Proceedings of Machine Learning Research 147:132–139, 2021 ISIPTA 2021

Credal Sets of Coherent Conditional Probabilities Defined by Hausdorff Measures

Serena Doria
Department of Engineering and Geology, University G.d'Annunzio, Chieti-Pescara, Italy

Abstract

Credal sets containing coherent conditional probabilities defined by Hausdorff measures on the Borel sigma-field of metric spaces with bi-Lipschitz equivalent metrics, are proven to represent merging opinions with increasing information.

Keywords: Coherent conditional probabilities, Hausdorff measures, bi-Lipschitz equivalent metrics, topological equivalent metrics, absolutely continuity

1. Introduction

A new model of coherent upper conditional previsions defined in a metric space by Hausdorff outer measures has been introduced to represent partial knowledge ([8], [9],[10], [11], [14]). The conditioning event represents the piece of information we have and the complexity of information is expressed in terms of Hausdorff dimension of the conditioning event. A natural question is to investigate the relation between partial knowledge produced in different metric spaces by the same piece of information. Hausdorff dimension where introduced in probability theory [2] to compute the dimensions of various sets where the strong low of large number is violated in a Markov chain. Hausdorff measures of subsets with respect to different metrics can be very different and the same subset can have different Hausdorff dimensions. If the metrics are bi-Lipschitz equivalent then a set has the same Hausdorff dimension in the two metric spaces and the Hausdorff measures are proven to be mutually absolutely continuous.

We prove that, if the metric space (Ω, d), where Ω is a set with positive and finite Hausdorff outer measure in its Hausdorff dimension, and d' is any metric bi-Lipschitz equivalent to d, and so topologically equivalent to d, then the conditional probabilities defined by Hausdorff measures on (Ω, d) and on (Ω, d') are mutually absolutely continuous. So given a metric space (Ω, d) and a set B with positive and finite Hausdorff outer measure in its Hausdorff dimension, we can consider the credal set $\mathbf{K}_B$ of conditional probability measures defined by Hausdorff measures, which are mutually absolutely continuous with respect to the conditional probability defined by Hausdorff measures with respect the metric d. In [3] it is proven that distance between two conditional probabilities $P(\cdot|G_n)$ and $Q(\cdot|G_n)$ defined on the same σ-field goes to zero, except on a Q-probability zero set, if Q is absolutely continuous with respect to P. This result, based on martingale convergence theorem, establishes merging of opinions with increasing information. Weak merging notions have been investigated in [18]. The aim of this paper is to verify if a similar result holds for coherent conditional prevision defined by Hausdorff measures. with respect to bi-Lipschitz equivalent metrics.

2. Coherent Upper Conditional Previsions

Let Ω be a non empty set, let $\mathbf{B}$ be a partition of Ω and denote by $\wp(\Omega)$, the family of all subsets of Ω . A random variable is a function $X : \Omega \to \widehat{\Re} = \Re \cup \{-\infty, +\infty\}$, let $\mathscr{U}(\Omega)$ be the class of all random variables. $\mathscr{U}(\Omega)$ is not a linear space in fact, if random variables take values $-\infty, +\infty$ then the sum between two of them can be not defined (when for the same ω one takes value $+\infty$ and the others $-\infty$). Let $L(\Omega) \subset \mathscr{U}(\Omega)$ be the linear space of all bounded random variables defined on Ω; for every $B \in \mathbf{B}$ denote by $X|B$ the restriction of X to B and by $\sup(X|B)$ the supremum of values that X assumes on B. Let $L(B)$ be the linear space of all bounded random variables $X|B$. Denote by I_A the indicator function of any event $A \in \wp(B)$, i.e. $I_A(\omega) = 1$ if $\omega \in A$ and $I_A(\omega) = 0$ if $\omega \in A^c$. For every $B \in \mathbf{B}$ let $\mathscr{K}(B)$ be a linear space of random variables $X|B$ with $X \in \mathscr{U}(\Omega)$. Coherent upper conditional previsions $\overline{P}(\cdot|B)$ are real valued functionals defined on a linear space $\mathscr{K}(B)$.

Definition 1 *Coherent upper conditional previsions are functionals $\overline{P}(\cdot|B)$ defined on a linear space $\mathscr{K}(B)$ with values in the real number, such that the following axioms of coherence hold for every X and Y in $\mathscr{K}(B)$ and every strictly positive constant λ:*

1) $\overline{P}(X|B) \le \sup(X|B)$;

2) $\overline{P}(\lambda X|B) = \lambda \overline{P}(X|B)$ (positive homogeneity);

3) $\overline{P}(X+Y|B) \le \overline{P}(X|B) + \overline{P}(Y|B)$ (subadditivity).

If $\mathscr{K}(B)$ coincides with $L(B)$ the previous definition is the definition of coherent upper conditional prevision given in Walley [23], [24]. Suppose that $\overline{P}(X|B)$ is a coherent upper conditional prevision on $\mathscr{K}$. Then its conjugate coherent lower conditional prevision is defined by the conjugacy property $\underline{P}(X|B) = -\overline{P}(-X|B)$. If for every X belonging to $\mathscr{K}(B)$ we have $P(X|B) = \underline{P}(X|B) = \overline{P}(X|B)$ then $P(X|B)$ is called a coherent *linear* conditional prevision and if $\mathscr{K} = L(B)$ it is a linear, positive and positively homogenous functional in the sense of de Finetti [5] [6], Regazzini [20][21] and Walley [24, Corollary 2.8.5].

