心理学統計法　　　松本治彦

１．　はじめに

１－１　事前に知っておくこと

論文の書式を知る

目的；何を問題にするか、何を明らかにしたいのかを述べる。そして、以下に行う実験・調査の仮説や予測を提出する。

方法；目的に沿って、どのような実験・調査を行うかを述べる。その際、次の「小見出し」を用いて、簡潔に要領よく述べる。実験計画、実験期日、被験者、材料、道具、装置、手続き、なお、調査研究の場合、前３者は調査計画、調査期日、対象者となる。また、研究対象が人間でない場合は、被検体、対象とする。

結果；実験・調査によって入手したデータを図表にして掲載する。そして、その分析の結果を示す。

考察；データ分析の結果から何がいえるのか、何が明らかになったかを述べる。それによって、仮説の採択・棄却を判断したり、今後の課題や研究の方向を示唆したりする。

実験・調査の方法を知る

実験とは、ある事象に対して、あらかじめ特定の考え方（理論）を持っている場合に、その理論から導いた仮説が正しいかどうかを確かめようとするもの。実験は仮説検証型の研究手段。

これに対して調査とは、ある事象に対しての特定の考え方を持っていない場合に、その事象の性質や状態を探ろうとするもの。調査は探索・発見型の研究手段。

両者は明確に区別できない点もあるが、普通は研究の進展につれて、調査（探索・発見）から実験（仮説検証）へと移行し、さらに妥当な理論を求めて、調査と実験を繰り返してゆく。

比較可能なデータを入手する。

データは比較可能でなければ処理する意味がない。３つの方法がある。

①　条件設定によって比較の観点を持ちこむ。

実験・調査の場面の中に、２つ以上の条件を設定する。例えば、学習場面において、騒音条件と音楽条件と無音条件を設定して学習成績を比較する。

②　標本抽出によって比較の観点を持ちこむ。

標本抽出とは、被験者または対象者を集めること。このとき、２種類以上の被験者（例えば小学生・中学生・高校生）を集めるなら、データは比較可能となる。

③　比較を限定する。

①の例のように学習成績に及ぼす騒音・音楽・無音の影響を比較したい場合、この条件以外の性質や状態は等質になるようにする（例えば、照明の明るさ、内容の困難度、被験者の学習能力）。すなわち、比較したい条件を１つだけ変化させ、あとは一定に保つ。この一定に保つことを統制するという。統制を行わないと、データには複数の比較の観点が入り込み、分析結果の一義的な解釈が困難になる。

１－２　尺度の判定

データの尺度の判定

データの値がどのような性質の尺度であるかによって、処理の仕方が異なる。

①　間隔尺度と比率尺度（別名；距離尺度と比例尺度）

間隔尺度とは、個々の値の間に等間隔が保証されている尺度である。比例尺度とは、等間隔性に加えて、ゼロを基点とすることができる尺度である。

例えば、温度や時間の値は等間隔である。すなわち、温度の１度と２度の間隔は２度と３度の間隔に等しい。また、時間の１秒と２秒の間隔は２秒と３秒の間隔に等しい。ただし、温度は基点となる値をもたないので間隔尺度どまり（この見解は、誤りで、絶対零度、つまり摂氏－２７３度がゼロと定義できる）。なぜなら、ゼロはあるが、温度０度と言っても温度がなくなるわけではなく、ある一定の高さの温度が存在する

からである。これに対して、時間のゼロは、間隔尺度の上に比率尺度でもある。時間のゼロは時間が存在しないことであり、０秒の２倍は、０秒であるという比率関係が成立する。しかし、間隔尺度と比率尺度の区別はデータ処理上は必要ない。どちらの尺度でも、用いる統計手法は同一。例としては、時間、距離、速度、振動数など特定単位を持つ物理的尺度。正答数、誤答数、試行回数、反応回数など、カウントした値。正答率、誤答率、再生率、発生率などパーセンテージ。知能指数、テスト得点、理解度など得点化された値。

② 順位尺度（別名；順序尺度）

順位尺度とは、個々の値の間に等間隔性が保証されない尺度。この尺度は、値の大小または後先しか問題にできない。例えば、１，２，３位という順位は１位と２位が「鼻の差」でも「ぶっちぎり」でも、１，２という値をとる。このように順位尺度は間隔を問題にしないので、1.2 とか 2.8 とか間隔をきざむ小数値もない。例としては、順位としての値（位の単位をつけることができるもの）。L・M・S、優・良・可、１０段階評定点など等級レベル。

③ 名義尺度

名義尺度とは、間隔の概念も、大小・後先の概念もない尺度。数値ではなく、性質や属性であり、具体的な分類名（カテゴリー）を値とする。例えば、「性別」という名義尺度の値は男女の２値である。このように、値数は少なく、一般に 2～5 値程度である。例としては、参加・不参加、所有・未所有、喫煙・非喫煙など被験者または対象者における特定の反応や特徴のあり・なし。一般に「あり」を１、「なし」を０とコード化するため、このような名義尺度データを「１・０（いち・ぜろ）」データと呼ぶことがある。男子・女子、内向性・外向性、右側・中央・左側など、被験者または対象者における対立する反応や特徴。快・不快、賛成・反対、A 党支持・B 党支持・C 党支持・支持政党なしなど、複数の選択肢のある質問の回答結果。

評定尺度得点の判定についての注意

評定尺度得点は、質問紙法の一種である評定尺度によって得られる。評定尺度法とは、形容詞対（例：好き－嫌い）または程度副詞（例：非常に・かなり・まあまあ）を横一列に並べて以下のような「ものさし（すなわち尺度）」をつくり、その上に○を付けさせて回答を得る方法。集計の際、それぞれ得点化され、数値データとなる。この得点化は、一般に消極的・否定的な回答から積極的・肯定的な回答へと値を一点きざみで増やす。したがって、5 段階なら 0～4 あるいは 1～5 の値を得る。問題は評定尺度得点についての尺度の判定である。順位尺度と判定するのに問題はないが、間隔尺度と判定してよいかどうかとなると難しい。なぜなら、この尺度は間隔の概念に乏しいからである。

２．間隔・比率尺度データの処理

２－１　データ表示

4 ステップからなる。

a.データ・リストの作成
b.ヒストグラムの作成
c.平均と標準偏差の計算
d.作表・作図

b.ヒストグラムの作成

① 極端値のチェック

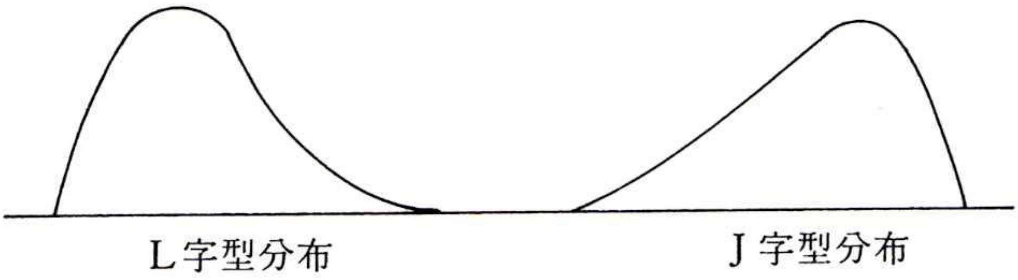
極端値は不良値であることが多いので、その極端値が出た原因を個別に追求し、不良値であることがわかったら、それを捨てて先に進む。

② データの分布形のチェック

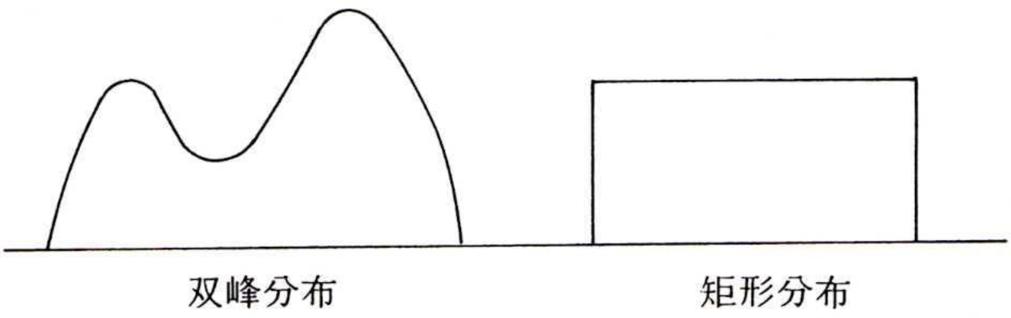
大部分、正規分布であることを前提としている。例えば「平均」の計算がそうである。正規分布していないデータから平均を求めても意味がない。そこでヒストグラムをみて、データが正規分布をしているかどうかチェックする必要がある。

（この部分に、正規分布とみなしてよい例、正規分布とみなせない例を図示する）

◆正規分布とみなしてよい例



◆正規分布とみなせない例



平均値

平均とは、複数個のデータを代表する一つの統計量である（データの分布が正規分布とみなせることが前提）。統計用語では、「代表値」の一種である。

例えば、平均が 50 であるとすれば 50 がこのデータ集団の代表である。実際に 50 という値のデータがなくても、この付近にデータが密集し、50 から±方向に離れるにつれて分布がまばらになってゆく。

平均を用いてはならない場合としては、データの尺度が間隔・比率尺度でない場合、データの中に極端値が存在する場合、データの分布が正規分布とみなせない場合である。このときには、名義尺度データあるいは順位尺度データの処理へ進む。

偏差平方和、標準偏差、標準誤差

数値データのばらつきの程度を調べるのに使われる。

偏差平方和（残差平方和ともいう）

偏差を平方した合計である。偏差とは、平均からの隔たりであり、測定値－平均値である。したがって、
偏差＝測定値－平均値、偏差平方＝偏差×偏差、偏差平方和＝Σ偏差平方となる。

標準偏差 SD（standard deviation ）s，σ

標準偏差とは、データが平均からどの程度ズレているかを表す統計量である。すなわち、標準偏差の値は、データ 1 個分の標準的なズレ幅を示す。統計用語では「散布度」の一種。もし、データの分布が正規分布であれば「平均±標準偏差」の範囲にデータ全体の約 68%がおさまる。例えば、100 個のデータがあり、平均が 50、標準偏差が 10 であるとすると、50±10（つまり、40～60）の範囲にほぼ 68 個のデータがおさまる。

標準偏差の計算式としては、SD＝√｛(偏差平方和) / (N－1)｝＝√｛(データの平方和) / (N－1) －N×（平均値）^{2}/ (N－1)｝＝不偏推定値（unbiased estimate）

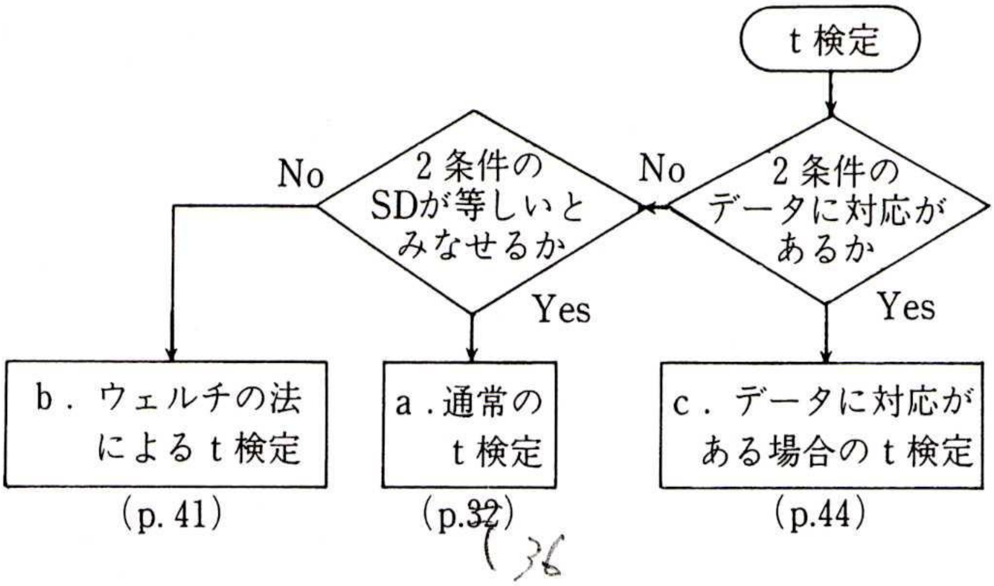
または、SD＝√（偏差平方和/N）＝√｛データの平方和/N－（平均値）^{2}｝

データ処理上は、どちらか一方を一貫して用いるなら支障はない。なお、SD^2は分散（varinace）。
標準誤差　SE
標準誤差は、標準偏差（不偏推定値の方）を√N で割ったもの。
標準誤差（SE）＝標準偏差の不偏推定値/√N＝√｛偏差平方和/N（N－1)｝
数値データのばらつきの程度を表す場合、1 群のみのときは、（平均値±標準偏差)、2 群以上で平均値の比較を行うときには、（平均値±標準誤差）である。
平均と標準偏差の意味
本来、平均とは真の値のことである。すなわち、データにまったく誤差が加わらないときには、すべてのデータが集中する一点の値を表している。したがって、真の値が一点に定まらないときの平均は意味がない。正規分布は真の値が一点に定まることの証である。また、そのときデータの値は、「データ＝真の値±偶然誤差」として定義することができる。この真の値の推定が平均値、偶然誤差の推定はSD であり、ある観察場面における標準的データは、（平均値±SD）という値をとることを意味している。

２－２　データ分析

データ表示が終了したら、次にデータ分析に移る。データ分析には、2 つのコースがある。①有意差を分析する 3 コースと②相関・予測を分析するこニースの二つ（②は省略する)。
①有意差を分析するコース
有意差の分析は、各条件の平均の差または分散の差を問題にする。その差が、偶然の数値のユレにすぎないのか、あるいは偶然ではなく、各条件の性質の違いであるかを判定する。この判定は、統計的検定という手続きを踏む。有意差の分析は、平均の差または分散の差を分析する。一般に、平均の有意差を分析する研究が圧倒的に多い。
平均の有意差を分析する方法としては、 t 検定と分散分析がある。 t 検定は 2 条件の平均の差しか扱えない。分散分析は 2 条件の場合を含め、もっと多くの条件の平均の差を一括して処理できる。

（ t 検定のフローチャート）



見た目の判断で確信がもてない場合は、「分散の有意差検定」を用いる。
2 条件のデータのばらつきの程度を比較するには、「偏差平方和/自由度」＝分散（不偏分散）を利用する。
2 条件の分散の比を計算するが、この時、必ず 2 条件の分数のうち、大きい方の分散を分母にもってくる。この 2 条件の分散の比を危険率 5%の F 表の F 値と比較する。そして、この値よりも 2 条件の分散の比の方が大き

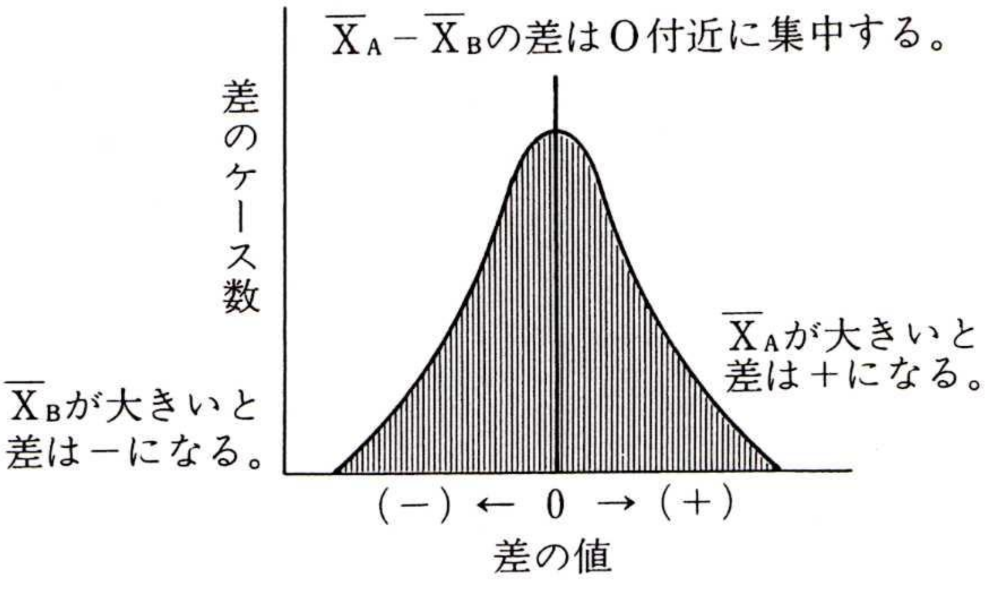
ければ、2 条件の分散に差があることになり、ウェルチの方法による検定を行う。逆に、この値よりも 2 条件の分散の比の方が小さければ、対応のない場合の通常の t 検定を行う。

＊データ分析における検定方法全般にあてはまること

実験・調査の研究仮説は、普通、条件間に平均の差が生じることを予測する。しかし、もしもデータ上に表れた平均の差が、そんなに大きくないなら、数値の上では差があっても、単に偶然によって生じたユレである可能性が強い。その場合、条件間に本当の差があると言うことはできない。そこで、目の前の差が、偶然によって生じたものか否かを判定する必要がある。もし偶然に生じる程度の小さな差なら、その差に意味がない。しかし、偶然に生じるよりも大きい差なら、その差は意味がある（すなわち有意差　significant difference)。この偶然か偶然でないか（意味があるか意味がないか）の判定を、有意差検定または有意性検定という。

有意差検定は背理法である。すなわち、最初に「目の前の差は偶然の差である」と仮定しておいて、偶然とは言えない反証を示し、最初の仮定をくつがえす論法である。この最初の偶然の仮定を専門用語で「帰無仮説（null hypothesis)」と呼ぶ。検定の結果、帰無仮説がくつがえされて、文字通り無に帰したときに「目の前の差は偶然に生じたのではない（すなわち有意である)」と認める。 t 検定の手順も、以上の背理法に従っている。まず、最初に「2 条件の平均の差は偶然に生じた差である」と仮定する。すなわち、まったく偶然が加わらなければ、条件 A と条件 B の平均の差はゼロになるはずである｛(A グループの平均) － (B グループの平均) =0｝。このように、帰無仮説は偶然の仮説である。ここで、この帰無仮説を正しいものとしておいて、理論的に偶然を作用させてみる。そして、偶然が加わったとき、平均の差の値がどう出るかを多数回想定してみる。もちろん、帰無仮説｛(A グループの平均) － (B グループの平均) =0｝が正しいと仮定しているのであるから、差の値は 0 になるケースが最も多く、0 の近辺に集中するケースが大部分であろう。たまたま大きな差が出現しても、偶然の結果であるから、ケース数はそう多くならない。このような偶然下で出現しうる差のケースを多数回想定していくと、理論的には、以下のような｛(A グループの平均) － (B グループの平均)｝の偶然分布を描くことができる。

(｛(A グループの平均) － (B グループの平均)｝の偶然分布の図)

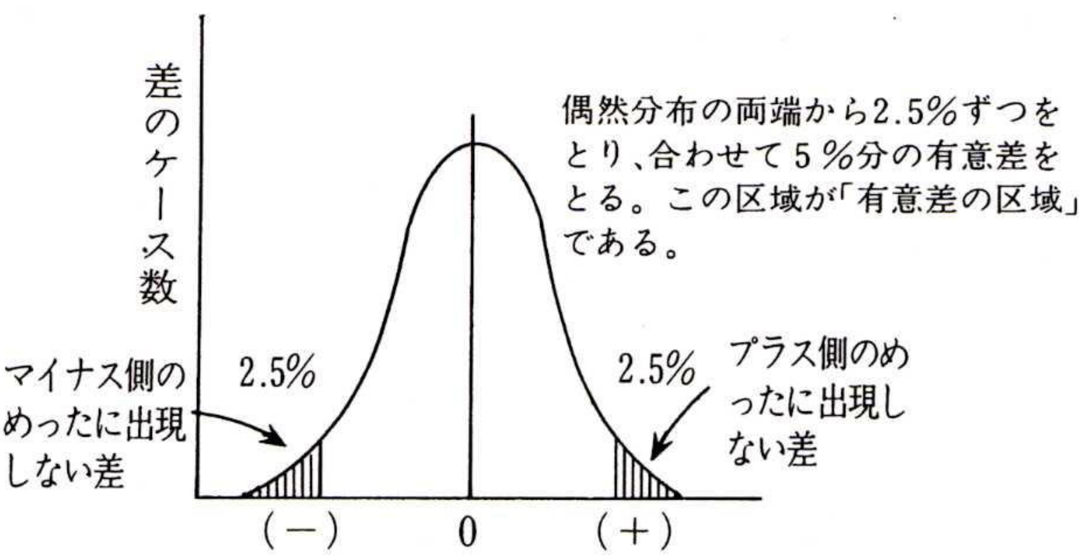


ここで、有意差検定は、実際のデータに表れた平均の差を上の偶然分布に当てはめる。そして、実際の差を同程度の差が偶然下ではどれだけ出現するのかをみようとする。このとき、もし、実際の平均の差が偶然分布の中心（0）からあまり遠くないところに当てはまるなら、同程度の差は偶然下でも相当数出現しているので、実際の平均の差も、それらと同様、偶然によるユレにすぎないと考えられる。したがって、帰無仮説を取り下げる理由がなく、本当に差があるとは認められない。これに対して、もし、実際のデータに表れた平均の差が 0 から大

きく離れて、偶然分布の端のほうへ当てはまるなら、この差は偶然に出現したとは考えにくい。事実、分布の端にくるような極端に大きな差は、偶然下ではそう多く出現しないはずである。そのような、偶然下では「めったに出現しない」大きさの差が、現実のデータに出現しているなら、この差は偶然のユレではない（有意差である）と認めてよいであろう。

このとき、統計学では「めったに出現しない」基準として、あらかじめ偶然分布の両端に全体の 5%分の「有意差の区域」を決めておく。そして、データの平均の差が、その区域に当てはまるなら、偶然には出現しない差（すなわち有意差）と認めることにしている。データの平均の差が、この 5%の「有意差の区域」に飛び込んだ場合、帰無仮説｛(A グループの平均) － (B グループの平均) ＝0｝を棄却する（このため有意差の区域は専門用語では棄却域と言われてる)。これによって、帰無仮説に代るべき対立仮説｛(A グループの平均) － (B グループの平均) ≠0｝を採択する。すなわち、データに表れた平均の差は実質的な差であると判定する。かくして、この差を予測した研究仮説を支持するのである。以上が有意差検定の考え方である。

（有意差の区域「棄却域」の分布）



２－３　ｔ検定

有意差検定を行うためには、あらかじめ平均の差の偶然分布を用意する必要がある。これは統計学では「ｔ分布」として用意されている。ここでは、ｔとは平均の差を表す統計量である。平均の差そのものは、データの数値の寸法によって大きくなったり、小さくなったりする。そこで、データの平均の差を汎用性のある統計量に換算する。

$$t = | X_A - X_B | / \sqrt{(SD_A^2 + SD_B^2) / (N-1)}$$

上式の分子は条件の平均の差を計算している。分母は、この平均の差に加わる偶然誤差の大きさを推定する式である。すなわち、ｔ値は平均の差が偶然誤差の何倍であるかを表す統計量となっている。

<u>データに対応がない場合のｔ検定</u>

表８　テスト得点の平均と標準偏差

	条件Ａ	条件Ｂ
N	6	6
X	48.0	70.0
SD	14.5	12.9

①　t 値の計算

$t = |X_A - X_B| / \sqrt{(SD_A^2 + SD_B^2)/(N-1)}$

$= |48.0 - 70.0| / \sqrt{(14.5^2 + 12.9^2)/(6-1)}$

$= 2.53$

t=2.53 とは、平均の差の大きさが偶然誤差の大きさの約 2 倍半の大きさであることを示している。

②　自由度の計算

degree of freedom　→　df

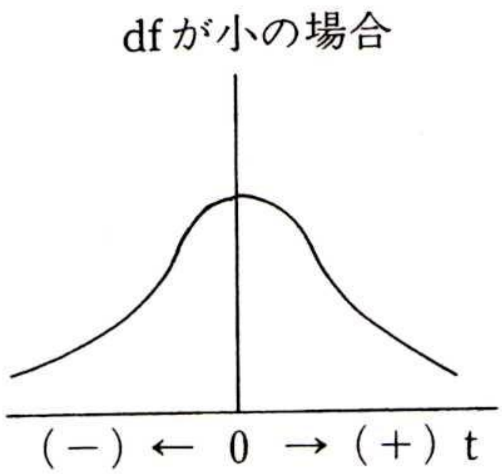
$df = (N_A - 1) + (N_B - 1) = (6-1) + (6-1) = 10$

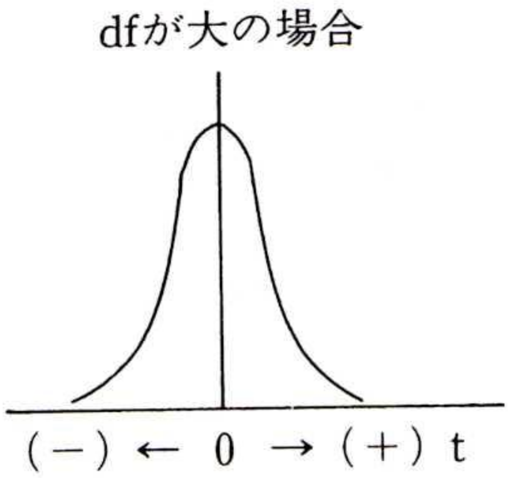
自由度は各条件のデータの個数から 1 を引いた合計。

自由度が小　→　データの個数が少ない。この場合、t 値の偶然分布は、偶然の影響が強く出て、0 から離れた値が出現しやすい。t 分布はスソが長くのびた形になる。

自由度が大　→　データの個数が多い。この場合、追加するデータによって偶然の影響が相殺され、極端に大きい値は出にくい。t 分布は 0 に寄って密集した形になる。

（df が小の場合）　　　　　　　　　　　　（df が大の場合）





＊自由度は t 分布の形を変える

③t 値の出現確率（危険率）p

①で求めた t=2.53 と②で求めた df=10 から、t 分布表をみて p を求める。

表 9　t 分布表

df	t 値		
9	1.833	2.262	3.250
10	1.813	2.228	3.169
11	1.796	2.201	3.106
出現確率	0.10	0.05	0.01
有意水準	有意傾向	5%	1%

t=2.53 の出現確率は、$0.05 > p > 0.01$（5%と 1%の間）である。

④　有意性の判定

データから計算で求めた t の値 2.53 は、理論上の偶然分布の中では 5%未満の小さな出現確率しかもたない。この出現確率は公認の有意性の基準をクリアしている。したがって、t=2.53 は「有意である」（偶然には出現しない）と認める。

t 値は、平均の差の統計量である。t 値が有意とは、平均の差が有意である（偶然の数値のユレではない）と言うことを示している。その差は、2 条件を設定した効果と推論する。

⑤　検定結果を論文に記載

「表 8 は条件 A と条件 B の平均と標準偏差を示したものである。t 検定の結果、両条件の平均の差は有意であ

った（両側検定 : t（10）＝2.53、p ＜0.05)。したがって、条件Ａより条件Ｂのほうが優れた成績をあげると言える。」

第１文は「データ表示」、第２文が「データ分析」、第３文は分析結果に基づく「推論」である。論文に必ず記載するものは、両側検定、t、df、ｐの４つである。

統計的検定の判定

表 10　出現確率と有意水準と論文記載上の表現

出現確率	有意水準	論文記載上の表現
ｐ＞0.10	―	「有意でない」
0.05＜ｐ＜0.10	―	「有意傾向である」
ｐ＜0.05	5％	「有意である」
ｐ＜0.01	1％	「有意である」

出現確率＝危険率？

例えば、t 値の出現確率 5％未満で帰無仮説を棄却、逆に言えば、5％未満であるが、帰無仮説下で偶然に出現する確率が残っている。したがって、この 5％未満という確率は、帰無仮説の棄却が誤りである危険性を示している。

ウェルチの法によるｔ検定

データに対応がない場合、かつ、２条件の SD の値が等しいと見なせない場合は、通常のｔ検定を用いることはできない。なぜなら、検定に利用するｔ検定はもともと条件Ａと条件Ｂの真の平均と標準偏差が等しいという仮定のもとに標準偏差を一定として、平均にのみ偶然が作用したときの分布になっているからである。したがって、データの SD が等しいと見なせない場合、通常の計算式によるｔ値はこのｔ分布にしたがわなくなる。そのような場合には、コクラン・コックス（Cochran Cox）の法やウェルチ（Welch）の法によるｔ検定を行わなければならない。ここでは、ウェルチの法を説明する。

①　下の表 11 に示したデータに対してｔ検定を行う。

表 11　テスト得点の平均と標準偏差

	条件Ａ	条件Ｂ
N	26	12
X	42.1	30.5
SD	13.5	19.5

SD の値をみると 6 ポイントの差であるが、これを等しいと見なすことができるかどうかは見た目だけでは判定できない。このような場合は、分散の有意差検定を用いる。それによってここでは仮の話で、この 6 ポイントの差は偶然の数値のユレではなく、有意差とする。すなわち、両条件の SD を等しいと見なすことはできない。そこで、ウェルチの法によるｔ検定を行う（ずっと後で分散の有意差を求める例を示す)。

②　下式によってｔ値を計算する（添字Ａ・Ｂは条件を表わす）

$t = | X_A - X_B | / \sqrt{(SD_A^2 / (N_A - 1) + SD_B^2 / (N_B - 1))}$

$= | 42.1 - 30.5 | / \sqrt{(13.5^2 / (26 - 1) + 19.5^2 / (12 - 1))}$

$= 1.79$

③　下式によって、いったん定数Ｃを求めて df（自由度）を計算する。

$c = SD_A^2 / (N_A - 1) / (SD_A^2 / (N_A - 1) + SD_B^2 / (N_B - 1))$

$= 13.5^2 / (26 - 1) / (13.5^2 / (26 - 1) + 19.5^2 / (12 - 1))$

$= 0.174$

$df = (N_A - 1) \cdot (N_B - 1) / ((N_A - 1) \cdot (1 - c)^2 + (N_B - 1) \cdot c^2)$

$= (26 - 1) \times (12 - 1) / ((26 - 1) \times (1 - 0.174)^2 + (12 - 1) \times 0.174^2)$

$= 15.8$

df＝15.8 となったが、小数点以下は切り捨てて、df＝15 とする。かりに、もし通常のｔ検定を適用することがで

きるなら、df は 36 となっていた。自由度が小さくなることによって、検定は厳しくなる。なお、各条件のデータの個数が等しい場合（N_A＝N_B)、上式は以下のように簡単になる。

$c＝SD_A^2/（SD_A^2＋SD_B^2）$、$df＝（N－1）/（c^2＋（1－c）^2）$

④ 計算した t 値の出現確率を t 分布表において調べる。②で計算した t＝1.79 を③で計算した df＝15 の t 分布に対照して、その出現確率を求める。

表 12 t 分布表

df	t 値		
14	1.761	2.145	2.977
15	1.753	2.132	2.947
16	1.746	2.120	2.921
出現確率	0.10	0.05	0.01
有意水準	有意傾向	5%	1%

上表において、t＝1.79 の偶然による出現確率は、10%未満、5%より大きいとなり、有意水準をクリアできないが、それをクリアする傾向にある。この有意性の一歩手前の段階を有意傾向という。

⑤ 検定の結果を論文に記載する。

「表 11 は、条件 A と条件 B のテスト得点の平均と標準偏差を示したものである。分散の大きさが等質と見なせなかったので、ウェルチの法による t 検定を行った。その結果、両条件の平均の差は、有意傾向であった（両側検定：t（15）＝1.79、0.05＜ p ＜0.10)。したがって、条件 A は条件 B よりも高得点をあげる傾向にある。」

上文中、「分散の大きさが、等質とみなせなかった」ことがウェルチの法を用いた釈明となっている。実際にウェルチの法を用いたかどうかは、自由度の値が通常よりも小さいことによってわかる。

データに対応がある場合の t 検定

2 条件のデータに対応がある場合は、その対応するデータを一組ずつ比べる。

① 対応するデータどうしを一組ずつ比較し、差をとる。

表 13 問題解決時間のデータ・リスト（単位：秒）

被験者	条件 A	条件 B	その差（A－B)
イ	112	125	－13
ロ	95	105	－10
ハ	103	98	5
ニ	90	97	－7
ホ	124	125	－1
ヘ	100	105	－5
ト	108	113	－5
X	104.6	109.7	－5.14
SD	10.5	10.8	5.46

上表で条件 B の問題解決のほうが長い時間を要したことがわかる。ここで、対応のある t 検定の場合、これらの差が－5.14 を中心に正規分布することを前提とする（そうみなせばよい)。

② 差の統計量から t 値を計算

差の個数（N＝7)、差の平均（X＝－5.14)、差の標準偏差（SD＝5.46）下式に代入し、t 値を求める。

$t = |X| / \sqrt{SD^2/(N-1)}$

$t = |-5.14| / \sqrt{5.46^2/(7-1)}$

$t = 2.31$

③ 自由度の計算

df=N－1=7－1=6

④　t 値の出現確率

②で計算した t＝2.31 と③で求めた df＝6 から t 分布表をみて p を求める。

表 14　t 分布表

df	t 値		
6	1.943	2.447	3.707
出現確率	0.10	0.05	0.01
有意水準	有意傾向	5%	1%

t＝2.31 を df＝6 の数列より　　0.10＞p＞0.05 となる。

⑤　有意性の判定

出現確率 10%未満 5%より大きいということは、有意傾向となる。

⑥　検定の結果を論文に記載

「表 15 は、条件 A と条件 B における問題解決時間の平均と標準偏差を示したものである。

表 15　各条件の問題解決時間の平均と標準偏差（秒）（N=7）

	条件 A	条件 B
X	104.6	109.7
SD	10.5	10.8

t 検定の結果、両条件の平均の差は有意傾向であった（両側検定：t（6）＝2.31、0.05＜ p ＜0. 10.)。

したがって、条件 A のほうが条件 B よりも問題解決に要する時間が少ない傾向にあると言える。」

この文の前半が「データ表示」、表 15 から「データ分析」、最後が分析結果に基づく「推論」になる。

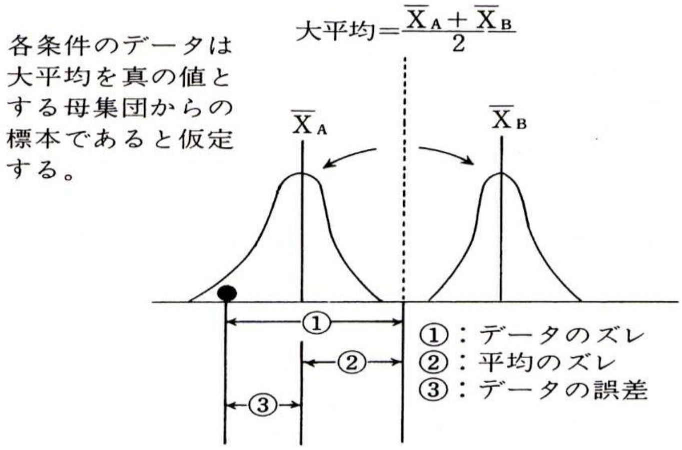
3　分散分析（Analysis of Variance）

<u>概要：</u>

英語の原語を略して通称 ANOVA（アノヴァ）と呼ばれる。<u>分散分析も t 検定と同様、平均の有意差検定の方法</u>である。ただし、t 検定は 2 個の平均しか比べることができなかったが、分散分析は<u>3 個以上の平均</u>を比べることができる。

<u>分散分析の考え方：</u>

はじめに 2 条件 A・B の平均の有意差検定から説明する。分散分析では、条件 A と条件 B のデータを同一の母集団からの標本とみなし（帰無仮説）、この母集団の真の値の推定値として条件 A と条件 B の「平均の平均」を求める。すなわち、(X_A＋X_B）/2。これを大平均と呼ぶ。分散分析はこの大平均からのズレに注目する。



上図で、条件Ａに所属するデータ１個（●）の、大平均からのズレ①に注目する。ズレ①は２つの部分に分解されることがわかる。すなわち、「ズレ①＝ズレ②＋ズレ③」である。ここで、ズレ②は、このデータが所属する条件Ａの中心（X_A）がそこまで移動したので、それにつれてデータが一緒に移動した分である。すなわち、ズレ②は、条件の平均の差（X_A と X_B との差）で生じている。一方、ズレ③は、何の影響で生じたかよくわからないので、このデータに加わった偶然の影響とみるしかない。したがって、「ズレ①＝ズレ②＋ズレ③」という等式は、下のように書き換えることができる。

データの大平均からのズレ＝平均の差の影響によって生じたズレ＋偶然の影響によって生じたズレ

分散分析は、このようなズレの分解を全データについて行う。そして、データのズレ（すなわちデータの値）を支配したのは、平均の差の影響力なのか、それとも偶然の影響力なのかを比較する。もし、前者の影響力が大きければ、平均の差は文字どおり偶然以上であることになる（有意である）。

このとき、ズレを表す統計量として分散（ズレの値を二乗したもの）を用いる。なぜなら、ズレの値は＋・－の方向性を示すからである（例えば、大平均の左にあるデータは－のズレを生じる）。そのまま全データについてズレを合計すると値が相殺され、最終的に「ズレの大きさ＝０」となってしまう。値の＋・－を消すために二乗値（分散）を用いるのは、統計計算の常套手段である。このように、ズレを分散に換算して分解しようとするので、「分散の分解」すなわち分散分析と呼ばれている。

計算例：

簡単な例を用いて、実際に分散分析を行ってみる。下の表 19 は、2 条件Ａ・Ｂのデータ・リストである。分散分析を用いた平均の有意差検定を行う。

表 19　実験結果のデータ・リスト

被験者	条件Ａ	被験者	条件Ｂ
イ	2	ヘ	5
ロ	3	ト	5
ハ	4	チ	7
ニ	5	リ	9
ホ	6	ヌ	9

①データの表示を行う。

各条件のデータ数（N）、平均（X）、標準偏差（SD）を計算し、作表によって掲載する。

表 20　各条件の平均と標準偏差

	条件A	条件B
N	5	5
X	4.0	7.0
SD	1.4	1.8

②　大平均（㊥）を計算する。

大平均＝（4.0+7.0）/2＝5.5　ただし、各条件のＮが等しくない場合は、加重平均による。

③　各データと大平均のズレを計算して、合計する。

ズレは、差として求めるが、値の＋・－を消すため二乗する。この合計をデータの全分散と呼ぶことにする。

データの全分散＝｛(データ①－大平均㊥）^{2}＋（データ②－大平均㊥）^{2}＋（データ③－大平均㊥）^{2}＋・・・・・

＝　$(2-5.5)^2+(3-5.5)^2+(4-5.5)^2+\cdots\cdots\cdot$

＝　48.5

被験者	条件Ａ	被験者	条件Ｂ
イ	2	ヘ	5
ロ	3	ト	5
ハ	4	チ	7
ニ	5	リ	9
ホ	6	ヌ	9

④　平均の差によって生じたズレを計算する。

各条件の平均と大平均のズレを計算して、データの個数分だけ倍にすればよい。

平均の差による分散＝（4.0－5.5）^{2}×5＋（7.0－5.5）^{2}×5＝22.5

⑤　偶然によるズレを計算する。

条件Aに所属するデータはこの所属する平均からのズレ、また、条件Bに所属するデータはこの所属する平均からのズレを計算して足す。

偶然による分散＝（データ①－X_A）^{2}＋（データ②－X_A）^{2}＋・・・・・
＋（データ⑥－X_B）^{2}＋（データ⑦－X_B）^{2}＋・・・
＝　（2－4.0）^{2}＋（3－4.0）^{2}＋・・・
＋（5－7.0）^{2}＋（5－7.0）^{2}＋・・・
＝26.0

以上で、データの全分散は2つに分解された。

$\boxed{\text{データの全分散（48.5）}}$＝$\boxed{\text{平均の差による分散（22.5）}}$＋$\boxed{\text{偶然による分散（26.0）}}$

⑥　分散分析表を作成する。

分散分析表とは、計算結果を掲載し、かつ、有意差検定に必要なF比という統計量を求めるための表である。

表21　分散分析表

要因（SV）	平方和（SS）	自由度（df）	平均平方（MS）	F
条　件	22.50	1	22.50	6.92
誤　差	26.00	8	3.25	
全　体	48.50	9		

各項目の見出しは、カッコ内の略号を用いてもよい。

要因（source of variance）とは、データの分散を生じさせた原因のことである。名称は、一般に表中のように簡略化する。読み方は、表をタテにみて、「条件」の平均の差と偶然の「誤差」がデータ「全体」の分散を生じさせたとする（条件＋誤差＝全体)。すなわち、前述したズレの分散式と同じである。

平方和（sum of square）とは、平方（二乗）の総和という意味である。すなわち、ズレの分散値の合計のことである。したがって、この欄に計算結果の数値を書き込む。

自由度（degree of freedom）とは、一般に、統計量の出方を左右する指数であり、要因ごとに固有の値をとる。ここでは平方和の大きさに直接影響するので無視できない。

条件の自由度＝条件－1　　(2－1＝1)

誤差の自由度＝（N_A－1）＋（N_B－1）　　(4＋4＝8)

全体の自由度＝データの総個数－1　　(10－1＝9)

なお、　条件の自由度＋誤差の自由度＝全体の自由度　となる。

平均平方（mean square）とは、自由度1個分に平均した平方和である。自由度の大きい要因は平方和も大きくなるので、要因どうしの平方和と比べるときに、そのままでは公平な比較にならない。そこで、自由度1個分の値に換算する。平均平方の計算は、表中のヨコの並びを利用して、割り算を行う。なお、「全体」については求めない。

要因	平方和÷自由度＝平均平方
条件	22.50÷　1　＝22.50
誤差	26.00÷　8　＝3.25

最後に、F比（F-ratio）を求める。これも表中のタテの並びを利用して、分散にみなすとよい。

F比は条件の平均平方（分子）が、誤差の平均平方（分母）の何倍あるかを示す統計量になっている。すなわち、条件の平均の差によるズレが、偶然の誤差によるズレより大きいかどうかを端的に比較している。この例では、平均の差によるズレは偶然の誤差によるズレより約7倍も大きい。このように、F比の値がかなり大きいなら、平均の差は偶然に生じるより大きい（すなわち偶然には生じない）と認めてもよいであろう。しかし、もしF＝1であるなら、平均の差は偶然誤差と同程度であり、単なる偶然の数値のユレにすぎないことになる。

⑦　F比の有意性を検定する。

分散分析の最終ステップは、F比の有意性の検定である。あらかじめ、F比がどの程度の大きさであれば有意と認めるかを決めておく必要があるが、この基準はt検定と同一である。すなわち、偶然によって生じる確率が5%未満しかないような大きさであればよい。ここで、統計学では、偶然下で生じる無限個のF比の理論的分布

を用意している（F 分布という）。それをみれば、偶然下では、どの程度の値がどの程度の確率で出現するのかがわかるので、計算した F 比の出現確率を知ることができる。ただ、F 分布は、F 比の分子と分母の自由度によって分布形を変えるので、該当する F 分布を探すためには、2 つの自由度をもっていなければならない

この場合、F 分布表をみると、df①＝1、df②＝8 で F 比が有意水準 5%で 5.35 であり、分散分析表より求めた F＝22.5÷3.25＝6.92 の（偶然による）出現確率は 5%未満しかなく、有意性の基準をクリアしている。このとき、分散分析表の F 比の値に特別な記号を付けるのが慣習である。

表 24　分散分析表（再掲）

要因（SV）	平方和（SS）	自由度（df）	平均平方（MS）	F
条　件	22.50	1	22.50	6.92
誤　差	26.00	8	3.25	
全　体	48.50	9		* p ＜.05

表中の“＊”は「アスタリスク」という。有意性 5%水準で 1 つ、1%水準で 2 つ付ける。また、欄外には必ず「* p ＜.05」または「** p ＜.01」という注釈を添える。

表 25　出現確率による有意性の判定とその表現

F 比の出現確率	有意水準	分散分析表における表記	論文での表現
p ＞.10	―	n.s.	「有意でない」
.05＜ p ＜. 10	有意傾向	†	「有意傾向がある」
p ＜.05	5%	＊	「有意である」
p ＜.01	1%	＊＊	「有意である」

上記の“n.s.”は“non-significant”（有意でない）の略である。また、“†”は「ダッカー」といい、有意傾向のときに付ける。有意傾向とは、有意性の一段手前の段階であり、有意になる可能性があることを表現している。

<u>分散分析の結果の論文記載例：</u>

分散分析は、F 比の検定を持って終了する。その結果は、次のように論文に記載する（F 比の 2 つの自由度の書き方に注意すること）。

「表 20 は、各条件の平均と標準偏差を示したものである。分散分析を行った結果、表 24 にみられるように条件の要因は有意であった（F（1，8）＝6.92、p ＜.05）。したがって、条件 B は条件 A を上回る成績をあげると言える。」

第 1 文がデータ表示、第 2 文がデータ分析の記述である。第 3 文は、分析の結果に基づいた推論となる。このように、「条件の要因が有意であった」という言い方をする。その上で、条件の諸要因のうち平均の差が有意であったと推論する。なぜなら、F 比が有意になる理由としては条件の「平均の差」以外の可能性も考えられなければならないからである。したがって、別の可能性が無視できないほど大きい場合は、分散分析を純粋に平均の有意差検定として用いることはできない。なお、分散分析による平均の有意差検定は両側検定にしかならないので、「両側検定」・「片側検定」の区別を付記する必要はない。

<u>分散分析を使用できない場合：</u>

①各条件のデータの分布が正規分布とみなせない場合

②各条件の分散（SD^2）の大きさが等しいとみなせない場合

以上の場合、F 比の大きさが影響を受けるので、分散分析を平均の有意差検定として用いることはできない。

<u>分布の非正規性と分散の非等質性への対処：</u>

以下の 3 つの対処の仕方がある。

①　なにもしない。

実際には、何も対処を行わない例も少なくない。というのは、分散分析法は、分布の非正規性と分散の非等質性に対して「頑健である」（あまり異なる結果をださない）ことが知られている。

②　有意水準を上げる。

各条件の分布が正規形でなく、分散の大きさが等しくない場合は、（平均の有意差によらない）不当に多くの有意な F 比が得られる傾向がある。したがって、そのような恐れがある場合は、有意な F 比を得る確率を低くしておけばよい。そこで、有意水準を通常の 5%から 1%に上げる。こうすると、F 比の出現率 p ＜.01 が「有意」、.01＜ p ＜.05 が「有意傾向」、p ＞.05 が「有意でない」となる。

③　データの数値を変換する。

一般には、この方法が、もっとも多く用いられている。なぜなら、数値の変換は、正規形から偏った分布の山を中央に移動させ、同時に、ある程度分散の大きさをおさえるからである。

<u>分布の非正規性・分散の非等質性に対処できない場合：</u>

次の場合は、どのような対処も妥当性がない。

①　各条件のデータの分布が2つ以上の山をもつ場合

典型的には、1つの条件のなかに、2つ以上の正規分布が見出される場合である。この場合、そもそも平均を求めること自体が無意味である。たとえ条件間に平均の有意差が認められても、平均自体がなにを代表しているか解釈できない。そのような場合は、引き続き第二実験あるいは第二調査を企画すべきである。すなわち、この2山を分離するような新たな条件を導入して、1条件1山になるように条件設定を組みなおしたほうがよい。

②　理論的・経験的根拠がまったくない場合

例えば、「各条件のデータは正規分布する性質をもっている」というような理論的根拠や、「先行研究では各条件の分散はおおむね等しく出ていた」というような経験的確証が、ぜんぜんない場合は、どんな対処も行わないほうがよい。それを無理に正規分布・当分散とみなすことは、事実を捻じ曲げることになる。そのような場合は、分散分析による平均の有意差検定に固執しないで、むしろ積極的に各条件の分散（SD^2）の非等質性を認め、その分散の差に条件の違いが表れているのではないかとみるほうが自然であり、かつ、生産的である。

<u>備考①：F比とt値</u>

2条件の場合については、分散分析の結果とt検定の結果は完全に一致する。すなわち、分散分析のF比と、t値との間には以下のような等式が成立する。

$$F_{(1,n)} = t_{(n)}{}^2$$

例えば、前述の分散分析の例ではF（1,8）＝6.92となったが、もし、t検定を行うと、t（8）＝2.63となる（2.63^2＝6.92）。このときのt値は、df＝8のt分布においてやはり5%未満の出現確率しかもたず、分散分析の結果と一致する。

<u>備考②：分散分析と実験計画</u>

本例では、分散分析の基礎概念を解説するため、データに対応がない2条件の比較を取り上げたが、もちろん、分散分析では、データに対応がある場合にも適用することができる。また、3個以上の平均の比較や「2条件×2条件」というような交差した比較も分析することができる。こうした、色々な条件の組み合わせ方によって、分散分析の計算の仕方も少しずつ異なってくる。そこで、統計学では、特定の条件設定のパターンと、それに見合う分散分析の手続きをセットとしてまとめている。そして、これらのセットを総称して「実験計画法」と呼んでいる。

分散分析の事例１；数値の変換を行った例

<u>事例</u>

論文名「バスケット・ボールにおける人気生徒と孤立生徒とのシュート率の比較」

方法　被験者：中学1年生、男子20人（人気生徒10人、孤立生徒10人）。5クラスの生徒200人（男女）にソシオメトリー・テスト（この意味は後に説明している）を施行し、その結果に基づいて各クラスの男子から人気生徒2人、孤立生徒2人ずつ抽出した。

