情報理論　#2

第2回講義　4月25日

情報科学研究科　　井上純一

http://chaosweb.complex.eng.hokudai.ac.jp/~j_inoue/

情報量とは何か？

前回の復習

情報量： $-\log p$　　$p(E)=p$　はある事象Eが起こる確率

抽象的概念である「情報」の量の数学的な定義

例1)　明日の天気に関し、p(晴れ)=p(雨)=1/2の場合、「明日は晴れ」という通報を受けた場合に得られる情報量

$$-\log p(\textbf{晴れ}) = \log 2 = 1(\textbf{ビット})$$

例2)　「犬が人間に噛みつく確率」p(犬→人間)=2^{-3}、「人間が犬に噛みつく確率」p(人間→犬)=2^{-1000}

「犬が人間に噛みついた」という通報を受けた際に得られる情報量　$-\log p(\textbf{犬} \to \textbf{人間}) = 3(\textbf{ビット})$

「人間が犬に噛みついた」という通報を受けた際に得られる情報量　$-\log(\textbf{人間} \to \textbf{犬}) = 1000(\textbf{ビット})$

めったに起こらない事象に対する情報量は大きい

⇒　情報量とはある通報で我々が驚く度合いを表している

エントロピー

1つのコイン投げを考える

$p(\text{表}) = 1/2$　　表が出たことを知って得られる情報量

$$-\log p(\text{表}) = -\log 2^{-1} = 1(\text{ビット})$$

$p(\text{裏}) = 1/2$　　裏が出たことを知って得られる情報量

$$-\log p(\text{裏}) = -\log 2^{-1} = 1(\text{ビット})$$

⇒ 実際にコインを投げる際にどの程度の情報量の増加が期待できるか？

$$H = -\sum_{i=\text{表},\text{裏}} p(i)\log p(i) = -\frac{1}{2}\log\frac{1}{2} - \frac{1}{2}\log\frac{1}{2} = 1(\text{ビット})$$