From axioms 1)-3) and by the conjugacy property we have that

$$1 \le \underline{P}(I_B|B) \le \overline{P}(I_B|B) \le 1$$

so that

$$\underline{P}(I_B|B) = \overline{P}(I_B|B) = 1$$

In Walley [24] the functionals $\overline{P}(X|B)$ defined for $B \in \mathbf{B}$ and $X \in L(B)$ satisfying axioms 1)-3) and such that $\overline{P}(I_B|B) = 1$ are called *separately coherent*.

The unconditional coherent upper prevision $\overline{P} = \overline{P}(\cdot|\Omega)$ is obtained as a particular case when the conditioning event is Ω. Coherent upper conditional probabilities are obtained when only 0-1 valued random variables are considered.

Definition 2 *Given a partition $\mathbf{B}$ and a random variable $X \in L(\Omega)$, a coherent upper conditional prevision $\overline{P}(X|\mathbf{B})$ is a random variable on Ω equal to $\overline{P}(X|B)$ if $\omega \in B$.*

Definition 3 *A bounded random variable $X \in L(\Omega)$ is called $\mathbf{B}$-measurable or measurable with respect to a partition $\mathbf{B}$ of Ω if it is constant on the atoms of the partition.*

The following necessary condition for coherence holds [24, p. 292]:

Proposition 1 *If $P(X|B)$ is a coherent linear previsions for every B that belongs to a partition $\mathbf{B}$ of Ω then $P(X|\mathbf{B}) = X$ for all random variables $X \in L(\Omega)$ that are $\mathbf{B}$-measurable.*

2.1. Coherent Conditional Prevision and Conditional Expectation

In the axiomatic approach [1, Section 34] conditional expectation is defined with respect to a σ-field of conditioning events by the Radon-Nikodym derivative.

Let $\mathbf{F}$ and $\mathbf{G}$ be two σ-field of subsets of Ω with $\mathbf{G}$ contained in $\mathbf{F}$ and let X be an integral random variable. Let P be a probability measure on $\mathbf{F}$; define a measure ν on $\mathbf{G}$ by $\nu(G) = \int_G X dP$. This measure is finite and absolutely continuous with respect to P. Thus there exists a non-negative function, the Radon-Nikodym derivative denoted by $E[X|\mathbf{G}]$, defined on Ω, i) $\mathbf{G}$-measurable, ii) integrable and satisfying the functional equation:

iii) $\int_G E[X|\mathbf{G}]_\omega dP = \int_G X dP$ with $G \in \mathbf{G}$

The Radon-Nikodym is a non-negative function but it is not restrictive to use it to define conditional expectation for any random variable. In fact if X is non-positive, it can be decomposed in $X = X^+ - X^-$ where X^+ is its *positive part* and X^- is its *negative part* which are non-negative functions given by:

$$X^+ = 0 \vee X; \quad X^- = (-X)^+$$

and $\vee$ is the maximum. The function $E[X|\mathbf{G}] = E[X^+|\mathbf{G}] - E[X^-|\mathbf{G}]$ satisfies the properties i), ii), iii). This function is unique up to a set of P-measure zero and it is a version of the conditional expected value.

The next theorem, proven in [12] shows that every time the σ-field $\mathbf{G}$ is properly contained in $\mathbf{F}$ and it contains all singletons of $[0,1]$ then the conditional prevision defined by the Radon-Nikodym derivative is not coherent. It occurs because one of the defining properties of the Radon-Nikodym derivative is to be measurable with respect to the σ-field of the conditioning events and this requirement contradicts the necessary condition for the coherence of a linear conditional prevision recalled in Proposition 1.

Theorem 1 *Let $\Omega = [0,1]^n$ and let $\mathbf{F}$ and $\mathbf{G}$ be two σ-field of subsets of Ω such that $\mathbf{G}$ is properly contained in $\mathbf{F}$ and it contains all singletons of Ω. Let $\mathbf{B}$ be the partition of singletons and let X be the indicator function of an event A belonging to $\mathbf{F} - \mathbf{G}$. If we define the conditional prevision $P(X|\mathbf{B})$ equal to the Radon-Nikodym derivative with probability 1, that is*

$$P(X|\mathbf{B}) = E[X|\mathbf{G}]$$

except on a subset N of $[0,1]^n$ of P-measure zero, then the conditional prevision $P(X|\mathbf{B})$ is not coherent.

3. The Model

A new model of coherent upper conditional probability based on Hausdorff outer measures on a metric space has been introduced for bounded and unbounded random variables [14].

Hausdorff outer measures are examples of outer measures defined on a metric space.

Let (Ω, d) be a metric space. The topology $\mathscr{T}$ induced by the metric d contains the empty set and the sets which are countable or finite unions of the sets $D_r(x) = \{\omega \in \Omega : d(\omega, x) < r\}$ with $r \ge 0$ and $x \in \Omega$. These sets in the topology $\mathscr{T}$ are called open sets. The Borel σ-field $\mathscr{B}$ is the smallest σ-field containing all open sets of Ω.

The diameter of a non-empty set U of Ω is defined as $|U| = \sup\left\{d(x,y) : x, y \in U\right\}$ and if a subset A of Ω is such that $A \subseteq \bigcup_i U_i$ and $0 \le |U_i| < \delta$ for each i, the countable class $\left\{U_i\right\}$ is called a δ-cover of A.

Let s be a non-negative number. For $\delta > 0$ we define $h_{s,\delta}(A) = \inf \sum_{i=1}^{\infty} |U_i|^s$, where the infimum is over all δ-covers $\{U_i\}$.

