第 7 章　離散隨機變數及其常用的機率分配

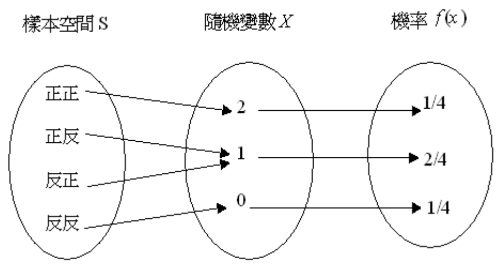
隨機變數的意義

- 某些隨機實驗的結果會產生數值(數量資料)，如擲一顆骰子的可能結果為 1、2、3、4、5、6 點；某些隨機實驗的結果卻是出現一些文字、符號(類別資料)，如擲銅板的可能結果為正面、反面。
  - ➢ 對於隨機實驗的結果，可依某種特性、或研究的重點，以某一變數(例如 X)來表示，稱為**隨機變數**(random variable)。
- **隨機變數**的意義：隨機變數是隨機實驗中對應於樣本點的實數值函數。[通常以大寫英文字母 X、Y、Z、...等表示，隨機變數的可能出現數值稱為**隨機變量**，以小寫英文字母 x、y、z、...表示]
- 例子 1：擲一顆骰子，令擲出的點數為 X，骰子點數為 1 時 $X=1$，骰子點數為 2 時 $X=2$，...，骰子點數為 6 時 $X=6$，則 X 為一隨機變數，隨機變量為 1、2、3、4、5、6。隨機變數 X 出現某特定數值(隨機變量)的機率，取決於擲骰子的結果。

- 例子 2：擲一枚銅板 1 次，當銅板出現正面時令 $X=1$，當銅板出現反面時令 $X=2$，則 X 為一隨機變數，隨機變量為 1、2。
- 例子 3：擲一枚銅板 2 次，令 X 為「出現正面的次數」，則其可能出現數值為 $x=0,1,2$。
  - ➢ X 就是隨機變數，隨機變量為 0,1,2。
  - ➢ 隨機變量是指其變量 x 的發生是『隨機的』；而隨機是指各變量的發生是『服從(follow)』(跟隨)某一機率分配。
  - ➢ 擲一枚銅板 2 次共有 4 個樣本點，每個樣本點會對應到一個隨機變量，隨機變數 X 與樣本點的函數關係如下

$$X(\omega)=\begin{cases}0 & 若\ \omega=(反,反)\\ 1 & 若\ \omega=(正,反)\ 或\ (反,正)\\ 2 & 若\ \omega=(正,正)\end{cases}$$

這就是『隨機變數是將隨機實驗的樣本點對應到實數值的函數』之意義，亦即 $X:S\to\mathbb{R}$。[參見下圖]



- ➢ 換句話說，隨機變數是隨機實驗結果的數值表現。而隨機變數出現某特定變量的機率取決於其背後的樣本點出現之機率。例如：當樣本點為(正,反)或(反,正)時，$X=1$，因此 $X=1$ 出現的機率就是 $\frac{1}{2}$ [(正,反)與(反,正)出現之機率]。

樣本點	正面的個數（x）	相對次數（機率）
（反，反）	0	$1/4=0.25$
（正，反）（反，正）	1	$2/4=0.50$
（正，正）	2	$1/4=0.25$
$N=4$		1.00

隨機變數的種類

- **間斷隨機變數**(discrete random variable)：隨機變數的變量個數是有限的，或個數是無限但可數的，稱為間斷或不連續隨機變數。
  - ➢ 間斷隨機變數的變量既然可數，就可將變量可能出現的數值明確列出，下表是幾個例子

隨機實驗	隨機變數	隨機變數 X 可能的值
1 枚銅板擲 2 次	出現正面的次數	0,1,2
抽取 10 個蘋果機檢查品質	不良品的個數	0,1,2,⋯,10
購買電扇顧客的性別	性別	0 爲男性，1 爲女性
出售的皮包數	銷售量	0,1,2,⋯
公車 263 路線 1 天的顧客	乘客人數	0,1,2,⋯

- **連續隨機變數**(continuous random variable)：隨機變數的變量個數為無限且不可數的，稱為連續隨機變數。
  - ➢ 由於連續隨機變數的變量個數不可數，因此通常以一個區間來代表變量的可能數值(可能出現的範圍)，下表是幾例

隨機實驗	隨機變數	隨機變數 X 可能的值
詢問陳先生的月薪	薪資收入	$x \geq 0$
觀察醫院病人候診時間	等候時間	$x \geq 0$
抽取 1 家電腦廠的年生產量	產量	$x \geq 0$
抽取 1,250ml 瓶裝汽水	汽水容量 ml	$0 \leq x \leq 1{,}250$

- 隨機變數可依變數的個數分為
  - **單一隨機變數**(univariate r.v.)：隨機變數個數只有一個。例如：令隨機變數 X 是丟兩次銅板出現正面的次數，由於變數只有 X 一個，故將 X 稱為單一隨機變數。
  - **二元隨機變數**(bivariate r.v.)：變數的數目為 2 時。例如抽取一個員工，令 X 為其薪資，Y 為其年資，X 與 Y 即稱為二元隨機變數。
  - **多元隨機變數**(multivariate r.v.)：變數數目為 3 個以上時。例如抽取一個員工，令 X 為其薪資，Y 為其年資，Z 為性別，D 為學歷，X、Y、Z、D 即稱為多元隨機變數。

單一間斷隨機變數的機率分配

- 間斷隨機變數的**機率分配**(probability distribution)：單一間斷隨機變數的機率分配是表示，間斷隨機變數的各個變量的發生機率(或相對次數)的分布情形。
  - 知道了間斷隨機變數的機率分配後，可預期各變量發生的機率，以及隨機變數在某一值域內的機率。
- 例子：擲一枚銅板 2 次，令 X 為「出現正面的次數」，則隨機變數 X 的機率分配如下

$$X(\omega)=\begin{cases}0 & \text{機率}=\frac{1}{4}\\ 1 & \text{機率}=\frac{1}{2}\\ 2 & \text{機率}=\frac{1}{4}\end{cases}$$

  - 根據這個分配可知，(1)隨機變數 X 出現數值 1 的機率為 $\frac{1}{2}$；(2)隨機變數 X 出現的數值小於等於 1 的機率為 $\frac{1}{4}+\frac{1}{2}=\frac{3}{4}$。

- 例子：台北市大安交通分隊想要了解配置多少人力才足以應付轄區內的交通事故案件，因此調查了去年每天的交通事故發生次數，整理出了底下的次數分配表

	A	B
1	每天交通事故件數	相對次數
2	0	0.37
3	1	0.31
4	2	0.18
5	3	0.09
6	4	0.04
7	5	0.01
8	合計	1

  - 由這個次數分配表可以得知每個交通事故件數(隨機變量)的機率分配，例如不超過 3 件事故的機率為 0.95，大安交通分隊即可依據這個數字來判斷是否需要增加人手。
  - **註**：你是否有個疑問，『利用過去的資訊來做為未來決策的參考是否恰當？』統計學的回應是：的確可能不恰當，但除了這個資訊，你應該找不到其他資訊可供參考了。

一些觀念補充：

- 隨機變數既然是將樣本空間中的樣本點對應到實數值的函數，那麼隨機變數的機率分配所描述的應該就是『**母體的行為**』，亦即長期實驗後才會具有的機率規則。
- 例如：擲一枚銅板 2 次，令 X 為「出現正面的次數」。一次隨機實驗指的是『擲一枚銅板 2 次』，做完一次隨機實驗，隨機變數的實現值一定是隨機變量(0,1,2)之一[此時沒有所謂的機率問題]；我們之前所描述之機率分配

