[2018スタンダードⅠⅡAB受 基本問題126]

生徒10人の試験結果が次のようになった。

85, 80, 25, 0, 65, 55, 40, 55, 75, 70 (点)

このデータの平均値，中央値，最頻値，分散を求めよ。

また，生徒全員に10点ずつ加点したときの平均値および分散を求めよ。

小さい順に並べると

0 , 25 , 40 , 55 , 55 | 65 , 70 , 75 , 80 , 85

中央値 $= \dfrac{55+65}{2} = 60$

最頻値 $= 55$

| $x_i$ | $x_i - \bar{x}$ | $(x_i - \bar{x})^2$ |
|---|---|---|
| 0 | $-55$ | 3025 |
| 25 | $-30$ | 900 |
| 40 | $-15$ | 225 |
| 55 | 0 | 0 |
| 55 | 0 | 0 |
| 65 | 10 | 100 |
| 70 | 15 | 225 |
| 75 | 20 | 400 |
| 80 | 25 | 625 |
| 85 | 30 | 900 |
| 550 | 0 | 6400 |

平均 $\bar{x} = \dfrac{550}{10} = 55$

分散 $= \dfrac{6400}{10} = 640$

以上より

平均値 $= 55$ , 中央値 $= 60$ , 最頻値 $= 55$ , 分散 $= 640$

10点ずつ加点したとき,

平均値 $= 65$ , 分散 $= 640$

[2018スタンダードⅠⅡAB受 基本問題127]

次のデータは，ある商店におけるＡ弁当とＢ弁当の７日間の販売数である。

Ａ弁当　22，28，16，24，33，27，21　(個)

Ｂ弁当　18，22，17，13，28，35，32　(個)

データの散らばりの度合いが大きいのは，Ａ弁当，Ｂ弁当のうちどちらか。四分位範囲に基づいて調べよ。

小さい順に並べると

A : 16 , 21 , 22 , 24 , 27 , 28 , 33

四分位範囲 $= 28 - 21 = 7$

B : 13 , 17 , 18 , 22 , 28 , 32 , 35

四分位範囲 $= 32 - 17 = 15$

よって

(Aの四分位範囲) < (Bの四分位範囲)

となるので

B弁当の方が散らばりの度合いが大きい.

[2018スタンダードⅠⅡAB受問題A395]

(1) 次のデータの四分位偏差を求めよ。

71, 72, 73, 74, 75, 76, 77, 79, 80, 93, 108, 125, 144, 165

(2) データ 2, 8, 1, 9, 4, $a$ がある。このデータの平均値が7であるような $a$ の値を求めよ。また，平均値と中央値が等しくなるような $a$ の値をすべて求めよ。

((1) 2017 藤田保健衛生大　(2) 2017 福岡大)

(1)

71, 72, 73, 74, 75, 76, 77 | 79, 80, 93, 108, 125, 144, 165

四分位範囲 $= 108 - 74 = 34$

四分位偏差 $= \frac{34}{2} = \underline{\underline{17}}$

(2)

平均値 = 7 のとき

$$\frac{2+8+1+9+4+a}{6} = 7$$

$24 + a = 42$

$\underline{\underline{a = 18}}$

平均値 = 中央値 のとき

平均値 $= \frac{24+a}{6}$

$a$ を除いて小さい順に並べると

1, 2, 4, 8, 9

中央値 は

(i) $a \leqq 2$ のとき

中央値 $= \frac{2+4}{2} = 3$

$\frac{24+a}{6} = 3$

$24 + a = 18$

$a = -6$

これは $a \leqq 2$ を満たす。

(ii) $2 \leqq a \leqq 8$ のとき

中央値 $= \frac{a+4}{2}$ より

$\frac{24+a}{6} = \frac{a+4}{2}$

$24 + a = 3a + 12$

$-2a = -12$

$a = 6$

これは $2 \leqq a \leqq 8$ を満たす。

(iii) $8 \leqq a$ のとき

中央値 $= \frac{4+8}{2} = 6$ より

$\frac{24+a}{6} = 6$

$24 + a = 36$

$a = 12$

これは $8 \leqq a$ を満たす。

(i)〜(iii) より

$\underline{\underline{a = -6,\ 6,\ 12}}$

[2018スタンダードⅠⅡAB受 問題A396]

