On general conditional random quantities

Veronica Biazzo, Angelo Gilio, Giuseppe Sanfilippo

vbiazzo@dmi.unict.it, gilio@dmmm.uniroma1.it, sanfilippo@unipa.it

1. Outline

- *We consider the notion of* general conditional prevision *of the form* $\mathbb{P}(X|Y)$*, where both* X *and* Y *are random quantities, introduced in (Lad and Dickey, 1990).*
- *We integrate the analysis of Lad and Dickey by properly managing the case* $\mathbb{P}(Y) = 0$
- *We propose a definition of coherence for the conditional prevision of 'X given Y'*
- *We obtain some results on coherence of a conditional prevision assessment* $\mathbb{P}(X|Y) = \mu$ *in the finite case*

2. Basic notions

In the setting of coherence, given any r. q. X and any events E, H, with $P(E|H) = p$ and $\mathbb{P}(X|H) = \mu$, if you pay p (resp., μ) you receive $E|H$ (resp., $X|H$); then, *operatively*, it is

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

A general conditional r. q. $X|Y$ is obtained by replacing in the last formula the event H (and its indicator) by a r. q. Y.

Definition 1. (Lad & Dickey)
Given two r. q. X and Y, 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 $G = sY[X - \mathbb{P}(X|Y)]$.

Definition 2. (Lad & Dickey)
Having asserted your conditional prevision $\mathbb{P}(X|Y) = \mu$, the c. r. q. $X|Y$ is defined as

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

By computing the prevision on both sides, it follows (*generalized compound prevision theorem*)

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

Some remarks.

1) if $Y \equiv 0$, you always receive the same amount $\mu = \mathbb{P}(X|Y)$ that you have payed (the net gain is always 0). To avoid this trivial case we will assume that $(Y = 0) \neq \Omega$.

2) if X and Y are uncorrelated, it is $\mathbb{P}(XY) = \mathbb{P}(X)\mathbb{P}(Y)$; then, assuming $\mathbb{P}(Y) \neq 0$, it follows $\mathbb{P}(X|Y) = \mathbb{P}(X)$.

In other words, *under the hypothesis* $\mathbb{P}(Y) \neq 0$*,* X *and* Y *are uncorrelated if and only if the prevision of 'X given Y' coincides with the prevision of* X*.*

3) $\mathbb{P}(Y) = 0 \;\not\Rightarrow\; \mathbb{P}(XY) = 0$; **then, it may happen that doesn't exists any finite value of** $\mathbb{P}(X|Y)$ **which satisfies the equality**

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

A critical example

(where $\mathbb{P}(Y) = 0\,,\quad \mathbb{P}(XY) \neq 0$)

$(X, Y) \in \{(0,-1), (0,1), (1,-1), (1,1)\}$;
we set $p(x, y) = P(X = x, Y = y)$, with

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

We have $Y \in \{-1, 1\}\,,\; XY \in \{-1, 0, 1\}$, with

$$P(Y = -1) = P(Y = 1) = \frac{1}{2}\,,\quad P(XY = 0) = \frac{1}{2}\,,$$

$$P(XY = -1) = \frac{1}{6}\,,\quad P(XY = 1) = \frac{1}{3}\,;$$

so that $\mathbb{P}(Y) = 0$ and $\mathbb{P}(XY) = \frac{1}{6}$; hence, *the equation* $\frac{1}{6} = \mathbb{P}(X|Y) \cdot 0$ *has no solutions.*

What about coherence of μ when $\mathbb{P}(Y) = 0$?

To properly manage the case $\mathbb{P}(Y) = 0$, we integrate the work of Lad and Dickey
(i) by using an explicit definition of coherence for any given assessment $\mathbb{P}(X|Y) = \mu$;
(ii) by discarding, in the definition of coherence, the value 0 of the net gain associated with the case $Y = 0$.

