

数量化と予測に関する根本概念

林 知 己 夫

(1959年5月受付)

Fundamental Concept of the Theory of Quantification and Prediction

HAYASHI CHIKIO

In the present paper, the fundamental ideas of the theory of quantification and prediction are synthetically described from our mathematico-statistical point of view. The methodological details will be found in the papers of reference. Institute of Statistical Mathematics

拙いものでありますが末綱先生の還暦を記念して捧げるものであります。

本論述々においてはわれわれが従来研究を行つてきた数量化や現象予測法に関して、その根本的な考え方を中心にして述べようとするものである。新しい方法をここで論述するのではなしに、今まで行つてきた諸研究を位置づけ、更にわれわれがこの研究を続けるに当つて現在の、また近い将来も、われわれが採るとされる立場を統一的に取り纏めて明らかにしようとしたものである。われわれなりの総合報告というべきか。

§ 1. 前 置 き

現象を何等かの意味で数量化して捉えようとする事が科学の基本的方法であることは更めて言うまでも無い。自然科学・社会科学孰れの分野に於ても、一定の方法によつて測定を実施し、資料を獲得し、これを纏め上げ、推論を下して行くことに変わりはない。このような場合、現象の本質に喰い入り、その表現を考察し、如何に現象に形を与えてわれわれの処理可能な形に持ち込むか（フォーミュレーションと名づける）、如何様実験・調査を行つて資料を獲得するか、これを如何に表現し、解析し、推論・予測を与えて行くか、が考えられねばならないし、これらの一連の操作が、如何にして、妥当性・信頼性・再現性・実際的可能性あるものたらしめ得るかについて、考えを致さなければならない。この様な場合、統計数理の考え方が非常に重要な役割を演ずることになるのであるが、これについては今は省略する*。さて数量化に関して、上述のように広義に述べること

* 林 知己夫、現象解析における統計的方法、I, II, 科学基礎論研究、第2, 第3号参照。

資料があり、これを解析して結論を出そうとするならば、前述の様に必ず統計の考えや方法が必要なものとなつて来る。何もこれは統計的方法が万能であつて何もかもこれに依つて説明されると主張しているのではない、唯これが極めて有効妥当な道筋をつけていると言ふことなのである。

統計数理の根本の考え方は、常に現象解析と言うことを第一義と考え、根源に遡つて本質的に考えを進めることを重視する。さらに理論的であると同時に実証的であること、実証的であると同時に理論的であること——実際に我々の操作によつて得られる（把握し得られる）もののみを以て理論を構成することになる、我々は単なる説明概念として理論を考えるのではなく、予測と言うことを念頭におき、我々が実際に把握し得る限りの事象と事象との関連性を全体的に追求するのである——を志向し、この意味で、本来の意味での「操作的」な立場、なお深く我々の行為と云うものを中心に又根本にして考えを進め理論を凝出せしめる立場をとつてゆくのである。また現象解析の妥当性（validity: 我々にとって何が肝要であるかを考え、我々の目的をまさしく有用性を以て達成せしめることを第一義とする）、有効性を根本にして考えを進めること、逐次近似の考えでゆくこと、常に問題発見の形で考えを進めること、を規範として重視するのである。

は、問題が余りにも龐大に過ぎ、漠として、持味を出し得ないので、以下においては、主としてわれわれが常々考え、方法論的に発展させ来つた一つの面、一つの方向に限定して話を進めて見ることにしよう。一般論を抽象的に述べるよりはこの様な仕方の方が、寧ろより明確に科学の基礎に繋がるものが読み取れるであろうと思つている。勿論われわれは統計に携り、統計数理を考えているものであるから、当然この立場からの論議となるのである。

通常統計の方法を論ずる場合、その取扱うものの標識は必ず数量であるとされている。標識 X_1, X_2, \dots, X_N は数量とされ、これに確率分布が付与される場合は確率変数とされているのである。これが認められたとして、種々の技術が理論と称されて展開されているのである。しかし統計の技術を用いて、論を進めようとする場合、先ずここから注意をしてみる必要がある。当然のことながら、数量として表現する場合、何等かの意味で等質化が行われているものである。このような等質化が妥当な意味を持つか否かが注意されなければならない。これはあるものの表現の問題である。如何なるものでも、実態的には、恐らくは無限の標識でなければ表現し切れないことであろう。このようなものを幾らかの標識によつて表現して把えようというのであるから、十分考えを致さなければならぬ。ここに「標識を与える」ということが、われわれの目的に対して妥当であるか否か・有効であるか否かということと相対的關係になくはならない、つまり、それが操作的機能的でなくてはならないのである。われわれの目的が異なれば同じものであつても、当然与えらるべき標識は異つて来なくてはならないのである。標識はものそのものに内在する「与えられている」属性ではなく、我々が「与える」ものなのである。当り前のことであるが、われわれは顧みる必要はないであろうか。ある測定方法によつて得られたものが数量的定量的のものであらうと定質的（關係表現を含む）のものであらうと同様である。前者は所謂数値として表現され、後者には範疇的表現が採られることに成るに過ぎない。現在の統計的方法は線型的な關係を基礎に置いて發展されたものであり且つこれ以上に出ているものは尠いといつてよい。このような現状であるから、以上のような数量的に測定されたものであつても、それをその儘統計の技術に掛けて妥当な推論を下し得るとは限らない。測定された数量的標識を、線型的な基礎の考えの下に操作して妥当なものであると考えられて始めて、意味を持つのであつて、それ以外であれば科学的に信を置けるものとはならない。素朴な測定量そのものを標識としてよいか否かの配慮が必要である。 X が測定された儘の数量的標識であるとした時(確率変数とする)、この X の変換された量 $Y=f(X)$ が一般に妥当なものなのである。ここに X, Y はベクトルでもマトリックスでもよく、 f はそれに応じた函数であつてよい。 Y ではなしに、 X を統計的方法にかけても何等妥当な結果を得ることは出来ないのである。 E を待望値を取る記号とする時 f が線型關係に無いならば一般に $f(E(X)) \neq E f(X)$ であることも、注意しなければならない。孰れにしても我々の目的に対しては、妥当な f の形そのものを見出すことが大切な意味を持つのである。しばしば統計においては、統計の技術のためにのみ $Y=f(X)$ なる変換を加える場合がある。例えば X の分布はガウス分布をしないとき、ただガウス分布をさせるためにのみ f を使う場合があるが——しかもわれわれにとつては X の操作が妥当な意味を持つ——これ等は、われわれはこれまで幾度も指摘しているように大変な誤りである。このような機械的な方法が跡を絶たない以上議論を繰返すことも無駄ではないと思われる。敢て繰返すが、先ず以てわれわれは測定量そのものを、統計の操作によつて妥当な意味をもつ量に変換して置く、即ち f の形を求めて置くことが極めて大事なことであり、これを措いては真剣な分析は出来ないといつてよい*。

* この変換の一つとして、最近流行の情報理論 (information 理論、我々は現在工学的なものに於て特に通信のコーディングの問題に於て妥当性を持つものと考へている) における測度、つまり $H = -\sum_i^R p_i \log p_i$ (p_i は i なる状態の確率、実際的には相対頻度とその生起の独立性の假定、 R は状態の数) が用いられることがあるが、これが我々の目的に対して、果して我々が内容的に表現しようとしている情報に対して妥当なものとなつているか否かを検討しなければならない。通信理論→ある工学的な意味をもつ量を情報と命名→命名された名前「情報」をもとの意義から離れて自分流に広義に解釈→変換式を今自分の目的とする情報と言ふものに対する妥当な道具と思ひこむ、と言つた誤りが多いのではないかと思はれる。

このように如何なる f の形を求むべきかの問題は極めて肝要なものであり、これについては、以下述べようとする主要な論題でもあるので今はそのままにして置く。斯様に考えて来るならば、測定されたものが数値であつたとしても、それに所謂四則算法を任意に加えても妥当なものは出てこないのであるから、測定されたものが数値であろうと範疇化されたものでであろうと本質的には何等変る所はなく全く同じものと見てよいのである。この点を十二分に強調したいのである。ただ測定対象の性格や測定方法の特性によつて、測定されたものの性格は異つてはいるのである。例えば一般の物理的測定における測定結果とある社会的態度測定における測定結果との違いに見られるような差異がある。しかし、これとても問題への接近方法であつて時には、物理における測定においても態度測定の場合に似た様な配慮が必要なることも起り得ることに思いを致さねばならない。測定方法ということさえも、それによつて得られる測定値の性格を媒介としてわれわれの目的と相対関係にあるのである。如何なる場合であつても、目的と方法と妥当性との関係を慎重に考察しなければならないのである。

扱、測定結果が数量的のものであつても次のようにして範疇化することができる。測定したものが 51 瓦であつたとする。40 瓦以上であるか否か。45 瓦以上であるか否か。50 瓦以上であるか否か。55 瓦以上であるか否か。以上の問いによつて 5 瓦の範囲内で範疇化されることになる。つまり問いに対する解答の範疇として表明されることになる。5 瓦の範囲がわれわれの目的に対して粗すぎて妥当なものでないとするならば——ここでも絶対的な数値を考えることは意味をなさない。天体観測の時間の誤差における 1/100 秒と列車の到着時刻における誤差の 10 秒とはその妥当性からいつて同一乃至は後者がより精確であるといひ得るのである——さらに範疇化を細かくして行けばよいのである。これが目的と測定方法（誤差を含む）から考えて妥当な桁だけ細かくして扱えばよいのである。きつちりした 51 瓦ということに意味があるのではなく、目的と測定方法とから言つてそれが範囲内にあるものとして範疇化しておくことが寧ろ本質的に大事なこととなつていのである。ここを根本に分析を進めるのが第一義である。斯様な配慮は統計的検定論における差の検出と妥当な内容的に意味のある差ということの関係においても表われる大切なものである*。われわれの得た数値は単独の一つの点としての数値で意味を持つのではなく、ある幅を持つて妥当化され得るものであるということそれ自身が本質的なものであることを銘記しなければならない。

以上とは逆に範疇化されて扱えたものを、無理に数値化して処理しようとする場合もある。例えば有限の範疇に 0, 1, 2, 3...のごとき整数の数値を与え数量的測定と全く同様に考えて操作しようとするような場合である。これ等は全く任意なものと言へるのであつて、等間隔、単調等ということが真に内的意味を持たぬ限り、信を置けるものではなく、数値の与え方次第でどのような結論をも導き出すことが出来る訳である。これについては多くの人が反つて注意しているために——あまりにも形式的な扱い方であるために——容易に気が付いて検討が行われている所と思われる。

この前置きで述べたことは、事の内容如何を問わず、素朴な測定結果そのものを統計的技術に掛けて扱う前に、その内的意味・妥当性を検討して、測定結果のもつ性格や「目的と妥当な関係」を持つように変換したり、あるいはそれを範疇化して表現し、これを再び立場を明らかにして数量化し直さなければならないという点である。非常に一般化して論ずるならば、測定結果をすべて範疇化して扱えるものと考え、これに如何にして妥当な数量化を施すかを考へると言うことになる。この数量化の基礎として、われわれの操作的に明確にされた具体的目的・測定方法・測定結果の性格を思い、表現と分析と推論に最も妥当性あらしめるには如何にすべきかの考察があるべきであり、且つその根本の考え方としては操作的・機能的な立場に立たねばならないということである。

