

Definition: Conditional Mutual Information

Applying the definition of mutual information to the conditional distribution $P_{XY|\mathcal{A}}$ naturally defines $I(X; Y|\mathcal{A})$, the mutual information of X and Y conditioned on the event \mathcal{A} :

Definition: Conditional mutual information

Let X, Y, Z be random variables. Then the conditional mutual information of X and Y given Z is defined as

$$I(X; Y|Z) = \sum_z P_Z(z) I(X; Y|Z=z),$$

with the convention that the corresponding argument in the summation is 0 for z with $P_Z(z) = 0$.

Conditional mutual information has properties similar to the ones we saw for mutual information:

$$\begin{aligned} I(X; Y|Z) &= I(Y; X|Z) \\ I(X; Y|Z) &\geq 0 \\ I(X; Y|Z) &= 0 \text{ iff } X \text{ and } Y \text{ are independent given } Z. \end{aligned}$$

Furthermore, the previous bounds $H(X) \geq 0$, $H(X|Y) \geq 0$, and $I(X; Y) \geq 0$, can all be seen as special cases of $I(X; Y|Z) \geq 0$. These bounds, and any bound they imply, are called **Shannon inequalities**. It is important to realize that $I(X; Y|Z)$ may be larger or smaller than (or equal to) $I(X; Y)$.

The following is sometimes used as definition of $I(X; Y|Z)$: verify it for yourself using the definition above.

Alternative definition

Let X, Y, Z be random variables. Then

$$I(X; Y|Z) = H(X|Z) - H(X|YZ).$$

Proof

$$\begin{aligned} I(X; Y|Z) &= \sum_{z \in \mathcal{Z}} P_Z(z) I(X; Y|Z=z) && \text{(by definition)} \\ &= \sum_{z \in \mathcal{Z}} P_Z(z) (H(X|Z=z) - H(X|Y, Z=z)) && \text{(definition of mutual information)} \\ &= \sum_{z \in \mathcal{Z}} P_Z(z) H(X|Z=z) - \sum_{z \in \mathcal{Z}} P_Z(z) H(X|Y, Z=z) \\ &= H(X|Z) - H(X|YZ). \end{aligned}$$