確率・統計 A

若木宏文

wakaki@math.sci.hiroshima-u.ac.jp
http://home.hiroshima-u.ac.jp/ wakaki/lecture/index.shtml

2015.6.2

目 次

3.1. 平均の定義

離散型確率変数の平均

定義 3.1

確率変数 X は離散型で, その確率関数を $f(x)$ とする.
X のとりうる値が有限個で, 取り得る値の集合が $D = \{x_1, \ldots, x_n\}$ の場合

$$\mathrm{E}(X) = \sum_{j=1}^{n} x_j \mathrm{P}(X = x_j) = \sum_{j=1}^{n} x_j f(x_j)$$

を X の平均, あるいは期待値という.

とりうる値の集合が可算集合 $D = \{x_1, x_2, \ldots\}$ の場合には,
級数

$$\mathrm{E}(X) = \sum_{j=1}^{\infty} x_j \mathrm{P}(X = x_j) = \sum_{j=1}^{\infty} x_j f(x_j)$$

が絶対収束するとき, X の平均は存在するといい,
その極限値を X の平均, あるいは期待値という.

連続型確率変数の平均

定義 3.2

確率変数 X は連続型で，その確率密度関数を $f(x)$ とする．

$$\int_{-\infty}^{\infty} |x| f(x) dx < \infty$$

であるとき，X の平均は存在するといい，

$$\mathsf{E}(X) = \int_{-\infty}^{\infty} x f(x) dx$$

を X の平均，あるいは期待値という．

平均の例

例 3.1

(1) $X \sim B(1,p)$ のとき, $\mathsf{E}(X) = p$

(2) $X \sim B(n,p)$ のとき, $\mathsf{E}(X) = np$

例 3.2

(1) $X \sim U(a,b]$ のとき, $E(X) = \dfrac{b-a}{2}$

(2) $X \sim N(\mu,\sigma^2)$ のとき, $\mathsf{E}(X) = \mu$.

(証明は板書で)

3.2. 基本的性質

確率変数の関数の期待値

定理 3.1

確率変数 X が離散型のとき, X のとりうる値を $x_j,\ j=1,2,\ldots$ とし, 確率関数を $f(x_j)$ とする. また, X が連続型のとき, その確率密度関数を $f(x)$ とする. このとき

$$\mathsf{E}\{g(X)\} = \begin{cases} \sum_{j=1}^{\infty} g(x_j)f(x_j), & X \text{ が離散型のとき} \\ \int_{-\infty}^{\infty} g(x)f(x)dx, & X \text{ が連続型のとき} \end{cases}$$

である.

(証明 (離散型の場合) は板書で)

注 3.3

この定理は, $g(X)$ の平均は本来 $g(X)$ の分布から求められるべきであるが, X の分布から求められることを述べている.

例 3.3

X は一様分布 $U(0,1]$ に従うとする．このとき, $Y = X^2$ の平均を考える．定理 3.1 を利用すると

$$\mathsf{E}(X^2) = \int_0^1 x^2 \cdot 1 dx = \frac{1}{3}$$

となる.
一方, Y の分布関数は $0 \leq y \leq 1$ のとき

$$F_Y(y) = \mathsf{P}(Y \leq y) = \mathsf{P}(X \leq \sqrt{y}) = \sqrt{y}$$

であるから, Y の確率密度関数は

$$f_Y(y) = \begin{cases} \frac{1}{2} y^{-1/2}, & 0 < x \leq 1 \text{ のとき} \\ 0, & \text{その他のとき} \end{cases}$$

となるので

$$\mathsf{E}(Y) = \int_0^1 y \cdot \frac{1}{2} y^{-1/2} dy = \frac{1}{3}$$

例 3.4

X : 連続型, $f(x)$: 確率密度関数, $F(x)$: 分布関数
g : 微分可能な狭義の単調減少関数,
$a = \inf\{g(X(\omega));\ \omega \in \Omega\}$, $b = \sup\{g(X(\omega));\ \omega \in \Omega\}$

$a < y < b$ のとき

$$F_Y(y) = \mathsf{P}(g(X) \le y) = \mathsf{P}(X \ge g^{-1}(y)) = 1 - F(g^{-1}(y))$$
$$\frac{d}{dy}F_Y(y) = -\frac{dg^{-1}(y)}{dy} f(g^{-1}(y)) = \left|\frac{dg^{-1}(y)}{dy}\right| f(g^{-1}(y))$$

$$\mathsf{E}(Y) = \int_a^b y \left|\frac{dg^{-1}(y)}{dy}\right| f(g^{-1}(y))dy$$

$x = g^{-1}(y)$ と変数変換すると

$$\mathsf{E}(Y) = \int_{-\infty}^{\infty} g(x)f(x)dx$$

となる. ($x > g^{-1}(a)$ または $x < g^{-1}(b)$ で $f(x) = 0$)

