# 主数入講義用スライド (2024/10/05)

講義内容

前回の記号の復習

- 1.4 主な1次元分布モデル
- 1.5 2次元分布
- 1.6 多次元分布と i.i.d. 標本

i.i.d. = identically and indepently distributed

## 記号の復習

$(\Omega, \mathcal{A}, \Pr)$: 確率空間

ただし, $\mathcal{A}$ は σ加法族, $\Pr : \mathcal{A} \longrightarrow [0, 1]$ は確率測度.
(Ωの部分集合族)

- $\mathcal{A}$ は σ加法族

$\Leftrightarrow$ (1) $\Omega \in \mathcal{A}$; (2) $A \in \mathcal{A} \Rightarrow A^c \in \mathcal{A}$

(3) $A_n \in \mathcal{A} \Rightarrow \bigcup_{n=1}^{\infty} A_n \in \mathcal{A}$

- $P_r : \mathcal{A} \ni A \mapsto P_r(A) \in [0, 1]$ は以下をみたす.

(1) $P_r(\Omega) = 1$; (2) $A \in \mathcal{A}$ に対し, $P_r(A^c) = 1 - P_r(A)$;

(3) 互いに排反な $\{A_n\}_{n=1}^{\infty} \subset \mathcal{A}$ に対し

$$P_r\left(\bigcup_{n=1}^{\infty} A_n\right) = \sum_{n=1}^{\infty} P_r(A_n).$$

$X : \Omega \to \mathbb{R}$ は確率変数

$\Longleftrightarrow$ ${}^\forall B \in \mathcal{B}(\mathbb{R})$ に対して,

$$X^{-1}(B) = \{\omega \in \Omega ; X(\omega) \in B\} \in \mathcal{A}$$

$\Longleftrightarrow$ $\forall r \in \mathbb{R}$ に対して

$$\{\omega \in \Omega ;\ X(\omega) > r\} \in \mathcal{A}$$

が成立.

ただし, $\mathcal{B}(\mathbb{R})$ は $\mathbb{R}$ の開集合族を含む最小の $\sigma$ 加法族. Borel集合族

・確率変数 $X$ の c.d.f.

$$F^X(x) = \Pr(X \leq x) \quad (x \in \mathbb{R})$$

- $F^X(\cdot)$ が階段関数のとき, $X$を離散型確率変数という
- $F^X(\cdot)$ が連続関数のとき, $X$を連続型確率変数という

$X$ の p.m.f. (Xは離散型)

$$p^X(x) := F^X(x) - F^X(x-) \quad (x \in \mathbb{R})$$

ただし, $F^X(x-) = \lim_{y \to x-0} F^X(y)$ (左極限)

$X$ の p.d.f.（$X$ は連続型）

$$F^X(x) = \int_{-\infty}^{x} p^X(\tau)\, d\tau$$ （存在するならば）

# 1.4 主な1次元モデル

## 1.4.1 離散型確率変数

**Bernolli 分布** $0 \leq \theta \leq 1$ とする.

$$P(x \mid \theta) = \begin{cases} \theta^{x}(1-\theta)^{1-x} & (x=0,1) \\ 0 & (\text{otherwise}) \end{cases}$$

すなわち, $\Pr(X=0) = 1-\theta$, $\Pr(X=1) = \theta$

$X \sim \mathrm{Ber}(\theta)$ と記す.

2項分布　$n \in \mathbb{N}$, $0 \leq \theta \leq 1$ とする.

$$p(x|\theta) = \begin{cases} \binom{n}{x} \theta^x (1-\theta)^{n-x} & (x = 0, 1, \dots n) \\ 0 & (\text{otherwise}) \end{cases}$$

ただし, $\binom{n}{x} = {}_nC_x = \frac{n!}{x!(n-x)!}$, $0! = 1$

$X \sim \mathrm{Bino}(n, \theta)$ と書く.

・**幾何分布** $0 < \theta < 1$ とする

$$p(x|\theta) = \begin{cases} \theta(1-\theta)^{x-1} & (x = 1, 2, \ldots) \\ 0 & (\text{otherwise}) \end{cases}$$

$X \sim Geo(\theta)$ と言う.

・Poisson 分布 $\theta > 0$ とする.

$$p(x|\theta) = \begin{cases} e^{-\theta} \dfrac{\theta^x}{x!} & (x = 0, 1, \ldots) \\ 0 & (\text{otherwise}) \end{cases}$$

$X \sim P_o(\theta)$ と言う.

## 1.4.2 連続型確率変数

- 正規分布　$\mu \in \mathbb{R}$, $0 < \sigma < \infty$ とする.

$$p(x \mid \mu, \sigma^2) = \frac{1}{\sqrt{2\pi\sigma^2}} \exp\left(-\frac{(x-\mu)^2}{2\sigma^2}\right)$$

$$(-\infty < x < \infty)$$

$X \sim N(\mu, \sigma^2)$ と書く.

