第１部
「論文を読むために必要な統計知識」

ここではまず，論文を読むために必要な統計知識を身に付けましょう。といっても，各種統計全般に渡ってということは難しいですから（私も知らない統計手法がゴロゴロしています），この講義の目的にそった，尺度を使った質問紙調査で主に使われるものを中心に解説します。

ちなみに，この授業では統計知識がゼロの人を対象としてはいません。でもたくさんは要求していません。「平均値」「中央値」「標準偏差」「分散」くらいは知っていて欲しいと思います（どうやって計算するのかなど）。図書館で１日勉強すれば大丈夫です。特に，数字に弱いという人は，ウォーミングアップのつもりで復習しておいてください。また本文中では，ほとんど数式を使っていません。この点については是非の両面がありますが，できるだけとっつきやすいものということで除いておきました。当然，知らなくてもいいというものではありませんので念のため。

なおここでの紹介は，論文を読んで，それを理解するために必要な知識を紹介するものであって，これを理解したからといって統計手法が使えるかという問題とは別です。使うときには，論文を読むときよりもさらに詳しい統計知識（スキルといってもいいかもしれない）が必要になります。これについては，第３部で話をします。なお，特に大学院進学を考えている人は，各自さらに勉強しなければなりません*[1]。院試も厳しいからね…

次に，紹介する順番について記しておきましょう。ここでは，質問紙法を使った研究の流れにそって紹介するつもりです。そのため，統計についての各種書籍の記述順とはかなり違ったものになっていますので注意してください。

*[1]　時に臨床を志望する人には，統計はわからなくても…，と考えたり，人間の心理は質問紙などではかることはできないと考えている人がいるようですが，大学院に行きたいのなら最低限これくらいはきちんと理解しておいてほしいです。まあ，これくらいの知識しかないようでは，大学院進学自体が難しいですが…。

０．尺度についての基礎知識

尺度を使った研究を理解するためには，どうしても尺度について少し知っておく必要があります。まずこれから始めましょう。

尺度（ひらたくいえば「ものさし」です）には，４種類あります。これは水準という観点から整理したものです。それらは，**名義尺度，順序尺度，間隔尺度，比率尺度**です。

(１)名義尺度

まず，名義尺度というのは，男女とか，所属学部とかを数値化したものです。男女の場合，男性を１，女性を２と数字をわりあてても構いませんし，男性を10，女性を15でも構わない，男性をM，女性をFという記号だっていいのです。数字が使われた場合でも，それは分類のために任意にわりふられた数字であるので，和を計算したり平均値を出したりしても，出てきた数字は意味を持ちません。

(２)順序尺度

次に順序尺度ですが，典型的なものとして，ものの大小の順を表したり，主要５教科の好きな順番などを表す場合の数字があります。この場合，数字はその順序を示すことになります。心理学の測定では，「いつもする」「ときどきする」「たまにする」「まったくしない」などが使われることがあります。“えっ”と思うかもしれませんが，これも順序尺度です（程度の順番ですね）。厳密に言えば，順序尺度では平均値は意味を持ちませんから，中央値などを使いますし，関連をあらわす指標として順位相関係数などを使うべきです*[2]。しかし，これらの選択肢の間隔は等しいと“見なし”て，次の間隔尺度の水準として用いている場合が多いです。

*[2]　順位相関係数については，相関係数の記述を参照してください。

(３)間隔尺度

間隔尺度は，各数字の間の間隔が等しいことが特徴です。温度を示す摂氏などがこれにあたります。つまりこの尺度で測定された場合，３と２の間の１という間隔は，５と４の間の１に等しいということが成り立つということです。こうなれば，平均値などが意味を持ってきます。ただし原点は不定であり，任意に設定されるものです。

心理学の尺度でよく使われる，「そう思わない」「どちらかといえば，そう思わない」「どちらともいえない」「どちらかといえば，そう思う」「そう思う」などの選択肢を用意しておいて回答を求めるものも，この尺度水準を狙ったものです。こういうと，論文を読むときの注意点がわかるでしょう。つまり，「そう思わない」から「そう思う」までの選択肢に，できるだけ等間隔になるような表現が使われていなければならないという点です。

なお，論文を読んでいると，「『○○』から『●●』の５件法を用いた」などと記載されていることがあります。ここでいう５件法というのは，選択肢を５つ準備したということです。

(４)比率尺度

最後に比率尺度ですが，長さの測定などがこれにあたります。原点が存在し，数値の間隔は等しいことが特徴です。またその名前が示すように，この数値は比を計算することができます。

先程，間隔尺度の代表として摂氏をとりあげましたが，摂氏100度は摂氏50度の２倍と言っていいでしょうか？　これはまずいでしょう。摂氏は水の氷点と沸点を便宜的に０と100としてありますから，その０は任意に設定されたものです。熱量には絶対温度という基準があり，摂氏-273度を0とするものですが，この０は任意ではないため比率尺度となります。つまり摂氏50度は323度となり，その２倍は絶対温度646度，つまり摂氏373度が摂氏50度の２倍の熱量ということになります。このように，比率尺度による数値は比を使った計算ができ

ます。

以上のように，尺度で測定された数値には，もとの尺度がどのようなものであるかによって特徴が出てきます。つまり平均値を出しても意味がなかったり，Aの得点はBの得点の２倍であるなどということができるかどうかは，もとの尺度によって決まるのです。審査されている論文（学会の出している雑誌などに掲載されているもの）では，このあたりを間違っている記述はないと思いますが，大学紀要とかでは，ときどきミスがあったりしますので注意しておいてください。

１．クロス集計

質問紙調査を行った場合，まず最初に被験者の属性について記述する場合が多いです。なぜならこれらの情報から，その結果がいかなる集団から得られたものなのか，言い換えればその結果をいかなる集団に適応することができるのかを考えることができるからです。例えば，大学生を対象にした研究でも，それが文科系学部を中心としたものなのか，理科系を中心に得られたものなのか，もしくは両方からバランスよくデータを得ているのかで，考察や一般化の度合いが変わってくるでしょう。論文の読み手としては，この点についてきちんと押さえておかなければならないでしょう。

記述されている内容としては，被験者の男女別人数や，年齢，所属などについてです。論文中の文章としては，「大学生250名（男性134名，女性116名）を対象に調査を行った。平均年齢は20.6歳（年齢最小値19，年齢最大値23，標準偏差.66）であった」などとなっているでしょう。同様に，被験者の背景となる変数（例えば，学部・学科，居住形態など）についても，できるだけ多くの情報を記載されています（？）。

このときによく出てくる言葉が，**クロス集計**とか**クロス表**です。クロス集計を表にしたものがクロス表です。クロス表は，**クロス集計表**とか**関連表**とかとも呼ばれます。２つ以上の異なる項目に対する回答を，関連づけて示したものです。２次元の紙の上に表現する以上，３つは何とかなりますが，４以上のクロス表はかなり難しい（読みにくい）ので注意してください。

性別と所属学校別を例にとってみましょう。以下に４つの表を作ってみました。Table 1は，素データです。人数がそのまま入力してあります。Table 2は，所属学校別に性別の割合をカッコの中に%表示してあります。Table 3は，性別に所属学校の割合をカッコの中に%表示してあります。そして，Table 4は，各セル[*3]の全体人数に占める割合をカッコの中に%表示してあります。いずれの表の形態をとらなければならないというものではありません。各セルにばらつきが少ないことを言いたい場合はTable 4などを使えばいいですし，男女比がほぼ一定していることを示したければTable 2を使うなどと，利用の方針にしたがって使えばいいのです。なお，カッコの中の数字の意味については，ちゃんと記しておかなければなりません（ここでは，ずぼらをして書いていません）。

また，このような分布に対しては，偏りがあるかどうかを検定する方法もあります。これについては，後のχ^2検定のところで少し解説します。

*3　セルとは要因の組み合わせによってできる最小単位のことです。つまり表の１マスのことです。

Table 1

	中学生	高校生	大学生	合計
男子	52	75	60	187
女子	49	78	70	197
合計	101	153	130	384

Table 2

	中学生	高校生	大学生	合計
男子	52	75	60	187
	(51.5)	(49.0)	(46.2)	(48.7)
女子	49	78	70	197
	(48.5)	(51.0)	(53.8)	(51.3)
合計	101	153	130	384

Table 3

	中学生	高校生	大学生	合計
男子	52	75	60	187
	(27.8)	(40.1)	(32.1)	
女子	49	78	70	197
	(24.9)	(39.6)	(35.5)	
合計	101	153	130	384
	(26.3)	(39.8)	(33.9)	

Table 4

	中学生	高校生	大学生	合計
男子	52	75	60	187
	(13.5)	(19.5)	(15.6)	
女子	49	78	70	197
	(12.8)	(20.3)	(18.2)	
合計	101	153	130	384

２．因子分析

既成の尺度を使った場合でも，新たに尺度を作成した場合でも，結果の最初の方で出てくる場合が多いと思います。後で例として利用する堀野・市川(1997)の研究を参考にしながら，尺度作成における因子分析の利用を考えてみましょう。この測定目的は英語の学習方略を測定することです。どのような方略をどの程度用いているのかを把握したいのです。単純に考えれば，周りの人から，また自分自身の経験から英語の学習方略を聞き集める。そしてそれをリストにして，調査対象にそれぞれをどの程度使っているかを聞けばよいでしょう。しかし，もしリストに100もあがってくると，なんとも使い勝手が悪くなります。項目が多すぎるのです。また100も出てくると，何となく似たようなものも含まれるかもしれません。しかし，似ているのか似ていないのかの判断は，これがまた難しいことになります。そこで考えれば，たくさんあがってきたものを，統計的（主観を極力省いて）に要約すればよいというアイデアが出てくるでしょう。そこで登場するのが因子分析です。また**因子**という概念を導入したほうが，理論構成としてよりすっきりしたりすることもあります。因子分析は，同じような反応を受ける項目をグループ化しようとする統計手法です。まずは，よく出てくる用語について把握しておきましょう。

(１)回答の偏りについての検討

因子分析を行う前に，各項目への回答に過度な偏りが無いかどうかを検討している場合があります。分散を用いた分析はいずれもそうなのですが，各項目への回答に適当な散らばりがなければ結果が歪んでしまうことがあります。全く分散がなければ（例えば，100人が100人すべて同じ回答をする場合など），計算ができません。そこで，各項目について適当な散らばりがあるかどうかを，分析者が基準を定めて検討している場合があるのです。過度な偏

りと評価する絶対的な基準はないようですが，例えば，どちらかの極に全体の80%以上が集中するような項目は除外する，などの対応がとられているケースもあります。

(２)因子

直接測定することはできないが，各項目に影響を与えていると仮定される要素のことです。つまり因子分析は，各項目に対する反応から，その背後にあるだろう因子を推定しようとするものです。例えば，数学と理科の点数はいいけれども，国語や社会は苦手という人がいると，われわれは理系の才能はあるが文系の才能は？だと考えることがあります。この考え方そのものが因子分析の考え方にあてはまるのです。なぜなら，直接測定することのできない理系の才能というものが，理科や数学の点数にあらわれていると考えているからです。つまりここでいう，理系の才能＝因子，理科や数学の点数＝項目，と考えればわかりやすいのではないでしょうか。

(３)因子抽出方法と回転

因子分析の記述には，必ず**因子抽出方法**が記述されています（たぶん）。また抽出された結果に回転を施したものもあります。**主成分分析**とか**主因子法**というのが因子抽出方法であり，**バリマックス回転**などが回転方法です。回転[*4]は，抽出された因子を解釈しやすくするために行う手続きであり，大きく分けると**直交回転**と**斜交回転**があります。直交回転は，因子間は無相関であることを仮定しており，バリマックス回転が代表的なものです。斜交回転（プロマックス回転とかプロクラステス回転など）は，因子間に相関を仮定したものです。これらの手法には様々なものがあるので，見知らぬものに出会ったときは調べておくこと。

[*4] 回転については，各自で必ず統計の書籍を確認しておいてください。図でもって解説してあるものが多いと思います。計算式がわからなくても，ビジュアル的な理解だけはしておいてください。

