

## 單元 35: 邊際及條件機率分布 (課本 §5.3)

定義. (a) 設聯合離散隨機變數

$$(Y_1, Y_2) \sim \text{joint pmf } p(y_1, y_2)$$

則  $Y_1$  的邊際機率質量函數 (marginal pmf)

$$\begin{aligned} p_1(y_1) &\stackrel{\text{def}}{=} P(Y_1 = y_1) \\ &= \sum_{y_2} P(Y_1 = y_1, Y_2 = y_2) \\ &= \sum_{y_2} p(y_1, y_2) \end{aligned}$$

其中累加和的範圍是涵蓋所有可能的  $y_2$ ; 同理,  $Y_2$  的 marginal pmf

$$\begin{aligned} p_2(y_2) &\stackrel{\text{def}}{=} P(Y_2 = y_2) \\ &= \sum_{y_1} P(Y_1 = y_1, Y_2 = y_2) \\ &= \sum_{y_1} p(y_1, y_2) \end{aligned}$$

(b) 設聯合連續隨機變數

$$(Y_1, Y_2) \sim \text{joint pdf } f(y_1, y_2)$$

則  $Y_1$  的邊際機率密度函數 (marginal pdf)

$$f_1(y_1) \stackrel{\text{def}}{=} \int_{-\infty}^{\infty} f(y_1, y_2) dy_2$$

且  $Y_2$  的 marginal pdf

$$f_2(y_2) \stackrel{\text{def}}{=} \int_{-\infty}^{\infty} f(y_1, y_2) dy_1$$

例 1. 由 3 位共和黨員, 2 位民主黨員, 與 1 位獨立黨員中, 任選 2 人組成一委員會. 令  $Y_1 =$  委員會中共和黨員數且  $Y_2 =$  委員會中民主黨員數. 試求  $Y_1$  與  $Y_2$  的 joint pmf, 並求  $Y_1$  的 marginal pmf.

<解> 根據多項分布,

$$\begin{aligned} p(1, 1) &= P(Y_1 = 1, Y_2 = 1) \\ &= \frac{\binom{3}{1} \binom{2}{1} \binom{1}{0}}{\binom{6}{2}} \\ &= \frac{3 \cdot 2}{15} = \frac{6}{15} \end{aligned}$$

同理, 可得其它的機率值如下述,

$$\begin{aligned} p(0, 0) &= P(\text{選出兩位獨立黨員}) \\ &= P(\emptyset) = 0 \end{aligned}$$

$$p(0, 1) = \frac{\binom{3}{0} \binom{2}{1} \binom{1}{1}}{\binom{6}{2}} = \frac{2}{15}$$

$$p(0, 2) = \frac{\binom{3}{0} \binom{2}{2} \binom{1}{0}}{\binom{6}{2}} = \frac{1}{15}$$

$$p(1, 0) = \frac{\binom{3}{1} \binom{2}{0} \binom{1}{1}}{\binom{6}{2}} = \frac{3}{15}$$

$$p(2, 0) = \frac{\binom{3}{2} \binom{2}{0} \binom{1}{0}}{\binom{6}{2}} = \frac{3}{15}$$

以及

$$p(1, 2) = p(2, 1) = p(2, 2) = P(\emptyset) = 0$$

此乃因為委員會只由 2 人組成, 並可列表如下:

$y_2$	$y_1$			Total
	0	1	2	
0	0	3/15	3/15	6/15
1	2/15	6/15	0	8/15
2	1/15	0	0	1/15
Total	3/15	9/15	3/15	

又根據邊際機率質量函數的定義,  $Y_1$  的 marginal pmf

如下:

$$\begin{aligned} p_1(0) &= p(0, 0) + p(0, 1) + p(0, 2) \\ &= 0 + \frac{2}{15} + \frac{1}{15} = \frac{3}{15} \end{aligned}$$

$$\begin{aligned} p_1(1) &= p(1, 0) + p(1, 1) + p(1, 2) \\ &= \frac{3}{15} + \frac{6}{15} + 0 = \frac{9}{15} \end{aligned}$$

且

$$\begin{aligned} p_1(2) &= p(2, 0) + p(2, 1) + p(2, 2) \\ &= \frac{3}{15} + 0 + 0 = \frac{3}{15} \end{aligned}$$

均為各行項的和. 因為置於表格的旁邊 (底邊), 故稱作邊際機率質量函數 (marginal pmf). 同理, 各列項的和為  $Y_2$  的 marginal pmf, 且位於表格的右邊.

