# 第八章

## 二元隨機變數及其機率分配

# 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。

# 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。

# 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。

# 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。

# 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
  - 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
    - 相關係數、共變數與獨立的關係
  - Excel 的使用

# 二元間斷隨機變數

- **聯合機率函數**
  設 $X$、$Y$ 為間斷隨機變數，$X$ 之值為 $x_1, x_2, \ldots , x_n$，$Y$ 之值為 $y_1, y_2, \ldots , y_n$，則 $f(x , y)$ 為二元聯合機率函數。

# 二元間斷隨機變數

- **聯合機率函數**

  設 $X$、$Y$ 為間斷隨機變數，$X$ 之值為 $x_1, x_2, \ldots, x_n$，$Y$ 之值為 $y_1, y_2, \ldots, y_n$，則 $f(x, y)$ 為二元聯合機率函數。

- **邊際機率函數**

  $X$ 的邊際機率函數

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

# 二元間斷隨機變數

- **聯合機率函數**

  設 $X$、$Y$ 為間斷隨機變數，$X$ 之值為 $x_1, x_2, \ldots, x_n$，$Y$ 之值為 $y_1, y_2, \ldots, y_n$，則 $f(x, y)$ 為二元聯合機率函數。

- **邊際機率函數**

  $X$ 的邊際機率函數

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

  $Y$ 的邊際機率函數

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

# $X$ 與 $Y$ 的聯合機率分配與邊際機率分配表

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

# 二元間斷隨機變數

- **條件機率函數**

  在 $Y = y_j$ 的條件下，發生 $x_i$ 的條件機率表為：

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

# 二元間斷隨機變數

- **條件機率函數**

  在 $Y=y_j$ 的條件下，發生 $x_i$ 的條件機率表為：

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

在 $Y=y_j$ 的條件下，發生 $x_i$ 的條件機率表為：

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

# 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 |

# $X$ 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 |

# $f(x \mid y)$的條件機率

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

## $f(y \mid x)$的條件機率

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

# 二元間斷隨機變數的期望值

- $X$ 的期望值

$$
\begin{aligned}
E(X) &= \sum_x \sum_y x f(x,y) \\
&= \sum_x x \sum_y f(x,y) = \sum_x x f_x(x)
\end{aligned}
$$

# 二元間斷隨機變數的期望值

- $X$ 的期望值

$$
\begin{aligned}
E(X) &= \sum_x \sum_y x f(x,y) \\
&= \sum_x x \sum_y f(x,y) = \sum_x x f_x(x)
\end{aligned}
$$

- $Y$ 的期望值

$$
\begin{aligned}
E(Y) &= \sum_x \sum_y y f(x,y) \\
&= \sum_y y \sum_x f(x,y) = \sum_y y f_y(y)
\end{aligned}
$$

# 二元間斷隨機變數的變異數

- $X$ 的變異數

$$\begin{aligned} V(X) &= \sum_x \sum_y (x-\mu_X)^2 f(x,y) \\ &= \sum_x (x-\mu_X)^2 f_x(x) = \sum_x x^2 f_x(x) - \mu_X^2 \end{aligned}$$

# 二元間斷隨機變數的變異數

- $X$ 的變異數

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

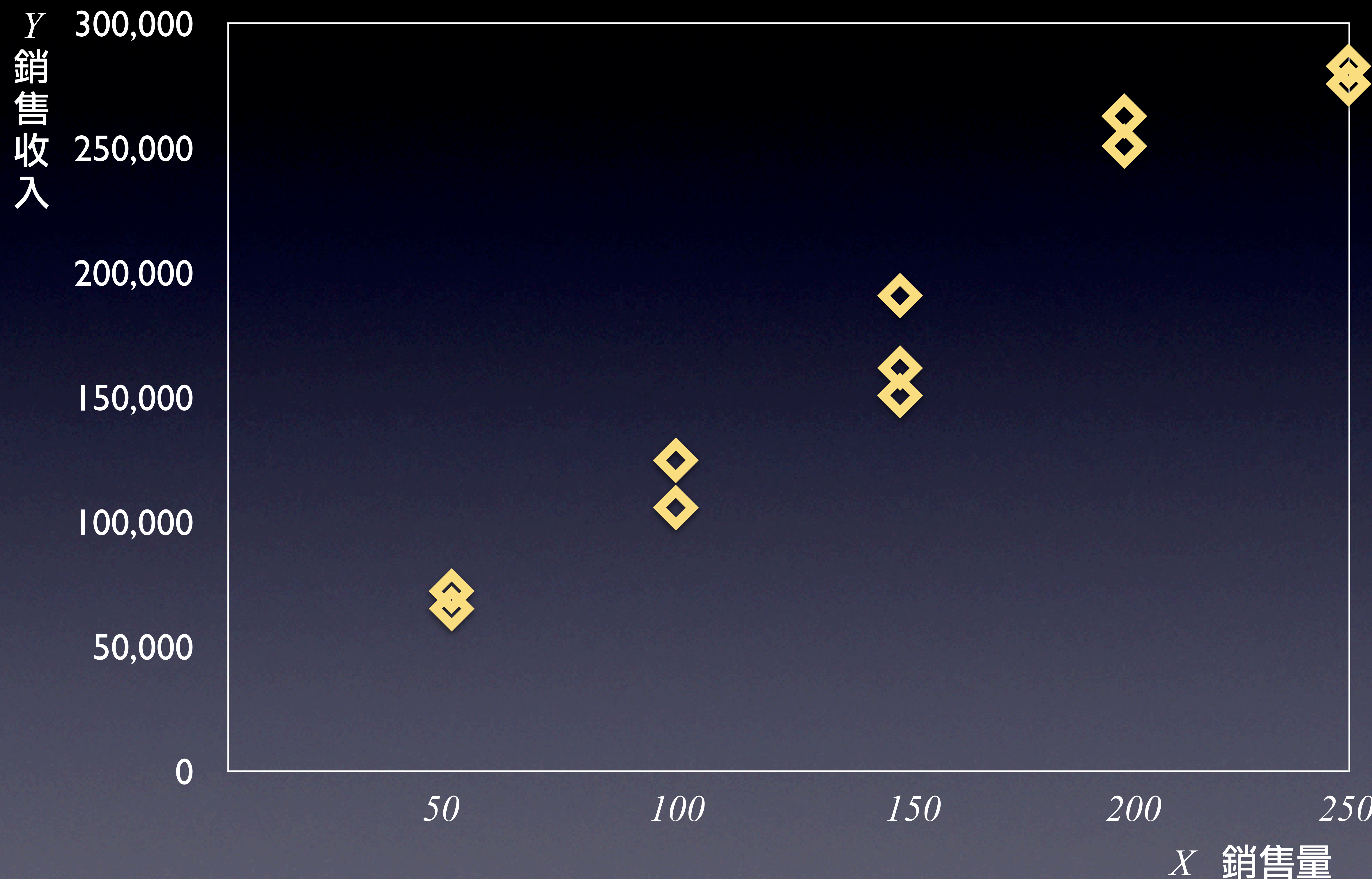
- $Y$ 的變異數

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

# 型錄發放量與銷售收入

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

# 型錄發放量與銷售收入的散佈圖



# 兩變數間的關係

- 兩變數獨立的條件
  設 $X, Y$ 為二元隨機變數，若 $X$ 與 $Y$ 之值均滿足下列任一條件，則 $X, Y$ 獨立。

# 兩變數間的關係

- **兩變數獨立的條件**
  設 $X, Y$ 為二元隨機變數，若 $X$ 與 $Y$ 之值均滿足下列任一條件，則 $X, Y$ 獨立。

1. $f(x, y) = f_x(x) \cdot f_y(y)$

# 兩變數間的關係

- **兩變數獨立的條件**
  設 $X, Y$ 為二元隨機變數，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. $f(x \mid y) = f_x(x)$

# 兩變數間的關係

- **兩變數獨立的條件**
  設 $X, Y$ 為二元隨機變數，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. $f(y \mid x) = f_y(y)$

# 繫安全帶與車禍受傷程度統計表

| $X$（受傷程度） | $Y$（繫安全帶） | | 合計 |
|---|---|---|---|
| | 0（無） | 1（有） | |
| 1（輕傷） | 380 | 170 | 550 |
| 2（重傷） | 240 | 50 | 290 |
| 3（死亡） | 140 | 20 | 160 |
| 合計 | 760 | 240 | 1,000 |

