連続一様分布する確率変数の和
(Sums of Continuous Uniform Random Variables)

緑川章一*

$X = x_i (i = 0,\ 1,\ 2, \cdots,\ n)$ は、それぞれ独立で、区間 $0 \leqq x_i \leqq 1$ における一様連続分布 $f(X)$ の確率変数とする。

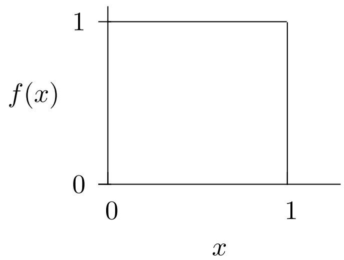

図 1: 確率密度関数 $f(x)$

このとき、$z = x_1 + x_2 \cdots + x_n$ の確率密度関数を求める。

1　$n = 2$ の場合

$$f_2(z) = \int_0^1 dx_2 \int_0^1 \delta(z - x_1 - x_2)\, dx_1 \tag{1}$$

まず最初に、x_1 についての積分をおこなう。次に、x_2 についての積分をおこなう。$x_1 = z - x_2$ の積分領域は、$0 \leqq x_1 \leqq 1$ だから、$0 \leqq z - x_2 \leqq 1$。ゆえに、$z - 1 \leqq x_2 \leqq z$。これと $0 \leqq x_2 \leqq 1$ を比較して、

$$max\{0,\ z-1\} \leqq x_2 \leqq min\{1,\ z\}$$

すなわち、$0 \leqq z \leqq 1$ のときは $0 \leqq x_2 \leqq z$, $11 \leqq z \leqq 2$ のときは $z - 1 \leqq x_2 \leqq 1$ となる。ゆえに、

$$f_2(z) = \begin{cases} \displaystyle\int_0^z dx_2 = z & 0 \leqq z \leqq 1 \\ \displaystyle\int_{z-1}^1 dx_2 = 2 - z & 1 \leqq z \leqq 2 \end{cases} \tag{2}$$

を得る。

*Shoichi Midorikawa

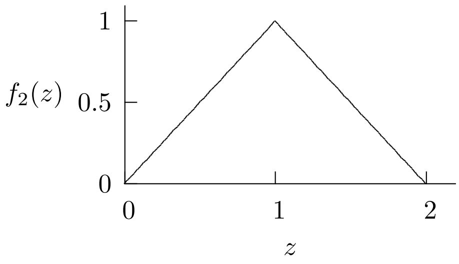

図 2: 確率密度関数 $f_2(z)$

2　$n=3$ の場合

$z=x_1+x_2+x_3$ とおくと、$0\leqq z\leqq 3$

$$f_3(z)=\int_0^1 dx_3\int_0^1 dx_2\int_0^1 \delta(z-x_1-x_2-x_3)\,dx_1 \tag{3}$$

まず最初に、x_1 についての積分をおこなう。次に、x_2 についての積分をおこなう。$x_1=z-x_2-x_3$ の積分領域は、$0\leqq x_1\leqq 1$ だから、$0\leqq z-x_2-x_3\leqq 1$。ゆえに、$z-1-x_3\leqq x_2\leqq z-x_3$。ところが、$0\leqq x_2\leqq 1$ だから、

$$max\{z-1-x_3,\ 0\}\leqq x_2\leqq min\{z-x_3,\ 1\}$$

すなわち、

(A) $x_3\leqq z-1$ のとき、$z-1-x_3\leqq x_2\leqq 1$
(B) $z-1\leqq x_3$ のとき、$0\leqq x_2\leqq z-x_3$

積分領域が存在するためには、$z-1-x_3\leqq 1$、かつ、$0\leqq z-x_3$。すなわち、$z-2\leqq x_3\leqq z$ でなければならない。これと、$0\leqq x_3\leqq 1$ を比較して、

$$max\{z-2,\ 0\}\leqq x_3\leqq mini\{z,\ 1\}$$

また、$z\leqq 1$ のとき、(A) を満たす x_3 は存在しない。$1\leqq z-1$ のときには、(B) を満たす x_3 は存在しない。

$0\leqq z\leqq 1$ の場合 $z-1\leqq x_3$ かつ $0\leqq x_3\leqq z$ だから、$0\leqq x_3\leqq z$。

$$f_3(z)=\int_0^z dx_3\int_0^{z-x_3} dx_2=\frac{z^2}{2} \tag{4a}$$

$1\leqq z\leqq 2$ の場合 $0\leqq z_3\leqq 1$ だから、

$$f_3(z)=\int_0^{z-1} dx_3\int_{z-1-x_3}^1 dx_2+\int_{z-1}^1 dx_3\int_0^{z-x_3} dx_2=-z^2+3z-\frac{3}{2} \tag{4b}$$

