

意識調査と数量化

宅地開発公団 企画用地部
 経営課長 上本 仁 士
 (前・総理府 広報室)

1. 昭和52年5月に、総理府が実施した「道路に関する世論調査」の中に、次のような二つの質問がありました。

質問1. あなたは、道路が整備されたために、生産地から新鮮な魚や野菜が消費者の食卓に届くようになったと思ったことがありますか、そういうふうには思いませんか

(答)	思ったことがある	思ったことはない	わからない	計
	53%	38%	9%	100%

質問2. (ここ数年間では、鉄道による貨物輸送よりも自動車(道路)による貨物輸送の方が増えていることを説明した上で) あなたは、自動車による貨物輸送をもっと増やした方がよいと思いますか、現在の程度でよいと思いますか、それとももっと減らした方がよいと思いますか

(答)	もっと増やした方がよい	現在の程度でよい	もっと減らした方がよい	わからない	計
	13%	46%	19%	22%	100%

一読したところでは、質問1を尋ねられた人は、意味がわからなくてとまどうかも知れませんが、質問の作成者である私の意図としては、道路整備が、貨物輸送の面で貢献していることを人々が認識しているか否かを観察したかった

わけです。

私の当初の予想からすれば、道路が整備されれば、それによって例えば、生鮮食料品もそれだけ早くかつ遠方の消費者まで届くようになる。と一度でも考えたことのある人は、質問2に対する反応においても、自動車による貨物輸送に好意的であると考えたわけです。

この私の予想は、質問1と質問2の関係を表わした次のクロス集計表をみれば、的外れでなかったことが言えると思います。

「道路が整備されたために、生鮮食品が消費者に届きやすくなったと思うか」

質問1 \ 質問2	(回答者数)	思が つあ た る こ と	思は つな いた い こ と	わ か ら な い	計
総 数	(2,507人)	53%	38%	9%	100%
[自動車による貨物輸送]					
もっと増やした方がよい	(317)	69	28	3	100
現在の程度でよい	(1,164)	55	38	7	100
もっと減らした方がよい	(473)	48	45	7	100
わからない	(553)	45	38	17	100

2. ところで、当然のように私の予想が的外れではなかったと断定していますが、果たしてこのクロス表からそういうことが言えるのでしょうか。

生鮮食品の輸送に果たす道路の効用を認める人(思ったことがあるという回答者)は、自動車による貨物輸送を増やした方がよいという意見(自動車による貨物輸送に好意的)が段々と多くなっている(48%→55%→69%)のに反して、これとは全く逆に、生鮮食品の輸送に果たす道路の効用に否定的な人(思ったことがないという回答者)は、自動車による貨物輸送をもっと減らした方がよいという意見(自動車による貨物輸送に悪意的)が段々と多くなっている(28%→38%→45%)ために、そう考えるのはもっともでしょう。

しかし、私は、ここでふとしたことから道草をくってしまった。この稿の主題も、全くその道草が内容となるものです。

引用したクロス表を眺めてみると、普通には上述のように考えるものでしょう。そのときに、当人は認識しているか否かにかかわらず当然のように前提としていることは、自動車による貨物輸送に好意的にせよ悪意的にせよ、いずれかの方向で人々の反応を当人が評価しているということです。どちらの方向にせよ自動車による貨物輸送をもっと減らした方がよい、現在の程度でよい、もっと増やした方がよいという三つの反応は、話を簡単にするために、等間隔で並んでいるものとしましょう。本当は、このこと自体が一つの問題であります。今は問わないことにします。そうすると、この三つの反応(カテゴリー)に、それぞれ1, 2, 3という数字を与えると、この三つのカテゴリーが、数字の大きさの順で一定の方向をもっていることが表わせることになります。好意的な人は、3という数字(増やした方がよい)が多いことを好むでしょうし、悪意的な人は、1という数字(減らした方がよい)が多いことを好むでしょう。

う。

ですから、例えば好意的な人についてだけ眺めてみると(1, 48%)、(2, 55%)、(3, 69%)