# $X$ 與 $Y$ 的聯合機率分配表

| $X$（受傷程度） | $Y$（繫安全帶） | | 合計 $f_X(x)$ |
|---|---|---|---|
| | 0（無） | 1（有） | |
| 1（輕傷） | 0.38 | 0.17 | 0.55 |
| 2（重傷） | 0.24 | 0.05 | 0.29 |
| 3（死亡） | 0.14 | 0.02 | 0.16 |
| 合計 $f_Y(y)$ | 0.76 | 0.24 | 1.00 |

# 兩變數間的關係

- 共變數

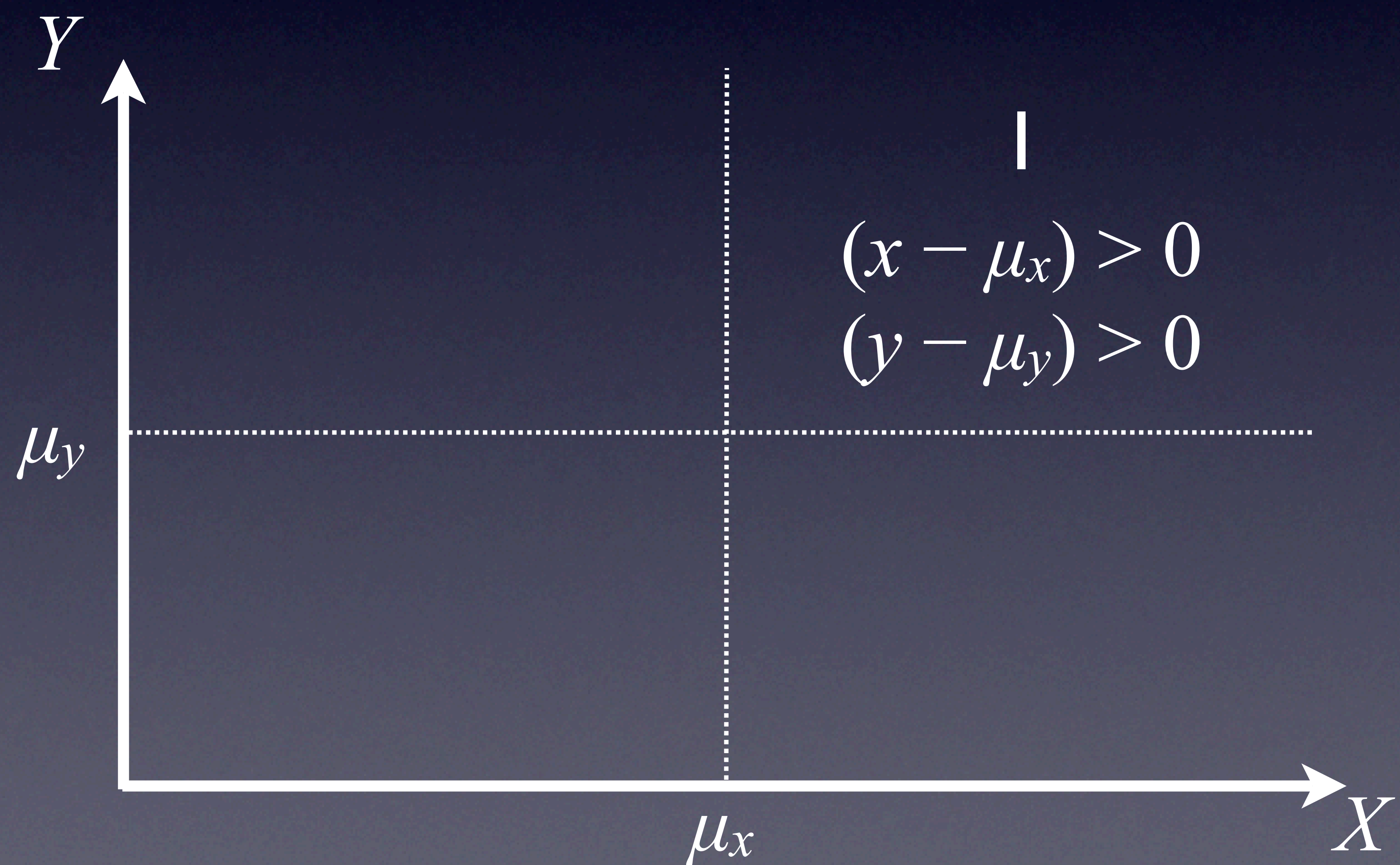
$$Cov(X, Y) = E[(X - \mu_x)(Y - \mu_y)]$$

