Stochastic causality and orthomodular lattices

Mária Bohdalová, Martin Kalina, Oľga Nánásiová

Comenius University, Slovakia
Slovak University of Technology, Bratislava, Slovakia

FSTA 2014, Liptovský Ján, January 27 - 31, 2014

Causality and stochastic causality

Cause and effect

Cause: I throw dice.
Effect: Dice falls down. It is not important which face.

Cause and stochastic effect

Cause: I throw dice.
Stochastic effect: Dice falls down. The face is a matter of stochasticity.
Theory of probability.

Causality and stochastic causality

Stochastic cause and stochastic effect

Stochastic cause: I throw dice only with probability 0.2.
Stochastic effect: Dice falls down with probability 0.2.
Theory of probability

Stochastic cause and stochastic effect

Stochastic cause: I throw dice only with probability 0.2
Stochastic effect: Dice falls down with probability 0.2. The face is a matter of stochasticity.
Theory of probability and conditional probability.

etc . . .

C. W. J. Granger: *Investiganting causal relations by econometric models and cross-spectral methods.* Econometrica, **37**, (1969)

Suppose we have two stationary time series

$$X = \{X(t)\}_{t \in Z} \qquad Y = \{Y(t)\}_{t \in Z}$$

and we intend to study whether X causes Y or not. Granger causality analysis is based on two principles:

- The cause happens prior to the effect.
- The cause makes unique changes in the effect. In other words, the causal series contains unique information about the effect series that is not available otherwise.

Granger causality

Let

- $I(t)$ is the set of all information in the universe up to time t
- $I_X(t)$ is the set of all information in the universe excluding X up to time t

Suppose all the information have been recorded on equally spaced time stamps $t \in Z$. Now given the two principles, the conditional distribution of future values of Y given $I_X(t)$ and $I(t)$ should differ.
X causes Y if

$$P(Y(t+1) \in A | I(t)) \neq P(Y(t+1) \in A | I_X(t))$$

for some measurable set $A \subseteq R$ and all $t \in Z$.

P.O. Amblard, O.J.J. Michal: *The relation between Granger Causality and directed information theory: A Review.* Entropy **15**, (2013)

Granger causality measures a stochastics dependence between the past of a process and the present of another. In this respect the word causality in the sense of Granger has the usual meaning that a cause occurs prior to its effect.

Granger causality is based on the ussual concept of conditional probability theory.

Let X, Y be random variables,

$$X \text{ be a cause and } Y \text{ be its effect.}$$

There exist two probability spaces

$$\mathcal{P}_X = (\Omega_X, \mathcal{F}_X, P_X) \qquad \mathcal{P}_Y = (\Omega_Y, \mathcal{F}_Y, P_Y)$$

How to model this situation?

- Within Cartesian products $\mathcal{P}_X \times \mathcal{P}_Y$ and $\mathcal{P}_Y \times \mathcal{P}_X$:
  we need two joint distributions $F_{X,Y} = F_X.F_Y$ and $F_{Y,X} \neq F_X.F_Y$.
- Within OML we need only one s-map p.

Let A, B be two random events, A be a cause and B be its effect. How is it possible to describe this causality via probability measure

$$p(effect, cause) = p(B, A) = ?$$

An orthomodular lattice (OML)

Definition

Let $(L, \mathbf{0}_L, \mathbf{1}_L, \vee, \wedge, \perp)$ be a lattice with the greatest element $\mathbf{1}_L$ and the smallest element $\mathbf{0}_L$. Let $\perp\colon L \to L$ be a unary operation on L with the following properties:

1. for all $a \in L$ there is a unique $a^\perp \in L$ such that $(a^\perp)^\perp = a$ and $a \vee a^\perp = \mathbf{1}_L$;
2. if $a, b \in L$ and $a \leq b$ then $b^\perp \leq a^\perp$;
3. if $a, b \in L$ and $a \leq b$ then $b = a \vee (a^\perp \wedge b)$ (orthomodular law).

Then $(L, \mathbf{0}_L, \mathbf{1}_L, \vee, \wedge, \perp)$ is said to be *an orthomodular lattice*.

A map $m : L \to [0,1]$ is called a σ-additive state on L, if for arbitrary, at most countable, system of mutually orthogonal elements $a_i \in L$, $i \in I \subset N$, the following holds

$$m(\vee_{i \in I} a_i) = \sum_{i \in I} m(a_i)$$

and $m(1_L) = 1$.

