青森公立大学経営経済学研究
第17巻第２号（2012）

数段階のリッカート型データを間隔データとして使う場合の理論的根拠について

村尾　　博※

１．はじめに

リッカートの簡便法に基づく数段階の評定尺度データにも関わらず、何の理論的根拠を示すことなく量的データとして取り扱い、平均や標準偏差を計算している学術論文が見られる。身近な例では、授業評価に関する5段階評価に対して1,2,3,4,5といった点数化を行い、何の検証もないまま平気で平均や標準偏差を報告している大学もある。このような評定尺度データは本来的に順序データであり、したがって質的データである。手元のデータが量的データであるといった理論的根拠を示した上で、平均や標準偏差を報告するのであれば問題がない。しかし、手元のデータは順序データと思われるにも関わらず、いきなり平均や標準偏差を報告する姿勢は問題である。最高の教育を受けていると思われる人々に、このような現象が見られることは摩訶不思議である。

具体的な議論を始める前に議論の対象となっている「数段階の評定尺度データ」について、もう少し明確にしておこう。社会科学の分野におけるアンケート調査では次のようなタイプのデータを作ることが多い。「○○についてどう思うか」という質問（意見項目）に対し、「強く反対」「反対」「どちらでもない」「賛成」「強く賛成」といったように数個の反応カテゴリーを準備しておき、いずれか一つの反応カテゴリーを被験者に選択させる。そして賛否の順序に従って反応カテゴリーに点数をつける。例えば「強く反対」に１をつけ、順次2,3,4と点数化してゆき、「強く賛成」に5をつける。点数の付け方は「強く反対」に－2をつけ、順次－1,0,＋1,とし、「強く賛成」に＋2をつける場合もある。反応カテゴリーの個数は5個や7個となっている場合が一般的である。4個や6個という場合もある。このような点数化は「評定尺度」と呼ばれ、さらに「リッカート尺度」とも呼ばれている。このような数段階の評定尺度データが議論の対象である。レンシス・リッカート（Rensis Likert1932）が提唱した本来のリッカート尺度は、もっと精緻な尺度構成法であり、「相加評定尺度」や「加算評定尺度」と呼ばれ、一つのことに関連した10個から30個ぐらいの質問（意見項目）から構成され、被験者の得点は連続型データに近いような多数段階の整数値になる。一方、ここで述べているような1個の質問から構成され、被験者の得点が数段階の整数値になる評定尺度は、本来のリッカート尺度と異なるが、同じくリッカート尺度と呼ばれている。したがって本稿においても「数段階の評定尺度データ」と「数段階のリッカート型データ」とは入れ替え可能な用語として使う。そして実際に行われているアンケート調査では、多数個の質問から構成する本来のリッカート尺度よりも、1個の質問から構成する数段階のリッカート尺度の方が広く一般的に用いられている（Clason and Dormody1994）。

本稿における議論の出発点はインターネット上で見つけた次の記載に端的に現われている。

（１）「５段階のリッカートスケールは順序尺度か？それとも間隔尺度か？　ナンバリングに意味を持たないので順序尺度かと思ってましたが、知人が間隔尺度と申しており混乱しています[1)]。」

（２）「調査票で主に使われるのはリッカート尺度ですが、研究者は5段階や7段階の尺

※　青森公立大学准教授

度を間隔尺度として仮定し、測定を行っています。しかし、それが間隔尺度であるかどうかを厳密に検討していることは稀です。仮に、その尺度の等間隔性が確認されないならば、リッカート尺度は順序尺度データとなります[2)]。」

（3）「このリッカートの簡便法は一見して簡単な作業なので安易に使われがちです。そのうえ、この方法によって得られたデータはもっぱら間隔尺度のデータと見なして分析されますが、私は簡便法による値に等間隔性があるとはいい難いと考えています（つまり間隔尺度として扱うのは好ましくない）[3)]。」

（4）「連続量への願望は、よほど強いと見える。なにしろ、アンケート調査で、ある意見を提示して、賛否の程度を調べたデータが、あっという間に、『非常に賛成』5点、『どちらでもない』3点、・・・とされてしまう。しかも、一流といわれる学術雑誌の掲載論文すら、平均3.45 、相関係数0.789といった数値が、小数点以下2桁、3桁の精度で、恥ずかしげもなく報告されている[4)]。」

多くの記載者は5段階や7段階のリッカート型データが順序データであることを認識しているが、その一方において、多くの研究者が厳密な検討のないまま間隔データ（量的データ）として使っていることも記述している。この種の情報はインターネット上において多く見られる。つまり、冒頭のような現象や混乱は大学生から研究者まで広く見られることである。

一方、数段階のリッカート型データを間隔データとして使うことに関し、「理論的根拠」や「正当性」を示していると思われる情報を整理すると、次のようになる。

（1）多くの研究者が使っているから。

（2）順序が5段階以上になっている場合は間隔データとして使っても良いと言われているから。

（3）順序が5段階以上になっており、回答がひとつの反応カテゴリーに集中することなく全ての反応カテゴリーに適度に分散している場合は、間隔データとして使っても良いと言われているから。

（4）順序の間隔が一定であると仮定する。しかし、その仮定の根拠は何ら示されていない。

残念ながら、この程度の情報が多い。一方、まともな情報としては次の相反する2つがある。

（1）名義尺度・順序尺度・間隔尺度・比例尺度の順に尺度水準が高くなっており、高い水準の尺度で定義された測定値を低い水準の尺度の値に変換することは可能であるが、その逆はできない。

