

第二章

常見分佈

2.3 連續型分佈

2.3.1 均勻分佈

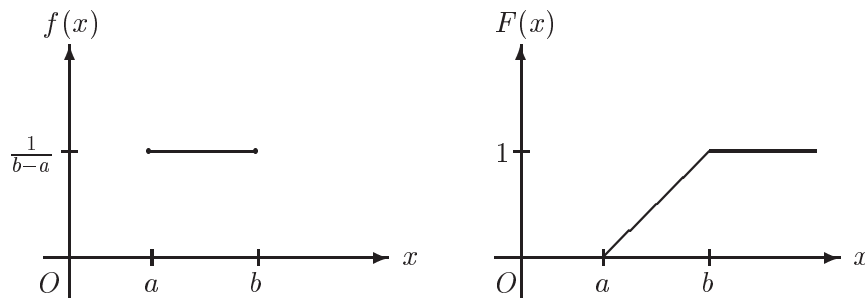
連續型在區間 $[a, b]$ 之**均勻分佈**，以 $\mathcal{U}[a, b]$ 表之。設隨機變數 X 有 $\mathcal{U}[a, b]$ 分佈，其中 $a < b$ ，則 X 之p.d.f.為

$$(3.1) \quad f(x) = \frac{1}{b-a}, x \in [a, b].$$

分佈函數則為

$$F(x) = \begin{cases} 0 & , x < a, \\ (x-a)/(b-a) & , a \leq x \leq b, \\ 1 & , x \geq b, \end{cases}$$

在區間 $[a, b]$ 中為線性函數。圖3.1給出 $\mathcal{U}[a, b]$ 分佈之p.d.f. $f(x)$ ，及分佈函數 $F(x)$ 之圖形。

圖3.1 $\mathcal{U}[a, b]$ 分佈之p.d.f. $f(x)$ 及分佈函數 $F(x)$ 之圖形

又

$$E(X) = \int_a^b \frac{x}{b-a} dx = \frac{b+a}{2},$$

$$\begin{aligned} \text{Var}(X) &= \int_a^b \frac{x^2}{b-a} dx - \left(\frac{b+a}{2}\right)^2 \\ &= \frac{1}{3} \frac{b^3 - a^3}{b-a} - \frac{(b+a)^2}{4} \\ &= \frac{(b-a)^2}{12}. \end{aligned}$$

$$\phi(s) = E(e^{-sX}) = \int_a^b e^{-sx} \frac{1}{b-a} dx = \begin{cases} \frac{e^{-sa} - e^{-sb}}{s(b-a)}, & s > 0, \\ 1, & s = 0. \end{cases}$$

通常自一有限的區間 $[a, b]$ 隨機地取一點，所得的點 X 便有 $\mathcal{U}[a, b]$ 分佈。有時區間可以是 (a, b) ， $[a, b)$ 或 $(a, b]$ ，分別對應 $\mathcal{U}(a, b)$ 分佈， $\mathcal{U}[a, b)$ 分佈及 $\mathcal{U}(a, b]$ 分佈，但不可以是無限的區間。

對一實數上的集合 A ，令

$$I_A(x) = \begin{cases} 1, & \text{若 } x \in A, \\ 0, & \text{若 } x \notin A, \end{cases}$$

則 I_A 稱為 A 之**指示函數** (在第一章例3.4，我們曾定義事件的指示函數)。若採用此符號， $\mathcal{U}[a, b]$ 之p.d.f. 可寫成 $\frac{1}{b-a} I_{[a, b]}(x)$ ，而不用在逗號後附記 $x \in [a, b]$ 。

指示函數的符號尚有不同的寫法, 如

$$I_A(x) = I_{\{x \in A\}}.$$

仍表 $x \in A$ 時, $I_{\{x \in A\}} = 1$, 否則為 0. 若採用此符號, $\mathcal{U}[a, b]$ 之 p.d.f. 便可寫成 $I_{\{x \in [a, b]\}}$, 或 $I_{\{a \leq x \leq b\}}$.

2.3.2 Gamma 分佈

隨機變數 X , 若其 p.d.f. 為

$$(3.2) \quad f(x) = \frac{x^{\alpha-1} e^{-x/\beta}}{\Gamma(\alpha) \beta^\alpha}, x > 0,$$

其中 $\alpha, \beta > 0$ 為二正數, 便稱為有參數 α, β 之 **gamma 分佈** (gamma distribution), 以 $\Gamma(\alpha, \beta)$ 表之, 在此 **gamma 函數** (gamma function) 之定義為

$$(3.3) \quad \Gamma(\alpha) = \int_0^\infty t^{\alpha-1} e^{-t} dt, \alpha > 0.$$

在微積分裡, 我們曾證明 (3.3) 式中之積分, 在 $\alpha > 0$ 處收斂, gamma 函數之定義域即取為 $(0, \infty)$. Gamma 函數為應用數學中一重要的函數, 它有不少性質, 例如,

$$(3.4) \quad \Gamma(\alpha + 1) = \alpha \Gamma(\alpha), \alpha > 0,$$

$$(3.5) \quad \Gamma(1) = 1,$$

$$(3.6) \quad \Gamma(n) = (n-1)!, n \text{ 為正整數}.$$

如何驗證 (3.2) 式的確定義出一 p.d.f. 呢? 首先由 (3.3) 式, 立即得知

$$(3.7) \quad g(t) = \frac{t^{\alpha-1} e^{-t}}{\Gamma(\alpha)}, t > 0,$$

對每一 $\alpha > 0$ 皆為一 p.d.f. 經由變數代換, 假設隨機變數 T , 具有如 (3.7) 式之 p.d.f., 令 $X = \beta T$, 其中 $\beta > 0$, 則得 X 之 p.d.f. 如 (3.2) 式所給.

由 $\int_0^\infty f(x) dx = 1$, 我們亦得下述公式:

$$(3.8) \quad \int_0^\infty x^{\alpha-1} e^{-x/\beta} dx = \Gamma(\alpha) \beta^\alpha, \alpha, \beta > 0.$$

4 第二章 常見分佈

Gamma分佈中的參數 α , 稱為**形狀參數**(shape parameter), β 稱為**尺度參數**(scale parameter). 此因 α 影響p.d.f.圖形之陡峭程度, 而 β 影響散佈程度. 圖3.2給出一些gamma分佈之圖形.

