【中学数学】 第 5 章 データの活用

体系数学 2/数研出版/代数編

Riku Sugawara

12.2025

第 1 節　データの整理

1.1　度数分布表とヒストグラム

Definition 1.1　度数分布表

データの範囲を適当に区切ったとき, 各区間に含まれるデータの個数を **度数** といい,
各区間にその区間の度数を対応させて整理した右のような表を **度数分布表** という .

階級（cm）	度数（人）
135 以上〜140 未満	2
140〜145	4
145〜150	5
150〜155	8
155〜160	11
160〜165	9
165〜170	7
170〜175	4
計	**50**

表 1.1　身長の度数分布表

Definition 1.2　階級

度数分布表において, 区切られた各区間を **階級** , 区間の幅を **階級の幅** ,
各階級の中央の値を **階級値** という .

Definition 1.3　ヒストグラム

度数分布表を, 柱状のグラフで表したものを **ヒストグラム** という .

ヒストグラムの各長方形の横の長さは階級の幅を表し, 高さは各階級の度数を表す .

Definition 1.4　度数折れ線

ヒストグラムの各長方形の上の辺の中点を結んでできる折れ線グラフを **度数折れ線** という .

ただし, 度数折れ線をつくる時は, ヒストグラムの左右の両端に度数が 0 の階級があるものと考える .

例題 【教 p.114 追加例題 1】

表 1.1 において, 次の問いに答えなさい．

(1) 階級の幅は何か．

(2) 階級の個数はいくつか．

(3) 階級 140cm 以上 145cm 未満の度数はいくつか．

(4) 階級 140cm 以上 145cm 未満の階級値はいくつか．

1.2 相対度数

Definition 1.5 相対度数

度数の合計に対する各階級の度数の割合を **相対度数** という．

相対度数はふつう小数を用いて表す．

Theorem 1.1 相対度数

相対度数は次の式で求められる．

$$\text{(相対度数)} = \frac{(\qquad\qquad\qquad)}{(\qquad\qquad\qquad)}$$

Definition1.5 より, 相対度数の合計は必ず　　　　になる．

また, 相対度数を使用するメリットは, 度数の合計が異なる複数の分布について,

　　　　を比較することができる．

1.3 累積度数

Definition 1.6 累積度数

度数分布表において, 各階級以下または各階級以上の階級の度数をたし合わせたものを **累積度数** という．

また, 累積度数を表にまとめたものを **累積度数分布表** という．

Definition 1.7　累積相対度数

累積度数についても, 度数の合計に対する各階級の累積度数の割合を考えることがある．

この割合を **累積相対度数** という．

例題 【教 p.119 追加例題 2】

表 1.1 において, 次の問いに答えなさい．

(1) 各階級の相対度数を求めなさい．

(2) 累積度数の列を加え, 累積度数分布表を作成しなさい．

(3) 累積相対度数を求めなさい．

(4) 身長が 150cm 以上の人の割合は何 % か求めなさい．

Proof. □

第 2 節　データの代表値

データの特徴を 1 つの値で表すとき，その値を**代表値**といいます．

代表値には，平均値，中央値（メジアン），最頻値（モード）があります．

2.1　平均値

Definition 2.1

平均値（mean）すべてのデータの値を足して，データの個数で割った値．

$$平均値 = \frac{データの値の合計}{データの個数}$$

記号では，データを $x_1, x_2, \ldots, x_n$ とすると，平均値 $\overline{x}$ は次のように表される．

$$\overline{x} = \frac{x_1 + x_2 + \cdots + x_n}{n} = \frac{1}{n}\sum_{i=1}^{n} x_i$$

特徴：

(i) すべてのデータが平均値に関わる

(ii) 極端に大きい値や小さい値の影響を受けやすい

例題 3

次のデータの平均値を求めなさい．

$$5,\ 8,\ 6,\ 9,\ 7,\ 10,\ 6,\ 8,\ 7,\ 9$$

【解答】

$$\begin{aligned}\overline{x} &= \frac{5+8+6+9+7+10+6+8+7+9}{10} \\ &= \frac{75}{10} \\ &= 7.5\end{aligned}$$