（2）下位の尺度から上位の尺度への変換には一定の理論的根拠が必要である。

後者の場合、一定の理論的根拠が必要であるとしながらも、そのための具体的な情報が示されていない。このような状況において、下位の尺度から上位の尺度への変換はできないと結論づけるのはたやすい。しかし、それでは冒頭のような現象や混乱は後を絶たない。冒頭に示すような現象や混乱は次のような「灰色のケース」に関わっていることが容易に想像できる。授業評価に関する5段階評価を例として考えると良い。5段階の順序は間隔がどこでも同等であると証明はできないが、間隔が極端に異なるとも考えにくいような灰色のケースである。このようなリッカート型データに直面した場合、間隔データ（量的データ）として使いたいといった心情は理解できる。

そこで本稿では次のような基本的な立場を取ることにする。下位の尺度から上位の尺度への変換には一定の理論的根拠が必要である。この立場を取り、本稿では順序データなのか間隔データなのか迷うような数段階のリッカート型データを間隔データとして使うことができる理論的根拠や条件を探る。

そのような理論的根拠のひとつは「順序尺度→間隔尺度」の変換器として正規分布を用いる

ものである。例えば標準正規分布はパーセント点という順位情報とｚ値という量的情報とが１対１で対応し、しかも理論的な確率分布として最も広く一般的に用いられている。したがって順位情報から量的情報への変換器として最適である。当然のことながら、このような変換を行うには手元のデータが正規分布に近い形で分布しているといった前提が必要である。このアプローチはリッカート尺度の生みの親であるリッカート（1932）に見られる。

正規分布を利用して「順序尺度→間隔尺度」の変換を行うと明言はしていないが、リッカートが提唱した手法はこのような変換に基づく。そして「順序尺度→間隔尺度」の変換を厳密に行う手法をシグマ法（sigma method）と呼び、彼が提唱した簡便法と区別している。彼が提唱した簡便法は1,2,3,4,5のように等間隔の整数値で行う点数化である。シグマ法から得たデータと簡便法から得たデータとが統計的に極めて良く似ていることから、簡便法を使って態度を測定することを提唱している。このリッカートの簡便法が混乱の始まりであると思われる。しかし、簡便法から得た数段階の整数値データを間隔データとして使いたいのであれば、「順序尺度→間隔尺度」の変換器としての正規分布を通じて出てきたデータであることを示す必要がある。言い換えると、手元のデータが正規分布に近い形で分布していることを示す必要がある。このようなことが認識されず、「順序が5段階以上になっている場合は間隔データとして使っても良い」といった通説になったと思われる。冒頭のような現象や混乱を防ぐにはリッカートのシグマ法を使うことを強調するのが良いかも知れないが、シグマ法は手間隙がかかり、一般的でない。そうであるからこそ、リッカートも簡便法を提唱したのである。

そこで本稿では順序データなのか間隔データなのか迷うような数段階のリッカート型データが間隔データとして使うことができるか否かを検証する実用的な手法を紹介する。それは「適合度の検定」と呼ばれるものである。まず、数段階のリッカート型データに関して次のような仮説を立てる。

帰無仮説：手元の観測データは正規分布からのデータである。

対立仮説：手元の観測データは正規分布からのデータでない。

もし、帰無仮説が棄却されない観測結果になれば、それは手元の観測データを間隔データ（量的データ）として使ってもよい理論的根拠になる。このような手法であれば容易に実行でき、実用的である。ただし、次のような認識も必要になる。連続型潜在変数としての正規確率変数に対して数段階の離散型観測値を得る想定になっているので、観測誤差が含まれた離散型データになる。このような点を留意点として述べることも本稿の重要ポイントになる。

これ以降における本稿の構成は次のようになっている。セクション2ではリッカートのシグマ法や簡便法に関わる論理を説明する。そこではリッカート（1932）が使った数値例を用い、冒頭のような誤解や混乱の元になっている点も具体的に示す。セクション3では適合度の検定に基づき、数段階のリッカート型データを間隔データとして使うことができるか否かを検証する手法を紹介する。ここでもリッカートの数値例を用い、仮説検定を行う。仮説検定において間隔データとして使える結果になっても、観測データには観測誤差が含まれる。このような点を含め、幾らかの留意点をセクション4で示す。セクション2でも3でも、潜在変数の分布は正規分布である。正規分布以外の分布でも利用可能かをセクション5で考察する。そして本稿のまとめをセクション6で行う。

２．リッカートのシグマ法と簡便法

ここではリッカート（1932）に従い、彼が提案したシグマ法と簡便法を要約すると共に、その論理を明確にする。シグマ法にせよ、簡便法にせよ、「順序尺度→間隔尺度」の変換器として正規分布を利用するというアイデアの下での尺度構成法であることを示す。

例えば「国際主義のスケール」といった一つの量的特性を測定するとしよう。まず、国際主

義に関連した具体的な内容の賛否程度を聞くような質問（意見項目）を10個から30個ぐらい準備する。質問の一つは「正しいか否かに基づき、我々は国のために戦うべきか」といった内容である。各質問には「強く反対」「反対」「どちらでもない」「賛成」「「強く賛成」といったように数個の反応カテゴリーを設けておく。被験者は各質問に対して一つの反応カテゴリーを選ぶ。そして反応カテゴリーに対して1,2,3,4,5のように等間隔で整数値を与える。被験者ごとに質問の得点を合計し、それを量的特性「国際主義スケール」に関する被験者の得点にする。多くの質問の得点を加算することから、彼が提案した尺度は「相加評定尺度」や「加算評定尺度」とも呼ばれている。したがって一つの量的特性「国際主義スケール」に関する被験者の得点は連続型データのような多数段階の整数値となる。これが本来のリッカート尺度である。