The *Hausdorff s-dimensional outer measure* of A ([22], [17]) denoted by $h^s(A)$, is defined as

$$h^s(A) = \lim_{\delta \to 0} h_{s,\delta}(A).$$

This limit exists, but may be infinite, since $h_{s,\delta}(A)$ increases as δ decreases.

A subset A of Ω is called *measurable* with respect to the outer measure h^s defined on $\wp(\Omega)$ if it decomposes every subset of Ω additively, that is if

$$h^s(E) = h^s(A \cap E) + h^s(A^c \cap E)$$

for all sets $E \subseteq \Omega$.

Hausdorff outer measures are *metric outer measures*, that is if E and F are positively separated, i.e.

$$d(E,F) = \inf\{d(x,y) : x \in E, y \in F\} > 0.$$

then

$$h^s(E \cup F) = h^s(E) + h^s(F).$$

By Theorem 1.5 of Falconer [17] since Hausdorff outer measures are metric outer measures then all Borel subsets of Ω are measurable.

For any set E the Hausdorff outer measure $h^s(E)$ is non-increasing as s increases from 0 to $+\infty$

The *Hausdorff dimension* of a set A, $dim_H(A)$, is defined as the unique value, such that

$$h^s(A) = \infty \text{ if } 0 \leq s < dim_H(A),$$
$$h^s(A) = 0 \text{ if } dim_H(A) < s < \infty.$$

The following theorem has been proven in [8].

Theorem 2 *Let (Ω, d) be a metric space and let $\mathbf{B}$ be a partition of Ω. For $B \in \mathbf{B}$ denote by s the Hausdorff dimension of the conditioning event B and by h^s the Hausdorff s-dimensional outer measure. Let m_B be a 0-1 valued finitely additive, but not countably additive, probability on $\wp(B)$. Thus, for each $B \in \mathbf{B}$, the function defined on $\wp(B)$ by*

$$\overline{P}(A|B) = \begin{cases} \frac{h^s(A\cap B)}{h^s(B)} & if \quad 0 < h^s(B) < +\infty \\ m_B(A \cap B) & if \quad h^s(B) \in \{0, +\infty\} \end{cases}$$

is a coherent upper conditional probability.

The coherent upper unconditional probability $\overline{P} = \mu^*_{\Omega}$ defined on $\wp(\Omega)$ is obtained for B equal to Ω.

If $B \in \mathbf{B}$ is a set with positive and finite Hausdorff outer measure in its Hausdorff dimension s the set function μ^*_B defined for every $A \in \wp(B)$ by $\mu^*_B(A) = \frac{h^s(A\cap B)}{h^s(B)}$ is a coherent upper conditional probability, which is

a) monotone, that is $\mu^*_B(\oslash) = 0$ and if A, $B \in \mathscr{S}$ with $A \subseteq B$ then $\mu^*_B(A) \leq \mu^*_B(B)$,

b) *submodular* or *2-alternating*, that is $\mu(A \cup E) + \mu(A \cap E) \leq \mu(A) + \mu(E)$ for every $A, E \in \wp(B)$,

c) *continuous from below* that is $\lim_{i\to\infty} \mu(A_i) = \mu(\lim_{i\to\infty} A_i)$ for any increasing sequence of sets $\{A_i\}$, with $A_i \in \wp(B)$

and such that its restriction to the σ-field of all μ^*_B measurable sets is a Borel regular countably additive probability.

If $B \in \mathbf{B}$ is such that $h^s(B) \in \{0, +\infty\}$ then the coherent upper conditional probability is defined by a 0-1 valued finitely additive, but not countably additive, probability m_B on $\wp(B)$. The existence of m_B is a consequence of the prime ideal theorem and any m_B is coherent. 0-1 valued finitely additive probabilities are in correspondence one-to-one with ultrafilter $\mathscr{A}$.

Definition 4 *An ultrafilter $\mathscr{A}$ is a class of subsets of $\wp(B)$ such that*

a) $\oslash \notin \mathscr{A}$

b) $A, E \in \mathscr{A} \Rightarrow A \cap E \in \mathscr{A}$

c) $A \in \mathscr{A}; A \subset E \subset \Omega \Rightarrow E \in \mathscr{A}$

d) $\forall A \in \wp(\Omega)$ *either* $A \in \mathscr{A}$ *or* $A^c \in \mathscr{A}$

Given an ultrafilter $\mathscr{A} \subset \wp(B)$ a 0-1 valued finitely additive probability m_B on $\wp(B)$ can be defined by $m_B(A) = 1$ if $A \in \mathscr{A}$ and $m_B(A) = 0$ if $A^c \in \mathscr{A}$.

Example 1 *Let $\mathscr{A}$ be the ultrafilter of Ω of sets whose complement is a finite set. Then $m_{\Omega}(A) = 0$ if A is any finite set and $m_{\Omega}(A) = 1$ otherwise.*

Example 2 *Let $\Omega = \mathscr{N}$, let $\mathscr{A}$ be the ultrafilter of Ω of sets whose complement is a finite set and let $A = \{2n : n \in \mathscr{N}\}$. Then by property d) of Definition 5 we can assume $m_{\Omega}(A) = 1$ if $A \in \mathscr{A}$ and $m_{\Omega}(A^c) = 0$ or $m_{\Omega}(A) = 0$ if $A^c \in \mathscr{A}$ and $m_{\Omega}(A^c) = 1$.*

3.1. Coherent Upper Conditional Previsions for Bounded and Unbounded Random Variables

Given a non-empty set Ω and denoted by $\wp(\Omega)$, the family of all subsets of Ω, let $\mathscr{S}$ a class properly contained in $\wp(\Omega)$ containing Ω.

