詳細解答

# STEP III 2017，第 13 題

概率與統計 • 期望值與方差

> **STEP III 2017，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/17-s3-q13

## 關鍵思路

把 $X - x$ 寫成 $(X - \mu) + (\mu - x)$。第一部分的平均值（mean）為零，所以平方後取期望值（expectation）時，交叉項（cross term）會消失。剩下的就是方差（variance）加上一個常數，這與「$\mathrm{E}(X^2) = \mathrm{Var}(X) + \mu^2$」是同一條恆等式。至於均勻分佈（uniform distribution）的情況，$Y$ 是 $X$ 的函數，但並非一對一（one-to-one），所以求它的概率密度函數（probability density function）最穩妥的方法，是經由它的累積分佈函數（cumulative distribution function）入手。

## 解答

### 函數 V

對任何固定的數 $x$，

$$\begin{aligned}\mathrm{V}(x) = \mathrm{E}((X-x)^2) &= \mathrm{E}(((X-\mu)+(\mu-x))^2) \\ &= \mathrm{E}((X-\mu)^2) + 2(\mu-x)\,\mathrm{E}(X-\mu) + (\mu-x)^2 \\ &= \sigma^2 + (x-\mu)^2,\end{aligned}$$

這裏用了期望值的線性性質（linearity of expectation）（此處 $\mu - x$ 是常數）以及 $\mathrm{E}(X - \mu) = 0$。

### $Y$ 的平均值

由於 $\mathrm{V}(x) = \sigma^2 + (x - \mu)^2$ 對每個實數 $x$ 都成立，我們可以代入 $x = X$：

$$Y = \mathrm{V}(X) = \sigma^2 + (X-\mu)^2.$$

**想一想** 既然 V 是由期望值定義的，$\mathrm{V}(X)$ 豈不是只是一個數？

不是：V 是一個普通的實變數函數（一個關於 $x$ 的二次式）。取期望值時 $x$ 是固定不變的。把隨機變量（random variable）$X$ 代入這個二次式，便得到一個新的隨機變量。所以

$$\mathrm{E}(Y) = \sigma^2 + \mathrm{E}((X-\mu)^2) = \sigma^2 + \sigma^2 = 2\sigma^2,$$

這就是 $(*)$。

### 均勻分佈的情況

若 $X$ 在 $[0,1]$ 上均勻分佈，則 $\mu = \frac{1}{2}$，而 $\sigma^2 = \int_0^1 x^2\,\mathrm{d}x - \frac{1}{4} = \frac{1}{3} - \frac{1}{4} = \frac{1}{12}$。因此

$$\mathrm{V}(x) = \tfrac{1}{12} + (x - \tfrac{1}{2})^2 = x^2 - x + \tfrac{1}{3}.$$

**$Y$ 的分佈。** 現在 $Y = \frac{1}{12} + (X - \frac{1}{2})^2$。當 $X$ 取遍 $[0,1]$ 時，$|X - \frac{1}{2}|$ 取遍 $[0, \frac{1}{2}]$，所以 $Y$ 的取值範圍是 $[\frac{1}{12}, \frac{1}{12} + \frac{1}{4}] = [\frac{1}{12}, \frac{1}{3}]$。

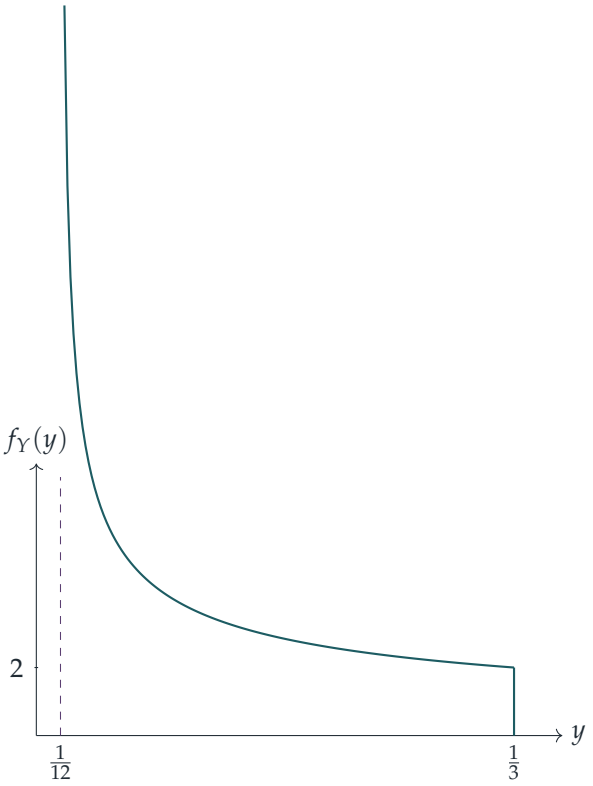
映射 $X \mapsto Y$ 是二對一的（點 $\frac{1}{2} \pm h$ 給出相同的 $Y$），所以我們改用分佈函數處理。對於 $\frac{1}{12} \le y \le \frac{1}{3}$，

