分散分析アルゴリズム

心理学統計法資料

1　被験者間要因

1.1　一要因の分散分析

まずは被験者間計画の、一要因の分散分析を考えよう。被験者間計画は、内計画に比べて計算が理解しやすい。

例として、学校で三つのクラス、A,B,C において、それぞれ異なる教科書、甲、乙、丙を使って教育し、期末テストの結果に差があるかどうかを検証するというシーンを考えて欲しい。もちろん、この検証の前提条件として、三つのクラスは検証前は成績が同じであり、同じ教師が同じように教える、と考えてもらいたい。つまり、他の要因から影響がないように、十分にコントロールされているわけである。また、各クラスは 4 人からなるグループで、計 12 人である。各クラスの人数が異なる場合でも、もちろん分散分析はできるが、計算の仕方が面倒になる。しかし、ここでは分散分析の原理的な理解をすすめることが目的なので、説明は割愛する。原理さえ理解していれば、後はその応用であり、現実問題としては、分析をしてくれるアプリケーションが、適当な修正をかけてくれるからである。

前置きが長くなった。さて、成績が以下のようになったとしよう (表 1)。全体の平均点は 30.42 点である。しかし、クラス毎の平均点を見てみると、A 組が 27.75 点、B 組が 29.25 点、C 組は 34.25 点であることがわかる。

前提条件にあげたように、この三つの組は、本来同程度の成績であった。つまり、各組の平均は、理想的には 30.42 点であるはずなのだ。それをクラス毎に、教科書を変えるという操作をしたので、成績に差が出てしまった。実に、A 組は全体平均よりも 2.67 点も低いし、B 組は 1.17 点も低い。C 組は平均より 3.83 点高い。

分散分析では、これを**要因の効果**だと考える。つまり、教科書を変えたことで、A 組は –2.67 点、B 組は –1.17 点、C 組は +3.83 点分の効果が現れたのである。このクラス別の効果の大きさをどのように測定すればいいだろうか。表 2 の ‘効果’ 列をそのまま加算したのでは 0.0 になってしまう。そこで、これを二乗して足し合わせることを考える。分散と同じ発想をするわけだ。すると、全体的な大きさは 23.167 になる。これが効果による変動、と考えられる。

しかし、このままでは全体の分散を説明する大きさにはならない。なぜなら、各クラスは 4 人いたわけで、クラスの効果 (A 組は –2.67 点、など) はクラス中の一人一人にかかる効果である。

表 1　被験者間・一要因分散分析の例

被験者	クラス	成績
学生 a	A	28
学生 b	A	29
学生 c	A	29
学生 d	A	25
学生 e	B	30
学生 f	B	29
学生 g	B	28
学生 h	B	30
学生 i	C	35
学生 j	C	32
学生 k	C	38
学生 l	C	32

表 2　平均値と効果

群	平均	全体平均	効果
A	27.75	30.42	-2.67
B	29.25	30.42	-1.17
C	34.25	30.42	3.83

これを考えると、正確に効果の計算をするには、これを 4 倍する必要があることがわかる (表 13)。

表 3　人数倍して正しく評価

群	平均	全体平均	効果	効果の二乗	人数倍
A	27.75	30.42	-2.67	7.129	28.516
B	29.25	30.42	-1.17	1.369	5.476
C	34.25	30.42	3.83	14.669	58.676
				23.167	92.667

さてさて。しかしながら、A 組の中にも優秀な人、不出来な人がいる。こういった差は、要因の効果ではなく、個々人特有の問題であり、要因では理解できないことなので、**誤差**であると考える。例えば、学生 a は、平均 27.75 点のクラスにいるが、28 点取れている。この +0.25 点分は誤差である。こういった考え方をもとに、表 1 を書き換えてみたのが、表 4 である。

さて、分散分析では、効果が誤差に比べて十分に大きいかどうか、を判定材料に使うことは既

表 4　効果と誤差を書いてみる

被験者	クラス	成績	全体平均	クラス平均	効果	誤差
学生 a	A	28	30.42	27.75	-2.67	0.25
学生 b	A	29	30.42	27.75	-2.67	1.25
学生 c	A	29	30.42	27.75	-2.67	1.25
学生 d	A	25	30.42	27.75	-2.67	-2.75
学生 e	B	30	30.42	29.25	-1.17	0.75
学生 f	B	29	30.42	29.25	-1.17	-0.25
学生 g	B	28	30.42	29.25	-1.17	-1.25
学生 h	B	30	30.42	29.25	-1.17	0.75
学生 i	C	35	30.42	34.25	3.83	0.75
学生 j	C	32	30.42	34.25	3.83	-2.25
学生 k	C	38	30.42	34.25	3.83	3.75
学生 l	C	32	30.42	34.25	3.83	-2.25
					二乗和	38.25

に述べた。効果の分散、誤差の分散はどのようにして求められるだろうか。

誤差の話からしよう。表 4 には、一人一人につけられた誤差が示されている。これの総和は、もちろん 0 になってしまうので、二乗和する、すなわち $S_e = \sum e_i^2 = 38.25$ で、これが誤差の大きさである。

これをこのまま比較に用いるわけにはいかない。分散分析は操作や誤差によって生じた分散を比較する。分散である限り、二乗和をデータ数 (あるいはデータ数 −1) で割る必要があるだろう。しかし、38.25 という大きさの効果をもたらしたものは、N 人それぞれというより、三つのクラスであり、クラスの中の一人一人である。「どこが違うんだ」と言われてしまいそうであるが、もう少し考えてみよう。

