

「キャリア教育研究」掲載の論文を読む
—基本的統計資料の理解を中心に—

資料

日本キャリア教育学会 研究推進委員会

2016年11月6日

対象論文

『キャリア教育研究』第34巻第1号掲載

白井利明著 「高校生のキャリア・デザイン形成における回想展望法の効果」

本資料は、2016年11月6日に開催された日本キャリア教育学会 研究推進委員会企画の研修会の資料です。著作権は日本キャリア教育学会 研究推進委員会にあります。日本キャリア教育学会会員もしくは非会員が自身のためにこれを印刷することは自由に行うことができます。学会員が、学校、勉強会等で非会員に配布するために印刷、複製する場合には、日本キャリア教育学会事務局（jssce-post@bunken.co.jp）まで、連絡者氏名、会員番号、利用目的、部数をご連絡ください。事務局より特段の返信はいたしません。この連絡をもって複製、配布を許可いたします。なお、電子データの複製、再配布はご遠慮ください。

論文の利用および本資料の公開を快諾していただきました白井利明先生に、ここに記してお礼申し上げます。

1. 論文とは

まずは、論文が書かれる目的は何か、どのような特徴をもっているのかというところからスタートします。この点をしっかり理解できれば、論文を読むことはもちろん、統計の意味の理解や、ひいては自分で研究をすることなどが容易になると思います。

次の2つの質問に答えてみてください。

質問1：論文が書かれる目的は何でしょうか？

質問2：論文とよばれる文書の特徴は何でしょうか？

論文とは

辞書では、学術研究の成果を理論的に述べた文章、などといった説明もされていますが、「論文」とは概ね、研究を通して得られた知識・知見を他者（読者）に伝えるために、特定の形式で記述された文書といえると思います。ここで「研究を通して得られた知識・知見を伝える」という部分の理解の仕方が、ひとつのポイントになります。それは、結果として得られた「知識・知見だけ」ではありません。「どのような研究を通して得られた知識・知見なのか」という部分を含んでいます。この点を開示することは、結果のねつ造を防いだり、追試やさらなる研究の発展に役立つ情報を提供することになります。

そしてもうひとつ、「他者（読者）は専門家・研究者が想定される」ということです。これは「論文は読みにくい」と評される原因の一つかもしれませんが、論文の執筆者には読者として専門家・研究者を想定するという共通理解があります。そのため、専門家や研究者には読みやすく（ムダがなく理解が早くできる）、想定に入っていない読者には読みにくい（知らない用語が多いのに説明不足）ものと映るでしょう。『キャリア教育研究』に載っている論文も同じです。それゆえ、論文を読むには知識が不可欠です。初心者用の論文というものはありません。

論文の表記的、内容的な特徴は、以上のような目的によってある程度規定されます。すなわち、専門家や研究者が、研究の過程を踏まえて結論を理解できるよう示すことに適した形式が特徴になります。以下はその特徴の一例です。

- ・ 図表を含め、必要不可欠な量で書かれる
- ・ 余談はほぼ無い（読者の推測を求めるような記述はない）
- ・ 形式は、研究方法が同じであれば、体裁も似てくる／用いている研究方法によって異なる

今回は、対象論文のような、介入を伴う実験的研究（研究法）の論文に焦点をあて、以下でその構造について触れます。他の実験的研究はもちろん、質問紙等を用いた調査研究など（仮説検証をねらったもの）も同様な構造をもつ論文になることが多いです。

論文の構造

介入を行うような実験的研究や、仮説の検証を目的とする調査を扱った論文では、以下のような構造になることが一般的でしょう。

問題→目的→方法→結果→考察 +引用文献

問題…どこに問題や課題があり、そのために何が明らかにされなければならないのかということ、従前の研究成果等をふまえて明確にし、本研究で明らかにするポイント（仮説）を整理する部分。

目的…問題の指摘から導かれる、この研究の目的を端的に述べる部分。

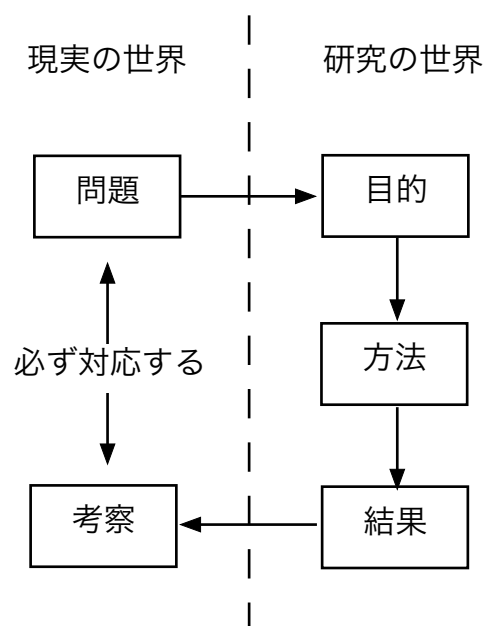
方法…どのような方法を用いて目的にアプローチしたのかを紹介する部分。後の研究者が追試を行って検討できるだけの必要十分な情報を示す。

