確率論 II

石綿 元

第３回講義

目次

1　確率変数と確率分布 I

1.0　((復習項目)) 微分積分学からの準備 1

1.0.1　順列と組み合わせ

■階乗関数　階乗を次で定義する。

$$n! = n \cdot (n-1) \cdot (n-2) \cdots 2 \cdot 1$$

ゼロの階乗を次で定義する。

$$0! = 1$$

■２項係数　２項係数を次で定義する。

$${}_nC_r = \binom{n}{r} = \frac{n \cdot (n-1) \cdot (n-2) \cdots (n-r+1)}{r!}$$

また、

$${}_nC_0 = 1, \quad {}_0C_0 = 1, \quad {}_nC_n = 1$$

■順列　n 個のものから r 個を取って、順番に 1 列に並べたもの。

$${}_nP_r = \frac{n!}{(n-r)!}$$

■重複順列　繰り返しを許して n 個のものから r 個を取る順列。

$$n^r = n \cdot n \cdot n \cdots n$$

■組み合わせ　n 個のものから順序を考えずに r 個を取り分ける。

$${}_nC_r = \frac{n \cdot (n-1) \cdot (n-2) \cdots (n-r+1)}{r!}$$

■重複組み合わせ　繰り返しを許して n 個のものから順序を考えずに r 個を取る組み合わせ。

$${}_nH_r = \frac{n \cdot (n+1) \cdot (n+2) \cdots (n+r-1)}{r!}$$

1.0.2　２項定理

２項の冪乗を考えると、組み合わせの記号で表現でき、

$$(\alpha+\beta)^n = {}_nC_0\alpha^n\beta^0 + {}_nC_1\alpha^{n-1}\beta^1 + {}_nC_2\alpha^{n-2}\beta^2 + \cdots + {}_nC_x\alpha^{n-x}\beta^x + \cdots + {}_nC_{n-1}\alpha^1\beta^{n-1} + {}_nC_n\alpha^0\beta^n$$

が成立している。これを 2 項定理と呼び、各係数に着目すると、パスカルの三角形を作っている。

1.0.3　組み合わせの記号の性質

■その１

$$ {}_{k+1}C_r = {}_kC_r + {}_kC_{r-1} $$

Proof) 左辺：

$$ \begin{aligned} {}_{k+1}C_r &= \frac{(k+1)\cdot k\cdot(k-1)\cdot(k-2)\cdots(k+1-r+1)}{r!} \\ &= \frac{(k+1)\cdot k\cdot(k-1)\cdot(k-2)\cdots(k-r+2)}{r!} \end{aligned} $$

右辺：

$$ \begin{aligned} {}_kC_r + {}_kC_{r-1} &= \frac{k\cdot(k-1)\cdot(k-2)\cdots(k-r+1)}{r!} + \frac{k\cdot(k-1)\cdot(k-2)\cdots(k-(r-1)+1)}{(r-1)!} \\ &= \frac{k\cdot(k-1)\cdot(k-2)\cdots(k-r+1)+r\cdot k\cdot(k-1)\cdot(k-2)\cdots(k-r+2)}{r!} \\ &= \frac{k\cdot(k-1)\cdot(k-2)\cdots(k-r+2)\{(k-r+1)+r\}}{r!} \\ &= \frac{k\cdot(k-1)\cdot(k-2)\cdots(k-r+2)\{k+1\}}{r!} \\ &= \frac{\{k+1\}\cdot k\cdot(k-1)\cdot(k-2)\cdots(k-r+2)}{r!} = {}_{k+1}C_r \end{aligned} $$

■その２

$$ \sum_{x=0}^{n}({}_MC_x)\cdot({}_{N-M}C_{n-x}) = {}_NC_n $$

Proof) まず、

$$ (1+\beta)^M\cdot(1+\beta)^{N-M} = (1+\beta)^N $$

であることは簡単である。これを用いて、２項定理における展開式の係数となっていることに着目すると

$$ \begin{aligned} (1+\beta)^M &= {}_MC_0\beta^0 + {}_MC_1\beta^1 + {}_MC_2\beta^2 + \cdots + {}_MC_x\beta^x + \cdots + {}_MC_{M-1}\beta^{M-1} + {}_MC_M\beta^M \\ (1+\beta)^{N-M} &= {}_{N-M}C_0\beta^0 + {}_{N-M}C_1\beta^1 + {}_{N-M}C_2\beta^2 + \cdots + {}_{N-M}C_x\beta^x + \cdots + \\ &\quad {}_{N-M}C_{N-M-1}\beta^{N-M-1} + {}_{N-M}C_{N-M}\beta^{N-M} \\ (1+\beta)^N &= {}_NC_0\beta^0 + {}_NC_1\beta^1 + {}_NC_2\beta^2 + \cdots + {}_NC_x\beta^x + \cdots + {}_NC_{N-1}\beta^{N-1} + {}_NC_N\beta^N \end{aligned} $$

よって、n を β^N の各指数 $0, 1, \cdots, N$ ごとに考えると

$$ \begin{aligned} ({}_NC_0)\,\beta^0 &= ({}_MC_0)\,\beta^0\cdot({}_{N-M}C_0)\,\beta^0 \\ ({}_NC_0) &= ({}_MC_0)\cdot({}_{N-M}C_0) \end{aligned} $$

また、

$$ \begin{aligned} ({}_NC_1)\,\beta^1 &= ({}_MC_0)\,\beta^0\cdot({}_{N-M}C_1)\,\beta^1 + ({}_MC_1)\,\beta^1\cdot({}_{N-M}C_0)\,\beta^0 \\ ({}_NC_1) &= ({}_MC_0)\cdot({}_{N-M}C_1) + ({}_MC_1)\cdot({}_{N-M}C_0) \end{aligned} $$

さらに、

$$
\begin{aligned}
({}_NC_2)\,\beta^2 &= ({}_MC_0)\,\beta^0 \cdot ({}_{N-M}C_2)\,\beta^2 + ({}_MC_1)\,\beta^1 \cdot ({}_{N-M}C_1)\,\beta^1 + ({}_MC_2)\,\beta^2 \cdot ({}_{N-M}C_0)\,\beta^0 \\
({}_NC_2) &= ({}_MC_0) \cdot ({}_{N-M}C_2) + ({}_MC_1) \cdot ({}_{N-M}C_1) + ({}_MC_2) \cdot ({}_{N-M}C_0)
\end{aligned}
$$