(４)固有値

いくつの因子を抽出するかを決定する，１つの基準を提示してくれるものです。0より大きな数字が，因子分析に使った項目の数だけ算出されます。「固有値１を基準として○個の因子を抽出した」などと記述されることがあるように，１を基準とし，１以上の固有値がいくつあるかで因子数を決定することがあります（**ガットマン・カイザーの基準**）。「何で１なの？」と思う人は，自分で少し勉強してみてください。他には，固有値に急激な落ち込みが見られるところを基準（**スクリー基準**）として因子数を決定することもあります。ただし，どちらも決定的な基準ではありません。

(５)因子負荷量

因子負荷量は，それぞれの項目が，ある因子を反映している程度（絶対値）と向き（正負）を示す数値です。１から-１までの値をとります（ただし回転方法によっては，これ以外の場合もありますが）。因子分析を行った場合にはほぼ間違いなく，負荷量行列が表として掲載してあります。それを見れば，どの項目がどの因子を，どの程度反映しているのかがわかります。例えば，項目１は第１因子に .756[*5]の負荷量を示し，項目２は.212の負荷量を示し

たとしましょう。この負荷量の絶対値の大きさから，項目１の方が第１因子を反映しているといえます。さらに，項目３が第１因子に -.722の負荷量を示したとしましょう。項目が因子を反映している程度は絶対値で見るので，この項目も第１因子を強く反映しているとみなすことができます。ただし，符号がマイナスなのです。これは，意味的に項目１とは逆方向と考えられます。項目１と３は，例えば「長い」－「短い」のような意味内容的に逆の関係にあると考えられるのです。

*[5]　この因子負荷量に限らず，心理学の論文では0.123のように一の位が０の小数の場合，この０を省略することが多いです。

(６)因子の解釈・命名

抽出された因子に対して，それに高く負荷する項目内容を参考にしながら，その因子の内容を解釈し，妥当な名前をつけていく作業です。研究者のセンスが問われる作業でもあります。つまり，違う研究者が同じ結果を得た場合，もしかしたら違った名前になるかもしれないというものです。

なお先に回転のところで，回転は解釈を行いやすくするために実施すると記しておきました。それぞれの項目が，ある１つの因子には高い負荷を示し，それ以外の因子には低い負荷を示す場合には，因子の特徴がつかみやすく解釈を行いやすい。逆に，いずれの項目も，それぞれの因子にある程度の負荷量を示すようでは，因子の特徴を見いだしにくい。因子を回転すると，この負荷量のメリハリがはっきりすると考えておいてください（**単純構造**と言ったりします）。

(７)探索的因子分析と確認的因子分析

因子分析には，利用の方向としてこの２つがあります。論文中に，はっきり明示されている場合はまれでしょう。ある概念に従い，それを測定しうるだろうと考えられる項目を集め，データを収集し，そこにどのような因子が見られるかを探索的に検討する場合に用いられるのが**探索的因子分析**です。新たに尺度を作る場合などはこれに該当します。また既存の尺度を用いてデータを収集し，そのデータが先行研究での因子構造と合致するか否かを検討するような場合，また確固たる因子の仮定があって，それにしたがって分析を行うようなときに用いるのが**確認的因子分析**です。この資料の最後の方に書いている，共分散構造分析を使った因子分析が代表的ですが，回転にプロクラステス法が使われていたら，これは仮説的回転方法なので，確認的因子分析をやっていると判断してもよいでしょう

－因子分析を使った論文記述例－

高校生の英語学習における学習動機と学習方略

（堀野　緑・市川伸一　1997　教育心理学研究　45，140-147．より抜粋・一部加筆）

尺度項目の作成

上記とは異なる神奈川県内の一高校に在籍する高校３年生250名（女子132名，男子128名）に，英語の授業中，英語担当教員の立ち会いの下，英語学習方略リストにあげられた（23の）方略をどのくらい使用しているかを「非常によく使用している」から「全然使用していない」までの７件法で回答

してもらった。主成分解を初期解としバリマックス回転する因子分析を行った。固有値の変化は，第１固有値と第２固有値，第３固有値と第４固有値の間で大きかったが，因子解釈可能性から３因子解を選択した。

どの因子にも負荷量が.40に満たない項目を省いて，17項目により再度，主成分解・バリマックス回転により因子分析した結果をTable 1に示す。第１因子は，「１つの単語のいろいろな形を関連させて覚える」，「同意語，類義語，反意語をピックアップしてまとめて覚える」などに負荷が高く，英単語を体制化して記憶しようという傾向であることから，「体制化方略」と命名した。第２因子は，「単語のスペルを頭の中に印刷の文字印刷の文字ごと浮かぶようにイメージする」，「単語をながめながらアルファベットの配列の雰囲気をつかむ」などに負荷量が高く，語のイメージやニュアンスをつかもうとする傾向であることから，「イメージ化方略」と命名した。第３因子は，「手と頭が完璧に覚えるまで何度も書く」，「英語から日本語，日本語から英語へと何度も書き換える」など，くり返しを重視する方略に負荷量が高いことから，「反復方略」と命名した。これらの因子に負荷の高い項目を用いて，項目平均点からなる尺度得点を算出した場合，α係数は，「体制化方略」が.83，「イメージ化方略」が.72，「反復方略」が.77であった。このように，因子構造の明確さと信頼性の高さは十分に確認されたといえよう。

Table 1　学習方略の因子分析（主成分解・バリマックス回転後）の負荷量

No.	項目内容	F1	F2	F3	h^2
3	１つの単語のいろいろな形（名詞形・動詞形〕を関連させて覚える	.86	.02	.11	.75
9	同意語，類義語，反意語をピックアップしてまとめて覚える	.79	-.08	.17	.67
7	同一場面で使える関連性のある単語をまとめて覚える	.76	.06	.09	.59
21	動詞の変化をまとめる	.72	.03	.13	.53
5	スペルが似ている単語，意味が似ている単語はまとめて一緒に覚える	.71	-.03	-.12	.52
18	動詞の分類化（自動詞，他動詞）をする	.66	.06	.10	.45
15	その単語を使っている熟語を覚える	.57	.22	.29	.45
1	単語のスペルを頭の中に印刷の文字ごと浮かぶようにイメージする	.06	.87	-.14	.77
8	単語をながめながらアルファベットの配列の雰囲気をつかむ	-.08	.81	-.08	.67
10	頭の中に単語がイメージできるように何度も見る	.13	.79	-.07	.65
12	何か他の単語と関連させて連想できるようにして覚える	.28	.54	-.12	.39
16	発音が何か他の別の言葉（日本語）に似ていたら語呂合わせをする	-.10	.51	.19	.30
2	手と頭が完璧に覚えるまで何度も書く	.00	-.18	.80	.67
17	英語から日本語，日本語から英語へと何度も書き換える	.23	.05	.71	.56
23	新しいわからない単語にラインをひいておく	.02	.13	.64	.42
20	発音しながら単語を書く	.12	-.25	.62	.46
13	わからない単語をチェックペンとシートを使って意味と単語をくりかえし覚える	.13	-.03	.56	.34
	二乗和	3.93	2.77	2.49	9.19
	寄与率	23.11	16.29	14.65	54.06

F1:体制化方略　F2:イメージ化方略　F3:反復方略

…ちょっと解説…

用語的には，先に記述したものでなんとかなるでしょう。補足としては，**主成分解**というのが出てきますが，これは主成分分析で得られた解のことですので，抽出方法としては主成分分析が使われています。主成分解の直後に，**初期解**というものが出てきますが，これは回転前の推定値のことです。

またTable 1中の用語についても説明しておきましょう。まず「h^2」は**共通性**とも呼ばれます。算出方法は，各項目の各因子への負荷量の自乗和*[6] を求めればOKです（つまり表の横の自乗和）。この計算式からわかるように，共通性は各項目の分散のうちの因子によって説明される部分の比率です。これが非常に小さい項目があるとすれば，その項目は今回の因子分析で得られたいずれの因子とも関連が浅いことを示します。また**二乗和**（自乗和）は，各因子ごとに負荷量の自乗和を求めたものです

（つまり表の縦の自乗和）。**寄与**は，二乗和（自乗和）を項目数で割ってたもので，**寄与率**はそれを%表示したものです。寄与率は，測定された変数全体の分散のうち，各因子が説明できる分散の大きさを示します。つまりこの研究では，３つの因子で全体の分散の半分以上（54.06%）を説明できるのです。因子分析を用いるメリットを要約機能と考えると，少ない因子数でできるだけ多くを説明する必要があります。これを判断するのに，寄与・寄与率が役に立ちます。なお，因子分析の結果を表示するときには，共通性と寄与率（寄与）は必ず書くようにしましょう。

もう一度本文に戻りますが，抽出する因子の数は，「固有値の変化は，第１固有値と第２固有値，第３固有値と第４固有値の間で大きかったが…」とあるように，スクリー基準を適用しようとする姿勢が見られますが，最終的には「因子解釈可能性から」決定しています。ではなぜ，「因子解釈可能性から」決定しているのでしょうか。通常因子分析を行う際には，抽出方法や抽出する因子数を変化させながら最も妥当な解釈が可能な結果を探索します。先に固有値のところで記した，固有値１やスクリー基準は，妥当な因子数を決定する際の目安なのです。ですから，最終的には因子解釈の妥当性から決定することが望ましいでしょう。

また，この研究では，まず23項目を用いて因子分析を行っています。そして「どの因子にも負荷量が.40に満たない項目を省いて，17項目により再度，主成分解・バリマックス回転により因子分析」を行っています。「どの因子にも負荷量が.40に満たない項目」というのは，今回抽出されたいずれの因子もあまり反映していない項目と考えられるのです。このような項目はよく出てくるのですが，これらの項目への対処方法としては大きく２つあるように見受けられます。１つはこの研究のように，それを除いて再度因子分析を行うという対処です。もう１つは，全項目を用いた因子分析結果を記載しておき，いずれの因子にも一定以上の負荷量をもたないような項目は，表の下の方にまとめ，残余項目として記載しておくというものです。因子分析を用いた論文をいくつか読んでいくと，大体この２つの方法が使われていることがわかると思います。しかし，どちらの対処が良いかということに関しては一概に答えを出すことはできません。また，いずれかが間違っているというものでもないのです。因子分析をどのように利用しようとしているのか，という分析者の意図に従って記載されていると考えていいでしょう。

*6　自乗和と二乗和は，結局は同じことです。ここでは用語としては，自乗に統一しています。

３．因子分析結果を基に尺度を構成する

因子分析では項目を要約しました。その要約をするときの柱となるのが因子です。このそれぞれの因子を代表するような項目をいくつか選出すれば，それらの項目の合計得点は各因子を代表する値となるはずです。これが因子分析結果を基に尺度を作成する作業で，つくられた尺度は**下位尺度**と呼ばれます。先の堀野・市川(1997)論文では，作ろうとした尺度は学習方略についての尺度で，その項目プールの中に３つの因子が見いだされています。つまり，学習方略についての尺度は，それぞれの因子に対応した３つの下位尺度から構成される尺度となったわけです。このような因子分析結果を基に尺度を構成するときにも，いくつかの統計用語や指標が使われます。ここではそれについて解説しましょう。

なお，因子分析を行っていないものでも，新たに尺度を構成する場合には以下の(２)，(３)のような手続きをふみます。

(１)各下位尺度を構成する項目のピックアップ

通常，各因子に高い負荷量を示す項目を選出して各下位尺度を構成する項目とします。こ

のときに，自分なりに抽出基準を設定して行うことになります（例えば，.400や.500（ただし，絶対値）を基準としてそれ以上の負荷量をその因子に示すもの）。また，各項目はすべての因子に対していくらかの負荷量を持ちますから，複数の因子に対してこれらの基準を満たすものがでてくることもあります（例えば，第１因子に.520，第２因子に.502というように）。このようなケースの場合では，さらに基準を加えることがあります。例えば，ある１つの因子に.500の負荷量を示し，その他の因子に対しては.400以下の負荷量を示すものを選出するとか，複数因子に.500以上を示す項目は採用しないなどといったものです。

