

數學 2

進階
講義

相關係數

景美女中 · 莊嘉銘 老師



信望愛文教基金會



$\frac{3}{4}$



7-2-2 相關係數

定理證明或說明

1. 相關係數

$$(1) r = \frac{\sum (x_i - \bar{x})(y_i - \bar{y})}{\sqrt{\sum (x_i - \bar{x})^2 \cdot \sum (y_i - \bar{y})^2}} \text{ 用於：平均數為整數。}$$

$$(2) r = \frac{\sum x_i y_i - n \cdot \bar{x} \cdot \bar{y}}{\sqrt{[\sum x_i^2 - n \cdot \bar{x}^2][\sum y_i^2 - n \cdot \bar{y}^2]}} \text{ 用於：平均數不為整數或資料未知。}$$

(3) 若 r 表 x, y 之相關係數，則 $-1 \leq r \leq 1$ 。

(4) 設 $P_i = ax_i + b$ ， $Q_i = cy_i + d$ ，則：

(I) 若 $ac > 0$ ，則 $r_{PQ} = r_{xy}$ 。 (II) 若 $ac < 0$ ，則 $r_{PQ} = -r_{xy}$ 。

2. (1) $r = 1$ ： X 與 Y 為完全正相關，其散佈圖為一斜率為正的直線圖形。

(2) $0.7 \leq r < 1$ ： X 與 Y 為高度正相關。

(3) $0.3 \leq r < 0.7$ ： X 與 Y 為中度正相關。

(4) $0 < r < 0.3$ ： X 與 Y 為低度正相關。

(5) $-0.3 \leq r < 0$ ： X 與 Y 為低度負相關。

(6) $-0.7 \leq r < -0.3$ ： X 與 Y 為中度負相關。

(7) $-1 < r < -0.7$ ： X 與 Y 為高度負相關。

(8) $r = -1$ ： X 與 Y 為完全負相關，其散佈圖為一斜率為負的直線圖形。

關鍵字

相關係數、相關程度

例題 1

五位學生的數學成績 X 與英文成績 Y 如表，求相關係數為何？

數學成績 X	81	75	77	69	73
英文成績 Y	86	76	83	80	75

Ans :

x	y	$x - \mu_x$	$y - \mu_y$	$(x - \mu_x)(y - \mu_y)$	$(x - \mu_x)^2$	$(y - \mu_y)^2$
81	86	6	6	36	36	36
75	76	0	-4	0	0	16
77	83	2	3	6	4	9
69	80	-6	0	0	36	0
73	75	-2	-5	10	4	25
375	400			52	80	86

$$\mu_x = \frac{375}{5} = 75$$

$$\mu_y = \frac{400}{5} = 80$$

$$\text{故 } r = \frac{52}{\sqrt{80 \times 86}} = \frac{52}{82.95} = 0.63$$

例題 2

若 $n = 12$, $\sum_{i=1}^{12} x_i = 120$, $\sum_{i=1}^{12} x_i^2 = 3700$, $\sum_{i=1}^{12} y_i = 240$, $\sum_{i=1}^{12} y_i^2 = 8400$,

$\sum_{i=1}^{12} x_i y_i = 3600$, 則相關係數為何 ?

Ans :

$$\text{可由題意知 } \mu_x = \frac{120}{12} = 10$$

$$\mu_y = \frac{240}{12} = 20$$

$$\begin{aligned} \text{所以相關係數 } r &= \frac{\sum_{i=1}^{12} x_i y_i - 12 \mu_x \cdot \mu_y}{\left(\sqrt{\sum_{i=1}^{12} x_i^2 - 12 \mu_x^2} \right) \left(\sqrt{\sum_{i=1}^{12} y_i^2 - 12 \mu_y^2} \right)} \\ &= \frac{3600 - 12 \times 10 \times 20}{(\sqrt{3700 - 12 \times 100})(\sqrt{8400 - 12 \times 400})} \\ &= \frac{1200}{\sqrt{2500} \cdot \sqrt{3600}} = 0.4 \end{aligned}$$

