Bayesian Updating under Incomplete or Imprecise Information in Finite Spaces

Giulianella Coletti[1], Davide Petturiti[2], and Barbara Vantaggi[2]

[1] Dip. Matematica e Informatica, Università di Perugia, Italy
giulianella.coletti@unipg.it
[2] Dip. S.B.A.I., Università di Roma "La Sapienza", Italy
{barbara.vantaggi,davide.petturiti}@sbai.uniroma1.it

Abstract. We provide (in a finite setting) a closed form expression for the lower envelope of the set of all the possible Bayesian posteriors derivable from a possibly incomplete or imprecise prior distribution (giving rise to a 2-monotone capacity) and a likelihood function.

Keywords: Bayesian updating, coherence, conditional probability, belief function, 2-monotone capacity.

1 Introduction

The classical Bayesian paradigm relies on a precise and complete probabilistic *prior* and *likelihood* assessment $\{P(H_i), P(E|H_i)\}_{i=1,...,n}$ and gives rise to a unique *posterior* distribution $\{P(H_i|E)\}_{i=1,...,n}$, whenever $P(E) > 0$. However, in real applications (e.g., medical diagnosis, forensic analysis and legal processes, to cite some) the prior knowledge could be imprecise (e.g., a belief function) or, even if precise, it could be only partially specified or defined on different hypotheses. At the same time, the expert could be interested in Bayesian queries on events more complex than the $H_i|E$'s.

The cases described above induce a (convex) set of prior probabilities whose lower envelope turns out to be a *belief function* [12,20,14,6]. Hence, the problem of non-unicity of the posterior needs to be dealt referring to the entire class of probabilistic extensions, and a characterization of the envelopes of such set is desirable, especially with a *sensitivity analysis* in view.

The main aim of this paper is to prove a generalized version of Bayes' theorem for finite spaces when the prior information is expressed by a 2-monotone capacity on the algebra spanned by the H_i's and the statistical model is still a likelihood function on the events $E|H_i$'s. Actually, our results can be generalized (see [5]) in order to extend results proved in [25,26], by allowing conditioning to any event in the algebra $\mathcal{A}$ spanned by E and the H_i's, without any positivity assumption on the corresponding (lower or upper) probability. This aim is in line with that of Walley [24].

Our contribution consists in providing a closed form expression for the lower envelope of the set of full conditional probabilities on $\mathcal{A}$ extending a complete

P. Grzegorzewski et al. (eds.), *Strengthening Links between Data Analysis & Soft Computing*,
Advances in Intelligent Systems and Computing 315, DOI: 10.1007/978-3-319-10765-3_9

and precise prior probability and a likelihood function. Then we characterize the lower envelope of the coherent conditional probability extensions of a prior probability referring to events different from those where the likelihood is given. Finally, a generalization of the first result is proved, by considering a prior 2-monotone capacity and a likelihood function. We show that the "lower posterior probability" may fail 2-monotonicity: in the case the lower posterior probability is a 2-monotone capacity, then the updating procedure can be iterated.

2 Framework of Reference

Let $\mathcal{A}$ be a Boolean algebra of *events*, endowed with the usual Boolean operations of contrary $(\cdot)^c$, disjunction $\vee$, and conjunction $\wedge$, and the partial order of implication $\subseteq$. We denote with Ω and $\emptyset$, respectively, the *sure event* and the *impossible event* which coincide with the top and the bottom elements of $\mathcal{A}$, respectively. A subset $\mathcal{H} \subseteq \mathcal{A}^0 = \mathcal{A} \setminus \{\emptyset\}$ is said an *additive class* if it is closed under finite disjunctions.

We refer to the following axiomatic definition of *conditional probability* [7] which is equivalent to [10,9].

Definition 1. *Let $\mathcal{A}$ be a Boolean algebra and $\mathcal{H} \subseteq \mathcal{A}^0$ an additive class. A function $P : \mathcal{A} \times \mathcal{H} \to [0,1]$ is a* **conditional probability** *if it satisfies the following conditions:*

(i) $P(E|H) = P(E \wedge H|H)$, *for every* $E \in \mathcal{A}$ *and* $H \in \mathcal{H}$*;*
(ii) $P(\cdot|H)$ *is a finitely additive probability on* $\mathcal{A}$, *for any* $H \in \mathcal{H}$*;*
(iii) $P(E \wedge F|H) = P(E|H) \cdot P(F|E \wedge H)$, *for any* $H, E \wedge H \in \mathcal{H}$ *and* $E, F \in \mathcal{A}$.

