，後述するように，「主効果」・「交互作用」という**要因効果**に関する情報がすべて取り出せるからである）．したがって，たとえば A 因子3水準，B 因子4水準，C 因子2水準の「3因子実験」では，$3\times4\times2=24$ 通りの処理組合せを少なくとも1回（以上）実験することになる．第8章以降で述べる「直交表実験」もまた多因子実験ではあるが，これは要因実験の**一部実施**にあたることが多い．

表 1-1　1因子実験の例

A_1 100°	A_2 120°	A_3 140°
○	○	○
○	○	○
○	○	○
○	○	○

表 1-2　2因子実験の例

	A_1 100°	A_2 120°	A_3 140°
B_1 4%	○	○	○
B_2 5%	○	○	○
B_3 6%	○	○	○
B_4 7%	○	○	○

因子の選び方

いま，「温度」と「触媒量」を A, B 2因子とし，それぞれの最適条件を知りたいとき，次の3通りの実験計画が考えられる（比較の公平を期するために，いずれも 24 回の実験をすることにする）．

（1）A, B 各因子について，それぞれ「1因子実験」を行なう（表 1-3）．

表 1-3　2つの1因子実験

A_1	A_2	A_3
○	○	○
○	○	○
○	○	○
○	○	○

（B_0 に固定）

B_1	B_2	B_3	B_4
○	○	○	○
○	○	○	○
○	○	○	○

（A_0 に固定）

（2）A, B 両因子について，「2因子実験」を行なう（表 1-2 の組合せを2回ずつ実施したと考える）．

（3）A, B 両因子の水準の組合せのうち，技術的に判断して適当なものだけを実施する（たとえば，A_1B_3, A_1B_4, A_2B_2, A_2B_3, A_2B_4, A_3B_1 の 6 組合せ

を各4回ずつ実施する).

これら3通りの計画の優劣を判断するために，結果の数字（収率とする）がかりに表 1-4 に示すような値であったとしよう．まず，表 1-3 の2つの1因子実験からはどういう情報が得られるであろうか．ここで，A因子3水準の比較をするとき，B因子の水準は B_0 に固定されている．この B_0 の取り方について，次の2通りの場合を考えてみる：

（i） B_0 を $B_1=4\%$ としたとき，実験が十分精度よく実施されたとすれば，Aの最適条件は表 1-4 より $A_3=140°$ であることがわかる．次に，この A_3 を A_0 としてB因子の実験を行なったとすれば，B_2 の収率が最高であることがわかるから，A_3B_2 が最適条件で，その期待収率は 80% であることを知る†.

表 1-4 A, B の組合せにおける結果の想定値（収率 %）

温度 ＼ 触媒量	A_1 100°	A_2 120°	A_3 140°
B_1 4%	40	60	70
B_2 5%	60	85	80
B_3 6%	70	90	70
B_4 7%	85	80	55

（ii） B_0 を $B_4=7\%$ としたとき，上と同様の手続きを経て，まず $A_1=100°$ を見つけ，次に $B_4=7\%$ が最高と知る．よって，A_1B_4 が最適条件で，その期待収率は 85% となる††.

このようにして，最初に固定した B_0 の条件によって結論がまったく異なるばかりではなく，このときにはそのいずれもが真の最適条件（A_2B_3 で 90%）ではない（もっとも B_0 を B_3 にとっておれば A_2B_3 に到達できたであろうが，B_3 にとるべきことが実験する前からわかっていたとするのでは，議論として不適当である).

これに反して，（2）の2因子実験の方法では，表 1-4 の 12 個の組合せのすべてで実験することになるから，それらに対応するデータが得られ，最適条件を見のがす危険は少ないと考えられる．よって，一般に，**1因子実験をいくつ**

† ここで B_2 の条件下で，再び A 因子の比較実験を行なうと，A_2 が最高であることがわかり，次に，A_2 の下で B 因子を比較すると，A_2B_3 に到達する．このようにして，この方法も逐次近似的には真の最適条件に到達する可能性をもっている.

†† このときには，逐次実験をしようとしても A_1B_4 から動けない．ここは，A因子の実験でも，B 因子の実験でも最適条件を与える点になっている．ちょうど，山の尾根道に沿って登って行くと，東西の方向にも南北の方向にも，そこが最高点で風通しがよく，一息入れたくなる場所にあたる．しかし，山登りのときには，たとえば，その西北の方向に頂上が見えて，それを見失うことはない.