Probability versus state two random events

- Probability space: $P(E) = P(F) = 1 \quad \Rightarrow \quad P(E \cap F) = 1$
- OML: $m(a) = m(b) = 1$ *does not imply* $m(a \wedge b) = 1$
- In the special case — Jauch-Piron states: $m(a) = m(b) = 1 \quad \Rightarrow \quad m(a \wedge b) = 1$

As it was proved by R.Greechie, there exist orthomodular lattices with no state.

Let L be an OML and let B be a Boolean algebra. A map $h : L \to B$ fulfilling the following properties:

- $h(0_L) = 0_B$, $h(1_L) = 1_B$;
- $h(a^\perp) = h(a)^\perp \ \forall a \in L$;
- if $a \perp b$ then $h(a \vee b) = h(a) \vee h(b)$.

will be called a morphism from L to B.

If $a, b \in L$

$$h(a \vee b) \geq h(a) \vee h(b)$$

$$h(a \wedge b) \leq h(a) \wedge h(b)$$

If $a \leftrightarrow b$

$$h(a \vee b) = h(a) \vee h(b)$$
$$h(a \wedge b) = h(a) \wedge h(b)$$

If $a \leq b$ then $h(a) \leq h(b)$.

Let μ be an additive measure on B and h be a morphism $h : L \to B$. Then

a) $\mu_h : L \to [0,1]$ such that $\mu_h(a) = \mu(h(a)) \quad \forall a \in L$ is a state on L;

b) $p_h : L \times L \to [0,1]$ such that $p_h(a,b) = \mu(h(a) \wedge h(b)) \quad \forall a, b \in L$ induces a joint distribution on L;

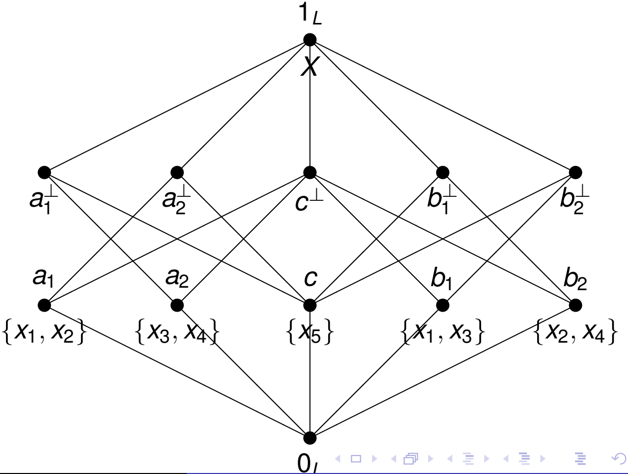
c) $d_h : L \times L \to [0,1]$ such that $d_h(a,b) = \mu(h(a) \triangle h(b)) \quad \forall a, b \in L$ is a measure of symmetric difference on L;

d) Let $B_0 = \{E \in B; \mu(E) > 0\}$ and $L_0 = \{e \in L; h(e) \in B_0\}$. Then $f_h : L \times L_0 \to [0,1]$ such that

$$f_h(a,b) = \frac{\mu(h(a) \wedge h(b))}{\mu(h(b))}$$

is a conditional state on L.

$X = \{x_1, x_2, x_3, x_4, x_5\}$ and $B = (X, 2^X, \mu)$:
$p_h(a_1, b_1) = \mu(\{x_1\})$
$d_h(a_1, b_1) = \mu(\{x_3, x_2\})$



Example: $L = \{a_1, a_2, c, b_1, b_2, a_1^\perp, a_2^\perp, c^\perp, b_1^\perp, b_2^\perp\}$

Let $X = \{x_1, x_2, x_3, x_4, x_5\}$ and $B = 2^X$. Let μ be a measure on B. Then $h : L \to B$ can be defined as follows

$$
\begin{array}{ll}
h(a_1) = \{x_1, x_2\} & h(a_2) = \{x_3, x_4\} \\
h(b_1) = \{x_1, x_3\} & h(b_2) = \{x_2, x_4\} \\
h(c) = \{x_5\} & h(c^\perp) = \{x_1, x_2, x_3, x_4\}
\end{array}
$$

s-map, symmetric difference, conditional state

a) $p_h(a_1, b_1) = \mu(h(a_1) \wedge h(b_1)) = \mu(\{x_1\})$

b)$d_h(a_1, b_1) = \mu(\{x_3, x_2\})$

c)$f_h(a_1, b_1) = \frac{\mu(\{x_1\})}{\mu(\{x_1, x_3\})} = P_\mu(\{x_1\}|\{x_1, x_3\})$

