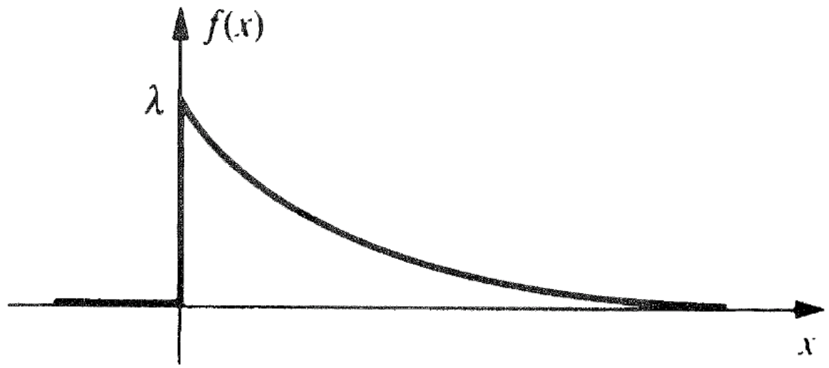
指數分佈

- 應用於間隔或等待時間的連續隨機分佈
  - 例：燈泡使用到壞掉的時間、顧客排隊等待的時間、機器發生故障的時間 ... 等
- 隨機變數為**非負數**值
- 通常該隨機變數超過其平均數的機率較小
- 當量測觀測事件在某一段時間裡的發生次數呈**卜瓦松**分佈，在此情況下探討兩次事件發生的間隔時間即可使用**指數分佈**

指數分佈 (Cont.)

- 若隨機變數 X 有一機率密度函數為
  - $f(x)=\lambda e^{-\lambda x}, x \geq 0$
  - 則 $E(X)=\frac{1}{\lambda}, V(X)=(\frac{1}{\lambda})^2$
- 則稱隨機變數 X 具有參數 λ 的指數分佈 (exponential distribution)
- 圖形



指數分佈 (Cont.)

- 應用：等待時間
  - 設微粒子的到達是獨立發生的
  - 令 λ 是**到達率**，亦即每分鐘到達的平均微粒子個數
  - 令 X 是等待下個微粒子到達所需的等待時間，亦即兩個微粒子抵達的間隔時間
  - 令 $F(x)$ 是 X 的分佈函數 (i.e. 先求分佈函數 $F(x)$ 再求密度函數 $f(x)$ 會比較簡單)
  - 因為**等待時間**不可能是**負數**，因此勢必 $X \geqq 0$，於是
    $F(x)=0 \text{ if } x\leq 0$

指數分佈 (Cont.)

- 針對 $x \geq 0$, 則 $F(x) = P(X \leq x) = 1 - P(X > x)$
- 若問：『等待下一個微粒子到達必須等待超過 x 分鐘』，則其所指的是在這等待的 x 分鐘裡沒有任何微粒子抵達，可用式子表示為
- $F(x) = 1 - P(\text{在等待的}\,x\,\text{分鐘裡沒有微粒子抵達})$
- 還記得嗎？討論『在 x 分鐘裡微粒子的抵達數』是具**卜瓦松分佈**的離散隨機變數，其中使用到的參數是在 x 分鐘裡微粒子的抵達數量，可標示為 λx
- 於是，分佈函數可表示為

$$F(x) = 1 - \frac{e^{-\lambda x}(\lambda x)^2}{0!} = 1 - e^{-\lambda x}, \text{ for } x \geq 0$$

指數分佈 (Cont.)

- Poisson distribution (卜瓦松分佈)
  - $P(n\text{次試驗有}\,k\,\text{次成功}) \approx \dfrac{e^{-\lambda}\lambda^{k}}{k!}$
- 接續上頁，密度函數$f(x)$可利用下列式子求得
- $f(x) = F'(x) = \lambda e^{-\lambda x}$ for $x \ge 0$
- 由此可知，在**等待時間**的問題上的是具**指數分佈**
- 這邊所提到的參數 λ 是指**到達率**(arrival rate)，同時它也可以被表示成**平均等待時間**(average waiting time)
- 因為到達率與平均等待時間互為倒數
- 例：若微粒子平均等待時間是2分鐘，則可知道每分鐘有1/2個微粒子抵達

指數分佈 (Cont.)