The definition of Choquet integral given in [7] is recalled.

A monotone set function $\mu : \mathscr{S} \to \overline{\Re}_+ = \Re_+ \cup \{+\infty\}$ is such that $\mu(\oslash) = 0$ and if A, $B \in \mathscr{S}$ with $A \subseteq B$ then $\mu(A) \leq \mu(B)$. Given a monotone set function μ on $\mathscr{S}$ the *outer set function* of μ is the set function μ^* defined on the whole power set $\wp(\Omega)$ by

$$\mu^*(A) = \inf\{\mu(B) : B \subseteq A; B \in S\}, A \in \wp(\Omega)$$

The inner set function of μ is the set function μ_* defined on the whole power set $\wp(\Omega)$ by

$$\mu_*(A) = \sup\{\mu(B) | B \subseteq A; B \in S\}, A \in \wp(\Omega)$$

Let $X : \Omega \to \overline{\Re} = \Re \cup \left\{-\infty, +\infty\right\}$ a random variable on Ω. Then the function

$$G_{\mu,X}(x) = \mu\left\{\omega \in \Omega : X(\omega) > x\right\}$$

is decreasing and it is called *decreasing distribution function* of X with respect to μ. If μ is continuous from below then $G_{\mu,X}(x)$ is right continuous.

In particular the decreasing distribution function of X with respect to the Hausdorff outer measures is right continuous since these outer measures are continuous from below.

A function $X : \Omega \to \overline{\Re}$ is called upper μ-measurable if $G_{\mu^*,X}(x) = G_{\mu_*,X}(x)$ [7]. Given an upper μ-measurable random variable $X : \Omega \to \overline{\Re}$ with decreasing distribution function $G_{\mu,X}(x)$, the Choquet integral of X with respect to μ is defined if $\mu(\Omega) < +\infty$ through

$$\int X d\mu = \int_{-\infty}^{0}(G_{\mu,X}(x) - \mu(\Omega))dx + \int_{0}^{+\infty} G_{\mu,X}(x)dx$$

where the integrals in the second member of the formula are Riemann integrals.

The integral can assume real values or can assume the values $-\infty$, $+\infty$ or it cannot exist.

Definition 5 *Let Ω be a set with positive and finite Hausdorff outer measure in its Hausdorff dimension s and let $\mathscr{S}$ be the σ-field of the h^s-measurable subsets of Ω. An $\mathscr{S}$-measurable random variable X is Choquet integrable with respect to a monotone set function μ_B^* if the Choquet integral is finite, that is*

$$-\infty < \int X d\mu_B^* < +\infty$$

Let $L^*(B)$ be the class of random variables which are Choquet integrable with respect to μ_B^* and with respect to the dual $\overline{\mu}_B^*$ defined by

$$\overline{\mu}_B^*(A) = \mu_B^*(B) - \mu_B^*(A^c).$$

In Theorem 11 of [13] $L^*(B)$ has been proven to be a linear space.

Since Hausdorff outer measures are submodular by Denneberg [7, Proposition 9.3] $L^*(B)$ coincides with the linear space of all absolutely Choquet integrable random variables on B, i.e. the random variables X such that

$$-\infty < \frac{1}{h^s(\Omega)} \int_B |X| dh^s < +\infty \text{ if } 0 < h^s(B) < +\infty$$

Since B has positive and finite Hausdorff outer measure in its Hausdorff dimension s then $L^*(B)$ contains also all constants. In [14] the following theorem has been proven:

Theorem 3 *Let (Ω, d) be a metric space and let $\mathbf{B}$ be a partition of Ω. For $B \in \mathbf{B}$ denote by s the Hausdorff dimension of the conditioning event B and by h^s the Hausdorff s-dimensional outer measure. Let m_B be a 0-1 valued finitely additive, but not countably additive, probability on $\wp(B)$. Then for each $B \in \mathbf{B}$ the functional $\overline{P}(X|B)$ defined on the linear space $L^*(B)$ by*

$$\overline{P}(X|B) = \begin{cases} \frac{1}{h^s(B)} \int_B X dh^s & if \quad 0 < h^s(B) < +\infty \\ \int_B X dm_B & if \quad h^s(B) \in \{0, +\infty\} \end{cases}$$

is a coherent upper conditional prevision if B has positive and finite Hausdorff measure in its Hausdorff dimension and it is a linear prevision whose restriction to events assumes only the values 0-1 if B has Hausdorff outer measure in its Hausdorff dimension equal to zero or infinity.

In Theorem 7 of [14] it is proven that coherent upper conditional previsions defined by Hausdorff outer measures as in Theorem 2 satisfy the disintegration property $\overline{P}(\overline{P}(X|\mathbf{B})) = \overline{P}(X)$ for every random variable $X \in L^*(\Omega)$ and for every partition, whose atoms are h^s-measurable where s is the Hausdorff dimension of Ω.

4. Absolute Continuity of Coherent Conditional Probability Measures Defined by Hausdorff Measures

In this section probability measures defined on the Borel σ-field of a metric space (Ω, d) by Hausdorff measures as in Theorem 1, are proven to be absolutely continuous with respect to any probability measure defined by Theorem 1 in a metric space (Ω, d') where d' is a bounded metric bi-Lipschitz equivalent to the metric d. It occurs because events which have zero Hausdorff measure in a metric space have also Hausdorff measure equal to zero in a metric space with a bi-Lipschitz equivalent metric.

An example is given to show that probability measures defined by Hausdorff measures in metric spaces that are topological equivalent are not absolutely continuous.

