

第一章

基本概念

1.6 期望值

對一隨機變數, 我們常想粗略地知道其值究竟多大, **期望值**(expectation, 或稱expected value, mean), 就是常被拿來扮演這種以一單一的值, 來代表一隨機現象中之變數大小的角色.

設隨機變數 X 之p.d.f.為 f , 則其期望值以 $E(X)$ 表示, 定義為

$$(6.1) \quad E(X) = \begin{cases} \int_{-\infty}^{\infty} xf(x)dx & , \text{連續型}, \\ \sum_x xf(x) & , \text{離散型}, \end{cases}$$

只要上述積分或和存在, 此處 $\sum_x xf(x)$ 表對所有可能的 x (即 $f(x) \neq 0$)相加. 通常以 μ_X (或只以 μ)表 $E(X)$.

假設有10個數: 1, 1, 2, 2, 2, 3, 3, 3, 4, 4, 則其平均為

$$\frac{1}{10}(1 + 1 + 2 + 2 + 2 + 3 + 3 + 3 + 4 + 4) = 2.5.$$

由於數字中有重複: 有2個1, 3個2, 3個3, 2個4, 所以平均亦可表為

$$\frac{1}{10}(1 \cdot 2 + 2 \cdot 3 + 3 \cdot 3 + 4 \cdot 2) = 2.5,$$

2 第一章 基本概念

所得相同. 第二種作法是我們常採用的, 特別是在數字較大時, 又上式左側可改寫為

$$1 \cdot \frac{2}{10} + 2 \cdot \frac{3}{10} + 3 \cdot \frac{3}{10} + 4 \cdot \frac{2}{10},$$

其中 $2/10$, $3/10$, $3/10$, $2/10$, 可分別視為10個數字中, 1,2,3,4所佔的份量, 所以第二種作法便是一種加權平均. 隨機變數期望值的定義, 便是基於加權平均的想法. 這點由離散型隨機變數期望值的定義, 特別可以看出來.

至於 X 之一函數 $g(X)$ 的期望值, 則定義為

$$(6.2) \quad E(g(X)) = \begin{cases} \int_{-\infty}^{\infty} g(x)f(x)dx, & \text{連續型,} \\ \sum_x g(x)f(x) & \text{離散型.} \end{cases}$$

亦可經由變數代換, 先求出 $Y = g(X)$ 之p.d.f., 再利用(6.1)式而求得 $E(Y)$.

底下給幾個例子.

例6.1 設離散型隨機變數 X 之p.d.f.為

$$\begin{aligned} f(0) &= \frac{1}{16}, & f(1) &= \frac{2}{16}, & f(2) &= \frac{3}{16}, & f(3) &= \frac{4}{16}, \\ f(4) &= \frac{4}{16}, & f(5) &= \frac{1}{16}, & f(6) &= \frac{1}{16}. \end{aligned}$$

則

$$\begin{aligned} E(X) &= \sum_{x=0}^6 xf(x) \\ &= 0 \cdot \frac{1}{16} + 1 \cdot \frac{2}{16} + 2 \cdot \frac{3}{16} + 3 \cdot \frac{4}{16} + 4 \cdot \frac{4}{16} + 5 \cdot \frac{1}{16} + 6 \cdot \frac{1}{16} \\ &= \frac{47}{16}. \end{aligned}$$

又

$$\begin{aligned} E((X-3)^2) &= \sum_{x=0}^6 (x-3)^2 f(x) \\ &= 9 \cdot \frac{1}{16} + 4 \cdot \frac{2}{16} + 1 \cdot \frac{3}{16} + 0 \cdot \frac{4}{16} + 1 \cdot \frac{4}{16} + 4 \cdot \frac{1}{16} + 9 \cdot \frac{1}{16} \\ &= \frac{37}{16}. \end{aligned}$$

在求 $E((X - 3)^2)$ 時, 也可利用變數代換, 令 $Y = (X - 3)^2$, 則 Y 取值 $0, 1, 4, 9$, 且

$$\begin{aligned}f_Y(0) &= P(X = 3) = \frac{4}{16}, \\f_Y(1) &= P(X = 2) + P(X = 4) = \frac{7}{16}, \\f_Y(4) &= P(X = 1) + P(X = 5) = \frac{3}{16}, \\f_Y(9) &= P(X = 0) + P(X = 6) = \frac{2}{16}.\end{aligned}$$

故

$$\begin{aligned}E((X - 3)^2) &= E(Y) = \sum_y y f_Y(y) \\&= 0 \cdot \frac{4}{16} + 1 \cdot \frac{7}{16} + 4 \cdot \frac{3}{16} + 9 \cdot \frac{2}{16} \\&= \frac{37}{16}.\end{aligned}$$

其中 $\sum$ 是對所有的 y 相加, 即 $0, 1, 4, 9$.

