2.3 $\boxed{\text{習題}}$ 說明 $\int_0^\infty e^{-x^2}\,dx = \frac{\sqrt{\pi}}{2}$.

例. $\int_0^\infty x^{-\frac{1}{2}}e^{-x}\,dx = \sqrt{\pi}$.

變換變數 $x = u^2$, $dx = 2u\,du$, 則原式

$$\int_0^\infty x^{-\frac{1}{2}}e^{-x}\,dx = \lim_{b\to\infty}\int_0^b x^{-\frac{1}{2}}e^{-x}\,dx = \lim_{b\to\infty}\int_0^{\sqrt{b}}\left(u^{-1}e^{-u^2}\right)\cdot(2u)du$$
$$= 2\lim_{b\to\infty}\int_0^{\sqrt{b}} e^{-u^2}du = 2\int_0^\infty e^{-u^2}\,du = \sqrt{\pi}$$

□

例. (Γ 函數) 定義 $\Gamma(\alpha) = \displaystyle\int_0^\infty x^{\alpha-1}e^{-x}\,dx$.

由前面計算知:

$$\Gamma(1) = \int_0^\infty e^{-x}\,dx = \lim_{b\to\infty}(-e^{-x})|_0^b = 1$$
$$\Gamma(\frac{1}{2}) = \int_0^\infty x^{-\frac{1}{2}}e^{-x}\,dx = \sqrt{\pi}$$

又當 $\alpha > 1$ 時,

$$\begin{aligned}\Gamma(\alpha) &= \int_0^\infty x^{\alpha-1}e^{-x}\,dx = \lim_{b\to\infty}\int_0^b x^{\alpha-1}\,d(-e^{-x})\\ &= \lim_{b\to\infty}\left(x^{\alpha-1}(-e^{-x})|_0^b + (\alpha-1)\int_0^b x^{\alpha-2}e^{-x}\,dx\right)\\ &= 0 + (\alpha-1)\cdot\Gamma(\alpha-1)\end{aligned}$$

由此可知, 當 $n \in \mathbb{N}$ 時

$$\begin{aligned}\Gamma(n) &= (n-1)\cdot\Gamma(n-1) = (n-1)\cdot(n-2)\cdot\Gamma(n-2)\\ &= \cdots = (n-1)!\end{aligned}$$

□

2.4 $\boxed{\text{習題}}$ 設$n \in \mathbb{N}$, 說明

$$\Gamma(n+\frac{1}{2}) = \frac{1\cdot 3\cdot 5\cdots(2n-1)}{2^n}\sqrt{\pi}$$

- -

2.5 $\boxed{\text{習題}}$ 計算 (1) $\int_{-\infty}^{\infty} xe^{-x^2}\,dx$ (2) $\int_{-\infty}^{\infty} x^2e^{-x^2}\,dx.$

(Ans. (1) 0, (2) $\frac{\sqrt{\pi}}{2}$.)

2.6 $\boxed{\text{習題}}$

(1) 驗證

$$\int_{-\infty}^{\infty} e^{\lambda x}e^{-x^2}\,dx \;=\; \sqrt{\pi}\cdot e^{\frac{\lambda^2}{4}}, \quad \lambda \text{ 為常數}.$$

(2) 將上式想成 λ 的等式, 兩邊對 $\lambda=0$ 作λ 的泰勒展式, 比較 λ^k 的係數, 説明

$$\int_{-\infty}^{\infty} x^{2j+1}e^{-x^2}\,dx = 0, \quad \int_{-\infty}^{\infty} x^{2j}e^{-x^2}\,dx = \sqrt{\pi}\cdot\frac{(2j)!}{4^j\cdot j!}, \quad j=0,1,2,\cdots$$

2.7 $\boxed{\text{習題}}$ 計算 $\int_{-\infty}^{\infty} e^{-\lambda t^2}\,dx,\ \lambda>0.$ (Ans. $\sqrt{\frac{\pi}{\lambda}}$.)

2.8 $\boxed{\text{習題}}$ 計算 $\displaystyle\int_{-\infty}^{\infty} e^{-\frac{(x-\mu)^2}{2\sigma^2}}\,dx,\ \mu,\sigma$ 為常數且 $\sigma>0.$ (Hint: 變數變換 $u=\frac{x-\mu}{\sqrt{2}\sigma}$)

(Ans. $\sqrt{2\pi}\sigma$.)

2.9 $\boxed{\text{習題}}$ (續上題) 計算 (1) $\int_{-\infty}^{\infty} x\cdot e^{-\frac{(x-\mu)^2}{2\sigma^2}}\,dx$ (2) $\int_{-\infty}^{\infty} x^2\cdot e^{-\frac{(x-\mu)^2}{2\sigma^2}}\,dx$

(Ans. (1) $\sqrt{2\pi}\sigma\mu$, (2) $\sqrt{2\pi}\sigma(\mu^2+\sigma^2)$.)

2.10 $\boxed{\text{習題}}$ 定義

$$\tilde{\Gamma}(\alpha,\beta) = \int_0^{\infty} x^{\alpha-1}e^{-\beta x}\,dx, \text{ 其中 } \alpha>1,\ \beta>0.$$

證明 $\tilde{\Gamma}(\alpha,\beta)=\Gamma(\alpha)\cdot\beta^{-\alpha}$. (Hint: 變數變換)

3 連續型機率

第一節中的機率是所謂的離散型機率, $\mathrm{P}(X=k)$ 只在某些「離散」的 k 值上面才有定義. 例如在二項分配中, 如果考慮 $Z=\frac{1}{2}$ 的事件, 要嘛, 我們就説這是機率為 0 的空事件, 也許更多人會認為這真是荒唐的問題. 不過在本節中, 我們要開始考慮, 隨機變數的取值在整個實數區間上的情形. 前面所談到的身高、體重、壽命都是連續隨機變數的例子.

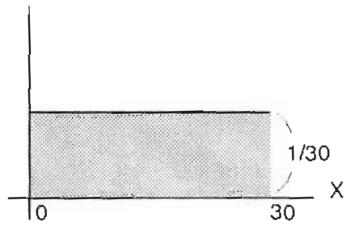
例. 在某公車牌候車30分鐘, 由於我們對公車時刻表毫無概念, 只知道每30分鐘一班車. 我們可以用下述(均勻分配)的函數來描述我們等車的過程:

令 T 表示何時等到車的隨機變數, T 遵循「機率函數」: $f(t) = \frac{1}{30}, 0 \le t \le 30$ 的意思是說:

(1) $\mathrm{P}(10 \le T \le 20) = \int_{10}^{20} \frac{1}{30}\, dt = \frac{1}{3}$. 表示有 $\frac{1}{3}$ 的機會, 車子會在第 10 分鐘到第20分鐘之間到站.

(2) $\mathrm{P}(0 \le T \le 30) = \int_{0}^{30} \frac{1}{30}\, dt = 1$. 表示在 30 分鐘內車子總會來的事實.

(3) $\mathrm{P}(T = t_0) = 0$. 車子剛好在 t_0 時到站的機率 $= 0$.



因此連續型與離散型機率的差異並不大, 只是用積分來取代取和, 以及用 $a \le T \le b$ 取代 $Z = k$ 來描述「有意義」的事件.

定義 3.1 (連續型隨機變數) 若 X 為取值在 $\mathbb{R}$ 的隨機變數. 如果可以找到函數 $\mathrm{f}_X(t) \ge 0$, $t \in \mathbb{R}$, 使得

$$\mathrm{P}(a \le X \le b) = \int_a^b \mathrm{f}_X(t)\, dt$$

而且

$$\mathrm{P}(X \in \mathbb{R}) = \mathrm{P}(-\infty < X < \infty) = \text{瑕積分} \int_{-\infty}^{\infty} \mathrm{f}_X(t)\, dt = 1$$

我們稱 X 為連續型隨機變數, f_X 稱為對應於 X 的機率密度函數.

例. 前節提供了許多可能的機率密度函數.

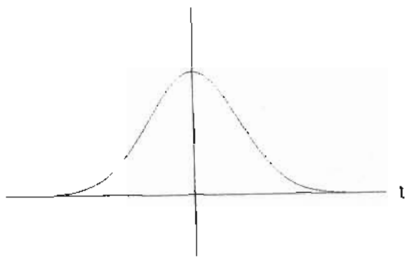
(1) $\dfrac{1}{\sqrt{\pi}} e^{-t^2}$ 常態分配.

(2) $\begin{cases} e^{-t} & t \ge 0 \\ 0 & t < 0 \end{cases}$ 指數分配.

(3) $\begin{cases} \dfrac{1}{\sqrt{2\pi}} t^{-\frac{1}{2}} e^{-\frac{t}{2}} & t > 0 \\ 0 & t \le 0 \end{cases}$ $\chi^2(1)$ 分配.

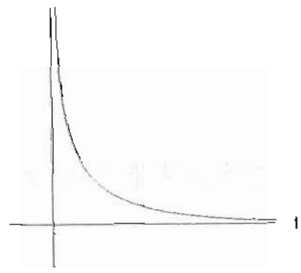
(4) $\begin{cases} \dfrac{\beta^\alpha}{\Gamma(\alpha)} t^{\alpha-1} e^{-\beta t} & t > 0 \\ 0 & t \le 0 \end{cases}$ $\Gamma(\alpha, \beta)$ 分配, $\alpha > 0$, $\beta > 0$.

