読めば必ずわかる分散分析の基礎

第 2 版 2003 年 12 月 5 日

小野 滋

この解説書は，分散分析の基礎について，

可能な限りわかりやすく，かつ詳しく

説明することを目的としています。

簡潔さは犠牲にし，長くてくどいかわりに，

読めばわからずにはいられない

説明を目指したいと思います。

なお，説明中に用いる記号は，後藤ほか (編)「心理学マニュアル 要因計画法」(北大路書房) に準じています。

目次

第 I 部

はじめに

1 予備知識

　この解説書では，全くの初学者を念頭において，できるかぎり易しい説明を試みます。それでも，説明の都合上，データ解析と実験研究について，ある程度の知識が必要です。

　そこで，読み進めるのにどうしても必要だと思われる予備知識を，17 項目にまとめてみました。以下のリストに目を通して，もし理解できない箇所があったら，その箇所を復習してから，先に進んで下さい。

■量的データの記述

1.1 量的データの全体的な大きさをあらわす指標として，**平均**が用いられることが多い。データ $x_1, x_2, \ldots, x_n$ の平均 $\bar{x}$(「エックス・バー」) は,

$$\bar{x} = \frac{1}{n}\sum_{i=1}^{n} x_i$$

として求められる。

1.2 量的データのばらつきをあらわす指標として，**分散**と**標準偏差**(SD ともいう) が用いられることが多い。データ $x_1, x_2, \ldots, x_n$ の 分散 s^2 は,

$$s^2 = \frac{1}{n}\sum_{i=1}^{n}(x_i - \bar{x})^2$$

として求められる。また標準偏差 s は,

$$s = \sqrt{s^2}$$

として求められる。

■母集団と標本

2.1 ある変量について，分析者が関心を持っている値の全体を，**母集団** と呼ぶ。

2.2 いっぽう，手元にあるデータの集まりを，**標本** と呼ぶ。標本のなかに含まれている値の数を，**標本のサイズ**と呼ぶ。

2.3 標本はいわば，母集団から取り出した (抽出した) 値の集まりである，と考えることができる。標本の性質をもとに，母集団の性質を推測するためには，標本は次の 2 つの性質を備えていなければばならない:

不偏性：母集団から偏りなく抽出されていること

独立性：個々のデータが，互いに影響を及ぼしていないこと

これらの性質を備えている標本のことを，**無作為標本**と呼ぶ。

■確率分布

3.1 とりうる実現値にそれぞれ確率が割りふられている変数のことを，**確率変数**という。また，それぞれの実現値に確率が割りふられているようすのことを，**確率分布**という。

3.2 重要な確率分布のひとつに，**正規分布**がある。平均 0, 分散 1 の正規分布を，とくに**標準正規分布**と呼ぶ。

■母集団特性の推定

サイズ n の無作為標本から，母集団の性質について推定するとき，

4.1 **母平均** μ (「ミュー」) の推定のためには，標本平均 $\bar{x}$ を用いるとよい。

4.2 **母分散** σ^2 (「シグマの二乗」) の推定のためには，標本分散 s^2 を少し大きめに修正した**不偏分散**

$$u^2 = \frac{1}{n-1}\sum_{i=1}^{n}(x_i - \bar{x})^2$$

を用いるとよい。

■仮説検定

5.1 **仮説検定**と呼ばれる手法は，次の 4 つの段階からなる。

1. **帰無仮説**(H_0) を設定する。
2. **検定統計量**を定める。
3. 決められた有意水準のもとでの**棄却域**を定める。
4. 標本から検定統計量の値を求め，棄却域と比較して，帰無仮説の**棄却**の有無を決定する。

5.2 **有意水準**は，「帰無仮説が真のとき，誤って帰無仮説を棄却してしまう」確率をあらわしている。5% ないし 1% がよく用いられる。

■実験研究の基礎概念

6.1 **実験**とは，いくつかの変数の値を研究者が操作し，それが別の変数にどう影響するか，を調べる研究のことである。

6.2 したがって実験研究では，変数は次の 3 つのどれかに分類されることになる。

従属変数　測定される変数。“原因-結果” という文脈でいえば，結果の側。

独立変数　研究者が操作する変数。**要因**, **処理**, **説明変数**, などともいう。

剰余変数　従属変数に影響を与えるかもしれないのに，研究者が操作していない変数。

6.3 独立変数のとる値は，いくつかに限られるのがふつうである。このとき，それぞれの値を**水準**という。

6.4 独立変数が複数ある実験の場合，水準と水準の組み合わせのことを**セル**という。

6.5 あるセルのなかにある測定値の数のことを，**繰り返し数**と呼ぶ。

6.6 心理学での実験研究においては，独立変数 (要因) の操作のしかたを，つぎの 2 種類におおまかにわけることができる。

被験者間要因：要因の各水準ごとに，異なる被験者が用意される場合

被験者内要因：各被験者が，その要因のすべての水準の下で実験を行う場合

2 なぜ分散には 2 種類あるのか?

予備知識 4.2 として挙げた「不偏分散」については，多くの人が納得のいかない思いをするようです。なぜ，本来の分散 (標本分散) のほかに，不偏分散が必要なのでしょうか? この 2 つはどのように使いわければ良いのでしょうか?

そこで，以下に 3 通りの説明 (梅，竹，松) を用意しました。先に進むほど，突っ込んだ議論になります。

すくなくとも，梅コースの内容については，きちんと理解してください。竹コース・松コースは，読み飛ばしてもかまいません。

2.1 梅コース

データ $x_1, x_2, \cdots, x_n$ について，全体的な大きさをあらわす指標としては，平均

$$\text{平均}\ \ \bar{x} = \frac{1}{n}\sum_{i=1}^{n} x$$

がよく用いられる。

また，値のばらつきをあらわす指標としては

$$\text{標本分散}\ \ s^2 = \frac{1}{n}\sum_{i=1}^{n}(x-\bar{x})^2$$

$$\text{不偏分散}\ \ u^2 = \frac{1}{n-1}\sum_{i=1}^{n}(x-\bar{x})^2$$

の 2 種類がよくもちいられる。

データについて述べる際，標本分散 s^2(ないし標本標準偏差 s) を用いるべきか，それとも不偏分散 u^2(ないし不偏標準偏差 u) を用いるべきかは，記述の目的によって決まる問題である。