Definition of coherence. Given two r. q. 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 $s \neq 0$, and $H = (Y \neq 0)$. 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.
(without loss of generality, we can set $s = 1$)

Remark. If Y is the indicator $|H|$ of an event H, then $X|Y = X|(|H|)$ and $(Y \neq 0) \equiv (H$ *true*); then, the coherence of the assessment $\mathbb{P}(X|Y) = \mu$ reduces to the notion of coherence for the assessment $\mathbb{P}(X|H) = \mu$.

3. Some examples.

We continue the study of the critical example, by examining the coherence of a given assessment $\mathbb{P}(X|Y) = \mu$.
We recall that $(X, Y) \in \{(0,-1), (0,1), (1,-1), (1,1)\}$;
moreover $H = (Y \neq 0) = \Omega\,,\quad G|H = G = Y(X - \mu)$.
The values of $G|H$ associated with the values of (X, Y) are respectively:

$$g_1 = \mu,\;\; g_2 = -\mu,\;\; g_3 = -1 + \mu,\;\; g_4 = 1 - \mu;$$

hence: $\inf G|H \cdot \sup G|H \leq 0\,,\;\; \forall \mu$.

Another example.

$$(X, Y) \in \{(0,-1), (1,1)\}\,,\quad \mathbb{P}(X|Y) = \mu\,.$$

We have: $H = (Y \neq 0) = \Omega\,,\quad G|H = G = Y(X - \mu)$;
the values of $G|H$ are: $g_1 = \mu\,,\;\; g_2 = 1 - \mu\,;$

then: $\inf G|H \cdot \sup G|H \leq 0 \iff \mu \notin (0, 1)$.

Notice that, with each μ it is associated a probability distribution on (X, Y), say $(p, 1-p)\,,\; 0 \leq p \leq 1$, where

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

By requiring that the prevision of the random gain be 0, i.e. $p\mu + (1-p)(1-\mu) = 0$, one has $p = f(\mu) = \frac{1-\mu}{1-2\mu}$, with

$$\frac{1}{2} < p \leq 1\,,\;\text{ if }\; \mu \leq 0;\qquad 0 \leq p \leq \frac{1}{2}\,,\;\text{ if }\; \mu \geq 1\,.$$

Notice that $\mu = f^{-1}(p) = \frac{1-p}{1-2p}$; i.e., $f^{-1} = f$.

As shown by this example, the set of **coherent** *assessments* μ *may be* **not convex.**

A strong generalized compound prevision theorem

We recall that $H = (Y \neq 0)\,,\; \mu = \mathbb{P}(X|Y)$.
We assume that $\mu\,,\; \mathbb{P}(Y|H)\,$, and $\mathbb{P}(XY|H)$ are finite; then, we remark that

(i) we pay μ and we receive $X|Y$, under the hypothesis H true; then, *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, we have $\mu = \mu + \mathbb{P}[(XY - \mu Y)|H]$ and by linearity of prevision it follows

$$\mathbb{P}(XY|H) = \mathbb{P}(X|Y)\mathbb{P}(Y|H)\,. \tag{1}$$

Remark. If Y is a finite discrete r. q., with $Y \geq 0$, or $Y \leq 0$, it is $\mathbb{P}(Y|H) \neq 0$; then, by (1) it follows

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

As $H^c = (Y = 0)$, it is $\mathbb{P}(Y|H^c) = \mathbb{P}(XY|H^c) = 0$; hence,

$$\mathbb{P}(Y) = \mathbb{P}(Y|H)P(H) + \mathbb{P}(Y|H^c)P(H^c) = \mathbb{P}(Y|H)P(H)\,, \tag{2}$$

$$\mathbb{P}(XY) = \mathbb{P}(XY|H)P(H) + \mathbb{P}(XY|H^c)P(H^c) = \mathbb{P}(XY|H)P(H) \tag{3}$$

Then, by (1), (2), and (3), one has

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

(the formula of Lad & Dickey, which we call *weak* generalized compound prevision theorem).