# 兩變數間的關係

- **共變數**

$$Cov(X,Y) = E[(X-\mu_x)(Y-\mu_y)]$$

$Cov(X, Y)$ 的符號

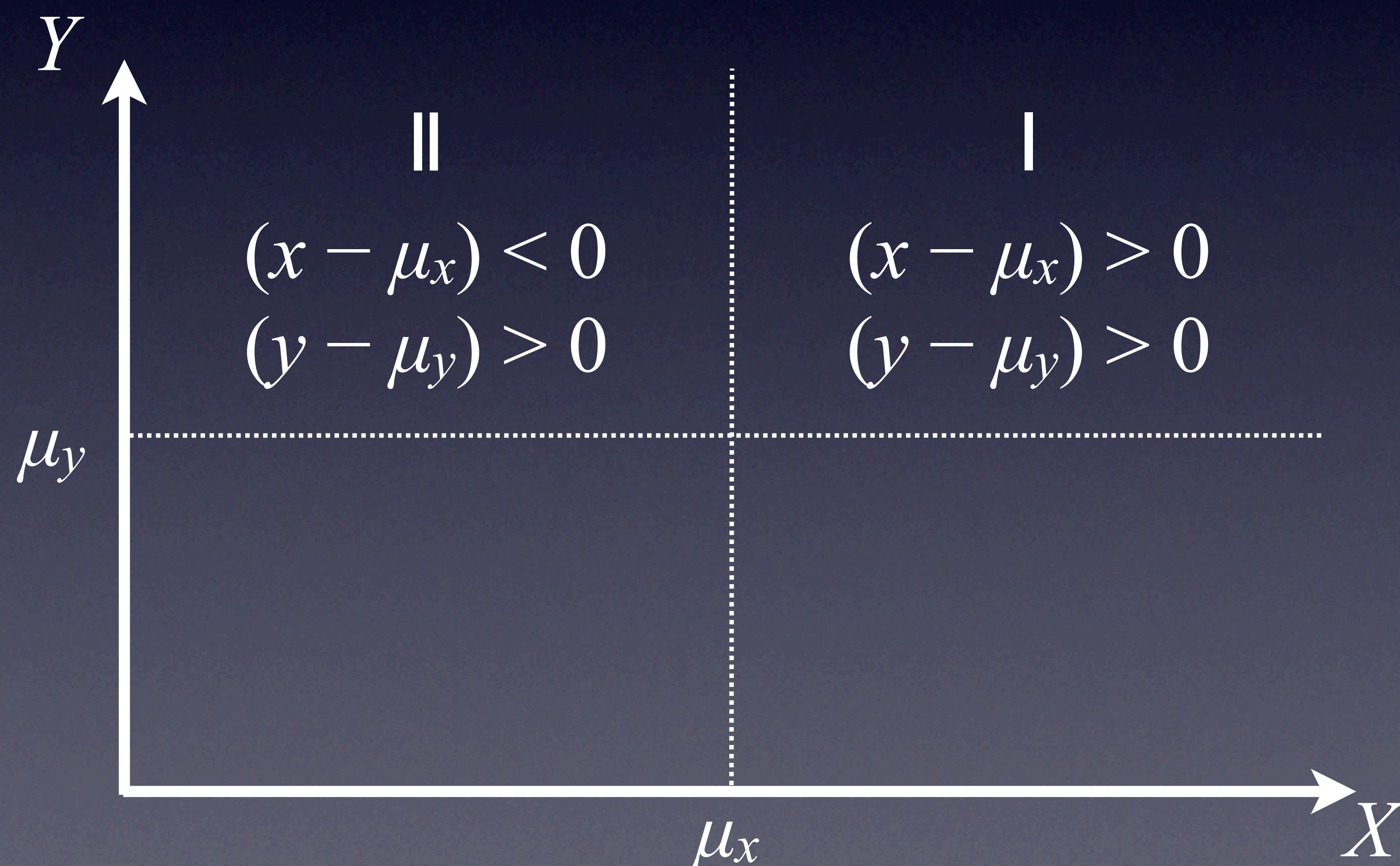


# 兩變數間的關係

- **共變數**

$$Cov(X,Y) = E[(X-\mu_x)(Y-\mu_y)]$$

$Cov(X, Y)$ 的符號



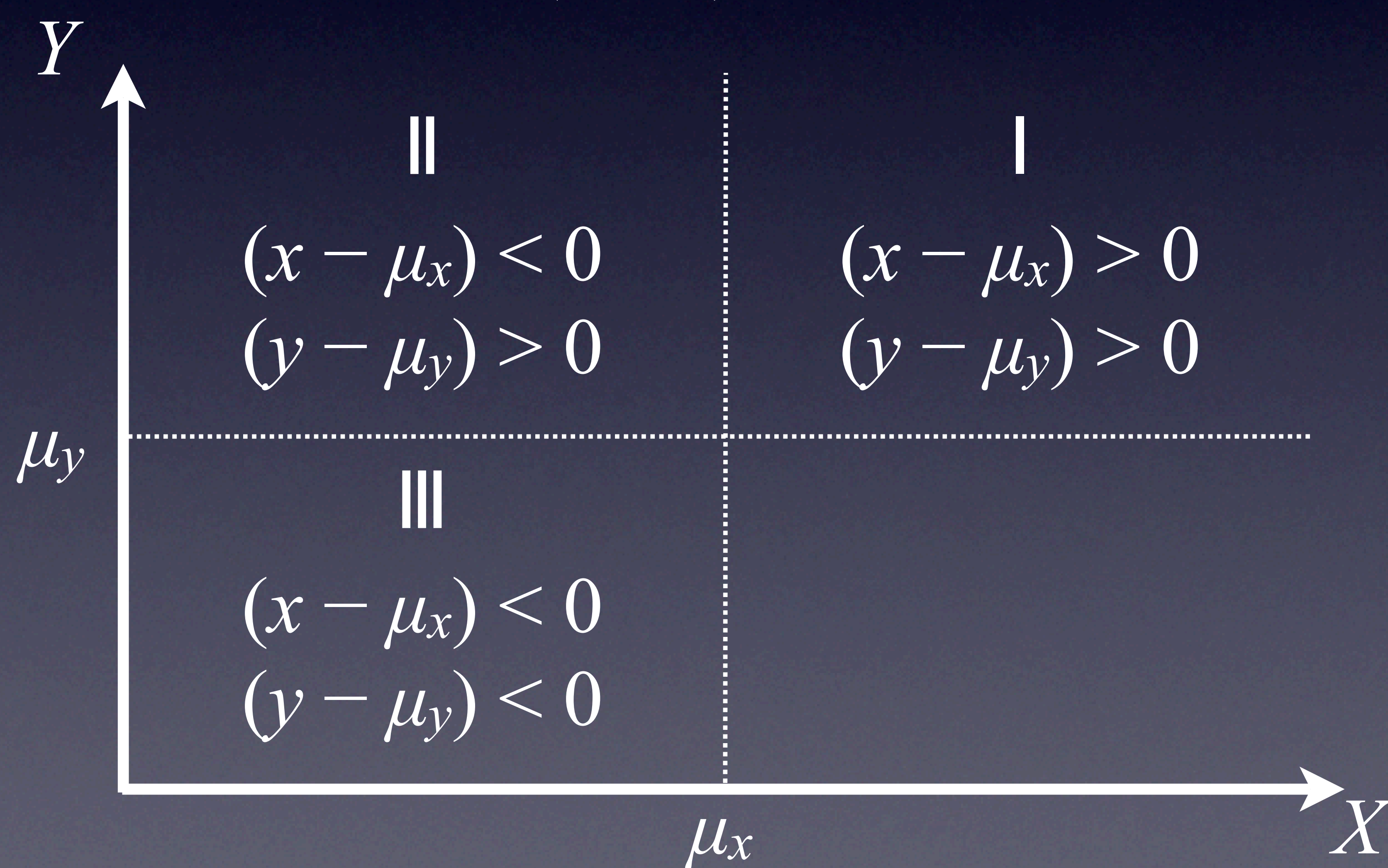

應用統計學 林惠玲 陳正倉著 雙葉書廊發行 2006

# 兩變數間的關係

- **共變數**

$$Cov(X,Y) = E[(X-\mu_x)(Y-\mu_y)]$$

$Cov(X, Y)$ 的符號

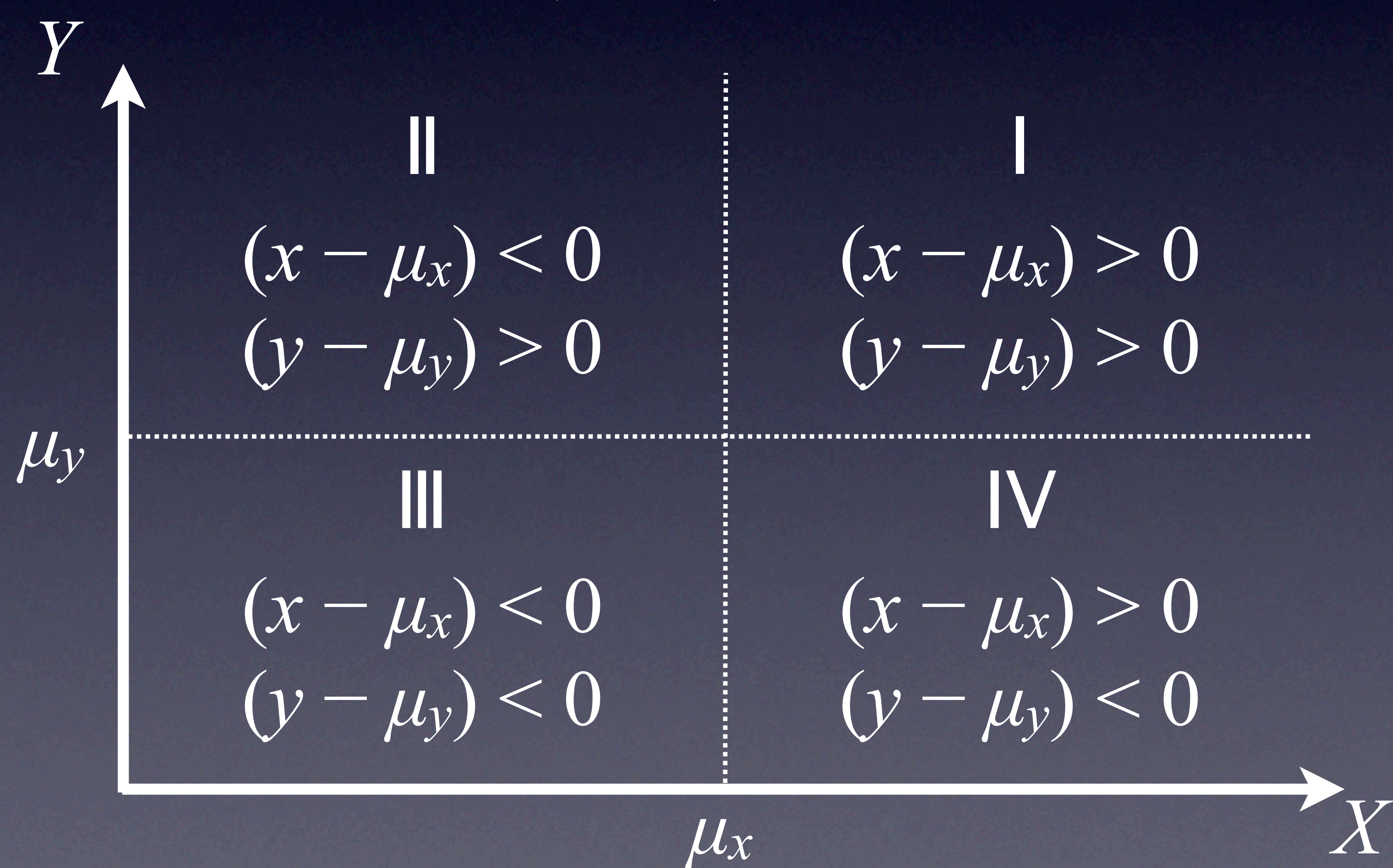


# 兩變數間的關係

- **共變數**

$$Cov(X,Y) = E[(X-\mu_x)(Y-\mu_y)]$$

$Cov(X, Y)$ 的符號

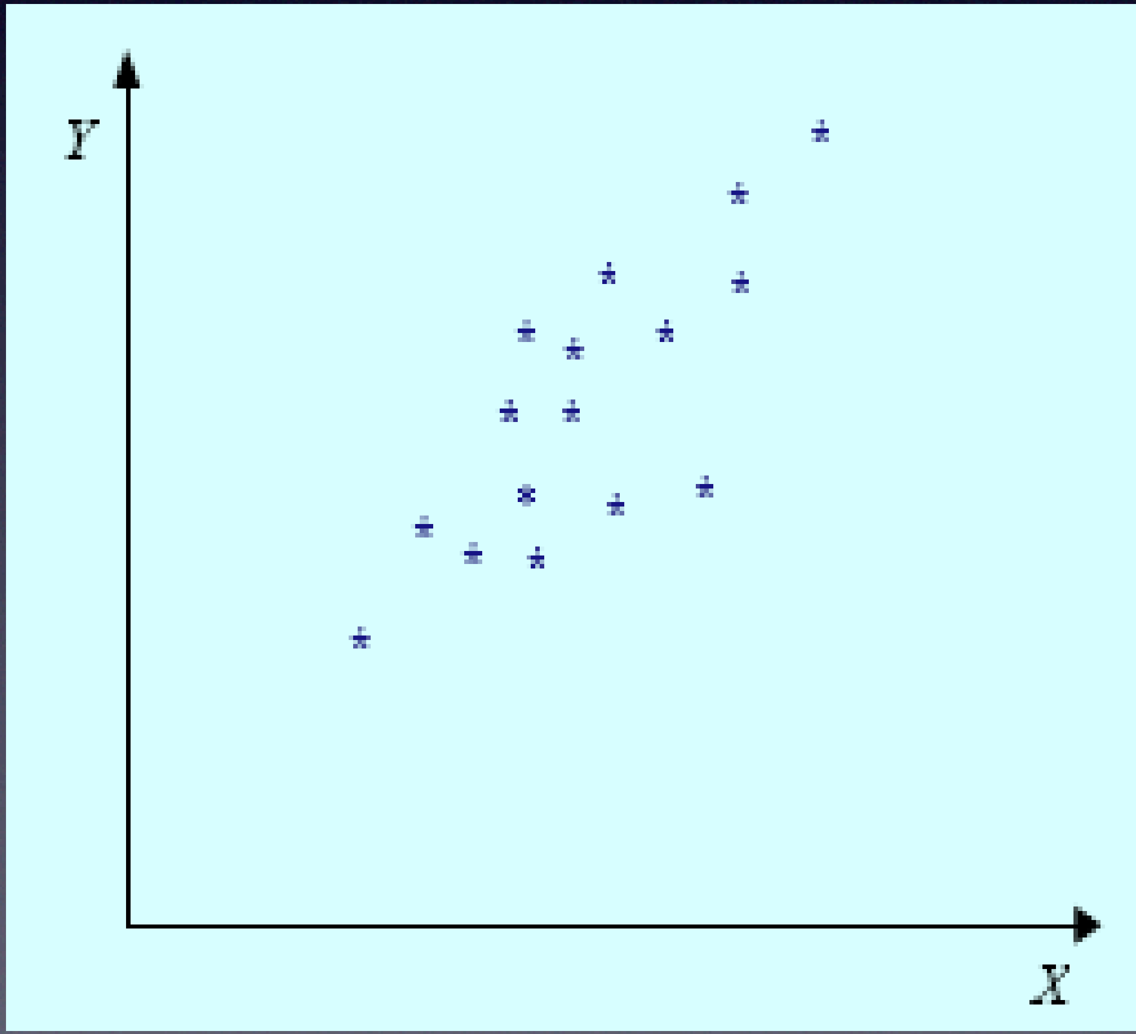


# 兩變數間的關係

- **共變數**

$$Cov(X,Y) = E[(X-\mu_x)(Y-\mu_y)]$$

正向共變

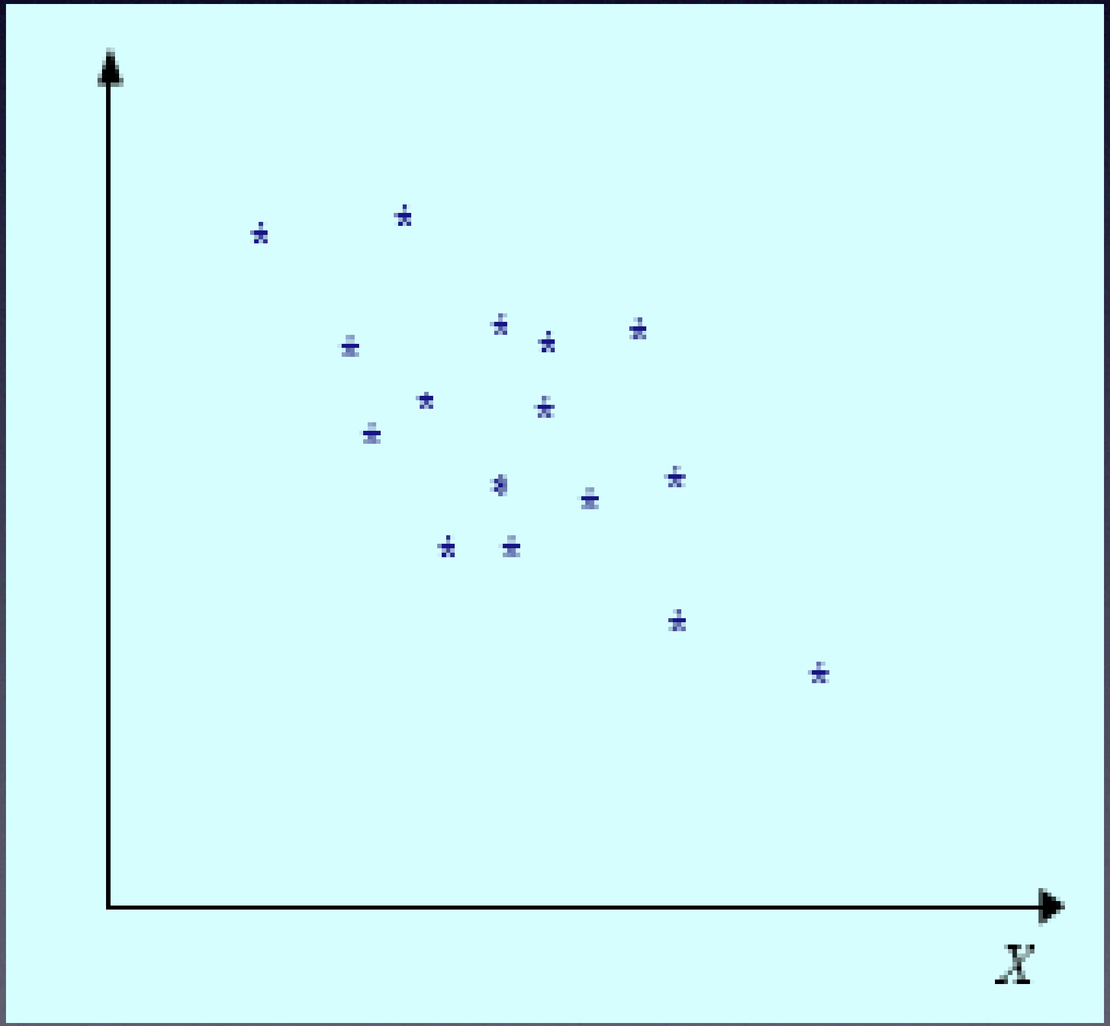


# 兩變數間的關係

- **共變數**

$$Cov(X,Y) = E[(X-\mu_x)(Y-\mu_y)]$$

