L02 2 次元正規分布

樋口さぶろお

龍谷大学 先端理工学部 数理・情報科学課程

理論物理学特論 L02(2021-09-28 Tue)

最終更新: Time-stamp: "2021-09-29 Wed 18:57 JST hig"

今日の目標

- 確率変数の和の確率密度関数が求められる
- $\mathrm{e}^{x,y}$ の 2 次式 の分布の母平均値と母 (共) 分散が求められる
- 指定の母平均値と母 (共) 分散の 2 次元正規分布の確率密度関数が書ける

2 次元連続型 永田棟方 多変量解析法入門 2.2(4)

X, Y を赤白 2 個の**連続**サイコロを振ったときに出る目 $(1 \leq x < 7)$ とする (この場合, たまたま X, Y は独立同分布にしたがう).
同時分布の確率密度関数は

$$f(x, y) = \begin{cases} \frac{1}{36} & (1 \leq x < 7, 1 \leq y < 7) \\ 0 & (\text{他}) \end{cases}.$$

グラフで描いた方が早い.
周辺分布 X の周辺分布の確率密度関数 $f_{\mathrm{X}}(x)$ は?
確率 $P(c \leq x < d) = \int_{-\infty}^{+\infty} \mathrm{I}_{[c \leq x < d]}(x) f_{\mathrm{X}}(x)\,\mathrm{d}x$ となるべき.
左辺の定義 $\iint_{-\infty}^{+\infty} \mathrm{I}_{[c \leq x < d]}(x, y) f(x, y)\,\mathrm{d}y\mathrm{d}x$
ってことは

周辺分布 (連続)

$$f_{\mathrm{X}}(x) = \int_{-\infty}^{+\infty} f(x, y)\,\mathrm{d}y.$$

L02-Q1

Quiz(2 次元の連続型確率分布の周辺分布)

連続型確率変数 X, Y の同時確率密度関数 $f(x, y)$ が次で与えられる.

$$f(x,y) = \begin{cases} \frac{1}{3} & (0 \leq x < 3, 0 \leq y < 2 - \frac{2}{3}x) \\ 0 & (\text{他}) \end{cases}$$

X, Y の周辺分布を求めよう.

和の分布

$S = X + Y$ の確率密度関数 $f_{\mathrm{S}}(s)$ は?
確率 $P(c \le S < d) = \int_{-\infty}^{+\infty} \mathrm{I}_{[c \le s < d]}(x) f_{\mathrm{S}}(s)\ \mathrm{d}s$ となるべき.

$$\begin{aligned} \text{左辺} &\overset{\text{定義}}{=} \iint_{-\infty}^{+\infty} \mathrm{I}_{[c \le x+y < d]}(x,y) f(x,y)\ \mathrm{d}y\mathrm{d}x \\ &= \iint_{-\infty}^{+\infty} \mathrm{I}_{[c \le s < d]}(s) f(t, s-t)\ \mathrm{d}s\mathrm{d}t \quad \text{変換 } s = x+y,\ t = x.\ |J| = 1 \end{aligned}$$

よって, $f_{\mathrm{S}}(s) = \int_{-\infty}^{+\infty} f(t, s-t)\mathrm{d}t$.
($c = -\infty$ とおいて累積分布関数 $F(d) = P(S < d)$ で話すことも多い)

和の確率密度関数

連続型 X, Y の確率密度関数:$f(x,y)$. 和 $S = X + Y$ の確率密度関数は,

$$f_{\mathrm{S}}(s) = \int_{-\infty}^{+\infty} f(x, s-x)\mathrm{d}x.$$

たたみ込み convolution

和の確率密度関数

連続型確率密度関数が独立で, X, Y の確率密度関数 $f(x,y)$ が $f(x,y)=f(x)g(y)$ と書けるとき, 和 $S=X+Y$ の確率密度関数 $f_{\mathrm{S}}(s)$ は,

$$f_{\mathrm{S}}(s)=\int_{-\infty}^{+\infty} f(x)g(s-x)\mathrm{d}x.$$

この $f_{\mathrm{S}}(s)$ (や右辺の積分の形) を, 2 つの関数 $f(x), g(y)$ の**たたみ込み合成積 convolution** といい, $(f*g)(s)$ と書くことがある.

合成積 $(f*g)$ のフーリエ変換は, f, g のフーリエ変換の積　　フーリエ解析

L02-Q2

Quiz(2 つの連続型確率変数の和の分布)

連続型確率変数 X, Y は独立で, 同分布 $\mathrm{U}(1,7)$ にしたがう. すなわち確率密度関数は

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

和 $S = X + Y$ の確率密度関数を求めよう.

L02-Q3

Quiz(2 つの連続型確率変数の和の分布)

連続型確率変数 X, Y は同時分布

$$f(x,y) = \begin{cases} \frac{1}{18} & (0 \leq x < 6, 0 \leq y < 6 - x) \\ 0 & (\text{他}) \end{cases}$$

にしたがう.
和 $S = X + Y$ の確率密度関数を求めよう.

母期待値が, X, Y で計算したときと, S で計算したときで一致するか, 検算しよう.

