第8章 二元隨機變數及其機率分配

林惠玲 陳正倉著 雙葉書廊發行 2004

◎概念

許多的事件是同時發生且相互關聯的

酒後開車與車禍

廣告支出與商品銷售量

→ 二個或二個以上相互有關的事件

→ 二元或多元隨機變數 (bivariate or multivariate random variable)

學習目的

1. 定義或了解二元間斷隨機變數與連續隨機變數的意義及其機率分配。
2. 了解邊際機率分配與條件機率分配。
3. 了解兩變數間的關係。
4. 了解二元隨機變數函數的期望值。
5. 了解多元隨機變數。

本章結構

- 二元隨機變數及其機率分配
  - 二元間斷隨機變數
    - 聯合機率函數
    - 邊際機率函數
    - 條件機率函數
    - 二元間斷隨機變數的期望值與變異數
    - X與Y的條件平均數與變異數
  - 二元連續隨機變數
    - 二元機率密度函數
    - 邊際機率密度函數
    - 條件機率密度函數
    - 二元連續隨機變數的期望值與變異數
    - X與Y的條件平均數與變異數
  - 兩變數間的關係
    - 兩變數獨立
    - 兩變數不獨立—共變與相關
    - 共變數
    - 相關係數
    - 共變數與相關係數的比較
    - 獨立與共變的關係
  - 二元隨機變數函數的期望值
  - 多元隨機變數
  - Excel的使用

8.1 二元間斷隨機變數

8.1.1 聯合機率函數

設 X,Y 爲二元間斷隨機變數，X 之值爲 $x_1,x_2,\cdots,x_n$，Y 之值爲 $y_1,y_2,\cdots y_m$，若 $f(x_i,y_j)$ 滿足機率的二條件：

① $0\le f(x_i,y_j)\le 1$

② $\sum_{i=1}^{n}\sum_{j=1}^{m} f(x_i,y_j)=1$

則 $f(x_i,y_j)$ (簡單表示爲 $f(x,y)$)爲聯合機率函數。

求 X與Y某一範圍的機率

$a \le x \le b \qquad c \le y \le d$

$$P(a \le X \le b, c \le Y \le d) = \sum_{a \le x \le b} \sum_{c \le y \le d} f(x, y)$$

二元累加機率函數

$$F(x, y) = P(X \le x, Y \le y) = \sum_{X \le x} \sum_{Y \le y} f(x, y)$$

Ex8.1 設隨機變數X與Y分別為王先生與李先生一週內出售的房屋數, 已知X與Y的聯合機率分配如表8.1

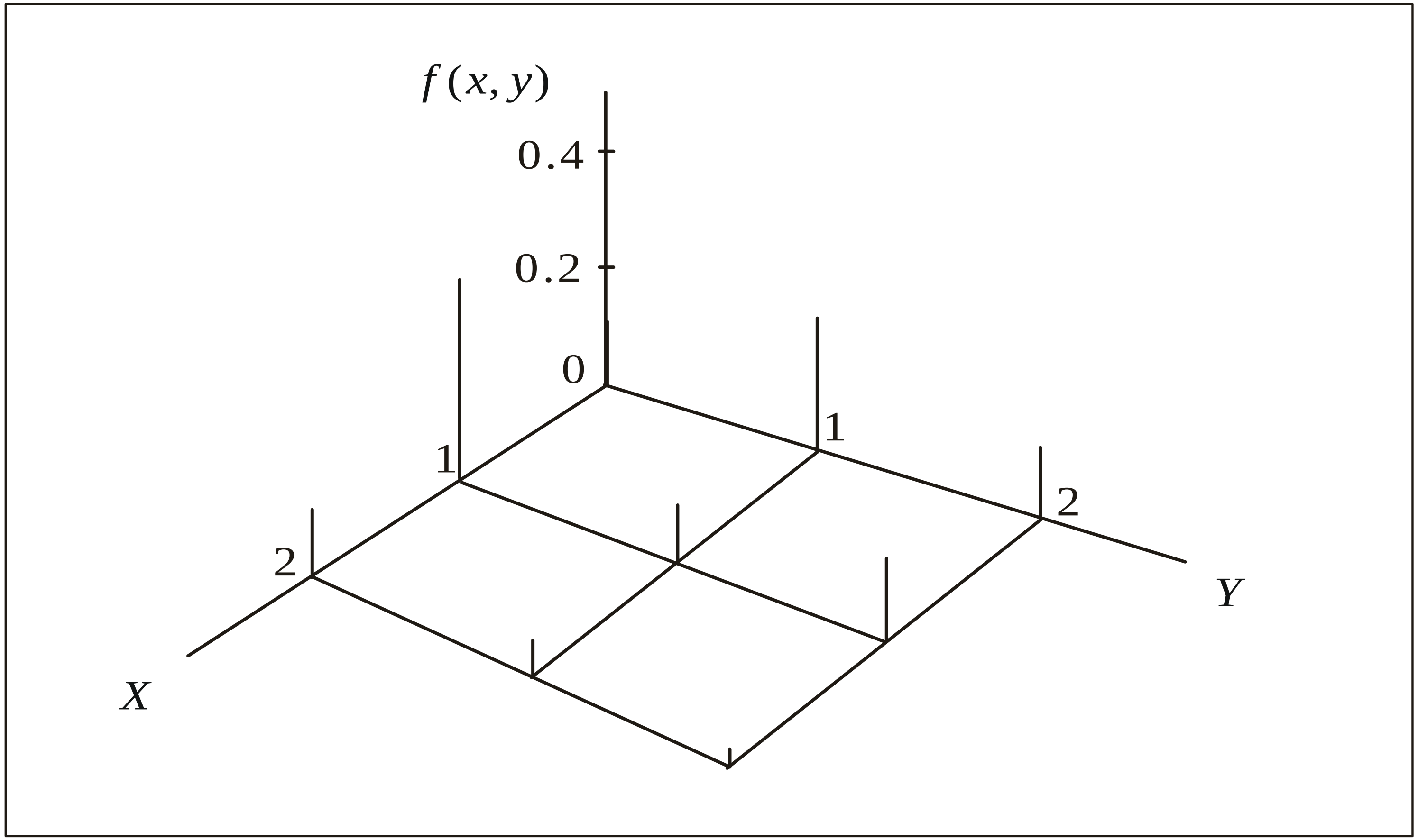
1. 王先生和李先生各賣一間房屋的機率為何?
2. 王先生和李先生各至少賣一間房屋的機率為何?
3. 王先生和李先生合起來賣二間房屋的機率為何?
4. 試求累加機率函數F(x,y)

表8.1 王先生與李先生售屋的機率

		Y		
		0	1	2
X	0	0.12	0.20	0.08
	1	0.32	0.07	0.01
	2	0.06	0.03	0.01

林惠玲 陳正倉著 雙葉書廊發行 2004

圖8.1 聯合機率分配圖



Solution:

1. P(X=1,Y=1)=0.07

2. $$P(X \geq 1, Y \geq 1) = \sum_{1 \leq x \leq 2} \sum_{1 \leq y \leq 2} f(x, y)$$

= f(1,1)+f(1,2)+f(2,1)+f(2,2)

=0.07+0.11+0.03+0.01=0.22

3. P(X+Y=2)=f(2,0)+f(1,1)+f(0,2)

=0.06+0.07+0.08=0.21

4. 累加機率F(x,y) 見下頁

林惠玲 陳正倉著 雙葉書廊發行 2004

表8.2 王先生與李先生售屋-累加機率函數

		Y		
		0	1	2
X	0	0.12	0.32	0.40
	1	0.44	0.71	0.90
	2	0.50	0.08	1.00

8.1.2 邊際機率函數

- X **的邊際機率函數**

$$f_x(x)=\sum_y f(x,y)=f(x,y_1)+f(x,y_2)+\cdots+f(x,y_n)$$

上式必須滿足下列兩條件：

①$0\leq f_x(x_i)\leq 1$

②$\sum_{i=1}^{n} f_x(x_i)=1$

二元間斷隨機變數

- Y 的邊際機率函數

$$f_y(y)=\sum_x f(x,y)=f(x_1,y)+f(x_2,y)+\cdots+f(x_n,y)$$

上式必須滿足下列兩條件：

①$0\le f_y(y_j)\le 1$

②$\sum_{j=1}^{m} f_y(y_j)=1$

林惠玲 陳正倉著 雙葉書廊發行 2004

表8.3 X 與 Y 的聯合機率分配與邊際機率分配表

$X \backslash Y$	y_1	y_2	$\cdots$	y_j	$\cdots$	y_m	$f_x(x_i)$
x_1	$f(x_1,y_1)$	$f(x_1,y_2)$	$\cdots$	$f(x_1,y_j)$	$\cdots$	$f(x_1,y_m)$	$f_x(x_1)$
x_2	$f(x_2,y_1)$	$f(x_2,y_2)$	$\cdots$	$f(x_2,y_j)$	$\cdots$	$f(x_2,y_m)$	$f_x(x_2)$
$\vdots$	$\vdots$	$\vdots$	$\cdots$	$\vdots$	$\cdots$	$\vdots$	$\vdots$
x_i	$f(x_i,y_1)$	$f(x_i,y_2)$	$\cdots$	$f(x_i,y_j)$	$\cdots$	$f(x_i,y_m)$	$f_x(x_i)$
$\vdots$	$\vdots$	$\vdots$	$\cdots$	$\vdots$	$\cdots$	$\vdots$	$\vdots$
x_n	$f(x_n,y_1)$	$f(x_n,y_2)$	$\cdots$	$f(x_n,y_j)$	$\cdots$	$f(x_n,y_m)$	$f_x(x_n)$
$f_y(y_j)$	$f_y(y_1)$	$f_y(y_2)$	$\cdots$	$f_y(y_j)$	$\cdots$	$f_y(y_m)$	1

8.1.3 條件機率函數

條件機率函數

設 $f(x,y)$ 爲二元機率函數，則在 $Y=y_j$ 條件下，發生 x_i 的條件機率表爲：

$$f(x_i \mid Y=y_j)=\frac{f(x_i,y_j)}{f_y(y_j)}$$

在 $X=x_i$ 條件下，發生 y_j 的條件機率表爲：

$$f(y_j \mid X=x_i)=\frac{f(x_i,y_j)}{f_x(x_i)}$$

Xi的條件機率必須滿足下列兩個條件

$$1.\ 0 \le f(x_i \mid Y = y_j) \le 1$$

$$2.\ \sum_{i=1}^{n} f(x_i \mid Y = y_j) = 1$$

yi的條件機率必須滿足下列兩個條件

$$1.\ 0 \le f(y_j \mid X = x_i) \le 1$$

$$2.\ \sum_{j=1}^{n} f(y_j \mid X = x_i) = 1$$

EX8.2 松青超市興隆店有3個結帳櫃檯. 現有2個顧客要進入櫃檯結帳(假設現時沒有顧客在結帳), 該2個顧客可隨機且獨立選擇櫃檯. 令X爲選擇1號櫃檯的顧客人數, Y惟選擇2號櫃檯的顧客人數. 試求