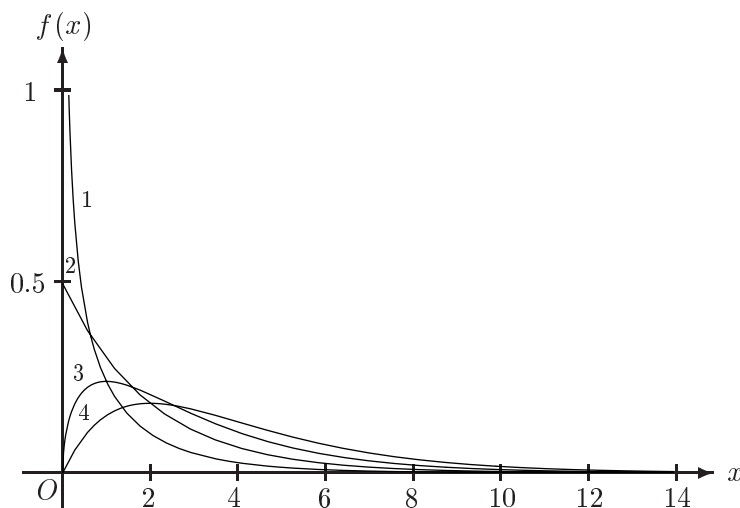


圖3.2 Gamma 分佈p.d.f.之圖形, 1,2,3,4分別對應
 $\alpha = 0.5, 1, 1.5, 2$, β 則皆為2

$$\begin{aligned}
 \text{Gamma分佈之動差極易求出對}\forall k \geq 1, \\
 E(X^k) &= \int_0^{\infty} x^k f(x) dx \\
 &= \int_0^{\infty} x^k \frac{x^{\alpha-1} e^{-x/\beta}}{\Gamma(\alpha) \beta^{\alpha}} dx \\
 &= \int_0^{\infty} \frac{\beta^k y^{\alpha+k-1} e^{-y}}{\Gamma(\alpha)} dy \\
 &= \frac{\beta^k}{\Gamma(\alpha)} \int_0^{\infty} y^{\alpha+k-1} e^{-y} dy \\
 &= \frac{\beta^k \Gamma(\alpha + k)}{\Gamma(\alpha)},
 \end{aligned}$$

此處用到積分的變數代換, 令 $y = x/\beta$, 以及因 $\alpha + k > 0$, 故由(3.3)式得

$$\int_0^{\infty} y^{\alpha+k-1} e^{-y} dy = \Gamma(\alpha + k).$$

附帶一提, 在上述推導中, k 其實不須為整數, 可看出只要實數 $k > -\alpha$, 便

有

$$(3.9) \quad E(X^k) = \frac{\beta^k \Gamma(\alpha + k)}{\Gamma(\alpha)} = (\alpha + k - 1) \cdots (\alpha + 1) \alpha \beta^k.$$

由(3.9)式得

$$E(X) = \alpha\beta, \quad E(X^2) = \alpha(\alpha + 1)\beta^2,$$

因此

$$\text{Var}(X) = \alpha\beta^2.$$

又

$$\begin{aligned} \phi(s) &= E(e^{-sX}) = \int_0^\infty e^{-sx} f(x) dx \\ &= \int_0^\infty \frac{1}{\Gamma(\alpha)\beta^\alpha} e^{-sx} x^{\alpha-1} e^{-x/\beta} dx \\ &= \frac{1}{\beta^\alpha} \int_0^\infty \frac{1}{\Gamma(\alpha)} x^{\alpha-1} e^{-(s+1/\beta)x} dx \\ &= \frac{1}{\beta^\alpha} (s + 1/\beta)^{-\alpha} = \frac{1}{(1 + \beta s)^\alpha}, s \geq 0. \end{aligned}$$

註3.1 有些書的gamma分佈, 其p.d.f.寫成

$$(3.10) \quad h(x) = \frac{\beta^\alpha x^{\alpha-1} e^{-\beta x}}{\Gamma(\alpha)}, x > 0.$$

由於 α, β 只是參數, 本質上此二形式並無差異, 只是此時期望值及變異數分別要改成 α/β , 及 α/β^2 , 而拉普拉斯轉換成爲

$$\frac{\beta^\alpha}{(\beta + s)^\alpha}, s \geq 0.$$

Gamma分佈與波松分佈之間, 有一有趣的關係, 見下例.

例3.1 設 X 有 $\Gamma(\alpha, \beta)$ 分佈, 且 α 爲一正整數. 則

$$(3.11) \quad P(X \leq x) = P(Y \geq \alpha), x > 0,$$

其中 Y 有 $\mathcal{P}(x/\beta)$ 分佈, 上式之另一表示法為

$$(3.12) \quad F_X(x) = 1 - F_Y(\alpha - 1).$$

(3.11)式可反覆地利用**分部積分**(integration by parts)而得: 首先因 α 為正整數, $\Gamma(\alpha) = (\alpha - 1)!$, 因此

$$\begin{aligned} P(X \leq x) &= \frac{1}{(\alpha - 1)!\beta^\alpha} \int_0^x t^{\alpha-1} e^{-t/\beta} dt \\ &= \frac{1}{(\alpha - 1)!\beta^\alpha} (-t^{\alpha-1} \beta e^{-t/\beta} \Big|_0^x + \int_0^x (\alpha - 1) t^{\alpha-2} \beta e^{-t/\beta} dt) \\ &= -\frac{1}{(\alpha - 1)!\beta^{\alpha-1}} x^{\alpha-1} e^{-x/\beta} + \frac{1}{(\alpha - 2)!\beta^{\alpha-1}} \int_0^x t^{\alpha-2} e^{-t/\beta} dt \\ &= \frac{1}{(\alpha - 2)!\beta^{\alpha-1}} \int_0^x t^{\alpha-2} e^{-t/\beta} dt - P(Y = \alpha - 1). \end{aligned}$$

此處用到分部積分, 取 $u = t^{\alpha-1}$, $v = e^{-t/\beta}$, 又因 Y 有 $\mathcal{P}(x/\beta)$ 分佈, 故

$$P(Y = \alpha - 1) = \frac{e^{-x/\beta} (x/\beta)^{\alpha-1}}{(\alpha - 1)!} = \frac{x^{\alpha-1} e^{-x/\beta}}{(\alpha - 1)!\beta^{\alpha-1}}.$$

在上述推導中, 我們其實假設 $\alpha \geq 2$. 若 $\alpha = 1$, 可如下看出(3.11)式成立:

$$\begin{aligned} P(X \leq x) &= \frac{1}{\beta} \int_0^x e^{-t/\beta} dt \\ &= -e^{-t/\beta} \Big|_0^x = 1 - e^{-x/\beta} \\ &= 1 - P(Y = 0) = P(Y \geq 1). \end{aligned}$$

現對 $\int_0^x t^{\alpha-2} e^{-t/\beta} dt$ 繼續利用分部積分(只要 $\alpha - 2 \geq 1$), 最後可得

$$\begin{aligned} P(X \leq x) &= \frac{1}{\beta} \int_0^x e^{-t/\beta} dt \\ &\quad - (P(Y = \alpha - 1) + P(Y = \alpha - 2) + \cdots + P(Y = 1)) \\ &= 1 - e^{-x/\beta} - \sum_{i=1}^{\alpha-1} P(Y = i) \\ &= 1 - \sum_{i=0}^{\alpha-1} P(Y = i) = P(Y \geq \alpha). \end{aligned}$$

得證(3.11)式.

由於與波松分佈有如(3.11)式的關係, 使得gamma分佈在**等待理論**(queueing theory), 及**可靠度理論**(reliability theory)裡非常重要, 細節我們在此不多談.

