関西大学心理学研究　2011 年　第 2 号　pp.1-6

項目と潜在変数との相関を使った項目分析
── 因子負荷量，因子構造そして因子パターンと関係の再考察 ──

清　水　和　秋　関西大学社会学部

Item Analysis Using the Correlations between Items and Latent Variable: Revisiting the Relationships among Factor Loading, Factor Structure, and Factor Pattern

Kazuaki SHIMIZU (Faculty of Sociology, Kansai University)

In the research filed of the item analysis, Richardson (1936) proposed the correlations between items and a test for the estimates of the factor loadings in a test. Guilford (1953) proposed the correction formula to get the correlation between the item and the sum of the remaining n − 1 items. Although the total score of such items is not sufficient for the criterion of the item selection, this rational of the item analysis has been used for the selection the items for scale construction while applying the factor analysis to the same items. In this note, it is suggested that latent variable like the factor score is more suitable than such sum score for the item analysis. Revisiting the definitions or descriptions of the factor loading, the factor structure and the factor pattern by Thurstone (1935), Thomson (1939), and Holzinger & Harman (1941), the estimation method of the reliability in the context of the factor analysis is also discussed.

Key words: factor analysis, scale development, item selection, construct

Kansai University Psychological Research
2011, No.2, pp.1–6

心理尺度の信頼性の推定値に関して，因子分析法の文脈においては，McDonald（1999）のωと Cattell & Tsujioka（1964）あるいは辻岡（1964）の構成尺度の共通性とが一致することを清水（2010）で議論した。この小論ではこれに引き続き，項目分析の手法としていまだに使われることがある項目（item）と総点（total）との相関係数に検討を加えてみたい。

よく知られていないようであるが，これは因子負荷量を推定する別法として Richardson（1936）がテスト項目とテストの全体の相関を計算するという提案からはじまったものである。この手法が採用した総点の意味を問い直すなかで，項目分析の方法として項目と潜在変数との相関である因子構造（あるいは因子負荷量）とこれによる信頼性の推定について，考察を加えてみたい。

IT 相関の計算

心理学での尺度構成では，研究対象である構成概念を測定するのに適切な項目を収集・作成し，適切な標本で収集したデータに因子分析を適用し，その結果から尺度を構成し，構成した尺度の信頼性を報告する手順が確立されている。わが国の学術雑誌では，因子解の数値の横に，IT（item-total）相関の数値も掲載し，これらの数値がよく似た傾向を示した

ので項目分析が適切であったと報告していることがある。これに加えて，情報量を捨てることにもなるGP分析による項目分析の結果を報告する例もある。

心理測定法では技術の進歩と共に概念を探求する方法や道具をより精緻なものとする努力が続けられてきた。例えば，因子分析での初期の解の推定に使われてきたセントロイド法は，コンピュータソフトの普及とともに姿を消し，主因子法や最尤法が使われるようになってきた。SASやSPSSそしてAmosのようなPCの統計解析ソフトが応用研究の場に普及し，最尤法での解の推定が簡単にできるようになったことだけが理由ではなく，数理統計学的な知識とこれがもたらす統計量の魅力が，このような動きの根底にあると考えている。

項目分析の現場は，少々事情が異なるようである。項目分析の手法としてのIT相関は，セントロイド法の手計算時代のままである。以下では，IT相関のはじまりの頃に遡り，何を求めて，この方法が提案されたのか，そして，因子分析との関係について再考してみることにする。

議論の手始めとして，心理測定法の代表的なテキストであった『精神測定法』を紐解いてみることにする。この本の中でGuilford（1959）は，「15.2.4.5 項目の因子分析」と項のタイトルを付けた箇所で「項目の心理学的等質性に関して疑問がある場合，あるいは項目・全体相関が低い傾向を示す場合には，テストをさらに，比較的等質性の大きいサブテストに分けるという手順を踏むことが必要となる。……途中略……この場合に，もっともよい方法は，通常あまり行われてはいないけれども，項目の因子分析を行うことである。（一部略，pp.545-546）」とし，Wherry, Campbell, & Perloff（1951）による継続項目分析法（method of successive item analysis）を「因子分析以外の方法で達成することを目指して，考え出された別法」として紹介している。項目（item）と項目分析の対象となる項目の総得点（total）との相関からより等質な新しい尺度を構成しようという提案である。総得点の中に項目分析の対象となる項目の得点が含まれるために相関がつり上げられることを指摘して，この見せかけの値を修正する式をGuilford（1953）が提案している。

