条件付き確率, 条件付き母期待値

樋口さぶろお https://hig3.net

龍谷大学大学院 理工学研究科 数理情報学専攻

理論物理学特論　L03 (2022-10-05 Wed)

最終更新: Time-stamp: "2022-10-07 Fri 12:40 JST hig"

今日の目標

- 2 次元正規分布を例に条件付き確率密度関数, 条件付き母期待値を計算できる
- 混合ガウス分布の定義を説明できる

略解 I

L02-Q1
Quiz 解答:2 次元正規分布

1. 指数関数の引数を平方完成する.

$$
\begin{aligned}
& -x^2-4x-2y^2+12y-5 \\
=& -\frac{1}{2}(2x^2+8x+4y^2-24y+10) \\
=& -\frac{1}{2}\left(\frac{(x+2)^2}{(1/\sqrt{2})^2}-8+\frac{(y-3)^2}{(1/2)^2}-36+10\right)
\end{aligned}
$$

より, $\mathrm{E}[X]=-2, \mathrm{E}[Y]=3, \mathrm{V}[X]=1/2, \mathrm{V}[Y]=1/4$. $f(x,y)$ は x 部分と y 部分の積だから X,Y は独立であり, $\mathrm{Cov}[X,Y]=0$.

略解 II

2. $f(x,y)$ は次の 2 つの形に書ける.

$$C{\cdot}\mathrm{e}^{-\frac{(x+2)^2}{2(1/\sqrt{2})^2}-\frac{(y-3)^2}{2(1/2)^2}}\mathrm{e}^{-\frac{1}{2}(-34)} = \frac{1}{\sqrt{2\pi(1/2)}}\frac{1}{\sqrt{2\pi(1/4)}}\mathrm{e}^{-\frac{(x+2)^2}{2(1/\sqrt{2})^2}-\frac{(y-3)^2}{2(1/2)^2}}$$

よって, $C = \frac{1}{\sqrt{2\pi(1/2)}}\frac{1}{\sqrt{2\pi(1/4)}}\mathrm{e}^{-17}$.

L02-Q2

Quiz 解答:2 次元正規分布

略解 III

1. 指数関数の引数を平方完成する.

$$
\begin{aligned}
&-4x^2 - \tfrac{1}{6}y^2 + 2y \\
=& -\frac{1}{2}(8x^2 + \tfrac{1}{3}y^2 - 4y) \\
=& -\frac{1}{2}\left(\frac{(x-0)^2}{(1/2\sqrt{2})^2} + \frac{(y-6)^2}{(\sqrt{3})^2} - 12\right)
\end{aligned}
$$

より, $\mathrm{E}[X] = 0, \mathrm{E}[Y] = 6, \mathrm{V}[X] = \frac{1}{8}, \mathrm{V}[Y] = 3, \mathrm{Cov}[X, Y] = 0.$

2. $f(x, y)$ は次の 2 つの形に書ける.

$$
C \cdot \mathrm{e}^{-\frac{(x-0)^2}{2(1/2\sqrt{2})^2} - \frac{(y-6)^2}{2(\sqrt{3})^2}} \mathrm{e}^{-\frac{1}{2}(-12)} = \frac{1}{\sqrt{2\pi(1/8)}} \frac{1}{\sqrt{2\pi \cdot 3}} \mathrm{e}^{-\frac{(x-0)^2}{2(1/2\sqrt{2})^2} - \frac{(y-6)^2}{2(\sqrt{3})^2}}
$$

であるべきだから, $C = \frac{1}{\sqrt{2\pi(1/8)}} \frac{1}{\sqrt{2\pi \cdot 3}} \mathrm{e}^{-6}$.