よって，平均値は **7.5** である．

問題 1 平均値の計算

次の各データの平均値を求めなさい．

（1） 3, 5, 7, 9, 11

（2） 12, 15, 18, 14, 16, 13, 17

（3） 20, 25, 22, 28, 30, 24, 26, 23

2.2 度数分布表からの平均値

度数分布表が与えられている場合, 階級値を使って平均値を求めます．

Definition 2.2

度数分布表からの平均値度数分布表において, 各階級の階級値を x_i, 度数を f_i とすると, 平均値は次のように求められる．

$$\overline{x} = \frac{x_1 f_1 + x_2 f_2 + \cdots + x_n f_n}{f_1 + f_2 + \cdots + f_n} = \frac{\sum x_i f_i}{\sum f_i}$$

計算手順：

（1）各階級の階級値と度数の積を求める

（2）すべての積の合計を求める

（3）合計を総度数で割る

例題 4

次の度数分布表から, 平均値を求めなさい．

階級（点）	階級値（点）	度数（人）
50 以上～60 未満	55	2
60 以上～70 未満	65	6
70 以上～80 未満	75	9
80 以上～90 未満	85	3
合計	-	**20**

【解答】

各階級値と度数の積を計算する．

$$\begin{aligned}\overline{x} &= \frac{55 \times 2 + 65 \times 6 + 75 \times 9 + 85 \times 3}{20} \\ &= \frac{110 + 390 + 675 + 255}{20} \\ &= \frac{1430}{20} \\ &= 71.5\end{aligned}$$

よって, 平均値は **71.5 点** である．

2.3　中央値（メジアン）

Definition 2.3

中央値（median）データを大きさの順に並べたとき, 中央に位置する値．

求め方：

（i）データの個数が奇数のとき：真ん中の値

（ii）データの個数が偶数のとき：真ん中の 2 つの値の平均

特徴：

（i）データを 2 等分する位置の値

（ii）極端な値の影響を受けにくい

例題 5

次のデータの中央値を求めなさい．

(1) 3, 7, 5, 9, 6, 8, 4

(2) 12, 15, 10, 18, 14, 16

【解答】

(1) データを大きさの順に並べる：3, 4, 5, **6**, 7, 8, 9

データは 7 個（奇数）なので, 4 番目の値が中央値．

よって, 中央値は **6** である．

(2) データを大きさの順に並べる：10, 12, **14**, **15**, 16, 18

データは 6 個（偶数）なので, 3 番目と 4 番目の値の平均が中央値．

$$\text{中央値} = \frac{14+15}{2} = \frac{29}{2} = 14.5$$

よって, 中央値は **14.5** である．

問題 2 中央値の計算

次の各データの中央値を求めなさい．

（1）8, 12, 15, 10, 14, 11, 13

（2）20, 25, 18, 22, 24, 19, 21, 23

2.4 最頻値（モード）

Definition 2.4

最頻値（mode）データの中で最も多く現れる値．度数分布表では, 度数が最も大きい階級の階級値．

特徴：

(i) データの中で最も頻繁に現れる値
(ii) 複数存在する場合もある
(iii) 存在しない場合もある

例題 6

次のデータの最頻値を求めなさい．

$$5, 7, 8, 5, 9, 6, 5, 8, 7, 10$$

【解答】

各値の度数を数える：

値	5	6	7	8	9	10
度数	**3**	1	2	2	1	1

5 が 3 回で最も多く現れる．

よって, 最頻値は **5** である．

2.5 3 つの代表値の比較

代表値のまとめ

代表値	**特徴**	**使い分け**
平均値	すべてのデータを使う	全体的な傾向を知りたいとき
中央値	極端な値の影響を受けにくい	データに極端な値があるとき
最頻値	最も多く現れる値	どの値が多いか知りたいとき

