

下書き 再生性

二つの確率変数が独立に同一の確率分布に従うとし、その和も同じ確率分布に従うとき、その確率分布は再生性を持つという。

離散確率分布 二項分布

基本的な確率分布であるベルヌーイ分布の再生性を調べる。ここでは二変数はベルヌーイ分布に従う $X, Y \sim \text{Ber}(p)$ とする。この和を $z = x + y$ の畳み込みを行う。 $y = z - x$ を踏まえて計算すると

$$\begin{aligned} f(z) &= \sum_{x=0}^z f(x)f(z-x) \\ &= \sum_{x=0}^z p^x(1-p)^{1-x}p^{z-x}(1-p)^{1-z+x} \\ &= p^z(1-p)^{2-z} \sum_{x=0}^z \\ &= (z-0+1)p^z(1-p)^{2-z} \\ &= (z+1)p^z(1-p)^{2-z} \end{aligned}$$

となる。これはベルヌーイ分布の確率質量関数に $z = x + y$ を代入したもの

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

と異なる。この畳み込みでは二項分布となるが、確率質量関数

$$f(s) = \binom{n}{s} p^s (1-p)^{n-s}$$

とも異なる。上記の計算には誤りがあるため異なる一致しなかった。定義域は x も y も $\{0, 1\}$ だが、その和 z は $\{0, 1, 2\}$ であり、総和の引数を x とし上限を z としたところに誤りがある。またベルヌーイ分布の確率質量関数は定義域の外では零になるようにしなければならない。そのため定義関数を用いる。これは指示関数とも特性関数*1とも呼ばれる。数式の表記もいくつかあり、ここでは $I_A(x)$ と表す。これは

$$I_A(x) := \begin{cases} 1 & x \in A \\ 0 & x \notin A \end{cases}$$

と定義される。定義関数を含んだベルヌーイ分布の確率質量関数を

$$f(x) = p^x(1-p)^{1-x}I_{\{0,1\}}(x)$$

$$f(y) = p^y(1-p)^{1-y}I_{\{0,1\}}(y)$$

*1 統計学においては、この名前を特殊な期待値で用いるため、このように呼ばない。

とする。また総和の上限下限については定義関数を含めることで定義域外では零になる。ここで任意の整数 \mathbb{Z} について足し合わせる。

$$\begin{aligned}
f(z) &= \sum_{x \in \mathbb{Z}} f(x) f(z-x) \\
&= p^z (1-p)^{2-z} \sum_{x \in \mathbb{Z}} I_{\{0,1\}}(x) I_{\{0,1\}}(z-x) \\
&= p^z (1-p)^{2-z} \left(\cdots + I_{\{0,1\}}(-1) I_{\{0,1\}}(z+1) + I_{\{0,1\}}(0) I_{\{0,1\}}(z) + I_{\{0,1\}}(1) I_{\{0,1\}}(z-1) \right. \\
&\quad \left. + I_{\{0,1\}}(2) I_{\{0,1\}}(z-2) + \cdots \right) \\
&= p^z (1-p)^{2-z} \left(\cdots + 0 + 1 \times I_{\{0,1\}}(z) + 1 \times I_{\{0,1\}}(z-1) + 0 \times I_{\{0,1\}}(z-2) + 0 + \cdots \right) \\
&= p^z (1-p)^{2-z} \left(I_{\{0,1\}}(z) + I_{\{0,1\}}(z-1) \right)
\end{aligned}$$

括弧の中は z の値により異なり

$$\begin{aligned}
I_{\{0,1\}}(0) + I_{\{0,1\}}(-1) &= 1 + 0 = 1 & (z=0) \\
I_{\{0,1\}}(1) + I_{\{0,1\}}(0) &= 1 + 1 = 2 & (z=1) \\
I_{\{0,1\}}(2) + I_{\{0,1\}}(1) &= 0 + 1 = 1 & (z=2)
\end{aligned}$$