$\exp(x) = e^x$

$\mu = 0, \sigma = 1$ のとき,

$$p(x \mid 0, 1) = \frac{1}{\sqrt{2\pi}} \exp\left(-\frac{x^2}{2}\right) \iff X \sim N(0, 1)$$

を標準正規分布という

・ガンマ分布 $\alpha > 0, \beta > 0$ とする

$$p(x|\alpha,\beta) = \begin{cases} \dfrac{1}{\beta^{\alpha}\Gamma(\alpha)} x^{\alpha-1} e^{-x/\beta} & (x > 0) \\ 0 & (\text{otherwise}) \end{cases}$$

ただし，Eulerのガンマ関数，

$$\Gamma(\alpha) = \int_0^{\infty} x^{\alpha-1} e^{-x} dx \quad (\alpha > 0)$$

$X \sim Ga(\alpha, \beta)$ と記す。

$Ga(1, \frac{1}{\lambda})$ を母数 $\lambda$ $(\lambda > 0)$ の指数分布

$p \in \mathbb{N}$ とする. $Ga\left(\frac{p}{2}, 2\right)$ を自由度 $p$ の

$\chi^2$ 分布 という.

# 1.5 2次元の分布

## 1.5.1 同時確率関数と確率密度関数

$X, Y : \Omega \to \mathbb{R}$ を確率変数とする.

**離散型の場合**

たとえば.

$$P_r(X=0, Y=0) = \frac{1}{9}; \quad P_r(X=0, Y=1) = \frac{2}{9}$$

$$P_r(X=1, Y=0) = \frac{2}{9}, \quad P_r(X=1, Y=1) = \frac{4}{9}$$

| $X \backslash Y$ | 0 | 1 | |
|---|---|---|---|
| 0 | $1/9$ | $2/9$ | $\frac{1}{3}$ |
| 1 | $2/9$ | $4/9$ | $\frac{2}{3}$ |
| | $\frac{1}{3}$ | $\frac{2}{3}$ | $1$ |

$p(x,y) = \Pr(X=x, Y=y)$ と表す、

← 同時 p.m.f.

# 連続型の場合

**定義 1.45** $(X, Y)$ の同時 p.d.f. $p(\cdot, \cdot)$

(1) $p(x, y) \geq 0 \quad ((x, y) \in \mathbb{R}^2)$

(2) $\int_{-\infty}^{\infty} \int_{-\infty}^{\infty} p(x, y)\, dx\, dy = 1$

(3) ${}^\forall A \in \mathcal{B}(\mathbb{R}^2)$ に対して

$$P_r\left((X, Y) \in A\right) = \iint_A p(x, y)\, dx\, dy$$

$(X, Y)$ の同時 c.d.f. $F(\cdot, \cdot)$

$$F(x, y) := \Pr(X \leq x, Y \leq y) \quad ((x, y) \in \mathbb{R}^2)$$

$(X, Y)$ の同時分布 $P(\cdot)$

$$P(A) := \Pr((X, Y) \in A) \quad (A \in \mathcal{B}(\mathbb{R}^2))$$

注意

$$F(x, y) = P((-\infty, x] \times (-\infty, y])$$

$F$ がわかる $\iff$ $P$ がわかる.

**例1.47** $(X, Y)$は連続型.

$$p(x, y) = \begin{cases} x + y & (0 < x < 1,\ 0 < y < 1) \\ 0 & (\text{otherwise}) \end{cases}$$

すると

$$\int_0^1 \int_0^1 p(x, y)\, dx\, dy = 1$$ となる.

# 1.5.2 周辺分布

例1

| $X \backslash Y$ | 0 | 1 | |
|---|---|---|---|
| 0 | $\frac{1}{9}$ | $\frac{2}{9}$ | $\frac{1}{3}$ |
| 1 | $\frac{2}{9}$ | $\frac{4}{9}$ | $\frac{2}{3}$ |
| | $\frac{1}{3}$ | $\frac{2}{3}$ | |

同時分布

Xの周辺分布

Yの周辺分布

## 定義 1.48

(1) 離散型

$X$ の周辺 p.m.f. $p^X$

$$p^X(x) = \Pr(X=x) = \sum_{y\in S_Y} \Pr(X=x, Y=y)$$

$$= \sum_{y\in S_Y} p(x,y) \qquad (x\in S_X)$$

ただし、$S_X := \{ x\in\mathbb{R} ;$ ある $y\in\mathbb{R}$ に対して、$p(x,y)>0 \}$

$S_Y := \{ y\in\mathbb{R} :$ ある $x$ に対して、$p(x,y)>0 \}$

(2) 連続型

$X$ の周辺 p.d.f. $p^X$

$$p^X(x) = \int_{-\infty}^{\infty} p(x, y)\, dy.$$

$(X, Y)$ の同時 c.d.f. を $F$ とおく

注意 $X$ の周辺 c.d.f. を $F^X$ とすると

$$F^X(x) = P_r(X \leq x) \quad (x \in \mathbb{R})$$

ここで

$$F^X(x) = \lim_{y \to +\infty} F(x, y)$$ となることがわかる。

## 1.5.3 (確率的)独立性と条件付き分布

### 定義1.51

確率変数 X と Y は独立

$$\Leftrightarrow \Pr(X \in A, Y \in B) = \Pr(X \in A)\Pr(Y \in B)$$

$\forall A, B \in \mathcal{B}(\mathbb{R})$

そうでないとき, X と Y は従属という

## 定理 1.52

$p(\cdot,\cdot)$ : $(X, Y)$ の同時 p.m.f. / p.d.f.

