Lecture 6

Source Entropy for two random variables X and Y.

Joint Entropy: Uncertainty about the joint sources X and Y.

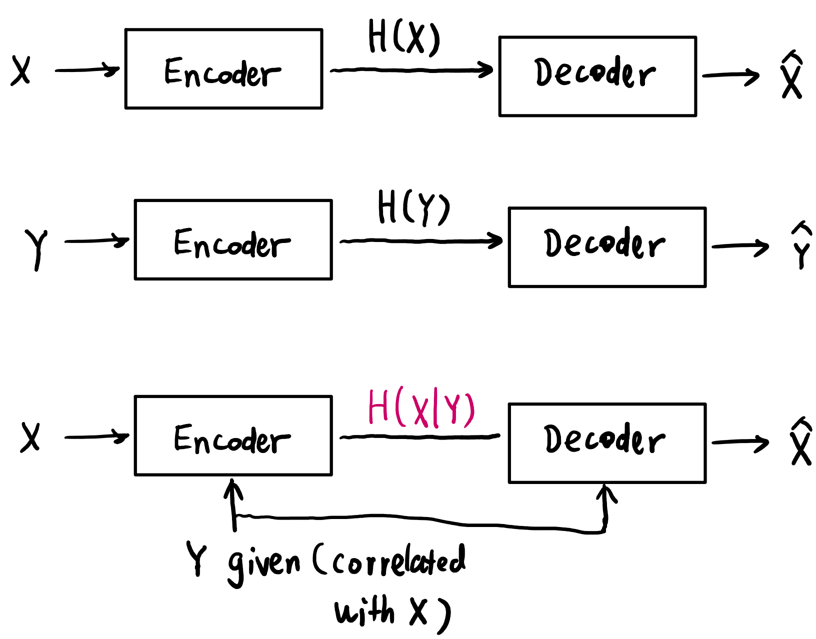
$$H(X,Y) = -\sum_{x \in X}\sum_{y \in Y} p(x,y) \log_2 p(x,y)$$

Joint Distribution P(X,Y):

$$\sum_{x \in X}\sum_{y \in Y} p(x,y) = 1$$

Y \ X	1	2	3	4	Σ
1	$\frac{1}{8}$	$\frac{1}{16}$	$\frac{1}{32}$	$\frac{1}{32}$	$P(Y=1) = \frac{1}{4}$
2	$\frac{1}{16}$	$\frac{1}{8}$	$\frac{1}{32}$	$\frac{1}{32}$	$P(Y=2) = \frac{1}{4}$
3	$\frac{1}{16}$	$\frac{1}{16}$	$\frac{1}{16}$	$\frac{1}{16}$	$P(Y=3) = \frac{1}{4}$
4	$\frac{1}{4}$	0	0	0	$P(Y=4) = \frac{1}{4}$
Σ	$P(X=1)$	$P(X=2)$	$P(X=3)$	$P(X=4)$	
	$\Downarrow$	$\Downarrow$	$\Downarrow$	$\Downarrow$	
	$\frac{1}{2}$	$\frac{1}{4}$	$\frac{1}{8}$	$\frac{1}{8}$	

- Conditional Entropy: H(X|Y) -- uncertainty about X given Y;
  H(Y|X) -- uncertainty about Y given X.



$H(X|Y) < H(X)$

Definition:

$$H(X|Y) = -\sum_{x\in X}\sum_{y\in Y} p(X=x, Y=y)\log_2 p(X=x|Y=y)$$

$$H(Y|X) = -\sum_{x\in X}\sum_{y\in Y} p(X=x, Y=y)\log_2 \underbrace{p(Y=y|X=x)}_{?}$$

How about $H(Y|X)$?

$P(Y=y \mid X=x) = ?$

Joint Distribution P(X,Y):

Y \ X	1	2	3	4	Σ
1	$\frac{1}{8}$	$\frac{1}{16}$	$\frac{1}{32}$	$\frac{1}{32}$	$P(Y=1)=\frac{1}{4}$
2	$\frac{1}{16}$	$\frac{1}{8}$	$\frac{1}{32}$	$\frac{1}{32}$	$P(Y=2)=\frac{1}{4}$
3	$\frac{1}{16}$	$\frac{1}{16}$	$\frac{1}{16}$	$\frac{1}{16}$	$P(Y=3)=\frac{1}{4}$
4	$\frac{1}{4}$	0	0	0	$P(Y=4)=\frac{1}{4}$
Σ	$P(X=1)$ ⇓ $\frac{1}{2}$	$P(X=2)$ ⇓ $\frac{1}{4}$	$P(X=3)$ ⇓ $\frac{1}{8}$	$P(X=4)$ ⇓ $\frac{1}{8}$	

$$P(Y=1 \mid X=1) = \frac{P(Y=1, X=1)}{P(X=1)} = \frac{\frac{1}{8}}{\frac{1}{2}} = \frac{1}{4}$$

$$P(Y=2 \mid X=1) = \frac{P(Y=2, X=1)}{P(X=1)} = \frac{\frac{1}{16}}{\frac{1}{2}} = \frac{1}{8}$$

