

§ 4.2 多次元の確率変数

例えば、ある工場で生産されるある製品を何個か取り出してその製品の重量を測るとする。この工場で生産されるこの製品の重量がどのように分布するか定まっている、かつ、製品の取り出し方が決まっているとする。このときこの工場で生産されるこの製品の1個の重量を表わす変数 X を確率変数と考えることができる。この製品を順番に1000個取り出してそれらの製品の重量を測るとする。1番めの製品の重量を表す変数を X_1 とおき、2番めの製品の重量を表す変数を X_2 とおき、3番めの製品の重量を表す変数を X_3 とおき、というように、1000番めの製品の重量を表わす変数 X_{1000} まで考えると、1000個の確率変数 $X_1, X_2, X_3, \dots, X_{1000}$ の組 $(X_1, X_2, X_3, \dots, X_{1000})$ ができる。このような1000個の確率変数の組を1000次元の確率変数という。

おおまかにいうと、正の自然数 n に対して、 n 次元の確率変数とは1番から n 番までの番号が付いた確率変数の組である。但し、数学的には以下のようにやや条件が付く。

[定義] n 個 (n は正の自然数)の確率変数 $X_1, X_2, X_3, \dots, X_n$ の組 $(X_1, X_2, X_3, \dots, X_n)$ が n 次元の確率変数であるとは次の条件を満たすことである：任意の実数 $a_1, a_2, a_3, \dots, a_n$ に対して、

$$X_1 \leq a_1 \text{ かつ } X_2 \leq a_2 \text{ かつ } X_3 \leq a_3 \text{ かつ } \cdots \text{ かつ } X_n \leq a_n$$

となる確率が定まる。

n 個 (n は正の自然数)の確率変数 $X_1, X_2, X_3, \dots, X_n$ の組 $(X_1, X_2, X_3, \dots, X_n)$ が n 次元の確率変数であるとき、任意の実数 $a_1, a_2, a_3, \dots, a_n$ に対して、“ $X_1 \leq a_1$ かつ $X_2 \leq a_2$ かつ $X_3 \leq a_3$ かつ $\cdots$ かつ $X_n \leq a_n$ ” となる確率を

$$P(X_1 \leq a_1, X_2 \leq a_2, X_3 \leq a_3, \dots, X_n \leq a_n)$$

と書き表す。

[定義] n 個 (n は正の自然数)の確率変数 $X_1, X_2, X_3, \dots, X_n$ が互いに独立であるとは次の条件を満たすことである： $(X_1, X_2, X_3, \dots, X_n)$ が n 次元の確率変数であり、任意の実数 $a_1, a_2, a_3, \dots, a_n$ について

$$P(X_1 \leq a_1, X_2 \leq a_2, \dots, X_n \leq a_n) = P(X_1 \leq a_1) P(X_2 \leq a_2) \cdots P(X_n \leq a_n) .$$

2次元の確率変数と同様に次の定理が成り立つ。

[定理 4.2.1] n 個 (n は正の自然数)の確率変数 $X_1, X_2, X_3, \dots, X_n$ の組 $(X_1, X_2, X_3, \dots, X_n)$ は n 次元の確率変数であるとする。変数 Y を $Y = X_1 + X_2 + X_3 + \cdots + X_n$ とおく。 Y は(1次元の)確率変数であり、

$$E[Y] = E[X_1 + X_2 + X_3 + \cdots + X_n] = E[X_1] + E[X_2] + E[X_3] + \cdots + E[X_n] .$$

更に、 n 個の確率変数 $X_1, X_2, X_3, \dots, X_n$ が互いに独立であるならば、

$$V[Y] = V[X_1 + X_2 + X_3 + \cdots + X_n] = V[X_1] + V[X_2] + V[X_3] + \cdots + V[X_n] .$$

この定理より直ぐに次の定理が導かれる。

[定理 4.2.2] 定数 μ, σ は実数で $\sigma > 0$ とする。 n 個 (n は正の自然数)の確率変数 $X_1, X_2, X_3, \dots, X_n$ の組 $(X_1, X_2, X_3, \dots, X_n)$ は n 次元の確率変数であり、

$$\begin{aligned} E[X_1] &= E[X_2] = E[X_3] = \cdots = E[X_n] = \mu , \\ V[X_1] &= V[X_2] = V[X_3] = \cdots = V[X_n] = \sigma^2 \end{aligned}$$

とする。 $X_1, X_2, X_3, \dots, X_n$ の平均を表わす確率変数

$$\overline{X} = \frac{X_1 + X_2 + X_3 + \cdots + X_n}{n} = \frac{1}{n} \sum_{k=1}^n X_k$$

について、 $E[\overline{X}] = \mu$, n 個の確率変数 $X_1, X_2, X_3, \dots, X_n$ が互いに独立であるならば $V[\overline{X}] = \frac{\sigma^2}{n}$.

