# アンケート調査における漠度概念とエントロピー分析

廣田　　薫

## 1. はじめに

一般大衆あるいは特定の集団に属する人々が，ある事柄に関し，どのような見解をもっているかを知りたいとき，アンケート方式の調査がよく行なわれていることは，周知の事実である．新聞やテレビ等のマスコミでもよく用いる調査方式ではあるが，自分がその被験者になり好意的に回答をした場合でも，あんな回答が正しいデータとして全体に反映しているのかと，不信の念をいだいた経験を有する読者諸兄も少なくないと想像する．

最近，一例として，研究室の学生が，就職の推薦書の一部として，その学生の人物調査アンケート用紙に記入を求めてきた．その一部分は，図1に示すように，その学生が行動派か否かを，5値で評価するようになっているものである．もちろん，好意的に回答を寄せたいとは思うが，ここで次の2つの疑問をもつに至った．

①その学生の表面的態度しか知識がなく，正しい評価ができかねる．すなわち，正確な回答をするための情報が不足しており，曖昧であること．

②それにもかかわらず，5値のいずれかを無理に選ぶとすると，評価の状態数が5値と細かすぎて，同一の質問を再度うけても同一の評価をするか否かに関して自信がもてないこと．一般にアンケート調査では，簡単なものは，yes と no の2値だが，学術的なものは3値・5値あるいは7値等を用いる．状態数が多ければ詳しい情報が得られると錯覚しやすいが，情報が曖昧な場合は，逆に再現性に乏しくなる危険があるのではないか？

以上2つの疑問は，この例のみならず，一般のアンケート調査でも生ずるものであろう．本稿では，この2つの問題点に関して，1つの理論的解析を試みる．基本となる数学的理論は，確率論，Zadeh の Fuzzy 理論［1］，

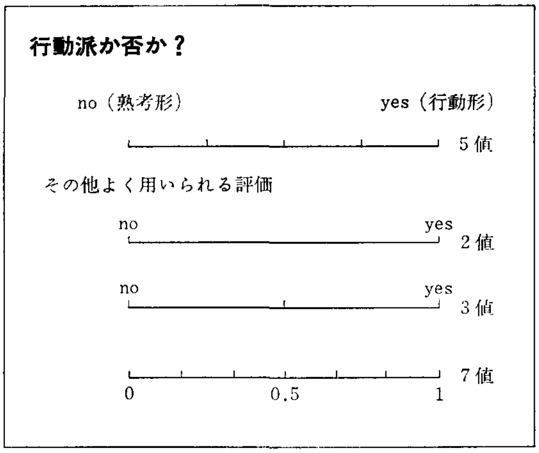


**図 1**　アンケート調査の一例

Shannon の Entropy 理論(情報理論)［2］，および，それらを総合的に取り入れて構成された確率集合論［3］等のいわゆる曖昧さの理論である．以下の議論は，束論や測度論の観点から詳しく展開することも可能だが，考え方を中心にするため，数学的厳密さは重視せず，本質を失わぬ程度に留めておく．

上記第1の疑問点に関しては，漠度関数（vagueness function）という新しい評価尺度を導入することにより，曖昧な情報でもより正確に表現できるという意味で，解決できることを示す．また，第2の疑問である状態数に関しては，(主観)エントロピーの観点から最適な状態数を決定し，3値が最適であり，5値も良好であるが，7値になると2値よりもかえって悪くなることを示す．

## 2. 漠度(vagueness)概念の提案

最初に，議論を定量的にするため，評価の数値定量化を試みる，図1に示すように，肯定的(yes) の場合を数値の1に，逆に，否定的 (no) の場合を数値の0に対応

ひろた　かおる　相模工業大学　情報工学科

させ，それ以外の曖昧な中間状態は，0から1までの中間値に対応させることにする．この [0, 1] 区間への数値定量化は，Fuzzy 集合論［1］のそれと同一であり，一般によく用いられている．

被験者が，具体的な質問を受けた場合，しかもその質問に対して曖昧な評価しか持ち合わせていない場合，評価値は [0, 1] 区間上でばらつくはずである．そこで，そのばらつきの度合いを確率分布でとらえることにする．すなわち，被験者の下す評価は，[0,1] 区間上の確率分布に従い，具体的な評価値はその1つの標本値として得られるものと考えるのである．したがって，同一の質問に同一の被験者が，場合によっては異なる評価を下すことも当然考えられることになる．そこで，被験者が有する [0,1] 区間上の確率分布を $f(\alpha)$ で表わすことにする．

