به نام زندگی

آنتروپی منبع $H(X)$

$$H(X) = E\left\{\log \frac{1}{P(X)}\right\} = -E\left\{\log P(X)\right\}$$

$$= \sum_{x} p(x) \log \frac{1}{p(x)} = -\sum_{x} p(x) \log p(x)$$

مثال: یک منبع تولید اطلاعات باینری را در نظر بگیرید که در آن

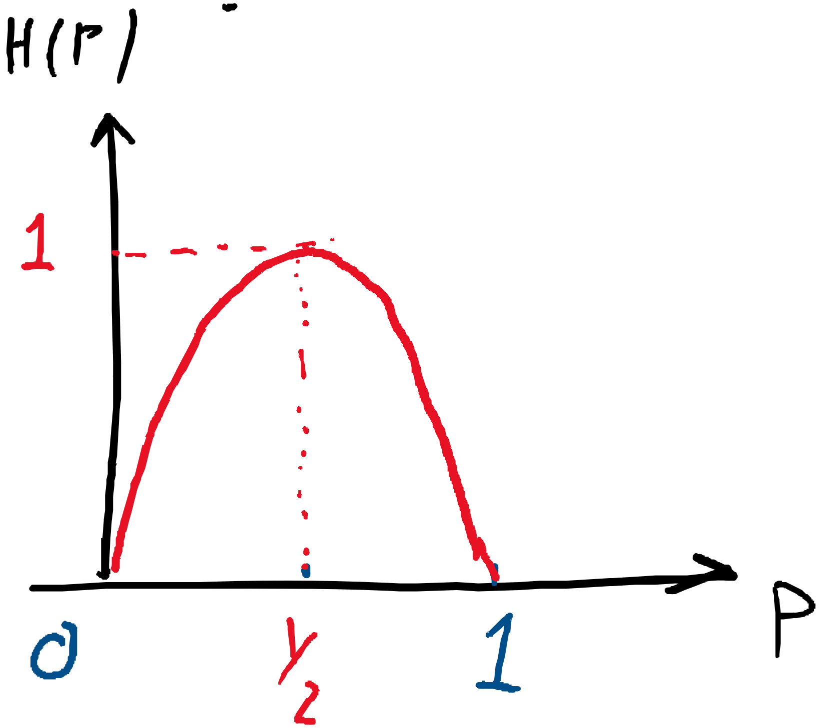
$$\mathcal{X} = \{\overset{A_1}{0}, \overset{A_2}{1}\} \quad , \quad x \in \mathcal{X}$$

$$P = \{\underset{p_1}{1-p}, \underset{p_2}{p}\} \longrightarrow \begin{array}{l} P_r\{X=1\} = p \\ P_r\{X=0\} = 1-p \end{array}$$

$X \sim$ Bernolli (p)

$$H(X) = -\sum_{x} p(x) \log p(x) = -p \log p - (1-p) \log (1-p) \triangleq H(p)$$

۲ اگر $H(p)$ را برحسب p رسم کنیم به نمودار زیر می‌رسیم



با توجه به نمودار می‌بینیم که به ازای $p=0$ یا $p=1$ ، آنتروپی منبع برابر صفر است زیرا در این دو حالت خروجی منبع همیشه برابر '0' یا همیشه برابر '1' است (اصلاً دمن نیست). آنتروپی به ازای $p=\frac{1}{2}$ بیشترین مقدار است زیرا در این حالت بیشترین ابهام را در مورد خروجی منبع داریم.

قضیه ماکزیمم آنتروپی

برای یک منبع گسسته با تابع جرم احتمال $P(x)$ که خروجی های آن از مجموعه $\mathcal{X}$ هستند، همواره داریم $H(X) \leq \log|\mathcal{X}|$ و آنتروپی زمانی ماکزیمم است که توزیع $P(x)$ یک توزیع یکنواخت باشد.

