

# 分散分析

**分散分析**：3群以上の平均の差の検定。要因効果の存在を、誤差に対する相対的な大きさから判定する。

**交互作用**：複数の要因間において、ある特定の水準同士が組み合わさったときだけに現れる相乗効果。

## 8.1 3群以上の平均の差の検定

本章では、農学をはじめ実験を行う分野に欠かせない**分散分析** (analysis of variance : ANOVA) を扱います。近年は社会科学でも実験が行われるようになつたので、あらゆる分野で必須の分析手法といえるでしょう。また、第10章で学ぶ実験計画法では、誤った判定を下さないためにはデータをどのように収集すべきかについて、この分散分析を土台にして考えます。

さて、分散分析を一言でいうと、実験の**要因**（因子；factor）の効果があるかないかを判定するための検定手法です。要因というと難しいですが、実験の結果として観測される**特性** (characteristic) に影響を及ぼしていると思われる処理条件（**水準**；level）の違いのことです。いいかえれば、**処理の水準を変えることで生じた複数の群の母平均に差があるかどうかを見る**のが分散分析です。

もちろん、前章で学んだt検定でも要因の効果は検証できます。しかし、2群だけを対象とした検定でした。例えば肥料（要因）が植物の成長（特性）を促す効果を持っていることを確認する場合を考えて見てください。肥料を使った／使わないの2水準の処理だけではなく、施肥量を3段階以上設定したい場合もあるでしょうし、3種類以上の肥料の効果を見てみたい場合もあるでしょう。そのような**3群以上の平均に対しても、差があるかどうかを検証できる**ように考え出されたのが分散分析なのです（2群にも使えますがt検定と同じ結果となります）。

処理水準は、10g／20g／30gのように**単調性**があつても構いませんし、薬

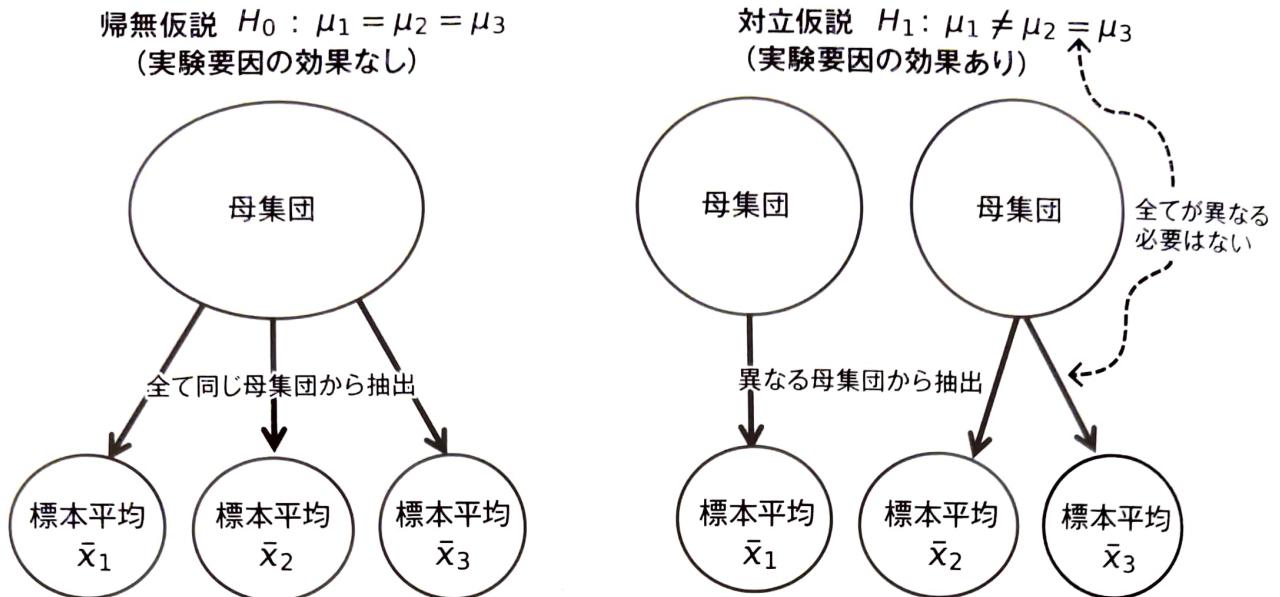


図 8.1 「分散分析」の仮説の考え方

A／薬B／薬Cのようなカテゴリカルなものでも構いません。

分散分析の仮説の考え方は図8.1のようになります ( $t$ 検定同様、全ての母集団は正規分布に従い、分散は等しいという前提です)。

帰無仮説は、要因効果“なし”で、全ての群の母平均が同じという内容になります (3群ならば  $H_0 : \mu_1 = \mu_2 = \mu_3$ )。一方の対立仮説には注意が必要です。帰無仮説とは逆の内容なので、要因効果“あり”なのですが、必ずしも全ての母平均が異なっているということではありません。いずれかの群間に差があれば効果ありといえるのです (例えば図8.1では水準1と水準2・3の母平均が異なるので  $H_1 : \mu_1 \neq \mu_2 = \mu_3$ )。

このように、たとえ対立仮説が採択されても、どの群とどの群の間の母平均に差があるのかわからないのが欠点ですが、次章で学ぶ多重比較法を用いることで、差のある群間を特定できます。

## 8.2 (対応のない) 一元配置分散分析

### ● 効果を確かめたい要因が1つの場合

分散分析は、 $F$ 分布の名前の由来にもなったR・フィッシャー (トピックス⑥) によって考案されました。その名の通り、実験で観測したデータの“分散を分析”するのですが、なぜ分散を使うと複数の平均間の差を検定できるのでしょうか？ そのあたりを本節ではわかりやすく説明していきたいと思います。なお、分散分析といってもいろいろあります、効果の存在を確認したい要因の数によって「○元配置分散分析」と呼びます。確認したい要因が1つな

8  
● 分散分析

表 8.1 肥料と小麦の収量			
対照群	肥料なし (水準1)	肥料A (水準2)	肥料B (水準3)
特性 (単収, t/ha)	1 3	8 6	10 14
群(標本) 平均	$\bar{x}_1 = 2$	$\bar{x}_2 = 7$	$\bar{x}_3 = 12$
			$\bar{x} = 7$

要因  
(処理水準)  
反復 (2回/水準)  
総平均

らば一元配置分散分析、2つならば二元配置分散分析という具合です。

また、*t*検定同様、データの取り方として、群間で個体に対応関係がある場合とない場合とでは検定統計量の内容が異なります。まずは、要因が1つだけで、群間に対応関係のない、**対応のない一元配置分散分析** (one-way ANOVA) から始めましょう。

理解しやすいように、実際に数値の入った仮想事例で説明しましょう。表8.1は、肥料の種類を変えたことが、小麦の収量を変化させたかどうかを検証するための圃場（農場）実験です。表の値（グレー部分）は、3種類の肥料それぞれにつき独立した（ちゃんと区切った）2区画で小麦を栽培し、その单収(1haあたり何t収穫できたか)を表したものです。事例では、手計算しやすいように2区画／水準からしかデータを取っていませんが、本来は検出力分析（事前分析）で、適切な反復数を計画しなければいけません（独立した実験を複数回実施することを反復と呼びます）。

なお、今回は「肥料なし」という**対照群** (control group) を設定し、施肥自体の効果量を推定できるようにしています。こうした実験を**対照実験** (control experiment) と呼び、とくに人への薬の効果を見る場合は偽薬 (placebo) の対照群を設定することで、偽薬効果も排除できるので大切です（2群の実験でも同様です）。

さて、この事例では、肥料なしの水準（群）の平均 $\bar{x}_1$ は2t/ha、肥料Aの平均 $\bar{x}_2$ は7t/ha、肥料Bの平均 $\bar{x}_3$ は12t/haとなり、肥料を変えたという要因の効果があるよう見えますが、本当にそうでしょうか？母集団においても各群の平均单収に差があるといつて良いかどうかを、(対応のない) 一元配置分散分析で確かめてみましょう。

## ● 要因効果の捉え方

当然ながら、事例で観測されたデータは完璧ではありません。というのも、屋外での実験ですから、気象条件（日の当たり方や風の吹き方など）や土壤条件など、多くの要因が影響しているのです。

件（肥沃度や水はけの良さなど），全ての環境を圃場全面において完全に統制することは不可能だからです。いいかえれば，観測されたデータには，検証したい（実験的目的となる）要因の効果だけでなく，その他の雑多な要因による効果も含まれてしまっているのです。

そこでフィッシャーは，それら（研究目的以外の雑多な要因による効果）を全て誤差要因による効果として，データから分離すればよいと考えました。そうすれば，その残りが，実験目的である要因効果になるからです。つまり，分散分析では，観測データを目的である要因と目的ではない誤差要因の2つの効果から構成されていると考えるのです。そこで問題になるのが，どのようにこれら2つの要因効果を分離するかということです。

