

單元 43: 條件期望值

(課本 §5.11)

條件期望值的定義方式與單變量期望值相同，只不過是將邊際機率密度函數（或機率質量函數，marginal pdf，或 pmf）換成條件機率密度函數（或條件機率質量函數，conditional pdf，或 pmf），如下述。

定義．若雙變量隨機變數

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

則給定 $Y_2 = y_2$ ， $g(Y_1)$ 的條件期望值 (conditional expectation)

$$E[g(Y_1)|Y_2 = y_2] \stackrel{\text{def}}{=} \int_{-\infty}^{\infty} g(y_1)f(y_1|y_2)dy_1$$

其中 $f(y_1|y_2)$ 為給定 y_2 ， Y_1 的 conditional pdf，亦即，

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

且

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

為 Y_2 的 marginal pdf.

若雙變量隨機變數

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

則給定 $Y_2 = y_2$, $g(Y_1)$ 的條件期望值

$$E[g(Y_1)|Y_2 = y_2] \stackrel{\text{def}}{=} \sum_{y_1} g(y_1)p(y_1|y_2)$$

其中 $p(y_1|y_2)$ 為給定 y_2 , Y_1 的 conditional pmf, 亦即,

$$p(y_1|y_2) = \frac{p(y_1, y_2)}{p_2(y_2)}$$

且

$$p_2(y_2) = \sum_{y_1} p(y_1, y_2)$$

為 Y_2 的 marginal pmf.

例 1. 設 (Y_1, Y_2) 的 joint pdf

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

試求 $E(Y_1|Y_2 = 1.5)$.

<解> (1) 求 $f(y_1|y_2)$: 首先, joint pdf $f(y_1, y_2)$ 在區域

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

上取值 $\frac{1}{2}$, 其它地方取值 0, 故對於 $0 \leq y_2 \leq 2$,

$$\begin{aligned} f_2(y_2) &= \int_{-\infty}^{\infty} f(y_1, y_2) dy_1 \\ &= \int_0^{y_2} \frac{1}{2} dy_1 \\ &= \frac{1}{2} y_1 \Big|_0^{y_2} = \frac{1}{2} y_2 \end{aligned}$$

當 $y_2 < 0$ 或 $y_2 > 2$ 時,

$$f_2(y_2) = \int_{-\infty}^{\infty} 0 dy_1 = 0$$

所以, Y_2 的 marginal pdf

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

又 conditional pdf $f(y_1|y_2)$ 只定義在 $f_2(y_2) \neq 0$ 時, 故當 $0 < y_2 \leq 2$, 給定 $Y_2 = y_2$, Y_1 的 conditional pdf

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

當 $y_2 \leq 0$ 或 $y_2 > 2$ 時, 因為 $f_2(y_2) = 0$, 故 conditional pdf $f(y_1|y_2)$ 未定義.

(2) 由條件期望值的定義, 對任意的 $0 < y_2 \leq 2$,

$$\begin{aligned} E(Y_1|Y_2 = y_2) &= \int_{-\infty}^{\infty} y_1 f(y_1|y_2) dy_1 \\ &= \int_0^{y_2} y_1 \left(\frac{1}{y_2} \right) dy_1 \\ &= \frac{1}{y_2} \cdot \frac{1}{2} y_1^2 \Big|_{y_1=0}^{y_2} \\ &= \frac{1}{y_2} \cdot \frac{1}{2} y_2^2 = \frac{1}{2} y_2 \end{aligned}$$

其中第二個等號成立乃因為前一式中 conditional pdf $f(y_1|y_2)$ 僅在 $0 \leq y_1 \leq y_2$ 上取值 $\frac{1}{y_2}$, 其它地方的取值為 0 所致. 因此, 給定 $Y_2 = 1.5$ 時, Y_1 的條件期望值

$$\begin{aligned} E(Y_1|Y_2 = 1.5) &= E(Y_1|Y_2 = y_2) \Big|_{y_2=1.5} \\ &= \frac{1}{2} y_2 \Big|_{y_2=1.5} \\ &= \frac{1.5}{2} = .75 \end{aligned}$$

註 1. 給定 $Y_2 = y_2$, $E(Y_1|Y_2 = y_2)$ 本身就是一個 y_2 的函數, 如上例的

$$E(Y_1|Y_2 = y_2) = \frac{1}{2} y_2$$

因此, 考慮 Y_2 的所有可能值時, 可視條件期望值

$$E(Y_1|Y_2)$$

為隨機變數 Y_2 的函數, 如上例的

$$E(Y_1|Y_2) = \frac{1}{2}Y_2$$

故 $E(Y_1|Y_2)$ 亦為一隨機變數, 其期望值與變異數如下述的二定理.

定理 5.14. 令 Y_1 與 Y_2 為二隨機變數, 則

$$E(Y_1) = E[E(Y_1|Y_2)]$$

亦即, 給定任一固定的隨機變數 Y_2 時, Y_1 的條件期望值的期望值等於 Y_1 的 (無條件) 期望值.

