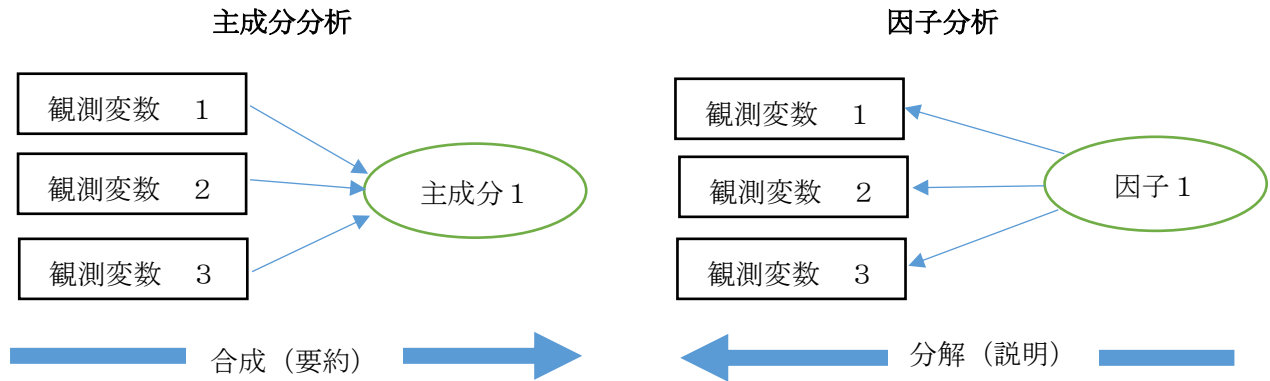


因子分析

1. 主成分分析と因子分析の基本的概念の違い

主成分分析と因子分析は多数の変数から少数の変数を得ることを目的とした、いわば標本が持つ情報を要約・説明するための探索型分析手段である。両分析は以下のようなモデルで示すことが出来る。

