

デルファイ予測の分析手法と実施例

Methods for Measuring the Consensus in the Delphi Forecasts

江藤 肇* Hajime Etô

This paper proposes statistical methods to measure the consensus among the experts in the Delphi, and reports their applications to the technology and social forecasts. No statistical method has been applied to the Delphi mainly because of two gaps; one between the subjectivity of the Delphi opinions and the objectivity of the technological development processes, and the other between the distributed opinions and a single event. However, the fact that hundreds of experts in influential positions take part in the Japanese Delphi forecasts allows the Delphi opinions to be considered as one of the leading factors in making a national decision. The consensus among experts is interpreted as reflecting the inter-subjective probability of the national decision, and is measured with respect to between-group similarities.

1 緒 言

デルファイ法は技術予測のために開発された手法で、たとえば技術専門家に、ある技術の実現時期を質問し、その回答集計結果を回答者に示し、再び同一質問をすることによって回答者に意見修正の機会を与えるものである。すなわち、このプロセスを数回反復し、意見収束を得るコンセンサス（意見一致）獲得手法と解されよう。アメリカでは1960年代はじめにデルファイ法による技術予測を実行し、月に人類を送り込めるのは1970年ごろという予測結果を得ていたと言われる。わが国では科学技術庁が昭和45年に、紀元2,000年までの技術予測を試みたのをはじめ、その後は大小合わせて毎年20回以上引き続き行なわれているものと推測される。

しかし、従来デルファイ回答集計結果を定量的に分析することはきわめて困難と思われてきた。その第一の理由は、技術予測の求めるものは客観事象としての技術実現であるのに対し、デルファイ法の得るものは主観的意見であるというギャップである。第二の理由は、技術予測の求めるものは一回的事象としての技術実現であるのに対し、デルファイ法の得るものはばらつきのある意見であるというギャップである。たとえば、平均という初等的統計量でさえも、実現時期に関する意見の平均という概念は意味不明であるとの理由から、捨て去られている。

しかし、デルファイ主宰者が官庁など政策決定・実行機関であり、回答者が権威ある専門家・オピニオンリーダーとして政策に強い影響力を持つほか、関係機関指導者（研究所長、学部長など）として実行力もある場合、その回答集計結果は政策決定の一要因をなすと考えられる。すなわち、デルファイ結果は主観的意見であるが、客観事象の要因、または先行事象として客観的意味を有する。

このようにデルファイ結果を参考データに政策が決定されるとき、これらオピニオンリーダーの意見一致度（コンセンサス度）、特に各界代表意見の一致が重視される。一致度は分散の小ささにより、代表意見はたとえば平均意見により表

現されよう。政策決定行動においては、意見の平均や分散は有意義な統計量と言える。

また、技術開発・社会開発計画において重視されるのは波及効果である。そのため、回答パターンを分析して波及効果を推定する方法を提案する。

2 政策決定要因としてのデルファイ法

デルファイ法がもし新聞社など第三者的機関により主宰されるならば、その予測が適中するか否かに価値が求められる。そのとき技術実現という客観事象は、専門家の意見から独立したものと考えられる。しかし政府・行政機関が政策決定資料の目的のため、デルファイ調査を行なう場合、もしある技術課題が計画年内に実現できると各界専門家の意見が一致すれば、それがプロジェクトに採用される可能性が強くなる。そして採用されれば十分な資源を得て、実現される可能性が強くなる。またオピニオンリーダーである専門家の意見が一致すれば、国民のコンセンサスも得やすく、プロジェクトメンバーの士気も高揚され、プロジェクトが成功する可能性が強くなる。このように、デルファイ調査回答結果は、技術実現の先行要因としての客観的意味をもつ。天気予報において、求めるものは晴雨、気温であるのに、気圧などの先行要因を分析するのと同じである。

一回的事象である技術実現に、ばらつきのある意見を対応させるのも、後者がコンセンサスの程度を表現すると考えれば全く自然である。事実、専門家の意見一致度（コンセンサス度）を、1回の事象の主観確率と考えることができる。たとえば、アメリカにおける天気予報が、1回の事象である明日の天気確率（主観確率）を与えるのと同じである。

政策決定プロセスにおいて、各界代表意見として多数意見や平均意見が重視される。その理由は、現代の社会は多数決制を原理としており、また平均という概念は公平妥当という概念に近いものと考えられているからである。すなわち、政

