International Institute for
Applied Systems Analysis
IIASA www.iiasa.ac.at

# Random Semicontinuos Functions

**Salinetti, G. and Wets, R.J.-B.**

**IIASA Working Paper**

**WP-86-047**

**August 1986**

# Working Paper

## RANDOM SEMICONTINUOUS FUNCTIONS

*Gabriella Salinetti*[*]
*Roger J-B. Wets*[**]

August 1986
WP-86-47

International Institute for Applied Systems Analysis
A-2361 Laxenburg, Austria

NOT FOR QUOTATION
WITHOUT THE PERMISSION
OF THE AUTHORS

# RANDOM SEMICONTINUOUS FUNCTIONS

*Gabriella Salinetti**
*Roger J-B. Wets***

August 1986
WP-86-47



INTERNATIONAL INSTITUTE FOR APPLIED SYSTEMS ANALYSIS
A-2361 Laxenburg, Austria

## FOREWORD

The study of dynamic stochastic optimization problems is often hampered by a number of technical complexities due to the "classical" mathematical framework for Stochastic Processes. Conceptually, as well as technically, the classical set-up is inappropriate for studying infima, allowing for approximations, etc. Here the authors introduce an alternative approach which smooths out most of these difficulties and gives the study of stochastic processes, in particular the study of functions of stochastic processes, another perspective. This work serves as background to IIASA's efforts in developing algorithmic procedures for stochastic programming and stochastic control problems.

Alexander B. Kurzhanski
Chairman
System and Decision Sciences Program

**CONTENTS**

# RANDOM SEMICONTINUOUS FUNCTIONS

*Gabriella Salinetti*[*] *and Roger J-B. Wets*[**]

These notes introduce a new approach for the description, and the analysis of stochastic phenomena. It parts company with the classical approach when the realizations are infinite dimensional in nature. We shall be mostly concerned with questions of convergence and the description of the probability distributions associated to such phenomena.

We begin with a brief review of the classical theory for stochastic processes, bringing to the fore some of the shortcomings of such an approach. In the second part of the paper we deal with the epigraphical approach that relies on the modeling of the "paths" of the stochastic phenomena by semicontinuous functions. We conclude with a discussion and a comparison of the two theories, and the application to the convergence of stochastic processes.

## 1. STOCHASTIC PROCESSES: THE CLASSICAL VIEW

A *stochastic process*, with values in the extended reals, is a collection $\{X_t, t \in T\}$, of extended real-valued random variables indexed by T and defined on a probability space $(\Omega, \mathcal{A}, \mu)$. Here, and in the next few sections, we take T to be a subset of $\mathbf{R}$. It is a *discrete process*, if T is a discrete subset of $\mathbf{R}$, in which case, without loss of generality we can always identify T with $\mathbf{Z}$ (the integers) or $\mathbf{N}$ (the natural numbers).

The probability measure associated to $\{X_t, t \in T\}$ is usually defined in terms of its finite dimensional distributions. For any finite subset $\{t_1, \dots, t_q\} \subset T$, the q-dimensional random vector

$$(X_{t_1}, X_{t_2}, \dots, X_{t_q})$$

defined on $(\Omega, \mathcal{A}, \mu)$ with values in $\overline{\mathbf{R}}^q = [-\infty, \infty]^q$ has the probability measure de-

*Università di Roma - La Sapienza, I-00185 Roma, Italy.
**University of California, Davis, California 95616, USA.

fined by the correspondence

$$P_{t_1,\dots,t_q}(B) := \mu\{\omega \in \Omega \mid (X_{t_1}(\omega), \dots, X_{t_q}(\omega)) \in B\}$$

where $B \in \bar{\mathcal{B}}^q$ is a Borel subset of $\bar{\mathbf{R}}^q$. The family of probability measures

$$\{P_I, I \in \mathcal{I}(T)\}$$

where $\mathcal{I}(T)$ is the collection of all finite subsets of T, is the family of ***finite dimensional distributions*** of the stochastic process $\{X_t, t \in T\}$.

This approach is attractive for a number of reasons, in particular because of its immediate simplicity, at least as far as the definition is concerned. But in many cases, the price must be paid at a later stage, and sometimes there are technical, and even conceptual, difficulties that can be directly traced back to this "finite dimensional" approach to stochastic processes.

In a functional setting, the classical approach leads to the following framework. To every $\omega \in \Omega$, there corresponds a function (***sample path, realization***):

$$t \mapsto X_t(\omega) : T \to \bar{\mathbf{R}} \ .$$

The stochastic process $\{X_t, t \in T\}$ can be viewed as map from $\Omega$ into $\bar{\mathbf{R}}^T$; we now identify $\bar{\mathbf{R}}^T$ with the space of all extended real-valued functions defined on T. The family of finite dimensional distributions assigns a probability to all subsets of the type

$$B_I := \{x \in \bar{\mathbf{R}}^T \mid (x(t_1), \dots, x(t_q)) \in B\}$$

where $B \in \bar{\mathcal{B}}^q$, $I = \{t_1, \dots, t_q\} \in \mathcal{I}(T)$. The sets $B_I$ are ***cylinders*** (with finite dimensional base) and they form a field on $\bar{\mathbf{R}}^T$. The finite dimensional distributions assign a measure to each set of this field through the identity.

$$P(B_I) = P_I(B) \ .$$

It can be shown, as done by Daniell and Kolmogorov, that this measure P can be uniquely extended to the $\sigma$-field, denoted by $\bar{\mathcal{R}}^T$, generated on $\bar{\mathbf{R}}^T$ by the family of cylinders. We can thus pin down a unique probability measure associated to the stochastic process $\{X_t, t \in T\}$. From this viewpoint, two stochastic processes are then ***equivalent*** if they have the same finite dimensional distributions, they identify the same probability measure on $\bar{\mathbf{R}}^T$.

## 2. SOME QUESTIONS, SOME EXAMPLES

One of the shortcomings of this approach is that no attention is paid to (possible) topological properties of the realizations of the process. In many applications, we may be interested in developing a calculus for processes that have very specific properties, whose paths may very well belong to a subset of $\overline{\mathbf{R}}^T$ of measure zero. The two following examples illustrate many of the difficulties.

EXAMPLE 2.1 Suppose $V : \Omega \to (0, \infty)$ is a random variable with continuous distribution function. For all $t \in \mathbf{R}$, prob $[V = t] = 0$. Let $T = \mathbf{R}_+$ and $\{Y_t, t \in T\}$, $\{Y'_t, t \in T\}$ be two stochastic processes such that

$$\text{for all } \omega \in \Omega: \quad Y_t(\omega) = Y'_t(\omega) = 0$$

$$\text{except that:} \quad Y_t(\omega) = -1 \quad \text{if} \quad V(\omega) = t$$

These two processes are equivalent, although the realizations of $\{Y'_t\}$ are continuous with probability 1, and those of $\{Y_t\}$ are continuous with probability 0.

One may be templed to view the phenomena illustrated by Example 2.1 as just another example of the fact that random variables that have the same distribution are not necessarily almost surely equal. But in this case there is something more that enters into play. Let C(T) denote the ***set of continuous functions*** defined on T and values in $\overline{\mathbf{R}}$. Thus, we could reformulate our earlier observation, in the following terms:

$$\mu[Y'_{\cdot} \in C(T)] = 1 \ , \quad \text{and} \ \mu[Y_{\cdot} \in C(T)] = 0 \ ,$$

but, as we shall now see, neither C(T) nor its complement – the space of functions with discontinuities – belong to $\mathcal{R}^T$. The preceding expressions make sense only because

$$\{\omega \in \Omega | Y'_{\cdot}(\omega) \in C(T)\} = \Omega \in \mathcal{A} \quad \text{and} \quad \{\omega \in \Omega | Y_{\cdot}(\omega) \notin C(T)\} = \Omega \in \mathcal{A} \ .$$

But in terms of the probability distributions P and P′ on $\mathcal{R}^T$ induced by $\{Y_t\}$ and $\{Y'_t\}$ respectively, the expressions $P'(C(T)) = 1$, and $P(C(T)) = 0$ do not make sense because neither P′ nor P are defined for the set C(T). To see this, simply observe that since $P = P'$, the above would imply

$$1 = P(\overline{R}^T) = P(C(T) \cup (\overline{\mathbf{R}}^T \setminus C(T))$$

$$= P'(C(T)) + P(\overline{\mathbf{R}}^T \setminus C(T)) = 1 + 1 = 2 \ !$$

Observe also that the paths of both processes are bounded. But again in terms of P, or equivalently P′, we cannot characterize boundedness since

$$\{x \in \mathbf{R}^T | 0 \le x(t) \le 1 \ , \text{for all } \ t \in T\} \notin \mathcal{R}^T \ .$$

EXAMPLE 2.2 *The Poisson Process.* Let $\{X_n, n = 1,...\}$ be a stochastic process ($T = \mathbf{N}$) where

$X_1$ is the waiting time for the first event,

and for $n = 2,...$

$X_n$ is the waiting time between $(n-1)$-th and n-th event.

