解答例

STEP I 2003、第 14 問

確率・統計 ・ ポアソン分布

> **STEP I 2003、問 14**
>
> 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/03-s1-q14

鍵となる考え方

Jane が会う友人の人数は、電話をかけてきた友人の人数そのものではなく、その人数を 2 で打ち切ったものである。打ち切った変数で問題になるのは三つの場合だけである。誰も電話をかけてこない場合、ちょうど一人がかけてくる場合、二人以上がかけてくる場合である。したがって、ポアソン分布（Poisson distribution）の級数全体を和として計算する必要はない。$P(0)$ と $P(1)$ さえ求めれば、「二人以上」の確率は残りの確率として得られる。新しい友人については、その友人が電話をかけてくるかどうかで条件付け（conditioning）を行う。

解答

もとからの友人

ある日に電話をかけてくる友人の人数を N とする。すると $N \sim \mathrm{Po}(2)$ であり、

$$P(N=k) = \frac{\mathrm{e}^{-2}2^k}{k!}, \qquad k = 0,1,2,\ldots$$

である。Jane が会う友人の人数は $S = \min(N,2)$ である。すなわち、$N=0$ のとき $S=0$、$N=1$ のとき $S=1$、$N \geq 2$ のとき $S=2$ である。ここで

$$P(N=0) = \mathrm{e}^{-2}, \qquad P(N=1) = 2\mathrm{e}^{-2}, \qquad P(N \geq 2) = 1 - 3\mathrm{e}^{-2},$$

であるから、

$$\mathrm{E}(S) = 0 \cdot \mathrm{e}^{-2} + 1 \cdot 2\mathrm{e}^{-2} + 2\left(1 - 3\mathrm{e}^{-2}\right) = 2 - 4\mathrm{e}^{-2},$$

となり、求める結果が得られた。

同じ計算をもっとすっきり書く方法があり、以下でも再び用いる。$\{0,1,2,\dots\}$ に値をとる任意の確率変数（random variable）X に対して

$$\mathrm{E}(\min(X,2)) = 2 - 2P(X=0) - P(X=1),$$

が成り立つ。なぜなら、$\min(X,2)$ は $X=0$ のとき 2 より 2 だけ小さく、$X=1$ のとき 1 だけ小さく、それ以外のときは 2 に等しいからである。$X=N$ とすると、再び $2-2\mathrm{e}^{-2}-2\mathrm{e}^{-2}$ が得られる。

新しい友人が加わった場合

新しい友人が電話をかけてくるとき $B=1$、そうでないとき $B=0$ とする。すると $P(B=1)=p$ である。問題文により、B は N と独立（independent）である。電話をかけてくる人数の合計は $M=N+B$ となり、Jane が会う友人の人数は $\min(M,2)$ である。

> **考えてみよう** M の分布全体が必要だろうか？

必要ない。上の恒等式により、$P(M=0)$ と $P(M=1)$ だけがあればよい。独立性を用いて確率を掛け合わせると、

$$P(M=0) = P(N=0)\,P(B=0) = (1-p)\mathrm{e}^{-2},$$
$$P(M=1) = P(N=1)\,P(B=0) + P(N=0)\,P(B=1) = 2(1-p)\mathrm{e}^{-2} + p\,\mathrm{e}^{-2}.$$

よって

$$\begin{aligned}\mathrm{E}(\min(M,2)) &= 2 - 2(1-p)\mathrm{e}^{-2} - 2(1-p)\mathrm{e}^{-2} - p\,\mathrm{e}^{-2} \\ &= 2 - (4-3p)\mathrm{e}^{-2}.\end{aligned}$$

別解として、B で条件付けてもよい。新しい友人が電話をかけてこない場合（確率 $1-p$）、平均は先ほどと同じく $2-4\mathrm{e}^{-2}$ である。電話をかけてくる場合（確率 p）、Jane が会う人数は $\min(N+1,2)$ であり、これは $N=0$ のとき 1、それ以外のとき 2 であるから、その平均は $\mathrm{e}^{-2}+2(1-\mathrm{e}^{-2}) = 2-\mathrm{e}^{-2}$ である。したがって平均は

$$(1-p)(2-4\mathrm{e}^{-2}) + p(2-\mathrm{e}^{-2}) = 2-(4-3p)\mathrm{e}^{-2}.$$

となる。以上より、Jane が一日に会う友人の人数は平均して $2-(4-3p)\mathrm{e}^{-2}$ 人となり、$3p\,\mathrm{e}^{-2}$ だけ増加する。

検算

- $p=0$ のとき、新しい友人は決して電話をかけてこないので、$2-4\mathrm{e}^{-2} \approx 1.4587$ に戻る。

- $p = 1$ のとき、$2 - e^{-2} \approx 1.8647$ となり、これは当然そうあるべきとおり $\min(N+1, 2)$ の平均に一致する。
- 答えは p について増加し、常に 2 より小さい。これは理にかなっている。
- 正確な分布に対して $\min(k+b, 2)$ の和をとった場合も、$p = 0.3$ として 4×10^5 日分のシミュレーションを行った場合も、どちらも約 1.580 となり、$2 - 3.1e^{-2}$ と一致する。

よくある間違い

- 二人という上限を忘れて、$E(N) = 2$ や $2 + p$ と答えてしまう。
- $N + B$ をポアソン分布に従うものとして扱ってしまう。ベルヌーイ（Bernoulli）変数はポアソン分布に従わないので、その和もポアソン分布には従わない。
- ちょうど一人が電話をかけてくる二通りの場合（もとからの友人一人がかけてきて新しい友人はかけてこない場合と、もとからの友人は誰もかけてこず新しい友人がかけてくる場合）のうち、一方を見落としてしまう。

STEP のその先へ

恒等式 $E(\min(X, 2)) = P(X \geq 1) + P(X \geq 2)$ は、非負整数値をとる確率変数に対する裾和公式（tail-sum formula）$E(X) = \sum_{k \geq 1} P(X \geq k)$ の特別な場合である。これは生存関数（survival function）を積分することの離散版である。このような打ち切られた期待値（expectation）は、待ち行列（queueing）モデルや在庫（inventory）モデルに現れ、処理能力に限りがあるときにサービスを受ける客の人数の期待値を与える。Grimmett and Welsh, *Probability: An Introduction*（確率論入門）では、ここで用いた二つの道具である裾和公式と条件付き期待値（conditional expectation）が扱われている。

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。