次は1個の質問（意見項目）に対する点数化を考える。具体例で考えた方が理解しやすいので、リッカート（1932）の表Ⅱに示されている数値例を用いる。反応カテゴリーは「強く賛成」「賛成」「どちらでもない」「反対」「強く反対」の5段階になっている。100人からなる標本において、反応カテゴリーの相対度数は「強く賛成」から「強く反対」への順に[0.13,0.43,0.21,0.13,0.10]となっている。「強く賛成」が左端、「強く反対」が右端に位置するように並べ、相対度数に従って標準正規分布を5区間に分割すると、図1のようになる。

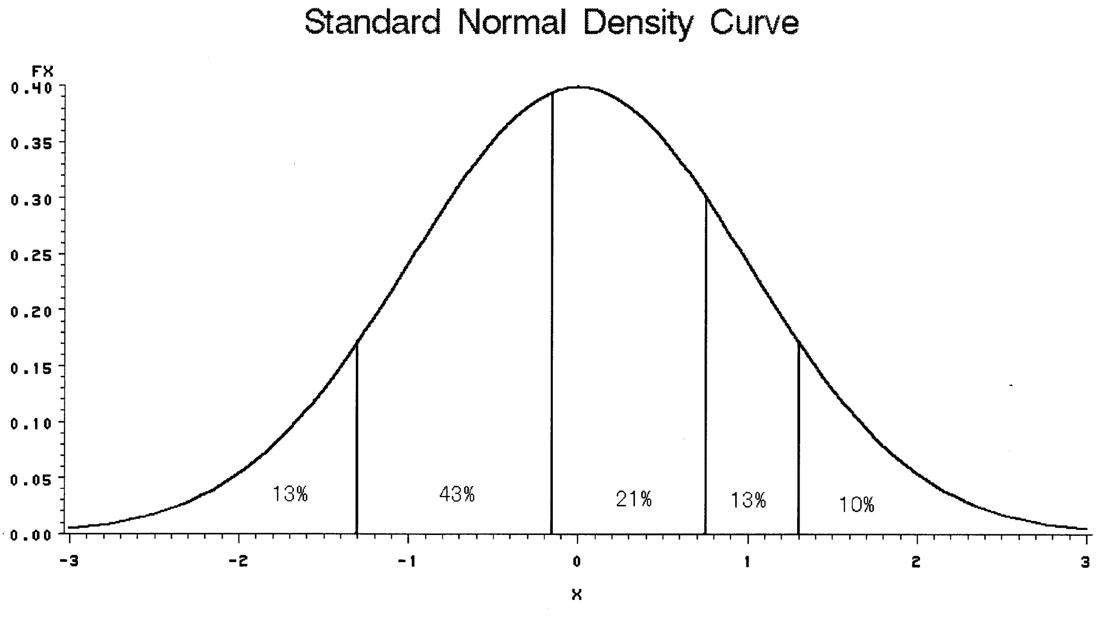

図１．反応カテゴリーの相対度数に従って標準正規分布を分割した図

図1のように分割された標準正規分布において、エリアの平均値を順位の得点にする。これは標準正規分布を用いて「順序尺度→間隔尺度」の変換を行うことを意味している。第1の「強く賛成」のエリアは13％の相対度数であるから、エリヤの平均値（シグマ値と言う）は－1.63となる。このようにして全ての順位に対してシグマ値を計算すると、［－1.63, －0.43, ＋0.43, ＋0.99, ＋1.76］が得られる。これがリッカートのシグマ法（sigma method）である。ここまでのことを整理すると、次の表のようになる。

表１．リッカートのシグマ法と簡便法との関係を示す数値例

反応カテゴリー	強く賛成	賛成	どちらでもない	反対	強く反対
相対度数	13％	43％	21％	13％	10％
シグマ法の値	－1.63	－0.43	＋0.43	＋0.99	＋1.76
原点を移動させた場合の値	1.00	2.20	3.06	3.62	4.39
簡便法の値	1	2	3	4	5

確かにシグマ法は優れた尺度構成法であるが、複雑な計算に手間暇がかかる。リッカート（1932）はシグマ法の計算式を示さなかったが、参考までにシグマ法の計算式（Golden and Brockett 1987）を書くと、次のようになる。

$$Z_i^* = \frac{1}{f_i\sqrt{2\pi}}\left\{\exp\left(-\tfrac{Z_{i-1}^2}{2}\right) - \exp\left(-\tfrac{Z_i^2}{2}\right)\right\} \qquad (1)$$

ただし、Z_i^*は反応カテゴリーiのシグマ値、f_iは反応カテゴリーiの相対度数、Z_{i-1}とZ_iは標準正規分布において当該反応カテゴリーiが占めるエリアの下限z値と上限z値である。