略解 IV

L02-Q3

Quiz 解答:2 次元正規分布の確率密度関数

1. $\boldsymbol{\mu} = \begin{pmatrix} -4 \\ 5 \end{pmatrix}$,
   $\Sigma = \begin{pmatrix} 2 & -1 \\ -1 & 3 \end{pmatrix}$
2. $\det \Sigma = 5$, $\Sigma^{-1} = \frac{1}{5}\begin{pmatrix} 3 & 1 \\ 1 & 2 \end{pmatrix}$ より,

$$f(x, y) = \frac{1}{(2\pi)^{2/2} \cdot 5} \mathrm{e}^{-\frac{1}{2}\frac{3}{5}(x+4)^2 - \frac{1}{2}\cdot 2 \cdot \frac{2}{5}(x+4)(y-5) - \frac{1}{2}\frac{2}{5}(y-5)^2}$$

L02-Q4

Quiz 解答:2 次元正規分布

略解 V

1. $\mathrm{E}[X] = 2.$
   $\mathrm{E}[Y] = -3.$

$$\Sigma = (\Sigma^{-1})^{-1} = \begin{pmatrix} 4 & -6 \\ -6 & 14 \end{pmatrix}^{-1} = \frac{1}{20}\begin{pmatrix} 14 & +6 \\ +6 & 4 \end{pmatrix}$$

   $\mathrm{V}[X] = \frac{14}{20}, \mathrm{Cov}[X, Y] = \frac{6}{20}, \mathrm{V}[Y] = \frac{4}{20}.$
2. $C = \dfrac{1}{\sqrt{(2\pi)^2 \det \Sigma}}.$

L02-Q5
Quiz 解答:2 次元正規分布

略解 VI

1. 指数関数の引数の 2 次式を対称行列を用いて書く. x, y の 1 次の項はないので, X, Y の母平均値は 0 であり,

$$-x^2 + 4xy - 7y^2 = -\frac{1}{2}\begin{pmatrix} x-0 & y-0 \end{pmatrix}\begin{pmatrix} 2 & -4 \\ -4 & 14 \end{pmatrix}\begin{pmatrix} x-0 \\ y-0 \end{pmatrix}$$

よって,

$$\Sigma^{-1} = \begin{pmatrix} 2 & -4 \\ -4 & 14 \end{pmatrix}, \quad \Sigma = \frac{1}{2 \cdot 14 - (-4)(-4)}\begin{pmatrix} 14 & 4 \\ 4 & 2 \end{pmatrix}.$$

2. $f(x, y) = C\mathrm{e}^{-x^2+4xy-7y^2} = \frac{1}{2\pi\sqrt{\det \Sigma}}\mathrm{e}^{-\frac{1}{2}{}^t\boldsymbol{x}\Sigma^{-1}\boldsymbol{x}}$. よって $C = \frac{1}{2\pi\sqrt{1/12}}$.

略解 VII

L02-Q6

Quiz 解答:2 次元正規分布の周辺分布

1.

$$
\begin{aligned}
f_X(x) &= \int_{-\infty}^{+\infty} f(x, y)\,\mathrm{d}y \\
&= C \int_{-\infty}^{+\infty} \mathrm{e}^{-x^2 - 7(y - \frac{2}{7}x)^2 + \frac{4}{7}x^2}\,\mathrm{d}y \\
&= C'' \mathrm{e}^{-\frac{3}{7}x^2} \\
&= \frac{1}{\sqrt{2\pi(7/6)}} \mathrm{e}^{-\frac{3}{7}x^2}.
\end{aligned}
$$

略解 VIII

2.

$$
\begin{aligned}
f_{\mathrm{Y}}(y) = &\int_{-\infty}^{+\infty} f(x, y)\,\mathrm{d}x \\
=&C \int_{-\infty}^{+\infty} \mathrm{e}^{-(x-2y)^2-7y^2+4y^2} \\
=&C'' \mathrm{e}^{-3y^2} \\
=&\frac{1}{\sqrt{2\pi(1/6)}} \mathrm{e}^{-3y^2}.
\end{aligned}
$$

ここまで来たよ

条件付き確率 岩薩林 確率・統計 §2.3

定義 (条件付き確率 岩薩林 確率・統計 (2.12))

条件 (事象) A のもとでの, 事象 B の条件付き確率 'P of A given B'

$$P(B|A) = \frac{P(A \cap B)}{P(A)}.$$

縦棒「|」の前後は不平等: P(確率を考える事象|条件の事象).
$P(B|A) \neq P(A|B), \neq P(A \cup B)$.
意味 A が起きる場合に限定した, B が起きる確率.
縦棒の前だけ見る (縦棒の後ろを固定する) と, ただの確率.

