

母比率の推定

母集団から抽出した標本をもとに母集団の比率（母比率）を区間推定する

- 選挙前の候補者の支持率（支持する／支持しない）を推定できる
- 好き嫌いのようなアンケート調査から全体の傾向を推定できる

正規分布による近似（標本数の多い場合）

- 二項分布（ある事象が起こるか起こらないかの確率の分布）は、試行回数 n が十分大きい場合、正規分布に近似できることを利用
- 母集団のある事象について、 n 回の試行（標本の大きさが n ）の標本の比率（標本比率）を $\bar{p}$ とするとき
- 母比率 p の信頼度 $100(1-\alpha)\%$ の信頼区間は次のとおり

$$\bar{p} - z_{(\alpha/2)} \sqrt{\frac{\bar{p}(1-\bar{p})}{n}} \leq p \leq \bar{p} + z_{(\alpha/2)} \sqrt{\frac{\bar{p}(1-\bar{p})}{n}}$$

- なお、標本比率を $\bar{p} = \frac{X}{n}$ とすると、その平均 $E(\bar{p})$ と分散 $V(\bar{p})$ は、次のようになる

$$\begin{aligned} E(\bar{p}) &= E\left(\frac{X}{n}\right) = \frac{1}{n}E(X) = \frac{1}{n}np \\ &= p \end{aligned}$$

$$\begin{aligned} V(\bar{p}) &= V\left(\frac{X}{n}\right) = \frac{1}{n^2}V(X) = \frac{1}{n^2}np(1-p) \\ &= \frac{p(1-p)}{n} \end{aligned}$$

正規分布による近似（標本数の少ない場合）

- 大きさが N の母集団のある事象について、大きさが n の標本の比率（標本比率）を $\bar{p}$ とするとき
- 母比率 p の信頼度 $100(1-\alpha)\%$ の信頼区間は次のとおり

$$\bar{p} - z_{(\alpha/2)} \sqrt{\frac{\bar{p}(1-\bar{p})}{n}} \sqrt{\frac{N-n}{N-1}} \leq p \leq \bar{p} + z_{(\alpha/2)} \sqrt{\frac{\bar{p}(1-\bar{p})}{n}} \sqrt{\frac{N-n}{N-1}}$$

- 最初の式（標本数が多い場合の式）より近似の精度が良い（母比率に近い値になる）

F分布から算出（標本数の少ない場合）

- 母集団のある事象について、 n 回の試行（標本の大きさが n ）の標本の比率（標本比率）を $\frac{x}{n}$ とするとき
- 母比率 p の信頼度 $100(1-\alpha)\%$ の信頼区間は次のとおり

- 信頼上限：

$$\frac{m_1 F_U}{m_1 F_U + m_2}, \quad m_1 = 2(x+1), \quad m_2 = 2(n-x)$$

- 第1自由度 m_1 、第2自由度 m_2 に対応するF分布の値を F_U とする

- 信頼下限：

$$\frac{n_2 F_L}{n_1 F_L + n_2}, \quad n_1 = 2(n - x + 1), \quad n_2 = 2x$$

- 第1自由度 n_1 、第2自由度 n_2 に対応するF分布の値を F_L とする