Then, the time of occurrence of the n-th event is

$$S_n := X_1 + \cdots + X_n \ .$$

Under the assumption that the event

$$0 =: S_0 < S_1 < \cdots < S_n < \cdots \ , \quad \sup_n S_n = \infty \tag{2.1}$$

has probability 1, on this subset of $\Omega$, we define the random variables

$$N_t := \max[n : S_n \le t] \tag{2.2}$$

that records the (random) number of events that occur in the interval $[0, t]$; if $\omega$ is not in the set specified by (2.1), we set $N_t(\omega) := 0$. It is well known, see [3] for example, that if the $\{X_n, n = 1,...\}$ are independent with the same exponential distribution, then $\{N_t, t \ge 0\}$ is the *Poisson* stochastic process.

For every $\omega$, the realization

$$t \mapsto N_t(\omega) : \Omega \to R_+$$

is a nondecreasing, integer-valued function.

Let $\mathbf{Q} \subset R$ be the rationals, and let $\varphi : [0, \infty) \to \mathbf{Q}$ be such that $\varphi(t) := t$ if $t \in \mathbf{Q}$, and $\varphi(t) := 0$ otherwise. Now, define

$$M_t(\omega) := N_t(\omega) + \varphi(t + X_1(\omega)) \ .$$

For all $t \in [0, \infty)$

$$\mu\{\omega|\varphi(t + X_1(\omega)) \neq 0\} = \mu[X_1 \in \mathbf{Q} - t] = 0 \ ,$$

since $\mathbf{Q} - t$ is a countable subset of $\mathbf{R}_+$ and $X_1$ is absolutely continuous (with respect to the Lebesgue measure). Thus

$$\mu\{\omega|M_t(\omega) = N_t(\omega)\} = 1$$

and the stochastic process $\{M_t, t \in \mathbf{R}_+\}$ has the same family of finite dimensional distributions as $\{N_t, t \in \mathbf{R}_+\}$. However, for all $\omega$, the realizations $t \mapsto \mathbf{M}_t(\omega)$ are everywhere discontinuous, neither monotone nor integer-valued!

We are basically in the same situation as in Example 2.1. The realizations of $\{N_t\}$ all lie in

$$\{x \in \mathbf{R}^{[0,\infty)}|x : [0, \infty) \to \mathbf{N}\,,\ x(s) \leq x(t) \quad \text{whenever} \quad s \leq t\}$$

which does not belong to $\mathcal{R}^T$.

All of this comes from the fact that *a subset* $B$ *of* $\mathbf{R}^T$ *cannot lie in* $\mathcal{R}^T$ *unless there exists a countable subset* $S$ *of* $T$ *with the property: if* $x \in B$ *and* $x(t) = y(t)$ *for all* $t$ *in* $S$ *then* $y \in B$ [3, Theorem 36.3].

This means that any set of the type

$$\{x \in \mathbf{R}^T|x(t) \in F \quad \text{for all} \quad t \in T' \subset T\}$$

where $F \subset \mathbf{R}$ is closed, are not necessarily in $\mathcal{R}^T$, since they usually cannot be obtained as *countable* intersections of sets in $\mathcal{R}^T$. This is especially important when it comes to the study of functionals of stochastic processes. For a stochastic process $\{X_t, t \in T\}$, let

$$J(\omega) = \inf_{t \in T} X_t(\omega) \ ,$$

then for all $\alpha \in \mathbf{R}$,

$$\{\omega|J(\omega) \geq \alpha\} = \{\omega|X_.(\omega) \in S_\alpha\}$$

where

$$S_\alpha = \{x \in \overline{\mathbf{R}}^T| \quad \text{for all} \quad t \in T,\ x(t) \geq \alpha\} \ ,$$

but $S_\alpha \notin \overline{\mathcal{R}}^T$, and thus J is not even measurably related to the stochastic process $\{X_t\}$. This point is brought home by considering the two equivalent processes of Example 2.1. Here, both

$$J_1 := \inf_{t \in T} Y_t, \quad \text{and} \quad J_1' : \inf_{t \in T} Y_t'$$

turn out to be measurable functions from $\Omega$ into [0, 1] but in no way "equivalent", since

$$J_1 \equiv -1, \quad \text{and} \quad J_1' \equiv 0 \ .$$

These are some of the simplest examples we know that clearly suggest that the class $\mathcal{R}^T$ is often too small to obtain an appropriate probabilistic description of stochastic processes. The applications should, of course, dictate the framework to use in any particular situation. In the next sections, we show that there is a rather general approach that allows us to avoid some of the objections that one may have to this "simple" definition of stochastic processes.

## 3. SOME TOPOLOGICAL CONSIDERATIONS

From a topological viewpoint, the shortcomings of the "finite dimensional distributions" description of stochastic processes come from the fact that $\mathbf{R}^T$ does not take into account the underlying topology of T. The $\sigma$-field $\mathcal{R}^T$ is not in general a Borel field, although the first step in the construction of $\mathcal{R}^T$ is topological in nature. We can think of $\mathcal{R}^T$ as generated by the class of measurable rectangles

$$\{x \in \mathbf{R}^T \,|\, (x(t_1), \ldots, x(t_k)) \in G_1 \times \cdots \times G_k\}$$

as $(t_1, \ldots, t_k)$ ranges over $I(T)$ and the $G_i$ ranges over $G(\mathbf{R})$, the open subsets of R.

This class of measurable rectangles is the base for the product topology on $\mathbf{R}^T$ but in general $\mathcal{R}^T$ is not the Borel field with respect to the product topology. Unless T is a countable space, the product topology has never a countable base [6, Theorem 6]. If $B_\pi$ denotes the Borel field generated by the open sets of the product topology, we have that

$$\mathcal{R}^T \subset B_\pi$$

with equality if T is countable. For example, if $T \subset \mathbf{R}$ is an open interval, let $A$ be the subset of $\mathbf{R}^T$ that consists of the constant functions with values in [0.1]. Then $A$ belongs to $B_\pi$ but not to $\mathcal{R}^T$.

The "classical" approach essentially ignores the topology with which T is endowed, in favor of the discrete topology. And since, with respect to the discrete topology, all functions in $\mathbf{R}^T$ are continuous, there is no way to distinguish between those realizations that we identify as continuous (with respect to the usual topology on $\mathbf{R}$) and any other realizations, that are also "continuous" but now with respect to the discrete topology.

One general approach, that allows us to include (at least to our knowledge) all interesting stochastic processes, and which skirts around all of the inherent difficulties of the "classical" approach, is to think of stochastic processes as random lower (or upper) semicontinuous functions. The realizations of such processes are then lower (or upper) semicontinuous functions, a rather large class of functions that should include nearly all possible applications. And for this class, there is a natural choice of topology, and an approach that avoids most of the pitfalls of the "finite dimensional distributions" approach.

For any function $x : T \to \overline{\mathbf{R}}$, the *epigraph* of x is the subset of the product space $T \times \mathbf{R}$ defined by

$$\text{epi } x = \{(t, \alpha) \mid \alpha \geq x(t)\} .$$

To any stochastic process $\{X_t, t \in T\}$ we can associate its *epigraphical representation*, i.e., the set-valued map defined as follows:

$$\omega \mapsto \text{epi } X_{.}(\omega) = \{(t, \alpha) \mid \alpha \geq X_t(\omega)\} .$$

For any finite set $I = \{(t_1, \alpha_1), \ldots, (t_q, \alpha_q)\}$ in $T \times \mathbf{R}$, we have

$$\{x \in \mathbf{R}^T \mid (x(t_1) > \alpha_1, \ldots, x(t_q)) > \alpha_q\} = \{x \in \mathbf{R}^T \mid \text{epi } x \cap I = \phi\} . \qquad (3.1)$$

Since $R^T$ is the minimal $\sigma$-filed generated by sets of the type

$$R^T = \sigma - \{\{x \in \mathbf{R}^T \mid x(t_1) \leq \alpha_1, \ldots, x(t_q) \leq \alpha_q\}, [(t_1, \alpha_1), \ldots, (t_q, \alpha_q)] \in I(T \times \mathbf{R})\}$$

with $I(T \times \mathbf{R})$ the finite subsets of $T \times \mathbf{R}$. From (3.1) it also follows that

$$R^T = \sigma - \{\{x \in \mathbf{R}^T \mid \text{epi } x \cap I \neq \phi\}, I \in I(T \times \mathbf{R})\} .$$

The sets of $I(T \times \mathbf{R})$ form a base for the discrete topology of $T \times \mathbf{R}$, and they are also compact with respect to this topology.

