数学Ⅰ　第５章　データの分析　No.2

学習のねらい

四分位数、四分位範囲、箱ひげ図を理解しよう！！

1.四分位数、四分位範囲

データの最大値と最小値の差を範囲という。しかし、範囲では、散らばり具合を比較する量として適切でない場合もある。

そこで、データの値の大きさの順に並べて４等分する。４等分する位置にくる値を四分位数という。四分位数は、小さい方から、第１四分位数、第２四分位数、第３四分位数といい、これらをQ_1、Q_2、Q_3と表す。Q_2は中央値である。

例）以下のデータを考えると、

下位のデータ　　上位のデータ

2　3　5　7　11　13　17　19　23

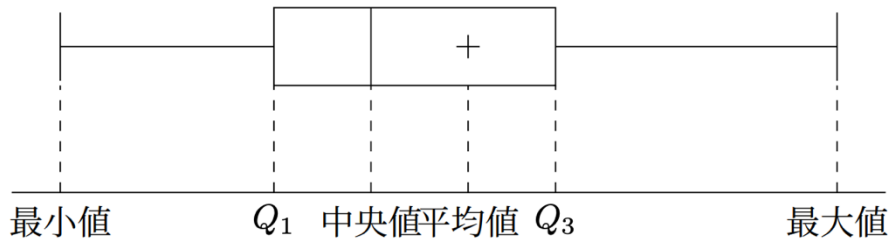
↑ Q_1　↑ Q_2　↑ Q_3

中央値を境にして左側を「下位のデータ」右側を「上位のデータ」と言い、下位のデータの中央値をQ_1とし、上位のデータの中央値をQ_3とする。$Q_1=\frac{3+5}{2}=4$、$Q_2=11$、$Q_3=\frac{17+19}{2}=18$となる。

また、Q_3-Q_1の値を四分位範囲といい、$\frac{Q_3-Q_1}{2}$を四分位偏差という。

2.箱ひげ図

データの分布を見るために、図のような図を考えることがある。最小値、第一四分位数、中央値、第三四分位数、最大値、(平均値)を視覚的に確認できる図である。

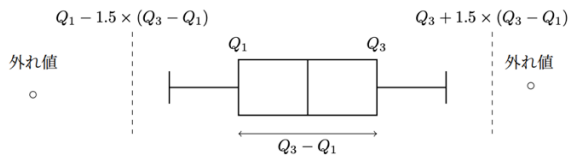


3.外れ値

例えば、クラスの人のスマホを見ている時間を調べたとしよう。

2、2、3、3、4、4、4、5、5、18

このようになったとき、平均は、5時間になる。「じゃあ！このクラスはみんないっぱい触っている！！」という結論にならないのは直感的に分かるだろうか。18という他の値から極端にかけ離れた外れ値があるせいで、平均時間が高くなってしまっているのだ。外れ値の基準は様々あるが、例えば、「$Q_1-1.5\times(Q_3-Q_1)$以下」または「$Q_3+1.5\times(Q_3-Q_1)$以上」とする。箱ひげ図は、以下のようになる。



このようなとき、外れ値を○で表す。

©算数・数学のアトリエ

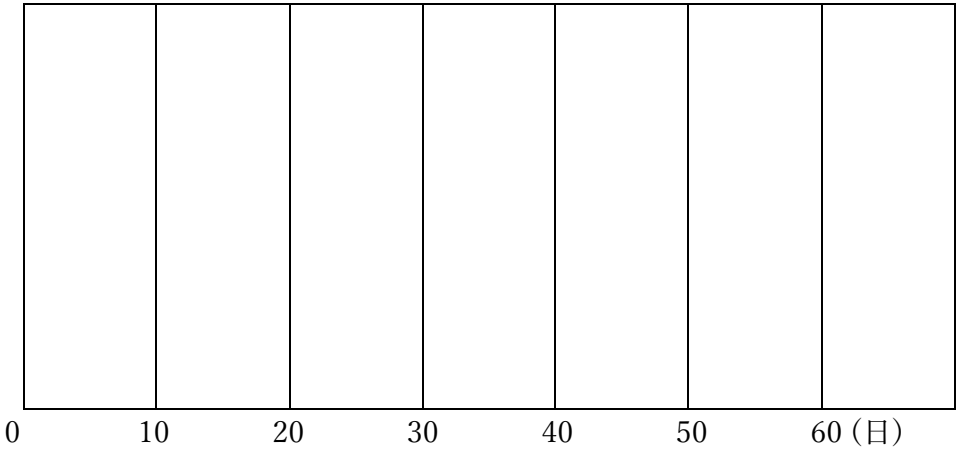
◊問題

1.　次のデータは、2 つの都市 A、B について、最低気温が 25℃以上であった日の日数を 1 年ごとに集計した 12 年分の結果である。(単位は日)。

A	19	31	46	41	25	33	17	41	31	20	31	25
B	25	46	37	57	48	48	30	51	46	44	44	42

このデータから、下の表を埋め、箱ひげ図を完成させよ。

	最小値	Q_1	中央値	Q_3	最大値
A					
B					



©算数・数学のアトリエ