$$\int_0^1 f(\alpha)d\alpha=1,\ f(\alpha)\geqq 0 \tag{1}$$

もちろん，この $f(\alpha)$ は，被験者が潜在的に有するものであり，存在は仮定できても，外部から明確にはとらえることができないものと考えるのが自然である．$f(\alpha)$ がデルタ関数による一点分布のようにシャープな形をしている場合は，常に同一の評価が得られるので問題はないが，逆に一様分布のように裾の広がったものであれば，調査をやるたびに別の評価がなされるという，データ収集側からはきわめて憂慮すべき事態も予想される．現実には，この両極端の中間の分布をもつ場合が多いのであろうが，いずれにしても，得られた評価値を鵜呑みにして，計算機処理をやったのでは，信憑性に乏しい結果しか得られない．そこで，正しい評価を得るために，1カ月程度の期間をおいて，同一被験者に数回の繰り返し調査をする方法もとられているが，手数と無駄という観点から，あまり推奨できるものではない．

以上のモデリングにおいて大切なことは，確率分布 $f(\alpha)$ のもつ情報を，正確にかつ適切に知ることである．そのために，モーメント解析を展開する．まず，$f(\alpha)$ が [0,1] 上の確率分布であることから，すべての次数のモーメントが存在することに注意しよう．

$$m^n=\int_0^1 \alpha^n f(\alpha)d\alpha \qquad n=0,1,2,\cdots \tag{2}$$

次に，$f$ から $\varphi$ への(3)式による変換

$$\varphi(t)=\int_0^1 e^{it\alpha}f(\alpha)\,d\alpha \tag{3}$$

を考えれば，逆変換が存在して，

$$f(\alpha)=(1/2\pi)\int_{-\infty}^{\infty} e^{-it\alpha}\varphi(t)dt \tag{4}$$

で与えられるから，$f$ と $\varphi$ の対応は1対1であり，同一の情報量をもつことになる．この $\varphi(t)$ については，

$$d^n\varphi(t)/dt^n|_{t=0}=i^n m^n \tag{5}$$

が成立し，さらに常に，

$$\varphi(t)=\sum_{n=0}^{\infty}\frac{i^n}{n!}m^n t^n \tag{6}$$

と，Taylor 展開が可能であることが示される．(これらは，確率測度論を用いた厳密な証明も可能である［4］.)

また，ここで級数を中途で打ち切れば，たとえば $n=2$ までで打ち切ると，

$$\begin{aligned}\varphi(t)&=1+im^1t-(1/2)m^2t^2+0(t^2)\\&=(1/2)\{1+im^1t)^2+(1-vt^2)\}+0(t^2)\end{aligned} \tag{7}$$

と近似表現される．(6)式により，すべての次数のモーメントを与えれば，$\varphi(t)$ が再現できるので，$f(\alpha)$ のもつ情報とすべてのモーメントのもつ情報が等価なことがいえた．

$$\begin{aligned}&f(\alpha)\leftrightarrow\varphi(t)\leftrightarrow\\&\{m^1,m^2,m^3,\cdots\}\leftrightarrow\{m^1,v=m_0{}^2,m_0{}^3,\cdots\}\end{aligned} \tag{8}$$

(ここで，$m_0{}^n$ は平均値まわりの $n$ 次モーメント) しかも，重要な情報は，平均や分散などの低次モーメントに集約していることが，理論的にも［4］経験的にもわかっている．

以上では，確率の観点からモーメントを導出したが，逆に最初に適当な条件のもとで，

$$\{M,V=M^2,M^3,M^4,\cdots\} \tag{9}$$

なる可算個の情報を与えることによって，1つの [0,1] 上の確率分布を導出できることが示せる［4］．これは，分布を議論の出発点とする確率論の立場とは異なるものであり，先験的にメンバーシップ関数が与えられるものとして議論を展開する Fuzzy 集合論の考え方に近い．(9)は，確率集合論［3］の拡張Fuzzy表現 (extended fuzzy expression) とよぶものであり，$M^n$ は $n$ 次のモニターとよばれる．モニターは，理論上は可算個の集合から構成されるが，重要な情報をもつのは低次のモニターであり，特に，$M$ は級格関数(membership) $V$ は漠度関数(vagueness function) とよばれ，[0,1] の値を取る．

級格関数の値は，そのクラスに属する資格の程度(membershipness)を表わし，漠度関数の値は，級格関数の値をどの程度漠然と (あるいは，はっきり自信をもって)示唆したかを示すものである．(たとえば，図1の例において，100％の自信をもって行動形と評価したときは，級格関数=1，漠度関数=0を指定することになり，逆にまったくわからないときは，級格関数=0.5，漠度関数=1のように指示することになる．)

