

盛岡三高数学科通信

How do you solve?

How do you teach?

第16号

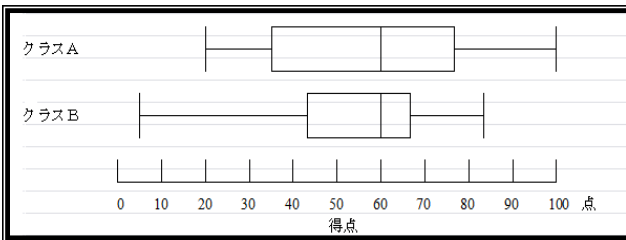
発行責任者
盛岡第三高等学校
下町壽男

箱ひげ図には気をつけろ

1年生が、数学Iで「箱ひげ図」の指導に入ったようなので、今回は箱ひげ図の話をしてみたいと思います。

以下の問題は、ある県が、「箱ひげ図」の導入に伴って2010年から先行指導を実施した際の、評価問題です。

【参考問題1】



上の箱ひげ図は、同じ人数のクラスに実施した中間テストの得点の分布を表している。以下の問いに答えよ。

- ①最高得点は、どちらのクラスの方が、高いですか。
○をつけよ。
- ②最低得点は、どちらのクラスの方が、低いですか。
○をつけよ。
- ③どちらのクラスの方が、30点以下の生徒の人数が多いですか。○をつけよ。
- ④どちらのクラスの方が、60点以上の生徒の人数が多いですか。○をつけよ。

【解答選択肢】

- a. クラスA
- b. クラスB
- c. 両方のクラスがほぼ等しい
- d. どちらということはいえない

研究では、2010年から2011年で正答率が次の表のように推移し、指導の成果が得られたと結んでいます。

質問項目	年度	a	b	c	d
①	2011	100	0	0	0
	2010	100	0	0	0
②	2011	0	100	0	0
	2010	10	90	0	0
③	2011	0	25	12.5	62.5
	2010	0	70	20	10
④	2011	25	0	75	0
	2010	70	0	20	10

表1：生徒の反応率(単位：%)

さて、この表を見ると、④の正解は「c」ということですが、本当にそれでよいのでしょうか。

例えば、2つのクラスの人数を15人とします。すると、第1四分位数、中央値、第3四分位数はデータをソートしたときに、それぞれ、4, 8, 12番目のデータとなります。

今、クラスA、クラスBのデータを次のようにしてみます。

クラスA
20 25 30 35 60 60 60 60 61 62 75 80 90 100
クラスB
5 10 30 45 50 55 57 60 61 62 65 70 75 85

上のデータでは、確かに「参考問題1」にあるような箱ひげ図ができます。

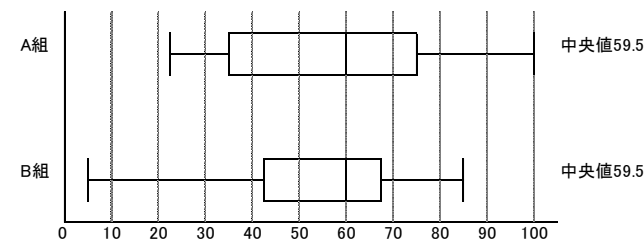
しかし、クラスAは60点以上の人数は11人(約73%)であり、クラスBの60点以上の人数は8人(約53%)です。

このような例はいくらでも作ることができるから、解答はcではなく、dとすべきでしょう。

今年度、岩手県が行った基礎力確認調査では、この問題を参考にして、次のような問題が出題されています。

【参考問題2】

次の図は、ある高校で行った2学年のA組、B組の数学のテストの結果を箱ひげ図に表したものである。A組、B組の生徒の数はどちらも40名で、中央値はどちらも59.5である。このとき、次の(1)～(3)の問いの答えとして、箱ひげ図から読み取れることで最も適当なものを下の選択肢①～④からそれぞれ1つ選び、その番号を書きなさい。ただし、テストの得点は0～100(点)の範囲の整数とする。



- (1) 最高得点は、どちらの組の方が高いですか。
- (2) 30点以下の生徒の人数はどちらの組の方が多いですか。
- (3) 60点以上の生徒の人数は、どちらの組の方が多いですか。

【選択肢】① A組である ② B組である

③ A組とB組の両方が等しいといえる
④ A組とB組のどちらともはいえない
ポイントは中央値が比整数になっていることです。つまりデータの20番目の得点と21番目の得点が変わっているんですね。そうすると、前の問題のようなケースは生じません。いやはや。問題を作る方もなかなか大変です(金濱先生ご苦労様でした)。

こうなってくると、四分位数の境界上の変量にばかりに目を奪われ、本来の、図からデータを分析するという方向に指導が進まないと思います。

元凶はデータの個数によって求め方を変える文科省の四分位数の算出法です。なぜ、このような矛盾を内包するような四分位数の定義にしたのか、恐らく、パソコンを用いずに、手作業で計算を行える(テストの問題や授業での説明に乗り易い)からだろうと推測できます。

本来ならエクセルのパーセンタイル関数を用いれば済むのですが、教室やセンター試験会場にPCを持ち込むわけにはいかないという問題があります。

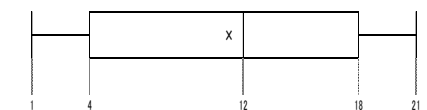
その結果として、手作業でソートできる程度の変量の個数が少ないデータを取り扱うことになります。それによる問題点は以下の通りです。

- ① 将来社会で扱う(かもしれない)箱ひげ図や四分位数はPCを用いて行うはず。それと違うことを生徒に教えることになる。
- ② そもそも手計算で算出できるようなレベルのデータならば何も箱ひげ図を持ち出すこともない。
- ③ 四分位数の境界上にある点の取り方によってはいくらかでも不合理な箱ひげ図を作ることができる。特に変数の数が少なければ尚更である。
- ④ 結局、データの個数が $4n, 4n+1, 4n+2, 4n+3$ という場合分けにこだわるような、およそデータの分析とは関係ない、不定方程式や整数問題が試験に出されることが予想できる。

特に、④については、私は、箱ひげ図が登場してきたときから言い続けてきました。そうしたら、やはり、昨年、河合塾の清先生が、大学入試の予想問題としてやはり次のような問題を取り上げていました。

【参考問題3】

次の11個の整数値からなるデータがある。
1, 3, 7, 9, 12, 14, 18, 21, a, b, c
ただし、 $a < b < c$ とする。このデータの平均は11であり箱ひげ図は下のようになる。



- (1) 考える a, b, c の組みをすべて求めよ。
- (2) さらに、このデータの数値のうち1つの数値を除き10個の整数値からなるデータを作ったところ、平均値と中央値が一致したという。このような a, b, c の値を求めよ。

これはデータ分析の問題というより、整数問題(不定方程式・不等式)ではなからうか。