به نام خدا

مروری بر مفاهیم تئوری اطلاعات — Information Theory

- آنتروپی — Entropy

مثال - مسابقه بله - خیر

یک بازی را در نظر بگیریم که در آن یکی از شرکت کنندگان، یکی از حروف $\{a, b, c, d\}$ را با احتمالات $\{\frac{1}{8}, \frac{1}{8}, \frac{1}{4}, \frac{1}{2}\}$ در نظر بگیرد. شرکت کننده‌ی دیگر باید با کمترین تعداد سؤالات (با جواب بله یا خیر) حرف مورد نظر را حدس بزند. راهکار پیشنهادی برای حل این مسأله چیست؟

$$
\begin{tikzcd}[column sep=large]
X \overset{?}{=} a \arrow[r, "\text{بله}"] \arrow[dr, "\text{خیر}"'] & X = a \; (\text{برنده شدن با یک سوال}) \; (1/2) & & \\
& X \overset{?}{=} b \arrow[r, "\text{بله}"] \arrow[dr, "\text{خیر}"'] & X = b \; (\text{برنده شدن با دو سوال}) \; (1/4) & \\
& & X \overset{?}{=} c \arrow[r, "\text{بله}"] \arrow[dr, "\text{خیر}"'] & X = c \; (\text{برنده شدن با سه سوال}) \; \left(\frac{1}{8} + \frac{1}{8}\right) \\
& & & X = d
\end{tikzcd}
$$

$$
E\left\{\begin{array}{c}\text{تعداد سوالات برای برنده شدن} \\ \text{با کمترین تعداد سوالات}\end{array}\right\} = \frac{1}{2} \times 1 + \frac{1}{4} \times 2 + \left(\frac{1}{8} + \frac{1}{8}\right) \times 3 = \frac{7}{4} = H(X)
$$

به صورت یک قانون کلی در این بازی همواره داریم

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

یعنی امید ریاضی تعداد سوالات لازم برای رسیدن به جواب همواره بزرگتر یا مساوی آنتروپی متغیر تصادفی مورد نظر است. با انتخاب بهینه سوالات می توانیم به میانگین برابر آنتروپی منبع برسیم.

مثال - فرض کنیم که شرکت کننده اول یک عدد را به صورت تصادفی از بین اعداد یک تا 50 در نظر بگیرد کمترین میانگین تعداد سوالات لازم برای رسیدن به جواب چقدر است؟

(توزیع یکنواخت)

$$X \in \{1, 2, \dots, 50\} = \mathcal{X}, \quad P_1 = P_2 = \dots = P_{50} = \frac{1}{50}$$

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

X : [illegible]

$$H(X) = \log |\mathcal{X}| = \log 50$$

$$\Rightarrow E\{\text{تعداد سوالات}\} \geq \log 50 \simeq 5.644$$

$$X \overset{?}{\geq} 25 \xrightarrow{\text{بله}} 25 - 50$$

$$X \overset{?}{\geq} 25 \xrightarrow{\text{خیر}} 1 - 25$$

آنتروپی مشترک یا توأم joint Entropy

برای دو متغیر تصادفی (X, Y) تنها می‌توان آنتروپی مشترک یا توأم تعریف کرد. منظور از آنتروپی مشترک دو متغیر تصادفی $H(X, Y)$، اطلاعاتی است که دو متغیر تصادفی با هم (به صورت توأم) به دست می‌دهند. آنتروپی مشترک دو متغیر تصادفی به صورت زیر تعریف می‌شود.

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

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

که در آن منظور از $p(x,y)$ تابع جرم احتمال توأم در متغیر تصادفی (X,Y) است.

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

* آنتروپی شرطی

آنتروپی شرطی متغیر تصادفی X به شرط معلوم بودن متغیر تصادفی Y، میزان اطلاعاتی است که متغیر تصادفی X به دست می‌دهد، به شرط آنکه متغیر تصادفی Y معلوم باشد. آنتروپی شرطی X به شرط معلوم بودن Y به صورت زیر تعریف می‌شود،

$$H(X|Y) = E_{P(y)}\Big\{ \underbrace{H(X|Y=y)}_{\text{تابعی از } y=y} \Big\} = \sum_{y} P(y)\, H(X|Y=y)$$

که در آن

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

$$= \sum_{x} P(x|y) \log \frac{1}{P(x|y)} = - \sum_{x} P(x|\overset{Y=y}{y}) \log P(x|y)$$

$$\Rightarrow \quad H(X|Y) = -\sum_{y} \sum_{x} \underbrace{P(y)\,P(x|y)}_{P(x,y)} \log P(x|y)$$

رابطه زنجیره‌ای

$$\Rightarrow H(X|Y) = -\sum_x \sum_y p(x,y)\log p(x|y) = -E_{p(x,y)}\{\log p(X|Y)\}$$

بطور مشابه می‌توانیم $H(Y|X)$ را نیز به صورت زیر تعریف کنیم.

