中心極限定理とその証明

(The Central Limit Theorem - A Direct Proof -)

2013.8.11 タイプミスの訂正

緑川章一*

$X_1,\ X_2,\ \cdots,\ X_n$ は、各々独立で同一の確率密度関数 $f(X)$ に従う確率変数 (independent and identically distributed (i.i.d) random variables) とする。簡単のために、$f(X)$ の平均を 0，分散を σ^2 とする。すなわち、

$$\int x f(x)\,dx = 0 \tag{1}$$

$$\int x^2 f(x)\,dx = \sigma^2 \tag{2}$$

確率変数 $\bar{X} = (X_1 + X_2 + \cdots + X_n)/\sqrt{n}\sigma$ の密度関数を $f_{\bar{X}}(\bar{x})$ と書くことにすると、

$$f_{\bar{X}}(\bar{x}) = \int \delta\left(\bar{x} - \frac{x_1 + x_2 + \cdots + x_n}{\sqrt{n}\sigma}\right) f(x_1) f(x_2)\cdots f(x_n)\,dx_1 dx_2 \cdots dx_n \tag{3}$$

で与えられる。ここで、デルタ関数の積分表示

$$\delta\left(\bar{x} - \frac{x_1 + x_2 + \cdots + x_n}{\sqrt{n}\sigma}\right) = \frac{1}{2\pi}\int_{-\infty}^{\infty} e^{ik\left(\bar{x} - (x_1 + x_2 + \cdots + x_n)/\sqrt{n}\sigma\right)}\,dk$$

を (3) 式に代入すると、

$$\begin{aligned} f_{\bar{X}}(\bar{x}) &= \frac{1}{2\pi}\int_{-\infty}^{\infty} dk\, e^{ik\bar{x}} \prod_{i=1}^{n} \int e^{-ikx_i/\sqrt{n}\sigma} f(x_i)\,dx_i \\ &= \frac{1}{2\pi}\int_{-\infty}^{\infty} dk\, e^{ik\bar{x}} \left[\int e^{-ikx/\sqrt{n}\sigma} f(x)\,dx\right]^n \end{aligned} \tag{4}$$

(4) 式のカッコ [ ] 内の関数

$$\varphi\left(\frac{k}{\sqrt{n}\sigma}\right) = \int e^{-ikx/\sqrt{n}\sigma} f(x)\,dx \tag{5}$$

は、確率分布 $f(X)$ の特性関数 (characteristic function) である。指数の符号が、一般の定義と違っているが、本質的にはまったく同じである。一般の特性関数の定義に一致させたければ、(4) 式において k を $-k$ に置き換えれば良い。

次に $e^{-ikx/\sqrt{n}\sigma}$ を、良く知られた部分積分の方法を用いて Taylor 展開する。

$$\begin{aligned} e^{-ikx/\sqrt{n}\sigma} &= 1 + i\frac{k}{\sqrt{n}\sigma}\int_0^x (x-t)' e^{-ikt/\sqrt{n}\sigma}\,dt \\ &= 1 + i\frac{k}{\sqrt{n}\sigma}\left\{\left[(x-t)e^{-ikt/\sqrt{n}\sigma}\right]_0^x + i\frac{k}{\sqrt{n}\sigma}\int_0^x (x-t) e^{-ikt/\sqrt{n}\sigma}\,dt\right\} \end{aligned}$$

*Shoichi Midorikawa

$$
\begin{aligned}
&= 1 - i\frac{k}{\sqrt{n}\sigma}\,x - \left(i\frac{k}{\sqrt{n}\sigma}\right)^2 \int_0^x \Big[\frac{(x-t)^2}{2!}\Big]' e^{-ikt/\sqrt{n}\sigma}\,dt \\
&= 1 - i\frac{k}{\sqrt{n}\sigma}\,x - \left(i\frac{k}{\sqrt{n}\sigma}\right)^2 \left\{ \Big[\frac{(x-t)^2}{2!}e^{-ikt/\sqrt{n}\sigma}\Big]_0^x + i\frac{k}{\sqrt{n}\sigma}\int_0^x \frac{(x-t)^2}{2!}e^{-ikt/\sqrt{n}\sigma}\,dt \right\} \\
&= 1 - i\frac{k}{\sqrt{n}\sigma}\,x - \frac{k^2}{2n\sigma^2}\,x^2 + i\frac{k^3}{n^{3/2}\sigma^3}\int_0^x \frac{(x-t)^2}{2!}e^{-ikt/\sqrt{n}\sigma}\,dt
\end{aligned}
$$

ここで、

$$
R(x) = \int_0^x \frac{(x-t)^2}{2!}e^{-ikt/\sqrt{n}\sigma}\,dt \tag{6}
$$

と書くことにすると、結局、

$$
e^{-ikx/\sqrt{n}\sigma} = 1 - i\frac{k}{\sqrt{n}\sigma}\,x - \frac{k^2}{2n\sigma^2}\,x^2 + i\frac{k^3}{n^{3/2}\sigma^3}R(x) \tag{7}
$$

