最適化理論

—連続型確率変数—

今回学ぶこと

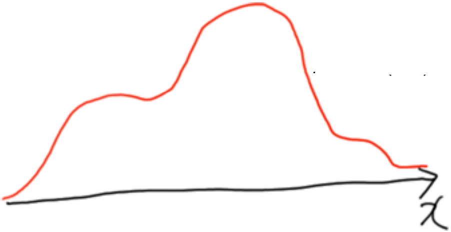
- 連続型確率変数
- 確率密度関数の復習
- 代表的な連続な確率密度関数について
- 連続型確率変数における確率推論

連続型確率変数

連続型確率変数 $=$ 実数値を実現値とする確率変数

- X: **確率変数**
- $D(X) = \mathbb{R}$: **ここで $\mathbb{R}$ は実数値全体の集合**
- $P_X(x)(x \in \mathbb{R})$: **確率密度関数**

確率密度関数の性質



- 確率密度関数 $P_X(x)$ は、確率変数 X に確率的振る舞いを表す一変数実数値関数である。
- "$P_X(x)$ の値 $=$ 確率" ではないことに注意！
- 確率密度関数 $P_X(x)$ は次の性質を満たす：

$$\forall x \in \mathbb{R}, \quad P_X(x) \geq 0$$

$$\int_{-\infty}^{\infty} P_X(x) dx = 1$$

確率密度関数と確率

X が区間 $[a, b]$ に含まれる確率

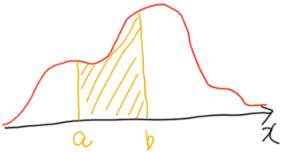
閉区間 $[a, b]$ **は**

$$[a, b] = \{x \in \mathbb{R} \mid a \leq x \leq b\}$$

と定義される (a, b **は** $a \leq b$ **を満たす実数)。確率変数** X **が閉区間** $[a, b]$ **の範囲に実現値を持つ確率を** $Prob[X \in [a, b]]$ **と書くと**

$$Prob[X \in [a, b]] = \int_a^b P_X(x)dx$$

となる。



期待値: 平均と分散

期待値

実数値関数 f の期待値

$$E[f(X)] = \int_{-\infty}^{\infty} f(x) P_X(x) dx$$

- **平均** ($f(x) = x$ **とする**)

$$m \triangleq E[X] = \int_{-\infty}^{\infty} x P_X(x) dx$$

- **分散** ($f(x) = (x - m)^2$ **とする**)

$$\sigma^2 \triangleq E[(X - m)^2] = \int_{-\infty}^{\infty} (x - m)^2 P_X(x) dx$$

ガウス確率変数

特に重要な連続確率変数として、ガウス確率変数が挙げられる。

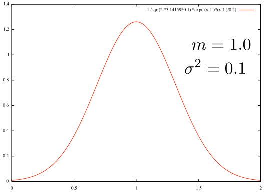
- 自然界にしばしば現れる、普遍的（中心極限定理との関係)
- 対数を取ると二次関数となり、取り扱いやすい

ガウス確率密度関数 (1 変数)