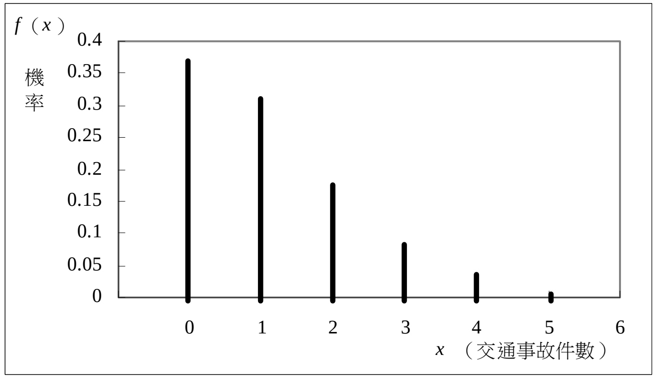
$$X(\omega)=\begin{cases}0 & \text{機率}=\frac{1}{4}\\ 1 & \text{機率}=\frac{1}{2}\\ 2 & \text{機率}=\frac{1}{4}\end{cases}$$

  的意義是：若我們進行無窮多次的隨機實驗，並記錄每一次實驗出現正面的次數；統計這些隨機實驗的結果後應可發現，正面次數為 0 的比例應接近 $\frac{1}{4}$，正面次數為 1 的比例應接近 $\frac{1}{2}$，正面次數為 2 的比例應接近 $\frac{1}{4}$ [長期的機率規則]

- 上例中的機率分配是根據**先驗機率理論**得到的，但未來的交通事故件數可不是先驗理論所能回答，因此我們以客觀機率理論(調查去年的交通事故發生件數)來得到機率分配。
- 若去年的每日交通事故發生件數的次數分配視為母體的理論分配[母體機率分配的估計]，定義隨機變數 X 為『每日交通事故發生件數』，隨機變數以 x 表示，並以 $f(x)$ 表示隨機變數 X 出現變量 x 的機率[視為 x 的函數]，可得機率分配

	A	B	C
1	隨機變量 x	相對次數	機率函數 $f(x)$
2	0	0.37	0.37
3	1	0.31	0.31
4	2	0.18	0.18
5	3	0.09	0.09
6	4	0.04	0.04
7	5	0.01	0.01
8	合計	$\Sigma=1.00$	$\Sigma f(x)=1.00$

- 機率分配亦可以圖形來表現



間斷隨機變數的機率函數

- 隨機變數是定義在樣本空間事件的實數值函數，間斷隨機變數的各變量代表某一事件。
  - 機率論中曾提及『事件的機率必須滿足機率的三個公理』，同樣的，隨機變數的各變量的機率亦必須滿足機率的公理

- **間斷隨機變數的機率函數**：設 X 為間斷隨機變數，其變量為 $x_1,\cdots,x_n$，對應 x_i 的每一數值有唯一機率與之對應，該機率值表為 $f(X=x_i)$ 或 $f(x_i)$，並滿足下列兩個條件：
  (1) $0\le f(x_i)\le 1$ [各變量的機率值介於 0 與 1 之間]
  (2) $\sum_{i=1}^{n} f(x_i)=1$ [各變量的機率總和等於 1]
  則 $f(x)$ 為 X 之**機率函數**或稱**機率分配**。
  - 間斷隨機變數各個變量所對應的機率均密集於各變量點上，故間斷的機率函數又稱為**機率密集函數**(probability mass function)。
- 例子：令 X 為「擲一枚銅板 2 次所出現的正面次數」，若定義：$f(X=0)=f(0)=\frac{1}{4}$、$f(X=1)=f(1)=\frac{1}{2}$、$f(X=2)=f(2)=\frac{1}{4}$。這的確就是機率函數，因 (1) 各變量定義唯一機率；(2) $0\le f(0)\le 1$、$0\le f(1)\le 1$、$0\le f(2)\le 1$；(3) $f(0)+f(1)+f(2)=1$

有時我們喜歡將機率函數表達為[各種變量 x 所對應的機率值]

$$f(x)=\begin{cases}\frac{1}{4} & 若\ x=0\\ \frac{1}{2} & 若\ x=1\\ \frac{1}{4} & 若\ x=2\end{cases}$$

間斷隨機變數的累加機率函數[課本上的定義是錯誤的]

- **累加機率函數**(cumulative probability function)或稱**累加分配函數**(cumulative distribution function)：隨機變數 X 的累加機率函數是由實數線 $\mathbb{R}$ 映射到 $[0,1]$ 的函數，其函數值為[若機率函數以小寫字母 f 表示，則累加機率函數以大寫字母 F 表示]
  $F(x)=P(X\le x)$
- **間斷隨機變數的累加機率函數**：設間斷隨機變數 X 之變量為 $x_1,\cdots,x_n$，則隨機變數 X 的累加機率函數為
  $F(x)=P(X\le x)=\sum_{i:x_i\le x} f(x_i)$
  [小於 x 的所有變量之機率加總]

- 累加機率函數 $F(x)$ 的特性(若變量 $x_1,\cdots,x_n$ 已由小至大排序)
  - $F(a)=0 \quad \forall a<x_1$ [小於最小變量的累加機率等於 0]
  - $F(b)=1 \quad \forall b>x_n$ [累加到最大變量 x_n 以上時的累加機率等於 1]
  - 如果 $c\ge d$，則 $F(c)\ge F(d)$。[累加機率函數為<u>非遞減函數</u>]
  - $f(x_i)=F(x_i)-F(x_{i-1})$，x_{i-1} 為 x_i 的前一個變量，$x_{i-1}<x_i$。[某變量的機率可以該變量的累加機率減掉前一個變量的累加機率求得]
- 例子：「擲一枚銅板 2 次所出現的正面次數」之累加機率函數

$$f(x)=\begin{cases}\frac{1}{4} & 若 x=0\\ \frac{1}{2} & 若 x=1\\ \frac{1}{4} & 若 x=2\end{cases} \qquad F(x)=\begin{cases}0 & 若 x<0\\ \frac{1}{4} & 若 0\le x<1\\ \frac{3}{4} & 若 1\le x<2\\ 1 & 若 2\le x\end{cases}$$

- 例子：交通事故件數的例子中，我們所定義的機率確實是機率函數，因為每個變量均有唯一的機率值與之對應、所有變量的機率值均介於 0 與 1 之間、且所有變量的機率值總和為 1。

$$\sum_{i=1}^{n} f(x_i)=f(0)+f(1)+f(2)+f(3)+f(4)+f(5)$$
$$=0.37+0.31+0.18+0.09+0.04+0.01=1$$

若知道機率函數，即可計算隨機變數位於某個範圍內的機率，如

$$P(2\le X\le 3)=f(2)+f(3)=0.18+0.09=0.27$$

	A	B	C
1	隨機變量 x	機率函數 $f(x)$	累加機率函數 $F(x)$
2	0	0.37	0.37
3	1	0.31	0.68
4	2	0.18	0.86
5	3	0.09	0.95
6	4	0.04	0.99
7	5	0.01	1
8	合計	$\sum=1.00$	

上表中的累加機率函數不正確，正確答案更正如下[圖 7.3 不正確，需重畫]

$$F(x)=\begin{cases}0 & 若 x<0\\ 0.37 & 若 0\le x<1\\ 0.68 & 若 1\le x<2\\ 0.86 & 若 2\le x<3\\ 0.95 & 若 3\le x<4\\ 0.99 & 若 4\le x<5\\ 1 & 若 5\le x\end{cases}$$

間斷隨機變數的期望值

- 隨機變數的平均數又稱為期望值。
- **期望值**(expected value)：期望值是指如果我們不斷的進行多次的實驗，預期會發生或觀察得到的數值或結果。[長期實驗中預期會發生的數值]

- **間斷隨機變數的期望值**：若間斷隨機變數 X 之變量為 $x_1,\cdots,x_n$，每個變量發生的機率分別為 $f(x_1),\ldots,f(x_n)$，則隨機變數 X 的期望值為(各個隨機變量以其發生之機率為權數的加權平均)

$$E(X)=\sum_{i=1}^{n} x_i f(x_i)=\mu$$

[符號說明] 不管隨機變數是間斷的或連續的，$E(X)$ 都是用來代表對隨機變數 X <u>取期望值(平均)的動作</u>；我們通常用符號 μ 來代表一個隨機變數的母體平均數(期望值)。

