工学教育における評価法への計測理論的アプローチ

A Measurement-Theoretic Approach to the Evaluation Methods in Engineering Education

森田 矢次郎

Yajiro MORITA

岡 雄三

Yuzo OKA

渡邊 啓史

Hiroshi WATANABE

著者らは評価の一般理論で話を始める。評価に関する議論のなかで、おそらく最も重要なのは尺度であろう。評価の対象が、ひとりの人間、ある団体、大きな計画などの、単純でない対象であるとき、どのような尺度を用いたらよいのであろうか？

計測科学、計測工学で従来あまり顧みられなかった名義尺度は単純でない、抽象的な計測対象に対して、もっと重要な役割を与えられてしかるべき、と思われる。工学の科目の一部に対して、最近試みられている新しい評価法のあるものは、著者らが、本論文で提唱しようとしている名義尺度、あるいは名義尺度と順序尺度の中間に位置する尺度の実例として解釈できる。

The writers begin with their general view and theory of evaluation. The scale of evaluation is perhaps the most important element in the whole setting of evaluation. When the object of evaluation is a person, an institution, a big project or the like no-simple object, what kind of scales should we employ?

The nominal scale which has been least frequently considered in the science and engineering of measurement, should be given a more prominent place in the evaluation of no-simple, abstract entities. Some new methods of evaluation in engineering courses are practical examples of the writers' measurement-theoretic approach for the choice of an appropriate scale in evaluating students.

1. はじめに

評価とは、われわれが日常遭遇するほとんどすべての事物についての、価値、重要度、量、程度、条件を勘案し、判断することである。われわれは、生まれつき評価の能力をもっているように思われる。嬰兒や乳児の段階から、人間は好悪の感情をあらわにする、これは評価のはじめ、といってよいであろう。

この論文では、筆者らは評価の問題を計測理論の観点でみてゆく。したがって、議論は意識的な評価に限られ乳幼児が玩具や子猫に対してしめす反応は、議論の対象には恐らくならない。しかし、成人が十人の候補者のなかから彼あるいは彼女の生涯の伴侶を決定する、という行為は、確実に評価の結果である、と解釈する。伝統的に評価の大部分は、数量的な尺度でなされてきた。4個の尺度、すなわち、名義尺度、順序尺度、間隔尺度、比例尺度は、久しく知られている^{(1)*註}。これらすべては、実数によってあらわされる。自然科学、工学で今日使われる変数は、数値で表現でき、尺度としては、ほとんどが、比例尺度である、といえる。わが国で慣行化している入

学試験、センター入試などの結果を偏差値で表現するやり方は、間隔尺度を用いている、と解釈され、その結果の合否を決める時には、得点順という、一種の順序尺度がものをいうことになる。これらの、得点の分布が論じられるところでは、正規分布をつくるような問題の出しかたが、理想的である、とされる。したがって『良い』問題とは、生徒の得点が正規分布をしている問題、ということになるが、学力をはかるのに、得点分布の統計的な性質が扱いやすいかどうか、問題の良否を決める、ということの指標になる、というのは、実用的な利点を別にすれば、再考の余地が多分にありそうに思われる。本論文では、間隔尺度、正規分布、などの精密な尺度で学生の学力評価をいかにみてゆくか、という方向ではなく、尺度そのものを学生の学力、というような単純ではない計測の対象に対して、どんなものと考えれば良いのか、ということの問題にする。

*註 よく使われる例を挙げると、名義尺度はスポーツ選手の背番号、順序尺度は碁将棋の段位、間隔尺度は摂氏温度、比例尺度は大部分の物理変数。

実際の学生の能力、努力、可能性、等々を評価し、採

点するということは、そんなに単純なことではない。長年の間われわれは0から100までの直線目盛りでテストの結果を出して、事終われり、としてきた。学力の尺度として、この比例尺度は適当であろうか？やがて一世紀になろうとする膨大な教育学での学力測定の世界や研究⁽²⁾を否定するものでは更々ないが、尺度として自然科学や工学のものさしが、そのまま使われ、その尺度の使用を前提として、問題を考え、点数の意味するところを思案する、という方向には、限界があるように見える。比例尺度は精密すぎ、学生の学力というような、あいまいな対象を測定するにはマッチしないのではないか？

