この見本について

高校数学の俯瞰 －演習セット－　数学Ⅰ A 第 4 章「データの分析」

これは**見本**です。実際の商品から **4 問**だけを抜き出して、問題と解説をそのまま載せています。文字を小さくしたり、解説を短くしたりはしていません。

選んだ 4 問

- 前半の 3 問 … **方針の候補が多い問題**や、**別解を複数紹介できる問題**を選びました。
- 最後の 1 問 … 章末の**【入試対策　実戦 Point 活用演習】**からです。入試レベルで、(1)(2)(3) の誘導が付いています。

商品の中身

- ①この教材のあつかいについて　②目次　③問題編　④答え一覧　⑤解説編　の順に入っています。
- 解説は**「問題を見て思うこと」→「解答」→「方針の候補（○△ ×）」→「別の方針でも解いてみる」→「この問題で身につくこと」→「使った Point」**の順です。
- 使った Point は**本編の該当箇所をそのまま載せている**ので、PDF 1 つで完結します。

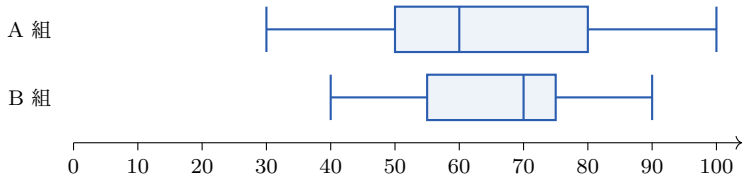
※ 見本は抜粋なので、「本文は問○参照」と書いてある問が、この見本に入っていないことがあります。商品版にはすべて収録されています。

※ 商品版には全ページに購入者のお名前と注文 ID が入ります。見本には入っていません。

※ 答え一覧は商品版にあります。見本では解説の中に答えが入っています。

問題編（見本）

問 5　A 組 40 人と B 組 40 人が同じテスト（100 点満点、得点は整数）を受けた。下の図は 2 つの組の得点の箱ひげ図である。



図：A 組・B 組の得点の箱ひげ図（点数の目盛りは 10 点きざみ）

次のア～カのうち、この図から確実にいえるものをいくつでも選べ。

ア 範囲は B 組のほうが大きい。

イ 四分位範囲は A 組のほうが大きい。

ウ 中央値以上の得点をとった生徒は、どちらの組にも 20 人以上いる。

エ A 組には 70 点以上の生徒が 10 人以上いる。

オ B 組には 80 点以上の生徒が 10 人以上いる。

カ A 組の平均点は B 組の平均点より低い。　（Point 4-0-2）

問 10　10 人の数学のテストの得点について、最初に集計したところ、平均値は 50 点、分散は 36 であった。

(1) あとで、45 点の生徒の得点が実は 55 点だったと分かった。訂正後の 10 人の平均値と分散を求めよ。

(2) 訂正前の 10 人のうち、ちょうど平均値と同じ 50 点だった生徒 1 人を除いた 9 人の平均値と分散を求めよ。　（Point 4-0-4）

問 16　x の分散は 0 でないとし、変量 y を $y = ax + b$（a, b は定数）で定める。

(1) $a = 2$, $b = 1$ のとき、s_{xy} を s_x^2 で表し、相関係数 r を求めよ。

(2) $a = -3$, $b = 5$ のとき、相関係数 r を求めよ。

(3) $a \neq 0$ のとき、一般に相関係数 r はどうなるか。また $a = 0$ のとき相関係数が定義できない理由を述べよ。　（Point 4-0-6）

【入試対策　実戦 Point 活用演習】

章末に置いている入試レベルの 6 問から、1 問を見本として載せます。

問 18　A 組 n_A 人、B 組 n_B 人が同じテストを受け、A 組の平均値・分散を m_A，s_A^2、B 組の平均値・分散を m_B，s_B^2 とする。$N = n_A + n_B$ とし、2 つの組を合わせた N 人全体の平均値を m、分散を s^2 とする。

(1) 各組の「2 乗の平均」を m_A，s_A^2 などで表すことで、

$$s^2 = \frac{n_A s_A^2 + n_B s_B^2}{N} + \frac{n_A n_B \left(m_A - m_B\right)^2}{N^2}$$

