

<データの分析>

16人の試験の結果が以下のデータである。

93 55 73 96 71 74 77 62 91 78 86 72 68 75 70 83

<基本知識>

最大値 最小値 平均値 (説明しなくていいね?)

求める時はデータを並び替える。

番目	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16
点数	55	62	68	70	71	72	73	74	75	77	78	83	86	91	93	96

最小値は55 最大値は96

平均値は

$$(93+55+73+96+71+74+77+62+91+78+86+72+68+75+70+83) \div 16 = 76.5$$

<仮の平均による求め方>

データを見て平均がいくつぐらいになるかを予想する。

仮に80点と設定しよう。元のデータに対し、80との差をとる。

元の点数	93	55	73	96	71	74	77	62	91	78	86	72	68	75	70	83
80との差	+13	-25	-7	+16	-9	-6	-3	-18	+11	-2	+6	-8	-12	-5	-10	+3

80との差の平均値は

$$(+13-25-7+16-9-6-3-18+11-2+6-8-12-5-10+3) \div 16 = -56 \div 16 = -3.5$$

同じ結果になったね

仮の平均80に-3.5 を足すと76.5

<データの分析>

16人の試験の結果が以下のデータである。

93 55 73 96 71 74 77 62 91 78 86 72 68 75 70 83

・度数分布表

データを複数の区間に分割し、それぞれの区間に存在するデータの個数をまとめた表

<用語>

階級 … データを分けたグループ(〇〇以上□□未満)

階級値 … 各階級の真ん中の値

度数 … 各階級に含まれるデータの個数

相対度数 … 各階級の度数をデータの総数で割った値

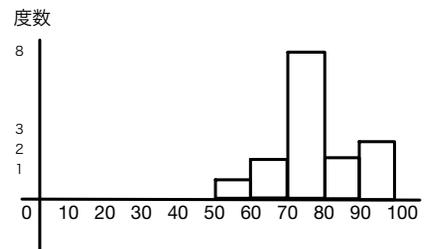
累積相対度数 … その階級以下の相対度数の合計

並べた結果をもとに

番目	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16
点数	55	62	68	70	71	72	73	74	75	77	78	83	86	91	93	96

階級	階級値	度数	相対度数	累積相対度数
50以上60未満	55	1	0.0625	0.0625
60以上70未満	65	2	0.125	0.1875
70以上80未満	75	8	0.5	0.6875
80以上90未満	85	2	0.125	0.8125
90以上100未満	95	3	0.1875	1
合計		16	1	

これをグラフにしたのがヒストグラム
学校のテストとかでもよく見るやつ



<データの分析>

16人の試験の結果が以下のデータである。

93 55 73 96 71 74 77 62 91 78 86 72 68 75 70 83

階級値はいつ使うの??

度数分布表における最頻値や平均値を聞かれた時に使う!

上のデータは一旦忘れて、下の度数分布表における

(1)最頻値、(2)平均値を求めなさいと言われたら?

階級	階級値	度数	相対度数	累積相対度数
50以上60未満	55	1	0.0625	0.0625
60以上70未満	65	2	0.125	0.1875
70以上80未満	75	8	0.5	0.6875
80以上90未満	85	2	0.125	0.8125
90以上100未満	95	3	0.1875	1
合計		16	1	

(1) 最頻値(モード) … 1番多く出現している値

度数が8の70以上80未満の階級が正解ではあるが、聞かれているのは範囲ではなく値。値を答えるために70以上80未満の代表となる値の階級値を使う

A. 75点

(2) 平均値

それぞれの階級に含まれる度数は分かるが、度数分布表だけではデータまでは分からない。そこで、その階級に含まれるデータを全て階級値に置き換えちゃおうという考え。

$$\begin{aligned}(55 \times 1 + 65 \times 2 + 75 \times 8 + 85 \times 2 + 95 \times 3) \div 16 \\ = 1240 \div 16 \\ = 77.5\end{aligned}$$

さっき平均値は76.5だったけど…?

これで大丈夫! 実際のデータの平均値と度数分布表での平均値は違うことがほとんど
それでも大体似たような値になっているのは階級値のおかげ

A. 77.5点

<データの分析>

16人の試験の結果が以下のデータである。

93 55 73 96 71 74 77 62 91 78 86 72 68 75 70 83

・箱ひげ図

データのばらつきを分かりやすく表現する図

<用語>

- 範囲 … (データの最大値) - (データの最小値)
- 四分位数 … データを4つに区切る値の総称
 - 中央値 … データを半分に区切る値(50%) (第二四分位数と呼ばれることも)
 - 第一四分位数 … 前半のデータを半分に区切る値(25%)
 - 第三四分位数 … 後半のデータを半分に区切る値(75%)
- 四分位範囲 … (第三四分位数) - (第一四分位数)

並べた結果をもとに

