解答例

STEP I 2007、第 14 問

確率・統計 ・ ポアソン分布の母数の一意性

> **STEP I 2007、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/07-s1-q14

鍵となる考え方

それぞれの情報は $h(\lambda) =$ 定数 の形の方程式を与える。ここで h は平均（mean）のある関数である。この方程式が λ をただ一つに定めるかどうかは、$\lambda > 0$ における h の形だけで決まる。狭義単調（strictly monotonic）な h は各値を高々 1 回しかとらない。一方、増加してから減少する h は、ほとんどの値を 2 回とり、最大値だけを 1 回とる。したがって、この問題の本質は、各関数がどこで単調であり、どこで最大となるかを調べることにある。

解答

以下では常に $\lambda > 0$（ポアソン分布（Poisson distribution）の平均は正である）とし、$\mathrm{P}(X = k) = \mathrm{e}^{-\lambda}\lambda^k/k!$ for $k = 0, 1, 2, \ldots$ である。

小問 (i)

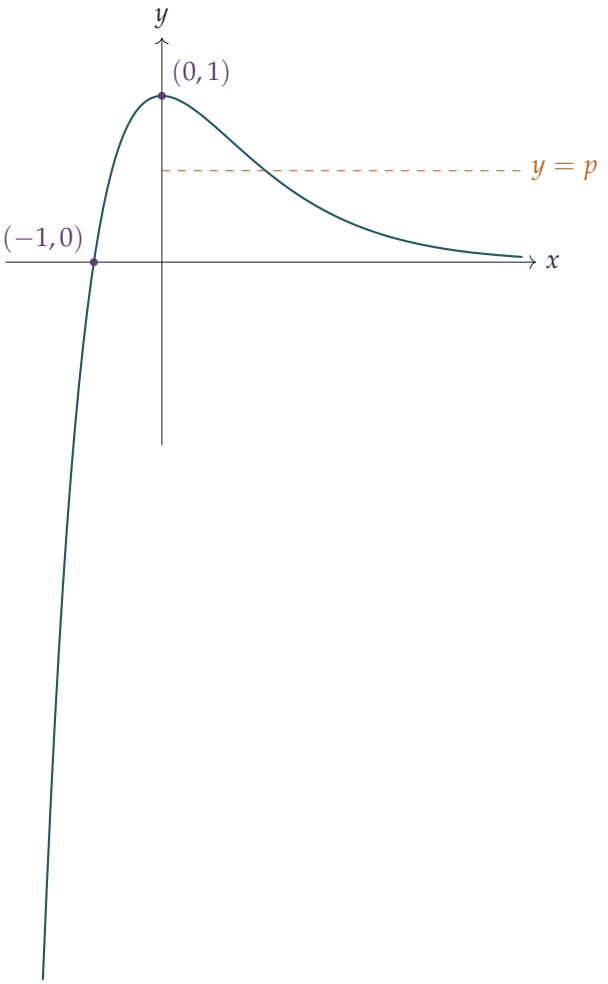
$f(x) = (x+1)\mathrm{e}^{-x}$ とおく。積の微分法（product rule）により、

$$f'(x) = \mathrm{e}^{-x} - (x+1)\mathrm{e}^{-x} = -x\,\mathrm{e}^{-x}.$$

$\mathrm{e}^{-x} > 0$ であるから、$f'(x)$ の符号は x の符号と逆である。すなわち、f は $x < 0$ で増加し、$x > 0$ で減少する。よって極値をとる点（turning point）はただ一つで、それは $(0, 1)$ における極大（maximum）であり、この点は曲線が y 軸と交わる点でもある。曲線は $x + 1 = 0$ となるところ、すなわち $(-1, 0)$ で x 軸と交わる。

両端の様子については次のとおりである。$x \to \infty$ のとき、e^{-x} が一次の因子に勝り、$f(x) \to 0$ となる（上から近づき、x 軸が漸近線（asymptote）となる）。$x \to -\infty$ のと

き、両方の因子の大きさが増大し、$f(x) \to -\infty$ となる。



次に確率を考える。

$$\mathrm{P}(X \geq 2) = 1 - \mathrm{P}(X = 0) - \mathrm{P}(X = 1) = 1 - \mathrm{e}^{-\lambda} - \lambda \mathrm{e}^{-\lambda} = 1 - f(\lambda),$$

