現代人の統計 5

実験計画法の基礎

林 知己夫 編

早川 毅 著

朝倉書店

実験計画法の基礎

早川　毅

現代人の統計
5

朝倉書店

著者略歴

1940年　千葉市に生まれる
1962年　名古屋大学理学部卒業
現　在　文部省統計数理研究所研究室長
　　理学博士

現代人の統計 5
実験計画法の基礎　　**定価 2000 円**

1977年12月20日　初版発行

著　者　早川毅（はやかわたけし）
発行者　朝倉鑛造
発行所　株式会社 朝倉書店
東京都新宿区新小川町2-10
郵便番号 162
電話 03 (260) 0141 (代)
振替口座 東京 6-8 6 7 3 番

〈検印省略〉

政弘印刷・渡辺製本

3341-111805-0032

はじめに

統計学において実験計画法は大きな地位を占めている．実際に観測値を得る時に，いかにして実験を計画し，設計するか，また得られた観測値をいかにして解析するかという統計における基本的考え方を含んでいる．

本書は実験計画法の基本的考え方，およびその解析についての説明を与え，より高度な理論，方法への橋渡しを与えることを目的としている．その為，多くのことを羅列的に述べることは避け，いくつかの基本的考え方，方法，解析計算について懇切丁寧に述べることにした．本書を執筆するにあたり設定した水準では取扱いにくい数理統計学の基本的な諸性質は，数値実験をし，グラフにより経験的に理解できる様に心がけた．

最後に，本書を執筆する機会をいただいた林 知己夫氏，松原 望氏に感謝の意を表し，また遅筆な著者を終始励まし，いろいろ御教示いただいた朝倉書店の方々に御礼を申し上げたい．

1977年11月

早川　毅

目　　次

1 実験計画法の考え方

1.1 実験計画法

実験計画法とは，ある目的のために実験をするとき，その目的にかなうようにいかにして実験を計画し，観測値を精度よく得るかを問題とする統計学における一手法である．

この方法は1920年代にフィッシャー(Fisher, R. A.)によってその基本的な考え方が作られ最初は農業試験において使われていた．しかし今日では実験科学の広範な分野(工学，医学，薬学等々)，また社会科学，心理学等の分野にもその適用が広げられている．特に，農学，医学実験においては得られる観測値の数は，工業実験の場合と比べて一般に少ない．工業実験においては比較的短時間に多くの観測値を繰り返し得ることができるから，知識の集積が容易であるが，一方，農業，医学関係の場合には個々の観測値を得るのに時間がかかったり，同じ環境を維持することがむずかしい．このため観測値の解析には特に統計学的な手法が重要になってくる．

さて，ある目的(たとえば，製品の品質の改良，薬品の薬効の検査)実験をするときには，どのような要因がその実験結果に本質的な影響を与えているかを調べて，その要因の比較をし，実験目的にかなう最適な要因条件を見つけ出すことを考える．また観測値を精度よく得るために，観測に伴って入ってくる実験誤差を除去するように実験を組み立てねばならない．ここで**'実験誤差'**とは，われわれがある目的のために実験をするとき，その観測結果に本質的に影響すると考えられるいくつかの要因を取り出してもなお，われわれが事前に知ることができない要因の影響が観測値の変動の中にあるかもしれない．このような構造的な知識の欠如による誤差と，実際に観測するときにでる観測値の変

動による誤差を組にしてわれわれは‘実験誤差’といい，このような構造的な知識を拡大するように実験結果を解析し，また観測に伴う変動を小さくするように実験を管理しなければならない．

いま，工場で製品の品質改良を目的とする実験を考えるとしよう．技術者は事前にいろいろな要因を，いままでの経験，知識，技術水準などを検討，整理して，仮に‘原料の種類’，‘作業条件(湿度)’，‘作業条件(温度)’などを取り上げたとする．実験結果に影響を与えると考え，取り上げて比較したい要因を**因子**(factor)という．各因子はいくつかの条件からなっているとする．

