

# 実験計画学入門

実験的研究の基礎

木村 朗

# 実験して得たデータの解釈

- 実験をして得たデータの解釈をどうするか

例

- 1) A, B, Cの3種類の方法で行ったROMの改善効果を調べた. どの方法がいちばんよいか?
- 2) 筋肉を温めるとROMはどうかを調べた.

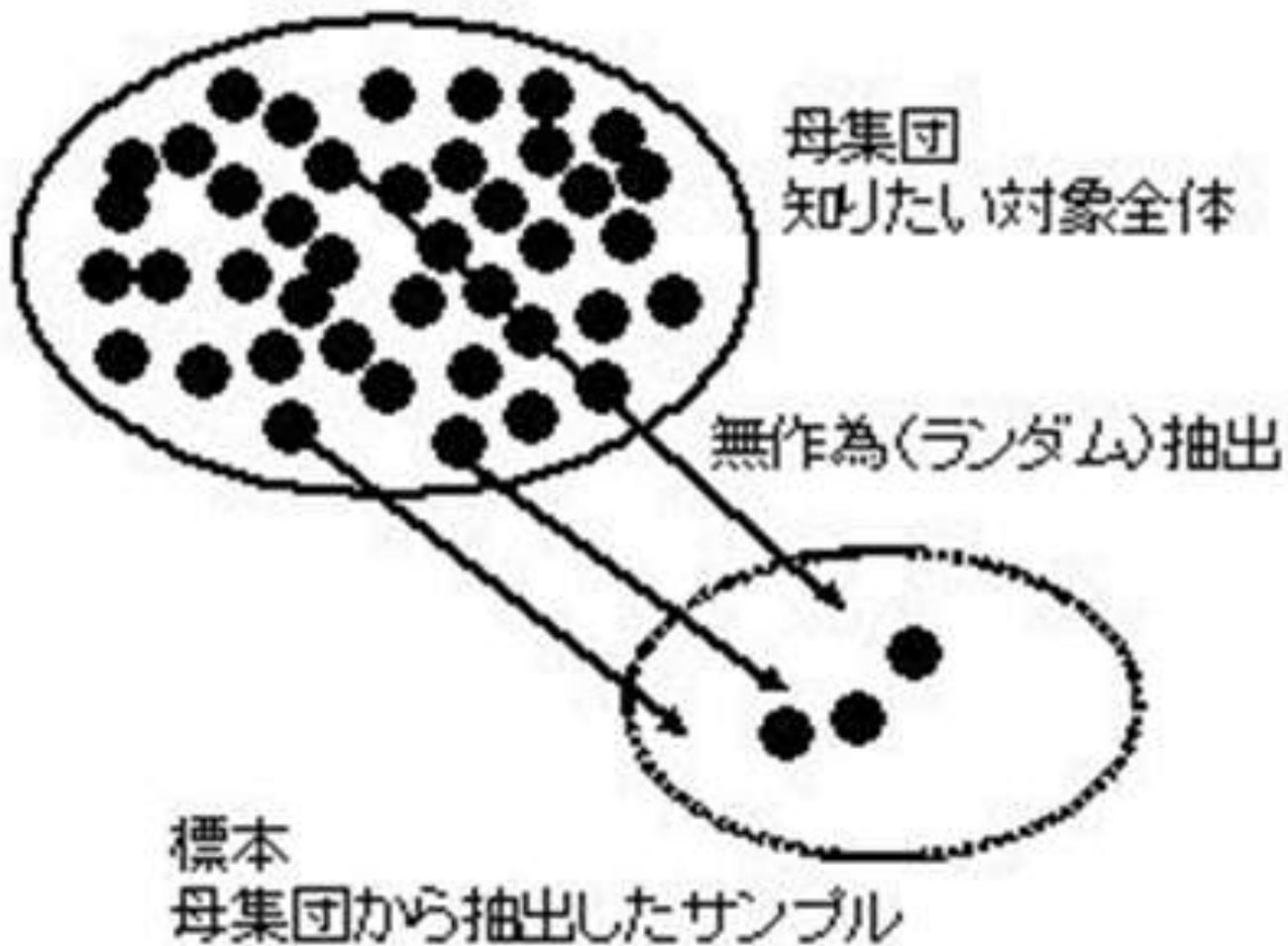
データにはばらつきがあるので, 1つや2つだけを調べてもそれが本当に正しいかは確信を持ってません. ではデータをどのように取ったら確信を持ってよいのでしょうか?

