L02 2 次元正規分布

樋口さぶろお　https://hig3.net

龍谷大学 先端理工学部 数理・情報科学課程

多変量解析☆演習　L02(2021-10-07 Thu)

最終更新: Time-stamp: "2021-10-06 Wed 17:34 JST hig"

今日の目標

- 2 次元正規分布の確率密度関数を母…から求められる
- 2 次元正規分布の母ナントカを確率密…から求められる
- 2 次元正規分布の周辺分布と再生性を説明できる

L00-Q1

Quiz 解答:2 次元の連続型確率分布の周辺分布

1.
$$f_{\mathrm{X}}(x) = \begin{cases} \frac{1}{6} & (1 \le x < 7) \\ 0 & (\text{他} \end{cases}$$
$$f_{\mathrm{Y}}(y) = \begin{cases} \frac{1}{6} & (1 \le y < 7) \\ 0 & (\text{他} \end{cases}$$

2. 任意の x, y に対して $f(x, y) = f_{\mathrm{X}}(x) \times f_{\mathrm{Y}}(y)$ が成立するので独立.
3. 任意の a に対して $f_{\mathrm{X}}(a) = f_{\mathrm{Y}}(a)$ が成立するので同分布.

L00-Q2

Quiz 解答:2 つの連続型確率変数の周辺分布

1.

$$f_{\mathrm{X}}(x) = \begin{cases} \frac{1}{18}(6-x) & (0 \le x < 6) \\ 0 & (\text{他}) \end{cases},$$

$$f_{\mathrm{Y}}(y) = \begin{cases} \frac{1}{18}(6-y) & (0 \le y < 6) \\ 0 & (\text{他}) \end{cases}.$$

2. $f(x,y) = f_{\mathrm{X}}(x) \times f_{\mathrm{Y}}(y)$ が成立しないので独立でない．反例．
   $(x,y) = (5,5)$ に対して，$0 \neq \frac{1}{18} \times \frac{1}{18}$.
3. 任意の a に対して $f_{\mathrm{X}}(a) = f_{\mathrm{Y}}(a)$ が成立するので同分布.

ここまで来たよ

- 確率変数の和
- 2 次元正規分布
  - 一般の 1 次元正規分布
  - 2 次元正規分布 (X, Y が独立)
  - 2 次元正規分布 (X, Y が独立でない)

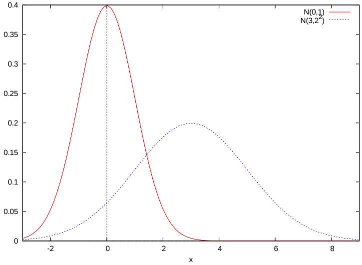
一般の 1 次元正規分布 normal distribution $\mathrm{N}(\mu, \sigma^2)$

一般の 1 次元正規分布 $\mathrm{N}(\mu, \sigma^2)$ 岩薩林 確率・統計 (4.23) 確率統計 I(2021)L08

母平均値 $\mathrm{E}[X] = \mu$, 母分散 $\mathrm{V}[X] = \sigma^2$ の正規分布 $\mathrm{N}(\mu, \sigma^2)$ の確率密度関数は,

$$f(x; \mu, \sigma^2) = \frac{1}{\sqrt{2\pi\sigma^2}} \mathrm{e}^{-\frac{(x-\mu)^2}{2\sigma^2}}.$$

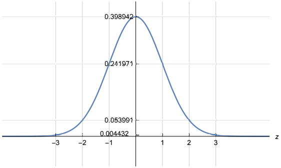
標準正規分布は, $\mu = 0, \sigma^2 = 1$ のときのこと. $Z \sim \mathrm{N}(0, 1^2)$.



標準正規分布の母期待値

$Z \sim \mathrm{N}(0, 1^2)$
k 次のモーメント (k:自然数)

$$
\begin{aligned}
\mathrm{E}[Z^{2k-1}] &= 0, \quad \text{奇関数} \\
\mathrm{E}[Z^{2k}] &= (2k-1)!! \\
&= 1 \cdot 3 \cdot 5 \cdot \cdots \cdot (2k-1) \quad \text{置換積分}
\end{aligned}
$$



全確率 $\mathrm{E}[Z^0] = 1,$ 岩薩林 確率・統計 (4.19) 微積分 II

母平均値 $\mathrm{E}[Z] = 0,$ 岩薩林 確率・統計 (4.20) 奇関数

母分散 $\mathrm{V}[Z] = 1.$ 岩薩林 確率・統計 (4.21)