* 水野・林・松下・青山，ノンパラメトリック検定，朝倉書店；林，心理学における統計的方法，応用心理学講座第 9 巻のうち，中山書店。

これからさらに細論に入ろう。ここ迄にしても未だ漠としたものであるから、次のような特種問題に絞つてわれわれの方法を概略的に記してみよう。

§ 2. 感覚・感覚的綜合判断・綜合判断・意見・態度・能力等に於ける数量化の問題と其等の要因分析

一般論から離れて、一見数量化が為し難いと思われているものを把えて、これに就いての研究を述べてみようと思う。ここに用いられる考え方は単に此処丈ではなく、§1の一般的なものにも広く用いられるのであるが、自ら色合いは異つてくることはあり得る。他の分野では既に常識化している所でもあるが、この分野ではこの所を十分考えて行かねばならないために、特に、いま、掘り返しているのである。従つて、反つて、それが常識化している分野においての反省の機会を与えるものでもある。蓋し、その分野においても、分析が細目に亘り或は目的や観点が変わるならば、常識化しているものを再び掘りかえて新しい立場に立たなければならない所であろう。扱て、ここで述べるようなものに関しても、他の分野と同様、われわれの目的を操作的に明確に決めることが第一に必要である。次にそれに対して如何様な方法でこのような対象に対して資料を獲得すべきかが問題になる。その獲得すべき方法に関しては、実験であろうと、調査であろうとその如何を問わず、測定方法が問題になつてくる。これによつて得られる素朴な測定結果は数値であろうと、範疇化されたものであろうとも差支へはない。測定方法にはそれぞれの持味・性格があり、長所もあり、欠点もあるのであるが目的に対して最も然るべきものが考案され用いられねばならない。この基本的なものに対しては心理学の分野で昔から種々考えられているのでその得失に関してはそれに譲ろう*。このうち特殊なものについては後述することにしよう。また対象の選定方法に関しても、われわれは結果を云々すべき個物(単位)として何を採用すべきか、また結論を下すべき対象の集団として何を思考するのが有効適切であるかに関しても標本調査法に関する考え方や社会調査(実験)的技術を用いて、有効なものを編み出さねばならないが、これについても広きに亘りすぎるので他に譲ろう**。さてこうして得られた資料を表現する段取りになるのであるが、範疇化されたものを数量としてではなく関係を関係として把えて論を進める行き方もあり、何等かの立場から数量化して把える行き方もある。前者のような関係表現による現象記述であつても、どこかの段階で数量的な表現をとらなければならないであろう。この両者の差は程度の差であつて、一言にして尽せば最も妥当性ある段階において数量的表現に移すのが望ましいのである。

ここで述べようとするものにあつて、信念の数量化というようなもので古くから試みられていたものは確率を用いての考え方である***。Laplace や所謂 von Mises の考えているような確率も当然これに該当する解釈や意味を持つものではあるが、今は措き****主観確率というような立場の

* J.P. Guilford, *Psychometric Method*, McGraw Hill, 1936 (新版がある); 高木貞二, 実験心理学提要の中の一編, 岩波書店; 高木貞二編, 心理学に於ける数量化の研究, 東大出版; 林 知己夫, 心理学事典の中の評価法の項目, 平凡社; C.H. Coombs, *Theory of Psychological Scaling*, *Engineering Research Bulletin*, No. 34, Univ. of Michigan, 及び *On the Multi-dimensional Analysis of Monotonic Single Stimuli Data* (がり刷); S.W. Torgerson, *Theory and Methods of Scaling*, John Wiley, 我國の文献についてはその他心理学者のものも幾多あるが、私の気持に合つたものとしては当然の事乍らこの様なものとなつた。

** 概括的のもので気持に合つたものとして、経済学大辞典(東洋経済)の中の標本調査法・市場調査法, 調査すべき対象の集り, 母集団の考え方に関しては、心理学における統計的方法(中山書店, 心理学講座第9巻)をあげておく。その他参考にすべきものとして、基本的教科書の他に興味ある内外書物や調査報告書がある。

*** R.A. Fisher, は確率は信念の表現として多くの場合不適当で、それは確率の法則に従はない。統計に於ける母集団の選択の場合にはそれではなしにライクリフッドと言うようなものがよいとしているが、彼の論旨からだけではその意義は十分首肯できない。本集報別稿参照。

**** この解釈に関するものとしては、例えば、前掲の現象解析に於ける統計的方法-I; 松下嘉米男, 確率の概念について, 科学基礎論研究第3号; 林知己夫; これくていふ序説, 統計数理研究所, 講究録第3巻, 第5号; 林知己夫, サンプルング調査はどう行ふか, 東大出版; 及び末尾の文献参照。

ものを持取りあげてみよう。この立場のものは直接的で特に信念の数量化の意図が強い。Keynes や Finnetti や Jeffreys もそのよう立場から考えているが、一番突き進めたものとして Koopmans の主観確率がある。これは確率が数値としてではなく『一方が一方より「より起り易い」と思う』というような関係に対する考えが主軸にとられているのである。この考えは「信ずる」ことの一つの表現であると見られるので、興味ある表現であつて、数量化の気持を十分に掬んでいるものと見てよい。この立場は、要因からする行為の予測に関して非常に興味深いものを与えているといえる。しかも、個人について、その要因の現れから行為を予測しようとする場合大切な考え方であつて、最終的には個人の集りとしての集団における的中率によつて客観的にその方法の効率が表現されることになる訳であろう。この主観的確率はある種の仮設を置くことによつて、客観的なものと結び付けなければならない。しかし、誤つても、単なる説明原理や膏葉のようにどこにも貼り付くようなものとして用いてはならないことである。結果から見て「一方が主観確率が高かつたのだ、一方がより起りそうだと判断されたのだ」とするだけではわれわれにとつては意味をなさない。形式的な理論家によくやる誤りである。またこれを平凡に用いるならば、例えば社会調査を行つて、何かについて希望を質問し、それによつて「ある条件の下では一方のことが一方のことより起り易い」つまりある条件の下では一方が一方より主観確率が高いとして予測をするならば、これまた余りにも機械的である。斯様な立場ではなく、より立ち入つて、主観確率表現の動きを種々の条件の絡み合いの下で分析し、条件の重みづけ——これにも主観確率の観点が必要である——を考え、それを、予測するような方向に向わねばならない。これから以下に述べるようなわれわれの立場からこの問題の妥当性を研究して行く必要もあろう。また von Neumann はゲームの理論*といわれる非常に興味ある考え方を示し、行動意欲の数量化を試みている。人は「ある定まつた合理的行動をしようと考えて行動するものである——rational behaviour——」という立場から如何なる行動を取るべきかを行動の種類に付与すべき確率——これは主観的確率ではない——という形において決定しているので、意欲の一つの数量化ということが出来る。またこれをもとにして、行動に関する色々な合理的思惟方法も明確に表現されている。これは、行為しようとする人が一人ではなく互に相手(他人)を慮つて互に最も合理的な行為(ゲームの場合は、相手が自分を最も損をさせようとしていると考えそれにも拘らず自分が最も損をすることないような行為)をしようすることがその特色であり斬新な所である。したがつてそれを行う人々が同一の土俵の上にあることが大切な所である。この考えを用いるに當つて、土俵の違う人々の行為を単純に数量化することは出来ない。如何に土俵の差を表現し、また行為の意欲の数量化されたものが如何に異つてくるかについて考えを進めてみたいものである。以上のようなものが、この§で述べようとする事の曙であることには間違いはない。これと形式は異なるが、非常に重要なものがある。これは、戦後、アメリカの海外兵士を、最も多くの人が望むように、逆にいえば誰もが同じように不満を持つがその程度を成可く少くするようにするには、どのような要因(条件)を持つ兵士から先に復員せしむべきかについて、Guttman (L.) が考えた方法である。これでは多くの人の希望を最もよく満そうとする条件の下に兵士の持つ要因の数量化が試みられたことになる。この考え方には劃期的なものが含まれているといつてよい。定質的なものであつて範疇化して把握されている要因に数量を与えるに際して、全く任意に与えずに、一定の明確な立場から操作的に数量を与えようとした最初のものである。

われわれは、このような考え方を基礎に置いて理論を展開して行つた。これ等に関する詳しい方法論的内容は文献に譲るとして、考え方の筋や理論の進め方について、述べてみようと思う。われわれの立場は主として定質的なものの単なる表現だけではなく、その要因解析・その数量化に重点を志向するものである。まず簡単なものから説明して行こう。例えば総合判断的なものを考えるこ

* この一番始めの非常に興味ある論文は 1928 年に発表されていたがさして問題にされていなかったが、10 年来この方面の流行には刮目すべきものがある。この理由は全く私には解らない。

とにすると、美しいもの、面白いもの、好きなものを如何に表現するかというような問題も勿論含まれる。単に感覚的なものの表現というような低次のものは言うまでもない。ある人があるものを美しいと感じる、この感じはどのようなものであろうか、これを表現するために、比較法の総合によつて重みづけを考えることもあるし、或る人々が夫々美しい・醜いと感じる幾つかのものに貫流する共通的なものとは各々異つた特殊的なものとを分離し、それらを分類して行くような方法をとることもあるし、またある人が美しいと感じる要因を直接的に解析する立場を採ることも出来る。後者のような場合は、ある人があるものを美しいと感じるか、感じないかをまず以て調べる。これを外的基準によつて決定され・表現される変数と呼んでいる。あるいは単に外的基準ということがある。こうして、各々のものには、ある人によつて美しいと感じられたか、感じられていないかの標識がつけられることになる。これが外的基準による排反的範疇への反応といわれるものである。次にそれが美しさを構成するであろうと思われる要因をわれわれの立場から考えるのである。勿論単純なものだけでなく各種の相が複合したものであつてよいのである。その要因は主観的のものであつてもよいが、その要因のどこの範疇に属するかが一義的に安定性を以て決定されるものでなくてはならない。この場合、範疇の作り方やその個数については当然 §1 で述べた妥当性の配慮が怠られてはならない。孰れにしても、美しさの要因といわれるものは、ある測定項目（アイテムと言う）とそれにおける排反的範疇への反応という形で測定され扱えられるようにして置かなくてはならない。このアイテムは幾つ有つても差支へはない。このようにして要因と考えられるものは目的に対して妥当な有限個のアイテムにおけるそれぞれの範疇への反応の様相（パタン）という形で表わされることになるのである。われわれはこのパタンをあるモデルによつて再構成して、如何なパタンが如何様にして、ある人に美しいとされるものであるか、美しいとされないものであるかを判定出来ればよいのである。これが要因の分析といった形のものである。このような判定は多くのものがある人（人々）に判断させ、その「最も大きくなるように表現された共通部分」として、その人（人々）の美しいという感じはどのような要因（アイテムに対するパタン）から成立しているものであるか、その要因の利き方は如何なるものであるか、この方法の効率はどの位のものであるか、各のもつ特殊要因はどの位のものであるか——多くのものを貫流する共通的のものでどの位美しさの感じが再構成されるか——をわれわれの方法によつて明らかにすることが出来るのである。これを行うのに数量化の方法を用いることによつてメトリックを導入して、外的基準によつて表現されることが最もよく再現されるように、各アイテムの範疇に数量を賦与して行くのである。優れた勘というようなものは、どこの世界にでも大事なものであるが、これを上に述べた立場から外的基準と要因という立場からわれわれの目的に対して妥当な有限個の排反的範疇への反応として表現し、その要因を解析し、そのパタンを数量化して、そのもつ力を万人の共有財産にして保存することも可能になつてくるわけである。ここにのべた外的基準によつて表現されることは、感覚的なものでなく、判断であつても、客観的に与えられる判定であつても差支へはないのである。この外的基準によつて表現されることは調査・実験の仕方や問題のフォーミュレーションの仕方によつて色々作り出されるものであり、アイテムやその範疇の作り方にも平盤的でなく様々な形があつてよいのである。このような形式で把える仕方は問題の性格や目的によつて大に異なるのであるから形式的な議論で終ることはないので注意を必要とする。孰れにせよ外的基準を最も良く表現する様に要因の反応パタンを数量化することが緊要のこととなるのである。一見、外的基準が無い場合もある。分類というようなものを考える場合がこれである。外的基準がないといつたが、これは前に述べた素朴な意味でのことである。このような場合であつても、アイテムに対する範疇への反応パタンは存在しているものとする。分類とは同じようなパタンをもつものを幾つかに組分けすることである。組分けすべき数はきまつていないし、また全く同じパタンをもつものも少いのである。そこでわれわれはある立場から、組分けの基準をつくることから始めなければならない。どのような立場から組分けを