となる。容易に分かるように、

$$
|R(x)| \leqq \int_0^{|x|} \frac{(x-t)^2}{2!}\,dt = \frac{|x|^3}{3!}
$$

である。

(5) 式を求めるにあたり、(7) を使い、$\displaystyle\int f(x)\,dx = 1$ および (1) 式、(2) 式を用いると、

$$
\begin{aligned}
\int e^{-ikx/\sqrt{n}\sigma} f(x)\,dx &= \int \left\{ 1 - i\frac{k}{\sqrt{n}\sigma}\,x - \frac{k^2}{2n\sigma^2}\,x^2 + i\frac{k^3}{n^{3/2}\sigma^3}R(x) \right\} f(x)\,dx \\
&= 1 - i\frac{k}{\sqrt{n}\sigma}\int xf(x)\,dx - \frac{k^2}{2n\sigma^2}\int x^2 f(x)\,dx + i\frac{k^3}{n^{3/2}\sigma^3}\int R(x)f(x)\,dx \\
&= 1 - \frac{k^2}{2n} + i\frac{k^3}{n^{3/2}\sigma^3}\int R(x)f(x)\,dx
\end{aligned} \tag{8}
$$

となる。ここで、右辺の積分を、$C = \displaystyle\int R(x)f(x)\,dx$ とおくと、

$$
|C| \leqq \int |R(x)|f(x)\,dx = \frac{1}{3!}\int |x|^3 f(x)\,dx
$$

である。ここでは、積分 $\displaystyle\int |x|^3 f(x)\,dx$ が有限、すなわち、$\displaystyle\int |x|^3 f(x)\,dx < \infty$ と仮定する。この場合、当然、$|C| < \infty$ である。このとき、

$$
\lim_{n\to\infty}\left[\int e^{-ikx/\sqrt{n}\sigma} f(x)\,dx\right]^n = \lim_{n\to\infty}\left[1 - \frac{k^2}{2n} + i\frac{k^3}{n^{3/2}\sigma^3}\,C\right]^n
$$

を求めよう。

まず、

$$
U = \frac{k^2}{2n} - i\frac{k^3}{n^{3/2}\sigma^3}\,C
$$

と置くと、

$$
\left[1 - \frac{k^2}{2n} + i\frac{k^3}{n^{3/2}\sigma^3}\,C\right]^n = \Big[1 - U\Big]^n = e^{n\ln(1-U)}
$$

ところで、

$$\ln(1-U) = -U - \frac{U^2}{2} - \frac{U^3}{3} - \frac{U^4}{4} - \cdots \qquad (|U|<1)$$

だから、

$$|\ln(1-U)+U| = \left|\frac{U^2}{2} + \frac{U^3}{3} + \frac{U^4}{4} + \cdots\right| \leqq \frac{|U|^2}{2} + \frac{|U|^3}{2} + \frac{|U|^4}{2} + \cdots \leqq \frac{1}{2}\,\frac{|U|^2}{1-|U|}$$

もしも、$|U| \leqq 1/2$ ならば、$1-|U| \geqq 1/2$ だから、

$$|\ln(1-U)+U| \leqq |U|^2$$

が成り立つ。すると、

$$\lim_{n\to\infty} |n\ln(1-U)+nU| \leqq \lim_{n\to\infty} n|U|^2 = 0$$

だから、

$$\lim_{n\to\infty} n\ln(1-U) = -\lim_{n\to\infty} nU = -\frac{k^2}{2}$$

となるので、結局、

$$\lim_{n\to\infty} \left[\int e^{-ikx/\sqrt{n}\sigma} f(x)\,dx\right]^n = e^{-k^2/2} \tag{9}$$

を得る。そこで、(9) 式の右辺を (4) 式に代入すると、

$$\begin{aligned}\lim_{n\to\infty} f_{\bar{X}}(\bar{x}) &= \frac{1}{2\pi}\int_{-\infty}^{\infty} dk\ \exp\left[ik\bar{x} - \frac{k^2}{2}\right] \\ &= \frac{1}{2\pi}e^{-\bar{x}^2/2}\int_{-\infty}^{\infty} dk\exp\left[-\frac{1}{2}(k-i\bar{x})^2\right] \end{aligned} \tag{10}$$

ところで、

$$\int_{-\infty}^{\infty} dk\exp\left[-\frac{1}{2}(k-i\bar{x})^2\right] = \int_{-\infty}^{\infty} dk' e^{-k'^2/2} = \sqrt{2\pi}$$

だから、結局 (10) 式は、

$$\lim_{n\to\infty} f_{\bar{X}}(\bar{x}) = \frac{1}{\sqrt{2\pi}}e^{-\bar{x}^2/2}$$

と書ける。これは、平均が 0, 分散が 1 の正規分布 $N(0,\ 1)$ である。

一般に、確率変数 $X_1,\ X_2,\ \cdots,\ X_n$ は各々独立で、平均が μ, 分散が σ^2 の同一の確率密度関数に従うものとすると、確率変数 $\bar{X} = (X_1 + X_2 + \cdots + X_n - n\mu)/\sqrt{n}\sigma$ は、n が十分に大きいときは、確率密度関数がどのようなものであっても、おおよそ正規分布 $N(0,\ 1)$ に従うとみなすことができる。

