On general conditional random quantities

Veronica Biazzo
Dip. Mat. Inf.
Viale A. Doria, 6
95125 Catania (Italy)
vbiazzo@dmi.unict.it

Angelo Gilio
Dip. Me. Mo. Mat.
Via A. Scarpa, 16
00161 Roma (Italy)
gilio@dmmm.uniroma1.it

Giuseppe Sanfilippo
Dip. Sc. Stat. Mat.
Viale delle Scienze, ed. 13
90128 Palermo (Italy)
sanfilippo@unipa.it

Abstract

In the first part of this paper, recalling a general discussion on iterated conditioning given by de Finetti in the appendix of his book, vol. 2, we give a representation of a conditional random quantity $X|HK$ as $(X|H)|K$. In this way, we obtain the classical formula $\mathbb{P}(XH|K) = \mathbb{P}(X|HK)P(H|K)$, by simply using linearity of prevision. Then, we consider the notion of general conditional prevision $\mathbb{P}(X|Y)$, where X and Y are two random quantities, introduced in 1990 in a paper by Lad and Dickey. After recalling the case where Y is an event, we consider the case of discrete finite random quantities and we make some critical comments and examples. We give a notion of coherence for such more general conditional prevision assessments; then, we obtain a strong generalized compound prevision theorem. We study the coherence of a general conditional prevision assessment $\mathbb{P}(X|Y)$ when Y has no negative values and when Y has no positive values. Finally, we give some results on coherence of $\mathbb{P}(X|Y)$ when Y assumes both positive and negative values. In order to illustrate critical aspects and remarks we examine several examples.

Keywords. conditional events, general conditional random quantities, general conditional prevision assessments, generalized compound prevision theorem, iterated conditioning, strong generalized compound prevision theorem.

1 Introduction

This paper takes as its starting point the definition of *general conditional prevision* introduced by Lad and Dickey in [16] and also considered by Lad in his book [17]. In these works, the authors propose a general theory of conditional prevision specifying its operational meaning. This theory, which considers conditional prevision of the form $\mathbb{P}(X|Y)$ where both X and Y are random quantities, generalizes the de Finetti's definition of a conditional prevision assertion $\mathbb{P}(X|H)$, where H is an event. We observe that, denoting the indicator of H by the same symbol, to assume "H true" amounts to assuming $(H = 1)$ true, that is $(H \neq 0)$ true. Then, in the approach of Lad and Dickey, $X|H$ can be looked at as $X|Y$, where Y is the indicator of H; hence, $\mathbb{P}(X|H) = \mathbb{P}[X|(H = 1)]$. Notice that we discard the case where Y is the constant 0, as it reduces to the case $X|H$ where $(H \neq 0)$ is impossible. We recall that, concerning (precise or imprecise) conditional probability or prevision assessments like $P(E|H)$ or $\mathbb{P}(X|H)$, where E and H are events and X is a random quantity, theoretical results and algorithms in the framework of coherence have been given by many authors (see, for instance, [2, 3, 4, 5, 6, 8, 9, 10, 19, 20, 21, 22]) The checking of coherence and the extension of precise conditional prevision assessments have been studied in [7].
In [16, 17] the general conditional prevision $\mathbb{P}(X|Y)$ is defined as a number that you specify asserting your willingness to engage any transaction yielding a suitable random net gain and it is shown that such a generalization answers to questions of decision problems involving "state dependent preferences". In his book ([17]), Lad introduces the notion of general conditional random quantity $X|Y$ from the definition of conditional prevision $\mathbb{P}(X|Y)$. Obviously, as usual in a subjective setting, engaging a transaction requires a coherency of your assertion. In [16, 17], the coherency of $\mathbb{P}(X|Y)$ requires that a generalized compound prevision theorem is satisfied, that is the quantities $\mathbb{P}(XY)$, $\mathbb{P}(Y)$ and $\mathbb{P}(X|Y)$ must be such that $\mathbb{P}(XY) = \mathbb{P}(X|Y)\mathbb{P}(Y)$. But, the general case is different from the case where Y is the indicator of an event H. In fact, $\mathbb{P}(H) = 0$ implies $\mathbb{P}(XH) = 0$, and using coherence ([15, 18]) we can directly assess $\mathbb{P}(X|H)$. On the contrary, $\mathbb{P}(Y) = 0$ doesn't imply that $\mathbb{P}(XY) = 0$ and it could happen that it doesn't exist a finite value of $\mathbb{P}(X|Y)$ which satisfies the generalized compound prevision theorem. Thus, in this paper we propose a notion of coherence in order to handle the case $\mathbb{P}(Y) = 0$, integrating the Lad's defi-

nition of $\mathbb{P}(X|Y)$. Then, we give a strong generalized compound prevision theorem which follows from our definition of coherence. The random quantities, like X and Y, considered in this paper are finite discrete. The paper is organized as follows. In section 2 we recall some preliminary concepts and results. In section 3 we deepen, in the setting of coherence, the operational meaning of the assessments $\mathbb{P}(X|H)$ and $\mathbb{P}(X|HK)$, where H and K are events and X is a random quantity; then, based on a general discussion on *iterated conditioning* given by de Finetti in ([12], Vol. 2, Appendix, section 13), we look at $B|AH$ and $X|HK$, respectively, as $(B|A)|H$ and $(X|H)|K$; then, we give a representation for $B|AH$ and $X|HK$ which allows to obtain the classical results $\mathbb{P}(AB|H) = \mathbb{P}(B|AH)P(A|H)$ and $\mathbb{P}(XH|K) = \mathbb{P}(X|HK)P(H|K)$, by simply applying the linearity of prevision. In section 4, we recall the definitions of conditional prevision $\mathbb{P}(X|Y)$ and conditional random quantity $X|Y$; then, we examine a critical example. In section 5, after some critical comments, we propose an explicit definition of coherence for the conditional prevision $\mathbb{P}(X|Y)$; then, we give a strong generalized compound prevision theorem; we also examine many examples to illustrate some further aspects. In section 6, we study the coherence of a conditional prevision assessment $\mathbb{P}(X|Y) = \mu$, when Y has no negative values, or Y has no positive values. In section 7, we give some results concerning the coherence of the assessment $\mathbb{P}(X|Y) = \mu$, where Y assumes both positive and negative values. In section 8, we show some results concerning the set of coherent prevision assessments on $X|Y'$, where Y' is a linear transformation of Y. Finally, in section 9 we give some conclusions and an outlook on future research, which should concern more in general the case of imprecise conditional prevision assessments on families of conditional random quantities.

2 Some preliminary notions

We assume that each random quantity has a finite set of possible values. We denote by Ω (resp., $\emptyset$) the sure (resp., impossible) event; moreover, we denote by A^c the negation of A and by $A \vee B$ (resp., AB) the disjunction (resp., the conjunction) of A and B. We use the same symbol to denote an event and its indicator. We recall that in the subjective approach to probability, your assessment $P(E|H) = p$ means that You accept a bet on the conditional event $E|H$ in which You pay an amount ps, with $s \neq 0$, by receiving the random quantity $sHE + psH^c$, so that your net random gain is

$$G = sHE + psH^c - ps = sH(E - p)\,.$$

By excluding trivial cases, the value of G is, respectively, $s(1-p)$, or $-ps$, or 0, according to whether EH is true, or E^cH is true, or H^c is true.
We recall that, considering the restricted random gain $G|H = s(E-p) \in \{s(1-p), -ps\}$, it is $\min G|H \cdot \max G|H = -s^2p(1-p)$. Then, the coherence of p is defined by the condition ([15, 18]): $\min G|H \cdot \max G|H \leq 0$; that is $p(1-p) \geq 0$, which amounts to: $0 \leq p \leq 1$.
We observe that, to determine the coherent values of p, we don't consider all the values of G, but only those of $G|H$; in other words the value 0 of G associated with the case "H false" is "discarded".
We also observe that, denoting by the same symbol the (conditional) events and their indicators, by choosing $s = 1$ we obtain

$$E|H = EH + pH^c = EH + (1-H)p\,,$$

where the indicator, or truth-value function, $E|H$ represents the quantity we receive when we pay the amount $p = P(E|H)$. Then, by the linearity of prevision, we obtain:

$$P(E|H) = P(EH) + [1 - P(H)]p\,,$$

that is: $P(EH) = P(H)P(E|H)$ (*compound probability theorem*). We recall that, starting with a pioneering work of de Finetti ([11]), the notion of conditional event as a three-valued (logical and/or numerical) entity has been proposed by many authors (see, e.g., [1], [13], [14]). Based on the betting scheme, the notions of conditional prevision and conditional random quantity are defined and widely exploited in [17]. Truth-values of conditional events and their extension to decomposable conditional measures of uncertainty, with the aim of finding reasonable axioms for a general theory, have been discussed in many papers by Coletti and Scozzafava, see e.g. [9].

