

2017 年度センター試験 数学 IA 解説

第 2 問

[2] ある陸上競技大会に出場した選手の身長（単位は cm）と体重（単位は kg）のデータが得られた。選手を男子短距離、男子長距離、女子短距離、女子長距離の四つのグループに分けると、それぞれのグループの選手数は以下の通りである。

種目	男子短距離	男子長距離	女子短距離	女子長距離
選手数	328	271	319	263

(1) 上で分けた四つのグループにおける身長のヒストグラムおよび箱ひげ図から読み取れる内容を選ぶ（ヒストグラムと箱ひげ図は省略）

① 四つのグループのうちで範囲が最も大きいのは、女子短距離グループである。

⇒箱ひげから四つのグループのうちで範囲が最も大きいグループを調べると、女子短距離グループは約 42 cm であるのに対し、男子長距離グループは約 50 cm であり、このグループが最も範囲が大きい。よってこの選択肢の内容は正しくない。

① 四つのグループのすべてにおいて、四分位範囲は 12 未満である。

⇒箱ひげ図において、四分位範囲は 12 未満であれば箱の左右の幅は 6 目盛り未満になる。四つのグループいずれも箱の左右の幅は 6 目盛り未満であるため、この選択肢の内容は正しい。

② 男子長距離グループのヒストグラムでは、度数最大の階級に中央値が入っている。

⇒男子長距離グループのヒストグラムでは、度数最大の階級は 170 以上 175 未満である。一方箱ひげ図では、男子長距離グループの中央値は 176 であるため、度数最大の階級に含まれていない。よってこの選択肢の内容は正しくない。

③ 女子長距離グループのヒストグラムでは、度数最大の階級に第 1 四分位

数が入っている。

⇒女子長距離グループのヒストグラムでは、度数最大の階級は 165 以上 170 未満である。一方箱ひげ図では、女子長距離グループの第 1 四分位数は約 161 であるため、度数最大の階級に含まれていない。よってこの選択肢の内容は正しくない。

- ④ すべての選手の中で最も身長の高い選手は、男子長距離グループの中にいる。

⇒すべての選手の中で最も身長の高い選手は男子短距離グループにいる 210 cmの選手である。よってこの選択肢の内容は正しくない。

- ⑤ すべての選手の中で最も身長の低い選手は、女子長距離グループの中にいる。

⇒すべての選手の中で最も身長の低い選手は女子短距離グループにいる約 145 cmの選手である。よってこの選択肢の内容は正しくない。

- ⑥ 男子短距離グループの中央値と男子長距離グループの第 3 四分位数は、ともに 180 以上 182 未満である。

⇒男子短距離グループの中央値と男子長距離グループの第 3 四分位数はいずれも約 181 cmであるため、180 以上 182 未満の範囲にある。よってこの選択肢の内容は正しい。

よって、2 つの図から読み取れる内容は ①、⑥ である。

(2) 身長を H 、体重を W とし、

$$X = \left(\frac{H}{100} \right)^2, \quad Z = \frac{W}{X}$$

で定義する。四つのグループにおける X と Z のデータの散布図、および Z の箱ひげ図から読み取れる内容を選ぶ（散布図と箱ひげ図は省略）

X と Z のデータの散布図には原点を通り、傾きが 15, 20, 25, 30 である四つの直線 l_1, l_2, l_3, l_4 も補助的に描いている。また、箱ひげ図は (a), (b), (c), (d) の四つあり、これらは四つのグループのいずれかの箱ひげ図に対応している。

まず、箱ひげ図 (a), (b), (c), (d) がどのグループのデータであるかを考える。 Z という値は散布図の点と原点を結ぶ直線の傾きを表す。たとえば散布図で記されている点が 2 直線 l_1, l_2 の間にあるとき、その点の Z の値は 15 以上 20 以下である。

男子短距離グループの散布図には、直線 l_4 に近い位置に点があるため、箱ひげ図は 30 まで範囲がある (a) がこのグループのものである。また女子長距離グループの散布図には、直線 l_3 より上の範囲に点がないため、箱ひげ図は最大値が 25 以下である (d) がこのグループのものである。女子短距離グループの散布図には、直線 l_1 に近い位置に点がある。よって、残りの箱ひげ図のうちでは、15 に近い範囲までである (b) がこのグループのものである。残りの (c) は男子長距離グループの箱ひげ図である。