例6.2 設隨機變數 X 之p.d.f.為

$$f(x) = 2x, x \in [0, 1].$$

則

$$\begin{aligned}E(X) &= \int_0^1 x \cdot 2x dx = \int_0^1 2x^2 dx = \frac{2}{3}, \\E(X^2) &= \int_0^1 x^2 \cdot 2x dx = \int_0^1 2x^3 dx = \frac{1}{2}.\end{aligned}$$

另外, 亦可利用變數代換, 令 $Y = X^2$, 先求出 Y 之p.d.f.為

$$g(y) = 2\sqrt{y} \cdot \frac{1}{2\sqrt{y}} = 1, y \in [0, 1].$$

如此

$$E(X^2) = E(Y) = \int_0^1 y \cdot 1 dy = \frac{1}{2},$$

仍得到相同的結果.

例6.3 在微積分裡有下述結果:

$$\sum_{i=1}^{\infty} \frac{1}{i^2} = \frac{\pi^2}{6}.$$

故

$$(6.3) \quad f(i) = \frac{6}{\pi^2} \frac{1}{i^2}, i = 1, 2, \dots,$$

爲一p.d.f. 但因

$$\sum_{i=1}^{\infty} \frac{1}{i} = \infty,$$

故若隨機變數 X 之p.d.f.如(6.3)式, 則 X 之期望值不存在.

另外, 因

$$\sum_{i=1}^{\infty} \frac{1}{i^4} = \frac{\pi^4}{90},$$

故

$$(6.4) \quad g(i) = \frac{90}{\pi^4} \frac{1}{i^4}, i = 1, 2, \dots,$$

爲一p.d.f. 又設隨機變數 Y 之p.d.f.如(6.4)式所給, 則

$$E(Y) = \frac{90}{\pi^4} \sum_{i=1}^{\infty} \frac{1}{i^3} < \infty,$$

$$E(Y^2) = \frac{90}{\pi^4} \sum_{i=1}^{\infty} \frac{1}{i^2} = \frac{90}{\pi^4} \frac{\pi^2}{6} = \frac{15}{\pi^2} < \infty.$$

即 $E(Y)$ 及 $E(Y^2)$ 皆存在, 當然 $E(Y^3)$ 便不存在了.

例6.4 設隨機變數 X 有 $\mathcal{C}(0, 1)$ 分佈, 即其p.d.f.爲

$$f(x) = \frac{1}{\pi(1+x^2)}, x \in R,$$

則 X 之期望值不存在, 此因

$$E(X) = \int_{-\infty}^{\infty} x \cdot \frac{1}{\pi(1+x^2)} dx,$$

爲一瑕積分(improper integral), 而依定義, 此瑕積分存在, 若且唯若

$$\int_0^{\infty} \frac{x}{\pi(1+x^2)} dx, \quad \int_{-\infty}^0 \frac{x}{\pi(1+x^2)} dx$$

皆要存在, 但此二積分皆不存在:

$$\int_0^{\infty} \frac{x}{\pi(1+x^2)} dx = \lim_{x \rightarrow \infty} \frac{1}{2\pi} \log(1+x^2) = \infty,$$

故 $E(X)$ 不存在.

要注意的是, 不可以爲因 $x/(\pi(1+x^2))$ 爲一奇函數, 所以在 $(0, \infty)$ 的積分與在 $(-\infty, 0)$ 的積分, 剛好正負消掉, 而得積分爲0. 要知道並沒有 $\infty - \infty = 0$ 的運算, ∞ 並非一實數, 實數裡的一些運算對它並不一定適用. 不過當一隨機變數 X 之p.d.f.爲偶函數, 則 $E(X)$ 若存在必爲0, 證明留在習題中. 因此由於標準的羅吉斯分佈之p.d.f. f_X 爲偶函數(見例4.10), 且

$$\begin{aligned} \int_0^{\infty} x f_X(x) dx &= \int_0^{\infty} x \cdot \frac{e^{-x}}{(1+e^{-x})^2} dx \leq \int_0^{\infty} x e^{-x} dx \\ &= - \int_0^{\infty} x d e^{-x} = -x e^{-x} \Big|_0^{\infty} + \int_0^{\infty} e^{-x} dx = 1 < \infty, \end{aligned}$$

故標準羅吉斯分佈之期望值存在且爲0.

附帶一提, 由上述瑕積分存在的定義, 可看出 $E(X)$ 存在, 若且唯若 $E(|X|)$ 存在(見習題第29題), 有些書便以 $E(|X|)$ 存在, 當做期望值存在的定義.

期望值有不少性質. 例如, 一**有界的**隨機變數, 期望值必存在. 在此對一隨機變數 X , 若存在一常數 $M > 0$, 使得 $P(|X| \leq M) = 1$, 便稱該隨機變數爲**有界的**. 若 X 有 $\mathcal{U}(0, 1)$ 分佈, 便爲有界, 有 $\mathcal{C}(0, 1)$ 分佈, 則非有界. 我們再列出一些性質如下.

