

科目：データサイエンス

注意：根号が現れた場合開く必要はないが、簡単にできるものは簡単にせよ。

問 6 に関しては、必ず導出過程を記述すること。

問 1 次の 10 名の体重データについて、①平均値、②中央値、③最頻値、④分散、⑤標準偏差を求めよ。解答は数値のみとし、単位は必要ない。

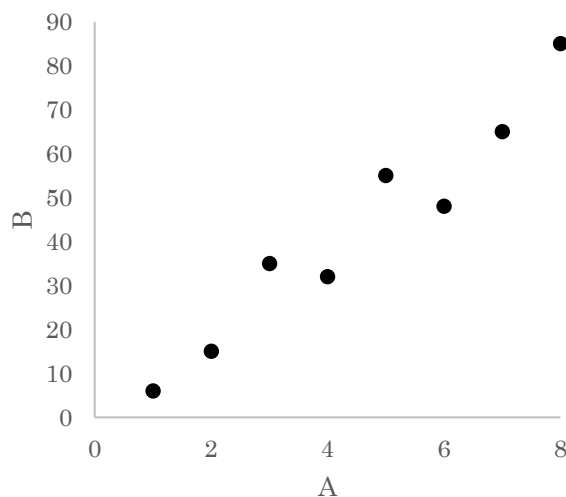
名前	佐藤	鈴木	高橋	田中	伊藤	渡辺	山本	中村	小林	加藤
体重[kg]	40	70	50	70	50	50	50	160	90	70

問 2 統計的仮説検定を使って、ある手法 A と手法 B の効果の比較をするとして、帰無仮説を「手法 A と手法 B の効果は同じである」とした時の次のものを答えよ。

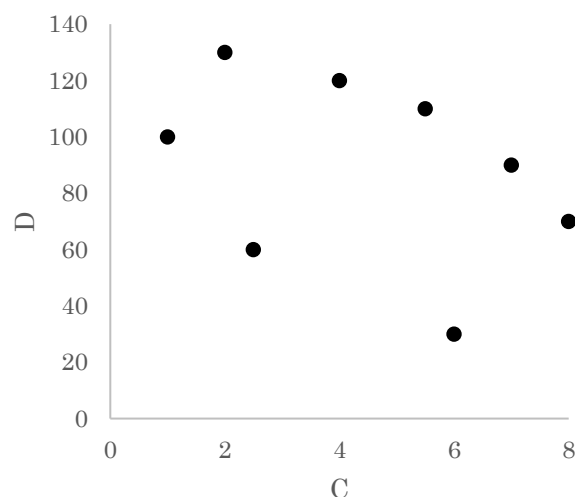
- ① 対立仮説
- ② 第 1 種の過誤
- ③ 第 2 種の過誤

問 3 次の相関に関する問いに答えよ。

- ① 次の 2 つのグラフの相関関係について答えよ。答えは「～の相関」という形式にせよ。ただし、グラフ (a) の相関係数の絶対値は約 0.97、グラフ (b) の相関係数の絶対値は約 0.39 である。



(a)



(b)

- ② 2 つの数 x , y について次のようなデータが得られた場合の、 x と y の相関係数を求めよ。

x	1	4	1	3	0	3
y	7	5	2	4	4	2

問 4 25 本のくじがある。このくじには当たりくじとして 1 等が 3 本, 2 等が 5 本あり, 残りは全てハズレくじである。このくじを同時に 3 本引くときを考える。次の問いに答えよ。

- ① 当たりくじが少なくとも 1 本引かれる確率を求めよ。
- ② 引いた当たりくじが 1 等, 2 等ともちょうど 1 本ずつ出る確率を求めよ。
- ③ 1 等を 1 本以上, かつ全て当たりくじを引く確率を求めよ。

問 5 XY 平面上の 5 点 $(-5, -6), (-3, -2), (1, 1), (4, 6), (6, 10)$ に最小二乗法を用いて, 適切な直線 $y = ax + b$ への当てはめを行う。次の問いに答えよ。

- ① 最小二乗法を適用したときの正規方程式は $\begin{bmatrix} x_1 & x_2 \\ x_2 & x_3 \end{bmatrix} \begin{bmatrix} a \\ b \end{bmatrix} = \begin{bmatrix} y_1 \\ y_2 \end{bmatrix}$ の形で表される。要素 x_1, x_2, x_3, y_1, y_2 の値を求めよ。
- ② 係数 a, b の値を求めよ。

問 6 行列 $A = [x_1 \ x_2 \ x_3] = \begin{bmatrix} 2 & 2 & 0 \\ 2 & 4 & 1 \\ 0 & 2 & 2 \end{bmatrix}$ の固有値を求めるため, QR 分解を行う。次の問いに答えよ。

- ① ベクトル x_1, x_2, x_3 の順に直交化を行い, 直交行列 Q を求めよ。
- ② $A = QR$ を満たす上三角行列 R を求めよ。