考えるかを定めなければならない。この場合であつても組分け・分類はある外的基準の明確な目的のために行うとするならば、分類のための分類ではなしに、直接外的基準と結び付けて考えるべきであつて、外的基準を最も良く表現するように要因のパタンを数量化することを考えればよいのである。しかし、ただ分類を考える場合は、分類されるもの同士の距離というようなものを考え、それが近いものを集めるということになるので、距離と言うものやその近さと言う様なものの定義を総合的立場からわれわれの分類の目的に応じて規定し、これを操作することによつて分類しようとするものを数量化し、この数量によつて組分けを行おうとするのである。こうした場合、アイテム・範疇の形で扱われている要因内容の数量化されたものは、分類に対する寄与を表現するものとして大切な意味を持つものである。分類以外のものであつても、同様であつて要因のパタンを例えばそれ自身の無矛盾性というものを最もよく表現するように総合し、要因のパタン及びそのもの自身を数量化することも出来るわけである。外的基準のない場合は、われわれが作ったある基準を最もよく表現するように数量化を行えばよいことになる。以上を纏めてみると

$$\left\{ \begin{array}{l} \text{外的基準が有る場合} \\ \text{外的基準が無い場合} \end{array} \right\} \left\{ \begin{array}{l} \text{それが数量である場合} \\ \text{それが質的な範疇である場合} \end{array} \right.$$

となる。外的基準がある場合はそれを、ない場合はわれわれがある立場から作った基準を最も良く表現するように妥当性を以て要因のパタン乃至はそれ自身を数量化することがわれわれの行うこととなるのである。あるものを表すのに O_i という記号を用いる。 i は i なるものという意味で $i=1, 2, \dots, N$ としておく。これに外的基準がある場合、それを y_s とする。これは数量であつても、範疇への反応であつてもよい。 $s=1, 2, \dots, K$ としておく。つまり O なるものはそれぞれ K 個の外的基準によつて判定されるものとしておく。この要因のアイテムは C_j としておく。ここに $j=1, 2, \dots, R$ としておく。つまり R 個のアイテムにおける範疇への反応として要因パタンが表わされることになる。外的基準にしろ、要因にしろ、一般的にアイテムに対する範疇への反応として与えられていると考えられるのであるから、それを書きあらわしてみよう。 s なる外的基準の方の範疇は α_s 個 ($s=1, 2, \dots, K$)、要因の方の j アイテムの範疇は β_j 個 ($j=1, 2, \dots, R$) あるものとする、 O_i はいずれかの範疇に反応するものとして表現される。これを表わすのに ${}^u\delta_i(s, \alpha_s)$, ${}^f\delta_i(j, \beta_j)$ なる記号を導入する。これらはいずれも

$$\left\{ \begin{array}{l} {}^u\delta_i(s, \alpha_s) = 1, O_i \text{ なるものか } s \text{ アイテムの } \alpha_s \text{ 範疇に反応しているとき} \\ \quad = 0, \text{ しからざるとき} \\ {}^f\delta_i(j, \beta_j) = 1, O_i \text{ なるものが } j \text{ アイテムの要因の } \beta_j \text{ 範疇に反応しているとき} \\ \quad = 0, \text{ しからざるとき} \end{array} \right.$$

として定義される。そうすると O_i は

$$\{ \{ ({}^u\delta_i(1, 1), {}^u\delta_i(1, 2), \dots, {}^u\delta_i(1, \alpha_1)), ({}^u\delta_i(2, 1), \dots, {}^u\delta_i(2, \alpha_2)), \dots, ({}^u\delta_i(K, 1), \dots, {}^u\delta_i(K, \alpha_K)) \} : \{ ({}^f\delta_i(1, 1), \dots, {}^f\delta_i(1, \beta_1)), ({}^f\delta_i(2, 1), \dots, {}^f\delta_i(2, \beta_2)), \dots, ({}^f\delta_i(R, 1), \dots, {}^f\delta_i(R, \beta_R)) \} \}$$

という 0, 1 からなるパタンによつて表現されることになる。これを §3 に述べるような方法によつて、 $i=1, 2, \dots, N$ を総合して外的基準（或はわれわれの作った基準、この場合は ${}^u\delta_i$ に関する項がなくなる）を最も良く表現するように各アイテムの範疇に数量を与えるということになる。

さて外的基準が数量である場合に一寸触れて置が、このような場合数量それ自身がわれわれの目的にとつて妥当なものでなくてはならないのはいうまでもない。この数量も幾度も繰返しているように範疇への反応として捉え得るように出来るわけである。またこの数量自身が「確率」である場合であつても一向に差支えはない。この場合は解釈上の問題が一応残り §3 で述べる考え方も必要になつて来よう。

扱わず、方面を変えて態度の数量化と言つたものについて述べてみようと思う。これは上述の議論の中に含まれるものでもあるが、従来より心理学の方面で試みられているものであつて、相当の遺産があるので、上記のものとは別個に論ずる方がよいと思われるので特に稍々詳しく取り上げてみることにする。これに於ては、如何なる調査方法をとるかにつての配慮が最も肝要なものとなり、これに付随して数量化の方法が考えられているといつてよい。後者に関しては、その多くは上に述べた枠内に入るものといつてよい。それ以外のものは上の数量化の方法ではなく、更に普通の統計的方法によつて処理し得るものとなつている。この重要なものとしては、**Thurstone (L.L.)**の態度数量化の方法があり、サーストン尺度とよばれるものである。この基礎にある考え方は、一次元的数量的態度軸の上に各人を位置づけようとするものであつて、このために等現間隔法の考え方——**Thurstone**の二者比較法の確率的モデルが基礎にある——が用いられている。**Guttman**はこれと異つた立場から人口に膾炙されている尺度分析法なるものを編み出している。これは態度を数量化するのに一次元的に考えず、多次元空間の中に位置づけることがより望ましい——われわれの目的つまり或る人の或る行動の予測に対して効率がよいという立場からこの考え方がとられている、目的にしたがつて質問の組み方は当然異つてくる——とされている。このためそれらに応じた各種の質問を編み出し、あることに対して好意的か非好意的かの態度の内容を現わすもの（コンテンツと呼ばれる）、その強さを現すもの（インテンシティと呼ばれる）、クロージャ(**closure**)と呼ばれるもの、インヴォリューション(**involution**)と呼ばれるもの、この四つをまず取りあげている。これらは、コンテンツに対して、それぞれ特色ある曲線を示すものであつて、これを見出すための質問の組み方、解答のとり方なども研究されている。一方これらのものを理論的に割りだそうとする考え方（態度の主要素理論）も作られ数量化の方法が用いられている。**Lazarsfeld (P. F.)**の潜在構造分析においては、人々は確率的に質問に答えるものという根底の上に立つて態度分析を考え、人々の集団には潜在的なクラスがあり、その中に属する人は各質問に対してある答に対して反応するある確率的な構造を持つ、そして各質問に対しては独立に解答するという立場から考えを進めている。はなはだ興味ある考え方ではあるが、このままでは十分妥当性あるものとは思われない。しかしこの考えは優れた著眼といわなければならない。また**Guttman**は態度調査をするに当り、質問を如何に組むべきかを考へている。彼は質問をつくるのにコンテンツ空間なるものを考へている。例えばある事柄に対する態度を知りたい、これを測りたいとする。このためには、種々の条件・状況下における質問とその回答の取り方を用意しなければならない。このすべての質問に対する回答を求めこの結果に統計的操作を加えるならば目的とする態度を知ることが出来るものと考へるのである。これがコンテンツ空間となるのである。これを明確に把握するためには諸条件や状況や解答の取り方をそれぞれいくつかの「操作的次元」に類別・分解し、更に分解されたそれぞれの次元においていくつかの範疇をつくりあげることが第一に必要となつて来る。これらのすべての組合わせとしてコンテンツ空間を作り上げるのである。この「操作的次元」をファセット(**facet**)と名づけている。それではその空間から、如何なる標本質問を抽出し、質問群をつくり調査にかけらば、コンテンツ空間のすべての質問に対する調査と等価になるか、あるいは如何様な信頼度でそれを推定できるかが考へられる。これについて思考をめぐらし、所謂実験計画法的な考えを用いたのがファセット理論と言われるものである。場合によつては、このような考え方それ自体に大いに疑問や難点があることを忘れてはならない。扱、この名前は**Guttman**が言い出したものであるが、既に早くわれわれはこの方法や考え方乃至は信頼度を加味した質問抽出方法を用いて一応の成果をあげていたのである*。

* 態度測定に就ての文献は後述するが邦文のものとしては、高木貞二編、心理学に於ける数量化の研究、東大出版；林知己夫、態度測定の現状と将来、輿論科学協会研究紀要 No. 4；林知己夫、諸外国と統計研究、科学基礎論研究、第 8, 9, 10 号；ファセット理論については上記の最後の二つに述べられてあるが、ファ
(次の頁の下に続く)

