Contrasting Two Laws of Large Numbers from Possibility Theory and Imprecise Probability*

Pedro Terán[1] and Elisa Pis Vigil[2]

[1] Departamento de Estadística e I.O. y D.M.
Universidad de Oviedo (Spain)
Email: teranpedro@uniovi.es

[2] Email: epis1@alumno.uned.es

Abstract. The law of large numbers for coherent lower previsions (specifically, Choquet integrals against belief measures) can be applied to possibility measures, yielding that sample averages are asymptotically confined in a compact interval. This interval differs from the one appearing in the law of large numbers from possibility theory. In order to understand this phenomenon, we undertake an in-depth study of the compatibility of the assumptions in those results. It turns out that, although there is no incompatibility between their conclusions, their assumptions can only be simultaneously satisfied if the possibility distributions of the variables are 0-1 valued.

1 The problem

This contribution is part of a systematic analysis of the relationships between the laws of large numbers in different uncertainty frameworks (plausibility/belief measures, upper/lower probabilities, upper/lower previsions, sublinear expectations) in the particular case that they are applied to possibility measures. The main part of that analysis is [13].
In [11, Theorem 2.6], the first author obtained the following law of large numbers for possibilistic variables.

Theorem 1. *Let X be a bounded variable in a possibility space $(\Omega, \mathcal{A}, \Pi)$ such that the possibility distribution π_X of X is upper semicontinuous. Let $\{X_n\}_n$ be a sequence of variables such that*
(i) X_n are product related,
(ii) X_n are identically distributed as X.
Then, for any fixed $\varepsilon > 0$,

$$N\left(\mathcal{M}[X] - \varepsilon < n^{-1}\sum_{i=1}^{n} X_i < \mathbb{M}[X] + \varepsilon\right) \to 1.$$

*The first author's research in this paper was partially funded by Asturias's *Consejería de Economía y Empleo* (FC-15-GRUPIN14-101) and by Spain's *Ministerio de Economía y Competitividad* (MTM2015–63971–P).

Here $\mathcal{A}$ is a σ-algebra, $\Pi : \mathcal{A} \to [0,1]$ a possibility measure with N its dual necessity measure, π_X is given by $\pi_X(x) = \Pi(X = x)$, and $\mathcal{M}[X], \mathbb{M}[X]$ are the infimum and supremum, respectively, of the 1-cut of π_X. The requirement that X_n are *product related* means

$$\Pi(X_1 = x_1, \ldots, X_n = x_n) = \Pi(X_1 = x_1) \ldots \Pi(X_n = x_n)$$

for all $n \in \mathbb{N}$ and $x_i \in \mathbb{R}$. Indeed, Theorem 1 in its original presentation also considers the more general situation that the product is generalized to a continuous Archimedian triangular norm.

This law of large numbers is in line with previous results in the literature of possibility measures [5,6,8,10]. However, it must be compared to the law of large numbers of De Cooman and Miranda [2, Theorem 2] developed in the context of coherent lower previsions as a generalization of the law of large numbers for belief measures [7] (see also [9]) in which a similar limit interval appears but involves Choquet integrals instead. For more information on Choquet integrals and lower previsions, the reader is referred to Denneberg and Walley's books, respectively [3,14].

Indeed, by observing that the Choquet integral against a necessity measure is a coherent lower prevision, and also rewriting their result in a way closer to Theorem 1 for ease of comparison, we obtain the following.

Theorem 2. (De Cooman and Miranda) *Let X be a bounded variable in a possibility space $(\Omega, \mathcal{P}(\Omega), \Pi)$. Let $\{X_n\}_n$ be a sequence of variables such that*

(i') $(X_1, \ldots, X_n)$ is forward factorizing for the Choquet integral E_N, for each $n \in \mathbb{N}$,

(ii') X_n are uniformly bounded and such that $E_N[X_n] = E_N[X]$ and $E_\Pi[X_n] = E_\Pi[X]$ for all $n \in \mathbb{N}$.