Following [13], we say that a conditional probability $P(\cdot|\cdot)$ is *full on* $\mathcal{A}$ if $\mathcal{H} = \mathcal{A}^0$. In order to deal with an assessment P on an *arbitrary* set $\mathcal{G}$ of conditional events we need to resort to the concept of *coherence* [7] (equivalent to [27,17]).

Definition 2. *Given an arbitrary set $\mathcal{G} = \{E_i|H_i\}_{i \in I}$ of conditional events, an assessment $P : \mathcal{G} \to [0,1]$ is a* **coherent conditional probability** *if and only if there is a conditional probability $\tilde{P} : \mathcal{A} \times \mathcal{H} \to [0,1]$ with $\mathcal{A} \times \mathcal{H} \supseteq \mathcal{G}$ extending the assessment P (i.e., $\tilde{P}_{|\mathcal{G}} = P$).*

By the conditional version [27,17] of *de Finetti's fundamental theorem for probabilities* [11], any coherent conditional probability P on $\mathcal{G}$ can be extended coherently to any further set $\mathcal{G}' \supset \mathcal{G}$ of conditional events. In general, the extension on $\mathcal{G}'$ is not unique thus we consider the set $\mathcal{P} = \{\tilde{P}(\cdot|\cdot)\}$ of all the coherent extensions of P. Such set is a compact subset of the space $[0,1]^{\mathcal{G}'}$ endowed with the product topology of pointwise convergence and is the Cartesian product of (possibly degenerate) closed intervals, which determine the *lower and upper envelopes* $\underline{P} = \min \mathcal{P}$ and $\overline{P} = \max \mathcal{P}$, where the minimum and the maximum are intended pointwise on the elements of $\mathcal{G}'$. The functions $\underline{P}$ and $\overline{P}$ on $\mathcal{G}'$ are *coherent lower and upper conditional probabilities* [7], respectively.

Notice that $\underline{P}$ and $\overline{P}$ are *dual*, i.e., $\overline{P}(E|H) = 1 - \underline{P}(E^c|H)$ if $E|H, E^c|H \in \mathcal{G}'$, thus, when $\mathcal{G}'$ is a structured set $\mathcal{A} \times \mathcal{H}$ the knowledge of $\underline{P}$ (simply called *lower conditional probability* in this case) is sufficient to recover $\overline{P}$.

Recall that a lower conditional probability $\underline{P}$ on $\mathcal{A} \times \mathcal{H}$ is such that for every $H \in \mathcal{H}$, $\underline{P}(\emptyset|H) = 0$, $\underline{P}(\Omega|H) = 1$, $\underline{P}(E|H) = \underline{P}(E \wedge H|H)$ and $\underline{P}(\cdot|H)$ is super-additive on $\mathcal{A}$. Furthermore, for $H \in \mathcal{H}$, $\underline{P}(\cdot|H)$ is said n-monotone ($n \geq 2$) on $\mathcal{A}$ if

$$\underline{P}\left(\bigvee_{i=1}^{n} E_i \middle| H\right) \geq \sum_{\emptyset \neq I \subseteq \{1,\ldots,n\}} (-1)^{|I|+1} \underline{P}\left(\bigwedge_{i \in I} E_i \middle| H\right), \qquad (1)$$

for every $E_1, \ldots, E_n \in \mathcal{A}$. In particular, for $H \in \mathcal{H}$, $\underline{P}(\cdot|H)$ is said a *belief function* [20] on $\mathcal{A}$ if it is n-monotone for every $n \geq 2$.

3 Precise and Complete Prior and Likelihood Function

Let $\mathcal{L} = \{H_1, \ldots, H_n\}$ be a finite partition of Ω, E an arbitrary possible event, and $\mathcal{A} = \langle \{E\} \cup \mathcal{L} \rangle$ the algebra spanned by $\{E\} \cup \mathcal{L}$, whose set of atoms is $\mathcal{C}_{\mathcal{A}}$.

A *likelihood function* f (see, e.g., [4]) is *any* map from $\{E\} \times \mathcal{L}$ to $[0, 1]$, with the only constraint that $f(E|H_i) = 0$ if $E \wedge H_i = \emptyset$ and $f(E|H_i) = 1$ if $E \wedge H_i = H_i$.

