Ch.4. 多次元確率変数の分布.

§4.1 同時確率分布と周辺分布

4.1.1 離散分布の場合.

- X : r.v. on $\mathcal{X} = \mathbb{Z}_{\geq 0}$,

  Y : r.v. on $\mathcal{Y} = \mathbb{Z}_{\geq 0}$.

  $$P(X=x, Y=y) =: f_{X,Y}(x,y), \quad (x,y) \in \mathcal{X}\times\mathcal{Y}.$$

- $C \subseteq \mathcal{X}\times\mathcal{Y}$,

  $$P(\underbrace{(X,Y)}_{\text{同時分布}} \in C) = \sum_{(x,y)\in C} \underbrace{f_{X,Y}(x,y)}_{\text{同時確率関数}}.$$

- $A \subseteq \mathcal{X}$ に対し, Xの周辺分布は

  $$P(X\in A) = P((X,Y) \in A\times\mathcal{Y})$$
  $$= \sum_{x\in A} \underbrace{\sum_{y\in\mathcal{Y}} f_{X,Y}(x,y)}_{=: f_X(x)}$$

  Xの周辺確率関数

- $$\mathbb{E}_{X,Y}[g(X,Y)] := \sum_{x\in\mathcal{X}} \sum_{y\in\mathcal{Y}} g(x,y) f_{X,Y}(x,y).$$

4.1.2 連続分布の場合

<u>Def.</u> (同時確率, 周辺確率)

$X, Y : \mathbb{R}$-r.v., $\quad C \subseteq \mathbb{R}^2$

$$P((X,Y) \in C) = \iint_{(x,y)\in C} f_{X,Y}(x,y)\,dx\,dy.$$

となるとき, $f_{X,Y}$ を 同時確率密度函数 という.

$$f_X(x) := \int_{-\infty}^{\infty} f_{X,Y}(x,y)\,dy, \quad f_Y(x) := \int_{-\infty}^{\infty} f_{X,Y}(x,y)\,dx$$

を, それぞれ X, Y の 周辺確率密度函数 という. $\square$

- 2次元の分布函数は,

$$F_{X,Y}(x,y) = P(X \leq x, Y \leq y) = \int_{-\infty}^{x}\int_{-\infty}^{y} f_{X,Y}(s,t)\,dt\,ds$$

で, $f_{X,Y}(x,y) = \dfrac{\partial^2}{\partial x \partial y} F_{X,Y}(x,y)$.

- $g(X,Y)$ の期待値.

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

§4.2 条件付き確率分布と独立性.

4.2.1 条件付き確率分布と条件付き期待値.

- X, Y : discrete r.v. のとき.

<u>Def.</u> (条件付き確率函数)

$f_X(x) \neq 0$ なる x に対し, $X=x$ を与えたときの $Y=y$ の

条件付き確率函数 を

$$f_{Y|X}(y|x) := P(Y=y \mid X=x) = \frac{f_{X,Y}(x,y)}{f_X(x)}$$

で定める. □

- $f_{Y|X}$ は確率分布.

<u>Def.</u> (条件付き期待値, 条件付き分散)

$X=x$ を与えたときの Y の条件付き期待値 は

$$\mathbb{E}_{Y|X}[Y \mid X=x] := \sum_{y=0}^{\infty} y f_{Y|X}(y|x) = \frac{\sum_{y=0}^{\infty} y f_{X,Y}(x,y)}{f_X(x)}$$

で定められる. また, Y の条件付き分散 は

$$\begin{aligned}\mathbb{V}_{Y|X}[Y \mid X=x] &:= \mathbb{E}_{Y|X}\left[\left(Y - \mathbb{E}_{Y|X}[Y \mid X=x]\right)^2 \mid X=x\right] \\ &= \mathbb{E}_{Y|X}[Y^2 \mid X=x] - \left(\mathbb{E}_{Y|X}[Y \mid X=x]\right)^2\end{aligned}$$