という数字の組合せ(前出クロス表参照)がどの程度の相関関係にあるかを計算すれば、これを数量的に表わしたことになります。この三組の数字の組合せの相関係数 r は、 $r=0.982$ と、正の完全相関である1にほぼ近い数字を示しました。我々はこうして数字の形に表わさないでも、先に掲げたクロス表から以上のような鮮明な関係を直観的(?)に把握してしまいます。

3. しかし、私が横道にそれたというのは、これらのクロス表から、0.982以外の別の相関係数も算出できるのではないかという疑いをもったからです。そもそもそのような疑念にとらえられたのは、一つの先例が私の念頭にあったからです。

それは、試験的な社会意識調査でした。家庭生活や暮しむき、生き甲斐等の様々な価値について、人間の幸福感を探るために30問くらいの質問を投げかけて、それに対する回答を(非常にそう感ずる)、(まあそう感ずる)、(どちらでもない)、(あまりそうは感じない)、(全然そうは感じない)、(わからない)の6個のカテゴリー、実質的には、(わからない)を除外すると5個のカテゴリーに区分して受入れられる構造の調査でした。

そうして、(非常にそう感ずる)から(全然そうは感じない)までに、1から5までの数字を与えた上で、30問あまりの質問の任意の2つの質問についての相関係数を算出して、これらのあらゆる2つの質問の組合せについての相関係数の集合(相関係数行列)から分析の基礎を行なったものでした。本稿はこの調査の紹介

が目的ではありませんから余談は避けませんが、要するにこの調査は、人間の幸福感を規定している要因が、主としてどのような項目によって左右されるものであるかを探り出そうとする意図に出たものであったことだけは、ご紹介しておく必要があると思います。この調査に限らず、重回帰分析、主成分分析、因子分析などいわゆる多変量解析を行なう場合の基礎的な分析手法としては、ある定性的または定量的反応と別のそれらの反応との相互関係をまずもって相関係数に表わして、これらの相関係数の行列から一定の発見をすることが行なわれるわけです。問題はその出発点となる相関係数の求め方ですが、この調査では、初めから因子分析の意図をもっていたために、どの質問の受け皿も、(非常にそう感ずる)から(どちらでもない)を経て(全然そうは感じない)までの肯定、否定の程度を明瞭に示す構造として用意されていました。従って例えば、Q1(自分が健康で元気に暮らしていて幸せだと感じますか)、Q2(やりたいことがまあまあできて毎日が充実していると感じますか)との質問についての相関関係は、回答者2,473人のQ1とQ2についてのクロス表を次表のとおり作成して

(表 1)

Q1 \ Q2	1	2	3	4	5
5	-	1	6	20	30
4	3	8	46	135	19
3	13	92	583	295	18
2	13	215	297	102	8
1	154	147	191	69	8

相関係数は、この2,473個の数字の組合せ、即ち(1, 1)が154個、(1, 2)が13個……(5, 5)が30個について求めるというものでした。この計算結果は、相関係数 $r=0.498$ でした。

4. さて、そこで、私が横道にそれた事情というのも少しづつ判って頂けたと思いますが、前項で述べた相関係数 $r=0.498$ を求める前提の数字の取扱い方と、その前に述べた $r=0.982$ を求めた時の数字の取扱い方が同一ではないということが、私が最初につき当たった疑問だったわけです。

もし、 $r=0.498$ のような求め方を、道路の効用の質問と自動車による貨物輸送の増減についての質問との相関関係にあてはめるなら、相関係数を求めるための数量の与え方としては、例えば次のようなものが考えられます。(表2)(わからない)と答えた者を除く回答者1,833人の質問1と質問2への回答状況に応じて数字を与えます。質問1については、生鮮食品輸送に関する道路整備の効用について考えたことのある者は1、(ない)者は2の数字を、質問2については、自動車による貨物輸送を(もっと増やす)という意見の者には1、(現在程度でよい)という意見の者には2、(もっと減らす)という意見の者には3の数字をそれぞれ与えて、(1, 1)という組合せの人が219人、(1, 2)という組合せの人が638人、……(2, 3)の組合せの人が212人という計1,833個の数字の組合せについて、相関係数を計算します。そうすれば、 $r=0.498$ の算出と同じ方法による相関係数の算出の仕方になります。