このようなシグマ法の計算を10個から30個ぐらいの全質問に対して実行することは容易でない。そこでリッカートは各質問における点数化では1,2,3,4,5といった等間隔の整数値を与えることを提案している。これがリッカートの簡便法である。シグマ法に代えて簡便法を用いる理論的根拠は多少長くなるが、次のようになる。まず、表1において、「強く賛成」に対応するシグマ値－1.63が1.00となるように原点を移動させると、「原点を移動させた場合の値」になる。その下側には「簡便法の値」を示している。これら2組の数値列を比較してほしい。手間暇かけて計算したシグマ法の数値と簡便法の数値との間に大差がない。これだけで簡便法を使う理由にはならないが、良く似た数値列であることがキーポイントになる。そして全ての質問の点数を合計して被験者の得点を得る。このようにして得た被験者全体の、シグマ法のデータと簡便法のデータは相関がほぼ完全である等、統計的に極めて良く似た性質を持っている。このようなことからリッカートは、手間隙のかかるシグマ法に代えて簡便法を提唱した。宝月（1989, p.64）も次のように述べている。「シグマ値法は、意見項目ごとの反応分布に標準正規分布（平均0、標準偏差1の正規分布）をあてはめて反応カテゴリーに応じるウェイトを求めるものであるが、採点方法が複雑であること、また算定された態度得点が簡便法によるものと高い相関があることから、簡便法で処理されることが多い。」

昨今のリッカート型アンケート調査では、シグマ法を用いず、単純に等間隔の整数値を与えるようになっている。しかも1個の質問から得た数段階のリッカート型データに基づく研究や調査報告が多くなっている。このことが冒頭のような現象や混乱につながっているように思える。しかし、リッカートが提唱したような、多数個の質問から得る加算評定尺度データと、1個の質問から得る評定尺度データとは次の重要な点で大きく異なる。まず、多数個の質問から得る前者は連続型のような多段階の整数値となり、正規分布の観測値として相応しい。これに対して1個の質問から得る後者は数段階の整数値であり、

正規分布の観測値として相応しいとは言いがたい。さらに前者の場合、正規分布に近い分布になるように、最終的に用いる質問の組み合わせを変えることもできる。この点は重要である。個々の質問における分布が左や右へ歪んでいても、それらの歪みが無くなるように質問を組み合わせることができるからである。しかし、後者には、そのようなメカニズムが存在しない。

このようなことを理解した上で、議論の本筋へ戻ろう。つまり、1個の質問から得る数段階のリッカート型データを間隔データとして使用するための理論的根拠や条件を探ろう。そのような理論的根拠としてリッカート（1932）の論理がそのまま利用できる。つまり、正規分布を「順序尺度→間隔尺度」の変換器として利用するアイデアである。そのためには手元のデータが正規分布に近い形で分布していることが前提になる。この点は重要である。手元のデータが正規分布に近い形で分布しているから、正規分布を「順序尺度→間隔尺度」の変換器として利用できるのである。したがって手元のデータの分布は何でも良いという具合にはいかない。「順序尺度→間隔尺度」の変換器として正規分布を利用するリッカートの論理に関しては村尾（2005）が詳しい。Borgatta and Bohrnsted (1980)も、順序データを間隔データとして使用する場合の理論的根拠にはベル型分布（正規分布）が必要であることを述べている。

このセクションでは「順序尺度→間隔尺度」の変換器として正規分布を利用するというアイデアの下で、リッカートのシグマ法や簡便法が生まれてきた経過を見てきた。そして順序データなのか間隔データなのか迷うような数段階のリッカート型データが間隔データとして使うことができる一定の理論的根拠も、正規分布による「順序尺度→間隔尺度」の変換に求めることが出来ることを示した。そのためには手元のデータが正規分布に近い形で分布していることが前提になる。

3．適合度の検定

前セクションの発想から少し離れ、ここでは別の方向の理論的根拠を考える。それは「適合度の検定」である。つまり、数段階のリッカート型データに関して次の仮説を立てる。

帰無仮説H_0：手元の観測データは正規分布からのデータである。

対立仮説H_1：手元の観測データは正規分布からのデータでない。

もし、帰無仮説が棄却されない検定結果になれば、それは手元の観測データを間隔データとして使ってもよい理論的根拠になる。つまり、正規分布に従う間隔データといった制限がつくが、数値間の間隔は一定であると仮定できる。恣意的に間隔が一定であると仮定するのと全く異なる。

このような適合度の検定は「順序尺度→間隔尺度」の変換といった発想を必要としないが、そのような発想の下でも、手元のデータが正規分布に近い形で分布しているか否かを検証するのに使われる。そういった意味では、適合度の検定は前セクションの延長線上にある具体的な検定法とも言える。先の村尾（2005）の論文でも適合度の検定のことを簡単に述べているが、そこでの議論の焦点は簡便法から起因する観測誤差の分析に向けられている。

それでは適合度の検定がどのようなものかを具体的に見ていこう。まず、分布の全区間をm個の区間に分け、次のような具体的な帰無仮説を立てる。

$$\text{帰無仮説 } H_0 ：\ f_i = p_i^0 \quad \text{for } i = 1,2,3,\cdots,m$$

ただし、f_iは観測データにおけるi番目の区間の相対度数（relative frequency）、p_i^0は帰無仮説の下での理論的な確率分布、この場合は正規分布の下での、当該区間の確率（probability）である。標本の大きさをnとするとき、i番目の区間における観測度数（observed frequency）は$O_i = nf_i$となり、確率p_i^0に対応する期待度数（expected frequency）は$E_i = np_i^0$で与えられる。標本の大きさnが十分大きいとき、次の結果が得られる。

$$\chi^2 = \sum_{i=1}^{m} \frac{(O_i - E_i)^2}{E_i} \overset{a}{\sim} \chi^2(m-1-k) \qquad (2)$$

ただし、k は期待度数 E_iを算出するために推定したパラメータの個数であり、$\chi^2(m-1-k)$は自由度が$(m-1-k)$のカイ2乗分布を示す。経験的なルールではあるが、(2) を使用するためには、区間の個数がm■5であり、各区間においてE_i■5を満たすことが必要と言われている。