2つの変量 $x$，$y$ に関するデータが右のように与えられていて，$y$ の平均値は4，分散は0.8 である。

(1) $x$ の平均値と分散を求めよ。

(2) $a$，$b$ の値を求めよ。ただし，$a<b$ とする。

(3) $x$ と $y$ の共分散を求めよ。

(4) $x$ と $y$ の相関係数を求めよ。

| 番号 | 1 | 2 | 3 | 4 | 5 |
|---|---|---|---|---|---|
| $x$ | 6 | 2 | 2 | 6 | 4 |
| $y$ | 5 | $a$ | $b$ | 5 | 3 |

(2017 西南学院大)

| $x$ | $y$ | $x-\bar{x}$ | $y-\bar{y}$ | $(x-\bar{x})^2$ | $(y-\bar{y})^2$ | $(x-\bar{x})(y-\bar{y})$ |
|---|---|---|---|---|---|---|
| 6 | 5 | 2 | 1 | 4 | 1 | 2 |
| 2 | $a$ | $-2$ | $a-4$ | 4 | $(a-4)^2$ | $-2(a-4)$ |
| 2 | $b$ | $-2$ | $b-4$ | 4 | $(b-4)^2$ | $-2(b-4)$ |
| 6 | 5 | 2 | 1 | 4 | 1 | 2 |
| 4 | 3 | 0 | $-1$ | 0 | 1 | 0 |
| 20 | | 0 | 0 | 16 | | |

(1)

$$\bar{x}=\frac{20}{5}=4$$

$$v_x=\frac{16}{5}=3.2$$

よって

$(x\text{の平均値})=4$

$(x\text{の分散})=3.2$

(2)

$$\begin{cases} 1+(a-4)+(b-4)+1-1=0 \\ \dfrac{1+(a-4)^2+(b-4)^2+1+1}{5}=0.8 \end{cases}$$

$$\Longleftrightarrow \begin{cases} a+b=7 & \text{―――①} \\ (a-4)^2+(b-4)^2=1 & \text{―――②} \end{cases}$$

② より

$a^2-8a+16+b^2-8b+16=1$

$a^2+b^2-8(a+b)=-31$

$(a+b)^2-2ab-8(a+b)=-31$

① を代入して

$49-2ab-56=-31$

$-2ab=-24$

$ab=12$

$a$，$b$ を解にもつ2次方程式は

$t^2-7t+12=0$

$(t-3)(t-4)=0$

$t=3,4$

$a<b$ より

$a=3,\ b=4$

(3)

$2-2(a-4)-2(b-4)+2$

$=2+2+2$

$=6$

よって

$(\text{共分散})=\dfrac{6}{5}=1.2$

(4)

$(x\text{の標準偏差})=\sqrt{3.2}=\sqrt{\dfrac{32}{10}}=\dfrac{4\sqrt{2}}{\sqrt{10}}$

$(y\text{の標準偏差})=\sqrt{0.8}=\sqrt{\dfrac{8}{10}}=\dfrac{2\sqrt{2}}{\sqrt{10}}$

より

$$(\text{相関係数})=\frac{\frac{12}{10}}{\frac{4\sqrt{2}}{\sqrt{10}}\times\frac{2\sqrt{2}}{\sqrt{10}}}$$

$$=\frac{12}{16}$$

$$=\frac{3}{4}$$

$$=0.75$$

[2018スタンダードⅠⅡAB受 問題A397]

ある集団はAとBの2つのグループで構成される。データを集計したところ，それぞれのグループの個数，平均値，分散は右の表のようになった。このとき，集団全体の平均値と分散を求めよ。

| グループ | 個数 | 平均値 | 分散 |
|---|---|---|---|
| A | 20 | 16 | 24 |
| B | 60 | 12 | 28 |

(2016 立命館大)

グループAのデータ：$x_1 \sim x_{20}$
グループBのデータ：$y_1 \sim y_{60}$
としておく。

Aの平均値 $=16$ より

$$\frac{1}{20}\sum_{i=1}^{20} x_i = 16$$

$$\sum_{i=1}^{20} x_i = 320 \quad \text{——①}$$

Aの分散 $=24$ より

$$\frac{1}{20}\sum_{i=1}^{20}(x_i-16)^2 = 24$$

$$\sum_{i=1}^{20}(x_i-16)^2 = 480$$

$$\sum_{i=1}^{20} x_i^2 - 32\sum_{i=1}^{20} x_i + 256\times 20 = 480$$

$$\sum_{i=1}^{20} x_i^2 - 32\times 320 + 256\times 20 = 480 \quad (\because ①より)$$

