中間試験のお知らせ

日時: 6 月 12 日 (金) 10:40〜
場所: いつもの講義室
範囲: 第 6 回講義まで
持ち込み禁止
遅刻・欠席はその原因を証明する公的書類を提出した場合だけ追試験または試験時間の延長を行う

2次元正規分布 (教科書の7.3節)

前回の煙の粒子の水平座標は正規分布に従ったが、高さが異なる 2 点で水平座標を測定すれば、異なる座標が 2 つ得られ、それらの周辺分布は正規分布に従うが、同時分布はなんだかわからない。今回は 2 つの正規確率変数の同時分布として 2 次元正規分布を説明する。

2 次元正規分布とは確率密度関数が以下の式で与えられる分布を言う

$$\frac{1}{2\pi\sqrt{\begin{vmatrix} \sigma_{11} & \sigma_{12} \\ \sigma_{12} & \sigma_{22} \end{vmatrix}}}\times$$

$$\exp\left(-\frac{1}{2}(x_1-m_1\ x_2-m_2)\begin{pmatrix} \sigma_{11} & \sigma_{12} \\ \sigma_{12} & \sigma_{22} \end{pmatrix}^{-1}\begin{pmatrix} x_1-m_1 \\ x_2-m_2 \end{pmatrix}\right) \tag{1}$$

但し行列 $\begin{pmatrix} \sigma_{11} & \sigma_{12} \\ \sigma_{12} & \sigma_{22} \end{pmatrix}$ は逆行列を持ち固有値が正であるとする。

これだけだとよくわからないので、独立な 2 つの標準正規確率変数 (平均 0, 分散 1 の正規分布に従う確率

変数) Y_1, Y_2 を変換して上記の確率密度関数を持つ確率変数を得られることを示す (146 ページ)。Y_1, Y_2 の同時確率密度関数は

$$
\begin{aligned}
& f_{Y_1Y_2}(y_1, y_2) \\
= \quad & f_{Y_1}(y_1) f_{Y_2}(y_2) \\
= \quad & \frac{1}{\sqrt{2\pi}} \exp(-y_1^2/2) \frac{1}{\sqrt{2\pi}} \exp(-y_2^2/2) \\
= \quad & \frac{1}{2\pi} \exp(-(y_1^2 + y_2^2)/2) \\
= \quad & \frac{1}{2\pi} \exp\left(-\frac{1}{2} (y_1 \ y_2) \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix}^{-1} \begin{pmatrix} y_1 \\ y_2 \end{pmatrix} \right) \quad (2)
\end{aligned}
$$

となっていて、式 (1) に $\sigma_{11} = \sigma_{22} = 1, \sigma_{12} = 0, m_1 = m_2 = 0$ を代入したものに等しいことに注意。この一致の意味を後で説明する。
確率変数 Z_1, Z_2 を

$$
\begin{pmatrix} Z_1 \\ Z_2 \end{pmatrix} = \begin{pmatrix} a & b \\ c & d \end{pmatrix} \begin{pmatrix} Y_1 \\ Y_2 \end{pmatrix}
$$

で定義する。Z_i は正規確率変数の線形和なので一次元正規分布に従う。このとき

$$
E\left[\begin{pmatrix} Z_1 \\ Z_2 \end{pmatrix} \right]
$$

$$
\begin{aligned}
&= E\left[\begin{pmatrix} a & b \\ c & d \end{pmatrix}\begin{pmatrix} Y_1 \\ Y_2 \end{pmatrix}\right] \\
&= \begin{pmatrix} a & b \\ c & d \end{pmatrix} E\left[\begin{pmatrix} Y_1 \\ Y_2 \end{pmatrix}\right] \\
&= \begin{pmatrix} a & b \\ c & d \end{pmatrix}\begin{pmatrix} 0 \\ 0 \end{pmatrix} \\
&= \begin{pmatrix} 0 \\ 0 \end{pmatrix}
\end{aligned}
$$

となる。すなわち $E[Z_1] = 0$, $E[Z_2] = 0$ である。
次に Z_1 の分散 $V[Z_1]$ を求める。

$$
\begin{aligned}
& V[Z_1] \\
= \ & V[aY_1 + bY_2] \\
= \ & V[aY_1] + V[bY_2] + 2\mathrm{Cov}[aY_1, bY_2] \\
= \ & V[aY_1] + V[bY_2] \\
= \ & a^2 + b^2
\end{aligned}
$$

