解答例

STEP II 1999、第 13 問

確率・統計 ・ 連続型確率変数

> **STEP II 1999、問 13**
>
> 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/99-s2-q13

鍵となる考え方

力学の分野に分類されてはいるが、これは連続確率変数（continuous random variable）に関する問題である。**短いほうの**破片が単純な分布に従うことに気づけば、すべてが易しくなる。折れる点が棒に沿って一様に分布するならば、短いほうの破片の長さは全長の $[0, \frac{1}{2}]$ 上の一様分布（uniform distribution）に従う。比 R は短いほうの破片の長さの増加関数であるから、「$R < r$」は短いほうの破片だけについての条件に書き換えられる。これにより分布関数（distribution function）が得られ、それを微分すれば密度が得られる。

解答

R の分布

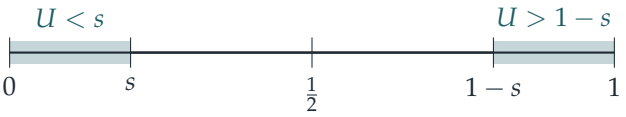
棒の長さを 1 とする。比は単位によらないからである。折れる点を一方の端から距離 U の位置とし、U は $[0,1]$ 上の一様分布に従うとする。2 つの破片の長さは U と $1-U$ であるから、短いほうの長さは

$$S = \min(U, 1-U),$$

であり、長いほうの長さは $1-S$ である。このとき $0 \le S \le \frac{1}{2}$ であり、

$$R = \frac{S}{1-S}.$$

$0 \le s \le \frac{1}{2}$ に対して、事象 $S < s$ が起こるのは、ちょうど $U < s$ または $U > 1-s$ となるときである。これらは互いに交わらない 2 つの区間で、長さの合計は $2s$ であるから、$\mathrm{P}(S < s) = 2s$ となる。すなわち、S は $[0, \frac{1}{2}]$ 上の一様分布に従う。



次に $0 \le r \le 1$ とする。関数 $s \mapsto \dfrac{s}{1-s}$ は $[0, \frac{1}{2}]$ 上で狭義単調増加である（導関数は $1/(1-s)^2 > 0$）から、不等式を次のように変形できる。

$$R < r \iff \frac{S}{1-S} < r \iff S < r(1-S) \iff S < \frac{r}{1+r},$$

ここで $1 - S \ge \frac{1}{2} > 0$ であるから、両辺に $1 - S$ を掛けても問題ない。$0 \le r \le 1$ のとき値 $\dfrac{r}{1+r}$ は $[0, \frac{1}{2}]$ に属するので、

$$\boxed{\mathrm{P}(R < r) = \frac{2r}{1+r} \qquad (0 \le r \le 1).}$$

短いほうの破片が長いほうより長くなることはないから、R はつねに $[0,1]$ に属する。よって $r \le 0$ のとき $\mathrm{P}(R < r) = 0$、$r \ge 1$ のとき $\mathrm{P}(R < r) = 1$ であり、これは $r = 1$ における公式の値 1 と一致する。

考えてみよう なぜ $R = U/(1-U)$ として U を直接扱わないのか？

それが正しいのは $U \le \frac{1}{2}$ のときだけである。$U > \frac{1}{2}$ のときは短いほうの破片は $1 - U$ であり、$R = (1-U)/U$ となる。この 2 つの場合に分けて考えることもでき、対称性からそれぞれが $\frac{r}{1+r}$ ずつ寄与する。S に移って考えれば、この場合分けの処理を一度で済ませられるのである。

密度、平均、分散

分布関数 $F(r) = \dfrac{2r}{1+r} = 2 - \dfrac{2}{1+r}$ を $(0,1)$ 上で微分すると、確率密度関数（probability density function）

$$f(r) = \frac{2}{(1+r)^2} \quad (0 \le r \le 1), \qquad f(r) = 0 \text{（それ以外）}.$$