以上の観点に立てば，これまで普通に行なわれてきたアンケート調査では，級格関数の値のみ答えてもらったことになる．本稿では，被験者が答えた評価値がどの程

度の自信をもって答えたものかという，第2の情報をも回収しようということを提唱するのである．こうすれば，たとえば図1の調査において，被験者がその学生をよく知っていて，自信をもって熟考形とも行動形とも思わなかった場合は，0.5という値を漠度0で答えればいいわけだし，逆にその学生についてほとんど知識がなくて，中間と答えておけば無難と考えたときは，0.5という値を漠度1で答えればよいことになる．ここで，前者の場合は，従来の級格関数のみを答える方式で，信頼性をあげるため1カ月程度の間隔でくり返し調査をしても，再び同一の0.5と答えるだけであるし，逆に後者の場合は，1カ月後の再調査で知識が増えて1.0と答えを変更するかも知れない（そのときの漠度は，たとえば0.5）．さらにもう1カ月後の再々調査においては，その間に学生との深いつき合いがあって，さらに意思変更をして0と答える可能性もあり得る(そのときの漠度=0)．このように，漠度という第2の情報も提供することにより，被験者にとっては，はじめに述べた第1の疑問をさほど感じないで，調査に“協力”することができるし，調査をする側にとっては，どの程度の確信で答えてくれたのかがひと目でわかるうえに，くり返し調査のような面倒な手数をかける必要もないという利点をもつことになる．

なお，従来の方法においても，被験者が質問に関して知識がないときは無回答にするという方式をとれば，第1の疑問をある程度解決できることになるが，漠度を3値以上の多値で質問することによりさらに詳しい情報を回収できることになり，さらに質問事項が多いときは，被験者が無回答にしたのか，あるいはミスによる記入漏れなのかをも，明確に区別できることになる．

確率論の立場にのみ固執すれば，本稿で提唱している級格関数は平均値に，また漠度関数は分散に対応する概念にすぎないという批判も可能であろう．しかし，被験者の心の中に潜在する確率分布 $f(\alpha)$ の，真の“平均値”と“分散”を，“級格関数”と“漠度関数”がそれぞれ示している必要は必ずしもない．それよりも，むしろ被験者の意思あるいは主観による評価値が優先する．くり返し調査によって，$f(\alpha)$ の平均値や分散を“推定”しようという“確率”分布を大前提とした受け身の立場でなく，積極的に級格や漠度を評価するという能動的な立場からの主張であることに注意したい．理論的橋渡しとして，平均や分散などのモーメントの解析法により，確率分布と可算個のモニターとの対応をつけた．また，被験者の潜在意識の世界においては，両者が一致しているのであろう．しかし，いずれにしても，現実にはそのすべての情報に介入することはできず，近似情報に甘んじなければならない．Fuzzy(あるいはその拡張)概念を確率概念より導出するという意味においては，まったく新しい概念を提唱しているわけではないが，(拡張) Fuzzy 概念を議論の出発点にすれば，必要以上の煩雑さをもち込まずに，理論展開の見通しをよくし，しかも応用上きわめて有効であることを示した．さらに，Zadeh による Fuzzy 概念［1］に，本質的に欠けている漠度概念の重要性をも指摘した．この意味で，確率概念のみに固執したり，Fuzzy 概念のみですべてを押し通そうとする考え方から脱皮し，現実の場面にふさわしい，新しい考え方の芽生えを示唆したことになる．

## 3. 最適状態数の決定

前節においては，[0,1] の実数連続区間に数値定量化を行ない，級格関数の値も漠度関数の値も，ともに，0から1までの値を自由に取れるものとした．しかし，計算機等による数値データ処理においても，また回答を要求される被験者にとっても，0から1までの数値を自由に選択することよりも，むしろ，図1に示したような，3値あるいは5値等の，[0,1] 区間のうちであらかじめ指定された有限個の数値(評価値)のうちから1つを指示するほうが馴じみやすい．その場合，あらかじめ指定する評価値の個数(状態数)をいくつにすべきかという点に関しては，あまり考察されていないようである．あまりに状態数を多くしても，被験者が選択するには迷いが生じてしまうので，調査をする側で経験にもとづいて，適当な状態数を決めて回答用紙を作成し，被験者はそのうちのいずれかをチェックすることにより回答をするという場合が多いようである．ここでは，曖昧さの指標としてよく知られた Shannon のエントロピー概念［2］を応用した主観エントロピーを用い，最適状態数の決定を試みる．

評価値として許容される数値の集合を，上記の理由により，[0,1] 区間の有限集合とし，$\{\alpha_i\}_{i=1}{}^n$ と記すことにする．

$$0 \leqq \alpha_1 < \alpha_2 < \cdots < \alpha_n \leqq 1 \qquad (10)$$

すると，被験者による評価の多様性を反映するために導入した(1)式の $f(\alpha)$ は，総和の形式で書けることになる．

$$\sum_{i=1}^{n} p_i = 1, \quad p_i \geqq 0 \qquad (11)$$

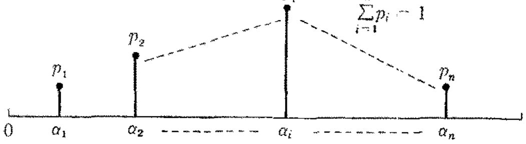


図 2　有限個の評価値とその選好確率分布

ここで，$p_i$ は，被験者が評価値 $\alpha_i$ を選ぶ確率である(図2).