この本の第15章「テストの作成」では，項目分析の項目間の相関行列からの因子分析のことは何も出てこない。因子分析法を詳細に説明した第16章でも，項目分析への言及はない。

計算が手で行われていた時代であった。do-littleの精神の下，より少ない計算量でベターな統計量を得ようとする工夫が行われた時代でもあった。項目分析の作業に因子分析が望ましいと分かっていても，経費のかからない方法で代替可能な方法に頼らざるをえなかった時代でもあった。

因子分析と項目分析とが交錯しているようでありながら，『精神測定法』では，このように独立して解説が行われている。この関係をさらに検討するために，次に，IT相関の考え方を詳細に説明してみることにする。

IT相関を式で表せば，次のように書くことができる。ここでは，対象とする構成概念をxとし，このxを測定する項目をn個作成したとする。そして，N人を対象に調査が行われたとする。個人iの項目全体の得点は，次のように表すことができる。

(1) $$x_{ti} = x_{1i} + x_{2i} + \cdots + x_{ji} + \cdots + x_{ni}$$

IT相関の素朴な計算は，x_tと各項目との相関係数（r_{jt}）を計算することである。そして，得られたn個の値から項目と尺度との関連性を評価しようとする方法であった。

この方法では，項目の分析の対象となる項目が，尺度に含まれている。このため，対象の項目を含む総点と項目分析の対象となる項目との相関は過大となることは明らかである。そこで，IT相関では，項目分析の対象の当該項目とこれを除いた残りの項目（remaining items）の総点との相関を計算することが必要となる。ここでj番目の項目を除いた残りの得点を，jを1からnまで変えながらn項目について，式で表してみることにする。

(2)
$$\begin{aligned}
x_{r(1)i} &= \phantom{x_{1i} +{}} x_{2i} + \cdots + x_{ji} + \cdots + x_{ni} \\
x_{r(2)i} &= x_{1i} + \phantom{x_{2i}} \cdots + x_{ji} + \cdots + x_{ni} \\
&\vdots \\
x_{r(j)i} &= x_{1i} + x_{2i} + \cdots + \phantom{x_{ji}} \cdots + x_{ni} \\
&\vdots \\
x_{r(n)i} &= x_{1i} + x_{2i} + \cdots + x_{ji} + \cdots
\end{aligned}$$

n個の項目を対象とする項目分析では，この(2)式の尺度得点をn個分計算し，これらとn個の項目の項目との相関（この場合では，x_jと$x_{r(j)}$のような組み

合わせでの相関）を計算する必要がある。尺度得点を n 個も別に計算することは，手計算の時代には計算ミスを起こす可能性を高めることにもなりかねない。Guilford（1953）の提案を修正と呼ぶのは正確ではない。$x_{r(1)}, x_{r(2)}, \ldots, x_{r(j)}, \ldots, x_{r(n)}$ の n 個得点を(2)式のように計算しなくても，評価対象の項目 j と尺度得点 x_t から対象項目の得点を除いた残りの総点との相関（$r_{jr(j)}$）を計算することができることを次の式で示しただけである。

$$r_{jr(j)} = \frac{\sum\left(X_j - \overline{X}_j\right)\left(X_{r(j)} - \overline{X}_{r(j)}\right)}{N \cdot S_j S_{r(j)}} = \frac{S_t r_{jt} - S_j}{\sqrt{S_j^2 + S_t^2 - 2S_j S_t r_{jt}}} \tag{3}$$

ここで，S_t は n 個の項目の総点の標準偏差であり，S_j は項目 j の標準偏差である。

心理測定法の黎明期に Psychometrika に掲載された論文には，Doolittle 法がよく出てくる。回帰方程式の展開や計算に，M. H. Doolittle によって考案された線形連立方程式の解法が使われたからである（Holzinger & Harman, 1941）。do-little には，怠け者という意味もあるが，より少ない計算労力で，より適切な結果を得たいという情熱が，この時代の論文には感じられる（例えば，Griffin, 1936 など）。そして，手計算でしかできないという制約の下で，理論の追求と同時にそれに近似する結果を手に入れることに努力した時代でもあった。このように考えて，この do-little をこの小論では使っている。