今、クラス A によって生み出された効果の大きさが V^2 だけあったとしよう。これは a,b,c,d の 4 人で生み出した効果だ。しかし、平均や分散は、総和したものを割るわけだから、$a+b+c+d = V^2$ である。今、V^2 はわかっているとすると、例えば c がどれほどの効果を生んだかは、$c = V^2-a-b-d$ で求められる。このように、誰か一人の値がわかっていなくても、全体さえあれば一人一人の効果は特定することが可能だ。言い方を変えると、V^2 の効果が得られるのは、実質 3 人分の働きである。このことから、クラス A において考慮すべき効果は $V^2/(4-1)$ で得る。

この「実質何人分の働きか」という考え方は、自由度と呼ばれるものである[*1]。各クラスの自由度は $4-1=3$ であり。これが三クラス分あるのだから、誤差全体の自由度は 3×3 より 9 であることがわかる。

[*1] 自由度は degrees of freedom であり、略して df と書く。

分散分析で比較すべきは、平均からの差を二乗和したものを自由度で割った、平均平方 (Mean Squares、MS) である。今回、誤差の平均平方は 38.25/9 = 4.25 である*2。

さて、では効果の大きさは先ほど表 13 に算出した。この SS を誤差の時と同じように、自由度で割らなければならない。効果の自由度は、三つのクラスに分けたところから来ているから、3 − 1 = 2 である。効果の MS は 92.666/2 = 46.333 になる。

これで計算は終わり。効果と誤差の比、すなわち 46.333/4.250 = 10.902 が F 値と呼ばれるものである。分散分析はふたつの群の比から得られる理論分布、F 分布を用いて、この F 値がどれほどの確率で生じるものかを見て検定する。今回は $p = 0.0039$ であり、5% 水準よりも小さい値なので、有意差アリといえそうだ。

ちなみに、分散分析は、計算過程を表した表を作ると、一見してわかりやすい結果になる。

表 5　被験者間要因・一要因分散分析の表

	SS	df	MS	F 値	prob.
クラスの効果	92.666	2	46.333	10.902	0.0039
誤差	38.25	9	4.25		
全体	130.916	11			

ここで、全体の分散 130.916 が効果の SS、92.666 と誤差の SS、38.25 に分けられていることに注目しよう*3。同様に、全体の自由度 11 も、効果の df、2 と誤差の df、9 とに分けられている*4。分散分析は、このように分散 (や自由度) を分解していくプロセスであることが一目瞭然である。

ここまでの計算プロセスだけまとめておこう。

1. 全体平均と各水準の平均値を求め、効果を算出する (表 2)。
2. 誤差による変動分を算出する (表 4)。
3. 全変動分を算出して、検算しておく。

$$\sum (a_{ij} - \bar{a}_{..})^2 = 130.916$$

4. 分散分析表を書く。

分散分析の考え方は、基本的に「要因の効果がなければ、全体の平均値は等しい」という考え方から始まる。水準毎の平均値が全体の平均値と違うことこそが、要因の効果であると考えるわけだ。また、平均値で説明できない分は、誤差であると考える。被験者 (あるいは被験体) がいくつかあったら、自然なばらつきをするものであり、注目すべきは誤差と効果ではどちらが大きいか、である。

*2 ちなみに、平均からの差を二乗和したもの、は Sum of Squares であり SS と書く。

*3 130.916 = 92.666 + 38.25

*4 11 = 9 + 2

1.2　二要因の分散分析

それでは要因がふたつに増えたとしよう。

先ほどの例を拡張して、学生への教育効果を考えよう。要因として、教師 A,B の二人と教科書甲、乙の二つがあるとする。二つの要因の組み合わせとして、教師 A が教科書甲で教える場合、教師 A が教科書乙で教える場合、教師 B が教科書甲で教える場合、教師 B が教科書乙で教える場合、の四パターンある。それぞれのパターンについて、能力が同程度と思われる (他の要因の効果がなさそうな) 三人の学生を割り振り、授業をやってもらって、成績を見たのが表 6 である。

表 6　被験者間・二要因分散分析の例

被験者	教師	教科書	成績
学生 a	A	甲	28
学生 b	A	甲	29
学生 c	A	甲	29
学生 d	A	乙	25
学生 e	A	乙	30
学生 f	A	乙	29
学生 g	B	甲	28
学生 h	B	甲	30
学生 i	B	甲	35
学生 j	B	乙	32
学生 k	B	乙	38
学生 l	B	乙	32

実は、数値的には先ほどと同じデータなので、全体の平均は 30.42、分散は 130.92 である。一要因の時と同じく、この全体の分散を効果別に分解していく作業が必要である。

まずは教師の効果を考えよう。教師 A の平均点は 28.33 で、平均より 2.09 点低い。教師 B は逆で、2.09 点高い。これが教師の効果で、どちらの教師につきたいか、と聞かれたら普通はこの結果から、教師 B が選ばれるだろう (表 7)。全体平均から、各群の平均値を引いて、その差を二

表 7　教師の効果

群	平均	全体平均	効果	二乗	人数倍
A	28.33	30.42	-2.08	4.34	26.04
B	32.50	30.42	2.08	4.34	26.04
				SS	52.08

乗したのが偏差平方の欄にある数値で、これを総和したもの (52.08) が教師によって生み出された効果の大きさである。

同様に、教科書の効果を見てみよう。学生 a〜c と g〜i の六人が甲の教科書を使っている。この六人の平均点は 29.83 で、平均点より 0.59 点低い。乙の教科書を使ってる学生の平均点は、全体の平均よりも 0.59 点高い。どちらの教科書がよいのか、という点から考えれば、乙の教科書を使う方が良さそうだ (表 8)。同様に、ここでの偏差平方は 4.08 になる。教科書の効果は教師の効果ほどではないらしい。

