ガンマ分布とベータ分布
(Gamma and Beta distributions)

緑川章一 *

1　ガンマ分布の導出

確率変数 $x_1,\ x_2,\ \cdots,\ x_n$ は、互いに独立で、指数分布 $f(x_i) = \lambda e^{-\lambda x_i}\quad (i = 1,\ 2,\ \cdots,\ n)$ に従うものとする。このとき、ガンマ分布 (gamma distribution) は、これら n コの独立変数の和 $z = x_1 + x_2 + \cdots + x_n$ の分布として導き出される。すなわち、

$$f_n(z) = \lambda^n \int \delta(z - x_1 - x_2 - \cdots - x_n) e^{-\lambda(x_1+x_2+\cdots+x_n)}\, dx_1 dx_2 \cdots dx_n \tag{1}$$

で与えられる。

(1) 式の積分は容易に実行することができる。$x_i \geqq 0\quad (i = 1,\ 2,\ \cdots,\ n)$ に注意すると、

$$f_n(z) = \lambda^n e^{-\lambda z} \int_0^z dx_n \int_0^{z-x_n} dx_{n-1} \cdots \int_0^{z-(x_4+\cdots+x_n)} dx_3 \int_0^{z-(x_3+x_4+\cdots+x_n)} dx_2 \tag{2}$$

となる。積分範囲が上記のようになることは、以下のようにして導くことができる。

$z = x_1 + x_2 + x_3 + \cdots + x_n$ より、$x_1 = z - (x_2 + x_3 + \cdots + x_n) \geqq 0$ なので、
$x_2 \leqq z - (x_3 + \cdots + x_n)$. ところが、$x_2 \geqq 0$ だから、$x_3 \leqq z - (x_4 + \cdots + x_n)$. 今度は $x_3 \geqq 0$ だから、$x_4 \leqq z - (x_4 + \cdots + x_n)$ 等々、と繰り返すと、最後に x_n の範囲として $0 \leqq x_n \leqq z$ を得る。

(2) 式の積分を実際に実行する。まず、

$$\int_0^{z-(x_3+x_4+\cdots+x_n)} dx_2 = (z - (x_4 + \cdots + x_n)) - x_3$$

次に、

$$\begin{aligned}
\int_0^{z-(x_4+\cdots+x_n)} \{(z - (x_4 + \cdots + x_n)) - x_3\}\, dx_3 &= \frac{1}{2!} (z - (x_4 + \cdots + x_n))^2 \\
&= \frac{1}{2!} ((z - (x_5 + \cdots + x_n)) - x_4)^2
\end{aligned}$$

更に、

$$\begin{aligned}
\frac{1}{2!} \int_0^{z-(x_5+\cdots+x_n)} \{(z - (x_5 + \cdots + x_n)) - x_4\}^2\, dx_4 &= \frac{1}{3!} (z - (x_5 + \cdots + x_n))^3 \\
&= \frac{1}{3!} ((z - (x_6 + \cdots + x_n)) - x_5)^3
\end{aligned}$$

*Shoichi Midorikawa

と繰り返して、最後に、

$$\begin{aligned}\frac{1}{(n-2)!}\int_0^z (z-x_n)^{n-2}\,dx_n &= \frac{1}{(n-1)!}\Big[-(z-x_n)^{n-1}\Big]_0^z \\ &= \frac{1}{(n-1)!}z^{n-1}\end{aligned}$$

を得るので、

$$f_n(z) = \frac{\lambda^n}{(n-1)!}z^{n-1}e^{-\lambda z}$$

となる。ところで、$\Gamma(n) = (n-1)!$ だから、ガンマ関数を用いて書き直すと、

$$f_n(z) = \frac{\lambda^n}{\Gamma(n)}z^{n-1}e^{-\lambda z} \tag{3}$$

となる。ところで、ガンマ関数の積分表示は、

$$\Gamma(n) = \int_0^\infty x^{n-1}e^{-x}\,dx$$

であった。(3) 式の積分から、

$$\begin{aligned}\lambda^n\int_0^\infty z^{n-1}e^{-\lambda z}\,dz &= \int_0^\infty x^{n-1}e^{-x}\,dx \qquad (x=\lambda z) \\ &= \Gamma(n)\end{aligned}$$