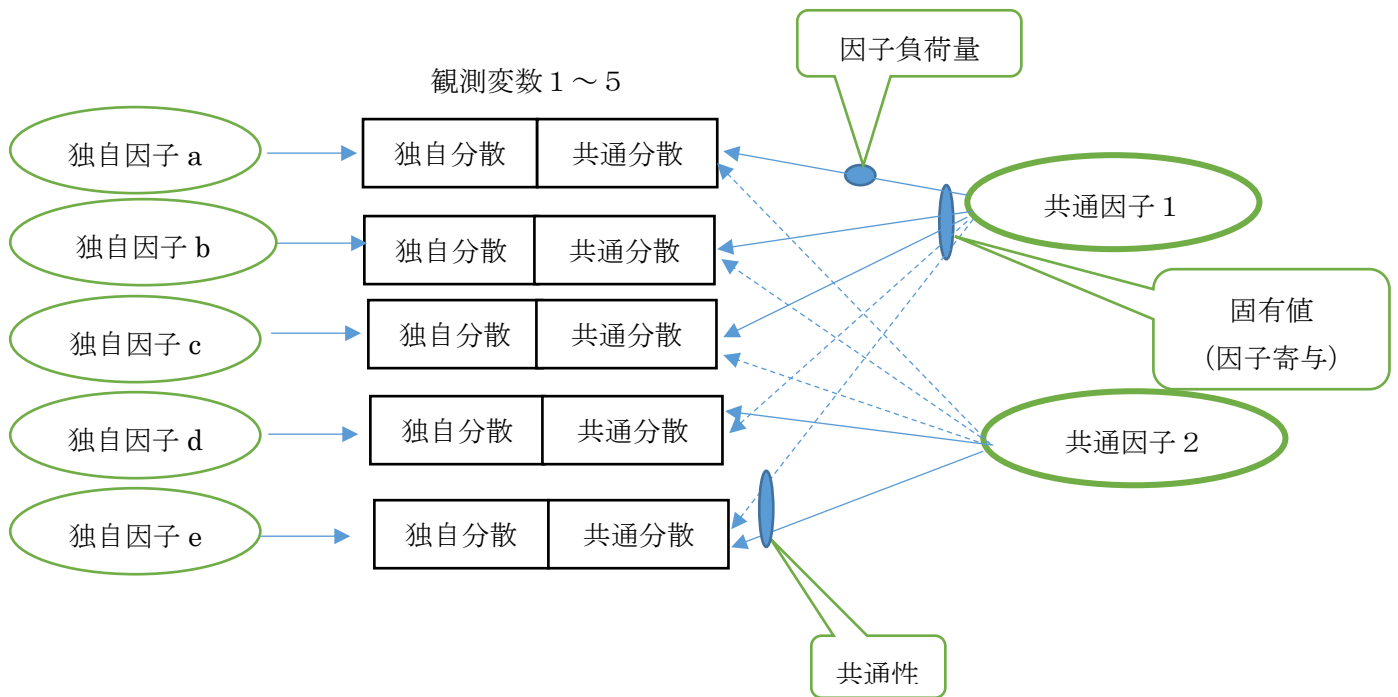


・複数の観測変数から主成分へ矢印が伸びている。観測変数を原因として扱い、主成分を合成し、要約的に記述する。多くの観測変数をまとめて数値化する社会科学や自然科学分野で利用されている。

・因子から観測変数へと矢印が伸びている。観測変数は結果であり、それらの背後にある共通性(因子)があると仮定し、それ観測変数間の相関行列から探し出し簡潔的に理解する手法。心理学、社会科学の分野で活用されることが多い。

2. 因子分析

- **ピアスマン**は1990年代初頭に、イギリスの中学生に対して語学、数学、音楽などの6科目の試験を行った。それらの得点に正の相関関係があることを発見した。(ある科目がよく出来る生徒は他の科目もよくでき、ある科目が苦手な生徒は他の科目も苦手) この事実から、ピアスマンは「知能」という全ての科目に共通する唯一の能力があると結論づけた。これが**共通因子 (common factor)** という考えかたの始まりである。
- 共通因子は観測変数の分散の**共通性 (communality)** を説明し、それらは**共通分散 (common variance)** と呼ぶ。**独自因子 (unique variance)** は観測変数に個別に影響し、測定された**独自分散 (unique variance)** から説明される。また、因子が観測変数にどのような影響を与えているかという**因子負荷量 (factor loading)** も考慮する必要がある。



補足説明

因子負荷量：因子と観測変数の関係性を示す。-1.00~+1.00 までの値を取り、.60 以上で高く強い関係性があると言える。.30 未満で低いとみなされる。

共通性：抽出された因子全体が個々の観測変数をどの程度説明しているか。観測変数を説明している各因子負担量の 2 乗和で算出される。共通性は 0~1 の値となり、1 から各観測変数の共通性の和を引くと独自性が分かる。

3. 探索的因子分析と検証的因子分析

因子分析には探索型因子分析と検証型因子分析に分けられる。

① 探索的因子分析 (exploratory factor analysis: EFA)

データ主導型分析。因子の意味、因子と観測変数の関係について先行する仮説や制約を分析内におかず、全ての観測変数の相関をもとに計算した結果から、因子を測定する。分析内の全ての観測変数と因子が関連するモデルを想定し、分析で得た固有値・因子負荷量・推定値を参考に分析者が因子数を決定し、それらの解釈を行う。探索的因子分析の目的

- (ア) 構成概念を探る
- (イ) 妥当性の高い質問紙を作成する
- (ウ) 変数をまとめる

② 検証的因子分析 (confirmatory factor analysis: CFA)

理論主導型分析。先行する理論にもとづき、因子数やその意味、因子と観測変数の関係を規定した仮説をモデルとして検証する。想定される因子は全ての観測変数に関連するとみなされず、因子と関連が強い観測変数に焦点をあてることになる。

