

「おまけの解説：因子分析」

■ 因子分析の目的

因子分析 (factor analysis) は、多数の意識変数がいくつかの根本的な概念によって形成されていると想定するような場合によく用いられる。統計的には、「複数の従属変数が、共通の独立変数による回帰式で説明できる」という枠組みでデータを整理している。ここでいう「複数の従属変数」は、実際にデータが得られている**顕在変数** (manifest variables; observed variables) であるのに対して、「共通の独立変数」はデータにはない想像上の**潜在変数** (latent variables; unobserved variables) である。

$$\hat{Y}_{(1)} = \alpha_{(1)} + \beta_{1(1)}X_1 + \beta_{2(1)}X_2 + \dots$$

$$\hat{Y}_{(2)} = \alpha_{(2)} + \beta_{1(2)}X_1 + \beta_{2(2)}X_2 + \dots$$

$$\hat{Y}_{(3)} = \alpha_{(3)} + \beta_{1(3)}X_1 + \beta_{2(3)}X_2 + \dots$$

例) ¹

$$\text{国語の得点} = 50.2 + 15.1 \times \text{才能} + 14.4 \times \text{やる気}$$

$$\text{数学の得点} = 10.9 + 22.4 \times \text{才能} + 3.7 \times \text{やる気}$$

$$\text{英語の得点} = 32.7 + 11.4 \times \text{才能} + 29.4 \times \text{やる気}$$

(網掛けが潜在変数)

すべての回帰式において同じ独立変数群 (X_1, X_2, \dots) が投入されていることがポイントであり、データ全体を少数の共通概念を説明しようとする姿勢を表している。複数の変数の分布を、どの変数にも共通する**共通因子[一般因子]** (common factor; general factor; 上式の X_1, X_2) と変数ごとの特殊因子 (specific factor; $\alpha_{(1)}, \alpha_{(2)}, \alpha_{(3)}$) に分けて説明するので、因子分析と呼ぶのだが、ほとんどの場合、共通因子に注目するので、共通因子だけを指して、「因子」と呼ぶことが多い。

実際的には次のような目的を満たすことができる。

- 1) 関連が強そうな複数の変数を整理するための妥当な共通因子を明らかにする (尺度の探索)
- 2) 1つの概念を確実に測るために設けた複数の変数が本当に1つの概念 (因子) で規定されているか確認する (尺度の検証)

それぞれ**探索的因子分析** (exploratory factor analysis)、**確証的因子分析** (confirmatory factor analysis) と呼ばれるが、基本は前者である。

すでに学習している共分散構造分析を用いれば、想定している因子をモデルの中に図示して、因子分析をおこなうことができる。さらに、そうして想定した因子 (潜在変数) と他の変数の関係も合わせて図式の中で考えることができる。しかし、そういった複雑なことを考えずに、単純に複数の項目の調査結果がどのような潜在変数に整理できそうか探索

¹ 才能とやる気は、科目ごとのものではなく、全科目で共通の因子 (数値が同じ)。

したい場合には、古典的な統計手法を用いた方が素直である。ここでは古典的な因子分析の手法について解説する。

■典型的な因子分析の対象

因子分析が適用できるデータは多様に考えられるのだが、典型的には多数の意識項目を少数の概念にまとめる場合によく用いられる。たとえば、就職活動のイメージについて尋ねた複数の意識項目の結果をまとめて、就活イメージを構成する基本的な要素（将来への希望、不安感、自立意識、……等）を析出したりする。

たとえば、表1は自分の生活の充足感に関係すると思われる10個の意識項目について、因子分析で要約しようとした結果である。一定の手続きに従うと3つの因子（生活前向き因子、現状否定因子、高望み因子）に要約することが適切と分析された。3つの因子は表に示されたような数値で10個の意識項目のいずれかとそれぞれ強く関係している（因子負荷量と呼ばれる）。±1に近いほど関係が強い。生活前向き因子等の名称は、それぞれの因子がどの意識項目と強く関係しているかという結果を参考にして、分析者が後付けで命名している。

表1 因子分析の結果を提示する典型的な表

表4 「生活に対する意識」項目の因子分析結果

	生活前向き	現状否定	高望み	共通性
私の人生は、きっと何とかやっていると	0.6739	-0.1700	0.0559	0.4862
いつも活躍していたい	0.6149	0.3355	0.0426	0.4925
毎日の生活の中に夢中になれるものがある	0.5610	-0.0667	0.1240	0.3346
私の周りの人たちはみんな信頼できる	0.5394	-0.1089	-0.0407	0.3045
私は人生の目標をうまく持つことができない	-0.2840	0.6035	-0.2218	0.4941
自分の適性にあった仕事をしたい	0.4575	0.4872	-0.0125	0.4469
だれも私の才能や能力をわかってくれない	0.0845	0.3800	0.2449	0.2115
人は私のことを裏切るかもしれない	-0.0902	0.3718	0.0120	0.1465
私は高望みするほうだ	0.1882	0.2664	0.7237	0.6301
私は高い理想を持っている	0.0729	0.1902	0.6782	0.5014
寄与率	24.5970	17.4580	15.0090	

因子抽出法：主因子法 プロマックス回転

出典：新井範子 2013「ソーシャルゲームにおけるユーザーの心理特性と課金行動の関連性について」 『上智経済論集』 58, pp. 277-287.

