

## 教育評価 — 尺度、絶対と相対 —

吉岡 昌紀

教職課程のカリキュラムでは、教育評価を扱うことになっており、筆者が担当する「教育方法論」で5回ないし6回分の授業をこれに充てている。本稿は、その中で、測定に用いる尺度に関する部分をまとめたものである。尺度が備えている基準点と目盛りの性質をもとに、比例尺度、間隔尺度、順序尺度についてまとめた。その後、絶対尺度と相対尺度について、絶対と相対の意味に着目しつつ、両者の違いと関連性について述べた。これは、絶対評価と相対評価を理解するための前提としても重要である。

### 1. 尺度

測定と評価の違いについては「教育評価 — 測定と評価 —」<sup>1</sup>で述べた。それを踏まえて、今回は、評価をするための基礎となる測定について考える。

測定をする際には、ふつう、何らかの道具を使う。たとえば、長さを測るためには定規を使い、重さを測るためにははかりを使い、時間を測るためにはストップウォッチを使う。道具はもっと細かく分化している。長さを測るためには定規を使うと書いたが、それ以外にも、短い長さを測るためのマイクロメーター、長い距離を測る巻尺、さらに長い距離を測るためのレーザー距離計など、さまざまな道具がある。

問1 このように、測ろうとする性質も、そのために使う道具もさまざまであるが、それら測定のための道具が共通に用意し、用いているものがある。それは何か。

問いの対象が広すぎ、考えにくいかもしれないが、その幅広い測定の道具は、何らかの目盛りを備えている。また、目盛りの中には基準となる特別の点がある。このような目盛りと基準点を尺度という。

先ほど述べた、長さ、重さ、時間、温度などを測定する道具に目盛りがついていることはわかりやすいだろう。典型的なのは身近にある物差しや温度計である。近年は、数値がディスプレイに表示される計器も多く、計器自体に目盛りが刻まれていない場合も多いが、その場合も、何らかの目盛りを当てはめ、その目盛り上の数値を読み取っているわけである。

目盛りに比べると意識しにくいかもしれないが、このような目盛りには基準点が存在している。小学校2年生の算数で定規の使い方を習うが、そのとき、正しく長さを測るためには、まず、定規のゼロを測定対象の一方の端とピタリと合わせ、その上で、測定対象のもう一方の端のところの目

盛りを読み取るように習うはずである。長さの測定ではゼロが基準点なので、測定に際してはまずそこを合わせなければならない。重さも同様である。まず、体重計の目盛りを0に合わせる必要がある。ストップウォッチならば、測定前にリセットして0秒に針を戻した上で、測定を開始する。

問2 次の測定／評価では、どのような目盛りを用いているか。また、基準となる点はその目盛りのどこにあるか。

- 1) 身長測定
- 2) 体温測定
- 3) 100 点満点の期末テスト
- 4) 授業の終わりに行う、8点以上取れたら合格という、10 点満点の小テスト
- 5) 上位 30 人まで特進クラスに入ることができる、塾のテスト
- 6) 全国規模の大学入試模試の偏差値

### 1) 身長測定

身長測定には身長計を用いるが、自明のことだがこれには長さの目盛りがついている。通常はミリメートル単位で目盛りが刻まれているだろう。基準となる点はゼロである。

### 2) 体温測定

体温測定には体温計を使うが、これも自明だが温度の目盛りがついている。通常は  $0.1^{\circ}\text{C}$  単位である。基準となる点はどこだろうか。たとえば体温が  $36.5^{\circ}\text{C}$  であった場合、摂氏で測定しているわけだから、基準は水の融点である  $0^{\circ}\text{C}$  と考えることもできる。しかし、通常、体温測定では平熱か発熱しているかが問題となる。一般的にこの基準は  $37^{\circ}\text{C}$  に置かれている（その証拠に、昔からある水銀体温計では、37 という数字だけ目立つように赤で書かれている）。このような用途を考えると、体温測定で基準となるのは一般に  $37^{\circ}\text{C}$  と考えることができる。もっとも、一人一人平熱が異なるから、その人の平熱を基準として考えると考えることもできる。