例題 3

有兩組變數 $X : x_1, x_2, \dots, x_n$, $Y : y_1, y_2, \dots, y_n$,

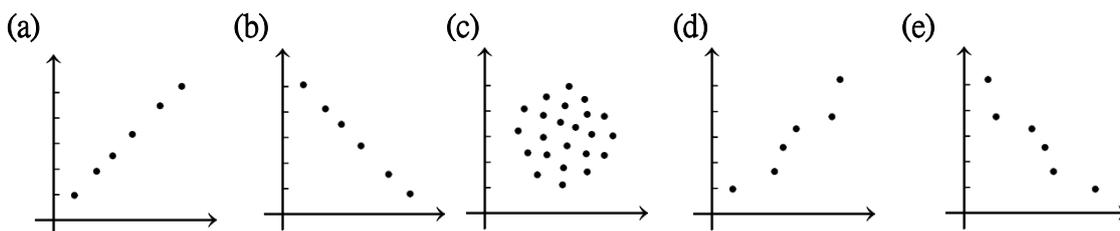
試求下列各條件下兩變數之相關係數

- (1) 若 $y_i = 0.8x_i - 1000$, $i = 1, 2, \dots, n$
- (2) 若 $y_i = 115 - 710x_i$, $i = 1, 2, \dots, n$
- (3) 若 $y_i = 1011$, $i = 1, 2, \dots, n$
- (4) 若已知 $r_{xy} = 0.75$, 則變數 $2X - 3$ 與 $5Y + 3$ 的相關係數
- (5) 若已知 $r_{xy} = 0.75$, 則變數 $2X - 3$ 與 $-35Y - 13$ 的相關係數

Ans :

- (1) 各點在一直線上，且斜率為正，故 $r_{xy} = 1$
- (2) 各點在一直線上，且斜率為負，故 $r_{xy} = -1$
- (3) 由於點在直線上 $y = 1011$ ，故 $r_{xy} = 0$
- (4) 由於 $2 \times 5 = 10 > 0$ ，故所求之相關係數與 r_{xy} 相同為 0.75
- (5) 由於 $2 \times (-35) = -70 < 0$ ，故所求之相關係數與 r_{xy} 相反為 -0.75

例題 4



請排出上面 5 個散佈圖中 x, y 的相關係數的大小次序（由小到大）。

Ans :

由圖知，(a)為完全正相關、(b)為完全負相關、(c)為零相關

(d)為正相關、(e)為負相關

故 $(b) < (e) < (c) < (d) < (a)$



習題 1

某廠商欲推出一種新產品，在上市以前以不同的單價 x (單位：10 元)，調查市場的需求量 y (單位：萬盒)，調查結果如下：

x	8	9	10	11	12
y	11	12	10	8	9

求單價與需求量的相關係數

習題 2

兩變量 X 與 Y ，若 $\sum_{i=1}^{10} x_i = 1250$ ， $\sum_{i=1}^{10} x_i^2 = 158500$ ， $\sum_{i=1}^{10} y_i = 800$ ，

$\sum_{i=1}^{10} y_i^2 = 64810$ ， $\sum_{i=1}^{10} x_i y_i = 101080$ ，求 X 與 Y 的相關係數

習題 3

設兩變量 X, Y 滿足 $\bar{X} = 10$ ， $\bar{Y} = 15$ ， $S_X = 1$ ， $S_Y = 2$ ， $r_{XY} = 0.72$ 且 $P = 3X + 2$ ， $Q = 2Y + 7$ ，則：

(1) $\bar{P} = \underline{\hspace{2cm}}$ 。 (2) $S_Q = \underline{\hspace{2cm}}$ 。 (3) $r_{PQ} = \underline{\hspace{2cm}}$ 。