命題 (乗法公式 岩薩林 確率・統計 (2.13))

$$P(A \cap B) = P(B|A) \times P(A).$$

連続型確率変数の条件付き分布

定義 (条件付き (分布の) 確率密度関数)

$f(x,y)$: 連続型確率変数 X, Y の同時確率密度関数.
$Y=y$ という条件のもとでの X の確率密度関数 $f_{\mathrm{X|Y}}(x|y)$ は,

$$f_{\mathrm{X|Y}}(x|y) = \frac{f(x,y)}{f_{\mathrm{Y}}(y)}$$

$|$ の前後は異質. y はパラメタ, x は確率変数の値.
分母は, $\mathrm{E}[X^0]=1$ にするための定数倍の調整.
意味: 同時確率密度関数のグラフを, $Y=y$ で切った断面の面積を 1 と思ったときの, 断面に見える X の確率密度関数.
もちろん次が成立することを示せる.

$$\forall y \quad \int_{-\infty}^{+\infty} f_{\mathrm{X|Y}}(x|y)\,\mathrm{d}x = 1.$$

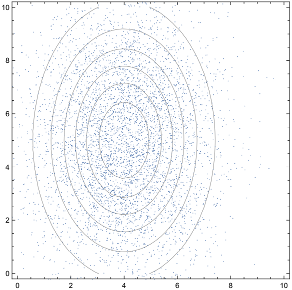
命題 (乗法公式相当)

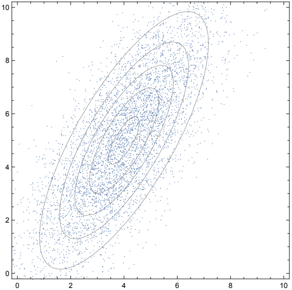
$$f(x,y) = f(x|y) \times f_{\mathrm{Y}}(y)$$

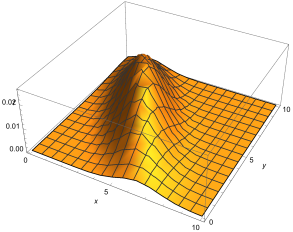
命題 (全確率の法則相当)

$$f_{\mathrm{X}}(x) = \int_{-\infty}^{+\infty} f_{\mathrm{X|Y}}(x|y) \times f_{\mathrm{Y}}(y)\mathrm{d}y$$

X の周辺分布と Y の周辺分布との関係を表す式.


0
2
4
6
8
10


0
2
4
6
8
10


0.02
0.01
0.00
0
5
x
10
0
5
y
10

条件付き母期待値

定義 (条件付き母期待値)

$Y = y$ という条件のもとでの $g(X)$ の母期待値

$$\mathrm{E}[g(X)|Y = y] = \int_{-\infty}^{+\infty} g(x) f_{\mathrm{X}|\mathrm{Y}}(x|y)\mathrm{d}x$$

条件付き母分散

$\mathrm{V}[X|Y = y]$ などが同様に定義される.

ここまで来たよ

2 次元正規分布の条件付き分布

命題 (2 次元正規分布の条件付き分布)

X, Y が (独立とはかぎらない)2 次元正規分布にしたがうとき,
$Y = y$ という条件のもとでの X の条件付き分布は 1 次元正規分布.

独立なとき

$$f_{\mathrm{X|Y}}(x|y) = \frac{C\mathrm{e}^{-\frac{1}{2\sigma_{\mathrm{X}}^2}(x-\mu_{\mathrm{X}})^2} \times \mathrm{e}^{-\frac{1}{2\sigma_{\mathrm{Y}}^2}(y-\mu_{\mathrm{Y}})^2}}{C'\mathrm{e}^{-\frac{1}{2\sigma_{\mathrm{Y}}^2}(y-\mu_{\mathrm{Y}})^2}}$$
$$= C''\mathrm{e}^{-\frac{1}{2\sigma_{\mathrm{X}}^2}(x-\mu_{\mathrm{X}})^2} = f_{\mathrm{X}}(x).$$

命題 (2 つの独立な確率変数の条件付き分布)

一般に, 独立なとき, 周辺分布=条件付き分布.
$f(x,y) = f_{\mathrm{X}}(x) \times f_{\mathrm{Y}}(y)$. $f_{\mathrm{X|Y}}(x|y) = f_{\mathrm{X}}(x)$. $f_{\mathrm{Y|X}}(y|x) = f_{\mathrm{Y}}(y)$.

