UPDATING PROBABILITIES

Ariel Caticha and Adom Giffin
Department of Physics, University at Albany - SUNY, Albany, NY, USA
(e-mail: ariel@albany.edu)

Abstract

The Method of Maximum (relative) Entropy (ME) has been designed for updating from a prior distribution to a posterior distribution when the information being processed is in the form of a constraint on the family of allowed posteriors. This is in contrast with the usual MaxEnt which was designed as a method to assign, and not to update, probabilities. The objective of this paper is to strengthen the ME method in two ways.

In [1] the axioms that define ME have been distilled down to three; here the design is improved by considerably weakening the axiom that refers to independent subsystems. Instead of the old axiom which read: "When a system is composed of subsystems that are *believed* to be independent it should not matter whether the inference procedure treats them separately or jointly" we now modify it by replacing the word 'believed' by the word 'known'. As pointed out by Karbelkar and by Uffink the modified axiom is a much weaker consistency requirement, which, in their view, fails to single out the usual (logarithmic) relative entropy as the unique tool for updating. It merely restricts the form of the entropy to a one-dimensional continuum labeled by a parameter η; the resulting η-entropies are equivalent to the Renyi or the Tsallis entropies. We show that further applications of the *same* modified axiom select a unique, universal value for the parameter η and this value corresponds to the usual (logarithmic) relative entropy. The advantage of our new approach is that it shows precisely how it is that the other η-entropies are ruled out as tools for updating.

Our second concern is mostly pedagogical. It concerns the relation between the ME method and Bayes' rule. We start by drawing the distinction between Bayes' *theorem*, which is a straightforward consequence of the product rule for probabilities, and Bayes' *rule*, which is the actual updating rule. We show that Bayes' rule can be derived as a special case of of the ME method. The virtue of our derivation, which hinges on translating the information in data into constraints that can be processed by ME, is that it is particularly clear. It throws light on Bayes' rule and it shows the complete compatibility of Bayes' updating with ME updating.

References:

[1] A. Caticha, "Relative Entropy and Inductive Inference," in *Bayesian Inference and Maximum Entropy Methods in Science and Engineering*, ed. by G. Erickson and Y. Zhai, AIP Conf. Proc. **707**, 75 (2004) (arXiv.org/abs/physics/0311093).

Key Words: relative entropy, Bayes rule