$p^X(\cdot)$ : $X$ の周辺 p.m.f. / p.d.f.

$p^Y(\cdot)$ : $Y$ の周辺 p.m.f. / p.d.f.

このとき、

$$X \text{ と } Y \text{ は独立} \iff p(x, y) = p^X(x)\, p^Y(y) \quad ({}^\forall x, y \in \mathbb{R})$$

__証明__ (⇒)は略し、(⇐)は測度論の議論が必要なので、省略。

__例 1.54__ $(X, Y)$は連続型で、$X$と$Y$は独立で、同じ p.d.f.

$$p(x) = \begin{cases} 2x & (0 < x < 1) \\ 0 & (\text{otherwise}) \end{cases}$$

を持つとする。すると$(X, Y)$の同時 p.d.f は

$$p^{(X,Y)}(x, y) = p(x)\,p(y)$$

$$= \begin{cases} 4xy & (0 < x < 1,\ 0 < y < 1) \\ 0 & (\text{otherwise}) \end{cases}$$

とする. たとえば

$$\Pr(X + Y \leq 1) = \iint_{x+y\leq 1} p^{(X,Y)}(x, y)\, dx\, dy$$

$$= 4 \int_0^1 x \left\{ \int_0^{1-x} y\, dy \right\} dx$$

$$= 4 \int_0^1 x \frac{(1-x)^2}{2}\, dx = \frac{1}{6}$$

//

## 定義 1.58 (1) 離散型

$p^Y(y) > 0$ なる $y$ に対して, $Y=y$ を与えたときの $X$ の条件付き p.m.f.

$$p^{X|Y}(x|y) = \Pr(X = x \mid Y = y)$$

$$= \frac{\Pr(X=x, Y=y)}{\Pr(Y=y)}$$

$$= \frac{p(x, y)}{p^Y(y)} \quad (x \in \mathbb{R})$$

(2) 連続型

$p^Y(y) > 0$ なる $y$ に対して, $Y=y$ を与えたときの

$X$ の条件付き p.d.f.

$$p^{X|Y}(x|y) = \frac{p(x,y)}{p^Y(y)} \quad (x \in \mathbb{R})$$

注意 1.59 連続型 $p^Y(y) > 0$ なる $y$ に対して

$Y=y$ を与えたときの事象 $\{X \in A\}$ の条件付き

確率を

$$\Pr(X \in A \mid Y = y) = \int_A p^{X|Y}(x|y)\,dx \quad (A \in \mathcal{B}(\mathbb{R}))$$

で形式的に定める.

例 1.60 $(X, Y)$ は連続型で, 同時 p.d.f

$$p(x,y) = \begin{cases} x+y & (0<x<1,\ 0<y<1) \\ 0 & (\text{otherwise}) \end{cases}$$

を持つとする. このとき

$$\Pr\left(X \leq \frac{1}{4} \,\middle|\, Y = \frac{1}{3}\right)$$ を求めよう.

$0 < y < 1$ のとき

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

以上より

$$p^Y(y) = \begin{cases} y + \frac{1}{2} & (0 < y < 1) \\ 0 & (\text{otherwise}) \end{cases}$$

よって, $0 < y < 1$ と $0 < x < 1$ において

$$p^{X|Y}(x|y) = \frac{p(x, y)}{p^Y(y)} = \frac{x + y}{y + \frac{1}{2}}$$

これより、$0 < y < 1$ のとき

$$p^{X|Y}(x|y) = \begin{cases} \dfrac{x+y}{y+\frac{1}{2}} & (0 < x < 1) \\ 0 & (\text{otherwise}) \end{cases}$$

よって

$$\Pr\left(X \leq \frac{1}{4} \,\middle|\, Y = \frac{1}{3}\right) = \int_0^{1/4} p^{X|Y}\left(x \,\middle|\, \frac{1}{3}\right) dx$$

$$= \int_0^{1/4} \frac{x + \frac{1}{3}}{\frac{1}{3} + \frac{1}{2}} dx = \frac{11}{80} \; //$$

# 1.6 多次元分布と i.i.d. 標本

$n \in \mathbb{N}$ とし, $X_1, X_2, \dots, X_n$ は確率空間 $(\Omega, \mathcal{A}, \mathrm{Pr})$ 上の確率変数列とする.

$$\underline{X} = (X_1, X_2, \dots, X_n)^T = \begin{pmatrix} X_1 \\ X_2 \\ \vdots \\ X_n \end{pmatrix}$$

とする.