反向共變

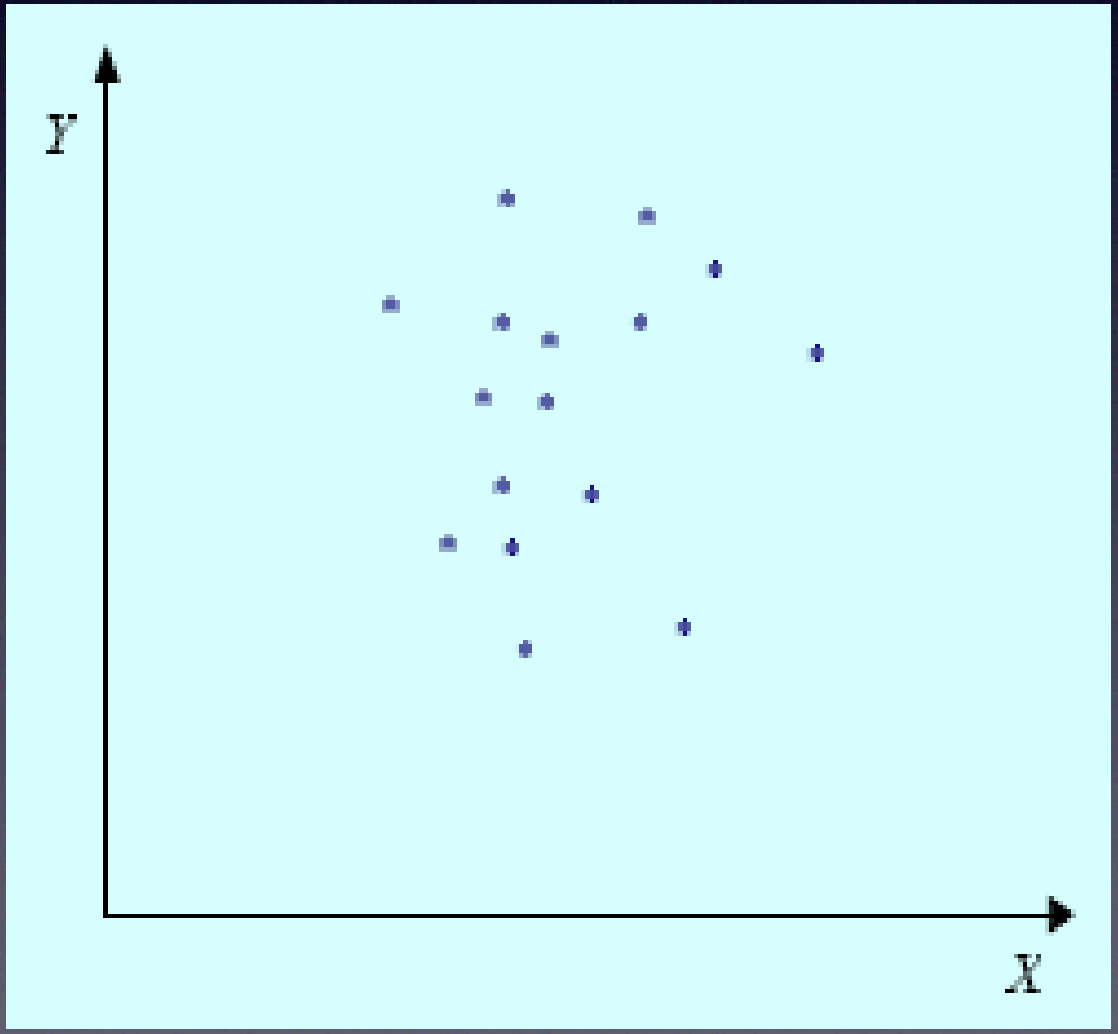


# 兩變數間的關係

- 共變數

$$Cov(X,Y) = E[(X-\mu_x)(Y-\mu_y)]$$

無關係



# 兩變數間的關係

- 共變數的優缺點

  1. 它可看出共變的方向與程度。

# 兩變數間的關係

- 共變數的優缺點

1. 它可看出共變的方向與程度。
2. $-\infty < Cov(X\ ,\ Y) < \infty$，其值域無限，不易根據 $Cov(X\ ,Y)$ 的大小來判斷其相關程度。

# 兩變數間的關係

- **共變數的優缺點**

1. 它可看出共變的方向與程度。
2. $-\infty < Cov(X\,,\,Y) < \infty$，其值域無限，不易根據 $Cov(X\,,Y)$ 的大小來判斷其相關程度。
3. 它易受衡量單位的影響。

# 兩變數間的關係

- 共變數的優缺點

1. 它可看出共變的方向與程度。
2. $-\infty < Cov(X, Y) < \infty$，其值域無限，不易根據 $Cov(X, Y)$ 的大小來判斷其相關程度。
3. 它易受衡量單位的影響。
4. 具有雙重的衡量單位。

# 利用 Excel 求共變數

## 手機客戶的資料

| | A | B |
|---|---|---|
| 1 | 是否看過型錄 | 是否購買手機 |
| 2 | 0 | 1 |
| 3 | 1 | 0 |
| 4 | 0 | 1 |
