

「因子分析 (1): 考え方」

■ 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とになる。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。

具体的には、下のよう<sup>3</sup>に 3 つの顕在変数  $Y_{(1)}, Y_{(2)}, Y_{(3)}$  を 2 つの潜在変数  $X_1, X_2$  で説明しようとしたりする。

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

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

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

例)

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

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

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

(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。上の例の場合、「才能」と「やる気」は科目ごとのものではなく、全科目で共通の勉強全体の才能とやる気である (数値が同じ)。

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。ただし、ほとんどの場合、共通因子に注目するので、共通因子のことだけを指して、「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。

多くの項目を調べたが、結局、これらの項目に共通の大切な概念群は何なのかよくわか

らない、という場面は多い。このようなときに、統計的に適切な整理の仕方を探索できる因子分析は魅力的である。とくに、行動に比べて意識・態度はその測定を直接的に行うことが難しいので、心理学的な研究では因子分析が多用される。因子分析には多くの応用方法があるが、ここでは古典的な探索的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 20 個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(たとえば、就業欲求、将来への不安感、自立意識の 3 つ) を析出したりする。

表 1 は、実際の論文での因子分析の例である。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を示している。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

	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思う	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、最低限、以下の 2 点について考える。

- ・ 因子はいくつが最適なのか。 → 固有値
- ・ それぞれの因子は、それぞれの調査項目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かと解釈できるのか)

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、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ことが多い。固有値とは、その因子によって分析に加えた顕在変数（調査項目）を何個ぶん説明できるかを意味する。固有値が1より小さいということは、つまり、その因子では調査項目1つぶんすら説明できていないということなので、データの要約として役立っていない。そのような因子には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、ふつう**最尤法** (maximum likelihood method) か**主因子法** (principal factor method) を用いる。現在の計算環境では最尤法の方が優れていると考えられているが、コンピュータが未発達な時代の簡便法として主因子法がよく広まっており、現在も論文にはよく出てくる。実用上、両者に大きな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。ただし、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「調査項目と共通因子の間の関係の強さ」は $-1 \sim +1$ の間の値で表されるので、相関係数のように結果を容易に読み取れる。これを**因子負荷量** (factor loading) と呼ぶ。

### (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、つまりそれぞれの調査項目が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。

この回転のルールには大きく分けて2通りの立場がある。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回転方法が複数あるが、通常は、直交回転の場合には**バリマックス回転** (varimax rotation) と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。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、独自性の強い因子群を析出する直交回転の方が基本的には意味があるので、まずはバリマックス回転を試すべきである。その上で、解釈が難しい場合には、プロマックス回転を試すとよい。

#### (4) 結果表のとりまとめ

理屈はやや難しいが、因子分析の中心的な結果は、「回転後の因子負荷量」という単純な表にから読み取れる（表 1 を参照）。結果表では調査項目を並び替えて、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、……といった順にしておく方が、それぞれの因子の意味が読み取りやすい。また、各項目がどの因子と一番強く結びついているか（因子負荷量のサイズが大きい）を見た目でわかるように、網掛けをしたり太い枠線で囲ったりすることが多い。

さらに、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うに注意しなければならない。

#### (5) 補足的な情報の追加

因子分析の結果表では、因子負荷量に加えて、各因子の「寄与率」と呼ばれる数値と、各調査項目の「共通性」と呼ばれる数値を表の周辺に記すことが多い（再び表 1 を参照）。

- ・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**
- ・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**

表の下部に記される各因子の**寄与率**（contribution）は、その因子によって、元々の調査項目の回答の個人差が何%くらい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 個の調査項目の回答個人差の 57%が説明できるのは有意義な要約といえる。

一方、表の右部に記す各調査項目についての**共通性**【**共通度**】（communality）は、析出した複数の因子を使ってその項目の回答個人差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多い（共通性が低い）ということがわかる。

また、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ので、表 1 のように、下に注釈として記しておくといよい。

### (練習)

結婚していない女性は子育てに対してどのようなイメージを持っているのだろうか。2004年1月に実施された第2回全国家族調査(NFRJ03)では、子どもを持つことや子育てについて12個の意識項目で意見を尋ねている(そう思う～そう思わない、の4点尺度)。

表2は、この調査の未婚女性のデータを用いて、12個の意識項目について因子分析を行った結果である。表の内容を読み取って、適切な語句や数値を記入・選択しなさい。

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 社会的疎外感	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
(ク) 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。