(２)信頼性を検討する

テストの**信頼性**は，そのテスト結果の正確性を表す概念です。内容的には，同じ対象に同じ尺度を繰り返したときには同じ結果にならなければならない，ということと，尺度に含まれる項目は同じものを測定していなければならない（**内部一貫性**と言ったりもします），という意味合いが含まれます。これらを包括して信頼性と呼ぶことが多いようです。

この信頼性を示す指標にはいくつかあります。まず**再テスト法**といって，同じ尺度を同じ対象に間隔をとって２度実施し，その２回の関連性から信頼性を検討する方法です。また**折半法**と呼ばれる方法もあります。また信頼性係数を算出する手法としては，**スピアマン・ブラウンの公式**，**キューダー・リチャードソンの公式（KR-20，KR-21）**，**アルファ（α）係数**などがあります。このあたりは各自で調べておくこと。

最近はα係数を使ったものが多いので，これについて少し説明を加えると，数値的には値が１に近づくほど信頼性（内部一貫性）が高いことになります。いくつかの書籍では，α係数が.7とか.8を越えることが望ましいと記してあります。なおα係数には特性があり，係数を算出する尺度の項目数が多くなれば係数は高くなり，逆に項目数が少なくなれば係数が低くなる傾向があります。そのため，項目数が２とか３の場合には，α係数はあまり意味を持たなくなり，それに対処するために各項目間の相関係数を算出し，その値から信頼性を確認するような方法も使われています（ここらあたりのことを知ってか知らずか，α係数を強引に記述してあることが時にあるので注意しておいてください）。先に紹介した堀野・市川(1997)論文の最後のところを，もう一度確認しておいてください。

最後に，これも信頼性の分析に入れてもよいと思うのですが，**I-T相関（Item-Total Correlation）**と**G-P(Good-Poor)分析**について書いておきましょう。まずI-T相関を算出しているケースですが，これは各項目得点と，その項目を除いた他の項目の合計得点との相関係数[*7]を求めるものです。この相関係数が高いということは，その項目と，その項目を除いた他の項目群の方向性が一致していることになります。逆に相関係数が低いと，その項目と，その項目を除いた他の項目群の方向性の方向性が一致していないことになります。つまりI-T相関が低い項目を含めると，尺度の内部一貫性が乱されることになるのです。

次にG-P分析ですが，これは尺度の総得点で被験者を３群（上位群，中位群，下位群），もしくは２群（上位群，下位群）に分け，その群ごとに各項目の平均値を求め，それらの得点間の比較を行うものです。項目と総得点が適切に対応していれば，各群における各項目の平均得点は，上位群で高く，下位群で低いという結果になるはずです。これを確認するのがG-P分析です。

[*7] 後にある「相関係数」のところを，ちょっと見ておいてください。ちなみに相関係数の高低は，係数の絶対

値で見ていきます。＋が高くて，－が低いということではない点に注意。

(3)妥当性を確認する

はっきりいって，この尺度は妥当性がある！と明言しているような論文は少ないと思います。**妥当性**は，測りたい属性をちゃんと測っているかどうかということを示す概念で，信頼性とともに心理測定で非常に重要なものです。でも，この妥当性を確認することは困難でもあるのです。

妥当性の主な内容には，**内容的妥当性**，**構成概念妥当性**，**基準関連妥当性**があります。内容的妥当性は，測定したいと考えられている対象を正しく測定しているかどうかという点での妥当性です。例えば，小学校４年生で学習する算数課題についての学力を測定したいと考えたとしましょう。そうすると，小学校４年生で習う全ての領域をカバーするような問題が必要となります。４年生の２学期で習うものだけではダメということです。内容的妥当性とは，このような内容的偏りがないかどうかといった点からチェックします。

次に構成概念妥当性ですが，心理学で扱う概念の多く（例えば，内向性とか自己受容など）は，研究者自身が構成した概念です。このように，つくられた特性を尺度がどの程度測定できているかという観点から構成概念妥当性は判断されます。この判断には，理論的に仮定される関連が，データの上でも現れるかどうかということを中心に確認していきます。内向性の尺度を作ることを考えてみましょう。例えば，理論的に「内向性の高い人は自己主張性が低い」と考えられるのであれば，内向性尺度と合わせて自己主張性（すでに妥当性が確認されている尺度で）を測定しておき，それらの測定値の間に理論通りの関連が認められるかどうかという検討を行います。そこで，もし理論通りの関連性が認められたならば，新たに作成した内向性の尺度は構成概念妥当性をもつものと判断してよいことになります。

最後に基準関連妥当性です。その尺度が測定しようとしているものが何らかの基準となる得点を持つものであれば，その基準となる得点と尺度の得点の関連が強いと，その尺度の基準関連妥当性は高いということになります。

以上のように，作成した尺度の妥当性を検討するためには，かなりの手順を踏むことが必要となります。そのため，尺度作成のみを目的とした研究もかなりの数があります。いくつかの論文を読みすすめるとわかってくると思いますが，この妥当性を保証するために，構成概念をできるだけ明確化する努力をしていたり，その論文を書いている研究者だけでなく複数人で項目の内容的妥当性を検討したり，項目内容の偏りを少なくするために最初は大量の項目を用意したりして対応している研究もあります。

最後に，以上の説明で信頼性と妥当性はまったくの別概念であることが理解していただけたことと思います。つまり，統計的に高い信頼性が確認されたからといって，妥当性があるとはいえないのです。論文を批判的に読むときにこのような点に気をつけて読んでいると，尺度の改良すべき点などが見えてくるでしょう。

４．尺度の決定と基礎統計量

以上の手続きをへて，やっと研究に使える材料（値）がそろいました。以後は，この値を

使って，性差があるかとか，学年進行にしたがって値がどのように変動するかとか，Aという値とBという値はどのような関係にあるのかとか，具体的な分析に入っていきます。

論文によっては，ここで各尺度・各下位尺度の平均値と標準偏差などが一覧できるように記載されているものもあります。その論文のどこかでこれらの値がわかればいいので，一覧表がなくてもかまわないのですが，卒論のように紙幅がある場合であれば掲載しておくのがよいでしょう。

５．個々の研究目的に応じた分析手法

ここからは，個々の研究目的に応じた分析手法が用いられます。この分析は使わなくてはならない，といったものはありません。以下，主な分析手法を個々に簡単に解説していきます。解説するものは，以下の通りです。もし自分で論文を読んでいて，ここには掲載されていない分析に出会ったら…　当然，自分で調べてください。調べれば，大抵のことは…　？？

A.　t検定
B.　１要因分散分析
C.　２要因分散分析
D.　相関係数
E.　偏相関係数
F.　重回帰分析
G.　パス解析
H.　共分散構造分析
I.　その他（χ^2検定）

A．t検定

２群の平均値の差が有意なものかどうかについての検定です。例えば，男女差があるかどうかを検定するときなどに使います。３群以上の平均の差については使えません。でも，例えば１年生と２年生と３年生の平均について，１年と２年，２年と３年，１年と３年というようにt検定を繰り返せば，３群以上あってもどの群の平均値が高いかわかるじゃないか，というかしこい発想をする人がいるかもしれませんが，これはやってはいけませんので注意してください。

またt検定には，対応のある場合と対応の無い場合の２種類があります。何が対応があるとか無いとかいうのかというと，標本間のことです。例えば男女であったら，これらの２群は全くの別物ですので対応はない，ということになります。しかし同じテストをある授業の前と後に行ったときの得点を比較したい場合，同じ被験者ですから，これは対応のある場合となります。これらは計算式が違いますが，読み方としては同じです。

論文には，両群の平均値，標準偏差，人数，が記載され，得られた**t値，自由度（df），有意水準（危険率，p）**が示されていることが望ましいですが，全部記載されているものはめずらしいかもしれません。文章中であれば「～の間に有意な差が認められた(t=5.48, df=350, p<.01)」などと記載されていることが多いです。t値や自由度，危険率の計算は専門書に任せますが，考え方は「２群の平均値に差が無い」という**帰無仮説**が成り立つ確率を求め，それがめったに無いような確率の現象であれば「差が無いとはいえない」，つまり差があるといってよいだろうということになります。ではいったいどのくらいの確率で起こりうるものであれば，それをめったに無いといえるのでしょうか。この起こりうる確率が先の有意水準であり，通常では .05（5%）や .01（1%）を目安とし，5%より低い確率のようならば，「有意な差がある（認められる）」と記述します。時に.10（10%）を使うときもありますが，10%～5%の間である場合は「**有意傾向**にある差がある（認められる）」と記述します。

有意水準の表記の仕方には，そのままp=.033などと表記されているときをはじめ様々ありますが，メジャーなものの１つに**アスタリスク**（*）*余談を使ったものがあります。通常，アスタリスク１つ（*）の場合5%水準をクリアしていること，２つ（**）の場合は1%水準をクリアしていることを示します（* p<.05，** p<.01　などという注記がついています）。また10%水準の時に十字みたいな記号（ダッカー）（⁺）を，0.1%水準の時にアスタリスク３つ（***）を使っている場合もあります。なお，有意な差が認められない場合，ns (no significantの略)とかN.S.とかと表記していることがあります。

*余談　アスタリスクのことを，俗に「お星さま」と呼ぶことがあります。「お星さま信仰」などと揶揄されることもあります。今後もよく出てきますが，これがなかなかくせ者なのです。自分たちでSPSSなどを使って分析を始めると，「お星さま信仰」の意味がわかってくると思います。お楽しみに。

－t検定を使った論文記述例－

「親となる」ことによる人格発達：生涯発達的視点から親を研究する試み

（柏木惠子・若松素子　1994　発達心理学研究　5，72-83．より抜粋）

では，こうした「親となる」ことの変化・発達は，父親・母親においてどのようであるかを各次元項目の得点平均によってみてみよう（Table 3）。

父母いずれでも，視野の広がり，生き甲斐・存在感，自己抑制などを筆頭に，最高で3.12，最低でも2.20と，「親になる前に比べて親になった後の」変化を認めている方向にある。得点順位に関して父母間には多少の違いがあるが，「自己の強さ」や「視野の広がり」などの変化は他の３次元での変化に比べて小さく，「運命・信仰・伝統の受容」「生き甲斐・存在感」「自己抑制」などでの変化が父母に共通して大きい。

（略）

もう一つ注目されるのは，いずれの次元でも父親に比べて母親での変化が有意に大きいことである。これは，子育てを通して「変化した」と回答する者の率が父親よりも母親でかなり高く，またその変化を「自分の成長」とするのも母親がずっと多いという牧野らの結果とも一致する。この差をもたらす要因として，子育てへの関与の差が考えられ，そこでは母親の職業の有無と関係していると予想される。（以下略）

Table 3　親となることによる成長・発達
－次元得点平均（標準偏差）－

	父		母	P
第I因子　柔軟さ	2.40 (0.74)	＜	2.83 (0.61)	***
第II因子　自己抑制	2.57 (0.72)	＜	2.99 (0.62)	***
第III因子　運命・信仰・伝統の受容	2.71 (0.73)	＜	3.12 (0.54)	***
第IV因子　視野の広がり	2.21 (0.67)	＜	2.60 (0.63)	***
第V因子　生き甲斐・存在感	2.82 (0.57)	＜	2.95 (0.53)	**
第VI因子　自己の強さ	2.35 (0.69)	＜	2.52 (0.58)	***

注.　** P <.01　*** P <.001

…ちょっと解説…

まず，表の記載から説明しましょう。といっても，説明するほどでもないかもしれません。父・母別の平均値，標準偏差，危険率，そしてどちらが大きいのかを明示する不等号が記されています。残念ながら，t値と自由度は表にも文章の中にも記載されていません。平均値と標準偏差，そしてそれぞれの群の人数さえわかっていれば計算することはできますが，t値くらいは示しておいたほうが親切ではないかと思います。自由度については計算は簡単*8 ですので，まあいいかとも思いますが。

*8　自由度の計算は，各群の人数から１を減じたものを足した数となります。この研究では346対の夫婦が調査対象ですから，父も母も346名ということです。ですから(346-1)+(346-1)で，自由度は690となります。ちなみみ第I因子では，t値は8.33くらいの値となります。