1. X與Y的聯合機率
2. 試求2號櫃檯有1位顧客的條件下, 1號櫃檯顧客人數的條件機率函數 f(x|Y=1) , 及 1號櫃檯有1位顧客的條件下, 2號櫃檯顧客人數的條件機率函數 f(y|X=1)

Solution:

1. 每位顧客有3個櫃台可選擇, 共有9個選擇方式
   每種選擇的機率為1/9, 可得聯合機率分配如下

表8.4 X 與 Y 的聯合機率分配表

		Y（2 號櫃台結帳人數）		
		0	1	2
X（1號櫃台結帳人數）	0	1/9	2/9	1/9
	1	2/9	2/9	0
	2	1/9	0	0

林惠玲　陳正倉著 雙葉書廊發行 2004

2. X與Y的邊際機率函數為

$$f_x(x)=\sum_{j=1}^{m} f(x,y_j)=f(x,0)+f(x,1)+f(x,2)$$

$$f_y(y)=\sum_{i=1}^{n} f(x_i,y)=f(0,y)+f(1,y)+f(2,y)$$

表8.5 X與Y的邊際機率分配表

		Y（2號櫃台結帳人數）			$f_x(x)$
		0	1	2	
X（1號櫃台結帳人數）	0	1/9	2/9	1/9	4/9
	1	2/9	2/9	0	4/9
	2	1/9	0	0	1/9
	$f_y(y)$	4/9	4/9	1/9	1

林惠玲 陳正倉著 雙葉書廊發行 2004

2號櫃檯有1位顧客的條件下, 1號櫃檯顧客人數的條件機率函數

$$f(x \mid Y=1)=\frac{f(x,Y=1)}{f_y(Y=1)} \qquad x=0,1,2$$

故

$$f(0 \mid Y=1)=\frac{f(0,1)}{f_y(Y=1)}=\frac{2/9}{4/9}=\frac{1}{2}$$

$$f(1 \mid Y=1)=\frac{f(1,1)}{f_y(Y=1)}=\frac{2/9}{4/9}=\frac{1}{2}$$

$$f(2 \mid Y=1)=\frac{f(2,1)}{f_y(Y=1)}=\frac{0}{4/9}=0$$

表8.6 華訊通訊最新型手機的銷售資料

		Y（是否看過型錄）		合計
		0（未看過）	1（看過）	
X（有否購買手機）	0（不買）	600	240	840
	1（購買）	120	240	360
合計		720	480	1,200

表8.7 X 與 Y 的聯合機率分配表

		Y （是否看過型錄）		合計 $f_x(x)$
		0（未看過）	**1**（看過）	
X（是否購買手機）	0（不買）	0.50	0.20	0.7
	1（購買）	0.10	0.20	0.3
合計 $f_y(y)$		0.60	0.40	1.0

林惠玲 陳正倉著 雙葉書廊發行 2004

表8.8 $f(x|y)$的條件機率

$f(x\|y)$	$f(x\|Y=0)$（未看過）	$f(x\|Y=1)$（看過）
$X=0$（不買）	0.833	0.5
$X=1$（購買）	0.167	0.5

表8.9　$f(y|x)$ 的條件機率

$f(y\|x)$	$f(y\|X=0)$（不買）	$f(y\|X=1)$（購買）
$Y=0$（未看過）	0.71	0.33
$Y=1$（看過）	0.29	0.67

林惠玲　陳正倉著 雙葉書廊發行 2004

EX8.4 設袋內有5黑球, 6白球 ,7紅球, 現隨機抽取4球,令X:白球個數, X=0,1,2,3,4; Y:紅球個數, Y=0,1,2,3,4; 試求X,Y的聯合機率函數, 邊際機率函數, 以及有一紅球(Y=1)的條件下X的條件機率函數

Solution:

1.

$$f(x,y)=\frac{C_x^6 C_y^7 C_{4-x-y}^5}{C_4^{18}} \quad x+y\le 4$$

2.X,Y的邊際機率函數

$$f_x(x)=\frac{C_x^6\sum_{y=0}^{4} C_y^7 C_{4-x-y}^5}{C_4^{18}}=\frac{C_x^6 C_{4-x}^{12}}{C_4^{18}}$$

$$f_y(y)=\frac{C_y^7\sum_{x=0}^{4}C_x^6C_{4-x-y}^5}{C_4^{18}}=\frac{C_y^7C_{4-y}^{11}}{C_4^{18}}$$

Y條件下X的條件機率函數

$$f(x\,|\,y)=\frac{f(x,y)}{f(y)}=\frac{C_x^6C_y^7C_{4-x-y}^5}{C_y^7C_{4-y}^{11}}=\frac{C_x^6C_{4-x-y}^5}{C_{4-y}^{11}}$$

$$f(x\,|\,Y=1)=\frac{C_x^6C_1^7C_{3-x}^5}{C_1^7C_2^{11}}=\frac{C_x^6C_{3-x}^5}{C_3^{11}}$$

8.1.4 二元間斷隨機變數的期望值與變異數

- X的期望值

$$E(X)=\sum_x\sum_y xf(x,y)$$
$$=\sum_x x\sum_y f(x,y)=\sum_x xf_x(x)$$

- Y的期望值

$$E(Y)=\sum_x\sum_y yf(x,y)$$
$$=\sum_y y\sum_x f(x,y)=\sum_y yf_y(y)$$

❍ X 的變異數

$$V(X)=\sum_x\sum_y(x-\mu_X)^2 f(x,y)$$
$$=\sum_x(x-\mu_X)^2 f_x(x)$$
$$=\sum_x x^2 f_x(x)-\mu_X^2$$

❍ Y 的變異數

$$V(Y)=\sum_x\sum_y(y-\mu_Y)^2 f(x,y)$$
$$=\sum_y(y-\mu_Y)^2 f_y(y)$$
$$=\sum_y y^2 f_y(y)-\mu_Y^2$$

EX8.5 華訊通訊的客戶是否看過型錄(Y)與客戶是否購買手機(X)的期望值與標準差各爲多少?

		Y（是否看過型錄）		合計
		0（未看過）	**1**（看過）	
X（有否購買手機）	0（不買）	600	240	840
	1（購買）	120	240	360
合計		720	480	1,200

		Y （是否看過型錄）		合計 $f_x(x)$
		0（未看過）	**1**（看過）	
X（是否購買手機）	0（不買）	0.50	0.20	0.7
	1（購買）	0.10	0.20	0.3
合計 $f_y(y)$		0.60	0.40	1.0

Solution:

$$E(X)=\sum x f_x(x)=0\times 0.7+1\times 0.3=0.3$$

$$E(Y)=\sum y f_y(y)=0\times 0.7+1\times 0.4=0.4$$

$$V(X)=\sum (x-\mu_X)^2 f_x(x)=\sum x^2 f_x(x)-\mu_X^2$$
$$=[0^2\times 0.7+1^2\times 0.3]-(0.3)^2=0.21$$

$$V(Y)=\sum (y-\mu_Y)^2 f_y(y)=\sum y^2 f_y(y)-\mu_Y^2$$
$$=[0^2\times 0.6+1^2\times 0.4]-(0.4)^2=0.24$$

8.1.5 X與Y的條件平均數與變異數

- X 的條件平均數與變異數

$$E(X|Y=y_j)=\sum_x xf(x|Y=y_j)$$

$$V(X|Y=y_j)=E\{[X-E(X|Y=y_j)]^2|Y=y_j\}$$
$$=E(X^2|Y=y_j)-[E(X|Y=y_j)]^2$$

- Y 的條件平均數與變異數

$$E(Y|X=x_i)=\sum_y yf(y|X=x_i)$$

$$V(Y|X=x_i)=E[(Y-E(Y|X=x_i))^2|X=x_i]$$
$$=E(Y^2|X=x_i)-[E(Y|X=x_i)]^2$$

EX8.6 例8.3中, 顧客在看過型錄後, 是否購買手機(X)的條件期望值,即E(X|Y=1)與變異數V(X|Y=1)為何?

	$f(x \mid Y=1)$
$X=0$	0.5
$X=1$	0.5

林惠玲　陳正倉著 雙葉書廊發行 2004

Solution:

E(X|Y=1)=Σxf(x|Y=1)=0*f(0|Y=1)+1*f(1|Y=1)

=0*0.5+1*0.5=0.5

$$V(X\,|\,Y=1)=E(X^2\,|\,Y=1)-[E(X\,|\,Y=1)]^2$$

$$=\sum x^2\cdot f(x\,|\,Y=1)-(0.5)^2$$

$$=0.5\text{-}(0.5)^2=0.25$$

標準差

$$\sigma_{X|Y=1}=\sqrt{0.25}=0.5$$

未看過型錄下是否購買的條件平均數與變異數

E(X|Y=0)=0.167

促銷有效

8.2 二元連續隨機變數

8.2.1 二元機率密度函數

設 X, Y 為二元連續隨機變數，$a \le x \le b$，$c \le y \le d$，若 $f(x,y)$ 滿足下列二條件：

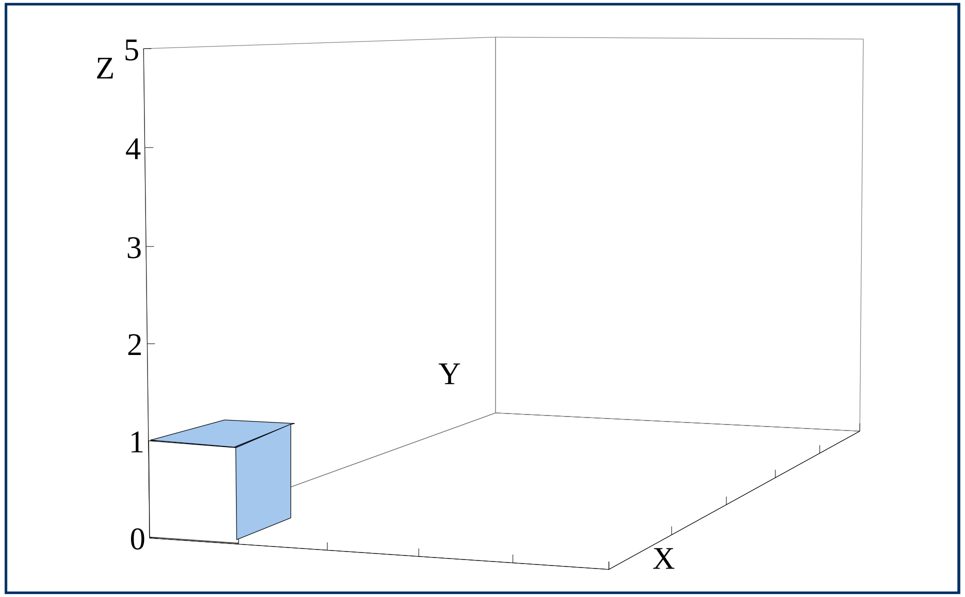
① $f(x,y) \ge 0$

② $\int_a^b \int_c^d f(x,y)\,dy\,dx = 1$

則 $f(x,y)$ 為一個二元機率密度函數。

EX8.7 設f(x,y)=1, 0≦x ≦ 1; 0 ≦ y ≦ 1; f(x,y)=0, 其他. 因 f(x,y) ≧0, 且 $\int_x \int_y f(x,y)\,dxdy = 1$, 因此f(x,y)爲一二元機率密度函數, 如下圖, 求 1. P(0≦x ≦ 0.5; 0≦y≦0.5)
2. F(x,y) 3. F(1,0.5)

