

以下の問題を解答するにあたっては、与えられたデータに対して、次の値を外れ値とする。

「(第1四分位数)  $- 1.5 \times$  (四分位範囲)」以下の値

「(第3四分位数)  $+ 1.5 \times$  (四分位範囲)」以上の値

水泳部に所属する太郎さんは、1500m 自由形におけるペース配分を考えるために、2021 年に開催された東京オリンピックの男子 1500m 自由形に関するデータを分析することにした。なお、自由形とは、どのような泳ぎ方で泳いでもよい競技のことである。

分析で用いるデータは、28 人の選手における、予選で計測された記録 (以下、タイム) とする。ここでは、タイムは秒単位で表すものとする。例えば、15 分 23 秒 46 であれば、 $60 \times 15 + 23.46 = 923.46$  (秒) である。そして、公式順位 (以下、順位) は、タイムの値が小さい方が上位となる。また、28 人の選手それぞれのタイムについて、スタートから 750m までのタイムを  $T_{\text{前}}$  とし、750m からゴールまでのタイムを  $T_{\text{後}}$  とする。さらに、 $T_{\text{前}}$  と  $T_{\text{後}}$  の平均値を  $T_{\text{前後}}$  とする。

なお、以下の図や表については、World Aquatics の Web ページをもとに作成している。

- (1) 太郎さんは、 $T_{\text{前}}$ 、 $T_{\text{後}}$ 、 $T_{\text{前後}}$  の関係を調べることにした。図 1 は  $T_{\text{前}}$  と  $T_{\text{後}}$  の散布図、図 2 は  $T_{\text{前}}$  と  $T_{\text{前後}}$  の散布図である。なお、これらの散布図には、完全に重なっている点はない。また、図 1 と図 2 において、A を付している点は、同じ選手であることを表している。

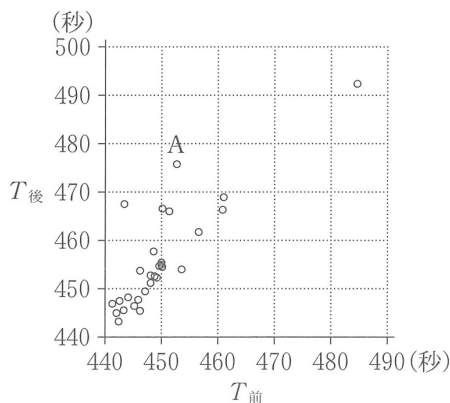


図 1  $T_{\text{前}}$  と  $T_{\text{後}}$  の散布図

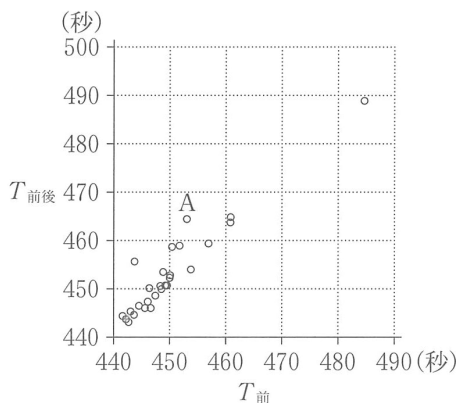


図 2  $T_{\text{前}}$  と  $T_{\text{前後}}$  の散布図

次の (a), (b) は、図 1 と図 2 に関する記述である。

- (a)  $T_{\text{前}}$  が 470 秒未満である選手について、 $T_{\text{後}}$  が 460 秒以上である選手の人数と、 $T_{\text{前後}}$  が 460 秒以上である選手の人数は等しい。
- (b) A を付している点が表す選手について、 $T_{\text{前}}$  の値は  $T_{\text{前後}}$  の値より小さく、かつ  $T_{\text{後}}$  の値は  $T_{\text{前後}}$  の値より大きい。

(a), (b) の正誤の組合せとして正しいものは ソ である。

ソ の解答群

	①	②	③	④
(a)	正	正	誤	誤
(b)	正	誤	正	誤

- (2) 太郎さんは、 $T_{\text{前}}$  と  $T_{\text{前後}}$  の相関係数を計算するために、表 1 のように、平均値、標準偏差および共分散を求めた。

表 1  $T_{\text{前}}$  と  $T_{\text{前後}}$  の平均値、標準偏差、共分散

	平均値	標準偏差	共分散
$T_{\text{前}}$	450	8.3	72.9
$T_{\text{前後}}$	453	9.3	

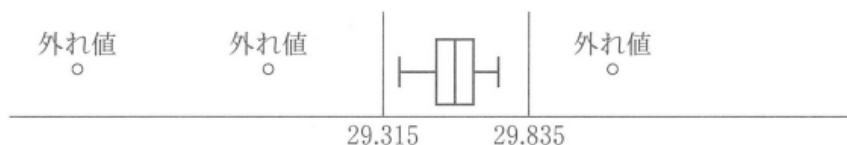
表 1 を用いると、 $T_{\text{前}}$  と  $T_{\text{前後}}$  の相関係数は タ である。