表 8　教科書の効果

群	平均	全体平均	効果	二乗	人数倍
甲	29.83	30.42	-0.58	0.34	2.04
乙	31.00	30.42	0.58	0.34	2.04
				SS	4.08

では、教師 A で教科書乙の場合と、教師 B で教科書甲の場合はどちらがよいのだろうか？

二要因以上の分散分析には、こういった組み合わせの効果を考える必要が出てくる。これを特に交互作用 (interaction) という[*5][*6]。

教師 A は平均点を 2.09 点下げる。甲の教科書はさらに 0.59 点下げる。ということは、全体平均 30.42 に二つの効果が加わると、30.42 − 2.09 − 0.59 = 27.74 である。ところが、教師 A かつ教科書甲で授業を受けた、学生 a〜c の平均点は 28.67 である。なぜこのような差ができてしまうのだろうか？実は、27.74 − 28.67 = −0.93 で算出される効果こそ、交互作用と呼ばれるものである。組み合わせたことによって、特別な効果が生まれたのである。

さて、それらの効果を計算したのが表 9 である。ここでの偏差平方、10.08 が交互作用による変動である。

最後に、誤差の計算をしよう。もう一度組み合わせの平均を見てもらいたい。これは、教師の

表 9　組み合わせの効果:交互作用

	全体平均	教師	教科書	結果	平均	交互作用	二乗	人数倍
A で甲	30.42	-2.08	-0.58	27.75	28.67	-0.91	0.83	2.50
A で乙	30.42	-2.08	0.58	28.92	28.00	0.92	0.85	2.54
B で甲	30.42	2.08	-0.58	31.92	31.00	0.92	0.85	2.54
B で乙	30.42	2.08	0.58	33.09	34.00	-0.91	0.83	2.50
							SS	10.08

[*5] これに対して、単体での効果を主効果 (Main Effect) という。
[*6] 社会心理学の専門用語としては、interaction は “相互作用” と呼ぶが、混同しないように。

効果、教科書の効果、交互作用の三つが組み合わさったものである[*7]。しかし、それでも学生 a の成績 28 点と組み合わせ平均 28.67 の間には 0.67 のズレがある。これは全ての効果を考えた上で、まだずれていると思われるところなのだから、もはや学生 a に生じる誤差としか考えようがない。こうして誤差を算出し、その平方和を計算したのが表 10 である。

表 10　誤差を取り出す

	成績	組合平均	誤差	二乗和
学生 a	28	28.67	-0.67	0.44
学生 b	29		0.33	0.11
学生 c	29		0.33	0.11
学生 d	25	28.00	-3.00	9.00
学生 e	30		2.00	4.00
学生 f	29		1.00	1.00
学生 g	28	31.00	-3.00	9.00
学生 h	30		-1.00	1.00
学生 i	35		4.00	16.00
学生 j	32	34.00	-2.00	4.00
学生 k	38		4.00	16.00
学生 l	32		-2.00	4.00
全体平均	30.42		誤差 SS	64.67

これで分散分析の準備は整った。自由度は、教師の効果が $2 - 1 = 1$、教科書の効果が $2 - 1 = 1$、交互作用の自由度は、二つの効果の自由度を掛け合わせることによって得られるから、教師 df× 教科書 df= $1 \times 1 = 1$ となる。全体の自由度が $N - 1 = 12 - 1 = 11$ なので、誤差の自由度は残りの $11 - 1 - 1 - 1 = 8$ になる。

さあ、分散分析表を作ってみよう。比較すべき平均平方 MS は、SS を df で割ったものである。最後に、教師の MS と誤差の MS の比を考え、$52.083/8.083 = 6.443$ となり、これが F 値である。これがこの自由度 ($F(1, 8)$) で出てくる確率は 0.035% であり、これは危険率 5% よりもちいさいから、効果があるといって良い。同様に、教科書の MS と誤算の MS の比、交互作用の MS と誤差の MS の比で算出される F 値は、危険率よりも大きくなるので、効果があるといってはいけない、という結果になる。

ここでも、全体の分散 130.916 は、効果、交互作用、誤差の分散に分解されていることがわかる ($130.916 = 52.083 + 4.083 + 10.083 + 64.666$)。普通、交互作用の分散は、誤算の計算途中で組み合わせ平均を算出しているプロセスが同じなので、交互作用 = 全体 − 効果 A− 効果 B− 誤差、

[*7] 交互作用 = 組み合わせ平均 −(教師の効果 + 教科書の効果) で算出したのだから、移項して組み合わせ平均 = 教師の効果 + 教科書の効果 + 交互作用である。

表 11　被験者間要因・二要因分散分析の表

要因	SS	df	MS	F	prob.
教師	52.083	1	52.083	6.443	0.035
教科書	4.083	1	4.083	0.505	0.497
交互作用	10.083	1	10.083	1.247	0.297
誤差	64.666	8	8.083		
全体	130.916	11			