圖8.2 二元機率密度函數的例子



Solution:

1．$P(0 \le x \le 0.5; 0 \le y \le 0.5) = \int_0^{0.5}\int_0^{0.5} 1 dxdy$

$$= \int_0^{0.5}\int_0^{0.5} \mathrm{dx} = \int_0^{0.5} (0.5-0)dx$$

$$= 0.5x\Big|_0^{0.5} dx = 0.25 - 0 = 0.25$$

2.

$$F(x,y) = P(X \le x; Y \le y) = \int_0^x\int_0^y 1 dtds = \int_0^x t\Big|_0^y ds = ys\Big|_0^x = yx$$

3. 根據F(x,y)=yx, 可求F(1,0.5)=1*0.5=0.5, 亦即x累積至1, y累積至0.5, 發生的機率為0.5

林惠玲 陳正倉著 雙葉書廊發行 2004

8.2.2 邊際機率密度函數

❍ X與Y的邊際機率密度函數

X邊際機率密度函數 $f_x(x)$為：

$$f_x(x)=\int_y f(x,y)dy$$

它必須滿足機率的二條件① $f_x(x)\geq 0$ ② $\int_x f_x(x)dx=1$。

Y邊際機率密度函數 $f_y(y)$為：

$$f_y(y)=\int_x f(x,y)dx$$

它必須滿足機率的二條件① $f_y(y)\geq 0$ ② $\int_y f_y(y)dy=1$。

EX8.8 麥當勞想了解中壢店與楊梅店來得速(drive in)買漢堡的情形. 現隨機抽取某一天, 令X與Y分別代表中壢店與楊梅店顧客利用汽車駛入的比例, 且假設X與Y的二元機率密度函數爲 f(x, y)= 2/3 (x+2y) 0≦x ≦ 1; 0 ≦ y ≦ 1

=0 其他

1. 求X的邊際機率密度函數
2. 求Y的邊際機率密度函數
3. 中壢店顧客利用汽車駛入的比例大於1/2的機率

Solution:

1.

$$f_x(x) = \int_0^1 \frac{2}{3}(x+2y)dy = \frac{2}{3}(xy+y^2)\,|_0^1 = \frac{2}{3}(x+1)$$

林惠玲 陳正倉著 雙葉書廊發行 2004

2.

$$f_y(y)=\int_0^1 \frac{2}{3}(x+2y)dx=\frac{2}{3}(\frac{1}{2}x^2+2xy)\Big|_0^1=\frac{2}{3}(\frac{1}{2}+2y)=\frac{1}{3}+\frac{4}{3}y$$

3.

$$P(x>\frac{1}{2})=\int_{\frac{1}{2}}^1 \frac{2}{3}(x+1)dx=\frac{2}{3}(\frac{1}{2}x^2+x)\Big|_{\frac{1}{2}}^1$$

$$=\frac{2}{3}(\frac{1}{2}+1)-\left[\frac{2}{3}(\frac{1}{2}\cdot\frac{1}{4}+\frac{1}{2})\right]=1-\frac{2}{3}\times\frac{5}{8}=\frac{7}{12}$$

林惠玲 陳正倉著 雙葉書廊發行 2004

8.2.3 條件率密度函數

- **X 與 Y 的條件機率密度函數**

$Y=y$條件下，X的條件機率密度函數$f(x\,|\,y)$為：

$$f(x\,|\,y)=\frac{f(x,y)}{f_y(y)}$$

它必須滿足機率的二條件 ① $f(x\,|\,y)\geq 0$ ② $\int_x f(x\,|\,y)dx=1$。

$X=x$條件下，Y的條件機率密度函數$f(y\,|\,x)$為：

$$f(y\,|\,x)=\frac{f(x,y)}{f_x(x)}$$

它必須滿足機率的二條件 ① $f(y\,|\,x)\geq 0$ ② $\int_y f(y\,|\,x)dy=1$。

林惠玲　陳正倉著 雙葉書廊發行 2004

EX8.9 大正電子想了解電腦電源供應器壽命與硬碟機壽命的關係. 設X與Y分別代表電源供應器壽命與硬碟機的壽命, 且假設已知X與Y的聯合機率密度函數爲

$$f(x,y)=e^{-(x+y)} \quad ,x>0,y>0)$$
$$=0 \qquad 其他$$

1. 求電源供應器壽命2年的條件下, 硬碟壽命的機率函數 f (y|x=2)
2. 硬碟機壽命超過2年的條件下, 電源供應器的壽命機率函數 f(x|y>2)
3. 硬碟壽命超過2的條件下, 電源供應器壽命小於1年的機率 f(0≦x<1|y<2)

Solution:

1. $$f_x(x)=\int_0^\infty e^{-(x+y)}dy=e^{-(x+y)}\Big|_0^\infty=0-(-e^{-x})=e^{-x}$$

$$f(y\,|\,x=2)=\frac{f(y,x=2)}{f_x(x=2)}=\frac{e^{-(2+y)}}{e^{-2}}=e^{-y}$$

2. $$f_y(y)=\int_0^\infty e^{-(x+y)}dx=e^{-(x+y)}\Big|_0^\infty=0-(-e^{-y})=e^{-y}$$

$$P(y>2)=\int_2^\infty e^{-y}dy=-e^{-y}\Big|_2^\infty=0-(-e^{-2})=e^{-2}$$

$$P(x,y>2)=\int_2^\infty e^{-(x+y)}dy=-e^{-(x+y)}\Big|_2^\infty=e^{-(x+2)}$$

林惠玲 陳正倉著 雙葉書廊發行 2004

$$f(x \mid y>2)=\frac{f(x,y>2)}{P(y>2)}=\frac{e^{-(x+2)}}{e^{-2}}=e^{-x}$$

3.

$$f(x \mid y)=\frac{f(x,y)}{f_y(y)}=\frac{e^{-(x+y)}}{e^{-y}}=e^{-x}$$

$$P(0 \le x \le 1 \mid y=2)=\int_0^1 f(x \mid y=2)dx$$

$$=\int_0^1 e^{-x}dx=-e^{-x}\Big|_0^1=-e^{-1}+1$$

$$=-0.3679+1=0.632$$

8.2.4 二元連續隨機變數的期望值

- X與Y的期望值

$$E(X)=\int_x\int_y xf(x,y)dydx=\int_x xf_x(x)dx$$

$$E(Y)=\int_x\int_y yf(x,y)dydx=\int_y yf_y(y)dy$$

二元連續隨機變數的變異數

- X 與 Y 的變異數

$$V(X)=\int_x\int_y(x-\mu_x)^2 f(x,y)dydx=E(X^2)-(E(X))^2$$

$$V(Y)=\int_x\int_y(y-\mu_y)^2 f(x,y)dydx=E(Y^2)-(E(Y))^2$$

EX8.10 衛生署想了解男女性別運動與休閒的比例是否有不同. 設X與Y分別為男性與女性一天運動時間佔休閒時間比例, 且機率密度函數如下

$f(x,y)=8xy \quad 0 \leqq x \leqq 1;\ 0 \leqq y \leqq x$

$= 0$ 其他

求 E(X), V(X), E(Y), V(Y)

Solution:

$$E(X)=\int_0^1\int_0^x x8xydydx=\int_0^1 \frac{1}{2}8x^2y^2 \Big|_0^x d=\int_0^1 4x^4dx=\frac{4}{5}x^5 \Big|_0^1=\frac{4}{5} \quad \#$$

$$E(X^2)=\int_0^1\int_0^x x^2 8xydydx=\int_0^1 \frac{1}{2}8x^3y^2 \Big|_0^x dx=\int_0^1 4x^5 \Big|_0^1=\frac{4}{6}x^6 \Big|_0^1=\frac{2}{3}$$

林惠玲 陳正倉著 雙葉書廊發行 2004

因此

$$V(X)=E(X^2)-[E(X)]^2=\frac{2}{3}-(\frac{4}{5})^2=\frac{2}{75}$$

$$E(Y)=\int_0^1\int_0^x y8xydydx=\int_0^1\frac{8}{3}8xy^3\Big|_0^x dx=\int_0^1\frac{8}{3}x^4dx=\frac{8}{15}x^5\Big|_0^1=\frac{5}{18}$$

$$E(Y^2)=\int_0^1\int_0^x y^2 8xydydx=\int_0^1\frac{8}{4}xy^4\Big|_0^x dx=\int_0^1 2x^5\Big|_0^1=\frac{1}{3}x^6\Big|_0^1=\frac{1}{3}$$

$$V(Y)=E(Y^2)-[E(Y)]^2=\frac{1}{3}-(\frac{8}{15})^2=\frac{11}{225}$$

林惠玲 陳正倉著 雙葉書廊發行 2004

8.2.5 X與Y的條件平均數與變異數

- X 的條件平均數與變異數

$$E(X\mid y)=\int_x x f(x\mid y)dx$$

$$\begin{aligned}V(X\mid y)&=E\left[(x-E(X\mid y))^2\mid y\right]\\&=\int_x (x-E(X\mid y))^2 f(x\mid y)dx\\&=\int_x x^2 f(x\mid y)dx-\left[E(X\mid y)\right]^2\end{aligned}$$

二元連續隨機變數的條件平均數與變異數

- Y 的條件平均數與變異數

$$E(Y|x)=\int_y y f(y|x)dy$$

$$\begin{aligned}V(Y|x)&=E\left[(y-E(Y|x))^2|x\right]\\&=\int_y (y-E(Y|x))^2 f(y|x)dy\\&=\int_y y^2 f(y|x)dy-\left[E(Y|x)\right]^2\end{aligned}$$

EX8.11 國光加油站每天一早的存油量為隨機變數Y, 每天賣出的油量為隨機變數X, 每天的存油量在0~1萬公升, 且在白天時不會再進油量, 即 X≦Y, 假設已知X與Y的聯合機率函數為

f(x,y)=2 　0<x<y, 0<y<1

=0 　其他

1. 試求 f(x|y)　2. $P(\frac{1}{4}<x<\frac{1}{2}\mid y=\frac{3}{4})$　3. E(X|Y)

4. V(X|Y)

Solution:

1. $$f_y(y)=\int_0^y 2dx = 2x\,|_0^y = 2y \qquad ,0<y<1$$

$$f(x|y)=\frac{f(x,y)}{f_y(y)}=\frac{2}{2y}=\frac{1}{y} \qquad ,0<x<y<1$$

2.

$$P(\frac{1}{4} < x < \frac{1}{2} \mid y = \frac{3}{4}) = \int_{\frac{1}{4}}^{\frac{1}{2}} f(x \mid y = \frac{3}{4})dx$$

$$= \int_{\frac{1}{4}}^{\frac{1}{2}} \frac{1}{0.75} dx = \frac{4}{3} x \Big|_{\frac{1}{4}}^{\frac{1}{2}} = \frac{4}{3} \times \frac{1}{2} - \frac{4}{3} \times \frac{1}{4} = \frac{1}{3}$$

3.

$$E(X \mid Y) = \int_0^y x f(x \mid y) dx = \int_0^y x \cdot \frac{1}{y} dx = \frac{1}{2}\frac{1}{y} x^2 \Big|_0^y = \frac{1}{2} y$$