- 1) 平均が知りたいなら何ケース調べたらよいでしょうか？
  - これに対する答えを見つけるには統計的推定を用いる
- >もう一度医療統計学を見る

2) 同じ方法でROM制限のある患者にROM訓練をした2組の患者群があった。しかし、一方は柔軟性を高めるスポーツをしていた者が多くROMが大きくなったようだ。この仮説を証明するには2つの群からそれぞれ何人を調べたらよいだろうか？

これに対する答えを見つけるには統計的検定を用いる

>もう一度医療統計学を見る



# 実験を成功に導く鍵

- フィッシャーの3原則を考慮した実験計画を立てる
  1. 繰り返すこと(反復) > その意味: 偶然によるバラつきを少なくするため(偶然誤差を少なくする)
  2. 無作為化 > その意味: 偏りのない標本抽出をすること(系統誤差を少なくする)
  3. 局所化(ブロック化) > その意味: 調べる要因以外のすべての要因を可能な限り取り除く(少なくする・一定にする) > > 例) 年齢が影響するようなら、年齢の影響を除くために、標本の年齢をできるだけ等しくする。

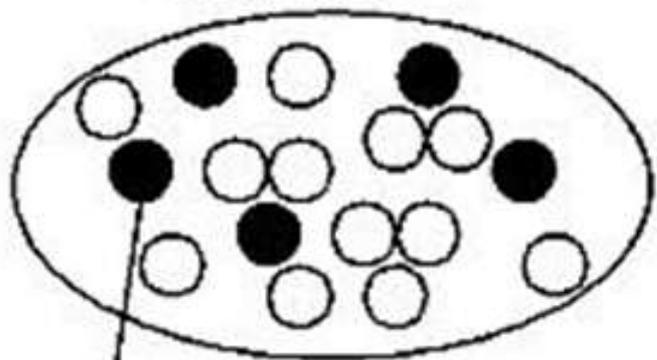
# 少ない実験でもできるだけの情報を 取り出したい

- 数が多くても偏った標本では正しい結論は得られません。
- 上の図のように無作為に抽出された標本からは統計的手法を用いて、少数の標本であっても何らかの結論を得ることができます。
- 標本数が少ないと確実さは小さくなりますが、結論自体は一定の範囲で出すことができます。

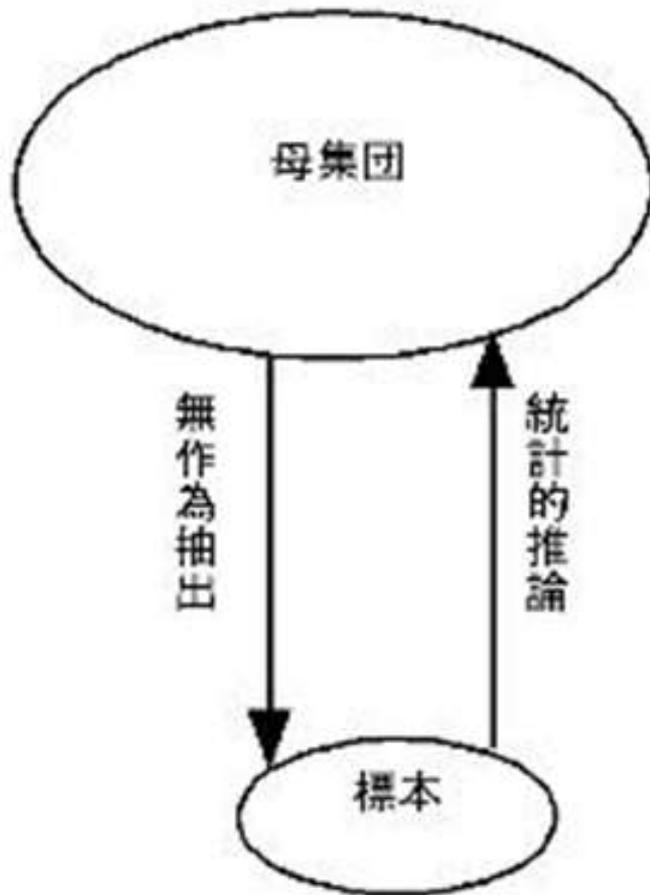
# 実験計画学のメリットは大きく3つ

1. 実験回数を少なくできる.
2. 精度がよくなる. あるいは精度がわかる.
3. 実験データの変動を解析できる.

