3　２次元確率変数

3.1　離散型の 2 次元確率変数

同時分布（joint distribution）　確率変数 X，Y を組にした (X, Y) を 2 次元確率変数という。X，Y はともに離散型確率変数とし，$X = x_i$ かつ $Y = y_j$ となる確率を $p(x_i, y_j)$ とすると，(X, Y) の確率分布は次のように表される。

$X \setminus Y$	y_1	y_2	$\cdots$	y_m	計
x_1	$p(x_1, y_1)$	$p(x_1, y_2)$	$\cdots$	$p(x_1, y_m)$	$p(x_1, \cdot)$
x_2	$p(x_2, y_1)$	$p(x_2, y_2)$	$\cdots$	$p(x_2, y_m)$	$p(x_2, \cdot)$
$\vdots$	$\vdots$	$\vdots$		$\vdots$	$\vdots$
x_ℓ	$p(x_\ell, y_1)$	$p(x_\ell, y_2)$	$\cdots$	$p(x_\ell, y_m)$	$p(x_\ell, \cdot)$
計	$p(\cdot, y_1)$	$p(\cdot, y_2)$	$\cdots$	$p(\cdot, y_m)$	1

この分布を X，Y の同時確率分布といい，$p(x_i, y_j)$ は同時分布の確率関数を表す。

$$p(x_i, y_j) = P\left(X = x_i, Y = y_j\right), \qquad \sum_{i=1}^{\ell} \sum_{j=1}^{m} p(x_i, y_j) = 1 \tag{1}$$

周辺分布（marginal distribution）　上の同時確率分布において，横方向に加えた確率は $P\left(X = x_i\right)$，縦方向に加えた確率は $P\left(Y = y_j\right)$ になる。

$$p(x_i, \cdot) = \sum_{j=1}^{m} p(x_i, y_j) = P\left(X = x_i\right), \qquad \sum_{i=1}^{\ell} p(x_i, \cdot) = 1 \tag{2}$$

$$p(\cdot, y_j) = \sum_{i=1}^{\ell} p(x_i, y_j) = P\left(Y = y_j\right), \qquad \sum_{j=1}^{m} p(\cdot, y_j) = 1 \tag{3}$$

同時確率分布と対比して，X のみまたは Y のみの分布のことを周辺確率分布という。$p(x_i, \cdot)$ は X の周辺分布，$p(\cdot, y_j)$ は Y の周辺分布の確率関数を表す。

X	x_1	$\cdots$	x_ℓ
P	$p(x_1, \cdot)$	$\cdots$	$p(x_\ell, \cdot)$

Y	y_1	$\cdots$	y_m
P	$p(\cdot, y_1)$	$\cdots$	$p(\cdot, y_m)$

独立な確率変数　X，Y は離散型確率変数とする。X，Y のすべての値 x_i，y_j について $p(x_i, y_j) = p(x_i, \cdot) \times p(\cdot, y_j)$ が成り立つとき，X，Y は独立であるという。

$$\begin{aligned} X，Y \text{ が独立} &\iff p(x_i, y_j) = p(x_i, \cdot) \times p(\cdot, y_j) \\ &\iff P\left(X = x_i, Y = y_j\right) = P\left(X = x_i\right) \times P\left(Y = y_j\right) \end{aligned} \tag{4}$$

次の分布表の Y と Z は周辺分布としては同一である。しかし，左表の同時分布に従う X，Y は独立であるが，右表の同時分布に従う X，Z は独立でない。

$X \setminus Y$	1	2	3	計
1	0.09	0.15	0.06	0.3
2	0.12	0.20	0.08	0.4
3	0.09	0.15	0.06	0.3
計	0.3	0.5	0.2	1

$X \setminus Z$	1	2	3	計
1	0.15	0.12	0.03	0.3
2	0.09	0.26	0.05	0.4
3	0.06	0.12	0.12	0.3
計	0.3	0.5	0.2	1