例 2. 設聯合連續隨機變數  $(Y_1, Y_2)$  的 joint pdf

$$f(y_1, y_2) = \begin{cases} 2y_1, & 0 \leq y_1 \leq 1, 0 \leq y_2 \leq 1 \\ 0, & \text{其它} \end{cases}$$

試繪  $f(y_1, y_2)$  的圖形, 並求  $Y_1$  與  $Y_2$  的 marginal pdf.

<解>  $f(y_1, y_2)$  的圖形如下所示.

觀察:

- (1) 累積方向為“平行於  $y_2$ -軸”時,  $Y_1$  的 marginal pdf 為三角形分布, 如圖所示.
- (2) 累積方向為“平行於  $y_1$  軸”時,  $Y_2$  的 marginal pdf 為均勻分布, 如圖所示.

實際求法如下:

- (1)  $Y_1$  的 marginal pdf: 對於  $0 \leq y_1 \leq 1$ ,

$$\begin{aligned} f_1(y_1) &= \int_{-\infty}^{\infty} f(y_1, y_2) dy_2 \\ &= \int_0^1 2y_1 dy_2 \\ &= 2y_1 y_2 \Big|_{y_2=0}^1 = 2y_1 \end{aligned}$$

對於  $y_1 < 0$  或  $y_1 > 1$ ,

$$\begin{aligned} f_1(y_1) &= \int_{-\infty}^{\infty} f(y_1, y_2) dy_2 \\ &= \int_{-\infty}^{\infty} 0 dy_2 = 0 \end{aligned}$$

因此, 綜合上述,  $Y_1$  的 marginal pdf

$$f_1(y_1) = \begin{cases} 2y_1, & 0 \leq y_1 \leq 1 \\ 0, & \text{其它} \end{cases}$$

如圖所示.

(2)  $Y_2$  的 marginal pdf: 對於  $0 \leq y_2 \leq 1$ ,

$$\begin{aligned} f_2(y_2) &= \int_{-\infty}^{\infty} f(y_1, y_2) dy_1 \\ &= \int_0^1 2y_1 dy_1 \\ &= y_1^2 \Big|_{y_1=0}^1 = 1 \end{aligned}$$

另  $y_2 < 0$  或  $y_2 > 1$  時,

$$\begin{aligned} f_2(y_2) &= \int_{-\infty}^{\infty} f(y_1, y_2) dy_1 \\ &= \int_{-\infty}^{\infty} 0 dy_1 = 0 \end{aligned}$$

因此, 綜合上述, 得  $Y_2$  的 marginal pdf

$$f_2(y_2) = \begin{cases} 1, & 0 \leq y_2 \leq 1 \\ 0, & \text{其它} \end{cases}$$

如圖所示.

定義. 設聯合離散隨機變數

$$(Y_1, Y_2) \sim \text{joint pmf } p(y_1, y_2)$$

且  $Y_1$  與  $Y_2$  的 marginal pmf 分別為  $p_1(y_1)$  與  $p_2(y_2)$ . 則給定  $Y_2 = y_2$  下,  $Y_1$  的條件機率質量函數 (conditional pmf)

$$\begin{aligned} p(y_1|y_2) &\stackrel{\text{def}}{=} P(Y_1 = y_1|Y_2 = y_2) \\ &= \frac{P(Y_1 = y_1, Y_2 = y_2)}{P(Y_2 = y_2)} \\ &= \frac{p(y_1, y_2)}{p_2(y_2)} \end{aligned}$$

當  $p_2(y_2) > 0$  時.

註. 若  $p_2(y_2) = 0$  時, 則  $p(y_1|y_2)$  未定義.

同理, 給定  $Y_1 = y_1$  下,  $Y_2$  的 conditional pmf

$$\begin{aligned} p(y_2|y_1) &\stackrel{\text{def}}{=} P(Y_2 = y_2|Y_1 = y_1) \\ &= \frac{P(Y_1 = y_1, Y_2 = y_2)}{P(Y_1 = y_1)} \\ &= \frac{p(y_1, y_2)}{p_1(y_1)} \end{aligned}$$

當  $p_1(y_1) > 0$  時.

註. 若  $p_1(y_1) = 0$  時, 則  $p(y_2|y_1)$  未定義.

