6　離散型確率変数

導入

これまで一般の空間でも確率を定義するために色々な準備を行ってきた．特に抽象的な議論から出発したので，何をしているのかを見失ってしまった方もいるかもしれない．今回からはまた具体的な対象を扱う，これまで通りの数学に戻る．これらの具体例を通して，前回までで展開してきた議論を再探訪してみてほしい[1)]．

6.1　確率変数の復習

確率変数 X とは，ある確率空間 $(\Omega, \mathcal{F}, P)$ から $(\mathbb{R}, \mathcal{B}(\mathbb{R}))$ への可測関数 $X\colon \Omega \to \mathbb{R}$ のことである．これは，一般の空間上の確率を，$\mathbb{R}$ 上の確率と思うことに対応する．有限集合上のものだと，以下のような確率分布表を用いて表せる．

X の値	x_1	x_2	$\cdots$	x_n	計
確率	p_1	p_2	$\cdots$	p_n	1

確率変数 X に付随する分布関数 F_X の値域が離散集合のとき，確率変数 X は**離散型**であるという．離散型の確率分布の具体例をいくつか見ていこう．

6.2　離散型一様分布

簡単のため，確率変数の取りうる値は $\{1, 2, \ldots, n\}$ であるとする．

定義 6.1

n 個の起こりうる事象の確率がすべて等しい分布，すなわち

$$P(X = k) = \frac{1}{n} \quad (k = 1, 2, \ldots, n)$$

となる分布を**一様分布**という．

注意. ここでの定義のように，元の空間 Ω については具体的に与えられないことも多い．これは例えばサイコロを振った目を読み取るという試行と 1 から 6 が書かれたカードから一枚取り出す試行はどちらも同じ分布になるように，確率変数という写像を通すことで一見異なった事象についても同じ枠組みで考察することが可能となる．

一様分布の例としては，サイコロを振った時の出目やトランプの山からカードを一枚引くといった試行が挙げられる．

6.3　二項分布

ある試行の結果起こる事象を A とし，その起こる確率は $p > 0$ であるとする．この試行を独立に n 回繰り返すとき，n 回のうち A の起こった回数を X とすると X は確率変数である．この X の取りうる値は $0, 1, \ldots, n$ で，その確率は

$$P(X = k) = \binom{n}{k} p^k (1-p)^{n-k} \quad (k = 0, 1, \ldots, n)$$

となる．

定義 6.2

$p > 0$ とする．$k = 0, 1, 2, \ldots, n$ の値をとる確率変数 X が

$$P(X = k) = \binom{m}{\ell} p^k (1-p)^{n-k}$$

を満たすとき，その分布をパラメータ p を持つ**二項分布**といい，$B(n, p)$ と表す．

確率変数 X の分布が二項分布 $B(n, p)$ であるとき，**確率変数 X は二項分布 $B(n, p)$ にしたがう**ということも多い．またさらに略して $X \sim B(n, p)$ などと書くこともある．

分布である条件は次の式変形で確認できる．

$$\begin{aligned} \sum_{k=0}^{n} P(X = k) &= \sum_{k=0}^{n} \binom{n}{k} p^k (1-p)^{n-k} \\ &= (p + 1 - p)^n = 1 \end{aligned}$$

例題 6.3

表が出る確率が p のコインを計 4 回投げる試行を考える．この 4 回のうち表が出る回数を X とすると，これは確率変数になる．以下の問いに答えよ．

(1)　確率分布表を作成せよ．

(2)　表が出る回数が偶数回になる確率を求めよ．

---

1)　3 回も講義時間をかけたわりに，講義内で扱う確率測度はどれも基本的なものばかりである．「大山鳴動して鼠一匹」のようにも思えるが，実はその鼠は金の鼠なのである．なんとなく扱っているものにしっかりとした理論的な裏付けを与えることが，数学の重要な役割である．

6.3.1　二項分布の応用

例題 6.4

ある製品の生産ラインの不良率は 0.02 である．このラインで生産された製品からランダムに 100 個を取り出すことを考える．このとき不良品の個数が 4 個以上ある確率はいくらになるだろうか．