[証明] 定理 4.2.1 より、 $X_1 + X_2 + X_3 + \cdots + X_n$ は確率変数で、

$$\begin{aligned} E[\overline{X}] &= E\left[\frac{X_1 + X_2 + X_3 + \cdots + X_n}{n}\right] = E\left[\frac{1}{n}(X_1 + X_2 + X_3 + \cdots + X_n)\right] \\ &= \frac{1}{n}E[X_1 + X_2 + X_3 + \cdots + X_n] \\ &= \frac{1}{n}(E[X_1] + E[X_2] + E[X_3] + \cdots + E[X_n]) \\ &= \frac{1}{n}\underbrace{(\mu + \mu + \mu + \cdots + \mu)}_{n \text{ 個の和}} = \frac{1}{n}(n\mu) \\ &= \mu . \end{aligned}$$

$X_1, X_2, X_3, \dots, X_n$ が互いに独立であるならば、

$$\begin{aligned} V[\overline{X}] &= V\left[\frac{X_1 + X_2 + X_3 + \cdots + X_n}{n}\right] = V\left[\frac{1}{n}(X_1 + X_2 + X_3 + \cdots + X_n)\right] \\ &= \left(\frac{1}{n}\right)^2 V[X_1 + X_2 + X_3 + \cdots + X_n] \\ &= \frac{1}{n^2}(V[X_1] + V[X_2] + V[X_3] + \cdots + V[X_n]) \\ &= \frac{1}{n^2}\underbrace{(\sigma^2 + \sigma^2 + \sigma^2 + \cdots + \sigma^2)}_{n \text{ 個の和}} = \frac{1}{n^2}(n\sigma^2) \\ &= \frac{\sigma^2}{n} . \end{aligned}$$

証明終了

互いに独立な n 個の確率変数 $X_1, X_2, X_3, \dots, X_n$ が総て例えば一様分布に従うときでも、確率変数 $\overline{X} = \frac{X_1 + X_2 + X_3 + \cdots + X_n}{n}$ が一様分布に従うとは限らない。しかし特に正規分布については次の定理が成り立つ。

[定理 4.2.3] 定数 μ, σ は実数で $\sigma > 0$ とする。正規分布 $N(\mu, \sigma^2)$ に従う n 個 (n は正の自然数)の確率変数 $X_1, X_2, X_3, \dots, X_n$ が互いに独立であるならば、これらの平均を表わす確率変数 $\overline{X} = \frac{X_1 + X_2 + X_3 + \cdots + X_n}{n} = \frac{1}{n} \sum_{k=1}^n X_k$ は正規分布 $N\left(\mu, \frac{\sigma^2}{n}\right)$ に従う。

[例題] 正規分布 $N(70, 90)$ に従う5個の確率変数 X_1, X_2, X_3, X_4, X_5 は互いに独立であるとする。これらの平均を表わす確率変数

$$\overline{X} = \frac{X_1 + X_2 + X_3 + X_4 + X_5}{5}$$

について、確率 $P(\overline{X} \geq 75)$ を求める。

確率変数 $\overline{X}$ は正規分布 $N\left(70, \frac{90}{5}\right) = N(70, 18)$ に従う。よって、 $\overline{X}$ を標準化した確率変数

$$Z = \frac{\overline{X} - 70}{\sqrt{18}}$$

は標準正規分布 $N(0, 1)$ に従う。

$$\overline{X} \geq 75 \iff \frac{\overline{X} - 70}{\sqrt{18}} \geq \frac{75 - 70}{\sqrt{18}} \iff Z \geq \frac{5}{\sqrt{18}} .$$

標準正規分布の累積分布関数を $\Phi(x)$ とおく。

$$\begin{aligned} P(\overline{X} \geq 75) &= P\left(Z \geq \frac{5}{\sqrt{18}}\right) = 1 - P\left(Z \leq \frac{5}{\sqrt{18}}\right) \\ &= 1 - \Phi\left(\frac{5}{\sqrt{18}}\right) \doteq 1 - \Phi(1.179) \doteq 1 - 0.88 \\ &= 0.12 . \end{aligned}$$

終

[問 4.2] 正規分布 $N(90, 112)$ に従う7個の確率変数 $X_1, X_2, X_3, \dots, X_7$ は互いに独立であるとする。これらの平均を表わす確率変数

$$\overline{X} = \frac{X_1 + X_2 + X_3 + \cdots + X_7}{7}$$

について、確率 $P(\overline{X} \leq 87)$ と $P(88 \leq \overline{X} \leq 95)$ とを求めよ。