

第83回：教実研しポート 「統計を通して言語活動を鍛える」

平成 24 年 12 月 1 日
野幌高校教諭 佐川 大樹

1 はじめに

今度の学習指導要領ではいたるところに「言語活動の充実」というフレーズが入ってきています。これは思考力・判断力・思考力を問う読解力や、記述式問題、知識・技能を活用する問題に課題があるという PISA などの調査結果をふまえたもので、要するに、自分の言葉で自分の思いを誤解なく相手に伝えていく能力を、さまざまな教科を通して鍛えていきたいと思いますということだ、と自分では認識しています。

本校では、生徒の実態を考えるとどうしても計算力の向上に目が行き、表現力・記述力などについては考えもしなかったというのが本音です。社会に出ればルーチンワークが確実にできるということが当たり前のように求められる中、それすらも危ういと思われる生徒も多数いるのが現状です。せめて計算力を鍛えることで、決められたルールで確実に作業するという訓練をさせていくことが、本校の数学教育では優先すべきことがらでした。

今回、自分のクラス（3 年生の理系クラス）で「言語活動」を意識した、もっとざっくばらんに言えば、相手に説明するというのを大切に授業を展開してきました。残念ながらその評価はテストでしか測ることができませんでしたが、相手に伝えることの難しさを改めて認識しました。その様子をレポートにまとめてご報告いたします。

2 対象クラスと指導内容

対象となるクラスは前述のように、私が担任をしている理系のクラスです。本校ではフィールド制を実施しているので、正確にはナチュラルサイエンスのコースを選んだ生徒たちの集まりということになります。これまではクラス編成に関しては自然学級（生徒がどのフィールドを選んだかに関係なくクラスを作っている）にしていますが、この学年はフィールドごとにクラス編成をしています。したがって、ナチュラルを選んだ生徒を集めてできたのがこのクラスです。理系（ナチュラル）だけのクラスにしたため、結構時間割などで自由に動けるといったメリットはありますが、学年全体で見ると生徒指導が難しい生徒が同じにクラスに集中するなどのデメリットも見られます。

さて、ナチュラルフィールドの教育課程ですが、なぜか数学 B が必修ではなく（2 年の選択科目）、数学 C が 3 年生で必修となっています。この教育課程を編成した当時の先生に尋ねると、それは担当する先生が自由にやれるようにそのようにしたとのことで、それならばお言葉に甘えて好き勝手にさせていただきますということで、以下のような授業内容で進めていきました。

前期中間範囲 数列（数 B）
後期中間範囲 確率分布（数 C）

前期期末範囲 統計とコンピュータ（数 B）
学年末範囲 統計処理（数 C）

すべての生徒が、将来自分が勤める職場でデータを取って分析する、または分析したデータを読んで今後の判断材料にする、ということが求められます。だから、統計の最低限の知識は身に付けてもらいたいと考えました。もう3年生の後半になると、進路が決まり「今さら数学なんて……」という態度を取る生徒も出てきます。しかし、受験で利用する、しないに関わらず、統計を知らなければ損をするんだということを、口を酸っぱくして言い続けてきたつもりです。

3 実際の問題と生徒の解答

夏休み明けの実力テストと前期の期末考査では、数Bの「統計とコンピュータ」が範囲でした。この単元ではさまざまな用語が登場します。ちなみに本校で使用している教科書だと、この単元の最初の5ページだけで13個の太字の用語（階級、相対度数、度数分布、ヒストグラムなど）が出てきます。これらの中でも特に重要なものについては考査で説明させてみることにしました。次の問題は、夏休み明けの実力テストで出題した問題です。

(問題. 1)

次の問いに答えよ。

- (1) n 個の数 x_1, x_2, \dots, x_n の平均を求める式を書け。
- (2) メジアン（中央値）とはどのような値のことをいうか。定義を述べよ。
- (3) モード（最頻値）とはどのような値のことをいうか。定義を述べよ。