B．１要因分散分析

t検定は，２群の平均値の差についての検定でした。そして，３群以上の場合は使えないと書いておきました。１要因分散分析は，この３群以上の平均値の差の検定に使う分析です。例えば，野球部とテニス部と弓道部員間の差とか，小学校１年生から６年生までの各学年ごとにおける，何らかの指標の差を検討するなどの時に使います。

この分析を行った結果は，まず分散分析の結果が有意かどうか，そして有意ならば多重比較を行って，どの群間に有意な差があるのかを示す，という流れになっていると思います。なぜこのような２段階の手順を踏むかというと，まず分散分析は「分析に用いられるすべての群の平均値は等しい」という帰無仮説を使い，これが成り立つ確率を求めるのです。ですから，この確率が一定の基準（５%や１%）を下回るときは，すべての群の平均値は等しいとは言えない，という結論を導くことになるのです。読んだらわかるように，「すべての群の平均値は等しいとは言えない」としか分析結果はいっていないのだから，どの群間に差があるのかはこの段階ではわかりません。ですから，**多重比較**と呼ばれる方法を使って，いずれの群間に差があるのかをさらに検討するわけです。

こう書いてくると，「言っていることが矛盾している」という鋭い指摘を受けるかもしれ

ません。t検定では，３群以上に対してそれを繰り返してはいけないといいました。でもこの多重比較はそれと同じことをやっているのではないか…。

ちょっと話はそれるのですが，ここで統計手法を利用した推論の誤りについて，ちょっと触れておきましょう。このような誤りには２種類があります。**第１種の誤り（タイプI エラー）**と呼ばれるものは，帰無仮説が本当は正しいのにそれを棄却してしまうことです。簡単に言うと，本当は差がないのに有意差を検出してしまう誤りのことです。もう１つ**第２種の誤り（タイプ II エラー）**と呼ばれるものは，帰無仮説は本当は間違っているのに，それを採択してしまうことです。つまり本当は差があるのに有意差を検出しないという誤りです。

話を多重比較にもどしましょう。t検定にしろ分散分析にしろ，この推論の誤りから逃れることはできません。有意かそうではないかを判断する分かれ目の５％基準を使って，本当は差のない３群の平均値を検定することを考えてみましょう。５％基準を用いるのですから，本当は差がないのに，差があると帰結してしまう（第１種の誤りを犯す）確率は５％ということになります。しかし，３群間の差を検討するのに，t検定を３回繰り返した場合ではどうでしょう。t検定１回について５％の誤る可能性があり，それを３回繰り返すので，結果３回のうち少なくとも１回は誤った結果を導く可能性は約14％（$1-0.95^3$）にもなってしまいます。ですから，まず一括して差の有無を検討し，もしあれば順次検討していくという分散分析＋多重比較という手続きを踏むわけです。また多重比較に用いられる方法には，t検定よりも第１種の誤りを犯す可能性が低い方法が用いられます。

では結果の読み取りについての話をしましょう。ほとんどの場合，**F値（F比）**，自由度（df），有意水準（危険率，p）が記載されていると思います。t検定結果との大きな違いは自由度が２つあるところでしょうか。pが５％水準や１％水準をクリアしていれば群間に有意な差があると判断できます。

次に多重比較に行きましょう。先にも述べたように，どの群間に差があるのかの検討です。多重比較の方法には**テューキー(Tukey)法**，**LSD法**，**ダンカン(Duncan)法**，**シェッフェ**(Scheffê)**法**，**ライアン(Ryan)法**など様々あります。

最後に，文中での結果記述の方法ですが，$F(1, 27)=5.64$，$p<.05$　といった表記がされていると思います。Fに続くカッコの中が自由度，＝の後がF値です。そしてそれが表れる確率（つまり有意水準）がその後に記されています。

－１要因分散分析を使った論文記述例－

父親になる意識の形成過程

（小野寺敦子・青木紀久代・小山真弓　1998　発達心理学研究　9，121-130．より抜粋）

（3）父親になる意識と親和性／自律性との関連：親和性に関する５項目と自律性に関する５項目の素点を合計し，親和性得点と自律性得点を算出した。両得点間の相関係数を求めたところ.281（$p<.001$）という有意な値を得た。これは，親和性得点が高くなるほど，自律性得点も筒くなる傾向が本研究の男性にはみられることを示している。しかし，男性の中には親和性得点は高いが自律性得点は低いもの，あるいはその逆の傾向を示す男性も存在するはずである。そこで，親和性得点と自律性得点の高低の組み合わせから４群を構成した。４群を構成するにあたっては，親和性得点と自律性得点が共に平均値よりも高い群を親和性High－自律性High群（48人，以下，親和性H－自律性H群と略す），親和性得点は平均値より高いが自律性得点は低い群を親和性High－自律性Low群（25人，親和性H－自律性L群）また，親和性得点は平均値よりも低いが自律性得点は平均値よりも高い群を親和性Low－自

律性High群（36人，親和性L－自律性H群），両得点がともに平均値よりも低い群を親和性Low－自律性Low群（58人，親和性L－自律性L群）とし，これら４群によって上記の６次元からなる父親になる意識がどのように異なっているのかを一元配置の分散分析と多重比較を行って検討した（Table 4）。

Grossmann et al. は，親和性と自律性その両方をバランスよく兼ね備えていることが親となった男性の健康な人格発達を促すとのべている。そこで，両群の組み合わせから親和性H－自律性H群の特徴をみると次のようなことが明らかになった。まず，この群の男性は「父親になる自信」得点と「父親になる実感・心の準備」得点が４群中最も高く，親和性L－自律性L群との間に有意な差が認められた。また「制約感」は４群中最も低くなっていた。このことから子どもを暖かく見守り，育てていこうとする親和性と社会の中で独立してやっていこうとする自律性の両側面が高い男性は，親になることに肯定的であり父性意識が妻の妊娠期にかなり確立されているといえよう。これに対し親和性L－自律性L群は，「制約感」得点が一番高く経済的，精神的に一家を支えることおよび家事の手伝いにかなりの負担意識をもっており，父性意識の確立もまだ十分でないと考えられる。「人間的成長・分身感」では親和性H－自律性L群の得点が一番高く，自律性ではなく親和性の高い男性が親になることによって自分は人間的に成長し変化するととらえ，また，子どもは自分の分身であると考える傾向がみられた。上記の結果から，男性が父性意識を確立していくには，単に親和性だけでなく社会人としての自律意識も高くもっていることが重要であることが明らかになった。

Table 4　親和性／自律性の高低と父親になる意識（各群の平均値とSDおよび多重比較の結果）

項　　目	1. 親和性 High 自律性 High	2. 親和性 High 自律性 Low	3. 親和性 Low 自律性 High	4. 親和性 Low 自律性 Low	*F* 値	
制約感	11.25	11.68	11.72	13.05	2.91	*
4＞1	(3.30)	(3.07)	(3.21)	(3.40)		
人間的成長・分身感	13.48	15.40	13.33	13.01	3.39	*
2＞4	(3.46)	(2.62)	(3.61)	(2.77)		
心配・不安	8.46	8.60	8.69	9.09	0.85	
	(1.97)	(1.60)	(2.48)	(2.10)		
時間・心の準備	6.56	6.28	6.03	5.64	4.00	**
1＞4	(1.24)	(1.40)	(1.38)	(1.49)		
父親になる喜び	7.31	7.32	7.00	6.78	2.89	*
1＞4	(0.92)	(0.88)	(1.16)	(1.11)		
父親になる自信	6.81	6.52	6.00	5.48	9.13	***
1＞3　1＞4　2＞4	(1.44)	(1.10)	(1.53)	(1.25)		

* p<.05　** p<.01　*** p<.001

…ちょっと解説…

論文中に，**一元配置の分散分析**とありますが，これが１要因分散分析のことです。要因が１つで一次元上に配置されているので，このように呼ばれることもあります。見方については簡単でしょう。

また表中には各群の平均値，標準偏差(SD)，F値，有意水準，多重比較の結果が示されています。これで十分だと思うのですが，先の記述のところで触れた，自由度についての情報は欠落しています。それほど多量な情報ではないので，やはりこれくらいの情報は入れるべきだと思います。

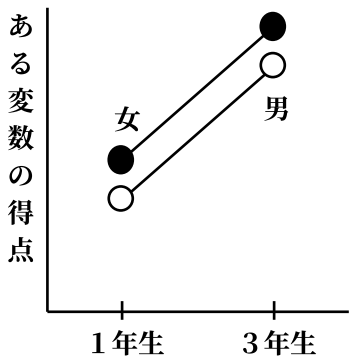
C．２要因分散分析

２要因分散分析は，その名の通り２つの要因によって被験者を分類し，各群間の平均値の

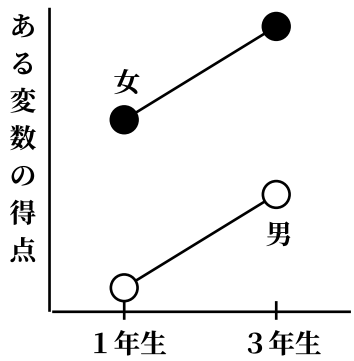
差について検討するものです。例えば，性と学年によって被験者を分類し，ある指標の平均値を検討するなどです。なお，３要因以上の分散分析を行うことも可能ですし，それを使った研究もあります。こうなってくると解釈も難しくなってきます。しかし，その多くは実験法を使った研究であり，質問紙調査をもとにした研究では，あまりお目にかかりません。せいぜい２要因まででしょう。特に実験法を使った研究を参照したり，自分でもやってみようと思う人は，分析の中心的手法になりますので必ず自分で勉強しておいてください。

この分析を行ったときには，**主効果**，**交互作用**についての記述があるはずです。まず，この意味から記しておきましょう。性と学年を要因とする分散分析で説明すると，この場合主効果は，性の主効果と学年の主効果の２つとなります。性の主効果は，学年の差を込みにした場合に性差があるかどうかについての検討というかたちで行われます。つまり性の主効果が見られたということは，性差があるということになります。次に交互作用ですが，言葉で説明すると難しいのですが（やはり図示がわかりやすいでしょう），一方の要因の違いによって，もう一方の要因の大きさが異なることを言います。例えば，女性の場合１年生の得点が３年生よりも高く，男性の場合は３年生の得点が１年生の得点よりも高いといった時に交互作用が検出されたりします。主効果や交互作用が見られる場合（２×２の分散分析の場合*9）の典型的なかたちを以下に図示しておくので，感覚的にでもいいですから主効果，交互作用の考え方をつかんでおいてください。

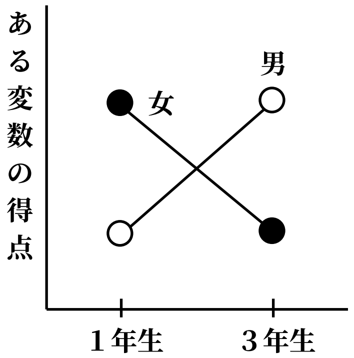
*9　分散分析の場合よく「２×２の分散分析」とか「２×４の分散分析」とかといった表記がされます。また「２(男，女)×４(中１，中３，高３，大３)の分散分析」といった表記もされます。○×○という○が要因を示し，そこに記入される数字が水準数を示します。２要因なら○×○，３要因なら○×○×○という感じです。なお水準は要因内の相違の数であり，性の場合は２（男女だから当然ですね），１・２・３年生の比較なら３ということになります。



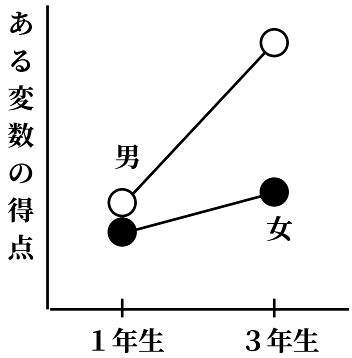
学年の主効果があるような場合



学年と性の両方の主効果があるような場合



交互作用があるような
場合１



交互作用があるような
場合２

後は，検定結果の記載の仕方ですか。主効果，交互作用のそれぞれについて有意差検定を行うことができます。結果は，多くの場合，主効果，交互作用のそれぞれについて，F値（F比），自由度（df），有意水準（危険率，p）が記載されていると思います。読み方は１要因分散分析の場合とほぼ同じです。pが５％水準や１％水準をクリアしていれば，有意な主効果や交互作用があると判断できます。

