

提出締切：2016 年 1 月 26 日 講義終了時

復習問題 12.1 表の出る確率 p が分からない硬貨がある。この確率 p を推定するために、以下のアルゴリズムを実行する。まず、この硬貨を n 回投げる。任意の $i \in \{1, \dots, n\}$ に対して、確率変数 X_i を

$$X_i = \begin{cases} 0 & (i \text{ 回目に投げたとき裏が出る}) \\ 1 & (i \text{ 回目に投げたとき表が出る}) \end{cases}$$

として定義し、 $X = X_1 + \dots + X_n$ とする。そして、 $\frac{X}{n}$ を p の推定値として出力する。以下の問いに答えよ。

1. 任意の $i \in \{1, \dots, n\}$ に対して $E[X_i]$ は何か？
2. $E[\frac{X}{n}] = p$ であることを証明せよ。
3. 任意の $i \in \{1, \dots, n\}$ に対して $E[X_i^2]$ は何か？
4. 任意の異なる $i, j \in \{1, \dots, n\}$ に対して $E[X_i X_j]$ は何か？
5. $E[(\frac{X}{n} - p)^2]$ は何か？
6. 任意の正実数 $\varepsilon > 0$ に対して、次の不等式

$$\Pr\left(\left|\frac{X}{n} - p\right| \geq \varepsilon\right) \leq \frac{1}{\varepsilon^2} \frac{p(1-p)}{n}$$

が成り立つことを証明せよ。

7. このアルゴリズムによって、確率 p を推定しようとしたとき、誤差 $|\frac{X}{n} - p|$ が ε 以上になる確率を δ 以下にするには、 n をどれほど大きくすれば十分か？ ただし、 $\delta > 0$ は正実数である。

復習問題 12.2 問題 12.1 と同じ設定の問題に対して、次のようなアルゴリズムを考える。投げる回数 n は自然数 k, t を用いて $n = (2k - 1)t$ と書けるものとする。そして、硬貨を n 回投げる。確率変数 X_i は問題 12.1 と同じように定義する。そして、任意の $j \in \{1, \dots, 2k - 1\}$ に対して、確率変数 Y_j を

$$Y_j = \frac{X_{(j-1)t+1} + \dots + X_{(j-1)t+t}}{t}$$

として定義する。出力は $Y = \text{med}\{Y_1, \dots, Y_{2k-1}\}$ 、すなわち、 $\{Y_1, \dots, Y_{2k-1}\}$ の中央値である。

以下の問いに答えよ。(問題 12.1 の結果を用いてもよい。)

1. 任意の $j \in \{1, \dots, 2k - 1\}$ に対して

$$\Pr(|Y_j - p| \geq \varepsilon) \leq \frac{1}{8}$$

が成り立つためには、 t をどれほど大きくすれば十分か？

2. この小問以降、 t は小問 1 の不等式を満たすように選ばれているとする。そのとき、次の不等式

$$\Pr(k \text{ 個の } j \text{ に対して, } |Y_j - p| \geq \varepsilon) < \left(\frac{3}{4}\right)^k$$

を証明せよ。

3. 不等式 $\Pr(|Y - p| \geq \varepsilon) < \left(\frac{3}{4}\right)^k$ を証明せよ。(問題 12.3 の結果を用いてもよい。)
4. 不等式 $\Pr(|Y - p| \geq \varepsilon) < \delta$ が成り立つためには、 k をどれほど大きくすれば十分か？
5. 以上を踏まえて、推定値 Y が p から ε 以上離れる確率を δ 以下にするには、 n をどれだけ大きくすれば十分か、答えよ。

補足問題 12.3 k を自然数として、 $2k - 1$ 個の実数 $x_1, x_2, \dots, x_{2k-1}$ を考える。これらは $x_1 \leq x_2 \leq \dots \leq x_{2k-1}$ という大小関係を満たしているとする。つまり、 $\text{med}\{x_1, \dots, x_{2k-1}\} = x_k$ である。任意の閉区間 $[a, b] \subseteq \mathbb{R}$ を考える。このとき、 $\{x_1, \dots, x_{2k-1}\}$ の中で、 $[a, b]$ の要素でないものの個数が k 未満であるならば、 $x_k \in [a, b]$ が成り立つことを証明せよ。また、この逆が成り立たないことを証明せよ。(反例を挙げよ。)

追加問題 12.4 自然数 $n \geq 2$ に対して、互いに独立な確率変数 X_1, \dots, X_n は、ある実数 p を用いて、次のように定められるとする(ただし、 $0 \leq p \leq 1$)。

$$\Pr(X_i = 0) = 1 - p, \quad \Pr(X_i = 1) = p.$$

このとき、 $E[X_i] = p$ である。しかし、確率変数 $X = \text{med}\{X_1, \dots, X_n\}$ に対して、 $E[X] = p$ が成り立つとは限らない。これが成り立たないような p と n の値の組を 1 つ発見せよ。(なぜ成り立たないのかも説明せよ。)

追加問題 12.5 互いに独立な確率変数 X_1, \dots, X_n は、ある実数 p を用いて、次のように定められるとする(ただし、 $0 \leq p \leq 1$)。

$$\Pr(X_i = 0) = 1 - p, \quad \Pr(X_i = 1) = p.$$

このとき、 X_1, \dots, X_n の最小値の期待値

$$E[\min\{X_1, \dots, X_n\}]$$

は何か？ 定めよ。(ヒント：直接計算せよ。)