3 Representation of conditional random quantities

We remark that the general formula $P(AB|H) = P(A|H)P(B|AH)$ can be obtained by using the general coherence condition for conditional probability assessments. The same formula can be obtained, based on the linearity of prevision, by the following refined reasoning. Let $\mathcal{P} = (x, y, z)$ a probability assessment on $\mathcal{F} = \{A|H, B|AH, AB|H\}$. We observe that representing the indicator $B|AH$ as

$$B|AH = ABH + (1 - AH)y\,,$$

we obtain

$$P(B|AH) = y = P(ABH) + [1 - P(AH)]y\,,$$

from which it follows: $P(ABH) = P(AH)y$, i.e. $zP(H) = xyP(H)$; hence, to reach the conclusion we need to assume $P(H) > 0$. To bypass this obstacle, based on the general discussion on *iterated conditioning* given by de Finetti in ([12], Appendix of Vol. 2, section 13), we can look at $B|AH$ as $(B|A)|H$. Moreover, defining $p = P(B|A)$, we have $B|A = AB + (1-A)p$. Of course, when we pass from $B|A$ to $B|AH$, we must replace p by y. Then

$$\begin{aligned} B|AH &= (B|A)|H = AB|H + [(1-A)|H]y = \\ &= AB|H + (A^c|H)y = (AB + yA^c)|H\,. \end{aligned} \tag{1}$$

The representation above is not surprising, as shown by the following remarks:
(i) with the family $\mathcal{F}$ we can associate the partition $\{ABH, AB^cH, A^cH, H^c\}$;
(ii) under the hypothesis "H true", the random quantities $B|AH$ and $(AB + yA^c)|H$ coincide, as they always assume the same value, that is 1, or 0, or y, according to whether ABH is true, or AB^cH is true, or A^cH is true.
Hence, it must be: $\mathbb{P}(B|AH) = \mathbb{P}[(AB + yA^c)|H]$, with $\mathbb{P}(B|AH) = P(B|AH) = y$ and

$$\begin{aligned} \mathbb{P}(AB + yA^c)|H &= \mathbb{P}(AB|H) + \mathbb{P}(yA^c|H) = \\ &= P(AB|H) + yP(A^c|H) = z + y(1-x)\,. \end{aligned}$$

Then, we obtain: $y = z + y(1-x)$, i.e. $z = xy$.
Notice that, based on this result, we have that $B|AH$ and $(AB + yA^c)|H$ coincide also when H^c is true. In fact, the value of $B|AH$ (resp., $(AB+yA^c)|H$) associated with H^c is y (resp., $z+y(1-x) = y+z-xy = y$).
Now, by generalizing the previous reasoning, given an event H and a discrete finite random quantity $X \in \{x_1, x_2, \ldots, x_n\}$, in the subjective approach the conditional prevision assessment $\mu = \mathbb{P}(X|H)$ is the amount to be payed in order to receive the random quantity $X|H = XH + (1-H)\mu$. The random gain is $G = X|H - \mu = XH - \mu H$ and, as before, the coherence condition for μ is: $\min G|H \cdot \max G|H \leq 0$, which amounts to: $\min X|H \leq \mu \leq \max X|H$.
Of course, we have

$$\begin{aligned} \mathbb{P}(X|H) &= \mu = \mathbb{P}[XH + (1-H)\mu] = \\ &= \mathbb{P}(XH) + \mathbb{P}(1-H)\mu = \mathbb{P}(XH) + \mu - P(H)\mu\,, \end{aligned}$$

from which it follows the well known formula: $\mathbb{P}(XH) = P(H)\mu = P(H)\mathbb{P}(X|H)$.
More in general, given two events H, K and a random quantity X, let $\mathcal{M} = (x, y, z)$ a conditional prevision assessment on $\mathcal{F} = \{H|K, X|HK, XH|K\}$.
By the same kind of reasoning, we have

$$\begin{aligned} X|HK &= (X|H)|K = [XH + (1-H)y]|K = \\ &= XH|K + yH^c|K\,. \end{aligned} \tag{2}$$

In fact, as for the case of conditional events, we can show that the conditional random quantities $X|HK$ and $[XH + (1-H)y]|K$ coincide by the following remarks:
(i) we denote by $\{x_1, \ldots, x_n\}$ the set of possible values of X and, for the sake of simplicity by $\{x_1, \ldots, x_r\}$ (resp., $\{x_1, \ldots, x_r, \ldots, x_t\}$) the set of values of X compatible with HK (resp., with K), where $r \leq t \leq n$; moreover, we set $E_i = (X = x_i)$ and with the family $\mathcal{F}$ we associate the partition (of the sure event Ω) $\{E_1HK, \ldots, E_rHK, H^cK, K^c\}$;
(ii) we have $X = \sum_{i=1}^n x_iE_i$ and $XH = \sum_{i=1}^n x_iE_iH$; then

$$X|HK = \sum_{i=1}^r x_iE_iHK + (1-HK)y\,;$$

$$\begin{aligned} XH|K + yH^c|K = \sum_{i=1}^r x_iE_iHK + (1-K)z\, + \\ + yH^cK + (1-K)y(1-x)\,; \end{aligned}$$

(iii) assuming "K true", if H is true, then $X = x_i$ for some $i \leq r$ and $X|HK = [XH + (1-H)y]|K = x_i$; if H is false, then $X = x_i$ for some i, with $r < i \leq t$, and $X|HK = XH|K + yH^c|K = y$; hence, under the hypothesis "K true", $X|HK$ and $[XH + (1-H)y]|K$ coincide. Then

$$\begin{aligned} \mathbb{P}(X|HK) &= y = \mathbb{P}([XH + (1-H)y]|K) = \\ &= \mathbb{P}(XH|K) + yP(H^c|K) = z + y(1-x)\,, \end{aligned}$$

from which it follows: $z = xy$, that is:

$$\mathbb{P}(XH|K) = \mathbb{P}(X|H)P(H|K)\,.$$

Notice that, by the previous formula, if K is false we have $X|HK = y$ and

$$XH|K + yH^c|K = z + y(1-x) = y + z - xy = y\,.$$

Therefore, the conditional random quantities $X|HK$ and $XH|K + yH^c|K = (XH + yH^c)|K$ coincide in all cases.

4 General conditional random quantities

Let be given two random quantities X and Y. In [17] it is proposed the notion of general conditional random quantity $X|Y$ based on the following definition for the prevision of $X|Y$, introduced in [16].

Definition 1. The conditional prevision for X given Y, denoted $\mathbb{P}(X|Y)$, is a number you specify with the understanding that you accept to engage any transaction yielding a random net gain $G = sY[X - \mathbb{P}(X|Y)]$.

The following definition is given for the conditional random quantity $X|Y$.