証明． この試行は，確率が $p = 0.02$ である試行を，ランダムに $n = 100$ 回繰り返すものであるから，100 個のうちの不良品の個数を X と置けば，X は二項分布 $B(n,p) = B(100, 0.02)$ に従うことになる．ここで

$$P(X = k) = \binom{100}{k}(0.02)^k(1-0.02)^{100-k}$$

で特に

$$\begin{aligned} P(X = 0) &\fallingdotseq 0.1326195559 \\ P(X = 1) &\fallingdotseq 0.2706521548 \\ P(X = 2) &\fallingdotseq 0.2734139115 \\ P(X = 3) &\fallingdotseq 0.1822759411 \end{aligned}$$

なので

$$\begin{aligned} P(X \geqq 4) &= 1 - P(X < 4) \\ &= 1 - \sum_{k=0}^{3} P(X = k) \\ &\fallingdotseq 1 - 0.8589615633 \\ &\fallingdotseq 0.1410384367 \end{aligned}$$

となり，約 14.1% ということになる．　□

6.4　ポアソン分布

「まれにしか起こらないことが単位時間あたりに何件起こったか」という確率に当てはまるものがポアソン分布である．たとえば，一定時間内に観測される回数の分布（駅の改札の通過人数，病院における外来患者数など）はポアソン分布にしたがうことが知られている．

定義 6.5

$\lambda > 0$ として，$k = 0, 1, 2, \ldots$ の値をとる確率変数 X が

$$P(X = k) = e^{-\lambda}\frac{\lambda^k}{k!} \quad (k = 0, 1, 2, \ldots)$$

を満たすとき，その分布をパラメータ λ を持つ**ポアソン分布**と呼び，$\mathrm{Po}(\lambda)$ と表す．

ポアソン分布が分布の条件を満たしていることは，指数関数のテイラー展開を利用して

$$\begin{aligned} \sum_{k=0}^{\infty} P(X = k) &= \sum_{k=0}^{\infty} e^{-\lambda}\frac{\lambda^k}{k!} \\ &= e^{-\lambda}\sum_{k=0}^{\infty}\frac{\lambda^k}{k!} = e^{-\lambda}\cdot e^{\lambda} = 1 \end{aligned}$$

となることより確認できる．

例題 6.6

ある病院への外来患者数は 1 時間に 10 人である．外来患者数がポアソン分布（$\mathrm{Po}(10)$）にしたがうとすると，1 時間に 15 人以上の患者が来院する確率を求めよ．

解． 1 時間に来院する患者の人数の確率変数 X とすると，$X \sim \mathrm{Po}(10)$ である．よって

$$P(X = k) = \frac{e^{-10}10^k}{k!} \quad (k = 0, 1, 2, \ldots).$$

これより

$$\begin{aligned} P(X \geqq 15) &= 1 - P(X < 15) \\ &= 1 - \sum_{k=0}^{14} P(X = k) \fallingdotseq 1 - 0.917 \\ &= 0.083 \end{aligned}$$

となり，約 8.3% である．　□

演習問題 6

(1) 1 枚の硬貨を 6 回投げるとき，表が 4 回出る確率を求めよ．

(2) n を自然数とする．1 枚の硬貨を $2n$ 回投げるとき，表が n 回出る確率を $p(n)$ とする．次の問いに答えよ．

(a) $p(1)$, $p(2)$, $p(3)$ を求めよ．

(b) $\frac{p(n)}{p(n+1)}$ を n を用いて表せ．

(c) $p(n) > p(n+1)$ が成り立つことを示せ．

(3) ある家に 1 日にかかってくる電話の件数を X をする．X はポアソン分布 Po(1.5) に従うものとして，次の確率を求めよ．ただし値は小数第五位を四捨五入せよ．

(a) 1 日にかかってくる電話の件数が 0 である確率

(b) 1 日にかかってくる電話の件数が 1 件である確率

(c) 1 日にかかってくる電話の件数が 3 件以上である確率

(4) 二項分布を用いて次の確率の近似値を求めよ．値は小数第五位を四捨五入せよ．

(a) ある機械から生産される製品には 0.2% の割合で不良品がある．この製品を箱に 100 個詰めるとき，この箱の中に不良品が 1 個以上入る確率．

(b) ある予防注射によって副作用を起こす確率は 0.1% であるとする．800 人の人にこの予防注射をするとき，2 人以上が副作用を起こす確率．

(5) n が十分大きいとき，二項分布はポアソン分布で近似できることが知られている．すなわち二項分布 $B(n,p)$ において $\lambda = np$ としたとき Po(λ) で近似できる．上の演習問題でそのことを確認せよ．