

## 9.1 被験者内(対応のある)1 要因分散分析

先週の例から発展させて、「対応のある」条件設定がおこなわれている(すなわち、被験者内要因計画の)実験をおこなった場合の分析はどうなるかを考えてみよう。

	実施時点ごとのテストの得点		
	トレーニング前 (test1)	3ヶ月後 (test2)	半年後 (test3)
1	30	52	73
2	32	48	75
3	25	50	73
4	28	49	76
5	31	55	80
6	35	58	72
7	30	50	80
8	27	47	78

先週の分散分析の結果、タイピング能力の向上にもっとも効果があるトレーニングメソッドはメソッド B であることがわかった。そこで、今度は1~8までの8名の被験者に、メソッド B を用いたトレーニングを半年間おこなった。

左表は、トレーニング前(test1)、トレーニング開始 3ヶ月後(test2)、そして半年後(test3)のタイピングテストの得点を示している。トレーニングによって得点は伸びているだろうか。つまり、長期間のメソッド B によるトレーニングは、効果をもっていたらどうか。

この場合、分析に際して従属変数となるのは、8名が受験した3回のテストの得点である。

先週の例とは異なり、要因の水準ごとのデータは8名について各3回ずつ繰り返して取られたものである。よって、平均値の差の検定で言えば「対応のある」ものと同様の扱いをする必要がある。分散分析の場合はこのような要因のことを被験者内要因と呼ぶ。

被験者内要因を用いた分散分析の場合、条件の効果を想定した要因に加えて被験者要因を組み込む必要がある。なぜならトレーニングの効果を正確に取り出すためには、被験者それぞれがもつであろう、個人内でトレーニング継続時間とは無関係に生じるズレが従属変数の値にもたらす影響を考慮する必要があるからである。被験者内要因の場合、データのズレは、

$$\text{全体のズレ} = \text{偶然のズレ(誤差} = \text{個人内のズレ} + \text{条件内のズレ)} + \text{条件によるズレ}$$

であると考えなければならない。

SAS による被験者内要因の分散分析をおこなう場合、2つの方法がある。いずれも PROC GLM を用いることは同じだが、データの記述スタイルにより、PROC GLM でのオプションの記述の仕方が異なり、結果的に出力も異なるスタイルでなされることになる(結果に違いが生じないのは当然のことである)。

```
data anova2;
input sub test1-test3;
cards;
1 30 52 73
2 32 48 75
3 25 50 73
4 28 49 76
5 31 55 80
6 35 58 72
7 30 50 80
8 27 47 78
;
proc glm;
model test1-test3= /ss3;
repeated test 3;
run; quit;
```

## 1) 被験者 1 人を 1 オブザベーションとみなすデータ記述の場合

これまでと同じように、被験者それぞれに番号を振り、個人 = 1 行でデータを記述する場合は、次のようなプログラムを書くことになる。

**class オプション**: なし

要因として指定しうる変数(テストをおこなった時点)データの中に記述されていないことに注意しよう。

**model オプション**

従属変数は test1 から test3 までの 3 時点でのテスト得点である。前述したように要因として指定しうる変数がデータの中に記述されていないので、右辺は何も書かなくてもよい。

**repeated オプション**: 繰り返し測定されている従属変数とその回数を指定

この形のデータの場合、このオプションで繰り返し要因とその水準数を明示する必要がある。左例では test\* という変数が 3 回繰り返し、という記述である。このように、繰り返し測定の従属変数は規則性のある命名をする方がよい。

## output の見方

たくさんのアウトプットが出力されて面食らうかもしれないが、重要なところは次の箇所である。

## The GLM Procedure

## Repeated Measures Analysis of Variance

## Univariate Tests of Hypotheses for Within Subject Effects

Source	DF	Type III SS	Mean Square	F Value	Pr > F	Adj Pr > F	
						G - G	H - F
test	2	8525.250000	4262.625000	423.99	<.0001	<.0001	<.0001
Error(test)	14	140.750000	10.053571				

1 行目: 被験者内要因 test (テストをおこなった時点) に関する自由度, 平方和, 平均平方, F 比および p 値

2 行目: test (テストをおこなった時点に関する) 誤差に関する自由度, 平方和

つまり, テストをおこなった時点により, 得点に有意な差があると言うことができる。

**問題点**

ただし, このデータ記述 & プログラムのスタイルの場合, 被験者要因に関する統計量は出力されない。また, 従属変数がそれぞれ異なる変数となっているため, means オプションを用いて多重比較をおこなうこともできない。前者はともかく, 後者は考察に必須なので,

- 1) SAS に頼らず自分で計算する
- 2) もう一つのデータ入力スタイルによる分析をおこなう

という 2 つからいずれかを選択することになる。ここでは 2) の方法を紹介する<sup>1</sup>。

## 2) do ループを用いたデータの記述方法

```
data anova_within;
  do sub=1 to 8;
    do time=1 to 3;
      input score @@; output;
    end;
  end;
cards;
30 52 73
32 48 75
25 50 73
28 49 76
31 55 80
35 58 72
30 50 80
27 47 78
;
run;
proc print;
run;
```