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

Let $\mathcal{C}_X, \mathcal{C}_Y$ and $\mathcal{C}$ be, respectively, the finite sets of possible values of X, Y and (X, Y).

$$X^0 = \{x_h \in \mathcal{C}_X \,:\, \exists\, (x_h, y_k) \in \mathcal{C} \,:\, y_k \neq 0\}\,,\quad \begin{cases} x_0 = \min\, X^0\,, \\ x^0 = \max\, X^0\,. \end{cases}$$

Theorem 1 Given two finite r. q. X, Y, with $Y \geq 0$ or $Y \leq 0$, the prevision assessment $\mathbb{P}(X|Y) = \mu$ is coherent iff $x_0 \leq \mu \leq x^0$.

Example. $(X, Y) \in \mathcal{C} = \{(0,1), (1,0), (1,1), (2,2)\}$. One has

$$X^0 = X\,,\;\; x_0 = \min\, \mathcal{C}_X = 0\,,\;\; x^0 = \max\, \mathcal{C}_X = 2\,;$$

the values of $G|H$, where $H = (Y \neq 0)$, are

$$g_1 = -\mu\,,\;\; g_2 = 1 - \mu\,,\;\; g_3 = 2(2 - \mu)\,;$$

such values are *all positive* (resp., *all negative*) when $\mu < 0$ (resp., $\mu > 2$);
hence every $\mu \notin [x_0, x^0] = [0, 2]$ is *not coherent.*
Finally, when $\mu \in [0, 2]$ one has $-\mu(2 - \mu) \leq 0$ and the condition $\inf G|H \cdot \sup G|H \leq 0$ holds.

5. The case $\min\, Y < 0 < \max\, Y$.

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

$\mu_0 = \min\,(\max\, X^-, \max\, X^+)\,,\quad \mu^0 = \max\,(\min\, X^-, \min\, X^+)\,,$
if $\mu_0 < \mu^0$, we set $I = (\mu_0, \mu^0)$; moreover, we set

$X^- < X^+\,,$ if $\max\, X^- < \min\, X^+;\quad X^- > X^+\,,$ if $\min\, X^- > \max\, X^+;$

$X^- \nsim X^+$, otherwise. Then, we obtain

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 2 Let be given two r. q. X, Y, with $\min\, Y < 0 < \max\, Y$.
If case 1, or case 2, holds, then $X^- \cap X^+ = \emptyset$ and the 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.

Example. We determine the set Π of coherent prevision assessments $\mathbb{P}(X|Y) = \mu$ on $X|Y$, where

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

We have: $X^- = \{1\}\,,\quad X^+ = \{0\}$, so that $X^- \cap X^+ = \emptyset$ and $X^- \nsim X^+$.
Then, $I = (0, 1)$ and, by Theorem 2, $\Pi = \Re \setminus (0, 1)$; that is, μ is coherent if and only if $\mu \notin (0, 1)$.

The same result follows, by observing that:
(i) $G|H = G$;
(ii) given any μ, the values of G are:

$$g_1 = -\mu,\;\; g_2 = -2\mu,\;\; g_3 = -1 + \mu,\;\; g_4 = -2 + 2\mu;$$

(iii) if $\mu \in (0, 1)$, the values of G are all negative; if $\mu \notin (0, 1)$, it is: $\min\, G < 0\,,\; \max\, G > 0$.

6. Final comments

- the notion of general conditional prevision, $\mathbb{P}(X|Y)$, was introduced by Lad & Dickey, in the setting of the operational subjective theory of coherent previsions, to solve decision problems involving "state dependent preferences";
- in particular, it was applied to a "currency exchange problem" suggested by Jay Kadane.

Further developments of the research concern:
(i) coherence of a conditional prevision assessment $\mathcal{M}_n = (\mu_1, \ldots, \mu_n)$ on a family of n conditional random quantities $\mathcal{F}_n = \{X_1|Y_1, \ldots, X_n|Y_n\}$;
(ii) study of general properties and methods for the checking of coherence;
(iii) generalized coherence of imprecise conditional prevision assessments, for instance interval-valued assessments like $\mathcal{A}_n = ([l_1, u_1], \ldots, [l_n, u_n])$, on $\mathcal{F}_n$.

Some results concerning (i) and (ii) have been obtained in a paper which will be presented on September at WUPES 2009.