$$\sum_{i=1}^{20} x_i^2 = 5600 \quad \text{——②}$$

Bの平均値 $=12$ より

$$\frac{1}{60}\sum_{j=1}^{60} y_j = 12$$

$$\sum_{j=1}^{60} y_j = 720 \quad \text{——③}$$

Bの分散 $=28$ より

$$\frac{1}{60}\sum_{j=1}^{60}(y_j-12)^2 = 28$$

$$\sum_{j=1}^{60}(y_j-12)^2 = 1680$$

$$\sum_{j=1}^{60} y_j^2 - 24\sum_{j=1}^{60} y_j + 144\times 60 = 1680$$

$$\sum_{j=1}^{60} y_j^2 - 24\times 720 + 144\times 60 = 1680 \quad (\because ③より)$$

$$\sum_{j=1}^{60} y_j^2 = 10320 \quad \text{——④}$$

集団全体の平均値は ①，③ より

$$\frac{1}{80}\left(\sum_{i=1}^{20} x_i + \sum_{j=1}^{60} y_j\right)$$

$$= \frac{1}{80}(320+720)$$

$$= \frac{1040}{80}$$

$$= \underline{\underline{13}}$$

集団全体の分散は ①～④ より

$$\frac{1}{80}\left(\sum_{i=1}^{20}(x_i-13)^2 + \sum_{j=1}^{60}(y_j-13)^2\right)$$

$$= \frac{1}{80}\left(\sum_{i=1}^{20} x_i^2 - 26\sum_{i=1}^{20} x_i + 169\times 20 + \sum_{j=1}^{60} y_j^2 - 26\sum_{j=1}^{60} y_j + 169\times 60\right)$$

$$= \frac{1}{80}(5600 - 26\times 320 + 169\times 20 + 10320 - 26\times 720 + 169\times 60)$$

$$= \frac{1}{80}(15920 - 27040 + 13520)$$

$$= \frac{1}{80}\times 2400$$

$$= \underline{\underline{30}}$$

[2018スタンダードⅠⅡAB受 例題50]

10個の値からなるデータがあり，そのうちの4個の値の平均値は30，分散は16であり，残りの6個の値の平均値は40，分散は25であった。このデータの平均値と分散を求めよ。 （福島県立医科大）

10個のデータを $x_1 \sim x_{10}$ とすると

$\frac{1}{4}\sum_{i=1}^{4} x_i = 30$ より

$\sum_{i=1}^{4} x_i = 120$ ——— ①

$\frac{1}{4}\sum_{i=1}^{4} x_i^2 - \left(\frac{1}{4}\sum_{i=1}^{4} x_i\right)^2 = 16$ より

$\frac{1}{4}\sum_{i=1}^{4} x_i^2 - 900 = 16$ （∵①より）

$\frac{1}{4}\sum_{i=1}^{4} x_i^2 = 916$

$\sum_{i=1}^{4} x_i^2 = 3664$ ——— ②

$\frac{1}{6}\sum_{i=5}^{10} x_i = 40$ より

$\sum_{i=5}^{10} x_i = 240$ ——— ③

$\frac{1}{6}\sum_{i=5}^{10} x_i^2 - \left(\frac{1}{6}\sum_{i=5}^{10} x_i\right)^2 = 25$ より

$\frac{1}{6}\sum_{i=5}^{10} x_i^2 - 1600 = 25$ （∵③より）

$\frac{1}{6}\sum_{i=5}^{10} x_i^2 = 1625$

$\sum_{i=5}^{10} x_i^2 = 9750$ ——— ④

①～④より

$$\begin{aligned}(\text{平均値}) &= \frac{1}{10}\sum_{i=1}^{10} x_i \\ &= \frac{1}{10}\left(\sum_{i=1}^{4} x_i + \sum_{i=5}^{10} x_i\right) \\ &= \frac{1}{10}(120+240) \\ &= \frac{360}{10} \\ &= \underline{\underline{36}}\end{aligned}$$

$$\begin{aligned}(\text{分散}) &= \frac{1}{10}\sum_{i=1}^{10} x_i^2 - \left(\frac{1}{10}\sum_{i=1}^{10} x_i\right)^2 \\ &= \frac{1}{10}\sum_{i=1}^{4} x_i^2 + \frac{1}{10}\sum_{i=5}^{10} x_i^2 - \left(\frac{1}{10}\sum_{i=1}^{10} x_i\right)^2 \\ &= \frac{3664}{10} + \frac{9750}{10} - 1296 \\ &= 1341.4 - 1296 \\ &= \underline{\underline{45.4}}\end{aligned}$$