註. 等號右邊內層的期望值是給定 Y_2 , 針對 Y_1 的條件分布求 Y_1 的期望值, 亦即, 先求

$$E(Y_1|Y_2 = y_2) = \int_{-\infty}^{\infty} y_1 f(y_1|y_2) dy_1$$

得一 y_2 的函數, 再將 y_2 換成 Y_2 . 外層的期望值是針對 Y_2 的分布求隨機變數 $E(Y_1|Y_2)$ (一 Y_2 的函數) 的期望值, 亦即,

$$E[E(Y_1|Y_2)] = \int_{-\infty}^{\infty} E(Y_1|Y_2 = y_2) f_2(y_2) dy_2$$

<證> 設 (Y_1, Y_2) 的 joint pdf 爲 $f(y_1, y_2)$, 且 Y_1 與 Y_2 的 marginal pdf 分別爲 $f_1(y_1)$ 與 $f_2(y_2)$. 則根據期望值的定義以及條件 pdf 的定義所導出的恆等式

$$f(y_1, y_2) = f(y_1|y_2)f_2(y_2)$$

得

$$\begin{aligned} E(Y_1) &= \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} y_1 f(y_1, y_2) dy_1 dy_2 \\ &= \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} y_1 f(y_1|y_2) f_2(y_2) dy_1 dy_2 \end{aligned}$$

再根據 Fubini 定理, 將上式的雙變數積分 (double integral) 改爲先對 y_1 積分, 再對 y_2 積分的重積分 (iterated integral), 得上式相當於

$$E(Y_1) = \int_{-\infty}^{\infty} f_2(y_2) \left[\int_{-\infty}^{\infty} y_1 f(y_1|y_2) dy_1 \right] dy_2$$

接著, 根據條件期望值的定義, 上式的中括號部分爲給定 $Y_2 = y_2$ 下, Y_1 的條件期望值, 故得

$$E(Y_1) = \int_{-\infty}^{\infty} E(Y_1|Y_2 = y_2) f_2(y_2) dy_2$$

最後, 根據隨機變數函數的期望值公式, 上式的被積函數爲函數值 $E(Y_1|Y_2 = y_2)$ 與產生此值的 pdf $f_2(y_2)$ 的乘積, 因而等於 $E(Y_1|Y_2)$ 的期望值, 亦即,

$$E(Y_1) = E[E(Y_1|Y_2)]$$

例 2. 設一條組裝線的品管計劃為, 每天隨機抽樣出 $n = 10$ 個產品並計數瑕疵的產品數 Y , 若每天的產量是大的 (亦即, 遠大於 $n = 10$), 且令 $p =$ 觀察到瑕疵的機率 (亦即, 瑕疵率), 則

$$Y \sim b(n, p)$$

但 p 不是固定的值卻是每天在 $[0, 1/4]$ 間均勻地變動, 亦即,

$$p \sim \text{unif}\left(0, \frac{1}{4}\right)$$

試求 Y 的期望值.

<解> 根據題意, 在已知瑕疵率 p 的條件下,

$$Y \sim b(n, p)$$

故

$$E(Y|p) = np$$

因此, 由條件期望值的角度 (給定 p 之下), 根據定理 5.14, 並將上式代入, 可得 Y 的無條件期望值

$$\begin{aligned} E(Y) &= E[E(Y|p)] \\ &= E(np) \\ &= nE(p) \\ &= n \cdot \frac{0 + 1/4}{2} = \frac{n}{8} \end{aligned}$$

其中倒數第二個等號成立乃因爲

$$p \sim \text{unif}\left(0, \frac{1}{4}\right)$$

所致. 當 $n = 10$ 時,

$$E(Y) = \frac{10}{8} = 1.25$$

所以在長期觀察下, 每天平均會找出 1.25 個瑕疵品.

註 2. 給定 $Y_2 = y_2$, Y_1 的條件變異數可以類似方式定義, 亦即, 將 Y_1 的 marginal pdf (或 pmf) 換成 Y_1 的 conditional pdf (或 pmf), 如下述:

$$\begin{aligned} \text{Var}(Y_1|Y_2 = y_2) \\ = E(Y_1^2|Y_2 = y_2) - [E(Y_1|Y_2 = y_2)]^2 \end{aligned}$$

亦即, “條件二階動差” 減掉 “條件一階動差的平方”.

註 3. 給定 $Y_2 = y_2$, $\text{Var}(Y_1|Y_2 = y_2)$ 本身就是一個 y_2 的函數. 因此, 考慮 Y_2 的所有可能值時, 可視 $\text{Var}(Y_1|Y_2)$ 爲隨機變數 Y_2 的函數, 亦即, 若

$$g(y_2) = \text{Var}(Y_1|Y_2 = y_2)$$

則

$$\text{Var}(Y_1|Y_2) = g(Y_2)$$

故其本身亦是一隨機變數，而且其期望值

$$E[\text{Var}(Y_1|Y_2)]$$

可用於計算（無條件） $\text{Var}(Y_1)$ ，如下定理所述。

定理 5.15. 令 Y_1 與 Y_2 為二隨機變數，則

$$\text{Var}(Y_1) = E[\text{Var}(Y_1|Y_2)] + \text{Var}[E(Y_1|Y_2)] \quad (1)$$

<證> 首先，由註 2 與註 3 知，

$$\text{Var}(Y_1|Y_2) = E(Y_1^2|Y_2) - [E(Y_1|Y_2)]^2$$

因此，

$$\begin{aligned} E[\text{Var}(Y_1|Y_2)] \\ = E[E(Y_1^2|Y_2)] - E\{[E(Y_1|Y_2)]^2\} \end{aligned} \quad (2)$$