$$\mathcal{X} = \{A_1, A_2, \dots, A_n\} \quad , \quad P_X(x) = \Pr\{X = x\}$$

$$= \begin{cases} P_i & x = A_i \\ 0 & \text{oth.} \end{cases}$$

$$H(X) \leq \log |X|$$

همواره داریم

اگر توزیع منبع توزیع یکنواخت باشد داریم

$$P_1 = P_2 = \cdots = P_n \quad , \quad \sum_{i=1}^{n} P_i = 1 \implies P_i = \frac{1}{n} \, , \, i = 1, \ldots, n$$

$$\implies H(X) = \sum_{x} P(x) \log \frac{1}{P(x)} = \sum_{i=1}^{n} \frac{1}{n} \log n = n \frac{1}{n} \log n$$

$$\implies H(X) = \log n = \log |X|$$

* در حالت پیوسته، اگر خروجی منبع را با متغیر تصادفی پیوسته X نمایش دهیم قضیه ماکزیمم آنتروپی با در نظر گرفتن محدودیت توان P $(\text{Power } X \leq P)$ برای متغیرهای تصادفی گوسی $N(0,P)$ برقرار است.

مثال: فرض کنیم خروجی یک منبع، یک متغیر تصادفی X به صورت زیر باشد

$$X = \begin{cases} a & \frac{1}{2} \\ b & \frac{1}{4} \\ c & \frac{1}{8} \\ d & 1/8 \end{cases}$$

$$P_r\{X = a\} = \frac{1}{2}$$

$\vdots$

آنتروپی منبع را به دست بیاورید:

$$H(X) = \sum_{x} p(x) \log \frac{1}{p(x)}$$

$$= \frac{1}{2}\underbrace{\log 2}_{1} + \frac{1}{4}\underbrace{\log 4}_{2} + 2 \times \frac{1}{8}\underbrace{\log 8}_{3} = \frac{7}{4} \text{ bits}$$

(آنتروپی منبع برابر با میانگین تعداد بیت لازم برای بیان خروجی منبع نیز هست)

مثال (مسابقه بله - خیر)

فرض کنیم که در یک مسابقه بله - خیر ، شرکت کننده ی اول یکی از حروف $\{a, b, c, d\}$ را با احتمالات $\{\frac{1}{2}, \frac{1}{4}, \frac{1}{8}, \frac{1}{8}\}$ در نظر می گیرد. شرکت کننده ی دوم باید این حرف را حدس بزند و زمانی برنده است که بتواند با کمترین تعداد سؤالات به جواب درست برسد. بهترین راه حل برای رسیدن به جواب چیست؟

$X \stackrel{?}{=} a \xrightarrow{\text{بله}}$ برنده شدن با یک سوال $\left(\frac{1}{2}\right)$

$\xrightarrow{\text{خیر}} X \stackrel{?}{=} b \xrightarrow{\text{بله}}$ برنده شدن با دو سوال $\left(\frac{1}{4}\right)$

$\xrightarrow{\text{خیر}} X \stackrel{?}{=} c \xrightarrow{\text{بله}}$ برنده شدن با 3 سوال $\left(\frac{2}{8}\right)$

$\xrightarrow{\text{خیر}} X = d$

میانگین تعداد سوالات پرسیده شده برای حدس زدن جواب $= 1 \times \frac{1}{2} + 2 \times \frac{1}{4} + 3 \times \frac{2}{8} = \frac{7}{4} = H(X)$

به راحتی می‌توان نشان داد که در این گونه مسائل همواره

$$E\{\text{تعداد سؤالات لازم}\} \geq H(X)$$

کمترین میانگین تعداد سؤالات برای رسیدن جواب ، برابر آنتروپی آن متغیر تصادفی است .

مثال: اگر در مسئله قبل، شرکت کننده اول یک عدد صحیح بین 1 تا 50 را در نظر بگیرد، کمترین میانگین تعداد سؤالات لازم برای رسیدن به جواب چقدر است؟

در این حالت X یک متغیر تصادفی یکنواخت است،

$$X \in \{1, 2, \ldots, 50\} = \mathcal{X} \quad , \quad P_i = \frac{1}{50}$$

$$\Rightarrow \quad H(X) = \log|\mathcal{X}| = \log 50 \cong 5.64$$

$$E\left\{\text{تعداد سوالات}\right\} \leqslant H(X) = 5.64$$

$$X > 25 \begin{array}{l} \xrightarrow{\text{بله}} \\ \xrightarrow[\text{خیر}]{} \end{array}$$

* مفهوم آنتروپی مشترک یا توام دو متغیر تصادفی (joint Entropy)