因子分析で12個の項目を少数の共通因子に要約すると、固有値が1以上になる有意義な因子が( )個析出された。因子の抽出方法としては( )法を用いて、軸の回転は( 直交・斜交 )回転のバリマックス回転を適用した。軸の回転によって、各因子はいずれかの調査項目と強く結びついて解釈しやすいものになる。

因子の内容は、各調査項目との間の( )の数値から読み取られる。1つ目の因子は、子育てへの肯定的な意識群と結びついているので、「肯定感因子」と名付けた。たとえば、「子どもとのふれあいが楽しい」との関係を表す数値は( )である。同様にして、第2因子は、子育てに対する負担を強調する否定的な意識なので、「負担感因子」と名付けた。最後の第3因子は、否定的な意識の中でも、社会からの疎外や無理解による苦しみを表す調査項目と関係しているため、「社会的疎外感因子」と名付けることにする。

寄与率に注目すると、12個の調査項目の個人差のうち第1因子で30.6%、第2因子で19.2%、第3因子で10.7%がそれぞれ説明できるので、3つの因子だけで12個の項目の個人差が( )%も説明できることになる。これは有意義である。ただし、3つの因子で説明できる程度は項目間で異なる。( )の数値が一番低い「サ」では、3つの因子で説明できるのは39.6%に留まる。この数値が極端に小さい場合は、他の調査項目とかなり意味合いが異なる内容なので、因子分析から外す方がよい。39.6%はそこまで低くない。

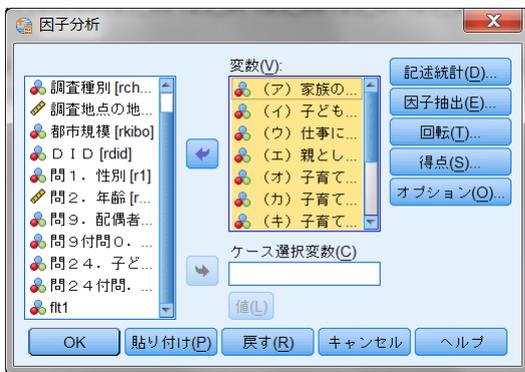
「因子分析 (2) : SPSS で実践」

■ SPSS でやってみよう

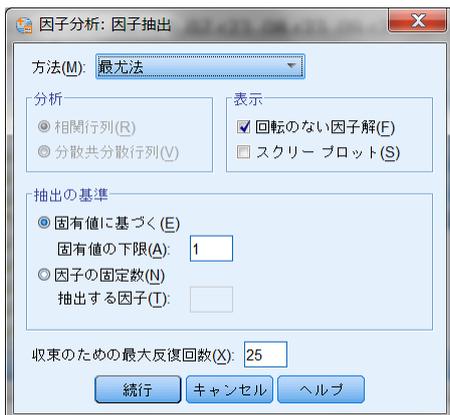
前回は、因子分析の考え方と典型的な結果表の示し方、読み取り方を解説した。SPSS でこの通りの結果を出すためには操作しないとイケないボックスがやや多い。難しくはないが、漏れがあるとまったく意図しない結果が出力されるので、注意しながら操作しよう。

因子分析の操作

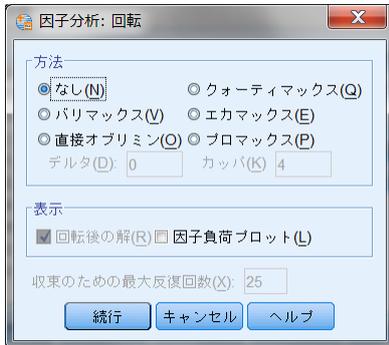
- ① メニューから、分析 → 次元分解 → 因子分析
- ② 分析に用いる調査項目 (顕在変数) をすべて [変数] に移動



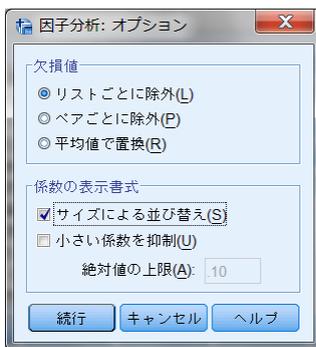
- ③ **因子抽出** ボタンで新しい窓を開いて、[方法] に「最尤法」を指定して **続行** ボタン (③' 因子の数を自動ではなくて自分で決める場合には、[因子の固定数] を指定