として計算されることが多い。しかし計算方法が楽になると、直接的な意味がわかりにくくなるので、本書ではわざと回り道をして計算した。

最後に計算プロセスだけまとめておこう。

1. 全体平均と各水準の平均値を求め、効果を算出し、二乗和して人数倍。これで平方和 (SS) を算出する (表 7、8)。
2. 組み合わせ平均と全体平均に効果を加えたものの偏差平方として、交互作用を算出する (表 9)。
3. 誤差による変動分を算出する (表 10)。
4. 全変動分を算出して、検算しておく。

$$\sum (a_{ijk} - \bar{a}_{...})^2 = 130.916$$

5. 分散分析表を書く。

また、これらの分割ができることは、分散分析の限界を示していることにもなる。実験を計画するとき、色々な要因を考えてみたくなるのが人の常だが、全体の分散はひとつしかない。要因を増やすと、ひとつの全分散を細かく分割していくことになり、どうしてもひとつの要因が占めることのできる分散が小さくなる。経験的に、実験は四要因が限界であると言われている。また、交互作用は様々な組み合わせが考えられるから、例えば A,B,C の三要因の実験であれば、$A \times B$、$A \times C$、$B \times C$、$A \times B \times C$ の四種類がある。四要因計画になれば、$A \times B$、$A \times C$、$A \times D$、$B \times C$、$B \times D$、$C \times D$、$A \times B \times C$、$A \times B \times D$、$B \times C \times D$、と九種類に増える。主効果と誤差を含めると十四種の分散に分解することになるわけで、相対的に各分散が小さくなり、有意な結果が出にくい。交互作用は解釈も難しく、三次以上の交互作用は考えるだけでも頭が痛くなる。

これらの点をよく踏まえて、しっかりとした要因計画を練るようにしよう。

2　被験者内計画

2.1　一要因の分散分析

次に対応のある場合の分散分析を考えよう。対応がある、すなわち被験者内計画は、一人の人間が二回以上、要因の影響を受けるような実験計画である。

例えば、学生の成績の変化を見たいとしよう。まず学期の始めにテストをし、次に中間テスト、最後に期末テストをしたとする。この間の成績の推移に統計的な意味があるかどうか、を検定するといった場合がこれに当たる。有意な結果であれば、教師の教育効果が目に見えて現れたということになろう。

例として、表 12 のデータを用いる。

表 12　一要因・被験者内計画の例

被験者	前期	中期	後期
学生 a	25	25	32
学生 b	23	30	29
学生 c	22	25	32
学生 d	26	30	32
学生 e	28	23	35
学生 f	32	34	37
各期の平均	26.00	27.83	32.83
全体平均	28.89	28.89	28.89
差	-2.89	-1.06	3.94

表中に既に書いておいたが、このときの全体平均は 28.89 点である。各テスト時の平均点とその差がそれぞれ、–2.89、–1.06、3.94 となっている。これは前期や中期は平均よりも低いが、後期は平均よりも上がっていることを示している。右肩上がりなので、なんだか教育効果がありそうだ。

さて、効果の大きさを算出したいので、符号をなくすために平均との偏差を二乗し、六人のデータから作られたことを考慮して、六倍するとしよう。すると前期の変動分は $(-2.89)^2 \times 6 = 50.07$ であり、同様に中期は 6.69、後期は 93.35 となる。この三つの数字を足したもの、すなわち 150.11 が要因の平方和である (表 13)。

被験者内計画はここからが特殊である。誤差の算出にあたり、個々人の個体差というものを考慮しなければならない。例えば、学生 a は学期を通じて平均点が 27.33 点であった。また学生 f は 34.33 である。もともと f 君は出来がよいようだが、「学期を通じての効果」を考えるときは、この個人平均に対する効果を見るべきであろう。つまり、a 君の後期得点 32 点は、彼の平均点

表 13　内要因でも効果の算出方法は同じ

群	平均	全体平均	効果	二乗	人数倍
t1	26.00	28.89	-2.89	8.35	50.07
t2	27.83	28.89	-1.06	1.11	6.69
t3	32.83	28.89	3.94	15.56	93.35
				SS	150.11

表 14　個人平均と全体平均からの偏差

個人平均	全体平均との差	二乗	回数倍
27.33	-1.56	2.42	7.26
27.33	-1.56	2.42	7.26
26.33	-2.56	6.53	19.59
29.33	0.44	0.20	0.59
28.67	-0.22	0.05	0.15
34.33	5.44	29.64	88.93
		個人差 SS	123.78

27.33 から +4.67 であるのに対し、f 君の後期得点 37 点は、彼の平均点 34.33 から +2.67 なのである。これを見ると、a 君は後期にグッと伸びているのだから、時系列的な効果は大きかったと言えるだろう。被験者間計画とは違って、誤差が全体平均ではなく、個人平均に対して生じるものと考える必要があるのだ。

それでは各人の平均を算出してみよう。表 14 に、個人の平均点と、その全体平均 28.89 との差を算出してみた。ここに現れる変動分は、個人差であって、効果とは関係のない変動分である。この偏差を二乗し、さらに前・中・後の三段階による変化だから三倍たもの (表中の 123.78)、これが個人差による変動分である。

さて、では誤差はどのように算出するのであろうか。これはもちろん、個人差や要因の効果で説明できない分になる。例えば学生 a の前期の成績は、全体平均 + 前期の効果 + 個人差 = 28.89 − 2.89 − 1.56 = 24.44 であるはずだった。が、実際には 25 点である。このズレ、すなわち 24.44 − 25 = 0.556 が取り出すべき誤差なのだ。こうして、各人・各時点での誤差が算出される (表 15)。これを平方和したもの (67.889) が誤差による変動分である。

表 12 の素点と全体平均との偏差平方和は 341.78 であるが、これはこうして、効果 150.11 と個人差 123.78、そして誤差 67.89 に分解されたわけだ。

さあ、分散分析表を作ろう (表 16)。

いつも面倒な自由度。まず個人差は六人で作った変動なので、6 − 1 = 5 である。要因は三水準だったから、3 − 1 = 2 である。誤差は六人が三回の計測の中で生み出したものだから、