3.2　連続型の 2 次元確率変数

X，Y は連続型確率変数とする。2 次元確率分布が確率密度関数 $f(x, y)$ をもつとき，同時分布は次のように表される。

$$P\left(a_1 \leqq X \leqq a_2,\ b_1 \leqq Y \leqq b_2\right) = \int_{b_1}^{b_2} \int_{a_1}^{a_2} f(x, y)\, dxdy$$

周辺分布の確率密度関数は次のように表される。

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

X，Y は連続型確率変数とする。X，Y のすべての値 x，y について $f(x, y) = f(x, \cdot) \times f(\cdot, y)$ が成り立つとき，X，Y は独立であるという。

$$X，Y \text{ が独立} \iff f(x, y) = f(x, \cdot) \times f(\cdot, y) \tag{5}$$

3.3　確率変数の和の期待値

(X, Y) は 2 次元確率変数，$g(x, y)$ は実数値 2 変数関数とする。離散型または連続型の場合の $g(X, Y)$ の期待値を次のように定める。

$$E\left[g(X, Y)\right] = \sum_i \sum_j g(x_i, y_j) p(x_i, y_j)$$
$$E\left[g(X, Y)\right] = \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} g(x, y) f(x_i, y_j) dxdy$$

離散型確率変数 X，Y について，その和 $X + Y$ の期待値を求める。$g(x, y) = x + y$ として上の定義式を用いる（X，Y が連続型の場合も同様である）。

$$\begin{aligned} E\left[X + Y\right] &= \sum_{i=1}^{\ell} \sum_{j=1}^{m} (x_i + y_j) p(x_i, y_j) \\ &= \sum_{i=1}^{\ell} x_i \sum_{j=1}^{m} p(x_i, y_j) + \sum_{j=1}^{m} y_j \sum_{i=1}^{\ell} p(x_i, y_j) \\ &= \sum_{i=1}^{\ell} x_i\, p(x_i, \cdot) + \sum_{j=1}^{m} y_j\, p(\cdot, y_j) = E\left[X\right] + E\left[Y\right] \end{aligned}$$

X，Y は離散型または連続型の確率変数，a は定数とする。2.3 節の結果と合わせて，$E[\ ]$ は次の性質をもつ。

$$E[aX] = aE[X], \quad E[a] = a, \quad E[X+Y] = E[X] + E[Y] \tag{6}$$

3.4 確率変数の和の分散

分散　離散型確率変数 X の期待値を $\mu = E[X]$ とおく。$(X-\mu)^2$ の期待値を X の分散といい，$V[X]$ と表す（連続型の場合も同様である）。

$$V[X] = E\left[(X-\mu)^2\right] = \sum_{i=1}^{\ell}(x_i-\mu)^2\,p(x_i,\cdot) \tag{7}$$

$$V[X] = E\left[X^2\right] - \{E[X]\}^2 \tag{8}$$

共分散　離散型確率変数 X，Y の期待値を各々$\mu = E[X]$，$\nu = E[Y]$ とおく [*i]。$(X-\mu)(Y-\nu)$ の期待値を X，Y の共分散といい，$C[X,Y]$ と表す（連続型の場合も同様である）。 [*i]

$$C[X,Y] = E[(X-\mu)(Y-\nu)] = \sum_{i=1}^{\ell}\sum_{j=1}^{m}(x_i-\mu)(y_j-\nu)\,p(x_i,y_j) \tag{9}$$

離散型確率変数の場合も $C[X,Y] = E[XY] - E[X]\,E[Y]$ が成り立つ。

$$\begin{aligned}
C[X,Y] &= \sum_i\sum_j(x_i-\mu)(y_j-\nu)\,p(x_i,y_j)\\
&= \sum_i\sum_j x_iy_j\,p(x_i,y_j) - \mu\sum_j y_j\sum_i p(x_i,y_j)\\
&\qquad - \nu\sum_i x_i\sum_j p(x_i,y_j) + \mu\nu\sum_i\sum_j p(x_i,y_j)\\
&= \sum_i\sum_j x_iy_j\,p(x_i,y_j) - \mu\sum_j y_jp(\cdot,y_j) - \nu\sum_i x_ip(x_i,\cdot) + \mu\nu\\
&= E[XY] - \mu E[Y] - \nu E[X] + \mu\nu\\
&= E[XY] - \mu\nu
\end{aligned}$$