タ については、最も適当なものを、次の ①～⑨ のうちから一つ選べ。

① 0.01	② 0.24	③ 0.47	④ 0.59	⑤ 0.72
⑥ 0.83	⑦ 0.94	⑧ 1.06	⑨ 1.38	⑩ 4.14

- (3) 太郎さんは、順位とペース配分の関係調べるために、前半と後半という二分割だけではなく、より細かく分割されたタイムを用いて分析することにした。1500m 自由形のタイムは、スタートから 50m までのタイム、50m から 100m までのタイムのように、ゴールまで 50m ごとの 30 個に分けて計測されている。そこで、これら 30 個のタイムを用いて分析する。

- (i) 1 位の選手の 30 個のタイムについて考えると、外れ値かどうかを判断する二つの値である 29.315 と 29.835 が算出され、29.315 以下の 2 個のタイムと 29.835 以上の 1 個のタイムが外れ値と判断された。このとき、1 位の選手の 30 個のタイムの四分位範囲は 0. チツ 秒である。



参考図

- (ii) 太郎さんは 28 人の選手それぞれについて、30 個のタイムを用いて、選手ごとの箱ひげ図を作成し、分散を計算した。図 3 は上から分散が小さい順になるように、28 人の選手それぞれの箱ひげ図を並べたものであり、30 個のタイムにおける外れ値は、白丸で示されている。なお、分散が等しい選手はいなかった。

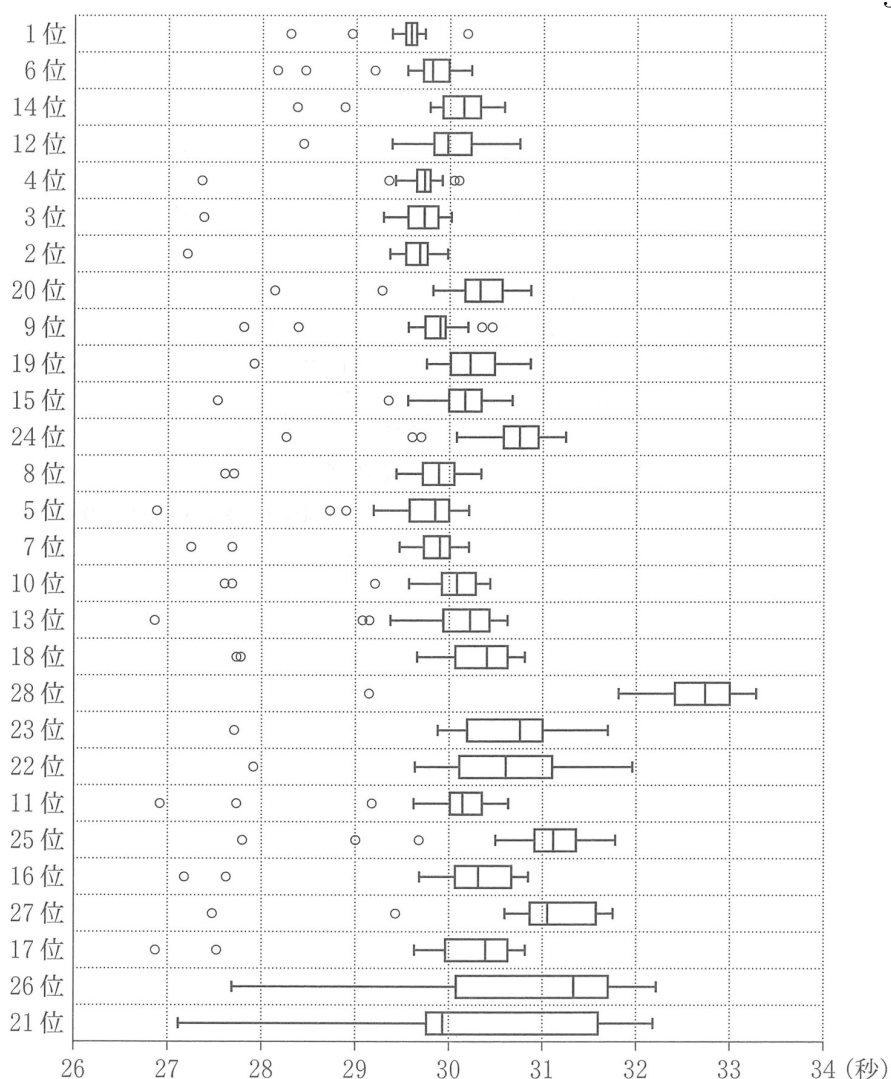


図3 28人の選手の順位と30個のタイムの箱ひげ図(上から分散の小さい順)

次の(a), (b)は, 図3に関する記述である.

