< 母比率の区間推定 1 >

ある集団の内閣支持率を調べる場合、母集団は支持する人と支持しない人に分かれる。この比率 $\dfrac{\text{支持する人の人数}}{\text{母集団の全人数}} = p$ を**母比率**と呼ぶ。この母比率を求めたい。この母集団から n 人を復元抽出し、

$$X_k = \begin{cases} 1 & : \ k \text{ 番目に選んだ人が支持者の場合} \\ 0 & : \ \text{そうでないとき} \end{cases}$$

とおくと、和 $\displaystyle\sum_{k=1}^{n} X_k$ は 2 項分布 $B(n,p)$ に従う確率変数である。従って平均 $\overline{X} = \dfrac{1}{n}\displaystyle\sum_{k=1}^{n} X_k$ は $E\left[\,\overline{X}\,\right] = p,\ E\left[\left(\overline{X} - p\right)^2\right] = \dfrac{p(1-p)}{n}$ であるので、その標準化 $\overline{X}^* = \left(\overline{X} - p\right)/\sqrt{p(1-p)/n}$ は近似的に標準正規分布に従う。すなわち

$$P\left(\left|\frac{\overline{X} - p}{\sqrt{\frac{p(1-p)}{n}}}\right| \leqq \alpha\right) \doteqdot \int_{-\alpha}^{\alpha} \frac{1}{\sqrt{2\pi}} e^{-\frac{x^2}{2}} dx = f(\alpha)$$

である。従って信頼度 $f(\alpha)$ の信頼区間は

$$\boxed{\overline{X} - \alpha\sqrt{\frac{p(1-p)}{n}} \leqq p \leqq \overline{X} + \alpha\sqrt{\frac{p(1-p)}{n}}} \qquad (****)$$

である。この式 $(****)$ は p に関する 2 次不等式であり、これを解くと

$$\boxed{\frac{b - \sqrt{D}}{2a} \leqq p \leqq \frac{b + \sqrt{D}}{2a}} \qquad (****)'$$

$$\left(a = \frac{1}{n} + \frac{1}{\alpha^2}\ ,\quad b = \frac{1}{n} + \frac{2\overline{X}}{\alpha^2}\ ,\quad c = \frac{\overline{X}^2}{\alpha^2}\ ,\quad D = b^2 - 4ac\right)$$

である。

(注) n が十分大きい場合は $(****)'$ のかわりに $(****)$ 式の $p(1-p)$ を $\overline{X}(1-\overline{X})$ で代用した式

$$\overline{X} - \alpha\sqrt{\frac{\overline{X}(1-\overline{X})}{n}} \leqq p \leqq \overline{X} + \alpha\sqrt{\frac{\overline{X}(1-\overline{X})}{n}}$$

を用いることもある。