

第 3 章 資料の整理

§ 1 1次元のデータ

1.1 代表値

データ, 量的変数, 代表値とは?

- 観測, 調査, 実験などの測定結果として得られた**変数**の集まりを, **データ**と呼ぶことにする。
- データには, 体重や温度など数量的に扱える**量的データ**と性別や満足度など数量的に扱えない**質的データ**があります。(最終頁【参考】も参照)
- データの特色を1つの数値で表したものを**代表値**という。

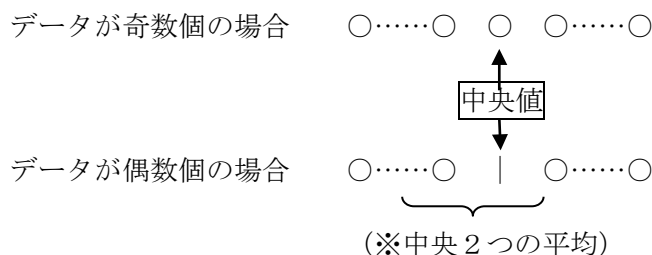
平均値, 中央値, 最頻値とは?

○代表値には, **平均値**(Mean), **中央値**(Median), **最頻値**(Mode)があります。

- 1) 変数 x におけるデータが, n 個の値 $x_1, x_2, x_3, \dots, x_n$ であるとき, それらの総数を n で割ったものをデータの**平均値**といい, μ で表す。

$$\mu = \frac{x_1 + x_2 + x_3 + \dots + x_n}{n} \quad (\mu \text{ の読み: ミュウ})$$

- 2) データをその値の大きさの順に並べたときに中央にくる値を**中央値**といいます。ただし, データが偶数個のときは, 中央2つのデータの平均値を中央値とします。



- 3) データの中でもっとも個数の多い値をそのデータの**最頻値**といいます。

例題 次の表は男子学生 15 名の身長データのデータである。(単位 : cm)

179 166 173 171 177 180 168 172

174 177 172 185 173 169 172

このデータについて、次の問いに答えよ。

- (1) 最頻値を求めよ。 (2) 中央値を求めよ。 (3) 平均値を求めよ。

[解答] 与えられたデータを小さいものから順に並べると

166 168 169 171 172 172 172 173

173 174 177 177 179 180 185

となる。

- (1) 最頻値は 172 [※データ数 3 個]
 (2) 中央値は 173 [※奇数なので、8 個目のデータ]
 (3) 平均値は 173.9 [※実際の計算は、次式を参照]

$$\begin{aligned}\mu &= \frac{166+168+169+171+172+172+172+173+173+174+177+177+179+180+185}{15} \\ &= 173.86\dots\end{aligned}$$

問 3.1 次のデータの平均値を求めよ。

10 25 60 30 45

問 3.2 次のデータの中央値を求めよ。

(1) 10 20 20 30 40 60 60 70 90

(2) 10 15 20 25 35 40 40 45 90 95

問 3.3 次のデータの最頻値を求めよ。

12 14 16 16 16 18 19 20 20 20 21

【注意：代表値の特色】

平均値 全データがもつ情報を有効利用。外れ値の影響を受けやすい。
平均値は 1 つだけ存在。

中央値 データは並べ替えのために使用。外れ値の影響を受けにくい。
中央値は 1 つだけ存在。

最頻値 データの値の個数調査のために使用。外れ値の影響を受けにくい。
最頻値が複数ある場合がある。

〔※データ全体の中で、他のデータの値に比べて極端に大きかったり
小さかったりする少数のデータの値を**外れ値**という。
また、外れ値の影響を受けにくい性質を**抵抗力がある**といいます。〕

