詳細解答

STEP II 1999，第 13 題

概率與統計 • 連續隨機變量

> **STEP II 1999，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/99-s2-q13

關鍵思路

這道題雖然歸入力學（mechanics）一類，其實是一道關於連續隨機變量（continuous random variable）的題目。只要留意到較短的一段有簡單的分佈，一切便迎刃而解：若折斷點在棒上均勻分佈（uniform），則較短的一段在棒長的 $[0, \frac{1}{2}]$ 上均勻分佈。比值 R 是較短一段的遞增函數，所以「$R < r$」可化為只涉及較短一段的命題。由此得出分佈函數（distribution function），再求導便得到密度（density）。

解答

R 的分佈

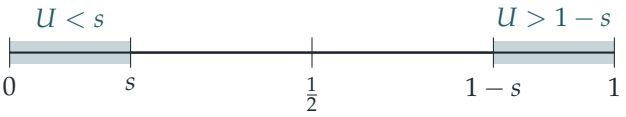
設棒長為 1；比值與所用單位無關。設折斷點與其中一端的距離為 U，其中 U 在 $[0,1]$ 上均勻分佈。兩段的長度分別為 U 和 $1-U$，所以較短一段的長度為

$$S = \min(U, 1-U),$$

而較長一段的長度為 $1-S$。於是 $0 \le S \le \frac{1}{2}$，並且

$$R = \frac{S}{1-S}.$$

對於 $0 \le s \le \frac{1}{2}$，事件 $S < s$ 恰好在 $U < s$ 或 $U > 1-s$ 時發生。這是兩個互不相交（disjoint）的區間，總長度為 $2s$，所以 $\mathrm{P}(S < s) = 2s$。換句話說，S 在 $[0, \frac{1}{2}]$ 上均勻分佈。



現設 $0 \le r \le 1$。函數 $s \mapsto \dfrac{s}{1-s}$ 在 $[0, \frac{1}{2}]$ 上嚴格遞增（strictly increasing）（其導數為 $1/(1-s)^2 > 0$），所以我們可以把不等式反解出來：

$$R < r \iff \frac{S}{1-S} < r \iff S < r(1-S) \iff S < \frac{r}{1+r},$$

其中兩邊同乘 $1-S$ 是穩妥的，因為 $1 - S \ge \frac{1}{2} > 0$。當 $0 \le r \le 1$ 時，$\dfrac{r}{1+r}$ 的值落在 $[0, \frac{1}{2}]$ 之內，所以

$$\boxed{\mathrm{P}(R < r) = \frac{2r}{1+r} \qquad (0 \le r \le 1).}$$

由於較短的一段永遠不會比較長的一段長，R 必定落在 $[0,1]$ 之內。所以當 $r \le 0$ 時 $\mathrm{P}(R < r) = 0$，當 $r \ge 1$ 時 $\mathrm{P}(R < r) = 1$，這與公式在 $r = 1$ 處取值 1 吻合。

想一想 為甚麼不直接用 U 來做，即取 $R = U/(1-U)$？

這只在 $U \le \frac{1}{2}$ 時才正確。當 $U > \frac{1}{2}$ 時，較短的一段是 $1-U$，而 $R = (1-U)/U$。你可以分成這兩種情況來處理，根據對稱性，每種情況各貢獻 $\frac{r}{1+r}$。改用 S 則一次過把這些細節處理妥當。

密度、平均值與方差

在 $(0,1)$ 上對分佈函數 $F(r) = \dfrac{2r}{1+r} = 2 - \dfrac{2}{1+r}$ 求導，得到概率密度函數（probability density function）

$$f(r) = \frac{2}{(1+r)^2} \quad (0 \le r \le 1), \qquad f(r) = 0 \text{ 其他情況.}$$

驗算一下：$\int_0^1 f(r)\,\mathrm{d}r = F(1) - F(0) = 1$。

求平均值（mean）時，寫成 $r = (1+r) - 1$：

$$\begin{aligned}\mathrm{E}(R) &= \int_0^1 \frac{2r}{(1+r)^2}\,\mathrm{d}r = 2\int_0^1 \left(\frac{1}{1+r} - \frac{1}{(1+r)^2}\right)\mathrm{d}r \\ &= 2\Big[\ln(1+r) + \frac{1}{1+r}\Big]_0^1 = 2\left(\ln 2 + \tfrac{1}{2} - 1\right) \\ &= 2\ln 2 - 1 \approx 0.386.\end{aligned}$$

求二階矩（second moment）時，寫成 $r^2 = (1+r)^2 - 2(1+r) + 1$，於是

$$\frac{r^2}{(1+r)^2} = 1 - \frac{2}{1+r} + \frac{1}{(1+r)^2}.$$

那麼

$$\begin{aligned}\mathrm{E}(R^2) &= 2\int_0^1 \left(1 - \frac{2}{1+r} + \frac{1}{(1+r)^2}\right) \mathrm{d}r = 2\Big[r - 2\ln(1+r) - \frac{1}{1+r}\Big]_0^1 \\ &= 2\left(1 - 2\ln 2 - \tfrac{1}{2} + 1\right) = 3 - 4\ln 2.\end{aligned}$$

因此方差（variance）為

$$\begin{aligned}\mathrm{Var}(R) &= \mathrm{E}(R^2) - (\mathrm{E}(R))^2 = 3 - 4\ln 2 - (4(\ln 2)^2 - 4\ln 2 + 1) \\ &= 2 - 4(\ln 2)^2 \approx 0.0782.\end{aligned}$$

含 $\ln 2$ 的項互相抵消，剩下這個簡潔的形式。

驗算

- 在 $r = 1$ 處概率為 1，理應如此。在 $r = \frac{1}{2}$ 處概率為 $\frac{2}{3}$：較長一段超過較短一段的兩倍，恰好是折斷點落在棒兩端各三分之一的範圍內，而這兩部分的總長度為 $\frac{2}{3}$。
- 平均值 0.386 介乎 0 與 1 之間，並且小於 $\frac{1}{2}$，這與密度由 $r = 0$ 處的 2 遞減至 $r = 1$ 處的 $\frac{1}{2}$ 相符。
- 模擬 10^6 次隨機折斷，得到樣本平均值 0.3862 及樣本方差 0.0783，相比之下理論值為 $2\ln 2 - 1 = 0.38629$ 及 $2 - 4(\ln 2)^2 = 0.07819$。

常見錯誤

- 對所有 U 都取 $R = U/(1-U)$。這個量可以大於 1，所以它並不是較短一段與較長一段之比。
- 忘記折斷點可以靠近任何一端，因而得出 $\mathrm{P}(R < r) = \frac{r}{1+r}$。
- 用較短一段的平均值 $\frac{1}{4}$ 來得出 $\mathrm{E}(R) = \frac{1/4}{3/4} = \frac{1}{3}$。一般而言 $\mathrm{E}\big(g(S)\big) \neq g\big(\mathrm{E}(S)\big)$。

STEP 之外

這裏用的是變換隨機變量的標準方法，即「分佈函數法」。若 $Y = g(X)$ 而 g 為遞增函數，則 $F_Y(y) = F_X\big(g^{-1}(y)\big)$，求導後便得到變量代換公式（change-of-variables formula）$f_Y(y) = f_X\big(g^{-1}(y)\big)\left|\mathrm{d}g^{-1}/\mathrm{d}y\right|$。同一想法可以引申到順序統計量（order statistics），以及經典的折棒問題，例如三段能構成三角形的概率。Grimmett 與 Welsh 所著的 *Probability: An Introduction*（概率論入門）是研讀連續隨機變量變換的好去處。

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。