單元 21: 控制變量的運用
(課本 §8.2)

令模擬的輸出資料爲 X, 但

$$\theta = E(X)$$

未知, 爲欲估計的參數.

假設可模擬另一輸出資料 Y 且

$$E(Y) = \mu_Y$$

爲已知.

問. 如何根據 X 與 Y 生成另一輸出資料滿足此輸出資料的

$$\text{期望值} = \theta$$

亦即, 不偏, 且

$$\text{變異數} < \mathsf{Var}(X)$$

亦即, 有較小的變異數.

答. 可採用下述的控制變量法 (control variates approach).

令 c 爲任一實數, 考慮隨機變數

$$X + c(Y - \mu_Y)$$

則

$$\begin{aligned} E[X + c(Y - \mu_Y)] &= E(X) + cE(Y - \mu_Y) \\ &= \theta + c \cdot 0 = \theta \end{aligned} \tag{1}$$

亦即, 不偏, 且

$$\begin{aligned} &\mathrm{Var}[X + c(Y - \mu_Y)] \\ &= \mathrm{Var}(X + cY) \\ &= \mathrm{Var}(X) + \mathrm{Var}(cY) + 2\mathrm{Cov}(X, cY) \\ &= \mathrm{Var}(X) + c^2\mathrm{Var}(Y) + 2c\mathrm{Cov}(X, Y) \end{aligned} \tag{2}$$

乃一 c 的函數,

令

$$f(c) \stackrel{\text{def}}{=} \mathrm{Var}(X) + c^2\mathrm{Var}(Y) + 2c\mathrm{Cov}(X, Y)$$

目的: 找 c 最小化 $f(c)$, 亦即, 最小化

$$\mathrm{Var}[X + c(Y - \mu_Y)]$$

因爲

$$f'(c) = 2c\mathrm{Var}(Y) + 2\mathrm{Cov}(X, Y)$$

故由 $f'(c) = 0$, 得臨界數

$$c = -\frac{\mathrm{Cov}(X,Y)}{\mathrm{Var}(Y)}$$

又

$$f''(c) = 2\mathrm{Var}(Y)$$

恆正. 所以, $f(c)$ 在

$$c^* \stackrel{\text{def}}{=} -\frac{\mathrm{Cov}(X,Y)}{\mathrm{Var}(Y)}$$

有最小值.

或因爲 $f(c)$ 爲一抛物線, 且 c^2 的係數

$$\mathrm{Var}(Y) > 0$$

故開口向上, 而在

$$c^* \stackrel{\text{def}}{=} -\frac{\mathrm{Cov}(X,Y)}{\mathrm{Var}(Y)}$$

有最小值.

因此, 取 $c = c^*$ 時, 由 (2) 式, 得最小值

$$\begin{aligned}
&\mathrm{Var}[X + c^*(Y - \mu_Y)] \\
&= \mathrm{Var}(X) + \frac{[\mathrm{Cov}(X,Y)]^2}{\mathrm{Var}(Y)} - 2\frac{[\mathrm{Cov}(X,Y)]^2}{\mathrm{Var}(Y)} \\
&= \mathrm{Var}(X) - \frac{[\mathrm{Cov}(X,Y)]^2}{\mathrm{Var}(Y)} < \mathrm{Var}(X) \qquad (3)
\end{aligned}$$

最後, 綜合 (1) 式與 (3) 式, 得隨機變數

$$X + c^*(Y - \mu_Y)$$

的期望值為 θ, 亦即, 不偏, 且變異數小於 Var(X), 並且在這種對於任一實數 c,

$$X + c(Y - \mu_Y)$$

的組合中有最小的變異數.

定義. (i) 上述中的 Y 稱為控制變量 (control variate).

(ii) 這種以

$$X + c^*(Y - \mu_Y)$$

為輸出資料的模擬法稱為控制變量法 (control variate approach).

(iii) θ 的控制變量估計量 (control variate estimator, 或控制估計量) 為

$$\overline{X} + c^*(\overline{Y} - \mu_Y)$$

亦即,

$$\frac{1}{n}\sum_{i=1}^{n}[X_i + c^*(Y_i - \mu_Y)] \tag{4}$$

其中

$$X_1, \ldots, X_n$$

爲輸出資料 X 的 n 個模擬值, 且

$$Y_1, \ldots, Y_n$$

爲輸出資料 Y 的 n 個模擬值.

