Recall

Ω: discrete sample space, $P : \Omega \to [0,1]$ probability distribution on Ω

$X : \Omega \to \mathbb{R}$ random variable

$p_X = P \circ X^{-1}$: probability induced by X on $\mathbb{R}$.

For r.v.s $X : \Omega_1 \to \mathbb{R}$ and $Y : \Omega_2 \to \mathbb{R}$, the joint distribution $p_{X,Y} : \Omega_1 \times \Omega_2 \to \mathbb{R}$

$$\underbrace{p_{X,Y}}_{\text{joint distribution}} = \underbrace{p_X}_{\text{marginal}} \times \underbrace{p_{Y|X}}_{\text{conditional}} = \underbrace{p_Y}_{\text{marginal}} \times \underbrace{p_{X|Y}}_{\text{conditional}}.$$

X_1, X_2 random variables are independent if

$$p_{X_1|X_2} = p_{X_1}$$

Then

$$p_{X_1 X_2} = p_{X_2} \, p_{X_1|X_2} = p_{X_2} \, p_{X_1} = p_{X_1} \, p_{X_2|X_1}$$

if and only if

$$P_{X_2} = P_{X_2|X_1}.$$

$X_1, X_2, \dots, X_n$ are mutually independent if

$$p_{X_1,\dots,X_n} = p_{X_1} \times p_{X_2} \times \dots \times p_{X_n}.$$

$X_1, \dots, X_n$ are pairwise independent if $\forall i, j$ $i \neq j$

$$p_{X_i, X_j} = p_{X_i} \, p_{X_j}$$

Entropy

"expected number of bits to represent an element in range (X)"

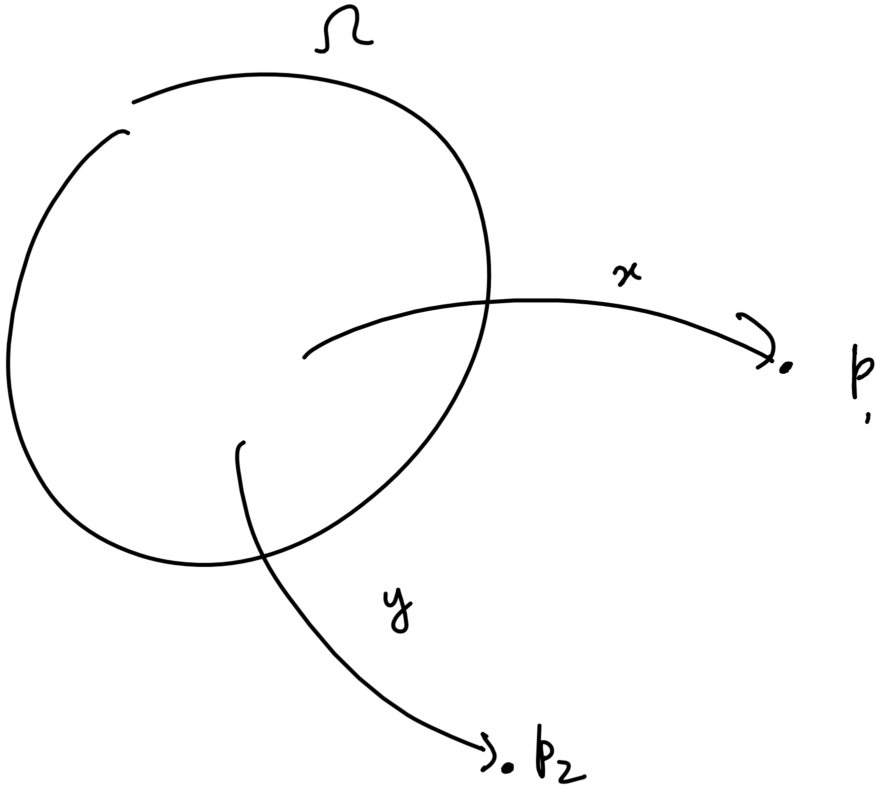
$$H(X) = -\sum_{x \in \text{range}(X)} p_X(x) \log_2 p_X(x)$$

$$H(X,Y) = -\sum_{x \in \text{range}(X)} \sum_{y \in \text{range}(Y)} p_{X,Y}(x,y) \log p_{X,Y}(x,y).$$

$$H(Y|X) = -\sum_{x \in \text{range}(X)} \sum_{y \in \text{range}(Y)} \underbrace{p_{X,Y}(x,y)}_{\text{joint}} \log \underbrace{p_{Y|X}(y|x)}_{\text{conditional}}$$

Note: 1) $H(X,Y) = H(X) + H(Y|X) = H(Y) + H(X|Y)$.

2) $H(X,Y) = H(Y,X)$.



Estimate of size of Ω	# bits required to represent size of Ω
$1/p_1$	$\log_2 (1/p_1)$
	$= -\log_2 p_1$

Weighted average for estimate of $|\Omega|$ $= p_1(-\log p_1) + p_2(-\log p_2)$

Why?

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

$$= -\sum_x \sum_y p_X(x)\, p_{Y|X}(y|x) \log p_{Y|X}(y|x)$$

$$= \sum_x p_x \left[-\sum_{y \in \mathrm{range}(Y)} p_{Y|X}(y|x) \log p_{Y|X}(y|x) \right]$$

$$= \sum_x p_x\, H(Y|X=x).$$

Information about X contained in Y is $I(X:Y) = H(X) - H(X|Y)$.

Lemma Symmetry of information

$$I(X:Y) = I(Y:X)$$

Proof

$$\begin{aligned} I(X:Y) - I(Y:X) &= H(X) - H(X|Y) - H(Y) + H(Y|X) \\ &= H(X) + H(Y|X) - [H(Y) + H(X|Y)] \\ &= H(X,Y) - H(Y,X) \\ &= 0. \end{aligned}$$

Jensen's Inequality

Fundamental Inequality

Basic Tools in Shannon Theory