解答例

# STEP I 1987、第 16 問

確率・統計 ・ ベイズの定理と正規近似

> **STEP I 1987、問 16**
>
> 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/87-s1-q16

## 鍵となる考え方

小問 (i) は、形を変えたベイズの定理（Bayes' theorem）である。2 回続けて同じ投票をしたという事実は、その議員がどちらの党に属するかについての証拠（evidence）になる。保存党（Preservative）の議員は必ずそうするが、進歩党（Progressive）の議員はときどきしかそうしないからである。そこで、まずその議員が保存党員である確率を更新し、それから 3 回目の投票を予測する。小問 (ii) では保存党員の投票は定まった数なので、ランダム性はすべて 40 人の進歩党員から生じる。彼らの賛成票の数は二項分布（binomial distribution）に従い、これを正規分布（normal distribution）で近似するのである。

## 解答

### 小問 (i)

選ばれた議員が保存党員であるという事象を $R$ とすると、$P(R) = \frac{60}{100} = \frac{3}{5}$、$P(R') = \frac{2}{5}$ である。その議員が最初の 2 回で同じ投票をしたという事象を $S_2$、3 回目の投票もそれと一致するという事象を $S_3$ とする。

保存党員は決して考えを変えないので、$P(S_2 \mid R) = 1$ である。進歩党員はランダムに投票し、各投票は独立にそれぞれ確率 $\frac{1}{2}$ で賛成または反対となる。1 回目の投票がどちらであっても、2 回目がそれと一致する確率は $\frac{1}{2}$ なので、$P(S_2 \mid R') = \frac{1}{2}$ である。したがって

$$P(S_2) = \tfrac{3}{5} \cdot 1 + \tfrac{2}{5} \cdot \tfrac{1}{2} = \tfrac{4}{5}, \qquad P(R \mid S_2) = \frac{P(R)\,P(S_2 \mid R)}{P(S_2)} = \frac{3/5}{4/5} = \frac{3}{4}.$$

よって、観察した事実のもとでは、その議員が保存党員である確率は $\frac{3}{4}$、進歩党員である確率は $\frac{1}{4}$ である。

**考えてみよう** 議員はランダムに選ばれたのだから、単に $P(R \mid S_2) = \frac{3}{5}$ としてはいけないのはなぜか？

情報 $S_2$ は進歩党員よりも保存党員のほうに起こりやすいので、オッズ（odds）を変化させるからである。観察前のオッズは保存党員側に $3:2$ であった。$S_2$ を観察した後は $3\cdot 1 : 2\cdot \frac{1}{2} = 3:1$ となる。

さて、議員の所属政党が与えられれば、3 回目の投票は、最初の 2 回の投票が政党について教えてくれる情報以上には、それらに依存しない。保存党員は確実に同じ投票を繰り返し、進歩党員は確率 $\frac{1}{2}$ で一致させる。よって

$$P(S_3 \mid S_2) = \tfrac{3}{4}\cdot 1 + \tfrac{1}{4}\cdot\tfrac{1}{2} = \frac{7}{8}.$$

（同値な計算として、$P(S_3 \mid S_2) = \dfrac{P(S_2 \cap S_3)}{P(S_2)} = \dfrac{\frac{3}{5} + \frac{2}{5}\cdot\frac{1}{4}}{\frac{4}{5}} = \dfrac{7/10}{4/5} = \dfrac{7}{8}$ である。）

**小問 (ii)**

保存党の政策案に関する採決を考え、100 人の議員全員が投票するものとする。60 人の保存党員のうち割合 $\alpha$、すなわち $60\alpha$ 人は常に反対票を投じ、残りの $60(1-\alpha)$ 人は常に賛成票を投じる。賛成票を投じる進歩党員の人数を $X$ とする。各進歩党員は独立に確率 $\frac{1}{2}$ で賛成するので

$$X \sim \mathrm{B}\left(40, \tfrac{1}{2}\right), \qquad E(X) = 20, \qquad \mathrm{Var}(X) = 40\cdot\tfrac{1}{2}\cdot\tfrac{1}{2} = 10.$$

100 票の単純過半数（simple majority）とは、賛成が 51 票以上であることを意味する：

$$60(1-\alpha) + X \geq 51 \iff X \geq 60\alpha - 9.$$

$n = 40$ はかなり大きく、$p = \frac{1}{2}$ により二項分布は対称なので、$X$ を $Y \sim \mathrm{N}(20, 10)$ で近似する。ここで $60\alpha - 9$ は整数（議員の人数）であるから、連続性補正（continuity correction）を用いて

$$P(\text{過半数}) = P(X \geq 60\alpha - 9) \approx P(Y > 60\alpha - 9.5) = P\left(Z > \frac{60\alpha - 29.5}{\sqrt{10}}\right),$$

となる。ただし $Z \sim \mathrm{N}(0,1)$ である。

この確率は $\alpha$ が増加すると減少するので、選挙が行われるのは、$\alpha$ がこの確率を 0.8 未満に押し下げるほど大きいとき、かつそのときに限る。数表より $P(Z > -0.8416) = 0.8$

なので、

$$\frac{60\alpha - 29.5}{\sqrt{10}} > -0.8416 \iff 60\alpha > 29.5 - 0.8416\sqrt{10} = 26.84$$
$$\iff \alpha > 0.447.$$

が必要である。$60\alpha$ は議員の人数として整数でなければならないので、とりうる最小の値は $60\alpha = 27$ である。よって、選挙が行われることになる $\alpha$ の最小値は

$$\alpha = \frac{27}{60} = 0.45.$$

これが正しいことを確かめよう。造反者が 26 人のとき過半数となる近似確率は $P(Z > -1.107) \approx 0.866$ であり、27 人のときは $P(Z > -0.791) \approx 0.785 < 0.8$ である。

## 検算

- (i) において、$\frac{7}{8}$ は当然あるべきとおり、$\frac{1}{2}$（進歩党員とわかっている場合）と 1（保存党員とわかっている場合）の間にある。ランダムに選んだ議員 400,000 人のシミュレーションでは 0.875 が得られた。
- (ii) において、二項分布による厳密な確率は $P(X \geq 17) = 0.8659$（造反者 26 人）、$P(X \geq 18) = 0.7852$（造反者 27 人）である。これらは正規近似の値と小数第 3 位まで一致し、境目が造反者 26 人と 27 人の間、すなわち $\alpha = 0.45$ にあることを裏づけている。

## よくある間違い

- (i) で $\frac{3}{5} \cdot 1 + \frac{2}{5} \cdot \frac{1}{2} = \frac{4}{5}$ と答えてしまう。これは、議員がすでに 2 回同じ投票をしているという情報を無視している。
- 単純過半数を「50 票以上」として扱ってしまう。50 対 50 の同数は過半数ではないので、51 票が必要である。
- 連続性補正を省いてしまう。補正をしないと境界は $\alpha > 0.456$ となり、造反者は 27 人ではなく 28 人となるが、これは二項分布による厳密な計算によれば誤りである。

## STEP のその先へ

小問 (i) は**ベイズ更新**（Bayesian updating）の小さな例である。データが隠れた「タイプ」についての信念を変化させ、その後の予測は更新された信念にわたる平均（事後予測分布（posterior predictive distribution））として与えられる。小問 (ii) の正規近似

は中心極限定理（central limit theorem）の一例であり、ここでほぼ完全に一致するのは偶然ではない。$p = \frac{1}{2}$ のとき二項分布は対称だからである。これら二つの考え方は、Blitzstein and Hwang, *Introduction to Probability*（確率論入門）で丁寧に展開されている。

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。