態度測定の問題に関しては、われわれはすでに上に述べた数量化の方法を用い種々態度の数量化を行つているが、さらに角度の異つたものとして次のようなものが試みられた。Guttman の尺度解析法は非常に優れたものであるが、ある質問群が尺度を作らない場合、不都合なことも起り得る。尺度をつくるようにする種々の方法が考えられているが（前頁脚注及び末尾参考文献参照）十分とは言い難い。ここに著目して、Guttman の方法に導入されていなかった質問に対するウェイトの概念を導入する方法を考えた。このために、各質問ごとに別々に解答を求める——Guttman の場合と同じ——他に、質問全体を通しての、つまり質問群に対するいくつかの解答模様 (response pattern) を与え、孰れに近い態度を持つかを併せ調査した。同一人に対して得られた二通りの解答を突き合わせ、各人が各質問に対してもつと考えられるウェイトと言うようなものを集団的に明らかにすることを行つた。この他バイアスをかけた質問を利用し態度のより深い分析を行う方法を考えた。また Lazarsfeld による潜在構造分析における確率的解答の考えを生かすと共に、——これはある事柄に対しては実験的に験し得るもので、ある面では非常に優れた考え方である——その好ましくない点を除くために確率的解答をしない部分、即ち強固にある態度 (意見) を保持する人々の群を選別し、その他の確率的解答を行うものに対しては、そのあるカテゴリーに対する解答確率の分布の特性を明らかにしようとするのを試みた。選別する方法と各特性を示すパラメーターを見出す方法がその中心的なものであるが、これによると各個人の特性を明らかにすることは出来ないが、集団ごとの特性は特徴づけることが出来る*。

なお態度測定の問題に関して、因子分析法の考えが用いられることがある。このためには測定されたものが数量でなければならないことが問題であると共に求められる因子 (軸) が如何なる意味をもつか、それ自身の中から求め難いという欠点がある。しかし、軸の意味は別として、以下に述べる多次元解析の立場から与えられた質問群の個数より低い個数の次元で全次元 (全体) を表現しようとするものであるとするならばそれ自身、統計的操作的には興味あるものといわねばならない (これについては §3 参照)。測定されたものが数量でなければならない、しかもこの数量が最後の結果に重大な影響を及ぼすことを考えるとき、その測定結果の妥当性は深酷に反省されねばならない。これについても Guttman は興味ある考え方を示している**。測定結果を数量として扱わないとするならば、因子分析の考え方は Lazarsfeld の潜在構造分析や Coombs の考えに近いものがある。以上のものについて因子 (軸) の意味は任意的であるが、それを明確にしようとする立場から Guttman はラデックス (radex) 理論を展開している**。

§ 3. 数量化における統計的方法

前に掲げたパターンに数量を与えるに当つて、全体を総合して考える必要がある。各要因や外的基準を一つ宛取り上げて独立に行うことは意味がない***。伝統的には前にも触れたように 0, 1, 2, 3 ……等の整数値をアイテム中の範疇に与えることがなされているが、各範疇が等間隔であるという

セツ的な考へ方としては、N. H. K (方法論は水野坦、西平重喜によつて考えられた)、放送の言語的条件と理解度との関係の実証的研究、放送文化研究所 (前掲のサンプリング調査はどう行ふか、にも書いてある); 堀川直義、文章のわかりやすさの研究、朝日新聞社調査研究室報告; 林知己夫、評価における統計数理的問題、評価と測定、教育心理学講座、金子書房; C. Hayashi, Note on Sampling from Sociometric Pattern, Annals of the Institute of Statistical Mathematics, Vol. 9 がすでに我国にもあり、且つ用いられて居り、国民性の統計数理的研究 (林知己夫、国民性の研究、教育統計第 30 号) の一部にも部分的に用いられていた。

* これら我々の行つた方法の詳細については末尾文献参照。

** 前掲 態度測定の現状と将来; L. Guttman, Metricizing Rank-order or Unordered Data for a Linear Factor Analysis, Center for Advanced Study in the Behavioral Sciences (Stanford, Calif.) のガリ刷; C.H. Coombs, Non-metric Factor Analysis (with R.C. Kao), Engineering Research Bulletin No. 38 Univ. of Michigan; ラデックス理論については末尾の文献参照

*** 例えば、林知己夫、色彩と統計、色彩科学講座中「工業と色彩」の中、修道社

必然性もなく、アイテム間を見てもそれが常に 0, 1, 2……と等しかるべき積極的理由もない。これは全く任意なものである。任意に数値を与えたのでは如何なる結果をも引き出すことが出来て了う。前にも述べたように数量は先験的にあるものではなく、われわれは目的に応じて、操作的・機械的に与えるものである。目的が変れば同じものであつても与える数値は異つてくるのである。これは測定値が測定方法に依存するのと同じく——先験的に測定値があるのではない、測定値は測定方法と相対的關係にあり、それから来る制限は当然あり、目的に応じて用いるべきことは周知の通りである——与えるべき数量はわれわれの目的と相対的關係にあり、妥当性を得しむることがその根本原理である。さて数量を与えると云つたが、ここでわれわれが与える数量には任意に演算を施してよいものではない。即ち自づと使い方があるものである。目的と方法とに依存してその持つ性格が異り、意味が違つてくるのである。ある時は差だけが意味を持ち、あるときは比・積のみが意味をもつことがある。われわれが与える数量の意味も数量化の方法に相対的であることを忘れてはならない。最後の目的に到達する中途の操作としては、直接的に数量を与える場合だけではなく、種々の定質的分類と言つた操作を重ねて行くこともあり得るが、この時も分類の中の等価なもの——外的或は我々の基準に照して——を見出すために数量化することもあり得る。要因のパタンだけ見れば異つていても、上述の基準に照してみると、その点に関しては全く差異を認め得ぬものもあるのであり、この意味でそれは同質化されるといい得るのであり、これを見出すことは非常に大切なことであり、われわれの方法の重要な成果でもある。所謂実態的に眺めるならば、存在するものは全て異つている、同じものは先ず一つも無いといつてよい。しかし、それ等を幾つかのアイテムに投影してみるならば、簡単にそれぞれの有限個の範疇の孰れかへ分類されて了う。先ず外的或はわれわれの基準に投影したとしよう。こうして孰れかの範疇に分類されたものを取扱うのである。次に要因のそれぞれのアイテムに投影したものを総合するのに範疇への反応パタンとして把えることが出来る。こうすると、実態そのものが問題となるのではなく、範疇への反応パタンだけが問題となるので、一見異つているものも、それぞれに応じて等質化されて来るのである。目的とすることが前者のように把えられるならば、これと関係づけようとする要因の選択に当つて、われわれが意味をもつと考へ得る諸要因——このためにただ一つの仮説の上に立つて要因を選択するのはわれわれの立場ではない、寧ろ問題発見を意図して先ず以て現象を記述し描き出すことが大切なものであると考へるので、多くの仮説を立て（場合により同時に反対仮説をたてる）、この上に立つて要因を選択する、われわれとしては爾後最も妥当な仮説を操作的に採り上げることを考へているのである——を採り上げ、これをアイテム化し、範疇の中に把えるのである。斯くして、目的とするものと関係づけるのである。つまり実態的なものとしてでなく、われわれの立場から作りあげられたパタンをもとにして、外的乃至はわれわれの基準を最も良く表現するように、これ等のものを再構成して行くことを考へているのである。それがどの程度まで行くかが、われわれの方法の効率、妥当性となつて来るのである。この意味からいつて、われわれの目的に対して等価なパタンを見出すことが極めて肝要なことになるのである。

一見はなほだしく異つた要因のパタンを示していたとしても、目的（前述のように範疇化されている）に対しては等価であるものも多いので、この等価の意味づけから要因パタンを媒介として現象のより深い構造を剔抉することが出来てくるのである。

さて一つ一つの要因でなくパタンを問題にし、これと外的乃至はわれわれの基準と関係づけるのであるから当然多次元解析の技法を援用しなければならなくなつてくる。勿論現在の多次元解析の技術*は満足すべきものではなく、寧ろ極めて薄弱なものであり、次元の場合の一つの偏つた拡張に過ぎないと思われるので、臆はわれわれの目的に添つたものを作り上げねばならないと考へ**

* これについては集報の次号に全体的な総合報告がなされる。

** 例へば、多次元の場合の測度の入れ方、多次元行列（式）の作り換へ（今までの満足すべきものではない）等である。

ているが、現在の所は一応多次元解析の考えを利用しなければならない。こうして要因の範疇や O_i や基準に対して数量を与えることになるのである。一度数量が与えられるならば、多次元における種々の統計量が、われわれの目的に応じて解釈されてくることになる。ここで翻つて考えてみるに、われわれの方法はまた一つの多次元解析の自然な拡張と見られ得るのである。つまり、予め数量が与えられていたとせず、質的パターンそのものを扱い、これに数量を与えようとするからである。前者はパラメーターの推定が問題になるが、後者ではそれを含めて、パターンへ与える数量に対する推定が問題になつていくからである。ただしここには非常な差異がある。前者の場合は空間に座標をとることが出来、前以て O_i の位置づけをすることが出来、その関係表現操作が問題になるのである。位置づけられたものとしての相互関係が数量的に問題になるのである。われわれの場合は予め空間の中に位置づけることは出来ない。数値を与えたものとして、座標を決定することが出来るが、位置が決つていないので、アイテムに関して O_i の空間配置をして置くことは出来ない。この空間的配置それ自体をパターンに表現された相互関係を媒介として定めることがわれわれの目的となるのである。われわれの目的に応じて（つまり外的或はわれわれの基準）、空間配置そのものが変化してくる所が面白いのであつて、目的の差異と空間配置の差異とを関係づけてくるならば、ますます要因（アイテム）の持つ意味が明確になつてくるのであり、より分析を深めることが出来るものといえよう。これもわれわれの方法の特質である。

なお、例えば、要因のアイテムがまず測定されたままで数量であつた場合、更めてこの考え方に依つて数量が与えられる場合、両者を関係付けてみるならば、もとの数値の変換されたものが新しい数値となつていくのであるから、§1 で述べた f を求めたことになつて、素朴な数量のもつ特性が明らかにされることになる。見かけのスケールが変換されて、妥当な意味をもつスケールとなると考えてよいのである。この時の f の形に注目すべきであつて、敢て単調な関係のみでなく、非連続に近いこと、例えばある点・ある点で飛躍し階段函数 (step function) の形で与えられることもある。このような場合には、その要因の持つより深い意味内容や関連性を見出すことが出来る（定質的な範疇の場合であつても、直観的な関係ではなく何かある所で飛躍的な関係が生じたり、単調でない場合等も起るがこの時は看過することなく、同様により深い洞察を必要とする）。例えば f が二次曲線であれば目的に対して折れ曲る点があり、 f が三次曲線であれば、山や谷を介して、素朴な数量のもつ意味が明確になる。Guttman の強度・クロージュア・インポリューションについて考えてみよう。コンテンツとして測定されたものを一応数量としておくならば——尺度解析法の立場から数量化されたものとしておく——また外的基準があるもののあることへの予測ということにして置くならば、この外的基準によつて表現されることを最もよく表現するように上述の三つの定質的内容をコンテンツと共に総合的に多次元的に数量化して、コンテンツの数量に対して変換の f を求めるならば、上述の三つのものはそれぞれの基準に応じて 2 次、3 次、4 次曲線によつてよく表現されるというようなことにもなると言えるであろう。