其圖形如下:

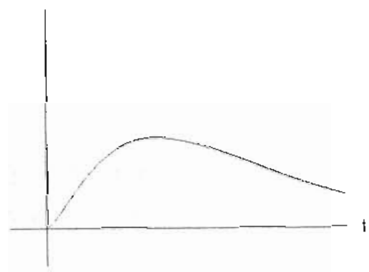


(1)

(2)



(3)



(4) $\Gamma(2,1)$

$\square$

註. 讀者如果覺得加上 0 的部份有些不自然, 也可以將相對應的隨機變數取值在恰當的範圍, 例如 (2) (3) (4) 可以想成取值在正實數的隨機變數.

3.1 $\boxed{\text{習題}}$ 驗證以上 (1)-(4) 都是機率密度函數.

3.2 $\boxed{\text{習題}}$ 令 $\mathrm{f}_X(t) = \dfrac{1}{\sqrt{2\pi}\sigma} e^{-\frac{(t-\mu)^2}{2\sigma^2}}$, $t \in \mathbb{R}$, μ, σ 是常數, $\sigma > 0$. 驗證 $\mathrm{f}_X(t)$ 是機率密度函數.

定義 **3.2** (獨立隨機變數) 我們說連續型隨機變數 X, Y 是獨立的, 如果對任意的 $a < b$, $c < d$, 滿足

$$\begin{aligned} \mathrm{P}(a \le X \le b, c \le Y \le d) &= \mathrm{P}(a \le X \le b) \cdot \mathrm{P}(c \le Y \le d) \\ = \int_a^b \mathrm{f}_X(t)\, dt \cdot \int_c^d \mathrm{f}_Y(s)\, ds &= \int_a^b \int_c^d (\mathrm{f}_X(t) \cdot \mathrm{f}_Y(s))\, ds\, dt \end{aligned}$$

這其實就是離散機率情況的連續版, 讀者可以比較一下.

$$\begin{aligned} \mathrm{P}(i \le X \le j, l \le Y \le m) &= \mathrm{P}(i \le X \le j) \cdot \mathrm{P}(l \le Y \le m) \\ = (\sum_{k=i}^{j} \mathrm{P}(X = k)) \cdot (\sum_{n=l}^{m} \mathrm{P}(Y = n)) &= \sum_{k=i}^{j} \sum_{n=l}^{m} (\mathrm{P}_X(k) \cdot \mathrm{P}_Y(n)) \end{aligned}$$

因此 $f_X(t), f_Y(s)$ 扮演的就是 $P_X(k), P_Y(n)$ 的角色, 而離散的聯合機率函數 $P_{XY}(k,n)$, 現在應代以連續型聯合機率密度函數 $f_{XY}(t,s)$, 其中在 X, Y 獨立時, $f_{XY}(t,s) = f_X(t)f_Y(s)$ (後面會用到這個性質).

定義 3.3 (期望值與變異數) X 爲連續型隨機變數, 定義 X 的期望值與變異數如下.

$$\begin{aligned} E(X) &= \int_{-\infty}^{\infty} t f_X(t)\, dt \\ \mathrm{Var}(X) &= \int_{-\infty}^{\infty} (t - E(X))^2 f_X(t)\, dt = E(X^2) - E^2(X) \end{aligned}$$

性質 3.1 第一節中, $E(X)$, $\mathrm{Var}(X)$ 的基本性質(性質 1.2, 1.4) 與柴比雪夫不等式(定理 1.1), 在連續型也一樣成立.

例. 令 $f_X(t) = \frac{1}{\sqrt{\pi}} e^{-t^2}$, $t \in \mathbb{R}$. 求 $E(X)$ 與 $\mathrm{Var}(X)$.

由習題 2.5,

$$\begin{aligned} E(X) &= \frac{1}{\sqrt{\pi}} \int_{-\infty}^{\infty} t e^{-t^2}\, dt = 0 \\ \mathrm{Var}(X) &= E(X^2) - E^2(X) \\ &= \frac{1}{\sqrt{\pi}} \int_{-\infty}^{\infty} t^2 e^{-t^2}\, dt \;-0 \\ &= \frac{1}{2} \end{aligned}$$

□

3.3 習題 令 $f_X(t) = \frac{1}{\sqrt{2\pi}\sigma} e^{-\frac{(t-\mu)^2}{2\sigma^2}}$, $t \in \mathbb{R}$, μ, $\sigma > 0$ 是常數. 求 $E(X)$ 與 $\mathrm{Var}(X)$.

(Ans. μ, σ^2.)

3.4 習題 (續上題) 若定義隨機變數 $Y = \frac{X-\mu}{\sigma}$, 利用性質3.1 與習題3.3, 説明 $E(Y) = 0$, $\mathrm{Var}(Y) = 1$.

問題是怎麼求習題3.4 中隨機變數 Y 的機率密度函數 $f_Y(t)$? 底下我們就來處理這個問題.

分配函數

與離散機率不同的是, 在連續型機率時, 機率密度函數的某個不定積分扮演重要的角色. 對於機率密度函數 $f_X(t)$, 定義分配函數

$$F_X(t) \equiv \int_{-\infty}^{t} f_X(s)\, ds = P(X \le t)$$

顯然

(1) $\mathrm{F}_X(\infty)=1$ (即 $\lim\limits_{b\to\infty}\mathrm{F}_X(b)=1$).

(2) $\mathrm{F}_X(t)$ 遞增. (why?)

(3) $\mathrm{F}'_X(t)=\mathrm{f}_X(t)$. (微積分基本定理)

$\mathrm{F}_X(t)$ 在計算機率密度函數時特別有用.

例. $\mathrm{f}_X(t)=\dfrac{1}{\sqrt{2\pi}\sigma}e^{-\frac{(t-\mu)^2}{2\sigma^2}}, t\in\mathbb{R}$. 若令隨機變數 $Y=\dfrac{X-\mu}{\sigma}$, 求 $\mathrm{f}_Y(t)$.

$$\begin{aligned}\mathrm{F}_Y(t) &= \mathrm{P}(Y\le t)=\mathrm{P}\Big(\frac{X-\mu}{\sigma}\le t\Big)\\ &= \mathrm{P}(X\le \mu+\sigma t)\\ &= \int_{-\infty}^{\mu+\sigma t}\frac{1}{\sqrt{2\pi}\sigma}e^{-\frac{(s-\mu)^2}{2\sigma^2}}\,ds\end{aligned}$$

由微積分基本定理, 得

$$\mathrm{f}_Y(t)=F'_Y(t)=\frac{1}{\sqrt{2\pi}}e^{-\frac{t^2}{2}}$$

□

注意. 對任何 $\mu,\sigma>0$, 上例 X 經過變換$Y=\frac{X-\mu}{\sigma}$ 後, 都會變成一標準形式:

$$\mathrm{f}_Y(t)=\frac{1}{\sqrt{2\pi}}e^{-\frac{t^2}{2}}$$

性質 **3.2** 若 X,Y 獨立, 且 $Z=X+Y$, 則

$$\mathrm{f}_Z(t)=\int_{-\infty}^{\infty}\mathrm{f}_X(t-s)\mathrm{f}_Y(s)\,ds=\int_{-\infty}^{\infty}\mathrm{f}_X(s)\mathrm{f}_Y(t-s)\,ds$$

註. 各位應注意到此公式與離散時類似:

$$\mathrm{P}(X+Y=k)=\sum_{i+j=k}\mathrm{P}_X(i)\mathrm{P}_Y(j)=\sum_i\mathrm{P}_X(i)\mathrm{P}_Y(k-i)=\sum_j\mathrm{P}_X(k-j)\mathrm{P}_Y(j)$$

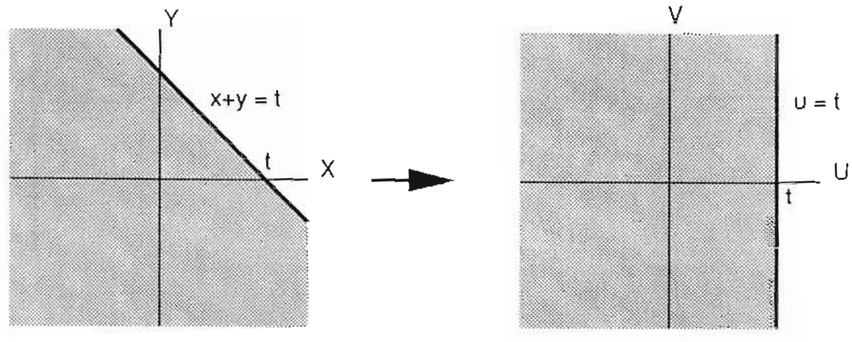
證明.

$$\begin{aligned}\mathrm{F}_Z(t) &= \mathrm{P}(X+Y\le t)\\ &= \iint_{x+y<t}\mathrm{f}_{XY}(x,y)\,dA\\ &= \int\int_{x+y\le t}\mathrm{f}_X(x)\mathrm{f}_Y(y)\,dx\,dy \quad \text{(這裡用到獨立的假設)}\end{aligned}$$

作變數變換 $u=x+y,\ v=x$ (見下圖)

$$
\begin{aligned}
&= \int_{-\infty}^{t}\int_{-\infty}^{\infty} \mathrm{f}_X(v)\mathrm{f}_Y(u-v)|J(u,v)|\,dv\,du \\
&= \int_{-\infty}^{t}\left(\int_{-\infty}^{\infty} \mathrm{f}_X(v)\mathrm{f}_Y(u-v)\,dv\right)\,du
\end{aligned}
$$

所以 $\mathrm{f}_Z(t)=\mathrm{F}'_Z(t)=\int_{-\infty}^{\infty}\mathrm{f}_X(v)\mathrm{f}_Y(t-v)\,dv$. (微積分基本定理)



□

例. $\mathrm{f}_X(t)=\mathrm{f}_Y(t)=\frac{1}{\sqrt{\pi}}e^{-t^2},\ t\in\mathbb{R},\ X,Y$ 獨立, 求 $\mathrm{f}_Z(t)$, 其中 $Z=X+Y$.

