第 9 章　二元隨機變數及其機率分配

- 世界上有許多事件是同時發生且相互關聯的。
  - 例如：『運動與高血壓』、『酒駕與車禍』、『廣告費支出與商品銷售量』、『工作年資、教育程度與薪資所得』。
  - 二個(或二個以上)相互有關的事件可定義為二元(或多元)隨機變數(bivariate [or multivariate] random variables)。
- 二元(或多元)隨機變數亦可區分為『二元(或多元)間斷隨機變數』與『二元(或多元)連續隨機變數』。
  - 二元間斷隨機變數：隨機抽取一名大學生，觀察其「性別」與「是否抽菸」，令隨機變數 $X=1$ 為男性、$X=0$ 為女性，令隨機變數 $Y=1$ 代表抽菸、$Y=0$ 代表未抽菸，則 X、Y 為定義在樣本空間中的實數值函數，為二元間斷隨機變數。
  - 二元連續隨機變數：隨機抽取一製造業工人，觀察其年齡(X)與薪資(Y)，則 X、Y 為二元連續隨機變數。

- 間斷＋連續隨機變數：隨機抽取一小學生，令 X 代表其身高(連續)、Y 代表其年級(間斷)，則 X、Y 為二元隨機變數，但一個是連續、一個是間斷。

二元間斷隨機變數(bivariate discrete random variable)

- 二元間斷隨機變數的**聯合機率函數**(joint probability function)：設 X、Y 為二元間斷隨機變數，X 之可能值為 $x_1,x_2,\cdots,x_n$，Y 之可能值為 $y_1,y_2,\cdots y_m$，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 $f(x_i,y_j)$ [簡單表示為 $f(x,y)$]為聯合機率函數。

  - $f(x_i,y_j)$ 代表隨機變數 X 發生 $X=x_i$，且隨機變數 Y 發生 $Y=y_j$ 的機率，亦即 $f(x_i,y_j)=\boxed{P(X=x_i \text{ 且 } Y=y_j)}$。

- 下表完整列出 X、Y 的聯合機率函數：例如，$f(x_1,y_1)$ 代表 $P(X=x_1$ 且 $Y=y_1)$，$f(x_n,y_m)$ 則代表 $P(X=x_n$ 且 $Y=y_m)$

$X \backslash Y$	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

- 二元間斷隨機變數的**二元累加分配函數**：由於現在有兩個隨機變數，累加分配函數應計算兩個隨機變數小於某特定變量的機率

$$F(x,y)=P(X\le x \text{ 且 } Y\le y)$$

  - 若 X、Y 為二元間斷隨機變數，則其累加分配函數定義為

$$F(x,y)=\sum_{x_i\le x}\sum_{y_j\le y} f(x_i,y_j)$$

  - $F(x,y)$ 具有非遞減的性質：若 $a\le b$，則 $F(a,y_j)\le F(b,y_j)$ [不管 y_j 為何]；若 $c\le d$，則 $F(x_i,c)\le F(x_i,d)$ [不管 x_i 為何]
- 二元累加分配函數描述的是兩個隨機變數的聯合隨機行為，而邊際分配則是這兩個隨機變數中任一個變數的個別隨機行為。
- 二元間斷隨機變數的**邊際機率函數**(marginal probability function)：
  - X 的邊際機率函數：我們以 $f_x(x)$ 來表示隨機變數 X 出現變量 $X=x$ 的機率[不管 Y 出現何種變量，僅關心 X]

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

    - ■ 上式必須滿足下列兩條件：

$$0\le f_x(x_i)\le 1$$

$$\sum_{i=1}^{n} f_x(x_i)=1$$

- Y 的邊際機率函數：我們以 $f_y(y)$ 來表示隨機變數 Y 出現變量 $Y=y$ 的機率[不管 X 出現何種變量，僅關心 Y]

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

  - 上式必須滿足下列兩條件：

$$0\le f_y(y_j)\le 1$$

$$\sum_{j=1}^{m} f_y(y_j)=1$$

- 根據邊際機率函數也可以計算**邊際累加分配函數** (marginal cumulative distribution function)

X 的邊際累加分配函數：

$$F_x(x)=\sum_{x_i\le x} f_x(x_i)=\sum_{x_i\le x}\sum_{j=1}^{m} f(x_i,y_j)$$

[$F_x(x)$ 其實就等於 $F(x,\infty)$：累加分配函數中未對 Y 進行範圍限制所得到的累加機率]

Y 的邊際累加分配函數：

$$F_y(y)=\sum_{y_j\le y} f_y(y_j)=\sum_{y_j\le y}\sum_{i=1}^{n} f(x_i,y_j)$$

[$F_y(y)$ 其實就等於 $F(\infty,y)$：累加分配函數中未對 X 進行範圍限制所得到的累加機率]

- **條件機率函數**(conditional probability function)：設 $f(x,y)$ 為二元聯合機率函數，則在 $Y=y_j$ 條件下，X 發生 x_i 的條件機率表為：

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

在 $X=x_i$ 條件下，Y 發生 y_j 的條件機率表為：

$$f(y_j\,|\,X=x_i)=\frac{P(x=x_i\ 且\ Y=y_j)}{P(x=x_i)}=\boxed{\frac{f(x_i,y_j)}{f_x(x_i)}}$$