そもそも要因効果とは，どのような形で捉えられるのでしょうか。

フィッシャーは，こうした要因の効果は，データのバラツキで捉えることができると言いました。というのも，観測されたデータは，必ずバラついています。似た値にはなるかもしれません，環境が完全に統制された宇宙ステーションで栽培でもしない限り，小数点まで完全に一致することはありません。では，そのバラツキはどうして生じたのでしょうか。もちろん，いろいろな影響を受けてバラついたのでしょうか。その影響こそが2つの要因（実験の目的となる要因と，それ以外の要因をひっくりめた誤差）による効果であることにフィッシャーは気がついたのです。また，都合の良いことに，バラツキの指標の中でも偏差平方和（分散分析では**変動**（variation）と呼びます）ならば足し算や引き算ができるため，データ全体の変動も2つの要因効果による変動に分離できるのです。

ですから，分散分析の基本的な考え方は，図8.2のようになります。最終的に，（実験目的となる）要因効果の有無を，これら2つの変動を比較することで検証します。

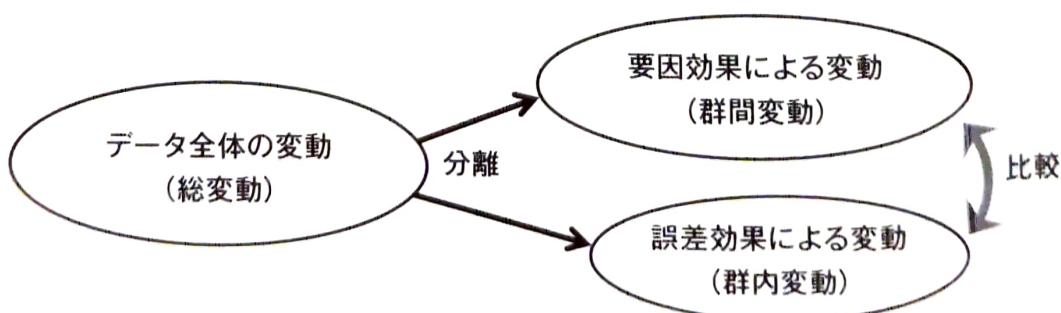


図 8.2 一元配置分散分析の考え方

## ● 変動の計算

それでは、実際に小麦の事例を使って、それぞれの変動を求めてみましょう（分散分析を理解するには手を動かすのが一番です）。順番ですが、まずはデータ全体の変動（総変動）を求め、次に要因効果によって生じた変動（群間変動）を求め、最後に誤差効果によって生じた変動（群内変動）を求めます。事例でいえば、表8.2のように横のバラツキが群間変動、縦のバラツキが群内変動となります（なお、表の $\leftrightarrow$ は、変動のもとになる偏差を示しています）。

表 8.2 肥料と小麦の収量（再掲）

水準1	水準2	水準3	
1	8	10	
3	6	14	
2	7	12	7

群間変動（要因効果）

群内変動（誤差効果）

### ①データ全体の変動（総変動）

データ全体のバラツキである**総変動**（total variation）を捉えるには、図8.3のように、まず6つある個別の特性値 $x$ から全体の平均（総平均： $\bar{x}$ ）である“7”を引いて偏差 $(x - \bar{x})$ を計算します。次に、偏差を2乗して負の符号を消した偏差平方 $(x - \bar{x})^2$ を計算します。最後に、全ての偏差平方を足し合わせて変動 $\sum \sum (x - \bar{x})^2$ を計算します（行方向と列方向の総和なので $\sum$ が2つ）。すると、小麦の事例の総変動は“112”となるはずです。

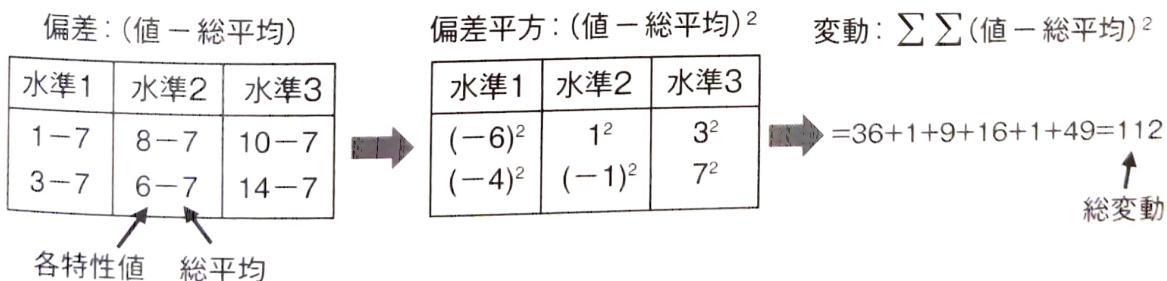


図 8.3 総変動の計算

### ②要因効果による変動（群間変動）

群間のバラツキである**群間変動**（inter-group variation；群を“級”と呼ぶこともあるので**級間変動**inter-class variationとも）は、要因効果の影響で生まれ

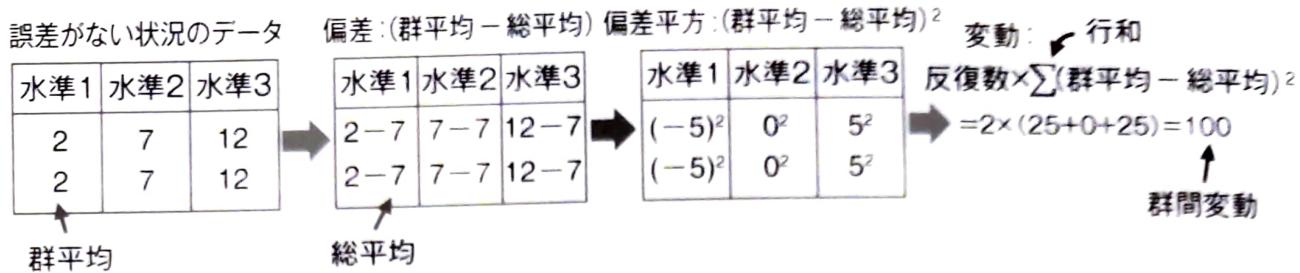


図 8.4 群間変動の計算

ます。この群間変動の捉え方ですが、群内の環境が統制された状況を考えれば明白です。もし、誤差要因が存在しない完璧な環境ならば、それぞれの群内の値は全て同じになるはずです（群内で値がバラついているのは誤差要因の影響だからです）。ですから、それぞれの群内の値を全て同じにして、そこから総平均を引けば群間変動の偏差が求まるというわけです。ただし、各水準の真の値（母平均）は誰にもわかりませんから、各群の標本平均で代用します。

よって図8.4のように、まず群内で誤差のない状況のデータ行列を想定し、群 $j$ の平均 $\bar{x}_j$ から総平均 $\bar{\bar{x}}$ を引いて偏差 $(\bar{x}_j - \bar{\bar{x}})$ を計算します。次に、それらを2乗して偏差平方 $(\bar{x}_j - \bar{\bar{x}})^2$ を計算し、全ての偏差平方を足し合わせた $2 \times \sum(\bar{x}_j - \bar{\bar{x}})^2$ が群間変動となります（この $\sum$ は行和なので反復数である2を乘じます）。小麦の事例では、群間変動は“100”となるはずです。

### ③誤差効果による変動（群内変動）

群内のバラツキである**群内変動** (within-group variation : 級内変動 within-class variation) は、誤差要因の影響で生まれます（誤差がなければ①で想定したように群内は全て同じ値となるでしょう）。ですから、誤差要因の効果は群内に生じた変動を足し合わせれば捉えることができます。

とはいえて変動は引き算ができるので、総変動（①で求めた112）から群間変動（②で求めた100）を引いた、残り“12”が小麦の事例の群内変動となります。実は、ソフトウェアも、この引き算のアルゴリズムで計算しているのですが、それでは勉強にならないので、ちゃんと計算してみましょう。

まず、群ごとに図8.5のように各特性値 $x$ から群平均 $\bar{x}_j$ を引いて偏差 $(x - \bar{x}_j)$ を計算して求めます。次に、その偏差を2乗して負の符号を消した偏差平方 $(x - \bar{x}_j)^2$ を計算します。最後に、全ての群の偏差平方を足し合わせれば、群内変動（偏差平方和） $\sum \sum (x - \bar{x}_j)^2$ が求まります。



図 8.5 群内変動の計算

## 8.3 分散分析におけるF検定

### ● 検定統計量の計算

群間変動（要因効果）と群内変動（誤差効果）が求まりましたが、これら2つの変動をそのまま割り算して比をとっても、検定統計量としては使えません。なぜなら変動と変動の比は何の確率分布にも従わないので、帰無仮説の下での確率を計算できないからです。

そこで、2つの変動をそれぞれの自由度で割って不偏分散にしましょう。不偏分散と不偏分散の比ならばF分布に従うため、検定に使えるからです。つまり、分散分析は等分散の検定と同様、F検定の1つなのです。早速、2つの変動を不偏分散にして、検定統計量となるそれらの比(F)を求めてみましょう。

