Thai Journal of Mathematics
Volume 5(2007) Number 1 : 137–142
www.math.science.cmu.ac.th/thaijournal

A Continuous Lattice Approach to Random Sets

Hung T. Nguyen

Abstract : This paper is about modeling of coarse data in statistics using set-valued random elements. We consider also the case of fuzzy set-valued random elements for modeling of perception-based information. These are typically random elements with values in some functional spaces. Unlike the standard approach in stochastic processes, we emphasize the lattice approach leading to compact, separable metric spaces.

Keywords : Choquet theorem, continuous lattices, random sets, random fuzzy sets, upper semicontinuous random functions.

1 Introduction

In statistics, set-valued mappings appear at the very beginning of the theory, namely as *sampling designs in finite populations. Formally, they are finite random sets.* However, the emphasis was on different types of samplings rather than formulating the concept of random sets rigorously as random elements.

In stochastic geometry, random sets appear naturally. A satisfactory theory of random closed sets on locally compact, Hausdorff and second countable (LCHS) was developed by G. Matheron in 1975 in the context of integral geometry and morphology.

With the advances of technology, more complex types of data are considered for statistical analysis. These include coarse data , i.e. low quality data, such as censored or grouped data in survival analysis, indirect observations in bioinformatics, and perception-based information in intelligent systems.

While random closed sets generalize random vectors in multivariate statistical analysis, a more general concept of random fuzzy sets provides models for imprecise observation processes. The latter are nothing else than upper semicontinuous functions taking values in the unit interval. Efforts to define upper semicontinuous random functions in the literature (see e.g. Li et al, 2002) was not satisfactory as they only led to a restrictions of the space of upper semicontinuous functions (usc) with compact supports. This due to the fact that the approach was based on the search for a metric on usc space , similar to $C[0,1]$, $D[0,1]$ (see e.g. Billingsley, 1968), in order to toplologize it. Now, the hit-or-miss topology on the space of closed sets $\mathcal{F}(X)$ of a LCHS space X is *compact*, Hausdorff and second countable (and hence metrizable) , it is expected that $USC(X)$, the space of upper semicontinuous functions, defined on X with values in $[0,1]$, should be similar.

This happens to be the case as we will proceed to show .

2. What is a random set ?

A random set is a set obtained at random ! If we view a set obtained at random as the outcome or realization of a random phenomenon, then we need to model the observation mechanism as a random element, defined on some probability space, $(\Omega, \mathcal{A}, P)$, with values as subsets of some set X. This requires the specification of some appropriate $\sigma-$field on that collection of subsets. For finite X, the situation is clear. For LCHS spaces such as $\mathbb{R}^d$, the situation is not so clear ! Now since singletons are closed sets, random closed sets generalize multivariate statistical situation. Thus we consider $\mathcal{F}(X)$. The problem is to specify a "good" σ-field on $\mathcal{F}(X)$. Of course the standard way is to topologize $\mathcal{F}(X)$ and then talke its Borel σ-field $\sigma(\mathcal{F})$. Then a random closed set is a $\mathcal{A}-\sigma(\mathcal{F})-$ measurable mapping in this framework. Keeping in mind that, for random vectors, probability measures on $\sigma(\mathbb{R}^d)$ are in one-to-one correspondence with distribution functions (via the Lebesgue-Stieltjes theorem), one should consider $\sigma(\mathcal{F})$ so that a counter-part of this theorem exists. This was acheived by Matheron (1975) as we recall here briefly.