手続き：男子のバスケット・ボールのゲームにおいて、各被験者が敵陣でボールを所持した回数およびそのときのパスとシュートの回数を調べた。

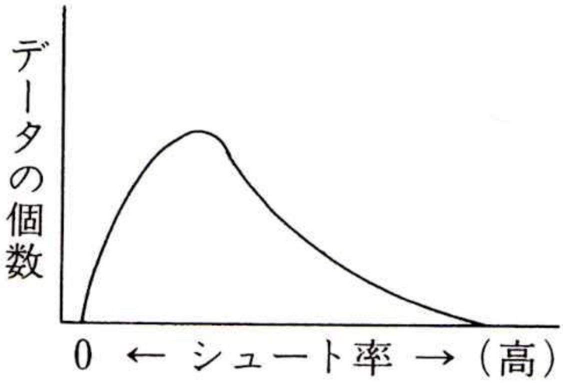
結果　ボール所持回数全体から、敵に奪取された回数とルールによって放棄した回数を引いて、シュート率を算出した（シュート率＝シュートの回数/ボール所持回数）。表26は、人気、孤立生徒別にシュート率の平均と標準偏差を示したものである。

表 26　人気生徒と孤立生徒のシュート率

	人気生徒	孤立生徒
N	10	10
X	.352	.166
SD	.173	.154

数値に角変換（後に説明がある）を施して、分散分析を行った結果、生徒の要因は有意であった（F (1,18) ＝6.78、p ＜.05）。したがって、ボールを所持した場合、人気生徒はシュートする率が高く、孤立生徒は相対的にパスをする傾向が強いといえる。

解説　「手続き」における「ソシオメトリック・テスト」とは、学級集団における各生徒の対人的地位の高さを調べるテストである。この得点に基づいて、各クラスの男子生徒の上位 2 名を「人気生徒」、下位 2 名を「孤立生徒」として抽出している。シュート率に対して行った「角変換」は、%のような数値に対してよく用いられる方法である。その目的は、データの分布を正規分布に整形することである。なぜなら、表 26 をみるとわかるように、どちらの平均の値も低めに出ており、「平均－標準偏差」の値は、かなり 0 に近いからである。すなわち、データの分布は、0 方向に詰った「L 字型分布」であると予想される。



したがって、数値変換の処理は賢明であった。分散分析は、変換後の数値で行う。もし変換前の分散分析が有意であって、変換後の分散分析では有意でないとしても、前者を採用してはならない。平均の差ではなく、分布の非正規性が F 比を有意に押し上げたのかもしれないからである。なお、変換後の数値によって X、SD を計算してデータ表示に用いてもよいが、もとのデータの X、SD を優先して示すこと。変換方法さえ明示してあれば、だいたい変換後どうなるかわかるので。

数値の変換（角変換と対数変換）

適用

数値の変換は、L 字型・J 字型のデータ分布に対して、分布の正規化と分散の圧縮のために用いる。変換方法としては、一般に、角変換と対数変換がよく用いられる。なお、もともと正規分布しているデータを変換すると、かえって分布形をゆがめることになるので、適用には注意が必要である。

角変換

角変換は、「閉じた尺度」のデータに対して適用する。閉じた尺度とは、**値に上限と下限を持つような尺度**である。例えば、正答率は 0%－100%の範囲でしか値をとることができない。また、テスト得点は（50 点満点とすれば）0 点～50 点の範囲内でしか値をとることができない。すなわち、正答率もテスト得点も値に上限と下限があり、尺度は閉じている。閉じた尺度では、データの平均（分布の中心）が上限または下限に近づくにつれ、分布の「すそ野」が詰ってくる傾向があり、L 字型分布や J 字型分布ができやすい。そこで、個々のデータの値を、下式によって角度に変換する。

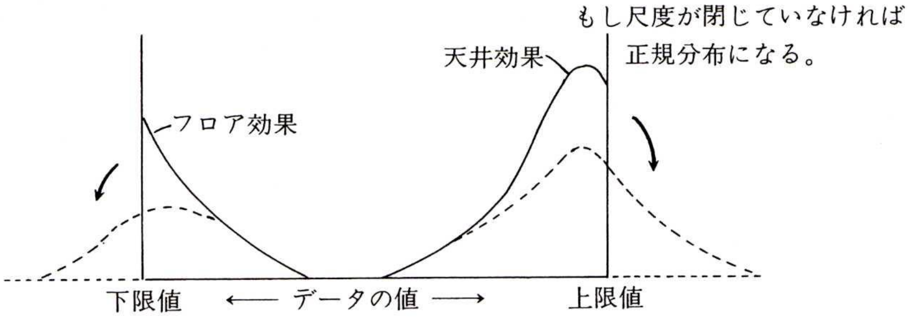
角変換値＝$\sin^{-1}\sqrt{\ }$（データの値／上限値）

上式の $\sin^{-1}$ は三角関数の 1 つ「逆正弦関数」である。また√内は 0～1 までの比率をとる（正答率 35%なら

0.35、50 点満点のテストで 20 点なら 0.40)。ただし、√内が 0 または 1 になる場合は便宜的比率を用いる（以下参照、N は上限値)。

$$\sin^{-1}\sqrt{(0/N)} \rightarrow \sin^{-1}\sqrt{(1/4N)},\quad \sin^{-1}\sqrt{(N/N)} \rightarrow \sin^{-1}\sqrt{((N-0.25)/N)}$$

例えば、解答数 10 で正答数 0 の被験者の便宜的比率は（√内）0.025、その角変換値は 9.10 となる。解答数 20 で正答数 20 の被験者の便宜的比率は 0.988、その角変換値は 83.71 となる。



対数変換

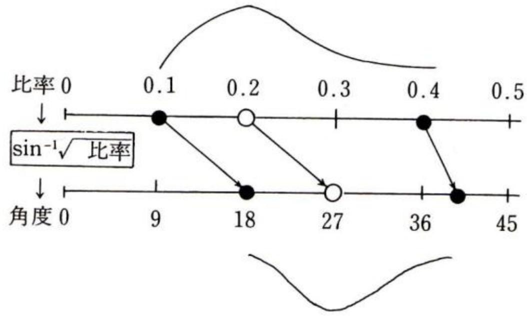
対数変換は、時間や試行数などの「開いた尺度」のデータに対して適用する。開いた尺度とは、「上限がなく、無限に大きな値をとり得る尺度のことである（例えば、時間は 0 秒から何千・何万秒まで際限なく値をとり続けることが出来る。)。このような（正確に言うと片方に）開いた尺度は、開いている方向に分布の「すそ野」がひっぱられて、L 字型分布をつくりやすい。そこで、下式によって対数値（常用対数）に変換する。

$$対数変換値 = \log_{10}(データの値)$$

なお、データの値が 1 未満の場合、変換値はマイナスとなり、変換効率もよくない。したがって、例えば 0.5 秒などのデータがある場合、データ全体の単位をミリ秒に直して 500 を代入する。

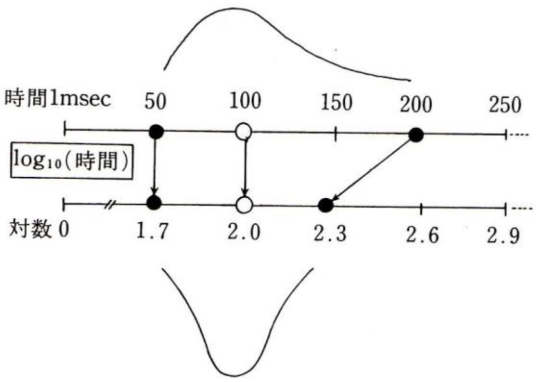
角変換と対数変換の事例

右図は、L 字型分布する比率及び時間のデータに対して、それぞれ角変換または対数変換を適用した例である。図中には、データの分布の両端（●）と中心（○）だけを示したが、最初は左端に寄っていた中心が、変換後は中央へ移動したことを確認していただきたい。また、分散も多少圧縮されている(●−●の距離を比較してみよう)。



<u>その他の変換とポアソン分布</u>

角変換と対数変換は、データ分布が L 字型あるいは J 字型のときに正規分布に近づけるための変換法であった。この 2 つの分布のうち L 字型分布は、めったに起こらない現象についての分布であり、発見者の名をとってポアソン分布と呼ばれる。この種の偏った分布に対しては、角変換、対数変換のほかにも開平変換や逆数変換が用いられる。これら 4 つの変換法の変換効率の大きさは、大きい順に逆数変換、対数変換、角変換、開平変換となるので、L 字型分布の偏りの程度にあわせて使い分ける。



開平変換値＝√（データの値)、データの値が 10 以下のときは、

$$=\sqrt{(データの値)} + \sqrt{(データの値+1)}$$

逆数変換値＝1/（データの値)、データの値に 0 があるときは

$$=1/(データの値+1)$$

どの変換法がよいかは試行錯誤的に決めるしかない。すなわち変換後のデータ分析ができるだけ正規形になるように、また各条件で等分散になるように、いくつかの変換法を試みることになる。ただし、これは技術的処置にすぎず、もともと、そのデータが正規分布するという理論的・経験的な根拠が必要であることを重ねて強調しておく。もし変数変換が適当ではなく、かえって真の分布をゆがめてしまう恐れがあるような場合は、間隔尺度データの処理方法をあきらめて、順位尺度あるいは名義尺度データの処理方法へ移行する方法も試みるべきである。

分散分析の事例Ⅱ：処理の方針を誤った例

事例

論文名「指示代名詞アノ・ソノの距離感について」

目的　指示代名詞アノ・ソノに対する心理的距離感を比較してみる。

方法　被験者：大学生 30 人

手続き：実験者が、ある一定の場所に立つ。そこから前方に一本のテープを張っておく。実験者は、被験者にコップを渡し、「私の前に張ってあるテープの「アノ」場所に、コップを置いてください」と指差しをせずに教示する。被験者がコップを置いてから、そこまでの距離を cm 単位で測定する。教示中の指示代名詞は被験者ごとに「アノ」と「ソノ」を順番に入れ替える。

結果　表 27 は、各代名詞に対する被験者の反応距離の平均と標準偏差を示したものである。分散分析の結果、指示代名詞の要因は有意であった（$F(1, 28) = 4.67$、$p < .05$）。したがって、「アノ」に対する距離感のほうが、「ソノ」に対する距離感よりも大きいと言える。

表 27　指示代名詞に対する反応距離（cm）

	アノ	ソノ
N	15	15
X	96	52
SD	38	66

解説

表 27 の SD に注目すると、かなり開きがあり、等しいとみなせるかどうか、確信がもてない。このような場合は、分散の有意差検定を行ってみるとよい。それによれば、実際は「アノ」と「ソノ」の分散には有意差がある。したがって、分散分析の結果、F 比は有意であったが、それは平均の差によるものでない可能性がある。かくして、この分散分析の結果は多義的であり、一定の結論を主張することはできない。ここで、とるべき道は二つある。一つは、両代名詞の分散の差が、なにか非本質的な原因（例；教示の誤解）によってもたらされたと考えることである。この場合は、たんなる実験の失敗であるので、例えば「アノの距離感のほうがソノの距離感よりも大きい」というような研究仮説を取り下げる必要はない。今回のデータの質が悪かっただけのことである。前述した対処の方法を取るか、データを再収集すればよい。しかしながら、もう一つには、これが事実である、と考えることである。その場合、研究仮説を根本的に修正しなければならない。そもそも、指示代名詞には、それぞれ固定した距離感があると思い込んだので、X に注目して分析したわけであるが、そのような理論的・経済的根拠は薄弱である。むしろ、「アノ」のほうが比較的一定の距離感をもつが、「ソノ」のほうは人によって距離感がまちまちであり、流動的であるとみたほうが正しいようである。もしそうなら、分散分析の適用による平均の有意差検定はデータ処理の方針を誤ったことになる。最初から、両代名詞の SD を比較し、その差を検定するべきであった。

分散の有意差の分布

データの処理の主流は、平均の有意差の分析である。このため、最初から SD に注目して分析するというケースは、あまり出て来ない。分散（SD）の有意差検定の使用頻度は少ない。例えば、t 検定の前提（2 条件の分散が等しいとみなせるか）などの確認のために利用される程度である。ここでは、もっとも単純な 2 つの分散の有意差検定を解説する。

2 つの分散の有意差検定

この検定は、データに対応がない場合の 2 条件件の分散の差を検定する。

概要

分散の有意差検定には F 比を利用する。この点、分散分析と同様である。ただし、分散分析の場合は、F 比の分子は平均の差による分散値、分母は偶然誤差による分散値であり、平均の差が偶然以上かどうかを判定していたため、実質的には平均の有意差検定になっていた。しかし、基本的な考え方は同じである。すなわち、一方の分散が他方の分散の何倍であるか（F 比）を求めて、この F 比の値が偶然に出現する確率を調べる。そして、それが 5%未満の基準をクリアするかどうかをみればよい。

検定例

「分散分析の事例Ⅱ　処理の方針を誤った例」を、分散の有意差検定を用いて再分析している。

表 27　指示代名詞に対する反応距離（cm）

	アノ	ソノ
N	15	15
X	96	52
SD	38	66

この表において、前例では、すぐに分散分析を試みて平均（X）の有意差検定を行った。しかし、ここで「距離感が大きいか小さいか」という観点から X を比べるのでなく、「距離感が一定であるかバラついているか」という観点から分散を比べる。

①各条件の自由度 1 個分の分散を計算する。

F 比の計算には、「自由度 1 個分の分散」を用いる。分散分析における F 比も、自由度 1 個分の分散「平方和/自由度」を用いた。ここで、データ表示の SD は「データ 1 個分の偏差」であるので、それを二乗した SD^2 が「データ 1 個分の分散」になる。これを「自由度 1 個分の分散」に変換するには、下式による。

自由度 1 個分の分散＝SD^2・N/（N－1）

上式の分子では、分散（SD^2）にデータの個数（N）をかけ、いったんデータの全分散を出している。これを自由度（N－1）で割って、自由度 1 個分を求める。ちなみに、この「自由度 1 個分の分散」は、データの背後に存在する無限データ集団（母集団）の分散の推定値である（不偏分散）。

②F 比を計算する。

F 比の計算は、「自由度 1 個分の分散」を分子・分母にとって、その比を計算する。このとき、必ず値の大きいほうを分子におく。本例の場合、「ソノ」のほうが分子になる。

F 比＝（ソノの自由度 1 個分の分散）/（アノの自由度 1 個分の分散）＝（66^2×15/（15-1））/（38^2×15/（15-1））＝3.02

③F 比の有意性を検定する。

F 比の分子の自由度と分母の自由度をもって F 分布表にあたる。本例の場合、どちらも 14 であるが、この自由度は表にないので、それよりも小さい df①＝12、df②＝12 を代用する。F 分布表において、df①＝12、df②＝12 の数列に F＝3.02 を対照にしてゆくと、.05＜p＜.10 となる。F＝3.02 は「ソノ」の分散が「アノ」の分散の 3 倍であることを示している。検定の結果、この F 比の偶然下での出現確率は 5%未満 1%より大きいであった。これは有意である（両側検定）。したがって、「ソノ」のほうが「アノ」よりも、データのバラつきが大きいといえる。

結果の論文記載例

「表 27 は、各代名詞に対する被験者の反応距離の平均と標準偏差を示したものである。標準偏差の大きさがかなり異なることから、両代名詞の違いは分散の差に表れているとみて、検定を行った結果、有意であった（両側検定：F（14、14）＝3.02、.01＜p＜.05）。したがって、「アノ」の距離感が比較的一定であるのに対して、「ソノ」の距離感は人によってまちまちであり、広い使われ方をする代名詞であることが示唆される。」

検定結果において、「両側検定」のコトバを必ず付記すること。この点、t 検定と同様、分散の有意差にも片側・

両側の区別がある。

なお、F（14,14）＝3.02 の検定は、実際には df①＝12、df②＝12 の F 分布を代用したのであるが、本来の自由度を書くようにする。また、代用の場合は、そのように小さい自由度の分布を用いること。

分散の有意差検定の事例；N がかなり大きい場合

事例

論文名「個性重視の指導プログラムの特性」

目的　生徒の個性を重視する指導プログラムには、どのような特性が必要かを、ケース・スタディとして探索した。その際、参考資料として、実際に対象校で用いられているプログラムが個性的な生徒を産出しているかどうかを調べてみた。

方法　対象者：4 つの高等学校の 3 年生 984 人。このうち、3 校はケース・スタディの対象校、1 校は非対象校である。

手続き　自作の生活スタイル・テストを対象者に課した。これは生活スタイルが合理化されている程度を測定するための質問紙である。50 項目の質問に「はい」か「いいえ」で回答する形式をとり、合理的回答 1 点、非合理的回答 0 点として得点化した。

結果　回答不備の者 12 人を集計から除外した。表 29 はケース・スタディの対象校 3 校と非対象校 1 校の生活スタイル特典を示したものである。得点が高いほど、生活スタイルが合理化されていることを示す。

表 29　各校の生活スタイル得点（満点 50）

	対象校（3 校）	非対象校（1 校）
N	640	332
X	33.6	32.5
SD	12.5	10.1

分散の有意差検定の結果、対象校と非対象校の分散の差は、有意であった（両側検定：F（639,331）＝1.53、p＜.05）。したがって、対象校の生徒のほうが生活スタイルの幅が大きいと言える。すなわち、きわめて合理的な生徒から経験的な生徒、試行錯誤的な生徒まで、多種多様な個人が見出されるということになろう。

解説

実証的研究としては、不備が目立つ。例えば、データの無作為抽出が仮定できない（学校規模や入学者の傾向が等質でない恐れがある）。しかしながら、問題論が先行した研究なので、それらの点は割り引いて評価すべきであろう。分散の有意差検定の適用例としてはユニークである。さて、検定結果は有意であったが、注意すべきは N の大きさである。全ての有意差検定または有意性検定に当てはまることであるが、N が大きいと、ほんのわずか差でもすぐに有意になりやすい。したがって、本例のように N がかなり大きい場合は、有意差で満足せず、実践的な見方をとることができる。例えば、表 29 の SD をみると、対象校は非対称校よりも 2.4 ポイント大きい。すなわち、比喩的に言えば、対象校のプログラムは、非対称校の生徒の約 7 割の分布を＋・－両方向に 2.4 ポイントだけ広げる効果を持っているといえる（50 項目中 5 項目程度の質問に異なる回答が出ることを意味する）。この効果が今回の対象者数に比べて実践的に価値のある大きさであるかどうかを検討してみなければ、ほんとうに対象校と非対称校との間に差異があるとはいえない。有意差とは、たんに偶然には生じない差をいうのである。偶然には生じなくても、取るに足らない、つまらない差もある。

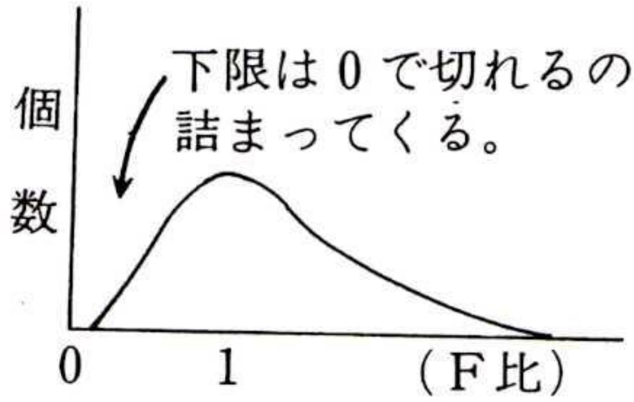
片側検定と両側検定

分散分析で平均の有意差検定を行う場合に、両側検定で統一していたが、本来、分散の有意差検定や t 検定は、片側検定と両側検定の 2 通りの方法をもっている。この区別を知るには、検定に利用する偶然分布（F 分布や t 分布）を理解しなければならない。ここでは、F 分布を取り上げる。例えば、分散の有意差検定の場合、F 比は下式によって計算した。

$$F = SD_A^2 / SD_B^2$$

ここで、F 分布は、両方の分散の差が偶然によって生じていると仮定している（帰無仮説）。したがって、まったく偶然の影響が無ければ、両分散は等しく（$SD_A^2 = SD_B^2$）、F 比は 1 となる。すなわち、F 比は 1 に収束するので、F 分布は 1 の値の近辺に頂点を持つ分布となる。たまたま、偶然に分子の SD_A^2 の値が大きくなるなら、F 比は 1 以上になるが、反対に分母の SD_B^2 の値が大きくなるなら、F 比は 1 未満の 0 に近づいた値になるであろう。

このＦ比の偶然分布において、公認の有意水準である出現確率 5%未満の値をあらかじめ選び、「有意差の区域」を設けておくとする。これをどこに設けるかが、両側検定・片側検定の区別になる。ここで、両側検定は、分布の左右から 2.5%ずつをとり、両方あわせて全体の 5%分の値を（有意な値として）用意する。これに対して、片側検定は、分布の片方の端から全体の 5%の値を選び、片側だけに「有意差の区域」を置く。



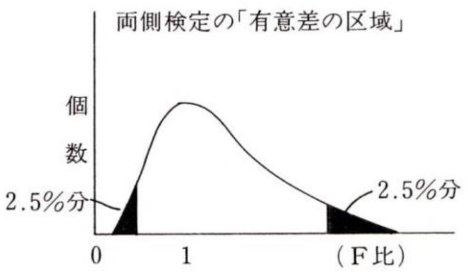
このように、前者の両側検定では、両分散のいずれが大きくなるのかを、前もって限定しない。したがって、F 比が 1 未満の方向へ落ちても、1 以上の方向へ落ちても対処できるようにしている。これに対して、片側検定では、データの F 比が、F 分布の片方だけに落ちるとみている。すなわち、F 比の分子・分母どちらかの分散のほうが大きくなるはずである。

両側検定に統一している理由

上の図からわかるように、分布の片方だけについてみれば、片側検定のほうが「有意差の区域」を両側検定の 2 倍も広く用意している。このため、データの F 比は相対的に小さな値で有意性を得ることができる。その意味では片側検定は「得な検定」と言える。しかし、もし片側検定が「有意差の区域」を用意しておかなかった方向へデータの F 比が落ちてしまったら、まったくお手上げになる。どんなに大きな F 比があっても、そこには「有意差の区域」がないので、有意性を認めようがない。このように、片側検定は「甘い成功」か「完全な敗北」かの賭けに似ている。科学研究に似合わない印象がある。

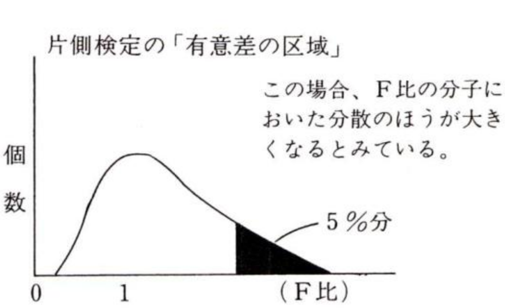
現実には、一つの F 比は分布の両側に落ちることはなく、必ずどちらか一方にしか落ちない。したがって、データをみてから F 比の落ちる方向へ片側 5%の有意な値を用意すれば完敗はないのであるが、こうしたドロナワ式の検定も科学に似つかわしくない。

以上の理由によって、ここでは、片側検定を全く用いず、両側検定に統一している。この方針による実用上の障害は、ほとんどない。片側検定で有意になる値は、両側検定でも有意傾向に含めて拾うことができる。



片側・両側検定の区別のある方法とない方法

ここでは、片側検定と両側検定の区別がある方法については、両側検定だけを一貫して採用する。例えば、上述した分散の有意差検定や t 検定、直接確率計算法などがそうである。これに対して、両側検定・片側検定の区別がない方法もある。分散分析や χ^2 検定がそうである。実際は、これらの方法に



おける検定は当該分布の片側にしか「有意差の区域」をもうけないので、分布の両側を使用しているか片側を使用しているかという観点から言えば「片側検定」である。例えば、分散分析の F 比の分子は「平均の差による分散」、分母は「偶然誤差による分散」であったが、分析結果としては「分子＞分母」（平均の差は偶然誤差以上である）と「分子＜分母」（平均の差が偶然誤差の範囲内である）の二通りしか想定することができない。すなわち、「有意差の区域」は分布の右側だけにおかざるをえない（左側はすべて n.s.）。

しかし、分布の使い方でなく、無帰仮説に対する対立仮説を両側に設定しているか片側に設定しているかとう観点から言えば、分散分析は「両側検定」である。例えば、分散分析による 2 条件の平均の有意差検定の場合、検定上の帰無仮説「$X_A = X_B$」に対する対立仮説「$X_A \neq X_B$」である。すなわち、実質的に「$X_A > X_B$」と「$X_A < X_B$」の両方向の差を検定している。

このような紛らわしさを避けて、例えば分散分析は「片側分布の両側仮説」であると表現することもあるが、研究の実用上は片側・両側の区別のない方法であると考えたほうがわかりやすいであろう。

以上、検定結果の記述に「両側検定」のコトバを必要としない方法は、実質的に両側検定にしかならない方法なのである。

４　実験計画法と分散分析

前章では、分散分析の基本概念を理解するため、データに対応がない場合の単純な２条件比較の分散分析だけを解説した。本章では、実験計画法と組み合わせて、もっと複雑な分散分析を解説する。その前に、ここで、実験計画法と分散分析がセットで取り扱われている理由を述べる。

実験計画法と分散分析の関係

実験計画法とは、計画的な条件設定によって、計画的にデータの値を変化させる方法である。これに、合わせて、データ分析の方針も、入手したデータの計画的変化分と無計画的変化分（偶然のユレ）とを区別することになる。そして、研究者の意図した計画的変化分が、無計画的な偶然のユレよりも大きいかどうかを判定する。このような方針にそったデータ分析の方法が、分散分析である。

実験計画法：$\boxed{\text{データの総変化}}$＝$\boxed{\text{計画的変化}}$＋$\boxed{\text{無計画的変化分}}$

分散分析：　$\boxed{\text{データの全分散}}$＝$\boxed{\text{条件設定の効果}}$＋$\boxed{\text{偶然の誤差}}$

以上のように、実験計画法と分散分析は一貫した方法の連係をなしている。したがって、各種の実験計画法に対応したいろいろなタイプの分散分析がある。このうち、本章では専門的に言えば母数模型、直交計画、加法モデルに基づく３要因までの分散分析を解説する。事前に、実験計画法の用語を紹介する。

①要因（factor）

要因とは、データの値を変化させる原因のことである。ただし、なにも条件設定をおこなわないときもデータの値は偶然のユレによって変化するが、偶然要因ということはない。ふつう、要因とは、人為的な条件設定による「実験要因」のことだけである。

②水準（level）

水準とは、要因を構成する条件のことである。実験計画上、１要因は必ず２水準以上によって構成されていなければならない。例えば、「性別」という要因は男女２水準、また、「学習環境」という要因は「騒音」・「音楽」・「無音」の３水準によって構成される。前者は水準数が固定している例であり、後者は研究目的によって任意に水準数を増減できる例である。なお、１要因の実験計画では、その要因の水準数がそのまま条件の数になる。２要因を組み合わせた実験計画では、両要因の水準数を掛け合わせた数が条件の数になる（上例の性別２水準と学習環境３水準を組み合わせた場合は２×３＝６の６条件）。

③独立変数と従属変数

独立変数とは要因又は水準、従属変数とはデータの値（解答時間やテスト得点など）のことである。両者の関係は「独立変数が従属変数を変える」となる。「独立変数はなにか」、「従属変数はなにか」というふうに自問自答や質疑応答に用いると便利である。

４－１　実験計画の手順

さて、計画的にデータの値を変化させ、分散分析に持ち込むには、周到な実験計画を組まなければならない。その手順は、以下のようになる。

①要因数を決める。

研究仮説の事情によって異なるが、実験計画に持ち込む要因数は最大３個までにおさえるべきである。４要因以上になると、高次の交互作用が有意に出た場合、解釈困難に陥る（我々の研究対象である人間は特に交互作用を起こしやすい）。「現象は複雑である」という結論は、単に実験計画の失敗を意味するに過ぎない。むしろ、３要因の場合でさえ、３要因の実験１回なら２要因の実験２回とするほうが賢明である。

②水準数を決める。

各要因の事情によって異なるが、水準数は、１要因計画の場合は５水準まで、２要因計画や３要因計画の場合は各要因３水準までに抑える。これ以上に水準数が増えると、経験的な印象であるが、分析と解釈に明快さを欠く傾向がある。「多水準小実験」よりも「小水準多実験」を心がけるようにする。

さて、要因数と水準数が決まると、必然的に条件の数が決まる。例えば、１要因の３水準なら３条件、２要因

の2水準と3水準なら6条件、3要因の「2×2×3」なら12条件となる。
③1条件当たりのデータの個数を決める。

データの個数は、1条件20個以上を努力目標とする。実際に、この程度の個数がないと、データのヒストグラムを描いても分布の形を判定できない。ただし、もしデータの分布形や分散の大きさについて事前知識があるなら、10個前後でもよい。なお、各条件のデータの個数は等しくするのが原則である。そのほうが、データ分析の検定力も高い。しかし、実際には実験の不備や試行の失敗によって不ぞろいになる場合も多い。
④募集する被験者数を決める。
実験計画には、被験者間計画と被験者内計画がある。この選択によって、必要となる被験者数は大きく異なってくる。ここで、**被験者間計画**とは、実験条件を異なる被験者に割り当ててゆく計画である。したがって、実験条件は、いわば被験者と被験者の間に配置されることになる（被験者間配置)。これに対して、**被験者内計画**とは、実験条件を同じ被験者に割り当ててゆく計画である。この場合、実験条件は、被験者個々人のなかに配置されることになる（被験者内配置)。したがって、例えば「1要因3水準」で「1条件20個」のデータを確保しようとする場合、被験者間計画では、この3条件を別々の被験者に割り当てるため合計60人の被験者が必要になる。しかし、被験者内計画では3条件を1人に割り当てるので、20人の被験者ですむ。このように、被験者募集の労力だけを考えると、被験者内計画のほうが経済的であるが、それなりの短所もある。

被験者間計画と被験者内計画

実験計画を組む場合、被験者間計画（between subjects design）をとるか、被験者内計画（within subjects design）をとるかは、非常に重大な選択である。そこで、両者の長所・短所を比較して、どこに選択の基準を置けばよいかを考えてみる。
①被験者内計画は被験者数が少なくてすむ（説明省略)。
②被験者内計画は実験前の等質化の必要がない。

公平な実験的比較のためには、各条件の実験前の特性や状態は等質にしておかなければならない。この点、被験者内計画では、各条件を同じ被験者に割り当てるので、条件の比較は同一ベース上でおこなわれることになり、問題は無い（等質化の必要がない)。しかし、被験者間計画では、各条件を異なる被験者に割り当てるため、各条件の実験前の特性や状態は基本的に等質ではない。そこで、等質化の手続き（労力がかかる）または無作為抽出の仮定（危険をともなう）が必要となる。
③被験者内計画の分散分析は検定力が高い。

被験者内計画によって入手したデータの分析には、それ専用の被験者内計画の分散分析を適用する。もし、これに被験者間計画の分散分析を誤って適用すると、有意な結果が有意でなくなる場合がある。そのように、被験者内計画の分散分析の方が検定力が高い。さて、前頁は被験者内計画の長所ばかりであったが、次に短所を挙げてみる。
①被験者内計画は攪乱要因が混入しやすい。

攪乱要因とは、データをかき乱す原因のことである。攪乱要因が混入すると、データは計画通り変化しない。または、質が落ちる（分布が正規形から逸脱する。各条件の分散がふぞろいになる)。例えば、被験者内計画では、同じ被験者が全条件を試行するため実験時間が長く、活動性のレベルが低下しやすい（疲労効果)。また、前の条件の反応と同一方向に次の条件の反応が引きずられて差異が表れにくくなったり（平均化傾向)、条件設定の意図を察知して、被験者が作為的に反応を調整したりする（作為的反応)。これらは、被験者内計画に特有の攪乱要因である。特に最後の点は、いわば実験計画が「バレる」ことであり、バレてはいけない研究の場合は被験者内計画は避けるべきである。
②被験者内計画は順序効果、練習効果を相殺する労力がかかる。

被験者内計画では、実験前の各条件の特性や状態は等質であるとしても、被験者が各条件を試行する順序については、等質であるとはいえない。ここで、条件の試行順序が被験者の反応に及ぼす影響を順序効果と呼ぶ。順序効果は、何とは特定せず、不測の影響全般を含めている。このうち、被験者の反応が改善されるような順序効果を特に練習効果と呼ぶ。これらを除去するためには、カウンター・バランスまたはランダマイゼーションの手続きをおこなわなければならない。

カウンター・バランスと（相殺）とは、各条件がすべての試行順序に当たるように配慮することである。完全なカウンター・バランスをおこなうために必要な被験者の人数は、「条件の階乗」（順列）になる。例えば、3 条件の場合は、3 の階乗すなわち 6 人が必要である。

被験者/試行順序	1回目	2回目	3回目
イ	条件A	条件B	条件C
ロ	条件A	条件C	条件B
ハ	条件B	条件A	条件C
ニ	条件B	条件C	条件A
ホ	条件C	条件A	条件B
ヘ	条件C	条件B	条件A

なお、上例においてデータの個数を増やした場合、6 の倍数ずつとってゆくことになる。

さて、完全なカウンター・バランスが困難であるときは、ランダマイゼーションをおこなうことになる。ランダマイゼーション（無作為化）とは、乱数を用いて、無作為に条件の試行順序を決めることである。例えば、3 条件の場合、あらかじめ各条件に１、２、３と番号をつけておく。そして、被験者ごとにサイコロをふったり、二桁の乱数を 3 で割った余り（０，１，２）に 1 を足した数を求めたりして、試行順序を決める。一般的には、（乱数÷条件数）の余り＋１、を用いる。実際的な例で言えば、本来、質問紙や評定用紙の各質問項目も、その掲載順序または提示順序をランダムに入れ替え、複数のバージョンを作成するべきである。しかし、作成と集計の手間がかかるので、一定順序のワン・バージョンで済まされているのが実情である。

以上、被験者間計画と被験者内計画の長所・短所を比べてきた。どちらを選ぶかは、基本的には研究目的によるが、一般論としては、「等質化の労力」が大きすぎる場合は、被験者内計画、「攪乱の要因の混入」が大きすぎる場合は被験者間計画を選択すればよい。

折衷型の実験計画：混合計画 mixed design

被験者間計画と被験者内計画の折衷型として混合計画がある。これは 2 要因以上の場合に、下表のように一方の要因（A）を異なる被験者に配置し、他方の要因（B）を同じ被験者に配置する方法である。

表 30　2×2　の混合計画のデータ・リスト

被験者	A①		被験者	A②	
性・名	B①	B②	性・名	B①	B②
男山田	26	58	女島崎	43	87
女高野	64	52	男黒木	65	71
男村井	39	77	女川谷	71	66
・・・・・	・・・	・・・	・・・・・	・・・	・・・

実験計画の論文記載例

以上の実験計画は、以下のような書式で論文に記載する。

「実験計画：１要因 5 基準の被験者間計画。被験者：成人 40 人。」

「実験計画：2×2 の被験者内計画。被験者：幼児 20 人。」

「実験計画：2×3 の混合計画。第一要因は被験者間配置、第二要因は被験者内配置。被験者：高校生 50 人」

これらの記載をみたら、1 条件に何個のデータがはいるか（上から 8 個、20 個、25 個）。

実験計画と分散分析のタイプ

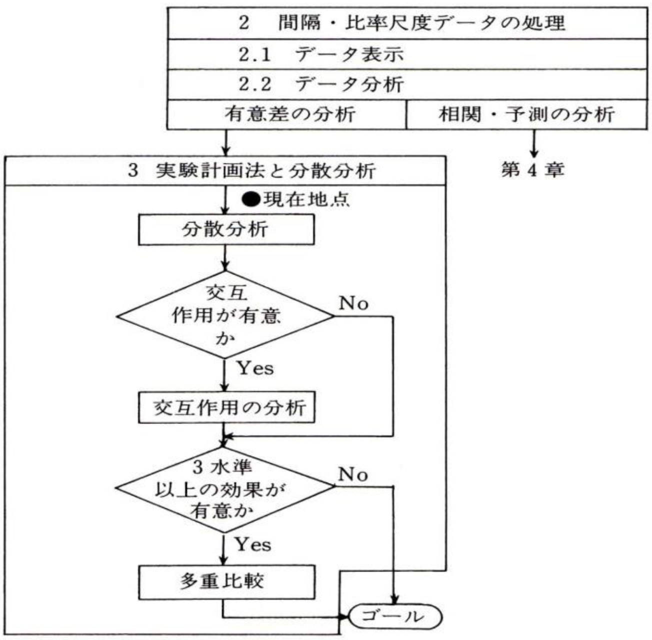
分散分析は、実験計画のタイプに合わせて少しずつ計算のしかたが異なる。したがって、以下、実験計画に分散分析の方法を解説する。下表は、分散分析の一覧であるが，解説は 3 要因までである。それ以上は手計算のレベルでなく、また、結果の解釈に困難を招く恐れも多い。

表 31　分散分析の一覧

実験計画の記述		分散分析の略号
1 要因	被験者間計画	AS タイプ
	被験者内計画	SA タイプ
2 要因	被験者間計画	ABS タイプ
	被験者内計画	SAB タイプ
	混合計画	ASB タイプ
3 要因	被験者間計画	ABCS タイプ
	被験者内計画	SABC タイプ
	混合計画 I	ABSC タイプ
	混合計画 II	ASBC タイプ

表中の分散分析の「略号」は、A、B、C がそれぞれ独立の 1 要因を示す。S は被験者（subjects）を示す。S の左側に書かれた要因は被験者間配置である（その要因の各水準は異なる被験者に割り当てられる）。S の右側に書かれた要因は被験者内配置である（その要因の各水準は同じ被験者に割り当てられる）。これらの略号は「1 被験者 1 行」でデータを入力するときの書式に一致しており、使い慣れると便利である。

なお、実際には分散分析そのものよりも、分散分析の結果に関する事後処理のほうが重要であり、複雑である。そこで、あらかじめ「事後処理のフローチャート」を掲載する。まだ説明していない用語もあるが、これまでのコースもふり返りながら、よく覚えていただきたい。



４－２　一要因の分散分析

a. 被験者間計画（AS タイプ）

適用

一要因の分散分析は、その要因を構成する各水準の平均どうしの有意差検定になる。この場合、ふつう要因名はおもてに出ない。ここでは、被験者間計画であるので、各水準を異なる被験者に割り当てる（下段のデータ・リスト参照)。なお、一要因被験者間計画の分散分析（AS タイプ）のモデルは、次のとおりである。

データの全分散（総変化分）＝各水準の平均の差による説明分＋偶然の影響による説明分（誤差）

計算例

実験計画は、「一要因 3 水準の被験者間計画」（要因名 A）とする。また、データの個数は、各条件 4 個とする。したがって、被験者は 12 人必要である。以下に、データ・リストを掲載し、データ表示をおこなっておく。

表 32　データ・リスト

A①は要因 A の第 1 水準を示す。

Sub.A①	Sub.A②	Sub.A③
イ　1	ホ　3	リ　3
ロ　3	ヘ　3	ヌ　4
ハ　2	ト　4	ル　5
ニ　3	チ　4	ヲ　6

表 33　データ表示

	A①	A②	A③
N	4	4	4
X	2.25	3.50	4.50
SD	0.83	0.50	1.12

①　大平均㊥を計算する。

　大平均㊥＝3.42

②　各条件の平均によって生じる分散を計算する。

条件の分散＝（2.25－3.42）^{2}×4＋（3.50－3.42）^{2}×4＋（4.50－3.42）^{2}×4＝10.17

③　偶然によって生じる分散（誤差）を計算する。

偶然の誤差＝0.83^2×4＋0.50^2×4＋1.12^2×4＝8.77

④　分散分析表を作成し、F 比を計算する。

表 34　分散分析表

要因	平方和	df（カッコ内計算式）	平均平方	F
条件	10.17	2（条件数－1）	5.09	5.25*
誤差	8.77	9（全データ数－条件数）	0.97	
全体	18.94	11（全データ数－1）		＊p＜.05

⑤　F 比を検定し、検定結果を分散分析表に書き込む。

分散分析表は、すでに F 比の検定結果を記載しているが、この場合の F＝5.25 は有意である（p＜.05）。

⑥　F 比が有意である場合は多重比較へ進む。

F 比の有意性は、各条件の平均どうしの間に有意差が存在することを示唆している。ただし、上例の場合、平均の差は 1 つではなく、A①－A②、A②－A③、A①－A③、の 3 つである。したがって、どの平均の差が有意であるのか、わからない。このように 3 条件以上（すなわち 3 水準以上）の分散分析において F 比が有意であった場合は、どの平均の差が有意であるかを調べるため、「多重比較」と呼ばれる事後処理を行わなければならない。1 要因 2 水準の場合、その必要はなく、ここで終わる（F 比の有意性はそのまま 2 条件の有意差である）。

多重比較

定義

多重比較（multiple comparison）とは、3 つ以上の平均がある場合に 2 平均ずつを対にした「多数回の比較」という意味である。多重比較の方法には、テューキーの法やシェッフェの法などがあるが、これらはタイプ I エラー（ほんとうは差がないのに有意差を検出する）への防御力が高い。しかし、その分だけ、タイプⅡエラー（ほんとうは差があるのに有意差を検出しない）を過小評価するきらいがある。そこで、ここでは、分散分析の結果が有意なときに限って用いることを条件に、検定力が高く計算の容易な LSD 法を紹介する。

備考：多重比較に t 検定を用いない理由

多重比較は、結局、2 平均の有意差検定になるので、これに t 検定を用いてもよさそうである。しかし、 t 検定は上述したタイプ I エラーにあまいことが知られているので用いない。簡単に言えば、 t 検定は、1 回 1 回の検定のたびに偶然誤差を推定するので、その大きさは一通りに決まらない。すると、多数回の推定のうちには、たまたま小さな偶然誤差が出現することもある（小さな偶然誤差は t 値を大きくする）。いわば、クジを引く回数が多いと、それだけ有意性に当たるチャンスも大きくなる。ただし、特定の 2 条件について特定の仮説があるなら、分散分析の事後に、その 2 条件に限って t 検定を適用することが出来る。この場合の 2 条件の有意差は、分散分析における要因の有意差とは全然関係がなく、全く別の仮説に基づく主張となる。要するに、 t 検定を分散分析の事後処理に使用した場合、 t 検定は「一仮説・一偶然誤差」の原則を守れないのである。

LSD 法

定義

LSD（Least Significant Difference）とは「最小有意差」という意味である。すなわち、LSD 法は、有意水準をクリアする最小の幅（LSD）を求める。そして、2 条件の平均の差が、その幅より大きければ有意、小さければ有意でないと判定する。

計算例

前例のデータ表示と分散分析の結果を再掲する。

表 33　データ表示

	A①	A②	A③
N	4	4	4
X	2.25	3.50	4.50
SD	0.83	0.50	1.12

表 34　分散分析表

要因	平方和	df（カッコ内計算式）	平均平方	F
条件	10.17	2（条件数－1）	5.09	5.25*
誤差	8.77	9（全データ数－条件数）	0.97	
全体	18.94	11（全データ数－1）		＊p＜.05

① ｔ値を用意する。

LSD の計算には有意水準 5%のｔ値が必要である。これには、分散分析の「誤差」の自由度（df＝9）に相当する 5%点のｔ値を求める（ｔ分布表より自由度 9、5%のｔ値＝2.262 である)。

② 「誤差」の平均平方（MSe）を用意する。

表 34 の分散分析表から「誤差」の平均平方（MSe＝0.97）を読み取っておく。なお、“MSe”は誤差（error）の平均平方（MS）を表す専用の記号である。

③ ｔ＝2.262 と MSe＝0.97 を下式に代入し、LSD を計算する。

$$LSD = t値 \cdot \sqrt{(2 \cdot MSe / 1平均あたりのデータ数)} = 2.262 \times \sqrt{(2 \times 0.97/4)} = 1.58$$

なお、各水準または各条件のデータの個数が等しくない場合は下式による（N_1、N_2 は比較したい 2 条件のデータの個数)。

$$LSD = t値 \cdot \sqrt{(MSe \cdot (1/N_1 + 1/N_2))}$$

④ 検定を行う。

各２条件ずつ平均の差を求め、LSD＝1.58 より大きいか小さいかを判定する。本例の場合、N が等しいので、一つの LSD をすべての比較に用いることが出来る。しかし、N が等しくない場合は、各比較ごとに異なる LSD を計算しなければならない。

$$X_1 vs X_2 : |2.25-3.50| = 1.25 < LSD \Rightarrow n.s.$$
$$X_2 vs X_3 : |3.50-4.50| = 1.00 < LSD \Rightarrow n.s.$$
$$X_1 vs X_3 : |2.25-4.50| = 2.25 > LSD \Rightarrow 有意$$

LSD＝1.58 は、有意水準 5%をクリアする最小の差である。これより大きければ有意であり、小さければ有意でない（n.s.)。

結果の論文記載例

これまでの分散分析と多重比較の結果は、以下のように論文に記載する。

「表 33 は、各条件の平均と標準偏差を示したものである。分散分析の結果、表 34 に示したように、条件の効果は有意であった（$F(2,9) = 5.25$、$p < .05$)。LSD 法を用いた多重比較によれば、条件 A①と条件 A③の間に有意差があった（MSe=0.97、5%水準)。しかしながら、条件 A①と条件 A②、および条件 A②と条件 A③の間の差は有意でなかった。」

上の文章中、第 1 文がデータ表示、第 2 文が分散分析、第 3 文以降が分散分析の事後処理としての多重比較である。ここで、多重比較の結果の記述において、「MSe＝0.97」が付記されていることに注意する。これは、LSD の計算に用いた偶然誤差の平均平方である。論文の読み手は、この MSe の値をみれば LSD の大きさを算定できる。もちろん、LSD そのものを記載してもよいが、あまりスマートではない（各条件の N が異なると、一つではすまなくなるので)。

また、LSD の有意水準は、上文の「5%水準」のように必ず明記すること。一般には、公認の有意水準に合わせて 5%水準だけを採用する。それ以外の有意傾向や 1%水準はとらない。なお、多重比較の検定は、分散分析の事後処理なので両側検定のみ可能である。片側検定はありえない。したがって、「両側検定」のコトバを付記する必要ない。

LSD 法の制約

LSD 法の使用には、以下の制約がある。

- 分散分析の結果が有意でない場合は使用できない。
  この制約は LSD 法に限らない。一般に多重比較は、分散分析の結果が有意でない場合に行ってはならない。あくまでも、先行の分析が有意であった場合の事後処理である。
- 有意水準は 5%のみを採用する。
  LSD 法の有意水準は、つねに 5%とする。例えば、計算の上では、ｔ値を用意する段階で 10%点（有意傾向）や 1%点のｔ値を求めれば、有意水準を任意に与えることができる。しかし、そうしないこと。たとえ分散分析の結果が有意傾向であっても、また、1%水準の有意であっても、LSD 法は 5%水準に固定すること（分散分析の結果が有意傾向の場合は多重比較へ移行しないこともある)。

備考：多重比較の結果のわかりやすい表記

多重比較の結果を、以下のような表にまとめるのも、わかりやすい。特に、条件数が多い場合は適切である。

表 37　多重比較の結果

左項vs.右項	条件A②	条件A③
条件A①	＝	＜
条件A②		＝
	不等号　p＜.05、等号　n.s.	

b．被験者内計画（SA タイプ）

適用

1 要因の被験者内計画は、「乱塊法」（randomized block design）と呼ばれる実験計画の典型である。全水準を同じ被験者に割り当てるため、いわゆる「対応のあるデータ」が得られる。この場合、分散分析（SA タイプ）のモデルは次のとおりである。

データの全分散（総変化分）＝各水準の平均の差による説明分＋個人差による説明分＋残差

上式からわかるように、被験者内計画の分散分析（SA タイプ）では、偶然誤差から「個人差」によるユレ分を取り出す。したがって、その分だけ偶然誤差を小さくすることができる（残りの偶然誤差は「残差」と呼ぶ）。このことが検定力の高い理由である。

計算例

比較のため、前述の AS タイプで用いたデータ例を 1 要因 3 水準の被験者内計画にみたてて用いる。したがって、被験者数は 12 人から 4 人に減る。以下はデータ・リストとデータ表示である。

表 38　データ・リスト

Sub.	A①	A②	A③	合計
イ	1	3	3	7
ロ	3	3	4	10
ハ	2	4	5	11
ニ	3	4	6	13

表 39　データ表示（N＝4）　A①は要因 A の第 1 水準のこと。

	A①	A②	A③
X	2.25	3.50	4.50
SD	0.83	0.50	1.12

① 大平均（㊥）を計算する。

㊥＝（2.25＋3.50＋4.50）/3＝3.42

なお、被験者内計画では N が等しくない場合はありえない。被験者が 1 水準でも失敗したら、その被験者の全データを除外する。

② 各条件の平均によって生じる分散を計算する。

条件の分散＝（2.25－3.42）^{2}×4＋（3.50－3.42）^{2}×4＋（4.50－3.42）^{2}×4＝10.17

③ 偶然によって生じる分散（誤差）を計算する。

偶然の誤差＝0.83^2×4＋0.50^2×4＋1.12^2×4＝8.77

ここまでは被験者間計画の分散分析（AS タイプ）と同一手順・同一結果である。ここから偶然誤差 8.77 を個人差と残差に分ける。

④ 個人差によって生じる分散を計算する。

個々の被験者のデータの「合計」を用いる。合計の値は、表 38 のデータ・リストにおいてすでに計算済みである。

個人差＝（合計^2＋合計^2＋合計^2＋合計^2）/1 人当たりのデータ数－（合計＋合計＋合計＋合計）^{2}/全データ数

＝（7^2＋10^2＋11^2＋13^2）/3－（7＋10＋11＋13）^{2}/12＝6.25

なお、④の計算式は簡易式である。定義どおり計算するなら、「個々の被験者の平均」（合計/3）を求め、それと大平均とのズレを 1 人のデータ数分だけ倍にして、足し上げてゆけばよい。

個人差＝（7/3－㊥）^{2}×3＋（10/3－㊥）^{2}×3＋（11/3－㊥）^{2}×3＋・・・・・

⑤ 残差を計算する。

残差＝偶然の誤差－個人差＝8.77－6.25＝2.52

ここで、偶然の誤差 8.77 のうち、その一部分 6.25 は被験者個人個人の差異として取り出すことができたが、

残り 2.52 は特になんの影響と認定することができないので、やはり偶然の影響とみるしかない(残差と呼ぶ)。すなわち、この残差が偶然の大きさを表している。したがって、F 比の計算の際は、これが分母になる。

⑥ 各分散値を分散分析表に記入し、F 比を計算する。

自由度の計算は、下表のように「残差」の df を最後に出すようにするとまちがいがない。なお、「個人差」は今回の被験者ではたまたま 6.25 であったが、時や人を変えると値も変わるので、一般的意味がなく、検定しない。

表 40　分散分析表

要因	SS	df（カッコ内計算式）	MS	F
条件	10.17	2（水準数－1）	5.09	12.12**
個人差	6.25	3（被験者数－1）	2.08	
残差	2.52	6（2×3）	0.42	
全体	18.94	11（全データ数－1）		** p＜.01

⑦ F 比を検定し、検定結果を分散分析表に書き込む。

検定の仕方は、F 比（5.09/0.42）の分子の自由度（df①＝2）と分母の自由度（df②＝6）を用いて以下のように行う。

表 41　F 分布表

df①	df②	F 比		
2	5	・・・	・・・	・・・
	6	3.46	5.14	10.93
	7	・・・	・・・	・・・
出現確率		.10	.05	.01
有意水準		有意傾向	5%	1%

F＝12.12 を、df①＝2、df②＝6 の数列に対照してゆく。右端の値を超えるので、 p＜.01 である。

⑧ F 比が有意であったので、多重比較へ進む。

LSD を計算するため、まず「残差」の自由度（df＝6）に相当する 5%点の t 値を求める（ t ＝2.447)。次に、分散分析表から「残差」の平均平方を読み取る（MSe＝0.42)。それらを下式に代入する。

LSD＝ t 値・√（2・MSe/1 平均あたりのデータ数）＝2.447×√（2×0.42/4）＝1.12

⑨ LSD＝1.12 を用いて検定をおこなう。

X_1 vs. X_2 ：｜2.25－3.50｜＝1.25 ＞ LSD ⇒ 有意

X_2 vs. X_3 ：｜3.50－4.50｜＝1.00 ＜ LSD ⇒ n.s.

