＜母平均の区間推定 5＞

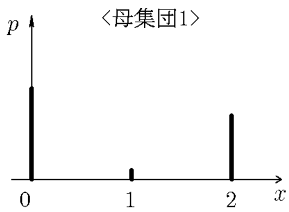
◎ 母集団の分布型がわからない場合、標本数が小さいと区間推定した場合に信頼性が全くないので、この場合は区間推定できない。

◎ 母集団の分布がわからなくても、標本数が多ければ区間推定できる。

(1) [復元抽出した場合]　中心極限定理より、標本平均 $\overline{X}$ の分布は正規分布に近づくので、30 ページ $(*)'$ が適用できる。

(2) [非復元抽出した場合]　母平均 μ, 母分散 σ^2, 大きさ N の母集団から n 個の標本 $X_1,\ X_2,\ \cdots,\ X_n$ を非復元抽出した場合、標本平均 $\overline{X}$ の確率変数としての平均と分散は定理 15 より

$$E\left[\,\overline{X}\,\right]=\mu \quad , \quad E\left[\left(\overline{X}-\mu\right)^2\right]=\frac{N-n}{N-1}\times\frac{\sigma^2}{n}$$



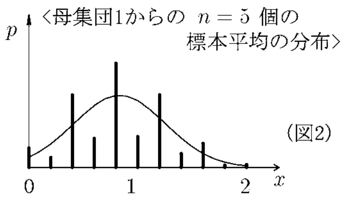
(図1)

である。さらに $N,\ n$ が共に大きいとき

(☆) $\boxed{\overline{X} \text{ は近似的に正規分布 } N\left(\mu,\ \frac{N-n}{N-1}\times\frac{\sigma^2}{n}\right) \text{ に従う}}$

ことが知られている。従って $\overline{X}$ の標準化

$$\left(\overline{X}-\mu\right)\Big/\sqrt{\frac{N-n}{N-1}\times\frac{\sigma^2}{n}}$$



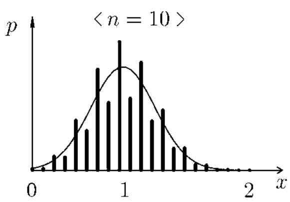
(図2)

は近似的に標準正規分布に従うから

$$P\left(\left|\left(\overline{X}-\mu\right)\Big/\sqrt{\frac{N-n}{N-1}\times\frac{\sigma^2}{n}}\right|\leqq\alpha\right)\fallingdotseq\int_{-\alpha}^{\alpha}\frac{1}{\sqrt{2\pi}}e^{-\frac{x^2}{2}}dx\left(=f(\alpha)\right)$$

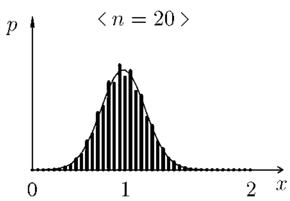
が成り立つ。よって信頼度 $f(\alpha)$ の信頼区間は

$$\boxed{\overline{X}-\alpha\sqrt{\frac{N-n}{N-1}\times\frac{\sigma^2}{n}}\leqq\mu\leqq\overline{X}+\alpha\sqrt{\frac{N-n}{N-1}\times\frac{\sigma^2}{n}}\quad(***)}$$



(図3)

なお母分散 σ^2 が未知のときは、$N,\ n$ が共に大きくなると、定理 16 より不偏分散 $V=\dfrac{N-1}{N(n-1)}\displaystyle\sum_{i=1}^{n}\left(X_i-\overline{X}\right)^2$ が σ^2 に近づくので、σ^2 のかわりに V で代用する。



(図4)

(注) 右の図は (☆) の例を示している。母集団が図 1 のように 3 つの値をもつ集合

$\{0,0,0,\cdots,1,1,\cdots,2,2,\cdots,2,2\}$ (0 が 100 個, 1 が 10 個, 2 が 70 個)

の場合、この母集団から無作為に非復元抽出した標本 $X_1,\ X_2,\ \cdots,\ X_n$ の平均 $\overline{X}=\dfrac{1}{n}\displaystyle\sum_{i=1}^{n}X_i$ の分布が図 2(n=5), 図 3(n=10), 図 4(n=20) である。n=20 ぐらいになると、$\overline{X}$ の分布はほぼ正規分布といってもよいであろう。