## 定義 1.67

$X_1, X_2, \dots, X_n$ は互いに独立

$\Longleftrightarrow$ $\forall A_1, \dots, A_n \in \mathcal{B}(\mathbb{R})$ に対して

$$\Pr(X_1 \in A_1, X_2 \in A_2, \dots, X_n \in A_n) = \prod_{j=1}^{n} \Pr(X_j \in A_j)$$

## 注意 1.63

$p(\cdot, \dots, \cdot)$ : $\underline{X}$ の同時 p.d.f. / p.m.f.

$p^{X_j}(\cdot)$ : $X_j$ の周辺 p.d.f. / p.m.f. $(j=1, \dots, n)$

$X_1, \dots, X_n$ は互いに独立

$$\Longleftrightarrow \quad p(x_1, x_2, \dots, x_n) = \prod_{j=1}^{n} p^{X_j}(x_j)$$

$$({}^\forall x_1, \dots, x_n \in \mathbb{R})$$

**定義 1.64** $X_1, X_2, \ldots, X_n$ は互いに独立で各 $X_j$ は同じ c.d.f. $F$ $(j=1, \ldots, n)$ を持つとき、$X_1, X_2, \ldots, X_n$ は独立同一分布に従うといい、

$$X_1, X_2, \ldots, X_n \overset{\text{i.i.d.}}{\sim} F$$

と記す。

# 2.1 期待値

$(\Omega, \mathcal{A}, \Pr)$：確率空間

$X : \Omega \to \mathbb{R}$：確率変数

$p$：$X$ の p.m.f. / p.d.f.

p.m.f. $p$ に対して、

$$\{x \in \mathbb{R} ; p(x) > 0\} = \{x_1, x_2, \ldots \}$$

と書くことにする。

**定理 2.1** $g: \mathbb{R} \to \mathbb{R}$ を可測関数とする

確率変数 $g(X)$ の期待値を (1), (2) の順で

定めて行く

(1) $g(x) \geq 0 \ (x \in \mathbb{R})$ のとき

$$E[g(X)] = \begin{cases} \displaystyle\sum_{n=1}^{\infty} g(x_n)\, p(x_n) & (\text{離散型}) \\ \displaystyle\int_{-\infty}^{\infty} g(x)\, p(x)\, dx & (\text{連続型}) \end{cases}$$

と定める. $\infty$ を許せば, 右辺は必ず存在する.

(2) 一般の可測関数 $g$ に対して

$$g^+(x) = \max\{g(x), 0\}; \quad g^-(x) = \max\{-g(x), 0\}$$
$(x \in \mathbb{R})$

と定めると、 $g^+(x) \geq 0$; $g^-(x) \geq 0$ $(x \in \mathbb{R})$

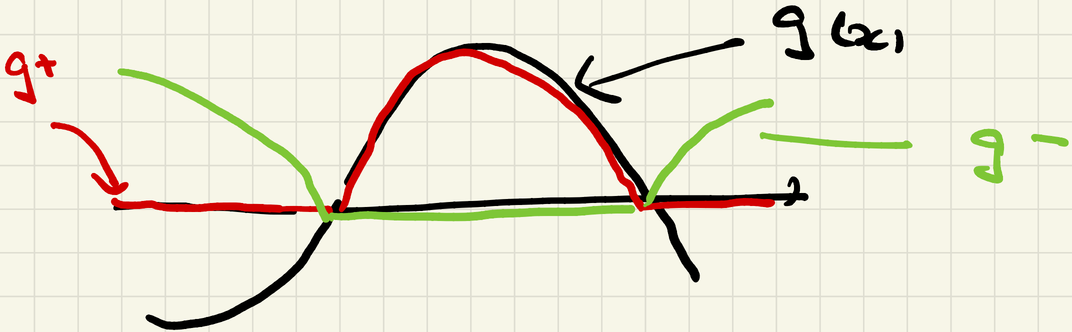
となる. $E[g^+(X)]$ または $E[g^-(X)]$ の

いずれかが有限ならば

$$E[g(X)] := E[g^+(X)] - E[g^-(X)]$$

と定める. $E[g^+(X)] = E[g^-(X)] = \infty$ のとき, $g(X)$ の期待値は定義されない.

$E[g^+(X)] < \infty$ かつ $E[g^-(X)] < \infty$ のとき, $E[g(X)]$ は有限となる




$E[|g(X)|] = E[g^+(X)] + E[g^-(X)]$ である.

