解答例

STEP I 2017、第 12 問

確率・統計 ・ 期待値と条件付け

> **STEP I 2017、問 12**
>
> 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/17-s1-q12

鍵となる考え方

主催者が受け取る売上は常に同じ Nc であり、不確かなのは賞金を支払う必要があるかどうかだけである。したがって期待利益（expected profit）は、売上から、J と「誰かが当選番号を選んだ」確率との積を引いたものであり、すべては「誰も選ばなかった」という余事象（complementary event）に帰着する。その確率は累乗 $(1-p)^N$ であり、これはまさに近似式 $(*)$ が扱うために用意された形である。

解答

小問 (i)

当選番号を W とし、どの参加者も W を選ばないという事象を E とする。W が何であっても、各参加者は独立に（independently）確率 $1-\frac{1}{N}$ でそれを避けるので、

$$\mathrm{P}(E) = \left(1-\frac{1}{N}\right)^N.$$

主催者はどの場合でも Nc を受け取る。E が起これば当選者はおらず、利益は Nc である。そうでなければ賞金は当選者の間で分けられるが、当選者が何人であっても支払総額は J なので、利益は $Nc-J$ である。よって

$$\text{期待利益} = Nc - J\,\mathrm{P}(E^c) = Nc - J\left[1-\left(1-\frac{1}{N}\right)^N\right].$$

$a=1$ として $(*)$ を用いると $(1-\frac{1}{N})^N \approx \mathrm{e}^{-1}$ なので、期待利益はおよそ $Nc-J(1-\mathrm{e}^{-1})$

である。$J = 2Nc$ ならば、これは

$$Nc - 2Nc\left(1 - \mathrm{e}^{-1}\right) = Nc\left(\frac{2}{\mathrm{e}} - 1\right).$$

となる。$\mathrm{e} > 2$ より $2/\mathrm{e} < 1$ なので、これは負である。すなわち、主催者は損失を出すと期待される。

小問 (ii)

人気のある番号は γN 個、人気のない番号は $(1-\gamma)N$ 個ある。各参加者はちょうど 1 つの番号を選ぶので、N 個すべての番号にわたる確率の和は 1 である：

$$\gamma N \cdot \frac{a}{N} + (1-\gamma)N \cdot \frac{b}{N} = 1 \quad \Longrightarrow \quad \gamma a + (1-\gamma)b = 1.$$

主催者は依然として W を一様に（uniformly）選ぶので、W は確率 γ で人気のある番号、確率 $1-\gamma$ で人気のない番号である。

> **考えてみよう** 参加者はもはや一様には選んでいない。そのことは $\mathrm{P}(E)$ の求め方を変えるだろうか？

変わるのは W を避ける確率だけであり、それはいまや W がどちらの種類の番号であるかに依存する。そこで、その種類で条件付けをする。W が人気のある番号ならば、各参加者は独立に確率 $1 - \frac{a}{N}$ でそれを避け、人気のない番号ならば確率 $1 - \frac{b}{N}$ で避ける。全確率の公式（law of total probability）と $(*)$ により、

$$\mathrm{P}(E) = \gamma\left(1 - \frac{a}{N}\right)^N + (1-\gamma)\left(1 - \frac{b}{N}\right)^N \approx \gamma\,\mathrm{e}^{-a} + (1-\gamma)\,\mathrm{e}^{-b}.$$

小問 (i) と同様に期待利益は $Nc - J(1 - \mathrm{P}(E))$ なので、

$$\text{期待利益} \approx \gamma J\,\mathrm{e}^{-a} + (1-\gamma)J\,\mathrm{e}^{-b} + (Nc - J),$$

となり、これは

$$A = \gamma J, \qquad B = (1-\gamma)J, \qquad C = Nc - J.$$

とした求める形である。

$\gamma = \frac{1}{8}$、$a = 9b$ の場合。 関係式から $\frac{1}{8} \cdot 9b + \frac{7}{8}b = 1$、すなわち $2b = 1$ となるので、

$$b = \tfrac{1}{2}, \qquad a = \tfrac{9}{2}.$$

$J = 2Nc$ のとき $A = \frac{1}{4}Nc$、$B = \frac{7}{4}Nc$、$C = -Nc$ なので、期待利益はおよそ

$$Nc\left(\tfrac{1}{4}\mathrm{e}^{-9/2} + \tfrac{7}{4}\mathrm{e}^{-1/2} - 1\right).$$

である。括弧内の第 1 項は正なので、$\frac{7}{4}e^{-1/2} > 1$、すなわち $\sqrt{e} < \frac{7}{4}$ を示せば十分である。$e < 3$ より

$$\sqrt{e} < \sqrt{3} < \frac{7}{4},$$

であり、最後の不等式は両辺が正で $3 < \frac{49}{16}$（すなわち $48 < 49$）であることから成り立つ。よって括弧内は正であり、主催者は利益を出すと期待される。賞金は小問 (i) と同じだが、当選番号が人気のない番号であることが多いため、受け取り手のいないことがより多く起こるのである。

検算

- $J = 2Nc$ のとき、小問 (i) の期待利益はおよそ $(2/e - 1)Nc \approx -0.264\,Nc$ であり、最後の部分では $\left(\frac{1}{4}e^{-4.5} + \frac{7}{4}e^{-0.5} - 1\right)Nc \approx 0.064\,Nc$ となる。主張どおり、わずかな利益である。
- $N = 800$ のとき、正確な累乗 $(1 - a/N)^N$ を用いると $0.0640\,Nc$ となるので、ここでは近似 $(*)$ は非常によい。
- $\gamma = 1$（すべての番号が「人気」）ならば、関係式から $a = 1$ となり、小問 (ii) の式は小問 (i) の $Je^{-1} + Nc - J$ に帰着する。
- $N = 4$ の場合に（参加者の選び方 4^4 通りすべてを）コンピュータで正確に数え上げると、近似前の 2 つの式の両方と一致する。

よくある間違い

- 賞金の分け方を気にしてしまうこと。当選者が少なくとも 1 人いれば、その人数にかかわらず、主催者は合計 J を支払う。
- 当選番号が人気のある番号である確率が a や b に依存すると考えてしまうこと。主催者は一様に選ぶので、その確率は γ である。
- 各種類の番号がいくつあるかを数えるのを忘れて、関係式を $a + b = 1$ や $\gamma a + b = 1$ と書いてしまうこと。

STEP のその先へ

当選番号を選ぶ参加者の人数は、パラメータ N と p の二項分布（binomial distribution）に従う。そして $(*)$ は、N が大きく Np が一定のとき、それが近似的にポアソン分布（Poisson distribution）に従うことを理解するための第一歩である。$(1 - p)^N \to e^{-Np}$

は、ちょうどポアソン分布で値が o となる確率にほかならない。キャリーオーバーのある実際の宝くじもまさにこの方法で分析されており、人気のない番号を選ぶプレイヤーは実際に期待収益が高い。Grimmett と Welsh の *Probability: An Introduction*（確率論入門）では、ポアソン極限や、ここで用いたような条件付けによる議論が展開されている。

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。