International Journal of Approximate Reasoning 50 (2009) 627–641

Contents lists available at ScienceDirect

International Journal of Approximate Reasoning

journal homepage: www.elsevier.com/locate/ijar

An extension of chaotic probability models to real-valued variables

P.I. Fierens

Department of Physics and Mathematics, Instituto Tecnológico de Buenos Aires (ITBA), Av. Eduardo Madero 399, Buenos Aires, C1106ACD Capital Federal, Argentina

ARTICLE INFO

Article history:
Received 17 December 2007
Accepted 30 September 2008
Available online 11 October 2008

Keywords:
Imprecise probabilities
Foundations of probability
Chaotic probability models
Frequentist interpretation
Simulation

ABSTRACT

In a recent series of papers, Fine and colleagues [P.I. Fierens, T.L. Fine, Towards a frequentist interpretation of sets of measures, in: G. de Cooman, T.L. Fine, T. Seidenfeld (Eds.), Proceedings of the Second International Symposium on Imprecise Probabilities and Their Applications, Shaker Publishing, 2001; P.I. Fierens, T.L. Fine, Towards a chaotic probability model for frequentist probability, in: J. Bernard, T. Seidenfeld, M. Zaffalon (Eds.), Proceedings of the Third International Symposium on Imprecise Probabilities and Their Applications, Carleton Scientific, 2003; L.C. Rêgo, T.L. Fine, Estimation of chaotic probabilities, in: Proceedings of the Fourth International Symposium on Imprecise Probabilities and Their Applications, 2005] have presented the first steps towards a frequentist understanding of sets of measures as imprecise probability models which have been called *chaotic models*. Simulation of the *chaotic variables* is an integral part of the theory.

Previous models, however, dealt only with sets of probability measures on finite algebras, that is, probability measures which can be related to variables with a finite number of possible values. In this paper, an extension of chaotic models is proposed in order to deal with the more general case of real-valued variables. This extension is based on the introduction of real-valued test functions which generalize binary-valued choices in the previous work.

1. Introduction

In a series of papers [5,6], we presented the first steps towards a frequentist interpretation of sets of measures as probability models, which we have called *chaotic probability models* in order to distinguish them from other plausible interpretations. This work was coherently presented in [4] and extended by Rêgo and Fine in [10]. In our previous work, we presented chaotic models as simply sets of probability measures whose domain is a finite set of events. In this sense, we may associate chaotic probability models to discrete "random"[1] variables with finite range (e.g., the outcome of the flipping of a coin or the tossing of a die). In this paper, we present a simple approach to the extension of chaotic probability models to real-valued variables (e.g., tomorrow's minimum temperature).

The paper is organized as follows: Section 2 presents some concepts of the previous work which are needed for this paper. In Section 3, we provide the basic motivation behind the model which is described in Section 4. In the latter Section, we also show that such a model is plausible. Section 5 is devoted to present extensions of this framework to include the concepts of visibility and temporal homogeneity defined in previous works. In Section 6, we provide an example of modelling and simulation with chaotic probabilities. Finally, in Section 7 we discuss the results presented in this paper and suggest future lines of work. The Appendices contain proofs of cited results.

E-mail address: pfierens@itba.edu.ar

[1] We use quotation marks to denote the difference between these *chaotic* variables and the usual understanding of random variables.

0888-613X/$ - see front matter © 2008 Elsevier Inc. All rights reserved.
doi:10.1016/j.ijar.2008.09.003

2. Previous work: variables with finite range

We need to recall the interpretation of chaotic probability models for variables with finite range [6,4,10].

2.1. An instrumental description of the model

Chaotic probabilities are intended to provide a frequentist interpretation of sets of probability measures as models. However, the problem is that we find it difficult to step outside the frame of mind of standard probability. How can we describe a time series of "random" (which are not "random" in a standard sense) variables which "reveals", from a frequentist point of view, a set of probabilities instead of a single distribution? Furthermore, where can we find such a time series in real world data? In order to face this problem, we have found it useful to employ an instrumental (that is, without commitment to reality) description of chaotic probability models which allows us to use a number of already-developed mathematical tools from the area of standard probability. In this section, we review the instrumental description that was presented in earlier works (see, e.g. [6,10,9]) and is basically preserved in this paper.

