問題 1. 一日のスマホの利用時間 (単位: 時間), 睡眠時間 (単位: 時間) とストレス度 (A から C で評価, C が最大) の関係を調べるために, ある大学の学生 10 名を対象に調査を行い, 結果を以下のデータ行列にまとめた.

students

事例	スマホ利用時間	睡眠時間	ストレス度
1	3	8	A
2	4.5	7	B
3	4	7.5	A
4	6	7	C
5	8	5	A
6	2	7	B
7	5	6.5	C
8	5	6	C
9	6.5	5.5	A
10	5.5	6.5	B

$-1 \leq R \leq 1$.

$-1 \leq R < 0$

(a) この研究は観察研究と実験研究のどちらか.

(b) この研究における 3 つの変数, スマホ利用時間, 睡眠時間, ストレス度の中でカテゴリカル変数はどれか.

(c) スマホ利用時間が睡眠時間に影響を与えていると考えるとき, 説明変数と目的変数はそれぞれどちらか.

(d) この研究により問題 (c) における 2 つの変数の<u>因果</u> 関係を検証することは可能か, 可能でないか.

(a) 観察研究.

(b) ストレス度.

(c) 説明変数:スマホ利用時間.

目的変数:睡眠時間.

(d) 因果関係を検証することは可能でない.

(∵ 観察研究では相関関係までしか検証できないから).

(e) 各事例 i のスマホ利用時間と睡眠時間をそれぞれ x_i, y_i とする．スマホ利用時間と睡眠時間の平均 $\bar{x}$ と $\bar{y}$ をそれぞれ求めよ．

(f) スマホ利用時間と睡眠時間の<u>標準偏差</u> s_x と s_y をそれぞれ求めよ．ただし分散の式は n 個のデータ $\{x_i\}$ に対して，

$$s_x^2 = \underbrace{\frac{1}{n-1}}_{\frac{1}{n}} \sum_{i=1}^{n} (x_i - \bar{x})^2$$

(e) $\bar{x} = \frac{1}{n}\sum_{i=1}^{n} x_i = \boxed{4.95}$.

$\bar{y} = \frac{1}{n}\sum_{i=1}^{n} y_i = \boxed{6.6}$.

$s_x^2 = 2.97$, $s_x = 1.72$.

$s_y^2 = 0.82$, $s_y = 0.91$

$$\boxed{\begin{aligned} s_x &= 1.72 \\ s_y &= 0.91 \end{aligned}}$$

(f) $s_x^2 = \frac{1}{n-1}\sum_{i=1}^{n}(x_i - \bar{x})^2$.

$n=10$

$= \frac{n}{n-1}\frac{1}{n}\sum_{i=1}^{n}(x_i-\bar{x})^2 = \frac{n}{n-1}\left\{\frac{1}{n}\sum_{i=1}^{n}x_i^2 - (\bar{x})^2\right\}$

(g) スマホ利用時間と睡眠時間の相関係数 R を求めよ．ただし n 個のデータ (x_i, y_i) に対して，相関係数は以下の式で与えられるものとする:

$$R = \underbrace{\frac{1}{n-1}}_{\frac{1}{n}} \sum_{i=1}^{n} \frac{x_i - \bar{x}}{s_x} \frac{y_i - \bar{y}}{s_y}.$$

問題の場合:

負の相関, $R < 0$.

一般に, $-1 \leq R \leq 1$.

$$R = \frac{n}{n-1} \left\{ \frac{1}{n} \sum_{i=1}^{n} \frac{x_i - \bar{x}}{s_x} \cdot \frac{y_i - \bar{y}}{s_y} \right\}$$

$$= \frac{n}{n-1} \left\{ \frac{1}{n} \sum_{i=1}^{n} x_i y_i - \bar{x} \cdot \bar{y} \right\} \frac{1}{s_x s_y}.$$

$$= \boxed{-0.80}.$$

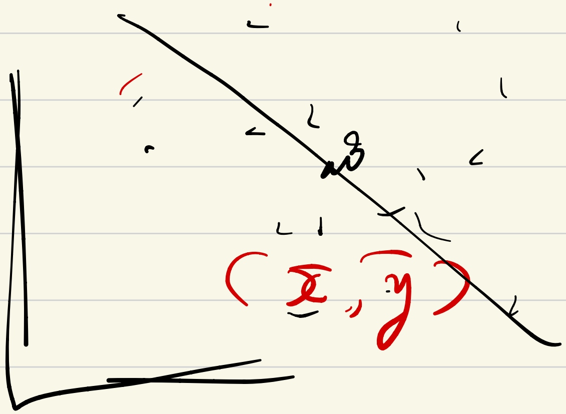
(h) スマホ利用時間 (x) から睡眠時間 ($\hat{y}$) を予測する最小二乗線

$$\hat{y} = b_0 + b_1 x$$

の切片 b_0 と傾き b_1 を求めよ.

(i) 問題 (h) の最小二乗線によって, スマホ利用時間と睡眠時間の関係を何%程度説明できているか.