- 例子：某電訊公司最近發送簡訊給曾向該公司買過 2G 手機的 600 個顧客，進行 3G 手機的促銷；發送之後 30 天依客戶『買或不買 3G 手機』與『是否仔細看過簡訊』進行分類統計，統計數據如下表。請問，發送簡訊的促銷方法是否有效？

		Y（是否看過簡訊）		合計
		0（未看過）	1（看過）	
X（有否購買 3G 手機）	0（不買）	198	132	330
	1（購買）	108	162	270
合計		306	294	600

- 依據上表，可得底下的聯合機率函數

		Y（是否看過簡訊）		合計 $f_x(x)$
		0（未看過）	1（看過）	
X（有否購買 3G 手機）	0（不買）	0.33	0.22	0.55
	1（購買）	0.18	0.27	0.45
合計 $f_y(y)$		0.51	0.49	1.00

- 隨機變數 X 僅有兩種變量 $X=0,1$，隨機變數 Y 亦僅有兩種變量 $Y=0,1$。由表中數據即可知道 X 與 Y 的聯合機率，例如『不買 3G 手機、但看過簡訊』[$X=0$、$Y=1$]的機率為 $f(0,1)=0.22$。
- 最右邊一行列出了 X 的邊際機率[$f_x(0)=0.55$、$f_x(1)=0.45$]，最下面一列則列出 Y 的邊際機率[$f_y(0)=0.51$、$f_y(1)=0.49$]。
- 看過簡訊($Y=1$)後的顧客購買手機($X=1$)的機率為

$$P(X=1\,|\,Y=1)=f(1\,|\,Y=1)=\frac{f(1,1)}{f_y(1)}=\frac{0.27}{0.49}=0.55$$

看過簡訊($Y=1$)後的顧客不買手機($X=0$)的機率為

$$P(X=0\,|\,Y=1)=f(0\,|\,Y=1)=\frac{f(0,1)}{f_y(1)}=\frac{0.22}{0.49}=0.45$$

未看過簡訊($Y=0$)的顧客購買手機($X=1$)的機率為

$$P(X=1|Y=0)=f(1|Y=0)=\frac{f(1,0)}{f_y(0)}=\frac{0.18}{0.51}=0.36$$

未看過簡訊($Y=0$)的顧客不購買手機($X=0$)的機率為

$$P(X=0|Y=0)=f(0|Y=0)=\frac{f(0,0)}{f_y(0)}=\frac{0.33}{0.51}=0.64$$

以上計算可整理成底下的條件機率 $f(x|y)$

$f(x\|y)$	$f(x\|Y=0)$（未看過）	$f(x\|Y=1)$（看過）
$X=0$（不買）	0.64	0.45
$X=1$（購買）	0.36	0.55

➢ 由上表可清楚得知,『看過簡訊的顧客購買 3G 手機的機率』(0.55)高於『未看過簡訊的顧客購買手機的機率』(0.36)；因此，發送簡訊的促銷方法確實有效。

➢ 我們亦可依條件機率的定義，計算 $f(y|x)$ 如下表

[給定顧客購買手機($X=1$)或未購買手機($X=0$)的情況下，顧客看過簡訊($Y=1$)與未看過簡訊($Y=0$)的機率]

$f(y\|x)$	$f(y\|X=0)$（不買）	$f(y\|X=1)$（購買）
$Y=0$（未看過）	0.60	0.40
$Y=1$（看過）	0.40	0.60

從表中數據可知,『購買手機的顧客中看過簡訊的機率』(0.60)比『購買手機的顧客中未看過簡訊的機率』(0.40)高；因此亦可斷定『發送簡訊的促銷方法確實有效』。

二元間斷隨機變數的期望值

- X 的期望值：[僅需使用 X 的邊際機率函數計算即可]

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

$$=\sum_x x\sum_y f(x,y)=\boxed{\sum_x xf_x(x)}$$

- Y 的期望值：[僅需使用 Y 的邊際機率函數計算即可]

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

$$=\sum_y y\sum_x f(x,y)=\boxed{\sum_y yf_y(x)}$$

二元間斷隨機變數的變異數

- X 的變異數：[僅需使用 X 的邊際機率函數計算即可]

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

$$=\boxed{\sum_x(x-\mu_X)^2 f_x(x)}$$

$$=\sum_x(x^2-2\mu_X x+\mu_X^2)f_x(x)$$

$$=\sum_x x^2 f_x(x)-2\mu_X\sum_x xf_x(x)+\mu_X^2\sum_x f_x(x)$$

$$=\boxed{\sum_x x^2 f_x(x)-\mu_X^2}$$

$$=\boxed{E(X^2)-[E(X)]^2}$$

- Y 的變異數：[僅需使用 Y 的邊際機率函數計算即可]