### 3) 100 点満点の期末テスト

テストもまた、測定の道具の一つであると考えられる。ある教科のある分野について、どのような知識、能力、技能等をどのくらい持っているのかを測っているのである。尺度である以上、目盛りがある。この場合、0から 100 までの数値（多くの場合、整数値）が目盛りである。基準となる点はどこにあるだろうか。これにはいくつかの考え方がありうる。教師は問題を作成するときに、「このくらいの点をとってほしい」「この点をとらないと不合格」という点を、0から 100 までのどこかに想定しているであろう。そのような、教師が設定した点を基準点と考えることができる。試験作成時ではなく、試験をしたあとで、その結果を見て、たとえば平均 60 点だった場合にそれを基準と考えることもありうる。この点についてはのちに再び考える。

### 4) 授業の終わりに行う、8 点以上取れたら合格という、10 点満点の小テスト

目盛りは0から 10 の数値（整数）。基準点は合格のボーダーラインとして明示された8点と考えられる。

### 5) 上位 30 人まで特進クラスに入ることができる、塾のテスト

テストだから得点がつくとはいえ、この場合は、得点自体よりも得点によって定まる順位に目が向くことだろう。順位は、1 から順に整数を割り振られた順位が目盛りとなる（順序尺度という）。基準となる点は 30 位と考えられる。

### 6) 全国規模の大学入試模試の偏差値

偏差値については、のちに改めて述べる。偏差値は標準偏差と平均から算出する。偏差値の目盛りは標準偏差をもとに、偏差値の基準点は平均をもとに算出する。

以上、いくつかの例を見てきたが、そこで確認できたように、測定では何らかの道具を用いるが、道具には尺度があり、尺度には目盛りと特別な基準点が備わっている。

## 2. 比例尺度、間隔尺度、順序尺度

尺度は、目盛りと基準点の性質から、いくつかに分けられる。

上の例題でも考えたが、長さ、重さ、時間の尺度はわかりやすい。いずれも基準となる点はゼロである。長さや重さや時間が存在しない、ゼロという状態は直感的にも理解できる。測定の際には、この状態を確認し、そこを基準として測定を始めることになる。また、これらの目盛りは等間隔であるという特徴を持つ。物差しが目盛りは等しい感覚に打ってある。体重計の目盛りも時計の目盛りも同様である。さらに、目盛りはとびとびに打ってあるが、実際の測定値は目盛りの間のどの数値も取りうる。このように、絶対のゼロを基準とし、目盛りが等間隔についており、測定値がその間のあらゆる値を取りうるような尺度を、比例尺度という。比例尺度は便利である。測定結果に四則演算を施して処理を加えたり、測定結果同士を比較したりすることを、正確にわかりやすくできる。また、測定結果を組み合わせて別の数値を得る（たとえば、長さ(の3乗)と重さから比重を計算する)こともたやすい。

温度も等間隔の目盛りを持つ尺度で測定するが、ゼロの意味が異なっている。長さなどのゼロは長さなどが無い状態だが、摂氏(℃)温度計のゼロは温度が無い状態ではない（マイナスの気温が存在することからも、このことはわかるだろう）。温度がゼロの状態は、絶対温度(K)の0(-273.15.℃)である。℃は、絶対温度のゼロから始まる目盛りの途中にある、水の融点を0と表示し直した尺度である。このような、目盛りが等間隔で測定値がさまざまな値をとり、基準点が絶対のゼロではない尺度を、間隔尺度という。℃や標高などがその例である。体温の測定の場合は、間隔尺度の℃を用い、さらに 37℃という、℃で表してもゼロではないところに基準点を設けていることになる。

また、順序尺度がある。これは、1位、2位、3位というように、何らかの性質の測定結果を、その性質の多寡の順に並べていった順番を示す。多い少ないに注目しているが、差の大きさは反映しない。たとえば、スポーツ大会の種目の一つに走り幅跳びがあったとしよう。1位のA君は5メートル 80 センチ、2位のB君はわずかに及ばず5メートル 78 センチ、3位のC君は少し離されて5メートル 50 センチだったとする。長さでは1位と2位の差は2センチ、2位と3位の差は 28 センチと異なるが、順位では、1位と2位、2位と3位はいずれも1つ違いとなる。