由定理知

$$
\begin{aligned}
\mathrm{f}_Z(t) &= \int_{-\infty}^{\infty}\mathrm{f}_X(t-s)\mathrm{f}_Y(s)\,ds \\
&= \int_{-\infty}^{\infty}\frac{1}{\sqrt{\pi}}e^{-(t-s)^2}\cdot\frac{1}{\sqrt{\pi}}e^{-s^2}\,ds \\
&= \int_{-\infty}^{\infty}\frac{1}{\pi}e^{-(2s^2-2st+t^2)}\,ds \\
&= \frac{1}{\pi}\cdot\int_{-\infty}^{\infty}e^{-2(s-\frac{t}{2})^2-\frac{t^2}{2}}\,ds \\
&\overset{u=s-\frac{t}{2}}{=} e^{-\frac{t^2}{2}}\cdot\frac{1}{\pi}\cdot\int_{-\infty}^{\infty}e^{-2u^2}\,du \\
&= e^{-\frac{t^2}{2}}\cdot\frac{1}{\pi}\cdot\sqrt{\frac{\pi}{2}}=\frac{1}{\sqrt{2\pi}}e^{-\frac{t^2}{2}} \quad (\text{習題}2.7)
\end{aligned}
$$

□

例. 若 $\mathrm{f}_X(t)=\mathrm{f}_Y(t)=e^{-t},\ t\geq 0;\ =0,\ t<0$. 設 X,Y 獨立, 求 $\mathrm{f}_{X+Y}(t)$.

由於函數有一不連續點 $t=0$, 使用定理比較囉嗦, 讓我們仔細檢查原來推導過程. 令 $Z=X+Y$

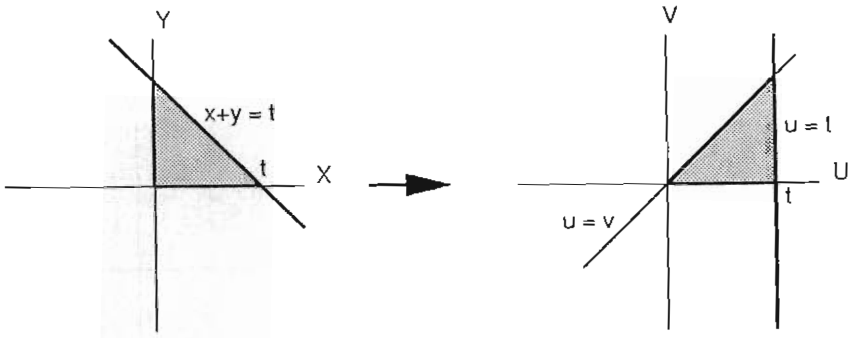
$$
\mathrm{F}_Z(t)=\mathrm{P}(X+Y\leq t)=\int\int_{x+y<t}\mathrm{f}_X(x)\mathrm{f}_Y(y)\,dx\,dy
$$

當 $t<0$, 顯然 $F_Z(t)=0$, 所以 $f_Z(t)=F'_Z(t)=0,\ t<0$; 當 $t>0$ 時, 積分區域是

$$x\ge 0,\quad y\ge 0,\quad x+y\le t$$

在作變數變換 $u=x+y,\ v=x$ 後, 得到新的積分區域為 (下圖)

$$u-v\ge 0,\quad v\ge 0,\quad u\le t$$



所以原積分 $=\displaystyle\int_0^t\int_0^u f_X(u-v)f_Y(v)\,dvdu$, 因此

$$f_Z(t)=F'_Z(t)=\int_0^t f_X(t-v)f_Y(v)\,dv$$

現代入函數e^{-t}, 得到

$$f_Z(t)=\int_0^t e^{v-t}\cdot e^{-v}\,dv=\int_0^t e^{-t}\,dv=te^{-t},\quad t>0$$

讀者可以自行檢查這是一個機率密度函數, 眼尖的人可能注意到這是一個 $\Gamma(2,1)$ 分配. (見習題 3.11)

□

這個例子也告訴我們, 真正關鍵的是分配函數的計算.

3.5 習題 驗證 $f_X(t)=\lambda e^{-\lambda t},\ t\ge 0,\ \lambda>0$ 是機率密度函數, 並求出 $E(X)$ 與 $Var(X)$.

(Ans. $\frac{1}{\lambda}$, $\frac{1}{\lambda^2}$.)

3.6 習題 求 $\Gamma(\alpha,\beta)$ 機率密度函數

$$f_X(t)=\frac{\beta^\alpha}{\Gamma(\alpha)}t^{\alpha-1}e^{-\beta t},\quad t>0,\alpha>0,\beta>0$$

之 $E(X)$, $Var(X)$. (Hint: 別忘了 $\Gamma(\alpha)=(\alpha-1)\Gamma(\alpha-1)$), $\alpha>1$.)

(Ans. $\frac{\alpha}{\beta}$, $\frac{\alpha}{\beta^2}$.)

3.7 $\boxed{\text{習題}}$ 若隨機變數 X 的機率密度函數為 $f_X(t)$, $t \in \mathbb{R}$.

(1) 說明隨機變數 $X+\alpha$ (α 為常數) 的機率密度函數為 $f_X(t-\alpha)$, $t \in \mathbb{R}$.

(2) 說明隨機變數 αX (α 為非零常數) 的機率密度函數 為 $\frac{1}{\alpha}f_X(\frac{t}{\alpha})$, $t \in \mathbb{R}$.

3.8 $\boxed{\text{習題}}$

(1) 若 $f_X(t) = f_Y(t) = \frac{1}{\sqrt{2\pi}\sigma}e^{-\frac{-(t-\mu)^2}{2\sigma^2}}$, X, Y 獨立, 令 $Z = X+Y$, 求 $f_Z(t)$.

(2) 由 (1), 若 $f_{X_i}(t) = \frac{1}{\sqrt{2\pi}\sigma}e^{-\frac{-(t-\mu)^2}{2\sigma^2}}$, X_i 互相獨立, 令 $Z = X_1 + \cdots X_n$, 求 $f_Z(t)$.

3.9 $\boxed{\text{習題}}$ 若 $f_X(t) = \frac{1}{\sqrt{2\pi}}e^{-\frac{t^2}{2}}$, 令隨機變數 $Y = X^2$, 求 $f_Y(t)$.
(Hint:寫出 $F_Y(t)$ 與微積分基本定理) (Ans. $\chi^2(1)$ 分配).

3.10 $\boxed{\text{習題}}$ 設 $m, n \in \mathbb{N}$.

(1) 利用分部積分法証明

$$\int_0^t (t-s)^m s^n \, ds = \frac{m}{n+1}\int_0^t (t-s)^{m-1} s^{n+1} \, ds$$

(2) 利用 (1), 說明

$$\int_0^t (t-s)^m s^n \, ds = \frac{n! \;\; m!}{(m+n+1)!} t^{m+n+1}$$

3.11 $\boxed{\text{習題}}$ 若 $X \sim \Gamma(m,\beta)$ 分配, $Y \sim \Gamma(n,\beta)$ 分配 (見 309頁), m, $n \in \mathbb{N}$, 利用上習題與上例, 證明 $X+Y \sim \Gamma(m+n,\beta)$ 分配.

4 Poisson 分配與指數分配

4.1 Poisson 分配

考慮下列現象: 每小時服務台訪客的人數, 每天家中電話的通數, 一本書中每頁的錯字數, 某條道路上每月發生車禍的次數, 生產線上的疵品數, 學生到辦公室找老師的次數...... 大致上都有一些共同的特徵: 在某時間區段內, 平均會發生若干次「事件」, 但是有時候很少, 有時又異常地多, 因此事件發生的次數是一個隨機變數, 它所對應的機率函數稱為 Poisson 分配.

一個 Poisson 過程有三個基本特性:

(1) 在一個短時間區間 Δt 內, 發生一次事件的機率與 Δt 成正比: $\lambda\Delta t$.

(2) 在短時間內發生兩次以上的機率可以忽略.

(3) 在不重疊的時間段落裡, 事件各自發生的次數是獨立的.

