2022 年 08 月 25 日

# たたみこみとポアソン分布と指数分布

新潟工科大学 基礎教育・教養系 竹野茂治

## 1 はじめに

講義では、ポアソン分布と指数分布が「表裏」の関係にあることを簡単に説明し、証明らしきものも紹介したが、証明を略した部分もあった。

本稿では、その証明の補足もかねて、

- たたみこみ (離散分布、連続分布) の定義、性質といくつかの例
- ポアソン分布のスケール不変性
- ポアソン分布の裏が指数分布となること
- 指数分布の裏がポアソン分布となること

について説明する。

なお、本稿では現代的な公理的確率論ではなく、古典的確率論の範疇で考える。

## 2 離散確率分布のたたみこみ

本稿では、離散確率分布 $P$ を、標本空間 $\Omega$ と、その上の確率関数 $p(x):\Omega\to[0,1]$ をセットにして、$P=(\Omega,p)$ のように表す。なお、$\Omega$ としては、基本的に $\mathrm{Z}_+=\{0,1,2,\ldots\}$ を考える。

離散確率分布 $P=(\mathrm{Z}_+,p), Q=(\mathrm{Z}_+,q)$ のたたみこみを紹介する。確率変数 $x\sim P$, $y\sim Q$ が独立であるとみて、その和 $z=x+y$ を考える。独立なので、2 次元確率関数 $r(x,y)$ は $r(x,y)=p(x)q(y)$ となり、その写像 $(x,y)\to x+y$ として $z=x+y$ が定まるが ([1])、この $z=x+y$ の確率関数 $s(z)$ は、$x$ と $y$ の独立性より、

$$s(n) \;=\; \mathrm{Prob}\{z=n\} \;=\; \mathrm{Prob}\{x+y=n\}$$

$$
\begin{aligned}
&= \sum_{k=0}^{n} \mathrm{Prob}\{x=k \text{ かつ } y=n-k\} \;=\; \sum_{k=0}^{n} r(k,n-k) \\
&= \sum_{k=0}^{n} p(k)q(n-k)
\end{aligned}
$$

となる。

このように、2 つの確率分布 $P=(\mathrm{Z}_+,p)$, $Q=(\mathrm{Z}_+,q)$ に対して、

$$
(p*q)(n) = \sum_{k=0}^{n} p(k)q(n-k) \quad (n \in \mathrm{Z}_+) \tag{1}
$$

で決まる関数 $p*q$ を確率関数とする確率分布を $P, Q$ の「たたみこみ」といい、$P*Q$ と書く。なお、$p*q$ の値は当然 0 以上であり、

$$
\begin{aligned}
\sum_{n=0}^{\infty}(p*q)(n) &= \sum_{n=0}^{\infty}\sum_{k=0}^{n} p(k)q(n-k) \;=\; \sum_{k=0}^{\infty}\sum_{n=k}^{\infty} p(k)q(n-k) \\
&= \sum_{k=0}^{\infty} p(k) \sum_{m=0}^{\infty} q(m) \;=\; 1
\end{aligned}
$$

となるので、確かに $p*q$ は $\mathrm{Z}_+$ 上の確率関数となる。たたみこみ $P*Q$ は、上に見たように $P, Q$ に従う独立な確率変数の和が従う確率分布、となる。

たたみこみに関しては、以下が成り立つ。

## 命題 1

1. $p*q = q*p$
2. $(p*q)*r = p*(q*r)$
3. $\mathrm{Z}_+$ 上の確率関数 $p, q$ に対して、$p(0)>0$ のとき、$p*r=q$ となる $\mathrm{Z}_+$ 上の関数 $r$ が一意に決定する (が、確率関数になるとは限らない)

証明

1. $m = n - k$ とすると、

$$(p * q)(n) = \sum_{k=0}^{n} p(k)q(n-k) = \sum_{m=0}^{n} p(n-m)q(m) = (q * p)(n)$$

となり成り立つ。

2.

$$\begin{aligned}((p * q) * r)(n) &= \sum_{k=0}^{n}(p * q)(k)r(n-k) = \sum_{k=0}^{n} r(n-k)\sum_{j=0}^{k} p(j)q(k-j) \\ &= \sum_{j=0}^{n} p(j)\sum_{k=j}^{n} q(k-j)r(n-k) = \sum_{j=0}^{n} p(j)\sum_{m=0}^{n-j} q(m)r(n-j-m) \\ &= \sum_{j=0}^{n} p(j)(q * r)(n-j) = (p * (q * r))(n)\end{aligned}$$