あいまい理論の基本には、順序構造だけを生かした、比例、間隔、にはそぐわない複雑であいまいな対象を測定(ないし表現)しようという考え方がある⁽³⁾。順序尺度を学力の評価に使おうとすると、メンバーシップ関数をどうきめ、どう使うかで従来にならぬ解釈で評価ができる可能性がある。

本論文では、順序尺度よりさらにゆるい位相構造をもつ尺度を考えて行く。

第一に、名義尺度である。これは伝統的な尺度の一つではあるものの、筆者らがその性質を違った角度から見直し、人間の評価行動に直結する尺度として認識するまでは⁽⁴⁾、実数直線を尺度の究極の姿とする慣行が災いして、数字を符号として使った、あるものを他のものと区別するだけの機能しかない、いわば退化した、つまらない尺度としてしか見られていなかった。

本論文は名義尺度を使って行われる計測が、計測される対象に、名義あるいは何かの単語(名詞、形容詞、などを含む)が対応する一種の写像である、と解釈する。計測、測定、表現、評価などのすべてに通用する広い概念でこの写像をあらわすとすると、『・・・は・・・である』の形式となるであろう。したがって本論文では、評価を対象一述語システムとして解釈し直す。ここで述語とは広義の意味をもち、対象すなわち評価されるものを特定するのに役立つ要素の集まりである。

続くふたつの章は、この理論の要点を述べる。

名義尺度は数字を使わなくても表現できる。と、言うことは数字のもつ性質の全部をもっていない尺度であり例えば、推移性を満足するかしないかは、名義尺度では問題にならない。実際われわれの日常行う評価行動は、推移性とはそぐわないものが沢山ある。評価の対象が多く、評価の基準が単純でないとき、われわれの評価は局所的に行われ、全域にわたって推移性のある順序尺度では、到底評価にうまく適合しないというのが一般的である。それでは、推移性がなくて局所的な評価が積み重なった評価行動の全体をまとめて眺める方法はあるのだろうか？

結論の直前の章では、この問題を取り上げる。伝統的

な一対比較法が、本論文の論旨に従って、非推移的であるが、名義尺度の対象一述語システムの枠組みの中では順序尺度よりの評価として、扱われている。

最終章は結論である。

筆者らは通常のあいまい理論では扱えない循環構造、即ち一対比較の尺度論的解釈を参考文献⁽¹⁰⁾で展開した。その後、参考文献⁽⁴⁾で計測を一般的に写像として捉える見方で名義尺度の新解釈を提唱した。これは参考文献^{(7),(8)}の述語による分類論に発展して現在に至っている。これら一連の研究を通して計測と評価に関する新しい展望を得て、工学教育の場での応用を初めて論じたのが本論文である。

2. 評価写像

本論文では、評価のプロセスは、評価される対象から評価に用いられる尺度への写像とみなされる。通常の理工学的な計測においても同様な写像としての解釈が成り立つ。計測される物理変数は、ほとんどの場合数値をもつ尺度に写像される。この写像の背後には、計測を行う人間、使用される計測器、そしてこの種の写像が目的にかなう、適当である、という斯界の暗黙の了解がある、とすることができよう。

我々は、1クラスの学生のもつ学力、課題の達成度などを、計測あるいは評価しようとするとき、ペーパーテストを行って〔0, 100〕の比例尺度を使い採点をする。これは、伝統的なやりかただが、工学部の学生にとって、あるばあい不適切である。これに代わる方法が、たとえば、総合的に学生を評価するために学生に提出させるポートフォリオ^{*註}による評価とか、ビデオ記録による観察での評価など、のように時折提案されている^{(5),(6)}。

*註 もともとは『書類入れ』と訳され、記録、資料、レポートなどを入れた鞆状のケース。学生は一年または半年ぐらいかけて、教授の出したテーマについて、多くはグループで検討、議論、調査しながら各自ポートフォリオを作ってゆく。教授は途中の活動と最後に提出されたポートフォリオで評価する。結論として80点というような簡単な評価にはしない。一種のミニ卒論のようなもの。

