数理統計学まとめ（その 9）：第 5 章 推定

4 区間推定

大きさ n の標本が得られ，その値が $(x_1, \ldots, x_n)$ であったとき，このデータを基に母数 θ の推定量 $T = T(x_1, \ldots, x_n)$ を得たとする．このとき，

「母数 θ の真の値は確率 95 % で区間 $[A, B]$ の中にある。」

という主張をすることを考える．95 % を 99 %まであげることも考えられる．このような結論の出し方を**区間推定**という．

4.1 母平均の区間推定: 母分散が既知の場合

まず，母平均 μ に関する区間推定を行う．母集団は $N(\mu, \sigma^2)$ に従うとして，σ^2 は既知の値として μ を区間推定する．標本平均

$$\bar{X} = \frac{1}{n}\sum_{j=1}^{n} X_j$$

は $N(\mu, \sigma^2/n)$ に従うので，

$$Z = \frac{\bar{X} - \mu}{\frac{\sigma}{\sqrt{n}}}$$

は標準正規分布 $N(0,1)$ に従う．z_α を標準正規分布の両側 α 点（ 100α %点）とすると，このとき

$$P(|Z| \geq z_\alpha) = \alpha$$

となっている（これが両側 α 点の定義）．したがって

$$P(|Z| < z_\alpha) = 1 - \alpha$$

ここで，

$$|Z| < z_\alpha \Longleftrightarrow |\bar{X} - \mu| < \frac{z_\alpha \sigma}{\sqrt{n}} \Longleftrightarrow \bar{X} - \frac{z_\alpha \sigma}{\sqrt{n}} < \mu < \bar{X} + \frac{z_\alpha \sigma}{\sqrt{n}}$$

だから，実際のデータ $x_1, \ldots, x_n$ に対して $\bar{x} = \frac{1}{n}\sum_{j=1}^{n} x_j$ とおくと，

$$P\left(\bar{x} - \frac{z_\alpha \sigma}{\sqrt{n}} < \mu < \bar{x} + \frac{z_\alpha \sigma}{\sqrt{n}}\right) = 1 - \alpha$$

となる．$1-\alpha$ を**信頼係数**，区間

$$[\bar{x}-\frac{z_\alpha\sigma}{\sqrt{n}},\bar{x}+\frac{z_\alpha\sigma}{\sqrt{n}}] \tag{1}$$

を**信頼区間**という．

95 %の信頼係数のときの信頼区間を考えてみよう．$z_{0.05}=\pm1.96$（教科書の標準正規分布表で $0.95/2=0.475$ に対応するところ）だから，母平均 μ は 95 %の確率で区間

$$\left(\bar{x}-\frac{1.96\sigma}{\sqrt{n}},\bar{x}+\frac{1.96\sigma}{\sqrt{n}}\right)$$

にあることが言える．これが信頼区間になる．

例 4.1（教科書 p.106，例題 5.6）

母分散 $\sigma^2=25$ の正規母集団から 10 個の無作為標本をとり，$\bar{x}=12.8$ を得た．母平均の 95 %信頼区間を求めよ．

解　未知母数は母集団の平均 μ で,

$$\frac{\bar{X}-\mu}{\frac{\sigma}{\sqrt{n}}}$$

が標準正規分布に従うので，その両側 5 %点は正規分布表では $0.5-0.025=0.475$ に対応する値 1.96 をつかって ±1.96 であることがわかる．したがって 95 %の確率で

$$\left|\frac{\sqrt{n}(\bar{X}-\mu)}{\sigma}\right|<1.96$$

が成り立っている．いま，$n=10$ で $\bar{X}=\bar{x}=12.8$ と $\sigma=\sqrt{25}=5$ がわかっているので，求める信頼区間は上の式にこれらの値を代入して μ について解くことにより

$$12.8-1.96\times5/\sqrt{10}<\mu<12.8+1.96\times5/\sqrt{10}.$$

これを小数点以下第 1 位まで求めて $(9.7,15.9)$ が μ の 95 %信頼区間である．(公式 (1) をそのまま利用して求めて良い）.

4.2　母平均の区間推定: 母分散が未知の場合

母分散が未知の場合は，ティー分布を使う．今，無作為標本 $X_1, \ldots, X_n$ が得られたとして，このとき

$$T = \frac{\bar{X} - \mu}{\frac{U}{\sqrt{n}}}$$

が自由度 $n-1$ のティー分布なので，信頼度 $1-\alpha$ の信頼区間を求めようとすると, ティー分布表で自由度 $n-1$ の両側 α 点 $t_{n-1}(\alpha)$ を求めて

$$\left| \frac{\bar{X} - \mu}{\frac{U}{\sqrt{n}}} \right| < t_{n-1}(\alpha)$$

を満たす μ の区間

$$\left(\bar{X} - t_{n-1}(\alpha) \frac{U}{\sqrt{n}}, \bar{X} + t_{n-1}(\alpha) \frac{U}{\sqrt{n}} \right) \tag{2}$$

が求める信頼区間となる．U^2 の代わりに標本分散

$$S^2 = \frac{1}{n} \sum_{j=1}^{n} (X_j - \bar{X})^2$$

を使う場合は，$S = \sqrt{S^2}$ として $U = \frac{n}{n-1} S$ だから上の式に代入して，信頼度 $1-\alpha$ の μ の信頼区間は

$$\left(\bar{X} - t_{n-1}(\alpha) \frac{S}{\sqrt{n-1}}, \bar{X} + t_{n-1}(\alpha) \frac{S}{\sqrt{n-1}} \right) \tag{3}$$

ともかける．

例 4.2　（教科書 p.108, 例題 5.8, (2)）正規母集団から 20 個の無作為標本を抽出して $\bar{x} = 125.8$ を得た．不偏分散による推定値が $u^2 = 4.57^2$ であるとき母平均の 95 %信頼区間と 99 %信頼区間を求めよ．
解　公式 (2) が使える．$U = u = 4.57, n = 20, \bar{X} = \bar{x} = 125.8$ を代入して，$t_{19}(0.05) = 2.093, t_{19}(0.01) = 2.861$ だから，95 %信頼区間は $(123.7, 127.9)$ で, 99 %信頼区間は，$(122.9, 128.7)$ となる．

4.3　母分散の推定

正規母集団からの大きさ n の無作為標本 $X_1, \ldots, X_n$ を得たとき，母分散が区間推定できる．これには

$$X^2 = \frac{1}{\sigma^2}\sum_{j=1}^{n}(X_j - \bar{X})^2$$

が自由度 $n-1$ のカイ 2 乗分布に従うので，上側 $\alpha/2$ 点 $\chi_{n-1}(\alpha/2)$ と上側 $1-\alpha/2$ 点 $\chi_{n-1}(1-\alpha/2)$ を求めて

$$\chi_{n-1}(1-\alpha/2) < X^2 < \chi_{n-1}(\alpha/2)$$

となる確率が $1-\alpha$ となるので，

$$X^2 = \frac{nS^2}{\sigma^2} = \frac{(n-1)U^2}{\sigma^2}$$

であことから信頼度 $1-\alpha$ の母分散 σ^2 の信頼区間を

$$\left(\frac{nS^2}{\chi_{n-1}(\alpha/2)}, \frac{nS^2}{\chi_{n-1}(1-\alpha/2)}\right)$$

と求めることができる．