Gamma 分佈有二特例也很重要.

2.3.3 卡方分佈

如果取 $\alpha = p/2$, 其中 p 為一正整數, $\beta = 2$, 便得到**卡方分佈**(chi-squared distribution). 此分佈只有一參數 p , 稱為其**自由度**(degrees of freedom), 而p.d.f.為

$$(3.13) \quad f(x) = \frac{1}{\Gamma(p/2)2^{p/2}} x^{p/2-1} e^{-x/2}, x > 0.$$

此分佈以 χ_p^2 表示. 圖3.2中之gamma分佈p.d.f.之圖形, 由於對應 $\beta = 2$, 故亦皆為卡方分佈p.d.f.之圖形, 分別為 χ_1^2 , χ_2^2 , χ_3^2 , 及 χ_4^2 .

當 X 有 χ_p^2 分佈, 利用關於gamma分佈的結果, 可得

$$\begin{aligned} E(X) &= p, \\ \text{Var}(X) &= 2p, \\ E(e^{-sX}) &= \frac{1}{(1+2s)^{p/2}}, s \geq 0, \end{aligned}$$

及

$$(3.14) \quad E(X^r) = \frac{2^r \Gamma(p/2 + r)}{\Gamma(p/2)}, r > -p/2,$$

其中 r 並不須為整數.

例3.2 設 X 有 $\Gamma(\alpha, \beta)$ 分佈. 則對 $\forall a > 0$, $Y = aX$ 有 $\Gamma(\alpha, a\beta)$ 分佈. 此由拉普拉斯轉換可唯一決定分佈即可得知:

$$E(e^{-sY}) = E(e^{-(sa)X}) = \frac{1}{(1+a\beta s)^\alpha}, s \geq 0.$$

因此若 U 有 $\Gamma(k, \beta)$ 分佈, 其中 k 為一正整數, 則 $(2/\beta)U$ 有 χ^2_{2k} 分佈.

卡方分佈在**統計推論**(statistical inference)中扮演著重要的角色, 我們稍後會說明.

2.3.4 指數分佈

在 $\Gamma(\alpha, \beta)$ 分佈裡, 若取 $\alpha = 1$, 就會得到gamma分佈之另一特例: **指數分佈**. 不過為了簡便, 通常再取 $\beta = 1/\lambda$, 而得p.d.f.

$$(3.15) \quad f(x) = \lambda e^{-\lambda x}, x > 0.$$

此為參數 λ 之**指數分佈**, 以 $\mathcal{E}(\lambda)$ 表之, 我們在1.5節曾引進. 可以符號 $\Gamma(1, 1/\lambda) = \mathcal{E}(\lambda)$ 表指數分佈與gamma分佈之關係. 又其p.d.f.之圖形如圖3.3.

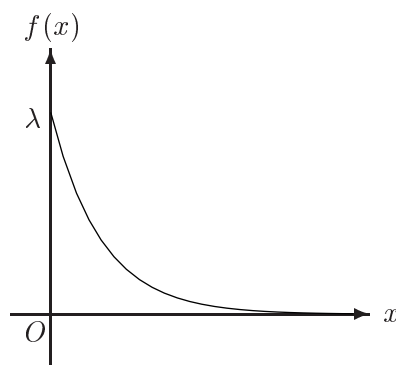


圖3.3 $\mathcal{E}(\lambda)$ 分佈p.d.f.之圖形

指數分佈之分佈函數形式簡單:

$$F(x) = 1 - e^{-\lambda x}, x > 0.$$

又由gamma分佈之結果可得

$$E(X) = \frac{1}{\lambda},$$

$$\text{Var}(X) = \frac{1}{\lambda^2},$$

$$\phi(s) = E(e^{-sX}) = \frac{\lambda}{\lambda + s}, s \geq 0.$$

由 $\phi(s)$ 立即得

$$E(X^n) = \frac{n!}{\lambda^n}, n \geq 1.$$

指數分佈亦有無記憶性質:

$$\begin{aligned} P(X > s+t | X > s) &= \frac{P(X > s+t)}{P(X > s)} = \frac{e^{-\lambda(s+t)}}{e^{-\lambda s}} \\ &= e^{-\lambda t} = P(X > t), s, t > 0. \end{aligned}$$

由於同有無記憶性質, 指數分佈與幾何分佈, 常有許多平行的結果同時成立. 指數分佈也常被用來當做壽命的模式. 在某些條件下, 連續的隨機變數若有無記憶性質, 可證明必有指數分佈. 諸如間隔多久會有某特定的意外發生, 知更鳥的壽命等, 往往有無記憶性質, 因此以指數分佈來當做機率模式, 便很恰當.

有些書對 $\mathcal{E}(\lambda)$ 分佈之p.d.f.定義為 $f(x) = \lambda^{-1}e^{-x/\lambda}$, $x > 0$. 若採用此p.d.f., 則 $\mathcal{E}(\lambda) = \Gamma(1, \lambda)$, 且 $E(X) = \lambda$, $\text{Var}(X) = \lambda^2$, 形式較簡潔.

2.3.5 韋伯分佈

韋伯分佈(Weibull distribution)以指數分佈為一特例, 其p.d.f.為

$$(3.16) \quad f(x) = \alpha\beta x^{\beta-1}e^{-\alpha x^\beta}, x > 0,$$

其中 $\alpha, \beta > 0$. 以 $\mathcal{W}(\alpha, \beta)$ 表此分佈, 有二參數 α, β , α 為尺度參數, β 為形狀參數. 有些書 $\mathcal{W}(\alpha, \beta)$ 分佈之p.d.f., 是將(3.16)式中的 α , 以 α^β 取代. 若取 $\beta = 1$, 則得 $\mathcal{E}(\alpha)$ 分佈, 以 $\mathcal{W}(\alpha, 1) = \mathcal{E}(\alpha)$ 表之. $\mathcal{W}(\alpha, 2)$ 分佈即參數 α 之瑞萊分佈. 又若令 $Y = X^\beta$, 則 Y 有 $\mathcal{E}(\alpha)$ 分佈(習題第6題). 韋伯分佈族與gamma分佈族的交集, 就是指數分佈族. 圖3.4給出一些韋伯分佈p.d.f.之圖形.

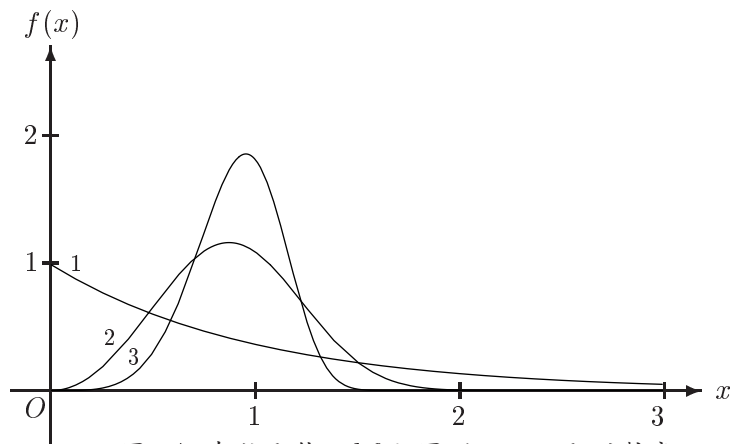


圖3.4 韋伯分佈p.d.f.之圖形, 1,2,3,分別對應
 $\beta = 1, 3, 5$, α 則皆為1.