これらの新しい評価法の骨子は学生の勉学達成度を全体的にとらえ、単に目にみえる結果だけではなく、彼等が将来のエンジニアとしての仕事の上で重要な、文献を根気よくしらべる姿勢、仲間の学生たちとの協調性、独創性、勉学中に獲得した実用的な技術、などをも含めた評価にしてゆこう、というものである。

これらの方法に共通して仮定できることは、一種のチェックリストがあって、全学生にあてはめるチェック項目、例えば『勤勉性』『協調性』『独創性』…があり、教師はそのおのおのについて〔0, 100〕の尺度を使うか、または、単に『yes』『no』でチェックする、と考えれば良いであろう。後者の簡単な方法でも項目の選び方、そ

の数などによって、さまざまな評価のできる可能性がある。この論文では、これを対象—述語の関係としてとらえる。『yes』『no』のチェックリストは対象を述語で表現する関係、例えば『Aさんは勤勉である』『B君は協力的でない』…の一群の肯定または否定の対象—述語システムと等価であるとみなすのである。これは計測の尺度にあてはめれば、名義尺度を使用していることと等価である。名義尺度は計測されている対象に数値をあてはめて、それを他の対象と区別する。われわれの評価システムでは、一群の述語に関して対象A, B, C, …のおおのは、0010…のような数値(符号としての数値)をもつ。それによって、A, B, C, …は互いに区別され評価される。区別のつかぬ対は、まったく等しいといえるであろう。[0, 100]の尺度の同点に似ているが、名義尺度には順序性がない、という点で決定的に違っている。教育の現場で学生を総合的に評価する必要のある読者は、さまざまな具体例に遭遇すると思われるが、この論文の4ページの表1を作ること、5ページの表2から5ページの図2図3を作ること、を自分の問題に当てはめ、解釈し直して使っていただければ幸いである。

もともと名義尺度はその他の尺度、実数 R を使った尺度を表現できる。しかし、逆は成立しない。図1はこの包含関係を表す。同図に破線で示す尺度の存在する可能性を後章で説明する。これは非推移性をもつ尺度で実数 R の性質ではカバーできない性質をもつ。

図1 評価写像

以上の評価システム、あるいは名義尺度による図1のような評価写像の基本的な形式化を以下に示す。

$X = \{x_1, x_2, \dots, x_n\}$ 計測される対象の集合

$Y = \{y_1, y_2, \dots, y_m\}$ チェックリストの項目の集合

$f: X \rightarrow Y$ X から Y への写像

$\nu = \{1, 0\}$ 価値(2値)の集合

$\langle X, Y, f, \nu \rangle$ 評価システム

評価の数値には0と1の二つの値しかない処に注意されたい。もしこれが暗黙の了解事項に入るならば上記の評価システムの表現は、4項組ではなく、3項組

$$\langle X, Y, f \rangle$$

として書ける。表1はこの評価システムのもうひとつの表現である。

表1 評価システム

	y_1	y_2	y_3	$\dots$	y_m
x_1	1	0	0	$\dots$	1
x_2	0	1	0		1
x_3	1	1	0		0
$\vdots$	$\vdots$			$\ddots$	$\vdots$
x_{n-1}	0	0	1		0
x_n	0	1	0	$\dots$	1

3. 評価の理論

前の章において定義された評価システムは、評価される対象をひとつの名義尺度によって分類したもの、と考えることもできる。この名義尺度はチェックリストの各項目の述語によって、評価者の心理的な基準を表現している。また例えば、表1で分類される X は対象の集合だが、 x_2 と x_n は同じ分類上のクラスに入るように見えるこの種の直感を、前の章で定義した記号と新たに導入する記号で形式化しておく。

いま集合 Y のなかの部分集合 $A = \{a_1, a_2, \dots\} \subset Y$ をとり、肯定集合(affirmative subset) A_i を次式で定義しよう。

$$A_i \triangleq \{a \in A : x_i \text{ affirms } a\}$$

集合 Y は評価項目の集合であり、 A はその部分集合だから A_i は、ある評価対象 x_i を肯定する項目のすべてを集めたもの、といえる。したがって任意に択んだ二つの対象 x_i と x_j とは共に肯定集合を持ちそれぞれ A_i , A_j と書ける。このときこの二つの対象 x_i と x_j とが、 A に関して、あるいは A のもとで、識別不能(indiscernible)であるとは $A_i = A_j$ が成立するときであり、またそのときに限る。これを式で書くと次のようになる。