が得られる。そこで、(3) 式をガンマ分布 (gamma distribution) と言い、記号 $Ga(n,\ \lambda)$ で表す。

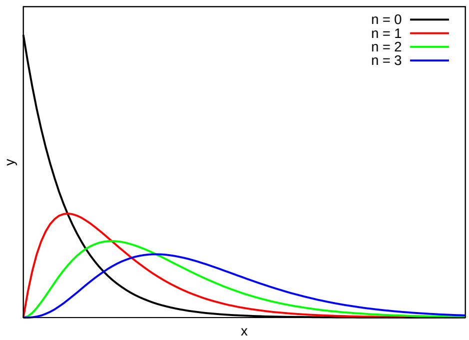

図 1: ガンマ分布 $Ga(n,\ \lambda)$ $n = 0,\ 1,\ 2,\ 3$ の場合

1.1 ガンマ分布の平均と分散

平均

$$\begin{aligned}\mu &= \int_0^\infty z f_n(z)\,dz \\ &= \frac{\lambda^n}{\Gamma(n)}\int_0^\infty z^n e^{-\lambda z}\,dz\end{aligned}$$

ここで、$x = \lambda z$ とおくと、

$$
\begin{aligned}
\mu &= \frac{1}{\lambda\Gamma(n)}\int_0^\infty x^n e^{-x}\,dx \\
&= \frac{\Gamma(n+1)}{\lambda\Gamma(n)} \\
&= \frac{n}{\lambda}
\end{aligned}
$$

分散

$$
\begin{aligned}
\sigma^2 &= \int_0^\infty z^2 f_n(z)\,dz - \mu^2 \\
&= \frac{\lambda^n}{\Gamma(n)}\int_0^\infty z^{n+1}e^{-\lambda z}\,dz - \mu^2 \\
&= \frac{\Gamma(n+2)}{\lambda^2\Gamma(n)} - \frac{n^2}{\lambda^2} \\
&= \frac{n}{\lambda^2}
\end{aligned}
$$

2　ガンマ分布のモーメント母関数

ガンマ分布のモーメント母関数 $M_Z(t)$ を求めよう。

$$
M_Z(t) = \int_0^\infty e^{tz} f_n(z)\,dz
$$

に、(1) 式を代入すると、

$$
M_Z(t) = \left(\lambda\int_{-\infty}^\infty e^{(t-\lambda)x_1}\,dx_1\right)\left(\lambda\int_0^\infty e^{(t-\lambda)x_2}\,dx_2\right)\cdots\left(\lambda\int_0^\infty e^{(t-\lambda)x_n}\,dx_n\right)
$$

ところで、$i = 1,\ 2,\ \cdots n$ に対して、

$$
\begin{aligned}
M_i(t) &= \left(\lambda\int_0^\infty e^{(t-\lambda)x_i}\,dx_i\right) \\
&= \frac{\lambda}{\lambda - t} \qquad (\text{ただし、}t < \lambda\text{とする})
\end{aligned}
$$

となるので、

$$
M_Z(t) = \left(\frac{\lambda}{\lambda - t}\right)^n \qquad (\text{ただし、}t < \lambda)
$$

を得る。

3　ガンマ分布と χ^2 分布

自由度 n の χ^2 分布は、

$$
T_n(z) = \frac{1}{2\Gamma(n/2)}\left(\frac{z}{2}\right)^{\frac{n}{2}-1} e^{-\frac{z}{2}}
$$

であった。この式と (3) 式を比べると、自由度 n の χ^2 分布は、$Ga(n/2,\ 1/2)$ であることが分かる。

4　逆ガンマ分布

逆ガンマ分布 (inverse gamma distribution) とは、(3) 式の変数 z を $y = \dfrac{1}{z}$ で置き換えたものである。

$$\begin{aligned} f_Y(y) &= \frac{\lambda^n}{\Gamma(n)} \int_0^\infty \delta\left(y - \frac{1}{z}\right) z^{n-1} e^{-\lambda z}\, dz \\ &= \frac{\lambda^n}{\Gamma(n)} \int_0^\infty z^2 \delta\left(z - \frac{1}{y}\right) z^{n-1} e^{-\lambda z}\, dz \\ &= \frac{\lambda^n}{\Gamma(n)} y^{-n-1} \exp\left(-\frac{\lambda}{y}\right) \end{aligned}$$

すなわち、

$$f_Y(y) = \frac{\lambda^n}{\Gamma(n)} y^{-n-1} \exp\left(-\frac{\lambda}{y}\right)$$