母集団n個



x個の標本を無作為抽出



氣象要因: 氣溫, 日射量

栽培: 育苗, 施肥, 農藥

土壤要因

收量

遺傳的要因(品種)

生物的要因(病害虫)

# 実験計画を立てる流れ1

第1番目 調べたい現象(結果)に影響を与える候補(因子)を挙げる

- ある状態(結果、事件の結果)に影響を与えるもの(犯人)は何(誰)？
- 因子(犯人の候補者(容疑者))を挙げる

# 実験計画を立てる流れ2

- 第2番目 調べたい現象(結果)の測定方法、候補(因子)の測定方法のめどを立てる
- 測定・検査方法と代表値を決める

＞代表値とは、何か？ 忘れてたら---

医療統計学のサイトを見直す

# 実験計画を立てる流れ3

- 第3番目 調べたい現象(結果)の代表値、候補(因子)の代表値について、分析するための統計的推定・統計的検定方法のめどを立てる。

＞複数の因子を検討する場合、たいてい分散分析を用いることになる

＞分散分析でなければできないこと＞反復を繰り返すデータから誤差を計算し、データの精度(質)を判定することができるというメリットがある

# 実験計画を立てる流れ4

- 第4番目 調べたい現象(結果)の代表値、候補(因子)の代表値を得るためのデータの集め方のめどを立てる

＞無作為化(ランダム化)につながる重要なプロセス

実験の対象が声を掛けやすい人だと、すでに偏り(バイアス)が生じる可能性が高い

万が一、そうだったとしても、実験(解析)で用いるデータはランダムに選ぶことが大切(それでも限界はある)

# 肝

- 実験研究では分散分析のあらましを理解することに尽きる
- また、因子(容疑者)が犯人である可能性が高くないと、結局逃がしてしまう。
  - 因子の特徴的な種類(分類)を知っておくべき

# 要因(因子)分類には種類がある

## 1) 制御因子

その最適条件(水準)を知るために取り上げる因子で、実験の場ではもとより、その結論を適用すべき場(より多くの物に適応する場、臨床)においても、その条件を制御できる因子

>ゆで卵を作るのに“ゆでる温度“が影響するだろう(この温度は、私たちが自由に変えられる。これを制御因子という)、“その程度”(実際の温度のレベル、幅)のことを水準という

\* 犯人候補ともいえる(アリバイを消す作業が待っている)---もしくはうまみスープを作るのに最もベストな調味料の調合程度(条件)と考えても良い

## 2) 標示因子

生産の場において、特性を最適化する因子を制御因子という

標本の少ない実験では、結果に影響を及ぼす因子であるが、大量かつ多くの反復を行うと、その影響が少なくなる性質をもつ因子のこと

＞実験研究では重要であるが、大量生産や大規模な疫学研究では分析上考慮しなくてもよいもの

### 3) ブロック因子

実験の精度を高めるために、実験の場の局所管理に用いる因子で、その水準自身は特性値に若干の影響を与えるかもしれないが、他の(制御・標示)因子とは交互作用を持たないと考えられる因子のことブロック因子といいます。

ブロック化(局所化)を行う時に決める因子のこと  
例) 年齢の水準(細かい分け)

#### 4) 層別因子

実験の場でも、生産・臨床・適用の場で制御できない因子を層別因子といいます。

＞治療法の割付において、予後に影響があると予測され、治療群の割付決定に用いる項目のこと

年齢・性別など、遺伝素因によって決まっているようなもの（実験計画者が操作できないもの）

# 実験の計画のコツ

ふつうは要因を複数同時に取り扱う要因実験の方が利点が多いです。複数の要因を同時に実験すると、

- A. 実験の精度自体を高めること、
- B. 交互作用を見積もれること、
- C. 実験のデータのばらつきをいろいろな角度から評価できる分散分析ができることなど有利な点が多いです。