In the epigraphical view, the "classical" approach defines a stochastic process $\{X_t, t \in T\}$ with domain $(\Omega, A, \mu)$ as a measurable map from $(\Omega, A)$ into $(\mathbf{R}^T, R^T)$, where measurability means that

$$\{\omega \in \Omega | \text{epi } X_{\cdot}(\omega) \cap K \neq \phi\} \in A \tag{3.2}$$

for all subsets K of $T \times \mathbf{R}$, that are compact with respect to the discrete topology. This highlights the source of the limitations of the classical approach, it is not able to identify the topological properties of the realizations beyond those that can be identified by the discrete topology. The preceding relation also suggest the remedy to use, in order to bring the topology of T into the probabilistic description of the process. Instead of working with the discrete topology on $T \times \mathbf{R}$, we could equip $T \times \mathbf{R}$ with a topology that would be more appropriate for the application at hand.

Let us return to Example 2.1 with $T = \mathbf{R}_+$. If P and P′ denote the probability measures induced by Y and Y′ respectively, then

$$\mu\{\omega | \text{epi } Y_{\cdot}(\omega) \cap K = \phi\} = P\{x \in \mathbf{R}^T | \text{epi } x \cap K = \phi\}$$

$$= P'\{x \in \mathbf{R}^T | \text{epi } x \cap K = \phi\}$$

$$= \mu\{\omega | \text{epi } Y'_{\cdot}(\omega) \cap K = \phi\}$$

for all subsets K of $T \times \mathbf{R}$ that are compact for the discrete topology. The situation is completely different if compact refers to the "natural" topology, i.e. the usual topology on $\mathbf{R}^2$ relative to $T \times \mathbf{R}$. It is easy to verify that for any $\beta \in (-1, 0)$ and $[\alpha_1, \alpha_2] \subset T$, we have that

$$\mu\{\omega | \text{epi } Y'_{\cdot}(\omega) \cap ([\alpha_1, \alpha_2] \times [-1, \beta]) = \phi\} = 1$$

and

$$\mu\{\omega | \text{epi } Y_{\cdot}(\omega) \cap ([\alpha_1, \alpha_2] \times [-1, \beta]) = \phi\} = \mu\{\omega | V(\omega) \not\in [\alpha_1, \alpha_2]\} .$$

This time, the "induced" probability measures will be different but of course they cannot be defined on $R^T$, that in the classical approach is the "universal" functional space for dealing with stochastic processes.

## 4. SEPARABILITY, MEASURABILITY AND STOCHASTIC EQUIVALENCE

The epigraphical approach focuses its attention on the sets of the type:

$$\{\omega \in \Omega \mid \text{epi } X_{.}(\omega) \cap K \neq \phi\} \ ,$$

to define measurability, as well as to serve as building block in the definition of the probability measure associated with the process $\{X_t, t \in T\}$. Let $K_\tau$ denote the class of compact subsets of $T \times \mathbf{R}$ where $\tau$ is the product topology generated by $\tau_1$ on T and the usual topology on $\mathbf{R}$. ***Measurability*** of the process $\{X_t, t \in T\}$ will now mean: ***for all*** K ***in*** $K_\tau$,

$$\{\omega \in \Omega \mid \text{epi } X_{.}(\omega) \cap K \neq 0\} \in A \ . \tag{4.1}$$

This condition is closely related to the classical notion that the process is measurable, which means that

$$(\omega, t) \mapsto X_t(\omega) \text{ is } A \otimes B(T) - \text{measurable} \tag{4.2}$$

where $B(T)$ is the Borel field on T generated by the $\tau_1$- open sets. In Section 6, we shall show that for stochastic processes with lower semicontinuous realizations, these two conditions are equivalent. We bring this fact to the fore at this time, because to require that a process be measurable is a standard condition used to overcome some of difficulties created by the classical definition. By definition any stochastic process is $R^T$-measurable, but not necessarily in terms of (4.2) or (4.1). This follows from the fact all sets that are compact with respect to the discrete topology are also $\tau$- compact.

Closely related to the notion of measurability of a stochastic process is that of the ***separability*** of a stochastic process, as introduced by Doob. Among the major shortcomings of the class $R^T$ is the fact that subsets of the type

$$\{x \in \overline{\mathbf{R}}^T \mid x(t) \in F, \ t \in T' \subset T\} \tag{4.3}$$

where F is a closed subset of $\mathbf{R}$, do not necessarily belong to $\overline{R}^T$. One circumvents the potential difficulties by requiring that the stochastic process $\{X_t, t \in T\}$ be ***separable***, i.e. there exists an everywhere dense countable subset D of T and a $\mu$-null set $N \subset \Omega$ such that for every open set $G \subset T$ and closed subset F of $\mathbf{R}$, the sets

$$\{\omega \in \Omega \mid X_t(\omega) \in F \ \text{ for all } \ t \in G \cap D\}$$

and

$$\{\omega \in \Omega | X_t(\omega) \in F \text{ for all } t \in G\}$$

differ from each other at most on a subset of N [5].

In terms of the realizations of the stochastic process $\{X_t, t \in T\}$, separability means that for all $\omega \in \Omega \setminus N$, the function $t \mapsto X_t(\omega)$ is D-*separable* [3], i.e. for every t in T there exists a sequence $\{t_n, n = 1,\ldots\}$ such that

$$t_n \in D, \; t = \lim_n t_n, \text{ and } X_t(\omega) = \lim_n X_{t_n}(\omega) \; ,$$

in other words, for every $\omega \in \Omega \setminus N$, the realization is completely determined by its values on D. A stochastic process separable with respect to D is $\mathcal{R}^D$- measurable, and one may reasonably assume that the fact that D is countable removes the "discrepancies" connected with "uncountabilities". Of course not all stochastic processes are separable. Process $\{Y_t\}$ of Example 2.1 is not separable, although the equivalent stochastic process $\{Y_t'\}$ is separable. In fact, given any finite-valued process there always exists an equivalent process defined on the same probability space that is separable [3, Theorem 38.1].

At first, it may appear that it is possible to restrict the study of stochastic processes to those that are separable, but there is some hidden difficulty. Separability is defined in term of a reference set D. For the convergence of stochastic processes, it would be necessary to prove first that there exists a set D with respect to which all elements of the sequence (or net), as well as the limit process, are separable. Moreover, the existence of an equivalent separable process does not mean that the functionals defined on these processes will in any way be comparable; think about the processes $\{Y_t\}$ and $\{Y_t'\}$ of Example 2.1 and the sup functional, see Section 2. Separability only guarantees that sets of the type (4.1) are measurable and that their probability can be determined by the family of finite dimensional distributions. If $\{X_t, t \in T\}$ is not separable, nothing can be said a priori about sets of the type (4.1), and no additional information is gained from the fact that there is an equivalent separable stochastic process. Thus a functional of the stochastic process involving sets of type (4.1) cannot be analyzed in terms of the same functional defined on an equivalent stochastic process.

Roughly speaking, separability is an attempt at recovering the topological structure of T, at posteriori. The approach developed in the next sections takes the topological structure of T directly into account.

## 5. THE EPIGRAPHICAL APPROACH

The earlier sections have pointed out the shortcomings of the "classical" approach by reformulating it in terms of the epigraphical representation of the process. We have seen that the inherent weaknesses of this approach can be overcome by requiring that the stochastic process satisfy the stronger measurability condition

$$\{\omega \in \Omega \,|\, \text{epi}\ X_{\cdot}(\omega) \cap K \neq \phi\} \in \mathcal{A} \quad \text{for all}\ K \in \mathcal{K}_T \tag{5.1}$$

which take into account the topological structure of T.

All that follows is devoted to the study of stochastic processes that satisfy condition (5.1) and have lower semicontinuous (l.sc.) realizations, i.e.,

$$t \mapsto X_t(\omega) \ \text{ is l.sc. on T, for all } \ \omega \in \Omega \ . \tag{5.2}$$

Such stochastic processes, with possibly the values $+\infty$ and $-\infty$, are called ***random l.sc. functions***. In another setting, such functions are known as ***normal integrands***, and much of the theory developed by Rockafellar [9, 10] for normal integrands can be transposed to the present context. Many of the questions raised in the earlier sections seem to find their natural formalization in terms the properties of random l.sc functions and the associated epigraphical behavior. This leads us also to consider the associated ***random closed set***

$$\omega \mapsto \text{epi}\ X_{\cdot}(\omega) : \Omega \rightrightarrows \mathbf{R} \ . \tag{5.3}$$

For each $\omega$, the set epi $X_{\cdot}(\omega)$ is a closed subset of $T \times \mathbf{R}$ since the functions $t \mapsto X_t(\omega)$ are l.sc., and the measurability of this set-valued function follows from condition (5.1).

All of this suggests defining a topology for the space of (extended real-valued) l.sc. functions in terms of the epigraphs, the ***epi–topology***. We shall see that the corresponding Borel field provides us with the desired interplay between topological properties and measurability. We follow the development that was initiated in [12] and review here some of the main features of that theory.

