Towards the Entropy-Limit Conjecture

Jürgen Landes, Soroush Rafiee Rad and Jon Williamson

Draft of July 5, 2019.

Abstract

The maximum entropy principle is widely used to determine non-committal probabilities on a finite domain, subject to a set of constraints. However, its application to infinite domains is notoriously problematic. Two strategies have been put forward for applying the maximum entropy principle on a first-order predicate language: (i) applying it to finite sublanguages and taking the pointwise limit of the resulting probabilities as the size n of the sublanguage increases; (ii) selecting a probability function on the language as a whole whose entropy on finite sublanguages of size n dominates that of other probability functions for sufficiently large n.

The entropy-limit conjecture says that, where the former approach yields determinate probabilities, the latter approach yields the same probabilities. If this conjecture is found to be true, it would provide a boost to the project of seeking a single canonical inductive logic—a project which faltered when Carnap's attempts in this direction succeeded only in determining a continuum of inductive methods.

Hitherto, the entropy-limit conjecture has been verified for languages which contain only unary predicate symbols and also for the case in which the constraints can be captured by a categorical statement of Σ_1 quantifier complexity. This paper shows that the entropy-limit conjecture also holds for categorical statements of Π_1 complexity, for various non-categorical constraints, and in certain other general situations.

Keywords: Inductive Logic, Uncertain Inference, Maximum Entropy, Objective Bayesianism, Predicate Languages

Inductive logic. Inductive logic seeks to determine how certain a conclusion proposition ψ is, on the basis of premiss propositions $\varphi_1, \ldots, \varphi_k$ that may themselves be uncertain. That is, the main task is to find Y such that

$$\varphi_1^{X_1}, \ldots, \varphi_k^{X_k} \approx\!\!\!\!| \;\psi^Y,$$

where $\approx\!\!\!\!|$ signifies the inductive entailment relation and $X_1, \ldots, X_k, Y$ are measures of certainty, such as probabilities or sets of probabilities. There are many possible semantics for inductive logic (Haenni et al., 2011). One key approach stems from the work of Carnap, who provided a continuum of inductive entailment relations (Carnap, 1945, 1971, 1980; Paris and Vencovská, 2015).[1] An

[1]While the discussion is often phrased in terms of probabilities in an inductive entailment relation, one may also interpret the probabilities as degrees of belief representing an agent's

alternative approach, which is the focus of this paper, is to apply the maximum entropy principle of Jaynes (1957a,b). According to this approach, one should consider, from all the probability functions that satisfy the premisses, all those with maximum entropy, and let Y be the set of probability values that these functions give to ψ.

If the underlying logical language is a finite propositional language then this latter proposal is rather straightforward to implement and has many nice properties (Paris, 1994). However, if the language is a first-order predicate language $\mathcal{L}$ with infinitely many constant symbols, certain intriguing questions arise. In particular, there are two main ways to implement the proposal in the predicate-language case, and it is not entirely clear as to whether the resulting inductive logics agree.

The entropy-limit approach. Barnett and Paris (2008) proceed as follows: (i) reinterpret the premisses as constraints on the probabilities of sentences of a *finite* predicate language $\mathcal{L}_n$ that has n constant symbols; (ii) determine the function P^n that maximises entropy on this finite language, subject to constraints imposed by the reinterpreted premisses; (iii) draw inductive inferences using the *entropy limit* function P^∞ defined by $P^\infty(\theta) \stackrel{\text{df}}{=} \lim_{n\to\infty} P^n(\theta)$ for sentences θ of $\mathcal{L}$.

The maximal-entropy approach. Williamson (2008, 2017), on the other hand, proceeds differently: (i) consider probability functions defined on the language $\mathcal{L}$ as a whole; (ii) deem one probability function P to have greater entropy than another function Q if $H_n(P)$, where H_n is the entropy function on the finite sublanguage $\mathcal{L}_n$, dominates $H_n(Q)$ for sufficiently large n; (iii) draw inductive inferences using those *maximal entropy* functions $P^\dagger$, from all those probability functions on $\mathcal{L}$ that satisfy the premisses, that have maximal entropy (i.e., no other function satisfying the premisses has greater entropy).

The entropy-limit approach has the advantage that it is more constructive, so it is typically easier to calculate the probabilities required for inductive inference. The maximal-entropy approach is apparently more general. This is because the entropy-limit approach faces what is known as the finite model problem: premisses can become inconsistent when reinterpreted as applying to a finite domain. Thus, there are cases where the maximal-entropy approach gives a solution but the entropy-limit approach does not.

These approaches to inductive logic would be strengthened if it could be shown that they give the same results where they are both applicable. Then one could use the maximal entropy approach to provide a general semantics for inductive logic, but use the entropy-limit approach where a more constructive approach is helpful.

The entropy-limit conjecture. This is the conjecture that

> where P^∞ exists and satisfies the constraints, it is the unique function with maximal entropy, i.e., $P^\dagger = P^\infty$ (Williamson, 2017, p. 191).

credal state for the purposes of a maximally accurate representation and/or maximally good decisions.