結果…その方法を用いて研究を行った結果はどうであったのかを、分析方法を含めて示し、どういうことが明らかになったのかを説明する部分。基本的に事実のみを述べる。

考察…結果が、問題としたことに対してどのように寄与するのか、またどのような示唆を与えるものなのかを考察する部分。結果や先行研究をもとにした妥当な推測を含む。

引用文献…本文中に引用した文献の出所を明記した部分。掲載誌によって形式は多様。参考文献は、記載しない場合が多い。

なお、論文は現実の世界と研究の世界を行き来します。「現実が現実、研究は研究」といった認識がなされるのは、このイメージからだと思います。しかし、研究は現実から生まれ、かならず現実に戻されるという構造になっています。考察は研究成果を現実に戻す作業になる場合が多いといえるでしょう。



「高校生のキャリア・デザイン形成における回想展望法の効果」の構造

「問題と目的」部分

- 第1段落 時間的展望という鍵概念と過去からのつながりへの着目
- 第2段落 従来の介入法の紹介と問題点の指摘
- 第3段落 白井（1994, 2010）による提案
- 第4段落 残されている課題
- 第5段落 本研究の目的

ポイントは…

- ・ 高校生を対象としている
- ・ 回想展望法を授業の中で展開している
- ・ 自己の気づきと職業選択の促しに、回想展望法がどのように関連しているのかを検討する

「方法」部分

- ・ 誰に、どのように実施したのかをしっかりと説明（1と2）
- ・ どのような内容を実施したのかをしっかりと説明（3） ←実践研究の特徴
（このような実践研究の場合、授業の目的と研究の目的が異なる場合も多い。
授業の目的とその達成度についても3に書かれている）
- ・ 関連について検討するための指標についてしっかりと説明（3と4）

ポイントは…

- ・ 目的部分の表現と、測定されている指標から、プレ・ポストで変化を記述する研究ではないことに気づきましたか？
- ・ 方法部分で、「自己の気づき」「職業選択の促し」に関する指標がない（明示されていない）ことに気づきましたか？

★論文を読む際には、一度「方法」までで止めて、「この後どのような分析が行われるのだろう」と想像してから結果を読むという姿勢をおすすめします。

2. 本論文を読むための統計的知識

一口に統計といっても、多様な内容が含まれますし、分析の手順のようなものも含まれます。そこで、まずは用語や手順の確認の意味をこめて、その概略をまとめておきます。

1. 質問紙を使ってデータを集め、数値化する

一般に「測定」とよばれる過程です。あまり知られていないことかもしれませんが、数値には4種類あり、その種類によって利用可能な分析方法が異なります。今回は、その種類についても取り上げます。

2. 因子の得点化が意図されるならば、因子分析を行う

これは、必要であれば行うことであり、論文によっては含まれない場合もありますが、今回の論文では用いられています。キャリア研究でもよく言及される「因子」とは何か、どのように把握し、数値化するのかということについて紹介します。ここも「測定」に関わる部分とってよいでしょう。この段階を経て、分析に使える数値が全てそろふこととなります。

3. 以後の分析に使う変数すべてについて、基礎的な統計量を開示する

基礎的な統計量とは、平均値などの代表値と、標準偏差などの散布度の指標のことを指す場合が多いです。論文では、この点はほとんどの場合で明示してありますが、その値に対して積極的に言及されることは少ないでしょう。ここは軽く触れるに留めます。

4. 目的に応じた分析を行う

出そろった数値を用いて、目的に応じた分析（統計処理、検定など）が行われます。今回の論文では相関係数と重回帰分析が用いられていますので、その説明を行います。また検定が用いられているので、検定の概略についても説明を加えます。

「測定」とは

測定とは、測定したい特徴に対して、ある基準を適用して数値を与えることです。長さ（測定したい特徴）を、モノサシ（基準）を使ってはかることは測定の一例です。すなわち測定をするためには、「測定したい特徴」と、用いる「基準」が不可欠です。論文では、主に「方法」の部分で、何を（測定したい特徴）、何で（基準）測定したのかが記述されますが、あまり明文的ではなく、「問題」部分から読者が適切に推測する必要がある場合もあります。

測定したい特徴は、直接観測できるものと、直接は観測できないものに大別されます。前者は「長さ」や「重さ」、「回数」、「タイム」などが該当します。これを顕在変数（観測変数）とよびます。後者には心の動き、機能などの直接観測することができないもの（構成概念）、たとえば「キャリア発達の程度」や「能力」、「学力」などが該当します。これらは潜在変数とよばれます。