くり返すことになるが，(11)式は，被験者が与えられた質問に関して潜在的にもつ評価の確率分布を与える．そこで，その確率分布を，確率空間の記号を用いて，より詳しく記述しよう．すなわち，全体の確率空間を $\Omega$ とし，その部分集合 $\Omega^i$ が評価 $\alpha_i$ に関与する確率 $p_i$ を与えるものとする．

$$\Omega=\bigcup_{i=1}^{n}\Omega^i,\ \ \Omega^i\cap\Omega^j=\phi\ \ (i\neq j) \tag{12}$$

$$P(\Omega^i)=p_i \tag{13}$$

この確率空間 $\Omega$ は，確率集合論［3］では，パラメータ空間 (parameter space) とよばれる．パラメータ空間のより詳しい構造を記述するために，次に純粋状態とよばれる概念を導入する．

(11)式の分布をもつような，被験者 A の多様な評価は，パラメータ空間 $\Omega$ から $[0,1]$ への(詳しくは，その部分集合 $\{\alpha_1,\alpha_2,\cdots,\alpha_n\}$ への）写像 $\mu_A$ としてとらえることができる．

$$\begin{array}{rcl}\mu_A:\Omega & \longrightarrow & [0,1] \\ \cup\!\!\shortmid & & \cup\!\!\shortmid \\ \omega & \longrightarrow & \mu_A(\omega)\end{array} \tag{14}$$

つまり，1つのパラメータ $\omega(\in\Omega)$ が指定されるごとに，1つの評価値 $\mu_A(\omega)\,(\in\{\alpha_1,\alpha_2,\cdots,\alpha_n\}\subset[0,1])$ が定まるものとし，被験者がどのパラメータをどの程度頻繁に採用するかは，(11)・(13)式にしたがうものとするのである．したがって，

$$\mu_A{}^{-1}(\alpha_i)=\Omega^i(\subset\Omega) \tag{15}$$

等が成立していることになる．

被験者が異なれば，当然別の評価をすることが期待されるが，その別な評価もまた(14)式と同様な写像で記述されることになる．たとえば，評価者Bの評価を $\mu_B$

$$\mu_B:\Omega\longrightarrow[0,1] \tag{15}$$

と記すことにする．常に同一の評価値を指定する被験者の評価は，定数値写像で与えられることになり，評価の単純さあるいは複雑さは，(14)・(15)式の写像の構造に依存することになる．

ここで，(14)・(15)式の2つの評価 $\mu_A,\mu_B$ に対して，$\lambda$ 和［3］とよぶ演算により生成される新しい評価 $\mu_C$ を考えよう．

$$\mu_C(\omega)=\lambda\mu_A(\omega)+(1-\lambda)\mu_B(\omega)\ \ (\lambda\in[0,1]) \tag{16}$$

この $\mu_C$ は，二通りの評価 $\mu_A$ と $\mu_B$ をもとにして，それぞれを $\lambda$ 対 $(1-\lambda)$ の比でとり入れて合成された新しい評価といえる．他方，(16)式を逆に解釈すれば，$\mu_C$ と

