STOCHASTICA, Vol. IV, nº 1 (1980)

SUR LES INFORMATIONS CONDITIONNELLES SANS F-COMPOSABILITE

Llorenç Valverde

ABSTRACT

In the framework of Generalized Information Theory of J. Kampé de Fériet, a system of axioms for the conditional information is studied without assuming additivity.

0. Introduction

Dans la Théorie de l'Information "a priori", essentiellement on a étudié les informations F-composables, c'est-à-dire, les informations telles que si $A \cap B = \emptyset$, on a $I(A \cup B) = F(I(A), I(B))$, propriété qui est satisfaite par l'information de Wiener-Shannon. Dans les premiers travaux de J. Kampé de Fériet et B. Forte [4] sur ce sujet-là et dans ceux de J. Kampé de Fériet [2, 3], on remarque que cette propriété n'admet pas de justification intuitive, mais que du point de vue mathématique elle offre des grands avantages. Par contre la relation entre $I(A)$, $I(B/A)$ et $I(A \cap B)$, est plus intuitive et cette relation fait intervenir la notion d'indépendance.

CLASSIFICATION AMS (1980) 94 A 15.

Dans cette note on traite de cette question, en postulant une propriété de <u>consécution</u> qui donne une relation entre I(A/B), I(B) et I(A $\cap$ B) sans présupposer la F-composabilité (cela laisse la possibilité de travailler dans des inf-demi-treillis). La technique qu'on a suivi est, en quelque sorte, parallèle a celle de la formulation de Rényi pour les probabilités conditionnelles (voir [1], p. 320), c'est-à-dire elle consiste à définir l'information conditionnelle axiomatiquement. En plus de la donnée de ces axiomes, on définit l'indépendance et le produit, finalement on confronte ces deux concepts via un théorème de representation similaire à celui de Banach-Marczewski pour les probabilités.

1. Axiomatique

Soit $R \equiv (R, \wedge, 0)$ un inf-demi-treillis avec minimum 0. Comme d'habitude, on écrira $R^* = R - \{0\}$ et $\bar{R}^+ = [0, +\infty]$. Une <u>information conditionnelle</u> sur R est une application J de $R \times R^*$ dans $\bar{R}^+$ vérifiant les axiomes suivants:

1) <u>Monotonie:</u> Si $a \leqslant a'$, alors $J(a,b) \geqslant J(a',b)$.

2) <u>Comparabilité:</u> $J(a \wedge b,b) = J(a,b)$.

3) <u>Valeurs universelles:</u> a) $J(a,b) = +\infty$, si et seulement si $a \wedge b = 0$.
   b) $J(a,a) = 0$.

4) <u>Consécution:</u> Il existe une opération binaire commutative G dans $\bar{R}^+$ telle que si $b \wedge c \neq 0$, alors $J(a \wedge b,c) = G(J(a,b \wedge c), J(b,c))$.

On appelle <u>espace d'information conditionnelle</u> (en abrégé EIC) le triplet (R,J,G). Le théorème suivant nous permettra de caractériser les opérations G liés à un EIC:

<u>Théorème 1.1.-</u> Soit (R,J,G) un EIC; on suppose que pour chaque $b \in R^*$, l'application $J(\cdot,b)$ de R dans $\bar{R}^+$ telle que

$J(,b)(a) = J(a,b)$, soit surjective, et que si $x \leqslant x'$ dans $\bar{R}^+$ alors ils existent $a \geqslant a'$ dans R tels que $x = J(a,b) \leqslant J(a',b) = x'$. Dans ces conditions $(\bar{R},G,\leqslant)$ est un semi-groupe ordonné, commutatif avec unité 0, et zéro $+\infty$.

En vertu de ce théorème nous restreindrons notre attention aux EIC (R,J,G) tels que G vérifie les propriétés ci-dessus et, en plus, nous postulerons l'hypothèse additionnelle d'archimédianité sur G (c'est-à-dire, G est continue et telle que pour chaque $x \in (0, +\infty)$ on a $G(x,x) > x$), cela nous permet de prendre en considération la représentation de fonctions associatives archimédiennes (Aczel [1], Ling [5]) que, d'après l'analyse de l'équation fonctionnelle fournie par l'axiome IV, nous mène au théorème suivant:

<u>Théoreme 1.2.-</u> Soit g une fonction de $\bar{R}^+$ dans $\bar{R}^+$, continue, strictement croissante et telle que $g(0) = 0$. Soit $f: R \to \bar{R}^+$, décroissante et telle que $f^{-1}(+\infty) = 0$. Alors l'application J_f de $R \times R^*$ dans $\bar{R}^+$, définie par

$$J_f(a,b) = \begin{cases} +\infty, & \text{si } a \wedge b = 0 \\ g^{[-1]}(g(f(a \wedge b)) - g(f(b)), & \text{sinon} \end{cases}$$

est une information conditionnelle relative à l'opération $G(x,y) = g^{[-1]}(g(x)+g(y))$.

