單元 9 二維數據分析 習題 一年____班 座號：____ 姓名：

1.以下各小題對的打「O」，錯的打「×」:

___(1)散布圖上的點愈多，其相關程度愈高

___(2)若 x 與 y 的相關係數 $r>0$，則 y 對 x 的迴歸直線的斜率 $m>0$

___(3)若 x 與 y 的標準差相等，則其相關係數與 y 對 x 迴歸直線的斜率也相等

___(4)相關係數 r 滿 $-1\le r\le 1$

___(5)若散布圖的所有點都在直線 $y=2x-1$ 上，則相關係數為 1

解：(1) ×：散布圖上的點愈接近在一條斜直線上，相關程度愈高

(2) O：因為 $m=r\times\dfrac{\sigma_y}{\sigma_x}$，且 $\sigma_x>0$，$\sigma_y>0$，所以 m 與 r 性質符號相同

(3) O：因為 $m=r\times\dfrac{\sigma_y}{\sigma_x}$，且 $\sigma_x=\sigma_y$，$\therefore m=r$

(4) O：相關係數 r 必滿足 $-1\le r\le 1$

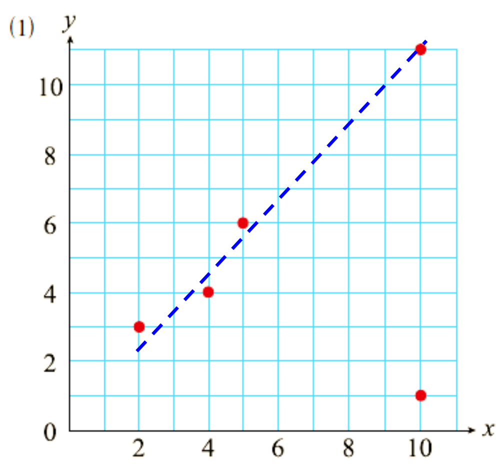
(5) O：散布圖中各點恰落在一條斜率為正的直線上，$r=1$

2.有五筆數據資料如下表：

代號	A	B	C	D	E
x	2	4	5	10	10
y	3	4	6	1	11

(1)繪出此數據的散布圖

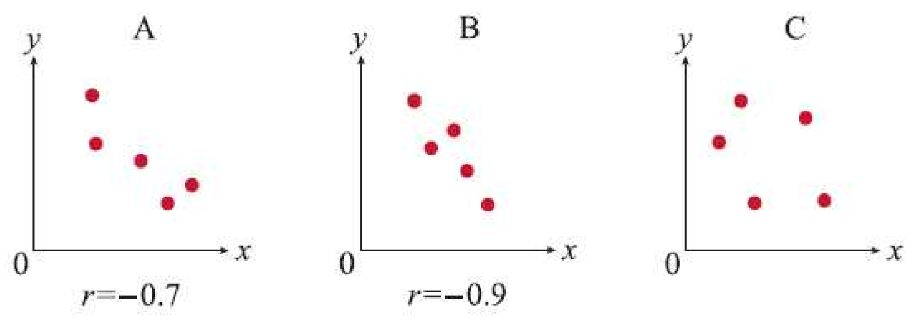
解：如右圖



(2)試問去掉哪一筆數據後，剩下來四筆數據的相關係數最大？

解：由散布圖知：迴歸直線如圖中虛線，∴去掉資料 D 後，剩下來的四筆數據的相關係數最大

3.兩組資料 A 與 B 的散布圖與相關係數如下圖所示。



下列哪一個選項最可能是資料 C 散布圖的相關係數？(1) -1.1 (2) -0.8 (3) -0.4 (4) 0.2 (5) 0.4

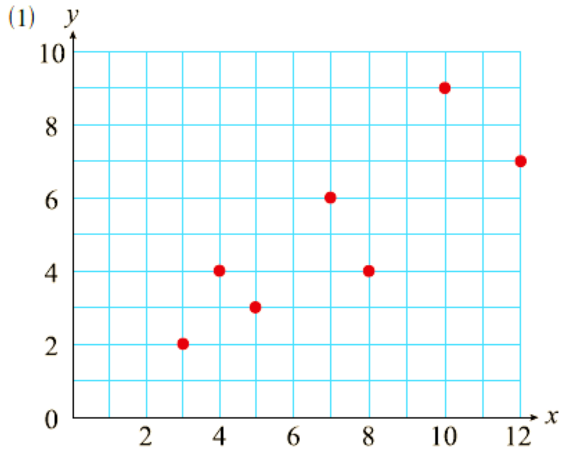
解：因為資料 C 為負相關，且其散布圖較資料 A 的散布圖分散，所以 C 的相關係數大於 -0.7。故選 3

4.七位同學的性向測驗 x 與成就測驗 y 的成績如下表：

性向測驗 x	5	3	7	4	8	12	10
成就測驗 y	3	2	6	4	4	7	9

(1)繪出此數據的散布圖

(2)求 x 與 y 的相關係數



解：(1)如右圖

(2)製表格如下： $r=\dfrac{S_{xy}}{\sqrt{S_{xx}}\sqrt{S_{yy}}}\dfrac{40}{\sqrt{64}\sqrt{36}}=\dfrac{40}{48}=\dfrac{5}{6}$

x	y	$x-\mu_x$	$y-\mu_y$	$(x-\mu_x)^2$	$(y-\mu_y)^2$	$(x-\mu_x)(y-\mu_y)$
5	3	-2	-2	4	4	4
3	2	-4	-3	16	9	12
7	6	0	1	0	1	0
4	4	-3	-1	9	1	3
8	4	1	-1	1	1	-1
12	7	5	2	25	4	10
10	9	3	4	9	16	12
$\mu_x=7$	$\mu_y=5$	0	0	$S_{xx}=64$	$S_{xy}=36$	$S_{xy}=40$