韋伯分佈是瑞典物理學家Waloddi Weibull, 為發展強化材料的理論, 於西元1939年所引進, 是一較“新”的分佈. 在可靠度理論及有關壽命檢定問題裡, 常少不了韋伯分佈的影子.

$\mathcal{W}(\alpha, \beta)$ 分佈的分佈函數為

$$F(x) = 1 - e^{-\alpha x^\beta}, x > 0.$$

期望值及變異數則可利用(3.9)式而得: 當 X 有 $\mathcal{W}(\alpha, \beta)$ 分佈, 我們已指出 $Y = X^\beta$ 有 $\mathcal{E}(\alpha) = \Gamma(1, 1/\alpha)$ 分佈. 因此(在(3.9)式中令 $\alpha = 1, \beta = 1/\alpha, k = 1/\beta$)

$$E(X) = E(Y^{1/\beta}) = \frac{(1/\alpha)^{1/\beta} \Gamma(1 + 1/\beta)}{\Gamma(1)} = \alpha^{-1/\beta} \Gamma(1 + \frac{1}{\beta}).$$

又(在(3.9)式中令 $\alpha = 1, \beta = 1/\alpha, k = 2/\beta$)

$$E(X^2) = E(Y^{2/\beta}) = \frac{(1/\alpha)^{2/\beta} \Gamma(1 + 2/\beta)}{\Gamma(1)} = \alpha^{-2/\beta} \Gamma(1 + \frac{2}{\beta}).$$

故

$$\text{Var}(X) = \alpha^{-2/\beta} (\Gamma(1 + \frac{2}{\beta}) - (\Gamma(1 + \frac{1}{\beta}))^2).$$

拉普拉斯轉換則不易求出.

2.3.6 常態分佈

常態分佈(normal distribution)又稱**高斯分佈**(Gaussian distribution)德國的10馬克紙幣,以**高斯**(Gauss, 1777-1855)為人像,人像左側有一常態分佈之p.d.f.及其圖形.高斯在數學上有諸多貢獻,但在10馬克的紙幣,挑出來與他相隨的,是常態分佈.可見常態分佈不只在統計上,在數學上亦很重要.不過高斯倒不是第一位提出此分佈者.法國數學家**棣美弗**(Abraham De Moivre, 1667-1754)早於他寫出此分佈.甚至一般認為**丹尼爾·伯努力**(Daniel Bernoulli, 1700-1782, 為2.2.1節之James Bernoulli的姪兒)更早就發現了.有人稱這種現象為**誤稱定律**(Law of Misnomer).要知數學上的命名,往往並非以實際發現者.

常態分佈之所以重要,原因很多,我們給出三個主要的原因:首先是常態分佈在分析上較易處理(不過乍看之下,可能不易體會).其次是常態分佈之p.d.f.的圖形為**鐘形曲線**(bell-shaped curve),再加上對稱性,使得很適合當做不少母體之機率模式.當然底下我們會看到鐘形且具對稱的分佈也有不少,但通常不像常態分佈,在分析上如此容易駕馭.第三個原因是由於下一章要介紹的**中央極限定理**(Central Limit Theorem),使得在不太強的條件下,常態分佈可當做不少大樣本的近似分佈.

常態分佈有二參數 μ 及 σ^2 , $\mu \in R$, $\sigma > 0$, 以 $\mathcal{N}(\mu, \sigma^2)$ 表此分佈, p.d.f.為

$$(3.17) \quad f(x) = \frac{1}{\sqrt{2\pi}\sigma} e^{-(x-\mu)^2/(2\sigma^2)}, x \in R.$$

若 X 有 $\mathcal{N}(\mu, \sigma^2)$ 分佈, 則

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

有 $\mathcal{N}(0, 1)$ 分佈, 稱為**標準常態**(standard normal)分佈.其p.d.f.為

$$(3.18) \quad f(x) = \frac{1}{\sqrt{2\pi}} e^{-x^2/2}, x \in R.$$

這只要利用變數代換便立即可看出. μ 稱為**位置參數**(location parameter), 此因 μ 影響圖形的位置, σ 則為**尺度參數**.由 X 至 Z 的一變換稱為

將 X 標準化, $\mathcal{N}(\mu, \sigma^2)$ 分佈經標準化後, 便成為 $\mathcal{N}(0, 1)$ 分佈(習題第11題). 由於有此性質, 雖有二參數, 常態分佈之機率值皆可藉由標準常態分佈之機率值(見附表4)而得. 通常令

$$(3.19) \quad \Phi(x) = \int_{-\infty}^x \frac{1}{\sqrt{2\pi}} e^{-u^2/2} du, x \in R,$$

表標準常態分佈之分佈函數. 則對任意 $a < b$,

$$\begin{aligned} (3.20) \quad & P(a < X \leq b) \\ &= P\left(\frac{a - \mu}{\sigma} < \frac{X - \mu}{\sigma} \leq \frac{b - \mu}{\sigma}\right) \\ &= P\left(Z \leq \frac{b - \mu}{\sigma}\right) - P\left(Z \leq \frac{a - \mu}{\sigma}\right) \\ &= \Phi\left(\frac{b - \mu}{\sigma}\right) - \Phi\left(\frac{a - \mu}{\sigma}\right). \end{aligned}$$

又因 Z 之p.d.f.為偶函數, 故

$$(3.21) \quad \Phi(-x) = 1 - \Phi(x), x > 0.$$

利用附表4, 可得常態分佈的機率值. 在此設 Z 有 $\mathcal{N}(0, 1)$ 分佈, 對 $\forall 0 < p < 1$, 令 z_p 滿足

$$(3.22) \quad \Phi(z_p) = P(Z \leq z_p) = p.$$

可看出 $z_p = \Phi^{-1}(p)$, 為標準常態分佈函數的第 p 分位數(見1.5節).

例3.3 由附表4得 $\Phi(1.12) \doteq 0.8686$, $\Phi(2.58) \doteq 0.9951$. 又若 $\Phi(x) \doteq 0.9750$, 則 $x \doteq 1.96$.

另外, $\Phi(-1.35) = 1 - \Phi(1.35) \doteq 1 - 0.9115 = 0.0885$.

給定一 p 值, $0 < p < 1$, 也可由附表4找出 z_p , 使得 $\Phi(z_p) = p$. 例如, 若 $p = 0.9850$, 則 $z_p \doteq 2.17$. 若 $p = 0.3409$, 因 $p < 1/2$, 取 $1 - p = 0.6591$, 由附表4, 得 $z_{1-p} \doteq 0.41$, 故 $z_p \doteq -0.41$.

