

数学 資料の散らばりぐあい ()年()組()番()

しぶんいすう

四分位数

データを代表する値が同じであっても、資料の分布が違ふことが多い。

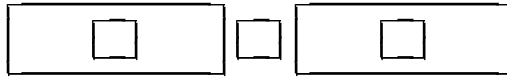
そのデータの散らばりぐあいを表す方法に()がある。

データを昇順(小さい順)に並べたとき、25%の値を第1四分位数 Q_1 、50%の値(中央値)を第2四分位数 Q_2 、75%の値を第3四分位数 Q_3 という。

データが奇数個($2m - 1$)の場合、 m 番目が(第四分位数)になる。

1番目から $m - 1$ 番目の中央値が(第四分位数)になる。

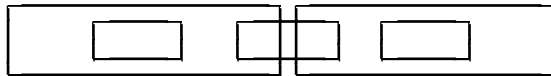
$m + 1$ 番目から $2m - 1$ 番目の中央値が(第四分位数)になる。



データが偶数個($2m$)の場合、 m 番目と $m + 1$ 番目の平均が(第四分位数)になる。

1番目から m 番目の中央値が(第四分位数)になる。

$m + 1$ 番目から $2m$ 番目の中央値が(第四分位数)になる。



問題 A 読書時間を調べたとき、次の表になった。四分位数を求めよ。

A グループ

No.	1	2	3	4	5	6	7	8	9	10
ぶん分	5	6	8	10	10	15	15	20	25	30

第1四分位数

第2四分位数

第3四分位数

B グループ

No.	1	2	3	4	5	6	7	8	9
ぶん分	0	0	10	10	15	20	20	25	25

第1四分位数

第2四分位数

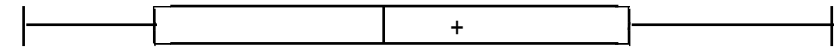
第3四分位数

箱ひげ図

データの最小値、第1四分位数、第2四分位数、第3四分位数、最大値を5数要約と

いい、この値を用いて()でデータの分布をわかりやすく表す。

最小値 第1四分位数 第2四分位数 第3四分位数 最大値

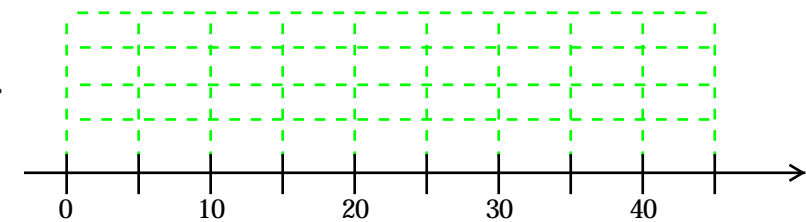


平均値(ないこともある)