[公式說明] 若你手上的 N 筆資料確實是隨機變數 X 長期實驗的結果，則這 N 筆資料應該有 $N\cdot f(x_1)$ 筆是 x_1，有 $N\cdot f(x_2)$ 筆是 x_2，…，有 $N\cdot f(x_n)$ 筆是 x_n；所以這 N 筆資料的平均為

$$\frac{N\cdot f(x_1)\cdot x_1+N\cdot f(x_2)\cdot x_2+\cdots+N\cdot f(x_n)\cdot x_n}{N}$$
$$=x_1 f(x_1)+x_2 f(x_2)+\cdots+x_n f(x_n)=\sum_{i=1}^{n} x_i f(x_i)$$

- 例子：預期每天會有多少件交通事故？
  - 每日交通事故的期望值為

$$E(X)=\sum_{i=1}^{n} x_i f(x_i)$$
$$=0\times 0.37+1\times 0.31+2\times 0.18+3\times 0.09+4\times 0.04+5\times 0.01=1.15$$

  - 交通分隊預期平均每天應處理 1.15 件交通事故。

	A	B	C
1	隨機變量 x	機率函數 $f(x)$	$xf(x)$
2	0	0.37	0
3	1	0.31	0.31
4	2	0.18	0.36
5	3	0.09	0.27
6	4	0.04	0.16
7	5	0.01	0.05
8	合計	$\sum=1.00$	$\sum xf(x)=1.15$

- 機率函數是經過長期試行或先驗、主觀而獲得，因此代表母體的分配，故根據機率函數所計算的期望值為『母體平均數』。

間斷隨機變數的變異數

- **間斷隨機變數的變異數**：若間斷隨機變數 X 之變量為 $x_1,\cdots,x_n$，變量發生的機率分別為 $f(x_1),\ldots,f(x_n)$，則隨機變數 X 的變異數

$$\sigma^2=V(X)\equiv E[(X-\mu)^2]=\sum_{i=1}^{n}(x_i-\mu)^2 f(x_i)$$

[符號說明] σ^2 與 $V(X)$ 是用來代表隨機變數 X 之變異數的符號，變異數定義為離均差平方的平均，亦即 $E[(X-\mu)^2]$。

[公式說明] 若有 N 筆資料確實是隨機變數 X 長期實驗的結果，則這 N 筆資料應該有 $N\cdot f(x_1)$ 筆是 x_1，有 $N\cdot f(x_2)$ 筆是 x_2，…，有 $N\cdot f(x_n)$ 筆是 x_n；所以這 N 筆資料離均差平方的平均為

$$\frac{N\cdot f(x_1)\cdot(x_1-\mu)^2+N\cdot f(x_2)\cdot(x_2-\mu)^2+\cdots+N\cdot f(x_n)\cdot(x_n-\mu)^2}{N}$$
$$=f(x_1)\cdot(x_1-\mu)^2+f(x_2)\cdot(x_2-\mu)^2+\cdots+f(x_n)\cdot(x_n-\mu)^2$$
$$=\sum_{i=1}^{n}(x_i-\mu)^2 f(x_i)$$

- 變異數又可表達為 $V(X)=E[(X-\mu)^2]=E(X^2)-[E(X)]^2$

證明：

$$V(X)=E[(X-\mu)^2]=\sum_{i=1}^{n}(x_i-\mu)^2 f(x_i)$$
$$=\sum_{i=1}^{n}[(x_i^2-2\mu x_i+\mu^2)f(x_i)]$$
$$=\sum_{i=1}^{n}x_i^2 f(x_i)-2\mu\sum_{i=1}^{n}x_i f(x_i)+\mu^2\sum_{i=1}^{n}f(x_i)$$
$$=\sum_{i=1}^{n}x_i^2 f(x_i)-2\mu\cdot\mu+\mu^2=\sum_{i=1}^{n}x_i^2 f(x_i)-\mu^2$$
$$=E(X^2)-[E(X)]^2$$

式中利用 $\sum_{i=1}^{n} f(x_i)=1$，且以符號 $E(X^2)$ 表示 $\sum_{i=1}^{n}x_i^2 f(x_i)$，以 $E(X)$ 表示 μ。

- 間斷隨機變數的標準差：將變異數開根號即得標準差

$$\sigma=\sqrt{\sigma^2}=\sqrt{\sum_{i=1}^{n}(x_i-\mu)^2 f(x_i)}$$

- 例子：每日交通事故的分散程度。欲計算每日交通事故的變異數，可利用

$$\sigma^2=E(X^2)-[E(X)]^2=\sum_{i=1}^{n}x_i^2 f(x_i)-\mu^2$$

由下表可知 $\mu=1.15$、$\sum_{i=1}^{n}x_i^2 f(x_i)=2.73$，故變異數為

$$\sigma^2=\sum_{i=1}^{n}x_i^2 f(x_i)-\mu^2$$
$$=2.73-1.15^2=1.407$$

而標準差則為 $\sigma=\sqrt{\sigma^2}=\sqrt{1.407}=1.186$

	A	B	C	D	E
1	x	$f(x)$	$xf(x)$	x^2	$x^2f(x)$
2	0	0.37	0	0	0
3	1	0.31	0.31	1	0.31
4	2	0.18	0.36	4	0.72
5	3	0.09	0.27	9	0.81
6	4	0.04	0.16	16	0.64
7	5	0.01	0.05	25	0.25
8	合計	$\sum=1.00$	$\sum xf(x)=1.15$		$\sum x^2f(x)=2.73$

標準化隨機變數

- 不同的隨機變數的中心位置與分散程度不一，其平均數、標準差的數值並不相同，因此在比較上會產生困難。
  - ➢ 若能將各個隨機變數標準化，將有助於判斷各變量與平均數、標準差之間的關係，同時亦能顯示各變量與平均數的距離及各變量的機率。
  - ➢ 標準化：『透過函數將隨機變數轉換成另一個平均數為 0、變異數為 1 的隨機變數』；將隨機變數減掉其平均數後，再除以其標準差就可完成標準化的動作。
- 標準化隨機變數(Z 變數)：設 X 為一隨機變數，其平均數為 μ，變異數為 σ^2，令

$$Z=\frac{X-\mu}{\sigma}$$

則 Z 為一標準化隨機變數。

- 例子：每日交通事故發生件數的標準化。
  - ➢ 由於每日交通事故發生件數的平均數為 1.15、標準差為 1.19，因此要將每日交通事故發生件數這個隨機變數標準化，必須利用

$$Z=\frac{X-\mu}{\sigma}=\frac{X-1.15}{1.19}$$

  - ➢ 對應於隨機變量 $X=5$ 的標準化變量為 $Z=3.325$，表示隨機變量 $X=5$ 位於平均數 $\mu=1.15$ 右方 3.325 個標準差的位置。

隨機變數 X	標準化隨機變數 $Z=\frac{X-\mu}{\sigma}=\frac{X-1.15}{1.19}$
0	-0.966
1	-0.126
2	0.714
3	1.555
4	2.395
5	3.235

隨機變數之函數的期望值

- 若 X 為一隨機變數，則該隨機變數之函數 $h(X)$ 亦為隨機變數。
  - ➢ 若 X 為間斷隨機變數，變量為 $x_1,\cdots,x_n$，各變量的機率分別為 $f(x_1),\ldots,f(x_n)$。則 $h(X)$ 的各變量分別為 $h(x_1),\cdots,h(x_n)$，$h(X)$ 之各變量發生的機率亦為 $f(x_1),\ldots,f(x_n)$。[隨機變數 X 之函數 $h(X)$ 的機率性質，與隨機變數 X 之機率性質相同]
  - ➢ 例子：假設間斷隨機變數 X 的各變量發生之機率如下左所示，定義 $h(X)=1+X^2$，則 $h(X)$ 之各變量及其發生之機率則如下右所示

$$X=\begin{cases}0 & \text{機率} f(0)=\frac{1}{4}\\ 1 & \text{機率} f(1)=\frac{1}{2}\\ 2 & \text{機率} f(2)=\frac{1}{4}\end{cases}\qquad h(X)=\begin{cases}=1+0^2=1 & \text{機率} f(0)=\frac{1}{4}\\ =1+1^2=2 & \text{機率} f(1)=\frac{1}{2}\\ =1+2^2=5 & \text{機率} f(2)=\frac{1}{4}\end{cases}$$