また主効果／交互作用が有意であったため，多重比較を行った，という記載にであうこともあります。この考え方は１要因の時と同じじととらえてもらって結構です。分散分析は，主効果があるかないか，交互作用があるかないかは検定してくれますが，どの群間に差があるのかは示してくれません。そのため，主効果もしくは交互作用が有意であり，いずれかの要因が３水準以上である場合は，多重比較で，どこの水準間に有意な差があるのかを検討するのです（２水準だったら見ればわかります）。多重比較の方法には１要因分散分析と同じように，テューキー法，LSD法，ダンカン法，シェッフェ法，ライアン法などが用いられます。なお，分散分析の考え方や計算式は，丁寧な解説書がありますから必ず自分でチェックしておいてください（絶対やってください）。

－２要因分散分析を使った論文記述例－

大学生活不安尺度の作成および信頼性・妥当性の検討

（藤井義久　1998　心理学研究　68，441-448．より抜粋）

大学生活不安の性差と学年差

抽出された三つの下位尺度のそれぞれについて，性と学年を要因とする２（男性，女性）×４（１年，２年，３年，４年）の分散分析を行った。その結果をTab1e 3に示す。それによると，“日常生活不安”と“評価不安”に有意な性差がみられ（日常生活不安：F〔1, 2775〕＝5.64，$p<0.05$；評価不安：F〔1, 2775〕＝13.05，p＜0.0l），いずれも女性の方が有意に得点が高かった。

また，同じく“日常生活不安”と“評価不安”には学年差がみられ（日常生活不安：F〔3, 2775〕＝2.99，$p<0.05$；評価不安：F〔3, 2775〕＝5.00，$p<0.01$）。Tukey法による多重比較を行った結果，いずれも学年が上がるにつれて得点が下がり，３年生以上が１年生に比べて有意に得点が低かった。また，本尺度の合計得点についても性差と学年差の検討を行ったが，いずれも有意差が認められた（性差：F〔1, 2775〕＝7.05，$p<0.01$；学年差：F〔1, 2775〕＝3.94，$p<0.05$）。さらに，すべての下位尺度および合計得点について有意な交互作用はみられなかった。これらのことから，女子学生の方が男子学生よりも不安を強く感じており，学年が上がり大学生活に馴染めば馴染むほど一般

に下がる傾向があるといえる。

Table 3　大学生活不安尺度の学年別および性別の平均値と分散分析

大学生活不安尺度の平均値および標準偏差									分散分析		
学　年	１年生		２年生		３年生		４年生		性差	学年差	交互作用
性　別	男性	女性	男性	女性	男性	女性	男性	女性			
被験者数	456	406	278	541	239	361	219	282			
日常生活不　安	5.07 (3.73)	5.41 (2.78)	4.23 (3.03)	5.76 (3.31)	2.21 (1.39)	4.12 (4.28)	2.03 (1.29)	3.02 (1.35)	5.64*	2.99*	*ns*
評価不安	5.57 (2.31)	6.22 (2.60)	3.77 (2.45)	6.33 (2.48)	3.30 (2.21)	3.86 (2.61)	2.05 (1.31)	2.99 (1.84)	13.05**	5.00**	*ns*
大　学不適応	1.14 (1.70)	0.70 (1.08)	1.23 (1.36)	0.97 (1.33)	1.30 (1.71)	0.71 (1.50)	0.93 (1.02)	0.54 (1.33)	*ns*	*ns*	*ns*
合計得点	11.79 (5.70)	12.33 (5.06)	9.23 (5.63)	13.06 (5.34)	6.80 (4.13)	8.57 (7.56)	5.01 (3.55)	6.55 (4.52)	7.05**	3.94*	*ns*

（注）　（　）内は標準偏差，** : p<0.01，* p<.05.

…ちょっと解説…

読み方は１要因分散分析の時と同じです。Fに続くカッコの中が自由度，＝の後がF値，そしてそれが現れる確率（つまり有意水準）の順に記されています。思い出しましたか。日常生活不安における性の主効果は，F［1, 2775］＝5.64，p＜0.05ですから，５％水準で有意な性差があることになります。性差は２水準ですから，一方が高ければ他方が低いというシンプルな関係なので，多重比較をしなくても表を見れば「女性の方が有意に得点が高い」ことがわかります。学年差については，「Tukye法による多重比較」が行われています。性差とは違って４水準ありますから，どの水準間に差があるのかをさらに検討したわけです。詳しい結果記載はないのですが，文章を読むかぎりでは，１年生と３年生，１年生と４年生の得点間に有意な差が認められていると考えられます。

なお統計をある程度知っていたら読み間違いはしないけれども，初心者だとちょっと引っかかるかなと思う文章があるので追記しておきます。それは，「Tukey法による多重比較を行った結果，いずれも学年が上がるにつれて得点が下がり，３年生以上が１年生に比べて有意に得点が低かった」という文章です。あなたはこの文章を，どこまで深読みしますか。特に２年生の得点のポジションについてです。Table 3にもどって，１年生の平均値と２年生の平均値を見てみましょう。学年差の主効果なので，この場合，性の別はつぶしてしまいます（先の説明の通りです）。表中には性をつぶした記載はありませんが，男女それぞれの人数が記載されているので，電卓で計算すると，１年生の平均値（日常生活不安）は5.23（｛(5.07×456)＋(5.41×406)｝÷(456＋406)），２年生の平均値（日常生活不安）は5.24（｛(4.23×278)＋(5.76×541)｝÷(278＋541)）となります。あれれ，あれれ，「学年が上がるにつれて得点が下が」っていない？　ここらあたりが書き方の難しいところで，多重比較で認められた差は，１年生と３年生，１年生と４年生の得点間です。つまり２年生は，１年生との差がないし，３年や４年とも差がない位置にあるのです*[10]。この分析から言えることは，文中の「３年生以上が１年生に比べて有意に得点が低かった」ということであり，「いずれも学年が上がるにつれて得点が下がり」というところは，単にその傾向を説明しているに過ぎません。でもこの文が入っていることにより，１年よりも２年生の方が低いのだろうと読んでしまう人が出てくる可能性を高めていると思います（個人的には，不要な挿入だと思います）。こういった，統計を少し知っている人は間違えないが，初心者だとうっかりしてしまうような記述が時にあるので，注意しながら読んでください。

*[10]　平均値で見ると１年生よりも高い値なのに，なぜ低い３・４年生と差がないのか，と思う人もいるでしょう。おそらく分散の状況などが影響しているものと考えられ

ますが，統計的には時々起こる現象です。

D．相関係数

２つの変数間の関連性を示す指標であり，最も頻繁に使われる分析方法の１つでしょう。それゆえに誤解も招きやすいので，分析の読み方だけでなく，使い方もちょっと勉強しておきましょう。

まず「関係」というものを，もう一度考えてみましょう（服部・海保,1996参照）。一口に関係と言っても，そこにはさまざなな種類が考えられます。

(１)関数関係

中学校（たぶん）で習った，一次関数や二次関数を思い出してください。基本的な式は書けますか？　xとyが，どういう関係にあるか説明できますか？　関数関係とは，右辺の変数の値が決まれば，左辺の変数の値が自動的に決まってしまうような関係です。E=IRという式がありましたよね。電流と抵抗か決まれば，電圧が決まる。そんな関係です。

(２)因果関係

対象となる２変数のうち，片方が原因，他方が結果である関係です。心理学の分野ではよく仮定される関係です（そしてよく問題にもなります）。因果関係が仮定できるためには，(a)原因変数が結果変数よりも時間的に先行していること，(b)理論的に必然性，整合性のある因果であること，(c)他の変数の影響を除いても原因変数と結果変数の間に共変関係があることが必要です。

(３)因果連鎖

「風が吹けば，桶屋がもうかる」式の，Aが起きればBがおきる，BがおきればCがおきる，といった連鎖の関係です。

(４)双方向の因果関係

その名前の通り，互いに一方が他方に影響を与える関係です。例えば，英語の成績と英語の学習意欲の関係。英語の成績が上がればより学習意欲がでてくるでしょう。そして学習意欲が出てくると，成績も上がるでしょう（たぶん）。このような関係のことをいいます。

(５)共通の原因によって生じる共変関係

自分が理科系に向いているかどうかを判断するときに，理科と数学の学力を参考にしませんか？　では，この理科と数学はどのような関係にあるのでしょうか。因子分析の考え方を思い出してください。そうすれば，因子として理系能力というものを仮定でき，それが理科と数学という異なったものに影響していると考えることができるでしょう。このような考え方を採用すれば，理科と数学は，理系能力因子という共通の原因によって生じている共変関係にあると考えられます。

(６)相関関係

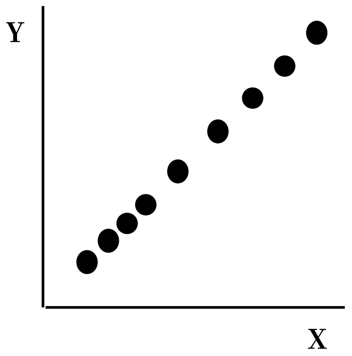
やっと本題が出てきました。相関関係とは…　これまで紹介したすべての関係を包括する関係です。万能選手ですから，非常に使いやすいものです。しかし逆に言えば，相関関係にあるといえば，それは因果関係かもしれないし，共通の原因によって生じる共変関係かもしれないという，不明確な表現でもあります。

では，相関係数の話に戻りましょう。関連を求める変数が量的変数である場合，最もよく使われるのが**ピアソンの積率相関係数（r）**です（一般的に「相関係数」といえばこれを指し

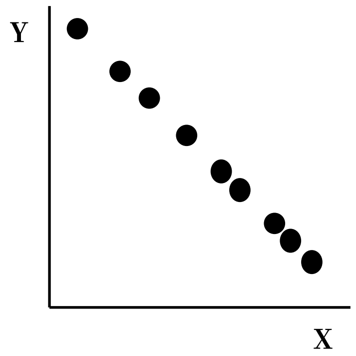
ます）。なお，質的変数の関連性の場合には**ファイ（φ）係数**や，スピアマンやケンドールの**順位相関係数**（**スピアマンのロー(ρ)**，**ケンドールのタウ(τ)**）を使ったりします。これらについては，各自でお勉強してください。

ピアソンの積率相関係数は-1から1までの値をとります（以下の図を参考にしてみてください。相関係数と分布状況の関係をイメージしてもらえればと思います）。また相関係数に対しては，**無相関検定**を行うことができます。これは得られた相関が０である確率を求め，それが小さければ（5%とか1%であれば），得られた相関係数は０ではないことになります。このような場合，それを有意な相関係数といいます。t検定の時と同じように，アスタリスクで表示されていることが多いです。

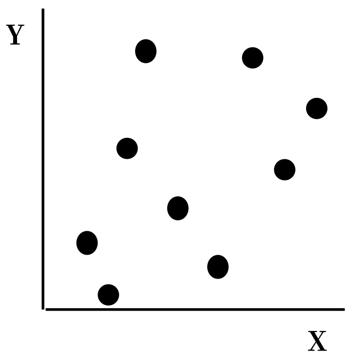
なお相関の強さの文章表現ですが，一応の目安としては，相関係数の絶対値が.2よりも小さいときは「ほとんど相関がない」，.2から.4くらいのときは「弱い相関がある」，.4から.7くらいのときは「中程度の相関がある」，.7を越えるときは「強い相関がある」と表現することが多いようです。