であるから、与えられた情報はちょうど

$$f(\lambda) = (1 + \lambda)\mathrm{e}^{-\lambda} = p.$$

ということである。$\lambda > 0$ において関数 f は連続かつ狭義単調減少であり（$f'(\lambda) = -\lambda \mathrm{e}^{-\lambda} < 0$ による）、$f(0) = 1$ から、$\lambda \to \infty$ のとき 0 に向かって減少する。したがって、$(0, 1)$ のすべての値をちょうど 1 回ずつとる。実際、中間値の定理（intermediate value theorem）により解が少なくとも一つ存在し、狭義単調性により二つ目の解は存在しない。グラフでいえば、$0 < p < 1$ のとき水平線 $y = p$ は右側の部分とちょうど 1 回交わる。よって λ はただ一つに定まる。

小問 (ii)

ここでは $\mathrm{P}(X=1)=\lambda \mathrm{e}^{-\lambda}=q$ である。$g(\lambda)=\lambda \mathrm{e}^{-\lambda}$ とおく。すると

$$g'(\lambda)=(1-\lambda)\mathrm{e}^{-\lambda},$$

となるので、g は $(0,1]$ で狭義単調増加、$[1,\infty)$ で狭義単調減少であり、最大値は $g(1)=\mathrm{e}^{-1}$ である。また、$\lambda \to 0^+$ のときも $\lambda \to \infty$ のときも $g(\lambda)\to 0$ である。(実は $g(\lambda)=\mathrm{e}^{-1}f(\lambda-1)$ であるから、g のグラフは小問 (i) の曲線を右に 1 だけ平行移動し、縦方向に e^{-1} 倍したものである。)

考えてみよう 頂点まで増加してから減少する曲線と、水平線は何回交わりうるか？

単調な各部分とは高々 1 回しか交わらないので、次のようになる。

- $0<q<\mathrm{e}^{-1}$ ならば、$(0,1)$ に解が一つ、$(1,\infty)$ にもう一つあり、λ の候補は二つある。
- $q>\mathrm{e}^{-1}$ ならば、解は存在しない。
- $q=\mathrm{e}^{-1}$ ならば、解は頂点そのもの $\lambda=1$ だけである。

したがって、この情報が λ をただ一つに定めるのは、q がちょうど一つの値をとるとき、すなわち

$$q=\mathrm{e}^{-1}, \qquad \lambda=1.$$

のときに限る。

小問 (iii)

条件付き確率（conditional probability）の定義により、また事象 $\{X=1\}$ は $\{X\le 2\}$ に含まれるから、

$$\mathrm{P}(X=1\mid X\le 2)=\frac{\mathrm{P}(X=1)}{\mathrm{P}(X\le 2)}=\frac{\lambda \mathrm{e}^{-\lambda}}{\mathrm{e}^{-\lambda}\left(1+\lambda+\frac{1}{2}\lambda^2\right)}=\frac{2\lambda}{2+2\lambda+\lambda^2}.$$

これを $k(\lambda)$ とおく。分子と分母を λ で割ると（$\lambda>0$ であるから許される）、

$$k(\lambda)=\frac{2}{\lambda+\dfrac{2}{\lambda}+2}.$$

考えてみよう $\lambda>0$ のとき、$\lambda+2/\lambda$ の最小値は何か？

相加平均と相乗平均の不等式（AM–GM inequality）により $\lambda + \dfrac{2}{\lambda} \geq 2\sqrt{\lambda \cdot \dfrac{2}{\lambda}} = 2\sqrt{2}$ であり、等号が成り立つのは $\lambda = 2/\lambda$、すなわち $\lambda = \sqrt{2}$ のときに限る。さらに、$\lambda + 2/\lambda$ は $(0, \sqrt{2}]$ で狭義単調減少、$[\sqrt{2}, \infty)$ で狭義単調増加であり（その導関数 $1 - 2/\lambda^2$ の符号が変わるのは $\sqrt{2}$ においてのみである）、両端で ∞ に発散する。よって k は小問 (ii) の g と同じく、増加してから減少する形をしており、最大値は