多変量の場合

定理 3.2

n 次元確率ベクトル $\boldsymbol{X}$ が離散型のとき，そのとりうる値を $\boldsymbol{x}_1, \boldsymbol{x}_2, \ldots$ とし確率関数を $f(\boldsymbol{x}_j)$ とする．また，$\boldsymbol{X}$ が連続型のとき，その確率密度関数を $f(\boldsymbol{x})$ とする．このとき

$$\mathsf{E}(g(\boldsymbol{X})) = \begin{cases} \sum_{j=1}^{\infty} g(\boldsymbol{x}_j) f(\boldsymbol{x}_j), & \boldsymbol{x} \text{ が離散型のとき} \\ \int_{-\infty}^{\infty} \cdots \int_{-\infty}^{\infty} g(\boldsymbol{x}) f(\boldsymbol{x}) d\boldsymbol{x}, & \boldsymbol{x} \text{ が連続型のとき} \end{cases}$$

である．

(離散型の場合の証明は板書で)

定理 3.2, 連続型の例

定理 3.3

n 次元確率ベクトル $\boldsymbol{X}$ は連続型で, 確率密度関数 $f(x_1,\ldots,x_n)$ をもつとする. また, $\boldsymbol{X}$ のとりうる値は領域 $\mathcal{X}=\{\boldsymbol{x}; f(\boldsymbol{x})>0\}\subset\mathbb{R}^n$ であるとする. $\boldsymbol{x}=(x_1,\ldots,x_n)'$ から $\boldsymbol{y}=(y_1,\ldots,y_n)'$ への変換

$$\begin{cases} y_1 = u_1(\boldsymbol{x}) = u_1(x_1,\ldots,x_n) \\ \qquad\vdots \\ y_n = u_n(\boldsymbol{x}) = u_n(x_1,\ldots,x_n) \end{cases}$$

は領域 $\mathcal{X}$ を領域 $\mathcal{Y}$ に写し, 次の性質 (1) $\sim$ (3) をみたすとする.

(1) $1-1$ である.

(2) 逆変換

$$\boldsymbol{y}\to\boldsymbol{x}: \quad x_i = v_i(\boldsymbol{y}) = v_i(y_1,\ldots,y_n),\ i=1,\ldots,n$$

は C^1 クラスである.

定理 3.3 (続き)

(3) ヤコビ行列式

$$J = \begin{vmatrix} \frac{\partial x_1}{\partial y_1} & \cdots & \frac{\partial x_1}{\partial y_n} \\ \vdots & \ddots & \vdots \\ \frac{\partial x_n}{\partial y_1} & \cdots & \frac{\partial x_n}{\partial y_n} \end{vmatrix}$$

の値が任意の $\boldsymbol{y} \in \mathcal{Y}$ に対してゼロでない.

この変換により定義される確率変数を

$$Y_i = u_i(\boldsymbol{X}) = u_i(X_1, \ldots, X_n), \quad i = 1, \ldots, n$$

とする. このとき, $\boldsymbol{Y} = (Y_1, \ldots, Y_n)'$ の確率密度関数は

$$h(y_1, \ldots, y_n) = \begin{cases} f(v_1(\boldsymbol{y}), \ldots, v_n(\boldsymbol{y}))|J|, & \boldsymbol{y} \in \mathcal{Y} \\ 0, & \boldsymbol{y} \in \mathcal{Y}^c \end{cases}$$

となる.

