

タテプリ データの分析・B問題 (改訂版)

1

次のデータは、A市とB市のある年の月ごとの平均湿度を並べたものである。

A市 58, 62, 62, 58, 60, 67, 70, 63, 61, 64, 59, 58

B市 41, 60, 61, 62, 60, 67, 70, 67, 68, 68, 56, 50 (単位は%)

- (1) A市, B市のデータの四分位範囲, 四分位偏差を求めよ。
- (2) A市, B市のデータについて, 四分位範囲によってデータの散らばりの度合いを比較せよ。

2

次のデータは, ある店舗で1日あたりに売れた缶コーヒーの本数である。ただし, a の値は0以上の整数である。

109, 102, 91, 97, 89, 117, 105, 111, 100, a (単位は本)

- (1) 1日あたりの本数の平均値が103本であるとき, a の値を求めよ。
- (2) a の値がわからないとき, 1日あたりの本数の中央値として何通りの値が考えられるか答えよ。

3

次のデータは, ある遊園地の「迷路」に挑戦した8人の高校生について何分で抜け出すことができたかを調べたものである。

7, 16, 11, 8, 12, 15, 10, 9 (単位は分)

- (1) このデータの平均値を求めよ。
- (2) データには一部に記録ミスがあり, 正しくは16分が15分, 11分が8分, 9分が13分であった。この誤りを修正したとき, データの平均値, 中央値, 分散について, 修正前より増加するか, 減少するか, または修正前と一致するかをそれぞれ答えよ。

4

20 個の値からなるデータがあり，そのうちの 9 個の値の平均値は 22，分散は 16，残り 11 個の値の平均値は 2，分散は 12 である。

- (1) このデータの平均値を求めよ。 (2) このデータの分散を求めよ。

5

40 人の生徒に行った 2 科目の試験の得点を x ， y とすると，次のようであった。

	満点	最高点	最低点	平均点	標準偏差
x	40	38	10	25	4.5
y	25	23	5	18	2.0

以下，小数の形で解答する場合，指定された桁数の一つ下の桁を四捨五入し，解答せよ。
途中で割り切れた場合は，指定された桁まで 0 を記入すること。

どちらの試験も，満点を 100 点，最低点を 40 点に揃えるように，得点を 1 次式 $x' = 2x + 20$ ， $y' = 3y + 25$ で変換した。

このとき， x' の平均点は 点， x' の標準偏差は . 点となる。

また， x と y の共分散が 7.65 のとき， x' と y' の共分散は .， x' と y' の相関係数は 0. となる。

1

解答 (1) 四分位範囲, 四分位偏差の順に

A市 5%, 2.5% B市 9.5%, 4.75%

(2) B市の方がデータの中央値の周りの散らばりが大きい

2

解答 (1) $a = 109$ (2) 6通り

3

解答 (1) 11分 (2) (平均値) 一致する (中央値) 増加する (分散) 一致する

4

解答 (1) 11 (2) 112.8

5

解答 (アイ) 70 (ウ).(エ) 9.0 (オカ).(キ) 45.9 0.(クケ) 0.85

1

(1) A市, B市のデータを値の順に並べると

A市: 58, 58, 58, 59, 60, 61, 62, 62, 63, 64, 67, 70

B市: 41, 50, 56, 60, 60, 61, 62, 67, 67, 68, 68, 70

A市のデータの四分位数について $Q_1 = \frac{58+59}{2} = 58.5$, $Q_3 = \frac{63+64}{2} = 63.5$

ゆえに, 四分位範囲は $63.5 - 58.5 = 5$ (%), 四分位偏差は $\frac{5}{2} = 2.5$ (%)

B市のデータの四分位数について $Q_1 = \frac{56+60}{2} = 58$, $Q_3 = \frac{67+68}{2} = 67.5$

ゆえに, 四分位範囲は $67.5 - 58 = 9.5$ (%), 四分位偏差は $\frac{9.5}{2} = 4.75$ (%)

(2) 四分位範囲は B市の方が大きいから, B市の方がデータの中央値の周りの散らばりが大きい。

2

a 以外の本数を小さい順に並べると

89, 91, 97, 100, 102, 105, 109, 111, 117

(1) 本数の平均値が 103 本であるから

$$\frac{1}{10}(89 + 91 + 97 + 100 + 102 + 105 + 109 + 111 + 117 + a) = 103$$

これを解いて $a = 109$

(2) データの大きさは 10 日分であるから, 小さい方から 5 番目と 6 番目の本数の平均値が中央値となる。

[1] $a \leq 100$ のとき

5 番目の本数は 100 本, 6 番目の本数は 102 本であるから, 中央値は

$$\frac{100 + 102}{2} = 101 \text{ (本)}$$

[2] $a \geq 105$ のとき

5 番目の本数は 102 本, 6 番目の本数は 105 本であるから, 中央値は

$$\frac{102 + 105}{2} = 103.5 \text{ (本)}$$

[3] $101 \leq a \leq 104$ のとき

5 番目, 6 番目の本数は a 本か 102 本のいずれかであり, 中央値は a の値によってすべて異なる。

a の値は, 101, 102, 103, 104 (本) の 4 通りあるから, 中央値も同様に 4 通りある。

また, これらの中央値は [1], [2] のものと異なる。

以上から, 中央値は 6 通りの値がありうる。

3

(1) $\frac{1}{8}(7+16+11+8+12+15+10+9) = \frac{1}{8} \times 88 = 11$ (分)

(2) [1] 平均値

修正によって、1分減り、3分減り、4分増えるから、全体の総和は変わらない。
よって、平均値は修正前と一致する。

[2] 中央値

修正前の小さい方から4, 5番目の値はそれぞれ10, 11であり、修正後の小さい方から4, 5番目の値はそれぞれ10, 12となる。
よって、中央値は修正前より増加する。

[3] 分散

[1]より、修正の前後で平均値は変わらず11であるから、修正した3つのデータの偏差の2乗の和について考える。

修正前は $(16-11)^2 + (11-11)^2 + (9-11)^2 = 29$,

修正後は $(15-11)^2 + (8-11)^2 + (13-11)^2 = 29$

となる。

よって、偏差の2乗の総和は変化しないから分散は修正前と一致する。

4

(1) $\frac{22 \times 9 + 2 \times 11}{20} = 11$

(2) 9個の値の2乗の和を a とすると $\frac{1}{9}a - 2^2 = 16$

よって $a = 4500$

残りの11個の値の2乗の和を b とすると $\frac{1}{11}b - 2^2 = 12$

よって $b = 176$

ゆえに、20個の値の2乗の和は $a + b = 4676$

したがって、20個の値の分散は

$$\frac{4676}{20} - 11^2 = 233.8 - 121 = 112.8$$

5

x' の平均点は $2 \times 25 + 20 = 70$

x' の標準偏差は $2 \times 4.5 = 9.0$

また、 x と y の共分散が 7.65 のとき、

x' と y' の共分散は $2 \times 3 \times 7.65 = 45.9$

y' の標準偏差は $3 \times 2.0 = 6.0$

よって、 x' と y' の相関係数は $\frac{45.9}{9.0 \times 6.0} = 0.85$