* 日立製作所システム開発研究所

策決定や入試選抜においては、公平妥当の目的から平均概念を用いていることが多い。意見や点数の平均は論理的な意味はもたないにしても、行動論的意味は深いものがあると言える。

各界代表意見が不一致の場合でも、それが対立を招来する可能性が弱ければ実行できるかもしれず、対立を招来する可能性が強ければ必要なことも避けるかもしれない。コンセンサス度は、代表意見の一致度だけでなく対立の可能性の程度をも含むことが望ましい。

政策にプライオリティをつけるうえで重要な要因は、波及効果である。経済の分野では産業連関分析を用いるが、技術開発や社会開発では、これに似たクロスインパクト表(図1参照)を用いることがある。政策*i*を実行した場合、政策*j*が促進されるならば、(*i*, *j*)の位置に↑印を入れ、抑止されるとき↓印を入れる。この表を得るには、これ自身をデルファイ回答してもらうのが望ましい。しかしデルファイ法は、専門家のベストメンバーを回答者に選ぶため、特定者に毎回大きな負担をかける結果になりやすい。この負担を軽減するため、他の質問への回答からクロスインパクトを推定する方法を提案する。さらに進んで、適切な波及効果をもたらすポリシィミックス(複数政策の組合せ)を求める方法を提案する。

3 コンセンサスの計量

3.1 代表意見の一致

代表意見の有力なものは、多数意見、中間意見および平均意見であろうが、これらすべてが同じ選択肢(し)にあれば、その階層内のコンセンサスが得られていると言える。二つの階層の間でもそれが一致していれば、両階層間のコンセンサスが得られていると言える。これは定性的評価に近い。

3.2 代表意見の一致度

二つの階層間の意見一致度を定量的に表現するには、両階層の意見分布の相関比の概念を用いるのが最良と思われる。これは、コンセンサス度を*C*、オピニオンギャップ度を*g*とし、

$$C = 1 - g \dots\dots\dots(1)$$

$$g^2 = N_1 N_2 (a_1 - a_2)^2 / \{(N_1 + N_2) S^2\} \dots\dots\dots(2)$$

とする。

ここに、*N*₁, *N*₂: 階層1, 階層2の人数, *a*₁, *a*₂は階層1, 階層2の平均意見, *S*²は総分散

(2)式は、二階層の平均意見のずれを人数や総分散で正規化したものとも解釈できる。この正規化の方法は、平均意見については相関比という深い意味をもつ。多数意見や中間意見を代表意見とするときは、もっと簡単に、