となる。これは二項係数 $\binom{2}{z}$ と同じ値である。そのため

$$f(z) = \binom{2}{z} p^z (1-p)^{2-z}$$

である。二項係数は下側の添え字が零より下になるときに、上側の添え字より大きくなるときに零になる。そのためこの式の二項係数は定義関数を用いて

$$\binom{2}{z} I_{\{0,1,2\}}(z)$$

と書いても同じ値を出す。定義関数を用いて

$$f(z) = \binom{2}{z} p^z (1-p)^{2-z} I_{\{0,1,2\}}(z)$$

とも書ける。

二項分布で $n=1$ とするとベルヌーイ分布になる。

$$\text{Bin}(s|p, 1) = \text{Ber}(s|p)$$

上記の畳み込みを $n=1$ の二項分布とベルヌーイ分布との畳み込みと見做し、これを繰り返すことで一般の二項分布が導かれることを数学的帰納法示す。つまり独立な $X_1, X_2, \dots, X_n \sim \text{Ber}(p)$ の和 $S_n = X_1 + X_2 + \dots + X_n$ は $\{0, 1, \dots, n\}$ の範囲であり、確率質量関数は

$$\text{Bin}(s_n|p, n) = \binom{n}{s_n} p^{s_n} (1-p)^{n-s_n} I_{\{0,1,\dots,n\}}(s_n)$$

になる。\$n = 2\$ のとき \$S_2 = X_1 + X_2\$ の畳み込みは上の計算を参考にすればいい。\$n = m\$ のとき成立すると仮定する。\$n = m + 1\$ のときについては

$$\begin{aligned}
& \sum_{s_m \in \mathbb{Z}} \text{Bin}(s_m|p, m) \text{Ber}(s_{m+1} - s_m|p) \\
&= \sum_{s_m \in \mathbb{Z}} \binom{m}{s_m} p^{s_m} (1-p)^{m-s_m} \mathbb{I}_{\{0,1,\dots,m\}}(s_m) p^{s_{m+1}-s_m} (1-p)^{1-s_{m+1}+s_m} \mathbb{I}_{\{0,1\}}(s_{m+1} - s_m) \\
&= p^{s_{m+1}} (1-p)^{m+1-s_{m+1}} \sum_{s_m \in \mathbb{Z}} \binom{m}{s_m} \mathbb{I}_{\{0,1,\dots,m\}}(s_m) \mathbb{I}_{\{s_{m+1}-1, s_{m+1}\}}(s_m) \\
&= p^{s_{m+1}} (1-p)^{m+1-s_{m+1}} \left(\dots + \binom{m}{s_{m+1}-1} \mathbb{I}_{\{0,1,\dots,m\}}(s_{m+1}-1) \mathbb{I}_{\{s_{m+1}-1, s_{m+1}\}}(s_{m+1}-1) \right. \\
&\quad \left. + \binom{m}{s_{m+1}} \mathbb{I}_{\{0,1,\dots,m\}}(s_{m+1}) \mathbb{I}_{\{s_{m+1}-1, s_{m+1}\}}(s_{m+1}) + \dots \right) \\
&= p^{s_{m+1}} (1-p)^{m+1-s_{m+1}} \left(\dots + 0 + \binom{m}{s_{m+1}-1} + \binom{m}{s_{m+1}} + 0 \dots \right) \\
&= \binom{m+1}{s_{m+1}} p^{s_{m+1}} (1-p)^{m+1-s_{m+1}}
\end{aligned}$$

となる。最後の式変形はパスカルの法則を用いた。更に二項係数は定義関数を用いて

$$\binom{m+1}{s_{m+1}} \mathbb{I}_{\{0,1,\dots,m,m+1\}}(s_{m+1})$$

と書くこともできる。そのため \$n = m + 1\$ でも畳み込みは

$$\sum_{s_m \in \mathbb{Z}} \text{Bin}(s_m|p, m) \text{Ber}(s_{m+1} - s_m|p) = \text{Bin}(s_{m+1}|p, m+1)$$