表 15　誤差を算出

前期	中期	後期
0.556	-1.278	0.722
-1.444	3.722	-2.278
-1.444	-0.278	1.722
-0.444	1.722	-1.278
2.222	-4.611	2.389
0.556	0.722	-1.278
	誤差 SS	67.889

表 16　被験者内要因・一要因分散分析の表

要因	SS	df	MS	F 値	prob
個人差	123.78	5	24.76		
テストをした時期	150.11	2	75.06	11.056	0.003
誤差	67.89	10	6.79		
全体	341.78	17			

$(6-1)\times(3-1)=10$ となる。わかりにくければ、全体の自由度 $6\times3-1=17$ から、個人差や要因の自由度を引いたと考えても良い。

さて、検定においては、個人差を見たいのではなくて、前期・中期・後期という時期の効果 (時系列的変化) を見たかったのだから、その変動 MS75.06 と誤差の変動 MS6.79 の比、11.056 が理論的 F 分布 (自由度 $F(2,10)$) のなかでどれぐらいの出現確率かを調べることになる。結果、5% の危険率より下回っているので、有意差が認められた。つまり、時期による効果があったことになり、各時点での平均値から、だんだん成績が伸びていることが統計的に確認された、となる。

最後に計算プロセスだけまとめておこう。

1. 全体平均と各水準の平均値を求め、効果の大きさを算出する (表 13)。
2. 個人差を算出する (表 14)。

$$\{(-1.56)^2+(-1.56)^2+\cdots+(5.44)^2\}\times3=123.78$$

3. 誤差による変動分を算出する (表 15)。

$$(0.556)^2+(-1.278)^2+\cdots+(0.722)^2+(-1.278)^2=67.889$$

4. 全変動分を算出して、検算しておく。

$$\sum(a_{ijk}-\bar{a}_{\dots})^2=341.78$$

5. 分散分析表を書く。

このように、被験者内の分散分析は誤差の算出がこれまでと異なる。二要因以上になると、交互作用も考えなければならないので、更に複雑になってくる。

2.2　二要因の分散分析

では二要因の例を考えてみよう。あるクラスを実験対象にして、一学期の前半を教師 A が、後半を教師 B が教えるとする。教師が教えに来る前と後でテストをして、成績の変化を見たいとしよう。要因は教師と教育効果 (事前・事後の変化) である。表 17 のような結果であったとしよう。全体の平均点は 29.13 点で、各列の平均点は表中に記してある。ここから、各列の全体平均からの偏差が求まるし、これを平方して人数分をかけたもの、すなわち $\{(-1.29)^2 + (0.71)^2 + (-0.79)^2 + (1.38)^2\} \times 6 = 28.125$ が要因による変動分だ。

表 17　二要因・被験者内計画のデータ例

	教員 A		教員 B	
被験者	事前	事後	事前	事後
学生 a	22	27	30	31
学生 b	27	32	25	27
学生 c	25	25	29	32
学生 d	30	28	26	29
学生 e	31	33	29	32
学生 f	32	34	31	32
平均	27.83	29.83	28.33	30.50
全体平均からの偏差	-1.29	0.71	-0.79	1.38

しかし、注意して欲しいのは、これには教員の変動と教育効果の変動、および交互作用が含まれているということである。

表 18 にまとめたように、それぞれの要因による変動分というのが算出できる。教員 B の方が教え方がうまいようだし、どちらの教員についたにしろ、教育効果が出ている (事後テストの方が得点が高い) ことがわかる。

これらの数値を平方し、人数倍 (六人が二回なので 12 人であることに注意！) すると、教員の要因による変動は $(-0.29)^2 \times 12 + (0.29)^2 \times 12 = 1.02 + 1.02 = 2.04$ となる。同様に、教育効果は $(-1.04)^2 \times 12 + (1.04)^2 \times 12 = 13.02 + 13.02 = 26.04$ である。

また、この二つの効果による変動分、$2.04 + 26.04 = 28.08$ は、全体の変動 28.125 よりも少ない。この差分、$28.125 - 28.08 = 0.04$ が交互作用による変動分である[*8]。もちろん、このような

[*8] この等式は正確ではないが、小数点三桁以下を丸めたことによる誤差である。

表 18　各要因による変動分

	平均値	効果	二乗	人数倍
教員Ａ平均	28.83	-0.29	0.09	1.02
教員Ｂ平均	29.42	0.29	0.09	1.02
			SS	2.04

	平均値	効果	二乗	人数倍
事前テストの平均	28.08	-1.04	1.09	13.02
事後テストの平均	30.17	1.04	1.09	13.02
			SS	26.04

手抜き計算をしないで、表 19 にあるように、各組み合わせの平均値を算出し (この場合列平均)、要因の効果との差を求め (表 19 中の「残差」)、それを二乗し、足し合わせたもの (合計) に人数 6 を欠けたもの (平方和、SS) として算出しても同じ結果になる。

表 19　正しい交互作用の算出方法

全体平均	教師	教育効果	結果	列平均	偏差	平方	人数倍
29.13	-0.29	-1.04	27.79	27.83	-0.04	0.002	0.01
29.13	-0.29	1.04	29.88	29.83	0.04	0.002	0.01
29.13	0.29	-1.04	28.38	28.33	0.04	0.002	0.01
29.13	0.29	1.04	30.46	30.50	-0.04	0.002	0.01
						SS	0.04