At first it may appear that the requirement that the process has l.sc. paths is a rather serious limitation. At least if we use this framework for the study of general stochastic processes. This is not the case. Of course, stochastic processes with continuous realizations fit into this class, but also any cád-lág process (continuous from the right, limits from the left) admits a trivial modification that makes it a stochastic process with l.sc. paths. Although, we restrict ourselves to the l.sc. case, it is clear that all the results have their counterpart in the upper semicontinuous (u.sc.) case, replacing everywhere epigraph by ***hypograph***.

Crucial to the ensuing development is the fact that for stochastic processes that are l.sc. random functions, we can introduce a notion of convergence which is not only the appropriate one if we are interested in the extremal properties of the process, as well as for many related functionals, but also provides in many situations a more satisfactory approach to the convergence of stochastic processes as the standard functional approach.

## 6. THE EPIGRAPHICAL RANDOM SET

Henceforth, we work in the following setting

- $(\Omega, \mathcal{A}, \mu)$ a *complete* probability space,
- $(T, \tau_1)$ a locally compact separable metric space,
- $(\omega, t) \mapsto X_t(\omega) : \Omega \times T \to \overline{\mathbf{R}}$ a *random l.sc. function*.

By this we mean that

(i) for every $\omega$, the realization $t \mapsto X_t(\omega)$ is l.sc. with values in the extended reals

(ii) the map $(\omega, t) \mapsto X_t(\omega)$ is $\mathcal{A} \otimes \mathcal{B}_1$- measurable, where $\mathcal{B}_1$ is the Borel field on T.

The associated epigraphical random set, is the map

$$\omega \mapsto \text{epi}\, X_{.}(\omega) : \Omega \rightrightarrows R$$

that takes values in the closed subsets of $T \times \mathbf{R}$, including the empty set.

The product space $T \times \mathbf{R}$ is given the product topology of $\tau_1$ with the natural topology on $\mathbf{R}$, we denote it by $\tau$. Thus $(T \times \mathbf{R}, \tau)$ is a locally compact separable metric space. Let

- $\mathcal{F} = \mathcal{F}(T \times \mathbf{R})$ denote the closed subsets of $T \times \mathbf{R}$,
- $\mathcal{G} = \mathcal{G}(T \times \mathbf{R})$ denote the open subsets of $T \times \mathbf{R}$,
- $\mathcal{K} = \mathcal{K}(T \times \mathbf{R})$ denote the compact subsets of $T \times \mathbf{R}$,

For any subset C of $T \times \mathbf{R}$, let

$$\mathcal{F}^C := \{F \in \mathcal{F} \mid F \cap C = \phi\},\ \ \mathcal{F}_C := \{F \in \mathcal{F} \mid F \cap C = \phi\}\ .$$

The topology $\mathcal{T}$ generated by the subbase of open sets

$$\{\mathcal{F}^K,\ K \in \mathcal{K}\},\quad \text{and}\quad \{\mathcal{F}_G,\ G \in \mathcal{G}\} \tag{6.1}$$

makes the topological (hyper)space $(\mathcal{F}, \mathcal{T})$ regular and compact, see e.g. [4, Proposition 3.2]. If T has a countable base, so does $(\mathcal{F}, \mathcal{T})$, see e.g. [7, Theorem 1-2-1] and [4] in which case a base for $\mathcal{T}$ is given by the open sets of the type

$$\{\mathcal{F}^{\mathrm{cl}\,B_1 \cup \cdots \cup \mathrm{cl}\,B_s} \cap \mathcal{F}_{B_{s+1}} \cap \cdots \cap \mathcal{F}_{B_q},\ q \text{ finite}\} \tag{6.2}$$

where cl C denoted the ***closure*** of C, and the

$$\{B_i,\ i = 1, \ldots, q\}$$

come from a countable base of open sets for $T \times \mathbf{R}$. The ***Borel field***, generated by the $\mathcal{T}$-open subsets of $\mathcal{F}$, will be denoted by $\mathcal{B}(\mathcal{F})$. It is easy to see that it can be generated from the subbase of open sets (6.1), and in the countable-base case by the restricted class (6.2), cf., [7, 11].

We can also view the epigraphical random set as a random variable defined on $\Omega$ and values in $\mathcal{E}$, the subset of $\mathcal{F}$, ***consisting of the sets that are epigraphs.*** It is easy to verify that $\mathcal{E}$ is a closed subset of $\mathcal{F}$, and thus with the $\mathcal{T}$-relative topology, it inherits all the properties of $\mathcal{F}$. The map

$$\omega \mapsto \text{epi}\ X_{\cdot}(\omega) : \Omega \to \mathcal{E}$$

is ***measurable*** (is a ***random set***), if for all $K \in \mathcal{K}$,

$$(\text{epi}\ X_{\cdot})^{-1}(K) = \{\omega \in \Omega \mid \text{epi}\ X_{\cdot}(\omega) \cap K \neq \phi\} \in \mathcal{A}\ . \tag{6.3}$$

This is equivalent [9, 11] to any one of the following conditions:

$(\text{epi}\ X_{\cdot})^{-1}(F) \in \mathcal{A}$ for all $F \in \mathcal{F}$,

$(\text{epi}\ X_{\cdot})^{-1}(B) \in \mathcal{A}$ for all closed balls B of $T \times \mathbf{R}$,

$\omega \mapsto \text{epi}\ X_{\cdot}(\omega)$ admits a Castaing representation (see below), graph (epi $X_.$) $\in \mathcal{A} \otimes \mathcal{B}_1$,

$\omega \mapsto$ epi $X_.(\omega) : \Omega \to \mathcal{F}$ is $\mathcal{B}(\mathcal{F})$-measurable.

Each one of these characterizations catches a special aspect of the measurability of the epi $X_.$. To have measurable graph corresponds to having $\{X_t, t \in T\}$ a measurable stochastic process. The fact that the random (closed) sets admits a Castaing representation generalizes the notion of separability of a stochastic process. And the last one induces on $(\mathcal{F}, \mathcal{B}(\mathcal{F}))$, more precisely on $(\mathcal{E}, \mathcal{B}(\mathcal{E}))$, a distribution. From the definitions, it is immediate to verify [10, Proposition 1] that

THEOREM 6.1 *The stochastic process* $\{X_t, t \in T\}$ *with l.sc. realizations is measurable if and only if* $(\omega, t) \mapsto X_t(\omega)$ *is a random l.sc. function, or still, if and only if* $\omega \mapsto$ epi $X_.(\omega)$ *is a random closed set.*

A countable collection of measurable functions $\{x_k, \alpha_k) : \Omega \to T \times \mathbf{R}, k = 1, \ldots\}$ is a *Castaing representation* [9] of epi $X_. \in \mathcal{A}$ if

$$\{\omega \mid \text{epi } X_.(\omega) \neq \phi\} := \text{dom epi } X_. \in \mathcal{Q} \ ,$$

and for all $\omega \in$ epi $X_.$,

$$\text{cl}\,(\textstyle\bigcup_k \{x_k(\omega), \alpha_k(\omega)\}) = \text{epi } X.(\omega) \ .$$

We now show that the fact that the random closed set epi $X_.$ admits a Castaing representation is an extension of the notion of separability for the stochastic process $\{X_t, t \in T\}$. The key fact is the following:

THEOREM 6.2 *Any real-valued separable stochastic process* $\{X_t, T \in T\}$ *with l.sc. realizations is a measurable process.*

PROOF In view of Theorem 6.1, and the equivalent definitions of measurability (for a random set), it suffices to exhibit a countable collection of measurable functions

$$(x_k, \alpha_k) : \Omega \to T \times \mathbf{R}, \ k = 1, \ldots,$$

such that for all $\omega \in \Omega$,

$$\text{epi } X_.(\omega) = \text{cl}\,(\textstyle\bigcup_k \{x_k(\omega), \alpha_k(\omega)\}) \ .$$

Suppose $D = \{d_i, i \in I\} \subset T$ is the countable set with respect to which $\{X_t, t \in T\}$ is separable, and let $A = \{a_j, j \in J\}$ be a countable dense subset of $\mathbf{R}$. Then $D \times A$ is a countable dense subset of $T \times \mathbf{R}$. Let $\{(x_i, \alpha_j): \Omega \to T \times \mathbf{R}, i \in I, j \in J\}$ be a countable collection of random functions defined by

$$x_i(\omega) := d_i, \ \alpha_j(\omega) = a_j \quad \text{for all} \ \omega \in \Omega \ .$$

Since $\{X_t, t \in T\}$ is a stochastic process, for all $(i, j)$

$$\{\omega | (x_i(\omega), \alpha_j(\omega)) \in \text{epi } X_.(\omega)\} = \{\omega | X_{d_i}(\omega) \leq a_j\} \in \mathcal{A} \ .$$