**補題 2.2** $g, h : \mathbb{R} \to \mathbb{R}$ は可測関数で

$E[|g(X)|] < \infty$ かつ $E[|h(X)|] < \infty$ とする

(1) $a, b \in \mathbb{R}$ に対して

$$E[ag(X) + h(X)] = aE[g(X)] + bE[h(X)].$$

(2) $g(x) \leq h(x) \ (x \in \mathbb{R})$ ならば

$$E[g(X)] \leq E[h(X)].$$

**補題 2.3** $0 < q < r$ に対して

$$E[|X|^r] < \infty \Rightarrow E[|X|^q] < \infty$$

**証明** Young の不等式 $a > 0,\ b > 0$ に対して

$$ab \leq \frac{a^s}{s} + \frac{b^t}{t} \quad \left(\frac{1}{s} + \frac{1}{t} = 1\right)$$

を用いる.

$$qs = r \Leftrightarrow s = \frac{r}{q} < 1$$

$$|X|^q \cdot 1 \leq \frac{|X|^r}{s} + \frac{1}{t} \qquad a = |X|^q,\ b = 1$$

と補題2.2より

$$E[|X|^q] \leq \frac{q}{r} E[|X|^r] + \text{constant}$$

定義2.4 (1) $k=1,2,\ldots$ に対して $E[|X|^k]<\infty$ のとき，$E[X^k]$ を $X$ の $k$ 次モーメント(積率)という

(2) $E[|X|]<\infty$ のとき，$E[X]$ を $X$ の平均値ともいう

(3) $E[X^2]<\infty$ のとき，$X$ の分散 $Var[X]$ を

$$Var[X]=E[\{X-E[X]\}^2]$$

で定める．

(4) 部分集合 $A \subset \mathbb{R}$ に対して, $\mathbb{1}_A : \mathbb{R} \to \mathbb{R}$ を

$$\mathbb{1}_A(x) = \begin{cases} 1 & (x \in A) \\ 0 & (x \notin A) \end{cases}$$

で定める. これを $A$ の指示関数という

$A \in \mathcal{B}(\mathbb{R})$ のとき

$$E[\mathbb{1}_A(X)] = \Pr(X \in A)$$

となる.

**注意 2.6** $g: \mathbb{R} \to \mathbb{R}$ は狭義単調関数とする。

$X$: 連続型確率変数で p.d.f. $p^X$ を持つとする

$$S_X := \{x \in \mathbb{R} ; \ p^X(x) > 0\}$$

$Y = g(X)$ の p.d.f. は

$$p^Y(y) = \begin{cases} p^X(g^{-1}(y)) \dfrac{1}{|\dot{g}(g^{-1}(y))|} & (y \in S_X) \\ 0 & (\text{otherwise}) \end{cases}$$

となる。但し

$$S_Y = g(S_X) \;; \quad \dot{g}(x) = \frac{dg}{dx}(x).$$

したがって

$$E[Y] = \int_{-\infty}^{\infty} y \, p^Y(y) \, dx$$

$$= \int_{-\infty}^{\infty} y \, p^X(g^{-1}(y)) \frac{1}{|\dot{g}(g^{-1}(y))|} dy$$

$$= \int_{-\infty}^{\infty} g(x) \, p^X(x) \, dx \quad \left( \begin{array}{l} x = g^{-1}(y) \\ \Leftrightarrow y = g(x) \\ dy = \dot{g}(x) \, dx \end{array} \right)$$

## 2.2 確率ベクトルの期待値

$(X, Y)$: 確率空間 $(\Omega, \mathcal{A}, P_r)$ 上の確率ベクトル

$p(\cdot, \cdot)$: $(X, Y)$ の p.d.f. / p.m.f.

$g : \mathbb{R}^2 \longrightarrow \mathbb{R}$ : 可測関数

**定義 2.7** $g(X, Y)$ の期待値を (1), (2) の順で定める.

(1) $g \geq 0$ に対して

$$E[g(X,Y)] = \begin{cases} \displaystyle\sum_{x,y} g(x,y)\,p(x,y) & (\text{離散型}) \\ \displaystyle\int_{-\infty}^{\infty}\int_{-\infty}^{\infty} g(x,y)\,p(x,y)\,dx\,dy & (\text{連続型}) \end{cases}$$

で定める

(2) 一般の $g$ に対して

$$g^{+}(x,y) = \max\{g(x,y),\ 0\}$$

$$g^{-}(x,y) = \max\{-g(x,y),\ 0\}$$

とおく. $E[g^+(X,Y)]$ または $E[g^-(X,Y)]$ のいずれかが有限のとき

$$E[g(X,Y)] = E[g^+(X,Y)] - E[g^-(X,Y)]$$

と定める.

**定理 2.8** $X_1, X_2, \dots, X_n$ は確率変数で

$E[|X_j|] < \infty$ $(j = 1, 2, \dots, n)$ とする。

定数 $a_1, a_2, \dots, a_n$ に対して

$$E\left[\sum_{j=1}^{n} a_j X_j\right] = \sum_{j=1}^{n} a_j E[X_j]$$

注意 $n = \infty$ のときは、成立するとはかぎらない

$$E\left[\sum_{j=1}^{\infty} X_j\right] = \sum_{j=1}^{\infty} E[X_j]$$

?

例 2.10　$0 < p < 1$, $j = 1, 2, \ldots, n$ に対して

$$X_j \sim \mathrm{Ber}(p)$$

とする。$S = \sum_{j=1}^{n} X_j$ としたとき

$$E[S] = \sum_{j=1}^{n} \underbrace{E[X_j]}_{=p} = n \times p$$

注意　独立性は必要ない。一般に、$S$ は 2項分布ではない。独立性があれば $S \sim \mathrm{Bino}(n, p)$.

