On Random Sets Independence and Strong Independence in Evidence Theory

Jiřina Vejnarová

Abstract. Belief and plausibility functions can be viewed as lower and upper probabilities possessing special properties. Therefore, (conditional) independence concepts from the framework of imprecise probabilities can also be applied to its sub-framework of evidence theory. In this paper we concentrate ourselves on random sets independence, which seems to be a natural concept in evidence theory, and strong independence, one of two principal concepts (together with epistemic independence) in the framework of credal sets. We show that application of strong independence to two bodies of evidence generally leads to a model which is beyond the framework of evidence theory. Nevertheless, if we add a condition on resulting focal elements, then strong independence reduces to random sets independence. Unfortunately, it is not valid no more for conditional independence.

1 Introduction

Imprecise probabilities is a general concept comprising different theories dealing with imprecise information. These theories can be partially ordered with respect to their generality and evidence theory belongs to the most specific ones. More precisely, belief and plausibility functions can be viewed as lower and upper probabilities, respectively, possessing special properties.

Independence belongs to the most important concepts within any theory dealing with uncertainty and therefore it has been studied in the evidential framework from the very beginning [11]. Because of reasons stated above, the application of independence concepts from imprecise probabilities to belief plausibility functions is, in principle, possible and their relationship to "natural" independence concepts in evidence theory is an interesting question, as already suggested in [5, 6, 8].

Jiřina Vejnarová
Institute of Information Theory and Automation of the AS CR, Pod Vodárenskou věží 4, Prague, Czech Republic
e-mail: vejnar@utia.cas.cz

T. Denœux & M.-H. Masson (Eds.): Belief Functions: Theory & Appl., AISC 164, pp. 247–254.
springerlink.com

In this paper we confine ourselves to random sets independence and strong independence and will not deal with epistemic irrelevance and independence, as they are based on conditional probabilities/beliefs and there does not exist a uniquely accepted conditioning rule [7] in the framework of evidence theory.

The paper is organized as follows. Section 2 is an overview of basic concepts from evidence theory and form credal sets and in Section 3 random sets independence and strong independence are introduced and their relationship in the framework of evidence theory is studied.

2 Basic Concepts

In this section we will briefly recall basic concepts from evidence theory [11] concerning sets and set functions and from the framework of credal sets [10].

2.1 Set Projections and Joins

For an index set $N = \{1,2,\dots,n\}$ let $\{X_i\}_{i\in N}$ be a system of variables, each X_i having its values in a finite set $\mathbf{X}_i$. In this paper we will deal with *multidimensional frame of discernment* $\mathbf{X}_N = \mathbf{X}_1 \times \mathbf{X}_2 \times \ldots \times \mathbf{X}_n$, and its *subframes* (for $K \subseteq N$)

$$\mathbf{X}_K = \times_{i\in K}\mathbf{X}_i.$$

When dealing with groups of variables on these subframes, X_K will denote a group of variables $\{X_i\}_{i\in K}$ throughout the paper.

A *projection* of $x = (x_1,x_2,\dots,x_n) \in \mathbf{X}_N$ into $\mathbf{X}_K$ will be denoted $x^{\downarrow K}$, i.e. for $K = \{i_1,i_2,\dots,i_k\}$

$$x^{\downarrow K} = (x_{i_1},x_{i_2},\dots,x_{i_k}) \in \mathbf{X}_K.$$

Analogously, for $M \subset K \subseteq N$ and $A \subset \mathbf{X}_K$, $A^{\downarrow M}$ will denote a *projection* of A into $\mathbf{X}_M$:

$$A^{\downarrow M} = \{y \in \mathbf{X}_M \mid \exists x \in A : y = x^{\downarrow M}\}.$$

In addition to the projection, in this text we will also need an opposite operation, which will be called a join. By a *join* [1] of two sets $A \subseteq \mathbf{X}_K$ and $B \subseteq \mathbf{X}_L$ $(K,L \subseteq N)$ we will understand a set

$$A \bowtie B = \{x \in \mathbf{X}_{K\cup L} : x^{\downarrow K} \in A \ \& \ x^{\downarrow L} \in B\}.$$

Let us note that for any $C \subseteq \mathbf{X}_{K\cup L}$ naturally $C \subseteq C^{\downarrow K} \bowtie C^{\downarrow L}$, but generally $C \neq C^{\downarrow K} \bowtie C^{\downarrow L}$, i.e., a join is, in a sense, a generalization of a rectangle — so called $X^{\downarrow K\cap L}$*-layered* rectangle [3].