で定められる. □

Prop. $\mathbb{E}_{X,Y}[g(X,Y)] = \mathbb{E}_X[\mathbb{E}_{Y|X}[g(X,Y)|X]]$

$= \mathbb{E}_Y[\mathbb{E}_{X|Y}[g(X,Y)|Y]]$ □

pf. $\mathbb{E}_X[\mathbb{E}_{Y|X}[g(X,Y)|X]]$

$$= \sum_{x\in\mathcal{X}} \mathbb{E}_{Y|X}[g(X,Y)|X=x] f_X(x)$$

$$= \sum_{x\in\mathcal{X}} \frac{\sum_{y\in\mathcal{Y}} g(x,y) f_{X,Y}(x,y)}{f_X(x)} f_X(x)$$

$$= \sum_{x\in\mathcal{X}} \sum_{y\in\mathcal{Y}} g(x,y) f_{X,Y}(x,y)$$

$$= \mathbb{E}_{X,Y}[g(X,Y)].$$ ▨

∘ X,Y : $\mathbb{R}$-r.v. のとき.

Def. (条件付き確率関数)

$f_X(x) > 0$ なる x に対し, $X=x$ を与えたときの $Y=y$ の

条件付き確率密度関数を

$$f_{Y|X}(y|x) := \frac{f_{X,Y}(x,y)}{f_X(x)}$$

で定める. □

・ $f_{Y|X}$ は確率分布.

Def. (条件付き期待値, 条件付き分散)

$X=x$ を与えたときの $g(x,Y)$ の条件付き期待値は

$$\mathbb{E}_{Y|X}[g(x,Y)|X=x] := \int_{-\infty}^{\infty} g(x,y) f_{Y|X}(x,y)\,dy = \frac{1}{f_X(x)} \int_{-\infty}^{\infty} g(x,y) f_{X,Y}(x,y)\,dy$$

で定められる. また, Y の条件付き分散は

$$\mathbb{V}_{Y|X}[Y|X=x] := \mathbb{E}_{Y|X}\left[(Y-\mathbb{E}_{Y|X}[Y|X=x])^2 \,\middle|\, X=x\right] = \mathbb{E}_{Y|X}[Y^2|X=x] - (\mathbb{E}_{Y|X}[Y|X=x])^2$$

で定められる. □

4.2.2 確率変数の独立性

Def. (独立)

X, Y が独立 $\overset{\text{def.}}{\iff} {}^\forall x \in \mathcal{X}, {}^\forall y \in \mathcal{Y}.\ f_{X,Y}(x,y) = f_X(x) f_Y(y).$

□

Prop. $X \perp\!\!\!\perp Y \Rightarrow \mathbb{E}_{X,Y}[g(X)h(Y)] = \mathbb{E}_X[g(X)]\mathbb{E}_Y[h(Y)].$ □

4.2.3 共分散と相関係数.

・ $X \not\perp\!\!\!\perp Y$ のとき, X と Y の関係をとらえるのに用いる.

<u>Def.</u> (共分散, 相関係数)

$\mu_X := \mathbb{E}_X[X]$, $\mu_Y := \mathbb{E}_Y[Y]$, $\sigma_X^2 := \mathbb{V}_X[X]$, $\sigma_Y^2 := \mathbb{V}_Y[Y]$.

X, Y の共分散を

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

と定める. X, Y の相関係数を

$$\mathrm{Corr}(X,Y) := \frac{\mathrm{Cov}(X,Y)}{\sigma_X \sigma_Y}$$

と定める. □

<u>Prop.</u> (1) $a, b, c, d \in \mathbb{R}$,

スケールに依存

$$\mathrm{Cov}(aX+b, cY+d) = ac\,\mathrm{Cov}(X,Y).$$

スケールに依存しない

$$\mathrm{Corr}(aX+b, cY+d) = \frac{ac}{|ac|}\mathrm{Corr}(X,Y).$$

(2) $\mathrm{Cov}(X,Y) = \mathbb{E}_{X,Y}[XY] - \mathbb{E}_X[X]\mathbb{E}_Y[Y]$.

(3) $|\mathrm{Corr}(X,Y)| \leq 1$.

等号成立 $\Leftrightarrow \exists a, b \in \mathbb{R}, a \neq 0, P(Y = aX+b) = 1$.

(4) $X \perp\!\!\!\perp Y \Rightarrow \mathrm{Corr}(X,Y) = 0$. □

Prop. $a, b \in \mathbb{R}$.

$$\mathbb{E}_{X,Y}[aX+bY] = a\mathbb{E}_X[X] + b\mathbb{E}_Y[Y].$$

$$\mathbb{V}_{X,Y}[aX+bY] = a^2\mathbb{V}_X[X] + 2ab\,\mathrm{Cov}(X,Y) + b^2\mathbb{V}_Y[Y]. \quad \square$$

4.2.4 階層モデルと混合分布.