若Y=0.75, 則賣出的油量爲E(X|Y=0.75)=0.375萬公升

4.

$$V(X|Y)=E(X^2|Y)-[E(X|Y)]^2$$

$$=\int_0^y x^2 f(x|y)dy-(\frac{1}{2}y)^2$$

$$=\int_0^y x^2\frac{1}{y}dx-\frac{1}{4}y^2=\frac{1}{3}\frac{1}{y}x^3|_0^y-\frac{1}{4}y^2$$

$$=\frac{1}{3}y^2-\frac{1}{4}y^2=\frac{1}{12}y^2$$

Y條件下X的變異數爲(1/12)y^2爲Y的函數,若Y=0.75

V(X|Y=(3/4))=(1/12)*(3/4)^2= 3/64

表8.11 電腦產品壽命的條件機率密度函數

產品壽命（Y）	顧客謹慎程度（X）		
	謹慎（1）	中等（2）	粗心（3）
1 年以下	0.10	0.20	0.40
1~3 年	0.15	0.30	0.20
3~5 年	0.30	0.10	0.10

資料來源：新軒電腦維修部

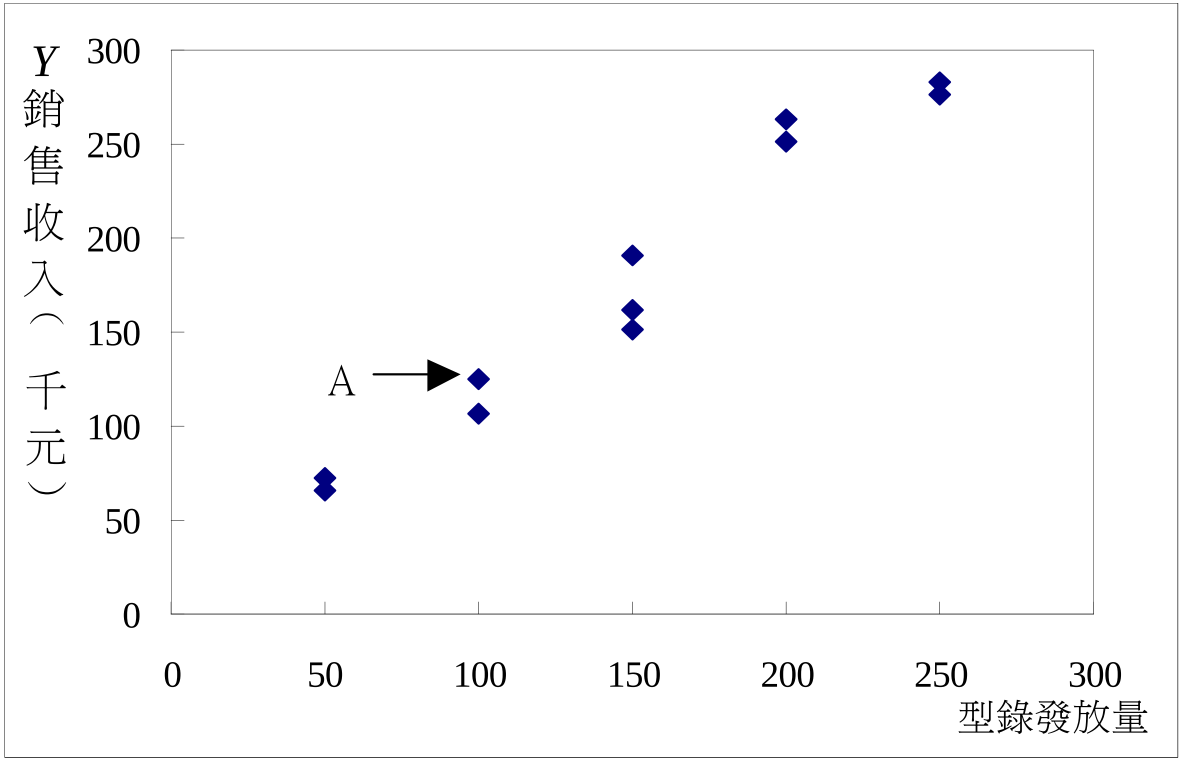
8.3 兩變數間的關係

用來表示兩個變數間關係的統計表稱為二項表
用來表示兩個變數資料散佈情形的統計圖稱為散佈圖(scatter diagrams)

表8.12 型錄發放量與銷售收入

	A	B	C
1	星期	型錄發放量 x	銷售收入 y
2	1	100	125,000
3	2	250	283,000
4	3	50	72,500
5	4	150	191,000
6	5	200	251,000
7	6	50	65,500
8	7	250	276,000
9	8	150	151,000
10	9	200	263,000
11	10	100	106,000
12	11	150	162,000
13	12	200	263,000

圖8.3 型錄發放量與銷售收入的散佈圖



X變數與Y變數間有何關係?? 獨立與不獨立

8.3.1 兩變數獨立

兩變數獨立的條件

設 X, Y 爲二元隨機變數，若 X 與 Y 之值均滿足下列任一條件，則 X, Y 獨立。

① $f(x,y) = f_x(x) \cdot f_y(y)$

② $f(x \mid y) = f_x(x)$

③ $f(y \mid x) = f_y(y)$

1. X,Y的聯合機率等於X與Y的邊際機率的乘積
2. Y條件下出現X的條件機率等於X邊際機率
3. X條件下出現Y的條件機率等於Y的邊際機率

EX8.14 251家多層次傳銷公司按地區及有無福利措施區分的統計表如下. 問傳銷公司之有無福利措施與其所在地是否有關?

表8.13　多層次傳銷之地區別與福利措施統計表

X（公司所在地區）	*Y*（福利措施） 0（無）	1（有）	合計
1（北部）	68	78	146
2（中部）	33	24	57
3（南部）	20	28	48
合計	121	130	251

資料來源：《中華民國91年台灣地區多層次傳銷事業經營概況調查報告》，行政院公平交易委員會，2003年6月。

表8.14 X與Y的聯合機率分配表

X（公司所在地區）	Y（有無福利措施）		合 計 $f_x(x)$
	0（無）	**1**（有）	
1（北 部）	0.27	0.31	0.58
2（中 部）	0.13	0.10	0.23
3（南 部）	0.08	0.11	0.19
合 計 $f_y(y)$	0.48	0.52	1.00

沒有福利措施的傳銷公司其位在北部地區的機率為
$f(1,0)=0.27$ $f_y(0)=0.48$

故 $f(X=1|Y=0)= f(1,0)/ f_y(0) = 0.27/0.48 =0.56 \neq f_x(1)=0.58$
表示已知沒有福利措施的條件下，該公司位於北部地區的機率，此一機率小於位於北部地區的機率$f_x(1)=0.58$，故有關係

8.3.2 兩變數不獨立

不獨立

→ 有關係, 最常見爲線性關係

亦即 X變動時y跟著變動, 可爲同方向或反方向

→ X, Y間線性共變關係的兩種統計測量數

共變數 (covariance)

相關係數 (correlation coefficient) 教好 較常用

林惠玲 陳正倉著 雙葉書廊發行 2004

8.3.3 共變數

公式 $$Cov(X,Y)=E[(X-\mu_X)(Y-\mu_Y)]$$

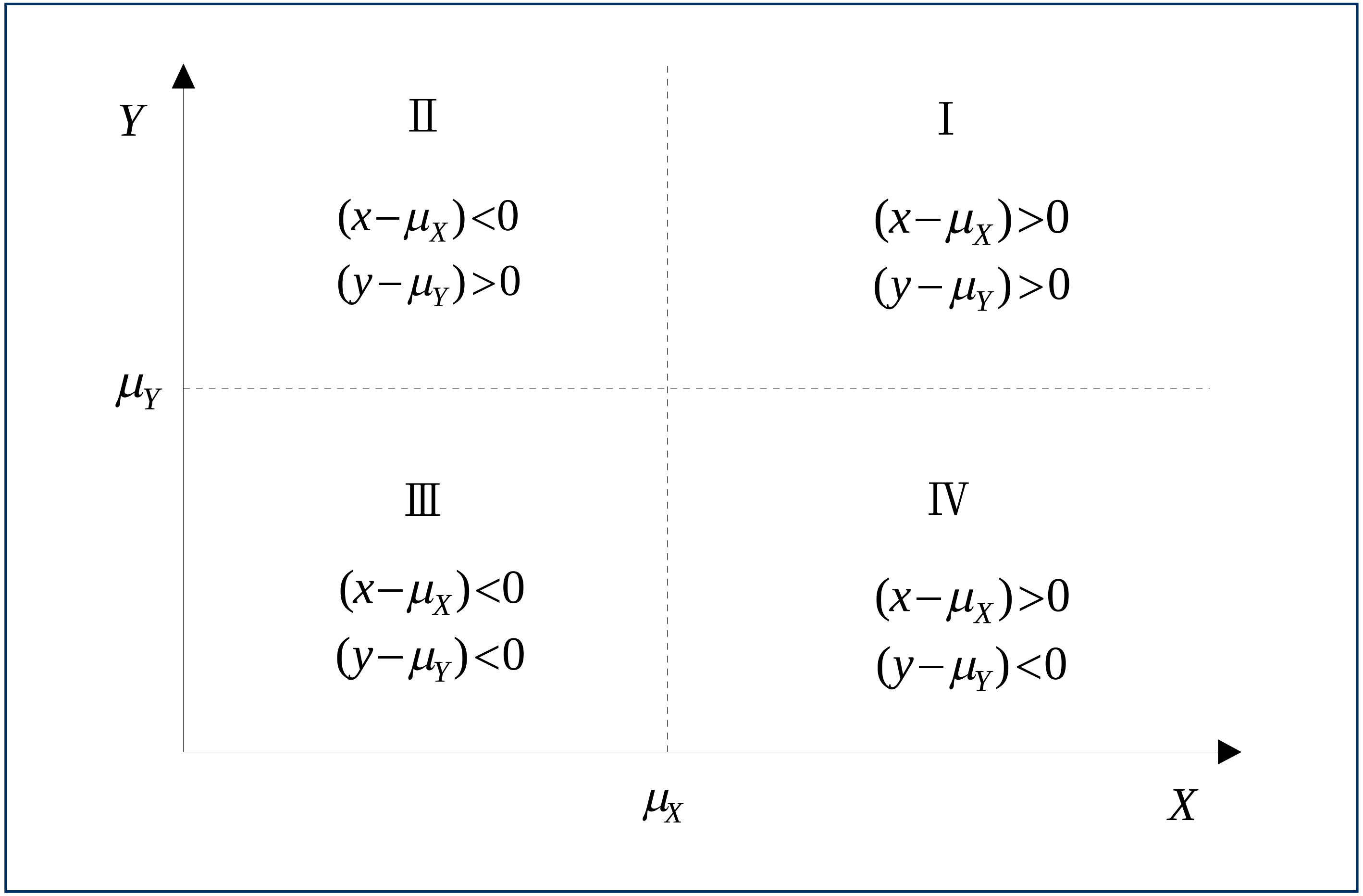
$$\begin{aligned}Cov(X,Y)&=E[(X-\mu_X)(Y-\mu_Y)]\\&=\sum_x\sum_y(x-\mu_X)(y-\mu_Y)\cdot f(x,y)\\&=\sum_x\sum_y(xy)f(x,y)-\left[\sum_x xf_x(x)\right]\left[\sum_y yf_y(y)\right]\\&=E(XY)-E(X)E(Y)\end{aligned}$$