X_1 vs. X_3 ：｜2.25－4.50｜＝2.25 ＞ LSD ⇒ 有意

結果の論文記載例

「表 39 は、各条件の平均と標準偏差を示したものである。分散分析の結果、表 40 に示したように、条件の効果は有意であった（F（2,6）＝12.12、p＜.01)。LSD 法を用いた多重比較によれば、各条件の平均の大小関係は「条件 A①＜条件 A②＝条件 A③」であった（MSe＝0.42、5%水準)。」

第 3 文中の「条件 A①＜条件 A②＝条件 A③」は、多重比較の結果を要約した表記である。このような表記が不可能なケースもある。全体としては、前述の被験者間計画の分散分析（AS タイプ）の記載例とほとんど同じである。

４－３　二要因の分散分析

種類

2 要因の場合、被験者の割り当て方によって 3 種類の実験計画ができあがる。それぞれの計画によって分散分析のタイプも異なってくる。例として、要因 A（2 水準）と要因 B（3 水準）の 2 要因計画を取り上げ、被験者の割り当て方を示す。

a.　2×3 の被験者間計画（ABS タイプ）

A①			A②		
B①	B②	B③	B①	B②	B③
田中	安部	鈴木	橋口	堀井	倉沢
山際	浜本	吉川	染谷	植竹	西須
遠藤	金森	岡崎	野瀬	高野	小島

データの個数は、1 条件につき 3 個とする。被験者はどれか 1 つの条件に割り当てられる。被験者数は 18 人。

b.　2×3 の被験者内計画（SAB タイプ）

A①			A②		
B①	B②	B③	B①	B②	B③
田中	田中	田中	田中	田中	田中
山際	山際	山際	山際	山際	山際
遠藤	遠藤	遠藤	遠藤	遠藤	遠藤

被験者は全条件に割り当てられる。被験者数は 3 人。

c.　2×3 の混合計画（ASB タイプ）

A①			A②		
B①	B②	B③	B①	B②	B③
田中	田中	田中	安部	安部	安部
山際	山際	山際	浜本	浜本	浜本
遠藤	遠藤	遠藤	金森	金森	金森

要因 A は被験者間配置、要因 B は被験者内配置、被験者数は 6 人。

a.　2×3 の被験者間計画（ABS タイプ）

適用

2 要因以上の分散分析は、各要因の水準間の平均の有差検定であると同時に、「交互作用」という要因の有意性検定でもある。例えば、以下に述べる 2 要因被験者間計画の分散分析（ABS タイプ）のモデルは次のようになる。

データの全分散（総変化分）　＝

要因 A の各水準の平均の差による説明分　＋　要因 B の各水準の平均の差による説明分

＋　要因 A と要因 B の交互作用による説明分　＋　偶然の影響による説明分（誤差）

上式からわかるように、分散分析（ABS タイプ）は、データの全分散を 4 つの部分に分割する。これを専門用語に置き換えると、上式は以下のようになる。

データの全分散＝要因 A の主効果＋要因 B の主効果＋交互作用＋偶然誤差

ここで、主効果（main effect）とは、その要因だけによる単独効果のことである。これに対して、交互作用（interaction）とは、複数の要因の相乗効果のことである。この交互作用が有意であるかどうかによって、分散分析の事後処理の方法が大きく異なることを、あらかじめ念頭においておく。

計算例

実験計画は「2×3 の被験者間計画」、前者は要因 A（2 水準）、後者は要因 B（3 水準）とする。データの個数は、1 条件 3 個とする（N＝3）。全部で 6 条件あるので、被験者数は 18 人必要である。データ・リストは省略し、データ表示のみを以下に示す。

表 42　データ表示

	A①			A②		
	B①	B②	B③	B①	B②	B③
N	3	3	3	3	3	3

X	4.33	3.67	5.33	3.33	6.00	4.33
SD	0.47	0.47	0.94	0.47	0.82	0.47

○　Nが等しくない場合は便宜上のNを計算する。

本例は、各条件のデータ個数は等しく N=3 である。しかし、もし、N が等しくない場合は便宜上の「等しいN」を計算する。この点、1 要因の分散分析（AS タイプ）とは異なる。例えば、上例の 6 条件のうち一つの条件（A②・B③）が 1 個のデータを欠いた場合、便宜上の「等しい」N は次のようになる。

N＝条件数/（$1/N_{11}$＋$1/N_{12}$＋$1/N_{13}$＋$1/N_{21}$＋$1/N_{22}$＋$1/N_{23}$）＝6/（1/3＋1/3＋1/3＋1/3＋1/3＋1/2）＝2.8

この N は調和平均といわれる。そして、これ以後、各条件はみなデータの個数が等しく N=2.8 であると仮定して計算を行う。

ここで、以下の計算手順を理解しやすくするため、2×3 の条件配置をタテ・ヨコにとって、平均を掲載しておく。

表 43　各条件の平均

	B①	B②	B③
A①	4.33	3.67	5.33
A②	3.33	6.00	4.33

①　大平均（㊥）を計算する（添字は条件の番号）。

㊥＝（4.33＋3.67＋5.33＋3.33＋6.00＋4.33）/6＝4.50

②　要因 A の 2 水準によって生じる分散を計算する。

要因 B の水準をつぶして、A①水準と A②水準の平均を求め、それと大平均とのズレを計算する。そして、データの個数分だけ倍にする。なお、A①水準と A②水準のデータ数は要因 B の 3 水準をつぶしているので

3N＝3×3＝9 である。

要因 A の分散＝（（4.33＋3.67＋5.33）/3－4.50）^{2}×9＋（（3.33＋6.00＋4.33）/3－4.50）^{2}×9＝0.06

③　要因 B の 3 水準によって生じる分散を計算する。

上と同様に、要因 A の 2 水準をつぶして、要因 B の各水準の平均を求め、大平均とのズレを計算する。これをデータの個数分だけ倍にする（要因 B の各水準のデータ数は 2N＝2×3＝6 である）。

要因 B の分散＝（（4.33＋3.33）/2－4.50）^{2}×6＋（（3.67＋6.00）/2－4.50）^{2}×6＋（（5.33＋4.33）/2－4.50）^{2}×6＝4.02

④　要因 A と要因 B の交互作用によって生じる分散を計算する。まず、全 6 条件の平均によって生じる分散を計算する。すなわち、各条件の平均と大平均とのズレを計算し、データの個数分だけ倍にして、足し上げる。

全 6 条件の分散＝（4.33－4.50）^{2}×3＋（3.67－4.50）^{2}×3＋（5.33－4.50）^{2}×3＋・・・＝15.16

次に、この「全 6 条件の分散」（15.16）から、要因 A と要因 B の分散を引く。残りが、要因 A・要因 B 単独では説明できない分散すなわち両者の相乗効果（交互作用）によって生じた分散となる。

交互作用＝全 6 条件の分散－要因 A の分散－要因 B の分散＝15.16－0.06－4.02＝11.08

⑤　最後に偶然誤差を計算する。

全条件の SD の二乗を合計し、データの個数をかける。

偶然誤差＝（SD_{11}^2＋SD_{12}^2＋SD_{13}^2＋SD_{21}^2＋SD_{22}^2＋SD_{23}^2）・N

＝（0.47^2＋0.47^2＋0.94^2＋0.47^2＋0.82^2＋0.47^2）×3＝7.32

ただし、各条件の N が等しくない場合、偶然誤差の計算には、便宜上の「等しい N」を用いず、実際のデータの個数を用いて下式による（添字は条件の番号）。

偶然誤差＝SD_{11}^2・N_{11}＋SD_{12}^2・N_{12}＋SD_{13}^2・N_{13}＋・・・＋SD_{23}^2・N_{23}

⑥　以上の各分散を分散分析表に記入し、F 比を計算する。

表 44　分散分析表

要因	平方和	df（カッコ内は計算要領）	平均平方	F
要因 A	0.06	1（水準数－1）	0.06	＜1
要因 B	4.02	2（水準数－1）	2.01	3.29†
A×B	11.08	2（df_A×df_B）	5.54	9.08**
誤差	7.32	12（全体－df_A－df_B　－df_A×df_B）	0.61	

全体	22.48	17（全データ数－1）		

†p＜.10　　**p＜.01

「全体」の分散（22.48）は各分散値の合計である。直接に計算するなら、(データー大平均）^{2}、を全データについて足し上げる。ただし、便宜上の「等しいN」を用いた場合には一致しない。自由度（df）の計算は、「誤差」の df を最後に出すのがコツ。平均平方は、平方和÷df、であった。F 比の計算は、「誤差」の平均平方（0.61）を共通の分母とする。

⑦　F 比を検定し、検定結果を分散分析表に書き込む。

F 比を F 分布表に対照するには 2 つの自由度が必要であった。その際、分母の自由度は共通に「誤差」の df になるが、分子の自由度はそれぞれ異なる。検定結果だけを列挙する。

要因A：F<1　（df①＝1、df②＝12）⇒　F<1 なので検定放棄

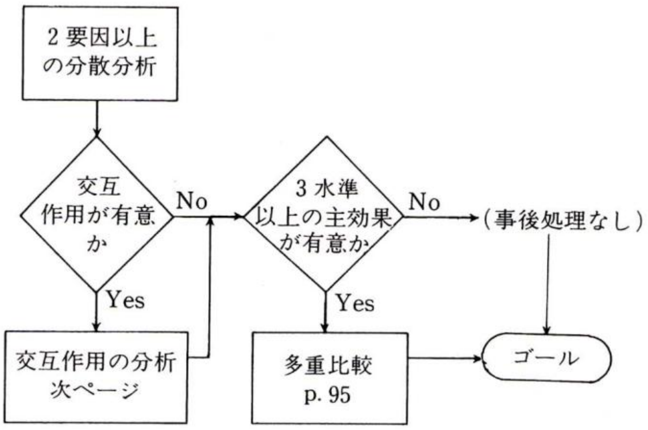
要因B：F＝3.29（df①＝2、df②＝12）⇒　有意傾向

交互作用：F＝9.08（df①＝2、df②＝12）⇒　1%水準で有意

さて、主効果（要因Aと要因B）および交互作用がどれも有意でなかった場合は、ここで終わる。一つでも有意（有意傾向）であった場合は、以下の事後処理に進む。

事後処理のコース

2 要因以上の分散分析の事後処理には、「交互作用の分析」と、前述した「多重比較」がある。処理の順序としては交互作用優先である。



<u>**交互作用の分析**</u>

定義

交互作用（interaction）とは、2 要因以上の実験計画において、一方の要因が他方の要因に及ぼす影響の「大きさ」または「方向」が一様でないことをいう。以下、2×2 の実験計画を例にとって、解説する。

①　一方の要因が他方の要因に及ぼす影響の大きさが一様でない例

最初の例は、単語の記憶実験である。

実験計画は、2×2 の被験者間計画。第一の要因は単語の有意味度（高・低）、第二の要因は単語の提示方法（視覚提示・聴覚提示）である。入手するデータは、提示した単語の再生率とする。下図 4 は、各条件の再生率の平均を示したものである。それによれば、単語の有意味度がそれぞれ視覚提示・聴覚提示に及ぼしている影響の大きさは各水準によって異なり、一様でない。

②　一方の要因が他方の要因に及ぼす影響の方向が一様でない例

さて、次は、教育心理学においてはよく知られている適性処遇交互作用（aptitude-treatment interaction：ATI）の例である。実験計画は、2×2 の被験者間計画、第一の要因は生徒の適性として対人不安の程度（高不安、低不安）、第二の要因は教育的処遇として指導方法（一斉指導・グループ指導）である。入手するデータは、指導後のテストにおける各生徒の得点とした。

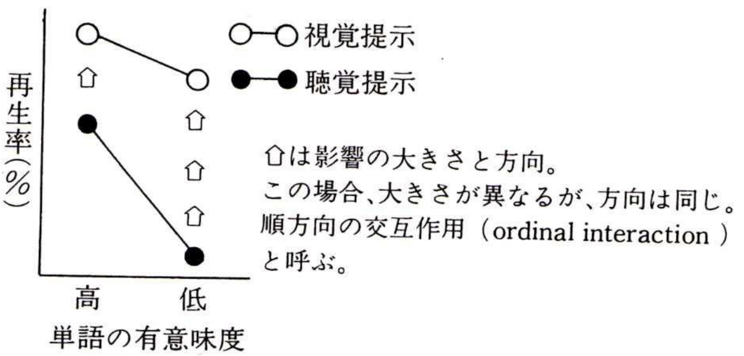


図4　単語の有意味度と提示方法の交互作用

下図 5 は、条件別に得点の平均を示したものである。この場合、対人不安の程度が指導方法の水準に及ぼす影響は、方向が異なっているとみることができる。

以上の例でわかるように、交互作用がある場合、各条件の平均を線グラフで図示すると線どうしが平行にならない。このことを「平均のプロフィールが交差する」と表現する（順方向の交互作用の場合は実際には交差しないが線を延長すれば交差する）。
そのように、交互作用をチェックする簡単な方法は、平均の線グラフを描いてみることである。

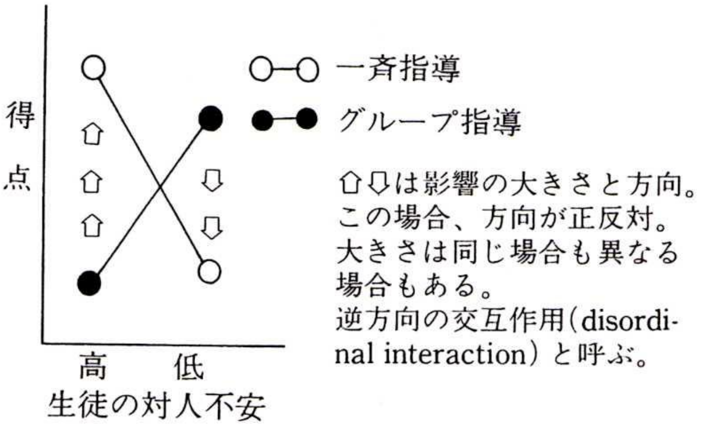


図5　生徒の対人不安と指導方法との交互作用

<u>交互作用の分析手順</u>

基本手順は、次のとおりである。
①平均の線グラフを描く。
②各要因の水準別に分散を再計算する。
③交互作用の分析結果を分散分析表に追加する。

以下、前例の2×3の交互作用を取り上げて解説する。

①　平均の線グラフを描く。

交互作用が有意であったら、そのパターンを確認するため、すぐに平均の線グラフを描いてみる。

表45　各条件の平均

	B①	B②	B③
A①	4.33	3.67	5.33
A②	3.33	6.00	4.33

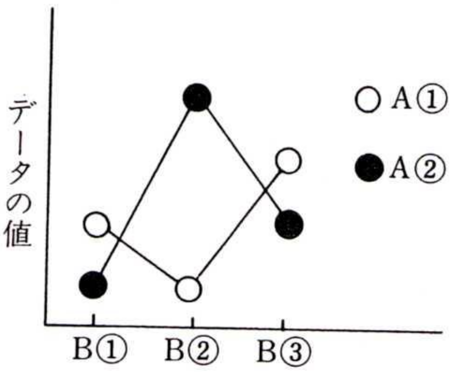


図6　平均のプロフィール

図6をみると、グラフは複雑に交差しており、部分的に逆方向の交互作用が認められる。このように交互作用がある場合は、各要因の単独効果（主効果）を取り上げても意味がない。なぜなら、その要因の効果の大きさと方向には一般性がないからである。そこで、データの全分散を要因単位で分けるのではなく、個々の水準別に分けることにする。

②　各要因の分散を水準別に再計算する。

最初に、要因Aの分散を要因Bの各水準別に計算する。まず、「B①水準における要因A」の分散を計算する。

B①水準の要因Aの分散＝$(X_{11}^2+X_{21}^2-(X_{11}+X_{21})^2/\text{平均の個数})\cdot N$

$=(4.33^2+3.33^2-(4.33+3.33)^2/2)\times 3=1.50$

同様にして「B②水準における要因A」の分散は8.14、「B③水準における要因A」の分散は1.50となる。以上は、簡易式であるが、定義式では「水準の平均」を求めて、それと個々の平均とのズレを計算して、合計する。

B①水準の要因Aの分散＝$(X_{11}-(X_{11}+X_{21})/2)^2\cdot N+(X_{21}-(X_{11}+X_{21})/2)^2\cdot N$

さて、次に要因Bの分散を要因Aの水準別に再計算する。まず、「A①水準における要因B」の分散は、下式のように計算する。

A①水準の要因Bの分散＝$(X_{11}^2+X_{12}^2+X_{13}^2-(X_{11}+X_{12}+X_{13})^2/\text{平均の個数})\cdot N$

$=(4.33^2+3.67^2+5.33^2-(4.33+3.67+5.33)^2/3)\times 3=4.19$

同様にして、A②水準における要因Bの分散は、10.92となる。

③　各要因の水準別の分散を分散分析表に追加する。上で求めた水準別の分散を、前出の分散分析表に以下のように書き加える。

表 46　交互作用の分析結果を書き加えた分散分析表

要因	SS	df	MS(SS/df)	F
要因Ａ	0.06	1	0.06	＜1
Ｂ①水準：	1.50	1	1.50	2.46
Ｂ②水準：	8.14	1	8.14	13.34**
Ｂ③水準：	1.50	1	1.50	2.46
要因Ｂ	4.02	2	2.01	3.30†
Ａ①水準：	4.19	2	2.10	3.44†
Ａ②水準：	10.92	2	5.46	8.95**
Ａ×Ｂ	11.08	2	5.54	9.08**
誤差	7.32	12	0.61	
全体	22.48	17		†p＜.10　**p＜.01

上表の分散値（SS）については、以下のような関係が成立する。

主効果＋交互作用＝水準別分散の合計

0.06＋11.08＝1.50＋8.14＋1.50

4.02＋11.08＝4.19＋10.92（←まるめによって等しくならない）。

なお、水準別の分散の自由度は、各要因の df と同じである。またＦ比を計算するときの分母は、すべて「誤差」の 0.61 を用いる。

さて、検定結果からわかるように、要因Ａの主効果（0.06）が有意でなくても、水準別に分けてみると、Ｂ②水準に限っては要因Ａの効果（8.14）は有意である（主効果に対して単純主効果と呼ぶ）。このように、交互作用がある場合は主効果の一般的意味がなくなる。

④　平均の多重比較をおこなう。

多重比較の対象は、分散分析または交互作用の分析において有意であった主効果または単純主効果のうち、水準数３以上（自由度２以上）の効果である。

本例の場合、交互作用が有意であったので主効果の多重比較は行わず、交互作用の分析によって検出した単純主効果の多重比較を行うことになる。

表 47　各条件の平均とタテの平均

	Ｂ①	Ｂ②	Ｂ③
Ａ①	4.33	3.67	5.33
Ａ②	3.33	6.00	4.33
タテ平均	3.83	4.84	4.83

要因Ｂは３水準なので多重比較する。Ａ①水準（傾向）はこの３つのを比較する。Ａ②水準（有意）はこの３つを比較する。主効果（傾向）はこの３つを比較するが、今回、交互作用有意なので比較しない。

上表からわかるように、多重比較は３個以上の平均が並ぶ場合に適用される。要因Ａのように水準数２つの場合は、主効果や単純効果の有意性はそのまま２平均の有意差を意味するので、多重比較の必要はない（2×2 の実験計画に多重比較はない）。そこで、上例の要因Ｂ（３水準）の単純主効果について、ＬＳＤ法による多重比較を行うことにする。「誤差」の自由度（df＝12）に相当する 5%点のｔ値をｔ分布表にて探し（ｔ＝2.179）、また、分散分析表から MSe＝0.61 を読み取って、下式に代入する。

ＬＳＤ＝ｔ値・√（2・MSe/1 平均あたりのデータ数）＝2.179×√（2×0.61/3）＝1.39

ただし、各条件のデータの個数が等しくない場合、上式の「1 平均あたりのデータ数」には便宜上の「等しいＮ」を用いる。さて、ＬＳＤ＝1.39 を最小有意差の幅として、各２平均の差がそれを越えたら有意（5%水準）、越えなかったら有意でないと判定する。検定の結果は次のとおりである。なお、＜、＞は有意差の方向、≒は有意差でないことを表す。

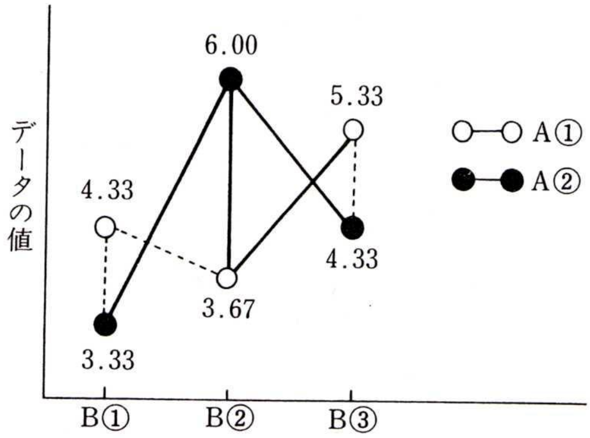
Ａ①水準の要因Ｂの多重比較：4.33≒3.673＜5.33（ただし 4.33≒5.33）

Ａ②水準の要因Ｂの多重比較：3.33＜6.00＞4.33（ただし 3.33≒4.33）

以上の多重比較の結果と前述の交互作用の分析結果から、本例の平均のプロフィールは、以下のように確定し

た（実線は有意差、破線は有意差でない）。

右図における要因Aの２水準の差（タテ線）は、交互作用の分析結果をそのまま用いたものである。これに対して、要因Bの３水準は、A①水準にて有意傾向、A②水準にて 1%有意であったので、さらに、多重比較の事後処理を行った結果である。分散分析の事後処理は、ここまで行わなければならない。



結果の論文記載例

「図６は、各条件の平均を図示したものである。分散分析の結果、交互作用が有意であった（F（2、12）＝9.08、ｐ＜.01）。そこで、各水準ごとに単純主効果を分析した結果、表 46 の分散分析表に示すとおりとなった。すなわち、要因AはB②水準において有意であるが、B①水準とB③水準においては有意でない。また、要因BはA①水準において有意傾向、A②水準において 1%水準で有意であった。

ＬＳＤ法による多重比較の結果、A①水準ではB③の平均がB②の平均よりも有意に大きく、A②水準ではB②の平均が他の２平均よりも有意に大きかった（MSe＝0.61、5%水準）。」

上の文章中、データ表示を図にしている点、、および、主効果についての言及がまったくない点に注意する。いずれも、交互作用が有意であったための処置である。

備考：交互作用の分析を多重比較で代用することについて

上述したような交互作用の分析を行わず、その代わりに、多重比較だけによって平均のプロフィールを確定することもできる。実際、本例のＬＳＤ（1.39）を直接に要因Ａ・要因Ｂの各２平均間すべてに適用すれば、先ほどと同じ確定図が得られる。この便宜的方法は多用されるが、本格的でない。例えば、交互作用の分析において分散が有意でなかった単純主効果に対しても、多重比較がどこかに有意差を拾ってしまう場合がある。これは誤りである。なぜなら、多重比較は、あくまで分散分析または交互作用の分析の事後処理であって、独立の有意性を主張できないからである。

b. 被験者内計画（ＳＡＢタイプ）

適用

２ 要因被験者内計画の分散分析（ＳＡＢタイプ）は、やや複雑である。データの全分散は、以下のように、7 つの部分に分けられる。

データの全分散＝要因Ａの主効果＋要因Ｂの主効果＋交互作用＋個人差＋要因Ａに対する誤差＋要因Ｂに対する誤差＋交互作用に対する誤差

上式からわかるように、前述した被験者間計画の分散分析（ABS タイプ）と比べて、偶然誤差の項目が 4 つに増えている。それ以外（2 つの主効果と交互作用）は同じである。

計算例

比較のため前例の被験者間計画のデータを、そのまま被験者内計画にみたてて用いる。すなわち、「2×3 の被験者内計画」であり、全条件を同じ被験者に割り当てる。データの個数は１条件につき３個とするので、被験者数も３人でよい。下表は、データ表示である。

表 48　データ表示（N＝3）

	A①			A②		
	B①	B②	B③	B①	B②	B③
X	4.33	3.67	5.33	3.33	6.00	4.33
SD	0.47	0.47	0.94	0.47	0.82	0.47

以下の計算は、第５ステップの偶然誤差の計算まで、前述した被験者間計画の分散分析（ABS タイプ）と全く同一手順・同一結果である。したがって、簡略化して示す。

① 大平均（㊥）を計算する。

㊥＝（4.33＋3.67＋5.33＋3.33＋6.00＋4.33）/6＝4.50

② 要因Aの２水準によって生じる分散を計算する。

要因Aの分散＝（(4.33＋3.67＋5.33）/3－4.50）^{2}・3N＋（(3.33＋6.00＋4.33）/3－4.50）^{2}・3N＝0.06

③ 要因Bの３水準によって生じる分散の計算をする。

要因Bの分散＝（(4.33＋3.33）/2－4.50）^{2}・2N＋（(3.67＋6.00）/2－4.50）^{2}・2N＋・・・＝4.02

④ 交互作用を計算する。

全６条件の分散＝（(4.33－4.50）^{2}＋（3.67－4.50）^{2}＋・・・）・N＝15.16

交互作用＝全６条件の分散－要因Aの分散－要因Bの分散＝11.08

⑤ 偶然誤差を計算する。

偶然誤差＝（0.47^2＋0.47^2＋0.94^2＋0.47^2＋0.82^2＋0.47^2）・N＝7.32

ここから、被験者内計画（SAB タイプ）特有の計算手順になる。計算の必要上、データ・リストを掲載する。被験者内計画であるので、各被験者はすべての条件に割り当てられている。

表 49 データ・リスト

Sub.	A①			A②			合計
	B①	B②	B③	B①	B②	B③	
田中	4	3	6	3	5	5	26
山際	4	4	6	4	6	4	28
遠藤	5	4	4	3	7	4	27

⑥ 個人差によって生じる分散を計算する。

個人差とは、今回たまたま集まった被験者個人と個人の差異である。個人差の計算には、被験者一人ひとりのデータの合計を用いる。なお、計算の便宜上、最初に「修正項」（K）を求めておく。

修正項（K）＝（全データの合計）^{2}/全データ数＝（26＋28＋27）^{2}/18＝364.50

個人差＝（田中の合計^2＋山際の合計^2＋遠藤の合計^2）/1 人あたりのデータ数－K

＝（26^2＋28^2＋27^2）/6－364.50＝0.33

⑦ 要因Aに対する誤差を計算する。

上の「個人差」は被験者と被験者の間の差異であるが（被験者間の偶然誤差）、これ以降は、被験者の個人内のユレを計算する（被験者内の偶然誤差）。まず、被験者個人が要因Aの２水準に対して生じる誤差を計算する。これには、次のような補助表をつくる。

Sub.	A①水準の合計	A②水準の合計
田中	4＋3＋6＝13	3＋5＋5＝13
山際	4＋4＋6＝14	4＋6＋4＝14
遠藤	5＋4＋4＝13	3＋7＋4＝14

上表は、要因Bの 3 水準をつぶして「被験者×要因A」のマスをつくり、データの合計を求めたものである。この 3×2 表の 6 個の平均が生じる全分散を計算して、そこから個人差（0.33）と要因A（0.06）の説明分を引いた残りが、説明しきれない誤差になる。

要因Aに対する誤差＝（(合計^2＋合計^2＋合計^2＋・・・）/1 マスあたりのデータ数）－K－個人差－要因A

＝（13^2＋13^2＋14^2＋14^2＋13^2＋14^2）/3－364.50－0.33－0.06＝0.11

⑧ 要因Bに対する誤差を計算する。

今度は要因Aの２水準をつぶして、「被験者×要因B」の補助表をつくる。計算の最後に、個人差と要因B（4.02）の分散を引く。

Sub.	B①水準の合計	B②水準の合計	B③水準の合計
田中	4＋3＝7	3＋5＝8	6＋5＝11
山際	4＋4＝8	4＋6＝10	6＋4＝10
遠藤	5＋3＝8	4＋7＝11	4＋4＝8

要因Bに対する誤差＝（(合計^2＋合計^2＋合計^2＋・・・）/1 マスあたりのデータ数）－K－個人差－要因B

＝（7^2＋8^2＋11^2＋8^2＋10^2＋10^2＋8^2＋11^2＋8^2）/2－364.50－0.33－4.02＝4.65

⑨　交互作用に対する誤差を計算する。

今まで、偶然誤差（7.32）のなかみを分けて、被験者間の個人差（0.33）と要因Aに対する個人内の誤差（0.11）と要因Bに対する個人内の誤差（4.65）を計算してきた。あと残りは、要因Aと要因Bの交互作用に対して被験者が生じた個人内の誤差である。

交互作用に対する誤差＝偶然誤差－個人誤差－要因Aに対する誤差－要因Bに対する誤差

＝7.32－0.33－0.11－4.65＝2.23

⑩　以上の要因分散・誤差分散を分散分析表に記入する。

下表の見出しの順番に注意。一般に、実験要因の分散とそれに対応する誤差の分散とをセットにして書く。

表 50　分散分析表（SAB タイプ）

要因	SS	df（カッコ内は計算式）	MS(MS/df)	F
個人差（S）	0.33	2（人数－1）	0.17	
要因A	0.06	1（水準数－1）	0.06	1.00
S×A	0.11	2（df_S×df_A）	0.06	
要因B	4.02	2（水準数－1）	2.01	1.73
S×B	4.65	4（df_S×df_B）	1.16	
A×B	11.08	2（df_A×df_B）	5.54	9.89*
S×A×B	2.23	4（df_S×df_A×df_B）	0.56	
全体	22.48	17（全データ数－1）		* p ＜.05

上表の見出しにおいて、「S」のからんだ記号はすべて偶然誤差である。例えば、「S×A」は「被験者（S）が要因Aに対して生じた誤差」を表している。一般にそのように書く。分散分析表を上段の見出しから順番に説明する。個人差は被験者間の差異である。これも偶然誤差の一部分であるが、今回は被験者内計画であり、実験要因は被験者間に配置されていないので出番がない。以下、すべて被験者内の分散である。要因AはA①水準とA②水準の平均の差によって生じた分散である（主効果）。これに対する偶然誤差はS×A である。すなわち、被験者（S）が要因Aに対して生じた誤差である。前者を分子、後者を分母にとったF比（1.00）は、「要因A」の分散が「S×A」の偶然誤差の大きさとあまりちがわないことを示している（平均平方による比較）。これは、検定するまでもなく、有意でない。

同様に、要因Bの主効果に対してはS×Bが偶然誤差となる。この場合、F比（1.73）をみると、「要因B」の分散は「S×B」の偶然誤差より 1.73 倍だけ大きい。そこで、このF比の分子の自由度（df_B）と分母の自由度（$df_{S\times B}$）をとって、df①＝2、df②＝4 のF分布に対照してみると、F＝1.73 で有意でない。

最後に、A×B の交互作用に対しては S×A×B が偶然誤差となる。「S×A×B」とは、被験者（S）が交互作用（A×B）に対して生じた誤差を表している。F比を計算した結果、交互作用は、この偶然誤差よりも約 10 倍の大きさがあった（F＝9.83）。F分布に対照してみると、(df①＝2、df②＝4)、5%水準で有意である。このように、被験者内計画の分散分析（SAB タイプ）では、それぞれの実験要因に、それぞれの偶然誤差がある。この対応づけを誤らないようにしなければならない。

なお、「全体」の SS（データの全分散）は以上の分散値の合計である。確かめは、(データ－大平均）^{2}、の総和をとればよい。

⑪　交互作用の分析をおこなう。

交互作用が有意であった場合は、ABS タイプと同様、各要因の分散を水準別に計算しなおすことになる。異なる点は、この水準別の分散（単純主効果）を検定する偶然誤差もまた水準別に計算しなければならないということである。すなわち、B①水準の要因Aの単純主効果AatB①を検定するにはB①水準の偶然誤差 S×AatB①を用いる、というように水準を一致させなければならない。

表 51　A×B の交互作用の分析表

要因	SS	df	MS	F
AatB①	1.50	1	1.50	3.00
(S×AatB①	0.99	2	0.50)	

AatB② (S×AatB②	8.14 0.35	1 2	8.14 0.18)	45.22*
AatB③ (S×AatB③	1.50 0.98	1 2	1.50 0.49)	3.00
BatA① (S×BatA①	4.19 3.75	2 4	2.10 0.94)	2.23
BatA② (S×BatA②	10.92 3.12	2 4	5.46 0.78)	7.00*
				＊p＜.05

付表　B①水準表のデータ

	A①	A②	計
	B①	B①	
田中	4	3	7
山際	4	4	8
遠藤	5	3	8
			23

B①水準の平均は23/6＝3.83
例えばB①水準の偶然誤差（0.99）の計算は、上のように付表を作成して被験者全体の平均（3.83）を求め、下式による。

$0.99 = 0.47^2 \times 3 + 0.47^2 \times 3$　　←　「B①水準のSDとN（表48から）」
　－（7/2－3.83）^{2}×2－（8/2－3.83）^{2}×2－（8/2－3.83）^{2}×2　←「各被験者のデータ2個分の偶然誤差」

⑫　多重比較をおこなう。
　水準数3以上の有意な効果（すなわち自由度2以上の有意なF比）がある場合には、多重比較をおこなう。
　本例の場合、A②水準における要因Bの単純主効果（自由度2）が有意であった（F（2,4）＝7.00、p＜.05）。そこで、LSD法を用いて、下の3つの平均の差をそれぞれ検定する。

表52　各条件の平均

	B①	B②	B③
A①	4.33	3.67	5.33
A②	3.33	6.00	4.33

A①水準は有意でなかった。A②の3つが多重比較の対象。
　ここで、LSDの計算に用いるt値は、単純主効果BatA②の偶然誤差であるS×BatA②の自由度（df＝4）に相当するt値となる（t分布表よりdf＝4の5%点はt＝2.776）。同様に、MSeもS×BatA②の平均平方を用いる（MSe＝0.78）。
LSD＝t値・√（2・MSe/1平均あたりのデータ数）
　＝2.776×√（2×0.78/3）＝2.00
上表のA②水準において、LSD＝2.00を越える差はB①とB②の差である（｜3.33－6.00｜＞2.00）。それだけが有意である。

結果の論文掲載例

「図6は、各条件の平均を示したものである。分散分析の結果、交互作用が有意であった（F（2,4）＝9.89、p＜.05）。そこで、各要因の単純主効果を分析した結果、表51に示すとおりとなった。なお、A②水準における要因Bの単純主効果については、LSD法による多重比較の結果、B①とB②の平均の差のみが有意であった（MSe＝0.78、5%水準）。」

c. 混合計画（ASB タイプ）

適用

2 要因の混合計画の場合、一方の要因（A）は被験者間、他方の要因（B）は被験者内に配置されている。すなわち、要因Aの水準は異なる被験者に割り当て、要因Bの水準は同じ被験者に割り当てる。したがって、混合計画の分散分析（ABS タイプ）は、以下のモデルのように、データの全分散を被験者間と被験者内に分ける。

データの全分散＝要因Aの主効果＋被験者間の偶然誤差（個人差）　←　「被験者間に生じる分散」
　　　　　　　　＋要因Bの主効果＋交互作用＋被験者内の偶然誤差　←　「被験者内に生じる分散」

計算例

前例と同じデータを混合計画として用いる。すなわち「２×3 の混合計画、第一の要因は被験者間配置、第二の要因は被験者内配置」となる。データの個数は１条件３個であるので、被験者数は、被験者間要因の水準数をかけて、３個×２水準＝６人となる。以下に、データ・リストを掲載し、データ表示を行う。

表５３　データ・リスト　（各被験者の合計は後の計算に用いる）

Sub.	A①			合計	Sub.	A②			合計
	B①	B②	B③			B①	B②	B③	
田中	4	3	6	13	小野	3	5	5	13
山際	4	4	6	14	高橋	4	6	4	14
遠藤	5	4	4	13	鈴木	3	7	4	14

表５４　データ表示

	A①			A②		
	B①	B②	B③	B①	B②	B③
N	3	3	3	3	3	3
X	4.33	3.67	5.33	3.33	6.00	4.33
SD	0.47	0.47	0.94	0.47	0.82	0.47

○　Nが異なる場合は便宜上の「等しいN」を計算する。

被験者間要因（A）の各水準のNが等しくない場合、調和平均による便宜上の「等しいN」を計算する。

$$N＝条件数/（1/N_{11}＋1/N_{12}＋1/N_{13}＋1/N_{21}＋1/N_{22}＋1/N_{23}）$$

さて、以下の主効果と交互作用の計算は、ＡＢＳタイプおよびＳＡＢタイプの分散分析とまったく同一である。したがって、③と⑥を除いては計算式のみを示す。

① 大平均を計算する。

大平均＝（4.33＋3.67＋5.33＋3.33＋6.00＋4.33）/6＝4.50

② 要因Aの主効果を計算する。

要因Aの主効果

＝（(4.33＋3.67＋5.33）/３－大平均）^{2}×3N＋（(3.33＋6.00＋4.33）/3－大平均）^{2}×3N＝0.06

③ 被験者間の偶然誤差（個人差）を計算する。

各被験者のデータの合計（表 53）から各被験者の平均を求め、大平均とのズレを合計する。そして要因Aの分散（0.06）を引く。

個人差＝（田中の合計/データ数－大平均）^{2}×データ数＋・・・＋（鈴木の合計/データ数－大平均）^{2}×データ数－要因A

＝（13/3－4.5）^{2}×3＋・・・＋（14/3－4.5）^{2}×3－0.06＝0.44

④ 要因Bの主効果を計算する。

要因Ｂの主効果＝（(4.33＋3.33）/2－大平均）²・2N＋・・・＝4.02

⑤　交互作用を計算する。

全６条件の分散＝（(4.33－大平均）²＋（3.67－大平均）²＋（5.33－大平均）²＋・・・）・N＝15.16

交互作用＝全６条件の分散－要因Ａの主効果－要因Ｂの主効果＝11.08

⑥　被験者内の偶然誤差を計算する。

すでに、③において被験者間の個人差（0.44）を求めているので、全体の偶然誤差を計算して、そこから0.44を引けばよい。

全体の偶然誤差＝（ＳD²＋ＳD²＋ＳD²＋・・・）・N＝7.32

被験者内の偶然誤差＝全体の偶然誤差－個人差＝6.88

ただし、各条件のNが等しくない場合は、下式による。

被験者内の偶然誤差＝$SD_{11}^2 \cdot N_{11} + SD_{12}^2 \cdot N_{12}$＋・・・－個人差

⑦　以上の分散値を分散分析表に記入する。

表５５　分散分析表（ＡＳＢタイプ）

要因	ＳＳ	df（ｶｯｺ内は計算式）	MS（SS/df）	F
要因Ａ	0.06	1（水準数－1）	0.06	0.55
個人差（Ｓ）	0.44	4（人数－要因Ａの水準数）	0.11	
要因Ｂ	4.02	2（水準数－1）	2.01	2.34
Ａ×Ｂ	11.08	2（$df_A \times df_B$）	5.54	6.44*
Ｓ×Ｂ	6.88	8（$df_S \times df_B$）	0.86	
全体	22.48	17（全データ数－1）		＊ｐ＜.05

表中の上段２つの見出しが被験者間の分散、下段３つの見出しが被験者内の分散である。一般に、そのように分けて書く。ここで、見出しの「個人差」は被験者間の偶然誤差である。

また、「Ｓ×Ｂ」は被験者内の偶然誤差であり、被験者（Ｓ）が要因Ｂに対して生じる個人内のユレである。２要因混合計画では、個人内の誤差はそれしか生じない。なぜなら、各被験者のデータは要因Ｂの水準とだけ一緒に変化するからである。要因Ａの水準とは独立である。このため、「Ｓ×Ａ」や「Ｓ×Ａ×Ｂ」のような要因Ａがらみの個人内の誤差は生じえない。

Ｆ比の計算にあたっては、被験者間の実験要因（要因Ａ）に対しては被験者間の偶然誤差（個人差Ｓ）を分母にする。同様に、被験者内の実験要因（要因Ｂと交互作用）に対しては被験者内の偶然誤差（Ｓ×Ｂ）を分母にする。

したがって、Ｆ分布表に対照するときも、要因ＡのＦ比（0.55）の自由度はdf①＝１、df②＝４となるが、要因ＢのＦ比（2.34）と交互作用のＦ比（6.44）の自由度はdf①＝２、df②＝８となる。

⑧　交互作用の分析を行う。

交互作用が有意であった場合の分析は、ＡＢＳタイプの場合と同様である。ただし、本例のＡＳＢタイプ（混合計画）では、要因Ａの単純主効果AatＢ○については水準別の個人差ＳatＢ○を新たに計算するが、要因Ｂの単純主効果ＢatＡ○については、Ｓ×Ｂをそのまま用いる。

表５６　Ａ×Ｂの交互作用の分析表

要因	ＳＳ	df	ＭＳ	Ｆ
AatＢ①：	1.50	1	1.50	4.50
（ＳatＢ①：	1.33	4	0.33）	
AatＢ②：	8.14	1	8.14	12.15*
（ＳatＢ②：	2.68	4	0.67）	
AatＢ③：	1.50	1	1.50	1.81
（ＳatＢ③：	3.31	4	0.83）	
ＢatＡ①：	4.19	2	2.10	2.44
ＢatＡ②：	10.92	2	5.46	6.35*
（Ｓ×Ｂ：	6.88	8	0.86）	

＊ｐ＜.05

例えば、Ｂ①水準の個人差（1.33）の計算は、Ｂ①水準のＳＤを分散値に変換し（二乗する）、それを被験者の人数倍して、合計すればよい。なぜならＳＤ（標準偏差）は、その条件内での個々人のデータの標準的バラツキを表しているからである。

したがって、データ表示の表５４から、Ｂ①水準のＳＤ（0.47 と 0.47）を取り出して下式のように計算する。

Ｂ①水準の個人差（ＳatＢ①）＝0.47^2×3＋0.47^2×3＝1.33

⑨　多重比較を行う。

先の分散分析表において、水準数３以上の有意な効果すなわち自由度２以上の有意なＦ比がある（Ａ②水準における要因Ｂの単純主効果：Ｆ（2,8）＝6.35、ｐ＜.05）。そこで、この単純主効果Ｂの平均に対して、ＬＳＤ法による多重比較を適用する。

ＬＳＤを計算するため、まず、ｔ値を用意する。この場合の偶然誤差は、「Ｓ×Ｂ」であるので、その自由度（df＝８）に相当する 5%点の値をｔ分布表から求める（ｔ＝2.306）。また、この「Ｓ×Ｂ」の平均平方を読み取っておく（MSe＝0.86）。

以上のｔ値と MSe を下式に代入し、ＬＳＤを計算する。なお、本例の場合は「1 平均あたりのデータ数」はＮである。ここまで便宜上の「等しいＮ」を用いてきた場合はそれを用いる。

ＬＳＤ＝ｔ値・√￣（２・MSe/1 平均あたりのデータ数）＝2.306×√￣（2×0.86/3）＝1.75

Ａ②水準における要因Ｂの３つの平均（3.33,6.00,4.33）のうち、ＬＳＤ＝1.75 を越える差は、3.33 と 6.00 の差だけである。すなわち、Ａ②水準では、Ｂ①とＢ②の平均の間に有意差がある。

結果の論文記載例

「図６は、各条件の平均を図示したものである。分散分析の結果、交互作用が有意であった（Ｆ（2,8）＝6.44、ｐ＜.05）。そこで、各要因の単純主効果を分析した結果、表５６に示すとおりとなった。なお、Ａ②水準におけえる要因Ｂの単純主効果については、ＬＳＤ法による多重比較の結果、Ｂ①とＢ②の平均の差のみが有意であった（MSe＝0.86、5%水準）。」

2 要因分散分析の各タイプの比較

下記にこれまで解説してきた 2 要因の分散分析表を列挙した。これを相互に比較してみる。

まず、各タイプの等しい点は、「全体」の分散である（22.48）。これは同一のデータを異なる実験計画に見立てて用いたので当然である。同様に、「要因Ａ」、「要因Ｂ」、「Ａ×Ｂ」の分散も等しい。このように、主効果と交互作用の分散は共通である。

タイプによって異なるのは、偶然誤差の分け方である。それによって、主効果・交互作用のＦ比の分母および自由度が異なり、検定結果が違ってくる。ただし、どのタイプも偶然誤差の合計は、等しく、みな 7.32 である。

備考：Ｎが等しくない場合の分散分析について

被験者間計画と混合計画の分散分析では、各条件のデータの個数が等しくない場合がある。この場合、調和平均による便宜上の「等しいＮ」を計算する。

もし、実際の等しくない個数をそのまま用いると、データの個数のアンバランスが交互作用の分散に出てしまう（ＳＳを大きくする。）。便宜的に等しくしたＮを用いるのは、このためである。

ただし、調和平均による「等しいＮ」は簡易の対処であり、あまりにデータの個数のアンバランス（くいちがい）が大きい場合は不適当である。その場合、データの個数のアンバランスに応じた重みづけを各要因の分散に施すことになるが、手計算のレベルではない。大型計算機における専用プログラムを用いる。ＳＡＳのＧＬＭでは、ＴＹＰＥⅢＳＳがこの重みづけを施した分散にあたる。

表 44　被験者間計画の分散分析表（ＡＢＳタイプ）

要因	平方和	df（カッコ内は計算要領）	平均平方	Ｆ
要因Ａ	0.06	1（水準数−1）	0.06	＜1
要因Ｂ	4.02	2（水準数−1）	2.01	3.29†
Ａ×Ｂ	11.08	2（df_A×df_B）	5.54	9.08＊＊
誤差	7.32	12（全体−df_A−df_B　−df_A×df_B）	0.61	
全体	22.48	17（全データ数−1）		

表５０　被験者内計画の分散分析表（ＳＡＢタイプ）

要因	SS	df（カッコ内は計算式）	MS(MS/df)	F
個人差（S）	0.33	2（人数−1）	0.17	
要因A	0.06	1（水準数−1）	0.06	1.00
S×A	0.11	2（$df_S \times df_A$）	0.06	
要因B	4.02	2（水準数−1）	2.01	1.73
S×B	4.65	4（$df_S \times df_B$）	1.16	
A×B	11.08	2（$df_A \times df_B$）	5.54	9.89*
S×A×B	2.23	4（$df_S \times df_A \times df_B$）	0.56	
全体	22.48	17（全データ数−1）		* $p < .05$

表５５　混合計画の分散分析表（ＡＳＢタイプ）

要因	ＳＳ	df（ｶｯｺ内は計算式）	MS（SS/df）	F
要因A	0.06	1（水準数−1）	0.06	0.55
個人差（S）	0.44	4（人数−要因Aの水準数）	0.11	
要因B	4.02	2（水準数−1）	2.01	2.34
A×B	11.08	2（$df_A \times df_B$）	5.54	6.44*
S×B	6.88	8（$df_S \times df_B$）	0.86	
全体	22.48	17（全データ数−1）		* $p < .05$

４－４　三要因の分散分析

a.被験者間計画の事例：ＡＢＣＳタイプ

論文名「音声化と形態化が単語と数字の再認に及ぼす効果」

方法

実験計画：2×2×4 の被験者間計画。第一の要因は記憶の補助動作（音声化・形態化）。第二の要因は材料（単語・数字）、第三の要因は年齢段階（6 歳・10 歳・14 歳・18 歳）。

被験者：小学一年生・四年生、中学二年生、大学生、計 177 人。

手続き：実験は個別実験である。被験者に３文字の単語または３ケタの数字を提示して、音声化または形態化（指でひざに書く）によって記憶させた。緩衝課題によるインタバルをおき、直後に再認試行を課して各被験者の正再認数を記録した。

結果

表５７は、条件別に正再認数の平均と標準偏差を示したものである。

表 57　音声化・形態化による単語・数字の正再認数（max.20）

		単語				数字			
		6 歳	10 歳	14 歳	18 歳	6 歳	10 歳	14 歳	18 歳
音声化	N	8	8	12	15	10	8	10	15
	X	5.2	8.9	12.5	15.2	6.3	10.2	13.4	16.0
	ＳD	3.8	4.7	4.4	4.6	3.2	5.7	5.0	5.5
形態化	N	10	11	10	15	9	8	13	15
	X	5.6	9.2	11.4	15.5	5.0	13.7	15.8	16.6
	ＳD	3.3	5.1	4.8	5.2	3.0	5.7	4.5	5.8