さて、基準によつて範疇化されたもの、要因のアイテムの範疇に対して与えるべき数量は単に一次元的な数値である必要は毫もないのである。これで十分目的を果している——妥当性がある——ならばそれで良いのであるが、そうは行かない場合も多い。このような時は多次元的な数量（一種のベクトル）であつてもよいのである。外的基準がただ一つの場合であつても、その範疇が 3 個以上である様な場合、基準が多い場合、いわば範疇間で移動律の成立しないような場合、複雑な分類を考えなくてはならないような場合は、一つの数値では十分ではなく、多次元的な数値を与える必要が起つてくる。例えば、二者比較法による調査による場合、 $O_1 \square O_2 \square O_3$, $O_1 \square O_3$ (\square , \square の符号は……より……を選択するというを示す) のごとき場合は極めて多いが、この時 O_1 , O_2 , O_3 を一列に並べることは出来ないが、二次元的な配置を与えることによつて問題の解決が可能となる

のである*。

このように多次元的な数量を与えることも多次元解析の重要な拡張であるが、数量のもつ性格が実数の性格と異なることには注意しなければならない。

次には具体的に如何にしてパタンから数量化を行つて数量を紡ぎ出すかを決めなければならない。外的基準乃至はわれわれの基準によつて表現されたものと要因のパタンとを関係づける仕方であるが、脚注に述べたような考え方も一つの方法である。一般的には、一次的に考えることの出来る場合は、両者が数量化されたものとして考えるならば一般的なモデルとして相関係数、これが好ましくない時はそれに相当する適切な測度をとることがあるし、基準によつて表現されるものを範疇のまま取扱うとすれば、一次的には相関比或は、外的基準のある数量的分割境界（判定境界）に従つて予測するとき的中・弁別の成功率（確率によつて表現される）乃至はそれに相当する測度がとられる。両者多次元的に考える場合は、多次元の場合如上のものに相当するもの、一般化された分散、これと類似する他の測度、成功率というようなものもとられるのである。この他この頁の脚注の場合と同様な考え方——ミニマックスの考え方もその一例である——もとられてよいのである。孰れにしても、この様にして作られる関係づけのための量が最も望ましいものになるように、相関係数・相関比・弁別の成功率というものであれば、それが最も高くなるように、他の測度であれば、それが最もよくわれわれの欲するものになるように要因のパタン（アイテムの各範疇）や基準によつて表現される範疇やまた O_i に数量を賦与することが試みられねばならない。また「考え方」を中心にするものであれば、それが最もよく満足されるように、操作的に、数量化がなされねばならない。こうすることによつて得られた数量がどの程度妥当性を持つかは、その測度や基準との合致度の持つ意味によつて効果的に表現されることになるわけである。

このようにして得られた数量の意味を考えることが大事である。数量化によつて作られた数量は、多次元的要因パタンのまにまに然るべき方法を基として得られたものであるから、その示す操作的内容にしたがつて解釈を下し、現象のより深い構造を剔抉・把握することが可能になり、これより進んでさらに妥当なモデル建立に移り得るのである。

なお数量化を実際に行うには龐大な数値計算を必要とする場合が多いので、その計算法についても、数式の内容や計算機の特性にしたがつて最も適合したものが作成せられねばならない。

以上ははなはだ概念的な説明であるが、これらの考え方を目的に応じて、種々効果的に使い分けることが非常に大切なことになつてくるのであるが、これについて詳論をする暇はない。目的の種類と方法との対応を統一につけて行かねばならないのであるが、これはさらに理論が発展した暁に、振返つてなさねばならないところであろう。しかし、当面は寧ろ、様々の問題に応じてわれわれが考えを進め新しい方法をつくり出す途上にあるといつてよく、徒らな統一化より個別に徹しな

* 二者比較法で 10 人の人に 7 個のものの好嫌を判定させたとして。この時各人の結果をつみあげたマ

B \ A	O_1	O_2	O_3	O_4	O_5	O_6	O_7	計	トリックスを左に示す。横にみた場合、Aに	第1軸	第2軸	第3軸	
O_1	8	4	4	8	8	4	4	36	書かれてあるものと比べ、それよりも横のもの	O_1	1	7	4
O_2	2	6	6	8	10	6	4	38	のを好きだとした人数を示す。このような場	O_2	2	2	5
O_3	6	4	4	8	10	6	4	38	合 O_1, \dots, O_7 を全体的でよいが好嫌の順にな	O_3	3	4	3
O_4	6	4	6	4	10	6	4	36	らべようとしても無理が起る。そこで、例え	O_4	5	3	2
O_5	2	2	2	6	10	6	4	28	ば、ある人は O_i をある軸に投影した所だけ	O_5	4	1	6
O_6	2	0	0	0	0	6	4	8	に着目して好嫌いを見、ある人はまた O_i を	O_6	6	5	7
O_7	6	4	4	4	4	4	4	26	に着目して好嫌をみるものと仮定してみよう。この	O_7	7	6	1

ある軸に投影した所だけに着目して好嫌をみるものと仮定してみよう。この着目する軸は操作的にはなるべく少い方が望ましいわけである。各 O_1, \dots, O_7 に与へる数量は順序の意味でよく、夫々の軸で順序の若い方がより好ましいとされるものとしておこう。こうすると O_1, \dots, O_7 の夫々の面における順序は右の様になり、これによつて空間的配置がきまつてくることになる。第1軸にのみ着目する人は4人、第2軸のみは2人、第3軸のみは4人となる。なおまた、この方法をより深めることもできる。さらにこれから突き進んで、軸の意味内容の解明に手をのばして行く必要がある。

これなども全くの例示に過ぎないが多次元の数量を与える一つの考え方である。数量と言つてもこの場合は順序の意味しか持たない事に注意しなければならない。

ければならない。これら個別のものに関してはその一々の文献を参照せられたい。

現象解析に当つては、数量化に際して用いる集団を時間的・空間的（特性的）に分割したり、得られた数量を比較し、意味づけ、内容分析してより深く現象構造をフォーミュレイトすることも大切なことである。更に数量化の方法もただそれだけを見ていたのでは、現象解析には十分ではない。現象解析中の重要な一方法であるが、総合的に他の統計数理の方法と絡み合わせて行かなければならない。現象解析の達人は諸の考え方、諸の技法の運用に長ずるものといつてよい。これらに関するいくつかの例も後掲する文献にあるのでそれ等を参照せられたい。

§ 4. 予測についての統計数理的考察——その一般論と数量化との関係

予測ということは非常に大切なものであつて広く解釈するならば科学の目的は予測にあるといつてよい。

過去のデータの分析や原理（理論）も予測的知識を与えるのでなくては意味がない。これに統計的方法を用いるにあつても然りである。

われわれは予測の問題を単純に形式的に見ようとする考え方ではない。単なる確率論的モデルの作成に終るものでもない。確率論的に把えようとするならば、そのフォーミュレーションと現実との対応の妥当性がまず第一に検討されねばならない。問題をそこまで帰著せしめる配慮が——したがつてその有効性の範囲が限定され明確にされる——最も大切な所である。このように問題解決の近接の仕方の研究が一番最初に行われねばならない。何者について、何をいかように予測しようとするかという、その予測すべき「何者」の決定（行動単位の決定と名づける、事象を単位とするかある集団中の個人を単位とするか、或はある集団そのものを単位とするか等、これが究極の目的とする単位のこともあるし、途次操作的に作り上げられることもある）、それから次に、頭初にも述べた様に、現象解析に対する目の付け所、これのフォーミュレーション（科学的統計的に取り扱うことが可能であり、且つ妥当性を持つようにする）、解析の基礎となる理論的考察（仮設・作業仮設の構成）、実験・調査による諸の資料の獲得、その表現、この分析・総合・推論の方法、これの有効・有用性の検討（実際へのあてはまり工合の検討、反復によるその妥当性の検討）これらの方法論を実際に即して展開する、これに統計的方法が捧げられることになるのである。

予測である以上、先づ予測すべきもののフォーミュレーションとリダクションとが概念的でなく、具体的に明確になされねばならない。これらは現象予測すべきものが何らかの意味で範疇化して把えられねばならないことを意味する（我々は、これによつてそれが操作的に明確に表現されていることを意味するものとする）。数量的表現が採られようと定質的な分類で把えられようといずれにしても差支えない。予測しようとするものがどのように複雑なものであつても、またダイナミックなものであつてもそれをそれなりに有限個で範疇化して置くのである。このことについては異論があり統計的方法による予測を狭いものとする考え方があるかも知れないが、ここを明確にして置かねば科学的に予測を取り扱うことは不可能になつて了う。この範疇化の行き方も良く考えてみるならば——また範疇化の方法がわれわれの目的に対して適切妥当であるならば——そう狭いものではない。この範疇化されたものを組み合わせて、われわれの目的の適切な表現をとることを試みるのである。なお予測しようとするものが数量であつても前 §2 で述べたように考えるべきである。また予測しようとするものが決定論的なものであつてもよいし「確率」であつても差支えないのである（後者の場合その中の検討には深い配慮を必要とする）。いまは予測しようとするものについて考察したのであるがその予測を与えるために用いる要因その物についても前 § で詳述したようなことがある得るのである*。

* われわれは、ここでも実態的なものに就て全面的に予測を行おうと言うのではない。われわれの行動として決定せられる現実には、やはり妥当性をもつて排反的有限個の範疇として把える事が出来るのである。

さて予測を考える以上その精度の点を明確に表現しておかねばならない。これを二つに分けて考えるのが得策である。一つは得られた資料をランダム・サンプルと考えてその場合の予測の誤差、もう一つはその分析の根底に用いられた仮設の恒常性に関する予測の誤差である。これについてはこれから更に突込んで行こうと思う。先ずその第一の方から取り扱つてみる。いちばん単純に考えるならば過去の資料から得られる行動単位群についての知識は、過去・現在・将来を含む母集団（多くの場合等しい抽出確率を与えられた行動単位の集まり）からのランダム・サンプルと考えられるとする行き方である。つまり、それがランダム・サンプルと考えられるような未来に対して適用する時、その精度が理論的に求められるという考えである。勿論この時、過去の資料間に想定される関係式は未来においても恒常的なものであるとされるのである。この考え方は、社会現象に対しては、それが絶対にくり返さないものである限り、問題であるかも知れない。しかし、これは実際的にはそのようであるかも知れないが、われわれのごとく実際の目的をさだめ、機能的・操作的にものを考える立場をとるならば、首肯される。即ちわれわれの現実的立場から、要因なり予測すべきものなりを採り上げているので、いわばわれわれによつて捨象され残つたものについて見る時、同じことの繰り返されるあるいは近似的意味で同じことの繰り返されると見做され得ることも極めて多いものと考えられる。この意味において、同質的な様相を持つ過去・現在・未来を考え且つ把握することが妥当性をもつて可能となるであろう。これだけでは狭すぎる。ある条件の時々の資料が採られた。将来はその条件が変わるといふことも有り得る。このような場合には条件とその時の反応の関係状況の恒常性（過去・現在・未来における）が仮定せられるならば、その条件変化が前の図式で予測せられるときには、問題とする予測も同様の考え方によつて可能となるのである。このように、過去・現在・未来を通しての何等かの恒常性が保証せられるならば、そしてまた条件の予測も母集団-サンプルのような立場から考え得るものならば、その精度は確率論をもつて表現せられることになる訳である。以上のような考え方には、予測を空間的に把握するという、比喩的な表現が与えられてよかろう。つまり予測が推定の問題に還元されたと考えられることになる。いわゆるインターポレーションの考え方に相通ずるものがあり、インターポレーション自身この考え方によつて処理され得ることにならう。この第一の場合においては、更に後述する様に、予測的の度合（これが高いほど望ましいもの）、——的中比率ということにする——その的中比率のサムプリング分散（その小なるほど安定性が高い）を分けて考える必要がある。孰れにしても母集団

