解答例

STEP I 1993、第 15 問

確率・統計 ・ ベイズの定理

> **STEP I 1993、問 15**
>
> 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/93-s1-q15

鍵となる考え方

これは物語の衣をまとったベイズの定理（Bayes' theorem）である。要点は、**真の**方向で条件付けし、それぞれの場合に住民の一連の答えがどれほど起こりやすいかを考えることにある。Biscii は常に嘘をつくので、Biscii の繰り返しの答えは常に一致し、しかも常に誤った方向を指す。したがって、同じ答えが繰り返されるほど、相手が Biscii である可能性が高まる。もう一度尋ねるだけで判断が逆転しうるのは、このためである。

解答

Ascii だけの場合

Ascii は各質問に確率 $\frac{3}{4}$ で正直に答え、それは以前の答えとは独立である。質問の答えは 2 通り（左か右）しかないので、2 つの答えが一致するのは、両方とも真実であるか、両方とも嘘であるときに限る。よって

$$\mathrm{P}(\text{同じ答えを 2 回}) = \left(\tfrac{3}{4}\right)^2 + \left(\tfrac{1}{4}\right)^2 = \tfrac{9}{16} + \tfrac{1}{16} = \tfrac{10}{16} = \tfrac{5}{8},$$

となり、求める結果が得られる。このうち「両方とも真実」の場合が $\frac{9}{16}$ を占めるので、条件付き確率（conditional probability）の定義より

$$\mathrm{P}(\text{真実を言っている} \mid \text{同じ答えを 2 回}) = \frac{9/16}{10/16} = \frac{9}{10},$$

となる。これが 2 つ目の結果である。

「左」という答えが 2 回の場合

左が正しい方向であるという事象を L、右が正しい方向であるという事象を R とする。また、住民が Ascii である事象を A、Biscii である事象を B とし、$\mathrm{P}(A) = \frac{11}{16}$、$\mathrm{P}(B) = \frac{5}{16}$ で、これは地理とは独立に選ばれるものとする。船長は初めは方向について何も知らないので、$\mathrm{P}(L) = \mathrm{P}(R) = \frac{1}{2}$ とおく。住民が「左」と 2 回答える事象を E_2 とする。

考えてみよう なぜ $\mathrm{P}(L \mid E_2)$ よりも $\mathrm{P}(E_2 \mid L)$ のほうが求めやすいのか。

真の方向がわかれば、各部族のふるまいは完全に決まるので、答えの確率は容易に書き下せる。そのうえでベイズの定理を用いて、条件を逆向きにすればよい。

左が正しい場合、Ascii が「左」と 2 回答えるのは 2 回とも真実を言うときであり、その確率は $\frac{9}{16}$ である。Biscii は嘘をついて「右」と言うので、E_2 を生じえない。よって

$$\mathrm{P}(E_2 \mid L) = \frac{11}{16} \cdot \frac{9}{16} + \frac{5}{16} \cdot 0 = \frac{99}{256}.$$

右が正しい場合、Ascii が「左」と 2 回答えるのは 2 回とも嘘をつくときであり、その確率は $\frac{1}{16}$ である。Biscii は必ず「左」と 2 回答える。よって

$$\mathrm{P}(E_2 \mid R) = \frac{11}{16} \cdot \frac{1}{16} + \frac{5}{16} \cdot 1 = \frac{11+80}{256} = \frac{91}{256}.$$

L と R の事前確率（prior）は等しいので、ベイズの定理において約分される。

$$\mathrm{P}(L \mid E_2) = \frac{\frac{1}{2} \cdot \frac{99}{256}}{\frac{1}{2} \cdot \frac{99}{256} + \frac{1}{2} \cdot \frac{91}{256}} = \frac{99}{99+91} = \frac{99}{190}.$$

これは $\frac{1}{2}$ より大きいので、左のほうが可能性の高い方向であり、船長は左へ進むべきである。そのとき正しい確率は $\frac{99}{190}$ である。