か並行して実施するよりは，それらの水準を組み合わせて多因子実験を構成する方がよいといえる．

ところで，（3）の方法からはどんな情報が与えられるであろうか．「温度が低いときには触媒量は多くしなければならないが，温度を高くするときには触媒量はあまり多くしない方がよい」というようなことがわかっておれば，（2）のように両因子の水準のあらゆる組合せを実験するよりは，（3）の方法をとる方が賢明のように見える．ただ（3）の方法では，A，B 両因子のそれぞれ単独の効果，およびその相互作用といったような情報を得ることはむずかしく，候補に上がった6組合せのうち，どれがいちばんよいかという知識しか得られない．すなわち，6水準をもつ「1因子実験」をしたのと同じになる．

以上要するに，過去の技術的情報を整理してそれから学ぶことは重要であるが，これを誤りのない前提として，実験を組むのが（1）および（3）の方法であり，（2）はどちらかというと，過去の知識をあまり信用しない計画である．しかし，（2）の計画でも，その「水準の組合せ方」について，後述するように，過去の知識を十分用いるならば，非常に実際的になるし，また，信頼できる技術的情報が十分でないときには，（2）による以外に正しい結論を得ることはできないのである．さらにまた，（2）の個々のデータは（1）のデータの**2倍の情報を含んでいる**といえる．なぜなら，（2）のデータはいずれも，A，B 両因子の比較に用いることができるが，（1）の各数字は，どちらか一方の因子の比較にしか用いることができない．これを各因子の水準の繰返し数でみると，表 1-3 では，Aは4回，Bは3回であるのに対し，表 1-2 を2回実施すれば，Aの水準は8回，Bのそれは6回繰り返されることになる．

このようにして，いくつかの因子の最適条件を知りたいときには（12·1「因子の分類」も参照せよ），それらの因子を全部同時に取り上げる「多因子要因実験」を採用すべきことがわかる（もちろん，これは原則論であって，個々の場合には，その実験目的と，技術的情報の信頼度によって，1因子実験や2因子実験を行なうこともありうる）．ところが，因子の数がふえると，その水準の組合せ数は加速度的に増大する．たとえば，各因子の水準数をすべて2としても，5因子で $2^5=32$，6因子で $2^6=64$ となり，3水準因子ばかりとすると，4因子で $3^4=81$ となる．こんなに多くの組合せでは，その1回ずつを実施することも困難であるから，その 1/2 あるいは 1/3 の組合せだけを実施する**一部実施法**という計画が考案されている．この種の計画は，第8章以下で述べる**直**

交表を用いることによって，容易に構成することができる．

水準の選び方

まず，**水準数**はどうして決めるか．「質的因子」については取り上げるべき水準があらかじめ決まっていることが多い．たとえば，原料を3社から購入しているとすれば，この因子の水準数は3にとる以外にない．比較される品種・銘柄が10個あれば，その水準数は10に決まる．これに反して，「温度」や「触媒量」や「反応時間」のような「量的因子」の水準数はどのようにも決めることができる．このとき，前段で結論したように，1つの実験になるべく多くの因子を同時に取り上げる（多因子実験にする）べきであるとするなら，実験規模を大きくしないためには，各因子の水準数はできるだけ少ない方が望ましい．そこで，まず2水準にとることにしよう．図1·3(a)に，温度因子を2水準にとったときの結果が示されている．これからわかることは，「温度140°の方が100°より高い収率を与える」ということだけであって，温度の最適条件

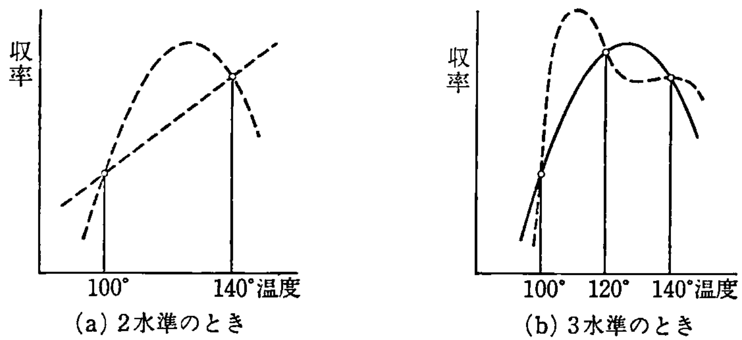

図 1·3　温度に対する収率の応答曲線の想定図

は，図に画いた2本の点線のいずれを想定するかによって，140°よりもっと高い温度であるとも，100°と140°の中間にあるともいえることになり，決定できない．一方，中間の120°をも含めた3水準をとる場合を考えてみよう．このときは，図の(b)に実線で示した**2次曲線**（**放物線**）が想定され，最適条件は計算によって120°と140°の間にあることが知られる（6·5参照）．もっとも，このときでも図に点線で示したような曲線を想定すれば，最高収率は100°と120°の間で得られることになるが，だれもそれを主張し続けることはできないであろう．なぜなら，このような収率の曲線は，温度因子の水準

の変域をあまり大きくしない限りは，すなわち，**局部的には**，1次式あるいは2次式で近似できることが知られているからである．2次式をあてはめるには3点あれば十分なのであるから，3水準の実験をすることに特別の意義がある．よって，実験計画法では，量的因子の水準数は2または3にとるのを基本とし，4水準以上にとることは特殊の目的のある場合に限ると考えてよい．

