03機率分配

陳欣得

靜宜大學企管系

1

03機率分配

機率分配（***probability distribution***）可以想像成相對次數表的特殊版本，其分組所依據的變數為有好好定義的隨機變數（*random variable*），多數情況機率函數（***probability function***）可以寫成數學函數，或者說，存在可以描述該分配的機率空間（*probability space*）。本章介紹幾個常用的機率分配，包含這些分配之隨機變數的意義、以及計算相關統計量值的方法。

章節安排：

2

§3.1 機率分配概論

機率分配（***probability distribution***）是隨機變數的機率函數。一個機率分配有以下四個特性：

(1)機率分配可以想像為相對次數分配表，並以隨機變數的可能數值為分組依據；

(2)機率可以寫成隨機變數的數學函數；

(3)離散隨機變數若可能數值太多，或連續隨機變數有無限多可能數值，則無法寫出相對次數表，只能想像；

(4)我們以隨機變數的區間來定義事件。

以下是一個離散隨機變數的機率分配，寫成相對次數分配表如下表

x	*P(x)*
0	1/16
1	4/16
2	6/16
3	4/16
4	1/16
	1

機率函數（***probability function***）寫成函數形式如：$P(x)=C_x^4\times(1/2)^4$，$x=0,1,\ldots,4$。連續隨機變數的機率分配有無限多分組，無法寫出相對次數分配表的形式，只能想像；機率值寫成$P(x)=f(x)dx$，這裡$f(x)$稱為**機率密度函數**（***probability density function***）。

令$E(X)$為隨機變數X的**期望值**（***expected value***），$\mathrm{var}(X)$為其**變異數**（***variance***），又令$\mathrm{cov}(X,Y)$為隨機變數X與隨機變數Y的**共變數**（***covariance***）。

定義 3.1　離散隨機變數的期望值、變異數與共變數

若 X、Y 為離散隨機變數，則

$$E(X)=\mu_x=\sum_{\forall x} xP(x)$$

$$\operatorname{var}(X)=\sigma_x^2=E\left((x-\mu_x)^2\right)=\sum_{\forall x}(x-\mu_x)^2 P(x)$$

$$\operatorname{cov}(X,Y)=\sigma_{xy}=E\left((x-\mu_x)(y-\mu_y)\right)=\sum_{\forall y}\sum_{\forall x}(x-\mu_x)(y-\mu_y)P(x,y)$$

其中， $P(x,y)$ 為 X、Y 的**聯合機率函數**（***joint probability function***）。

定義 3.2　連續隨機變數的期望值、變異數與共變數

若 X、Y 為連續隨機變數，則

$$E(X)=\mu_x=\int_x xf(x)dx$$

$$\operatorname{var}(X)=\sigma_x^2=E\left((x-\mu_x)^2\right)=\int_x (x-\mu_x)^2 f(x)dx$$

$$\operatorname{cov}(X,Y)=\sigma_{xy}=E\left((x-\mu_x)(y-\mu_y)\right)=\iint_{x,y}(x-\mu_x)(y-\mu_y)f(x,y)dxdy$$

其中，$f(x,y)$ 為 X、Y 的**聯合機率密度函數**(***joint probability density function***)。

§3.1.1 期望值、變異數與共變數的操作

機率分配（***probability distribution***）可以想像成相對次數表的特殊版本，其分組所依據的變數為有好好定義的隨機變數（*random variable*），多數情況機率函數（***probability function***）可以寫成數學函數，或者說，存在可以描述該分配的機率空間（*probability space*）。本章介紹幾個常用的機率分配，包含這些分配之隨機變數的意義、以及計算相關統計量值的方法。

定理3.3　變異數與共變數的運算

對於隨機變數 X、Y，以下等式成立：

$\operatorname{var}(X)=E(X^2)-E(X)^2$；

$\operatorname{cov}(X,Y)=E(XY)-E(X)E(Y)$。

我們不提供定理 3.3 的證明。但可以這樣想，對於每一數值的出現機率都相等之離散隨機變數，也就是有 N 個出現機率皆為 $1/N$ 數值，那麼我們可以用次數分配表的方式來計算變異數與共變數。計算過程如下：

$$\operatorname{var}(X)=\frac{\Sigma(x-\bar{x})^2}{N}=\frac{\Sigma x^2-(\Sigma x)^2/N}{N}=\Sigma\tfrac{1}{N}x^2-\left(\Sigma\tfrac{1}{N}x\right)^2=E(X^2)-E(X)^2,$$

$$\operatorname{cov}(X)=\frac{\Sigma(x-\bar{x})(y-\bar{y})}{N}=\Sigma\tfrac{1}{N}xy-\left(\Sigma\tfrac{1}{N}x\right)\left(\Sigma\tfrac{1}{N}y\right)=E(XY)-E(X)E(Y)。$$

可以得到定理 3.3 的結果。

定理3.4　一些隨機變數函數之期望值運算E()結果

對於隨機變數 X、Y，令 $a,b\in R$，f、g 為 X、Y 的函數，以下等式成立：

$E(a)=a$；

$E(aX)=aE(X)$；

$E(aX+bY)=aE(X)+bE(Y)$；

$E\left[f(X)+g(X)\right]=E\left[f(X)\right]+E\left[g(X)\right]$。

定理3.5　一些隨機變數函數之變異數運算var()結果

對於隨機變數 X、Y，令 $a,b\in R$，以下等式成立：

$\operatorname{var}(a)=0$；

$\operatorname{var}(aX)=a^2\operatorname{var}(X)$；

$\operatorname{var}(aX+bY)=a^2\operatorname{var}(X)+b^2\operatorname{var}(Y)+2ab\operatorname{cov}(X,Y)$。

定理3.6　一些隨機變數函數之共變數運算cov()結果

對於隨機變數 X、Y、Z，令 $a,b\in R$，以下等式成立：

$\operatorname{cov}(X,X)=\operatorname{var}(X)$；

$\operatorname{cov}(a,X)=\operatorname{cov}(X,a)=0$；

$\operatorname{cov}(aX,bY)=ab\operatorname{cov}(X,Y)$；

$\operatorname{cov}(X+Y,Z)=\operatorname{cov}(X,Z)+\operatorname{cov}(Y,Z)$。

範例3.1　（期望值E()的操作）

已知 $E(X)=8$、$E(Y)=6$，求 $E(3X-2Y)$。

【解】

$$E(3X-2Y)=3E(X)-2E(Y)=3\times 8-2\times 6=12$$

範例3.2　（變異數var()的操作）

已知 $\text{var}(X)=8$ 、 $\text{var}(Y)=6$ ，且 X 與 Y 互相獨立，求 $\text{var}(3X-2Y)$ 。

【解】

X 與 Y 互相獨立隱含 $\text{cov}(X,Y)=0$ ，因此

$$\text{var}(3X-2Y)=3^2\,\text{var}(X)+(-2)^2\,\text{var}(Y)=9\times 8+4\times 6=96$$

13

範例3.3　（共變數cov()的操作）

已知 $\text{var}(X)=3$ 、 $\text{var}(Y)=5$ ， $\text{cov}(X,Y)=2$ ，求 $\text{cov}(X+4Y,X+2)$ 。

【解】

先計算

$$(X+4Y)\times(X+2)=X^2+4XY+2X+8\text{Y}$$

等號兩邊取共變數操作，則

$$\begin{aligned}\text{cov}(X+4Y,X+2)&=\text{cov}(X,X)+\text{cov}(4X,Y)+\text{cov}(2,X)+\text{cov}(8,Y)\\&=\text{var}(X)+4\,\text{cov}(X,Y)+0+0=3+4\times 2=11\end{aligned}$$

14

§3.2 離散隨機變數之機率分配

本章介紹離散隨機變數的機率分配，下一章則為連續隨機變數。我們會介紹以下 7 個離散機率分配：

(1)白努力分配，

(2)二項分配，

(3)超幾何分配，

(4)卜瓦松分配，

(5)多項分配，

(6)負二項分配，

(7)幾何分配。

討論機率分配有以下幾個應注意的重點。

(1)有哪些參數，參數的意義為何？

(2)隨機變數的意義為何？

(3)分配的機率函數（機率密度函數）為何？

(4)隨機變數的期望值、變異數為何？

§3.2.1 白努力分配

白努力分配(***Bernoulli distribution***)又稱為**點二項分配**(***point binomial distribution***)，是所有機率分配的基礎。

白努力分配描述的是每次實驗成功的機率為 p，每組樣本只作 $n=1$ 次實驗；隨機變數定義為樣本中有 x 次成功出象，其中 $x\in\{0,1\}$ 。白努力分配的機率函數如下：

$$P(x)=f(x;p)=p^{x}(1-p)^{1-x}\text{，}\qquad x\in\{0,1\}\text{，}0<p<1\text{。}$$

其期望值與變異數分別為

$$E(x)=p\text{，}\qquad \operatorname{var}(x)=p(1-p)\text{。}$$

§3.2.2 二項分配

二項分配（***binomial distribution***）是複數樣本的白努力分配。也就是，二項分配描述的是每次實驗成功的機率為 p，每組樣本作 n 次實驗；隨機變數定義為樣本中有 x 次成功出象，其中 $x \in \{0,1,\ldots,n\}$ 。二項分配的機率函數如下：

$$P(x)=f(x;n,p)=C_x^n p^x (1-p)^{n-x}\text{，}\qquad x \in \{0,1,\ldots,n\}\text{，}0<p<1\text{。}$$

