二項分布の最頻値

二項分布の最頻値（モード）について調べる．一般的な 2 項分布について次のような関数を考える．

$$f(r) ={}_n C_r\, p^r(1-p)^{n-r} \tag{1}$$

$f(r)$ が最大となるのは r がどのような値のときか，つまり $f(r)$ の増減を調べればよいわけであるが，$f(r)$ を微分するというのはいかにも無謀である．また

$$f(r) - f(r-1) \geqq 0 \tag{2}$$

という不等式を解き，その解の中で最大の r を求めるという方法も考えられる．しかしこの計算は通分と因数分解がわずらわしい．関数の値は全て正なので，

$$\frac{f(r)}{f(r-1)} \geqq 1 \tag{3}$$

という不等式を解くことにする．この計算は約分するだけだから楽である．(3) の分子は

$$\frac{n!}{r!(n-r)!} \cdot p^r(1-p)^{n-r} \tag{4}$$

分母は

$$\frac{n!}{(r-1)!(n-r+1)!} \cdot p^{r-1}(1-p)^{n-r+1} \tag{5}$$

であるから (4) ÷ (5) を行うことにより，(3) は

$$\frac{(n-r+1)p}{r(1-p)} \geqq 1 \tag{6}$$

分母は正なので，分母をはらって，

$$\begin{aligned}(n-r+1)p &\geqq r(1-p)\\ np-pr+p &\geqq r-rp\\ r &\geqq (n+1)p \end{aligned} \tag{7}$$

となり，この不等式を満たす最大の整数値が $f(r)$ に最頻値を与える値である．ただし (7) の等号が成り立つ場合は注意が必要で，答えは 2 つある．(r と $r-1$)

例　$n = 40, p = \dfrac{1}{7}$ の場合,(7) は

$$r \leqq \frac{41}{7} = 5.85714286$$

であるから，$r = 5$ が最頻値を与える．p が同じ値で $n = 41$ のとき，$r = 5$ と $r = 6$ の両方が答えとなる．