2.2 Set Functions

In evidence theory [11] (or Dempster-Shafer theory) two dual measures are used to model the uncertainty: belief and plausibility measures. Both of them can be defined with the help of another set function called a *basic (probability* or *belief) assignment* m on $\mathbf{X}_N$, i.e., $m : \mathscr{P}(\mathbf{X}_N) \longrightarrow [0,1]$, where $\mathscr{P}(\mathbf{X}_N)$ is power set of $\mathbf{X}_N$ and $\sum_{A\subseteq \mathbf{X}_N} m(A) = 1$. Furthermore, we assume that $m(\emptyset) = 0$. A set $A \in \mathscr{P}(\mathbf{X}_N)$ is a *focal element* if $m(A) > 0$.

Belief and *plausibility measures* are defined for any $A \subseteq \mathbf{X}_N$ by the equalities

$$Bel(A) = \sum_{B\subseteq A} m(B), \qquad Pl(A) = \sum_{B\cap A\neq\emptyset} m(B),$$

respectively. It is well-known (and evident from these formulae) that for any $A \in \mathscr{P}(\mathbf{X}_N)$

$$Bel(A) \leq Pl(A), \qquad Pl(A) = 1 - Bel(A^C), \tag{1}$$

where A^C is the set complement of $A \in \mathscr{P}(\mathbf{X}_N)$.

Because of (1) belief and plausibility functions may be viewed as lower and upper probabilities, respectively. Furthermore, basic assignment can be computed from belief function via Möbius inversion:

$$m(A) = \sum_{B\subseteq A} (-1)^{|A\setminus B|} Bel(B), \tag{2}$$

i.e. any of these three functions is sufficient to define values of the remaining two.

For a basic assignment m on $\mathbf{X}_K$ and $M \subset K$, a *marginal basic assignment* of m on $\mathbf{X}_M$ is defined (for each $A \subseteq \mathbf{X}_M$):

$$m^{\downarrow M}(A) = \sum_{\substack{B\subseteq \mathbf{X}_K \\ B^{\downarrow M}=A}} m(B).$$

Analogously we will denote by $Bel^{\downarrow M}$ marginal belief measure on $\mathbf{X}_M$.

2.3 Credal Sets

A *credal set* $\mathscr{M}(X)$ about a variable X is defined as a closed convex set of probability measures about the values of this variable. In order to simplify the expression of operations with credal sets, it is often considered [10] that a credal set is the set of probability distributions associated to the probability measures in it. Under such consideration a credal set can be expressed as a *convex hull* of its extreme distributions

$$\mathscr{M}(X) = \mathrm{CH}\{\mathrm{ext}(\mathscr{M}(X))\}.$$

Any lower probability $\underline{P}$ can be associated with a credal set of probabilities dominating it:

$$\mathcal{M}(\underline{P}) = \mathrm{CH}\{P : P(A) \geq \underline{P}(A), A \subseteq \mathbf{X}\}.$$

As belief measure is a lower probability, this association can be done also for it, as suggested in both examples in the next section.

3 Independence Concepts

3.1 Random Sets Independence

Let us start this section by recalling the notion of random sets independence [4].

Definition 1. Let m be a basic assignment on $\mathbf{X}_N$ and $K, L \subset N$ be disjoint. We say that groups of variables X_K and X_L are *independent with respect to basic assignment* m if

$$m^{\downarrow K \cup L}(A) = m^{\downarrow K}(A^{\downarrow K}) \cdot m^{\downarrow L}(A^{\downarrow L}) \tag{3}$$

for all $A \subseteq \mathbf{X}_{K \cup L}$ for which $A = A^{\downarrow K} \times A^{\downarrow L}$, and $m(A) = 0$ otherwise.