X,Y為連續變數

$$\begin{aligned}Cov(X,Y)&=\int_x\int_y(x-\mu_X)(y-\mu_Y)f(x,y)dxdy\\&=\int_x\int_y xyf(x,y)dydx-\mu_X\mu_Y\\&=E(XY)-E(X)E(Y)\end{aligned}$$

林惠玲 陳正倉著 雙葉書廊發行 2004

圖8.4 $Cov(X,Y)$的符號



說明:

I. III象限 + II, IV象限 -- 散落？

→ 共變數符號可衡量X,Y間共變的方向, 絕對值大小衡量共變程度的大小

◎ 共變數符號的說明

1. Cov(X,Y)>0 X,Y為正方向共變關係

2. Cov (X,Y) <0 X, Y為反方向共變關係

3. Cov (X,Y)=0 X, Y無共變關係 (但不能下結論說X, Y無關係或獨立, 可以為拋物線或圓形關係等)

圖8.5 正向共變

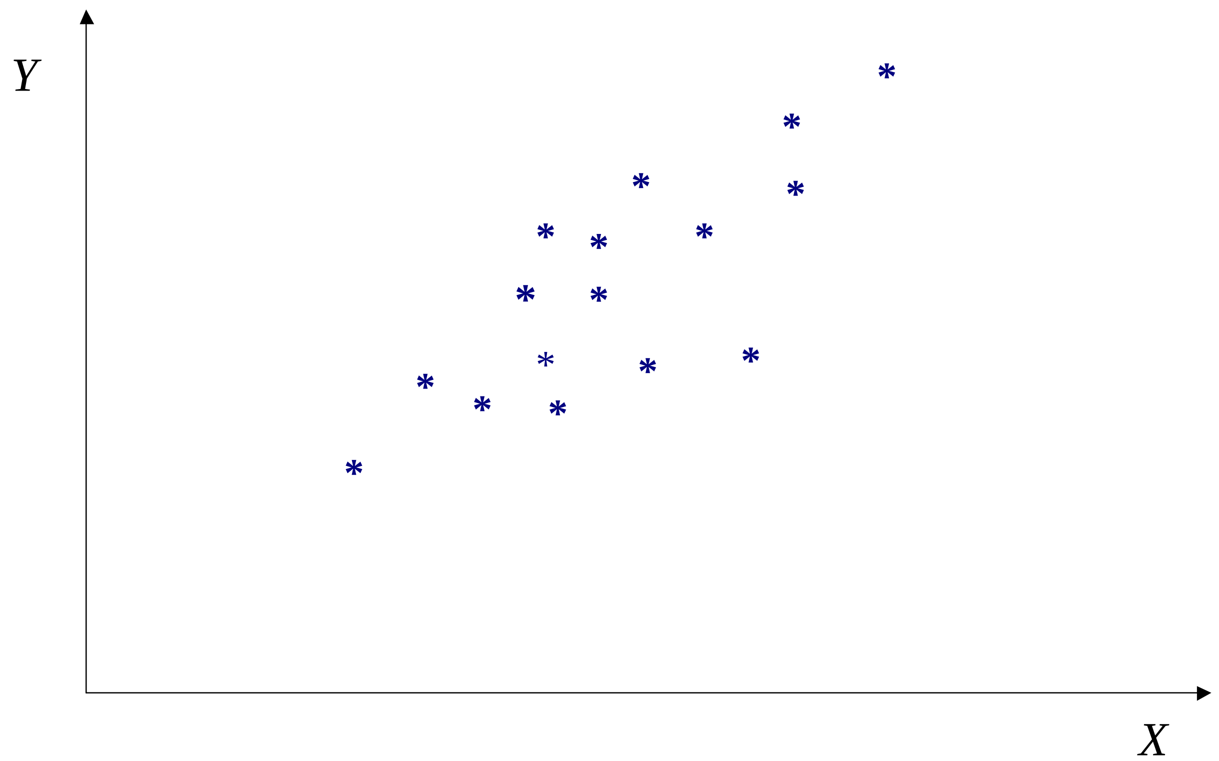


圖8.6 反向共變

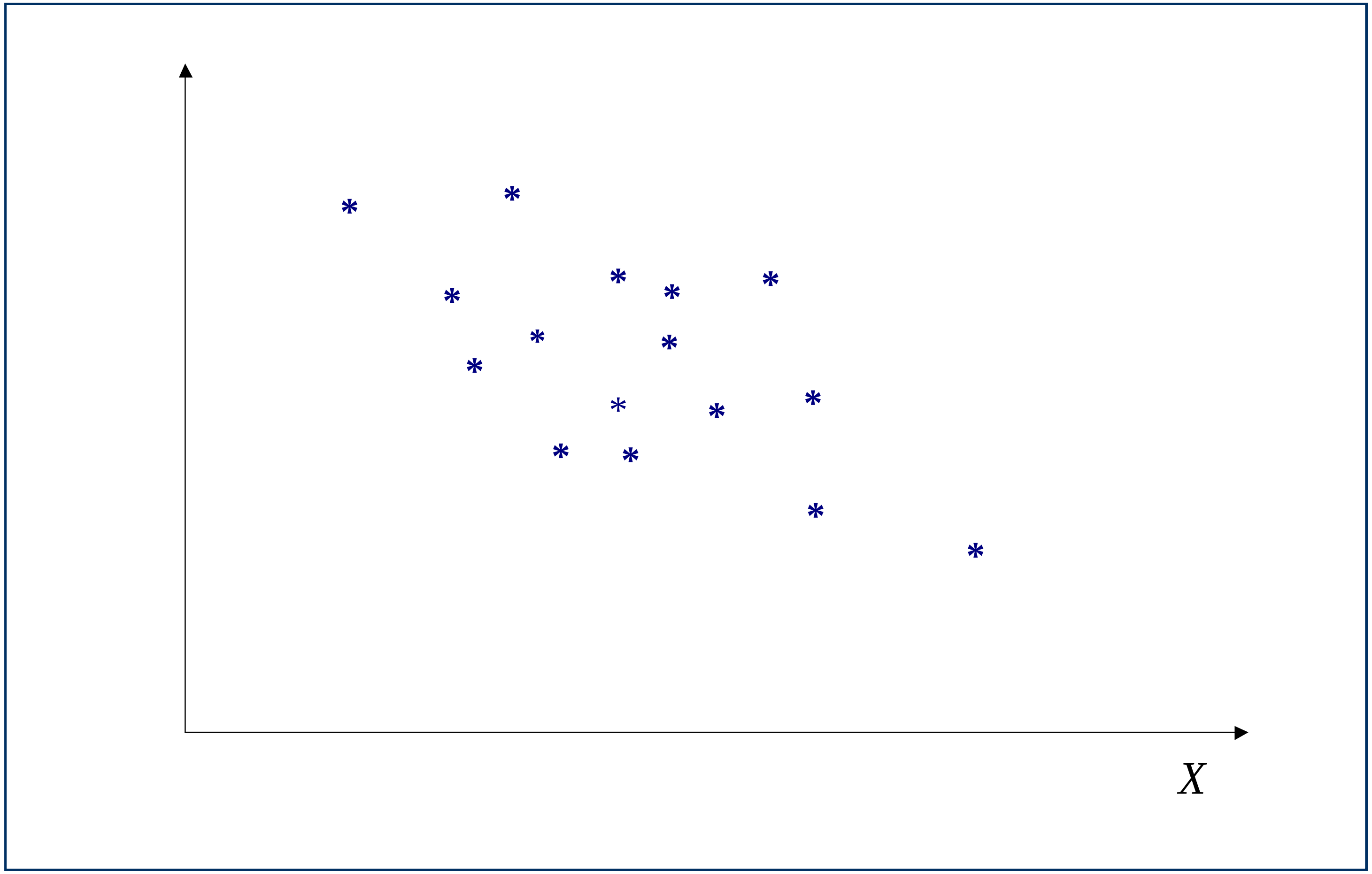
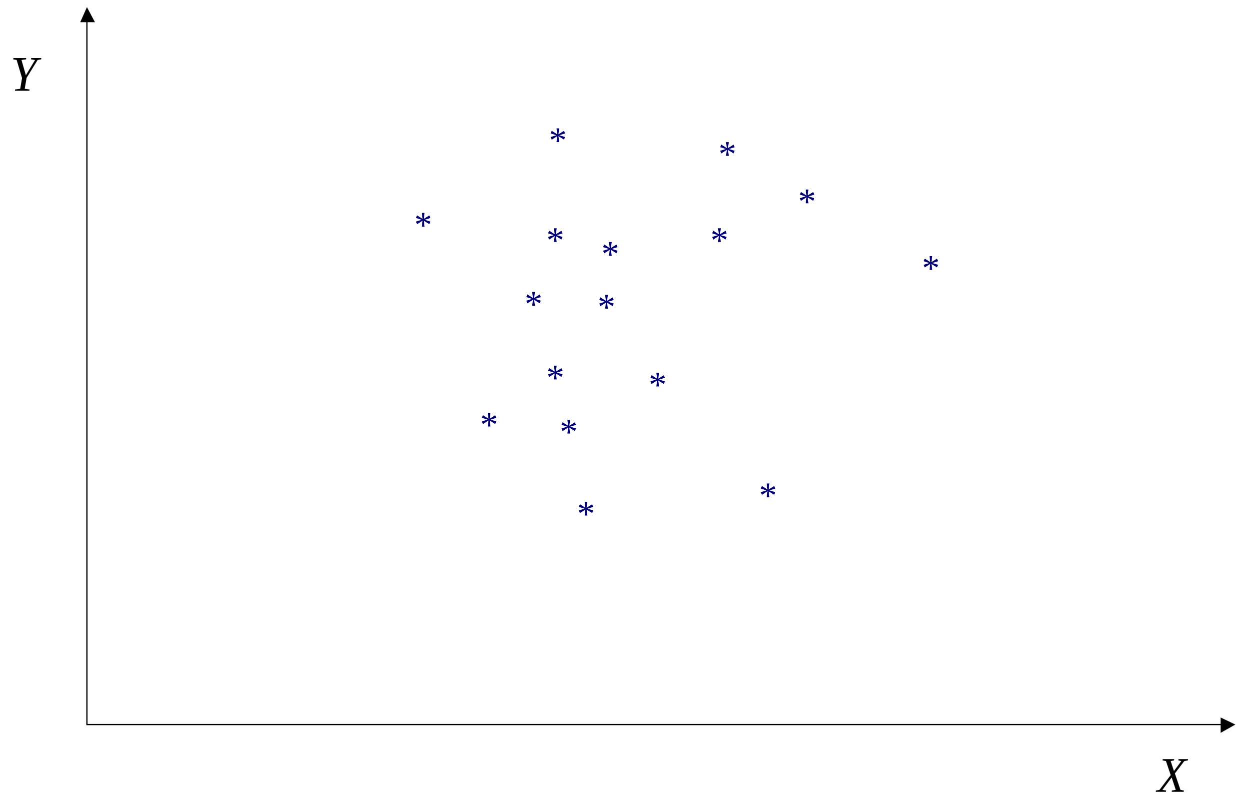