# 統計解析法の決定

- どのような統計解析をするかを決めます  
データを取り終わってからどういう解析をするかを決めるのは本当は正しくありません。

2つの母集団の平均値の有意差検定 t検定

3つ以上の母集団の平均値の有意差検定 分散分析

分散分析の後、質的因子の場合は多重検定、量的因子の場合は回帰分析を行い、最適な水準がどれかを定めることができる。

# 3. データを集めたら

## 1. 基礎統計量を計算する

### A. データの数が多いとき

データが30以上あれば、度数分布、ヒストグラムを書いてデータの分布を調べる。

次に平均、分散、標準偏差などを計算する。

データが正規分布から大きく外れると判断したら、メジアン、モードなども有用な統計量である。

### B. 異常値をチェックする

データの中に異常に大きい値、あるいは小さい値があるときの対処法は以下に述べるような方法がある。

a. は必ず行うべき。 b. ~d. はどれを用いるべきかは実験の目的、データの性質などを検討して、できる限り実験を開始する前に決めておく。

## a. 異常値の原因が明らかなきとき ☆☆☆☆☆

異常値の原因を調べ、測定におかしな点があるときは除去する。

## b. 反復数を増やす☆☆☆

もし可能であれば、実験を繰り返して、データを増やすとそのデータが異常値であるかより明確になり、かつ異常値が平均値に及ぼす影響も軽減される。

## c. 異常値の除去☆☆

異常値を除去したいときはスミルノフ法(スミルノフ・グラブス法)\*<sup>1</sup>あるいはディクソンのQテスト\*<sup>2</sup>で検定し、異常値であると認められるときは除去できる。根拠もなく、不都合なデータを捨てるのは好ましくない。

\*1 データの母集団が正規分布に従うという仮定の下で、データの中で特に大きい値や特に小さい値が外れ値であるかどうかを検定する手法。

\*2  $n=4$ から25の試料中に1個の異常値があった場合に棄却する検定

## d. 内部平均を用いる方法☆☆

集めたデータのうち両極端な値，すなわち最大値と最小値を除いたデータから得た平均を内部平均という．内部平均を用いるとデータのばらつきはかなり小さくなることが多い．データが3つのときの内部平均はメジアンと一致する．

# データ間の相関を分析する

得られたデータについてはどのデータとどのデータに相関があるかを散布図と相関係数を計算することで調べる。

はじめは想定しなかった関係を見つけることもあるでしょうし、最初に考えていたような関係がないこともある。

データをグラフのように見える形にすることは相関関係を知るだけでなく、異常値やデータのグループ分けを見つける上でも重要。

# 目的に応じた統計解析

- 実験計画で決めた分散分析をここで行います.