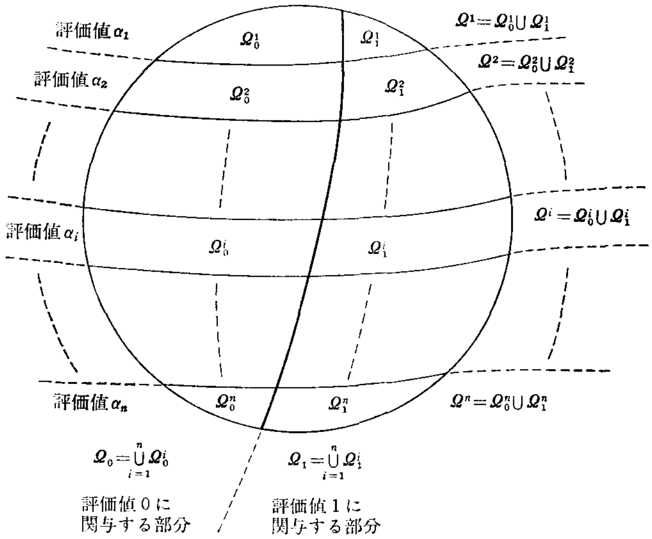


**図 3** パラメータ空間 $\Omega$ の $2n$ 分割

いう評価の原因を追求すると，2つのより基本的な評価 $\mu_A$ と $\mu_B$ に分解され，それらの合成による結果として得られたものとみることができる．一般に，多様で複雑な評価であっても，それを分解してみれば（その分解の操作は，被験者個人の精神構造にまで介入することになるから，不可能かも知れないが，仮に分解できたとすれば）より基本的な評価があり，それらが合成されたものとみなすことができよう．そこで，それ以上より基本的な評価に分解できないという，最も基本的な極限状態になっているとき，その評価は，純粋状態 (pure state) にあるとよぶことにする．

この純粋状態の具体的構造を調べてみると，評価値が0か1しかとらない評価であることが証明される（文献［5］の定理1）．すなわち，yes(=1)・no(=0) のいずれかしかとらない評価が最も基本的なものであり，一般の曖昧な中間値の評価は，それらが(16)式のようにして合成されたものであるという，きわめて自然な結果を，理論的に導いたことになる．

そこで，(10)〜(13)式で，評価値 $\alpha_i$ が出現頻度確率 $p_i$ で得られる原因となるパラメータ空間 $\Omega$ の部分集合 $\Omega^i$ を，（被験者の精神構造にまで介入したモデルを考えて）さらに二分し，評価値1に関与する部分 $\Omega_1{}^i$ と0に関与する部分 $\Omega_0{}^i$ の2つから構成されているものとする．

$$\Omega^i=\Omega_0{}^i\cup\Omega_1{}^i\ \ (\Omega_0{}^i\cap\Omega_1{}^i=\phi) \tag{17}$$

$$P(\Omega_1{}^i)/P(\Omega^i)=\alpha_i \tag{18}$$

すなわち図3のように，パラメータ空間 $\Omega$ が互いに素な $2n$ 個の部分に分割され，

$$\Omega=\Omega_0\cup\Omega_1 \tag{19}$$

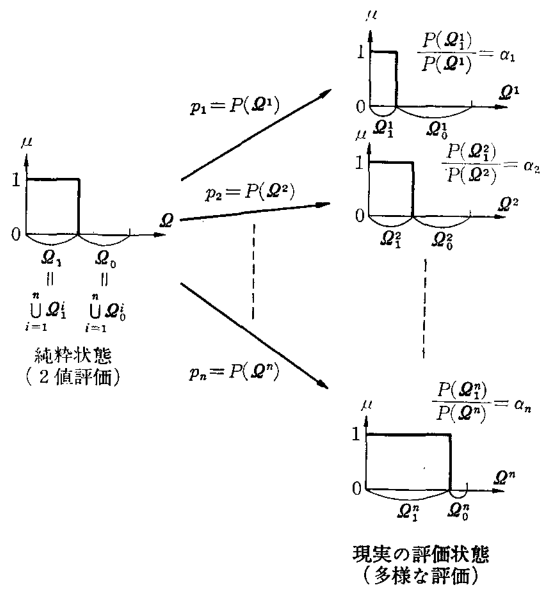


図 4　分割されたパラメータ空間の集約平均化により得られる多様な評価

$$\Omega_0=\bigcup_{i=1}^{n}\Omega_0{}^i,\quad \Omega_1=\bigcup_{i=1}^{n}\Omega_1{}^i \tag{20}$$

各 $i$ に対して(17)式のようにパラメータが集約され，その中で平均化され，その結果(18)式の評価値 $\alpha_i$ が (13)式の頻度 $p_i$ で得られると考えるのである(図4).

図4において，集約平均化の前段階においては $\{0,1\}$ 2値の純粋状態にあり，集約平均化された後は状態数が $n$ で複雑な評価をすることになるが，ともに級格関数の値 $M$ は，

$$M=P(\Omega_1)=\sum_{i=1}^{n}P(\Omega_1{}^i)=\sum_{i=1}^{n}\alpha_i p_i \tag{21}$$

と一致していることに注意しよう.

この場合，全体としては $\alpha_1$ から $\alpha_n$ までの $n$ 種類の評価値が可能になり，評価も複雑になるが，その複雑さの根元は(19)・(20)式あるいは図3のパラメータ空間 $\Omega$ の $2n$ 分割に依存することになる.