- 手元のデータそのものについての要約に重点がある場合には，標本分散を
- 母集団についての推測に重点がある場合には，不偏分散を

用いるのが理にかなっている。もっとも，どちらを使ってもおかしくないケースも多い。

2.2　竹コース

手元にあるデータ $x_1, x_2, \cdots, x_n$ が，ある母集団からの無作為標本だとみなせる場合について考える。母集団のなかには無限個の (ないし，非常に多くの) 値が含まれていると考えられるが，それら無限個の値にも，平均や分散があると考えることができるだろう。ここで，母平均 (母集団の平均) を μ，母分散 (母集団の分散) を σ^2 と表記することにする。

では，手元にあるデータから，母集団の性質を推測する方法について考えてみよう。

■母平均の推定量　まず，母平均 μ を推定するためには，標本のどのような性質に着目すればよいだろうか。いろいろな考え方がありうるが，一般的にいって，標本平均 $\bar{x}$ に着目するやり方が，一番優れていることがわかっている。そこで，**母平均 μ の推定のためには，標本平均 $\bar{x}$ を用いる**。

■母分散の推定量　ところが，母分散 σ^2 の推定という問題は，さほど簡単ではない。標本の分散 s^2 は，一般的にいって，σ^2 よりも少し小さめの値になってしまう。なぜか?

もともと分散とは,「それぞれの値と平均との距離 (偏差) の二乗の平均」をあらわすものである。だから，σ^2 の推定量としては，本来は $\frac{1}{n}\sum(x_i - \mu)^2$ がふさわしいのである。

しかし現実には，母平均 μ の値はわからないので，標本平均 $\bar{x}$ で代用せざるを得ない。ところが $\frac{1}{n}\sum(x_i - \bar{x})^2$ は，本来の推定量 $\frac{1}{n}\sum(x_i - \mu)^2$ よりも，少し小さめになってしまう。なぜなら，いま任意の値 c について $\sum(x_i - c)^2$ を求めることにすると，その値が一番小さくなるのは，c が $\bar{x}$ に一致するときだからである。

そこで，s^2 を少し大きめに修正したものを，σ^2 の推定量にすればいい，という考え方が登場する。この修正された分散を「不偏分散」と呼んでいる。ここで，

$$\text{不偏分散}\ u^2 = \frac{1}{n-1}\sum_{i=1}^{n}(x_i - \bar{x})^2$$

であるということがわかっている (2.3 参照)。**母分散 σ^2 の推定のためには，この不偏分散 s^2 を用いる**。なお，不偏分散と区別するために，本来の分散を「標本分散」と呼ぶことがある。

2.3 松コース

では，なぜ不偏分散 u^2 の分子は $n-1$ なのだろうか? どうしても気になってしかたがないあなたのために，徹底的な説明をお送りしよう。

2.3.1 確率変数と期待値

まず，期待値という概念を導入する。少し抽象的な話になるので，ゆっくり読み進めてほしい。

数学の世界では，取りうる値 (実現値) に確率が割り振られているような変数のことを，確率変数と呼んでいる。ある確率変数 Y について，その確率分布の平均を，Y の期待値 $E(Y)$ と呼ぶ。

たとえば，「サイコロを振ったときに出る目」という変数 X は，実現値 $(1,2,3,4,5,6)$ に確率が割り振られているので (すべて $1/6$)，確率変数だということができる。その期待値 $E(X)$ は，サイコロを無限回振って手にはいる，無限個の目 $(1,1,1,\ldots,2,2,2,\ldots,6,6,6)$ の平均値，すなわち 3.5 である。

■ある変量の期待値　いま手元に，ある変量についての n 個のデータ $x_1, x_2, \cdots, x_n$ があるとしよう。これらのデータは，いわば X という謎のサイコロを n 回振って手に入れた値だ，とみなすことができる。つまり，変量 X は，確率変数だとみなすことができるわけである。

その期待値 $E(X)$ とは，「データサイズ n が無限大にまで大きくなったときに，そこから得られる平均」のことである。手元のデータがなんらかの母集団の無作為標本であるならば，「無限大の大きさの標本」とは，すなわち母集団のことになる。だから，これは母平均 μ をあらわしている。すなわち，

$$E(X) = \mu \tag{1}$$

■ある変量のばらつきの期待値　つぎに，変量 X のあるひとつのデータと，その母平均 μ とのずれの大きさについて考えてみたい。そのためには，ずれの絶対値 $|X-\mu|$ ついて考えればよいだろう。しかし，絶対値は数学的に扱いが面倒なので，そのかわりに，ずれの二乗 $(X-\mu)^2$ について考えることにする。

その期待値 $E[(X-\mu)^2]$ とは，「無限大のサイズの標本について，すべてのデータからそれぞれの $(X-\mu)^2$ を求めた，その平均」のことである。さきにみたように，「無限大の大きさの標本」は

母集団に相当するから，結局これは母分散 σ^2 のことである。すなわち，

$$E[(X-\mu)^2]=\sigma^2 \tag{2}$$

■データの平均の期待値　では，上の n 個のデータから求める統計量，たとえば平均 $\bar{X}$ について考えてみよう。この値は，その値が確率的に決まるという意味で，いわば $\bar{X}$ という謎のサイコロを 1 回振って手に入れた値だ，とみなすことができる。つまり，標本平均 $\bar{X}$ もまた，確率変数だとみなすことができる。

その期待値 $E(\bar{X})$ とは，「もし標本抽出を無限回繰り返し，標本平均が無限個手に入ったら，それらの平均はなにか」を意味する。当然それは，母平均 μ に一致する。すなわち，

$$E(\bar{X})=\mu \tag{3}$$

である。

ところでこの式は，「標本平均 $\bar{X}$ は母平均 μ の不偏推定量 (偏りのない推定量) だ」ということに対応している。このように，

> 「標本から得られる統計量〇〇は，母集団の特性××の不偏推定量だ」ということを，E(〇〇) $=$ ×× とあらわすことができる。

■データの平均のばらつきの期待値　さて，標本平均 $\bar{X}$ は，母平均 μ からさほど遠くない推定を与えてくれることもあれば，大きく外してしまうこともあるだろう。そのばらつきの程度について考えてみたい。そのためには，推定のずれの絶対値 $|\bar{X}-\mu|$ の期待値について考えればよいだろう。しかし，絶対値は数学的に扱いが面倒なので，そのかわりに，推定のずれの二乗の期待値 $E[(\bar{X}-\mu)^2]$ について考えることにしよう。

