Knowledge-Based Systems 52 (2013) 107–113

Contents lists available at ScienceDirect

Knowledge-Based Systems

journal homepage: www.elsevier.com/locate/knosys

On the interrelation between Dempster–Shafer Belief Structures and their Belief Cumulative Distribution Functions

Fabrizio Durante [a], Juan Fernández Sánchez [b], Wolfgang Trutschnig [c,*]

[a] *School of Economics and Management, Free University of Bozen-Bolzano, Bolzano, Italy*
[b] *Grupo de Investigación de Análisis Matemático, Universidad de Almería, La Cañada de San Urbano, Almería, Spain*
[c] *Research Unit for Intelligent Data Analysis, European Centre for Soft Computing, Edificio Científico Tecnológico, Calle Gonzalo Gutiérrez Quirós, 33600 Mieres, Asturias, Spain*

ARTICLE INFO

Article history:
Received 22 March 2013
Received in revised form 23 May 2013
Accepted 17 July 2013
Available online 26 July 2013

Keywords:
Dempster–Shafer belief structure
Belief cumulative distribution function
Random set
Interval data
Probability measure

ABSTRACT

We consider Dempster–Shafer belief structures (DSBSs) $\mathfrak{m}$ having finitely many non-empty compact intervals as focal elements and prove several results describing the interrelation between DSBSs and their so-called Belief Cumulative Distribution Functions (BCDFs) $x \mapsto \mathbf{F}(x) = (\underline{F}(x), \overline{F}(x))$ induced by the corresponding belief and plausibility measures. In particular, we answer a question by Ronald R. Yager on the injectivity of the assignment $\mathfrak{m} \mapsto \mathbf{F}$.

1. Introduction

Suppose that $(\Omega, \mathcal{A}, \mathcal{P})$ is a probability space and that $X : \Omega \to \mathbb{R}$ is an unobservable random variable. Instead of $X(\omega)$, however, it is possible to observe a compact interval $\mathbf{X}(\omega) = [\underline{X}(\omega), \overline{X}(\omega)]$ containing the true value $X(\omega)$ for every $\omega \in \Omega$ (think, for instance, of a measurement device rounding to a certain digit or general interval-censored data, see [7,9]). Assuming that $\mathbf{X}$ is a random compact interval (i.e. that both $\underline{X}$ and $\overline{X}$ are random variables) according to Dempster [1] $\mathbf{X}$ induces the so-called *lower and upper probability* $\underline{\pi}$ and $\bar{\pi}$ (also referred to as *belief and plausibility measure*) respectively via

$$\begin{aligned} \underline{\pi}(B) &= \mathcal{P}(\{\omega \in \Omega : \mathbf{X}(\omega) \subseteq B\}) \\ \bar{\pi}(B) &= \mathcal{P}(\{\omega \in \Omega : \mathbf{X}(\omega) \cap B \neq \emptyset\}) \end{aligned} \tag{1}$$

for every Borel set $B \in \mathcal{B}(\mathbb{R})$. Obviously $\underline{\pi}(B)$ and $\bar{\pi}(B)$ can be used as lower and upper bound for $\mathcal{P}(\{\omega \in \Omega : X(\omega) \in B\})$. For important properties of $\underline{\pi}$ and $\bar{\pi}$ as set functions (also in the general setting of random closed sets) see, for instance, Matheron [3] and Molchanov [4]. In case the range of $\mathbf{X}$ only consists of (pairwise different) intervals $[a_1, b_1], \ldots, [a_n, b_n]$ the random interval $\mathbf{X}$ is fully characterized by the quantities $m_i = \mathcal{P}(\{\omega \in \Omega : \mathbf{X}(\omega) = [a_i, b_i]\})$ for $i \in \{1, \ldots, n\}$. Defining $\mathfrak{m} : 2^{\mathbb{R}} \to [0,1]$ by $\mathfrak{m}(A) = 0$ for $A \notin \{[a_1, b_1], \ldots, [a_n, b_n]\}$ and $\mathfrak{m}(A) = m_i$ for $A = [a_i, b_i]$ induces a Dempster–Shafer belief structure (DSBS, see [8]) which we will denote by $(\mathbb{R}, \mathfrak{m}, ([a_i, b_i])_{i=1}^n)$. Yager [8] studied the interrelation between DSBSs and their so-called Belief Cumulative Distribution Function (BCDF) $\mathbf{F}_{\mathfrak{m}}$, defined by (notation as before)