独立でないとき

簡単のため $f(x,y) = C\mathrm{e}^{-ax^2-by^2+cxy}$ とする.
気分を出すため, $f(x|y_0)$ を考える.
$f_\mathrm{Y}(y) =$ 先週 $= C'\mathrm{e}^{-b'y^2}$ より,

$$
\begin{aligned}
f_{\mathrm{X}|\mathrm{Y}}(x|y_0) =& \frac{f(x,y_0)}{\text{周辺分布の } y_0 \text{ での値}} \\
=& C'(y_0)\mathrm{e}^{-ax^2-by_0^2+cxy_0} \\
=& C''(y_0)\mathrm{e}^{-ax^2+cxy_0} \\
=& C'''(y_0)\mathrm{e}^{-a(x-\frac{c}{2a}y_0)^2}.
\end{aligned}
$$

すなわち, $f(x,y_0)$ を定数倍で調整したもの.
X の母平均値 $-\frac{c}{2a}y_0$, 母分散 $\frac{1}{2a}$ の正規分布.
$\mathrm{E}[X|Y=y] = -\frac{c}{2a}y$.
母平均値のみが y_0 による.

L03-Q1

Quiz(2 次元正規分布の条件付き)

次の 2 変数確率密度関数は (X, Y) の 2 次元正規分布を定める.

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

1. $X = 2$ という条件のもとでの Y の条件付き分布を求めよう.
2. $X = 2$ という条件のもとでの Y^1 の母期待値を求めよう.
3. $Y = -1$ という条件のもとでの X の条件付き分布を求めよう.
4. $Y = -1$ という条件のもとでの X^2 の母期待値を求めよう.

ここまで来たよ

離散, 連続型確率変数の同時分布

X: 連続型, Y: 離散型確率変数 の多次元分布 岩薩林 確率・統計 §3.3
同時分布 $f(x,y)$. x については確率密度, y について確率.

$$\mathrm{E}[g(X,Y)] = \int_{-\infty}^{+\infty} \sum_y g(x,y) f(x,y)\,\mathrm{d}x.$$

$$P(c \le X < d, Y = y_0) = \int_{-\infty}^{+\infty} \sum_y \mathrm{I}_{[c \le X < d, Y = y_0]}(x,y) f(x,y) \mathrm{d}x = \int_c^d f(x,y_0) \mathrm{d}x$$

気分を出すために, 表や場合分けで $f(x,y)$. ただし, $Y = 0, 1$. x についての確率密度関数 $h_0(x), h_1(x)$.

$$f(x,y) = \begin{cases} h_0(x) & (y=0) \\ h_1(x) & (y=1) \\ 0 & (y \neq 0,1) \end{cases}$$

$y \backslash x$	$-\infty < x < +\infty$
0	$h_0(x)$
1	$h_1(x)$

全 $= \int_{-\infty}^{+\infty} h_0(x)\,\mathrm{d}x + \int_{-\infty}^{+\infty} h_1(x)\,\mathrm{d}x = 1.$

混合ガウス分布 GMM=Gaussian mixture model

正規分布 normal distribution = ガウス分布 Gaussian distribution

岩薩林 確率・統計 §4.5

混合する = mix

混合ガウス分布を定義する準備として, $\pi_0+\pi_1=1, \pi_y \geq 0, \pi_{\text{他}}=0$ として, 次の同時分布 $f(x,y)$ を考える.

$$f(x,y)=\pi_y\frac{1}{(2\pi\sigma_y^2)^{1/2}}\mathrm{e}^{-\frac{(x-\mu_y)^2}{2\sigma_y^2}}=\begin{cases}\pi_0\cdot\frac{1}{(2\pi\sigma_0^2)^{1/2}}\mathrm{e}^{-\frac{(x-\mu_0)^2}{2\sigma_0^2}} & (y=0)\\ \pi_1\cdot\frac{1}{(2\pi\sigma_1^2)^{1/2}}\mathrm{e}^{-\frac{(x-\mu_1)^2}{2\sigma_1^2}} & (y=1)\\ 0 & (y\neq 0,1)\end{cases}$$

X,Y は独立ではない ($\sigma_0=\sigma_1, \mu_0=\mu_1$ でない限り).