$$x_i \sim x_j \Leftrightarrow A_i = A_j$$

命題 1

評価システムの評価項目 Y のなかの任意の二つの部分集合 A と B をとろう。すると次式が成立する。

$$x_i \sim^B x_j \Leftrightarrow x_i \sim x_j, x_i \sim^A x_j$$

この命題は二つの集合 A と B との和集合に関する識別不

能性を述べている。

証明

まず簡単な検算で $A_i \cup B_i = (A \cup B)_i$ であることがわかる。もし $x_i \overset{A}{\sim} x_j$ かつ $x_i \overset{B}{\sim} x_j$ が成りたてば、識別不能性の定義により $A_i = A_j$ かつ $B_i = B_j$ 。ゆえに次式が成り立つ。

$$(A \cup B)_i = A_i \cup B_i = A_j \cup B_j = (A \cup B)_j$$

これは $x_i \overset{A \cup B}{\sim} x_j$ を意味する。逆に $x_i \overset{A \cup B}{\sim} x_j$ が成立すると、 $(A \cup B)_i = (A \cup B)_j$ 。従って次式を得る。

$$A = A \cap (A \cup B)_i = A \cap (A \cup B)_j = A_j$$

これは $x_i \overset{A}{\sim} x_j$ を意味する。 B についても同様で $x_i \overset{B}{\sim} x_j$ が得られる。■

評価される対象 X が $A \subset Y$ によって分類されるとも解釈できることは前に述べた。この分類の表記を $C(A)$ としよう。同じ X の別の分類 $C(B)$ が $C(A)$ よりも精しい、とは $x_i \overset{B}{\sim} x_j \Rightarrow x_i \overset{A}{\sim} x_j$ が成り立つときで、これを次のように書く。

$$C(A) < C(B)$$

さらに、 $C(A) < C(B)$ かつ $C(B) < C(A)$ なら我々は二つの分類は同等である、と言い、次のように書く。

$$C(A) \sim C(B)$$

命題 2

一般に $A \subset B \Rightarrow C(A) < C(B)$ はかならず成り立つが、逆 $C(A) < C(B) \Rightarrow A \subset B$ は必ずしも成り立つとはいえない。

証明

前段は命題 1 からただちに言える。後段は簡単な矛盾例を作って証明できる。■

命題 3

$$C(A) < C(B) \Leftrightarrow C(A \cup B) \sim C(B)$$

この命題の証明は簡単ではないが、済んでいる。

以上命題 1 ないし 3 は著者のひとりによって得られた結果の一部である^{(7),(8)}。命題 2 および 3 についての詳しい証明は紙数の関係から省略するが、その含意は重要である。

4. 名義尺度と順序尺度の間

評価システム $\langle X, Y, f, \nu \rangle$ は多くの形式を取り得るし、また広い応用面をもっている⁽⁹⁾。前章でとりあげたのは、述語、特に yes と no による答が意味のある述語だけ、

による評価をうける対象の分類であった。もしその述語が対象の質に関する述語であるということになると、分類は『順序づけ』『ランク付』『選好』…等々の方向に動いてゆく。しかし、実数 R を用いた順序尺度にいたるまでには多くの段階があり、筆者らは名義尺度(分類)から順序尺度(線形のものさし)までの間にある興味深い尺度にもっと注意をはらうべきであろう、と考えている。我々の日常生活での判断は多くの場合、はっきりしていない、そうした判断の基準をどのように合理的に表現できるかは、評価、あるいは計測の基本になる事がらとして常に問題を提起し続けるであろう。

日本においては、まだ大部分の大学で卒業論文を書くことが要求されている。通常四年生になると、指導する教授のもとで自分自身の研究あるいは調査をやらなければならない。教授のほうは数人の学生を受け持って同じ研究室、ゼミで一年間さまざまなコンタクトをもつ。半年くらいで教授は総合的に判断して、どの学生がどの方面の仕事に向いているかを、企業から来た求人担当者に告げることができるようになるであろう。

この際、教授は自分の指導している何人かの学生について一対比較法(paired comparison)を用いることができる。前章で述べた評価システム $\langle X, Y, f, \nu \rangle$ は次の表 2 のような形をとる。