[2018スタンダードⅠⅡAB受 問題B398]

$n$ を2以上の自然数とする。$n$ 人の得点が $x_1=100$，$x_i=99$ $(i=2,\ 3,\ \cdots\cdots,\ n)$ であるとき，$n$ 人の得点の平均 $\overline{x}$，分散 $v$ を求めよ。また，得点 $x_i$ $(i=1,\ 2,\ 3,\ \cdots\cdots,\ n)$ の偏差値 $t_i$ が $t_i=50+\dfrac{10(x_i-\overline{x})}{\sqrt{v}}$ によって計算されることを利用して，$t_1$ が100以上となる最小の $n$ を求めよ。 (2017 福岡大)

$$\overline{x}=\frac{x_1+x_2+x_3+\cdots+x_n}{n}$$

$$=\frac{100+99(n-1)}{n}$$

$$=\frac{99n+1}{n}$$

$$v=\frac{x_1^2+x_2^2+x_3^2+\cdots+x_n^2}{n}-\overline{x}^2$$

$$=\frac{10000+99^2(n-1)}{n}-\left(\frac{99n+1}{n}\right)^2$$

$$=\frac{10000n+9801n^2-9801n-9801n^2-198n-1}{n^2}$$

$$=\frac{n-1}{n^2}$$

$t_1 \geqq 100$ より

$$50+\frac{10(x_1-\overline{x})}{\sqrt{v}}\geqq 100$$

$$\frac{10(x_1-\overline{x})}{\sqrt{v}}\geqq 50$$

$$x_1-\overline{x}\geqq 5\sqrt{v}$$

$$100-\frac{99n+1}{n}\geqq 5\sqrt{\frac{n-1}{n^2}}$$

$$100-\frac{99n+1}{n}\geqq 5\times\frac{\sqrt{n-1}}{n}$$

$$100n-99n-1\geqq 5\sqrt{n-1}$$

$$n-1\geqq 5\sqrt{n-1}$$

$\sqrt{n-1}>0$ より

$$\sqrt{n-1}\geqq 5$$

両辺ともに正なので

$$n-1\geqq 25$$

$$n\geqq 26$$

$n\in\mathbb{N}$ なので

$$\min n=26$$

[2018スタンダードⅠⅡAB受 問題B399]

2つの変量 $x$, $y$ のデータが，$n$ 個の $x$, $y$ の値の組として

$$(x_1,\ y_1),\ (x_2,\ y_2),\ \cdots\cdots,\ (x_n,\ y_n)$$

のように与えられているとする。

(1) $x$, $y$ の平均値をそれぞれ $\overline{x}$, $\overline{y}$ とするとき，変量 $x$ と $y$ の共分散 $s_{xy}$ は $s_{xy}=\dfrac{1}{n}\left(\displaystyle\sum_{k=1}^{n}x_ky_k\right)-\overline{x}\ \overline{y}$ であることを示せ。

(2) これらのデータの間には，$y_k=ax_k+b$ ($k=1,\ 2,\ \cdots\cdots,\ n$) という関係があるとする。ただし，$a$, $b$ は実数で，$a\neq0$ である。変量 $x$ の標準偏差 $s_x$ は $0$ でないとする。このとき，$x$ と $y$ の相関係数を求めよ。(2016 信州大)

(1) (証明)

$$S_{xy}=\frac{1}{n}\{(x_1-\overline{x})(y_1-\overline{y})+(x_2-\overline{x})(y_2-\overline{y})+\cdots+(x_n-\overline{x})(y_n-\overline{y})\}$$

$$=\frac{1}{n}\sum_{k=1}^{n}(x_k-\overline{x})(y_k-\overline{y})$$

$$=\frac{1}{n}\sum_{k=1}^{n}(x_ky_k-\overline{y}x_k-\overline{x}y_k+\overline{x}\overline{y})$$

$$=\frac{1}{n}\left(\sum_{k=1}^{n}x_ky_k\right)-\frac{1}{n}\overline{y}\sum_{k=1}^{n}x_k-\frac{1}{n}\overline{x}\sum_{k=1}^{n}y_k+\frac{1}{n}\sum_{k=1}^{n}\overline{x}\overline{y}$$