Y の周辺分布は離散型でベルヌイ分布 $B(1,\pi_1)$ 岩薩林 確率・統計 §3.4

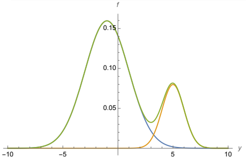
$$f_Y(y) = \int f(x,y)\mathrm{d}x = \int_{-\infty}^{+\infty} \pi_y \cdot \frac{1}{(2\pi\sigma_y^2)^{1/2}} \mathrm{e}^{-\frac{(x-\mu_y)^2}{2\sigma_y^2}} \mathrm{d}x = \pi_y.$$

X の周辺分布は連続型で GMM

定義 (GMM=混合ガウス分布)

連続型確率変数 X がパラメタ $(\pi_0, \pi_1, \mu_0, \mu_1, \sigma_0^2, \sigma_1^2)$ の GMM に従うとは, 次の確率密度関数にしたがうことをいう.

$$f_X(x) = \sum_y f(x,y) = \pi_0 \cdot \frac{1}{(2\pi\sigma_0^2)^{1/2}} \mathrm{e}^{-\frac{(x-\mu_0)^2}{2\sigma_0^2}} + \pi_1 \cdot \frac{1}{(2\pi\sigma_1^2)^{1/2}} \mathrm{e}^{-\frac{(x-\mu_1)^2}{2\sigma_1^2}}.$$



therophys-L03-gmm.ipynb

一般の混合ガウス分布: y の値が 2 個と限らず有限個
scipy.stats には GMM はないので, scipy.stats.norm を使って自分で計算する.
sklearn.mixture.GaussianMixture
一般の混合分布: ガウス分布とは限らない分布を重み π_y で「重ね合わせ」
注: 混合 $\pi_0 f_0(x) + \pi_1 f_1(x)$ と, 確率関数の和 $S = X + Y$ はもちろん異なる.

L03-Q2

Quiz(混合ガウス分布の確率密度関数)

X は混合ガウス分布 $(\pi_0 = \frac{3}{10}, \pi_1 = \frac{7}{10}, \mu_0 = 2, \mu_1 = 6, \sigma_0 = 2, \sigma_1 = \frac{1}{2})$ にしたがう．確率密度関数 $f(x)$ のグラフの概形を描こう．
手で, または, scipy.stats.norm の浮動小数点計算で．

L03-Q3

Quiz(混合ガウス分布の確率)

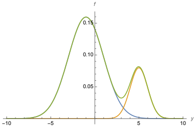
X は混合ガウス分布 $(\pi_0 = \frac{3}{10}, \pi_1 = \frac{7}{10}, \mu_0 = 2, \mu_1 = 6, \sigma_0 = 2, \sigma_1 = \frac{1}{2})$ にしたがう.

1. X の確率密度関数の $x = 4$ における値を求めよう.
2. 確率 $P(X \leq 4)$ を, 標準正規分布の累積分布関数 $F(x)$ で表そう.

さらに scipy.stats.norm で浮動小数点数で求めてみると安心できるかも.

混合ガウス分布のモーメント，母期待値を周辺分布から

$$
\begin{aligned}
\mathrm{E}[X^k] &= \int \sum_y x^k \pi_y \cdot \frac{1}{(2\pi\sigma_y^2)^{1/2}} \mathrm{e}^{-\frac{(x-\mu_y)^2}{2\sigma_y^2}} \mathrm{d}x \\
&= \pi_0 \int x^k \frac{1}{(2\pi\sigma_0^2)^{1/2}} \mathrm{e}^{-\frac{(x-\mu_0)^2}{2\sigma_0^2}} \mathrm{d}x + \pi_1 \int x^k \frac{1}{(2\pi\sigma_1^2)^{1/2}} \mathrm{e}^{-\frac{(x-\mu_1)^2}{2\sigma_1^2}} \mathrm{d}x. \\
\mathrm{E}[X^0] &= \pi_0 1 + \pi_1 1 = 1. \\
\mathrm{E}[X^1] &= \pi_0 \mu_0 + \pi_1 \mu_1. \\
\mathrm{V}[X] &= \text{着実な計算} = \pi_0 \sigma_0^2 + \pi_1 \sigma_1^2 + \pi_0 \pi_1 (\mu_0 - \mu_1)^2. \\
\mathrm{E}[Y^k] &= \text{ベルヌイ分布} = \pi_1 \quad (k = 1, 2, 3, \ldots).
\end{aligned}
$$