⑥ 四つのグループのすべてにおいて、 X と W に負の相関がある。

⇒ 負の相関がみられるとき、散布図の点が右下がりの方向に近いが、この問題の散布図ではいずれも右下がり方向に近いものがない。よってこの選択肢の内容は正しくない。

⑦ 四つのグループのうちで Z の中央値が一番大きいのは、男子長距離グループである。

⇒ 箱ひげ図から Z の中央値が一番大きいのは (a) にあたる男子短距離グループである。よってこの選択肢の内容は正しくない。

⑧ 四つのグループのうちで Z の範囲が最小なのは、男子長距離グループである。

⇒ 箱ひげ図から Z の範囲が最小なのは (d) にあたる女子長距離グループである。よってこの選択肢の内容は正しくない。

⑨ 四つのグループのうちで Z の四分位範囲が最小なのは、男子短距離グループである。

⇒ 男子短距離グループの箱ひげ図は (a) である。この図から四分位範囲は 2 以上ある。一方、箱ひげ図 (c) の四分位範囲は 2 未満である。つまり四分位範囲が最小なのは、男子短距離グループではない。よってこの選択肢の内容は正しくない。

⑩ 女子長距離グループはすべての Z の値は 25 より小さい。

⇒ 前述の通り、女子長距離グループはすべての Z の値は 25 より小さい。

よってこの選択肢の内容は正しい。

⑤ 男子長距離グループの Z の箱ひげ図は (c) である。

⇒ 前述の通り、男子長距離グループの Z の箱ひげ図は (c) である。よってこの選択肢の内容は正しい。

よって、2つの図から読み取れる内容は ④, ⑤ である。

(3) 次の表 1 は、設問 (2) で定義された X , W について、女子長距離グループの平均値、標準偏差および共分散を計算したものである。

表 1 平均値、標準偏差および共分散

X の 平均値	W の 平均値	X の 標準偏差	W の 標準偏差	X と W の 共分散
2.75	51.1	0.200	5.36	0.754

この表から女子長距離グループの X と W の相関係数を求める。相関係数は以下の式で求められる。

$$\frac{\Sigma(\bar{X} - X_i)(\bar{W} - W_i)}{\left\{\Sigma(\bar{X} - X_i)^2\right\}^{\frac{1}{2}} \left\{\Sigma(\bar{W} - W_i)^2\right\}^{\frac{1}{2}}}$$

この式を変形すると

$$\begin{aligned} & \frac{\Sigma(\bar{X} - X_i)(\bar{W} - W_i)}{\left\{\Sigma(\bar{X} - X_i)^2\right\}^{\frac{1}{2}} \left\{\Sigma(\bar{W} - W_i)^2\right\}^{\frac{1}{2}}} \\ &= \frac{1}{263} \Sigma(\bar{X} - X_i)(\bar{W} - W_i) \times \left\{\frac{1}{263} \Sigma(\bar{X} - X_i)^2\right\}^{-\frac{1}{2}} \times \left\{\frac{1}{263} \Sigma(\bar{W} - W_i)^2\right\}^{-\frac{1}{2}} \end{aligned}$$

となる。よって、相関係数は

$$X \text{ と } W \text{ の共分散} \times \frac{1}{X \text{ の標準偏差}} \times \frac{1}{W \text{ の標準偏差}}$$

によって求められる。表 1 より相関係数は

$$0.754 \times \frac{1}{0.200} \times \frac{1}{5.36} = \frac{0.754}{1.072} \cong \mathbf{0.703}$$

である。

(4) n を自然数とする。実数値のデータ $x_1, x_2, \dots, x_n$ に対して、平均値 $\bar{x}$ を

$$\bar{x} = \frac{x_1 + x_2 + \dots + x_n}{n}$$

とおくと、分散 s^2 は

$$s^2 = \frac{x_1^2 + x_2^2 + \dots + x_n^2}{n} - (\bar{x})^2$$

で計算できることが知られている。

女子長距離グループのデータについて考える。

設問 (3) の表 1 の数値により、このグループの身長データを各々 2 乗した値の平均値を求める。まず、設問 (2) で定義された X と H の関係を用いると、

$$X = \left(\frac{H}{100} \right)^2 \Rightarrow H^2 = 100^2 X$$

である。女子長距離グループの身長データを $h_1, h_2, \dots, h_{263}$ 、各 i に対して、 $x_i = (h_i/100)^2$ とおくと、身長データを各々 2 乗した値の平均値は

$$\begin{aligned} & \frac{1}{263} (h_1^2 + h_2^2 + \dots + h_{263}^2) \\ &= \frac{1}{263} (100^2 x_1 + 100^2 x_2 + \dots + 100^2 x_{263}) \end{aligned}$$

$$\begin{aligned}
 &= 100^2 \times \frac{1}{263} (x_1 + x_2 + \cdots + x_{263}) \\
 &= 100^2 \times \text{女子長距離グループの } X \text{ の平均値}
 \end{aligned}$$

となる。設問 (3) の表 1 からこの値は

$$100^2 \times 2.75 = \mathbf{27500}$$

である。

次に、このグループの身長の平均値が **165.7** のとき、このグループの身長の分散は

$$27500 - 165.7^2 = 27500 - 27456.49 = 43.51 \cong \mathbf{43.5}$$

である。

第 4 問の正解

ア・イ	ウ・エ	オ	カ	キ
1, 6	4, 5	2	7	2