詳細解答

STEP II 2006，第 14 題

概率與統計 • 連續隨機變量

> **STEP II 2006，第 14 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/06-s2-q14

關鍵思路

整道題都建基於一個原函數（antiderivative）。分母 $x\ln x$ 是 $\ln x$ 乘以 x，而 $1/x$ 正是 $\ln x$ 的導數，所以被積函數具有 $g'(x)/g(x)$ 的形式，其中 $g(x)=\ln x$。因此

$$\int \frac{1}{x\ln x}\,\mathrm{d}x = \ln|\ln x| + C.$$

此後，每一部分都只需運用「總概率為 1」，並小心留意正負號，以及密度（density）實際上在哪裏不為零。

解答

草圖

設 $y=\dfrac{1}{x\ln x}$，其中 $x>0$，$x\neq 1$。

正負號。由於 $x>0$，y 的正負號與 $\ln x$ 相同：在 $0<x<1$ 上 $y<0$，在 $x>1$ 上 $y>0$。曲線與兩條坐標軸都不相交。

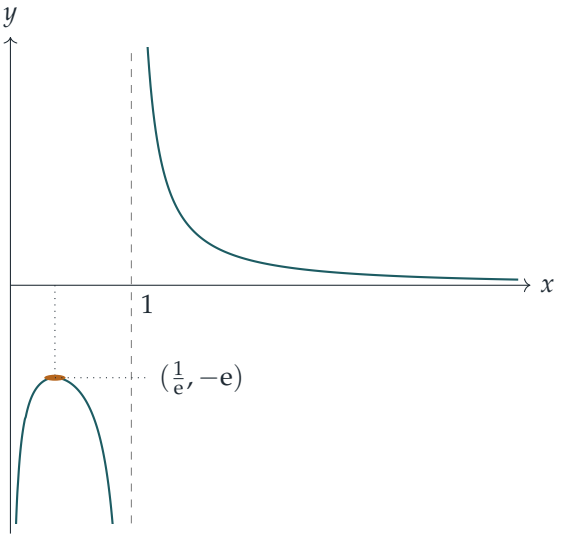
兩端的變化。

- 當 $x\to 0^+$ 時：題目告訴我們 $x\ln x\to 0$，而且在該處 $x\ln x<0$，所以 $y\to-\infty$。
- 當 $x\to 1$ 時：$x\ln x\to 0$，從左邊趨近時為負，從右邊趨近時為正，所以當 $x\to 1^-$ 時 $y\to-\infty$，當 $x\to 1^+$ 時 $y\to+\infty$。直線 $x=1$ 是垂直漸近線(vertical asymptote)。
- 當 $x\to\infty$ 時：$x\ln x\to\infty$，所以 $y\to 0^+$，而 x 軸是漸近線。

轉向點（*turning points*）。由商法則（quotient rule）或鏈式法則（chain rule），

$$\frac{\mathrm{d}y}{\mathrm{d}x} = -\frac{\ln x + 1}{(x \ln x)^2},$$

它只在 $\ln x = -1$ 時為零，即 $x = 1/\mathrm{e}$，此時 $y = \dfrac{1}{(1/\mathrm{e})(-1)} = -\mathrm{e}$。當 $0 < x < 1/\mathrm{e}$ 時，我們有 $\ln x + 1 < 0$，所以 y 遞增；當 $x > 1/\mathrm{e}$（且 $x \neq 1$）時，函數遞減。因此 $(1/\mathrm{e}, -\mathrm{e})$ 是左邊分支的局部極大值（local maximum）點，而右邊分支則由 $+\infty$ 一直遞減至 0。



第 (i) 部分

這裏 $a = \frac{1}{4}$ 而 $b = \frac{1}{2}$，兩者都在 $(0,1)$ 內，而在該區間上 $x \ln x < 0$。利用上面的原函數，

$$\begin{aligned}\int_{1/4}^{1/2} \frac{\mathrm{d}x}{x \ln x} &= \Big[\ln|\ln x|\Big]_{1/4}^{1/2} = \ln|\ln \tfrac{1}{2}| - \ln|\ln \tfrac{1}{4}| \\ &= \ln(\ln 2) - \ln(2 \ln 2) = -\ln 2.\end{aligned}$$

總概率必須為 1，所以 $\lambda \cdot (-\ln 2) = 1$，由此得

$$\lambda = -\frac{1}{\ln 2}.$$

這個負號正是我們所需要的：在 $[\frac{1}{4}, \frac{1}{2}]$ 上 $x \ln x < 0$，所以 λ 必須為負，f 在該處才會非負。

第 (ii) 部分

當 $\lambda = 1$ 時，密度必須非負，所以區間 $[a,b]$ 必須位於 $x\ln x > 0$ 的範圍內；由於 $b > a > 1$，這一點確實成立。於是 $\ln a, \ln b > 0$，並且

$$1 = \int_a^b \frac{\mathrm{d}x}{x\ln x} = \ln(\ln b) - \ln(\ln a) = \ln\frac{\ln b}{\ln a}.$$

