解答例

STEP I 1988、第 14 問

確率・統計 ・ 打ち切り正規分布の期待値

> **STEP I 1988、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/88-s1-q14

鍵となる考え方

変数 X_M は単に $\min(X, M)$ である。すなわち、正規確率変数（normal random variable）のうち M を超える部分をすべて M に押しつぶしたものである。その期待値（expectation）は $x < M$ での積分と、定数と裾確率（tail probability）の積とに分かれる。そして $x\phi(x)$ は符号を除けば $\phi(x)$ の導関数（derivative）なので、この積分は易しい。観光客の話も、同じ考え方が姿を変えて現れたものである。バスの車両群が運ぶ乗客数は $\min(\text{需要}, \text{定員})$ であり、正規近似（normal approximation）を施すと、これはまさに X_M の形になる。

解答

X_M の期待値

$\{X < M\}$ 上では $X_M = X$、$\{X \geq M\}$ 上では $X_M = M$ であるから、

$$\mathrm{E}(X_M) = \int_{-\infty}^{M} x\,\phi(x)\,\mathrm{d}x \;+\; M\,\mathrm{P}(X \geq M) = \int_{-\infty}^{M} x\,\phi(x)\,\mathrm{d}x \;+\; M(1 - \Phi(M)).$$

> **考えてみよう** 導関数が $x\phi(x)$ となる関数はあるだろうか。

ある。$\phi(x) = \frac{1}{\sqrt{2\pi}}\mathrm{e}^{-x^2/2}$ に対して、合成関数の微分法（chain rule）により $\phi'(x) = -x\phi(x)$ となる。したがって

$$\int_{-\infty}^{M} x\,\phi(x)\,\mathrm{d}x = \Big[-\phi(x)\Big]_{-\infty}^{M} = -\phi(M) + 0,$$

である。ここで $x \to -\infty$ のとき $\phi(x) \to 0$ であることを用いた。よって

$$\mathrm{E}(X_M) = -\phi(M) + M(1 - \Phi(M)),$$

となり、示すべき式が得られた。

妥当性の確認として、$M \to \infty$ とすると $M(1-\Phi(M)) \to 0$ であるから $\mathrm{E}(X_M) \to 0 = \mathrm{E}(X)$ となる。また $M \to -\infty$ とすると $\mathrm{E}(X_M) \approx M$ となり、当然そうあるべき結果である。

バス

1 回の来訪において、バス乗り場へ案内される観光客の人数を N とする。1024 人の観光客はそれぞれ独立に、どちらの方向へも確率 $\frac{1}{2}$ で案内されると仮定する。すると

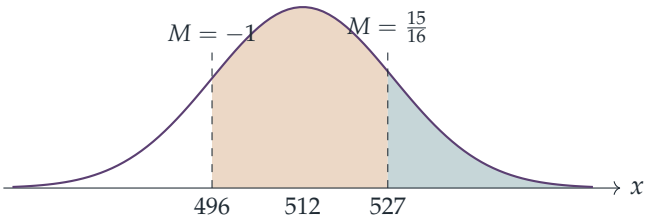
$$N \sim \mathrm{B}(1024, \tfrac{1}{2}), \qquad \mathrm{E}(N) = 512, \qquad \mathrm{Var}(N) = 1024 \cdot \tfrac{1}{2} \cdot \tfrac{1}{2} = 256 = 16^2.$$

である。$n = 1024$ とこれほど大きいので、N は X を標準正規分布（standard normal distribution）に従う変数として $512 + 16X$ で非常によく近似される。

バスが 16 台なら定員は $16 \times 31 = 496$、17 台なら 527 である。定員 C の車両群は $\min(N, C)$ 人の乗客を運ぶので、1 回の来訪あたりの利益（単位はヴロスカ）は

$$\begin{aligned}\min(N, C) &\approx \min(512 + 16X,\ C) \\ &= 512 + 16\min\Big(X, \frac{C - 512}{16}\Big) = 512 + 16\,X_M, \qquad \text{ただし} M = \frac{C - 512}{16}.\end{aligned}$$