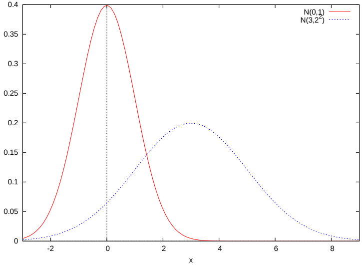
復習:(1 次元の) 正規分布

(一般の) 正規分布 $\mathrm{N}(b, a^2)$ の確率密度関数

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

b, a^2:パラメタ ($a > 0$).

https://www.geogebra.org/#probability
https://ja.wolframalpha.com ⇝ 正規分布



⇝ 正規分布

2 次元正規分布 (のうち X, Y が独立な簡単なケース)

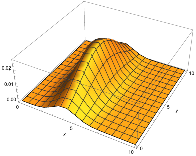
$$f_{\mathrm{XY}}(x,y) = \frac{1}{\sqrt{2\pi\sigma_{\mathrm{X}}^2}} \mathrm{e}^{-\frac{(x-\mu_{\mathrm{X}})^2}{2\sigma_{\mathrm{X}}^2}} \times \frac{1}{\sqrt{2\pi\sigma_{\mathrm{Y}}^2}} \mathrm{e}^{-\frac{(y-\mu_{\mathrm{Y}})^2}{2\sigma_{\mathrm{Y}}^2}}.$$

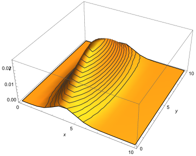
$$\mathrm{E}[X] = \mu_{\mathrm{X}}, \quad \mathrm{E}[Y] = \mu_{\mathrm{Y}}, \quad \mathrm{V}[X] = \sigma_{\mathrm{X}}^2, \quad \mathrm{V}[Y] = \sigma_{\mathrm{Y}}^2,$$

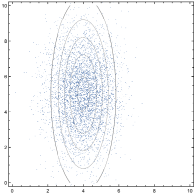
$$\text{母共分散 } C_{\mathrm{XY}} = (\text{独立なので}) = 0,$$

$$\text{母相関係数 } r_{\mathrm{XY}} = \frac{\mathrm{Cov}[X,Y]}{\sqrt{\mathrm{V}[X]}\sqrt{\mathrm{V}[Y]}} = 0$$

2 変量データ 確率統計☆演習 I(2015)L04　　同時分布 確率統計☆演習 I(2015)L09







等　高　線　は $(\mu_{\mathrm{X}}, \mu_{\mathrm{Y}})$ を中心

この分布の周辺分布は? 独立だから…

L02-Q4

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-Q5

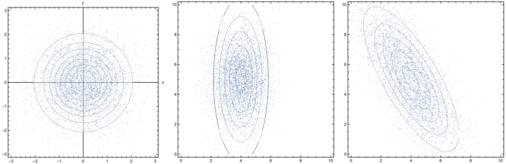
Quiz(2 次元正規分布)

次の 2 変数確率密度関数は 2 次元正規分布を定める.

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

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.

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

$$D = x^2 + y^2 = \begin{pmatrix} x & y \end{pmatrix} \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix}^{-1} \begin{pmatrix} x \\ y \end{pmatrix}$$

から出発して, 回転と拡大縮小する.

$$D = \begin{pmatrix} x' & y' \end{pmatrix} \begin{pmatrix} \cos\theta & -\sin\theta \\ \sin\theta & \cos\theta \end{pmatrix} \begin{pmatrix} 1/\sigma_{\mathrm{X}} & 0 \\ 0 & 1/\sigma_{\mathrm{Y}} \end{pmatrix} \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix}^{-1} \times \begin{pmatrix} 1/\sigma_{\mathrm{X}} & 0 \\ 0 & 1/\sigma_{\mathrm{Y}} \end{pmatrix} \begin{pmatrix} \cos\theta & \sin\theta \\ -\sin\theta & \cos\theta \end{pmatrix} \begin{pmatrix} x' \\ y' \end{pmatrix}$$

考え方 2 指数関数の肩に xy の項も入れておけばいい.

2 次元正規分布

2 次元正規分布の確率密度関数

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

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

ただし, パラメタは,

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

$$\text{母共分散行列}\Sigma = \begin{pmatrix} \sigma_{\mathrm{X}}^2 & C_{\mathrm{XY}} \\ C_{\mathrm{XY}} & \sigma_{\mathrm{Y}}^2 \end{pmatrix} \left(= \text{実は} \begin{pmatrix} \mathrm{V}[X] & \mathrm{Cov}[X, Y] \\ \mathrm{Cov}[X, Y] & \mathrm{V}[Y] \end{pmatrix}\right)$$

行列 V は大文字だが, 確率変数ではなく, パラメタ. σ^2 が 2 次元になったもの. V は対称行列での固有値は 2 個とも正.

パラメタの V の成分が母 (共) 分散になることはわりに簡単に証明で

Example (2 次元正規分布の確率密度関数)

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

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

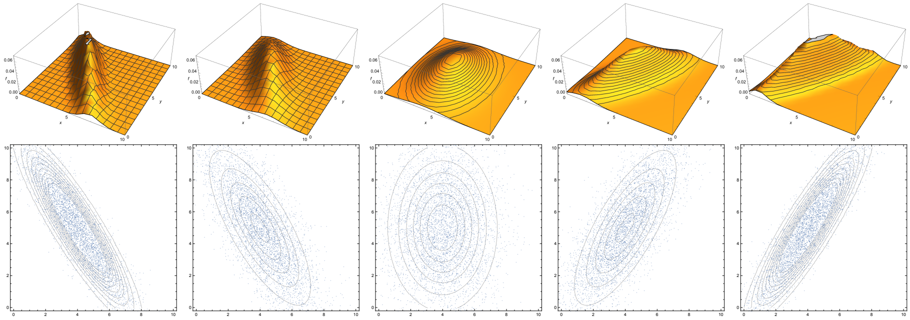
のとき, f を (行列やベクトルを使わずに) 具体的に書こう.

超便利なファミリー: 条件付き分布, 周辺分布は 1 次元正規分布

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

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

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



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.

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

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.