例 3. 承接例 1, 試求給定  $Y_2 = 1$  下,  $Y_1$  的條件機率質量函數 (conditional pmf), 亦即, 當委員會中有一人爲民主黨員時, 共和黨員數的條件分布.

<解> 由例 1, joint pmf  $p(y_1, y_2)$  及 marginal pmf  $p_1(y_1), p_2(y_2)$  如下表:

	$y_1$			
$y_2$	0	1	2	Total
0	0	3/15	3/15	6/15
1	2/15	6/15	0	8/15
2	1/15	0	0	1/15
Total	3/15	9/15	3/15	

接著, 求給定  $Y_2 = 1$ ,  $Y_1$  的 conditional pmf  $p(y_1|Y_2 = 1)$  時, 只需看 ( $Y_2 = 1$ ) 所對應的列即可, 如下述.

$$p(0|Y_2 = 1) = \frac{p(0, 1)}{p_2(1)} = \frac{2/15}{8/15} = \frac{1}{4}$$

以及

$$p(1|Y_2 = 1) = \frac{p(1, 1)}{p_2(1)} = \frac{6/15}{8/15} = \frac{3}{4}$$



和

$$p(2|Y_2 = 1) = \frac{p(2, 1)}{p_2(1)} = \frac{0}{8/15} = 0$$

因此, 給定  $Y_2 = 1$ ,  $Y_1$  的 conditional pmf

$$p(y_1|Y_2 = 1) = \begin{cases} \frac{1}{4}, & y_1 = 0 \\ \frac{3}{4}, & y_1 = 1 \end{cases}$$

同理, 給定  $Y_2 = 0$ ,  $Y_1$  的 conditional pmf

$$p(y_1|Y_2 = 0) = \begin{cases} \frac{1}{2}, & y_1 = 1 \\ \frac{1}{2}, & y_1 = 2 \end{cases}$$

以及給定  $Y_2 = 2$  時,  $Y_1$  的 conditional pmf

$$p(y_1|Y_2 = 2) = 1, \quad y_1 = 0$$

註 1. 固定給定的  $Y_2 = y_2$  時, conditional pmf  $p(y_1|y_2)$  本身為一可能值為  $y_1$  的隨機變數的 pmf.

註 2. 條件機率質量函數 (conditional pmf)  $p(y_1|y_2)$  隨著  $Y_2$  的觀察值而變.

註 3. 給定  $Y_1 = y_1$ ,  $Y_2$  的 conditional pmf 如下:

$$p(y_2|Y_1 = 0) = \begin{cases} \frac{2}{3}, & y_2 = 1 \\ \frac{1}{3}, & y_2 = 2 \end{cases}$$

$$p(y_2|Y_1 = 1) = \begin{cases} \frac{1}{3}, & y_2 = 0 \\ \frac{2}{3}, & y_2 = 1 \end{cases}$$

以及

$$p(y_2|Y_1 = 2) = 1, \quad y_2 = 0$$

亦反應出固定給定的  $Y_1 = y_1$ ,  $p(y_2|y_1)$  本身為一可能值為  $y_2$  的隨機變數的 pmf, 以及隨著  $Y_1$  的觀察值而變.

註 4. 若  $Y_1$  與  $Y_2$  為聯合連續隨機變數, 則無法定義  $P(Y_1 = y_1|Y_2 = y_2)$  或  $P(Y_2 = y_2|Y_1 = y_1)$ , 因為此時  $P(Y_2 = y_2)$  與  $P(Y_1 = y_1)$  均恆為 0. 但可透過如下定義的條件累積分布函數 (conditional cdf) 著手.

定義. 設聯合連續隨機變數

$$(Y_1, Y_2) \sim \text{joint pdf } f(y_1, y_2)$$

則給定  $Y_2 = y_2$ ,  $Y_1$  的 conditional cdf

$$F(y_1|y_2) \stackrel{\text{def}}{=} P(Y_1 \leq y_1|Y_2 = y_2)$$

註. 固定給定值  $Y_2 = y_2$ , conditional cdf  $F(y_1|y_2)$  為一  $y_1$  的函數, 且本身為一可能值為  $y_1$  的隨機變數的 cdf.

