解答例

STEP I 2006、第 13 問

確率・統計 ・ ポアソン混合分布

STEP I 2006、問 13

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-q13

鍵となる考え方

ダイヤモンドが一定の平均率で独立にばらまかれているとき、これはポアソン過程（Poisson process）であり、m キログラムのチョコレートに含まれるダイヤモンドの個数は平均 m のポアソン分布に従う。チョコレートの量そのものが確率的に決まるので、その量で条件付け（condition）を行う。すなわち、量を固定したときの答えを求め、それをさいころの目について全確率の法則（law of total probability）（および期待値に対するその仲間の法則）を用いて平均するのである。こうすると各小問は級数の和の計算に帰着する。(i) では有限和、(ii) では等比級数である。

解答

私が見つけるダイヤモンドの個数を D とする。ダイヤモンドは 1 キログラムあたり平均 1 個の割合で独立に埋め込まれているので、質量 $m\,\mathrm{kg}$ の板チョコに含まれる個数は平均 m のポアソン分布（Poisson distribution）に従う。特に、質量が $m\,\mathrm{kg}$ であるとき、

$$\mathrm{P}(D=0)=\mathrm{e}^{-m}, \qquad \mathrm{E}(D)=m.$$

全体を通して、$100k$ グラムは $0.1k$ キログラムである。

小問 (i)

さいころの目を N とすると、$n=1,\dots,6$ に対して $\mathrm{P}(N=n)=\frac{1}{6}$ である。$N=n$ のとき、私は $0.1n\,\mathrm{kg}$ を買うので、$D \mid N=n$ は平均 $0.1n$ のポアソン分布に従う。全確

率の法則により、

$$\begin{aligned}\mathrm{P}(D=0) &= \sum_{n=1}^{6} \mathrm{P}(N=n)\,\mathrm{P}(D=0 \mid N=n) = \frac{1}{6}\sum_{n=1}^{6} \mathrm{e}^{-0.1n} \\ &= \frac{1}{6}\left(\mathrm{e}^{-0.1} + \mathrm{e}^{-0.2} + \cdots + \mathrm{e}^{-0.6}\right).\end{aligned}$$

これは初項 $\mathrm{e}^{-0.1}$、公比 $\mathrm{e}^{-0.1}$、項数 6 の等比級数（geometric series）であるから、

$$\mathrm{P}(D=0) = \frac{\mathrm{e}^{-0.1}}{6}\left(\frac{1-\mathrm{e}^{-0.6}}{1-\mathrm{e}^{-0.1}}\right),$$

となり、示すべき式が得られた。

期待値（expectation）についても同様に条件付けを行う。

$$\mathrm{E}(D) = \sum_{n=1}^{6} \mathrm{P}(N=n)\,\mathrm{E}(D \mid N=n) = \frac{1}{6}\sum_{n=1}^{6} 0.1n = \frac{0.1\times 21}{6} = 0.35.$$

言い換えれば $\mathrm{E}(D) = 0.1\,\mathrm{E}(N) = 0.1\times 3.5$ であり、これが示すべきことであった。

小問 (ii)

今度は、初めて 6 が出るのが何回目の投げかを T とする。各回の投げは独立であり、$T=t$ とは 6 以外が $t-1$ 回続いた後に 6 が出ることであるから、

$$\mathrm{P}(T=t) = \left(\tfrac{5}{6}\right)^{t-1}\tfrac{1}{6}, \qquad t = 1,2,3,\ldots$$

$T=t$ のとき、私は $0.1t\,\mathrm{kg}$ を買うので、$\mathrm{P}(D=0 \mid T=t) = \mathrm{e}^{-0.1t}$ である。したがって

$$\mathrm{P}(D=0) = \sum_{t=1}^{\infty}\left(\tfrac{5}{6}\right)^{t-1}\tfrac{1}{6}\,\mathrm{e}^{-0.1t} = \frac{\mathrm{e}^{-0.1}}{6}\sum_{t=1}^{\infty}\left(\tfrac{5}{6}\mathrm{e}^{-0.1}\right)^{t-1}.$$