Let N be a $\mu$-null subset of $\Omega$ such that every realization of $X_.(\omega)$ is D-separable for all $\omega \in \Omega \setminus N$. We have that for all $\omega \in \Omega \setminus N$,

$$\text{epi } X_.(\omega) = \text{cl} (\text{epi } X_.(\omega) \cap \{(x_i(\omega), \alpha_j(\omega)), i \in I, j \in J\} \ .$$

For all $\omega \in \Omega \setminus N$ and all $(x, \alpha) \in \text{epi } X_.(\omega)$, by D-separability of $\{X_t, t \in T\}$, there exists $\{d_n \in D, n = 1, \ldots\}$ such that

$$x = \lim_n d_n, \quad \text{and} \quad \alpha \geq X_t(\omega) = \lim_n X_{d_n}(\omega) \ .$$

Since A is dense in $\mathbf{R}$ and $t \mapsto X_t(\omega)$ is l.sc., we can always find a sequence $\{\alpha_n \in A, n = 1, \ldots\}$ such that $\alpha_n \geq X_{d_n}(\omega)$ and $\alpha = \lim_n \alpha_n$. This means that for all $\omega \in \Omega \setminus N$

$$\text{epi } X_.(\omega) \subset \text{cl} \{\text{epi } X_.(\omega) \cap \{(x_i(\omega), \alpha_j(\omega)), i \in I, j \in J\} \ .$$

But this yields equality since the reverse direction is trivially satisfied.

The stochastic process $\{X'_t, t \in T\}$ having epigraphical representation

$$\omega \mapsto \text{epi } X'_.(\omega) = \begin{cases} \text{epi } X_.(\omega) & \text{if} \ \ \omega \in \Omega \setminus N \\ \text{cl} (\text{epi } X_.(\omega) \cap \{x_i, (\omega), \alpha_j(\omega)), i \in I, j \in J\} & \text{if} \ \ \omega \in N \end{cases}$$

is measurable, by Theorem 6.1, i.e., for all $K \in \mathcal{K}$

$$\{\omega \in \Omega | \text{epi } X'_.(\omega) \cap K \neq 0\} \in \mathcal{A} \ .$$

For the process $\{X_t, t \in T\}$, we have that

$$\{\omega \in \Omega | \text{epi } X_.(\omega) \cap K \neq \phi\}$$

$$= \{\omega \in \Omega \setminus N | \text{epi } X'_.(\omega) \cap K \neq \phi\} \cup \{\omega \in N | \text{epi } X_.(\omega) \cap K \neq \phi\}$$

also belongs to $\mathcal{A}$. The latter set is of measure 0 and belongs to $\mathcal{A}$, since $(\Omega, \mathcal{A}, \mu)$ is complete by assumption.□

The converse of this theorem does not hold. A counterexample would be the process $\{Y_t, t \in T\}$ as defined in Example 2.1 with $T = \mathbf{R}_+$.

REMARK 6.3 In Section 4, we indicated that separability was introduced to recover the measurability of the sets

$$\{\omega \in \Omega \mid X_t(\omega) \in F, \quad \text{for all } t \in G \subset T\}$$

where $F \subset \mathbf{R}$ is closed and G is $\tau_1$-open, we should note that there are of course no measurability problems if $(\omega, t) \mapsto X_t(\omega)$ is a random l.sc. function. And thus in that context, separability is mostly an irrelevant concept.

## 7. DISTRIBUTIONS AND DISTRIBUTION FUNCTIONS

In section 6, we have seen that to each random l.sc. function we can associate an epigraphical random closed set. As we shall show now, to each random closed set there corresponds a distribution function, which in turn will allow us to define the "distribution function" of a random l.sc. function. Let us denote by $\Gamma$ a random closed set, defined on $\Omega$ and with values in the closed subsets of $T \times \mathbf{R}$. Let P denote the distribution of $\Gamma$, i.e., the probability measure induced on $\mathcal{B}(\mathcal{F})$ by the relation

$$P(B) = \mu\{\omega \mid \Gamma(\omega) \in B\} \tag{7.1}$$

for all $B \in \mathcal{B}(\mathcal{F})$.

Since the topologial space $(\mathrm{F}, \mathcal{B}(\mathcal{F}))$ is metrizable, see Section 6, every probability measure defined on $\mathcal{B}(\mathcal{F})$ is regular [2, Theorem 1.1], and thus is completely determined by its values on the open (or closed) subsets of $\mathcal{F}$. If we assume that $\mathcal{F}$ has a countable base – and for this it suffices that T has a countable base – every open set in $\mathcal{F}$ is the countable union of elements in the base, obtained by taking finite intersection of the elements in the subbase. Thus, it will certainly be sufficient to know the values of P on the subbase (6.1) to completely determine P. This observation will bring us to the notion of a distribution function for the random closed set $\Gamma$ [12].

First observe that the restriction of P to the class $\{\mathcal{F}_K, K \in \mathcal{K}\}$ defines a function D on $\mathcal{K}$ through the relation:

$$D(K) = P(\mathcal{F}_K) = \mu\{\omega \mid \Gamma \cap K \neq \phi\} \tag{7.2}$$

for all $K \in \mathcal{K}$. This function has the following properties:

$$D(\phi) = 0 \ ; \tag{7.3}$$

for any decreasing sequence $\{K_\nu, \ \nu = 1, \ldots\}$ in $\mathcal{K}$, the sequence

$$\{D(K_\nu), \ \nu = 1, \ldots\} \quad \text{decreases to} \ D(\lim K_\nu) \ ; \tag{7.4}$$

for any sequence of sets $\{K_\nu, \ \nu = 0, \ldots\}$, the functions $\{\Delta_n, \ n = 0, 1, \ldots\}$ defined recursively by

$$\Delta_0(K_0) = 1 - D(K_0) \ ,$$

$$\Delta_1(K_0; K_1) = \Delta_0(K_0) - \Delta_0(K_0 \cup K_1) \ , \tag{7.5}$$

and for $n = 2, \ldots$

$$\Delta_\nu(K_0; K_1, \ldots, K_\nu) = \Delta_{\nu-1}(K_0; K_1, \ldots, K_{\nu-1}) - \Delta_{\nu-1}(K_0 \cup K_\nu; K_1, \ldots, K_{\nu-1}) \ ,$$

take on their values $[0, 1]$.

The properties of D on $\mathcal{K}$ are essentially the same as those of the distribution function of a 1– or n-dimensional random variable. Property (7.4) is the same as right-continuity, whereas (7.3) corresponds to the continuity at $-\infty$ for a distribution function on the real line. Property (7.5) can be viewed as an extension of the notion of monotonicity. In view of this, and the fact [12, Choquet's Theorem 1.3] that any function $D: \mathcal{K} \to [0, 1]$ that satisfies the conditions (7.3), (7.4), (7.5) uniquely determines a probability measure on $\mathcal{B}(\mathcal{F})$, we call D the ***distribution function*** of $\Gamma$.

The fact that we can restrict the domain of definition of D to the subclass $\mathcal{K}^{ub}$ of $\mathcal{K}$ is very useful in a number of applications, where

$\mathcal{K}^{ub}$ = {finite union of closed balls with positive radii};

note that $\phi \in \mathcal{K}^{ub}$ as the union of an empty collection. This comes from the fact that the properties of $(\mathcal{F}, \mathcal{T})$ enables us to generate $\mathcal{B}(\mathcal{F})$ from the family

$$\{\mathcal{F}_K, \ K \in \mathcal{K}^{ub}\} \ ;$$

in fact, for all $K \in \mathcal{K}$, we have

$$K = \bigcap \{K' \mid K' \supset K, \ K' \in \mathcal{K}^{ub}\}$$

and

$$\mathcal{F}_K = \bigcap_{\substack{K' \supset K \\ K' \in \mathcal{K}^{ub}}} \mathcal{F}_{K'} \ ;$$

and consequently

$$D(K) = P(\mathcal{F}_K) = \inf_{\substack{K' \supset K \\ K' \in \mathcal{K}^{ub}}} P(\mathcal{F}_{K'}) = \inf_{\substack{K' \supset K \\ K' \in \mathcal{K}^{ub}}} D(K') \ .$$

The (probability) ***distribution function*** of a random lower semicontinuous function $(\omega, t) \mapsto X_t(\omega)$ is the distribution function of its epigraphical random set. Since the random set takes its values in the (hyper)space of epigraphs we could reformulate it in the following terms: let C be a $\tau_1$-compact subset of T, and $\alpha \in \mathbf{R}$, then

$$D(C, \alpha) := \mu\{\omega \,|\, \inf_{t \in C} X_t(\omega) \leq \alpha\}$$

defined on (the compact subsets of T) $\times$ $\mathbf{R}$ can be used instead of the usual definition of D on the compact subsets of T $\times$ $\mathbf{R}$.