さて次に、個人差を算出しよう。個人ごとの変動は誤差ではないので、これを差し引いておく必要があった。

各個人の平均値を算出し、それと全体平均との偏差平方を求める。このとき、四回 (二水準 × 二水準) の分だけ影響されているはずだから、偏差平方を四倍しておくことを忘れずに。これを表したのが表 20 である。これを総和したものが個人差で、今回は 85.88 である。

二要因以上になると次がややこしい。なぜかというと、誤差がひとつではないからである。すなわち、教師要因に伴って生じる誤差、教育効果要因によって生じる誤差、交互作用によって生じる誤差、があるからだ。

順に見ていこう。教師要因に伴って生じる誤差を算出したい。まず、それぞれの教師において事前・事後のテストが行われているが、この二つのテストの平均点を算出する (表 21)。ここで出てきた数値と全体平均との偏差、およびその平方は、「事前・事後の教育効果」を排した、教師要

表 20　個人差を取り出す

個人平均	全体平均	偏差	二乗	回数倍
27.5	29.13	-1.63	2.64	10.56
27.75	29.13	-1.38	1.89	7.56
27.75	29.13	-1.38	1.89	7.56
28.25	29.13	-0.88	0.77	3.06
31.25	29.13	2.13	4.52	18.06
32.25	29.13	3.13	9.77	39.06
			個人差	85.88

表 21　教育効果の要因を潰した表

被験者	教師 A 平均	教師 B 平均
学生 a	24.50	30.50
学生 b	29.50	26.00
学生 c	25.00	30.50
学生 d	29.00	27.50
学生 e	32.00	30.50
学生 f	33.00	31.50

因だけに関連する誤差だと考えられる。これを算出し、教育効果要因である前・後の二水準を潰した分を二倍することでカヴァーした数値、171.13 が、教師要因だけから算出される変動分である (表 22)。

もっとも、この 171.13 には、教師要因の効果と個人差が含まれている。そこで、これらの変動を差し引くと、171.13 – 2.04 – 85.88 = 83.21 が残る。これが教師要因に伴って生じる誤差である。

同様に、教師要因を潰して教育効果要因だけの変動分を算出し、教育効果だけに伴って生じる誤差を算出しよう (表 23)。変動全体の平方和は 118.13 だが、個人差と効果を引いた残りは 6.21 であり、これが誤差である。

最後に交互作用に伴って生じる誤差がある。これはもう、ほとんど全てのデータが出そろっているので、演繹的に求めるのがいいだろう。

まず、各セルと全体平均との偏差を求める。これは個人差を始め、全ての誤差を含んだものになる (表 24)。この偏差平方和、190.50 は個人差、教師要因の誤差、教育効果要因の誤差、交互作用の誤差を含んでいるので、差し引くことによって交互作用の誤差が求められる。190.50 – 85.88 – 83.21 – 6.21 = 15.21 である。もちろん、組み合わせ平均と各セルとの偏差平方和を求め、そこから交互作用の効果、個人差、教師要因の誤差、教育効果要因の誤差を引いたも

表 22　教師要因だけの変動や誤差を取り出す

教師 A 偏差	教師 B 偏差	教師 A 平方	教師 B 平方
-4.63	1.38	21.39	1.89
0.38	-3.13	0.14	9.77
-4.13	1.38	17.02	1.89
-0.13	-1.63	0.02	2.64
2.88	1.38	8.27	1.89
3.88	2.38	15.02	5.64
		平方和	85.56
		二倍	171.13
		教師要因についての誤差	83.21

表 23　教育効果だけの変動や誤差を算出

被験者	事前	事後	事前偏差	事後偏差	事前偏差	事後偏差
学生 a	26.00	29.00	-3.13	-0.13	9.77	0.02
学生 b	26.00	29.50	-3.13	0.38	9.77	0.14
学生 c	27.00	28.50	-2.13	-0.63	4.52	0.39
学生 d	28.00	28.50	-1.13	-0.63	1.27	0.39
学生 e	30.00	32.50	0.88	3.38	0.77	11.39
学生 f	31.50	33.00	2.38	3.88	5.64	15.02
					平方和	59.06
					二倍	118.13
					教育効果についての誤差	6.21

のとしても算出できる。

こうして全ての変動分が揃った。では分散分析表を作成しよう (表 25)。自由度は、個人差が人数 -1 で 5、教師と教育効果は二水準 -1 で 1、交互作用は $1 \times 1 = 1$ である。要因に対する誤差は人数 -1 で 5、全体の自由度が 24 セル -1 で 23 なので、交互作用の自由度が残る 5 となる。

注意すべきは、それぞれの要因と、それに伴って起きる誤差とで比較することである。教師の要因は、教師の誤差と比べて十分大きかったかどうか。教育効果はその誤差と比べてどうだったか。交互作用はどうか。である。こうしてみると、教育効果は危険率より低いので、前後の効果はあるんだな、ということがわかる。教師は効果がないので、別にどちらの教員が優れているというわけではないのだ。ともかく教育をすれば、成績が上がるという結果になっている。

最後に計算プロセスだけまとめておこう。

表 24　各セルと全体平均との偏差

-5.83	-2.83	1.67	0.50
-0.83	2.17	-3.33	-3.50
-2.83	-4.83	0.67	1.50
2.17	-1.83	-2.33	-1.50
3.17	3.17	0.67	1.50
4.17	4.17	2.67	1.50

表 25　二要因・被験者内計画の分散分析表

要因	SS	df	MS	F 値	Prob
個人差	85.875	5	17.175		
教師	2.042	1	2.0417	0.122684	0.7404
教師誤差	83.208	5	16.642		
教育効果	26.042	1	26.042	20.97316	0.0059
教育効果誤差	6.208	5	1.2417		
交互作用	0.042	1	0.0417	0.013702	0.9114
交互作用誤差	15.208	5	3.0417		
全体	218.625	23			