エントロピー、あるいは**平均情報量**と呼ぶ

表、裏のどちらが出るかが予めわかっている場合　$p(\text{表}) = 1$, or, $p(\text{裏}) = 1$

⇒ エントロピーはゼロ

エントロピーの定義と性質

ある事象の集合を $\mathrm{A}=\{a_1, a_2, \ldots, a_m\}$ とし、$P_X(x)$ を $x \in \mathrm{A}$

なる事象が生起される確率とすると、確率変数 X のエントロピーは

$$H = -\sum_{x \in \mathrm{A}} P_X(x) \log P_X(x)$$ で与えられる

エントロピーの定義

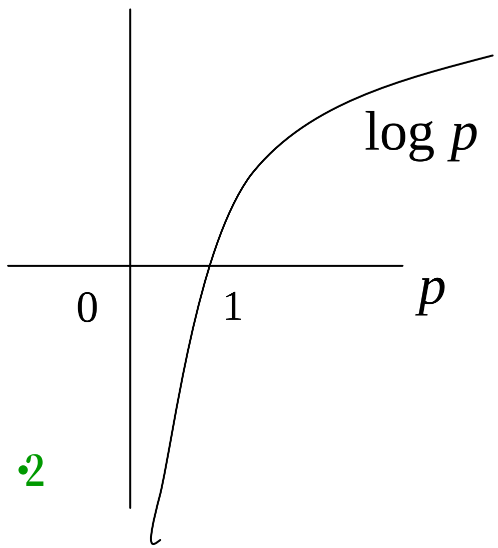
エントロピーの性質

(1) $H(x) \geq 0$

(2) $P(x)=1$ となる $x \in \mathrm{A}$ があれば $H(x)=0$

エントロピーは確率変数Xに関する我々の知識のあいまいさを表す

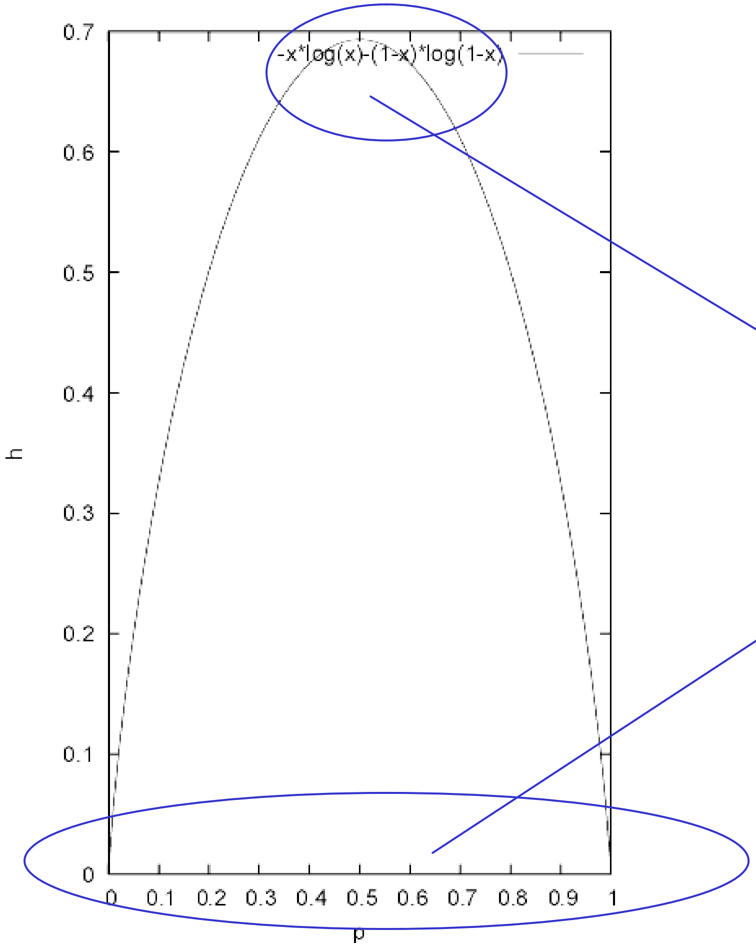
(3) $0 < P(x) < 1$ なる $x \in \mathrm{A}$ があれば $H(x) > 0$



具体例は教科書 p. 15、例2･2

2値エントロピー関数

$P(a) = p, p(b) = 1 - P(a) = 1 - p, \mathrm{A} = \{a, b\}$ のときのエントロピーは



$$\boxed{h(p) = -p\log p - (1-p)\log(1-p)}$$

2値エントロピー関数

最大値は $h(p = 1/2) = 0.69$

(左図では自然対数でプロットしていることに注意)

$h(p) \geq 0,\ h(p = 0) = h(p = 1) = 0$

2値エントロピー関数を一般化する問題
⇒ **例題2**

複合事象のエントロピー

2つの異なる事象の集合を$X=\{a_1,a_2,..,a_n\},Y=\{b_1,b_2,..,b_n\}$ とするとき

$$\boxed{H(X,Y)=-\sum_{x\in X}\sum_{y\in Y}P_{XY}(x,y)\log P_{XY}(x,y)}$$

複合事象のエントロピー 結合エントロピー)

(例)

a_1**(晴れ)**, a_2**(雨)**, b_1**(気温：摂氏0〜10度)**, b_2**(気温: 摂氏10〜20度)**

$$H(X,Y)=-P_{XY}(a_1,b_1)\log P_{XY}(a_1,b_1)-P_{XY}(a_1,b_2)\log P_{XY}(a_1,b_2)$$
$$-P_{XY}(a_2,b_1)\log P_{XY}(a_2,b_1)-P_{XY}(a_2,b_2)\log P_{XY}(a_2,b_2)$$

雨で気温が摂氏10度〜20度である確率

条件付きエントロピー

$X = x$　という条件下での確率変数　Y　についてのエントロピー

$$H(Y \mid X = x) = -\sum_{y \in Y} P_{Y|X}(y \mid X = x) \log P_{Y|X}(y \mid X = x)$$

$X = x$　の値に依存するので $P_X(x)$　で平均する

$$\begin{aligned} H(Y \mid X) &= \sum_{x \in X} P_X(x) H(Y \mid X = x) \\ &= -\sum_{x \in X} P_X(x) \sum_{y \in Y} P_{Y|X}(y \mid X = x) \log P_{Y|X}(y \mid X = x) \\ &= -\sum_{x \in X} \sum_{y \in Y} P_X(x) P_{Y|X}(y \mid x) \log P_{Y|X}(y \mid x) \end{aligned}$$

確率に関する積の公式より
$P_{XY}(x, y)$

条件付きエントロピー

$$\boxed{H(Y \mid X) = -\sum_{x \in X} \sum_{y \in Y} P_{XY}(x, y) \log P_{Y|X}(y \mid x)}$$

⇒　X が与えられた条件下で、Y についてのあいまいさを表す