となる。よって $C = 496$ のとき $M = -1$、$C = 527$ のとき $M = \frac{15}{16}$ である。



需要が琥珀色の帯に入るとき、17 台目のバスは 496 を超えた分の乗客を運ぶ。需要が青緑色の裾に入るとき、17 台目のバスは満員の 31 人を運ぶ。

$\phi(1) \approx 0.24197$、$\Phi(1) \approx 0.84134$ を用い、$1 - \Phi(-1) = \Phi(1)$ に注意すると、

$$\mathrm{E}(X_{-1}) = -\phi(1) - \big(1 - \Phi(-1)\big) = -\phi(1) - \Phi(1) \approx -0.24197 - 0.84134 = -1.08332.$$

となる。$\phi(0.9375) \approx 0.25707$、$1 - \Phi(0.9375) \approx 0.17425$ を用いると、

$$\mathrm{E}(X_{15/16}) \approx -0.25707 + 0.9375 \times 0.17425 \approx -0.25707 + 0.16336 = -0.09371.$$

となる。したがって、追加のバスによって 1 回の来訪あたりの期待利益は

$$16\big(\mathrm{E}(X_{15/16}) - \mathrm{E}(X_{-1})\big) \approx 16 \times 0.98961 \approx 15.83 \text{ ヴロスカ}.$$

だけ増える。来訪は年に 50 回あるので、追加のバスの価値はおよそ

$$50 \times 15.83 \approx 792 \text{ ヴロスカ（年あたり）}.$$

である。これが検討に値する賄賂の年額の上限である。期待される追加利益を超える額を支払えば、協同組合は平均として損をすることになる。

考えてみよう　N は離散的なのだから、連続性補正（continuity correction）を用いるべきではないのか。

この場合は必要ない。連続性補正が役立つのは、$\mathrm{P}(N \le c)$ のような確率を近似するとき、つまり平均をとる関数が整数のところで跳ぶときである。関数 $k \mapsto \min(k, C)$ には跳びがなく、補正なしの正規近似がこの関数に対しては極めて正確である。厳密な二項分布（binomial distribution）による計算では 791.68 となり、上の値と小数第 2 位まで一致する。定員を $\frac{1}{2}$ ずらすとかえって悪くなり、775 となってしまう。

検算

- $\int_{-\infty}^{M} x\phi(x)\,\mathrm{d}x + M(1-\Phi(M))$ を中点則（midpoint rule）で数値積分すると、いくつかの M の値について公式と小数第 8 位まで一致する。4×10^5 個の標本を用いた $\min(X, M)$ のモンテカルロ・シミュレーション（Monte Carlo simulation）も一致する。
- $\mathrm{B}(1024, \frac{1}{2})$ にわたる厳密な和をとると、1 回の来訪あたりの追加期待利益は 15.8337、すなわち年あたり 791.68 となり、正規近似と一致する。
- 大まかな妥当性：追加のバスは $N > 496$ のときには必ず何らかの利益を生み、これは約 0.84 の確率で起こり、しかもしばしば満員になる。したがって、1 回の来訪あたり 31 の半分を少し上回る値は妥当である。

よくある間違い

- 尺度の掛け忘れ：$\min(N, C)$ は $512 + 16X_M$ であるから、$\mathrm{E}(X_M)$ の差には標準偏差（standard deviation）16 を掛け、さらに 50 を掛けなければならない。
- $M = -1$ での符号の誤り：項 $M(1-\Phi(M))$ は $-(1-\Phi(-1)) = -\Phi(1)$ であり、$-(1-\Phi(1))$ ではない。

STEP のその先へ

量 $\mathrm{E}\min(X, M)$ は、**打ち切り**（censored）正規確率変数の期待値である。$\mathrm{E}(X-M)^{+} = \phi(M) - M(1-\Phi(M))$ のような密接に関連した式は、在庫理論（inventory theory、「新聞売り子問題（newsvendor problem）」）や、オプションのブラック–ショールズ（Black–Scholes）価格の根幹をなしている。恒等式 $\mathrm{E}\min(N, C+1) - \mathrm{E}\min(N, C) = \mathrm{P}(N > C)$ は、定員を 1 単位増やすことの価値を見通すうまい方法である。これらの考え方についてさらに知りたければ、Grimmett と Welsh の *Probability: An Introduction*（確率論入門）を参照するとよい。

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。