考えてみよう なぜこの和を無限まで取ってよいのか？

これは公比 $r = \frac{5}{6}\mathrm{e}^{-0.1}$ の等比級数であり、$0<r<1$ であるから収束（converge）し、その和は $1/(1-r)$ である。よって

$$\mathrm{P}(D=0) = \frac{\mathrm{e}^{-0.1}}{6}\cdot\frac{1}{1-\frac{5}{6}\mathrm{e}^{-0.1}} = \frac{\mathrm{e}^{-0.1}}{6-5\mathrm{e}^{-0.1}},$$

ここで分子と分母に 6 を掛けた。これが求める結果である。

ダイヤモンドの個数の期待値は

$$\mathrm{E}(D) = \sum_{t=1}^{\infty}\mathrm{P}(T=t)\,\mathrm{E}(D \mid T=t) = \sum_{t=1}^{\infty}\left(\tfrac{5}{6}\right)^{t-1}\tfrac{1}{6}\times 0.1t = \frac{0.1}{6}\sum_{t=1}^{\infty} t\left(\tfrac{5}{6}\right)^{t-1}.$$

ここでヒントが役に立つ。$|x| < 1$ のとき、二項展開（binomial expansion）により

$$(1-x)^{-2} = 1 + 2x + 3x^2 + 4x^3 + \cdots = \sum_{t=1}^{\infty} t x^{t-1}.$$

$x = \frac{5}{6}$ とすると和は $\left(1 - \frac{5}{6}\right)^{-2} = 36$ となるので、

$$\mathrm{E}(D) = \frac{0.1}{6} \times 36 = 0.6.$$

これもまた $0.1\,\mathrm{E}(T)$ に等しい。ここで $\mathrm{E}(T) = 6$ は、6 が出るまでに必要な投げの回数の平均としてよく知られた値である。

検算

- 数値的には、(i) の確率は約 0.7150、(ii) の確率は約 0.6131 である。ダイヤモンドの個数をポアソン分布として買い物客 2×10^5 人をシミュレーションしたところ、0.7151 と 0.6136 が得られ、ダイヤモンドの個数の平均は 0.350 と 0.599 であった。
- (ii) の確率は (i) の確率より小さい。これは理にかなっている。平均して 350 g ではなく 600 g を買うからである。
- どちらの期待値も、確率的に決まる倍率の平均の 0.1 倍になっている。質量が与えられたときのダイヤモンドの個数の期待値は質量について線形であるから、ダイヤモンドの個数の期待値はキログラム数の期待値に等しいのである。

よくある間違い

- ダイヤモンドが 1 個もない確率として $\mathrm{e}^{-0.35}$（あるいは $\mathrm{e}^{-0.6}$）を計算してしまう。これは平均の質量しか使っていないが、e^{-m} は m について線形ではないので、平均を取るのは e^{-m} を計算した後でなければならず、前ではない。期待値は線形であり、それゆえ $\mathrm{E}(D)$ についてはこの近道がうまくいくのである。
- グラムをキログラムに換算し忘れて、平均を $0.1n$ ではなく $100n$ としてしまう。
- $\mathrm{P}(T = t) = (\frac{5}{6})^t \frac{1}{6}$ と書いてしまう。これは 6 以外の目を 1 回多く数えている。

STEP のその先へ

ここではポアソン分布のパラメータそのものが確率変数であり、そのため D は**混合分布**（mixture distribution）となる。用いた道具、すなわち全確率の法則と塔の法則（tower law）$\mathrm{E}(D) = \mathrm{E}\big(\mathrm{E}(D \mid T)\big)$ は、大学 1 年次の確率論の中心的な内容である。こ

のような問題は確率母関数（probability generating function）を使うと見通しよく扱える。$\mathrm{P}(D = 0)$ は T の母関数に $\mathrm{e}^{-0.1}$ を代入した値である。試行の回数を確率的に定め、さらに各試行での個数も確率的に定めるという考え方は、複合分布（compound distribution）や分岐過程（branching process）へとつながっていく。Grimmett and Welsh, *Probability: An Introduction*（確率論入門）は、これらすべてを扱っている。

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。