原料の種類　　：A_1, A_2
作業条件(湿度)：40%, 50%, 60%
作業条件(温度)：20°, 30°, 40°, 50°

として，各要因(因子)の持つ条件を**水準**という．このとき，‘原料の種類’は2水準，‘作業条件(湿度)’は3水準，‘作業条件(温度)’は4水準という．

われわれはどの因子が実験結果に強い影響を与えているのか，また影響を与えていると考えられる因子があるならどの水準が本質的なのか，また取り上げた因子どうしが互いに影響しあうことはないか，等々を知りたいし，またそれらの具体的な大きさをも推測したい．

品質実験で，特に‘原料’の差が品質の良さに決定的に影響していると考えて，‘原料の種類’のみを取り上げて実験する場合はこれを**一因子実験**という．また‘原料の種類’，‘作業条件(湿度)’が重要として実験をすれば**二因子実験**ということになる．一つの実験を行うときに，同時に多くの因子を取り上げて行う実験を**多因子実験**といい，取り上げた因子の水準のすべての組み合わせについて実験する方法を**要因実験**とよんでいる．たとえば前出の3因子実験では $2\times3\times4=24$通りの水準の組み合わせすべてについて行うとき要因実験という．要因実験ではすべての水準の組み合わせを行うため，因子の数，水準の数が少し多くなるとその組み合わせの数は膨大になる．実験は各組み合わせの水準において同じ環境(条件)で行われねばならないから，すべてを一定の条件にしておくことは大変である．農業，医学実験などでは因子数，水準数を小さくして

行われる．一方，工業実験では多くの因子を同時に取り扱うが，その水準数は比較的少ない．このような場合には多因子実験となる．事前の技術的，理論的考察からすべての組み合わせについて実験しなくても必要な情報が得られることがある．水準の組み合わせの一部を用いて行う**一部実施法，交絡法**を用いることにより多因子実験はその労力から解放される．要因実験については第2～6章，直交表を用いる一部実施法については第7章以降で扱う．

各要因の水準の数の定め方はたとえば‘原料の種類’というような質的な要因に対してはその数を変えることもないし，また場合によってはその数をふやすこともできないかもしれない．しかし‘作業条件(湿度)’，‘作業条件(温度)’というような量的な要因はその水準の数をいくらでも変えられるし，また水準の間の幅も簡単に変えることができよう．このような場合にはなるべく実験として可能と考えられうるだけ幅を広く取っておくとよい．そして実験を繰り返すことにより得られる結果を最適なものにするように水準の変更をしていけばよいであろう．

フィッシャーは実験の精度を高めるために**三つの原則**を提案している．

(i) 反復の原則： 観測誤差の評価．

(ii) 完全無作為化の原則： 系統誤差の偶然誤差への転化．

(iii) 局所管理の原則： 系統誤差の除去．

ある実験を行うとき，ただ1回だけの観測では，その観測値が真の値からどれくらい離れて観測されているのか評価はできない．このとき2回以上の観測をすれば観測値のバラツキの程度を見ることにより観測値の真の値からの変動がどれくらいあるかが評価できる．つまりわれわれは得られた観測値の全体からしかその観測値の精度について議論をすることはできないから，同じ状況での観測値が多ければ多いほど，観測値が得られている状態を知ることができるわけである(**反復の原則**)．

次に3個の‘処理’A, B, Cを3回ずつ行う実験を考えてみよう．全部で9回の実験がなされるが，1日ではできず3日に分けて，毎日3個ずつ行うとする．最初に，処理を

(i)　$\underbrace{AAA}_{1日目}$　$\underbrace{BBB}_{2日目}$　$\underbrace{CCC}_{3日目}$

のように行ったとする．この場合には毎日同じ処理が3回ずつなされることになるが，同じ処理を続けて行うことによる‘なれ’の系統的な誤差が入ってくるので，本来知りたい処理による効果と‘なれ’による効果とが分離できなくなり好ましくない．そこでこの‘処理’の順序をランダムな順序にして次のようにしてみる．