を逆ガンマ分布という。

5　ベータ分布

2 つの独立な確率変数 x と y を考える。確率変数 x はガンマ分布 $Ga(m,\ \lambda)$ に、確率変数 y は $Ga(n,\ \lambda)$ に従うとする。すなわち、

$$f_m(x) = \frac{\lambda^m}{\Gamma(m)} x^{m-1} e^{-\lambda x} \tag{4a}$$

$$f_n(y) = \frac{\lambda^n}{\Gamma(n)} y^{n-1} e^{-\lambda y} \tag{4b}$$

とする。

5.1　$x/(x+y)$ の分布

確率変数 $z = \dfrac{x}{x+y}$ の分布は、

$$f_Z(z) = \int \delta\left(z - \frac{x}{x+y}\right) f_m(x) f_n(y)\, dxdy$$

で与えられる。この積分を求めるのだが、y についての積分を先に実行することにすると、

$$\delta\left(z - \frac{x}{x+y}\right) = \frac{x}{z^2}\delta\left(y - \frac{(1-z)x}{z}\right)$$

だから、

$$f_Z(z) = \frac{1}{z^2} \int_0^\infty x\, f_m\,(x)\ f_n\left(\frac{1-z}{z}x\right)\, dx \tag{5}$$

となる。この式に (4a), (4b) 式を代入して計算すると、

$$
\begin{aligned}
f_Z(z) &= \frac{\lambda^{m+n}}{\Gamma(m)\Gamma(n)} z^{-n-1}(1-z)^{n-1} \int_0^\infty x^{m+n-1} e^{-\lambda x/z}\,dx \\
&= \frac{1}{\Gamma(m)\Gamma(n)} z^{m-1}(1-z)^{n-1} \int_0^\infty t^{m+n-1} e^{-t}\,dt \qquad \left(t = \frac{\lambda}{z}x\right) \\
&= \frac{\Gamma(m+n)}{\Gamma(m)\Gamma(n)} z^{m-1}(1-z)^{n-1} \\
&= \frac{z^{m-1}(1-z)^{n-1}}{B(m,n)} \qquad (0 \leqq\ z \leqq\ 1)
\end{aligned}
$$

ここで、

$$
\Gamma(m+n) = \int_0^\infty t^{m+n-1} e^{-t}\,dt
$$

および、

$$
B(m,n) = \frac{\Gamma(m)\Gamma(n)}{\Gamma(m+n)}
$$

を用いた。

ベータ関数 $B(m,n)$ の積分表示は、

$$
B(m,n) = \int_0^1 z^{m-1}(1-z)^{n-1}\ dz
$$

なので、確率密度関数

$$
f_Z(z) = \frac{z^{m-1}(1-z)^{n-1}}{B(m,n)} \qquad (0 \leqq\ z \leqq\ 1) \tag{6}
$$

をベータ分布と呼び、$Be(m,n)$ で表す。

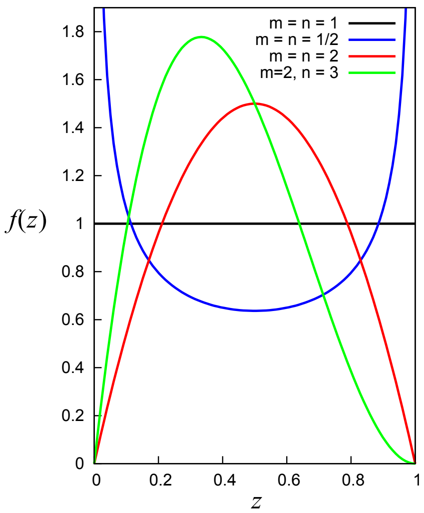

図 2: ベータ分布 $Be(m,n)$

5.2 ベータ分布の平均と分散

平均

$$\begin{aligned}
\mu &= \int_0^1 z f_Z(z)\,dz \\
&= \frac{1}{B(m,n)}\int_0^1 z^m(1-z)^{n-1}\,dz \\
&= \frac{B(m+1,n)}{B(m,n)} \\
&= \frac{\Gamma(m+1)\Gamma(n)}{\Gamma(m+n+1)}\frac{\Gamma(m+n)}{\Gamma(m)\Gamma(n)} \\
&= \frac{m}{m+n}
\end{aligned}$$

