详细解答

STEP I 2006，第 12 题

概率与统计 · 条件概率

> **STEP I 2006，第 12 题**
>
> 继续阅读之前，请先看题目（原题为英文）。题目载于 STEP Support Programme 的试题资料库：
>
> step.maths.org/questions/06-s1-q12

关键思路

一条道路要畅通，必须路上每一段都畅通；而两镇要被隔绝，则必须每一条道路都受阻。因此我们逐条道路处理，利用独立性（independence），并且每一步都借助补事件（complement）来计算。第二部分是一个条件概率（conditional probability）问题，当中选了哪条路是有影响的：只有一段的道路比有两段的道路更可能畅通，所以发现自己所选的路畅通，其实透露了你选的是哪条路的一些信息。我们用定义 $\mathrm{P}(B \mid A) = \mathrm{P}(A \cap B)/\mathrm{P}(A)$ 来处理，并按所选的道路分情况计算。

解答

被隔绝的概率

把两条有两段的道路称为 R_1 和 R_2，只有一段的道路称为 R_3。设 $q = 1 - p$ 为单独一段畅通的概率。

一条有两段的道路畅通，当且仅当它的两段都畅通；由独立性，其概率为 $q^2 = (1-p)^2$。所以

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

三条道路共用到五个不同的路段，所以“R_1 受阻”、“R_2 受阻”和“R_3 受阻”这三个事件各自取决于互不相交的几组独立路段，因而互相独立。Oxtown 与 Camville 被隔绝，当且仅当三条道路全部受阻，所以

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

正如所求。

条件概率

设 C_k 为我选择道路 R_k 的事件，则对每个 k 都有 $\mathrm{P}(C_k) = \frac{1}{3}$，而且我的选择与水浸情况独立。设 O 为我所选的路畅通的事件，B 为另外两条路都受阻的事件。我们要求的是 $\mathrm{P}(B \mid O) = \mathrm{P}(B \cap O)/\mathrm{P}(O)$。

> **想一想** 根据独立性，答案不就是 P(另外两条路受阻) 吗?

假如我们知道自己选了哪条路，答案确是如此。但“另外两条路”是哪两条，要视乎我们的选择而定；而且观察到自己的路畅通之后，选中较短的 R_3 的可能性就比原先高了。所以我们必须按所选的道路分情况，并配上正确的权重，而下面的公式做的正是这件事。

分母。对所选的道路运用全概率公式（law of total probability），

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

分子。如果我选了 R_1，就需要 R_1 畅通而 R_2、R_3 受阻；由各条道路的独立性，其概率为 $(1-p)^2 \cdot p(2-p) \cdot p$。选 R_2 的结果相同。如果我选了 R_3，就需要 R_3 畅通而两条有两段的道路都受阻：$(1-p) \cdot p^2(2-p)^2$。因此

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

答案。两式相除，并约去因子 $\frac{1}{3}(1-p)$（当 $p < 1$ 时它不等于零），得

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

因子 p、$2-p$、$4-3p$ 和 $3-2p$ 全都是一次式，所以这已经是完全因式分解的形式。

极限 $p \to 1$

当 $p \to 1$ 时，每个因子都趋向 1，所以

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

这是合理的。当水浸几乎必然发生时，我自己的路畅通纯属侥幸；而其他道路是否水浸与我的路独立，它们几乎肯定没有同样的运气：知道我的路畅通，并不能说明它们是否畅通。所以我应该几乎肯定，换了任何其他选择都会行不通。（在 $p = 1$ 本身，作为条件的事件概率为零，条件概率没有定义；只有极限才有意义。）

验算

- 当 $p = 0$ 时，没有任何路段会受阻，而公式给出 0，正该如此。
- 当 $p = \frac{1}{2}$ 时，公式给出 $\dfrac{\frac{1}{4} \cdot \frac{3}{2} \cdot \frac{5}{2}}{2} = \frac{15}{32}$。把路段受阻的全部 2^5 种情况和三种道路选择逐一精确列举，得到的数值相同；对其他几个 p 值也是如此，例如 $p = \frac{2}{3}$ 时得 $\frac{32}{45}$。
- 对 $0 \leq p < 1$，答案总是介乎 0 与 1 之间：例如当 $p = 0.99$ 时，它约为 0.9996。

常见错误

- 把有两段的道路受阻的概率当作 p^2（两段都受阻），而不是 $1 - (1 - p)^2$（至少一段受阻）。
- 忘记取条件：另外两条路受阻的概率，与已知我随机选中的路畅通时它的条件概率，并不相同。
- 在观察之后仍把三个选择看成权重相等；相等的是先验概率（prior probability）$\frac{1}{3}$，而不是后验概率（posterior probability）。

STEP 之外

多条道路并联、每条道路又由若干路段串联而成，是可靠性网络（reliability network）最简单的例子；一个网络能把两点连通的概率，其一般情况是渗流理论（percolation theory）的研究对象，而当 p 变化时网络行为出现的急剧转变，正是该理论的核心课题。第二部分其实是乔装了的贝叶斯定理（Bayes' theorem）：“我的路畅通”这项观察，更新了选中各条道路的概率。Grimmett 与 Welsh 合著的 *Probability: An Introduction*（概率论入门）正是以这种思路讲解条件概率和独立性。

Beyond STEP 详细解答：https://beyondstep.keithmcnulty.org/zh-hans/。解答文字采用 CC0 1.0；题目本身版权属 Cambridge University Press & Assessment 所有（©），本文件不予转载。由机器从英文翻译而成；尚未经审阅。