ⒸNSCA JAPAN
Volume 21, Number 2, pages 9-13

特集 feature

論文の書き方5

分散分析(その1)

星川佳広 CSCS, NSCAジャパン編集委員, 東海学園大学准教授

本稿は、「論文の書き方1：投稿論文を書こう」(2013年10月号)の続編パート4です。NSCAジャパンウェブサイトに掲載中の「投稿論文(事例報告)の書き方」を参照しながら読み進めてください。

【投稿論文(事例報告)の書き方】
NSCAジャパンウェブサイトTOP → ［指導者の育成］ → ［事例報告・研究論文］、「投稿要領」部分

前回解説したt検定は、レギュラーとサブの比較、あるいはトレーニング前後の比較のように、2つのグループ間の差異を検定するときに使うものでした。しかし、私たちS&C専門職にとっては2つ以上のグループ間を比較したいということもよくあります。例えば、ポジション間(FW、MF、DF、GKなど)、学年間(高1、高2、高3など)、シーズン間(オフシーズン、プレシーズン、インシーズン)で比較したい場合などです。このようなときに便利な検定法が分散分析です。英語では「Analysis of Variance」なので、ANOVAと略して言われます。

分散分析は、t検定と比べると理解するためのハードルがぐっと高くなります。以下では、本連載の目的である「論文を書く」ための入り口としての知識を整理していきますが、短い文章では自ずと説明に限界があります。実際に分散分析を使ってデータ分析する場合は、本記事のみならず次ステップの知識が必要になります。まず今回は、分散分析の基礎と一元配置分散分析について、次回は二元配置分散分析について基本的な事項を解説します。

1. 分散分析の考え方

分散分析にはいくつかの種類があって、データの構造によって利用する方法が異なります[1]。その中で最も基本となり利用頻度も高いのが**一元配置**の分散分析です。一元配置分散分析は、1つの**要因**(要因Ⅰ)により分類された複数のグループ間を比較します。例えば、身長をポジションという要因で比較する場合などです。ポジションという要因の中のFW、MF、DF、GKといった複数のグループは、**水準**といいます。

分散分析の帰無仮説(前号参照)は、例えば「身長はポジション間で差がない」です。つまり、分析しようとしているデータ(標本)が同一の正規分布する母集団から得られたものと仮定したときに、実際に手元にあるデータのようにバラつく確率がどれくらいかを計算するわけです。そしてその確率が5％(もしくは1％)未満しかありえないならば、帰無仮説を棄却し、「身長はポジション間で有意に異なる」と考えるわけです。

表1のポジション間の身長データ例で見てみましょう。FW、MF、DF、GKの4水準に各4名の被検者の身長データがあります。ポジションごとの平均値は、それぞれ177.5、173.3、175.3、180.5cmで、全体の平均値は176.6cmです。

もし帰無仮説が正しいならば、標

1 一元配置、二元配置等の他に、ノンパラメトリックな方法がいくつかあります。ノンパラメトリックな方法は、グループ間の分散が異なる場合や標本データの分布が正規分布とみなせない場合などに使います。

本（各データ）は、176.6cm（全体平均値）を中心とする同一の母集団から抽出されていると仮定されるわけですから、16個（4水準×4名）のデータは176.6を中心に一定の誤差範囲内で適当に（ランダムに）バラついて分布するはずです。しかし今、**表1**のMF、GKそれぞれの4つのデータは、その平均値がそれぞれ173.3、180.5cmであるように、全体平均の176.6cmからは偏って分布しているようにみえます。帰無仮説を前提としたとき、このように偏ってしまう確率がどれくらいなのか？　その確率が5％未満と低ければ帰無仮説を棄却して、ポジション間に有意差ありとしよう、というのが分散分析になります。つまり、分散分析は個々のデータを、「**個々のデータ＝全体の平均＋誤差**」とみるべきか、「**個々のデータ＝全体の平均＋ポジションの効果＋誤差**」[2]とみるべきかの判断基準を与えます。前者はポジションの影響は小さく、個々のデータは単に誤差として、ポジションとは無関係に全体平均まわりにバラついているだけであるとみなします。すなわち、MF、GKの各身長データは、たまたま高かったり、低かったりしただけだということです。一方で後者は身長にポジションの影響が存在するとみなす場合で、個々のデータはポジションの影響を受けた上でバラついていると考えます。ここでポジションの影響があることを、分散分析では「ポジションの効果がある」とか「要因Ⅰの**主効果**」（上記下線部）と呼びます。つまり身長はポジション間に有意差があるということになります。