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

$$=\boxed{\sum_y(y-\mu_Y)^2 f_y(y)}=\sum_y(y^2-2\mu_Y y+\mu_y^2)f_y(y)$$

$$=\sum_y y^2 f_y(y)-2\mu_Y\sum_y yf_y(y)+\mu_Y^2\sum_y f_y(y)$$

$$=\boxed{\sum_y y^2 f_y(y)-\mu_Y^2}$$

$$=\boxed{E(Y^2)-[E(Y)]^2}$$

- 例子：預期會購買手機的顧客($X=1$)比率為

$$E(X)=\sum_x xf_x(x)=0\times0.55+1\times0.45=0.45$$

而 X 的變異數為

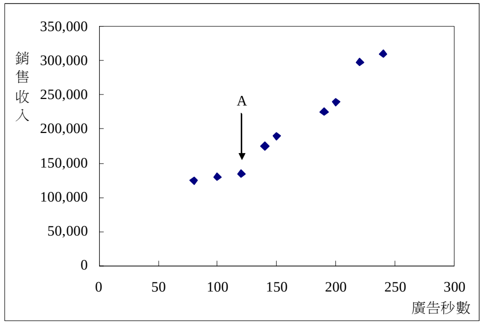
$$V(X)=\sum_x x^2 f_x(x)-\mu_X^2$$

$$=[0^2\times0.55+1^2\times0.45]-(0.45)^2=0.2475$$

兩變數間的關係

- **散佈圖**(scatter plot)可用來表示兩變數之資料的散佈情形。
- 例子：某公司統計了過去 9 星期廣告秒數與銷售收入資料如下表

	A	B	C
1	星期	廣告秒數	銷售收入
2	1	120	135,000
3	2	200	240,000
4	3	100	130,000
5	4	150	190,000
6	5	240	310,000
7	6	80	125,000
8	7	220	298,000
9	8	140	175,000
10	9	190	225,000



  - ➢ 將每個星期的廣告秒數標示於橫軸(X)、銷售收入標示於縱軸(Y)，畫出散佈圖：如 A 點標示出第 1 星期的資料位置($X=120$, $Y=13{,}500$)。由散佈圖可以看出，廣告秒數越高，銷售收入越高，兩變數呈正向關係。

- 變數間之相關性的分類：假設有兩個隨機變數 X 與 Y
  - ➢ 正向關係：當 X 越大(小)時，Y 也傾向越大(小)。
  - ➢ 負向關係：當 X 越大(小)時，Y 卻傾向越小(大)。
  - ➢ 無相關：若 X 與 Y 並無上述描述的關係。
- 一個比無相關更強的條件：**獨立**(independence)
  - ➢ 設 X,Y 為二元隨機變數，若知道 X 發生的數值並不影響我們對 Y 的機率分配之看法[$f(y\,|\,X=x)=f_y(y)$]，或是知道 X 發生的數值並不影響我們對 Y 的機率分配之看法[$f(x\,|\,Y=y)=f_x(x)$]，則稱 X,Y 為<u>獨立隨機變數</u>。
- 兩變數**獨立的條件**：設 X,Y 為二元隨機變數，若 X 與 Y 之值均滿足下列條件[滿足其中之一即會滿足全部]，則 X,Y 獨立。

(1) $f(x,y)=f_x(x)\cdot f_y(y)$ [聯合機率等於邊際機率的乘積]

(2) $f(x\,|\,y)=f_x(x)$ [給定 Y 時 X 的條件機率為 X 的邊際機率]

(3) $f(y\,|\,x)=f_y(y)$ [給定 X 時 Y 的條件機率為 Y 的邊際機率]

- 例子：顧客購買 3G 手機的行為(X)與看過簡訊(Y)兩者是否獨立？
  - ➢ 由於 $f(0,0)=0.33$，而

    $$f_x(0)\cdot f_y(0)=0.55\times 0.51=0.2805$$

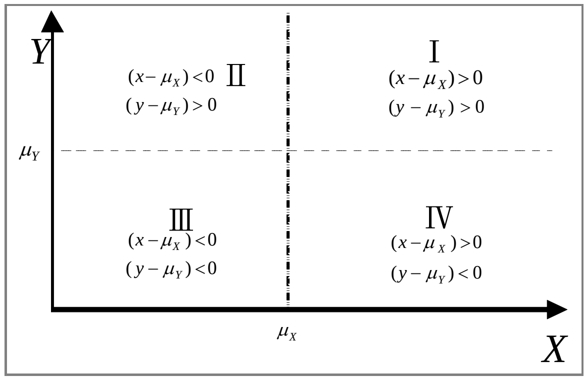
    故 $f(0,0)\neq f_x(0)\cdot f_y(0)$。因此購買 3G 手機的行為($X$)與看過簡訊($Y$)兩者並不獨立。[照理而言，我們應該檢驗聯合機率函數所有的函數值是否滿足獨立的條件；但事實上只要檢驗出有一個函數值違反獨立條件即可]
  - ➢ 兩個不獨立的隨機變數稱為**相依**(dependent)。
- 當 X,Y 兩個隨機變數為相依時(不互為獨立)，則 X,Y 間互有關係。變數間的關係可能是線性的，亦可能是非線性的；若變數間的關係為線性時，則變數間的關係可分為：
  - ➢ **正向關係**：當 X 變大(小)時，Y 也變大(小)。
  - ➢ **負向關係**：當 X 變大(小)時，Y 卻變小(大)。
  - ➢ 欲衡量變數間的線性相關程度，可用**共異數與相關係數**。

