平成 13 年 6 月 22 日

中心極限定理の証明

新潟工科大学 情報電子工学科 竹野茂治

一般の連続分布に対する中心極限定理:

> $x_1, x_2, \ldots, x_n, \ldots$ が独立同分布で、その平均を μ, 分散を σ^2 とするとき、
>
> $$\bar{x}_n = \frac{x_1 + x_2 + \cdots x_n}{n},\ \ y_n = \sqrt{n}\frac{\bar{x}_n - \mu}{\sigma}$$
>
> とすると y_n の分布は $n \to \infty$ のとき標準正規分布 $N(0,1)$ に収束する。すなわち、n が大きいとき、$\bar{x}_n$ の分布は $N(\mu, \sigma^2/n)$ で近似できる。

の証明はかなり面倒であるが (畳み込み積分、そのフーリエ変換、n 重積分の類似積分変換などの準備が必要)、次の意味でのド · モアブル = ラプラスの定理ならばスターリングの公式を使って導き出すことができる。

定理 1

$0 < p < 1$, u に対して $\mu = np$, $\sigma = \sqrt{npq}$ $(q = 1 - p)$ とし、$x = \sigma u + \mu$ $(u = (x - \mu)/\sigma)$ とするとき、p, u を固定したまま $n \to \infty$ とすると

$$\sigma \binom{n}{x} p^x q^{n-x} \to \frac{1}{\sqrt{2\pi}} e^{-u^2/2}$$

となる。よって、n が大きいとき、二項分布 $B(n,p)$ は正規分布 $N(\mu, \sigma^2)$ $(\mu = np, \sigma = \sqrt{npq})$ で近似できることになる。

この証明には、スターリングの公式:

$$n! \sim n^n \sqrt{2\pi n} e^{-n}\ \ (n \to \infty) \tag{1}$$

を用いる。なお、この $\sim$ の意味は両辺の比が 1 に収束することを意味する。また、以下ではランダウの記号と呼ばれる O (ラージオー), o (スモールオー) も使用するので、まず、その説明をしておく。

定義 2

$n \to \infty$ のとき、$a_n = o(b_n)$ (スモールオー) は、

$$\frac{a_n}{b_n} \to 0$$

となることを意味し、$a_n = O(b_n)$ (ラージオー) は、a_n/b_n が有界であることを意味する。

$a_n = o(b_n)$ のときは、a_n は b_n より小さいもの、$a_n = O(b_n)$ のときは、a_n は b_n よりほぼ同等、ということになる。例えば $a_n \to 0$ はこの記号を使えば $a_n = o(1)$ と書け、スターリングの公式は、この記号を使えば

$$n! = O(n^n \sqrt{2\pi n} e^{-n}), \quad n! = n^n \sqrt{2\pi n} e^{-n}(1 + o(1))$$

のように書けることになる。

定理 1 の証明

$x = u\sigma + \mu = u\sqrt{npq} + np$ により x も n によって変わることに注意する。また、$0 < p < 1$ とする。

$$
\begin{aligned}
&\log \sigma \begin{pmatrix} n \\ x \end{pmatrix} p^x q^{n-x} \\
&= \log \sqrt{npq}\frac{n!}{x!(n-x)!}p^x q^{n-x} \\
&= \frac{1}{2}\log npq + \log n! - \log x! - \log(n-x)! + x\log p + (n-x)\log q
\end{aligned}
$$

となる。今、$p > 0, q > 0$ なので、$x, n-x$ は $n \to \infty$ のとき

$$
\begin{aligned}
x &= np + u\sqrt{npq} = n\left(p + u\sqrt{\frac{pq}{n}}\right) \to \infty \\
n - x &= n - np - u\sqrt{npq} = nq - u\sqrt{npq} = n\left(q - u\sqrt{\frac{pq}{n}}\right) \to \infty
\end{aligned}
$$

となる。また、スターリングの公式より $m \to \infty$ のときは

$$
\begin{aligned}
\log m! &= \log m^m \sqrt{2\pi m} e^{-m}(1 + o(1)) \\
&= m\log m + \frac{1}{2}\log 2\pi + \frac{1}{2}\log m - m + \log(1 + o(1)) \\
&= \frac{1}{2}\log 2\pi + \left(m + \frac{1}{2}\right)\log m - m + o(1)
\end{aligned}
$$