今回の論文では、「方法」と「性別」、「回想された職業数」が顕在変数、残りが潜在変数といえるでしょう。しかし、「日常的回想」や「回想の使用」など、潜在変数とはいえど、実際に測定されています。つまりこれらは顕在変数でもあるのです。

次に与えられる数値ですが、これらにも種類があります。これらの数字には、それが持つ情報量、精密さから、一般的に4つの種類（水準）が想定されています。

呼称	操作の特徴	条件
名義水準（尺度）	比較も演算も不可	
順序水準（尺度）	大小比較	順序性がある
間隔水準（尺度）	+、-のみ可能	等間隔である
比率水準（尺度）	四則演算が可能	原点0をもつ

名義水準

この水準の数値には、 $1 < 2$ といった大小関係はなく、また $1+3=4$ といった計算にも使えません。 $1 \neq 2$ といった、それぞれの数は他の数とは違うということだけを意味します。この水準は数値として把握されるものの質的変数（定性的な変数）であり、以下のもの（量的変数）とは区別されて扱われます。今回の論

文では、「方法」と「性別」が該当します。

順序水準

順序水準では、 $1 < 2$, $2 < 5$ といった大小関係（前後関係）が生まれます。なお、昇順でも降順でもよく、ポイントは特定の順序を数値の順が示すという点にあります。またこの水準の数値は、 $1 \neq 2$ といった名義水準にあったそれぞれの数値は他の数値とは違うという意味も持っています。すなわち名義水準の数値が持つ情報に、大小関係（前後関係）情報が加わったものです。 $1+3=4$ といった計算には使えず、書籍によっては質的変数に分類される場合もあります。今回の論文では使われていません。

間隔水準

間隔水準での測定には、等間隔である選択肢（等間隔であるという前提の選択肢）を準備し、当てはまる場所にチェックを求めるという方法がよく使われます。等間隔という前提をもつ選択肢を使って数値化しているところがポイントです。この数値は、順序水準の数値が持っていた情報に、さらに等間隔であるという特徴が追加されます。この等間隔という特徴のため、「複数項目の数値を合計する」といった操作が可能になります。それゆえ、項目の平均値を計算することができます。

今回の論文では、「日常的回想」、「回想の使用」、「一貫性の予想」、「一貫性の理解度」、および事後質問で用いられた 11 の項目が該当します。平均値が算出されている場合、また次の比率尺度で注目される原点の存在が考えられない場合、間隔水準と見なされていることとなります。

比率水準

比率水準は、その名が示す通り、得られた数値で比が計算できるという特徴を持ちます。比を計算するためには、数値が等間隔であることに加え、原点を持っていることが不可欠です。比率水準の数値は、原点をもちます。例としては、身長や体重、何かに要した時間、試行などの回数等があります。今回の論文では「回想された職業数」が該当します。

測定された結果の数値は、それ自体をデータとして分析が行われますが、潜在変数の測定の場合、さらに処理を施した後に次の分析へと進む場合も多くあります。これは、潜在変数をより適切にとらえ、測定するためになされます。そこで用いられる代表的な方法が因子分析です。

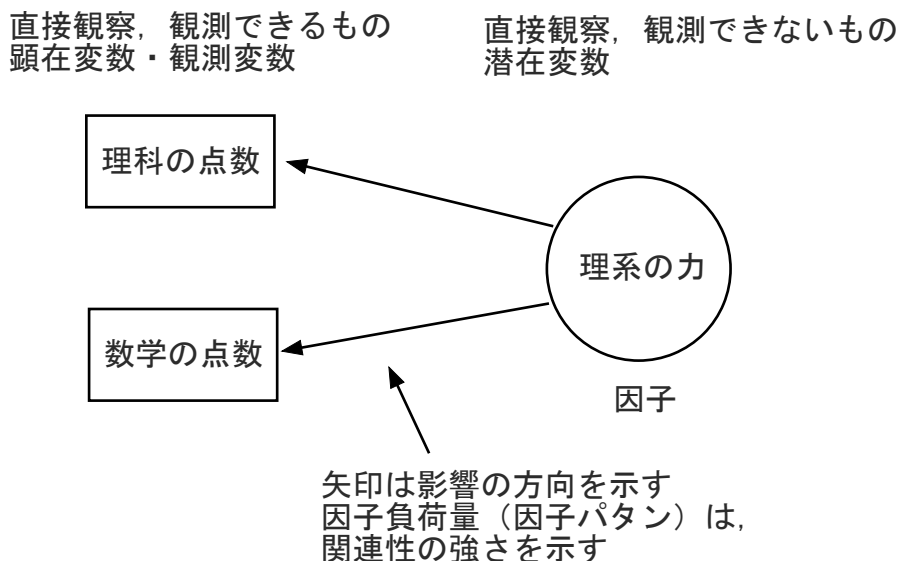
因子について

「因子 (factor)」とは、各項目の背景にあり、項目への回答に影響を与えているものと仮定されます。因子は、直接観察したり把握することはできないものです (すなわち、潜在変数です)。実際の現象を説明するために、その存在が仮定されたものです。