Two dimensional state - s-map

Definition

Let L be an OML. A map $p : L \times L \to [0,1]$ will be called **an s-map on** L if the following conditions are fulfilled:

(s1) $p(1_L, 1_L) = 1$;

(s2) for all $a, b \in L$ if $a \perp b$ then $p(a,b) = 0$;

(s3) for all $a, b, c \in L$ if $a \perp b$ then

$$p(a \vee b, c) = p(a,c) + p(b,c) \quad p(c, a \vee b) = p(c,a) + p(c,b).$$

In general $p(a,b) = p(b,a)$ **is not true.**
If $a \leftrightarrow b$, then $p(a,b) = p(a \wedge b, a \wedge b)$.

If $\mu_p : L \to [0,1]$ such that $\mu_p(a) = p(a,a)\ \forall a \in L$, then

(p1) μ_p is a state.

(p2) $p(a,b) \leq \mu_p(a)$ for all $a, b \in L$.

(p3) $p(a,b) = \mu_p(a \wedge b)$ for $a \leftrightarrow b$.

Jauch-Piron's property: Let $a, b \in L$.

$$p(a,a) = p(b,b) = 1 \text{ iff } p(a,b) = p(b,a) = 1$$

$$p(a,c) = p(b,c) \qquad \forall c \in L.$$

Let $d(a,b) = p(a^{\perp}, b) + p(a, b^{\perp})$. If $\forall a, b, c \in L$

$$p(a,b) = p(a,b), \qquad d(a,b) \leq d(a,c) + d(b,c)$$

then OML L with p looks like as classical probability space (virtual probability space).

Let L be a σ-OML. A σ-homomorphism x from Borel sets to L ($\mathcal{B}(R)$), such that $x(R) = 1_L$ is called an observable on L.

Let L be a σ-OML. Observables x, y are called compatible ($x \leftrightarrow y$) if $x(A) \leftrightarrow y(B)$ for all $A, B \in \mathcal{B}(R)$.

Loomis-Sikorsky Theorem

Let L be a σ-OML and x, y be compatible observables on L. Then there exists a σ-homomorphism H and real functions f, g such that $x(A) = H(f^{-1}(A))$ and $y(A) = H(g^{-1}(A))$ for each $A \in \mathcal{B}(R)$ (briefly $x = f \circ H$ and $y = g \circ H$).

Let x be an observable and m be a σ-additive state on L. Then the expectation of the observable x in the state m ($E_m(x)$) is defined by

$$E_m(x) = \int_R t\, m(x(dt)),$$

if the integral exists.

Let $(\Omega, \mathcal{F}, P)$ be a probability space.

- Hence $\mathcal{F}$ is a σ-OML. Furthermore, if ξ is a random variable on $(\Omega, \mathcal{F}, P)$, then ξ^{-1} is an observable.
- If we have an observable x on a σ-OML L, we are in the same situation as in the classical probability space. We use only another language for the standard situation.
- Problems occur if we have two causal observales.

Joint distribution

Let L be a σ-OML and $x, y \in \mathcal{O}$. Then a map $p_{x,y} : \mathcal{B}(\mathbb{R})^2 \to [0,1]$, such that $p_{x,y}(t,s) = p(x((-\infty,t)), y((-\infty,s)))$ is called a joint p-distribution of the observables x, y.

Causality

Let L be a σ-OML, x, y be observables and p be s-map. We say that: x is causal to y with respect to p if there exist $A, B \in \mathcal{B}(R)$ such that

$$p(x(A), y(B)) \neq p(y(B), x(A))$$

X cause Y its effect

Strong causality

Let L be a σ-OML, x, y be observables and p be s-map. We say that: x is causal to y with respect to p if $\exists E, F \in \mathcal{B}(R)$ such that

$$p(x(E), y(F)) \neq p(y(F), Y(F)).p(x(E), x(E))$$

and $\forall A, B \in \mathcal{B}(R)$

$$p(y(B), x(A)) = p(y(B), Y(B)).p(x(A), x(A))$$

X cause Y its effect

Let p be s-map. Conditional state:

$$f_p(a,b) = \frac{p(a,b)}{p(b,b)} \qquad p(b,b) \neq 0$$

Conditional expectation

Let L be a σ-OML, p be an s-map, x an observable and $\mathcal{B}$ be a Boolean sub-σ-algebra of L. A version of conditional expectation of the observable x with respect to $\mathcal{B}$ is an observable z (notation $z = E_p(x|\mathcal{B})$) such that $R(z) \subset \mathcal{B}$ and moreover

$$E_{f_p}(z|a) = E_{f_p}(x|a)$$

for arbitrary $a \in \{u \in \mathcal{B}; \mu_p(u) \neq 0\}$.