註. 由 (3) 及 (4) 式, 得

$$\begin{aligned}
&\text{Var}[\overline{X} + c^*(Y - \mu_Y)] \\
&\qquad = \frac{1}{n}\left\{\text{Var}(X) - \frac{[\text{Cov}(X, Y)]^2}{\text{Var}(Y)}\right\}
\end{aligned}$$

討論 1. 控制變量的基本概念如下述.

情況 1. X 與 Y 爲正相關.

根據正相關的定義,

$$\text{Cov}(X, Y) > 0$$

且

$$c^* = -\frac{\text{Cov}(X, Y)}{\text{Var(Y)}} < 0$$

另外, 若

$$Y \text{ 的觀察值} > (<) \ \mu_Y$$

由正相關知, 也可能

$$X \text{ 的觀察值} > (<) \ \theta$$

故一個合理的修正為以較小 (大) 的觀察值

$$X + c^*(Y - \mu_Y)$$

估計 θ.

為何會較小 (大)? 因為

$$c^*(Y - \mu_Y) = \begin{cases} (-)(+) < 0 \\ (-)(-) > 0 \end{cases}$$

所致.

情況 2. X 與 Y 為負相關.

同理, 亦成立, 自行驗證.

討論 2. 即使

$$\mathrm{Var}(Y)$$

已知, 但通常

$$\text{Cov}(X, Y)$$

未知, 故

$$c^* = -\frac{\text{Cov}(X, Y)}{\text{Var}(Y)}$$

在未模擬前是未知. 所以, 可估計 c^* 如下:

(i) 生成 n 組的輸出資料

$$(X_1, Y_1), \ldots, (X_n, Y_n)$$

(ii) 以

$$\overline{\text{Cov}(X, Y)} \stackrel{\text{def}}{=} \frac{1}{n-1} \sum_{i=1}^{n} (X_i - \overline{X})(Y_i - \overline{Y})$$

稱作 X 與 Y 的樣本共變異數, 估計

$$\text{Cov}(X, Y) = E[(X - \mu_X)(Y - \mu_Y)]$$

(iii) 若 $\text{Var}(Y)$ 未知, 則以

$$S^2 = \frac{1}{n-1} \sum_{i=1}^{n} (Y_i - \overline{Y})^2$$

估計 Var(Y).

(iv) 以

$$\widehat{c}^* = \begin{cases} -\dfrac{\sum_{i=1}^{n}(X_i-\overline{X})(Y_i-\overline{Y})}{\sum_{i=1}^{n}(Y_i-\overline{Y})^2}, & \text{若 Var}(X,Y)\text{ 未知} \\ -\dfrac{\overline{\text{Cov}(X,Y)}}{\text{Var}(Y)}, & \text{若 Var}(X,Y)\text{ 已知} \end{cases}$$

估計 c^*.

最後, 以

$$\overline{X} + \widehat{c}^*(\overline{Y} - \mu_Y)$$

估計 $\theta = E(X)$.

討論 3. 由 (3) 式知, 變異數降低率為

$$\frac{\text{Var}(X) - \text{Var}[X + c^*(Y - \mu_Y)]}{\text{Var}(X)} = \frac{[\text{Cov}(X,Y)]^2}{\text{Var}(X)\text{Var}(Y)} = \text{corr}^2(X,Y) \times 100\%$$

其中

$$\text{corr}(X, Y) = \frac{\text{Cov}(X, Y)}{\sqrt{\text{Var}(X)}\sqrt{\text{Var}(Y)}}$$

爲 X 與 Y 的相關係數, 亦即, X 與 Y 愈相關, 則變異數降低率愈大.