| 5 | 0 | 0 |
| 6 | 1 | 1 |
| 7 | 1 | 0 |
| 8 | 0 | 1 |
| 9 | 0 | 0 |
| 10 | 1 | 0 |
| 11 | 0 | 1 |

# 利用 Excel 求共變數

手機客戶的資料

| | A | B |
|---|---|---|
| 1 | 是否看過型錄 | 是否購買手機 |
| 2 | 0 | 1 |
| 3 | 1 | 0 |
| 4 | 0 | 1 |
| 5 | 0 | 0 |
| 6 | 1 | 1 |
| 7 | 1 | 0 |
| 8 | 0 | 1 |
| 9 | 0 | 0 |
| 10 | 1 | 0 |
| 11 | 0 | 1 |

工具 → 資料分析 → 共變數

# 利用 Excel 求共變數

看過型錄與購買手機之共變數

| | A | B | C |
|---|---|---|---|
| 1 | | 看過型錄 | 購買手機 |
| 2 | 看過型錄 | 0.2402 | |
| 3 | 購買手機 | 0.0775 | 0.2085 |

# 兩變數間的關係

- **相關係數**

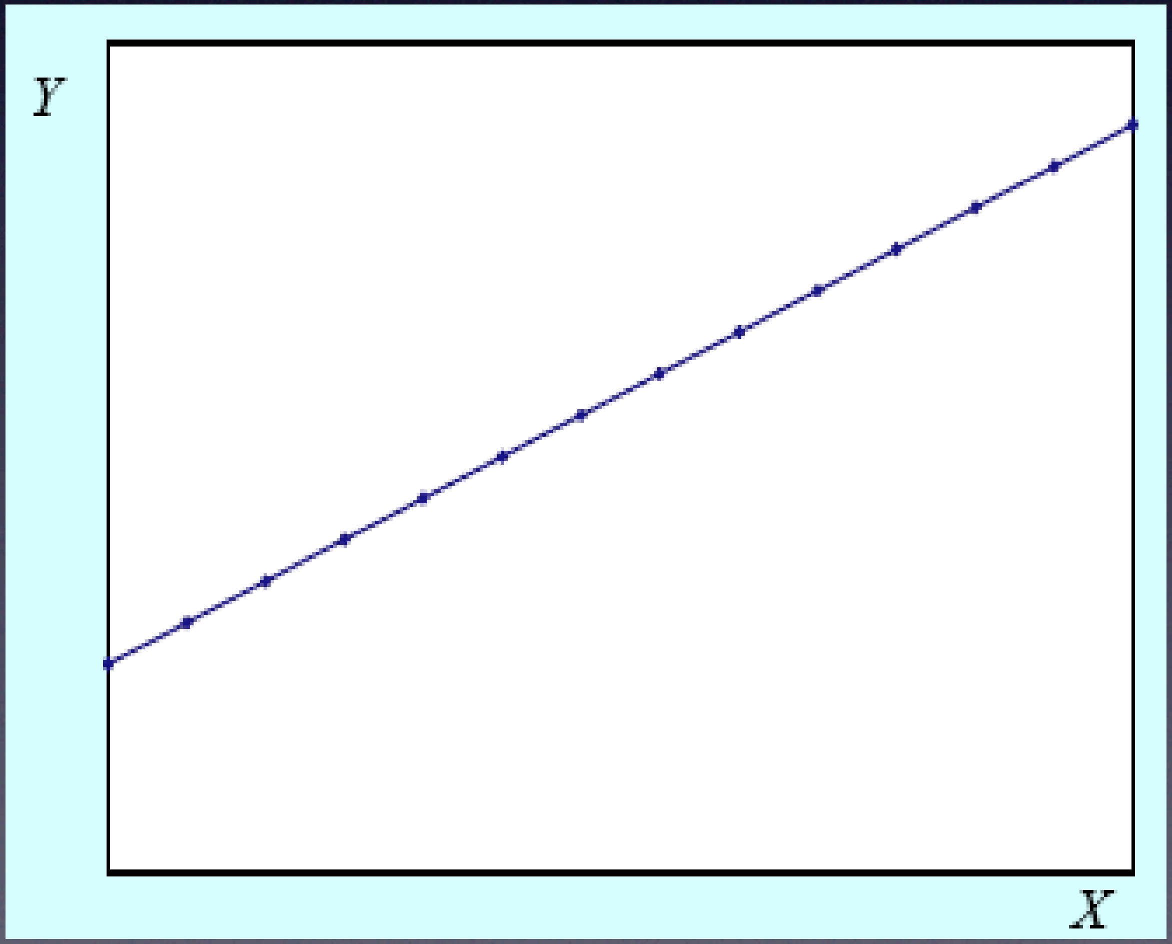
$$\rho_{XY} = \frac{Cov(X,Y)}{\sigma_X \sigma_Y} = E\left(\frac{X-\mu_X}{\sigma_X}\right)\left(\frac{Y-\mu_Y}{\sigma_Y}\right)$$