即使對附表4中沒有的 p , 也可藉由內插法, 而求出 z_p 之近似值. 例如, 取 $p = 0.95$, 則因 $\Phi(1.64) \doteq 0.9495$, $\Phi(1.65) = 0.9505$, 且

$$\frac{0.95 - 0.9495}{0.9505 - 0.9495} = \frac{0.0005}{0.001} = 0.5,$$

$$0.5 \cdot (1.65 - 1.64) = 0.005,$$

$$1.64 + 0.005 = 1.645,$$

故得 $\Phi(1.645) \doteq 0.95$.

對一般的常態分佈, 則可先利用(3.20)式, 轉換成標準常態分佈而得其機率值. 例如, 設 X 有 $\mathcal{N}(1, 9)$ 分佈, 欲求 $P(-2 < X < 7.3)$. 首先

$$P(-2 < X < 7.3) = P\left(\frac{-2-1}{3} < \frac{X-1}{3} < \frac{7.3-1}{3}\right) = \Phi(2.1) - \Phi(-1).$$

又 $\Phi(2.1) \doteq 0.9821$, $\Phi(-1) = 1 - \Phi(1) \doteq 1 - 0.8413 = 0.1587$. 故得

$$P(-2 < X < 7.3) \doteq 0.9821 - 0.1587 = 0.8234.$$

最後利用(3.21)式, 對 Z 有 $\mathcal{N}(0, 1)$ 分佈, 立即可得

$$(3.23) \quad P(|Z| \leq c) = 2\Phi(c) - 1, c \geq 0.$$

X 之期望值及變異數, 也可藉由 Z 之期望值及變異數而得:

$$\begin{aligned} E(Z) &= \int_{-\infty}^{\infty} \frac{1}{\sqrt{2\pi}} z e^{-z^2/2} dz \\ &= -\frac{1}{\sqrt{2\pi}} e^{-z^2/2} \Big|_{-\infty}^{\infty} \\ &= 0, \\ E(Z^2) &= \int_{-\infty}^{\infty} \frac{1}{\sqrt{2\pi}} z^2 e^{-z^2/2} dz \\ &= -\frac{1}{\sqrt{2\pi}} \int_{-\infty}^{\infty} z d e^{-z^2/2} \\ &= -\frac{1}{\sqrt{2\pi}} z e^{-z^2/2} \Big|_{-\infty}^{\infty} + \int_{-\infty}^{\infty} \frac{1}{\sqrt{2\pi}} e^{-z^2/2} dz \\ &= 1. \end{aligned}$$

故

$$\text{Var}(Z) = E(Z^2) - (E(Z))^2 = 1.$$

即標準常態分佈之期望值為0, 變異數為1. 由此即得

$$\begin{aligned} E(X) &= E(\mu + \sigma Z) = \mu + \sigma E(Z) = \mu, \\ \text{Var}(X) &= \text{Var}(\mu + \sigma Z) = \sigma^2 \text{Var}(Z) = \sigma^2. \end{aligned}$$

故 X 之期望值為 μ , 變異數為 σ^2 . 也就是在 $\mathcal{N}(\mu, \sigma^2)$ 分佈裡, 二參數 μ, σ^2 , 分別為 X 之期望值及變異數.

我們給出標準常態之圖形如圖3.5.

常態分佈之特徵函數存在, 我們列出如下, 推導過程則略去:

$$(3.24) \quad \psi(t) = E(e^{itX}) = e^{i\mu t - \sigma^2 t^2/2}, t \in R.$$

動差母函數亦存在:

$$(3.25) \quad M(t) = E(e^{tX}) = e^{\mu t + \sigma^2 t^2/2}, t \in R.$$

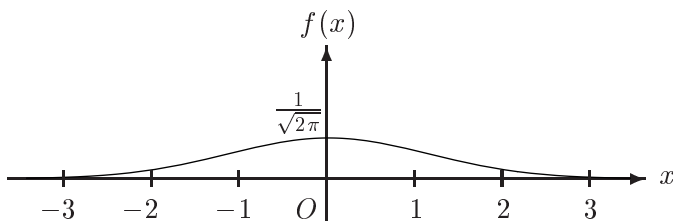


圖3.5 $\mathcal{N}(0, 1)$ 分佈p.d.f.之圖形

這只要利用型如(3.17)式所給之函數, 在 $(-\infty, \infty)$ 的積分等於1便可求得, 我們留給各位當習題.

利用(3.25)式, 可得到 $\mathcal{N}(\mu, \sigma^2)$ 分佈之所有動差. 對於標準常態分佈, 其所有動差皆有簡潔的公式: 設 Z 有 $\mathcal{N}(0, 1)$ 分佈, 利用分部積分可得

$$(3.26) \quad E(Z^k) = (k-1)E(Z^{k-2}), k \geq 2.$$

又因 $E(Z) = 0$, 且 $E(Z^0) = 1$, 故

$$(3.27) \quad E(Z^{2k}) = \prod_{i=1}^k (2i-1), k \geq 1,$$

且

$$(3.28) \quad E(Z^{2k+1}) = 0, k \geq 0.$$

大家應注意到, 我們其實尚未檢驗(3.17)式所定義的 f 的確為一p.d.f. 由於可利用標準化的變換, 我們只要檢驗

$$\frac{1}{\sqrt{2\pi}} \int_{-\infty}^{\infty} e^{-z^2/2} dz = 1$$

即可. 而因 $g(z) = e^{-z^2/2}$, $z \in R$, 為一偶函數, 故只要證出

$$(3.29) \quad \int_0^{\infty} e^{-z^2/2} dz = \frac{\sqrt{2\pi}}{2} = \sqrt{\frac{\pi}{2}}$$

即可. 雖然 $e^{-z^2/2}$ 之反導數(antiderivative)並非初等函數(elementary function), 但(3.29)式中的定積分之值卻可求出, 只是要用到重積分:

$$\begin{aligned} \left(\int_0^{\infty} e^{-z^2/2} dz \right)^2 &= \int_0^{\infty} e^{-x^2/2} dx \int_0^{\infty} e^{-y^2/2} dy \\ &= \int_0^{\infty} \int_0^{\infty} e^{-(x^2+y^2)/2} dx dy \\ &= \int_0^{\infty} \int_0^{\pi/2} r e^{-r^2/2} d\theta dr \\ &= \frac{\pi}{2} \int_0^{\infty} r e^{-r^2/2} dr \\ &= \frac{\pi}{2} \left(-e^{-r^2/2} \Big|_0^{\infty} \right) = \frac{\pi}{2}, \end{aligned}$$

此處用到極坐標變換, 令 $x = r \cos \theta$, $y = r \sin \theta$, $0 < \theta < \pi/2$, $0 < r < \infty$. 因此(3.29)式成立. 附帶一提, (3.29)式等價於

$$(3.30) \quad \int_0^{\infty} e^{-z^2} dz = \frac{\sqrt{\pi}}{2}.$$

(3.30)式之積分與gamma函數關係密切. 令 $w = z^2$, 則(3.30)式成為

$$\int_0^{\infty} \frac{1}{2} w^{-1/2} e^{-w} dw = \frac{\sqrt{\pi}}{2},$$

即得

$$(3.31) \quad \Gamma\left(\frac{1}{2}\right) = \int_0^{\infty} w^{-1/2} e^{-w} dw = \sqrt{\pi}.$$

由此可得(利用(3.4)式) $\Gamma((2k+1)/2)$ 之值, $k = 0, 1, 2, \dots$.

