

### 3.3 条件付き期待値

定義 3.10  $X = x$  が与えられたときの  $Y$  の条件付確率関数または条件付確率密度関数を  $f_{Y|X}(y|x)$  とする .  $g: \mathbb{R} \rightarrow \mathbb{R}$  としたとき ,  $X = x$  が与えられたときの  $g(Y)$  の条件付期待値を

$$\mathbb{E}[g(Y)|x] = \begin{cases} \sum_y g(y)f_{Y|X}(y|x), & (\text{離散型}), \\ \int_{-\infty}^{\infty} g(y)f_{Y|X}(y|x) dy, & (\text{連続型}) \end{cases}$$

で定める . ただし , 条件付期待値は  $\mathbb{E}[|g(Y)||x] < \infty$  のときに存在するものとする .

命題 3.5 (条件付期待値の性質)  $a_1, a_2, b$  を定数 ,  $g_1: \mathbb{R} \rightarrow \mathbb{R}, g_2: \mathbb{R} \rightarrow \mathbb{R}$  とする .

- (1)  $\mathbb{E}[a_1g_1(Y) + a_2g_2(Y) + b|x] = a_1\mathbb{E}[g_1(Y)|x] + a_2\mathbb{E}[g_2(Y)|x] + b.$
- (2)  $g_1(y) \geq 0$  ならば ,  $\mathbb{E}[g_1(Y)|x] \geq 0.$
- (3)  $a_1 \leq g_1(y) \leq a_2$  ならば ,  $a_1 \leq \mathbb{E}[g_1(Y)|x] \leq a_2.$
- (4)  $\mathbb{E}[g_1(X)g_2(Y)|x] = g_1(x)\mathbb{E}[g_2(Y)|x].$

が成立する . ただし , 上の条件付期待値はすべて存在するものとする .

証明 積分の性質からわかる . □

定義 3.11  $X$  と  $Y$  を確率変数とし ,  $\mathbb{E}[Y^2] < \infty$  とする .  $X = x$  が与えられたときの  $Y$  の条件付分散を

$$\text{VAR}[Y|x] = \mathbb{E}[Y^2|x] - \{\mathbb{E}[Y|x]\}^2$$

で定義する .  $v(x) := \text{VAR}[Y|x]$  においてとき ,  $\text{VAR}[Y|X] := v(X)$  で定める . 以上の定義から  $\text{VAR}[Y|X] = \mathbb{E}[Y^2|X] - \{\mathbb{E}[Y|X]\}^2$  となることに注意する .

例 3.4 連続型確率ベクトル  $(X, Y)$  は同時確率密度関数

$$f_{X,Y}(x,y) = \begin{cases} e^{-y}, & 0 < x < y < \infty, \\ 0, & \text{その他} \end{cases}$$

を持つとき ,  $X = x$  が与えられたときの  $Y$  の条件付確率密度関数は  $x > 0$  の場合のみに定義され , 各  $x > 0$  に対して

$$\begin{aligned} f_{Y|X}(y|x) &= \frac{f_{X,Y}(x,y)}{f_X(x)} = \frac{e^{-y}}{e^{-x}} = e^{-(y-x)}, & y > x, \\ f_{Y|X}(y|x) &= \frac{f_{X,Y}(x,y)}{f_X(x)} = \frac{0}{e^{-x}} = 0, & y \leq x \end{aligned}$$

であった .  $X = x (x > 0)$  が与えられたときの  $Y$  の条件付期待値は

$$\mathbb{E}[Y|x] = \int_{-\infty}^{\infty} y f_{Y|X}(y|x) dy = \int_x^{\infty} y e^{-(y-x)} dy = 1 + x$$

となる . 同様に ,  $X = x (x > 0)$  が与えられたときの  $Y^2$  の条件付期待値は

$$\mathbb{E}[Y^2|x] = \int_{-\infty}^{\infty} y^2 f_{Y|X}(y|x) dy = \int_x^{\infty} y^2 e^{-(y-x)} dy = \int_0^{\infty} (t+x)^2 e^{-t} dt = x^2 + 2x + 2$$

となる . したがって ,  $X = x (x > 0)$  が与えられたときの  $Y$  の条件付分散は

$$\text{VAR}[Y|x] = \mathbb{E}[Y^2|x] - \{\mathbb{E}[Y|x]\}^2 = 1$$

となる .