定理6.1 設 X 爲一隨機變數, a, b, c 爲常數, g_1, g_2 爲二函數, 使得 $E(g_1(X)), E(g_2(X))$ 皆存在, 則

6 第一章 基本概念

- (i) $E(ag_1(X) + bg_2(X) + c) = aE(g_1(X)) + bE(g_2(X)) + c$;
- (ii) 若 $g_1(x) \geq 0, \forall x \in R$, 則 $E(g_1(X)) \geq 0$;
- (iii) 若 $g_1(x) \geq g_2(x), \forall x \in R$, 則 $E(g_1(X)) \geq E(g_2(X))$;
- (iv) 若 $a \leq g_1(x) \leq b, \forall x \in R$, 則 $a \leq E(g_1(X)) \leq b$.

由上定理之(i), 特別地, 只要 $E(X)$ 存在, 則

$$(6.5) \quad E(aX + b) = aE(X) + b,$$

$$(6.6) \quad E(b) = b.$$

即一常數之期望值等於該常數. 由於 $E(X)$ 為一常數(已非隨機了), 故其期望值仍為本身. 即

$$E(E(X)) = E(X).$$

期望值這個名詞也許會令有些人感到困擾. 投擲一公正的骰子一次, 令 X 表所得之點數. 則 X 之 p.d.f. 為

$$f(x) = \frac{1}{6}, x = 1, 2, \dots, 6.$$

因此

$$E(X) = \sum_{i=1}^6 i \cdot \frac{1}{6} = \frac{21}{6} = 3.5.$$

投擲一骰子, 點數之期望值為 3.5. 點數為整數, 無論怎麼投擲都得不到 3.5. 那怎會期望得到 3.5 呢? 唐詩三百首裡, 元稹遣悲懷:

同穴窅冥何所望, 他生緣會更難期.

3.5 的期望值, 可說較他生緣更難期了.

所以暫時不要去想期望值字面上的含義, 是不易想通的. 不過期望值是一隨機變數之“最佳”常數的代表值. 所謂最佳, 指的是誤差平方的期望值最小. 見下定理.

定理 6.2 設隨機變數 X 滿足 $E(X^2) < \infty$. 則對所有 $a \in R$, $E((X - a)^2)$ 之極小值發生在 $a = E(X)$.

證明.由

$$E((X - a)^2) = E(X^2) - 2aE(X) + a^2,$$

將 $E(X^2) - 2aE(X) + a^2$ 視為 $-a$ 的二次式,經由微分可得極小值發生在 $a = E(X)$.得證.

在3.8節我們會看到若 $E(X^2)$ 存在,則導致 $E(X)$ 存在.

想以一常數 a 來代表一隨機變數 X ,很難說那一個 a 最好.誤差 $X - a$ 仍是隨機的.但在誤差的平方 $(X - a)^2$ 之期望值最小下,上定理指出期望值 $E(X)$ 便是最好的選擇.

其次我們定義**動差**(moment).設 n 為一整數,則若 $E(X^n)$ 存在,便稱此為隨機變數 X 之 **n 次動差**.而 $E((X - \mu)^n)$ 稱為 X 之 **n 次中心動差**(n th central moment),其中 $\mu = E(X)$.特別地,二次中心動差稱為 X 之**變異數**(variance),以 σ_X^2 , $\text{Var}(X)$ 或簡單地以 σ^2 表之,即

$$(6.7) \quad \text{Var}(X) = E((X - E(X))^2).$$

$\text{Var}(X)$ 之正的平方根,以 σ_X 或 σ 表之,稱為 X 之**標準離差**(standard deviation),或**標準差**、**均方差**.另外,如同期望值,對於 n 次動差仍有 $E(X^n)$ 存在,若且唯若 $E(|X|^n)$ 存在,其證明留在習題中.

如同期望值,標準離差此一名詞可能令人感到迷惑.這並不表示隨機變數 X 有較大的機率與期望值之差距不超過一個標準離差.常有人會有這樣的誤解.如見目前通行的某版本高中數學甲(上)(即自然組高三上學期).我們給一例子如下,讀者不難造出更多的例子.

設 $P(X = 0) = 0.1$, $P(X = 2) = P(X = -2) = 0.45$.則

$$\mu = E(X) = 0,$$

$$\sigma^2 = \text{Var}(X) = E(X^2) = 0.1 \cdot 0^2 + 0.45 \cdot 2^2 + 0.45 \cdot (-2)^2 = 3.6.$$

因此 X 之標準離差 $\sigma = \sqrt{3.6} \doteq 1.8972 < 2$.但有0.9之機率, X 取值與 $\mu = 0$ 之差距超過 σ .