となる。依って任意の整数で成立する。

畳み込みは二項分布同士でも成立し、二項分布となる。そのため再生性を満たす。改めて最初の畳み込みの誤りを考える。定義関数を用いない場合、二項分布の確率質量関数は

$$\text{Bin}(s|p, n) = \binom{n}{s} p^s (1-p)^{n-s}$$

と書ける。この式は試行回数 \$n\$ に対し成功回数 \$s\$ を求めるものである。例えば籤を十回引いて三十回当たりが出ることはあり得ないように、成功回数が試行回数を上回るような現象は無いが、あったとしても二項分布で表すことはできないだろう。そのため定義関数で範囲を指定しなければ畳み込みが正しい結果とならない。このことから分かるように任意の確率分布ではそのまま畳み込みは成立しない。一部の確率分布では定義関数を用いて畳み込みをしなければならない。但し二項分布は定義関数を用いなければ再生性を証明できないということではなく、確率母関数を経由すれば定義関数をなしでも証明できる。多項分布についても確率母関数を用いて再生性を示せ、

$$\begin{aligned}
G_{\vec{X}}(\vec{a}) &= (p_1 a_1 + \dots + p_K a_K)^n \\
G_{\vec{Y}}(\vec{a}) &= (p_1 a_1 + \dots + p_K a_K)^m
\end{aligned}$$

とすると積は

$$G_{\vec{X}}(\vec{a}) G_{\vec{Y}}(\vec{a}) = (p_1 a_1 + \dots + p_K a_K)^{n+m} = G_{\vec{X}+\vec{Y}}(\vec{a})$$

となる。

確率質量関数を用いる場合は二項分布と同じく定義関数が必要である。 $\vec{X} \sim \text{Mult}(\vec{p}, n)$ について $n = 1$ のときをカテゴリカル分布といい $\vec{Y} \sim \text{Cat}(\vec{p})$ の和 $\vec{Z} = \vec{X} + \vec{Y}$ が多項分布となることを示す。つまり二項分布の際は二項分布とベルヌーイ分布を組み込むことで示したように、多項分布にカテゴリカル分布を組み込むことで示す。定義関数を含めない確率質量関数は

$$\text{Mult}(\vec{x} | \vec{p}, n) = \binom{n}{x_1, x_2, \dots, x_K} p_1^{x_1} p_2^{x_2} \cdots p_K^{x_K}$$

である。この定義関数を考える。 \vec{X} は成功回数であり、その和は $n = x_1 + x_2 + \cdots + x_K$ を満たす。またどの次元の \vec{X} の値も零以上であり、ある値以外の全てが零のとき最大値 n となる。そのため

$$A = \{x_1 + x_2 + \cdots + x_K = n\} \cap \{0 \leq x_i \leq n\}$$

とする A の範囲で一となる定義関数 $I_A(\vec{x})$ で制限する。もう一方の確率質量関数は $\vec{Y} = \vec{Z} - \vec{X}$ より

$$\text{Cat}(\vec{y} | \vec{p}) = \text{Cat}(\vec{z} - \vec{x} | \vec{p}) = \binom{1}{z_1 - x_1, z_2 - x_2, \dots, z_K - x_K} p_1^{z_1 - x_1} p_2^{z_2 - x_2} \cdots p_K^{z_K - x_K}$$

である。多項係数は上側の引数が一であるため一である。定義域は

$$B = \{z_1 - x_1 + z_2 - x_2 + \cdots + z_K - x_K = 1\} \cap \{0 \leq z_i - x_i \leq 1\}$$