## 8. ... AND FINITE DIMENSIONAL DISTRIBUTIONS!

Let us consider $\{X_t, t \in T\}$ a measurable stochastic process with l.sc. realizations, then epi $X_\cdot : \Omega \rightrightarrows T \times \mathbf{R}$ is a closed random set with distribution function $D : \mathcal{K} \rightarrow [0, 1]$. Any finite set $I = \{(t_1, \alpha_1), \ldots, (t_h, \alpha_h)\} \subset T \times \mathbf{R}$ is $\tau$-compact, and thus we have

$$D(I) = \mu\{\omega \in \Omega \,|\, \text{epi } X_\cdot(\omega) \neq \phi\} \ .$$

In particular, if we fix t, then for all $\alpha \in R$

$$D(\{(t, \alpha)\}) = \mu\{\omega \in \Omega \,|\, X_t(\omega) \leq \alpha\} = P_t((-\infty, \alpha])$$

where $P_t$ refers to the 1-dimensional probability measure of the random variable $X_t$. Similarly, if we fix $t_1, \ldots, t_q$, then

$$P_{t1,\ldots,tq}((-\infty, \alpha_1] \times \cdots \times (-\infty, \alpha_q]) = \mu\{\omega \,|\, X_{t_1}(\omega) \leq \alpha_1, \ldots, X_{tq}(\omega) \leq \alpha_q\}$$

$$= \sum_{i=1}^{q} D(\{(t_i, \alpha_i)\}) - \sum_{\substack{i,j \\ i \neq j}} D(\{t_i, \alpha_i), (t_j, \alpha_j)\})$$

$$+ \cdots + (-1)^{h+1} D(\{(t_1, \alpha_1), \ldots, (t_q, \alpha_q)\})$$

It is now immediate that

THEOREM 8.1 *If* $\{X_t, t \in T\}$ *is a measurable stochastic process with l.sc. realizations, the finite dimensional distributions are completely determined by* D, *or equivalently by the restriction of* D *to the finite subsets of* $T \times \mathbf{R}$.

Of course, the converse of this theorem does not necessarily hold. Take for example the process $\{Y_t, t \in T\}$ of Example 2.1 with $T = \mathbf{R}_+$ and let $K = [t_1, t_2] \times [-\frac{1}{2}, -\frac{3}{4}]$, where $0 < t_1 < t_2$. Then $D(K) = \mu\{\omega | V(\omega) \in [t_1, t_2]\} > 0$, but $D(I) = 0$ for any finite subset I of K. The family of finite dimensional distributions, that assigns a value to D for every finite subset of K, does not allow us to make any inference about the value to assign to D(K).

REMARK 8.2 Note that the standard consistency conditions for the family of finite dimensional distribution could actually be derived from the "monotonicity" property (7.5) of the distribution function D. Thus, we can think of this family of finite dimensional distributions itself as a distribution function, but defined on the finite subsets of $T \times \mathbf{R}$. This suggests another approach to Kolmogorov's Consistency Theorem via Choquet's Theorem.

The fact that a compact set $K \subset T \times \mathbf{R}$ cannot be obtained as a countable union of finite sets is a topological fact that leads to a probabilistic discrepancy in the example involving the process $\{Y_t, t \in T\}$.

DEFINITION 8.3 *The distribution function of a random l.sc. function is said to be* inner separable, *if to any* $K \in \mathcal{K}$ *and* $\varepsilon > 0$, *there corresponds a finite set* $I_\varepsilon \subset K$ *such that* $D(K) < D(I_\varepsilon) + \varepsilon$.

The basic difference between separability of a stochastic process and the inner separability of its distribution is that separability is aimed at the reconstruction of sets through "finite sets", whereas inner separability is aimed at the reconstruction of the probabilistic content of the sets in terms of the probability associated to finite sets.

PROPOSITION 8.4 [12, Proposition 4.6]. *Suppose* $\{X_t, t \in T\}$ *is a measurable stochastic process with l.sc. realizations. If it is separable, then its distribution function is inner separable. Moreover, if its distribution function is inner separable, its values on* $\mathcal{K}$ *are completely determined by its values on the finite subsets of* $T \times \mathbf{R}$.

This last assertion is an immediate consequence of the definition of inner separability.

## 9. WEAK CONVERGENCE AND CONVERGENCE IN DISTRIBUTION

We show that for random l.sc. functions, weak convergence of the probability measures corresponds to the convergence of the distribution functions at the "continuity" sets.

By $\nu$, we index the members of a sequence of stochastic processes, the induced probability measures on $B(E)$, or the corresponding distribution functions on $K = K(T \times \mathbf{R})$; by $B(E)$ we mean the Borel field $B(F)$ restricted to $E$. With $\nu = \infty$, or simply without index, we refer to the limit element of the sequence. We have seen that for every $K \in K$:

$$D^\nu(K) = P^\nu(E_K) = \mu\{\omega \in \Omega | \text{epi } X^\nu(\omega) \cap K \neq \phi\} .$$

Since $E_K$ is a closed subset of $F$--$E$ is a closed subset of $F$--, we can easily obtain from the Portemanteau Theorem [2] that

PROPOSITION 9.1 *If* $P^\nu$ *converges weakly to* P, *then for all* $K \in K$

$$\limsup_{\nu \to \infty} D^\nu(K) \leq D(K) . \tag{9.1}$$

Unless $P(\text{bdy } E_K) = 0$, the probability measure attached to the ***boundary*** of $E_K$, we cannot guarantee that

$$\liminf_{\nu \to \infty} D^\nu(K) \geq D(K) , \tag{9.2}$$

i.e. unless K is a "continuity" point of D in a sense to be defined below. Note that "continuity sets" of D must correspond to P-continuity sets and that the class of sets for which this continuity is defined must at least be a convergence determining class [2].

DEFINITION 9.2 *An increasing sequence* $\{K^n, n = 1,\ldots\}$ *of compact sets is said to regularly converge to* K *if*

$$K = \text{cl} \bigcup_{n=1}^\infty K^n \quad \text{and} \quad \text{int } K \subset \bigcup_{n=1}^\infty K^n ; \tag{9.3}$$

*where* int S *denotes the interior of the set* S.

DEFINITION 9.3 *A distribution function* $D : \mathcal{K} \to [0,1]$ *is* distribution-continuous *at* K, *if for every regularly increasing sequence* $\{K^n, n = 1, \dots\}$ *to* K,

$$D(K) = \lim_{n \to \infty} D(K^n) \tag{9.4}$$

The *distribution-continuity set* $C_D$ of D, is the subset of $\mathcal{K}$ on which D is distribution-continuous

PROPOSITION 9.4 [12, Lemma 1.11]. *For any* $K \in \mathcal{K}$,

(i) *if* $(P(\text{bdy } E_K) = 0$, *then* $K \in C_D$;

(ii) $K \in C_D$ *and* $K = \text{cl}(\text{int } K)$, *then* $P(\text{bdy } E_K) = 0$.

Assuming that $(T, \tau_1)$ has a countable base, let $\mathcal{K}^{ub}_{\mathbf{Q}} \subset \mathcal{K}^{ub}$ be such that $\mathcal{K}^{ub}_{\mathbf{Q}}$ is the finite union of balls that determine a countable basis for $(T \times \mathbf{R}, \tau)$. We have

$$\mathcal{K}^{ub} \cap C_D = \mathcal{K}^{ub} \cap \{K \mid P(\text{bdy } E_K) = 0\} \ ,$$

and if T has a countable base

$$\mathcal{K}^{ub}_{\mathbf{Q}} \cap C_D = \mathcal{K}^{ub}_{\mathbf{Q}} \cap \{K \mid P(\text{bdy } E_K) = 0\} \ .$$

This allows us to rephrase weak-convergence of probability measures in terms of the pointwise convergence of the distribution functions.

THEOREM 9.5 [12, Theorem 1.15] *For the family of random l.sc. functions* $\{X_\cdot, X^\nu_\cdot, \nu = 1, \dots\}$, *equivalently of measurable stochastic processes with l.sc. realizations, we have that the* $P^\nu$ *converge weakly to* P *if and only if for all* $K \in \mathcal{K}^{ub} \cap C_D$, *(and if* $(T, \tau_1)$ *has a countable base, for all* $K \in \mathcal{K}^{ub}_{\mathbf{Q}} \cap C_D$*):*

$$D(K) = \lim_{\nu \to \infty} D^\nu(K) \ .$$

We refer to this type of convergence, as *convergence in distribution* of the stochastic processes $\{X^\nu_t, t \in T\}$ to $\{X_t, t \in T\}$, and denote it by $X^\nu_\cdot \xrightarrow{l.d} X_\cdot$.