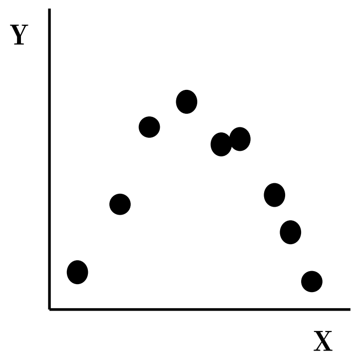
r=1 である場合



r=-1 である場合



r=0 になるような場合



相関係数は
どうなるでしょう

－相関係数を使った論文記述例－

ソーシャル・サポートの互恵性が青年の心身の健康に及ぼす影響

（周玉慧・深田博己　1996　心理学研究　67，33-41．より抜粋・一部加筆）

相関分析

各変数の平均と標準偏差をTable 1に示した。変数間の関係を検討するために，相関分析を行い，そ

の結果もTable lに示した。

４種類のSS（ソーシャル・サポートの略記）間の関係については，求められたSSと提供したSSとの間に，あるいは求めたSSと受け取ったSSの間に極めて高い正の相関関係が見られ，求められた程度に応じてSSを提供し，求めた程度に応じてSSを受け取っていることがわかった。また，求められたSSと求めたSSとの間に，あるいは提供したSSと受け取ったSSとの間に非常に高い正の相関関係が見られ，青年のSSのやりとりはかなり互恵的な状態にあることが示された。なお，4種類のSSの指標間の相関関係はすべて高く，これらのSSは相互に独立していないことも示された。

４種類のSSと負債感などの変数との間の相関関係はあまり顕著でなく，有意な相関関係が見られたところについてもその相関係数は小さかった。４種類のSSは負債感などの変数にあまり影響を与えていないことが示された。

負債感や負担感と不適応度や精神的自覚症状あるいは身体的自覚症状との間に有意な正の相関関係が存在し，負債感や負担感が強いほど，不適応度が増し，精神的自覚症状あるいは身体的自覚症状が大きくなることが示された。なお，不適応度，精神的自覚症状，身体的自覚症状の３変数間にもかなり強い正の相関関係が示された。

Table 1　各変数の得点および変数感の相関係数（一部抜粋）

	1.	2.	3.	4.	5.	6.	7.	8.
1. 求められたサポート	—							
2. 提供したサポート	.89 ***	—						
3. 求めたサポート	.73 ***	.70 ***	—					
4. 受け取ったサポート	.76 ***	.76 ***	.87 ***	—				
5. 負債感	.08	.03	.13 **	.13 **	—			
6. 負担感	.06	.01	.07	-.03	.33 ***	—		
7. 不適応感	-.05	-.09 *	.01	-.07	.27 ***	.45 ***	—	
8. 精神的自覚症状	.06	.03	.13 **	.06	.29 ***	.47 ***	.77 **	—
9. 身体的自覚症状	.14 **	.14 **	.12 **	.13 **	.20 ***	.28 ***	.52 **	.62 ***

*** p<.001, ** p<.01, * p<.05

…ちょっと解説…

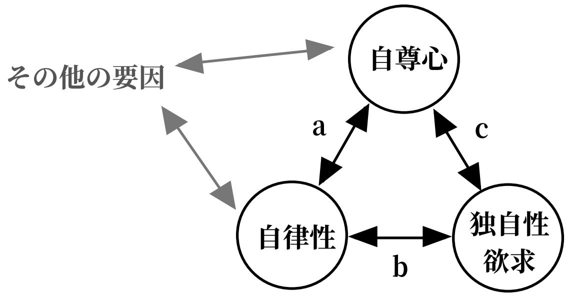
まあ，解説するまでもないでしょう。相関分析という言葉が出てきますが，相関係数を使った分析ということでいいですね。この論文では，危険率をアスタリスク１つから３つまでで表記してあります。このような場合には，ほぼ間違いなく，表の下の方に注記があって，記号と危険率の対応が示してあります。

もう少し補足しておくと，この論文では $r = -.09$ が５％水準で有意な相関であることが示されています。相関係数は，それを算出する集団の人数が増えれば増えるほど，小さい係数で有意なものになります。先に，相関係数の絶対値が .2よりも小さいときは「ほとんど相関がない」と記載することが多いと書きました。つまり，「有意な相関である」が「ほとんど相関がない」と一見矛盾したようなことになります。でもこれは矛盾していないのです。無相関検定は，得られた係数が０であるという帰無仮説を否定することになります。つまり帰結は「相関係数は０ではない」ということですので，基本的に相関が強いか弱いかということを示すものではありません。ごくまれですが，無相関検定の意味を取り違えている記述がありますので注意しておいてください。

E．偏相関係数

論文ではあまりお目にかからない分析ですが，私個人としては，心理学の分野では使い道の多いものだと思うのですが…。例えば，自尊心と自律性と独自性欲求の関連を考えてみます。これらが，下図のように三つどもえ状態にあることは容易に想像できます（自尊心が高

ければ自律性が高いだろう・自律性が高ければ独自性欲求が高いだろう・独自性欲求か高ければ自尊心が高いだろう（これらの逆回りもアリ））。



この中から，任意に２つの要素を取り出します。とりあえず，自尊心と自律性にします。自尊心と自律性の相関係数を求めることは，データがあれば簡単にできます。しかし，自尊心と自律性の間の相関係数は，aの強さを示しているわけではありません。三つどもえ状態を仮定した場合，そこには独自性欲求を経由した関連性も含まれます。つまり，bとcの経路を経由した影響も相関係数の中に含まれてします。さらにこの図の要因には取り入れられていない他の要因を経由した影響も含まれてきます。つまり自尊心と依存心の間に認められた相関係数は，純粋にaの関連の強さだけを示しているわけではないのです。

ではbやcの関連性を，自尊心と自律性の相関係数から外すことはできないか。これをやってくれるのが，偏相関係数です。簡単に表現すれば，bとcの関連性を一定にしたときのaの関連性を算出するのです。ここで注意しなければならないのは，bとcの関連性を一定にしたときのaの関連性を算出するのですから，測定されていないその他の要因の影響は取り除くことはできないことです。

なお，偏相関係数，無相関検定の読み方は，相関係数と同様です。

－偏相関係数を使った論文記述例－

女子短大生の進路選択に対する自己効力と職業不決断－Taylor & Betz(1983)の追試的検討－

（浦上昌則　1995　進路指導研究　16，40-45．より抜粋・一部修正）

職業不決断尺度得点間の相関係数，および，係数が求められる２下位尺度以外の下位尺度得点を統制した場合の，偏相関係数を算出した。続いて不決断各下位尺度と進路選択自己効力得点の相関係数，および，係数が求められる２尺度以外の尺度得点を統制した場合の，偏相関係数を算出した。算出された相関係数，偏相関係数を合わせてTable 5に示す。

まず，職業不決断の各下位尺度間の関連について検討する。Table 5に示された結果より，この不決断因子では，「情報・自信不足」の因子が中心になっていると考えられる。そしてこの因子を取り巻くように，他の４因子が配置されているのではないだろうか。また「モラトリアム」と「相談希求」の間には，有意な負の偏相関係数が認められた。不決断を測定する下位尺度間には低い値ではあるが負の偏相関係数が認められたことは，妥当性の観点から問題を含むと考えられる。しかしながら，「相談希求」は相談を欲しているという不決断の積極的な側面を示し，一方「モラトリアム」は決定を猶予もしくは回避しようとする不決断の消極的な側面を示していると考えれば，無相関もしくは弱い負の関連にあることは納得できるものといえよう。

次に，進路選択自己効力と不決断の関連について検討する。Table 5に示されるように，進路選択自己効力と不決断各下位尺度の関連は，相関係数と偏相関係数の間でいくつかの違いが認められる。相

関係数が示すところでは，進路選択自己効力は「情報・自信不足」，「葛藤」，「モラトリアム」と有意な負の関連にある。しかしながら偏相関係数では，進路選択自己効力は「情報・自信不足」とのみ，有意な負の相関を持つことが示された。「情報・自信不足」，「葛藤」，「モラトリアム」の３者は，中程度の相関関係にあり，かつ「情報・自信不足」が最も強く進路選択自己効力と関連するため，このような結果が得られたものと考えられる。

また，因子間の相関係数や偏相関係数より，「情報・自信不足」は不決断の各側面の中心的な存在であることが示唆された。これらのことから，進路選択に対する自己効力は，自己や職業についての情報不足や選択への自身の不足を通して，不決断と関連していると考えられる。

さらに，「相談希求」は，相関係数では進路選択自己効力と無相関であるが，偏相関係数では，低いながらも有意な正の相関関係にあることが示されている。この「相談希求」については，上述したように，不決断状態の一側面ではあるが，そこから脱却しようとする積極的な面を示していると考えられる。この点から，進路選択自己効力は「相談希求」と，本質的には正の関連を持つと考えることが可能であろう。このような関連が，偏相関係数として現れたものと考えられる。

Table 5　進路選択に対する自己効力および職業不決断各下位尺度間の相関係数，偏相関係数

	情報・自身不足	希望関連不安	相談希求	葛藤	モラトリアム	自己効力
情報・自身不足	―	.004	.329 **	.466 **	.334 **	-.482 **
希望関連不安	.217 **	―	.328 **	.143 *	.022	-.023
相談希求	.353 **	.392 **	―	.007	-.169 **	.170 **
葛藤	.570 **	.255 **	.262 **	―	-.011	.086
モラトリアム	.410 **	.051	-.008	.216 **	―	-.016
進路選択に対する自己効力	-.519 **	-.028	-.024	-.223 **	-.253 **	―

注）左下が相関係数，右上が偏相関係数
* p<.05，** p<.01

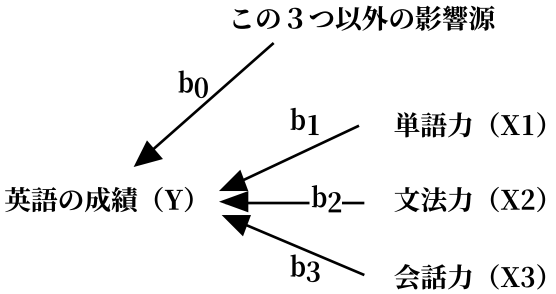
…ちょっと解説…

相関係数と偏相関係数の相違に注意して見てください。相関係数では11の有意なものが認められていますが，偏相関係数では８に減少しています。その分，関連がすっきりしていると思います（手前みそでしょうか）。わかりにくかったら，図にしてみましょう。少しすっきり理解できると思います。

F．重回帰分析

いくつかの変数を使って１つの変数を説明しようとするとき，また，ある変数に影響を与えている変数を，いくつかの候補の中から選出するときなどに使うのが重回帰分析です。

誤解をおそれず，重回帰分析を非常に単純に説明しましょう。例えば，英語の成績（要因Y）は，単語力（要因X1），文法力（要因X2），会話力（要因X3）によって推定することができると予測したとしましょう。でも，これらの要因すべてが英語の成績に同じ強さの影響を与えているとは考えにくいでしょう。そこで，それぞれの影響の強さを，単語力からの影響（b_1），文法力からの影響（b_2），会話力からの影響（b_3）とします。さらに，この３つでは説明できない部分の影響力をb_0としておきます。これを図に書くと以下のようになります。



数式にすると，要因Yの値をy，要因X1の値をx_1，要因X2の値をx_2，要因X3の値をx_3とすると，その式は　$y=b_1x_1+b_2x_2+b_3x_3+b_0$　と表現することができます。

重回帰分析においては，説明される要因（ここでは英語の成績）を**従属変数**もしくは**被説明変数**，あるいは**目的変数**と呼び，説明する要因（ここでは，単語力，文法力，会話力）を**独立変数**もしくは**説明変数**と呼びます。そして，それぞれの影響力の強さを示す値を**偏回帰係数**，これを標準化したものを**標準偏回帰係数（β）**と呼びます（ここではb_1からb_3）。なぜ標準化するかといえば，それぞれの独立変数はその測定単位が同じであるとは言えない状況にあるからです（ここがよくわからないという人は，**標準化**の意味について勉強しておいてください）。単位が違うと比較することが困難になるので，これを統一し，相互に比較可能なものにするのです。