例 1. 試以控制變量法模擬穩定性函數

$$r(p_1, \ldots, p_n) = E[\phi(S_1, \ldots, S_n)]$$

其中

$$p_1, \ldots, p_n$$

已知, 且對於 $1 \le i \le n$, 第 i 個零件的狀態

$$S_i = \begin{cases} 1, & \text{若 } U_i < p_i, \\ 0, & \text{其它.} \end{cases}$$

<解> (i) 找控制變量 Y. 因爲穩定性函數與正常零件數有關, 且對於 $1 \le i \le n$, 由

$$E(S_i) = P(U_i < p_i) = p_i$$

已知, 得

$$E\left(\sum_{i=1}^{n} S_i\right) = \sum_{i=1}^{n} E(S_i) = \sum_{i=1}^{n} p_i$$

已知, 故根據上述的討論 3, 可考慮控制變量

$$Y \stackrel{\text{def}}{=} \sum_{i=1}^{n} S_i$$

亦即, 正常零件的總數.

(ii) 估計

$$c^* = -\frac{\text{Cov}(X, Y)}{\text{Var}(Y)}$$

首先, 根據零件間的獨立性以及各自的分布, 得

$$\begin{aligned}
\text{Var}(Y) &= \text{Var}\left(\sum_{i=1}^{n} S_i\right) \\
&= \sum_{i=1}^{n} \text{Var}(S_i) \\
&= \sum_{i=1}^{n} p_i(1 - p_i)
\end{aligned}$$

已知, 其中最後一個等號成立乃因爲對於 $1 \le i \le n$,

$$P(S_i = 1) = p_i$$

亦即, S_i 爲 Bernoulli 隨機變數, 故

$$\text{var}(S_i) = p_i(1 - p_i)$$

所致.

接著, 模擬 r 組資料並估計 Cov(X, Y), 如下述.

生成第一組資料

$$(S_1^1, \ldots, S_n^1)$$

並求出

$$X_1 = \phi(S_1^1, \ldots, S_n^1)$$

以及

$$Y_1 = \sum_{i=1}^{n} S_i^1$$

生成第二組資料

$$S_1^2, \ldots, S_n^2$$

並求出

$$X_2 = \phi(S_1^2, \ldots, S_n^2)$$

以及

$$Y_2 = \sum_{i=1}^{n} S_i^2$$

$$\vdots$$

生成第 r 組資料

$$S_1^r, \ldots, S_n^r$$

並求出

$$X_r = \phi(S_1^r, \ldots, S_n^r)$$

以及

$$Y_r = \sum_{i=1}^{n} S_i^r$$

由此得樣本共變異數

$$\overline{\mathrm{Cov}(X, Y)} = \frac{1}{r-1} \sum_{i=1}^{r} (X_i - \overline{X})(Y_i - \overline{Y})$$

並以此估計 $\mathrm{Cov}(X, Y)$.

最後, 以

$$\begin{aligned} \widehat{c}^* &= -\frac{\overline{\mathrm{Cov}(X, Y)}}{\mathrm{Var}(Y)} \\ &= -\frac{\dfrac{1}{r-1} \displaystyle\sum_{i=1}^{r} (X_i - \overline{X})(Y_i - \overline{Y})}{\displaystyle\sum_{i=1}^{n} p_i(1 - p_i)} \end{aligned}$$

估計 c^*.

(iii) 以

$$\overline{X}+\widehat{c}^*\left(\overline{Y}-\sum_{i=1}^{n}p_i\right)$$

估計穩定性函數

$$r(p_1,\ldots,p_n)$$

例 2. 在一佇列系統中, 假設

(1) 顧客的到達時間 $\sim$ NPP$[\lambda(s)]$, $s>0$.

(2) 對於 $1\le i\le n$, S_i 爲第 i 個顧客接受服務的時間, 且

$$S_1,\ldots,S_n\overset{\text{iid}}{\sim}G$$

並與顧客的到達時間相互獨立.

對於 $1\le i\le n$, 令 W_i 爲第 i 個顧客停留在系統中的時間, 且 $N(t)$ 爲在時間區間 $[0,t]$ 內進入系統的顧客總數. 試以控制變量法模擬

$$\theta=E(X)$$

其中

$$X = \sum_{i=1}^{N(t)} W_i$$

亦即, 在時間 t 之前, 所有顧客停留時間的總合.

註. 根據 NPP 的性質,

$$N(t) \sim \text{Poisson}\left(\int_0^t \lambda(s)ds\right)$$

故 X 爲一<u>隨機項和</u>所形成的隨機變數.