Let X be a LCHS space. We denote by $\mathcal{F}(X), \mathcal{G}(X), \mathcal{K}(X)$ or simply $\mathcal{F}, \mathcal{G}, \mathcal{K}$ the spaces of closed, open and compact subsets of X, respectively. The so-called *hit-or-miss topology* on $\mathcal{F}$ is generated by the base consisting of

$$\mathcal{F}^K_{G_1,\ldots,G_n} = \mathcal{F}^K \cap \mathcal{F}_{G_1} \cap \mathcal{F}_{G_2} \cap \ldots \cap \mathcal{F}_{G_n} \text{ , for } n \in \mathbb{N} \text{ , } K \in \mathcal{K} \text{ , } G_i \in \mathcal{G}$$
$$\text{where } \mathcal{F}_A = \{F \in \mathcal{F} \text{ , } F \cap A \neq \varnothing\} \text{ , } \mathcal{F}^A = \{F \in \mathcal{F} \text{ , } F \cap A = \varnothing\}$$

With this topology, $\mathcal{F}$ is a *compact*, Hausdorff and second countable topological space and hence metrizable. When $X = \mathbb{R}^d$, a compatible metric on $\mathcal{F}$ is the stereographical metric (see Rockaffelar and Wets, 1984) . For general LCHS spaces, concrete metrics are obtained similarly by using Alexandroff compactification (see Wang and Wei, 2007).

The *Choquet theorem* on $\mathcal{F}(X)$ is this. A set-function $T : \mathcal{K} \to [0,1]$ is called a *capacity functional* if it satisfies :

(i) $T(\varnothing) = 0$

(ii) If $K_n \searrow K$ in $\mathcal{K}$ then $T(K_n) \searrow T(K)$

(iii) T is alternating of infinite order, i.e., T is monotone increasing and for $K_i \in \mathcal{K}$, $i = 1, 2, \ldots n \geq 2$,

$$T(\cap_{i=1}^n) \leq \sum_{\varnothing \neq I \subseteq \{1,2,\ldots,n\}} (-1)^{|I|+1} T(\cup_{i \in I} K_i)$$
where $|I|$ denotes the cardinality of the set I .

Theorem (Choquet): Let $T : \mathcal{K} \to [0,1]$. There exists uniquely a probability measure Q on $\sigma(\mathcal{F})$ satisfying $Q(\mathcal{F}_K) = T(K)$ for $K \in \mathcal{K}$, if and only if T is a capacity functional.

3. Another look at the hit-or-miss topology.

For any topological space X, $(\mathcal{F}, \subseteq)$ is a complete lattice with

$$\wedge\{F_i : i \in I\} = \cap\{F_i : i \in I\}, \text{ and } \vee\{F_i : i \in I\} = \text{the closure of } \cup\{F_i : i \in I\}$$

Also, $(\mathcal{F}, \supseteq)$,with reverse order, is a complete lattice. However, the difference is that, for locally compact X, while the lattice $(\mathcal{F}, \subseteq)$ is not *continuous*, $(\mathcal{F}, \supseteq)$ is .

This concept of continuity for lattices was introduced by Gierz et al in their first edition of *A Compendium of Continuous Lattices* (Springer-Verlag, 1980), after Matheron's book (1975). The basic novel concept in their work is to realize that to every partial order R (or $\leq$), one can associate another relation (not reflexive), denoted as $\ll$, called the "*way below*" relation.

Specifically, on a complete lattice $(L, \leq)$, x is said to be *way below* y, denoted as $x \ll y$, if for directed set $D \subseteq L$, the relation $y \leq \sup D$ always implies $\exists d \in D$ such that $x \leq d$.

With this associated "way below" relation, the complete lattice $(L, \leq)$ is said to be *continuous* if it satisfies the following approximation property :

$$x = \sup\{u \in L : u \ll x\} \text{ for all } x \in L.$$

One of the main results in "The Compendium" is this. If $(L, \leq)$ is a continuous lattice, then there exists a canonical topology, called the *Lawson topology*, denoted as $\lambda(L)$, making L a compact and Hausdorff topological space. This topology is generated by sets of the form (a subbase):

$$\Uparrow x = \{y \in L : x \ll y\} \text{ or } L\backslash \uparrow x = \{y \in L : x \nleq y\} \qquad \text{for } \ x \in L.$$

It turns out that our $\mathcal{F}(X)$, with appropriate order relation, is continuous (for locally compact X) whose Lawson topology coincides with its hit-or-miss topology. The additional property of second countability of X implies the same property for the Lawson topology.