$X = x$  が与えられたときの  $g(Y)$  の条件付期待値 (存在するならば)  $\mathbb{E}[g(Y)|x]$  は  $x$  の関数であるので,  $h(x) = \mathbb{E}[g(Y)|x]$  とおくと, 確率変数  $\mathbb{E}[g(Y)|X]$  を  $\mathbb{E}[g(Y)|X] = h(X)$  で定めることにする. すなわち,  $X = x$  のとき, 確率変数  $\mathbb{E}[g(Y)|X]$  の値は  $\mathbb{E}[g(Y)|x]$  である. 記号の読み方であるが,  $\mathbb{E}[g(Y)|X]$  に期待値の記号  $\mathbb{E}$  が使われているが,  $\mathbb{E}[g(Y)|X]$  は  $X$  に依存する確率変数である.

**定理 3.3**  $X$  と  $Y$  を確率変数とし,  $Y$  の期待値は存在するとする. このとき,

$$\mathbb{E}[Y] = \mathbb{E}[\mathbb{E}[Y|X]]$$

が成立する.

**証明**  $X$  と  $Y$  が連続型確率変数の場合の証明を与える.  $(X, Y)$  の同時確率密度関数を  $f_{X,Y}(x, y)$  とする.  $X = x$  が与えられたときの  $Y$  の条件付確率密度関数  $f_{Y|X}(y|x)$  および条件付期待値の定義から

$$\begin{aligned} \mathbb{E}[Y] &= \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} y f_{X,Y}(x, y) dx dy \\ &= \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} y \frac{f_{X,Y}(x, y)}{f_X(x)} f_X(x) dx dy \\ &= \int_{-\infty}^{\infty} \left[ \int_{-\infty}^{\infty} y f_{Y|X}(y|x) dx \right] f_X(x) dx \\ &= \int_{-\infty}^{\infty} \mathbb{E}[Y|x] f_X(x) dx \\ &= \mathbb{E}[\mathbb{E}[Y|X]] \end{aligned}$$

となる. □

**例 3.5** 確率変数  $X$  と  $Y$  は同時確率密度関数

$$f_{X,Y}(x, y) = \begin{cases} 2 & x + y < 1, x > 0, y > 0, \\ 0 & (\text{その他}) \end{cases}$$

を持つとする. このとき,  $X$  の周辺確率密度関数は,  $0 < x < 1$  のとき,

$$f_X(x) = \int_{-\infty}^{\infty} f_{X,Y}(x, y) dy = \int_0^{-x+1} 2 dy = 2(1-x)$$

となる. したがって,  $X$  の周辺確率密度関数は

$$f_X(x) = \begin{cases} 2(1-x) & 0 < x < 1, \\ 0 & (\text{その他}), \end{cases}$$

である. また,  $0 < y < 1$  に対して

$$f_Y(y) = \int_{-\infty}^{\infty} f_{X,Y}(x, y) dx = \int_0^{-y+1} 2 dx = 2(1-y)$$

となることから  $Y$  の周辺確率密度関数は

$$f_Y(y) = \begin{cases} 2(1-y) & 0 < y < 1, \\ 0 & (\text{その他}), \end{cases}$$

となる．これから  $X$  の二次までの積率を求めると

$$\begin{aligned}\mathbb{E}[Y] &= \int_{-\infty}^{\infty} x f_Y(y) dy = \int_0^1 y 2(1-y) dy = \frac{1}{3}, \\ \mathbb{E}[Y^2] &= \int_{-\infty}^{\infty} y^2 f_Y(y) dy = \int_0^1 y^2 2(1-y) dy = \frac{1}{6}, \\ \text{VAR}[Y] &= \mathbb{E}[Y^2] - \{\mathbb{E}[Y]\}^2 = \frac{1}{18}\end{aligned}$$

