6　第 6 回目講義の補足

6.1　ベルヌーイ試行

コインを n 回投げて表裏の出方を観測する試行を考える。コインは歪んでいるかも知れないものとする。この試行のポイントは

(i) 表が出る確率は各回で共通である;

(ii) 各回の結果は互いに独立である

の 2 点にある。数学的に書くため、

$$E_i = \{ i \text{ 回目に表が出る } \} \quad (i = 1, 2, \cdots, n) \tag{1}$$

と置けば、

(i) $P(E_1) = P(E_2) = \cdots = P(E_n) = p$;

(ii) $E_1, E_2, \cdots, E_n$ は独立

となる。

結果が 2 通りしかない試行は全てコイン投げと同一視出来るから、これをもう少し抽象的な言葉に直しておくと便利である。以後、コインの表裏をそれぞれ「成功」「失敗」と呼べば、p は「1 回当たりの成功確率」と解釈出来る。条件 (i)(ii) を満たす試行を**「長さ n、成功確率 p のベルヌーイ試行」**と言う。

6.2　2 項分布 $B(n, p)$ の導出

長さ n のベルヌーイ試行において、

$$X = n \text{ 回中の表 (成功) の回数}$$

とおくと、$X \sim B(n, p)$ である。即ち、

$$P(X = x) = {}_nC_x p^x (1-p)^{n-x} \quad (x = 0, 1, \cdots, n).$$

(証明) 各回の試行は独立であることに注意しよう。

$$\begin{aligned}
P(X=0) &= P\Big(\{ \text{ 全て裏 } \}\Big) && (1)\\
&= P\Big(E_1^c \cap E_2^c \cap \cdots \cap E_n^c\Big) && (2)\\
&= P\Big(E_1^c\Big) \times P\Big(E_2^c\Big) \times \cdots \times P\Big(E_n^c\Big) \quad (\text{独立性より}) && (3)\\
&= (1-p) \times (1-p) \times \cdots \times (1-p) && (4)\\
&= (1-p)^n \ \Big(= {}_nC_0 p^0 (1-p)^{n-0}, \ \text{何故なら}\ {}_nC_0 = 1 \Big) && (5)
\end{aligned}$$

となる。次に

$$\begin{aligned}
P(X=1) &= P\Big(\{1 \text{ 回目のみ表 }\} \cup \{2 \text{ 回目のみ表 }\} \cup \cdots \cup \{n \text{ 回目のみ表 }\}\Big)\\
&= P\Big(\Big[E_1 \cap E_2^c \cap \cdots \cap E_n^c\Big] \cup \Big[E_1^c \cap E_2 \cap E_3^c \cdots \cap E_n^c\Big]\\
&\qquad \cup \cdots \cup \Big[E_1^c \cap \cdots \cap E_{n-1}^c \cap E_n\Big]\Big) && (6)\\
&= P\Big(E_1 \cap E_2^c \cap \cdots \cap E_n^c\Big) + P\Big(E_1^c \cap E_2 \cap E_3^c \cdots \cap E_n^c\Big)\\
&\qquad + \cdots + P\Big(E_1^c \cap \cdots \cap E_{n-1}^c \cap E_n\Big) \quad (\text{排反}) && (7)
\end{aligned}$$

これは確率の n 個の足し算だが、実はおのおのの確率は全て等しい。実際、

$$\begin{aligned}
P\Big(E_1 \cap E_2^c \cap \cdots \cap E_n^c\Big) &= p \times (1-p) \times \cdots \times (1-p) = p(1-p)^{n-1} && (8)\\
P\Big(E_1^c \cap E_2 \cap E_3^c \cdots \cap E_n^c\Big) &= (1-p) \times p \times \cdots \times (1-p) = p(1-p)^{n-1} && (9)\\
&\ \ \vdots\\
P\Big(E_1^c \cap \cdots \cap E_{n-1}^c \cap E_n\Big) &= (1-p) \times (1-p) \times \cdots \times p = p(1-p)^{n-1} && (10)
\end{aligned}$$