在機率論裡我們引進了隨機變數的概念.例如,學生不再如小學的練習裡,每位體重40公斤,求25位學生的總體重,而是讓每位學生的體重有一機率分佈.但因覺得無法掌握隨機的量之大小,我們才想要有一代表值.由於有很多好的性質,期望值遂常被拿來當做隨機變數之代表值.期望值像是隨機變數分佈之一個核心,隨機變數的可能值,散佈在期望值的左右.其他亦常被拿來當做隨機變數之代表值的尚有**中位數**(median)及**眾數**(mode)等.眾數早期稱為**密集數**.在此一實數 m 若滿足

$$(6.8) \quad P(X \leq m) \geq \frac{1}{2}, \quad \text{且} \quad P(X \geq m) \geq \frac{1}{2},$$

便稱為隨機變數 X (或其分佈)之一**中位數**.依此定義中位數不一定唯一.當然如果 X 之分佈函數為連續且嚴格漸增,則中位數便唯一.此時中位數為滿足 $F(x) = 0.5$ 之 x ,即 X 之第0.5分位數.有些書便將中位數定義為 X 之第0.5分位數.依其定義中位數便唯一存在了.若 X 之分佈對稱於 m ,則 X 之中位數為 m (習題第16題).另外,若 $f(a)$ 為 X 之機率密度函數的一絕對極大值,便稱 a 為 X (或其分佈)之一**眾數**.對 X 為離散型的隨機變數,其眾數表 X 最可能取的值.眾數也不一定唯一.中位數及眾數有時雖存在但卻不易求出(見習題第16題).

例6.5 設 X 為一離散型的隨機變數,且 $P(X = 0) = 1/4$, $P(X = 1) = 1/2$, $P(X = 2) = 1/4$.則 X 有唯一的中位數1:

$$\begin{aligned} P(X \leq 1) &= P(X = 0) + P(X = 1) = \frac{3}{4} \geq \frac{1}{2}, \\ P(X \geq 1) &= P(X = 1) + P(X = 2) = \frac{3}{4} \geq \frac{1}{2}. \end{aligned}$$

故1為一中位數.又任一 $a < 1$,

$$P(X \leq a) \leq P(X = 0) = \frac{1}{4} < \frac{1}{2},$$

任一 $a > 1$,

$$P(X \geq a) \leq P(X = 2) = \frac{1}{4} < \frac{1}{2},$$

皆非中位數.

至於 X 之眾數亦為1.

例6.6 設 X 為一離散型的隨機變數, 且 $P(X = 0) = 0.1$, $P(X = 1) = 0.4$, $P(X = 2) = 0.4$, $P(X = 3) = 0.1$. 則任一 $a \in [1, 2]$ 皆為中位數, 而1,2皆為眾數.

例6.7 設隨機變數 X 之分佈函數為

$$F(x) = \begin{cases} 0 & , x < 0, \\ x^2 & , 0 \leq x < 1, \\ 1 & , x \geq 1. \end{cases}$$

令 $x^2 = 1/2$, 解出 $x = \sqrt{2}/2 \doteq 0.707$. 即 X 之中位數為 $\sqrt{2}/2$. 也可以說 F 之中位數為 $\sqrt{2}/2$. 若取 X 之p.d.f.為 $f(x) = 2x$, $0 \leq x \leq 1$, 則 X 之眾數為1. 若取 X 之p.d.f.為 $f(x) = 2x$, $0 \leq x < 1$, 則眾數不存在.

我們常要藉助統計做各種決策. 投資某項事業, 可能成功而獲利若干, 可能失敗而損失若干. 那到底要不要投資呢? 不能只看到成功機率很大(說不定獲利很小), 也不能只看到獲利很大(說不定成功機率很低), 算一算所得(若損失所得便為負)之期望值, 是一簡單的決策依據. 我們平常在判斷出門要不要帶雨傘, 也是先盤算一下期望損失(先估計下雨的機率, 下雨不帶傘的損失等). 期望值在做決策時, 常扮演重要的角色.

例6.8 某賭局的賭法如下: 投擲一公正的銅板, 若第 n 次投擲才首度出現正面, 則賭徒可獲得 1.95^n 元. 試依期望所得, 決定賭徒每次賭之前, 該付多少元, 此賭局才是公平.

解. 不難看出第 n 次投擲才首度出現正面之機率為 2^{-n} , $n \geq 1$. 令 X 表賭徒賭一次之所得, 則

$$E(X) = \sum_{n=1}^{\infty} 1.95^n \cdot 2^{-n} = \sum_{n=1}^{\infty} \left(\frac{1.95}{2}\right)^n$$

$$= \frac{1.95/2}{1 - 1.95/2} = 39.$$

故每次該付39元.

不論期望值是多好的一個隨機變數之代表值, 難免會偏離該隨機變數, 如何量測偏差究竟有多大呢? 這就是變異數與標準差的功能.