Two different notions of equivalence can be considered for metrics: Bi-Lipschitz equivalence [15] and topological equivalence.

Definition 6 *Let (Ω, d) be a metric space; a metric d' on Ω is bi-Lipschitz equivalent to the metric d if there exist two positive real constants α, β such that*

$$\alpha d'(x,y) \leq d(x.y) \leq \beta d'(x,y)$$

Definition 7 *Let (Ω, d) be a metric space and let d' be a metric on Ω; d and d' are topological equivalent if they induce the same topology.*

Proposition 2 *Let (Ω, d) be a metric space and let d' be a metric on Ω bi-Lipschitz equivalent to d, then d and d' are topological equivalent.*

The following example shows that the converse is not true.

Example 3 *Let $(\Re^n, d)$ be the Euclidean metric space and let d' be a metric on $\Re^n$ defined $\forall \bar{x}, \bar{y} \in \Re^n$ by*

$$d'(\bar{x}, \bar{y}) = \frac{d(\bar{x}, \bar{y})}{1 + d(\bar{x}, \bar{y})};$$

d' is topological equivalent to the Euclidean metric d but it is not bi-Lipschitz equivalent to d since there not exist two positive real constants constants α, β such that $\alpha d'(x,y) \leq d(x.y) \leq \beta d'(x,y)$

Theorem 4 *Let (Ω, d) be a metric space, let d and d' be two metrics on Ω bi-Lipschitz equivalent and let h^s and h_1^s be the s-dimensional Hausdorff measures defined respectively in the metric space (Ω, d) and (Ω, d'), then there exist two positive real constants α, β such that*

$$\alpha h_1^s(E) \leq h^s(E) \leq \beta h_1^s(E)$$

Proof The result follows by the definition of Hausdorff outer measures and by the fact that the metrics are bi-Lipschtz equivalent (see Lemma 1.8 of [17]). ■

Theorem 5 *Let (Ω, d) be a metric space and let d' be a metric on Ω bi-Lipschitz equivalent to d. Then the Hausdorff dimension of any set $A \in \wp(\Omega)$ is invariant in the two metric spaces (Ω, d) and (Ω, d')*

The Hausdorff dimension of any set $A \in \wp(\Omega)$ is not invariant with respect to two topological equivalent metrics which are not bi-Lipschitz equivalent.

Example 4 *Let $\Omega = [0,1]$ and let d be the Euclidean metric*

$$d(\omega_1, \omega_2) = |\omega_1 - \omega_2|$$

and let d' the discrete distance, that is

$$d'(\omega_1, \omega_2) = \begin{cases} 0 & if \quad \omega_1 = \omega_2 \\ 1 & \quad otherwise. \end{cases}$$

d and d' are not topologically equivalent; in fact all subsets of Ω are open sets in the topology induced by d' since $D_r(x) = \{\omega \in \Omega : d(\omega, x) < r\} = \{x\}$ if $r < 1$ and $D_r(x) = \{\omega \in \Omega : d(\omega, x) < r\} = \Omega$ if $r \geq 1$, while singletons are not open sets in the topology induced by the Euclidean metric.

Example 5 *Let $(\Re^2, d)$ be the Euclidean metric space and let d'' be the metric defined by*

$$d''(x,y) = max\{|x_1 - y_1|; |x_2 - y_2|\}$$

where $x = (x_1, x_2)$ and $y = (y_1, y_2)$.

Then d and d'' are topologically equivalent.

The notion of boundness of a set depends on the metric.

Definition 8 *A metric on Ω is bounded if $diam(\Omega)$ is bounded. A metric space (Ω, d) is bounded if d is bounded.*

Proposition 3 *Let (Ω, d) be a metric space where Ω has positive and finite Hausdorff measures in its Hausdorff dimension then (Ω, d) be a bounded metric space.*

Definition 9 *Let μ and ν be two probabilities measures on the same σ-field $\mathscr{F}$ then ν is absolutely continuous with respect to μ, $\mu << \nu$, iff $\mu(A) = 0 \Rightarrow \nu(A) = 0$ for every $A \in \mathscr{F}$.*

Theorem 6 *Let (Ω, d) be a bounded metric space where Ω is a set with positive and finite Hausdorff measure in its Hausdorff dimension, let $\boldsymbol{B}$ be a partition of Ω and let $\mathscr{B}$ be the Borel σ-field induced by the metric d. Let d' be a bounded metric on Ω bi-Lipschitz equivalent to d. Then for every $B \in \boldsymbol{B}$ with positive and finite Hausdorff outer measures in its dimensions in both metric spaces, the restrictions μ_B and ν_B on $\mathscr{B}$ of the coherent upper conditional probabilities defined respectively in (Ω, d) and (Ω, d') as in Theorem 1, are countably additive probabilities which are mutually absolutely continuous.*

Proof Since d' is bi-Lipschitz equivalent, and so topologically equivalent, to d then d and d' induce the same topology and the same Borel σ-field $\mathscr{B}$. Since d' is a bounded metric then by Theorem 9 Ω has positive and finite Hausdorff measure in its Hausdorff dimension also in (Ω, d'). Let s be the Hausdorff dimension of B , let h^s and h_1^s be the s-dimensional Hausdorff measure in the two metric spaces and let μ_B and ν_B be the two probability measures on $\mathscr{B}$ defined by

$$\mu_B(A) = \frac{h^s(A \cap B)}{h^s(B)} \text{ and } \nu_B(A) = \frac{h_1^s(A \cap B)}{h_1^s(B)}.$$

