

連絡事項: 期末テスト (7月18日) の範囲は教科書の以下の部分とします。電卓は持ち込み可です。

第3章 2.3 統計量と標本分布,

第4章 1.2 母平均の区間推定 (1), 1.3 母平均の区間推定 (2) の例題 2, 問 5,

1.5 母比率の区間推定, 練習問題 1-A 3, 4, 5,

2.1 仮説と検定, 2.2 母平均の検定 (1), 2.7 母比率の検定, 練習問題 2-A 1, 6

1.2 母平均の区間推定 (1)

未知母数がある確率 $1 - \alpha$ で (これを信頼係数または信頼度という) 入る区間を推定する。これを区間推定といい、推定された区間を $100(1 - \alpha)\%$ 信頼区間あるいは信頼係数 $100(1 - \alpha)\%$ の信頼区間という。

$0 < \alpha < 1$ である値 α と標準正規分布に従う確率変数 Z について $P(Z \geq z_\alpha) = \alpha$ となる z_α を標準正規分布の上側 α 点という。(教科書 p.98 のグラフを参照のこと。) 以下の数値がよく用いられる。

$$z_{0.05} = 1.645, \quad z_{0.025} = 1.960, \quad z_{0.01} = 2.326, \quad z_{0.005} = 2.576.$$

定理 (母平均の区間推定 (母分散が既知の場合)) 正規母集団 $N(\mu, \sigma^2)$ から大きさ n の無作為標本の標本平均の実現値を \bar{x} とすると、母平均 μ の $100(1 - \alpha)\%$ 信頼区間は

$$\bar{x} - z_{\alpha/2} \sqrt{\frac{\sigma^2}{n}} \leq \mu \leq \bar{x} + z_{\alpha/2} \sqrt{\frac{\sigma^2}{n}}. \quad (1)$$

証明: 正規母集団 $N(\mu, \sigma^2)$ から大きさ n の無作為標本の標本平均を \bar{X} とすると、 \bar{X} は $N\left(\mu, \frac{\sigma^2}{n}\right)$ に従う。

このとき、その標準化 $Z = \frac{\bar{X} - \mu}{\sqrt{\sigma^2/n}}$ は $N(0, 1)$ に従うから、

$$P\left(-z_{\alpha/2} \leq \frac{\bar{X} - \mu}{\sqrt{\sigma^2/n}} \leq z_{\alpha/2}\right) = 1 - \alpha.$$

括弧内の不等式を μ について解くと、 $\bar{X} - z_{\alpha/2} \sqrt{\frac{\sigma^2}{n}} \leq \mu \leq \bar{X} + z_{\alpha/2} \sqrt{\frac{\sigma^2}{n}}$ となるが、これは μ がこの区間に含まれる確率が $1 - \alpha$ であることを示している。この \bar{X} にその実現値 \bar{x} を代入することで (1) 式を得る。□

注意 母集団分布が未知でも、標本の大きさ n が大きい場合は、母分散 σ^2 を不偏分散の実現値 u^2 に代用して上記の公式 (1) は成立する (cf. 教科書 p.100)。これは中心極限定理による。また、不偏分散 U^2 に対して $E[U^2] = \sigma^2$ が成立するため、標本分散より不偏分散のほうが σ^2 をより反映する数値であることに注意する。

例題 2 ある学校の生徒 50 人を無作為に選び、1 週間あたりのテレビ視聴時間 (単位 時間) を聞いたところ、50 人の平均 \bar{x} は 18.2, 不偏分散 u^2 は 30.25 であった、母平均 μ の 95% 信頼区間を求めよ。

解: $100(1 - \alpha) = 95$ より $\alpha = 0.05$. 標本の大きさ $n = 50$ は十分に大きいと考えると、

$$\bar{x} \pm z_{\alpha/2} \sqrt{\frac{u^2}{n}} = 18.2 \pm 1.960 \sqrt{\frac{30.25}{50}} = 18.2 \pm 1.5245 \dots = \begin{cases} 19.7245 \dots \\ 16.6754 \dots \end{cases}$$