それが複雑なものであつても、われわれがその予測の目的を明確に限定するならばその様に表現し得るものである（これを実際に行うのは容易であるというのではない。非常な努力と苦心とを要するのであるが考えを深めることに依つてまた技術を練磨することに依つて可能となし得るとの意味である）。考え回らしていることは、如何に幅輻していようと行為として決断に依り現実化された行動様式——操作的には、われわれはこれのみしか把える事は出来ない、従令それが心の中のものであろうともそれが操作的に把えるものであればこの種の行動となつて表に現われているものと考え得るからである——そのものは範疇化されるのであり、これをわれわれは問題にするのである。これを明確な形で組み合わせて、予測すべきものとして解析してゆくのである。これに依つて始て、これに依つてこそ、これに依つてのみ予測の現実的・直接的有用性が確保せられると言えるのである。これが予測においてわれわれが何を何の為に予測しようとするかという“目的の限定”の必要である所似である。

われわれの“目的の適切な表現”といつたが、これには次のような考察をも含めるものである。例えばある物財の需要予測をしようとするのを考える。十年後の予測も考えようとするれば、この精度のよい数量的予測はおそらくは不可能であろう。翻つて何の為にこの予測を行おうとするか、を考えてみる必要がある。それは、今如何なる施策を採りたいかを知ろうとする為の事が多いであろう。そうであるとするならば、問題のフォーミュレーションを変え、いま、われわれの考え得るいつれの施策をとるならば、もつとも損失が少ないかが解るようなフォーミュレーションが望ましいものとなる。この實際的解決には需要予測も勿論必要であるが、これを精度よく知り得ないと言うことを基にして、この需要の傾向的予測の知識を用い、明確な解決を与えることが出来る。これが問題のリダクションである。ここに述べたものは、最も突込んだ、最も実際のな所で、これを行おうとする考え方である。これは前§2で触れた所謂ミニマックスの理論（ゲームの理論）に通ずる考え方であり、統計的行為決定法（statistical decision）の考えにしたがうものと言つてよい。これに限らず問題のリダクションは大切な所である。

を考へることによつてこれらを計算することは可能なのである。以上第一番目の誤差の問題は全く確率論的に計量可能であり、予測の精度は確率の数値によつて現わされ得ることになる。この前提が統計的方法を用いる予測の科学的保証となるのである。さて二番目の誤差の処理は聊か異なる。これは、過去の資料が過去・現在・未来よりのランダム・サンプルと考えられない場合（あるいはその適用を誤つた場合）、何らかの意味の恒常性が保持されなくなつた場合、つまり予測モデルそのもののずれがおこつた場合のものである。これは前にも述べた同質の様相を持つ母集団を考へるところに問題がある。この時多くの場合、資料の得られた背景の差異が挙げられるであろう（背景が変化しても、それが、採り上げた要因に関してはランダムな働きしかなさぬものと見做し得るか否かによつて、その取扱ひ方が変わつて来るのはいうまでもない）。この背景そのものをすべて条件として考へ尽せるものならば、そして条件と現象の生起とを理論的・実証的に対応づけ得るものならば、さらにその恒常性が保持され、将来の変化する条件が予測し得るものであるならば、問題の処理は前述したように可能となる訳である。しかし上述の要素のうち一つが欠けてもこういう分析は不可能となる。背景そのものに関する知識が不十分な場合も多々あるものである。これにも拘らず、前に述べた同質性の母集団と考へるところに誤差が起つてくる訳である。そのため、可及的に同質的な母集団に限り予測が行われるよう、条件を限り、予測の適用を範囲を制限して（母集団に制約を与えて）考へを進めて行くことが必要になる。これがどの程度、近似的に当て嵌り、現実的に妥当であるかは予測研究を積み上げ、これを理論的に検討しなければならない問題である。この第二番目の誤差を評価し、それを少なくするには、同一種類の研究に対して、体験を積み、資料を重ね、また問題点は事例研究によつて深く掘り下げ、長期にわたつて研鑽が積み重ねなければならない。予測の研究は一回切りのものではなく、フォローアップが続けられなければならないのである。重ねて述べるが、第一のものは確率論的に処理でき第二のものは多くの経験を経てフォーミュレーションがなされ、あるものは再び第一のものの中に含ましめることが出来る（つまり、採り上げる要因の中に含ましめられわれの予測の方式を作り上げることが出来ることになる）が、その他のものは母集団の制約として形を与えられて、併せて予測の実質的精度が評価せられてゆくべきものである。なお、要因のうち、あまり、母集団を考へた時の背景（環境・条件）に左右されないもののみを採り上げるならば可及的に同質的な母集団の範囲を拡大出来ることも注意する必要がある。雖就そのような要因を採り上げた時第一の意味の予測的中率が聊か低くなろうとも、同質的と考へ得る母集団が広くなれば、特殊時点にのみ利く要因を加えて狭い範囲の同質的な母集団に対する的中率の高い予測を与えるよりも有利な場合も多いのである。同質的な母集団の範囲は、採り上げる要因と予測すべきものとの関係の恒常性に依存するものと考えなければならない。

これまで外形的な枠について話を進めてきたが、予測の当り・外れということ进行分析してみよう。これは厳密に述べれば“予測的中の割合”ということになる。また、何について予測を与えるのかの行動単位を明確にしておかねばならない。これが予測的中の割合をきめる基ともなるからである。“予測は当らねばならない”これは至上のことである。先ずこの“当る”という内容を明確にしてゆかねばならない。前に予測しようとするものが範疇化されておらねばならない——それはもちろん数量そのものであつてもよいのである——と書いたが、今ある行動単位に関して予測したものが、この範疇化されたものに的中しているか否かが“当る”“当らぬ”の意味である。勿論それより前にこの範疇化の方法自体が現実的妥当性を持つように考へて置かなくてはならないのは言う迄もない。この意味での中・非的中が考へられねばならない。いま範疇化されたものがある数値であるとしよう。この数値そのものをピッタリと当てねばならぬか否かを現実的に考へてみる必要がある。例えば ± 2 は許せるとすれば、その数値およびその範囲を範疇化されたものとするのである。したがつて唯名的に物理的に単位が同一であつても、知ろうとする内容が異なるならばその数値のもつ妥当性の意義が異なるのである（この意味で、ものによつて異なるので、そのものを見

るためのスケールという)。前にも述べた様に天体観測における 1/100 秒の差とある人の集合時刻の遅速における 1 秒の差とではその内容が非常に異なっているであろう。このように当るといふことの相対的意味を考えて、前§で詳述した様に問題のフォーミュレーション、範疇化を考えなくてはならない。この目の粗さによつて、予測の方法自体の方式が決定されてくるわけである(現実的負担一定のもとに最大の効率をあげるという観点を必ずわれわれは念頭におかねばならないから)。この予測的中する比率(予測を与えた全行動単位中、予測的中した行動単位の割合)は母集団における確率的な表現として与えるのである*。

また、予測を行おうとする場合、的中の精度を高めなければならないのは言うまでもないが、その精度が低い場合、これらの材料による分析においては“予測判定困難”ということをはっきりすることも今後の研究のため、非常に意義あることを忘れてはならない。

次には予測に関する研究法の問題である。幾度も話をしているように、われわれの立場は実際的なものであるから、抽象的な議論を行わない。まず行動単位ごとに過去の資料・それを包む環境条件についても資料を一定の方式にしたがって蒐集するのが第一歩である。それは予測に対して有効な要因と考えられる諸種の資料——この現象を構成する要因別決のための理論的考察と「調査という操作」による要因に関する資料との関係、事例調査・統計的準備調査による有効な要因の決定及び資料の組み合わせ方の問題等をも考慮に入れた上での話である——とその結果(これが予測すべきものであつて、寧ろ“現われ”といつた方が良いかも知れない)とを行動単位ごとに対応せしめることから始めねばならない。この二種のものが予測研究の手掛りとなるのであつて、この二種のものの突き合わせによつて要因から、如何様なことが起るか、如何に予測可能であるかを検討することになるのである。ここに設定された予測方式を用い、将来は要因の方を調べ上げ、如何様なことが起るかを予測することになる訳である。この資料は一時点のもので無く可能なる限り多くの時点のものが得られることが望ましいのである。

現象の予測に関する研究では、現象が明確なものを把えて方法論を鍛えることから始めるのが得策であろうと思う。現象が明確であるとは、予測が容易なものを意味するのではない。予測すべきものが明確に判定出来るものをいうのである。そのため社会現象では経済行動の予測、選挙予測、犯罪予測、教育効果の予測、学力の予測、結婚予測、マスコミ接触状況の予測等の分明したものから取り扱つて考えを進めるのが適当である。これによつて方法を鍛え、これが完備すれば、それに限らず広く応用して、確信を以て多くの部面に施そうと考えるのである。

さて予測の図式にいかなるものがあるかを考えてみよう。一般的な考察は他に譲るとして、ここでは次のような立場から考えを進めることにする。

第一のものは、局限された場面において、現象の構造を抽象し(現象の内容を確率を以て表現することもあり得る)、単純化し、形式化し、ストカスティックな函数関係・確率を中心とした函数関係(確率過程)を建て、形式的には微積分方程式、あるいは階差方程式の形にもちこみ、あるいは電氣的なアナログ・確率的なアナログによつて問題を解明しようとする立場である。この方法によると取りあげる要因は比較的簡単であり、複雑な現象パターンを問題にせず、大局的に筋を通す

- * 1. 範疇化されたものが厳しいものであればあるだけ、予測的中の比率が低くなるから、許容出来る範囲内で範疇化すべきものは甘く考えておくのが実際的である。
2. 予測的中の比率は 100% になるものは殆ど無いといつてよい。出来る丈少数の要因で可及的に高い中比率を得ようとするのが大切なことである。この場合、他の要因はランダム要因と見なしてゆくことになる。
3. 前に述べた様に、如何なる施策をとるべきかの問題のときは、この施策を有限個のものとして表現し、いずれの施策をとるとき、損失がいちばん少ないかということ予測する問題となる。この“損失の程度的中”も今の様な立場から閉らかにしておかねばならない。具体的にこの予測の外れが大きい場合であつても、当時の状態から考えて、他の施策に比べて最も良かったと言うのであれば少なくとも満足すべきであろう。この外れの要因を分析するのは、もとより必要であるが、他の施策も考えるに到らなかつた所が問題点である。

考え方である。この方法はいわゆる“理論”的考察を主とするものであつて、形式的に流れ易く欠点としては、複雑な現実での予測が従となつているものが多いと思われるので*、われわれとしては新たな観点に立つて、現象の本質に喰いつたものを、この立場から考えようとしている（本集報の他の論文参照）。即ち従来の単純な形式的立場を離れて、現象の本質的に肝要な部分を総合的にモデル化する確率過程の統計数理的な理論（単なる定常過程ではなしに、色々に発展する機構を考えてのモデルに関するもの、例えばエピソード——epidemic——のものをも含んでいる）を考えている。より突込んで考えるならば、この中には当然前§で述べた数量化についての考え方や方法も採り入れられることが出来るのである。

