11　多次元分布（1）

導入

これまでの議論では 1 つの確率変数に対して行ってきたが，2 つ以上の確率変数を同時に扱うことも必要になる．今回は 2 つ以上の確率変数についての数学的定式化を行い，そのときに重要になる確率変数の独立性について述べていく．

11.1　多次元分布

まずは具体例を通してみていこう．

例 11.1. サイコロを 2 個同時に投げる試行を考える．X を二つの目のうち大きいほうの値とし，Y を大きい目から小さい目を引いた値とする．例えばサイコロの出目が 3 と 5 であった場合は $X=5, Y=5-3=2$ となる．このとき確率 $P(X=i, Y=j)$ は

$X\backslash Y$	0	1	2	3	4	5	計
1	$\frac{1}{36}$	0	0	0	0	0	$\frac{1}{36}$
2	$\frac{1}{36}$	$\frac{1}{18}$	0	0	0	0	$\frac{1}{12}$
3	$\frac{1}{36}$	$\frac{1}{18}$	$\frac{1}{18}$	0	0	0	$\frac{5}{36}$
4	$\frac{1}{36}$	$\frac{1}{18}$	$\frac{1}{18}$	$\frac{1}{18}$	0	0	$\frac{7}{36}$
5	$\frac{1}{36}$	$\frac{1}{18}$	$\frac{1}{18}$	$\frac{1}{18}$	$\frac{1}{18}$	0	$\frac{1}{4}$
6	$\frac{1}{36}$	$\frac{1}{18}$	$\frac{1}{18}$	$\frac{1}{18}$	$\frac{1}{18}$	$\frac{1}{18}$	$\frac{11}{36}$
計	$\frac{1}{6}$	$\frac{5}{18}$	$\frac{2}{9}$	$\frac{1}{6}$	$\frac{1}{9}$	$\frac{1}{18}$	1

これを確率変数 (X,Y) の**同時分布** (または**多次元分布**) と呼ぶ．各行・各列の確率の総和は X, Y それぞれの分布を表すことになる．これらをそれぞれ X, Y の**周辺分布**という．この例における X の周辺分布は

X の値	1	2	3	4	5	6	計
確率	$\frac{1}{36}$	$\frac{1}{12}$	$\frac{5}{36}$	$\frac{7}{36}$	$\frac{1}{4}$	$\frac{11}{36}$	1

であり，Y の周辺分布は

Y の値	0	1	2	3	4	5	計
確率	$\frac{1}{6}$	$\frac{5}{18}$	$\frac{2}{9}$	$\frac{1}{6}$	$\frac{1}{9}$	$\frac{1}{18}$	1

演習問題 11.2. 確率変数 X, Y の同時確率分布が次のように与えられている．

$X\backslash Y$	1	2	3
1	0.15	0.06	0.09
2	0.35	0.14	0.21

(1)　$P(X \leqq 1, Y \leqq 2)$ を求めよ．

(2)　確率変数 X, Y の周辺分布を求めよ．

(3)　確率変数 X, Y の期待値及び分散を求めよ．

同様の考え方で，連続型の場合についても同時分布が定義される．

定義 11.3

X, Y を二つの連続型確率変数とするとき，

$$P(a \leqq X \leqq b,\ c \leqq Y \leqq d) = \int_c^d \int_a^b f(x,y)\,dx\,dy$$

を 2 次元確率変数 (X,Y) の**同時確率分布**という．ただし $f(x,y)$ は

$$f(x,y) \geqq 0, \quad \int_{-\infty}^{\infty}\int_{-\infty}^{\infty} f(x,y)\,dx\,dy = 1$$

を満たす 2 変数の非負可積分関数である．この $f(x,y)$ を**同時確率密度関数**という．X, Y の**周辺確率密度関数** $f_X(x)$, $f_Y(y)$ はそれぞれ

$$f_X(x) = \int_{-\infty}^{\infty} f(x,y)\,dy$$
$$f_Y(y) = \int_{-\infty}^{\infty} f(x,y)\,dx$$

で与えられる．

演習問題 11.4. 確率変数 X, Y の同時確率密度関数が，以下で与えられている．

$$f(x,y) = \begin{cases} ce^{-3x-y} & (x \geqq 0,\ y \geqq 0), \\ 0 & (\text{その他}). \end{cases}$$

(1)　c の値を求めよ．

(2)　確率変数 X, Y それぞれの周辺確率密度関数 $f_X(x)$, $f_Y(y)$ を求めよ．

(3)　X, Y それぞれの期待値及び分散を求めよ．

11.2　条件付き確率

確率の重要な概念に条件付き確率や独立性というものがあった．これらを連続型の確率変数に対しても定義をしよう．まずは条件付き確率について考え

る．B という条件のもとで A が起こる確率は

$$P(A|\,B)=\frac{P(A\cap B)}{P(B)}$$

で与えられるのであった．そこで，次のように定義をする．

定義 11.5

確率変数 X,Y の同時確率密度関数を $f(x,y)$ とする．$X=x$ という条件付きの Y の確率密度関数を次で定義する．

$$f(y|\,x)=\frac{f(x,y)}{f_X(x)}\quad(f_X(x)\neq 0)$$

同様に，$Y=y$ という条件付きの X の確率密度関数を次で定義する．

$$f(x|\,y)=\frac{f(x,y)}{f_Y(y)}\quad(f_Y(y)\neq 0)$$

実際，たとえば $Y=y_*$ という条件[1] の下での X の確率を考える際には，$A=\{(x,y);\,a\leqq x\leqq b\}$ などとすれば

$$\begin{aligned}B&=\{(X,Y);\,Y=y_*\}\,,\\ P(B)&=\int_{-\infty}^{\infty}f(x,y_*)\,dx=f_Y(y_*)\\ P(A\cap B)&=\int_a^b\int_{y=y_*}f(x,y)\,dy\,dx\\ &=\int_a^b f(x,y_*)\,dx\end{aligned}$$