(ii)　$\underbrace{ABB}_{1日目}$　$\underbrace{CAB}_{2日目}$　$\underbrace{CCA}_{3日目}$

というようにA, B, Cの並べ方を無作為化することにより‘なれ’による効果は全体として消し合うようになり，またこの無作為化した実験を繰り返して行えば各処理効果は平均的には平等になってくる．このように‘処理’を完全無作為な順に並べて，系統誤差を除去するようにする(**完全無作為化の原則**)．

さて，作業の能率は最初のうちは作業になれないために低く，2日目，3日目という順で能率が良くなるとしよう．このような場合，並べ方(ii)では作業能率の悪い初日に‘処理B’が2回なされ，作業能率の良い第3日に‘処理C’がやはり2回なされている．この場合‘処理B’は本来の‘処理B’による効果が作業の不能率のために低く評価され，一方‘処理C’は不必要に高く評価されることになる．このような状態をさけるには，同じような作業能率の状態で比較したい3‘処理’を同時に行えばよい．たとえば，

(iii)　$\underbrace{BAC}_{1日目}$　$\underbrace{ABC}_{2日目}$　$\underbrace{CBA}_{3日目}$

としてみる．毎日‘処理’の順序は完全無作為に並べられて行うとする．このように同じような‘実験の場’をまとめて，その中で比較したいものを完全無作為な順序で行うことを**局所管理**といい，同じような‘実験の場’を**ブロック**という．このようにして行う実験方法を**乱塊法**とよんでいる．

‘日’による作業能率への影響をブロックを作ることにより除去できるとして，次に，もし毎日の作業で1回目の作業よりは2回目，3回目の作業のほうが能率が高いというようなことが作業者に考えられるとしたらどうであろうか．乱塊法ではブロック内の作業能率は作業順序には無関係であるとした．今

度は，これをもっと細分化して考えようというものである．作業順序に差があるとすれば，(iii)では‘処理B’は2回とも中程度の能率でなされ，‘処理C’は2回とも高程度の能率でなされている．そこで，この順序についても同等にしようとすれば次のような並べ方が考えられる．

(iv) $\underbrace{\mathrm{ACB}}_{1\text{日目}}$ $\underbrace{\mathrm{CBA}}_{2\text{日目}}$ $\underbrace{\mathrm{BAC}}_{3\text{日目}}$

このように並べ方に‘日’によるブロックと‘順序’によるブロックの2種類を入れて作られる方法を**ラテン方格法**とよんでいる．

以上の考察より，実験を行い観測をする場合には，‘実験をする場’を完全無作為化して系統誤差が入ってくるのを防ぐこと，また同じような‘実験の場’を作り系統的な誤差を除去する局所管理をすることの二つが重要なことが理解されよう．

われわれは本書において，観測精度を高めるために‘実験の場’を‘完全無作為化’し‘局所管理’していくための手法のいくつかについて勉強することになる．

1.2 確率変数

われわれは以下において本書を読むのに必要な数学的道具とその使い方について述べることにする．問題を簡単にするために，次のような例を考えてみよう．たとえば祭のときなどに売っている‘タンキリアメ’の長さを観測する場合を考えよう．熟練した職人が包丁を使って棒状のアメを切って作るのであるが，彼がアメの長さを正確に2 cmにしようとして切っても必ず少し長いものや短いものができてくる．しかし腕の良い職人ほど大体2 cmに近い長さのアメを多く切り，極端に長いものや短いものはそんなに多くは切らないと考えら