Given a likelihood function $f(E|\cdot)$ and a *prior probability distribution* $p(\cdot)$ on $\mathcal{L}$, the joint assessment $\{p, f\}$ is a coherent conditional probability on $\mathcal{G} = \{E|H_i, H_i\}_{i=1,\ldots,n}$ [18,7,22] which determines a unique coherent extension P on $\mathcal{G}' = \mathcal{A} \times (\{\Omega\} \cup \mathcal{L})$. Nevertheless, the further extension of P on $\mathcal{A} \times \mathcal{A}^0$ is not unique in general so we need to consider the set

$$\mathcal{P} = \{\tilde{P} : \text{full conditional probability on } \mathcal{A} \text{ s.t. } \tilde{P}_{|\mathcal{G}'} = P\}.$$

The following theorem provides a closed form expression for $\underline{P} = \min \mathcal{P}$.

Theorem 1. *Given a likelihood function $f(E|\cdot)$ and a prior probability distribution $p(\cdot)$ on $\mathcal{L}$, for every $F|K \in \mathcal{A} \times \mathcal{A}^0$, $\underline{P}(F|K) = 1$ when $F \wedge K = K$, and if $F \wedge K \neq K$, then:*

(i) if $P(K) > 0$ then

$$\underline{P}(F|K) = \frac{P(F \wedge K)}{P(K)}; \qquad (2)$$

(ii) if $P(K) = 0$, then if $I \neq \emptyset$ and $H_j \wedge F \wedge K \neq \emptyset$ for all $j \in J$ and $F^c \wedge K \wedge \left(\bigvee_{i \in I} H_i\right)^c = \emptyset$, where $I, J \subseteq \{1, \ldots, n\}$ are, respectively, the maximum and minimum index set such that $\bigvee_{i \in I} H_i \subseteq K \subseteq \bigvee_{j \in J} H_j$, then

$$\underline{P}(F|K) = \min \left\{ \min_{\substack{E \wedge H_i \subseteq F \wedge K \\ E^c \wedge H_i \subseteq F^c \wedge K}} f(E|H_i), \min_{\substack{E^c \wedge H_i \subseteq F \wedge K \\ E \wedge H_i \subseteq F^c \wedge K}} (1 - f(E|H_i)) \right\}; \qquad (3)$$

otherwise $\underline{P}(F|K) = 0$.

Proof. The proof is trivial in the case $F \wedge K = K$ or $P(K) > 0$. Assume $F \wedge K \neq K$ and $P(K) = 0$. Denote with $E^\bullet$ either E or E^c and let $\mathcal{C}_1 = \{E^\bullet \wedge H_i \in \mathcal{C}_\mathcal{A} : P(E^\bullet \wedge H_i) = 0\}$. The lower bound $\underline{P}(F|K)$ can be computed by solving the optimization problem (see [7,1]) with non-negative unknowns x_j^1 for $E \wedge H_j \in \mathcal{C}_1$, $j \in J$, and y_j^1 for $E^c \wedge H_j \in \mathcal{C}_1$, $j \in J$,

$$\text{minimize} \left[\sum_{E \wedge H_j \subseteq F \wedge K} x_j^1 + \sum_{E^c \wedge H_j \subseteq F \wedge K} y_j^1 \right]$$

$$\begin{cases} x_j^1 = f(E|H_j) \cdot (x_j^1 + y_j^1) & \text{if } E \wedge H_j \in \mathcal{C}_1 \text{ and } E^c \wedge H_j \in \mathcal{C}_1 \text{ and } j \in J, \\ \sum\limits_{E \wedge H_j \subseteq K} x_j^1 + \sum\limits_{E^c \wedge H_j \subseteq K} y_j^1 = 1. \end{cases}$$

The unknowns in the system are divided in independent groups corresponding to each H_j with $j \in J$ and are constrained together only by the last equation. If $I = \emptyset$ or there exits $j \in J$ s.t. $H_j \wedge F \wedge K = \emptyset$ or $F \wedge K \wedge \left(\bigvee_{i \in I} H_i\right)^c \neq \emptyset$, one can always build a solution such that $\sum\limits_{E \wedge H_j \subseteq F \wedge K} \mathbf{x}_j^1 + \sum\limits_{E^c \wedge H_j \subseteq F \wedge K} \mathbf{y}_j^1 = 0$ and $\sum\limits_{E \wedge H_j \subseteq F^c \wedge K} \mathbf{x}_j^1 + \sum\limits_{E^c \wedge H_j \subseteq F^c \wedge K} \mathbf{y}_j^1 = 1$, which implies $\underline{P}(F|K) = 0$. In the opposite case the minimum is achieved in correspondence of those solutions such that $\mathbf{x}_i^1 + \mathbf{y}_i^1 = 1$ for $E^\bullet \wedge H_i \subseteq F \wedge K$ and $(E^\bullet)^c \wedge H_i \subseteq F^c \wedge K$, thus the conclusion follows. □