Definition 2. Having asserted your conditional prevision $\mathbb{P}(X|Y) = \mu$, the conditional random quantity $X|Y$ is defined as

$$X|Y = XY + (1-Y)\mu = \mu + Y(X-\mu)\,. \qquad (3)$$

Notice that, if Y assumes only the value 0, that is $Y \equiv 0$, you can pay every real number $\mu = \mathbb{P}(X|Y)$, as you always receive the same amount μ; in fact, the net gain is always 0. To avoid this trivial case we will assume that $(Y = 0) \neq \Omega$.
We remark that such a general notion of conditional random quantity reduces to the classical one $X|H = XH + (1-H)\mu$ when Y coincides with an event H. Lad remarks that the direction of the net gain (or loss) depends on the difference $(X-\mu)$, while the scale depends on the numerical value of Y. Lad also remarks that for $Y = 0$ (resp., $Y = 1$) the net gain is 0 (resp., $s(X-\mu)$), i.e. the possible net gains obtained when Y is an event. Then, by computing the prevision on both sides of (3), Lad obtains

$$\mu = \mu + \mathbb{P}[Y(X-\mu)] = \mu + \mathbb{P}(XY) - \mu\mathbb{P}(Y)\,,$$

so that $\mathbb{P}(XY) = \mathbb{P}(X|Y)\mathbb{P}(Y)$, which becomes $\sum_j p_j y_j \mathbb{P}[X\,|\,(Y = y_j)] = \mathbb{P}(X|Y)\sum_j p_j y_j$, where $p_j = P(Y = y_j)$. This condition, which we call "generalized compound prevision theorem", generalizes the classical one $\mathbb{P}(XH) = \mathbb{P}(X|H)P(H)$, where H is an event. Then, when $\mathbb{P}(Y) \neq 0$ it immediately follows $\mathbb{P}(X|Y) = \frac{\mathbb{P}(XY)}{\mathbb{P}(Y)}$ (actually, we will see that the generalized compound prevision theorem holds in a *stronger* sense). Several properties are obtained by Lad, under the condition $\mathbb{P}(Y) \neq 0$. We also notice that, when X and Y are uncorrelated, i.e. $Cov(X,Y) = 0$, it is $\mathbb{P}(XY) = \mathbb{P}(X)\mathbb{P}(Y)$; then, under the hypothesis $\mathbb{P}(Y) \neq 0$, it follows $\mathbb{P}(X|Y) = \mathbb{P}(X)$. We can say that, *under the condition $\mathbb{P}(Y) \neq 0$, X and Y are uncorrelated if and only if the prevision of 'X given Y' coincides with the prevision of X.*
We examine below an example, in which Y is not an event, to illustrate a critical aspect.

Example 1. We recall that by the formula $\mathbb{P}(XH) = P(H)\mathbb{P}(X|H)$, when $P(H) > 0$ it follows $\mathbb{P}(X|H) = \frac{\mathbb{P}(XH)}{P(H)}$. Moreover, if $P(H) = 0$, then $\mathbb{P}(XH) = 0$; in this case, based on coherence principle ([15, 18]) and assuming $\emptyset \neq H \neq \Omega$, it can be proved that the assessment $(0, 0, \mu)$ on $\{H, XH, X|H\}$ is coherent if and only if: $\min X|H \leq \mu \leq \max X|H$. But, replacing H by a random quantity Y, we are in a *very different situation*, as $\mathbb{P}(Y) = 0$ *doesn't imply* $\mathbb{P}(XY) = 0$. To illustrate this aspect, let us consider a random vector

$$(X,Y) \in \mathcal{C} = \{(0,-1),(0,1),(1,-1),(1,1)\}\,,$$

with

$$p(0,-1) = \frac{1}{3},\ p(0,1) = \frac{1}{6},\ p(1,-1) = \frac{1}{6},\ p(1,1) = \frac{1}{3},$$

where $p(x,y) = P(X = x, Y = y)$. We denote the joint distribution of (X,Y) by the vector $(\frac{1}{3}, \frac{1}{6}, \frac{1}{6}, \frac{1}{3})$. We have

$$Y \in \mathcal{C}_Y = \{-1, 1\}\,,\ XY \in \mathcal{C}_{XY} = \{-1, 0, 1\}\,,$$

with $P(Y = -1) = P(Y = 1) = \frac{1}{2}$, and with $P(XY = -1) = \frac{1}{6}$, $P(XY = 0) = \frac{1}{2}$, $P(XY = 1) = \frac{1}{3}$, so that $\mathbb{P}(Y) = 0$ and $\mathbb{P}(XY) = \frac{1}{6}$. In this case, it doesn't exist any finite value $\mathbb{P}(X|Y)$ which satisfies the equality $\mathbb{P}(XY) = \mathbb{P}(X|Y)\mathbb{P}(Y)$. In fact, given any assessment $\mathbb{P}(X|Y) = \mu$, the values of $Y(X-\mu)$ associated with that of (X,Y) are, respectively, $\mu, -\mu, -1+\mu, 1-\mu$; then, assuming (for the sake of simplicity) $s = 1$, one has

$$\mathbb{P}(G) = \mathbb{P}[Y(X-\mu)] =$$
$$= \frac{1}{3}\mu + \frac{1}{6}(-\mu) + \frac{1}{6}(-1+\mu) + \frac{1}{3}(1-\mu) = \frac{1}{6} \neq 0\,,\ \ \forall\,\mu\,.$$

Hence, by starting with a joint probability distribution on (X,Y), it may happen that the equation $\mathbb{P}(XY) = \mathbb{P}(X|Y)\mathbb{P}(Y)$ has no finite solutions in the unknown $\mathbb{P}(X|Y)$.
If we assign the joint distribution $(\frac{1}{3}-\varepsilon, \frac{1}{6}+\varepsilon, \frac{1}{6}, \frac{1}{3})$ on (X,Y), with $\varepsilon \in [-\frac{1}{6}, 0) \cup (0, \frac{1}{3}]$, we obtains

$$P(Y = -1) = \frac{1}{2} - \varepsilon\,,\ P(Y = 1) = \frac{1}{2} + \varepsilon\,,\ \ \mathbb{P}(Y) = 2\varepsilon\,,$$

$$P(Y = -1) = \frac{1}{2} - \varepsilon\,,\ P(Y = 1) = \frac{1}{2} + \varepsilon\,,\ \ \mathbb{P}(Y) = 2\varepsilon\,,$$

while the distribution of XY doesn't change; moreover,

$$\mathbb{P}(G) = (\frac{1}{3}-\varepsilon)\mu + (\frac{1}{6}+\varepsilon)(-\mu) + \frac{1}{6}(-1+\mu) + \frac{1}{3}(1-\mu) =$$
$$= \frac{1}{6} - 2\varepsilon\mu = \mathbb{P}(XY) - \mathbb{P}(Y)\mathbb{P}(X|Y)\,,$$

and imposing $\mathbb{P}(G) = 0$, it follows

$$\mu = \mathbb{P}(X|Y) = \frac{1}{12\,\varepsilon}\,,\quad \varepsilon \in [-\frac{1}{6}, 0) \cup (0, \frac{1}{3}]\,.$$

In particular, for $\varepsilon \in [-\frac{1}{6}, 0)$ it is $\mu \in (-\infty, -\frac{1}{2}]$, while for $\varepsilon \in (0, \frac{1}{3}]$ it is $\mu \in [\frac{1}{4}, +\infty)$.
Finally, if we assign a uniform distribution on (X,Y), that is

$$p(0,-1) = p(0,1) = p(1,-1) = p(1,1) = \frac{1}{4}\,,$$

it follows $\mathbb{P}(Y) = \mathbb{P}(XY) = 0$; then, the equality $\mathbb{P}(XY) = \mathbb{P}(Y)\mathbb{P}(X|Y)$ becomes $0 = 0 \cdot \mathbb{P}(X|Y)$. In this case, we need a *direct* assessment of $\mathbb{P}(X|Y)$ and the problem of *coherence* arises. This basic problem will be addressed in the next section.

5 Coherence of general conditional prevision assessments

A crucial problem arises when $\mathbb{P}(Y)=0$; what can be said about coherence of a given assessment $\mathbb{P}(X|Y)=\mu$? We remark that this case has not been examined in the book of Lad. We also observe that when Y equals 0 Lad notices that the net gain is 0 without further comments. But, concerning the classical case of a conditional random quantity $X|H$, in order to check the coherence of the assessment $\mathbb{P}(X|H)=\mu$, as is well known the value 0 of the net gain associated with the case $H=0$ is discarded by the set of values of the net gain G, i.e. coherence checking is based on the values of $G|H$. Hence, in order to integrate the analysis of Lad by properly managing the case $\mathbb{P}(Y)=0$, we propose:
(i) to give an explicit definition of coherence for a given assessment $\mathbb{P}(X|Y)=\mu$;
(ii) to discard, in the definition of coherence, the value 0 of the net gain associated with the case $Y=0$.
Then, based on [15, 18], we give the following

Definition 3. Given two random quantities X, Y and a conditional prevision assessment $\mathbb{P}(X|Y)=\mu$, let $G=s(X|Y-\mu)=sY(X-\mu)$ be the net random gain, where s is an arbitrary real quantity, with $s\neq 0$. Defining the event $H=(Y\neq 0)$, the assessment $\mathbb{P}(X|Y)=\mu$ is coherent if and only if: $\inf G|H\cdot\sup G|H\leq 0$, for every s.