- 例子：假設某公司電話總機有電話進來的時事件是獨立的，且知道其為每小時平均有2通電話進來，問，至少等3個小時才會接到下一通電話的機率為何?
  - 令x是等待電話進來的時間
  - 由題目得知參數$\lambda=2$
  - 於是可以求得至少等待3個小時才會接到下一通電話的機率為
  - $P(X\geq 3)=\int_3^{\infty} 2e^{-2x}dx=-e^{-2x}\Big|_3^{\infty}=e^{-6}$

指數分佈 (Cont.)

- 應用：生命週期
  - 假設高腳杯破掉的情況發生是獨立的，不會受到高腳杯使用的年限因素所影響
  - 假設在某場合需要使用到高腳杯，一旦高腳杯破掉，則換上另一個高腳杯
  - 令X是高腳杯的生命週期，即高腳杯被用在該場合直到破掉的時間
  - 換個解釋：X是等待杯子破掉的時間
  - 於是，本問題可對應到等待時間的應用問題上，因此生命週期問題也是具指數分佈

$$\lambda = \frac{1}{\text{高腳杯生命週期平均}}$$

指數分佈的無記憶性

- 總機平均每小時會接到兩通撥入的電話，假設目前時間已是下午四點鐘，而上一通電話進來是下午的兩點三十分，問，到晚間七點前還是沒有電話進來的機率為何?
  - 從兩點三十分到下一通電話進來至少等了4 ½ 小時，而目前時間是4:00，表示已經等了1 ½ 小時，必須再等3小時才會有電話進來的機率
  - 丟擲銅板的試驗是獨立發生的，不管執行到第幾次丟擲，丟到正面的機率一樣是 ½
  - 相同概念，總機接到外面撥入的電話也是獨立的，不會彼此受影響

指數分佈的無記憶性 (Cont.)

- 不管你已經等了多久的時間，下一通電話撥入並不會隨著等的時間長短而提前後延後
- 所以 P(兩通撥入電話相隔超過$4\frac{1}{2}$|已經等待$1\frac{1}{2}$小時)

  $=P$(到少要再等3小時|已經等$1\frac{1}{2}$小時)

  $=P$(從現在開始必要再等≥ 3小時)

  $=\int_3^\infty 2e^{-2x}dx$　　等待時間呈指數分佈，且$\lambda=2$

  $=e^{-6}$
- 換言之，若X具有指數分佈，則

$$P\left(X \geq 4\frac{1}{2} \middle| X \geq 1\frac{1}{2}\right)=P\left(X \geq 3\right)$$

指數分佈的無記憶性 (Cont.)

- 更進一步而言 $P(X \geq x+t \mid X \geq t) = P(X \geq x)$
- 即，若到達的等待時間是指數分佈，已經等了t個小時後必須再等x小時的條件機率，是與一開始至少要等x小時的機率一樣
- 微粒子例子，微粒子的存活時間具有指數分佈，則微粒子至少再存活x小時的機率，將不管是新微粒子或舊微粒子都有同樣的機率
- 這樣的情況稱之為指數分佈的**無記憶性**
- 此外，指數分佈是*唯一*一種無記憶性的連續分佈

指數分佈的無記憶性 (Cont.)

- 從指數分佈的公式推導證明其具有無記憶性
- 首先，我們知道 $P(X \geq x) = \int_{x}^{\infty} \lambda e^{-\lambda x} dx = e^{-\lambda x}$
- 再者 $P(X \geq x+t \mid X \geq t) = \dfrac{P(X \geq x+t \text{ and } X \geq t)}{P(X \geq t)}$

$$= \frac{P(X \geq x+t)}{P(X \geq t)}$$

$$= \frac{\int_{x+t}^{\infty} \lambda e^{-\lambda x} dx}{\int_{t}^{\infty} \lambda e^{-\lambda x} dx}$$

$$= \frac{e^{-\lambda(x+t)}}{e^{-\lambda t}} = \frac{e^{-\lambda(x-t)}}{e^{-\lambda t}} = e^{-\lambda x - \lambda t + \lambda t}$$

$$= e^{-\lambda x}$$

指數分佈的無記憶性

令 $u=-\lambda x$ 則 $\dfrac{du}{dx}=-\lambda, du=-\lambda dx, dx=-\dfrac{1}{\lambda}du$

$$\begin{aligned}\int_t^{\infty}\lambda e^{-\lambda x}dx &= \int_t^{\infty} e^{u}(-\frac{1}{\lambda})du \\ &= \int_t^{\infty} e^{u}du \\ &= e^{u}\Big|_t^{\infty} \\ &= e^{-\lambda\infty}-e^{-\lambda t} \\ &= 0-e^{-\lambda t}\end{aligned}$$