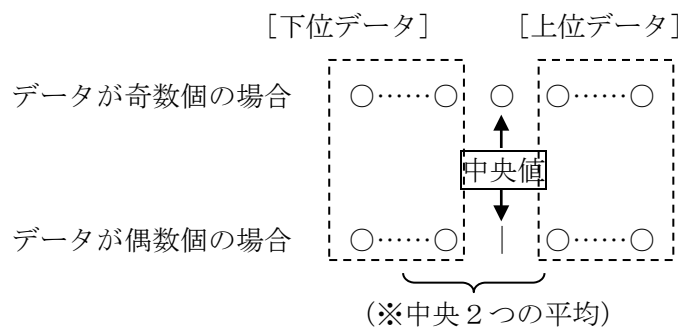
1.2 箱ひげ図

箱ひげ図とは？

次の 9 個の内容を組み込んだ図を、**箱ひげ図**と呼びます。

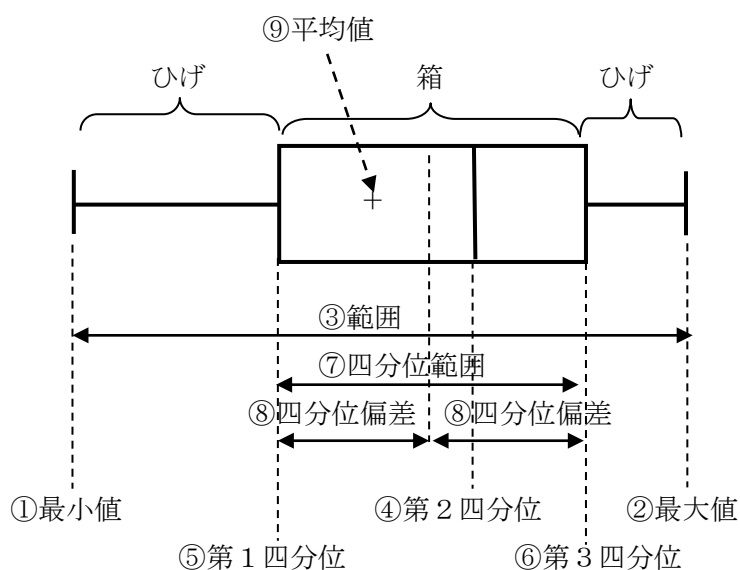
- ①最小値：データの値の中で一番小さな値
- ②最大値：データの値の中で一番大きな値
- ③範囲：データが存在する①から②までの幅

④中央値(第 2 四分位数)：小さい順に並べたときの中央にくる値



- ⑤第 1 四分位数：下位データの中央値
- ⑥第 3 四分位数：上位データの中央値
- ⑦四分位範囲：⑥から⑤を引いた「箱」の長さ
- ⑧四分位偏差：四分位範囲を半分にした長さ

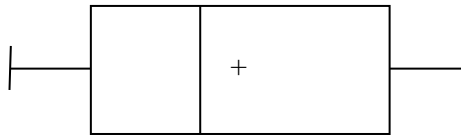
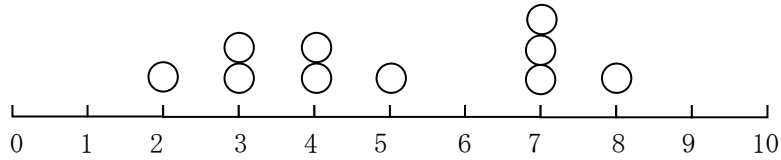
⑨平均値：データ全体の平均値(「+」で図中に表記)。



例題 次のデータに基づき、箱ひげ図を作成せよ。

小テストの点数：2, 3, 3, 4, 4, 5, 7, 7, 7, 8

[解答]



- ①最小値 2 ②最大値 8 ③範囲 6
- ④中央値 4.5 ⑤第1四分位数 3 ⑥第3四分位数 7
- ⑦四分位範囲 4 ⑧四分位偏差 2 ⑨平均值 5

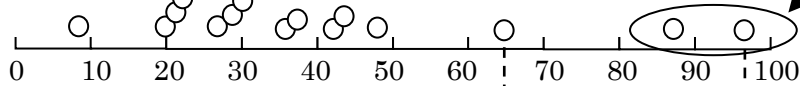
問3.4 次のデータに基づき、箱ひげ図を作成せよ。

小テストの点数：2 2 4 4 4 7 7 7 8

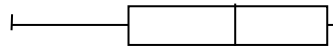
【研究：外れ値の検討】 「ひげの長さ」は、「四分位範囲の1.5倍」までが目安です。

[※外れ値?]

例) テストの成績：9, 20, 21, 22, 27, 29, 30, 36, 38, 42, 43, 48, 65, 88, 97



最小値 9, 第1四分位 22, 第2四分位 36, 第3四分位 48, 最大値 97



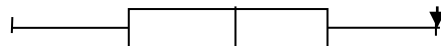
[※右側のひげが長過ぎ?]

適切なひげの長さ (第3四分位 - 第1四分位) × 1.5

$$(48 - 22) \times 1.5 = 39$$

つまり、今回は第3四分位に理想のひげの長さ 39 を足すと 48 + 39 = 87

よって、「87」以内に含まれる最大値を「65」でひげを書き直します。



(※「88」「97」は、ひげの外に残し、今回は「特に優秀」として検討する。)