其期望值與變異數分別為

$$E(x)=np\text{，}\qquad \operatorname{var}(x)=np(1-p)\text{。}$$

定理3.7　二項分配與白努力分配的關係

若 $X_1, X_2, \ldots, X_n$ 為相同（成功機率為 p）且互相獨立的白努力分配，令

$$Y = X_1 + X_2 + \cdots + X_n$$

則 Y 為二項分配。

範例3.4　（二項分配）

若有 10%的網友謊報其性別，則 10 個網友中，請計算

(1)恰好有 2 位謊報性別的機率，　(2)沒有人謊報的機率，　(3)2 位以上謊報的機率。

【解】

資料：$p=10\%=0.1$，$n=10$，二項分配。

機率函數：$P(x)=f(x;n,p)=C_x^n p^x(1-p)^{n-x}$。

(1)

$$P(x=2)=C_2^{10}\times0.1^2\times0.9^8=\frac{10\times9}{2!}\times0.01\times0.43047=0.1937$$

(2)

$$P(x=0)=C_0^{10}\times0.1^0\times0.9^{10}=1\times1\times0.3487=0.3487$$

(3)

$$P(x>2)=1-P(x=0)-P(x=1)-P(x=2)=1-0.3487-0.3874-0.1937=0.0702$$

§3.2.3 超幾何分配

超幾何分配（***hyper-geometric distribution***）如同二項分配，只是成功機率在抽取樣本的過程會變動。

超幾何分配描述在總共有 N 個球，其中有 $S\le N$ 個球標示為『成功』球的球池中作抽球實驗，也就是第一次實驗抽取出『成功』球的機率為 $p=S/N$，接下來的成功機率會因前面出象結果的不同而有些許變動。如同二項分配，超幾何分配的每組樣本作 n 次實驗；隨機變數定義為樣本中有 x 次成功出象，其中 $x\in\{0,1,\ldots,n\}$；其機率函數如下：

$$P(x)=f(x;n,N,S)=\frac{C_x^S C_{n-x}^{N-S}}{C_n^N}，\qquad x\in\{0,1,\ldots,n\}，0\le S\le N。$$

其期望值與變異數分別為

$$E(x)=n\frac{S}{N}，\quad \mathrm{var}(x)=n\frac{S}{N}\left(1-\frac{S}{N}\right)\frac{N-n}{N-1}。$$

若令 $p=S/N$，則以上期望值、變數數與二項分配幾乎完全相同，僅有的差異為乘在變異數後的 $(N-n)/(N-1)$，此稱為**有限母體修正項**（***finite population correction***，***FPC***）；留意這個修正項，後面章節會有角色。

範例3.5 （超幾何分配）

某一 12 人的旅行團中，有 8 個人睡覺會打鼾，若 3 人一個房間，則某一房間

(1)恰好有 2 人會打鼾的機率為何？ (2)沒有人打鼾的機率？

【解】

資料：$N=12$ ，$S=8$ ，$n=3$ ，超幾何分配。

機率函數：$P(x)=f(x;n,N,S)=\dfrac{C_x^S C_{n-x}^{N-S}}{C_n^N}$ 。

(1)

$$P(x=2)=\frac{C_2^8C_1^4}{C_3^{12}}=\frac{\frac{8\times7}{2!}\times\frac{4}{1}}{\frac{12\times11\times10}{3!}}=\frac{8\times7\times2}{2\times11\times10}=\frac{28}{55}$$

(2)

$$P(x=0)=\frac{C_0^8C_3^4}{C_3^{12}}=\frac{1\times4}{\frac{12\times11\times10}{3!}}=\frac{4}{2\times11\times10}=\frac{1}{55}$$

超幾何分配與二項分配的關係

在超幾何分配之實驗球池中，即使依序抽取 n 個實驗球，我們仍然要求『**抽取後不放回**』；若實驗規則是『**抽取後放回**』，則每次實驗抽出成功球的機率會維持 $p=S/N$ ，這時它的機率函數應該是二項分配。

當 $N\to\infty$ 的時候，修正項 $(N-n)/(N-1)$ 會趨近於一；也就是說，當 $N\to\infty$ ，超幾何分配會趨近於二項分配。以下是機率函數的推導，供參考。

$$\begin{aligned}\lim_{N\to\infty}\frac{C_x^S C_{n-x}^{N-S}}{C_n^N}&=\lim_{N\to\infty}\frac{S!}{x!(S-x)!}\times\frac{(N-S)!}{(n-x)!(N-S-n+x)!}\times\frac{n!(N-n)!}{N!}\\&=\lim_{N\to\infty}\frac{n!}{x!(n-x)!}\times\frac{\overbrace{S(S-1)\cdots}^{x\text{ 項}}\times\overbrace{(N-S)(N-S-1)\cdots}^{n-x\text{ 項}}}{\underbrace{N(N-1)\cdots}_{n\text{ 項}}}\\&=\frac{n!}{x!(n-x)!}\times\frac{S^x(N-S)^{n-x}}{N^n}\\&=C_x^n\left(\frac{S}{N}\right)^x\left(1-\frac{S}{N}\right)^{n-x}\end{aligned}$$

若令 $p=S/N$ ，則會等於二項分配的機率函數 $C_x^n p^x(1-p)^{n-x}$ 。

§3.2.4 卜瓦松分配

卜瓦松分配（***Poisson distribution***）為每組樣本中實驗次數非常多（ $n\to\infty$ ）的二項分配；相對於實驗次數非常多，卜瓦松分配的另一特色，為每次實驗成功的機率會非常小，或說趨近於零， $p\to 0$ ；最後，卜瓦松分配的第三個特色為每組樣本成功次數的期望值倒是一個固定的數值， $np=\lambda$ 。

卜瓦松分配只有一個參數，即每組樣本的期望成功次數 λ。隨機變數定義為樣本中有 x 次成功出象，其機率函數如下：

$$P(x)=f(x;\lambda)=\frac{\lambda^x e^{-\lambda}}{x!}，\qquad x\in\{0,1,\ldots\}，\lambda>0。$$

其期望值與變異數分別為

$$E(x)=\lambda，\quad \operatorname{var}(x)=\lambda。$$

請留意，期望值與變異數相等， $E(x)=\operatorname{var}(x)=\lambda$ ，是卜瓦松分配非常特殊的特徵。

卜瓦松分配與二項分配的關係

考慮二項分配下列的極限狀況：

$$n\to\infty，p\to 0，且\ np=\lambda\ （\lambda 為常數）；$$

在極限狀況下，二項分配會趨近於卜瓦松分配。以下是機率函數的推導，供參考。

$$\begin{aligned}\lim_{n\to\infty} C_x^n p^x(1-p)^{n-x} &= \lim_{n\to\infty}\frac{n!}{x!(n-x)!}\left(\frac{\lambda}{n}\right)^x\left(1-\frac{\lambda}{n}\right)^{n-x}\\ &= \lim_{n\to\infty}\frac{\lambda^x}{x!}\frac{\overbrace{n\times(n-1)\times\cdots\times(n-x+1)}^{x\text{ 項}}}{n^x}\left(1-\frac{\lambda}{n}\right)^n\left(1-\frac{\lambda}{n}\right)^{-x}\\ &= \frac{\lambda^x}{x!}e^{-\lambda}\end{aligned}$$

其中，

$$\lim_{n\to\infty}\frac{\overbrace{n\times(n-1)\times\cdots\times(n-x+1)}^{x\text{ 項}}}{n^x}=1，\ \lim_{n\to\infty}\left(1-\frac{\lambda}{n}\right)^n=e^{-\lambda}，\ \lim_{n\to\infty}\left(1-\frac{\lambda}{n}\right)^{-x}=1。$$

範例3.6 （卜瓦松分配）

若 0.15%的車子有保颱風洪水險，則 100 輛泡水車中，

(1)恰好有 2 輛有保洪水險的機率為何？ (2)沒有車輛保洪水險的機率？

【解】

資料：$p = 0.15\% = 0.0015$ ，$n = 100$ ，二項分配。

機率函數：$P(x) = f(x;n,p) = C_x^n p^x (1-p)^{n-x}$ 。

(1) $P(x=2) = C_2^{100} \times 0.0015^2 \times 0.9985^{98} = \cdots$

(2) $P(x=0) = C_0^{100} \times 0.0015^0 \times 0.9985^{100} = \cdots$

或者

資料：$\lambda = np = 100 \times 0.15\% = 0.15$ ，卜瓦松分配。

機率函數：$P(x) = f(x;\lambda) = \dfrac{\lambda^x e^{-\lambda}}{x!}$ 。

(1) $P(x=2) = \dfrac{0.15^2}{2!} e^{-0.15} = 0.01125 \times 0.8607 = 0.00968$

(2) $P(x=0) = \dfrac{0.15^0}{0!} e^{-0.15} = e^{-0.15} = 0.8607$

範例3.7 （卜瓦松分配）

某打字員平均每兩頁出現一次錯誤，則該員在 5 頁的文稿中，

(1)恰好有 2 個錯誤機率為何？ (2)都沒有錯誤的機率？

【解】

資料：$\lambda = \frac{1}{2} \times 5 = 2.5$ ，卜瓦松分配。

機率函數：$P(x) = f(x;\lambda) = \dfrac{\lambda^x e^{-\lambda}}{x!}$ 。

(1)

$$P(x=2) = \frac{2.5^2}{2!} e^{-2.5} = 3.125 \times 0.0821 = 0.2565$$

(2)

$$P(x=0) = \frac{2.5^0}{0!} e^{-2.5} = e^{-2.5} = 0.0821$$

§3.2.5 多項分配

多項分配（***multinomial distribution***）是二項分配的延伸。相較於二項分配，每次實驗只有成功、失敗兩種出象，多項分配會有多項出象；例如三項分配的實驗會有三種可能出象。