この計算結果は、相関係数 $r=0.124$ という

(表 2)

質問2 自動車による貨物輸送は	質問1	質問1			計
		1 (あり)	2 (なし)	3 (わからない)	
わからない	4	248	209	96	553
もっと減らす	3	229人	212人	32	473
現在程度	2	638人	447人	79	1,164
もっと増やす	1	219人	88人	10	317
計		1,334	956	217	2,507人

極めて低いもので、これは質問1と質問2との回答結果には殆んど相関関係が認められないという数値になっています。ちなみに、(わからない)という回答に3(質問1)及び4(質問2)という数字を与えた2,507人全員についての相関係数も $r=0.172$ と大差ありません。どうして直観的(?)な判断である相関関係の強さ(数値的には $r=0.982$)とこうも極端にちがってくるのでしょうか。

5. 3個の数字の組合せ((1, 48)、(2, 55)、(3, 69))と1,833個の数字の組合せ((1, 1)が219個、(1, 2)が638個、……(2, 3)が212個)とでは、計算の基礎が異なるのだから結果が異なるのは当然だという説明は、現象を示しているだけであって原因の指摘とはいえないでしょう。

問題は、どうしてこのような異なる数字の組合せ方を与えるのかという点にあります。この問題について示唆に富むのは「市場調査の計画と実際」という本の中の次のような指摘です。

・数量は与えるもの
さて、数量はそのもの自身に内在するものではないことにまず注意されたいのである。ここを

つきりさせるために、測定したものを「アイテムにおけるカテゴリーへの反応という形であらわす」ことにしたのである。数量は、われわれが目的に応じて与えるものなのである。目的とは、われわれがそれを知ることによって、「妥当性ある結果を獲得できるようにすること」を意味する。数量を与えて解析して得た結果がわれわれの行為に、判断に科学的意味で役立つようにしなければならないのである。逆にそのように数量を与えて、解析をしなければならないのである。数量を与えることは、一番はじめにすべきことであるが、これは最終の結果を見とおし、そこから漕ぎ上がって、考えなければならないことである。数量は、このように目的に応じて、相対的に与えられるべきものであるから、目的が変われば、たとえ同一のもの——測定されたもの、アイテムにおけるあるカテゴリーへの反応——であっても、与えられるべき数値は変わってくるのである。(後略)

・統計的分析・数量化における2つの立場

これは、寺田寅彦の言葉を借りれば「最初の言葉」「最後の言葉」に相当するものである。1つは最初のさぐりを入れる消息子の働きをするものであり、一方は直接行動の指針、その効果を直接的にさすもの(これそのものが最後の結末を与えるもの)であるといえよう。この2つは、はなはだ異なる機能をもつものである、それぞれの特色があるので、取り違えては大変な誤りとなってしまう。統計分析の例でいえば、前者は因子分析法などに見られ後者は回帰分析法などに見られる。

別の例でもう少し説明してみよう。

病気の分類すなわち病名を与えることは最初の言葉となる。病名を与えることは、とり上げられたいいくつかの計測値の似たものをあつめて分類することに対応する。

(中略)

一方、第二の立場をのべた「最後の言葉」のほうは、ある計測値を示すものに対して、甲という治療法がよい結果を与えるか、あるいはよい結果を与えないか、これこれの計測値パターンをもつものは予後がよい、あるいは予後が悪いか、などをいおうとするときに用いられるものである。これは病名の決定というものではなく「どうすればよいか」をさし示すものとなるのである。したがって、よい結果を与えるかどうかという「はっきりした」ことを可能な限る精度高く予測することが目的となるので、どうすれば、精度よい予測ができるかどのくらいの的中率があるかが示される方法を考えることになるのである。

