

数学 I 「データの分析」について

秋田県立横手清陵学院高等学校 今川 直行
※北海道・北東北3県教員人事交流派遣

数学 I 「データの分析」の授業を担当して気がついたことや、教科書にあったらいいなあ、授業で扱えばよかったなあと感じたデータの散布度(散らばりの度合い)や代表値を記します。

数学 I の教科書にはおおよそ次のように記載されています。

- ・データの平均値の周りに、データの各値がどのように分布しているかを示す値として、まず各値と平均値との差を考えてみる。
- ・データの値 x_1, x_2, \dots, x_n の平均値を \bar{x} とするとき、各値と平均値との差を偏差という。偏差の総和は 0 になるから、偏差の平均値も 0 である。
- ・よって、偏差の平均値では、散らばりの度合いを表すことはできない。そこで、偏差をそのまま用いずに、偏差の 2 乗の平均値を考えよう。
- ・【偏差の 2 乗の平均値(分散)を求める例題】
- ・【分散と標準偏差についてのまとめ】
- ・データの値が平均値の周りに集中しているほど、それぞれの偏差の絶対値は小さくなり、分散、標準偏差も小さくなる傾向にある。

この記事を授業前に読んで気がついたことは「偏差の絶対値の平均値」が散らばりの度合いを表すものにならないのか、ということです。

そこで、この「偏差の絶対値の平均値」を平均偏差ということや、さらに調べたことを次のようにまとめました。

【平均偏差】

偏差 $x - \bar{x}$ の絶対値をとって平均した値

$$\frac{1}{n} \left(|x_1 - \bar{x}| + |x_2 - \bar{x}| + \dots + |x_n - \bar{x}| \right)$$

参考にした文献には、平均偏差についておおよそ次のように記載されています。

- ・データの散布度を計るものとしては自然なものであるが、絶対値の扱いが面倒である。
- ・全てのデータを用いて散らばりの度合いを測るという点では優れているが、絶対値の扱いが面倒なためあまり使われない。

分散や標準偏差を記載する前に、あまり使われなく扱いが面倒であっても、散らばりの度合いを考える過程として、「偏差の絶対値の平均値」の記載が教科書にあったらいいなあと思いました。

次に、この単元の授業時に扱えばよかったなあと後になって感じたデータの代表値を記します。

トリムド・ミーンと呼ばれるデータの代表値で、例えばスキージャンプ競技の飛型点を採点するときに用いられます。数人の審査員が採点した点数の中で、最高点と最低点を省きます。これは公正な採点をするためであると考えます。極端な例ではありますが、好きな選手には実力に関係なく満点にしよう、または嫌いだから点数を低くしよう、ということを防ぐことができます。

【トリムド・ミーン】

データの値 x_1, x_2, \dots, x_n を小さい順に並べたとき、最大値 x_n と最小値 x_1 を取り除いた残りの平均値

$$\frac{1}{n-2} (x_2 + x_3 + \dots + x_{n-1})$$

札幌などでスキージャンプ競技が行われます。現在秋田県の学校に勤めており、そこでよく北海道にかかわる話をするのですが、この単元を授業で扱ったときに、この代表値の話ですればよかったなあと思いました。

【参考文献】

- 1) 新編数学 I 数研出版
- 2) 入門統計学 共立出版
- 3) 例解統計学 S I 企画