解答例

STEP III 2005、第 14 問

確率・統計 ・ 連続型確率変数

STEP III 2005、問 14

先に進む前に、まず問題を読んでおこう。問題は STEP Support Programme の問題データベースにある：

step.maths.org/questions/05-s3-q14

鍵となる考え方

ここでの議論はすべて、引用してよいとされている一つの積分に基づいている。したがって、作業の中心は各積分にどの m と n が必要かを見抜くことにある。置換（substitution）$x = k^2/u$ は、被積分関数（integrand）の形を変えずに区間 $[0, k]$ と $[k, \infty)$ を入れ替える。この対称性から中央値（median）がただちに定まる。時間 $T = s/V$ についても、同じ逆数の置換によって、T が V とまったく同じ族の確率密度関数（probability density function）をもち、k が s/k に置き換わっただけであることがわかる。そうなれば、中央値と平均は V についての結果からそのまま読み取れる。

解答

定数 C

与えられた結果で $m = a$、$n = 2a$ とする。a は正の整数であり $2a \geq a$ であるから、これは許される。このとき $n + 2 = 2a + 2$、$n - m + 1 = a + 1$ であるから、

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

確率密度関数をその定義域全体で積分すると 1 になるので、

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

となり、示すべき式が得られた。

置換と中央値

左辺の積分において $x = k^2/u$ とおくと、$\mathrm{d}x = -\dfrac{k^2}{u^2}\,\mathrm{d}u$ である。x が 0 から v まで動くとき、u は ∞ から k^2/v まで減少する。また

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

であるから、

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

$\mathrm{d}x$ を掛けると k^2/u^2 が打ち消され、負号によって積分の上端と下端が入れ替わる：

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

これが示すべき恒等式である。

> **考えてみよう** v をどの値にとれば、この恒等式が中央値について何かを語るだろうか。

$v = k$ とすると、$k^2/v = k$ にもなる。両辺に Ck^{a+1} を掛けると、恒等式は次のようになる。

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

V は連続型であるから $\mathrm{P}(V = k) = 0$ であり、この二つの確率の和は 1 である。よってそれぞれ $\frac{1}{2}$ であり、V の中央値は k である。

V の期待値

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

今度は $m = a+1$、$n = 2a$ とする。これには $2a \geq a+1$ が必要であるが、$a \geq 1$ よりこれは成り立つ。このとき $n - m + 1 = a$ であり、

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

となる。これが期待値（expected value）である。

T の分布

速さは正であるから、$t > 0$ のとき、不等号の向きを変えずに両辺に $V/t > 0$ を掛けることができる：

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

したがって

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

となり、示すべき式が得られた。

次に、$(*)$ において $u = s/x$ と置換する。すると $x = s/u$、$\mathrm{d}x = -\dfrac{s}{u^2}\,\mathrm{d}u$ である。x が s/t から ∞ まで動くとき、u は t から 0 まで減少する。$x + k = (s + ku)/u$ であるから、

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

これはすべての $t > 0$ について成り立ち、また $t \le 0$ のとき $\mathrm{P}(T < t) = 0$ である。よって t で微分すると T の確率密度関数が得られる：

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

これを f と比較するため、分母から k^{2a+2} をくくり出し、$K = s/k$ とおく：

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

これはまさに確率密度関数 f の k を $K = s/k$ に置き換えたものである。よって T は V と同じ分布族（family of distributions）に属する。

中央値と平均

T はパラメータ $K = s/k$ をもつ同じ形の分布に従うので、先の結果がそのまま適用できる：

$$\text{中央値}T = \frac{s}{k}, \qquad \mathrm{E}(T) = \frac{(a+1)}{a} \cdot \frac{s}{k}.$$

したがって、時間の中央値と速さの中央値の積は

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

であり、主張どおり距離に等しい。時間の期待値と速さの期待値の積は

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

である。ここで、a は正の整数であるから $1 + 1/a > 1$ となることを用いた。以上で問題は完結する。

検算

- $a = 1$、$k = 1$ のとき $C = 3!/1 = 6$ であり、$6x/(x+1)^4$ を $[0,\infty)$ で数値積分（numerical integration）すると 1.000000 となる。$[0,1]$ で積分すると 0.500000 となり、中央値が確かめられる。
- $a = 2$、$k = 0.7$ などいくつかの組について数値積分すると、$\mathrm{E}(V) = 1.05000 = \frac{3}{2}(0.7)$ が得られる。$s = 3.2$ とすると $\mathrm{E}(T)\mathrm{E}(V) = 7.20000 = 3.2 \times \frac{9}{4}$ となる。
- $a = 2$、$k = s = 1$ として 2×10^5 個の速さをシミュレーションすると、速さの中央値 × 時間の中央値 ≈ 1.000001 となったが、速さの平均は ≈ 1.504（理論値：1.5）であった。
- 中央値の積が s に等しいことは、式を使わなくてもわかる。$T = s/V$ は V の減少関数であるから、T の中央値は s を V の中央値で割ったものになる。

よくある間違い

- 与えられた積分を引用するときに $n \geq m$ の確認を忘れること。$\mathrm{E}(V)$ については $a \geq 1$ が必要であり、これこそ a が正の整数でなければならない理由である。より高次の積率（moment）では、いずれこの公式は使えなくなり、実際 $\mathrm{E}(V^{a+1})$ は無限大になる。
- T の確率密度関数を求めるために $(*)$ を微分するとき、因子 s/t^2（合成関数の微分法、chain rule）を落とすこと。
- $\mathrm{E}(s/V) = s/\mathrm{E}(V)$ と思い込むこと。最後の小問の要点は、まさにこれが誤りであるということにある。

STEP のその先へ

最後の不等式はイェンセンの不等式（Jensen's inequality）の一例である。$v > 0$ において $1/v$ は凸（convex）であるから $\mathrm{E}(1/V) \geq 1/\mathrm{E}(V)$ が成り立ち、等号は V が定数のときに限る。これに対して中央値は、任意の単調関数（monotonic function）でそのまま移される。置換 $V/(V+k)$ によって V は $\frac{1}{2}$ に関して対称なベータ分布（Beta distribution）$\mathrm{Beta}(a+1, a+1)$ に従う確率変数に変わり、これも中央値が k である理由を理解するもう一つの方法である。Blitzstein と Hwang の *Introduction to Probability*（確率論入門）は、イェンセンの不等式、変数変換、ベータ分布族をたいへん読みやすく解説している。

Beyond STEP 解答集：https://beyondstep.keithmcnulty.org/ja/。解答本文は CC0 1.0 である。問題そのものは © Cambridge University Press & Assessment に帰属し、ここには掲載していない。英語版から機械翻訳したものであり、まだ校閲を経ていない。