解答例

STEP II 1996、第 12 問

確率・統計 ・ 最大値と最小値の分布

> **STEP II 1996、問 12**
>
> 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/96-s2-q12

鍵となる考え方

いくつかの独立（independent）な確率変数の最大値は直接扱いにくい。しかし「最大値が小さい」という事象は扱いやすい。この事象が起こるのは、**すべての**変数が小さいときであり、かつそのときに限る。しかも独立性により確率を掛け合わせることができる。そこで、まず累積分布関数（cumulative distribution function）を求め、それを微分して確率密度関数（probability density function）を得る。小問 (ii) はその裏返しであり、**最小値**が大きいのは、すべての変数が大きいとき、かつそのときに限る。

解答

小問 (i)

各 X_j は $[0,1]$ 上の一様分布（uniform distribution）に従うので、$0 \le y \le 1$ に対して $\mathrm{P}(X_j < y) = y$ である。この範囲の y について、$Y < y$ が成り立つのは、すべての j について $X_j < y$ が成り立つとき、かつそのときに限る。X_j は独立であるから、

$$F(y) = \mathrm{P}(Y < y) = \prod_{j=1}^{n} \mathrm{P}(X_j < y) = y^n, \qquad 0 \le y \le 1.$$

もちろん $y < 0$ では $F(y) = 0$、$y > 1$ では $F(y) = 1$ である。微分すると、Y の確率密度関数は

$$f(y) = \begin{cases} ny^{n-1}, & 0 \le y \le 1, \\ 0, & \text{それ以外.} \end{cases}$$

となる（確認として、$\int_0^1 ny^{n-1}\,\mathrm{d}y = 1$ である）。

次にモーメント（moments）を求める。

$$\mathrm{E}(Y) = \int_0^1 y \cdot n y^{n-1}\,\mathrm{d}y = \frac{n}{n+1},$$
$$\mathrm{E}(Y^2) = \int_0^1 y^2 \cdot n y^{n-1}\,\mathrm{d}y = \frac{n}{n+2}.$$

したがって

$$\mathrm{Var}(Y) = \frac{n}{n+2} - \frac{n^2}{(n+1)^2} = \frac{n\left[(n+1)^2 - n(n+2)\right]}{(n+2)(n+1)^2} = \frac{n}{(n+2)(n+1)^2},$$

である。ここで $(n+1)^2 - n(n+2) = n^2 + 2n + 1 - n^2 - 2n = 1$ を用いた。これが求める結果である。

考えてみよう いきなり $\mathrm{P}(Y = y)$ や Y の密度関数を書き下そうとするのではなく、$\mathrm{P}(Y < y)$ を扱うのはなぜだろうか。

連続型の確率変数では、密度関数の値は確率ではない。また「最大値が y に等しい」という事象には、どの X_j が最大であるかということと、他の変数がどうなるかということが入り混じっている。一方、事象 $\{Y < y\}$ は n 個の独立な事象の共通部分なので、その確率は単純な積になる。分布関数を求めてから微分するというのが、最大値や最小値を扱う標準的な方法である。

小問 (ii)

j 番目の電球が故障する時刻を T_j とすると、$\mathrm{P}(T_j \le t) = 1 - \mathrm{e}^{-t/\lambda}$ であり、

$$\mathrm{P}(T_j > t) = \mathrm{e}^{-t/\lambda}, \qquad t > 0.$$

電球は互いに独立なので、T_j も独立である。最初の故障が起こる時刻を $T = \min(T_1, \ldots, T_n)$ とする。今度は事象 $T > t$ を考えるのが便利である。これが起こるのは、時刻 t においてすべての電球がまだ点いているとき、かつそのときに限る。

$$\mathrm{P}(T > t) = \prod_{j=1}^{n} \mathrm{P}(T_j > t) = \left(\mathrm{e}^{-t/\lambda}\right)^n = \mathrm{e}^{-nt/\lambda}.$$

よって $\mathrm{P}(T \le t) = 1 - \mathrm{e}^{-nt/\lambda}$ であり、微分すると密度関数

$$g(t) = \frac{n}{\lambda}\,\mathrm{e}^{-nt/\lambda}, \qquad t > 0.$$

が得られる。言い換えれば、T は電球 1 個と同じ種類の分布に従い、ただ λ が λ/n に置き換わっている。$u = t$ として部分積分（integration by parts）を行うと、

$$\begin{aligned}\mathrm{E}(T) &= \int_0^\infty t\,\frac{n}{\lambda}\,\mathrm{e}^{-nt/\lambda}\,\mathrm{d}t = \left[-t\,\mathrm{e}^{-nt/\lambda}\right]_0^\infty + \int_0^\infty \mathrm{e}^{-nt/\lambda}\,\mathrm{d}t \\ &= 0 + \left[-\frac{\lambda}{n}\,\mathrm{e}^{-nt/\lambda}\right]_0^\infty = \frac{\lambda}{n},\end{aligned}$$

となる。ここで、$t \to \infty$ のとき $te^{-nt/\lambda} \to 0$ であるから、境界項は 0 になる。したがって、最初の故障までの時間の期待値（expected value）は λ/n であり、これが求める結果である。

検算

- $n = 1$ のとき、Y は単に $[0,1]$ 上の一様分布に従い、公式は $\frac{1}{3\cdot4} = \frac{1}{12}$ を与える。これはよく知られた一様分布の分散（variance）である。
- n が大きくなると、分散は $1/n^2$ のように振る舞う。多数の一様変数の最大値は 1 の近くに押し寄せられるので、予想どおりである。
- 小問 (ii) で $n = 1$ とすると、電球 1 個の寿命の期待値は λ となる。これはパラメータ $1/\lambda$ の指数分布（exponential distribution）の平均である。
- 5 個の一様変数の最大値を 200 000 回シミュレーションしたところ、分散は 0.01993 となり、$\frac{5}{7\cdot36} \approx 0.01984$ とよく合った。また $\lambda = 3$ として 4 個の電球をシミュレーションしたところ、最初の故障時刻の平均は 0.749 となり、$\frac{3}{4}$ とよく合った。

よくある間違い

- $\mathrm{P}(T > t)$ を微分するときに負号を忘れ、負の「密度関数」を得てしまう。
- $(\mathrm{E}(Y))^2$ を引くべきところを、$\mathrm{Var}(Y) = \mathrm{E}(Y^2) - \mathrm{E}(Y)$ と書いてしまう。
- n 個の電球はそれぞれ平均して時刻 λ に故障するのだから、最初の故障も λ に起こると考えてしまう。n 個の変数の最小値は、そのどれか 1 つよりもずっと小さい。

STEP のその先へ

最大値と最小値は**順序統計量**（order statistics）の両極端の場合である。標本を小さい順に並べたとき、n 個の一様変数のうち k 番目に小さいものはベータ分布（beta distribution）に従い、ここでの Y は $k = n$ の場合にあたる。小問 (ii) は、独立な指数分布に従う変数の最小値は再び指数分布に従い、その率（rate）は足し合わされるという重要な事実を示している。この事実は、ポアソン過程（Poisson process）や連続時間マルコフ連鎖（continuous-time Markov chain）の理論の原動力となっている。Grimmett と Welsh の *Probability: An Introduction*（確率論入門）や、Blitzstein と Hwang の *Introduction to Probability*（確率論入門）は、いずれもこれらの考え方を展開している。

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。