解答例

STEP III 2022、第 11 問

確率・統計 ・ 二項分布の平均偏差

STEP III 2022、問 11

先に進む前に、まず問題を読んでおこう。問題は、OCR が配布している 2022 年の STEP 過去問一式（zip ファイル）のうち、STEP III の問題用紙にある：

www.ocr.org.uk/Images/696587-2022-past-papers.zip

鍵となる考え方

硬貨が公平であるため、X の分布は平均（mean）に関して対称（symmetric）になり、平均より下の確率と上の確率は一致する。この対称性によって、絶対値（absolute value）を含む和が片側だけの和に変わる。さらに (i)(c) の恒等式（identity）によって $\sum r\binom{2n}{r}$ がより小さな二項分布（binomial distribution）の裾の和（tail sum）に変わり、これも再び対称性で求めることができる。

解答

以下では常に $X \sim \mathrm{B}(N, \frac{1}{2})$ とする。したがって $\mathrm{P}(X = r) = \binom{N}{r}2^{-N}$ であり、$\mu = \frac{1}{2}N$ である。$\binom{N}{r} = \binom{N}{N-r}$ であるから $\mathrm{P}(X = r) = \mathrm{P}(X = N - r)$ が成り立つ。すなわち、分布は $\frac{1}{2}N$ に関して対称である。

小問 (i)(a)

$N = 2n$ のとき、写像 $r \mapsto 2n - r$ は $\{0, \ldots, n-1\}$ を $\{n+1, \ldots, 2n\}$ に移す。よって対称性より

$$\mathrm{P}(X \leq n-1) = \mathrm{P}(X \geq n+1).$$

3 つの事象 $X \leq n-1$、$X = n$、$X \geq n+1$ はすべての結果を尽くすので、$2\,\mathrm{P}(X \leq n-1) + \mathrm{P}(X = n) = 1$ であり、これから

$$\mathrm{P}(X \leq n-1) = \tfrac{1}{2}(1 - \mathrm{P}(X = n)).$$

小問 (i)(b)

ここでは $\mu = n$ であるから

$$\delta = \sum_{r=0}^{2n} |r-n| \binom{2n}{r} \frac{1}{2^{2n}}.$$

$r = n$ の項は 0 である。$r \geq n+1$ の項で $r = 2n - s$ と置き換えると、$0 \leq s \leq n-1$ で $|r-n| = n-s$、$\binom{2n}{r} = \binom{2n}{s}$ となる。したがって、平均より上の項は平均より下の項をちょうど繰り返す：

$$\delta = 2\sum_{r=0}^{n-1} (n-r)\binom{2n}{r}\frac{1}{2^{2n}} = \sum_{r=0}^{n-1}(n-r)\binom{2n}{r}\frac{1}{2^{2n-1}},$$

となり、示すべき式が得られた。

小問 (i)(c)

$1 \leq r \leq 2n$ に対して

$$r\binom{2n}{r} = \frac{r\,(2n)!}{r!\,(2n-r)!} = 2n \cdot \frac{(2n-1)!}{(r-1)!\,(2n-r)!} = 2n\binom{2n-1}{r-1}.$$

ここで (b) の和を $2^{2n-1}\delta = A - B$ と分ける。ただし

$$A = n\sum_{r=0}^{n-1}\binom{2n}{r}, \qquad B = \sum_{r=1}^{n-1} r\binom{2n}{r}$$

である（B の $r=0$ の項は 0 なので省いた）。

和 A。(a) より $\sum_{r=0}^{n-1}\binom{2n}{r} = 2^{2n}\,\mathrm{P}(X \leq n-1) = 2^{2n-1} - \frac{1}{2}\binom{2n}{n}$ であるから

$$A = n\,2^{2n-1} - \tfrac{n}{2}\binom{2n}{n}.$$

? **考えてみよう** s が 0 から $n-2$ までのとき、$\binom{2n-1}{s}$ の和はどのように求めればよいか。

和 B。恒等式を用いて $s = r-1$ とおくと

$$B = 2n\sum_{s=0}^{n-2}\binom{2n-1}{s}.$$

$0 \leq s \leq 2n-1$ にわたる $\binom{2n-1}{s}$ の列は項の個数が偶数であり、総和は 2^{2n-1} である。この列は $s \mapsto 2n-1-s$ に関して対称であり、この写像は $\{0,\ldots,n-1\}$ と $\{n,\ldots,2n-1\}$

を入れ替えるので、前半の和は 2^{2n-2} となる。いま考えている和はその前半より 1 項少ないところで止まるので

$$\sum_{s=0}^{n-2}\binom{2n-1}{s} = 2^{2n-2} - \binom{2n-1}{n-1} = 2^{2n-2} - \tfrac{1}{2}\binom{2n}{n},$$