一般に，曖昧な確率事象の曖昧さを数量化する評価尺度としては，Shannon によるエントロピー概念[2]がきわめてよく用いられている．確率空間が有限分割されている場合のエントロピーは，各々の確率の逆数の対数を平均化したもので定義され，“場合の数の複雑さ”を反映する量を与える．この場合は，(19)・(20) 式の $\Omega$ の $2n$ (有限) 分割は，被験者の精神構造にまで介入して得たものであり，被験者の“主観”にもとづくものであるから，対応するエントロピーを，特に主観エントロピー(subjective entropy)と名づけ，記号 $H$ で表わすものとする．すなわち，(19)・(20)式にもとづく主観エントロピー $H$ は，

$$H=-\sum_{i=1}^{n}\{P(\Omega_0{}^i)\log_2 P(\Omega_0{}^i)+P(\Omega_1{}^i)\log_2 P(\Omega_1{}^i)\} \tag{22}$$

で与えられることになる.

しかし，(19)・(20)式の分割は，被験者の精神構造にまで介入したモデルとして得られたものであり，現実にはとらえることができないから，(22)式はそのままでは計算不可能である．けれども，(13)・(18)式を用いれば，

$$P(\Omega_1{}^i)=\alpha_i\cdot p_i \tag{23}$$

$$P(\Omega_0{}^i)=(1-\alpha_i)\cdot p_i \tag{24}$$

の関係を得て，現実にとらえうる評価値 $\alpha_i$ とその出現頻度確率 $p_i$ のみで (22)式を推定できることになる．すなわち，(23)・(24)式を(22)式に代入整理して次式を得る.

$$H=-\sum_{i=1}^{n}\{\alpha_i p_i\log_2\alpha_i p_i+(1-\alpha_i)p_i\log_2(1-\alpha_i)p_i\} \tag{25}$$

$$=-\sum_{i=1}^{n}p_i\log_2 p_i+\sum_{i=1}^{n}p_i\cdot h(\alpha_i) \tag{26}$$

ここで，$h(\cdot)$ は $\{0,1\}$ 2値の純粋状態の主観エントロピーであり，

$$h(\alpha)=-\{\alpha\log_2\alpha+(1-\alpha)\log_2(1-\alpha)\} \tag{27}$$

yes($\alpha=1$) または no($\alpha=0$) のはっきりした場合は0，D. K. ($\alpha=0.5$) の最も曖昧な場合最大値1(ビット)を与えるよく知られた Shannon のエントロピー関数[2]である．(22)式によりパラメータ空間という確率空間の $2n$ 個の分割のエントロピーによって定義をした主観エントロピーも，(26)式をみれば，$n$ 種の異なった評価値 $\alpha_i$ それ自身がもつエントロピー $h(\alpha_i)$ の出現頻度 $p_i$ による平均値(第2項)と，その出現頻度自身のもつエントロピー(第1項)との和として与えられていることがわかる.

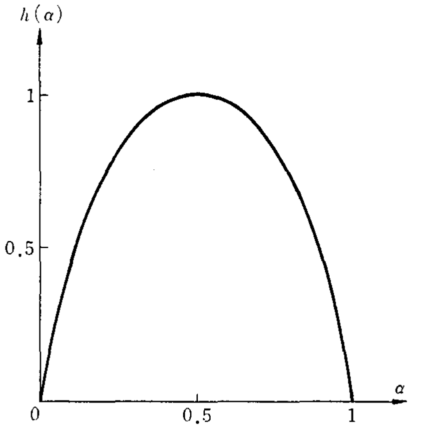


図 5　Shannon のエントロピー関数 $h(\cdot)$
(純粋状態の主観エントロピー)