- **共變數**[共變異數](covariance)：兩隨機變數 X,Y 的共變數定義為

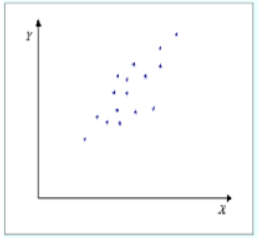
$$Cov(X,Y)=E[(X-\mu_X)(Y-\mu_Y)]$$

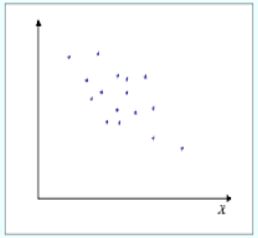
其中 μ_X、μ_Y 分別代表 X,Y 的平均數。

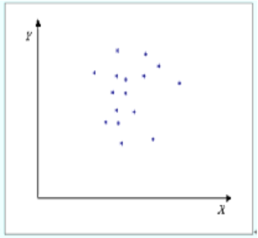
  - ➢ 若在 X 軸標示出 μ_X，經過該點畫一條垂直線，在 Y 軸標示出 μ_Y，經過該點畫一條水平線，則可將平面切割成四塊



- 當資料點 (X,Y) 落在 I、III 象限時，表示 X 小時 Y 也小(或 X 大時 Y 也大)，此時 $(X-\mu_X)(Y-\mu_Y)>0$，故 $Cov(X,Y)>0$
  因此，若 $Cov(X,Y)>0$，代表 X、Y 間具有正向的線性關係
- 當資料點 (X,Y) 落在 II、IV 象限時，表示 X 小時 Y 大(或 X 大時 Y 小)，此時 $(X-\mu_X)(Y-\mu_Y)<0$，故 $Cov(X,Y)<0$。
  因此，若 $Cov(X,Y)<0$，代表 X、Y 間具有反向的線性關係
- 若 X 與 Y 無關係，則資料點應平均散佈於四個象限，因此 $Cov(X,Y)=0$。故 $Cov(X,Y)=0$ 代表 X 與 Y 無線性關係。







- 若 X,Y 為間斷隨機變數，則其共變數之計算如下

$$Cov(X,Y)=E[(X-\mu_X)(Y-\mu_Y)]$$
$$=\sum_x\sum_y (x-\mu_X)(y-\mu_Y)f(x,y)$$

- 共變數的另一個公式[以間斷隨機變數為例]
  $Cov(X,Y)=E(XY)-E(X)E(Y)$
  證明：

$$Cov(X,Y)=E[(X-\mu_X)(Y-\mu_Y)]$$
$$=\sum_x\sum_y (x-\mu_X)(y-\mu_Y)f(x,y)$$
$$=\sum_x\sum_y (xy-\mu_X y-\mu_Y x+\mu_X\mu_Y)f(x,y)$$
$$=\sum_x\sum_y xyf(x,y)-\mu_X\sum_x\sum_y yf(x,y)$$
$$-\mu_Y\sum_x\sum_y xf(x,y)+\sum_x\sum_y \mu_X\mu_Y f(x,y)$$
$$=\sum_x\sum_y xyf(x,y)-\mu_X\sum_y y\sum_x f(x,y)-\mu_Y\sum_x x\sum_y f(x,y)+\mu_X\mu_Y$$

$$=\sum_x\sum_y xyf(x,y)-\mu_X\sum_y yf_y(y)-\mu_Y\sum_x xf_x(x)+\mu_X\mu_Y$$
$$=\sum_x\sum_y xyf(x,y)-\mu_X\mu_Y-\mu_X\mu_Y+\mu_X\mu_Y$$
$$=\sum_x\sum_y xyf(x,y)-\mu_X\mu_Y$$
$$=E(XY)-E(X)E(Y)$$

- 若 X 與 Y 獨立，則其共變數為 0。
  證明：若 X 與 Y 獨立，則 $f(x,y)=f_x(x)f_y(y)$，因此

$$E(XY)=\sum_x\sum_y xyf(x,y)=\sum_x\sum_y xyf_x(x)f_y(y)$$
$$=\sum_x xf_x(x)\sum_y yf_y(y)=\mu_Y\sum_x xf_x(x)=\mu_X\mu_Y$$

  故 $Cov(X,Y)=E(XY)-\mu_X\mu_Y=\mu_X\mu_Y-\mu_X\mu_Y=0$

- 若 X,Y 兩個隨機變數的變異數分別為 σ_X^2、σ_Y^2，兩者的共變數為 σ_{XY}。令 $W=aX+bY$，則 W 的變異數為

$$V(W)=\sigma_W^2=a^2\sigma_X^2+b^2\sigma_Y^2+2ab\sigma_{XY}$$

證明：因為 $E(W)=E(aX+bY)=a\mu_X+b\mu_Y$，故

$$V=E[W-E(W)]^2=E[aX+bY-E(aX+bY)]^2$$
$$=E[aX+bY-(a\mu_X+b\mu_Y)]^2=E[a(X-\mu_X)+b(Y-\mu_Y)]^2$$
$$=E[a^2(X-\mu_X)^2+b^2(Y-\mu_Y)^2+2ab(X-\mu_X)(Y-\mu_Y)]$$
$$=a^2E(X-\mu_X)^2+b^2E(Y-\mu_Y)^2+2abE[(X-\mu_X)(Y-\mu_Y)]$$
$$=a^2\sigma_X^2+b^2\sigma_Y^2+2ab\sigma_{XY}$$