- **隨機變數之函數的期望值**：設 X 為間斷隨機變數，其機率函數為 $f(x)$。令 $h(X)$ 為 X 的函數，則 $h(X)$ 的期望值表示為 $E[h(X)]$ 或 $\mu_{h(x)}$：

$$\mu_{h(x)}=E[h(X)]=\sum_{i=1}^{n}h(x_i)f(x_i)$$

- 一些有用的性質：設 C 為常數，$h(X)$ 為 X 的函數，$h_1(X),h_2(X),\cdots,h_k(X)$ 亦均為 X 的函數，則
  - ➢ $E(C)=C$ [常數的期望值為該常數本身]

    證明：$E(C)=\sum_{i=1}^{n}C\cdot f(x_i)=C\sum_{i=1}^{n}f(x_i)=C$
  - ➢ $E[C\cdot h(X)]=C\cdot E[h(X)]$ [常數項與隨機變數函數的乘積之期望值，等於該常數與該隨機變數函數之期望值的乘積]

    證明：$E[C\cdot h(X)]=\sum_{i=1}^{n}C\cdot h(x_i)\cdot f(x_i)=C\sum_{i=1}^{n}h(x_i)\cdot f(x_i)$
    $=C\cdot E[h(X)]$

- ➢ $E[h_1(X)+h_2(X)+\cdots+h_k(X)]=E[h_1(X)]+\cdots+E[h_k(X)]$

  證明：

$$
\begin{aligned}
&E[h_1(X)+h_2(X)+\cdots+h_k(X)]\\
&=\sum_{i=1}^{n}[h_1(x_i)+h_2(x_i)+\cdots+h_k(x_i)]f(x_i)\\
&=\sum_{i=1}^{n}h_1(x_i)f(x_i)+\sum_{i=1}^{n}h_2(x_i)f(x_i)+\cdots+\sum_{i=1}^{n}h_k(x_i)f(x_i)\\
&=E[h_1(X)]+E[h_2(X)]+\cdots+E[h_k(X)]
\end{aligned}
$$

- 例子：X 為間斷隨機變數，證明 $V(X)=E[(X-\mu)^2]=E(X^2)-\mu^2$

$$
\begin{aligned}
V(X)&=E[(X-\mu)^2]=E(X^2-2\mu X+\mu^2)\\
&=E(X^2)-E(2\mu X)+E(\mu^2)\\
&=E(X^2)-2\mu\cdot E(X)+\mu^2\\
&=E(X^2)-2\mu\cdot\mu+\mu^2\\
&=E(X^2)-\mu^2
\end{aligned}
$$

- ➢ **一個有用的性質**：因 $\sigma^2=E(X^2)-\mu^2$，所以 $E(X^2)=\sigma^2+\mu^2$
- 例子：設 X 為間斷隨機變數，其平均數為 5，變異數為 2，試求 $E(5)$、$E(5X)$、$E(X^2+2X+1)$。
  - ➢ $E(5)=5$
  - ➢ $E(5X)=5E(X)=5\times5=25$
  - ➢ 由於 $E(X^2)=\sigma^2+\mu^2=2+5^2=27$，所以

$$
\begin{aligned}
E(X^2+2X+1)&=E(X^2)+E(2X)+E(1)\\
&=E(X^2)+2E(X)+1=27+2\times5+1=38
\end{aligned}
$$

- **隨機變數之函數的變異數**：設 X 為間斷隨機變數，其機率函數為 $f(x)$，令 $h(X)$ 為 X 的函數，則 $h(X)$ 的變異數為：

$$\sigma^2_{h(X)}=V[h(X)]=E[h(X)-E[h(X)]]^2$$

  - ➢ 依然照隨機變數之變異數的定義，只不過現在的隨機變數是 $h(X)$，其平均數為 $E[h(X)]$。

特例：線性函數的期望值與變異數

- 現在假設函數 $h(X)$ 為線性，亦即

$$h(X)=a+bX$$

- **線性函數的期望值**：設 $Y=a+bX$，則 Y 的期望值(平均數)為：

$$E(Y)=E(a+bX)=a+bE(X)$$

  [對隨機變數 X 的線性函數取期望值時，將期望值符號 E 直接移往隨機變數 X 之前即可]

  證明：假設 X 為間斷隨機變數

$$
\begin{aligned}
E(Y)&=E(a+bX)\\
&=\sum_{i=1}^{n}(a+bx_i)\cdot f(x_i)=a\sum_{i=1}^{n}f(x_i)+b\sum_{i=1}^{n}x_i\cdot f(x_i)=a+bE(X)
\end{aligned}
$$

- **線性函數的變異數**：設 $Y=a+bX$，則 Y 的變異數為：

$$V(Y)=V(a+bX)=V(bX)=b^2V(X)$$

  [隨機變數 X 之線性函數的常數項 a 並不影響變異數，X 之係數 b 則將變異數放大成 b^2 倍]

證明：假設 X 為間斷隨機變數

$$
\begin{aligned}
V(Y)&=V(a+bX)=E[(a+bX)-E(a+bX)]^2\\
&=E[(a+bX)-a-E(bX)]^2=E[bX-E(bX)]^2=\boxed{V(bX)}\\
&=E[bX-bE(X)]^2=E[b(X-E(X))]^2\\
&=b^2E[X-E(X)]^2=\boxed{b^2V(X)}
\end{aligned}
$$

- 標準化隨機變數 Z 的平均數與變異數分別為

$$
\begin{aligned}
E(Z)&=E\left(\frac{X-\mu}{\sigma}\right)=E\left(-\frac{\mu}{\sigma}+\frac{X}{\sigma}\right)=-\frac{\mu}{\sigma}+E\left(\frac{X}{\sigma}\right)\\
&=-\frac{\mu}{\sigma}+\frac{E(X)}{\sigma}=-\frac{\mu}{\sigma}+\frac{\mu}{\sigma}=0
\end{aligned}
$$

$$V(Z)=V\left(\frac{X-\mu}{\sigma}\right)=V\left(-\frac{\mu}{\sigma}+\frac{X}{\sigma}\right)=V\left(\frac{X}{\sigma}\right)=\frac{V(X)}{\sigma^2}=\frac{\sigma^2}{\sigma^2}=1$$

- 例子：定義 $h(X)=1+3X$ ，則 $h(X)$ 之各變量及其發生之機率

$$X=\begin{cases}0 & \text{機率} f(0)=\frac{1}{4}\\ 1 & \text{機率} f(1)=\frac{1}{2}\\ 2 & \text{機率} f(2)=\frac{1}{4}\end{cases} \qquad h(X)=\begin{cases}=1+3\cdot 0=1 & \text{機率} f(0)=\frac{1}{4}\\ =1+3\cdot 1=4 & \text{機率} f(1)=\frac{1}{2}\\ =1+3\cdot 2=7 & \text{機率} f(2)=\frac{1}{4}\end{cases}$$

$E(X)=0\times\frac{1}{4}+1\times\frac{1}{2}+2\times\frac{1}{4}=1$

$V(X)=(0-1)^2\times\frac{1}{4}+(1-1)^2\times\frac{1}{2}+(2-1)^2\times\frac{1}{4}=\frac{1}{2}$

$E[h(X)]=E[1+3X]=1+3E(X)=1+3\times 1=4$

驗證： $E[h(X)]=E[1+3X]=1\times\frac{1}{4}+4\times\frac{1}{2}+7\times\frac{1}{4}=4$

$V[h(X)]=V(1+3X)=3^2\cdot V(X)=9\times\frac{1}{2}=\frac{9}{2}$

驗證：$V[h(X)]=(1-4)^2\times\frac{1}{4}+(4-4)^2\times\frac{1}{2}+(7-4)^2\times\frac{1}{4}=\frac{9}{2}$