立ち帰って因子分析について説明してみよう。この方法はある仮説によって得られたデータを分類し、説明を得ようとするのが目的となるので「最初の言葉」を与える方法となるのである。そこで得られたデータを解釈することが非常に必要となるのである。しかし解釈そのものはデータからそのまま得られるものではなく、われわれが因子分析によって位置づけられたパターンから、仮説をたてつつ考えるものなのである。このため解釈は一義的にきまらないのである。いくら解釈をしても、これは「文学」である。科学的論述の結末ではないのである。

解釈が壮麗に書きあげられた論文をよくみるが、これは1つの想像フィクションなのである。これをたしかめるためには、得られた解釈をもとにして、再び調査・実験を組み目的に対する検証(これには「最後の言葉」による方法が用いられる)をしなくてはならない。

市場調査の計画と実際(昭和39年8月、日刊工業新聞社刊、林知己夫、村山孝喜共著)中の付録「数量化理論」(300頁、301頁、303頁)から抜粋

ここでは、因子分析や主成分分析の結果の解

釈は、一義的にはきまらないで、あくまで一つの「文学」であるとされています。例えば、仮りに7割の増悪と2割の脅威と1割の親愛とをつきまぜた感情を何と称するかと言うときに、それを「嫉妬」と解釈するのは解釈する者の自由だというわけです。

(注、この例示では、理解の便宜上、説明要因に「増悪」「脅威」「親愛」という数量化出来ないものを引用しましたが、因子分析や主成分分析は、説明要因として数量で表示できるものについて行ないます(逆に厳密な意味では、分類要因も、数量要因も、ある項目(アイテム)へのある範疇(カテゴリー)への反応であるとされます=前掲「市場調査の計画と実際」298頁)から、例示の場合はむしろ数量化第3類の分析手法を用いるべき事例でしょう。)

6. さて、道路整備が生鮮食料品の輸送に効果があると考える者は、そんなことは特に想起しない者よりも、自動車による貨物輸送を増やした方がよいとする者が明らかに多い。……このように主張する人がいます。その人は、前掲のクロス集計表や、同表から道路整備の効果を認識している者について算出される一つの相関係数0.982(ちなみに、認識していない者についても同様の方法で相関係数を算出すると0.995と大差がありません)を引用して、その主張に客観的な科学的根拠があると言うでしょう。

では、もう一方の相関係数0.124を主張する側には正当性はないのでしょうか。考えようによってはこちらの方が、百分率表示の前掲クロス集計表よりもなまの回答者実数を両質問でクロスしているのですから、一方向(横向き)にのみ100%としている百分率表示よりも正確であるとも言えますし、また、相関係数の算出も、道路整備の効用を認識している人のみではなく認識していない人の反応も含めて行なわれ

ているのです。

そこで、これら二つのどちらの判断に軍配を上げるべきかということが問題になります。二つの判断は、いずれも科学的なもってもらしい根拠のもとに主張されていますが、その結論は、正反対といえるほど対立しているのです。

7. 私のもう1つの試みは、二つの要因が相互に独立であるか、何らかの関連性がみとめられるかについてカイ平方(χ^2)の値を算出して、いわゆる独立性の検定を行ってみることでした。末尾に算出課程を示しましたが、 χ^2 の値は104.872と異常ともいえる程大きく、二つの要因が独立であるとは言えない(つまり、何らかの相関関係がある)と判断しても、0.5%以下の危険(1000回のうち5回判断を誤る危険)しかおかないことがわかりました。

これとて、教科書どおりの手法を機械的に適用したものにすぎず、その根源的な理解は、私には必ずしも十分ではありませんが、要は、二つの要因が独立と考えるには極めて異例の数字であり、例示の二要因は相互に高い相関が認められるという結論になります。