よって共分散は次のように表せる。

$$C[X,Y] = E[XY] - E[X]\,E[Y] \tag{10}$$

分散と共分散　(7) と (9)，あるいは (8) と (10) の比較により，X，X の共分散は X の分散に等しい。

$$C[X,X] = V[X]$$

[*i] μ（ミュー），ν（ニュー）はギリシャ文字。

独立性と共分散　離散型確率変数 X，Y は独立とする。(9)，(4) より

$$\begin{aligned}
C[X,Y] &= \sum_i \sum_j (x_i-\mu)(y_j-\nu)\, p(x_i,y_j) \\
&= \sum_i \sum_j (x_i-\mu)(y_j-\nu)\, p(x_i,\cdot)p(\cdot,y_j) \\
&= \sum_i (x_i-\mu)p(x_i,\cdot) \times \sum_j (y_j-\nu)p(\cdot,y_j) = 0 \times 0
\end{aligned}$$

*ii　よって独立ならば無相関となる [*ii]（連続型の場合も同様である）。

$$X,\ Y \text{ が独立} \implies C[X,Y]=0 \tag{11}$$

ただし逆（無相関ならば独立）は成り立たない。たとえば次の表の確率変数 X，Y は無相関であるが，独立ではない。

$X \setminus Y$	-1	0	1	計
-1	0	0.25	0	0.25
0	0.25	0	0.25	0.50
1	0	0.25	0	0.25
計	0.25	0.50	0.25	1

確率変数の和の分散　$X+Y$ の分散を求める。$\mu=E[X]$，$\nu=E[Y]$ とおく。

$$\begin{aligned}
V[X+Y] &= E\left[\{(X+Y)-E[X+Y]\}^2\right] \\
&= E\left[\{(X+Y)-(\mu+\nu)\}^2\right] = E\left[\{(X-\mu)+(Y-\nu)\}^2\right] \\
&= E\left[(X-\mu)^2\right] + 2E[(X-\mu)(Y-\nu)] + E\left[(Y-\nu)^2\right] \\
&= V[X] + 2C[X,Y] + V[Y]
\end{aligned}$$

X，Y は離散型または連続型の確率変数，a は定数とする。2.3 節の結果と合わせて，$V[\ \ ]$ は次の性質をもつ。

$$\begin{gathered}
V[aX] = a^2V[X], \qquad V[a]=0 \\
V[X+Y] = V[X] + 2C[X,Y] + V[Y]
\end{gathered} \tag{12}$$

特に，独立の場合は次が成り立つ。

$$X,\ Y \text{ が独立} \implies V[X+Y] = V[X] + V[Y] \tag{13}$$

[*ii] 無相関は，$C[X,Y]=0$ の言い換えである。

本稿の参考文献

統計学入門（基礎統計学）		
東京大学教養学部統計学教室（編）	東京大学出版会	978-4-13-042065-5
統計学		
久保川 達也 （著）	東京大学出版会	978-4-13-062921-8
確率統計 新版 （新版数学シリーズ）		
岡本 和夫 （ほか著）	実教出版	978-4-407-32171-5
統計学序論 改訂版		
山本 義郎 （著）	東海大学出版部	978-4-486-02133-9
確率統計（高専テキストシリーズ）		
上野 健爾（監修）	森北出版	978-4-627-05561-2
新統計入門		
小寺 平治 （著）	裳華房	978-4-7853-1099-8
Schaum's Outline of Introduction to Probability and Statistics		
Seymour Lipschutz （著）	McGraw-Hill Education	978-0-07-176249-6

`www5e.biglobe.ne.jp/~emm386/statistics/`