分散

$$\begin{aligned}
\sigma^2 &= \int_0^1 z^2 f_Z(z)\,dz - \mu^2 \\
&= \frac{1}{B(m,n)}\int_0^1 z^{m+1}(1-z)^{n-1}\,dz - \mu^2 \\
&= \frac{B(m+2,n)}{B(m,n)} - \mu^2 \\
&= \frac{\Gamma(m+2)\Gamma(n)}{\Gamma(m+n+2)}\frac{\Gamma(m+n)}{\Gamma(m)\Gamma(n)} - \mu^2 \\
&= \frac{(m+1)m}{(m+n+1)(m+n)} - \frac{m^2}{(m+n)^2} \\
&= \frac{mn}{(m+n)^2(m+n+1)}
\end{aligned}$$

5.3 x/y の分布

確率変数 $u = \dfrac{x}{y}$ の分布は、

$$f_U(u) = \int \delta\left(u - \frac{x}{y}\right) f_m(x) f_n(y)\, dx dy$$

で与えられる。この積分を求めるときに、今度は x についての積分を先に実行することにすると、

$$\delta\left(u - \frac{x}{y}\right) = y\delta\,(x - uy)$$

だから、

$$f_U(u) = \int y f_m(uy) f_n(y)\, dy \tag{7}$$

となる。この式に (4a), (4b) 式を代入して計算すると、

$$f_U(u) = \frac{\lambda^{m+n}}{\Gamma(m)\Gamma(n)} u^{m-1} \int_0^\infty y^{m+n-1} e^{-\lambda(1+u)y}\, dy \tag{8}$$

ここで、$t = \lambda(1+u)y$ とおくと、

$$\begin{aligned} f_U(u) &= \frac{1}{\Gamma(m)\Gamma(n)} \frac{u^{m-1}}{(1+u)^{m+n}} \int_0^\infty t^{m+n-1} e^{-t}\, dt \\ &= \frac{\Gamma(m+n)}{\Gamma(m)\Gamma(n)} \frac{u^{m-1}}{(1+u)^{m+n}} \\ &= \frac{1}{B(m,n)} \frac{u^{m-1}}{(1+u)^{m+n}} \end{aligned} \tag{9}$$

を得る。

ところで、ベータ関数の積分表示は、

$$B(m,n) = \int_0^1 z^{m-1}(1-z)^{n-1}\, dz = \int_0^\infty \frac{u^{m-1}}{(1+u)^{m+n}}\, du \qquad u = \frac{z}{1-z}$$

なので、(9) 式も、また、一種のベータ分布とみなすことができる。

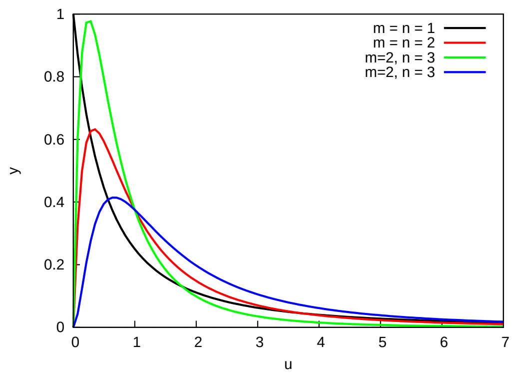

図 3: 分布 $f_U(u)$

$z^{m-1}(1-z)^{n-1}$ と $\dfrac{u^{m-1}}{(1+u)^{m+n}}$ の関係は、

$$z^{m-1}(1-z)^{n-1} = \int_0^\infty \delta\left(z - \frac{u}{1+u}\right) \frac{u^{m-1}}{(1+u)^{m+n}}\, du$$

で与えられる。このことは、デルタ関数の公式、

$$\delta\left(g(u)\right) = \frac{1}{|g'(\alpha)|}\delta(u-\alpha) \qquad \text{ここに、}\alpha\text{は } g(u)=0 \text{ の解である。}$$

を用いれば、容易に証明できる。今の場合は、$\delta\left(z - \dfrac{u}{1+u}\right) = \dfrac{1}{(1-z)^2}\delta\left(u - \dfrac{z}{1-z}\right)$ である。

ゆえに (6) 式と (9) 式の関係は、

$$f_Z(z) = \frac{1}{(1-z)^2} f_U\left(\frac{z}{1-z}\right)$$

となる。