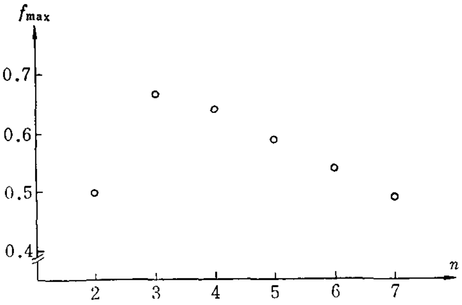


**図 6** 最大平均主観エントロピー $f_{\max}$ と状態数 $n$ の関係

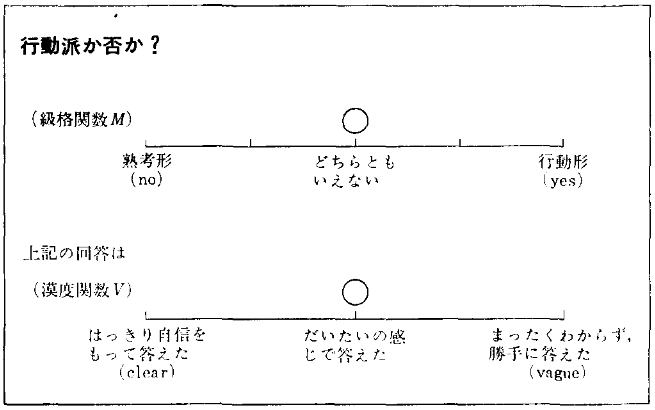


**図 7** 漠度を考慮した 3 値回答方式

この主観エントロピーの定義の妥当性等については，文献 5 にゆずることにして，これをもとにして最適状態数 $n$ を決定することにしよう．

まず，[0,1] 上の評価値 $\{\alpha_i\}_{i=1}{}^n$ の配置に関しては，

$$\alpha_i = \frac{i-1}{n-1} \tag{28}$$

と仮定する．すなわち，$\alpha_1=0$, $\alpha_n=1$ とし，その間を等間隔に分割するものとしよう．そのうえで $n$ を決定するわけであるが，その際の決定基準は，次のように考えることにする．一般に状態数 $n$ が多くなれば，きめのこまかい評価が可能になるが，反面，はじめに疑問点の②で述べたように，被験者にとっては，評価値の隣同士($\alpha_{i-1}$ と $\alpha_i$) の明確な区別がつきにくくなって，評価そのものが曖昧になってしまうというジレンマに陥ることになる．そこで，評価値の状態数を 1 つ増したとき，それが情報量(あるいはエントロピー)の観点からみて，どの程度の価値をもつかに注目することにしよう．すなわち，(22) あるいは (25)・(26) 式の主観エントロピー $H$ を状態数 $n$ で割れば，1 つの評価値あたりの平均的な情報量的価値を与えることになるので，その値の大小により判定することにする．ここで $H/n$ の値は，(11) 式の確率分布 $\{p_i\}_{i=1}{}^n$ によって（しかもその分布のみに依存して)変化するので，$f(p_1, p_2, \cdots, p_n)$ と記すことにする．

$$f(p_1, p_2, \cdots, p_n) = H/n$$
$$= -\frac{1}{n}\sum_{i=1}^{n} p_i \log_2 p_i + \frac{1}{n}\sum_{i=1}^{n} p_i \cdot h\left(\frac{i-1}{n-1}\right) \tag{29}$$

この式の値の大小により $n$ を決定するわけであるが，最も効率よく情報を表現している場合，すなわち $p_1$ から $p_n$ を変化させたときの (29) 式の最大値によって比較をすることにする．(29)式の最大値は，ラグランジェの未定係数法により求めることができる．すなわち，$\lambda$ を任意の定数とし，

$$f(p_1, p_2, \cdots, p_n) = -\frac{1}{n}\sum_{i=1}^{n} p_i \log_2 p_i + \frac{1}{n}\sum_{i=1}^{n} p_i \cdot h\left(\frac{i-1}{n-1}\right) + \lambda\left(\sum_{i=1}^{n} p_i - 1\right) \tag{30}$$

とおき，$i=1 \sim n$ について，

$$\partial f/\partial p_i = 0 \tag{31}$$

を計算すればよい．その結果(29)式の最大値 $f_{\max}$ は，

$$p_i = 2^{h\left(\frac{i-1}{n-1}\right)} \Big/ \sum_{i=1}^{n} 2^{h\left(\frac{i-1}{n-1}\right)} \tag{32}$$

のとき，

$$f_{\max} = \frac{1}{n}\log_2 \sum_{i=1}^{n} 2^{h\left(\frac{i-1}{n-1}\right)} \tag{33}$$

と与えられることが判明する．

この $f_{\max}$ と $n$ の関係を図示すれば，図 6 を得る．この図を一見してわかることは，最適な状態数は 3 値であり，4 値・5 値・6 値とそれに続くが，7 値になると yes・no の 2 値よりもかえって悪くなっているという事実である．

以上の解析結果および前節で述べた漠度概念の提案をもとにすれば，たとえば図 1 で示したアンケート調査は，図 7 で示すような方式で行なえばよいことになる．すなわち，その学生が行動形の人間か否かということに関して，まず最初に「そうである（行動形)」「そうではない（熟考形)」「どちらでもない」の 3 値のうちのいずれかを級格関数の値として選び，次にそれを選んだことに関して「はっきり自信をもって選んだ」「さほど自信はもてないがだいたいの感じで選んだ」「まったくわからないので適当に選んだ」の 3 値のいずれかを漠度関数の値として記入すればよい．漠度まで答えることにより，質問に関しての情報不足あるいは認識不足による不正確なデータを提供する“後ろめたさ”を，解消できることになるし，評価の状態数も 3 値であれば，漠度がかなり大きい場合であっても回答の再現性や信頼性が，か