検定の数値例としてはリッカート（1932）の表Ⅱに示されている数値例をそのまま用いる。分布の全区間を m 個に分割するが、反応カテゴリー数が5個であるから、m ＝5となる。一方、母集団が正規分布であるから、母平均と母分散を推定する必要があり、k ＝2となる。したがって検定に使うカイ2乗分布の自由度は2となる。一方、E_i■5を満たすことは、各反応カテゴリーに5個以上の回答が得えられるようなアンケート調査を実行することであり、そのようなことは困難でない。われわれの数値例では、最小の相対度数が10％であるから、それを最小の p_i として用いて逆算すると標本数 n が50以上であることが求められる。われわれの数値例では n＝100であるから、各区間においてE_i■5を満すであろうことは容易に想像できる。事実、後で示すように各区間においてE_i■5を満たしている。いずれにせよ、適合度の検定に必要な前提条件を満たすことは困難なことではない。

それでは数値例を用い、適合度の検定を実行していこう。「強く賛成」から「強く反対」への順に相対度数が [0.13,0.43,0.21,0.13,0.10]、標本の大きさがn＝100であるから、観測度数O_iは[13,43,21,13,10] となる。先の簡便法と同じく得点は「強く賛成」から「強く反対」への順に1,2,3,4,5とする。このようなデータを標準化したと想定する。それに対応する標準正規分布の方も5つの区間に分割する。このとき標準正規分布の全範囲、つまり－∞から＋∞の範囲を等間隔で5分割するという具合にいかない。このことに関連し、標準正規確率変数のほとんどのデータが含まれる (－3.0,＋3.0) の範囲を等間隔で5分割するのか、それとも(－2.5,＋2.5)の範囲を等間隔で5分割するのかといった疑問が生じる。シグマ法のように相対度数に従って標準正規分布を5分割すれば、このような問題は生じないが、それでは全ての区間においてO_i＝E_iとなってしまう。シグマ法と異なり、観測値が等間隔の整数値となっていることから、標準正規分布の横軸を等間隔で区切る必要がある。ただし、標準正規分布は－∞から＋∞に広がる分布であるから、両端の区間は別にする必要がある。このようなことを考慮すると、次のアイデアが良さそうだ。

両端の相対度数が13％と10％であるので、平均の11.5％を採用し、11.5％に相当する区間を分布の両端に作る。その結果、(－1.20,＋1.20)の区間を3等分することになり、階級の境界値は[－1.20,－0.40,＋0.40,＋1.20] となる。この区分けに従って期待度数 E_i を計算すると、「強く賛成」から「強く反対」への順に [12,23,30,23,12] が得られる。ただし、四捨五入して期待度数も整数値となるようにしている。観測度数O_iが [13,43,21,13,10]、期待度数 E_i が [12,23,30,23,12] であることから、検定統計量の観測値は χ^2 ＝24.86となる。

この検定の場合は帰無仮説を採択したいので、帰無仮説を採択するときの誤りの確率が小さくなるように有意水準を通常よりも高めの20％に設定しよう。一方、検定に使う分布は $\chi^2(2)$である。その結果、棄却点は3.22となり、20％の有意水準で帰無仮説は棄却される。なお、検定のP値はほぼゼロである。したがって有意水準が多少変化したとしても、検定結果が逆転することはない。

リッカート（1932）が使っていた数値例であるから、帰無仮説が採択されると思っていたが、そうはならなかった。そこで観測データのヒストグラムと正規分布とを重ね合わせて比べてみよう。それを図2に示す。両者間で平均と分散がそれぞれ等しくなるようにしている。