証明は省くが，次の式が成り立つことがわかっている。

$$E[(\bar{X}-\mu)^2]=\frac{\sigma^2}{n} \tag{4}$$

この式は，「母平均 μ を，標本平均 $\bar{X}$ を用いて推定するとき，その推定のずれは，母集団の値のばらつき σ^2 が大きいときに大きく，標本サイズ n が大きいときに小さい」という，ごくあたりまえの事柄に対応している。

2.3.2　なぜ $n-1$ か

では，いよいよ本題に戻ろう。まず， u^2 の分子の部分を変形する。

$$\begin{aligned}
&\sum(x_i-\bar{x})^2\\
=&\sum[(x_i-\mu)+(\mu-\bar{x})]^2\\
=&\sum(x_i-\mu)^2+2\sum(x_i-\mu)(\mu-\bar{x})+\sum(\bar{x}-\mu)^2\\
=&\sum(x_i-\mu)^2-2(\bar{x}-\mu)\sum(x_i-\mu)+n(\bar{x}-\mu)^2\\
=&\sum(x_i-\mu)^2-2(\bar{x}-\mu)(\sum x_i-n\mu)+n(\bar{x}-\mu)^2\\
=&\sum(x_i-\mu)^2-2(\bar{x}-\mu)(n\bar{x}-n\mu)+n(\bar{x}-\mu)^2\\
=&\sum(x_i-\mu)^2-2n(\bar{x}-\mu)^2+n(\bar{x}-\mu)^2\\
=&\sum(x_i-\mu)^2-n(\bar{x}-\mu)^2
\end{aligned}$$

第1項 $\sum(x_i-\mu)^2$ の期待値は,

$$\begin{aligned}
E[\sum(x_i-\mu)^2]=&E[(x_1-\mu)^2]+E[(x_2-\mu)^2]+\cdots+E[(x_n-\mu)^2]\\
=&\sigma^2+\sigma^2+\cdots+\sigma^2 \quad \leftarrow\text{式 (2)}\\
=&n\sigma^2
\end{aligned}$$

第2項 $n(\bar{x}-\mu)^2$ の期待値は,

$$\begin{aligned}
E[n(\bar{x}-\mu)^2]=&n\times E[(\bar{x}-\mu)^2]\\
=&n\times\frac{\sigma^2}{n} \quad \leftarrow\text{式 (4)}\\
=&\sigma^2
\end{aligned}$$

従って

$$E\left[\sum(x_i-\bar{x})^2\right]=n\sigma^2-\sigma^2=(n-1)\sigma^2$$
$$E\left[\frac{1}{n-1}\sum(x_i-\bar{x})^2\right]=\sigma^2$$

であり，不偏分散 u^2 が母分散 σ^2 の不偏推定量であることがわかる。

3　平方和，自由度，平均平方

今後の説明の都合上，いくつかの用語を紹介しておく。

■平方和　2 種類の分散

$$\text{標本分散}\ \ s^2 = \frac{1}{n}\sum_{i=1}^{n}(x_i - \bar{x})^2$$
$$\text{不偏分散}\ \ u^2 = \frac{1}{n-1}\sum_{i=1}^{n}(x_i - \bar{x})^2$$

は，分子 $\sum(x-\bar{x})^2$ が共通している。この部分は，偏差の平方 (二乗のこと) の合計なので，偏差平方和と呼んだり，単に**平方和**(SS と略記する) と呼んだりする。**変動**と呼ぶこともある。

■自由度　いっぽう，不偏分散の分子の部分 $n-1$ を，この平方和の**自由度**(df と略記する) と呼ぶ。

自由度とは，自由に値をとることができる変数の数を指す用語である。たとえば，3 つの変数 X_1, X_2, X_3 があるとしよう。これらの変数の値について，平均と平方和を求める式は，

$$\text{平均}\ \ \bar{X} = \frac{X_1 + X_2 + X_3}{3}$$
$$\text{平方和}\ \ SS = (X_1 - \bar{X})^2 + (X_2 - \bar{X})^2 + (X_3 - \bar{X})^2$$

となる。さて，平方和の式の右辺には，3 つの変数が登場するが，$\bar{X}$ が決まっているとすると，自由に動ける変数は 2 つしかない (もし $\bar{X} = 10, X_1 = 9, X_2 = 10$ ならば，X_3 の値は 11 に決まってしまう)。このことを指して，この平方和の自由度は 2 である，と言う。

■平均平方　平方和を自由度で割ったもののことを，**平均平方**と呼ぶ (MS と略記する)。従って，2 章で示したのは，「標本の平均平方は母分散の不偏推定量である」ということであった，といいかえることができる。

4 なぜ分散分析が必要か?

4.1 水準が 3 つ以上のときに必要だ

たとえば，次のような問題について考えてみよう。

例題 1 (後藤ほか編, p.30)

生徒の学習形態のちがいが，課題の達成に影響するかどうかを調べるために，あらかじめ学力の等しい生徒をランダムにわけて，3 つのグループを構成した。グループ 1 では一斉指導，グループ 2 では体験学習，グループ 3 では仲間による討議学習をおこなった。授業終了後，課題の到達度テストを実施したところ，次の得点 (略) が得られた。3 つの学習形態のあいだに差はあるか。

この例題について検討する際には，2 つの路線がある。

■多重比較 ひとつの路線は，この問題を，次の 3 つの問題と，それに対応する帰無仮説 (H_0) に分割する考え方である。

- 一斉指導と体験学習のあいだで，得点のちがいはあるか? ($H_0 : \mu_1 = \mu_2$)
- 体験学習と討議学習のあいだで，得点のちがいはあるか? ($H_0 : \mu_2 = \mu_3$)
- 一斉指導と討議学習のあいだで，得点のちがいはあるか? ($H_0 : \mu_1 = \mu_3$)