$$g^2 = (r_1 - r_2)^2 / m^2 \dots\dots\dots(3)$$

としてよい。

図1 クロスインパクトマトリックスの例

政策*i*を実行したとき政策*j*が促進されるならば、上から*i*番め、左から*j*番めの欄に↑、抑止されるならば↓を記入する。

Fig. 1 Cross Impact Matrix

項目	1	2	3	4	5
1		↑	↑	↓	↓
2	↑		↓	↓	↑
3	↑	↓		↑	↑
4	↓	↓	↑		↓
5	↓	↑	↑	↓	

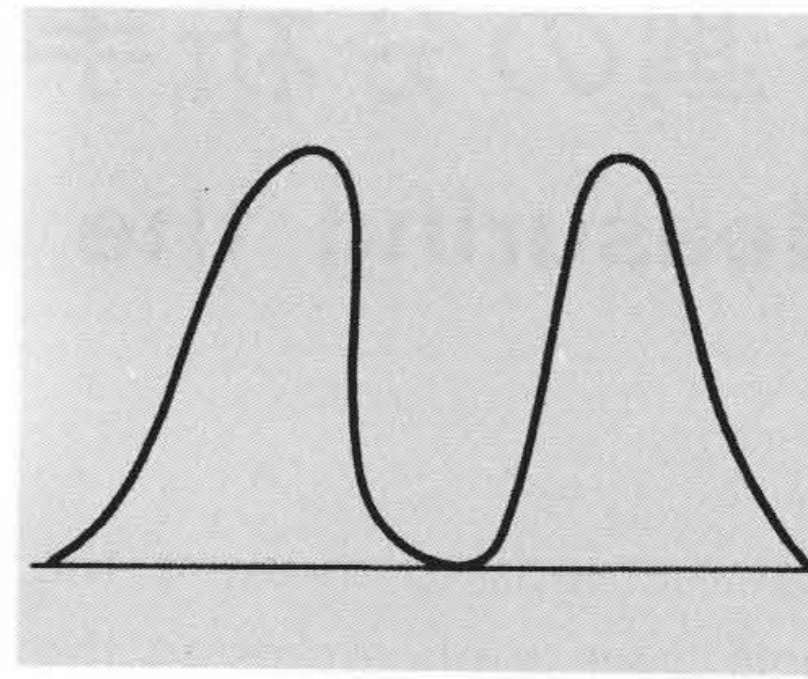


図2 分離した分布 二つの分布形(山)に共通部分がない。

Fig. 2 Separated Distributions

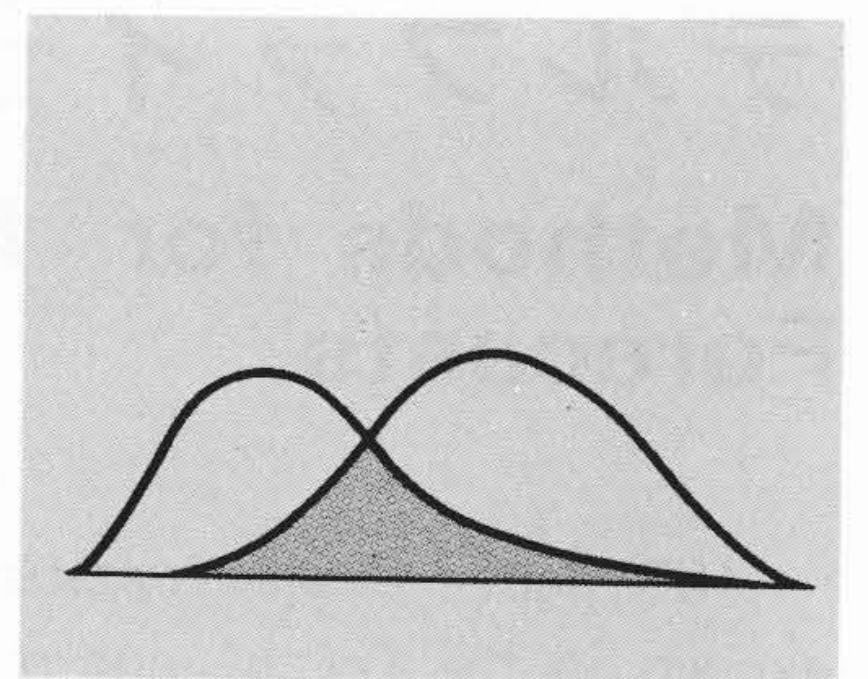


図3 相互浸透した分布 二つの分布形(山)が入り組んでいる。

Fig. 3 Penetrating Distributions

ここに、*r*₁, *r*₂は階層1, 階層2の代表意見, *m*は選択肢の数。労働省が経営者代表・労働組合代表を対象にしたコンセンサスデルファイ(1972年)で、(1)式および(2)式による計量を行ない、定性的にも合理的な結果を得ることができた。

3.3 意見相互浸透度

もし二階層の意見が完全に分離していれば(図2参照)、両階層は相互に理解できず、たとえ平均意見や多数意見が近くとも、わずかの不一致が対立をもたらす可能性が強い。図3のように、両階層の意見分布が相互に入り込んでいるとき、相互浸透部分(図3の濃い部分)が仲介者の役割を果たして、対立に至らない可能性が強い。特に両階層がよく組織されていて、中間部分が脱退して第三勢力を作る可能性がないとき、仲介者の役割は大きい。たとえば、地方公共団体間の関係、労使間の関係などがそうである。

労働省の労使間コンセンサスデルファイでは、3.2で述べた代表意見間一致が高いときは、この意見相互浸透性もかなり高くなることがわかった。科学技術庁技術予測デルファイでは、階層が組織化されていないので意見相互浸透性を特に計量しなかった。なお、これは判別分析と逆の概念と言える。