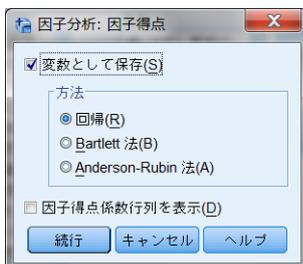
- ④ **回転** ボタンを押して軸の回転方法を指定して**続行** ボタン  
直交回転にする場合（因子同士が無相関）→「バリマックス」  
斜交回転にする場合（因子同士の相関を許容）→「プロマックス」



- 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

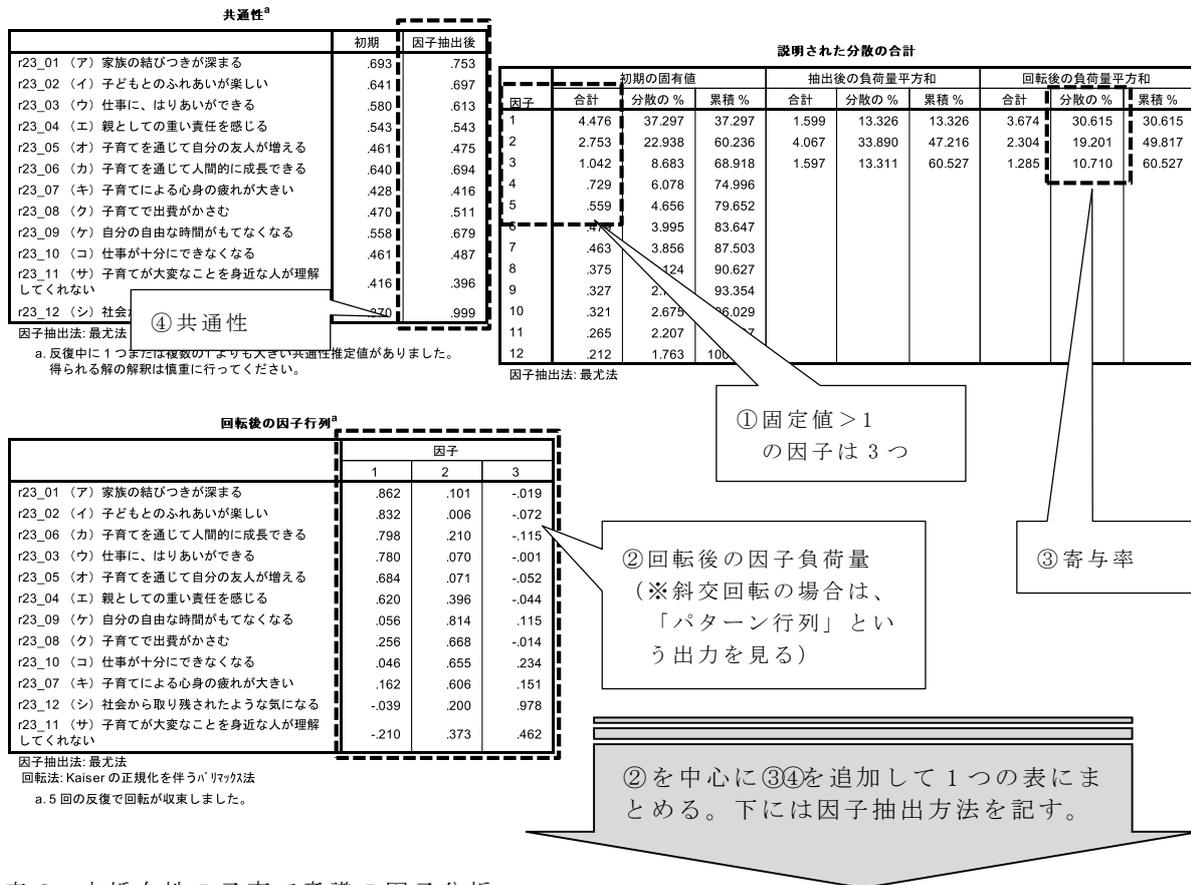


表 3 未婚女性の子育て意識の因子分析

	因子 1 肯定感	因子 2 負担感	因子 3 社会的疎外感	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。