- 共變數的優缺點：共變數在衡量共變關係時，有下列優缺點
  - 它可看出共變的方向與程度。
  - $-\infty<\mathrm{Cov}(X,Y)<\infty$，其值域無限，不易根據 $Cov(X,Y)$ 的大小來判斷 X 與 Y 的相關程度。
  - 它易受衡量單位的影響，例如當以公斤改為以磅衡量體重時，所得的共變數會改變。
  - 具有雙重的衡量單位(同時有 X 與 Y 的衡量單位不易解釋)

- 例子：購買 3G 手機的行為(X)與看過簡訊(Y)的共變數

$$Cov(X,Y)=E(XY)-E(X)E(Y)$$
$$=\sum_x\sum_y xyf(x,y)-(\sum_x xf_x(x))(\sum_y yf_y(y))$$
$$=0\times0\times0.33+0\times1\times0.22+01\times0\times0.18+1\times1\times0.27-(0.45)(0.49)$$
$$=0.0495$$

- **相關係數(coefficient of correlation)**：用來衡量變數間相關程度

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

式中 $Cov(X,Y)$ 為 X 、Y 的共變數，σ_X 、σ_Y 為 X 與 Y 的標準差

➢ 相關係數亦可計算如下

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

- 相關係數的一些特性：

➢ 相關係數的數值可證明是介於 -1 與 $+1$ 之間。

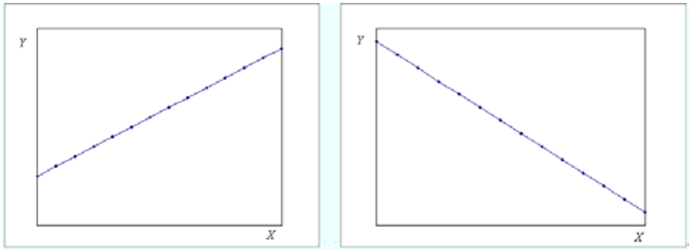
證明：由於

$$V(X-aY)=\sigma_X^2+a^2\sigma_Y^2-2a\sigma_{XY}\geq 0$$

若令 $a=\sigma_{XY}/\sigma_Y^2$ ，將其代入上式可得

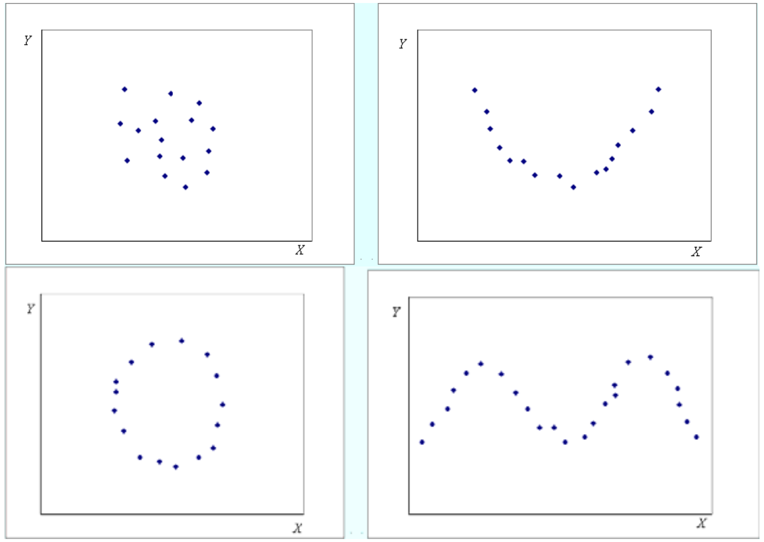
$$\sigma_X^2+\frac{(\sigma_{XY})^2}{(\sigma_Y^2)^2}\sigma_Y^2-2\frac{\sigma_{XY}}{\sigma_Y^2}\sigma_{XY}\geq 0$$
$$\Rightarrow \sigma_X^2+\frac{(\sigma_{XY})^2}{\sigma_Y^2}-2\frac{(\sigma_{XY})^2}{\sigma_Y^2}\geq 0$$
$$\Rightarrow \sigma_X^2-\frac{(\sigma_{XY})^2}{\sigma_Y^2}\geq 0\Rightarrow 1-\frac{(\sigma_{XY})^2}{\sigma_X^2\sigma_Y^2}\geq 0\Rightarrow \frac{(\sigma_{XY})^2}{\sigma_X^2\sigma_Y^2}\leq 1$$
$$\Rightarrow \left|\frac{\sigma_{XY}}{\sigma_X\sigma_Y}\right|\leq 1\Rightarrow -1\leq\frac{\sigma_{XY}}{\sigma_X\sigma_Y}\leq 1\Rightarrow \boxed{-1\leq\rho_{XY}\leq 1}$$

➢ 當 $\rho_{XY}=+1$ 時，我們稱 X、Y 具正的完全直線關係(**完全正相關**)，此時 (X,Y) 的點會落在一條斜率為正的直線上[如下左圖]。當 $\rho_{XY}=-1$ 時，我們稱 X 、Y 具負的完全直線關係(**完全負相關**)，此時 (X,Y) 的點會落在一條斜率為負的直線上[如下右圖]。