一方現象を解明するとき複雑な条件を考えに入れ、場合によつては、行動のダイナミックなパターンをも問題にしてモデルをつくり、逐次的に一步一步予測を行おうとする立場がある。これにおいては、如何に現象をフォーミュレートし、如何なる所に目を付けて行けばよいか、如何なる妥当な実験・調査を如何様に行つて現象構造を精度高く調べて行けばよいか（勿論サンプリングによる誤差・測定誤差なども量的に評価しておく）、その妥当性・信頼性・客観性・伝達可能性・再現性・実際性は十分保持されるようにしなければならない。このようにして得られる多相多元現象を如何に表現し、数量化し（§2,3の考えを十分活用し、予測すべきものを、外的基準によつて明らかにせられるものとして扱う）、如何に分類・総合して的中比率の高い予測を与えてゆか、これらのことが重要な意味を持つてくるのである。即ち現象をフォーミュレートし、要因相互の内的関連性、目的とするものとの関連性を考慮し、この中にメトリックを入れ、高い予測的中比率を得るように多元解析の立場から総合して行くことが中心問題となるのである（単純な要因の力を並列的につみあげるとき全く的中率の低いものが出てしまう。要因間の関係性を考えないからである）。自ずと前者の予測方法と理論の発展方向が異つてくるのである。社会現象や自然現象でも複雑なものは

* ここで述べた様な伝統的方法によるモデルのたて方と要因の数量化の任意性に問題がある場合が多い。前者については、いかなる行動単位に対するモデルか、このように要因を捨象できるか、また行動単位がすべて同一のモデルにしたがうものであるか、それがはたして首肯できるものかに問題がある。この方法は社会現象に関しては、エコノメトリックス・心理学における数理的研究（例えば Mosteller と Bush による Stochastic Model of Learning, John Wiley, 情報理論の心理学への応用の論文、その他数学的心理学の論文）においてよく用いられるものであるが、これらのモデルには疑問のものが多いと思われる。つまり単純な函数関係の仮設作成（純化された条件下での函数関係の積みあげ）に妥当性と言う点で無理があるものと思われるからである。したがつてより批判的に見れば、ここで出されている結果は現実的にはどうでもよい様な事で更に大事な問題が他に沢山あると言えるのではない。また、そこに用いられている量は、本当に操作的にとらえられる量か否かの問題もある。単なる概念を符号であらわし、その関係を類推的に数理的に表現したものが多いのではないかと考えられる。実際には評価方法によるメトリックの入れ方が問題であるにも拘らず、任意に——概念と操作的な量とが一對一に対応しない、一對多の関係が多い——これを入れて、理論と実際の付け合わせが行われているものも多い。この立場は「理論の設定」→「実験による理論の確認」→「理論を用いての予測」、の経路をとるのであるが、実際に大切なメトリックの関係ですべてに筋が立つていなければならない。この一義的対応を付けることは非常に困難な問題であろうし、また理論とその確認も非常に単純なものについてしか実証し得ないことも多いであろう。この単純なものに就ての理論の積み上げで複雑なものを予測しようと試みるのである。然し、単純な場合の積み上げで複雑なものが予測出来ると考へるのは大変な誤り（前§でのべた多次元解析の必要に就て述べた所の脚註の文献参照）である。問題は説明的な理論が大切か、予測が大切かの点であるが、現象予測に役立つ理論は意味がないのであつて、実際的な予測に重点をおいて理論をたたきあげねばならないと考へる。然しこの方法の良さはその抽象性と大局的類推の点にあるのである。この抽象によつて他種の現象が繋ぎ合わされ開眼の態で、予測に有効な思ひぬ理論の発展に資することがある（動物実験の結果とそれの人間に対する結果の予測の問題に似ている。生物的抽象性がこの効用をささえている訳であるので全面的に否定すべきではない。然し、この点社会現象の予測にはこの方法にのみ頼らぬことが肝要である。）

猶この方法の他、単純なものとして傾向性・趨向性・週期性・生長曲線等をさまざまな曲線の当て嵌めに依り見出し、これに依り将来を予測する伝統的方法もあるがこれは条件の斉一性、現在までの状況の素直な推移を仮定しているものである。非常に包括的には前述の立場を行つたものとして考へてよい。

後者を主軸とし前者を参考にするのがよいのではないかと思われる*。

以上の他、予測において、概念的に行く行き方があるが、多くは説明的ではあるが、前向きのものではないので敢てとりあげないことにする。

扱、以上において予測の大略の考え方を述べたのであるが、これらの方法についてやや注意を添えておこうかと思う。例えば予測を行おうとする時、行動単位たる個物（個人）の行動を予測しようとする場合と個物（個人）の集りとしての集団の状況（この時は、行動単位をそれぞれに応じ別様に考えねばならない）を予測しようとする場合とがある。この二つでは方法論的にも異つた近接の仕方が取られることになる。例えば選挙予測について考えてみよう。これに必要なのは誰が当選するか、どの位の票を取るかを知ることであり、選挙人が誰に投票したかを個人別に予測することではない。個人が誰に投票したかは、調査によつて本当のことを知る由もない。そこで予測の方法は自ずと集団的——選挙において公表された（地域別に判明することもある）数字のみが客観的に知り得られるものである——なものとなり、候補者の得票数字そのもの、あるいは順位あるいは当落、を予測することになる。これも個々にではなく衆・参議院の場合なら党派別当選者の予測ということが問題になる場合がある。これがそれぞれ“行動単位”の予測すべきものとなり、この点からの中が取り扱われねばならない。この時には、予測は選挙人たる個人の調査における反応を積み上げるにあつて、集団的に、ある配慮がなされねばならない。調査における回答者の嘘つきの比率とか有名人に対する回答が調査では多目に出るとか、共産党の回答が少なめに出るとか、調査不能の投票に行く比率はどの位とか、一種の社会常数的なものを見出すのが得策となる。これが時期的にどうなるかを調べておくことも大切なことになる。このほかパネル調査の方法の併用（パネル調査の誤差評価も大事である）も考えられるし、前回の投票傾向、工夫を凝らした質問への回答パターンにより公表結果に近いものをくみ上げる仕方、また投票・棄権の動向判断の仕方も手がかりにしようのである。例えば、こうして、候補者をもとにした時（この場合、候補者が行動単位となる）その人に対する集団的調査結果の状況を総合してその当落・得票率等をしつかり予測する仕方も考えられてくるのである。如何に着眼し、如何に調査を行い、その誤差を読み取るか、調査結果をいかに総合するか、また如何なるものを候補者の操作的要因・パターンとしてとりあげ、いかに精度のよい予測式を作成して行くか（例えば候補者を単位にした場合、§2, 3 に述べたような考え方により数量化の技法が有効となり、次に述べる場合に還元される場合もある）、調査時と投票日との状況

* この方法に依るとき複雑なものの中から如何なる要因を考えこれをとりあげ別決するかが非常に重要な問題となる。この為、統計的に見て妥当な事例調査も必要になつて来るし、単純な操作で測定可能な要因を調べ、これらを組み上げて有効な要因を作りあげてゆくことも大切な事と成る。調査方法による誤差も明確に表現出来得る様にして置く事も大切な事である。予測の問題では、如何なる状況変化においても決定力をもつ要因の発見に第一に力を入れる可きであるが、その他、状況変化に敏感であり、それに応ずる仕方を明確に為し得る要因を見付けることも次に大切なことである。予測しようとする時点前に知りえられる要因、時点後にはじめて知りえられる要因とを分けておくのも大切である。また、個人的要因、環境的要因、その交錯による文模様の状況、社会的・文化的要因、心理的要因、自的要因、他的要因、われわれの変化させざる要因、変化させ得ない要因等々に分けてその取扱い方を考えることもある。これらは問題に応じて構成されなければならない。要因の安定性の問題、これらを慮つた上で前§で述べた要因間の相互関連性・相互制約性、要因の数量化も考えねばならない。ここでは要因の利き方の単調性・非単調性も検討を加えて考察しなければならないし、態度の様なものも量的に取り扱われねばならない。またこの立場で予測しようとするれば緻密な計画でフォローアップもくり返さねばならない。また非合理的な姿を示しているならば非合理的な儘に取り扱つて単純な理論で割り切らないのも大切である。「われわれの論理」ではなく、“行動単位”の持つている論理の儘に解析を施すことが大事である。前述の予測方法との大きな差異である。これらの方法の筋としては、繰り返して調査や外的基準ある場合の数量化の方法（予測すべき現れそれ自身を“外的基準によつて表現されること”と替へればよいのである）。統計推論の考え方が採用されるのである。ここでは理論構成はあくまでも実際の予測的中を第一義と考え、考えの仕方は多元的であり、理論・概念も操作的な要因のパターンとして把握し、表現し、構成し、それを機能的に取り扱うのである。ここで述べる様な考え方は概念的に純化された簡単な場に於て、考えを進める——制約された条件下で実験によつて検証されたとしようと——単純な理論に拘み尽され無い多くのものを含んでいる。ここを支点として更に分析を加え、考察を深めることになるのである。

の違いはどのように評価するか、調査時点の動きと調査結果の推移をいかに最終目標に対して評価するか、これを集団的に如何に組み上げるか、これらも、年を追ひ、回を重ねて研究しなければならぬ。これに似た集団的な予測には景気動向の予測や集団的にみたまある意見のパターンや商品の市場予測、伝染病の伝播の状況、確率過程的モデルによる集団の予測等があげられる*。ただしこの種のものであつても、問題のフォーミュレーションによつて次のような取扱いも可能になつてくるのである。学生の能力の進歩の状況、病気の予後の予測、結婚の予後の予測、犯罪者の将来予測、選挙の場合でも、取り上げ方を変えて如何なる要因を持つ候補者が当選するかを予測するような場合等の問題は集団的にではなく、個物個物（個人個人）がいかに判定を与えられ、それがいかに的中するかが大事なことになつてくるのである（行動単位は個物または個人）。しかもこの場合、個人個人・個物個物の将来は必ず突き止められるのである。ここが大切な所である、このような場合は個人（個物）および個物（個人）を取り巻く要因、それとの交錯し互に影響しあう状況等肌理細かい要因の分析、それら総合が問題になつてくるのである。さらに面倒な配慮が必要なのであつて、少なくとも前述した態度測定法や数量化・多元解析の方法・統計推論の技法が十分に駆使されねばならなくなつてくる。これは評価方法を如何に作成するか、その評価法に依つて扱えられる諸資料を如何に重みづけかつ総合して行くか、その精度は如何に表現されるかを考えることになる。問題に応じて自ずとこの二つの予測の面があらうかと思う。問題の処理の仕方ははなはだしく異なつてくるものである。行動単位たる個物（個人）の定義の仕方は色々あり、前述の様に同じ現象でも二つの取り上げ方は勿論可能となる場合もあり、われわれの目的に対してより適切なものを選択すべきことを繰返し注意しておこう。素朴に言えば、前者は個物（個人）の積み上げた結果を問題にし——個物（個人）の行動の予測は不問なのである——後者は積み上げではなく個物（個人）そのものの予測が問題でありその積み上げは予測の効率において示されるのである。