となる． $0 < x < 1$  としたとき， $X = x$  が与えられたときの  $Y$  の条件付確率密度関数は

$$f_{Y|X}(y|x) = \begin{cases} \frac{1}{1-x} & 0 < y < 1-x, \\ 0 & (\text{その他}) \end{cases}$$

となる．したがって， $X = x$  ( $0 < x < 1$ ) が与えられたときの  $Y$  の条件付き期待値は

$$\mathbb{E}[Y|x] = \int_{-\infty}^{\infty} y f_{Y|X}(y|x) dy = \int_0^{1-x} y \frac{1}{1-x} dy = \frac{1}{1-x} \left[ \frac{y^2}{2} \right]_0^{1-x} = \frac{1-x}{2}$$

となる．さらに，

$$\mathbb{E}[\mathbb{E}[Y|X]] = \int_{-\infty}^{\infty} \mathbb{E}[Y|x] f_X(x) dx = \int_0^1 \frac{1-x}{2} 2(1-x) dx = \frac{1}{3}$$

となる．

次に， $X = x$  ( $0 < x < 1$ ) が与えられたときの  $Y$  の条件付き分散を求める．そのために，

$$\mathbb{E}[Y^2|x] = \int_{-\infty}^{\infty} y^2 f_{Y|X}(y|x) dy = \int_0^{1-x} y^2 \frac{1}{1-x} dy = \frac{1}{1-x} \left[ \frac{y^3}{3} \right]_0^{1-x} = \frac{(1-x)^2}{3}$$

となる．よって， $X = x$  ( $0 < x < 1$ ) が与えられたときの  $Y$  の条件付き分散は

$$\text{VAR}[Y|x] = \mathbb{E}[Y^2|x] - \{\mathbb{E}[Y|x]\}^2 = \frac{(1-x)^2}{3} - \frac{(1-x)^2}{4} = \frac{(1-x)^2}{12}$$

となる．

**定理 3.4**  $X, Y$  を確率変数とし， $\mathbb{E}[Y^2] < \infty$  とする．このとき，

$$\text{VAR}[Y] = \mathbb{E}[\text{VAR}[Y|X]] + \text{VAR}[\mathbb{E}[Y|X]]$$

である．ただし， $h(x) = \text{VAR}[Y|x]$  とおいたとき， $\text{VAR}[Y|X] = h(X)$  で定めた．

**証明** まず，

$$\begin{aligned}\text{VAR}[Y] &= \mathbb{E}\{\{Y - \mathbb{E}[Y]\}^2\} = \mathbb{E}\{\{Y - \mathbb{E}[Y|X] + \mathbb{E}[Y|X] - \mathbb{E}[Y]\}^2\} \\ &= \mathbb{E}\{\{Y - \mathbb{E}[Y|X]\}^2\} + \mathbb{E}\{\{\mathbb{E}[Y|X] - \mathbb{E}[Y]\}^2\} \\ &\quad + 2\mathbb{E}\{\{Y - \mathbb{E}[Y|X]\}\{\mathbb{E}[Y|X] - \mathbb{E}[Y]\}\}\end{aligned}$$

に注意する．上の式の最右辺の 3 項目は

$$\mathbb{E}\{\{Y - \mathbb{E}[Y|X]\}\{\mathbb{E}[Y|X] - \mathbb{E}[Y]\}\} = \mathbb{E}[\mathbb{E}\{\{Y - \mathbb{E}[Y|X]\}\{\mathbb{E}[Y|X] - \mathbb{E}[Y]\}|X]]$$

となる．ここで， $X$  が与えられたとき， $\mathbb{E}[Y|X]$  と  $\mathbb{E}[Y]$  は定数であることに注意して，上の式の右辺の期待値の中を評価する：

$$\begin{aligned}\mathbb{E}\{\{Y - \mathbb{E}[Y|X]\}\{\mathbb{E}[Y|X] - \mathbb{E}[Y]\}|X\} &= \{\mathbb{E}[Y|X] - \mathbb{E}[Y]\}\mathbb{E}\{Y - \mathbb{E}[Y|X]|X\} \\ &= \{\mathbb{E}[Y|X] - \mathbb{E}[Y]\}\{\mathbb{E}[Y|X] - \mathbb{E}[Y|X]\} \\ &= 0\end{aligned}$$