➢ 當 $\rho_{XY}=0$ 時， X 、Y 無線性相關[但不代表沒有非線性關係]。下頁四個圖都表示 X 、Y 無關。左上圖表示 X、Y 真的沒有任何關係[無線性與非線性相關]；右上圖表示 X、Y 沒有線性關係，但為拋物線關係。左下圖表示 X 、Y 無線

性相關，但具圓形關係。右下圖則顯示出 X、Y 具有明顯的非線性關係，但無線性關係。



- ➢ $|\rho_{XY}|$愈大表示 X、Y 的線性相關程度愈大。一般而言，若相關係數的數值在 0.5 左右稱為中度相關；若相關係數大於 0.7，則認為高度相關；若小於 0.3 則為低度相關。
- ➢ 若兩變數互為獨立，則其相關係數為 0 [因其共變數為 0]

- 例子：購買 3G 手機的行為(X)與看過簡訊(Y)的

由於$\sigma_X^2 = 0.2475$、$\sigma_Y^2 = 0.2499$、$\sigma_{XY} = 0.0495$，故

$$\rho_{XY} = \frac{\sigma_{XY}}{\sigma_X \sigma_Y} = \frac{0.0495}{\sqrt{0.2475}\sqrt{0.2499}} = \frac{0.0495}{0.4975 \times 0.4999} = 0.1990$$

相關係數、共變數與獨立的關係

- 若兩變數獨立，則兩個變數不但沒有線性關係，亦無非線性關係。兩變數獨立時，隱含其共變數為 0、相關係數亦為 0。
- 相關係數是衡量變數間的線性關係，無法衡量非線性關係。若兩變數的相關係數為 0，僅代表變數間無線性關係，不代表連非線性關係也沒有(獨立)。
- 相關係數的正負號與共變數相同

條件期望值(conditional expectation)與條件變異數

- 在$Y = y_j$ 條件下，X 發生 x_i 的條件機率為

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

除非 X、Y 互相獨立，否則條件機率函數 $f(x_i | Y = y_j)$ 應與邊際機率函數 $f_x(x_i)$ 不同。因此，在$Y = y_j$ 的條件下，X 的**(條件)期望值** $E(X | Y = y_j)$ 應與 X 的期望值 $E(X)$ 不同。[變異數亦應不同]

- 在 $X = x_i$ 條件下，Y 發生 y_j 的條件機率應不等於邊際機率 $f_y(y_j)$

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

因此，在 $X = x_i$ 的條件下，Y 的**(條件)期望值** $E(Y | X = x_i)$ 應與 Y 的期望值 $E(Y)$ 不同。[變異數亦應不同]

- **條件期望值**與**條件變異數**的定義

在 $Y = y_j$ 條件下，X 的條件期望值與條件變異數為

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

$$V(X | Y = y_j) = \sum_x [x - E(X | Y = y_j)]^2 f(x | Y = y_j)$$

在 $X = x_i$ 條件下，Y 的條件期望值與條件變異數為

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

$$V(Y | X = x_i) = \sum_y [y - E(Y | X = x_i)]^2 f(y | X = x_i)$$

- 例子：顧客看過簡訊($Y = 1$)後購買 3G 手機(X)的條件期望值為

$$E(X | Y = 1) = \sum_x x f(x | Y = 1) = 0 \times f(0 | Y = 1) + 1 \times f(1 | Y = 1)$$
$$= 0 \times 0.45 + 1 \times 0.55 = 0.55$$

條件變異數為$V(X | Y = 1) = \sum_x [x - E(X | Y = 1)]^2 f(x | Y = 1)$

$$= (0 - 0.55)^2 \times f(0 | Y = 1) + (1 - 0.55)^2 \times f(1 | Y = 1)$$
$$= (0 - 0.55)^2 \times 0.45 + (1 - 0.55)^2 \times 0.55 = 0.2475$$

補充：二元連續隨機變數

- **聯合累加分配函數**：若 X、Y 為連續隨機變數[$-\infty < X < \infty$、$-\infty < Y < \infty$]，其聯合累加分配函數定義為

$$F_{X,Y}(x, y) = P(X \le x \text{ 且 } Y \le y)$$

- **聯合機率密度函數**：若 $F_{X,Y}(x, y)$ 為連續函數且二次可微，則聯合機率密度函數定義為

$$f_{X,Y}(x, y) = \frac{\partial^2 F_{X,Y}(x, y)}{\partial x \partial y}$$

- ➢ 聯合機率密度函數必須滿足
  (1) $f_{X,Y}(x, y) \ge 0$
  (2) $\int_{-\infty}^{\infty} \int_{-\infty}^{\infty} f_{X,Y}(x, y) dxdy = 1$
- ➢ 聯合累加分配函數可由聯合機率密度函數求算

$$F_{X,Y}(x, y) = \int_{-\infty}^{x} \int_{-\infty}^{y} f_{X,Y}(r, s) drds$$

- **X 的邊際機率密度函數**：

$$f_X(x)=\int_{-\infty}^{\infty} f_{X,Y}(x,y)dy$$

X 的邊際累加分配函數

$$F_X(x)=\int_{-\infty}^{x} f_X(r)dr$$

- **Y 的邊際機率密度函數**：

$$f_Y(y)=\int_{-\infty}^{\infty} f_{X,Y}(x,y)dx$$

Y 的邊際累加分配函數

$$F_Y(y)=\int_{-\infty}^{y} f_Y(s)ds$$

- X 的**期望值**：

$$E(X)=\mu_X=\int_{-\infty}^{\infty}\int_{-\infty}^{\infty} xf_{X,Y}(x,y)dxdy$$

$$=\int_{-\infty}^{\infty} x\left(\int_{-\infty}^{\infty} f_{X,Y}(x,y)dy\right)dx=\boxed{\int_{-\infty}^{\infty} xf_X(x)dx}$$

