詳細解答

STEP III 2015，第 13 題

概率與統計 • 均勻分佈變量的函數

STEP III 2015，第 13 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/15-s3-q13

關鍵思路

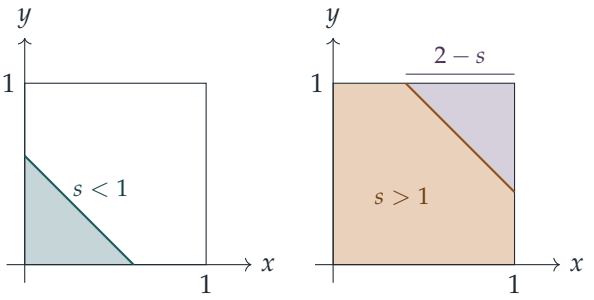
由於 X 與 Y 互相獨立 (independent)，且都在 $[0,1]$ 上均勻分佈 (uniform distribution)，數對 (X,Y) 便在單位正方形上均勻分佈，因此關於 X 和 Y 的每一個概率，其實就是一個面積。像 $X+Y \leq s$ 或 $Y \leq rX$ 這類事件，都是用一條直線把正方形切開，而直線一側的面積不是一個三角形，就是正方形減去一個三角形。之後，只要把 $1/(X+Y)$ 和 $X/(X+Y)$ 的事件改寫成我們已經量度過的事件，就能得出它們的分佈。

解答

全文中，(X,Y) 在面積為 1 的單位正方形 $[0,1]^2$ 上均勻分佈，所以對正方形內任何區域 A，$\mathrm{P}((X,Y) \in A)$ 都等於 A 的面積。

第 (i) 部分

設 $S = X+Y$，其取值範圍為 $[0,2]$，並設 $F(s) = \mathrm{P}(S \leq s)$。事件 $S \leq s$ 就是正方形中位於直線 $x+y=s$ 上或其下方的部分。



若 $0 \le s \le 1$，該直線與兩軸相交於 $(s,0)$ 和 $(0,s)$，其下方的區域是一個兩直角邊長均為 s 的等腰直角三角形。若 $1 \le s \le 2$，則直線上方的區域是位於角落 $(1,1)$ 的三角形；它的兩條直角邊由直線延伸至該角，長度為 $1-(s-1) = 2-s$。因此

$$F(s) = \mathrm{P}(X+Y \le s) = \begin{cases} 0 & s < 0, \\ \frac{1}{2}s^2 & 0 \le s \le 1, \\ 1 - \frac{1}{2}(2-s)^2 & 1 \le s \le 2, \\ 1 & s > 2. \end{cases}$$

中間兩條公式在 $s = 1$ 時一致（都等於 $\frac{1}{2}$），這是理所當然的。

第 (ii) 部分

設 $T = 1/S$。由於 $0 < S \le 2$（$S = 0$ 的概率為零），T 的取值範圍為 $[\frac{1}{2}, \infty)$。對 $t > 0$，因為 $s \mapsto 1/s$ 在 $s > 0$ 上是遞減的，

$$\mathrm{P}(T \le t) = \mathrm{P}\big(S \ge \tfrac{1}{t}\big) = 1 - F\big(\tfrac{1}{t}\big),$$

這裏用到 S 是連續隨機變量（continuous random variable），所以 $\mathrm{P}(S = 1/t) = 0$。

? **想一想** 應該用 F 的哪一條公式？這取決於 $1/t$ 落在哪裏，而不是 t 落在哪裏。

若 $t \ge 1$，則 $0 < 1/t \le 1$，所以用第一個分支：$\mathrm{P}(T \le t) = 1 - \frac{1}{2}t^{-2}$。若 $\frac{1}{2} \le t \le 1$，則 $1 \le 1/t \le 2$，所以用第二個分支：$\mathrm{P}(T \le t) = \frac{1}{2}\big(2 - t^{-1}\big)^2$。而當 $t < \frac{1}{2}$ 時，$\mathrm{P}(T \le t) = 0$。

求導便得到概率密度函數（probability density function）。當 $t \ge 1$，

$$f(t) = \frac{\mathrm{d}}{\mathrm{d}t}\Big(1 - \tfrac{1}{2}t^{-2}\Big) = t^{-3},$$