分散分析の結果、単語・数字の主効果（$F(3,161) = 24.28$、$p < .01$）が有意であった。しかし、それ以外の主

効果と交互作用は有意でなかった。

解説

本例は、3 要因の被験者間計画（ＡＢＣＳタイプ）である。すなわち、各条件を異なる被験者に割り当てている。このため、各条件のデータの個数が不ぞろいになっている。

結果としては、主効果のみが有意なケースである。分散分析の計算手順とともに、多要因の場合の主効果に対する多重比較の要領を学ぶ。

○　データの個数が等しくないので調和平均のＮを計算する。

調和平均の計算式は以下のとおりである（添え字は条件の番号）。

N＝全条件数/（$1/N_{111}+1/N_{112}+1/N_{113}+1/N_{114}+1/N_{121}+1/N_{122}+\cdots$）

＝16/（1/8＋1/8＋1/12＋1/15＋1/10＋1/8＋1/10＋1/15＋・・・）

＝10.46

以下、どの条件も、データの個数はＮ＝10.46 であるとみなす。

①　大平均を計算する（添え字は条件の番号）。

大平均＝（$X_{111}+X_{112}+X_{113}+X_{114}+X_{121}+X_{122}+X_{123}+\cdots$）/平均の個数

＝（5.2＋8.9＋12.5＋15.2＋6.3＋10.2＋13.4＋・・・）/16

＝11.28

②　「音声化・形態化」の主効果の分散を計算する。

「単語・数字」の 2 水準と「年齢」の 4 水準を無視して、「音声化」の平均と「形態化」の平均を計算する。そして、それぞれの平均と大平均との分散を計算して、データ数分（８Ｎ）だけ倍にする。

音声化・形態化＝（(5.2＋8.9＋12.5＋15.2＋6.3＋・・・＋16.0）/平均の個数（8）－大平均）^{2}・8N

＋（(5.6＋9.2＋11.4＋15.5＋5.0＋・・・＋16.6）/平均の個数（8）－大平均）^{2}・8N

＝17.00

③　「単語・数字」の主効果の分散を計算する。

今度は、「音声化・形態化」の 2 水準と「年齢」の 4 水準を無視して、「単語」の平均と「数字」の平均を再計算する。

単語・数字＝（(5.2＋8.9＋12.5＋15.2＋5.6＋・・・＋15.5）/平均の個数（8）－大平均）^{2}・8N

＋（(6.3＋10.2＋13.4＋16.0＋5.0＋・・・＋16.6）/平均の個数（8）－大平均）^{2}・8N

＝119.15

④　同様に、「年齢」の主効果の分散を計算する。

各年齢のデータの個数は、それぞれ 4Nとなることに注意せよ。

年齢＝（(5.2＋5.6＋6.3＋5.0）/平均の個数（4）－大平均）^{2}・4N＋・・以下同様に 10 歳、14 歳、18 歳の各分散を計算して足していく。

＝2442.01

⑤　「音声化・形態化」×「単語・数字」の交互作用を計算する。

「年齢」の 4 水準をつぶして、「音声化・形態化」×「単語・数字」の補助表をつくる。表中の各マスは平均が 4 個ずつ入る。

	単語	数字
音声化	5.2,8.9,12.5,15.2	6.3,10.2,13.4,16.0
形態化	5.6,9.2,11.4,15.5	5.0,13.7,15.8,16.6

上表の各マスの平均を再計算して、各マスの平均が生じる全分散を計算する。そこから「音声化・形態化」と「単語・数字」の主効果分を引いた残りが、単独要因では説明できない分散すなわち交互作用の分散となる。1 マスあたりのデータの個数は４Ｎ。

音声化形態化×単語数字＝（(5.2＋8.9＋12.5＋15.2）/平均の個数（4）－大平均）^{2}・4N

＋（(6.3＋10.2＋13.4＋16.0）/平均の個数（4）－大平均）^{2}・4N

＋（(5.6＋9.2＋11.4＋15.5）/平均の個数（4）－大平均）^{2}・4N

＋（(5.0＋13.7＋15.8＋16.6）/平均の個数（4）－大平均）^{2}・4N

－音声化形態化の主効果－単語数字の主効果

＝18.36

⑥　「単語・数字」×「年齢」の交互作用を計算する。

上と同様に今度は「音声化・形態化」の 2 水準をつぶして、「単語・数字」×「年齢」の補助表をつくる。表中の各マスには、平均が 2 個ずつ、データが２N個ずつ入る。計算では、各マスの新たな平均を再計算し、大平均との分散を求め、データの個数（２N）倍する。そこから主効果分を引く。

	6 歳	10 歳	14 歳	18 歳
単語	5.2,5.6	8.9,9.2	12.5,11.4	15.2,15.5
数字	6.3,5.0	10.2,13.7	13.4,15.8	16.0,16.6

単語数字×年齢＝（(5.2＋5.6）/2－大平均）²・2N＋（(8.9＋9.2）/2－大平均）²・2N

＋・・・以下、同様にして残りの 6 マスの各分散を計算して足してゆく。

－単語数字の主効果（119.15）－年齢の主効果（2442.01）

＝52.37

⑦　「音声化・形態化」×「年齢」の交互作用を計算する。

「単語・数字」の 2 水準をつぶし、「音声化・形態化」×「年齢」の補助表をつくる。各マスには平均 2 個、データ２N個が入る。

	6 歳	10 歳	14 歳	18 歳
音声化	5.2,6.3	8.9,10.2	12.5,13.4	15.2,16.0
形態化	5.6,5.0	9.2,13.7	11.4,15.8	15.5,16.6

音声化形態化×年齢＝（(5.2＋6.3）/2－大平均）²・2N＋（(8.9＋10.2）/2－大平均）²・2N

＋・・・以下、同様にして残りの 6 マスの各分散を計算して足してゆく。

－音声化形態化の主効果（17.00）－年齢の主効果（2442.01）

＝29.41

⑧　3 重交互作用の分散を計算する。

3 重交互作用とは、「音声化・形態化」×「単語・数字」×「年齢」の 3 要因による交互作用のことである。別名、二次の交互作用と呼ぶ。したがって、これまでの 2 要因の交互作用は、一次の交互作用である（一次・二次は×の数と思えばよい）。

計算は最初に全 16 条件の 16 平均によって生じる総分散を計算して、そこから、以上で計算したすべての主効果・交互作用を引く。

全 16 条件の分散＝（5.2－大平均）²・N＋（8.9－大平均）²・N＋（12.5－大平均）²・N

＋・・・以下、同様にして残り 13 平均の各分散を計算して足してゆく。

＝2726.55

3 重交互作用＝全 16 条件の分散－音声化形態化の主効果－単語数字の主効果－年齢の主効果－音声化形態化と単語数字の交互作用－単語数字と年齢の交互作用－音声化形態化と年齢の交互作用

＝2726.55－17.00－119.15－2442.01－18.36－52.37－29.41

＝48.25

⑨　偶然誤差を計算する。

偶然誤差の計算には便宜上の「等しいN」ではなく、実際のデータの個数を用いる（添え字は条件の番号)。

偶然誤差＝$SD_{111}^2 \cdot N_{111} + SD_{112}^2 \cdot N_{112} + SD_{113}^2 \cdot N_{113}$＋・・・

＝$3.8^2 \times 8 + 4.7^2 \times 8 + 4.4^2 \times 12 + 4.6^2 \times 15$・・・・

＝4047.81

⑩　分散分析表を作成する。

表５８　分散分析表（ＡＢＣＳタイプ）

要因	ＳＳ	df（ｶｯｺ内は計算式）	MS（SS/df）	F
音声形態（A）	17.00	1 （水準数－１）	17.00	＜１
単語数字（B）	119.15	1 （水準数－１）	119.15	4.74*
年齢（C）	2442.01	3 （水準数－１）	814.00	32.38**
A×B	18.36	1 （$df_A \times df_B$）	18.36	＜１

B×C	52.37	3 (df_B×df_C)	17.46	＜1
A×C	29.41	3 (df_A×df_C)	9.80	＜1
A×B×C	48.25	3 (df_A×df_B×df_C)	16.08	＜1
誤差	4047.81	161（全データ数－条件数）	25.14	
全体	6774.36	176（全データ数－１）	＊p＜.05	＊＊p＜.01

「全体」の分散（6774.36）は以上の分散値の合計である。Ｆ比の検定の際、分母の自由度は df②＝161 であるが、Ｆ分布表に載っていないので、df②＝120 を代用する。

さて、ここから事後処理に入る。上表の結果は、二つの主効果が有意であるが、まず、「単語・数字」の主効果から検討してみる。「単語・数字」の各水準の平均を再計算すると、以下のようになる。

単語	数字
5.2＋8.9＋12.5＋15.2＋（5.6＋9.2＋11.4＋15.5）/8＝10.4	6.3＋10.2＋13.4＋16.0＋（5.0＋13.7＋15.8＋16.6）/8＝12.1

以上のように単語条件の平均は 10.4、数字条件の平均は 12.1 である。この場合、水準は 2 つしかないので、主効果の有意性はそのまま両平均の有意差となる。すなわち、「単語＜数字」である。しかし、「年齢」の要因は 4 水準あるので、多重比較を適用してみなければならない。手順は以下のとおり。

① 「年齢」の各水準の平均を再計算する。

6 歳	10 歳	14 歳	18 歳
5.2＋5.6＋（6.3＋5.0）/4＝5.5	8.9＋9.2＋（10.2＋13.7）/4＝10.5	12.5＋11.4＋(13.4＋15.8）/4＝13.3	15.2 ＋ 15.5 ＋ （16.0 ＋ 16.6）/4＝15.8

② ＬＳＤを計算する。

ｔ値は、分散分析表の「誤差」の自由度（161）に相当する 5%点の値を求める。ただし、df＝161 はｔ分布表に載っていないので、df＝120 の分布を代用する（その 5%点はｔ＝1.980)。また、「誤差」の平均平方を読み取っておく（MSe＝25.14)。

以上を下式に代入し、ＬＳＤを計算する。

なお、「１平均あたりのデータ数」は、１マス４個の平均をつぶしているので、４Ｎである。

ＬＳＤ＝ｔ値√（２・MSe）/（１平均あたりのデータ数)

＝1.980×√（2×25.14）/（4×10.46）＝2.2

③ ＬＳＤを基準として各２平均の差を検定する。

平均の総当り数をつくってまとめると、わかりやすい。本例の場合、その必要はなく、隣り合う年齢どうしの平均の差はすべてＬＳＤ＝2.2 を越えている。したがって、各年齢の平均は「6 歳＜10 歳＜14 歳＜18 歳」という線型的大小関係にある（MSe=25.14、5%水準)。したがって、上の年齢ほど再認成績が良好であると言える。

b.被験者内計画の事例：ＳＡＢＣタイプ

論文名「スピーチの話題と特性が話し手の人物印象に及ぼす影響」

方法

実験計画：3×2×2 の被験者内計画。第一の要因は話題（自分、他人、事件)、第二の要因はスピーチの声の性別（男性の声、女性の声)、第三の要因は声の高さ(高い、低い)。

被験者：大学生２０人

手続き：実験は個別実験である。被験者にヘッドホーンを通して、全 12 条件のそれぞれのスピーチのテープを無作為な順序で聞かせた。そして、一回ごとにスピーチの話し手の性格（外向的か内向的か）について５項目６ポイントの評定尺度をあたえ、回答をさせた。

結果

回答結果は、ダミー項目をはずし、４項目合計として０～２０に得点化した。１０点が中立であり、それより得点が大きいほど外向的である。これを「外向性得点」と呼ぶことにする。表５９は、各条件の平均と標準偏差を示したもの。

表５９　各条件の外向性得点の平均と標準偏差（満点２０）（N＝２０）

話題		自分	他人	事件

声の性別			男の声	女の声	男の声	女の声	男の声	女の声
声の高さ	高	X	12.9	13.4	13.0	16.0	11.9	10.8
		ＳＤ	5.1	4.7	5.4	6.0	5.2	5.1
	低	X	6.7	8.4	9.3	10.9	9.9	9.7
		ＳＤ	4.6	5.3	4.1	4.9	4.6	4.2

分散分析の結果、話題の主効果（$F(2,38)=3.62$、$p<.05$）および声の高さの主効果（$F(1,19)=30.55$、$p<.01$）が有意であったが、声の性別の主効果は有意でなかった。また、「話題×声の高さ」の交互作用が有意傾向にあった（$F(2,38)=3.17$、$.05<p<.10$）。このほかの交互作用は有意でなかった。

下表は、分散分析表である。

解説

本例は、３要因の被験者内計画（ＳＡＢＣタイプ）である。被験者は全 12 条件すべてに割り当てられる。計算は極めて複雑である。また、Ｆ比の計算と検定の際に、特定の実験要因と特定の偶然誤差との対比を誤らないように注意する必要がある。

さて、本例の結果の記述は「主効果→交互作用」という順番であるが、これは不適切である。なぜなら、「Ａ×Ｃ」の交互作用が有意傾向にあったので、主効果Ａと主効果Ｃの有意性には独自の意味がないからである。すなわち、交互作用が有意である場合、それに関わる主効果に言及してはならない。

この事例では、3 要因計画の場合の交互作用の分析の仕方を学ぶ。

① 大平均を求め、主効果・交互作用を計算する。

被験者間計画の分散分析と被験者内計画の分散分析の異なる点は、偶然誤差の分け方だけである。したがって、実験要因（主効果と交互作用）の計算は、いずれのタイプも同一手順・同一結果になる。ここでは、前述した 3 要因被験者間計画の分散分析（ＡＢＣＳタイプ）をそっくり実行すればよい。

計算結果は以下のようになる。

話題（A）の主効果＝182.10、声の性別（B）の主効果＝50.42、声の高さ（C）の主効果＝889.35、Ａ×Ｂの交互作用＝88.03、Ｂ×Ｃの交互作用＝0.82、Ａ×Ｃの交互作用＝173.10、Ａ×Ｂ×Ｃの交互作用＝20.23

② 個人差（S）の分散を計算する。（考え方だけを示す。）

個人差（S）は各被験者の平均と、大平均との差である。被験者ごとにデータ（1 人 12 個）の平均を出し、それと大平均との差を求めて、二乗する。これを全被験者について合計する。

Ｓ＝（被験者のデータの合計/12－大平均）^{2}×12＋・・・・

表６０　分散分析表（ＳＡＢＣタイプ）

要因	ＳＳ	ｄf(ｶｯｺ内計算式)	MS（SS/df）	F
個人差（S）	443.16	19（被験者数－１）	23.32	
話題（A）	182.10	2（水準数－１）	91.12	3.62*
Ｓ×Ａ	954.63	38（$df_S \times df_A$）	25.12	
声の性別（B）	50.42	1（水準数－１）	50.42	2.50
Ｓ×Ｂ	382.72	19（$df_S \times df_B$）	20.14	
声の高さ（C）	889.35	1（水準数－１）	889.35	30.55**
Ｓ×Ｃ	553.11	19（$df_S \times df_C$）	29.11	
Ａ×Ｂ	88.03	2（$df_A \times df_B$）	44.02	2.15
Ｓ×Ａ×Ｂ	778.86	38（$df_S \times df_A \times df_B$）	20.50	
Ｂ×Ｃ	0.82	1（$df_B \times df_C$）	0.82	＜1
Ｓ×Ｂ×Ｃ	658.77	19（$df_S \times df_B \times df_C$）	34.67	
Ａ×Ｃ	173.10	2（$df_A \times df_C$）	86.55	3.17†
Ｓ×Ａ×Ｃ	1039.01	38（$df_S \times df_A \times df_C$）	27.34	
Ａ×Ｂ×Ｃ	20.23	2（$df_A \times df_B \times df_C$）	10.12	＜1
Ｓ×Ａ×Ｂ×Ｃ	1093.34	38（$df_S \times df_A \times df_B \times df_C$）	28.77	
全体	7307.65	239（全データ数－１）		

		† p＜.10	* p＜.05	** p＜.01

③　S×Aの分散を計算する。

S×Aは被験者（S）が要因Aの３水準に対して生じた誤差である。被験者ごとに要因A各水準のデータの平均を出し（被験者１人あたり３個の平均が出る）、それぞれと大平均との差を求めて二乗する。これを合計してゆく。そして、最後に、この合計値から個人差Sと要因Aの分散を引く。

S×A＝（(被験者のA①水準のデータの合計）/データ数（４）－大平均）^{2}×4＋・・・
　　－個人差S（443.16）－要因A（182.10）

④　同様にしてS×B、S×Cの分散を計算する。

S×B、S×Cでは、被験者があたえるデータ数は各水準６個ずつになる。

S×B＝（(被験者のB①水準のデータの合計）/データ数（６）－大平均）^{2}×6＋・・・
　　－個人差S（443.16）－要因B（50.42）

S×C＝（(被験者のC①水準のデータの合計）/データ数（６）－大平均）^{2}×6＋・・・
　　－個人差S（443.16）－要因C（889.35）

⑤　S×A×Bの分散を計算する。

S×A×Bは被験者（S）が要因Aと要因Bに対して生じる誤差である。１人の被験者は要因A（3 水準）と要因B（2 水準）に対して 3×２＝６個の平均をあたえる。それぞれと大平均との差を求め、二乗して、合計する。これを全被験者について合計する。最後に、個人差（S）、要因A、要因Bの分散を引く。

S×A×B
＝（(被験者のA①・B①水準のデータの合計）/データ数（２）－大平均）^{2}×2＋・・・
　－個人差S（443.16）－要因A（182.10）－要因B（50.42）

⑥　同様にしてS×B×C、S×A×Cを計算する。

S×B×C
＝((被験者のB①・C①水準のデータの合計)/データ数（３）－大平均)^2×3＋・・・
　－個人差S（443.16）－要因B（50.42）－要因C（889.35）

S×A×C
＝((被験者のA①・C①水準のデータの合計)/データ数（２）－大平均)^2×2＋・・・
　－個人差S（443.16）－要因A（182.10）－要因C（889.35）

⑦　S×A×B×Cを計算する。

各条件のSDを二乗して、N倍する。それを全 12 条件について合計する。その合計から、上で計算した誤差をすべて引く。

S×A×B×C＝SD^2＋SD^2＋SD^2＋SD^2＋・・・・
　－S－S×A－S×B－S×C－S×A×B－S×B×C－S×A×C

以上で分散分析は終了する。その結果、「話題×声の高さ」の交互作用（A×C）が有意傾向であったので、交互作用の分析に移る。

⑧　水準別の分散（単純主効果）を計算する。

「話題×声の高さ」（A×C）の補助表をつくり、それぞれのマスに入る平均２個から新たな平均を求める。

表６１「話題×声の高さ」の平均（データ数は１マスあたり 2N個）

話題		自分	他人	事件
声の性別		男声　女声	男声　　女声	男声　女声
声の高さ	高	(12.9＋13.4）/2＝13.15	(13.0＋　16.0)/2＝14.50	(11.9＋10.8）/2＝11.35
	低	(6.7＋　8.4）/2＝7.55	(9.3＋　10.9）/2＝10.10	(9.9　＋9.7）/2＝9.80

「声の高さ」高水準の「話題」の３つの平均を、さらに平均する（**水準平均**と呼ぶ）。そして各平均の、この水準平均からの差を計算して、二乗する。これを３平均について合計する。

水準平均＝（13.15＋14.50＋11.35）/3＝13.00

「声の高さ」高水準の「話題」の分散

＝（13.15－1300）^{2}×2N＋（14.50－13.00）^{2}×2N＋（11.35－13.00）^{2}×2N＝199.80

同様にして「声の高さ」低水準の「話題」の分散は下式による。

水準平均＝（7.55＋10.10＋9.80）/3＝9.15
「声の高さ」低水準の「話題」の分散
＝（7.55−9.15）^{2}×2N＋（10.10−7.55）^{2}×2N＋（9.80−9.15）^{2}×2N＝155.40
次に、「話題」の水準別に「声の高さ」の分散を計算する。
「話題」自分水準の「声の高さ」の分散は下式による。
水準平均＝（13.15＋7.55）/2＝10.35
「話題」自分水準の「声の高さ」の分散
＝（13.15−10.35）^{2}×2N＋（7.55−10.35）^{2}×2N＝627.20
同様にして「話題」他人水準の「声の高さ」の分散は 387.20、「話題」事件水準の「声の高さ」の分散は 48.05 となる。

⑨　以上の水準の分散（単純主効果）を検定するため、新たに水準別の偶然誤差を計算する。
例えば、「話題」事件水準の「声の高さ」の分散を検定するには、この事件水準だけの S×C を計算する。
　前掲の分散分析表では要因 C の分散に偶然誤差 S×C を対応させていたが、それと同様に、水準別の分散 CatA③には水準別の偶然誤差 S×CatA③を対応させなければならない。
この水準別の偶然誤差の計算は、事件水準（A③）のデータだけを用いて、上述の S×C の計算式に準ずる。ただし、その計算式中の個人差（S）と要因 C の分散を代入する部分があるが、その計算にも A③水準のデータだけを用いることになる（SatA③、CatA③を計算する必要がある）。
　このように水準別の分散（単純主効果）の検定に用いる水準別の偶然誤差を水準別誤差項と呼ぶ。この新たな偶然誤差の計算が必要となるのは、被験者内計画と混合計画の交互作用の分析においてである。被験者間計画ではあらわれない。

⑩　交互作用の分析表をつくる。
　水準別誤差項の実際の計算は省略する。
　以上の結果を交互作用の分析表にまとめる。
　表中の上段が「話題」の単純主効果の検定、下段が「声の高さ」の単純主効果の検定である。例えば「話題」の「声の高さ・高」（C①）水準の単純主効果 199.80 は偶然誤差 S×AatC①の 958.39 と比べて、平均平方で 3.96 倍あり、有意である。
　なお、A×C の交互作用は 173.10 であったが、表において、単純主効果の合計＝主効果＋交互作用（173.10）、が成立する。
　また、S×AatC○の合計＝S×A＋S×A×C、および S×CatA○の合計＝S×C＋S×A×C、が成立する。

表６２　「話題×声の高さ」の交互作用の分析表

要因	SS	df	MS	F
話題（A）	182.10	2	91.05	3.62*
声の高さ・高（C①）：	199.80	2	99.90	3.96*
S×AatC①：	958.39	38	25.22	
声の高さ・低（C②）：	155.40	2	77.70	2.85†
S×AatC②：	1035.25	38	27.24	
声の高さ（C）	889.35	1	889.35	30.55**
話題・自分（A①）：	627.20	1	627.20	21.10**
S×CatA①：	564.70	19	29.72	
話題・他人（A②）：	387.20	1	387.20	14.34**
S×CatA②：	513.24	19	27.01	
話題・事件（A③）：	48.05	1	48.05	1.78
S×CatA③：	514.18	19	27.06	
		† p ＜.10	* p ＜.05	** p ＜.01

⑪　多重比較を行う。
　先ほどの「話題」の二つの単純効果が有意または有意傾向であり、かつ、3 水準であるので、さらに多重比較を行わなければならない。そこで、LSD を計算するため t 値と MSe を用意する。

まず、ｔ値は、「話題」の偶然誤差「S×A」の自由度（df＝38）に相当する5%点を求める（df＝35のｔ分布を代用してｔ＝2.030）。次にMSeは、「声の高さ・高」水準と「声の高さ・低」水準では異なるので別々に計算する。まず、S×AatC①を表６２から読み取る（25.22）。なお、下式の「１平均あたりのデータ数」は40である（表６１の１マスのデータ数は２N）。

LSD＝ｔ値・√（2・MSe）/（1平均あたりのデータ数）＝2.030×√（2×25.22÷40）＝2.28

同様にして、「声の高さ・低」水準では、S×AatC②＝27.24を読み取って、上式のMSeに代入する(LSD＝2.37)。そして、「声の高さ・高」水準は2.28、「声の高さ・低」水準は2.37を基準として各2平均どうしの差を検定する（下表）。

表６３　「話題×声の高さ」の平均

話題		自分	他人	事件	
声の高さ	高	13.15	14.50	11.35	この水準では「他人＞事件」のみ有意
	低	7.55	10.10	9.80	この水準では「自分＜他人」のみ有意

交互作用の分析と以上の多重比較によって「話題×声の高さ」の交互作用のプロフィールは完全に確定した。（図7）。

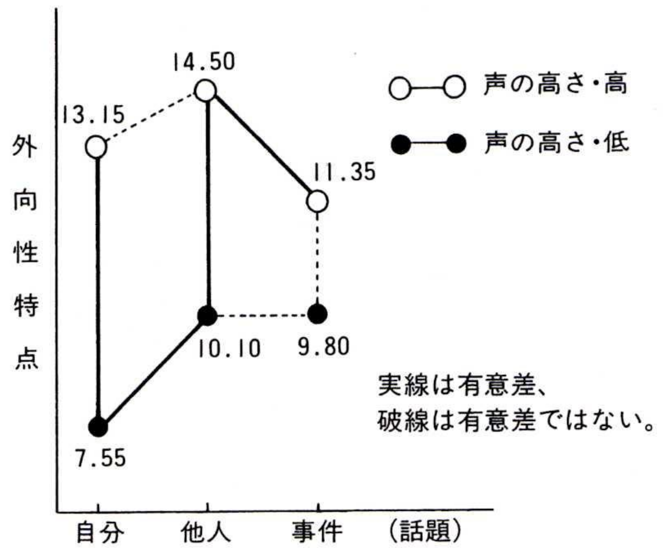


図7　「話題×声の高さ」の交互作用

論文記載例
「・・・(前略)・・・分散分析の結果、表60の分散分析表にみられるように、「話題×声の高さ」の交互作用が有意傾向であった（F (2,38) =3.17、.05＜ｐ＜.10)。そこで、各要因の水準別に単純効果を分析した結果、表６２（交互作用の分析表）に示したように、話題が事件の場合は声の高さの効果は有意でなかった。また、話題の単純主効果の多重比較によれば、高い声の場合に「他人＞事件」(MSe＝25.22、5%水準)、低い声の場合に「自分＜他人」(MSe＝27.24、5%水準) という平均の有意差があった。したがって、特に高い声で他人を話題にすると外向性の印象が強く、低い声で自分を話題にすると内向性の印象が強くなると言える。なお、声の性別の効果は有意でなかった。」

c.　混合計画の事例Ｉ：ABSC タイプ
論文名「漫画の吹き出しにおける対自性のセリフについて」
方法
調査計画：2×3×2 の混合計画。第一の要因は少年漫画・少女漫画（被検体間の配置)、第二の要因はジャンル（ストーリー・アクション・ギャグ：被検体間の配置)、第三の要因はセリフの長さ（1 文未満・1 文以上：被検体内の配置)。
被検体：少年・少女漫画雑誌からストーリー・アクション・ギャグの各ジャンルの漫画を無作為に 10 本ずつ、計60本を抽出した。
手続き：各漫画の吹き出しのなかのセリフについて、それが他者に向けられたセリフ（対他性セリフ）であるか、自己に向けられたセリフ（対自性セリフ）であるかを判定した。感動語は除外した。
結果
セリフは、1 文未満のセリフと 1 文以上のセリフとに分けた。そして、各漫画ごとに、セリフの総数に占める対自性セリフ数の割合を算出した。図８は、その平均を図示したものである。

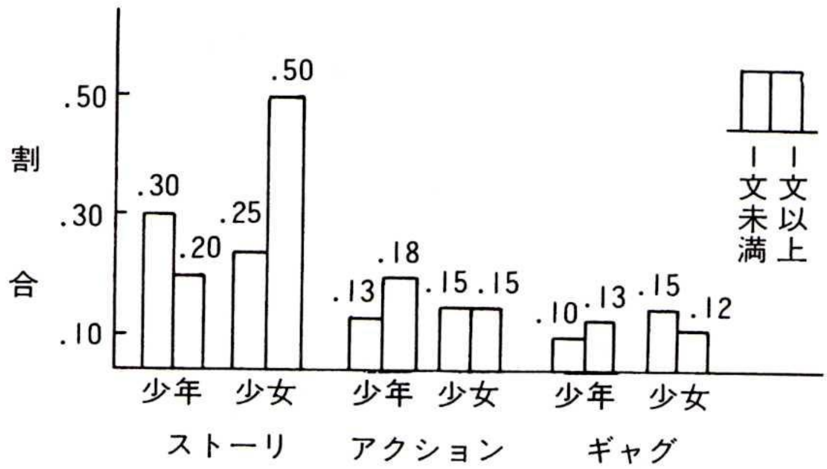


図８　少年・少女漫画別の各ジャンルの対自性セリフの割合

分散分析を行った結果（表６４)、二次の交互作用が有意であった（F (2,54) =13.70、ｐ＜.01)。図 8 によれば、ストーリー物において、少年漫画は特に 1 文未満の対自性のセリフが多く、少女漫画は特に 1 文以上の対自性のセリフが多いと言える。
解説
本例は、ABSC タイプの分散分析である。分析単位が漫画であるため、「被験者」でなく「被検体」を用いているが、「対象」でもよい。要因配置は被検体間２つ、被検体内１つである。3 要因混合計画の分散分析には、このほかに、後述する ASBC タイプがある（被験者間１つ、被験者内２つ)。
さて、データ表示の図８をみると、多くの平均が０に接近した値をとっているので数値の変換を施すべきであ

ったかもしれない（この場合データの値は割合なので角変換が適切である）。

表６４　３要因混合計画Ｉの分散分析表（ABSC タイプ）

要因	SS	df（ｶｯｺ内計算式）	MS（SS/df）	F
少年少女（A）	0.065	1（水準数－１）	0.065	5.00*
ジャンル（B）	0.820	2（水準数－１）	0.410	31.54**
A×B	0.095	2（$df_A \times df_B$）	0.048	3.69*
個体差（S）	0.685	54（検体数－A2×B2）	0.013	
ｾﾘﾌの長さ（C）	0.033	1（水準数－1）	0.033	3.30†
A×C	0.048	1（$df_A \times df_C$）	0.048	4.80*
B×C	0.029	2（$df_B \times df_C$）	0.015	1.50
A×B×C	0.274	2（$df_A \times df_B \times df_C$）	0.137	13.70**
S×C	0.515	54（$df_S \times df_C$）	0.010	
全体	2.565	119（全データ数－1）		
		† $p < .10$	* $p < .05$	** $p < .01$

分散分析の結果、ほとんどの要因が有意または有意傾向であったが、「二次の交互作用」（A×B×C）が有意であるため、「一次の交互作用」（A×B、 B×C、 A×C）も主効果も取り上げることができない。このように高次の交互作用が有意であると、低次の交互作用と主効果の有意性はすべて従属的な意味しかもたなくなる。この点、本例では、主効果や一次の交互作用の有意性に触れず、二次の交互作用を、データ表示の図８との整合性に基づいてサラリと解釈している。この処理は事後処理の煩雑さを避けて賢明である。

しかし、「見ればわかる」ような結果の場合は幸いであるが、そうでない場合は、高次の交互作用を一段ずつバラしていかなければならない。その結果は、たいてい各条件ごとのケース・バイ・ケースの結果となり、実験計画を組んだ価値が薄れるのが通例である。

さて、以下に、二次の交互作用の事後分析を含めて、ABSC タイプの分散分析の手順を示す。

①　大平均を求め、主効果・交互作用の分散を計算する。

実験要因（主効果・交互作用）の分散については、どのタイプの分散分析も計算手順は同一である。前述した３要因被験者間計画の分散分析（ABCS タイプ）を参照していただきたい。これによって、主効果 A,B,C、交互作用 A×B、 B×C、 A×C 、A×B×C のそれぞれの分散を求める。結果は、表６４のようになる。

なお、各条件のデータの個数が等しくない場合は、調和平均による便宜上の「等しい N」を計算することを忘れずに。

②　被検体間の個体差（S）を計算する。

これは総計 60 本の漫画の差異である。人間を研究対象としたときの「個人差」に相当する。この計算には、各被検体のデータの合計を用いる（下のような補助表をつくるとわかりやすい）。

A 少年少女	B ジャンル	S 被検体	C １文未満	 １文以上	合計
少年向け	ストーリ	漫画１	0.43	0.24	0.67
少女向け	ギャグ	漫画２	0.12	0.05	0.17
少女向け	ストーリ	漫画３	0.10.	0.34	0.44

個体差（S）＝

（$0.67^2+0.17^2+0.44^2$＋・・・）/（1 漫画あたりのデータ数）(2) －（0.67＋0.17＋0.44＋・・）^{2}/全データ数（120）－主効果 A－主効果 B－交互作用 A×B

③　被検体内の偶然誤差（S×C）を計算する。

（S×C）とは、個々の漫画（S）が「セリフの長さ」（C）によって生じる誤差である。この計算には全体の偶然誤差を計算して、そこから、上で求めた個人差（S）を引く。ただし、下式の N は各条件の実際のデータの個数とする。偶然誤差の計算には、調和平均による「等しい N」は用いない。

S×C＝（SD^2・N＋SD^2・N＋・・・・）－個体差 S

以上の結果を分散分析表に書き込む。

④　交互作用の分析へ移る。

検定の結果、二次の交互作用（A×B×C）が有意であったので（F（2,54）＝13.70、p＜.01）、それを優先して分析する。この場合、一次の交互作用と主効果に言及してはならない。

分析の仕方は、3 要因のうち任意の一つの要因に注目し、その水準別に一段低次の交互作用に落とす。ここでは、のちの解釈のしやすさを考慮して要因Ｂ（ジャンル）に注目する。

B①水準におけるA×C
A×B×C → 要因Ｂの３水準に注目して分ける → B②水準におけるA×C
B③水準におけるA×C

上の方針にそって、要因Ｂの各水準ごとに「A×C」の分散を再計算する。準備として要因Ｂの各水準ごとに平均をまとめておく。

表６５　各条件の平均

ジャンル	B①（ストーリー）				B②（アクション）				B③（ギャグ）			
少年少女	A①		A②		A①		A②		A①		A②	
セリフの長さ	C①	C②	C①	C②	C①	C②	C①	C②	C①	C②	C①	C②
X	.30	.20	.25	.50	.13	18	.15	.15	.10	.13	.15	.12

まず、B①水準における「Ａ×Ｃ」の分散を計算する。これには、以下のような補助表をつくり、ヨコ計とタテ計を求めておく。

	C①	C②	ヨコ計
A①	.30	.20	.50
A②	.25	.50	.75
タテ計	.55	.70	1.25

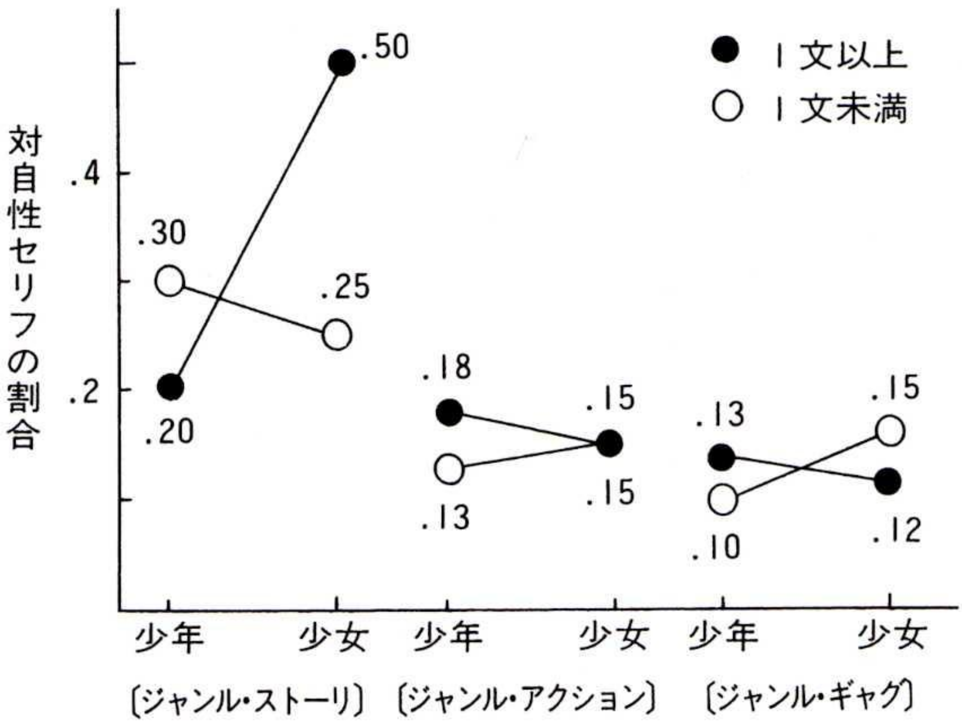
基本的に、この補助表において、「Ａ×Ｃ」の２要因の分散分析を行えばよい。
大平均＝平均の総計／平均の個数＝1.25／4＝.3125
主効果Ａ＝（(A①水準の平均－大平均）^{2}＋（A②水準の平均－大平均）^{2})・１水準あたりのデータ数
＝（(.50/2－.3125）^{2}＋（.75/2－.3125）^{2}) ×（2×10)
＝.1563
主効果Ｃ＝（(C①水準の平均－大平均）^{2}＋（C②水準の平均－大平均）^{2})・１水準あたりのデータ数
＝（(.55/2－.3125）^{2}＋（.70/2－.3125）^{2}) ×（2×10)
＝.0563
Ａ×Ｃ＝（(X－大平均）^{2}＋（X－大平均）^{2}＋（X－大平均）^{2}＋（X－大平均）^{2})・N－A－C
＝（(.30－.3125）^{2}＋（.20－.3125）^{2}＋（.25－.3125）^{2}＋（.50－.3125）^{2}) ×10－.1563－.0563
＝.306
以上、Ｂ①水準における「Ａ×Ｃ」の分散は、0.306 である。同様にして補助表をつくり計算すると、Ｂ②水準の「Ａ×Ｃ」の分散は 0.006、Ｂ③水準の「Ａ×Ｃ」の分散は 0.009 となる。
ここで、以上の分析結果を分散分析表に書き込む。
ここで、水準別の一次の交互作用の合計（0.306＋0.006＋0.009）は「Ａ×Ｃ」と「Ａ×Ｂ×Ｃ」の分散の合計（0.048＋0.274）に一致する。
検定の結果は、「ジャンル」ストーリー水準の「Ａ×Ｃ」の交互作用だけが有意であった（Ｆ（1,54）＝30.60、ｐ＜.01)。そこで、これ以降の分析は、さらに、この有意な交互作用を単純主効果へバラしてゆくことになるが、その手順は前述した「(一次の）交互作用の分析」と同様なので省略する。
ここでは、以上の結果から、二次の交互作用の意味について考える。
さて、要因Ｂの各水準ごとに「Ａ×Ｃ」(少年少女×セリフの長さ）の平均のプロフィールを描いてみる。
図の線グラフの交差の程度が、交互作用のＦ比の大きさになる。あきらかに「ジャンル」ストーリー水準の「Ａ×Ｃ」のプロフィールが激しく交差しているが、そのとおりに分析結果は有意であった。他のアクション水準とギャグ水準の交互作用はそれほどでなく、有意でない。

このように、二次の交互作用の有意性は、一次の交互作用のプロフィールが等しくないことを意味する。なお、今回は結果が簡明であるので、要因Ｂ（ジャンル）に注目して１種類の交互作用（Ａ×Ｃ）だけにバラしたが、注目する要因を順番に変えてゆき、全種類の一次の交互作用を分析してみる場合もある。

最後に、以上の事後分析の結果の論文記述例を示す。
「ジャンルに注目して二次の交互作用を分析した結果、ストーリー物においてのみ「少年少女×文の長さ」の交互作用が有意であった（F（1,54）=30.60、p＜.01)。すなわち、ストーリー物の少年漫画には1文未満、ストーリー物の少女漫画には1文以上の対自性のセリフが多かった。他のジャンルには、このことは言えない。」

表６６　二次の交互作用の分析結果を書き込んだ分散分析表

要因	SS	df	MS	F
少年少女（A)	0.065	1（水準数－１）	0.065	5.00*
ジャンル（B)	0.820	2（水準数－１）	0.410	31.54**
A×B	0.095	2（df_A×df_B)	0.048	3.69*
個体差（S)	0.685	54（検体数－A2×B2)	0.013	
ｾﾘﾌの長さ（C)	0.033	1（水準数－1)	0.033	3.30†
A×C	0.048	1（df_A×df_C)	0.048	4.80*
ｼﾞｬﾝﾙ･ｽﾄｰﾘｰ	0.306	1	0.306	30.60**
ｼﾞｬﾝﾙ･ｱｸｼｮﾝ	0.006	1	0.006	＜1
ｼﾞｬﾝﾙ･ｷﾞｬｸﾞ	0.009	1	0.009	＜1
B×C	0.029	2（df_B×df_C)	0.015	1.50
A×B×C	0.274	2（df_A×df_B×df_C)	0.137	13.70**
S×C	0.515	54（df_S×df_C)	0.010	
全体	2.565	119（全データ数－1)		
		† p＜.10	* p＜.05	** p＜.01



ｄ．混合計画の事例Ⅱ：ＡＳＢＣタイプ
ここでは、JavaScript-STARの分散分析メニューを用いて、実際にデータ処理の演習を行ってみる。
事例
　実験計画は、3×2×2の混合計画。第一の要因Ａ（3水準）は被験者間配置、第二の要因Ｂ（2水準）と第三の要因Ｃ（2水準）は被験者内配置である。
　データリストは以下のとおりである。被験者間配置の要因Ａの水準には異なる被験者が割り当てられているが、人数が等しくないことに注意すること。

表６７　ＡＳＢＣタイプのデータ・リスト

	Ｓｕｂ.	B①		B②		合計
		C①	C②	C①	C②	
A①	山際	60	25	30	40	155
	田中	45	45	50	60	200
	遠藤	40	35	40	35	150
A②	小野	42	44	32	64	182
	高橋	66	30	50	40	186
	竹村	56	52	54	46	208
A③	服部	75	35	70	45	225
	古川	50	60	45	55	210
	石橋	65	40	50	70	225
	倉沢	70	60	55	60	245

①　JavaScript-STAR の分散分析メニューを選ぶ。

「STAR」のフォルダを開き、index.htm をクリックする。

JavaScript-STAR のＴＯＰページが表れるので、画面左の分散分析メニューから、「AsBC（3 要因混合)」をクリックする。タイトルと「データ枠」「結果の枠」があらわれる。

②　データを入力する。

画面の中央にある「Q&A 入力」のボタンをクリックする。以下、コンピュータの質問に答えを返すやり方でデータを入力する。なお、データ数が多い場合は、直接「データ枠」に書き込んだり貼り付けたりする方法が便利なので、「データ枠」右肩の「Data Type」をクリックして参照するか、田中・中野著「クイック・データアナリシス」（新曜社）を参照してください。

- 「要因Aの名前を入力してください」と聞いてくる　→　Group などの何か適当な名前を入力する（ＯＫボタンを押す)。表示されている「A」のままでもよい。
- 「要因Aの水準数を入力してください」と聞いてくる。　→　３　と入力する。数字は必ず半角モードで入力すること。
- 「要因Aの第一水準の被験者数を入力してください」と聞いてくる　→　３　と入力する。A①水準は山際、田中、遠藤の３人。
- 「要因Aの第二水準の被験者数を入力してください」と聞いてくる　→　３　と入力する。ここは小野、高橋、竹村の３人。
- 「要因Aの第三水準の被験者数を入力してください」と聞いてくる　→　４　と入力する。ここは、服部、古川、石崎、倉沢の４人。
- 「要因Bの名前を入力してください」と聞いてくる　→　ここでは、表示されている「B」のままにする。
- 「要因Bの水準数を入力してください」と聞いてくる　→　２　と入力する。
- 「要因Cの名前を入力してください」と聞いてくる　→　表示されている「C」のままにする。
- 「要因Cの水準数を入力してください」と聞いてくる　→　２　と入力する。

ここまで計算を制御する情報をあたえている。質問が AsBC の順になっている（s は被験者 subjects)。「よろしいですか？」と確認が入るので、ＯＫを押す。以下、データを１個ずつ入力してゆくことになる。

- 「要因A第一水準　被験者 No.1　要因B第一水準　要因C第一水準？」と聞いてくる。　→　６０　と入力する。これは、山際のデータの最初の値である。
- 「要因A第一水準　被験者 No.1　要因B第一水準　要因C第二水準？」と聞いてくる。　→　２５　と入力する。質問は「要因C第二水準」を指定しているので、山際のデータの次の値を答える。
- 「要因A第一水準　被験者 No.1　要因B第二水準　要因C第一水準？」と聞いてくる。　→　３０　と入力する。質問は「要因B」に移ったが、山際のデータを左から順番にあたえていくようにすればよい。入力の間違いに気が付いてもそのままにして、次の入力に進み、最後に見直しがあるのでそこで直すようにする。その方が楽。
- 「データの作成が終了しました」と確認が入る　→　ＯＫを押す。「データ枠」にデータが表示される。

表 67 と同じ形式のデータリスト部分があるので、ここで入力ミスを訂正する。直接、間違った数字を消して書き換える。

③　データを保存する。

すぐに計算に行かないで、いったん、データを保存する。

- 「データ枠」のどこかをクリックし、アクティブにする。
- 画面上部で「編集」→「すべてを選択」を選ぶ。
- 続けて、「編集」→「コピー」を選ぶ。
- いったん、JabaScript-STAR の画面を最小化する。
- デスクトップ画面でマウスを右クリックして、「新規作成」→「テキスト文書」を選ぶ→「新規テキスト文書」が作られる。
- 「新規テキスト文書」をクリックして、開く。
- 「編集」→「貼り付け」（ペースト）を選び、この文書にデータを貼り付けておく。

④　分散分析を実行する。

Java Script-STAR の画面を元に戻し、中央にある「計算」ボタンを押すと、分散分析が実行される。結果は「結果の枠」にあらわれる。これもコピー&ペーストで、上の新規文書に貼り付けて、データと結果を一緒にしておく。最終的に、適当な名前を付けて保存しておく。

⑤　結果の出力を見てデータ表示の図表をつくる。

下に結果の出力例を示す。これからまずデータ表示の図表を作成する。ここでは、表にしてNと平均と標準偏差を、読み取って掲載する（表 68)。なお、平均の記号はMでもよい。

Mean&SDs(of samples')
[A=A]
[B=B]
[C=C]

A	B	C	N	Mean	S.D.
1	1	1	3	48.3333	8.4984
1	1	2	3	35.0000	8.1650
1	2	1	3	40.0000	8.1649
1	2	2	3	45.0000	10.8012
2	1	1	3	54.6667	9.8432
2	1	2	3	42.0000	9.0921
2	2	1	3	45.3333	9.5685
2	2	2	3	50.0000	10.1980
3	1	1	4	65.0000	9.3542
3	1	2	4	48.7500	11.3880
3	2	1	4	55.0000	9.3541
3	2	2	4	57.5000	9.0139

＊カク・セルノNガフゾロイデス。
＊Unweighted-Mean ANOVA ヲオコナイマス。
＊N=3.27（チョウワ・ヘイキン）トカテイシマス。

Analysis of Variance
[A(3)=A]
[B(2)=B]
[C(2)=C]

S.V.	SS	df	MS	F
A	1387.5038	2	693.7519	7.69*
Sub	631.8527	7	90.2647	

B	0.2292	1	0.2292	0.00ns
A×B	4.7765	2	2.3883	0.05ns
S×B	345.0179	7	49.2883	
C	246.8201	1	246.8201	0.89ns
A×C	17.0492	2	8.5246	0.03ns
S×C	1939.8600	7	277.1229	
B×C	807.5928	1	807.5928	7.86*
A×B×C	0.8674	2	0.4337	0.00ns
S×B×C	719.0125	7	102.7161	
Total	6100.5822	39	+p＜.10 *p＜.05	**p＜.01

表 68　各条件の得点の平均と標準偏差（満点は 100）

		A①		A②		A③	
		B①	B②	B①	B②	B①	B②
C①	N	3	3	3	3	4	4
	M	48.3	40.0	54.7	45.3	65.0	55.0
	SD	8.5	8.2	9.8	9.6	9.4	9.4
C②	N	3	3	3	3	4	4
	M	35.0	45.0	42.0	50.0	48.8	57.5
	SD	8.2	10.8	9.1	10.2	11.4	9.0

⑥　分散分析表をつくる。

表 69　3 要因混合計画Ⅱ（ＡＳＢＣタイプ）の分散分析表

要因	ＳＳ	ｄｆ（括弧内計算式）	ＭＳ（ＳＳ/ｄｆ）	Ｆ
要因A	1387.50	2（水準数−1）	693.75	7.69*
個人差（S）	631.85	7（人数−Aの水準数）	90.26	
要因B	0.23	1（水準数−1）	0.23	＜1
A×B	4.78	2（$df_A \times df_B$）	2.39	＜1
S×B	345.02	7（$df_S \times df_B$）	49.29	
要因C	246.82	1（水準数−1）	246.82	＜1
A×C	17.05	2（$df_A \times df_C$）	8.52	＜1
S×C	1939.86	7（$df_S \times df_C$）	277.12	
B×C	807.59	1（$df_B \times df_C$）	807.59	7.86*
A×B×C	0.87	2（$df_A \times df_B \times df_C$）	0.43	＜1
S×B×C	719.01	7（$df_S \times df_B \times df_C$）	102.72	
全体	6100.58	39（全データ数−1）		*p＜.05

⑦　交互作用の分析表をつくる。

下の表は、結果の出力に続きである。本例の結果は交互作用Ｂ×Ｃが有意であるので、STAR は自動的に交互作用の分析へ移行して結果を出力した。もし交互作用が有意でない場合は、この出力はない。最初に、交互作用Ｂ×Ｃの平均が新たに計算されているので、交互作用のプロフィールを描いておく。この図のなかに検定の結果を書き込んでゆけば理解しやすい。

さて、交互作用は各要因の水準別の単純主効果と、それに対応する水準別の偶然誤差との計算となる。その結果は、表 72 のようにまとめられる。この結果から、図のようなプロフィールが確定する。この表のＦ比の検定結果が、図中の、どの平均間の差を検定しているかを確認する。

Analysis of B×C　Interaction

[B(2)=B]

[C(2)=C]

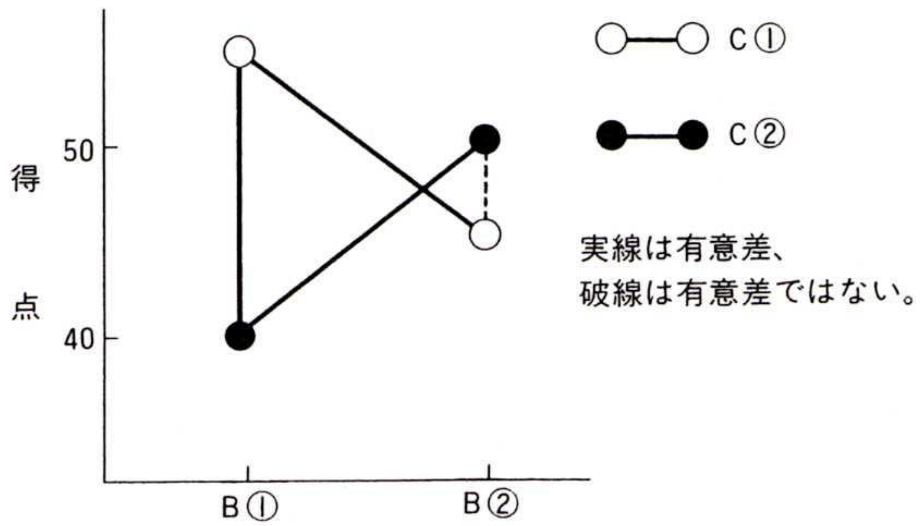
B	C	N	Mean
1	1	10	56.0000
1	2	10	41.9167
2	1	10	46.7778
2	2	10	50.8333

S.V.	SS	df	MS	F
B at C1	417.5152	1	417.5152	6.39*
(S×B at C1	457.6559	7	65.3794)	
B at C2	390.3068	1	390.3068	4.51+
(S×B at C2	606.3745	7	86.6249)	
C at B1	973.6705	1	973.6705	4.67+
(S×C at B1	1460.0385	7	208.5769)	
C at B2	80.7424	1	80.7424	0.47ns
(S×C at B2	1198.8341	7	171.2620)	

表 72　「B×C」の交互作用の分析表

要因	ＳＳ	ｄｆ	ＭＳ	Ｆ
B at C①	417.52	1	417.52	6.39*
S×B at C①	457.66	7	65.38	
B at C②	390.31	1	390.31	4.51†
S×B at C②	606.37	7	86.62	
C at B①	973.67	1	973.67	4.67†
S×C at B①	1460.04	7	208.58	
C atB②	80.74	1	80.74	0.47
S×C at B②	1198.83	7	171.26	

† p＜.10　＊p＜.05



⑧　論文に記載する文章を書く。

データ分析としては、あと要因Aの主効果の多重比較が残っている。その必要があるときは、STAR は自動的に処理をして以下のように結果を出力する。これも参照して、これまで作成した図表に言及しながら、結果を文章に書く。

Multiple Comparisions by LSD
[MSe=90.2647, *p<.05]
[LSD=8.78105]

[Main Effect of Factor A]		
A	N	Mean
1	13	42.0833
2	13	48.0000
3	13	56.5625

A1=A2 n.s.
A1<A3 *
A2=A3 n.s.