# 兩變數間的關係

- **相關係數**

$$\rho_{XY} = \frac{Cov(X,Y)}{\sigma_X \sigma_Y} = E\left(\frac{X-\mu_X}{\sigma_X}\right)\left(\frac{Y-\mu_Y}{\sigma_Y}\right)$$

相關係數 $\rho_{XY} = 1$

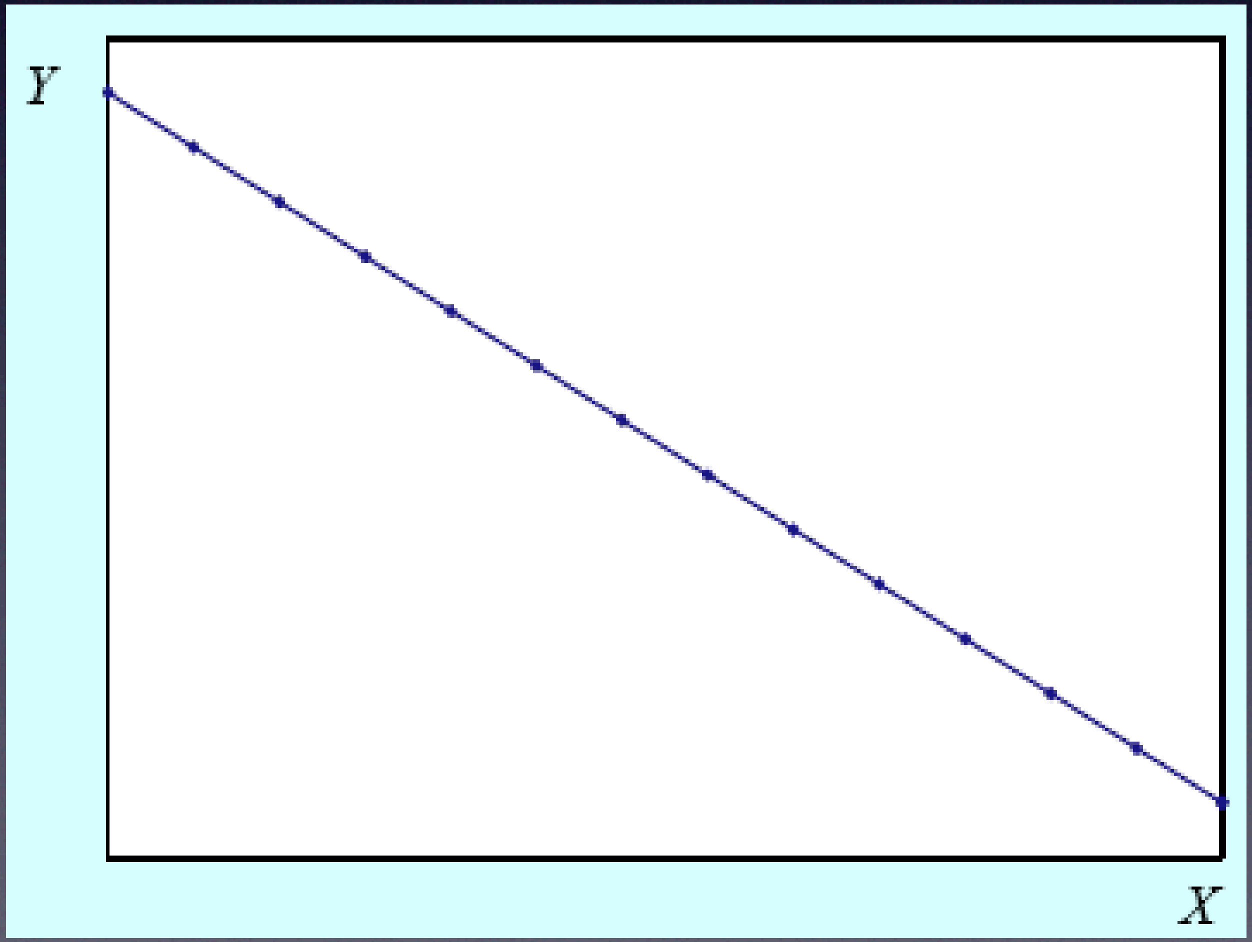


# 兩變數間的關係

- **相關係數**

$$\rho_{XY} = \frac{Cov(X,Y)}{\sigma_X \sigma_Y} = E\left(\frac{X-\mu_X}{\sigma_X}\right)\left(\frac{Y-\mu_Y}{\sigma_Y}\right)$$

