## §7.4　大標本による母比率の区間推定

母比率が $\rho$ である無限2母集団から無作為抽出した標本の大きさ $n$ が充分大きいとします．中心極限定理より導かれる定理5.4より，標本比率を表す確率変数 $R$ は近似的に正規分布 $\mathrm{N}\left(\rho, \dfrac{\rho(1-\rho)}{n}\right)$ に従います．このことを用いて，$0<\alpha<1$ である実数 $\alpha$ に対して，母比率 $\rho$ の信頼度 $\alpha$ の信頼区間を求めます．

標準正規分布 $\mathrm{N}(0,1)$ の累積分布関数を $\Phi(x)$ とおきます：

$$\Phi(x) = \int_{-\infty}^{x} \frac{1}{\sqrt{2\pi}} \exp\left(-\frac{t^2}{2}\right) dt \ .$$

定理7.2より，正規分布に従う確率変数 $X$ の標準偏差 $\sigma[X]$ が分かるとき，$X$ の実現値 $x$ に対して，$X$ の平均値 $\mathrm{E}[X]$ の信頼度 $\alpha$ の信頼区間は

$$\left[x-\sigma[X]\,\Phi^{-1}\left(\frac{1+\alpha}{2}\right),\, x+\sigma[X]\,\Phi^{-1}\left(\frac{1+\alpha}{2}\right)\right] \ .$$

標本比率を表す確率変数 $R$ は近似的に正規分布に従いますから，$R$ の標準偏差 $\sigma[R]$ が分かるとき，$R$ の実現値 $r$ に対して，$R$ の平均値 $\mathrm{E}[R]$ の信頼度 $\alpha$ の信頼区間は

$$\left[r-\sigma[R]\,\Phi^{-1}\left(\frac{1+\alpha}{2}\right),\, r+\sigma[R]\,\Phi^{-1}\left(\frac{1+\alpha}{2}\right)\right] \ .$$

標本比率を表す確率変数 $R$ は母比率 $\rho$ の不偏推定量です：$\mathrm{E}[R]=\rho$ ．これより，$R$ の標準偏差 $\sigma[R]$ が分かるとき，$R$ の実現値 $r$ に対して，母比率 $\rho$ の信頼度 $\alpha$ の信頼区間は

$$\left[r-\sigma[R]\,\Phi^{-1}\left(\frac{1+\alpha}{2}\right),\, r+\sigma[R]\,\Phi^{-1}\left(\frac{1+\alpha}{2}\right)\right] \ .$$

標本比率を表す確率変数 $R$ の標準偏差 $\sigma[R]$ を考えます．$R$ は近似的に正規分布 $\mathrm{N}\left(\rho, \dfrac{\rho(1-\rho)}{n}\right)$ に従いますから，$\mathrm{V}[R] \fallingdotseq \dfrac{\rho(1-\rho)}{n}$ ，$\sigma[R] \fallingdotseq \sqrt{\dfrac{\rho(1-\rho)}{n}}$ ．母比率 $\rho$ の値が分かりません（母比率の値が分かっていれば区間推定などしません）が，ここでは標本の大きさ $n$ が充分に大きい場合を考えていますから，標本比率が母比率とほぼ一致すると考えて差し支えありません．そこで，普通は母比率 $\rho$ を標本比率 $r$ で近似します：$\sigma[R] \fallingdotseq \sqrt{\dfrac{r(1-r)}{n}}$ ．故に，$R$ の実現値 $r$ に対して，母比率 $\rho$ の信頼度 $\alpha$ の信頼区間は

$$\left[r-\sqrt{\frac{r(1-r)}{n}}\,\Phi^{-1}\left(\frac{1+\alpha}{2}\right),\, r+\sqrt{\frac{r(1-r)}{n}}\,\Phi^{-1}\left(\frac{1+\alpha}{2}\right)\right] \ .$$

このようにして次の公式が成り立ちます．

**公式** 7.4　標準正規分布 $\mathrm{N}(0,1)$ の累積分布関数を $\Phi(x)$ とおく．無限2項母集団から無作為抽出した標本の大きさ $n$ が充分大きいとする．標本比率が $r$ であるとき，母比率の信頼度 $\alpha$ の信頼区間は

$$\left[r-\sqrt{\frac{r(1-r)}{n}}\,\Phi^{-1}\left(\frac{1+\alpha}{2}\right),\, r+\sqrt{\frac{r(1-r)}{n}}\,\Phi^{-1}\left(\frac{1+\alpha}{2}\right)\right]$$

で与えられる．

例題 7.4　大きさが約73万の2項母集団から大きさ1000の標本を無作為に非復元抽出する．標本比率が0.6であるとする．母比率の信頼度99%の信頼区間を求める．

母集団の大きさ約73万に対する標本の大きさ1000の比率は非常に小さいので，母集団を無限母集団とみなす．標準正規分布 $\mathrm{N}(0,1)$ の累積分布関数を $\Phi(x)$ とおく．母比率の信頼度99%の信頼区間は，

$$\begin{aligned}
&\left[0.6-\sqrt{\frac{0.6\times 0.4}{1000}}\,\Phi^{-1}\left(\frac{1+0.99}{2}\right),\, 0.6+\sqrt{\frac{0.6\times 0.4}{1000}}\,\Phi^{-1}\left(\frac{1+0.99}{2}\right)\right] \\
&= \left[0.6-\sqrt{\frac{6}{25000}}\,\Phi^{-1}(0.995),\, 0.6+\sqrt{\frac{6}{25000}}\,\Phi^{-1}(0.995)\right] \\
&\fallingdotseq \left[0.6-\frac{\sqrt{15}}{250}\cdot 2.576,\, 0.6+\frac{\sqrt{15}}{250}\cdot 2.576\right] \\
&\fallingdotseq [0.560, 0.640] \ .
\end{aligned}$$

終

問題 7.4　大きさが約38万の2項母集団から大きさ400の標本を無作為に非復元抽出します．標本比率が0.8であるとします．母比率の信頼度95%の信頼区間を求めなさい．