In what follows, without loss of generality, we will set $s=1$.

5.1 A strong generalized compound prevision theorem

Based on Definition 3, we will obtain a stronger version of the generalized compound prevision theorem. We recall that H is the event $(Y\neq 0)$; then, we make the following reasoning (where we assume that $\mu, \mathbb{P}(Y|H)$, and $\mathbb{P}(XY|H)$ are finite):
(i) by Definition 3, μ is the quantity to be payed, in order to receive $X|Y$, under the hypothesis H true; hence, *operatively* μ is the prevision of $X|Y$, *conditional on* H; (ii) hence, a more appropriate representation of $X|Y$ is given by: $X|Y=[\mu+Y(X-\mu)]|H$; (iii) then, by computing the prevision on both sides of the previous equality, we have:

$$\mu=\mathbb{P}(X|Y)=\mathbb{P}[\mu+Y(X-\mu)|H]=\mu+\mathbb{P}[Y(X-\mu)|H],$$

so that $\mathbb{P}[Y(X-\mu)|H]=\mathbb{P}[(XY-\mu Y)|H]=0$; then, by the linearity of prevision, it follows

$$\mathbb{P}(XY|H)=\mathbb{P}(X|Y)\mathbb{P}(Y|H). \tag{4}$$

Notice that, if Y is a finite discrete random quantity, with $Y\geq 0$, or $Y\leq 0$, surely it is $\mathbb{P}(Y|H)\neq 0$; then, by (4) it follows $\mathbb{P}(X|Y)=\frac{\mathbb{P}(XY|H)}{\mathbb{P}(Y|H)}$.
We recall that H^c is the event $(Y=0)$; moreover, we observe that $\mathbb{P}(Y|H^c)=\mathbb{P}(XY|H^c)=0$; hence,

$$\begin{aligned}\mathbb{P}(Y)&=\mathbb{P}(Y|H)P(H)+\mathbb{P}(Y|H^c)P(H^c)=\\&=\mathbb{P}(Y|H)P(H)=\mathbb{P}(YH),\end{aligned} \tag{5}$$

$$\begin{aligned}\mathbb{P}(XY)&=\mathbb{P}(XY|H)P(H)+\mathbb{P}(XY|H^c)P(H^c)=\\&=\mathbb{P}(XY|H)P(H)=\mathbb{P}(XYH).\end{aligned} \tag{6}$$

Then, by (4), (5), and (6), one has

$$\mathbb{P}(XY|H)P(H)=\mathbb{P}(X|Y)\mathbb{P}(Y|H)P(H),$$

that is, the formula $\mathbb{P}(XY)=\mathbb{P}(X|Y)\mathbb{P}(Y)$, given in [16] and [17], which we call *weak* generalized compound prevision theorem.

5.2 Some examples and remarks

In the finite case, denoting respectively by $\mathcal{C}_X, \mathcal{C}_Y$ and $\mathcal{C}$ the sets of possible values of X, Y and (X,Y), with each $(x_h, y_k)\in\mathcal{C}$ it is associated for the net gain G the value $g_{hk}=y_k(x_h-\mu)$. We set $\mathcal{C}_0=\{(x_h,y_k)\in\mathcal{C} : y_k\neq 0\}$; of course $\mathcal{C}_0\subseteq\mathcal{C}$. Then, by Definition 3, the assessment μ is coherent if and only if: $m\leq 0\leq M$, where

$$m=\min_{(x_h,y_k)\in\mathcal{C}_0} y_k(x_h-\mu),\quad M=\max_{(x_h,y_k)\in\mathcal{C}_0} y_k(x_h-\mu).$$

We denote by Π the set of coherent assessments μ; then, we remark that, assuming $\mathcal{C}_0\neq\emptyset$, the assessment $\mu=x_h$ is coherent, as it trivially satisfies the condition of coherence (it is $g_{hk}=0, \forall\,(x_h,y_k)\in\mathcal{C}_0$). Hence, $\mathcal{C}_X\subseteq\Pi$.

Example 2. Given a random vector $(X,Y)\in\mathcal{C}=\{(-1,0),(1,1)\}$, consider the assessment $\mathbb{P}(Y|X)=\mu$ on the conditional random quantity $X|Y$. We have $H=(Y\neq 0)$; hence $\mathcal{C}_0=\{(1,1)\}$. Moreover, one has $G=Y(X-\mu)\in\{0,1-\mu\}$, with $G|H=1-\mu$. We observe that Y coincides with the indicator of H, so that $X|Y=X|H$. Then, by Definition 3, μ is coherent if and only if $1-\mu=0$, that is $\mu=1$. Notice that this result is consistent with the usual approach to the notion of conditional prevision.

Remark 1. Notice that in Example 2, while the coherence condition $\inf G|H\cdot\sup G|H\leq 0$ is satisfied uniquely with $\mu=1$, the condition $\inf G\cdot\sup G\leq 0$ is satisfied for every μ. Then, if the condition $\inf G|H\cdot\sup G|H\leq 0$ were replaced by $\inf G\cdot\sup G\leq 0$, it would follow that every assessment $\mathbb{P}(X|Y)=\mu$ would be coherent, which is clearly unreasonable (however, as we will show by other examples, still applying the condition $\inf G|H\cdot\sup G|H\leq$

0, it may be $\Pi = \mathbb{R}$). Example 2 confirms that, in order to look at $X|Y$ as $X|H$ in the usual sense, when checking coherence we must discard the value 0 of the random gain G associated with the case $Y = 0$. In this way, we can look at the family of conditional random quantities like $X|H$, where H is an event, as a sub-family of the family of general conditional random quantities like $X|Y$, where Y is a random quantity.

We recall that, given any event $H \neq \emptyset$, if X is a constant, say $X = c$, then $\mathbb{P}(X|H) = c$. The following example shows that, if $X = c$ and Y is a random quantity, with $\min Y < 0 < \max Y$, then the assessment $\mathbb{P}(X|Y) = \mu$ is coherent for every $\mu \in \mathbb{R}$.

Example 3. Given $(X, Y) \in \mathcal{C} = \{(c, -y_1), (c, y_2)\}$, with $c \in \mathbb{R}$ and $y_1, y_2 > 0$, consider the coherence of any assessment $\mathbb{P}(X|Y) = \mu$. We have $\mathcal{C}_0 = \mathcal{C}$, so that $H = (Y \neq 0) = \Omega$ and $G|H = G = Y(c - \mu)$. The values of $G|H$ are: $-y_1(c-\mu), y_2(c-\mu)$, and the coherence condition $\inf G|H \cdot \sup G|H \leq 0$ is satisfied for every $\mu \in \mathbb{R}$. Moreover, given a joint distribution on (X, Y), say $(p, 1-p)$, where

$$p = P(X = c, Y = -y_1)\,, \quad 1-p = P(X = c, Y = y_2),\,,$$

with $0 \leq p \leq 1$, we have $\mathbb{P}(Y) = y_2 - p(y_1 + y_2)$ and

$$\mathbb{P}(XY) = c\mathbb{P}(Y) = c[y_2 - p(y_1 + y_2)]\,.$$

Then, if $p \neq \frac{y_2}{y_1+y_2}$, one has $\mathbb{P}(Y) \neq 0$ and c is the unique coherent value of μ associated with the distribution $(p, 1-p)$. Whereas, if $p = \frac{y_2}{y_1+y_2}$, then $\mathbb{P}(Y) = \mathbb{P}(XY) = 0$, and the assessment $\mathbb{P}(X|Y) = \mu$, associated with the distribution $(\frac{y_2}{y_1+y_2}, \frac{y_1}{y_1+y_2})$, is coherent for every $\mu \in \mathbb{R}$.