IT 相関での総点

IT 相関のオリジナルな提案は，あまり知られていないようであるが，*Psychometrika* の第 1 巻に掲載された Richardson（1936）による。少し長くなるが引用してみる。“The first factor loading on the centroid is a measure of the correlation between a test and the sum or average of the tests in the battery. A similar interpretation may be made in the item analysis situation. The item-test coefficient measures the correlation between a variable（the item）and the sum or average of many such variables. In this context, the item-test coefficient is the “factor” loading of the item with an arbitrary test variable which is the sum of the items.（p.71）”

この時代の初期因子解の計算方法は，セントロイド法であった。Holzinger & Harman（1941）によると，相関行列から因子解を計算するセントロイド法は C. Burt が 1917 年に C. Spearman の一般因子を得るための方法として最初に使った方法であり，1931 年に L. L. Thurstone によって，多因子の因子解の抽出の方法として完成をみた（Thomson, 1939）。この手計算の時代にあっても Thurstone（1935）は，セントロイド法と主軸法（主因子法）との 2 つの方法の計算手順を詳細に説明している。計算量が比較的少なくて済むセントロイド法は，その後，先にも紹介したように Guilford（1959）による『精神測定法』でも解説があり，コンピュータで因子分析が行われるようになる 1960 年代までは，do-little の精神の下で使われてきた。なお，セントロイド法については芝（1972）が主因法の近似を与えることを負荷量に 1 の符号を与えて計算がはじまるという点を明らかにしながら，Fortran プログラムともに詳細に解説している。

上で引用した Richardson（1936）は，分析対象のテストバッテリに含まれるテストとバッテリ全体の総点との相関が，セントロイドの第 1 因子の負荷量となることを指摘し，これを項目分析へと応用することを提案している。そして，項目−総点の相関係数の値を因子負荷量の推定値と仮定したわけである。

Mosire（1936）は，この論文を取り上げながら，何らかの外的な基準とこれによる項目分析という大きな枠組みからの議論を行っている。項目分析として項目と尺度との相関がセントロイドの因子負荷量と一致するであろうとして，Richardson（1936）の項目全体の総点も 1 つの基準であるとしながら，項目の全体の意味することへの注意を喚起している。項目分析では，構成概念を測定するものとして収集・作成された項目の全体を項目分析の出発尺度と呼ぶこともある。この出発尺度に 2 つあるいはそれ以上のものが含まれている場合には，項目分析の実際に問題が発生するという議論である。

項目分析の準拠枠

収集・作成した項目の全体を基準とする IT 相関による項目方法は，Mosire（1936）が警告するように，出発点となる項目の全体に含まれる質によって，

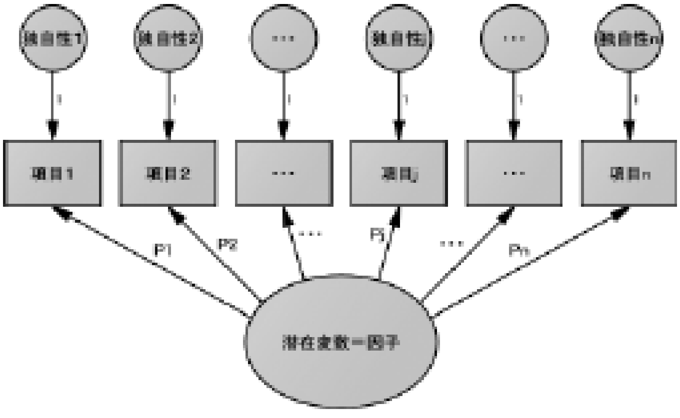

図１　１次元の因子分析モデル
注：P1 ～ Pn は因子パターンである。Amos などの解の推定では，識別性を確保するために，P1 ～ Pn のいずれかの値を１で固定する必要がある。