さあ，ここらあたりまで説明すると，漠然とわかってきたのではないでしょうか。大きい標準偏回帰係数をもつ独立変数は，強く従属変数に影響を与えている。逆に，標準偏回帰係数が小さい独立変数は，あまり従属変数に影響を与えないということです。例えば，標準偏回帰係数が$b_2＞b_1＞b_3$という大小関係にあれば，英語の成績に最も強く影響するのは文法力であり，その次が単語力，最も影響力が小さいのが会話力ということがわかってきます。なお，この偏回帰係数については，有意性の検定を行うことができます。つまり検定の結果，有意と判定されたものは影響力が０ではないといえるものなのです。そして論文の結果解釈は，有意な偏回帰係数について解釈してあるはずです。

しかし，ここで問題になるのがb_0，つまり３つの独立変数では説明できない部分の存在です。このb_0がやたらと大きくなったらどうでしょう。たとえ有意な標準偏回帰係数が見つかったとしても，b_0の影響にかき消されてしまう可能性があります。その時に見るべき指標として，**重相関係数（R）**があります。この重相関係数についても有意性の検定を行うことができます。これが有意なもの（これが０であるという帰無仮説を棄却できる）であってはじめて，先の式の意味がでてきます。つまり重相関係数が有意でなければ，今回の予測は妥当ではなかったことになります。論文の結果では，重相関係数が有意であるということを踏まえてから，有意な偏回帰係数について解釈するという順序になっているはずです。なお，論文によっては，Rではなく，**重相関係数の自乗値（R^2）**（**決定係数**ともいう）や，**自由度調整済み重相関係数の自乗値（Adj. R^2）**が記載されているものもあります。

以上，いくつかの変数を使って１つの変数を説明しようとするときの流れに沿って重回帰分析を説明してきました。いくつかの候補の中から重要な要因を選出するときに重回帰分析を使う考え方については，もうわかりますよね。

－重回帰分析を使った論文記述例－

高校生の自我同一性に及ぼす信頼感の影響

（天貝由美子　1995　教育心理学研究　43，364-371．より抜粋）

結果　1）信頼感の各下位尺度と同一性地位判別尺度との関連

信頼感各下位尺度が，同一性地位を構成する３変数「現在の自己投入」「危機」「将来の自己投入の希求」をどの程度予測しうるかを調べるため，３変数各々の得点を目的変数，信頼感各下位尺度得点を説明変数として重回帰分析を行った（Table 3）。その結果，自分への信頼と他人への信頼は「現在の自己投入」「将来の自己投入の希求」に対して影響力を持つことが明らかとなった。また，「危機」には不信と他人への信頼が影響していることが示されたが，その説明率は相対的に低かった。

Table 3　同一性地位３変数を目的変数，信頼感下位尺度を説明変数とする重回帰分析（標準偏回帰係数）

	危機	現在の自己投入	将来の自己投入への希求
不信	-.221 **	.061 +	.039
自分への信頼	-.008	.325 **	.310 **
他人への信頼	.124 **	.142 **	.226 **
説明率（R^2）	.040 **	.200 **	.246 **

（** p<.01，+ p<.1）

註）以下「不信」尺度得点として記す場合には，値が高いほど不信が強いことを示すよう数値が変換されている。

…ちょっと解説…

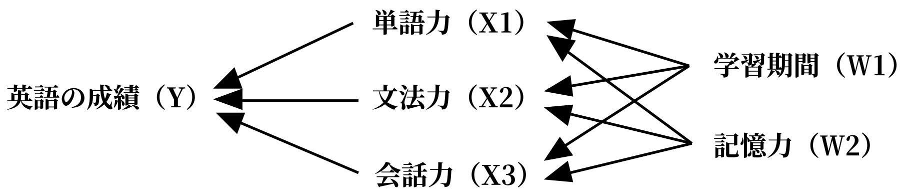
非常にシンプルな重回帰分析を使った結果の記述を示してみました。この表ではR^2が**説明率**という名前で記載されています。なぜこれが説明率なのかというと，それが，従属変数の分散のうち何%くらいが独立変数によって説明されているかを示す数値でもあるからです。Table 3でいうと，「危機」の分散のうち4%が「不信」「自分への信頼」「他者への信頼」で説明できることになります。このように説明すると4%分しか説明していないのに，それが有意なのか？と思う人も出てくるでしょう。そういう人は，もう一度重相関係数Rの検定の考え方を思い出してください。この検定は，重相関係数が０である確率がどの程度であるかということを問題にしています。検定で有意になれば，そこで言える帰結は「重相関係数は０ではない」ということです。つまり，わずかしか説明できない場合でも，それは０とは言えない場合は有意となるのです。

かしこい人はもうわかってきたと思いますが，重相関係数Rが有意と判断された場合でも，それが極めて小さい場合には，独立変数がたいした説明力を持っていない，つまりもう少しモデルを考え直したほうがいいのではないかと考えられる，ということがわかるはずです。小さいRでも有意であるからそれを強く主張するのか，有意であっても小さいから他の説明要因の存在を考えていくのか，このあたりのスタンスの取り方は研究者次第です。どの程度の値をどのように解釈しているのかという点に注意して読んでいくと，それぞれの研究者の考え方がわかってきて，それはそれで面白いと思いますし，その研究者と違う考え方を取るならば，さらなる発展のきっかけとなるでしょう。

G．パス解析

先の重回帰分析のところで，英語の成績は，単語力，文法力，会話力によって推定するこ

とができるとの予測を使いました。ここではちょっと表現を変えて，これを因果関係，つまり英語の成績は，単語力，文法力，会話力に影響を受けるとします。この因果についてさらに考えていくと，単語力などは記憶力に影響を受けるのではないかとか，英語を勉強した期間に影響を受けているのではないかと考えることができます。そうすると，次のような因果連鎖（**因果モデル**）を仮定することができるでしょう。



この図ような要因間を矢印で結んで表現したモデルを，**パス・ダイアグラム（パス図）**と呼びます。パスとはpath，すなわち道筋のことで，モデルの要因間をつないでいる矢印です。パス解析は，モデル中のパスの値を求めることを第１目的としています。一般的に，パス係数には標準偏回帰係数を代用します。つまり，単語力を従属変数，学習期間と記憶力を独立変数として重回帰分析を行い，その結果得られた偏回帰係数が，学習期間と単語力，記憶力と単語力をつなぐパス係数とされます。同様に文法力，会話力についてもそれぞれパス係数を求めます。次に，英語の成績を従属変数，単語力，文法力，会話力，学習期間，記憶力を独立変数として重回帰分析を行い，パス係数を決定していきます。ここで，英語の成績を従属変数としたとき，学習期間，記憶力も独立変数としなければならないのはなぜかという疑問が出てくるのではないでしょうか。先の図では，パスがひかれていません。しかし，それは意図的に引かなかっただけであり，パスがないことを踏まえたうえでのものではありません。つまり，計算をするときには，従属変数に先行する変数すべてを独立変数として計算を行います。もし，学習期間から英語の成績へダイレクトの有意なパス係数（偏相関係数）が認められたならば，モデル自体を考え直すことになります（最初に仮定したモデルに合致しないから，これは当たり前ですね）。

さて，このパス解析の方法がわかったら，論文を読むことはたやすいでしょう（ただし，重回帰分析を理解しているならばですが）。論文では，ほとんどの場合パス・ダイアグラムが掲載されており，その中にパス係数（偏相関係数）と重相関係数／その自乗値，および検定結果が示されています。

なお，パス解析を行って予想通りの結果が得られたからといって，それはパスの流れの方向性が証明された，または因果関係が証明されたとは言えません。なぜだかは各自で考えてください。ここのところを，微妙に取り違えたような記述もないことはないので注意しておいてください。基本的には，因果の妥当性は理論を組み立てる中で押さえておくべきものです（例に示したモデルでは，単語力などを英語の試験に先行して調査しておけば大丈夫だと思います。たぶん）。

－パス解析を使った論文記述例－

女子短大生の職業選択過程についての研究

－進路選択に対する自己効力・就職活動・自己概念の関連から－

（浦上昌則　1996　教育心理学研究　44，195-203．より抜粋・一部加筆）

次に，進路選択に対する自己効力，２つの就職活動，自己概念の変化の４変数を説明変数，職業的自己概念の変化を被説明変数として，パス解析を行った。専攻別の変数間の相関係数をTable 4（ここではTable 4は割愛）に，重相関係数，パス係数（標準偏回帰係数）をFigure 2およびFigure 3に示す。なおパス・ダイアグラムには，有意もしくは有意傾向が認められるパスのみを記入した。

まず幼児教育科生（Figure 2）の傾向について検討する。幼児教育科生の場合，進路選択に対する自己効力から「自己と職業の理解・統合」行動と「就職活動の計画・実行」行動への，有意，もしくは有意傾向にある正のパス（それぞれβ＝.371，.203）が認められた。これは，進路選択に対する自己効力が積極的な就職活動を導くという仮説について，それを支持する結果といえよう。また，自己概念の変化から職業的自己概念の変化への正のパス（β＝.420）も認められた。しかしながら，「自己と職業の理解・統合」行動，「就職活動の計画・実行」行動を説明変数，自己概念の変化を被説明変数とした場合，「就職活動の計画・実行」行動からの有意なパスは認められず，また「自己と職業の理解・統合」行動からのパス係数（β＝.282）は有意であったものの，重相関係数（R）は有意ではなかった。すなわち，就職活動から自己概念の明確化への影響については，モデルに示されるような連続性を積極的に支持することはできない。この結果は，本研究におけるパスの仮定に合致するものではなく，モデルの再検討の必要性を示すものといえよう。

次に教養学科生（Figure 3）の傾向について検討する。教養学科生の場合，「自己と職業の理解・統合」行動，「就職活動の計画・実行」行動に，進路選択に対する自己効力からの有意な正パスが認められる（それぞれβ＝.528，.423）。これは，進路選択に対する自己効カが，積極的な就職活動を導くという仮説を支持するものである。また自己概念の変化は，進路選択に対する自己効力，および「就職活動の計画・実行」行動からの有意な正のパスが認められ（それぞれβ＝.282，.247），職業的自己概念の変化は，「自己と職業の理解・統合」行動と自己概念の変化から，有意な正のパスが認められる（それぞれβ＝.301，.524）。この教養学科生の場合の結果は，Figure 1に示される本研究の仮説を，ほぼ支持するものと考えられる。