つまり、

$$\sum_{x=0}^{n} ({}_MC_x) \cdot ({}_{N-M}C_{n-x}) = {}_NC_n$$

1.0.4 家計簿の原理

日付		食費	交通費	交際費	雑費	支出計
		x_{i1}	x_{i2}	x_{i3}	x_{i4}	$\sum_{j=1}^{4} x_{ij}$
4月1日	x_{1j}	$x_{11} = 3250$	$x_{12} = 420$	$x_{13} = 500$	$x_{14} = 800$	$\sum_{j=1}^{4} x_{1j} = 4970$
4月2日	x_{2j}	$x_{21} = 2840$	$x_{22} = 420$	$x_{23} = 18000$	$x_{24} = 280$	$\sum_{j=1}^{4} x_{2j} = 21540$
4月3日	x_{3j}	$x_{31} = 3100$	$x_{32} = 420$	$x_{33} = 1000$	$x_{34} = 300$	$\sum_{j=1}^{4} x_{3j} = 4820$
4月4日	x_{4j}	$x_{41} = 2930$	$x_{42} = 420$	$x_{43} = 3000$	$x_{44} = 250$	$\sum_{j=1}^{4} x_{4j} = 6600$
4月5日	x_{5j}	$x_{51} = 3330$	$x_{52} = 420$	$x_{53} = 200$	$x_{54} = 680$	$\sum_{j=1}^{4} x_{5j} = 4630$
4月6日	x_{6j}	$x_{61} = 2720$	$x_{62} = 420$	$x_{63} = 4500$	$x_{64} = 900$	$\sum_{j=1}^{4} x_{6j} = 8540$
4月7日	x_{7j}	$x_{71} = 8000$	$x_{72} = 420$	$x_{73} = 13200$	$x_{74} = 180$	$\sum_{j=1}^{4} x_{7j} = 21800$
一週間計	$\sum_{i=1}^{7} x_{ij}$	$\sum_{i=1}^{7} x_{i1}$	$\sum_{i=1}^{7} x_{i2}$	$\sum_{i=1}^{7} x_{i3}$	$\sum_{i=1}^{7} x_{i4}$	$\sum_{i=1}^{7}\sum_{j=1}^{4} x_{ij} = 72900$
		$= 26170$	$= 2940$	$= 40400$	$= 3390$	$\sum_{j=1}^{4}\sum_{i=1}^{7} x_{ij} = 72900$

つまり、

$$\sum_{i=1}^{m}\sum_{j=1}^{n} x_{ij} = \sum_{j=1}^{n}\sum_{i=1}^{m} x_{ij}$$

1.0.5 極限の性質

$$\lim_{n\to\infty} \{f(x) \cdot g(x)\} = \left(\lim_{n\to\infty} f(x)\right) \cdot \left(\lim_{n\to\infty} g(x)\right)$$

1.0.6 指数関数

■e の定義

$$e = \lim_{n\to\infty} \left(1 + \frac{1}{n}\right)^n$$

■e^x の定義

$$e^x = \lim_{n\to\infty} \left(1 + \frac{x}{n}\right)^n$$

1.0.7 Rolle の定理

関数 $y = f(x)$ が $[a, b]$ で連続、(a, b) で微分可能なとき、$f(a) = f(b)$ ならば、

$$f'(\xi) = 0$$

となる点 ξ $(a < \xi < b)$ が存在する。

1.0.8 平均値の定理

関数 $y = f(x)$ が $[a, b]$ で連続、(a, b) で微分可能ならば、

$$\frac{f(b) - f(a)}{b - a} = f'(\xi)$$

となる点 ξ $(a < \xi < b)$ が存在する。

1.0.9 Taylor の定理

関数 $y = f(x)$ が (a, b) で n 回微分可能ならば、

$$f(b) = f(a) + (b-a)\frac{f'(a)}{1!} + (b-a)^2\frac{f''(a)}{2!} + \cdots + (b-a)^{n-1}\frac{f^{(n-1)}(a)}{(n-1)!} + (b-a)^n\frac{f^{(n)}(\xi)}{n!}$$

を満たす ξ $(a < \xi < b)$ が存在する。

■剰余項 最終項について、

$$R_n = (b-a)^n\frac{f^{(n)}(\xi)}{n!}$$

とおき、これをラグランジュの剰余項と呼ぶ。

また、一般化した形として求めておくと、

$$R_n = (b-a)^p(b-\xi)^{n-p}\frac{f^{(n)}(\xi)}{(n-1)! \cdot p}$$

この形の剰余項をロッシュ・シュレミルヒの剰余項と呼ぶ。$p = n$ とすると、ラグランジュの剰余項となる。どんな形をしていても当然、同じ R_n であり同一の値である。

1.0.10 Cauchy の平均値の定理

２つの関数 $f(x),\ g(x)$ がともに (a, b) で微分可能かつ $g'(x) \neq 0$ ならば、

$$\frac{f(b) - f(a)}{g(b) - g(a)} = \frac{f'(\xi)}{g'(\xi)}$$

をみたす ξ $(a < \xi < b)$ が存在する。

1.0.11 L'Hopital の定理

$$\lim_{x\to a}\frac{f(x)}{g(x)} = \frac{0}{0}\text{もしくは}\lim_{x\to a}\frac{f(x)}{g(x)} = \frac{\infty}{\infty}\text{このとき、次が成り立つ。}\lim_{x\to a}\frac{f(x)}{g(x)} = \lim_{x\to a}\frac{f'(x)}{g'(x)}$$

1.0.12 Taylor 級数展開

■冪級数 a を定数 β_n を定数 $(n = 0, 1, 2, \cdots)$ とするとき、関数項級数

$$\beta_0 + (x-a)\beta_1 + (x-a)^2\beta_2 + (x-a)^3\beta_3 + \cdots + (x-a)^n\beta_n + \cdots = \sum_{n=0}^{\infty}(x-a)^n\beta_n$$

を a を中心とするべき級数と定義する。

■冪級数展開　$y=f(x)$ が無限回微分可能で、点 a を中心とするべき級数で表せたとすると、

$$f(x)=\beta_0+(x-a)\beta_1+(x-a)^2\beta_2+\cdots+(x-a)^n\beta_n+\cdots$$

このとき、$x=a$ とすると β_0 が求まる。次に、$f(x)$ を一回微分して $x=a$ を代入すると、β_1 が求まる。これを繰り返すことで、

$$\begin{aligned}f(x)&=\beta_0+(x-a)\beta_1+(x-a)^2\beta_2+\cdots+(x-a)^n\beta_n+\cdots\\&=f(a)+(x-a)f'(a)+(x-a)^2\frac{f''(a)}{2!}+\cdots+(x-a)^n\frac{f^{(n)}(a)}{n!}+\cdots\end{aligned}$$

これを点 a を中心とする冪級数展開と定義する。

■テイラー展開　テイラーの定理

$$f(b)=f(a)+(b-a)\frac{f'(a)}{1!}+(b-a)^2\frac{f''(a)}{2!}+\cdots+(b-a)^{n-1}\frac{f^{(n-1)}(a)}{(n-1)!}+(b-a)^n\frac{f^{(n)}(\xi)}{n!}$$

において、$f(x)$ が無限回微分可能で、

$$\lim_{n\to\infty}R_n=0$$

となるなら、

$$f(x)=f(a)+(x-a)f'(a)+(x-a)^2\frac{f''(a)}{2!}+\cdots+(x-a)^n\frac{f^{(n)}(a)}{n!}+\cdots$$

これは、点 a 中心の冪級数展開そのものである。この冪級数展開を $f(x)$ の $x=a$ におけるテイラー展開といい、a を $f(x)$ のテイラー展開における中心という。

■マクローリン展開　中心が 0 のときのテイラー展開を特にマクローリン展開という。

$$f(x)=f(0)+x\frac{f'(0)}{1!}+x^2\frac{f''(0)}{2!}+\cdots+x^n\frac{f^{(n)}(0)}{n!}+\cdots$$

