解答例

STEP II 2003、第 14 問

確率・統計 ・ 仮説検定の誤り

STEP II 2003、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/03-s2-q14

鍵となる考え方

最初の 6 が出るまでに必要な投げる回数は幾何分布（geometric distribution）に従い、その裾（tail）は驚くほど簡単な形をしている。すなわち、「M 回より多く投げる」という事象は、単に「最初の M 回がすべて失敗する」という事象にほかならない。したがって、2 つの誤りの確率はどちらも純粋な M 乗の形になり、両者から M を消去するには対数をとればよい。

解答

2 つの誤りの確率

各回の投げは独立（independent）であり、どの回も同じ確率 p で 6 が出るとモデル化する。このとき、最初の 6 が出る回までを含めた投げる回数 X は

$$X > M \iff \text{最初の } M \text{ 回のどれも 6 でない}, \qquad \text{ゆえに} \qquad \mathrm{P}(X > M) = (1-p)^M,$$

を満たす。ここで各回の投げの独立性を用いた。（幾何級数 $\sum p(1-p)^{k-1}$ の和を計算する必要はない。事象を直接考えるほうが早い。）

この検定は、ちょうど $X > M$ のときに H_1 を採択する。したがって、実際には $p = B$ であるときに H_1 を採択する確率 α は

$$\alpha = (1-B)^M,$$

であり、実際には $p = A$ であるときに H_0 を採択する（すなわち $X \le M$ となる）確率 β は

$$\beta = 1 - \mathrm{P}(X > M \mid p = A) = 1 - (1-A)^M.$$

である。

M の消去

$(1-A)^M$ を $\alpha = (1-B)^M$ の累乗として表したい。$0 < B < 1$ であるから $\ln(1-B)$ は定義され、かつ 0 ではない。そこで α の対数をとると

$$M = \frac{\ln \alpha}{\ln(1-B)}.$$

を得る。すると

$$(1-A)^M = \mathrm{e}^{M\ln(1-A)} = \exp\left(\frac{\ln(1-A)}{\ln(1-B)}\ln\alpha\right) = \alpha^K, \qquad \text{ただし} \quad K = \frac{\ln(1-A)}{\ln(1-B)}.$$

である。よって

$$\beta = 1-\alpha^K, \qquad K = \frac{\ln(1-A)}{\ln(1-B)},$$

となり、求められていたとおり、K は A と B のみに依存し、M には依存しない。

考えてみよう K の大きさについて何がいえるだろうか。

$0 < A < B < 1$ より $1 > 1-A > 1-B > 0$ であるから、2 つの対数はどちらも負であり、$\ln(1-B) < \ln(1-A) < 0$ である。割り算をすると、K は正であって

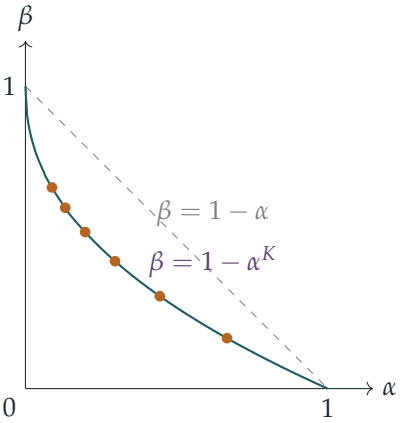
$$0 < K < 1.$$

となる。これがグラフの形を決定づける。

α に対する β のグラフ

M が $1, 2, 3, \ldots$ と動くとき、$\alpha = (1-B)^M$ は $(0,1)$ の値をとりながら 0 に向かって減少していく。厳密にはグラフは点の列であり、それらはすべて曲線 $\beta = 1-\alpha^K, 0 \le \alpha \le 1$ の上にある。この曲線について次のことがいえる。

- 点 $(0,1)$ と $(1,0)$ を通る。
- $\dfrac{\mathrm{d}\beta}{\mathrm{d}\alpha} = -K\alpha^{K-1} < 0$ であるから、α が増加すると β は減少する。つまり、一方の誤りを起こりにくくすると、もう一方の誤りが起こりやすくなる。
- $K - 1 < 0$ であるから、$\alpha \to 0^+$ のとき傾きは $-\infty$ に発散し（曲線は $(0,1)$ から垂直に出発する）、$\alpha = 1$ では傾きは -1 と 0 の間にある $-K$ に等しい。
- $\dfrac{\mathrm{d}^2\beta}{\mathrm{d}\alpha^2} = K(1-K)\alpha^{K-2} > 0$ であるから、曲線は下に凸（convex）である。
- $0 < \alpha < 1$ かつ $0 < K < 1$ のとき $\alpha^K > \alpha$ であるから、$\beta < 1-\alpha$ である。すなわち、曲線は $(0,1)$ と $(1,0)$ を結ぶ線分より下にある。



曲線は $A = \frac{1}{6}, B = \frac{1}{3}$（したがって $K \approx 0.45$）の場合について描いたものであり、$M = 1, \ldots, 6$ に対応する点を記してある。M を増やすと、点は左上へ移動する。

検算

- $A = \frac{1}{6}, B = \frac{1}{3}, M = 4$ とする。直接計算すると $\alpha = (\frac{2}{3})^4 \approx 0.1975$、$\beta = 1 - (\frac{5}{6})^4 \approx 0.5177$ である。$K = \ln \frac{5}{6} / \ln \frac{2}{3} \approx 0.4497$ とすると、$1 - \alpha^K \approx 0.5177$ ともなり一致する。この検定を 2×10^5 回実行したシミュレーションでは、0.198 と 0.518 が得られた。
- 極端な場合も妥当な振る舞いをする。A が B に近ければ $K \approx 1$ かつ $\beta \approx 1 - \alpha$ となり、この検定は投げた結果を無視してでたらめに決めるのとほとんど変わらない。A が非常に小さければ $K \approx 0$ となり、どんな妥当な α に対しても β を小さくできる。

よくある間違い

- どちらの仮説がどちらかを取り違えること。ここでは H_0 が**大きいほうの**値 B であるから、6 が出るまで長く待つことは H_0 に反する証拠となる。
- K の分子と分母を逆にして $\ln(1 - B) / \ln(1 - A)$ としてしまうこと。これは 1 より大きい。$\beta < 1 - \alpha$（まともな検定ならあてずっぽうより優れている）という確認をすれば、この誤りに気づける。

STEP のその先へ

α に対する β の曲線は、2 種類の誤りの間のトレードオフ曲線であり、統計学や機械学習の至るところで用いられる ROC 曲線（ROC curve）と密接に関係している。ネイマン–ピアソンの補題（Neyman–Pearson lemma）によれば、今回のような 2 つの単純仮説に対しては、尤度比（likelihood ratio）に基づく検定が、各 α に対して可能な限り最小の β を与える。ここでは尤度比が X について単調であるから、「X が大きいときに H_0 を棄却する」という検定は実際に最適である。これらの考え方についてさらに学ぶには、基礎として Grimmett and Welsh, *Probability: An Introduction*（確率論入門）を読み、その後、数理統計学の入門的な講義や教科書に進むとよい。

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。