同様に $V[Z_2] = c^2 + d^2$ がわかる。
共分散は

$$
\begin{aligned}
& \mathrm{Cov}[Z_1, Z_2] \\
= \ & E[Z_1 Z_2] - E[Z_1]E[Z_2] \\
= \ & E[(aY_1 + bY_2)(cY_1 + dY_2)]
\end{aligned}
$$

$$
\begin{aligned}
&= E[acY_1^2 + bdY_2^2 + (ad - bc)Y_1Y_2] \\
&= acV[Y_1] + bdV[Y_2] + (ad - bc)E[Y_1Y_2] \\
&= ac + bd + (ad - bc)E[Y_1]E[Y_2] \\
&= ac + bd
\end{aligned}
$$

Z_1, Z_2 の同時確率密度

次に Z_1, Z_2 の同時確率密度関数を求める。まず Z_1, Z_2 の共分散は非零なので Z_1 と Z_2 は統計的に独立ではない。従って $f_{Z_1Z_2}$ を f_{Z_1} と f_{Z_2} の積として求めることはできない。Δz_1, Δz_2 を小さい正数とすると $(z_1,\ z_2)$, $(z_1+\Delta z_1,\ z_2)$, $(z_1+\Delta z_1,\ z_2+\Delta z_2)$, $(z_1,\ z_2+\Delta z_2)$ で囲まれる四角形で確率密度関数 $f_{Z_1Z_2}$ の値はだいたい一定である。従ってこの四角形に Z_1, Z_2 の実現値が入る確率は

$$\Delta z_1 \Delta z_2 f_{Z_1Z_2}(z_1, z_2) \tag{3}$$

に近い (第４回講義で既にやった)。
一方

$$\begin{pmatrix} y_1 \\ y_2 \end{pmatrix} = \begin{pmatrix} a & b \\ c & d \end{pmatrix}^{-1} \begin{pmatrix} z_1 \\ z_2 \end{pmatrix} \tag{4}$$

とすると $Y_1 = y_1$ かつ $Y_2 = y_2$ となる事象と $Z_1 = z_1$ かつ $Z_2 = z_2$ となる事象は同じである。また $(z_1,\ z_2)$, $(z_1+\Delta z_1,\ z_2)$, $(z_1+\Delta z_1,\ z_2+\Delta z_2)$, $(z_1,\ z_2+\Delta z_2)$ で囲まれる面積 $\Delta z_1 \Delta z_2$ の四角形を上記変換で移した四角形の面積は

$$\frac{\Delta z_1 \Delta z_2}{\begin{vmatrix} a & b \\ c & d \end{vmatrix}}$$

になる。従って Y_1, Y_2 がこの四角形に入る確率はおおよそ

$$\frac{\Delta z_1 \Delta z_2 f_{Y_1 Y_2}(y_1, y_2)}{\begin{vmatrix} a & b \\ c & d \end{vmatrix}} \tag{5}$$

である。但し分母の行列式が負のときはその絶対値をとるものとする。式 (3) と式 (5) が等しいことから

$$f_{Z_1 Z_2}(z_1, z_2) = \frac{f_{Y_1 Y_2}(y_1, y_2)}{\begin{vmatrix} a & b \\ c & d \end{vmatrix}}$$

である。ここで上の右辺を z_1 と z_2 を用いて表現したいから式 (2) に式 (4) を代入すると $f_{Y_1 Y_2}(y_1, y_2)$ は

$$\frac{1}{2\pi} \exp\left(-\frac{1}{2} (z_1 \ z_2) \left(\begin{pmatrix} a & b \\ c & d \end{pmatrix}^{-1} \right)^T \begin{pmatrix} a & b \\ c & d \end{pmatrix}^{-1} \begin{pmatrix} z_1 \\ z_2 \end{pmatrix} \right)$$

と書き直せる。式の中の行列の積を整理して

$$\begin{aligned} & \left(\begin{pmatrix} a & b \\ c & d \end{pmatrix}^{-1} \right)^T \begin{pmatrix} a & b \\ c & d \end{pmatrix}^{-1} \\ = \ & \frac{1}{(ad - bc)^2} \begin{pmatrix} d & -b \\ -c & a \end{pmatrix}^T \begin{pmatrix} d & -b \\ -c & a \end{pmatrix} \end{aligned}$$

$$
\begin{aligned}
&= \frac{1}{(ad-bc)^2}\begin{pmatrix} d & -c \\ -b & a \end{pmatrix}\begin{pmatrix} d & -b \\ -c & a \end{pmatrix} \\
&= \frac{1}{(ad-bc)^2}\begin{pmatrix} c^2+d^2 & -ac-bd \\ -ac-bd & a^2+b^2 \end{pmatrix} \qquad (6)
\end{aligned}
$$