3.4 意見分布の類似度

代表意見だけでなく、少数意見、例外意見も含めてコンセンサスを考える。公害をはじめ、現代の社会的病根は早くから少数者によって指摘されながら、少数意見として放置され最後に一挙に爆発したものである。たとえば、水利権や廃棄物処理をめぐる地方公共団体間の対立も、最初は少数者によって問題点が指摘され、多数者は問題が顕在化してから感情的対立にまで至ったのである。多数意見や平均意見はそれが一致していても、対立の芽は潜在しているかも知れない。反対に、代表意見は必ずしも一致せずとも、意見分布状況が近似していれば両階層は相互に理解され、対立に至らぬ可能性が強いと期待される。もし分布が一致していれば、そもそも階層区分したこと自体が無意味であったと判断される。たとえば技術予測につき、職業別、専門別、年代別に意見差があると考えることが無意味かもしれない。

分布形の一致性についての従来理論は、分布の山の高い所、すなわち多数部分に焦点を合わせていた。分布の裾(すそ)、すなわち少数部分を扱うときは分布が広範囲に及び、一つ一つの出現に順序がつけられる場合に限られていた。しかし、政策選択においては広範囲にわたる順序をつけることは困難で、少数の選択肢から回答してもらうことが多い。この場合、多数部分と少数部分をまとめて比較するための方法を新たに提案する。多数意見、少数意見の差を消すため、図4の階層差*d_j*を(4)式により正規化した量*e_j*を扱う。図4で*d_j*の大きさ

は矢印の長さにより、符号は矢印の向きにより表現される。 d_j は多数意見選択肢で大きくなるが、 e_j はそうでない。

$$e_j = d_j / (f_{1j} + f_{2j}) \dots\dots\dots(4)$$

ここに図4が示すように、

j は選択肢、 f_{1j} と f_{2j} は選択肢 j への階層1と階層2の回答数(正規化済み)

e_j の絶対値は0から1までの間の実数である。選択肢の数を m とすれば、オピニオンギャップ g は、

$$g = \sum_{j=1}^m e_j^2 / m \dots\dots\dots(5)$$

で表現される。コンセンサス度 C は(1)式で得られる。 g も、したがって C も0と1の間の実数である。複数政策にコンセンサス度に関する順位をつけるにはこれでよい。

明確な定性的意味をもつようにコンセンサス度を定義したいときは、「階層分類することの無意味さ」、「両階層が同質であることの確からしさ」、「両階層が同一集団のランダム分割と仮定することの信頼度」を計量する。このとき、

$$C = \frac{\sum (F_1 \dots F_m)}{(f_{11} \dots f_{1m}) \sum R} \left(\frac{F_1}{f_{11}} \right) \left(\frac{F_2}{f_{12}} \right) \dots \left(\frac{F_m}{f_{1m}} \right) \bigg/ \left(\frac{N}{N_1} \right) \dots\dots(6)$$

ここで、

$$F_j = f_{1j} + f_{2j} \quad (j = 1, \dots, m)$$

$$R = \{ (f_{11}, \dots, f_{1m}) \mid \sum_{j=1}^m e_j^2 > \sum_{j=1}^m \underline{e}_j^2 \}$$

$$\underline{e}_j : e_j \text{の実現値} ((4) \text{式参照})$$

(6)式を用いる方法は、概念としては正確だが、和の範囲 R を求める手順が複雑である。

もしデルファイ回答が、不特定多数の母集団からのランダムサンプルと考えられるならば、

$$X^2 = \sum_{j=1}^m \{ d_j^2 / (f_{1j} + f_{2j}) \} \dots\dots\dots(7)$$

として、 X^2 分布または F 分布に属する量にすることができる。 F 分布にするときは、

$$F' = X^2 / (m - 1)$$

が、自由度 $m - 1$ 、 ∞ の F 分布に属することを利用する。しかしデルファイ法は専門家につき全数調査したものであり、また専門家の意見は安定している(各ラウンドにおいて)と前提して活用されるものである。たとえば選挙結果もサンプルと考えず、全数調査であり、次の選挙までは安定した民意と考えるのが制度となっている。選挙結果を変数と考えるのはクリティクの場合にのみ許される。デルファイも、これをクリティクするときには(7)式を用いる計量法が有用であろう。