従って (7) に戻れば、

$$P\Big(X=1\Big) = n \times p(1-p)^{n-1} \ \Big(= {}_nC_1 p^1 (1-p)^{n-1},\ 何故なら {}_nC_1 = n\Big) \tag{11}$$

が得られる。同様に考えて、$P(X=x)$ を求めよう。まず

$$P\Big(\{\text{ 最初の } x \text{ 回は表、あとは裏 }\}\Big) = P\Big(E_1 \cap \cdots E_x \cap E_{x+1}^c \cap \cdots \cap E_n^c\Big) = p^x(1-p)^{n-x} \tag{12}$$

は明らか。従って

$$P\Big(X=x\Big) = \text{「} n \text{ 回中 } x \text{ 回 } 1 \text{ が出る場合の数」} \times\ p^x(1-p)^{n-x} = {}_nC_x p^x (1-p)^{n-x} \tag{13}$$

が分かる。(証明終)

6.3 2 項分布に従う確率変数の期待値と分散

定理 3. $X \sim B(n,p)$ とする。

(1) $E(X) = np$

(2) $V(X) = np(1-p)$

(証明) 証明には、2 項定理を用いる：

定理 4.（2 項定理） A、B を任意の実数、m を正の整数とすると、次式が成立する。

$$(A+B)^m = {}_mC_0 A^0 B^m + {}_mC_1 A^1 B^{m-1} + {}_mC_2 A^2 B^{m-2} + \cdots + {}_mC_m A^m B^0 = \sum_{k=0}^{m} {}_mC_k A^k B^{m-k}$$

2 項定理を用いて、$E(X) = np$ を示してみよう。

$$\begin{aligned}
E(X) &= \sum_{x=0}^{n} x\ {}_nC_x p^x (1-p)^{n-x} && (14)\\
&= \sum_{x=0}^{n} x\ \frac{n!}{x!(n-x)!}\ p^x(1-p)^{n-x} \quad (\text{2 項係数の定義}) && (15)\\
&= \sum_{x=1}^{n} x\ \frac{n!}{x!(n-x)!}\ p^x(1-p)^{n-x} \quad (x=0 \text{ の項は明らかに } 0) && (16)\\
&= \sum_{x=1}^{n} \frac{n!}{(x-1)!(n-x)!}\ p^x(1-p)^{n-x} \quad \left(\frac{x}{x!} = \frac{1}{(x-1)!}\right) && (17)\\
&= \sum_{x=1}^{n} \frac{n \times (n-1)!}{(x-1)![(n-1)-(x-1)]!}\ p \times p^{x-1}(1-p)^{n-x} \quad \big(n! = n \times (n-1)!,\ p^x = p \times p^{x-1}\big) && (18)\\
&= np \sum_{x=1}^{n} \frac{(n-1)!}{(x-1)![(n-1)-(x-1)]!}\ p^{x-1}(1-p)^{(n-1)-(x-1)} \quad (np \text{ を外へ出した}) && (19)\\
&= np \sum_{k=0}^{m} \frac{m!}{k!(m-k)!} p^k (1-p)^{m-k} \quad (m = n-1,\ k = x-1 \text{ と置き換え}) && (20)\\
&= np \sum_{k=0}^{m} {}_mC_k p^k (1-p)^{m-k} && (21)\\
&= np\,[p + (1-p)]^m \quad (\text{2 項定理で } A = p, B = 1-p \text{ とする}) && (22)\\
&= np \times 1^m && (23)\\
&= np && (24)
\end{aligned}$$

分散については

$$V(X) = E(X^2) - \{E(X)\}^2$$

や、

$$V(X) = E\{X(X-1)\} + E(X) - \{E(X)\}^2 \tag{25}$$

などの公式が便利である。上と同様の方法で、

$$E\{X(X-1)\} = n(n-1)p^2$$

が示せる[1]。これを (25) に代入して、

$$V(X) = n(n-1)p^2 + np - (np)^2 = np(1-p)$$

が得られる。(証明終)

