Baez, John C.; Fritz, Tobias

A Bayesian characterization of relative entropy. (English) Zbl 1321.94023

Theory Appl. Categ. 29, 422-456 (2014).

This paper gives a new characterization of the concept of relative entropy, aka *relative information*, *relative gain* or *Kullback-Leibler divergence*. Whenever we have two probability distributions p and q on the same set X, we define the information of q relative to p as

$$S(q,p)=\sum_{x\in X}q_x\ln\left(\frac{q_x}{p_x}\right)$$

where $q_x\ln\left(\frac{q_x}{p_x}\right)$ is set equal to ∞ when $p_x=0$, unless q_x is also 0, in which case it is set equal to 0.

Bayesian probability theory emphasizes the role of the prior so that relative entropy naturally lends itself to a Bayesian interpretation [*P. Baldi* and *L. Itti*, Neural Netw. 23, No. 5, 649–666 (2010; Zbl 1401.62225)]. The goal of this paper is to make this precise in a mathematical characterization of relative entropy. The authors consider a category $\mathtt{FinStat}$, where an object (X,q) is a finite set X gifted with a probability distribution $x\mapsto q_x$, while a morphism $(f,s):(X,q)\to(Y,r)$ is a measure-preserving function $f:X\to Y$ hand in hand with a probability distribution $x\mapsto s_{xy}$ on X for each element $y\in Y$ with the property $s_{xy}=0$ unless $f(x)=y$.

Intuitively speaking, an object of $\mathtt{FinStat}$ is to be thought of a system with some finite set of states as well as a probanility distribution on it. A morphism $(f,s):(X,q)\to(Y,r)$ is a deterministic measuring process $f:X\to Y$ mapping states of some system under measurement to those of a measuring apparatus as well as a hypothesis s meaning the probability s_{xy} that the system under measurement is in the state x given any measurement outcome $y\in Y$.

Given a morphism $(f,s):(X,q)\to(Y,r)$ in $\mathtt{FinStat}$, the authors define

$$\mathrm{RE}(f,s)=S(q,p)$$

where

$$p_x=s_{xf(x)}r_{f(x)}$$

and s is said to be *optimal* as long as the above equation gives a prior p equal to the true probability distribution q on the states of the system under measurement. It is nontrivial and rather interesting to establish the fact that

$$\mathrm{RE}:\mathtt{FinStat}\to[0,\infty]$$

where $[0,\infty]$ is thought of a category with one object, the nonnegative real numbers with ∞ as morphisms whose composition is simply addition. The functoriality of RE claims that, given

$$(X,q)\xrightarrow{(f,s)}(Y,r)\xrightarrow{(g,t)}(Z,u)$$

we have

$$\mathrm{RE}\left((g,t)\circ(f,s)\right)=\mathrm{RE}(g,t)+\mathrm{RE}(f,s)$$

The main result of this paper (Theorem 3.1), which was inspired by *D. Petz* [Acta Math. Hung. 59, No. 3–4, 449–455 (1992; Zbl 0765.46045)] in both its formulation and its proof, is that RE is, up to constant multiples, the unique functor from $\mathtt{FinStat}$ to $[0,\infty]$ obeying the following three conditions:

1. RE vanishes on morphisms with an optimal hypothesis.
2. RE is lower semicontinuous.
3. RE is convex linear.

Reviewer: Hirokazu Nishimura (Tsukuba)

MSC:

94A17 Measures of information, entropy
62F15 Bayesian inference
18B99 Special categories

Keywords:

relative entropy; Kullback-Leibler divergence; measures of information

Full Text: EMIS arXiv