なり高くなることが期待される．したがって，はじめに述べた2つの疑問点も解決されたことになり，結論として，アンケート調査の標準的な方式は図7のような形式を採用すればよいことになる．

しかし，図7の形式は，あくまで一般的な標準形式として提案するものであり，個々の場合に応じて多少の修正はすべきである．たとえば，質問に関する問題意識がきわめて明確である集団を調査するのに，わざわざ手数をかけて漠度まで答えてもらう必要はない．また，事例研究の結果，ある程度漠度が小さいと予想される集団に，もう少し詳しい調査をしたいときは，級格関数を5値，漠度関数を3値にした調査(図7の破線)も良好である．状態数に関しては，図6を念頭に置いて設定すべきであり，必ずしも3値のみ推奨するわけではなく，場合によっては4値5値のほうが良いこともある．しかし，事例研究の結果からも，7値や9値にするとかえって混乱をまねきやすくなるようである．

## 4. おわりに

アンケート調査に関する問題点を2つほど指摘し，いわゆる"曖昧さの理論"を用いてそれらを解析し，アンケート調査の標準形式ともいうべき方式を提案した．本稿で示した結果を要約すれば，漠度概念の提案と評価状態数の最適化曲線図6の2点になる．

なお，実際に事例研究をやってみると，理論にはなじみにくい，現実の泥臭い問題にも多くぶつかる．たとえば，アンケートを実施しても回答を拒否されたり，回答を得るまでに対人間的な問題点を生ずることも少なくない．また，仮に回答を得たとしても，それが"好意的"なものでない場合もある．すなわち，虚栄心その他の理由で，事実とは正反対の回答を，故意に寄せてくるのである．これらは，いずれも扱いにくいけれど，情報化社会を進歩させるにあたって，ORワーカーが避けて通ることができない問題点であろう．

本稿で駆使することになった"曖昧さの理論"も，避けて通ることができない問題点として，ここ20年ほどの間に脚光を浴びるようになった学問分野の1つである．

一般に，1つの事柄を表現するにも，いくつかの方法がある場合が多い．たとえば，1つの建造物を表現するのに，その実物縮尺モデルによる方法，正面図・平面図・側面図による3面製図による方法，あるいは見取り図による方法等が考えられる．曖昧さという扱いにくい概念も，分布による確率表現，級格・漠度あるいは高次のモニターによる拡張 Fuzzy 表現，そして主観エントロピーによる表現など，いくつかの方法で定量化し記述できることを示した．本稿では，これらを，アンケート調査で"好意的"な回答を寄せてくれる場合に限定し，理論解析に応用したわけである．しかし，上述のもう少し"泥臭く"高度な意思決定問題への応用も，研究されつつあることを指摘して結びとする．

### 参考文献


[1] L. A. Zadeh: "Fuzzy Sets" *Information & Control*, Vol. 8, pp. 338-353, 1965.

[2] C. E. Shannon: "A Mathematical Theory of Communication" *Bell Syst. Tech. J.*, 27, p. 370, p. 623, 1948.

[3] 廣田，飯島："確率集合論の論理学的基礎付け" 電子通信学会論文誌, Vol. 62-D No. 2, pp. 73-80, 1979.

[4] K. Hirota: "Extended Fuzzy Expression of Probabilistic Sets" in Advances in Fuzzy Set Theory & Applications (edited by M. M. Gupta et al.), North-Holland Publ. Co., pp. 201-214, 1979.

[5] 廣田，飯島："単一確率集合の主観エントロピーによる解析" 電子通信学会論文誌, Vol. 62-D No. 2, pp. 81-88, 1979.


●ミニミニ● ●OR●

## 千葉大の頃

電力中研に在籍していた当時，先輩に紹介されて千葉大学で線形計画の話をすることになった．学生は経済学科と数学科から聴きにきていた．僕は学生を数字で評価するのが嫌だったので，max(ノート，レポート，試験) で採点すると宣告しておいた．期末の試験は定式化の問題と単体法の計算問題を出し，皆で相談して答案を作るようにと言った．プロジェクト・チームに入って仕事をする場合には，皆で知恵を出しあうことが必要だからだ．相談して答を出すようにと言うと，皆かえって神妙にとりくむのであった．計算問題をまっ先に提出してさっさと帰ってしまうのは，たいてい経済の女子学生で，最後までやっているのはたいてい数学科の男子学生だった．

評点はほとんど優にしておいたが，1人だけ，ふだんあまり見かけずに試験の時だけ現われて，隣の女子学生の答案をのぞき込んでいる男がいたので良にした．彼も今頃はどこかの課長さんくらいになっているかも知れない． (小野勝章)