設有一隨機變數 X , 且設期望值 $E(X)$ 存在, 則 X 與 $E(X)$ 之離差為 $X - E(X)$, 此值有正有負, 但期望值

$$E(X - E(X)) = E(X) - E(X) = 0.$$

離差之期望值為0, 此亦為期望值之一性質, 但我們往往是對離差之絕對值較感興趣, 此正如射飛鏢時, 有時上偏有時下偏, 總不能得意地說平均命中紅心.

以 $|X - E(X)|$ 表離差之絕對值, 此量仍為一隨機變數, 如何衡量其大小呢? 取期望值為一辦法, 也就是求量測離差之絕對值的“平均”, 但如果你微積分學得還不錯的話, 應知道有關絕對值函數之求和或積分, 往往有些麻煩(想想求 $\int_{-3}^3 |\sin x - 0.3x| dx$ 的例子). 對於 $E(|X - E(X)|)$, 要討論何時 $X \geq E(X)$, 何時 $X < E(X)$, 因此通常的作法, 是考慮離差平方之期望值, 而這就是變異數. 由於是平方的期望值, 像是距離的平方, 因此若將變異數開平方, 就像使其回到距離, 標準差便產生了.

在統計裡是允許誤差的存在, 我們不要求百分之百的精準, 我們只希望誤差在可接受的範圍, 而這誤差其實也是隨機的, 因此我們才考慮“平均式”(取期望值)的誤差. 大家常說, 雖不中亦不遠矣, 但不遠到底是多遠? 標準差的功能, 就是對誤差之大小, 提供一參考值. 如果說期望值是描述一隨機變數之分佈的“集中處”, 標準差便是描述分佈對該集中處之偏離程度. 標準差愈小(或等價地說變異數愈小), 表隨機變數之分佈較集中在期望值附近; 標準差愈大, 表隨機變數之分佈, 散佈的較開. 人們常說命運, 命就有點像是期望值, 運則像是變異數. 命也許不太好(期望值不大), 但由於有變異的關係, 使得有時好運會發生.

假設兩組學生參加測驗，每組各有三人，平均成績皆為60分，但第一組之成績為61, 60, 59；第二組之成績為90, 60, 30。第一組學生的程度很接近，而第二組學生的程度則有很大差異。光由平均成績，並無法了解兩組學生程度差異之情況。這也是對隨機變數，為何除了期望值，我們尚須給出其標準差。給了此二值，對其“中心”及“散佈範圍”，才有一較清晰的概念。

例6.9 假設有 A, B 二項投資。投資 A 於一年後，有 $1/2$ 的機率賺一倍的錢，有 $1/4$ 的機率賠一半的錢，有 $1/4$ 的機率不賺不賠。投資 B 一年後，有 $3/4$ 的機率賺一半的錢，有 $1/4$ 的機率不賺不賠。若放10,000元於投資 A ，則一年後之獲利 X ，為一離散型的隨機變數，p.d.f.為(略去“元”，底下同)

$$f_X(x) = \begin{cases} 1/2, & x = 10,000, \\ 1/4, & x = -5,000, \\ 1/4, & x = 0. \end{cases}$$

若放10,000於投資 B ，則一年後之獲利 Y ，亦為一離散型的隨機變數，p.d.f.為

$$f_Y(y) = \begin{cases} 3/4, & y = 5,000, \\ 1/4, & y = 0. \end{cases}$$

依上述條件，可得

$$E(X) = \sum_x x f_X(x) = \frac{1}{2} \cdot 10,000 + \frac{1}{4} \cdot (-5,000) + \frac{1}{4} \cdot 0 = 3,750.$$

$$E(Y) = \sum_y y f_Y(y) = \frac{3}{4} \cdot 5,000 + \frac{1}{4} \cdot 0 = 3,750.$$

即投資 A 與投資 B 獲利之期望值相同。雖然有相同的期望值，但 X 之分佈較 Y 之分佈，散佈的較開，所以我們猜測 X 之變異數較大。底下來驗證此猜測。

$$\begin{aligned} \text{Var}(X) &= \sum_x (x - 3,750)^2 f_X(x) \\ &= \frac{1}{2}(10,000 - 3,750)^2 + \frac{1}{4} \cdot (-5,000 - 3,750)^2 + \frac{1}{4}(0 - 3,750)^2 \end{aligned}$$

$$\begin{aligned}
&= 42,187,500. \\
\text{Var}(Y) &= \sum_y (y - 3,750)^2 f_Y(y) \\
&= \frac{3}{4}(5,000 - 3,750)^2 + \frac{1}{4}(0 - 3,750)^2 \\
&= 4,687,500.
\end{aligned}$$

由於 $\text{Var}(X) > \text{Var}(Y)$, 表示投資 A 之變異較大. 若喜愛冒險者, 說不定會採投資 A , 有時會有較高獲利. 個性保守者, 可能會採投資 B , 穩紮穩打.