が成り立つことを示せ。

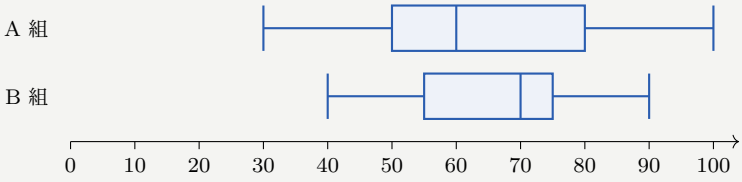
(2) $n_A = 20$，$m_A = 60$，$s_A^2 = 80$，$n_B = 30$，$m_B = 70$，$s_B^2 = 70$ のとき、m と s^2 を求めよ。

(3) B 組の人数 30 と分散 70 は変えず、B 組の平均値だけを m_B（$m_B > 60$）に取りかえる。A 組はそのままで、全体の分散がちょうど 80 になるような m_B を求めよ。　(Point 4-0-4)

解説編（見本）

問 5

A 組 40 人と B 組 40 人が同じテスト（100 点満点、得点は整数）を受けた。下の図は 2 つの組の得点の箱ひげ図である。次のア～カのうち、この図から確実にいえるものをいくつでも選べ。



ア 範囲は B 組のほうが大きい。　イ 四分位範囲は A 組のほうが大きい。
ウ 中央値以上の得点をとった生徒は、どちらの組にも 20 人以上いる。
エ A 組には 70 点以上の生徒が 10 人以上いる。
オ B 組には 80 点以上の生徒が 10 人以上いる。　カ A 組の平均点は B 組の平均点より低い。

問題を見て思うこと

箱ひげ図の 5 つの値は、データを**約 4 等分する位置**を表す。40 人なら、Q_1，Q_2，Q_3 で区切られた各区間に約 10 人ずつ。　▶Point 4-0-2

→ まず 5 数要約を読む。A 組 (30, 50, 60, 80, 100)，B 組 (40, 55, 70, 75, 90)。

→ 範囲と四分位範囲は引き算で決まる。人数に関する文は、**何番目の人がどの値以上か**に直して確かめる。

→ 図に「+」がなく、平均点は読み取れない。**確実にいえる**かどうかは、ほかの並びが成り立つ例を作れるかで判断する。

解答

読み取った 5 数要約は、A 組が (30, 50, 60, 80, 100)，B 組が (40, 55, 70, 75, 90)。40 人なので、Q_2 は 20 番目と 21 番目の平均、Q_3 は 30 番目と 31 番目の平均である。

ア　範囲は A 組が $100-30=70$、B 組が $90-40=50$。A 組のほうが大きいので**誤り**。

イ　四分位範囲は A 組が $80-50=30$、B 組が $75-55=20$。A 組のほうが大きいので**正しい**。

ウ　中央値は 20 番目と 21 番目の平均だから、21 番目は中央値以上。21 番目から 40 番目までの 20 人は中央値以上であり、どちらの組についても言えるので**正しい**。

エ　A 組の $Q_3=80$ は 30 番目と 31 番目の平均だから、31 番目は 80 点以上。31 番目から 40 番目までの 10 人は 80 点以上なので、70 点以上の生徒も 10 人以上いる。**正しい**。

オ　B 組の $Q_3=75$ から言えるのは、31 番目から 40 番目までの 10 人が 75 点以上ということまで。80 点以上が 10 人とは限らない。たとえば得点が次の表のとおりのとき、5 数要約は図と同じ (40, 55, 70, 75, 90) で、80 点以上は 1 人だけである。よって**確実にはいえない**。

得点	40	55	56	70	72	75	90	計
人数	1	10	8	2	8	10	1	40

カ　箱ひげ図に平均点（「＋」）が描かれていないので、平均点の大小は読み取れない。**確実にはいえない**。

以上より、確実にいえるのは **イ，ウ，エ**。

この問題で身につくこと

・箱ひげ図の 5 つの値は、データを**約 4 等分する位置**を表す。
・Q_3 以上が約 $\frac{1}{4}$、中央値以上が約半分、というように**割合**なら読み取れる。
・範囲・四分位範囲は**引き算で決まる**ので比べられる。
・「○点以上が何人」は、Q **の値そのもの**にあたる点数でないと決まらないことが多い。
・平均点は「＋」がなければ読み取れない。