各位可以驗證上述各種實際的例子，是不是相當符合 Poisson 過程的定義？

現令 $P(k,T)$ 表示在時間區間 T 中發生 k 次事件的機率 (注意T 表示時間區間的長度，而不是絕對時間)，由 (1) (2) 知$P(1,\Delta t)=\lambda\Delta t$，且 $P(k,\Delta t)=0$, $k\geq 2$. 現將 T 分割成 N 個短時間區段 (即 $T=N\Delta t$)，利用 (3) 各時間區段出現之事件是獨立的條件，可知

$$\begin{aligned}P(k,T) &\approx C_k^N(\lambda\Delta t)^k(1-\lambda\Delta t)^{N-k} \qquad (\text{二項分配})\\ &= \frac{N(N-1)\cdots(N-k+1)}{k!}\cdot\lambda^k\cdot\frac{T^k}{N^k}\cdot\frac{(1-\frac{(\lambda T)}{N})^N}{(1-\frac{(\lambda T)}{N})^k}\\ &= \frac{(\lambda T)^k}{k!}\cdot\frac{N(N-1)\cdots(N-k+1)}{N\cdot N\cdots N}\cdot\frac{(1-\frac{\lambda T}{N})^N}{(1-\frac{\lambda T}{N})^k}\end{aligned}$$

固定 k，當 $N\to\infty$ 時

$$P(k,T)=\frac{(\lambda T)^k}{k!}e^{-(\lambda T)} \quad (\text{記得}\,(1+\frac{\alpha}{N})^N\to e^{\alpha})$$

由上可知 Poisson 分配是二項分配 $B(N,p,q)$ 的一種極限，其中$Np=$ 常數 λT，再讓 $N\to\infty$. 另外，我們通常將 λT 記為 m，表示在時間區間 T 中，平均的發生次數(見下面習題).

4.1 $\boxed{\text{習題}}$

(1) 驗證 $\sum_{k=0}^{\infty}P(k,T)=1$.

(2) 令 $P_X(k)=\frac{m^k}{k!}e^{-m}$, $k=1,2,3,\cdots$. 求 $E(X)$ 與 $\text{Var}(X)$.

(Ans. m, m.)

4.2 $\boxed{\text{習題}}$ 設隨機變數 X 遵守 Poisson 分配，且

$$P_X(k)=\frac{m^k}{k!}e^{-m}, \quad k=1,2,3,\cdots$$

若 $X_1,X_2\sim X$, X_1, X_2 獨立，且令 $Y=X_1+X_2$，說明

$$f_Y(k)=\frac{(2m)^k}{k!}e^{-2m}, \quad k=1,2,3,\cdots$$

例. 一公司之電話通數大約每小時 20 通，求在 5 分鐘內沒有一通電話的機率？

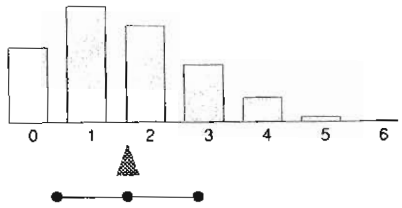
每小時 20 通，表示每分鐘平均 $\lambda=\frac{1}{3}$ 通/分. 因此在 5 分鐘的時間區間中，平均的電話通數為 $m=\lambda T=\frac{1}{3}\times 5=\frac{5}{3}$. 所以

$$P(k)\equiv P(k,5)=\frac{(\frac{5}{3})^k}{k!}e^{-\frac{5}{3}}, \quad k=0,1,2,\cdots$$

所以沒有一通電話的機率 $P(0)=e^{-\frac{5}{3}}\approx 0.19$. 有了 $P(k)$，我們可以回答許多類似的問題：在 5 分鐘內有 4 通電話的機率是 $P(4)=\frac{\frac{5}{3}^4}{4!}e^{-\frac{5}{3}}\approx 0.06$，大概每十六次才有一次. 在 5 分鐘內有超過 3 通電話的機率是

$$\sum_{k=4}^{\infty}\frac{(\frac{5}{3})^k}{k!}e^{-\frac{5}{3}}=1-\sum_{k=0}^{3}\frac{(\frac{5}{3})^k}{k!}e^{-\frac{5}{3}}\approx 0.09$$

由習題4.1, 這個機率分配的期望值 $=\frac{5}{3}$, 標準差 $=\sqrt{\frac{5}{3}}\approx 1.29$. 右圖是 $\mathrm{P}(k)$ 的圖形, 當然由於 $k=0,1,\cdots$, 所以這只是部分圖形. 讀者可與 300 頁二項分配的圖形比較.



$\mathrm{E}(X)=\frac{5}{3}\approx 1.67$

$\sqrt{\mathrm{Var}(X)}=\sqrt{\frac{5}{3}}\approx 1.29$

□

4.3 $\boxed{\text{習題}}$ 固定 m. 說明函數 $\mathrm{P}(k)=\frac{m^k}{k!}e^{-m}$, $k=1,2,\cdots$, 在 $m-1\le k\le m$ 時為最大值.

例. 下表是 1910 年 Rutherford 觀察放射性物質放射 α 粒子的記錄, 每次觀察 7.5秒, 共觀察 2608 次.

粒子數	次數	頻率	$\mathrm{P}(k)$
0	57	0.022	0.021
1	203	0.078	0.081
2	383	0.147	0.156
3	525	0.201	0.201
4	532	0.204	0.195
5	408	0.156	0.151
6	273	0.105	0.097
7	139	0.053	0.054
8	45	0.017	0.026
9	27	0.010	0.011
≥ 10	16	0.006	0.007

$m=\frac{10094}{2608}=3.87$

這裡 $\mathrm{P}(k)=\mathrm{P}(k,7.5)$, 其中 $\mathrm{P}(k)=\frac{m^k}{k!}e^{-m}$, $m=3.87$(見表最末欄), 為 7.5 秒中 α 粒子放射之平均個數. 可以看到, 如果假設 α 粒子的放射是一Poisson過程, 結果相當吻合.

□

例. 令一放射性物質在時間 t 時所含之放射性粒子總量為 $N(t)$, 如果假設放射粒子是一 Poisson 過程, 則在短時間 Δt 後,

$$N(t+\Delta t) - N(t) = -(\lambda \Delta t) N(t)$$

注意到 $(\lambda \Delta t)N(t)$ 是一期望值的形式. 所以

$$N'(t) \approx \frac{N(t+\Delta t) - N(t)}{\Delta t} = -\lambda N(t)$$

這就是上學期輻射定律的「證明」.

□

4.2 指數分配

令 W 表示在 Poisson 過程中, 由開始到第一次事件發生的時間 (這是一隨機變數). 由前例知

$$\begin{aligned} \mathrm{P}(W > t) &= \mathrm{P}(\text{在 } [0,t] \text{ 中無事件發生}) \\ &= e^{-\lambda t}, \quad t > 0 \end{aligned}$$

但

$$\mathrm{F}_W(t) = \mathrm{P}(W \le t) = 1 - \mathrm{P}(W > t) = 1 - e^{-\lambda t}$$

所以

$$\mathrm{f}_W(t) = \mathrm{F}'_W(t) = \lambda e^{-\lambda t}, \quad t > 0$$

這個機率分配稱為指數分配. 由習題3.5知$\mathrm{E}(W) = \frac{1}{\lambda}$, 這就是第一次事件發生的平均時間. 另外, $\mathrm{Var}(W) = \frac{1}{\lambda^2}$.

注意. 讀者應與第一節 293 頁之離散隨機變數 W 比較.

指數分配與排隊理論

排隊的現象無所不在: 買各種票, 吃自助餐, 超商, 百貨公司,... 等. 顧客揣度「應該排那一服務櫃台會比較快?」「到底還要排多久?」是城市生活的基本問題; 相對的, 商家也要盤算到底在何時要開幾個窗口櫃台才符合成本, 探討這個問題的數學理論通稱為排隊理論, 而指數分配經常被用到排隊理論, 當作服務客人時間 (這是一隨機變數) 的機率密度函數.

讓我們假設某櫃台, 服務客人的平均時間為 μ, 想像在服務結束後, 櫃員會亮燈請下一位客人進來, 則亮燈的平均時間是 μ. 若將「燈亮」視為一事件發生, 則亮燈的過程近似於一 Poisson 過程. 而且前面定義的 W 正好表示兩次亮燈間的間隔. 所以 W 的機率密度函數是指數分配:

$$\mathrm{f}_W(t) = \frac{1}{\mu} e^{-\frac{1}{\mu}t}, \quad t > 0$$

例. 現假設一櫃台平均服務時間為 3 分鐘, 設等待時間的機率密度函數為

$$f_W(t)=\frac{1}{3}e^{-\frac{t}{3}},\quad t>0$$

(1)等候時間超過6分鐘的機率是多少?

$$\begin{aligned}P(W>6) &= \int_6^\infty \frac{1}{3}e^{-\frac{t}{3}}\,dt=\lim_{b\to\infty}(-e^{-\frac{t}{3}})|_6^b\\ &= e^{-2}\approx 0.14\end{aligned}$$

事實上, 等候超過 T 分鐘的機率是 $e^{-\frac{T}{3}}$.

(2) 另一個合理的問題是, 如果在我前面還有另一個客人, 則我怎麼描述, 我等待時間的機率分配呢?