Example 4. We continue the study of Example 1, by examining the coherence of a given assessment $\mathbb{P}(X|Y) = \mu$. We recall that $(X, Y) \in \mathcal{C} = \{(0, -1), (0, 1), (1, -1), (1, 1)\}$; moreover, we observe that $\mathcal{C}_0 = \mathcal{C}$, as $H = (Y \neq 0) = \Omega$ and hence $G|H = G = Y(X - \mu)$. With the values of (X, Y) are associated respectively the following values of $G|H$: $\mu, -\mu, -1 + \mu, 1 - \mu$; hence, the coherence condition $\inf G|H \cdot \sup G|H \leq 0$ is satisfied for every μ.

Example 5. We assume that $(X, Y) \in \mathcal{C} = \{(0, -1), (1, 1)\}$, by examining the coherence of a given assessment $\mathbb{P}(X|Y) = \mu$. We have $\mathcal{C}_0 = \mathcal{C}$; so that $H = (Y \neq 0) = \Omega$ and we have $G|H = G = Y(X - \mu)$. The values of $G|H$ are: $\mu, 1 - \mu$ and, as it can be verified, the coherence condition $\inf G|H \cdot \sup G|H \leq 0$ is satisfied if and only if $\mu \notin (0, 1)$, that is μ is coherent if and only if $\mu \in (-\infty, 0] \cup [1, +\infty)$. In this example with each coherent assessment μ it is associated a unique joint distribution on (X, Y), say $(p, 1-p)$, where

$$p = P(X = 0, Y = -1)\,,$$

$$1 - p = P(X = 1, Y = 1)\,, \quad 0 \leq p \leq 1\,.$$

The parameter p is determined by requiring that the prevision of the random gain be 0, that is

$$p\mu + (1 - p)(1 - \mu) = 0\,. \tag{7}$$

As it can be verified, one has

$$p = f(\mu) = \frac{1 - \mu}{1 - 2\mu}\,;$$

moreover, when $\mu \leq 0$ it is $\frac{1}{2} < p \leq 1$; when $\mu \geq 1$ it is $0 \leq p \leq \frac{1}{2}$. Notice that

$$\mu = f^{-1}(p) = \frac{1 - p}{1 - 2p}\,;$$

that is: $f^{-1} = f$. This result depends on the symmetry of the equation (7) with respect to p and μ.

As shown by Example 5, the set Π *of the coherent assessments* μ *may be not convex.*
To better analyze this aspect, in what follows we examine separately two cases:
(i) $Y \geq 0$, or $Y \leq 0$; (ii) $\min Y < 0 < \max Y$.

6 The case $Y \geq 0$, or $Y \leq 0$.

We assume $X \in \mathcal{C}_X = \{x_1, \ldots, x_n\}$ and $Y \in \mathcal{C}_Y = \{y_1, \ldots, y_r\}$, with $y_k \geq 0\,, \forall\, k$. Moreover, we denote by X^0 the subset of $\mathcal{C}_X$ such that for each $x_h \in X^0$ there exists $(x_h, y_k) \in \mathcal{C}_0$. Then, we set

$$x_0 = \min X^0\,, \quad x^0 = \max X^0\,. \tag{8}$$

We first consider the case $Y \geq 0$; we have

Theorem 1. Given two finite random quantities X, Y, with $Y \geq 0$, the prevision assessment $\mathbb{P}(X|Y) = \mu$ is coherent if and only if $x_0 \leq \mu \leq x^0$.

Proof. Given any μ, with each pair $(x_h, y_k) \in \mathcal{C}_0$ we associate the inequality $y_k(x_h - \mu) \geq 0$. Under the hypothesis $Y \neq 0$ it is $y_k > 0$; then the inequality is satisfied if and only if $\mu \leq x_h$. We observe that, for each $x_h \in X^0$, there exists (at least) a value $y_k > 0$ such that $(x_h, y_k) \in \mathcal{C}_0$. Then, we distinguish three cases: (i) $\mu < x_0$; (ii) $\mu > x^0$; (iii) $x_0 \leq \mu \leq x^0$. In the first case it is $y_k(x_h - \mu) > 0$ for every $(x_h, y_k) \in \mathcal{C}_0$, so that $\inf G|H \cdot \sup G|H > 0$ and hence μ is not coherent. In the second case it is $y_k(x_h - \mu) < 0$ for every $(x_h, y_k) \in \mathcal{C}_0$, so that $\inf G|H \cdot \sup G|H > 0$ and hence μ is not coherent. In the third case, denoting by y_k and y_s two positive values of Y such that $(x_0, y_k) \in \mathcal{C}_0$, $(x^0, y_s) \in \mathcal{C}_0$, it is $y_k(x_0 - \mu) \leq 0\,,$ $y_s(x^0 - \mu) \geq 0$, so that $\inf G|H \leq 0\,,$ $\sup G|H \geq 0$ and hence $\inf G|H \cdot \sup G|H \leq 0$. Therefore, for every $\mu \in [x_0, x^0]$, μ is coherent. □

We illustrate the previous result by the following

Example 6. Given a random vector $(X,Y) \in \mathcal{C} = \{(0,1),(1,0),(1,1),(2,2)\}$, let us determine the set Π of coherent prevision assessment $\mathbb{P}(X|Y) = \mu$ on $X|Y$. We observe that $X^0 = X$, so that $x_0 = \min \mathcal{C}_X = 0$, $x^0 = \max \mathcal{C}_X = 2$; moreover, it is $\mathcal{C}_0 = \{(0,1),(1,1),(2,2)\}$ and the values of $Y(X-\mu)$, under the restriction $(X,Y) \in \mathcal{C}_0$ are, respectively, $-\mu, 1-\mu, 2(2-\mu)$; such values are all positive (resp., all negative) when $\mu < 0$ (resp., $\mu > 2$); hence each $\mu \notin [0,2]$ is not coherent. Finally, when $\mu \in [0,2]$ one has $-\mu(2-\mu) \leq 0$, so that the condition $\inf G|H \cdot \sup G|H \leq 0$ is satisfied. Hence, we have $\Pi = [x_0, x^0] = [0,2]$.

We now consider the case $Y \leq 0$; we have

Theorem 2. Given two finite random quantities X, Y, with $Y \leq 0$, the conditional prevision assessment $\mathbb{P}(X|Y) = \mu$ is coherent if and only if $x_0 \leq \mu \leq x^0$.

Proof. We observe that, as $-Y \geq 0$, by Theorem 1 the assessment $\mathbb{P}(X|-Y) = \mu$ is coherent if and only if $x_0 \leq \mu \leq x^0$. On the other hand, defining $G'|H = -Y(X-\mu)|H$, we have $G|H = Y(X-\mu) = -G'|H$. Then

$$\inf G|H = -\sup G'|H\,, \quad \sup G|H = -\inf G'|H\,,$$

and hence: $\inf G|H \cdot \sup G|H = \inf G'|H \cdot \sup G'|H$; thus, the assessment $\mathbb{P}(X|H) = \mu$ is coherent if and only if $x_0 \leq \mu \leq x^0$. $\square$

7 The case $\min Y < 0 < \max Y$.

We now examine the general case in which there exist positive and negative values of Y. We set

$$X^- = \{x_h \in \mathcal{C}_X : \exists (x_h, y_k) \in \mathcal{C}_0, y_k < 0\}\,,$$

$$X^+ = \{x_h \in \mathcal{C}_X : \exists (x_h, y_k) \in \mathcal{C}_0, y_k > 0\}\,;$$

$$\mathcal{C}^- = \{(x_h, y_k) \in \mathcal{C}_0 : y_k < 0\}\,,$$

$$\mathcal{C}^+ = \{(x_h, y_k) \in \mathcal{C}_0 : y_k > 0\}\,.$$

Of course, $\mathcal{C}^- \cap \mathcal{C}^+ = \emptyset$ and $\mathcal{C}^- \cup \mathcal{C}^+ = \mathcal{C}_0$. We have

Theorem 3. Let be given two random quantities X, Y, with $\min Y < 0 < \max Y$. If $X^- \cap X^+ \neq \emptyset$, then the conditional prevision assessment $\mathbb{P}(X|Y) = \mu$ is coherent, for every real number μ.