「表 68 は各条件の平均と標準偏差を示したものである。分散分析の結果、表 69 の分散分析表にみられるように要因Aの主効果（F（2,7）＝7.69、p＜.05）および要因Bと要因Cの交互作用（F（1,7）＝7.86、p＜.05）が有意であった。そのほかの主効果と交互作用は有意でなかった。

そこで、要因Aの各水準の平均をＬＳＤ法によって多重比較した結果、A①水準とA③水準との間に有意差があった（MSe=90.26、p＜.05）。

また要因Bと要因Cの交互作用を分析した結果、表 72 に示したように要因Bの単純主効果はC①水準、C②水準で共に有意または有意傾向であった。C①水準では、B①条件の平均がB②条件の平均よりも有意に大きく（F（1,7）＝6.39、p＜.05）、またC②水準では、B②条件の平均がB①条件の平均よりも有意に大きい傾向があった（F（1,7）＝4.51、.05＜p＜.10）。

他方、要因Cの単純主効果は、B①水準では有意傾向であったが、（F（1,7）＝4.67、.05＜p＜.10）、B②水準では有意でなかった。B①水準では、C①条件の平均がC②条件の平均よりも大きい傾向があると言える。」

上の文章中、第１文がデータ表示、第２文・第３文が分散分析の結果である。

第２段落は、要因Aの主効果についての多重比較の結果であり STAR の出力を参照して書かれている。

第３段落以下は交互作用の分析結果である。最初に、単純主効果の有意性を記載している。次に、有意であった場合に各平均の大小関係を記載している。

なお、実際の文章ではA，B，Cや①、②、③などの記号が使われることはありえず、すべて具体的な要因名、水準名が置き換わる。

以上で、「実験計画法と分散分析」を終わる。分散分析の計算には、実際はコンピュータ・プログラムを使用する場合が多いが、交互作用や多重比較の事後処理までフォローしてくれるプログラムは少ない。

４－５　まとめ：分散分析の結果を読み取る方針

分散分析の結果の読み取りは、それぞれの効果のＦ比が有意であったかどうかを見ることである。このときの見方は次の方針にしたがう。

- 高次の効果から見てゆく。
- そして高次の効果が有意のときは、それより低次の効果は見ない。

ここで、「高次」、「低次」は要因の掛け合わせの数である。例えば、単独の要因Aの効果は（他の要因と掛け合わせていないので）０次である。他の要因１つと掛け合わせられた交互作用A×BやB×Cは１次であり、さらに交互作用A×B×C は他の要因２つと掛け合わせられているので２次である。

上記の読み取りの方針にしたがえば、２要因計画の分散分析では交互作用A×B が有意であれば、もはや主効果は（有意であろうがなかろうが）見ないということになる。それでよい。交互作用が有意であることによって

主効果の意味がなくなったので、水準別の単純主効果を再計算して検定したのである。
では、３要因計画の分散分析はどうであろうか。その場合、もし最高次の交互作用 A×B×C が有意であれば、それ以下の１次の交互作用も、０次の主効果もすべて見ない。
実際の方針としては、分散分析表の読み取りは上の段からではなく、下の段から見てゆくことになる。
さて、以下の２つの例では、どのように読み取るべきか。

例①：３要因ＡＢＳＣタイプ

A	＊
B	＊
A×B	ns
S	
C	＊
B×C	ns
A×C	＊
A×B×C	ns
S×C	
ns　p ＞.10　　＊　p ＜.05	

例②：３要因ＡＳＢＣタイプ

A	＊
S	
B	ns
A×B	＊
S×B	
C	＊
A×C	ns
S×C	
B×C	ns
A×B×C	ns
S×B×C	
ns　p ＞.10　　＊　p ＜.05	

まず、例①から読み取る。第一ステップとして、２次の交互作用 A×B×C を見る。これは有意でない。そこで、第二ステップとして、１次の交互作用 A×B、B×C、A×C を見る。このなかでは A×C だけが有意である。これによって 0 次の要因 A・要因 C の主効果は（有意であっても）意味がないが、第三ステップとして要因 B の主効果を見ることができる。したがって、例①の結果として読み取られるものは、A×C の有意性と主効果 B の有意性だけである。
同様に、例②では A×B の有意性と主効果 C の有意性を読み取ることができる。主効果 A の有意性は取り上げてはならない。

備考：被験者間計画における等質化の方法
等質化の方法
等質化の必要性があるのは被験者間計画である。被験者内計画にはない（その代わり順序の相殺をおこなわなければならないが）。すなわち、被験者間計画は、各条件を異なる被験者に割り当てるため、実験前の各条件の状態や特性は等質でないおそれがある。もし等質でない場合は、正当な実験的比較を行うことができない。なぜなら、実験前の非等質性がそのまま実験後の有意差になってあらわれたり、その反対に有意差を打ち消したりするからである。そこで、各条件を等質化するために、一般に二つの方法をとる。

① ランダム・サンプリング（無作為抽出）を実行する。

これは各条件をでたらめに被験者に割り当てることである。例えば、3 条件 A・B・C を設定する場合、個別実験ならば実験室にやって来た順を利用して、A、B、C、A、B、C、。。。と割り当ててゆく。また、質問紙による条件設定ならば、異なる質問紙をたがいちがいにA、B、C、A、B、C、。。。と積んでおいて上から順番に配ってゆく。これがランダム・サンプリングにならない例は、条件 A を 20 人実施した次は条件 B を 20 人実施するとか、同条件の質問紙だけを山積みしておいて配るとか、という場合である。

ランダム・サンプリングによって、(データ特性上の）平均的被験者も異常な被験者もだいたい同数ずつ各条件にはいる。したがって、全体としてみれば、どの条件も同様の状態・同様の特性を備えることになるであろう。ただし、これは「仮定」の意味が強い。

② 実験前のデータをとり、有意差検定をおこなう。

これはランダム・サンプリングよりも確実な方法である。すなわち、すべての被験者から実験前の状態や特性についてのデータをとり、実験に各条件の平均を計算して有意差があるか否かを判定する。ただし、実験前のデータは実験後のデータと相関するものでなければならない。また、このときの平均の有意差検定の方法は、つねに 1 要因の分散分析（AS タイプ）を用いる（たとえ実験計画が 2 要因以上であっても)。その結果、もし有意差がなければ、各条件は等質であるとみなしてそのまま実験に移ってよい。しかし、もし有意差があった場合は、以下のいずれかの対処をおこなわなければならない。

- 各条件の被験者を入れ替えて平均の有意差が出なくなるように「群構成」を行う。これが、最も正当な方法であるが、学校現場等の実験ではなかなか困難である。
- 各被験者ごとに実験前のデータと実験後のデータの差をとって、この数値を分散分析する。ただし、実験前・実験後のデータは同種類のものでなければならない。いわゆる「プリテスト・ポストテスト法」である。また、この「プリ・ポスト」を被験者内の 1 要因として追加し、混合計画の分散分析に持ち込んでもよい。
- 実験前・実験後のデータの種類が異なる場合は、共分散分析を用いる。例えば、等質性のチェックのため知能指数をとり、実験結果の指標として正答率をとったような場合である。しかし、共分散分析は使用上の制約がきつく、計算が複雑かつ多量である。

5. 相関・予測の分析

目的

有意差の分析が条件間の差異をみるのに対して、相関・予測の分析は条件間の関連をみる。したがって、各条件のデータどうしは独立であってはならず、対応のあるデータでなければならない。対応のあるデータを入手するには、1 人の被験者または対象者から 2 個以上のデータをとればよい。例えば、1 被験者を 2 条件以上に割り当てたり、1 対象者に複数の心理検査を施行したりする。このとき、データの種類や単位は違っていてもかまわない。例えば、数学のテストの得点と技術科の作品の評価点、あるいはテニスのサーブの練習回数（本）とその成功率（%)、など。また、必ずしも同一の被験者（同一の対象者）からとったデータでなくとも、データどうしを対応づけることができればよい。例えば、生徒と親の知能指数、あるいは、友人どうしの社会的地位得点（ソシオメトリック・テストによる)、など。相関・予測の分析は、このような対応するデータどうしの間に、次のような 3 点の規則的関係を見出そうとしている。

① 正の相関
いわゆる正比例の関係があるかどうかを見る。すなわち、一方のデータの数値が大きい場合に、それに対応する他方のデータの数値も大きい（小さい場合には小さい）なら、正の相関がある。

② 負の相関
いわゆる逆比例の関係があるかどうかをみる。一方のデータの数値が大きい場合に、それと反対に他方のデータの数値が小さい（小さい場合に大きい）なら、負の相関がある。

③ 予測式
データどうしの相関関係を、一次方程式として表現する。この方程式に一方のデータの任意の数値を代入

すれば、他方のデータの数値を予測的に求めることができる。

ここで、分析の具体的なイメージをもっておくために、以下の事例をあげる。

下の図9は、10人の生徒を対象に、テニスのサーブの練習本数と成功率を調べ、図示したものである。この場合、相関・予測の分析は正の相関を見出し、成功率（%）＝0.17×練習本数＋16、という予測式を立てることができる（図中の直線は予測直線）。

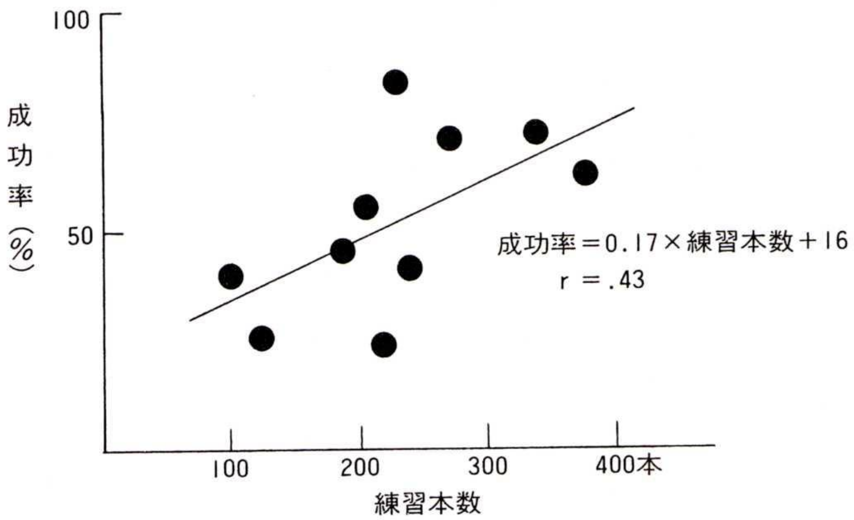


図9　サーブの練習本数と成功率の相関

用語

相関・予測の分析に特有の専門用語を、以下に解説する。

① 変数

変数（variable）とは「変化する数値」という意味である。相関・予測の分析では、データを変数とみなす。前頁の例では、サーブの練習本数のデータが一つの変数であり、成功率のデータがもう一つの変数である。したがって前例の場合は「2 変数の相関と予測を分析した」という言い方をする。このように、変数という言い方をするのは、相関・予測の分析がデータどうしの対応関係を数学の関数関係になぞらえて分析しようとするからである。

② 予測と回帰

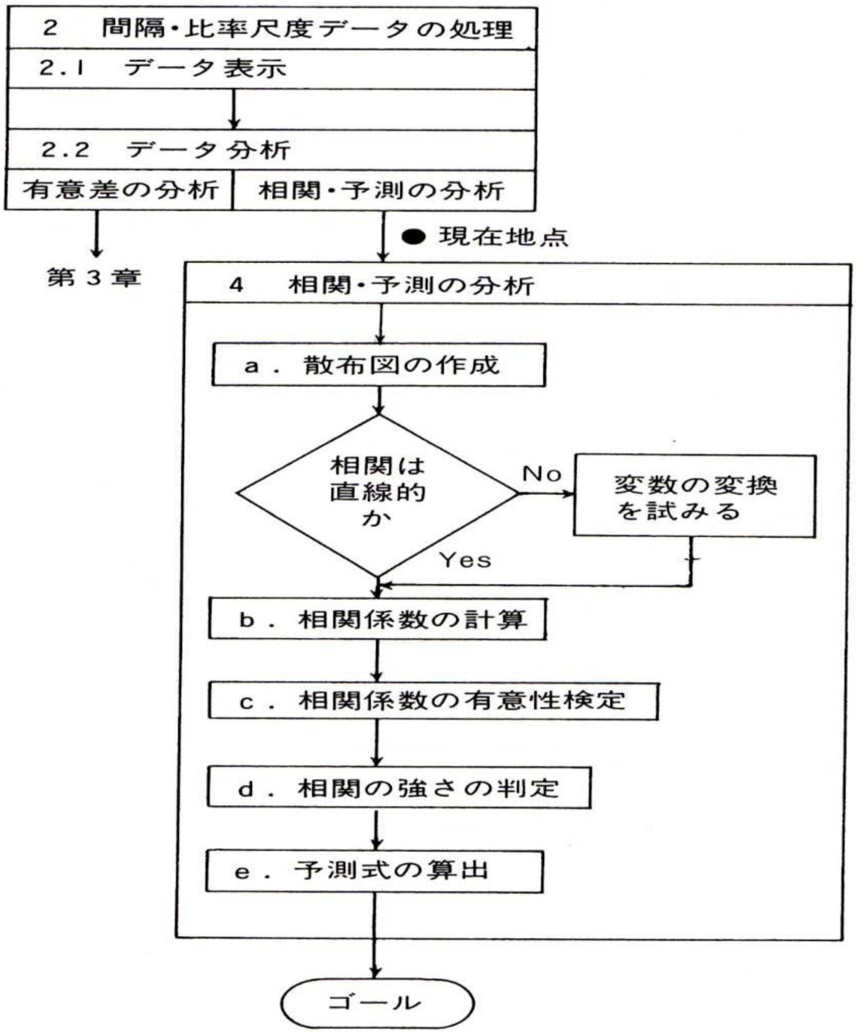
予測は回帰ともいわれる。回帰（regression）とは「もどす」ことであり、予測関係を逆方向から見たときの言い方である。（下表参照）

用語	見る方向	言い方・読み方の例
予測	練習本数→成功率	練習本数から成功率を予測する。
回帰	練習本数←成功率	練習本数から成功率を回帰させる。

このように、予測も回帰も、言っていることは同じであり、見る方向が異なるだけである。ここでは、「予測」という言い方を多用するが、専門的には「回帰」という言い方が好まれる。

5．1　分析の手順

相関・回帰の分析は、以下のフローチャートにおけるa〜eの5ステップを踏む。



a. <u>散布図の作成</u>

<u>散布図を描く目的</u>

散布図とは、対応する 2 変数の数値をタテ軸・ヨコ軸の座標とみなして、データを平面上の点として図示したものである。例えば、表 73 のような変数 A と変数 B のデータ・リストからは、図 10 のような散布図を描くことができる。

表 73　データ・リスト

被験者	変数A	変数B
イ	4	5
ロ	7	9
ハ	2	2
ニ	5	4
ホ	8	6

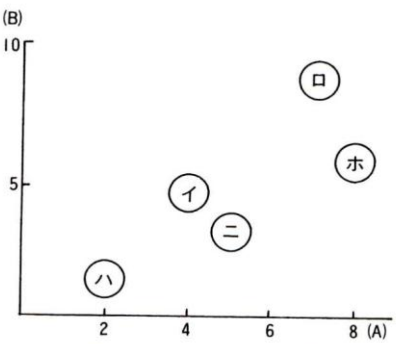


図10　変数Aと変数Bの散布図

散布図を描く目的は、変数間の相関・予測が直線的であるかどうかをチェックすることである。なぜなら、相関・予測の分析は直線相関と一次方程式の予測を前提としているからである。図 10 の例のように、散布図上のデータが一本の直線を想定して寄り集まる傾向にあるなら、相関・予測は直線的であると言える。なお、3 変数以上の散布図は立体的となるが、2 変数ずつ平面的に描くようにすればよい。この場合も、すべての散

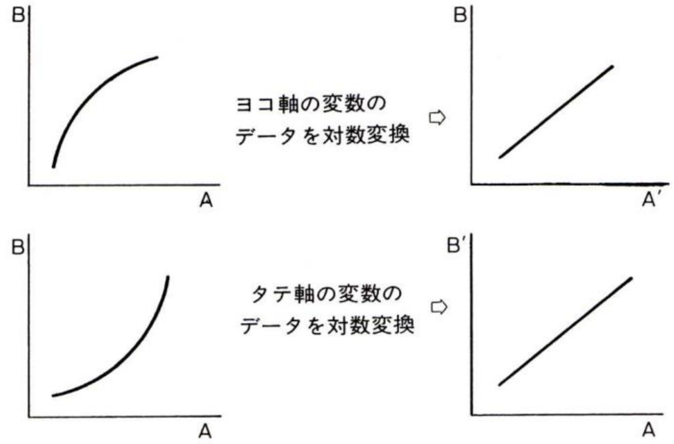
布図に直線性が見出せることが相関･予測の分析をおこなう前提となる。直線性が見出せない場合の対処を、次に述べる。

相関・予測が直線的でない場合の対処

相関・予測が直線的でなく、曲線相関・曲線予測が想定される場合、以下の3つの対処の方法がある。

① 数値を変換する。

数値の変換によって曲線相関を直線相関に変える。ただし、どのような数値の変換を用いたらよいかは試行錯誤的に決めるしかない。下例は、2変数の一方の変数を対数変換した例である。



この対処をおこなったら、変換後の数値をもって分析を進める。

② 変数をいくつかに分割して、複数の直線を当てはめてゆく。分割は、やはり試行錯誤的である。あまりおこなわれない。

③ 相関比を用いる。

① 、②は、曲線相関をなんとか直線相関に持ち込もうとする対処であったが、曲線相関を曲線のまま扱うとしたら「相関比」がある。ただし、利用のしかたは難しい。

b. 相関係数の計算

定義

変数間の相関・予測が直線的であることをチェックしたら、次に相関係数を計算する。相関係数とは、2 変数の相関の強さと方向を表す統計量である（3 変数以上の場合には重相関係数がある)。この統計量は K.Pearson の考案によるので、特に「ピアソンの相関係数」と呼ぶこともある。

記号

相関係数の記号は“ｒ”である。なお、ｒはデータの相関係数である。このデータの背後に存在する無限データ集団（母集団）の相関係数を表すときには、ｒに相当するギリシャ文字“ρ”を用いる。

ｒの性質

ｒの値は0をはさんで、－1～＋1の範囲で変化する。ｒの値が－の場合は、「負の相関」であり、散布図上のデータは右下がりの直線に収束する傾向を示す。これに対して、ｒの値が＋の場合は「正の相関」であり、散布図上のデータは右上がりの直線に収束する傾向を示す。完全相関（ｒ＝－1、ｒ＝＋1）の場合は、全データが一本の直線上に乗るが、それ以外は、直線への収束には程度の違いがある。なお、ｒ＝0 は「無相関」であり、データはどのような直線へ収束するきざしもみせない。

相関係数の計算

変数Ａと変数Ｂのデータを、下表のように入手したとする。

表 74　データ・リスト

対象者	変数Ａ	変数Ｂ
Y.Y.	1	6
T.S.	2	4
E.K.	3	5
O.M.	4	3
N.K.	5	1

表 75　データ表示(N=5)

	変数Ａ	変数Ｂ
X	3.0	3.8
SD	1.4	1.7

① 偏差積和を計算する。

偏差とは、データと平均との差のことである。偏差は各対象者ごとに2個できるので、それを掛け合わせて（積)、全被験者について合計した値が「偏差積和」となる。

偏差積和＝（1−3.0）×（6−3.8）＋（2−3.0）×（4−3.8）＋（3−3.0）×（5−3.8）＋（4−3.0）×（3−3.8）＋（5−3.0）×（1−3.8）

＝（−4.4）＋（−0.2）＋（0）＋（−0.8）＋（−5.6）

＝−11.0

ここでは、たまたまマイナスの積が多く出たので、和もマイナスになった。この「偏差の積」のイメージとしては各データと平均とのズレを面積として表現したものと考えることができる。

② データ１組分の偏差積和に変換する。

上で計算した偏差積和は、変数Ａと変数Ｂの対応するデータ５組分の値である。これを１組分の値にする。

データ１組分の偏差積和＝偏差積和/N＝−11.0/5＝−2.2

③ 下式によってｒを計算する。

ｒ＝（データ１組分の偏差積和）/（SD_A×SD_B）＝−2.2/（1.4×1.7）＝−.92

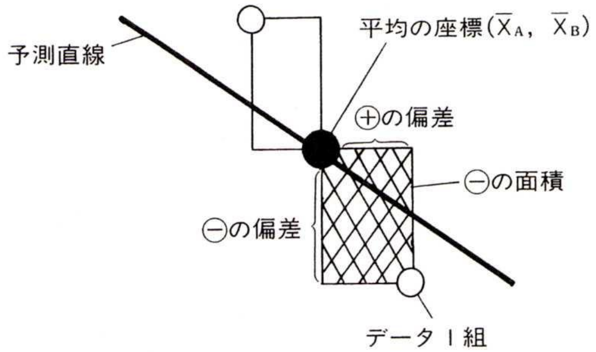
ここで、相関係数ｒ＝−.92は二つの情報をもっている。一つは値の大きさ（.92）が意味することであり、もう一つは値の符号（マイナス）が意味することである。以下に、順番に述べる。

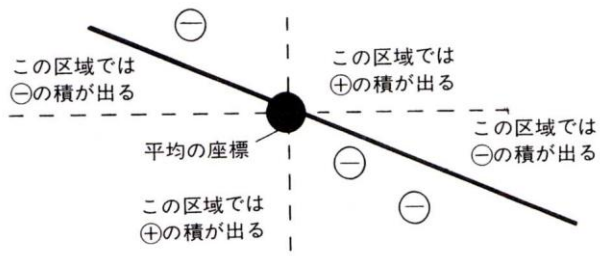
ｒの値の大きさが意味すること

上のｒの計算式における分子の「データ１組分の偏差積和」とは、変数Ａと変数Ｂのデータ１組分の偏差で出来る平均的面積であった。これに対して、分母の「SD_A×SD_B」は各変数の標準偏差を組みとなしたときに出来る面積である。したがって、データの各組の値のとり方が「SD_A：SD_B」と同比率であれば、分子は分母に収束し、「ｒ＝±1」となる。このとき、散布図上のデータの点は完全に直線上に並ぶ。このように変数Ａと変数Ｂの標準的な座標（SD_A、SD_B）を仮定して、実際のデータの組みがそれにどの程度の近似した値のとり方をするか、という程度を面積比とした値がｒである。ｒ＝−.92 とは完全な直線に対する９割程度の近似を意味している。

ｒの値の符号が意味すること

次に、ｒ＝−.92 はマイナスに出ているが、これは偏差積和がマイナスになっていたからである。すなわち、マイナスの積を出すデータの組が多かった（①においてデータ５組のうち４組がマイナスの積を与えた）。このように、マイナスの積が多く出るということは、平均の座標からみて、その左上と右下にデータが集まっていることを意味している。





右図からわかるように、ｒの符号がマイナスの場合は逆比例の関数直線、プラスの場合は正比例の関数直線に収束する傾向がある。分析上の用語では、それぞれ「負の相関」、「正の相関」と呼ぶ。かくして、ｒ＝−.92 のマイナスの符号は、散布図において左上から右下へ向かう直線に収束する様な相関（負の相関）を意味している。

以上、ｒ＝−.92から、変数Ａと変数Ｂは「完全な（または標準的な）関数直線に９割程度近似した負の相関関係」にある、ということを読み取ることができる。

ただし、この相関は本物の相関でなく、偶然によって生じた見せ掛けの相関であるかもしれない。この判定を、次に行わなければならない。

ｃ．相関関係の有意性検定

検定の目的

前の項目で求めた相関係数（−.92）のように、実際にはｒの値は±1（完全な正負の相関）や０（完全な無相関）になることはなく、たいてい中間的な値をとる。こうした中間的な値は、多かれ少なかれ偶然によって生じる可能性を含んでいる。すなわち、実質的に無相関であっても、偶然の数値のユレとしてｒが多少の値をもつこともある。そこで、データから計算したｒの値が、そのよう

に偶然に出現した値にすぎないのか否かを判定しなければならない。この判定を相関係数の有意性検定という。

偶然分布の用意

ここで、「真の相関係数は0である」（帰無仮説）としたときの偶然分布が必要になる。そして、その偶然分布中の出現確率が 5%未満しかないような「めったに出現しない」値であれば、そのときのｒは偶然に生じたものではないと考えることにする（有意であると認める)。これには、F分布を利用する。F分布は、分散分析におけるF比の検定の際にも偶然分布として用いたが、もともと、そのように汎用性のある分布である。ただし、ｒの値をF分布にあてはめるためには、下式によってｒをF比に換算しなければならない（下式のNはデータの組数)。

$$F = r^2 / (1 - r^2) / (N - 2)$$

検定の手続き

前例のｒ＝－.92 の有意性検定をおこなってみる。

1） ｒとNの値を用意して、F比を計算する。

ｒ＝－.92、N=5 をF比の計算式に代入する。

$F = r^2 / (1 - r^2) / (N - 2) = 0.8464/0.1536/3 = 16.53$

2）F比の分子・分母の自由度を計算する。F比をF分布に対照するには、F比の分子・分母の自由度が必要であった。ここで、求めておく。

分子の自由度；df①＝変数の個数－1＝2－1＝1

分母の自由度；df②＝N－変数の個数＝5－2＝3

3）F比をF分布表に対照し、出現確率を求める。

F=16.53 を、df①＝1、df②＝3 のF分布に対照する。

表 76　F分布表

df①	df②	F比		
1	3	5.54	10.13← ・	→34.12
出現確率		.10	.05	.01
有意水準		有意傾向	5%	1%

F＝16.53 を数列に対照していく。落ちた所の下段をみてｐ＜.05 と読む。

ここで、もしF比が有意でない場合は、次のステップへ進むことができない。

ｄ． 相関の強さの判定

相関の強さを判定する理由

さて、検定の結果は、有意であった（F（1,3）＝16.53、ｐ＜.05)。すなわち、ｒ＝－.92 は「真の相関係数0」＋「偶然のユレ」としてはめったに出現することのない値である。したがって、実質的に相関係数は0でないと認めてよいであろう。しかし、この有意性は、その相関が強い相関であることまで保証しない。有意であっても、取るに足らない弱い相関もある。そこで、有意性検定において有意であった場合は、次の相関の強さを判定する必要がある。

判定の基準

相関の強さの判定は経験的であり、一般に、以下の基準を用いている。

負の相関	相関の強さの判定	正の相関
－1		＋1
	強い相関がある	
－.7		＋.7
	中程度の相関がある	
－.4		＋.4
	弱い相関がある	
－.2		＋.2
	ほとんど相関がない	
0		0

注意：相関係数の値と相関の強さの関係

先の判定基準からわかるように、相関係数の値と相関の強さとは直線的関係にない。例えば、「弱い相関」の段階は 0.2～0.4 の2ポイントの幅であるが、「中程度の相関」の段階は 0.4～0.7 の3ポイントの幅にとってある。したがって、相関係数の値が二倍であるからといって、相関の強さも二倍であると言えない。この点、相関の強さを正確に理解したいなら、説明率または決定係数（coefficient of determination）といわれる指数を計算しておくほうが、誤ったイメージをもたないですむ。

説明率の計算

説明率とは、一方の変数が他方の変数をどのていど説明するかを表す統計量である。この説明率と相関の強さとは同義である。専門用語では「決定係数」と呼ぶこともある。説明率の詳細については次の「予測式の算出」で述べる。

説明率の計算は、下式による。

$$説明率（\%）= r^2 \times 100$$

例えば、r＝.35 の説明率は.35²×100＝12%である。r＝.35 はr＝.70 と比べると相関係数としては二分の一であるが、説明率を計算してみると、実は相関の強さとしてはもっと弱く、四分の一以下であることがわかる（r＝.70 の説明率は 49%)。

ちなみに、先の判定基準を説明率に置き換えてみると、強い相関は 50%以上、中程度の相関は 15~50%、弱い相関は 5~15%、ほとんど無相関は 5%以下となる。

e. 予測式の算出

分析の目的

予測の分析の目的は、一方の変数の値から他方の変数の値を予測する方程式を求めることである。この方程式を予測式または回帰式という。予測と回帰のちがいは、たんに見方の違いである。例えば、「変数Ａを変数Ｂから予測する」は「変数Ａを変数Ｂに回帰させる」ことでもある。以下「予測」に統一するが、「回帰」に読み替えてもよい。

予測式の説明

さて、変数Ａを変数Ｂから予測する数式は、一般に、次のような形になる。

A`＝k・B＋y

この予測式は、散布図上では、データが収束すべき直線を表している（この直線を予測直線という)。式中の A` は、変数Ａの予測された値であり、必ずしも実際の値と一致しないので、A とは区別して表している。k はいわゆる「傾き」である。k が＋なら予測直線は右上がり、k がーならば予測直線は右下がりになる。専門用語では、「回帰係数」と呼び、変数Ａが変数Ｂに依存する(回帰する)程度とみている。なぜなら、k の値が大きければ大きいほど、変数Ａの予測値（A`）は変数Ｂが１ポイント上下するたびに大きく変化するからである。また、このため k を「重み」と表現することもある。重みが大きいことは、変数Ａに及ぼす変数Ｂの影響が重大であることを意味している。予測式の最後の y は定数であり、いわゆる「y 切片」である。すなわち、予測直線がタテ軸を切る点を表している。

注意；予測の方向の問題

変数Ａと変数Ｂから予測する場合と、その反対に変数Ｂを変数Ａから予測する場合とでは、予測式の k、y の値が異なる。したがって、予測式を求めるときは、予測の方向をあらかじめ決めておくことが前提である。

この点、統計学では、予測される側の変数と予測する側の変数を別々の名称を用いて、言い分けている。

予測される側　←　予測する側

目的変数			予測変数		
従属変数	＝	k×	独立変数	＋	y
基準変数			説明変数		

ここでは、一番上の「目的変数」・「予測変数」という言い方をとる。すなわち、前の例では変数Ａが目的変数、変数Ｂが予測変数である。

予測式の算出例

下表のような変数Ａと変数Ｂのデータがあるとする。

表 77　変数Ａ、Ｂの平均と標準偏差（N＝20)

	変数Ａ	変数Ｂ
X	12.5	13.7
SD	6.9	5.5

このとき、変数Ａを目的変数、変数Ｂを予測変数とする予測式を求める。

1）相関係数を計算しておく。

すでに計算してあり（r＝.837)、検定の結果、有意であるとする。

2）下式によってkを計算する。

$k = r \times (SD_A/SD_B) = .837 \times (6.9/5.5) = 1.05$

3）下式によってyを計算する。

$y = X_A - k \cdot X_B = 12.5 - 1.05 \times 13.7 = -1.89$

4）以上のkとyの値を下式に代入して、予測式を完成させる。

$A` = k \cdot B + y$

$A` = 1.05B - 1.89$

<u>相関・予測の分析の論文記載例</u>

「表77は、変数Aと変数Bの平均および標準偏差を示したものである。また、図11（省略）は、その散布図である。相関係数は.837であり、有意であった（$F(1, 18) = 42.11$、$p < .01$）。説明率は70.1%であり、両変数の間には強い相関があると言える。なお、変数Aを目的変数、変数Bを予測変数として、予測式を求めると、$A` = 1.05B - 1.89$となった。予測の標準誤差は3.7である。」

第1文と第2文が「データ表示」、第3文以降が「データ分析」である。相関・予測の分析の場合、データ表示は、各変数の統計量N、X、SDと散布図を掲載する。ただし、散布図はスペースをとるので、たんに相関・予測の直線性をチェックするだけなら省略してもよい。

分析結果の記述は、相関係数→その有意性検定→説明率→相関の強さの判定、と進む。途中の有意性検定は「両側検定」であるが、この場合は付記する必要がない。F分布を利用したことが、そのまま両側検定であることを示している。なお、相関係数の両側検定とは、相関係数が＋に出ても、－に出ても対処できる検定のことである。

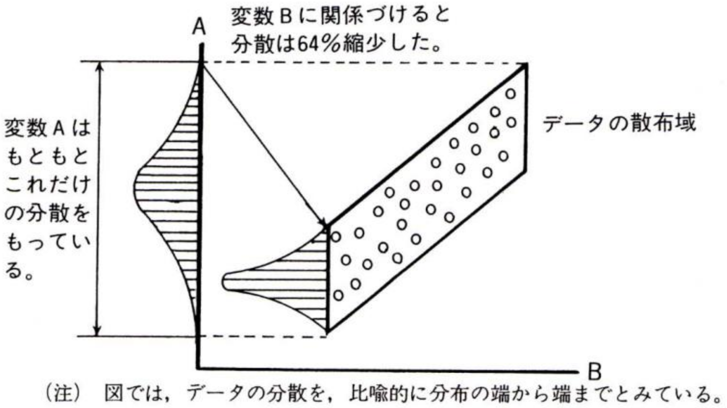
第2段落は予測式の算出である。一応、予測式を求めたので掲載した。ここで、「予測の標準誤差」とは、予測式の精度を表す指標である。すなわち、この予測式の精度は、$A` \pm 3.7$の範囲に実際値の68%がおさまる程度であることを意味している。もちろん、実用に耐えられる予測式は$N = 20$程度では無理である。

<u>備考1；相関の説明率と予測の誤差について</u>

相関の説明率は、予測式の的中率と等しい。反対に言えば、相関の説明率の残り分が、予測式の誤差である。例えば、$r = .80$の説明率は64%であるが、残り36%が説明できない率（すなわち予測がはずれる率）となる。ここで、説明率の意味は、目的変数Aのデータの分散が、予測変数Bとの間の相関関係によって64%縮小した、ということである（下図参照）。

右図からわかるように、変数Aの総分散のうち、64%は変数Bとの相関によって生じた規則的変化である。あとの残りの36%は規則性を持たない変化であり、偶然の影響による誤差とみなすしかない。このことを等式によって表現すると次のようになる。

目的変数の分散＝予測変数との相関によって説明される分＋残った分（偶然誤差）



（注） 図では，データの分散を，比喩的に分布の端から端までとみている。

さらに、上の等式を数式化すると次のようになる。なお、SD_Aは目的変数Aの標準偏差である。

$$SD_A^2 = SD_A^2 \cdot r^2 + SD_A^2 \cdot (1 - r^2)$$

前例をあてはめれば、上式のr^2が64%、$(1 - r^2)$が36%に相当する。すなわち、説明分と誤差分である。この誤差分の$SD_A^2 \cdot (1 - r^2)$の分散値がデータ本来の寸法に合わせて√した値、すなわち$SD_A \cdot \sqrt{(1 - r^2)}$を**<u>予測の標準誤差</u>**（standard error of estimate）と呼ぶ。予測の標準誤差の意味は、平均に対する標準偏差と等しい。すなわち、予測値±予測の標準誤差の範囲に、目的変数の実際値の約68%が収まることを意味している。したがって、予測の標準誤差が小さいほど、予測直線への収束の程度が大きく、それだけ予測の精度は高いと言える。

<u>備考2；相関・予測の分析をおこなえない場合</u>

次の場合は、相関係数や予測式を求めることができない。すなわち、各変数のデータの分布が正規分布でない場合、および、目的変数が予測変数に対して等散布性を示さない場合、である。等散布性（homoscedasty）とは、

予測直性上のどの点においても、「予測値±予測の標準誤差」の範囲に目的変数の実際値の約 68%が均等に集まっていることである。ただし、これらの制約は「みなし」がきけばクリアできる。処理を誤って結論を誤るのもよくないが、処理の誤りを恐れて結論を見送るのもつまらない。研究は一回で終わると考えないことである。誤った結論は批判を受けて消え、有望な結論は確証されて残る。

備考３；相関係数の有意性検定とＦ分布

相関係数の有意性検定に、Ｆ分布を利用した経緯を説明する。Ｆ比は２つの分散を分子・分母にとったときの比であったが、相関係数の有意性検定においてはｒからＦ比を計算した。

$$F = r^2 / (変数の個数-1) \; / \; (1-r^2) \; / \; (N-変数の個数)$$

上式は、一般式であり、前述したＦ比の計算は「変数の個数」を２とした場合の特殊式である。ここで、分子のr^2は相関の強さを表す説明率である。これに対して、分母の $1-r^2$は相関によって説明できない偶然誤差の率である。それぞれに目的変数の分散（SD_A^2）をかけてみれば、この式の分子・分母は「相関によって説明される分散」と「偶然誤差による分散」となって、説明分散が誤差分散の何倍あるかを示す分散比になる（SD_Aをかけなくても比の値は変わらない）。したがって、上の計算式の値は、Ｆ分布にしたがう。ただし、上式分子のr^2は「変数の個数－1」によって、また、分母の $1-r^2$は「データの組数－変数の個数」によってそれぞれ値の出方を左右される。そこで、分子・分母を「自由度１個分の値」に換算して公平な比較を期しているのである。もし、相関による説明分が偶然誤差のユレより大きければ、分子は分母より大きく、Ｆ比は１以上の値を示す。さらに、それが偶然には「めったに出現しない」（出現確率 5%未満の）大きさであれば、実質的な相関関係があるということになる。

相関・予測の分析における注意事項

１）相関関係・予測関係は因果関係ではない。

例えば、学習意欲の得点と教科のテスト得点との間に有意な正の相関があったとしても、「学習意欲が教科の成績を向上させる」とは言えない。また、その逆も言えない。なぜなら、相関は、たんに変数間の規則的な対応を示しているにすぎないからである。変数間の因果関係は、別途、主張すべきものである（理論的モデルによって、あるいは、実験計画法に基づく研究によって）。

２）第３の変数による見かけの相関に注意せよ。

変数Ａと変数Ｂが、第３の変数Ｃになんらかの影響を被るときは、ＡとＢの相関は「見かけの相関」であり、額面どおり解釈するわけにはいかない。対処は、重相関の分析に移る。

３）相関の強さの判定は研究目的によって異なる。

相関の強さの判定については、先に経験的な基準を掲載したが、基本的に、相関の強さは研究目的に応じて判定するべきである。例えば、知能テストの信頼性（安定した結果が得られること）の確認のため、同一の被験者に対する一か月前と一か月後の得点の相関をとるような場合（再テスト法）、ｒ＝.7 程度であっても弱い相関とみるべきである。これに対して、２種類の性格特性の関連を問題とするような場合、それらの性格テストの得点の相関がｒ＝.4 程度であっても両特性間には相当の関連があり、むしろ、高すぎる相関は同一のものを測っているにすぎないと解釈するべきであろう。

相関・予測の分析例：散布図の作成を怠った例

事例

論文名「読みと内的発声の関係」

目的

読みにおける内的発声が、日常のスピーチの内的発声と類似の過程であるかどうかを検討する。

方法

被験者：大学生５人。（少なすぎるが解説の便宜上、ここではこのようにする）

材料：簡単な和文の文章と英文の文章。

手続き：まず、くつろいだ雰囲気で被験者と５分間の雑談をおこない、その会話を録音する。次に、被験者に文章材料を提示し、２分間黙読させる。この試行を和文と英文について繰り返す。

結果

和文・英文の読みの速度として、２分間に読んだ単語の音節数をカウントした。また、スピーチの速度として、

雑談の録音資料から話題転換のない10秒以上の発声部分を取り出し、生産された音節数をカウントした。そして、これらの読みとスピーチの音節数を10秒当たりの個数に換算した。表78（省略）は、通常のスピーチの速度と和文・英文の読みの速度の平均及び標準偏差を示したものである。相関係数を計算した結果、スピーチ×和文の読みは、.803（F（1,3）＝5.45、p ＞.10）。スピーチ×英文の読みは－.498（F＜1）であり、いずれも有意でなかった。したがって、和文・英文を問わず、読みにおける内的発声は、日常のスピーチの発声と類似した過程であるとは言えないであろう。

解説

「スピーチ×和文の読み」も「スピーチ×英文の読み」も有意でなかった。しかし、データ・リストを入手して散布図を描いてみると、実は前者の相関は曲線相関であったことが分かる。

表79　データ・リスト（音節数/10秒）

被験者	スピーチ	和文
イ	25.5	30.8
ロ	46.3	39.1
ハ	26.3	35.4
ニ	31.6	36.9
ホ	18.5	19.5

そこで、スピーチの速度を対数変換して、和文の読みの速度との相関係数を再計算してみた。その結果、r ＝.881となり、これは、はっきりと有意である（F（1,3）＝10.40、p ＜.05）。ちなみに、和文の読みの速度を目的変数（J）、スピーチの速度を予測変数（S）としたときの予測式は「$J' = 47.0 \times \log_{10} S - 35.9$」である。かくして、和文の読みは、日常のスピーチと相関した速さで進行する類似の発声過程を含んでいる可能性が示唆される。本例は、散布図の作成を怠り、相関の直線性をチェックしなかった。そのため、分析の結果に妥当性を欠き、結論を誤った。なお、曲線相関への対処として、変数の変換は試行錯誤的である。この例では、スピーチの速度を対数変換してうまくいったが、もし、ノの字形の曲線相関ならタテ軸の変数を対数変換するのがよい。

3変数以上の相関・予測の分析

これまで、2変数の相関を分析してきた。今度は、もっと変数の多い場合、すなわち3変数以上の相関を分析してみる。

コースの選択

3変数以上の相関を分析するには、以下の2つのコースがある。

コースⅠ：「相関マトリスクの作成」→　「因子分析」

このコースは、変数間に特定の**予測関係が存在しない場合**に選択する。基本的には、2変数ずつを取り出して、一組ずつ相関係数を計算する（相関マトリスクの作成）。また、その結果に基づいて、多くの変数を少数の因子にまとめる「因子分析」へ発展することもできる。

コースⅡ：「重相関係数の計算」→　「回帰分析」

このコースは、変数間に特定の**予測関係が存在する場合**に選択する。すなわち、全変数のうち、ある1個が目的変数、あとの残りがすべて予測変数と最初から決まっていなければならない。このコースを選択すると、まず「重相関係数」を計算し、次に予測式を算出する。ただし、この予測式の算出は「回帰分析」として行われるのが一般的である。すなわち、目的変数の予測に貢献する予測変数と貢献しない予測変数を、複数の候補の中から選択しようとする。

５．２　３変数以上の分析Ⅰ：相関マトリスクの作成

予定

このコースは、まず「相関マトリスク」を作成し、次に「因子分析」へ移行することができる。ただし、因子分析への移行の際には、変数の個数が10個以上（できるなら20個以上）は、ほしい。まず、相関マトリスクか

ら説明する。

相関マトリスクの作成

相関マトリスクとは「相関行列」という意味であり、多くの相関係数をヨコ（行）とタテ（列）に並べた表である。相関マトリスクを作成するには、いわば全変数の「総当たり表」を作って、この表のなかで「ぶつかる」2 変数どうしの相関係数を一つずつ計算してゆけばよい。例えば、下の表 80 は 4 変数の相関マトリスクの例である。

表 80　4 変数の相関マトリスク（N＝42）

	B	C	D
変数A	.432**	.610**	.279†
変数B		.214	.186
変数C			.378*
†p<.10 *p<.05 **p<.01（両側）			

表中の各相関係数に対しては、個別に有意性検定（両側検定）を行う。そして、その結果を記号を付けて表すこと。

因子分析の事例：主因子法とバリマスク回転

用途

因子分析（Factor Analysis）は、多くの変数の相関関係を、少数の因子に要約するための方法である。ただし、そのような多変数の動きを支配する共通の因子がみつかるかどうかはケース・バイ・ケースであり、探索的である。

実用上は、因子分析を手計算することはなく、ほとんど大型計算機のプログラムを用いる。ここでは、その出力結果を読む時のポイントを簡単に解説する。以下に上げる事例は、ＳＡＳの因子分析プログラムを使用した研究例である。

事例

論文名「青年期の友人選択に関する因子分析的研究」

目的　青年期の友人選択が、どのような因子に規定されているかをあきらかにする。

方法

対象者：大学生 90 人（男子 48 人、女子 42 人）

道具：7 項目の質問紙（表 81 参照）

手続き：調査は集団で行った。記入は匿名として、特定の人物についての記述を提示した。回答は「ぜんぜん友人にしたくない」～「ひじょうに友人にしたい」の 6 段階であり、1 点から 6 点までの得点化をおこなって、提示した人物ごとに集計した。

表 81　質問紙と 7 つの人物記述

次のような人をどの程度友人にしたいかを判断し、下のあてはまる番号に○を付けてください。

人物 1.　いろいろな相談にのってくれる人。

1　2　3　4　5　6

ぜんぜん友人にしたくない |　|　|　|　|　| ひじょうに友人にしたい

以下、同形式で人物 2～人物 7 を提示した。文章のみを示す。

人物 2.　自分の誤りを指摘してくれる人。

人物 3.　落ち込んだとき元気づけてくれる人。

人物 4.　一緒にいると楽しい人。

人物 5.　自分にないものを持っている人。

人物 6.　なにかに熱中している人。

人物 7.　みんなから信頼されている人。

結果

表 82 は、各人物に対する得点の平均と標準偏差、および、人物間の得点の相関係数を示したものである。得

点が大きいほど友人にしたいと思われる人物であることを示す。

表 82　各人物の得点と人物間の得点の相関係数（N＝90）

	人物 1	人物 2	人物 3	人物 4	人物 5	人物 6	人物 7
X	4.34	4.20	3.25	3.71	3.15	3.36	3.45
ＳＤ	0.97	1.16	1.13	1.13	1.07	1.00	1.08
人物 2	r=0.56**						
人物 3	0.23*	0.39**					
人物 4	0.18†	0.28**	0.12				
人物 5	0.11	0.31**	0.16†	0.40**			
人物 6	-0.14	0.04	-0.11	0.23*	0.59**		
人物 7	-0.12	0.08	-0.04	0.22*	0.56**	0.62**	
†p<.10　*p<.05　**p<.01（両側）							

因子分析（主因子法、バリマスク回転）の結果、表 83 に示したように 2 つの因子を抽出した。

表 83　7 人物の得点の因子分析の結果

項目	因子負荷量		共通性
	第一因子	第二因子	
人物 6	0.853	-0.162	0.754
人物 7	0.836	-0.111	0.710
人物 5	0.826	0.262	0.751
人物 2	0.174	0.836	0.729
人物 1	-0.098	0.785	0.626
人物 3	-0.023	0.632	0.400
人物 4	0.475	0.394	0.381
寄与	2.374	1.977	

このうち、第Ⅰ因子は、人物 6（何かに熱中）・人物 7（みんなから信頼）・人物 5（自分にないものを持っている）の負荷が大きいことから、人物の特徴や力量などに基準をおいて友人を選択する程度を表わしているとみられる。したがって、「力量基準因子」と命名する。

