

SPSS による探索的因子分析について、出力内容の順に、何が出力されたのか、それはどのような意味なのかについて解説していく。

< 1 > 相関係数行列

相関行列

	Q1_1	Q2_4	Q4_8	Q7_16	Q8_17	Q9_20	Q13_25	Q14_27	Q15_28	Q16_29	Q17_31	Q19_34
相関 Q1_1	1.000	.560	.470	.498	.159	.170	.102	.172	.186	.144	.267	.228
Q2_4	.560	1.000	.444	.478	.213	.190	.219	.201	.181	.167	.255	.260
Q4_8	.470	.444	1.000	.439	.231	.390	.136	.216	.289	.210	.343	.226
Q7_16	.498	.478	.439	1.000	.266	.325	.200	.202	.182	.217	.303	.360
Q8_17	.159	.213	.231	.266	1.000	.427	.386	.208	.207	.180	.503	.267
Q9_20	.170	.190	.390	.325	.427	1.000	.371	.279	.292	.172	.409	.312
Q13_25	.102	.219	.136	.200	.386	.371	1.000	.332	.239	.166	.357	.256
Q14_27	.172	.201	.216	.202	.208	.279	.332	1.000	.332	.360	.261	.441
Q15_28	.186	.181	.289	.182	.207	.292	.239	.332	1.000	.403	.257	.324
Q16_29	.144	.167	.210	.217	.180	.172	.166	.360	.403	1.000	.244	.496
Q17_31	.267	.255	.343	.303	.503	.409	.357	.261	.257	.244	1.000	.327
Q19_34	.228	.260	.226	.360	.267	.312	.256	.441	.324	.496	.327	1.000

- 1) <1>の相関行列はピアソンの積率相関係数である。
- 2) 因子分析の計算は、<1>の相関行列の対角成分(相関が1になっているところ)に共通性(communality)の推定値(SPSSで使用しているのは重相関係数の2乗(squared multiple correlation coefficient: SMC))を代入して、行列の固有値(eigen value)と固有ベクトル(eigen vector)を計算している。
- 3) 固有値と固有ベクトルとは、相関行列を R、ベクトルを x、スカラー量(1つの値のこと)を λ としたとき $R * x = \lambda * x$ となるような λ と x を求める。それぞれを固有値と固有ベクトルという(※実は、後に示す因子寄与(factor contribution)が固有値、因子負荷量(factor loadings)の数値の集まりが固有ベクトルに相当する)。
- 4) 行列の対角線の成分の合計値をトレース(Tr)といい、固有値の値の合計はトレースを越えない。すなわち、左の行列では、トレースは12なので固有値の合計は12となる。(※もし、共通性の推定値を入れると、それは1よりも小さくなるので、固有値の合計は12よりも小さくなる)。

< 2 > 共通性の推定値

結果に記載

共通性

	初期	因子抽出後
Q1_1	.427	.621
Q2_4	.409	.505
Q4_8	.398	.419
Q7_16	.398	.460
Q8_17	.344	.486
Q9_20	.366	.425
Q13_25	.282	.318
Q14_27	.292	.342
Q15_28	.262	.293
Q16_29	.331	.512
Q17_31	.370	.458
Q19_34	.400	.503