## 2.3 分散と共分散

$E[X^2] < \infty$ を仮定する.

$X$ の分散 $Var[X]$ は

$$Var[X] = E[\{X-\mu\}^2]; \quad \mu = E[X]$$

で定義した. さらに, $\sqrt{Var[X]}$ を $X$ の標準偏差という.

## 定理 2.11

(1) (分散公式) $Var[X] = E[X^2] - \{E[X]\}^2$

よって. $E[X^2] \geq \{E[X]\}^2$

Xの2乗の期待値 ≧ Xの期待値 の2乗

(2) 定数 $a, b \in \mathbb{R}$ に対して

$$Var[aX+b] = a^2 Var[X]$$

注意 $a=0$ のときも「=」は成立.

(3) $X$と$Y$は独立で $E[|XY|]<\infty$ とする。このとき

$$E[XY]=E[X]E[Y]$$

(4) $X_1, X_2, \ldots, X_n$は互いに独立とする。実数 $a_1, \ldots, a_n \in \mathbb{R}$ に対して

$$\mathrm{Var}\left[\sum_{j=1}^{n} a_j X_j\right]=\sum_{j=1}^{n} a_j^2 \,\mathrm{Var}[X_j]$$

例 2.12 $X_1, X_2, \ldots, X_n$ は互いに独立で
$X_j \sim Ber(p)$ $(0 < p < 1,\ j = 1, \ldots, n)$ とする。

$$
\begin{aligned}
Var[X_j] &= E[X_j^2] - \{E[X_j]\}^2 \\
&= 0^2 \times (1-p) + 1^2 \times p - \{0 \times (1-p) + 1 \times p\}^2 \\
&= p(1-p).
\end{aligned}
$$

定理 2.11 (4) から

$$\mathrm{Var}\left[\sum_{j=1}^{n} X_j\right] = \sum_{j=1}^{n} \mathrm{Var}[X_j] = np(1-p).$$

定理2.13 $\underbrace{X_1, X_2, \dots, X_n}_{n \geq 2 \text{とする.}}$ を i.i.d. 確率変数で

$$E[X_1] = \mu ; \quad \mathrm{Var}[X_1] = \sigma^2 \ (0 < \sigma < \infty)$$

とする.

$$\overline{X}_n = \frac{1}{n}\sum_{j=1}^{n} X_j ; \quad S_n^2 = \frac{1}{n-1}\sum_{j=1}^{n}\{X_j - \overline{X}_n\}^2$$

と定める. このとき.

(1) $E[\overline{X}_n] = \mu$; (2) $Var[\overline{X}_n] = \frac{\sigma^2}{n}$

(3) $E[S_n^2] = \sigma^2$.

(3)の証明

$$\sum_{j=1}^{n} (X_j - \overline{X}_n)^2 = \sum_{j=1}^{n} \{ (X_j - \mu) - (\overline{X}_n - \mu) \}^2$$

$$= \sum_{j=1}^{n} (X_j - \mu)^2 - n(\overline{X}_n - \mu)^2$$

定理2.11(1)より

$$E\left[\sum_{j=1}^{n}(X_j-\mu)^2\right]=\sum_{j=1}^{n}E\left[(X_j-\mu)^2\right]=n\sigma^2.$$

さらに，(1)，(2) より

$$E\left[(\overline{X}_n-\mu)^2\right]=E\left[\{\overline{X}_n-E[\overline{X}_n]\}^2\right]$$

$$=Var[\overline{X}_n]=\frac{\sigma^2}{n}$$

これよって

$$E[S_n^2]=\frac{1}{n-1}\left\{\sum_{j=1}^{n}E\left[(X_j-\mu)^2\right]-nE\left[\{\overline{X}_n-\mu\}^2\right]\right\}$$

$$= \frac{1}{n-1}\left\{ n\sigma^2 - n\cdot\frac{\sigma^2}{n} \right\} = \sigma^2.$$

**定義2.14** $X, Y$を確率変数とする

$E[X] = \mu_X$, $Var[X] = \sigma_X^2$

$E[Y] = \mu_Y$, $Var[Y] = \sigma_Y^2$

$0 < \sigma_X < \infty$, $0 < \sigma_Y < \infty$

とする.

$X$ と $Y$ の共分散を

$$\mathrm{Cov}[X, Y] = E[(X - \mu_X)(Y - \mu_Y)]$$

で定義し，$X$ と $Y$ の (Pearson) の相関係数を

$$\rho[X, Y] = \frac{\mathrm{Cov}[X, Y]}{\sigma_X \sigma_Y}$$

で定める.