$$\mathbf{F}_{\mathfrak{m}}(x) = [\underline{\pi}((-\infty, x]), \bar{\pi}((-\infty, x])]$$

for every $x \in \mathbb{R}$, and asked the question under which (necessary and sufficient) conditions two DSBSs induce the same Belief Cumulative Distribution Function. In the current note we give an answer to this question and prove several related results. In fact we present two main theorems: Firstly, for each pair of non-decreasing right-continuous step functions F_1, F_2 with $F_1 \leqslant F_2$ we can find a (not necessarily unique) DSBS $(\mathbb{R}, \mathfrak{m}, ([a_i, b_i])_{i=1}^n)$ such that $\mathbf{F}_{\mathfrak{m}} = [F_1, F_2]$. And secondly we show that $\mathfrak{m}$ is unique if and only if its focal elements $([a_i, b_i])_{i=1}^n$ fulfill a simple intersection-condition.

As main contribution the paper provides a concise description of the interrelation between DSBSs and BCDFs. Such a description is not only interesting from the theoretical mathematical point of view since it generalizes the one-to-one relationship between discrete probability measures and discrete distribution functions, which is, for instance, implicitly utilized in statistics whenever the empirical distribution function F_n of a sample $x_1, \ldots, x_n$ instead of the sample itself is considered. In fact it is also essential from the applied perspective since whenever working with DSBSs and BCDFs it has to be understood completely under which conditions,

* Corresponding author. Tel.: +34 985456545; fax: +34 985456699.
E-mail addresses: fabrizio.durante@unibz.it (F. Durante), juanfernandez@ual.es (J. Fernández Sánchez), wolfgang@trutschnig.net (W. Trutschnig).

0950-7051/$ - see front matter © 2013 Elsevier B.V. All rights reserved.
http://dx.doi.org/10.1016/j.knosys.2013.07.012

for instance, considering the BCDF instead of the DSBS does not imply a loss of (relevant) information.

2. Notation and preliminaries

For every set E the cardinality of E will be denoted by $\#E$. In the sequel $\mathcal{K}(\mathbb{R})$ will denote the family of all non-empty compact (i.e. closed and bounded) intervals in $\mathbb{R}, \mathcal{K}([0,1])$ the family of all non-empty compact subintervals of [0, 1]. $\mathcal{B}(\mathbb{R})$ denotes the Borel σ-field in $\mathbb{R}, \delta_x$ the Dirac measure in $x, \mathcal{P}$ the family of all probability measures on $\mathcal{B}(\mathbb{R})$ and $\mathcal{P}_d$ the family of all elements $\mu \in \mathcal{P}$ for which there exists a finite set $\Lambda \subseteq \mathbb{R}$ with $\mu(\Lambda) = 1$. $\mathcal{F}$ will denote the family of all distribution functions on $\mathbb{R}$. It is well known that there is a one-to-one correspondence between $\mathcal{F}$ and $\mathcal{P}$: In fact, given $F \in \mathcal{F}$ and defining $\mu_F : \mathcal{B}(\mathbb{R}) \to [0,1]$ by setting

$$\mu_F((-\infty, x]) = F(x)$$

for every $x \in \mathbb{R}$ and extending μ_F in the standard way (see for instance [2,5]) to full $\mathcal{B}(\mathbb{R})$ yields $\mu_F \in \mathcal{P}$. Moreover, given $\mu \in \mathcal{P}$ and defining $F_\mu(x) = \mu((-\infty, x])$ for every $x \in \mathbb{R}$, it is easily verified that $F_\mu \in \mathcal{F}$. Altogether the maps $\Phi : \mathcal{P} \to \mathcal{F}$ and $\Psi : \mathcal{F} \to \mathcal{P}$, given by