Then, for any fixed $\varepsilon > 0$,

$$N\left(E_N[X] - \varepsilon < n^{-1}\sum_{i=1}^{n} X_i < E_\Pi[X] + \varepsilon\right) \to 1.$$

We emphasize that this version is weaker than [2, Theorem 2] in several respects, but will be better suited to our purpose. In it, E_N and E_Π denote the Choquet integrals with respect to N and Π, respectively. Condition (ii') is obviously satisfied if X_n are identically distributed as X, i.e. if (ii) holds. The restriction to the σ-algebra of parts $\mathcal{P}(\Omega)$ is inessential. Therefore, the substantial difference in the assumptions is condition (i') of *forward factorization*, namely the property that

$$E_N[g(X_1, \ldots, X_{n-1})(h(X_n) - E_N[h(X_n)])] \geq 0 \tag{1}$$

for all $n \in \mathbb{N}$ and bounded functions $g : \mathbb{R}^{n-1} \to [0, \infty)$ and $h : \mathbb{R} \to \mathbb{R}$.

Condition (i') is rather different from condition (i), and any relationship between them is not obviously visible. In this communication, we aim at clarifying their relationships or lack thereof, in view of the fact that different conclusions appear in Theorems 1 and 2. Both results claim that the averages $n^{-1}\sum_{i=1}^{n} X_i$ tend to be asymptotically confined inside a compact interval, but each yields a different interval: $[\mathcal{M}[X], \mathbb{M}[X]]$ in Theorem 1 and $[E_N[X], E_\Pi[X]]$ in Theorem 2.

There are three important remarks to be made. Firstly, both intervals have a special significance in fuzzy and possibility theory, making their study particularly relevant. Indeed, if the possibility distribution π_X is a fuzzy interval then $[\mathcal{M}[X], \mathbb{M}[X]]$ is its *core* and $[E_N[X], E_\Pi[X]]$ is its *mean value* in the sense of Dubois and Prade [4] (as follows immediately from the fact that $E_N[X]$ and $E_\Pi[X]$ are the infimum and the supremum of all expectations of X against probability measures dominated by Π, see [1, Lemma A.2] or [12, Proposition 3.5]).
Secondly, there is no incompatibility between both conclusions, since it is always the case that $[\mathcal{M}[X], \mathbb{M}[X]] \subset [E_N[X], E_\Pi[X]]$. Therefore the task of contrasting (i) and (i') is not a trivial one.
And thirdly, it may happen that $[E_N[X], E_\Pi[X]]$ is significantly larger than $[\mathcal{M}[X], \mathbb{M}[X]]$, which is reduced to a point if there is a unique point $x \in \mathbb{R}$ such that $\Pi(X = x) = 1$.
We will proceed by analyzing a specific type of function depending on two events, which eventually leads to 625 systems of equations and inequations, at least one of which must be satisfied if (i,i') hold simultaneously. Patient work reduces those systems to 14, which are finally shown to have solutions only under restrictive conditions, yielding the result stated in the abstract (see Corollary 5 below).

2 Forward factorization and product relatedness

In this section we will prove that conditions (i) and (i') are compatible only under very special circumstances. To that end it is enough to consider the situation of a couple of variables X, Y instead of a whole sequence.
Our first result shows that certain functions of X and Y must have Choquet integrals of opposite signs under those conditions. Below, I_A and I_B denote the indicator functions of events A, B. We also use the notation $\vee$ for the maximum.

Proposition 3. *Let X, Y be bounded variables in a possibility space $(\Omega, \mathcal{P}(\Omega), \Pi)$. Then, for any $A, B \subset \Omega$,*
(a) If (X, Y) is forward factorizing for the Choquet integral E_N, then

$$E_N[I_A(X)(I_B(Y) - N(Y \in B))] \geq 0.$$

(b) If X and Y are product related, then

$$E_N[I_A(X)(I_B(Y) - N(Y \in B))] \leq 0.$$

Proof. Part (a) follows directly from (1), taking $g = I_A$ and $h = I_B$ and observing

$$E_N[I_B(Y)] = E_N[I_{\{Y \in B\}}] = N(Y \in B).$$

As regards part (b), set

$$\kappa = E_N[I_A(X)(I_B(Y) - N(Y \in B))]$$

and the variable

$$Z = \begin{cases} 1, & X \in A, Y \in B \\ N(Y \in B), & X \notin A \\ 0, & X \in A, Y \notin B. \end{cases}$$