8. それでは、相関係数から論じていた二つの試算結果の対立は、 $r=0.982$ の方に軍配が上ったといえるのでしょうか。

このような問題に対するある程度の示唆は、すでに一つの文献から引用したところですが、それに若干の私の意見も加えて一応の結論に近づくことにします。

「市場調査の計画と実際」は、「これは、いくら解釈をしても文学である。科学的論述の結末ではないのである」と明解に断案を下します。この命題は先にふれたように、因子分析や主成

分分析について、その結果の解釈は全くの「文学」であり、表現方法の工夫であるというものです。が、(3)及び(4)で述べたような、ある意見分布の解釈にあたって特定の反応カテゴリーに数字を附与する場合にも、同様の意味合いにおいてあてはまる指摘であると私は考えます。

電子計算機は、大量の計算を短時間でこなしてくれます。しかし人間の思考作用の大きな特徴である創造性ないし原始性(originality)を計算機に求めることは不可能です。数量は、まさに人間が与えるものであり、与える人間が考えることだからです。

$r=0.982$ が算出された僅か三個の数字の組合せ1つを例にとっても、(もっと減らした方がよい)、(現在の程度でよい)、(もっと増やした方がよい)というカテゴリーに応じて、(1, 48)、(2, 55)、(3, 69)と並べる必然性はないとすら言えなくもないのです。疑問はいろいろな方向から出てくるに違いありません。私が思いついたものとしては、前記のように、何故道路整備の効用を認識している人についてのみ数字をとり出すのか、つまり、認識していない人とのからみをどう考えるのかという点、また、自動車輸送の増減の反応ごとの回答者数の違いは、百分率表示では考慮されないという点でした。更に、1, 2, 3という等間隔かつ等比ではなく、数字の与え方に例えば、1, 5, 6というような重みを与えることは考えなくてもよいであろうかという問題もあります。この問題に対応して、分析手段を準備するための尺度のとり方が、決定を迫られるわけです。

サーストンやリッカート、ガットマン等が尺度の値の求め方についていろいろの考案をしているのもこのせいです。尺度論は、こうして尺度の客観化の方法を模索します。

9. ところで、私は、 $r=0.982$ と $r=0.124$ とのいずれを採るべきか、結論を急がなければなりません。私の考え至らない色々の要素があるかも知れないと知り乍ら、乱暴でもここで中継地点として一応の結論を置いた方が、その次の考察に新しい出発点を与えるかも知れません。そこで私は、敢えて、ここで $r=0.982$ の方向（どちらかと言えば相関関係が強いと判断する立場）をとることにします。

というのも、前記のクロス集計表を見て私が最初に理解したことは、現在の私の常識からすれば、二つの要因に明らかに高い相関関係を認めるものだったからです。ですから、その後に行った考察において、この最初の私の常識をくつがえすものは、例えば $r=0.124$ であれば、すでにそのことが私の疑惑の対象となるばかりで、その点、検証の意味は最初から与えられていない考察になるわけで、まるで1つのトートロジー（同語反覆）にすぎないのです。

しかし考えてみれば、「私の常識」とは奇妙な言葉であり、それ自体では説得性に乏しい理屈です。私だけがそう考えるなら、「私の認識」にすぎないものであって、人を納得させることは出来ません。私だけでなく、大多数の人々がそう考えるなら、そういう意味での常識として提起した問題の結論は、二要因の間には、相関関係強しということになります。そうして、その場合には、単に「常識」であって「私の常識」とは言葉の矛盾になります。

10. それでは、二要因の相関が与えられたデータに照らして高いか低いかの判断を、単に多数決で決定してよいものでしょうか。「常識」なるものを、数の上で大多数の人々の考えることと定義するならば、あるいはそれでもよいかも知

れません。しかし、そこには、人間はいなくて、生命のない多数の論理があるにすぎません。そもそもそのような意味での常識は、何らの手続きも経ないで、忽然として、かつ自明のこととして存在するものでしょうか。私の答えは否です。