Example 1. Consider two basic assignments m_X and m_Y on $\mathbf{X} = \{x, \bar{x}\}$ and $\mathbf{Y} = \{y, \bar{y}\}$, respectively, specified in Table 1 together with their beliefs and plausibilities. Under the assumption of random sets independence we get the joint basic assignment m, values of which are contained in the second column of Table 2. In third and fourth columns one can find beliefs and plausibilities of the corresponding sets, respectively. ◇

Table 1 Basic assignments m_X and m_Y.

$A \subseteq \mathbf{X}$	$m_X(A)$	$Bel_X(A)$	$Pl_X(A)$
$\{x\}$	0.3	0.3	0.8
$\{\bar{x}\}$	0.2	0.2	0.7
$\mathbf{X}$	0.5	1	1

$A \subseteq \mathbf{Y}$	$m_Y(A)$	$Bel_Y(A)$	$Pl_Y(A)$
$\{y\}$	0.6	0.6	0.9
$\{\bar{y}\}$	0.1	0.1	0.4
$\mathbf{Y}$	0.3	1	1

There exist numerous generalizations [3, 9, 12] of this notion to the conditional case. For the reasons presented e.g. in [9], we use the following one.

Definition 2. Let m be a basic assignment on $\mathbf{X}_N$ and $K, L, M \subset N$ be disjoint, $K \neq \emptyset \neq L$. We say that groups of variables X_K and X_L are *conditionally independent given X_M with respect to m* (and denote it by $K \perp\!\!\!\perp L | M\ [m]$), if the equality

$$m^{\downarrow K \cup L \cup M}(A) \cdot m^{\downarrow M}(A^{\downarrow M}) = m^{\downarrow K \cup M}(A^{\downarrow K \cup M}) \cdot m^{\downarrow L \cup M}(A^{\downarrow L \cup M}) \tag{4}$$

holds for any $A \subseteq \mathbf{X}_{K \cup L \cup M}$ such that $A = A^{\downarrow K \cup M} \bowtie A^{\downarrow L \cup M}$, and $m(A) = 0$ otherwise.

Table 2 Results of application of random sets independence (Col. 2–4) and strong independence (Col. 5–7).

$C \subseteq \mathbf{X} \times \mathbf{Y}$	$m_R(C)$	$Bel_R(C)$	$Bel_R(C)$	$\underline{P}_{XY}(C)$	$\overline{P}_{XY}(C)$	$\tilde{m}_{XY}(C)$
$\{xy\}$	0.18	0.18	0.72	0.18	0.72	0.18
$\{x\bar{y}\}$	0.03	0.03	0.32	0.03	0.32	0.03
$\{\bar{x}y\}$	0.12	0.12	0.63	0.12	0.63	0.12
$\{\bar{x}\bar{y}\}$	0.02	0.02	0.28	0.02	0.28	0.02
$\{x\} \times \mathbf{Y}$	0.09	0.3	0.8	0.3	0.8	0.09
$\{\bar{x}\} \times \mathbf{Y}$	0.06	0.2	0.7	0.2	0.7	0.06
$\mathbf{X} \times \{y\}$	0.3	0.6	0.9	0.6	0.9	0.3
$\mathbf{X} \times \{\bar{y}\}$	0.05	0.1	0.4	0.1	0.4	0.05
$\{xy, \bar{x}\bar{y}\}$	0	0.2	0.85	0.34	0.74	0.14
$\{x\bar{y}, \bar{x}y\}$	0	0.15	0.8	0.26	.66	0.11
$\mathbf{X} \times \mathbf{Y} \setminus \{\bar{x}\bar{y}\}$	0	0.72	0.98	0.72	0.98	−0.11
$\mathbf{X} \times \mathbf{Y} \setminus \{\bar{x}y\}$	0	0.37	0.88	0.37	0.88	−0.14
$\mathbf{X} \times \mathbf{Y} \setminus \{x\bar{y}\}$	0	0.68	0.97	0.68	0.97	−0.14
$\mathbf{X} \times \mathbf{Y} \setminus \{xy\}$	0	0.28	0.82	0.28	0.82	−0.11
$\mathbf{X} \times \mathbf{Y}$	0.15	1	1	1	1	0.4

3.2 *Strong Independence*

From numerous definitions of independence for credal sets [4] we have chosen strong independence, as it seems to be most proper for multidimensional models.

