

第三章

多維隨機變數

3.4 混合分佈

有時一隨機變數之參數,亦允許為隨機變數。我們先看一個例子。

例4.1 假設高雄市每天發生之交通事故的次數 Y 有 $\mathcal{P}(\lambda)$ 分佈。每次交通事故有 p 的機率會有死亡。令 X 表每天有死亡之交通事件。則 $X|Y$ 有 $\mathcal{B}(Y, p)$ 分佈, 而 Y 有 $\mathcal{P}(\lambda)$ 分佈。本模式即二項分佈試驗的次數亦為一隨機變數 Y 。給定 $Y = y$, 則 X 有 $\mathcal{B}(y, p)$ 分佈。至於 Y 則有 $\mathcal{P}(\lambda)$ 分佈。當一隨機

2 第三章 多維隨機變數

變數 X 之參數亦為隨機變數時，我們便說 X 有一**混合分佈**(mixture distribution)。

在本例中， X 之**非條件分佈**如何求？由假設，且利用第一章定理2.1，得

$$\begin{aligned}
 P(X = x) &= \sum_{y=0}^{\infty} P(X = x|Y = y)P(Y = y) \\
 &= \sum_{y=0}^{\infty} \binom{y}{x} p^x (1-p)^{y-x} \frac{e^{-\lambda} \lambda^y}{y!} \\
 &= \frac{(\lambda p)^x e^{-\lambda}}{x!} \sum_{y=x}^{\infty} \frac{((1-p)\lambda)^{y-x}}{(y-x)!} \\
 &= \frac{(\lambda p)^x e^{-\lambda}}{x!} \sum_{i=0}^{\infty} \frac{((1-p)\lambda)^i}{i!} \\
 &= \frac{(\lambda p)^x e^{-\lambda}}{x!} \cdot e^{\lambda(1-p)} = \frac{e^{-\lambda p} (\lambda p)^x}{x!}。
 \end{aligned}$$

因此 X 之非條件分佈為 $\mathcal{P}(\lambda p)$ 。

在上例中，每天死亡人數之期望值為 $E(X) = \lambda p$ 。又顯然 $E(X|Y) = Yp$ ， $E(Y) = \lambda$ ，故得 $E(X) = E(E(X|Y)) = E(Yp) = pE(Y) = \lambda p$ 。這就是**總機率法則**(Law of total probability)：條件期望值之期望值，等於原期望值。我們列成如下的定理。

定理4.1 對任二隨機變數 X, Y ，只要下述各期望值皆為有限，便有

$$(4.1) \quad E(X) = E(E(X|Y))。$$

證明。我們只證連續型隨機變數的情況。設 (X, Y) 之聯合p.d.f.為 $f(x, y)$ 。由定義

$$\begin{aligned}
 E(X) &= \int \int x f(x, y) dx dy \\
 &= \int \left(\int x f(x|y) dx \right) f_Y(y) dy \\
 &= \int E(X|y) f_Y(y) dy \\
 &= E(E(X|Y)),
 \end{aligned}$$

其中第一個等號右側的積分是在 $A = \{(x, y) | f(x, y) > 0\}$, 其餘積分也分別在適當的範圍內。

由例4.1或定理4.1知, 所謂混合分佈, 乃是分層來看隨機變數。有時這樣子描述一機率模式, 會較容易理解。像例4.1, 對每天發生的交通事故, 給它一個分佈。每次事故又有一定的機率會有死亡。這樣子分兩層來想, 似乎容易多了。有時還可分好幾層, 見下例。

例4.2 承例4.1。假設不同的城市, 每天發生交通事故的次數雖皆可以波松分佈當模式, 但參數可能不同。即設 $Y|\Lambda$ 有 $\mathcal{P}(\Lambda)$ 分佈, 又設 Λ 有 $\mathcal{E}(\beta)$ 分佈。則

$$\begin{aligned} E(X) &= E(E(X|Y)) = E(pY) \\ &= pE(Y) = pE(E(Y|\Lambda)) \\ &= pE(\Lambda) = p/\beta。 \end{aligned}$$