とする。これは

$$B = \{z_1 + z_2 + \cdots + z_K = n + 1\} \cap \{z_i - 1 \leq x_i \leq z_i\}$$

と書き換えられる。まず $n = 1$ のとき畳み込みを実行する。これは事実上カテゴリカル分布同士の畳み込みであり、任意の整数からなるベクトルを \vec{Z} と置くと

$$\begin{aligned} f(\vec{z}) &= \sum_{\vec{x} \in \vec{Z}} \text{Mult}(\vec{x} | \vec{p}, 1) \text{Cat}(\vec{z} - \vec{x} | \vec{p}) \\ &= p_1^{z_1} p_2^{z_2} \cdots p_K^{z_K} \sum_{\vec{x} \in \vec{Z}} I_A(\vec{x}) I_B(\vec{z} - \vec{x}) \\ &= p_1^{z_1} p_2^{z_2} \cdots p_K^{z_K} \sum_{\vec{x} \in A} I_B(\vec{z} - \vec{x}) \end{aligned}$$

この総和は範囲が

$$\begin{aligned} A &= \{x_1 + x_2 + \cdots + x_K = 1\} \cap \{0 \leq x_i \leq 1\} \\ B &= \{z_1 + z_2 + \cdots + z_K = 2\} \cap \{x_i \leq z_i \leq x_i + 1\} \end{aligned}$$

となることから

$$\sum_{\vec{x} \in A} I_B(\vec{z} - \vec{x}) = I_B \begin{pmatrix} z_1 - 1 \\ z_2 \\ \vdots \\ z_K \end{pmatrix} + I_B \begin{pmatrix} z_1 \\ z_2 - 1 \\ \vdots \\ z_K \end{pmatrix} + \cdots + I_B \begin{pmatrix} z_1 \\ z_2 \\ \vdots \\ z_K - 1 \end{pmatrix}$$

様々な z の組を代入してみる。まず $z_1 = 2, z_{i \neq 1} = 0$ のとき

$$B = \{z_1 - x_1 + z_2 - x_2 + \cdots + z_K - x_K = 1\} \cap \{0 \leq z_i - x_i \leq 1\}$$

より

$$\begin{aligned}\sum_{\vec{x} \in A} I_B(\vec{z} - \vec{x}) &= I_B \begin{pmatrix} 1 \\ 0 \\ \vdots \\ 0 \end{pmatrix} + I_B \begin{pmatrix} 2 \\ -1 \\ \vdots \\ 0 \end{pmatrix} + \cdots + I_B \begin{pmatrix} 2 \\ 0 \\ \vdots \\ -1 \end{pmatrix} \\ &= 1 + 0 + \cdots + 0 \\ &= 1\end{aligned}$$

となる。次に $z_1 = 1, z_2 = 1, z_{i \neq 1,2} = 0$ のとき

$$\begin{aligned}\sum_{\vec{x} \in A} I_B(\vec{z} - \vec{x}) &= I_B \begin{pmatrix} 0 \\ 1 \\ \vdots \\ 0 \end{pmatrix} + I_B \begin{pmatrix} 1 \\ 0 \\ \vdots \\ 0 \end{pmatrix} + \cdots + I_B \begin{pmatrix} 1 \\ 0 \\ \vdots \\ -1 \end{pmatrix} \\ &= 1 + 1 + \cdots + 0 \\ &= 2\end{aligned}$$

となる。対称性より $1 \leq a, b \leq K$ に於いて $z_a = 2, z_{i \neq a} = 0$ のとき

$$\sum_{\vec{x} \in A} I_B(\vec{z} - \vec{x}) = 1$$

となり、 $z_a = 1, z_b = 1, z_{i \neq a,b} = 0$ のとき

$$\sum_{\vec{x} \in A} I_B(\vec{z} - \vec{x}) = 2$$

となる。この結果は

$$\binom{2}{z_1, z_2, \dots, z_K} = \frac{2!}{z_1! z_2! \cdots z_K!}$$

と一致する。故に

$$\begin{aligned}f(\vec{z}) &= \sum_{\vec{x} \in \vec{Z}} \text{Mult}(\vec{x} | \vec{p}, 1) \text{Cat}(\vec{z} - \vec{x} | \vec{p}) \\ &= \binom{2}{z_1, z_2, \dots, z_K} p_1^{z_1} p_2^{z_2} \cdots p_K^{z_K} \\ &= \text{Mult}(\vec{z} | \vec{p}, 2)\end{aligned}$$