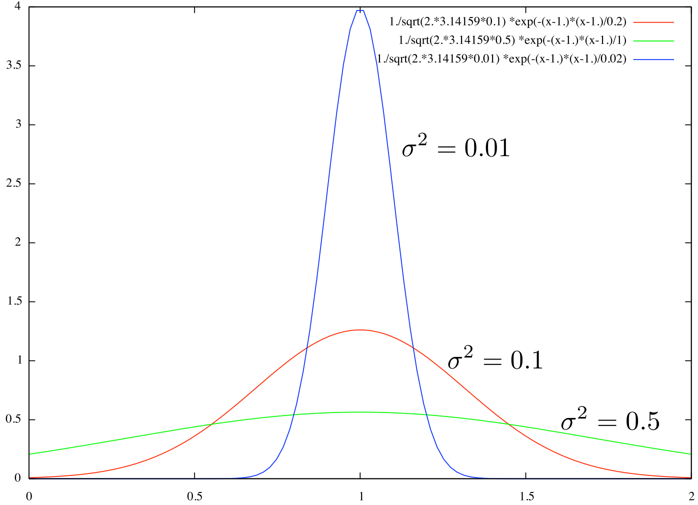
- μ:**平均**, σ^2:**分散**

$$P_X(x) = \frac{1}{\sqrt{2\pi\sigma^2}} \exp\left(-\frac{(x-\mu)^2}{2\sigma^2}\right)$$

確率変数 X **がこのガウス分布に従う場合、**$X \sim \mathcal{N}(\mu, \sigma^2)$ **と書く。**



分散の異なるガウス分布



同時確率密度関数 (joint PDF)

与えられた 2 つの実数値確率変数 X, Y について、同時 PDF を

$$P_{XY}(x, y)$$

と書く。3 変数以上の場合も同様。次の 2 つの条件を満たす必要がある。

$$\forall (x, y) \in \mathbb{R}^2,\ P_{XY}(x, y) \geq 0$$

$$\int_{-\infty}^{\infty} \int_{-\infty}^{\infty} P_{XY}(x, y) dx dy = 1$$

同時確率密度関数と確率 (1)

与えられた $a_1, a_2, b_1, b_2 \in \mathbb{R}(a_1 \leq a_2, b_1 \leq b_2)$ **に対して閉長方形** C **を**

$$C = \{(x, y) \in \mathbb{R}^2 \mid a_1 \leq x \leq a_2, b_1 \leq y \leq b_2\}$$

とする。

(X, Y) が C に含まれる確率

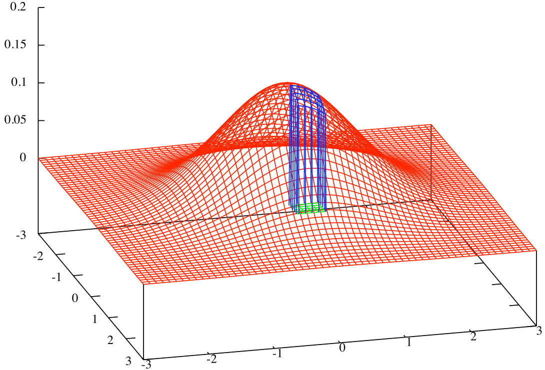
(X, Y) **の組の実現値が長方形** C **に含まれる確率** $Prob[(X, Y) \in C]$ **は**

$$Prob[(X, Y) \in C] = \int_{b_1}^{b_2} \int_{a_1}^{a_2} P_{XY}(x, y) dxdy$$

と与えられる。

同時確率密度関数と確率 (2)

$$Prob[(X, Y) \in C] = \int_{b_1}^{b_2} \int_{a_1}^{a_2} P_{XY}(x, y) dxdy$$



周辺化・周辺確率密度関数

$P_{XY}(x, y)$: **同時確率密度関数**

周辺化

$$P_X(x) = \int_{-\infty}^{\infty} P_{XY}(x, y) dy$$

$$P_Y(y) = \int_{-\infty}^{\infty} P_{XY}(x, y) dx$$

条件付確率密度関数

条件付確率密度関数の定義

$$P_{Y|X}(y|x) = \frac{P_{XY}(x, y)}{P_X(x)}$$

任意の $x \in \mathbb{R}$ に対して、 $P_X(x) > 0$ でないと定義できないことに注意。

さまざまな確率密度関数：一様分布

一様分布

一様分布は、区間 $[\alpha, \beta]$ において、一定値 $t > 0$ を値として持つ確率密度関数である。すなわち、X をそのような一様分布に従う確率変数とすると

$$P_X(x) = \begin{cases} t & \alpha \leq x \leq \beta \\ 0 & \text{otherwise} \end{cases}$$