Si (R,J,G) est un EIC et Ψ une application de $\bar{R}^+$ dans $[0,1]$ telle que

1.1.1) Ψ est continue,

1.1.2) strictement décroissante, (1.1)

1.1.3) $\Psi(0) = 1$, $\Psi(+\infty) = 0$,

alors $H = \Psi \circ J$ est une application de $R \times R^*$ dans $[0,1]$ vérifiant les propriétés:

1.2.1.) Si $a \leqslant a'$, alors $H(a,b) \leqslant H(a',b)$.

1.2.2.) $H(a \wedge b,b) = H(a,b)$. (1.2.)

1.2.3.) a) $H(a,b) = 0$ si et seulement si $a \wedge b = 0$;

b) $H(a,a) = 1$.

1.2.4.) $K = \Psi \circ G \circ (\Psi^{-1} \times \Psi^{-1})$ est une opération dans $[0,1]$ telle que, si $b \wedge c \neq 0$, alors $H(a \wedge b,c) = K(H(a,b \wedge c), H(b,c))$.

Evidemment $([0,1], K, \leqslant)$ est un semi-groupe ordonné commutatif avec unité 1, et zéro 0. En particulier, si (Ω, A, P) est un espace de probabilité $(\tilde{A},\tilde{P})$ est la probabilité induite sur l'algèbre quotient $\tilde{A}=A/P$ et $\tilde{A}^* = \tilde{A} - \{\tilde{\emptyset}\}$. L'application H de $\tilde{A} \times \tilde{A}^*$ dans $[0,1]$ définie par $H_P(\tilde{A},\tilde{B}) = P(A/B)$ vérifie les propriétés (1.2) ci-dessus, avec l'opération $K(x,y) = x.y$. Par suite, pour toute application Ψ de $\bar{R}^+$ dans $[0,1]$ vérifiant les propriétés (1.1) ci-dessus, il résulte que $J = \Psi^{-1} \circ H_P$ est une information conditionnelle relative à l'opération $G(x,y)=[\Psi^{-1} \circ K \circ (\Psi \times \Psi)](x,y)$. En particulier, pour $\Psi(x) = e^{-x/c}$ on obtient l'information de Wiener-Shannon.

2. Produit d'espaces d'information conditionnelle.

Dans la suite, les inf-demi-treillis considérés auront minimum et maximum. Soient $R_i \equiv (R_i, \wedge, 0_i, u_i)$, $i = 1,2$; deux inf-demi-treillis avec minimum 0_i, et maximum u_i. L'ensemble $R_1 \otimes R_2 = (R_1^* \times R_2^*) \cup \{(0_1, 0_2)\}$, muni de l'opération:

$$(a_1,a_2) \wedge (a_1',a_2') = \begin{cases} (a_1 \wedge a_1', a_2 \wedge a_2'), & \text{si } a_1 \wedge a_1' \neq 0_1 \text{ et } a_2 \wedge a_2' \neq 0_2; \\ (0_1,0_2), & \text{sinon;} \end{cases}$$

a une structure d'inf-demi-treillis avec minimum $(0_1,0_2)$ et maximum (u_1,u_2). Si $(R_1 \otimes R_2, J, G)$ est un EIC, les applications

$$J_1(a_1,a_1') = \begin{cases} +\infty, & \text{si } a_1 \wedge a_1' = 0_1 \ ; \\ J((a_1,u_2),(a_1',u_2)), & \text{sinon;} \end{cases} \qquad (2.1)$$

$$J_2(a_2,a_2') = \begin{cases} +\infty, & \text{si } a_2 \wedge a_2' = 0_2 \ ; \\ J((u_1,a_2),(u_1,a_2')), & \text{sinon ;} \end{cases}$$

définissent des informations conditionnelles sur R_1 et R_2 respectivement, relatives à la même operation G. J_1 et J_2 seront appelées <u>informations conditionnelles marginales</u> de J.

<u>Théoreme 2.1.-</u> Soient (R_i,J_i,G_i), i = 1,2; deux EIC et F une application de $\bar{R}^+ \times \bar{R}^+$ dans $\bar{R}^+$. L'application J de $(R_1 \otimes R_2) \times (R_1 \otimes R_2)^*$ dans $\bar{R}^+$ définie par

$$J((a_1,a_2),(a_1',a_2'))=F(J_1(a_1,a_1'),J_2(a_2,a_2')) \qquad (2.2)$$

est une information conditionnelle sur $R_1 \otimes R_2$ relative à une opération G, si et seulement si $F(0,0) = 0, F(+\infty,+\infty) = +\infty$, F est croissante et

$$G(F(x,y),F(z,t))=F(G_1(x,z),G_2(y,t)) \qquad (2.3)$$

Dans ces conditions, J_1 et J_2 définies par (2.1) sont les informations conditionnelles marginales de l'information J définie par (2.2) si et seulement si

$$F(x,0) = F(0,x) = x \text{ et } F(x,+\infty) = F(+\infty,x) = +\infty,$$

et dans ce cas la seule solution de (2.3) est $F = G = G_1 = G_2$.