3. $q(0) = p(0)r(0)$ なので、$p(0) > 0$ より $r(0)$ は $r(0) = q(0)/p(0)$ と一意に決定する。

また、$r(k)$ が $k = 0, 1, \ldots, n-1$ まで決定したとすると、

$$q(n) = \sum_{k=0}^{n} p(k)r(n-k) = p(0)r(n) + p(1)r(n-1) + \cdots + p(n)r(0)$$

より、

$$r(n) = \frac{q(n)}{p(0)} - \frac{1}{p(0)}\left(p(1)r(n-1) + \cdots + p(n)r(0)\right)$$

によって $r(n)$ も一意に決定するから、$r$ が一意に決定することになる。

ただし、$r(n)$ の値が負にならないとは言えないので、$r$ が確率関数になるとは限らない。■

この命題 1 の 2. より、$x_i \sim P_i = (\mathrm{Z}_+, p_i)$ $(1 \le i \le n)$ に対するたたみこみ $P_1 * \cdots * P_n$ を考えることもできる。これは、順にたたみこんだものであるが、

$$(p_1 * \cdots * p_n)(m) = \sum_{k_1=0}^{m} p_1(k_1)(p_2 * \cdots * p_n)(m-k_1)$$

$$\begin{aligned}
&= \sum_{k_1=0}^{m} p_1(k_1) \sum_{k_2=0}^{m-k_1} p_2(k_2)(p_3 * \cdots * p_n)(m-k_1-k_2) \;=\; \ldots \\
&= \sum_{k_1=0}^{m} p_1(k_1) \sum_{k_2=0}^{m-k_1} p_2(k_2) \cdots \sum_{k_{n-1}=0}^{m-k_1-\cdots-k_{n-2}} p_{n-1}(k_{n-1}) \\
&\quad \times p_n(m-k_1-\cdots-k_{n-2}-k_{n-1}) \\
&= \sum_{i_1+\ldots+i_n=m} p_1(i_1)p_2(i_2)\cdots p_n(i_n)
\end{aligned}$$

と書ける。これは、$x_1,\ldots,x_n$ が独立であるとみたときの $z=x_1+\ldots+x_n$ が $m$ に等しい確率になるので、$P_1*\cdots*P_n$ はこの和の確率分布となる。

また、この $P_j$ がすべて $P=(\mathrm{Z}_+,p)$ に等しい場合は、それを本稿では $P^{(n)}=(\mathrm{Z}_+,p^{(n)})$ と書くことにする。命題 1 の 3. と同様、これに対しても次が成り立つ。

## 命題 2

$\mathrm{Z}_+$ 上の確率関数 $q$ と $n\geq 2$ に対して、$q(0)>0$ のとき、$p^{(n)}=q$ となる $\mathrm{Z}_+$ 上の関数 $p$ が一意に決定する (が、確率関数になるとは限らない)

証明

$q(0)=p(0)^n$ より、$q(0)>0$ であれば $p(0)=q(0)^{1/n}$ と一意に $p(0)$ が決定する。

また、$p(k)$ が $k=1,\ldots,m-1$ まで決定したとすると、

$$q(m) = \sum_{i_1+\cdots+i_n=m} p(i_1)\cdots p(i_n) = np(0)^{n-1}p(m) + \sum_{i_1+\cdots+i_n=m, i_k<n} p(i_1)\cdots p(i_n)$$

となるので、$p(m)$ は、

$$p(m) = \frac{q(m)}{np(0)^{n-1}} - \frac{1}{np(0)^{n-1}} \sum_{i_1+\cdots+i_n=m, i_k<n} p(i_1)\cdots p(i_n)$$

により $p(m)$ が一意に決定する。

ただし、$p(m)$ の値が負にならないとは言えないので、$p$ が確率関数になるとは限らない。$\blacksquare$

ここで、たたみこみの例を 2,3 紹介する。

## 例 3

2 項分布 $B(m,r)*B(n,r)$ のたたみこみを計算する。2 項分布 $B(m,r)$ の確率関数を

$$p_{m,r}(x) = \begin{pmatrix} m \\ x \end{pmatrix} r^x(1-r)^{m-x} \quad (0<r<1)$$

とする。ただし、$x<0,$ または $x>m$ では

$$\begin{pmatrix} m \\ x \end{pmatrix} = 0$$

と考え、標本空間は $\mathrm{Z}_+$ とする。

$$\begin{aligned}(p_{m,r}*p_{n,r})(k) &= \sum_{j=0}^{k} p_{m,r}(j)p_{n,r}(k-j) \\ &= \sum_{j=0}^{k} \begin{pmatrix} m \\ j \end{pmatrix} r^j(1-r)^{m-j} \begin{pmatrix} n \\ k-j \end{pmatrix} r^{k-j}(1-r)^{n-k+j} \\ &= r^k(1-r)^{m+n-k}\sum_{j=0}^{k} \begin{pmatrix} m \\ j \end{pmatrix} \begin{pmatrix} n \\ k-j \end{pmatrix}\end{aligned}$$