<解> (i) 找控制變量 Y. 因爲第 i 個顧客停留在系統中的時間 W_i 與其接受服務的時間 S_i 有相關性, 且 S_i 的分布 G 爲已知, 故可考慮

$$Y \stackrel{\text{def}}{=} \sum_{i=1}^{N(t)} S_i$$

接著, 根據 Y 的定義以及以 $N(t)$ 條件化,

$$\begin{aligned} E(Y) &= \sum_{k=1}^{\infty} E(Y|N(t)=k)P(N(t)=k) \\ &= \sum_{k=1}^{\infty} E\left(\left.\sum_{i=1}^{N(t)} S_i \right| N(t)=k\right) \cdot \\ &\qquad\qquad P(N(t)=k) \quad (5) \end{aligned}$$

其中以 $N(t)$ 條件化的目的乃在於當 $N(t)$ 爲給定的 k 值後, 形成 Y 的隨機項和就成爲有限的 k 項和, 而可根據期望值的線性性質繼續計算; 另外, $N(t)$ 爲卜松分布, $k=0$ 亦爲可能值, 而此處卻由 $k=1$ 開始條件化, 乃因爲當 $k=0$ 時, 形成 Y 的項數爲 0 項, 故 Y 爲 0, 不影響期望值的計算結果, 可不予考慮.

在給定 $N(t)=k$, 且 $N(t)$ 乃由顧客的到達時間所決定, 而顧客接受的服務時間 S_i 又與到達時間相互獨立, 得 $N(t)$ 與 S_i 互爲獨立, 故由 (5) 式, 得

$$
\begin{aligned}
E(Y) &= \sum_{k=1}^{\infty} E\left(\sum_{i=1}^{k} S_i \middle| N(t)=k\right) \cdot \\
&= \qquad\qquad\qquad P(N(t)=k) \\
&= \sum_{k=1}^{\infty} E\left(\sum_{i=1}^{k} S_i\right) P(N(t)=k)
\end{aligned}
$$

再根據 S_i 間的同分布, 由上式得

$$
\begin{aligned}
E(Y) &= \sum_{k=1}^{\infty} kE(S)P(N(t)=k) \\
&= E(S)\sum_{k=1}^{\infty} kP(N(t)=k) \\
&= E(S)E[N(t)] \qquad (6)
\end{aligned}
$$

其中

$$S \sim G$$

由假設條件, $E(S)$ 爲已知, 且

$$E[N(t)] = \int_0^t \lambda(s)ds$$

亦爲已知, 故可明確地計算出 $E(Y)$.

因此, 可取

$$Y = \sum_{i=1}^{N(t)} S_i$$

爲控制變量.

(ii) 估計

$$c^* = -\frac{\mathrm{Cov}(X,Y)}{\mathrm{Var}(Y)}$$

首先, 根據上述求 $E(Y)$ 的條件化過程, 得

$$\begin{aligned} E(Y^2) &= \sum_{k=1}^{\infty} E(Y^2|N(t)=k)P(N(t)=k) \\ &= \sum_{k=1}^{\infty} E\left[\left.\left(\sum_{i=1}^{N(t)} S_i\right)^2 \right| N(t)=k\right] \cdot \\ &\qquad P(N(t)=k) \end{aligned}$$

在給定 $N(t)=k$, 且 $N(t)$ 與 S_i 相互獨立下, 由上式得

$$
\begin{aligned}
E(Y^2) &= \sum_{k=1}^{\infty} E\left[\left(\sum_{i=1}^{k} S_i\right)^2 \middle| N(t)=k\right] \cdot \\
&\qquad P(N(t)=k) \\
&= \sum_{k=1}^{\infty} E\left[\left(\sum_{i=1}^{k} S_i\right)^2\right] P(N(t)=k) \\
&= \sum_{k=1}^{\infty} E\left(\sum_{i=1}^{k} S_i^2 + \sum_{1\le i\ne j\le k} S_i S_j\right) \cdot \\
&\qquad P(N(t)=k)
\end{aligned}
$$

其中最後一個等號乃是展開第二個等號右邊平方項的結果.