従って、 $16.6 \leq \mu \leq 19.8$. □

注意 例題 2 解の最後の端数処理において、下限の数値の端数は切り捨て、上限の数値の端数は切り上げた。通常、区間推定では安全第一で、区間が狭くなりすぎないようにこのような端数処理を行う。次ページの例題 4 も同様。(下限上限ともに四捨五入する場合があります。教科書は四捨五入しているようです。小数点以下何位まで求めたらよいかを含め、問題文等の指示に従って処理してください。)

1.5 母比率の区間推定

定理 (母比率の区間推定) 二項母集団から大きさ n の無作為標本の標本比率の実現値を \hat{p} とすると、母比率 p の $100(1 - \alpha)\%$ 信頼区間は

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

解説: 母比率 p の二項母集団から大きさ n の無作為標本の標本比率を \hat{P} とすると、 $E[\hat{P}] = p$, $V[\hat{P}] = \frac{p(1 - p)}{n}$.

よって、 $Z = \frac{\hat{P} - p}{\sqrt{p(1 - p)/n}}$ は $N(0, 1)$ に従うから、

$$P\left(-z_{\alpha/2} \leq \frac{\hat{P} - p}{\sqrt{p(1 - p)/n}} \leq z_{\alpha/2}\right) = 1 - \alpha.$$

括弧の中の不等式は、 $\hat{P} - z_{\alpha/2} \sqrt{\frac{p(1 - p)}{n}} \leq p \leq \hat{P} + z_{\alpha/2} \sqrt{\frac{p(1 - p)}{n}}$ と変形できる。この \hat{P} とその両辺の根号内の p をすべて \hat{P} の実現値 \hat{p} に置き換えて、(2) 式を得る。□

例題 4 A 市で、あるテレビ番組の視聴率を調べるために、成人 500 人を無作為抽出したところ、45 人が見ていることがわかった。A 市におけるこの番組の成人の視聴率の 95% 信頼区間を求めよ。

解: $\alpha = 0.05$, $n = 500$, $\hat{p} = \frac{45}{500} = 0.09$ だから

$$\hat{p} \pm z_{\alpha/2} \sqrt{\frac{\hat{p}(1 - \hat{p})}{n}} = 0.09 \pm 1.960 \sqrt{\frac{0.09 \times 0.91}{500}} = 0.09 \pm 0.02508 \dots = \begin{cases} 0.11508 \dots \\ 0.06491 \dots \end{cases}$$

従って、 $0.064 \leq p \leq 0.116$. □

例題 5 ある都市で、有権者の内閣支持率 p を調べるのに標本調査を行うことになった。信頼区間の幅が 0.04 以下になるように、信頼係数 95% で p を区間推定したい。抽出する有権者の数を何名以上にすればよいか。

解: 抽出する人数を n とする。信頼区間は (2) 式で与えられるので、その幅は

$$\left(\hat{p} + z_{\alpha/2} \sqrt{\frac{\hat{p}(1 - \hat{p})}{n}}\right) - \left(\hat{p} - z_{\alpha/2} \sqrt{\frac{\hat{p}(1 - \hat{p})}{n}}\right) = 2 \times z_{\alpha/2} \sqrt{\frac{\hat{p}(1 - \hat{p})}{n}}.$$

したがって、 $2 \times z_{\alpha/2} \sqrt{\frac{\hat{p}(1 - \hat{p})}{n}} \leq 0.04$. $\alpha = 0.05$ なので、これを代入し変形すると

$$n \geq \left(\frac{2 \times z_{0.025}}{0.04}\right)^2 \hat{p}(1 - \hat{p})$$

これがどんな \hat{p} に対しても成立すればよい。ここで、 $\hat{p}(1 - \hat{p}) = -\left(\hat{p} - \frac{1}{2}\right)^2 + \frac{1}{4} \leq \frac{1}{4}$ より

$$n \geq \left(\frac{2 \times 1.960}{0.04}\right)^2 \times \frac{1}{4} = 2401.$$

以上より 2401 名以上とすればよい。□