Let us note that if $P(K) > 0$, $\underline{P}(\cdot|K)$ is a probability measure (and so a belief function) on $\mathcal{A}$. However the following example shows that for some $K \in \mathcal{A}^0$ with $P(K) = 0$, the lower envelope $\underline{P}(\cdot|K)$ can fail even 2-monotonicity.

Example 1. Let $\mathcal{L} = \{H_1, H_2, H_3, H_4\}$ be a partition of Ω and E an event logically independent of $\mathcal{L}$. Consider the likelihood $f(E|H_i) = \frac{1}{2}$, $i = 1, 2, 3, 4$, and the prior probability distribution $p(H_1) = 1$ and $p(H_i) = 0$, $i = 2, 3, 4$.

Let $K = H_2 \vee H_3 \vee H_4$ and $F = (E \wedge H_2) \vee (E^c \wedge H_3) \vee (E \wedge H_4)$. It holds $\underline{P}(E \vee F|K) = \underline{P}(E|K) = \underline{P}(F|K) = \frac{1}{2}$ and $\underline{P}(E \wedge F|K) = 0$, which implies $\underline{P}(\cdot|K)$ is not 2-monotone on $\mathcal{A} = \langle \{E\} \cup \mathcal{L} \rangle$ since it is $\underline{P}(E \vee F|K) < \underline{P}(E|K) + \underline{P}(F|K) - \underline{P}(E \wedge F|K)$.

4 Imprecise or Partial Prior Information

Consider two finite Boolean algebras of events $\mathcal{A}, \mathcal{A}'$, and a probability measure P on $\mathcal{A}$. If the algebra of interest is $\mathcal{A}'$ we can consider the set of coherent extensions on $\mathcal{G}' = (\mathcal{A} \times \{\Omega\}) \cup (\mathcal{A}' \times \mathcal{A}'^0)$

$$\mathcal{P} = \{\tilde{P} : \text{coherent conditional probability on } \mathcal{G}' \text{ s.t. } \tilde{P}_{|\mathcal{A} \times \{\Omega\}} = P\}$$

with its lower envelope $\underline{P} = \min \mathcal{P}$. Next theorem provides a closed form expression for $\underline{P}$ on $\mathcal{A}' \times \mathcal{A}'^0$, relying on the lower and upper probabilities $\underline{P}(\cdot) = \underline{P}(\cdot|\Omega)$ and $\overline{P}(\cdot) = \overline{P}(\cdot|\Omega)$ on $\mathcal{A}'$, obtained extending P on $\mathcal{A} \cup \mathcal{A}'$, which are known to be, respectively, a belief function and a plausibility function [14].

Theorem 2. *Let $\mathcal{A}, \mathcal{A}'$ be two finite Boolean algebras, P a probability measure on $\mathcal{A}$, and $\underline{P}(\cdot|\cdot)$ the lower envelope of the set of coherent extensions of P on $\mathcal{G}'$. The following statements hold:*

(i) $\underline{P}(\cdot|K)$ *is a belief function on* $\mathcal{A}'$, *for every* $K \in \mathcal{A}'^0$;
(ii) *for every* $F|K \in \mathcal{A}' \times \mathcal{A}'^0$, $\underline{P}(F|K) = 1$ *when* $F \wedge K = K$, *and if* $F \wedge K \neq K$, *then we have*

$$\underline{P}(F|K) = \begin{cases} \frac{\underline{P}(F \wedge K)}{\underline{P}(F \wedge K) + \overline{P}(F^c \wedge K)} & \text{if } \underline{P}(F \wedge K) + \overline{P}(F^c \wedge K) > 0, \\ 0 & \text{otherwise.} \end{cases} \quad (4)$$

Proof. We prove condition *(ii)* first. If $F \wedge K = K$, for every $\tilde{P} \in \mathcal{P}$, $\tilde{P}(F|K) = 1$, so $\underline{P}(F|K) = 1$. Hence assume $F \wedge K \neq K$. By Proposition 3.1 in [14], $\underline{P}(\cdot)$ is a belief function on $\mathcal{A}'$, so Theorem 7.2 in [23] implies equation (4) when $\underline{P}(F \wedge K) + \overline{P}(F^c \wedge K) > 0$. Finally, in the case $\underline{P}(F \wedge K) + \overline{P}(F^c \wedge K) = 0$ equation (4) follows by Proposition 3 in [1].