$2\leqq z\leqq 3$ の場合 $x_3\leqq z-1$ かつ $z-2\leqq x_3\leqq 1$ だから、$z-2\leqq z_3\leqq 1$。

$$f_3(z)=\int_{z-2}^1 dx_3\int_{z-1-x_3}^1 dx_2=\frac{1}{2}(z-3)^2 \tag{4c}$$

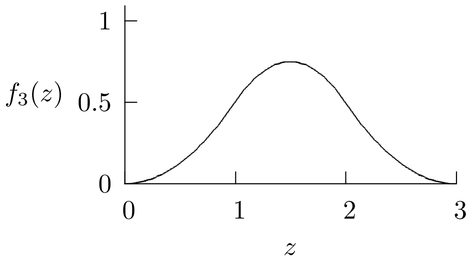

図 3: 確率密度関数 $f_3(z)$

3　$n = 4$ の場合

(2) 式を用いると、

$$
\begin{aligned}
f_4(z) &= \int_0^2 dy \int_0^2 \delta(z-x-y) f_2(x) f_2(y)\, dx \\
&= \int_0^2 f_2(z-y) f_2(y)\, dy \qquad (5)
\end{aligned}
$$

となる。$z = x + y$ で、$0 \leqq x \leqq 2,\ 0 \leqq y \leqq 2$ なので、当然のことながら、$0 \leqq z \leqq 4$ である。ところで、$0 \leqq y \leqq 2$ より、$0 \leqq z - x \leqq 2$。すなわち、$z - 2 \leqq y \leqq z$。これと、$0 \leqq y \leqq 2$ を比較すると、

$$
max\{z-2,\ 0\} \leqq y \leqq min\{z,\ 2\}
$$

$0 \leqq z \leqq 2$ のときは、$0 \leqq y \leqq z$ なので、

$$
f_4(z) = \int_0^z f_2(z-y) f_2(y)\, dy \qquad (6)
$$

$2 \leqq z \leqq 4$ のときは、$z - 2 \leqq y \leqq 2$ なので、

$$
f_4(z) = \int_{z-2}^2 f_2(z-y) f_2(y)\, dy \qquad (7)
$$

(6) 式は、更に 2 つの場合に分けられる。すなわち、$0 \leqq z \leqq 1$ の場合と、$2 \leqq z \leqq 2$ の場合である。

$0 \leqq z \leqq 1$ の場合

$$
f_4(z) = \int_0^z (z-y) y\, dy = \frac{z^3}{6} \qquad (8)
$$

$1 \leqq z \leqq 2$ の場合

$$
\begin{aligned}
f_4(z) &= \int_0^{z-1} (2-z+y) y\, dy + \int_{z-1}^1 (z-y) y\, dy + \int_1^z (z-y)(2-y)\, dy \qquad (9) \\
&= -\frac{1}{2} z^3 + 2z^2 - 2z + \frac{2}{3} \qquad (10)
\end{aligned}
$$

(7) 式は、更に 2 つの場合に分けられる。すなわち、$2 \leqq z \leqq 3$ の場合と、$3 \leqq z \leqq 4$ の場合である。

$2 \leqq z \leqq 3$ の場合

$$\begin{aligned} f_4(z) &= \int_{z-2}^{1} (2-z+y)y\,dy + \int_{1}^{z-1} (2-z+y)(2-y)\,dy + \int_{z-1}^{2} (z-y)(2-y)\,dy \\ &= \frac{1}{2}z^3 - 4z^2 + 10z - \frac{22}{3} \end{aligned}$$

$3 \leqq z \leqq 4$ の場合

$$f_4(z) = \int_{z-2}^{2} (2-z+y)(2-y)\,dy = \frac{1}{6}(4-z)^3 \tag{11}$$

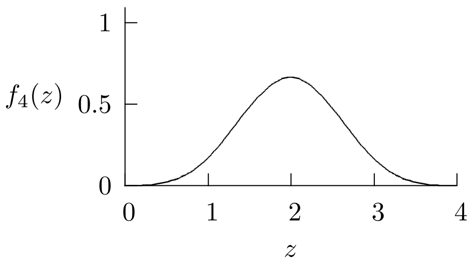

図 4: 確率密度関数 $f_4(z)$

4　中心極限定理 (central limit theorem)