(h).



$$b_1 = \frac{S_y}{S_x} R = \frac{0.91}{1.72} \times (-0.8)$$

$$= -0.42$$

$$b_0 = \bar{y} - b_1 \bar{x}$$

$$= 6.6 - (-0.42) \times 4.95$$

$$= 8.68$$

$$\boxed{\hat{y} = 8.68 - 0.42x}$$

(i). R^2 : 決定係数 を使って.

$$100 \times R^2 = 100 \times (-0.8)^2 = \boxed{64\%}$$

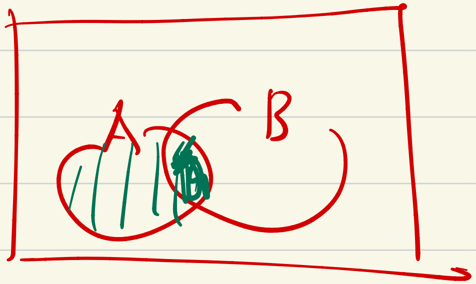
問題 2. 事象 A と事象 B の確率をそれぞれ $P(A) = 0.2$, $P(B) = 0.8$ とする.

(a) 確率 $P(A \cap B) = 0.1$ のとき, 確率 $P(A \cup B)$ を求めよ.

(b) 確率 $P(A \cap B) = 0.1$ のとき, 事象 A と事象 B は独立か, 独立でないか.

(c) 確率 $P(A \cap B) = 0.1$ のとき, 事象 A を条件とする事象 B の条件付き確率 $P(B|A)$ を求めよ.

(c). $P(B|A) := \dfrac{P(B \cap A)}{P(A)}$.



$= \dfrac{0.1}{0.2} = \underline{0.5}$ ⇒ 独立でない.

(b). 定義. 事象 A と B が独立 $\Leftrightarrow P(A)P(B) = P(A \cap B)$

$P(A \cap B) = 0.1 \neq$

$P(A) \cdot P(B) = 0.2 \times 0.8 = 0.16$

$\Rightarrow P(B) = \dfrac{P(A \cap B)}{P(A)} = P(B|A)$.

問題 3. 全国民のうち 10%がある素質 X を持っているという．ある遺伝子テストでは，素質があれば 90%の精度で陽性が出て，素質が無ければ 95%の精度で陰性が出る．陽性が出たときに実際に素質を持っている確率 P(素質あり|陽性) はいくつか．

与えられているもの

$P(\text{素質あり}) = 0.1$.　　$P(+\mid\text{素質あり}) = 0.9$.

$P(\text{素質なし}) = 0.9$.　　$P(-\mid\text{素質あり}) = 0.1$.

求めるもの

$P(-\mid\text{素質なし}) = 0.95$

$P(+\mid\text{素質なし}) = 0.05$.

$P(\text{素質あり}\mid +)$

$$P(\text{素質あり}\mid +) = \frac{P(\text{素質あり})P(+\mid\text{素質あり})}{P(\text{素質あり})P(+\mid\text{素質あり}) + P(\text{素質なし})P(+\mid\text{素質なし})}$$

問題 4. 確率変数 X の 5 つの値 $\{x_1, x_2, \ldots, x_5\}$ の確率分布が下記の表によって与えられている．このとき X の平均 μ, 分散 σ^2, 標準偏差 σ を求めよ．

i	1	2	3	4	5
x_i	2	3	4	-2	0
P(X=x_i)	0.15	0.25	0.1	0.2	0.3

0.95 1.95 2.95 −3.05 −1.05

$$= \frac{0.1 \times 0.9}{0.1 \times 0.9 + 0.9 \times 0.05}$$

$$= \frac{9}{9 + 4.5} = \frac{9}{13.5} = 0.67$$

$$\mu = E(X) = \sum_{i=1}^{5} x_i P(x_i)$$

1.05

$$\sigma = \sqrt{\sigma^2}$$

2.04

$$\sigma^2 = V(X) = \sum_{i=1}^{5} (x_i - \bar{x})^2 P(x_i)$$

4.15

問題 5. 太郎は A 社と B 社の株にそれぞれ 5000 ドルと 2000 ドルを投資した．A 社と B 社のひと月の株価の上昇率をそれぞれ X %と Y %とする．ここで X と Y はそれぞれ $E(X) = 2.5, V(X) = 4.0$ と $E(Y) = 2.0, V(Y) = 1.0$ の期待値と分散を持つ確率変数である．このときのひと月での儲け W ドルについて以下の問いに答えよ.

(a) 確率変数 W を確率変数 X と Y を使った式で表せ.

(b) 確率変数 W の期待値 $E(W)$ を求めよ.

(c) 確率変数 W の分散 $V(W)$ と標準偏差 $\sqrt{V(W)}$ を求めよ．ただし, 確率変数 X と Y は独立と仮定する.

(a). $W = 5000 \times \frac{X}{100} + 2000 \times \frac{Y}{100}$ (ドル)

$= 50X + 20Y$

(b). $E(W) = E(50X + 20Y)$

$$= 50\,E(X) + 20\,E(Y)$$

$$= 50 \times 2.5 + 20 \times 2.0$$

$$= 125 + 40$$

$$= 165$$

(c) $V(W) = V(50X + 20Y)$

$$= V(50X) + V(20Y)$$

$$= 50^2 \times \underset{4}{V(X)} + 20^2 \times \underset{1}{V(Y)}$$

$$= 10000 + 400$$

$$= 10400 \quad (\text{ドル}^2)$$

$$\sqrt{V(x)} = 101.98$$

この月のもうりは 165 ± 102 ドル。

授業のかだい … 40%

中間 + 期末 … 60%

60/100 以上が合格.