de Moivre-Laplaceの定理

Theorem. 確率変数 X が二項分布 $B(n,p)$ に従う, つまり $P(X=k)=\binom{n}{k}p^k q^{n-k}$ $(q=1-p)$ で表現されるとき

$$\lim_{n\to\infty} P\left(\alpha \le \frac{X-np}{\sqrt{npq}} \le \beta\right) = \frac{1}{\sqrt{2\pi}}\int_\alpha^\beta e^{-\frac{x^2}{2}}\,dx$$

が成り立つ.

Proof. n を十分大きな自然数とし, 整数 α_n, β_n を

$$\alpha_n = \min\{k\in\{0,\dots,n\} \mid np+\alpha\sqrt{npq} \le k \le np+\beta\sqrt{npq}\}$$
$$\beta_n = \max\{k\in\{0,\dots,n\} \mid np+\alpha\sqrt{npq} \le k \le np+\beta\sqrt{npq}\}$$

とする. このとき

$$P\left(\alpha \le \frac{X-np}{\sqrt{npq}} \le \beta\right) = \sum_{k=\alpha_n}^{\beta_n} P(X=k) = \sum_{\alpha_n\le k\le\beta_n}\binom{n}{k}p^k q^{n-k}$$

が成り立つ. ここで, $\delta_k = k-np$ とすると $\delta_k \ge 0$ であり, $k=np+\delta_k, n-k=nq-\delta_k$ となるから

$$k \ge \alpha_n \ge n\left(p+\frac{\alpha\sqrt{pq}}{\sqrt{n}}\right) > \frac{p}{2}n$$
$$n-k \ge n-\beta_n \ge n\left(q-\frac{\beta\sqrt{pq}}{\sqrt{n}}\right) > \frac{q}{2}n$$

となることがわかる. 以上より, $c=\min\left\{\frac{p}{2},\frac{q}{2}\right\}$ とすると, $\alpha_n\le k\le\beta_n$ に対して

$$cn\le k, cn\le n-k, \alpha\sqrt{npq}\le\alpha_n-np\le\delta_k\le\beta_k-np\le\beta\sqrt{npq}$$

が成り立つ. よって, $\frac{(\alpha\sqrt{pq})^3}{\sqrt{n}} \le \frac{\delta_k^3}{n^2} \le \frac{(\beta\sqrt{pq})^3}{\sqrt{n}}$ より, $\max_{\alpha_n\le k\le\beta_n}\frac{|\delta_k^3|}{n^2} \le \frac{C_1}{\sqrt{n}}$ (C_1は定数) が成り立ち, 同様に, $\max_{\alpha_n\le k\le\beta_n}\frac{|\delta_n|}{n} \le \frac{C_2}{\sqrt{n}}$ (C_2は定数) が成り立つ.

これらの式と, Stirling の公式 $n! = \sqrt{2\pi}n^{n+\frac{1}{2}}e^{-n+r_n}$, $\frac{1}{12n+1} < r_n < \frac{1}{12n}$ を用いると

$$\begin{aligned}
\binom{n}{k}p^k q^{n-k} &= \frac{n!}{k!(n-k)!}p^k q^{n-k} \\
&= \frac{\sqrt{2\pi}n^{n+\frac{1}{2}}e^{-n+r_n}}{\sqrt{2\pi}k^{k+\frac{1}{2}}e^{-k+r_k}\cdot\sqrt{2\pi}(n-k)^{n-k+\frac{1}{2}}e^{-(n-k)+r_{n-k}}}p^k q^{n-k} \\
&= \frac{1}{2\pi}\sqrt{\frac{n}{k(n-k)}}\left(\frac{np}{k}\right)^k\left(\frac{nq}{n-k}\right)^{n-k}e^{\theta_{n,k}}\ (\theta_{n,k}=r_n-r_k-r_{n-k}) \\
|\theta_{n,k}| &\le \frac{1}{12}\left(\frac{1}{n}+\frac{1}{k}+\frac{1}{n-k}\right) \le \frac{C_3}{n}\ (C_3\text{は定数})
\end{aligned}$$

が成り立つことがわかる. また

$$\sqrt{\frac{n}{k(n-k)}} = \sqrt{\frac{n}{(np+\delta_k)(nq-\delta_k)}} = \frac{1}{\sqrt{npq}}\cdot\frac{1}{\sqrt{\left(1+\frac{\delta_k}{np}\right)\left(1-\frac{\delta_k}{nq}\right)}}$$