しかしデルファイ結果を尊重し活用するときには、(5)式または(6)式を用いる方法がよいと思われる。

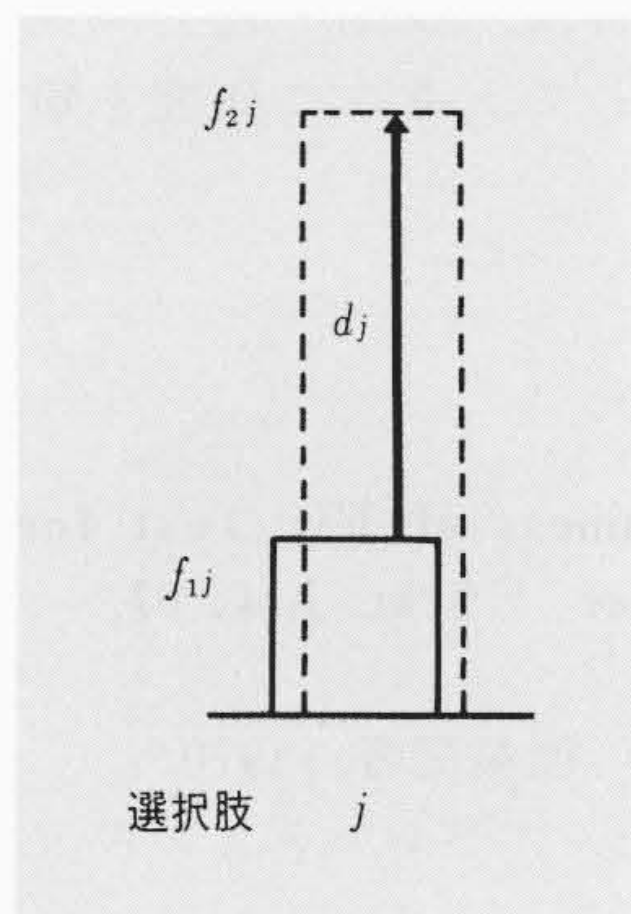


図4 階層差の定義
選択肢 j における二つの分布の差を d_j とする。

Fig. 4 Strata Gap

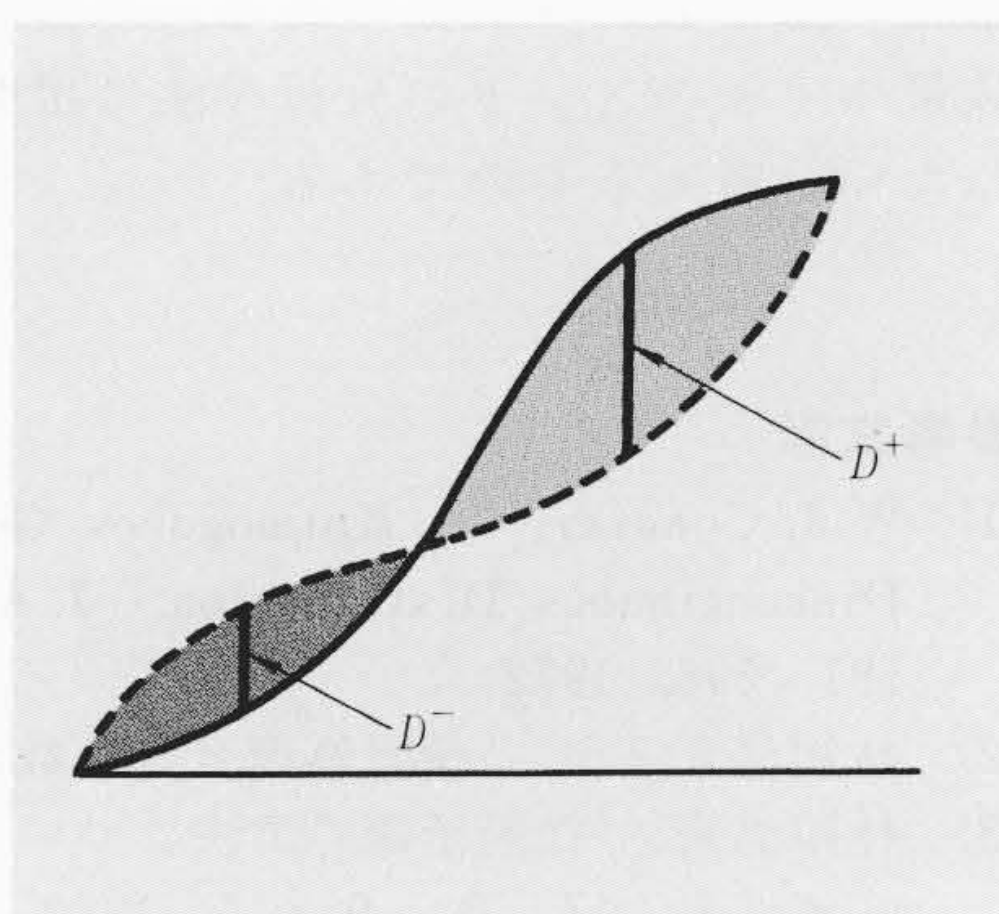


図5 累積相違度 二つの累積分布の最大差(コルモゴロフ統計量)を示す。

Fig. 5 Kolmogorov's Statistics

少数意見選択肢と多数意見選択肢の差が少ないとき、たとえばたが同じときは、コルモゴロフ・スミルノフ統計量を用いて、両階層意見分布の差を求めてもよい。それは(図5)

$$\left. \begin{aligned} D &= D^+ + D^- \\ D^+ &= \frac{1}{m} \sum_{j=1}^m \sum_{i=1}^j (f_{1i} - f_{2i}) \\ D^- &= \frac{1}{m} \sum_{j=1}^m \sum_{i=j+1}^m (f_{2i} - f_{1i}) \end{aligned} \right\} \dots\dots\dots(8)$$

を定義し、離散の場合の信頼度計算法⁽¹⁾を用い、信頼度をコンセンサス度とする。

選択肢の数が多い場合は、サンプル $j = 1, \dots, m$ につき f_{1j} と f_{2j} の間の相関係数を求め、これをコンセンサス度とすればよく、この場合は回答数を正規化しなくてよい。

労働省のコンセンサス デルファイには(6)式を用いる方法を応用してみた。3.2および3.3による計量法では、学識経験者は経営者代表に近く、(6)式による計量法では学識経験者は労働組合代表に近かった。学識経験者が主要指標においては経営者代表に近いにもかかわらず、労働組合代表との間に対立を起こさず、かつ中立者としての機能を発揮できる理由はここにあると思われる。

3.5 意見体系の類似度

3.4までは個々の質問項目ごとにコンセンサスを計量していた。一方、たとえば公害防止技術につき一連の質問項目があり、これについての意見を総合的に考察して意見一致度を計量するときには、回答パターンをベクトルで表現し、ベクトル相関係数を用いるとよい。以下の議論では、選択肢の数がすべての質問項目に同一であると前提する。もし同一でないときは適当にカテゴライズする。通常の技術実現時期に関するデルファイでは、この前提が成立している。