2. F分布

分散分析で確率を計算する際にはF分布を利用します。今、同じ正規分布する母集団から2組のデータ群（標本）を抽出したとします。そして、それぞれのデータ群の分散（S_1^2,S_2^2）を求め、その比をとります（S_1^2/S_2^2）。この比をF値といいます。F値（$=S_1^2/S_2^2$）は2組のデータ群をどのように抽出するかによって変わるはずです。そこでこの作業を何回も繰り返します。その結果、F分布と呼ばれる分布が出現します。2組のデータ群は同じ母集団から得られているので、$S_1^2=S_2^2$、つまり$S_1^2/S_2^2=1$（**F=1**）になる確率が高いはずです。一方で、F（$=S_1^2/S_2^2$）が1よりずっと大きくなる確率は低いはずです。一元配置分散分析では、個々のデータのばらつき（総変動）を、①要因Ⅰによる変動（グループ間変動）と②誤差による変動（グループ内変動）に分解します。そして、①要因Ⅰによる分散（S_1^2、グループ間分散）と②誤差による分散（S_2^2、グループ内分散）の比（F値）を計算します。そしてF値が1よりずっと大きければ、要因Ⅰによる変動（グループ間変動）が大きい――要因Ⅰの主効果が存在すると考えます。

図1左は、**表1**のデータをグラフ化したものです。一方、**図1右**はポジションごとの平均値は**図1左**と同一ですが、個々のデータはバラつきが大き

表1　ポジション間の身長データ例

ポジション	FW	MF	DF	GK
（cm）	174.0	169.0	172.0	179.0
	177.0	173.0	174.0	180.0
	178.0	175.0	175.0	181.0
	181.0	176.0	180.0	182.0
ポジションの平均	177.5	173.3	175.3	180.5
全体の平均	176.6			

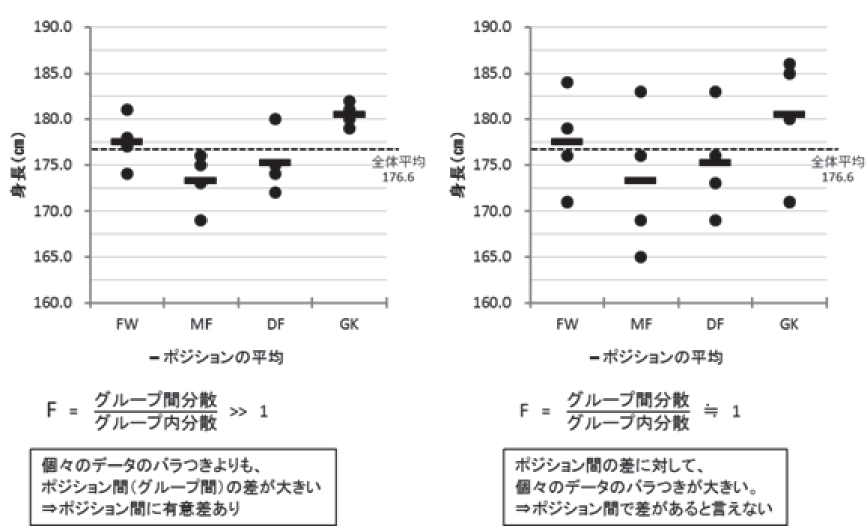

図1　分散分析の考え方

2　例えば最後のデータ（GK4番目）の182.0cmで考えてみましょう。この182.0cmを全体平均176.6cm＋誤差4.4cmとみるべきか、全体平均176.6cm+GKポジションの効果3.9cm（GK平均180.5cm－全体平均176.6cm）＋誤差1.5cmとみるべきか。

