解答例

STEP I 2006、第 12 問

確率・統計 ・ 条件付き確率

> **STEP I 2006、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/06-s1-q12

鍵となる考え方

道路が通行可能であるのは、その道路上の**すべての**区間が通行可能なときに限り、二つの町が孤立するのは、**すべての**道路が不通になったときに限る。そこで、独立性（independence）を用い、常に余事象（complement）を通して、道路ごとに考えていく。後半は、どの道路を選んだかが効いてくる条件付き確率（conditional probability）の問題である。1 区間の道路は 2 区間の道路よりも通行可能である確率が高いので、自分の道路が通行可能だとわかれば、どの道路を選んだかについての情報が得られる。これを定義 $\mathrm{P}(B \mid A) = \mathrm{P}(A \cap B)/\mathrm{P}(A)$ を用い、選んだ道路で場合分けして処理する。

解答

孤立する確率

2 区間の道路を R_1, R_2、1 区間の道路を R_3 とする。また、1 つの区間が通行可能である確率を $q = 1 - p$ とおく。

2 区間の道路が通行可能であるのは、その 2 つの区間がともに通行可能であるときに限り、独立性よりその確率は $q^2 = (1-p)^2$ である。よって

$$\mathrm{P}(R_1 \text{ が不通}) = \mathrm{P}(R_2 \text{ が不通}) = 1 - (1-p)^2 = p(2-p),$$
$$\mathrm{P}(R_3 \text{ が不通}) = p.$$

3 本の道路は互いに異なる 5 つの区間からなるので、事象「R_1 が不通」「R_2 が不通」「R_3 が不通」は、独立な区間の互いに素な集合にそれぞれ依存しており、したがって独立である。オックスタウン（Oxtown）がキャムヴィル（Camville）から孤立するのは 3

本の道路がすべて不通のときに限るので、

$$\mathrm{P}(\text{孤立}) = p(2-p)\cdot p(2-p)\cdot p = p^3(2-p)^2,$$

となり、示された。

条件付き確率

道路 R_k を選ぶ事象を C_k とすると、各 k について $\mathrm{P}(C_k) = \frac{1}{3}$ であり、道路の選択は冠水とは独立である。自分の道路が通行可能である事象を O、他の 2 本の道路がともに不通である事象を B とする。求めるのは $\mathrm{P}(B \mid O) = \mathrm{P}(B \cap O)/\mathrm{P}(O)$ である。

> **考えてみよう** 独立性から、答えは単に P(他の 2 本の道路が不通) ではないのか？

どの道路を選んだかがわかっていれば、そのとおりである。しかし「他の 2 本の道路」は選択に応じて異なる組になり、さらに自分の道路が通行可能だという観察によって、短い道路 R_3 を選んだ可能性が以前より高くなる。したがって、選んだ道路で場合分けし、正しく重み付けしなければならない。まさにそれを行うのが以下の式である。

分母。道路の選択について全確率の法則（law of total probability）を用いると、

$$\begin{aligned}\mathrm{P}(O) &= \tfrac{1}{3}(1-p)^2 + \tfrac{1}{3}(1-p)^2 + \tfrac{1}{3}(1-p)\\ &= \tfrac{1}{3}(1-p)[2(1-p)+1] = \tfrac{1}{3}(1-p)(3-2p).\end{aligned}$$

分子。R_1 を選んだ場合、R_1 が通行可能で R_2, R_3 が不通である必要があり、道路の独立性よりその確率は $(1-p)^2\cdot p(2-p)\cdot p$ である。R_2 を選んだ場合も同じである。R_3 を選んだ場合、R_3 が通行可能で 2 区間の道路が両方とも不通である必要があり、その確率は $(1-p)\cdot p^2(2-p)^2$ である。よって

$$\begin{aligned}\mathrm{P}(B\cap O) &= \tfrac{1}{3}\left[2p^2(2-p)(1-p)^2 + p^2(2-p)^2(1-p)\right]\\ &= \tfrac{1}{3}\,p^2(1-p)(2-p)[2(1-p)+(2-p)]\\ &= \tfrac{1}{3}\,p^2(1-p)(2-p)(4-3p).\end{aligned}$$

答え。割り算をして、$p<1$ のとき 0 でない因数 $\frac{1}{3}(1-p)$ を約分すると、

$$\mathrm{P}(B \mid O) = \frac{p^2(2-p)(4-3p)}{3-2p}.$$

因数 $p, 2-p, 4-3p, 3-2p$ はいずれも 1 次式なので、これで完全に因数分解されている。

極限 $p \to 1$

$p \to 1$ のとき各因数は 1 に近づくので、

$$\mathrm{P}(B \mid O) \to \frac{1 \cdot 1 \cdot 1}{1} = 1.$$

これは理にかなっている。冠水がほぼ確実なとき、自分の道路が通行可能であったのは幸運な偶然であり、自分の道路とは独立に冠水する他の道路は、ほぼ確実にその幸運に恵まれていない。自分の道路が通行可能だという知識は、他の道路が通行可能かどうかについて何も教えてくれないのである。したがって、他のどの道路を選んでいても通れなかったことはほぼ確実だと考えるべきである。（$p = 1$ そのものでは条件となる事象の確率が 0 となり、条件付き確率は定義されない。意味をもつのは極限（limit）だけである。）

検算

- $p = 0$ のときは何も不通にならず、式は 0 を与える。当然そうなるべきである。
- $p = \frac{1}{2}$ のとき、式は $\dfrac{\frac{1}{4} \cdot \frac{3}{2} \cdot \frac{5}{2}}{2} = \frac{15}{32}$ を与える。不通区間のパターン 2^5 通りすべてと 3 通りの道路の選択すべてを厳密に数え上げても同じ値が得られ、$p = \frac{2}{3}$（このとき $\frac{32}{45}$）など他のいくつかの p の値についても同様である。
- $0 \leq p < 1$ に対して答えは常に 0 と 1 の間にある。例えば $p = 0.99$ のとき、約 0.9996 である。

よくある間違い

- 2 区間の道路が不通になる確率を、$1 - (1 - p)^2$（少なくとも 1 区間が不通）ではなく p^2（両区間とも不通）としてしまう。
- 条件付けを忘れる。他の 2 本の道路が不通である確率は、無作為に選んだ自分の道路が通行可能であるという条件の下での条件付き確率とは同じではない。
- 観察の**後で** 3 通りの選択を等しく重み付けしてしまう。等しいのは事前確率（prior probability）$\frac{1}{3}$ であって、事後確率（posterior probability）ではない。

STEP のその先へ

直列につながった区間からなる道路を並列に並べたものは、**信頼性ネットワーク**（reliability network）の最も簡単な例である。ネットワークが 2 点を連結する確率は、

一般にはパーコレーション理論（percolation theory）で研究されており、そこでは p の変化に伴う振る舞いの急激な変化が中心的なテーマとなっている。後半は姿を変えたベイズの定理（Bayes' theorem）である。「自分の道路が通行可能である」という観察によって、どの道路が選ばれたかの確率が更新されるのである。Grimmett と Welsh の *Probability: An Introduction*（確率論入門）は、まさにこの精神で条件付き確率と独立性を扱っている。

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。