以上総括的議論を繰返してきたが、最後に、予測に関しての精度について、もう一度考えを整理しておこう。予測的中の比率が 90% というとき、これは予測を与える全行動単位中その予測的中する確率が 0.90 と言うことであるが、この 90% のサンプリングによる分散を考えに入れること（大サンプルではほとんど問題がないものとなる）——如何に予測するかを具体的に決定するために用いる過去の資料をサンプルと考え、これによる分散を考慮に入れる必要がある。この分散を用いた的中確率の最低も考慮に入れること——、これはわれわれの採つた予測方式を同一母集団に対して適用したとき得られる的中比率の安定性を考えることである。次に、予測的中比率そのものも母集団における確率モデルとして表現されるのであるが、この大小は資料獲得におけるサンプリングの問題としてよりも、要因の採り上げ方、この要因パターン総合の方法によるものなのである。さらに、この予測方式そのものの恒常性がどの程度保持されるか否か、つまりこの予測方式を適用すべきものが同質的な母集団であるか否か、を確かめねばならない。予測的中比率の値そのものが低ければ安定性があつても意味はない。予測的中比率の値が高く、この安定性のあること（その分散の小さいこと）、適用すべき同質的な母集団の範囲の可及的に広がるべきこと、これらが予

* このことは古典物理学の気体分子運動論に似ている。気体分子個々の行動は知る由もないが、これに対してある確率論的モデルを建てることにより、そこからわれわれが観測出来る様な物理量を計算する事が出来る。これと実験との対応が問題になる。この様な仕方は統計の問題でも多く、エビデミックに於ける現象数量化に対する確率的モデルによる予測、交通現象に於ける確率過程的取扱い、待合せ行列(queue)の問題、その他過程事象数量化に関する確率論的・統計的取扱い等に現れる。この場合、その仮定したモデルが現象に適合しない場合もあるが、そこから計算されるものがわれわれが観測出来る量と結果的にあまり差異が無いことも多い。仮定したモデルの仮説自身が妥当であるか否かを、われわれが直接検証出来る様な場合はこれを行い、もしそれが不相当とする場合は当然反省しなければならない（これが気体分子運動の場合の様に検証出来ない場合は止むを得ない）。例えばこれにはエビデミックの問題の単純な取扱い方や態度測定に於ける確率的反応モデルの場合等挙げれば切りがない。然しそれが適切でない場合でも結果的にはよく合致している場合もあるので、その近似の妥当性については考察を進めなければならない。

測の精度が高く妥当性が大きいということ、つまり、前にも述べた有用な予測であることを示しているのである*。

統計数理研究所

参考文献

此処では、この論文の主題に関するものでも、自分の気持や立場にびつたりせず、また何等かの意味でわれわれの研究を進めるのに寄与していないものはすべてこれを除外した。然し従令、実際に研究をして行く立場は反対でも、その考える仕方そのものに感銘を受けたものは記しておいた。この方面に関する文献はまだいくらもあり、その網羅的文献集をつくる人は他に適任者があろう。以下は私の覚書に残しておきたい文献である。

数量的表現をこころみたものとして、われわれにとつて興味あるものとしては次のものをあげておこう。

J. von Neumann: *Gesellschaftsspiele*, *Math. Ann.* 1928, (この紹介は遊戯論視見として講究録, 第3巻 第6号にある) これ以後、盛に論じられているがわれわれにとつて大事なことはここに述べられていることに尽きている。

R. von Mises: *Probability, Statistics and Truth*, Hodge, 1939. この新版も McMillan から出ている。 *Wahrscheinlichkeit, Statistik und Wahrheit*, Springer (Wien) も同様な内容である。この系統の文献は林知己夫, サンプルング調査はどう行ふか, 東大出版の付録に詳しい。また *Positivism*, Harvard Univ. Press も面白い。

E. Tornier: *Wahrscheinlichkeitsrechnung und Allgemeine Integrations Theorie*, Teubner

B.O. Koopmans: *The Axioms and Algebra of Intuitive Probability*, *Annals of Mathematics*, Ser. 2, Vol. 41; *Intuitive Probabilities and Sequences*, *Annals of Mathematics*, Ser. 2, Vol. 42

A.H. Copeland: *Probabilities, Observations and Prediction*, *Third Berkeley Symposium*, これは proposition を要素とした finite sequence of probability の理論で実際的に興味がある。これと一応関聯づけてみる必要のあるものに H. Blume, *Mathematische Begründung und Entwicklung einer Wahrscheinlichkeitsrechnung mit Finiten Kollektiven*, *Zeitschrift der Physik*, Vol. 92 及び *Zur Anwendung der Wahrscheinlichkeitsrechnung mit Finiten Kollektiven*, *Zeitschrift der Physik*, Vol. 94, と前掲の Koopmans の論文がある。

J. L. Savage: *The Foundations of Statistics*, John Wiley

K. Arrow: *Social Choice and Individual Values*, John Wiley

能度測定に関しては次のものが直接われわれを裨益しわれわれに感銘を与えた。

L.L. Thurstone: *Vector of Mind*, Chicago Univ. Press, *The Measurement of Attitude: A Psychological Method*, with E.J. Chave, Chicago Univ. Press

L. Guttman: S. Stouffer 編, *Measurement and Prediction*, Princeton Univ. の中の尺度解析の部分 (*A Basis for Scaling Qualitative Data*, *American Sociological Review* Vol. 9, の他従来多くの論文があるが、これに一通りまとめられてある); P.F. Lazarsfeld 編, *Mathematical Thinking in Social Research*, Free Press, 中の態度の主要素, ラデックス (radex) 理論についての部分; Israel Institute of Applied Social Research から出されている態度の主要素, 及びファセット理論に関する実際の報告 (ガリ刷)。

その他, Lazarsfeld の潜在構造分析 (上述の二冊の書の中にある) 及び Rutger Univ. Press の Riley, Toby 編, *Sociological Studies in Scale Analysis*

数量化に関しては次のものが直接われわれの考えを刺戟し、且つ感銘を与えている。

H. Horst 編: *The Prediction of Personal Adjustment*, Bulletin No. 48, Social Science Research Council, New York

L. Guttman: *An Approach for Quantifying Paired Comparison and Rank Order*, *Annals of Mathematical Statistics*, Vol. 17, No. 2

* 社会現象の予測に関する文献の一部は林, 予測, 現代教育心理学大系, 第9巻, 中山書店にある。

W.S. Torgerson: *Multidimensional Scaling I*, Psychometrika, Vol. 17 (彼の書 *Theory and Methods of Scaling*, John Wiley, 中の第 11 章として詳しく論じられている).

因子分析法として Thurstone の factor analysis の書も面白いものであるが, 方法的には P. Whittle: *On Principal Components and Least Square Methods of Factor Analysis*, Statistcs Upsala, Vol. 9, に面白い所がある.

われわれがこの方面に於て新しい考えの下に行つたものの参考文献をあげておこうと思う.

(イ) 態度測定

新しい方法に関して

林, 池内, 水原, 大塩, 佐野: 態度数量化の一方法 I, 彙報, 第 2 卷, 第 1 号, これについては高木貞二編, 心理学における数量化の研究, 東大出版, にある.

林, 高倉, 牧田, 齋藤: 態度数量化の方法 II, 彙報, 第 6 卷, 第 1 号

林: 尺度点(目盛り)決定, 測定基準作成に於ける一つの統計的考え方, 彙報, 第 6 卷, 第 1 号

(ロ) われわれの数量化理論とその応用に関しては

C. Hayashi: *On the Quantification of Qualitative Data from the Mathematico-Statistical Point of View*, Ann. Inst. Stat. Math. Vol. II, No. 1, *On the Prediction of Phenomena from Qualitative Data and the Quantification of Qualitative Data from the Mathematico-Statistical Point of View*, Ann. Inst. Stat. Math. Vol. III, No. 2, *Multidimensional Quantification I, II*, The Proceedings of the Japan Academy, Vol. 30, No. 2, No. 3, *Multidimensional Quantification*, Ann. Inst. Stat. Math. Vol. 5, No. 2, 数量化と応用例 I, II, III, IV. [彙報 Vol. 2, No. 1, Vol. 4, No. 2, Vol. 5, No. 1, Vol. 6, No. 1, 現象解析に於ける統計的方法, 科学基礎論研究, 第 3 号

林知己夫・植松俊夫: 多次元解析・数量化理論とその応用, 統計数理研究所養成所講義録その他工業統計・教育統計にもこの種のものがある.

西村克彦・林知己夫: 仮釈放の研究, 東大出版

水野 坦・林知己夫・青山博次郎: 数量化と予測, 丸善出版

青山博次郎: 数量化の一問題, 講究録, 第 8 卷, 第 4 号

青山博次郎: 数量化に於ける標本誤差, 彙報, 第 2 卷, 第 2 号

H. Aoyama: *On the Evaluation of the Risk Index of the Railroad Crossing*, Ann. Inst. Stat. Math. Vol. 10, No. 2

崎野滋樹・巷野悟郎: 数量化による予後の判定, 彙報, 第 1 卷, 第 1 号

崎野滋樹・巷野悟郎: 症状による病名の予測, 彙報, 第 2 卷, 第 1 号

S. Sakino and G. Kono: *On the Forecasting of Prognosis in Pediatrics by a Quantifying Method*, Ann. Inst. Stat. Math., Vol. 6, No. 2

小池重夫・他: 発塵作業場の衛生学的評価, 労働科学, 第 30 卷, 第 12 号

山中二郎: *Readership Score* 予測の一方法, 調査と技術, 4 月号, 1958.

日野原昌: 累犯少年に関する研究, 司法研究報告書, 第 8 輯, 第 8 号.

上田尚一: 審判後に於ける少年の成績とその統計的予測, 最高裁判所; 少年審判における統計的予測の方法, 最高裁判所

久保舜一: 学力調査, 福村書店

久保舜一: 双生児法による形態盤曲線検査法の吟味(国立教研紀要)

国立国語研究所・統計数理研究所: 敬語と敬語意識, 秀英出版

佐藤敬之輔: 好きときらい(1)-(4)——調査と分析の小手調べ——日経広告手帖; 日本字デザイン, 丸善出版

朝日新聞社世論調査室: 昭和 33 年 5 月総選挙, 選挙調査報告書

池内・林・藤本・佐久間・竹内・岡田: 事実確認法によるメディア内容選択の立体的分析, 朝日新聞東京本社広告部

橋爪・梅津・中村・林: *Troublesome Hospital Managers Problem*, 第 54 回生物統計学会例会

(c) 計算法に関して

石田正次：数量化に於ける計算法，統計数理研究所輯報，第8号

多賀保志：リレー計算機による線型計算について，彙報，第5巻，第1号

H. Akaike: On a Computation Method for Eigenvalue Problems and its Application to Statistical Analysis, Ann. Inst. Stat. Math., Vol. 10, No. 1

T. Uematu: Note on the Numerical Computation in the Discriminant Problem, Ann. Inst. Stat. Math., Vol. 10, No. 2