再根據 S_i 間的獨立同分布以及期望值的線性性質, 由上式得

$$
\begin{aligned}
E(Y^2) &= \sum_{k=1}^{\infty} \left\{kE(S^2) + k(k-1)[E(S)]^2\right\} \cdot \\
&\qquad P(N(t)=k) \\
&= E(S^2) \sum_{k=1}^{\infty} kP(N(t)=k) + \\
&\qquad [E(S)]^2 \sum_{k=1}^{\infty} k(k-1)P(N(t)=k)
\end{aligned}
$$

最後, 根據期望值的定義, 由上式得

$$\begin{aligned} E(Y^2) &= E(S^2)E[N(t)] + \\ &\qquad [E(S)]^2 E[N(t)(N(t)-1)] \quad (7) \end{aligned}$$

因此, 由變異數的公式以及 (6) 式與 (7) 式, 並化簡, 得

$$\begin{aligned} &\mathrm{Var}(Y) \\ &= E(Y^2) - [E(Y)]^2 \\ &= E(S^2)E[N(t)] + [E(S)]^2 E\{[N(t)]^2\} - \\ &\qquad [E(S)]^2 E[N(t)] - [E(S)]^2\{E[N(t)]\}^2 \\ &= \{E(S^2) - [E(S)]^2\}E[N(t)] + \\ &\qquad [E(S)]^2(E\{[N(t)]^2\} - \{E[N(t)]\}^2) \\ &= \mathrm{Var}(S)E[N(t)] + [E(S)]^2\mathrm{Var}[N(t)] \quad (8) \end{aligned}$$

又因爲

$$N(t) \sim \text{Poisson}\left(\int_0^t \lambda(s)ds\right)$$

故

$$E[N(t)] = \mathrm{Var}[N(t)]$$

並由 (8) 式以及變異數的公式, 得

$$\begin{aligned} \mathrm{Var}(Y) &= \{\mathrm{Var}(S) + [E(S)]^2\}E[N(t)] \\ &= E(S^2)E[N(t)] \quad (9) \end{aligned}$$

可由假設條件求得.

另外, 需模擬 r 組輸出資料並估計 Cov(X, Y), 如下述.

生成第一組資料

$$(W_1^1, \ldots, W_{N_1(t)}^1, S_1^1, \ldots, S_{N_1(t)}^1)$$

並計算

$$X_1 = \sum_{i=1}^{N_1(t)} W_i^1$$

以及

$$Y_1 = \sum_{i=1}^{N_1(t)} S_i^1$$

生成第二組資料

$$(W_1^2, \ldots, W_{N_2(t)}^2, S_1^2, \ldots, S_{N_2(t)}^2)$$

並計算

$$X_2 = \sum_{i=1}^{N_2(t)} W_i^2$$

以及

$$Y_2 = \sum_{i=1}^{N_2(t)} S_i^2$$

$\vdots$

生成第 r 組資料

$$(W_1^r, \ldots, W_{N_r(t)}^r, S_1^r, \ldots, S_{N_r(t)}^r)$$

並計算

$$X_r = \sum_{i=1}^{N_r(t)} W_i^r$$

以及

$$Y_r = \sum_{i=1}^{N_r(t)} S_i^r$$

由此得樣本共變異數

$$\overline{\text{Cov}(X,Y)} = \frac{1}{r-1} \sum_{i=1}^{r} (X_i - \overline{X})(Y_i - \overline{Y})$$

並以此估計 $\text{Cov}(X,Y)$.

因此, 根據上式以及 (9) 式, 以

$$\begin{aligned}
\widehat{c}^* &= -\frac{\overline{\text{Cov}(X,Y)}}{\text{Var}(Y)} \\
&= -\frac{\dfrac{1}{r-1} \displaystyle\sum_{i=1}^{r} (X_i - \overline{X})(Y_i - \overline{Y})}{E(S^2)E[N(t)]}
\end{aligned}$$

估計 c^*.

(iii) 以控制變量估計量

$$\overline{X} + \widehat{c}^*\{\overline{Y} - E(S)E[N(t)]\}$$

估計

$$\theta = E(X)$$