相關係數 $\rho_{XY} = -1$

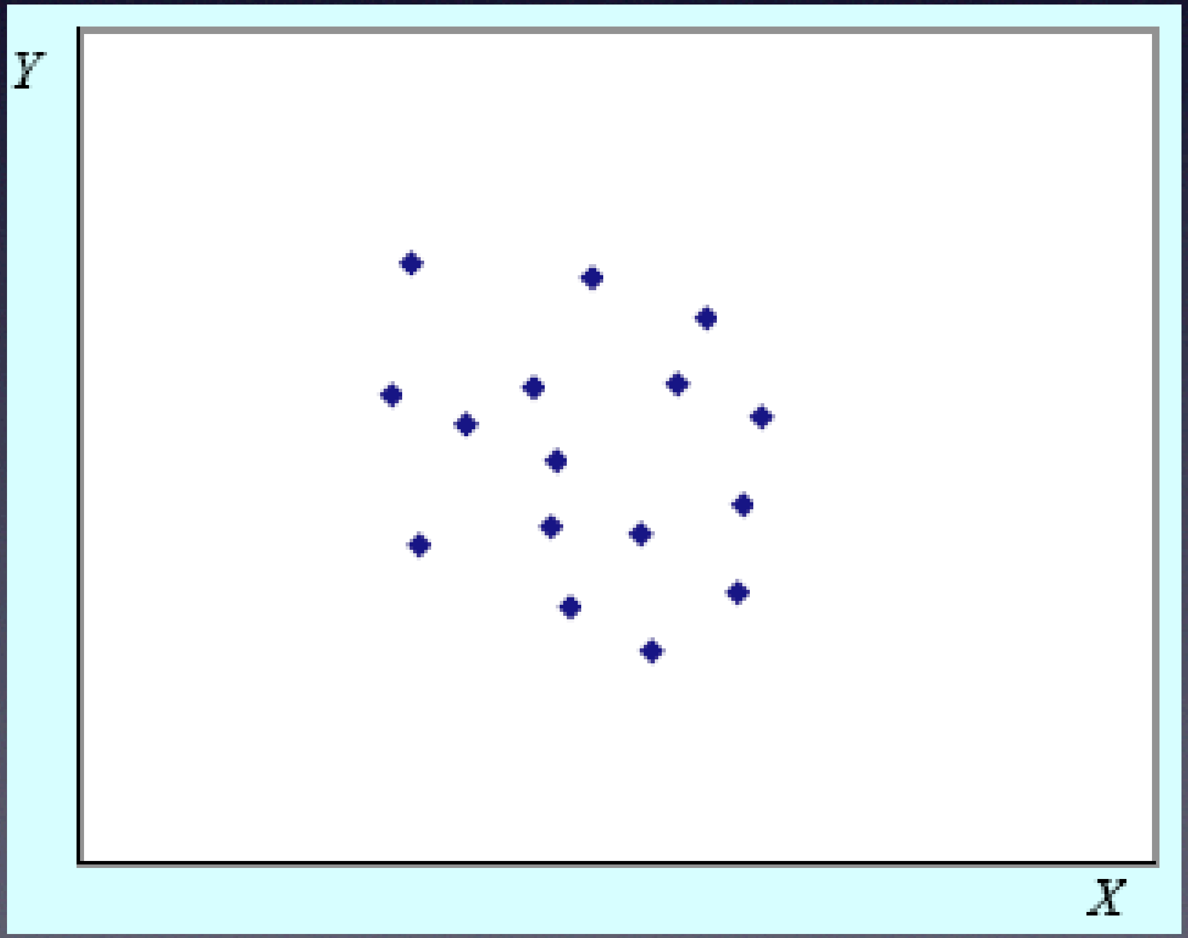


# 兩變數間的關係

- **相關係數**

$$\rho_{XY} = \frac{Cov(X,Y)}{\sigma_X \sigma_Y} = E\left(\frac{X-\mu_X}{\sigma_X}\right)\left(\frac{Y-\mu_Y}{\sigma_Y}\right)$$

$X$、$Y$ 無關

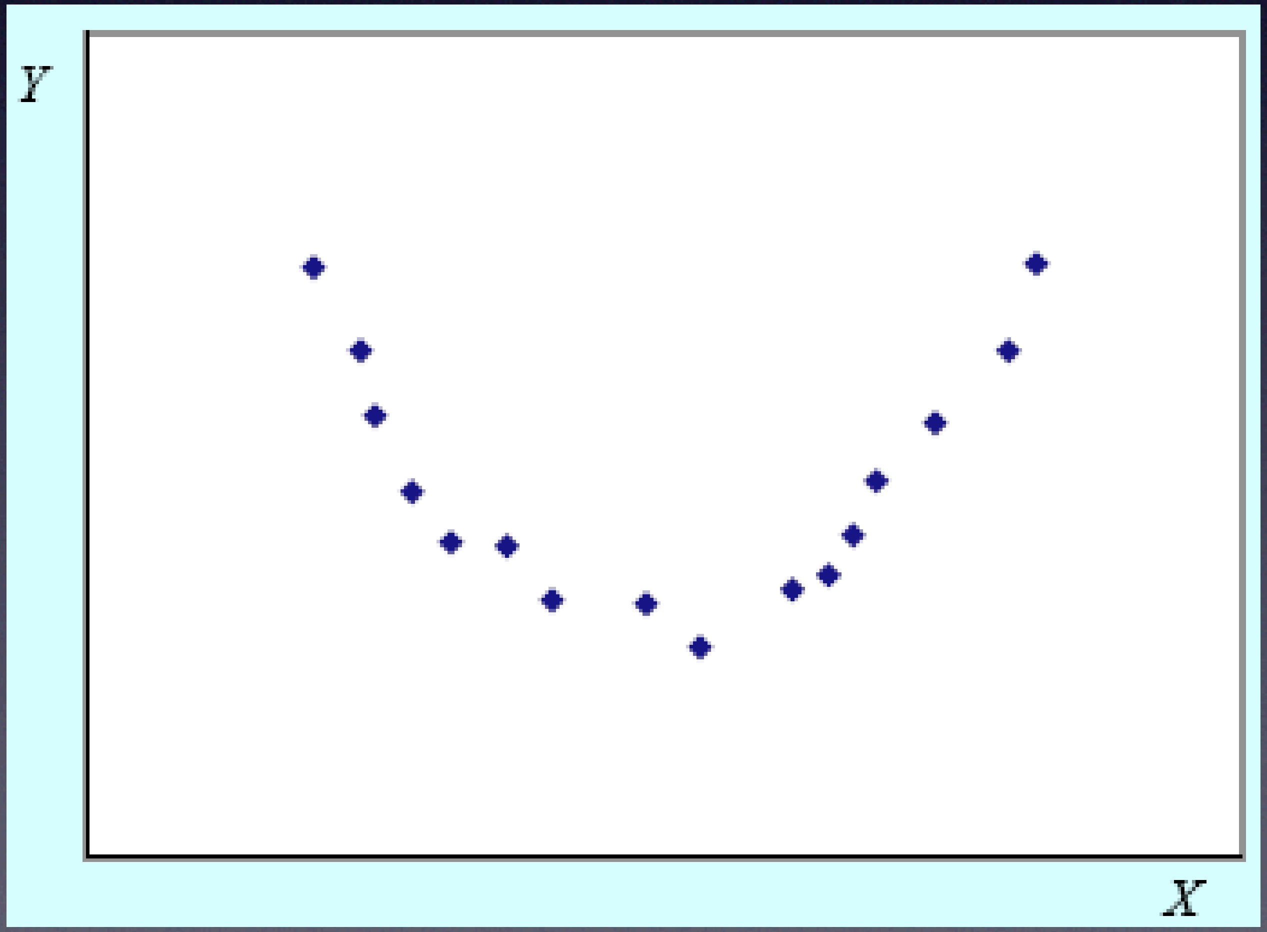


# 兩變數間的關係

- 相關係數

$$\rho_{XY} = \frac{Cov(X,Y)}{\sigma_X \sigma_Y} = E\left(\frac{X-\mu_X}{\sigma_X}\right)\left(\frac{Y-\mu_Y}{\sigma_Y}\right)$$

$X$、$Y$ 為拋物線觀念

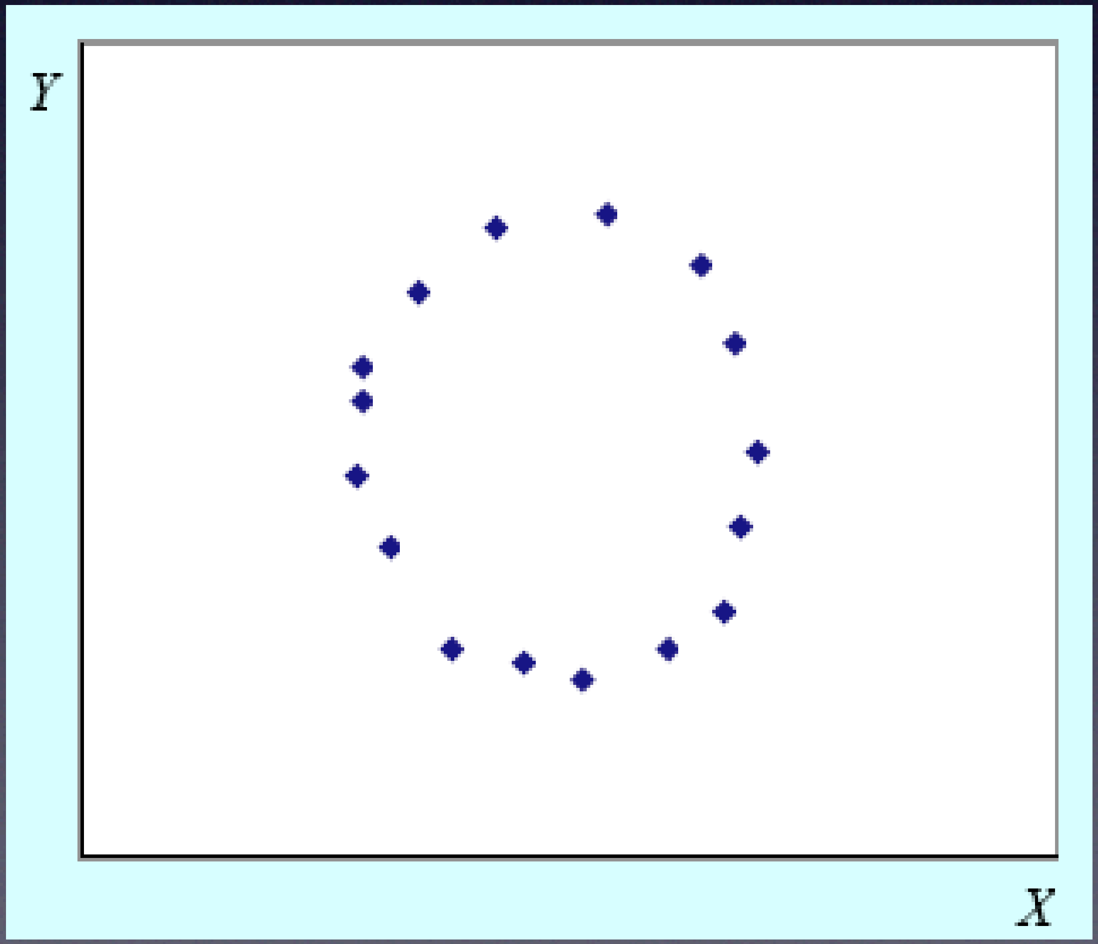


# 兩變數間的關係

- **相關係數**

$$\rho_{XY} = \frac{Cov(X,Y)}{\sigma_X \sigma_Y} = E\left(\frac{X-\mu_X}{\sigma_X}\right)\left(\frac{Y-\mu_Y}{\sigma_Y}\right)$$