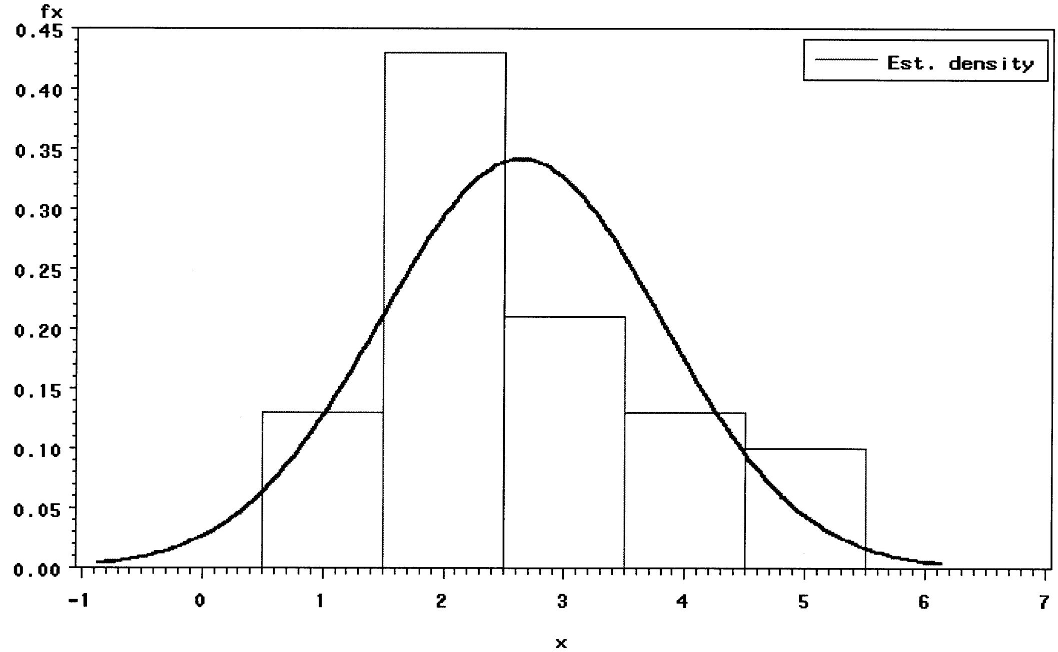

図２．(1,2,3,4,5) の点数化を用いた場合のヒストグラムと正規分布との比較

図2に見られるように左端から2番目「賛成」の相対度数が飛び抜けて大きい。リッカート型アンケート調査では、歪みをもった分布になることが良く知られている。つまり、人々はポジティブ（賛成）側の反応カテゴリー、しかし一番端でない反応カテゴリーを選ぶ傾向があることが知られている（Rowan 2001)。それが如実に現れる結果になっている。このことが帰無仮説の棄却に深く関わっていることは明白である。このような典型的パターンを考えると、1個の質問から得る簡便法のデータが間隔データとして使える可能性は極めて低いと言えるかも知れない。

次は観測度数O_iを [13,43,21,13,10] から [13, 28,31,18,10]へ変更してみよう。両端の観測度数が変わっていないことから期待度数E_iは変わらない。その結果、検定統計量の観測値がχ^2＝2.62となり、20％の有意水準でも帰無仮説は採択されるようになる。この場合は手元のデータを間隔データとして使える。さらに正規分布からのデータであると仮定できるので、様々な統計的推測においても便利である。

このセクションを終えるにあたり、リッカート（1932）が使っていた数値例のデータを間隔データとして使えるようにする救い道がないのかを考えてみた。再び点数化を行う必要があるが、シグマ法があるではないか。既に見てきたように、シグマ法は母集団が正規分布である仮定の上に成り立っている。簡便法から得たデータのヒストグラムがM字型のような場合は母集団が正規分布であると仮定するには無理があるが、図2に示す程度の山形になっているのであれば、母集団が正規分布であると仮定することに躊躇する必要はない。何故ならば、シグマ法には次のような「調整機能」が備わっているからである。先の数値例の場合、[1,2,3,4,5] と点数化すれば間隔データにならないが、シグマ法によって [1.00, 2.20, 3.06, 3.62, 4.39] と点数化すれば間隔データになる。この違いは次のように考えれば良い。例えば相対度数が大きい反応カテゴリー「賛成」では、正規分布のデータとなるように、得点が2.00から2.20へ「調整」されたと考えれば良い。それに応じて反応カテゴリー「賛成」の階級幅も広くなっている。その他の順位も同様な形で調整される。整数値の点数化では歪んだヒストグラムを示すような場合でも、正規分布を反映するように数値を修正するのがシグマ法である。

このように考えると、シグマ法は優れた尺度構成法と言える。簡便法から得る数段階のリッカート型データを間隔データとして使う場合の理論的根拠を探るといった議論の本筋から少し逸脱したが、シグマ法のこのような性質は特記すべきものがある。

4．間隔データとして使う場合の留意点

本稿ではリッカート簡便法からなる数段階のデータを間隔データとして使う場合の理論的根拠を探ってきたが、次のような大きな考え方の上に成り立っている。観測変数が潜在変数の尺度を反映していると推測できる場合のみ、順序尺度の観測変数を間隔尺度とみなして良いとする考え方である。そして潜在変数の分布は正規分布であった。ところが、潜在変数としての正規確率変数は連続型であるの対し、観測変数は数個の整数値を取る離散型である想定になっている。このことから離散型の観測値は観測誤差が含まれたものになる。例えば潜在変数の値（真の値）が第2階級「賛成」内の2.4であっても、当該階級「賛成」の階級値2として観測されるといった意味の観測誤差が生じる。このような観測誤差の認識は大切である。例えば回帰分析において、説明変数が観測誤差を含んでいる場合、最も一般的な通常最小２乗推定法は不偏性も、一致性も保証せず、散々な推定結果になる。この点に関しては村尾（2004）が詳しい。このように間隔データとしての使用が可能な場合でも、リッカート簡便法からなる数段階のデータは純粋な間隔データのようにはいかない。

一方、1個の質問から得る数段階のシグマ値データはどうであろうか。各階級において観測誤差がゼロとなる設計になっているが、潜在変数としての正規確率変数が連続型であるの対して観測変数は数段階の離散型である想定に変わりはなく、観測誤差は生じる。例えば潜在変数の値（真の値）が第2階級「賛成」内の2.4であっても、当該階級「賛成」の平均値2.2として観測されるといった意味の観測誤差が発生する。つまり、各被験者についての真の値と観測値との間に誤差が生じる点は解消されていない。

ここで述べているような観測誤差は、正確なデータが得られないのが普通になっている分野では大きな問題にならないかも知れない。とは言うものの、観測誤差が存在するといった認識も必要であり、それに応じたデータ分析手法が求められる。さらに述べるならば、間隔データとしての使用が可能であっても、本稿で議論しているようなデータであれば、間隔データに基づく分析と順序データに基づく分析との双方を実行することにより、より有用な情報が得られると考える。例えば相関分析であれば積率相関係数と順位相関係数との双方を用いるといった具合である。

5．正規分布以外の確率分布も使えるのか

リッカート簡便法からなる数段階のデータを間隔データとして使う場合の理論的根拠を探ってきたが、重要なチェックポイントは観測変数の分布が潜在変数の分布を反映しているか否かである。ここまでの議論では潜在変数の分布は正規分布であった。正規分布以外にも、そのような潜在変数として使える確率分布があると考えるのが自然である。そこで、この点について考えてみよう。