これらの帰無仮説 (H_0) のそれぞれについて，仮説検定の手法を用いて検討すればよい。

この路線はわかりやすいし，アイデアそれ自体はまちがっていない。しかし，この路線に沿って，単純に t 検定を繰り返すのは，統計学的にみて，深刻な誤りである (コラム参照)。このような場合には，**多重比較**と呼ばれる手法を用いなければならない。

■分散分析 もうひとつの路線は,

- 3 種類の学習形態の間に, 得点のちがいはあるか? ($H_0 : \mu_1 = \mu_2 = \mu_3$)

という問題ひとつだけについて，仮説検定の手法を用いて検討することである。これを可能にしてくれるのが**分散分析**である。

たいていの場合，多重比較よりも分散分析のほうが簡単だし，結果も解釈しやすい。

■分散分析から多重比較へ　分散分析路線の欠点は，仮に「3 つの学習形態の間に 得点のちがいがある」という結果が得られたとしても，それではどれとどれの間にちがいがあるのかはわからない，という点である。

そこで，まず分散分析をおこない，「3 つの学習形態の間に 得点のちがいがあるか」という点を調べ，ちがいがあることがわかったら，こんどは多重比較によって，「どれとどれの間にちがいがあるか」を調べる，という方法が広く用いられている。このとき，後半の多重比較のことを，とくに**下位検定**と呼ぶ。

4.2　要因が 2 つ以上あるときに必要だ

この例題では，要因がひとつしかない。しかし，実験研究では，複数個の要因を同時に制御することも多い。そのような場合には，**分散分析**の考え方がどうしても必要になる。

コラム：なぜ検定を単純に繰り返してはいけないのか

有意水準 5% で検定をおこなうとする。いま帰無仮説 H_0 が真であるとすると，誤って H_0 を棄却する確率 (タイプ I エラーの確率) は 0.05 である。さて，ひとつの論文のあちこちで，いろいろな問題について別々のデータ解析がおこなわれているとする。検定が 3 回おこなわれているとしよう。いま，検討されている 3 つの H_0 がすべて真であるときに，「論文のなかのどこか 1 箇所以上でタイプ I エラーを犯す確率」は, $1 - 0.95^3 = 0.14$ と，意外に高くなる。10 回のときには，実に 0.40 である。

さらに，異なる検定が同じデータに基づいている場合には，より深刻な問題が生じる。たとえば，A, B, C の 3 群間で，A vs. B, B vs. C, C vs. A の 3 つの t 検定をおこなったとしよう。いま，A の標本平均が，運悪く真の平均よりもずっと高かったとすると，その場合，A vs. B の t 検定でも，A vs. C の t 検定でも，H_0 が棄却されやすくなる。従って，検討されている 3 つの H_0 がすべて真であるときに，「どこか 1 箇所以上でタイプ I エラーを犯す確率」は，$1 - 0.95^3 = 0.14$ よりも高くなり，予想がつかなくなる。

このように，単純に検定を繰り返すと，

- 全体を通じたタイプ I エラーの確率が高くなる。
- データが独立でない場合，タイプ I エラーの確率がわからなくなる。

このような場合には，多重比較のための特別な検定手法を用いなければならない。

第 II 部

基礎編

それではいよいよ，分散分析の考え方についての説明をはじめよう。次の例題を用いて説明することにする。

例題 1 (後藤ほか編, p.30)

生徒の学習形態のちがいが，課題の達成に影響するかどうかを調べるために，あらかじめ学力の等しい生徒をランダムにわけて，3 つのグループを構成した。グループ 1 では一斉指導，グループ 2 では体験学習，グループ 3 では仲間による討議学習をおこなった。授業終了後，課題の到達度テストを実施したところ，次の得点が得られた。3 つの学習形態のあいだに差はあるか。

学習形態	一斉指導	体験学習	討議学習	
	5	8	7	
	4	4	6	
	6	3	8	
	3	3	9	
	3	7	10	
	7	9	9	
	6	8	8	
	5	7	9	
	3	3	7	
	5	4	8	
平均	4.7	5.6	8.1	全平均 6.1
サイズ	10	10	10	

実際の数値を書いているとわかりにくいので，説明文中では下の記号を用いることにする。

要因 A	水準 A_1	水準 A_2	水準 A_3	
	x_{11}	x_{12}	x_{13}	
	x_{21}	x_{22}	x_{23}	
	x_{31}	x_{32}	x_{33}	
	$\vdots$	$\vdots$	$\vdots$	
	x_{n1}	x_{n2}	x_{n3}	
平均	$\bar{T}_1$	$\bar{T}_2$	$\bar{T}_3$	全平均 $\bar{G}$
サイズ	n	n	n	

5 構造モデル

まず，例題 1 の特徴を確認しておこう。**独立変数(要因)** はひとつ，3 **水準**，水準間にデータの対応がない (いわゆる**被験者間**要因)。各水準での標本サイズ (**繰り返し数**) は等しい。

例題 1 のデータについて，「一斉指導群 1 番さんの得点 (5) は，一斉指導を受けた被験者が本来示す得点 (μ_1) に，なんらかの影響 (ε_{11}) が加わったものだ」というふうに考えてみよう。ここでいう "なんらかの影響" とは，学習形態とは無関係な要因すべて，つまり，(この被験者の努力といった) 剰余変数がもたらす影響や，測定の誤差，偶然に生じる値のばらつきなどが含まれる。これをひとことで，**誤差**と呼ぶことにする。

$$
\begin{aligned}
\text{一斉指導群 1 番さんの得点 (5)} &= \text{一斉指導群の母平均}\ (\mu_1) + \text{誤差}\ (\varepsilon_{11}) \\
\text{一斉指導群 2 番さんの得点 (4)} &= \text{一斉指導群の母平均}\ (\mu_1) + \text{誤差}\ (\varepsilon_{21}) \\
&\vdots \\
\text{体験学習群 1 番さんの得点 (8)} &= \text{体験学習群の母平均}\ (\mu_2) + \text{誤差}\ (\varepsilon_{12}) \\
&\vdots \\
\text{討議学習群 1 番さんの得点 (7)} &= \text{討議学習群の母平均}\ (\mu_3) + \text{誤差}\ (\varepsilon_{13}) \\
&\vdots
\end{aligned}
$$