これに対して、第Ⅱ因子は、人物 2（誤りを指摘してくれる）・人物 1（相談にのってくれる）・人物 3（元気づけてくれる）の負荷が大きいことから、人物から自分への関わりに基準をおいて友人を選択する程度を表わしているとみられる。したがって、「関係基準因子」と命名する。

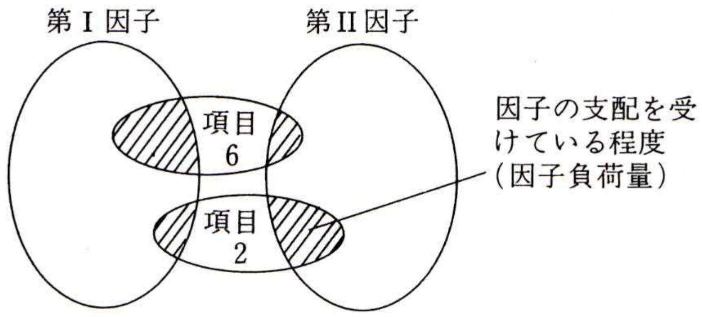
このように、青年期の友人選択は、人物のもつ力量および人物と自分との関係によって規定されていると考えられる。

解説

因子分析にはいくつかの方法があるが、本例の「主因子法」が主流である。主因子法によって解を求めた後、バリマスク回転といわれる因子軸の回転をおこなっている。なお、抽出する因子数については多くの打ち切り基準があるが、SAS ではオプションなしなら自動的に固有値（因子の説明率）1 以上の因子を抽出する。

さて、因子分析の結果を読むときの最大のポイントは、抽出された因子の解釈である。これには、表 83 における各項目（人物）の因子負荷量（factor loading）を参考にする。因子負荷量とは、その項目が因子に支配される程度のことで、相関係数と同様、値が大きければ大きいほど強く支配を被っていることになる（正の支配、負の支配）。そこで、因子負荷量の大きい項目に共通する要素を推理して、その因子が何を支配する因子なのかを解釈するのである（下図の因子負荷量のイメージ参照）。

因子解釈の最後に、因子に名前を付ける。本事例では、第Ⅰ因子が「力量基準因子」、第Ⅱ因子が「関係基準因子」と命名されたが、これによって因子が多変数の相関の要約らしくなる。なお、人物４は因子解釈の参考にしなかったが、両因子に対して負荷が均等であるためと共通性の低いためである。



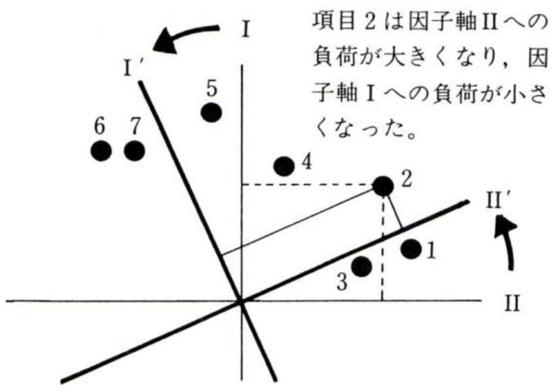
備考①：因子の回転

因子軸の回転は、因子の意味を解釈しやすくするための計算方法である。事例では「主因子法」によって２因子を抽出した後、「バリマスク回転」をおこなっている。これによって、各項目（人物）の因子負荷量はどちらか一方の因子に対して大きくなるので、因子の解釈が容易になる。下表 84 において、回転前と回転後の因子負荷量を比較してください。

表 84　因子軸の回転前と回転後の因子負荷量

項目	回転前		回転後	
	第Ⅰ因子	第Ⅱ因子	第Ⅰ因子	第Ⅱ因子
人物 6	0.710	-0.446	0.853	-0.162
人物 7	0.715	-0.446	0.836	-0.111
人物 5	0.860	-0.103	0.826	0.262
人物 2	0.504	0.690	0.174	0.836
人物 1	0.235	0.756	-0.098	0.785
人物 3	0.240	0.585	-0.023	0.632
人物 4	0.595	0.163	0.475	0.394

なお、下図に因子軸の回転のイメージを示しておく。

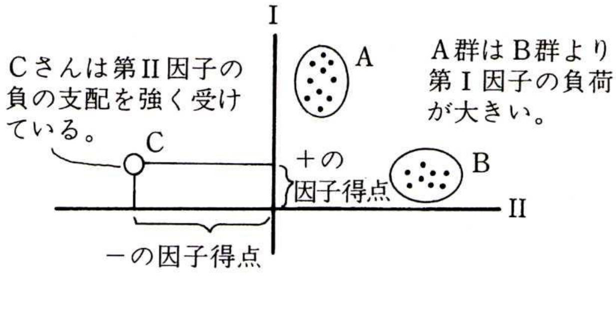


備考②：共通性について

ある項目のデータは、因子によって共通に規定される部分（共通性 communality）と、そのような規定をうけない独自の部分（独自性 uniqueness）とをもっている。後者は誤差といってもよい。この共通性は各因子負荷量の二乗値を合計したものである。例えば、人物６のデータにおける共通性は、第Ⅰ因子の負荷量の二乗値 $(0.853)^2$ ＋第Ⅱ因子の負荷量の二乗値 $(-0.162)^2 = 0.754$ となる。事例では、因子の解釈に人物４を参考にしなかったが、共通性が低く、因子の特徴を反映している部分が小さかったためである。すなわち、特定の因子に要約するには独自性が大きすぎるのである。なお、項目数が多い場合は、そのような共通性の小さい項目を除いて、再度、因子分析を行うこともある。

備考③：因子得点の利用

因子分析の結果の応用として、個々の被験者のデータに対する各因子の規定分を計算することができる（手計算では無理であるが）。これを因子得点（factor score）と呼ぶ。因子得点のイメージは、個々の被験者を因子という「ものさし」の上に並べたときの位置である。したがって、右図のように、因子平面または因子空間のなかに個々の被験者を座標として置いてみることができる。



また、因子得点を従属変数として、例えば対象者を男女に分けてｔ検定や分散分析を適用してみることもできる（下表 85 を参照)。

表 85　友人選択における男女差の検討

	力量基準因子		関係基準因子	
	男	女	男	女
N	48	42	48	42
X	−0.12	0.14	−0.05	0.16
SD	1.10	0.87	0.97	1.03

このような有意差の分析は、例えば力量基準因子の得点の場合、その因子への負荷の大きい人物 6・人物 7・人物 5 への各回答をデータとした計 3 回の男女差の検定を一括して行うような意味がある。表 85 の場合、女子のほうが友人選択の際に相手の力量や自分との関係を重視する傾向が強いと言える。

さらに、因子得点を相関・予測の分析に応用して、例えば同一の対象者から別途に内向性・外向性の得点を入手し、それらと力量基準因子・関係基準因子の得点との相関を求めることもできる。

因子分析の限界

因子分析の使用に当たって、もっとも注意しなければならないことは、因子分析の解は一つでなく、無数に存在するということである。このため、いくつかの条件を持ちこんで解を得ているのである。したがって、抽出した因子を断定的に扱うことは避けなければならない。特に、因子の命名を行うと、そのような因子が実在するように考えがちであるが、一つの仮説の可能性を示唆するにすぎない。あくまでも、因子分析は要約的・探索的方法であり、結論に至るには今後の確証を予定しなければならない。

「生活科学のための多変量解析　第４章　因子分析より抜粋

１．因子分析とは

（1）因子分析の性質と特徴

因子分析は、主成分分析と同様、多数の測定項目からなるデータの情報を要約あるいは縮小する方法である。応用という点からすると主成分分析とよく似ているが、主成分分析が変数の変換あるいは新たな合成変数の作成を目的としているのに対して、因子分析は測定された項目に潜む潜在的構造を数学モデルとして次のように仮定している点で基本的に異なったものとなっている。

因子分析では、測定された変数のすべてに共通の要因（これを共通因子と呼ぶ）と、それぞれの変数にだけ特有の要因（独自因子と呼ぶ）に分解されるというモデルを仮定する。多変数の持つ情報を共通因子と独自因子という二つの成分に分解するわけである。

因子分析のモデル；変数の持つ情報＝共通因子＋独自因子

ところが、このようにして定義される潜在的因子が実際上、本当に測定できることはまれなことが多いので、この因子が示す意味については計算結果を「解釈」することによって意味づけられ、「因子名」が与えられている。解釈に当たっては共通因子として抽出された一種の「合成変数」とそれぞれの測定項目との相関係数の大小が用いられる。共通因子と測定項目との相関係数のことを因子負荷量と言っている。

共通因子に適当な解釈がなされ、因子に命名されると（例えば、「生活水準の因子」や「体の大きさの因子」など)、その解釈された因子の性質に基づいて、実際に測定することのできる具体的な要因を探索したり推察したりする手掛かりを得ることができる。あるいは共通因子とそれぞれの変数との相関の程度から、測定項目や個人、会社、地域など類別化をすることもできる。つまり、因子分析というのは、多変数の中から潜在的な共通因子を取り出し、それによって複雑なデータをより簡潔な構造に書き直す統計解析法であるといえる。では、因子分析はどのような場合に利用できるのであろうか。

①　観測変数が多く、これを要約したい時。

②　主要概念について、過去数年間、研究者間に一致が見られない時。

③　複雑な相互作用があり、実験的操作によっては、要因を分離し得ない場合。

④　多くの変数を組み合わせた総合的包括的変数を作り出したい時。
⑤　多くの変数から中心的変数と周辺的変数を分離したい時。
⑥　多くの変数を分類したい時。
⑦　多くの変数を要約して、より単純な変数を構成して、それによって個人や企業、地域などを評価したい時。
などに非常に便利な方法である。このほかにもこれを利用する人の工夫によって多方面に応用しうると言える。
一般的には、因子分析を使用目的に応じて次のように整理することができる。
（2）因子分析のいろいろ
仮説検証的因子分析
①　仮説的モデルの検証
探索的因子分析
②　一般的モデル（因子構造）の推定
③　データの要約的記述
④　観察対象の分類または類型化
確証的因子分析

①　仮説的モデルの検証のための因子分析
因子分析の創始者スピアマンは 1904 年に、知能の研究において「全ての精神能力には、知能と称すべきただ一つの“一般因子”が存在する」という仮説を提示した。スピアマンの考え方はまず最初に、二つのテストを考え、両者に相関があると仮定する。これは両者に共通な 1 個の因子と 2 個の特殊因子とがあるという意味だと仮定することができる。いま、二つのテストを a および b とし、共通因子を G、2 個の特殊因子を S1 および S2 で表わす。そうすると、テスト a および b は、共通要素 G についての 2 個の測度であって、それらが S1 および S2 という二つの残差を持つというものである。
スピアマンは、人間の精神能力における認識過程を含むテストにおいては、その共通要素はすべて同一だという遠大な結論を引き出そうという意図を持っていた。そこで彼は、全てのテストが、その全部に共通な一般因子（G）と、それぞれのテストだけに見いだされる特殊因子（S）とから成り立っていると考えた。すなわち、ある選ばれた集団の人々に、彼らの精神的能力を調べる P 種類のテストを施す時、それらテキストの得点を変数ｘ1、ｘ2、ｘ3、・・・・、ｘｐとするとそれらの変数は次式で表わされるような構造を持つと仮定したのである。これをスピアマンの 2 因子説（two-factor theory）と言う。
ｘ1＝a1G＋S1
ｘ2＝a2G＋S2
・・・・・・・
ｘp＝apG＋Sp
ここで、G は一般因子（general factor）、a1,a2,……,ap は因子負荷量（factor loading）、S1,S2,….,Sp は特殊因子（specific factor）と呼ばれる。このモデルでは、一般因子 G は潜在的（観測不可能という意味）ではあるが、その集団の個体に対して一定の分布に従う変動であると考えられ、各テスト（変数）ごとに定まる因子負荷量 ai を重みとして、各変数 xi に関与する。ｘi の変動のうち、この一般因子では説明されない残りの部分は、各テストに固有の特殊因子と称される変動で、一般因子およびほかの特殊因子とは無相関であると仮定される。ここで一般因子が特殊因子に対して何らかの基準より十分に大きければ仮説モデルが検証されたと考えるのである。
一般因子と特殊因子の図解；2 因子配列型は、図４－１（ｂ）に示されている。
「スピアマンの一般因子は、中央の大きな円 G で示され、特殊因子は、G の周囲に小さな円で描かれている。長円は、それぞれ一つの精神検査を表わす。G についての“負荷量”が、テストによってそれぞれ異なっていることを表わすために、長円と G との重なりの程度がいろいろに描かれている。任意の二つのテスト間の相関の大きさは、その二つのテストが G を含む程度によって決定される。それゆえ、テストａとｂとは、共に G を含んでいるので、比較的高い相関を示すだろう。またテストａとｃとは、ともに G についての負荷量が小さいので、ほとんど相関しないだろう」
しかし、分析の対象となる変量の間の相関は、一つだけの一般因子で説明し尽くされるようなことはめったに起こらないことがやがて明らかにされたので、二つ以上の共通因子を想定するようになった。これを多因子モデ

ル、その解析法を多因子分析（multi-factor analysis）と言う。その一つは、共通因子の数をあらかじめ固定せず、次に紹介するような一般的な（多因子分析）モデルを前提として、m 個の因子をデータから推定し、かつその潜在因子を抽出しようとするのである。

② 一般的モデル（因子構造）の推定のための因子分析

スピアマン以来、いくつかの因子分析法が提案されている。それらはすべて、相関行列または分散共分散行列から出発する。因子を抽出する手続きの中で主なものは、ホテリングの主成分分析ケリーの主軸法、バートの総和法、およびサーストンのセントロイド法（重心法）である。主成分分析と主軸法は多くの共通点を有し、総和法とセントロイド法も類似している。因子を抽出する手続きのほかに、サーストンは特に回転法を付け加えている。因子分析とはそもそもサーストンの名付けた言葉である。ここでは、一般的によく用いられるサーストンの方法を紹介する。

かれは次のような多因子モデルを考えた。

ｘ1＝a11f1＋a12f2＋・・・・・＋a1mfm＋u1

ｘ2＝a21f1＋a22f2＋・・・・・＋a2mfm＋u2

・・・・・・

ｘp＝ap1f1＋ap2f2＋・・・・・＋apmfm＋up

ここで、x1,x2,….,xp は観測特性である p 個の変数であり、f1,f2,…..,fm は m 個の（潜在）共通因子、u1,u2,….,up は特殊因子または独自因子と呼ばれる潜在的な変数である。係数 aik は、第 i 特性に対する k 共通因子の因子負荷量と呼ばれる定数である。

このモデルは、観測可能な p 個の変数が、その数 m も未知である共通因子の一次結合によって説明される共通部分と、各変数に固有な部分 ui とから構成されると仮定している。この式で、その説明変数に対応する個々の共通因子は観測不可能な仮説的な変数で、その数 m も未知である。もう少し分かりやすくいうと次のようになる。「知能」の測定という状況を想定してみる。

知能に関する測定値は、被験者がテストによって様々な値をとる変数である。この変数の値は 2 種類の源泉（共通因子と独自因子）を持っている。一般にその源泉を知能の（因子）という。この知能因子に由来する被験者の得点を知能因子得点という。この因子得点はテストと言うフィルターを通して、間接的に測定される。このフィルターを媒介することによって、被験者の因子得点は推定される。そしてこの因子得点が、被験者のあるテストの成績になる。因子得点を導出するフィルターとしてのテストの作用の程度を表わす程度を（因子負荷量）と言う。そして、当の被験者のあるテストの成績は（因子負荷量）と（因子得点）の積として把握される。すなわち、次のように表現できる。

（人間の知能に関する得点）＝（因子負荷量）×（因子得点）

ここに因子負荷量は被験者のいかんにかかわらず、テストごとに一定の値をとる定数である。また因子得点はテストのいかんにかかわらず、被験者ごとに一定の値をとる定数である。ここで、先に挙げた共通因子とは、被験者のどのテストの成績に対しても、共通してその変数の源泉となりうる因子のことである。例えば、人間の知能が、言語能力、数量的能力、記憶力、推理力、知能能力という五つの能力の合成として把握されるものとすると、テストがこのような性質の知能を測定するのにふさわしいものであれば、どのテストをとろうとも、そのテストの成績には、これら五つの知的能力が影響している、という点からこの五つの能力を共通因子と言う。

これに対して特殊因子または独自因子は被験者のテストの成績の変動に影響を与える源泉であるが、あるテストに固有であるという点において、共通因子とは区別される。そして因子モデルでは、これがテストごとに一個ずつ設定されている。この特殊因子または独自因子にかかわる因子負荷量と因子得点を、それぞれ（特殊因子（独自因子）負荷量）、（特殊因子（独自因子）得点）と言う。

そして最終的には、テスト j に関する被験者 i の成績 Zij は（4－3）式のように分解される。すなわち、因子分析においては、テストの成績は、テストごとに一定の値を持つ因子負荷量と被験者ごとに一定の値をとる因子得点の積和、すなわち線型結合として把握される。一般に共通因子は m 個ある。この時の（因子モデル）は（4－4）式のようになる。

Zji=aj1F1i＋aj2F2i＋aj3F3i＋aj4F4i＋aj4F5i＋ujUji　　（４－３）

（テスト j における被験者 i の成績）＝（テスト j における第一共通因子負荷量）×（被験者 i の第一共通因子得点）＋（テスト j における第二共通因子負荷量）×（被験者 i の第二共通因子得点）＋（テスト j における第三共

通因子負荷量）×（被験者 i の第三共通因子得点）＋（テスト j における第四共通因子負荷量）×（被験者 i の第四共通因子得点）＋（テスト j における第五共通因子負荷量）×（被験者 i の第五共通因子得点）＋（テスト j における独自因子負荷量）×（テスト j における被験者 i の独自因子得点）

Zji=aj1F1i＋aj2F2i＋・・・・ajmFmi＋ujUji　　（4－4）

ただし、測定値 Zji の変動に影響を与える第 3 の因子として（誤差因子）を設け、次式を一般的な因子モデルとみなす論者もいる。

Zji=aj1F1i＋aj2F2i＋・・・・ajmFmi＋ujUji＋ejEji　　（4－5）

ここに、ej は誤差因子負荷量、Eji は誤差因子得点

誤差項 ejEji を設ける論拠として、ケンドールは次の二つを挙げる。(4－4) 式の因子モデルは、m 個の共通因子が想定されている。もしも共通因子が（m＋1）個あるいはそれ以上あるとしたら、このモデルは不完全である。この不完全性を救うべく、(4－5) 式では誤差項が置かれている。これが第一の論拠である。第二の論拠として、得点 Zji が確率的に発見する観測誤差を含みうることを挙げている。ケンドールはいずれの論拠から誤差項を設けるかを峻別すべきという（しかし、数学的には、これを峻別することはできない）。いずれの場合にも、モデル構造の推定は、表 4－1 に示したようなデータ（n 個の観測対象のおのおのについて測った p 個の変数の値）から行われる。ところが、(4－4) 式の右辺には、観測可能な変数も既知の定数も含まれず、まったく未知のものばがかりである。そこで、手さぐり的な状況から通常取り上げられる典型的な問題は、次のとおりである。

①　共通因子数の決定

②　因子負荷量 a の推定

③　因子の解釈とその軸の回転

④　因子得点の推定

これらの問題に対して、それを解く基準が一意的には定まらないので、種々様々な手法が提案されている。その問題と手法名を図 4－2 にまとめて示す。

図 4－2　因子分析の適用のプロセス

モデルの設定→実験の計画（変量の選択、標本の抽出）→標本、多変量データ（正規分布か）→相関行列（分散・共分散行列）、(共通性の推定）→因子負荷量の推定（最尤推定、主成分分析、主因子分析、セントロイド法、アルファ因子分析、Image 法）→因子の解釈と回転（Varimax　　回転、コーティマックス回転、直交か斜交か）→因子得点の推定

ここで、特に注意することは、このモデル自身の一意性の問題などもいまだに解決されていないことである。実際、この種の問題は、確率的誤差を含まない純粋に数学上の行列理論の問題に帰着するのであるが、それがまだ解けていない問題も多い。

③　データを記述する立場としての因子分析

多変量解析には特にモデルを想定しないで表 4－1 に示したような多変量のデータの持つ構造を簡潔にまとめて記述しようとする立場がある。後述する主成分分析という解析手法もそういう立場の手法である。これに対して因子分析法はモデルを仮定して解析を行う推測的な立場であるとも言える。しかし、因子分析にもデータを要約する機能はあると考えるべきである。

いま、15 人の生徒に対する 12 個の人体計測値（変量）がある（表 4－1）。このデータをこれらの変量の平均からの偏差で表わして zij（i=1,2,…..,15 番目の生徒。J-1,2,….,12 番目の変数）としておく。ここで、それぞれのデータは標準化されているので（平均＝0、標準偏差＝1）とするので（後出 4－8 式を参照)、

Σzij（i=1～15）＝0　　(4－6）となる。

このデータ行列（15 人分のデータ×12 個の変数）を Z（15×12）＝（xij）と置く。ここで因子得点行列（F）と因子負荷量行列（A）を

F（15×m）＝fik　（i=1,…,15、k=1,…,m　この m は因子の数）

A（12×m）＝ajk　（j=1,…,12、k=1,…,m）とするとき、

F と A の転置行列 A´の積 FA´によって、もとのデータ行列 Z が十分に近似されるような F と A を求める。

Z（15×12）＝F（15×m）A´（m×12）＋E（15×12）　　(4－7)

ここで、E を残差とする時、これがある程度小さくなるような解を求めるのである。もちろん、データを要約す

るのがこの解析手法の立場だから可能な限り m を小さくする。つまり、因子分析では因子の数が少ないほどその目的にかなうのである。このようにデータを簡潔、節約的に記述する（A の m を小さくする）ことが因子分析の機能である。簡単に言えばここでは、12 個の変数を因子という新たな、より少ない変数に減らそうというのが解析の目的である。

④　観測対象の分類と類似化のための因子分析

表 4－1 では 15 人の観測対象を示したが、この 15 人に対してそれぞれ因子得点が推定される。この因子得点を用いて 15 人を適当に分類する。あるいはいくつかの類型にパターン化するというのがこの立場である。もし二つの共通因子によって表現されるとすれば、この 2 因子（新たに合成された二つの変数と考えてもよい）に対応する二つの因子得点によってある個人が類型化されることになるので、簡単に個人を分類したり、特徴づけたりすることが可能になる。このデータの場合には後述するように「体の形態の高さの因子」と「皮下脂肪厚に関する因子」というように簡潔に特徴づけられることになり、12 個の変数で個人を分類するよりもはるかに簡単に個人を特徴づける

⑤　確証的因子分析

この立場はデータを単に要約するというのではなく、因子数 m はもとより因子負荷量の一部についても、実質科学上の考察からある程度の構造を具体的に仮定したうえで、それをデータから検証しようという立場である。近年、米国を中心にして新たな因子分析の立場として脚光を浴びている。

（3）因子分析の計算とその解説

因子分析ではまず表 4−1 のようなデータ行列（X）から出発する。ここですべてのデータ（xi）を zi＝（xi−x（バー））/s　（x（バー）＝平均、s＝標準偏差）　　（4−8）

によって、標準得点（z）に変換しておくと z（バー）＝0、sz＝1.0 となるのでこれ以後の計算に都合がよい。この標準化されたデータ行列を Z とする。いま、個人を i＝1,2,…,N、測定項目を j=1,2,…,n とすれば Zji は

		個人	
		1,	2,…… N
Zji＝変量	1	z11	z12…..z1N
	2	z21	z22….z2N　　　（4−9）
	………		
	n	zn1	zn2….znN

という形式のデータ行列である。(説明の都合上 Zij でなく Zji としておく)、因子分析では先に述べたようにこの Zji が

Zji=aj1F1i＋aj2F2i＋・・・＋ujUji

のように構成分解されるモデルを考えるところに最大の特徴がある。この Zji という具体的なデータ行列から因子分析という解析的手続きを経て

aji＝(共通)因子負荷量、Fji＝（共通）因子得点、uj＝独自因子負荷量、Uj＝独自因子得点を抽出する（実際には独自因子負荷量と独自因子得点は計算の途中で消えてしまうので無視してよい)。

① 因子行列（因子負荷量行列）の数学的性格　　省略

変数 j と k の相関係数、同じく標準得点を用いて分散は、いずれも因子得点間の相関を示している。そこで、因子分析では重要な仮定を置く。

仮定 1　因子得点間の相関は 0 である。つまり因子得点ベクトルは直交している。

仮定 2　因子得点の標準偏差は 1.0 である。

② 共通性の推定

因子分析法には共通性の推定という重要な作業がある。

③ 因子の抽出

④ 因子回転

行列を回転させると、初期の因子分析結果はより理解しやすい構造に近づく。この構造を単純構造という。単純構造の基礎には「自然は単純である」というニュートン流の思想が存在するといわれる。このことは、因子分析の中では、因子数が出来るだけ少なく、少数の変数にのみ負荷を持つこと、ということになる。これは得られた回転後の因子行列のうちに出来るだけ 0 に近い負荷量を見出すようにすることである。因子分

析的研究のエッセンスはこの単純構造を達成することにある。回転法は、コーティマックス、バリマックス、オブリマックス、コーティミン、コバリミン、バイコーティング、オブリミン等の回転法など数多くが考察されている。コーティマックスは因子行列の各行、すなわち各変量の解釈を容易にするための単純化を図っているが、バリマックスは各列、すなわち抽出された因子の解釈を容易にするように回転されるなど、それぞれ特徴がある。最近はバリマックス回転法が多用されている。

カイザー（1958）の考案によるこの回転法は、因子負荷量の高い変数と低い変数に出来るだけ分化極大化させようとする（varimax とはバリアンス・マキシマムという意味）。これは、因子の解釈を容易にするためである。そのために、変数間の因子負荷量の平方の差異が最大になるように回転するわけである。

⑤ 因子の解釈と命名

因子分析の最後の作業は得られた因子を解釈し命名することである。共通因子というものは複数の変数間に潜在的に含まれている変数であるから、同じ因子に高い負荷（高い相関）を示した複数の変数間の関係からその共通となる性質を考慮して因子解釈を行う。

⑥ 斜交因子をどう解釈するか

斜交因子回転とは因子同士が直交している（相関しない）という前提を否定して、因子間に何らかの関係があるべきだと考える立場に立った回転。例えば、「体の大きさ」と「皮下脂肪厚」がまったく無相関であるとか「理科的な能力」と「文化的な能力」が相関ゼロであるという信念があるならばともかく、少しは関係があるだろうと考えるなら、斜交回転をするべきである。一般に斜交回転ではオブリミンやオブリマックスが利用されている。直交解では、何となくあいまいだった因子解が、よりすっきりすることが多いので、もっと利用してよい。

（4）因子分析を行う時の重点事項と注意

1．標本数はどれくらい必要か

一般にはよく 100 例以上が基準的に用いられているが、これは因子解が安定的であるために必要な最低限の数である。

2．標本の偏りは因子構造に影響するか

因子構造をはっきりとさせるためには分散の大きいデータが必要であり、そのためにはなるべく広い範囲にわたって標本を収集した方が良い結果が得られる。似通った標本ばかりで解析を行っても結局、明瞭な構造は得られないことが多いのである。

3．因子は何個まで抽出するのか（固有値について）

因子数は多ければ情報の取りこぼしが少なくなる。しかし情報を節約してモデルを説明するのが因子分析の一つの目的であることからして因子数が多すぎても冗長な結論になって具合が悪い。そこで、一般に基準として固有値と固有値寄与率を利用する。

1） 一般には固有値が 1.0 を超える因子を採用する。
この理由は単一の変数が行列の中にあって抽出可能な全分散を上回るくらいの分散を持つべきとの考えに基づいている。

2） 固有値の累積寄与率が 50%以上を下限とし、望ましくは 70%以上の分散を説明できることが望ましいが、これも相対的な問題であるので、目的に応じて加減してよい。

3） 固有値のグラフを作り、大きく固有値が減少したところで因子数を決定する。

4） 因子負荷量に 0.30 以上のものがない時は因子としては採用しない方が良い。因子数の決定は多方面にわたる検討が必要である。ある一つの基準だけで回転を施し、良い結果が示されたところで最終的に因子数を決めてやるのがよい。

4．共通性の推定値は何を用いるべきか

相関係数行列の対角成分に共通性の推定値を入れないと因子分析は始まらない。共通性とは、変数の分散を因子で説明できる割合で、0 を最小、1.0 を最大とする。最大の 1.0 を用いた場合（主成分分析）には因子数は最多となり、もし 0.1 などのきわめて小さい値を入れると因子数は少なくなる。一般的には、最大の 1.0 を代入して因子数を最大限に抽出し、取りこぼしがないように拾い出してから回転を与えて、再度因子数を決定するという方法、あるいは重相関係数の 2 乗値を用いる方法などがある。

5．共通性と因子寄与率をどのように解釈すべきか

因子負荷行列の中で、共通性は因子負荷量の 2 乗値の和、因子寄与は因子負荷量を二乗して全変数について合計する、この因子寄与を変数の数で割ると因子寄与率が得られる。さらに、この因子寄与率を累積すると共通性の平均値と一致する。

ここで、因子寄与率は全分散に対してその因子がどの程度の説明力を持っているかを示したものである。もし全変数がどれかの因子に帰属しているならば累積寄与率が著しく低くなることはない。しかし、累積寄与率が 30%以下であれば因子の回収性があまりに低いということになり、変数の選択に誤りがあったとか、因子数を少なくしすぎたとか、もともとの相関行列が低すぎるとか、さらには標本にひどい偏りがあったとか、様々な原因が考えられる。

６．どのような回転法を使うべきか

変数を解釈する場合にはコーティミン回転、因子を解釈する目的ではバリマックス回転が良く用いられている。いずれにせよ回転した後から因子を解釈するのであるから事後的に因子に意味を与えることが多いので、もし直交回転の結果が良好な結果を与えなければ斜交回転をしてみるとよい。実際のデータでは斜交回転の方が良い結果を与えることが多いかもしれない。

７．データの分析はどのようなものであるべきか

分布が正規分布から大きく歪んでいるものを用いないこと、極端な値をもつデータを除外しておく、分布の山が複数存在しないこと、０−１データなどの場合、比率がどちらかにひどく偏っていないこと、ある二つの変数の間に曲線の相関関係が存在しないこと

８．変数の独立性を保つとはどういうことか

Ａ、Ｂという変数を（Ａ＋Ｂ）などの合成変数と一緒に用いないこと、例えば最大血圧や最小血圧と一緒に脈圧（＝最大血圧−最小血圧）を用いるのは誤りである。また、一つの選択肢を二つの変数として利用するのも正しくない。

９．因子の解釈と命名にはどのような注意が必要か

因子の命名、解釈には最低複数個の変数が必要、もし、一つの変数しか用いないとするならば、それは因子分析をした意味を持たない。複数の因子にまたがる変数はどのように扱うのか、そのような変数はなるべく少なくすること、そのような変数は「因子的に複合的である」と解釈される。ＡとＢの因子にまたがる変数に対して、Ａのみ、あるいはＢのみに高い負荷を示す変数とよく比較する必要がある。因子階層の低い水準で、少数子の変数から因子が構成される場合、それらの変数がきわめて類似した内容の変数のことがある。この場合にはどちらかの変数を削るなどの工夫をし、よく吟味すること。

１０．異質集団の混在した因子分析

男性と女性、あるいはサラリーマンと自由業などの質的に違った集団を一緒に因子分析してよいか、という質問をよく受ける。分析する内容によっては、因子構造が異なっている可能性が大きい場合が少なくない。このような場合に一緒に因子分析をしてはいけない。同時に分析するには、あらかじめ異質集団の因子構造の類似を確かめておく必要がある。

１１．そのほかの注意

主成分分析か因子分析か、共通因子の抽出が目的ならば因子分析、単にデータを要約したいなら主成分分析。おおよその見当をつけるのが目的ならまず主成分分析をするし、モデル的な枠組みや仮説的構造を考えている場合には、因子分析を施すのが一般的である。

因子概念の妥当性は、因子の解釈は因子分析とは別に得られる事実によって、検証されるべきである。したがって、因子の抽出は研究の開始点を意味するのであって到着地点ではない。

因子分析の応用例：

ＳＤ法とその応用例

生活科学や心理学をはじめ、各方面でＳＤ法（Semantic Differential 法あるいは意味分化法）が頻繁に用いられている。言葉にはものごとを他と区別して定義づける表示的意味と同時に「好ましい」とか「良くない」といったような意味がある。オズグッドは我々が何かある概念（例えば、先生、ディナーなど）を提示された時、それによって脳裡に呼び起こされる「意味」（情緒的意味）内容には、どんなものがどれだけあるか吟味した。概念には、その外延と内包の相反する両側面がある。内包を引き出す基本的枠組みとして、相反する意味の形容語を

対とした、双極性評定尺度を多数用意する。そして、その概念をその尺度上で評価させる。初期の研究では約 50 対が用いられた。評定したデータをＮ人分集めて形容詞対がｎ対ならｎ×ｎのこれらの評定の間の相関係数行列を計算し､因子分析を行う｡オズグットの研究によるとＳＤ法では多くの場合次の３本の因子が見出されている。概念の種類や被験者の属性（姓、年齢、文化、言語など）の差にはほとんど影響されることなく、「評価」（良い－悪い)、「力量」(強い－弱い)、「活動」(速い－遅い）という３因子、あるいはこれらの変種が抽出されたことから、オズグッドは、これを「情動的意味体系における交差文化的普遍性」と名付けている。これは、文化人類学や比較言語学で広く知られているホワーフの仮説に対して提起された学説で、ホワーフの仮説が、言語の違いによって人間の認知システムも違ってくるとする言語的相対性を主張するのに対し、「美しい」ものは「良い」、「大きい」ものは「強い」、「動的」なものは「速い」と感ずる感じ方、すなわち情動的意味体系の構造は、文化や言語の差によってあまり変わらないという心理言語学的普遍性を提唱したものである。オズグッドはこの方法を言語系統の異なる諸民族に適用したり、男女差、その他を吟味して、この三次元が普遍的であると考えた。彼はまた、元来の色の明るさに適用されるべき、「明るい－暗い」といった評定尺度が、人間のイメージやにおいのもたらす感覚などにも十分適用できることの理論的根拠を、「共感覚」現象に求めた。我々が感覚について論ずる場合にもこの方法の有用性に注目してよいであろう。基本的次元の数が確定すれば、各次元を代表する形容詞対を適当な数選ぶことにより、少数の尺度だけで、十分的確なプロファイルを作ることができる。ＳＤ法の長所はその合理的整理法を提案したことにある。ＳＤ法は言語心理学だけでなく医学、政治学、生活諸科学などで広く利用することができる。この方法は我々の周囲を取り巻く、あらゆる種類の対象を言語で形容するには最小限何個の共通因子が必要が、そして、それによって、いかに形容されるかを見ることにある。形式的にみれば、この考え方は、正統的Rテクニックの因子分析と全く同じものである。しかし、あらゆる対象を若干の共通な言葉で表現しようとすれば、興味ある問題を取り上げたものということができる反面、それぞれの領域の専門的研究者からすれば、それぞれの領域における、微妙でデリケートなニュアンスの差をとらえるような専門性や特殊性を拭い去ってしまう結果にもなりかねない。したがって、実際上の形容詞対の選択には、それぞれの専門家に従う方が、より妥当性の高い結果になると思われる。

また、言語学上、双極尺度が、はたして反対語の組み合わせか否かが問題にされることがあるが、これらの場合にも単極性尺度としたり、最も単純化した場合にはチェックリストの方にしても、ほぼ同じ目的を達成することが出来よう。

被験者によっては「非常に～かなり～やや」といった「程度」を評定する際に誇張しがちな人もあれば、控えめな人もある。また一般に、評定者は両端の極端な評定を避け（中心化傾向)、また自分をとかく良い方に評定し（寛大化傾向)、また、ある点で優れた人は、他の点でも良いように評定しやすい（ハロー効果)。また、意見や態度の質問でも、回答者は自分がよく見られたい願望から社会的に望ましい意見の方に答えやすいという傾向を示すことがあるので、評定された内容を個人ごとにチェックしておくことも必要である。」

ここまで、抜粋。

「多変量解析のはなし（日科技連）より抜粋

４．相関の変わり者

尺度が変わると

表 4.1 をみると、東京 116 人、大阪 60 人、福岡 24 人の計 200 人に巨人、広島、大洋、阪神の中から一番ひいきのチームをあげてもらい集計したデータがそこにあります。

表 4.1　こんなデータがあるとする（実現値）

	巨	広	洋	神	小計
東京	66	12	20	18	116
大阪	4	10	4	42	60
福岡	14	6	2	2	24
小計	84	28	26	62	200

問題は、ひいきチームと都市との間に相関があるだろうか、ということである。各チームのファンは3つの都市に平均的に分布しているのではなく、都市によってファンチームに偏りがありそうだが、その偏りがどのくらい激しいものだろうかということです。この尺度は名義尺度なので、その相関を求めてみる。

もう一度、表4.1をみると、東京、大阪、福岡をならしていうと、巨、広、洋、神のファンが84：28：26：62の割合で存在している。また、巨、広、洋、神をこみにしてみると、東京、大阪、福岡に116：60：24の割合で住んでいる。そうすると、都市とひいきチームの間にまったく関係がなく、どのチームのファンも3つの都市に平均的に住んでいると仮定すると、東京の巨人ファンは調査人数200人のうち200×116÷200×84÷200＝48.72人となるのが公平なところ。同じように大阪の広島ファンは
200×60÷200×28÷200＝8.40 でなければならない。こうして、もし、都市とひいきチームの間に全く相関がないならば、表4.2のようになる。このような値は「こうなるはず」という値なので「期待値」と呼ばれる。

表4.2　相関がなければ、こうなるはず（期待値）

	巨	広	洋	神	計
東京	48.72	16.24	15.08	35.96	116
大阪	25.20	8.40	7.80	18.60	60
福岡	10.08	3.36	3.12	7.44	24
計	84	28	26	62	200

ところが、現実の調査では、表4.1のように、なってしまったのです。対比すると、表4.1と4.2とはだいぶ食い違いがある。この食い違いが小さければ小さいほど都市とひいきチームの相関はゼロに近く、食い違いが大きければ大きいほど相関は大きい。そこで、期待値と実現値との食い違いの大きさを求めることになります。求める値は、
（実現値－期待値）^{2}/（期待値）です。期待値で割ったのは、重みをそろえるためです。これを計算すると表4.4になり、これらを合計すると、75.21となり、χ^2である。

表4.4　食い違いの大きさを求める

	巨	広	洋	神	計
東京	6.13	1.11	1.61	8.97	75.21
大阪	17.83	0.30	1.85	29.44	
福岡	1.52	2.07	0.40	3.98	

しかし、行や列が多いほど大きな値になりやすいので、正規化する。そのためには、χ^2が取り得る最小値と最大値を調べる必要がある。χ^2が最小になるのは、実現値と期待値がすべて等しいときで$\chi^2=0$、χ^2が最大になるのは、実現値のうちゼロでないすべての値が、それに対応する行と列の周辺度数と等しい場合で、そのときχ^2はN（Nmin－1）となる。ここで、Nはデータの全数、Nminは行の数と列の数のうちの小さいほうの値。χ^2をこの値で割った値が、クラメールの関連指数。」
ここまで、抜粋。

5.3　3変数以上の分析Ⅱ：重相関係数の計算

予定

このコースは以下の3つのステップを予定している。

a　重相関係数の計算と有意性検定
b　偏相関係数の計算と有意性検定
c　予測式の算出と回帰分析

なお、ここでは、以上のステップを3変数の場合に限って解説する。4変数以上の場合は複雑であり、専門的な問題が絡んでくるので、実際にデータ処理をおこなうときには専門家とコンピュータを頼んだ方が無難である。

重相関と重相関係数の定義

重相関（multiple correlation）とは、変数が3個以上の場合で、かつ、この変数間にあらかじめ一定の予測関係が存在する場合の相関関係のことである。すなわち、全ての変数のうち、特定の1個が目的変数、あとの残り

が予測変数と最初から決まっていなければならない。また、重相関係数とは、そのような３変数以上の間の相関関係の強さを表す統計量である。重相関係数の記号には一般に“R”が用いられる。これは２変数の場合の相関係数の記号”ｒ“の大文字である。

ａ．重相関係数の計算と有意性検定

分析の第一ステップは、重相関係数の計算と有意性検定である。

Ｒの計算

３変数Ａ・Ｂ・Ｃがあり、予測関数が明確に存在する（大前提）。ここでは、Ａが目的変数、Ｂ・Ｃが予測変数であるとする。

１）データ表示を行う。

表 86　各変数の平均と標準偏差（満点各 20 点、（N＝50）

	条件Ａ	条件Ｂ	条件Ｃ
X	14.9	15.7	13.8
SD	7.6	5.2	6.9

２）２変数ずつ取り出し、相関係数ｒを計算する。

ここでは、すでに計算ずみであり、以下のとおりとする。なお、条件Ａの全データを変数Ａということにする。以下同様。

変数Ａと変数Ｂの相関係数；r_{AB}＝.512

変数Ｂと変数Ｃの相関係数；r_{BC}＝.356

変数Ａと変数Ｃの相関係数；r_{AC}＝.434

３）ｒの値を下式に代入し、重相関係数Ｒを計算する。

$$R=\sqrt{\left(r^2_{AB}+r^2_{AC}-2\cdot r_{AB}\cdot r_{BC}\cdot r_{CA}\right)/\left(1-r^2_{BC}\right)}$$

$$=\sqrt{\left(.512^2+.356^2-2\cdot .512\cdot .356\cdot .434\right)/\left(1-.434^2\right)}$$

$$=.533$$

Ｒの有意性検定

Ｒを計算したら、次に、その有意性検定をおこなう。これには、2 変数の相関係数の有意性検定と同様、F 分布を利用する。

１）R、変数の個数、データの組数（N）の値を用意する。本例では R＝.533、変数の個数は 3、N＝50 である。

２）以上の値を下式に代入し、F 比を計算する。

$$F=R^2/(\text{変数の個数}-1)/(1-R^2)/(N-\text{変数の個数})$$

$$=.533^2/(3-1)/(1-.533^2)/(50-3)=9.33$$

３）F 比の分子・分母の自由度を計算する。

分子の自由度；df①＝変数の個数－1＝3－1＝2

分母の自由度；df②＝N－変数の個数＝50－3＝47

４）F＝9.33 を F 分布に対照して、出現確率を求める。

df①＝2、df②＝47 であるが、（2,40）の F 分布を代用する。

表 87　F 分布表（一部）

df①	df②	F 比		
2	40	2.44	3.23	5.18
出現確率		.10	.05	.01
有意水準		有意傾向	5%	1%

F＝9.33 は 5.18 を越えている。したがって、$p<.01$ である。ここで、R＝.533 は有意であった（$F(2,47)=9.33$、$p<.01$）。しかし、もし有意または有意傾向でなかったら、次のステップへ進むことはできない。

重相関係数Ｒの性質

上に計算したＲの性質を以下に列挙する。これらの性質のうち①と②は、2 変数の相関係数ｒには見られないものであり、注意が必要。

①　Ｒは０から１までの正の値しかとらない。

重相関は３変数以上の対応関係であるため、その方向性をプラス・マイナスだけに限定することはできない。したがって、R は、重相関の強さだけを表す。そのように、R は強さの情報しか持っていないので、論文には R そのままでなく、強さの意味を表す R^2 を掲載するのが通例である。

② 予測関係が変わると R の値も変わる。
例えば、「目的変数 A、予測変数 B・C」としたときの R の値と「目的変数 B、予測変数 A・C」としたときの R の値とは一致しない。この点、変数間の予測関係が事前に確定していないと、なにを求めているのかわからなくなる。

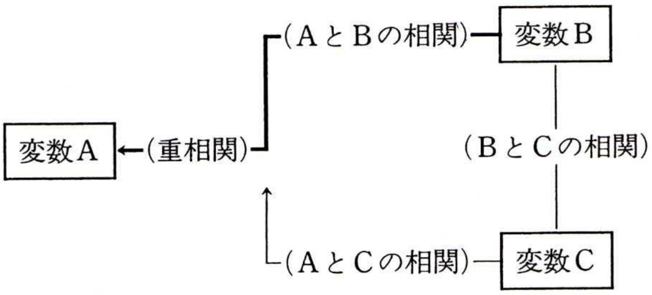
③ 使用上の制約は r と同様である。
各変数間の相関の直線性、データ分布の正規性、目的変数の（各予測変数に対する）等散布性が満たされないと、R は使用できない。

b. 偏相関係数の計算と有意性検定

重相関係数が有意であった場合、分析の第二ステップは、偏相関係数といわれる統計量の計算と有意性検定である。

偏相関の定義

偏相関（partial correlation）とは「部分的相関」という意味である。すなわち、重相関における複数の予測変数のうち、特定の１個だけと目的変数との相関関係を表す用語である。例えば、変数 A・B・C があり、A が目的変数、B・C が予測変数である時、この３変数は以下のような相関関係によって結ばれている（矢印は予測の方向）。



このとき、例えば、目的変数 A に対する予測変数 B だけの純粋な相関（上図の太線部分）をみるには、このAとBの相関に加わる変数 C の影響力（A と C の相関、B と C の相関）を取り除かなければならない。偏相関とは、そうした他の変数の影響を取り除いたときの、目的変数と予測変数１個との部分的相関を意味している。なお、上図からわかるように、３変数の場合、偏相関係数は２個求めることができる（A と B、および、A と C）。

偏相関係数の記号

偏相関係数の記号は、通常の相関係数の記号 r に特別な添え字を付ける。例えば、“$r_{AB.C}$”のように書く。この添え字「AB.C」の意味は、「変数 A と変数 B の偏相関・変数 C の影響を取り除いたときの」ということである。

計算例

重相関係数を計算したときの例を用いて、さらに偏相関係数を計算してみる。前例では、目的変数 A、予測変数 B・C であった。

① 各２変数の相関係数を計算する。
$r_{AB}=.512$、$r_{BC}=.434$、$r_{AC}=.356$

② 各相関係数を下式に代入し、偏相関係数を計算する。
まず、変数 A と変数 B の偏相関係数を計算する。

$$r_{AB.C}=(r_{AB}-r_{AC}\cdot r_{BC})\ /\ (\sqrt{(1-r^2_{AC})}\cdot\sqrt{(1-r^2_{BC})})$$
$$=(.512-.356\times.434)\ /\ (\sqrt{(1-.356^2)}\times\sqrt{(1-.434^2)})=.425$$

次に、同様にして、変数 A と変数 C の偏相関係数を計算する。

$$r_{AC.B}=(r_{AC}-r_{AB}\cdot r_{BC})\ /\ (\sqrt{(1-r^2_{AB})}\cdot\sqrt{(1-r^2_{BC})})$$
$$=(.356-.512\times.434)\ /\ (\sqrt{(1-.512^2)}\times\sqrt{(1-.434^2)})=.173$$

偏相関係数の有意性検定

偏相関係数の有意性検定は、通常の相関係数や重相関係数と同様、F 分布を利用する。ここでは、先に計算した $r_{AB.C}=.425$ を検定してみる。

① $r_{AB.C}$、変数の個数、データの組数（N）を用意する。
$r_{AB.C}=.425$、変数の個数は３、N＝50 であった。

② 以上の値を下式に代入し、F比を計算する。

$F = r^2_{AB.C} / (2-1) / (1 - r^2_{AB.C}) / (N - 変数の個数)$

$= .425^2/1/(1 - .425^2)/(50-3) = 10.36$

式中、分子では説明率を自由度1個分に換算しているが、このときの自由度は、「変数の個数−1」すなわち3−1でなく、2−1である。なぜなら、偏相関は目的変数1個と予測変数1個の2変数相関になるからである。ただし、偏相関は重相関のなかの部分相関であるので、偶然誤差（分母）は重相関の偶然誤差と同一のものを用いる。したがって、計算したF=10.36は、変数Aと変数Bの偏相関の強さ（説明率）が重相関の偶然誤差の約10倍の大きさであることを示している。

③ F比の分子・分母の自由度を計算する。

分子の自由度：df①＝2−1＝1（←偏相関の場合はつねに1）

分母の自由度：df②＝N−変数の個数＝50−3＝47

④ F=10.36をF分布表に対照する。

df①＝1、df②＝47であるが、表に載っていないので、それより小さい自由度の（1,40）のF分布表を代用する。

表88 F分布表

df①	df②	F比		
1	30			
	40	2.84	4.08	7.31
	60			
出現確率		.10	.05	.01
有意水準		有意傾向	5%	1%

F=10.36は右端の7.31を越えている。したがって、$p < .01$である。

なお、前例のもう一つの偏相関$r_{AC.B}$＝.173は、同様の検定の結果、有意でなかった（$F(1,47) = 1.45$、$p > .10$）。

偏相関係数の性質

偏相関係数の性質は、通常の2変数の相関係数rと同一である。

値は−1~＋1まで変化する。

値の＋・−は相関の正・負を表す。

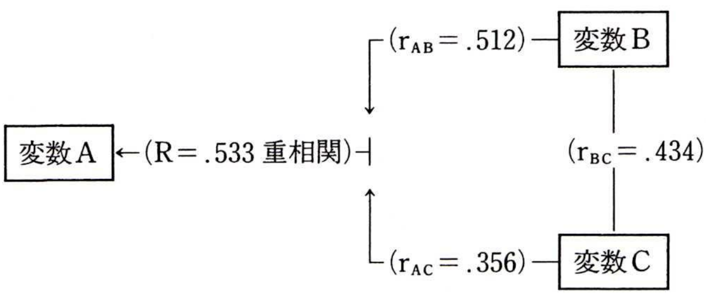
値の大きさは相関の強さを示す（二乗値は説明率になる）。

例えば、上例では$r_{AB.C}$＝.425であり、有意であった。したがって、変数Cの影響を取り除いたときの変数Aと変数Bの純粋な相関は、「中程度の強さの正の相関」であるとわかる。

備考：偏相関と見かけの相関

前例では、変数A・B・Cは、次のような相関関係になっていた（矢印は予測の方向を表す）。

右図からわかるように、変数Aと変数Bの相関r_{AB}＝.512は、いわゆる第三の変数Cに影響された「見かけの相関」である。このことは、r_{AC}＝.356にもあてはまる（第三の変数はB）。偏相関係数は、このような第三の変数の影響を取り除いたときの純粋な相関係数である。その結果、値は以下のように変わった。