1. 効果を算出し、全要因による変動分を算出する (表 17)。

$$\{(-1.29)^2 + (0.71)^2 + (-0.79)^2 + (1.38)^2\} \times 6 = 28.125$$

2. 各要因による変動と交互作用による変動を分離する。

$$(-0.29)^2 \times 6 + (0.29)^2 \times 6 = 2.04$$

$$(-1.04)^2 \times 6 + (1.04)^2 \times 6 = 26.04$$

$$28.125 - 2.04 - 26.04 = 0.04$$

3. 個人差を算出する (表 20)。

$$p \times q \times \sum_{i=1}^{n} \sum_{t_1=1}^{p} \sum_{t_2=1}^{q} (a_{it_1t_2} - \overline{a_{i..}})^2 = \{(-5.50)^2 + (-0.50)^2 + \cdots + (-0.25)^2\} \times 4 = 85.88$$

4. 一方の要因を潰した表を作り (表 21)、ある要因だけから生じる変動分を算出する。

$$\{(-4.63)^2 + (1.38)^2 + (0.38)^2 + \cdots + (2.38)^2\} \times 2 = 171.13$$

5. これから要因の効果と個人差を引いたものが、その要因に対する誤差として算出できる。

$$171.13 - 2.04 - 85.88 = 83.21$$

6. 同様に、もう一方の要因についても誤差を計算する。

$$118.13 - 26.04 - 85.88 = 6.21$$

7. 全変動分を算出する。

$$\sum (a_{ijk} - \bar{a}_{...})^2 = 190.50$$

8. 交互作用に伴って生じる誤差を算出する。

$$190.50 - 85.88 - 83.21 - 6.21 = 15.21$$

9. 分散分析表を書く。

被験者間計画以上に、内計画は誤差の分解で苦労することが理解できただろうか。分散分析は、ともかく分散を分解していくので、あまり要因や水準を増やすべきではない。これはどのような計画でも同じことである。

ここまでの分析プロセスを、まとめておく。

被験者間計画

一要因の場合

$$\text{全変動} = \text{要因} + \text{誤差}$$

二要因の場合

$$\text{全変動} = \text{要因 } A + \text{要因 } B + \text{交互作用} + \text{誤差}$$

被験者内計画

一要因の場合

$$\text{全変動} = \text{要因} + \underbrace{\text{個人差} + \text{誤差}}_{\text{偶然誤差}}$$

二要因の場合

$$\text{全変動} = \text{要因 } A + \text{要因 } B + \text{交互作用}$$
$$+ \underbrace{\text{個人差} + \text{要因 } A \text{ の誤差} + \text{要因 } B \text{ の誤差} + \text{交互作用の誤差}}_{\text{偶然誤差}}$$

ああ、ややこしい。

さらに、内要因と間要因が組み合わさることだってあるのだ。次はそれについてみてみよう。

3　混合計画

要因 A が被験者間計画、要因 B が被験者内計画だったとしよう。このように、間と内の両方のデザインが入った実験計画は、特に混合計画という。この場合、データはどのように分解されるかというと、

$$\text{全変動} = \underbrace{\text{要因 } A + \text{被験者間の偶然誤差}}_{\text{被験者間に生じる変動}}$$

$$+ \underbrace{\text{要因 } B + \text{交互作用} + \text{被験者内の偶然誤差}}_{\text{被験者内に生じる変動}}$$

となる。

例をあげてみてみよう。クラス A、B はそれぞれ 6 人からなる学級で、ここで学期前、学期末の前後でテストをしたとしよう (表 26)。各列の平均は表中に示した通り。全体平均は 31.71 点で

表 26　混合計画のデータ例

被験者	クラス	事前	事後
学生 a	A	28	33
学生 b	A	29	27
学生 c	A	29	30
学生 d	A	25	31
学生 e	A	30	30
学生 f	A	29	28
学生 g	B	28	32
学生 h	B	30	32
学生 i	B	35	35
学生 j	B	32	41
学生 k	B	38	36
学生 l	B	32	41
		30.42	33.00

ある。まず、各要因ごとの平均値を算出して、全体平均とのズレから、要因がどれほどの変動を生み出したのかを見ることができる (表 27)。A 組は平均より低いし、その分 B 組は平均より高い。テストは授業後にする方が点数がいい。この変動分ー全体平均からの偏差は、平方和にしてデータの数 (二水準 ×6 人) 分膨らませてやろう。これが各要因の効果で、クラスに分けたことによる変動分が $(-2.63)^2 \times 2 \times 6 + (2.63)^2 \times 2 \times 6 = 165.375$ である。同様に、時間的前後の効果は $(-1.29)^2 \times 2 \times 6 + (1.29)^2 \times 2 \times 6 = 40.042$ である。

表 27　各要因による変動分

群	平均	全体平均	効果	二乗	人数倍
A 組	29.08	31.71	-2.63	6.89	82.69
B 組	34.33	31.71	2.63	6.89	82.69
					165.38

群	平均	全体平均	効果	二乗	人数倍
事前	30.42	31.71	-1.29	1.67	20.02
事後	33.00	31.71	1.29	1.67	20.02
					40.04

ここまでは今まで通り。次に、被験者間に生じる誤差を考えよう。これは換言すれば、被験者内の効果を消し去るということだから、前後の平均値をデータにした一要因被験者間データのようにして考える。表 28 に示したのが、学生の前後平均と、クラスの平均値である。この偏差が