問題 3 代表値の計算

次のデータについて, 平均値, 中央値, 最頻値を求めなさい．

$$6, 8, 7, 9, 6, 10, 7, 8, 6, 9, 7, 8$$

第 3 節　データの散らばりと四分位範囲

データの代表値だけでは, データの特徴を十分に表せないことがあります．

データがどれだけ散らばっているかを表す値として, 範囲, 四分位数, 四分位範囲などがあります．

3.1　範囲（レンジ）

Definition 3.1

範囲（range）データの最大値と最小値の差．

$$\text{範囲} = \text{最大値} - \text{最小値}$$

特徴：

(i) 計算が簡単
(ii) 極端な値の影響を受けやすい
(iii) データの散らばり具合の大まかな目安になる

例題 7

次のデータの範囲を求めなさい．

$$12,\ 18,\ 15,\ 22,\ 14,\ 20,\ 16,\ 19$$

【解答】

最大値は 22, 最小値は 12 なので,

$$範囲 = 22 - 12 = 10$$

よって, 範囲は **10** である．

3.2 四分位数

データを 4 等分する位置の値を四分位数といいます．

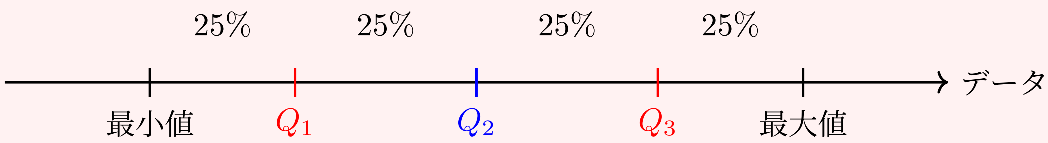
Definition 3.2

四分位数（quartile）データを小さい順に並べたとき, データを 4 等分する 3 つの値．

（i）**第 1 四分位数**（Q_1）：下位 25% の位置の値

（ii）**第 2 四分位数**（Q_2）：下位 50% の位置の値（中央値と同じ）

（iii）**第 3 四分位数**（Q_3）：下位 75% の位置の値



3.3 四分位数の求め方

四分位数の求め方

手順：

（1）データを小さい順に並べる

（2）中央値（Q_2）を求める

（3）中央値より小さいデータの中央値が Q_1

（4）中央値より大きいデータの中央値が Q_3

注意：

（i）データの個数が奇数の場合, 中央値は含めない

（ii）データの個数が偶数の場合, 半分に分けて考える

例題 8

次のデータの四分位数を求めなさい．

$$8, 12, 15, 10, 18, 14, 20, 11, 16, 13, 19$$

【解答】

Step 1: データを小さい順に並べる

$$8, 10, 11, 12, 13, \mathbf{14}, 15, 16, 18, 19, 20$$

データは 11 個（奇数）なので, 6 番目の値が中央値．

$$Q_2 = 14$$

Step 2: Q_1 を求める（中央値より小さいデータ）

$$8, 10, \mathbf{11}, 12, 13$$

5 個（奇数）なので, 3 番目の値が Q_1．

$$Q_1 = 11$$

Step 3: Q_3 を求める（中央値より大きいデータ）

$$15, 16, \mathbf{18}, 19, 20$$

5 個（奇数）なので, 3 番目の値が Q_3．

$$Q_3 = 18$$

よって, $Q_1 = \mathbf{11}$, $Q_2 = \mathbf{14}$, $Q_3 = \mathbf{18}$ である．

問題 4 四分位数の計算

次の各データの四分位数 Q_1, Q_2, Q_3 を求めなさい．