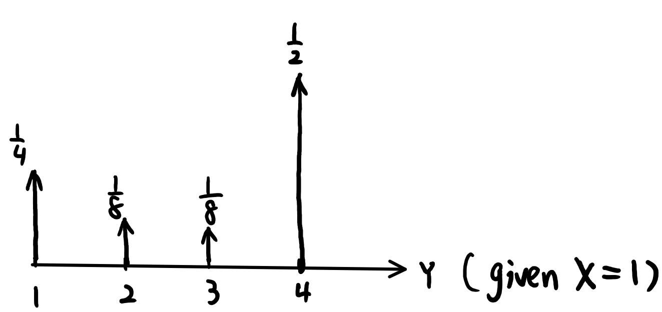
$$P(Y=3 \mid X=1) = \cdots = \frac{\frac{1}{16}}{\frac{1}{2}} = \frac{1}{8}$$

$$P(Y=4 \mid X=1) = \cdots = \frac{\frac{1}{4}}{\frac{1}{2}} = \frac{1}{2}$$

P(AB) = P(A) x P(B|A) = P(B) x P(A|B): Bayes Rule

$$\Rightarrow P(B|A) = \frac{P(AB)}{P(A)} \quad \text{or} \quad P(A|B) = \frac{P(AB)}{P(B)}$$

Conditional PMF: $P(Y|X=1)$



⋮

Alternatively,

$$H(Y|X) = -\sum_{x \in X} \sum_{y \in Y} \underbrace{p(X=x, Y=y)}_{P(Y=y|X=x) \cdot \underbrace{p(X=x)}_{\text{Fixed for the group}}} \log_2 p(Y=y|X=x)$$

e.g., define

$$H(Y|X=1) = -\{ P(Y=1|X=1) \cdot \log_2 P(Y=1|X=1) + P(Y=2|X=1) \cdot \log_2 P(Y=2|X=1)$$

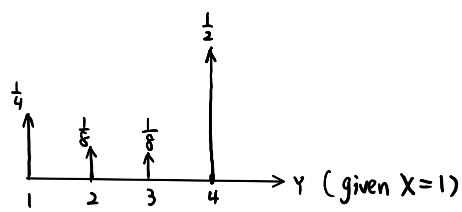
$$+ P(Y=3|X=1) \cdot \log_2 P(Y=3|X=1) + P(Y=4|X=1) \cdot \log_2 P(Y=4|X=1) \}$$

P(AB) = P(A) x P(B|A) = P(B) x P(A|B): Bayes Rule

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

Conditional PMF: $P(Y|X=1)$ $\Rightarrow$

$$= \tfrac{7}{4} \text{ bits}$$



Thus for the 4 groups: ← 1st group

$$H(Y|X) = P(X=1) \cdot H(Y|X=1)$$

$$+ P(X=2) \cdot H(Y|X=2)$$

$$+ P(X=3) \cdot H(Y|X=3)$$

$$+ P(X=4) \cdot H(Y|X=4)$$

$$= \tfrac{13}{8} \text{ bits (Verify this!)}$$

PMF of Y:

$P(Y=1) = \frac{1}{4}$
$P(Y=2) = \frac{1}{4}$
$P(Y=3) = \frac{1}{4}$
$P(Y=4) = \frac{1}{4}$

Compared with $H(Y) = H\left(\tfrac{1}{4}, \tfrac{1}{4}, \tfrac{1}{4}, \tfrac{1}{4}\right)$

$$= -4 \times \tfrac{1}{4} \log_2 \tfrac{1}{4} = 2 \text{ bits}$$

Check if X and Y are dependent on each other?

If A and B are independent, then P(AB) = P(A) x P(B)

Similarly,

$$H(X|Y) = -\sum_{x \in X}\sum_{y \in Y} p(X=x, Y=y) \log_2 p(X=x|Y=y)$$

$$= p(Y=1)\cdot H(X|Y=1) + p(Y=2)\cdot H(X|Y=2) + p(Y=3)\cdot H(X|Y=3) + p(Y=4)\cdot H(X|Y=4)$$

$$= \frac{11}{8} \text{ bits (Verify this!)}$$

Compared with $H(X)$

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

$$= \frac{7}{4} \text{ bits}$$

$P(X=1)$	$P(X=2)$	$P(X=3)$	$P(X=4)$
$\Downarrow$	$\Downarrow$	$\Downarrow$	$\Downarrow$
$\frac{1}{2}$	$\frac{1}{4}$	$\frac{1}{8}$	$\frac{1}{8}$

$$H(X|Y) < H(X)$$

Joint Distribution P(X,Y):

Y \ X	1	2	3	4	Σ
1	$\frac{1}{8}$	$\frac{1}{16}$	$\frac{1}{32}$	$\frac{1}{32}$	$P(Y=1)=\frac{1}{4}$
2	$\frac{1}{16}$	$\frac{1}{8}$	$\frac{1}{32}$	$\frac{1}{32}$	$P(Y=2)=\frac{1}{4}$
3	$\frac{1}{16}$	$\frac{1}{16}$	$\frac{1}{16}$	$\frac{1}{16}$	$P(Y=3)=\frac{1}{4}$
4	$\frac{1}{4}$	0	0	0	$P(Y=4)=\frac{1}{4}$
Σ	$P(X=1)$ $\Downarrow$ $\frac{1}{2}$	$P(X=2)$ $\Downarrow$ $\frac{1}{4}$	$P(X=3)$ $\Downarrow$ $\frac{1}{8}$	$P(X=4)$ $\Downarrow$ $\frac{1}{8}$	

Next, check if X and Y are independent?

If X and Y are independent, then P(XY) = P(X) x P(Y)

$$P(X=1, Y=1) = P(X=1)\cdot P(Y=1) = \frac{1}{8}$$

$$P(X=2, Y=2) = \frac{1}{8} \neq P(X=2)\cdot P(Y=2), \text{ Since}$$

$$P(X=2)\cdot P(Y=2) = \frac{1}{4}\cdot\frac{1}{4} = \frac{1}{16}$$

Thus X and Y are NOT independent.