接著，因為 $E(Y_1|Y_2)$ 為 Y_2 的函數，故可視其為一隨機變數，並由變異數的定義，得

$$\begin{aligned} \text{Var}[E(Y_1|Y_2)] \\ = E\{[E(Y_1|Y_2)]^2\} - \{E[E(Y_1|Y_2)]\}^2 \end{aligned} \quad (3)$$

最後，將 (2) 式與 (3) 式相加，得

$$\begin{aligned} E[\text{Var}(Y_1|Y_2)] + \text{Var}[E(Y_1|Y_2)] \\ = E[E(Y_1^2|Y_2)] - \{E[E(Y_1|Y_2)]\}^2 \\ = E(Y_1^2) - [E(Y_1)]^2 = \text{Var}(Y_1) \end{aligned}$$

得證，其中第三個等號成立乃根據變異數的定義；第二個等號成立乃因為以 Y_2 條件化 (conditioning on Y_2) 後，將定理 5.14 簡單的推廣，亦即，將定理及證明中的 Y_1 換成 Y_1^2 以及 y_1 換成 y_1^2 ，可得

$$E[E(Y_1^2|Y_2)] = E(Y_1^2)$$

剛好就是前一式中等號右邊的第一項，以及直接根據定理 5.14，可得等號右邊第二項大括號內的

$$E[E(Y_1|Y_2)] = E(Y_1)$$

所致。

註 4. 因為

$$\text{Var}(Y_1|Y_2) \geq 0$$

故

$$E[\text{Var}(Y_1|Y_2)] \geq 0$$

因此，由 (1) 式，得

$$\text{Var}[E(Y_1|Y_2)] \leq \text{Var}(Y_1)$$

此不等式的重要意義如下述： Y_1 的條件期望值 $E(Y_1|Y_2)$ (一 Y_2 的函數，故為一隨機變數) 與隨機變數 Y_1 有相同的期望值，亦即，

$$E[E(Y_1|Y_2)] = E(Y_1)$$

但隨機變數 $E[Y_1|Y_2]$ 卻有較小的變異數, 亦即,

$$\text{Var}[E(Y_1|Y_2)] \leq \text{Var}(Y_1)$$

故以給定訊息下的隨機變數 $E(Y_1|Y_2)$ 描述隨機現象比直接以隨機變數 Y_1 描述來得穩定.

例 3. 承接例 2, 其中給定 p 時,

$$Y \sim b(n, p)$$

且

$$p \sim \text{unif}\left(0, \frac{1}{4}\right)$$

試求 $\text{Var}(Y)$.

<解> 根據題意, 給定 p 時,

$$Y \sim b(n, p)$$

且以

$$Y|p \sim b(n, p)$$

表示之, 則條件期望值

$$E(Y|p) = np$$

且條件變異數

$$\text{Var}(Y|p) = np(1 - p)$$

故由定理 5.15, 得

$$\begin{aligned}\text{Var}(Y) &= E[\text{Var}(Y|p)] + \text{Var}[E(Y|p)] \\ &= E[np(1-p)] + \text{Var}(np)\end{aligned}$$

接著, 根據線性組合的期望值和變異數公式, 上式相當於

$$\begin{aligned}\text{Var}(Y) &= nE[p(1-p)] + n^2\text{Var}(p) \\ &= n[E(p) - E(p^2)] + n^2\text{Var}(p) \quad (4)\end{aligned}$$

最後, 當

$$p \sim \text{unif}\left(0, \frac{1}{4}\right)$$

時, p 的期望值

$$E(p) = \frac{0 + 1/4}{2} = \frac{1}{8}$$

變異數

$$\text{Var}(p) = \frac{1}{12} \left(\frac{1}{4} - 0\right)^2 = \frac{1}{192}$$

以及二階動差

$$\begin{aligned}E(p^2) &= \text{Var}(p) + [E(p)]^2 \\ &= \frac{1}{192} + \frac{1}{64} \\ &= \frac{1+3}{192} = \frac{1}{48}\end{aligned}$$

並將它們代入 (4) 式, 得

$$\begin{aligned}\text{Var}(Y) &= n \left[\frac{1}{8} - \frac{1}{48} \right] + n^2 \cdot \frac{1}{192} \\ &= \frac{5n}{48} + \frac{n^2}{192}\end{aligned}$$

因此, 當 $n = 10$ 時,

$$\text{Var}(Y) = \frac{50}{48} + \frac{100}{192} = \frac{300}{192} = \frac{25}{16}$$