詳細解答

STEP III 1991，第 16 題

概率與統計 • 聯合分佈與邊際分佈

STEP III 1991，第 16 題

繼續閱讀之前，請先看題目（原題為英文）。題目載於 STEP Support Programme 的試題資料庫：

step.maths.org/questions/91-s3-q16

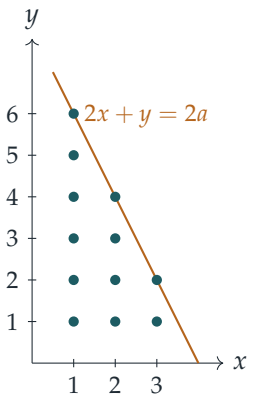
關鍵思路

這裏的一切都是在一個由格點（lattice points）組成的三角形上的有限和。先把區域畫出來：對每個 x，我們要知道哪些 y 是容許的；對每個 y，又有哪些 x 是容許的。一旦求和的上下限定對了，每個邊際分佈（marginal distribution）都是一個等差數列（arithmetic progression）之和。唯一要留神的地方，是對給定的 y，容許的最大 x 是 $\lfloor (2a-y)/2 \rfloor$，而這取決於 y 是偶數還是奇數。

解答

區域

我們需要 $x \geq 1$、$y \geq 1$ 和 $2x + y \leq 2a$。由於 $y \geq 1$，必須有 $2x \leq 2a - 1$，所以 $x \leq a - 1$。對於固定的 x（$1 \leq x \leq a - 1$），容許的值是 $y = 1, 2, \ldots, 2a - 2x$。同樣地，因為 $x \geq 1$，所以 $y \leq 2a - 2$。



圖中所示為 $a = 4$ 的情況：第 x 列有 $2a - 2x$ 個點。

X 的邊際分佈與 c 的值

固定 x，並記 $n = 2a - 2x$。把等差數列 $2x + 1, \ldots, 2x + n$ 求和，

$$\begin{aligned} \mathrm{P}(X = x) &= c \sum_{y=1}^{n} (2x + y) = c \left(2xn + \frac{n(n+1)}{2} \right) = c \, \frac{n}{2} \, (4x + n + 1) \\ &= c \, (a - x)(2a + 2x + 1), \qquad x = 1, 2, \ldots, a - 1, \end{aligned}$$

而在其他情況下 $\mathrm{P}(X = x) = 0$。

所有概率之和必須等於 1。設 $k = a - x$，它由 1 取到 $a - 1$；於是 $2a + 2x + 1 = 4a + 1 - 2k$，並且

$$\begin{aligned} \sum_{k=1}^{a-1} k(4a + 1 - 2k) &= (4a + 1) \frac{(a-1)a}{2} - 2 \cdot \frac{(a-1)a(2a-1)}{6} \\ &= a(a-1) \left(\frac{4a+1}{2} - \frac{2a-1}{3} \right) = \frac{a(a-1)(8a+5)}{6}. \end{aligned}$$

令 c 乘以此和等於 1，便得

$$c = \frac{6}{a(a-1)(8a+5)},$$

正如所求。

Y 的邊際分佈

固定 y，其中 $1 \le y \le 2a - 2$。容許的 x 是 $1, 2, \ldots, N$，其中 N 是滿足 $2N \le 2a - y$ 的最大整數。對這 N 項求和，

$$\sum_{x=1}^{N} (2x + y) = N(N + 1) + yN = N(N + 1 + y).$$

y **為偶數**。此時 $2a-y$ 是偶數，$N=\frac{1}{2}(2a-y)$，所以 $N+1+y=\frac{1}{2}(2a+2+y)$，並且

$$\mathrm{P}(Y=y)=c\cdot\frac{(2a-y)(2a+2+y)}{4}=\frac{3(2a-y)(2a+2+y)}{2a(a-1)(8a+5)},$$

其中用了 c 的值。這就是所要求的結果。

想一想 當 y 是奇數時，為甚麼不能用 $N=\frac{1}{2}(2a-y)$？

當 y 是奇數時，$\frac{1}{2}(2a-y)$ 不是整數。滿足 $2x\le 2a-y$ 的最大整數 x 是 $N=\frac{1}{2}(2a-y-1)$：向下取整把那半個捨去了。

y **為奇數**。取 $N=\frac{1}{2}(2a-1-y)$，便得 $N+1+y=\frac{1}{2}(2a+1+y)$，所以

$$\mathrm{P}(Y=y)=c\cdot\frac{(2a-1-y)(2a+1+y)}{4}=\frac{3(2a-1-y)(2a+1+y)}{2a(a-1)(8a+5)}.$$