例6.10 某項食品是以機械填裝, 每包標示16公斤. 若容量不足, 顧客可能會抱怨, 若容量過多, 公司也有損失. 令 X 表每包之容量, 假設 X 有 p.d.f.

$$f(x) = 2.5, \quad 15.8 \leq x \leq 16.2,$$

即 X 有 $\mathcal{U}[15.8, 16.2]$ 分佈, 則

$$E(X) = \int_{-\infty}^{\infty} xf(x)dx = \int_{15.8}^{16.2} 2.5xdx = 1.25x^2 \Big|_{15.8}^{16.2} = 16.0.$$

所以平均而言, 機器填裝的容量是正確的. 但

$$P(X < 16) = P(X > 16) = \frac{1}{2},$$

即各有一半的機率容量不足及容量超過. 我們來求其變異數.

$$\begin{aligned}
\text{Var}(X) &= \int_{-\infty}^{\infty} (x - 16)^2 f(x)dx \\
&= \int_{15.8}^{16.2} 2.5(x - 16)^2 dx = \frac{2.5}{3}(x - 16)^3 \Big|_{15.8}^{16.2} = 0.0133.
\end{aligned}$$

標準差為 $\sqrt{\text{Var}(X)} \doteq 0.115$.

爲了降低變異, 公司引進新機器, 新的容量 Y 之 p.d.f. 爲

$$g(y) = \begin{cases} 25(x - 16) + 5, & 15.8 \leq x < 16, \\ 25(16 - x) + 5, & 16 \leq x < 16.2. \end{cases}$$

則仍有(習題第1題), $E(Y) = 16$, 但 $\text{Var}(Y) = 1/150 \doteq 0.00\overline{6}$, 標準差 $\doteq 0.0816$. 變異數下降一半以上, 標準差則下降約29%.

圖6.1給出 X 及 Y p.d.f.之圖形.由圖形亦可看出 Y 之變異數較 X 之變異數小:較多的機率配置在 Y 之期望值的近傍,但在可能取的值為區間 $[15.8, 16.2]$,且期望值為16之隨機變數中,上述 X 倒並非變異數最大的隨機變數,此點做了習題第33題便可了解.

我們先給變異數之一些性質.

定理6.3 設 X 為一隨機變數,則下述二式成立,其中 a, b 為常數.

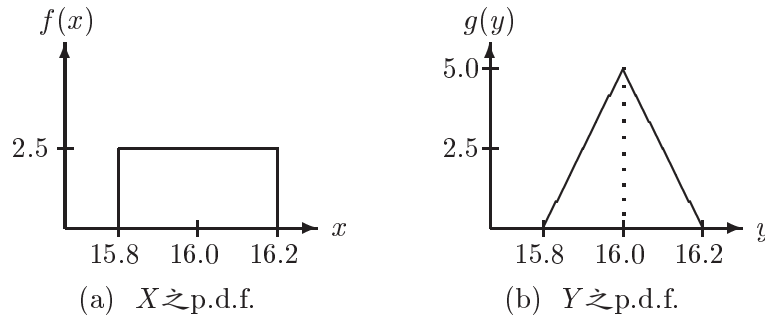


圖6.1 例6.10中 X 及 Y p.d.f.之圖形

$$(6.9) \quad \text{Var}(aX + b) = a^2 \text{Var}(X),$$

$$(6.10) \quad \text{Var}(X) = E(X^2) - (E(X))^2.$$

證明.首先因 $E(aX + b) = aE(X) + b$, 由定義得

$$\begin{aligned} \text{Var}(aX + b) &= E((aX + b - (aE(X) + b))^2) \\ &= E((aX - aE(X))^2) \\ &= a^2 E((X - E(X))^2) \\ &= a^2 \text{Var}(X), \end{aligned}$$

其中第三個等號成立, 仍藉助(6.5)式, 得證(6.9)式.

其次

$$\text{Var}(X) = E((X - E(X))^2)$$

$$\begin{aligned}
&= E(X^2 - 2XE(X) + (E(X))^2) \\
&= E(X^2) - 2(E(X))^2 + (E(X))^2 \\
&= E(X^2) - (E(X))^2,
\end{aligned}$$

此處用到(6.6)式($E(X)$ 及 $(E(X))^2$ 皆為常數),得證(6.10)式.

由於變異數一定不是負的(非負隨機變數 $(X - E(X))^2$ 之期望值必為非負,見定理6.1之(ii)),故由(6.10)式得

$$(6.11) \quad E(X^2) \geq (E(X))^2.$$

又由(6.9)式可看出(取 $a = 0$)常數的變異數為0.當然這點由(6.10)式亦可看出.即若存在一常數 b ,使得 $P(X = b) = 1$,則因 $E(X^2) = b^2$, $E(X) = b$,故 $\text{Var}(X) = 0$.這當然是合理的.既然是常數(可視為一退化的隨機變數),就沒有變異,每次觀測皆得到該常數.反之,若 $\text{Var}(X) = 0$,則存在常數 b ,使得 $P(X = b) = 1$.此部分的證明我們留在習題中.