4. 因子分析の前提

(1) サンプルサイズ

絶対的な基準はなく、データの質、観測変数や因子数、因子と観測変数の関連の強さ、共通性の大きさなどにより、適切なサンプルサイズは異なる。少なくとも測定する観測変数の数の10倍程度のサンプルを確保する。**KMOの標本妥当性** (Kaiser-Meyer-Olkin measure of sampling adequacy) でサンプルサイズの妥当性を検証すること。1に近いほど相関関係が適切に算出されたことを意味する。

基準値：.50未満（不十分） / .50~.70（中程度） / .70~.80（良い） / .80~.90（非常に良い） / .90（優秀）

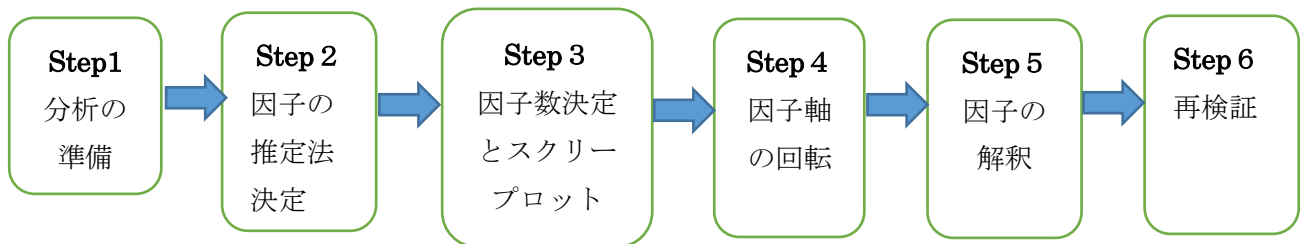
(2) データの種類 間隔尺度（5件法のデータも含む）・比率尺度

(3) 観測変数の数 3~4の観測変数が1つの因子に対し高い因子負荷量を示すことが目安となる。

(4) 観測変数間の相関

因子の推定に用いられる相関係数が.30以上を示すこと。.90以上のかなり高い場合は、多重共線性や単一性の恐れがある。

5. 因子分析の主な手順



5. 1 因子の推定法 (step 2)

- 因子の推定方法にはいくつかの種類がある。SPSSからは以下のような方法から選択することができる。仮説検証を行う前提として母集団を一般化しない場合は1・2・7を、母集団まで結果を拡張し一般化する場合は5・6が適している。
- 因子抽出の段階で共通性が1以上の値（独自性がゼロもしくは負）を示す不適解（Heywood case）が起こることがある。原因は①推定する因子負荷量の数に対してサンプルサイズが不十分な時、②データに適合する因子のモデルが定まらない場合

	抽出方法	特徴
1	主成分分析	複数の観測変数からできるだけ少数の合成変数を作成。
2	主因子法	第1因子から順に因子寄与が最大となるように因子を抽出。 不適解が出にくい。
3	重み付けのない最小2乗法	データと因子分析のモデルから算出される共通分散行列の間の差を最少にするように行う。
4	重み付けされた（一般化さ	最小2乗法に重み付けをし、尺度の単位に影響されないように行

	れた) 最小2乗法	う。適合度の検定が可能。
5	最尤法	因子分析に関わるパラメーターから尤度指標を算出し、これを最大になるように行う。適度の検定が可能。現在よく使用される傾向にある。
6	アルファ因子法	各因子のアルファ信頼係数が最大になるように因子を抽出。
7	イメージ因子法	ガットマンのイメージ理論に基づいた方法。