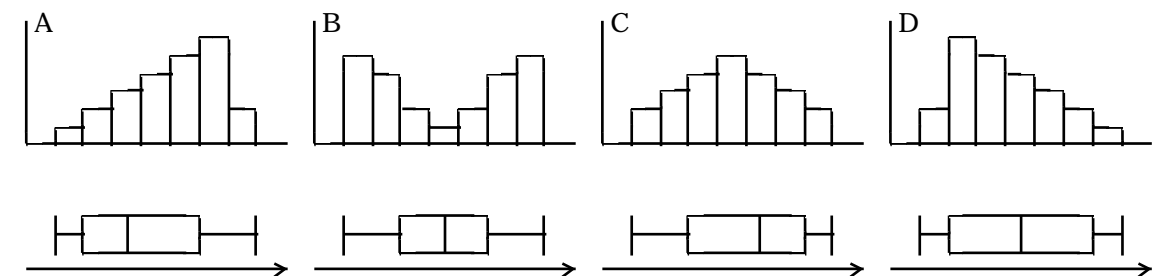
問題 B 問題 A のデータの箱ひげ図を作みなさい。

A グループ

B グループ



問題 C 次のヒストグラム(A ~ D)に対応する箱ひげ図を選びなさい。



しぶんいはんい

四分位範囲

データの最大値(max) - 最小値(min)をデータの分布の範囲(range)という。

また、第3四分位数と第1四分位数との差を四分位範囲といい、四分位範囲 $\div 2$ を四分位偏差という。

問題 D 問題 A の A グループの範囲、四分位範囲、四分位偏差を求めよ。

数学

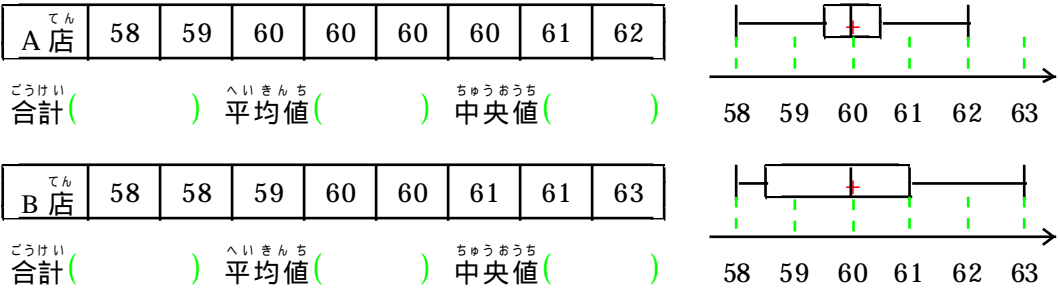
分散

()年()組()番()

分散

四分位数を利用して箱ひげ図を作るとデータの散らばりぐあいを表すことができる。
平均値をもとにした()によっても散らばりぐあいを表すことができる。

次の表はA, B の2 店での販売しているたまごの重さ(g)である。



A 店とB 店を比べると、(店の散らばりが大きい)。この散らばりを数値で表す。
(データの値) - (平均値) を()という。偏差の2 乗の平均値を()という。
また、分散の正の平方根を標準偏差といい、s で表す。

問題 A 次のデータ(平均 60)の偏差、分散(偏差の2 乗の平均)を求めよ。

A 店				B 店			
データ	2 乗	偏差	(偏差) ²	データ	2 乗	偏差	(偏差) ²
58	3364			58	3364		
59	3481			58	3364		
60	3600			59	3481		
60	3600			60	3600		
60	3600			60	3600		
60	3600			61	3721		
61	3721			61	3721		
62	3844			63	3969		
合計	28810			合計	28820		
平均	3601.25			平均	3602.5		

A, B とも偏差の合計が()になるので、偏差の平均値も()になる。

偏差の絶対値の平均では、計算が難しいので、偏差の2 乗を使っている。

分散の計算式

n 番目のデータを X_n とすると、n 番目のまでのデータの和 $X_1 + X_2 + X_3 + \cdots + X_n$ を和を表す記号 を用いて、 $\sum_{i=1}^n X_i$ と表すことができる。

n 番目までのデータの平均 \bar{X} は $\frac{1}{n} \sum_{i=1}^n X_i$ になる。

n 番目までのデータの分散 s^2 は $\frac{1}{n} \sum_{i=1}^n (X_i - \bar{X})^2$ になる。

この定義式では、データを追加したとき、分散の計算をやり直す必要がある。

データの分散 s^2 の定義式を変形する。

$$\begin{aligned} s^2 &= \frac{1}{n} \sum_{i=1}^n (X_i - \bar{X})^2 = \frac{1}{n} \sum_{i=1}^n (X_i^2 - 2 X_i \bar{X} + \bar{X}^2) \\ &= \frac{1}{n} \sum_{i=1}^n X_i^2 - 2 \bar{X} \frac{1}{n} \sum_{i=1}^n X_i + \bar{X}^2 \frac{1}{n} \sum_{i=1}^n 1 \\ &= \frac{1}{n} \sum_{i=1}^n X_i^2 - 2 \bar{X} \bar{X} + \bar{X}^2 = \left(\frac{1}{n} \sum_{i=1}^n X_i^2 - \right. \end{aligned}$$

この式を分散 s^2 の計算式という。

問題 B 次のデータ(平均 3)の分散と標準偏差を求めよ。

(偏差)² × 度数の平均 または(階級値)² × 度数の平均 - (資料の平均)²

(1)

かいきゅうち 階級値	どうすう 度数	
0	6	
1	6	
2	6	
3	6	
4	6	
5	6	
6	6	
ごうけい 合計	42	

(2)

かいきゅうち 階級値	どうすう 度数	
1	14	
2	7	
3	6	
4	5	
5	4	
6	3	
7	2	
8	1	
ごうけい 合計	42	