であるので、

$$
\begin{aligned}
&\log \sigma \begin{pmatrix} n \\ x \end{pmatrix} p^x q^{n-x} \\
&= \frac{1}{2}\log npq + \left\{\frac{1}{2}\log 2\pi + \left(n + \frac{1}{2}\right)\log n - n + o(1)\right\} \\
&\quad - \left\{\frac{1}{2}\log 2\pi + \left(x + \frac{1}{2}\right)\log x - x + o(1)\right\} \\
&\quad - \left\{\frac{1}{2}\log 2\pi + \left(n - x + \frac{1}{2}\right)\log(n-x) - (n-x) + o(1)\right\} + x\log p + (n-x)\log q \\
&= \frac{1}{2}\log npq - \frac{1}{2}\log 2\pi - \frac{1}{2}\log n + \left(x + \frac{1}{2}\right)\log\frac{n}{x} + \left(n - x + \frac{1}{2}\right)\log\frac{n}{n-x} \\
&\quad + x\log p + (n-x)\log q + o(1) \\
&= -\frac{1}{2}\log 2\pi + \frac{1}{2}\log p\frac{n}{x} + \frac{1}{2}\log q\frac{n}{n-x} + x\log p\frac{n}{x} + (n-x)\log q\frac{n}{n-x} + o(1)
\end{aligned}
$$

となる。ここで、上で見たように

$$
\frac{x}{n} = p + u\sqrt{\frac{pq}{n}}, \quad \frac{n-x}{n} = q - u\sqrt{\frac{pq}{n}}
$$

であるので、

$$
\log p\frac{n}{x} \to \log 1 = 0, \quad \log q\frac{n}{n-x} \to \log 1 = 0
$$

となるので

$$
\begin{aligned}
&\log \sigma \begin{pmatrix} n \\ x \end{pmatrix} p^x q^{n-x} \\
&= \quad -\frac{1}{2}\log 2\pi - x\log \frac{1}{p}\frac{x}{n} - (n-x)\log \frac{1}{q}\frac{n-x}{n} + o(1) \\
&= \quad -\frac{1}{2}\log 2\pi - x\log \left(1 + u\sqrt{\frac{q}{np}}\right) - (n-x)\log \left(1 - u\sqrt{\frac{p}{nq}}\right) + o(1)
\end{aligned}
$$

となる。テイラー展開

$$
\log(1+x) = x - \frac{x^2}{2} + O(x^3) \quad (x \to 0)
$$

を用いると、

$$
\begin{aligned}
&x\log\left(1 + u\sqrt{\frac{q}{np}}\right) = (np + u\sqrt{npq})\left\{u\sqrt{\frac{q}{np}} - \frac{1}{2}u^2\frac{q}{np} + O\left(\frac{1}{n\sqrt{n}}\right)\right\} \\
&= \quad u\sqrt{npq} + u^2 q - \frac{u^2}{2}q + O\left(\frac{1}{\sqrt{n}}\right) + O\left(\frac{1}{n}\right) \\
&= \quad u\sqrt{npq} + \frac{u^2}{2}q + O\left(\frac{1}{\sqrt{n}}\right) \\
&(n-x)\log\left(1 - u\sqrt{\frac{p}{nq}}\right) = (nq - u\sqrt{npq})\left\{-u\sqrt{\frac{p}{nq}} - \frac{1}{2}u^2\frac{p}{nq} + O\left(\frac{1}{n\sqrt{n}}\right)\right\} \\
&= \quad -u\sqrt{npq} + u^2 p - \frac{u^2}{2}p + O\left(\frac{1}{\sqrt{n}}\right) + O\left(\frac{1}{n}\right) \\
&= \quad -u\sqrt{npq} + \frac{u^2}{2}p + O\left(\frac{1}{\sqrt{n}}\right)
\end{aligned}
$$

ゆえに

$$
\begin{aligned}
&\log \sigma \begin{pmatrix} n \\ x \end{pmatrix} p^x q^{n-x} = -\frac{1}{2}\log 2\pi - \frac{u^2}{2}(p+q) + O\left(\frac{1}{\sqrt{n}}\right) + o(1) \\
&\longrightarrow \quad -\frac{1}{2}\log 2\pi - \frac{u^2}{2} = \log \frac{1}{\sqrt{2\pi}} e^{-u^2/2}
\end{aligned}
$$

∎