X 的**變異數**：

$$V(X)=\int_{-\infty}^{\infty}\int_{-\infty}^{\infty}(x-\mu_X)^2 f_{X,Y}(x,y)dxdy$$

$$=\int_{-\infty}^{\infty}(x-\mu_X)^2\left(\int_{-\infty}^{\infty} f_{X,Y}(x,y)dy\right)dx=\boxed{\int_{-\infty}^{\infty}(x-\mu_X)^2 f_X(x)dx}$$

- Y 的**期望值**：

$$E(Y)=\mu_Y=\int_{-\infty}^{\infty}\int_{-\infty}^{\infty} yf_{X,Y}(x,y)dxdy$$

$$=\int_{-\infty}^{\infty} y\left(\int_{-\infty}^{\infty} f_{X,Y}(x,y)dx\right)dy=\boxed{\int_{-\infty}^{\infty} yf_Y(y)dy}$$

Y 的**變異數**：

$$V(Y)=\int_{-\infty}^{\infty}\int_{-\infty}^{\infty}(y-\mu_Y)^2 f_{X,Y}(x,y)dxdy$$

$$=\int_{-\infty}^{\infty}(y-\mu_Y)^2\left(\int_{-\infty}^{\infty} f_{X,Y}(x,y)dx\right)dy=\boxed{\int_{-\infty}^{\infty}(y-\mu_Y)^2 f_Y(y)dy}$$

- X 與 Y 的**交叉動差** $E(XY)=\int_{-\infty}^{\infty}\int_{-\infty}^{\infty} xyf_{X,Y}(x,y)dxdy$

- X 與 Y 的**共變數**

$$Cov(X,Y)=E[(X-\mu_X)(Y-\mu_Y)]$$

$$=\boxed{\int_{-\infty}^{\infty}\int_{-\infty}^{\infty}(x-\mu_X)(y-\mu_Y)f_{X,Y}(x,y)dxdy}$$

可進一步變為

$$Cov(X,Y)=\int_{-\infty}^{\infty}\int_{-\infty}^{\infty}(xy-\mu_Y x-\mu_X y+\mu_X\mu_Y)f_{X,Y}(x,y)dxdy$$

$$=\int_{-\infty}^{\infty}\int_{-\infty}^{\infty} xyf_{X,Y}(x,y)dxdy-\int_{-\infty}^{\infty}\int_{-\infty}^{\infty}\mu_Y xf_{X,Y}(x,y)dxdy$$

$$-\int_{-\infty}^{\infty}\int_{-\infty}^{\infty}\mu_X yf_{X,Y}(x,y)dxdy+\int_{-\infty}^{\infty}\int_{-\infty}^{\infty}\mu_X\mu_Y f_{X,Y}(x,y)dxdy$$

$$=E(XY)-\mu_Y\int_{-\infty}^{\infty}\int_{-\infty}^{\infty} xf_{X,Y}(x,y)dxdy-\mu_X\int_{-\infty}^{\infty}\int_{-\infty}^{\infty} yf_{X,Y}(x,y)dxdy$$

