母比率の検定

母比率の検定では、「**母比率と標本比率との差の程度**」を調べる。

- 帰無仮説 H_0 は「母比率と標本比率が等しい」: $p=\hat{p}$
- 対立仮説 H_1 は「母比率と標本比率が等しくない」:
  - 両側検定の場合は $p \neq \hat{p}$
  - 片側検定の場合は $p < \hat{p}$ または $p > \hat{p}$

二項検定

- 二項定理を使って、母比率に対する標本比率の統計値を直接計算し、有意水準と比較する
- 理論的には正当な方法だが計算が複雑なため、コンピュータによる統計処理が登場するまでは、正規分布に近似する方法(後述)などが使われていた。

考え方、

- 母比率 p_0 の事象を、n 回試行するとき、
- r 回起きる確率は次のようになる

$$P_r = {}_nC_r {p_0}^r (1-p_0)^{n-r}$$

- 「r 回以上」起きる確率は、次のような確率の和から算出できる

$$\begin{aligned} P &= {}_nC_r {p_0}^r (1-p_0)^{n-r} + {}_nC_{r+1} {p_0}^{r+1} (1-p_0)^{n-r+1} + \cdots + {}_nC_n {p_0}^n (1-p_0)^0 \\ &= \sum_{i=r}^{n} {}_nC_i {p_0}^i (1-p_0)^{n-i} \\ &= 1 - \sum_{i=0}^{r-1} {}_nC_i {p_0}^i (1-p_0)^{n-i} \end{aligned}$$

- 算出した確率(P値)と有意水準を比較する
  - 片側検定
    - 帰無仮説 H_0 を棄却: $P < \alpha$
    - 帰無仮説 H_0 を採択: $P \geq \alpha$
  - 両側検定
    - 帰無仮説 H_0 を棄却: $2P < \alpha$
    - 帰無仮説 H_0 を採択: $2P \geq \alpha$

正規分布に近似

- 特定の条件の時にだけ使える方法である
  - 標本数を n 、母比率を p_0 とするとき、
  - $np_0 > 5,\ p_0 < 1-p_0$ 、または、 $n > 25$ の場合

考え方

- 標本の値を x 、標本比率を $\hat{p_0} = \frac{x}{n}$ とする
- 標準正規分布に近似される、検定統計量 z_0 を次の式から算出する

$$z_0 = \frac{\widehat{p_0} - p_0}{\sqrt{\frac{p_0(1-p_0)}{n}}}$$

- 検定統計量 z_0 を使って判定をする

連続補正をする場合

- 二項分布は離散型の分布であるため、正規分布のような連続型の分布に近似すると、その精度はあまりよくない
- そこで、連続補正（イエーツ(Yates)の補正）をすることで、精度をよくする

$$z_{0c} = \frac{|x - np_0| - 0.5}{\sqrt{np_0(1-p_0)}}$$

- 補正した検定統計量 z_{0c} を使って判定をする