さて、変動を自由度で割れば不偏分散になるのですが、分散分析の場合は自由度が少し面倒です。第3章の復習になりますが、自由度は不偏推定量の計算に不可欠な概念で、データ数から、計算に使った標本平均の数を引いた値です。前章までの統計量では、標本平均は1つしか使わなかったので、自由度はデータ数から1を引いたものばかりでしたが、分散分析の場合はそうはいきません。

それでは、自由度に注意して、要因効果で生じた群間変動の不偏分散（以降、**要因分散**と略します）から求めてみましょう。もう一度、図8.4を見てください。最初に、誤差がない状況のデータを想定しましたが、そこで使用した値は群平均で代用したものでしたので、データ数（値の種類）は水準（群）の数と同じ3となります。一方、偏差を計算するのに用いた標本平均は総平均の1つだけでした。よって、自由度（Fの分子になる第1自由度 $v_1$ ）は、水準数（3）から総平均数（1）を引いた“2”となります。先に求めた群間変動は100でしたので、要因分散はそれを2で割った“50”となります。

一方、誤差要因で生じた群内変動の不偏分散（以降、**誤差分散**と略しますが、第3章の標本平均の分散とは区別してください）についても、図8.5を再度見

てください。群ごとに変動を求めてそれを足し合わせていたので、全体のデータ数は3水準（群）×2反復（区画）で6。偏差の計算に使った標本平均は群平均でしたので、その数は水準（群）数の3となります。よって、自由度（第2自由度 $\nu_2$ ）は、水準数（3）×反復数（2）－水準数（3）で“3”となります。先に求めた群内変動は12でしたので、誤差分散は、それを3で割った“4”となります。

要因分散と誤差分散が求まつたら、それを割り算するだけです。従いまして、小麦の事例の検定統計量Fは、要因分散（50）÷誤差分散（4）で“12.5”となります。式に整理すると次のようになります。 $\nu$ は自由度、 $n$ は群あたりの反復数、 $j$ は水準（群）です。

### （対応のない）分散分析の検定統計量

$$F_{(\nu_1, \nu_2)} = \frac{\text{要因分散}}{\text{誤差分散}} = \frac{\frac{\text{群間変動}}{\text{自由度}}}{\frac{\text{群内変動}}{\text{自由度}}} = \frac{\frac{\text{反復数} \times \sum(\text{群平均} - \text{総平均})^2}{\text{水準数} - 1}}{\frac{\sum \sum (\text{値} - \text{群平均})^2}{\text{データ総数} - \text{水準数}}}$$

$$= \frac{\frac{n \sum (\bar{x}_j - \bar{\bar{x}})^2}{\nu_1}}{\frac{\sum \sum (x - \bar{x}_j)^2}{\nu_2}}$$

ここで1つ注意があります。分散分析の検定統計量は、要因分散を（たとえ誤差分散より小さくても）分子に持ってくるようにしてください。なぜならば、分散分析は要因効果の有無を検証するのが目的なので、等分散の検定のように分子と分母を入れ替えOKにしてしまうと（簡便化のために分散の大きい方を分子としました），結果の解釈が難しくなってしまうからです。常に要因分散を分子にしておけば、F値は誤差効果に対する要因効果の大きさと定義できるので、F分布の上側だけで要因効果の有無を判定すれば済むのです。

## ● 仮説の検定

いま一度、分散分析の検定統計量Fの内容を確認しておくと「要因効果の誤差効果に対する相対的な大きさ」です。ですので、図8.6の右のように、存在を検証したい要因効果があればFの値は大きくなります（対立仮説）。

逆に、図左のように、要因効果がなければFの値は小さくなります（帰無仮説）。とはいえる、要因効果が全くなくとも0に近づくほど小さくなることは滅多にありません。なぜならば、処理水準以外の群間環境を完全に統制すること

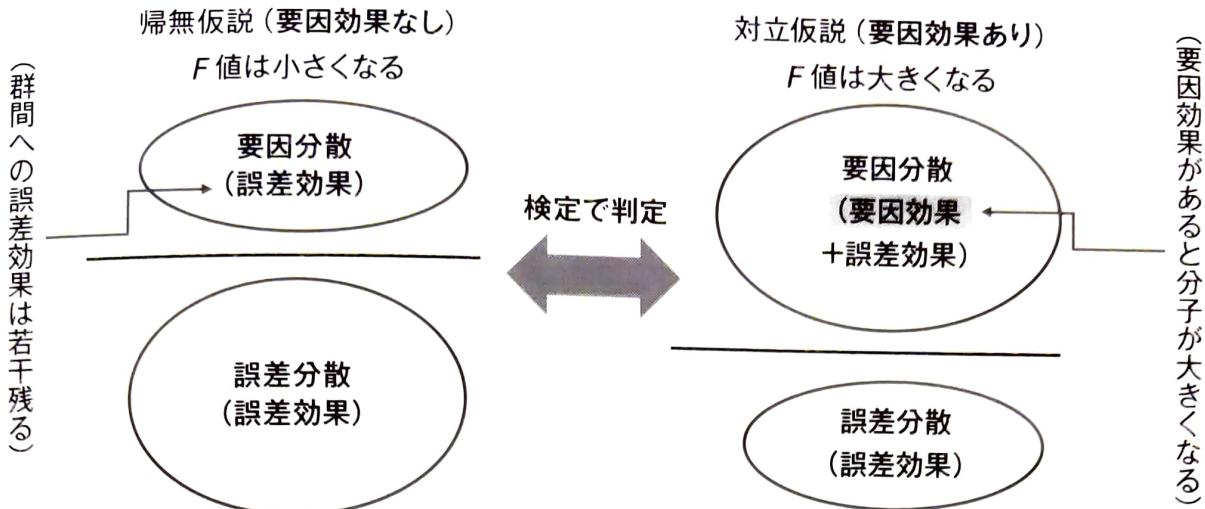


図 8.6 分散分析の検定統計量Fの性質（分数のポンチ絵）

はできないため、分子となる要因分散にも（群間変動を生じさせる様々な）誤差効果が残っているからです。

検定手順は、これまで学んできた検定と同じです。図8.7のように、帰無仮説の分布において、事前に決めておいた有意水準（上側だけで $\alpha$ とします）に対応する限界値（分布表から読み取ります）と、検定統計量Fを比較して、検定統計量の方が大きい場合に帰無仮説を棄却するのです（小さい場合には受容）。

事例におけるF検定の限界値を、上側確率が5%のF分布表（付録IV）から読み取ると、分子となる要因分散の第1自由度である2の列と、分母となる誤差分散の第2自由度である3の行がクロスする“9.55”であることがわかります。これと、先ほど計算した検定統計量Fの値である“12.5”を比較すると、検定統計量の方が大きいことから、「検定統計量は棄却域に入るので、有意水準5%で有意である」といえます。従って、「肥料の種類によって小麦の単収は変

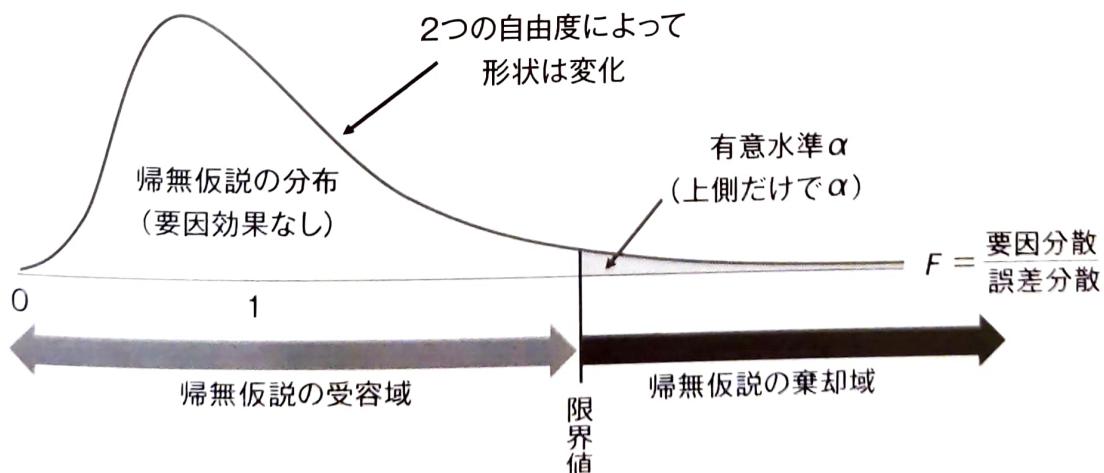


図 8.7 分散分析におけるF検定

化した（肥料という要因効果はあった）」ことになります。

「稲は土で穫り、麦は肥料で穫る」といわれるほど、小麦の栽培に肥料は欠かせません。ただし、小麦は冷涼で乾燥した気候を好むため、ヨーロッパの8t/haに比べて、日本では北海道でさえ4t/haと半分程度の単収にとどまるようです。

## ● ソフトウェアによる一元配置分散分析と検出力分析

Excelの分析ツールでは〔分散分析：一元配置〕が、（対応のない）一元配置分散分析になります。事例を分析してみると、次のような分散分析表が出力されると思います。「グループ」とは群のこと、「観測された分散比」が検定統計量F、「F境界値」が限界値のことです。ここまで手計算でやってきた内容と同じになっていることが確認できましたでしょうか。なお、ソフトウェアを用いると、p値“0.035”も計算してくれるのがありがたいですね（有意水準とp値を比較して判定する方法でも構いません）。