What is known so far. Rafiee Rad (2009, Theorem 29) shows that the entropy-limit conjecture is true in the case of a predicate language that contains only unary predicate symbols. Moreover, Rafiee Rad (2019, Corollary 1) shows that the conjecture is true in the categorical Σ_1 case, i.e., the case in which the premiss propositions $\varphi_1, \ldots, \varphi_k$ are all Σ_1 statements and no uncertainty attaches to these propositions, $X_1 = \cdots = X_k = 1$. These premisses can equivalently be expressed by $\bigwedge_{i=1}^k \varphi_i^1$ with $\bigwedge_{i=1}^k \varphi_i \in \Sigma_1$.

Furthermore, Rafiee Rad (2009, §4.3) shows that there exist cases in which no function has maximum entropy: for any probability function satisfying the premiss $\exists x \forall y Rxy \in \Pi_2$ there is another probability function with greater entropy that also satisfies that premiss. On the other hand, Rafiee Rad (2009, §4.1) shows that there are cases with a single premiss $\varphi \in \Pi_2$ in which P^∞ does not exist but $P^\dagger$ does. Rafiee Rad (2009, Theorem 29) shows that if the premisses involve only unary predicate symbols, then both P^∞ and $P^\dagger$ exist, are unique and coincide, i.e., the entropy-limit conjecture holds in such a situation.

This paper. This paper contains three main sets of results.

First, we show that for $X = 1$ and all $\varphi \in \Pi_1$ such that $P_=(\varphi) > 0$ (where $P_=$ is the equivocator function assigning every state description on each finite sublanguage the same probability), the entropy-limit conjecture holds, $P^\dagger = P^\infty$. This is achieved by constructing P^∞ and showing that it is the unique function with maximum entropy. For $X = 1$ and $\varphi \in \Pi_1$ such that $P_=(\varphi) = 0$ we also show that the conjecture holds. Unfortunately, we do not (yet!) know how to construct P^∞ in this case.

Second, we show that, if the entropy limit P^∞_φ exists for φ^1 and $P^\infty_{\neg\varphi}$ exists for $\neg\varphi^1$ for an arbitrary sentence $\varphi \in S\mathcal{L}$, then the entropy limit for the non-categorical premisses $\varphi^c, \neg\varphi^{1-c}$ (where $c \in (0,1)$) exists and is obtained by Jeffrey updating with the categorical entropy limits: $P^\infty = c \cdot P^\infty_\varphi + (1-c) \cdot P^\infty_{\neg\varphi}$. On top of that, if $P^\dagger_\varphi$ is the unique maximum entropy function for the premiss φ^1 and if $P^\dagger_{\neg\varphi}$ is the unique maximum entropy function for the premiss $\neg\varphi^1$, then $P^\dagger = P^\infty = c \cdot P^\dagger_\varphi + (1-c) \cdot P^\dagger_{\neg\varphi}$ for the non-categorical premisses $\varphi^c, \neg\varphi^{1-c}$.

Third, we show that the entropy-limit conjecture is true under some rather general conditions: if P^n converges sufficiently rapidly to P^∞ and either dominates P^∞ in n-entropy or P^∞ is the equivocator function. The proof of this appeals to several basic information-theoretic results. These results can be used to check whether the entropy-limit conjecture holds in a particular case without constructing $P^\dagger$.

Conclusions. These results extend the range of situations in which the entropy-limit conjecture in known to hold and provide inductive support for the truth of the conjecture in general. They thus provide a boost to a new research programme for inductive logic that differs in important ways from Carnap's programme.

References

Barnett, O. and Paris, J. (2008). Maximum entropy inference with quantified knowledge. *Logic Journal of the IGPL*, 16(1):85–98.

Carnap, R. (1945). On inductive logic. *Philosophy of Science*, 12(2):72–97.

Carnap, R. (1971). A basic system of inductive logic part 1. In Carnap, R. and Jeffrey, R. C., editors, *Studies in inductive logic and probability*, volume 1, pages 33–165. University of California Press, Berkeley CA.

Carnap, R. (1980). A basic system of inductive logic part 2. In Carnap, R. and Jeffrey, R. C., editors, *Studies in inductive logic and probability*, volume 2, pages 7–155. University of California Press, Berkeley CA.

Haenni, R., Romeijn, J.-W., Wheeler, G., and Williamson, J. (2011). *Probabilistic logics and probabilistic networks*. Synthese Library. Springer, Dordrecht.

Jaynes, E. T. (1957a). Information theory and statistical mechanics. *The Physical Review*, 106(4):620–630.

Jaynes, E. T. (1957b). Information theory and statistical mechanics II. *The Physical Review*, 108(2):171–190.

Paris, J. and Vencovská, A. (2015). *Pure inductive logic*. Cambridge University Press, Cambridge.

Paris, J. B. (1994). *The uncertain reasoner's companion*. Cambridge University Press, Cambridge.

Rafiee Rad, S. (2009). *Inference Processes for on First Order Probabilistic Languages*. PhD thesis, Manchester Institute for Mathematical Sciences.

Rafiee Rad, S. (2019). On Probabilistic Models π_1 of Theories. forthcoming.

Williamson, J. (2008). Objective Bayesian probabilistic logic. *Journal of Algorithms in Cognition, Informatics and Logic*, 63:167–183.

Williamson, J. (2017). *Lectures on Inductive Logic*. Oxford University Press, Oxford.