此一三層的模式, 亦可表為兩層的模式。我們來求 Y 之非條件分佈如下。對 $\forall y = 0, 1, 2, \dots$,

$$\begin{aligned} P(Y = y) &= \int_0^\infty f_{Y,\Lambda}(y, \lambda) d\lambda \\ &= \int_0^\infty f_{Y|\Lambda}(y|\lambda) f_\Lambda(\lambda) d\lambda \\ &= \int_0^\infty \frac{e^{-\lambda} \lambda^y}{y!} \beta e^{-\beta\lambda} d\lambda \\ &= \frac{\beta}{y!} \int_0^\infty \lambda^y e^{-(1+\beta)\lambda} d\lambda \\ &= \frac{\beta}{y!} \frac{\Gamma(y+1)}{(1+\beta)^{y+1}} \\ &= \frac{\beta}{1+\beta} \left(\frac{1}{1+\beta} \right)^y。 \end{aligned}$$

可看出 Y 有自0開始之幾何分佈, 每次成功的機率為 $\beta/(1+\beta)$ 。知道 Y 的分佈, 就可以不用管 Λ 了(也可以說就不知道有 Λ 存在)。但以三層的方式來描述此機率模式, 顯然讓人較容易理解。

4 第三章 多維隨機變數

大家是否注意到, 在上例中, 我們引進了一與以前不太一樣的模式, 即二隨機變數一為連續一為離散。不過倒不至於有問題。例如, 在定義 $f_{Y,\Lambda}(y, \lambda)$ 時, 只要記住在求機率或期望值時, 若是離散型的隨機變數便用加的, 連續型的隨機變數則用積分。

底下再給一例。

例4.3 設 $X|P = p$ 有 $Ber(p)$ 分佈, 而 P 有 $Be(\alpha, \beta)$ 分佈。則仿例4.2, 對 $x = 0, 1$, 得

$$\begin{aligned} P(X = x) &= \int_0^1 p^x (1-p)^{1-x} \frac{\Gamma(\alpha+\beta)}{\Gamma(\alpha)\Gamma(\beta)} p^{\alpha-1} (1-p)^{\beta-1} dp \\ &= \frac{\Gamma(\alpha+\beta)}{\Gamma(\alpha)\Gamma(\beta)} \int_0^1 p^{\alpha+x-1} (1-p)^{\beta-x+1-1} dp \\ &= \frac{\Gamma(\alpha+\beta)}{\Gamma(\alpha)\Gamma(\beta)} \frac{\Gamma(\alpha+x)\Gamma(\beta-x+1)}{\Gamma(\alpha+\beta+1)} \\ &= \frac{\Gamma(\alpha+x)\Gamma(\beta-x+1)}{(\alpha+\beta)\Gamma(\alpha)\Gamma(\beta)} \\ &= \begin{cases} \frac{\alpha}{\alpha+\beta}, & x = 1, \\ \frac{\beta}{\alpha+\beta}, & x = 0. \end{cases} \end{aligned}$$

可看出 X 之非條件分佈為 $Ber(\alpha/(\alpha+\beta))$, 即仍有伯努力分佈, 只是參數改變了: 由 P 變成 P 的期望值 $\alpha/(\alpha+\beta)$ 。由此即得

$$E(X) = \frac{\alpha}{\alpha+\beta}。$$

但若只是要求 $E(X)$, 則可不用求出 X 之非條件分佈, 而利用總機率法則:

$$E(X) = E(E(X|P)) = E(P) = \frac{\alpha}{\alpha+\beta},$$

所得相同。

總機率法則以分兩層來求期望值。對於求變異數則有下述定理。

定理4.2 對任二隨機變數 X, Y , 只要下述各期望值及變異數皆存在, 便有

$$(4.2) \quad \text{Var}(X) = E(\text{Var}(X|Y)) + \text{Var}(E(X|Y))。$$

證明.由定義

$$\begin{aligned}
 (4.3) \quad \text{Var}(X) &= E((X - E(X))^2) \\
 &= E((X - E(X|Y) + E(X|Y) - E(X))^2) \\
 &= E((X - E(X|Y))^2) + E((E(X|Y) - E(X))^2) \\
 &\quad + 2E((X - E(X|Y))(E(X|Y) - E(X)))。
 \end{aligned}$$