模範解答

- (2) 「変数（データ）を大きい順番（小さい順番）に並べかえたとき、ちょうど中央にくる値のこと」
- (3) 「度数が最も大きい階級の階級値のこと」

これに対する生徒の解答は以下の通りです。「○」は正解、「△」は減点（部分点を与えた）、「×」は不正解を表しています。

(問題. 1) (2) の生徒の解答

- △：数を順番にならべた時の中央の値。 △：データを順番に並べた時の中央にある値。
 一応○：はずれ値がある場合に使用され、大きい数から並べ中間の数を示した数。
 ×：ある数の中のちょうど真ん中の値。 △：データの値を順番にならべたまん中の数。
 ×：何個かある数の中央の値。 ×：中央の値。 ×：まん中の数を値にする。 ×：まんなかの数。
 ×：全体の数の中心の数。 ○：大きい順にならびかえたときちょうど中央の値。 ×：中央にある値。
 ○：データを大きい順に並べた時の中央の値。 ×：真ん中の数を2つたして割る。
 ×：偶数個あるうちの真ん中にある2つの数をたして2で割った値。 △：順番にならべた数の真ん中の数。
 ×：個数のちょうどまん中にあたる値。 ×：はずれ値がなく、平均値より、より近い値。
 ×：度数の合計の中心の値。 ○ 数字を大きい順に並べかえたときのちょうどまんなかの値のこと。
 ○：データを大きい順に並びかえたときの中央の値。 ×：ちょうど真ん中の値の真ん中。
 ×：データのまんなかにある値。 ○：大きさの順に並べて真ん中の数のこと。

中央値については、データの個数の偶奇によって場合分けしなければなりません、そこまでは問いませんでした。「何をどのような順番に並べかえたときどういう値なのか」ということがきちんと書かれているかどうか見ました。採点した当時は「何を」の部分がなくても正解にしていました（おまけしておきました）。「中央値」だからとりあえず「真ん中」とか「中央」って書いておけばという心理だったのかもしれませんが、ただ中央としか書いていないと、データの並び方によって中央値が変わってしまいます。定義というのは誰が読んでも誤解のないようにしなければならないので、あいまいな表現は相手に理解されないといういい教訓になりました。

(問題. 1) (3)の生徒の解答

△：一定の範囲で区切ったときに、最も多くのデータが解答した範囲の中央値。

×：最も人数が多いときに使用される値。 △：データの値の度数が一番多い階級値。

△：度数が一番高い階級の値。 △：もっとも度数の大きい値。 ×：一番値が高い数。

△：度数が最も大きいものごと。 ×：いちばんはしにある数。 ○：度数の最も大きい階級の階級値。

△：度数が一番大きいところの値。 △：度数の最も多い値。 ×：最大の数。 ×：一番大きな値。

×：最も多い数。 ○：最も度数が多いところの階級の階級値。 ○：度数が最も多い階級の階級値のこと。

○：度数が一番高い階級の階級値。 ×：階級に分けた時の1番多い値の真ん中。

×：データの最も大きい値。 ×：度数を階級で並べたときの階級値。

「度数」「階級」「階級値」などの用語が入ってくるので、用語を知らない生徒は苦戦していたようです。一生懸命説明しようとしている生徒もいますが、よくよく読んでみると最頻値とはまったく関係のない値の説明をしているものもありました。また、度数が最も多いところまではわかっている、その「階級」なのか「階級値」なのかが触れられていないと、読み手はどっちにも解釈してしまいます。

次の(問題. 2)も夏休み明けの実力テストで出題した問題です。このときは(3)の解答としては、例えば「149という値が他に比べて極端に大きいので、平均を計算すると149の影響を受けて高めに出来てしまい、このグループの特徴を表す数値としてはふさわしくないから。」なんていうのを模範解答としていましたが、実際には「149がはずれ値であるから」という解答が非常に多かったです(17人)。別に間違いでもなく、こちらもしっかりと指示していなかったのも、これでも正解にしました。しかし、はずれ値の本当の意味がどれだけわかっているのだろうかという疑問から、9月の前期期末考査では(問題. 3)のようにはずれ値の説明をさせる問題にしました。

(問題. 2)

8個のデータ 149, 51, 47, 42, 40, 36, 35, 32 について、次の問いに答えよ。

(1) 平均を求めよ。

(2) メジアンを求めよ。

(3) この8個の特徴を表す代表値としてふさわしいのは、この場合はメジアンであると言われている。その理由をわかりやすく説明せよ。

(問題. 3)

次の数値はある会社の社員 10 人の年棒である (単位: 万円). 次の問いに答えよ.