たとえば、理科や数学の点数が他の教科に比べて良い生徒がいるとします。こういう生徒に対して、「理系の力がある」と評したりします。このことを例にとります。

「理系の力がある」という評は、換言すれば理科や数学の点数が良いことを説明しようとしたものです。「理系の力」という概念の存在を仮定し、理科や数学の点数が良いという現象の説明に採用しているということです。ここで、「理科や数学の点数」は観察されるもの (顕在変数) ですが、「理系の力」は観察されていないこと。また、それは直接把握することもできないもの (潜在変数) であることを確認してください。以上の様相を図示すると下図のようになります。

因子分析とは、複数の直接観測された変数の背後にある要因 (因子) を見つけだそうとする方法といえるでしょう。因子はいくつあるのか、またそれらの因子はどの項目に影響しているのか、どのような特徴の因子なのか、などといった検討に用いられます。今回の論文では、事後質問で用いられた 11 の項目の背後にある因子を見つけ出し、それを測定するため (それに数値を与えるため) に用いられています。



論文では、「結果」の最初に因子分析について記述されています。おおまかな因子分析の流れは、①実際に利用した項目群の背景にいくつの因子を仮定することが妥当かを判断する。②決めた数の因子を抽出し、軸を回転させ、その結果を表にする。③表になった因子負荷量から解釈し、それぞれ因子に名前を付ける。④特定の因子と対応の強い項目を用いて、それらの項目の合計や平均を算出し、これを因子の得点とみなす。というものです。以下では、論文の記載と対応させながら、各段階について説明します。

①については「主因子法」を用い、「固有値の推移」から、2つの因子を抽出するという判断がなされています。前者が因子を抽出する方法の名称で、他には「最尤法」「最小二乗法」などがあります。後者は抽出する因子数を決定する判断に関する基準で、他には「スクリー基準」「固有値1以上の数」「平行分析の結果」「因子の解釈のしやすさ」などが参考にされることがあります。

②に関しては、軸の回転に「バリマックス回転」が用いられたことが記されています。軸の回転のメカニズムについては今回は触れませんが、それは因子分析の結果を解釈しやすくするために施されます。他には「プロマックス回転」などがあります。また論文には、項目の取捨選択を行ったことも記載されています。「(因子) 負荷量」とは、前ページの図のように、因子と項目の対応関係の強さを示します。基本的には-1.00から1.00までの値をとり、絶対値で1に近いほど関係が強く、0に近づくと関係が弱くなります。因子と項目の関係はより明確にしておきたいので、負荷量の値を参考に、因子との関係が弱い項目や、複数の因子から同程度に影響を受けるような項目を排除しています。このような操作を行い、表3に示される結果を得ています。なお、回転に「プロマックス回転」などが用いられる場合は、「因子パタン (パターン)」が掲載されます。

表3から、項目5, 4, 3, 6は第1因子 (F1) との関連が強く、項目1, 7, 2, 9は第2因子 (F2) との関連が強いことがわかります。③では、このような関係の強い項目との対応から、各因子がどのようなものと考えられるのかを検討します。前ページの図にあるように、関係の方向性 (因子が項目に影響してる) も踏まえつつ、それぞれ因子に適切な名前を付けます。

それぞれの因子と関連の強い項目群を選出して、平均値と標準偏差を算出する過程が④になります。今回の論文では、先の手順から関連の弱い項目はすでに除いてありますので、それぞれの因子で4つの項目が使われています。特定の因子と対応の強い項目を用いてその合計や平均値を算出し、これを因子の得

点とするわけですが、この操作により、項目それぞれへの反応ではなく、その背後にあってこれらの項目へ影響しているもの、つまり因子を測定しているとみなすのです。なお、それぞれの4項目の和を求める前提として、それらの項目群に内部一貫性が認められるかどうかの確認が必要です。 α 係数は内部一貫性の指標です。絶対的な基準はありませんが、概ね.70以上を目安とすることが多いようです。そのため、この論文でも、.75と.68では異なった評価表現が使われています。

以上の手続きを経て、事後質問で用いられた11の項目の背景にある2つの因子に数値を与えています。

ここで再度確認をしますが、この事後質問で用いられた11の項目は、今回の授業が、将来の職業選択を考える際のきっかけになったかどうかを確認するために用意されたものでした。すなわち、何のきっかけになったか、どの程度なったかを測定していることになります。何のきっかけになったかということは、2つの因子、すなわち「自己への気づき」「職業目標の認識促進」ということのきっかけになったといえます。そして、どの程度きっかけになったかということが、それぞれの得点として得られたわけです。

