2 項分布と正規分布 (Binomial and Gaussian distributions)

2 項分布 (binomial distribution)$B(n,\ p)$ は、

$$p(x) = {}_nC_x p^x q^{(n-x)} = \frac{n!}{x!(n-x)!} p^x q^{(n-x)}, \qquad \text{但し、} q = 1 - p \tag{1}$$

と表される。ここで、変数 n, x $(n \gg x)$ が十分大きい場合を考える。

見通しをよくするために、(1) 式において $y = n - x$ と置いて、y についての和をとる。

$$p(x) = \sum_y \frac{n!}{x!y!}\, p^x q^y\, \delta_{x+y,\ n}$$

ここで、$\delta_{x+y,\ n}$ は、クロネッカーのデルタである。スターリングの公式

$$n! = \sqrt{2\pi n}\, n^n e^{-n}, \quad x! = \sqrt{2\pi x}\, x^x e^{-x}, \quad y! = \sqrt{2\pi y}\, y^y e^{-y}$$

を用いて、$p(x)$ を書き直すと、

$$\begin{aligned} p(x) &\sim \sum_y \frac{1}{\sqrt{2\pi n}} \left(\frac{x}{n}\right)^{-x-\frac{1}{2}} \left(\frac{y}{n}\right)^{-y-\frac{1}{2}} p^x q^y\, \delta_{x+y,\ n} \\ &= \sum_y \frac{1}{\sqrt{2\pi npq}} \left(\frac{x}{np}\right)^{-x-\frac{1}{2}} \left(\frac{y}{nq}\right)^{-y-\frac{1}{2}} \delta_{x+y,\ n} \end{aligned}$$

ここで、$x = t + np$、$y = u + nq$ とおいて、変数 x、y の代わりに t、u を用いることにすると、

$$p(t+np) = f(t) \sim \sum_u \frac{1}{\sqrt{2\pi npq}} \left(1 + \frac{t}{np}\right)^{-\left(np+t+\frac{1}{2}\right)} \left(1 + \frac{u}{nq}\right)^{-\left(nq+u+\frac{1}{2}\right)} \delta_{t+u,0}$$

さらに、$np \gg t \gg 1/2$、$np \gg u \gg 1/2$ として展開すると、

$$\begin{aligned} \left(1 + \frac{t}{np}\right)^{-\left(np+t+\frac{1}{2}\right)} &= \exp\left\{-\left(np + t + \frac{1}{2}\right) \ln\left(1 + \frac{t}{np}\right)\right\} \\ &= \exp\left\{-\left(np + t + \frac{1}{2}\right) \left(\frac{t}{np} - \frac{t^2}{2(np)^2} + \cdots\right)\right\} \\ &\simeq \exp\left(-t - \frac{t^2}{2np}\right) \end{aligned}$$

同様に、

$$\left(1 + \frac{u}{nq}\right)^{-\left(nq+u+\frac{1}{2}\right)} \simeq \exp\left(-u - \frac{u^2}{2nq}\right)$$

と近似できるので、

$$f(t) = \sum_u \delta_{t+u,0} \frac{1}{\sqrt{2\pi npq}} \exp\left(-(t+u) - \frac{t^2}{2np} - \frac{u^2}{2nq}\right)$$

と書ける。もはや t も u も連続変数として扱えるので、$\sum_u \delta_{t+u,0}$ も $\int du\, \delta(t+u)$ で置き換える。ここで、$\delta(t+u)$ は、ディラックのデルタ関数である。すると、

$$\begin{aligned} f(t) &= \int du\, \delta(t+u) \frac{1}{\sqrt{2\pi npq}} \exp\left(-(t+u) - \frac{t^2}{2np} - \frac{u^2}{2nq}\right) \\ &= \frac{1}{\sqrt{2\pi npq}} \exp\left(-\frac{t^2}{2np} - \frac{t^2}{2nq}\right) \\ &= \frac{1}{\sqrt{2\pi npq}} \exp\left(-\frac{t^2}{2npq}\right) \end{aligned}$$

となる。最後に、$\sigma^2 = npq$ と置いて、

$$f(t) = \frac{1}{\sqrt{2\pi\sigma^2}} \exp\left(-\frac{t^2}{2\sigma^2}\right)$$

これは、平均値が 0, 分散が σ^2 の正規分布 (ガウス分布) である。