$X$、$Y$ 為圓形觀念

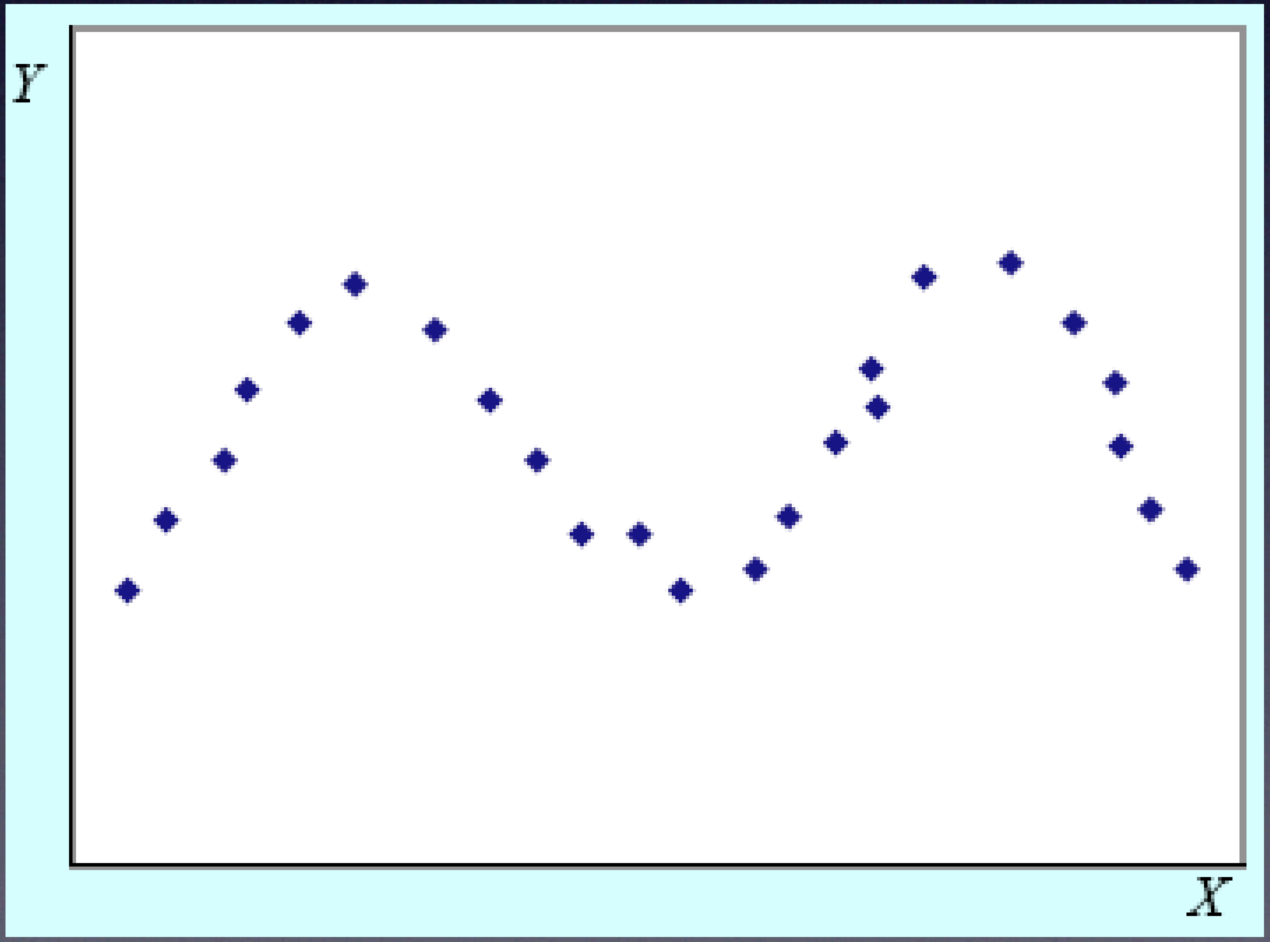


# 兩變數間的關係

- **相關係數**

$$\rho_{XY} = \frac{Cov(X,Y)}{\sigma_X \sigma_Y} = E\left(\frac{X-\mu_X}{\sigma_X}\right)\left(\frac{Y-\mu_Y}{\sigma_Y}\right)$$

$X$、$Y$ 為非線性關係



# 兩變數間的關係

- 相關係數的數值

1. 相關係數的數值可證明世界於 $-1$ 與 $+1$ 之間。

# 兩變數間的關係

- 相關係數的數值

 1. 相關係數的數值可證明世界於 $-1$ 與 $+1$ 之間。
 2. 當 $\rho_{XY} = +1$ 時，我們稱 $X$、$Y$ 具正的完全直線關係，此時 $(X\ ,\ Y)$ 的點會落在一條斜率正的直線上。

# 兩變數間的關係

- 相關係數的數值

1. 相關係數的數值可證明世界於 $-1$ 與 $+1$ 之間。
2. 當 $\rho_{XY} = +1$ 時，我們稱 $X$、$Y$ 具正的完全直線關係，此時 $(X, Y)$ 的點會落在一條斜率正的直線上。
3. 當 $\rho_{XY} = 0$ 時，$X$、$Y$無線性旅遊

# 兩變數間的關係

- 相關係數的數值

1. 相關係數的數值可證明世界於 $-1$ 與 $+1$ 之間。
2. 當 $\rho_{XY} = +1$ 時，我們稱 $X$、$Y$ 具正的完全直線關係，此時 $(X\ ,\ Y)$ 的點會落在一條斜率正的直線上。
3. 當 $\rho_{XY} = 0$ 時，$X$、$Y$無線性旅遊
4. $|\,\rho_{XY}\,|$ 愈大表示 $X$、$Y$ 的線性相關程度愈大。

# 利用 Excel 求相關係數

## 手機客戶的資料

| | A | B |
|---|---|---|
| 1 | 是否看過型錄 | 是否購買手機 |
| 2 | 0 | 1 |
| 3 | 1 | 0 |
| 4 | 0 | 1 |
| 5 | 0 | 0 |
| 6 | 1 | 1 |
| 7 | 1 | 0 |
| 8 | 0 | 1 |
| 9 | 0 | 0 |
| 10 | 1 | 0 |
| 11 | 0 | 1 |

# 利用 Excel 求相關係數

## 手機客戶的資料

| | A | B |
|---|---|---|
| 1 | 是否看過型錄 | 是否購買手機 |
| 2 | 0 | 1 |
| 3 | 1 | 0 |
| 4 | 0 | 1 |
| 5 | 0 | 0 |
| 6 | 1 | 1 |
| 7 | 1 | 0 |
| 8 | 0 | 1 |
| 9 | 0 | 0 |
| 10 | 1 | 0 |
| 11 | 0 | 1 |

工具 → 資料分析 → 相關係數

# 利用 Excel 求相關係數

看過型錄與購買手機之相關係數

| | A | B | C |
|---|---|---|---|
| 1 | | 看過型錄 | 購買手機 |
| 2 | 看過型錄 | 1 | |
| 3 | 購買手機 | 0.356348323 | 1 |