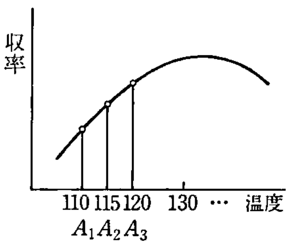

図 1·4 水準のとり方の悪い例

次に，量的因子の**水準の幅**はどのように決めればよいか．再び，上の温度因子3水準を例にとると，それらを，たとえば100°, 105°, 110°としたのでは狭すぎる．なぜなら，このときには収率に，はっきりした差が認められないかもしれないし，また，温度管理が悪ければ，±5°ぐらいの差は打ち消されてしまうかもしれない．反対に，水準の幅をもっと広げて，100°, 130°, 160°とすると，160°のときには予期通りの化学反応が進行すると保証できないかもしれない．要するに，水準の幅が狭すぎると有用な情報は何も得られないが，広すぎると危険が伴うということになる．しかしながら，実験をする目的は，現状に満足せず，よりよいものへの飛躍を探求することであるから，**水準の幅は，技術的に可能と考えられる限り，はじめはなるべく広くとるべきである**．実験を繰り返し，技術的情報が積み重ねられるにつれて，水準の幅はだんだん狭くしてもよい．警戒すべきことは，図 1·4 に示すように，応答曲線のふもとの方でコソコソと条件を変えて実験するような結果になることである．

ところで，1つの因子（たとえば温度）の水準だけを大幅に変え，他の原因系の条件はすべて現状のままにしておくと，一般によい結果が期待できないのは，多くの現場の技術者の知っているところである．しかし，第2，第3の因子の水準をも同時に変えると，どのような結果が得られるかについては，現場の技術者も的確に答えられない場合が多い．それは，人間の直観や勘は1変数の世界ではよく働くが，多変数（多次元）についてはあまり鋭くないのが通常だからである．この観点からも「多因子実験」を採用して，多

表 1-5 2つの因子の水準の組合せ方

	A_1 100°	A_2 120°	A_3 140°
B_1	5	4	3
B_2	6	5	4
B_3	7	6	5
B_4	8	7	6

くの因子の水準の組合せを自動的に実験する意義があるが，さらに前段で述べたような「温度」と「触媒量」の効果についての予備知識があれば，表 1-2 のように機械的に 2 つの因子の水準を組み合わせる代わりに，表 1-5 のように触媒量の 4 水準を温度因子の水準ごとに**ずらせて**選ぶこともできる．すなわち，温度 120° のときには原案の 4, 5, 6, 7% を比較するが，100° のときには，それぞれ 1% ずつ高くし，140° のときは，逆に 1% ずつ引き下げるのである．こうすると，図 1・5 に×印で示した条件——温度低く触媒量も少ない組合せ（100° で 4%）や，温度高く触媒量も多い組合せ（140° で 7%）——は採用しなくてよいことになる．すなわち，技術的にはナンセンス（ときには危険！）と思われる組合せを排除して，それこそ（3）の計画のように，いずれも最適条件の候補となりうるような組合せばかりを実験することができるのである．しかも，これは，本質的には A 因子 3 水準， B 因子 4 水準の「2 因子実験」と考えられるから，各因子の主効果とその交互作用（第 6 章参照）を評価することができる．このような水準の決め方を，**水準の組合せ方**または**ずらせ方**と呼ぶことにしよう．水準をずらすと，一般に「交互作用」項の変動が小さくなり，後述のように，「一部実施法」の採用が容易になる．

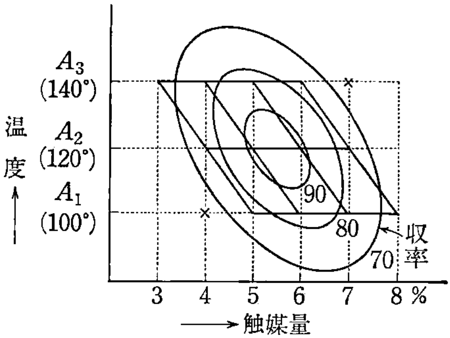

図 1・5　2 因子についての収率等高線

1・3　統計的判定の考え方

1・1 の (2) で述べたように，本書で対象とする実験では，同じ条件の下で得られたデータにも若干のバラツキが伴うということであった．これは，実験をいかに精密に行なっても管理しきれない「誤差」(error) が存在することを意味する．この「誤差」は，「過誤」(mistake) や失敗とは本質的に区別すべきである．実験がいかなる過誤も失敗もなしに遂行されても，なおかつデータに付随するバラツキの原因の総称が「誤差」要因である．「誤差要因」とは，その実験では制御しきれなかった無数の微小な原因の集まりである．ちょうど，品質管理図上でのバラツキがこれにあたる．それは，一定の条件の下で生産が