（1）5, 8, 12, 15, 18, 20, 22, 25, 28

（2）10, 14, 16, 18, 20, 22, 24, 26, 28, 30

3.4 四分位範囲と四分位偏差

Definition 3.3

四分位範囲（IQR）第 3 四分位数と第 1 四分位数の差．

$$四分位範囲 = Q_3 - Q_1$$

データの中央 50% の散らばり具合を表す．

特徴：

(i) 極端な値の影響を受けにくい
(ii) データの散らばり具合を安定的に表せる
(iii) 箱ひげ図で視覚的に表現できる

Definition 3.4

四分位偏差四分位範囲の半分の値．

$$四分位偏差 = \frac{Q_3 - Q_1}{2}$$

データの散らばりを表す別の指標．

例題 9

例題 8 のデータについて, 四分位範囲と四分位偏差を求めなさい．

【解答】

例題 8 より, $Q_1 = 11$, $Q_3 = 18$ なので,

$$\begin{aligned} 四分位範囲 &= Q_3 - Q_1 \\ &= 18 - 11 \\ &= 7 \end{aligned}$$

$$\begin{aligned} 四分位偏差 &= \frac{Q_3 - Q_1}{2} \\ &= \frac{18 - 11}{2} \\ &= \frac{7}{2} \\ &= 3.5 \end{aligned}$$

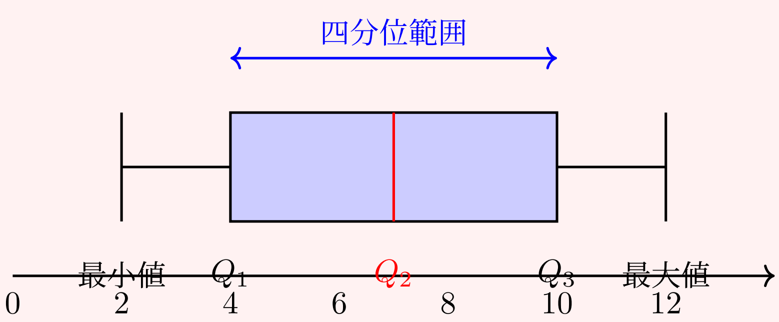
よって, 四分位範囲は **7**, 四分位偏差は **3.5** である．

3.5 箱ひげ図

四分位数を使って, データの分布を視覚的に表したグラフを箱ひげ図といいます．

Definition 3.5

箱ひげ図（box plot）データの最小値, Q_1, Q_2, Q_3, 最大値の 5 つの値を使って, データの分布を表すグラフ．



読み取れる情報：

（i）データの中央値と散らばり

（ii）データの分布の偏り

（iii）複数のデータの比較

問題 5 四分位範囲と箱ひげ図

2 つのクラス A, B で数学のテストを行った．結果は次の通りである．

クラス A： 最小値 50 点, $Q_1$65 点, $Q_2$72 点, $Q_3$80 点, 最大値 95 点

クラス B： 最小値 55 点, $Q_1$62 点, $Q_2$70 点, $Q_3$78 点, 最大値 90 点

(1) 各クラスの四分位範囲を求めなさい．

(2) 各クラスの箱ひげ図を描き, 成績の分布を比較しなさい．

3.6 まとめ：データの散らばりを表す指標

散らばりの指標のまとめ

指標	計算式	特徴
範囲	最大値 − 最小値	簡単だが極端な値の影響を受けやすい
四分位範囲	$Q_3 - Q_1$	極端な値の影響を受けにくい
四分位偏差	$\dfrac{Q_3 - Q_1}{2}$	四分位範囲の半分

問題 6 総合問題

次のデータは, ある学校の生徒 10 人の読書時間（分/日）である．

$$30, 45, 20, 60, 35, 50, 25, 55, 40, 70$$

(1) 平均値, 中央値を求めなさい．

(2) 四分位数 Q_1, Q_2, Q_3 を求めなさい．

(3) 範囲と四分位範囲を求めなさい．

（4）箱ひげ図を描きなさい．