14.4 アナログ情報源のエントロピー

確率密度関数 $f(x)$で規定されるアナログ情報源 X のエントロピーは離散値の場合と同様に定義される．

$$H(X) = E[-\log f(X)] = -\int_{-\infty}^{\infty} f(x)\log_2 f(x)dx$$

２つの確率変数 X, Y が関わる場合については，確率密度関数 $f(x, y)$に対して，

$$f_1(x) = \int_{-\infty}^{\infty} f(x,y)dy \qquad f_2(y) = \int_{-\infty}^{\infty} f(x,y)dx$$

を使い，

$$H(X,Y) = E[-\log f(X,Y)] = -\int_{-\infty}^{\infty}\int_{-\infty}^{\infty} f(x,y)\log_2 f(x,y)dxdy$$

$$H(X) = E[-\log_2 f_1(X)] = -\int_{-\infty}^{\infty} f_1(x)\log_2 f_1(x)dx$$

$$H(Y) = E[-\log_2 f_2(Y)] = -\int_{-\infty}^{\infty} f_2(y)\log_2 f_2(y)dy$$

$$H(X|Y) = E[-\log_2 f_y(X|Y)] = -\int_{-\infty}^{\infty}\int_{-\infty}^{\infty} f(x,y)\log_2 \frac{f(x,y)}{f_2(y)}dxdy$$

$$H(Y|X) = E[-\log_2 f_x(Y|X)] = -\int_{-\infty}^{\infty}\int_{-\infty}^{\infty} f(x,y)\log_2 \frac{f(x,y)}{f_1(x)}dxdy$$

が定義され，相互情報量

$$I(X;Y) = \int_{-\infty}^{\infty}\int_{-\infty}^{\infty} f(x,y)\log_2 \frac{f(x,y)}{f_1(x)f_2(y)}dxdy$$

も導入される．相互情報量については，次のように非負であることが示される．

$$\begin{aligned}
I(X;Y) &= \int_{-\infty}^{\infty}\int_{-\infty}^{\infty} f(x,y)\log_2 \frac{f(x,y)}{f_1(x)f_2(y)}dxdy \\
&= -\int_{-\infty}^{\infty}\int_{-\infty}^{\infty} f(x,y)\log_2 \frac{f_1(x)}{f(x|y)}dxdy \\
&\geq -\int_{-\infty}^{\infty}\int_{-\infty}^{\infty} \log_2 e\cdot f(x,y)\left(\frac{f_1(x)}{f(x|y)}-1\right)dxdy \\
&= -\int_{-\infty}^{\infty}\int_{-\infty}^{\infty} \log_2 e\cdot f_2(y)f_1(x)dxdy + \int_{-\infty}^{\infty}\int_{-\infty}^{\infty} \log_2 e\cdot f(x,y)dxdy \\
&= -\log_2 e + \log_2 e = 0
\end{aligned}$$

さらに多次元にする場合は，

$$\begin{aligned}
H(X_1,\cdots,X_n) &= E[-\log_2 f(X_1,\cdots,X_n)] \\
&= -\frac{1}{n}\int_{-\infty}^{\infty}\cdots\int_{-\infty}^{\infty} f(x_1,\cdots,x_n)\log_2 f(x_1,\cdots,x_n)dx_1\cdots dx_n
\end{aligned}$$

によりエントロピーが定義される．

ただし，この定式化には数学的な困難があり，

(1) エントロピーが負になることがある．

(2) エントロピーが無限大になることがある．

(3) 座標系の取り方によって，エントロピーが不変であるとは限らない．
ことが知られている[Reza 1961, p. 269].

【例 1】区間$[-A, A]$で一様分布をする確率変数Xで規定される情報源のエントロピー：

$$H(X) = -\int_{-A}^{A} \frac{1}{2A} \log_2 \frac{1}{2A} dx = \log_2 2A$$

【例 2】確率密度関数$f(x)$が

$$f(x) = \frac{1}{\sqrt{2\pi\sigma^2}} e^{-\frac{(x-\mu)^2}{2\sigma^2}}$$

となる確率変数Xで規定される白色ガウス情報源のエントロピー：

$$\begin{aligned}
H(X) &= -\int_{-\infty}^{\infty} f(x) \log_2 f(x) dx \\
&= -\int_{-\infty}^{\infty} \frac{1}{\sqrt{2\pi\sigma^2}} e^{-\frac{(x-\mu)^2}{2\sigma^2}} \cdot \log_2 \left(\frac{1}{\sqrt{2\pi\sigma^2}} e^{-\frac{(x-\mu)^2}{2\sigma^2}} \right) dx \\
&= \int_{-\infty}^{\infty} \frac{1}{\sqrt{2\pi\sigma^2}} e^{-\frac{(x-\mu)^2}{2\sigma^2}} \cdot \left(\log_2 \sqrt{2\pi\sigma^2} + \frac{(x-\mu)^2}{2\sigma^2} \cdot \log_2 e \right) dx \\
&= \log_2 \sqrt{2\pi\sigma^2} + \frac{1}{2} \log_2 e = \log_2 \sqrt{2\pi e \sigma^2}
\end{aligned}$$