圖8.7 無關係



林惠玲 陳正倉著 雙葉書廊發行 2004

◎ 共變數的優缺點

1. 可看出共變的方向和程度

2. $-\infty < \mathrm{Cov}(X,Y) < \infty$, 其值域無限, 不易根據Cov(X,Y)的大小來判斷其相關程度

3. 易受衡量單位的影響

4. 具有雙重的衡量單位 (同時有X與Y的衡量單位不易解釋)

林惠玲　陳正倉著 雙葉書廊發行 2004

◎ 計算共變數的常用定理

1. Cov(X,Y)=Cov(Y,X)
2. Cov(X,X)=V(X)
3. Cov(X,C)=0
4. Cov(aX, bY)= ab Cov(X,Y)
5. Cov (X+c, Y+d) = Cov (X,Y)
6. 隨機變數線性函數共變數的定理

   a. 設aX+c, bY+d 為X,Y 隨機變數的線性函數, 則

   $$\mathrm{Cov}\,(aX+c,\ bY+d) = ab\,\mathrm{Cov}\,(X,Y)$$

   b.設aX+bY, cX+dY 為X,Y 隨機變數的線性函數, 則

   $$\mathrm{Cov}\,(aX+bY,\ cX+dY) = acV(X)+(ad+bc)\mathrm{Cov}\,(X,Y) + bdV(Y)$$

EX8.15 華訊通訊的型錄廣告與客戶購買手機有關, 然而看過型錄(Y)與購買手機 (X) 兩者呈現何種關係? 其共變數爲何?

Solution:

$$\text{Cov}(X,Y)=E(XY) - E(X)\,E(Y)$$
$$= \Sigma\Sigma xy\, f(x,y) - (\Sigma x f(x))\,(\Sigma y\, f(y))$$
$$= 0 * 0 * 0.5 + 0 * 1 * 0.2 + 1 * 0 * 0.1 + 1 * 1 * 0.2 - 0.3 * 0.4$$
$$= 0.08$$

共變數爲0.08, 其值爲正. 因此看過型錄 (Y)與購買手機 (X) 兩者呈正向的共線關係

EX 8.16 在例8.11中, (1) 問汽油的存量與銷售量是有關呢? 還是獨立? (2) 請求算 Cov (X,Y)

Solution:

(1) 爲驗證是否獨立, 先求算邊際機率 $f_x(x)$, $f_y(y)$

$$f_x(x)=\int_x^1 2dy=2y|_x^1=2(1-x),\ 0<x<1$$

$$f_y(y)=\int_0^y 2dx=2x|_0^y=2y,\ 0<y<1$$

$$f_x(x)\cdot f_y(y)=2(1-x)\times 2y\neq f(x,y)=2$$

因此, X與Y不獨立

(2) 先求算

$$E(XY)=\int_0^1\int_0^Y xy\cdot 2dxdy=\int_0^1 x^2y|_0^y\ dy=\int_0^1 y^3dy=\frac{1}{4}y^4\ |_0^1=\frac{1}{4}$$

$$E(X)=\int_0^1\int_0^Y x\cdot 2dxdy=\int_0^1 x^2\ |_0^y\ dy=\int_0^1 y^2dy=\frac{1}{3}y^3\ |_0^1=\frac{1}{3}$$

$$E(Y)=\int_0^1\int_0^Y y\cdot 2dxdy=\int_0^1 2yx|_0^y\ dy=\int_0^1 2y^2dy=\frac{2}{3}y^3\ |_0^1=\frac{2}{3}$$

Cov (X,Y)=E(XY) – E(X)E(Y)

= (1/4)-(1/3)(2/3) = (1/4)-(2/9)=1/36

8.3.4 相關係數

相關係數

$$\rho_{XY} = \frac{Cov(X,Y)}{\sigma_X \sigma_Y}$$

式中：$Cov(X,Y)$ 爲共變數，σ_X，σ_Y 爲標準差。

林惠玲 陳正倉著 雙葉書廊發行 2004

◎ 相關係數的性質

1. 相關係數的數值可證明是介於 -1與 +1之間
   a. ρ_{XY} =+1時, X,Y爲具正的完全直線相關
   b. ρ_{XY} =0時, X,Y無線性相關

2. $|\rho_{XY}|$ 愈大表示X,Y的線性相關程度愈大

   一般而言: >0.7 高度相關
   0.3 ~ 0.7 中度相關
   <0.3 低度相關

圖8.8 相關係數ρ_{XY}= 1

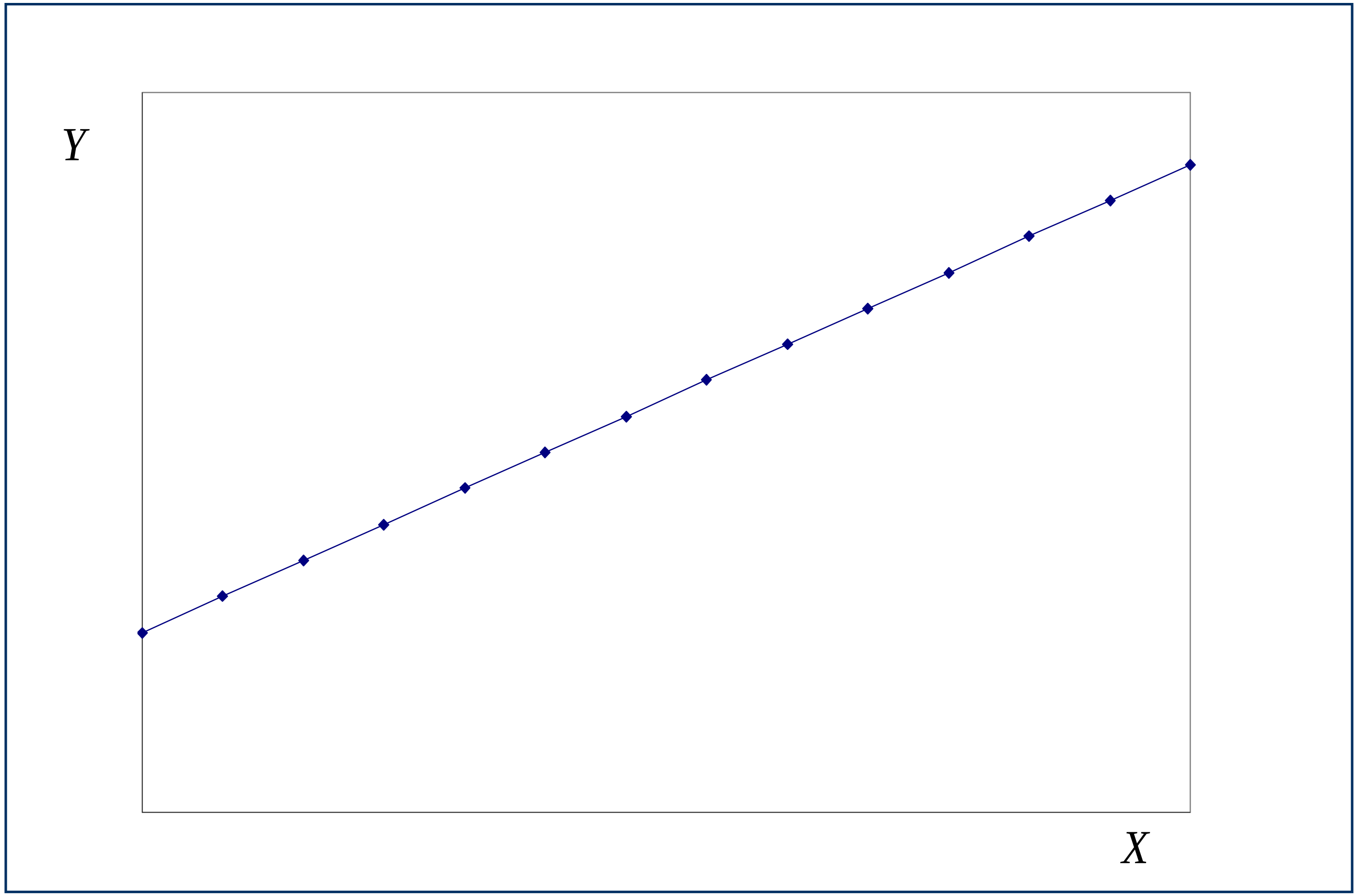
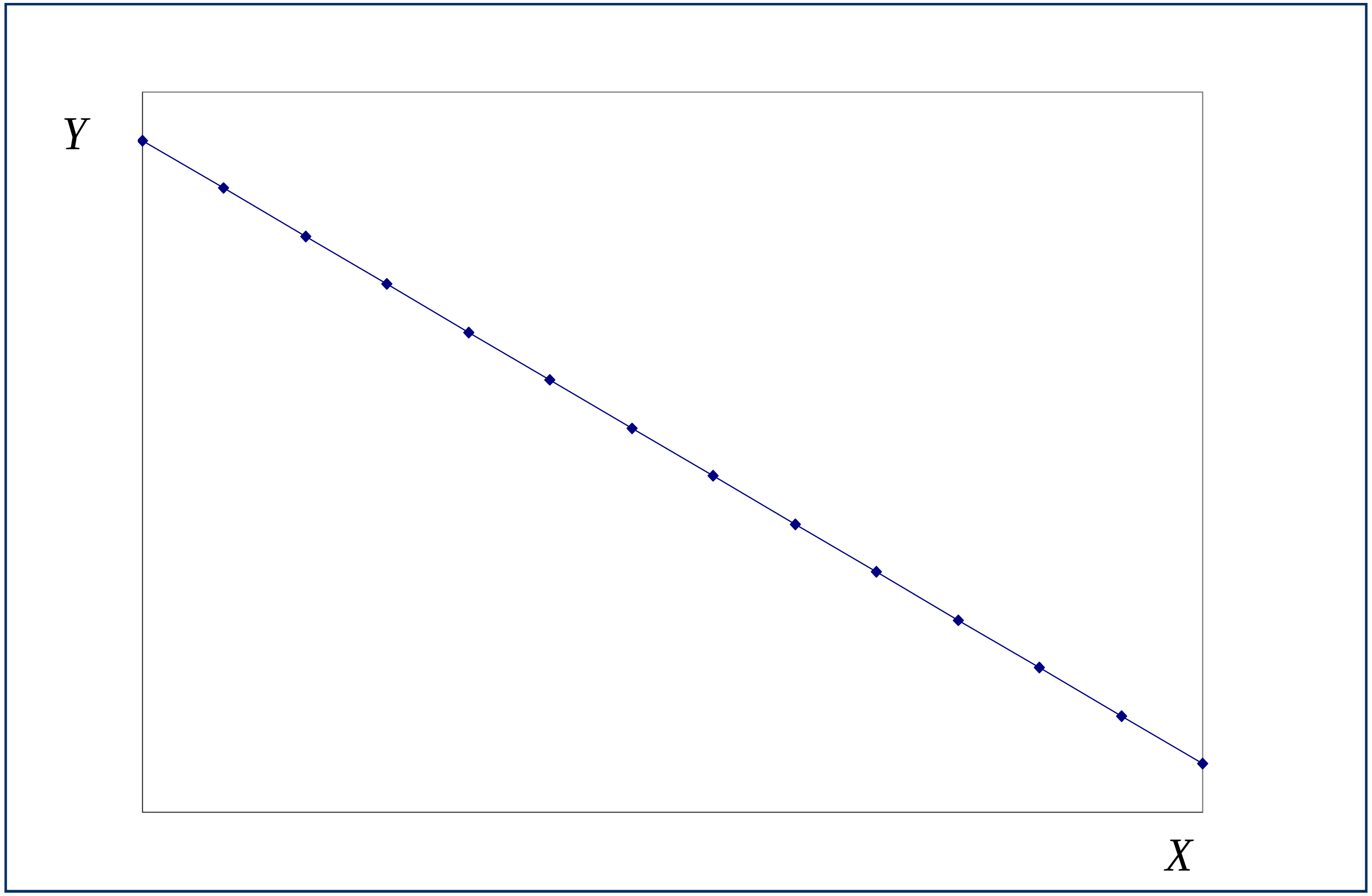


