演習問題 2.01

ベルヌーイ分布（ 2.2 ）が次の性質を満たすことを確かめよ。

$$\sum_{x=0}^{1} p(\ x \mid \mu\ ) = 1$$

$$\cdots\ (\ 2.257\ )$$

$$\mathrm{E}[\ x\ ] = \mu$$

$$\cdots\ (\ 2.258\ )$$

$$\mathrm{var}[\ x\ ] = \mu\ (\ 1 - \mu\ )$$

$$\cdots\ (\ 2.259\ )$$

ベルヌーイ分布に従う二値確率変数 x のエントロピー $\mathrm{H}[\ x\ ]$ が

$$\mathrm{H}[\ x\ ] = -\mu \ln \mu - (\ 1 - \mu\ ) \ln(\ 1 - \mu\ )$$

$$\cdots\ (\ 2.260\ )$$

で与えられることを示せ。

［ 期待値 ］

ある関数 $f\,(\,x\,)$ の確率分布 $p\,(\,x\,)$ の下での平均値のこと。

$$\mathrm{E}[\ f\ ] = \sum_{n=1}^{N} p\,(\,x_n\,)\,f\,(\,x_n\,)$$

$$\cdots\ (\ 1.33\ )'$$

［ 分散 ］

$f\,(\,x\,)$ がその平均値 $\mathrm{E}[\,f\,(\,x\,)\,]$ の周りでどれぐらいバラつくかを表す尺度のこと。

$$\mathrm{var}[\ f\ ] = \mathrm{E}[\ (\,f\,(\,x\,) - \mathrm{E}[\,f\,(\,x\,)\,]\,)^2\ ]$$

$$\cdots\ (\ 1.38\ )$$

$$= \mathrm{E}[\,f\,(\,x\,)^2\,] - \mathrm{E}[\,f\,(\,x\,)\,]^2$$

$$\cdots\ (\ 1.39\ )$$

［ エントロピー ］

情報量の確率平均のこと。分布 $p(\,x_n\,)$ に関して期待値をとる。

$$\mathrm{H} = -\sum_{n=1}^{N} p(\,x_n\,) \ln p(\,x_n\,)$$

$$\cdots\ (\ 1.98\ )'$$

［ ベルヌーイ分布 ］

二値確率変数 $x \in \{\ 0\ ,\ 1\ \}$ に対し、

$$\mathrm{Bern}(\ x \mid \mu\ ) = \mu^{x}\,(\ 1 - \mu\ )^{1-x}$$

$$\cdots\ (\ 2.2\ )$$

となる確率分布のこと。

［ 解 ］

まず、式 (2.257) について確かめる。ベルヌーイ分布は式 (2.2) で与えられるから、式の左辺は、

$$\sum_{x=0}^{1} p(\, x \mid \mu\,) = \sum_{x=0}^{1} \mathrm{Bern}(\, x \mid \mu\,) = \sum_{x=0}^{1} \mu^{x}(\, 1 - \mu\,)^{1-x}$$
$$= \mu^{0}(\, 1 - \mu\,)^{1-0} + \mu^{1}(\, 1 - \mu\,)^{1-1} = (\, 1 - \mu\,) + \mu = 1$$

となり、式 (2.257) が満たせた。

> ベルヌーイ分布は、二値確率変数 $x \in \{\, 0, 1 \,\}$ であるため、x の範囲は、 $0 \leq x \leq 1$ であることに注意する。

次に、式 (2.258) について確かめる。ベルヌーイ分布が二値確率変数 $x \in \{\, 0, 1 \,\}$ であることと、式 (1.33)′ より、ベルヌーイ分布 (2.2) の期待値は、

$$\mathrm{E}[\, x \,] = \sum_{x=0}^{1} p(\, x \mid \mu\,)\, x = \sum_{x=0}^{1} \mathrm{Bern}(\, x \mid \mu\,)\, x = \sum_{x=0}^{1} \mu^{x}(\, 1 - \mu\,)^{1-x}\, x$$
$$= \mu^{0}(\, 1 - \mu\,)^{1-0} \times 0 + \mu^{1}(\, 1 - \mu\,)^{1-1} \times 1 = 0 + \mu = \mu$$

となり、式 (2.258) が満たせた。

さらに、式 (2.259) について確かめる。式 (2.258) と式 (1.39) より、ベルヌーイ分布 (2.2) の分散は、

$$\mathrm{var}[\, x \,] = \mathrm{E}[\, x^2 \,] - \mathrm{E}[\, x \,]^2$$

… ※

と書き表せる。ここで、

$$\mathrm{E}[\, x^2 \,] = \sum_{x=0}^{1} p(\, x \mid \mu\,)\, x^2 = \sum_{x=0}^{1} \mathrm{Bern}(\, x \mid \mu\,)\, x^2 = \sum_{x=0}^{1} \mu^{x}(\, 1 - \mu\,)^{1-x}\, x^2$$
$$= \mu^{0}(\, 1 - \mu\,)^{1-0} \times 0^2 + \mu^{1}(\, 1 - \mu\,)^{1-1} \times 1^2 = 0 + \mu = \mu$$

であるから、式 ※ は、

$$= \mu - \mu^2 = \mu(\, 1 - \mu\,)$$

となり、式 (2.259) が満たせた。

最後に、ベルヌーイ分布に従う二値確率変数 x のエントロピー $\mathrm{H}[\, x \,]$ が、式 (2.260) で与えられることを示す。式 (1.98)′ より、ベルヌーイ分布のエントロピー $\mathrm{H}[\, x \,]$ は、

$$\mathrm{H}[\, x \,] = -\sum_{x=0}^{1} p(\, x \mid \mu\,) \ln p(\, x \mid \mu\,) = -\sum_{x=0}^{1} \mathrm{Bern}(\, x \mid \mu\,) \ln \mathrm{Bern}(\, x \mid \mu\,)$$
$$= -\sum_{x=0}^{1} \mu^{x}(\, 1 - \mu\,)^{1-x} \ln \mu^{x}(\, 1 - \mu\,)^{1-x}$$
$$= -\sum_{x=0}^{1} \mu^{x}(\, 1 - \mu\,)^{1-x} \ln \mu^{x}(\, 1 - \mu\,)^{1-x}$$

$$= -\mu^0(1-\mu)^{1-0}\ln\mu^0(1-\mu)^{1-0} - \mu^1(1-\mu)^{1-1}\ln\mu^1(1-\mu)^{1-1}$$
$$= -\mu\ln\mu - (1-\mu)\ln(1-\mu)$$

となる。よって、ベルヌーイ分布に従う二値確率変数 x のエントロピー $\mathrm{H}[x]$ が、式（2.260）で与えられることが示せた。