混合ガウス分布と関係する条件付き分布

一般に, $Y=y_0$ という条件のもとでの $X=x$ の条件付き確率 岩薩林 確率・統計 p.59

$$P(X=x|Y=y_0)=f_{\mathrm{X|Y}}(x|y_0)=\frac{f(x,y_0)}{\sum_{x'}f(x',y_0)}$$

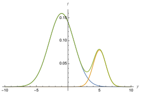
以下, 略記 $f(x;\mu,\sigma^2)=\frac{1}{(2\pi\sigma^2)^{1/2}}\mathrm{e}^{-\frac{(x-\mu)^2}{2\sigma^2}}$.

$X=x_0$ であるという条件のもとでの y の条件付き確率

$$P(Y=1|X=x_0)=f_{\mathrm{Y|X}}(1|x_0)=\frac{f(x_0,1)}{\sum_{y'}f(x_0,y')}$$
$$=\frac{\pi_1 f(x_0;\mu_1,\sigma_1^2)}{\pi_0 f(x_0;\mu_0,\sigma_0^2)+\pi_1 f(x_0;\mu_1,\sigma_1^2)}$$

$Y=y$ であるという条件のもとでの x の条件付き確率密度

$$p(X=x|Y=y)=\frac{\pi_y f(x;\mu_y,\sigma_y^2)}{\pi_y\int_{-\infty}^{+\infty}f(x';\mu_y,\sigma_y^2)\mathrm{d}x'}=f(x;\mu_y,\sigma_y^2)$$



混合ガウス分布の標本抽出のアルゴリズム

(π_y, μ_y, σ_y): 既知とする (仏の立場).

道具 1 コイン (二項分布) scipy.stats.binom(n=1,p=π_1).rvs(size=1)
道具 2 正規分布連続サイコロ
scipy.stats.norm(loc=μ_y,scale=σ_y).rvs(size=1)

アイデア
$f_{\mathrm{X}}(x)$ は難しい.
$f(x,y) = f_{\mathrm{X}|\mathrm{Y}}(x|y) \cdot f_{\mathrm{Y}}(y)$ で (x,y) を得た後, 周辺分布を作る (y を無視する).
アルゴリズム

- n 回繰り返す.
  - コインを投げて $y = 0, 1$ を決定
  - 次に μ_y, σ_y に調節したサイコロを投げて x を得る
  - 組み合わせた (x, y) がひとつのデータ.

2 次元の混合ガウス分布

2 次元の混合ガウス分布

x が 2 次元ベクトル $\boldsymbol{x} = {}^{\mathrm{t}}(x_1, x_2)$ になっただけ.

$$f(\boldsymbol{x}, y) = \begin{cases} \dfrac{1}{(2\pi)^{2/2}(\det\Sigma_0)^{1/2}} \mathrm{e}^{-\frac{1}{2}\,{}^{\mathrm{t}}(\boldsymbol{x}-\boldsymbol{\mu}_0)\Sigma_0^{-1}(\boldsymbol{x}-\boldsymbol{\mu}_0)} \cdot \pi_0 & (y=0) \\ \dfrac{1}{(2\pi)^{2/2}(\det\Sigma_1)^{1/2}} \mathrm{e}^{-\frac{1}{2}\,{}^{\mathrm{t}}(\boldsymbol{x}-\boldsymbol{\mu}_1)\Sigma_1^{-1}(\boldsymbol{x}-\boldsymbol{\mu}_1)} \cdot \pi_1 & (y=1) \\ 0 & (y\text{が他}) \end{cases}$$