We say that X_K and X_L are *strongly independent* with respect to $\mathcal{M}(X_K X_L)$ iff (in terms of probability distributions)

$$\mathcal{M}(X_K X_L) = \mathrm{CH}\{P_1 \cdot P_2 : P_1 \in \mathcal{M}(X_K), P_2 \in \mathcal{M}(X_L)\}. \tag{5}$$

Again, there exist several generalizations of this notion to conditional independence, see e.g. [10], but as the following definition is suggested by the authors as the most appropriate for the marginal problem, it seems to be a suitable counterpart of random sets independence.

Given three variables X, Y and Z we say that X and Y are *independent on the distribution* given Z under global set $\mathcal{M}(X,Y,Z)$ iff

$$\mathcal{M}(X,Y,Z) = \{(p_1 \cdot p_2)/p_1^{\downarrow Z} : p_1 \in \mathcal{M}(X,Z), p_2 \in \mathcal{M}(X,Z), p_1^{\downarrow Z} = p_2^{\downarrow Z}\}.$$

From the term "strong independence" one could deduce that it should imply random sets independence. Nevertheless, it is not true, as can be seen from the following simple example.

Example 1. (Continued) From values contained in Table 1 we obtain credal sets about variables X and Y:

$$\mathcal{M}(X) = \mathrm{CH}\{(0.3,0.7),(0.8,0.2)\}, \quad \mathcal{M}(Y) = \mathrm{CH}\{(0.6,0.4),(0.9,0.1)\}.$$

Under the assumption of strong independence we get

$$\begin{aligned}\mathcal{M}(XY) = \mathrm{CH}\{&(0.18,0.12,0.42,0.28),(0.27,0.03,0.63,0.07),\\ &(0.48,0.32,0.12,0.08),(0.72,0.08,0.18,0.02)\}.\end{aligned}$$

Let us compute lower and upper probabilities of all nonempty subsets of $\mathbf{X} \times \mathbf{Y}$. Their values can be found in fifth and sixth columns of Table 2.

In the last column one can find hypothetical values of basic assignment corresponding the these lower and upper probabilities taken as beliefs and plausibilities computed via formula (2). From this column one can see that X and Y do not satisfy random set independence, as m_{XY} assigns positive values also to subsets which are not of the form $A = B \times C$. Furthermore, negative values are assigned to some sets, which violates the nonnegativity of basic assignment, i.e. we are beyond the limits of evidence theory. ◇

This result led us to the conclusion that strong independence cannot be applied in the framework of evidence theory. Nevertheless, under specific conditions it can be done as the following theorem[1] holds true.

Theorem 1. *Let X_K and X_L $(K \cap L \neq \emptyset)$ be two groups of variables with basic assignments $m^{\downarrow K}$ and $m^{\downarrow L}$, respectively. Let $Bel^{\downarrow K \cup L}$ and $\underline{P}^{\downarrow K \cup L}$ denote the joint belief function under random sets independence and joint lower probability under strong independence, respectively, and let A be a subset of $\mathbf{X}_K \times \mathbf{X}_L$ such that $A = A^{\downarrow K} \times A^{\downarrow L}$. Then*

$$Bel^{\downarrow K \cup L}(A) = \underline{P}^{\downarrow K \cup L}(A). \tag{6}$$

Proof. It is well-known[2] that for random sets independence the following equality holds true for any $A = A^{\downarrow K} \times A^{\downarrow L}$:

$$Bel^{\downarrow K \cup L}(A) = Bel^{\downarrow K}(A^{\downarrow K}) \cdot Bel^{\downarrow L}(A^{\downarrow L}).$$