■代表的なテイラー級数展開の公式

$$\begin{aligned}\sin x&=x-x^3\frac{1}{3!}+x^5\frac{1}{5!}-x^7\frac{1}{7!}+\cdots=\sum_{n=0}^{\infty}(-1)^n\frac{x^{2n+1}}{(2n+1)!}=\sum_{n=1}^{\infty}(-1)^{n+1}\frac{x^{2n-1}}{(2n-1)!}\\\cos x&=1-x^2\frac{1}{2!}+x^4\frac{1}{4!}-x^6\frac{1}{6!}+\cdots=\sum_{n=0}^{\infty}(-1)^n\frac{x^{2n}}{(2n)!}=\sum_{n=1}^{\infty}(-1)^{n+1}\frac{x^{2n-2}}{(2n-2)!}\\e^x&=1+x+x^2\frac{1}{2!}+x^3\frac{1}{3!}+x^4\frac{1}{4!}+\cdots=\sum_{n=0}^{\infty}x^n\frac{1}{n!}=\sum_{n=1}^{\infty}\frac{x^{n-1}}{(n-1)!}=\sum_{n=2}^{\infty}\frac{x^{n-2}}{(n-2)!}\\\log(1+x)&=x-\frac{x^2}{2}+\frac{x^3}{3}-\frac{x^4}{4}+\cdots=\sum_{n=1}^{\infty}(-1)^{n+1}\frac{x^n}{n}\quad(-1<x\leq 1)\\(1+x)^\alpha&=\binom{\alpha}{0}+x\binom{\alpha}{1}+x^2\binom{\alpha}{2}+\cdots=\sum_{n=0}^{\infty}x^n\binom{\alpha}{n}\quad(-1<x<1)\end{aligned}$$

1.1　確率変数の導入

1.1.1　確率変数の定義

標本空間 Ω における関数 $X:\Omega \to \boldsymbol{R}$ とする確率空間 $(\Omega,\ \mathscr{F},\ P)$ で定義された可測関数 X を確率変数と定義する。

1.1.2　標本空間と確率変数

標本空間 Ω は、試行結果である根元事象 ω_i の全体集合である。このとき、標本空間 Ω の各根元事象 ω_i それぞれがその事象特有の対応する確率で起こるとき、各根元事象 ω_i を変数と扱うために、適当な数値 x_i を対応させておく。すなわち、この対応させた数値である変数を確率変数と言うのであって、標本空間 Ω について、その根元事象 ω_i とすると、

$$\Omega = \{\omega_1, \omega_2, \omega_3, \cdots\}$$

この根元事象 ω_i に対して適当な数値を対応させるので、

$$\begin{aligned} X &= \{X(\omega_1), X(\omega_2), X(\omega_3), \cdots\} \\ &= \{x_1, x_2, x_3, \cdots\} \end{aligned}$$

つまり、根元事象 ω_i に数値 x_i を対応させるルールであり、標本空間・根元事象をどのように取るかによって、対応は変わる。

例）さいころを振って出る目の数の標本空間において確率変数をどう取るか。（出る目の数そのものに興味がある場合、出る目の偶奇に興味がある場合）

例）２枚のコインを投げるときの標本空間において確率変数をどう取るか。

例）2 個の区別できるさいころの出る目の標本空間において確率変数をどう取るか。

1.2　確率関数

■離散型　確率変数を導入すると、その変数についての関数が定義できる。

■離散型確率関数の定義　離散型確率変数 X について

$$\begin{aligned} f(x_i) &= P(X = x_i) \quad && (x_i = x_1, x_2, \cdots) \\ &= 0 && (x_i \neq x_1, x_2, \cdots) \end{aligned}$$

ここで、$P(X = x_i)$ とは、確率変数 X が数値 x_i を取る場合の確率である。このとき、関数 f を X の確率関数と定義する。当然、

$$0 \leq f(x_i) \leq 1, \qquad \sum_{i=1}^{n} f(x_i) = 1$$

を満たしている。確率関数で表される分布を確率分布という。

1.3　確率分布の性質

「第 1 回講義　**データの視覚化**」では、度数分布表とヒストグラムを用いて観測データを記述した。このときの、ヒストグラムは観測データを数え上げげた実際の分布であり、経験分布と呼ぶ。これに対して、確率分

布とは、数学的なモデルである理論分布である。「第 1 回講義　**データを代表値で読む**」における「位置の尺度」と「散らばりの尺度」について相対度数を確率に取り替えてまったく同様に考えることにする。

1.3.1　期待値・平均値

理論分布である確率分布について「平均」を考える。

「第 1 回講義　**データを代表値で読む**」における「2.1.1　**位置の尺度**」で定義された平均とは、ヒストグラムを作成したときの**重心**であって、度数分布表のみが与えられている場合、データの総数 n で、r 個の各階級に属する度数を m_i、各階級の階級値を x_i として

$$\bar{x} = \frac{\sum_{i=1}^{r} \mathrm{x}_i m_i}{n} = \frac{1}{n}\sum_{i=1}^{r} \mathrm{x}_i m_i = \sum_{i=1}^{r} \mathrm{x}_i \frac{m_i}{n}$$

であった。

ここで、記号の意味を揃えるため、ノーテーションの変更を行う。データの総数 m で、n 個の各階級に属する度数を u_i、各階級の階級値を x_i として

$$\bar{x} = \frac{\sum_{i=1}^{n} \mathrm{x}_i u_i}{m} = \frac{1}{m}\sum_{i=1}^{n} \mathrm{x}_i u_i = \sum_{i=1}^{n} \mathrm{x}_i \frac{u_i}{m}$$

と表現することにする。

今、考えている確率関数 $f(x_i)$ は $\sum_{i=1}^{n} f(x_i) = 1$ を条件としているわけだから、相対度数 $\frac{u_i}{m}$ を $f(x_i)$ と考えることができる。

つまり、確率関数 $f(x_i)$ においての平均、

$$E[X] = \mu = \sum_{i=1}^{n} x_i f(x_i)$$

を離散型確率変数 X の期待値と定義する。

1.3.2　分散

同様に確率分布の「分散」を考える。

「第 1 回講義　**データを代表値で読む**」における「2.2.1　**散らばりの尺度**」で定義された分散とは、度数分布表のみが与えられている場合、データの総数 n で、r 個の各階級に属する度数を m_i、各階級の階級値を x_i として

$$s^2 = \frac{1}{n}\sum_{i=1}^{r} (\mathrm{x}_i - \bar{x})^2 m_i = \sum_{i=1}^{r} (\mathrm{x}_i - \bar{x})^2 \frac{m_i}{n}$$

であった。

ここでも記号の意味を揃えるため、ノーテーションの変更を行う。データの総数 m で、n 個の各階級に属する度数を u_i、各階級の階級値を x_i として

$$s^2 = \frac{1}{m}\sum_{i=1}^{n} (\mathrm{x}_i - \bar{x})^2 u_i = \sum_{i=1}^{n} (\mathrm{x}_i - \bar{x})^2 \frac{u_i}{m}$$

と表現することにする。

今、考えている確率関数 $f(x_i)$ は $\sum_{i=1}^{n} f(x_i) = 1$ を条件としているわけだから、相対度数 $\frac{u_i}{m}$ を $f(x_i)$ と考えることができる。

つまり、確率関数 $f(x_i)$ において、

$$V[X] = \sigma^2 = \sum_{i=1}^{n} (x_i - \mu)^2 f(x_i)$$

を離散型確率変数 X の分散と定義する。

1.3.3　標準偏差

標準偏差についても同様に

$$\sigma = \sqrt{\sigma^2} = \sqrt{V[X]}$$

と定義する。

1.4　積率

1.4.1　積率 (モーメント)