例6.11 投擲一公正的骰子,令 X 表所得之點數,試求 X 之變異數及標準差.

解.前面已求出 $E(X) = 3.5 = 7/2$.又

$$E(X^2) = \sum_{i=1}^6 \frac{1}{6} \cdot i^2 = \frac{91}{6}.$$

故

$$\begin{aligned}
\text{Var}(X) &= E(X^2) - (E(X))^2 \\
&= \frac{91}{6} - \frac{49}{4} = \frac{35}{12},
\end{aligned}$$

且 X 之標準差為 $\sqrt{35/12} \doteq 1.7078$.

例6.12 對一事件 A ,如例3.5,令 I_A 表 A 之指示函數, I_A 為一只取0,1兩個值之隨機變數.且 $I_A = 1$,若 A 發生, $I_A = 0$,若 A 不發生.故

$$E(I_A) = 1 \cdot P(I_A = 1) + 0 \cdot P(I_A = 0) = P(A).$$

$$\begin{aligned}
\text{Var}(I_A) &= E(I_A^2) - (E(I_A))^2 \\
&= E(I_A) - (E(I_A))^2 \\
&= P(A) - (P(A))^2.
\end{aligned}$$

例6.13 承例6.2, 由於 $E(X) = 2/3$, 依定義得

$$\begin{aligned}
\text{Var}(X) &= E((X - E(X))^2) \\
&= \int_0^1 \left(x - \frac{2}{3}\right)^2 \cdot 2x dx \\
&= \int_0^1 \left(2x^3 - \frac{8}{3}x^2 + \frac{8}{9}x\right) dx = \frac{1}{18}.
\end{aligned}$$

另外, 因已求出 $E(X^2) = 1/2$, 利用 (6.10) 式得

$$\text{Var}(X) = E(X^2) - (E(X))^2 = \frac{1}{2} - \frac{4}{9} = \frac{1}{18},$$

所得相同.

例6.9及6.10中之變異數, 亦皆可以(6.10)式求之, 所得必相同. 我們再給一例.

例6.14 設隨機變數 X 之 p.d.f. 為

$$f(x) = \frac{1}{x \log 2}, \quad 1 \leq x \leq 2.$$

則

$$\begin{aligned}
E(X) &= \int_1^2 x \cdot \frac{1}{x \log 2} dx = \frac{1}{\log 2} \int_1^2 1 dx = \frac{1}{\log 2} \doteq 1.443. \\
E(X^2) &= \int_1^2 x^2 \cdot \frac{1}{x \log 2} dx = \frac{1}{\log 2} \int_1^2 x dx = \frac{3}{2 \log 2}.
\end{aligned}$$

故

$$\begin{aligned}
\text{Var}(X) &= E(X^2) - (E(X))^2 \\
&= \frac{3}{2 \log 2} - \frac{1}{(\log 2)^2} \doteq 0.08267.
\end{aligned}$$

標準差為 $\sqrt{\text{Var}(X)} = 0.2875$.

附帶一提, 因 X 之分佈函數

$$F(x) = \int_1^x f(t)dt = \frac{\log x}{\log 2}, 1 \leq x \leq 2,$$

解 $F(x) = 1/2$, 得中位數為 $\sqrt{2}$. 至於眾數則為 1.

對一隨機變數, 於給出一次及二次動差(變異數是基於二次動差)後, 對其分佈有了基本的了解. 至於若繼續知道更多次的動差, 對其分佈就愈了解. 所謂分佈, 顧名思義, 就是一隨機變數的値之散佈情況. 由於散佈著, 我們才又想以一些基本的量來描述散佈的情況. 這就是動差的功能.

另外, 亦可以一些轉換, 來描述隨機變數的分佈. 我們介紹幾個常見且可以唯一決定隨機變數之分佈的轉換. 設 X 為一隨機變數.

若 X 只取非負整數值, 且 $P(X = k) = a_k, k = 0, 1, \dots$, 我們常考慮其母函數(generating function, 或稱生成函數)

$$(6.12) \quad g(s) = E(s^X) = \sum_{k=0}^{\infty} a_k s^k, |s| \leq 1.$$

可看出 $g(1) = 1$. 又

$$(6.13) \quad a_k = \frac{1}{k!} \frac{d^k g(s)}{ds^k}, k = 0, 1, \dots.$$

若 X 取值在一可數的集合 $\{\lambda_0, \lambda_1, \dots\}$, 亦可類似地定義其母函數.