もっと簡潔に表現してみよう。水準 $j(j = \{1, 2, 3\})$ の母平均を μ_j とすると，水準 j の i 番目の測定値 X_{ij} は $X_{ij} = \mu_j + \varepsilon_{ij}$ とあらわすことができる。

さて，各水準の母平均 μ_1, μ_2, μ_3 の平均を μ とあらわすことにし，$\mu_1 = \mu + \tau_1$, $\mu_2 = \mu + \tau_2$, $\mu_3 = \mu + \tau_3$ とする。ここで μ は，すべての得点の母平均，つまり，学習形態によるちがいを除去した得点の母平均をあらわしている。また τ_1, τ_2, τ_3 は，3 種類の学習形態が持っている，得点への (プラスないしマイナスの) **効果**をあらわしている。すると，上の式は次のように書き直すことができる。

> 全体の母平均を μ, 水準 j の効果を τ_j とする。水準 j の i 番目の測定値 X_{ij} は
>
> $$X_{ij} = \mu + \tau_j + \varepsilon_{ij}$$

この数式を，分散分析の**構造モデル**という。

6　分散分析の前提

さて，分散分析では，**誤差 ε_{ij} が平均 0 の正規分布に従い，その分散は等しい**，と仮定する。

いいかえれば，

- $\{x_{11}, x_{21}, \cdots, x_{n1}\}$ は，平均 $\mu + \tau_1$ の正規分布に従う
- $\{x_{12}, x_{22}, \cdots, x_{n2}\}$ は，平均 $\mu + \tau_2$ の正規分布に従う
- $\{x_{13}, x_{23}, \cdots, x_{n3}\}$ は，平均 $\mu + \tau_3$ の正規分布に従う
- この 3 つの正規分布の分散は等しい

と仮定する。

この仮定は，データの性質としては

- 各水準の内側でのデータの分布が，正規分布に近いこと (**正規性**)
- 各水準の内側でのデータの分散が，だいたい同じであること (**等分散性**)

に対応する。

例題 1 のデータについてみると

- 3 枚のヒストグラムは，どれもおおまかにいって，左右対称な山形であり，
- 3 群の標準偏差は 1.35, 2.29, 1.14 であり，あまり大きな差はない。

したがって，誤差についての仮定には無理がなさそうだ。

7　分散分析の発想

■分散の分析とは?　さて，いま知りたいのは，ガソリンによって燃費に差があるかどうかである。仮説検定の枠組みに従えば，帰無仮説 $H_0 : \tau_1 = \tau_2 = \tau_3$ を棄却できるかどうか，を検討することになる。

この問題について検討するためには，τ_1, τ_2, τ_3 のそれぞれについて推定値を求め，その差を調べればいいのではないか? ... という方向に話を進めないのが，分散分析の面白いところである。分散分析では，τ_1, τ_2, τ_3 そのものについての推定をおこなうのではなくて，この 3 つの効果の分散を推定しようとする。これが「分散分析」という名前の由来である。

ここで，構造モデルの各項の分散について，呼び名と表記を決めておこう。

- X_{ij} の分散，すなわち母集団全体の分散 (**全分散**) を，σ^2_{Total} と表記する。
- τ_j の分散 (つまり $\{\tau_1, \tau_2, \tau_3\}$ の分散) を，**要因分散**と呼ぶ。σ^2_A と表記する[*1]。
- ϵ_{ij} の分散を，**誤差分散**と呼ぶ。σ^2_{Error} と表記する。

$$
\begin{array}{ccccccc}
\text{測定値} & & \text{全平均} & & \text{要因の効果} & & \text{誤差} \\
X_{ij} & = & \mu & + & \tau_j & + & \epsilon_{ij} \\
\downarrow & & & & \downarrow & & \downarrow \\
\text{全分散} & & & & \text{要因分散} & & \text{誤差分散} \\
\sigma^2_{Total} & & & & \sigma^2_A & & \sigma^2_{Error}
\end{array}
$$

さて，

- もし学習形態によって得点に差がないならば，$\tau_1 = \tau_2 = \tau_3 = 0$ なので，$\sigma^2_A = 0$ である。
- もし学習形態によって得点に差があるならば，τ_1, τ_2, τ_3 がなんであれ，$\sigma^2_A \neq 0$ である。

だから，τ_1, τ_2, τ_3 についての推定をおこなわなくても，**要因分散 σ^2_A が 0 かどうかを判断すれば，用が足りるのである**。

■要因分散についての検討とは?　ところが，σ^2_A の大きさについての検討は，一筋縄ではいかない。

まず，構造モデルの各項について，標本から推定する方法を考えてみると

- 全平均 μ の推定量は $\bar{G}$

[*1] 後藤ほか (編) では σ^2_{Treat} と表記している。なお，$\sigma^2_A = \dfrac{n \sum \tau_j}{3-1}$ と定義しておく。

- 各水準の効果 τ_j の推定量は $(\bar{T}_j - \bar{G})$
- 誤差 ε_{ij} の推定量は $(x_{ij} - \bar{T}_j)$

以上の推定量を用いて，手元のデータに構造モデルをあてはめると

$$x_{ij} = \bar{G} + (\bar{T}_j - \bar{G}) + (x_{ij} - \bar{T}_j)$$

となる。

$$\begin{array}{lcccccccc}
\text{母集団} & X_{ij} & = & \boxed{\mu} & + & \boxed{\tau_j} & + & \boxed{\epsilon_{ij}} \\
 & & & \Uparrow \text{推定} & & \Uparrow \text{推定} & & \Uparrow \text{推定} \\
\text{標本} & x_{ij} & = & \boxed{G} & + & \boxed{(T_j - G)} & + & \boxed{(x_{ij} - T_j)}
\end{array}$$

ところで，

(a) 測定値 x_{ij} の平均平方 (平方和を自由度で割ったもの) は，全分散 σ^2_{Total} の不偏推定量となる (3 章参照)。

ならば，

(b) $(\bar{T}_j - \bar{G})$ の平均平方は，要因分散 σ^2_A の不偏推定量となるのではないか?