對一 $\mathcal{N}(\mu, \sigma^2)$ 分佈, 二參數 μ, σ^2 分別提供此分佈之位置及形狀的資訊:
 μ 不但為此分佈之期望值、中位數及眾數, 且p.d.f.之圖形對稱於 $x = \mu$, 反曲點發生在 $x = \mu + \sigma$ 及 $x = \mu - \sigma$. 而 σ^2 為變異數, σ 愈大表圖形散的愈開. 要知p.d.f.之極大值發生在 $x = \mu$, 且極大值為 $1/(\sigma\sqrt{2\pi})$. 當 σ 愈小, 極大值便愈大, 因此分佈愈集中在 $x = \mu$ 附近; 當 σ 愈大, 極大值便愈小, 圖形也愈平緩. 圖3.6給三個有相同的 μ , 但 σ^2 不同之p.d.f.的圖形, 以做比較.

當 X 有 $\mathcal{N}(\mu, \sigma^2)$ 分佈, X 會在 μ 的1個, 2個及3個標準差內之機率, 分別約為

$$P(|X - \mu| \leq \sigma) = P(|Z| \leq 1) \doteq 0.6826,$$

$$P(|X - \mu| \leq 2\sigma) = P(|Z| \leq 2) \doteq 0.9544,$$

$$P(|X - \mu| \leq 3\sigma) = P(|Z| \leq 3) \doteq 0.9974,$$

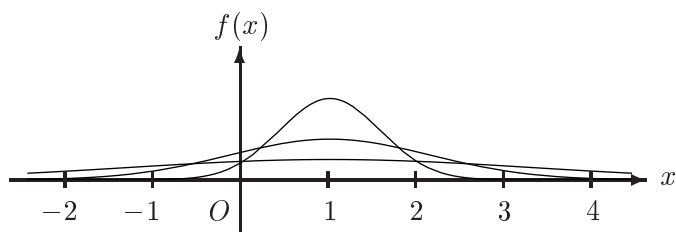


圖3.6 $\mathcal{N}(1, 2)$ 、 $\mathcal{N}(1, 1)$ 及 $\mathcal{N}(1, 1/2)$ 分佈p.d.f.圖形之比較

其中 Z 有 $\mathcal{N}(0, 1)$ 分佈. 品質管制裡有所謂**三個標準差規則**(three-sigma rule), 就是源自於此. 對於常態分佈, 觀測值會超過三個標準差之可能性(約為0.0026)是很低的.

例3.4 假設6歲小孩之智商(Intelligence Quotient, 簡稱IQ)可以 $\mathcal{N}(104.5, 263.66)$ 分佈當做模式. 則任一6歲小孩, 其智商介於 $(104.5 - 3\sqrt{263.66},$

$104.5 + 3\sqrt{263.66}$)間, 即介於(55.79, 153.21)間之機率約為0.9974. 此區間長度為6個標準差, 而任一6歲小孩, 智商超過160之機率為

$$\begin{aligned} P(X > 160) &= P\left(\frac{X - 104.5}{\sqrt{263.66}} > \frac{160 - 104.5}{\sqrt{263.66}}\right) \\ &\doteq P(Z > 3.418) \doteq 0.0003. \end{aligned}$$

差不多是萬分之三.

最後, 常態分佈與gamma分佈亦有關係. 當 X 有 $\mathcal{N}(0, 1)$ 分佈, 則 X^2 有 χ_1^2 分佈(習題第13題). 又因當 Y 有 $\mathcal{N}(\mu, \sigma^2)$ 分佈時, $(Y - \mu)/\sigma$ 有 $\mathcal{N}(0, 1)$ 分佈, 故 $(Y - \mu)^2/\sigma^2$ 有 χ_1^2 分佈(此亦為常態分佈之一重要性質). 由於有此關係, 如前所述, 使得卡方分佈, 在統計推論中很重要.

2.3.7 Beta 分佈

Beta分佈(beta distribution)定義在區間 $(0, 1)$, 有二參數 $\alpha, \beta > 0$, 以 $Be(\alpha, \beta)$ 表之, 其p.d.f.為

$$(3.32) \quad f(x) = \frac{\Gamma(\alpha + \beta)}{\Gamma(\alpha)\Gamma(\beta)} x^{\alpha-1} (1-x)^{\beta-1}, 0 < x < 1.$$

在微積分裡, 可證明下述積分:

$$\int_0^1 x^{\alpha-1} (1-x)^{\beta-1} dx$$

對 $\alpha, \beta > 0$ 皆存在. 先定義函數 $B(\alpha, \beta)$ 為

$$(3.33) \quad B(\alpha, \beta) = \int_0^1 x^{\alpha-1} (1-x)^{\beta-1} dx, \alpha, \beta > 0,$$

稱為**beta函數**(beta function). Beta分佈就是因beta函數而命名. 利用重積分, 可證明

$$(3.34) \quad \int_0^1 x^{\alpha-1} (1-x)^{\beta-1} dx = \frac{\Gamma(\alpha)\Gamma(\beta)}{\Gamma(\alpha + \beta)}, \alpha, \beta > 0.$$

即

$$(3.35) \quad B(\alpha, \beta) = \frac{\Gamma(\alpha)\Gamma(\beta)}{\Gamma(\alpha + \beta)}, \alpha, \beta > 0.$$

因此(3.32)的確定義出一p.d.f.

Beta分佈, 是少數取值在一有限區間的常見分佈, 可用來當做取值在0至1的母體之機率模式. $\mathcal{U}(0, 1)$ 分佈也是一種beta分佈, 即為 $\mathcal{B}e(1, 1)$.

由於有二參數, 參數的改變, 可使圖形有很大的變化(見圖3.7). 例如, 因p.d.f.的圖形與 x 軸間之面積為1, 大部分常見的分佈, 其p.d.f.均為有界, 也就是存在一常數 $k > 0$, 使得 $|f(x)| \leq k, \forall x \in R$. 但beta分佈之p.d.f.卻不一定有界. 如 $\mathcal{B}e(1/2, 1)$, 或 $\mathcal{B}e(1, 1/3)$ (讀者是否知道我們之前介紹的分佈, 何者之p.d.f.也不一定有界?). 甚至對 $\mathcal{B}e(1/2, 1/3)$, p.d.f.的圖形為兩端(較靠近 $x = 0$ 及 $x = 1$)較高, 中間較低, 這也是其他分佈所少見的. 雖然p.d.f.定義在 $(0, 1)$, 但若令 $Y = aX, a > 0$, 則 Y 之p.d.f.定義在 $(0, a)$. 因此beta分佈很適合用來作為取值在有限區間之分佈的機率模式.