$$+\mu_X\mu_Y\int_{-\infty}^{\infty}\int_{-\infty}^{\infty} f_{X,Y}(x,y)dxdy$$

$$=E(XY)-\mu_X\mu_Y-\mu_X\mu_Y+\mu_X\mu_Y=\boxed{E(XY)-E(X)E(Y)}$$

- **條件機率密度函數**：在 $Y=y$ 的條件下，X 的條件機率密度函數

$$f_{X|Y}(x|Y=y)=\frac{f_{X,Y}(x,y)}{f_Y(y)}$$

在 $X=x$ 的條件下，Y 的條件機率密度函數

$$f_{Y|X}(y|X=x)=\frac{f_{X,Y}(x,y)}{f_X(x)}$$

- **獨立**的定義：若 X 與 Y 獨立，在 $Y=y$ 的條件下，X 的條件機率密度函數應與 X 的邊際機率密度函數相同

$$f_{X|Y}(x|Y=y)=\frac{f_{X,Y}(x,y)}{f_Y(y)}=f_X(x)$$

因此，若 X 與 Y 獨立，則其<u>聯合機率密度函數為邊際機率密度函數之乘積</u>

$$f_{X,Y}(x,y)=f_X(x)f_Y(y)$$

- 條件期望值與條件變異數

$$E(X\,|\,Y=y)=\int_{-\infty}^{\infty} x f_{X|Y}(x\,|\,Y=y)dx$$

$$V(X\,|\,Y=y)=\int_{-\infty}^{\infty}[x-E(X\,|\,Y=y)]^2 f_{X|Y}(x\,|\,Y=y)dx$$

$$E(Y\,|\,X=x)=\int_{-\infty}^{\infty} y f_{Y|X}(Y\,|\,X=x)dy$$

$$E(Y\,|\,X=x)=\int_{-\infty}^{\infty}[y-E(Y\,|\,X=x)]^2 f_{Y|X}(Y\,|\,X=x)dy$$

補充：多元隨機變數

- 若 $X_1,\ldots,X_n$ 為 n 個隨機變數，令 $f_{X_1,\ldots,X_n}(x_1,\ldots,x_n)$ 代表其**聯合機率函數**(或**聯合機率密度函數**)，令 $F_{X_1,\ldots,X_n}(x_1,\ldots,x_n)$ 代表其**多元累加分配函數**，則 $F_{X_1,\ldots,X_n}(x_1,\ldots,x_n)$ 可由聯合機率函數加總或聯合機率密度函數積分而得

$$F_{X_1,\ldots,X_n}(x_1,\ldots,x_n)=P(X_1\le x_1 \text{ 且 } X_2\le x_2 \text{ 且 } \cdots \text{ 且 } X_n\le x_n)$$

- **獨立**的條件：若 $X_1,\ldots,X_n$ 間相互獨立，則聯合機率函數(聯合機率密度函數)為邊際機率函數(邊際機率密度函數)的乘積

$$f_{X_1,\ldots,X_n}(x_1,\ldots,x_n)=f_{X_1}(x_1)f_{X_2}(x_2)\cdots f_{X_n}(x_n)$$

- 隨機變數之線性函數的期望值與變異數：若 $X_1,\ldots,X_n$ 為 n 個隨機變數，令 $W=a_1X_1+\cdots+a_nX_n$，則

$$E(W)=E(a_1X_1+\cdots+a_nX_n)=a_1E(X_1)+\cdots+a_nE(X_n)=a_1\mu_1+\cdots+a_n\mu_n$$

$$V(W)=V(a_1X_1+\cdots+a_nX_n)$$

$$=\sum_{i=1}^n a_i^2\sigma_i^2+\sum_{i=1}^n\sum_{\substack{j=1\\ j\ne i}}^n a_ia_j\sigma_{ij}$$

$\boxed{\text{證明}}$：根據變異數的定義

$$V(W)=E[(a_1X_1+\cdots+a_nX_n)-E(a_1X_1+\cdots+a_nX_n)]^2$$

$$=E[(a_1X_1+\cdots+a_nX_n)-(a_1\mu_1+\cdots+a_n\mu_n)]^2$$

$$=E[a_1(X_1-\mu_1)+\cdots+a_n(X_n-\mu_n)]^2$$

中括號內的平方展開後，有底下中括號中的 n^2 項

$$\begin{array}{c|cccc|}
 & a_1(X_1-\mu_1) & a_2(X_2-\mu_2) & & a_n(X_n-\mu_n)\\
\hline
a_1(X_1-\mu_1) & a_1^2(X_1-\mu_1)^2 & a_1a_2(X_1-\mu_1)(X_2-\mu_2) & \cdots & a_1a_n(X_1-\mu_1)(X_n-\mu_n)\\
a_2(X_2-\mu_2) & a_2a_1(X_2-\mu_2)(X_1-\mu_1) & a_2^2(X_2-\mu_2)^2 & \cdots & a_2a_n(X_2-\mu_2)(X_n-\mu_n)\\
\vdots & \vdots & \vdots & \vdots & \vdots\\
a_n(X_n-\mu_n) & a_nX_1(X_n-\mu_n)(X_1-\mu_1) & a_na_2(X_n-\mu_n)(X_2-\mu_2) & \cdots & a_n^2(X_n-\mu_n)^2
\end{array}$$

將中括號內的每一項取期望值可得

$$\begin{bmatrix}
a_1^2\sigma_1^2 & a_1a_2\sigma_{12} & \cdots & a_1a_n\sigma_{1n}\\
a_2a_1\sigma_{21} & a_2^2\sigma_2^2 & \cdots & a_2a_n\sigma_{2n}\\
\vdots & \vdots & \vdots & \vdots\\
a_na_1\sigma_{n1} & a_na_2\sigma_{n2} & \cdots & a_n^2\sigma_n^2
\end{bmatrix}$$

將對角線部分加總可得 $\sum_{i=1}^n a_i^2\sigma_i^2$

將非對角線部分加總可得 $\sum_{i=1}^n\sum_{\substack{j=1\\ j\ne i}}^n a_ia_j\sigma_{ij}$

因此 $V(W)=\sum_{i=1}^n a_i^2\sigma_i^2+\sum_{i=1}^n\sum_{\substack{j=1\\ j\ne i}}^n a_ia_j\sigma_{ij}$

- 特例 1：若 $X_1,\ldots,X_n$ 間相互獨立，則 $V(W)=\sum_{i=1}^n a_i^2\sigma_i^2$
- 特例 2：若 $X_1,\ldots,X_n$ 均為常態分配，則

  $W\sim N(E(W),V(W))$

- 其他與共變數有關的公式
  - $Cov(a+bX,c+dY)=cd\cdot Cov(X,Y)$
  - $Cov(aX+bY,cX+dY)$
    $=ac\cdot V(X)+bd\cdot V(Y)+(ad+bc)\cdot Cov(X,Y)$