詳細解答

# STEP III 2005，第 14 題

概率與統計 • 連續隨機變量

> **STEP III 2005，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/05-s3-q14

## 關鍵思路

這一題由始至終只靠題目容許我們直接引用的那一個積分，所以功夫全在於看出每個積分需要取哪個 $m$ 和 $n$。代換（substitution）$x = k^2/u$ 把區間 $[0, k]$ 與 $[k, \infty)$ 對調，而被積函數（integrand）的形狀保持不變，這種對稱性馬上確定了中位數（median）。至於時間 $T = s/V$，同樣的倒數技巧說明 $T$ 的概率密度（density）與 $V$ 的屬於完全相同的族，只是 $k$ 換成了 $s/k$。之後，中位數和平均值（mean）都可以從 $V$ 的結果直接讀出。

## 解答

### 常數 $C$

在已知結果中取 $m = a$ 和 $n = 2a$。這是容許的，因為 $a$ 是正整數，而且 $2a \geq a$。這時 $n + 2 = 2a + 2$，$n - m + 1 = a + 1$，所以

$$\int_0^\infty \frac{x^a}{(x+k)^{2a+2}}\,\mathrm{d}x = \frac{a!\,a!}{(2a+1)!\,k^{a+1}}.$$

密度函數在其取值範圍上的積分為 1，所以

$$1 = Ck^{a+1} \cdot \frac{a!\,a!}{(2a+1)!\,k^{a+1}} \qquad \Longrightarrow \qquad C = \frac{(2a+1)!}{a!\,a!},$$

正如所求。

### 代換與中位數

在左邊的積分中，設 $x = k^2/u$，於是 $\mathrm{d}x = -\dfrac{k^2}{u^2}\,\mathrm{d}u$。當 $x$ 由 0 增至 $v$ 時，$u$ 由 $\infty$ 減至 $k^2/v$。此外

$$x + k = \frac{k^2}{u} + k = \frac{k(u+k)}{u},$$

所以

$$\frac{x^a}{(x+k)^{2a+2}} = \frac{k^{2a}/u^a}{k^{2a+2}(u+k)^{2a+2}/u^{2a+2}} = \frac{u^{a+2}}{k^2\,(u+k)^{2a+2}}.$$

乘以 $\mathrm{d}x$ 後，$k^2/u^2$ 剛好約去，而負號則把積分上下限對調：

$$\int_0^v \frac{x^a}{(x+k)^{2a+2}}\,\mathrm{d}x = \int_\infty^{k^2/v} \frac{u^a}{(u+k)^{2a+2}}\,(-\mathrm{d}u) = \int_{k^2/v}^\infty \frac{u^a}{(u+k)^{2a+2}}\,\mathrm{d}u,$$

這正是所要證明的恆等式。

**想一想** $v$ 取哪個值時，這條恆等式會告訴我們關於中位數的事情？

取 $v = k$，於是 $k^2/v = k$ 也成立。把兩邊同乘 $Ck^{a+1}$ 後，恆等式變成

$$\begin{aligned}\mathrm{P}(V < k) &= \int_0^k \mathrm{f}(x)\,\mathrm{d}x \\ &= \int_k^\infty \mathrm{f}(u)\,\mathrm{d}u = \mathrm{P}(V > k).\end{aligned}$$

$V$ 是連續（continuous）隨機變量，所以 $\mathrm{P}(V = k) = 0$，這兩個概率之和為 1。因此兩者各為 $\frac{1}{2}$，$V$ 的中位數就是 $k$。

### $V$ 的期望值

期望值（expected value）為

$$\mathrm{E}(V) = Ck^{a+1}\int_0^\infty \frac{x^{a+1}}{(x+k)^{2a+2}}\,\mathrm{d}x.$$

現在取 $m = a+1$ 和 $n = 2a$。這需要 $2a \geq a+1$，而由於 $a \geq 1$，這是成立的。這時 $n - m + 1 = a$，並且

$$\mathrm{E}(V) = \frac{(2a+1)!}{a!\,a!}\,k^{a+1}\cdot\frac{(a+1)!\,(a-1)!}{(2a+1)!\,k^a} = k\cdot\frac{(a+1)!}{a!}\cdot\frac{(a-1)!}{a!} = \frac{(a+1)k}{a}.$$

### $T$ 的分佈

速率是正數，所以對於 $t > 0$，我們可以把不等式兩邊同乘 $V/t > 0$ 而不必改變不等號的方向：

$$T < t \iff \frac{s}{V} < t \iff V > \frac{s}{t}.$$

因此

$$\mathrm{P}(T<t)=\mathrm{P}\Big(V>\frac{s}{t}\Big)=\int_{s/t}^{\infty}\frac{Ck^{a+1}x^{a}}{(x+k)^{2a+2}}\,\mathrm{d}x, \qquad (*)$$

