解答例

STEP II 1999、第 12 問

確率・統計 ・ ベイズの定理

> **STEP II 1999、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/99-s2-q12

鍵となる考え方

これはパラメータを残したままのベイズの定理（Bayes' theorem）である。全確率の公式（law of total probability）によって $\mathrm{P}(S)$ を求めてしまえば、2 つの条件付き確率（conditional probability）はどちらも p の簡単な有理関数（rational function）になり、許される区間上でそれぞれ単調（monotonic）である。したがって、各検査員の手順における「最大値」は区間 $0 \le p \le \frac{1}{3}$ の端にあり、興味深いのは、それがただ 1 回の観測から推定することについて何を語っているかという点である。

解答

条件付き確率

無作為に選んだチップがそれぞれ A、B、C によって製造されたという事象を A、B、C とし、そのチップが良品であるという事象を S とする。これら 3 つの事象は標本空間（sample space）を分割（partition）し、

$$\mathrm{P}(A) = 2p, \quad \mathrm{P}(B) = p, \quad \mathrm{P}(C) = 1 - 3p,$$
$$\mathrm{P}(S \mid A) = \tfrac{7}{10}, \quad \mathrm{P}(S \mid B) = \tfrac{8}{10}, \quad \mathrm{P}(S \mid C) = \tfrac{9}{10}.$$

である。全確率の公式により、

$$\begin{aligned}\mathrm{P}(S) &= 2p \cdot \tfrac{7}{10} + p \cdot \tfrac{8}{10} + (1 - 3p) \cdot \tfrac{9}{10} \\ &= \tfrac{1}{10}(14p + 8p + 9 - 27p) = \tfrac{1}{10}(9 - 5p).\end{aligned}$$

$0 \le p \le \frac{1}{3}$ の範囲全体で $9 - 5p > 0$ であるから、これで割ってよいことに注意する。ベイズの定理により、

$$\mathrm{P}(A \mid S) = \frac{\mathrm{P}(S \mid A)\,\mathrm{P}(A)}{\mathrm{P}(S)} = \frac{\frac{14p}{10}}{\frac{9-5p}{10}} = \frac{14p}{9-5p},$$

であり、同様に

$$\mathrm{P}(C \mid S) = \frac{\frac{9}{10}(1-3p)}{\frac{1}{10}(9-5p)} = \frac{9(1-3p)}{9-5p} = \frac{9-27p}{9-5p}.$$

である。

1 人目の検査員の推定値

p が 0 から $\frac{1}{3}$ まで増加するとき、分子 $14p$ は増加し、正の分母 $9-5p$ は減少するので、$\mathrm{P}(A \mid S)$ は狭義単調増加である。（同じことだが、その導関数は $126/(9-5p)^2 > 0$ である。）したがって最大値は右端 $p = \frac{1}{3}$ でとり、その値は

$$\frac{14/3}{9-5/3} = \frac{14/3}{22/3} = \frac{7}{11}.$$

である。よって 1 人目の検査員の推定値は $\hat{p} = \frac{1}{3}$ である。

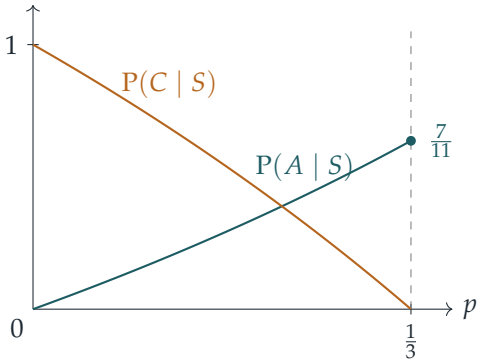
2 人目の検査員の推定値

> **考えてみよう** $\mathrm{P}(C \mid S)$ は増加するか、それとも減少するか。p が大きくなると、分子も分母もともに減少する。

分子と分母が同じ向きに動くときは答えをそのまま読み取ることはできないので、商の微分法（quotient rule）を用いて微分する：

$$\frac{\mathrm{d}}{\mathrm{d}p}\frac{9-27p}{9-5p} = \frac{-27(9-5p)+5(9-27p)}{(9-5p)^2} = \frac{-243+135p+45-135p}{(9-5p)^2} = \frac{-198}{(9-5p)^2} < 0.$$

よって $\mathrm{P}(C \mid S)$ は $[0, \frac{1}{3}]$ 上で狭義単調減少であり、最大値は $p = 0$ でとり、その値は 1 である。2 人目の検査員の推定値は $\hat{p} = 0$ である。



コメント

2 人の検査員は、同等に妥当な 1 回ずつの観測に同じ方法を適用しているのに、許される範囲の正反対の両端にたどり着く。すなわち $\hat{p} = \frac{1}{3}$（C 製のチップは 1 つもない）と $\hat{p} = 0$（すべてのチップが C 製）である。どちらの推定値も単に「自分のチップの製造元ができるだけ多くなるような p の値」であり、標本が 1 つしかなければ必然的にそうなる。2 人目の推定値は 1 人目の検査員のチップによって否定されさえする。$p = 0$ は A がチップを 1 つも製造しないことを意味するからである。どちらの推定値も信用できない。1 回の観測では p を推定するには情報が少なすぎ、より大きな標本（各製造元ごとの個数を伴うもの）が必要である。

また、チップが良品であることはほとんど何の役割も果たしていないことにも注目する価値がある。「良品であり、かつ A 製である」ことの尤度（likelihood）は $\frac{7}{10} \cdot 2p$ であり、これは $2p$ 自身と同じ場所で最大になる。係数 $\frac{7}{10}$ は p に依存しないからである。

検算

- $p = 0$ のときはすべてのチップが C 製なので、$\mathrm{P}(C \mid S)$ は 1、$\mathrm{P}(A \mid S)$ は 0 となるはずである。どちらの式もこれと一致する。
- 同様に $\mathrm{P}(B \mid S) = 8p/(9 - 5p)$ であり、3 つの条件付き確率の和は

$$\frac{14p + 8p + 9 - 27p}{9 - 5p} = 1.$$

  となる。
- $p = 0.2$ のとき、式からは $\mathrm{P}(A \mid S) = 0.35$、$\mathrm{P}(C \mid S) = 0.45$ が得られる。400 000 個のチップのシミュレーションでは 0.3505 と 0.4505 が得られた。

よくある間違い

- 導関数を 0 とおいて極値を探すこと。極値は存在しない。どちらの最大値も $[0, \frac{1}{3}]$ の端点でとる。
- $\mathrm{P}(C \mid S)$ には慎重な議論が必要であることを忘れること。分子も分母もともに減少するので、「分子が小さくなる」だけでは不十分である。

STEP のその先へ

2 人の検査員はどちらも、実質的には最尤推定（maximum likelihood estimation）を行っている。すなわち、自分が観測したことが最も起こりやすくなるようなパラメータの値を選んでいるのである。観測が 1 回だけだと、最尤推定値はしばしばパラメータ空間（parameter space）の境界に落ちる。これは、最尤推定がうまく機能するには十分なデータが必要であることへの警告である。ベイズ主義者（Bayesian）であれば、代わりに p に事前分布（prior distribution）を置き、チップを 1 つ観測するごとにそれを更新するだろう。そうすれば、1 回の観測からこのような極端な結論が導かれることは避けられる。Blitzstein と Hwang の *Introduction to Probability*（確率論入門）は、この両方の立場を論じている。

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。