First, the lattice $(\mathcal{F}, \subseteq)$, for locally compact X, is not *continuous* . This can be checked using the fact that : $x \ll y$ iff for any $A \subseteq L$, $y \leq \sup A$ implies the existence of some finite subset $B \subseteq A$ such that $x \leq \sup B$.

To see that, take $X = \mathbb{R}$, we notice that if $A \ll \mathbb{R}$, then for any subset B of A i.e $B \subseteq A$, we also have $B \ll \mathbb{R}$ (just use the equivalent condition of the way-below relation). Then any singleton closed set, e.g $\{0\}$, is not way-below $\mathbb{R}$. Indeed, $\bigvee_{n \in \mathbb{N}} \{(-\infty, -1/n] \cup [1/n, \infty)\} = \mathbb{R}$, but we can not find any finite subset A of $\{(-\infty, -1/n] \cup [1/n, \infty)\}_{n \in \mathbb{N}}$ such that $\{0\} \subseteq \bigvee A$.Therefore, the only closed set that is way-below $\mathbb{R}$ is the empty set. Then

$\sup\{A \in \mathcal{F}(\mathbb{R}) : A \ll \mathbb{R}\} = \sup\{\emptyset\} = \emptyset \neq \mathbb{R}$.

However, $(\mathcal{F}, \supseteq)$, with X locally compact, is a continuous lattice. For details of proofs, see Nguyen and Tran (2007). If we examine its Lawson topology, we recognize that it is generated by the subbase consisting of sets of the form $\mathcal{F}^G$, $\mathcal{F}_K$, for $K \in \mathcal{K}$ and $G \in \mathcal{G}$, so that it coincides with the hit-or-miss topology.

Thus, random closed sets on LCHS spaces are random elements when equiping $\mathcal{F}$ with its Lawson topology. The upshot is this. When extending $\mathcal{F}(X)$ to $USC(X)$, we should look at lattice structures of $USC(X)$ in the search for a "good" topology on it.

4. USC(X) as a continuous lattice.

Biased by our interests in modeling of perception-based data (in problems such as "social dynamics modeling") which are expressed in natural languages where meaning representation needs to be modeled, we consider the space $USC(X)$ of upper semicontinuous functions, defined on a LCHS space X, with values in $[0, 1]$. The motivation is this. The mathematical model for fuzzy concepts in our natural language is *fuzzy set theory*. Each fuzzy concept is characterized as a membership function $f : X \to [0, 1]$. These membership functions generalize indicator functions of ordinary sets. Since indicator functions of closed sets are usc, we are led to consider the whole space $USC(X)$.

The need to define rigorously random elements with values in $USC(X)$ has started way back in the mid 80's . However, in one hand, there was no mentioned of Matheron's work on closed sets as a guidance for extension, and, on the other hand, viewing $USC(X)$ as a function space, the approach to topologize it followed the same path in stochastic processes theory, namely searching directly for metrics, like for $C[0, 1], D[0, 1]$,see e.g. Billingsley (1968). As a result, the space $USC(X)$ has to be restricted to the subset of usc functions with compact support, see e.g. Li et al (2002).

We are going to take a continuous lattice approach to this problem leading to a more general and satisfactory theory of the so-called random fuzzy sets in the field of *Soft Computing.* Specifically, with the Lawson topology on the continuous lattice $(USC(X), \geq)$, the space $USC(X)$ is a *compact*, Hausdorff and second countable (hence separably metrizable) topological space whose compatible metrics can be identified concretely. Moreover, $\mathcal{F}(X)$ is embedded into $USC(X)$.

Remark. Recall that a fuzzy subset of X is a mapping f from X to the unit interval $[0, 1]$. For such mappings to generalize closed sets, they have to be upper semicontinuous (usc) so that their level-sets $A_\alpha(f) = \{x \in X : f(x) \geq \alpha\}$, $\alpha \in [0, 1]$, are closed. Thus, formally, by a random fuzzy set, we mean a fuzzy subset whose membership function is usc. To be rigorous, we need to topologize the space $USC(X)$ (from now on we suppose functions takes values in $[0, 1]$).