- (a) 28人の選手において, 29秒より速いタイムはすべて外れ値である.  
 (b) 28人の選手から2人を選んだとき, 分散の大きい選手の四分位範囲は, 分散の小さい選手の四分位範囲より小さいことがある.

(a), (b)の正誤の組合せとして正しいものは テ である.

テの解答群

	①	②	③
(a)	正	誤	誤
(b)	正	正	誤

- (iii) 順位が1位から8位までの選手のグループを決勝進出グループ, 9位から28位までの選手のグループを予選敗退グループと呼ぶことにする. 決勝進出グループ

であり、かつ 30 個のタイムの分散が小さい方から 14 番目までの選手の人数を  $n$ (人) とすると、表 2 のようになる。

表 2 順位と分散の表 (単位は人)

		分散 (小さい値)		計
		1 番～14 番	15 番～28 番	
順位	決勝進出グループ	$n$	$8 - n$	8
	予選敗退グループ	$14 - n$	$6 + n$	20
計		14	14	28

このとき、図 3 から  $n = \boxed{\text{ト}}$  であることがわかる。このことから、決勝進出グループにおいて分散が小さい方から 14 番目までの選手が占める割合を  $P$ 、予選敗退グループにおいて分散が小さい方から 14 番目までの選手が占める割合を  $Q$  とすると、 $P \boxed{\text{ナ}} Q$  であることがわかる。

$\boxed{\text{ナ}}$  の解答群

① <	① =	② >
-----	-----	-----

(26 共通テスト 本試験 I・A 第 2 問 [2]，I 第 4 問 [1])

【答】

ソ	タ	チツ	テ	ト	ナ
2	6	13	2	7	2

【解答】

- (1) (a)  $T_{\text{前}}$  が 470 秒未満である選手について

$T_{\text{後}}$  が 460 秒以上である選手の人数  $N_{\text{後}}$  は図 1 より  $N_{\text{後}} = 7$ (人)，

$T_{\text{前後}}$  が 460 秒以上である選手の人数  $N_{\text{前後}}$  は図 2 より  $N_{\text{前後}} = 3$ (人)

であるから

$N_{\text{後}} > N_{\text{前後}}$  であり、(a) の記述は 誤り である。

- (b) A を付している点が表す選手について

$T_{\text{前}}$  の値は約 452(秒)， $T_{\text{前後}}$  の値は約 465(秒)， $T_{\text{後}}$  の値は約 476(秒)

であるから

$T_{\text{前}} < T_{\text{前後}}$  かつ  $T_{\text{後}} > T_{\text{前後}}$  であり、(b) の記述は正しい。

よって、(a)、(b) の正誤の組合せとして正しいものは ② である。

……(答)

- (2)  $T_{\text{前}}$  と  $T_{\text{前後}}$  の相関係数  $r$  は

$$r = \frac{(\text{共分散})}{(T_{\text{前}} \text{の標準偏差})(T_{\text{前後}} \text{の標準偏差})} = \frac{72.9}{8.3 \times 9.3} = 0.944 \dots \quad \text{⑥} \quad \dots\dots(\text{答})$$

である。

- (3) (i) 第 1 四分位数, 第 3 四分位数をそれぞれ  $Q_1$ ,  $Q_3$  とおくと, 四分位範囲は  $Q_3 - Q_1$  である. 外れ値かどうかを判断する二つの値が 29.315 と 29.835 であるから

$$\begin{cases} Q_1 - 1.5(Q_3 - Q_1) = 29.315 \\ Q_3 + 1.5(Q_3 - Q_1) = 29.835 \end{cases}$$

が成り立つ. 辺々ひくと

$$Q_3 - Q_1 + 3.0(Q_3 - Q_1) = 0.520$$

であり, 四分位範囲  $Q_3 - Q_1$  は

$$Q_3 - Q_1 = \frac{0.520}{4} = 0.13 \text{ (秒)} \quad \dots\dots(\text{答})$$

である.

- (ii) (i) 21 位の人の最速タイムは約 27.1 秒であるが外れ値ではないので

(a) の記述は誤りである.

- (ii) 6 位と 4 位の選手を選んだとき, 分散の大きい 4 位の選手の四分位範囲は分散の小さい 6 位の選手の四分位範囲より小さいから

(b) の記述は正しい.

よって, (a), (b) の正誤の組合せとして正しいものは ② である. \dots\dots(\text{答})

- (iii) 図 3 において上から 14 番目までの中に 1 位~8 位の選手は 7 人いるから

$$n = 7 \quad \dots\dots(\text{答})$$

である. このことから,

決勝進出グループにおいて分散が小さい方から 14 番目までの選手が占める割合  $P$  は

$$P = \frac{7}{8},$$

予選敗退グループにおいて分散が小さい方から 14 番目までの選手が占める割合  $Q$  は

$$Q = \frac{14-7}{20} = \frac{7}{20}$$

であり

$$P > Q \quad \text{②} \quad \dots\dots(\text{答})$$

であることがわかる.