因子抽出法: 最尤法

- 5) 2)の共通性とは、ある変数(項目)が、他の変数と共通な情報をどの程度持っているか(共通の因子によって説明できるか)を表す量である。因子分析では、 $\text{変数の情報} = \text{共通性の成分} + \text{独自性の成分}$ と仮定し、共通の部分の構造、すなわち因子構造を明らかにする。この場合、全体の情報を1という量で考え、共通性がそのうちの程度を示すのかを数値化している(※独自性の成分を0にして分析する場合を主成分分析(principal component analysis)という)。
- 6) 初期の共通性とは、計算のために最初に推定値(重相関係数の2乗など)を入れて固有値と固有ベクトルを計算するときの値、「因子抽出後」の共通性とは、計算を繰り返した結果、最終的に得られた共通性の値である。結果の表示の際にはこれを使う。
- 7) <2>の右欄をみると、Q1_1 や Q16_29 などの共通性が高く、Q15_28 がかなり低くなっている。
- 8) 共通性の最大値は1である。ということは、Q1_1の場合、 $1 - 0.621 = 0.379$ の分は、他の11の変数と共通でないような情報である。これを独自性(uniqueness)という。「共通性 + 独自性 = 1」。別の見方をすると、Q1_1 は、あとで示す3つの因子によって、Q1_1に関する情報の62.1%を説明(表現)しており、37.9%の情報は3因子による分析結果では表現されていないのである。

< 3 > 説明された分散の合計

因子	初期の固有値			抽出後の負荷量平方和			回転後の負荷量平方和		
	合計	分散の%	累積%	合計	分散の%	累積%	合計	分散の%	累積%
1	4.173	34.775	34.775	3.609	30.078	30.078	1.972	16.433	16.433
2	1.539	12.823	47.598	1.039	8.661	38.739	1.733	14.441	30.875
3	1.239	10.325	57.923	.694	5.783	44.522	1.638	13.647	44.522
4	.807	6.728	64.651						
5	.748	6.233	70.884						
6	.672	5.597	76.480						
7	.594	4.947	81.428						
8	.503	4.191	85.619						
9	.484	4.032	89.651						
10	.448	3.735	93.386						
11	.408	3.396	96.783						
12	.386	3.217	100.000						

因子抽出法: 最尤法

回転無し及び Promax 回転結果に記載

Varimax 回転の結果に記載

9) <3>の一番左の初期の固有値の合計欄は、相関行列の対角成分を1(トレースは12になる、独自性を考えないことに相当)にした場合の固有値である。したがって、合計欄にある12個の値を合計すると12になる。ちなみにこの固有値は主成分分析の結果になる(成分寄与)。

10) この欄をみて、因子数をいくつにするかを判断する。よく使われるのは合計(=固有値)欄が1以上という基準(カイザー・ガットマン基準)。<3>の例では、3因子かな、と判断できる。

11) 初期の固有値の合計欄の値が極端に減少するところまでを因子数にするという基準もある。これをスクリー基準という。

12) 9)で示したように、「初期の固有値」の合計欄を合計すると12になる。これはトレースの値、ここでは変数の数に相当する。隣の分散の%欄は12に対する各固有値の割合である(1因子目では4.173/12*100=34.775)。累積%欄は、その因子までの分散%の合計を表す(たとえば2因子目では34.775+12.823=47.598)。因子数を定めるもうひとつの基準として、この累積%が70~80%になる因子数、という方法もある。

13) この累積%の値は、たとえば、3因子まででデータ全体が持つ分散(情報の量と考えてよい)の57.923%を説明している(表現している、含めている、などでもよい)(ただし、この値はあくまでもトレースが12の場合である)。

14) そして、因子数決定として最も重要なのは、何因子になると考えていたか、因子に命名できるか、である。このデータは、3因子にした。3因子とした場合の固有値などの情報が、「抽出後の負荷量平方和」欄に示されている。

15) 「抽出後の負荷量平方和欄」のしくみも「初期の固有値」欄と同じ。この欄の合計を因子寄与(factor contribution)、分散の%を因子寄与率(proportion of factor contributionあるいは proportion of variance explained)、累積%を累積因子寄与率(cumulative ~)という。こ3因子に設定したので、新たに3因子として計算した結果を示している。

16) 「回転後の負荷量平方和」欄は、<5>の因子軸の回転後に同じ計算をした値である。合計欄と分散の%欄は抽出後の負荷量平方和欄とは異なる値になっているが、第3因子までの累積%は44.522%で同じになっていることに注目。因子軸の回転については<5>の説明参照。

