2.0　多次元確率変数 (続き)

連続型の場合: 2 次元確率変数 (X,Y) が連続型であるとは、ある関数 $f(x,y)$

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

と表されるときをいい、この。このとき、

$$f(x,y) \ge 0, \qquad \int_{-\infty}^{\infty}\int_{-\infty}^{\infty} f(x,y)\,dxdy = 1$$

が成り立つ。また、これが成り立つとき、$f(x,y)$ は密度関数となる。

事象 $a \le X \le b$ は $a \le X \le b, -\infty < Y < \infty$ で表される事象と考えられるから、

$$P(a \le X \le b) = \int_a^b \Big(\int_{-\infty}^{\infty} f(x,y)\,dy \Big)\,dx = \int_a^b f_X(x)\,dx \qquad \text{ただし} \quad f_X(x) = \int_{-\infty}^{\infty} f(x,y)\,dy$$

となる。これを X の周辺分布、$f_X(x)$ は X の周辺密度関数という。同様に、

$$P(c \le Y \le d) = \int_c^d f_Y(y)\,dy \qquad \text{ただし} \quad f_Y(y) = \int_{-\infty}^{\infty} f(x,y)\,dx$$

を Y の周辺分布、$f_Y(x)$ は Y の周辺密度関数という。

(確率変数の独立性) X,Y が独立であるとは、$f(x,y) = f_X(x)f_Y(y)$ が成立するときにいう。

2.1　確率変数の関数

(X,Y) の同時密度関数が $f(x,y)$ をもつとき、X,Y の関数 $\phi(X,Y)$ の平均 $E[\phi(X,Y)]$ を

$$E[\phi(X,Y)] = \int_{-\infty}^{\infty}\int_{-\infty}^{\infty} \phi(x,y)f(x,y)\,dxdy$$

と定め、分散 $V[\phi(X,Y)]$ を

$$V[\phi(X,Y)] = E[\Big(\phi(X,Y) - E[\phi(X,Y)]\Big)^2]$$

で定める。また、$\mu_1 = E[X]$, $\mu_2 = E[Y]$ とし、X,Y の共分散 $\mathrm{Cov}[X,Y]$ を次で定める。

$$\mathrm{Cov}[X,Y] = E[(X-\mu_1)(Y-\mu_2)] = E[XY] - \mu_2 E[X] - \mu_1 E[Y] + \mu_1\mu_2 = E[XY] - \mu_1\mu_2.$$

定理 2.1 X,Y が独立であれば、$E[XY] = E[X]E[Y]$. 特に、$\mathrm{Cov}[X,Y] = 0$.

証明: (X,Y) の同時密度関数が $f(x,y)$ とすると、$f(x,y) = f_X(x)f_Y(y)$ なので、

$$\begin{aligned} E[XY] &= \int_{-\infty}^{\infty}\int_{-\infty}^{\infty} xyf(x,y)\,dxdy = \int_{-\infty}^{\infty}\int_{-\infty}^{\infty} xyf_X(x)f_Y(y)\,dxdy \\ &= \int_{-\infty}^{\infty} xf_X(x)\,dx \int_{-\infty}^{\infty} yf_Y(y)\,dy = E[X]E[Y]. \quad \square \end{aligned}$$

定理 2.2 (平均と分散の性質) 定数 a,b,c に対して、$E[aX+bY+c] = aE[X] + bE[Y] + c$ となる。
さらに、X,Y が独立であれば、$V[aX+bY+c] = a^2V[X] + b^2V[Y]$ となる。

証明: $$\begin{aligned} E[aX+bY+c] &= \int_{-\infty}^{\infty}\int_{-\infty}^{\infty} (ax+by+c)f(x,y)\,dxdy = a\int_{-\infty}^{\infty} xf_X(x)\,dx + b\int_{-\infty}^{\infty} yf_Y(y)\,dy + c \\ &= aE[X] + bE[Y] + c. \end{aligned}$$
分散については離散のときとまったく同様なので省略する。　□

例 2.3 (2 次元正規分布) 2 次元確率変数 (X, Y) の同時密度関数が次で与えられるとき、(X, Y) は 2 次元正規分布に従うという。ただし $-\infty < \mu_1, \mu_2 < \infty$, $\sigma_1, \sigma_2 > 0$, $-1 < \rho < 1$ である。

$$f(x,y) = \frac{1}{2\pi\sigma_1\sigma_2\sqrt{1-\rho^2}} \exp\left[-\frac{1}{2(1-\rho^2)}\left\{\frac{(x-\mu_1)^2}{\sigma_1^2} - \frac{2\rho(x-\mu_1)(y-\mu_2)}{\sigma_1\sigma_2} + \frac{(y-\mu_2)^2}{\sigma_2^2}\right\}\right]$$

