§7.3　大標本による母平均の区間推定

母平均が μ であり母標準偏差が σ である無限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た定理5.2.3より，標本平均を表す確率変数 $\overline{X}$ は近似的に正規分布 $\mathrm{N}\left(\mu, \dfrac{\sigma^2}{n}\right)$ に従います．このことを用いて，$0<\alpha<1$ である実数 α に対して，母平均の信頼度 α の信頼区間を求めます．

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

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

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 α の信頼区間は

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

標本平均を表す確率変数 $\overline{X}$ は近似的に正規分布に従うので，$\overline{X}$ の標準偏差 $\sigma\left[\overline{X}\right]$ が分かるとき，$\overline{X}$ の実現値 $\overline{x}$ に対して，$\overline{X}$ の平均値 $\mathrm{E}\left[\overline{X}\right]$ の信頼度 α の信頼区間は

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

標本平均を表す確率変数 $\overline{X}$ は母平均 μ の不偏推定量です：$\mathrm{E}\left[\overline{X}\right]=\mu$．これより，$\overline{X}$ の標準偏差 $\sigma\left[\overline{X}\right]$ が分かるとき，$\overline{X}$ の実現値 $\overline{x}$ に対して，母平均 μ の信頼度 α の信頼区間は

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

標本平均を表す確率変数 $\overline{X}$ の標準偏差 $\sigma\left[\overline{X}\right]$ を考えます．$\overline{X}$ は近似的に正規分布 $\mathrm{N}\left(\mu, \dfrac{\sigma^2}{n}\right)$ に従いますから，$\mathrm{V}\left[\overline{X}\right] \fallingdotseq \dfrac{\sigma^2}{n}$，$\sigma\left[\overline{X}\right] \fallingdotseq \sqrt{\dfrac{\sigma^2}{n}}$．標本調査では母分散 σ^2 の値は普通は分かりませんが，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，標本の不偏分散が母分散とほぼ一致すると考えて差し支えありません．そこで，普通は母分散 σ^2 を標本の不偏分散 u で近似します：$\sigma\left[\overline{X}\right] \fallingdotseq \sqrt{\dfrac{u}{n}}$．故に，$\overline{X}$ の実現値 $\overline{x}$ に対して，母平均の信頼度 α の信頼区間は

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

このようにして次の公式が成り立ちます．

公式 7.3　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．実数 α について $0<\alpha<1$ とする．標本平均が $\overline{x}$ であり，標本の不偏分散が u であるとき，母平均の信頼度 α の信頼区間は

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

で与えられる．

例題 7.3　大きさが約73万の母集団から大きさ 500 の標本を無作為に非復元抽出する．標本の平均が 570 で不偏分散が 180 であるとする．標本の不偏分散を母分散として代用して，母平均の信頼度 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 500 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．また，標本の大きさ 500 は充分大きいので，標本の不偏分散 180 を母分散とみなす．標準正規分布 $\mathrm{N}(0,1)$ の累積分布関数を $\Phi(x)$ とおく．母平均の信頼度 99% の信頼区間は，

$$\begin{aligned}
&\left[570-\sqrt{\frac{180}{500}}\,\Phi^{-1}\left(\frac{1+0.99}{2}\right),\, 570+\sqrt{\frac{180}{500}}\,\Phi^{-1}\left(\frac{1+0.99}{2}\right) \right] \\
&= \left[570-\sqrt{\frac{9}{25}}\,\Phi^{-1}(0.995),\, 570+\sqrt{\frac{9}{25}}\,\Phi^{-1}(0.995) \right] \\
&\fallingdotseq \left[570-\frac{3}{5}\cdot 2.576,\, 570+\frac{3}{5}\cdot 2.576 \right] \\
&\fallingdotseq [568.45, 571.55]
\end{aligned}$$

問題 7.3　大きさが約67万の母集団から大きさ 400 の標本を無作為に非復元抽出します．標本の平均が 380 であり不偏分散が 144 であるとします．標本の不偏分散を母分散として代用して，母平均の信頼度 95% の信頼区間を求めなさい．