(c) $(x_{ij} - \bar{T}_j)$ の平均平方は，誤差分散 σ^2_{Error} の不偏推定量になるのではないか?

$$\begin{array}{lcccccc}
\text{母集団} & \boxed{\begin{array}{c} X_{ij} \\ \downarrow \\ \sigma^2_{Total} \end{array}} & = \mu + & \boxed{\begin{array}{c} \tau_j \\ \downarrow \\ \sigma^2_{A} \end{array}} & +8 & \boxed{\begin{array}{c} \varepsilon_{ij} \\ \downarrow \\ \sigma^2_{Error} \end{array}} \\
 & \Uparrow \text{推定 } (a) & & \Uparrow \text{推定?}(b) & & \Uparrow \text{推定?}(c) \\
\text{標本} & \boxed{\begin{array}{c} MS_{Total} \\ \uparrow \\ x_{ij} \end{array}} & = \bar{G} + & \boxed{\begin{array}{c} MS_{A} \\ \uparrow \\ (\bar{T}_j - \bar{G}) \end{array}} & + & \boxed{\begin{array}{c} MS_{Error} \\ \uparrow \\ (x_{ij} - \bar{T}_j) \end{array}}
\end{array}$$

先に結論を紹介しておくと，**(c)** は正しいが，**(b)** は正しくない。しかし，この発想じたいは優れているので，このまま話を先に進めてみよう。

8 平方和の分解

まず，各項の平方和を求めてみよう。

$$\text{全体の平方和}\ \ SS_{Total} = \sum_{j=1}^{3}\sum_{i=1}^{n}(x_{ij} - \bar{G})^2$$

$$\text{要因の平方和}\ \ SS_A = \sum_{j=1}^{3}\sum_{i=1}^{n}\{(\bar{T}_j - \bar{G}) - 0\}^2 = n\sum_{j=1}^{3}(\bar{T}_j - \bar{G})^2$$

$$\text{誤差の平方和}\ \ SS_{Error} = \sum_{j=1}^{3}\sum_{i=1}^{n}\{(x_{ij} - \bar{T}_j) - 0\}^2 = \sum_{j=1}^{3}\sum_{i=1}^{n}(x_{ij} - \bar{T}_j)^2$$

ここで，

$$SS_{Total} = SS_A + SS_{Error}$$

という関係が成り立っている (コラム参照)。つまり，ここでおこなっているのは，**測定値の平方和を分解する**作業なのである。

例題 1 の場合。わかりやすいように，全平均 $\bar{G}$ を左辺に移項している。

(得点 − 全平均) $(x_{ij} - \bar{G})$		(水準の平均 − 全平均) $(\bar{T}_j - \bar{G})$		(得点 − 水準の平均) $(x_{ij} - \bar{T}_j)$
$(5 - 6.1)$	$=$	$(4.7 - 6.1)$	$+$	$(5 - 4.7)$
$(4 - 6.1)$	$=$	$(4.7 - 6.1)$	$+$	$(4 - 4.7)$
$\vdots$		$\vdots$		$\vdots$
$(8 - 6.1)$	$=$	$(5.6 - 6.1)$	$+$	$(8 - 5.6)$
$(4 - 6.1)$	$=$	$(5.6 - 6.1)$	$+$	$(4 - 5.6)$
$\vdots$		$\vdots$		$\vdots$
$(7 - 6.1)$	$=$	$(8.1 - 6.1)$	$+$	$(7 - 8.1)$
$(6 - 6.1)$	$=$	$(8.1 - 6.1)$	$+$	$(6 - 8.1)$
$\vdots$		$\vdots$		$\vdots$
$\Downarrow$ 二乗して合計 $SS_{Total} = 145.47$		$\Downarrow$ 二乗して合計 $SS_A = 62.07$		$\Downarrow$ 二乗して合計 $SS_{Error} = 83.4$

ここで行ったのは，得点のばらつき 145.47 を，学習形態に由来するばらつき 62.07 と，それ以外のばらつき 83.4 とに分解する作業であった，ということができる。

コラム: なぜ平方和は分解できるのか

構造モデル

$$x_{ij} = \bar{G} + (\bar{T}_j - \bar{G}) + (x_{ij} - \bar{T}_i)$$

の $\bar{G}$ を左辺に移項して

$$x_{ij} - \bar{G} = (\bar{T}_j - \bar{G}) + (x_{ij} - \bar{T}_i)$$

両辺を 2 乗して

$$(x_{ij} - \bar{G})^2 = (\bar{T}_j - \bar{G})^2 + (x_{ij} - \bar{T}_i)^2 + 2(\bar{T}_j - \bar{G})(x_{ij} - \bar{T}_i)$$

合計して

$$\sum_j \sum_i (x_{ij} - \bar{G})^2 = \sum_j n(\bar{T}_j - \bar{G})^2 + \sum_j \sum_i (x_{ij} - \bar{T}_j)^2 + \sum_j \sum_i 2(\bar{T}_j - \bar{G})(x_{ij} - \bar{T}_j)$$

第三項は

$$\begin{aligned} \sum_j \sum_i 2(\bar{T}_j - \bar{G})(x_{ij} - \bar{T}_i) &= 2\sum_j \{(\bar{T}_j - \bar{G}) \sum_i (x_{ij} - \bar{T}_i)\} \\ &= 2\sum_j \{(\bar{T}_j - \bar{G}) \times 0\} \\ &= 0 \end{aligned}$$

従って，

$$\begin{aligned} \sum_j \sum_i (x_{ij} - \bar{G})^2 &= \sum_j n(\bar{T}_j - \bar{G})^2 + \sum_j \sum_i (x_{ij} - \bar{T}_j)^2 \\ SS_{Total} &= SS_A + SS_{Error} \end{aligned}$$

であることがわかる。

なお，構造モデルがもっと複雑なものになっても，上記と同じように，全体の平方和を各項の平方和の和に分解することができる。

9 平均平方の算出

次に，それぞれの平方和が持つ自由度について考えておこう。自由度とは，自由に動くことができる値の数なので (3 章参照),

- 全体の平方和 SS_{Total} の自由度 =(値の個数 -1) $= 3n-1$
- 要因の平方和 SS_A の自由度 =($\bar{T}_j$ の個数 -1) $= 3-1$
- 誤差の平方和 SS_{Error} の自由度 $=$ 水準数$\times$ (水準内の値の個数-1) $= 3(n-1)$