Since d' is bi-Lipschitz equivalent to d by Theorem 9 we have that there exist two positive real constants α and β such that

$$\alpha \nu(A) = \alpha \frac{h_1^s(A)}{h^s(\Omega)} \leq \mu(A) = \frac{h^s(A)}{h^s(\Omega)} \leq \beta \frac{h_1^s(A)}{h_1^s(\Omega)} = \beta \nu(A)$$

so that $\nu(A) = 0$ implies $\mu(A) = 0$ and $\mu(A) = 0$ implies $\nu(A) = 0$. ■

5. Credal Sets of Coherent Countably Additive Conditional Probabilities Defined by Hausdorff Measures with Respect to Bounded Bi-Lipschitz Equivalent Metrics

In this section we prove that distance between coherent conditional probabilities defined by Hausdorff measures with respect to metrics which are bi-Lipschitz equivalent goes to zero when the information increases. So the credal set [19] containing all these coherent conditional probabilities represents opinions which merge when information increases [3].

Definition 10 *Let (Ω, d) and (Ω, d_i) be two metric spaces and let $\boldsymbol{B}$ be a partition of Ω. Let $B \in \boldsymbol{B}$ be a set with positive and finite Hausdorff outer measures in its dimensions in both metric spaces and denote by μ_B and ν_B^i the coherent conditional probabilities defined on the Borel σ-field $\mathscr{B}$ by Theorem 1 in the two metric spaces. The distance between μ_B and ν_B^i is defined by*

$$\sup |\mu_B(D) - \nu_B^i(D)|$$

where the supremum is token over $D \in \mathscr{B}$

In the paper of Blackwell and Dubins [3] it is shown that, given a monotone increasing or monotone decreasing sequence of σ-fields $\{\mathbf{G}_n\}$, the distance between two conditional probabilities defined in the axiomatic way on the same σ-field,, $P(\cdot|\mathbf{G}_n)$ and $Q(\cdot|\mathbf{G}_n)$, goes to zero as n goes to $+\infty$ except on a Q-probability zero set, if Q is absolutely continuous with respect to P. This result is an application of the Radon-Nikodym derivative and the generalized martingale convergence theorem. In the next section martingales with respect to coherent conditional previsions defined by Hausdorff measures are introduced.

5.1. Martingales with Respect to Coherent Conditional Previsions Defined by Hausdorff Measures

Martingales are defined when coherent conditional previsions are defined by Hausdorff measures and not by the Radon-Nykodym derivative. Some generalized martingale convergence theorems are proven.

The σ-field $\mathbf{F}$ generated by a finite or countable partition $\mathbf{B}$ of Ω contains sets that are finite or countable union of the atoms of the partition. It is the smallest σ-field contains the partition $\mathbf{B}$. Then the coherent upper conditional prevision $\overline{P}(X|\mathbf{F})$ is the random variable defined on Ω that associates to each $\omega \in \Omega$ the value $\overline{P}(X|\mathbf{F}) = \overline{P}(X|B)$ if ω belongs to B.

Definition 11 *Let (Ω, d) be a metric space, where Ω is a set with positive and finite Hausdorff outer measures in its Hausdorff dimension s. Let $\{\boldsymbol{B}_n\}$ be a sequence of Borel finite or countable partitions of Ω and let $\boldsymbol{F}_n$ be the σ-field generated by $\boldsymbol{B}_1, \boldsymbol{B}_2, .., \boldsymbol{B}_n$. We have that $\boldsymbol{F}_n \subseteq \boldsymbol{F}_{n+1}$ for all $n \in N$ and $\mathscr{F} = \bigcup_n \boldsymbol{F}_n$ the σ-field generated by all $\boldsymbol{F}_n$ then $\mathscr{F} = \mathscr{B}$. Let $X_1, X_2, \ldots$ be a sequence of Borel-measurable random variables in $L^*(\Omega)$. The sequence $\{(X_n, \boldsymbol{F}_n) : n = 1, 2, \ldots\}$ is a martingale if*

$$P(X_{n+1}|\boldsymbol{F}_n) = X_n.$$

Example 6 *Let $Z \in L^*(\Omega)$ and $\boldsymbol{F}_n$ non- decreasing Borel-σ-fields. Then*

$$\{(X_n, \boldsymbol{F}_n) : n = 1, 2, \ldots\} = \{P(X|\boldsymbol{F}_n), n = 1, 2, \ldots\}$$

is a martingale relative to $\{\boldsymbol{F}_n . n = 1, 2, \ldots\}$.

Remark 12 *The difference with the axiomatic definition (see for example [1, Section 35]) is that in Definition 16 the random variables X_n are not required to be measurable with respect to the σ-field of the conditioning events $\boldsymbol{F}_n$. Coherent conditional probabilities defined by Hausdorff measure are countably additive since they defined on the Borel σ-field of the metric space (Ω, d) and coherent conditional previsions are defined for Borel measurable random variables and so the Choquet integral coincides with the Lebesgue integral.*

5.2. Merging for Coherent Conditional Probabilities Defined by Hausdorff Measures

In this section we investigate if coherent conditional probabilities assigned by Hausdorff measures in different metric spaces, whose metrics are bi-Lipschitz, merge with each other. The following results hold since the restrictions, to the class of Borel-measurable random variables, of coherent conditional previsions defined in Theorem 4 are considered. These restrictions are linear.

Denote by $H_n(\omega)$ the atom of the σ-field $\mathbf{F}_n$ containing ω.

