解答例

STEP I 1993、第 14 問

確率・統計 ・ 指数分布に従う待ち時間

> **STEP I 1993、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/93-s1-q14

鍵となる考え方

重要なのはトード氏（Mr Toad）の速さではなく、最後の 1 km を進むのにかかる**時間** T であり、$T = 1/V$ である。V の確率密度関数（density）は扱いにくそうに見えるが、その累積分布関数（cumulative distribution function）は非常に単純であり、これを T についての主張に書き換えると、T がパラメータ α の指数分布（exponential distribution）に従うことがわかる。そうすれば、青信号の各期間に対する確率は二つの指数関数の差となり、青信号の期間すべてについて足し合わせると等比級数（geometric series）が現れる。

解答

到着時刻の分布

V を分速（km 毎分）で表した速さとし、$T = 1/V$ を、信号が青に変わった瞬間から測って、彼が信号に到着するまでの時間（分）とする。

$$\frac{\mathrm{d}}{\mathrm{d}v}\,\mathrm{e}^{-\alpha/v} = \alpha v^{-2}\mathrm{e}^{-\alpha/v},$$

であること、および $v \to 0^+$ のとき $\mathrm{e}^{-\alpha/v} \to 0$ であることに気づけば、V の分布関数（distribution function）は

$$\mathrm{P}(V \le v) = \int_0^v \alpha u^{-2}\mathrm{e}^{-\alpha/u}\,\mathrm{d}u = \mathrm{e}^{-\alpha/v} \qquad (v > 0).$$

となる（$v \to \infty$ のときこれは 1 に収束するので、確かに確率密度関数になっている）。$V > 0$ であるから、$t > 0$ に対して $T > t \iff 1/V > t \iff V < 1/t$ が成り立ち、

したがって

$$\mathrm{P}(T > t) = \mathrm{P}\left(V < \tfrac{1}{t}\right) = \mathrm{e}^{-\alpha t}.$$

である。よって T はパラメータ α の指数分布に従う。

考えてみよう なぜ V ではなく T を扱うのか？

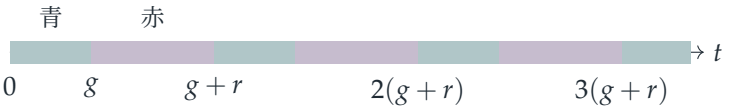
信号の状態は時間によって記述されるので、考えたい事象はすべて T の値の区間として表される。V の確率密度関数を手計算で T の確率密度関数に書き換えるには変数変換（change of variable）が必要になるが、$\mathrm{P}(T > t)$ を経由すればそれを避けられ、答えが一行で得られる。

青信号の一つの期間

時刻 0 に信号が青に変わる。信号は $[0, g)$ で青、$[g, g+r)$ で赤、$[g+r, (g+r)+g)$ で青となり、以下同様に続く。$(n+1)$ 番目の青信号の期間は $n(g+r) \le t < n(g+r)+g$ である。T は連続型なので、端点を含めるかどうかは結果に影響せず、

$$\begin{aligned}\mathrm{P}(n(g+r) < T < n(g+r)+g) &= \mathrm{P}(T > n(g+r)) - \mathrm{P}(T > n(g+r)+g) \\ &= \mathrm{e}^{-\alpha n(g+r)} - \mathrm{e}^{-\alpha(n(g+r)+g)},\end{aligned}$$

となり、求める式が得られる（問題文では二つ目の指数において閉じ括弧の位置が誤っている。指数はここに示したとおり $-\alpha(n(g+r)+g)$ でなければならない）。この計算は $n = 0$、すなわち信号が一度も赤になる前の最初の青信号の期間に対しても同様に成り立つ。



青信号で通過する確率

彼が青信号で通過するのは、T がこれらの青信号の期間のいずれかに入るとき、かつそのときに限る。これらの期間は互いに交わらないので、$n = 0$ も含めてそれぞれの確率を足し合わせればよい：

$$\begin{aligned}\mathrm{P}(\alpha) &= \sum_{n=0}^{\infty}\left(\mathrm{e}^{-\alpha n(g+r)} - \mathrm{e}^{-\alpha n(g+r)}\mathrm{e}^{-\alpha g}\right) = (1-\mathrm{e}^{-\alpha g})\sum_{n=0}^{\infty}(\mathrm{e}^{-\alpha(g+r)})^n \\ &= \frac{1-\mathrm{e}^{-\alpha g}}{1-\mathrm{e}^{-\alpha(g+r)}},\end{aligned}$$