■n 次積率　確率変数 X について偏差 $X - \mu$ の n 乗の期待値を

$$E[(X-\mu)^n] = \sum_{i=1}^{n} (x_i - \mu)^n f(x_i)$$

を確率変数 X の期待値周りの n 次モーメントまたは n 次積率と定義する。

■原点周りの n 次積率　$\mu = 0$ である場合を、確率変数 X についての n 乗の期待値を確率変数 X の原点周りの n 次モーメントまたは n 次積率と定義する。

$$E[X^n] = \sum_{i=1}^{n} x_i^n f(x_i)$$

■n 次標準化モーメント　確率変数 X の期待値周りの n 次モーメントを標準偏差 σ の n 乗である σ^n で割ったもの

$$\alpha_n = \frac{E[(X-\mu)^n]}{\sigma^n}$$

を n 次の標準化モーメントと定義する。

1.4.2　期待値・分散を積率で表現する

■原点周りの 1 次積率

$$E[X^1] = E[X] = \sum_{i=1}^{n} x_i f(x_i)$$

期待値そのものである。

■原点周りの 2 次積率

$$E[X^2] = \sum_{i=1}^{n} x_i^2 f(x_i)$$

■期待値と分散の関係

$$\begin{aligned}
V[X] = \sigma^2 &= \sum_{i=1}^{n} (x_i - \mu)^2 f(x_i) \\
&= \sum_{i=1}^{n} (x_i^2 - 2\mu x_i + \mu^2) f(x_i) \\
&= \sum_{i=1}^{n} (x_i^2 f(x_i) - 2\mu x_i f(x_i) + \mu^2 f(x_i)) \\
&= \sum_{i=1}^{n} x_i^2 f(x_i) - \sum_{i=1}^{n} 2\mu x_i f(x_i) + \sum_{i=1}^{n} \mu^2 f(x_i) \\
&= \sum_{i=1}^{n} x_i^2 f(x_i) - 2\mu \sum_{i=1}^{n} x_i f(x_i) + \mu^2 \sum_{i=1}^{n} f(x_i) \\
&= \sum_{i=1}^{n} x_i^2 f(x_i) - \mu^2 \\
&= E[X^2] - (E[X])^2
\end{aligned}$$

2　確率変数と確率分布 II

2.1　離散型一様分布

2.1.1　離散型一様分布の確率関数

確率関数 $f(x_i)$ が、

$$f(x_i) = P(X = x_i) = \frac{1}{n}$$

と表せるとき、確率変数 X は、一様分布に従うといい、DU(n) と表す。

例）さいころを一つ投げるときの出る目の値について

1 の目	2 の目	3 の目	4 の目	5 の目	6 の目
$\frac{1}{6}$	$\frac{1}{6}$	$\frac{1}{6}$	$\frac{1}{6}$	$\frac{1}{6}$	$\frac{1}{6}$

出る目の数に興味があるならば、目の数そのものを確率変数とすれば良く、各変数に対応する確率はすべて $\frac{1}{6}$ である。

2.1.2　DU(n) の期待値

$$\begin{aligned} E[X] &= \sum_{i=1}^{n} x_i f(x_i) \\ &= \sum_{i=1}^{n} x_i \frac{1}{n} \\ &= \frac{1}{n} \sum_{i=1}^{n} x_i \end{aligned}$$

2.1.3　DU(n) の分散

$$\begin{aligned} V[X] &= \sum_{i=1}^{n} (x_i - \mu)^2 f(x_i) \\ &= \sum_{i=1}^{n} (x_i - \mu)^2 \frac{1}{n} \\ &= \frac{1}{n} \sum_{i=1}^{n} (x_i - \mu)^2 \end{aligned}$$

2.2 2項分布

2.2.1 ベルヌーイ試行と2項分布の確率関数

事象が「成功」か「失敗」のみのような二種類だけである試行をベルヌーイ試行という。ベルヌーイ試行で、n 回試行するような場合、確率変数 X を試行回数とする。n 回試行した結果として確率変数 X は、

$$X = \{0, 1, 2, 3, \cdots, n\}$$

となる。このとき、成功する確率を p とすると失敗する確率は $1-p$ であり、成功事象が x 回とすると n 回試行したとして確率関数 $f(x)$ が、

$$f(x) = P(X = x) = {}_nC_x p^x (1-p)^{n-x}$$

と表せるとき、確率変数 X は、2項分布に従うといい、B(n,p) と表す。

例）さいころを5回転がすとき、1の目がちょうど2回出る確率は?

さいころは1の目が出る確率は $\frac{1}{6}$ を成功確率 p とする。このとき5回の試行で成功の選ばれる組み合わせは、

0回	1回	2回	3回	4回	5回
${}_5C_0$	${}_5C_1$	${}_5C_2$	${}_5C_3$	${}_5C_4$	${}_5C_5$
1	5	10	10	5	1

$$f(2) = 10 \cdot \left(\frac{1}{6}\right)^2 \cdot \left(\frac{5}{6}\right)^{5-2} = 10 \times \frac{1}{36} \times \frac{125}{216} = \frac{1250}{7776} \approx 0.16$$

2.2.2 B(n,p) の期待値

$$\begin{aligned}
E[X] &= \sum_{x=0}^{n} x\, {}_nC_x p^x (1-p)^{n-x} \\
&= \sum_{x=1}^{n} x \frac{n!}{x!(n-x)!} p^x (1-p)^{n-x} \\
&= \sum_{x=1}^{n} \frac{n!}{(x-1)!(n-x)!} p^x (1-p)^{n-x} \\
&= np \sum_{x=1}^{n} \frac{(n-1)!}{(x-1)!(n-x)!} p^{x-1} (1-p)^{n-x}
\end{aligned}$$

ここで、$y = x - 1$ とおくと

$$= np \sum_{y=0}^{n-1} \frac{(n-1)!}{y!(n-y-1)!} p^y (1-p)^{n-y-1}$$

ここで、$\sum_{y=0}^{n-1} \frac{(n-1)!}{y!(n-y-1)!} p^y (1-p)^{n-y-1}$ について考えると確率変数 y の2項分布の総和なので1

$$= np$$

$$E[X] = np$$

2.2.3 B(n,p) の分散

$$V[X] = E[X^2] - (E[X])^2$$

を用いる。まず、$E[X^2]$ を計算すると、

$$\begin{aligned}
E[X^2] &= \sum_{x=0}^{n} x^2 {}_nC_x p^x (1-p)^{n-x} \\
&\text{ここで、}x^2 = x(x-1) + x \text{ とおくと} \\
&= \sum_{x=0}^{n} (x(x-1)+x) \frac{n!}{x!(n-x)!} p^x (1-p)^{n-x} \\
&= \sum_{x=0}^{n} x(x-1) \frac{n!}{x!(n-x)!} p^x (1-p)^{n-x} + \mu \\
&\text{ここで、}x = 0 \text{ および } x = 1 \text{ はどちらも } 0 \\
&= \sum_{x=2}^{n} \frac{n!}{(x-2)!(n-x)!} p^x (1-p)^{n-x} + \mu \\
&= n(n-1)p^2 \sum_{x=2}^{n} \frac{(n-2)!}{(x-2)!(n-x)!} p^{x-2} (1-p)^{n-x} + \mu \\
&\text{ここで、}z = x - 2 \text{ とおくと} \\
&= n(n-1)p^2 \sum_{z=0}^{n-2} \frac{(n-2)!}{z!(n-z-2)!} p^z (1-p)^{n-z-2} + \mu \\
&\text{ここで、}\sum_{z=0}^{n-2} \frac{(n-2)!}{z!(n-z-2)!} p^z (1-p)^{n-z-2} \text{について考えると確率変数 } z \text{ の 2 項分布の総和なので } 1 \\
&= n(n-1)p^2 + \mu \\
&= n(n-1)p^2 + np \\
&= n^2p^2 - np^2 + np
\end{aligned}$$