- 若定義 $h(X)=1+X^2$ ，則

  $E[h(X)]=E(1+X^2)=1+E(X^2)=1+[V(X)+[E(X)]^2]=1+[\frac{1}{2}+1^2]=\frac{5}{2}$

$V[h(X)]=V(1+X^2)=V(X^2)=E[X^2-E(X^2)]^2$

$=E[X^4-2X^2E(X^2)+[E(X^2)]^2]=E[X^4]-[E(X^2)]^2$

$=[0^4\times\frac{1}{4}+1^4\times\frac{1}{2}+2^4\times\frac{1}{4}]-[\frac{1}{2}+1^2]^2=\frac{9}{4}$

- 柴比氏定理：設 X 為一隨機變數，其平均數為 μ，變異數為 σ^2 ，則對任何大於 1 的正數 $k>1$ 而言

$$P(|X-\mu|<k\sigma)\geq 1-\frac{1}{k^2} \quad \text{或} \quad P(|x-\mu|\geq k\sigma)\leq\frac{1}{k^2}$$

證明：假設 X 為間斷隨機變數。根據變異數的定義可知

$\sigma^2=\sum_{i=1}^{n}(x_i-\mu)^2 f(x_i)$

我們將加總的變量拆成兩部分，一部分的變量與平均數 μ 的差距至少 k 個標準差，亦即屬於集合 $S=\{x:|x-\mu|\geq k\sigma\}$ 的所有變量，另一部分的變量與平均數 μ 的差距小於 k 個標準差，亦即屬於集合 S 的補集 $\overline{S}=\{x:|x-\mu|<k\sigma\}$ 的所有變量。根據這個拆解

$$\begin{aligned}\sigma^2&=\sum_{i=1}^{n}(x_i-\mu)^2 f(x_i)\\ &=\sum_{x_i\in s}(x_i-\mu)^2 f(x_i)+\sum_{x_i\in\overline{s}}(x_i-\mu)^2 f(x_i)\\ &\geq\sum_{x_i\in s}(x_i-\mu)^2 f(x_i)\\ &\geq\sum_{x_i\in s}k^2\sigma^2 f(x_i) \qquad \text{若} x_i\in S,\text{則}(x_i-\mu)^2\geq k^2\sigma^2\\ &=k^2\sigma^2\sum_{x_i\in s}f(x_i)\\ &=k^2\sigma^2\cdot P(|x-\mu|\geq k\sigma)\end{aligned}$$

因此 $P(|x-\mu|\geq k\sigma)\leq\frac{1}{k^2}$ ，又由於 $\overline{S}=\{x:|x-\mu|<k\sigma\}$ 為 $S=\{x:|x-\mu|\geq k\sigma\}$ 的補集，故

$P(|X-\mu|<k\sigma)=1-P(|x-\mu|\geq k\sigma)\geq 1-\frac{1}{k^2}$

二項機率分配(binomial probability distribution)

- **伯努利(Bernoulli)隨機實驗**：只有兩種可能出象的隨機實驗。
  - **伯努利隨機變數**：只有兩個隨機變量的隨機變數；通常根據類別資料定義出來，變量則定義為 0 [失敗]與 1 [成功]。
  - 例如：擲一枚銅板，僅有正面、反面兩種可能出象，所以丟銅板為一個伯努利隨機實驗；若出現反面定義 X=0 [失敗]，出現正面則定義 X=1 [成功]，則 X 為伯努利隨機變數。
  - 若伯努利隨機變數出現變量 $X=1$ 的機率為 p，出現變量 $X=0$ 的機率為 $(1-p)$ ，則伯努利隨機變數的

    期望值為 $E(X)=1\times p+0\times(1-p)=p$

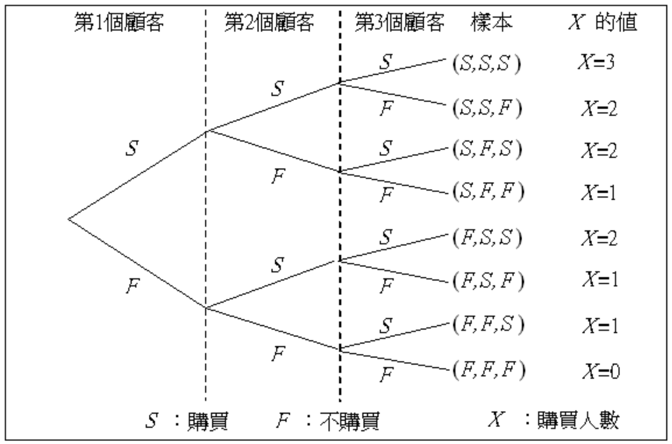
    變異數為 $V(X)=(1-p)^2\times p+(0-p)^2\times(1-p)$

    $=p(1-p)^2+p^2(1-p)$

    $=p(1-p)[(1-p)+p]=p(1-p)$

- **二項隨機實驗**：若一個隨機實驗包含 n 個獨立相同的伯努利隨機實驗，則稱該實驗為二項隨機實驗(binomial random experiment)。
- **二項隨機實驗的特性**：
  - 實驗中包含 n 次相同的試行(伯努利隨機實驗)
  - 每一次試行只有二種互斥的可能結果，不是成功(表示為 S)，就是失敗(表示為 F)。
  - 成功的機率為 $P(S)=p$，失敗的機率為 $P(F)=1-p$ (或表為 q)，且每次試行的機率均相同。
  - 每一次試行是獨立的。
  - 隨機變數定義為 n 次試行中成功的次數。
- **二項隨機變數**(binomial random variable)：若隨機變數 X 定義成『二項隨機實驗中成功的次數』，隨機變數 X 稱為二項隨機變數。其分配則稱為二項機率分配。
  - 例子：(1)從手機外殼成品中抽取 20 個，計算不良品個數；(2)餐廳人員從顧客中抽取 15 位，調查滿意全餐的人數。

- 例子：根據最近的經驗，汽車公司劉先生推銷小客車成功的機率為 0.2。現在假設有三個客戶來店參觀，則劉先生推銷成功的顧客數為 0 個、1 個、2 個、3 個的機率各為何？
  - 3 位潛在客戶買與不買的行為有 8 種可能，如樹枝圖所示



- 將隨機變數 X 定義為『3 次推銷中客戶購買的次數』，則 X 為二項隨機變數。
- 每次試行成功的機率為 p=0.2，失敗的機率為 $q=1-p=0.8$
- 0 個顧客購買的機率：僅有 (F,F,F) 這 1 [C_0^3]種狀況，該狀況發生的機率為 $0.8\times0.8\times0.8=0.2^0\times0.8^3=0.512=p^0q^3$，因此

$$f(X=0)=C_0^3p^0q^{3-0}=\tfrac{3!}{0!3!}\times0.2^0\times0.8^3=0.512$$

- 1 個顧客購買的機率：有 (S,F,F)、(F,S,F)、(F,F,S) 等 3 [C_1^3]種狀況，3 種狀況發生的機率都相同

(S,F,F) ：$0.2\times0.8\times0.8=0.2^1\times0.8^2=0.128=pq^2$

(F,S,F) ：$0.8\times0.2\times0.8=0.2^1\times0.8^2=0.128=pq^2$

(F,F,S) ：$0.8\times0.8\times0.2=0.2^1\times0.8^2=0.128=pq^2$

所以

$$f(X=1)=C_1^3p^1q^{3-1}=\tfrac{3!}{1!(3-1)!}\times0.2^1\times0.8^2=0.384$$

- 2 個顧客購買的機率：有 (S,S,F)、(S,F,S)、(F,S,S) 等 3 [C_2^3]種狀況，3 種狀況發生的機率都相同

(S,S,F) ：$0.2\times0.2\times0.8=0.2^2\times0.8^1=0.032=p^2q$

(S,F,S) ：$0.2\times0.8\times0.2=0.2^2\times0.8^1=0.032=p^2q$

(F,S,S) ：$0.8\times0.2\times0.2=0.2^2\times0.8^1=0.032=p^2q$

所以

$$f(X=2)=C_2^3p^2q^{3-2}=\tfrac{3!}{2!(3-2)!}\times0.2^2\times0.8^1=0.096$$

- 3 個顧客購買的機率：僅有 (S,S,S) 這 1 [C_0^3]種狀況，該狀況發生的機率為 $0.2\times0.2\times0.2=0.2^3\times0.8^0=0.008=p^3q^0$，因此

