解答例

STEP II 1992、第 16 問

確率・統計 ・ 条件付き確率

> **STEP II 1992、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/92-s2-q16

この問題は、*Beyond STEP* の第 12.17 節「**バナッハのマッチ箱**」で詳しく解説されている：https://beyondstep.keithmcnulty.org/ja/ch12/matchboxes.html

鍵となる考え方

これはバナッハ（Banach）のマッチ箱問題（matchbox problem）の一つの変形である。不足に気づいたのが**いつか**（N 回目に手を伸ばしたとき）が与えられているので、これは条件付き確率（conditional probability）の問題である。そこで二つの同時確率（joint probability）を計算する。一つは最初の失敗が N 回目に起こり、それがトフィーである確率、もう一つは最初の失敗が N 回目に起こり、それがミントである確率である。どちらも「最初の $N-1$ 回の中に各種類がちょうどこれだけ含まれ、その次に特定の種類が来る」という計算であり、二項分布（binomial distribution）によるものである。

解答

手を伸ばす各回は独立（independent）であり、各回でトフィーに手を伸ばす確率は p、ミントに手を伸ばす確率は q であるとする。ただし $0 < p < 1$ である。T を、N 回目に彼女がトフィーに手を伸ばしてトフィーの袋が空であると気づき、かつそれより前の回ではどの袋も空だと気づいていない、という事象とする。M を、トフィーをミントに置き換えた同じ事象とする。求めるものは

$$\mathrm{P}(\text{トフィー} \mid \text{初めて気づいたのが第} N) = \frac{\mathrm{P}(T)}{\mathrm{P}(T) + \mathrm{P}(M)}.$$

である。

T の確率

T が起こるためには、$N-1$ 回手を伸ばした後にトフィーの袋が空になっていなければならない。それまでに失敗した回はないので、それらの各回ではすべて菓子を取り出している。したがって最初の $N-1$ 回のうちちょうど n 回がトフィーに手を伸ばした回であり（もう一回あれば失敗になっていたはずである）、残りの $N-1-n$ 回はミントに手を伸ばした回である。そして N 回目はトフィーに手を伸ばす。

逆に、最初の $N-1$ 回の中に、順序は任意で、トフィーに手を伸ばした回がちょうど n 回、ミントに手を伸ばした回が $N-1-n \le m$ 回含まれているとする。このとき、それより前の回で失敗することはありえない。どの段階でも、トフィーに n 回より多く、あるいはミントに m 回より多く手を伸ばしてはいないからである。そのうえで N 回目がトフィーであれば、T が起こる。

> **考えてみよう** 早い段階で使い切ってしまわないように、最初の $N-1$ 回の順序が問題になるのではないか？

問題にならない。彼女が袋が空であることに気づくのは、空になった袋に手を伸ばしたときだけである。トフィーの袋を早い段階で（n 個目のトフィーで）空にしてしまっても、N 回目より前に再びトフィーに手を伸ばさなければ差し支えない。そしてこれは自動的に成り立つ。最初の $N-1$ 回の中にトフィーに手を伸ばした回は n 回しかないからである。よって $\binom{N-1}{n}$ 通りの順序はすべて許され、

$$\mathrm{P}(T) = \binom{N-1}{n} p^n q^{N-1-n} \cdot p = \binom{N-1}{n} p^{n+1} q^{N-1-n}, \qquad n \le N-1 \le n+m,$$

であり、それ以外の場合は $\mathrm{P}(T)=0$ である。$N-1<n$ ならば、トフィーを使い切ることはできない。$N-1>n+m$ ならば、m 個より多くのミントが必要になってしまう。

M の確率

役割を入れ替えれば、まったく同様にして

$$\mathrm{P}(M) = \binom{N-1}{m} q^{m+1} p^{N-1-m}, \qquad m \le N-1 \le n+m,$$

であり、それ以外の場合は $\mathrm{P}(M)=0$ である。

答え

気づくのは $\min(n,m)+1$ 回目から $n+m+1$ 回目のどこかでなければならない。$n+m+2$ 回目はありえない。その時点では両方の袋が空になっており、すでにどこかで失敗が起きているはずだからである。