所以 $\dfrac{\ln b}{\ln a} = \mathrm{e}$，即 $\ln b = \mathrm{e}\ln a = \ln\left(a^{\mathrm{e}}\right)$；又因為 ln 是一對一（one-to-one）的，故 $b = a^{\mathrm{e}}$，正如所求。

第 (iii) 部分

取 $a = \mathrm{e}$，由第 (ii) 部分得 $b = \mathrm{e}^{\mathrm{e}}$。由於 $1 < \frac{3}{2} < 2 < \mathrm{e}$，整個區間 $[\mathrm{e}^{3/2}, \mathrm{e}^2]$ 都位於 $[a,b] = [\mathrm{e}, \mathrm{e}^{\mathrm{e}}]$ 之內，所以

$$\mathrm{P}(\mathrm{e}^{3/2} \le X \le \mathrm{e}^2) = \Big[\ln(\ln x)\Big]_{\mathrm{e}^{3/2}}^{\mathrm{e}^2} = \ln 2 - \ln\tfrac{3}{2} = \ln\tfrac{4}{3}.$$

要把這個結果化成分數，可使用麥克勞林級數（Maclaurin series）$\ln(1+t) = t - \frac{t^2}{2} + \frac{t^3}{3} - \frac{t^4}{4} + \cdots$，取 $t = \frac{1}{3}$，並保留四項：

$$\ln\tfrac{4}{3} \approx \frac{1}{3} - \frac{1}{18} + \frac{1}{81} - \frac{1}{324} = \frac{108 - 18 + 4 - 1}{324} = \frac{93}{324} = \frac{31}{108}.$$

想一想 我們怎樣知道取到第四項就停下來是合理的？

這個級數正負項交替出現，而且各項的大小遞減，所以取四項之後的誤差小於第五項，即 $\frac{1}{5\cdot 3^5} = \frac{1}{1215} < 0.001$。因此 $\mathrm{P} = \ln\frac{4}{3} \approx \frac{31}{108}$，正如所求。

第 (iv) 部分

現在 $a = \mathrm{e}^{1/2}$，所以由第 (ii) 部分得

$$b = a^{\mathrm{e}} = \mathrm{e}^{\mathrm{e}/2}.$$

想一想 在積分之前，先想想：密度實際上在哪裏不為零？

由於 $\mathrm{e} < 3$，我們有 $\mathrm{e}/2 < \frac{3}{2}$，所以 $b = \mathrm{e}^{\mathrm{e}/2} < \mathrm{e}^{3/2}$。區間 $[\mathrm{e}^{3/2}, \mathrm{e}^2]$ 完全位於 b 的右邊，而在那裏 $\mathrm{f}(x) = 0$。因此

$$\mathrm{P}(\mathrm{e}^{3/2} \le X \le \mathrm{e}^2) = 0.$$

（若不加思索地把 $1/(x\ln x)$ 由 $\mathrm{e}^{3/2}$ 積分至 e^2，便會再次得到 $\ln\frac{4}{3}$，而這是錯的。）

驗算

- 以數值計算，$e^{-1} \approx 0.368$，而在 $(0,1)$ 上逐點檢視 y 的值，可見它在該區間上的最大值 $\approx -2.718 = -e$。
- 辛普森法則(Simpson's rule)給出 $\int_{1/4}^{1/2} \frac{dx}{x \ln x} \approx -0.6931 = -\ln 2$，所以 $\lambda \approx -1.4427$。
- 對於若干個 $a > 1$ 的值（例如 $a = 2$、$a = 7.3$），$\int_a^{a^e} \frac{dx}{x \ln x}$ 的結果都是 1.0000。
- $\ln \frac{4}{3} \approx 0.28768$ 而 $\frac{31}{108} \approx 0.28704$，兩者相差不超過 0.0007。
- 在 (iv) 中，$b = e^{e/2} \approx 3.89$，而 $e^{3/2} \approx 4.48$。

常見錯誤

- 在第 (i) 部分寫成 $\ln(\ln x)$，但那裏 $\ln x < 0$：原函數需要加上絕對值（modulus），即 $\ln|\ln x|$。
- 答 $\lambda = \frac{1}{\ln 2}$：正的 λ 會令「密度」變成負數。
- 在第 (iv) 部分，沒有先檢查 $[e^{3/2}, e^2]$ 是否位於支撐集（support）$[a,b]$ 之內，便直接在這個區間上積分。

STEP 之外

代換(substitution)$u = \ln x$ 把這個密度變成 $[\ln a, \ln b]$ 上的 $1/u$，而條件 $b = a^e$ 其實就是 $\int_{\ln a}^{\ln b} \frac{du}{u} = 1$。像這樣對隨機變量（random variable）作變換（這裏是 $U = \ln X$），是大學一年級概率論的核心工具；在那裏，隨機變量的函數的密度是由鏈式法則求出的。同一個函數 $1/(x \ln x)$ 在分析學中也十分有名：$\sum 1/(n \ln n)$ 發散（diverges），這可以透過與這個積分比較而得出，儘管它的各項趨向零。關於隨機變量的變換，可參閱 Grimmett 與 Welsh 的 *Probability: An Introduction*。

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。