★実はここに至るまで、目的部分にあった「自己の気づき」「職業選択の促し」という記述が指し示すものが読者には理解できない構造になっています。因子分析を伴うような研究では、途中まで重要な概念の中身が明確にわからない論文もあります。

因子分析結果を読む際の注目点・注意点

・因子分析の結果から研究者は因子名をつけます。ここはその研究者の考え方が尊重される場所ですが、唯一の正解があるわけでもありません。それゆえ、異論も出やすい部分です。以後、因子名を使って結果が記述されるので、読者なりに因子の意味するところを把握しておくといよいでしょう。

平均値と標準偏差

平均値 (M)

間隔，比率水準で測定されたデータの代表値のひとつです。外れ値の影響を受けやすく，分布の様子が正規分布に近い場合に使われることが多いものです。

代表値とは，データを要約し，その中心的位置を表す値のことで，平均値の他に中央値，最頻値などがあります。それぞれに特徴があるため，データの特徴により使い分けます。

標準偏差 (SD もしくは S) ・分散 (S^2)

データ全体の散らばりの程度をあらわす散布度のひとつです。平均値と個々の値との差 (偏差) を用いて計算するため，平均値とともに用いられます。

ダミー変数

性別のような質的変数に，0 と 1 (他の場合もあり，この論文では 1 と 2 が与えられている) を与えたもののことです。

今回の研究で用いられている相関係数や重回帰分析は分散，共分散を使って計算されます。そのため名義水準で測定された変数は，当然用いることができません。そこで，たとえば「女性ではない--とても女性である」といった連続性を想定し，男性と女性をその両極付近にあるとみなして，0 と 1 を与えたものがダミー変数です。このような見なしにより，0 と 1 は比率 (もしくは間隔) 水準となります。これによって平均値や分散を計算することができます (意味をもちます) が，あくまでもこのような操作によって作られている変数ということには留意が必要です。

補足：論文の数値

これは慣例的なものですが，論文では「0.15」と「.15」という小数の表記が混在します。どちらも「0.15」の意味なのですが，たとえば相関係数のように，-1 から 1 の間しかとらないような数値もあります。このような数値の場合は，小数点前の 0 を省き，そうでない数値の場合は 0 を省かないことが一般的です。

相関係数

2変数間の直線的関連性の強さを表す指標です。単に相関係数というと、間隔もしくは比率水準の変数間の強さを表す「ピアソンの積率相関係数 (r)」を表すことが一般的です。論文では「ピアソンの相関係数」と表記されていますが、同じものです。他には、2変数がともに順位水準の場合に使われる、スピアマンの順位相関係数やケンドールの順位相関係数など多種あります。

確認：「強い直線的関係」と「弱い直線的関係」の例を以下に図示してください。



「ピアソンの積率相関係数」(以後は単に相関係数)は、2変数の共分散をそれぞれの標準偏差の積で割ったものです。

それぞれの標準偏差の積で割ることで共分散の単位がそろえられ、 r は-1から1までの値をとることになります。 r が正の値、すなわち一方の変数が増加すると他方も増加する関係の場合を「正の相関」と、 r が負の値、すなわち一方の変数が増加すると他方は減少する関係の場合を「負の相関」とよびます。

相関係数を読む際の注目点・注意点

- ・曲線的な関係の場合には適用できないので、相関係数が用いられている場合は、少なくとも研究者は変数間に直線的な関係を仮定していたと考えてよいでしょう。
- ・相関係数自体は間隔水準なので、たとえば .30

共分散
の公式
$$S_{xy} = \frac{\sum (X - \bar{X})(Y - \bar{Y})}{N}$$

相関係数
の公式
$$\frac{S_{xy}}{S_x S_y}$$

と .60 という 2 つの相関係数を使って、「後者は前者の 2 倍強い関連にある」などと考えることは間違いです。

・強い相関係数が認められたことは、必ずしも特定の関係（たとえば因果関係）にあることを意味するわけではありません。相関関係とは、「何らかの関係」というような、特定の関係を意味しない表現です。

・一般的に、相関係数が .30 とか .40 であれば「弱い相関」と表現されたり、表記されたりします。確かに値としては弱いのですが、扱うデータの特徴などの影響から各分野での「相場」のようなものもあり、強弱の表現にはブレがあります。文中の表現のみならず、得られた相関係数もしっかり確認する必要があります。

無相関検定 → 検定については p. 18 へ

サンプルにおける相関係数から、「母集団における相関係数は 0 か、そうでないか」を判断するための検定です。これは、母集団においては無相関（相関係数が 0）であることを仮定（帰無仮説）した上で、実際にサンプルで得られた相関係数が出現する確率を計算して判断します。