卜瓦松 vs. 指數

- 設微粒子平均每分鐘有2顆到達，且微粒子的到達是獨立發生的，彼此不互相干擾
- 討論『在一分鐘內抵達的微粒子數量』是一種離散隨機變數，且其呈現具$\lambda=2$ 的卜瓦松分佈
- 換句話說，假設以7分鐘為觀測單位，則在7分鐘內的微粒子到達數量是具有$\lambda=14$ 的卜瓦松分佈
- 例子：問，在7分鐘內有10個微粒子抵達的機率為何?
- $P(\text{任一個7分鐘爲單位的時間裡有10微粒子抵達})=\dfrac{e^{14}14^{10}}{10!}$

卜瓦松 vs. 指數 (Cont.)

- 另外一方面，討論『等待下一個微粒子抵達的時間』是一種連續隨機分佈且其為具有參數 λ=2 的指數分佈
- 例如：

$$P(\text{下一個微粒子抵達需再等3到8分鐘}) = \int_3^8 \lambda e^{-\lambda x}\,dx = \int_3^8 2e^{-2x}\,dx$$

令 $u = -2x$, 則 $\frac{du}{dx} = -2, dx = \frac{-1}{2}du$

$$\begin{aligned}
\int_3^8 2e^{-2x}\,dx &= 2\int_3^8 e^u\left(-\frac{1}{2}du\right) \\
&= -\int_3^8 e^u\,du \\
&= -e^u\Big|_3^8 \\
&= -e^{-2x}\Big|_3^8 \\
&= -e^{-16} + e^{-6}
\end{aligned}$$

加瑪分佈 (Gamma Distribution)

- 設微粒子抵達是獨立事件，且每秒平均到達率為λ
- 令X是第n個微粒子抵達所須等待的時間，則X的密度為 $f(x)=\dfrac{\lambda^n e^{-\lambda x} x^{n-1}}{(n-1)!}$ for $x\geq 0$
- 這樣稱之為**加瑪分佈** (Gamma distribution)
- 加瑪分佈有兩個參數n與λ
- 當n=1時，便是**指數分佈**，因此可以說**指數分佈**是**加瑪分佈**的特殊情況

加瑪分佈　(Gamma Distribution) (Cont.)

- 加瑪分佈的推導
  - 首先 $F(x)=0$ for $x\leq 0$
  - 當 $x\geq 0$ 則

$$
\begin{aligned}
F(x) &= P(X\leq x)\\
&= 1-P(X\geq x)\\
&= 1-P(\text{等待第}\,n\,\text{個微粒子到達的時間是}\geq x)\\
&= 1-P(\text{在前面}\,x\,\text{秒抵達的微粒子數量是}\,0\ \text{or}\ 1\ \text{or}\ \ldots\ \text{or}\ n-1)\\
&= 1-[P(x\,\text{秒有}0\text{個微粒子抵達})+\ldots+P(x\,\text{秒有}\,n-1\,\text{個微粒子抵達})]
\end{aligned}
$$

加瑪分佈 (Gamma Distribution) (Cont.)

- 在x秒內到達的微粒子數是具有參數 λx 的卜瓦松分佈，因此

$$F(x)=1-e^{-\lambda x}[1+\lambda x+\frac{(\lambda x)^2}{2!}+\ldots+\frac{(\lambda x)^{n-1}}{(n-1)!}]$$

- 所以，使用乘法原則可得

$$f(x)=F'(x)=-e^{-\lambda x}[\lambda+\frac{2\lambda^2 x}{2!}+\frac{3\lambda^3 x^2}{3!}+\ldots+\frac{(n-1)\lambda^{n-1}x^{n-2}}{(n-1)!}]$$
$$+\lambda e^{-\lambda x}[1+\lambda x+\frac{(\lambda x)^2}{2!}+\ldots+\frac{(\lambda x)^{n-1}}{(n-1)!}]$$