例 6.1. (コイン投げ) コインを 100 回投げ、表の出た回数を X で表す。このとき、$\mathcal{R}(X) = \{0, 1, \cdots, n\}$ であり、

$$X \sim B(100, 0.5)$$

である。即ち、

$$P(X = x) = {}_{100}C_x (0.5)^x (1-0.5)^{100-x} \quad (= {}_{100}C_x / 2^{100}).$$

平均や分散などについては、公式より

$$\begin{aligned} E(X) &= np = 100 \times 0.5 = 50 \\ V(X) &= np(1-p) = 100 \times 0.5 \times 0.95 = 47.5 \\ D(X) &= \sqrt{V(X)} \approx 6.9. \end{aligned}$$

と求まる ($n = 100$、$p = 0.5$ とおいている)。例えば、$P(X \geq 56)$ となる確率はどれくらいだろうか (予想してみよう)。実際に計算してみると次の通り、

$$\begin{aligned} P(X \geq 56) &= P(X = 56) + \cdots + P(X = 100) \\ &= \sum_{x=56}^{100} {}_{100}C_x (0.5)^x (1-0.5)^{100-x} \\ &\approx 0.136 \end{aligned}$$

である。同様に、$P(X \leq 44) \approx 0.136$、$P(45 \leq X \leq 55) \approx 0.729$ なども求まる。//

6.3.1 Poisson 分布 $P_O(\lambda)$ についての補足

例 6.2. (交通事故件数)

(1) 人口 100 万人の都市があり、この都市では 1 日当りおよそ 2 件の交通事故が発生するものとする。この都市の第 i 番目の住人 ($i = 1, 2, \cdots, 100$ 万) が 1 日に事故に遭うか遭わないかを表す確率変数を

$$X_i = \begin{cases} 1 & (\text{事故に遭う}) \\ 0 & (\text{事故に遭わない}) \end{cases}$$

と定義する ($i = 1, 2, \cdots, 100$ 万) と、これは

$$P(X_i = 1) = \frac{2}{100\,\text{万}} = p,\ P(X_i = 0) = 1 - \frac{2}{100\,\text{万}} = 1 - p$$

なる長さ $n(= 100$ 万) のベルヌーイ試行と考えることが出来る。従って、

$$X = \sum_{i=1}^{100\,\text{万}} X_i = \text{この都市の 1 日の事故件数}$$

とおけば、

$$X \sim Bi\left(100\,\text{万}, \frac{2}{100\,\text{万}}\right)$$

[1] 概略は次の通り。$E\{X(X-1)\} = \sum_{x=0}^{n} x(x-1) \frac{n!}{x!(n-x)!} p^x (1-p)^{n-x} = n(n-1)p^2 \sum_{x=2}^{n} {}_{n-2}C_{x-2} p^{x-2} (1-p)^{(n-2)-(x-2)} = n(n-1)p^2 \{p + (1-p)\}^{n-2} = n(n-1)p^2$.

である。従って、$E(X)=np=2$ であり、また

$$
\begin{aligned}
P(X=0) &= {}_{100\,万}C_0\left(\frac{2}{100\,万}\right)^0\left(1-\frac{2}{100\,万}\right)^{100\,万-0}\\
P(X=1) &= {}_{100\,万}C_1\left(\frac{2}{100\,万}\right)^1\left(1-\frac{2}{100\,万}\right)^{100\,万-1}\\
P(X=2) &= {}_{100\,万}C_2\left(\frac{2}{100\,万}\right)^2\left(1-\frac{2}{100\,万}\right)^{100\,万-2}\\
&\ \vdots
\end{aligned}
$$

等として確率を求めることが出来るが、計算効率、数理モデルとしての扱いやすさなどの観点から見て合理的なモデルとは言い難い。

(2) $X \sim P_O(2)$ と考えれば、

$$
\begin{aligned}
P(X=0) &= e^{-2}\frac{2^0}{0!}\approx 0.135,\quad P(X=1)=e^{-2}\frac{2^1}{1!}\approx 0.271,\quad P(X=2)=e^{-2}\frac{2^2}{2!}\approx 0.271,\\
P(X=3) &= e^{-2}\frac{2^3}{3!}\approx 0.180,\quad P(X=4)=e^{-2}\frac{2^4}{4!}\approx 0.090,\quad \cdots
\end{aligned}
$$