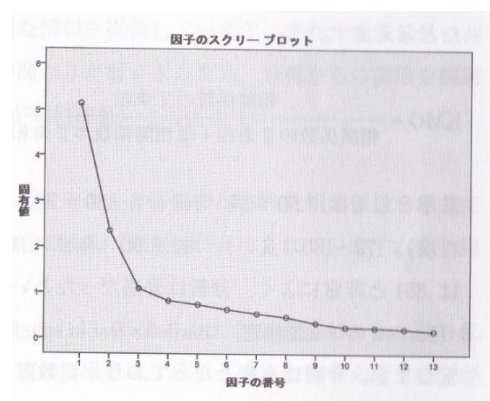
5. 2 因子数決定とスクリープロット (step 3)

①固有値 (eigenvalue) : 回転前の各因子の寄与の度合いを表す。この値の合計は項目数に一致し、高ければ因子寄与が大きい。

解釈 SPSSでは[固有値の下限]に[1]と入力されているため、結果は固有値が1以上の因子により解釈が進められる。次の図からは13個の因子が抽出されているが、固有値1以上に限ると3つに限定される。

因子	初期の固有値			抽出後の負荷量平方和		
	合計	分散の%	累積%	合計	分散の%	累積%
1	5.138	39.522	39.522	4.816	37.042	37.042
2	2.350	18.078	57.599	1.697	13.054	50.097
3	1.085	8.346	65.945	.789	6.072	56.169
4	.820	6.307	72.252			
5	.739	5.681	77.933			
6	.632	4.861	82.795			
7	.543	4.180	86.975			
8	.473	3.640	90.615			
9	.340	2.612	93.227			
10	.251	1.932	95.159			
11	.241	1.853	97.012			
12	.214	1.650	98.662			
13	.174	1.338	100.000			

因子抽出法: 最尤法



②スクリープロット (scree plot) : 固有値をプロットした図。視覚的に固有値の落差が最も大きいところまで因子と解釈できる。上記の右側の図。

解釈 固有値が大きいものから順にプロットされ、第3因子までで落差がとまっているため、3因子が適切であると判断できる。

③適合度 (goodness-of-fit) : 最尤法を用いるとデータとして抽出した因子によるモデルの適合度が算出される。値はカイ2乗値で表され、有意確率5%より大きければ「データとモデルに差がない」と解釈できる。

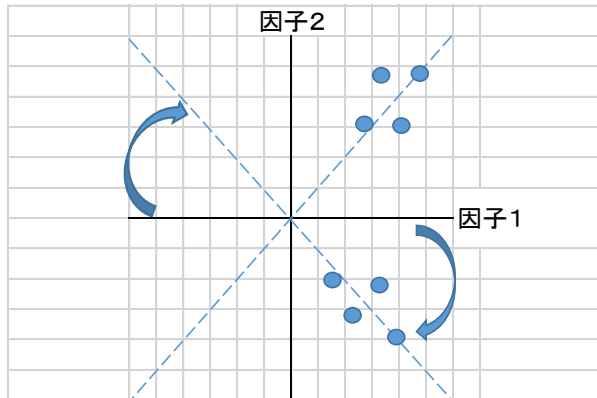
④因子構造 (factor structure) : 抽出された因子に強く負荷する観測変数のパターンから、理論的に妥当な結果であるか判断。

5. 3 因子軸の回転

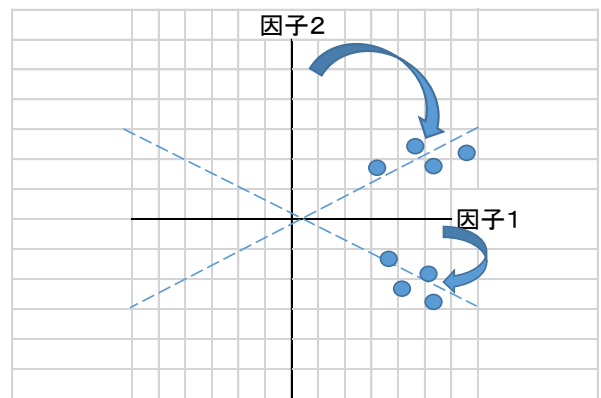
因子分析では、因子の解釈を容易にするために、因子の回転を行う。通常、因子分析を行うと因子が抽出されるが、因子と観測変数がうまく合致していない。そのため因子軸を引き、それらを観測変数が密集しているところまで回転させ、傾向が似たいくつかの観測変数の固まりを解釈できるようにする。回転には、**直交回転**（orthogonal rotation）と**斜交回転**（oblique rotation）がある。