■因子分析の手順

因子分析では、通常、以下のような点を結果から読み取る。下記3つの読み取りは必須である。

- ・因子はいくつが最適なのか。 → **固有値**
- ・それぞれの因子は、それぞれの従属変数をどう説明できるのか。 → **因子負荷量**
（それぞれの因子はどんな概念を表していると解釈できるのか）
- ・それぞれの因子は、全体的にどの程度の説明力をもつのか。 → **寄与率**

また、以下の点も読み取る時がある。

- ・それぞれの従属変数は、この因子群でどの程度説明できるのか。 → **共通性**
- ・もし因子が測定できれば、各回答者はどのような値を持つのか → **因子得点**

考え方の過程を順番に説明しよう。

(1) 因子の数の決定

因子の数は分析者が指定することもできるが、ふつうは固有値が1以上の因子が何個導けるかで自動的に決定する。固有値は、その因子によって分析に加えた従属変数何個分を説明できるかを意味する。固有値が1より小さい（つまり、従属変数1つ分すら説明できない）因子はほとんど価値がないわけである。

(2) 初期解の算出

測定されていない因子の想定はさまざまに可能なので、同等に最適な解が無限に存在する。手続きとしては、まず、最適な解の1つをとりあえず出してみる。そのための統計的な方法は多様にあるが、現在では**最尤法** (maximum likelihood method) による推定がよいと考えられている。古典的には**主因子法** (principal factor method) と呼ばれる方法が有名であるが、これはコンピュータが未発達な時代の簡便法である。現在はメリットがないと考えられるが、論文にはよく出てくるし、実用上、最尤法との違いが出ることはほとんどないので、いずれでもかまわない。分析結果を示すときには、形式上、どの推定方法を使ったのかを記すことになっている（最尤法を用いた、等）。

無限にある解の中でも、最低限の決まりとして、因子は平均が0で分散が1の標準正規分布の変数であると考えられる。特に必然性がない場合、それがもっともわかりやすいからである。同じ理由で、従属変数も平均0、分散1に標準化して考える。つまり、因子分析で算出される回帰係数は-1~+1の間の値で関係の強さを表わす標準化回帰係数であり、結果を容易に読み取れる。これは**因子負荷量** (factor loading) と呼ばれる。

(3) 回転後の解の算出

最初に導かれる解は、ふつうそのままでは解釈しにくい（それぞれの因子がどのような概念を表すのかわかりにくい）。そこで無限にある同等に最適な解の中から解釈しやすいものを選び出し、解を置き換える。解釈しやすい解とは、なるべく単純なもの、つまりそれぞれの従属変数が1つの因子に規定されるような解である。この作業の図形的な意味は、解釈しやすいように各因子の影響力（因子負荷量）を示すものさしの軸を**回転** (rotation) させているものと考えられることができる。また、別の見方をすれば、それぞれの因子がそれなりに意味を持つように（寄与率が高くなるように）、定められた因子の個数の中で割り振りをやり直していることになる。

この回転のルールは大きく分けて2つある。1つは、因子同士の相関が0であるという制約の中で解釈しやすいように回転を行う場合である。もう1つはそのような制約は設けずに、因子同士が相関関係をもってもよいという前提で回転を行う場合である。これらは、2つの因子の軸が直角に交わるという前提で回転を行うのか、斜めに交わってもよいという前提で回転を行うのかという違いに相当するので、それぞれ**直交回転** (orthogonal

rotation)、**斜交回転** (oblique rotation) と呼ぶ。それぞれの中でまた具体的な方法が複数あるが、通常は、直交回転の場合には**バリマックス回転**と (varimax rotation) 呼ばれるものを用い、斜交回転の場合には**プロマックス回転** (promax rotation) と呼ばれるものを用いることが多い。斜交回転の方が制約が弱い分、当然、もっともらしい解を出しやすくなるが、通常の回帰分析ではそれぞれの独立変数の効果を他の独立変数とは関係ないものとして表している (つまり直交回転と同じ制約を課している) というのを考えれば、まずは直交回転が自然であり (互いに独自性の強い因子を析出している、という意義が認められる)、斜交回転の多用には一定の注意が必要である。