変数Aと変数Bの相関：r_{AB}＝.512**→$r_{AB.C}$＝.425**　（*$p < .05$、**$p < .01$）

変数Aと変数Cの相関：r_{AC}＝.356*→$r_{AC.B}$＝.173

この値の変化は、「見かけの相関」から「実質的な相関」への変化であると言える。これによって、変数Aと変数Cの相関は、有意でなくなったが、以前の相関は変数Bに影響されたうわべのものにすぎなかった。かくして、目的変数Aを予測するには、予測変数Bは有力であるが、予測変数Cは有力でないということがわかる。

c. 予測式の算出と回帰分析

目的

重相関係数と偏相関係数を計算したら、最後は予測式を算出する。3 変数以上の場合に、予測式を求める目的は二つある。１）目的変数を予測すること、および２）予測変数を選別することである。このうち、最初の１）はふつうの目的であり、目的変数の値の変化に関心がある。しかし、一般に研究例は少ない。これに対して、次の２）の目的は、予測変数どうしの有効性の比較に関心がある。すなわち、目的変数を予測するのに、どの予測変数が有効か、どの予測変数が有効でないかをより分ける。この目的をとった場合は実験計画法に似ており、けっきょく「効く要因」の検討になる。3 変数の場合を例にとると、そのときの予測式の形は、例えば以下のようになる。

A‘＝1.38B＋0.77C＋11.6

上式の傾き（1.39、0.77）が予測変数 B と C の効き具合を表すことになる。ここでは、変数 B のほうが有力である。すなわち、1 ポイント変化したときに目的変数 A の値を上下する度合いが大きい（1.39＞0.77）。

なお、予測変数 B・C が共に 0 のときも、目的変数 A は 11.6 程度の値をとる。これは、なんの説明もできないので、偶然誤差によるユレであるとみなすしかない。かくして、上式は、実験計画法における分散の分解式に類似した意味を持つ（下式）。

データ A‘＝変数 B の効果＋変数 C の効果＋偶然誤差

このように予測式をデータの模型にみたてて、予測変数どうしの重み（傾き）の大きさを比較する方法を回帰分析と呼ぶ。なぜなら、分析の焦点が目的変数でなく予測変数にあるので、視点は「目的変数←予測変数」（予測）という方向でなく、「目的変数→予測変数」（回帰）という方法になるからである。したがって、回帰分析を目的とした場合の予測式は「回帰式」と呼ぶ方が適切であり、専門家も、その呼び方を好む。

重回帰式の算出例

さて、予測変数どうしの有効性を比較するという目的をもって、回帰式を求める。この場合、比較すべき予測変数は 2 個以上あるので、前述した予測変数 1 個の予測式（単回帰式）と区別して、特に重回帰式といわれる。まず、目的変数 A、予測変数 B・C とした場合の重回帰式の一般形を示す（下式）。

$$A' = k_{AB}\cdot B + k_{AC}\cdot C + y$$

上式の k_{AB} と k_{AC} はいわゆる「傾き」であるが、専門用語では偏回帰係数と呼ぶ。偏回帰係数は、目的変数 A に対する各予測変数の単独の影響を表す。例えば、k_{AB} は目的変数 A に及ぼす予測変数 B だけの影響力（他の予測変数 C の影響分を取り除いた）を表す。ちなみに、偏相関係数と偏回帰係数は同一主旨の別表現である。すなわち、目的変数と予測変数 1 個だけの単独の相関関係を、偏相関関係は標準的な－1~＋1 の範囲の値に換算し、偏回帰係数は目的変数の数値の寸法に合わせて表現しているに過ぎない。なお、y は定数であり、タテ軸（変数 A）の切片である。以上の k_{AB} 、k_{AC} 、y の値を計算すればよい。

１）計算に必要な統計量をそろえる。

各変数の X と SD、および、各 2 変数の相関係数 r_{AB}、r_{AC}、r_{BC} または、偏相関係数 $r_{AB.C}$、$r_{AC.B}$ を用意する。前例を用いることにして、以下に各統計量を掲載する。

表 86　各変数の平均と標準偏差（満点各 20 点、（N＝50）

	条件 A	条件 B	条件 C
X	14.9	15.7	13.8
SD	7.6	5.2	6.9

相関係数：r_{AB}＝.512、r_{BC}＝. 356、r_{AC}＝. 434

偏相関係数：$r_{AB.C}$＝.425、$r_{AC.B}$＝.173

重相関係数：R＝.533

２）偏回帰係数 k_{AB} 、k_{AC} を計算する。

相関係数だけを用いる計算式と、偏相関係数を用いる計算式を示す。いずれを用いてもよい。まず k_{AB} を計算する。

k_{AB}＝（$r_{AB} - r_{AC}\cdot r_{BC}$）/（$1 - r^2_{BC}$）・$SD_A/SD_B$＝$r_{AB.C}\sqrt{}$（$1 - r^2_{AC}$）/$\sqrt{}$（$1 - r^2_{BC}$）・$SD_A/SD_B$＝（.512 －.356×.434）/（1－.434²）×7.6/5.2＝.425×$\sqrt{}$（1－.356²）/$\sqrt{}$（1－.434²）×7.6/5.2＝0.644　　（負の相関係数がある場合は－の値になる。）

次に、同様にして、k_{AC} を計算する。

k_{AC}＝（$r_{AC} - r_{AB}\cdot r_{BC}$）/（$1 - r^2_{BC}$）・$SD_A/SD_C$＝$r_{AC.B}\sqrt{}$（$1 - r^2_{AB}$）/$\sqrt{}$（$1 - r^2_{BC}$）・$SD_A/SD_C$＝（.356

$-.512\times.434)$ / $(1-.434^2)\times7.6/6.9=.173\times\sqrt{}(1-.512^2)/\sqrt{}(1-.434^2)\times7.6/6.9=0.182$

３）切片 y を計算する。

$y=X_A-k_{AB}\cdot X_B-k_{AC}\cdot X_C$

$=14.9-0.644\times15.7-0.182\times13.8=2.278$

４）以上の k_{AB} 、k_{AC}、y の値を重回帰式の一般形に代入する。

$$A' = k_{AB}\cdot B + k_{AC}\cdot C + y$$

$$\downarrow$$

$$A' = 0.64B + 0.18C + 2.28$$

偏回帰係数の有意性検定

さて、重回帰式において予測変数の有効性を特定するためには、偏回帰係数の有意性検定を行う。しかしながら、これは、偏相関係数の有意性検定と全く同一結果になるので、本例では、すでに下表のように済んでいる。

予測変数	偏回帰係数	偏相関係数	有意性検定の結果
B の場合	$k_{AB}=0.64$	$r_{AB.C}=.425$	F（1,47）＝10.36、$p<.01$
C の場合	$k_{AC}=0.18$	$r_{AC.B}=.173$	F（1,47）＝1.45、$p>.10$

重相関係数・偏相関係数・重回帰式のまとめ

3 変数以上の「相関・予測の分析」は、実際の研究例では、複数の予測変数の比較すなわち回帰分析に帰着することが多い。ただし、分析の結果を論文に記載する場合には、偏相関係数を用いた書き方と偏回帰係数を用いた書き方があり、やや論調が異なる。

１）偏相関係数を用いた論文記載例

「表 86 は、変数 A・B・C の平均と標準偏差を示したものである。変数 A を目的変数、変数 B・C を予測変数として重相関係数を計算した結果、$R^2=.284$ であり、有意であった（F（2,47）＝9.33、$p<.01$）。

ここで、変数 C を一定としたときの変数 A と変数 B の偏相関係数.425（F（1,47）＝10.36、$p<.01$）。また、変数 B を一定としたときの変数 A と変数 C の偏相関係数は.173（F（1,47）＝1.45、$p>.10$）であった。したがって、変数 A と変数 B の間に実質的な相関関係のあることが示唆される。説明率（偏相関係数の二乗値で$.425^2$）は 18%であり、相関の強さは中程度である。」

上の文章中、重相関係数 R は、相関の強さの情報しか持っていないので、強さの正味を表すために二乗値として掲載している。第二段落から、予測変数どうしの比較にはいる。「変数 C を一定にしたときの」という言い方は、「変数 C の影響を取り除いたときの」という意味と同じであるが、実際には変数 C の影響を 0 にすることはできないので厳密な表現をとっている。そのように言う方がよい。最後に説明率に言及しているが、偏相関係数は相関の強さを理解しやすい利点がある。

２）偏回帰係数を用いた論文記載例

「表 86 は、変数 A・B・C の平均と標準偏差を示したものである。変数 A を目的変数、変数 B・C を予測変数とした回帰分析の結果、変数 B の偏回帰係数は 0.64（両側検定：t（47）＝3.22、$p<.01$）。また、変数 C の偏回帰係数は 0.18（両側検定：t（47）＝1.20、$p>.10$）であった。したがって、変数 A に及ぼす変数 B の効果は有意であるが、変数 C の効果は実質的なものであるとは言えない。

なお、このときの回帰式全体の説明率は $R^2=.284$ であり、有意であった（F（2,47）＝9.33、$p<.01$）。」

偏回帰係数を用いた書き方の場合は、はっきりと予測変数の取捨選択に焦点が絞られる。重回帰式の全体形は掲載されないことが多く、偏回帰係数の有意性検定が興味の中心となる。上文では、偏回帰係数の有意性検定に t 値を用いているが、コンピュータ・プログラムによっては、F 比でなく t 値を出力する場合があるので参考例として示した。結果は、F 比による検定と同一である（上例の t 値を二乗すると前に記載した F 比になる）。ただし、t 値を用いた検定の場合には「両側検定」と「片側検定」の区別を付記しなければならないので、どちらの t 値であるかをよく確認すること。

なお、ここで解説した回帰分析は一回きりの分析であったが、実際の回帰分析は、もっと多くの予測変数を扱うので、予測変数どうしをいろいろ組み合わせて何回も回帰式を算出する（ステップ・ワイズ方式と呼ぶ）。

重回帰式と回帰分析の制約

以下の場合、重回帰式の算出および回帰分析をおこなうことはできない。

１）重相関係数が有意でない場合

重相関係数は、重相関の強さを示すが、それはそのまま重回帰式の予測力である。したがって、重相関係数が有意でない場合、重回帰式は偶然にあたる程度の予測力しかもたず、求める意味がない。

2）いずれかの変数間に曲線相関が想定される場合

重相関係数・偏相関係数に限らず、すべての相関の統計量は直線相関を前提としている。また、予測式（回帰式）も単純加算の線型モデルに基づいて立てられるので、曲線相関には対処できない。したがって、2 変数ずつ散布図を描いてみると、相関が直線的でない場合は以上の統計量や予測式・回帰式を用いることができない（用いても有意性を得ることができない）。

備考；多重共線性の問題

ここでは、2 変数の場合しか扱わなかったが、多数の予測変数を回帰分析に持ち込む場合、もし、これらの予測変数どうしの相関が高いと、偏相関係数・偏回帰係数が不当に高くでることが知られている（この場合、計算式の分母を検討する）。専門的には、「多重共線性」の問題と呼ばれる。内容や性質がよく似た予測変数を持ち込む場合に起こりやすい。

回帰分析の事例：ステップ・ワイズ方式

論文名「青年期の自我の安定に貢献する要因について」

目的

青年期の自我の安定度が、どのような要因に依存しているかをあきらかにする。そのような要因として考えられるものは、対人適応性（対人関係における適応性）、学業成就度（学業上の成功・失敗の度合い）、未来志向性（未来に対する見通しの大きさ）などである。それぞれの要因が自我の安定に貢献するかどうかを検討してみる。

方法

対象者：高校 2 年生 64 人（男子 31 人、女子 33 人）

道具：自我安定度、対人適応性、学業成就度、未来志向性についてそれぞれ 15 項目ずつの質問を載せた調査用紙。

手続き：調査は集団で行った。各質問項目（例：なにもかも放り出したくなることがありますか？）に対する回答は、「よくある・ときどきある・あまりない・ほとんどない」の 4 肢選択として 3 点から 0 点までに得点化した。これを合計して、各変数 45 点満点とした。なお、学業成就度については、過去の教科成績と照合して整合性のない者を除外した。

結果

表 90 は、各変数の平均と標準偏差を示したものである。

表 90　各変数の平均と標準偏差（満点 45）　　(N=59)

	自我安定度	対人適応性	学業成就度	未来志向性
X	27.5	32.6	22.9	23.3
SD	11.3	12.5	15.4	10.8

自我安定度を目的変数としてステップ・ワイズ方式の回帰分析を行った結果、偏回帰係数が有意であった変数は選出順に対人適応性（$F(1,54)=11.58$、$p<.01$）と未来志向性（$F(1,55)=6.30$、$p<.05$）であった。したがって、青年期の自我の安定度は、対人関係の円滑さと未来に対する展望に依存していると言える。

表 91　回帰分析の結果

ステップ	予測変数	R^2	回帰係数	F 比
1	対人適応性	.61	0.40	11.85**
2	未来志向性	.69	0.29	6.30*

解説

本例は、予測変数 3 個に対するステップ・ワイズ方式の回帰分析の例である。ステップ・ワイズとは「一段一段」という意味である。すなわち、最初から全部の予測変数を用いて回帰式を求めるのではなく、1 変数ずつ、その重み（偏回帰係数）の有意性を確認しながら回帰式のなかへ入れてゆくのである。例えば、表 91 によれば、第 1 ステップとして、3 つの予測変数のうち対人適応性の回帰係数（0.40）が最大の有意性を示したので（F=11.85）、

まずこれを用いて回帰式を求めた。そのときの回帰式の予測力は、.61 であったことがわかる。次に、第 2 ステップとして、残る 2 つの予測変数のうち未来志向性の（偏）回帰係数（0.29）のほうが回帰式のなかで大きな有意性を示したので（F＝6.30）、これを追加することにした（回帰式の予測力は R^2＝.69 となった）。なお、すでに回帰式のなかには、第 1 ステップにおいて選出された対人適応性が存在している。したがって、未来志向性を追加したことによって、この対人適応性の有意性が低下したかもしれないが、いったん入れた変数は出さないという方針をとる。ただし、オプションとして、新たな変数の追加によって有意水準を割り込んだ既存の変数は排除するという方針をとることもできる（その場合は表 91 の結果の書き方が異なる）。

さて、第 2 ステップが終了した時点で、残る予測変数（学業成就度）を回帰式に組み入れ、偏回帰係数の有意性を検定したが、有意水準に達しなかったため組み入れを中止した。したがって、第 3 ステップはない。

このようにステップ・ワイズ方式の回帰分析は、先に選出された予測変数が存在する時点で、残る予測変数の回帰係数を計算しなおし、そのうち最大の有意性（最大の F 比）を示す変数を 1 個ずつ選出してゆく。そして、残りの予測変数のどれもが有意水準に達しなくなった時点で分析を終了するのである。結果として、この方法は、必ずしも回帰式全体の予測力（R^2）を最大にしない。この点、予測そのものへの関心は二次的であり、予測変数の取捨選択に主目的がある。

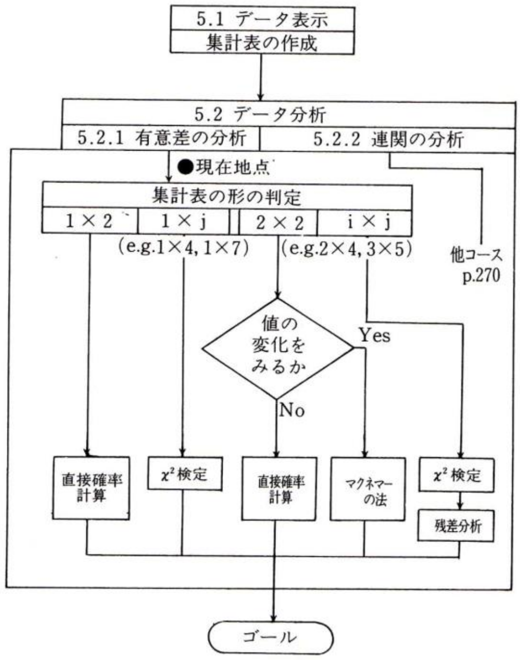
ところで、ステップ・ワイズ方式にも、いくつかの方法がある。本例は、一般によく用いられる順次追加型（forward selection）であったが、いったん全変数を回帰式に投入しておいて有意でない変数を 1 ステップ 1 個ずつ排除してゆく順次排除型（back-ward elimination）もある。また、やや目的を異にするが、回帰式の予測力本位に R^2 を最大にするように予想変数を選別する方法もある。いずれにしろ、これらは手計算のレベルではない。

6. 名義尺度データの処理

実際のデータ処理の頻度から言えば、これまで述べてきた間隔・比率尺度データを扱う例が圧倒的に多い。次に名義尺度データである。順位尺度は、実際にはあまり扱われない。ここでは、名義尺度データの処理を解説する。

予定

名義尺度データの処理も、基本手順は、第一にデータ表示、第二にデータ分析である。



6.1 データ表示；集計表の作成

名義尺度データの表示は、集計表を作成することである。ただし、名義尺度データの値は、数量でなく、性質や状態である（このため質的データまたはカテゴリカル・データといわれる）。したがって、平均や標準偏差を計算することができない。例えば、下の表は、左が間隔・比率尺度であり、右が名義尺度データである。

表 92　間隔・尺度データ

被験者	条件A	被験者	条件B
イ	77	ヘ	58
ロ	64	ト	39
ハ	80	チ	61
ニ	52	リ	48
ホ	73	ヌ	55

表 94　名義尺度

被験者	条件A	被験者	条件B
ル	できた	タ	できない
ヲ	できた	レ	できた
ワ	できない	ソ	できない
カ	できた	ツ	できた
ヨ	できた	ネ	できない

表 93　X、ＳＤの掲載（得点）

	条件Ａ	条件Ｂ
N	5	5
X	69.2	52.2
SD	10.1	7.9

表 95　度数の掲載（人数）

	条件Ａ	条件Ｂ
できた人	4	2
できない人	1	3

この場合、名義尺度のデータの値は、「できた」・「できない」の 2 値である。これは数量でないので、唯一可能なデータ表示の方法は、表 95 のように各値をもつ被験者の人数をカウントしてゆくことだけである。カウントした数を度数、その掲載表を集計表と呼ぶ。

集計表の例

幾つかの集計表の例を示す。

1）1×2 集計表、1×3 集計表

「1×2」とは、「1 条件×2 値」という意味、同様に「1×3」とは、「1 条件×3 値」という意味である。

表 96　講演会の出席者数

男性	女性
59	84

表 97　新執行部の支持者数

支持	中立	不支持
234	47	566

表 96 は講演会という 1 つの条件において出席者の性別（男性・女性の 2 値）をカウントしたものである。また、表 97 は新執行部という 1 つの条件への各会員の意向（支持・中立・不支持の 3 値）をカウントしたものである。これらの例では、データ分析へ進んだ場合、1 条件内の値どうしの多い少ないを比較することになる。このように、名義尺度データの場合は、条件が一つしかなくても比較可能である。

もし、これらを「男性講師の講演会 vs.女性講師の講演会」とか「新執行部 vs.旧執行部」という 2 条件にすれば、条件間の比較となる。その場合、集計表は「2×2」または「2×3」となる。なお、条件と値はどちらがタテでもヨコでもよい。

表 98　講演会の出席者数

	男性	女性
男性講師	59	84
女性講師	72	95

表 99 新旧執行部の支持者数

	支持	中立	不支持
新執行部	234	47	566
旧執行部	291	38	518

2）2×2×3 集計表

「2×2×3」とは、「2 条件×2 条件×3 値」という意味である。このように条件を組み合わせた比較の場合は、条件どうしをまとめてヨコに並べるのがよい。

表 100　災害映画の視聴と援助行動の意志（人）

街頭募金活動への参加の意志	映画視聴群		統制群	
	男子	女子	男子	女子
全面的参加	9	12	4	3
一時的参加	21	32	11	16
参加しない	16	7	31	29

上表では、「映画視聴群 vs.統制群」で 2 条件である。そして「男子の場合 vs.女子の場合」で 2 条件である。あわせて 2×2 の 4 条件である。このとき、各被験者のもつ値は、街頭募金の活動に「全面的に参加する」・「一時的なら参加してもよい」・「参加しない」の 3 値である。計、2×2×3 となる。

なお、「統制群」とは実験的処遇をなにもあたえていない被験者集団のことである。すなわち、統制群の被験者は災害映画をみせられず、たんに街頭募金への参加の意志を問われただけである。

ところで、名義尺度において、2×2×3 というような交差した条件間の比較は最近までかなり困難であった。今では、ログ・リニア・アナリシスの統計プログラムの開発によって簡単に対処できるようになっている。

6.2　データ分析

データ表示が終わったら、次に、データ分析に移る。名義尺度データの分析には、1）有意差の分析と、2）連関の分析の 2 つのコースがある。

1）有意差の分析コース

このコースでは、集計表に掲載した人数（または度数）を相互に比較して、有意な差があるかどうかを判定する。この判定には、直接確率計算やχ^2検定を用いる。

6.2.1　有意差の分析

方法の選択

有意差の分析をする方法としては、直接確率計算とχ^2検定がある。また、特殊用途として、マクネマーの法がある。これらの使い分けは、集計表の形によっておこなう。

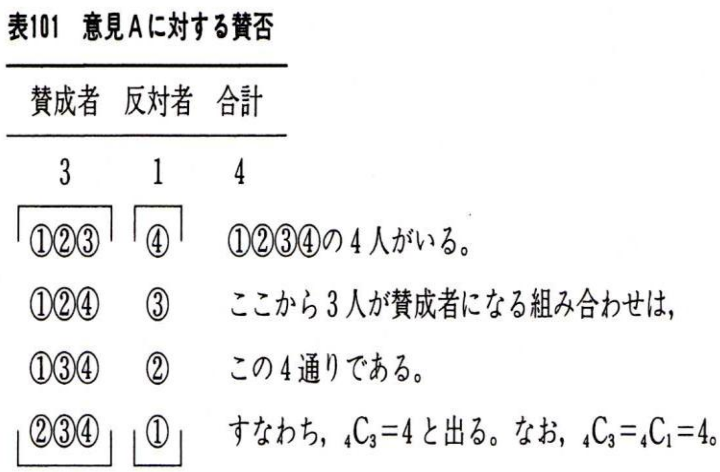
a.1×2 表の分析；直接確率計算

適用

1×2 表とは、「1 条件×2 値」の集計表である。この形の分析は 2 値間の有意差の分析となる。

予備知識

1×2 表の有意差の分析には直接確率計算をおこなうが、これに用いる「組み合わせ関数」について、事前に解説しておく。組み合わせ関数は、電卓のボタン表示で ${}_{N}C_{r}$ である。“C”は、“combination”（組み合わせ）の頭文字であり、「N 個から r 個を取り出す組み合わせの数」を出力する。例えば、下の集計表において 4 人から 3 人を賛成者として取り出す組み合わせは、全部で 4 通りである。



すなわち、4 人に賛否をたずねる場合 (N＝4)、3 人が賛成にまわる（ｒ＝3）というケースは全部で 4 通りあることがわかる。なお、「4 から 3 を取り出す」場合と「4 から 1 を取り出す」場合の組み合わせの数は、上の例からあきらかなように等しい。

ちなみに、組み合わせ関数 ${}_{N}C_{r}$ の定義式は以下のとおりである。

さて、1×2 表の有意差の分析には、この組み合わせ関数を用いて、直接確率計算をおこなう。以下に、事例によって説明する。

直接確率計算の事例

町内の空き地の有効利用に関する私案Ａに対して、無作為に 9 人の住人を選び、賛成・反対をたずねた。表 102 は、その結果である。賛成者と反対者の人数に実質的な差があるかどうか、直接確率計算によって検定してみる。

表 102　試案Ａに対する賛否（人）

賛成者	反対者	合計
2	7	9

もちろん、住人全員に賛否を尋ねることができれば検定の必要はない。その場合は、1 人、2 人の差があっても多いほうが絶対的に多い。しかし、上の例の場合は 9 人の住人に限定した結果であるので、5 人の差があっても真に反対者が多いとは即断できない。このように、統計的検定とは母集団の一部分のデータしか入手できない場合に全体傾向を推測するための方法である。この点、データ（9 人）が無作為抽出であるということが重要である。

① N＝9 の時の出現可能な組み合わせをすべて数え上げる。

まず検定上の帰無仮説を立てる。すなわち「賛成 2、反対 7」は偶然の人数のユレであり、ほんとうは賛成者・反対者の出現確率は 50%・50%であると仮定する。すなわち、N＝9 を賛成・反対の二つに分けるすべての組み合わせは等確率で出現するとみておく。そこで、考えられるすべての組み合わせを列挙し、その数をカウントし

てみる（表 103）。

表103　N＝9の賛否の組み合わせ数

	賛成(r)	反対	${}_NC_r$
	0	9	1通り
実際に	1	8	9
生じた→	$\boxed{2}$	$\boxed{7}$	$\boxed{36}$
組み合わせ	3	6	84
	4	5	126
	5	4	126
	6	3	84
	7	2	36
	8	1	9
	9	0	1
		合計	512通り

専門用語で，
この ${}_NC_r$ の数の出方を二項分布と呼ぶ。
　データの2値すなわち二項（賛成・反対）の出現確率によって数の出方が決まってくる。

②　実際に生じた組み合わせ「2,7」の出現確率を計算する。

上の表からわかるように、偶然の出現（すなわち等確率の出現）を仮定するなら「4,5」や「5,4」がもっとも出現しやすい。極端に人数が偏るケースは偶然には生じにくい。ここで、実際に生じた組み合わせ「2,7」の数は 36 通りあるので、その出現確率は 36/512＝.070 となる。すなわち、「2,7」の組み合わせは偶然下では 7%の確率でしか生じない。

③　「2,7」以外の帰無仮説に反する組み合わせをすべて数える。

実際の「2,7」という結果は偶然下では 7%しか生じないことが分かった。しかし、それを反証として「2,7」は偶然の結果ではないと主張するなら、「1,8」と「0,9」も偶然の結果ではないことになる。そこで、それらの出現数も、偶然でないケースとして数え入れなければならない（36＋9＋1＝46 通りになる）。同様に、人数の偏りが「2,7」方向だけでなく、それと同程度に反対の「7,2」方向へ偏る場合も偶然でないケースとみなす必要がある。そこで、また「7,2」・「8,1」。「9,0」の数も加え入れなければならない（便宜的に上で求めた 46 を二倍すればよいので、総計 46×2＝92 通りになる）。かくして、「2,7」に代表される人数の偏りが偶然下で出現する確率は、92/512＝.180 である（p＝.180 とおく）。

④　p＝.180 を有意水準に対照して有意性を判定する。

統計的検定における公認の有意水準は 5%である。これに対照すると、p＝.180 は有意傾向の範囲（5~10%）にも入らない。したがって、9 人の住人が賛成 2・反対 7 に割れるケースは有意でない（両側検定：このコトバを必ず付けること）。すなわち、今回の割れ方は偶然にも十分起こりうるので、とくに反対多数であるとは断言できない。試案 A の帰趨（きすう：帰着の意）はさだかでないとみるべきである。ところで、もし、賛成 1・反対 8 であったなら有意に反対者が多いと言えた（これを、確かめてください）。

<u>直接確率計算の別方法</u>

1×2 表はデータの値が 2 値であるので、表中の人数または度数の出力は二項分布の確率密度にしたがう。そこで、二項分布の確率密度の計算式を利用して、ある人数または度数が出現する確率を直接に計算することもできる。

$p={}_NC_r \cdot P_1{}^r \cdot P_2{}^{N-r}$

上式の ${}_NC_r$ は組み合わせ関数であり、N から r を取り出す組み合わせの数を与える。前例では N=9、r＝2 であった。P_1 は一方の値（賛成者）の帰無仮説上の出現確率であり、前例の場合は偶然の 50%・50%を仮定していたので、P_1=.50 である。同様に P_2 も他方の値（反対者）の出現確率であり、$P_2=1-P_1$=.50 である。したがって、前例の「2,7」の出現確率は下のように計算できる。

$p={}_NC_r \cdot P_1{}^r \cdot P_2{}^{N-r}={}_9C_2\times(.50)^2\times(.50)^{9-2}=.070$

これに、「1,8」・「0,9」の出現確率を加え、さらに、反対方向の同程度の偏り「7,2」・「8,1」・「9,0」の出現確率を加えると、前例の p＝.180 を得ることができる。

「2，7」の出現確率：$p={}_9C_2\times(.50)^2\times(.50)^{9-2}=.070$

「1，8」の出現確率：$p={}_9C_1\times(.50)^1\times(.50)^{9-1}=.018$

「0，9」の出現確率：$p={}_9C_0\times(.50)^0\times(.50)^{9-0}=.002$

小計 .090

さらに，「7，2」・「8，1」・「9，0」の出現確率の加算として便宜的に上の小計を二倍する：　$p=.090\times2=.180$

さらに、「7,2」・「8,1」・「9,0」の出現確率の加算として便宜的に上の小計を二倍する：p＝.180

この計算の利点は 2 値の帰無仮説上の出現確率がフィフティ・フィフティでない場合でも対処できることである。以下に事例をあげる。

<u>1×2 表の事例：帰無仮説上の出現確率が</u>

0.50 でない例

勝率 3 割の少年野球チームがある。夏休みを使って、臨時コーチをたのみ、1 か月間の特訓をおこなった。その後の終期リーグの成績は 6 勝 2 敗であった。特訓の成果があったと言えるか。

表 104　秋期リーグの勝敗

勝ち	負け	試合数
6	2	8

この場合の帰無仮説は「勝率は以前と変わらない」である。したがって、以前の勝率は 3 割であるから「勝ち」の帰無仮説上の出現確率は.3.「負け」の帰無仮説上の出現確率は.7 となる。そこで、直接確率計算をおこなってみる。

6 勝 2 敗の出現確率：$_8C_6 \times (.3)^6 \times (.7)^{8-6} = .0100$

7 勝 1 敗の出現確率：$_8C_7 \times (.3)^7 \times (.7)^{8-7} = .0012$

8 勝 0 敗の出現確率：$_8C_8 \times (.3)^8 \times (.7)^{8-8} = .0001$

小計 .0113

$\therefore\ p = .0113 \times 2 = .0226$

以上より、6 勝 2 敗という勝敗の偏りは有意である（両側検定：$p < .05$）。したがって、特訓の成果があったと言える。かりに、もし有意または有意傾向でなかったら（$p > .10$）、その程度の勝敗の偏りは偶然に起こりうる範囲であり、とくに特訓の成果というほどのことではないことになる。

二項検定の表

以上のような、組み合わせ関数や二項分布の計算式を用いた直接確率計算を「二項検定」と呼ぶことがある。この検定を使う時に役立つように、2 値の出現確率が 50%・50%のときの有意な組み合わせの表を以下に掲載する（N＝36 まで)。表の見方は、データ N（人数または度数の合計）と一致する行を左から右に見てゆく。集計表の度数の別れ方が下表のパターンと同じか、それよりも偏っているところで、真下の出現確率を読み取る。どのパターンよりも偏りが小さいなら、有意でない（$p > .10$)。

二項検定の表（$p_1 = .5$, $p_2 = .5$）

N	Nの有意な分かれ方のパターン			N	Nの有意な分かれ方のパターン		
5	0:5	—	—	21	6:15	5:16	4:17
6	—	0:6	—	22	6:16	5:17	4:18
7	—	0:7	—	23	7:16	6:17	4:19
8	1:7	—	0:8	24	7:17	6:18	5:19
9	—	1:8	0:9	25	—	7:18	5:20
10	—	1:9	0:10	26	8:18	7:19	6:20
11	2:9	1:10	0:11	27	8:19	7:20	6:21
12	—	2:10	1:11	28	9:19	8:20	6:22
13	3:10	2:11	1:12	29	9:20	8:21	7:22
14	3:11	2:12	1:13	30	10:20	9:21	7:23
15	—	3:12	2:15	31	10:21	9:22	7:24
16	4:12	3:13	2:14	32	10:22	9:23	8:24
17	—	4:13	2:15	33	11:22	10:23	8:25
18	5:13	4:14	3:15	34	11:23	10:24	9:25
19	5:14	4:15	3:16	35	12:23	11:24	9:26
20	—	5:15	3:17	36	12:24	11:25	9:27
出現確率	.10	.05	.01	出現確率	.10	.05	.01
有意水準	有意傾向	5％	1％	有意水準	有意傾向	5％	1％

（注） N＞36のときはp.242の χ^2 検定を代用してもよい。

備考：集計表におけるパーセント表示の問題

人数や度数などのデータ表示は集計表によっておこなうが、特に割合を示したいときにパーセントを併記することが多い。その場合、あくまで実際の度数の表示を優先して、パーセント表示はカッコ書きとして併記する。なぜなら、パーセントに対して検定を行ってはならないからである。例えば同じように「75%と 25%」という割合であっても「度数 6 と度数 2」は有意でないが、「度数 15 と度数 5」は有意である。このように、小規模データではパーセント表示は有意性の判定を誤らせる。したがって、どうしてもパーセントだけを表示したい場合は、必ず度数の合計（N）を明記することであるが、原則として実際の度数の表示を優先する方が望ましい。

また、2×2 表や 2×3 表など、一般に i×j 表の場合は、特にパーセント計算に注意が必要である。すなわち、タテのカテゴリーとヨコのカテゴリーの意味を考えて、どちらの割合を計算するかを決める。例えば下表は、初めて新しい服を着たときに、他人の目が不快であるかどうかを男女に訪ねたときの回答結果であるが、この場合、

男女数が異なるので割合でないと比較しにくい。そこでパーセントを併記する。これによって男性よりも女性の方が快であると答える人が多いということが分かる。タテについてパーセントを求めると解釈を誤る。

表　新しい服装を見られることの快・不快回答

	快	どちらでもない	不快	合計
男性	10(25)	20(50)	10(25)	40
女性	100(67)	40(27)	10(7)	150
(注) カッコ内%				

b.　1×j 表の分析：χ^2検定

適用

1×j 表は「1 条件×3 値以上」の集計表である。(j は 3 以上の任意の数)。この形の有意差の分析にはχ^2検定を用いる。“χ^2”は「カイ自乗」と読む。“χ”は英語のエックスではなく、それに相当するギリシャ語のカイである。

χ^2検定の考え方

1×3 以上の表の分析であるが、χ^2検定の説明を単純にするため、最初に 1×2 表を例として取り上げる。

表 105　1 条件内 2 値の度数

	値 A	値 B	合計
観測度数	4	16	20
期待度数	10	10	20

上表の観測度数とは、実際に値 A・値 B をカウントした数である。また、期待度数とは、帰無仮説が期待する数であり、この場合は偶然を仮定して値 A と値 B も等比率で出現するとみているので、期待度数は同数ずつの 10 となる。このように、表 105 はすでに検定に備えている。すなわち、検定の焦点は、観測度数と期待度数のズレが大きいかどうかにある。例えば、値 A は偶然下では 10 出ると期待されたが、実際には 4 しか観測されなかった。逆に、値 B は期待度数 10 のところ 16 観測された。このズレの大きさが偶然に出現する範囲におさまっているか否かを判定するのが、χ^2検定である。

このためには、まず観測度数と期待度数のズレを数量化しなければならない。ここで、χ^2検定では、χ^2と言われるズレの統計量を計算する。

$$\chi^2 = \overbrace{\frac{(\text{観測度数}-\text{期待度数})^2}{\text{期待度数}}}^{\text{値Aのズレ}} + \overbrace{\frac{(\text{観測度数}-\text{期待度数})^2}{\text{期待度数}}}^{\text{値Bのズレ}}$$

$$= \frac{(4-10)^2}{10} + \frac{(16-10)^2}{10}$$

$$= 7.2$$

右式の分子は、直接に観測度数と期待度数のズレ、すなわち差を求めている。ただし、差の＋・－を消すために二乗している。分母に再び期待度数を置いているのは、分子で求めたズレの大きさを度数 1 個分に換算するためである。こうして、値 A の度数 1 個につき生じたズレと、値 B の度数 1 個につき生じたズレを足し合わせた数量（7.2）がχ^2の値となる。そこで、このχ^2=7.2 というズレの大きさが偶然に出現する確率を調べてみて、それが 5%未満しかなければ偶然に出現する大きさではないと認める（有意）。これにはχ^2分布と言われる偶然分布を用いて、それに 7.2 を対照すればよい。ところで、この「χ^2という統計量」は度数のズレを表すが、本来の「χ^2分布」は正規分布をベースにした連続量のズレの偶然分布である。したがって、上の計算式で求める「χ^2という統計量」は近似的にしか本来の「χ^2分布」にしたがわない。N が小さいと、この近似が悪くなると言われている。

1×j 表の検定例Ⅰ：期待度数が等しい場合

以下に、事例において、1×j 表のχ^2検定を説明する。小学 4 年生 86 人を対象に、先生にほめられるとしたら、どんなほめられ方を好むかを 4 つの選択肢（名前発表・能力賞賛・努力賞賛・賞状授与）のなかから選ばせた。したがって、データは 4 値の名義尺度であり、結果は 1×4 の集計表となった。この結果から、特に小学 4 年生の好むほめられ方があると言えるか。

表 106　小学 4 年生のほめられ方の好み（人）

名前発表	能力賞賛	努力賞賛	賞状授与	合計
11	17	25	33	86

1）期待度数を算定する。

この場合、ほめられ方の好みに差はないと仮定する。したがって、各ほめられ方を好む児童は同数ずつの 21.5 人であると期待される（期待度数＝合計/値数＝86/4＝21.5）。

2）χ^2を計算する。

$$\chi^2=\frac{(\text{観測度数}-\text{期待度数})^2}{\text{期待度数}}+\cdots$$

$$=\frac{(11-21.5)^2}{21.5}+\frac{(17-21.5)^2}{21.5}+\frac{(25-21.5)^2}{21.5}+\frac{(33-21.5)^2}{21.5}$$

$$=12.79$$

3）下式を用いて自由度（df）を計算する。

df＝値数－1＝4－1＝3

ここで、自由度はχ^2 の大きさを左右する指数である。例えば、上の2）のχ^2の計算では、各ほめられ方ごとに度数 1 個分のズレを計算して、それらを合計した。したがって、χ^2は最終的に度数 4 個分のズレの大きさになっている。このように、入手する名義尺度データの値数が多ければ多いほどχ^2 は大きくなる。そこで、χ^2 を対照すべきχ^2 分布も、それぞれの自由度に応じた適切なものを選ばなければならないのである。なお、「自由度＝値数－1」とせずに、値数をそのまま指数としてもよいのであるが、値数＝1 の場合は比較にならない。そこで値数＝2 を最小の指数（1）に換算し、以下、それに準じて 1 を引いている。このやり方は一般に統計量の出方を左右する指数（自由度）を求める際に妥当である。

4）χ^2分布表を調べ、χ^2の出現確率を読み取る。

χ^2＝12.79 を df＝3 のχ^2分布に対照する。

表107　χ^2 分布（一部：全体表は p. 299）

df	χ^2 値		
3	6.251	7.815	11.345
出現確率	.10	.05	.01
有意水準	有意傾向	5％	1％

χ^2＝12.79 は、←右端の値を超える。そこで真下の出現確率を見て、p＜.01 と読む。

以上より、12.79 という大きさは偶然下では 1%未満の出現確率しかもたないことがわかった。したがって、偶然に出現した大きさではないと考える方が自然であろう（有意であると認める）。

<u>検定結果の論文記載例</u>

「表 106 は、各ほめられ方ごとに、それをもっとも好む児童の人数を示したものである。χ^2検定の結果、人数の偏りは有意であった（χ^2（3）＝12.79、p＜.01）。表 106 によると「名前発表」はあまり好まれず、「賞状授与」は特に好まれると言える。」

上の文章中、第 1 文がデータ表示、第 2 文がデータ分析である。「χ^2（3）＝12.79」のカッコ内の数字は自由度である（df＝3）。χ^2検定の結果の記述は、「有意差があった」というよりも、上文のように、「人数の偏りが有意であった」というほうが適切である。なぜなら、χ^2 検定が指摘することは集計表全体としてのアンバランス（度数の出現比率が等しくないこと）であり、各観測度数ごとにみれば期待度数とあまり差がないものも含まれているからである。そこで、第 3 文はそれぞれの観測度数についての検討となる。この検討は、基本的には、各観測度数と期待度数とのズレの大きさを基にした推論になる。本例のように、期待度数の同数ずつの場合は、観測度数を直接に比べてみることができる。なお、χ^2 検定は実質的に両側検定であるので、そのコトバを付記する必要はない（計算式の分子を二乗して、ズレの方向性を消しているので）。

<u>χ^2検定の制約</u>

ここでχ^2検定の制約について述べておく。前に少しふれたが、実は、χ^2（統計量）はχ^2分布に近似的にしかしたがわない。すなわち、χ^2統計量とχ^2分布とは、本来別物である。したがって、χ^2統計量の偶然分布として、近似的にχ^2分布を用いることには一定の制約がある。一般に、集計表において、5 以下の期待度数があると近似が悪いと言われる。ただし、1×5 以上の集計表の場合は 5 以下の期待度数が一つあってもよい（二つ以上あると悪い）。残念ながら、この制約に抵触した場合の簡単な対処法は確立されていない。イェーツの修正を施せばよいという意見もあるが、1×3 以上の表には適用できないという意見もあり、決着をつけがたい。また、そもそも、近似が悪化する基準についても疑問がないとは言えない。このように、χ^2統計量とχ^2分布との近似をめぐる問題は複雑な状況にある。

そこで、ここではひたすら観測度数を増やして期待度数を 6 以上にあげることをすすめる。例えば、期待度数の出現を等比率とした場合、集計表の大きさに 6 をかけた数だけ観測度数を確保すれば問題はない（1×3 表の場合なら 1×3×6＝18、1×4 表の場合なら 1×4×6＝24)。それができないとき（1×4 表で N＝20 のようなとき）は、例えば 1×4 表の 4 値を 2 値ずつまとめて、1×2 表に作りかえるという方法がある（期待度数は 10 ずつになる)。ただし、内容の妥当性や研究のねらいがくずれるようなら好ましくない。

1×j 表の検定例Ⅱ：期待度数が等しくない場合

成人スピーチの資料を分析し、言い誤りを拾い出した。表 108 は、スピーチの構文単位ごとに言い誤りの発生数を集計したものである。言い誤りは特定の箇所に発生しやすいと言えるか。

表 108　スピーチにおける言い誤りの発生数

	主部	述部	その他	合計
観測度数	14	24	35	73
期待度数	17	28	28	73

1）期待度数を算定する。

別の集計において、スピーチの資料は、主部 23.3%、述部 38.3%、その他の句 38.3%からなっていることがわかったとする。したがって、もし言い誤りがそれらの構文単位に等比率で発生するなら、これらの%で発生するはずである。そこで、言い誤りの合計数 73 に各構文単位の%をかけて期待度数とした。なお、「その他」は期待度数 5 以下の構文単位を含めた「寄せ集め」である。このようにして、χ^2の制約をかわす方法もある。

2）χ^2を計算する。

$$\chi^2=\frac{(\text{観測度数}-\text{期待度数})^2}{\text{期待度数}}+\cdots$$
$$=\frac{(14-17)^2}{17}+\frac{(24-28)^2}{28}+\frac{(35-28)^2}{28}$$
$$=2.85$$

3）df を計算する。

df＝値数－1＝3－1＝2

4）χ^2分布表を調べて、χ^2の出現確率を求める。

χ^2＝2.85 を df＝2 のχ^2分布に対照する（表 109)。

表109　χ^2分布（一部：全体表は p. 299)

df	χ^2 値		
2	4.605	5.991	9.210
出現確率	.10	.05	.01
有意水準	有意傾向	5％	1％

χ^2＝2.85 は，←左端の値を超えられない。したがって，p>.10 と読む。

検定の結果、2.85 は偶然分布内で 10%以上の出現確率をもつことがわかった。これでは「めったに出現しない」というわけにはいかない。すなわち、χ^2＝2.85 という程度の大きさは偶然のユレとしても十分にあらわれうる。

結果の論文記載例

「表 108 は、成人のスピーチに発生した言い誤りを構文単位ごとに集計したものである。なお、表中の期待度数は、スピーチの資料に含まれる各単位の割合に基づいて算定してある。χ^2 検定の結果、度数の偏りは有意ではなかった（$\chi^2(2)=2.85$、$p>.10$)。したがって、言い誤りは、主部・述部などの構文単位にとらわれず随時に発生する現象であることが示唆される。」

① 2 条件×2 値のデータ

被験者	条件	観測値
イ	女性	喫煙
ロ	男性	非喫煙
ハ	女性	非喫煙
ニ	男性	喫煙
ホ	男性	喫煙

⇩

	喫煙	非喫煙
男性	2	1
女性	1	1

② 2 値×2 値のデータ

被験者	前の値	後の値
ヘ	喫煙	非喫煙
ト	喫煙	非喫煙
チ	非喫煙	非喫煙
リ	喫煙	喫煙
ヌ	非喫煙	非喫煙

⇩

前＼後	喫煙	非喫煙
喫煙	1	3
非喫煙	0	1

2×2 表の分析

コースの分岐

2×2 表の場合には、同じ形をしていても二種類の性格の異なる表があり、分析の目的も異なる。一つは、下の①のような「2 条件×2 値」の集計表である。もう一つは、②のような、どちらも観測値であるような「2 値×2 値」の集計表である。

左の①の場合、分析の目的は、男性・女性の 2 条件間で喫煙・非喫煙の割合を比較することである。これには直

接確率計算を用いる。右の②の場合、分析の目的は、前の値から後の値への二通りの変化（喫煙→非喫煙 vs.非喫煙→喫煙）を比較することである。すなわち、上表の例では、3 と 0 の度数のみを比べる。これには、マクネマーの法を用いる。

c.　2×2 表の分析Ⅰ：条件間の比較

適用

　この場合の 2×2 表とは「2 条件×2 値」の集計表のことである。値の変化を分析するための「2 値×2 値」の集計表は後述する。

用語の紹介

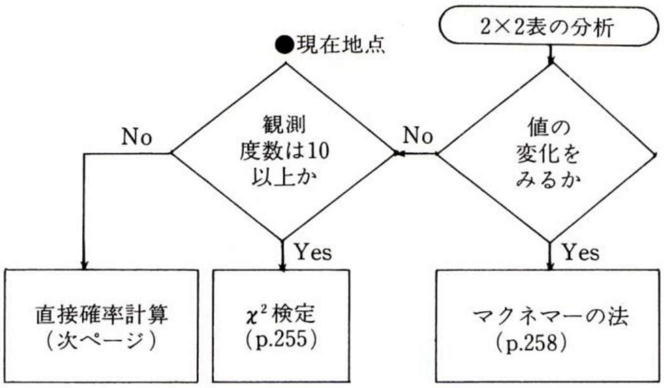
　今後、2×2 以上の集計表では、セルと周辺度数を名指す必要がある。「セル」とは表中のマスのことである。「周辺度数」とは表外の小計のことである。

	喫煙	非喫煙	計
男性	セル	セル	周辺度数
女性	セル	セル	周辺度数
計	周辺度数	周辺度数	総計

表からわかるように、2×2 はセルの数になる。

方法の選択

2×2 表の分析の方法としては、観測度数 10 以下のセルがある場合は直接確率計算、それ以外は χ^2 検定を用いる。



直接確率計算：観測度数 10 以下のセルがある場合

直接確率計算は、考案者の名にちなみ、「フィッシャーの検定」（Fisher's exact test）とも呼ぶ。χ^2 検定は近似であるが、直接確率計算は真正の出現確率を計算するので、計算量が膨大にならない限りは、これを用いるほうが結論を誤らない。以下、事例によって説明する。さて、19 人の生徒を無作為に選び、英語担当の教師個人について関心があるかないかを尋ねた。さらに、英語の授業が面白いかどうかを尋ねた。表 110 は、その結果を集計したものである。教師への関心の有無が授業への興味の有無と関係していると言えるか。

表 110　教師への関心と授業への興味の関係

教師/授業	興味あり	興味なし	計
関心あり	5	2	7
関心なし	4	8	12
計	9	10	19

（注）データ表示としては周辺度数は必要ない。

集計表の形としては、まず教師への「関心あり・なし」で生徒を 2 条件に分け、次に各生徒について授業への「興味あり・なし」の 2 値を観測したいという結果である（逆の見方も可）。すなわち 2 条件×2 値の条件間の比較である。ここで、検定上は「条件に差がない」とみる（帰無仮説）。したがって、関心あり・なし条件の生徒（周辺度数の 7 人と 12 人）は、どちらも 9 : 10（興味あり・なしの周辺度数）で異なる値に分かれるであろう。すなわち、各セルの度数は、周辺度数の相対比率にしたがって観測されるはずである。

　実際に、特定の人数の分かれ方が何通りあるかを計算してみる。

　1）実際に生じた人数の分かれ方は何通りあるのかを計算する。

　組み合わせ関数 ${}_{N}C_{r}$ を用いて以下のように計算する。

２）実際に生じた結果より偏っている分かれ方も計算する。

周辺度数を固定した状態で、最小度数のセルを１ずつ減らして極端に偏る組み合わせをつくり、それが何通りあるかを計算する。

興味なし	計
2	7
8	12

（i）　7から2を取り出す組み合わせ：${}_{7}C_{2}=21$ 通り
（ii）　12から8を取り出す組み合わせ：${}_{12}C_{8}=495$ 通り
（iii）　iの1通りそれぞれに対して，iiの495通りすべてが組み合わさるので，左表の集計になるケースは全部で 21×495＝10395 通りある。

以上、実際の結果を含めて「偏った分かれ方」をする総数は、これらを足しあげ、10395＋1540＋66＝12001 通りあることがわかった。なお、本例では、たまたま「ありーあり」・「なしーなし」の度数に偏ったが、同程度に偶然にさからう偏り方として「ありーなし」・「なしーあり」の度数が大きくなるという場合も想定できるので最初から一方に限定できない（両側検定)。このため、上の総数の２倍をもって偶然への総反証例数とする（12001
×2
24002)。

