第10回 数理統計 レポート

小森 一輝

2020年12月3日

定義 4.12. n次元連続型確率変数, 同時確率密度関数
n 次元確率変数 $\mathbb{X}$ に対して, $\mathbb{R}^n$ 上の非負関数 $f_{\mathbb{X}}$ が存在し, 確率分布 $P_{\mathbb{X}}$ が

$$P_X(\mathbf{B}) = \int_{\mathbf{B}} f_{\mathbb{X}}(x)dx \qquad (\mathbf{B} \in \mathscr{B}^n)$$

によって表されるとき, $\mathbb{X}$ は **n次元連続型確率変数 (n dimensional continuous random variable)** と呼ばれ, $f_{\mathbb{X}}$ は $\mathbb{X}$ の **同時確率密度関数 (join probability density function)** とよばれる. ここで, $x = (x_1, x_2, \dots x_n), dx = dx_1 dx_2 \cdots dx_n, \int_{\mathbf{B}} = \int_{\mathbf{B}_1} \int_{\mathbf{B}_2} \cdots \int_{\mathbf{B}_n}$ である.
このとき, $\mathbb{X}$ の同時確率分布関数は,

$$F_{\mathbb{X}}(x) = \int_{-\infty}^{x_1} \int_{-\infty}^{x_2} \cdots \int_{-\infty}^{x_n} f_{\mathbb{X}}(x)dx$$

で表される.

補足

$f_{\mathbb{X}}(x_1, x_2, \dots, x_n) \cdots \mathbb{X}$ の同時確率密度関数

定義 4.13. 周辺確率密度関数
n 次元連続確率変数 $\mathbb{X} = (X_1, X_2, \dots, X_n)$ を定義 4.4 と同様に $\mathbb{X} = (\mathbb{X}_{(p)}, (\mathbb{X})_{(q)})$ と 2 つに分ける. このとき,

$$f_{\mathbb{X}_{(p)}}(x_{(p)}) = \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} \cdots \int_{-\infty}^{\infty} f_{\mathbb{X}}(x)dx_{(q)}$$

を $\mathbb{X}_{(p)}$ の **周辺確率密度関数 (marginal probability density function) とよぶ.**

定義 4.14. 条件付き確率密度関数
n 次元連続確率変数 $\mathbb{X}=(X_1,X_2,\dots,X_n)$ を定義 4.4 と同様に $\mathbb{X}=(\mathbb{X}_{(p)},(\mathbb{X})_{(q)})$ と 2 つに分ける .このとき,

$$f_{\mathbb{X}_{(p)}|\mathbb{X}_{(q)}}(x_{(p)}\mid x_{(q)})=\begin{cases}\frac{f_{\mathbb{X}}(x)}{f_{\mathbb{X}_{(q)}}(x_{(q)})} & (f_{\mathbb{X}_{(q)}}(x_{(q)})>0)\\ 0 & (f_{\mathbb{X}_{(q)}}(x_{(q)})=0)\end{cases}$$

例 4.6. 2 次元連続型確率変数 (X,Y) の同時確率分布関数 $F_{X,Y}(x,y)$ は,

$$F_{X,Y}(x,y)=\int_{-\infty}^{x}\int_{-\infty}^{x}f_{X,Y}(s,t)dt\,ds \qquad (f_{X,Y}(s,t)>0)$$

で表され, このとき $f_{X,Y}(x,y)$ が同時確率密度関数となる.また, 周辺確率密度関数 f_X,f_Y, 条件付き確率密度関数 $f_{X|Y}$は以下で与えられる.

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

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

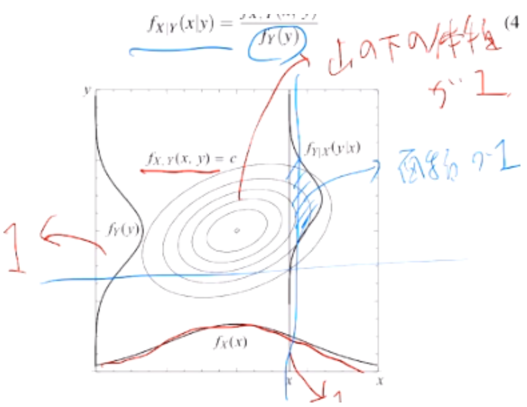
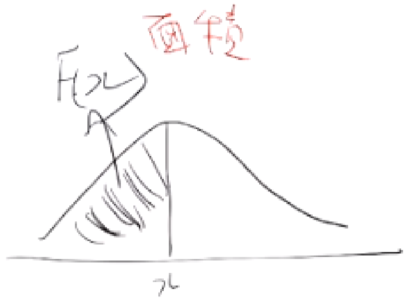
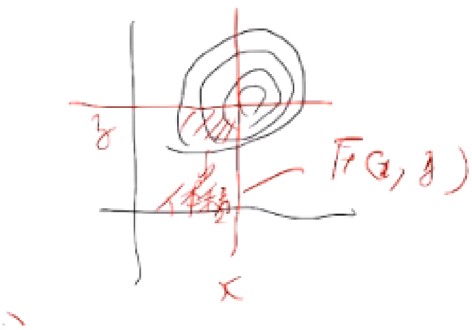


