

二項分布

$$\frac{n!}{k!l!} p^k q^l \text{ (ただし } p+q=1, k+l=n \text{)}$$

が、 $n$  が巨大であり、 $k$  が  $np$  に近いとき、正規分布

$$\frac{1}{\sqrt{2\pi npq}} e^{-\frac{(k-np)^2}{2npq}}$$

に近づくことを納得したい。

前提とする事柄は、Stirling の公式と、「 $h$  が小さいとき  $1+h \approx e^{h(1-\frac{h}{2})}$ 」なる近似である。後者は  $\ln(1+h) = h - (h^2/2) + \dots$  という展開から頷けるであろう。

Stirling の公式  $t! \approx \sqrt{2\pi t}(t/e)^t$  により

$$\frac{n!}{k!l!} p^k q^l \approx \frac{\sqrt{2\pi n}(n/e)^n}{\sqrt{2\pi k}(k/e)^k \sqrt{2\pi l}(l/e)^l} p^k q^l$$

$\sqrt{\quad}$  の部分をまとめ、さらに  $(n/e)^n$  の肩の上の  $n$  にだけ  $n = k+l$  を用いて、 $k$  乗、 $l$  乗ごとに整理すると、 $e$  は約されて消え、下のように整理される。

$$\sqrt{\frac{n}{2\pi kl}} \left(\frac{np}{k}\right)^k \left(\frac{nq}{l}\right)^l$$

先頭の  $\sqrt{\quad}$  部は、 $k \approx np$ ,  $l \approx nq$  と考えて置き換えれば (強引!)  $1/\sqrt{2\pi npq}$  と見ることができ、正規分布の式の前半部に一致することが分かる。

残りの部分を見るための準備として、 $k:l$  と  $p:q$  とのわずかなズレを  $\Delta$  とすれば、 $k = n(p + \Delta)$ ,  $l = n(q - \Delta)$  と書くことができる。この  $\Delta$  はあとで使うことにして、さらに  $k = n(p + \Delta) = np(1+h)$  (つまり  $h = \Delta/p$ ) と置き、 $(np/k)^k$  の部分を変形してみよう。

$$\left(\frac{np}{k}\right)^k = \left[\frac{np}{np(1+h)}\right]^{np(1+h)} = (1+h)^{-np(1+h)}$$

ここで  $1+h \approx e^{h(1-\frac{h}{2})}$  を用いれば

$$\approx e^{h(1-\frac{h}{2})[-np(1+h)]}$$

指数部を計算すると

$$-nph \left(1 - \frac{h}{2}\right) (1+h) = -nph \left(1 + \frac{h}{2} - \frac{h^2}{2}\right)$$

$h^2$  の項を落とし、 $h = \Delta/p$  と置いたことを思い出して書き戻せば

$$\approx -n\Delta \left(1 + \frac{\Delta}{2p}\right)$$

となり、結局

$$\left(\frac{np}{k}\right)^k \approx e^{-n\Delta(1+\frac{\Delta}{2p})}$$

が得られる。まったく同様にして

$$\left(\frac{nq}{l}\right)^l \approx e^{-n\Delta(-1+\frac{\Delta}{2q})}$$

が得られるので、これらの積は

$$\left(\frac{np}{k}\right)^k \left(\frac{nq}{l}\right)^l \approx e^{-n\Delta(1+\frac{\Delta}{2p})-n\Delta(-1+\frac{\Delta}{2q})}$$

となる。再び指数部を取り出し、 $p + q = 1$  を用いて整理すると

$$\begin{aligned} -n\Delta \left(1 + \frac{\Delta}{2p}\right) - n\Delta \left(-1 + \frac{\Delta}{2q}\right) &= -n\Delta \left(\frac{\Delta}{2p} + \frac{\Delta}{2q}\right) \\ &= -\frac{n\Delta^2}{2} \left(\frac{1}{p} + \frac{1}{q}\right) = -\frac{n\Delta^2}{2} \cdot \frac{p+q}{pq} = -\frac{n\Delta^2}{2pq} = -\frac{(n\Delta)^2}{2npq} \end{aligned}$$

$n\Delta = k - np$  であることに注意すれば、正規分布の式が得られていることに気付くであろう。