Let $\mathbf{X}$ be a sample space. We denote by $\mathbf{X}^*$ the set of all finite sequences of elements taken in $\mathbf{X}$. A particular sequence of n samples from $\mathbf{X}$ is denoted by $x^n = \{x_1, x_2, \ldots, x_n\}$. $\mathbf{P}$ denotes the set of all measures on the power set of $\mathbf{X}$. A chaotic probability model $\mathbf{M}$ is a subset of $\mathbf{P}$ and models the "marginals" of some process generating sequences in $\mathbf{X}^*$.

Let $F : \mathbf{X}^* \to \mathbf{M}$ be a function that, for each finite string x^{k-1}, returns a measure $\nu \in \mathbf{M}$. Given any $n \in \mathbb{N}$, consider the generation of a sequence x^n of length n by the following algorithm:[2]

```
FOR k = 1 TO k = n
(1) Choose ν_k = F(x^{k-1}) ∈ M;
(2) Generate x_k according to ν_k.
```

For any $k \leqslant n$, F determines the probability distribution of the *potential* kth outcome X_k of the sequence,

$$(\forall \mathbf{A} \subseteq \mathbf{X}) \; P(X_k \in \mathbf{A} | X^{k-1} = x^{k-1}) = \nu_k(X_k \in \mathbf{A}).$$

The probability of a particular realization x^n of a sequence of random variables X^n is given by

$$P(X_1 = x_1, \ldots, X_n = x_n) = \prod_{k=1}^{n} \nu_k(X_k = x_k).$$

We denote by $\mathbf{M}^*$ the family of all such process measures P, one for each possible function F. From the analysis of a single data sequence x^n of any finite length n, we do not expect in general to be able to infer a single $P \in \mathbf{M}^*$ or even a small subset of $\mathbf{M}^*$, what we call a **fine-grained picture** of the source. On the contrary, we expect our knowable **operational quantities to be** (**large**) **subsets of** $\mathbf{M}^*$ which provide an appropriate **coarse-grained** description of the source. In terms of our instrumental understanding of chaotic probability models, it may be possible relate the gap between the coarse-grained and the fine-grained descriptions to the complexity of the function F, as suggested in [10].

While it is true that, under the description given here, $\mathbf{M}$ can be understood as a set of Markov–Kernels together with a set of start-distributions, we insist in that the description is instrumental and we do not place any emphasis on either the probability or provability of the reality of that description.

2.2. Data analysis and estimation

We begin the study of a sequence $x^n \in \mathbf{X}^*$ by decomposing it into several subsequences. These subsequences are selected by rules that satisfy the following:

Definition 1. A computable function $\psi : \mathbf{X}^* \to \{0, 1\}$ is a **causal subsequence selection rule** (also known as a Church place selection rule) if for any $x^n \in \mathbf{X}^*$, x_k is the jth term in the generated subsequence $x^{\psi,n}$, of length $\lambda_{\psi,n}$, whenever

$$\psi(x^{k-1}) = 1, \quad \sum_{i=1}^{k} \psi(x^{i-1}) = j, \quad \lambda_{\psi,n} = \sum_{k=1}^{n} \psi(x^{k-1}).$$

The introduction of causality makes good sense when we think of the data sequence as one indexed by physical time (e.g., a record of daily maximum temperatures). Causality may not make much sense in other contexts. Let $\Psi = \{\psi_\alpha\}$ be a set of causal subsequence selection rules. For each $\psi \in \Psi$, we study the behavior of the relative frequency of marginal events along the chosen subsequence. That is, given x^n and a selection rule $\psi \in \Psi$ we determine the **frequentist empirical** (**relative frequency**) **measure** $\bar{\mu}_{\psi,n}$ along the subsequence $x^{\psi,n}$ through

[2] We denote the empty string by x^0.