Definition 13 *Let (Ω, d) and (Ω, d_i) be two metric spaces where Ω is a set with positive and finite Hausdorff outer measures in its Hausdorff dimension s. Let $\{\boldsymbol{B}_n\}$ be a sequence of Borel finite or countable partitions of Ω and let $\boldsymbol{F}_n$ be the σ-field generated by $\boldsymbol{B}_1, \boldsymbol{B}_2, .., \boldsymbol{B}_n$. Let $H_n \in \boldsymbol{F}_n$ be a set with positive and finite s-Hausdorff outer measures in both metric spaces and denote by μ_{H_n} and $\nu_{H_n}^i$ the coherent conditional probabilities defined on $\mathscr{B}$ by Theorem 2 in the two metric spaces. Then μ_{H_n} merges to $\mu_{H_n}^i$ along $\{\boldsymbol{F}_n\}_{n=1}^{\infty}$ if for all $\varepsilon > 0$ there exists $N = N(\varepsilon, \omega)$ such that for all $n > N$ such that $H_n \in \boldsymbol{F}_n$ is a set with positive and finite s-Hausdorff outer measures in both metric spaces and all $\omega \in \Omega$*

$$\left|\mu(A|H_n(\omega)) - \mu^i(A|H_n(\omega))\right| < \varepsilon \text{ for all } A \in \mathscr{B}.$$

Suppose that $\mathbf{F}_n$ are σ-fields satisfying $\mathbf{F}_1 \subset \mathbf{F}_2 \subset \ldots \subset \mathbf{F}_n$. If the union $\bigcup_{n=1}^{\infty} \mathbf{F}_n$ generates the σ-field $\mathbf{F}_\infty$, this is expressed by $\mathbf{F}_n \uparrow \mathbf{F}_\infty$

Theorem 7 *Let Ω be a set with positive and finite Hausdorff outer measure in its Hausdorff dimension and let Z is a Borel-measurable random variable belonging to $L^*(\Omega)$. If $\mathbf{F}_n \uparrow \mathbf{F}_\infty$, then $P(Z|\mathbf{F}_n) \to P(Z|\mathbf{F}_\infty)$ with probability 1 with respect to μ_Ω.*

Proof Since Z is a Borel-measurable random variable in $L^*(\Omega)$ then $P(Z|\mathbf{F}_n)$ by Definition 11 is a linear coherent conditional prevision and by Example 6 the random variables $X_n = P(Z|\mathbf{F}_n)$ form a martingale, according to Definition 11. Since $P(|X_n|) \leq P(|Z|)$ then (by Theorem 35.4 [1]) the X_n converge to an integrable X. We have to identify X with $P(Z|\mathbf{F}_\infty)$. Let H be an atom of the σ-field F_∞ with positive and finite Hausdorff measure in its Hausdorff dimension s equal to the Hausdorff dimension of Ω. By the uniform integrability it is possible to integrate to the limit so that $\int_H X d\mu_\Omega = lim_{n\to+\infty} \int_\omega X_n d\mu_\Omega$; since the atoms H of the σ-fields F_n are Borel-measurable the coherent conditional prevision satisfy the disintegration property and we obtain

$$lim_{n\to+\infty} \int_H X_n d\mu_\Omega = \int_H P(Z|\mathbf{F}_n) d\mu_\Omega = \int_H Z d\mu_\Omega$$

Therefore $\int_H X d\mu_\Omega = \int_H Z d\mu_\Omega$ for all atoms H of $\mathbf{F}_\infty$ with positive Hausdorff measure μ_Ω. ■

The previous theorem is applied in the proof of Theorem 2 of the paper of Blackwell and Dubins [3] to assure that conditional probabilities, defined by probability measures which are absolutely continuous, merge when the cardinality of the σ-field of the conditioning events increases.

Theorem 8 *Let (Ω, d) be a metric space where Ω is a set with positive and finite Hausdorff outer measure in its Hausdorff dimension s and let d_i be a metric bi-Lipschitz equivalent to d. Let $\mathscr{B}$ the Borel σ-field of the metric spaces (Ω, d) and (Ω, d_i). Let $H_n \in \mathbf{F}_n$ with Hausdorff dimension equal to s. Denote by h^s and by h_i^s respectively the s-dimensional Hausdorff measure with respect to the metric d and d_i. If H has positive and finite s-dimensional Hausdorff outer measure in the two metric spaces then define on $\mathscr{B}$ coherent countably additive probability measures by*

$$\mu_H(A) = \frac{h^s(A)}{h^s(H)} \text{ and } \mu_H^i(A) = \frac{h_i^s(A)}{h_i^s(H)}$$

Then μ_H^i merges to μ_H.

Proof Since H has positive and finite Hausdorff outer measure in its Hausdorff dimension s equal to the Hausdorff dimension of Ω, the conditional probabilities μ_H^i and μ_H are defined, on the Borel σ-field $\mathscr{B}$ of the two metric spaces, by the s-dimensional Hausdorff measures which are countably additive and mutually absolutely continuous by Theorem 6. Then by Theorem 7 of this paper and Theorem 2 of [3] the distance between μ_H^i and μ_H goes to 0 when H belongs to $\mathbf{F}_\infty$, that is μ_H^i merges to μ_H. ■

Remark 14 *We can observe that if H has Hausdorff measure equal to zero or infinity, for instance if H is a countable set, then a coherent conditional probability is defined in Theorem 2 by a 0-1 valued finitely additive , but not countably additive, probability and in this case Theorem 8 does not hold; in fact there exist 0-1 valued conditional probabilities which are mutually absolutely continuous but they do not merge. An example is given in Section 6 of [3].*

For each H atom of $\mathbf{F}_\infty$ with positive and finite Hausdorff outer measure in its Hausdorff dimension s let μ_H be the probability measure defined by Theorem 2 in (Ω, d) and let $\mathbf{K}_H = [\mu_H^1, \mu_H^2, \ldots]$ be the credal set of all coherent countably additive probabilities μ_i, defined on $\mathscr{B}$ such that each μ_H^i is defined by a distance which is bi-Lipschitz with respect to d. So each μ_B^i is absolutely continuous with respect to μ_B. Then the credal set $\mathbf{K}_H$ represent the class of the opinions of the individuals, which agree when the information increases.

