解答例

STEP I 1997、第 14 問

確率・統計 ・ 期待費用の最小化

STEP I 1997、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/97-s1-q14

鍵となる考え方

費用 T は確率変数（random variable）X の関数であるから、その期待値（expectation）はその関数に確率密度関数（density）を掛けて積分したものになる。積分を $x = y$ で分割すると、$\mathrm{E}(T)$ が y の普通の関数として得られる。あとはそれを最小化するだけで、これは大学 1 年の微積分の演習問題である。すなわち、微分して 0 とおき、停留点（stationary point）が本当に最小値を与えることを確かめればよい。

解答

準備量が負になることは意味をなさないので、$y \geq 0$ とする。$T = g(X)$ と書く。ここで

$$g(x) = \begin{cases} y & x \leq y, \\ y + r + s(x - y) & x > y. \end{cases}$$

である。すると

$$\begin{aligned} \mathrm{E}(T) &= \int_0^\infty g(x)\,\mathrm{e}^{-x}\,\mathrm{d}x \\ &= \int_0^\infty y\,\mathrm{e}^{-x}\,\mathrm{d}x + \int_y^\infty (r + s(x - y))\mathrm{e}^{-x}\,\mathrm{d}x, \end{aligned}$$

となる。ここでは、何が起ころうと支払う費用 y を分けて取り出した。確率密度関数の積分は 1 であるから、第 1 の積分は y である。

第 2 の積分については $u = x - y$ と置換する。このとき $\mathrm{e}^{-x} = \mathrm{e}^{-y}\mathrm{e}^{-u}$ であり、u は 0

から ∞ まで動く：

$$\begin{aligned}\int_y^\infty (r+s(x-y))\mathrm{e}^{-x}\,\mathrm{d}x &= \mathrm{e}^{-y}\int_0^\infty (r+su)\,\mathrm{e}^{-u}\,\mathrm{d}u \\ &= \mathrm{e}^{-y}\Big(r\int_0^\infty \mathrm{e}^{-u}\,\mathrm{d}u + s\int_0^\infty u\,\mathrm{e}^{-u}\,\mathrm{d}u\Big) \\ &= \mathrm{e}^{-y}(r+s),\end{aligned}$$

ここで $\int_0^\infty u\,\mathrm{e}^{-u}\,\mathrm{d}u = [-u\mathrm{e}^{-u}]_0^\infty + \int_0^\infty \mathrm{e}^{-u}\,\mathrm{d}u = 1$（部分積分（integration by parts））を用いた。したがって

$$C(y) = \mathrm{E}(T) = y + (r+s)\,\mathrm{e}^{-y}.$$

考えてみよう 積分せずに因子 $\mathrm{e}^{-y}(r+s)$ を見抜く方法はあるだろうか？

ある。洪水被害が起こる確率は $\mathrm{P}(X > y) = \mathrm{e}^{-y}$ である。被害が起こったという条件のもとで、超過分 $X - y$ は再び平均 1 の指数分布（exponential distribution）に従う（指数分布は**無記憶性**（memoryless）をもつ）。よって被害額の期待値は $r + s\cdot 1$ である。これらを掛け合わせると $(r+s)\mathrm{e}^{-y}$ が得られる。

次に最小化する。微分すると

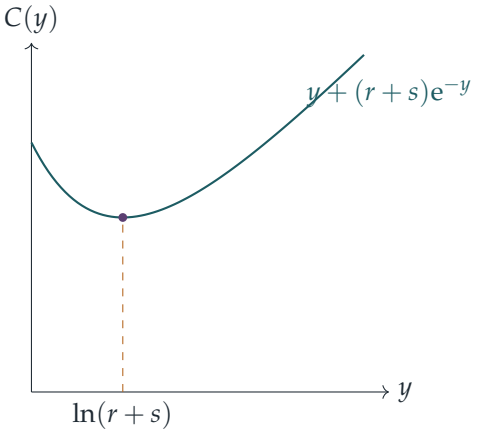
$$C'(y) = 1 - (r+s)\mathrm{e}^{-y}, \qquad C''(y) = (r+s)\mathrm{e}^{-y} > 0.$$

よって $C'(y) = 0$ となるのは $\mathrm{e}^y = r+s$、すなわち $y = \ln(r+s)$ のときに限る。すべての y で $C'' > 0$ であるから、C は凸（convex）であり、C' は増加関数である。したがって C は $y < \ln(r+s)$ で減少し、$y > \ln(r+s)$ で増加する。ゆえにこの停留点は大域的最小点（global minimum）であり、総費用の期待値は

$$y = \ln(r+s),$$

とすることで最小になる。これが示すべきことであった。最小の期待費用は $C(\ln(r+s)) = \ln(r+s) + 1$ である。

これには $\ln(r+s) \geq 0$、すなわち $r+s \geq 1$ が必要であり、問題文ではこれが暗に仮定されている。$r+s < 1$ ならば、すべての $y \geq 0$ に対して $C'(y) > 0$ となるので、最善の選択は $y = 0$ である。つまり被害があまりに安くつくので、そもそも準備する価値がないのである。



($r+s=3$ として描いた概形。)

検算

- $y=0$ のときは必ず被害を支払うことになり、$\mathrm{E}(T) = r + s\,\mathrm{E}(X) = r+s$ である。これは $C(0)$ と一致する。
- $r=2, s=3$ のとき、公式は $y = \ln 5 \approx 1.609$、最小費用 $1+\ln 5 \approx 2.609$ を与える。$C(y)$ を細かい格子で探索しても同じ最小点が見つかり、また $y=1$ で 4×10^5 年分のシミュレーションを行うと平均費用は 2.844 となり、公式の値 $1+5\mathrm{e}^{-1} \approx 2.839$ とよく合う。
- $r+s$ が大きくなると y も増加する。洪水の被害が高くつくほど多くの準備が正当化されるのであり、常識とも合致する。

よくある間違い

- 準備費用 y は $X \le y$ のときだけでなく、**毎年**支払われることを忘れる。
- $sx\,\mathrm{e}^{-x}$ を積分する際に $-sy\,\mathrm{e}^{-x}$ の項を落とし、$\mathrm{e}^{-y}(r+s)$ の代わりに $\mathrm{e}^{-y}(r+s+sy)$ を得てしまう。
- 停留点を求めただけで、それが最小値であることを示さない。

STEP のその先へ

これは不確実性のもとでの意思決定（decision-making under uncertainty）の小さな例である。すなわち、期待損失（expected loss）を最小にする行動を選ぶ問題である。確実にかかる費用と期待される損失とを釣り合わせるという同じ構造は、オペレーションズ・リサーチ（operations research）における「新聞売り子問題」（newsvendor problem）や、統計学におけるベイズ決定（Bayes decision）の考え方の根底にもある。上で用いた無記憶性は、連続分布の中で指数分布を特徴づける性質である。Grimmett and Welsh, *Probability: An Introduction*（確率論入門）では、これが確率変数の関数の期待値とともに扱われている。

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。