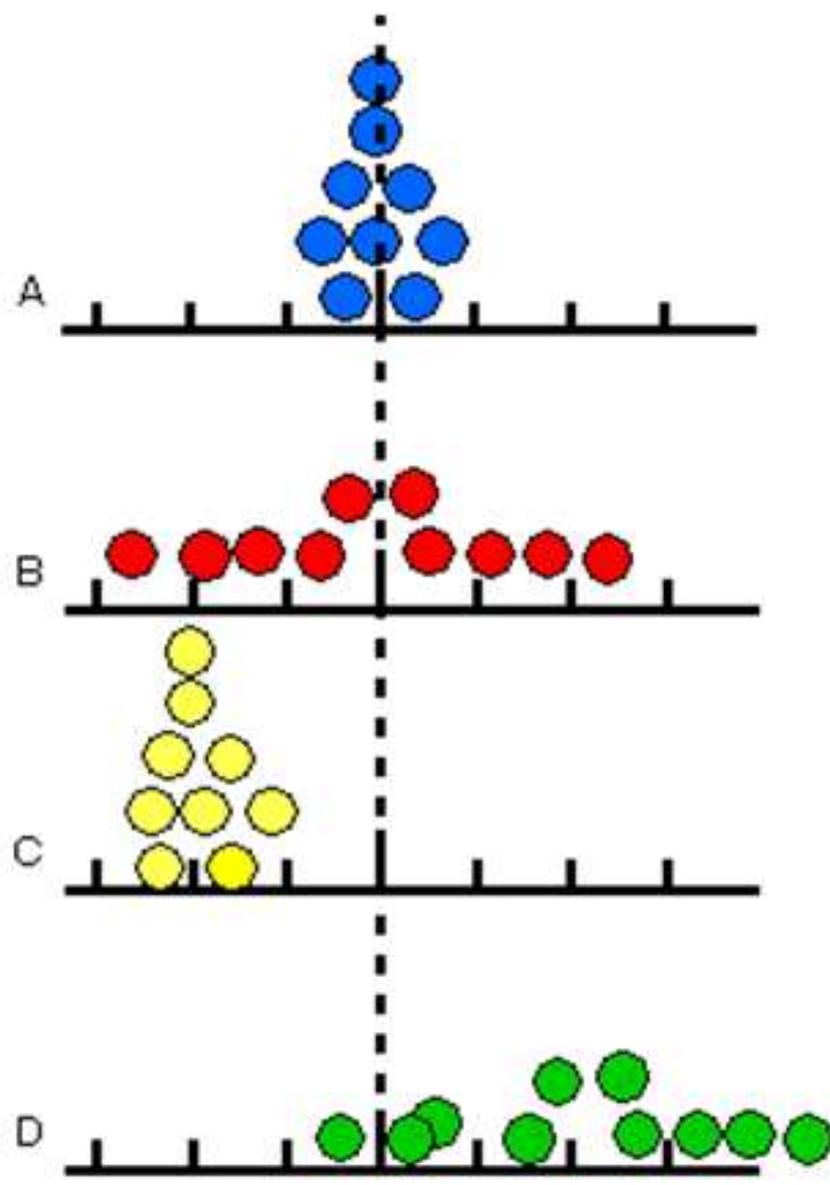
# 分散分析と実験計画法

- 実験にともなう誤差をどう制御するか？

データを取り扱う実験には誤差が必ず存在する。

- この誤差をいかに制御するかが実験で得られたデータの正確さ, 信頼性を高める上で重要.
- 誤差は大別すると, 系統誤差と偶然誤差に分けられる.

＞ 系統誤差は偏りを持った誤差



- 図を見てください。A, B, C, Dの4人がある測定を何回か繰り返しました。
- 真の値は点線で示してあります。
- A, Bは平均は真の値に近いのですが, Bはデータのばらつきが大きいようです。一方, C, Dは平均も真の値から外れています。さて, **BとCではどちらの分析がよいのでしょうか。**