たとえば、「日常的回想」と「回想の使用」の間の相関係数は .35 です。もし、母集団において「日常的回想」と「回想の使用」の間の相関係数が 0 であったならば、今回認められた .35 という値が出現する確率は 1% (.01) 以下の、めったにないこととなります。そこで帰無仮説（母集団において無相関である）が間違っていると判断し、母集団において「日常的回想」と「回想の使用」の間の相関係数は 0 ではない、と結論するのです。無相関検定の結果、帰無仮説を棄却できる場合を「有意な関連がある」と表現します。

以上のように、「有意である」と判定されれば、「関連がある」と見なし、そのような記述をします。有意でない場合、母集団において相関があるとはいえないことになるので、大抵はそれ以上の言及を控えます。

なお「有意な関連」は、「強い（明確な）関連」とはまったく意味が異なる点に注意が必要です。無相関検定は、「母集団における相関係数は 0 か、そうでないか」を判断するだけで、強さは考慮に入れていません。関連の強さは、相関係数自体（0 に近いのか、±1 に近いのか）で判断します。

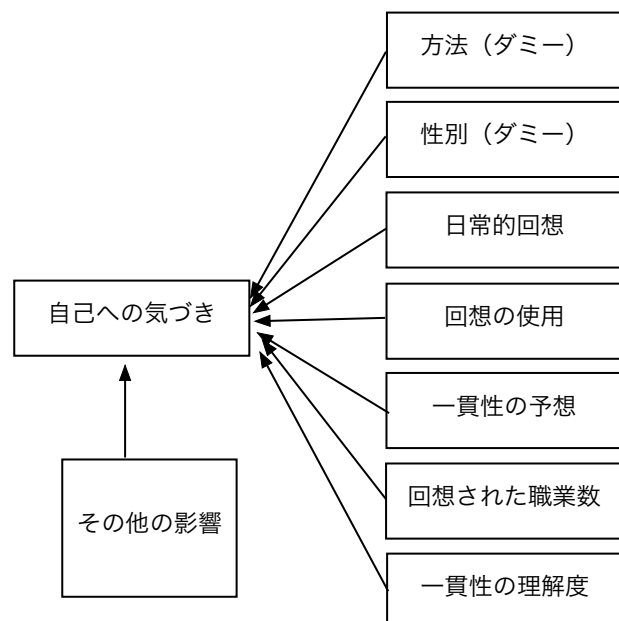
重回帰分析

ある変数を別の変数で予測する分析です。説明される変数を従属変数、説明する変数を独立変数と呼びます。ある従属変数に対して独立変数が2つ以上の場合は重回帰分析とよびますが、1つの独立変数で予測する単回帰分析もあります。

読み方としては、まず説明率 (R^2) に注目します。説明率は、従属変数の分散のうち、独立変数が説明している割合を表しており、一般的には重決定係数とよばれます。つまり、「自己への気づき」の分散は、「方法 (ダミー変数)」から「一貫性の理解度」の7つの独立変数によって、その約10% (.10) が説明されているということです。逆にいえば、残りの部分はその他の何らかの影響によるものといえます (下図を参照)。

この R^2 については、有意性の検定が行われます。それは、独立変数が説明している部分がある (その割合が0ではない) ことを確認するために行われるもので、帰無仮説は、母集団において「独立変数が説明している割合は0である (独立変数から従属変数へのパスすべてが0である)」です。「 $R^2 = .10, p < .05$ 」と記載されていますが、「帰無仮説の元で、今回 (このサンプル) のように R^2 が.10となる確率は.05 (5%) 以下である」ことを表現しています。5%以下というめったにないことが起きているので、前提となる帰無仮説を棄却し、「独立変数が説明している割合は0ではない」と判断するのです。なお、 R^2 は R を2乗したものですが、この R は重相関係数とよばれます。重回帰分析の結果 (重回帰式) によって予測される値と、実際の値の相関係数になります。

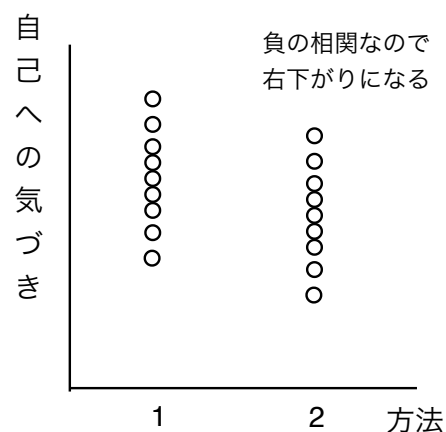
この R^2 が有意であれば、このモデル自体が意味を持つので、次の確認に入ります。それは各独立変数から従属変数へのパス (図の矢印) の大きさの確認です。この値は偏回帰係数とよばれますが、多くの場合、独立変数の数値の単位は異なっているため、比較するには単位をそろえる必要が出てきます。単位をそろえたものは標準偏回帰係数 (β)です。