となり $n = 1$ で成り立つ。次に $n = m$ で成り立つとして $n = m + 1$ を考える。

$$\begin{aligned}f(\vec{z}) &= \sum_{\vec{x} \in \vec{Z}} \text{Mult}(\vec{x} | \vec{p}, m) \text{Cat}(\vec{z} - \vec{x} | \vec{p}) \\ &= p_1^{z_1} p_2^{z_2} \cdots p_K^{z_K} \sum_{\vec{x} \in \vec{Z}} \binom{m}{x_1, x_2, \dots, x_K} I_A(\vec{x}) I_B(\vec{z} - \vec{x}) \\ &= p_1^{z_1} p_2^{z_2} \cdots p_K^{z_K} \sum_{\vec{x} \in A} \binom{m}{x_1, x_2, \dots, x_K} I_B(\vec{z} - \vec{x})\end{aligned}$$

となる。このとき定義域 B は

$$B = \{z_1 - x_1 + z_2 - x_2 + \cdots + z_K - x_K = 1\} \cap \{z_i - 1 \leq x_i \leq z_i\}$$

となる。これを満たす定義関数の領域は

$$f(\vec{z}) = p_1^{z_1} p_2^{z_2} \cdots p_K^{z_K} \left\{ \binom{m}{z_1 - 1, z_2, \dots, z_K} + \binom{m}{z_1, z_2 - 1, \dots, z_K} + \cdots + \binom{m}{z_1, z_2, \dots, z_K - 1} \right\}$$

である。この括弧の中の式は一般化したパスカルの法則より

$$\binom{m+1}{z_1, z_2, \dots, z_K}$$

となる。故に

$$f(\vec{z}) = \sum_{\vec{x} \in \vec{Z}} \text{Mult}(\vec{x} | \vec{p}, m) \text{Cat}(\vec{z} - \vec{x} | \vec{p}) = \text{Mult}(\vec{z} | \vec{p}, m+1)$$

となり $n = m + 1$ で組み込み $\vec{Z} \sim \text{Mult}(\vec{p}, m+1)$ が成り立つ。

ポアソン分布

定義関数を用いた証明は離散分布に関しては二項分布、多項分布などがある。定義関数が必要ないものとしてポアソン分布、負の二項分布などがある。ポアソン分布は二項分布の n の極大化

$$\lim_{n \rightarrow \infty} \text{Bin}(s|p, n) = \text{Po}(s|\lambda)$$

として得られることから、二項分布で想定した成功回数が試行回数を上回るようなことは起こらない。 $X_1 \sim \text{Po}(\lambda_1)$ と $X_2 \sim \text{Po}(\lambda_2)$ との和 $Z = X_1 + X_2$ を畳み込む。ポアソン分布の確率質量関数は

$$f(s) = \text{Po}(s|\lambda) = \frac{\lambda^s}{s!} \exp(-\lambda)$$

であることから

$$\begin{aligned} f(z) &= \sum_{x_1=0}^z \text{Po}(x_1|\lambda_1) \text{Po}(z-x_1|\lambda_2) \\ &= \sum_{x_1=0}^z \frac{\lambda_1^{x_1}}{x_1!} \exp(-\lambda_1) \frac{\lambda_2^{z-x_1}}{(z-x_1)!} \exp(-\lambda_2) \\ &= \frac{1}{z!} \exp(-(\lambda_1 + \lambda_2)) \sum_{x_1=0}^z \frac{z!}{x_1!(z-x_1)!} \lambda_1^{x_1} \lambda_2^{z-x_1} \\ &= \frac{(\lambda_1 + \lambda_2)^z}{z!} \exp(-(\lambda_1 + \lambda_2)) \\ &= \text{Po}(z|\lambda_1 + \lambda_2) \end{aligned}$$