270, 330, 280, 390, 320, 310, 200, 360, 350, 2600

- (1) 平均値を求めよ.
- (2) 中央値を求めよ.
- (3) このデータの場合, 平均値よりも中央値や最頻値の方が代表値としてふさわしいとされている. その理由は「2600 万がはずれ値だから」であるが, これはどういう意味か. わかりやすく説明せよ.

模範解答

2600 万円という値が他に比べて極端に大きい値なので, 平均を計算するとその影響で高い値が出てしまい, 10 人のうち 9 人が平均以下となってしまう. したがってこの場合, 平均は代表値としてふさわしくないという意味である.

(問題. 3) (3)の生徒の解答

△: 2600 万を入れて平均を出すと 541 万で, 2600 万の次に高い値が 390 万ですごい差が出てしまうため代表値とは言えなくなり, 中央値は 325 万で 2600 万をのぞく 9 人に近い数字のため代表値となる.

△: はずれ値が平均に影響して, データが平均の値を 2600 万しか上回れず, 代表値としてふさわしくないから.

○: 2600 (万) は他の数とくらべて極端に大きいので, 平均を求めた場合 2 番目に大きい 390 (万) よりも上回るため代表値としてはふさわしくない.

○: 2600 万が極端に大きい数字のため, 平均を計算したとき, 2 番目に大きい数より大きくなるため, 代表値としてふさわしくない.

○: 平均の計算をすると, はずれ値の 2600 (万) が全体の平均を上げてしまい, 2 番目に大きい 390 (万) を超えてしまい, 代表値にふさわしい値にならないという意味.

△: 2600 万は他の数よりすごく大きいので数字が上がってしまうので, 中央値や最頻値などの値が代表値としてふさわしい.

○: 2600 万が他のと比べて極端に大きい値なので, 平均を計算すると, その影響で 2 番目に大きい 390 万よりも大きくなるので, 代表値としてふさわしくないという意味.

○: 2600 万という数値がとても大きく, 平均を出した時に 2 番目に大きい 390 万より平均が大きくなってしまふから.

△: 平均をしたとき 2600 万によって平均が 2 番目の 390 万より大きい値になって平均にふさわしくないから. それに比べて中央値の方がふさわしいから.

△: 2600 (万) が他に比べて極端に大きい値なので, 計算してもその影響で 2 番目に大きい 390 より大きいので代表値にはふさわしくない.

×: 2600 を入れて平均を出すと, 2 番の大きい値との差が大きくなってしまふから, 代表値としてふさわしくない.

○: 2600 万が大きいので, 平均を求めたとき, 次に大きい 390 万よりも大きくなってしまふからという意味.

○: はずれ値である 2600 万の影響で, 平均が 2600 万以外の全ての値を上回っているから.

- △：一番大きい2600（万）は、この中でずば抜けて大きい値なので、平均を計算すると、2番目に大きい390（万）よりも大きくなってしまい、このグループにはふさわしくないから。
- ：2600万円という数値が極端に大きく、平均を上げてしまい、2番目に大きい390万円という数値よりも平均が大きくなってしまうという意味。
- ：2600（万）は他の値に比べてすごく大きい値なので、平均を計算すると2番目に大きい値より平均の方が大きいから代表値としてふさわしくないということ。
- △：このデータをそのまま平均値で出すと出た値が2番目に大きくなるので、代表値としてふさわしくない、中央値や最頻値の方がはずれ値の影響を受けにくいので代表値としてふさわしい。
- △：2600（万）という数は極端に大きい数なので、平均を計算するときの影響で、次に大きい390（万）よりも大きくなってしまうので、代表値としてふさわしくないという意味。
- △：2600万は極端に大きいので、その影響で2番目に大きい390万よりも大きくなるため、代表値としてふさわしくないから。
- ：2600万が他の値に比べ極端に大きく、平均を計算したとき、2番目に大きい390万よりも大きくなり、代表値としてふさわしくないという意味。
- △：2600万が2番目に多い390万と数値が極端にかけはなれていて、平均値の代表値としてふさわしくないため。
- ：2600万が他よりも極端に大きくて、平均を出してもそれに影響されて、2番目に大きい390万よりも大きくなり、代表値としてふさわしくないという意味。
- ×：2600万ははずれ値のため、2番目に大きい数よりも大きくなってしまいう影響があるため、はずれ値の2600万はふさわしくない。
- △：2600（万）は大きい値なので平均がどうしても大きくなり、二番目の値よりも平均が大きいので2600（万）はふさわしくない。
- ：2600（万）が他の値よりも極端に大きく、平均で求めるとその値に影響され2番目に大きい390よりも大きくなってしまいうので、代表値としてふさわしくないという意味である。
- ：2600（万）という値がずばぬけて大きいので、平均を計算した時に2番目に大きい390（万）よりも大きくなってしまい代表値としてふさわしくないという意味。
- ：2600（万）は他に比べて極端に大きい値なので、平均を計算するとその影響で、2番目に大きい390（万）よりも大きくなり、代表値としてふさわしくないということ。
- △：2600（万）が他に比べて1番大きい値であり平均を求めるとき、その影響で2番目に大きい390よりも大きくなってしまいうから、代表値としてふさわしくないということ。