となる。

問題
上記の一様分布を t の値を α, β で書け。

さまざまな確率密度関数：指数分布

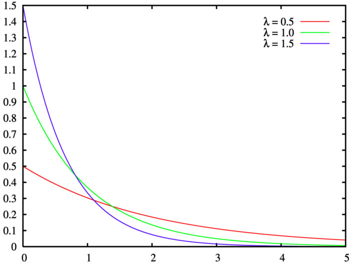
指数分布

確率密度関数

$$P_X(x) = \begin{cases} \lambda e^{-\lambda x} & x \geq 0 \\ 0 & x < 0 \end{cases}$$

を指数分布と呼ぶ。ここで、パラメータ λ は正の実数である。

まれにしか生じない事象 (ポアソン分布に従う) の生起間隔をモデル化する場合などに利用される。



平均ベクトルと共分散行列

- n 次元ベクトル型の確率変数

$$X = (X_1, X_2, \ldots, X_n)^T$$

- $D(X) = \mathbb{R}^n$
- $\mu = (\mu_1, \mu_2, \ldots, \mu_n) \in \mathbb{R}^n$: **平均値ベクトル**
- $S \in \mathbb{R}^{n \times n}$: **共分散行列**

平均値ベクトル

$$\mu_i = E[X_i], \quad i = 1, 2, \ldots, n$$

共分散行列 $S = \{S_{i,j}\}$

$$S_{i,j} = E[(X_i - \mu_i)(X_j - \mu_j)]$$

さまざまな確率密度関数：多次元ガウス分布

- n 次元ベクトル型の確率変数

$$X = (X_1, X_2, \ldots, X_n)^T$$

- $D(X) = \mathbb{R}^n$
- $\mu = (\mu_1, \mu_2, \ldots, \mu_n) \in \mathbb{R}^n$: **平均値ベクトル**
- $S \in \mathbb{R}^{n \times n}$: **共分散行列**

多次元ガウス分布

次の確率密度関数を多次元ガウス分布と呼ぶ:

$$P_X(x) = \frac{1}{(\sqrt{2\pi})^n \sqrt{|S|}} \exp\left(-\frac{1}{2}(x-\mu)^T S^{-1}(x-\mu)\right)$$

ここで $x = (x_1, x_2, \ldots, x_n)^T$ **である。**

確率変数 X **がこのガウス分布に従う場合、** $X \sim \mathcal{N}(\mu, S)$ **と書く。**

連続型確率変数における事後確率の計算

- 基本的な考え方は、離散確率変数の場合と同様。
- 周辺化: 和 → 積分

同時確率密度関数が既知のとき

同時確率密度関数が既知の場合:

- $P_{XY}(x, y)$ **が既知。**
- Y **の実現値** y^* **が得られている。**

の事後確率の計算は次のようになる。

事後確率密度関数

$$P_{X|Y}(x|y^*) = \frac{P_{XY}(x, y^*)}{P_Y(y^*)} = \frac{P_{XY}(x, y^*)}{\int_{-\infty}^{\infty} P_{XY}(x, y^*)dx}$$

と計算できる。

事前確率密度関数と条件付確率密度関数が既知のとき

次に事前確率密度関数と条件付確率密度関数が既知の場合:

- $P_X(x)$ **が既知。**
- $P_{Y|X}(y|x)$ **が既知。**
- Y **の実現値** y^* **が得られている。**

このとき、観測できない確率変数 X **に関する事後確率は次のように与えられる。**

事後確率密度関数

$$
\begin{aligned}
P_{X|Y}(x|y^*) &= \frac{P_{Y|X}(y^*|x)P_X(x)}{P_Y(y^*)} \\
&= \frac{P_{Y|X}(y^*|x)P_X(x)}{\int_{-\infty}^{\infty} P_{Y|X}(y^*|x)P_X(x)dx}
\end{aligned}
$$