三項分配描述的是每次實驗有三種可能出象，其出現機率分別為 p_1 、 p_2 、 p_3 ， $p_1+p_2+p_3=1$ ，每組樣本作 n 次實驗；隨機變數定義為樣本中三種出象各有 x_1 、 x_2 、 x_3 次， $x_1+x_2+x_3=n$ ，其中 $x_1,x_2,x_3\in\{0,1,\ldots,n\}$ 。三項分配的機率函數如下：

$$P(x_1,x_2,x_3)=f(x_1,x_2,x_3;n,p_1,p_2,p_3)=\frac{n!}{x_1!x_2!x_3!}p_1^{x_1}p_2^{x_2}p_3^{x_3}$$

其期望值與變異數分別為

$$E(x_1)=np_1 \text{，} E(x_2)=np_2 \text{，} E(x_3)=np_3 \text{，}$$
$$\operatorname{var}(x_1)=np_1(1-p_1) \text{，} \operatorname{var}(x_2)=np_2(1-p_2) \text{，} \operatorname{var}(x_3)=np_3(1-p_3) \text{。}$$

§3.2.6 負二項分配

負二項分配（***negative binomial distribution***）與二項分配互為對偶關係。二項分配以每組樣本中之實驗次數為參數，成功次數為隨機變數；而負二項分配則以每組樣本中之實驗次數為隨機變數，成功次數為參數。

負二項分配描述的是每次實驗成功的機率為 p，目標為樣本中有 r 次成功出象；隨機變數定義為每組樣本需作 x 次實驗，其中 $x\in\{r,r+1,\ldots\}$ 。負二項分配的機率函數如下：

$$P(x)=f(x;r,p)=C_{r-1}^{x-1}p^r(1-p)^{x-r} \text{，} \qquad x\in\{r,r+1,\ldots\} \text{，} 0<p<1 \text{。}$$

其期望值與變異數分別為

$$E(x)=\frac{r}{p} \text{，} \qquad \operatorname{var}(x)=\frac{r(1-p)}{p^2} \text{。}$$

請留意，負二項分配的最後一次實驗一定是成功出象；順著這個訊息就可以看懂上面的機率函數，那是前 $n-1$ 次實驗成功 $r-1$ 次，加上最後一次為成功出象的結果。

§3.2.7 幾何分配

幾何分配（***geometric distribution***）為 $r=1$ 的負二項分配。

幾何分配描述的是每次實驗成功的機率為 p，目標為樣本中出現第一次成功出象；隨機變數定義為每組樣本需作 x 次實驗，其中 $x\in\{1,2,\ldots\}$。幾何分配的機率函數如下：

$$P(x)=f(x;p)=p(1-p)^{x-1}，\qquad x\in\{1,2,\ldots\}，0<p<1。$$

其期望值與變異數分別為

$$E(x)=\frac{1}{p}，\quad \operatorname{var}(x)=\frac{1-p}{p^2}。$$

請留意，幾何分配與白努力分配互為對偶關係。

範例3.8 （幾何分配與負二項分配）

假設在某廟裡擲出『笑杯』的機率為 0.4，則連續擲杯 5 次後，

(1)出現第一個笑杯的機率為何？ (2)出現第三個笑杯的機率？

【解】

資料：。

機率函數：。

(1)

資料：$p=0.4$，幾何分配。

機率函數：$P(x)=f(x;p)=p(1-p)^{x-1}$。

$P(x=5)=0.4\times0.6^4=0.0518$

(2)

資料：$p=0.4$，$r=3$，負二項分配。

機率函數：$P(x)=f(x;r,p)=C_{r-1}^{x-1}p^r(1-p)^{x-r}$。

$P(x=5)=C_2^4\times0.4^3\times0.6^2=6\times0.064\times0.36=0.1382$

§3.2.8 離散分配間之關係

圖 0.1 呈現七個離散機率分配之間的關係。除了每次實驗的成功機率 p 之外，第二個參數可能是橫軸的實驗次數、或者縱軸的成功次數；其中，前者以成功次數為隨機變數，如列於圖上方白努力、二項分配等五個機率分配，後者以實驗次數為隨機變數，如列於圖右方的負二項、幾何分配。

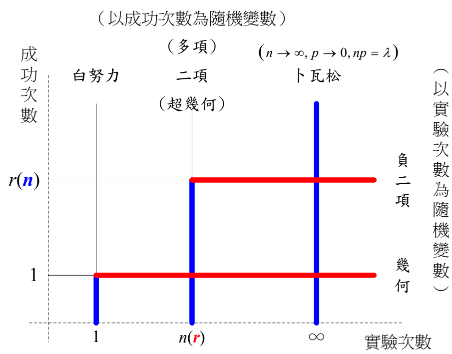


圖 0.1　離散分配間之關係

上方以成功次數為隨機變數的分配中，**二項分配**為其核心，有 4 種變化。(1)實驗次數只有 1 次者（$n \to 1$）成為**白努力分配**；(2)實驗次數很多者（$n \to \infty$，並搭配 $p \to 0$、$np = \lambda$）成為**卜瓦松分配**；(3)實驗成功機率不固定，而是隨抽出成功球或失敗球會有些微變化，則成為超幾何分配；(4)實驗出象不是只有如成功、失敗兩種，而是有多種可能出象，則成為**多項分配**。

右方以實驗次數為隨機變數的分配，要以對偶分配來理解。其中，負二項分配是二項分配的對偶分配，而幾何分配則為白努力分配的對偶分配。

圖 3.1 中所畫的粗線表示隨機變數的可能區間，縱線由左至右分別為白努力分配、二項分配、以及卜瓦松分配的成功次數的區間，橫線由下至上則分別為幾何分配、負二項分配的需要實驗次數的區間。

§3.3 連續隨機變數之機率分配

連續隨機變數的可能數值有無限多，造成單一數值出現的機率趨近於零，$P(x)=0, \forall x$，因此，我們不以機率函數，而是以**累積分配函數**（***cumulative distribution function***，***CDF***），$F(a)=P(x\le a)$，來表示連續隨機變數的機率。累積分配函數的微分版本為**機率密度函數**（***probability density function***，***PDF***），

$$f(x)=\frac{dF(x)}{dx} \quad 或 \quad F(a)=P(x\le a)=\int_{-\infty}^{a} f(x)\,dx，$$

機率密度函數也是連續機率分配的重要特徵。

本章會介紹 3 個連續機率分配，均等分配、指數分配、以及常態分配。

§3.3.1 均等分配

均等分配（***uniform distribution***）為隨機變數的每個數值出現的機率均相等的機率分配。若均等分配的隨機變數只適用於區間$[a,b]$，則其機率密度函數為

$$f(x)=f(x;a,b)=\frac{1}{b-a}， \qquad a\le x\le b；$$

期望值、變異數、與累積分配函數分別為

$$E(x)=\frac{a+b}{2}，\ \operatorname{var}(x)=\frac{(b-a)^2}{12}，\ F(x)=(x-a)/(b-a)。$$

範例3.9 （均等分配）

公車每 30 分鐘來一班，假設乘客不看時刻表，隨機來等車，請計算

(1)平均等候時間， (2)等候時間的標準差， (3)等候超過 10 分鐘的機率。

【解】

資料： $a=0$ ， $b=30$ ，均等分配。

累積分配函數： $F(x)=x/30$ ， $0\le x\le 30$ 。

(1)

$$E(x)=(a+b)/2=30/2=15$$

(2)

$$\sigma=\sqrt{\operatorname{var}(x)}=\sqrt{(b-a)^2/12}=\sqrt{30^2/12}=8.66$$

(3)

$$P(x>10)=1-P(x\le 10)=1-F(10)=1-10/30=2/3$$

§3.3.2 指數分配

指數分配（*exponential distribution*）需與卜瓦松分配互相參照。兩者擁有相同的參數，即每組樣本的平均成功次數 λ ；兩者隨機變數互為對偶，

卜瓦松分配以樣本的成功次數為隨機變數，

指數分配以兩次成功間的時間間隔為隨機變數。

以上如同二項分配（卜瓦松分配為二項分配的極限版本）與負二項分配的對偶關係，前者以成功次數為隨機變數，相對於後者以所需實驗次數為隨機變數；又若每次實驗的時間相同，則所需實驗次數等同於全部實驗所需的時間。事實上，指數分配對應的離散分配的版本為尋求第一次成功的幾何分配。

　　指數分配面臨單位時間的平均成功次數為λ，隨機變數為經過時間x後發生第一次成功出象，其機率密度函數為

$$f(x)=f(x;\lambda)=\lambda e^{-\lambda x}\text{，}\qquad x>0,\lambda>0\text{；}$$

期望值、變異數、與累積分配函數分別為

$$E(x)=\frac{1}{\lambda}\text{，}\operatorname{var}(x)=\left(\frac{1}{\lambda}\right)^2\text{，}F(x)=1-e^{-\lambda x}\text{。}$$