となるが、

$$\sum_{j=0}^{k} \begin{pmatrix} m \\ j \end{pmatrix} \begin{pmatrix} n \\ k-j \end{pmatrix}$$

は、$m+n$ 個の玉のうち、$k$ 個が白玉、$(m+n-k)$ 個が赤玉であるときにそれを 1 列に並べたときの並べかえの個数

$$\begin{pmatrix} m+n \\ k \end{pmatrix}$$

に等しい。実際、その 1 列の並びの最初の $m$ 個と後の $n$ 個に分けて考えると、その並び換えの総数は、最初の $m$ 個に $j$ 個の白玉が含まれ、後の $n$ 個に $(k-j)$

個の白玉が含まれるときの組み合わせの総数を、$j$ に関して加えたものになっているからである。

よって、

$$p_{m,r} * p_{n,r}(k) = \begin{pmatrix} m+n \\ k \end{pmatrix} r^k (1-r)^{m+n-k} = p_{m+n,r}$$

となり、よって $B(m,r) * B(n,r) = B(m+n,r)$ となることがわかる。一般に、

$$B(n_1,r) * \cdots * B(n_k,r) = B(n_1 + \cdots + n_k, r)$$

となる。

## 例 4

ポアソン分布 $P(\lambda) * P(\mu)$ のたたみこみを計算する。ポアソン分布 $P(\lambda)$ の確率関数を

$$p_\lambda(x) = \frac{\lambda^x}{x!} e^{-\lambda}$$

とする。

$$\begin{aligned}(p_\lambda * p_\mu)(n) &= \sum_{k=0}^{n} p_\lambda(k) p_\mu(n-k) = \sum_{k=0}^{n} \frac{\lambda^k}{k!} \frac{\mu^{n-k}}{(n-k)!} e^{-\lambda-\mu} \\ &= e^{-\lambda-\mu} \frac{1}{n!} \sum_{k=0}^{n} \begin{pmatrix} n \\ k \end{pmatrix} \lambda^k \mu^{n-k} = e^{-\lambda-\mu} \frac{(\lambda+\mu)^n}{n!} = p_{\lambda+\mu}(n)\end{aligned}$$

となるので、$P(\lambda) * P(\mu) = P(\lambda+\mu)$ がわかる。一般に、

$$P(\lambda_1) * \cdots * P(\lambda_n) = P(\lambda_1 + \cdots + \lambda_n)$$

となる。

# 3 ポアソン分布のスケール不変性

次は、ポアソン分布の裏が指数分布となることを示す際に使われる、ポアソン分布のスケール不変性を示す。具体的には、以下の通り。

> 1 時間の間にある事象が独立にいくつか起きるときに、その事象の起きる回数 $x$ がポアソン分布 $P(\lambda) = (\mathrm{Z}_+, p_\lambda)$ に従うとすると、$T$ 時間 $(T > 0)$ にその事象が起きる回数 $y_T$ は、$P(\lambda T)$ に従う。

実際、講義ではこれを証明なしに認めた上で、ポアソン分布の裏が指数分布となることを簡単に紹介した。

本節では、これをいくつかの段階に分けて示していく。まず、求めるべき $y_T$ の従う分布を $Q(T) = (\mathrm{Z}_+, q_T)$ と書くことにする。目標は、$Q(T) = P(\lambda T)$ を示すことである。

$T$ が正の整数 $T = m$ の場合、最初の 1 時間に起きる回数を $x_1$, 次の 1 時間に起きる回数を $x_2$, ..., 最後の 1 時間に起きる回数を $x_m$ とすれば、$m$ 時間で起きる回数は $y_m = x_1 + \cdots + x_m$ となるので、$x_1, \ldots, x_m$ の独立性と、2 節の例 4 より

$$y_m \sim P(\lambda) * \cdots * P(\lambda) = P(\lambda)^{(m)} = P(\lambda m)$$

となり、よって $T = m$ の場合には $Q(m) = P(\lambda m)$ が成り立つ。

次は、$T = n/m$ の有理数の場合を考える。この場合は、逆に $Q(n/m)$ を $m$ 回繰り返せば $n$ 時間の回数となるので、

$$Q\left(\frac{n}{m}\right)^{(m)} = P(\lambda n)$$

となることになる。一方、例 4 より、当然

$$P\left(\frac{\lambda n}{m}\right)^{(m)} = P(\lambda n)$$

であるが、命題 2 より、そのような分布は一意に決まるので、よって $Q(n/m) = P(\lambda n/m)$ が成り立つ。

最後は実数 (無理数) $T$ の場合であるが、$0 < T_1 < T_2$ に対しては、当然 $T_1$ 時間に起きる回数よりも $T_2$ 時間に起きる回数の方が多いので、$T_1$ 時間に $n$ 回以上