使った Point（本編より。例題は省略）

Point 4-0-2 四分位範囲・箱ひげ図
（本文は問 3 参照）

問 10

10 人の数学のテストの得点について、最初に集計したところ、平均値は 50 点、分散は 36 であった。

(1) あとで、45 点の生徒の得点が実は 55 点だったと分かった。訂正後の 10 人の平均値と分散を求めよ。

(2) 訂正前の 10 人のうち、ちょうど平均値と同じ 50 点だった生徒 1 人を除いた 9 人の平均値と分散を求めよ。

問題を見て思うこと

データそのものは分からないが、平均値と分散から**合計と 2 乗の合計**は復元できる。 ▶Point 4-0-4

→ 合計 $= n\overline{x} = 500$。2 乗の合計 $= n\left(\overline{x}^2 + s^2\right) = 10\left(50^2 + 36\right) = 25360$（分散 $=$ 2 乗の平均 $-$ 平均の 2 乗 を逆に使う）。

→ (1) 差し替えた 1 人分だけ、合計と 2 乗の合計を直す。(2) 取り除いた 1 人分を引き、個数を 9 にする。

→ (2) は、平均値と同じ値の人の偏差が 0 であることに注目すると、偏差の 2 乗の合計 360 が変わらないとも見える。

解答

訂正前について、合計は $10 \cdot 50 = 500$、2 乗の合計は $10\left(50^2 + 36\right) = 25360$ である。

(1)　45 を 55 に直すと、合計は $500 - 45 + 55 = 510$、2 乗の合計は $25360 - 45^2 + 55^2 = 25360 - 2025 + 3025 = 26360$。よって

$$\text{平均値} = \frac{510}{10} = \mathbf{51}, \qquad \text{分散} = \frac{26360}{10} - 51^2 = 2636 - 2601 = \mathbf{35}$$

(2)　50 点の生徒を除くと、合計は $500 - 50 = 450$、2 乗の合計は $25360 - 50^2 = 22860$、個数は 9。よって

$$\text{平均値} = \frac{450}{9} = \mathbf{50}, \qquad \text{分散} = \frac{22860}{9} - 50^2 = 2540 - 2500 = \mathbf{40}$$

方針の候補

○ 平均値と分散から**合計と 2 乗の合計**に戻して、変わる分だけ直す。 ▶Point 4-0-4

○ ((2) のみ) 平均値と同じ値は偏差が 0 なので、**偏差の 2 乗の合計 360 が変わらない**と見る。

× 分散 36 のまま、個数だけ変えて答える ··· 分散は**個数で割った値**なので、個数が変われば変わる。

× 訂正で 45 が 55 に変わっただけなので分散は変わらないと考える ··· 2 乗の合計が変わるので、**分散も変わる**。

この問題で身につくこと

・平均値・分散から、**合計** $= n\bar{x}$、2 **乗の合計** $= n\left(\bar{x}^2 + s^2\right)$ に戻せる。
・値を差し替えたら、**合計と** 2 **乗の合計だけ**を直せばよい。
・平均値と同じ値を除くと、**平均値は変わらず**、分散は 36 から 40 に**大きくなる**（個数が減るため）。
・偏差が 0 のデータは、偏差の 2 乗の合計に寄与しない。
・データが手元になくても、**平均値と分散があれば**計算できる。

使った Point（本編より。例題は省略）

Point 4-0-4 分散と標準偏差
（本文は問 8 参照）

問 16

x の分散は 0 でないとし、変量 y を $y = ax + b$（a, b は定数）で定める。
(1) $a = 2$, $b = 1$ のとき、s_{xy} を s_x^2 で表し、相関係数 r を求めよ。
(2) $a = -3$, $b = 5$ のとき、相関係数 r を求めよ。
(3) $a \neq 0$ のとき、一般に相関係数 r はどうなるか。また $a = 0$ のとき相関係数が定義できない理由を述べよ。

問題を見て思うこと

$y = ax + b$ なら、全部の点が**傾き a の直線上**に並ぶ。**完全な直線的関係**なので、$r = \pm 1$ と予想できる。 ▶Point 4-0-5
→ 確かめは s_{xy} と s_y を s_x で表すこと。y の偏差は $y - \overline{y} = a(x - \overline{x})$（$b$ は消える）。
→ すると $s_{xy} = a\,s_x^2$, $s_y = |a| s_x$。→ 比をとると $\dfrac{a}{|a|}$。$a = 0$ のときは $s_y = 0$ で割れない。
▶Point 4-0-6

解答

$y = ax + b$ のとき $\overline{y} = a\overline{x} + b$ であるから、$y - \overline{y} = a(x - \overline{x})$。

