（0-1）分布参数的区间估计

已知(0-1)分布的均值和方差分别为

$$\mu = p,\quad \sigma^2 = p(1-p)$$

设 $X_1, X_2, \cdots, X_n$ 是一个样本，因样本容量 n 比较大，由中心极限定理，知

$$\frac{\sum_{i=1}^{n} X_i - np}{\sqrt{np(1-p)}} = \frac{n\bar{X} - np}{\sqrt{np(1-p)}}$$

近似地服从 $N(0,1)$ 分布，于是有

$$P\left\{ -z_{\alpha/2} < \frac{n\bar{X} - np}{\sqrt{np(1-p)}} < z_{\alpha/2} \right\} \approx 1-\alpha$$

而不等式

$$-z_{\alpha/2} < \frac{n\bar{X} - np}{\sqrt{np(1-p)}} < z_{\alpha/2} \tag{1}$$

等价于

$$(n + z_{\alpha/2}^2)p^2 - (2n\bar{X} + z_{\alpha/2}^2)p + n\bar{X}^2 < 0$$

记

$$p_1 = \frac{1}{2a}(-b - \sqrt{b^2 - 4ac})$$
$$p_2 = \frac{1}{2a}(-b + \sqrt{b^2 - 4ac})$$

此处 $a = n + z_{\alpha/2}^2, b = -(2n\bar{X} + z_{\alpha/2}^2), c = n\bar{X}^2$

于是由(1)式得 p 的一个近似的置信水平为 $1-\alpha$ 的置信区间为

$$(p_1, p_2)$$