表 2 一対比較

	x_1	x_2	x_3	$\cdots$	x_{n-1}	x_n
x_1	*	1	0	$\cdots$	0	1
x_2	0	*	1		0	1
x_3	1	0	*		1	0
$\vdots$	$\vdots$			$\ddots$		$\vdots$
x_{n-1}	1	1	0		*	0
x_n	0	0	1	$\cdots$	1	*

この一対比較表から、行の 1 の数を加え合わせた優越数(dominance number)を求めることもできようが一対比較表には単なる優越数よりも多くの情報が含まれている。

図 2 と図 3 は 10 人の学生(即評価対象)の主道線図(main path diagram)とよばれる一対比較表現である。これは 3 すくみや、もっと大きなサイクルを含む、そして 10 人全部を含むサイクルが少なくとも 1 個ある例である。両図とも同じ一対比較を表す。矢印は選好(preference)すなわち表 2 でいえば 1 を表す。図中、例えば x_1 は x_{10} より好まれることを示す。直線部の矢印と同じ向き矢印は隣り合わせ以外はすべて省略し、逆向きのときだけ、半円弧に矢印をつけた枝で示す。主道線図の特徴のひとつは、それが多くの、等価な主道線図に変換できる

図2 主道線図(MPD)

図3 上図と等価なMPD

ということである。

両図とも単なる例に過ぎないが、図3の方は、3要素のグループ(x_9, x_7, x_8)と、6要素のグループ($x_1, x_2, \dots, x_{10}$)との間にある種の順序関係があることが見てとれる。これは順序尺度的な比較への一つの接近であるといえよう。この種の推移性のない尺度を『疑似区間順序(pseudo interval order scale)』と名づける⁽¹⁰⁾。

実際、求人に応じて自分の研究室にいる10人の学生の誰を推薦するか、という時、教授は頭の中にあるいろいろな条件(相手の会社の性格、職種、学生の性格、学力、得意な分野、…等)を勘案しながら、まず表2を作る。それを図2、図3…の形にして行く過程で、10人学生を比較するわけだが、この手法には、

- 1) 総合的な判断が含まれている。
 - 2) 見落としている因子が減る。
 - 3) 10人は本当に誰でも推薦できるか、あるいは多少差があるのかを自信をもって決められる。
 - 4) 条件を変えた時、どうなるかがはっきりする。
- 等々の利点がある。

この一対比較は、古くからある手法だが、主道線図のほうは著者らの発案による。

5. 結 論

以上この論文では、工学教育における最近の評価法として提案されている総合的な評価法(参考文献^{(5),(6)}など)を名義尺度の一般化の観点からとらえなおし、従来から採用されている比例、間隔、順序に比較して、名義尺度が評価法として十分意味のある尺度であり、理論的にも

応用的にも更に研究されるべき対象であることを明らかにした。また、学生の総合的な評価について、著者らの計測理論的な研究結果が適用可能であることを明らかにした。

名義尺度と順序尺度中間にある尺度等に例示されているように、疑似区間順序のような複雑な尺度の構成や解析は、手計算の及ぶ範囲を越えている。しかしコンピュータ時代に住む我々には、評価にかかわる新しい方法を模索し、実験し、確立してゆく大きな可能性が与えられている。いままでの評価の尺度についての議論は、ほとんど実数直線をもとにした議論であった、一対比較に現れる3すくみやサイクルは、実数の世界からみると、誤り、乃至不確実性の現れとして処理、あるいは無視されてきた。この論文で著者らが提示した名義尺度の再解釈とともに、工学教育は、尺度に関する応用例や野心的な実験が潜在的に豊富な分野であると信じている。