## 10. CONVERGENCE IN DISTRIBUTION AND CONVERGENCE OF THE FINITE DIMENSIONAL DISTRIBUTIONS

In the classical approach to the study of stochastic processes, convergence of stochastic processes is defined in terms of the convergence of the finite dimensional distributions, that we denote by

$$X^{\nu}_{.} \xrightarrow{f.d.} X_{.} \ .$$

In view of the comments in Section 8, we cannot expect that $X^{\nu}_{.} \xrightarrow{f.d} X_{.}$ implies that $X^{\nu}_{.} \xrightarrow{l.d} X_{.}$, but the converse could reasonably be conjectured, see Theorem 8.1. However, in general also this implication fails. The reason is that for finite sets $K \subset \mathcal{K}$, the notions of distribution-continuity and continuity of the corresponding finite dimensional distribution do not coincide.

REMARK 10.1 This can all be traced back to the relationship between the epi-topology and the pointwise-topology. Equivalence is obtained in the presence of equi-semicontinuity [12, Section 3], see also [4] for details.

The passage from convergence in distribution to convergence of the finite dimensional distributions and vice-versa, is based on the possibility of "approximating" the values of the distribution function for compact sets K by finite sets, *independent of* $\nu$, and conversely.

DEFINITION 10.2 *The family of distribution functions* $\{D; D^{\nu} = 1, \dots\}$ *on* $\mathcal{K}$ *is* equi-outer regular *at the finite set* $I \subset T \times \mathbf{R}$, *if to every* $\varepsilon > 0$ *there corresponds a compact set* $K_{\varepsilon} \in \mathcal{K}^{ub} \cap \mathcal{C}_D$ *with* $K_{\varepsilon} \supset I$ *such that for* $\nu = 1, \dots$

$$D^{\nu}(K_{\varepsilon}) < D^{\nu}(I) + \varepsilon, \quad \text{and} \quad D(K_{\varepsilon}) < D(I) + \varepsilon \ .$$

Now, let $\mathcal{C}_{f.d.}$ denote the finite subsets of $T \times \mathbf{R}$, i.e.

$$\mathcal{C}_{f.d} \subset \{I = \{(t_1, \alpha_1), \dots, (t_q, \alpha_q)\}, \ q \text{ finite}\} \ ,$$

such that the distribution function of the vector $(X_{t_1}, \dots, X_{t_q})$ is continuous at $(\alpha_1, \dots, \alpha_q)$.

DEFINITION 10.3 *The family of distribution functions* $\{D; D^\nu, \nu = 1,\dots\}$ *on* $K$ *is* equi-inner separable *at* $K \in K$, *if to every* $\varepsilon > 0$, *there corresponds a finite set* $I_\varepsilon$ *such that*

$$D(K) < D(I_\varepsilon) + \varepsilon, \quad \text{and} \quad D^\nu(K) < D^\nu(I_\varepsilon) + \varepsilon$$

*for* $\nu = 1,\dots$; *see Definition 8.3.*

THEOREM [12, Corollary 4.6] *Suppose* $\{X; X^\nu, \nu = 1,\dots\}$ *is a collection of random l.sc. functions. Then* $X^\nu_\cdot \xrightarrow{\text{l.d.}} X$ *implies* $X^\nu_\cdot \xrightarrow{\text{f.d}} X_\cdot$ *if and only if* $\{D, D^\nu, \nu = 1,\dots\}$ *is equi-outer regular on* $C_{\text{f.d.}}$. *And* $X^\nu_\cdot \xrightarrow{\text{f.d}} X_\cdot$ *implies* $X^\nu_\cdot \xrightarrow{\text{l.d.}} X_\cdot$ *if and only if* $\{D; D^\nu, \nu = 1,\dots\}$ *is equi-inner separable.*

## 11. BOUNDED RANDOM L.SC. FUNCTIONS

Applications usually requires us to restrict our attention to a subclass of processes that possess further properties beside lower (or upper) semicontinuity. From the point of view of the eqigraphs, this means that, the realizations now belong to $E'$ a subset of $E$. Let $T'$ be the relative $T$-topology on $E'$. Then the topological space $(E', T')$ inherits a number of the properties of $(F, T)$ [6]. In particular, if $(F, T)$ is metric with countable base, then $(E', T')$ is metric with countable base. Thus, in principle all the earlier results still apply to $(E', T')$, and the theory of weak-convergence on separable metric spaces can be used to obtain convergence criteria. In particular, recall that:

THEOREM 11.1 Prohorov. *The sequence* $\{P^\nu, \nu = 1,\dots\}$ *of probability measures on* $B(E')$ *is tight if and only if every subsequence contains a further subsequence that weakly converges to a probability measure.*

This means that the sequence $\{P^\nu, \nu = 1,\dots\}$ is relatively compact. A subset $S$ of $E'$ is $T'$-compact if and only if it is a $T$-closed subset of $E$, see Section 6.

We now deal with bounded processes. We use this class to illustrate the potential application of the "epigraphical" approach to specific classes of stochastic processes. To begin with, let us observe:

LEMMA 11.2 *For all* $\alpha \in \mathbf{R}_+$

$$E_\alpha = \{\text{epi } x \,|\, \sup_{t\in T}|x(t)| \le \alpha\} \subset E$$

*is T-compact. And hence, any collection of probability measures P on* $B(E')$ *such that for every* $\varepsilon > 0$, *there exists* $\alpha \ge 0$ *such that for all* $P' \in P$

$$P'(E_\alpha) > 1 - \varepsilon \ ,$$

*is tight.*

PROOF. The first assertion follows from [4, Section 4] and the second one from the definition of tightness [8].□

Let

$$E^\#: = \{\text{epi } x \,|\, \sup_{t\in T}|x(t)| \le \alpha^\#\}$$

be the space of epigraphs associated to l.sc. functions that are bounded below and above by $\alpha^\#$. From Lemma 11.2, and Theorem 11.1, it follows directly that

PROPOSITION 11.3 *Any collection P of probability measures on* $B(E^\#)$ *is tight, and hence every subsequence has a convergent subsequence.*

## 12. AN APPLICATION TO GOODNESS-OF-FIT STATISTICS

Let us consider the basic case of independent observations $(\xi_1, \xi_2, \ldots, \xi_\nu)$ from the uniform distribution on [0, 1]. Let us define the empirical process

$$U_t^\nu(\omega) = \begin{cases} F^\nu(\omega, t) - t, & \text{if } 0 < t < 1, \\ 0 & \text{otherwise} . \end{cases}$$

where for every $\omega$, $F^\nu(\omega, \cdot)$ is the empiral distribution (taken left-continuous) determined by the sample $(\xi_1, \ldots, \xi_\nu)$. The realizations $U_\cdot^\nu(\omega)$ are l.sc. on [0, 1] (with respect to the natural topology on $\mathbf{R}$); this comes from the fact that $F^\nu$ is a left-continuous piecewise constant distribution function on $\mathbf{R}$. It is also easy to verify that the function

$$(\omega, t) \mapsto U_t^\nu(\omega) : [0, 1]^\nu \times [0, 1] \to [-1, 1]$$

is measurable. Redefining the underlying sample space to be $[0,1]^\infty$, and making the obvious identifications, we have that for all $\nu = 1, \ldots$

$$(\omega, t) \mapsto U_t^\nu(\omega) = [0, 1]^\infty \times [0, 1] \to [-1, 1]$$

is a random l.sc. function. We are here in the case when for all $\nu = 1, \dots$

$$\text{epi } U_\cdot^\nu \subset \{\text{epi } x \mid -1 \le x(t) \le 1,\ t \in [0, 1]\} =: E' \ .$$

Moreover, for all $\nu$, the corresponding distribution functions $\{D^\nu, \nu = 1, \dots\}$ are inner separable at K, for all K in $\mathcal{K}^{ub}$. This follows from the inner-separability of the distribution function associated to the stochastic process $\{F^\nu(\cdot, t), t \in [0. 1]\}$. Since, we may as well take for balls the products of intervals, we see that epi $F^\nu(\omega) \cap ([t_1, t_2] \times [\alpha_1, \alpha_2])$ only if $F^\nu(\omega, t_2) \le \alpha_1$, since $F^\nu$ is monotone nondecreasing. Thus for any finite collection of balls, the value of the associated distribution function is determined by its values on some finite set.

By Proposition 8.4, and the fact that the values of $D^\nu$ on $\mathcal{K}^{ub}$ determine uniquely its values on $\mathcal{K}$, we know that the finite dimensional distributions completely determine $D^\nu$. Moreover from Proposition 11.3, since the $\{U_t^\nu, t \in T\}$ are (equi-) bounded, the associated probability measures are tight. This means that there always exists a subsequence

$$\{D^{\nu_k}, k = 1, \dots\} \text{ converging } D \ ,$$

(i.e. $U_\cdot^\nu \xrightarrow{\text{i.d.}} U_\cdot$).