読み方は、相関係数と同じです。-1から1までの間をとり、0は関連がないことを意味します。

この標準偏回帰係数についても、有意性の検定が行われます。有意でないものについては、そのパスはつながっていない（関連がない）と判断できます。この論文では、「自己への気づき」には、「方法」と「日常的回想」のみが、弱く関連していると考えられ、「職業目標の認識促進」には「方法」と「一貫性の理解度」が関連しているといえるのです。

なお、「方法（ダミー変数）」は、「自己への気づき」や「職業目標の認識促進」と負の相関係数が認められました。また重回帰分析でも負の標準偏回帰係数が認められています。表4の注記にあるように、方法Aに1、Bに2という数値が与えられていますので、方法Aの方が「自己への気づき」や「職業目標の認識促進」の点数が高いということになります（右図参照）。



重回帰分析を読む際の注目点・注意点

- ・重決定係数（説明率： R^2 ）をみて、現実をモデルがどの程度うまくすくい上げているかについてチェックするとよいでしょう。この値が小さいということは、現実とモデルの適合性がよくないともいえます。
- ・独立変数と従属変数間の相関係数と、標準偏回帰係数を比較してみることもよく行われます。「相関係数の検討から重回帰へ」という流れは一般的です（有意な相関が認められたから重回帰分析を行う）。今回の論文でもそうですが、概ね相関係数よりも標準偏回帰係数の方が小さく（0に近い）なりがちです。これは、それぞれの分析の特徴といえるかもしれません。相互にいくらかの関連をもつA、B、Cの3つの変数があったとき、AとBの相関係数には、AとBの直接的な関係と、Cを媒介した関係（A-C-Bという関係の強さ）が反映されます。他方、重回帰分析では「独立変数間は相互に関連を持たない」という仮定があるため、たとえばBを従属変数、AとCを独立変数とすると、AとBの直接的な関係のみが抽出されることとなります。相関係数と標準偏回帰係数を比較し、その差異からモデルについて検討してみるのもよいでしょう。

母集団とサンプル，そして検定

たとえば，キャリア発達に効果があるだろう新しい実習を開発したとしましょう。あるクラスでそれを実践し，効果を検討するために，実習の前後でキャリア発達の測定尺度を実施しました。実習前よりも実習後で，キャリア発達の測定尺度の得点（平均値）が高ければ効果があったといえることになります。

この時，実習前の平均値が 30.5，実習後が 34.0 だったとします。

さて，オリジナル実習の効果はあったといえるのでしょうか？

このような場合，結論について 2 つの考え方ができます。そのひとつは，「今回実習に参加した人たちだけにあてはまる結論」を導くことです。もうひとつは，「今回の実習に参加していない人にもあてはまる，より一般化された結論を導く」ことです。

後者のような結論を導く場合，この国や世界中に多くの対象者がいることが想定され，今回の実習に参加した人は，多くの対象者から選び出した一部の人たちという位置が与えられるでしょう。対象者として想定されるすべての人々を母集団，今回の協力者をサンプル（標本）とよびます。

今回実習に参加した人たちだけにあてはまる結論を導く場合だと，得られたデータは，いわゆる全数調査の結果になります。すなわち実習は，参加者の平均値を 30.5 から 34.0 へと上昇させたという結論になります。しかしこの知見は，「今回実習に参加した人たちだけにあてはまる結論」であることに注意しなければなりません。他の場合での効果の程は，統計的には不明なままです。

対して，今回の実習に参加していない人（母集団）にもあてはまる，より一般化された結論を導く場合，サンプルにおいて 30.5 から 34.0 へと平均値が上昇したことから，母集団で起き得る変化を適切に推測しなければなりません。こういう場合に用いられるのが推測統計とよばれるもので，いわゆる検定も含まれます。

今回の論文では使われていませんが，2 群の平均値に差があるのかどうかを検定することを例にして，簡単に検定の考え方を紹介します。

我が国において，A というテストの平均値に性差があるかどうかを検定することを考えます。すべての国民にテストができればよいのですが，それは無理な

ので一部に実施するしかありません。つまり、全国民が母集団、実施できた人たちがサンプルになります。

まず、母集団において差があるかどうかを判断するための道具を準備します。「差がない」ということが否定できれば、「差がある」といえるでしょう。これをまずは基本に据えます。

母集団において「差がない」ならば、サンプルにおいても「差がない」はずですが、抽出されたサンプルによって、多少の誤差は生じるはずですが。母集団において、男女ともに10点が平均になることがわかっているとしても、サンプルによっては、9.7になったり11.0になったりすることもあるでしょう。そのため、サンプルにおける男女の平均値の差を計算すると0にならない場合も多いと推測できます。しかし、確率的には0付近が多く、0から遠く離れるような差はめったに出ないだろうとも推測できるでしょう。数学的に、ある条件の元で、ある差がどの程度の確率で生じるのかを求めることができるのです。この、ある条件の元で、ある値（検定統計量）がどの程度の確率で生じるのかということがあらかじめ分かっている、というのが2つめの道具です。

