详细解答

STEP III 2005，第 14 题

概率与统计 · 连续随机变量

> **STEP III 2005，第 14 题**
>
> 继续阅读之前，请先看题目（原题为英文）。题目载于 STEP Support Programme 的试题资料库：
>
> step.maths.org/questions/05-s3-q14

关键思路

这一题由始至终只靠题目容许我们直接引用的那一个积分，所以功夫全在于看出每个积分需要取哪个 m 和 n。代换（substitution）$x = k^2/u$ 把区间 $[0,k]$ 与 $[k,\infty)$ 对调，而被积函数（integrand）的形状保持不变，这种对称性马上确定了中位数（median）。至于时间 $T = s/V$，同样的倒数技巧说明 T 的概率密度（density）与 V 的属于完全相同的族，只是 k 换成了 s/k。之后，中位数和平均值（mean）都可以从 V 的结果直接读出。

解答

常数 C

在已知结果中取 $m = a$ 和 $n = 2a$。这是容许的，因为 a 是正整数，而且 $2a \geq a$。这时 $n+2 = 2a+2$，$n-m+1 = a+1$，所以

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

密度函数在其取值范围上的积分为 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。

代换与中位数

在左边的积分中，设 $x = k^2/u$，于是 $\mathrm{d}x = -\dfrac{k^2}{u^2}\,\mathrm{d}u$。当 x 由 0 增至 v 时，u 由 ∞ 减至 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 $\mathrm{d}x$ 后，k^2/u^2 刚好约去，而负号则把积分上下限对调：

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

这正是所要证明的恒等式。

> **想一想** v 取哪个值时，这条恒等式会告诉我们关于中位数的事情?

取 $v = k$，于是 $k^2/v = k$ 也成立。把两边同乘 Ck^{a+1} 后，恒等式变成

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

V 是连续（continuous）随机变量，所以 $\mathrm{P}(V = k) = 0$，这两个概率之和为 1。因此两者各为 $\frac{1}{2}$，V 的中位数就是 k。

V 的期望值

期望值（expected value）为

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

现在取 $m = a+1$ 和 $n = 2a$。这需要 $2a \geq a+1$，而由于 $a \geq 1$，这是成立的。这时 $n - m + 1 = a$，并且

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

T 的分布

速率是正数，所以对于 $t > 0$，我们可以把不等式两边同乘 $V/t > 0$ 而不必改变不等号的方向：

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

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。

现在在 $(*)$ 中作代换 $u=s/x$，于是 $x=s/u$，$\mathrm{d}x=-\dfrac{s}{u^2}\,\mathrm{d}u$。当 x 由 s/t 增至 ∞ 时，u 由 t 减至 0。由于 $x+k=(s+ku)/u$，

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

这对每个 $t>0$ 都成立，而当 $t\le 0$ 时 $\mathrm{P}(T<t)=0$。所以对 t 求导，便得到 T 的密度函数：

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

为了与 f 比较，从分母抽出 k^{2a+2}，并记 $K=s/k$：

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

这正正是把 k 换成 $K=s/k$ 之后的密度函数 f。所以 T 与 V 属于同一个分布（distribution）族。

中位数与平均值

由于 T 的分布形式相同，只是参数（parameter）为 $K=s/k$，之前的结果可以直接套用：

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

因此，时间的中位数与速率的中位数之积为

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

这就是距离，与题目所述相符。时间的期望值与速率的期望值之积为

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

因为 a 是正整数，所以 $1+1/a>1$。至此全题完成。

验算

- 当 $a=1$，$k=1$ 时，我们得到 $C=3!/1=6$，而对 $6x/(x+1)^4$ 在 $[0,\infty)$ 上作数值积分（numerical integration）得 1.000000。在 $[0,1]$ 上则得 0.500000，印证了中位数。

- 对若干组数值作数值求积，例如 $a = 2$，$k = 0.7$，得 $\mathrm{E}(V) = 1.05000 = \frac{3}{2}(0.7)$。取 $s = 3.2$ 时，得 $\mathrm{E}(T)\mathrm{E}(V) = 7.20000 = 3.2 \times \frac{9}{4}$。
- 以 $a = 2$，$k = s = 1$ 模拟（simulation）2×10^5 个速率，结果是速率中位数 $\times$ 时间中位数 ≈ 1.000001，但平均速率 ≈ 1.504（理论值：1.5）。
- 中位数之积等于 s 这一点，不用任何公式也看得出来。$T = s/V$ 是 V 的递减函数，所以 T 的中位数就是 s 除以 V 的中位数。

常见错误

- 引用已知积分时忘记检查 $n \geq m$。对 $\mathrm{E}(V)$ 而言，这需要 $a \geq 1$，这正是 a 必须为正整数的原因。对于更高阶的矩（moment），公式最终会失效，事实上 $\mathrm{E}(V^{a+1})$ 是无限大。
- 对 $(*)$ 求导以求 T 的密度函数时，漏掉因子 s/t^2（链式法则，chain rule）。
- 以为 $\mathrm{E}(s/V) = s/\mathrm{E}(V)$。最后一部分的重点，正是要指出这是错的。

STEP 之外

最后的不等式是琴生不等式（Jensen's inequality）的一个例子：当 $v > 0$ 时 $1/v$ 是凸（convex）函数，所以 $\mathrm{E}(1/V) \geq 1/\mathrm{E}(V)$，等号仅在 V 为常数时成立。相比之下，中位数却可以直接穿过任何单调（monotonic）函数。代换 $V/(V+k)$ 把 V 变成一个 $\mathrm{Beta}(a+1, a+1)$ 变量，它关于 $\frac{1}{2}$ 对称，这是理解中位数为何是 k 的另一个途径。Blitzstein 与 Hwang 的 *Introduction to Probability*（概率论入门）对琴生不等式、变量代换和 Beta 分布族都有十分易读的讲解。

Beyond STEP 详细解答：https://beyondstep.keithmcnulty.org/zh-hans/。解答文字采用 CC0 1.0；题目本身版权属 Cambridge University Press & Assessment 所有（©），本文件不予转载。由机器从英文翻译而成；尚未经审阅。