その正否が決まることになる。対象となる項目間の相関行列（$n \times n$）から，内部の等質性を確認する方法としては因子分析に期待はあったが，n の数が大きくなると，実際の計算としては，心理測定法の黎明期にあっては，その実行がほとんど不可能であった。

IT 相関では，項目分析で計算する相関の数は（$n \times 1$）個と，（$n \times n$）という大きな相関行列と比べると計算量は確かに do-little ではある。その一方で，IT 相関が失ったものの１つが，n 個の項目間の関連性である。内的整合性を（$n \times 1$）個の相関係数から想像するというやり方を取らざるを得なかったわけである。

Guilford（1953）の(3)式による修正の提案は，計算量という点では do-little な結果を与えてくれた。その一方で，総点が対象とする項目に対応して異なるという状況をもたらした。総点に当該項目が含まれることによるこの項目と総点との相関の値のつり上がり部分を修正しようとすると，総点そのものが(2)式の $x_{r(1)}, x_{r(2)}, ..., x_{r(j)}, ..., x_{r(n)}$ のように対象となる項目によって異なることになるわけである。この結果，Richardson（1936）が準拠枠とした n 個の全体の総点とは異なった n-1 の数の総点が項目分析の対象となり，分析の全体にわたって一貫した準拠枠を失うことになる。

因子分析が身近になった現代では，（$n \times n$）の項目間相関を因子分析することは，ある程度以上の数の対象者数を集めれば，簡単に実行することができる。因子分析によって，内部構造が１次元かどうかを確認することができると同時に，因子と変数との関係である因子負荷量を手にすることができる。

現代の因子分析は，潜在変数である因子が原因となって，現象としての観測変数に影響を与えていると考えることができる。構造方程式モデリングの測定モデルとして，１次元構造となる n 個の項目を図１のようにパス図で描いてみることにする。この図の「潜在変数＝因子」が因子分析による項目分析での準拠枠となる。そして，項目の選択は，因子負荷量である P1 ～ Pn の標準化した値から行うことになる。

項目分析を行うような探索的な因子分析の場合には，構造方程式モデリングを使った解析で期待されるような高い適合度を得ることは容易ではないと考えるべきであろう（例えば，清水・山本，2007 など)。一般的には，項目分析を目的とした場合には，SPSS や SAS のような統計ソフトで探索的因子分析を実行し，１因子の因子解を推定することになる。

因子構造：因子負荷量と因子パターン

ここでは，Holzinger & Harman（1941）に従って，探索的な因子分析モデルを対象として議論を進めることにする。なお，彼らが体系化した因子分析は，構造方程式モデリングの変量モデルではなく記述モデルである（例えば，柳井・繁桝・前川・市川，1990）。

ここで，ある標本の N 人の対象者について、n 個の項目を測定したものとする。そして、ここでは一般的に議論を進めるために，m 個の次元として因子の数を特定することがきたものとする。

観測変数の標準得点行列 $\boldsymbol{Z}$ ($N \times n$) は、次のような因子分析モデルとして表すことができる。

$$(4)\quad \boldsymbol{Z} = \boldsymbol{F}\boldsymbol{V}'_{fp} + \boldsymbol{U}\boldsymbol{D}$$

ここで、$\boldsymbol{F}$ は ($N \times m$) 次の因子得点行列であり、$\boldsymbol{V}_{fp}$ は ($n \times m$) 次の因子パターン行列である。この共通因子空間とは独立した独自性に関するものが、($N \times n$) 次の独自性得点行列 $\boldsymbol{U}$ と独自性を対角項にもつ ($n \times n$) 次の対角行列 $\boldsymbol{D}$ である。

伝統的な因子分析モデルでの(4)式は、$\boldsymbol{R}$ を (n×n) 次の変数間の相関行列とすると、次のように展開することができる。

$$(5)\quad \begin{aligned} \boldsymbol{R} &= \frac{1}{N}\boldsymbol{Z}'\boldsymbol{Z} \\ &= \boldsymbol{V}_{fp}\boldsymbol{C}_f\boldsymbol{V}'_{fp} + \boldsymbol{D}^2 \end{aligned}$$

