§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します．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て，かつ，製品の取り出し方が決まっているとします．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ます．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します．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 $X_1, X_2, X_3, \ldots, X_{1000}$ の組 $(X_1, X_2, X_3, \ldots, X_{1000})$ ができます．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います．

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す．但し，数学的には以下のようにやや条件が付きます．

定義　n 個（n は正の自然数）の変数 $X_1, X_2, X_3, \ldots, X_n$ の組 $(X_1, X_2, X_3, \ldots, X_n)$ 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 $a_1, a_2, a_3, \ldots, a_n$ に対して，

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

となる確率が定まる．

n 次元（n は正の自然数）の確率変数 $(X_1, X_2, X_3, \ldots, X_n)$ 及び任意の実数 $a_1, a_2, a_3, \ldots, a_n$ に対して，$X_1 \le a_1$ かつ $X_2 \le a_2$ かつ $X_3 \le a_3$ かつ $\cdots$ かつ $X_n \le a_n$ となる確率を

$$\mathrm{P}(\, X_1 \le a_1 \,,\, X_2 \le a_2 \,,\, X_3 \le a_3 \,,\, \cdots \,,\, X_n \le a_n \,)$$

と書き表す．

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

$$\mathrm{P}(\, X_1 \le a_1 \,,\, X_2 \le a_2 \,,\cdots,\, X_n \le a_n \,) = \mathrm{P}(\, X_1 \le a_1 \,)\,\mathrm{P}(\, X_2 \le a_2 \,)\cdots\mathrm{P}(\, X_n \le a_n \,)\ .$$

2次元の確率変数と同様に次の定理が成り立ちます．

定理 4.2.1　n 次元（n は正の自然数）の確率変数 $(X_1, X_2, X_3, \cdots, X_n)$ に対して，$Y = X_1 + X_2 + X_3 + \cdots + X_n$ となる確率変数 Y は（1次元の）確率変数であり，

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

更に，n 個の確率変数 $X_1, X_2, X_3, \cdots, X_n$ が互いに独立であるとき，

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

この定理より直ぐに次の定理が導けます．

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 $X_1, X_2, X_3, \cdots, X_n$ に対して，$(X_1, X_2, X_3, \cdots, X_n)$ は n 次元の確率変数であり，

$$\mathrm{E}[X_1] = \mathrm{E}[X_2] = \mathrm{E}[X_3] = \cdots = \mathrm{E}[X_n] = \mu\ ,$$
$$\mathrm{V}[X_1] = \mathrm{V}[X_2] = \mathrm{V}[X_3] = \cdots = \mathrm{V}[X_n] = \sigma^2$$

とする．$X_1, X_2, X_3, \cdots, X_n$ 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，$\mathrm{E}[\overline{X}] = \mu$，$n$ 個の確率変数 $X_1, X_2, X_3, \cdots, X_n$ が互いに独立であるならば $\mathrm{V}[\overline{X}] = \dfrac{\sigma^2}{n}$ ．

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

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

$X_1, X_2, X_3, \cdots, X_n$ が互いに独立であるとき，

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

（証明終り）

互いに独立な n 個の確率変数 $X_1, X_2, X_3, \cdots, X_n$ が総て例えば一様分布に従うときでも，確率変数 $\overline{X} = \dfrac{X_1 + X_2 + X_3 + \cdots + X_n}{n}$ が一様分布に従うとは限りません．しかし特に正規分布については次の定理が成り立ちます．

定理 4.2.3　定数 μ, σ は実数で $\sigma > 0$ とする．互いに独立な n 個（n は正の自然数）の確率変数 $X_1, X_2, X_3, \cdots, X_n$ が総て正規分布 $\mathrm{N}(\mu, \sigma^2)$ に従うとき，これらの平均を表わす確率変数 $\overline{X} = \dfrac{X_1 + X_2 + X_3 + \cdots + X_n}{n} = \dfrac{1}{n}\sum_{k=1}^{n} X_k$ は正規分布 $\mathrm{N}\left(\mu, \dfrac{\sigma^2}{n}\right)$ に従う．

例題 4.2　5個の確率変数 X_1, X_2, X_3, X_4, X_5 が互いに独立で，総て正規分布 $\mathrm{N}(70, 90)$ に従うとする．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 $\mathrm{P}(\,\overline{X} \ge 75\,)$ を求める．

確率変数 $\overline{X}$ は正規分布 $\mathrm{N}\left(70, \dfrac{90}{5}\right) = \mathrm{N}(70, 18)$ 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 $\mathrm{N}(0, 1)$ に従う．

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

標準正規分布の累積分布関数を $\Phi(x)$ とおく．

$$\begin{aligned}
\mathrm{P}(\,\overline{X} \ge 75\,) &= \mathrm{P}\left(\, Z \ge \frac{5}{\sqrt{18}} \,\right) = 1 - \mathrm{P}\left(\, Z \le \frac{5}{\sqrt{18}} \,\right) \\
&= 1 - \Phi\left(\frac{5}{\sqrt{18}}\right) \fallingdotseq 1 - \Phi(1.179) \fallingdotseq 1 - 0.88 \\
&= 0.12\ .
\end{aligned}$$

終

問題 4.2　7個の確率変数 $X_1, X_2, X_3, \cdots, X_7$ が互いに独立で，総て正規分布 $\mathrm{N}(90, 112)$ に従うとします．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 $\mathrm{P}(\,\overline{X} \le 87\,)$ と $\mathrm{P}(\, 88 \le \overline{X} \le 95 \,)$ とを求めなさい．