が得られる。検算として、$\int_0^1 f(r)\,\mathrm{d}r = F(1) - F(0) = 1$ である。

平均（mean）を求めるために、$r = (1+r) - 1$ と書く。

$$\begin{aligned}
\mathrm{E}(R) &= \int_0^1 \frac{2r}{(1+r)^2}\,\mathrm{d}r = 2\int_0^1 \left(\frac{1}{1+r} - \frac{1}{(1+r)^2}\right)\mathrm{d}r \\
&= 2\left[\ln(1+r) + \frac{1}{1+r}\right]_0^1 = 2\left(\ln 2 + \tfrac{1}{2} - 1\right) \\
&= 2\ln 2 - 1 \approx 0.386.
\end{aligned}$$

2 次のモーメント（second moment）を求めるために、$r^2 = (1+r)^2 - 2(1+r) + 1$ と書くと、

$$\frac{r^2}{(1+r)^2} = 1 - \frac{2}{1+r} + \frac{1}{(1+r)^2}.$$

となる。すると

$$\begin{aligned}\mathrm{E}(R^2) &= 2\int_0^1 \left(1 - \frac{2}{1+r} + \frac{1}{(1+r)^2}\right) \mathrm{d}r = 2\Big[r - 2\ln(1+r) - \frac{1}{1+r}\Big]_0^1 \\ &= 2\left(1 - 2\ln 2 - \tfrac{1}{2} + 1\right) = 3 - 4\ln 2.\end{aligned}$$

したがって、分散（variance）は

$$\begin{aligned}\mathrm{Var}(R) &= \mathrm{E}(R^2) - \big(\mathrm{E}(R)\big)^2 = 3 - 4\ln 2 - \big(4(\ln 2)^2 - 4\ln 2 + 1\big) \\ &= 2 - 4(\ln 2)^2 \approx 0.0782.\end{aligned}$$

$\ln 2$ の 1 次の項が打ち消し合い、このすっきりした形が残る。

検算

- $r = 1$ のとき確率は 1 であり、当然そうなるべきである。$r = \frac{1}{2}$ のときは $\frac{2}{3}$ である。長いほうの破片が短いほうの 2 倍より長くなるのは、ちょうど折れる点が棒の両端の 3 分の 1 ずつの部分に入るときであり、その長さの合計は $\frac{2}{3}$ である。
- 平均 0.386 は 0 と 1 の間にあり、しかも $\frac{1}{2}$ より小さい。これは、密度が $r = 0$ での 2 から $r = 1$ での $\frac{1}{2}$ まで減少することと整合する。
- 10^6 回の無作為な折り取りをシミュレーションしたところ、標本平均は 0.3862、標本分散は 0.0783 となった。理論値は $2\ln 2 - 1 = 0.38629$ と $2 - 4(\ln 2)^2 = 0.07819$ である。

よくある間違い

- すべての U に対して $R = U/(1-U)$ とすること。この量は 1 を超えることがあるので、短いほうと長いほうの比にはならない。
- 折れる点がどちらの端の近くにもなりうることを忘れて、$\mathrm{P}(R < r) = \frac{r}{1+r}$ としてしまうこと。
- 短いほうの破片の平均 $\frac{1}{4}$ を使って $\mathrm{E}(R) = \frac{1/4}{3/4} = \frac{1}{3}$ とすること。一般に $\mathrm{E}\big(g(S)\big) \neq g\big(\mathrm{E}(S)\big)$ である。

STEP のその先へ

ここで用いた方法は、確率変数を変換するための標準的な「分布関数法（distribution function technique）」である。g が増加関数で $Y = g(X)$ ならば $F_Y(y) = F_X\left(g^{-1}(y)\right)$ であり、これを微分すると変数変換の公式（change-of-variables formula）$f_Y(y) = f_X\left(g^{-1}(y)\right)\left|\mathrm{d}g^{-1}/\mathrm{d}y\right|$ が得られる。同じ考え方は、順序統計量（order statistics）や、3 つの破片で三角形が作れる確率を問うような古典的な「折れた棒の問題（broken-stick problem）」へとつながっていく。連続確率変数の変換について学ぶには、Grimmett と Welsh の *Probability: An Introduction*（確率論入門）がよい。

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。