Proof. Let be given $x_h \in X^- \cap X^+, y_k \in \mathcal{C}_Y, y_t \in \mathcal{C}_Y$ such that $(x_h, y_k) \in \mathcal{C}^-$ and $(x_h, y_t) \in \mathcal{C}^+$; moreover, let μ be any real number. It is $g_{hk}g_{ht} = y_k(x_h - \mu) \cdot y_t(x_h - \mu) = y_k y_t (x_h - \mu)^2 \leq 0$, so that $\inf G|H \cdot \sup G|H \leq 0$. Therefore, for every $\mu \in \Re$, μ is coherent. $\square$

We illustrate the previous result by the following

Example 7. We determine the set Π of coherent prevision assessment $\mathbb{P}(X|Y) = \mu$ on $X|Y$, where $(X,Y) \in \mathcal{C} = \{(0,1),(0,-1),(1,-1),(1,1)\}$, as in Example 1. We have $X^- = X^+ = \{0,1\}$, so that $X^- \cap X^+ \neq \emptyset$; hence, by Theorem 3, $\Pi = \Re$.

In what follows, we examine the cases

$$\min X^- = \max X^+\,, \;\; \max X^- = \min X^+\,;$$

then, we study in depth the case $X^- \cap X^+ = \emptyset$. Given any $(x_h, y_s) \in \mathcal{C}^-$, $(x_k, y_t) \in \mathcal{C}^+$, we set

$$m_{hk} = \min\{x_h, x_k\}\,, \quad M_{hk} = \max\{x_h, x_k\}\,;$$

moreover, we denote by I_{hk} the open interval (m_{hk}, M_{hk}). Then, we set

$$I = \bigcap_{x_h \in X^-,\, x_k \in X^+} I_{hk}\,. \tag{9}$$

Notice that, defining

$$\begin{aligned} \mu_0 &= \max_{x_h \in X^-,\, x_k \in X^+} \; m_{hk}\,, \\ \mu^0 &= \min_{x_h \in X^-,\, x_k \in X^+} \; M_{hk}\,, \end{aligned} \tag{10}$$

one has $I \neq \emptyset$ if and only if $\mu_0 < \mu^0$ and, in this case, $I = (\mu_0, \mu^0)$. We have

Theorem 4. Let the quantities μ_0, μ^0 be defined as in (10); then $\mu_0 = \min(\max X^-, \max X^+)$ and $\mu^0 = \max(\min X^-, \min X^+)$.

Proof. We first prove that μ_0 coincides with $\min(\max X^-, \max X^+)$. Let be $x_k = \max X^+$, $x_h = \max X^-$. Then $x_r \leq x_k, \forall x_r \in X^+$ and $x_t \leq x_h, \forall x_t \in X^-$. Let be $\min(\max X^-, \max X^+) = x_h$. Then, there exists $x_r \in X^+$ such that $x_r \geq x_h$, i.e. there exist $(x_h, y_s) \in C^-$ and $(x_r, y_t) \in C^+$, such that $m_{hr} = x_h$. Suppose that $\mu_0 \neq x_h$, i.e. $\mu_0 \neq \min(\max X^-, \max X^+)$; then $\mu_0 > x_h$ and, as $x_h = \max X^-$, it must be $\mu_0 = x_t$ for some $x_t \in X^+$. Then, there exist $(x_v, y_r) \in C^-$, $(x_t, y_s) \in C^+$ such that $x_t \leq x_v$. From $x_v \leq x_h$, it is $x_t \leq x_v \leq x_h$, i.e. $\mu_0 \leq x_h$, which is absurd; hence $\mu_0 = \min(\max X^-, \max X^+)$. The proof is similar if $\min(\max X^-, \max X^+) = x_k$, where $x_k = \max X^+$. We now prove that $\mu^0 = \max(\min X^-, \min X^+)$. Let be $x_k = \min X^+$, $x_h = \min X^-$. Then $x_r \geq x_k, \forall x_r \in X^+$ and $x_t \geq x_h, \forall x_t \in X^-$. Let be $\max(\min X^-, \min X^+) = x_h$. Then, there exists $x_r \in X^+$ such that $x_r \leq x_h$, i.e. there exist $(x_h, y_s) \in C^-$ and $(x_r, y_t) \in C^+$, such that $M_{hr} = x_h$. Suppose that $\mu^0 \neq x_h$, i.e. $\mu^0 \neq \max(\min X^-, \min X^+)$; then $\mu^0 < x_h$ and, as $x_h = \min X^-$, it must be $\mu^0 = x_t$ for some $x_t \in X^+$. Then, there exist $(x_v, y_r) \in C^-$,

$(x_t, y_s) \in C^+$ such that $x_t \geq x_v$. From $x_v \geq x_h$, it is $x_t \geq x_v \geq x_h$, i.e. $\mu^0 \geq x_h$, which is absurd; hence $\mu^0 = \max(\min X^-, \min X^+)$.
The proof is similar if $\max(\min X^-, \min X^+) = x_k$, where $x_k = \min X^+$. □

Thus, if $\mu_0 < \mu^0$, it is $I = (\mu_0, \mu^0) = (\min(\max X^-, \max X^+),\ \max(\min X^-, \min X^+))$. We set $X^- < X^+$ (resp., $X^- > X^+$) if and only if $\max X^- < \min X^+$ (resp., $\min X^- > \max X^+$), otherwise we set $X^- \nsim X^+$. We have

Theorem 5. $I \neq \emptyset$ if and only if $X^- < X^+$, or $X^- > X^+$.

Proof. Obviously, $I \neq \emptyset$ if and only if $\mu_0 < \mu^0$. We prove that $\mu_0 \geq \mu^0$ if and only if $X^- \nsim X^+$. Such a situation happens if and only if $\mu_0 \in X^-$ and $\mu^0 \in X^-$ or $\mu_0 \in X^+$ and $\mu^0 \in X^+$. Suppose that $\mu_0 = x_h \in X^+$ and $\mu^0 = x_k \in X^+$. It is $\mu_0 = \max X^+$ and $\mu^0 = \min X^+$. From $\mu_0 = \min(\max X^-, \max X^+)$, there exists $x_s \in X^-$ such that $x_s \geq x_h$ and, from $\mu^0 = \max(\min X^-, \min X^+)]$, there exists $x_t \in X^-$ such that $x_t \leq x_k$, that is $X^- \nsim X^+$. Moreover, from $\mu_0 = \max X^+$, $\mu^0 = \min X^+$, it is $\mu_0 \geq \mu^0$ and $I = \emptyset$. If we suppose that $\mu_0 = x_h \in X^-$ and $\mu^0 = x_k \in X^-$, by a similar reasoning, we have that $X^- \nsim X^+$ and $\mu_0 > \mu^0$ so that $I = \emptyset$.
Suppose that $I \neq \emptyset$ that is $\mu_0 < \mu^0$. Thus, $\mu_0 = x_k \in X^+$ and $\mu^0 = x_h \in X^-$ or $\mu_0 = x_h \in X^-$ and $\mu^0 = x_k \in X^+$. In the first case it is $X^+ < X^-$, in the other case it is $X^+ < X^-$. Conversely, if $X^+ < X^-$, it is $\max X^+ < \max X^-$ and $\mu_0 = \max X^+$. Moreover, it is $\min X^+ < \min X^-$ and $\mu^0 = \min X^-$, with $\mu_0 < \mu^0$. If, $X^+ > X^-$ it is $\max X^+ > \max X^-$ and $\mu_0 = \max X^-$. Moreover, it is $\min X^+ > \min X^-$ and $\mu^0 = \min X^+$, with $\mu_0 < \mu^0$. □

Based on the previous result, we have the following three cases

1. $X^+ < X^- \Leftrightarrow I \neq \emptyset$ and $I = (\mu_0, \mu^0)$, with $\mu_0 = \max X^+$, $\mu^0 = \min X^-$.
2. $X^+ > X^- \Leftrightarrow I \neq \emptyset$ and $I = (\mu_0, \mu^0)$, with $\mu_0 = \max X^-$, $\mu^0 = \min X^+$.
3. $X^- \nsim X^+ \Leftrightarrow I = \emptyset$.