同理, 給定  $Y_1 = y_1$ ,  $Y_2$  的 conditional cdf

$$F(y_2|y_1) \stackrel{\text{def}}{=} P(Y_2 \leq y_2|Y_1 = y_1)$$

註. 固定給定的  $Y_1 = y_1$ ,  $F(y_2|y_1)$  本身為一可能值為  $y_2$  的隨機變數的 cdf.

註. 因為  $Y_2$  落在  $y_2$  附近小區間的機率近似於

$$f_2(y_2)dy_2$$

故由全機率律的概念,

$$\begin{aligned} F(y_1) &= P(Y_1 \leq y_1) \\ &= \int_{-\infty}^{\infty} P(Y_1 \leq y_1|Y_2 = y_2)f_2(y_2)dy_2 \\ &= \int_{-\infty}^{\infty} F(y_1|y_2)f_2(y_2)dy_2 \end{aligned} \quad (1)$$

又因為  $Y_1$  的 marginal pdf 為  $f_1(y_1)$ , 故

$$\begin{aligned} F(y_1) &= P(Y_1 \leq y_1) \\ &= \int_{-\infty}^{y_1} f_1(t_1)dt_1 \\ &= \int_{-\infty}^{y_1} \left[ \int_{-\infty}^{\infty} f(t_1, y_2)dy_2 \right] dt_1 \\ &= \int_{-\infty}^{\infty} \left[ \int_{-\infty}^{y_1} f(t_1, y_2)dt_1 \right] dy_2 \end{aligned} \quad (2)$$

比較 (1) 式與 (2) 式, 得

$$F(y_1|y_2)f_2(y_2) = \int_{-\infty}^{y_1} f(t_1, y_2)dt_1$$

或

$$F(y_1|y_2) = \int_{-\infty}^{y_1} \frac{f(t_1, y_2)}{f_2(y_2)}dt_1$$

所以, 由 cdf 與 pdf 之間的關係, 可稱上式中的被積函數

$$\frac{f(y_1, y_2)}{f_2(y_2)}$$

為給定  $Y_2 = y_2$  下,  $Y_1$  的 conditional pdf, 並以  $f(y_1|y_2)$  表示之. 正式定義如下.

定義. 設聯合連續隨機變數

$$(Y_1, Y_2) \sim \text{joint pdf } f(y_1, y_2)$$

且  $Y_1$  與  $Y_2$  的 marginal pdf 分別為  $f_1(y_1)$  與  $f_2(y_2)$ . 則對任意使得  $f_2(y_2) > 0$  的  $y_2$ , 給定  $Y_2 = y_2$ ,  $Y_1$  的 conditional pdf

$$f(y_1|y_2) \stackrel{\text{def}}{=} \frac{f(y_1, y_2)}{f_2(y_2)}$$

註. 固定給定的  $Y_2 = y_2$ ,  $f(y_1|y_2)$  本身為一可能值為  $y_1$  的隨機變數的 pdf. 又隨著  $Y_2$  的觀察值而變.

同理, 對任意使得  $f_1(y_1) > 0$  的  $y_1$ , 給定  $Y_1 = y_1$ ,  $Y_2$  的 conditional pdf

$$f(y_2|y_1) \stackrel{\text{def}}{=} \frac{f(y_1, y_2)}{f_1(y_1)}$$

註. 固定給定的  $Y_1 = y_1$ ,  $f(y_2|y_1)$  本身為一可能值為  $y_2$  的隨機變數的 pdf. 又隨著  $Y_1$  的觀察值而變.

註 1. 若  $f_2(y_2) = 0$ , 則  $f(y_1|y_2)$  未定義. 同理, 若  $f_1(y_1) = 0$ ,  $f(y_2|y_1)$  亦未定義.

註 2. 綜合 conditional cdf 與 conditional pdf 的定義, 得

$$F(y_2|y_1) = \int_{-\infty}^{y_2} f(t_2|y_1) dt_2$$

與

$$F(y_1|y_2) = \int_{-\infty}^{y_1} f(t_1|y_2) dt_1$$

例 4. 設一飲料機內含  $Y_2$  加侖的飲料且銷售量為  $Y_1$  加侖, 故  $Y_1 \leq Y_2$ . 又

$$(Y_1, Y_2) \sim f(y_1, y_2) = \begin{cases} \frac{1}{2}, & 0 \leq y_1 \leq y_2 \leq 2 \\ 0, & \text{其它} \end{cases}$$

試求 (a) 給定  $Y_2 = y_2$ ,  $Y_1$  的 conditional pdf, 以及 (b)  $P(\text{銷售量} \leq 1/2 | \text{機器內含1.5})$ .