(4) 結果のとりまとめ

ともかく分析者が定めた回転方法で解を置き換えたものが、最終的な分析結果であり、この因子負荷量を中心に結果を表にまとめて読み取る。表は、ふつう第1因子と関係が強い (因子負荷量のサイズが大きい) 項目、第2因子と関係が強い項目、.....といった順に並び替える。また、因子には結果にふさわしい名称を分析者が付ける。その方がわかりやすいからであるが、命名によって誤った印象を与えてしまわないよに注意しなければならない。

また、因子負荷量の下には、ふつうそれぞれの因子の**寄与率** (contribution) を示す。これはその因子によって、元々の調査項目の回答の何%くらいが説明できることになるのかを示す。当然、寄与率が高い因子はそれだけ有意義な概念を表わしているということである。表1の場合、3つの因子の寄与率は24.6%、17.5%、15.0%なので、第1因子の説明力がもっとも高い。3つ合わせると約57%なので、たった3つの数値で10個の調査項目の57% (5.7個分) が説明できるのは有意義な要約といえる。

寄与率に比べると重要度は下がるが、それぞれの調査項目について**共通性 [共通度]** (communality) を横に示すこともある。これは析出した因子を使えば、それぞれの項目がどのくらい説明できるかを表わす。たとえば、表1の場合、「私は高望みするほうだ」の共通性は0.6301と高いので、これら3つの因子で十分に表現できているが、「人は私のことを裏切るかもしれない」の共通性は0.1465と低いので、これらの因子では表せない部分が大きい (共通性が低い) ということがわかる。

表1のように、分析に用いた方法の設定 (主因子法、プロマックス回転) は、約束事として示しておかなければならない。とくに回転の方法は分析結果の解釈も左右する。

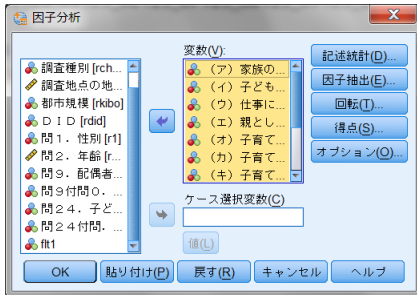
■ 特殊な回帰分析としての因子分析

因子分析では、いくつかの特殊な用語が使われるが、通常の回帰分析の枠組みで意味を理解すれば、難しいことはないので、対応を確認しておこう

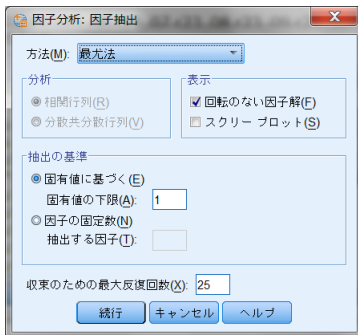
- 因子負荷量 → 回帰係数 (を標準化したもの)
- 共通性 → それぞれの調査項目を従属変数とする回帰分析の決定係数 (説明力)
- 寄与率 → 決定係数の要素を因子ごとに合計したもの (その因子での説明力合計)
- 固有値 → 寄与率 × 調査項目の個数 (その因子で調査項目何個分が説明できるか)
- 因子得点 → 回帰式の上で想定している因子の、各ケースでの推定得点

因子分析の操作

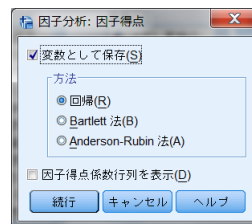
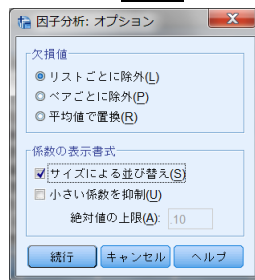
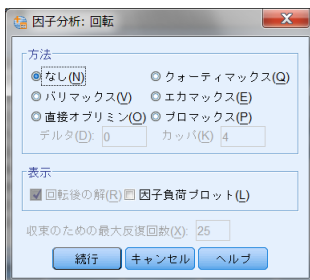
- ①メニューから、分析→次元分解→因子分析
- ②分析に用いる顕在変数をすべて「変数」に移動



- ③「因子抽出」ボタンで新しい窓を開いて、「方法」に「最尤法」を指定して「続行」ボタン
(③' 因子の数を自動ではなくて自分で決める場合には、「因子の固定数」を指定



- ④「回転」ボタンを押して軸の回転方法を指定して「続行」ボタン
直交回転にする場合（因子同士が無相関）→「バリマックス」
斜交回転にする場合（因子同士の相関を許容）→「プロマックス」
- ⑤「オプション」ボタンを押して「サイズによる並び替え」にチェックを入れて「続行」ボタン
(※分析結果が因子負荷量のサイズ順に並び替えられる)
(⑤' 調査項目が多い場合には、「小さい係数を抑制」にもチェックを入れた方が見やすい)
(⑥各ケースでの因子の得点を後の分析に使いたい場合、「得点」ボタンを押して「変数として保存」にチェックした上で、「続行」ボタン)