分散分析：一元配置

変動要因	変動	自由度	分散	観測された分散比	P-値	F境界値	
グループ間	100	2	50		12.5	0.03507	
グループ内	12	3	4	↑ 検定統計量 F		↑ 限界値	
合計	112	5					

さらに無料ソフトG\*powerの事後(Post hoc)分析を使って、検出力を計算してみましょう([F tests] → [ANOVA: Fixed effects, omnibus, one-way])。

分散分析の効果量にもいろいろな計算式があるのですが、ここでは、 $\sqrt{F \times (\nu_1 \div \nu_2)}$ で推定される効果量“2.89”を用います(Determineのウインドウに群間変動100と群内変動12を入力してもOK)。[Total sample size]に3群合わせた反復数の6と、[Number of groups]に群数の3を入れて[Calculate]を押せば、検出力(Power)は“0.93”と出力されます。総標本サイズが6と小さいにもかかわらず検出力が高かったのは、推定される効果量が極めて大きかったためです(なにせ著者が演習用に創作したデータですからね……)。

## ● 正規性と等分散性 (Rコマンダー)

t検定同様、分散分析においても、各群(水準)の母集団は、それぞれ正規性と等分散性を満たしていないなりません。正規性については、特性が量的尺度で、反復数がそこそこあれば(事例の2は少なすぎますが……)気にす

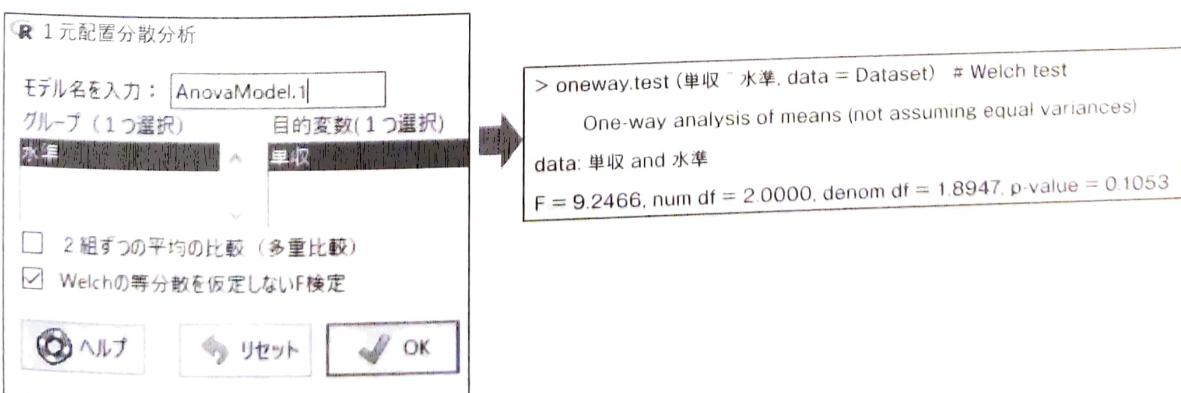
る必要はありません。ただし、やはり等分散性については、気にかけた方がよいでしょう。

異分散への対応方法もt検定と同じです。分散分析にも「ウェルチの検定」があるので、群間で反復数や分散が大きく（1.5～2倍以上）異なる場合、あるいは等分散性の検定を実施して、その結果をみて（等分散という帰無仮説が棄却されたら）ウェルチの検定を用いるようにすればよいでしょう。もちろん、何も考えずに最初からウェルチの検定を用いるのも“あり”です（むしろ近年はその方向にあります）。

ただし、残念ながら、Excelの分析ツールには、多群用のウェルチの検定も等分散の検定も搭載されていませんので、別途、高機能なソフトウェアを用いる必要があります。例えば無料のRコマンダー（インストールについては本書冒頭の「ソフトウェアについて」をお読みください）ならば、どちらも実施できますので紹介しておきましょう。

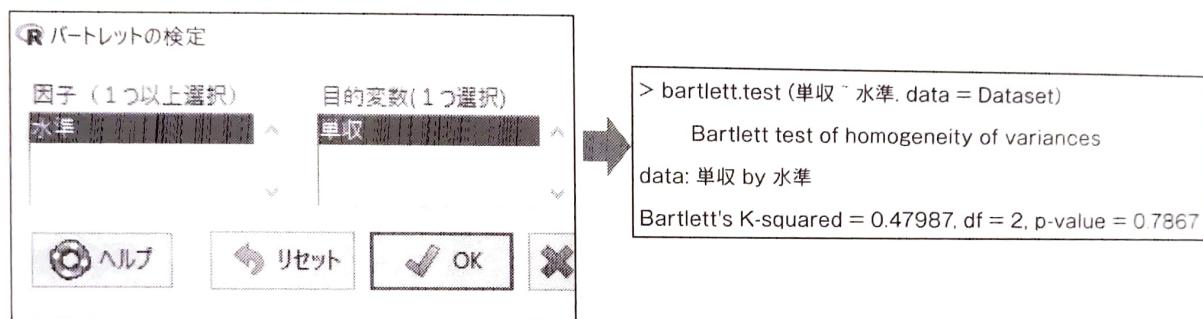
まず、R Console上で>記号の右に「library(Rcmdr)」と入力してRコマンダーを起動します。Rコマンダーのメニューから、[データ] → [データセットのロード]で、オーム社のWebページにアップされているデータファイル「対応のない一元配置分散分析（肥料と小麦）.RData」を読み込みます（小麦の事例が入っています）。もしくは[データ] → [新しいデータセット]でデータエディタを起動し、直接データを入力していただいても結構です（ただし、[アクティブデータセット内の編集の管理] → [数値変数を因子に変換]で適当な水準名を指定するようにしてください）。

ウェルチの検定は、[統計量] → [平均] → [1元配置分散分析]の中に、[Welchの等分散を仮定しないF検定]が選択できるようになっていますので、して実施するだけです（下図）。試しに小麦の事例で実施してみると、p値は0.1を超えており、5%水準では有意となりませんでした（ウェルチの検定は普通の分散分析よりも保守的な結果になります）。



予備的に等分散性を検証したい場合には、[統計量] → [分散] の中に、**バートレットの検定** (Bartlett's test) と**ルビーンの検定** (Levene's test) という2種類の手法があります。このうち基本的なのは、バートレットの検定で、分散の偏り度が $\chi^2$ 分布に従うことを利用して検定します（偏りが大きければ異分散となります）。ただし、正規性が満たされている必要があるので、それが前提できない場合にはルビーンの検定を用いることになります。

バートレットの検定を小麦の事例で実施してみると（小麦の事例は群内変動の偏差が同じ群があるのでルビーン検定は適していません）、下図のようにp値は0.79となり、等分散という帰無仮説は棄却されなかったので、（ウェルチの検定ではなく）普通の分散分析で問題はなさそうです。



## 8.4 対応のある一元配置分散分析

### ● 個人差や個体差を考慮する

ここまで、対応のない一元配置分散分析について説明してきましたが、**対応のある一元配置分散分析** (one-way repeated-measures ANOVA) では何が変わるのでしょうか？ 対応のある場合とは、対象となる個人や個体が、群間で対応関係にあることです。つまり、同じ人や個体に対して処理水準ごとに繰り返し測定 (repeated-measures) されている状況です。反復測定とも呼ばれますが、あくまで異なる処理水準の実験・測定が同じ人や個体で反復されることがあります。同水準内で独立した実験を反復させること (207ページ) とは区別しておいてください。なお、対象となるのは個人や個体だけでなく、場所の区画や工場のラインなども該当するので、より抽象的に、「塊」という意味でブロック (block) と呼ばれることが多いです。

こちらも早速、実際に値の入った事例で説明しましょう。表8.3は、禁煙外来における追加的なカウンセリングが、喫煙本数を変化させるかどうかを検証するため実施した実験の仮想データです。この実験では、喫煙者2名に、カ

表 8.3 カウンセリングと喫煙本数 (本/日)

	相談前 (水準1)	1回目 (水準2)	2回目 (水準3)	被験者平均	被験者間変動 (個人差)
Aさん	22 ↔ 14	14 ↔ 9		$\bar{x}_A = 15$	
Bさん	14 ↔ 8	8 ↔ 5		$\bar{x}_B = 9$	
	$\bar{x}_1 = 18$	$\bar{x}_2 = 11$	$\bar{x}_3 = 7$	$\bar{x} = 12$	

カウンセリングを受ける前、1回目に受けた後、2回目に受けた後の喫煙本数/日をそれぞれ記録してもらっています。つまり被験者ごとに3回の繰り返し測定をしています。カウンセリング前の群の平均  $\bar{x}_1$  は18本でしたが、1回目後の群の平均  $\bar{x}_2$  は11本、2回目後の群平均は  $\bar{x}_3$  は7本となっており、一見すると減っているようですが、本当にカウンセリング（要因）は禁煙（減煙）に効果があるといっても良いでしょうか？ 対応のある一元配置分散分析で、母平均にも差があるかどうかを確かめてみましょう。