【例 3】2 つの独立な確率変数X_1, X_2が与えられ，X_1は平均値 0，分散$\sigma_1{}^2$のガウス分布に，X_2は平均値 0，分散$\sigma_2{}^2$のガウス分布に従うとする．このとき，$Y = X_1 + X_2$によって規定される確率変数Yについて考えてみよう．

Yは平均値 0，分散$\sigma^2 = \sigma_1{}^2 + \sigma_2{}^2$のガウス分布に従う．ここでは，より一般的な次の命題を証明しよう．

X_1, X_2をそれぞれ平均値 0，分散$\sigma_1{}^2, \sigma_2{}^2$をもつ独立な確率変数，その確率密度関数をそれぞれ，

$$f(x_1) = \frac{1}{\sqrt{2\pi}\sigma_1} e^{-\frac{x_1{}^2}{2\sigma_1{}^2}} \qquad f(x_2) = \frac{1}{\sqrt{2\pi}\sigma_2} e^{-\frac{x_2{}^2}{2\sigma_2{}^2}}$$

とするとき，

$$Y = a_1 X_1 + a_2 X_2$$

によって規定される確率変数Yは平均値 0 をもち，分散$\sigma^2 = a_1{}^2\sigma_1{}^2 + a_2{}^2\sigma_2{}^2$をもつ正規分布をもつ．確率変数$Y_1 = a_1 X_1$, $Y_2 = a_2 X_2$を導入すると，その確率密度関数は，それぞれ，

$$\frac{1}{\sqrt{2\pi}\alpha_1} e^{-\frac{y_1{}^2}{2\alpha_1{}^2}} \qquad \frac{1}{\sqrt{2\pi}\alpha_2} e^{-\frac{y_2{}^2}{2\alpha_2{}^2}}$$

となる．ただし，$\alpha_1 = a_1\sigma_1,\ \alpha_2 = a_2\sigma_2$ である．$Y = Y_1 + Y_2$ の確率密度関数は，畳み込み (convolution)

$$g(y) = \int_{-\infty}^{\infty} f_1(y - y_2) f_2(y_2) dy_2$$

で得られる．ただし，$f_1(y_1),\ f_2(y_2)$ は $Y_1,\ Y_2$ の確率密度関数である．$g(y)$ の計算を進めよう．

$$\begin{aligned}
g(y) &= \int_{-\infty}^{\infty} \frac{1}{2\pi\alpha_1\alpha_2} \exp\left[-\frac{(y-y_2)^2}{2\alpha_1^{\,2}}\right] \exp\left[-\frac{y_2^{\,2}}{2\alpha_2^{\,2}}\right] dy_2 \\
&= \int_{-\infty}^{\infty} \frac{1}{2\pi\alpha_1\alpha_2} \exp\left[-\frac{(\alpha_1^{\,2}+\alpha_2^{\,2})y_2^{\,2} - 2\alpha_2^{\,2} y_2 y + \alpha_2^{\,2} y^2}{2\alpha_1^{\,2}\alpha_2^{\,2}}\right] dy_2 \\
&= \int_{-\infty}^{\infty} \frac{1}{2\pi\alpha_1\alpha_2} \exp\left[-\frac{\alpha_1^{\,2}+\alpha_2^{\,2}}{2\alpha_1^{\,2}\alpha_2^{\,2}}\left(y_2 - \frac{\alpha_2^{\,2} y}{\alpha_1^{\,2}+\alpha_2^{\,2}}\right)^2 - y^2\left(\frac{1}{2\alpha_1^{\,2}} - \frac{\alpha_2^{\,2}}{2\alpha_1^{\,2}(\alpha_1^{\,2}+\alpha_2^{\,2})}\right)\right] dy_2 \\
&= \int_{-\infty}^{\infty} \frac{1}{2\pi\alpha_1\alpha_2} \exp\left[-y^2\left(\frac{1}{2\alpha_1^{\,2}} - \frac{\alpha_2^{\,2}}{2\alpha_1^{\,2}(\alpha_1^{\,2}+\alpha_2^{\,2})}\right)\right] \exp\left[-\frac{\alpha_1^{\,2}+\alpha_2^{\,2}}{2\alpha_1^{\,2}\alpha_2^{\,2}}\left(y_2 - \frac{\alpha_2^{\,2} y}{\alpha_1^{\,2}+\alpha_2^{\,2}}\right)^2\right] dy_2 \\
&= \int_{-\infty}^{\infty} \frac{\sqrt{\alpha_1^{\,2}+\alpha_2^{\,2}}}{\sqrt{2\pi}\alpha_1\alpha_2} \frac{1}{\sqrt{2\pi(\alpha_1^{\,2}+\alpha_2^{\,2})}} \exp\left[-\frac{y^2}{2(\alpha_1^{\,2}+\alpha_2^{\,2})}\right] \exp\left[-\frac{\alpha_1^{\,2}+\alpha_2^{\,2}}{2\alpha_1^{\,2}\alpha_2^{\,2}}\left(y_2 - \frac{\alpha_2^{\,2} y}{\alpha_1^{\,2}+\alpha_2^{\,2}}\right)^2\right] dy_2
\end{aligned}$$