となる。ここで

$$(3n-1) = (3-1) + 3(n-1)$$

であり，**自由度もまた，平方和と同じように分解されている**。

それでは，各項の平均平方 (平方和を自由度で割った値) を求めよう。

$$\begin{aligned}
&\text{全体の平均平方} \quad MS_{Total} = SS_{Total}/(3n-1) \\
&\text{要因の平均平方} \quad MS_A = SS_A/(3-1) \\
&\text{誤差の平均平方} \quad MS_{Error} = SS_{Error}/3(n-1)
\end{aligned}$$

例題 1 の場合:

(得点 – 全平均) $(x_{ij}-\bar{G})$		(水準の平均 – 全平均) $(\bar{T}_j-\bar{G})$		(得点 – 水準の平均) $(x_{ij}-\bar{T}_j)$
$\Downarrow$ 二乗して合計 $SS_{Total}=145.47$	$=$	$\Downarrow$ 二乗して合計 $SS_A=62.07$	$+$	$\Downarrow$ 二乗して合計 $SS_{Error}=83.4$
$\Downarrow$ 自由度は $3n-1=29$		$\Downarrow$ 自由度は $3-1=2$		$\Downarrow$ 自由度は $3(n-1)=27$
$\Downarrow$ わり算して $MS_{Total}=5.02$		$\Downarrow$ わり算して $MS_A=31.03$		$\Downarrow$ わり算して $MS_{Error}=3.08$

10 平均平方の意義

この章も，少し面倒な内容を含んでいるので，3 通りの説明 (梅，竹，松) を用意しました。先に進むほど，突っ込んだ議論になります。すくなくとも，梅コースの内容については，きちんと理解してください。竹コース・松コースは，読み飛ばしてもかまいません。

10.1 梅コース

さて，いま私たちが目指しているのは，要因分散 σ_A^2 が 0 かどうかの判断である。そのためには，MS_A だけを調べていては不十分である。なぜなら，誤差分散 σ_{Error}^2 が大きいときにも，MS_A は大きくなってしまうからである。

そこで，MS_A を MS_{Error} で割った量

$$F = MS_A/MS_{Error}$$

を調べる。

例題 1 の場合は，

$$F = 31.03/3.08 = 10.04$$

F 値は，要因分散 σ_A^2 が 0 のときに 1 に近くなり，σ_A^2 が 0 でないとき (すなわち，要因の水準によって差があるとき) には 1 よりも大きな値になる。

10.2　竹コース

以上の内容を，別の角度から説明しよう。

図 1 は，例題 1 のデータを縦に並べ，プロットしたものである。図の上・中・下が，3 種類の学習形態に対応している。黒丸は測定値を，中央の縦の点線は全平均 $\bar{G}$ を，太線は各水準の平均 $\bar{T}_1, \bar{T}_2, \bar{T}_3$ を示している。MS_{Total} は黒丸のばらつき，MS_A は太線のばらつき，MS_{Error} は黒丸から太線までの垂線の長さのばらつきに相当する。

この図をみるだけで，被験者の属する群によって黒丸の位置が異なっていること，したがって要因の効果がみられることが，直感的にわかるだろう。

では，もしデータが**図 2** のようであったらどうだろうか。この図の太線は，**図 1** の太線とまったく同じである。しかしこの図の黒丸の布置をみても，要因の効果がみられるとはとても思えない。なぜなら，測定値のばらつきが大きいからである。たしかに，太線にもばらつきはみられるものの，それは単に測定値のばらつきのせいではないか，つまり，もうすこし測定値を増やせば，太線のかたちは簡単に変わってしまうのではないか — という気がするだろう。

このように，要因の効果があるかどうか ($\sigma_A \neq 0$ かどうか) の判断は，誤差 e_{ij} のばらつきと 比べて 水準の平均 $\bar{T}_j$ のばらつきが大きいかどうか，に基づいておこなわれるべきである。そ

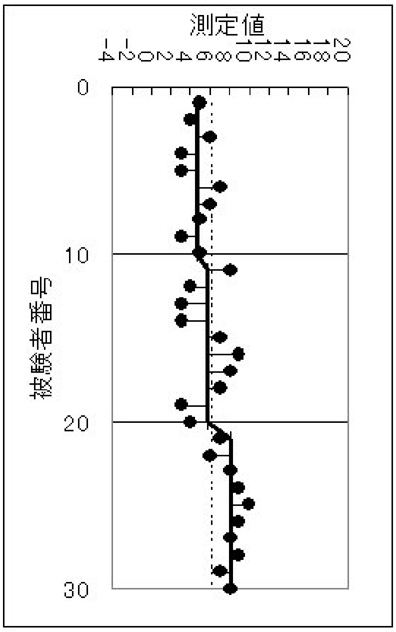

図 1: 例題 1 のデータ

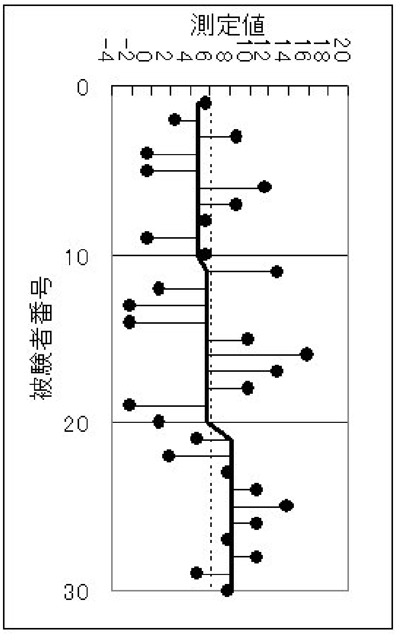

図 2: もしこんなデータなら…

こで，「水準の平均のばらつき MS_A が，誤差 e_{ij} のばらつき MS_{Error} の何倍あるか」，つまり $F = MS_A/MS_{Error}$ を求めるのである。もし要因の効果がなければ，MS_A と MS_{Error} は同程度となり，F は 1 に近くなるだろう。もし要因の効果があるのなら，F はもっと大きな値になるだろう。