図 1: 2 次元連続型確率変数の分布

補足

同時 : $f_{X,Y}(x,y)$
周辺 : $f_X(x),f_Y(y)$
条件付き : $f_{X|Y}(x\mid y)$, $\quad$ (y が与えられたときの x の . . .) $\quad f_{Y|X}(y\mid x)$





定義 4.15. n 次元確率変数の関数の期待値と分散

n 次元確率変数 $\mathbb{X}$ のボレル関数 $h : \mathbb{R}^n \to \mathbb{R}$ による変数返還 $h(\mathbb{X})$ を考える. このとき, $h(\mathbb{X})$ の期待値, 分散を以下で定義する.

$$E(h(\mathbb{X})) = \begin{cases} \sum^{\infty} \cdots \sum^{\infty} h(x) f_{\mathbb{X}}(x) & (\mathbb{X}\text{ が離散型}) \\ \int_{-\infty}^{\infty} \cdots \int_{-\infty}^{\infty} h(x) f_{\mathbb{X}}(x) dx & (\mathbb{X}\text{ が連続型}) \end{cases}$$
$$V(h(\mathbb{X})) = E\left((h(\mathbb{X})) - E(h(\mathbb{X}))^2\right)$$

補足

$h(X, Y) = aX + bY \cdots f_{X,Y}$

$$\begin{aligned} E(aX + bY) &= \int\int (ax + by) f_{X,Y}(x, y) dxdy \\ &(= aE(X) + bE(Y)) \\ &= a \int\int x f_{X,Y}(x, y) dxdy + b \int\int y f_{X,Y}(x, y) dxdy \\ &= a \int\int x f_X(x) dx + b \int\int y f_Y(y) dy \\ &= aE(X) + bE(Y) \end{aligned}$$

例 4.7. 確率変数 $X_1, X_2, \ldots, X_n$ および $a_i \in \mathbb{R}(i = 1, 2, \ldots n)$ に対して,

$$E\left(\sum_{i=1}^{n} a_i X_i\right) = \sum_{i=1}^{n} a_i E(X_i)$$

が成り立つ.

注) x_i は独立とは限らない.

確率変数の和の期待値は, 確率変数の期待値の和

Proof. 2 次元の場合を離散と連続に分ける.
$E(a_1X_1 + a_2X_2) = a_1E(X_1) + a_2E(X_2)$

$$
\begin{aligned}
E(a_1X_1 + a_2X_2) &= \sum_{x_1}\sum_{x_2}(a_1x_1 + a_2x_2)f_{X_1,X_2}(x_1,x_2) \\
&= \sum_{x_1}\sum_{x_2} a_1x_1f_{X_1,X_2}(x_1,x_2) + \sum_{x_1}\sum_{x_2} a_2x_2f_{X_1,X_2}(x_1,x_2) \\
&= a_1\sum_{x_1} x_1 \sum_{x_2} f_{X_1,X_2}(x_1,x_2) + a_2\sum_{x_1} x_2 \sum_{x_2} f_{X_1,X_2}(x_1,x_2) \\
&= a_1\sum_{x_1} x_1 f_{X_2}(x_2) + a_2\sum_{x_1} x_2 f_{X_2}(x_2) \\
&= a_1E(X_1) + a_2E(X_2)
\end{aligned}
$$

□

命題 4.8. 確率変数 $X_1, X_2, \ldots, X_n$ および $a_i \in \mathbb{R}(i = 1, 2, \ldots n)$ に対して,

$$
\begin{aligned}
V\left(\sum_{i=1}^{n} a_iX_i\right) &= \sum_{i=1}^{n}\sum_{j=1}^{n} a_ia_jCov(X_i, X_j) \\
&= \sum_{i=1}^{n} a_i^2V(X_i) + \sum_{i\neq j} a_ia_jCov(X_i, X_j)
\end{aligned}
$$

が成り立つ.ここで, $Cov(X_i, X_j)$ については, 定義 4.16 参照.