例えば、本稿で提起した二要因の相関関係をどう見るかについて、一般的な意味での常識が当然のこととして存在する筈はありえないのです。ここで始めて、人間の真骨頂としての「思考」がその真価である個性ないし独創性を発揮する場が与えられます。まず最初に私の認識ないし判断が求められました。そうして私は、改めて「私の常識」という漫然とした認識に、脅威を与える何物かの存在を意識せざるを得ませんでした。残念乍ら $r=0.982$ の方が、 $r=0.124$ の方にくらべて正当であると断言する自信がないのです。それに、私自身、理論的には、どちらかといえば $r=0.124$ の方が精緻に出来ているように思います。このように、私の思考はここに至って大いに動揺します。そこでは生命のない数の論理、多数が正しいという考え方は精彩を失います。

11. 事柄一般の認識や判断は、その人の到達している人間としての水準においてでしかなされず、また、ある事象の解釈は、解釈する人の人格において行われるものであるという一般的命題があります。私の場合もその例外ではありえず、「私の常識」的理解は、二つの要因の間に相当高い相関関係を認める立場に魅かれたまま、なお残念乍ら「思考」作用に至っていません。

今、私は、それに自ら疑問を提出するところまで進みました。いずれの考え方が正しいかという事は絶対者の決めることでしょう。むしろ、そのことよりも、以上の経過の中から例えば問題の相関関係については、両者の論理を一段止揚する論理、一般的には、より高い思考や思想に対して我々は虚心でなければならぬという教訓を抜き出すことが大切だと私は考えます。

私が、「私の常識」という不用意であり、かつ結果的に皮肉な表現をしたのも、今私なりに整理してみれば、「常識」なるものが人間の真の思考に立脚してほしいこと、そのためには、一人一人の「私」が考えることから始まるとい

ろ、そのことよりも、以上の経過の中から例えば問題の相関関係については、両者の論理を一段止揚する論理、一般的には、より高い思考や思想に対して我々は虚心でなければならぬという教訓を抜き出すことが大切だと私は考えます。

附. [質問1と質問2における独立性の検定]

$$f_{ij}$$

A \ B	あ り	な し	不 明	計
ふ や す	219	88	10	317
現 在	638	447	79	1,164
へ ら す	229	212	32	473
不 明	248	209	96	553
計	1,334	956	217	2,507

f'_{ij} (理論値)

A \ B	あ り	な し	不 明	計
ふ や す	169	121	27	317
現 在	619	444	101	1,164
へ ら す	252	180	41	473
不 明	294	211	48	553
計	1,334	956	217	2,507

$$\sum_{i=1}^h \sum_{j=1}^l \frac{(f_{ij} - f'_{ij})^2}{f'_{ij}} = \chi^2$$

$$\phi = (h-1)(l-1) = 6 \text{ (自由度)}$$

の χ^2 分布に従う。

A = 自動車による貨物輸送 (質問2)

B = 生鮮食品と道路との関係 (質問1)

但し、 $f'_{ij} = f_i \cdot f_j / n$

$$(f_{ij} - f'_{ij})^2 / f'_{ij} \rightarrow \chi^2$$

	あ り	な し	不 明	計
	14.793	9.000	10.704	
	0.583	0.020	4.792	
	2.099	5.689	1.976	
	7.197	0.019	48.000	
				$\chi^2_{\alpha} = 10.4872$

$\phi=6$ における χ^2 の値は $\alpha=0.005$ で $\chi^2=18.55$ 、 $\alpha=0.01$ で $\chi^2=16.81$ 、 $\alpha=0.05$ で $\chi^2=12.59$ と、いずれも $\chi^2_{\alpha}=10.4872$ をはるかに超えるので、仮に、危険率(α)0.5%でも「仮説：要因Aと要因Bは独立である」は棄却される。

うことを強調したいためと考えることが出来ます。

そうして、データの解釈・数量化に当たっても既成の権威の検証を、人間本来の特技である「思考」を活かして行ってゆきたいと考えるものです。

そうは申しましたが、ついに私は、 $r=0.124$ を棄却したいと思いつつ、何故棄却できるのかわからず終いでした。近い将来にこの壁のとれることを願いつつ筆をおきます。