くなるよう変更してあります。**図1**の左右2つのグラフの違いは、要因Ⅰの効果(各ポジション平均と全体平均の差異)に対して個々のデータのバラつきが小さいか大きいかです。**図1右**では、要因Ⅰの効果に対して個々のバラつきが大きすぎます。こういう場合、F値は1に近いはずです。つまり、個々のデータは要因Ⅰの効果とは関係がなく、単に誤差によって全体平均のまわりにバラついているだけで、今回はたまたま**図1右**のように分布した――このように分布する確率は、「身長はポジション間に差異がない」という帰無仮説の下ではよくある(5％以上)、とみなします。一方、**図1左**のようにF値が1より大きいということは、個々のデータのバラつきには、要因Ⅰによる変動のほうが誤差による変動よりもずっと大きく影響していることを意味します。帰無仮説の下では、個々のデータがこのように分布する確率はきわめて低い(5％以下)――つまり、個々のデータはポジションの効果が存在した上でバラついており、「身長はポジションで有意に異なる」と判断できるわけです。

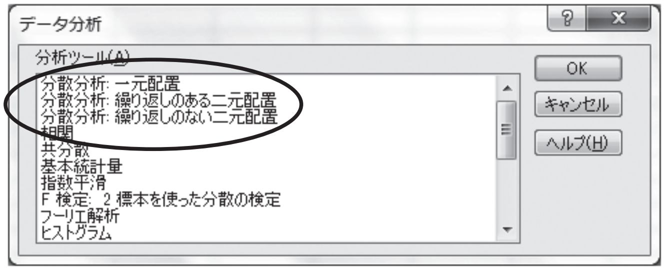

図2　Excelの分析ツール

表2　ワークシート例(一元配置)

	A	B	C
1	レッスンA	レッスンB	レッスンC
2	−1.5	0.7	−1.6
3	−0.2	−0.8	−1.2
4	0.9	−0.1	−1.8
5	0.3	0.7	−0.7
6	0.1	−0.9	0.2
7	−0.7	−0.2	0.5
8	−0.3	−0.9	−2.1
9	0.5	−0.4	−1.8
10		−1.6	−2
11		−1.2	

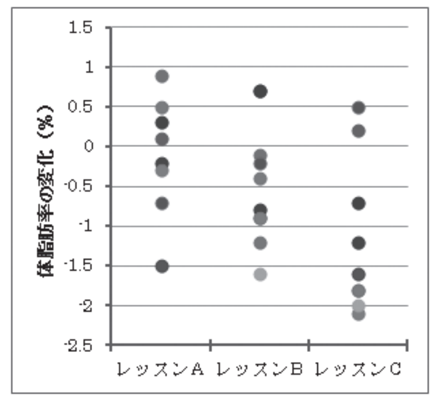

図3　表2データのグラフ

3. Excelによる分散分析

Excelには、分散分析を行なうためのアドイン(追加)機能「分析ツール」が用意されています。分析ツールを利用すれば、面倒くさい分散やF値、確率の計算を代行してくれます。Excel2007以上のバージョンでは、「ファイル→オプション→アドイン→分析ツール→OK」で、Excel2003では、「ツール→アドイン→分析ツール」をチェックすることで、分析ツールが利用できる状態になります。一旦アドインを追加したらそれ以降は、「データ→データ分析」(Excel2007以上)、「ツール→分析ツール」(Excel2003)で、**図2**が開けます。

「データ分析」には、一元配置分散分析、繰り返しのある二元配置分散分析、繰り返しのない二元配置分散分析の3種類の分散分析が用意されています。手持ちのデータの構造に合わせて選択します。今回は一元配置分散分析について説明します。

4. 一元配置分散分析

あなたはエアロビクスのインストラクターで、レッスン内容としてA、B、Cの3パターンを持っているとします。A、B、Cのうちどれが最も体脂肪率を下げる効果が大きいか？　自分としては、A、BよりはCのレッスンを体脂肪燃焼を狙った構成にしているつもりだが、本当にそう言って良いのかどうか。10週間の講座前後で体脂肪率を測定してみて、その変化が**表2**のようなデータであったとします。ここでレッスンA、B、Cは、それぞれ異なるクライアントが受けていたとします。また、体脂肪率が減少する程度には体脂肪率の初期値が影響するので、開講前のクライアントの体脂肪率はだいたい同じであったと仮定します。このように対応のない3群以上のデータ、つまり異なる3群以上の被検者間のデータを比較する[3]ときに、一元配置分散分析を利用します。

さて、**表2**をグラフ化したものが**図3**です。確かにレッスンCは、狙いどおり体脂肪率の減少程度が大きいようにもみえますが、そうだと強くも言いにくい結果です。当然ながら、体脂