もし, **Cの値を補正する手だてがなければ, Bの方が分析はよいことになります。** なぜなら, Bならば**実験回数を増やせば,**

真の値に平均が近づきますが, Cはいくら実験をしても真の値には近づきません。系統誤差とはこのような**Cの出す誤差だ**といえます。

いくら大数の法則でたくさんデータを集めれば真の値に近づくといっても, 系統誤差ではデータを増やしても平均からの誤差が小さくならないのです。

すなわち系統誤差があり、それに気づかなければ、分散分析して得られる誤差には系統誤差が入らないことになり、データの解釈を大きく間違ってしまう！

したがって、誤差全体を小さくする以前に、系統誤差を減らす、なくす、そして減らせない場合には偶然誤差に転化することが必要になる。

その方法を示すのが、フィッシャーの三原則だったのだ。

- 系統誤差の例

10頭の同じ品種の雄の羊を集めたとしても、齢、体重、親の遺伝その他いろいろな要素が実験を左右する。

実験するときには系統誤差を出すとわかっているものはあらかじめ除去するのがよいが、

体重、齢などの全く同一な羊を10頭も集めることは現実的ではない。

あるいは系統誤差は出るのだけれど、制御どころか、系統誤差の原因の定かでないものもある。

分析機械の日や時間による微妙な変動などは系統誤差を生むが、原因を詳しく調べようとするのは手数。

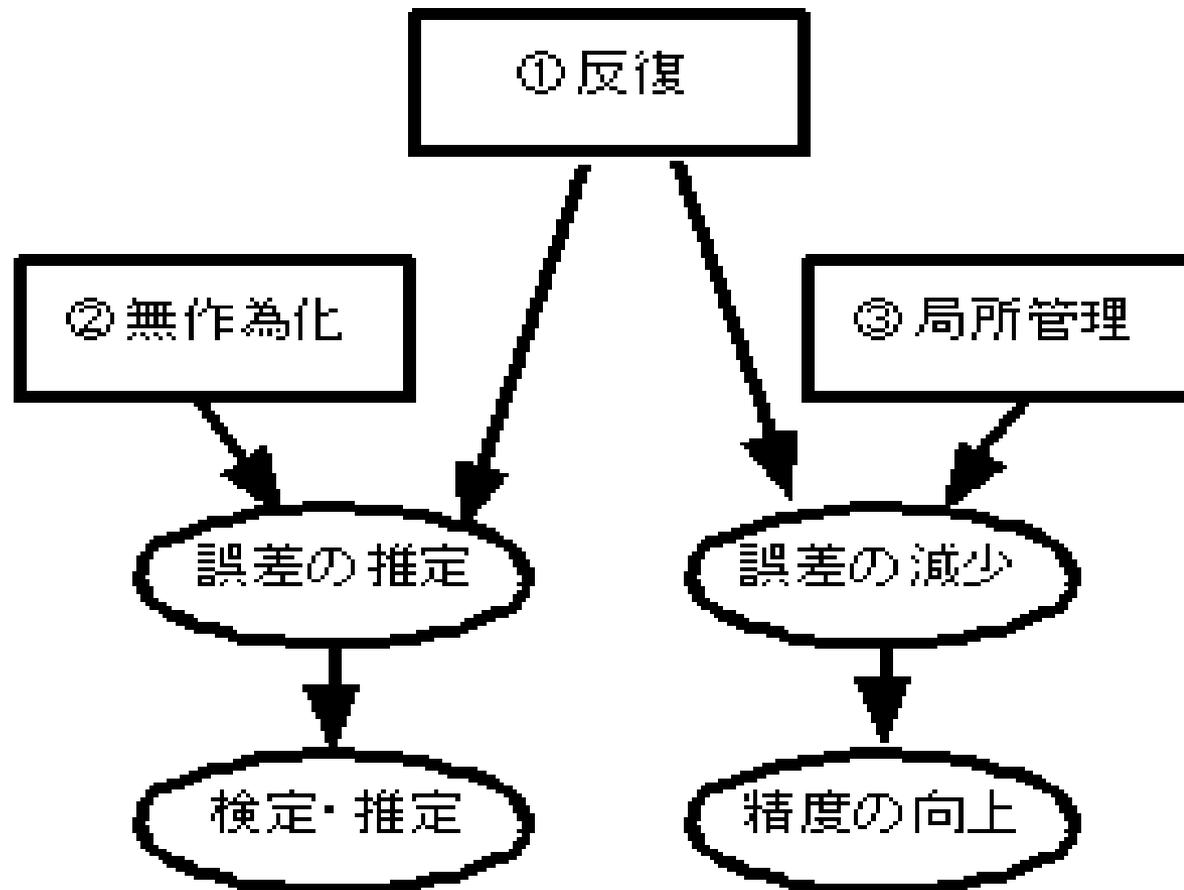
- フィッシャーの三原則ではこのような系統誤差を制御するためにまず

## 1. 反復

誤差を見積もることができないと、制御するにも制御できない。誤差を知るには少なくとも同じ条件で実験を複数繰り返す必要がある。これを反復という。

さらに反復を増やすと、平均についての誤差である標準誤差そのものも小さくなる。

すなわち、反復を設けることによって、誤差の評価と誤差の減少の2つが実現する。



- フィッシャーの三原則では図のように3つの原則がそれぞれ誤差の推定と減少に役立ち、誤差を推定することで分散分析のように統計的検定・推定が出来、誤差を減少することによって精度を向上することができる。

## 2. 無作為化

系統誤差のあるデータに分散分析などの統計的手法を適用しても正しい結果は得られない。

系統誤差をできるだけ偶然誤差に転化する方法を無作為化といいます(ムズカシク言うとな)。

例えば、羊は個体によって餌に対する反応が違いかもしれません。与える餌をどの羊に与えるかを無作為に(ランダムに)決めることによって、羊の個体間差異に基づく系統誤差は無作為化されます。

例えば羊の体重によって、結果が変わる場合

小さい体重の羊3頭にA, 大きい体重の羊3頭にBの薬を与えるなら, その結果には系統誤差が入り込む.

こういう系統誤差は先ほどの図のCの結果と同じく, データの見かけのばらつきを小さくする(なぜなら, 体重の近い羊のデータはよりばらつきが小さいから)からよけいに解釈を間違えることになりかねない.

体重に関係なく, 無作為に実験すると体重のちがいによる誤差(系統誤差)を偶然誤差に転化するので, データの見かけ上の誤差は大きくなる. しかし, 無作為化した結果, 偶然誤差を分散分析などで評価できるので, 実際には実験の検出力は高まる.

- 例えば、定量分析の場合でもA, B, Cの3つのサンプルを3回反復測定するときに、A, A, A, B, B, B, C, C, Cの順で分析する人が多い。
- これも見かけ上は誤差は小さくなる、しかし、**分析のときの周りの環境(温度, 試薬, 器具)の影響による系統誤差を偶然誤差に転化していない**から真の値から大きくずれてしまう危険がある。
- したがって、9回の分析を無作為に行うべき。
- その場合、見かけ上データの誤差は大きくなりますが、それは実験のどこかに系統誤差の原因があることを示すよい証拠であり、さらに実験の精度を高める機会を与えられたということにもなる

# 3. 局所管理

フィッシャーの三原則の最初は反復であるように、反復がない実験は誤差を評価しようがない

フィッシャー自身はそういうのは実験ではなく、ただ経験を増やしたにすぎないとコメントしたらしい、

反復は最も大事だが、反復を増やすと誤差を別の意味で増やしてしまう可能性がある。

無作為化の例で挙げたように、羊の頭数を増やすことは、反復を増やすことだが、全く同じ羊がこの世にいない以上、羊の個体による差は誤差を増加させることになりま  
す。