處理指數分配時要留意，指定參數λ之時間區間與隨機變數x的時間單位必須相同。

　　指數分配有另外一種表示法。上面的表示法中，參數λ與隨機變數x的內容或單位不同；新表示法以$\beta=1/\lambda$為參數，即平均兩次成功間的時間間隔為β。新表示法的機率密度函數為

$$f(x)=f(x;\beta)=\frac{1}{\beta}e^{-x/\beta}\text{，}\qquad x>0,\beta>0\text{；}$$

期望值、變異數、與累積分配函數分別為

$$E(x)=\beta\text{，}\operatorname{var}(x)=\beta^2\text{，}F(x)=1-e^{-x/\beta}\text{。}$$

比較多的學者使用λ為參數，可能這樣會與卜瓦松分配產生比較緊密的連結吧。

範例3.10　（指數分配）

某路段平均每年發生 10 件車禍，則

(1)平均多久發生一次車禍？　(2)某個月內沒有發生任何車禍的機率為何？

【解】

資料：$\lambda=10$，指數分配。

累計機率函數：$F(x)=1-e^{-\lambda x}$。

(1)

$$E(x)=\frac{1}{\lambda}=\frac{1}{10}\ (\text{年})$$

(2)

$$P(x>\text{一個月})=P\left(x>\tfrac{1}{12}\text{年}\right)=1-F\left(\tfrac{1}{12}\right)=e^{-10/12}=0.4346$$

§3.3.3 常態分配

常態分配（***normal distribution***）為抽樣分配的共同極限分配。

常態分配以平均數 μ 、標準差 σ 為參數，隨機變數為實數值 x，其機率密度函數為

$$f(x)=f(x;\mu,\sigma)=\frac{1}{\sigma\sqrt{2\pi}}e^{-\frac{1}{2}\left(\frac{x-\mu}{\sigma}\right)^2}\ ,\qquad x\in R,\ \sigma>0,\ \mu\in R\ ;$$

期望值與變異數分別為

$$E(x)=\mu\ ,\ \operatorname{var}(x)=\sigma^2\ 。$$

常態分配的累積分配函數無法表達成簡單的數學函數，需要由電腦執行積分運算，或者經由事先計算好的機率表來求取。

標準常態分配

標準常態分配（***normal distribution***）為常態隨機變數 x 取標準化分數後的分配。

標準常態分配的隨機變數為標準化分數 z， $z=(x-\mu)/\sigma$ ，其中 x 為期望值 μ 、標準差 σ 的常態分配，其機率密度函數為

$$f(z)=\frac{1}{\sqrt{2\pi}}e^{-\frac{1}{2}z^2}\ ,\qquad z\in R\ ;$$

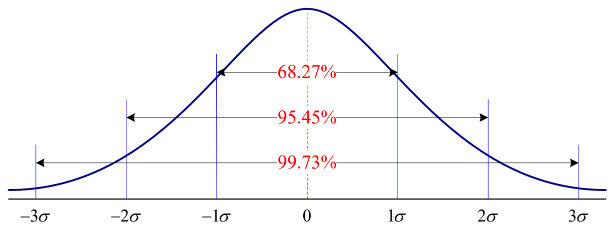
期望值與變異數分別為

$$E(z)=0\ ,\ \operatorname{var}(z)=1\ 。$$

標準常態分配的累積分配表需經由電腦積分計算，或者查機率表來求取。

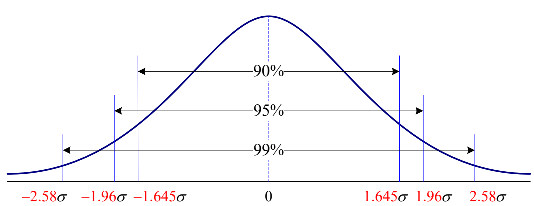
實務上，**我們只製作標準常態分配的機率表**；經由標準化分數的轉換，該機率表可以支援所有常態分配之機率的求取。

下面列出一些學習統計學的同學最好背誦起來，以節省查機率表時間的一些機率值以及其相關的分配圖。



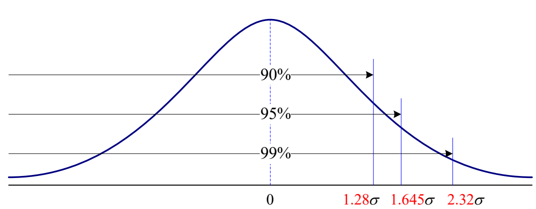
$P(-1 \le z \le 1) = 0.6827$ ， $P(-2 \le z \le 2) = 0.9545$ ， $P(-3 \le z \le 3) = 0.9973$ 。

圖 0.1　整數標準差區間的機率



$P(-1.645 \le z \le 1.645) = 0.9$ ， $P(-1.96 \le z \le 1.960) = 0.95$ ， $P(-2.58 \le z \le 2.576) = 0.99$

圖 0.1　常用區間機率(雙尾)的臨界值



$P(z \le 1.282) = 0.90$ ， $P(z \le 1.645) = 0.95$ ， $P(z \le 2.326) = 0.99$ 。

圖 0.2　常用累績機率(左尾)的臨界值

範例3.11 （常態分配）

已知 x 為 $\mu=12$ 、 $\sigma=4$ 的常態分配，則

(1)求 $P(8\le X\le 16)$ ；

(2)求 a 使 $P(X\le a)=0.9$ 。

【解】

(1)

$$z_1^*=\frac{8-\mu}{\sigma}=\frac{8-12}{4}=-1 \text{，} z_2^*=\frac{16-\mu}{\sigma}=\frac{16-12}{4}=1 \text{；}$$

$$P(8\le X\le 16)=P(z_1^*\le z\le z_2^*)=P(-1\le z\le 1)=0.6827$$

(2)

$$P(z\le 1.282)=0.9 \quad\Rightarrow\quad z^*=1.282$$

$$a=\mu+z^*\times\sigma=12+1.282\times 4=17.128$$

常態分配是二項分配的極限分配

若 n 越大，則二項分配（參數 n、p）與參數 $\mu=np$ 、 $\sigma=\sqrt{np(1-p)}$ 的常態分配越接近。當 n 的數值越大，計算二項分配的機率值變得越發困難（想像需要計算大數值的階乘 $n!$），此時我們會轉向以常態分配來求取近似值。

以常態分配（連續分配）來取得二項分配（離散分配）的近似值時，需特別注意 ±1/2 的修正項。假設 x 是離散隨機變數，y 是可以用來近似 x 的連續隨機變數，我們不能直接用 $P(y=a)$ 來近似 $P(x=a)$ ，因為只有單一數值之連續變數其機率值等於零，即 $P(y=a)=0$ ；應該以 $P(a-\frac{1}{2}\le y\le a+\frac{1}{2})$ 來近似 $P(x=a)$ 。近似的規則整理如下：

$$P_{\text{離散分配}}(x=a)=P_{\text{連續分配}}\left(a-\tfrac{1}{2}\le y\le a+\tfrac{1}{2}\right) \text{，}$$

$$P_{\text{離散分配}}(x\le a)=P_{\text{連續分配}}\left(y\le a+\tfrac{1}{2}\right) \text{，} \qquad P_{\text{離散分配}}(x\ge a)=P_{\text{連續分配}}\left(y\ge a-\tfrac{1}{2}\right) \text{，}$$

$$P_{\text{離散分配}}(x<a)=P_{\text{連續分配}}\left(y\le a-\tfrac{1}{2}\right) \text{，} \qquad P_{\text{離散分配}}(x>a)=P_{\text{連續分配}}\left(y\ge a+\tfrac{1}{2}\right) \text{。}$$

範例3.12　（以常態分配近似二項分配）

已知 x 為 $n=100$ 、 $p=0.2$ 的二項分配，y 為 $\mu=20$ 、 $\sigma=4$ 的常態分配，請計算 (1) $P(19\le X\le 21)$，(2) $P(19\le Y\le 21)$，(3) $P(18.5\le Y\le 21.5)$ 。

【解】

資料： $E(X)=np=20$ ， $var(X)=np(1-p)=16$ ， $\sigma_X=\sqrt{np(1-p)}=4$ 。

(1)　x 為 $n=100$ 、 $p=0.2$ 的二項分配， $P(x=a)=\dfrac{100!}{a!\times(100-a)!}\times 0.2^a\times 0.8^{100-a}$ 。

$$P(19\le x\le 21)=P(x=19)+P(x=20)+P(x=21)$$
$$=0.0981+0.0993+0.0946=0.2919$$

(2)　y 為 $\mu=20$ 、 $\sigma=4$ 的常態分配。

$$z_1^*=\frac{19-20}{4}=-0.25\text{ ， }z_2^*=\frac{21-20}{4}=0.25\text{ ；}$$

$$P(19\le Y\le 21)=P(-0.25\le z\le 0.25)=0.1974$$

(3)　y 為 $\mu=20$ 、 $\sigma=4$ 的常態分配。

$$z_1^*=\frac{18.5-20}{4}=-0.375\text{ ， }z_2^*=\frac{21.5-20}{4}=0.375\text{ ；}$$

$$P(18.5\le Y\le 21.5)=P(-0.375\le z\le 0.375)=0.2923$$

§3.4 機率表

為了求取某事件（隨機變數的某區間）的機率，除了直接從機率函數計算（離算分配），或從機率密度函數積分（連續分配），也可以直接查詢機率表中已經事先計算後的機率數值。

需要求取機率的區間有四類，簡稱為左尾、右尾、區間、以及雙尾，定義如下：

(1)**左尾**， $P(x\le a)=\alpha$ ，或稱以下累積機率；

(2)**右尾**， $P(x\ge b)=\alpha$ ，或稱以上累積機率；

(3)**區間**， $P(a\le x\le b)=\alpha$ ；

(4)**雙尾**， $P(x\le a$ 或 $x\ge b)=\alpha$ 。

在介紹連續機率分配時，因單一數值的機率為零，**我們提供累積機率** $F(a)=P(x\le a)$，這就是左尾。之所以這樣做，因為多數機率表提供的是左尾機率；雖然，初次使用機率表時還是要仔細閱讀說明，確定所提供的哪一種區間的機率。