- ⑦元の窓で「OK」ボタン

因子分析の読み取り箇所

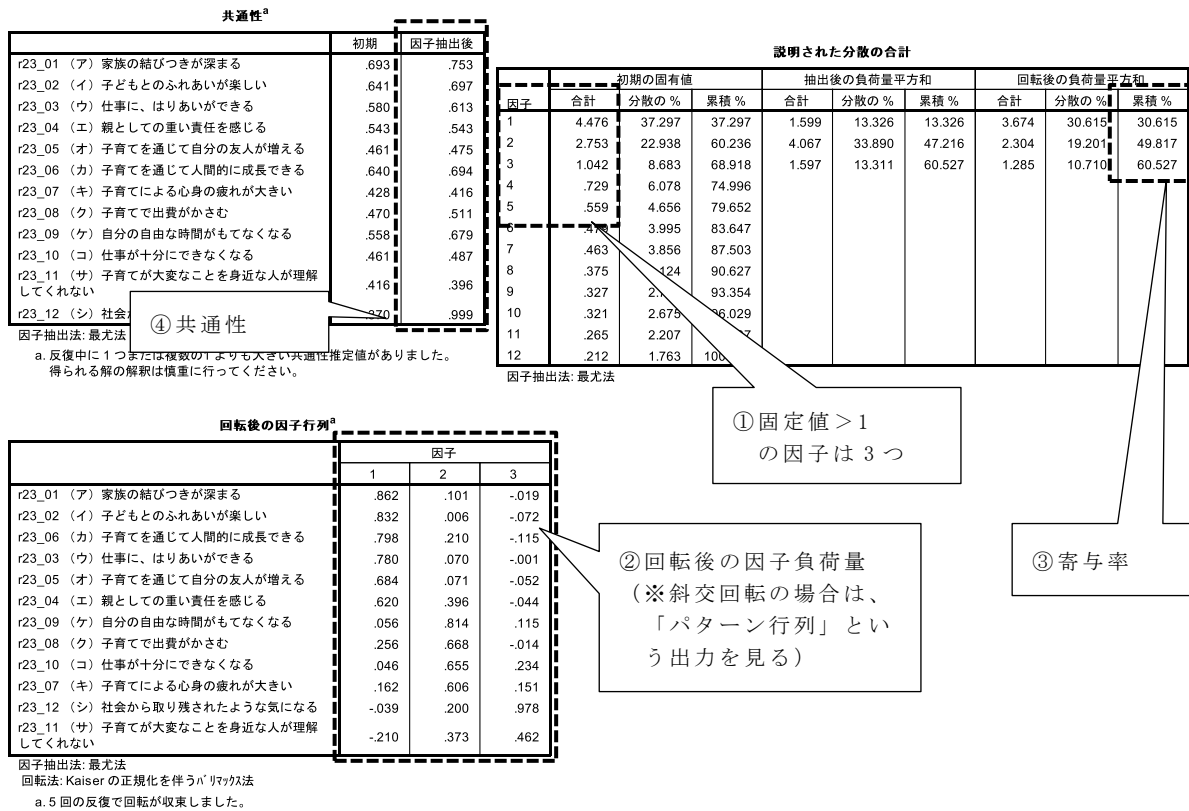


表 2 未婚女性の子育て意識の因子分析

	因子 1 肯定感	因子 2 負担感	因子 3 社会的受容	共通性
(ア) 家族の結びつきが深まる	.862	.101	-.019	.753
(イ) 子どもとのふれあいが楽しい	.832	.006	-.072	.697
(カ) 子育てを通じて人間的に成長できる	.798	.210	-.115	.694
(ウ) 仕事に、はりあいができる	.780	.070	-.001	.613
(オ) 子育てを通じて自分の友人が増える	.684	.071	-.052	.475
(エ) 親としての重い責任を感じる	.620	.396	-.044	.543
(ケ) 自分の自由な時間がもてなくなる	.056	.814	.115	.679
(ク) 子育てで出費がかさむ	.256	.668	-.014	.511
(コ) 仕事が十分にできなくなる	.046	.655	.234	.487
(キ) 子育てによる心身の疲れが大きい	.162	.606	.151	.416
(シ) 社会から取り残されたような気になる	-.039	.200	.978	.999
(サ) 子育てが大変なことを身近な人が理解してくれない	-.210	.373	.462	.396
寄与率	30.6%	19.2%	10.7%	

注: n=157。最尤法、バリマックス回転を使用。NFRJ03 若年票で未婚女性を分析。