起きる確率よりも、$T_2$ 時間に $n$ 回以上起きる確率の方が多くなる。なお、「$T_1$ 時間に丁度 $n$ 回起きる確率よりも、$T_2$ 時間に丁度 $n$ 回起きる確率の方が多くなる」とは言えない。よって、$0 < T_1 < T_2 < T_3$ と 0 以上の任意の整数 $n$ に対して

$$\sum_{k=n}^{\infty} q_{T_1}(k) \leq \sum_{k=n}^{\infty} q_{T_2}(k) \leq \sum_{k=n}^{\infty} q_{T_3}(k)$$

となる。ここで、$T_1, T_3$ を有理数 $T_1 = n_1/m_1, T_3 = n_2/m_2$ に取り、$T_2 = T$ とすると、$q_{T_1} = p_{\lambda n_1/m_1}, q_{T_3} = p_{\lambda n_2/m_2}$ となるので、

$$\sum_{k=n}^{\infty} p_{\lambda n_1/m_1}(k) \leq \sum_{k=n}^{\infty} q_T(k) \leq \sum_{k=n}^{\infty} p_{\lambda n_2/m_2}(k)$$

ここで、$n_1/m_1 \to T, n_2/m_2 \to T,$ すなわち有理数の値を取りながら $T$ に近付けていくと、

$$\sum_{k=n}^{\infty} p_{\lambda n_j/m_j}(k) = 1 - \sum_{k=0}^{n-1} \frac{(\lambda n_j/m_j)^k}{k!} e^{-\lambda n_j/m_j}$$

の有限和の形に書けるので、極限は $n_j/m_j$ を $T$ に置きかえたものになり、よってはさみうちの原理により

$$\sum_{k=n}^{\infty} q_T(k) = \sum_{k=n}^{\infty} p_{\lambda T}(k)$$

が成り立つ。これは、1 から引けば、

$$\sum_{k=0}^{n-1} q_T(k) = \sum_{k=0}^{n-1} p_{\lambda T}(k)$$

となるから、$n = 1, 2, \ldots$ と順番に代入していけば、すべての $n$ に対して $q_T(n) = p_{\lambda T}(n)$ となることがわかる。

よって、すべての正の実数 $T$ に対して、$Q(T) = P(\lambda T)$ となることが言えたことになる。

# 4 ポアソン分布の裏が指数分布となること

次は、講義でも紹介した、ポアソン分布 $P(\lambda)$ の「裏」が指数分布 $e(1/\lambda)$ となることを考える。指数分布 $e(1/\lambda)$ とは、

$$F_\lambda(d) = \begin{cases} 1 - e^{-\lambda d} & (d > 0) \\ 0 & (d < 0) \end{cases} \tag{2}$$

を分布関数とするような連続確率分布である。密度関数 $f_\lambda(d)$ は、

$$f_\lambda(d) = \begin{cases} \lambda e^{-\lambda d} & (d > 0) \\ 0 & (d < 0) \end{cases} \tag{3}$$

となる。これが、ポアソン分布の「裏」とはどういうことかを説明するが、この表現は講義の教科書 [3] によるもので、一般的な表現かどうかは知らない。

> 3 節で述べたように、ポアソン分布 $P(\lambda)$ が、1 時間の間にある事象が起きる回数の分布になっているとき、その事象と次に起きた事象の間の時間 $d$ は、指数分布 $e(1/\lambda)$ に従う。

実は教科書はこの逆の形で書いてあるのであるが、この形のものもどこかの本に書いてあった (図書館にあった本だと思うが、どの本かは忘れてしまった)。

これについて講義では、その本に書いてあった、以下のような説明を行った。

> $T > 0$ に対して、時間間隔 $d$ が $T$ よりも大きい確率 $\mathrm{Prob}\{d > T\} = 1 - F_\lambda(T)$ は、$T$ 時間の間に 1 回も起こらない確率に等しい。$T$ 時間の間に起きる回数 $x_T$ は $P(\lambda T)$ に従う (前節の結果) ので、
>
> $$\mathrm{Prob}\{d > T\} = 1 - F_\lambda(T) = \mathrm{Prob}\{x_T = 0\} = p_{\lambda T}(0)$$
>
> となり、よって
>
> $$F_\lambda(T) = 1 - p_{\lambda T}(0) = 1 - \frac{(\lambda T)^0}{0!}e^{-\lambda T} = 1 - e^{-\lambda T}$$
>
> となる。