注意　$E[|XY|] < \infty$ とする

## 定理 2.15

(1) (共分散公式)

$$\mathrm{Cov}[X, Y] = E[XY] - E[X]E[Y]$$

(2) $\mathrm{Cov}[X, Y] = \mathrm{Cov}[Y, X]$

(3) 定数 $a$, $b$ に対して

$$\mathrm{Cov}[aX + bY, Z]$$

$$= a\,\mathrm{Cov}[X, Z] + b\,\mathrm{Cov}[Y, Z]$$

(4) $-1 \leq \rho[X, Y] \leq 1$

(5) ある定数 $a, b$ が存在して，$Y = aX + b$ $(a \neq 0)$ と書けるとき

$$a > 0 \Rightarrow \rho[X, Y] = 1$$

$$a < 0 \Rightarrow \rho[X, Y] = -1$$

(6) $X$ と $Y$ が独立なとき

$$\mathrm{Cov}[X, Y] = 0.$$

注意 2.16　$E[X]=E[Y]=0$ とする.

Cauchy-Schwarz の不等式

$$|E[XY]| \leq \sqrt{E[X^2]}\sqrt{E[Y^2]}$$

注意 2.17

$X$ と $Y$ は独立 $\Rightarrow$ $Cov[X,Y]=0$

の逆は一般に真ではない

反例は 問 2.2

**定理 2.18** $E[X^2]<\infty$, $E[Y^2]<\infty$ のとき

$$\mathrm{Var}[X+Y] = \mathrm{Var}[X] + \mathrm{Var}[Y] + 2\,\mathrm{Cov}[X, Y].$$

**定義2.19** $d \geq 2$, $X_1, X_2, \ldots, X_d$ を有限な2次の積率を持つ確率変数列として,

$$\underline{X} = (X_1, X_2, \ldots, X_d)^T = \begin{pmatrix} X_1 \\ X_2 \\ \vdots \\ X_d \end{pmatrix}$$

と書く. このとき, 確率ベクトル $\underline{X}$ の期待値 $E[\underline{X}]$ を

$$E[\underline{X}] = (E[X_1], E[X_2], \ldots, E[X_d])^T =: \underline{\mu}$$

と約束する.

確率ベクトル $\underline{X}$ の共分散を

$$\mathrm{Var}[\underline{X}] = E\left[(\underline{X} - \underline{\mu})(\underline{X} - \underline{\mu})^T\right]$$

$$= \begin{pmatrix} \mathrm{Var}[X_1] & \mathrm{Cov}[X_1, X_2] & \cdots & \mathrm{Cov}[X_1, X_d] \\ \mathrm{Cov}[X_2, X_1] & \mathrm{Var}[X_2] & & \\ \vdots & & \ddots & \\ \mathrm{Cov}[X_d, X_1] & \cdots & \cdots & \mathrm{Var}[X_d] \end{pmatrix}$$

と約束する。

**注意 2.20** $\forall \underline{a} \in \mathbb{R}^d$ に対して

$$\underline{a}^T \mathrm{Var}[\underline{X}]\, \underline{a}$$

$$= E\left[\underbrace{\underline{a}^T (\underline{X} - \underline{\mu})}_{\text{スカラー}} \underbrace{(\underline{X} - \underline{\mu})^T \underline{a}}_{\text{スカラー}}\right]$$

$$= E\left[\underline{a}^T (\underline{X} - \underline{\mu}) \left\{ (\underline{X} - \underline{\mu})^T \underline{a} \right\}^T \right]$$

$$= E\left[\underline{a}^T (\underline{X} - \underline{\mu})\, \underline{a}^T (\underline{X} - \underline{\mu})\right]$$

$$= E\left[\underbrace{\{\underline{a}^T(\underline{X}-\underline{\mu})\}^2}_{\geq 0}\right] \geq 0$$

となるので、$Var[\underline{X}]$は半正定値。

補題 2.21 $\underline{X} = (X_1, \ldots, X_d)^T$の各成分は有限な2次の積率を持つ。

$$E[\underline{X}] = \underline{\mu}, \quad Var[\underline{X}] = \Sigma$$

とする。ただし $\underline{\mu} \in \mathbb{R}^d$、$\Sigma$は$d \times d$の半正定値行列

(1) 定数ベクトル $\underline{a} \in \mathbb{R}^d$ に対して

$$E[\underline{a}^T \underline{X}] = \underline{a}^T \underline{\mu}$$

$$Var[\underline{a}^T \underline{X}] = \underline{a}^T \Sigma \underline{a}$$

(2) $k \in \mathbb{N}$ とし, $A$ を $k \times d$ の定数行列

とする. このとき

$$E[A\underline{X}] = A\underline{\mu},$$

$$Var[A\underline{X}] = A \Sigma A^T$$

系 2.22　定数ベクトル $\underline{a} \neq 0$ ($\underline{a} \in \mathbb{R}^d$) に対して

$$\Pr(\underline{a}^T \underline{X} = 0) = 0$$

とする．このとき，$\Sigma$ は正定値

$$\underline{a}^T \Sigma \underline{a} > 0$$

# 2.4 条件付き期待値

**定義2.24** $g: \mathbb{R}^2 \to \mathbb{R}$ は可測とし，$E[g(X,Y)] < \infty$

とする．$p^{X|Y}(\cdot \mid y)$ を $Y=y$ を与えたときの $X$ の条件付き p.m.f / p.d.f. とする．

$Y=y$ を与えたときの $g(X,Y)$ の条件付き期待値 $E[g(X,Y) \mid Y=y]$ を

$$E[g(X,Y) \mid Y = y]$$

$$= \begin{cases} \displaystyle\sum_x g(x,y)\, p^{X|Y}(x|y) & (\text{離散型}) \\ \displaystyle\int_{-\infty}^{\infty} g(x,y)\, p^{X|Y}(x|y)\, dx & (\text{連続型}) \end{cases}$$

で定める.