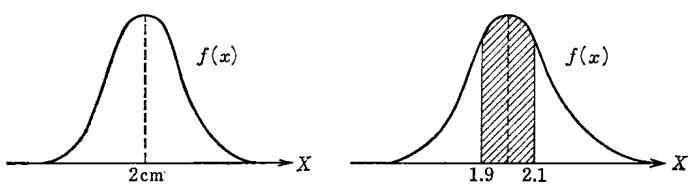

図 1

れる．アメの長さの出方(頻度)を示すグラフは大体図1(a)のように2cmを中心とした左右対称な形をしていると考えてよいであろう．

観測される‘タンキリアメ’の長さがどの範囲にあるかを示す変数をXで記し，**確率変数**という．そしてXの値の‘出やすさ’を示すグラフ$f(x)$をXの**密度関数**という．われわれはこの出やすさを示す密度関数$f(x)$とX軸で囲まれる面積を1とするようにグラフを定めることにしよう．そして，‘タンキリアメ’を1個取り上げてみて，その長さが1.8cmから2.1cmの間になっている可能性(確率)を図1(b)において，密度関数$f(x)$と$x=1.8$と$x=2.1$の線で囲まれた斜線の面積として考えることにする．

さて，このグラフからもわかるように‘タンキリアメ’の長さは2cmを中心に長かったり短かったりするから，その変動が職人の技術による変動であり，2cmからの誤差と考えられる．もちろん§1.1で見たようにこの変動分は単に‘アメ’を切るときに出る技術的な変動と同時に，いろいろな要因(たとえば，‘アメ’を切りはじめたときの‘なれ’の差，職人の体調等々)による変動も含まれていよう．この内容については特にここでは立ち入らないことにしておこう．これより，長さXは

$$X=2\,\text{cm}+\text{誤差}\ e \qquad (1.1)$$

と書くことができる．この誤差eは次のような性質を持っていると考えよう．

(i)　誤差eは正にも負にも平等に出やすい．

(ii)　誤差eは0に近い値を取りやすく，0から離れた値は取りにくい．

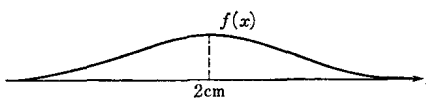

図2

また職人の体調，気分によっては2cmの長さに切るのにバラツキが大きくなることもある．このようなときには‘出かた’を示す密度関数は2cmを中心に少し平たい，すそを引く形のものとなろう．

以上よりわかることは‘タンキリアメ’の長さが2cmを中心として左右対称に出やすく，また場合によってはその出方が広い範囲にわたって出やすくなる(形が平たくなっている)ことである．このような状態を取り扱うのには，数学的に密度関数$f(x)$として

$$f(x)=\frac{1}{\sqrt{2\pi}\,\sigma}e^{-\frac{1}{2}\frac{(x-\mu)^2}{\sigma^2}} \qquad (-\infty<x<\infty) \tag{1.2}$$

という式をあてはめて考えると，いろいろと便利である．

(1.2) のグラフは，① $X=\mu$ を中心に左右対称である，② $X=\mu$ で最大となる，③ X が μ から離れれば0になる，④ $f(x)$ と X 軸とで囲まれる面積は1となる，という性質を持っていて，われわれの取り扱っている現象を説明しやすい．この密度関数を持つ確率変数 X のことを**正規確率変数**といい，'$X\sim N(\mu,\sigma^2)$' と書く(これは確率変数 X が正規変数で，その密度関数として (1.2) を持っていることを示す)．この (μ,σ^2) を変数 X の**母数**という．

