

■数Ⅱ・Bは、やっぱり確率統計を取ればよかった■

<数学Ⅰ・数学A 第3問 (15)>

●総評 明らかに矛盾があるものを見つけるというのは、統計の問題にふさわしい問題形式で今後の統計の問いのスタンダードになるであろう。特に、(3)は良問だと思う。考え方、矛盾の見つけ方に慣れてほしい。

主なテクニック 21, 39

[1]

(1) 人数が40人なので、第3四分位数は、大きい方から10番目と11番目の人の平均である。大きい方から10番目の人も11番目の人も「25m以上30m未満」の階級に含まれているので、第3四分位数が含まれる階級は「25m以上30m未満」である。

よって、答えは $\boxed{④}$ [3点]

(2) 第1四分位数は、小さい方から10番目と11番目の人の平均である。小さい方から10番目の人も11番目の人も「15m以上20m未満」の階級に含まれているので、第1四分位数が含まれる階級は「15m以上20m未満」である。

第1四分位数が含まれる階級が「15m以上20m未満」でないものは、②, ③, ⑤

第3四分位数が含まれる階級は「25m以上30m未満」でないものは、①, ②, ③

なので、矛盾する箱ひげ図は、 $\boxed{①, ②, ③, ⑤}$ [4点]

(3) 確実に言えることを押さえて、それと矛盾する箱ひげ図を探そう。

(A-a) 「どの生徒の記録も下がった」

から

「最大値、最小値、第1四分位数、平均、第3四分位数がすべて下がる」

と言える。しかし、最初「15m以上20m未満」であった第1四分位数が、aでは「20m以上25m未満」に上がっているので矛盾

(C-c) 「最初に取ったデータで上位 $\frac{1}{3}$ に入る生徒の記録が伸びた」

から

「最大値、第3四分位数が上がる」

と言える。しかし、最初「45m以上50m未満」だった最大値が、cでは「40m以上45m未満」になっているので矛盾。

答えは、 $\boxed{①, ②}$ [6点]

カ、キ

[2] 1回目の標準偏差を s_1 、2回目の標準偏差を s_2 、1回目と2回目の標準偏差を s_{12} とすると、相関係数 r は、

$$r = \frac{s_{12}}{s_1 s_2} = \frac{54.3}{8.21 \times 6.98} \div \frac{54.3}{8.2 \times 7.0} = \frac{54.3}{57.4} \div 0.95 \quad \boxed{\text{㉗}} \quad [2 \text{ 点}]$$

コメント 選択肢を見ると、上から二ケタまでを要求されているので、分母の掛け算を上から2ケタで計算しよう。

<数学Ⅱ・数学B 第5問 (20)>

●総評 どうだろう、この易しさは、初年度は基本的で平易な問題が並ぶと予想はしていたが、ここまで易しいとは筆者も予想していなかった。今年度は数列が難しかったから、この本のまえがきの言葉を信じて数列の代わりに確率統計を取った受験生は大きなアドバンテージを得ただろう。来年もこの傾向が続いて、確率統計を取った受験生が大いに報われることを願っている。

主なテクニック 17, 35, 44, 48

$$\begin{aligned} (1) \quad P(W=0) &= \frac{{}_3C_3}{{}_7C_3} = \frac{\boxed{1}}{\boxed{35}} \quad [2 \text{ 点}], \quad P(W=1) = \frac{{}_4C_1 \cdot {}_3C_2}{{}_7C_3} = \frac{\boxed{12}}{35} \quad [2 \text{ 点}], \\ P(W=2) &= \frac{{}_4C_2 \cdot {}_3C_1}{{}_7C_3} = \frac{\boxed{18}}{35} \quad [2 \text{ 点}], \quad P(W=3) = \frac{{}_4C_3}{{}_7C_3} = \frac{\boxed{4}}{35} \quad [2 \text{ 点}] \\ E(W) &= 0 \cdot \frac{1}{35} + 1 \cdot \frac{12}{35} + 2 \cdot \frac{18}{35} + 3 \cdot \frac{4}{35} = \frac{60}{35} = \frac{\boxed{12}}{\boxed{7}} \quad [2 \text{ 点}] \\ E(W^2) &= 0^2 \cdot \frac{1}{35} + 1^2 \cdot \frac{12}{35} + 2^2 \cdot \frac{18}{35} + 3^2 \cdot \frac{4}{35} = \frac{12+72+36}{35} = \frac{120}{35} = \frac{24}{7} \\ V(W) &= E(W^2) - \{E(W)\}^2 = \frac{24}{7} - \left(\frac{12}{7}\right)^2 = \frac{24 \cdot 7 - 12^2}{49} = \frac{12(14-12)}{49} = \frac{\boxed{24}}{\boxed{49}} \quad [3 \text{ 点}] \end{aligned}$$

