

§ 5.3 無限母集団と有限母集団

標本調査において、母集団の大きさに対する標本の大きさの割合がきわめて小さいとき、その母集団を**無限母集団** (infinite population) という。母集団が無限母集団といえないときは有限母集団ということがある。無限母集団において、抽出する前の状態と抽出した後の状態とはほとんど同じである。

例 大きさが 100 万である母集団から抽出した標本の大きさが 1000 であるとする。母集団の大きさに対する標本の大きさの割合は 1000 分の 1 できわめて小さい。母標準偏差を σ とおき、標本平均を表す確率変数を $\overline{X}$ とおく。 $\overline{X}$ の分散 $V[\overline{X}]$ は次のようになる：非復元抽出のとき、

$$V[\overline{X}] = \frac{1000000 - 1000}{1000000 - 1} \frac{\sigma^2}{1000} = \frac{\sigma^2}{1001} ;$$

復元抽出のとき、

$$V[\overline{X}] = \frac{\sigma^2}{1000} .$$

ここで

$$\frac{1}{1001} \doteq 0.999000999 \times 10^{-3} \doteq 10^{-3} = \frac{1}{1000} .$$

非復元抽出のときの $V[\overline{X}]$ の値と復元抽出のときの $V[\overline{X}]$ の値とはあまり差がない。 終

このように、無限母集団のとき、つまり母集団の大きさに対する標本の大きさの割合がきわめて小さいとき、標本平均を表す確率変数の確率分布は、非復元抽出と復元抽出とであまり差がない。

5.2 節において中心極限定理から次の定理を導いた。

[定理 5.2.3] 母平均が μ であり母標準偏差が σ である母集団から無作為に復元抽出された標本の標本平均を表す確率変数は、標本の大きさ n が充分大きいとき、近似的に正規分布 $N\left(\mu, \frac{\sigma^2}{n}\right)$ に従う。

無限母集団において、標本平均を表す確率変数の確率分布は、非復元抽出と復元抽出とであまり差がない。これより次の定理が成り立つ。

[定理 5.3] 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された標本の標本平均を表す確率変数は、標本の大きさ n が充分大きいとき、近似的に正規分布 $N\left(\mu, \frac{\sigma^2}{n}\right)$ に従う。

つまり、無限母集団においては、非復元抽出であっても、無作為抽出された標本の大きさが大きいとき標本平均を表す確率変数は近似的に正規分布に従う。

標本調査では通常は厳密な値を得られない。それ故、数学的根拠において確率が厳密な値か近似値かに拘ることにあまり意味がない。そこで、しばしば、定理 5.3 における“近似的に正規分布 $N\left(\mu, \frac{\sigma^2}{n}\right)$ に従う。”という記述の“近似的に”という断りを省く。

例題 母平均が 59 であり母標準偏差が 7 である大きさ約 37 万の母集団から無作為に非復元抽出された大きさ 100 の標本の標本平均を表す確率変数 $\overline{X}$ について、確率 $P(\overline{X} \geq 60)$ を求める。

母集団の大きさ約 37 万に対する標本の大きさ 100 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。確率変数 $\overline{X}$ の分散 $V[\overline{X}]$ は

$$V[\overline{X}] = \frac{7^2}{100} .$$

確率変数 $\overline{X}$ は正規分布 $N\left(59, \frac{7^2}{100}\right)$ に従うとする。 $\overline{X}$ を標準化した確率変数

$$Z = \frac{\overline{X} - 59}{\sqrt{\frac{7^2}{100}}} = \frac{10}{7}(\overline{X} - 59)$$

は標準正規分布 $N(0,1)$ に従う。

$$\overline{X} \geq 60 \iff \frac{10}{7}(\overline{X} - 59) \geq \frac{10}{7}(60 - 59) \iff Z \geq \frac{10}{7} .$$

標準正規分布の累積分布関数を $\Phi(x)$ とおく。

$$\begin{aligned} P(\overline{X} \geq 60) &= P\left(Z \geq \frac{10}{7}\right) \doteq 1 - \Phi\left(\frac{10}{7}\right) \doteq 1 - \Phi(1.43) \doteq 1 - 0.9236 \\ &= 0.0764 . \end{aligned}$$

終

問 5.3 母平均が 378 であり母分散が 162 である大きさ約 74 万の母集団から無作為に非復元抽出された大きさ 200 の標本の標本平均を表す確率変数 $\overline{X}$ について、確率 $P(377 \leq \overline{X} \leq 380)$ を求めよ。

母平均を μ とおき、母標準偏差を σ とおく。無限母集団から無作為抽出された大きさ n の標本の標本平均を表す確率変数 $\overline{X}$ について

$$E[\overline{X}] = \mu , \quad V[\overline{X}] = \frac{\sigma^2}{n} .$$

大きさ N の有限母集団から無作為に非復元抽出された大きさ n の標本の標本平均を表す確率変数 $\overline{X}$ について、

$$E[\overline{X}] = \mu , \quad V[\overline{X}] = \frac{N-n}{N-1} \frac{\sigma^2}{n} .$$

このように、無限母集団から無作為抽出された標本の標本平均を表す確率変数の分散に $\frac{N-n}{N-1}$ を掛けると、有限母集団から無作為に非復元抽出された標本の標本平均を表す確率変数の分散になる。

標本の大きさ n が充分大きいとき、無限母集団では、非復元抽出でも近似的に復元抽出とみなして中心極限定理を適用することができる。しかし有限母集団では本来はそのようなことができない。しかしながら、数学的には無理があることを承知で、便宜的に、標本平均を表す確率変数の確率分布を正規分布 $N\left(\mu, \frac{N-n}{N-1} \frac{\sigma^2}{n}\right)$ で近似することがある。