定量分析でも何回かの分析をするために何時間も実験するうちに、実験の場の微妙な変動によって、誤差が増大する可能性があります。

- このように反復を増やすときの誤差の変動を、羊であれば体重の近いものを、

分析であれば時間的に近いある部分（午前とか午後とか）に、

実験で比較したい1組を当てることによって、誤差をその部分の違いに転化して、除去することができます。これを局所管理といい、このようなできるだけ均一にした実験の場の一部をブロックといいます。

（不利な条件にさらされたら、みんな不利にしてしまえば、そのハンデがなくなるということ）

実験計画法では、系統誤差をこのような**局所管理**によって、ブロック間の差にするだけでなく、分散分析によって、ブロック間の誤差（すなわちブロックにした結果、除去できた誤差）も定量できます。

		施 肥 量		
		少	中	多
品 種	1	○	○	○
	2	○	○	○
	3	○	○	○

表 1. 2 因子（品種・施肥量 各 3 水準）実験

○をつけた組み合わせすべてで実験する。

		施 肥 量		
		少	中	多
品 種	1	A	B	C
	2	B	C	A
	3	C	A	B

表 2. 3 因子（品種・施肥量・圃場 各 3 水準）実験

A, B, C は圃場の水準を表す。

二元配置分散分析を用いて二要因3水準の繰り返しのある実験をする場合の配置

ラテン方格法を用いて、二要因3水準の繰り返しのある実験を $3 \times 3 \times 3 = 27$ 試行せずに $3 \times 3 = 9$ 試行で済ませるブロック配置

# 4. 乱塊法の分散分析

実験をする場合、反復と無作為化はかならず必要。

反復がなければ誤差を見積もることができないから、分散分析できない。

無作為化をしなければ、データの変動が処理によるものなのか、誤差によるものかがわからない。

系統誤差を処理変動から分離できないから。

フィッシャーの3原則のうち、反復と無作為化を満たすものを完全無作為化法という。

# 乱塊法とラテン方格法

**乱塊法**はブロックを持つ配置法で、すべての水準の組み合わせをブロックの数だけ実験するもの。

**ラテン方格法**は組合せを減らしながら各因子の各水準が他のすべての因子の水準と組合せられるような方法。種々編み出されている。**直交計画**(各因子を互いに直交するベクトルとして解釈できるため)と呼ばれている。

# 1. 実験計画法

- 実験計画学で取り扱う実験には次の2つの特徴がある。
  - (1) いろいろな条件を人為的に設定して、その結果の比較を目的とする実験である
  - (2) 同じ条件の下で実験を繰り返しても、結果(データで示される)は必ずしも一定ではなく、かなりのばらつきを示す。

(1)については、以下でどのような条件を設定することができるのか、さらに条件は1つだけとは限らず、複数の条件を設定する実験が必要な場合があり、そのときのそれぞれの条件の特徴を考える。

(2)については実験では設定する条件以外を均一に必ずできるわけではなく、そのような不均一さが原因で誤差が生じます。いかにして誤差を小さくするか、あるいは実験結果に偏りのないような誤差に転化するかを考える。

# 実験計画法の目的

- (1) 実験で取り扱う因子(要因)を適切に選択すること
- (2) 実験につきものである誤差を制御すること

## 2. 因子と水準

- 実験においてその条件を種々に変えて比較するものを因子といい、因子の取りうる条件を水準という

- 

例えば、花の栽培実験では、因子として、品種や施肥量などを考えることができる。

因子	水準
品種	コシヒカリ, ハナエチゼン, ヒノヒカリ
施肥量	0, 4, 8, 12g/m <sup>2</sup>

# 因子(再掲)

- (1) 制御因子 その最適条件(水準)を知るために取り上げる因子
- (2) 標示因子 制御因子と交互作用があるために取り上げる因子のうち、**実験の場では制御できる因子**
- (3) ブロック因子 局所管理に用いる因子。 他の因子と交互作用はない。
- (4) 層別因子 **実験の場でも制御できないが、他の因子と交互作用を持つ因子。**

- 実験で明らかにしたい因子と水準を明確にする
- その因子と交互作用(単独では影響は示さないが、他の因子があると効果を発揮すること)のある因子のうち、現実的に重要なものがないかをよく考える。

# 3. 誤差の制御

## (1) 分散分析と誤差

因子が効果があるのかを知る統計的手法が分散分析.