令 W_1 是第一個客人等待的時間, W_2 是第一個客人開始被服務後, 我所等待的時間, 則 $W_1\sim W_2\sim W$, 而且總等待時間 $U=W_1+W_2$, 另外顯然 W_1 與 W_2 是互相獨立的. 所以我們的問題就是要計算 $f_U(t)$, 由 313 頁例子的方法, 可以計算得

$$f_U(t)=\frac{1}{9}te^{-\frac{t}{3}},\quad t>0$$

或者, 如果將指數分配 $f_W(t)$ 想成是 $\Gamma(1,\frac{1}{3})$ 分配, 則此相當於

$$\Gamma(1,\frac{1}{3})+\Gamma(1,\frac{1}{3})\sim\Gamma(2,\frac{1}{3})\qquad (\text{習題}3.11)$$

因此如果我們想知道總等候時間不超過5分鐘的機率, 則

$$\begin{aligned}P(U\le 5) &= \int_0^5 \frac{1}{9}te^{-\frac{t}{3}}\,dt\\ &= \frac{1}{9}\left((-3te^{-\frac{t}{3}})|_0^5+3\int_0^5 e^{-\frac{t}{3}}\,dt\right)\\ &\approx 0.5\end{aligned}$$

有一半的機會.

(3) 如果前面有 $n-1$ 個客人時, 則可定義 $U=W_1+W_2+\cdots+W_n$, 其中 W_i 彼此獨立, 由習題 3.11知 $U\sim\Gamma(n,\frac{1}{3})$, 即

$$f_U(t)=\frac{1}{3^n(n-1)!}t^{n-1}e^{-\frac{t}{3}},\quad t>0$$

這告訴我們 $\Gamma(\alpha,\beta)$ 分配與排隊理論的關係. 我們將細節留作習題.

□

4.4 $\boxed{\text{習題}}$ 若 $\mathrm{P}_X(i) = \frac{m^i}{i!}e^{-m}$, $i = 0, 1, 2, \cdots$. $\mathrm{P}_Y(j) = \frac{l^j}{j!}e^{-l}$, $j = 0, 1, 2, \cdots$. 設 X, Y 獨立, 令 $Z = X + Y$, 求 $\mathrm{P}_Z(k)$?

4.5 $\boxed{\text{習題}}$ 若 $X_1, \cdots, X_n$ 皆遵守 Poisson 分配, $\mathrm{P}_{X_i}(k) = \frac{m^k}{k!}e^{-m}$, $k = 0, 1, 2, \cdots$, 且 X_i 彼此獨立, 求 $Z = X_1 + \cdots + X_n$ 的機率函數 $\mathrm{P}_Z(k)$?

4.6 $\boxed{\text{習題}}$ 若 X, Y 獨立並都遵守指數分配, $\lambda e^{-\lambda t}$, $t > 0$. 令 $Z = X + Y$, 求 $\mathrm{f}_Z(t)$.
(Ans. $\mathrm{f}_Z(t) = \lambda^2 t e^{-\lambda t}$, $t > 0$.)

4.7 $\boxed{\text{習題}}$ 若 X_i 遵守指數分配, $\lambda e^{-\lambda t}$, $t > 0$. 而且假設 $X_1, X_2, \cdots, X_n$ 相互獨立.
(1) 令 $Z = X_1 + X_2 + \cdots + X_n$. 說明

$$\mathrm{f}_Z(t) = \frac{\lambda^n}{(n-1)!}t^{n-1}e^{-\lambda t} \qquad (\Gamma(n, \lambda)\text{分配})$$

(2) 定義 $\bar{X} = \frac{X_1 + \cdots + X_n}{n}$, 說明 $\bar{X} \sim \Gamma(n, n\lambda)$ 分配.

4.8 $\boxed{\text{習題}}$ 校一本書, 假設每頁平均可以發現 1 個錯誤, 求在某頁發現1 個錯誤的機率? 求在 4 頁中發現 4 個錯誤的機率?
(Ans. $\frac{1}{e} \approx 0.368$; 0.195.)

4.9 $\boxed{\text{習題}}$ 一彎道平均每月發生 10 次車禍, 請問在 10 天內僅發生一件車禍的機率?
(Ans. 0.119.)

4.10 $\boxed{\text{習題}}$ (1) 超級市場一服務員平均服務時間為 2 分鐘, 設用指數分配當作等候時間之機率分配, 則機率密度函數是什麼?

(2) 如果他正開始服務一位客人, 而你前面還有一位客人在等候, 則你會等超過 6 分鐘的機率是多少?

(3) 若服務員甲平均服務時間為 2 分鐘, 而服務員乙之平均服務時間為 3 分鐘, 如果你選擇乙, 你朋友選擇甲, 且一起開始接受服務, 則你會比朋友快的機率是多少? (當然甲與乙的服務是相互獨立的) 你能給出一個一般的計算公式嗎?

4.3 應用: 可靠性

人類的工作環境非常依賴機器設備, 各式工廠的生產線, 家庭中的電器或電線配置 (例如簡單的電燈配置), 公司內的電話, 電腦, 影印機等的聯線配置, 汽車內必要零件的配置等. 但是設備總有故障的時候, 因此以最少成本裝配出盡量少故障的配置方式是機率理論的一種應用.

令 L 表示某裝置壽命的隨機變數, 一設備的可靠度定義成

$$R_L(t) = \mathrm{P}(L > t) \quad (\text{即到時間 } t \text{ 還未故障的機率})$$

注意到 $R_L(t)$ 與 $f_L(t)$ 的關係是

$$f_L(t) = F'_L(t) = (P(L \le t))' = (1 - P(L > t))' = -R'_L(t)$$

所以設備的平均壽命為

$$\begin{aligned} E(L) &= \int_0^\infty t f_L(t)\,dt \\ &= \int_0^\infty t(-R'_L(t))\,dt \\ &= -tR_L(t)|_0^\infty + \int_0^\infty R_L(t)\,dt \\ &= \int_0^\infty R_L(t)\,dt \end{aligned}$$

(常識說 $P(L >$ 很大$N)$ 的機率非常非常小, 不妨假設 $\lim\limits_{t\to\infty} tR_L(t) = 0$)

例. 一個常用的壽命機率分配是指數分配: $\mu e^{-\mu t}$, $t > 0$.

這相當於說, 在一定時間內, 一設備故障的次數遵守 Poisson 分配, 而壽命相當於等候它故障的時間. 因此這裡 L 其實就是前節的 W.

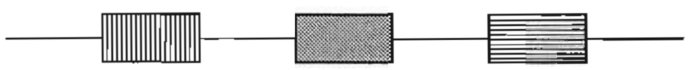
由前知

$$\begin{aligned} R_L(t) &= \int_t^\infty \mu e^{-\mu t} = e^{-\mu t} \\ E(L) &= \int_0^\infty R(t)\,dt = \frac{1}{\mu} \end{aligned}$$

□

配置有兩種最基本的型態: 串聯與並聯

(一) 串聯: 有一個設備或零件故障, 系統就故障.

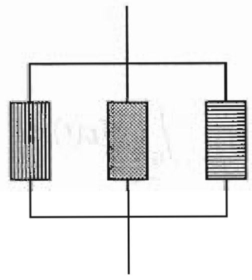


將串聯在一起的設備 $L_1, L_2, \cdots, L_n$, 看成一部大機器 $\hat{L}$. 則此大機器的可靠性 $R_{\hat{L}}(t)$ 是各零件設備可靠性$R_i(t) \equiv R_{L_i}(t)$ 的某種函數: 假設各設備之運作獨立, 則

$$\begin{aligned} R_{\hat{L}}(t) &= P(\hat{L} > t) \\ &= P(L_1 > t)\cdot P(L_2 > t)\cdots P(L_n > t) \\ &= R_1(t)\cdot R_2(t)\cdots R_n(t) \end{aligned}$$

例如假設各零件之可靠性為 $e^{-\lambda_i t}$, 則此串聯機器之可靠性為 $e^{-(\lambda_1+\cdots+\lambda_n)t}$.

(二) 並聯: 只要有一個設備不故障則整個系統就可以運轉.



$$
\begin{aligned}
R_{\hat{L}}(t) &= \mathrm{P}(\hat{L} > t) = 1 - \mathrm{P}(\hat{L} \le t) \\
&= 1 - \mathrm{P}(L_1 \le t) \cdot \mathrm{P}(L_2 \le t) \cdots \mathrm{P}(L_n \le t) \\
&= 1 - (1 - R_1(t)) \cdot (1 - R_2(t)) \cdots (1 - R_n(t))
\end{aligned}
$$

例如並聯兩上述設備之可靠度為:

$$1 - (1 - e^{-\lambda_1 t})(1 - e^{-\lambda_2 t}) = e^{-\lambda_1 t} + e^{-\lambda_2 t} - e^{-(\lambda_1 + \lambda_2) t}$$

例. 比較串聯與並聯兩設備(可靠度都是 $e^{-\frac{t}{\mu}}$) 之壽命.

$$
\begin{aligned}
&\text{串聯} \quad \mathrm{E}(\hat{L}) = \int_0^\infty R_{\hat{L}}(t) = \int_0^\infty e^{-\frac{2t}{\mu}}\, dt = \frac{\mu}{2} \\
&\text{並聯} \quad \mathrm{E}(\hat{L}) = \int_0^\infty (2e^{-\frac{t}{\mu}} - e^{-\frac{2t}{\mu}})\, dt = 2\mu - \frac{\mu}{2} = \frac{3}{2}\mu
\end{aligned}
$$

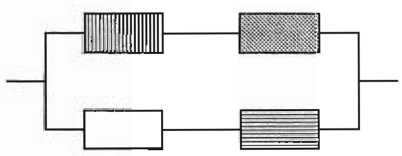
– –

4.11 習題 同上例考慮 n 部可靠度都是 $e^{-\frac{t}{\mu}}$ 的設備, 證明

$$\text{串聯}: \mathrm{E}(\hat{L}) = \tfrac{\mu}{n} \qquad \text{並聯}: \mathrm{E}(\hat{L}) = (1 + \tfrac{1}{2} + \cdots + \tfrac{1}{n})\mu$$

4.12 習題

(1) 一系統配置如右圖, 假設單一零件壽命皆遵守指數分配, 且平均壽命均為 6 個月, 求此配置系統之平均壽命?