よって、

$$\begin{aligned}
V[X] &= E[X^2] - (E[X])^2 \\
&= n^2p^2 - np^2 + np - (np)^2 \\
&= -np^2 + np \\
&= np(1-p) \\
V[X] &= np(1-p)
\end{aligned}$$

2.3 ポアソン分布

2.3.1 ポアソン分布の確率関数

２項分布において、「成功」事象が極めて稀にしか起こらない (成功の起こる確率がほとんどゼロに近い)ときを考える。とはいっても、ほとんどゼロに近くてもゼロでないなら試行回数が無限大に近づくと「成功」事象を無視できなくなってくる。このとき、確率関数として、

$$f(x) = \frac{\mu^x}{x!}e^{-\mu} \qquad (x = 0, 1, \cdots)$$

で近似的に表すことができる場合、確率変数 X は、ポアソン分布に従うといい、$\mathrm{Po}(\mu)$ と表す。

例）リンゴ 250 個が箱詰めされている。箱詰めしたリンゴは、平均 0.8 %腐ってしまう。箱詰めしたリンゴ１箱を開けたとき、３個以上腐っている確率は?

$n = 250$、$p = 0.008$。よって、$\mu = np = 2$。したがって、ポアソン分布の確率関数を用いれば、

$$f(0) = e^{-2} \quad f(1) = 2e^{-2} \quad f(2) = 2e^{-2}$$

よって、

$$P(X \geq 3) = 1 - (f(0) + f(1) + f(2)) = 1 - e^{-2} - 2e^{-2} - 2e^{-2} = 1 - 5e^{-2} \approx 0.32$$

■２項分布とポアソン分布の関係

$$\begin{aligned}
f(x) = \frac{n!}{x!(n-x)!}p^x(1-p)^{n-x} &= \frac{n(n-1)(n-2)(n-3)\cdots(n-x+1)}{x!}p^x(1-p)^{n-x} \\
&= n(n-1)(n-2)(n-3)\cdots(n-x+1)\frac{p^x}{x!}(1-p)^{n-x} \\
&= \frac{n(n-1)(n-2)(n-3)\cdots(n-x+1)}{n^x}\frac{n^x p^x}{x!}(1-p)^{n-x} \\
&= \left(1-\frac{1}{n}\right)\left(1-\frac{2}{n}\right)\cdots\left(1-\frac{x-1}{n}\right)\frac{(np)^x}{x!}(1-p)^{n-x} \\
&= \frac{\left(1-\frac{1}{n}\right)\left(1-\frac{2}{n}\right)\cdots\left(1-\frac{x-1}{n}\right)}{(1-p)^x}\frac{\mu^x}{x!}(1-p)^n
\end{aligned}$$

ここで、

$$(1-p)^n = \left((1-p)^{-\frac{1}{p}}\right)^{-np} = \left((1-p)^{-\frac{1}{p}}\right)^{-\mu}$$

について、$e = \lim_{z\to 0}(1+z)^{\frac{1}{z}}$ を用いて

$$\lim_{p\to 0}\left((1-p)^{-\frac{1}{p}}\right)^{-\mu} = e^{-\mu}$$

いま、試行回数 n が ∞ のとき、p が 0 に近づくときを考えているので、

$$\frac{\left(1-\frac{1}{n}\right)\left(1-\frac{2}{n}\right)\cdots\left(1-\frac{x-1}{n}\right)}{(1-p)^x} = 1$$

よって、

$$f(x) = \frac{\mu^x}{x!}e^{-\mu}$$

■きわめて稀とは　一般的に２項分布 B(n,p) が $n > 50,\ np \leq 5$ のとき、ポアソン分布で近似できる。

2.3.2　Po(μ) の期待値

e^μ についてのテイラー展開

$$e^\mu = \sum_{x=1}^{\infty} \frac{\mu^{x-1}}{(x-1)!}$$

を用いて、

$$\begin{aligned} E[X] &= \sum_{x=1}^{\infty} x\frac{\mu^x}{x!}e^{-\mu} \\ &= \mu\sum_{x=1}^{\infty} \frac{\mu^{x-1}}{(x-1)!}e^{-\mu} \\ &= \mu e^\mu e^{-\mu} \\ &= \mu \\ E[X] &= \mu \end{aligned}$$

2.3.3　Po(μ) の分散

e^μ についてのテイラー展開

$$e^\mu = \sum_{x=0}^{\infty} \frac{\mu^x}{(x)!} = \sum_{x=1}^{\infty} \frac{\mu^{x-1}}{(x-1)!} = \sum_{x=2}^{\infty} \frac{\mu^{x-2}}{(x-2)!}$$

を用いて、

$$\begin{aligned} V[X] &= E[X^2] - (E[X])^2 \\ &= \sum_{x=1}^{\infty} x^2\frac{\mu^x}{x!}e^{-\mu} - \mu^2 \\ &= \sum_{x=1}^{\infty} x^2\frac{\mu^x}{x!}e^{-\mu} - \sum_{x=1}^{\infty} x\frac{\mu^x}{x!}e^{-\mu} + \sum_{x=1}^{\infty} x\frac{\mu^x}{x!}e^{-\mu} - \mu^2 \\ &= \sum_{x=1}^{\infty} x(x-1)\frac{\mu^x}{x!}e^{-\mu} + \sum_{x=1}^{\infty} x\frac{\mu^x}{x!}e^{-\mu} - \mu^2 \end{aligned}$$

ここで、$x = 1$ のとき、$x(x-1)$ の項がゼロなので、総和は $x = 2$ からで取り替えて、

$$\begin{aligned} &= \sum_{x=2}^{\infty} x(x-1)\frac{\mu^x}{x!}e^{-\mu} + \mu - \mu^2 \\ &= \sum_{x=2}^{\infty} \frac{\mu^x}{(x-2)!}e^{-\mu} + \mu - \mu^2 \\ &= \sum_{x=2}^{\infty} \mu^2\frac{\mu^{x-2}}{(x-2)!}e^{-\mu} + \mu - \mu^2 \\ &= \mu^2 + \mu - \mu^2 \\ &= \mu \\ V[X] &= \mu \end{aligned}$$

2.4 幾何分布

2.4.1 幾何分布の確率関数

2 項分布においてベルヌーイ試行の試行回数を確率変数とした。このとき、試行回数ではなく、「成功」するまでの試行回数を確率変数 X とするときの確率分布を考える。

いま、ベルヌーイ試行の「成功」する確率を p とすると「失敗」する確率は $1-p$ なので

- $x=1$ とは試行回数 1 回目で「成功」したということ。よって、$f(1)=p$
- $x=2$ とは試行回数 2 回目で「成功」したということ。よって、$f(2)=(1-p)\cdot p$
- $x=3$ とは試行回数 3 回目で「成功」したということ。よって、$f(3)=(1-p)\cdot(1-p)\cdot p$
- $x=4$ とは試行回数 4 回目で「成功」したということ。よって、$f(4)=(1-p)\cdot(1-p)\cdot(1-p)\cdot p$
- $\cdots$