Now we prove condition *(i)*. Theorem 1 in [15] (or, equivalently, Theorem 4.1 in [21]) implies that $\underline{P}(\cdot|K)$ is a belief function on $\mathcal{A}'$ when $\underline{P}(K) > 0$, which implies $\underline{P}(F \wedge K) + \overline{P}(F^c \wedge K) > 0$. When $\underline{P}(K) = 0$, the claim follows by the monotonicity of $\underline{P}(\cdot|K)$ and since $\underline{P}(F|K) > 0$ only for events $F \in \mathcal{A}'$ such that $F \wedge K = K$. □

Previous theorem differs from Theorem 7.2 in [23], where $\underline{P}(F|K)$ is not defined when $\overline{P}(K) = 0$, moreover, in the case $\overline{P}(K) > 0$ and $\underline{P}(F \wedge K) + \overline{P}(F^c \wedge K) = 0$, $\underline{P}(F|K)$ is set equal to 1, which is not the minimum coherent value for $F|K$ (actually it is the maximum). The quoted result refers to the *regular extension* for lower previsions [24]. On the other hand, by considering the *natural extension*, a result equivalent to our Theorem 2 follows [24,16].

Let φ be a 2-monotone capacity on $\mathcal{A}'$ together with its dual $\overline{\varphi}$ and consider

$$\mathcal{P}_\varphi = \{\tilde{P} : \text{probability on } \mathcal{A}' \text{ s.t. } \varphi \leq \tilde{P} \leq \overline{\varphi}\}. \quad (5)$$

If φ is a belief function on a finite Boolean algebra $\mathcal{A}'$, Corollary 3.6 in [14] assures the existence of a finite algebra $\mathcal{A}$ and a probability measure P on $\mathcal{A}$, such that φ is obtained as the lower envelope on $\mathcal{A}'$ of the set of coherent extensions of P on $\mathcal{A} \cup \mathcal{A}'$. In this case, Theorem 2 characterizes the lower envelope of the set of full conditional probabilities obtained extending each $\tilde{P} \in \mathcal{P}_\varphi$ on $\mathcal{A}' \times \mathcal{A}'^0$. Hence, the same theorem characterizes also the lower envelope of the set of coherent extensions on $\mathcal{A}' \times \mathcal{A}'^0$ of a belief function (viewed as a lower probability on $\mathcal{A}'$).

Let $\mathcal{L} = \{H_1, \dots, H_n\}$ be a finite partition, a *partial prior probability distribution* is a coherent probability P on a set of incompatible events $\{K_1, \dots, K_m\} \subseteq \langle\mathcal{L}\rangle^0$. In [6] it has been shown that the lower envelope of the set of coherent extensions of P on $\langle\mathcal{L}\rangle$ is a belief function, thus also in this case Theorem 2 characterizes the lower envelope of the coherent extensions on $\langle\mathcal{L}\rangle \times \langle\mathcal{L}\rangle^0$.

5 2-monotone Prior Capacity and Likelihood Function

Given $\mathcal{L}$ and E as in Section 3, here we assume that our knowledge *a priori* is expressed by a 2-monotone capacity φ on $\langle\mathcal{L}\rangle$ while the statistical model is still represented by a likelihood function $f(E|\cdot)$ on $\mathcal{L}$. By Proposition 1 in [18] the assessment $\{\tilde{P}, f\}$ is a coherent conditional probability for every $\tilde{P} \in \mathcal{P}_\varphi$, thus the assessment $\{\varphi, f\}$ is a coherent lower conditional probability. Our aim is to provide a closed form expression for the lower envelope $\underline{P}$ of the set of coherent extensions of $\{\varphi, f\}$ on $\mathcal{A} \times \mathcal{A}^0$, with $\mathcal{A} = \langle \{E\} \cup \mathcal{L} \rangle$.