假設機率表提供左尾機率 $F(a)$，以下說明連續分配如何以左尾機率來求其它區間的機率。

右尾：$P(x \geq b) = 1 - F(b)$；
區間：$P(a \leq x \leq b) = F(b) - F(a)$；
雙尾：$P(x \leq a \text{ 或 } x \geq b) = F(a) + 1 - F(b)$。

對於離散分配，我們需要留意臨界值（a 或 b）屬於哪一邊。例如，

$$P(x \geq 9) = 1 - F(8)，\quad P(x > 9) = 1 - F(9) \quad 或 \quad P(4 \leq x \leq 9) = F(9) - F(3)。$$

這些東西很容易犯錯，需隨時提醒自己留意。

§3.4.1 二項分配與卜瓦松分配機率表

表 0.1 是二項分配與卜瓦松分配的機率表。上方標題列為參數，例如左方的二項分配提供了 $(n=10, p=0.05)$、$(n=10, p=0.1)$ 等 5 個機率表，右方卜瓦松分配則提供了 $\lambda=0.15$、$\lambda=0.2$ 等 5 個機率表。左側標題行為隨機變數的臨界值(當然是左尾臨界值)。表身的數值為累積機率（左尾機率）。

表 0.1　二項分配(左)與卜瓦松分配機率表

*二項*分配***左尾***機率表

x \ p	n = 10				
	0.05	0.1	0.15	0.2	0.25
0	0.5987	0.3487	0.1969	0.1074	0.0563
1	0.9139	0.7361	0.5443	0.3758	0.2440
2	0.9885	0.9298	0.8202	0.6778	0.5256
3	0.9990	0.9872	0.9500	0.8791	0.7759
4	0.9999	0.9984	0.9901	0.9672	0.9219
5	1.0000	0.9999	0.9986	0.9936	0.9803
6	1.0000	1.0000	0.9999	0.9991	0.9965
7	1.0000	1.0000	1.0000	0.9999	0.9996
8	1.0000	1.0000	1.0000	1.0000	1.0000
9	1.0000	1.0000	1.0000	1.0000	1.0000
10	1.0000	1.0000	1.0000	1.0000	1.0000
11					
12					
13					
14					

*卜瓦松項*分配***左尾***機率表

	λ				
x	0.15	0.2	2	2.5	3
0	0.8607	0.8187	0.1353	0.0821	0.0498
1	0.9898	0.9825	0.4060	0.2873	0.1991
2	0.9995	0.9989	0.6767	0.5438	0.4232
3	1.0000	0.9999	0.8571	0.7576	0.6472
4	1.0000	1.0000	0.9473	0.8912	0.8153
5	1.0000	1.0000	0.9834	0.9580	0.9161
6	1.0000	1.0000	0.9955	0.9858	0.9665
7	1.0000	1.0000	0.9989	0.9958	0.9881
8	1.0000	1.0000	0.9998	0.9989	0.9962
9	1.0000	1.0000	1.0000	0.9997	0.9989
10	1.0000	1.0000	1.0000	0.9999	0.9997
11	1.0000	1.0000	1.0000	1.0000	0.9999
12	1.0000	1.0000	1.0000	1.0000	1.0000
13	1.0000	1.0000	1.0000	1.0000	1.0000
14	1.0000	1.0000	1.0000	1.0000	1.0000

範例3.13 （查表求二項分配的機率）

假設 x 為 $n=10$ 、 $p=0.1$ 的二項分配，請計算：

(1) $P(x \le 2)$ ， (2) $P(x \le 1)$ ， (3) $P(x=2)$ ， (4) $P(x>2)$ 。

【解】

(1) $P(x \le 2) = F(2) = bin(2,\ n=10, p=0.1) = 0.9298$

(2) $P(x \le 1) = F(1) = bin(1,\ n=10, p=0.1) = 0.7361$

(3) $P(x=2) = F(2) - F(1) = bin(2) - bin(1) = 0.9298 - 0.7361 = 0.1937$

(4) $P(x>2) = 1 - P(x \le 2) = 1 - F(2) = 1 - bin(2) = 1 - 0.9298 = 0.0702$

範例3.14 （查表求卜瓦松分配的機率）

假設 x 為 $\lambda=2$ 的卜瓦松分配，請計算：

(1) $P(x \le 2)$ ， (2) $P(x \le 1)$ ， (3) $P(x=2)$ ， (4) $P(x>2)$ 。

【解】

(1) $P(x \le 2) = F(2) = poi(2, \lambda=2) = 0.6767$

(2) $P(x \le 1) = F(1) = poi(1, \lambda=2) = 0.4060$

(3) $P(x=2) = F(2) - F(1) = poi(2, \lambda=2) - poi(1, \lambda=2) = 0.6767 - 0.4060 = 0.2707$

(4) $P(x>2) = 1 - P(x \le 2) = 1 - F(2) = 1 - poi(2, \lambda=2) = 1 - 0.6767 = 0.3233$

§3.4.2 標準常態分配機率表

表 0.1 是標準常態的機率表，或稱為 z 表。該表左方標題行與上方標題列合起來表示左尾臨界值，例如，第 3 列的 0.2 與第 4 行的 0.03 組合成 $z^* = 0.23$。表身的機率值不完全是累積機率（左尾），需要加上 0.5 後才是；此例查表得 $z(0.23) = 0.0910$（第 3 列第 4 行），則左尾機率為

$$F(z^* = 0.23) = P(z \le 0.23) = 0.5 + 0.0910 = 0.5 + z(0.23)$$ 。

表 0.1 標準常態分配機率表

z 分配**左尾**機率表

z	0.00	0.01	0.02	0.03	0.04	0.05	0.06	0.07	0.08	0.09
0.0	0.0000	0.0040	0.0080	0.0120	0.0160	0.0199	0.0239	0.0279	0.0319	0.0359
0.1	0.0398	0.0438	0.0478	0.0517	0.0557	0.0596	0.0636	0.0675	0.0714	0.0753
0.2	0.0793	0.0832	0.0871	0.0910	0.0948	0.0987	0.1026	0.1064	0.1103	0.1141
0.3	0.1179	0.1217	0.1255	0.1293	0.1331	0.1368	0.1406	0.1443	0.1480	0.1517
0.4	0.1554	0.1591	0.1628	0.1664	0.1700	0.1736	0.1772	0.1808	0.1844	0.1879
0.5	0.1915	0.1950	0.1985	0.2019	0.2054	0.2088	0.2123	0.2157	0.2190	0.2224
0.6	0.2257	0.2291	0.2324	0.2357	0.2389	0.2422	0.2454	0.2486	0.2517	0.2549
0.7	0.2580	0.2611	0.2642	0.2673	0.2704	0.2734	0.2764	0.2794	0.2823	0.2852
0.8	0.2881	0.2910	0.2939	0.2967	0.2995	0.3023	0.3051	0.3078	0.3106	0.3133
0.9	0.3159	0.3186	0.3212	0.3238	0.3264	0.3289	0.3315	0.3340	0.3365	0.3389

因為 z 分配為以 $z = 0$ 為中心的對稱分配，如此一來機率表只需列出正 z 值之機率。負 z 值之機率如 $F(z^* = -0.23) = 1 - F(z^* = 0.23) = 0.5 - 0.0910 = 0.5 - z(0.23)$。再來看一個例子，查表得 $z(0.59) = 0.2224$，其機率為

$$F(z^* = 0.59) = 0.5 + z(0.59) = 0.5 + 0.2224$$ ；

$$F(z^* = -0.59) = 0.5 - z(0.59) = 0.5 - 0.2224$$ 。

其查表規則如下

若 $a \ge 0$，則 $F(a) = 0.5 + z(a)$；若 $a < 0$，則 $F(a) = 0.5 - z(|a|)$。

相同的道理，已知機率後逆向找臨界值的規則如下

若 $F(a) = p \ge 0.5$，則 $a = z^{-1}(p - 0.5)$；若 $F(a) = p < 0.5$，則 $a = -z^{-1}(0.5 - p)$。

例如，求 $F(b) = 0.8 > 0.5$，則 $b = z^{-1}(0.8 - 0.5) = z^{-1}(0.3) = 0.84$；求 $F(b) = 0.2 < 0.5$，則 $b = -z^{-1}(0.5 - 0.3) = -z^{-1}(0.3) = -0.84$。