Figure 1　本研究で仮定される要因間の関連

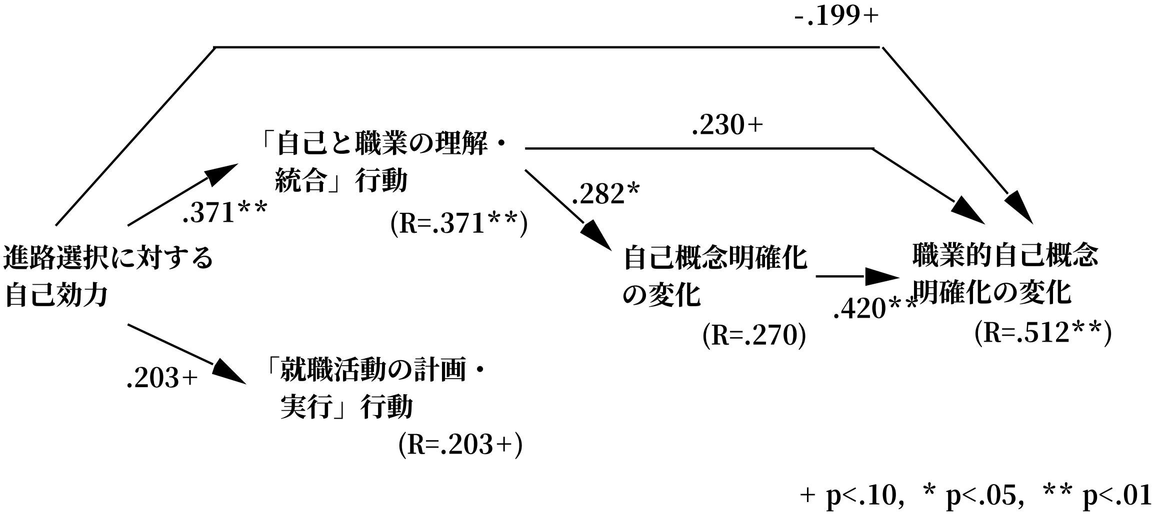


Figure 2　幼児教育科生におけるパス・ダイアグラム

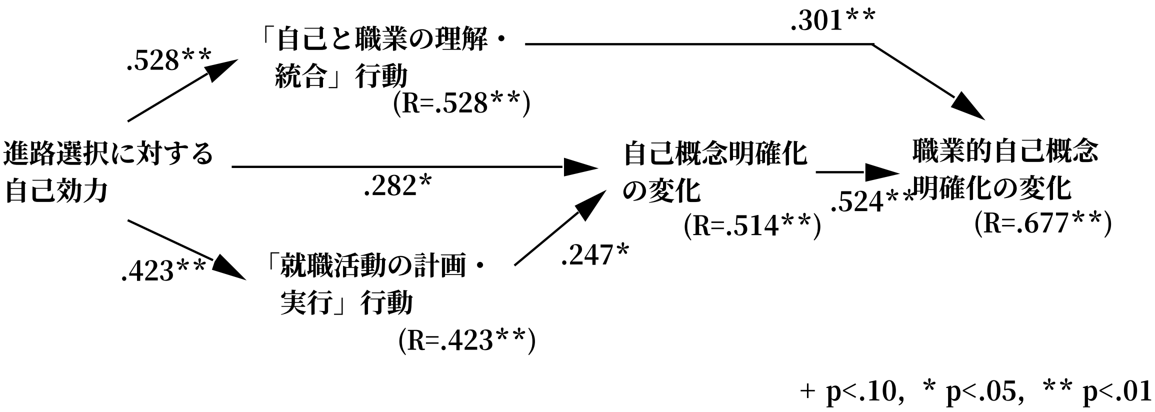


Figure 3　教養学科生におけるパス・ダイアグラム

…ちょっと解説…

この研究では，まずパス・ダイアグラムが仮説通りにつながるか，またそれは学科によって違うかどうかという点を検討しています。一見してわかるように，Figure 1には無かったパスがあったり，あるだろうと仮定されていたパスがなかったりしています。無かったパスがあるのは，先の説明からわかるでしょう。

この結果で最も問題となるのが，Figure 2の「自己と職業の理解・統合」行動からのパス係数（β＝.282）は有意であったものの，重相関係数（R）は有意ではなかったところです。重回帰分析のところで説明したように，重相関係数が有意ではないときは，標準偏回帰係数の意味はなくなってしまいますので，ここでもこのパス係数が意味を持ちません。つまり，「就職活動から自己概念の明確化への影響については，モデルに示されるような連続性を積極的に支持することはできない」ことになります。だから，「本研究におけるパスの仮定に合致するものではなく，モデルの再検討の必要性」が出てくるのです。

H．共分散構造分析

個人的には，これを使って論文を書いたこともあるし，おもしろい分析方法だと思います。論文集などを見ていても，けっこう頻繁に使われるようになっていますし，今後は論文を理解するためにどんな分析かを知っておく必要が出てくるでしょう（特に大学院に進学しようと思っている人は，ちょっとは触れておいたほうがいいかも）。

私がわかる範囲内で，この分析の利用について書いておきましょう。論文のいくつかを見てみると，やはりモデルの検証に使われていることが多いようです。分析に先立って１つ，あるいは複数のモデルを構築し，それが，もしくはそのうちのどれが一番データにあてはまるかを検討するなどです。このような利用として，よくでてくるのは次の２つの利用でしょう。

(１)確認的因子分析

探索的因子分析と確認的因子分析の違いは，因子分析のところでちょっと触れておきました。探索的因子分析については，因子分析のところで説明したような方法でかまわないのですが，確認的因子分析については，その方法として共分散構造分析を用いるものがいくつか見られます。確認的ですから，研究に先行して因子構造を理論的に確立しておいて，それにデータがどれくらいそってくるかというような検討がなされています。例えば，土肥の研究（土肥，1996）などを参考にしてみてください。

(２)因果モデルの検討

パス解析を思い出してください。パス解析は計算方法として重回帰分析を繰り返すということを覚えていますか。そしてパス解析の目的の１つがパス係数を算出することであったことも思い出してください。つまりパス解析では，その因果モデル自体がデータを反映したものかどうかという検討ではなく，一本一本のパスが有意であるかどうかという点に検討が集中しています。そのためモデル全体のデータへのあてはまり具合というものは二の次になっているのです。共分散構造分析では，GFI（適合度指標），AGFI（自由度調整済み適合度指標），AIC（赤池情報量基準）などの，モデル自体がデータに適合しているかどうかを示す指標が用意されており，それによって複数のモデルを比較検討して，いずれが最もあてはまり具合が良いかという判断を下すことができます*[11]。

なお，パス解析のところで引用した私の論文を共分散構造分析を使って再分析した豊田の記述（豊田,1997）もありますので，読み比べてみるのも面白いと思います。

他にもメリットは色々とあるのですが，これ以上の解説は，私には力不足です（教えてください，と私のところにやって来ても「自分で勉強しろ」と言うでしょう…）。とっつきやすい本（豊田・前田・柳井,1992）がありますので，ざっと読んで，こんな分析なんだと思うだけでも，論文を読みやすくなるのではないでしょうか。なおこの分析はSPSSではできませんので（大学のSPSSはR.4.0と古いので），SASを使う必要があります。

*[11]　GFI，AGFIについては，モデルとデータの一致度が高くなれば１に近づきます。いずれも0.9が適合性が高いかどうかを判断する１つの目安になっています。さらにGFI，AGFIの値の差が小さいほうがよいとされます。またAICは，複数のモデルの適合度を相対比較する時に使う指標です。値が小さいほどよいモデルとなります。例えばAというモデルのAICが36.44，Bは27.50だったと

したら，モデルBの方がよいモデルと判断されます。

I.その他

ここでは特に**χ^2（カイジジョウ）検定**について説明しておきます。これは**質的データ**（度数，頻度などのデータ）の分析によく使われる方法です。観察法などを使った乳幼児研究などでよくお目にかかりますが，質問紙法を使った研究では，それほど頻繁には使われていないようです。ただし，「はい」－「いいえ」といった，２段階で回答を求めるような調査方法をとった場合には使われているでしょう。

この検定の利用には，大きく３つほどの使い方があります。それらは，適合度の検定，独立性の検定，母集団の比率の等質性の検定です。いずれもχ^2分布を用いた検定であり，論文には「χ^2検定を行った」と書いてあることが多いと思います。

まず適合度の検定ですが，実際にデータとして得られた度数（**観測度数**）が，理論的に導かれる度数（**期待度数**）と一致するかどうか検討するものです。例えば，男子100人，女子200人から性別に関係なく30人を選ぶという選抜があったとすると，理論的には男子10人，女子20人が選ばれることになります。しかし，実際に得られたデータでは，男子13人，女子17人であった場合，これはよくあるブレ，誤差の範囲なのか，そうとは言えないのかについて検定します。

次に独立性の検定ですが，２つの属性の間に関連性があるかどうかについて検定するものです。例えば，１つ目の属性を性別，もう１つの属性を野球のファンかそうでないかとしておきます。野球のファンかそうでないかに性別は関係ないとすれば，この２つの属性は独立であり，関係があれば２つの属性は独立ではないといいます。このような関連性を調べるときにも，χ^2検定を使います。

最後に，母集団の比率の等質性の検定ですが，ある母集団（a）からA，B，Cという３つのカテゴリーに属する人を抽出した場合の，Aa，Ba，Caの人数の比率と，違う母集団（b）からA，B，Cという３つのカテゴリーに属する人を抽出した場合の，Ab，Bb，Cbの人数の比率が等しいかどうかを検定するものです。例えば，文学部学生と理学部学生で（これが母集団の違いです），所属する部活（カテゴリーです）に差があるかどうかを検討する場合などに使われます。差がなければ，例えば文学部の中で水泳部に属している者の比率と，理学部の中で水泳部に属している者の比率は同等になるはずです。

これらの検定結果には，**χ^2値**，自由度（df），有意水準（危険率，p）が記載されているはずです。文中では「χ^2=15.89，df=4，p<.01」とか，「χ^2(4)=15.89，p<.01」などとなっているはずです。ただし，結果の解釈にはそれぞれの用いられ方に対応した配慮が必要です。まず適合度の検定ですが，この場合の帰無仮説は「観測度数と期待度数が一致している」ですから，有意な結果が得られた場合には，理論値と実際のデータは一致していないということになります。独立性の検定では帰無仮説が「２つの属性は独立である」ですから，有意な結果が得られた場合には，その２つの属性は独立でない，つまり関連があるということになります。母集団の比率の等質性の検定では，帰無仮説が「等質である（比率は同じ）」ですから，有意な結果が得られた場合には，「等質ではない（比率は違う）」という帰結になります。多くの場合，「母集団の比率の等質性の検定を行った」とかという書き方ではなく，単に「χ^2検定を行った」という書き方になっていますので，読むときに注意しておいてください。

なお，この検定を行う場合，各セルの人数が少なすぎる（５を下回るようなセルがあれば）と，**イエーツの修正**を施した検定が勧められています。また０のセルがあれば，χ^2検定は使えません。また，χ^2検定の後に**残差分析**が行われている場合があります。これは，観測度数と期待度数の差についての検定ですので，これが有意であれば，そこは差があると言えることになります。

これまで，代表的な統計手法とその読み方について説明してきました。これだけ知っていれば，６～７割方の調査系論文は統計手法でつまることなく読めると思います*[12]。後は，時々お目にかかる**クラスター分析**，**数量化Ⅰ類**，**数量化Ⅲ類**あたりでしょうか。これらについては出てくる頻度が低いので，ここでは説明しません。論文を読んでいる中で出会ったときに調べてください。

このあたりをどうやって勉強したらよいのか，と聞かれることがあります。学問としての統計を勉強したい人は，間違いなく数式理解に向けた勉強が必要です。でも研究の手段として，また大学院入試に向けてなどという場合であれば，ある程度の数式理解とできるだけ多くの論文を読むこと。そして知らないものに出会ったら調べること。これを繰り返すしか無いでしょう。一気に身に付くものではありませんので，少しずつ自分の知識にしていってください。

*[12]　ここで紹介した統計手法は，特に調査系の論文読解に必要なものです。乳幼児研究に多い観察を主とした研究では，主に質的データを扱う**ノンパラメトリック検定**（今回触れたものの中ではχ^2検定がこれにあたります。他のものは**パラメトリック検定**と呼びます）についてさらに勉強する必要があります。また実験系の論文を中心に読んでいくのであれば，分散分析に関するさらに詳細な知識が必要になるでしょう。大学院志望の人は，当然両方を勉強しておいてください。くれぐれも，これで終わり，これで全部ということではありませんので！

参考文献

●本文中で取り上げたもの…

土肥伊都子　1996　ジェンダー・アイデンティティ尺度の作成　教育心理学研究　44,　187-194.

服部環・海保博之　1996　Q＆A 心理データ解析　福村出版

豊田秀樹　1997　測定・評価と共分散構造モデル　教育心理学年報　36　119-127.

豊田秀樹・前田忠彦・柳井晴夫　1992　原因をさぐる統計学－共分散構造分析入門－　講談社ブルーバックス　（請求番号：417L/1664）

●ベーシックな教科書…

海保博之（編著）　1985　心理・教育データの解析法10講（基礎編）　福村出版　（請求番号：140K/906/v.0-1　DB1/-001455　他あり）

海保博之（編著）　1986　心理・教育データの解析法10講（応用編）　福村出版　（請求番号：140K/906/v.0-2　DB1/-001385　他あり）

●分散分析をもうちょっと勉強したい人へ…

山際勇一郎・田中　敏　1989　ユーザーのための教育・心理統計と実験計画法　教育出版

田中　敏　1996　実践心理データ解析：問題の発想・データ処理・論文の作成　新曜社　（請求番号：140K/1095）

●辞典…

芝　祐順・渡部　洋・石塚智一（編）　1984　統計用語辞典　新曜社　（請求番号：R/417/1113/B）

●個人的に好きなもの…

古谷野　亘　1988　数学が苦手な人のための多変量解析ガイド　川島書店

●何となくわかったような気になれるもの…

石村貞夫（作）　1995　マンガ統計手法入門　CMC