詳細解答

STEP II 2009，第 12 題

概率與統計 • 半正態分佈

STEP II 2009，第 12 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/09-s2-q12

關鍵思路

指數 $-2x^2$ 其實是一個喬裝了的正態分佈密度函數（normal density）。把它寫成 $-2x^2 = -\dfrac{x^2}{2\sigma^2}$，便得 $\sigma = \frac{1}{2}$，所以 f 恰好是 $Y \sim \mathrm{N}(0, \frac{1}{4})$ 的密度函數的兩倍，並在 0 處截斷。換句話說，X 與 $|Y|$ 有相同的分佈。一旦看出這一點，歸一化常數（normalising constant）、$\mathrm{E}(X^2)$ 和中位數（median）全都可以直接由正態分佈的性質得出。只有 $\mathrm{E}(X)$ 需要計算一個（非常容易的）積分。

解答

以下一律設 $Y \sim \mathrm{N}(0, \frac{1}{4})$ 及 $Z \sim \mathrm{N}(0,1)$，於是 Y 與 $\frac{1}{2}Z$ 有相同的分佈，而它的密度函數為

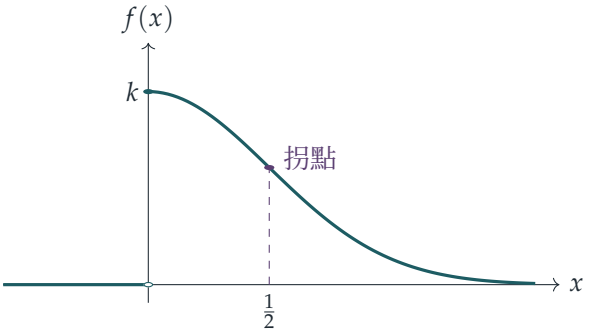
$$g(y) = \frac{1}{\frac{1}{2}\sqrt{2\pi}}\,\mathrm{e}^{-y^2/(2\cdot\frac{1}{4})} = \frac{2}{\sqrt{2\pi}}\,\mathrm{e}^{-2y^2}.$$

第 (i) 部分

當 $x \geq 0$ 時，$f'(x) = -4kx\,\mathrm{e}^{-2x^2} \leq 0$，所以 f 由 $x = 0$ 處的最大值 k 開始遞減，而該處的斜率為零。另外

$$f''(x) = k(16x^2 - 4)\,\mathrm{e}^{-2x^2},$$

它在 $x = \frac{1}{2}$ 處變號，所以在 $x = \frac{1}{2}$ 處有一個拐點（point of inflection）。當 $x \to \infty$ 時，$f(x) \to 0$，而且趨近得非常快；當 $x < 0$ 時，$f(x) = 0$。圖像是鐘形曲線的右半部分（由第 (ii) 部分可知 $k \approx 1.60$），並在 $x = 0$ 處由 0 跳升至 k。



第 (ii) 部分

總概率必須等於 1。由於 g 關於 0 對稱，而且在整條實數線上的積分為 1，所以它在 $[0,\infty)$ 上的積分為 $\frac{1}{2}$：

$$\int_0^\infty \frac{2}{\sqrt{2\pi}}\,\mathrm{e}^{-2x^2}\,\mathrm{d}x = \frac{1}{2} \quad\Longrightarrow\quad \int_0^\infty \mathrm{e}^{-2x^2}\,\mathrm{d}x = \frac{\sqrt{2\pi}}{4}.$$

因此 $k\cdot\dfrac{\sqrt{2\pi}}{4} = 1$，得

$$k = \frac{4}{\sqrt{2\pi}} = 2\sqrt{\frac{2}{\pi}} \approx 1.596.$$

留意當 $x \geq 0$ 時 $f(x) = 2g(x)$，這印證了 X 的分佈與 $|Y|$ 相同。

第 (iii) 部分

求平均值時，因子 x 令積分變得十分簡單，因為 $x\,\mathrm{e}^{-2x^2}$ 是 e^{-2x^2} 的導數的常數倍：

$$\mathrm{E}(X) = k\int_0^\infty x\,\mathrm{e}^{-2x^2}\,\mathrm{d}x = k\Big[-\tfrac{1}{4}\mathrm{e}^{-2x^2}\Big]_0^\infty = \frac{k}{4} = \frac{1}{\sqrt{2\pi}} \approx 0.399.$$

