解答例

STEP II 1995、第 13 問

確率・統計 ・ ポアソン分布と最適化

STEP II 1995、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/95-s2-q13

鍵となる考え方

データベース上の分類はともかく、これはビジネスの装いをした確率の問題である。来なかった客の人数は、試行回数が大きく成功確率がきわめて小さい二項分布（binomial distribution）に従う数であり、これはまさにポアソン分布（Poisson distribution）が取って代わる状況である。その後、整数 k についての最適化のこつは微分ではなく階差（*first difference*） $u_{k+1} - u_k$ を見ることである。チケットをもう 1 枚売ることで期待利益が増える限り、売り続ければよい。この階差が減少することが示されれば、最適な k は階差が初めて正でなくなるところである。

解答

ポアソンモデル

$N + k$ 人のチケット所持者がそれぞれ独立に確率 ϵ で現れないとする。このとき、現れない人数 R は二項分布 $\mathrm{B}(N+k, \epsilon)$ に従い、その平均は $(N+k)\epsilon$ である。k は N に比べて非常に小さいので、この平均はおよそ $N\epsilon = \lambda$ である。試行回数が大きく成功確率が非常に小さい二項分布は、同じ平均をもつポアソン分布でよく近似される。実際、r を固定し、$n\epsilon = \lambda$ を一定に保ったまま $n \to \infty$ とすると、

$$\binom{n}{r}\epsilon^r(1-\epsilon)^{n-r} = \frac{n(n-1)\cdots(n-r+1)}{n^r}\cdot\frac{\lambda^r}{r!}\left(1-\frac{\lambda}{n}\right)^{n-r} \to \frac{\lambda^r}{r!}\,\mathrm{e}^{-\lambda},$$

となる。最初の因子は 1 に収束し、$(1-\lambda/n)^n \to \mathrm{e}^{-\lambda}$ だからである。したがって、近似的に $R \sim \mathrm{Po}(\lambda)$ である。以下ではこれを厳密なものとして扱い、

$$p_r = \mathrm{P}(R = r) = \frac{\lambda^r}{r!}\,\mathrm{e}^{-\lambda}.$$

と書く。

利益

航空会社は $N+k$ 枚のチケットを 1 枚あたり £A の利益で販売するので、何が起ころうと $A(N+k)$ を得る。r 人の乗客が現れなければ、N 席に対して $N+k-r$ 人が到着する。

- $r \geq k$ ならば、到着するのは高々 N 人であり、全員が搭乗できる。補償を受ける者はおらず、$\max(0, k-r) = 0$ である。
- $r < k$ ならば、$k-r$ 人が搭乗できず、1 人につき £B の費用がかかり、$\max(0, k-r) = k-r$ である。

いずれの場合も、利益（ポンド）は $A(N+k) - B\max(0, k-r)$ である。

期待利益 $u_0, \ldots, u_3$

期待値（expectation）をとると、ペナルティに寄与するのは $r = 0, 1, \ldots, k-1$ のみであるから、

$$u_k = A(N+k) - B\sum_{r=0}^{k-1}(k-r)\,p_r.$$

$p_0 = \mathrm{e}^{-\lambda}$、$p_1 = \lambda\mathrm{e}^{-\lambda}$、$p_2 = \frac{1}{2}\lambda^2\mathrm{e}^{-\lambda}$ を用いると、

$$\begin{aligned}
u_0 &= AN,\\
u_1 &= A(N+1) - B\mathrm{e}^{-\lambda},\\
u_2 &= A(N+2) - B\mathrm{e}^{-\lambda}(2+\lambda),\\
u_3 &= A(N+3) - B\mathrm{e}^{-\lambda}\left(3 + 2\lambda + \tfrac{1}{2}\lambda^2\right).
\end{aligned}$$

階差 v_k

一般式どうしを引くと、

$$\begin{aligned}
v_k = u_{k+1} - u_k &= A - B\left[\sum_{r=0}^{k}(k+1-r)\,p_r - \sum_{r=0}^{k-1}(k-r)\,p_r\right]\\
&= A - B\left[p_k + \sum_{r=0}^{k-1}\big((k+1-r) - (k-r)\big)p_r\right]\\
&= A - B\sum_{r=0}^{k} p_r,
\end{aligned}$$

となる。ここで、第 1 の和の $r = k$ の項 $(k+1-k)p_k = p_k$ には、第 2 の和に対応する項がない。言葉で言えば、$v_k = A - B\,\mathrm{P}(R \leq k)$ である。これは納得のいく結果である。

$(k+1)$ 枚目の追加チケットは確実に $£A$ を稼ぎ、その乗客が搭乗を断られるとき、すなわち $R \le k$ のときに限って追加の $£B$ がかかるのである。

考えてみよう なぜ括弧の中は確率の単純な和にまとまるのだろうか。

各 $r \le k-1$ について、k が 1 増えると p_r の係数はちょうど 1 増える（その結果において搭乗を断られる乗客が 1 人増える）。また、結果 $r = k$ は新たに 1 人分の支払いを生じ始める。したがって、0 から k までのすべての r が p_r をちょうど 1 回ずつ寄与する。

すると

$$v_k - v_{k+1} = B\,p_{k+1} = B\,\frac{\lambda^{k+1}}{(k+1)!}\,\mathrm{e}^{-\lambda} > 0,$$

である。$B > 0$ かつ $\lambda > 0$ だからである。よって、すべての k について $v_k > v_{k+1}$ であり、階差は狭義単調減少する。