範例3.15 （查表求標準常態分配的機率）

假設 z 為標準常態分配，請計算：

(a) $P(0 \le z \le 1)$， (b) $P(z \le 1)$， (c) $P(-1 \le z \le 0)$， (d) $P(-1 \le z \le 1)$；

(e) $P(z \le -0.82)$， (f) $P(z \ge 0.26)$。

【解】

若 $a \ge 0$，則 $F(a) = 0.5 + z(a)$；若 $a < 0$，則 $F(a) = 0.5 - z(|a|)$。

(a)

查表 $z(1) = 0.3413$，$z(0) = 0$；

$P(0 \le z \le 1) = F(1) - F(0) = (0.5 + 0.3413) - (0.5 + 0) = 0.3413$。

(b)

查表 $z(1) = 0.3413$；$P(z \le 1) = F(1) = 0.5 + 0.3413 = 0.8413$。

範例3.15 （查表求標準常態分配的機率）(續)

(c)

查表 $z(0) = 0$，$z(1) = 0.3413$；

$P(-1 \le z \le 0) = F(0) - F(-1) = (0.5 + 0) - (0.5 - 0.3413) = 0.3413$。

(d)

查表 $z(1) = 0.3413$；

$P(-1 \le z \le 1) = F(1) - F(-1) = (0.5 + 0.3413) - (0.5 - 0.3413) = 0.6826$。

(e)

查表 $z(0.82) = 0.2939$；

$P(z \le -0.82) = F(-0.82) = 0.5 - 2939 = 0.2061$。

(f)

查表 $z(0.26) = 0.1026$；

$P(z \ge 0.26) = 1 - F(0.26) = 1 - (0.5 + 0.1026) = 0.3974$。

範例3.16 （查表求標準常態分配的臨界值）

假設 z 為標準常態分配，請求臨界值 b：

(a) $P(z \le b) = 0.22$， (b) $P(z \le b) = 0.84$， (c) $P(z \ge b) = 0.18$，

(d) $P(z \ge b) = 0.64$。

【解】

若 $F(a) = p \ge 0.5$，則 $a = z^{-1}(p - 0.5)$；若 $F(a) = p < 0.5$，則 $a = -z^{-1}(0.5 - p)$。

(a) $P(z \le b) = F(b) = 0.22 < 0.5$；

查表 $z^{-1}(0.5 - 0.22) = z^{-1}(0.28) = 0.77$；$b = -z^{-1}(0.28) = -0.77$。

(b) $P(z \le b) = F(b) = 0.84 > 0.5$；

查表 $z^{-1}(0.84 - 0.5) = z^{-1}(0.34) = 0.995$；$b = z^{-1}(0.34) = 0.995$。

(c) $P(z \ge b) = 1 - F(b) = 0.18 \Rightarrow F(b) = 0.82 > 0.5$；

查表 $z^{-1}(0.82 - 0.5) = z^{-1}(0.32) = 0.915$；$b = z^{-1}(0.32) = 0.915$。

(d) $P(z \ge b) = 1 - F(b) = 0.64 \Rightarrow F(b) = 0.36 < 0.5$；

查表 $z^{-1}(0.5 - 0.36) = z^{-1}(0.14) = 0.36$；$b = -z^{-1}(0.14) = -0.36$。

§3.5 計算隨機變數的期望值與變異數

連續隨機分配需要以積分來計算期望值、變異數等統計量值。本章先介紹一些基本的微積分公式，然後以範例的方式介紹如何計算各種分配的統計量值。

若函數 $f(x)$ 之(不定)積分的結果為 $F(x)$，兩者有以下關係：

$$\int f(x)dx = F(x) \quad \Leftrightarrow \quad f(x) = \frac{d}{dx}F(x) \quad \Leftrightarrow \quad f(x)dx = d\,F(x)$$

基本函數的微分、積分關係如**表 0.1** 所示。以下列出多項函數與指數函數的微分與積分關係。

多項函數：$\frac{d}{dx}x^a = ax^{a-1}$， $\int x^a dx = \frac{1}{a+1}x^{a+1}$， $a \ne 1$；

指數函數：$\frac{d}{dx}e^{bx} = be^{bx}$， $\int e^{bx} dx = \frac{1}{b}e^{bx}$， $b \ne 0$。

表 0.1 基本函數的微分、積分關係

$f(x)$	x^a	e^{bx}	c/x	$\sin\alpha x$	$\cos\beta x$
$F(x)$	$\frac{1}{a+1}x^{a+1}$	$\frac{1}{b}e^{bx}$	$c\ln x$	$-\frac{1}{\alpha}\cos\alpha x$	$\frac{1}{\beta}\sin\beta x$

另外，定積分與函數的和、積的積分關係如下。

定積分：　$\int_{\alpha}^{\beta} x^{a} dx = \frac{1}{a+1} x^{a+1} \Big]_{\alpha}^{\beta} = \frac{1}{a+1}\alpha^{a+1} - \frac{1}{a+1}\beta^{a+1}$，　$a \neq -1$，$\alpha < \beta$；

函數和的積分：　$\int [f(x)+g(x)]dx = \int f(x)dx + \int g(x)dx$；

函數純量積：　$\int af(x)dx = a\int f(x)dx$；

部分積分：　$\int F(x)g(x)dx = F(x)G(x) - \int f(x)G(x)dx$。

範例3.17　（基本積分計算）

請計算下列積分：

(a) $\int_{x=0}^{x=3} \frac{1}{3}x^2 dx$，　(b) $\int_0^{\infty} \lambda e^{-\lambda x} dx$，　(c) $\int_{y=0}^{y=3} \left(\frac{2}{3} - ay\right) dy$，　(d) $\int_0^{\infty} x\lambda e^{-\lambda x} dx$。

【解】

(a) $\int_{x=0}^{x=3} \frac{1}{3}x^2 dx = \frac{1}{3} \times \frac{1}{3}x^3 \Big|_{x=0}^{x=3} = \frac{1}{9} \times 27 = 3$

(b)

$$\int_0^{\infty} \lambda e^{-\lambda x} dx = \lambda \frac{1}{-\lambda} e^{-\lambda x}\Big|_0^{\infty} = 1$$

(c)

$$\int_{y=0}^{y=3} \left(\frac{2}{3} - ay\right) dy = \int_{y=0}^{y=3} \frac{2}{3} dy - \int_{y=0}^{y=3} aydy = \left(\frac{2}{3}y - a \times \frac{1}{2}y^2\right)\Big|_{y=0}^{y=3} = 2 - \frac{9}{2}a$$

(d)

令 $f = x, dg = \lambda e^{-\lambda x} dx \quad \Rightarrow \quad df = dx, g = -e^{-\lambda x}$

$$\int_0^{\infty} x\lambda e^{-\lambda x} dx = \int_0^{\infty} fdg = fg\Big|_0^{\infty} - \int_0^{\infty} gdf = -xe^{-\lambda x}\Big|_0^{\infty} + \int_0^{\infty} e^{-\lambda x} dx = -\frac{1}{\lambda} e^{-\lambda x}\Big|_0^{\infty} = \frac{1}{\lambda}$$

範例3.18 （以積分計算隨機變數之期望值、變異數）

隨機變數 X 的機率密度函數如下：

$$f(x)=\frac{1}{3},\ \ 0\le x\le 3$$

計算 $E(X)$ 、 $\operatorname{var}(X)$ 。

【解】

計算期望值：

$$E(X)=\int_{x=0}^{x=3} xf(x)dx=\int_{x=0}^{x=3} x\frac{1}{3}dx=\left.\frac{x^2}{6}\right|_{x=0}^{x=3}=\frac{9}{6}=\frac{3}{2}$$

計算變異數：

$$\operatorname{var}(X)=\int_{x=0}^{x=3}\left(x-E(x)\right)^2 f(x)dx=\int_{x=0}^{x=3}\left(x-\frac{3}{2}\right)^2\frac{1}{3}dx=\frac{3}{4}$$

或者

$$\operatorname{var}(X)=\int_{x=0}^{x=3} x^2 f(x)dx-\left(E(x)\right)^2=\int_{x=0}^{x=3}\frac{1}{3}x^2dx-\left(\frac{3}{2}\right)^2=3-\frac{9}{4}=\frac{3}{4}$$

範例3.19 （以積分計算隨機變數之期望值、變異數）

隨機變數 Y 的機率密度函數如下：

$$f(y)=\tfrac{2}{3}-ay,\ \ 0\le y\le 3$$

計算 $E(Y)$ 、 $\operatorname{var}(Y)$ 。

【解】

求未知參數 a：

$$\int_{y=0}^{y=3} f(y)dy=1\ \Rightarrow\ \int_{y=0}^{y=3}\left(\tfrac{2}{3}-ay\right)dy=\left.\left(\tfrac{2}{3}y-a\frac{y^2}{2}\right)\right|_{y=0}^{y=3}=2-\tfrac{9}{2}a=1\ \ \Rightarrow\ \ a=\frac{2}{9}$$

計算期望值：

$$E(Y)=\int_{y=0}^{y=3} yf(y)dy=\int_{y=0}^{y=3} y\left(\tfrac{2}{3}-\tfrac{2}{9}y\right)dy=\left.\left(\tfrac{1}{3}y^2-\tfrac{2}{27}y^3\right)\right|_{y=0}^{y=3}=1$$