Taking into account the fact that

$$Bel^{\downarrow K}(A^{\downarrow K}) = \underline{P}^{\downarrow K}(A^{\downarrow K}), \qquad Bel^{\downarrow L}(A^{\downarrow L}) = \underline{P}^{\downarrow L}(A^{\downarrow L}),$$

to get (6) it is enough to prove that for any $A \subseteq \mathbf{X}_K \times \mathbf{X}_L$ such that $A = A^{\downarrow K} \times A^{\downarrow L}$ the equality

---

[1] Let us note that the content of this theorem was already mentioned (without a proof) in [4].

[2] In [2] equality (7) together with an analogous one for plausibilities is used as a definition of *evidential independence* and Definition 1 is presented as an equivalent characterization.

$$\underline{P}^{\downarrow K\cup L}(A) = \underline{P}^{\downarrow K}(A^{\downarrow K}) \cdot \underline{P}^{\downarrow L}(A^{\downarrow L})$$

is satisfied.

Generally,

$$\underline{P}^{\downarrow K\cup L}(A) = \min_{P\in\mathcal{M}}\{\sum_{x\in A} P(x)\},$$

but as $\mathcal{M} = \{P_1 \cdot P_2 : P_1 \in \mathcal{M}_K, P_2 \in \mathcal{M}_L\}$, and $A = A^{\downarrow K} \times A^{\downarrow L}$ then

$$\begin{aligned}\underline{P}^{\downarrow K\cup L}(A) &= \min_{P\in\mathcal{M}}\{\sum_{x\in A} P(x)\} = \min_{P=P_1\cdot P_2, P_1\in\mathcal{M}_K, P_2\in\mathcal{M}_L}\{\sum_{x_K\in A^{\downarrow K}} P(x_K)\cdot \sum_{x_L\in A^{\downarrow L}} P(x_L)\}\\ &= \min_{P_1\in\mathcal{M}_K}\{\sum_{x_K\in A^{\downarrow K}} P(x_K)\}\cdot \min_{P_2\in\mathcal{M}_L}\{\sum_{x_L\in A^{\downarrow L}} P(x_L)\} = \underline{P}^{\downarrow K}(A^{\downarrow K}) \cdot \underline{P}^{\downarrow L}(A^{\downarrow L}),\end{aligned}$$

as requested (where the last but one equality holds thanks to the fact that we deal with non-negative numbers.) □

Unfortunately, for conditional independence an analogous result does not hold.

Example 2. Let X, Y and Z be three binary variables with values in $\mathbf{X} = \{x, \bar{x}\}$, $\mathbf{Y} = \{y, \bar{y}\}$ and $\mathbf{Z} = \{z, \bar{z}\}$, respectively, and m_{XZ} and m_{YZ} two basic assignments, both of them having only two focal elements:

$$m_{XZ}(\{(x,\bar{z}),(\bar{x},\bar{z})\}) = 0.5, \qquad m_{XZ}(\{(u,\bar{z}),(\bar{x},z)\}) = 0.5,$$
$$m_{YZ}(\{(y,\bar{z}),(\bar{y},\bar{z})\}) = 0.5, \qquad m_{YZ}(\{(y,\bar{z}),(\bar{y},z)\}) = 0.5.$$

Applying Definition 2 one can easily obtain the following joint assignment:

$$m(\mathbf{X}\times\mathbf{Y}\times\{\bar{z}\}) = 0.5, \qquad m(\{(x,y,\bar{z}),(\bar{x},\bar{y},z)\}) = 0.5.$$