Since $R(x)$ is Boolean sub-σ-algebra of L we will write simply $E_p(y|x) = E_p(y|R(x))$.

X cause Y its effect

In fact, the conditional expectation $z = E_p(x|\mathcal{B})$ is a projection of the observable z into the Boolean σ-algebra $\mathcal{B}$. This means, if we have $z = E_p(y|x)$ then we have $z \leftrightarrow x$. This property implies that the conditional expectation $E_p(y|x)$ behaves as we are used to from the conditional expectation of random variables in the Kolmogorovian probability theory.

Properties

(e1) $E_p(E_p(x|y)) = E_p(x)$,

(e2) $E_p(x|x) = x$,

(e3) $E_p(E_p(x|y)|y) = E_p(x|y)$,

(e4) $E_p(x, E_p(y|x)) = E_p(x, y)$.

X cause Y its effect

If $x \leftrightarrow y$ then $x = f \circ H$ and $y = g \circ H$. Applying L-S Theorem we get

$$x + y = (f + g) \circ H.$$

If x, y are non-compatible then we cannot apply this procedure and $x + y$ does not exist in this sense.

Let L be a σ-OML and $p \in \mathcal{P}$. A map $\oplus_p : \mathcal{O} \times \mathcal{O} \to \mathcal{O}$ is called *a summability operator* if the following conditions are fulfilled

(d1) $R(\oplus_p(x, y)) \subset R(y)$;

(d2) $\oplus_p(x, y) = E_p(x|y) + y$.

X cause Y its effect

Let us have a state $\mu : L \to [0,1]$ defined by the following

$$\mu(t) = \begin{cases} 1, & \text{if } t = \mathbf{1}_L, \\ 0, & \text{if } t = \mathbf{0}_L, \\ 0.5, & \text{otherwise.} \end{cases}$$

Let x be an observable whose spectrum is $\mathcal{B}_1$, and y be an observable whose spectrum is $\mathcal{B}_2$. Then $x \not\leftrightarrow y$. We may have an s-map $p : L^2 \to [0,1]$, achieving the following values for non-compatible elements $s, t \in L$:

$$p(s,t) = \begin{cases} 0.3, & \text{if } (s,t) \in \{(a,b),(a^\perp,b^\perp)\}, \\ 0.2, & \text{if } (s,t) \in \{(a^\perp,b),(a,b^\perp)\}, \\ 0.1, & \text{if } (s,t) \in \{(b,a),(b^\perp,a^\perp)\}, \\ 0.4, & \text{if } (s,t) \in \{(b^\perp,a),(b,a^\perp)\}. \end{cases}$$

X cause Y its effect

Definition

Let L be a σ-OML, $\mathcal{B}$ be a Boolean sub-σ-algebra of L, and $p \in \mathcal{P}$. A map $\oplus_p^{\mathcal{B}} : \mathcal{O} \times \mathcal{O} \to \mathcal{O}$ is called *a summability operator with respect to a condition* $\mathcal{B}$ if the following conditions are fulfilled

(a1) $R\left(\oplus_p^{\mathcal{B}}(x, y)\right) \subset \mathcal{B}$;

(a2) $\oplus_p^{\mathcal{B}}(x, y) = E_p(x|\mathcal{B}) + E_p(y|\mathcal{B})$.

Proposition. Let L be a σ-OML, $\mathcal{B}$ be a Boolean sub-σ-algebra of L, and $p \in \mathcal{P}$. Assume $x, y \in \mathcal{O}$. Then the following statements are satisfied

(e1) if $x \leftrightarrow y$ then $\oplus_p(x, y) \leftrightarrow \oplus_p(y, x)$;

(e2) $\oplus_p^{\mathcal{B}}(x, y) = \oplus_p^{\mathcal{B}}(y, x)$;

(e3) $E_p\left(\oplus_p^{\mathcal{B}}(x, y)\right) = E_p(\oplus_p(x, y)) = E_p(x) + E_p(y)$;

(e4) if $\sigma(x) = \{x_1, x_2, ..., x_n\}$ and $\sigma(y) = \{y_1, y_2, ..., y_k\}$ then

$$E_p(x) + E_p(y) = \sum_i \sum_j (x_i + y_j) p(x(\{x_i\}), y(\{y_j\})).$$

Thank you for your kind attention