$$\begin{aligned}\mathrm{P}(Y \le y) &= \mathrm{P}((X-\tfrac{1}{2})^2 \le y - \tfrac{1}{12}) = \mathrm{P}(\tfrac{1}{2} - s \le X \le \tfrac{1}{2} + s), \quad \text{其中} s = \sqrt{y - \tfrac{1}{12}},\\ &= 2s = 2\sqrt{y - \tfrac{1}{12}},\end{aligned}$$

因為 $0 \le s \le \frac{1}{2}$，所以該區間位於 $[0,1]$ 之內，長度為 $2s$。（在 $y = \frac{1}{3}$ 時，這給出 $2 \cdot \frac{1}{2} = 1$，理應如此。）求導，得

$$f_Y(y) = \frac{1}{\sqrt{y - \frac{1}{12}}} \quad \text{當} \tfrac{1}{12} < y \le \tfrac{1}{3}, \qquad f_Y(y) = 0 \text{ 其他情況。}$$

這個概率密度函數在 $y = \frac{1}{12}$ 附近無界，但仍然可積（integrable）：$Y$ 最有可能接近它的最小值，因為 $(X - \frac{1}{2})^2$ 在 $X = \frac{1}{2}$ 附近很平坦。




**驗證** $(*)$。 利用概率密度函數，並代換 $t = y - \frac{1}{12}$（於是 $t$ 由 0 變到 $\frac{1}{4}$），

$$\mathrm{E}(Y) = \int_{1/12}^{1/3} \frac{y}{\sqrt{y - \frac{1}{12}}}\,\mathrm{d}y = \int_0^{1/4} (t^{1/2} + \tfrac{1}{12}t^{-1/2})\,\mathrm{d}t = \left[\tfrac{2}{3}t^{3/2} + \tfrac{1}{6}t^{1/2}\right]_0^{1/4}$$
$$= \tfrac{2}{3}\cdot\tfrac{1}{8} + \tfrac{1}{6}\cdot\tfrac{1}{2} = \tfrac{1}{12} + \tfrac{1}{12} = \tfrac{1}{6}.$$

由於 $2\sigma^2 = 2\cdot\frac{1}{12} = \frac{1}{6}$，結果與 $(*)$ 吻合。

## 驗算

- 概率密度函數的積分等於 1：$\int_0^{1/4} t^{-1/2}\,\mathrm{d}t = 2\cdot\frac{1}{2} = 1$。
- $\mathrm{V}(x)$ 在 $x = \mu$ 處取最小值，最小值為 $\sigma^2$：平均值正是使期望平方距離最小的那一點。在均勻分佈的情況下，$\mathrm{V}(\frac{1}{2}) = \frac{1}{12}$，即 $Y$ 取值範圍的下端。
- 以 400 000 個均勻分佈的數值進行電腦模擬，得到 $Y$ 的樣本平均值約為 0.1667，與 $\frac{1}{6}$ 吻合。

## 常見錯誤

- 把 $\mathrm{V}(X)$ 當作常數 $\sigma^2$，結果得出 $\mathrm{E}(Y) = \sigma^2$。
- 使用一對一的變數變換公式 $f_Y = f_X \left|\mathrm{d}x/\mathrm{d}y\right|$ 時只計算了一個分支，結果只得到正確概率密度函數的一半。其實 $X = \frac{1}{2} + s$ 和 $X = \frac{1}{2} - s$ 兩者都有貢獻。

## STEP 之外

$\mathrm{V}(x) = \sigma^2 + (x - \mu)^2$ 是最簡單的偏差–方差分解（bias–variance decomposition）。如果你用 $x$ 去估計某個量，均方誤差（mean squared error）就會分拆為數據本身的離散程度加上偏差的平方。結果 $\mathrm{E}(Y) = 2\sigma^2$ 說明，$X$ 的兩個獨立副本之間的期望平方距離是 $2\sigma^2$。統計學家一再用到這個事實，例如在解釋為何樣本方差（sample variance）要除以 $n - 1$ 時。Blitzstein 與 Hwang 合著的 *Introduction to Probability*（《概率導論》）涵蓋了這兩個概念，也講解了如何求隨機變量的非單調函數的概率密度函數。

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。