最後に、$\sum_{r=0}^{k} p_r$ は $\mathrm{e}^{-\lambda}\sum_{r\ge 0}\lambda^r/r! = \mathrm{e}^{-\lambda}\mathrm{e}^{\lambda} = 1$ の部分和（partial sum）であるから、$k \to \infty$ のとき 1 に収束し、

$$v_k \to A - B \qquad \text{極限} k \to \infty,$$

となり、示すべきことが示された。$B > A$ であるから、この極限は負である。

Fly By Night 社への助言

$F(k) = \mathrm{P}(R \le k) = \sum_{r=0}^{k} p_r$ とおくと、$v_k = A - BF(k)$ であり、

$$u_{k+1} > u_k \iff v_k > 0 \iff F(k) < \frac{A}{B}.$$

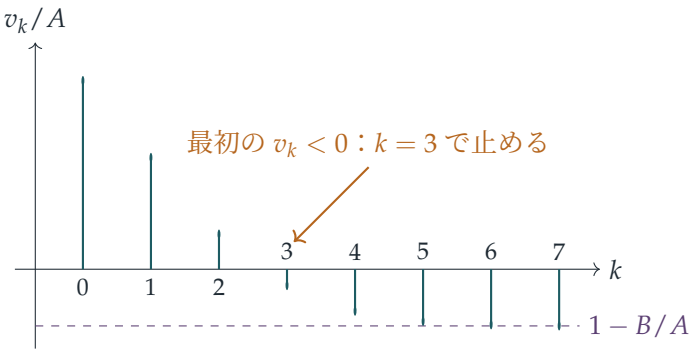
v_k は狭義単調減少で極限が負であるから、数列 $v_0, v_1, v_2, \ldots$ はある点までは正で、それ以降は負（または一度だけ 0）である。したがって u_k は増加したのち減少し、その最大値は、次のチケットがもはや利益を生まなくなる最初の k でとられる。

次を満たす最小のk を選ぶ：$F(k) = \displaystyle\sum_{r=0}^{k} \frac{\lambda^r}{r!}\mathrm{e}^{-\lambda} \;\ge\; \frac{A}{B}.$

特に、$\mathrm{e}^{-\lambda} \ge A/B$（乗客が非常に律儀であるか、ペナルティが非常に重い場合）ならば $v_0 \le 0$ であり、航空会社はオーバーブッキング（overbooking）をまったくすべきでない。すなわち $k = 0$ である。ちょうど $F(k) = A/B$ ならば $v_k = 0$ であり、k と $k+1$ は同じ期待利益を与える。$A/B < 1$ かつ $F(k) \to 1$ であるから、このような k は必ず存在する。

したがって助言は次のとおりである。高々 k 人しか現れない確率が、チケットの利益と補償額の比 A/B より小さい間は追加チケットを売り続け、その確率が初めて A/B に達した k で止めよ。

例えば、$\lambda = 2$、$B = 1.25A$（したがって $A/B = 0.8$）のとき、$F(0) \approx 0.135$、$F(1) \approx 0.406$、$F(2) \approx 0.677$、$F(3) \approx 0.857$ となるので、最適な選択は $k = 3$ である。図はこの場合の v_k/A を示している。



検算

- 一般式から、$u_1 - u_0 = A - Be^{-\lambda} = A - Bp_0$、$u_2 - u_1 = A - Be^{-\lambda}(1 + \lambda) = A - B(p_0 + p_1)$ であり、v_0 と v_1 に一致する。
- 極端な場合：$B \to A$ ならば $A/B \to 1$ となり、k は限りなく大きくなる（搭乗拒否の費用がほぼかからない）。$B \to \infty$ ならば $A/B \to 0$ となり、$k = 0$ である。
- 数値的には、N, A, B, λ を何百通りも無作為に選び、利益の式から直接 u_k を計算して、$u_0, \ldots, u_3$ と v_k の式、$v_k - v_{k+1} = Bp_{k+1}$ であること、および u_k を最大にする k が $F(k) \geq A/B$ を満たす最小の k であることを確かめた。$N = 400$、$k = 3$、$\epsilon = 0.01$ のとき、現れない人数が $r = 3$ となる二項分布とポアソン分布による確率は、それぞれ 0.194 と 0.195 である。

よくある間違い

- ペナルティの和に $r \geq k$ の項を含めてしまうこと。このときは現れた全員が座席を得るので、これらの項は何も寄与しない。
- u_k を k で微分してしまうこと。変数は整数であり、階差 v_k を用いるのが正しい道具である。
- 答えを「$v_k = 0$」と述べること。一般に v_k をちょうど 0 にする k は存在しない。答えは $v_k \leq 0$ となる最初の k である。

STEP のその先へ

「階差が減少するので、その符号が変わるところで止める」という議論は、凹関数（concave function）が導関数の 0 になる点で最大値をとることの離散版である。「$\mathrm{P}(R \leq k)$ が A/B に達したら止める」という規則は、オペレーションズ・リサーチ（operations research）における古典的な**新聞売り子問題**（*newsvendor*）の規則（臨界比率（critical fractile）規則）であり、航空会社のオーバーブッキングはその標準的な応用例である。上で厳密に示した二項分布のポアソン極限は、大学の確率論の講義で最初に学ぶ極限定理（limit theorem）の一つである。Grimmett と Welsh の *Probability: An Introduction*（確率論入門）は、ポアソン近似と、$\max(0, k - R)$ のような関数の期待値の両方を扱っている。

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。