である。ここで $\binom{2n}{n} = \binom{2n-1}{n-1} + \binom{2n-1}{n} = 2\binom{2n-1}{n-1}$ を用いた。よって

$$B = n\,2^{2n-1} - n\binom{2n}{n}.$$

まとめる。

$$2^{2n-1}\delta = A - B = \left(n - \frac{n}{2}\right)\binom{2n}{n} = \frac{n}{2}\binom{2n}{n}, \qquad \text{よって} \qquad \delta = \frac{n}{2^{2n}}\binom{2n}{n}.$$

小問 (ii)

今度は $N = 2n+1$、$\mu = n + \frac{1}{2}$ であり、これは X のとりうる値ではない。対称性 $r \mapsto 2n+1-r$ は $\{0,\dots,n\}$ と $\{n+1,\dots,2n+1\}$ を入れ替え、$|r-\mu|$ と $\binom{2n+1}{r}$ をともに保つ。したがって (i)(b) と同様に

$$\delta = 2\sum_{r=0}^{n}\left(n+\tfrac{1}{2}-r\right)\binom{2n+1}{r}\frac{1}{2^{2n+1}} = \frac{1}{2^{2n+1}}\sum_{r=0}^{n}(2n+1-2r)\binom{2n+1}{r}.$$

- 同じ対称性により、$\binom{2n+1}{r}$ の列の前半の和は 2^{2n+1} の半分である：$\sum_{r=0}^{n}\binom{2n+1}{r} = 2^{2n}$。
- 恒等式 $r\binom{2n+1}{r} = (2n+1)\binom{2n}{r-1}$（(i)(c) とまったく同様に証明できる）と (i)(a) により

$$\sum_{r=1}^{n} r\binom{2n+1}{r} = (2n+1)\sum_{s=0}^{n-1}\binom{2n}{s} = (2n+1)\left(2^{2n-1} - \tfrac{1}{2}\binom{2n}{n}\right).$$

したがって

$$\sum_{r=0}^{n}(2n+1-2r)\binom{2n+1}{r} = (2n+1)2^{2n} - (2n+1)\left(2^{2n} - \binom{2n}{n}\right) = (2n+1)\binom{2n}{n},$$

であり、$N = 2n+1$ のとき

$$\boxed{\delta = \frac{2n+1}{2^{2n+1}}\binom{2n}{n}}$$

となる。

検算

- $N=3$（$n=1$）：$|X-\frac{3}{2}|$ は確率 $\frac{1}{4}$ で $\frac{3}{2}$、確率 $\frac{3}{4}$ で $\frac{1}{2}$ となるので、$\delta=\frac{3}{8}+\frac{3}{8}=\frac{3}{4}$ である。公式は $\frac{3\cdot 2}{8}=\frac{3}{4}$ を与える。$n=0$（$N=1$）のときは $\frac{1}{2}$ を与え、これも正しい。
- $N=4$（$n=2$）：(i)(c) は $\frac{2\cdot 6}{16}=\frac{3}{4}$ を与え、$N=3$ のときと同じになる。実は $N=2n+1$ のときの δ は常に $N=2n+2$ のときの δ に等しい。$\binom{2n+2}{n+1}=\frac{2(2n+1)}{n+1}\binom{2n}{n}$ であるから、偶数の場合の公式で n を $n+1$ に置き換えると奇数の場合の公式に帰着する。
- どちらの公式も、$N\le 60$ のすべての N について δ の定義と完全に一致することを計算機で確かめた。たとえば $N=10$ のとき $\delta=\frac{315}{256}\approx 1.230$ である。

よくある間違い

- 対称性から生じる因数 2 を忘れること。(i)(b) で 2^{2n} ではなく 2^{2n-1} が現れるのはこの因数のためである。
- $s=r-1$ とずらした後の s の範囲を誤ること。和は $n-2$ で止まり、列の半分より 1 項少ない。その欠けた項 $\binom{2n-1}{n-1}$ こそが答えを 0 でなくしている。
- (ii) で μ を整数として扱うこと。消える項はなく、列の前半と後半はどちらも $n+1$ 項からなる。

STEP のその先へ

スターリングの公式（Stirling's formula）により $\binom{2n}{n}\approx 4^n/\sqrt{\pi n}$ であるから、N が大きいとき $\delta\approx\sqrt{N/(2\pi)}$ となる。$\sigma=\frac{1}{2}\sqrt{N}$ であるから、これは $\delta\approx\sigma\sqrt{2/\pi}$ であり、正規分布（normal distribution）の平均絶対偏差（mean absolute deviation）にちょうど一致する。これは中心極限定理（central limit theorem）が予言するとおりである。Grimmett と Welsh の *Probability: An Introduction*（確率論入門）は、二項分布の正規近似（normal approximation）とその背後にある中心極限定理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。