$Z \sim \mathrm{N}(0, 1^2)$ と $X \sim \mathrm{N}(\mu, \sigma^2)$ との関係は, $X = \sigma Z + \mu$. モーメント

$$1 = \mathrm{E}[X^0] = \int_{-\infty}^{+\infty} x^0 \cdot \frac{1}{\sqrt{2\pi\sigma^2}} \mathrm{e}^{-\frac{(x-\mu)^2}{2\sigma^2}}\, \mathrm{d}x$$

$$\mu = \mathrm{E}[X] = \int_{-\infty}^{+\infty} x^1 \cdot \frac{1}{\sqrt{2\pi\sigma^2}} \mathrm{e}^{-\frac{(x-\mu)^2}{2\sigma^2}}\, \mathrm{d}x$$

$$\sigma^2 + \mu^2 = \mathrm{V}[X^2] + \mathrm{E}[X]^2 = \mathrm{E}[X^2] = \int_{-\infty}^{+\infty} x^2 \cdot \frac{1}{\sqrt{2\pi\sigma^2}} \mathrm{e}^{-\frac{(x-\mu)^2}{2\sigma^2}}\, \mathrm{d}x$$

を公式のように使って計算する. 右辺の形にもっていって左辺を答える.
L02-Q1

Quiz(正規分布のモーメントを利用した定積分の計算)

$f(x) = \mathrm{e}^{-2x^2+24x}$ に対して, 次の定積分の値を求めよう.

1. $\int_{-\infty}^{+\infty} f(x)\, \mathrm{d}x.$
2. $\int_{-\infty}^{+\infty} x \cdot f(x)\, \mathrm{d}x.$
3. $\int_{-\infty}^{+\infty} x^2 \cdot f(x)\, \mathrm{d}x.$

ここまで来たよ

- 確率変数の和
- 2 次元正規分布
  - 一般の 1 次元正規分布
  - 2 次元正規分布 (X, Y が独立)
  - 2 次元正規分布 (X, Y が独立でない)

2 次元正規分布 (X, Y が独立でない) をプレビュー

一般の 2 次元正規分布 (あとからもっと美しい形で書く)

連続型確率変数 X, Y が 次の同時確率密度関数 確率統計 I(2021)L03 を持つとき，X, Y は 2 次元正規分布にしたがうという．

$$\begin{aligned} f(x,y) &= \text{定数} \times \mathrm{e}^{-a(x-\mu_\mathrm{X})^2 - b(y-\mu_\mathrm{Y})^2 + cxy} \\ &= \text{定数}' \times \mathrm{e}^{-ax^2 - by^2 + cxy + px + qy} \end{aligned}$$

$a, b > 0$. c と定数はある条件を満たす定数．

2 次元正規分布で $c = 0$(X, Y が独立) 岩薩林 確率・統計 §4.6, 例題 4.12

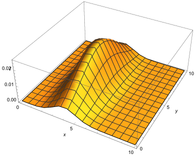
確率変数 X, Y に対して, 次の同時確率密度関数 確率統計 I(2021)L03 を考える.

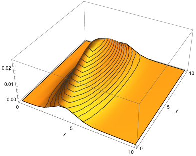
$$f(x, y) = \frac{1}{(2\pi\sigma_{\mathrm{X}}^2)^{1/2}} \mathrm{e}^{-\frac{(x-\mu_{\mathrm{X}})^2}{2\sigma_{\mathrm{X}}^2}} \times \frac{1}{(2\pi\sigma_{\mathrm{Y}}^2)^{1/2}} \mathrm{e}^{-\frac{(y-\mu_{\mathrm{Y}})^2}{2\sigma_{\mathrm{Y}}^2}}.$$

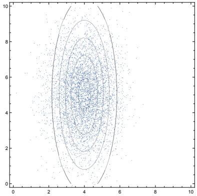
$c = 0$ のとき, X, Y の周辺分布は… よって X, Y は独立.

$$\mathrm{E}[X] = \mu_{\mathrm{X}}, \quad \mathrm{E}[Y] = \mu_{\mathrm{Y}}, \quad \mathrm{V}[X] = \sigma_{\mathrm{X}}^2, \quad \mathrm{V}[Y] = \sigma_{\mathrm{Y}}^2, \mathrm{E}[XY] = \mu_{\mathrm{X}} \cdot \mu_{\mathrm{Y}}$$

母共分散 $\mathrm{Cov}[X, Y] = 0$, 母相関係数 $r_{\mathrm{XY}} = \dfrac{\mathrm{Cov}[X, Y]}{\sqrt{\mathrm{V}[X]}\sqrt{\mathrm{V}[Y]}} = 0$







f の等高線は $(\mu_{\mathrm{X}}, \mu_{\mathrm{Y}})$ を中心とする楕円.