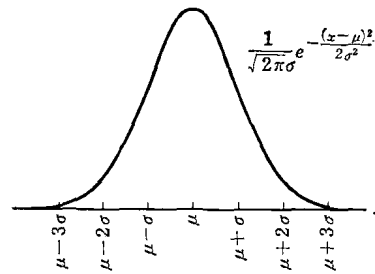

図 3

確率変数 X が取る平均的大きさと，その '出かた'，'広がり' を示す尺度として，われわれは**平均値(期待値)**と**分散**を次のように定める．

定義1 確率変数 X の密度関数を $f(x)$ とするとき，その平均値 $E(X)$ は，

$$E(X)=\int_{-\infty}^{\infty}xf(x)dx\equiv m \tag{1.3}$$

とし，また分散 $V(X)$ は

$$V(X)=E[(X-m)^2]=\int_{-\infty}^{\infty}(x-m)^2f(x)dx \tag{1.4}$$

として定める．ただし，これらの値が有限な値であるとする．

平均値 $E(X)$ は X の取る値とその値の出やすさ $f(x)$ との積をすべての場合について加えた形をしている．

また分散 $V(X)$ は，X の平均値 m からの距離の平方 $(X-m)^2$ とその値の出やすさ $f(x)$ との積をすべての場合について加えた形をしている．つまり分散は，平均値 m から観測される値がどれくらい離れているかを見る尺度といえよう．

平均値，分散の性質として，次のようなものがある．

(i) $E(aX+b)=aE(X)+b$ (1.5)

確率変数 X を定数 a 倍し，定数 b を加えた $aX+b$ の平均的出やすさは，X の平均値 $E(X)$ を a 倍し，定数 b を加えたものになっている．

これより，$a=0$ とすれば $E(b)=b$ となり，定数の平均値はその定数そのものになる．

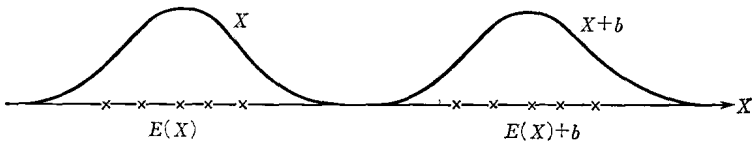

図 4

(ii)　$V(aX+b)=a^2V(X)$　　(1.6)

$a=1$ とすれば $V(X+b)=V(X)$ となる．すなわち確率変数 X の出方を示すグラフは X に b を加え平行移動した $X+b$ の出方を示すグラフとまったく同じであるから，その分散も変わらないことを意味する．

$b=0$ とすれば，$V(aX)=a^2V(X)$ となる．これは変数 X の出方の値を a 倍すると，その出方は分散の尺度ではかると a^2 倍となることを意味する．

(1.6) は (1.5) を用いて次のように示せる．

$$\begin{aligned} V(aX+b) &= E[\{aX+b-E(aX+b)\}^2] \\ &= E[\{aX-aE(X)\}^2] \\ &= a^2E[\{X-E(X)\}^2] \\ &= a^2V(X) \end{aligned} \qquad (1.7)$$