$$k(\sqrt{2}) = \frac{2}{2\sqrt{2}+2} = \frac{1}{\sqrt{2}+1} = \sqrt{2} - 1.$$

である。前と全く同様に、$0 < r < \sqrt{2} - 1$ のとき λ の値は二つ、$r > \sqrt{2} - 1$ のとき存在せず、$r = \sqrt{2} - 1$ のときちょうど一つである。したがって、λ がただ一つに定まるのは

$$r = \sqrt{2} - 1, \qquad \lambda = \sqrt{2}.$$

のときに限る。

別解として、$k(\lambda) = r$ の分母を払うと、二次方程式 $r\lambda^2 - 2(1-r)\lambda + 2r = 0$ が得られる。その解の積は $2 > 0$、和は $2(1-r)/r > 0$ であるから、実数解があればどちらも正である。解がちょうど一つとなるのは判別式（discriminant）$4\left[(1-r)^2 - 2r^2\right] = 4(1 - 2r - r^2)$ が 0 になるときであり、これより（正の解をとって）$r = \sqrt{2} - 1$、そのときの重解（double root）は $\lambda = (1-r)/r = \sqrt{2}$ である。

検算

- 小問 (ii)：$\lambda = 1$ のとき $\mathrm{P}(X = 1) = \mathrm{e}^{-1} \approx 0.3679$ である。$0 < \lambda < 40$ の範囲で $\lambda \mathrm{e}^{-\lambda}$ を数値的に調べると、この値を超えることはなく、$q = 0.3$ では交点が二つある。
- 小問 (iii)：$\lambda = \sqrt{2}$ のとき $\dfrac{2\sqrt{2}}{2 + 2\sqrt{2} + 2} = \dfrac{\sqrt{2}}{2 + \sqrt{2}} = \sqrt{2} - 1 \approx 0.4142$ であり、数値的に調べるとこれが最大値であることが確かめられる。$r = 0.35$ では λ の値が二つ、$r = 0.45$ では存在しない。
- 妥当性の確認：$\lambda \to 0$ のとき確率のほぼすべてが 0 に集まり、$\lambda \to \infty$ のとき $X \leq 2$ という条件の下での確率は 2 に集中する。よってどちらの極限でも $\mathrm{P}(X = 1 \mid X \leq 2) \to 0$ となり、式が示すとおりである。

よくある間違い

- $\lambda > 0$ であることを忘れる：小問 (i) の曲線は $x < 0$ でも存在するが、確率に関係するのは $x > 0$ の部分だけである。

- 小問 (i) で理由を述べずに一意性を主張する：解が存在すること（連続性と、値が 1 から 0 まで動くこと）と、二つ目の解がないこと（狭義単調性）の両方が必要である。
- 小問 (iii) で $\mathrm{P}(X = 1 \text{ かつ} X \leq 2) = \mathrm{P}(X = 1)\,\mathrm{P}(X \leq 2)$ を用いる：これらの事象は独立（independent）ではない。そうではなく、$\{X = 1\}$ が $\{X \leq 2\}$ に含まれているのである。

STEP のその先へ

これは**識別可能性**（identifiability）への最初の入り口である。統計学では、観測される量がモデルのパラメータを決定するかどうかを問題にする。その答えが「はい」となるのは、パラメータから観測量への写像が一対一（one-to-one）であるときにちょうど限る。ここでは $\mathrm{P}(X \geq 2)$ は λ を識別するが、$\mathrm{P}(X = 1)$ や $\mathrm{P}(X = 1 \mid X \leq 2)$ は一般には識別しない。同じ「増加してから減少する」形は、尤度（likelihood）を最大化するときにも再び現れる。$\lambda \mathrm{e}^{-\lambda}$ は観測値 $X = 1$ を一つ得たときの λ の尤度であり、その頂点 $\lambda = 1$ が最尤推定値（maximum likelihood estimate）である。ポアソン分布と条件付けについてさらに学ぶには、Blitzstein と Hwang による *Introduction to Probability*（確率論入門）が生き生きとした読み物である。

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。