・ $f_{X,Y}(x,y) = f_{X|Y}(x|y) f_Y(y)$.

$$\rightarrow \quad X|Y=y \sim f_{X|Y}(x|y)$$

$$Y \sim f_Y(y)$$

階層モデル

$$f_X(x) = \int f_{X|Y}(x|y) \underbrace{f_Y(y)\,d\mu_Y(y)}_{\text{重み}}$$

混合分布

ex) Y : discrete r.v. $f_Y(i) = P(Y=i) = p_i \quad (i=1,\cdots,k)$

$$f_i(x) = P(X=x \mid Y=i).$$

$$f_X(x) = \sum_{i=1}^{k} p_i f_i(x). \quad \leftarrow \text{混合分布}.$$

Prop. (1) $\mathbb{E}[X] = \mathbb{E}_Y[\mathbb{E}_{X|Y}[X|Y]]$.

(2) $\mathbb{V}[X] = \mathbb{E}_Y[\mathbb{V}_{X|Y}[X|Y]] + \mathbb{V}_Y[\mathbb{E}_{X|Y}[X|Y]]$. $\square$

ex) ベータ・2項分布.

$$X|Y \sim B(n, Y), \quad Y \sim Beta(\alpha, \beta).$$

$\Rightarrow$ X は ベータ・2項分布 に従い,

$$f_X(x) = \binom{n}{x} \frac{B(x+\alpha, n-x+\beta)}{B(\alpha, \beta)} \quad (x = 0, \cdots, n).$$

ex) ガンマ・Poisson分布.

$$X|Y \sim Po(Y), \quad Y \sim Ga(\alpha, \beta)$$

$\Rightarrow$ X は ガンマ・Poisson分布 に従い,

$$f_X(x) = \frac{\Gamma(x+\alpha)}{\Gamma(\alpha) x!} \frac{\beta^x}{(1+\beta)^{x+\alpha}}.$$

ex) 非心カイ2乗分布

$$V|J \sim \chi^2_{n+2J}, \quad J \sim Po(\lambda)$$

$\Rightarrow$ V は 自由度 n, 非心度 λ の 非心カイ2乗分布 $\chi^2_n(\lambda)$ に従い.

$$f_V(v) = \sum_{j=0}^{\infty} \frac{\lambda^j}{j!} e^{-\lambda} f_{n+2j}(v).$$

($f_{n+2j}(v)$: χ^2_{n+2j} の pdf).

§4.3 変数変換.

4.3.1. 変数変換の公式.

- X, Y : r.v. $\quad f_{X,Y}(x,y)$: joint pdf.

$$\begin{cases} S = g_1(X,Y) \\ T = g_2(X,Y) \end{cases}$$

$\to f_{S,T}(s,t)$?

$D \subseteq \mathbb{R}^2$, $C_D := \{(x,y) \mid (g_1(x,y), g_2(x,y)) \in D\}$ とすると.

$$P((S,T) \in D) = P((x,y) \in C_D).$$

- $$\begin{cases} X = h_1(S,T) \\ Y = h_2(S,T) \end{cases}$$

とかけるとき, Jacobian は

$$J(s,t) := J((s,t) \to (x,y)) = \det \begin{pmatrix} \frac{\partial}{\partial s} h_1(s,t) & \frac{\partial}{\partial t} h_1(s,t) \\ \frac{\partial}{\partial s} h_2(s,t) & \frac{\partial}{\partial t} h_2(s,t) \end{pmatrix}.$$

$$\therefore \iint_{(x,y) \in C_D} f_{X,Y}(x,y)\,dx\,dy$$

$$= \iint_{(s,t) \in D} f_{X,Y}(h_1(s,t), h_2(s,t))\,|J(s,t)|\,ds\,dt.$$

$$\therefore f_{S,T}(s,t) = f_{X,Y}(h_1(s,t), h_2(s,t))\,|J(s,t)|.$$

- $J((x,y)\to(s,t)) = \dfrac{1}{J((s,t)\to(x,y))}$.

ex) Box-Muller変換.