であることが容易に求まる。

1. Chebyshev の不等式についても復習しておこう。これによれば 2 シグマ範囲の確率は

$$P\Big(\frac{|X-\lambda|}{\sqrt{\lambda}}\le 2\Big)=P\Big(\lambda-2\sqrt{\lambda}\le X\le \lambda+2\sqrt{\lambda}\Big)\ge \frac{3}{4}=0.75$$

となるのであった。$\lambda=2$ のとき、

$$左辺 = P(2-2\sqrt{2}\le X\le 2+2\sqrt{2})=P(0\le X\le 4)$$

である。実際にこれを計算してみると 0.947 となる。//

6.4 幾何分布 $Ge(p)$ に関する補足

[1] 幾何分布の導出
(長さを定めない) ベルヌーイ試行において

$$X = 最初の表\ (成功)\ が現れる回$$

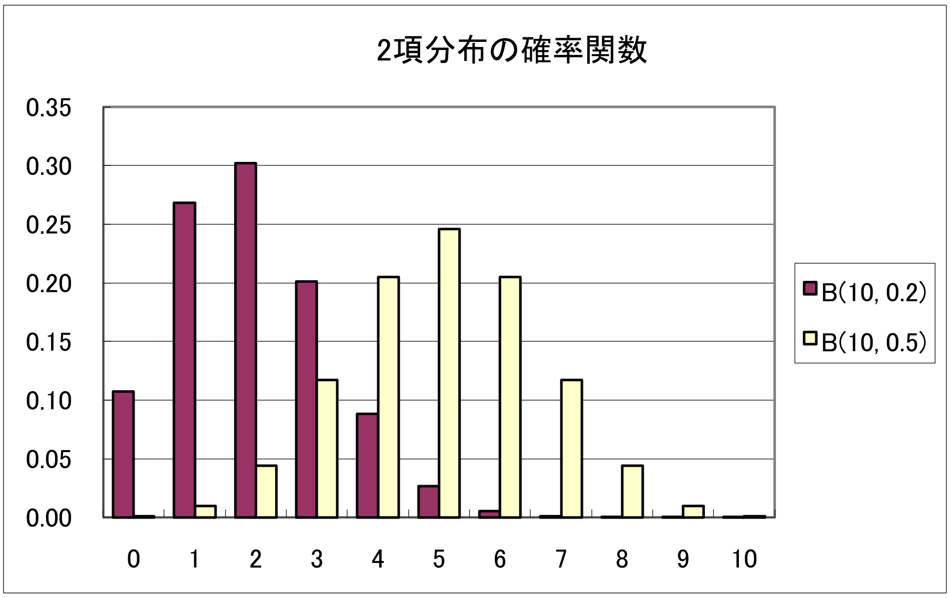
と置くと、$X \sim G_e(p)$ である。即ち、

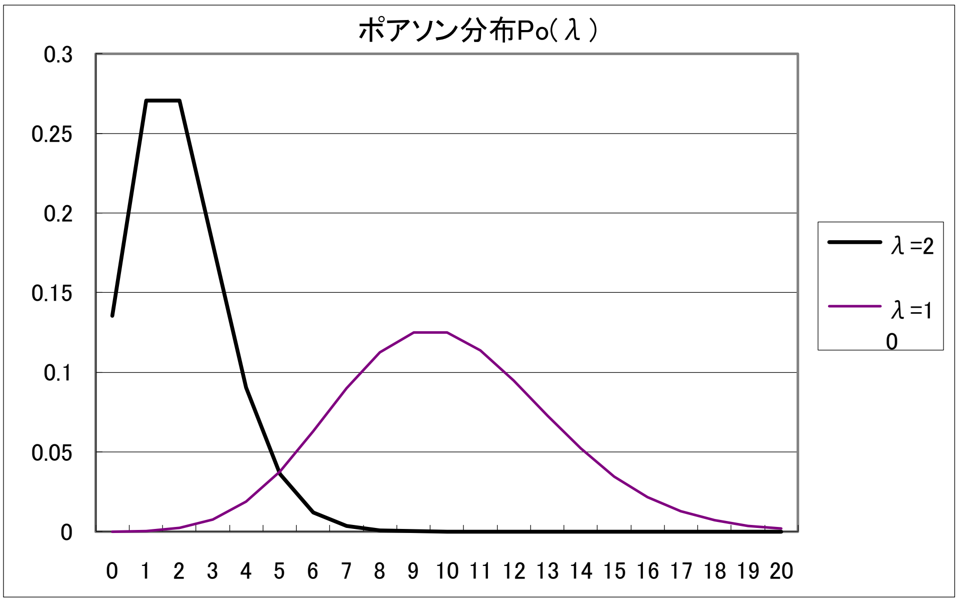
$$P(X=x)=p(1-p)^{x-1}\ (x=1,2,\cdots).$$