Now we work towards expressing κ in terms of possibilities:

$$\begin{gathered} \kappa = E_N[I_A(X)I_B(Y) + N(Y \in B)I_{A^c}(X) - N(Y \in B)] \\ = E_N[Z] - N(Y \in B) \\ = N(Y \in B)N(\{X \in A, Y \in B\} \cup \{X \notin A\}) \\ +(1 - N(Y \in B))N(X \in A, Y \in B) - N(Y \in B) \\ = N(Y \in B)N(\{X \in A, Y \in B^c\}^c) \\ +(1 - N(Y \in B))N(X \in A, Y \in B) - N(Y \in B) \\ = -N(Y \in B)(1 - N(\{X \in A, Y \in B^c\}^c)) \\ +(1 - N(Y \in B))N(X \in A, Y \in B) \\ = -(1 - \Pi(Y \in B^c))\Pi(X \in A, Y \in B^c) \\ +\Pi(Y \in B^c)(1 - \Pi(\{X \in A, Y \in B\}^c)). \end{gathered}$$

Let a, b, c, d be the possibilities of the events involved as summarized in the following table:

Π	$Y \in B$	$Y \in B^c$	
$X \in A$	a	b	$a \vee b$
$X \in A^c$	c	d	$c \vee d$
	$a \vee c$	$b \vee d$	

With that notation,

$$\kappa = -(1 - (b \vee d)) \cdot b + (b \vee d)(1 - (b \vee c \vee d)).$$

Since X and Y are product related,

$$b = \Pi(X \in A, Y \in B^c) = \sup_{x \in A, y \in B^c} \Pi(X = x, Y = y)$$

$$= \sup_{x \in A, y \in B^c} \pi_X(x)\pi_Y(y) = \Pi(X \in A)\Pi(Y \in B^c) = (a \vee b)(b \vee d),$$

whence

$$\begin{aligned} \kappa =& - (1 - (b \vee d)) \cdot (a \vee b) \cdot (b \vee d) + (b \vee d)(1 - (b \vee c \vee d)) \\ =&(b \vee d)[1 - (b \vee c \vee d) - (1 - (b \vee d))(a \vee b)]. \end{aligned} \tag{2}$$

Observing

$$1 = \Pi(\Omega) = a \vee b \vee c \vee d,$$

there are two possibilities:
CASE 1. If $a = 1$, then

$$\kappa = (b \vee d)[(b \vee d) - (b \vee c \vee d)] \leq 0.$$

CASE 2. If $b \vee c \vee d = 1$, then

$$\kappa = -(b \vee d)(1 - (b \vee d))(a \vee b) \leq 0.$$

Hence $\kappa \leq 0$ and the proof is complete. □

It is clear from Proposition 3 that forward factorization and product relatedness can occur simultaneously only if, in the notation of its proof, $\kappa = 0$. That has definite consequences for the possible distributions of X and Y, as our main result shows that at least one of them must be uniform, i.e. there is a set A such that $\pi_X = I_A$ or $\pi_Y = I_A$.

Theorem 4. *Let X and Y be bounded variables in a possibility space $(\Omega, \mathcal{P}(\Omega), \Pi)$. Conditions*
(I) (X, Y) is forward factorizing for the Choquet integral E_N
(II) X and Y are product related
cannot be simultaneously met unless at least one of the variables is uniform.

Proof. Let $A, B \subseteq \mathbb{R}$. By Proposition 3, if both (I) and (II) hold then it must be

$$E_N[I_A(X)(I_B(Y) - N(Y \in B)] = 0$$

and therefore for a, b, c, d in the notation of (2) we have

$$(b \vee d)[1 - (b \vee c \vee d) - (1 - (b \vee d))(a \vee b)] = 0, \tag{3}$$

whence

$$b \vee d = 0 \text{ (i.e. } b = d = 0)$$

or

$$1 - (b \vee c \vee d) = (1 - (b \vee d))(a \vee b).$$

Since $a \vee b \vee c \vee d = 1$, there are 3 possibilities for the latter equation:

. If $a = 1$, it becomes $1 - (b \vee c \vee d) = 1 - (b \vee d)$ i.e. $c \leq b \vee d$.
. If $b = 1$ or $d = 1$, then it always holds.
. If $c = 1$, it becomes $0 = (1 - (b \vee d))(a \vee b)$, whence $b \vee d = 1$ or $a = b = 0$.