正如所求。

現在在 $(*)$ 中作代換 $u=s/x$，於是 $x=s/u$，$\mathrm{d}x=-\dfrac{s}{u^2}\,\mathrm{d}u$。當 $x$ 由 $s/t$ 增至 $\infty$ 時，$u$ 由 $t$ 減至 0。由於 $x+k=(s+ku)/u$，

$$\mathrm{P}(T<t)=\int_0^t Ck^{a+1}\,\frac{s^a/u^a}{(s+ku)^{2a+2}/u^{2a+2}}\cdot\frac{s}{u^2}\,\mathrm{d}u=\int_0^t\frac{Ck^{a+1}s^{a+1}\,u^a}{(s+ku)^{2a+2}}\,\mathrm{d}u.$$

這對每個 $t>0$ 都成立，而當 $t\le 0$ 時 $\mathrm{P}(T<t)=0$。所以對 $t$ 求導，便得到 $T$ 的密度函數：

$$\mathrm{g}(t)=\frac{Ck^{a+1}s^{a+1}\,t^a}{(s+kt)^{2a+2}}\qquad(t>0).$$

為了與 f 比較，從分母抽出 $k^{2a+2}$，並記 $K=s/k$：

$$\mathrm{g}(t)=\frac{C\,(s/k)^{a+1}\,t^a}{(t+s/k)^{2a+2}}=\frac{CK^{a+1}t^a}{(t+K)^{2a+2}}.$$

這正正是把 $k$ 換成 $K=s/k$ 之後的密度函數 f。所以 $T$ 與 $V$ 屬於同一個分佈（distribution）族。

### 中位數與平均值

由於 $T$ 的分佈形式相同，只是參數（parameter）為 $K=s/k$，之前的結果可以直接套用：

$$\text{就中位數而言，}T=\frac{s}{k},\qquad \mathrm{E}(T)=\frac{(a+1)}{a}\cdot\frac{s}{k}.$$

因此，時間的中位數與速率的中位數之積為

$$\frac{s}{k}\cdot k=s,$$

這就是距離，與題目所述相符。時間的期望值與速率的期望值之積為

$$\mathrm{E}(T)\,\mathrm{E}(V)=\frac{(a+1)s}{ak}\cdot\frac{(a+1)k}{a}=s\Big(1+\frac{1}{a}\Big)^2>s,$$

因為 $a$ 是正整數，所以 $1+1/a>1$。至此全題完成。

## 驗算

- 當 $a=1$，$k=1$ 時，我們得到 $C=3!/1=6$，而對 $6x/(x+1)^4$ 在 $[0,\infty)$ 上作數值積分（numerical integration）得 1.000000。在 $[0,1]$ 上則得 0.500000，印證了中位數。

- 對若干組數值作數值求積，例如 $a = 2$，$k = 0.7$，得 $\mathrm{E}(V) = 1.05000 = \frac{3}{2}(0.7)$。取 $s = 3.2$ 時，得 $\mathrm{E}(T)\mathrm{E}(V) = 7.20000 = 3.2 \times \frac{9}{4}$。
- 以 $a = 2$，$k = s = 1$ 模擬（simulation）$2 \times 10^5$ 個速率，結果是速率中位數 $\times$ 時間中位數 $\approx 1.000001$，但平均速率 $\approx 1.504$（理論值：1.5）。
- 中位數之積等於 $s$ 這一點，不用任何公式也看得出來。$T = s/V$ 是 $V$ 的遞減函數，所以 $T$ 的中位數就是 $s$ 除以 $V$ 的中位數。

## 常見錯誤

- 引用已知積分時忘記檢查 $n \geq m$。對 $\mathrm{E}(V)$ 而言，這需要 $a \geq 1$，這正是 $a$ 必須為正整數的原因。對於更高階的矩（moment），公式最終會失效，事實上 $\mathrm{E}(V^{a+1})$ 是無限大。
- 對 $(*)$ 求導以求 $T$ 的密度函數時，漏掉因子 $s/t^2$（鏈式法則，chain rule）。
- 以為 $\mathrm{E}(s/V) = s/\mathrm{E}(V)$。最後一部分的重點，正是要指出這是錯的。

## STEP 之外

最後的不等式是琴生不等式（Jensen's inequality）的一個例子：當 $v > 0$ 時 $1/v$ 是凸（convex）函數，所以 $\mathrm{E}(1/V) \geq 1/\mathrm{E}(V)$，等號僅在 $V$ 為常數時成立。相比之下，中位數卻可以直接穿過任何單調（monotonic）函數。代換 $V/(V+k)$ 把 $V$ 變成一個 $\mathrm{Beta}(a+1, a+1)$ 變量，它關於 $\frac{1}{2}$ 對稱，這是理解中位數為何是 $k$ 的另一個途徑。Blitzstein 與 Hwang 的 *Introduction to Probability*（概率論入門）對琴生不等式、變量代換和 Beta 分佈族都有十分易讀的講解。

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。