以上のように、このとき確率関数は、

$$f(x)=p(1-p)^{x-1}$$

と表すことができる。このとき、確率変数 X は、幾何分布に従うといい、$\mathrm{G}(p)$ と表す。

例）裏が 4 倍出にくいコインについて 5 回目ではじめて裏が出る確率は？

成功を「裏が出る」とするとき、成功確率 $p=\frac{1}{5}$、失敗確率 $(1-p)=\frac{4}{5}$ であるから、

$$f(5)=\left(\frac{1}{5}\right)\left(1-\frac{1}{5}\right)^{5-1}=\left(\frac{1}{5}\right)\left(\frac{4}{5}\right)^{4}=(0.2)\times(0.8)^4=0.2\times0.4096=0.08192$$

2.4.2 G(p) の期待値

途中 $q=1-p$ としておく。また、等比級数の和の公式は、$\sum_{x=1}^{\infty}\alpha\beta^{x-1}=\frac{\alpha}{1-\beta}\quad(|\beta|<1)$

$$\begin{aligned}
E[X]&=\sum_{x=1}^{n}xp(1-p)^{x-1}\\
&\text{ここで、}1-p=q\text{ とおく。}\\
&=\sum_{x=1}^{n}xpq^{x-1}=p\sum_{x=1}^{n}xq^{x-1}=p(1+2\cdot q+3\cdot q^2+\cdots)\\
&=p(1+q+q^2+\cdots)(1+q+q^2+\cdots)\\
&=p(\sum_{x=0}^{n}q^x)(\sum_{x=0}^{n}q^x)=p(\sum_{x=0}^{n}q^x)^2\\
&\text{ここで、初項 1 公比 }q\text{ の等比級数の和として、公式を適用すると}\\
&=p\left(\frac{1}{1-q}\right)^2=p\left(\frac{1}{1-(1-p)}\right)^2=p\left(\frac{1}{p}\right)^2=p\frac{1}{p^2}\\
&=\frac{1}{p}\\
E[X]&=\frac{1}{p}
\end{aligned}$$

2.4.3 G(p) の分散

$V[X] = E[X^2] - (E[X])^2$ を用いる。まず、$E[X^2]$ を計算しておく。

$$
\begin{aligned}
E[X^2] &= \sum_{x=1}^{n} x^2 p(1-p)^{x-1} \\
&= \sum_{x=1}^{n} x^2 p q^{x-1} \\
&= p \sum_{x=1}^{n} x^2 q^{x-1} \\
&= p(1 + 4 \cdot q + 9 \cdot q^2 + \cdots) \\
&= p(1 + q + q^2 + \cdots)(1 + 3q + 5q^2 + 7q^3 + \cdots) \\
&= p\left(\frac{1}{1-q}\right)(1 + 3q + 5q^2 + 7q^3 + \cdots) \\
&= p\frac{1}{p}(1 + 3q + 5q^2 + 7q^3 + \cdots) \\
&= (1 + 3q + 5q^2 + 7q^3 + \cdots) \\
&= \sum_{x=1}^{n}(2x-1)q^{x-1} \\
&= 2\sum_{x=1}^{n} xq^{x-1} - \sum_{x=1}^{n} q^{x-1} \\
&= 2\left(\frac{1}{1-q}\right)^2 - \frac{1}{1-q} \\
&= \frac{2}{p^2} - \frac{1}{p} \\
&= \frac{2-p}{p^2}
\end{aligned}
$$

よって、

$$
\begin{aligned}
V[X] &= E[X^2] - (E[X])^2 \\
&= \frac{2-p}{p^2} - \left(\frac{1}{p}\right)^2 \\
&= \frac{1-p}{p^2}
\end{aligned}
$$

2.5 超幾何分布

2.5.1 超幾何分布の確率関数

全体集合の元の個数を N とする。このうち M 個が種類 A、残り $(N-M)$ 個が種類 B であるとする。この全体集合から無作為に n 個取り出すとすると、${}_NC_n$ の組み合わせが存在する。言い換えると、n 個の組み合わせは、すべて等確率 $\frac{1}{{}_NC_n}$ で選ばれる。当然、選ばれた n 個の組の中には A と B の二種類の性質が混ざっている。そこで、n 個のうち、種類 A である個数を x 個とすると、種類 B である個数は $(n-x)$ 個である。種類 A の x 個とは、もともと M 個のうちから選ばれたものであるから、組み合わせは ${}_MC_x$ である。一方、種類 B の $(n-x)$ 個は、$(N-M)$ 個から選ばれたものであるから、組み合わせは、${}_{N-M}C_{n-x}$ である。したがって、それぞれの選ばれ方は独立なので、$({}_MC_x)\cdot({}_{N-M}C_{n-x})$ である。このとき、確率関数 $f(x)$ が、

$$f(x)=\frac{({}_MC_x)\cdot({}_{N-M}C_{n-x})}{{}_NC_n}$$

と表せるとき、確率変数 X は、超幾何分布に従うといい、$p=\frac{M}{N}$ として HG(N,n,p) と表す。幾何分布との関係はない。

例）アメリカ合衆国は 50 州から成り立っているが、各州 2 人づつの上院議員が選出される。１００人の上院議員の中から無作為に５０人を選んだとき、その中にイリノイ州選出の議員が x 人はいっている確率は?

$N=100$、$n=50$、$M=2$ の超幾何分布として考えることができる。

$$f(x)=\frac{{}_2C_x\cdot{}_{100-2}C_{50-x}}{{}_{100}C_{50}}$$

$$f(0)=\frac{49}{198}\quad f(1)=\frac{50}{99}\quad f(2)=\frac{49}{198}$$

つまり、だれもいない 0 人の確率は $\frac{49}{198}$。1 人の確率は $\frac{50}{99}$、2 人の確率は $\frac{49}{198}$ である。

■２項分布との関係　$N>>n$ のとき、$p=\frac{M}{N}$ として 2 項分布で近似できる。

$$\begin{aligned}
f(x)&=\frac{({}_MC_x)\cdot({}_{N-M}C_{n-x})}{{}_NC_n}\\
&=\frac{\frac{M(M-1)(M-2)\cdots(M-x+1)}{x!}\,\frac{(N-M)(N-M-1)(N-M-2)\cdots(N-M-(n-x)+1)}{(n-x)!}}{\frac{N(N-1)(N-2)\cdots(N-n+1)}{n!}}\\
&=\frac{n!}{x!(n-x)!}\frac{(M\cdot M-1\cdot M-2\cdots M-x+1)(N-M\cdot N-M-1\cdot N-M-2\cdots N-M-(n-x)+1)}{N\cdot N-1\cdot N-2\cdots N-n+1}\\
&={}_nC_x\frac{(M\cdot M-1\cdot M-2\cdots M-x+1)(N-M\cdot N-M-1\cdot N-M-2\cdots N-M-(n-x)+1)}{N\cdot N-1\cdot N-2\cdots N-n+1}\\
&\approx{}_nC_x\frac{M^x(N-M)^{n-x}}{N^n}\\
&={}_nC_x\frac{M^x}{N^x}\frac{(N-M)^{n-x}}{N^{n-x}}\\
&={}_nC_xp^x(1-p)^{n-x}
\end{aligned}$$

2.5.2 HG(N,n,p) の期待値

$p = \frac{M}{N}$ とする。

$$
\begin{aligned}
E[X] &= \sum_{x=0}^{n} x \frac{({}_M C_x) \cdot ({}_{N-M} C_{n-x})}{{}_N C_n} \\
&= \frac{M}{{}_N C_n} \sum_{x=1}^{n} ({}_{M-1} C_{x-1}) \cdot ({}_{N-M} C_{n-x}) \\
&= \frac{M}{{}_N C_n}\, {}_{N-1} C_{n-1} \\
&= n \frac{M}{N} = np \\
E[X] &= np
\end{aligned}
$$

