



**Why****Definition**

Consider two distributions on the same finite set. The *cross entropy* of the first distribution *relative* to the second distribution is the expectation of the negative logarithm of the first distribution under the second distribution.

**Notation**

Let  $R$  denote the set of real numbers. Let  $A$  be a finite set. Let  $p : A \rightarrow R$  and  $q : A \rightarrow R$  be distributions. The cross entropy of  $p$  relative to  $q$  is

$$-\sum_{a \in A} q(a) \log(p(a)).$$

We denote the cross entropy of  $p$  relative to  $q$  by  $H(q, p)$ .