برای دو متغیر تصادفی X و Y با تابع جرم احتمال توام $P(x,y)$

آنتروپی توام یا مشترک به صورت $H(X,Y)$ در نظر گرفته می‌شود که

نشان دهنده میزان اطلاعاتی است که X و Y با هم به دست می‌دهند و

داریم:

$$H(X,Y) = E\left\{ \log \frac{1}{P(X,Y)} \right\}$$

$$= \sum_{x} \sum_{y} p(x,y) \log \frac{1}{p(x,y)}$$

$$= - \sum_{x} \sum_{y} p(x,y) \log p(x,y)$$

که در آن $P(x,y)$ تابع جرم احتمال توأم دو متغیر تصادفی گسسته‌ی X و Y

است و داریم

$$P(x,y) = P_r\{X=x \text{ و } Y=y\}$$

* مفهوم آنتروپی شرطی

آنتروپی شرطی متغیر تصادفی X به شرط معلوم بودن متغیر تصادفی Y را آنتروپی شرطی X به شرط Y می‌گوییم و با $H(X|Y)$ نمایش می‌دهیم.

$H(X|y)$ نشان دهنده‌ی اطلاعات متغیر تصادفی X است وقتی متغیر تصادفی y مشخص باشد.

$$H(X|y) = E_{p(y)}\left\{ H(X|y=y) \right\}$$

که در آن

$$H(X|y=y) = E_{p(x|y)}\left\{ \log \frac{1}{p(x|y)} \right\} = \sum_{x} p(x|y) \log \frac{1}{p(x|y)}$$

$$= - \sum_{x} p(x|y) \log p(x|y)$$

$$\Rightarrow \quad H(X|Y) = E_{p(y)}\left\{ H(X|Y=y) \right\}$$

$$= E_{p(y)}\left\{ E_{p(x|y)}\left\{ \log \frac{1}{p(x|y)} \right\} \right\}$$

$$= E_{p(x,y)}\left\{ \log \frac{1}{p(x|y)} \right\}$$

$$H(X|Y) = E_{p(y)}\left\{ H(X|Y=y) \right\}$$

$$= E_{p(y)}\left\{ E_{p(x|y)}\left\{ \log \frac{1}{p(x|y)} \right\} \right\}$$

$$= \sum_{y} p(y)\, H(X|Y=y) = \sum_{y} \left\{ E_{p(x|y)}\left\{ \log \frac{1}{p(x|y)} \right\} \right\}$$

$$= \sum_{y} p(y) \left\{ \sum_{x} p(x|y) \log \frac{1}{p(x|y)} \right\}$$

رابطه بیزی

$$= \sum_{y} \sum_{x} p(y)\, p(x|y) \log \frac{1}{p(x|y)} = \sum_{y} \sum_{x} p(x,y) \log \frac{1}{p(x|y)}$$

رابطه زنجیره‌ای (یادآوری)

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

$$\Longrightarrow \quad H(X|y) = E_{p(X,y)}\left\{ \log \frac{1}{p(X|y)} \right\}$$

$$= -E_{p(X,y)}\left\{ \log p(X|y) \right\}$$

به طریق مشابه داریم

$$H(y|X) = E_{p(x)}\left\{ H(y|X=x) \right\}$$

$$= E_{p(x)}\left\{ E_{p(y|x)}\left\{ \log \frac{1}{p(y|x)} \right\} \right\}$$

$$= E_{p(x,y)}\left\{ \log \frac{1}{p(y|x)} \right\}$$

$$= -E_{p(x,y)}\left\{ \log p(y|x) \right\}$$

* رابطه بین آنتروپی مشترک و مشروط (رابطه زنجیره‌ای برای آنتروپی مشترک)

$$H(X,Y) = E_{p(X,Y)} \left\{ \log \frac{1}{p(X,Y)} \right\}$$

رابطه زنجیره‌ای

$$= -\sum_{x} \sum_{y} p(x,y) \log \overbrace{p(x,y)}^{p(x)\,p(y|x)}$$

$$= -\sum_{x} \sum_{y} p(x,y) \left[\log p(x) + \log p(y|x) \right]$$

$$\Rightarrow H(X,Y) = -\sum_{x}\sum_{y} p(x,y)\log p(x) \underbrace{\underbrace{-\sum_{x}\sum_{y} p(x,y)\log p(y|x)}_{E_{p(x,y)}\left[\log \frac{1}{p(y|x)}\right]}}_{H(Y|X)}$$