The solution $c = 1$, $b \vee d = 1$ is already included in either case $b = 1$ or $d = 1$, whence (3) is rewritten as

$$\begin{aligned} b = d = 0 \quad &\textbf{or} \quad a = 1, c \leq b \vee d \quad \textbf{or} \quad b = 1 \\ \textbf{or} \quad &c = 1, a = b = 0 \quad \textbf{or} \quad d = 1. \end{aligned} \tag{4}$$

The same reasoning applies to the pairs of events (A^c, B), (A, B^c) and (A^c, B^c), from which the analogous conditions

$$\begin{aligned} d = b = 0 \quad &\textbf{or} \quad c = 1, a \leq d \vee b \quad \textbf{or} \quad d = 1 \\ \textbf{or} \quad &a = 1, c = d = 0 \quad \textbf{or} \quad b = 1; \end{aligned} \tag{5}$$

$$\begin{aligned} a = c = 0 \quad &\textbf{or} \quad b = 1, d \leq a \vee c \quad \textbf{or} \quad a = 1 \\ \textbf{or} \quad &d = 1, b = a = 0 \quad \textbf{or} \quad c = 1; \end{aligned} \tag{6}$$

$$\begin{aligned} c = a = 0 \quad &\textbf{or} \quad d = 1, b \leq c \vee a \quad \textbf{or} \quad c = 1 \\ \textbf{or} \quad &b = 1, d = c = 0 \quad \textbf{or} \quad a = 1. \end{aligned} \tag{7}$$

are derived. Thus conditions (4) through (7) might simultaneously be satisfied in $5^4 = 625$ different ways. Since a, b, c, d come from a possibility measure we have the restrictions

$$0 \leq a \leq 1, \quad 0 \leq b \leq 1, \quad 0 \leq c \leq 1, \quad 0 \leq d \leq 1, \quad a \vee b \vee c \vee d = 1 \tag{8}$$

as well as, by the product relatedness,

$$\begin{aligned} a &= (a \vee b)(a \vee c), \quad & b &= (a \vee b)(b \vee d), \\ c &= (a \vee c)(c \vee d), \quad & d &= (b \vee d)(c \vee d). \end{aligned} \tag{9}$$

The task of finding a, b, c, d is thus tantamount to solving these 625 systems of 13 to 17 equations and inequations in 4 unknowns.
We start by combining restrictions (4) through (7), adding one at a time and always using (8) to simplify the obtained conditions if convenient (thus, for example, $a = b = 0$ would replace $a \vee b = 0$).
Conditions (4) and (5) can be satisfied in 25 ways, of which the following 6 contain all others:

1. $b = d = 0$
2. $a = c = 1, b \vee d = 1$
3. $c = d = 0, a = 1$
4. $b = 1$
5. $a = b = 0, c = 1$
6. $d = 1$

Merging these with (6), conditions (4) through (6) can be satisfied in 30 ways, of which the following 14 contain all others:

1. $b = d = 0, a = 1$
2. $b = d = 0, c = 1$
3. $a = b = 1$
4. $a = c = 1, b \vee d = 1$
5. $c = d = 0, a = 1$
6. $a = c = 0, b = 1$
7. $b = 1, d \leq a \vee c$
8. $b = c = 1$
9. $a = b = 0, c = 1$
10. $a = c = 0, d = 1$
11. $b = d = 1, a \vee c = 1$
12. $a = d = 1$
13. $a = b = 0, d = 1$
14. $c = d = 1$

Merging these with (7), conditions (4) through (7) can be satisfied in 70 ways, of which the following 14 contain all others:

1. $b = d = 0, a = 1$
2. $b = d = 0, c = 1$
3. $a = b = 1$
4. $a = c = 1, b \vee d = 1$
5. $c = d = 0, a = 1$
6. $a = c = 0, b = 1$
7. $b = d = 1, a \vee c = 1$
8. $b = c = 1$
9. $c = d = 0, b = 1$
10. $a = b = 0, c = 1$
11. $a = c = 0, d = 1$
12. $a = d = 1$
13. $a = b = 0, d = 1$
14. $c = d = 1$

With a direct inspection of (9) in each of the fourteen cases, after eliminating redundancies and imposing (8) on the range of the variables we finally arrive at the following ten families of solutions:
1. $a = 0, b = 0, c = 1, d \in [0, 1]$.
2. $a = 0, b = 0, c \in [0, 1], d = 1$.
3. $a = 0, b = 1, c = 0, d \in [0, 1]$.
4. $a = 0, b \in [0, 1], c = 0, d = 1$.
5. $a = 1, b = 0, c \in [0, 1], d = 1$.
6. $a \in [0, 1], b = 0, c = 1, d = 0$.
7. $a = 1, b \in [0, 1], c = 0, d = 0$.
8. $a \in [0, 1], b = 1, c = 0, d = 0$.
9. $a = 1, b = 1, c = d \in [0, 1]$.
10. $a = b \in [0, 1], c = 1, d = 1$.

