Alberto Andrés Valdés González.
Degree: Mathematical Engineer.
Work position: Data Scientist.
Mail: anvaldes@uc.cl/alberto.valdes.gonzalez.96@gmail.com
Location: Santiago, Chile.

Mutual Information

First of all we have to introduce the concept of **entropy.**

Entropy: The entropy of a variable is a masure of the information, or alternatively, the üncertainty.ºf the variable's possibles values.

$$\boxed{H(X) = -\sum_{x \in X} p(x) \cdot log_2(p(x))}$$

Relative Entropy: The realtive entropy measures the distance between two distributions and it is also called Kullback-Leibler distance.

$$\boxed{P(p|q) = \sum_{x \in X} \sum_{y \in Y} p(x) \cdot log_2 \left(\frac{p(x)}{q(y)} \right)}$$

Mutual Information: Utilizing the relative entropy, we can now define the mutual information. We define the mutual information as the relative entropy between the joint distribution of the two variables and the product of their marginal distributions.

$$\boxed{I(X,Y) = \sum_{x \in X} \sum_{y \in Y} p(x,y) \cdot log_2 \left(\frac{p(x,y)}{p(x) \cdot p(y)} \right)}$$

Interpretation: Mutual information is always larger than or equal to zero. Meanwhile larger is the value, greater the relationship between the two variables. If the calculated result is zero, then the variables are independent.