5.某公司欲推出一種新電腦，在上市前以不同的單價 x (單位：千元)調查市場的需求量 y (單位：十萬臺)。

調查結果如下表：

x	16	17	18	19	20
y	9	10	8	6	7

(1)求 x 與 y 的相關係數

(2)求 y 對 x 的迴歸直線方程式

(3)利用迴歸直線預測：當上市的單價訂為 17.5 千元時，市場的需求量為多少？

解：(1)製表格如下：製表格如下： $r=\dfrac{S_{xy}}{\sqrt{S_{xx}}\sqrt{S_{yy}}}\dfrac{-8}{\sqrt{10}\sqrt{10}}=-\dfrac{4}{5}=-0.8$

x	y	$x-\mu_x$	$y-\mu_y$	$(x-\mu_x)^2$	$(y-\mu_y)^2$	$(x-\mu_x)(y-\mu_y)$
16	9	-2	1	4	1	-2
17	10	-1	2	1	4	-2
18	8	0	0	0	0	0
19	6	1	-2	1	4	-2
20	7	2	-1	4	1	-2
$\mu_x=18$	$\mu_y=8$	0	0	$S_{xx}=10$	$S_{xy}=10$	$S_{xy}=-8$

(2) y 對 x 的迴歸直線 $y-\mu_y=\dfrac{S_{xy}}{S_{xx}}(x-\mu_x)$，$\therefore y-8=-\dfrac{4}{5}(x-18)$，得 $y=-\dfrac{4}{5}x+\dfrac{112}{5}$

(3) $x=17.5$ 代入 $y=-\dfrac{4}{5}x+\dfrac{112}{5}=-\dfrac{4}{5}(17.5)+\dfrac{112}{5}=8.4$

6.已知變量 x 的平均數為 2，標準差為 3；變量 y 的平均數為 7，標準差為 8，且 y 對 x 的迴歸直線方程式為 $y=2x+k$，求實數 k 及 x 與 y 的相關係數

解：(1)迴歸直線方程式為 $y=2x+k$ 必過$(\mu_x，\mu_y)$，∴(2，7)代入 $y=2x+k$，∴$7=2\times2+k$，得 $k=3$

(2)設相關係數為 r，直線方程式斜率 $m=r\cdot\dfrac{\sigma_y}{\sigma_x}$，∴$2=r\cdot\dfrac{8}{3}$，∴相關係數為 $r=\dfrac{3}{4}$

7.有 20 筆數據$(x_i，y_i)$，$i=1，2，\cdots，20$，其平均數 $\mu_x=6$，$\mu_y=5$，變量 x 與 y 的相關係數 $r=-0.9$，且 y 對 x 的迴歸直線通過點(4，6)，選出正確的選項：

(1)迴歸直線通過點(6，5)　　(2)迴歸直線斜率為 -0.5

(3)迴歸直線通過點(2，7)　　(4) x 的標準差小於 y 的標準差

解：(1) O：迴歸直線必過$(\mu_x，\mu_y)=(6，5)$

(2) O：迴歸直線必過(6，5)，又過(4，6)，由兩點得斜率 $m=\dfrac{5-6}{6-4}=-0.5$

(3) O：迴歸直線 $y-5=-0.5(x-6)$，將(2，7)代入滿足，∴迴歸直線通過點(2，7)

(4) ×：斜率 $m=r\cdot\dfrac{\sigma_y}{\sigma_x}$，∴$-0.5=-0.9\cdot\dfrac{\sigma_y}{\sigma_x}$，∴$\dfrac{\sigma_y}{\sigma_x}=\dfrac{5}{9}$，得知 $\sigma_x>\sigma_y$，即 x 的標準差小於 y 的標準差

8.段考數學考題共有 20 題，每題 5 分，總分 100 分。經統計：全校數學最高分 80 分，且數學成績與英文成績的相關係數為 0.73。已知老師將數學每題 5 分的配分更改為每題 6 分，藉以調高數學成績，求：

(1)調高後的數學成績與英文成績的相關係數

(2)調高後的數學成績與原數學成績的相關係數

解：設原數學成績為 x，調高後的數學成績為 y，英文成績為 z

得知 x 與 z 的相關係數為 $r=0.73$，且 y 與 x 的關係為 $y=\dfrac{6}{5}x$

(1)∵y 與 x 的關係為 $y=\dfrac{6}{5}x$，其中 $\dfrac{6}{5}>0$，∴調高後的數學成績與英文成績的相關係數 $r=0.73$

(2)調高後的數學成績 y 與原數學成績 x 的關係式為 $y=\dfrac{6}{5}x$，表示散布圖中所有點都在直線 $y=\dfrac{6}{5}x$ 上

⇒調高後的數學成績與原數學成績的相關係數 $r=1$