(1) $a = 2$ のとき $y - \overline{y} = 2(x - \overline{x})$ であり

$$s_{xy} = \frac{1}{n}\sum (x - \overline{x}) \cdot 2(x - \overline{x}) = \boldsymbol{2s_x^2}, \qquad s_y^2 = 4s_x^2 \ \text{より} \ \ s_y = 2s_x$$

よって $r = \dfrac{2s_x^2}{s_x \cdot 2s_x} = \boldsymbol{1}$。

(2) $a = -3$ のとき $s_{xy} = -3s_x^2$, $s_y = |-3| s_x = 3s_x$ であるから、$r = \dfrac{-3s_x^2}{s_x \cdot 3s_x} = \boldsymbol{-1}$。

(3) $a \neq 0$ のとき $s_{xy} = a\,s_x^2$, $s_y = |a| s_x$ であるから

$$r = \frac{a\,s_x^2}{s_x \cdot |a| s_x} = \frac{a}{|a|}$$

すなわち、$a > 0$ のとき $r = \boldsymbol{1}$、$a < 0$ のとき $r = \boldsymbol{-1}$。$a = 0$ のときは $y = b$（定数）で、y の偏差がすべて 0 なので $s_y = 0$。相関係数の分母 $s_x s_y$ が 0 になり、0 **で割ることになるため定義できない**。

方針の候補

○ y の偏差が $a(x - \overline{x})$ になることから、s_{xy} **と** s_y **を** s_x **で表す**。 ▶Point 4-0-5
△ x と x 自身の相関係数が 1 であることに、変換の表（$u = x$, $v = ax + b$）を使う … 解けはするが、x **と** x **の相関係数を考える発想**が要る。
× $a = 0$ でも $r = 1$ や $r = -1$ だとする … $s_y = 0$ **で割れない**ので、r は定義されない。
× b が r に影響すると考える … 偏差をとると b **は消える**。

別の方針でも解いてみる（略解）

△ 変換の表で x と x の相関係数から導く道

x と x の共分散は $s_{xx} = s_x^2$ だから、相関係数は $\dfrac{s_x^2}{s_x s_x} = 1$。ここで $u = x$（$a = 1,\ b = 0$）、$v = ax + b$（$c = a,\ d = b$）と見ると、変換の表の $ac = a$ より

$$r_{uv} = \frac{a}{|a|} \cdot 1 = \frac{a}{|a|}$$

→ 同じ結論に到達する。

この問題で身につくこと

・$y = ax + b$ なら、点は**完全に直線上**に並び、$r = \pm 1$。
・符号は**傾き a の符号**：$a > 0$ で $+1$、$a < 0$ で -1。
・b（切片）は**相関係数に影響しない**。
・$a = 0$ のときは $s_y = 0$ で、**相関係数は定義されない**。
・$|r| \leqq 1$ の「等号が成り立つ場合」の正体が、この直線上の並びである。

使った Point（本編より。例題は省略）

Point 4-0-5 共分散と相関係数
（本文は問 12 参照）

Point 4-0-6 変量の変換まとめ
（本文は問 14 参照）

問 18

A 組 n_A 人、B 組 n_B 人が同じテストを受け、A 組の平均値・分散を m_A，s_A^2、B 組の平均値・分散を m_B，s_B^2 とする。$N = n_A + n_B$ とし、2 つの組を合わせた N 人全体の平均値を m、分散を s^2 とする。

(1) 各組の「2 乗の平均」を m_A，s_A^2 などで表すことで、

$$s^2 = \frac{n_A s_A^2 + n_B s_B^2}{N} + \frac{n_A n_B \left(m_A - m_B\right)^2}{N^2}$$

が成り立つことを示せ。

(2) $n_A = 20$，$m_A = 60$，$s_A^2 = 80$，$n_B = 30$，$m_B = 70$，$s_B^2 = 70$ のとき、m と s^2 を求めよ。

(3) B 組の人数 30 と分散 70 は変えず、B 組の平均値だけを m_B（$m_B > 60$）に取りかえる。A 組はそのままで、全体の分散がちょうど 80 になるような m_B を求めよ。

問題を見て思うこと

分散は 2 **乗の平均** − **平均の** 2 **乗**。各組の平均値と分散から、**合計と** 2 **乗の合計**に戻せる。

▶ **Point 4-0-4**

→ 各組の 2 乗の平均は $s_A^2 + m_A^2$，$s_B^2 + m_B^2$。全体の 2 乗の平均は、人数で重みをつけた平均 $\dfrac{n_A\left(s_A^2 + m_A^2\right) + n_B\left(s_B^2 + m_B^2\right)}{N}$。