ここで，

$$\int_{-\infty}^{\infty} \frac{1}{\sqrt{2\pi}\sigma} \exp\left[-\frac{x^2}{2\sigma^2}\right] dx = 1$$

であるので，

$$g(y) = \frac{1}{\sqrt{2\pi(\alpha_1^{\,2}+\alpha_2^{\,2})}} \exp\left[-\frac{y^2}{2(\alpha_1^{\,2}+\alpha_2^{\,2})}\right]$$

となり，Y は平均値 0，分散 $\sigma^2 = a_1^{\,2}\sigma_1^{\,2} + a_2^{\,2}\sigma_2^{\,2}$ の正規分布に従うことが示された．

以上から，

$$H(Y) = \log_2 \sqrt{2\pi e(\sigma_1^{\,2} + \sigma_2^{\,2})}$$

Y は X_1 の各実現値 x_1 を平均値とする分散 $\sigma_2^{\,2}$ のガウス分布に従うので，

$$p(y \mid x_1) = \frac{1}{\sqrt{2\pi\sigma^2}} e^{-\frac{(y-x)^2}{2\sigma_z^{\,2}}}$$

となる．これから，

$$H(Y \mid X_1) = \log_2 \sqrt{2\pi e \sigma_2^{\,2}}$$

が導かれる．ゆえに，

$$I(X_1;Y)=H(Y)-H(Y\,|\,X_1)=\log_2\sqrt{\frac{{\sigma_1}^2+{\sigma_2}^2}{{\sigma_2}^2}}=\frac{1}{2}\log_2\left(1+\frac{{\sigma_1}^2}{{\sigma_2}^2}\right)$$

が得られる．

最大エントロピー定理

$$\int_{-\infty}^{\infty}x^2p(x)dx\le P\quad \text{かつ}\quad \int_{-\infty}^{\infty}p(x)dx=1$$

という条件下で，

$$H(X)=-\int_{-\infty}^{\infty}p(x)\log_2 p(x)dx$$

を最大化する $p(x)$ は

$$p(x)=\frac{1}{\sqrt{2\pi P}}e^{-\frac{x^2}{2P}}$$

である．

【証明】未定乗数 λ, μ を導入し，ラグランジュの未定乗数法により，

$$F=-f(x)\log_2 f(x)+\lambda x^2 f(x)+\mu f(x)$$

を最大化する $f(x)$ を求める．

$$\frac{\partial F}{\partial f(x)}=-\log_2 f(x)-\log_2 e+\lambda x^2+\mu=0$$

から，

$$f(x)=\alpha e^{\beta x^2}$$

が導かれる．これを条件式

$$\int_{-\infty}^{\infty}\alpha e^{\beta x^2}dx=1$$

に代入すると，ガウス積分公式

$$\int_{-\infty}^{\infty}e^{-ax^2}dx=\sqrt{\frac{\pi}{a}}$$

より，

$$\frac{1}{\alpha}=\sqrt{\frac{\pi}{-\beta}}$$

が得られる．ゆえに

$$\beta=-\pi\alpha^2$$

また，エントロピー最大になるとき，

$$\int_{-\infty}^{\infty}x^2 f(x)dx=P$$

すなわち，

$$\int_{-\infty}^{\infty}x^2\alpha e^{\beta x^2}dx=P$$

になるはずである．ガウスの公式

$$\int_{-\infty}^{\infty} x^2 e^{-ax^2} dx = \frac{\sqrt{\pi}}{2a^{\frac{3}{2}}}$$

から

$$\frac{\sqrt{\pi}}{2(-\beta)^{\frac{3}{2}}} = \frac{P}{\alpha}$$

ここに，$\beta = -\pi\alpha^2$ を代入して，

$$\frac{\sqrt{\pi}}{2(\pi\alpha^2)^{\frac{3}{2}}} = \frac{P}{\alpha}$$

から，

$$\alpha = \frac{1}{\sqrt{2\pi P}}, \quad \beta = -\frac{1}{2P}$$

を得る．以上により，

$$f(x) = \frac{1}{\sqrt{2\pi P}} e^{-\frac{x^2}{2P}}$$

である．　■

参考文献

F.M. Reza. An Introduction to Information Theory, McGraw-Hill, 1961.

G. Gallager. Information Theory and Reliable Communication, Wiley, 1968

今井秀樹．情報理論，昭晃堂，1974.