<解> (a) 先求  $Y_2$  的 marginal pdf  $f_2(y_2)$ : 首先 joint pdf  $f(y_1, y_2)$  在區域

$$R : \begin{cases} 0 \leq y_2 \leq 2 \\ 0 \leq y_1 \leq y_2 \end{cases}$$

上取正值  $\frac{1}{2}$ , 其它地方取值為 0, 如圖示.

註. 因為  $R$  的面積為  $\frac{1}{2}(2)(2) = 2$ , 所以

$$(Y_1, Y_2) \sim \text{unif}(R)$$

接著, 當  $0 \leq y_2 \leq 2$  時,

$$\begin{aligned} f_2(y_2) &= \int_{-\infty}^{\infty} f(y_1, y_2) dy_1 \\ &= \int_0^{y_2} \frac{1}{2} dy_1 \\ &= \frac{1}{2} y_1 \Big|_{y_1=0}^{y_2} = \frac{1}{2} y_2 \end{aligned}$$

當  $y_2 < 0$  或  $y_2 > 2$  時,

$$\begin{aligned} f_2(y_2) &= \int_{-\infty}^{\infty} f(y_1, y_2) dy_1 \\ &= \int_{-\infty}^{\infty} 0 dy_1 = 0 \end{aligned}$$

因此, 綜合上述,  $Y_2$  的 marginal pdf

$$f_2(y_2) = \begin{cases} \frac{1}{2}y_2, & 0 \leq y_2 \leq 2 \\ 0, & \text{其它} \end{cases}$$

最後, 根據 conditional pdf 的定義, 對於  $0 < y_2 \leq 2$  (因為此時  $f_2(y_2) > 0$ ), 給定  $Y_2 = y_2$ ,  $Y_1$  的 conditional pdf

$$\begin{aligned} f(y_1|y_2) &= \frac{f(y_1, y_2)}{f_2(y_2)} \\ &= \begin{cases} \frac{1/2}{(1/2)y_2} = \frac{1}{y_2}, & 0 \leq y_1 \leq y_2 \\ \frac{0}{(1/2)y_2} = 0, & \text{其它} \end{cases} \quad (3) \end{aligned}$$

又  $y_2 \leq 0$  或  $y_2 > 2$  時,  $f(y_1|y_2)$  未定義, 因為此時  $f_2(y_2) = 0$ .

(b) 根據題意,

$$\begin{aligned} \text{所求} &= P(Y_1 \leq 1/2 | Y_2 = 1.5) \\ &= \int_{-\infty}^{1/2} f(y_1 | y_2 = 1.5) dy_1 \\ &= \int_0^{1/2} \frac{1}{1.5} dy_1 \\ &= \frac{1}{1.5} y_1 \Big|_0^{1/2} = \frac{0.5}{1.5} = \frac{1}{3} \end{aligned}$$

其中第二個等號成立乃根據註 2, 第三個等號成立乃根據 (3) 式, 得

$$f(y_1|y_2 = 1.5) = \frac{1}{1.5}, \quad 0 \leq y_1 \leq 1.5$$

所致.

註. 觀察不同  $Y_2$  值的條件機率, 如

$$\begin{aligned} P(Y_1 \leq 1/2|Y_2 = 2) &= \int_{-\infty}^{1/2} f(y_1|y_2 = 2)dy_1 \\ &= \int_0^{1/2} \frac{1}{2}dy_1 \\ &= \frac{1}{2}y_1 \Big|_{y_1=0}^{1/2} = \frac{1}{4} \end{aligned}$$

以及

$$\begin{aligned} P(Y_1 \leq 1/2|Y_2 = 1/3) &= \int_{-\infty}^{1/2} f(y_1|y_2 = 1/3)dy_1 \\ &= \int_0^{1/3} 3dy_1 \\ &= 3y_1 \Big|_0^{1/3} = 1 \end{aligned}$$

一合理的結果, 其中第二個等號成立乃因為

$$f(y_1|y_2 = 1/3) = 3, \quad 0 \leq y_1 \leq \frac{1}{3}$$



所致. 另外, 由此例亦反應出條件機率隨著  $Y_2$  的觀察值而變!!