我們先看上式第三個等號右側最後一項，它必須是0，原因如下：

$$\begin{aligned}
 (4.4) \quad &E((X - E(X|Y))(E(X|Y) - E(X))) \\
 &= E(E((X - E(X|Y))(E(X|Y) - E(X))|Y))。
 \end{aligned}$$

但給定 Y 之下， $E(X|Y)$ 及 $E(X)$ 皆為常數。故

$$\begin{aligned}
 (4.5) \quad &E((X - E(X|Y))(E(X|Y) - E(X))|Y) \\
 &= (E(X|Y) - E(X))E(X - E(X|Y)|Y) \\
 &= (E(X|Y) - E(X))(E(X|Y) - E(X|Y)) \\
 &= 0。
 \end{aligned}$$

代入(4.4)式，即得

$$E((X - E(X|Y))(E(X|Y) - E(X))) = E(0) = 0。$$

回到(4.3)式，因

$$(4.6) \quad \text{Var}(X|Y) = E((X - E(X|Y))^2|Y),$$

上式也是 $\text{Var}(X|Y)$ 之定義，故

$$\begin{aligned}
 (4.7) \quad E((X - E(X|Y))^2) &= E(E((X - E(X|Y))^2|Y)) \\
 &= E(\text{Var}(X|Y))。
 \end{aligned}$$

又因 $E(E(X|Y)) = E(X)$ ，故

$$(4.8) \quad E((E(X|Y) - E(X))^2) = \text{Var}(E(X|Y))。$$

得證(4.2)式。

上述證明稍嫌冗長，底下的推導較簡潔些：

由(4.6)式，得

$$(4.9) \quad \text{Var}(X|Y) = E(X^2|Y) - (E(X|Y))^2。$$

因此

$$\begin{aligned} E(\text{Var}(X|Y)) &= E(E(X^2|Y) - (E(X|Y))^2) \\ &= E(E(X^2|Y)) - E((E(X|Y))^2) \\ &= E(X^2) - (E(X))^2 - (E((E(X|Y))^2) - (E(X))^2) \\ &= \text{Var}(X) - \text{Var}(E(X|Y))。 \end{aligned}$$

要注意的是，不要將總機率法則誤用在 $\text{Var}(X)$ 。即雖然 $E(X) = E(E(X|Y))$ ，但由(4.2)式，一般而言

$$\text{Var}(X) \neq E(\text{Var}(X|Y)),$$

除非 $\text{Var}(E(X|Y)) = 0$ 。

例4.4 承例4.3，欲求 X 的變異數。

首先，設未求出 X 之非條件分佈。由(4.2)式，

$$(4.10) \quad \text{Var}(X) = \text{Var}(E(X|P)) + E(\text{Var}(X|P))。$$

而因 $X|P$ 有 $\text{Ber}(P)$ 分佈，故

$$E(X|P) = P, \quad \text{Var}(X|P) = P(1 - P)。$$

因此

$$\begin{aligned} \text{Var}(X) &= \text{Var}(P) + E(P(1 - P)) \\ &= E(P^2) - (E(P))^2 + E(P) - E(P^2) \end{aligned}$$

$$\begin{aligned} &= E(P)(1 - E(P)) \\ &= \frac{\alpha}{\alpha + \beta} \left(1 - \frac{\alpha}{\alpha + \beta}\right) \\ &= \frac{\alpha\beta}{\alpha + \beta} \circ \end{aligned}$$

另外, 如果已求出 X 之非條件分佈為 $\mathcal{Ber}(\alpha/(\alpha + \beta))$, 則得到相同的 $\text{Var}(X)$ 。