さて、この表を対応のない一元配置分散分析の小麦の事例（表8.1）と見比べてください。最右列の被験者平均  $\bar{x}_A$  と  $\bar{x}_B$ （グレー部分）が求められていることに気がつきましたでしょうか。これら各被験者平均の値はそれぞれ異なります。つまりバラつくのですが、このバラツキこそが個人差なのです（群間方向にも個人差はあるでしょうが、要因効果と一緒になってしまっているので捉えるのはあきらめます）。

このような個人差という要因によって発生する効果を**被験者間変動**（あるいは**標本間変動**；inter-subject variation）と呼びます。実験する立場から見れば、個人差は誤差の一部です。なぜならば、この実験の目的は「カウンセリングは汎用的に禁煙（減煙）の効果がある」ことを検証したいのであって、「カウンセリングの効果には個人差がある」と結論づけたいわけではないからです。つまり、対応のあるデータを用いるということは、図8.8のように、対応のない場合には誤差として一括して扱っていた群内変動から、個人差による変動を分離できるということなのです。そして、個人差が除去された新しい群内変動を検定統計量  $F$  の分母とすれば（分子は変わりません）、個人差が考慮された精度の高い検定となるというわけです。ですから、人間や動植物など、個人差や個体差が大きいと思われる標本を対象とした実験では、できる限り対応関係のあるデータがとれるように計画し、対応のある分散分析を実施するとよいでしょう。

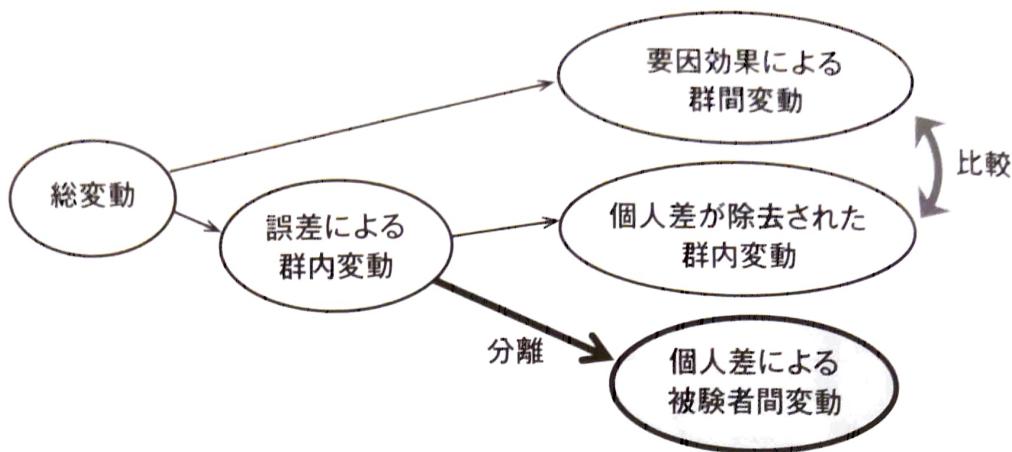


図 8.8 対応のある一元配置分散分析の考え方

## ● 個人差による変動の計算

それでは、実際に禁煙の事例を使って、個人差によって生じる被験者間変動を求めてみましょう。なお、図8.8をもう一度見ていただければわかるように、群間変動は対応のない場合と全く同じです（個人差があろうがなかろうが要因効果自体に変化はありません）。あくまで個人差は誤差ですので、群内変動の求め方が変わるだけです。なお、事例の群間変動は“124”、（個人差を分離する前の）群内変動は“58”となるはずです（後で使いますので計算しておいてください）。

被験者間変動は被験者平均のバラツキ（偏差平方和）のことですから、図8.9のように、被験者*i*の平均 $\bar{x}_i$ から総平均 $\bar{\bar{x}}$ を引いた偏差 $(\bar{x}_i - \bar{\bar{x}})$ を全ての群に入れます。もちろん、水準によって本当の（母集団の）個人差の大きさは異なるでしょうが、誰にもわからないので平均で代用するのです。次に、それを2乗した偏差平方 $(\bar{x}_i - \bar{\bar{x}})^2$ を求め、全て足し合わせた偏差平方和 $3 \times \sum(\bar{x}_i - \bar{\bar{x}})^2$ が被験者間変動となります（この $\sum$ は列和なので群数3を乗じます）。事例では、被験者間変動は“54”となるはずです。

なお、不偏分散の計算で必要になる自由度は、全ての群で同じ値を使っているので、被験者平均数（2）から総平均数（1）を引いた“1”となります。

偏差:(被験者平均 - 総平均)			偏差平方:(被験者平均 - 総平均) <sup>2</sup>			変動: $\sum(\bar{x}_i - \bar{\bar{x}})^2$ = 群数 $\times \sum(\bar{x}_i - \bar{\bar{x}})^2$ = $3 \times (9 + 9) = 54$
水準1	水準2	水準3	水準1	水準2	水準3	
15-12	15-12	15-12	3 <sup>2</sup>	3 <sup>2</sup>	3 <sup>2</sup>	
9-12	9-12	9-12	(-3) <sup>2</sup>	(-3) <sup>2</sup>	(-3) <sup>2</sup>	

被験者平均 総平均

被験者間変動

図 8.9 被験者間変動の計算

## ● 検定統計量の計算

対応のある場合の検定統計量  $F$  も、対応のない場合と同様、要因分散の誤差分散に対する大きさ（比）です。ただし、分母の誤差分散が、（個人差が分離された）新しい群内変動から求められている点が異なります。

変動は引き算ができるので、個人差を分離する前の群内変動（58）から、先ほど求めた被験者間変動（54）を引いた“4”が、個人差が除去された新しい群内変動になります。この4を自由度で割れば  $F$  の分母になるのですが、自由度はどのようになるでしょうか。

実は自由度も引き算になります。個人差を分離する前の群内変動に対する自由度は、水準数（3）×被験者数（2）－水準数（3）なので“3”でした。そこから被験者間変動に対する自由度（1）を引いた“2”が、新しい誤差分散の自由度（第2自由度  $\nu_2$ ）になります。従いまして、個人差が分離された新しい誤差分散（ $F$  の分母）は、 $4 \div 2$  で“2”となります。一方、 $F$  の分子になる要因分散は、対応のない場合と同じ求め方ですので、群間変動（124）を第1自由度  $\nu_1$ （2）で割った“62”となります。

よって、禁煙の事例の、対応のある一元配置分散分析の検定統計量  $F$  は、要因分散（62）÷新しい誤差分散（2）で“31”となります。式に整理すると次のようにになります。 $\nu$  は自由度、 $n$  は被験者（反復）数、 $k$  は水準（群）数、 $i$  は被験者、 $j$  は水準（群）です。

### 対応のある一元配置分散分析の検定統計量

$$\begin{aligned}
 F_{(\nu_1, \nu_2)} &= \frac{\text{要因分散}}{\text{個人差が除去された誤差分散}} = \frac{\overline{\text{群間変動}}}{\overline{\text{自由度}}} \\
 &= \frac{\frac{\text{被験者数} \times \sum (\text{群平均} - \text{総平均})^2}{\text{水準数} - 1}}{\frac{\sum \sum (\text{値} - \text{群平均})^2 - \text{水準数} \times \sum (\text{被験者平均} - \text{総平均})^2}{(\text{データ総数} - \text{水準数}) - (\text{被験者数} - 1)}} \\
 &= \frac{\frac{n \sum (\bar{x}_j - \bar{\bar{x}})^2}{\nu_1}}{\frac{\sum \sum (x - \bar{x}_j)^2 - k \times \sum (\bar{x}_i - \bar{\bar{x}})^2}{\nu_2}}
 \end{aligned}$$

早速、事例を検定してみましょう。分散分析は片側検定ですから、上側確率だけで有意水準 $\alpha$ とします。ですから5%のF分布表（付録IV）で第1自由度である2の列と、第2自由度である2の行のクロスする“19”がそのまま限界値となります。検定統計量F(31)は、有意水準5%の限界値(19)よりも大きいので「有意である」、つまりカウンセリングの禁煙（減煙）効果は「あった」といえます。

さて、著者は非喫煙者なので禁煙の難しさはわからないのですが、実際には事例のように徐々に本数を減らしていくよりも、きっぱりと喫煙を止める方が禁煙を継続しやすく、かつ効果的であることが、海外の大学の研究チームによって検証されているようです。

## ● ソフトウェアによる分析と検出力について

Excelの分析ツールには、対応のある一元配置分散分析はありません。しかし、[分散分析：繰り返しのない二元配置]で代用することができます。というのも、個人差も1つの要因と考えれば二元配置となるからです。なお、「繰り返しのない」とは、行要因と列要因の各水準の組合せでデータが1つしかないことを意味しています。事例では、被験者Aの水準1のデータは“22”のみでしたね。このあたりがわかりにくいのですが、この後、二元配置の事例で「繰り返しのある」場合を学べば区別できるようになるでしょう。