ただ、この説明では若干 $\mathrm{Prob}\{d > T\} = \mathrm{Prob}\{x_T = 0\}$ のところが気になる。むしろ、これは、次のように考えた方がわかりやすいような気がする。

最後にその事象が起きた時刻を 0 として、そこから $T$ 時間までの間に 1 回でも起きれば、この $T$ 時間の間には少なくとも 1 回以上起きることになるが、これは、最初の事象までの時間間隔で考えれば、$d \leq T$ を意味し、一方起きる回数 $x_T$ で言えば $x_T \geq 1$ を意味する。よって、$\mathrm{Prob}\{d \leq T\} = \mathrm{Prob}\{x_T \geq 1\}$ となり、これを 1 から引けば、$\mathrm{Prob}\{d > T\} = \mathrm{Prob}\{x_T = 0\}$ となる。

あとは上の計算と同じである。「最後にその事象が起きた時刻を 0」とすることで、どこに起点を置いて考えるか、どの時間間隔を見るかが明確になって、わかりやすくなるように思う。そしてこれなら $\mathrm{Prob}\{d > T\} = \mathrm{Prob}\{x_T = 0\}$ もある程度納得できると思う。

# 5 連続分布のたたみこみ

前節で、ポアソン分布の裏が指数分布であることを示したが、そこにも述べたように教科書 [3] には、実は「指数分布の裏がポアソン分布」と書いてある。次はそれを考えるために、まずは連続分布のたたみこみについて説明する。

本稿では、連続確率分布 $P$ を、標本空間 R と、分布関数 $F(x)$ をセットにして、$P = (\mathrm{R}, F)$ のように表す。密度関数 $f(x)$ は $f(x) = F'(x)$ である。

この場合も、独立な確率変数の和の分布で考える。$x \sim P = (\mathrm{R}, F)$, $y \sim Q = (\mathrm{R}, G)$ のとき、$x, y$ を独立として考えた 2 次元確率変数 $(x, y)$ の分布関数は $H(x, y) = F(x)G(y)$ で、密度関数は $h(x, y) = f(x)g(y)$ となる ($f = F'$, $g = G'$)。よって、$z = x + y$ は、以下の $K(z)$ を分布関数とする確率変数となる。

$$\begin{aligned} K(t) &= \mathrm{Prob}\{x + y \leq t\} = \iint_{\{x+y\leq t\}} h(x, y)dxdy \\ &= \int_{\mathrm{R}} dx \int_{-\infty}^{t-x} f(x)g(y)dy = \int_{\mathrm{R}} f(x)G(t - x)dx \end{aligned} \tag{4}$$

よって、$z$ の密度関数は、これを $t$ で微分した

$$k(t) = K'(t) = \int_{\mathrm{R}} f(x)g(t - x)dx \tag{5}$$

となる。逆に、$P = (\mathrm{R}, F)$,$Q = (\mathrm{R}, G)$ から (5) によって作った $k(t)$ $(= (f * g)(t)$ と書く$)$ を密度関数とするような連続確率分布を、$P$ と $Q$ の「たたみこみ」と呼び、$P * Q$ と書く。

なお、$f \geq 0,\, g \geq 0$ で $f, g \in L^1(\mathrm{R})$ なので $f * g \geq 0$ で、

$$\begin{aligned}\int_{\mathrm{R}} (f*g)(t)dt &= \int_{\mathrm{R}} dt \int_{\mathrm{R}} f(x)g(t-x)dx = \int_{\mathrm{R}} dx \int_{\mathrm{R}} f(x)g(t-x)dt \\ &= \int_{\mathrm{R}} f(x)dx \int_{\mathrm{R}} g(y)dy = 1\end{aligned}$$

となるので、確かに $f * g$ はある連続分布の密度関数となる。

また、(4) より、$P * Q = (\mathrm{R}, f * G)$ となるが、これは次の命題 5 により $P * Q = (\mathrm{R}, F * g)$ とも書ける。

## 命題 5

1. $f * g = g * f$
2. $(f * g) * h = f * (g * h)$

証明

1.

$$(f*g)(t) = \int_{\mathrm{R}} f(x)g(t-x)dx = \int_{\mathrm{R}} f(t-y)g(y)dx = (g*f)(t)$$

より O.K.