そもそもの出発点は、リッカート簡便法からなる数段階のデータを作り、出来ることならば間隔データとして使いたいといった動機であった。したがって潜在変数の分布としては、広く一般的に知られている確率分布であり、比較的簡単に目的が達成できる分布であることが求められる。少なくともリッカートのシグマ法よりも簡単に目的が達成できる条件は必要であろう。そして標本の大きさや自由度に応じて分布形が変化する分布は、考慮している潜在変数の分布としては問題外である。パラメータの調整によって歪みを持った分布形から左右対称の分布形まで様々な分布形へ変化するガンマ分布やベータ分布も、問題外である。このような変幻自在な確率分布を用いると、極論すると、どのような順序データでも間隔データとして使えるようになり、「順序の間隔が一定であると恣意的に仮定する」のと大差がなくなる。繰り返しになるが、

変幻自在なガンマ分布やベータ分布の使用は、異なった順序データを異なった基準で恣意的に数量化することに繋がるので問題外である。

一方、対数正規分布は微妙である。対数正規分布であれば、簡便法の点数化の下で歪んだ分布形を示すデータにも対応できる点が魅力的である。データのヒストグラムと対数正規分布とを重ね合わせて分布形が似ているか否かといった視覚的な比較は容易であっても、その適合度の検定となると計算が複雑になる。対数正規分布を使う適合度の仮説検定よりも、リッカートのシグマ法を使った方が簡単であろう。そういった意味において対数正規分布の使用は本末転倒といった感じがしないでもない。

以上のようなことを考慮すると、潜在変数の分布としては、正規分布、一様分布、三角分布ぐらいから選ぶのが良さそうだ。特に正規分布であれば、標準的に使われている分析道具や手法が豊富に存在し、何かと便利である。一方、１個の質問から得るリッカート簡便法のデータが、どのような典型的パターンの分布になるかといった現実的な側面を考えると、一様分布の出番は少ないようだ。また、正規分布と三角分布とは左右対称の山形分布といった意味で似ており、三角分布の使用を考えるような状況では、便利な正規分布の方を使うことになると思われる。そういった意味において三角分布の出番も少ないであろう。明確なボーダーラインを引くことは出来ないが、結論的には正規分布だけを考えれば良いようだ。

６．おわりに

リッカート簡便法から得る数段階の評定尺度データを間隔データとして使う場合の理論的根拠について考察した。その一つはリッカート（1932）が使った論理をそのまま使うものであり、正規分布を「順序尺度→間隔尺度」の変換器として利用するアイデアである。そのためには手元のデータが正規分布に近い形で分布していることが前提になる。別の一つは「順序尺度→間隔尺度」の変換という発想から離れ、適合度の検定を利用し、「手元の観測データは正規分布からのデータである」といった検定結果を得ることである。この手法は比較的簡単に実行できる利点もあり、実用的である。そして、この場合の適合度の検定を詳しく見てきた。

このような理論的根拠は、観測変数が潜在変数の尺度を反映していると推測できる場合のみ、順序尺度の観測変数を間隔尺度とみなして良いとする考え方に基づく。潜在変数としての正規確率変数は連続型であるの対し、観測変数は数個の整数値を取る離散型である想定になっていることから、離散型の観測値は観測誤差が含まれたものになる。このような認識は大切である。間隔データとしての使用が可能な場合でも、数段階のリッカート型データは純粋な間隔データのようにはいかない。したがって分析手法も限られてくる。さらに述べるならば間隔データとしての使用が可能な場合でも、間隔データに基づく分析と順序データに基づく分析との双方を実行することにより、より有用な情報が得られると考える。

最後に潜在変数となりうる確率分布に関し、対数正規分布やガンマ分布などを含め、幾らかの確率分布について考察した。一様分布や三角分布も潜在変数の分布になりうるが、その出番は少ないようだ。

（2011年11月25日受付、2012年２月３日受理）

注

1) ウェブサイト「Yahoo Japan 知恵袋」に記載された質問と回答「５段階のリッカートスケールは順序尺度か？それとも間隔尺度か？ナンバリングに意...」（http://detail.chiebukuro.yahoo.co.jp/qa/question_detail/q1359833538, 2011-11-21）
2) 清水裕士，ウェブページ「順序カテゴリカルデータでシミュレーション（その１）」（http://norimune.blog15.fc2.com/blog-entry-551.html, 2011-11-21)
3) 片所強，ウェブページ「リッカートの簡便法とシグマ法」（http://homepage2.nifty.com/nandemoarchive/toukei_hosoku/likert.htm,2011-11-21）

4) 杉万俊夫, ウェブページ「講義『グループ・ダイナミックス』シラバス」
（http://www.group-dynamics.org/sugiman/research/theoretical_study/T-012.htm, 2011-11-21）

参考文献

Abelson, Robert P., and John W. Tukey (1970), "Efficient Conversion of Non-Metric Information into Metric Information," *The Quantitative Analysis of Social Problem* (ed. Edward R. Tufte), Addison-Wesley, 407-417.

Agresti, Alan (1990), *Categorical Data Analysis*, John Wiley & Sons.

__________ (1996), *An Introduction to Categorical Data Analysis*, John Wiley & Sons.

Birkett, Nicholas J. (1986), "Selecting the Number of Response Categories for a Likert-type Scale," *Proceeding of the Survey Research Methods Section*, *American Statistical Association*, 488-492.

Borgatta, Edgar F., and Geoge W. Bohrnstedt (1980), "Level of Measurement: Once Over Again," *Sociological Methods and Research*, vol. 9 (November), 147-160.