禁煙の事例を分析してみると、次のような分散分析表（上表）が出力されると思います。行が個人差、列が要因（カウンセリング）効果となります。このように、二元配置分散分析なので個人差（行）の検定結果も出力されているの

分散分析：繰り返しのない二元配置

変動要因	変動	自由度	分散	観測された分散比	P-値	F境界値
行	54	1	54	27	0.035	18.51
列	124	2	62	31	0.031	19
誤差	4	2	2			
合計	182	5				

個人差を検定すると有意になる

(参考：対応のない一元配置分散分析の結果)

変動要因	変動	自由度	分散	観測された分散比	P-値	F境界値
グループ間	124	2	62	3.21	0.180	9.55
グループ内	58	3	19.33			
合計	182	5				

分母（誤差）に個人差が含まれているので  
Fが小さくなり、有意とならない

ですが（ちなみに有意差はあるようですね）、そちらは無視して要因効果（列）のみ確認すれば良いでしょう。試しに、対応のない一元配置の分散分析表（174ページの下表）を比較してみると、個人差が誤差に残っているため、検定統計量 $F$ は3.21と小さくなつて有意となりませんでした。このことからも、個人差が大きいと思われる場合には、対応のあるデータがとれる実験を心がけた方が良いことがわかります（もちろん個人差が小さい場合には、その限りではありません）。なお、Rコマンダーにも【1元配置反復測定ANOVA/ANCOVA】があるのでですが、大変使いにくないので本書では紹介しません。

検出力については、一般に、対応のない一元配置よりも上位の難しいモデルでは計算しません。一応、G\*powerですと【ANOVA: Repeated measures, within factors】が“対応のある”○元配置分析に該当する（要因内で反復測定の意味で繰り返しはなし）のですが、効果量・有意水準・標本サイズ以外にも、群間の相関係数や球面性（各群間の差の分散）など、数多くの未知のパラメータを入力する必要があるため、結局、不正確な推定となってしまうからです。

## 8.5 (対応のない) 二元配置分散分析

### ◎ 交互作用を捉える

最後に本節では、（実験で効果の存在を確認したい）要因が2つある二元配置を扱います。二元配置でも、それぞれの要因において対応関係がある場合となる場合があるのですが、紙幅も限られているので、本書では**対応のない二元配置分散分析**（two-way ANOVA）で終わりにしたいと思います。なお、二元以上の分散分析をまとめて**多元配置分散分析**と呼ぶことがあります。

みなさん方にしてみれば、要因が2つならば一元配置分散分析を2回実施すれば良いだろうと思われるでしょうが、実は二元配置分散分析でしか得られない大きなメリットがあります。それは**交互作用**（interaction）の有無を検証できるということです。交互作用とは、複数（ここでは2つ）の要因間で、ある特定の水準同士が組み合わさったときにだけに現れる相乗効果（特性値が大きくなること）や相殺効果（小さくなること）のことです。よって、二元配置分散分析で検証できる要因効果は、もともとの2つの要因の効果（**主効果**：main effect）に、交互作用が加わるので合わせて3つになります。

いつもの変動の分解のポンチ絵で示しますと図8.10のようになります。対応のない一元配置分散分析のときと比べると、1つだった要因変動が3つに分解され、それぞれが誤差変動と比較される点が異なります。つまり、3つの検

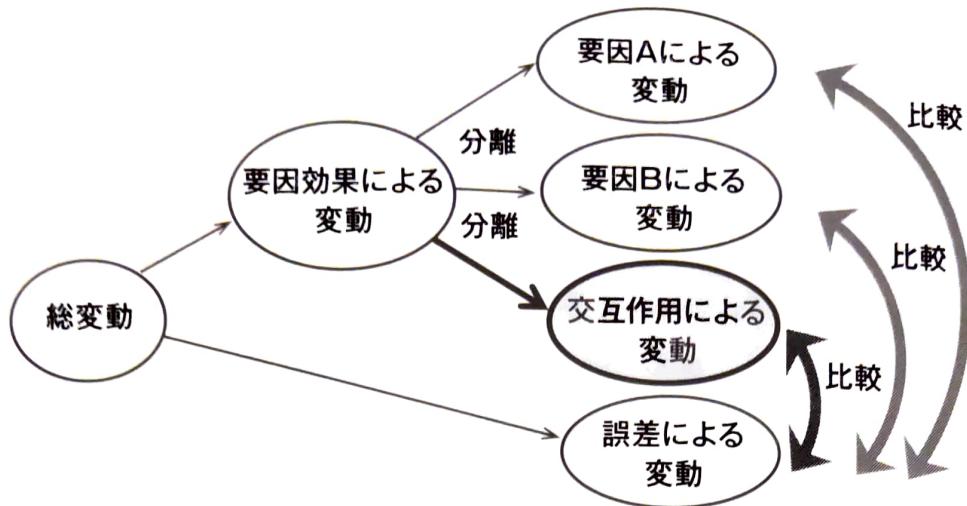


図 8.10 対応のない二元配置分散分析の考え方

定が一度に実施されるわけです。

二元配置分散分析についても、事例を使って説明していきましょう。表8.4は、アミラーゼという、デンブンを分解して糖にする（唾液などに含まれる）酵素を使った仮想実験です。反応速度（特性；グレー部分）に影響を与えると思われるpH（要因A）と温度（要因B）を2水準ずつ設定し、各水準の組合せごとに独立した実験を2回繰り返しています。もし、水準の組合せ内で実験を繰り返さない場合は、「繰り返しのない二元配置分散分析」となりますので、前節と同じ対応のある一元配置と同じ状態になります（個人差も要因と考えれば二元配置）。その場合、交互作用があっても、誤差に含まれたまま分離できなくなるため検証できません。つまり、前節の対応のある一元配置分析では、存在を検証したい要因と個人差との間に交互作用がないと仮定していたのです。なお、前節までは $x$ や $\bar{x}_1$ ,  $\bar{x}_A$ ,  $\bar{\bar{x}}$ などの記号を使いましたが、二元配置分散では式が複雑になりすぎるので止めておきます。

表 8.4 アミラーゼの反応速度  
温度(要因B)

pH (要因A)	10°C (水準1)		40°C (水準2)	行平均
	4.0 (水準1)	2	10	
7.0 (水準2)	6	22	15	
	8	24		
列平均	4	16	10	

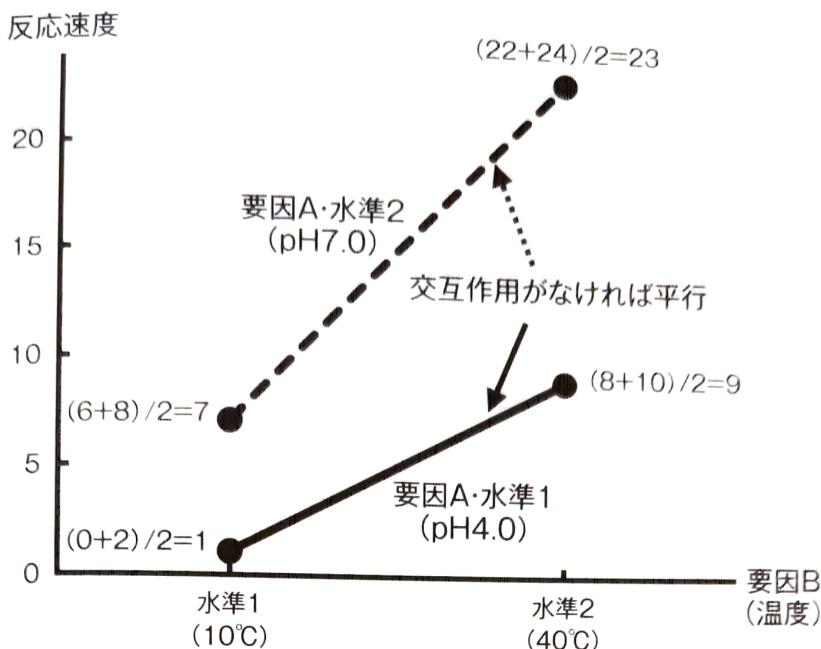


図 8.11 事例の交互作用の確認

この事例を線グラフにしたのが図8.11です。特性（反応速度）を縦軸、要因B（温度）の水準を横軸にして、要因Aの水準1（以降A1と表現します）を実線、A2を破線で表しています（4つの線端●の値は水準組合せの平均）。

これを見ると、要因A、Bともに水準1よりも水準2の（平均）反応速度が速くなっているため、どちらの主効果も存在しているようです。

また、B1のときのA1とA2の間隔よりも、B2のときのA1とA2の間隔の方が広くなっているため、交互作用（以降A×Bと表現します）の存在も十分予想できます。なぜならば、交互作用がなければ2本の線は平行になるはずだからです。

## ● 変動の計算

さて、変動の計算に入りますが、基本的な考え方としては、図8.10を再度見ていただければわかるように、まずは3つの要因変動が全て含まれる群間変動と、誤差による群内変動に分離し、その後、群間変動をそれぞれの要因変動に分離します。