? **想一想** 怎樣才能不用冗長的分部積分法（integration by parts）而求出 $\displaystyle\int_0^\infty x^2\,\mathrm{e}^{-2x^2}\,\mathrm{d}x$？

再次利用正態分佈。函數 $x^2g(x)$ 是偶函數（even function），所以

$$\mathrm{E}(X^2) = \int_0^\infty x^2\cdot 2g(x)\,\mathrm{d}x = \int_{-\infty}^\infty x^2 g(x)\,\mathrm{d}x = \mathrm{E}(Y^2) = \mathrm{Var}(Y) = \frac{1}{4},$$

因為 $\mathrm{E}(Y) = 0$。（取 $u = x$、$\mathrm{d}v = x\mathrm{e}^{-2x^2}\,\mathrm{d}x$ 作分部積分，亦得相同結果：$\int_0^\infty x^2\mathrm{e}^{-2x^2}\,\mathrm{d}x = \frac{1}{4}\int_0^\infty \mathrm{e}^{-2x^2}\,\mathrm{d}x = \frac{1}{4k}$。）因此

$$\mathrm{Var}(X) = \mathrm{E}(X^2) - (\mathrm{E}(X))^2 = \frac{1}{4} - \frac{1}{2\pi} = \frac{\pi-2}{4\pi} \approx 0.0908.$$

第 (iv) 部分

中位數 m 滿足 $\mathrm{P}(X \le m) = \frac{1}{2}$。由於 X 的分佈與 $|Y|$ 相同，而 Y 的分佈與 $\frac{1}{2}Z$ 相同，

$$\mathrm{P}(X \le m) = \mathrm{P}(|Z| \le 2m) = 2\Phi(2m) - 1,$$

這裏用到了標準正態分佈（standard normal distribution）的對稱性。（直接計算：$\mathrm{P}(X \le m) = \int_0^m 2g(x)\,\mathrm{d}x = 2(\mathrm{P}(Y \le m) - \frac{1}{2})$。）令上式等於 $\frac{1}{2}$，得

$$\Phi(2m) = \frac{3}{4}.$$

查表得 $\Phi^{-1}(0.75) = 0.6745$，所以 $2m = 0.6745$，而

$$m \approx 0.337 \quad \text{（準確至三位有效數字）}.$$

驗算

取 $k = 2\sqrt{2/\pi}$，以數值方法（辛普森法則，Simpson's rule）在 $[0, 10]$ 上對 f 積分，得總概率為 1.0000、$\mathrm{E}(X) = 0.39894$ 及 $\mathrm{Var}(X) = 0.090845$，與 $1/\sqrt{2\pi}$ 和 $(\pi - 2)/(4\pi)$ 吻合。以數值方法解 $\int_0^m f(x)\,\mathrm{d}x = \frac{1}{2}$，得 $m = 0.33724$。以 200 000 個標準差（standard deviation）為 $\frac{1}{2}$ 的正態隨機數模擬 $|Y|$，得平均值 0.399 及方差（variance）0.091，並有 50.0% 的數值小於 0.337。合理性檢查：中位數 0.337 小於平均值 0.399，對於一個右尾較長的分佈來說，理應如此。

常見錯誤

- 從指數誤取 $\sigma^2 = 2$ 或 $\sigma = \frac{1}{4}$：應把 $2x^2$ 與 $x^2/(2\sigma^2)$ 對照，得 $\sigma^2 = \frac{1}{4}$。
- 忘記因子 2：在半條實數線上，正態密度函數的積分只有 $\frac{1}{2}$。
- 解 $\Phi(2m) = \frac{1}{2}$，結果得出 $m = 0$。X 在 m 以下的那一半概率，對應的是 Y 中間的那一半概率，所以條件應為 $\Phi(2m) = \frac{3}{4}$。

STEP 之外

X 是一個半正態（half-normal）隨機變量（random variable），而當 Y 服從正態分佈時，$|Y|$ 就是隨機變量的變換（transformation of a random variable）最簡單的例子。一般而言，若 $X = h(Y)$，則 X 的密度函數可由 Y 的密度函數對所有原像（preimage）求和而得，這正是此處因子 2 的由來。同樣的想法可以證明 Z^2 服從自由度（degree of freedom）為一的卡方分佈（chi-squared distribution），而這正是許多統計學內容的起點。Grimmett

and Welsh, *Probability: An Introduction*（概率入門教材）對隨機變量的函數有仔細的論述。

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。