大事なことは、2600万という値が極端に大きいこと、平均を計算するとグループの実態を表す数字にはほど遠いこと、の2点です。後者の方は比較的指摘していましたが、前者の方は2600万がただ大きいとしか書いていないものは減点することにしました（平均よりも大きいデータが存在するのは当たり前のことだからです）。また、言わんとしていることはわかるけど、よく読めば何を書いてあるかよくわからないというものも減点することにしました。

最後の問題は、分散を使ってあなたならどちらを選ぶかという問題です。前期の期末考査で出題しました。一応、授業で扱って生徒達に考えさせましたが、下の生徒を見てもわかるように、「今の子どもたちって、やんちゃに見えて意外に安定志向なのね……」というのがわかりました。

(問題. 4)

あなたは陸上部の監督です。今、ある大会に向けて砲丸投げの選手を1人選ばなければなりません。候補となるのは次の2人の生徒A, Bです。次の値は、この2人の最近の5回の試技の結果です。2人とも平均は10mです。

A: 8, 11, 9, 9, 13

B: 6, 13, 15, 7, 9 (単位は m)

以下の問いに答えなさい。

- (1) A, Bの分散を、それぞれ単位をつけて計算しなさい。(各3点)
- (2) あなたならA, Bのどちらを選手に選びますか。(1)の結果を判断材料にして、あなたの考えを述べよ(この問題はA, Bのどちらを選んでも構いません。理由が明確に述べられているかどうかについて採点します)。(4点)

模範解答

(Aを選んだ場合)

Aの方が、分散が小さい。つまり、5回の結果が安定しているので、大会でも平均に近い記録は残せそうだから。

(Bを選んだ場合)

Bの方が、分散が大きい。つまり、5回の結果にばらつきが見られるので、大会では失敗するかもしれないが、好成績を収める可能性も期待できるから。

生徒の解答

- × A. 私はジャンプラーではないので、バランスの取れた結果を残しているA選手を選びます。
- △ B. 分散の値が大きいのであまり飛ばせないこともあるが、かなり飛ばせることもあり、上位をねらっているから。
- △ A. 理由は、Bよりも安定して平均に近い距離を飛ばせるから。
- △ A. Aの方が5回投げたときの全ての距離がだいたい安定しているから。
- △ A. Aの選手は分散の散らばりが小さく、大会でも安定した記録を出せるから。
- △ A. Aは分散しても数が小さく、記録にばらつきがないので、Aを選びます。
- A. Aの方が分散が小さいので、各回のばらつきが小さい。つまり、安定した記録を残せるのでAを選びました。
- △ A. 私はAを選びます。理由は、結果の上下が激しくないからです。
- △ A. AはBに比べて平均的に高く、散布の結果を見てもAの方が安定しているので、Aの方がいいと思う。
- B. Aは分散が小さく安定している。Bは分散が大きく不安定である。自分が監督ならばBを候補とする。なぜならば、Bの方が不安定ではあるが、分散が大きいので、よい結果を残すことができるかもしれないので。

