

第36回 データ分析①

データの代表値

データ全体の特徴を表す適当な一つの数値

・平均値 : データの総和を総度数で割った値

例) 1, 2, 4, 5 $\Rightarrow \frac{1+2+4+5}{4} = 3$

・中央値(メジアン) : データを値の大きさの順に並べたときの中央の位置にくる値

例) データの個数が奇数 1, 1, 7, 11, 16 データの個数が偶数 2, 5, 6, 9, 12, 12

・最頻値(モード) : 最も度数が多いデータの値

例) 2, 2, 3, 3, 3, 5, 5, 9, 10

最頻値 3

Pattern. 1 平均値・中央値・最頻値

★POINT★

・データは小さい川頁に並べる

・中央値はデータ個数が偶数か奇数かに注意!

(例題 1) 次の問いに答えよ。

(1) 次のデータの平均値、中央値、最頻値をそれぞれ求めよ。

3, 5, 5, 3, 2, 14, 3, 5

(2) 10 人がテストをした結果以下のような結果になりました。このとき、平均値、中央値、最頻値をそれぞれ求めよ。

階級	0 以上 20 未満	20 以上 40 未満	40 以上 60 未満	60 以上 80 未満	80 以上 100 未満	合計
階級値	10	30	50	70	90	
度数	0	1	2	2	5	10

(1) 平均値 $\frac{3+5+5+3+2+14+3+5}{8} = 5$

中央値 $\frac{3+5}{2} = 4$

最頻値 3 と 5

(2) 平均値 $\frac{10 \times 0 + 30 \times 1 + 50 \times 2 + 70 \times 2 + 90 \times 5}{10}$

$= 72$

(1) の順: 30, 50, 50, 70, 70, 90, 90, 90, 90, 90

中央値 $\frac{70+90}{2} = 80$

最頻値 90

データの散らばりと四分位範囲

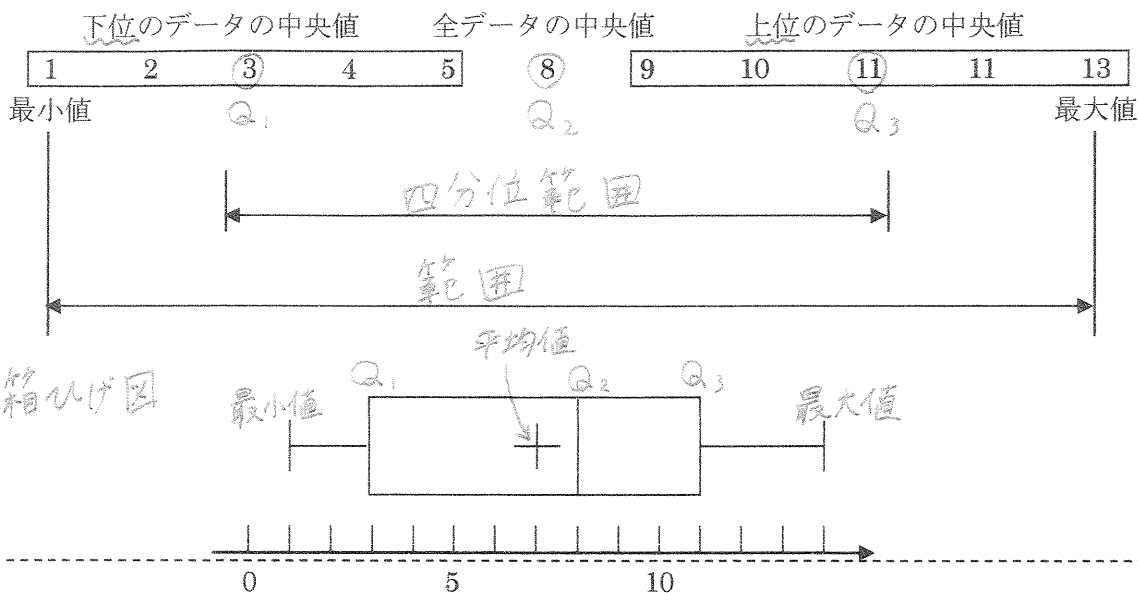
・**範囲**：データの最小値と最大値の差

・**四分位数**：データの値を小さい順に並べたとき、4等分する位置にくる3つの数字
しぶんい

小さいほうから第1四分位数(Q_1)、第2四分位数(Q_2)、第3四分位数(Q_3)

・**四分値範囲**： $Q_3 - Q_1$ ・**四分位偏差**： $\frac{Q_3 - Q_1}{2}$

・**5数要約**：最小値、第1四分位数、第2四分位数、第3四分位数、最大値



Pattern. 2 四分位数

★POINT★

第1四分位数：下位のデータの中央値 …… Q_1

第2四分位数：全データの中央値 …… Q_2

第3四分位数：上位のデータの中央値 …… Q_3

(例題2) 以下のデータについて5数要約、範囲、四分位範囲、四分位偏差を求めよ。また、5数要約を箱ひげ図に表せ。

20, 14, 9, 9, 13, 15, 16, 22, 7, 5 → 5, 7, ⑨, 9, 13, 14, 15, ⑪, 20, 22

最小値 5

$Q_1 = 9$

$Q_2 = \frac{13+14}{2} = 13.5$

$Q_3 = 16$

最大値 22

範囲：最大値 - 最小値 = $22 - 5 = 17$

四分位範囲： $Q_3 - Q_1 = 16 - 9 = 7$

四分位偏差： $\frac{Q_3 - Q_1}{2} = \frac{7}{2} = 3.5$

(平均値 = $\frac{5+7+9+9+13+14+15+16+20+22}{10} = 13$)

箱ひげ図