となる．したがって，

$$\mathbb{E}\{\{Y - \mathbb{E}[Y|X]\}\{\mathbb{E}[Y|X] - \mathbb{E}[Y]\}\} = 0$$

となり，定理は証明された． □

例 3.6  $X_1, X_2, \dots, X_n, \dots$  は独立に同一分布に従う確率変数の列で

$$\mathbb{E}[X_1] = \mu_X, \quad \text{VAR}[X_1] = \sigma_X^2 < \infty$$

とする． $N$  は非負整数値確率変数で確率変数列  $\{X_i\}_{i=1}^{\infty}$  とは独立し，

$$\mathbb{E}[N] = \mu_N, \quad \text{VAR}[N] = \sigma_N < \infty$$

とする．このとき， $N$  個までの確率変数列  $\{X_i\}_{i=1}^{\infty}$  の部分 and を

$$S_N = \sum_{i=1}^N X_i$$

と定義する．ただし， $S_0 = 0$  とする．定理 3.3 より

$$\mathbb{E}[S_N] = \mathbb{E}[\mathbb{E}[S_N|N]]$$

がわかる．しかし，

$$\mathbb{E}[S_N|n] = \mathbb{E}[X_1 + X_2 + \dots + X_n] = n\mu_X$$

より  $\mathbb{E}[S_N|N] = N\mu_X$  となる．これから

$$\mathbb{E}[S_N] = \mathbb{E}[N\mu_X] = \mathbb{E}[N]\mu_X = \mu_X\mu_N$$

がわかる．同様に，

$$\text{VAR}[S_N|n] = \text{VAR}[X_1 + X_2 + \dots + X_n|n] = \sum_{i=1}^n \text{VAR}[X_i|n] = n\text{VAR}[X_1|n] = n\sigma_X^2$$

から  $\text{VAR}[S_N|N] = N\sigma_X^2$  がわかる．定理 3.4 から

$$\begin{aligned}\text{VAR}[S_N] &= \mathbb{E}[\text{VAR}[S_N|N]] + \text{VAR}[\mathbb{E}[S_N|N]] \\ &= \sigma_X^2 \mathbb{E}[N] + \text{VAR}[N\mu_X] \\ &= \sigma_X^2 \mathbb{E}[N] + \mu_X^2 \text{VAR}[N] \\ &= \sigma_X^2 \mu_N + \mu_X^2 \sigma_N^2\end{aligned}$$

となる．

$N$  は母数  $p$  ( $0 < p \leq 1$ ) の幾何分布に従うとする：

$$f_N(n) = \begin{cases} p(1-p)^{n-1} & n = 1, 2, \dots \\ 0 & (\text{その他}) \end{cases}$$

とする。したがって

$$\mathbb{E}[N] = \frac{1}{p}, \quad \text{VAR}[N] = \frac{1-p}{p^2}, \quad m_N(t) = \frac{pe^t}{1-(1-p)e^t}$$

となる。これは

$$\begin{aligned} \mathbb{E}[e^{tN}] &= \sum_{n=1}^{\infty} e^{tn} p(1-p)^{n-1} = \sum_{n=0}^{\infty} pe^t \{(1-p)e^t\}^{n-1} \\ &= \frac{pe^t}{1-(1-p)e^t} \sum_{n=0}^{\infty} \{1-(1-p)e^t\} \{1-(1-p)e^t\}^{n-1} \\ &= \frac{pe^t}{1-(1-p)e^t} \end{aligned}$$

よりわかる。さらに、 $X_1$  は母数  $\lambda (\lambda > 0)$  の指数分布に従うとする：

$$f_X(x) = \begin{cases} \lambda e^{-\lambda x} & x > 0 \\ 0 & (\text{その他}) \end{cases}$$