Observe that independence did not play any role up to now. If the $\{\xi_k, k = 1, \dots\}$ are i.i.d, by the law of large numbers, for every $I = (t_1, \dots, t_q)$, the finite dimensional distributions converge in distribution to the q-dimensional distribution of the random vector identically zero. And thus the limit process $\{U_t, t \in T\}$ must be a stochastic process whose realizations are such that

$$U_t(\omega) = 0 \quad \text{for all } t \in [0, 1] \ ,$$

and for all $\omega \in \Omega \setminus N$ where N is a set of measure 0.

Actually a somewhat stronger result does hold. From, the strong law of large numbers, for every $t \in T$

$$U_t^\nu = F^\nu(\cdot, t) - t \to 0 \quad \text{a.s.} \ ,$$

i.e. there exists a set $N_t$ of measure 0, such that

$$U_t^\nu(\omega) = F^\nu(\omega, t) - t \to 0 \quad \text{for all } \omega \in \Omega \setminus N_t \ . \tag{12.1}$$

We shall show that almost surely

$$\text{epi } U_{\cdot} = \lim_{\nu \to \infty} \text{epi } U_{\cdot}^{\nu}$$

Let $S = \{t_1, t_2, \dots\}$ be a countable dense subset of $T = [0, 1]$. Then by (12.1), we have that

$$U_{t_k}^{\nu}(\omega) = F^{\nu}(\omega, t_k) - t_k \to 0 \quad \text{for all } \omega \in \Omega \setminus N$$

where N is the null set

$$N := \cup_{k=1}^{\infty} N_{t_k} \ .$$

Now, it is an exercise in epi-convergence to show that for every $\omega \in \Omega \setminus N$

$$\limsup_{\nu \to \infty} \text{epi } U_{\cdot}(\omega) \subset \text{epi } U_{\cdot} \subset \liminf_{\nu \to \infty} \text{epi } U_{\cdot}^{\nu}(\omega) \ ,$$

where $\lim \sup_{\nu \to \infty}$ and $\lim \inf_{\nu \to \infty}$ are the superior and inferior limits of sets [4, 12]. In fact it suffices to show that for all $\omega \in \Omega \setminus N$, $t \in [0, 1]$

$$\text{- for all } t_k \to t, \ (k) \subset (\nu) : \liminf_{k \to \infty} U_{t_k}^{k}(\omega) \geq 0 \ , \tag{12.2}$$

and

$$\text{- there exists } t_{\nu} \to t : \limsup_{\nu \to \infty} U_{t_{\nu}}^{\nu}(\omega) \leq 0 \ . \tag{12.3}$$

Condition (12.3) is immediate. For, let $t \in T$, $\varepsilon > 0$ and take $t_{\nu} \in S$ with $t_{\nu} \in [t, t + \varepsilon)$. We have

$$F^{\nu}(\omega, t) - t \leq F^{\nu}(\omega, t_{\nu}) - t_{\nu} + \varepsilon$$

Hence

$$\limsup_{\nu \to \infty} (U_t^{\nu}(\omega) = F^{\nu}(\omega, t) - t) \leq \varepsilon \ ,$$

and since $\varepsilon > 0$ is arbitrary, (12.3) follows.

Now let $t_k \to t$ and $(\nu_k)$ be a subsequence of $(\nu)$. For any $\varepsilon > 0$, fix $t_{\varepsilon} \in D$ such that $t_{\varepsilon} \in (t - \varepsilon, t]$. Since $t_k \to t$, there is $k_{\varepsilon}$ such that for all $k \geq k_{\varepsilon}$,

$$t_{\varepsilon} < t_k < t_{\varepsilon} + \varepsilon$$

Thus for all $\nu \in (\nu_k)$ with $\nu \geq \nu_{k_{\varepsilon}}$, for all $\omega \in \Omega \setminus N$, we have

$$F^{\nu}(\omega, t_k) - t_k \geq F^{\nu}(\omega, t_\varepsilon) - t_\varepsilon - \varepsilon \ ,$$

since $F^{\nu}$ is monotone increasing with respect to t. This implies that for all $\omega \in \Omega \setminus N$,

$$\liminf_{k \to \infty} U_{t_k}^{\nu_k}(\omega) \geq -\varepsilon$$

Since $\varepsilon > 0$ is arbitrary, it yields (12.2).

Almost sure epi-convergence implies convergence in distribution [12, Section 3] and thus

$$U_{\cdot}^{\nu} \xrightarrow{\text{l.d.}} U_{\cdot} \qquad (12.4)$$

Glivenko-Cantelli's Theorem is a corollary of epi-convergence in distribution, as we see next.

GLIVENKO–CANTELLI'S THEOREM 12.1 $\quad \text{Sup}_{t \in [0,1]} |U_t^{\nu}(\omega)| \to 0$, a.s.

PROOF Suppose to the contrary that for some $\omega \in \Omega \setminus N$, and $\varepsilon > 0$, there is a subsequence $(\nu_k)$ of $(\nu)$ such that $\sup_t |U_t^{\nu_k}(\omega)| > \varepsilon$. This means that there exists for each k, $t_k$ such that $|U_{t_k}^{\nu_k}(\omega)| > \varepsilon$. Passing to a subsequence if necessary, let t be the limit of $\{t_k, k = 1, \ldots\}$, then

$$\text{either} \quad U_{t_k}^{\nu_k}(\omega) > \varepsilon, \quad \text{or} \quad U_{t_k}^{\nu_k}(\omega) < -\varepsilon \ .$$

If the second inequality occurred infinitely often, then for some subsequence we would have that

$$\liminf_{k' \to \infty} U_{t_{k'}}^{k'}(\omega) < -\varepsilon \ .$$

which does contradict the epi-convergence of the $U_{\cdot}^{\nu}$ to $U_{\cdot}$. If $U_{t_k}^{\nu_k}(\omega) > \varepsilon$ infinitely often, then

$$\varepsilon \leq \limsup_{k \to \infty} U_{t_k}^{\nu_k}(\omega) \ .$$

If $t' = 1$ then $t_k \leq 1$ and $t_k > 1 - \varepsilon/2$ for k sufficiently large. The preceding inequality then implies that

$$\varepsilon \le \limsup_{k \to \infty} U_{t_k}^{\nu_k}(\omega) \le \limsup_{k \to \infty} U_1^{\nu_k}(\omega) + \frac{\varepsilon}{2} = \frac{\varepsilon}{2} \quad ,$$

recall that $U_1^{\nu}(\omega) = 0$, see the definition. If $t' \in [0, 1]$, there exists $\varepsilon' > 0$, $2\varepsilon' < \varepsilon$ such that for $k$ sufficiently large

$$t' - \varepsilon' < t_k < t' + \varepsilon' \quad .$$

Then, from the proof given for (12.3), it follows that

$$\varepsilon \le \limsup_{k \to \infty} U_{t_k}^{\nu_k}(\omega)$$

$$\le \limsup_{k \to \infty} [U_{t'+\varepsilon'}(\omega) + 2\varepsilon'] < \varepsilon \quad .$$

This is again a contradiction, and the proof is complete.□

## REFERENCES


1. Ash, R., *Real Analysis and Probability*, Academic Press, New York, 1972.
2. Attouch, H., *Variational Convergence for Functions and Operators*, Pitman Co., Boston, 1984.
3. Billingsley, P., *Probability and Measure*, J. Wiley & Sons, New York, 1979.
4. Dolecki, S., G. Salinetti, and R. Wets, "Convergence of functions: equi-semicontinuity," *Transaction American Mathematical Society*, vol. 276, pp. 409-429, 1983.
5. Gikhman, I.I. and A.V. Skorohod, *Introduction to the Theory of Random Processes*, W.B. Saunders Co., Philadephia, 1969.
6. Kelley, J.L., *General Topology*, Van Nostrand, New York, 1955.
7. Matheron, G., *Random Sets and Integral Geometry*, John Wiley & Sons, New York, 1975.
8. Parthasaraty, K.R., *Probability Measures on Metric Spaces*, Academic Press, New York, 1967.
9. Rockafellar, R.T., "Integrals functionals, normal integrands and measurable selections," in *Nonlinear Operators and Calculus of Variations*, ed. L. Waelbroeck, pp. 157-207, Springer Verlag, Berlin, 1976. Lecture Notes in Mathematics 543
10. Rockafelllar, R.T., "Convex integral functionals and duality," in *Contribution to Nonlinear Analysis*, ed. E. Zarantonello, pp. 215-236, Academic Press, New York, 1971.
11. Salinetti, G. and R. Wets, "On the convergence of closed-valued measurable multifunctions," *Transaction American Mathematical Society*, vol. 266, pp. 275-289, 1981.

12. Salinetti, G. and R. Wets, "On the convergence in distribution of measurable multifunctions (random sets), normal integrands, stochastic processes and stochastic infima," *Mathematics of Operations Research*, vol. 11, pp. 385–419, 1986.