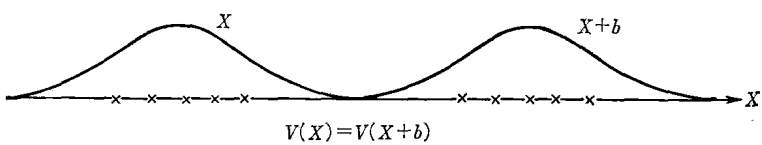

図 5

正規確率変数 $X\sim N(\mu,\sigma^2)$ に対しては，その平均値，分散は

$$E(X)=\mu, \qquad V(X)=\sigma^2 \qquad (1.8)$$

となることが示される．

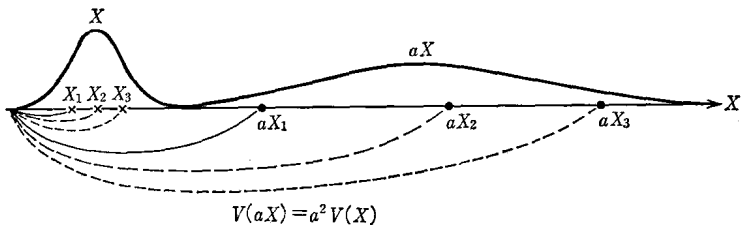

図 6

すなわち，正規確率変数の密度関数 (1.2) を定める母数はその平均値 μ と分散 σ^2 であることがわかる．

さて，$X\sim N(\mu,\sigma^2)$ とするとき，(1.1) の形に表現すると，

$$X-\mu=e$$

となり，(1.5), (1.6) を用いて，

$$E(e)=0, \qquad V(e)=\sigma^2 \tag{1.9}$$

となる．

正規変数 X を μ だけ平行移動させ $X-\mu=e$ としても，やはり正規変動をするので，誤差変動 e は，

$$e\sim N(0,\sigma^2)$$

となることがわかる．

次に分散の平方根 $\sigma=\sqrt{\sigma^2}$ (**標準偏差**という) を用いて，

$$\frac{X-\mu}{\sigma}=\frac{e}{\sigma}\equiv Z \tag{1.10}$$

とすれば，

$$E(Z)=0, \qquad V(Z)=1$$

となる．“確率変数 X が正規変数であれば (1.10) のように変換しても Z はやはり正規変数となる”という事実が確率論から示されるから，結局 $Z\sim N(0,1)$ となる．このように平均値 0，分散 1 を持つように変数を変換することを**標準化**といい，変換された変数を**標準正規変数**という．

$Z\sim N(0,1)$ より，(1.2) において $\mu=0$, $\sigma^2=1$ とすれば，標準正規変数の密度関数は

$$f(z)=\frac{1}{\sqrt{2\pi}}e^{-\frac{z^2}{2}} \qquad (1.11)$$

と書ける．

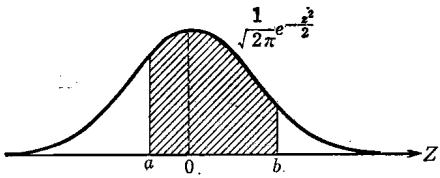

図 7

標準正規変数 Z が区間 $[a, b]$ の間にある可能性(確率)を $P(a\leq Z\leq b)$ で示す．これは図 7 で斜線で囲まれた面積として与えられる．

標準正規変数の確率については数表が作られている．付表 1 は，$P(0\leq Z\leq a)$ の値が $a=0$ から 4 まで 0.01 刻みで与えられている．

たとえば $P(0\leq Z\leq 1)=0.3413$.

また，$P(-1\leq Z\leq 2)=P(-1\leq Z\leq 0)+P(0\leq Z\leq 2)$ である．Z の密度関数は $Z=0$ に関して対称であるから，$-1\leq Z\leq 0$ での面積と $0\leq Z\leq 1$ での面積とは同じである．したがって $P(-1\leq Z\leq 0)=P(0\leq Z\leq 1)$ であり，前述の式は下記のようになる．

$$\begin{aligned}P(-1\leq Z\leq 2)&=P(0\leq Z\leq 1)+P(0\leq Z\leq 2)\\&=0.3413+0.4772\\&=0.8185\end{aligned}$$

任意の平均値 μ と分散 σ^2 を持つ $X\sim N(\mu, \sigma^2)$ に対して，$P(a\leq X\leq b)$ はどのようにして求めるか．

$$\begin{aligned}a\leq X\leq b &\Leftrightarrow a-\mu\leq X-\mu\leq b-\mu\\&\Leftrightarrow\frac{a-\mu}{\sigma}\leq\frac{X-\mu}{\sigma}\leq\frac{b-\mu}{\sigma}\end{aligned} \qquad (1.12)$$

は同じ不等式を意味し，第 3 番目で，$(X-\mu)/\sigma=Z$ は標準正規変数である．すなわち，区間 $[a, b]$ 内に X がある確率は標準正規変数 Z が区間 $[(a-\mu)/\sigma, (b-\mu)/\sigma]$ の中にある確率と同じであることを意味する．

$$P(a\leq X\leq b)=P\left(\frac{a-\mu}{\sigma}\leq Z\leq\frac{b-\mu}{\sigma}\right) \qquad (1.13)$$

たとえば，“$X\sim N(2, (1/2)^2)$” とする．このとき

$$P(1.5\leq X\leq 3)=P\left\{\frac{1.5-2}{\frac{1}{2}}\leq Z\leq\frac{3-2}{\frac{1}{2}}\right\}$$