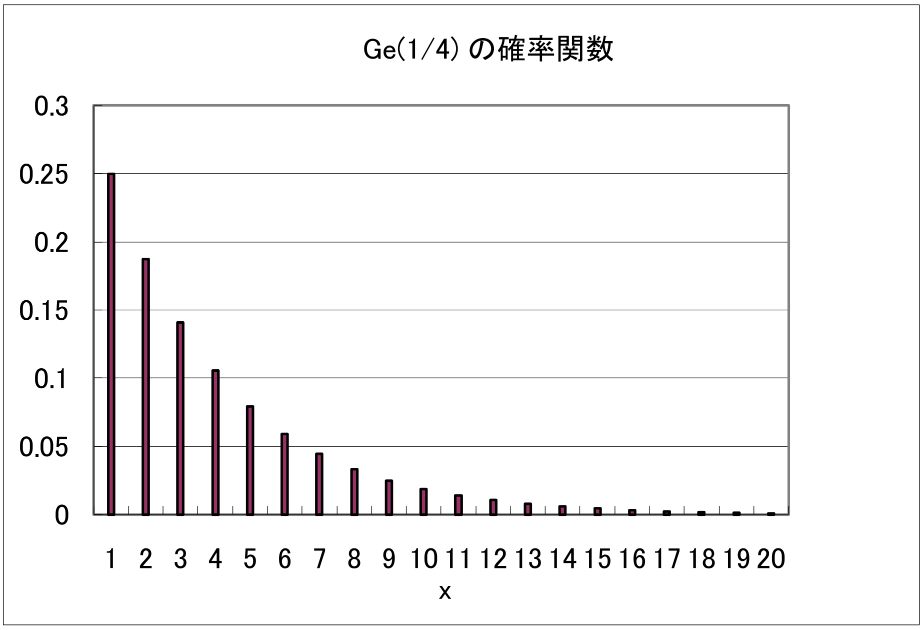
(証明) 証明は計算。

$$
\begin{aligned}
P(X=x) &= P\Big(E_1^c\cap\cdots\cap E_{x-1}^c\cap E_x\Big) &(26)\\
&= P\Big(E_1^c\Big)\times\cdots\times P\Big(E_{x-1}^c\Big)\times P\Big(E_x\Big)\ (独立性) &(27)\\
&= (1-p)^{x-1}p.\ \ (証明終) &(28)
\end{aligned}
$$

[2](幾何分布 $G_e(1/4)$ のグラフ)


2項分布の確率関数
0.35
0.30
0.25
0.20
0.15
0.10
0.05
0.00
0 1 2 3 4 5 6 7 8 9 10
B(10, 0.2)
B(10, 0.5)


ポアソン分布Po(λ)
0.3
0.25
0.2
0.15
0.1
0.05
0
0 1 2 3 4 5 6 7 8 9 10 11 12 13 14 15 16 17 18 19 20
λ=2
λ=10


Ge(1/4) の確率関数
0.3
0.25
0.2
0.15
0.1
0.05
0
1 2 3 4 5 6 7 8 9 10 11 12 13 14 15 16 17 18 19 20
x