表 28　間要因に生じる誤差

被験者	前後平均	クラス平均	偏差	平方
学生 a	30.50	29.083	1.4167	2.0069
学生 b	28.00	29.083	-1.0833	1.1736
学生 c	29.50	29.083	0.4167	0.1736
学生 d	28.00	29.083	-1.0833	1.1736
学生 e	30.00	29.083	0.9167	0.8403
学生 f	28.50	29.083	-0.5833	0.3403
学生 g	30.00	34.333	-4.3333	18.7778
学生 h	31.00	34.333	-3.3333	11.1111
学生 i	35.00	34.333	0.6667	0.4444
学生 j	36.50	34.333	2.1667	4.6944
学生 k	37.00	34.333	2.6667	7.1111
学生 l	36.50	34.333	2.1667	4.6944
			偏差平方和	52.542
			誤差	105.083

クラス要因に伴って生じる誤差なのだから、偏差平方和を算出する。これが 52.542 となるが、前後の変動を潰して過小評価している分、二倍する必要があろう。出てきた 105.083 が間要因に伴う誤差である。

では次に、交互作用による変動分を算出しよう (表 29)。A 組は平均より −2.63 点になるものだし、事前のテストは −1.29 点の変動があるはずだ。全体平均が 31.708 だから、両方の効果だけなら 31.708 − 2.63 − 1.29 = 27.792 点になるはずだ。しかし、A 組の前テスト平均を算出すると 28.333 点である。この差分こそ、交互作用による影響である。偏差平方和を算出し、各セクション 6 名であったことからこれを六倍して、交互作用による SS の 7.0417 が得られる。これが交互作用による変動分である。

表 29　交互作用を算出

	効果による値	組み合わせ平均	偏差平方	人数倍
A 組事前	27.79	28.33	0.29	1.76
A 組事後	30.38	29.83	0.29	1.76
B 組事前	33.04	32.50	0.29	1.76
B 組事後	35.63	36.17	0.29	1.76
			交互作用 SS	7.04

さて、最後は被験者内の誤差である。それ以外のデータが揃っているので、全変動からの差分としてこれを算出することができる。表 30 に、各セルの全体平均との偏差を示した。これの平方和、396.958 がこのデータから得られる全変動である。ここから、クラス要因の効果、前後要因の効果、交互作用、被験者間に生じる誤差を引くと、396.958 − 165.375 − 40.041 − 7.041 − 105.083 = 79.416 であり、これが求めているものである。

表 30　全変動の算出

被験者	クラス	事前	事後
学生 a	A	-3.708	1.292
学生 b	A	-2.708	-4.708
学生 c	A	-2.708	-1.708
学生 d	A	-6.708	-0.708
学生 e	A	-1.708	-1.708
学生 f	A	-2.708	-3.708
学生 g	B	-3.708	0.292
学生 h	B	-1.708	0.292
学生 i	B	3.292	3.292
学生 j	B	0.292	9.292
学生 k	B	6.292	4.292
学生 l	B	0.292	9.292

さあ分散分析表を書こう (表 31)。いつも悩ましい自由度だが、クラスは二水準 −1 で 1、各ク

表 31　混合計画の分散分析表

要因	SS	df	MS	F 値	prob
クラス	165.375	1	165.375	15.738	0.003
被験者間の誤差	105.083	10	10.508		
前後	40.041	1	40.041	5.042	0.049
交互作用	7.041	1	7.041	0.887	0.369
被験者内の誤差	79.416	10	7.942		
全体	396.958	23			

ラス 6 名で、$(6-1)+(6-1)=10$ が間要因の自由度である。内要因の自由度は今回二水準 −1 で 1、交互作用は 1×1 で 1 である。全体の自由度が 24 セル $-1 = 23$ だから、交互作用の自由度は 10 である。

注意すべきは、間要因の MS は間誤差の MS と比較すること。また、内要因の MS と交互作用 MS は内誤差の MS と比較することである。それぞれどの誤差と比べて効果があったのか、をはっきりさせなければばならないのだ。

最後に計算プロセスだけまとめておこう。

1. 効果を算出し、要因による変動分を算出する (表 27)。

$$(-2.63)^2 \times 2 \times 6 + (2.63) \times 2 \times 6 = 165.375$$

$$(-1.29)^2 \times 2 \times 6 + (1.29) \times 2 \times 6 = 40.042$$

2. 内要因を潰した表から、間要因に伴って生じる誤差を計算する (表 28)

$$p \times \sum_{j=1}^{m}\sum_{i=1}^{n}(\overline{a_i j.} - \overline{a_i..})^2 = 2 \times 52.542 = 105.083$$

3. 交互作用を算出する (表 29)。

$$\{(-0.542)^2 + (0.542)^2 + (0.542)^2 + (-0.542)^2\} \times 6 = 7.0417$$

4. 全変動分を算出する。

$$\sum(a_{ijk} - \overline{a}_{...})^2 = 396.958$$

5. 内要因に伴って生じる誤差を算出する。

$$396.958 - 165.375 - 40.041 - 7.041 - 105.083 = 79.416$$

6. 分散分析表を書く。

これより複雑な実験計画は、全てここまで論じてきた算術方法の応用でできる。とはいえ、なかなか全ての計算を自由自在に操るのは難しい。実際問題は、分散分析を計算してくれるアプリケーションの助けを借りることになるだろう。ただし、原理として「分散を分解していくこと」や「対応する誤差との比で検定すること」などは、ユーザーとして最低限理解しておかなければならない。