< 4 > 因子パターン行列 (回転前)

因子行列^{a)}

	因子		
	1	2	3
Q1_1	.596	-.512	-.063
Q2_4	.596	-.387	-.027
Q4_8	.603	-.232	.046
Q7_16	.630	-.250	-.002
Q8_17	.518	.254	.391
Q9_20	.549	.216	.278
Q13_25	.440	.273	.221
Q14_27	.483	.271	-.188
Q15_28	.463	.219	-.174
Q16_29	.474	.315	-.434
Q17_31	.594	.180	.270
Q19_34	.592	.268	-.285

因子抽出法: 最尤法

a. 3個の因子が抽出されました。4回の反復が必要です。

17) <4>の値は、因子負荷量(factor loadings)あるいは因子パターン(factor pattern)と言う。パス係数の値でもある。因子負荷量は、各変数が因子とどのくらい関係が強いを表す値である。相関係数の値に近い(主成分分析では相関係数の値になる)。最大値は1なので、因子1と関係が強いのは、Q7_16, Q1_1, Q2_4...の順である。別の見方をすると、因子1はこれら3つの変数などが持つ共通の特徴で説明される。

18) なお、ここでの項目の並び順は、raw dataの項目順である。第一因子での因子負荷量が大い順にすることもできる。これはSPSSの「オプション」→「係数の表示形式」→「サイズによる並べ替え」で設定。

19) 因子負荷量をみると、全体に第一因子に多くの負荷量(関係の強さ)が集まってしまい、「第一因子はどのような特徴?」という問いに答えにくくなっている。このような場合に、因子軸の回転を行う。

20) 各因子の因子負荷量それぞれを2乗して合計すると<3>で示した因子寄与の値(抽出後の負荷量平方和の合計欄)になる。因子1では3.609になる。

21) 因子負荷量が-1になっている(例: -0.512)は反対の意味に解釈する。たとえば、第2因子が「読み手意識」と命名されたならマイナスの値は、「読み手を意識しない」となる。

22) 一番下のa.の部分のメッセージに注意。とくに、もし「抽出が終了しました」などのメッセージが出ていたら、計算の反復回数が少なく、十分な精度で計算されていないことを示す(すなわち、この計算結果は精度不足)。なお、反復とは、固有値計算のときに、行列計算を何回繰り返したかを示している。このような場合には、分析→データの分解→因子分析→因子抽出の「収束のための最大反復回数」欄に大きな数値を入力。

< 5 > 因子パターン行列
(直交回転 : Varimax 回転)

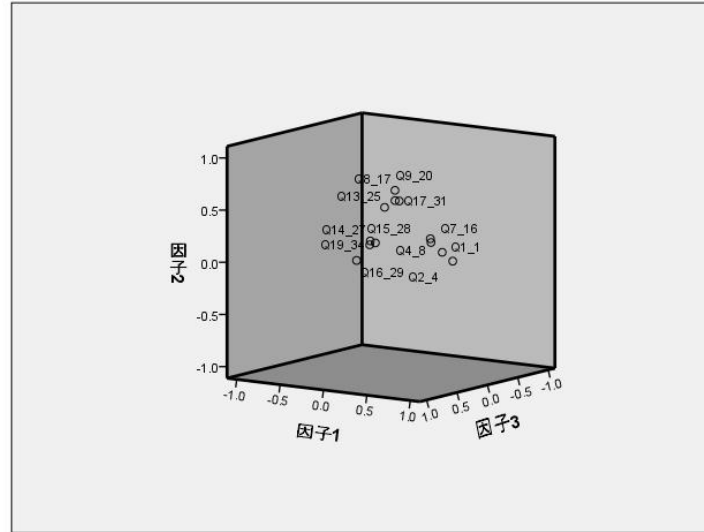
回転後の因子行列^{a)}

	因子		
	1	2	3
Q1_1	.780	.055	.098
Q2_4	.685	.136	.135
Q4_8	.570	.259	.166
Q7_16	.604	.232	.204
Q8_17	.134	.673	.126
Q9_20	.187	.591	.201
Q13_25	.076	.517	.213
Q14_27	.124	.244	.517
Q15_28	.149	.221	.471
Q16_29	.097	.081	.705
Q17_31	.244	.594	.213
Q19_34	.202	.232	.639