$$\Rightarrow H(X,Y) = -\sum_{x} \overbrace{\left(\sum_{y} p(x,y)\right)}^{p(x)} \log p(x) + H(Y|X)$$

$$= \underbrace{-\sum_{x} p(x)\log p(x)}_{H(X)} + H(Y|X)$$

$$\Longrightarrow H(X,Y) = H(X) + H(Y|X)$$ (مشابه رابطه زنجیره‌ای)

به طریق مشابه با جایگذاری $\left(P(x,y) = p(y)\,p(x|y) \right)$ (به عنوان تمرین انجام دهید)

داریم

$$H(X,Y) = H(Y) + H(X|Y)$$ (مشابه رابطه زنجیره‌ای)

(برای آنتروپی مشترک)

$$\Rightarrow \begin{cases} H(X,Y) = H(X) + H(Y|X) & \text{①} \\ H(X,Y) = H(Y) + H(X|Y) & \text{②} \end{cases}$$

$$\text{①} = \text{②} \Rightarrow \quad H(X) + H(Y|X) = H(Y) + H(X|Y)$$

بنابراین می توان نتیجه گیری کرد که

$$H(X|Y) = H(Y|X) \iff H(X) = H(Y)$$

* در حالت کلی $H(X|Y)$ و $H(Y|X)$ با هم مساوی نیستند.

مثال: تابع جرم احتمال توأم دو متغیر تصادفی X, Y به صورت زیر در نظر گرفته می‌شود

$p(x,y)$

$Y \backslash X$	1	2	3	4
1	$\frac{1}{8}$	$\frac{1}{16}$	$\frac{1}{32}$	$\frac{1}{32}$
2	$\frac{1}{16}$	$\frac{1}{8}$	$\frac{1}{32}$	$\frac{1}{32}$
3	$\frac{1}{16}$	$\frac{1}{16}$	$\frac{1}{16}$	$\frac{1}{16}$
4	$\frac{1}{4}$	0	0	0

$$P_r\{X=1, Y=1\} = \frac{1}{8}$$

$$P_r\{X=x\} = \sum_y p(x,y)$$

آنتروپی های $H(X)$ ، $H(Y)$ ، $H(X,Y)$ ، $H(X|Y)$ و $H(Y|X)$ را به دست بیاورید.

می خواهیم $H(Y|X)$ را به دست بیاوریم. می دانیم

$$H(Y|X) = E_{P(X)}\left\{ H(Y|X=x) \right\}$$

که در آن

$$H(Y|X=x) = E_{P(Y|X)}\left\{ \log \frac{1}{P(Y|X)} \right\}$$

برای حل مسئله لازم است که $P(x)$ و $P(y|x)$ را به دست بیاوریم

می‌دانیم:

$$P(y|x) = \frac{P(x,y)}{P(x)}$$

ابتدا $P(x)$ را به دست می‌آوریم:

$$P_r\{X=1\} = \frac{1}{8}+\frac{1}{16}+\frac{1}{16}+\frac{1}{4} = \frac{1}{2}$$

$$P_r\{X=3\} = \frac{1}{8}$$

$$P_r\{X=2\} = \frac{1}{4}$$

$$P_r\{X=4\} = \frac{1}{8}$$

$$P(y \mid X=1) = \frac{P(X=1, y)}{P(X=1)} = 2\left\{\frac{1}{8}, \frac{1}{16}, \frac{1}{16}, \frac{1}{4}\right\}$$

$$\Rightarrow H(y \mid X=1) = H\left(\frac{1}{4}, \frac{1}{8}, \frac{1}{8}, \frac{1}{2}\right)$$

$$P(y \mid X=2) = \frac{P(X=2, y)}{P(X=2)} = \left\{\frac{1}{4}, \frac{1}{2}, \frac{1}{4}, 0\right\}$$

$$\Rightarrow H(y \mid X=2) = H\left(\frac{1}{4}, \frac{1}{2}, \frac{1}{4}, 0\right)$$