Next theorem characterizes the lower envelope $\underline{P}(\cdot) = \underline{P}(\cdot|\Omega)$ on $\mathcal{A} \times \{\Omega\}$ as a Choquet integral with respect to φ and it generalizes a result given in [3]. For this aim, for every $F \in \mathcal{A}$ define the $\langle \mathcal{L} \rangle$-measurable function $G_F : \mathcal{L} \to [0, 1]$

$$G_F(H_i) = \begin{cases} 0 & \text{if } F \wedge H_i = \emptyset, \\ 1 & \text{if } F \wedge H_i = H_i, \\ f(E|H_i) & \text{if } F \wedge E \wedge H_i \neq \emptyset = F \wedge E^c \wedge H_i, \\ 1 - f(E|H_i) & \text{if } F \wedge E^c \wedge H_i \neq \emptyset = F \wedge E \wedge H_i. \end{cases} \tag{6}$$

Theorem 3. *Given a likelihood function $f(E|\cdot)$ on $\mathcal{L}$ and a 2-monotone capacity $\varphi(\cdot)$ on $\langle \mathcal{L} \rangle$, for every $F \in \mathcal{A}$ it holds*

$$\underline{P}(F) = \oint G_F \mathrm{d}\varphi = \int_0^{+\infty} \varphi\left(\bigvee \{H_i \in \mathcal{L} \,:\, G_F(H_i) \geq x\}\right) \mathrm{d}x.$$

Proof. For every $F \in \mathcal{A}$ and $\tilde{P} \in \mathcal{P}_\varphi$, the probability of F is the expectation of G_F with respect to $\tilde{P}$, so $\underline{P}(F)$ coincides with the minimum of the expectations varying $\tilde{P} \in \mathcal{P}_\varphi$. The proof follows by Proposition 3 in [19] which implies that the lower expectation of G_F with respect to the class of probabilities $\mathcal{P}_\varphi$ coincides with the Choquet integral of G_F with respect to φ. □

Theorem 3 characterizes also the dual upper envelope $\overline{P}(\cdot) = \overline{P}(\cdot|\Omega)$ on $\mathcal{A} \times \{\Omega\}$ as a Choquet integral with respect to $\overline{\varphi}$. Given $\underline{P}(\cdot), \overline{P}(\cdot)$ on $\mathcal{A}$, for every $F|K \in \mathcal{A} \times \mathcal{A}^0$ define

$$L(F \wedge K) = \min\left\{\int G_{F \wedge K} \mathrm{d}\tilde{P} \,:\, \tilde{P} \in \mathcal{P}_\varphi, \int G_{F^c \wedge K} \mathrm{d}\tilde{P} = \overline{P}(F^c \wedge K)\right\}, \tag{7}$$

$$U(F^c \wedge K) = \max\left\{\int G_{F^c \wedge K} \mathrm{d}\tilde{P} \,:\, \tilde{P} \in \mathcal{P}_\varphi, \int G_{F \wedge K} \mathrm{d}\tilde{P} = \underline{P}(F \wedge K)\right\}. \tag{8}$$

Note that it holds in general $\underline{P}(F \wedge K) \leq L(F \wedge K)$ and $U(F^c \wedge K) \leq \overline{P}(F^c \wedge K)$.

The min and max in equations (7) and (8) are attained in correspondence of the extreme points of the set $\mathcal{P}_\varphi$, characterized in [2], whose number is at most $n!$ (i.e., the permutations of $\mathcal{L}$).

Next theorem provides a complete characterization of $\underline{P}(\cdot|\cdot)$ on $\mathcal{A} \times \mathcal{A}^0$ in terms of $\underline{P}(\cdot)$, $\overline{P}(\cdot)$, $L(\cdot)$ and $U(\cdot)$.

Theorem 4. *Given a likelihood function $f(E|\cdot)$ on $\mathcal{L}$ and a 2-monotone capacity φ on $\langle \mathcal{L} \rangle$, for every $F|K \in \mathcal{A} \times \mathcal{A}^0$, $\underline{P}(F|K) = 1$ when $F \wedge K = K$, and if $F \wedge K \neq K$, then:*

(i) if $\underline{P}(K) > 0$ then

$$\underline{P}(F|K) = \min\left\{\frac{\underline{P}(F \wedge K)}{\underline{P}(F \wedge K) + U(F^c \wedge K)}, \frac{L(F \wedge K)}{L(F \wedge K) + \overline{P}(F^c \wedge K)}\right\}; \tag{9}$$