$U_1, U_2 \overset{\text{i.i.d.}}{\sim} U(0,1)$.

$$\begin{cases} r = \sqrt{-2\log U_1} \\ \theta = 2\pi U_2. \end{cases} \qquad \begin{cases} X = r\cos\theta \\ Y = r\sin\theta. \end{cases}$$

[Claim] $X, Y \overset{\text{i.i.d.}}{\sim} N(0,1)$.

pf. $$\begin{cases} r = \sqrt{X^2+Y^2} \\ \theta = \arctan\dfrac{Y}{X} \end{cases} \qquad \begin{cases} U_1 = \exp\left(-\frac{1}{2}(X^2+Y^2)\right) \\ U_2 = \dfrac{1}{2\pi}\arctan\dfrac{Y}{X} \end{cases}$$

$$f_{X,Y}(x,y) = f_{U_1,U_2}\left(\exp\left(-\tfrac{1}{2}(x^2+y^2)\right), \tfrac{1}{2\pi}\arctan\tfrac{y}{x}\right)|J(x,y)|$$

$$f_{U_1,U_2}(u_1,u_2) = f_{U_1}(u_1)f_2(u_2) = 1\cdot 1 = 1.$$

$$J(x,y) = \det\begin{pmatrix} \dfrac{\partial}{\partial x}\exp\left(-\frac{1}{2}(x^2+y^2)\right) & \dfrac{\partial}{\partial y}\exp\left(-\frac{1}{2}(x^2+y^2)\right) \\ \dfrac{\partial}{\partial x}\dfrac{1}{2\pi}\arctan\dfrac{y}{x} & \dfrac{\partial}{\partial y}\dfrac{1}{2\pi}\arctan\dfrac{y}{x} \end{pmatrix}$$

$$= \det\begin{pmatrix} -x\exp\left(-\frac{1}{2}(x^2+y^2)\right) & -y\exp\left(-\frac{1}{2}(x^2+y^2)\right) \\ -\dfrac{1}{2\pi}\dfrac{\frac{y}{x^2}}{1+(y/x)^2} & \dfrac{1}{2\pi}\dfrac{\frac{1}{x}}{1+(y/x)^2} \end{pmatrix}$$

$$= -\frac{1}{2\pi}\exp\left(-\frac{1}{2}(x^2+y^2)\right)$$

$$\therefore f_{X,Y}(x,y) = \frac{1}{2\pi} \exp\left(-\frac{1}{2}(x^2+y^2)\right)$$

$$f_X(x) = \int_{-\infty}^{\infty} \frac{1}{2\pi} \exp\left(-\frac{1}{2}(x^2+y^2)\right) dy$$

$$= \frac{1}{2\pi} \exp\left(-\frac{1}{2}x^2\right) \int_{-\infty}^{\infty} e^{-\frac{1}{2}y^2} dy$$

$$= \frac{1}{2\pi} \exp\left(-\frac{1}{2}x^2\right) \sqrt{2\pi}$$

$$= \frac{1}{\sqrt{2\pi}} \exp\left(-\frac{1}{2}x^2\right). \quad ▨$$

4.3.2 確率変数の和の分布.

- $X \sim f_X(x)$, $Y \sim f_Y(y)$, $X \perp\!\!\!\perp Y$.

→ $Z = X + Y$ の分布?

- $$\begin{cases} Z = X + Y \\ T = Y \end{cases} \Rightarrow \begin{cases} X = Z - T \\ Y = T \end{cases}$$

$$J(x,y) = \det\begin{pmatrix} 1 & -1 \\ 0 & 1 \end{pmatrix} = 1.$$

$$\therefore f_{Z,T}(z,t) = f_{X,Y}(z-t, t)$$

$$= f_X(z-t) f_Y(t).$$

畳み込み

$$\therefore f_Z(z) = \int_{-\infty}^{\infty} f_X(z-t) f_Y(t)\, dt =: f_X * f_Y(z).$$

$$\cdot\ \phi_Z(t) = \mathbb{E}_{X,Y}[e^{it(X+Y)}] = \mathbb{E}_X[e^{itX}]\mathbb{E}_Y[e^{itY}]$$
$$= \phi_X(t)\phi_Y(t).$$

<u>Prop.</u>（分布の再生性）

(1) $X \sim \mathcal{N}(\mu_X, \sigma_X^2),\ Y \sim \mathcal{N}(\mu_Y, \sigma_Y^2)$