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

که در آن

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

$$= -\sum_y p(y|x)\log p(y|x)$$

$$H(Y|X) = -\sum_x \sum_y \underbrace{p(x)\,p(y|x)}_{p(x,y)} \log p(y|x)$$

$$\Rightarrow \begin{cases} H(Y|X) = -\sum_x \sum_y p(x,y) \log p(y|x) = -E_{p(x,y)} \{ \log p(Y|X) \} \\ H(X|Y) = -\sum_x \sum_y p(x,y) \log p(x|y) = -E_{p(x,y)} \{ \log p(X|Y) \} \end{cases}$$

$$\begin{cases} H(Y|X) = E_{p(x)} \{ H(Y|X=x) \} \\ H(X|Y) = E_{p(y)} \{ H(X|Y=y) \} \end{cases}$$

رابطه زنجیره‌ای در آنتروپی

برای آنتروپی مشترک و شرطی، رابطه‌ی زنجیره‌ای به فرم زیر برقرار است:

$$H(X,Y) = H(X) + H(Y|X) \qquad (\text{I})$$

$$H(X,Y) = H(Y) + H(X|Y)$$

اثبات: برای رابطه (I)

$$H(X,Y) = -\sum_x \sum_y p(x,y) \log \overbrace{p(x,y)}^{p(x)\,p(y|x)}$$

$$= -\sum_x \sum_y p(x,y) \log p(x) \overbrace{- \sum_x \sum_y p(x,y) \log p(y|x)}^{H(Y|X)}$$

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

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

$$\Rightarrow H(X,y) = H(X) + H(y|X)$$

به طریق مشابه می‌توان نشان داد که

$$H(X,y) = H(y) + H(Y|X)$$

$$\Rightarrow \quad H(x,y) = H(x) + H(y|x) = H(y) + H(x|y)$$

$$\Rightarrow \quad \underline{H(x|y) = H(y|x)} \iff \underline{H(x) = H(y)}$$

(در حالت کلی $H(x|y)$ و $H(y|x)$ باهم برابر نیستند)

مثال - تابع چگالی احتمال توأم (pmf توأم) دو متغیر تصادفی (X, Y) مطابق جدول زیر مشخص شده است. آنتروپی های زیر را محاسبه کنید

$$H(X)\ ,\ H(Y)\ ,\ H(X,Y)\ ,\ H(X|Y)\ ,\ H(Y|X)$$

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

$X \in \{1,2,3,4\}\ ,\ Y \in \{1,2,3,4\}$

Y \ 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

$\rightarrow p(y=1) = \sum_x p(x, \overset{y=1}{y}) = \frac{1}{8} + \frac{1}{16} + \frac{1}{32} + \frac{1}{32} = \frac{1}{4}$

$\rightarrow p(y=2) = \sum_x p(x, \overset{y=2}{y}) = \frac{1}{16} + \frac{1}{8} + \frac{1}{32} + \frac{1}{32} = \frac{1}{4}$

$\rightarrow p(y=3) = \sum_x p(x, y=3) = 4 \times \frac{1}{16} = \frac{1}{4}$

$\rightarrow p(y=4) = \frac{1}{4}$

$p(X=1) = \frac{1}{8} + \frac{1}{16} + \frac{1}{16} + \frac{1}{4} = \frac{1}{2}$

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

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

$$p(x|y) = \frac{p(x,y)}{\underbrace{p(y)}_{1/4}} = 4\, p(x,y)$$

$$\Longrightarrow \begin{cases} p(x|Y=1) = \left(\frac{1}{2}, \frac{1}{4}, \frac{1}{8}, \frac{1}{8}\right) \\ p(x|Y=2) = \left(\frac{1}{4}, \frac{1}{2}, \frac{1}{8}, \frac{1}{8}\right) \\ p(x|Y=3) = \left(\frac{1}{4}, \frac{1}{4}, \frac{1}{4}, \frac{1}{4}\right) \\ p(x|Y=4) = (1, 0, 0, 0) \end{cases}$$

$$\Rightarrow \begin{cases} H(X \mid Y=1) = H\left(\frac{1}{2}, \frac{1}{4}, \frac{1}{8}, \frac{1}{8}\right) \\ H(X \mid Y=2) = H\left(\frac{1}{4}, \frac{1}{2}, \frac{1}{8}, \frac{1}{8}\right) \\ H(X \mid Y=3) = H\left(\frac{1}{4}, \frac{1}{4}, \frac{1}{4}, \frac{1}{4}\right) \\ H(X \mid Y=4) = H(1, 0, 0, 0) \end{cases}$$

$$\Rightarrow H(X \mid Y) = \sum_{y} p(y) \, H(X \mid Y = y)$$

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

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

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

به عنوان تمرین : $H(X)$ ، $H(Y)$ ، $H(X,Y)$ ، $H(Y|X)$ را بدست

بیاورید

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

$$H(Y) = -\sum_{y} p(y)\log p(y) = H\left(\frac{1}{4}, \frac{1}{4}, \frac{1}{4}, \frac{1}{4}\right)$$

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

$$H(Y|X) = \sum_{x} p(x) \, H(Y|X=x)$$

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

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