のようになるので

$$P(A|\,B)=\int_a^b\frac{f(x,y_*)}{f_Y(y_*)}\,dx$$

となり，定義の妥当性が確認できる[2]．

11.3　独立性

2 つの事象が独立であるという性質は，

$$P(A\cap B)=P(A)P(B)$$

により定義された．連続型の場合，これを確率密度関数の言葉で表現することができる．

定義 11.6

連続型の確率変数 X,Y が**独立**であるとは，それらの同時確率分布において，任意の x,y に関して確率密度関数の等式

$$f(x,y)=f_X(x)\,f_Y(y)$$

が成立することをいう．

X と Y が独立であるとき，条件付き確率密度関数は

$$f(x|\,y)=f_X(x)$$

となり，X の確率分布が Y の値に影響されない．このことが**独立**という言葉の意味である．

演習問題 11.7. 次で与えられる X,Y の同時確率密度関数について，以下の問いに答えよ．

(a)　$f(x,y)=\begin{cases}2-x-y & (0\leqq x\leqq 1,\ 0\leqq y\leqq 1),\\ 0 & (\text{その他}).\end{cases}$

(b)　$f(x,y)=\begin{cases}e^{-(x+y)} & (x\geqq 0,\ y\geqq 0),\\ 0 & (\text{その他}).\end{cases}$

(1)　周辺確率密度関数 $f_X(x)$, $f_Y(y)$ を求めよ．

(2)　条件付き確率密度関数 $f(x|\,y)$ を求めよ．

(3)　条件付き確率密度関数 $f(y|\,x)$ を求めよ．

(4)　X と Y が独立かどうか調べよ．

演習問題 11.8. 演習問題 11.2 の確率変数 X,Y について，以下の問いに答えよ．

- $P(X\leqq 1|Y\leqq 2)$ を求めよ．
- X と Y が独立かどうか調べよ．

1）　固定しているということを明確にするために，y ではなく y_* と書いている．

2）　$y=y_*$ の区間で積分をすれば値は 0 なので，少し誤魔化しが入っている．$B_\varepsilon=\{(x,y);\,|y-y_*|<\varepsilon\}$ で考えて $\varepsilon\to+0$ などとすればよい．

演習問題 11

(1) 確率変数 X, Y の同時確率分布が次のように与えられている．

$X\backslash Y$	1	2	3
1	0.15	0.06	0.09
2	0.35	0.14	0.21

(a) $P(X \leqq 1,\, Y \leqq 2)$ を求めよ．

(b) 確率変数 X, Y の周辺分布を求めよ．

(c) 確率変数 X, Y の期待値及び分散を求めよ．

(2) 確率変数 X, Y の同時確率密度関数が，以下で与えられている．

$$f(x,y) = \begin{cases} ce^{-3x-y} & (x \geqq 0,\ y \geqq 0), \\ 0 & (\text{その他}). \end{cases}$$

(a) c の値を求めよ．

(b) 確率変数 X, Y それぞれの周辺確率密度関数 $f_X(x)$, $f_Y(y)$ を求めよ．

(c) X, Y それぞれの期待値及び分散を求めよ．

(3) 次で与えられる X, Y の同時確率密度関数について，以下の問いに答えよ．

$$\text{(a)}\quad f(x,y) = \begin{cases} 2-x-y & (0 \leqq x \leqq 1,\ 0 \leqq y \leqq 1), \\ 0 & (\text{その他}). \end{cases}$$

$$\text{(b)}\quad f(x,y) = \begin{cases} e^{-(x+y)} & (x \geqq 0,\ y \geqq 0), \\ 0 & (\text{その他}). \end{cases}$$

(a) 周辺確率密度関数 $f_X(x)$, $f_Y(y)$ を求めよ．

(b) 条件付き確率密度関数 $f(x|\,y)$ を求めよ．

(c) 条件付き確率密度関数 $f(y|\,x)$ を求めよ．

(d) X と Y が独立かどうか調べよ．

(4) 演習問題 (1) の確率変数 X, Y について，以下の問いに答えよ．

- $P(X \leqq 1 | Y \leqq 2)$ を求めよ．
- X と Y が独立かどうか調べよ．

(5) $\Omega = \mathbb{R}$ 上の確率分布で，次の密度関数を持つものを考える．

$$f(\omega) = \begin{cases} \dfrac{1}{2\pi} & (0 \leqq \omega \leqq 2\pi) \\ 0 & (\text{それ以外}) \end{cases}$$

確率変数 X, Y をそれぞれ $X(\omega) = \cos\omega$, $Y(\omega) = \sin\omega$ $(\omega \in \Omega)$ により定義する．

(a) X, Y は独立ではないことを示せ．

(b) $\mathbb{E}[X]$, $\mathbb{E}[Y]$ および $\mathbb{E}[XY]$ をそれぞれ計算せよ．