ここで等比級数の和を求めた。$\alpha > 0$ かつ $g + r > 0$ より $0 < \mathrm{e}^{-\alpha(g+r)} < 1$ であるから、この級数は収束（converge）する。

極限

$\alpha \to \infty$ のとき、$g > 0$、$g + r > 0$ であるから $e^{-\alpha g} \to 0$ かつ $e^{-\alpha(g+r)} \to 0$ となり、

$$P(\alpha) \to \frac{1-0}{1-0} = 1.$$

である。これは納得のいく結果である。α が大きいことはトード氏が非常に速いことを意味し、その場合ほぼ確実に最初の青信号の期間が終わる前に通過できる。

$\alpha \to 0$ のときは分子と分母がともに 0 に収束するので、もっと注意が必要である。それぞれにおいて $1 - e^{-x} = e^{-x}(e^x - 1)$ と書き直すと、

$$P(\alpha) = \frac{e^{-\alpha g}\,(e^{\alpha g} - 1)}{e^{-\alpha(g+r)}\,(e^{\alpha(g+r)} - 1)} = e^{\alpha r} \cdot \frac{\dfrac{e^{\alpha g} - 1}{\alpha g}}{\dfrac{e^{\alpha(g+r)} - 1}{\alpha(g+r)}} \cdot \frac{g}{g+r},$$

となる。ここで分子と分母を α で割り、因子 g と $g + r$ を挿入した。$\alpha \to 0$ のとき $e^{\alpha r} \to 1$ であり、また $\alpha g \to 0$ かつ $\alpha(g + r) \to 0$ であるから、$(e^x - 1)/x \to 1$ により二つの分数はいずれも 1 に収束する。したがって

$$P(\alpha) \to \frac{g}{r+g} \qquad \text{ただし} \alpha \to 0.$$

である。この場合トード氏は非常に遅く、到着時刻が非常に多くの周期にわたって広がっているので、青信号である確率は、各周期のうち青である時間の割合にほかならない。

検算

- $\alpha = 0.7$、$g = 1$、$r = 2$ のとき、公式から $P(\alpha) \approx 0.5737$ が得られる。$e^{-\alpha/v}$ の逆関数を用いて V を生成し、4×10^5 回の走行をシミュレーションしたところ 0.5744 となり、$n = 0, 1, 2$ の各青信号の期間に対する確率も一致した。
- $g = 1$、$r = 2$ のとき：$P(10^{-4}) \approx 0.33337$ であり $\frac{1}{3}$ に近く、また $P(10) \approx 0.99995$ である。
- $r = 0$ のとき信号は常に青であり、公式はすべての α に対して $P(\alpha) = 1$ を与える。これは当然そうなるべき結果である。

よくある間違い

- $n = 0$ を含め忘れる。一度も赤になる前の最初の青信号の期間が、最も大きな確率をもつ。

- 問題文に印刷された指数 $-\alpha(n(g+r))+g$ をそのまま写してしまう。これは $-\alpha\big(n(g+r)+g\big)$ の誤植（misprint）である。
- V そのものを到着時刻として扱ったり、$\mathrm{E}(T)$ を求めようとしたりする。実はここでは $\mathrm{E}(V)$ は無限大であるが、この問題ではそれは一切必要ない。

STEP のその先へ

分布関数を通じて V の分布を $T=g(V)$ の分布に変換することは、確率変数（random variable）の変換の標準的な方法である。ここで V は逆指数分布（inverse-exponential distribution）に従っている。極限 $\alpha\to 0$ は、ある一般的な考え方の小さな例になっている。すなわち、多くの周期にわたって広がった確率的な時刻は、各局面にその局面の長さに比例した確率で入る。これが再生理論（renewal theory）の出発点である。Grimmett と Welsh の *Probability: An Introduction*（確率論入門）では、確率変数の変換と指数分布の両方が扱われている。

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。