となり、 $\lambda_1 + \lambda_2 = \lambda$ とすると

$$f(z) = \text{Po}(z|\lambda) = \frac{\lambda^z}{z!} \exp(-\lambda)$$

となり、これは元のポアソン分布の確率質量関数であり再生性を満たす。

スケラム分布

上記はポアソン分布の和 $Z = X_1 + X_2$ の畳み込みをした。ここでは差 $K = X_2 - X_1$ の畳み込みをする。

$$\begin{aligned} f(z) &= \sum_{x_1=0}^z \text{Po}(x_1|\lambda_1)\text{Po}(k+x_1|\lambda_2) \\ &= \sum_{x_1=0}^z \frac{\lambda_1^{x_1}}{x_1!} \exp(-\lambda_1) \frac{\lambda_2^{k+x_1}}{(k+x_1)!} \exp(-\lambda_2) \\ &= \exp(-(\lambda_1 + \lambda_2)) \sum_{x_1=0}^z \frac{\lambda_1^{x_1} \lambda_2^{k+x_1}}{x_1!(k+x_1)!} \\ &= \exp(-(\lambda_1 + \lambda_2)) \left(\frac{\lambda_1}{\lambda_2}\right)^{\frac{k}{2}} \sum_{x_1=0}^z \frac{1}{x_1!(k+x_1)!} (\lambda_1 \lambda_2)^{x_1 + \frac{k}{2}} \\ &= \exp(-(\lambda_1 + \lambda_2)) \left(\frac{\lambda_1}{\lambda_2}\right)^{\frac{k}{2}} I_k(2\sqrt{\lambda_1 \lambda_2}) \end{aligned}$$

負の二項分布

ここでは幾何分布も負の二項分布も失敗回数を確率変数としている。幾何分布の確率変数の和 $l + \lambda = x$ で畳み込みをする。幾何分布の確率質量関数は $\text{Ge}(l) = q^l p^1$ であるため

$$\begin{aligned} f(x) &= \sum_{l=0}^x \text{Ge}(l)\text{Ge}(x-l) \\ &= \sum_{l=0}^x q^l p^1 q^{x-l} p^1 \\ &= q^x p^2 (x - 0 + 1) \\ &= \binom{x+2-1}{1} q^x p^2 \end{aligned}$$

である。負の二項分布の確率質量関数は

$$\text{NB}(x) = \binom{x+s-1}{s-1} q^x p^s$$

であり、この式で $s = 2$ とした式は上の畳み込みの式に等しい。また $s = 1$ のとき幾何分布の確率質量関数と等しいことから、負の二項分布は幾何分布の畳み込みによって生じると考えられる。数学的帰納法を用いて $\text{Ge}(l) = \text{NB}(l|1)$ の畳み込みが $\text{NB}(l|s)$ になることを示す。つまり

$$f(x) = \sum_{l=0}^x \text{NB}(l|s)\text{Ge}(x-l) = \text{NB}(x|s+1)$$

が任意の s で成り立つことを示す。 $s = 1$ は上記を参考にすればいい。 $s = k$ で成り立つと仮定する。このとき

$$\begin{aligned} f(x) &= \sum_{l=0}^x \text{NB}(l|k)\text{Ge}(x-l) \\ &= \sum_{l=0}^x \binom{l+k-1}{k-1} q^l p^k q^{x-l} p^1 \\ &= \left\{ \binom{k-1}{k-1} + \binom{1+k-1}{k-1} + \dots + \binom{x+k-1}{k-1} \right\} q^x p^{k+1} \\ &= \binom{x+k}{k} q^x p^{k+1} \\ &= \text{NB}(x|k+1) \end{aligned}$$

となる。途中アイスホッケーの公式を用いた。^{*2}そのため任意の s で成立する。

^{*2} 参考 https://yakumido.blogspot.com/2014/05/blog-post_9044.html

連続確率分布

安定分布

二項分布の節でベルヌーイ分布から二項分布を導いた。逆に二項分布からベルヌーイ分布を導くことはできるだろうか。