先ほどのデータ記述方法では, 個人に被験者番号を振り, 各時点でのテスト得点に異なる変数名 (test1 ~ test3) をつけていた。

一方, 左に示すような **do ループ** を用いた記述方法<sup>2</sup> ならば, テスト得点をあらかず score という変数に各実施時点 3 回 (time) のテスト得点を順に割り当てることができるし, 被験者要因も組み込むことが可能である。

- do 文では「ある要因 (この場合は被験者要因と sub, テスト実施時点 time) がいくつの水準から構成されているか (この場合各 8, 3)」を指定する。
- input 文では, 要因のそれぞれの水準で測定されたデータ (この場合は得点 score) を指定する。
- @@; output; はとりあえず「呪文」だから必ずこう書くのだと覚えておこう。
- do 文は必ず最後に do の数だけ end; を持ってくるのがルール。
- このように書けば, 各オブザベーションの最初のデータは sub=1, time=1 の時の score 変数の値, 2 番目は sub=1, time=2 の時, 3 番目は sub=1, time=3 の時と順番に読み込んでくれる。1 行のデータが終了したら次の行に移って同様の順序でデータを読み込んでくれる。
- 外部データファイルを読み込む場合は, do ループの直前に infile ステートメントを置けばよい。

<sup>1</sup> 1) については森・吉田「心理学のためのデータ解析テクニカルブック」(北大路書房)などを参照せよ

<sup>2</sup> なお, このタイプのデータだとデータ記述パターン 1) と同じであるように見える (というか被験者番号の有無以外同じ並び) であるが, それぞれの変数の読み込みスタイルが根本的に異なることに注意しよう

よく分からない人は、このデータがどのように読み込まれているかを、PROC PRINT で確かめるとよい。プログラム中に登場した変数を見ると、分散分析に必要な「要因」「従属変数」の情報がすべて含まれている(と SAS が認識できる)形になっていることが否応なく分かるのではないだろうか。PROC PRINT の出力結果は右のようになる。

このデータ記述スタイルの場合、分散分析をおこなう PROC GLM は次のように記述することになる。

```
proc glm;
class sub time;
model score=sub time /ss3;
means time/tukey;
run; quit;
```

**class ステートメント**で、従属変数であるテスト得点 score の変化を説明する要因(独立変数)を指定する。この場合はテスト実施時点(time)と被験者要因(sub)である。

**model ステートメント**で、従属変数と要因を指定する。

**means ステートメント**で、テスト実施時点ごとの得点の平均値を求め、その差を比較(多重比較)する。tukey オプションは検定方法を指定している。

OBS	sub	time	score
1	1	1	30
2	1	2	52
3	1	3	73
4	2	1	32
5	2	2	48
6	2	3	75
7	3	1	25
8	3	2	50
9	3	3	73
10	4	1	28
11	4	2	49
12	4	3	76
13	5	1	31
14	5	2	55
15	5	3	80
16	6	1	35
17	6	2	58
18	6	3	72
19	7	1	30
20	7	2	50
21	7	3	80
22	8	1	27
23	8	2	47
24	8	3	78

#### output の見方

出力パターンは被験者間要因の場合と同じだが、被験者要因が増えていることに注意しよう

#### 主効果の検定

Source	DF	Type III SS	Mean Square	F Value	Pr > F
sub	7	94.500000	13.500000	1.34	0.3018 ... 被験者要因
time	2	8525.250000	4262.625000	423.99	<.0001

この結果から、テスト実施時点の効果は有意だが、被験者要因は有意ではない(すなわち個人内のばらつきはないとみなしてよい)ことがわかる。また、前述のデータ記述スタイル 1)の時の出力結果と見比べて、平方和、平均平方、F 比の値などが同じであることを確認しよう。また、自由度についても確認しよう。<sup>3</sup>

テスト実施時点のどことどの間に有意差があるのか、詳しく見るために多重比較の結果を見てみよう

#### 多重比較

	Mean	N	time
A	75.875	8	3
B	51.125	8	2
C	29.750	8	1

異なるアルファベットがマーキングされている群に有意差があると判断できる。つまりこの場合は、「トレーニング開始前」「3 ヶ月後」「半年後」のいずれの間にも有意差があるといえ、その得点は時間経過と共に向上していることがわかる。つまり、実験の結果、長期間トレーニングメソッド B を実施することの効果があがっていることが示されたといえる。

<sup>3</sup> 全体の自由度は、(被験者数 × 水準数) - 1 = 23、誤差の自由度は (被験者数 - 1) × (水準数 - 1) = 14 である

## 結果の書き方

分析が首尾良くできたとしても、レポートに必要な事項を正しく記述できなければ意味がない。ここでは、心理学分野における標準的な分散分析の結果記述方法を紹介する。

## 結果

タッチタイピング能力向上のために開発された3種のトレーニングメソッドのうち、予備実験でもっとも効果があることが確認されたメソッド B について、長期的な能力向上効果があるかどうかを検証するための実験結果について、被験者内1要因分散分析(3水準)をおこなった。