コメント ダミー変数を用いて、期待値を求めてみよう。

取り出した球を 1, 2, 3 として、

i が白球のとき $X_i=1$, i が赤球のとき $X_i=0$

となる確率変数を設定すると、 $W=X_1+X_2+X_3$ と表される。

$$E(W) = E(X_1+X_2+X_3) = E(X_1) + E(X_2) + E(X_3) = 3E(X_1) = 3 \cdot 1 \cdot \frac{4}{7} = \frac{12}{7}$$

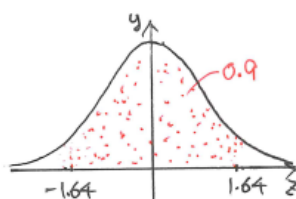
X_1, X_2, X_3 は独立でないので、同様に分散を計算しようとする、 $V(X_1X_2)$ などの値が必要になって必ずしも楽ではない。

(2) $0.99 \div 2 = 0.455$ なので、表から 0.455 に近い値を探すと、2.58 3 [2 点]

コメント なお,

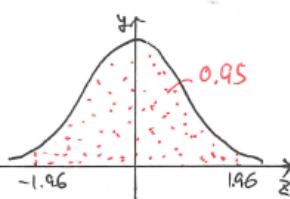
①

1.64 は両側 90% の値



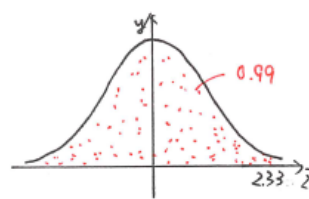
①

1.96 は両側 95% の値



②

2.33 は片側 99% の値



これくらいは問題を解いていく中で自然と覚えていただろう。

(3) 標本の平均を \bar{x} とする。

$0.95 \div 2 = 0.475$ なので、0.475 を表から探して 1.96

$P(-1.96 \leq Z \leq 1.96) = 0.95$ なので、信頼度 (信頼係数) 95% の信頼区間は,

$$[A, B] = [\bar{x} - 1.96 \frac{\sigma}{\sqrt{n}}, \bar{x} + 1.96 \frac{\sigma}{\sqrt{n}}] \quad \text{よって, } L_1 = B - A = 1.96 \frac{\sigma}{\sqrt{n}} \times 2$$

$P(-2.58 \leq Z \leq 2.58) = 0.99$ なので、信頼度 (信頼係数) 99% の信頼区間は,

$$[C, D] = [\bar{x} - 2.58 \frac{\sigma}{\sqrt{n}}, \bar{x} + 2.58 \frac{\sigma}{\sqrt{n}}] \quad \text{よって, } L_2 = D - C = 2.58 \frac{\sigma}{\sqrt{n}} \times 2$$

$$\frac{L_2}{L_1} = \frac{2.58 \frac{\sigma}{\sqrt{n}} \times 2}{1.96 \frac{\sigma}{\sqrt{n}} \times 2} = \frac{2.58}{1.96} = 1.31... = \overset{\text{チツ}}{\boxed{1.3}} \text{ [2 点]}$$

標本数が $4n$ のときの信頼度 (信頼係数) 95% の信頼区間は,

$$[E, F] = [\bar{\square} - 1.96 \frac{\square}{\sqrt{4\square}}, \bar{\square} + 1.96 \frac{\square}{\sqrt{4\square}}]$$

$$\text{よって, } L_3 = F - E = 1.96 \frac{\square}{\sqrt{4\square}} \times 2 = \frac{1}{2} \cdot 1.96 \frac{\square}{\sqrt{\square}} \times 2 = \frac{1}{2} L_1 \quad \frac{\square_3}{\square_1} = \overset{\text{テ.ト}}{\boxed{0.5}} \text{ [2 点]}$$

コメント 標本数が m 倍になると、信頼区間が $\frac{1}{\sqrt{m}}$ 倍になることは常識にしておきたい。