$$P(y \mid X=3) = \frac{P(X=3, y)}{P(X=3)} = \left\{\frac{1}{4}, \frac{1}{4}, \frac{1}{2}, 0\right\}$$

$$\Rightarrow H(y \mid X=3) = H\left(\frac{1}{4}, \frac{1}{4}, \frac{1}{2}, 0\right)$$

$$P(y \mid X=4) = \frac{P(X=4, y)}{P(X=4)} = \left\{\frac{1}{4}, \frac{1}{4}, \frac{1}{2}, 0\right\}$$

$$\Rightarrow H(Y|X=4) = H\left(\frac{1}{4}, \frac{1}{4}, \frac{1}{2}, 0\right)$$

$$H(Y|X=4) = \frac{1}{4}\log 4 + \frac{1}{4}\log 4 + \frac{1}{2}\log 2 + 0$$

$$\Rightarrow H(Y|X) = E_{p(x)}\{H(Y|X=x)\}$$

$$= \frac{1}{2}H(Y|X=1) + \frac{1}{4}H(Y|X=2) + \frac{1}{8}H(Y|X=3) + \frac{1}{8}H(Y|X=4)$$

$$\Rightarrow H(Y|X) = \frac{1}{4} H\left(\frac{1}{4}, \frac{1}{8}, \frac{1}{2}, \frac{1}{2}\right) +$$

$$\frac{1}{8} H\left(\frac{1}{4}, \frac{1}{4}, \frac{1}{2}, 0\right) +$$

$$\frac{1}{16} H\left(\frac{1}{4}, \frac{1}{2}, \frac{1}{4}, 0\right) +$$

$$\frac{1}{16} H\left(\frac{1}{4}, \frac{1}{4}, \frac{1}{2}, 0\right)$$

$$\Leftrightarrow H(Y|X) = \frac{1}{4} H\left(\frac{1}{4}, \frac{1}{8}, \frac{1}{2}, \frac{1}{2}\right) + \frac{1}{4} H\left(\frac{1}{4}, \frac{1}{4}, \frac{1}{2}, 0\right)$$

$H(X)$ ، $H(Y)$ ، $H(X,Y)$ ، $H(X|Y)$ را

زیر $H(X)$: $P(X)$ را می‌گیرید

زیر $H(Y)$: $P(Y)$ را می‌گیرید (جمع سطرها)

زیر $H(X,Y)$: از جدول

زیر $H(X|Y)$: به طریق مشابه مثال (۱) شده $P(y)$ ، $P(x|\bar{y})$ را از جدول به دست بیاورید

به عنوان تمرین به دست بیاورید.

$$H(X,Y) = -\sum_{x}\sum_{y} p(x,y)\log p(x,y)$$

$$H(X|Y) = E_{p(y)}\left\{ H(X|Y=y) \right\}$$

$$H(X|Y=y) = E_{P(X|Y)} \left\{ \log \frac{1}{P(X|Y)} \right\}$$

* مفهوم اطلاعات متقابل دو متغیر تصادفی $I(X;Y)$

منظور از اطلاعات متقابل دو متغیر تصادفی X و Y ، میزان اطلاعاتی است که یکی از این متغیرها در مورد دیگری به ما می دهد که با $I(X;Y)$ نمایش داده می شود و به صورت زیر تعریف می شود.

$$I(X;Y) = E_{p(x,y)}\left\{ \log \frac{p(x,y)}{p(x)\,p(y)} \right\}$$

با توجه به تعریف بالا، به راحتی می‌توان دید که

$$I(X;Y) = I(Y;X)$$

همچنین با توجه به تعریف اطلاعات متقابل، می‌توان گفت که اگر X و Y مستقل از هم باشند یعنی

$$\forall x,y \;;\quad p(x,y) = p(x)\,p(y)$$

آنگاه اطلاعات متقابل این دو متغیر تصادفی برابر صفر است یعنی

$$X \perp\!\!\!\perp Y \implies I(X;Y) = I(Y;X) = 0$$

به عبارت دیگر میزان اطلاعاتی از یکی از این متغیرها در دیگری وجود دارد، برابر صفر است. (هیچ اطلاعاتی از یک متغیر تصادفی در دیگری وجود ندارد)

www.dr-haghbin.info/Courses/Digital Commu.

ahaghbin@gmail.com