とする。したがって、

$$\mathbb{E}[X] = \frac{1}{\lambda}, \quad \text{VAR}[X] = \frac{1}{\lambda^2}, \quad m_X(t) = \frac{\lambda}{\lambda-t}, \quad t < \lambda$$

である。

定理 3.3 から

$$\mathbb{E}[S_N] = \mu_N \mu_X = \frac{1}{p} \frac{1}{\lambda} = \frac{1}{p\lambda}$$

となる。定理 3.4 から

$$\text{VAR}[S_N] = \mu_N \sigma_X^2 + \sigma_N^2 \mu_X^2 = \frac{1}{p} \frac{1}{\lambda^2} + \frac{1-p}{p^2} \frac{1}{\lambda^2} = \frac{1}{(p\lambda)^2}$$

がわかる。

$S_n$  の分布を求めてみよう。そのために、

$$I\{S_N \leq z\} = \begin{cases} 1 & (S_N \leq z \text{ が起きたとき}) \\ 0 & (\text{その他}) \end{cases}$$

とおけば、 $\mathbb{P}[S_N \leq z] = \mathbb{E}[I\{S_N \leq z\}]$  と  $\mathbb{P}[S_n \leq z | N = n] = \mathbb{E}[I\{S_N \leq z\} | n]$  となることと  $S_n$  は母数  $n, \lambda$  のガンマ分布

$$f_{S_n}(x) = \begin{cases} \frac{\lambda^n}{\Gamma(n)} (\lambda x)^{n-1} \lambda e^{-\lambda x} & x > 0 \\ 0 & (\text{その他}) \end{cases}$$

従うことに注意する .  $z > 0$  に対して

$$\begin{aligned}
 \mathbb{P}[S_N \leq z] &= \mathbb{E}[I\{S_N \leq z\}] \\
 &= \mathbb{E}[\mathbb{E}[I\{S_N \leq z\} | N]] \\
 &= \sum_{n=0}^{\infty} \mathbb{E}[I\{S_N \leq z\} | n] f_N(n) \\
 &= \sum_{n=1}^{\infty} \mathbb{E}[I\{S_N \leq z\} | n] f_N(n) \\
 &= \sum_{n=1}^{\infty} \mathbb{P}[S_N \leq z | N = n] f_N(n) \\
 &= \sum_{n=1}^{\infty} \left( \int_0^z \frac{\lambda}{\Gamma(n)} (\lambda x)^{n-1} \lambda e^{-\lambda x} dx \right) p(1-p)^{n-1} \\
 &= \sum_{n=1}^{\infty} \int_0^z \lambda p e^{-\lambda z} \frac{[\lambda x(1-p)]^{n-1}}{(n-1)!} dx \\
 &= \int_0^{\infty} \lambda p e^{-\lambda z} \sum_{n=1}^{\infty} \left( \frac{[\lambda x(1-p)]^{n-1}}{(n-1)!} \right) dx \\
 &= \int_0^{\infty} \lambda p e^{-\lambda z} e^{(1-p)\lambda x} dx \\
 &= \lambda p \int_0^{\infty} e^{-\lambda p x} dx \\
 &= 1 - e^{-\lambda p z}
 \end{aligned}$$

となる . 最後から 4 番目の等号における和の記号と積分記号の順序交換はつきから保障されることがわかる :

$$\int_0^z \frac{\lambda}{\Gamma(n)} (\lambda x)^{n-1} \lambda e^{-\lambda x} dx \leq 1$$

から

$$\left| \sum_{n=1}^{\infty} \left( \int_0^z \frac{\lambda}{\Gamma(n)} (\lambda x)^{n-1} \lambda e^{-\lambda x} dx \right) p(1-p)^{n-1} \right| \leq \sum_{n=1}^{\infty} p(1-p)^{n-1} = 1$$

よりわかる .

したがって ,  $S_N$  は母数  $p\lambda$  の指数分布に従うことがわかる .