圖8.9 相關係數ρ_{XY}= -1



林惠玲 陳正倉著 雙葉書廊發行 2004

圖8.10 X，Y 無關

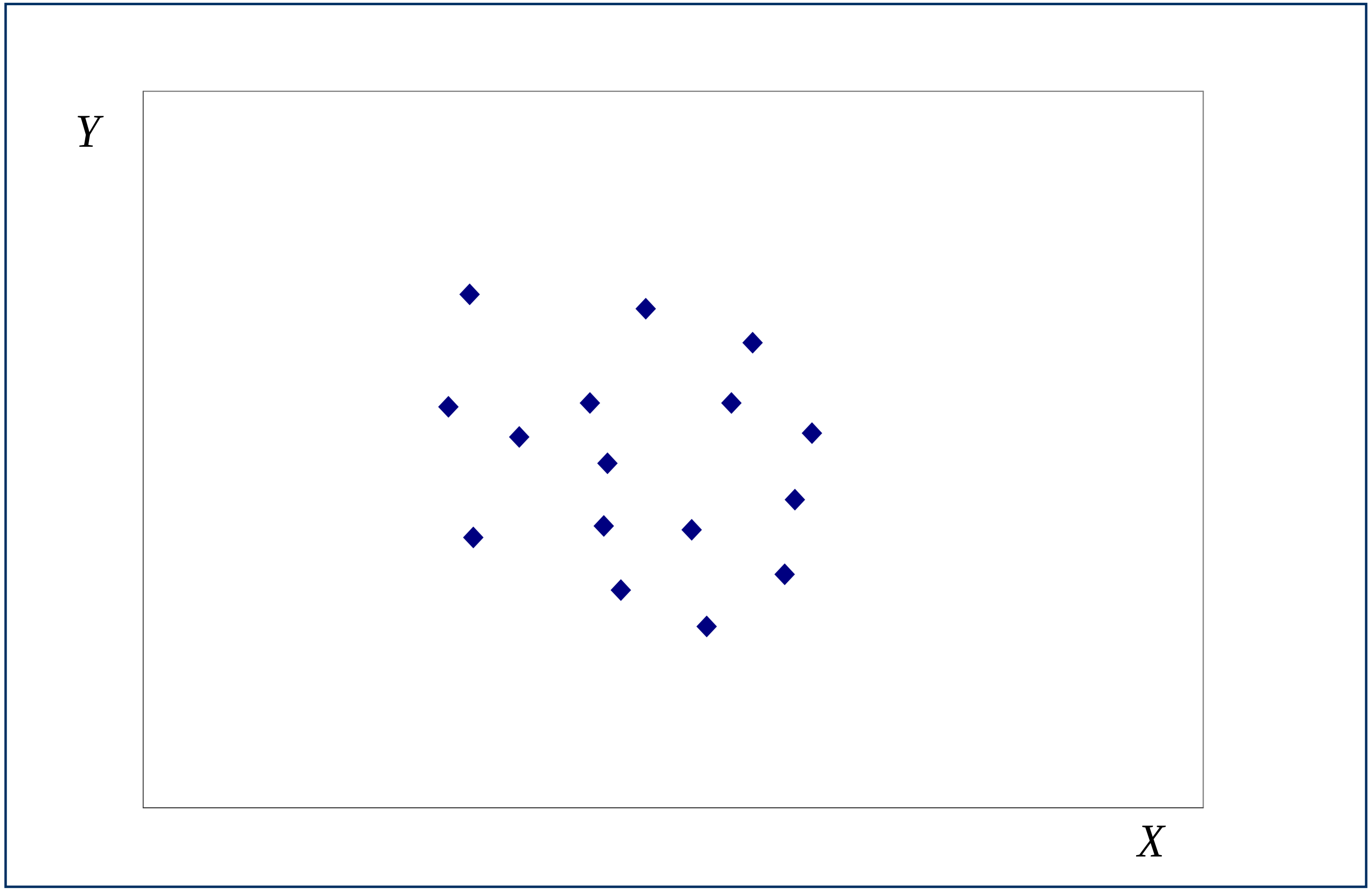


圖8.11　*X*，*Y* 為拋物線關係

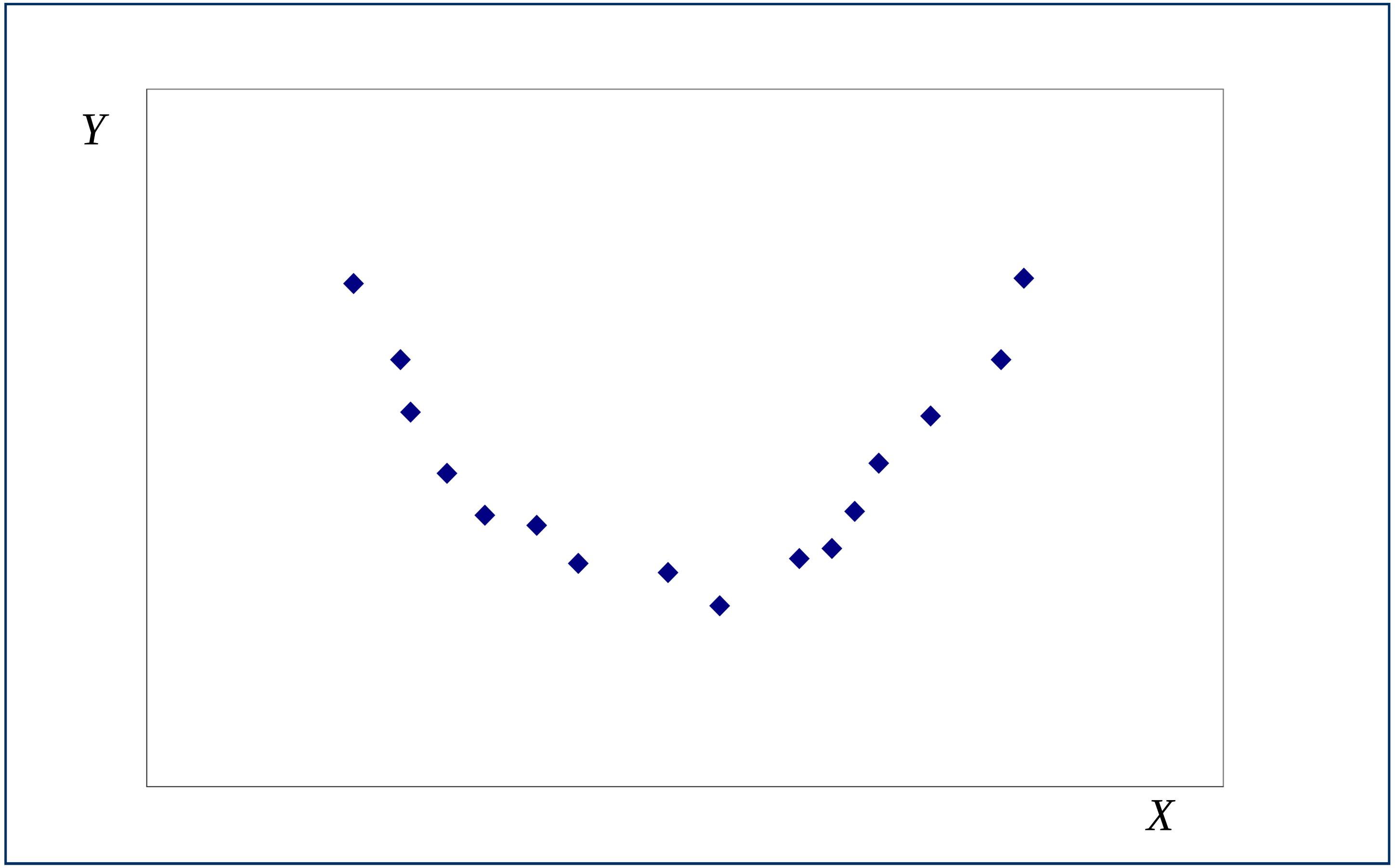


圖8.12 X，Y 為圓形關係

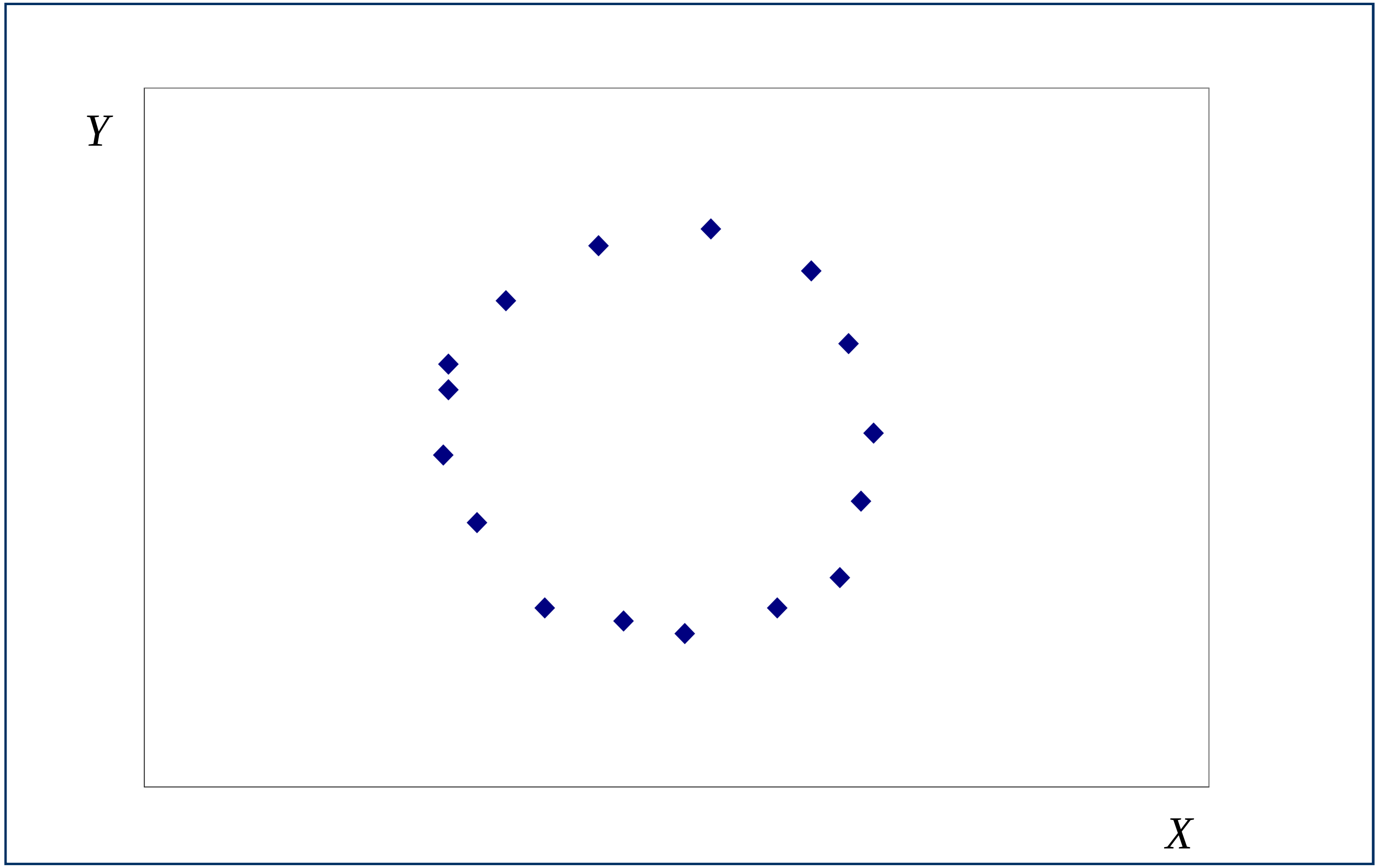
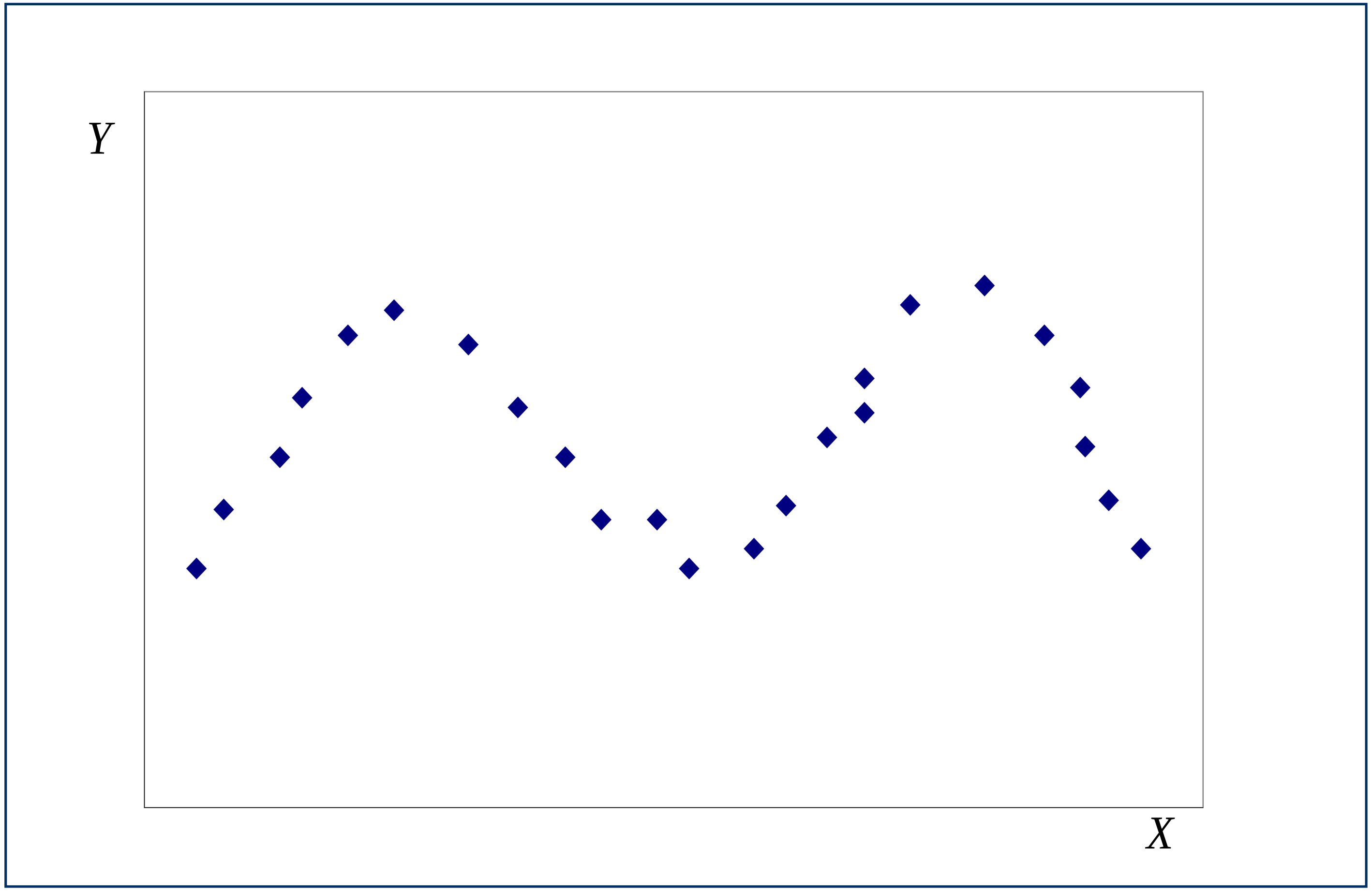


圖8.13 X，Y 為非線性關係



EX8.17 前面已知顧客購買型錄(Y)與購買手機(X)之間爲正向共變, 現問其相關爲何?

Solution:

在EX8.5已知

$$\sigma_X^2 = 0.16 \quad , \quad \sigma_Y^2 = 0.24$$

在EX8.15中已知 Cov (X,Y)=0.08, 故

$$\rho_{XY} = \frac{Cov(X,Y)}{\sigma_X \sigma_Y} = \frac{0.08}{\sqrt{0.16}\sqrt{0.24}} = \frac{0.08}{0.4 \times 0.4898} = 0.4083$$

其相關程度爲中等

林惠玲 陳正倉著 雙葉書廊發行 2004

8.3.5 共變數與相關系數的比較

1. 若 Cov(X,Y)>0, 則 $\rho_{XY} > 0$, 表示X,Y具正的相關
2. 若 Cov(X,Y)<0, 則 $\rho_{XY} < 0$, 表示X,Y具負的相關
3. 若 Cov(X,Y)=0, 則 $\rho_{XY} = 0$, 表示X,Y無線性相關

然而,若 Cov(X,Y因受衡量單位的影響, 共變數的值在 $(-\infty, \infty)$之間, 故不能從其值的大小來判斷其相關程度

ρ_{XY} 介於(-1, +1)之間 ,由ρ_{XY} 的大小可知兩個隨機變數的相關程度的大小

8.3.6 獨立與共變數(相關係數)的關係

- 當兩變數獨立(無相關)時, 共變數與相關係數應爲0; 但共變數與相關係數爲0時, X與Y不一定獨立(無關係).

→ 只能說X與Y無線性關係

→ 可能有非線性關係 (如二次式)

證明見

EX8.18 and 8.19

表8.15 擲銅板的聯合機率

		Y = 0	Y = 1	$f_x(x)$
X	0	0	1/4	1/4
	1	1/2	0	1/2
	2	0	1/4	1/4
$f_y(y)$		1/2	1/2	1

8.4 二元隨機變數函數的期望值

h(X,Y)期望值的定理

設 h(X,Y)為間斷隨機變數(X,Y)的函數, 則其期望值E(h(X,Y))為

$$E(h(X,Y)) = \sum_{i=1}^{n}\sum_{j=1}^{m} h(x_i, y_j) f(x_i, y_j)$$

設 h(X,Y)為連續隨機變數(X,Y)的函數, 則其期望值E(h(X,Y))為

$$E(h(X,Y)) = \int_x \int_y h(x,y) f(x,y) dydx$$

林惠玲 陳正倉著 雙葉書廊發行 2004

