演習問題 1.41

確率の加法・乗法定理を用いて、相互情報量 $\mathbf{I}[\mathbf{x},\mathbf{y}]$ が式 (1.121) の関係を満たすことを示せ。

[微分エントロピー]

$$\mathbf{H}[\mathbf{x}] = -\int p(\mathbf{x}) \ln p(\mathbf{x})\, d\mathbf{x}$$

… (1.104)

[$\mathbf{x}$ に対する $\mathbf{y}$ の**条件付きエントロピー**]

$$\mathbf{H}[\mathbf{y}|\mathbf{x}] = -\int\int p(\mathbf{y},\ \mathbf{x}) \ln p(\mathbf{y}|\mathbf{x})\, d\mathbf{y}\, d\mathbf{x}$$

… (1.111)

[条件付きエントロピーの関係式]

$$\mathbf{H}[\mathbf{x},\ \mathbf{y}] = \mathbf{H}[\mathbf{y}|\mathbf{x}] + \mathbf{H}[\mathbf{x}]$$

… (1.112)

[**相互情報量**]

$$\mathbf{I}[\mathbf{x},\mathbf{y}] = \mathrm{KL}(p(\mathbf{x},\mathbf{y}) \| p(\mathbf{x})p(\mathbf{y})) = -\int\int p(\mathbf{x},\mathbf{y}) \ln\left(\frac{p(\mathbf{x})p(\mathbf{y})}{p(\mathbf{x},\mathbf{y})}\right) d\mathbf{x}\, d\mathbf{y}$$

… (1.120)

$$\mathbf{I}[\mathbf{x},\mathbf{y}] = \mathbf{H}[\mathbf{x}] - \mathbf{H}[\mathbf{x}|\mathbf{y}] = \mathbf{H}[\mathbf{y}] - \mathbf{H}[\mathbf{y}|\mathbf{x}]$$

… (1.121)

[解]

式 (1.120) より、相互情報量 $\mathbf{I}[\mathbf{x},\mathbf{y}]$ は、

$$\mathbf{I}[\mathbf{x},\mathbf{y}] = -\int\int p(\mathbf{x},\mathbf{y}) \ln\left(\frac{p(\mathbf{x})p(\mathbf{y})}{p(\mathbf{x},\mathbf{y})}\right) d\mathbf{x}\, d\mathbf{y}$$

$$= -\int\int p(\mathbf{x},\mathbf{y}) \ln p(\mathbf{x})p(\mathbf{y})\, d\mathbf{x}\, d\mathbf{y} + \int\int p(\mathbf{x},\mathbf{y}) \ln p(\mathbf{x},\mathbf{y})\, d\mathbf{x}\, d\mathbf{y}$$

… ※

と展開できる。ここで、式 ※ の第二項は、式 (1.104) より、

$$\int\int p(\mathbf{x},\mathbf{y}) \ln p(\mathbf{x},\mathbf{y})\, d\mathbf{x}\, d\mathbf{y} = -\mathbf{H}[\mathbf{x},\mathbf{y}]$$

となる。さらに、式 (1.112) より、上記の式の第二項は、

$$= -(\mathbf{H}[\mathbf{y}|\mathbf{x}] + \mathbf{H}[\mathbf{x}])$$

… ※※

となる。次に、式 ※ の第一項を整理していくと、

$$-\iint p(\mathbf{x},\mathbf{y}) \ln p(\mathbf{x})p(\mathbf{y})\, d\mathbf{x}\, d\mathbf{y} = -\iint p(\mathbf{x},\mathbf{y})\{\ln p(\mathbf{x}) + \ln p(\mathbf{y})\} d\mathbf{x}\, d\mathbf{y}$$

$$= -\iint p(\mathbf{x},\mathbf{y}) \ln p(\mathbf{x})\, d\mathbf{x}\, d\mathbf{y} - \iint p(\mathbf{x},\mathbf{y}) \ln p(\mathbf{y})\, d\mathbf{x}\, d\mathbf{y}$$

と展開でき、ここで、確率の加法定理より、

$$= -\int p(\mathbf{x}) \ln p(\mathbf{x})\, d\mathbf{x} - \int p(\mathbf{y}) \ln p(\mathbf{y})\, d\mathbf{y}$$

となり、式（ 1.104 ）より、

$$= -\int p(\mathbf{x}) \ln p(\mathbf{x})\, d\mathbf{x} - \int p(\mathbf{y}) \ln p(\mathbf{y})\, d\mathbf{y}$$

$$= \mathbf{H}[\mathbf{x}] + \mathbf{H}[\mathbf{y}]$$

となる。よって、式 ※ は、

$$(※) = (-\mathbf{H}[\mathbf{x}] - \mathbf{H}[\mathbf{y}]) + (\mathbf{H}[\mathbf{y}|\mathbf{x}] + \mathbf{H}[\mathbf{x}])$$

$$= (\mathbf{H}[\mathbf{x}] + \mathbf{H}[\mathbf{y}]) - (\mathbf{H}[\mathbf{y}|\mathbf{x}] + \mathbf{H}[\mathbf{x}])$$

$$= \mathbf{H}[\mathbf{y}] - \mathbf{H}[\mathbf{y}|\mathbf{x}]$$

となり、式（ 1.121 ）の $\mathbf{I}[\mathbf{x},\mathbf{y}] = \mathbf{H}[\mathbf{y}] - \mathbf{H}[\mathbf{y}|\mathbf{x}]$ が満たせた。また、※※ での

$$\mathbf{H}[\mathbf{x},\mathbf{y}] = \mathbf{H}[\mathbf{y}|\mathbf{x}] + \mathbf{H}[\mathbf{x}]$$

の変形を、

$$\mathbf{H}[\mathbf{x},\mathbf{y}] = \mathbf{H}[\mathbf{x}|\mathbf{y}] + \mathbf{H}[\mathbf{y}]$$

とすることにより、式（ 1.121 ）の $\mathbf{I}[\mathbf{x},\mathbf{y}] = \mathbf{H}[\mathbf{x}] - \mathbf{H}[\mathbf{x}|\mathbf{y}]$ も同様に満たせる。以上より、相互情報量 $\mathbf{I}[\mathbf{x},\mathbf{y}]$ が式（ 1.121 ）の関係を満たすことを示せた。

[確率の加法定理]

$$p(x) = \sum_y p(x, y)$$

[確率の乗法定理]

$$p(x, y) = p(x|y)p(y) = p(y|x)p(x)$$