

中心極限定理と正規近似 問題 1 解答

- 1 コンピュータを使い 0 から 10 までの実数を、一様分布に従うものとしてランダムに 300 個生成した.

- (1) 中心極限定理を用いて、生成した実数の平均が近似的に従う正規分布を求めよ.
 (2) (1) で求めた近似正規分布を用いて、平均が 4.75 と 5.25 の間にある確率を求めよ.

[解]:

- (1) ランダムに生成された実数をそれぞれ $X_1, X_2, \dots, X_{300}$ とすると、これらは互いに独立に同一の連続一様分布 $U(0, 10)$ に従う. このとき一様分布 $U(0, 10)$ の期待値と分散は

$$E(X_i) = \frac{0+10}{2} = 5, \quad V(X_i) = \frac{(10-0)^2}{12} = \frac{25}{3}.$$

このとき標本平均 $\bar{X} = \frac{1}{300} \sum_{i=1}^{300} X_i$ の期待値と分散は、互いに独立であることから

$$E(\bar{X}) = 5, \quad V(\bar{X}) = \frac{\frac{25}{3}}{300} = \frac{25}{900} = \left(\frac{1}{6}\right)^2$$

と求められる.

ここで中心極限定理とは,

確率変数 $X_1, X_2, \dots, X_n$ は互いに独立で、平均 μ 、分散 σ^2 のある確率分布 $A(\mu, \sigma^2)$ に従うとする. このとき標本平均 $\bar{X} = \frac{1}{n} \sum_{i=1}^n X_i$ に対して、大きさ n が十分大きければ $\bar{X}$ の標本分布は限りなく正規分布 $N\left(\mu, \frac{\sigma^2}{n}\right)$ に近似される.

したがって中心極限定理により、 $\bar{X}$ は正規分布 $N(5, (\frac{1}{6})^2)$ に近似的に従う.

- (2) $\bar{X} \sim N(5, (\frac{1}{6})^2)$ より、確率変数を $Z = \frac{\bar{X}-5}{\frac{1}{6}} = 6(\bar{X} - 5)$ とおくと、 Z は標準正規分布に従う. したがって求める確率は、

$$P(4.75 \leq \bar{X} \leq 5.25) = P(-1.5 \leq Z \leq 1.5) = 2 \cdot 0.4332 = 0.8664.$$

- 2 裏と表の出る確率がちょうど $\frac{1}{2}$ であるコインを 400 回投げた. このとき、表が出る回数が 180 回以上 220 回以下となる確率を二項分布の正規近似を利用して求めよ.

[解]: コインを 400 回投げて、表が出た回数を X とおくと、 X は二項分布 $B(400, 0.5)$ に従う. ここで、二項分布の正規近似とは以下の確率分布の近似のことを言う:

$Y \sim B(n, p)$ ならば、 Y は近似的に正規分布 $N(np, np(1-p))$ に従う.

したがって X は近似的に正規分布 $N(200, 10^2)$ に従う. P_B を二項分布 $B(400, 0.5)$ における確率、 P_N を正規分布 $N(200, 10^2)$ における確率とすると、求める確率は

$$P_B(180 \leq X \leq 220) \div P_N(180 \leq X \leq 220)$$

となり、正規分布の確率で近似計算が行える.

離散分布をヒストグラムと見て考えると、離散分布である二項分布を連続分布へ近似するときには、半整数補正を行うと精度の良い近似となる。すなわち、 $P_N(180 \leq X \leq 220)$ の代わりに確率 $P_N(180 - 0.5 \leq X \leq 220 + 0.5)$ を求める。

ここで、確率変数の変数変換によって標準化 $Z = \frac{X - 200}{10} \sim N(0, 1)$ を行う。確率変数 Z は標準正規分布に従うので、

$$P_N(180 - 0.5 \leq X \leq 220 + 0.5) = P_N(-2.05 \leq Z \leq 2.05) = 2 \cdot P_N(0 \leq Z \leq 2.05) = 0.9596.$$

したがって、求める確率は 0.9596 である。

[正規近似を使わず二項分布のままで計算すると、求める確率は 0.959769 となる.]