- $N \le m$ のとき（このとき $\mathrm{P}(M) = 0$ だが $\mathrm{P}(T) > 0$）、彼女は確実にトフィーに手を伸ばしていた。確率は 1 である。
- $N \le n$ のとき（このとき $\mathrm{P}(T) = 0$）、確率は 0 である。
- $\max(n, m) + 1 \le N \le n + m + 1$ のとき、どちらも起こりうるので、

$$\mathrm{P}(\text{トフィー}) = \frac{\binom{N-1}{n} p^{n+1} q^{N-1-n}}{\binom{N-1}{n} p^{n+1} q^{N-1-n} + \binom{N-1}{m} q^{m+1} p^{N-1-m}}.$$

最後の場合は、分子と分母を $\mathrm{P}(T)$ で割ることで式を整理できる。二項係数から共通因数 $(N-1)!$ が約分され、累乗はまとめて $(q/p)^{n+m+2-N}$ となる。これより

$$\mathrm{P}(\text{トフィー}) = \left[1 + \frac{n!\,(N-1-n)!}{m!\,(N-1-m)!} \left(\frac{q}{p}\right)^{n+m+2-N}\right]^{-1}.$$

を得る。

検算

- 対称性：$n = m$ かつ $p = q = \frac{1}{2}$ ならば、二つの式は等しく、答えは当然 $\frac{1}{2}$ となる。
- $n = 2$、$m = 3$、$p = \frac{1}{2}$、$N = 5$ とする。このとき $\mathrm{P}(T) = \binom{4}{2}/2^5 = \frac{3}{16}$、$\mathrm{P}(M) = \binom{4}{3}/2^5 = \frac{1}{8}$ であるから、答えは $\frac{3/16}{5/16} = \frac{3}{5}$ である。トフィーのほうが少ないので、トフィーが尽きるほうが起こりやすい。
- $N = 1, \ldots, n + m + 1$ にわたって $\sum_N \big(\mathrm{P}(T) + \mathrm{P}(M)\big) = 1$ である。いずれはどこかの回で必ず失敗するからである。このことと公式とを、すべての $n, m \le 6$ といくつかの p の値について、一歩ずつの厳密な数え上げと照合して確かめた。$n = 4$、$m = 3$、$N = 7$、$p = 0.55$ での 2×10^5 回の試行によるシミュレーションでは 0.527 となり、厳密値 0.528 とよく一致した。

よくある間違い

- N 回目そのものに対応する最後の因数 p（または q）を忘れる。
- 条件付き確率 $\mathrm{P}(T)/(\mathrm{P}(T) + \mathrm{P}(M))$ ではなく $\mathrm{P}(T)$ を答えてしまう。気づいたのが N 回目であることは与えられているのである。
- 事象を「N 回目にトフィーが尽きる」ことだと考えてしまう。最後のトフィーを取り出すことは気づくことではない。気づくのはその次に手を伸ばしたときである。

STEP のその先へ

最初の失敗が起こる時刻は、二つの**負の二項分布**（negative binomial distribution）の混合である。負の二項分布は、与えられた回数の成功を得るまでに必要な試行の回数を数えるものである。バナッハの元のマッチ箱問題（$n = m$、$p = \frac{1}{2}$）は、気づいた瞬間にもう一方の箱にマッチが何本残っているかを問うものである。その答えはスターリングの公式（Stirling's formula）で近似でき、平均はおよそ $2\sqrt{n/\pi}$ となる。負の二項分布については Grimmett and Welsh, *Probability: An Introduction*（確率論入門）を参照されたい。フェラー（Feller）の古典的名著 *An Introduction to Probability Theory and Its Applications*, Volume 1（確率論とその応用第 1 巻）には、マッチ箱問題についての有名な議論がある。

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。