計算變異數：

$$\operatorname{var}(Y)=\int_{y=0}^{y=3}\left(y-E(y)\right)^2 f(y)dy=\int_{y=0}^{y=3}\left(y-1\right)^2\left(\tfrac{2}{3}-\tfrac{2}{9}y\right)dy=-\frac{9}{2}+10-7+2=\frac{1}{2}$$

或者

$$\operatorname{var}(Y)=\int_{y=0}^{y=3}\left(y-E(y)\right)^2 f(y)dy=\int_{y=0}^{y=3}\left(y-1\right)^2\left(\tfrac{2}{3}-\tfrac{2}{9}y\right)dy=-\frac{9}{2}+10-7+2=\frac{1}{2}$$

範例3.20 （計算條件機率、邊際機率）

隨機變數 X 、 Y 的聯合機率密度函數如下：

$$
\begin{aligned}
f(x,y) &= x+ay, \quad 0<x<1, 0<y<1 \\
&= 0 \qquad \text{otherwise}
\end{aligned}
$$

計算 $f(x)$ 、 $f(y)$ 、 $f(x|y)$ 、 $f(y|x)$ 、 $\text{var}(X)$ 、 $\text{cov}(X,Y)$ 。

【解】

求未知參數 a：

$$\int_{y=0}^{1}\int_{x=0}^{1}(x+ay)dxdy=\int_{y=0}^{1}\left(\tfrac{1}{2}+ay\right)dy=\tfrac{1}{2}+\tfrac{a}{2}=1 \quad \Rightarrow \quad a=1$$

計算邊際機率密度函數：

$$f(x)=\int_{y=0}^{1}(x+y)dy=x+\tfrac{1}{2}$$

$$f(y)=\int_{x=0}^{1}(x+y)dx=\tfrac{1}{2}+y$$

因 $f(x,y)=x+y\neq f(x)f(y)=\left(x+\tfrac{1}{2}\right)\left(\tfrac{1}{2}+y\right)$，故 X 、 Y 兩者不獨立。

範例3.20 （計算條件機率、邊際機率）(續)

計算條件機率密度函數：

$$f(x|y)=\frac{f(x,y)}{f(y)}=\frac{x+y}{y+1/2}, \qquad f(y|x)=\frac{f(x,y)}{f(x)}=\frac{x+y}{x+1/2} 。$$

計算期望值：

$$E(x)=E(y)=\int_0^1 y\left(\tfrac{1}{2}+y\right)dy=\frac{1}{2}\times\frac{1}{2}+\frac{1}{3}=\frac{7}{12}$$

$$E(x^2)=E(y^2)=\int_0^1 y^2\left(\tfrac{1}{2}+y\right)dy=\frac{1}{2}\times\frac{1}{4}+\frac{1}{4}=\frac{5}{12}$$

$$E(xy)=\int_0^1\int_0^1 xy(x+y)dxdy=\int_0^1\left(\tfrac{y}{3}+\tfrac{y^2}{2}\right)dy=\frac{1}{3}\times\frac{1}{2}+\frac{1}{2}\times\frac{1}{3}=\frac{1}{3}$$

計算變異數、共變數：

$$\text{var}(X)=E(x^2)-\left[E(x)\right]^2=\frac{5}{12}-\frac{7}{12}\times\frac{7}{12}=\frac{11}{144}$$

$$\text{cov}(X,Y)=E(xy)-E(x)E(y)=\frac{1}{3}-\frac{7}{12}\times\frac{7}{12}=-\frac{1}{144}$$

§3.6 動差與動差母函數

討論隨機變數 X 的機率分配性質時，常會計算 $E(x)$、$E(x^2)$、$E(x^3)$ 等隨機變數某次方的期望值，這些數值稱為**動差**（***moments***）或**零動差**（***zero moments***），n 次方的動差稱為 ***n* 級動差**，標記如下：

n 級動差：μ_n'、M_n'、$E(x^n)$。

相對於零動差，減去平均值後再取次方者稱為**中央動差**（***central moments***）或**主動差**（***principle moments***），標記如下：

n 級主動差：μ_n、M_n、$E\left[(x-\mu)^n\right]$。

事實上，具有統計意義的是中央動差，而零動差有容易計算的特性。例如變異數公式：

$$\sigma^2 = E\left[(x-\mu)^2\right] = E(x^2) - E(x)^2，$$

這是將中央動差表達成零動差的公式；先計算零動差，然後帶入公式求算統計量值。以下列出更多中央動差表達成零動差的所有公式：

$$M_2 = E(x-\mu)^2 = E(x^2) - 2\mu E(x) + \mu^2;$$
$$M_3 = E(x-\mu)^3 = E(x^3) - 3\mu E(x^2) + 3\mu^2 E(x) - \mu^3;$$
$$M_n = E(x-\mu)^n = \sum_{i=0}^{n} C_i^n (-\mu)^i E(x^{n-i})。$$

已經介紹過與這些中央動差有關的統計量有，平均數 $\mu = E(x)$，變異數 $\sigma^2 = M_2$，偏態係數 $\alpha_3 = M_3/\sigma^3$，以及峰態係數 $\alpha_4 = M_4/\sigma^4$。

另外，函數e^{tx}是一個特殊的函數，其中，x為隨機變數，$t \in R$。此函數的期望值，

$$m(t)=E\left(e^{tx}\right),$$

稱為**動差母函數**（***moment generating function***，**動差生成函數**）。意思是，我們可藉由動差母函數來產生所有各級的動差。

範例3.21 （由動差母函數產生各級動差）

將e^{tx}作泰勒展開：

$$e^{tx}=1+tx+\frac{(tx)^2}{2!}+\frac{(tx)^3}{3!}+\frac{(tx)^4}{4!}+\cdots$$

等號兩邊各取期望值，則

$$m(t)=E\left(e^{tx}\right)=1+tE(x)+\frac{t^2}{2!}E\left(x^2\right)+\frac{t^3}{3!}E\left(x^3\right)+\frac{t^4}{4!}E\left(x^4\right)+\cdots$$

可以驗證以下操作成立：

$$\left.\frac{d}{dt}m(t)\right]_{t=0}=\left.E(x)+tE\left(x^2\right)+\frac{t^2}{2!}E\left(x^3\right)+\cdots\right]_{t=0}=E(x)$$

$$\left.\frac{d^2}{dt^2}m(t)\right]_{t=0}=\left.E\left(x^2\right)+tE\left(x^3\right)+\frac{t^2}{2!}E\left(x^4\right)+\cdots\right]_{t=0}=E\left(x^2\right)$$

$$\left.\frac{d^k}{dt^k}m(t)\right]_{t=0}=\left.E\left(x^k\right)+tE\left(x^{k+1}\right)+\frac{t^2}{2!}E\left(x^{k+2}\right)+\cdots\right]_{t=0}=E\left(x^k\right)$$

範例3.22 （由動差母函數求期望值、變異數）

若隨機變數 x 的動差母函數為 $m(t)=e^{\lambda(e^t-1)}$ ，請計算 $E(x)$ 、 $\mathrm{var}(x)$ 。

【解】

由 $m(t)$ 求一、二級動差，過程如下

$$E(x)=m^{(1)}(0)=e^{\lambda(e^t-1)}\lambda e^t\Big]_{t=0}=\lambda$$

$$E(x^2)=m^{(2)}(0)=\left\{e^{\lambda(e^t-1)}(\lambda e^t)^2+e^{\lambda(e^t-1)}\lambda e^t\right\}\Big]_{t=0}=\lambda^2+\lambda$$

則

$$\mathrm{var}(x)=E(x^2)-E(x)^2=\lambda^2+\lambda-\lambda^2=\lambda$$

範例3.23 （白努力分配之動差母函數）

白努力分配的機率函數為

$$P(x)=f(x;p)=p^x(1-p)^{1-x} ,\qquad x\in\{0,1\} ，0<p<1 。$$

驗證機率函數

$$\sum_{x=0}^{1}p^x(1-p)^{1-x} = (1-p)+p = 1$$

其動差母函數 $m(t)$ 的求解過程如下

$$m(t)=E(e^{tx})=\sum_{x=0}^{1}e^{tx}p^x(1-p)^{1-x}=(1-p)+e^tp=pe^t+(1-p)$$

期望值

$$E(x)=m^{(1)}(0)=pe^t\Big]_{t=0}=p$$

變異數

$$E(x^2)=m^{(2)}(0)=pe^t\Big]_{t=0}=p$$

$$\mathrm{var}(x)=E(x^2)-[E(x)]^2=p-p^2=p(1-p)$$

範例3.24 （二項分配之動差母函數）

二項分配的機率函數為

$$P(x)=f(x;n,p)=C_x^n p^x(1-p)^{n-x}\text{，}\quad x\in\{0,1,\ldots,n\}\text{，}0<p<1\text{。}$$

驗證機率函數

$$\sum_{x=0}^{n} C_x^n p^x(1-p)^{n-x}=\left[p+(1-p)\right]^n=1 \qquad \text{（二項式定理）}$$

其動差母函數 $m(t)$ 的求解過程如下

$$m(t)=E\left(e^{tx}\right)=\sum_{x=0}^{n} e^{tx}C_x^n p^x(1-p)^{n-x}=\sum_{x=0}^{n} C_x^n\left(pe^t\right)^x(1-p)^{n-x}=\left[pe^t+(1-p)\right]^n$$