If $f : X \to [0, 1]$, then f can be identified with the level-sets $A_\alpha(f)$ for $\alpha \in \mathbb{Q}_1 = \mathbb{Q} \cap [0, 1]$ (rationals in $[0, 1]$). Thus, the mapping $\psi : USC(X) \to \mathcal{F}_\alpha$, the countable cartesian product of identical copies $\mathcal{F}_\alpha$ of $\mathcal{F}$, sending f to $(A_\alpha(f)$ $,\alpha \in \mathbb{Q}_1)$, is an embedding. Thus one hopes to induce a topology on $USC(X)$ from the product topology of $\mathcal{F}_\alpha$ which is a compact and second countable space. Unfortunately, the induced topology does not make $USC(X)$ a compact space. For a counter example, see Nguyen et al (2006).

First, like $\mathcal{F}(X)$, with the pointwise order $\leq$, $USC(X)$ is a complete lattice with

$$\underset{j\in J}{\wedge} f_j = \inf_{j\in J} f_j \text{ , where } f \leq g \text{ means } \nabla x \in X \text{ , } f(x) \leq g(x)$$

Now, noting that the function $1 : X \to [0,1] : x \to 1$ is usc. As a result, $(USC(X), \leq)$ is a $\wedge-$ semilattice with a top (1), hence it is a complete lattice, where

for any $A \subseteq USC(X)$, $\vee A = \wedge\{g \in USC(X) : g$ is an upper bound of $A\} = \wedge\{g \in USC(X) : g \geq f$ for some $f \in A\}$

Again, with this order, $(USC(X), \leq)$ is not continuous. And, as expected, the reverse order lattice $(USC(X), \geq)$ is continuous. Its Lawson topology has a subbase consisting of sets of the form

$$\{f : f(y) < r, \ \forall y \in K\} \text{ , for } r \in (0,1] \text{ and } K \text{ (compact)} \subseteq X$$

together with

$$\{f : \exists x \in X \text{ such that } g(x) < f(x)\}, \text{ where } g \in USC(X).$$

References.

[1] Billingsley, P. (1968). *Convergence of Probability Measures.* J.Wiley

[2] Gierz, G. et al (2003). *Continuous Lattices and Domains.* Cambridge University Press, UK.

[3] Li, S., Ogura,Y. and Kreinovich, K. (2002). *Limit Theorems and Applications of Set-Valued and Fuzzy Set-valued Random Variables.* Kluwer Academic.

[4] Matheron, G. (1975). *Random Sets and Integral Geometry.* J.Wiley.

[5] Nguyen, H.T., Ogura, Y., Tasena, S. and Tran, H. (2006). A note on random upper semicontinuous functions. *Soft Methods for Integrated Uncertainty Modeling* (Lawry, J. et al, Eds.), pp. 129-135. Springer-Verlag.

[6] Nguyen, H.T., Wang,Y. and Wei, G. (2007). On Choquet theorem for random upper semicontinuous functions. To appear in *International Journal of Approximate Reasoning.*

[7] Nguyen, H.T. and H. Tran (2007). On a continuous lattice approach to modeling of coarse data in systems analysis. To appear in *Journal of Uncertain Systems.*

[8] Norberg, T. (1989). Existence theorems for measures on continuous posets, with applications to random set theory. *Math. Scand. (64), pp. 15-51.*

[9] Rockaffelar, R.T. and Wets, R. (1984). Variational systems: An introduction. *Lecture Notes in Mathematics* 1091, pp. 1-54. Springer-Verlag.

[10] Wang, Y. and Wei,G. (2007). On metrization of Matheron topology using Alexandroff compactification. To appear in *International Journal of Approximate Reasoning.*

(Received 12 February 2007)

Hung T. Nguyen
Department of Mathematical Sciences,
New Mexico State University,
Las Cruces, NM 88003-8001 , USA
E-mail: hunguyen@nmsu.edu