興味なし	計
2	7
8	12
10	

①で計算ずみ
10395 通り

⇨

興味なし	計
1	7
9	12
10	

${}_{7}C_{1}\times{}_{12}C_{9}$
＝1540 通り

⇨

興味なし	計
0	7
10	12
10	

${}_{7}C_{0}\times{}_{12}C_{10}$
＝66 通り

３）「偏った分かれ方」が出現する確率を計算する。

さて、実際の結果を含めた偏った分かれ方の総数は、24002 通りあることがわかった。このとき、全生徒数 19 人のうちから 10 人が「興味なし」の値をとる分かれ方は、${}_{19}C_{10}$＝92378 通りである。この 92378 通りのなかに、上で計算した「偏った分かれ方」も含まれているわけである。そこで、「偏った分かれ方」の出現確率は「24002/92378＝.260」となる（p＝.260 とおく)。これは公認の有意水準に達しない。

<u>検定結果の論文記述例</u>

「表 110 は、教師への関心あり・なし別に、授業への興味がある生徒・ない生徒を集計したものである。直接確率計算を行った結果、人数の偏りは有意でない（両側検定：p＝.260)。したがって、教師への関心と授業への興味に関連性があるとは言えない。」

計算のまとめ

本例の手順を一つの計算式にまとめておく。なお、最小度数のセルを含む小計欄ならヨコの周辺度数を用いても p は同値になる。

両側検定なので
二倍することを忘れずに

$$p=\frac{\overbrace{{}_{7}C_{2}\cdot{}_{12}C_{8}+{}_{7}C_{1}\cdot{}_{12}C_{9}+{}_{7}C_{0}\cdot{}_{12}C_{10}}^{\text{最小度数が0になるまで計算する}}}{{}_{19}C_{10}}\times 2$$

総度数　周辺度数

なお、p を２倍する方法は各度数が対角線対称でない場合、正確でない。

<u>2×2 表の χ^2 検定：観測度数 10 以下のセルがない場合</u>

大学生 100 人を対象に向性検査をおこない、かつ、好きな色を回答させた。

表 111 は、内向性・外向性ごとに後退色（例；黒色）を好む者と前進色（例：黄色）を好む者の人数を示している。性格と色の好みとは関連すると言えるのか。

表 111　内向性・外向性による色の好み（人）

	後退色	前進色	計
内向性	17	13	30
外向性	23	47	70
計	40	60	100

１）χ^2 を計算する。

ここでは、期待度数を算定しない簡易式を紹介する。

２）自由度（df）を計算する。

2×2 表の場合は、つねに df=1 である。一般に、i×j 表の自由度は df=（i－１）×（j−1）である。χ^2分布表への対照は省略するが、内向性・外向性と後退色・前進色の好みには有意な関連性がある（χ^2（1）=4.96、p＜.05）。すなわち、表中の度数 17 と 47 が、23 と 13 に比べて相対的に大きい。

$$\chi^2 = \frac{\overbrace{(17\times47-13\times23)}^{\text{ななめどうしのセルをかけあわせ，引いて二乗する}}{}^2\times\text{総度数}}{\text{周辺度数}\times\text{周辺度数}\times\text{周辺度数}\times\text{周辺度数}}$$

$$=\frac{(17\times47-13\times23)^2\times100}{30\times70\times40\times60}$$

$$=4.96$$

直接確率計算を嫌った事例

事例

論文名「対象者の心理状態による心理検査の結果のちがい」

目的

対象者がフラストレーション状態で検査を受ける場合と、不満のない平衡状態で検査を受ける場合の結果を比較した。心理検査としては認知型テストを用いた。

方法

対象者：小学三年生 21 人（男子 13 人、女子 8 人）

道具：箱庭づくりの小道具および衝動・熟慮の認知度テスト。

手続き：検査は個別に行った。まず、対象者に箱庭作りをさせた。そして興味が盛り上がったところで強制的に中断して認知型の検査に移る条件（中断条件）と、最後まで箱庭を作らせてから認知型の検査に移る条件（完遂条件：かんすい）とを設定した。

結果

表 112 は、中断・完遂条件別に、衝動型または熟慮型と判定された対象者の人数を集計したものである。

表 112　中断・完遂条件別の衝動・熟慮型の人数

	衝動型	熟慮型
中断条件	7	2
完遂条件	4	8

χ^2検定の結果、人数の偏りは有意であった（χ^2（1）=4.07、p＜.05）。中断によるフラストレーションは、完遂による心理的平衡状態よりも、認知型テストの衝動型反応を増加させると言える。

解説

集計表は、2 条件（中断・完遂）×2 値（衝動・熟慮）の形である。この場合の分析方法は、基本的に直接確率計算であるが、計算量の多さと煩雑さを嫌って、本例ではχ^2検定を用いている。しかし、最小度数が 10 以下のセルがある場合には、めんどうでも直接確率計算を行う方が安全である。なぜなら、直接確率計算による出現確率は真正であるが、χ^2検定のそれは近似的にすぎず、特に総度数が小さい場合は、この近似が悪くなるからである。例えば、この例では検定の結果は有意であったが、これはχ^2検定の近似の悪さによる誤りである。直接確率計算によれば、真正の出現率は「p=.080（両側検定）」となり、有意でない。したがって、箱庭作りの中断・完遂による認知型反応の差異は偶然に起こる範囲であり、積極的な違いを認めることはできない。このように、χ^2検定の近似の悪さは出現確率を小さくする（すなわち結果を有意にする）方向に向かうので、注意が必要である。

d.　2×2 表の分析Ⅱ：値の変化の比較

適用

この場合の 2×2 表とは「2 値×2 値」の集計表のことである。前述の 2×2 表は独立の 2 条件の比較であったが、この場合、形は同じでも、特定の１条件における２つの値の変化の比較である。ここで用いるマクネマーの法は一つの考え方であり、新しい方法ではない。すなわち、「2 値×2 値」の 4 セルのうち、変化を表す 2 セルだけに注目して、1×2 表につくりかえるのである。

さて、新商品の宣伝フィルムを見る前と見た後で、その商品に対する購買動機が変化するかどうかを調べた。表

113 は、その結果の集計である。宣伝フィルムの効果があったと言えるか。

表 113　CM による購入動機の変化（人）

前/後	買いたい	買わない
買いたい	23	7
買わない	17	20

1）上の 2×2 表から 1×2 表をつくる。

この場合、宣伝フィルムが購入動機にどんな変化をもたらしたかを分析したいので、変化のないセルは注目しない。そこで、「買いたい→買わない」と「買わない→買いたい」のセルだけを取り出し、あらたに、1×2 表をつくる。

表 114　購入動機が変化した者の人数

買いたい→買わない	買わない→買いたい
7	17

2）1×2 表の分析方法を適用する。「二項検定の表」を利用すれば簡単である。

表 115　二項検定の表（P_1=.5、P_2＝.5）

N	Nの有意な分かれ方のパターン		
24	7 : 17	6 : 18	5 : 19＋
出現確率	.05＜p＜.10	p＜.05	p＜.01
有意水準	有意傾向	5%	1%

本例の場合、N＝24 であるので、上表の N＝24 の行をヨコにみて同一の度数のパターンをさがす。「7 : 17」があったので、真下の出現確率を読み、p＜.10 となる。これは有意傾向である。

<u>検定結果の論文の記載例</u>

「表 113 は、新商品の宣伝フィルムの視聴前後に購入動機をたずねて、その回答を集計したものである。購入動機の変化を検定した結果、「買わない」から「買いたい」へ変化した者の人数がその逆の人数よりも有意に多くなる傾向があった（両側検定：.05＜p＜.10）。したがって、この宣伝フィルムは、購入動機の低い者を高揚させるのに効果があると言える。」

このように、マクネマーの法は、「プラスの変化 vs.マイナスの変化」という 1×2 の比較である。もし、本例に 2×2 の分析方法を適用すると「変化あり vs.変化なし」という比較になる。この比較では、「変化あり」にプラス・マイナスの変化がコミにされるので、検定の結果が有意に出ても、どんな効果が認めたのか、わからない。

<u>e.　i×j 表の分析</u>

<u>適用</u>

i×j 表とは、「2 条件×3 値以上」の集計表のことである（i は 2 以上、j は 3 以上の任意の数でよい）。この形には、χ^2検定を適用する。その結果、有意であった場合には、さらに残差分析をおこなう。

<u>i×j 表のχ^2検定の例</u>

事例によって説明する。

一方通行の廊下において、壁に飾ってある絵画を鑑賞する人間を立たせた。そして、通行者が、この使用空間（鑑賞者と絵画の間）をヨコ切るかどうかを観察することにした。この際、絵画との距離を 3 条件設定し（1.5m、2m、2.5m）、それぞれ 10 分間のタイム・サンプリング（時間を限定したデータ入手）をおこなった。表 115 は、結果の集計である。使用空間の広さによって、侵入の心理的抵抗に差異があると言えるか。

表 115　使用空間の広さによる侵入者・非侵入者の人数

鑑賞距離	1.5m	2m	2.5m	計
侵入者	10（14.1）	8（11.3）	30（22.6）	48
非侵入者	15（10.9）	12（8.7）	10（17.4）	37
計	25	20	40	85

(注）カッコ内は期待度数、その算定は下式による。

あるセルの期待度数＝（そのセルのヨコの周辺度数）×（タテの周辺度数）/総度数

1）χ^2を計算する。表中では、期待度数を算定してあるが、その算定式を含む簡易式を紹介する。定義式に

よる計算と一致するか、確認する。

2）自由度（df）を計算する。

本例は「3条件×2値」であるので、自由度は以下の通り。df＝（i−1）×（j−1）＝2

3）χ^2分布表を調べ、χ^2の出現確率を求める。

χ^2＝10.55をdf＝2のχ^2分布に対照する。

$$\chi^2=\left(\frac{\text{セル度数}^2}{\text{ヨコの周辺度数}\times\text{タテの周辺度数}}+\cdots-1\right)\cdot\text{総度数}$$

$$=\left(\frac{10^2}{48\times25}+\frac{8^2}{48\times20}+\frac{30^2}{48\times40}+\frac{15^2}{37\times25}+\frac{12^2}{37\times20}+\frac{10^2}{37\times40}-1\right)\times85$$

$$=10.55$$

表116　χ^2分布表（一部：全体表はp. 299）

df	χ^2値		
2	4.605	5.991	9.210
出現確率	.10	.05	.01
有意水準	有意傾向	5％	1％

χ^2＝10.55は，←右端の値を超えている。したがって，p<.01である。

検定の結果は有意であったが、i×j表の場合、セル数が多いので、どのセルに期待以上（以下）の度数が集中しているのか把握しにくい。そこで、次に残差分析を行う。

χ^2検定の残差分析

用途

残差分析とは、一般に、i×jの集計表において、χ^2検定の結果が有意であった場合に、どのセルが、この有意性に貢献したのかを判定する方法である。ここで、残差とは観測度数の差のことであるが、以下の計算式においては、残差の値が標準正規分布にしたがうように調整する。これを「調整された残差」と呼ぶ。こうすることによって、標準正規分布を偶然分布とした検定をおこなうことができる。なお、ここでは自由度∞のｔ分布を使用する。

表115　使用空間の広さによる侵入者・非侵入者の人数

鑑賞距離	1.5m	2m	2.5m	計
侵入者	10（14.1）	8（11.3）	30（22.6）	48
非侵入者	15（10.9）	12（ 8.7）	10（17.4）	37
計	25	20	40	85

（注）　カッコ内は期待度数。その算定は下式による。

$$\text{あるセルの期待度数}=\frac{\text{そのセルのヨコの周辺度数}\times\text{タテの周辺度数}}{\text{総度数}}$$

事例

前例の2×3表を取り上げ、χ^2検定の結果が有意であったので、残差分析をおこなってみる。表115より考える。

1）各セルの「調整された残差」を計算する。

表115の観測度数15のセルの残差を計算してみる。

$$\text{観測度数15のセルの調整された残差}=\frac{\dfrac{\text{観測度数}-\text{期待度数}}{\sqrt{\text{期待度数}}}}{\sqrt{\left(1-\dfrac{\text{ヨコの周辺度数}}{\text{総度数}}\right)\cdot\left(1-\dfrac{\text{タテの周辺度数}}{\text{総度数}}\right)}}$$

$$=\frac{\dfrac{15-10.9}{\sqrt{10.9}}}{\sqrt{\left(1-\dfrac{37}{85}\right)\times\left(1-\dfrac{25}{85}\right)}}$$

$$=1.97$$

2）残差の一覧表を作成し、有意性を検定する。集計表の配置に合わせて、計算した残差を掲載する。

表117　表115の調整された残差

鑑賞距離	1.5m	2m	2.5m
侵入者	−1.97*	−1.70†	3.24**
非侵入者	1.97*	1.70†	−3.24**

表中ではすでに有意性の記号を付けたが、有意な残差の値は下の付表のように定数として決まっているので検定は容易である。

付表　残差の有意性検定

|残差|＞1.65	⇨ †	p＜.10
|残差|＞1.96	⇨ *	p＜.05
|残差|＞2.58	⇨ **	p＜.01

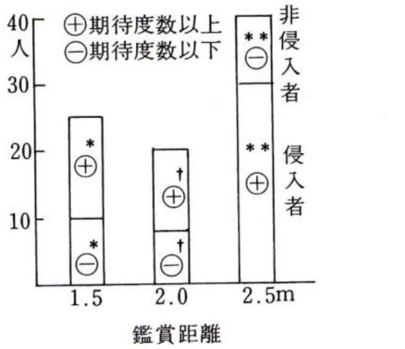
χ^2検定と残差分析の論文記載例

以下の分析結果は、次のように論文に記載する。

「表 115 は、鑑賞距離の条件別に侵入者と非侵入者の人数を集計したものである。χ^2検定の結果、人数の偏りは有意である（$\chi^2(2)=10.55$、$p<.01$）。そこで、残差分析をおこなった結果、表 117 にみられるように、2m 以下の鑑賞距離では侵入者は少なくなり、2.5m になると侵入者が増えるということがわかった。

したがって、使用空間が狭くなるにつれ、そこへの侵入にともなう心理的抵抗が大きくなると言える。」

上の文章中、第一段階の 3 文は、それぞれデータ表示、χ^2検定、残差分析の記述に相当する。ただし、χ^2検定が有意でなかった場合は、残差分析に進むことができないので、その記述もなくなる。残差分析の記述は、上例のように、本文外に「残差の一覧表」を掲載し、それに言及する形をとるほうがスマートである。その際、残差の読み取り方は、二つの観点から行う。すなわち、①残差が有意であるかどうか、②プラスの残差かマイナスの残差か。プラスの残差は期待以上の、また、マイナスの残差は期待以下の観測度数がそのセルにおちたことを意味する。段落が変わって、最後の文は、分析結果に基づく推論となる。なお、データ表示の集計表と残差の一覧表を一緒にして、作図による表示法をとることもできる。



（注）† p< .10, *p< .05, **p< .01（残差分析）

図12　鑑賞距離による侵入者・非侵入者数

i×j 表のχ^2検定の制約

以下の場合、χ^2（統計量）は、χ^2分布に近似しなくなると言われている。期待度数が 1 以下のセルが一つでもある場合。期待度数 5 以下のセルの数が全セル数の 20%を越える場合。例えば、3×4 の集計表では期待度数 5 以下のセルが 3 つ以上あると、この制約にひっかかる。（全セル数 12 の 20%は 2.4 個）。以上の場合、確立された対処法はないが、次のような対策が考えられる。①期待度数 6 以上になるまで観測度数を増やす。②小度数のセルどうしを併合する。なお、どうしても対処しきれない場合は、研究とデータの意味を考え直すほうがはやい。

間隔尺度データを名義尺度に変更した事例

事例

論文名「居住都市と地震の発生予想との関連性」

方法

対象者：地理的位置の異なる都市 A・B・C の成人以上の住民。無作為に都市 A129 人、都市 B104 人、都市 C101 人を抽出した。

手続き：各都市の街頭において、10 分間 6 セットのタイム・サンプリングをおこなった。各対象者に「もし地震がくるとしたら今から何年後に来ると思いますか」と質問し、年数で回答させた。「すぐに」とか「1 年以内」等の回答は、さらに月数をたずねて年数換算した。

結果

表 118 は、各都市の住民の地震発生予想年数を示したものである。

表 118　都市別の地震発生予想年数

	都市 A	都市 B	都市 C
N	129	104	101
X	9.4	8.9	5.8
SD	13.1	9.7	6.8

データは「5 年後」までの年数に集中し、しかも「0.6 年後」等の少数や「50 年後」・「100 年度」等の極端な値をかなり含んでいるため、等間隔な尺度とは言えない。たとえ、等間隔性を仮定しても、表中の SD からあきらかなように等分散性や分布の正規性に欠ける。そこで、予想年数を 4 段階に分けて、集計することにした。

表 119　都市別の地震発生予想（人）

	都市 A	都市 B	都市 C
1 年以内	20	19	28
1~2 年	35	23	36
3~4 年	38	31	19
5 年以上	36	31	18

表 120　各セルの調整された残差

	都市 A	都市 B	都市 C
1 年以内	−1.65	−0.55	2.30*
1~2 年	−0.33	−1.65	2.01*
3~4 年	1.02	0.97	−2.06*
5 年以上	0.82	1.23	−2.11*

上の表 119 における 3×4 の χ^2 検定の結果、人数の偏りは有意であった（χ^2（6）＝14.38、p＜.05)。そこで、残差分析をおこなった。残差の有意性によれば、都市 A・都市 B よりも、都市 C において地震発生を 1 年以内と予想する者が多く、相対的に 3 年以後と予想する者が少なかった。したがって、住民は、地震発生を居住都市の地理的位置と関係づけて予想していると言える。

解説

　本例は、間隔・比率尺度データとしての処理の予定を変更し、名義尺度データとして処理した例である。本来、データの尺度の議論や X、SD の表示は論文上にあらわれないものであるが、ここでは予定変更の経緯を示すために便宜的に記載した。表 118 の X と表 119 の度数の出方を比較すれば、この場合、X がいかに当てにならない代表値であるか理解できる。

誤って 2×3 表に組んだ事例

事例

論文名「リラクセーション用音楽の開発と効果の検証」

目的

心理的緊張をほぐして脱力と弛緩（しかん）によるリラクセーションを誘う音楽を試作し、カセットテープに吹き込んだ。その効果を検証するため、音楽なしの状態と比較することにした。

方法

被験者；大学生 129 人（男子 52 人、女子 77 人）

手続き；被験者にリラクセーション用音楽のテープか、または、無音のテープを配り、就眠時に安静状態で利用すると熟睡が得られると説明した（無音テープは可聴域外の周波数による録音であるとした。)。その翌日に、昨夜の睡眠の程度を 3 段階で回答させた。

結果

表 121 は、テープを利用しなかったものを除外して、音楽テープ・無音テープ別に睡眠の程度の回答を集計したものである。

表 121　テープ別の睡眠程度の回答結果（人）

	たいへん十分	十分	ふつう
音楽テープ	7	14	9
無音テープ	7	11	19

χ^2 検定の結果、回答の偏りは有意でなかった（χ^2（2）＝3.24、p＞.10)。したがって、今回のリラクセーション用音楽は、あまり効果のあるものとは言えなかった。

解説

本例は、データの集計を誤って 2×3 表に組んだ例であるが、以下のように 1×3 表とするのが正しく、その場合、有意となる。

「表 122 は、音楽テープを利用した者の回答の集計である。なお、期待度数は、無音テープ利用者の回答比率（表 123）に基づいて算定した（例：5.7＝音楽テープ利用者の総数 30 人×18.9%)。

表 122　音楽テープ利用者の回答（人）

	たいへん十分	十分	ふつう
観測度数	7	14	9
期待度数	5.7	8.9	15.4

表 123　無音テープ利用者の回答

	たいへん十分	十分	ふつう
観測度数	7	11	19
期待度数	18.9%	29.7%	51.4%

χ^2検定の結果、回答の偏りは有意傾向であった（χ^2（2）＝5.88、.05＜ｐ＜.10）。したがって、このリラクセーション用音楽は、睡眠の程度をより深めるのに効果的であると言える。」

本例の「無音テープ」は、プラセボ効果（偽薬効果）のような暗示的影響の出力を推定するための統制条件である。その回答比率よりも、音楽テープの回答が望ましい方向へ偏ればよいのであるから、上例のように期待度数を算定しての 1×3 に組むのが正しい。

しかし、もし「リラクセーション用音楽 vs.無音」でなく、「リラクセーション用音楽 vs.クラッシック音楽」であれば、2×3 に組むのが正しい。なぜなら、後者の場合は「効果がある・ない」の比較でなく、「効果に差がある・ない」の比較になるからである（有意でないときは効果なしでなく、効果は同程度という結論になる）。この違いは理解しにくいが、結論を変えるくらい重大である。要するに、実験条件と統制条件の比較なら期待度数を算定して、1×j に組む。2 つの実験条件の比較ならそのまま 2×j に組む。

6．2．2　連関の分析

定義

連関とは、集計表の条件と値の関連性のことである。簡単に言えば、集計表のタテとヨコの関連性である。実際のイメージとしては、連関が強い場合、集計表の対角線上のセルに度数が偏る。なお、連関を見ることができるのは、2×2 以上の集計表であり、1×2 表や 1×j 表には連関の概念はない。

χ^2と連関の関係

χ^2検定は、度数の有意差あるいは偏りを分析する。χ^2が有意であるということは、度数が各セルに偶然の期待どおりには配分されていないということである。すなわち、度数を特定のセルに引き寄せる何らかの関連性が集計表のタテ・ヨコにあることを示唆している。

このように度数の偏りと集計表のタテ・ヨコの関連性は同一現象の表裏である。χ^2が有意であれば、連関も有意であり、両者はまったく等価である。ただし、χ^2の値はデータの総度数に応じて、どこまでも大きくなるので、χ^2どうしの値の比較はできない。そこで、これを 0~1 の範囲におさめて標準化した値が連関である。連関どうしは同じ「ものさし」上にあるので相互に比較することができる。

このように、χ^2と連関とは、同一事象の異表現である。一般には、関連性の有無をみたいだけなら、χ^2検定、関連性の強さを評価または比較したいなら連関の分析をとる。連関を表す統計量には、2×2 表の場合にϕ係数（ファイけいすう）、一般に i×j 表の場合にクラメールの連関係数がある。

a.　2×2 表の連関：ϕ係数

定義

ϕ係数とは、「2 条件×2 値」の集計表において、条件と値の連関を表す統計量である。

事例

表 124 は、意見Ａと意見Ｂに対する賛否を集計したものである（ちなみに同一意見に対する賛否ではないので値の変化を表す 2×2 の表ではない）。意見Ａと意見Ｂに対する賛成・反対は、どの程度の関連性があるか。

表 124　意見Ａと意見Ｂの賛否（人）

意見 A/意見 B	賛成	反対	計
賛成	12	18	30
反対	16	4	20
計	28	22	50

1）χ^2を計算する（2×2 のχ^2検定参照）。

χ^2＝（12×4－18×16）^{2}×50/（30×20×22×28）＝7.79

ここで、χ^2が有意でなければ関連性がないので、ϕを計算しても無意味である。上例の場合は有意（χ^2（1）＝7.79、ｐ＜.01）。

２）下式によってφ係数を計算する（ただし、N は総度数)。

$\phi = \sqrt{}(\chi^2/N) = \sqrt{}(7.79/50) = .395$

$\phi = .395$ は、0~1 に対して中程度の強さであると言える。

b.　i×j 表の連関；クラメールの連関係数

定義

クラメールの連関係数は、一般に i×j 表の連関を表す統計量である。i と j は 2 つ以上の数なら任意である。例えば、クラメールの連関係数を 2×2 表に適用した場合、φ係数となる。なお、ここではクラメールの連関係数の記号として“C r”を用いる。

事例

表 125 は、クレッチマーの性格の類型論に基づいて、大学生の体格と所属学部との関連を調べたものである。両者の間の関連は強いと言えるか。

表 125　大学生の体格と所属学部との関連（人）

	工学部	教育学部	法学部	芸術学部	計
肥満型	35	27	29	9	100
細身型	14	19	22	15	70
闘士型	21	4	9	6	40
計	70	50	60	30	210

「体格と所属学部との関連があるかどうか」だけなら、χ^2 検定でよいが、関連性の強さをみたいので連関係数を計算する。

１） χ^2 を計算する（2×3 の検定参照)。

$$\chi^2=\left(\frac{35^2}{100\times70}+\frac{27^2}{100\times50}+\frac{29^2}{100\times60}+\frac{9^2}{100\times30}+\frac{14^2}{70\times70}+\frac{19^2}{70\times50}+\frac{22^2}{70\times60}+\frac{15^2}{70\times30}+\frac{21^2}{40\times70}+\frac{4^2}{40\times50}+\frac{9^2}{40\times60}+\frac{6^2}{40\times30}-1\right)\times210$$

$$=17.38$$

ここで、やはり χ^2 が有意でなければ連関係数を求めることができない。この場合、有意であった（χ^2（6）＝17.38、p＜.01)。

２）下式によって連関係数 C r を計算する。

$$C_r=\sqrt{\frac{\chi^2}{N\times(\min.i-1,j-1)}}$$

N：総度数

min.i－1, j－1：

i－1 と j－1 の小さいほうの値

$$=\sqrt{\frac{17.38}{210\times(3-1)}}$$

$$=.203$$

C r ＝.203 は、0~1 に対比して弱い連関であると言える。すなわち、体格と所属学部との関連性は全体として有意であるが（χ^2 検定の結果による)、連関係数の低さをみると、3×4 表の中で観測度数と期待度数が大きくズレたセルは、そんなに多くなかったとみられる。このように、連関の分析は、全体的関連性の評価になる。もし、個々のセルの検討を目的にするなら、χ^2 検定→残差分析、というコースをとったほうがよい。

連関の分析結果の論文記載例

前例の論文記載例を示す。まず、2×2 表のφ係数の例である。

「表 124 は、意見Ａと意見Ｂに対する賛否を集計したものである。φ係数は.395 であり、有意であった（χ^2（1）＝7.79、p＜.01)。連関の強さは中程度と言える。表 124 によれば、意見Ａに賛成する者は意見Ｂに反対し、逆に、意見Ａに反対する者は意見Ｂに賛成する、という傾向がみられる。両意見は、ある程度の対立的な主旨を含むものであると考えられる。」

上文からわかるように、φ係数の有意性はけっきょく χ^2 の有意性である。そのように、連関と χ^2 は、たんに表現様式の違いにすぎない。しかし、連関の利点は、強さの判定ができることである。すなわち、ここでのφ係数の強さ（中程度）が、そのまま第二段落の推論の強さ（ある程度の対立的主旨）を保証している。次は、クラメールの連関係数を用いた例である。

「表 125 は、大学生の体格と所属学部を調査した結果である。クラメールの連関係数を計算した結果、.166 で

あり、有意であった（χ^2（6）＝17.38、p＜.01)。しかし、連関の強さは弱く、体型と所属学部とは、それほどの関連性があるとは言えない。」

i×j 表の連関の場合は、必ずクラメールの連関係数であることを明記すること（連関係数は他にもあるので)。なお、結果の解釈としては「わずかながらも関連性がある。」と主張することもできる。

連関の分析の制約

連関の制約は、χ^2の制約と同じである。すなわち、χ^2の近似が悪くなる場合は、ϕ係数もクラメールの連関係数も意味を失う。χ^2の有意性が、それらの係数の有意性である。研究において、総度数が小さい（χ^2の近似が悪い）ときは、連関によって全体傾向を推測しても当てにならない感じである。経験的な目安としては、2×2 表で総度数 50 以上はないと連関を求める動機がわかない。

備考；ϕ係数の計算式の構造

χ^2自体の値は「無連関」で$\chi^2=0$（観測度数と期待度数が等しい)、「完全連関」で$\chi^2=N$（総度数）となる。以下の 2×2 表において実際にχ^2を計算してみれば確認できる。

そのように、0〜N まで変化するχ^2を、0~1 に標準化するためには、χ^2を N で割ればよい。ただし、χ^2はズレの二乗値であったので、このままでは 1 方向にスソの延びた L 字型分布となる。そこで、√をかけてならしている。

$\phi=\sqrt{(\chi^2/N)}$

無連関：$\chi^2=0$

10	10	20
10	10	20
20	20	40

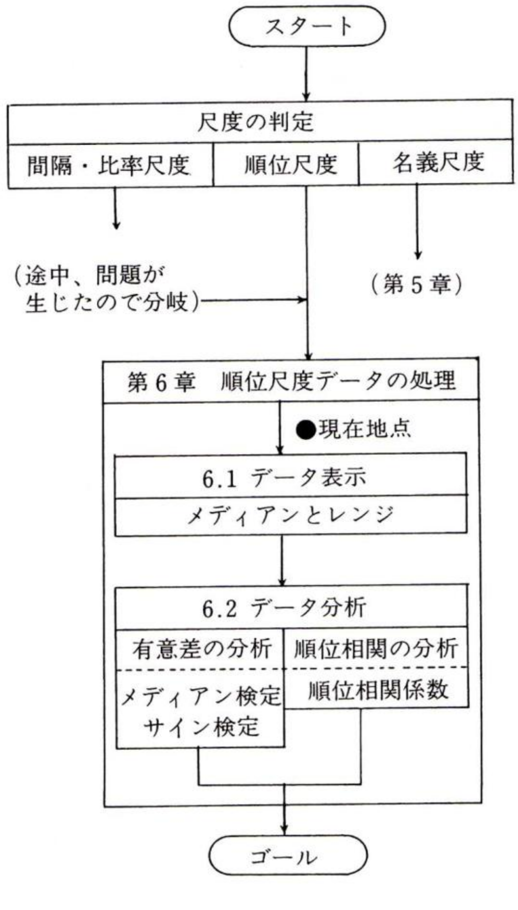
完全連関：$\chi^2=40$

0	20	20
20	0	20
20	20	40

7. 順位尺度データの処理

データの特徴

順位尺度は、間隔をもたない尺度である。したがって、データ自体は数値であるが、通常の四則演算の等式はすべて成立しない。このため、間隔・比率尺度なみの高度な処理は期待できない。もっと基本的には、平均を求めることさえ無意味である。しかし、その反面、順位尺度は「間隔フリー」であるがゆえに、次のようなデータ処理上の長所をもっている。①正規分布を前提にしない。②極端値に強い。③少数データを扱える。



データを順位尺度に落とす場合

このように、順位尺度は、データへの制約が少ない。そこで、間隔・比率尺度として問題のあるデータを、故意に順位尺度に「落として」処理する場合がある。以下のような場合である。

1）顕型尺度・元型尺度の問題にひっかかった場合

例えば、テスト得点の 70 点、80 点、90 点は数値上は等間隔であるが、70 点を 80 点に上げる努力と 80 点を 90 点に上げる努力が同程度でないとしたら、努力の測定値としては等間隔ではない。すなわち、データ（顕型尺度）の間隔と、測定対象（元型尺度）の間隔が等比率でない。これを顕型尺度・元型尺度の問題と呼ぶ。

2）分布の非正規性・分散の非等質性に対処できなかった場合

説明省略。

処理の予定

順位尺度データの処理も基本的には、データ表示→データ分析、と進む。

なお、順位尺度データの処理は、単純な 2 条件比較が中心であり、複数要因の交差作用の分析は困難である。この点、一概に、問題のある間隔・比率尺度データを順位尺度に落とすのも考えものである。

7.1　データ表示

用いる統計量

順位尺度データの表示に用いる統計量は、データの個数、メディアン、レンジ、の３つである。
これらは間隔・比率尺度データのN、X、SDにそれぞれ相当する。データの手順は、以下の４つのステップを踏む。

①　各条件のデータの個数を数える。
②　各条件のメディアンを算出する。
③　各条件のレンジを算出する。
④　以上の統計量を作表によって掲載する。

最後の④の統計量の掲載は、作表のみであり、作図はない。なぜなら、作図に用いる棒（バー）や線（ライン）は、順位尺度の数値間を埋めて、連続性を表現してしまうからである。一般に、作図による統計量の掲載は、順位尺度や名義尺度など、中間値の欠落した「離散変量」には不向きである。

なお、間隔・比率尺度から順位尺度に落ちてきたデータの場合、データ表示は、もともとの「連続変量」としての統計量（X、SD）によって行われる例が多い。その場合、間隔・比率尺度データを順位尺度に変換してデータ分析に移るのであるが、変換後のデータ表示を省略する。すなわち、「間隔・比率尺度としてのデータ表示」→「順位尺度としてのデータ分析」、という手順になる。

a.　メディアンの算出

定義

メディアン（median）とは、データを値の小さい順に並べたとき、最小値・最大値から数えてちょうどまんなかにくる値のことである。訳語では「中央値」または「中位値」という。統計学上の分類では、間隔・比率尺度データの平均と同じく、代表値の一種である。メディアンの記号は、一般には“Me”を用いるが、あまり普及してない。

算出例

メディアンの算出は、データの個数（N）が奇数の場合と偶数の場合で、若干、方法が異なる。

1）Nが奇数の場合の例
　データ｛7,12,38,41,75｝→メディアンは「38」
2）Nが偶数の場合の例
　データ｛7,12,38,41,75,86｝→メディアンは38,41または（38＋41）/2＝39.5

b.　レンジの算出

定義

レンジ（range）とは、データを値の小さい順に並べたとき、特定の値から特定の値までの幅のことである。この幅を100%にとると、最小値から最大値までの「フル・レンジ」になる。また、80%や50%の部分的範囲にとると「パーセンタイル・レンジ」になる。前者は極端値の影響を受けやすい。統計学の分類では、間隔・比率尺度データの標準偏差と同様、散布度の一種である。レンジの記号は“R”であるが、これは重相関係数の記号と重複するため、ここでは「レンジ」と書く。

算出例

まず、レンジの範囲を決める。標準偏差に準ずるなら70%がよいが、少数データの場合は50%程度が適当である。

データ｛2,6,8,9,10,11,12,15,19｝（100%）
　　　｛・6,8,9,10,11,12,15・｝（78%）
　　　｛・・8,9,10,11,12・・｝（56%）←ストップ

以上のように50%にいちばん近いところで止める。その結果、このデータのレンジ（56%）は
「8－12」である。なお、メディアンは「10」であるので、このデータは、10を中心に全体の56%の個数が8－12の範囲におさまっているとわかる。

データ表示の事例

事例

児童の創造力を発揮させるためには、協同場面と競争場面のどちらが効果的であるかを比較した。2人一組として各場面に12組ずつ割り当てて、5種類の廃物（空き缶や割り箸など）の再利用法を考えさせた。その結果、出されたアイディア数によって各組を1位～21位に順位づけした（協同場面の演出に失敗した3組は除外した）。

表126は、協同場面・競争場面ごとに順位のメディアンとレンジを求めたものである。順位の小さい方が創造力に優れている。

表126　共同・競争場面別のアイディア数による順位

	協同場面	競争場面
N	9	12
メディアン	10	13,14
レンジ	8－12（56%）	5－17（50%）

解説

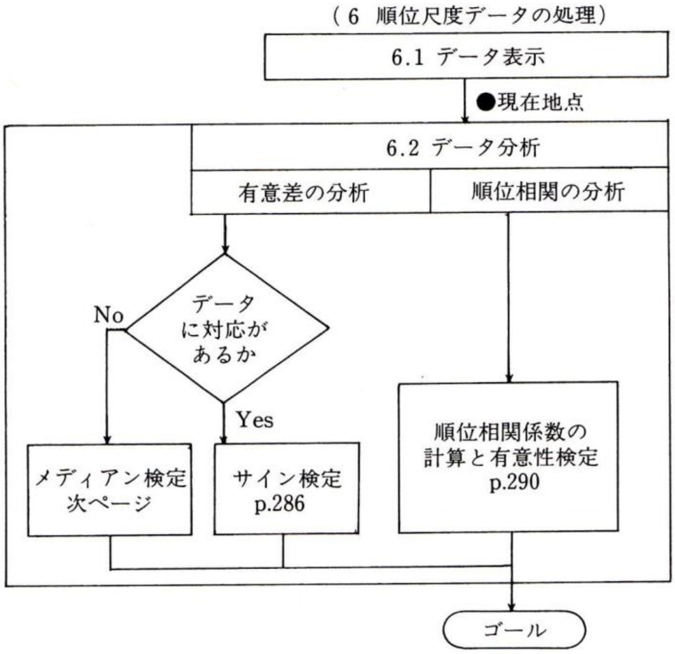
出されたアイディア数は創造的思考の量と直線的に関係しないが、優劣の指標にはなるであろう。そこで、ここでは順位づけ（ランキング）をおこなっている。上表からわかるように、Nが偶数の場合、メディアンは2つある。レンジの幅はふぞろいであるが、50パーセンタイル・レンジを基準にした結果である。メディアンの値によれば、順位として共同場面のほうが優位である。しかし、競争場面のレンジが大きく、競争が創造力の発揮にうまく作用する場合とそうでない場合の格差が大きいと言える。

7.2　データ分析

コースの選択

順位尺度のデータの分析には、①有意差の分析と、②順位相関の分析がある。有意差の分析は、2条件の差異をみる場合である。一方、順位相関の分析は、2変数の関係を見る場合である。

なお、以下に全体の予定をフローチャートとして示す。



7．2．1　有意差の分析

a.　メディアン検定

用途

メディアン検定は、条件間にデータの対応がない場合の有意差検定である。すなわち、各条件のデータは、すべて異なる被験者または異なる対象者から入手されていなければならない。なお、対応がある場合は、サイン検定を用いる。

考え方

順位尺度データの分析であるが、最終的に、名義尺度データの分析方法を適用することが出来るような形に持ち込む。

まず、各条件のデータをコミにして、共通のメディアンを求める（総メディアンと呼ぶ）。次に、各条件ごとに、この総メディアン以上と以下のデータを数え上げる。すると、「条件×2値（総メディアン以上・以下）」の度数の集計表をつくることができる。そこで、この集計表に対して、直接確率計算またはχ^2検定を適用する。このように、メディアン検定において利用するのは、各条件のデータ表示用のメディアンでなく、それらをコミにした「総メディアン」である。ここで、もし各条件のメディアンに差がなければ、どのメディアンも、この総メディアンに一致する（帰無仮説）。その場合、各条件とも、総メディアン以上・以下には同数ずつのデータが存在するはずである。この検定上の仮説は、各セルへの度数の偶然配分を意味しているので、名義尺度データの分析と同一主旨になるのである。

メディアン検定の事例

指導法Ａと指導法Ｂの効果を比較するため、被験者に、各指導法による処遇をあたえたあとで、満点 100（20 問×5 点）のテストを課した。その後、テスト 20 問の困難度にかなりバラつきがあることがわかった。また、軽微な減点も多用した。したがって、同じ 5 点であっても等価にならないおそれが出てきた。そこで、順位尺度データとして分析することにした。

表 127 は、データ・リストであり、表 128 は、そのメディアンと 50%パーセンタイル・レンジを示したものである。

表 127　データ・リスト

Sub.	指導法Ａ	Sub.	指導法Ｂ
イ	38	ヌ	32
ロ	48	ル	38
ハ	50	ヲ	45
ニ	55	ワ	46
ホ	65	カ	46
ヘ	72	ヨ	51
ト	74	タ	52
チ	80	レ	60
リ	90		
Sub.は、Subjects（被験者）の略			

表 128　　データ表示

	指導法Ａ	指導法Ｂ
N	9	8
メディアン	65	46
レンジ	50－74（56%）	45－51（50%）
指導法Ｂのメディアンは 2 つ、同値なので一方のみ掲載		

①　「総メディアン」を求める。
ここで、全体のデータ数（N_A+N_B）が偶数となると、総メディアンは 2 つになるが、そのときは両者の平均をとる（小数値になってもよい）。上例の場合は奇数であり、総メディアンは 51 である。
②　「条件×総メディアン以上・以下」の集計表をつくる。
総メディアンと同じ値のデータは「以下」のほうへ入れる。

表 129　メディアン検定のための集計表

	指導法Ａ	指導法Ｂ
＞51	6	2
≦51	3	6

③　直接確率計算または χ^2 検定をおこなう。
本例は 2 条件であり、2×2 表となる。この場合、直接確率計算を適用する。もし 3 条件の場合は、3 条件×2 値の表となるので、χ^2 検定を適用する。

上表の度数の組み合わせが偶然に出現する確率は、直接確率計算の結果、$p = .22$（両側検定）であった。これは、5%の有意水準をクリアすることができない。したがって、指導法Ａと指導法Ｂのメディアンの差は有意でなく、偶然に生じる程度でしかない。

以上、本例は、データの尺度の等間隔性に自信がもてないので、メディアン検定を用いた。その結果、指導法Ａ・Ｂの効果に大差はなかった。しかし、表 128 のレンジの幅を比べると、指導法Ａの恩恵を受けた被験者が一部存在する可能性がある。すなわち、指導法Ａの効果は普遍的でなく、選択的であることが示唆される。なお、本例はＸと SD を求め（$X_A = 63.6$、$SD_A = 16.0$: $X_B = 46.3$、$SD_B = 8.0$）、ｔ検定をおこなうと有意である。ただし、そ

のコースをとると、データの値の等間隔性を疑われながら、指導法Aの優秀性だけで終わる。本例の結果と考察は、もっと微妙な事実を示しており、これは処理の方針が正しいので疑いえない。

b.　サイン検定

用途

サイン検定（または符号検定）は、条件間にデータの対応がある場合の有意差検定である。順位尺度データの分析方法の中で、よく利用されるほうである。特に、間隔・比率尺度データにおいてNが少ない場合、対応するデータどうしの差が正規分布するかどうかを確認できないので、t検定にゆかず、サイン検定へ分岐してくることが多い。

サイン検定の基本的な考え方は、メディアン検定と同様である。すなわち、最終的に、データを度数の集計に置き換え、名義尺度データの分析方法を適用することができる形にする。なお、サイン検定は、2条件の場合にしか適用できない。3条件以上の場合は、2条件ずつ検定して、結果を総合的に考察すればよい。ちなみに、3条件以上の一括検定の方法としては「フリードマンの検定」がある。

事例

他者の存在が課題遂行に影響を及ぼすかどうかを調べるため、傍観者を立てた条件と立てない条件を設定し、被験者に3分間の掛け算または割り算をさせた。被験者は両方の条件を試行する。なお、傍観者のあり・なしと掛け算・割り算を試行する順序は、被験者ごとの乱数によって決定した。表130は、時間内に解答した問題の個数を示したものである（Nが小さいのでロー・データすなわち生データのまま掲載した)。

表130　条件別の解答個数

Sub.	傍観者あり	傍観者なし	ありーなし
イ	22	18	＋4
ロ	21	39	−18
ハ	18	25	−7
ニ	17	24	−7
ホ	24	29	−5
Sub.は、Subjects（被験者）の略			

①　サインを求める。

上表のように、対応するデータどうし（すなわち同一被験者のデータどうし）を引き算して、そのサインだけを求める（＋が1個、−が4個：順位尺度は間隔の概念がないので値の大きさは意味がない)。この作業は、全データを「通し順位」に変換してからおこなってもよいが、解答個数そのままでもかまわない。

②　サインの同数を1×2表に集計する。

表131　サイン検定のための集計表

＋	−	合計
1	4	5
±0は、除外し、集計しない。		

この表に対して、直接確率計算または二項検定を適用する（N＞30なら1×2のχ^2検定を代用してもよい)。検定の結果、p＞.10であり、有意でない（両側検定)。したがって、−のサインが多いので、傍観者の存在が課題遂行を慎重にさせ解答個数を減らしたようにみえるが、偶然のユレの範囲内である。

サイン検定の事例；3条件の場合の適用例

事例

論文名「自己開示に対する態度の変化」

目的

日本人は、対人関係における自己開示（本来の自分を表現すること）に消極的であると言われる。その心理的抵抗の特性をあきらかにするため、成人の被験者を対象に、自己開示する人物のビデオを見せ、この人物への好意度を10段階評定させた。ビデオは20分であり、評定は開始後5分後、10分後、20分後の3回おこなうことにした。

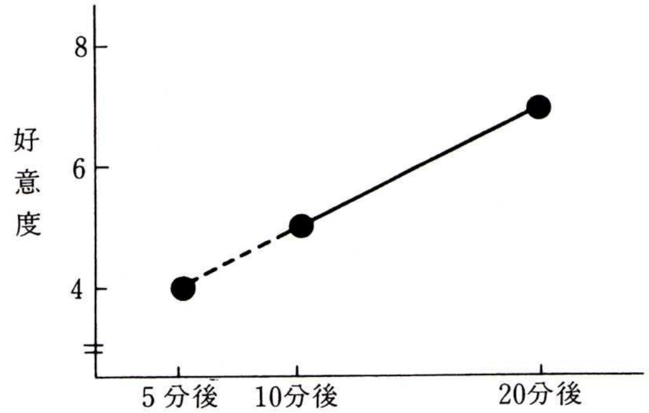
方法

被験者：大学生12人（男子6人、女子6人）

手続き：実験は集団実験である。

結果

図 12 はビデオの上映開始後の好意度評定の平均得点を示したものである（値が大きいほど好意度が高い）。



(注)破線はn. s., 実線は有意傾向を示す(サイン検定).

図12　自己開示の進行にともなう好意度の変化

サイン検定の結果、5 分後と 10 分後の評定点の差は有意でなかったが（両側検定：ｐ＞.10、N＝8）、10 分後と 20 分後の評定点の差は有意傾向にあった（両側検定：.05＜ｐ＜.10、N＝11）。したがって、自己開示への抵抗は最初は大きいが、固定的なものではなく、ビデオ視聴の進行につれて解消してゆくものであることが示唆された。

解説

データは、10 段階の評定点である。間隔尺度とみなしてもよさそうであるが、「間隔尺度としてのデータ表示」→「順位尺度としてのデータ分析」という折衷的処理を行っている。また、本例は 3 条件であり、サイン検定では一括分析できないため、時間軸に沿って隣り合う条件どうしを 2 つずつ検定している。これでも十分である。論文の「考察」に役立つ材料がそろったと言える。

なお、サイン検定は、上の例の場合、二項検定によっている。直接確率計算を用いた場合は、「両側検定：.05＜ｐ＜.10」でなく、例えば「両側検定：ｐ＝.07」というふうに計算結果の値を直接記載すればよい。また、χ^2検定を用いた場合は「両側検定」のコトバは不要であり（両側検定しかない）、「χ^2、df、ｐ」の三者を記載する。検定結果の記述において、カッコ内の「N=8」や「N＝11」は、対応するデータのうちサインを与えた組数を示している。例えば、5 分後と 10 分後の検定では「N=8」であったので、全 12 組のデータ中 4 組の引き算が±0 となり、サイン検定から除外されたということがわかる。

7．2．2　順位相関の分析

定義

順位相関（rank correlation）とは、順位尺度データにおける変数の関連度のことである。順位相関を表すための統計量としては、スピアマンの順位相関係数とケンドールの順位相関係数τ（タウ）が有名である。

性質

スピアマン（Spearman：記号はr_s）の順位相関係数は、2 変数の順位が完全に一致したときに 1 となり、完全に反対のときに、−1 となる。この点、間隔・比率尺度データの相関係数に準じた性質をもつように調整されている。

計算例

①　対応するデータどうしの差を求め、その二乗和を計算する。

被験者	変数A	変数B	差	差の二乗
ア	1	2	−1	1
イ	2	1	1	1
ウ	3	5	2	4
エ	4	4	0	0
オ	5	3	2	4
カ	6	6	0	0

総和10（←二乗和という）

(注)この場合,「通し順位」では計算できないので，各変数ごとに順位づけすること。

②　二乗和を下式に代入し、r_sを計算する（N はデータの組数）。

r_s＝1−（6×二乗和）/（N^3−N）＝1−（6×10）/（6^3−6）＝.71

「差の二乗和」は2変数の順位が完全一致で最小値0、完全に正反対で最大値（N^3−N）/3 になる。そこで、これを−1～+1 に置き換えるため、上の式のような調整を行っている。

③　r_sの有意性を検定する。

順位相関係数の有意性検定は、データの対応数が 20 を越える場合（N＞20）、近似的に F 分布を利用することができる。これは、ピアソンの相関係数（ｒ）と同様の検定になる。データの組数が 20 以下の場合（N≦20）、以下の表を用いる。

順位相関係数の制約
一つの条件の順位の中に同順位が含まれていると計算不可能になる（例：1,2.5,2.5,4,5.・・）。対処の方法はいくつかあるが、やや複雑。
順位相関の事例：曲線相関への応用
第４章で述べた間隔・比率尺度データの相関・予測の分析は、「直線相関」を仮定していた。このため、あらかじめ散布図によって相関の直線性をチェックしたり、もし曲線相関の場合には、何らかの対策を講じたりする必要があった。ところで、そのような場合に、順位相関を適用するのも効果的である。例えば、下表は、前に事例として取り上げた間隔尺度データのリストである。その際は、曲線相関であったため「スピーチ」のデータを対数変換することによって直線相関に整形した。

順位相関係数（r_s）の有意性検定の表（両側検定）

N	有意な r_s の値		N	有意な r_s の値	
5	.900		13	.478	.555
6	.771	.829	14	.459	.534
7	.679	.754	15	.443	.518
8	.619	.714	16	.427	.500
9	.583	.683	17	.409	.488
10	.552	.636	18	.399	.472
11	.527	.609	19	.388	.458
12	.500	.580	20	.379	.445
出現確率	.10	.05	出現確率	.10	.05
有意水準	有意傾向	5％	有意水準	有意傾向	5％

表 79　データ・リスト (音節数/10 秒)　　⇒　表 132　順位に変換した場合

被験者	スピーチ	和文
イ	25.5	30.8
ロ	46.3	39.1
ハ	26.3	35.4
ニ	31.6	36.9
ホ	18.5	19.5

被験者	スピーチ	和文
イ	4	4
ロ	1	1
ハ	3	3
ニ	2	2
ホ	5	5

しかし、ここで、順位相関を適用すれば、表 132 のように、完全に「スピーチ」と「和文」のデータ順位は一致する。すなわち、r_s＝1 となる。