2.

$$\begin{aligned}((f*g)*h)(t) &= \int_{\mathrm{R}} (f*g)(x)h(t-x)dx = \int_{\mathrm{R}} h(t-x)dx \int_{\mathrm{R}} f(y)g(x-y)dy \\ &= \int_{\mathrm{R}} f(y)dy \int_{\mathrm{R}} g(x-y)h(t-x)dx = \int_{\mathrm{R}} f(y)dy \int_{\mathrm{R}} g(z)h(t-y-z)dz \\ &= \int_{\mathrm{R}} f(y)(g*h)(t-y)dy = (f*(g*h))(t)\end{aligned}$$

より O.K. ■

この命題 5 より、離散の場合同様、$x_i \sim P_i = (\mathrm{R}, F_i)$ $(1 \le i \le n)$ に対するたたみこみ $P_1 * \cdots * P_n$ を考えることもできる。

$$
\begin{aligned}
(f_1 * \cdots * f_n)(t) &= \int_{\mathrm{R}} f_1(x_1)(f_2 * \cdots * f_n)(t - x_1)dx_1 \\
&= \int_{\mathrm{R}} f_1(x_1)dx_1 \int_{\mathrm{R}} f_2(x_2)(f_3 * \cdots * f_n)(t - x_1 - x_2)dx_2 = \ldots \\
&= \int_{\mathrm{R}} f_1(x_1)dx_1 \cdots \int_{\mathrm{R}} f_{n-1}(x_{n-1}) f_n(t - x_1 - \cdots - x_{n-1})dx_{n-1}
\end{aligned}
$$

なので、$t$ に関して $-\infty$ から $y$ まで積分すると、

$$
\begin{aligned}
&\int_{-\infty}^{y} (f_1 * \cdots * f_n)(t)dt \\
&= \int_{\mathrm{R}} f_1(x_1)dx_1 \cdots \int_{\mathrm{R}} f_{n-1}(x_{n-1})dx_{n-1} \int_{-\infty}^{y} f_n(t - x_1 - \cdots - x_{n-1})dt \\
&= \int_{\mathrm{R}} f_1(x_1)dx_1 \cdots \int_{\mathrm{R}} f_{n-1}(x_{n-1})dx_{n-1} \int_{-\infty}^{y - x_1 - \cdots - x_n} f_n(x_n)dx_n \\
&= \int \cdots \int_{\{x_1 + \cdots + x_n \le y\}} f_1(x_1) \cdots f_n(x_n) d\overrightarrow{x}
\end{aligned}
$$

となり、これは、$x_1, \ldots, x_n$ を独立と見た場合の $z = x_1 + \ldots + x_n$ の分布関数にほかならない。よって、$f_1 * \cdots * f_n$ はその $z$ の密度関数となる。

離散の場合と同様に、$P_j$ がすべて $P = (\mathrm{Z}_+, F)$ に等しい場合、本稿では $P$ の $n$ 重のたたみこみを $P^{(n)} = (\mathrm{Z}_+, F^{(n)})$ と書く。$n$ 階導関数ではないので注意すること。

## 例 6

正規分布 $N(\mu, \sigma^2)$ のたたみこみを計算する。正規分布 $N(\mu, \sigma^2)$ の密度関数を、

$$
f_{\mu,\sigma}(x) = \frac{1}{\sqrt{2\pi}\,\sigma} e^{-(x-\mu)^2/(2\sigma^2)}
$$

とする。このとき、

$$
(f_{\mu_1,\sigma_1} * f_{\mu_2,\sigma_2})(t) = \frac{1}{2\pi\sigma_1\sigma_2} \int_{\mathrm{R}} e^{-(x-\mu_1)^2/(2\sigma_1^2) - (t-x-\mu_2)^2/(2\sigma_2^2)} dx
$$

となるが、$x-\mu_1=y, t-x-\mu_2=t-\mu_1-\mu_2-y=a-y$ とすると、

$$\begin{aligned}\frac{(x-\mu_1)^2}{\sigma_1^2}+\frac{(t-x-\mu_2)^2}{\sigma_2^2} &= \frac{y^2}{\sigma_1^2}+\frac{(a-y)^2}{\sigma_2^2} = \frac{\sigma_1^2+\sigma_2^2}{\sigma_1^2\sigma_2^2}y^2-\frac{2a}{\sigma_2^2}y+\frac{a^2}{\sigma_2^2}\\ &= \frac{\sigma_1^2+\sigma_2^2}{\sigma_1^2\sigma_2^2}\left(y-\frac{\sigma_1^2 a}{\sigma_1^2+\sigma_2^2}\right)^2-\frac{\sigma_1^2a^2}{\sigma_2^2(\sigma_1^2+\sigma_2^2)}+\frac{a^2}{\sigma_2^2}\\ &= \frac{\sigma_1^2+\sigma_2^2}{\sigma_1^2\sigma_2^2}\left(y-\frac{\sigma_1^2 a}{\sigma_1^2+\sigma_2^2}\right)^2+\frac{a^2}{\sigma_1^2+\sigma_2^2}\end{aligned}$$