ここで、$\sum_{x=1}^{n} ({}_{M-1} C_{x-1}) \cdot ({}_{N-M} C_{n-x}) = {}_{N-1} C_{n-1}$ については、組み合わせ記号の性質：その 2 を用いた。

2.5.3 HG(N,n,p) の分散

$$
\begin{aligned}
V[X] &= E[X^2] - (E[X])^2 \\
&= E[X(X-1)] + E[X] - (E[X])^2
\end{aligned}
$$

を用いる。

まず、

$$
\begin{aligned}
E[X(X-1)] &= \sum_{x=0}^{n} x(x-1) \frac{({}_M C_x) \cdot ({}_{N-M} C_{n-x})}{{}_N C_n} \\
&= \frac{M(M-1)}{{}_N C_n} \sum_{x=2}^{n} ({}_{M-2} C_{x-2}) \cdot ({}_{N-M} C_{n-x}) \\
&= \frac{M(M-1)}{{}_N C_n}\, {}_{N-2} C_{n-2} \\
&= \frac{n(n-1)M(M-1)}{N(N-1)}
\end{aligned}
$$

よって、

$$
\begin{aligned}
V[X] &= E[X^2] - (E[X])^2 \\
&= E[X(X-1)] + E[X] - (E[X])^2 \\
&= \frac{n(n-1)M(M-1)}{N(N-1)} + n\frac{M}{N} - \left(n\frac{M}{N}\right)^2 \\
&= \frac{N-n}{N-1} np(1-p) \\
V[X] &= \frac{N-n}{N-1} np(1-p)
\end{aligned}
$$

2.6　負の２項分布

2.6.1　負の２項分布の確率関数：k 回成功するまでに必要な試行回数を確率変数 X と取る場合

ベルヌーイ試行による成功確率 p で k 回の成功が得られるまでの試行回数が x 回であった場合の確率関数は、全試行回数が x 回で、x 回目が k 回目の成功であるから、その直前までが、$x-1$ 回の試行のうち $k-1$ の成功で、その組み合わせは、${}_{x-1}C_{k-1}$ であるから、

$$f(x) = {}_{x-1}C_{k-1}p^k(1-p)^{x-k}$$

このとき、確率変数 X は、負の２項分布に従うといい、NB(k,p)$_+$ と表す。$k=1$ のとき、通常の幾何分布である。つまり、G(p)=NB(1,p)$_+$ である。

2.6.2　NB(k,p)$_+$ の期待値

$$\begin{aligned}
E[X] &= \sum_{x=0}^{x} x\,{}_{x-1}C_{k-1}p^k(1-p)^{x-k} \\
&= \sum_{x=0}^{x} x\frac{(x-1)(x-2)\cdots(x-1-(k-1)+1)}{(k-1)!}p^k(1-p)^{x-k} \\
&= \sum_{x=0}^{x} x\frac{(x-1)(x-2)\cdots(x-1-k+1+1)}{(k-1)!}p^k(1-p)^{x-k} \\
&= \sum_{x=0}^{x} x\frac{(x-1)(x-2)\cdots(x-k+1)}{(k-1)!}p^k(1-p)^{x-k} \\
&= \sum_{x=0}^{x} k\frac{x(x-1)(x-2)\cdots(x-k+1)}{k(k-1)!}p^k(1-p)^{x-k} \\
&= \sum_{x=0}^{x} k\frac{x(x-1)(x-2)\cdots(x-k+1)}{k!}p^k(1-p)^{x-k} \\
&= \sum_{x=0}^{x} k\,{}_xC_kp^k(1-p)^{x-k} \\
&= k\sum_{x=0}^{x} {}_xC_k\frac{p^{k+1}}{p}(1-p)^{(x+1)-(k+1)} \\
&= \frac{k}{p}\sum_{x=0}^{x} {}_xC_kp^{k+1}(1-p)^{x-k} \\
&\quad \text{ここで、全確率} \sum_{x=0}^{x} {}_xC_kp^{k+1}(1-p)^{x-k} = 1 \text{ であるから、} \\
&= \frac{k}{p} \\
E[X] &= \frac{k}{p}
\end{aligned}$$

2.6.3 NB(k,p)$_+$ の分散

$$V[X] = E[X^2] - (E[X])^2$$

を用いる。まず、$E[X^2]$ を計算すると、

$$\begin{aligned}
E[X^2] &= \sum_{x=0}^{x} x^2 \,{}_{x-1}C_{k-1} p^k (1-p)^{x-k} \\
&\text{ここで、}x^2 = x(x+1) - x \text{ とおく} \\
&= \sum_{x=0}^{x} (x(x+1) - x) \,{}_{x-1}C_{k-1} p^k (1-p)^{x-k} \\
&= \sum_{x=0}^{x} x(x+1) \,{}_{x-1}C_{k-1} p^k (1-p)^{x-k} - \sum_{x=0}^{x} x \,{}_{x-1}C_{k-1} p^k (1-p)^{x-k} \\
&\text{第 2 項目は期待値そのものだから、} \\
&= \sum_{x=0}^{x} x(x+1) \,{}_{x-1}C_{k-1} p^k (1-p)^{x-k} - \frac{k}{p} \\
&= \sum_{x=0}^{x} x(x+1) \frac{(x-1)(x-2)\cdots(x-1-(k-1)+1)}{(k-1)!} p^k (1-p)^{x-k} - \frac{k}{p} \\
&= \sum_{x=0}^{x} \frac{(x+1)x(x-1)(x-2)\cdots(x-k+1)}{(k-1)!} p^k (1-p)^{x-k} - \frac{k}{p} \\
&= \sum_{x=0}^{x} (k+1)k \frac{(x+1)x(x-1)(x-2)\cdots(x-k+1)}{(k+1)!} p^k (1-p)^{x-k} - \frac{k}{p} \\
&= (k+1)k \sum_{x=0}^{x} \,{}_{x+1}C_{k+1} p^k (1-p)^{x-k} - \frac{k}{p} \\
&= (k+1)k \sum_{x=0}^{x} \,{}_{x+1}C_{k+1} \frac{p^{k+2}}{p^2} (1-p)^{(x+2)-(k+2)} - \frac{k}{p} \\
&= \frac{(k+1)k}{p^2} \sum_{x=0}^{x} \,{}_{x+1}C_{k+1} p^{k+2} (1-p)^{x-k} - \frac{k}{p} \\
&\text{ここで、全確率} \sum_{x=0}^{x} \,{}_{x+1}C_{k+1} p^{k+2} (1-p)^{x-k} = 1 \text{ であるから、} \\
&= \frac{(k+1)k}{p^2} - \frac{k}{p}
\end{aligned}$$

よって、

$$\begin{aligned}
V[X] &= E[X^2] - (E[X])^2 \\
&= \frac{(k+1)k}{p^2} - \frac{k}{p} - \left(\frac{k}{p}\right)^2 \\
&= \frac{(k+1)k - pk - k^2}{p^2}
\end{aligned}$$

$$
\begin{aligned}
&= \frac{k^2 + k - pk - k^2}{p^2} \\
&= \frac{k - pk}{p^2} \\
&= \frac{k(1-p)}{p^2} \\
V[X] &= \frac{k(1-p)}{p^2}
\end{aligned}
$$