6. Conclusions

A natural interpretation of conditional probability is to represent subject's belief or subject's opinion about an event, given information represented by a sigma-field or a partition. The result proven in [3] can be interpreted to imply that if the opinion of two individuals, summarized by two conditional probabilities, agree on events, which have positive probability with respect to the first conditional probability if and only they have positive probability with respect to the second conditional probability, then after a finite number of observations, they will become and remain close to each other. If conditional probabilities are defined in a metric space, as proposed in the present paper, different individuals can define conditional probability in different metric spaces. The proposed results show that if different opinions are represented by coherent conditional probabilities defined on the Borel σ-field by Hausdorff measure in different metric spaces, with metrics bi-Lipschtz equivalent, then the distance between these conditional probabilities goes to zero when the cardinality of the σ-field of the conditioning events goes to infinity. It is an example of "'merging of opinions with increasing information"'. The result does not hold if the conditioning event has probability equal to zero or infinity because in this case conditional probability is defined by a 0-1 valued finitely additive, but not countably additive probability and there are examples of 0-1 valued probabilities, which are mutually absolutely continuous but they do not merge. This result confirms the idea, on which is based the model of conditional probability proposed in Theorem 2, that an individual really updates his opinion when the conditioning event is an unexpected, that is an event with zero probability. The future aim of this research is to investigate if a similar result holds for coherent conditional previsions defined on the class of all Choquet integrable random variables.

References

[1] Billingsley, P. *Probability and measure*. Wiley, 1986

[2] Billingsley, P. *Hausdorff dimension in Probability Theory*. paper presented to the Institute of Mathematical Statistics , August, 1958

[3] Blackwell, D., Dubins, E: Merging of opinions with increasing information. *The Annals of Mathematical Statistics*, 882-886, 1962

[4] Choquet, G. *Theory of capacity*. Ann. Inst. Fourier, 5, 131-295, 1953

[5] de Finetti, B. *Probability Induction and Statistics*. Wiley, New York, 1970

[6] de Finetti, B. *Theory of Probability*. Wiley, London, 1974

[7] Denneberg, D., *Non-additive measure and integral*. Kluwer Academic Publishers, 1994

[8] Doria, S. Probabilistic independence with respect to upper and lower conditional probabilities assigned by Hausdorff outer and inner measures. *International Journal of Approximate Reasoning*, 46, 617-635, 2007

[9] Doria, S. Characterization of a coherent upper conditional prevision as the Choquet integral with respect to its associated Hausdorff outer measure. *Annals of Operations Research*, 33-48, 2012

[10] Doria, S. Symmetric coherent upper conditional prevision by the Choquet integral with respect to Hausdorff outer measure. *Annals of Operations Research*, 377-396, 2014

[11] Doria, S. On the disintegration property of a coherent upper conditional prevision by the Choquet integral with respect to its associated Hausdorff outer measure. *Annals of Operations Research*, Volume 216, Issue 2, 253-269, 2017

[12] Doria, S., Dutta, B., Mesiar, R. Integral representation of coherent upper conditional prevision with respect to its associated Hausdorff outer measure: a comparison among the Choquet integral, the pan-integral and the concave integral. *International Journal of General Systems*, Volume 216, Issue 2, 569-592, 2018

[13] Doria, S., Coherent Upper Conditional Previsions Defined by Hausdorff Outer Measures for Unbounded Random Variables, ISIPTA 2019, Proceedings of Machine Learning Research 103:159-166, (2019)

[14] Doria, S. Disintegration property for coherent upper conditional previsions with respect to Hausdorff Outer Measures for unbounded random variables, published on line January 2021, *International Journal of General Systems*, 2021

[15] Drutu, C., Kapovich, M. Geometric Group Theory, Colloquium Publications, AMS, 2018

[16] Dubins, L.E. Finitely additive conditional probabilities, conglomerability and disintegrations. *The Annals of Probability*, Vol. 3, 89-99, 1975

[17] Falconer, K.J. *The geometry of fractal sets*. Cambridge University Press, 1986

[18] Lehrer, E., Smorodinsky, R. Merging and Learning. *Statistics, Probability and Game Theory* IMS Lecture Notes - Monograph Series, Voleme 30, 147-168, 1996

[19] Levi, I. *The Enterprise of Knowledge*. MIT Press, Cambridge, Massachusetts, 1980

[20] Regazzini, E. *Finitely additive conditional probabilities*. Rend. Sem. Mat. Fis., Milano, 5 5,69-89, 1985

[21] Regazzini, E. De Finetti's coherence and statistical inference. *The Annals of Statistics*, Vol. 15, No. 2, 845-864, 1987

[22] Rogers, C. A *Hausdorff measures*. Cambridge University Press, 1970

[23] Walley, P. *Coherent lower (and upper) probabilities*, Statistics Research Report, University of Warwick, 1981

[24] Walley, P. *Statistical Reasoning with Imprecise Probabilities*. Chapman and Hall, London, 1991