となるので、$b>0$ に対し、

$$\int_{\mathrm{R}} e^{-bx^2/2}dx=\int_{\mathrm{R}} e^{-t^2}dt\sqrt{\frac{2}{b}}=\sqrt{\frac{2\pi}{b}}$$

より、

$$\begin{aligned}(f_{\mu_1,\sigma_1}*f_{\mu_2,\sigma_2})(t) &= \frac{1}{2\pi\sigma_1\sigma_2}\sqrt{\frac{2\pi\sigma_1^2\sigma_2^2}{\sigma_1^2+\sigma_2^2}}\,e^{-a^2/(2(\sigma_1^2+\sigma_2^2))}\\ &= \frac{1}{\sqrt{2\pi}\sqrt{\sigma_1^2+\sigma_2^2}}\,e^{-(t-\mu_1-\mu_2)^2/(2(\sigma_1^2+\sigma_2^2))} = f_{\mu_1+\mu_2,\sqrt{\sigma_1^2+\sigma_2^2}}(t)\end{aligned}$$

となる。よって、

$$N(\mu_1,\sigma_1^2)*N(\mu_2,\sigma_2^2)=N(\mu_1+\mu_2,\sigma_1^2+\sigma_2^2)$$

がわかる。左辺は、$x_1+x_2$ の分布であるから、これは丁度 [2] の考察に対応する。

## 例 7

指数分布 $e(1/\lambda)$ のたたみこみを計算する。$e(1/\lambda)$ の密度関数 $f_\lambda$ は (3) の形なので、$z<0$ なら $(f_\lambda*f_\mu)(z)=0$ であり、$z>0$ なら

$$\begin{aligned}(f_\lambda*f_\mu)(z) &= \int_0^z f_\lambda(t)f_\mu(z-t)dt = \lambda\mu\int_0^z e^{-\lambda t}e^{-\mu(z-t)}dt\\ &= \lambda\mu e^{-\mu z}\int_0^z e^{(\mu-\lambda)t}dt\end{aligned}$$

より、$\lambda = \mu$ なら

$$(f_\lambda * f_\lambda)(z) = f_\lambda^{(2)}(z) = \lambda^2 z e^{-\lambda z} \quad (z > 0)$$

となり、$\lambda \neq \mu$ なら

$$(f_\lambda * f_\mu)(z) = \lambda\mu e^{-\mu z}\left[\frac{e^{(\mu-\lambda)t}}{\mu-\lambda}\right]_0^z = \frac{\lambda\mu}{\mu-\lambda}(e^{-\lambda z} - e^{-\mu z}) \quad (z > 0)$$

となる。よって、

$$(f_\lambda * f_\mu)(z) = \begin{cases} \lambda^2 z e^{-\lambda z}\chi_+(z) & (\lambda = \mu) \\ \lambda\mu\,\dfrac{e^{-\lambda z} - e^{-\mu z}}{\mu - \lambda}\chi_+(z) & (\lambda \neq \mu) \end{cases}$$

となる。ここで、$\chi_+(z)$ は、$z > 0$ では 1、$z < 0$ では 0 となる関数とする。これらはいずれも指数分布とは別の分布となる。

$f_\lambda^{(n)}$ については、

$$f_\lambda^{(n)}(z) = \lambda^n \frac{z^{n-1}}{(n-1)!} e^{-\lambda z}\chi_+(z) \tag{6}$$

となることを示す。$n = 1, 2$ については上の結果より成立する。$n - 1$ まで成り立つとすると $(n \geq 2)$、$z > 0$ に対し、

$$\begin{aligned} f_\lambda^{(n)}(z) &= (f_\lambda^{(n-1)} * f_\lambda)(z) = \frac{\lambda^{n-1}}{(n-2)!}\int_0^z t^{n-2} e^{-\lambda t} \lambda e^{-\lambda(z-t)} dt \\ &= \frac{\lambda^n}{(n-2)!} e^{-\lambda z}\int_0^z t^{n-2} dt = \frac{\lambda^n}{(n-2)!} e^{-\lambda z}\frac{z^{n-1}}{n-1} = \lambda^n \frac{z^{n-1}}{(n-1)!} e^{-\lambda z} \end{aligned}$$