### 3. 絶対尺度、相対尺度

尺度には別の分け方がある。

上で校内走り幅跳び大会の例をあげた際に、1位は5メートル 80 センチ、2位は5メートル 78 センチというように、順序尺度と比例尺度の二つを用いて結果を示し、両尺度には目盛りの間隔や基準点に違いがあると述べた。しかし、両尺度には、それ以外にさらに大きな違いがある。

この違いは、実際の競技の様子を考えるとわかりやすい。

100 メートル走などとは異なり、走り幅跳びの場合、選手は一斉に競技をするのではなく、1人ずつ順番に跳躍する。そして選手が跳躍すると跳んだ距離が測定され、〇メートル〇〇センチと測定結果が確定する。この作業が、選手が跳ぶたびに繰り返される。

では、順位はいつ確定するだろうか。

もちろん、選手が跳んだ時点で、それまでの選手の中での順位は確定する。しかしそれはあくまでもそれまでの選手の中での暫定的な順位であって、最終的な順位ではない。これは距離の測定と是对照的である。距離は測定するたびに、それが最終的な測定結果となった。ところが、順位はその時点で得られるのは暫定順位でしかなく、最終的な順位は、全選手が跳び終わってはじめて確定する。観客は、誰が首位になるのか最後までわからないから、はらはらしながら見続けることになる（ちなみに、スキージャンプ競技などでは、このはらはら感を増幅するような工夫がなされている。その時点で首位の選手が、首位の選手用に用意されたシートに座る。あとから跳んだ選手がさらに上の記録を出すたび、そのシートに座る選手が入れ替わる）。距離などのように、測定の都度、最終的な測定結果が得られる尺度と、順位のように、すべての測定データが揃ってからのはじめて、最終的な測定結果が得られる尺度がある。

測定の都度、確定した結果が得られる尺度は、なぜそのような結果が得られるかといえば、測定に用いる尺度（目盛りと基準点）が測定よりも前に存在しているからである。すでに存在している尺度をそのまま当てはめれば、結果が得られる。当たり前すぎることもかもしれないが、長さ、重さ、時間、温度など、ふつう何らかの道具を用いて測定する尺度はこのような尺度である。

一方、順位は、すべての結果が揃わないと測定結果（順位）が定まらない。これは、順位が、すべての結果の中でそれぞれの結果の位置を示すという、測定結果相互の関係にもとづくものだからである。まだ測定していないものがあるうちは、ある結果の位置が定まらないのは当然である。

前者のように、あらかじめ存在している尺度を当てはめれば測定結果が確定する尺度を絶対尺度という。一方、すべての測定結果が集まったのち、その相互関係の中で個々の結果の位置が明らかとなり測定結果が確定する尺度を、相対尺度という。

問3 次の測定／評価では、比例尺度、間隔尺度、順序尺度のどれが使われているか。また、絶対尺度、相対尺度のいずれが使われているか。

- 1) 身長測定
- 2) 体温測定
- 3) 上位 30 人まで特進クラスに入ることができる、塾のテスト
- 4) ある単元の終わりに行う、50 点を取れば合格という、100 点満点のテスト

- 5) ある単元の終わりに行う、平均点を取れば合格という、100 点満点のテスト
- 6) 全国規模の大学入試模試の偏差値

### 1) 身長測定

比例尺度。絶対尺度。これは特に説明を要しないだろう。

### 2) 体温測定

間隔尺度。絶対尺度。これも特に説明を要しないだろう。

### 3) 上位 30 人まで特進クラスに入ることができる、塾のテスト

上位 30 人という順位をもとにこの判定を行うのだから、ここで用いられているのは順序尺度である。順位は、全員の得点が明らかになるまで確定しない、すべての結果にもとづいて定まるものだから相対尺度である。順序尺度は相対尺度の代表例である。

### 4) ある単元の終わりに行う、50 点を取れば合格という、100 点満点のテスト

100 点満点のテストの場合、測定結果は 0 から 100 までの数値（ふつうは整数値）として示される。この 0 から 100 までの数字がふつうある尺度は、比例尺度なのだろうか、間隔尺度なのだろうか。