$$=\frac{1}{n}\left(\sum_{k=1}^{n}x_ky_k\right)-\frac{1}{n}\overline{y}\,n\overline{x}-\frac{1}{n}\overline{x}\,n\overline{y}+\frac{1}{n}\cdot n\overline{x}\overline{y}$$

$$=\frac{1}{n}\left(\sum_{k=1}^{n}x_ky_k\right)-\overline{x}\overline{y}-\overline{x}\overline{y}+\overline{x}\overline{y}$$

$$=\frac{1}{n}\left(\sum_{k=1}^{n}x_ky_k\right)-\overline{x}\overline{y}$$

■

(2)

$$S_{xy}=\frac{1}{n}\left(\sum_{k=1}^{n}x_ky_k\right)-\overline{x}\overline{y}$$

$$=\frac{1}{n}\left(\sum_{k=1}^{n}x_k(ax_k+b)\right)-\overline{x}(a\overline{x}+b)$$

$$=\frac{a}{n}\sum_{k=1}^{n}x_k^2+\frac{b}{n}\sum_{k=1}^{n}x_k-a\overline{x}^2-b\overline{x}$$

$$=a\left(\frac{1}{n}\sum_{k=1}^{n}x_k^2-\overline{x}^2\right)+b\overline{x}-b\overline{x}$$

$$=a\left(\frac{1}{n}\sum_{k=1}^{n}x_k^2-\overline{x}^2\right)$$

$$=a\,(x\text{の分散})$$

$$(y\text{の分散})=\frac{1}{n}\sum_{k=1}^{n}(y_k-\overline{y})^2$$

$$=\frac{1}{n}\sum_{k=1}^{n}(ax_k+b-a\overline{x}-b)^2$$

$$=\frac{1}{n}\sum_{k=1}^{n}a^2(x_k-\overline{x})^2$$

$$=\frac{a^2}{n}\sum_{k=1}^{n}(x_k-\overline{x})^2$$

$$=a^2\,(x\text{の分散})$$

よって，

$x$ と $y$ の相関係数 $r_{xy}$ は

$$r_{xy}=\frac{S_{xy}}{\sqrt{x\text{の分散}}\times\sqrt{y\text{の分散}}}$$

$$=\frac{a\,(x\text{の分散})}{\sqrt{x\text{の分散}}\,|a|\sqrt{x\text{の分散}}}$$

$$=\frac{a}{|a|}$$

つまり

$$r_{xy}=\begin{cases}1 & (a>0\text{のとき})\\ -1 & (a<0\text{のとき})\end{cases}$$

[2018スタンダードⅠⅡAB受 問題B400]

(1) 変量 $x$ のデータの値を $x_1$, $x_2$, ……, $x_n$ とし，$x$ の平均値を $\overline{x}$，標準偏差を $s\,(s\neq 0)$ とする。$p\geqq 1$ を満たす $p$ について，$|x_k-\overline{x}|\geqq ps$ を満たす $x_k$ の個数を $K(p)$ とするとき，$K(p)\leqq \dfrac{n}{p^2}$ が成り立つことを示せ。

(2) あるテストを受けた生徒300人について調べたところ，得点の平均値が63点，標準偏差が8点であった。得点が47点より高く79点より低い生徒は少なくとも何人いるか，(1) を利用して答えよ。

(1) (証明)

$x_1\geqq x_2\geqq\cdots\geqq x_{K(p)}>x_{K(p)+1}\geqq\cdots\geqq x_n$

としても一般性を失わない.

$$s^2=\frac{1}{n}\sum_{k=1}^{n}(x_k-\overline{x})^2$$
$$\geqq\frac{1}{n}\sum_{k=1}^{K(p)}|x_k-\overline{x}|^2$$
$$\geqq\frac{1}{n}\sum_{k=1}^{K(p)}(ps)^2$$
$$=\frac{1}{n}p^2s^2\cdot K(p)$$

つまり

$$s^2\geqq\frac{1}{n}p^2s^2K(p)$$
$$K(p)\leqq\frac{n}{p^2}$$

■

(2) $47<x_k<79$ より

$|x_k-63|<16$

ここで，(1) において，$n=300$, $s=8$ であり，
$\overline{x}=63$, $p=2$ とすると
$K(2)\leqq\dfrac{300}{2^2}=75$

求める人数 $N$ は

$N\geqq 300-75=225$

よって 少なくとも 225人 いる.