## 注意2.25

$$h(y) := E[X \mid Y = y]$$

とおいたとき

$$E[X \mid Y] := h(Y)$$

で定める。すると，$y = Y(\omega)$ とおくと

$$E[X \mid Y] : \Omega \ni \omega \mapsto E[X \mid Y(\omega)] = E[X \mid Y = y]$$

は可測となることが知られている。

例12.26 連続型確率変数 $Y$ の p.d.f. を

$$p^Y(y) = \begin{cases} 1 & (0 < y < 1) \\ 0 & (\text{otherwise}) \end{cases}$$

とし，$Y = y\ (0<y<1)$ を与えたときの $X$ の条件付き

分布は

$$X \mid Y = y \sim \mathrm{Unif}(y, 1)$$

$$\Leftrightarrow\ p^{X|Y}(x|y) = \begin{cases} \dfrac{1}{1-y} & (y < x < 1) \\ 0 & (\text{otherwise}) \end{cases}$$

[illegible], $0 < y < 1$ [illegible]

$$E[X \mid Y = y] = \int_y^1 x \, p^{X|Y}(x \mid y) \, dx$$

$$= \frac{1}{1-y} \int_y^1 x \, dx = \frac{1+y}{2}$$

[illegible]

$$E[X \mid Y] = \frac{1+Y}{2}$$

定理2.2) $E[|X|]<\infty$, $E[|Y|]<\infty$ とする)

(1) $E[X+Y|Z] = E[X|Z] + E[Y|Z]$

(2) $E[E[Y|X]] = E[Y]$

$E[E[X|Y]] = E[X]$

(3) $g: \mathbb{R}^2 \to \mathbb{R}$ が $E[|g(X,Y)|]<\infty$

のとき

$$E[E[g(X,Y)|Y]] = E[g(X,Y)].$$

(4) $E[XY \mid Y] = Y E[X \mid Y]$.

例 2.28 (例 2.26 の続き)

$$E[X] = E[E[X \mid Y]] = E\left[\frac{1+Y}{2}\right] = \frac{3}{4}.$$

一方, $0 < y < x < 1$ において

$$p^{(X,Y)}(x,y) = p^{X \mid Y}(x \mid y)\, p^{Y}(y)$$

$$= \frac{1}{1-y}.$$

よって

$$E[X] = \iint_{\mathbb{R}^2} x\, p^{(X,Y)}(x,y)\, dx\, dy$$

$$= \int_0^1 \left\{ \int_y^1 x \frac{1}{1-y}\, dx \right\} dy = \frac{3}{4}$$

## 2.5 積率母関数

**定義 2.31** $X$ を確率変数とし，ある $t_0 > 0$ が存在して，$E[e^{tX}] < \infty$ ($\forall |t| < t_0$) とする．このとき，$X$ の積率母関数 (m.g.f.) を

$$m^X(t) := E[e^{tX}] \quad (-t_0 < t < t_0)$$

と定義する．

注意2.32　X の m.g.f. が存在するとき

$$\dot{m}^X(0) = \frac{d}{dt} m^X(t) \Big|_{t=0} = E[X]$$

$$\ddot{m}^X(0) = \frac{d^2}{dt^2} m^X(t) \Big|_{t=0} = E[X^2]$$

となる.

例 2.33　$X \sim Exp(1)$ とする. $X$ の p.d.f.

は

$$p(x) = \begin{cases} e^{-x} & (x > 0) \\ 0 & (\text{otherwise}) \end{cases}$$

である. $t < 1$ に対して

$$m^X(t) = E[e^{tX}] = \int_0^\infty e^{tx} e^{-x} dx = \frac{1}{1-t}.$$

よって

$$\dot{m}^X(t) = \frac{1}{(1-t)^2} \; ; \; \ddot{m}^X(t) = \frac{2}{(1-t)^3}$$

$\tau_1$ の ?

$$E[X] = \dot{m}^{X}(0) = 1$$

$$E[X^2] = \ddot{m}^{X}(0) = 2$$

$$Var[X] = E[X^2] - \{E[X]\}^2 = 1$$

定理2.35 $t_0>0$ が存在して

$$m^X(t) = m^Y(t) \quad ({}^\forall |t| < t_0)$$

ならば

$$X \overset{d}{=} Y \iff F^X(x) = F^Y(x) \quad ({}^\forall x \in \mathbb{R})$$

$$\iff P_r(X \in B) = P_r(Y \in B) \quad ({}^\forall B \in \mathcal{B}(\mathbb{R}))$$