すぐに気づくのは、どのような学力テストであれ、0 点を取ったからといって、そのテストが測定しようとする知識や技能などがまったくないわけではない、という点である。たとえば、このテストが数学の四則演算を複雑に組み合わせた計算問題からできており、ある生徒が 0 点を取ってしまったとしよう。この生徒は、このテストが測定できる範囲では、知識や技能がゼロという結果となったわけだが、だからと言って、この生徒の計算能力がゼロとは言いきれない。何をゼロとしたらいいかも、高い方の数値をどう表したらいいかもわからない知識や技能のごく一部について、0 から 100 までの数値を当てはめて測定するのが、学校のテストと考えられる。また、その知識や技能は連続的な数値を取りうるものと思われるが、実際の得点としては整数値としてのみ表示されることが多い。

また、このテストには、見たところ 0 から 100 までの数値が等間隔に並んでいるが、測定される知識や技能は等間隔なのだろうか。たとえば、あるテストについて、作成した教師が、「きちんと勉強すれば 90 点を取ることができますが、100 点取るのはかなり難しいテストです。」などと言うことがあるのではないだろうか。このような場合、80 点から 90 点までの 10 点と比べて、90 点から 100 点までの 10 点の間隔は、同じ 10 点ではあるけれども、知識や技能としてはかなり幅が広い、と想定されていることになるだろう。

このように、学校で行われるテストは、0 点だからといってそれが完全なゼロを意味するわけではない。また、目盛りも、数値は等間隔につけられていても、測定されている知識、技能、資質などが等間隔である保証はない。しかしながら、実際には、学校でのテストは間隔尺度として（すなわち、数値は等間隔のものとして）扱われることが多い。

ところでこのテストは、絶対尺度、相対尺度という観点からはどのように考えたらよいだろうか。50 点以上が合格だというのであれば、一人一人のテストを採点するたびに、その生徒が合格か不合

格かを明確に決定することができる。このテストは絶対尺度として用いられていることになる。

#### 5) ある単元の終わりに行う、平均点を取れば合格という、100 点満点のテスト

上の 3.と同じく、この尺度は、通常は間隔尺度のように扱う。

では、絶対尺度、相対尺度という観点からは、この場合どう考えられるだろうか。

3.との違いは可否の判定に平均点が用いられているところである。平均はすべてのデータを合計して、それをデータの数で割って得られる。ということは、すべてのデータが得られてはじめて、平均点が確定することになる。平均は相対尺度の基準点としてよく用いられる。

相対尺度によく用いられる平均が基準点だとして、目盛りはどうだろうか。目盛りは 3.と同じく 0 から 100 の数値であり、得点は、その目盛りの中の数値として、採点の都度、確定していく。すなわち、得点は、絶対尺度と考えられる性質を持っている。しかしながら、得点は測定すればすぐに確定したとしても、3.とは異なって、可否は平均が算出されるまで確定しない。得点は確定しても、基準点としての平均が算出されるまで、その意味は一種の保留状態にあり、平均が算出されてはじめて、平均よりも○点上だった（から合格）、○点下だった（から不合格）といった具合に意味が確定する。このようにみえてくると、平均を基準点とすることによって、絶対尺度の目盛り上にあった得点の意味が、平均を基準にした相対尺度のものへと変換され解釈されていると考えられる。したがってこの尺度は、相対尺度として用いられていると考えられる。

3.と 4.の事例からわかるように、同一の測定値であっても、基準の置き方や測定値の捉え方、測定の目的などによって、絶対尺度上のものとも相対尺度上のものとも考えられることはよくあることである。のちに改めて述べるが、絶対評価、相対評価を考える際にも、両尺度のこのような関係を注意深く考慮する必要がある。

#### 6) 全国規模の大学入試模試の偏差値

間隔尺度で相対尺度。問2の 6.の説明で述べたように、偏差値は平均を基準点として使っている。このことから推測できるように偏差値は相対尺度である。偏差値で目盛りを作るときに用いる標準偏差もまた、すべての測定結果をもとに算出している。すなわち、偏差値は目盛りと基準点の両方をすべての測定結果をもとに算出している相対尺度である。なお、偏差値の目盛りは等間隔と考えられるので間隔尺度になる。