$$\Phi(\mu) = F_\mu \quad \text{and} \quad \Psi(F) = \mu_F, \tag{2}$$

fulfill $\Phi(\mu_F) = F$ for every $F \in \mathcal{F}$ as well as $\Psi(F_\mu) = F$ for every $\mu \in \mathcal{P}$ (again see [2,5] for details), implying bijectivity of Φ, Ψ and $\Psi = \Phi^{-1}$. $\mathcal{F}_d$ will denote the family all elements in $\mathcal{F}$ corresponding to $\mu \in \mathcal{P}_d$, i.e. all functions of the form $F(x) = \sum_{i=1}^n \alpha_i \mathbf{1}_{[x_i,\infty)}(x)$ with $n \in \mathbb{N}$, $\{x_1, \ldots, x_n\} \subseteq \mathbb{R}, \{\alpha_1, \ldots, \alpha_n\} \subseteq (0,1]^d$ and $\sum_{i=1}^n \alpha_i = 1$. Obviously Φ maps $\mathcal{P}_d$ in a one-to-one manner to $\mathcal{F}_d$ and Ψ maps $\mathcal{F}_d$ in a one-to-one-manner to $\mathcal{P}_d$. Due to monotonicity of every $F \in \mathcal{F}$ the left- and right limit of F at $x \in \mathbb{R}$ exist and will be denoted by $F(x-)$ and $F(x+)$ respectively.

Following Tavalera et al. [6] and Yager [8] we will only consider Dempster–Shafer belief structures (DSBS, for short) on $\mathbb{R}$ with pairwise different non-empty compact intervals $B_1 = [a_1, b_1]$, $B_2 = [a_2, b_2]$, ..., $B_n = [a_n, b_n]$ as focal elements, i.e. we consider mappings $\mathfrak{m} : 2^{\mathbb{R}} \to [0,1]$ fulfilling

(i) $\mathfrak{m}(A) = 0$ if $A \neq B_i$ for all $i \in \{1, \ldots, n\}$
(ii) $\sum_{i=1}^n \mathfrak{m}(B_i) = 1$
(iii) $B_i \in \mathcal{K}(\mathbb{R})$ for every $i \in \{1, \ldots, n\}$ and $B_i \neq B_j$ for $i \neq j$.

Note that, contrary to half open intervals, considering non-empty compact intervals does not exclude the possibility of having focal elements containing only a single point. Without loss of generality we will also assume $\mathfrak{m}(B_i) > 0$ for every $i \in \{1, \ldots, n\}$. Each such DSBS will be denoted in the form $(\mathbb{R}, \mathfrak{m}, (B_i)_{i=1}^n), \mathcal{D}$ will denote the family of all these DBSB. For every DSBS $(\mathbb{R}, \mathfrak{m}, (B_i)_{i=1}^n)$, the sets $L_\mathfrak{m}, R_\mathfrak{m}$ are defined by

$$L_\mathfrak{m} = \{a_1, \ldots, a_n\} \quad \text{and} \quad R_\mathfrak{m} = \{b_1, \ldots, b_n\}. \tag{3}$$

Obviously the cardinalities $\#L_\mathfrak{m}, \#R_\mathfrak{m}$ of $L_\mathfrak{m}, R_\mathfrak{m}$ fulfill $\#L_\mathfrak{m}, \#R_\mathfrak{m} \leqslant n$. Every DSBS $(\mathbb{R}, \mathfrak{m}, (B_i)_{i=1}^n)$ induces a *belief measure* $Bel_\mathfrak{m} : 2^{\mathbb{R}} \to [0,1]$ and a *plausibility measure* $Pl_\mathfrak{m} : 2^{\mathbb{R}} \to [0,1]$ by setting

$$Bel_\mathfrak{m}(A) = \sum_{i:\ B_i \subseteq A \neq \emptyset} \mathfrak{m}(B_i), \qquad Pl_\mathfrak{m}(A) = \sum_{i:\ B_i \cap A \neq \emptyset} \mathfrak{m}(B_i) \tag{4}$$