選択肢数を m 、質問項目数を n とし、階層 k の質問項目 i への回答頻(ひん)度を m 次元ベクトル

$$(h_{i1k}, h_{i2k}, \dots, h_{imk}) = H_{i \cdot k} \quad (i = 1, \dots, n)$$

で表現する。階層1と階層2の回答パターンベクトル $H_{i \cdot 1}$ と $H_{i \cdot 2}$ の間にベクトル相関係数を計算し($i = 1, 2, \dots, n$ 標本数 n)、一連の質問項目への意見体系の類似度とする。ベクトル相関係数 $r_{1, 2}$ は次の(2)式で得られる($0 \leq r_{1, 2} \leq 1$)。

$$r_{1, 2}^2 = 1 - \det \Sigma_{1, 2} / \det \Sigma_1 \det \Sigma_2 \dots\dots\dots(9)$$

ここで、 Σ :分散共分散行列

$\det$:行列式の値

回答頻度の代わりに、中位値と中位値から四分値への差から成る三次元ベクトルをとってもよい。中位値から二つの四分値への差は楽観側、悲観測へのばらつきを表わしている。ベクトル相関係数は類似性を求めるというデータ解析の目的に即した手法である。ベクトル相関係数の代わりに正準相関係数を用いてもよい。正準相関係数は数学的に興味深い概念で、計算サブルーチンが普及しており便利である。

科学技術庁技術予測デルファイに上記三次元ベクトルのベクトル相関数を計量した。社会開発部門の公害防止技術については、国公立研究所員が他のすべての階層に類似した回答パターンをもっている。国公立研究所員は研究者・技術者として大学の研究者や企業技術者に近く、公務員として行政者に近いと思われる。各界専門家を対象とする大規模デルファイを頻繁に行なうことはできないが、行政機関が国公立研究所員を対象に小規模デルファイを行なうことは比較的簡単と思われる。その結果が中位値についても、ばらつきについても、大規模デルファイに類似していれば、この小規模デルファイの利用価値は高いと言える。また、医療保健部門における環境問題関連技術に応用したところ、階層間の差がほとんどないことがわかった。これは、医療保健部門の回答