## 4. 絶対評価、相対評価

絶対尺度、相対尺度について述べてきたが、その中ですでに、絶対評価、相対評価にも言及していた。その際には測定と評価を明確に分けず述べたので、ここで改めて整理しておくと、絶対尺度を用いて評価を行うのが絶対評価であり、相対尺度を用いて評価を行うのが相対評価ということになる。

絶対評価と相対評価は対極にあるように受け止められているかもしれないが、そうではない。同一の測定結果を絶対尺度上で捉えたり相対尺度上で捉えたりすることができたように、絶対評価、相対評価も同一の測定結果に対して行いうる。もちろん両者には明確な違いがあるが、両者はかな

り近い関係にもある。

たとえば次のような例を考えてみよう。ある教師が、70 点を取れば合格と評価するつもりでテストを作成したとしよう。教師は、生徒の8割は 70 点を取ることができるようにつめたつもりであった。テスト後、教師は、A君は 72 点だから合格、B君は 60 点だから不合格という具合に採点を始めたのだが、採点を始めてしばらくして、どうやら 70 点を下回る生徒が予想よりも大幅に多く、このままでは 70 点を取れる生徒は5割ほどしかいそうにないと気づいた。そこで教師は、採点の途中で、対応策を考え始めた。

さて、教師はどのような対応策を考えることができるだろうか。いくつか例を示す。

- 1) 全部の答案を採点し、上位から8割の生徒を合格とする。
- 2) 設問の配点を変えたり部分点を多めに与えたりするなど、採点基準を甘くし、8割が 70 点となるようにする。
- 3) 70 点を取れないのに合格にするわけにはいかないので、70 点を取れない生徒は不合格にする（つまり、対応策は取らず当初の評価基準のままとする）。

採点を始めたとき、この教師は絶対評価を行っていた。このことは説明を要しないだろう。では、採点途中で考えた 1)から 3)の案は、絶対評価、相対評価のいずれであろうか。1)は、は相対尺度の代表である順序尺度を用いた相対評価である。3)は採点を始めた時と同じ絶対評価である。これらも説明は要しないだろう。ちなみにいずれも、可否の評価基準は変えたが、採点の仕方（各問に関する評価基準）は変えていない。

2)の場合は、1)や 3)とは異なって採点の仕方（各問に関する評価基準）を変えているが、70 点を取れば合格と評価しているところは採点開始時と同じであり、採点の都度、可否を判断できるのでから絶対評価と言えそうである。しかし、そうやってしまってもよいだろうか。8割の生徒の得点が 70 点以上になるように採点の仕方を変えた、ということはつまり上位8割を合格させようとしたのであり、順位で上から8割を合格させることにした 1) の場合と同様の意図を持ち同様の結果を得たことになりはしないだろうか。2)は、一見絶対評価を行なっているように見えるが、そのもとには相対評価の意識が存在していたと考えられる。

しかし、こういうことを言い始めると、そもそもテスト作成時にさかのぼって考え直す必要が出てくる。この教師は生徒の8割が 70 点を取れるようにテストを作成したのだったが、このとき、これまでの経験などに照らして、このくらいの難易度であればこのような結果になると考えたに違いない。すなわち、頭の中にある生徒集団に対して、上位8割が合格するというシミュレーションをしつつ問題を作成したことになる。これはつまり、頭の中で相対評価をしていたことになる。絶対評価をするためのテストを、頭の中に存在する集団に対して相対評価しつつ作成していたことになる。

このように、絶対評価の場合も、何らかの標準的な集団を想定し、その集団の中での平均や順位を想定して（すなわち相対尺度を用いて）評価を行うことが多い。絶対尺度と相対尺度、絶対評価と相対評価はこのように相互に関わりあっている。

- 
- <sup>1</sup> 吉岡昌紀(2017)『教育評価―測定と評価―』清泉女子大学教職課程紀要第1号、45-51