このように，散布図で，すべてのデータの点が負の傾きの直線上に存在するとき，その相関係数 r_{XY} は $r_{XY}=-1$ となる。これに対して，すべてのデータの点が正の傾きの直線上に存在するときは，その相関係数 r_{XY} は $r_{XY}=1$ となるんだね。これらのことも，もう 1 度頭に入れておこう。

それでは，このような特殊な場合ではない，一般的な 2 変数データの共分散 S_{XY} と相関係数 r_{XY} の問題を最後に 2 題解いておこう。

練習問題 41 共分散と相関係数 (II) CHECK 1 CHECK 2 CHECK 3

次の 7 組の 2 変数データがある。

$(\underset{x_1}{9},\ \underset{y_1}{6}),\ (\underset{x_2}{7},\ \underset{y_2}{1}),\ (\underset{x_3}{6},\ \underset{y_3}{9}),\ (\underset{x_4}{5},\ \underset{y_4}{7}),\ (\underset{x_5}{8},\ \underset{y_5}{4}),\ (\underset{x_6}{11},\ \underset{y_6}{3}),\ (\underset{x_7}{10},\ \underset{y_7}{5})$

ここで，2 変数 X，Y を，

$$\begin{cases} X=9,\ 7,\ 6,\ 5,\ 8,\ 11,\ 10 \\ Y=6,\ 1,\ 9,\ 7,\ 4,\ 3,\ 5 \end{cases}$$ **とおくとき，**

X と Y の共分散 S_{XY} と相関係数 r_{XY} を求めよ。

7 組の 2 変数データの問題だね。まず，表を利用して，X と Y の標準偏差 S_X，S_Y と共分散 S_{XY} を求めるんだね。そして，これから相関係数を $r_{XY}=\dfrac{S_{XY}}{S_X\cdot S_Y}$ により求める。

7 組の 2 変数データを

$(X,\ Y)=(\underset{x_1}{9},\ \underset{y_1}{6}),\ (\underset{x_2}{7},\ \underset{y_2}{1}),\ (\underset{x_3}{6},\ \underset{y_3}{9}),\ (\underset{x_4}{5},\ \underset{y_4}{7}),\ (\underset{x_5}{8},\ \underset{y_5}{4}),\ (\underset{x_6}{11},\ \underset{y_6}{3}),\ (\underset{x_7}{10},\ \underset{y_7}{5})$

とおく。よって，

$$\begin{cases} X=x_1,\ x_2,\ x_3,\ \cdots,\ x_7=9,\ 7,\ 6,\ \cdots,\ 10 \\ Y=y_1,\ y_2,\ y_3,\ \cdots,\ y_7=6,\ 1,\ 9,\ \cdots,\ 5 \end{cases}$$ より，

2 変数 X と Y の平均値 m_X と m_Y，標準偏差 S_X と S_Y，そして共分散 S_{XY} を次の表を利用して算出し，相関係数 r_{XY} を，

公式：$r_{XY}=\dfrac{S_{XY}}{S_X\cdot S_Y}$ から求めてみよう。

表 5　標準偏差 S_X, S_Y と共分散 S_{XY} の計算

データ X から平均値 $m_X=8$ を引いたもの

データ Y から平均値 $m_Y=5$ を引いたもの

データ No	データ X	偏差 x_k-m_X	偏差平方 $(x_k-m_X)^2$	データ Y	偏差 y_k-m_Y	偏差平方 $(y_k-m_Y)^2$	$(x_k-m_X)(y_k-m_Y)$
1	9	1	1	6	1	1	$1\ (=1\times1)$
2	7	-1	1	1	-4	16	$4\ (=-1\times(-4))$
3	6	-2	4	9	4	16	$-8(=-2\times4)$
4	5	-3	9	7	2	4	$-6(=-3\times2)$
5	8	0	0	4	-1	1	$0\ (=0\times(-1))$
6	11	3	9	3	-2	4	$-6(=3\times(-2))$
7	10	2	4	5	0	0	$0\ (=2\times0)$
合計	56	0	28	35	0	42	-15
平均	8		4	5		6	$-\frac{15}{7}$
	m_X		$S_X{}^2$	m_Y		$S_Y{}^2$	S_{XY}

以上より,

X の平均値 $m_X=8$, 分散 $S_X{}^2=4$, 標準偏差 $S_X=\sqrt{S_X{}^2}=\sqrt{4}=2$

Y の平均値 $m_Y=5$, 分散 $S_Y{}^2=6$, 標準偏差 $S_Y=\sqrt{S_Y{}^2}=\sqrt{6}$

X と Y の共分散 $S_{XY}=-\dfrac{15}{7}$ ………(答)

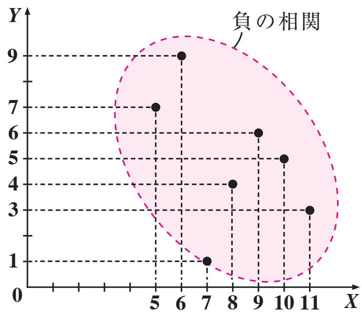
よって, X と Y の相関係数 r_{XY} は,

$$r_{XY}=\frac{S_{XY}}{S_X\cdot S_Y}=\frac{-\frac{15}{7}}{2\cdot\sqrt{6}}$$

$$=-\frac{15}{14\sqrt{6}}=-\frac{15\sqrt{6}}{84}$$

$$=-\frac{5\sqrt{6}}{28}\ (\fallingdotseq -0.437)$$ ………(答)

図 4　散布図



よって,散布図に示すように X と Y の間に負の相関があることが分かるんだね。