Entropy and Mutual Information

LET:

- X be a random variable (R.V).
- P(X) be the probability distribution of X.
- p(x) be the probability density of X.

The **entropy** of X, H(X) is defined by:

$$H(X) = -E_X[\ \log(P(X))\ ]$$

Entropy is a measure of randomness.
The more random a variable is, the more Entropy it has.

The **joint entropy** is a statistics that summarizes the degree of dependence of a RV X on an other RV Y. It is defined by:

$$H(X,Y) = -E_X\ [E_Y\ [\ \log(p(X,Y))\ ]\ ]$$

The **conditional entropy** is a statistics that summarizes the randomness of Y given knowledge of X. It is defined by:

$$H(X|Y) = -E_X\ [E_Y\ [\ \log(p(X|Y))\ ]\ ]$$

Two random variables are considered to be **independent** if:

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

The **Mutual Information**, MI, between two random variables X and Y is given by:

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

(It's thus a measure of the reduction of the entropy of Y given X.)