しかし、変動は引き算ができるので、実際には（ソフトウェアなどでは）図8.12のように、まず総変動を計算して、次に要因Aによる変動を計算します。そして総変動から要因Aによる変動を引いて、残った変動が要因Bと交互作用、そして誤差による変動と考えます。これを次々に繰り返し、最後に残るの用、そして誤差による変動と考えます。これを次々に繰り返し、最後に残るのが誤差による変動となります。第10章で出てくる「直交配列表」を理解するのが誤差による変動となります。第10章で出てくる「直交配列表」を理解するのも役立つので、本節では後者（引き算）の方法で計算しましょう。

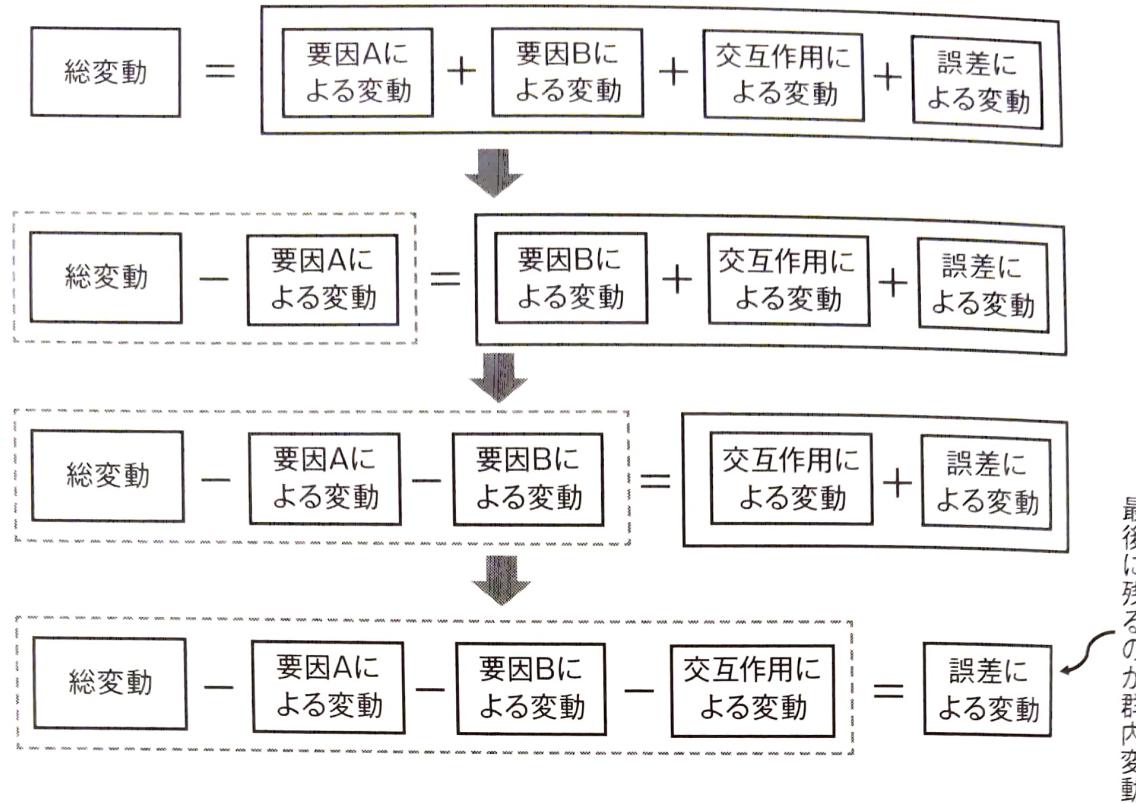


図 8.12 實際の変動の計算（対応のない二元配置）

## ①総変動

酵素の事例で総変動から計算してみましょう。総変動は、ここまで分散分析と同じようにするだけです。具体的には次ページの図8.13の左の表①のように、8個のデータ（特性値；反応速度）それぞれにおいて総平均（10）を引いて偏差を計算し、それを2乗した偏差平方を、全て足し合わせるだけです。そうすると、事例の総変動は“528”となるはずです。

なお、総変動の不偏分散は検定に使いませんが、自由度も要因間で引き算するので計算しておきましょう。総変動の自由度は、データ数（8）から変動の計算に使った標本平均（総平均）の数（1）を引いた“7”となります。

## ②要因Aによる変動

もし、値がバラつく原因が要因Aの主効果によるものだけならば、要因Aの各水準（行）内の4つの値は全て同じになるはずです。ですから、図8.13の真ん中の表②のように、要因Aの各水準（行）内において、行平均（水準1が5、水準2が15）から総平均（10）を引いた偏差を2乗して全て足し合せた“200”が、要因Aの効果によって生じた変動となります。自由度は、水準数（2）から総平均数（1）を引いた“1”です。

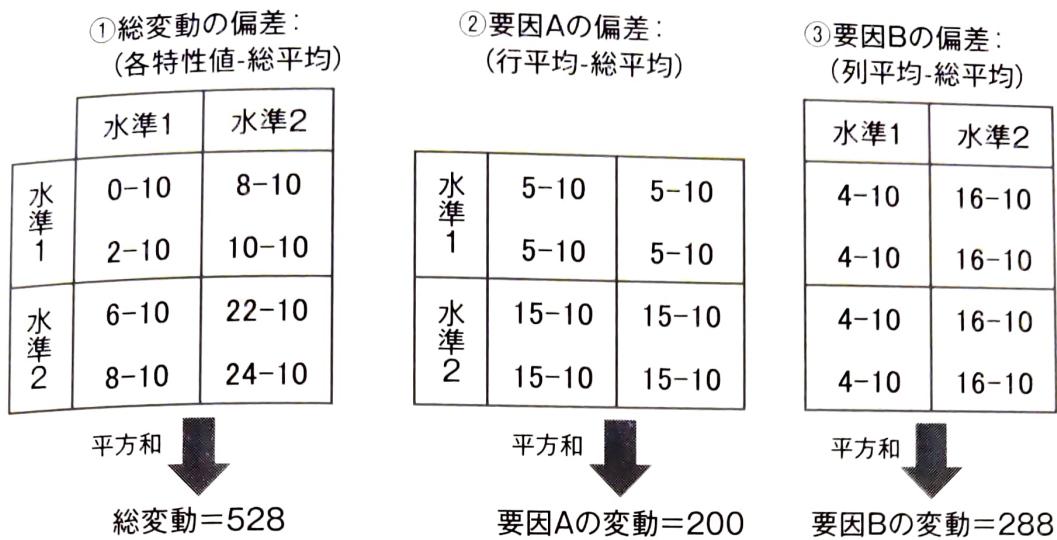


図 8.13 総変動と主効果変動の計算

### ③要因Bによる変動

こちらも、値がバラつく原因が要因Bの主効果によるものだけならば、要因Bの各水準（列）内の4つの値は全て同じになるはずです。ですから、図8.13の右の表③のように、要因Bの各水準（列）内において、列平均（水準1が4、水準2が16）から総平均（10）を引いた偏差を2乗して全て足し合わせた“288”が、要因Aの効果によって生じた変動となります。こちらの自由度も、水準数（2）から総平均数（1）を引いた“1”です。

### ④交互作用A×Bによる変動

要因Aと要因Bの交互作用A×Bによって生じた変動の計算は少し面倒です。図8.11のところで触れたように、もし交互作用がなかったら2本線は平行になるはずです。つまり、この2本線の平行からのズレが交互作用の大きさということになります。ということは、どちらかの要因の各水準を基準として、もう一方の要因水準の特性値の離れ具合を捉えれば良いのです。

次ページの図8.14は、Bの各水準を基準とした場合の交互作用の変動の計算手順を示したものです。左上のA1・B1のセルでは、(A1・B1の平均-A1の平均)-(B1の平均-総平均)で、B1におけるA1の作用の偏差を計算しています。同様に、右上でB2におけるA1の作用を、左下でB1におけるA2の作用を、右下でB2におけるA2の作用を捉えます。4セル全てにおいて、これら偏差を2乗して足し合わせれば変動ですので、事例では“32”が交互作用の大きさ（変動）となります。なお、交互作用A×Bの自由度は、要因Aの自由度（1）と要因Bの自由度（1）を乗じた“1”になります。

