演習問題 2.02

ベルヌーイ分布の式 (2.2) の表現では、x の 2 つの値 0 と 1 に関して対称ではない。場合によっては、対称な $x \in \{ -1 , 1 \}$ を用いた等価な表現の方が便利である。このとき分布は、

$$p(x \mid \mu) = \left(\frac{1 - \mu}{2} \right)^{(1-x)/2} \left(\frac{1 + \mu}{2} \right)^{(1+x)/2} \quad \cdots (2.261)$$

と書くことができる。ただし、$\mu \in [-1 , 1]$ である。この分布 (2.261) が正規化されていることを示し、その平均、分散、およびエントロピーを計算せよ。

[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。

$$\mathrm{E}[f] = \sum_{n=1}^{N} p(x_n) f(x_n) \quad \cdots (1.33)'$$

[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。

$$\mathrm{var}[f] = \mathrm{E}[(f(x) - \mathrm{E}[f(x)])^2] \quad \cdots (1.38)$$

$$= \mathrm{E}[f(x)^2] - \mathrm{E}[f(x)]^2 \quad \cdots (1.39)$$

[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。

$$\mathrm{H} = - \sum_{n=1}^{N} p(x_n) \ln p(x_n) \quad \cdots (1.98)'$$

[解]

基本的には、演習問題 2.01 と同じ要領で解く。まず、確率分布 (2.261) が正規化されていることを示す。ここで、x の値は $x \in \{ -1 , 1 \}$ の 2 つの値しかとらないことに注意し、式 (2.261) の総和を計算すると、

$$\sum_{x=-1}^{1} p(x \mid \mu) = \sum_{x=-1}^{1} \left(\frac{1 - \mu}{2} \right)^{(1-x)/2} \left(\frac{1 + \mu}{2} \right)^{(1+x)/2}$$

$$= \left(\frac{1 - \mu}{2} \right)^{\{1-(-1)\}/2} \left(\frac{1 + \mu}{2} \right)^{\{1+(-1)\}/2}$$

$$+ \left(\frac{1 - \mu}{2} \right)^{(1-1)/2} \left(\frac{1 + \mu}{2} \right)^{(1+1)/2}$$

$$= \left(\frac{1-\mu}{2} \right)^{1} \left(\frac{1+\mu}{2} \right)^{0} + \left(\frac{1-\mu}{2} \right)^{0} \left(\frac{1+\mu}{2} \right)^{1}$$

$$= \frac{1-\mu}{2} + \frac{1+\mu}{2} = 1$$

となる。よって、確率分布 (2.261) が正規化されていることが示せた。

次に、確率分布 (2.261) の平均 $\mathrm{E}[x]$ を計算する。平均は式 (1.33)′ より、

$$\mathrm{E}[x] = \sum_{x=-1}^{1} p(x \mid \mu)\,x = \sum_{x=-1}^{1} \left(\frac{1-\mu}{2} \right)^{(1-x)/2} \left(\frac{1+\mu}{2} \right)^{(1+x)/2} x$$

$$= \left(\frac{1-\mu}{2} \right)^{\{1-(-1)\}/2} \left(\frac{1+\mu}{2} \right)^{\{1+(-1)\}/2} (-1)$$

$$+ \left(\frac{1-\mu}{2} \right)^{(1-1)/2} \left(\frac{1+\mu}{2} \right)^{(1+1)/2} \cdot 1$$

$$= -\frac{1-\mu}{2} + \frac{1+\mu}{2} = \mu$$

となる。

さらに、確率分布 (2.261) の分散 $\mathrm{var}[x]$ を計算する。分散は式 (1.39) より、

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

… ※

と表される。ここで、$\mathrm{E}[x^2]$ は、

$$\mathrm{E}[x^2] = \sum_{x=-1}^{1} p(x \mid \mu)\,x^2 = \sum_{x=-1}^{1} \left(\frac{1-\mu}{2} \right)^{(1-x)/2} \left(\frac{1+\mu}{2} \right)^{(1+x)/2} x^2$$

$$= \left(\frac{1-\mu}{2} \right)^{\{1-(-1)\}/2} \left(\frac{1+\mu}{2} \right)^{\{1+(-1)\}/2} (-1)^2$$

$$+ \left(\frac{1-\mu}{2} \right)^{(1-1)/2} \left(\frac{1+\mu}{2} \right)^{(1+1)/2} \cdot 1^2$$

$$= \frac{1-\mu}{2} + \frac{1+\mu}{2} = 1$$

であるから、式 ※ は、

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

となる。今、$\mu \in [-1, 1]$ であり、$\mu = -1, 1$ のとき、分散 $\mathrm{E}[x^2]$ は 0 となるので、この結果は妥当である。

最後に、確率分布 (2.261) のエントロピー $\mathrm{H}[x]$ を計算する。

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

$$= -\sum_{x=-1}^{1} \left(\frac{1-\mu}{2}\right)^{(1-x)/2} \left(\frac{1+\mu}{2}\right)^{(1+x)/2} \ln\left(\frac{1-\mu}{2}\right)^{(1-x)/2} \left(\frac{1+\mu}{2}\right)^{(1+x)/2}$$

$$= -\left(\frac{1-\mu}{2}\right)^{\{1-(-1)\}/2} \left(\frac{1+\mu}{2}\right)^{\{1+(-1)\}/2} \ln\left(\frac{1-\mu}{2}\right)^{\{1-(-1)\}/2} \left(\frac{1+\mu}{2}\right)^{\{1+(-1)\}/2}$$

$$-\left(\frac{1-\mu}{2}\right)^{\{1-1\}/2} \left(\frac{1+\mu}{2}\right)^{\{1+1\}/2} \ln\left(\frac{1-\mu}{2}\right)^{\{1-1\}/2} \left(\frac{1+\mu}{2}\right)^{\{1+1\}/2}$$

$$= -\frac{1-\mu}{2} \ln\frac{1-\mu}{2} - \frac{1+\mu}{2} \ln\frac{1+\mu}{2}$$

となる。今、$\mu \in [-1, 1]$ であり、$\mu = -1, 1$ のとき、分散 $\mathrm{H}[x]$ は 0 となるので、この結果は妥当である。