補足

$$Cov(X_i, X_j) = E((X_i - E(X_i))(X_j - E(X_j)))$$

$$
\begin{aligned}
Cov(X, Y) &= E_{X,Y}((X\_E(X))(Y - E(Y))) \\
&= E_{X,Y}(XY) - E_X(X)E_Y(Y)
\end{aligned}
$$

Proof. X_iが独立なら, $V(\sum_{i=1}^{n} a_iX_i) = \sum_{i=1}^{n} a_iV(X_i)$

□

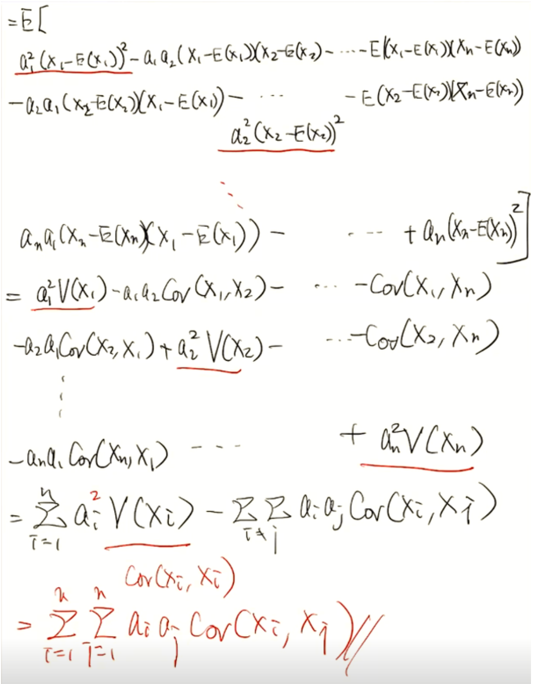


図 2: 証明

定理 4.9. 確率変数の和の分布
X, Y を離散型または連続型の確率変数とし, それぞれの同時確率関数または同時確率密度関数を $f_{X,Y}$ で表す. このとき, 確率変数 X, Y の和 $Z = X + Y$ の確率関数または確率密度関数 f_z は,

$$f_Z(z) = \begin{cases} \sum_{i=1}^{\infty} f_{X,Y}(x_i, z - x_i) & (X, Y \text{ が離散型}) \\ \int_{-\infty}^{\infty} f_{X,Y}(x, z - x)dx & (X, Y \text{ が連続型}) \end{cases}$$

で与えられる.

Proof. 離散型

$$f_Z(z) = \sum_{x_i + y_j} f_{X,Y}(x_i, y_j) = \sum_{i=1}^{\infty} f_{X,Y}(x_i, z - x_i)$$

連続型 U=X+Y, V=X とおく.

$$
\begin{aligned}
f_{u,v}(u,v) &= f_{X,Y}(v,u-v)|-1| \\
f_v(u) &= \int_{-\infty}^{\infty} f_{u,v}(u,v)dv \\
&= \int_{-\infty}^{\infty} f_{X,Y}(v,u-v)dv
\end{aligned}
$$

□

これをたたみこみという.

定理 4.10. 確率変数の積と商の分布
X, Y を連続型の確率変数とし, それぞれおの同時確率密度関数を $f_{X,Y}$で表す. このとき, 確率変数 $X.Y$ の積 $Z_1 = XY$, 商 $Z_2 = X/Y$ の確率密度関数 f_{Z_1}, f_{Z_2}は,
それぞれ,

$$
\begin{aligned}
f_{Z_1(z_1)} &= \int_{-\infty}^{\infty} f_{X,Y}\left(x, \frac{z_1}{x}\right)\frac{1}{|x|}dx = \int_{-\infty}^{\infty} f_{X,Y}\left(\frac{z_1}{y}, y\right)\frac{1}{|y|}dy \\
f_{Z_2(z_2)} &= \int_{-\infty}^{\infty} f_{X,Y}\left(x, \frac{x}{z_2}\right)\frac{|x|}{z_2^2}dx = \int_{-\infty}^{\infty} f_{X,Y}\left(z_2 y, y\right)|y|dy
\end{aligned}
$$

で与えられる.

補足

$$
\begin{aligned}
&1\text{ 行目} && \to U = XY,\ V = X(V = Y \text{ でも } OK) \\
&2\text{ 行目} && \to U = XY,\ V = X(V = Y \text{ でも } OK)
\end{aligned}
$$

4.3 2次元確率変数の特性値

定理 4.16. 2 次元確率変数 (X, Y) の積率
2 次元確率変数 (X, Y) の原点まわりの $i \neq j$ 次積率μ'_{ij}および, 平均まわりの $i \neq j$ 次積率μ_{ij}をそれぞれ以下のように定める.

$$
\mu'_{ij} = E(X^i Y^j), \qquad \mu_{ij} = E((X - E(X))^i (Y - E(Y))^j)
$$

特に,μ_{11}を X と Y の**共分散 (covariance)** とよび,$Cov(X,Y)$ と記す.

$$\mu_1 1 = Cov(X,Y) = E((X-E(X))(Y-E(Y)))$$

命題 4.11. 共分散の性質
共分散 $Cov(X,Y)$ について,以下が成り立つ.

i) $Cov(X,Y) = Cov(Y,X)$

ii) $Cov(X,X) = V(X), \quad Cov(Y,Y) = V(Y)$

iii) $Cov(X,Y) = E(XY) - E(X)E(Y)$

iv) $(Cov(X,Y))^2 \leq V(X)V(Y)$