期望值

$$E(x)=m^{(1)}(0)=n\left[pe^t+(1-p)\right]^{n-1}pe^t\Big]_{t=0}=np$$

變異數

$$E\left(x^2\right)=m^{(2)}(0)$$

$$=n(n-1)\left[pe^t+(1-p)\right]^{n-2}\left(pe^t\right)^2+n\left[pe^t+(1-p)\right]^{n-1}pe^t\Big]_{t=0}=n(n-1)p^2+$$

$$\operatorname{var}(x)=E\left(x^2\right)-E(x)^2=n(n-1)p^2+np-(np)^2=np(1-p)$$

範例3.25 （動差母函數之性質）

令 $X_1,X_2,\ldots,X_n$ 為一致且互相獨立的分配，其動差母函數皆為 $m(t)$，若

$$Y=X_1+X_2+\cdots+X_n$$

則 Y 之動差母函數為

$$E\left(e^{ty}\right)=E\left(e^{t(x_1+x_2+\cdots+x_n)}\right)=E\left(e^{tx_1}e^{tx_2}\cdots e^{tx_n}\right)=E\left(e^{tx_1}\right)E\left(e^{tx_2}\right)\cdots E\left(e^{tx_n}\right)=m(t)^n$$

白努力分配與二項分配有以上的關係，例如白努力分配的動差為

$$m(t)=pe^t+(1-p)$$

而二項分配的動差則為

$$m(t)^n=\left[pe^t+(1-p)\right]^n$$

範例3.26　（幾何分配之動差母函數）

幾何分配的機率函數為

$$P(x)=f(x;p)=p(1-p)^{x-1}\text{ ，}\quad x\in\{1,2,\ldots\}\text{ ，}0<p<1\text{ 。}$$

驗證機率函數

$$\sum_{x=1}^{\infty}p(1-p)^{x-1}=p\sum_{x=1}^{\infty}(1-p)^{x-1}=p\times\frac{1}{1-(1-p)}=1\qquad\left(\sum_{n=0}^{\infty}x^{n}=\frac{1}{1-x}\right)$$

其動差母函數 $m(t)$ 的求解過程如下

$$m(t)=E\left(e^{tx}\right)=\sum_{x=0}^{1}e^{tx}p(1-p)^{x-1}=pe^{t}\sum_{x=0}^{1}\left[e^{t}(1-p)\right]^{x-1}=\frac{pe^{t}}{1-e^{t}(1-p)}=\frac{pe^{t}}{1-(1-p)e^{t}}$$

期望值

$$E(x)=m^{(1)}(0)=\frac{pe^{t}}{1-(1-p)e^{t}}+\frac{pe^{t}\times(1-p)e^{t}}{\left[1-(1-p)e^{t}\right]^{2}}\Bigg]_{t=0}=\frac{pe^{t}}{\left[1-(1-p)e^{t}\right]^{2}}\Bigg]_{t=0}=\frac{1}{p}$$

範例3.26　（幾何分配之動差母函數）(續)

變異數

$$E\left(x^{2}\right)=m^{(2)}(0)$$

$$=\frac{pe^{t}}{\left[1-(1-p)e^{t}\right]^{2}}+\frac{pe^{t}\times2(1-p)e^{t}}{\left[1-(1-p)e^{t}\right]^{3}}\Bigg]_{t=0}=\frac{pe^{t}\times\left[1+(1-p)e^{t}\right]}{\left[1-(1-p)e^{t}\right]^{3}}\Bigg]_{t=0}=\frac{2-p}{p^{2}}$$

$$\operatorname{var}(x)=E\left(x^{2}\right)-E(x)^{2}=\frac{2-p}{p^{2}}-\left(\frac{1}{p}\right)^{2}=\frac{1-p}{p^{2}}=\frac{1}{p}\times\frac{1-p}{p}$$

範例3.27 （負二項分配之動差母函數）

負二項分配的機率函數為

$$P(x)=f(x;r,p)=C_{r-1}^{x-1}p^{r}(1-p)^{x-r}\text{ ，}\qquad x\in\{r,r+1,\ldots\}\text{ ，}0<p<1\text{ 。}$$

驗證機率函數

$$\sum_{x=r}^{\infty}C_{r-1}^{x-1}p^{r}(1-p)^{x-r}=p^{n}\sum_{x=r}^{\infty}\frac{\overbrace{(x-1)(x-2)\cdots}^{r-1\text{ 項}}}{(r-1)!}(1-p)^{x-r}=p^{r}\times\frac{1}{(r-1)!}\frac{(r-1)!}{\left[1-(1-p)\right]^{r}}=1$$

其動差母函數 $m(t)$ 的求解過程如下

$$m(t)=E\left(e^{tx}\right)=\left[\frac{pe^{t}}{1-(1-p)e^{t}}\right]^{r}$$

期望值

$$E(x)=m^{(1)}(0)=r\left[\frac{pe^{t}}{1-(1-p)e^{t}}\right]^{r-1}\frac{pe^{t}}{\left[1-(1-p)e^{t}\right]^{2}}\Bigg]_{t=0}=\frac{r}{p}$$

範例3.27 （負二項分配之動差母函數）(續)

變異數

$$E\left(x^{2}\right)=m^{(2)}(0)=r(r-1)\left[\frac{pe^{t}}{1-(1-p)e^{t}}\right]^{r-2}\left(\frac{pe^{t}}{\left[1-(1-p)e^{t}\right]^{2}}\right)^{2}\Bigg]_{t=0}$$

$$+r\left[\frac{pe^{t}}{1-(1-p)e^{t}}\right]^{r-1}\frac{pe^{t}\times\left[1+(1-p)e^{t}\right]}{\left[1-(1-p)e^{t}\right]^{3}}\Bigg]_{t=0}$$

$$=\frac{r(r-1)}{p^{2}}+\frac{r(2-p)}{p^{2}}=\frac{r(r+1-p)}{p^{2}}$$

$$\operatorname{var}(x)=E\left(x^{2}\right)-E(x)^{2}=\frac{r(r+1-p)}{p^{2}}-\left(\frac{r}{p}\right)^{2}=\frac{r(1-p)}{p^{2}}=r\times\frac{1}{p}\times\frac{1-p}{p}$$

範例3.28　（卜瓦松分配之動差母函數）

卜瓦松分配的機率函數為

$$P(x)=f(x;\lambda)=\frac{\lambda^x e^{-\lambda}}{x!}\text{，}\qquad x\in\{0,1,\ldots\}\text{，}\lambda>0\text{。}$$

驗證機率函數

$$\sum_{x=0}^{\infty}\frac{\lambda^x}{x!}e^{-\lambda} = e^{-\lambda}\sum_{x=0}^{\infty}\frac{\lambda^x}{x!} = e^{-\lambda}\times e^{\lambda} = 1 \qquad \left(e^{\lambda}=\sum_{x=0}^{\infty}\frac{\lambda^x}{x!}\right)$$

其動差母函數 $m(t)$ 的求解過程如下

$$m(t)=E\left(e^{tx}\right)=\sum_{x=0}^{\infty}e^{tx}\frac{\lambda^x e^{-\lambda}}{x!}=e^{-\lambda}\sum_{x=0}^{\infty}\frac{\left(\lambda e^t\right)^x}{x!}=e^{-\lambda}e^{\lambda e^t}=e^{\lambda\left(e^t-1\right)}$$

期望值

$$E(x)=m^{(1)}(0)=e^{\lambda\left(e^t-1\right)}\lambda e^t\Big]_{t=0}=\lambda$$

範例3.28　（卜瓦松分配之動差母函數）(續)

變異數

$$E\left(x^2\right)=m^{(2)}(0)=\left\{e^{\lambda\left(e^t-1\right)}\left(\lambda e^t\right)^2+e^{\lambda\left(e^t-1\right)}\lambda e^t\right\}\Big]_{t=0}=\lambda^2+\lambda$$

$$\operatorname{var}(x)=E\left(x^2\right)-E(x)^2=\lambda^2+\lambda-\lambda^2=\lambda$$

範例3.29 （指數分配之動差母函數）

指數分配的機率密度函數為

$$f(x)=f(x;\lambda)=\lambda e^{-\lambda x}\text{，}\quad x>0\text{，}\lambda>0\text{。}$$

驗證機率函數

$$\int_0^\infty \lambda e^{-\lambda x}dx=\lambda\frac{1}{-\lambda}e^{-\lambda x}\Big|_0^\infty=1$$

其動差母函數 $m(t)$ 的求解過程如下

$$E(x)=m^{(1)}(0)=\left.\frac{\lambda}{(\lambda-t)^2}\right]_{t=0}=\frac{1}{\lambda}$$

期望值

$$E(x)=m^{(1)}(0)=\left.e^{\lambda(e^t-1)}\lambda e^t\right]_{t=0}=\lambda$$

範例3.29 （指數分配之動差母函數）(續)

變異數

$$E(x^2)=m^{(2)}(0)=\left.\frac{2\lambda}{(\lambda-t)^3}\right]_{t=0}=\frac{2}{\lambda^2}$$

$$\mathrm{var}(x)=E(x^2)-E(x)^2=\frac{2}{\lambda^2}-\left(\frac{1}{\lambda}\right)^2=\frac{1}{\lambda^2}$$

The End

79

80