若 X 為非負, 則往往考慮其拉普拉斯轉換(Laplace transform)

$$(6.14) \quad \phi(s) = E(e^{-sX}), s \geq 0.$$

顯然 $\phi(0) = 1$. 若 X 為連續型, 有 p.d.f. f , 則

$$(6.15) \quad \phi(s) = \int_0^{\infty} e^{-sx} f(x) dx, s \geq 0.$$

若 X 為離散型, 取值在 $\{\lambda_0, \lambda_1, \dots\}$, 且 $P(X = \lambda_k) = p_k$, 則

$$(6.16) \quad \phi(s) = \sum_{k=0}^{\infty} p_k e^{-s\lambda_k}, s \geq 0.$$

特徵函數(characteristic function)則是對所有隨機變數均適用, 且必存在:

$$(6.17) \quad \psi(t) = E(e^{itX}), t \in R, i = \sqrt{-1}.$$

仍有 $\psi(0) = 1$.

上述這些轉換都具有一以貫之的功能, 能以一扼要的方式, 來包含關於隨機變數 X 分佈之所有資訊. 由這些轉換, 也可得到 X 之各次動差. 例如, 設 $\phi(s)$ 為 X 之拉普拉斯轉換, 則只要 X 的 n 次動差存在,

$$(6.18) \quad E(X^n) = (-1)^n \left. \frac{d^n \phi(s)}{ds^n} \right|_{s=0} = (-1)^n \phi^{(n)}(0).$$

又這些轉換只要存在, 彼此便有一簡單的關係. 例如,

$$(6.19) \quad \phi(s) = g(e^{-s}), s \geq 0,$$

$$(6.20) \quad g(s) = \phi(-\log s), 0 < s < 1.$$

最後

$$(6.21) \quad M(t) = E(e^{tX})$$

稱為 X 之**動差母函數**(moment generating function). 不過動差母函數不一定存在, 有時只在 t 屬於某一範圍才存在, 但 $M(0) = 1$ 必存在.

動差母函數若存在, 亦可經由微分得到動差:

$$(6.22) \quad E(X^n) = \left. \frac{d^n M(t)}{dt^n} \right|_{t=0} = M^{(n)}(0).$$

例6.15 設隨機變數 X 有 $\mathcal{E}(\lambda)$ 分佈, $\lambda > 0$. X 之 p.d.f. 為 $f(x) = \lambda e^{-\lambda x}$, $x > 0$. 則 X 之拉普拉斯轉換及動差母函數分別為

$$\begin{aligned} \phi(s) &= E(e^{-sX}) = \int_0^\infty e^{-sx} f(x) dx = \int_0^\infty e^{-sx} \lambda e^{-\lambda x} dx \\ &= \lambda \int_0^\infty e^{-(\lambda+s)x} dx = -\frac{\lambda}{\lambda+s} e^{-(\lambda+s)x} \Big|_0^\infty \\ &= \frac{\lambda}{\lambda+s}, s \geq 0. \end{aligned}$$

$$\begin{aligned}
M(t) &= E(e^{tX}) = \int_0^\infty e^{tx} f(x) dx = \int_0^\infty e^{tx} \lambda e^{-\lambda x} dx \\
&= \lambda \int_0^\infty e^{-(\lambda-t)x} dx = -\frac{\lambda}{\lambda-t} e^{-(\lambda-t)x} \Big|_0^\infty \\
&= \begin{cases} \lambda/(\lambda-t), & t < \lambda, \\ \infty & , t \geq \lambda. \end{cases}
\end{aligned}$$

即 $s \geq 0$ 時, $\phi(s)$ 皆存在, 但 $M(t)$ 在 $t \geq \lambda$ 時則不存在.

因

$$\begin{aligned}
\phi'(s) &= -\frac{\lambda}{(\lambda+s)^2}, \\
\phi''(s) &= \frac{2\lambda}{(\lambda+s)^3}.
\end{aligned}$$

故

$$\begin{aligned}
E(X) &= -\phi'(0) = \frac{1}{\lambda}, \\
E(X^2) &= \phi''(0) = \frac{2}{\lambda^2}.
\end{aligned}$$

由此又得

$$\text{Var}(X) = E(X^2) - (E(X))^2 = \frac{2}{\lambda^2} - \left(\frac{1}{\lambda}\right)^2 = \frac{1}{\lambda^2}.$$

另外, 若利用 $M(t)$ 亦可得到 $E(X)$ 及 $E(X^2)$:

$$M'(t) = \frac{\lambda}{(\lambda-t)^2}, \quad M''(t) = \frac{2\lambda}{(\lambda-t)^3}.$$

故

$$\begin{aligned}
E(X) &= M'(0) = \frac{1}{\lambda}, \\
E(X^2) &= M''(0) = \frac{2}{\lambda^2}.
\end{aligned}$$

結果相同.