（這兩種情況可以寫成 $3\big((2a+1)^2-(y+1)^2\big)$ 和 $3\big(4a^2-(y+1)^2\big)$ 除以同一個分母。）

Y 的期望值

把 $y\,\mathrm{P}(Y=y)$ 求和，就要把偶數和奇數的 y 分開處理。較簡潔的做法是回到聯合分佈（joint distribution），逐列求和。對於固定的 x，同前面一樣取 $n=2a-2x$，

$$\sum_{y=1}^{n}y(2x+y)=2x\cdot\frac{n(n+1)}{2}+\frac{n(n+1)(2n+1)}{6}=\frac{n(n+1)}{6}\big(6x+2n+1\big).$$

再次設 $k=a-x$，於是 $n=2k$，$6x+2n+1=6a-2k+1$。每列之和變成

$$\frac{2k(2k+1)(6a+1-2k)}{6}=\frac{k(2k+1)(6a+1-2k)}{3}.$$

展開得 $k(2k+1)(6a+1-2k)=(6a+1)(2k^2+k)-4k^3-2k^2$。利用求和至 $a-1$ 的標準公式，

$$\sum_{k=1}^{a-1}(2k^2+k)=\frac{a(a-1)(2a-1)}{3}+\frac{a(a-1)}{2}=\frac{a(a-1)(4a+1)}{6},$$

$$\sum_{k=1}^{a-1}k^3=\frac{a^2(a-1)^2}{4},\qquad\sum_{k=1}^{a-1}k^2=\frac{a(a-1)(2a-1)}{6}.$$

所以

$$\begin{aligned}&\sum_{k=1}^{a-1}k(2k+1)(6a+1-2k)\\&\quad=a(a-1)\left(\frac{(6a+1)(4a+1)}{6}-a(a-1)-\frac{2a-1}{3}\right)\\&\quad=\frac{a(a-1)}{6}\big(24a^2+10a+1-6a^2+6a-4a+2\big)\\&\quad=\frac{a(a-1)(6a^2+4a+1)}{2}.\end{aligned}$$

因此

$$\mathrm{E}(Y) = \frac{c}{3}\cdot\frac{a(a-1)(6a^2+4a+1)}{2} = \frac{6}{a(a-1)(8a+5)}\cdot\frac{a(a-1)(6a^2+4a+1)}{6}$$
$$= \frac{6a^2+4a+1}{8a+5}.$$

驗算

- 當 $a = 2$ 時只有兩個點，$(1,1)$ 和 $(1,2)$，權重分別為 $3c$ 和 $4c$，所以 $c = \frac{1}{7}$。這與 $6/(2\cdot 1\cdot 21)$ 一致。於是 $\mathrm{E}(Y) = \frac{3+8}{7} = \frac{11}{7}$，而公式給出 $\frac{33}{21} = \frac{11}{7}$。
- X 和 Y 的邊際分佈公式（奇偶兩種情況）以及 $\mathrm{E}(Y)$，已對 a 由 2 至 29 的每個值用分數作精確枚舉核對。例如，$a = 3$ 給出 $\mathrm{E}(Y) = \frac{67}{29}$。
- 當 a 很大時，$\mathrm{E}(Y) \approx \frac{3}{4}a$。權重 $2x + y$ 偏重靠近斜邊的點，所以 Y 明顯高於在三角形上取相等權重時所得的 $\frac{2}{3}a$。

常見錯誤

- 讓 x 一直取到 a：$x = a$ 這一點會迫使 $y \le 0$。
- 對奇數 y 使用 $N = \frac{1}{2}(2a - y)$，而它並不是整數。
- 試圖從分兩種情況的邊際分佈求 $\mathrm{E}(Y)$。這樣做是行得通的，但沿聯合分佈的各列求和，就可以完全避開奇偶之分。

STEP 之外

從聯合分佈求邊際分佈，就是把另一個變量求和消去，而難處總在於區域的上下限。對於連續變量，同樣的工作變成在二重積分（double integral）中選取積分限，而交換求和次序（正如我們求 $\mathrm{E}(Y)$ 時所做的）則變成交換積分次序。Grimmett 與 Welsh 的 *Probability: An Introduction*（概率論入門）涵蓋了離散和連續變量的聯合分佈與邊際分佈。

Beyond STEP 詳細解答：https://beyondstep.keithmcnulty.org/zh-hant/。解答文字採用 CC0 1.0；題目本身版權屬 Cambridge University Press & Assessment 所有（©），本文件不予轉載。由機器從英文翻譯而成；尚未經審閱。