$$f(X=3)=C_3^3p^3q^{3-3}=\tfrac{3!}{3!(3-3)!}\times0.2^3\times0.8^0=0.008$$

- 二項隨機變數各變量之發生機率：
  - 具有 $\underline{n}$ 個獨立伯努利試行的二項隨機變數 X，其隨機變量為 $x = 0,1,2,\ldots,n$ $[n+1$個$]$。
  - 假設每個獨立試行之成功機率為 p，失敗機率為 $q = 1-p$。
  - 隨機變量 $X = x$ 代表的是『n 次試行中，有 x 個成功，$(n-x)$ 個失敗』。n 次試行中，恰好有 x 個成功的情況共有 C_x^n 種，而每種情況發生的機率為 $p^x q^{n-x}$，故

$$P(X = x) = C_x^n p^x q^{n-x}$$

- **二項機率分配**：設 X 為一間斷隨機變數，若其機率函數 $f(x)$ 為：

$$f(x) = C_x^n p^x q^{n-x} = \frac{n!}{x!(n-x)!} p^x q^{n-x} \qquad x = 0,1,2,\cdots,n$$

則 $f(x)$ 為**二項機率分配**。式中，n：試行次數，x：成功的次數，p：成功的機率，$q = 1-p$：失敗的機率。

  - 二項機率分配有兩個參數：n 與 p。
  - 通常以符號 $B(X;n,p)$ 來表示隨機變數 X 具有參數 n 與 p 的二項機率分配。
  - 課本最後的附表一(p.776~778)列出二項機率分配個變量的機率值。例如：要得到汽車銷售例子中的機率值，可查表中的 $n=3$、$p=0.2$ 的機率值。
  - 除了查表外，亦可以 EXCEL 中的指令『BINOMDIST』計算出二項機率分配各變量的機率值(參見課本 p.198)。
- 上述定義的二項隨機變數之機率確實是機率函數：因為
  - $0 \le C_x^n p^x q^{n-x} \le 1$
  - $\sum_{x=0}^n f(x) = 1$

$$\sum_{x=0}^n f(x) = \sum_{x=0}^n C_x^n p^x q^{n-x} = (p+q)^n$$
$$= [p + (1-p)]^n = 1$$

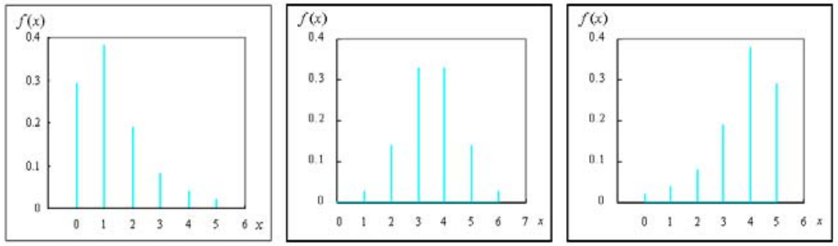
- 二項機率分配的**累加機率函數**：

$$F(x) = P(X \le x) = \sum_{x_i \le x} f(x_i)$$

  - 例子：至少有兩個客戶購買汽車的機率為

$$P(X \ge 2) = 1 - P(X \le 1) = 1 - F(1)$$
$$= 1 - [f(0) + f(1)] = 1 - [0.512 + 0.384] = 0.104$$

- 二項隨機變數之參數 n、p 與分配之型態
  - 當 $p = 0.5$ 時，無論 n 為多少，二項分配為一對稱分配(中圖)
  - 當 $p < 0.5$ 時，二項分配為一右偏分配(左圖)
  - 當 $p > 0.5$ 時，二項分配為一左偏分配(右圖)



  - 當 n 固定，p 越接近 0.5 時，二項分配越接近對稱分配。
  - 若 p 固定，當 n 持續增加時，二項分配越接近對稱分配。
  - 當 $n=1$ 時，二項分配又稱伯努利分配(或點二項分配)。
- **二項隨機變數的期望值**：$E(X) = np$

$$E(X) = \sum_{x=0}^n x f(x) = \sum_{x=0}^n x C_x^n p^x q^{n-x}$$
$$= \sum_{x=1}^n x C_x^n p^x q^{n-x} \qquad \text{因為當} x = 0 \text{時}, x C_x^n p^x q^{n-x} = 0$$
$$= \sum_{x=1}^n x \tfrac{n!}{x!(n-x)!} p^x q^{n-x}$$
$$= \sum_{x=1}^n \tfrac{n\cdot(n-1)!}{(x-1)!(n-x)!} p \cdot p^{x-1} q^{(n-1)-(x-1)}$$
$$= np \sum_{x=1}^n \tfrac{(n-1)!}{(x-1)![(n-1)-(x-1)]!} p^{x-1} q^{(n-1)-(x-1)}$$
$$= np \sum_{y=0}^{n-1} \tfrac{(n-1)!}{y![(n-1)-y)!} p^y q^{(n-1)-y} \qquad \text{定義} y = x-1$$
$$= np(p+q)^{n-1}$$
$$= np$$

- **二項隨機變數的變異數**：$V(X)=npq$

  若以 $\mu=np$ 代表二項隨機變數的平均數，則根據變異數的定義

  $$V(X)=\sum_{x=0}^{n}(x-\mu)^2 f(x)=\sum_{x=0}^{n}x^2 f(x)-\mu^2$$

  $$=\sum_{x=0}^{n}x(x-1)f(x)+\sum_{x=0}^{n}xf(x)-\mu^2$$

  其中 $\sum_{x=0}^{n}x(x-1)f(x)=\sum_{x=0}^{n}x(x-1)C_x^n p^x q^{n-x}$

  $$=\sum_{x=0}^{n}x(x-1)\frac{n!}{x!(n-x)!}p^x q^{n-x}$$

  $$=\sum_{x=2}^{n}x(x-1)\frac{n!}{x!(n-x)!}p^x q^{n-x}\qquad \text{因為當}x=0,1\text{時}, x(x-1)=0$$

  $$=\sum_{x=2}^{n}\frac{n(n-1)\cdot(n-2)!}{(x-2)![(n-2)-(x-2)]!}p^2\cdot p^{x-2}q^{(n-2)-(x-2)}$$

  $$=n(n-1)p^2\sum_{x=2}^{n}\frac{(n-2)!}{(x-2)![(n-2)-(x-2)]!}p^{x-2}q^{(n-2)-(x-2)}$$

  $$=n(n-1)p^2\sum_{y=0}^{n-2}\frac{(n-2)!}{y![(n-2)-y]!}p^y q^{(n-2)-y}\qquad \text{定義}y=x-2$$

  $$=n(n-1)p^2[p+q]^{n-2}=n(n-1)p^2$$

且 $\sum_{x=0}^{n}xf(x)=\mu=np$ ，故

$$V(X)=n(n-1)p^2+np-n^2p^2=[n^2p^2-np^2]+np-n^2p^2$$

$$=np-np^2=np(1-p)=npq$$

- **二項隨機變數的標準差**：$\sigma=\sqrt{npq}$
- 例子：劉先生預期會有幾個顧客買車。由於 $n=3$、 $p=0.2$，故

  $E(X)=np=3\times0.2=0.6$

  而顧客購車人數的變異數為

  $V(X)=npq=3\times0.2\times0.8=0.48$

  - 推廣：若劉先生向其推銷的人數有 100 人，則預期應會有 $E(X)=np=100\times0.2=20$ 個人會跟他買車。購車人數的變異數為 $V(X)=npq=100\times0.2\times0.8=16$。
- **補充**：若將隨機變數 X 定義為『二項隨機實驗中成功的比例』，則 X 的分配依然為二項分配，其隨機變量為 $x=0,\frac{1}{n},\frac{2}{n},...,\frac{n-1}{n},1$。

