演習問題 (4) 連続型確率変数について

連続型確率変数は、実現値の集合が非可算無限集合となる確率変数であり、その確率的振る舞いは確率密度関数により特徴づけられる。本講義では、$P_X(x)$ で確率変数 X の確率密度関数を表す。離散型変数の場合と同様に確率密度関数については次の条件が成立する。

> **確率密度関数の満たすべき条件**
> 確率密度関数 $P_X(x)$ について、
> $$\int_{-\infty}^{\infty} P_X(x)dx = 1$$
> かつ、任意の $x \in \mathbb{R}$ について $P_X(x) \geq 0$ が成り立つ。

上の 2 つの囲みを見て分かる通り、離散型における「和」は連続型における「積分」に対応していることに注意したい。

例を挙げる。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

$$P_X(x) = \begin{cases} t & \alpha \leq x \leq \beta \\ 0 & \text{otherwise} \end{cases} \tag{1}$$

となる。

> 問 1
> 上記の一様分布を t の値を α, β で書け。

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。

> 問 2
> 指数分布 $P_X(x)$ を $-\infty$ から ∞ の範囲で積分した場合、その結果は 1 になることを示せ。

ガウス分布: 最も重要な確率密度関数といっても過言ではない。1 次元ガウス分布の確率密度関数は次の通り与えられる：

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

期待値は $E[X] = m$ となり、分散は $Var[X] = \sigma^2$ となる。中心極限定理に現れる。(ガウス分布は様々な分野で必要となるので、関数の形をしっかり頭に刻みこんでおこう。)

連続型確率変数の場合も

- 同時確率密度関数
- 周辺確率密度関数

が離散型と同様に存在する。

> **周辺確率密度関数**
> 同時確率密度関数 $p_{XY}(x, y)$ が与えられているとき、X に関する周辺化は
> $$P_X(x) = \int_{-\infty}^{\infty} P_{XY}(x, y)dy$$
> となり、$p_X(x)$ は周辺確率密度関数と呼ばれる。

> 問 3
> 2 次元ガウス分布
> $$P_{XY}(x, y) = \frac{1}{2\pi} \exp\left(-\frac{x^2 + y^2}{2}\right) \tag{2}$$
> を周辺化し、$P_X(x)$ を求めよ。

確率密度関数の「意味」を確認しておこう。

> **区間に対応する確率**
> 確率密度関数は実数閉区間 $[a, b] = \{x \in \mathbb{R} : a \leq x \leq b\}$ に対応する確率を与えている。す

なわち、区間 $[a,b]$ に対応する確率を $Prob(X \in [a,b])$ と書くとき、

$$Prob(X \in [a,b]) = \int_a^b P_X(x)dx \qquad (3)$$

連続型確率変数における事後確率の計算

まずは簡単な場合として、系が連続型確率変数 X, Y の 2 変数から成る場合について考える。

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の事後確率の計算は次のようになる。

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

事後確率密度関数
この場合、事後確率 $P_{X|Y}(x|y^*)$ は

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

と計算できる。

上の場合の計算例を次に示す。

問 4　同時分布が 2 次元ガウス分布

$$P_{XY}(x,y) = \frac{1}{2\pi}\exp\left(-\frac{x^2+y^2}{2}\right)$$

で与えられる。Y の実現値 $y^* = 2$ が観測されたとき、事後確率密度関数 $P_{X|Y}(x|y^*)$ を計算せよ。

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を検討する。これも典型的な推定・推論問題である。ここでの仮定は

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

である。

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

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

問 5　事前確率密度関数が一様分布

$$P_M(m) = \begin{cases} 1/(2\alpha) & -\alpha \le m \le \alpha \\ 0 & \text{otherwise} \end{cases} \qquad (4)$$

条件付確率密度関数は次のガウス分布

$$P_{X|M}(x|m) = \frac{1}{\sqrt{2\pi}}\exp\left(-\frac{(x-m)^2}{2}\right) \qquad (5)$$

を仮定する。$\alpha = 5$ で $x^* = 2.3$ を観測したときの M に関する事後分布の概形を描け。