定理 3.3 の証明

$\boldsymbol{Y} = (Y_1, \ldots, Y_n)'$ の分布関数は

$$
\begin{aligned}
F_{Y_1,\ldots,Y_n}(a_1, \ldots, a_n) &= \mathsf{P}(Y_1 \le a_1, \ldots, Y_n \le a_n) \\
&= \mathsf{P}(u_1(\boldsymbol{x}) \le a_1, \ldots, u_n(\boldsymbol{x}) \le a_n)
\end{aligned}
$$

と表せる. $D_{\boldsymbol{a}} = \{\boldsymbol{x} \in \mathbb{R}^n;\ u_i(\boldsymbol{x}) \le a_i,\ i = 1, \ldots, n\}$ とおく. 変数変換 $y_i = u_i(\boldsymbol{x})$, $i = 1, \ldots, n$ を考えることによって

$$
\begin{aligned}
F_{Y_1,\ldots,Y_n}(a_1, \ldots, a_n) &= \int \cdots \int_{D_{\boldsymbol{a}}} f(x_1, \ldots, x_n) dx_1 \cdots dx_n \\
&= \int_{-\infty}^{a_1} \cdots \int_{-\infty}^{a_n} h(y_1, \ldots, y_n) dy_1 \cdots dy_n
\end{aligned}
$$

を得る. □

定理 3.2 の証明 (連続型)

$$Y = Y_1 = u_1(X_1, \ldots, X_n) = g(X_1, \ldots, X_n)$$

とし, (1)~(3) が成り立つような関数 $u_2(\boldsymbol{x}), \ldots, u_n(\boldsymbol{x})$ がとれる場合.
$Y = Y_1$ の確率密度関数は

$$h_1(y_1) = \int_{-\infty}^{\infty} \cdots \int_{-\infty}^{\infty} h(y_1, \ldots, y_n) dy_2 \cdots dy_n$$

で与えられ, 変数変換 $y_i = u_i(\boldsymbol{x})$, $i = 1, \ldots, n$ を考えることによって

$$\begin{aligned}\mathrm{E}(Y_1) &= \int_{-\infty}^{\infty} y_1 h_1(y_1) dy_1 = \int_{-\infty}^{\infty} \cdots \int_{-\infty}^{\infty} y_1 h(y_1, \ldots, y_n) dy_1 \cdots dy_n \\ &= \int_{-\infty}^{\infty} \cdots \int_{-\infty}^{\infty} g(\boldsymbol{x}) f(\boldsymbol{x}) d\boldsymbol{x}. \qquad \square\end{aligned}$$

例

X_1, X_2 は独立で, $X_1 \sim N(2,1),\ X_2 \sim N(1,1)$

(1) $Y_1 = X_1 + X_2,\ Y_2 = X_1 - X_2$ とすると
$$\begin{pmatrix} Y_1 \\ Y_2 \end{pmatrix} \sim N_2(\boldsymbol{\mu}, \Sigma), \quad \boldsymbol{\mu} = \begin{pmatrix} 3 \\ 1 \end{pmatrix}, \Sigma = \begin{pmatrix} 2 & 0 \\ 0 & 2 \end{pmatrix}$$

(2) $\mathsf{E}(X_1 + X_2) = 3$

(証明は板書で)

平均の線形性と単調性, 独立な確率変数の積の平均

定理 3.4

確率変数 X, Y は平均をもつとする.

(1) 任意の定数 a に対して, $\mathrm{E}(aX) = a\mathrm{E}(X)$.

(2) $\mathrm{E}(X+Y) = \mathrm{E}(X) + \mathrm{E}(Y)$.

(3) $X \leq Y \Rightarrow \mathrm{E}(X) \leq \mathrm{E}(Y)$.

(証明は板書で)

定理 3.5

確率変数 X, Y は平均をもち, 独立とする. このとき

$$\mathrm{E}(XY) = \mathrm{E}(X)\mathrm{E}(Y).$$

(証明は板書で)

注 3.5

g, h を $\mathbb{R}$ 上の任意のボレル関数とする．定理 3.2 を用いると定理 3.5 と同様に X, Y が独立ならば

$$\mathsf{E}\{g(X)h(Y)\} = \mathsf{E}\{g(X)\}\mathsf{E}\{h(Y)\} \tag{1}$$

を示すことができる．逆に，任意のボレル関数 g, h に対して (1) が成り立つならば，g, h をそれぞれ，区間 $J_a = (a_1, a_2],\ J_b = (b_1, b_2]$ の定義関数ととることにより
$\mathsf{P}(X \in J_a,\ Y \in J_b) = \mathsf{E}\{\mathrm{I}_{(a_1,a_2]}(X)\mathrm{I}_{(b_1,b_2]}(Y)\} = \mathsf{P}(X \in J_a)\mathsf{P}(Y \in J_b)$
が導かれるので，X, Y は独立である．ここで，一般に集合 S の部分集合 A に対して，A の定義関数とは

$$\mathrm{I}_A(s) = \begin{cases} 1, & s \in A \\ 0, & s \notin A \end{cases}$$

と定義される S 上の関数を意味する．