Beta分佈之動差可很容易地求出:

$$\begin{aligned}
 (3.36) \quad E(X^k) &= \frac{1}{B(\alpha, \beta)} \int_0^1 x^k x^{\alpha-1} (1-x)^{\beta-1} dx \\
 &= \frac{1}{B(\alpha, \beta)} \int_0^1 x^{\alpha+k-1} (1-x)^{\beta-1} dx
 \end{aligned}$$

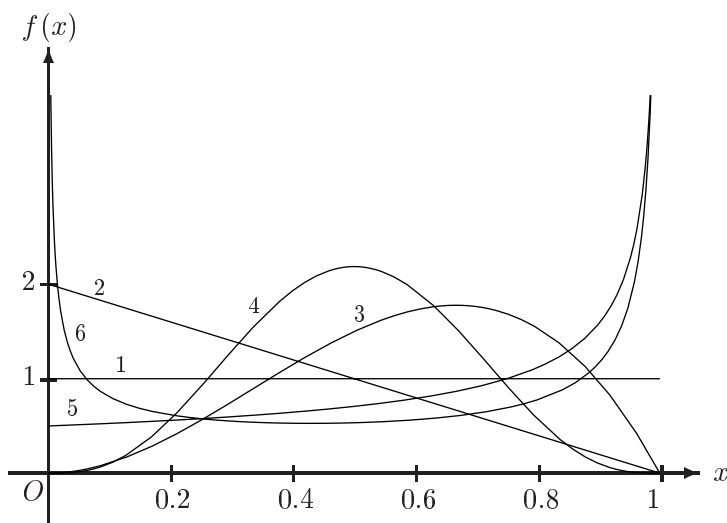


圖3.7 Beta分佈p.d.f.之圖形, 1,2,3,4,5,6分別對應
 $(\alpha, \beta) = (1, 1), (1, 2), (3, 2), (4, 4), (1, 1/2)$, 及 $(1/2, 1/3)$

$$\begin{aligned}
&= \frac{B(\alpha + k, \beta)}{B(\alpha, \beta)} \\
&= \frac{\Gamma(\alpha + k)\Gamma(\alpha + \beta)}{\Gamma(\alpha + \beta + k)\Gamma(\alpha)}.
\end{aligned}$$

上式對 $\alpha + k > 0$, 即 $k > -\alpha$, 皆成立. 由此即得

$$\begin{aligned}
E(X) &= \frac{\alpha}{\alpha + \beta}. \\
\text{Var}(X) &= \frac{\alpha\beta}{(\alpha + \beta)^2(\alpha + \beta + 1)}.
\end{aligned}$$

當 beta 分佈的二參數皆為整數時, 則其分佈函數與二項分佈有密切的關係, 我們留至 11.2 節再說明. 至於拉普拉斯轉換則不易求出.

2.3.8 柯西分佈

柯西分佈, 是因大數學家柯西而命名, 在 1.5 節我們曾介紹標準的 $\mathcal{C}(0, 1)$ 分佈. 在此隨機變數 X 稱為有 $\mathcal{C}(\theta, a)$ 分佈, $\theta \in R$, $a > 0$, 若其 p.d.f. 為

$$(3.37) \quad f(x) = \frac{a}{\pi(a^2 + (x - \theta)^2)}, x \in R.$$

因

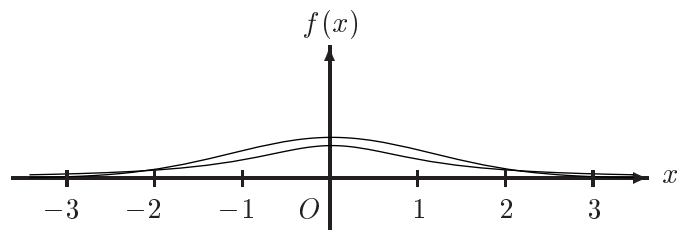
$$\int_{-\infty}^{\infty} \frac{a}{\pi(a^2 + (x - \theta)^2)} dx = \frac{1}{\pi} \arctan \frac{x - \theta}{a} \Big|_{-\infty}^{\infty} = 1,$$

故 (3.37) 式的確定義出一 p.d.f. 又分佈函數為

$$(3.38) \quad F(x) = \int_{-\infty}^x \frac{a}{\pi(a^2 + (x - \theta)^2)} dx = \frac{1}{2} + \frac{1}{\pi} \arctan \frac{x - \theta}{a}, x \in R.$$

對 X 有 $\mathcal{C}(\theta, a)$ 分佈, 令 $Y = (X - \theta)/a$, 則 Y 有 $\mathcal{C}(0, 1)$ 分佈. 常態分佈也有類似的性質.

柯西分佈有二參數 θ, a , p.d.f. 之圖形亦為鐘形, 不仔細看, 還不容易與常態分佈 p.d.f. 之圖形區別. 圖 3.8 將 $\mathcal{N}(0, 1)$ 及 $\mathcal{C}(0, 1)$ p.d.f. 之圖形放在一起比較. 希望讀者能正確指出何者對應 $\mathcal{N}(0, 1)$ 分佈, 何者對應 $\mathcal{C}(0, 1)$ 分佈.

圖3.8 $\mathcal{N}(0, 1)$ 與 $\mathcal{C}(0, 1)$ p.d.f.圖形之比較

雖然都是鐘形，與常態分佈p.d.f.之圖形有一很大的差異，就是柯西分佈p.d.f.之圖形下降至0的速度慢很多，可證明

$$\lim_{x \rightarrow \infty} \frac{1}{\sqrt{2\pi}} e^{-x^2/2} \bigg/ \frac{1}{\pi(1+x^2)} = 0.$$

雖然 $\mathcal{N}(0, 1)$ 分佈之任意次動差皆存在(見(3.27)及(3.28)式):

$$\int_0^\infty x^k \cdot \frac{1}{\sqrt{2\pi}} e^{-x^2/2} dx < \infty, k > 0,$$

但即使 $k = 1$,

$$\int_0^\infty x \cdot \frac{1}{\pi(1+x^2)} dx = \infty.$$

事實上對任一 $\mathcal{C}(\theta, a)$ 分佈，其任意次動差皆不存在，柯西分佈之拉普拉斯轉換，及動差母函數也皆不存在，但特徵函數存在(注意：任一分佈之特徵函數均存在)。即若 X 有 $\mathcal{C}(\theta, a)$ 分佈，則

$$(3.39) \quad E(e^{itX}) = e^{-a|t|+it\theta}, t \in R.$$

上式之推導要利用到複變函數之技巧，在此略去。

由於p.d.f.之圖形對稱於 $x = \theta$ ，因此 $\mathcal{C}(\theta, a)$ 分佈之中位數為 θ ，眾數當然也是 θ 。再加上與常態分佈p.d.f.圖形的表面上看起來相近，其實卻有顯著差異，使得柯西分佈在統計理論裡扮演著特別的角色。另外，柯西分佈有時會在我們料想不到的地方出現，我們下一章再說明。