(2) 求出整個系統壽命之機率密度函數並作圖.

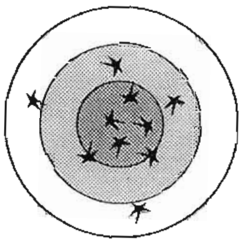
5 常態分配

5.1 常態分配

常態分配是 Gauss 為了解釋預測行星位置所發展出來的機率模型, 廣泛地使用於測量誤差的理論. 常態分配最重要的性質是: 做測量時, 經常會受到一系列微小的隨機因素影響而產生誤差, 這些誤差可能遵守不同的機率分配, 但是只要這些誤差因素彼此的相關性不大 (例如彼此皆獨立), 則誤差的總和會相當接近於一常態分配.

註. 除了誤差的應用, 常態分配最常被使用來近似地描述生物或社會上群體的統計現象, 例如人的壽命, 產品的壽命, 身高, 體重, 智商, 考試成績及度量種種變態或異常的人格, 犯罪心理等等. 將常態分配的鐘型曲線套用在社會分析上經常引起大規模的辯論. 其中最極端的例子是 「某某人種先天智力比較差」.

底下我們要用一個想法來導出常態分配, 想像用一把好槍去打靶, 從常識我們知道, 雖然射擊者瞄準靶心, 但是彈著點的分佈會近似於右圖: 它們一定會有誤差, 但是在靶心周圍的彈著點應該要比外環的點要密集, 另外彈著點集中的方式, 不會特別偏向於水平方向或是垂直方向, 事實上彈著分布應該只與到靶心的距離有關.



現令 (X, Y) 為彈著點的座標 (以靶心為原點), 顯然 X 與 Y 都是隨機變數, 而且彼此應該是獨立的. 由以上的常識觀察, 我們知道

(1) $\mathrm{f}_X(t) = \mathrm{f}_Y(t)$ (對稱性).

(2) $\mathrm{f}_{XY}(t,s) = \mathrm{f}_X(t) \cdot \mathrm{f}_Y(s)$ (因為 X, Y 獨立).

(3) $\mathrm{f}_{XY}(t,s) = g(\sqrt{t^2+s^2})$ (因為彈著分布只與到原點距離有關).

從 (1) (2) (3), 我們想要導出 $\mathrm{f}_X(t)$ 的表式, 現令 $f(t) = \mathrm{f}_X(t)$, $F(t,s) = \mathrm{f}_{XY}(t,s)$, 則由 (1) (2) (3)

$$g(\sqrt{s^2+t^2}) = F(t,s) = f(t) \cdot f(s)$$

又令 $s = 0$, 則 $g(|t|) = f(0) \cdot f(t)$, 因此 $f(t)$ 滿足 $f(t) = f(-t)$ (偶函數), $f(0) \neq 0$ 而且

$$g(\sqrt{s^2+t^2}) = f(0) \cdot f(\sqrt{s^2+t^2})$$

所以 $f(0) \cdot f(\sqrt{s^2+t^2}) = f(t) \cdot f(s)$, 兩邊同除 $f^2(0)$ 得

$$\frac{f(\sqrt{s^2+t^2})}{f(0)} = \frac{f(s)}{f(0)} \cdot \frac{f(t)}{f(0)}$$

現令 $\phi(t) = \ln\frac{f(t)}{f(0)}$, 則上式可改寫成

$$\phi(\sqrt{s^2+t^2}) = \phi(s) + \phi(t)$$

Claim: $\phi(t) = \phi(1)\cdot t^2$, $t \in \mathbb{R}$.

證明.

由上式, 可以導得

$$\begin{aligned}
\phi\left(\sqrt{t_1^2+t_2^2+\cdots+t_n^2}\right) &= \phi\left(\sqrt{\left(\sqrt{t_1^2+t_2^2+\cdots+t_{n-1}^2}\right)^2+t_n^2}\right) \\
&= \phi\left(\sqrt{t_1^2+t_2^2+\cdots+t_{n-1}^2}\right)+\phi(t_n) \\
&\vdots \qquad\quad \vdots \\
&= \phi(t_1)+\phi(t_2)+\cdots+\phi(t_n)
\end{aligned}$$

所以如果令 $t_1 = t_2 = \cdots = t_n = t$, 而且取 n 是平方數 $n = k^2$, 則

$$\begin{cases} \phi(kt) = k^2\phi(t) \\ \phi(k) = k^2\phi(1) \end{cases} \quad \text{其中 } k \in \mathbb{N}$$

令 $t = \frac{m}{k}$, $m, k \in \mathbb{N}$, 代入上第一式, 得到

$$\phi(m) = k^2\phi(\frac{m}{k})$$

但由第二式, $\phi(m) = m^2\phi(1)$, 所以

$$\phi(\frac{m}{k}) = (\frac{m}{k})^2\phi(1)$$

這表示對所有正有理數 t, $\phi(t)$ 滿足

$$\phi(t) = \phi(1)\cdot t^2$$

因此, 此式對所有實數都成立 (當然我們假設 $\phi(t)$ 連續).

□

所以原來的 $f(t)$ 滿足

$$f(t) = f(0)\cdot e^{\phi(t)} = f(0)\cdot e^{\phi(1)t^2} = \alpha e^{\beta t^2}$$

由於 $f(t)$ 是機率密度函數, 必須滿足 $\int_{-\infty}^{\infty} f(t)\, dt = 1$, 所以 $\beta < 0$, 令 $\beta = -\frac{1}{2\sigma^2}$, 則由前面瑕積分的例子, 我們知道 $f(t)$ 或 $\mathrm{f}_X(t)$ 必須等於

$$\mathrm{f}_X(t) = f(t) = \frac{1}{\sqrt{2\pi}\sigma}e^{-\frac{t^2}{2\sigma^2}}, \quad t \in \mathbb{R}$$

這正是常態分配的基本機率密度函數, 前面我們已經計算過

$$\mathrm{E}(X)=0, \quad \mathrm{Var}(X)=\sigma^2$$

這表示彈著點雖然有誤差, 不過其平均值應該是原點. 而 σ 是不能避免的參數, 可以拿來鑑別射手的技術. σ 越小, 表示此射手越近於一名神槍手.

定義 5.1 (常態分配; $N(\mu,\sigma^2)$) 一常態分配 $N(\mu,\sigma^2)$ 的機率密度函數爲

$$\frac{1}{\sqrt{2\pi}\sigma}e^{-\frac{(t-\mu)^2}{2\sigma}}, \qquad t\in\mathbb{R}$$

由第三節312頁的例子與習題3.3, 3.8, 我們知道

性質 5.1 (常態分配的性質)

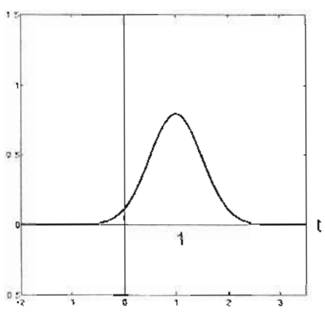
(1) $X\sim N(\mu,\sigma^2)$, 則 $\mathrm{E}(X)=\mu, \mathrm{Var}(X)=\sigma^2$.

(2) 若 $X_1, \cdots, X_n$ 獨立, $X_i\sim N(\mu,\sigma^2)$, 則

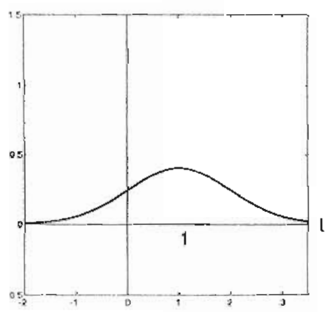
$$(X_1+\cdots+X_n) \quad \sim \quad N(n\mu, n\sigma^2)$$

(3) 設 $X\sim N(\mu,\sigma^2)$, 令 $Y=\frac{X-\mu}{\sigma}$, 則 $Y\sim N(0,1)$.

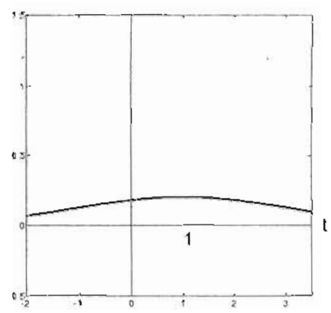
底下為 $N(1,\sigma^2)$ 在 $\sigma=\frac{1}{2}$, 1, 2 的圖形.



$\sigma=\frac{1}{2}$



$\sigma=1$



$\sigma=2$

5.1 $\boxed{\text{習題}}$ 若 $X\sim N(\mu_1,\sigma_1^2)$, $Y\sim N(\mu_2,\sigma_2^2)$, 且 X, Y 獨立, 驗証 $X+Y \sim N(\mu,\sigma^2)$, 並計算 μ, σ 的公式.