回転種類	説明	代表的な回転手法
直交回転	因子間の相関はないと仮定し、複数の因子軸を直交させたまま回転させる方法	バリマックス（回転）、バイコーティマックス、コーティマックス、エカマックス
斜交回転	因子間に何らかの相関があるとして、複数の因子軸を別々に回転させる方法 因子間の相関がないと仮定するのは不自然な場合が多いため、斜交回転を使うことが多い。	プロマックス（回転）、コバリミン、バイコーティミン、コーティミン、直接オブリミン

直交回転のイメージ



斜交回転のイメージ



因子回転の相違点

	直交回転	斜交回転
回転	2本の因子軸を直交させたまま動かす。因子軸の角度を保つために、軸と観測変数の重なりにより制約が生まれる。	2本に因子軸を別々に動かすために、観測変数の散らばりに対し柔軟に軸を重ねることができ、因子の解釈が行いやすい。
因子負荷量と相関係数	観測変数同士の相関を想定しないため、因子負荷量と相関係数は同等のものとして扱う。	因子負荷量と相関係数は異なるために、別の表に出力される。
因子寄与 因子寄与率 累積因子寄与率の算出	因子寄与・因子寄与率・累積因子寄与率が、初期・因子抽出後、回転後に表示される。	どの因子に対しても高い因子負荷量を示すことが可能で、最大値が定まらず、因子寄与の計算が行われない。回転後の因子寄与の値は表示されず、全体的にどの程度の貢献度かについては相対的比較にとどまる。

因子の回転の目的は、**単純構造**（simple structure）を探し出すために行う。単純構造とは、観測変数

が強い関係性をもつ因子に対してはより高い因子負荷量を、関係性の低い因子にはより低い因子負荷量を示し、各因子がより明確に構成されている状態である。

5. 4 結果の解釈と再検証

バリマックス回転の解釈を行う際には、[回転後の因子行列]の因子負荷量を確認する。一般的には、.30~.40が基準値となる（斜交回転も同じ基準値）。プロマックス回転では[パターン行列]と[構造行列]が出力される。[パターン行列]は他の変数からの影響を除いた観測変数と因子の関係性を表し、因子負荷量にあたり、因子の解釈に用いることができる。因子分析では、納得のいく因子構造を得るために、結果を検討し適切でない観測変数を分析から除外し再度分析を行う。

6. 論文への記載

- 因子分析の結果を論文に記載する場合は次のような情報を含める。
- アンケートの項目
- 記述統計：各観測変数の平均値と標準偏差、相関行列
- 抽出法と回転法、因子数の絞り方の説明（スクリープロット図）
- 因子負荷量の最低水準および問題が見られた観測変数の対処法などの途中経過
- プロマックス回転の最終パターン行列と因子相関：全項目の因子負荷量、因子抽出後の共通性、因子ごとのアルファ係数、命名した因子名など

【参考文献】

平井明代 (2014). 「教育・心理系研究のためのデータ分析入門」東京書籍

栗原伸一 (2011). 「入門 統計学 ―検定から多変量分析・実践計画まで―」オーム社

山口和範・高橋淳一・竹内光悦 (2004). 「図解入門 よくわかる多変量解析の基本としくみ」秀和システム

中部大学人文学部心理学科 小塩研究室 心理データ解析 8 (1)

http://psy.isc.chubu.ac.jp/~oshiolab/teaching_folder/datakaiseki_folder/08_folder/da08_01.html