(ii) if $\underline{P}(K) = 0$*, then if* $I \neq \emptyset$ *and* $H_j \wedge F \wedge K \neq \emptyset$ *for all* $j \in J$ *and* $F^c \wedge K \wedge \left(\bigvee_{i\in I} H_i\right)^c = \emptyset$*, where* $I, J \subseteq \{1,\ldots,n\}$ *are, respectively, the maximum and minimum index set such that* $\bigvee_{i\in I} H_i \subseteq K \subseteq \bigvee_{j\in J} H_j$*, then*

$$\underline{P}(F|K) = \min\left\{\min_{\substack{E\wedge H_i \subseteq F\wedge K\\ E^c\wedge H_i\subseteq F^c\wedge K}} f(E|H_i), \min_{\substack{E^c\wedge H_i\subseteq F\wedge K\\ E\wedge H_i\subseteq F^c\wedge K}}(1-f(E|H_i))\right\}; \quad (10)$$

otherwise $\underline{P}(F|K) = 0$*.*

Proof. Let $\mathcal{P} = \{\tilde{P}(\cdot|\cdot)\}$ be the set of full conditional probabilities on $\mathcal{A} \times \mathcal{A}^0$ such that $\tilde{P}_{|\{E\}\times\mathcal{L}} = f$ and $\varphi(\cdot) \leq \tilde{P}(\cdot|\Omega) \leq \overline{\varphi}(\cdot)$, with $\overline{\varphi}$ the dual capacity of φ. If $F \wedge K = K$, then, for every $\tilde{P} \in \mathcal{P}$, it follows $\tilde{P}(F|K) = 1$, which implies $\underline{P}(F|K) = 1$. Hence assume $F \wedge K \neq K$.

To prove condition *(i)*, suppose $\underline{P}(K) > 0$, which implies $\tilde{P}(K) = \tilde{P}(K|\Omega) > 0$ for every $\tilde{P} \in \mathcal{P}$, and so $\underline{P}(F|K) = \min\left\{\frac{\tilde{P}(F\wedge K)}{\tilde{P}(F\wedge K)+\tilde{P}(F^c\wedge K)} : \tilde{P} \in \mathcal{P}\right\}$. The conclusion follows since the real function $\frac{x}{x+y}$ is increasing in x and decreasing in y, so the minimum is attained in correspondence of $\frac{\underline{P}(F\wedge K)}{\underline{P}(F\wedge K)+U(F^c\wedge K)}$ or $\frac{L(F\wedge K)}{L(F\wedge K)+\overline{P}(F^c\wedge K)}$. Finally, condition *(ii)* is implied by the extension procedure described in [8] and Theorem 1. □

In particular, if $\underline{P}(E) > 0$, then for every $F \in \mathcal{A}$ we have $\underline{P}(F\wedge E) = L(F\wedge E)$ and $\overline{P}(F^c\wedge E) = U(F^c\wedge E)$, thus Theorem 4 implies $\underline{P}(F|E) = \frac{\underline{P}(F\wedge E)}{\underline{P}(F\wedge E)+\overline{P}(F^c\wedge E)}$, which coincides with the *lower posterior probability* defined in [25,26].

Note that for all $F|K \in \langle\mathcal{L}\rangle \times \langle\mathcal{L}\rangle^0$, if φ is a belief function and $\underline{P}(K) = \varphi(K) > 0$, then $\underline{P}(\cdot|\cdot)$ on $\langle\mathcal{L}\rangle \times \langle\mathcal{L}\rangle^0$ has the same characterization given in Theorem 2. As a further consequence, for all $F|K \in \mathcal{A} \times \langle\mathcal{L}\rangle^0$, $\underline{P}(F|K)$ can be expressed as the Choquet integral of G_F with respect to the restriction of $\underline{P}(\cdot|K)$ on $\langle\mathcal{L}\rangle$, that is $\underline{P}(F|K) = \oint G_F(\cdot)\mathrm{d}\underline{P}(\cdot|K)$.

Notice that also for the function $\underline{P}(\cdot|K)$ studied in this section (in particular for the lower posterior probability) 2-monotonicity may fail when $\underline{P}(K) = 0$ (see, again, Example 1). In the case the lower posterior probability is 2-monotone, previous results can be used in order to iterate the updating procedure by taking as new prior a lower posterior probability and considering a likelihood function related to another evidence.