このとき、$E[X] = \mu_1$, $E[Y] = \mu_2$, $V[X] = \sigma_1^2$, $V[Y] = \sigma_2^2$, $\mathrm{Cov}[X, Y] = \rho\sigma_1\sigma_2$ となる。
ここで、もし $\rho = 0$ であれば、

$$f(x,y) = f_X(x) f_Y(y) \qquad \text{ただし} \quad f_X(x) = \frac{1}{\sqrt{2\pi}\sigma_1} e^{-\frac{(x-\mu_1)^2}{2\sigma_1^2}}, \quad f_Y(y) = \frac{1}{\sqrt{2\pi}\sigma_2} e^{-\frac{(y-\mu_2)^2}{2\sigma_2^2}}$$

と表せる。ここで、$f_X(x)$ は正規分布 $N(\mu_1, \sigma_2^2)$, $f_Y(x)$ は $N(\mu_2, \sigma_2^2)$ の密度関数となる。従って、$\rho = 0$ のとき X, Y は独立で、X, Y の周辺確率密度関数がそれぞれ $f_X(x), f_Y(y)$ となることがわかる。

共分散は X, Y が互いに関連しながらそれぞれの平均 $E[X], E[Y]$ からばらつく程度を表している。特に、共分散が 0 のとき、X, Y は無相関であるという。定理 2.1 より X, Y が独立であれば X, Y は無相関であるが、逆は必ずしも成り立たない (正規分布の場合は成り立つが、これは特別な場合である)。

また、確率変数 X, Y の相関係数 $\rho[X, Y]$ を $\rho[X, Y] = \dfrac{\mathrm{Cov}[X, Y]}{\sqrt{V[X]V[Y]}}$ で定める。$-1 \leq \rho[X, Y] \leq 1$ となることが知られている。

3 つ以上の確率変数 $X_1, X_2, \ldots, X_n$ を考える。$X_1, X_2, \ldots, X_n$ が独立であるとは (離散型、連続型も含んだ形で)、任意の実数 $a_1, b_1, \ldots, a_n, b_n$ $(a_i \leq b_i,\ 1 \leq i \leq n)$ に対して

$$P(a_1 \leq X_1 \leq b_1, \cdots, a_n \leq X_n \leq b_n) = P(a_1 \leq X_1 \leq b_1) \cdots P(a_n \leq X_n \leq b_n)$$

が成り立つときと定義する。また、2 次元確率変数の場合と同様に、同時確率分布や同時密度関数を用いて、$X_1, X_2, \ldots, X_n$ の関数の平均や分散が定義でき、さらに、定理 2.2 と同様に次の定理が成立する。

定理 2.3 (平均と分散の性質) (1) $a_1, a_2, \ldots, a_n$ が定数のとき

$$E[a_1X_1 + a_2X_2 + \cdots + a_nX_n] = a_1E[X_1] + a_2E[X_2] + \cdots + a_nE[X_n].$$

(2) $a_1, a_2, \ldots, a_n$ が定数で $X_1, X_2, \ldots, X_n$ が独立のとき

$$V[a_1X_1 + a_2X_2 + \cdots + a_nX_n] = {a_1}^2V[X_1] + {a_2}^2V[X_2] + \cdots + {a_n}^2V[X_n].$$

2.2 母集団と標本 (各自教科書を読んでおいてください。)

2.3 統計量と標本分布

$X_1, X_2, \cdots, X_n$ を大きさ n の無作為標本とする。これは、数学的には $X_1, X_2, \cdots, X_n$ は同一の分布に従う独立な確率変数として定義される。この $X_1, X_2, \cdots, X_n$ の関数を統計量というが、よく用いられる統計量には次のものがある。

$$\overline{X} = \frac{1}{n}\sum_{i=1}^{n} X_i, \qquad S^2 = \frac{1}{n}\sum_{i=1}^{n}(X_i - \overline{X})^2, \qquad U^2 = \frac{1}{n-1}\sum_{i=1}^{n}(X_i - \overline{X})^2,$$

$\overline{X}, S^2, U^2$ をそれぞれ標本平均、標本分散、不偏分散という。このとき、次が成立する。

定理 2.4 $X_1, X_2, \cdots, X_n$ を母平均 μ, 母分散 σ^2 の母集団からの大きさ n の無作為標本とする。すなわち、$E[X_i] = \mu$, $V[X_i] = \sigma^2$ とする。このとき、標本平均 $\overline{X}$ と不変分散 U^2 に対して次が成立する。

$$E[\overline{X}] = \mu, \qquad V[\overline{X}] = \frac{\sigma^2}{n}, \qquad E[U^2] = \sigma^2.$$

証明: 標本平均 $\overline{X}$ については定理 2.3 より明らか。不偏分散については時間があれば授業中に説明する。 □