となって $n$ でも成立する。ちなみに、

$$\begin{aligned} \int_{\mathrm{R}} f_\lambda^{(n)}(z) dz &= \frac{1}{(n-1)!}\int_0^\infty \lambda^n z^{n-1} e^{-\lambda z} dz = \frac{1}{(n-1)!}\int_0^\infty y^{n-1} e^{-y} dy \\ &= \frac{1}{(n-1)!}\,\Gamma(n) = \frac{(n-1)!}{(n-1)!} = 1 \end{aligned}$$

となり、$f_\lambda^{(n)}$ が確かにある分布の密度関数であることがわかるが、これを密度関数として持つ分布はガンマ分布 $\Gamma(n, 1/\lambda)$ と呼ばれる。よって、$e(1/\lambda)^{(n)} = \Gamma(n, 1/\lambda)$ となる。

# 6 指数分布の裏がポアソン分布となること

本節では、4 節とは逆に、指数分布 $e(1/\lambda)$ の裏がポアソン分布 $P(\lambda)$ となることを考える。

4 節と同様に、最後にその事象が起きたところを時刻 0 として考える。時刻 0 から 1 回目の事象が起こるまでの時間を $d_1$、$(j-1)$ 回目の事象から $j$ 回目の事象までの時間を $d_j$ とする $(j=2,3,\ldots)$。

このとき、例えば 1 時間以内に 2 回の事象が収まる確率 $\mathrm{Prob}\{d_1+d_2\leq 1\}$ は、1 時間以内に起こる回数 $x$ で見ると $\mathrm{Prob}\{x\geq 2\}$ に対応する ($\mathrm{Prob}\{x=2\}$ ではない)。よって、一般に、

$$\mathrm{Prob}\{d_1+\cdots+d_m\leq 1\}=\mathrm{Prob}\{x\geq m\} \tag{7}$$

となる。

この左辺は、前節例 7 より独立な指数分布の和であるガンマ分布 $\Gamma(m,1/\lambda)$ に従うので、その密度関数 $f_\lambda^{(m)}$ の 1 までの積分で表される。

$$\mathrm{Prob}\{d_1+\cdots+d_m\leq 1\}=\int_0^1\lambda^m\frac{z^{m-1}}{(m-1)!}e^{-\lambda z}dz=\int_0^\lambda\frac{y^{m-1}}{(m-1)!}e^{-y}dy \tag{8}$$

この右辺を $K_m$ と書くことにする $(m\geq 1)$。同様にして、

$$\mathrm{Prob}\{d_1+\cdots+d_{m+1}\leq 1\}=K_{m+1}$$

となるから、(7) より、

$$\mathrm{Prob}\{x=m\}=\mathrm{Prob}\{x\geq m\}-\mathrm{Prob}\{x\geq m+1\}=K_m-K_{m+1} \tag{9}$$

となる。ここで部分積分により、

$$K_{m+1}=\int_0^\lambda\frac{y^m}{m!}(-e^{-y})'dy=-\frac{\lambda^m}{m!}e^{-\lambda}+\int_0^\lambda\frac{y^{m-1}}{(m-1)!}e^{-y}dy=-\frac{\lambda^m}{m!}e^{-\lambda}+K_m$$

となるので、よって (9) より

$$\mathrm{Prob}\{x=m\}=K_m-K_{m+1}=\frac{\lambda^m}{m!}e^{-\lambda}=p_\lambda(m)$$

の、ポアソン分布の確率関数 $p_\lambda$ の形になる ($m \geq 1$)。$m=0$ のときは、(7) で $m=1$ として 1 から引けば、

$$\begin{aligned}\mathrm{Prob}\{x=0\} &= \mathrm{Prob}\{d_1>1\}=1-K_1=1-\int_0^\lambda e^{-y}dy=1+e^{-\lambda}-1\\ &= e^{-\lambda}=p_\lambda(0)\end{aligned}$$

となる。

よって、指数分布 $e(1/\lambda)$ を時間間隔とする独立な事象の 1 時間の回数の分布はポアソン分布 $P(\lambda)$ であることが示された。

# 参考文献

[1] 竹野茂治「多次元確率分布と独立性」(2022) http://takeno.iee.niit.ac.jp/~shige/math/lecture/graduate/data/ndimrandvar1.pdf

[2] 竹野茂治「正規確率変数の一次式」(2022) http://takeno.iee.niit.ac.jp/~shige/math/lecture/graduate/data/normal1.pdf

[3] 和達三樹、十河清「キーポイント確率・統計」、岩波書店 (1993)