References

1. Capotorti, A., Vantaggi, B.: Locally strong coherence in inference processes. Ann. of Math. and Art. Int. 35(1-4), 125–149 (2002)
2. Chateauneuf, A., Jaffray, J.-Y.: Some characterizations of lower probabilities and other monotone capacities through the use of Möbius inversion. Math. Soc. Sci. 17(3), 263–283 (1989)
3. Coletti, G., Gervasi, O., Tasso, S., Vantaggi, B.: Generalized Bayesian inference in a fuzzy context: From theory to a virtual reality application. Comp. Stat. & Data Anal. 56(4), 967–980 (2012)

4. Coletti, G., Petturiti, D., Vantaggi, B.: Possibilistic and probabilistic likelihood functions and their extensions: Common features and specific characteristics. Fuzzy Sets and Sys. (in press), doi: 10.1016/j.fss.2013.09.010
5. Coletti, G., Petturiti, D., Vantaggi, B.: Bayesian inference: the role of coherence to deal with a prior belief function Stat. Meth. and App. (under review)
6. Coletti, G., Scozzafava, R.: Toward a General Theory of Conditional Beliefs. Int. J. of Int. Sys. 21, 229–259 (2006)
7. Coletti, G., Scozzafava, R.: Probabilistic Logic in a Coherent Setting. Trends in Logic, vol. 15. Kluwer Academic Publisher, Dordrecht (2002)
8. Coletti, G., Vantaggi, B.: Probabilistic reasoning with vague information. In: Proc. of the 2nd World Conf. on Soft. Comp., Baku, Azerbaijan, December 3-5 (2012)
9. Császár, Á.: Sur la structure des espaces de probabilitè conditionelle. Acta Math. Ac. Sci. Hung. 6, 337–361 (1955)
10. de Finetti, B.: Sull'impostazione assiomatica del calcolo delle probabilità. Annali Triestini 19(2), 29–81 (1949)
11. de Finetti, B.: Theory of Probability 1-2. John Wiley & Sons, London (1975)
12. Dempster, A.P.: Upper and lower probabilities induced by a multivalued mapping. Ann. of Math. Stat. 38(2), 325–339 (1967)
13. Dubins, L.E.: Finitely additive conditional probabilities, conglomerability and disintegrations. Ann. of Prob. 3(1), 89–99 (1975)
14. Fagin, R., Halpern, J.Y.: Uncertainty, belief, and probability. IBM Res. Rep., RJ 6191 (1988)
15. Jaffray, J.-Y.: Bayesian updating and belief functions. IEEE Trans. on Sys., Man and Cyb. 22(5), 1144–1152 (1992)
16. Miranda, E., Montes, I.: Coherent updating of 2-monotone previsions. In: 8th Int. Symp. on Imp. Prob.: Th. and App., Compiègne, France (2013)
17. Regazzini, E.: Finitely additive conditional probabilities. Rend. del Sem. Mat. e Fis. di Milano 55(1), 69–89 (1985)
18. Regazzini, E.: De Finetti's coherence and statistical inference. Ann. of Stat. 15(2), 845–864 (1987)
19. Schmeidler, D.: Integral representation without additivity. Proc. of the Am. Math. Soc. 97(2), 255–261 (1986)
20. Shafer, G.: A Mathematical Theory of Evidence. Princeton University Press, Princeton (1976)
21. Sundberg, C., Wagner, C.: Generalized Finite Differences and Bayesian Conditioning of Choquet Capacities. Adv. in App. Math. 13(3), 262–272 (1992)
22. Vantaggi, B.: Statistical matching of multiple sources: A look through coherence. Int. J. of Approx. Reas. 49(3), 701–711 (2008)
23. Walley, P.: Coherent lower (and upper) probabilities. Tech. Rep. n. 22, Department of Statistics, University of Warwick (1981)
24. Walley, P.: Statistical Reasoning with Imprecise Probabilities. Chapman and Hall, London (1991)
25. Wasserman, L.A.: Prior envelopes based on belief functions. Ann. of Stat. 18(1), 454–464 (1990)
26. Wasserman, L.A., Kadane, J.B.: Bayes' theorem for Choquet capacities. Ann. of Stat. 18(3), 1328–1339 (1990)
27. Williams, P.M.: Note on conditional previsions. Int. J. of Approx. Reas. 44, 366–383 (2007); (Unpublished manuscript 1975)