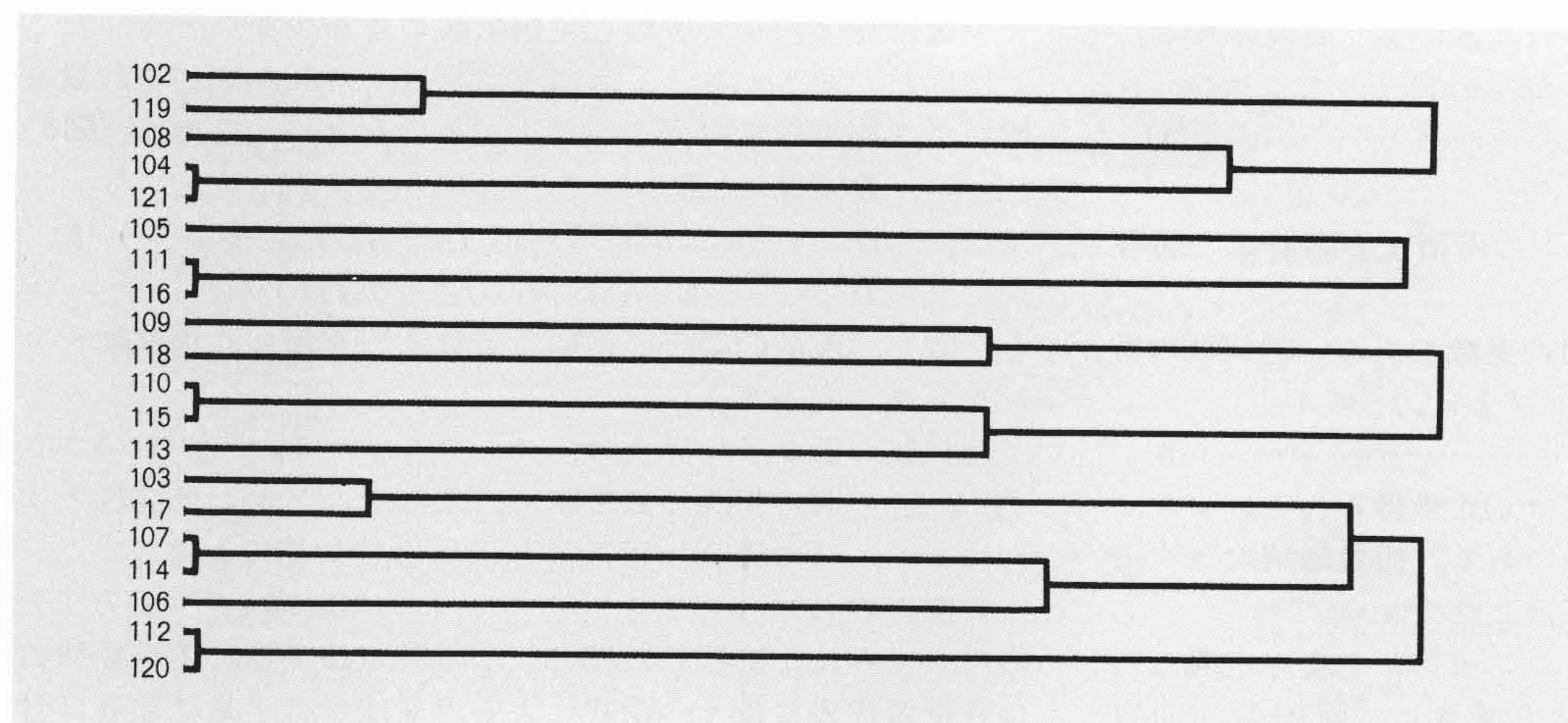


図6 公害防止技術のトリー
左の番号は技術課題番号⁽⁴⁾，結合の距離は類似度を表わす。

Fig. 6 Similarity Tree of Pollution Protection Technology

者はほとんど医薬部門の出身で同一学会に所属し相互コミュニケーションが深いと思われる。したがってこの部門では、任意に小規模デルファイを行なっても大勢においてさしつかえないと期待される。

