



## Why

1

## Definition

The mutual information of a joint distribution over two random variables is the entropy of the product of the marginal distributions relative to the joint distribution.

## Notation

Let  $A$  and  $B$  be two non-empty sets. Let  $p_{12} : A \times B \rightarrow \mathbf{R}$  be a distribution with marginal distributions  $p_1 : A \rightarrow \mathbf{R}$  and  $p_2 : B \rightarrow \mathbf{R}$ . The mutual information of  $p$  is  $d(p, p_1 p_2)$  where  $d$  denotes the relative entropy.

---

<sup>1</sup>Future editions will include.