「因子分析 (3) : 発展」

■ 因子得点の利用

因子分析で抽出される因子は、ただの抽象的な概念ではなく具体的な数値を持つ変数である。かりに抽出された因子 (潜在変数) を調査で測定できたとしたら、それぞれの回答者の各因子はどのような数値を持つはずだったのか、予測することができる。そのような数値を**因子得点** (factor score) と呼ぶ。手計算で因子得点を算出することはめんどろであるが、統計分析ソフトでは自動的にすべての因子得点を算出して、データに追加してくれるので手間はない。

因子得点は、通常の変数と同じようにそのまま他の分析に用いることができるので、分析を深めるために便利である。たとえば、家族規範意識についての 20 個の調査項目から性別役割意識と家父長制意識という 2 つの因子が抽出されたとする。これらの因子得点を用いて、結婚行動を説明する回帰分析を行ったり、因子得点の平均値を男女で比較したりすることができる。

■ 特殊な回帰分析としての因子分析

因子分析では、数々の特殊な用語法が使われるが、最初に記しているように因子分析は各調査項目を従属変数とする回帰分析の集まりにすぎない。通常回帰分析の枠組みで言えば、因子分析の特殊な言い回しは以下のようにとらえ直すことができる。回帰分析を深く理解していれば、因子分析の理解はより容易になる。

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

寄与率 → 決定係数を分解して因子ごとに合計し直したもの

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 (独立変数) の推定値

このことから、回帰分析の理解が他の多変量解析の理解にとって非常に重要なことがわかる。ほとんどの多変量解析は、回帰分析の考え方をベースにして、特殊な工夫を施したものになっている。回帰分析は実用上ももっとも多用されるので、ともかくしっかりと理解しなければならない。

\*\*\* 小課題 \*\*\*

Web ページ (<http://www2.itc.kansai-u.ac.jp/~tyasuda/>) から「小課題用データ」をダウンロードして因子分析をおこない、「小課題提出ファイル」のレポートを完成させなさい。(1月6日提出、1回遅れまでは減点で受け取ります)

\*\*\* 今後の予定 \*\*\*

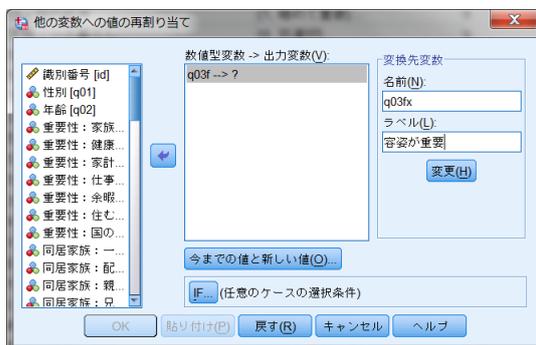
12/16: 小課題の説明 12/24(水): 質問受付 1/6, 13, 20: 最終課題取り組み 1/27: 提出

## ■（再掲）関連する SPSS の操作

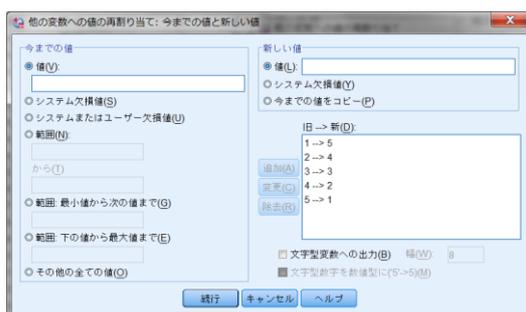
### リコーディング（値の再割り当て）

リコーディングは、既存の変数の数値（コード）を新しいルールで置き換えて、新しい変数を作る作業である。

- ①メニューから、変換 → 他の変数へ値の再割り当て
- ②リコーディングをしたい変数を左から選択して **→** ボタン  
（※複数の変数を同じルールでリコーディングする場合は、複数選択）



- ③リコーディング後の新変数の [名前] と [ラベル] を入力して、**変更** ボタン  
（※名前はアルファベットの形式変数名、ラベルは変数の内容がわかる日本語）
- ④**今までの値と新しい値** ボタンで変換ルールの窓を開く
- ⑤1つ1つの変換ルールについて [今までの値] と [新しい値] を入力して **追加** ボタン  
（※ルールの数だけ、これを繰り返す）



- ⑥すべてのルールがそろったら **続行** ボタン  
（※値を変更しない場合でも、ルールを入れないと空データになるので、必ず全部指定）
- ⑦元の窓で **OK** ボタン
- ⑧リコーディングで作成された 新しい変数 を使って、やりたかった分析を行なう