組	番号	名　　前

5章・1節　データの整理と分析

①　データの整理
②　データの代表値
③　データの散らばり

1 次の $\boxed{\quad}$ をうめよ。知

(1)　気温，湿度，降水量のように，ある特性を数量的に表すものを $\boxed{\text{変量}}$ という。データを整理するために用いる区間を階級といい，区間の幅を $\boxed{\text{階級の幅}}$，階級の真ん中の値を $\boxed{\text{階級値}}$，各階級に入っているデータの値の個数をその階級の $\boxed{\text{度数}}$ という。各階級に度数を対応させ，それを表にしたものを $\boxed{\text{度数分布表}}$ という。

(2)　データの特徴を表す数値を $\boxed{\text{代表値}}$ という。データの値の総和をデータの値の個数で割った値を $\boxed{\text{平均値}}$ といい，記号 $\overline{x}$ で表す。
また，データのすべての値を小さい順に並べたとき，中央の位置にある数値を $\boxed{\text{中央値}}$，度数分布表で，度数がもっとも多い階級の階級値を $\boxed{\text{最頻値}}$ という。

(3)　データの値を小さい方から順に並べ，中央値を境にしてデータの値の個数が等しくなるように 2 つの部分に分けたとき，最小値を含む方のデータの中央値を $\boxed{\text{第1四分位数}}$，中央値を $\boxed{\text{第2四分位数}}$，最大値を含む方のデータの中央値を $\boxed{\text{第3四分位数}}$ といい，これらを合わせて $\boxed{\text{四分位数}}$ という。また，データの最小値，最大値と四分位数を用いて要約する方法を $\boxed{\text{5数要約}}$ といい，これを，箱と線(ひげ)を用いて 1 つの図に表したものを $\boxed{\text{箱ひげ図}}$ という。

(4)　データの最大値から最小値を引いた値を $\boxed{\text{範囲}}$ または $\boxed{\text{レンジ}}$ という。
また，第3四分位数から第1四分位数を引いた値を $\boxed{\text{四分位範囲}}$，四分位範囲を2で割った値を $\boxed{\text{四分位偏差}}$ という。

(5)　データの値が x_1，x_2，…，x_n で平均値が $\overline{x}$ のとき，$x_1-\overline{x}, x_2-\overline{x}$，…，$x_n-\overline{x}$ をそれぞれの値の平均値からの $\boxed{\text{偏差}}$ という。また，
$\frac{1}{n}\{(x_1-\overline{x})^2+(x_2-\overline{x})^2+\cdots+(x_n-\overline{x})^2\}$ を $\boxed{\text{分散}}$ といい，s^2 で表す。
さらに，この値の正の平方根を $\boxed{\text{標準偏差}}$ といい，s で表す。

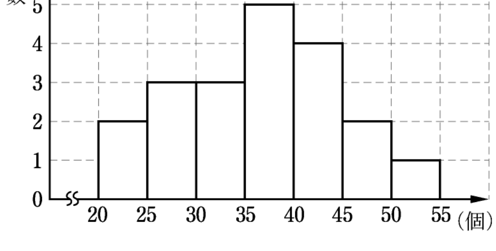
2 次のデータは，ある国語辞典から 20 ページを任意に選び，各ページの見出し語の個数を調べた結果である。技

32	46	22	38	28	49	36	24	27	40	
36	44	34	42	43	39	25	33	52	38	(単位　個)

(1)　このデータについて，下の度数分布表を完成させよ。

個数の階級 個以上　個未満	階級値 x	度数 f	xf	相対度数
20～25	22.5	2	45.0	0.10
25～30	27.5	3	82.5	0.15
30～35	32.5	3	97.5	0.15
35～40	37.5	5	187.5	0.25
40～45	42.5	4	170.0	0.20
45～50	47.5	2	95.0	0.10
50～55	52.5	1	52.5	0.05
計		20	730	1.00

(2)　度数分布表をもとに，このデータのヒストグラムをかけ。



3 **2** の見出し語の個数のデータについて，次の問に答えよ。技

(1)　このデータの平均値を次の 2 つの方法で求めよ。

①　20 個のデータの値の総和を利用する方法

［解］　$\frac{1}{20}(32+46+22+38+28+49+36+24+27+40$
$+36+44+34+42+43+39+25+33+52+38)=\frac{728}{20}=$ **36.4 (個)**

②　度数分布表を利用する方法

［解］　$\frac{1}{20}(22.5\times2+27.5\times3+32.5\times3+37.5\times5+42.5\times4+47.5\times2+52.5\times1)$
$=\frac{730}{20}=$ **36.5 (個)**

(2)　このデータの中央値を求めよ。また，**2** の(1)で作成した度数分布表から最頻値を求めよ。

［解］　20 個のデータを小さい順に並べたとき，中央値は 10 番目のデータ 36 と 11 番目のデータ 38 の平均値であるから $\frac{36+38}{2}=$ **37 (個)**
また，度数分布表において度数が最も多い階級は「35 個以上 40 個未満」であるから，最頻値はその階級の階級値　**37.5 個**

4 次のデータは，ある地点における湿度を 10 日間調べた結果である。技

51	39	48	42	77	62	59	84	53	66	(単位　%)

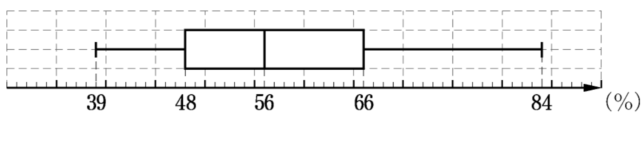
(1)　このデータの第1四分位数，第2四分位数，第3四分位数を求めよ。

［解］　データを小さい方から順に並べると　39，42，48，51，53，59，62，66，77，84
第1四分位数は前半5個のデータの中央値であるから　**48%**
第2四分位数はデータの中央値であるから　$\frac{53+59}{2}=$ **56 (%)**
第3四分位数は後半5個のデータの中央値であるから　**66%**

(2)　このデータの四分位範囲と四分位偏差を求めよ。

［解］　四分位範囲は第3四分位数から第1四分位数を引いた値であるから 66−48=**18(%)**
四分位偏差は四分位範囲を2で割った値であるから $\frac{18}{2}=$ **9 (%)**

(3)　このデータの箱ひげ図をかけ。



5 次の表は，A，B 2 人の 5 回のテストの得点である。技

回	1	2	3	4	5
A の得点	8	10	4	6	7

回	1	2	3	4	5
B の得点	7	9	5	6	9

(1)　A，B それぞれの得点の平均値を求めよ。

［解］　A の得点の平均値は　$\frac{1}{5}(8+10+4+6+7)=$ **7 (点)**
B の得点の平均値は　$\frac{1}{5}(7+9+5+6+9)=$ **7.2 (点)**

(2)　A，B それぞれの得点の分散を求めよ。

［解］　A の得点の各値の偏差は，順に 1，3，−3，−1，0 であるから，A の得点の分散は
$\frac{1}{5}\{1^2+3^2+(-3)^2+(-1)^2+0^2\}=\frac{20}{5}=$ **4**
B の得点の各値の偏差は，順に −0.2，1.8，−2.2，−1.2，1.8 であるから，B の得点の分散は　$\frac{1}{5}\{(-0.2)^2+1.8^2+(-2.2)^2+(-1.2)^2+1.8^2\}=\frac{12.8}{5}=$ **2.56**

(3)　A，B それぞれの得点の標準偏差を求めよ。

［解］　(2)より，A の得点の標準偏差は $\sqrt{4}=$ **2 (点)**
B の得点の標準偏差は $\sqrt{2.56}=$ **1.6 (点)**