超幾何分配(hypergeometric probability distribution)

- 假設有一有限母體，其元素共有 N 個，元素分為兩類，其中一類具有某種特質，共有 K 個，另一類不具有某種特質，共有 $N-K$ 個。現在自該母體中抽取 n 個(每次抽取 1 個，抽出後不放回，抽取 n 次)[$n\le N$]，若抽出的具有某種特質視為成功，抽出另一類則視為失敗。令 X 為 n 次中的成功次數，則此隨機實驗為**超幾何實驗**，X 為**超幾何隨機變數**，其機率函數為**超幾何分配**。

$N=$母體元素總數	隨機抽取	$n=$樣本數
$K=$成功的次數	$\rightarrow$	$x=$成功的次數
$N-K=$失敗的次數	抽出不放回	$n-x=$失敗的次數

  - 例如：一個籃子中有 N 個球，其中 K 個為紅色， $N-K$ 個為黑色。從籃子中抽取 n 個球[$n\le N$]，令 X 為 n 個球中紅球(成功)的個數，則 X 為超幾何隨機變數，其隨機變量為 $x=0,1,2,...,n$。

  - N 個元素抽取 n 個共有 C_n^N 種組合，若抽出的 n 個元素中成功次數為 $X=x$ [失敗的次數為 $n-x$]，則共有 $C_x^K C_{n-x}^{N-K}$ 種組合，因此 n 次中的成功次數為 x 的機率為

  $$P(X=x)=\frac{C_x^K C_{n-x}^{N-K}}{C_n^N}$$

- **超幾何分配**：設 X 為一間斷隨機變數，若其機率函數 $f(x)$ 為：

  $$f(x)=\frac{C_x^K C_{n-x}^{N-K}}{C_n^N}\text{ ， }x=0,1,\cdots,n\text{ ，但 }x\le K\text{ 且 }n-x\le N-K$$

  則 $f(x)$ 為**超幾何分配機率函數**。式中，N：母體元素個數，K：定義為成功的元素個數， $N-K$ ：定義為失敗的元素個數，n：抽出的元素個數，x：抽出的 n 個元素中成功的個數。

  - 若你所抽出的元素個數 $n>K$，則 n 個元素中成功的個數最多僅有 K 個(所以有 $x\le K$ 這個限制條件)，失敗的個數最多也僅有 $N-K$ 個(所以有 $n-x\le N-K$ 這個限制條件)

- 驗證超幾何分配之機率函數滿足 $\sum_{x=0}^{n} f(x)=1$
  - ➢ 若 $n\le K$ 且 $n\le N-K$，則 $\sum_{x=0}^{n} C_x^K C_{n-x}^{N-K} = \boxed{C_{x+(n-x)}^{K+(N-K)}} = \boxed{C_n^N}$

證明：我們已知道從 N 個元素中抽取 n 個，共有 C_n^N 種組合。而這 C_n^N 種組合，我們可依其所包含的成功個數細分為底下 $n+1$種情況：

$x=0$ 個成功、$n-x=n$ 個失敗：有 $C_0^K C_{n-0}^{N-K}$ 種組合

$x=1$ 個成功、$n-x=n-1$ 個失敗：有 $C_1^K C_{n-1}^{N-K}$ 種組合

$x=2$ 個成功、$n-x=n-2$ 個失敗：有 $C_2^K C_{n-2}^{N-K}$ 種組合

⋮

$x=n$ 個成功、$n-x=0$ 個失敗：有 $C_n^K C_0^{N-K}$ 種組合

這 $n+1$種情況已經考慮了所有可能情況，因此其組合數加起來應該等於 C_n^N，亦即

$$\sum_{x=0}^{n} C_x^K C_{n-x}^{N-K} = C_0^K C_{n-0}^{N-K} + C_1^K C_{n-1}^{N-K} + \cdots + C_n^K C_{n-n}^{N-K}$$
$$= C_{x+(n-x)}^{K+(N-K)} = C_n^N$$

- ➢ 超幾何分配之機率函數滿足 $\sum_{x=0}^{n} f(x)=1$

$$\sum_{x=0}^{n} f(x) = \sum_{x=0}^{n} \frac{C_x^K C_{n-x}^{N-K}}{C_n^N} = \frac{1}{C_n^N}\sum_{x=0}^{n} C_x^K C_{n-x}^{N-K} = \frac{C_n^N}{C_n^N} = 1$$

- 例子：假設 18 個應徵工作者中有 12 個為男性。隨機抽取 3 個，其中 1 個為男性的機率為何？
  - ➢ 猛一看會覺得這是一個二項隨機實驗，因為抽出的應徵者不是男性就是女性；但這個問題並不是二項隨機實驗，因為抽出的應徵者兩兩之間並非獨立(抽出第一個之後會影響到第二個是男性或女性的機率)。故須以超幾何分配求解

    註：超幾何實驗其實可視為一序列相依之伯努力隨機試驗，這也就是超幾何實驗與二項實驗就大的差別

- ➢ 1 個男性都沒被選中($x=0$)的機率：

$$f(0) = \frac{C_x^K C_{n-x}^{N-K}}{C_n^N} = \frac{C_0^{12} C_{3-0}^{18-12}}{C_3^{18}} = \frac{\frac{12!}{0!12!}\frac{6!}{3!3!}}{\frac{18!}{3!15!}} = \frac{1\times 20}{816} = 0.0245$$

- ➢ 1 個男性被選中($x=1$)的機率：

$$f(1) = \frac{C_x^K C_{n-x}^{N-K}}{C_n^N} = \frac{C_1^{12} C_{3-1}^{18-12}}{C_3^{18}} = \frac{\frac{12!}{1!11!}\frac{6!}{2!4!}}{\frac{18!}{3!15!}} = \frac{12\times 15}{816} = 0.2206$$

- ➢ 2 個男性被選中($x=2$)的機率：

$$f(2) = \frac{C_x^K C_{n-x}^{N-K}}{C_n^N} = \frac{C_2^{12} C_{3-2}^{18-12}}{C_3^{18}} = \frac{\frac{12!}{2!10!}\frac{6!}{1!5!}}{\frac{18!}{3!52!}} = \frac{66\times 6}{816} = 0.4853$$

- ➢ 3 個男性被選中($x=3$)的機率：

$$f(3) = \frac{C_x^K C_{n-x}^{N-K}}{C_n^N} = \frac{C_3^{12} C_{3-3}^{18-12}}{C_3^{18}} = \frac{\frac{12!}{3!9!}\frac{6!}{0!6!}}{\frac{18!}{3!52!}} = \frac{220\times 1}{816} = 0.2696$$

- **超幾何分配之平均數**：$E(X) = n\cdot\frac{K}{N}$
- **超幾何分配之變異數**：$V(X) = n\cdot\frac{K}{N}\cdot\frac{N-K}{N}\cdot\frac{N-n}{N-1}$
  - ➢ 超幾何分配之平均數與變異數的證明技巧跟二項分配差不多，請自行參閱光碟當中的(第 7 章)數學附錄。
- 例子：令 X 為前例中被選中的男性人數，則 X 的期望值為

  $E(X) = n\cdot\frac{K}{N} = 3\cdot\frac{12}{18} = 2$

  變異數為

  $V(X) = n\cdot\frac{K}{N}\cdot\frac{N-K}{N}\cdot\frac{N-n}{N-1} = 3\cdot\frac{12}{18}\cdot\frac{18-12}{18}\cdot\frac{18-3}{18-1} = \frac{30}{51} = 0.588$
- 當 $N\to\infty$ 時，超幾何分配趨近於二項分配。正式結論：當 $N\to\infty$ 時，若 $\lim_{N\to\infty}\frac{K}{N} = p$，則 $P(X=x) = \lim_{N\to\infty}\frac{C_x^K C_{n-x}^{N-K}}{C_n^N} = C_x^n p^x (1-p)^{n-x}$。
  - ➢ 超幾何隨機實驗是抽出後不放回的實驗，若母體元素個數 N 很多，抽出後不放回對成功的機率應該影響不大，因此當 $N\to\infty$ 時，把每次抽出都視為獨立相同的伯努利試驗。