4 波及効果分析

4.1 トリー形成

内容の類似性に基づき技術課題をトリー化することは実務上よく行なわれる。このとき内容理解が一面的だと、かたよったトリーが作られる。これを補うため、回答パターンによるトリーを作る。一つの技術課題は多くの面をもち、各階層回答者はそれらにさまざまなウエイトをつけて評価していくと考えられる。各面を総合して類似した課題には、各階層回答パターン間の関係も似てくると思われる。手法的には、3.5における階層と質問項目とを交換するだけでよい。数理的手法は普遍的だから、問題が若干異なっても応用できる。こうして得られた回答パターン類似度（ベクトル相関係数）からトリーを作るには、クラスター分析の最長距離法（別名Complete Linkage法）が良いと思われる。この方法を科学技術庁デルファイの社会開発部門公害防止技術に応用し、図6のトリーを得た。これによれば規制の社会的難易度が注目されていることがわかる。104自動車排ガスと121自然生態系解析は、不特定広域の汚染移動または移動汚染源に関する問題として類似する。111排水と116建設機器騒音は特定多数汚染源に関し類似する。また110污水处理場と115空港騒音は、場も管理者も特定少数の公共体という点で類似する。

4.2 クロス インパクト表の推定

質問項目間の関係を回答パターンから求めて、クロス インパクト表(図1参照)を作ることができる。偏相関係数を用いる方法、偏回帰係数を用いる方法および数量化理論IIの根の並ぶ順序による方法が考えられる。労働省デルファイでは、数量化理論IIIの根の並ぶ順序が一致すれば正インパクト、逆のときは負インパクト、それ以外の場合は特定インパクトなし(キャンセルアウト)と解釈して良い表を得た。数量化理論IIIを用いた理由は、独立要因と従属要因を区別しないモデルを用いたからである。

4.3 ポリシー ミックスの選定

目標とする波及効果を達成する最小のポリシー組合せをクロス インパクト表から得る方法を提案する。これは項目の中から基本項目を選別することに還元される。選定の基準は、重要度の高い順に、(1)目標波及効果を達成する、(2)最大の影

響力をもつ項目を含む、(3)基本項目は相互に独立である、(4)基本項目数はなるべく少ない、(5)上記基準を満たす組は一意的である、ことである。

労働省デルファイの場合、これら基準はすべて満たされた。その結果、(1)農業人口減少を止めること、(2)年金を充実すること、(3)労組の産業別連帯、(4)政府は賃上げ誘導目標を今後も提示しない。

以上の4政策が選定された。

5 結 言

専門家の意見を技術・社会開発政策の決定プロセスにおける先行要因として分析する視点を確立することによって、はじめてデルファイ結果の定量分析への道を開いた。特に重要な要因であるコンセンサスと波及効果を計量的に評価する方法を提案した。これを科学技術庁技術予測デルファイと労働省コンセンサス デルファイに応用し、有意義な分析結果を得た。この方法は、デルファイ主宰者にとっても、デルファイ結果を利用する企業などにとっても有用であることが確認された。

本研究は科学技術庁計画局および労働省労働経済課多数の担当官各位のご援助によるものであり、特に労働省浜事務官のご協力によるものが多大であった。また、立正大学稲越先生、立教大学松井先生、神戸商科大学河崎先生および東京大学竹内先生からそれぞれ有益なご示唆を得た。合わせて深く感謝の意を表わす次第である。今後筆者は、技術・社会開発政策のアセスメントの分析手法に拡張して、さらに研究を重ねたいと考えるものである。

参考文献

- (1) W.J. Conover, "A Kolmogorov Goodness of Fit Test for Discontinuous Distribution," J. Amer. Stat. Ass. 67, 591~596, 1972
- (2) 林知己夫ほか, 「情報処理と統計数理」産業図書, 1970
- (3) 林知己夫, 「技術予測の方法について」, オペレーションズ・リサーチ, 17, 3, 9~15, 1972
- (4) 科学技術庁, 「技術予測報告書」, 1971
- (5) 労働省, 「週休二日制, 余暇, 定年制」, 労働協会, 1973
- (6) 付記: 本稿は、昭和46年度財団法人未来工学研究所における筆者の研究結果の一部を引用した。