分散分析表を表1に示し、3時点(トレーニング前、トレーニング開始3ヶ月後、6ヶ月後)における参加者のタイピングテスト得点を図1に示す。

表1 分散分析表:タイピングメソッドの長期的能力向上効果

要因	自由度	平方和	平均平方	F値	有意水準
テスト時点	2	8525.25	4262.63	423.99	$p<.0001$
被験者	7	94.50	13.50	1.34	n.s.
誤差	14	140.75	10.05		
全体	23	8760.50			

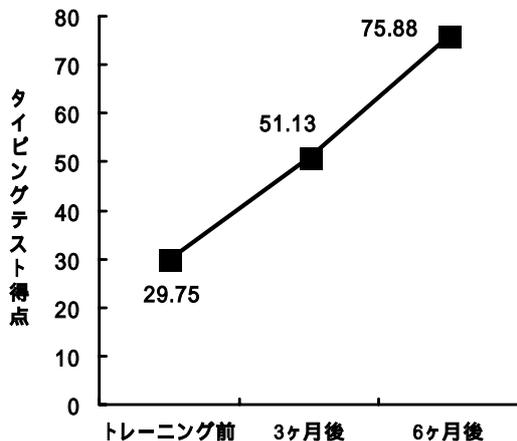


図1 3時点のタイピングテスト得点

分散分析の結果、テスト時点による有意な主効果( $F_{(2,14)}=423.99$ ,  $p<.0001$ )が見られた。Tukey法による多重比較の結果、3時点いずれの間にも有意差( $p<.05$ )が見られ、図1に示したように、メソッド B にはトレーニングを継続することによる長期的なタイピング能力向上効果が期待できることが明らかになった。

## 課題1

講義 Web に示された課題について、分散分析をおこなって、結果を比較せよ

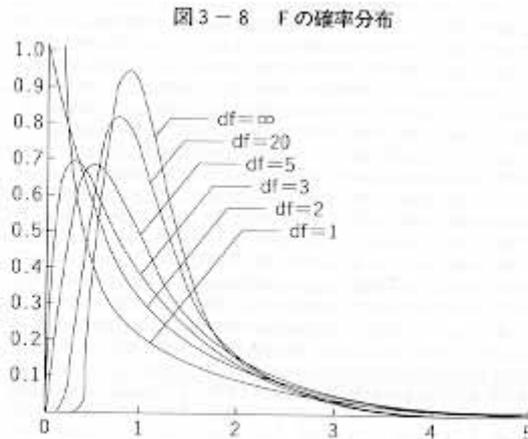
## 課題2

課題1の2種類の分散分析のうち、被験者間要因とみなして分析した結果について、上記例を参考にして文章と図表で結果を記述せよ(結果の記述はまったく同じというわけではない、どこに注意すればよいだろうか?)。

## 補足:なぜF分布?

先週解説したように、分散分析とは、従属変数における「要因(条件)による値のズレ」と「偶然による値のズレ」を比べることによって、要因(条件)の効果の有無を考えるものです。このとき検定の対象となる値は「条件による値のズレ」=「要因(条件)の平均平方」を「偶然による値のズレ」=「誤差の平均平方」で割った比の値です。

では比の値をなぜF分布にもとづいて検討するのでしょうか？ まず F 分布とはどういう統計分布であるかを思い出しましょう。



F分布とは、2つの不偏分散(平均からの差の自乗和をn-1で割ったもの)の比の確率分布です。Fは有名な統計学者 Fisher の頭文字を取って名付けられました。

上図にあるように、F分布は、正規分布やt分布のように左右対称の分布は取りません。分子に分母より大きい値が来る場合と、小さい値が来る場合が半々の確率で起きるはずなので、1 を中心にほぼ半々の出現確率になります。さらに、比の分布なので、分子、分子それぞれの値が異なることによっても分布の形が異なります。つまり、分布の形を決める値 = 自由度は2つあることになります。上図は、分子のサンプル数にもとづいて定まる自由度が10、分母のサンプル数にもとづいて定まる自由度が1～ までの確率分布を示したものです。

分散分析に当てはめて考えてみると、「分子 = 条件によるズレ」と「分母 = 偶然によるズレ」からなる比の値について、自由度(分子のサンプル数による自由度、分母のサンプル数による自由度)の F 確率分布のどこに位置するかによって、その比の値の大きさの有意性を検定することになります。比の値が1となる、すなわち分母と分子が等しい確率は50%で、分子の方が有意に大きくないと、要因(条件)の差があるとは言えません。

つまり、1 を中心とする確率分布の右側半分 (= 分子の方が大きい場合) で、全体の 0.05 にあたる面積を占めるような X 軸の値が「有意確率5%」点となります。求められた比の値がこの点よりも大きな値をとっていれば、面積は 0.05 より小さくなる、つまりその要因(条件)の効果は「偶然によるズレと同じ」である確率は5%未満であることが示されたこととなります(いわゆる片側検定)。このことが「要因(条件)による有意な効果がある」ことを示しているのです。