さらに、ひとつのルールを採用します。出現する確率が5%より小さい場合、それをめったにないことと判断しようというルールです。

以上が判断をするための道具に関するものです。これらをどう使って差の有無を判断するかといえば、背理法のような考え方をを用います。

背理法

「Aである」と仮定 …論を進める…矛盾に突き当たる→ 「Aである」は間違い

統計的検定では、母集団において「差がない」と仮定して論を進めます。そして、その仮定ではおかしい状況（その仮定の上では、めったにない状況）が生じた場合、先に置いた「差がない」という仮定を間違いだと判断する、というやり方を用います。これに、「差がない」ということが否定できれば、「差がある」といえる、という考え方を加えるのです。

手順のイメージとしては、母集団において「差がない」と仮定したうえで…

- ① サンプルで得られたデータから男女の平均値の差を求める。(たとえば, 3.7点だった)
- ② これを, 「差がない」という条件の元で得られる差の出現確率と照合し, サンプルのような差が生じる確率を求める。(たとえば, .043 (4.3%) だということがわかった)
- ③ 「出現する確率が5%より小さい場合, それをめったにないことと判断しようというルール」を用い, サンプルにおける差は, 「差がない」という条件のもとでめったにない状況といえるかどうかを判断する。(.043 は, .05 (5%) より小さく, めったにない状況といえる)
- ④ 5%より小さい確率でしか起こり得ない状況が生じているならば, 仮定した「差がない」ということを否定し, 「差がある」という結論を導く。(めったにない状況が起きているので, 「差がない」とはいえない。つまり「差がある」と結論)

なお実際は, この例のようには進められません。あくまでも考え方の流れを示したものです。

論文等では, 頻繁に検定に関する専門用語が用いられますので, この例を用いてそれらの用語を紹介しておきます。

帰無仮説… 検定のために準備する仮説。例の場合「差がない」。

対立仮説… 帰無仮説が棄却された場合に採択される仮説。例では「差がある」。

検定統計量… サンプルから計算された統計量。例では, 平均値の差

(平均値の差の検定には t 統計量が用いられることが多い)

有意水準… ルール (共通見解) として定められている, 「めったにない」と判断する基準となる確率。分野によって違いがあるともいわれるが, 一般的に5% (もしくは1%) が用いられる。10%~5%の間の場合, 「有意傾向にある」と表現することもある。「危険率」ともいう。

p 値… サンプルから計算された検定統計量と同じ値か, それよりも極端な (めったにない) 値が観測される確率。この値が, 有意水準を下回るかどうかで判断を分ける。 $p < .05$ の場合「*」, $p < .01$ の場合「**」と記号で表記することもある。

3. 本論文の示唆とクリティカルな検討

近年は、「クリティカル・シンキング」といった言葉もよく使われるようになってきました。クリティカル (critical) とは、「批判的」と訳されますが、相手を攻撃するような意味だけではありません。学術的には、「論理的かつ内省的に、対象について詳細に」という意味合いで用いられます。

ではこの論文をクリティカルに検討することから、あなたほどのような示唆を受け取りますか？

学び方について

・教えてもらう

心理学系の学部を卒業した人ならば、ある程度は、基礎を身につけていると思います。もし周りにそのような人がいれば教えてもらう、学生時代の資料などを見せてもらう、というのが一番ハードルが低いのではないのでしょうか。また本気で学びたいなら、大学の科目等履修生の制度（大学によっては、科目履修生などともよんでいる）を利用することも一案です。心理測定や心理統計に関する科目を履修することができます。

なお、経済や経営学部、理系の学部でも統計は学びますが、因子分析などはおそらくわからない（学んでいない）でしょう。今回の論文など、キャリア関係の研究では心理学の測定法（心理測定とよばれる専門分野：psychometric）が利用されることが多くあります。この心理測定に関しては、他の学問分野における統計では扱われないことが多いので、統計を学んだ人なら大抵は知っている、という内容ではありません。

・ 系統だった統計の書籍を使う

webで「心理」「統計」といったキーワードで検索すると、非常に多くの書籍がヒットします。しかし、内容も説明の仕方も非常に多様です。入門的なものから、専門性の高いものまでありますし、数式が多いものも、できるだけ数式を使わずに説明するものもあります。そのため、実際に手に取って中身を見てから購入されるのがよいと思います。

それには、心理系の学部・学科・専攻を擁する大学の図書館に行くのが便利だと思います。大きな書店でもよいですが、超がつくような大型書店でない限り、心理測定、心理統計の本はそれほど置かれていないでしょう。

また、統計の本には、統計の解説書と、統計処理ソフト（SPSSやR）のマニュアル本が混在していますので注意してください。中には統計処理ソフトを使いながら統計を学ぶというコンセプトのものもあります。（Rは無料の統計処理ソフトです）