であり, $x=0$ 付近で $\dfrac{1}{\sqrt{1+x}}=1-\dfrac{1}{2}x+\dfrac{3}{8}x^2-\cdots$ が成り立つことを用いると

$$\begin{aligned}\frac{1}{\sqrt{\left(1+\frac{\delta_k}{np}\right)\left(1-\frac{\delta_k}{nq}\right)}}&=\left(1+O\left(\frac{\delta_k}{n}\right)\right)\left(1+O\left(\frac{\delta_k}{n}\right)\right)\\&=1+2O\left(\frac{\delta_k}{n}\right)+O\left(\frac{\delta_k^2}{n^2}\right)\\&=1+O\left(\frac{\delta_k}{n}\right)\\\sqrt{\frac{n}{k(n-k)}}&=\frac{1}{\sqrt{npq}}\left(1+O\left(\frac{\delta_k}{n}\right)\right)\end{aligned}$$

となる. 次に, $x=0$ 付近で $\log(1+x)=x-\dfrac{1}{2}x^2+\dfrac{1}{3}x^3-\cdots$ が成り立つことを用いると

$$\begin{aligned}\log\left(\left(\frac{np}{k}\right)^k\left(\frac{nq}{n-k}\right)^{n-k}\right)&=-k\log\frac{k}{np}-(n-k)\log\frac{n-k}{nq}\\&=-(np+\delta_k)\log\left(1+\frac{\delta_k}{np}\right)-(nq-\delta_k)\log\left(1-\frac{\delta_k}{nq}\right)\\&=-(np+\delta_k)\left\{\frac{\delta_k}{np}-\frac{1}{2}\left(\frac{\delta_k}{np}\right)^2+O\left(\left(\frac{\delta_k}{n}\right)^3\right)\right\}\\&\qquad-(nq-\delta_k)\left\{-\frac{\delta_k}{nq}-\frac{1}{2}\left(\frac{\delta_k}{nq}\right)^2+O\left(\left(\frac{\delta_k}{n}\right)^3\right)\right\}\\&=-\frac{1}{2}\frac{\delta_k^2}{n}\left(\frac{1}{p}+\frac{1}{q}\right)+O\left(\frac{\delta_k^3}{n^2}\right)\\&=-\frac{\delta_k^2}{2npq}+O\left(\frac{\delta_k^3}{n^2}\right)\end{aligned}$$

となる. よって

$$\binom{n}{k}p^kq^{n-k}=\frac{1}{\sqrt{2\pi}}\cdot\frac{1}{\sqrt{npq}}e^{-\frac{\delta_k^2}{2npq}}E_{n,k}\qquad\left(E_{n,k}=\left(1+O\left(\frac{\delta_k}{n}\right)\right)e^{O\left(\frac{\delta_k^3}{n^2}\right)}e^{O\left(\frac{1}{n}\right)}\right)$$

と変形できる. ここで, $x=0$ 付近で $e^x=1+x+\dfrac{1}{2}x^2+\cdots$ が成り立つことを用いると

$$\begin{aligned}E_{n,k}&=\left(1+O\left(\frac{\delta_k}{n}\right)\right)\left(1+O\left(\frac{\delta_k^3}{n^2}\right)\right)\left(1+O\left(\frac{1}{n}\right)\right)\\&=1+R_{n,k}\end{aligned}$$

となり, $|R_{n,k}|\le\dfrac{C_4}{\sqrt{n}}$ (C_4は定数) がわかる.

$\phi(x)=\dfrac{1}{\sqrt{2\pi}}e^{-\frac{x^2}{2}}, x_k=\dfrac{\delta_k}{\sqrt{npq}}=\dfrac{k-np}{\sqrt{npq}}$ とすると

$$P\left(\alpha\le\frac{X-np}{\sqrt{npq}}\le\beta\right)=\sum_{\alpha_n\le k\le\beta_n}\frac{1}{\sqrt{npq}}\phi(x_k)(1+R_{n,k})=\sum_{\alpha_n\le k\le\beta_n}\frac{1}{\sqrt{npq}}\phi(x_k)+\sum_{\alpha_n\le k\le\beta_n}\frac{1}{\sqrt{npq}}\phi(x_k)R_{n,k}$$

が成り立つ. ここで, 第 2 項について

$$\begin{aligned}\left|\sum_{\alpha_n\le k\le\beta_n}\frac{1}{\sqrt{npq}}\phi(x_k)R_{n,k}\right|&\le\sum_{\alpha_n\le k\beta_n}\frac{1}{\sqrt{npq}}\cdot\frac{1}{\sqrt{2\pi}}\cdot\frac{C_4}{\sqrt{n}}\\&=\frac{\beta_n-\alpha_n+1}{n\sqrt{2\pi pq}}\\&\le\frac{(\beta-\alpha)\sqrt{npq}+1}{n\sqrt{2\pi pq}}\\&\to 0\ (n\to\infty)\end{aligned}$$

よって

$$P\left(\alpha \leq \frac{X-np}{\sqrt{npq}} \leq \beta\right) \to \int_{\alpha}^{\beta} \phi(x)\,dx = \frac{1}{\sqrt{2\pi}}\int_{\alpha}^{\beta} e^{-\frac{x^2}{2}}\,dx \ (n \to \infty)$$

が成り立つ.

以上より, 定理が示された. ■