5.2 $\boxed{\text{習題}}$ 若 $X_1, \cdots, X_n$ 獨立, $X_i\sim N(\mu,\sigma^2)$. 定義隨機變數 $\bar{X}=\frac{X_1+\cdots+X_n}{n}$, $n\in\mathbb{N}$, 說明 $\bar{X}\sim N(\mu,\frac{\sigma^2}{n})$.

5.3 $\boxed{\text{習題}}$ 設某桌長度為 140 cm, 某人測量此桌長, 記為 X cm (此為一隨機變數), 設 $X\sim N(140, 0.25)$. 若將結果改用公厘記為 Y mm, 請問 Y 遵守怎樣的機率分配? 注意到你不能理所當然地假設 Y 遵守常態分配.

設 $X \sim N(0,1)$, 定義 $\rho(x) = \mathrm{P}(0 \le X \le x) = \int_0^x \frac{1}{\sqrt{2\pi}} e^{-\frac{t^2}{2}}\, dt$, 下表是 $\rho(x)$ 的數值計算表.

$\rho(x)$	.00	.01	.02	.03	.04	.05	.06	.07	.08	.09
0.0	.0000	.0040	.0080	.0120	.0160	.0199	.0239	.0279	.0319	.0359
0.1	.0398	.0438	.0478	.0517	.0557	.0596	.0636	.0675	.0714	.0753
0.2	.0793	.0832	.0871	.0910	.0948	.0987	.1026	.1064	.1103	.1141
0.3	.1179	.1217	.1255	.1293	.1331	.1368	.1406	.1443	.1480	.1517
0.4	.1554	.1591	.1628	.1664	.1700	.1736	.1772	.1808	.1844	.1879
0.5	.1915	.1950	.1985	.2019	.2054	.2088	.2123	.2157	.2190	.2224
0.6	.2257	.2291	.2324	.2357	.2389	.2422	.2454	.2486	.2517	.2549
0.7	.2580	.2611	.2642	.2673	.2704	.2734	.2764	.2794	.2823	.2852
0.8	.2881	.2910	.2939	.2967	.2995	.3023	.3051	.3078	.3106	.3133
0.9	.3159	.3186	.3212	.3238	.3264	.3289	.3315	.3340	.3365	.3389
1.0	.3413	.3438	.3461	.3485	.3508	.3531	.3554	.3577	.3599	.3621
1.1	.3643	.3665	.3686	.3708	.3729	.3749	.3770	.3790	.3810	.3830
1.2	.3849	.3869	.3888	.3907	3925	.3944	.3962	.3980	.3997	.4015
1.3	.4032	.4049	.4066	.4082	.4099	.4115	.4131	.4147	.4162	.4177
1.4	.4192	.4207	.4222	.4236	.4251	.4265	.4279	.4292	.4306	.4319
1.5	.4332	.4345	.4357	.4370	.4382	.4394	.4406	.4418	.4429	.4441
1.6	.4452	.4463	.4474	.4484	.4495	.4505	.4515	.4525	.4535	.4545
1.7	.4554	.4564	.4573	.4582	.4591	.4599	.4608	.4616	.4625	.4633
1.8	.4641	.4649	.4656	.4664	.4671	.4678	.4686	.4693	.4699	.4706
1.9	.4713	.4719	.4726	.4732	.4738	.4744	.4750	.4756	.4761	.4767
2.0	.4772	.4778	.4783	.4788	.4793	.4798	.4803	.4808	.4812	.4817
2.1	.4821	.4826	.4830	.4834	.4838	.4842	.4846	.4850	.4854	.4857
2.2	.4861	.4864	.4868	.4871	.4875	.4878	.4881	.4884	.4887	.4890
2.3	.4893	.4896	.4898	.4901	.4904	.4906	.4909	.4911	.4913	.4916
2.4	.4918	.4920	.4922	.4925	.4927	.4929	.4931	.4932	.4934	.4936
2.5	.4938	.4940	.4941	.4943	.4945	.4946	.4948	.4949	.4951	.4952
2.6	.4953	.4955	.4956	.4957	.4959	.4960	.4961	.4962	.4963	.4964
2.7	.4965	.4966	.4967	.4968	.4969	.4970	.4971	.4972	.4973	.4974
2.8	.4974	.4975	.4976	.4977	.4977	.4978	.4979	.4979	.4980	.4981
2.9	.4981	.4982	.4982	.4983	.4984	.4984	.4985	.4985	.4986	.4986
3.0	.4987	.4987	.4987	.4988	.4988	.4989	.4989	.4989	.4990	.4990

由於 $f_X(t)$ 圖形對稱於 y-軸, 因此我們可以利用此表計算 $N(0,1)$ 的機率.

例. $\mathrm{P}(-1 \le X \le 2) = \rho(1) + \rho(2)$, $\mathrm{P}(1 \le X \le 2) = \rho(2) - \rho(1)$.

易知

$$\mathrm{P}(-\alpha \le X \le 0) = \int_{-\alpha}^{0} f_X(t)\, dt = \int_0^{\alpha} f_X(t)\, dt = \mathrm{P}(0 \le X \le \alpha) = \rho(\alpha)$$

所以

(1) $\mathrm{P}(-1 \le X \le 2) = \mathrm{P}(-1 \le X \le 0) + \mathrm{P}(0 \le X \le 2) = \rho(1) + \rho(2)$

(2) $\mathrm{P}(1 \le X \le 2) = \mathrm{P}(0 \le X \le 2) - \mathrm{P}(0 \le X \le 1) = \rho(2) - \rho(1)$

□

5.4 習題 說明 $\mathrm{P}(-\epsilon \le X \le \epsilon) = 2\rho(\epsilon)$.

5.5 習題 運用此表計算

$$\mathrm{P}(-2 \le X \le -1.5),\quad \mathrm{P}(-2 \le X \le 2),\quad \mathrm{P}(1 \le X),$$
$$\mathrm{P}(-3 \le X),\quad \mathrm{P}(X \le -1.44),\quad \mathrm{P}(X \le 2.5).$$

(Ans. 0.044, 0.9544, 0.1587, 0.9987, 0.0749, 0.9938.)

5.6 習題 驗證

$$\begin{aligned}
\mathrm{P}(|X| \le 1.64) &\approx 90\% \\
\mathrm{P}(|X| \le 1.96) &\approx 95\% \\
\mathrm{P}(|X| \le 2.57) &\approx 99\%
\end{aligned}$$

註. 本習題的數值在下節將有應用.

有趣的是, 本表也可以用來計算一般常態分配的機率.

例. 設 $X \sim N(-3, 4)$, 計算 $\mathrm{P}(-5 \le X \le 0)$.

由性質 5.1, 令 $Y = \frac{X-(-3)}{2}$, 則 $Y \sim N(0,1)$, 因此

$$\begin{aligned}
\mathrm{P}(-5 \le X \le 0) &= \mathrm{P}(-1 \le \frac{X+3}{2} \le \frac{3}{2}) \\
&= \mathrm{P}(-1 \le Y \le 1.5) = \rho(1) + \rho(1.5) \\
&\approx 0.7745
\end{aligned}$$

□

5.7 習題 設 $X \sim N(\mu, \sigma^2)$, 計算 $\mathrm{P}(\mu - (2.57)\sigma \le X \le \mu + (2.57)\sigma)$. (Ans. 0.99.)

5.8 習題 設某桌長度為 100 cm, 某人測量此桌長, 記為 X cm, 設 $X \sim N(100, 0.25)$. 現此人測量桌長 25 次, 再求其平均值記為 S cm, 則
(1) 誤差 $|S - 100| \le 0.1$ cm 的機率是多少?
(2) 若希望誤差小於 0.05 cm 的機率 $\ge 99\%$, 則至少要測量幾次?

(Ans. 0.68; 661 次.)

5.3 中央極限定理

為什麼這個表這麼重要呢? 這是因為下面這個類似於大數法則, 令人驚訝的定理.

定理 5.1 (中央極限定理) 若隨機變數 $X_1, X_2, \cdots, X_n \sim X$ 來自於同一機率分配, $\mathrm{E}(X) = \mu$, $\mathrm{Var}(X) = \sigma^2$, 假設 X_i 彼此獨立. 令 $\bar{X} = \frac{X_1+\cdots+X_n}{n}$, 則隨機變數 $\frac{\bar{X}-\mu}{\frac{\sigma}{\sqrt{n}}} \to N(0,1)$, 當 $n \to \infty$. 即

$$\mathrm{P}\Big(\frac{\bar{X}-\mu}{\frac{\sigma}{\sqrt{n}}} \le t\Big) \approx \int_{-\infty}^{t} \frac{1}{\sqrt{2\pi}} e^{-\frac{s^2}{2}}\, ds,\quad \text{當 } n \text{ 夠大時}$$

例. 若 $X_1,\cdots,X_n$ 獨立, $X_i\sim N(\mu,\sigma^2)$, 則由習題 5.2

$$\frac{\bar{X}-\mu}{\frac{\sigma}{\sqrt{n}}} \sim N(0,1), \quad n\in\mathbb{N}$$

□

注意到, 這個定理並不局限在常態分配, 以二項分配為例, 我們有

定理 5.2 (De Moivre-Laplace)
設 $X_i\sim B(1,p,q)$, $i=1,\cdots,n$, 且 $\bar{X}=\frac{X_1+\cdots+X_n}{n}$, 則

$$\frac{\bar{X}-p}{\sqrt{\frac{pq}{n}}} \longrightarrow N(0,1)$$

我們用下面的例子來說明, 如何運用中央極限定理到我們丟銅板的老問題.