泊松分配(Poisson distribution)

- **動機**：假設某工廠所生產的電腦不良率僅有 $p = 0.01$，若我們抽取 1000 台電腦加以檢測，令 X 為「不良品的個數」，則 X 為二項隨機變數，其期望值為 $\mu = np = 1000 \cdot 0.01 = 10$，且

$$P(X = x) = C_x^{1000} 0.01^x (1-0.01)^{1000-x} = \tfrac{1000!}{x!(1000-x)!} 0.01^x (1-0.01)^{1000-x}$$

欲進行這種計算非常麻煩[如 $x = 1$ 時，要計算 $(1-0.01)^{999}$]。

上述機率的**另一種解釋**：假設工廠一天生產 1000 台電腦，我們站在生產線前作電腦檢測，則 $P(X = x) = C_x^{1000} 0.01^x (1-0.01)^{1000-x}$ 亦可解釋成『一天(一段期間)內所發現的不良品個數為 x 的機率』

  - **泊松公式**：當二項隨機變數的期望值 $\mu = np$ 不大不小(且 n 很大、p 很小)，則二項機率 $P(X = x)$ 可用底下公式近似

$$P(X = x) = \frac{\mu^x e^{-\mu}}{x!}$$

- 我們經常需要衡量『某事件在一段時間(或空間)內發生』的機率
  - 一個小時內打進來的電話次數
  - 一天內通過泰山收費站的汽車數
  - 一小時內到銀行提款的人數

  這些事件都是在一定的連續區間(時間)內發生，且事件的發生彼此互相獨立(或可視為互相獨立)，稱為泊松隨機實驗。

- **泊松隨機實驗**具有底下三個特性：
  - 在某區間發生的事件個數，與另一區間發生的事件個數是獨立的。例如要研究 2:30~5:00 間來店喝咖啡的人數，則任何兩個不重疊的時段(如 2:30~3:00 與 3:30~5:00)事件的發生是獨立的。
  - 在一段區間發生事件個數的期望值(平均數)與區間大小成正比。例如：若 1 小時內來店喝咖啡的平均人數為 20 人，則 2 小時內來店喝咖啡的平均人數為 40 人。

  - 在很短的區間內，事件僅可能發生 1 次或不發生。例如：在 0.1 秒非常微小的時間內，來店喝咖啡的人數可能僅有 1 人或一個也沒有。
- **泊松隨機變數**：令隨機變數 X 代表『一段時間(或空間)內某事件發生的次數』，且事件的發生滿足泊松隨機實驗的特性，則 X 稱為泊松隨機變數，其機率分配為泊松分配。
- **泊松分配**：設已知在一定的區間發生事件 A 的期望值為 λ，令 X 為該區間發生事件的次數，若 X 具有底下之機率函數 $f(x)$：

$$f(x) = \frac{\lambda^x e^{-\lambda}}{x!}, \quad x = 0, 1, 2, \cdots, \infty$$

則稱 X 具有泊松分配，其參數為 λ。

- e^λ 的 MacLaurin Series：將 $f(\lambda) = e^\lambda$ 對 $\lambda = 0$ 作泰勒展開式

$$e^\lambda = f(0) + \tfrac{f'(0)}{1!}\lambda + \tfrac{f''(0)}{2!}\lambda^2 + \tfrac{f^{(3)}(0)}{3!}\lambda^3 +$$

$$= 1 + \tfrac{\lambda}{1!} + \tfrac{\lambda^2}{2!} + \tfrac{\lambda^3}{3!} + \cdots = \sum_{x=0}^{\infty} \tfrac{\lambda^x}{x!}$$

- 證明泊松分配之機率函數滿足 $\sum_{x=0}^{\infty} f(x) = 1$

$$\sum_{x=0}^{\infty} f(x) = \sum_{x=0}^{\infty} \tfrac{\lambda^x e^{-\lambda}}{x!} = e^{-\lambda} \sum_{x=0}^{\infty} \tfrac{\lambda^x}{x!} = e^{-\lambda} \cdot e^{\lambda} = 1$$

- 例子：若每 5 分鐘到達 ATM 的平均顧客人數約為 1.5 人，則在上午 9:00~9:30 分間，有 5 位顧客到達 ATM 交易的機率為何？
  - 每 5 分鐘平均有 1.5 人到達 ATM，30 分鐘平均有 $1.5 \times 6 = 9$ 人到達 ATM，若以 X 代表『30 分鐘內到達 ATM 的顧客人數』，則 X 應具泊松分配，其機率函數為

$$f(x) = \frac{\lambda^x e^{-\lambda}}{x!} = \frac{9^x e^{-9}}{x!}$$

  - 30 分鐘內有 5 位顧客到達 ATM 的機率 $P(X = 5)$ 為

$$P(X = 5) = f(5) = \frac{9^5 e^{-9}}{5!} = 0.0607$$

  - 這個機率可自行計算，或查課本後的附表二(p.779~780)。

- 例子：若在早上的上班尖峰時間 6:30~8:30，平均每小時發生 2 件車禍，則在此尖峰時間內發生 4 件車禍的機率為何？
  - 平均每小時發生 2 件車禍，兩小時內平均發生 $\lambda = 4$ 件車禍。令 X 為『尖峰時間內發生的車禍件數』，則 X 應具泊松分配，其機率函數為

    $$f(x) = \frac{\lambda^x e^{-\lambda}}{x!} = \frac{4^x e^{-4}}{x!}$$

    因此，發生 4 件車禍的機率為

    $$P(X=4) = f(4) = \frac{4^4 e^{-4}}{4!} = 0.1954$$

  - 至少發生兩件車禍的機率為

    $$P(X \ge 2) = 1 - P(X < 2) = 1 - [f(0) + f(1)]$$
    $$= 1 - [0.0183 + 0.0773] = 0.9044$$

- **泊松分配之期望值**：$E(X) = \lambda$

$$E(X) = \sum_{x=0}^{\infty} x \frac{e^{-\lambda}\lambda^x}{x!} = \lambda \sum_{x=1}^{\infty} \frac{e^{-\lambda}\lambda^{x-1}}{(x-1)!} \underbrace{=}_{令y=x-1} \lambda e^{-\lambda} \sum_{y=0}^{\infty} \frac{\lambda^y}{y!} = \lambda e^{-\lambda} e^{\lambda} = \lambda$$

- **泊松分配之變異數**：$V(X) = \lambda$

變異數可計算如下

$$V(X) = E[X^2] - [E(X)]^2 = E[X(X-1)] + E(X) - [E(X)]^2$$

而 $E(X) = \lambda$，且

$$E[X(X-1)] = \sum_{x=0}^{\infty} x(x-1) \frac{e^{-\lambda}\lambda^x}{x!} = \lambda^2 \sum_{x=2}^{\infty} \frac{e^{-\lambda}\lambda^{x-2}}{(x-2)!}$$

$$\underbrace{=}_{令y=x-2} \lambda^2 e^{-\lambda} \sum_{y=0}^{\infty} \frac{\lambda^y}{y!} = \lambda^2 e^{-\lambda} \cdot e^{\lambda} = \lambda^2$$

故

$$V(X) = E[X(X-1)] + E(X) - [E(X)]^2 = \lambda^2 + \lambda - \lambda^2 = \lambda$$

- 泊松分配與二項分配的關係：若二項分配的 n [試行次數]很大 p [成功機率]很小時，二項分配會趨近泊松分配[可用泊松分配來計算二項分配的機率值]。
  - n 該多大？p 該多小？

    $n \ge 20$ 時，$np \le 1$；$n \ge 50$ 時，$np \le 5$；$n \ge 100$ 時，$np \le 10$
  - 下圖比較二項分配($n = 100$ 、 $p = 0.1$)與泊松分配($\lambda = np = 10$)的差異，可看出差異其實不大。