B を選ぶ。

- × B. 自分は分散して 12 なら B の方がよいと思う。なぜならば、12m と多く飛んでいるので。
- △ B. B の方が分散が大きいから安定した記録ではないが、とびぬけて大きい記録が出る可能性があるから。
- A. 分散の値が小さいため、結果の値にばらつきが少なく、安定した結果を求めることができるため。
- A. A の選手を選びます。A は分散が少ないので安定した投球ができる。大会でいつもどおりに投げれたら大きな失敗はなく、そこそこいい順位をとることができる。しかし大きな成功もないのでやはりそこそこいい順位しかとれない。
- △ A. A は分散が小さいので、各回の記録をしっかりとれるし、安定感を得られるので、A を選んだ。
- △ A. A の方が分散が小さいし、計算のばらつきも小さいので安定した結果が残るから。
- △ A. なぜなら分散の値が小さいため、大会の結果で安定した記録が期待できるから。
- × B. B の方が A より大きな記録を出すことがあるので、これからも出してけると期待できるから。
- △ A. B の選手は試技の結果にばらつきがあり、本番で思わぬ好記録が出るかもしれない。だが逆に大失敗となる可能性もある。その点 A の選手は試技の結果から本番でも安定した記録が出せると思うので、A の選手を選びます。
- △ A. A の方が偏差も小さく、変化も小さくなる。安定した結果なので。
- △ B. B の方が分散が大きく、失敗すれば大きな損害となるが、その分成功したときは大きな利益が出るから。
- △ A. A と B、ともに分散した結果、A のが散らばり具合が少ない。なので、B は散らばり具合が大きいので好不調が激しいことがわかった。A は、いつもあまり変わらない記録を出せると思い、選びました。
- A. A の方が分散が小さく、記録のばらつきがより小さい。つまり、本番でも安定した記録が期待できるから A を選んだ。
- △ A. 私はいちかばちかの勝負は好きではないので、分散の小さい A を選びます。
- × A. 分散が小さいから、本番も同じくらいの記録が出せると予想できるから。

本来ならば「推定」の考えを利用して、信頼度 95% でここからここまでの範囲の記録が残せるとすべきなんだろうと思います（そのときは標本数が少ないからスチューデントの t 分布を使うのでしょうか？）。しかし、単純に分散が大きい、小さいでどんなことが読み取れるかということを理解させれば OK という考えていましたので、このような問題にしました。ちなみに、この問題の採点基準は A を選んだ場合で言うと、①分散が B よりも小さいこと、②分散が小さいとはどういうことか（散らばりが少ない、安定していること）、③だから本番ではどういうことが予想されるのか、の 3 つにしました。

4 まとめ

先日、ある討論番組で若者のコミュニケーション能力が低くなった要因を、家庭の会話で単語しか使わないことにあると指摘した方がいました。例えば、「ケーキ」と子どもが言えば、ケーキがおやつに出てくるので、いちいち文で言わなくても通じてしまう環境で育ったからだというのです。確かに一理あるなと思いました。確かに、中央値の説明でただ「まんなかの数」と答えたのはそういうことだったのかとうなずいてしまいました。もちろんそれだけがすべての原因ではないと思いますが、「何が」「何を」「何の」などのところ（5W1Hというのでしょうか）を意識していない生徒が多いように思います。

現在、3年生の担任をしているので、生徒の履歴書や願書などの志望動機を添削する機会が多いのですが、一文がものすごく長い、長いから主語と述語の対応が合っていない、おまけに「てにをは」を間違えているからますます読みにくい、などのさまざまなおかしい文章を目にします。数学の教師とはいえども、文章を書く指導というのは日ごろの生徒指導（進路指導）の中で当然求められてきます。

さて、今回テストでこのような問題を出すことができたのは、採点する人間が私1人だけであるからだと思います。複数で担当すると、採点基準を厳密にして採点する人間が誰であっても同じ基準で採点しないと不公平になるからです。その辺の問題をどうクリアするかは今後の課題ですが、統計の問題は他の分野に比べて実生活に例えて出題することができるので、（一般的な数学の記述式の解答とは別の）自分の考えを述べさせる問題を作るのに適していると思います。もちろん改善の余地はたくさんあると思いますが、数学という授業の中でも生徒の文章作成能力を鍛えることは十分にできると感じました。