例. 在301頁, 我們使用大數法則說明在拋擲銅板 250000 次後, 我們有 $\frac{99}{100}$ 的信心說, $\bar{X}$ 與 p 的差距不超過 $\frac{1}{100}$.

現由中央極限定理, 我們有

$$\mathrm{P}\Big(-\tau\le\frac{\bar{X}-p}{\sqrt{\frac{pq}{n}}}\le\tau\Big)\approx 2\rho(\tau)$$

如果我們希望維持 $\frac{99}{100}$ 的信心, 則 $2\rho(\tau)=99\%$, 由習題5.6 知 τ 要取成 2.57, 所以

$$|\bar{X}-p|\le\tau\cdot\frac{\sqrt{pq}}{\sqrt{n}}\le 2.57\cdot\frac{1}{2}\cdot\frac{1}{\sqrt{n}}$$

所以如果我們希望 $\bar{X}$ 與 p 的差距不超過 $\frac{1}{100}$,

$$2.57\cdot\frac{1}{2}\cdot\frac{1}{\sqrt{n}}\le\frac{1}{100} \quad\Rightarrow\quad \sqrt{n}\ge 128.5 \quad\Rightarrow\quad n\ge 16512.25$$

只要 16513 次(大約一天)就夠了!! 這回應我們前面的看法, 由於柴比雪夫不等式太粗糙, 我們事實上不需要投擲 17天的銅板.

□

5.9 $\boxed{\text{習題}}$ (續上例) 若我們將 $|\bar{X}-p|$ 誤差的標準放鬆為 $\le\frac{1}{10}$ 就可以, 則要投擲銅板幾次? 假設將信心調鬆成 90%, 再作一次.

在科學研究中經常會碰到這樣的問題:

「從實驗資料結果, 我們怎麼判斷要不要接受某個理論假設?」

由於實驗有各種因素的誤差，所以我們經常要重複多次實驗以提題高判斷的正確性，又因為是重複試驗，所以就會使用到中央極限定理．底下是一個例子．

例．317頁提到 Rutherford 的 α 粒子放射試驗，如果從理論預測 $m = 3.9$，則從 Rutherford 所獲得的資料可以驗證這個假設嗎？設容忍度為 $\frac{5}{100}$（容忍度即假設 $m = 3.9$ 正確，但實驗資料卻與它不合的機率）．

在 Rutherford 的 α 粒子試驗中，設 X_i 表第 i 次觀測 α 粒子的放射數，Rutherford 總共作了 $n = 2608$ 次試驗，若令 $\bar{X} = \frac{X_1+\cdots+X_n}{n}$（即平均值），我們希望找到一個門檻 δ，使得

$$P(\ |\bar{X} - 3.9|\ ) \ \geq\ \delta\ ) \ =\ \frac{5}{100}$$

這表示如果 $m = 3.9$ 的假設真的正確，但是由於隨機因素，實驗結果的平均值 $\bar{X}$ 與 3.9 的差距卻大於 δ 的可能性只有 5%．由於出錯的可能性很小 (5%)，這麼一來如果實驗結果 $\bar{X}$ 真的滿足 $|\bar{X} - 3.9| \geq \delta$，那麼我們就相當有理由懷疑本來假設的正確性了．所以如果找到這麼一個判準值 δ，我們就可以利用實驗結果作如下的判斷：

$$\begin{cases} |\bar{X} - 3.9| \geq \delta & \Rightarrow \quad \text{應該拒斥假設「}m = 3.9\text{」} \\ |\bar{X} - 3.9| < \delta & \Rightarrow \quad \text{應該接受假設「}m = 3.9\text{」} \end{cases}$$

但一個非常現實的問題是，如果用 Poisson 過程描述 α 粒子的放射現像，機率 $P(|\bar{X} - 3.9| \geq \delta)$ 的計算非常複雜，這時就是中央極限定理上場的時候了．

由習題4.1知，當 $X_i \sim X \sim$ Poisson 分配 $\frac{m^k}{k!}e^{-m}$，則 $E(X) = Var(X) = m.$，從中央極限定理

$$\frac{\bar{X} - 3.9}{\sqrt{\frac{3.9}{2608}}} \quad \text{近似於} \quad N(0,1)$$

由習題 5.6 得

$$P\left(\left|\frac{\bar{X} - 3.9}{\sqrt{\frac{3.9}{2608}}}\right| \geq 1.96\right) \approx \frac{5}{100}$$

原式化簡為 $P(|\bar{X} - 3.9| \geq 0.0758) \approx \frac{5}{100}$，因此判準值 δ 可取為 0.0758．但由 Rutherford 的實驗數據知 $\bar{X} = 3.87$，因為 $|3.87 - 3.9| = 0.03 \leq 0.0758$，所以我們可以接受這個假設「$m = 3.9$」．

□

中央極限定理也可以用來估計機率分配中的參數，例如 $N(\mu, \sigma^2)$ 中的 μ, σ，或 Poisson 分配中的 m，指數分配中的 λ 等等．

例. (續前例) 在 Rutherford α 粒子實驗中, 試估計 m 可能的在區間(假設 90% 的信心).

雖然不可能由實驗結果去斷言 m 的真確值, 但我們希望能夠從實驗數據去估計 m 的可能範圍. 而所謂 90% 信心的意思是: m 在這個範圍的機率是 90%. 仿前例, 由中央極限定理

$$\mathrm{P}\Big(\Big|\frac{\bar{X}-m}{\sqrt{\frac{m}{2608}}}\Big| \leq 1.64\Big) = 90\%$$

代入實驗數據 $\bar{X} = 3.87$, 則

$$\begin{aligned} & |m - 3.87| \leq 1.64 \cdot \sqrt{\frac{m}{2608}} \\ \Rightarrow\ & (m-3.87)^2 \leq \frac{1.64^2}{2608}\, m \\ \Rightarrow\ & m^2 - 7.7410\, m + 14.9769 \leq 0 \end{aligned}$$

解不等式得, $3.807 \leq m \leq 3.934$. 因此 m 有 90% 的可能會落在區間 $(3.807, 3.934)$ 之間.

當然我們必須記得, 這些推估都是統計式的, 出錯的可能是永遠存在的,

□

5.10 習題 假設在去年某第一年間出生的 258341 個嬰孩中有 132561 名是男孩, 如果將容忍度訂為 $\frac{5}{100}$, 我們是否該接受男女孩出生的機率各為 $\frac{1}{2}$ 的假設呢?

5.11 習題 在訪談過 2500 人後, 支持興建核四共有 1150 人, 試利用中央極限定理以及 95% 信心, 估計反核比率 p 的可能區間.
(1) 使用不等式 $pq \leq \frac{1}{4}$ 去估計.
(2) 直接使用二次不等式估計.
(Ans. (1) $0.5204 < p < 0.5596$, (2) $0.5212 < p < 0.5603$.)

5.12 習題 在某學校單位中(例如數學系), 共有 60 台電話分機, 假設每部分機可能有 $\frac{5}{100}$ 的時間打外線電話 (當然彼此是獨立的), 問該單位應該裝幾條外線電話, 以保證有 $\frac{99}{100}$ 的機會, 不是所有外線皆佔線?

(Hint: $X_i \sim X = \begin{cases} 1 & \text{外線} \\ 0 & \text{內線或無} \end{cases}$ 求 n, 使得 $\mathrm{P}(\sum_{i=1}^{60} X_i < n) \geq \frac{99}{100}$)

5.13 習題 某工廠製造的燈泡壽命大約服從常態分配 $N(100, \sigma^2)$, σ 越大表示品質越參差不齊, 如果希望 $\mathrm{P}(80 \leq 壽命 \leq 120) \geq 0.85$, 則 σ 應控制在多少以下?

5.14 $\boxed{\text{習題}}$ 某超商經理紀錄一服務員之服務時間, 共記錄 100 次, 且平均服務時間為 1 分 15 秒, 若用指數分配 $\frac{1}{\mu}e^{-\frac{t}{\mu}}$, $t > 0$ 當作等候時間的模型, 請以 95% 的信心估計 μ 的可能區間? (Ans. 1 分 3 秒 $< \mu <$ 1 分 33 秒)

6 短結

在本章關於基本機率統計的介紹裡, 我們完全沒有碰觸到所謂相關性的問題, 這主要是我們一直局限在討論獨立事件的種種性質, 而獨立當然就表示這些事件不相關.

當然相關性的問題是很重要的, 例如氣體體積與壓力是相關的, 某些癌症或高血壓與遺傳相關, 人格成長與成長環境有關, 或者, 成就與智商有關嗎? 體重與身高有關嗎? 社會地位與相貌有關嗎?... 由於某種現象如果有某種規律來解釋, 則實驗資料必然應呈現相關性. 因此如何分析統計資料間的相關性往往是揭開深層律則的第一步. 但是光憑相關, 便斷言其中存在規律, 則是一種常犯的統計謬誤.

在討論相關性時, 聯合機率(密度)函數的分析有很大的重要性, 就像本章所述一樣, 先從機率理論出發, 再用統計方法來處理資料, 其中最重要的方法仍然不脫大數法則, 中央極限定理為源的各種樣本理論. 這主要還是因為收集樣本是獨立的過程. 就這點而言, 本章各節仍然是重要的理論基礎.