L02-Q2

Quiz(2 次元正規分布)

次の同時確率密度関数は 2 次元正規分布を定める.

$$f(x,y) = C \cdot \mathrm{e}^{-x^2-4x-2y^2+12y-5}.$$

1. X, Y の母平均値, 母分散, 母共分散を求めよう.
2. $\mathrm{E}[1] = 1$ が満たされるように定数 C を定めよう.

2 次形式の標準化 (線形代数 I)

L02-Q3

Quiz(2 次元正規分布)

次の 2 変数確率密度関数は 2 次元正規分布を定める.

$$f(x,y) = C \cdot \mathrm{e}^{-4x^2 - \frac{1}{6}y^2 + 2y}$$

1. X, Y の母平均値, 母分散, 母共分散を求めよう.
2. $\mathrm{E}[1] = 1$ が満たされるように定数 C を定めよう.

再生性 reproductive property

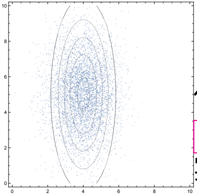
正規分布の再生性 岩薩林 確率・統計 定理 4.3

独立な連続型確率変数 $X \sim \mathrm{N}(\mu_{\mathrm{X}}, \sigma_{\mathrm{X}}^2)$, $Y \sim \mathrm{N}(\mu_{\mathrm{Y}}, \sigma_{\mathrm{Y}}^2)$ に対して,

$$S = aX + bY \sim \mathrm{N}(a\mu_{\mathrm{X}} + b\mu_{\mathrm{Y}}, a^2\sigma_{\mathrm{X}}^2 + b^2\sigma_{\mathrm{Y}}^2)$$

特に,

$$S = X + Y \sim \mathrm{N}(\mu_{\mathrm{X}} + \mu_{\mathrm{Y}}, \sigma_{\mathrm{X}}^2 + \sigma_{\mathrm{Y}}^2)$$

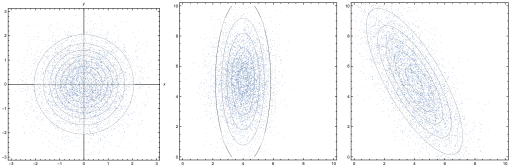


再生性とは, 独立な確率変数 X, Y の和 $S = X + Y$ が (パラメタは違うけど) 同じ分布 (今なら正規分布) にしたがうという, 稀なよい性質.
他に, 二項分布, ポアソン分布など.
岩薩林 確率・統計 例題 4.14 特定の a, b, μ, σ での証明
実は, 一般の (独立とかぎらない)2 次元正規分布に対しても $aX + bY$ は 1 次元正規分布にしたがうことがわかる).

ここまで来たよ

- 確率変数の和
- 2 次元正規分布
  - 一般の 1 次元正規分布
  - 2 次元正規分布 (X, Y が独立)
  - 2 次元正規分布 (X, Y が独立でない)

等高線が傾いた楕円の分布も作りたい!



考え方 1
等高線が円 $D = x^2 + y^2$
等高線が傾いた楕円 $D = x^2 + cxy + 2y^2$　　2 次形式の標準化 (線形代数 I,II)
考え方 2 独立じゃなくしたいから, 指数関数の引数に xy の項でもいれておけば?

一般の 2 次元正規分布

$$f(x,y) = \text{定数}' \times \mathrm{e}^{-ax^2-by^2+cxy+px+qy}$$

$\rightsquigarrow$

一般の 2 次元正規分布の同時確率密度関数

2 次元正規分布 $\mathrm{N}(\boldsymbol{\mu}, \Sigma)$ の同時確率密度関数は, 確率変数を $\boldsymbol{X} = {}^{\mathrm{t}}(X, Y)$ とするとき,

$$f(\boldsymbol{x}) = \frac{1}{(2\pi)^{2/2}\sqrt{\det \Sigma}} \mathrm{e}^{-\frac{1}{2}{}^{\mathrm{t}}(\boldsymbol{x}-\boldsymbol{\mu})\Sigma^{-1}(\boldsymbol{x}-\boldsymbol{\mu})}$$

$$\text{母平均値 (ベクトル)}\boldsymbol{\mu} = \begin{pmatrix} \mu_{\mathrm{X}} \\ \mu_{\mathrm{Y}} \end{pmatrix} \overset{\text{要証明}}{=} \begin{pmatrix} \mathrm{E}[X] \\ \mathrm{E}[Y] \end{pmatrix}$$

$$\text{母共分散行列}\Sigma = \begin{pmatrix} \sigma_{\mathrm{X}}^2 & C_{\mathrm{XY}} \\ C_{\mathrm{XY}} & \sigma_{\mathrm{Y}}^2 \end{pmatrix} \overset{\text{要証明}}{=} \begin{pmatrix} \mathrm{V}[X] & \mathrm{Cov}[X,Y] \\ \mathrm{Cov}[X,Y] & \mathrm{V}[Y] \end{pmatrix}$$