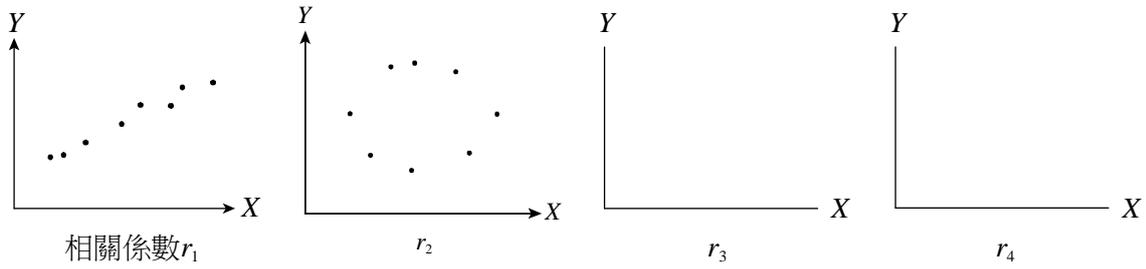
習題 4

關於兩變數 X 與 Y 的相關係數 r ，選出正確的選項：

- (1) r 的範圍為 $-1 \leq r \leq 1$
- (2) r 愈大，表示 X 與 Y 的相關程度愈高
- (3) 當各數據點都在一直線上時， $r = 1$
- (4) 將 X 與 Y 標準化後，其相關係數不變
- (5) 若 $X' = 5X$ ， $Y' = 3Y$ ，則 X' 與 Y' 的相關係數為 $15r$

習題 5

試將下列各散佈圖的相關係數 r_1, r_2, r_3, r_4 由大到小排列



解答與解析

習題 1 : -0.8

【詳解】可先求出 $\mu_x = 10$ 及 $\mu_y = 10$

x	8	9	10	11	12	和
y	11	12	10	8	9	
$x - \mu_x$	-2	-1	0	1	2	
$y - \mu_y$	1	2	0	-2	-1	
$(x - \mu_x)^2$	4	1	0	1	4	10
$(y - \mu_y)^2$	1	4	0	4	1	10
$(x - \mu_x)(y - \mu_y)$	-2	-2	0	-2	-2	-8

$$\begin{aligned}
 \text{所以相關係數 } r &= \frac{\sum_{i=1}^5 (x_i - \mu_x)(y_i - \mu_y)}{\sqrt{\sum_{i=1}^5 (x_i - \mu_x)^2 \cdot \sum_{i=1}^5 (y_i - \mu_y)^2}} \\
 &= \frac{-8}{\sqrt{10 \times 10}} = -0.8
 \end{aligned}$$

習題 2 : 0.8

【詳解】先求出 $\mu_x = \frac{\sum_{i=1}^{10} x_i}{10} = 125$ 及 $\mu_y = \frac{\sum_{i=1}^{10} y_i}{10} = 80$

$$\begin{aligned} \text{所以相關係數 } r &= \frac{\sum_{i=1}^{10} x_i y_i - 10\mu_x \mu_y}{\sqrt{(\sum_{i=1}^{10} x_i^2 - 10\mu_x^2) \cdot (\sum_{i=1}^{10} y_i^2 - 10\mu_y^2)}} \\ &= \frac{101080 - 10 \times 125 \times 80}{\sqrt{(158500 - 10 \times 125^2)(64810 - 10 \times 80^2)}} \\ &= \frac{1080}{\sqrt{2250 \times 810}} = 0.8 \end{aligned}$$

習題 3 : (1)32 (2)4 (3)0.72

【詳解】(1) 因為 $P = 3X + 2 \Rightarrow \bar{P} = 3\bar{X} + 2 = 3 \times 10 + 2 = 32$

(2) 因為 $Q = 2Y + 7 \Rightarrow S_Q = 2S_Y = 4$

(3) 因為 $3 \times 2 = 6 > 0$ ，所以 $r_{PQ} = r_{XY} = 0.72$

習題 4 : (1)(4)

【詳解】(2) 錯誤， r 的絕對值愈大才表示相關程度愈高

(3) 錯誤，可能是 -1 或 0

(5) 錯誤，伸縮變量時相關係數不變或為相反數

故選項(1)(4)正確

習題 5 : $r_1 > r_4 > r_2 > r_3$

【詳解】由圖中可知， $r_1 > r_4 > 0$ ， $r_2 \approx 0$ ， $r_3 < 0$

故 $r_1 > r_4 > r_2 > r_3$