因子抽出法: 最尤法
回転法: Kaiser の正規化を伴うバリマックス法

a. 4 回の反復で回転が収束しました。

回転後の因子空間の因子プロット



23) <5>で、回転(rotation)とは、各因子の特徴を解釈しやすくするため、通常、大きい負荷量はより大きく、小さい負荷量はより小さくするように因子軸を回転することである。

24) 回転は直交回転と斜交回転とがある。直交回転とは、因子軸を直角のまま回転させた場合である。直角とは90度だから $\cos \theta$ は0、すなわち因子間に相関がないと仮定している。よく使われる方法としてバリマックス回転(Varimax 回転)がある。<5>は直交回転の結果を示した。

25) <5>は、18) で示した「サイズによる並べ替え」を実施している。すなわち、第1因子で負荷量大きい順、次は第2因子、のように、解釈しやすいようになっている。

26) 因子ごとに因子負荷量の2乗を合計すると<3>の「回転後の負荷量平方和」欄の「合計」欄の値になる。また、項目ごとに負荷量の2乗和を求めると共通性の値になり、<2>で示した値である。すなわち、回転しても共通性の値は変わらない(Q1_1では0.621)。

27) 因子間の相関を仮定する場合もある。その場合は斜交回転を用いる。よく使われるのはプロマックス回転(Promax 回転)である。

< 6 > 説明された分散の合計 (斜交回転 : Promax 回転)

説明された分散の合計

因子	初期の固有値			抽出後の負荷量平方和			回転後の負荷量平方和 ^{a)}
	合計	分散の%	累積%	合計	分散の%	累積%	合計
1	4.173	34.775	34.775	3.609	30.078	30.078	2.755
2	1.539	12.823	47.598	1.039	8.661	38.739	2.797
3	1.239	10.325	57.923	.694	5.783	44.522	2.606
4	.807	6.728	64.651				
5	.748	6.233	70.884				
6	.672	5.597	76.480				
7	.594	4.947	81.428				
8	.503	4.191	85.619				
9	.484	4.032	89.651				
10	.448	3.735	93.386				
11	.408	3.396	96.783				
12	.386	3.217	100.000				

因子抽出法: 最尤法

a. 因子が相関する場合は、負荷量平方和を加算しても総分散を得ることはできません。

28) <6>は、因子軸の回転の際に、斜交回転(ここでは Promax 回転)を行ったときの、説明された分散の出力結果である。

29) 「初期の固有値」「抽出後の負荷量平方和」の2つの欄は、<3>と全く同じ形式、値である。違うのは、「回転後の負荷量平方和」欄。

30) この欄の値は、表の下に、「a. 因子が相関する場合は、負荷量平方和を加算しても総分散を得ることはできません」とあるように、分散の%や累積%などの値は意味をもたない、単なる平方和の値にすぎない。

31) したがって、Promax 回転による因子分析結果を報告する際には、各因子の情報として、真ん中の「抽出後の負荷量平方和」欄(赤枠)の値を示すことが多い。

< 7 > 因子パターン行列(Promax 回転)

	因子		
	1	2	3
Q1_1	.862	-.134	-.043
Q2_4	.726	-.025	-.006
Q7_16	.596	.090	.061
Q4_8	.558	.142	.014
Q8_17	-.038	.759	-.084
Q9_20	.029	.624	.021
Q17_31	.092	.612	.024
Q13_25	-.083	.554	.079
Q16_29	-.042	-.136	.800
Q19_34	.054	.046	.655
Q14_27	-.016	.115	.520
Q15_28	.027	.095	.468

