統計的な推測【母比率の推定】　p.98〜99

1 学習内容の説明 ⇒ 2 問題演習 ⇒ 3 振り返り(確認テスト・相互採点・リフレクションの記入)

【態度目標】しゃべる、質問する、説明する、動く、協力する、貢献する

【内容目標】前回の話を生かして標本比率から母比率を推定しよう

□母比率の推定

標本比率から母比率を推定する方法を調べよう。

ある特性 A の母比率が p である十分大きな母集団から，大きさ n の無作為標本を抽出し，それらに対して，X_1，X_2，……，X_n の値を特質の有無で 1 か 0 かを定めると，$T=X_1+X_2+\cdots\cdots+X_n$ は大きさ n の標本の中で特性 A をもつものの個数を表す確率変数であり，標本比率は $R=\dfrac{T}{n}$ である。

T は二項分布 $B(n,\ p)$ に従うから，n が大きいとき，近似的に正規分布 $N(np,\ np(1-p))$ に従う。

94ページ

母平均の推定の場合と同様に，

96ページの母平均の推定と同様にして…

$$P\left(-1.96\leqq\frac{T-np}{\sqrt{np(1-p)}}\leqq1.96\right)=0.95 \longleftarrow \boxed{Z=\frac{X-(\text{平均})}{(\text{標準偏差})}}$$

ここで，n が十分に大きいならば大数の法則により 標本比率 R は母比率 p に近いとみなしてよい。よって，根号の中の母比率 p の代わりに標本平均 R を用いてよいことが知られているから，次のように変形できる。

$$P\left(-1.96\leqq\frac{\dfrac{T}{n}-p}{\sqrt{\dfrac{R(1-R)}{n}}}\leqq1.96\right)=0.95 \quad\cdots\cdots ①$$

$$P\left(\frac{T}{n}-1.96\cdot\sqrt{\frac{R(1-R)}{n}}\leqq p\leqq\frac{T}{n}+1.96\cdot\sqrt{\frac{R(1-R)}{n}}\right)=0.95$$

$$P\left(R-1.96\cdot\sqrt{\frac{R(1-R)}{n}}\leqq p\leqq R+1.96\cdot\sqrt{\frac{R(1-R)}{n}}\right)=0.95 \quad \boxed{R=\frac{T}{n}}$$

すなわち，標本調査で得られる 1 つの標本比率 R を用いて，母比率は次の区間内にあると推定され

$$R-1.96\cdot\sqrt{\frac{R(1-R)}{n}}\leqq p\leqq R+1.96\cdot\sqrt{\frac{R(1-R)}{n}}$$

の成り立つ確率は 0.95 である。

標本比率を用いて、
母集団の比率を推定する！

> **母比率の推定**
>
> 標本の大きさ n が大きいとき，標本比率を R とすると，母比率 p に対する信頼度 95 %の信頼区間は
>
> $$\left[R-1.96\sqrt{\frac{R(1-R)}{n}},\ \ R+1.96\sqrt{\frac{R(1-R)}{n}}\right]$$

※ 信頼度 99 %では

$$R-2.58\sqrt{\frac{R(1-R)}{n}}\leqq p\leqq R+2.58\sqrt{\frac{R(1-R)}{n}}$$

統計的な推測【母比率の推定】　p.98～99

例題６） ある世論調査で，有権者から無作為抽出した400人についてA政党の支持者を調べたら144人いた。A政党の支持者の母比率 p に対して，信頼度95％の信頼区間を求めよ。

解答　標本比率 R は　$R=\dfrac{144}{400}=0.36$

$n=400$ であるから

$$1.96\sqrt{\frac{R(1-R)}{n}}=1.96\sqrt{\frac{0.36\times0.64}{400}}=1.96\times0.024\fallingdotseq0.047$$

よって，求める信頼区間は

$[0.36-0.047,\ 0.36+0.047]$

すなわち　$[0.313,\ 0.407]$

分数や指数の活用も考えた方がよいかも

$$\frac{196}{100}\sqrt{\frac{1}{400}\cdot\frac{36}{100}\cdot\frac{64}{100}}$$
$$=\frac{196}{100}\sqrt{\frac{1}{100}\cdot\frac{9}{100}\cdot\frac{64}{100}}$$
$$=\frac{196}{100}\cdot\frac{1}{10}\cdot\frac{3}{10}\cdot\frac{8}{10}$$
$$=(200-4)\cdot24\cdot10^{-5}$$
$$=(4800-96)\cdot10^{-5}$$
$$=4704\cdot10^{-5}$$
$$=0.04704$$

信頼度99％とすると　$2.58\sqrt{\dfrac{0.36\times0.64}{400}}=2.58\times0.024\fallingdotseq0.062$

よって，求める信頼区間は　$[0.36-0.062,\ 0.36+0.062]$　すなわち　$[0.298,\ 0.422]$

解説　標本における比率 $\dfrac{T}{n}$ を R とすれば n が十分に大きいとき $R\fallingdotseq p$ と見なしてよいことについて

$\left|\dfrac{T-np}{\sqrt{np(1-p)}}\right|\leqq1.96$ であり，両辺がともに正であることから２乗すると

$\left(\dfrac{T-np}{\sqrt{np(1-p)}}\right)^2\leqq1.96^2$　であるから　$(T-np)^2\leqq1.96^2\cdot np(1-p)$

$T=n\cdot R$ より　$n\cdot(R-p)^2\leqq1.96^2\cdot p(1-p)$　$\therefore(n+1.96^2)p^2-(2nR+1.96^2)p+nR^2\leqq0$

$=0$ とみて解の公式を用いると

$$p=\frac{2nR+1.96^2\pm\sqrt{(2nR+1.96^2)^2-4(n+1.96^2)\cdot nR^2}}{2(n+1.96^2)}$$

$$p=\frac{2nR+1.96^2\pm\sqrt{4nR\cdot1.96^2+1.96^4-4\cdot1.96^2\cdot nR^2}}{2(n+1.96^2)}$$

$$p=\frac{2nR+1.96^2\pm1.96\sqrt{4nR+1.96^2-4nR^2}}{2(n+1.96^2)}$$

$$p=\frac{2R+\dfrac{1.96^2}{n}\pm1.96\sqrt{\dfrac{4R}{n}+\dfrac{1.96^2}{n^2}-\dfrac{4R}{n}}}{2\left(1+\dfrac{1.96^2}{n}\right)}$$

n を次第に大きくすると，n の次数を考慮すると　$\dfrac{1.96^2}{n^2}$ が他より先に0になることから

$p\fallingdotseq\dfrac{2R\pm1.96\sqrt{\dfrac{4R}{n}-\dfrac{4R^2}{n}}}{2}$　より　$p=R\pm1.96\sqrt{\dfrac{R(1-R)}{n}}$

したがって　$R-1.96\sqrt{\dfrac{R(1-R)}{n}}\leqq p\leqq R+1.96\sqrt{\dfrac{R(1-R)}{n}}$