3　同じクライアントがレッスンA、B、Cをすべて受けて比較する場合(対応がある場合)は、繰り返しのない二元配置分散分析を利用します(次号解説)。

肪率の変化にはレッスン内容だけではなくて、レッスン以外での身体活動や食事、その他諸々の影響を受けるので、こういう何とも判断しにくい結果になることも多いはずです。こういう結果に対し、確率的に考えて判断基準を提供してくれるのが統計的検定です。

分析ツールの「分散分析：一元配置」を開きます(**図4**)。入力範囲には**表2**の例では、$A$1:$C$11 を指定します。「先頭行をラベルとして使用」をチェックすると入力範囲の先頭行(レッスンA～と書かれている行)がデータとしてではなくて、ラベルとして利用されます。出力先を$A$13に指定してOKをすると、**表3**のA13:G27までが出力されます。

出力された「概要」(A15)にはレッスンA～Cそれぞれの標本数(=被検者数)や平均値などが記されています。出力でより重要なのは「分散分析表」(A22)です。分散分析表には、前述のグループ間分散やF値、確率などが記されています。分散分析の結果を正しく理解するためには、この分散分析表の意味を十分に読み取る必要がありますが、現時点では割愛します[4]。セルF24に記されたP値が有意かどうかの判断基準になります。**表3**の例ではP値は0.043…なので、この分散分析の結果、5%水準でレッスン間に有意差があると判断することができます。

図4　分析ツール(分散分析：一元配置)

表3　一元配置分散分析の出力結果

	A	B	C	D	E	F	G
13	分散分析：一元配置						
14							
15	概要						
16	グループ	標本数	合計	平均	分散		
17	レッスンA	8	-0.9	-0.1125	0.56125		
18	レッスンB	10	-4.7	-0.47	0.582333333		
19	レッスンC	9	-10.5	-1.166666667	0.9275		
20							
21							
22	分散分析表						
23	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
24	グループ間	4.95987963	2	2.479939815	3.587670432	0.043331166	3.402826105
25	グループ内	16.58975	24	0.691239583			
26							
27	合計	21.54962963	26				
28							
29					95%信頼区間		
30	最小有意差法			平均の差	下限	上限	
31	t(0.05)		A,B間	0.3575	-0.45644284	1.17144284	
32	2.063898562		B,C間	0.696666667	-0.091754353	1.485087686	
33			A,C間	1.054166667	0.220368644	1.887964689	＊
34							
35					95%信頼区間		
36	シェッフェの方法			平均の差	下限	上限	
37	F(0.05)		A,B間	0.3575	-0.671322455	1.386322455	
38	3.402826105		B,C間	0.696666667	-0.299896247	1.693229581	
39			A,C間	1.054166667	0.000247293	2.108086041	＊

5. 事後検定

ここまで一元配置分散分析を使って、レッスン間に有意差があることを示しました。しかし、分散分析が示すのはレッスン間に差異があることまでで、ではどのレッスン間に差があるかどうかは、別途検討しなければなりません。このように分散分析によって全体としての有意な結果を得たのちに、レッスンAとB、レッスンAとC、レッスンBとCのように、ペアごとにさらに検定を行なうことを事後検定といいます[5]。

4　大まかに言えば、グループ間の分散(D24、レッスンA、B、Cの平均値のバラつき)とグループ内の分散(D25、各レッスン内での個々の被検者のバラつき)の比がセルE24(F値)。このF値をF分布から確率を求めたのがF24のP値。すなわち、E24が大きいほどグループ間分散がグループ内分散よりも大きいことを表す。このことはつまり、グループ間のバラつきが個々のデータのバラつきよりも大きいことを意味し、各グループは同一の母集団から得られたデータとは考えにくい→グループ間に差異がある、と判断する。

5　わざわざ分散分析で全体の有意性を検定しなくても、初めから個々のペアの有意差をt検定で調べればよいのでは？　→有意差検定を個々のペアで複数回行なうと、有意差が出現しやすくなる(危険率5%と言いながら5%以上になってしまう)。したがって、まずは分散分析で、「少なくとも1つのグループは他のグループと違う」ことを5%水準で検定した後、事後に個々のペアを検討しなければならない。