We have

Theorem 6. Let be given two random quantities X, Y, with $\min Y < 0 < \max Y$. If case 1, or case 2, holds, then $X^- \cap X^+ = \emptyset$ and the conditional prevision assessment $\mathbb{P}(X|Y) = \mu$ is coherent if and only if $\mu \notin I$. In the case 3, the assessment $\mathbb{P}(X|Y) = \mu$ is coherent for every real number μ.

Proof. Case 1. Suppose $\mu \leq \mu_0$. We prove that μ is coherent. It is $\mu \leq \mu_0 = \max X^+ < \min X^- = \mu^0$. Let $x_h \in X^+$, that is there exist $(x_h, y_s) \in C^+$ and $y_s > 0$. It is $x_h - \mu \geq 0$, then $g_{hs} = y_s(x_h - \mu) \geq 0$. Let $x_k \in X^-$, that is there exist $(x_k, y_t) \in C^-$ and $y_s < 0$. It is $x_k - \mu > 0$, then $g_{kt} = y_t(x_k - \mu) < 0$. It follows $\inf G|H \cdot \sup G|H \leq 0$, that is μ is coherent. By a similar reasoning, if $\mu \geq \mu^0$ it follows that μ is coherent.
Conversely, we prove that, if $\mu_0 = \max X^+ < \mu < \min X^- = \mu^0$, μ is not coherent. From $X^+ < X^-$, it is $x_k \leq \mu_0 < \mu < \mu^0 \leq x_h$ for each $x_h \in X^-$,$x_k \in X^+$. Hence, we have that for each $(x_h, y_s) \in \mathcal{C}^-$ one has $g_{hs} = y_s(x_h - \mu) < 0$, as $y_s < 0$ and $x_h - \mu \geq \mu_0 - \mu > 0$; moreover, for each $(x_k, y_t) \in \mathcal{C}^+$ one has $g_{kt} = y_t(x_k - \mu) < 0$, as $y_t > 0$ and $x_k - \mu \leq \mu^0 - \mu < 0$. Hence, for every $(x_h, y_k) \in \mathcal{C}$, it is $g_{hk} = y_k(x_h - \mu) < 0$. Then $\inf G|H \cdot \sup G|H > 0$, that is μ is not coherent.
Case 2. The proof is formally identical to the case 1.
Case 3. There exist $(x_h, y_t) \in C^-$, $(x_k, y_s) \in C^+$, $(x_u, y_r) \in C^-$, $(x_v, y_z) \in C^+$, such that $x_h < x_k$ and $x_u > x_v$. Let μ be a real number. Suppose that $g_{ht} = y_t(x_h - \mu) < 0$. Then, $(x_h - \mu) > 0$ and $(x_k - \mu) > 0$, hence $g_{ks} = y_s(x_h - \mu) > 0$ and μ is coherent.
Suppose that $g_{ht} = y_t(x_h - \mu) > 0$. Then, $(x_h - \mu) < 0$. Thus, suppose that $(x_k - \mu) > 0$. It is $x_h < \mu < h_k$. By absurd, suppose that $g_{ur} = y_r(x_u - \mu) > 0$ and $g_{vz} = y_z(x_v - \mu) > 0$. Thus, it is $x_u - \mu < 0$ and $x_v - \mu > 0$, that is $x_u < \mu < x_v$ and $x_u < x_v$, which is absurd, as $x_u > x_v$. Then, μ is coherent. □

Remark 2. We observe that Theorem 3 is a particular case of Theorem 6, as $X^- \cap X^+ \neq \emptyset$ implies $X^- \nsim X^+$.

We say that $X^+ \leq X^-$ if $\max X^+ = \min X^-$, and $X^+ \geq X^-$ if $\min X^+ = \max X^-$.
From the previous results, we can summarize the case $\min Y < 0 < \max Y > 0$ in the following way

- $X^+ < X^- \Leftrightarrow \mu_0 = \max X^+ < \min X^- = \mu^0$. Then μ is coherent if and only if $\mu \leq \mu_0$ or $\mu \geq \mu^0$.
- $X^+ > X^- \Leftrightarrow \mu_0 = \max X^- < \min X^+ = \mu^0$. Then μ is coherent if and only if $\mu \leq \mu_0$ or $\mu \geq \mu^0$.
- $X^- \nsim X^+$. If $X^+ \leq X^-$ or $X^+ \geq X^-$, then $\mu_0 = \mu^0$, otherwise $\mu_0 > \mu^0$ and in all such cases every real number μ is coherent.

We illustrate the previous result by the example below.

Example 8. We determine the set Π of coherent prevision assessment $\mathbb{P}(X|Y) = \mu$ on $X|Y$, where $(X,Y) \in \mathcal{C} = \{(0,1),(0,2),(1,-1),(1,-2)\}$. We have $X^- = \{1\}$, $X^+ = \{0\}$, so that $X^- \cap X^+ = \emptyset$; we have to consider a unique case: $x_h = 1, x_k = 0$, with the associated open interval $I_{hk} = (0,1)$. Then, $I = I_{hk} = (0,1)$ and, by Theorem 6, $\Pi = \Re \setminus (0,1)$; that is, μ is coherent if and only if $\mu \notin (0,1)$. The same result follows directly, by observing that: (i) $\mathcal{C}_0 = \mathcal{C}$, so that $G|H = G$;
(ii) given any μ, the values of G are: $-\mu, -2\mu, -1+\mu, -2+2\mu$; (iii) if $\mu \in (0,1)$, the values of G are all negative; if $\mu \notin (0,1)$, it is: $\min G < 0$, $\max G > 0$.

8 Linear transformations of Y.

In this section we examine the effect produced on the set Π (of coherent conditional prevision assessments on $X|Y$) by a linear transformation on the *conditioning* random quantity Y. Given two random quantities X, Y and two constants c, d, with $(c,d) \neq (0,0)$, we set $y_0 = min\, Y$, $y^0 = max\, Y$, $Y' = cY + d$ and, if $c \neq 0$, $Y^* = Y + \frac{d}{c}$,; moreover, we denote by Π' (resp., Π^*) the set of coherent prevision assessments on $X|Y' = X|(cY+d)$ (resp., $X|Y^* = X|(Y + \frac{d}{c})$). We show below, among other things, that: (a) for $d \neq 0$ both cases $\Pi^* = \Pi$, or $\Pi^* \neq \Pi$, are possible; (b) $\Pi' = \Pi^*$.

Theorem 7. Given two finite random quantities X, Y and two constants c, d, with $(c,d) \neq (0,0)$, we have:
1. if $c = 0, d \neq 0$, then $\mathbb{P}(X|Y') = \mathbb{P}(X|d) = \mathbb{P}(X)$ and $\Pi' = [min\, X, max\, X]$;
2. if $c \neq 0$, $\frac{d}{c} \notin (-y^0, -y_0)$, then $\Pi^* = [x_0, x^0]$, where the values x_0, x^0 are defined as in (8) with Y replaced by Y^*;
3. if $c \neq 0$, $\frac{d}{c} \in (-y^0, -y_0)$, then $\Pi^* = \mathbb{R} \setminus I$, where the (possibly empty) interval I is defined as in (9) with Y replaced by Y^*;
4. $\Pi' = \Pi^*$.