なお、$\boldsymbol{C}_f$ は (m×m) 次の因子間相関行列である。

因子パターンは，因子分析のはじまりの頃は Thurstone (1935) が使った「因子負荷量」と呼ばれていた。Thomson (1939) は, “(Very often in multiple-factor analysis the “saturation” of a test with a factor is called the “loading,” and this is a convenient place to introduce the new term.) (p.25)” という興味深い記述を残している。一般因子を提唱しながら因子分析をはじめた C. Spearman は，この値を全体に対する飽和の状態として表現していたようである。なお，saturation については，Spearman (1927) が理論と式を紹介している。Thomson (1939) は，これに従って相関行列から saturation の計算手順を詳細に解説している。

この因子負荷量を現在の用語で整理したのが，Holzinger & Harman (1941) である。彼らは，因子負荷量を因子パターンとして定義し，観測変数と因子との相関を因子構造として定義した (p.16)。上の因子分析の定義に合わせて，因子構造行列を $\boldsymbol{V}_{fs}$ ($n \times m$) 次とすると，次のように計算することができる。

$$(6)\quad \boldsymbol{V}_{fs} = \frac{1}{N}\boldsymbol{Z}\boldsymbol{F}'$$

あるいは

$$(7)\quad \boldsymbol{V}_{fs} = \boldsymbol{V}_{fp}\boldsymbol{C}_f$$

ここまで紹介してきた m 因子の因子分析モデルは，L. L. Thurstone による多因子モデルのことでもあり，断るまでもなく斜交因子の共通因子分析モデルでもある。一般的に，項目分析から 1 つの尺度を構成することを目的とする場合には，m の数は 1 と考えられる。この場合には，(7)式からも明らかなように因子パターンと因子構造とは一致する。1 因子の場合には，そして，因子軸を回転する必要がないことから，主因子法（古くはセントロイド法）あるいは最尤法で求めた因子解となる。

構成した尺度の信頼性

1 次元尺度の構成の場合には，このように，因子解の値から適切な項目を選択することになる。この作業から構成された尺度の信頼性は，McDonald (1999) の ω として，あるいは項目の因子分析の共通因子空間での構成した尺度の共通性（Cattell & Tsujioka, 1964; 辻岡，1964）として，次の式で計算することができる（詳細は清水（2010）参照）。

$$(8)\quad h_s^2 = \omega = \frac{\left(\sum_{j=1}^{p}\lambda_{j1}\right)^2}{\left(\sum_{j=1}^{p}\lambda_{j1}\right)^2 + \sum_{j=1}^{p}\theta_j^2}$$

ここで λ は因子解（あるいは因子負荷量）であり，θ は独自性である。図 1 では，P1 ～ Pn の標準化した値がこの λ に相当する。

多次元の因子分析結果では，項目の選択は，因子パターンの値から行うことになる。そして，その場合でも各因子の信頼性は，(8)を拡張して計算することができる。詳しくは，清水 (2010) を参照されたい。

最後に

因子分析の黎明期にあっても，構成概念の内部構造を分析する手がかりは，Spearman (1927) が彼の主張する知能の一般因子 g と関連変数の相関係数を丹念に追求しているように，因子と変数との相関にあった。知能構造に関する L. L. Thurstone との論争にはこの相関の呼び方をめぐって，saturation か loading かという論争も潜在していたようある。この二人から影響を受けたエジンバラ大学の G. Thomson は，先に引用したように，多因子説ばかりではなく，loading に賛同したようである。最終的には，C.

Spearmanとの共同論文もあるK. J. Holzingerがシカゴ大学へ移り，因子負荷量（factor loading）を因子構造（factor structure）と因子パターン（factor pattern）として，因子分析体系と用語の整理を行った。

Mosire（1936）のいう項目分析の外的基準として潜在変数である因子を想定することは，実は，1次元構造の場合には因子構造を項目分析の対象とするということである。因子構造という観測変数と潜在変数との相関という定義は，拡張性に優れている。例えば，上で紹介した構成尺度の共通性は，構成した尺度と項目の因子分析から得られた因子得点との相関である構成尺度の因子構造を手がかりとして計算することができた。また，Cattell & Tsujioka（1964）などによる尺度での因子分析から抽出した共通因子空間に項目を布置させる延長因子分析（例えば，辻岡・清水，1975; 清水，2005など）がその応用例である。項目応答理論では，潜在的変数θ（古典的テスト理論の真の得点）を想定する（例えば，池田（1994）など）。このθと項目との相関を計算することによって，この場合の因子構造を得ることができる。中山（2010）が試みたように項目応答理論からもωへの展開が可能である。