一方ここで $\sigma_{11} = a^2+b^2$, $\sigma_{12} = ac+bd$, $\sigma_{22} = c^2+d^2$ とおくと

$$
\begin{aligned}
&\begin{pmatrix} \sigma_{11} & \sigma_{12} \\ \sigma_{12} & \sigma_{22} \end{pmatrix}^{-1} \\
&= \frac{1}{\sigma_{11}\sigma_{22}-\sigma_{12}^2}\begin{pmatrix} \sigma_{22} & -\sigma_{12} \\ -\sigma_{12} & \sigma_{11} \end{pmatrix} \\
&= \frac{1}{(ad-bc)^2}\begin{pmatrix} c^2+d^2 & -ac-bd \\ -ac-bd & a^2+b^2 \end{pmatrix}
\end{aligned}
$$

となり式 (6) と等しくなる。このことは、$f_{Z_1Z_2}(z_1, z_2)$ が

$$
\begin{aligned}
& f_{Z_1 Z_2}(z_1, z_2) \\
= \quad & \frac{1}{2\pi \begin{vmatrix} a & b \\ c & d \end{vmatrix}} \exp\left(-\frac{1}{2}(z_1 \quad z_2) \frac{1}{(ad-bc)^2} \right. \\
& \left. \begin{pmatrix} c^2+d^2 & -ac-bd \\ -ac-bd & a^2+b^2 \end{pmatrix} \begin{pmatrix} z_1 \\ z_2 \end{pmatrix} \right) \\
= \quad & \frac{1}{2\pi \sqrt{\begin{vmatrix} \sigma_{11} & \sigma_{12} \\ \sigma_{12} & \sigma_{22} \end{vmatrix}}} \times \\
& \exp\left(-\frac{1}{2}(z_1 z_2) \begin{pmatrix} \sigma_{11} & \sigma_{12} \\ \sigma_{12} & \sigma_{22} \end{pmatrix}^{-1} \begin{pmatrix} z_1 \\ z_2 \end{pmatrix} \right)
\end{aligned}
$$

と書き直すことができて、これは式 (1) において $m_1 = m_2 = 0$ としたものに等しい。逆に言えば式 (1) の確率密度関数は (Z_1, Z_2) を (m_1, m_2) だけ平行移動した確率変数に等しい。

このことから式 (1) の確率密度関数に従う確率変数 X_1, X_2 は期待値 m_1, m_2, 分散 σ_{11}, σ_{22}, 共分散 σ_{12} を持つことがわかる。また X_1, X_2 は独立な標準正規確率変数 Y_1, Y_2 に行列を掛けて線形変換したあとに平行移動して得ることができる。

勉強とは公式の当てはめ方と解法の暗記だと誤解している人がいるが大間違いである。暗記するなら導き

方まで含めて暗記して下さい。本日レポートの東工大大学院入試問題はここまでの導き方を元にした問題です。

よく使う解法だけ暗記して考え方を理解していない人は、新しい状況に対して学んだことをどのように用いればいいか判断できないので現場で使いものにならない。

n 次元正規分布

以上の定義を一般化して n 次元正規確率変数 $X_1, \ldots, X_n$ を定義することができる。Σ を (i,j) 成分が σ_{ij} である可逆な実対称行列でその固有値はすべて正であるとする (すなわち Σ は正値対称行列 (斎藤正彦著「線形代数入門」の 152 ページ参照))。また $\vec{m} = (m_1, \ldots, m_n)^T$, $\vec{x} = (x_1, \ldots, x_n)^T$ とする。$X_1, \ldots, X_n$ の確率密度関数が以下の形を持つときに $X_1, \ldots, X_n$ を n 次元正規確率変数と呼ぶ。

$$f_{X_1 \cdots X_n}(x_1, \ldots, x_n) = \frac{1}{\sqrt{2^n \pi^n |\Sigma|}} \exp\left((\vec{x} - \vec{m})^T \Sigma^{-1} (\vec{x} - \vec{m}) \right)$$

このとき X_i の期待値は m_i, 分散は σ_{ii} であり X_i と X_j の共分散は σ_{ij} となる。$X_1, \ldots, X_n$ も n 個の独立な標準正規確率変数を線形変換した後に平行移動して得ることができる。したがってすべての i について X_i が一次元正規分布に従うことがわかる。
教科書 7.3 節と異なり本講義では行列を用いて n 次元正規分布を定義したが、行列を用いたおかげで $n \geq 3$ のときも密度関数を記述することが可能になっている。教科書の表記法では $n \geq 3$ の場合にどのように密度関数が一般化されるかまったく予想できないことに注意。

添付の演習問題を解答し、6 月 1 日 (月) の夜までに南 3 号館 3 階 311 号室前の箱に提出すること。また前回演習問題で間違え正しい解答が理解できない場合と、今回演習問題でわからない部分がある場合には講義時間内に講師に確認しておくこと。
この演習問題は過去の東工大大学院入試で 40 分程度で解答できることを目指して作成されている。従って大学院を受験する際は、この程度の問題は 40 分以内に処理できるように準備しておく必要がある。問題の内容は今回授業内容を 2 次元から N 次元に拡張したものになっている。