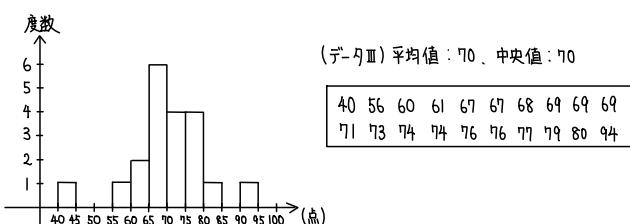
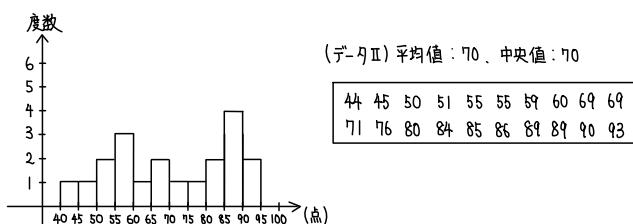
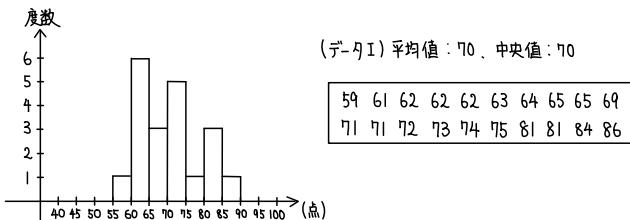


4

・データの散らばりと四分位範囲



平均値や中央値が同じでも散らばっているように見える

↳ どのくらい散らばっているか

四分位範囲 $Q_3 - Q_1$ (データの散らばりの度合いを表す量)

$$\text{四分位偏差 } \frac{Q_3 - Q_1}{2}$$

※ 四分位範囲は、中央50%のデータの範囲にほぼ等しく極端に離れた値の影響を受けにくい。

(例4)

(データI) 平均値: 70、中央値: 70、範囲: 27

59	61	62	62	63	64	65	65	69	
71	71	72	73	74	75	81	81	84	86

$$Q_1 = 62.5, Q_2 = 70, Q_3 = 74.5$$

$$Q_3 - Q_1 = 12$$

(データII) 平均値: 70、中央値: 70、範囲: 49

44	45	50	51	55	55	59	60	69	69	
71	71	76	80	84	85	86	89	89	90	93

$$Q_1 = 55, Q_2 = 70, Q_3 = 85.5$$

$$Q_3 - Q_1 = 30.5$$

(データIII) 平均値: 70、中央値: 70、範囲: 54

40	56	60	61	67	67	68	69	69	69
71	73	74	74	76	76	77	77	80	94

$$Q_1 = 67, Q_2 = 70, Q_3 = 76$$

$$Q_3 - Q_1 = 9$$

散らばりの度合い(中央50%のデータで比較)

III < I < II

範囲: データの最大値と最小値の差 (データの散らばりの度合いを表す量)

※ 極端に離れた値があると、範囲は大きく変わってしまう。

(例1) データI: $86 - 59 = 27$ (点) データII: $93 - 44 = 49$ (点) データIII: $94 - 40 = 54$ (点)

散らばりの度合い
I < II < III
本当にそう言える?

第1四分位数 Q_1 … 下位のデータの中央値

第2四分位数 Q_2 … 中央値と等しい

第3四分位数 Q_3 … 上位のデータの中央値

(例2) $\underbrace{\text{下位のデータ}}_{3 \ 3 \ 4} \ \underbrace{\text{上位のデータ}}_{5 \ 6 \ 7 \ 7 \ 8 \ 9}$

$$Q_1 = \frac{3+4}{2} = 3.5$$

$$Q_2 = 6$$

$$Q_3 = \frac{7+8}{2} = 7.5$$

(例3) $\underbrace{\text{下位のデータ}}_{3 \ 3 \ 4} \ \underbrace{\text{上位のデータ}}_{5 \ 6 \ 7 \ 7 \ 8 \ 9 \ 10}$

$$Q_1 = 4$$

$$Q_2 = \frac{6+7}{2} = 6.5$$

$$Q_3 = 8$$