8.4 二元隨機變數函數的期望值

- 二元隨機變數線性函數的期望值

$$E(aX \pm bY) = aE(X) \pm bE(Y)$$

- 二元隨機變數線性函數的變異數

$$V(aX \pm bY) = a^2V(X) + b^2V(Y) \pm 2abCov(X,Y)$$

表8.16 台積電、國建兩種股票的投資報酬率

期數（年）	X（台積電）	Y（國建）
88（景氣佳）	1.352	−0.475
89（景氣差）	−0.53	−0.292
90（景氣復甦）	0.115	0.128
91（景氣差）	−0.513	−0.096
92（景氣復甦）	0.491	0.682

資料來源：台灣證券交易所

某投資人將一筆資金等量投資於台積電與國建,試求該投資人一年投資報酬率的期望值與變異數(風險)為何?

Solution: 求投資報酬率的期望值與變異數, 即是求

E(0.5X+0.5Y)與 V(0.5X+0.5Y)

→ E(0.5X+0.5Y)= 0.5 E(X)+0.5E(Y)

V(0.5X+0.5Y)= 0.25V(X)+0.25V(Y)+2*0.5*0.5 Cov(X,Y)

而

$$E(X)=\frac{1.352+(-0.53)+0.115+(-0.513)+0.491}{5}=0.183$$

$$E(Y)=\frac{(-0.475)+(-0.292)+0.128+(-0.096)+0.682}{5}=-0.011$$

$$V(X)=E(X-\mu_X)^2=E(X^2)-\mu_X^2=\frac{\sum x_i^2}{5}-(0.183)^2=0.491$$

$$V(Y)=E(Y-\mu_Y)^2=E(Y^2)-\mu_Y^2=\frac{\sum y_i^2}{5}-(-0.011)^2=0.1602$$

林惠玲 陳正倉著 雙葉書廊發行 2004

共變數爲

$$Cov(X,Y)=E(X-\mu_X)(Y-\mu_Y)=\frac{\sum_{i=1}^{N}(x_i-\mu_X)(y_i-\mu_Y)}{N}=-0.0157$$

因此

E(0.5X+0.5Y)= 0.5 E(X)+0.5E(Y)=0.5(0.183)+0.5(-0.011)=0.086

V(0.5X+0.5Y)= 0.25V(X)+0.25V(Y)+2*0.5*0.5 Cov(X,Y)

= 0.25 (0.4918)+0.25(0.1602)+0.5 (-0.0157)

= 0.1552

多元機率分配的條件

① $0 \le f(x_1, x_2, \cdots, x_6) \le 1$

② $\sum_{x_1}\sum_{x_2}\cdots\sum_{x_6} f(x_1, x_2, \cdots, x_6) = 1$

多元隨機變數

- 多元隨機變數函數的期望值

$$E(W)=\sum_{i=1}^{K}a_iE(X_i)$$

- 多元隨機變數函數的變異數

$$V(W)=\sum_{i=1}^{K}a_i^2V(X_i)+\sum_{i=1}^{K}\sum_{j=1}^{K}a_ia_jCov(X_i,X_j)\quad,\quad i\neq j$$