因子抽出法: 最尤法
 回転法: Kaiserの正規化を伴うプロマックス法

a. 4 回の反復で回転が収束しました。

< 8 > 因子構造行列(Promax 回転)

	因子		
	1	2	3
Q1_1	.776	.272	.273
Q2_4	.710	.334	.309
Q7_16	.669	.422	.383
Q4_8	.635	.428	.347
Q8_17	.303	.692	.329
Q17_31	.408	.671	.413
Q9_20	.351	.651	.388
Q13_25	.230	.557	.356
Q19_34	.374	.444	.706
Q16_29	.253	.296	.704
Q14_27	.277	.402	.578
Q15_28	.287	.373	.534

因子抽出法: 最尤法
 回転法: Kaiserの正規化を伴うプロマックス法

< 9 > Promax 回転後の因子相関行列

因子	1	2	3
1	1.000	.500	.454
2	.500	1.000	.567
3	.454	.567	1.000

因子抽出法: 最尤法
 回転法: Kaiserの正規化を伴うプロマックス法

32) < 7 > と < 8 > が、Promax 回転による因子分析の結果である。表内数値は因子負荷量を示す。解釈のしかたも同じ。ただし、斜交回転の場合、因子パターン行列と因子構造行列の2つが出力される。

33) 通常は、因子パターン行列の結果を論文に書く。両者の違いは、項目(変数)のベクトルから因子軸上の座標値、すなわち因子負荷量の値を求めるときに、平行線座標値によるのが因子パターン行列、垂線座標値によるのが因子構造行列である。直交回転では両者が一致しているので、因子パターン行列しか出力されない。

34) 因子パターン行列では、因子軸の相関の程度によっては、因子負荷量が 1.0 を越えることもあるが、間違いではない。

35) < 7 > の因子パターン行列を解釈すると、因子1が「全体構成」「読み手意識」「表現選択」と解釈した。

36) < 9 > の因子間相関に注意。これは因子軸の間の相関を表す。この結果では 0.5 以上の相関がある。かなり高いようである。したがって、Varimax 回転の結果によって解釈するのはかなり危険であることがわかる。

< 1 0 > 因子得点

	第1因子 因子得点	第2因子 因子得点	第3因子 因子得点
1	1.133	-1.388	-0.522
2	-0.626	0.065	-0.279
3	0.698	0.991	0.656
.	.	.	.
.	.	.	.
.	.	.	.
257	0.108	-3.483	0.313
258	0.837	0.813	-0.165
259	-0.580	-0.094	0.667
260	0.397	0.218	1.223

37) <10>で、因子得点は、260人の対象者が、3つの因子の要素をどの程度持っているのかを表している。
 38) 因子得点は、因子分析の結果を基に算出している。算出は、分析→データの分解→因子分析→得点→変数として保存→続行で計算。計算結果は、データエディタの raw data の右側に追加される。
 39) 因子得点はz得点なので、平均0、SDは1。間隔尺度なので、これを data として殆どの分析が可能である。
 40) <10>の結果より、たとえば第1因子を「全体構成」とすると、#1の人は、かなり全体構成を意識して作文を書いており、その得点はz得点で1.133、偏差値では61.33である。逆に#2の人は全体構成を考える程度が低い、解釈できる。

< 1 1 > KMO および Bartlett の検定

KMO および Bartlett の検定

Kaiser-Meyer-Olkinの標本妥当性の測度	.841
Bartlettの球面性検定	867.979
近似が2乗	
自由度	66
有意確率	.000

41) KMO は、変数群に適切な共通因子が存在するかどうかの指標である。これが0.5未満なら因子分析の適用は不適切とされている。< 1 1 >は0.841なので因子分析を行ってよいといえる。
 42) また、Bartlettの球面性検定は相関係数が単位行列をなす母集団から得られているかどうか帰無仮説であり、例のように棄却されれば問題ない。