2020.6.9(火)

情報統計力学 学習8

6月9日 (火) 午後１時ー２時半

何回かにわたって、情報エントロピーと物理におけるエントロピーの関係について解説する。

1 情報エントロピー

X を離散型確率変数とする。簡単のため、X は有限個の値、$x_1, x_2, \cdots, x_n$ をとるとする。確率関数を $p(x)$ とすると、X の情報エントロピー H_X は次式で定義される。

$$H_X = -\sum_{i=1}^{n} p(x_i)\log_2 p(x_i) = E[\log_2\left(\frac{1}{p(X)}\right)]. \tag{1}$$

ここで、$\log_2$ は 2 を底とする対数。但し、$0\log_2 0 = 0$ とする。
エントロピーの単位は、ビット (bits)。
統計力学のエントロピー S は、状態数を W とすると、$S = k_{\mathrm{B}} \ln W$。ln は、自然対数で e(ネイピア数) を底とする。単位は、ナット (nats)。$\ln = \log_e,\ e = 2.7182818...$
情報エントロピー H_X は、直感的には、確率変数 X の不確定性を測る、missing information とも考えられる。
H_X が大きいほど、X についての**事前情報**が少ない。

X が確定しているとき。$X = \{x_1\},\ p(x_1) = 1,\ H_X = -\log_2 1 = 0,$
X のどの値も同じ確率で起こるとき。$p(x) = \dfrac{1}{|\Omega|}$, Ωは標本空間で、$|\Omega|$ はその要素数、n。

$$H_X = -\sum_{i=1}^{|\Omega|} \frac{1}{|\Omega|}\log_2\frac{1}{|\Omega|} = \log_2|\Omega|.$$

大雑把に言うと、H_X は、X がとりうる値の数の $\log_2$ となっている。
いくつか例をあげる。

コイン投げ。
表の出る確率を p、裏のでる確率を q とする。$X = 1$ が表、$X = 0$ が裏とする。これは、前に定義したベルヌーイ分布 $\mathrm{Ber}(p)$ である。

$$H_X = -p\log_2 p - q\log_2 q = -p\log_2 p - (1-p)\log_2(1-p) \equiv \mathcal{H}(p). \tag{2}$$

$p = \frac{1}{2}$ とすると、

$$H_X = -\frac{1}{2}\log_2\frac{1}{2} - \frac{1}{2}\log_2\frac{1}{2} = \log_2 2 = 1(\text{ビット}). \tag{3}$$

$\mathcal{H}(p)$ は、$p = \frac{1}{2}$ で最大値 1 を持ち、$p = \frac{1}{2}$ について対称で、$\mathcal{H}(0) = \mathcal{H}(1) = 0$ となる。

$p=\frac{1}{2}$ の同じコインが M 個あった場合

同時に投げたときに出る事象の個数は、全部で 2^M 個。$X=(X_1,X_2,\cdots,X_M)$ とする。$\boldsymbol{x}=(x_1,x_2,\cdots,x_M)$, $X_i=1,0$.

$$H_X=-\sum_{i=1}^{2^M}\frac{1}{2^M}\log_2\frac{1}{2^M}=-2^M\frac{1}{2^M}\log_2\frac{1}{2^M}=M\log_2 2=M. \tag{4}$$

M 個の面を持つサイコロ。どの面も同じ確率 $\frac{1}{M}$ ので出る場合

$$H_X=-\sum_{i=1}^{M}\frac{1}{M}\log_2\frac{1}{M}=-M\frac{1}{M}\log_2\frac{1}{M}=\log_2 M. \tag{5}$$

お金をかけるゲーム

確率変数 X がベルヌーイ分布 $\mathrm{Ber}(p)$ に従うとする。1 か 0 にお金を掛け、当たれば 2 倍になるとする。毎回、持金の p 倍を $X=1$ に、$(1-p)$ 倍を $X=0$ にかけるとする。W を最初に持っているお金の額とする。すると、１回めは、pW の額を 1 にかけ、$(1-p)W$ の額を 0 にかける。すると、１回目が終わったときの持金の期待値は、$2pW\times p+2(1-p)W\times(1-p)=2W(p^2+(1-p)^2)$. これを t 回行う。$X_1,X_2,\cdots,X_t$ なる確率変数列ができる。

$$w(X_i)=\begin{cases} p & X_i=1\\ 1-p & X_i=0\end{cases}$$

と定義する。すると、持金は以下のように変化する。

$$W\to 2Ww(X_1)\to 2^2Ww(X_1)w(X_2)\cdots\to 2^tWw(X_1)w(X_2)\cdots w(X_t)=W\prod_{t'=1}^{t}(2w(X_{t'})).$$

すると、倍率は、$\prod_{t'=1}^{t}(2w(X_{t'}))$. １回あたり 2^x 倍になったとすると、$2^x=2w(X_{t'})$. よって、$x=\log_2(2w(X_{t'}))$. これの期待値は、

$$\begin{aligned}\langle x\rangle &= \frac{1}{t}\sum_{t'}E[\log_2(2w(X_{t'}))]=\frac{1}{t}\sum_{t'}\Big(p\log_2(2w(1))+(1-p)\log_2(2w(0))\Big)\\ &= \frac{1}{t}\sum_{t'}\Big(p\log_2(2p)+(1-p)\log_2(2(1-p))\Big)\\ &= p\log_2 2+p\log_2 p+(1-p)\log_2 2+(1-p)\log_2(1-p)\\ &= 1+p\log_2 p+(1-p)\log_2(1-p)=1-\mathcal{H}(p).\end{aligned} \tag{6}$$

よって、$\mathcal{H}(p)$ が小さければ、儲けは大きい。$p=\frac{1}{2}$ なら、持金のまま。