而當 $\frac{1}{2} \le t \le 1$，由鏈式法則（chain rule），

$$f(t) = \big(2 - t^{-1}\big) \cdot t^{-2} = 2t^{-2} - t^{-3},$$

其他情況下 $f(t) = 0$。這就是所求的概率密度函數。

至於期望值（expectation），

$$\begin{aligned} \mathrm{E}\Big(\frac{1}{X+Y}\Big) = \int_{1/2}^{\infty} t f(t)\,\mathrm{d}t &= \int_{1/2}^{1} \big(2t^{-1} - t^{-2}\big)\,\mathrm{d}t + \int_{1}^{\infty} t^{-2}\,\mathrm{d}t \\ &= \Big[2\ln t + t^{-1}\Big]_{1/2}^{1} + \Big[-t^{-1}\Big]_{1}^{\infty} \\ &= (0+1) - (-2\ln 2 + 2) + 1 \\ &= 2\ln 2. \end{aligned}$$

這個反常積分（improper integral）收斂，因為 t^{-2} 在 $[1,\infty)$ 上可積。

第 (iii) 部分

設 $R = Y/X$，其取值範圍為 $[0,\infty)$（事件 $X = 0$ 的概率為零）。對 $r \geq 0$，事件 $R \leq r$ 即 $Y \leq rX$：正方形中位於通過原點的直線 $y = rx$ 上或其下方的部分。

若 $0 \leq r \leq 1$，這條直線從正方形右邊的 $(1, r)$ 離開，其下方的區域是一個底為 1、高為 r 的三角形，面積為 $\frac{1}{2}r$。若 $r \geq 1$，直線從頂邊的 $(1/r, 1)$ 離開，其上方的區域是一個沿頂邊的底為 $1/r$、高為 1 的三角形，面積為 $\frac{1}{2r}$。所以

$$H(r) = \mathrm{P}\Big(\frac{Y}{X} \leq r\Big) = \begin{cases} 0 & r < 0, \\ \frac{1}{2}r & 0 \leq r \leq 1, \\ 1 - \dfrac{1}{2r} & r \geq 1. \end{cases}$$

現設 $Z = \dfrac{X}{X+Y} = \dfrac{1}{1+R}$，其取值範圍為 $[0,1]$。對 $0 < z \leq 1$，由於 $r \mapsto 1/(1+r)$ 在 $r \geq 0$ 上是遞減的，

$$\mathrm{P}(Z \leq z) = \mathrm{P}\Big(1 + R \geq \frac{1}{z}\Big) = \mathrm{P}\Big(R \geq \frac{1-z}{z}\Big) = 1 - H\Big(\frac{1-z}{z}\Big).$$

同樣，用哪個分支取決於 $(1-z)/z$ 落在哪裏：它不大於 1 當且僅當 $z \geq \frac{1}{2}$。

- 當 $\frac{1}{2} \leq z \leq 1$：$\mathrm{P}(Z \leq z) = 1 - \dfrac{1-z}{2z} = \dfrac{3}{2} - \dfrac{1}{2z}$。
- 當 $0 < z \leq \frac{1}{2}$：$\mathrm{P}(Z \leq z) = 1 - \Big(1 - \dfrac{z}{2(1-z)}\Big) = \dfrac{z}{2(1-z)}$。

兩者在 $z = \frac{1}{2}$ 時都等於 $\frac{1}{2}$。求導（第二條可先寫成 $\frac{z}{2(1-z)} = \frac{1}{2}\big(\frac{1}{1-z} - 1\big)$），得 Z 的概率密度函數為

$$g(z) = \begin{cases} \dfrac{1}{2(1-z)^2} & 0 \leq z \leq \frac{1}{2}, \\ \dfrac{1}{2z^2} & \frac{1}{2} \leq z \leq 1, \\ 0 & \text{其他情況}. \end{cases}$$