2.3.9 對數常態分佈

隨機變數 X , 若滿足 $\log X$ 有 $\mathcal{N}(\mu, \sigma^2)$ 分佈, 便稱有**對數常態分佈**(log-normal distribution), 參數為 μ, σ^2 . 利用變數代換, 可得(習題第31題) X 之p.d.f.為

$$(3.40) \quad f(x) = \frac{1}{\sqrt{2\pi}\sigma x} e^{-(\log x - \mu)^2 / (2\sigma^2)}, \quad x > 0,$$

其中 $\mu \in R, \sigma > 0$. 乍看之下, X 之動差不太好求. 但由於 $Y = \log X$ 有 $\mathcal{N}(\mu, \sigma^2)$ 分佈, 利用常態分佈之動差母函數(見(3.25)式), 可得(注意 $X = e^Y$)

$$\begin{aligned} E(X) &= E(e^Y) = e^{\mu + \sigma^2/2}, \\ \text{Var}(X) &= E(X^2) - (E(X))^2 = E(e^{2Y}) - (e^{\mu + \sigma^2/2})^2 \\ &= e^{2(\mu + \sigma^2)} - e^{2\mu + \sigma^2}. \end{aligned}$$

甚至任意次動差皆可求出:

$$(3.41) \quad E(X^k) = E(e^{kY}) = e^{\mu k + \sigma^2 k^2/2}.$$

上式其實對 $\forall k \in R$ 皆成立. 至於拉普拉斯轉換則不易求出.

因 $\log X$ 有常態分佈, 所以稱 X 有**反對數常態分佈**可能較恰當(有如反正弦函數), 不過約定俗成, 我們仍採對數常態的稱呼.

在什麼情況下對數常態分佈適用? 當觀測的數據為**右傾**(skewed to the right), 常可以對數常態為其模式. 例如, 國民所得之分佈通常為右傾: 高收入的人較少, 低收入的人較多. 我們給一些對數常態p.d.f.的圖形如圖3.9.

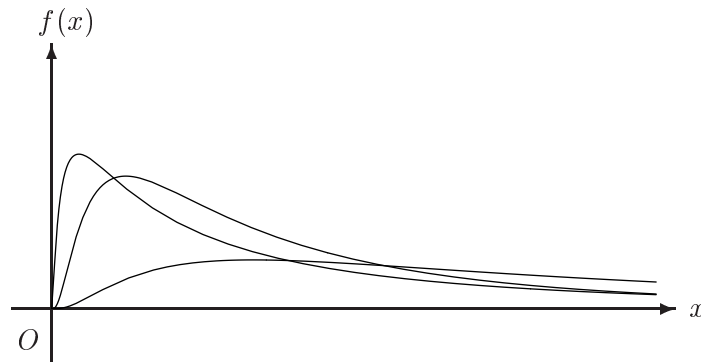


圖3.9 對數常態分佈一些p.d.f.的圖形

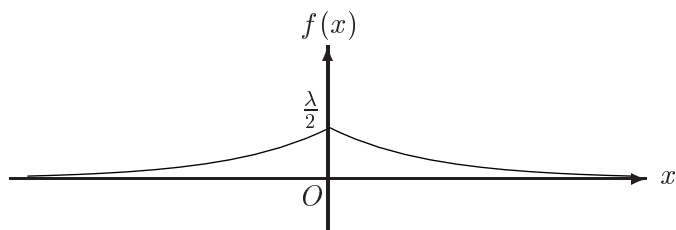
註3.2 一隨機變數 X , 若滿足 $E((X - \mu_X)^3) > 0$, 其中 $\mu_X = E(X)$, 則其分佈稱為**右傾**, 或稱**正傾**(positively skewed); 若 $E((X - \mu_X)^3) < 0$, 則稱其分佈為**左傾**(skewed to the left), 或稱**負傾**(negatively skewed). 若 X 之p.d.f. 為偶函數, 且 $E(X^3)$ 存在, 則 $E((X - \mu_X)^3) = 0$.

2.3.10 雙指數分佈

若隨機變數 X 有p.d.f.

$$(3.42) \quad f(x) = \frac{\lambda}{2} e^{-\lambda|x|}, x \in R,$$

其中 $\lambda > 0$, 則稱 X 有參數 λ 之**雙指數分佈**(double exponential distribution), 以 $\mathcal{D}\text{-}\mathcal{E}(\lambda)$ 表之. 雙指數分佈之p.d.f.為一偶函數. 簡單地講, 雙指數分佈就是將原來指數分佈一半的份量, 移至小於0之處. 其圖形約略如圖3.10. 不難看出 $Y = |X|$ 有 $\mathcal{E}(\lambda)$ 分佈(因此 $E(X^2)$ 與 $\mathcal{E}(\lambda)$ 分佈之二次動差必相同).

圖3.10 $\mathcal{D}\text{-}\mathcal{E}(\lambda)$ p.d.f.之圖形

$\mathcal{D}\text{-}\mathcal{E}(\lambda)$ 分佈p.d.f.之圖形並非鐘形, 事實上p.d.f.之圖形在 $x = 0$ 不可微. 由於p.d.f.含有 $|x|$ 項, 在做積分時, 要分 $x \geq 0$ 及 $x < 0$. 又可求出

$$(3.43) \quad E(X) = 0,$$

$$(3.44) \quad \text{Var}(X) = E(X^2) = \frac{2}{\lambda^2},$$

及動差母函數

$$\begin{aligned}
 (3.45) \quad E(e^{tX}) &= \int_{-\infty}^{\infty} \frac{\lambda}{2} e^{tx} e^{-\lambda|x|} dx \\
 &= \frac{\lambda}{2} \int_{-\infty}^0 e^{(\lambda+t)x} dx + \frac{\lambda}{2} \int_0^{\infty} e^{-(\lambda-t)x} dx \\
 &= \frac{\lambda}{2(\lambda+t)} e^{(\lambda+t)x} \Big|_{-\infty}^0 - \frac{\lambda}{2(\lambda-t)} e^{-(\lambda-t)x} \Big|_0^{\infty} \\
 &= \frac{\lambda}{2(\lambda+t)} + \frac{\lambda}{2(\lambda-t)} = \frac{\lambda^2}{\lambda^2 - t^2},
 \end{aligned}$$

對每一 $|t| < \lambda$ 皆存在.

雙指數分佈又稱**拉普拉斯分佈**(Laplace distribution), 較一般的形式之 p.d.f. 為

$$(3.46) \quad f(x) = \frac{\lambda}{2} e^{-\lambda|x-\alpha|}, \quad x \in R, \lambda > 0, \alpha \in R.$$

有二參數 α, λ . 此分佈之期望值為 α , 變異數為 $2/\lambda^2$.