→ そこから m^2 を引く。m_A^2，m_B^2 の項が $\left(m_A - m_B\right)^2$ にまとまるかが山場。

→ (3) は (1) の式を逆向きに読む。第 1 項は m_B によらず 74。

解答

(1)　A 組の合計は $n_A m_A$，2 乗の合計は $n_A\left(s_A^2 + m_A^2\right)$。B 組も同様である。したがって

$$m = \frac{n_A m_A + n_B m_B}{N}, \qquad \overline{x^2} = \frac{n_A\left(s_A^2 + m_A^2\right) + n_B\left(s_B^2 + m_B^2\right)}{N}$$

分散は $s^2 = \overline{x^2} - m^2$ であるから

$$s^2 = \frac{n_A s_A^2 + n_B s_B^2}{N} + \left\{\frac{n_A m_A^2 + n_B m_B^2}{N} - \left(\frac{n_A m_A + n_B m_B}{N}\right)^2\right\}$$

右辺の { } の中を通分すると、分子は

$$\begin{aligned} &N\left(n_A m_A^2 + n_B m_B^2\right) - \left(n_A m_A + n_B m_B\right)^2 \\ &= n_A n_B m_A^2 + n_A n_B m_B^2 - 2 n_A n_B m_A m_B = n_A n_B\left(m_A - m_B\right)^2 \end{aligned}$$

（$N = n_A + n_B$ を用いて展開し、$n_A^2 m_A^2$ と $n_B^2 m_B^2$ が消える。）分母は N^2 だから、示された。

(2)　$m = \dfrac{20 \cdot 60 + 30 \cdot 70}{50} = \mathbf{66}$。(1) の式より

$$s^2 = \frac{20 \cdot 80 + 30 \cdot 70}{50} + \frac{20 \cdot 30 \cdot (60 - 70)^2}{50^2} = 74 + 24 = \mathbf{98}$$

(3)　第 1 項は $\dfrac{20\cdot 80+30\cdot 70}{50}=74$ のまま。第 2 項は $\dfrac{20\cdot 30\cdot(60-m_B)^2}{50^2}=\dfrac{6}{25}(m_B-60)^2$ であるから

$$74+\frac{6}{25}(m_B-60)^2=80 \quad\Longleftrightarrow\quad (m_B-60)^2=25$$

$m_B>60$ より $m_B-60=5$、よって $m_B=\mathbf{65}$。

方針の候補

○ 各組を**合計と 2 乗の合計**に戻して合成する。 ▶ Point 4-0-4

△ 全体の平均 m からの偏差の 2 乗の合計を、各組の平均からの偏差に分けて考える ··· 解けるが、**交差項が 0 になること**に気づく発想が要る。

× 全体の分散を $\dfrac{s_A^2+s_B^2}{2}$ とする ··· **人数の重み**を無視している。

× 全体の分散を各組の分散の加重平均 74 とする ··· 平均値が異なると、**組どうしのずれ**の分だけ大きくなる（98）。

別の方針でも解いてみる（略解）

△ 偏差の 2 乗の合計を分ける道

A 組の人について $x-m=(x-m_A)+(m_A-m)$ と分けて 2 乗して足すと、交差項は $2(m_A-m)\sum(x-m_A)=0$ なので $\sum_A(x-m)^2=n_As_A^2+n_A(m_A-m)^2$。B 組も同様。ここで $m_A-m=\dfrac{n_B}{N}(m_A-m_B)$, $m_B-m=-\dfrac{n_A}{N}(m_A-m_B)$ だから、組どうしのずれの項の合計は

$$\frac{n_An_B^2+n_Bn_A^2}{N^2}(m_A-m_B)^2=\frac{n_An_B}{N}(m_A-m_B)^2$$

これを N で割ると (1) の式になる。
→ 同じ結論に到達する。

この問題で身につくこと

・合成した分散 ＝ **組の中の散らばり（分散の加重平均）＋ 組どうしのずれ**。
・平均値が同じなら、**ずれの項が 0**。全体の分散は分散の加重平均に一致する。
・平均値・分散から**合計と 2 乗の合計**に戻すと、データがなくても合成できる（問 10）。
・(3) のように、**式を逆向きに読む**と未知の平均値が決まる。
・式の形が複雑なときは、**具体的な数値で先に確かめる**と見通しがよい。

使った Point（本編より。例題は省略）

Point 4-0-4 分散と標準偏差
（本文は問 8 参照）