ここで紹介してきた因子分析の古典の序に，実際に手計算を行った関係者への謝辞が記されていることがある。理論的に望ましいことは分かっていても，計算結果を手に入れることが困難な時代であった。do-littleの精神の下，実証的証拠を近似的に推定値から求める工夫の1つがIT相関であった。PC時代の研究者は，Richardson（1936）からはじまった手計算時代のIT相関という遺物に，別れを告げなければならない。

ここで紹介してきた因子分析の黎明期は職業指導の特性・因子論の成立期とも重なる（清水，2008）。Thomson（1939）は，興味深いことに因子分析の応用としてvocational guidanceにも頁を割いている。

引用文献

Cattell, R. B., & Tsujioka, B. (1964). The importance of factor-trueness and validity, versus homogeneity and orthogonality, in test scales. *Educational and Psychological Measurement*, **24**, 3–30.

Griffin, H. D. (1936). A further simplification of the multiple and partial correlation process. *Psychometrika*, **1**, 219–228.

Guilford, J. P. (1953). The correlation of an item with a composite of remaining items in a test. *Educational and Psychological Measurement*, **13**, 87–93.

Guilford, J. P. (1954). *Psychometrics methods* (2nd Ed.). New York: McGraw-Hill. (秋重義治（監訳）(1959). 精神測定法　培風館）

池田　央（1994）. 現代テスト理論　朝倉書店.

Holzinger, K. J., & Harman, H. H. (1941). *Factor analysis*. Chicago, IL: University of Chicago Press.

McDonald, R. P. (1999). *Test theory: A unified approach*. Mahwah, NJ: Lawrence Erlbaum.

Mosier, C. I. (1936). A note on item analysis and the criterion of internal consistency. *Psychometrika*, **1**, 275–282.

中山皓平（2010）. 心理尺度構成方法論の比較研究―FAモデルとIRTモデルを対象として―　関西大学大学院心理学研究科修士論文（未刊行）.

Richardson, M. W. (1936). Notes on the rationale of item analysis. *Psychometrika*, **1**, 69–76.

芝　祐順（1972）. 因子分析法　東京大学出版会.

清水和秋（2005）. 因子分析によるテスト構成　日本テスト学会第3回大会シンポジウム「心理テストの効用をめぐって―21世紀を展望する」発表論文集，26–27.

清水和秋（2008）. 職業指導の成立　日本キャリア教育学会編　キャリア教育概説（pp.46–48）　東洋館出版社.

清水和秋（2010）. 項目因子分析で構成した尺度の因子パターン，共通性，信頼性そして因子的真実性　関西大学心理学研究，**1**，9–24.

清水和秋・山本理恵（2007）. 小包化した変数によるパーソナリティ構成概念間の関係性のモデル化―Big Five・不安（STAI）・気分（POMS）―　関西大学社会学部紀要，38(3)，61–96.

Spearman, C. (1927). *The abilities of man*. London: Macmillan.

Thomson, G. (1939). *The factor analysis of human ability*. London: University of London Press.

Thurstone, L. L. (1935). *The vectors of mind*. Chicago, IL: University of Chicago Press.

辻岡美延（1964）. テスト尺度構成における新しい原理　心理学評論，**8**，82–90.

辻岡美延・清水和秋（1975）. 項目分析における項目統計量と構成尺度の統計量―因子的真実性係数と因子的妥当性　関西大学社会学部紀要，**7**(1)，107–120.

Wherry, R. J., Campbell, J. T., & Perloff, R. (1951). An empirical verification of the Wherry-Gaylord iterative factor analysis procedure. *Psychometrika*, **16**, 67–74.

柳井晴夫・繁桝算男・前川眞一・市川雅教（1990）. 因子分析―その理論と方法―　朝倉書店.