



# Relative Entropy

## 1 Why

## 2 Definition

Consider two distributions on the same finite set. The *relative entropy* of the first distribution *relative* to the second distribution is the difference of the cross entropy of the first distribution relative to the second and the entropy of the second distribution.

### 2.1 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. Let  $H(q, p)$  denote the cross entropy of  $p$  relative to  $q$  and let  $H(q)$  denote the entropy of  $q$ . The entropy of  $p$  relative to  $q$  is

$$H(q, p) - H(q).$$

Herein, we denote the entropy of  $p$  relative to  $q$  by  $d(q, p)$ .

## 3 Distance between Distributions

**Proposition 1.** *Let  $q$  and  $p$  be distributions on the same set. Then  $d(q, p) \geq 0$  with equality if and only if  $p = q$ .*

Thus  $d$  is definite, the first property of a metric.

### 3.1 Asymmetry

However,  $d$  is not a metric; for example, it is not symmetric.

**Proposition 2.**  $d(q, p) \neq d(p, q)$

### 3.2 Optimization Perspective

if we want to find a distribution  $p$  to

$$\text{minimize } d(q, p)$$

then  $p = q$  is a solution.