From the values of the basic assignments m_{XZ} we will obtain credal set

$$\mathcal{M}(XZ) = \mathrm{CH}\{(0,1,0,0),(0,.5,0,.5),(0,.5,0.5,0),(0,0,.5,.5)\},$$

and credal set $\mathcal{M}(YZ)$ is identical. We can see, that the first two probability distributions are projective and the remaining two as well. Therefore under the assumption of strong conditional independence we will get the following joint credal set

$$\begin{aligned}\mathcal{M}(XYZ) = \mathrm{CH}\{&(0,1,0,0,0,0,0,0),(0,.5,0,.5,0,0,0,0),(0,.5,0,0,0,.5,0,0),\\ &(0,.25,0,.25,0,.25,0,.25),(0,.5,0,0,0,0,.5,0),\\ &(0,.0,0,.5,0,0,.5,0),(0,0,0,0,0,.5,.5,0),(0,0,0,0,0,0,.5,.5)\}.\end{aligned}$$

From $\mathcal{M}(XYZ)$ we can easily get values of lower and upper probabilities of all singletons as well as values of bigger subsets. For example, for the above mentioned focal elements we have

$$\underline{P}(\mathbf{X}\times\mathbf{Y}\times\{v\}) = 0.5, \quad \underline{P}(\{(u,u,v),(v,v,u)\}) = 0.25,$$

i.e., the latter is different from that obtained under random sets independence. ◇

4 Conclusions

The aim of this paper was to clarify the relationship between random sets independence and strong independence in the framework of evidence theory. Although evidence theory can be viewed as a special case of imprecise probabilities, application of strong independence may lead to models which are beyond the framework of evidence theory. If we confine ourselves to rectangles, values of joint belief function (under random sets independence) and those of joint lower probability (under strong independence) coincide. Nevertheless, an analogous result does not hold in the conditional case.

The problem of (epistemic) irrelevance was not discussed here, as the properties of irrelevance are dependent on the conditioning rule in question, and the problem of conditioning in evidence theory has not yet been satisfactorily solved.

Acknowledgements. The support of Grant GAČR P402/11/0378 is gratefully acknowledged.

References

1. Beeri, C., Fagin, R., Maier, D., Yannakakis, M.: On the desirability of acyclic database schemes. J. of the Association for Computing Machinery 30, 479–513 (1983)
2. Ben Yaghlane, B., Smets, P., Mellouli, K.: Belief functions independence: I. the marginal case. Int. J. Approx. Reasoning 29, 47–70 (2002)
3. Ben Yaghlane, B., Smets, P., Mellouli, K.: Belief functions independence: II. the conditional case. Int. J. Approx. Reasoning 31, 31–75 (2002)
4. Couso, I., Moral, S., Walley, P.: Examples of independence for imprecise probabilities. In: de Cooman, G., Cozman, F.G., Moral, S., Walley, P. (eds.) Proceedings of ISIPTA 1999, Ghent, June 29-July 2, pp. 121–130 (1999)
5. Couso, I., Moral, S.: Independence Concepts in Evidence Theory. Int. J. Approx. Reasoning 51, 748–758 (2010)
6. Destercke, S.: Independence concepts in evidence theory: some results about epistemic irrelevance and imprecise belief functions. In: Proceedings of BELIEF 2010 (2010)
7. Fagin, R., Halpern, J.Y.: A new approach to updating beliefs. In: Bonissone, et al. (eds.) Uncertainty in Artifiacial Intelligence, vol. VI, pp. 347–374. Elsevier (1991)
8. Fetz, T.: Sets of joint probability measures generated by weighted marginal focal sets. In: de Cooman, G., Cozman, F.G., Fine, T., Moral, S. (eds.) Proceedings of ISIPTA 2001, Ithaca, June 26-29, pp. 201–210 (2001)
9. Jiroušek, R., Vejnarová, J.: Compositional models and conditional independence in Evidence Theory. Int. J. Approx. Reasoning 52, 316–334 (2011)
10. Moral, S., Cano, A.: Strong conditional independence for credal sets. Ann. of Math. and Artif. Intell. 35, 295–321 (2002)
11. Shafer, G.: A Mathematical Theory of Evidence. Princeton University Press, Princeton (1976)
12. Shenoy, P.P.: Conditional independence in valuation-based systems. Int. J. Approx. Reasoning 10, 203–234 (1994)