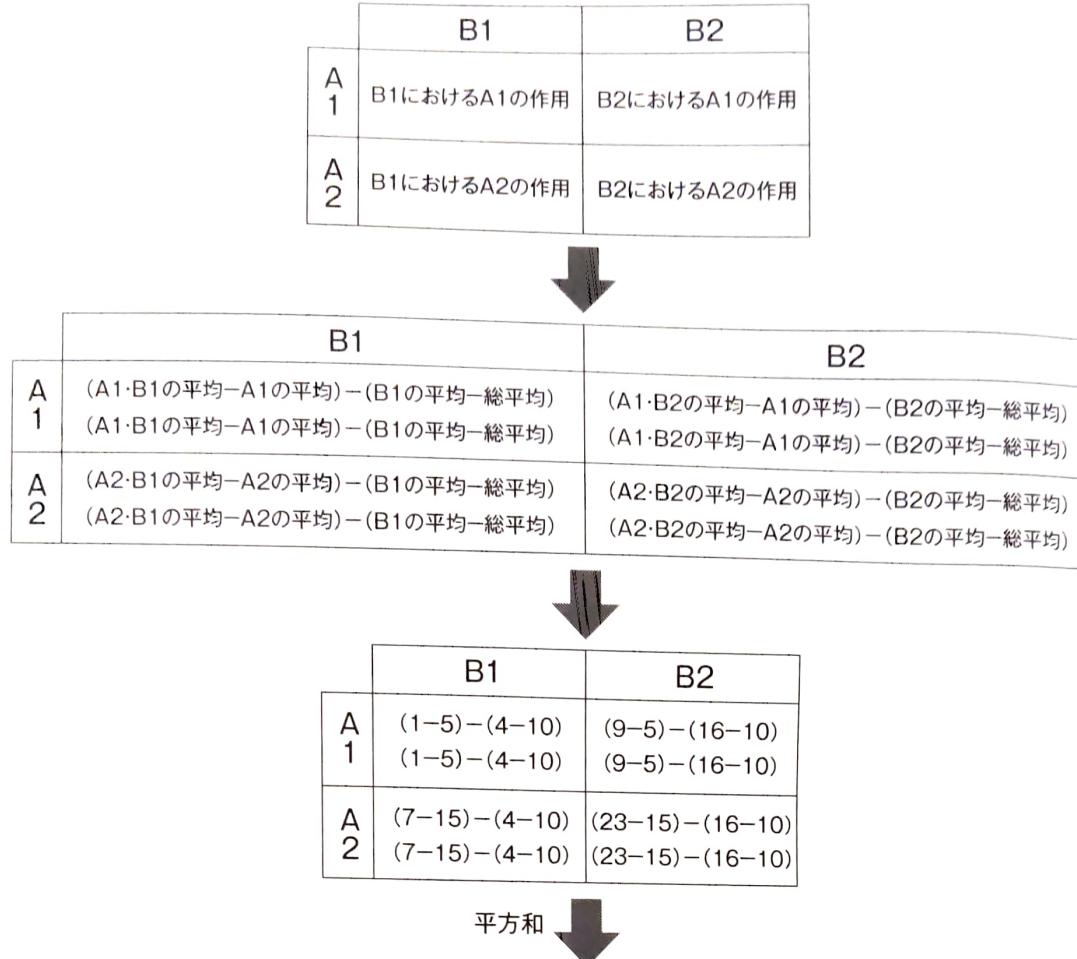


図 8.14 交互作用による変動の計算

## ⑤誤差による変動

誤差による変動は、図8.12の一番下のように、総変動からここまで求めてきた3つの要因（要因A、要因B、交互作用）の変動を引くだけです。事例では、「総変動（528）」－「要因A：pHの変動（200）」－「要因B：温度の変動（288）」－「交互作用：pH × 温度の変動（32）」で“8”が誤差変動となります。自由度も引き算で求めますので、総変動（7）－要因A（1）－要因B（1）－交互作用（1）で“4”となります。

## ● 検定統計量の計算と判定

二元配置分散分析では、要因Aと要因Bと交互作用という3つの効果を検定します。

再度、図8.10を見ていただければわかりますように、検定統計量Fの分母となる誤差分散は全て同じで、誤差による変動（8）÷自由度（4）で“2”となります。

一方、分子の不偏分散は、3つの要因それぞれを計算します。

まず、要因A（pH）は、変動（200）÷自由度（1）で“200”となるので、検定統計量Fは $200 \div 2$ で“100”となります。

次に、要因B（温度）は、変動（288）÷自由度（1）で“288”となるので、検定統計量Fは $288 \div 2$ で“144”となります。

最後に、交互作用A×B（pH×温度）は、変動（32）÷自由度（1）で“32”となるので、検定統計量Fは $32 \div 2$ で“16”となります。

これら3つの検定統計量が求まつたら、それぞれ限界値と比較して、帰無仮説が棄却できるか否かを判定します。ただし、事例では、どちらの要因とも2水準だったので分子の第1自由度は（交互作用も含め）全て1となるため（分母は誤差分散なので第2自由度は全て4），限界値をF分布表から読み取るのは1度で済みます。

今回も有意水準 $\alpha$ は5%で検定してみましょう。つまり、上側確率5%のF分布表（付録IV）で、1の列と4の行がクロスする“7.71”が限界値となります。

その結果、いずれの検定統計量Fも限界値より大きいため、要因Aの主効果も要因Bの主効果も、それらの交互作用A×Bの効果も有意であるといえます。

## ● ソフトウェアによる分析

Excelの分析ツールでは、[分散分析：繰り返しのある二元配置]が、対応のない二元配置分散分析になります。本事例で分析ツールを実施すると、次のような結果になります。表の「標本」は要因A（pH）のことを、「列」は要因B（温度）のことを指しています。2つの主効果と交互作用が、いずれも有意となっていることが確認できます。なお、Rコマンダーでは[多元配置分散分析]で二元以上の配置を扱うことができます。

分散分析：繰り返しのある二元配置

変動要因	変動	自由度	分散	観測された分散比	P-値	F境界値
標本	200	1	200	100	0.0006	7.71
列	288	1	288	144	0.0003	7.71
交互作用	32	1	32	16	0.0161	7.71
繰り返し誤差	8	4	2			
合計	528	7				

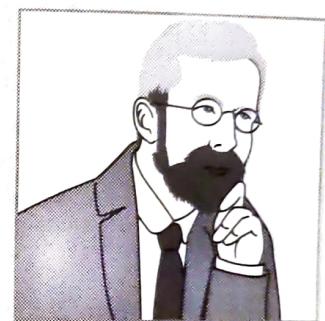
ところで、事例では、各組合せとも繰り返し数は2でそろっていましたが、これが異なっていると、2つの要因は直交している（相関なし）という、二元配置分散分析の前提条件が崩れてしまいます。高度なソフトウェアですと、「平方和のタイプ」というオプションがあり、繰り返し数が異なっていても相関を修正してくれます。これ以上の解説は止めておきますが、Type I以外を選んでおけば問題ありません（分析ツールには、このオプションはありませんが、RコマンダーはType IIで固定されています）。

## トピックス⑥

### 実験計画法と分散分析の歴史 —フィッシャー—

実験計画法は、推測統計学の父と呼ばれるロナルド・フィッシャーによって初めて提案されました。フィッシャーは1919年にイギリスの農事試験場に就職しましたが、彼がそこで目のあたりにしたのは、90年間にわたって観測されたままの試験データの山でした。その頃、農学分野では、実験をすること自体が目的となっていました。しかも実験の内容は問題のあるものばかりでした。そこでフィッシャーは、分析に適したデータを得るためにどのような実験を実施すべきかというルール集を一冊の本（『実験計画法』）にまとめ上げ、現在の科学的実験の基礎<sup>いしざえ</sup>を築いたのです。

そして分散分析もその頃フィッシャーによって考え出されました。試験場で蓄積されていたデータの内容は、異なる肥料と麦類の収量の関係、そして気象に関するものでした。フィッシャーは、こうしたデータを前に試行錯誤を繰り返し、収量のバラツキ（変動）を施肥などの要因に起因するものと誤差に起因するものに分解し、F検定を用いることで施肥効果の有無を判定できることを思いついたのです。この手法は、開発されてから100年以上経った現代においても、農学をはじめとした自然科学分野における分析ツールの主役となっています。このことからも、いかに優れた発想であったのかがわかります。



R.A. Fisher  
(1890~1962)

## 章末問題

**問1** 次のデータは、ある年の営農類型別農業所得です。具体的には、水田作経営、果樹作経営、施設野菜作経営、それぞれにおいて経営体（農家）2件の農業所得（農業粗収益 - 農業経営費）を整理したものです（仮想データ）。

このデータから、営農類型によって農業所得が異なることを検定しなさい（有意水準 $\alpha$ は5%とします）。

営農類型別農業所得(万円/経営体)

水田	果樹	施設野菜
80	200	600
40	300	400

**問2** 次のデータは、人種と血圧との関係を見るため実施した実験結果です（仮想データ）。アジア人、黒人、白人12名ずつを対象に、半数（6名）の人には1週間、塩分控えめの食事をとってもらい、もう半数には塩をふんだんに使った食事をとってもらいました。このデータから血圧に影響を与えていると統計的（有意水準5%）にいえる要因を探し出しなさい。

人種と血圧の関係（収縮期血圧、mmHG）

	アジア人	黒人	白人
塩負荷なし	128	110	127
	132	131	119
	118	130	140
	142	106	121
	130	100	106
	122	127	118
塩負荷あり	133	159	120
	140	160	141
	100	148	135
	150	178	109
	104	150	98
	111	138	129

問3 以下の図は、図8.11と同じように描いた二元配置実験の結果です。このうち、A1・B1に対してA2・B2が相乗効果があるものと相殺効果があるものをそれぞれ挙げなさい。また、要因Aの主効果のみあるものはどれか。