「左」という答えが 3 回の場合

「左」と 3 回答える事象を E_3 とする。同じ推論が、今度は Ascii からの 3 つの独立な答えについて成り立つ。

$$\mathrm{P}(E_3 \mid L) = \frac{11}{16}\left(\frac{3}{4}\right)^3 + \frac{5}{16} \cdot 0 = \frac{11}{16} \cdot \frac{27}{64} = \frac{297}{1024},$$
$$\mathrm{P}(E_3 \mid R) = \frac{11}{16}\left(\frac{1}{4}\right)^3 + \frac{5}{16} \cdot 1 = \frac{11}{1024} + \frac{320}{1024} = \frac{331}{1024}.$$

$\mathrm{P}(E_3 \mid R) > \mathrm{P}(E_3 \mid L)$ であり、事前確率は等しいので、今度は右のほうが可能性の高い方向となる。正確には

$$\mathrm{P}(R \mid E_3) = \frac{331}{297+331} = \frac{331}{628} > \frac{1}{2}.$$

したがって船長は**右**へ進むべきであり、そのとき正しい確率は $\frac{331}{628}$ である。

考えてみよう 3回目の「左」によって、なぜ左の可能性が下がるのか。

「左」が1回増えるごとに、$\mathrm{P}(E \mid L)$ における Ascii の寄与は $\frac{3}{4}$ 倍になる。一方、$\mathrm{P}(E \mid R)$ では Ascii の項は $\frac{1}{4}$ 倍に縮むが、Biscii の項は $\frac{5}{16}$ のまま変わらない。同じ答えが長く続くことはまさに Biscii が生み出すものなので、証拠はますます「相手は嘘つきだ」と告げるようになる。正直な Ascii も答えは一貫しているが、それは確率 $(\frac{3}{4})^n$ でしか起こらず、この値は急速に小さくなっていく。

検算

- 「左」が1回だけの場合、尤度（likelihood）は $\frac{11}{16} \cdot \frac{3}{4} = \frac{33}{64}$（$L$ のもと）と $\frac{11}{16} \cdot \frac{1}{4} + \frac{5}{16} = \frac{31}{64}$（$R$ のもと）であるから、$\mathrm{P}(L \mid \text{左}) = \frac{33}{64}$ である。数列 $\frac{33}{64} \approx 0.516$、$\frac{99}{190} \approx 0.521$、$\frac{297}{628} \approx 0.473$ はいったん増加してから減少しており、上の説明と合致する。
- 答えの回数が多くなると、Biscii の項が支配的になるので、$\mathrm{P}(L \mid E_n) \to 0$ となる。
- 4つの確率はすべて、分数を用いた厳密な数え上げと、10^6 回の会話のシミュレーションによって確認した。シミュレーションの値は 0.5214（厳密値 $\frac{99}{190} \approx 0.5211$）および 0.5266（厳密値 $\frac{331}{628} \approx 0.5271$）であった。

よくある間違い

- $\mathrm{P}(L) = \mathrm{P}(R) = \frac{1}{2}$ という仮定を明記し忘れること。方向について何らかの事前確率を置かなければ、この問題には答えられない。
- 最初の部分で得た $\frac{9}{10}$ を住民全体に当てはめること。これは Ascii だけについての主張であり、Biscii もまた毎回必ず同じ答えを2回返す。
- 住民が**誰であっても**、3つの答えが独立であるとみなすこと。答えが独立になるのは部族（および真の方向）がわかったときに限られる。だからこそ、まず条件付けを行うのである。

STEP のその先へ

この計算はベイズ更新（Bayesian updating）の一例である。比 $\mathrm{P}(E \mid R)/\mathrm{P}(E \mid L)$ は尤度比（likelihood ratio）またはベイズ因子（Bayes factor）と呼ばれ、証拠が一方の仮説をどれほど強く支持するかを測る量である。部族はすべての答えに共通する隠れ変数（hidden variable）であるため、答えは**条件付きで**独立（conditionally independent）であるにすぎず、一方を指しているように見える証拠が、結局は逆を指すことになり

うる。Blitzstein and Hwang, *Introduction to Probability*（確率論入門）は、条件付き確率と、まさにこの種のパラドックスについて優れた解説をしている。

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。