参 考 文 献

- (1) Stevens S.S.: On the theory of scales of measurement, Science 102, 677-680, 1946
- (2) 池田 央: 調査と測定, 新曜社, 1975
- (3) Zadeh L.A.: Fuzzy sets, Information and Control 8, 338-353, 1965
- (4) Watanabe H. et al.: A study of nominal scale, Proc. of IMEKO Int'l Symp. on AIMac '91, pp. 295-302, 1991
- (5) Kalonji G.: Alternative Assessment in Engineering Education: The Use of Journals in a Core Materials Science Subject, Proc. of the Engineering Foundation Conference, Santa Barbara, CA, pp.215-220, January 1992
- (6) Leifer L.: Evaluating Project-Based Learning Education: Int'l Symp. on Engineering Education Reform and Evaluation, Osaka, pp.71-78, November 1995
- (7) 渡邊啓史ほか: 述語による分類と計測・評価の基礎理論, 日本行動計量学会第23回大会発表論文集, pp. 102-105, 1995
- (8) Watanabe H.: Theory of classification of objects by predicates, Measurement: Journal of the International Measurement Confederation Vol.18, No.1, pp.59-69, 1996
- (9) Morita Y. et al.: A Theory of evaluation applied to human-environment systems, Journal of Thermal Biology 18, No.5/6, 443-447, 1993
- (10) 岡 雄三ほか: 評価の過程で現れる循環構造における分割と疑似区間順序, 計測自動制御学会論文集 17-2, pp.389-396, 1981

(平成9年2月24日受付)

森田 矢次郎

1953年 東京工業大学機械科卒業
 1956年 米国フロリダ大学大学院修士課程終了M.Sc.
 1959年 米国ペンシルバニア州立大学大学院博士課程終了Ph.D., 東京工業大学助手。助教授、教授を経て
 1992年 東京工業大学名誉教授
 現在 拓殖大学工学部機械システム工学科教授

岡 雄三

1967年 東京農工大学工学部卒業
 1969年 東京工業大学大学院修士課程終了
 1975年 東京工業大学大学院博士課程満期退学，東京工業大学制御工学科助手
 1988年 桐蔭学園横浜大学制御システム工学科助教授
 1993年 工学院大学機械工学科助教授
 現在 工学院大学機械システム工学科助教授

渡邊啓史

1982年 東京大学理学部数学科卒業，東京大学理学部情報科学科研究生
 1983年 東京工業大学工学部研究生
 現在 東京工業大学総合理工学部研究生

N 日工教ニュース

平成9年度工学教育連合講演会の開催

日本工学教育協会主催，日本土木学会担当により，「工学教育の国際化と評価認定制度」—大学と企業の教育の方向性—というテーマで，下記プログラムにより標記連合講演会を去る5月16日(金)に土木学会図書館講堂において開催し，盛況裡に終了した。

セッション 1.

主要国の専門技術者資格取得と工学強化課程の評価認定
 日本工学教育協会 専務理事 原田 耕作氏
 我が国における工学教育のアカレディテーションシステムの在り方について 工学院大学長 大橋 秀雄氏
 北海道工業大学教授 柴田 拓二氏
 基調講演 歴史的転換期の渦中にある工学
 名古屋大学教授 松尾 稔氏

セッション 2.

日本技術者の国際化対応
 大成建設(株)国際事業本部土木部長 草柳 俊二氏

セッション 3.

工学系大学における留学生の受入れの在り方について
 埼玉大学大学院政策科学研究科長 西野 文雄氏

技術的成長に果たす工学の役割

慶応義塾大学大学院政策・メディア研究科教授
 茅 陽一氏

シンポジウム「工学教育の現状と将来」の開催

日本工学教育協会は，東海工学教育協会との連合により，標記シンポジウム(副題—工学教育のカリキュラムとその周辺の問題—)を，去る5月20日(火)午後，名古屋大学シンポジオンホールにおいて下記プログラムにより開催，東海地域の大学及び企業関係者多数の参加を得て，盛況裡に終了することができた。

1. 工学教育のカリキュラムに関する大学での検討状況
 名古屋大学理工科学総合研究センター長 架谷 昌信氏
 2. 高等学校での新学習指導要領に基づく理科教育
 愛知県理科学科研究会，高等学校部会・物理化学部会長 愛知県立稲沢高等学校長 竹吉 廣一氏
 3. 工学教育プログラムのアカレディテーションと専門技術者資格取得に関する内外の動向
 日本工学教育協会 専務理事 原田 耕作氏
 4. 我が国の工学教育におけるアカレディテーションシステムの在り方について
 工学院大学 学長・日本学術会議会員 大橋 秀雄氏
- 以上の講演後，質疑応答と懇親会が開催された。