留意 $g(1-z) = g(z)$：這個概率密度函數關於 $\frac{1}{2}$ 對稱。

期望值。 我們無須積分便可直接寫出答案。由於 X 與 Y 互相獨立且分佈相同，數對 (X, Y) 與 (Y, X) 有相同的分佈。所以 $\dfrac{X}{X+Y}$ 與 $\dfrac{Y}{X+Y}$ 的期望值相同。兩者之和恰好是 1，所以

$$\mathrm{E}\Big(\frac{X}{X+Y}\Big) = \tfrac{1}{2}.$$

以積分驗證。

$$\mathrm{E}(Z) = \int_0^{1/2} \frac{z}{2(1-z)^2}\,\mathrm{d}z + \int_{1/2}^1 \frac{1}{2z}\,\mathrm{d}z.$$

第二個積分等於 $\frac{1}{2}\big[\ln z\big]_{1/2}^1 = \frac{1}{2}\ln 2$。第一個積分中，代入 $u = 1-z$，於是 $\mathrm{d}z = -\mathrm{d}u$，而 $z = 0, \frac{1}{2}$ 分別對應 $u = 1, \frac{1}{2}$：

$$\begin{aligned}\int_0^{1/2} \frac{z}{2(1-z)^2}\,\mathrm{d}z &= \frac{1}{2}\int_{1/2}^1 \frac{1-u}{u^2}\,\mathrm{d}u = \frac{1}{2}\Big[-\frac{1}{u} - \ln u\Big]_{1/2}^1 \\ &= \frac{1}{2}\big(-1-(-2+\ln 2)\big) = \frac{1}{2}(1-\ln 2).\end{aligned}$$

相加得 $\mathrm{E}(Z) = \frac{1}{2}(1-\ln 2) + \frac{1}{2}\ln 2 = \frac{1}{2}$，證實了上面的值。

驗算

- 每個分佈函數（distribution function）都由 0 上升至 1，並在各分支相接處連續：$F(1) = \frac{1}{2}$、$\mathrm{P}(T \le 1) = \frac{1}{2}$、$H(1) = \frac{1}{2}$、$\mathrm{P}(Z \le \frac{1}{2}) = \frac{1}{2}$。由於正方形關於對角線對稱，這些值都是 $\frac{1}{2}$。
- 概率密度函數 f 的積分等於 1：$\int_{1/2}^1 (2t^{-2} - t^{-3})\,\mathrm{d}t = 2 - \frac{3}{2} = \frac{1}{2}$，而 $\int_1^\infty t^{-3}\,\mathrm{d}t = \frac{1}{2}$。
- $2\ln 2 \approx 1.386$ 介乎 1 與 2 之間，合乎預期；而且它大於 $1/\mathrm{E}(X+Y) = 1$，正如對凸函數（convex function）$1/s$ 應用琴生不等式（Jensen's inequality）所預測的。直接以數值計算 $\iint (x+y)^{-1}\,\mathrm{d}x\,\mathrm{d}y$ 在正方形上的值，得 1.3863。各分佈函數以及 $\mathrm{E}(Z) = \frac{1}{2}$ 亦已用電腦模擬證實。

常見錯誤

- 按 t 或 z 而不是按 $1/t$ 或 $(1-z)/z$ 來選擇 F 或 H 的分支，結果把兩條公式對調了。
- 忘記 $1/s$ 和 $1/(1+r)$ 是遞減的，因而沒有把不等號反轉：$T \le t$ 的意思是 $S \ge 1/t$。
- 把第 (i) 部分角落三角形的直角邊弄錯：它是 $2-s$，不是 $s-1$。

STEP 之外

把關於 (X, Y) 的陳述化為面積，再對分佈函數求導，這是求隨機變量函數之分佈的一般方法，稱為*分佈函數法*（distribution function method）。到了大學，還會學到聯合概率密度函數的變量代換公式（雅可比行列式，Jacobian），它可以同時處理 $X+Y$ 與 $X/(X+Y)$；特別地，這裏 $X+Y$ 的概率密度函數正是兩個均勻分佈概率密度函數的*卷積*（convolution）。Grimmett 與 Welsh 的 *Probability: An Introduction*（《概率導論》）涵

蓋了這兩種方法，而 Blitzstein 與 Hwang 的 *Introduction to Probability*（《概率導論》）則有許多同類型的例子。

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。