詳細解答

# STEP II 1988，第 16 題

概率與統計 • 正態分佈

> **STEP II 1988，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/88-s2-q16

## 關鍵思路

關於兩根的一切，都由單一個隨機變量（random variable）$B$ 決定，所以每一個關於 $X_1$ 和 $X_2$ 的問題，都應該轉化為「$B$ 落在哪個區間」的問題。判別式（discriminant）告訴我們兩根何時為實數，而韋達定理（Vieta's formulas）（兩根之和為 $-2B$，兩根之積為 1）令 (i) 和 (ii) 的條件很容易轉換。之後，每個答案都是標準正態變量（standard normal variable）的條件概率（conditional probability）或條件期望值（conditional expectation），而求期望值只需計算一個簡短的積分。

## 解答

全題設 $B \sim \mathrm{N}(0,1)$，其密度函數（density）為 $\phi(z) = \frac{1}{\sqrt{2\pi}}\mathrm{e}^{-z^2/2}$，分佈函數（distribution function）為 $\Phi$。設 $R$ 為「兩根皆為實數」這個事件。

### 實根

$X^2 + 2BX + 1$ 的判別式為 $4B^2 - 4$，所以兩根為實數當且僅當 $B^2 \geq 1$，即 $|B| \geq 1$。由正態分佈（normal distribution）的對稱性，

$$\mathrm{P}(R) = \mathrm{P}(|B| \geq 1) = 2(1 - \Phi(1)) = 2(1 - 0.84134) = 0.31731\ldots \approx 0.317.$$

（是否包括等根的情況 $|B| = 1$ 並無分別，因為它的概率為零。）

當兩根為實數時，它們是

$$X_{1,2} = -B \pm \sqrt{B^2 - 1}, \qquad X_1 + X_2 = -2B, \qquad X_1 X_2 = 1.$$

### 第 (i) 部分

> **想一想** 怎樣把「兩根都大於 $\frac{1}{5}$」轉化為關於 $B$ 的條件？

首先，兩根必須都是正數。兩根之積為 $1 > 0$，所以它們同號，而這個符號就是兩根之和 $-2B$ 的符號。因此我們需要 $B < 0$，連同 $R$ 即表示 $B \le -1$。記 $b = -B \ge 1$；這時兩根為 $b \pm \sqrt{b^2-1}$，都是正數。

由於兩根之積為 1，較小的根是較大的根的倒數：

$$b - \sqrt{b^2-1} = \frac{1}{b+\sqrt{b^2-1}}.$$

所以較小的根大於 $\frac{1}{5}$ 當且僅當較大的根小於 5，而這時較大的根自然也大於 $\frac{1}{5}$。因此條件就是

$$\sqrt{b^2-1} < 5 - b.$$

這需要 $b < 5$，而當 $1 \le b < 5$ 時兩邊皆為非負數，所以可以把兩邊平方而不改變不等式：

$$b^2 - 1 < 25 - 10b + b^2 \iff 10b < 26 \iff b < 2.6.$$

（$b \ge 5$ 的值無論如何都不符合，因為 $2.6 < 5$。）因此，在兩根為實數的條件下，兩根都大於 $\frac{1}{5}$ 當且僅當 $-2.6 < B \le -1$，於是

$$\begin{aligned}\mathrm{P}(X_1, X_2 > \tfrac{1}{5} \mid R) &= \frac{\mathrm{P}(-2.6 < B \le -1)}{\mathrm{P}(|B| \ge 1)} = \frac{\Phi(2.6) - \Phi(1)}{2(1-\Phi(1))} \\ &= \frac{0.99534 - 0.84134}{0.31731} = 0.4853\ldots \approx 0.485.\end{aligned}$$

### 第 (ii) 部分

由於 $|X_1 + X_2| = 2|B|$，我們要求的是 $\mathrm{E}(2|B| \,|\, |B| \ge 1)$。要計算給定某事件下的條件期望值，我們在該事件上積分，再除以該事件的概率：

$$\mathrm{E}(|B| \,|\, |B| \ge 1) = \frac{1}{\mathrm{P}(|B| \ge 1)} \int_{|z| \ge 1} |z|\, \phi(z)\, \mathrm{d}z = \frac{2\int_1^\infty z\,\phi(z)\,\mathrm{d}z}{2(1-\Phi(1))},$$

這裏用到了 $\phi$ 的對稱性。這個積分很容易計算，因為 $z\,\phi(z)$ 正是 $\phi(z)$ 的導數的負值：

$$\int_1^\infty z\,\phi(z)\,\mathrm{d}z = \Big[-\phi(z)\Big]_1^\infty = \phi(1) = \frac{\mathrm{e}^{-1/2}}{\sqrt{2\pi}} = 0.24197\ldots$$

因此

$$\mathrm{E}(|X_1 + X_2| \,|\, R) = \frac{2\phi(1)}{1-\Phi(1)} = \frac{2 \times 0.24197}{0.15866} = 3.050\ldots \approx 3.05.$$

## 驗算

- 試取 $B = -2.6$：方程為 $X^2 - 5.2X + 1 = 0$，兩根恰好是 5 和 $\frac{1}{5}$，正是 (i) 的邊界情況。
- (ii) 的答案必須大於 2，因為在 $R$ 上 $2|B| \geq 2$。答案確實如此，而且合理：正態分佈在 1 以外的尾部很薄，所以 $|B|$ 通常只比 1 大一點（它的條件平均值約為 1.53）。
- 以 $2 \times 10^6$ 個標準正態的 $B$ 值進行模擬，得出 0.3177、0.4855 和 3.049，與 0.317、0.485 和 3.05 吻合。

## 常見錯誤

- 忘記取條件：在 (i) 和 (ii) 中，我們必須除以 $\mathrm{P}(R) = 0.317$。無條件概率 $\Phi(2.6) - \Phi(1) = 0.154$ 並不是答案。
- 在 (i) 中把 $B \geq 1$ 也計算在內。這時兩根都是負數，所以沒有一個大於 $\frac{1}{5}$。
- 把 $\mathrm{E}|X_1 + X_2|$ 當作 $|\mathrm{E}(X_1 + X_2)|$ 來計算，而後者由對稱性可知等於 0。

## STEP 之外

$\phi(1)/(1 - \Phi(1))$ 這個量是截尾（truncated）正態分佈的平均值，而比值 $\phi(a)/(1 - \Phi(a))$ 稱為逆米爾斯比率（inverse Mills ratio）；它出現於統計學（例如設限數據（censored data）的模型），也出現於可靠性理論（reliability theory），在那裏它是正態分佈的風險率（hazard rate）。關於隨機多項式（random polynomials）的問題，例如係數隨機的多項式平均有多少個實根，構成了概率論中一個豐富得出人意表的領域。關於條件期望值和連續分佈，可參閱 Grimmett 與 Welsh 的 *Probability: An Introduction*（概率論入門），或 Blitzstein 與 Hwang 的 *Introduction to Probability*（概率論導引）。

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。