分散分析では主効果や交互作用の大きさを誤差変動と比較するので誤差を小さくできれば検出力が高まる

どのようにしたら誤差を小さくできるか？

# 実験計画法とは

実験計画法は1925年頃に英国の農場試験場の遺伝学者、Sir Ronald Aylmer Fisherにより考え出された方法論

例えば農作物の収穫量がどのような要因（例えば肥料の量や農薬の量）によって増減するのか、そしてその要因をどのくらいの値に設定することが望ましいのか、というような事を客観的に把握する事を目的とします。

# 実験計画法で何ができるのか

- 例えばお米の収穫量を考えて見ます。同じ広さの田んぼからの収穫量は多い方が望ましいと言えます。収穫量に影響を与える要因は例えば与える水の量や肥料の量、そして農薬の量などいろいろと考えられます。実験計画法を用いると様々な要因の中から、より大きな影響を与える要因を把握することや、それぞれの要因について適切な値を、より効率的に且つ客観的に把握することが可能です。
- 例えばこの例では肥料と農薬が収穫量に大きな影響を与えていることや、肥料や農薬をそれぞれどのくらい与えると最も収穫が増えるのかを把握することが可能。

# 実験計画法の2つのステップ

実験計画法は以下の2つのステップで構成されます。

- 計画性を持ってデータを集めるステップ。
- 集められたデータを分析するステップ。

# 計画性を持ってデータを集めるステップ

- このステップではデータを分析するステップで良い解析ができるように 実験を計画します。この計画ステップでは以下のような事を 明確にする事が必要です。

## 実験の目的

- 例えば「収穫量を増やす」、「製品強度を増す」など実験を行う目的を明確にします。ここでの収穫量や製品強度を表す値は**特性値**と呼ばれます。

## 因子の選定

- 目的の特性に影響を及ぼすさまざまな要因の中から、実験で取り上げる要因(**因子**と呼ばれる)を決定します。数多くの要因をすべて実験することは通常現実的でないので、実験で取り上げる因子の選定が必要です。

## 水準の決定

- 選定された因子の水準(実験で設定する値)を決定します。例えば与える肥料の量は3水準(50g/平米, 100g/平米, 150g/平米等)とする等を決定します。水準は処置と呼ばれる場合もあります。

## ブロックの決定

- 実験を複数のブロックに分けて行うかどうかを決定します。例えば収穫量の実験では畑のいくつかの区画(ブロック)に分けて実験をするかどうかを決定します。

# 実験計画法の種類

因子の数や実験順序などにより、以下に示すよういくつかの種類があります。

- 一元配置法

因子が一つの場合に用いられます。

- 二元配置法

因子が二つの場合に用いられます。

多元配置法因子が三つ以上の場合に用いられます。

- 乱塊法

ブロックの種類(ブロック因子)が一つある場合に用いられます。ラテン方格法ブロックの種類(ブロック因子)が二つある場合に用いられます。

# 集められたデータを分析するステップ

- 計画に基づき行われた実験で集められたデータは、次のステップで分析されます。分析には分散分析、回帰分析、多重比較などの手法が用いられます。
- 分散分析を行うと分散分析表を得ることができます。分散分析表は例えば以下のようなものです。

# 分散分析を行うと分散分析表を得ることができます。

## 分散分析表

ソース	自由度	平方和	平均平方	F-値	F-有意度
処置	3	1636.5	545.5	5.4063429	0.0068759
残差	20	2018	100.9	14	42
合計	23	3654.5			

# 読み方

- ※平方和は偏差平方和、
- 平均平方は不偏分散、
- F値は分散比、
- F有意度はP値とそれぞれ呼ばれることもあります。
  
- **分散分析表に出力されるF-有意度はその値が小さければ小さいほど、実験で得られた結果が偶然ではないことが示されます。**
- 実際にどのくらいの有意度が得られた時に「偶然ではない」とみなすかは多くの場合、慣習的に0.01や0.05がそのボーダーラインとして使われます。有意度が0.01以下である場合のみを有意であるとみなす場合、0.05以下で有意であるとみなす場合、更にマーケティング分野などの分析では有意度が0.15以下でも有意であるとみなす場合もあります。
  
- 分散分析の結果は解釈され、要因の効果、交互作用などを読み取ります。

## 参考文献

- Winer B J, 『Statistical Principles in Experimental Design』
- G B Wetherill, 『Elementary Statistical Method』
- G W Snedecor and W G Cochran, 『Statistical Methods』
- Davis O L, 『The Design and Analysis of Industrial Experiments』

分散分析についてはkimualilaboサイトの分散分析編を必ず参照のこと