Bross, Irwin D. J. (1958), "How to Use Ridit Analysis," *Biometrics*, vol. 14 (March), 18-38.

Champion, Dean J. (1968), "'Some Observations on Measurement and Statistics': Comment," *Social Forces,* vol. 46 (June), 541.

Clason, Dennis L., and Thomas J. Dormody (1994), "Analyzing Data Measured in Individual Likert-Type Items," *Journal of Agricultural Education*, vol. 34 (4), 31-35.

Easthope, Gary (1974), *A History of Social Research Methods*, Longman. (=1982,阿久津昌三，他４名訳『社会調査方法史』慶應通信.)

Golden, Linda L. and Patrick L. Brockett (1987), "The Effect of Alternative Scoring Methods on the Analysis of Rank Order Categorical Data," *Journal of Mathematical Sociology*, vol. 12 (4), 383-414.

Henkel, Ramon E. (1975), "Part-whole Correlations and the Treatment of Ordinal and Quasi-interval Data as Interval Data," *Pacific Sociological Review*, vol. 18 (1), 3-26.

Kim, Jae-On (1975), "Multivariate Analysis of Ordinal Variables," *American Journal of Sociology*, vol. 81 (2), 261-298.

Labovitz, Sanford (1967), "Some Observations on Measurement and Statistics," *Social Forces*, vol. 46 (December), 151-160.

__________ (1968), "Reply to Champion and Morris," *Social Forces*, vol. 46 (June), 543-544.

__________ (1970), "The Assignment of Numbers to Rank Ordered Categories," *American Sociological Review*, vol. 35(June), 515-524.

__________ (1971), "In Defense of Assigning Numbers to Ranks," *American Sociological Review*, vol. 36 (June), 521-522.

__________ (1972), "Statistical Usage in Sociology: Sacred Cows and Rituals," *Sociological Methods and Research*, vol. 1 (August), 13-37.

__________ (1975), "Comments on Henkel's Paper: The Interplay Between Measurement and Statistics," *Pacific Sociological Review*, vol. 18 (January), 27-35.

Likert, Rensis (1932), *A Technique for the Measurement of Attitudes*, New York: Archives of Psychology.

Mayer, Lawrence S. (1971), "A Note on Treating Ordinal Data as Interval Data," *American Sociological Review*, vol. 36 (June), 519-520.

Morris, Raymond N. (1968), "'Some Observations on Measurement and Statistics': Further Comment," *Social Forces*, vol. 46 (June), 541-542.

Morrison, Donald G., and Norman E. Toy (1982), "The Effect of Grouping Continuous Variables on Correlation Coefficients," *Marketing Science*, vol. 1 (4), 379-389.

O'Brien, Robert M. (1979), "The Use of Pearson's R with Ordinal Data," *American Sociological Review*, vol. 44 (October), 851-857.

Roberts, James S., James E. Laughlin, and Douglas H. Wedell(1999)," Validity Issues in the Likert and Thurstone Approaches to Attitude Measurement," *Education and Psychological Measurement*, vol. 59 (2), 211-233.

Rowan, John (2001), "Heavier Than Ultramarine?" *The Psychologist*, vol. 14(12), 624.

Thorndike, Edward L. (1919), *An Introduction to the Theory of Mental and Social Measurements* (Second Edition), Columbia University Press.

Thurstone, L.L., and E. J. Chave (1929), *The Measurement of Attitude*, University Chicago Press.

Schulz, E. Matthew, and Anji Sun (2001), "Controlling for Rater Effect When Comparing Survey Items With Incomplete Likert Data," *CT Research Report Series*

2001-2, (February).
Vargo, Louis G. (1971), "Comment of 'The Assignment of Numbers to Rank Order Categories'," *American Sociological Review*, vol. 36 (June), 517-518.
Wilson, Thomas P. (1971), "Critique of Ordinal Variables," *Causal Models in the Social Sciences* (ed. H. M. Blalock), New York: Aldine, 415-431.
林知己夫（1993),『行動計量学序論』 朝倉書房.
福武直（1984),『社会調査　補訂版』 岩波全書.
宝月誠・中道寛・田中滋・中野正大（1989),『社会調査』 有陽閣Sシリーズ.
マン, P.H.（1982),『社会調査を学ぶ人のために』(中野正大　訳) 世界思想社.
村尾博（2004),「リッカート型項目データの回帰への使用と通常最小２乗推定量」『青森公立大学経営経済学研究』9(2),63-79.
村尾博（2005),「リッカート型項目データの間隔データとしての使用」『青森公立大学経営経済学研究』10(2),3-19.
メイ，ティム（2001),『社会調査の考え方』(中野正大監訳) 世界思想社.
安田三郎・原順輔（1982),『社会調査ハンドブック　第３版』 有陽閣双書.

Theoretical Bases of the Usage of Likert-type Data as Interval Data

Hiroshi MURAO

Abstract

The Likert-type of survey is very popular in the social research. Here the Likert-type of survey means that we get several integer values such as 1,2,3,4,5 based on the degree of agreement to a question, which are ordinal data by construction. However, many researchers or the like use such ordinal data as interval data without showing a theoretical base of the usage. This paper investigates theoretical bases for such usage, i.e., the usage of the ordinal data of the Likert-type as interval data. We can adapt the logic of Likert (1932) who uses the normal distribution as a transformation device from ordinal data to interval data. Another theoretical base is a statistical hypothesis test which is called as "the goodness-of-fit test." This paper explores the approach of the goodness-of-fit test in details for the given topic. Some other points are also discussed with regard to the given topic.