10.3 松コース

以下の説明は，2 章の松コース (2.3) を読了した人向けに書かれています。

3 章で述べたように，データの平方和を自由度で割ると，母分散の不偏推定量が手にはいる。これを期待値という概念を用いてあらわせば，

$$E(MS_{Total}) = \sigma^2_{Total}$$

さて，要因の平均平方 MS_A の期待値は

$$E(MS_A) = \sigma^2_{Error} + n\sigma^2_A$$

となる[*2]。つまり，MS_A は σ^2_A の不偏推定量ではなく，σ^2_A と σ^2_{Error} の両方を反映する統計量なのである。

いっぽう，誤差の平均平方 MS_{Error} の期待値は

$$E(MS_{Error}) = \sigma^2_{Error}$$

であり[*3]，MS_{Error} は誤差分散 σ^2_{Error} の不偏推定量である。

さて，

- もし要因の効果がないならば (H_0 が真ならば)，MS_A と MS_{Error} とは，ともに σ^2_{Error} の不偏推定量だから，近い値になるはずである。
- いっぽう，要因の効果があるならば (H_0 が偽ならば)，それがどのような効果であれ，MS_A は大きくなるはずである。

そこで，$F = MS_A/MS_{Error}$ を検定統計量として，$H_0 : \sigma^2_A = 0$ についての仮説検定をおこなうわけである。(σ^2_A の大きさの推定や，τ_1, τ_2, τ_3 の推定には，もはや関心が持たれていないことに注目してほしい。)

[*2] 高校までの数学で導出できる。お試しあれ。
[*3] 同上。

11　F 検定

それでは，F を検定統計量として，要因の効果の有無についての仮説検定をおこなうことにしよう。

予備知識 5.1 に挙げたように，仮説検定は 4 つの段階からなる。

■1. 帰無仮説 (H_0) を設定する　帰無仮説は:

$$H_0 : \text{要因の効果はない}\ (\tau_1 = \tau_2 = \tau_3 = 0,\ \sigma_A^2 = 0)$$

■2. 検定統計量を定める　すでに説明したように，検定統計量としては F を用いる。

■3. 決められた有意水準のもとでの棄却域を定める　さて，誤差の正規性と等分散性という仮定が成り立っているときに限り (6 章参照)，F には以下の性質がある。帰無仮説が真である場合には，F は「**自由度 (要因の自由度, 誤差の自由度) の F 分布**」と呼ばれる確率分布に従う。いっぽう，帰無仮説が偽の場合には，F は大きくなる。

そこで，F 分布の右 $\alpha\%$ の範囲を，有意水準 $\alpha\%$ の棄却域と定めることにする。

> 例題 1 では: 自由度 $(2, 27)$ の F 分布を用いる。1% 棄却域は $F > 5.49$ である。

■4. 棄却の有無を決定する　F 値が棄却域に含まれていた場合は，帰無仮説は棄却される。

> 例題 1 では: $F = 10.04$ は棄却域に含まれているので，棄却域は 1% 有意水準で棄却される。従って，学習形態という要因の効果が認められたと判断される。

ここで，F 値の大きさは効果の大きさをあらわしているわけではない，という点に注意してほしい。前章でみたように，F は効果の大きさ (σ^2_{Total}) をあらわす指標ではない。

12　まとめ: 1 要因の分散分析

どのようなデータであれ，分散分析を用いたデータ分析は，6 つの段階からなっている。

1. データの構造についてよく考え，構造モデルを構築する。
2. 誤差の分布についての仮定が，データにあてはまっているかどうか検討する。
3. 平方和を分解し，各項の平均平方を求める。
4. 検討したい要因について，F を求め，帰無仮説の棄却の有無を判断する。
5. それがなにを意味しているのか考えるために，グラフに戻ったり，下位検定に進んだりする。

この解説書では，このうち 1-5 の段階について，いわゆる被験者間 1 要因計画の実験データを例に挙げて，詳しく検討してきた。

ここまでの内容をまとめておこう。1 要因 (k 水準, 水準間にデータの対応なし) の実験の結果，測定値 x_{ij} を得た。ただし，i は各水準内での測定値の番号 ($1 \sim n_k$)，j は水準の番号 ($1 \sim k$) とする (下表)。

要因 A	水準 A_1	水準 A_2	$\cdots$	水準 A_k	
	x_{11}	x_{12}	$\cdots$	x_{1k}	
	x_{21}	x_{22}	$\cdots$	x_{2k}	
	x_{31}	x_{32}	$\cdots$	x_{3k}	
	$\vdots$	$\vdots$	$\cdots$	$\vdots$	
	$x_{n_1 1}$	$x_{n_2 2}$	$\cdots$	$x_{n_k k}$	
平均	$\bar{T}_1$	$\bar{T}_2$	$\cdots$	$\bar{T}_k$	全平均$\bar{G}$
サイズ	n_1	n_2	$\cdots$	n_k	

このとき，誤差の正規性と等分散性という仮定の下で，分散分析をおこなうことができる。

分散分析の計算過程は，下のような書式の表にまとめることが多い。これを**分散分析表**という。

1 要因 (水準間にデータの対応なし) の分散分析表

変動因	平方和 (SS)	自由度 (df)	平均平方 (MS)	F
要因 (A)	$\displaystyle\sum_{j=1}^{k} n_j(\bar{T}_j - \bar{G})^2$	$k-1$	$\dfrac{SS_A}{k-1}$	$\dfrac{MS_A}{MS_{Error}}$
誤差 ($Error$)	$\displaystyle\sum_{j=1}^{k}\sum_{i=1}^{n_j}(x_{ij} - \bar{T}_j)^2$	$\displaystyle\sum_{j=1}^{k}(n_j - 1)$	$\dfrac{SS_{Error}}{\sum_{j=1}^{k}(n_j-1)}$	
全体	$\displaystyle\sum_{j=1}^{k}\sum_{i=1}^{n_j}(x_{ij} - \bar{G})^2$	$\displaystyle\sum_{j=1}^{k} n_j - 1$		

つづく . . . かも