L02-Q4

Quiz(2 次元正規分布の確率密度関数)

2 次元正規分布の同時確率密度関数で,

$$\boldsymbol{\mu} = \begin{pmatrix} -4 \\ 5 \end{pmatrix},$$

$$\Sigma = \begin{pmatrix} 2 & -1 \\ -1 & 3 \end{pmatrix}$$

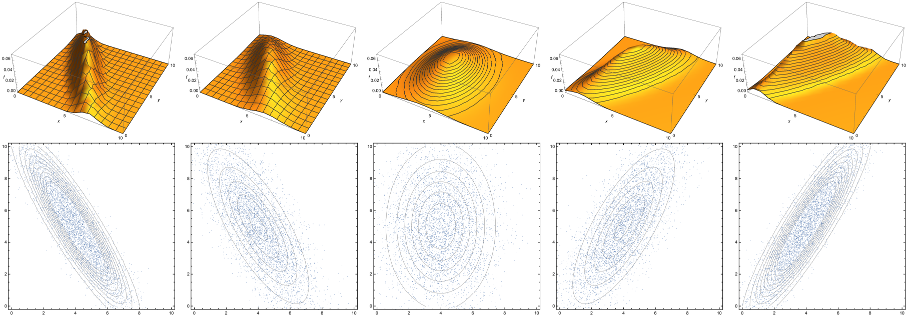
のとき,

1. 母平均値, 母 (共) 分散を答えよう.
2. 同時確率密度関数 f を (行列やベクトルを使わずに) 具体的に書こう.

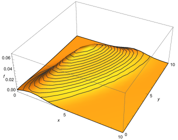
母相関係数 r, 共分散行列 Σ. $c = \frac{9\sqrt{21}}{10}$

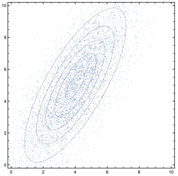
$r = -0.90$ $\quad -0.76$ $\quad 0$ $\quad +0.76$ $\quad +0.90$

$$\Sigma = \begin{pmatrix} 3 & -c \\ -c & 7 \end{pmatrix}, \begin{pmatrix} 3 & -2\sqrt{3} \\ -2\sqrt{3} & 7 \end{pmatrix}, \begin{pmatrix} 3 & 0 \\ 0 & 7 \end{pmatrix}, \begin{pmatrix} 3 & +2\sqrt{3} \\ +2\sqrt{3} & 7 \end{pmatrix}, \begin{pmatrix} 3 & +c \\ +c & 7 \end{pmatrix}$$



2 次元正規分布の周辺分布は 1 次元正規分布





このほか, $X = x_0$ の条件付き分布, $S = aX + bY$ の分布も 1 次元正規分布 → 再生性

確率密度関数から母ナントカ I

L02-Q5

Quiz(2 次元正規分布)

次の 2 変数確率密度関数は 2 次元正規分布を定める.

$$f(x,y) = C \cdot \mathrm{e}^{-2x^2+6xy-7y^2+26x-54y-7}.$$

1. 母平均値, 母共分散行列を求めよう.
2. $\mathrm{E}[1] = 1$ が満たされるように定数 C を定めよう.

L02-Q6

Quiz(2 次元正規分布)

次の 2 変数確率密度関数は 2 次元正規分布を定める.

$$f(x,y) = C \cdot \mathrm{e}^{-x^2+4xy-7y^2}.$$

1. 母共分散行列を求めよう.
2. $\mathrm{E}[1] = 1$ が満たされるように定数 C を定めよう.

2 次形式の標準化 (線形代数 I)

L02-Q7

Quiz(2 次元正規分布)

次の 2 変数確率密度関数は 2 次元正規分布を定める.

$$f(x,y) = C \cdot \mathrm{e}^{-2x^2-xy-\frac{1}{2}y^2}.$$

1. 母共分散行列を求めよう.
2. $\mathrm{E}[1] = 1$ が満たされるように定数 C を定めよう.

2 次形式の標準化 (線形代数 I)