$\Rightarrow X+Y \sim \mathcal{N}(\mu_X+\mu_Y, \sigma_X^2+\sigma_Y^2).$

(2) $X \sim B(m,p),\ Y \sim B(n,p)$

$\Rightarrow X+Y \sim B(m+n,p).$

(3) $X \sim Po(\lambda_1),\ Y \sim Po(\lambda_2)$

$\Rightarrow X+Y \sim Po(\lambda_1+\lambda_2).$

(4) $X \sim Ga(\alpha_1,\beta),\ Y \sim Ga(\alpha_2,\beta)$

$\Rightarrow X+Y \sim Ga(\alpha_1+\alpha_2,\beta)$

(5) $X \sim \chi_m^2,\ Y \sim \chi_n^2$

$\Rightarrow X+Y \sim \chi_{m+n}^2.$ □

<u>Prop.</u> $Z_i \overset{i.i.d.}{\sim} \mathcal{N}(0,1) \Rightarrow \sum_{i=1}^{k} Z_i^2 \sim \chi_k^2.$ □

§4.4 多次元確率分布.

4.4.1 多次元確率変数の分布.

$$\mathbb{X} = (X_1, \cdots, X_k)^T, \quad \mathbb{Y} = (Y_1, \cdots, Y_k)^T.$$

$$\mathbb{X} = \begin{pmatrix} h_1(\mathbb{Y}) \\ \vdots \\ h_k(\mathbb{Y}) \end{pmatrix} =: \mathbb{h}(\mathbb{Y})$$

Jacobian :

$$J(\mathbb{y}) := J(\mathbb{y} \to \mathbb{x}) = \det \begin{pmatrix} \frac{\partial}{\partial y_1} h_1(\mathbb{y}) & \cdots & \frac{\partial}{\partial y_k} h_1(\mathbb{y}) \\ \vdots & & \vdots \\ \frac{\partial}{\partial y_1} h_k(\mathbb{y}) & \cdots & \frac{\partial}{\partial y_k} h_k(\mathbb{y}) \end{pmatrix}$$

$$\to f_{\mathbb{Y}}(\mathbb{y}) = f_{\mathbb{X}}(\mathbb{h}(\mathbb{y}))\,|J(\mathbb{y})|.$$

4.4.2 多項分布.

Def. (多項分布)

$\mathbb{X} : \mathbb{Z}_{\geq 0}^k$-r.v., $n \in \mathbb{Z}_{>0}$, $\mathbb{p}$: prob.vec.

$\mathbb{x} = (x_1, \cdots, x_k)^T$ における joint pmf が

$$f_{\mathbb{X}}(\mathbb{x} | n, \mathbb{p}) = \binom{n}{x_1, \cdots, x_k} p_1^{x_1} \cdots\cdots p_k^{x_k}$$

のとき, $\mathbb{X}$ は 多項分布 に従うといい,

$\mathbb{X} \sim Mul(n, \mathbb{p})$ とかく. $\square$

Th. (多項定理)

$\mathcal{X}_n := \{(x_1,\cdots,x_k) \mid x_1+\cdots+x_k=n,\ x_i\in\{0,\cdots,n\}\}$ として

$$(a_1+\cdots+a_k)^n = \sum_{x\in\mathcal{X}_n}\binom{n}{x_1,\cdots,x_k} a_1^{x_1}\cdot\cdots\cdot a_k^{x_k}. \qquad \square$$

4.4.3 多変量正規分布.

Def. (多変量正規分布)

$\mathbb{X}:\mathbb{R}^k$-r.v., $\mu\in\mathbb{R}^k$, $\Sigma\in M_k(\mathbb{R})$: 正定値対称.

$x\in\mathbb{R}^k$ における joint pdf が

$$f_X(x\mid\mu,\Sigma) = \frac{1}{(2\pi)^{\frac{k}{2}}(\det\Sigma)^{\frac{1}{2}}}\exp\left(-\frac{1}{2}(x-\mu)^T\Sigma^{-1}(x-\mu)\right).$$

のとき, $\mathbb{X}$ は平均 μ, 共分散行列 Σ の多変量正規分布に従うという.

$\mathbb{X}\sim\mathcal{N}_k(\mu,\Sigma)$ とかく. $\square$

Prop. 多変量正規分布では,

X_i と X_j $(i\neq j)$ が無相関 $\iff$ X_i と X_j $(i\neq j)$ が indep. $\square$