2.6.4 負の２項分布の確率関数：k 回成功するまでに失敗した試行回数を確率変数 Y と取る場合

ベルヌーイ試行による成功確率 p で k 回の成功が得られるまでに失敗した回数が y 回であった場合の確率関数は、全試行回数が $x = k + y$ 回で、$k + y$ 回目が k 回目の成功であるから、その直前までが、$x - 1 = k + y - 1$ 回の試行のうち $k - 1$ の成功で、その組み合わせは、${}_{x-1}C_{k-1} = {}_{k+y-1}C_{k-1}$ であるから、

$$
f(y) = {}_{k+y-1}C_{k-1} p^k (1-p)^y = {}_{k+y-1}C_y p^k (1-p)^y
$$

このとき、確率変数 Y は、負の２項分布に従うといい、NB(k,p)$_-$ と表す。また、$k + y - 1$ 回のうち $k - 1$ 回の成功と y 回の失敗であるから、${}_{k+y-1}C_{k-1} = {}_{k+y-1}C_y$ どちらの組み合わせを用いて表しても同じである。

2.6.5 NB(k,p)$_-$ の期待値

$$
\begin{aligned}
E[Y] &= \sum_{y=0}^{y} y \, {}_{k+y-1}C_{k-1} p^k (1-p)^y \\
&= \sum_{y=0}^{y} y \frac{(k+y-1)(k+y-2)\cdots(k+y-1-(k-1)+1)}{(k-1)!} p^k (1-p)^y \\
&= \sum_{y=0}^{y} y \frac{(k+y-1)(k+y-2)\cdots(y+1)}{(k-1)!} p^k (1-p)^y \\
&= \sum_{y=0}^{y} k \frac{(k+y-1)(k+y-2)\cdots(y+1)y}{k!} p^k (1-p)^y \\
&= \sum_{y=0}^{y} k \, {}_{k+y-1}C_k p^k (1-p)^y \\
&= \sum_{y=0}^{y} \frac{k(1-p)}{p} {}_{k+y-1}C_k p^{k+1} (1-p)^{y-1} \\
&= \frac{k(1-p)}{p} \sum_{y=0}^{y} {}_{k+y-1}C_k p^{k+1} (1-p)^{y-1} \\
&\quad \text{ここで、全確率} \sum_{y=0}^{y} {}_{k+y-1}C_k p^{k+1} (1-p)^{y-1} = 1 \text{ であるから、} \\
&= \frac{k(1-p)}{p} \\
E[Y] &= \frac{k(1-p)}{p}
\end{aligned}
$$

2.6.6 NB(k,p)$_-$ の分散

$$V[Y]=E[Y^2]-(E[Y])^2$$

を用いる。まず、$E[Y^2]$ を計算すると、

$$\begin{aligned}
E[Y^2]&=\sum_{y=0}^{y}y^2\ {}_{k+y-1}C_{k-1}p^k(1-p)^y\\
&\text{ここで、}y^2=y(y-1)+y\text{ とおく}\\
&=\sum_{y=0}^{y}(y(y-1)+y)\ {}_{k+y-1}C_{k-1}p^k(1-p)^y\\
&=\sum_{y=0}^{y}y(y-1)\ {}_{k+y-1}C_{k-1}p^k(1-p)^y+\sum_{y=0}^{y}y\ {}_{k+y-1}C_{k-1}p^k(1-p)^y\\
&\text{第 2 項目は期待値そのものだから、}\\
&=\sum_{y=0}^{y}y(y-1)\ {}_{k+y-1}C_{k-1}p^k(1-p)^y+\frac{k(1-p)}{p}\\
&=\sum_{y=0}^{y}y(y-1)\frac{(k+y-1)(k+y-2)\cdots(k+y-1-(k-1)+1)}{(k-1)!}p^k(1-p)^y+\frac{k(1-p)}{p}\\
&=\sum_{y=0}^{y}(k+1)k\frac{(k+y-1)(k+y-2)\cdots(y+1)y(y-1)}{(k+1)!}\frac{p^{k+2}}{p^2}(1-p)^y+\frac{k(1-p)}{p}\\
&=\frac{(k+1)k(1-p)^2}{p^2}\sum_{y=0}^{y}\ {}_{k+y-1}C_{k+1}p^{k+2}(1-p)^{y-2}+\frac{k(1-p)}{p}\\
&\text{ここで、全確率}\sum_{y=0}^{y}\ {}_{k+y-1}C_{k+1}p^{k+2}(1-p)^{y-2}=1\text{ であるから、}\\
&=\frac{(k+1)k(1-p)^2}{p^2}+\frac{k(1-p)}{p}
\end{aligned}$$

よって、

$$\begin{aligned}
V[Y]&=E[Y^2]-(E[Y])^2\\
&=\frac{(k+1)k(1-p)^2}{p^2}+\frac{k(1-p)}{p}-\left(\frac{k(1-p)}{p}\right)^2\\
&=\frac{(k+1)k(1-p)^2+pk(1-p)-k^2(1-p)^2}{p^2}\\
&=\frac{k^2(1-p)^2+k(1-p)^2+pk(1-p)-k^2(1-p)^2}{p^2}\\
&=\frac{k(1-p)\{(1-p)+p\}}{p^2}\\
V[Y]&=\frac{k(1-p)}{p^2}
\end{aligned}$$

2.7　離散型多次元確率分布

2.7.1　多項分布

２項分布を多次元に拡張する。いま、r 個の事象が試行ごとにそれぞれ確率 $p_1, p_2, \cdots, p_r$ で起こるものとする。つまり、$p_1 + p_2 + \cdots + p_r = 1$ である。この試行を n 回繰り返すとき、最大 r 種類の事象が全部で n 個得られる。このとき、r 種類ごとの個々の事象の個数を $X_1, X_2, \cdots, X_r$ とすると、$X_1 + X_2 + \cdots + X_r = n$ であり、これをもとに確率変数を $\boldsymbol{X} = \{X_1, X_2, \cdots, X_r\}$ として導入すると、確率分布は r 個の変数を持つ多次元確率分布となり

$$\Omega = \frac{n!}{x_1!x_2!\cdots x_r!} p_1^{x_1} p_2^{x_2} \cdots p_r^{x_r}$$

と表せる。このとき、確率変数 $\boldsymbol{X}$ は、多項分布に従うといい、$\mathrm{M}_r(n; p_1, p_2, \cdots, p_r)$ と表す。

例）さいころを７回投げることを考える。正しいさいころならば、$p_1 = \frac{1}{6}, p_2 = \frac{1}{6}, \cdots, p_6 = \frac{1}{6}$ である。このとき、１の目が１回、２の目が０回、３の目が２回、４の目が１回、５の目が１回、６の目が２回であったときの確率は？

$$\begin{aligned}\Omega &= \frac{7\cdot6\cdot5\cdot4\cdot3\cdot2\cdot1}{1!\cdot0!\cdot2!\cdot1!\cdot1!\cdot2!}\left(\frac{1}{6}\right)^1\cdot\left(\frac{1}{6}\right)^0\cdot\left(\frac{1}{6}\right)^2\cdot\left(\frac{1}{6}\right)^1\cdot\left(\frac{1}{6}\right)^1\cdot\left(\frac{1}{6}\right)^2 \\ &= 0.0045\end{aligned}$$

索引