for every $A \subseteq \mathbb{R}$. Note that, using the interpretation of DSBS with compact intervals as focal elements given in the Introduction, the pair $(Bel_\mathfrak{m}, Pl_\mathfrak{m})$ coincides with the lower and upper probability $(\underline{\pi}, \overline{\pi})$ induced by the discrete random compact interval $\mathbf{X}$. Define functions $\underline{F}_\mathfrak{m}, \overline{F}_\mathfrak{m} : \mathbb{R} \to [0,1]$ by $\underline{F}_\mathfrak{m}(x) = Bel((-\infty, x])$ and $\overline{F}_\mathfrak{m}(x) = Pl((-\infty, x])$ for every $x \in \mathbb{R}$. Then the function $\mathbf{F}_\mathfrak{m} : \mathbb{R} \to \mathcal{K}([0,1])$,

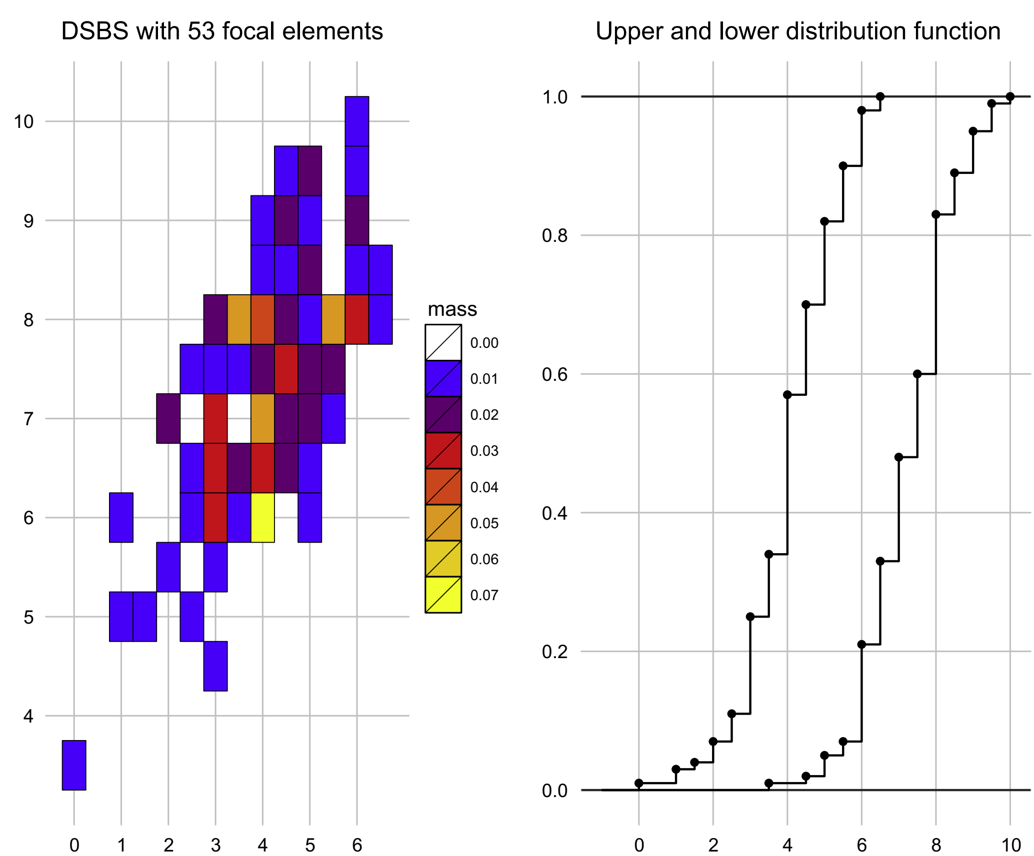

Fig. 1. DSBS with 53 intervals $[a_i, b_i]$ as focal elements (left) and the corresponding BCDF (right); all a_i, b_i are elements of $\mathbb{N}/2 = \{0, 1/2, 1, 3/2, 2, \ldots\}$.