残念ながら、Excelの分析ツールは事後検定まではやってくれません。自分でExcel関数を使いこなしつつ計算しなければいけません。事後検定には様々な方法[6]がありますが、ここでは、事後検定として最小有意差法とシェッフェの方法の2つについて、**表3**にあてはめた計算例を紹介します。

●最小有意差法（表3A30：F33の領域）

グループ間の差が有意になるための最小の値を求めて、実際のグループ間の差がそれ以上かどうか検討する方法です。

まず、セルA32にグループ内変動の自由度(=C25)におけるt分布の両側確率0.05でのt値をTINV関数によって求めておきます。

A32：=TINV(0.05,C25)

セルD31は、レッスンA、B間の平均値の差異で、セルE31、F31は、その95％信頼区間の下限と上限です。

D31：=＋D17－D18
E31：=D31－A32＊SQRT((1/B17＋1/B18)＊D25)
F31：=D31＋A32＊SQRT((1/B17＋1/B18)＊D25)

ここで、セルE31、F31は、セルD31±A32＊√(SQRT関数)の形になっています。√内は、グループ内変動の分散(D25)を被検者数で除していることを表します。レッスンB、C間、レッスンA、C間も同様にして、各平均値の差異の95％信頼区間を求めることができます。すなわち、

D32：=＋D18－D19
E32：=D32－A32＊SQRT((1/B18＋1/B19)＊D25)
F32：=D32＋A32＊SQRT((1/B18＋1/B19)＊D25)
D33：=＋D17－D19
E33：=D33－A32＊SQRT((1/B17＋1/B19)＊D25)
F33：=D33＋A32＊SQRT((1/B17＋1/B19)＊D25)

ここで95％信頼区間とは、各レッスン間の差は今回はたまたまセルD31～D33の値になったけれども、各レッスン内の個々のデータのバラつきを考慮すれば、その値はある幅をもって変動しうる――その変動は、95％の確率で下限(セルE31～E33)から上限(セルF31～F33)の間に位置することを意味しています。逆に考えると、両者の差が下限から上限の間以外になる確率は5％未満しかない、ことを意味します。

今、レッスンA、C間を見てみましょう。計算の結果、レッスンAとCの差の95％信頼区間は0.220～1.887(セルE33、F33)であることがわかりました。つまり、レッスンAとCの差は95％の確率をもって0以上であるわけです。すなわち、

レッスンA－レッスンC>0
⇕
レッスンA>レッスンC

(レッスンAとC間には5％水準で有意差がある)

一方、レッスンAとB、レッスンBとCでは、95％信頼区間内に0が含まれます。このことは、それぞれの差は0であることも起こりうる――両レッスン間には差がない、ということになります。

●シェッフェの方法（表3A36：F39の領域）

シェッフェの方法では有意判定するF値を求め、各レッスン間でのF値がそれと比較して大きいかという点をみます。

まず、セルA38に自由度がグループ内、グループ間それぞれの自由度のF分布の上側確率0.05のF値をFINV関数によって求めます。

A38：=FINV(0.05,C24,C25)

セルD37～D39は各レッスン間の差異で、セルD31～D33に同じです。

セルE37～F39には、各レッスン間の差異の95％信頼区間の下限と上限を求めます。

E37：=D37－SQRT(C24＊A38＊(1/B17＋1/B18)＊D25)
F37：=D37＋SQRT(C24＊A38＊(1/B17＋1/B18)＊D25)
E38：=D38－SQRT(C24＊A38＊(1/B18＋1/B19)＊D25)
F38：=D38＋SQRT(C24＊A38＊(1/B18＋1/B19)＊D25)
E39：=D39－SQRT(C24＊A38＊(1/B17+1/B19)＊D25)
F39：=D39+SQRT(C24＊A38＊(1/B17+1/B19)＊D25)

上式の√内は、グループ間変動の自由度(C24)、F値(A38)、グループ内分散(D25)の積を被検者数で除しています。

計算の結果、レッスン間の差異の95％信頼区間に0が含まれていないのは、やはりレッスンAとC間のみでした。つまり、シェッフェの方法でもレッスンAとC間には5％水準で有意差があることが導かれました。◆

6 事後検定は利用する方法により、ある方法では有意でも他の方法だと有意とならないということがありうる。一般に最小有意差法は有意差が出やすく、シェッフェの方法は有意差が出にくい、といわれます。