Reasoning by contradiction, assume now that X and Y were both not uniform. By definition, there would exist $x, y \in \mathbb{R}$ such that

$$p := \Pi(X = x) \in (0, 1), \qquad q := \Pi(Y = y) \in (0, 1).$$

Taking $A = \{x\}$ and $B = \{y\}$ above, using (8) and (9) we obtain the table

Π	$Y \in B$	$Y \in B^c$	
$X \in A$	pq	p	p
$X \in A^c$	q	1	1
	q	1	

representing a solution which nonetheless is not in any of the ten families above, a contradiction. Therefore, indeed X or Y must be uniform. □

As a consequence, for sequences of variables we obtain the following corollary.

Corollary 5. *Let $\{X_n\}_n$ be a sequence of identically distributed variables in a possibility space $(\Omega, \mathcal{P}(\Omega), \Pi)$. If both forward factorization and product relatedness, i.e. conditions (i) and (i'), hold, then the X_n must have uniform possibility distributions.*

3 Discussion

It is interesting that the conditions studied here are barely compatible, in the sense that a sequence of identically distributed variables satisfying both must have distributions giving possibility 0 or 1 to every event. Thus Theorems 1 and 2 are complementary as regards those assumptions.
The original laws of large numbers from which they have been simplified are also complementary in that both have content not covered by the other. The law from Possibility Theory covers the situation that the marginals of the X_n are linked by a triangular norm more general than the product, whereas the one from Imprecise Probability is of course applicable beyond possibility measures and also shows that the speed of the convergence is exponential.
It would be tempting to conclude that this 'almost incompatibility' is the explanation of the fact that both laws exhibit different limit intervals,

specially since $[\mathcal{M}[X], \mathbb{M}[X]] = [E_N[X], E_\Pi[X]]$ when both conditions apply (as follows from [13]).
However, it must be emphasized that such a conclusion is not warranted, i.e. it is unclear whether the larger interval in Theorem 2 is actually optimal under condition (i') when applied to possibility measures.

References

1. A. Castaldo, F. Maccheroni, M. Marinacci (2004). Random correspondences as bundles of random variables. Sankhya: Indian J. Statist. 66, 409-427.
2. G. de Cooman, E. Miranda (2008). Weak and strong laws of large numbers for coherent lower previsions. J. Statist. Plann. Inference 138, 2409–2432.
3. D. Denneberg (1994). Non-additive measure and integral. Kluwer, Dordrecht.
4. D. Dubois, H. Prade (1987). The mean value of a fuzzy number. Fuzzy Sets and Systems 24, 279–300.
5. R. Fullér (1992). A law of large numbers for fuzzy numbers. Fuzzy Sets and Systems 45, 299–303.
6. D. H. Hong, Y. M. Kim (1996). A law of large numbers for fuzzy numbers in a Banach space. Fuzzy Sets and Systems 77, 349–354.
7. F. Maccheroni, M. Marinacci (2005). A strong law of large numbers for capacities. Ann. Probab. 33, 1171–1178.
8. A. Puhalskii (2001). Large deviations and idempotent probability. Chapman & Hall/CRC, Boca Raton.
9. Y. Rébillé (2009). Laws of large numbers for continuous belief measures on compact spaces. Internat. J. Uncertain. Fuzziness Knowledge-Based Systems 17, 685–704.
10. P. Terán (2008). On convergence in necessity and its laws of large numbers. In: Soft methods for handling variability and imprecision, 289–296. Springer, Berlin.
11. P. Terán (2014). Law of large numbers for the possibilistic mean value. Fuzzy Sets and Systems 245, 116–124.
12. P. Terán (2014). Laws of large numbers without additivity. Trans. Amer. Math. Soc. 366, 5431–5451.
13. P. Terán. A unified approach to the laws of large numbers for possibility measures in the context of more general uncertainty theories. Submitted for publication.
14. P. Walley (1991). Statistical reasoning with imprecise probabilities. Chapman and Hall, London.