Proof. In case 1 it is $G = d(X - \mu)$; then, under coherence of $\mathbb{P}(X)$, from $\mathbb{P}(G) = 0$ it follows $\mu = \mathbb{P}(X) \in [min\, X, max\, X]$. In case 2, it is $Y^* \geq 0$, when $\frac{d}{c} \geq -y_0$, and $Y^* \leq 0$, when $\frac{d}{c} \leq -y^0$; then, by Theorems 1 and 2, it follows $\Pi^* = [x_0, x^0]$. In case 3, as $-y^0 \leq \frac{d}{c} \leq -y_0$, it is $min\, Y^* < 0 < max\, Y^*$; then, by Theorem 6, one has $\Pi^* = \mathbb{R} \setminus I$, with the interval I possibly empty.
In case 4 it is $Y' = cY^*$ and, denoting by G' (resp., G^*) the random gain associated with $X|Y'$ (resp., $X|Y^*$), we have $G' = cY^*(X - \mu) = cG^*$. Then

$$\inf G'|H \cdot \sup G'|H = c^2 \inf G^*|H \cdot \sup G^*|H\,,$$

and, being $c^2 \neq 0$, the assessment $\mathbb{P}(X|Y') = \mathbb{P}(X|cY^*) = \mu$ is coherent if and only if $\mathbb{P}(X|Y^*) = \mu$ is coherent; thus $\Pi' = \Pi^*$. □

We give below an example where $\Pi^* \neq \Pi$.

Example 9. As in Example 6, we consider the random vector $(X,Y) \in \mathcal{C} = \{(0,1),(1,0),(1,1),(2,2)\}$. We recall that $\Pi = [0,2]$. Given $Y' = 2Y - 2 = 2Y^*$, where $Y^* = Y - 1$, let us determine the set $\Pi' = \Pi^*$. It is

$$(X, Y^*) \in \mathcal{C}^* = \{(0,0),(1,-1),(1,0),(2,1)\}$$

$$X^{*-} = \{1\}\,,\quad X^{*+} = \{2\}\,,\quad X^{*-} \cap X^{*+} = \emptyset\,.$$

Then: $X^{*-} < X^{*+}$, $\mu_0 = 1$, $\mu^0 = 2$, and we have $I = (1,2)$; moreover

$$y_0 = 0\,,\quad y^0 = 2\,,\quad \frac{d}{c} = -1 \in (-2,0) = (-y^0, -y_0)\,.$$

Then, by Theorem 7, case 3, we obtain

$$\Pi' = \Pi^* = (-\infty, 1) \cup (2, +\infty) = \mathbb{R} \setminus (1,2) \neq \Pi\,.$$

9 Conclusions

In this paper, recalling a general discussion on iterated conditioning given by de Finetti in his book, vol. 2, Appendix, section 13, we have given a representation of a conditional random quantity $X|HK$ as $(X|H)|K$. In this way, we have obtained the classical formula $\mathbb{P}(XH|K) = \mathbb{P}(X|H)P(H|K)$, by simply using linearity of prevision. Then, we have considered the notion of general conditional prevision $\mathbb{P}(X|Y)$, where X and Y are two random quantities, introduced in 1990 in a paper by Lad and Dickey, also discussed by Lad in his book published in 1996. After recalling the case where Y is an event, we have considered the case of discrete finite random quantities and we made some critical comments and examples. We have given a notion of coherence for such more general conditional prevision assessments; then, we have obtained a strong generalized compound prevision theorem. We have studied the coherence of a general conditional prevision assessment $\mathbb{P}(X|Y)$ when Y has no negative values and when Y has no positive values. We gave some results concerning the set of coherent conditional prevision assessments of $X|Y'$, where Y' is a linear transformation of Y. Finally, we have given some results on coherence of $\mathbb{P}(X|Y)$ when Y assumes both positive and negative values. To better illustrate some critical points and remarks we have also examined several examples. Future research more in general should concern: (i) the coherence of a conditional prevision assessment $\mathcal{A} = (\mu_1, \ldots, \mu_n)$ on a family of n conditional random quantities $\mathcal{F} = \{X_1|Y_1, \ldots, X_n|Y_n\}$; (ii) the generalized coherence of imprecise conditional prevision assessments, for instance an interval-valued assessment $\mathcal{A} = ([l_1, u_1], \ldots, [l_n, u_n])$, on $\mathcal{F}$.

Acknowledgements
We are grateful to the anonymous referees for their very useful criticisms and suggestions.

References

[1] Baioletti M., and Capotorti A., A comparison between classical logic and three-valued logic for conditional events, *Proc. of IPMU'96*, Granada, July 1-5, 1996, Vol. 3, pp. 1217-1221.

[2] Biazzo V., and Gilio A., A generalization of the fundamental theorem of de Finetti for imprecise conditional probability assessments, *International Journal of Approximate Reasoning,* 24, 251-272, 2000.

[3] Biazzo V., and Gilio A., On the linear structure of betting criterion and the checking of coherence, *Annals of Mathematics and Artificial Intelligence,* 35, 83-106, 2002.

[4] Biazzo V., and Gilio A., Some results on imprecise conditional prevision assessments, *Proc. of ISIPTA'07*, Prague, Czech Republic, July 16-19, 2007, pp. 31-40.

[5] Biazzo V., Gilio A., and Sanfilippo G., Coherence Checking and Propagation of Lower Probability Bounds, *Soft Computing,* 7, 310-320, 2003.

[6] Biazzo V., Gilio A., and Sanfilippo G., Generalized coherence and connection property of imprecise conditional previsions, *Proc. of IPMU'08*, Malaga, Spain, June 22 - 27, 2008, pp. 907-914.

[7] Capotorti A., and Paneni T., An operational view of coherent conditional previsions, *Proc. of EC-SQARU'01*, Toulouse, France, September 19-21, 2001, pp. 132-143.

[8] Capotorti A., and Vantaggi B., Locally strong coherence in inference processes, *Annals of Mathematics and Artificial Intelligence,* 35, 125-149, 2002.

[9] Coletti G., and Scozzafava R., *Probabilistic logic in a coherent setting*, Kluwer Academic Publishers, 2002.

[10] de Cooman G., and Miranda E., Marginal extension in the theory of coherent lower previsions, *International Journal of Approximate Reasoning*, 46, 188-225, 2007.

[11] de Finetti B. (1935), La logique de la probabilité, *Actes du Congrès International de Philosophie Scientifique*, Paris 1935, Hermann: IV, 1-9, 1936.

[12] de Finetti B. (1970), *Teoria delle probabilità*, voll. 1-2, Einaudi, Torino, 1970 (Engl. transl.: *Theory of Probability*, voll. 1-2, Wiley, Chichester, 1974).

[13] Gilio A., Criterio di penalizzazione e condizioni di coerenza nella valutazione soggettiva della probabilità, *Bollettino Un. Matem. Ital.*, [7a] 4-B(3): 645-660, 1990.

[14] Gilio A., Incomplete probability assessments in decision analysis, *J. It. Statist. Soc.*, Vol. 1, N. 1, 67-76, 1992.

[15] Holzer S., On coherence and conditional prevision, *Boll. Un. Mat. Ital.* 4 (6), 441-460, 1985.

[16] Lad F., and Dickey J. M., A general theory of conditional prevision, $P(X|Y)$, and the problem of state-dependent preferences, *Economic Decision-Making: Games, Econometrics and Optimization, Essays in Honor of Jaques Dreze*, J.J. Gabsewicz, J.F. Richard, and L.A. Wolsey (eds.), Amsterdam: North Holland, 369-383, 1990.

[17] Lad F., *Operational Subjective Statistical Methods: A Mathematical, Philosophical, and Historical Introduction*. New York, Wiley, 1996.

[18] Regazzini E., Finitely additive conditional probabilities, *Rend. Sem. Mat. Fis. Milano* 55, 69-89, 1985.

[19] Vicig P., Zaffalon M., and Cozman F.G., Notes on "Notes on conditional previsions", *International Journal of Approximate Reasoning*, 44, 358-365, 2007.

[20] Walley P., *Statistical reasoning with imprecise probabilities*, Chapman and Hall, London, 1991.

[21] Walley P., Pelessoni R., and Vicig P., Direct Algorithms for Checking Coherence and Making Inferences from Conditional Probability Assessments, *Journal of Statistical Planning and Inference*, 126(1), 119-151, 2004.

[22] Williams P.M., Notes on conditional previsions, Technical report, University of Sussex, 1975. Reprinted in a revised form in: *International Journal of Approximate Reasoning*, 44(3):366-383, 2007.