いま、一様分布の密度関数 $f(X)$ の確率変数 X を平均値 $\mu = \dfrac{1}{2}$ が 0 になるように、変換 $Y = X - \dfrac{1}{2}$ をおこない、n コの独立な確率変数 $Y_1,\ Y_2,\ \cdots,\ Y_n$ の算術平均 $\bar{Y} = (Y_1 + Y_2 + \cdots + Y_n)/n$ の確率分布

$$f_{\bar{Y}}(\bar{y}) = \int_{-1/2}^{1/2} dy_1 \int_{-1/2}^{1/2} dy_2 \cdots \int_{-1/2}^{1/2} dy_n\, \delta\left(\bar{y} - \frac{y_1 + y_2 + \cdots + y_n}{n}\right) \tag{12}$$

を n が大きい場合に求める。

まず、デルタ関数の積分表示

$$\delta\left(\bar{y} - \frac{y_1 + y_2 + \cdots + y_n}{n}\right) = \frac{1}{2\pi} \int_{-\infty}^{\infty} e^{ik(\bar{y} - (y_1 + y_2 + \cdots + y_n)/n)}\,dk$$

を (12) 式に代入すると、

$$\begin{aligned} f_{\bar{Y}}(\bar{y}) &= \frac{1}{2\pi} \int_{-\infty}^{\infty} dk\, e^{ik\bar{y}} \prod_{i}^{n} \left(\int_{-1/2}^{1/2} e^{-iky_i/n}\,dy_i \right) \\ &= \frac{1}{2\pi} \int_{-\infty}^{\infty} dk\, e^{ik\bar{y}} \left(\frac{\sin \frac{k}{2n}}{\frac{k}{2n}} \right)^n \end{aligned}$$

ところで、$\left(\dfrac{\sin \frac{k}{2n}}{\frac{k}{2n}} \right)^n$ は、n が十分大きいときには、$k/2n \approx 0$ で 1 となる鋭いピークを持つ k についての偶関数である。$k/2n \approx 0$ 付近では、

$$\left(\frac{\sin\frac{k}{2n}}{\frac{k}{2n}}\right)^n \simeq 1-\frac{k^2}{24n}$$

となるので、指数関数 $\exp\left(-\frac{k^2}{24n}\right)$ で近似することにすると、

$$\begin{aligned}
f_{\bar{Y}}(\bar{y}) &\simeq \frac{1}{2\pi}\int_{-\infty}^{\infty} dk\,\exp\left[-\frac{k^2}{24n}+ik\bar{y}\right] \\
&= \frac{1}{2\pi}\int_{-\infty}^{\infty} dk\,\exp\left[-\frac{1}{24n}\left(k-12in\bar{y}\right)^2-6n\bar{y}^2\right] \\
&= \frac{1}{2\pi}\,e^{-6n\bar{y}^2}\int_{-\infty}^{\infty} dk\,\exp\left[-\frac{1}{24n}\left(k-12in\bar{y}\right)^2\right]
\end{aligned}$$

ところで、

$$\begin{aligned}
&\int_{-\infty}^{\infty} dk\,\exp\left[-\frac{1}{24n}\left(k-12in\bar{y}\right)^2\right] \\
&= \int_{-\infty}^{\infty} dk'\,\exp\left[-\frac{k'^2}{24n}\right] \\
&= \sqrt{24\pi n}
\end{aligned}$$

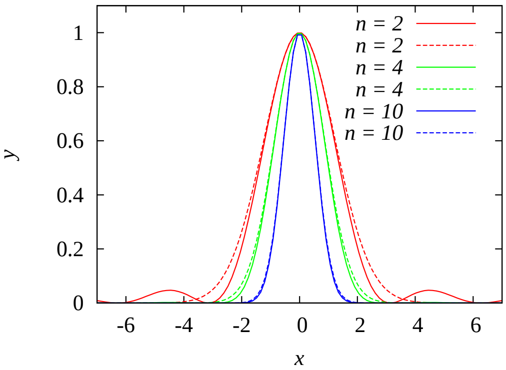

図 5: $y=(\sin x/x)^n$ のグラフ (実線) と $y=e^{-nx^2/6}$(点線) のグラフ

だから、

$$f_{\bar{Y}}(\bar{y}) \simeq \sqrt{\frac{6n}{\pi}}\,e^{-6n\bar{y}^2} \tag{13}$$

が成り立つ。これは、分散が $\sigma_n^2=\dfrac{1}{12n}$ の正規分布である。一様分布の分散は、

$$\sigma^2=\int_{-1/2}^{1/2} x^2\,dx=\frac{1}{12}$$

なので、

$$\sigma_n^2=\frac{\sigma^2}{n} \tag{14}$$

が成り立ち、(13) 式は、

$$f_{\bar{Y}}(\bar{y}) \simeq \frac{1}{\sqrt{2\pi\sigma_n^2}}e^{-\bar{y}^2/2\sigma_m^2} \tag{15}$$

と書ける。ここで、σ_n^2 は (14) 式で与えられる。