例 1　標準正規分布 $f(x) = \dfrac{1}{\sqrt{2\pi}}e^{-x^2/2}$ の場合

$$f_{\bar{X}}(\bar{x}) = \frac{1}{2\pi}\int_{-\infty}^{\infty} dk\, e^{ik\bar{x}}\left[\frac{1}{\sqrt{2\pi}}\int e^{-ikx/\sqrt{n}-x^2/2}\,dx\right]^n$$

$$
\begin{aligned}
&= \frac{1}{2\pi}\int_{-\infty}^{\infty} dk\, e^{ik\bar{x}} \left[\frac{e^{-k^2/2n}}{\sqrt{2\pi}} \int e^{-(x+ik/\sqrt{n})^2/2}\, dx\right]^n \\
&= \frac{1}{2\pi}\int_{-\infty}^{\infty} dk e^{-k^2/2+ik\bar{x}} \\
&= \frac{e^{-\bar{x}^2/2}}{2\pi}\int_{-\infty}^{\infty} dk e^{-(k-i\bar{x})^2/2} \\
&= \frac{1}{\sqrt{2\pi}} e^{-\bar{x}^2/2}
\end{aligned}
$$

確率変数 $\bar{X} = (X_1 + X_2 + \cdots + X_n)/\sqrt{n}$ は、厳密に正規分布 $N(0,\ 1)$ に従う。

例２　連続一様分布 $f(x) = \begin{cases} 1 & \text{for } -\frac{1}{2} \leqq x \leqq \frac{1}{2} \\ 0 & \text{otherwise} \end{cases}$　の場合

分散は、$\sigma^2 = \displaystyle\int_{-1/2}^{1/2} x^2\, dx = \frac{1}{12}$ となるので、

$\bar{X} = (X_1 + X_2 + \cdots + X_n)/\sqrt{n}\sigma = \sqrt{12/n}\,(X_1 + X_2 + \cdots + X_n)$

の分布は、

$$
\begin{aligned}
f_{\bar{X}}(\bar{x}) &= \frac{1}{2\pi}\int_{-\infty}^{\infty} dk\, e^{ik\bar{x}} \left[\int_{-1/2}^{1/2} e^{-i\sqrt{12/n}\,kx}\, dx\right]^n \\
&= \frac{1}{2\pi}\int_{-\infty}^{\infty} dk\, e^{ik\bar{x}} \left(\frac{\sin\sqrt{\frac{3}{n}}\,k}{\sqrt{\frac{3}{n}}\,k}\right)^n \\
&\approx \frac{1}{2\pi}\int_{-\infty}^{\infty} dk\, e^{ik\bar{x}} \left[1 - \frac{k^2}{2n}\right]^n \\
&\approx \frac{1}{2\pi}\int_{-\infty}^{\infty} dk\, e^{-k^2/2+ik\bar{x}} \\
&= \frac{1}{\sqrt{2\pi}} e^{-\bar{x}^2/2}
\end{aligned}
$$

すなわち、正規分布 $N(0,\ 1)$ で近似できる。

例３　指数分布 $f(x) = \lambda e^{-\lambda x}$ $(x \geqq 0)$ の場合

$$
\begin{aligned}
\mu &= \lambda\int_0^{\infty} x e^{-\lambda x}\, dx = \frac{1}{\lambda} \\
\sigma^2 &= \lambda\int_0^{\infty} x^2 e^{-\lambda x}\, dx - \mu^2 = \frac{1}{\lambda^2}
\end{aligned}
$$

だから、

$\bar{X} = (X_1 + X_2 + \cdots + X_n)/\sqrt{n}\sigma = \lambda(X_1 + X_2 + \cdots + X_n)/\sqrt{n}$

の分布は、

$$
\begin{aligned}
f_{\bar{X}}(\bar{x}) &= \frac{1}{2\pi}\int_{-\infty}^{\infty} dk\, e^{ik\bar{x}} \left[\lambda\int_0^{\infty} e^{-\lambda(1+ik/\sqrt{n})x}\, dx\right]^n \\
&= \frac{1}{2\pi}\int_{-\infty}^{\infty} dk\, e^{ik\bar{x}} \left[\frac{1}{1 + i\frac{k}{\sqrt{n}}}\right]^n
\end{aligned}
$$

ここで、

$$\left[\frac{1}{1+i\frac{k}{\sqrt{n}}} \right]^n = \left[\frac{1-i\frac{k}{\sqrt{n}}}{1+\frac{k^2}{n}} \right]^n \approx \left[1 - i\frac{k}{\sqrt{n}} - \frac{k^2}{n} \right]^n \approx e^{-i\sqrt{n}k} e^{-k^2/2}$$

と近似できるので、

$$f_{\bar{X}}(\bar{x}) \approx \frac{1}{\sqrt{2\pi}} \exp\left[-\frac{1}{2}\left(\bar{x}-\sqrt{n}\right)^2 \right]$$

となる。これは、正規分布 $N(\sqrt{n},\ 1)$ を表している。