3. Independance en information conditionnelle.

Définition 3.1.- Soit (R,J,G) un EIC, on dira que deux éléments a, b $\epsilon R - \{0,u\}$ sont J-indépendants si et seulement si $J(a,b) = J(a,u)$ et $J(b,a) = J(b,u)$. L'élément 0, minimum de R, est J-indépendant de tout autre. Si R_i, $i = 1,2$; sont deux sous-inf-demi-treillis de R tels que $0, u \epsilon R_i$, $i = 1,2$; on dira qu'ils sont J-indépendants si chaque élément de $R_1 - \{0,u\}$ est J-indépendant de chaque élément de $R_2 - \{0,u\}$.

En général une des deux égalités antérieures n'est pas suffisante pour affirmer la J-indépendance de deux éléments de R (voir [6], p. 44).

Théorème 3.1.- i) Si a et b sont J-indépendants, alors $J(a \wedge b,u) = G(J(a,u), J(b,u))$. Si G est strict la réciproque est aussi vraie.

ii) Si $J(a,b) = J(a,u)$ et G est strict, alors a et b sont J-indépendants.

iii) Si G est strict et si a et b sont J-indépendants, ainsi que b et c, a et $b \wedge c$, alors c et $a \wedge b$ sont aussi indépen-dants.

Les considérations suivantes seront utilisées pour la formulation d'un théorème qui met en relation la J-indépendance et le produit d'EIC, ressemblant, donc, à celui de Banach-Marczewski pour les probabilités.

Soit $(R_i)_{i \epsilon I}$ une famille finie de sous-inf-demi-treillis de l'inf-demi-treillis $R \equiv (R,\wedge,0,u)$ tels que 0, $u \epsilon R_i$ pour tout $i \epsilon I$, on dira qu'ils sont M-indépendants si et seulement si pour tout $a_i \epsilon R_i - \{0,u\}$, $i \epsilon I$, on a $\bigwedge_{i \epsilon I} a_i \notin \bigcup_{i \epsilon I} R_i$. On définit

$\bigwedge_{i \epsilon I} R_i = \{ \bigwedge_{i \epsilon I} a_i \ ; \ a_i \epsilon R_i\}$, comme étant le plus petit sous-inf-demi-treillis de R qui contienne tous les R_i. On l'appelera inf-demi-treillis engendré par la famille $(R_i)_{i \epsilon I}$.

<u>Théorème 3.2.-</u> Soient R_i, $i = 1,2$; deux sous-inf-demi-treillis M-indépendants et complets de l'inf-demi-treillis complet $R = (R,\wedge,0,u)$, tels que 0, $u \in R_i$, $(i=1,2)$ et que $R = R_1 \wedge R_2$. Si (R_1,J_1,G) et (R_2,J_2,G) sont EIC alors il existe une information conditionnelle J, sur R relative à l'opération G telle que:

i) $J|_{R_1\times R_1^*} = J_1$; $J|_{R_2\times R_2^*} = J_2$.

ii) R_1 et R_2 sont J-indépendants.

<u>Démonstration:</u> Soit f l'application de $R_1 \wedge R_2$ dans $R_1 \otimes R_2$ définie par $f(c) = (f_1(c),f_2(c))$, où $f_1(c)=\mathrm{Inf}\{a_1\in R_1; \exists\ a_2\in R_2$ et $a_1 \wedge a_2 = c\}$, $f_2(c)=\mathrm{Inf}\{a_2\in R_2; \exists\ a_1\in R_1$ et $a_1 \wedge a_2 = c\}$. f est bien définie, est une injection et $f_1(c) \wedge f_2(c) = c$ quel que soit $c\in R_1 \wedge R_2$. Soit maintenant $(R_1 \otimes R_2, J',G)$ l'EIC produit de (R_1,J_1,G) et (R_2,J_2,G) est J l'application de $R \times R^*$ dans $\bar{R}^+$ définie par $J(c,c') = J'(f(c),f(c'))$. Il est clair que (R,J,G) est un EIC. Soit $a_1\in R_1$; de la M-indépendance de R_1 et R_2, il résulte que $f(a_1) = \{(a_1,u)\}$, et par suite pour chaque a_1, $a_1'\in R_1$ on a

$$J(a_1,b_1) = J'(f(a_1),f(b_1)) = J'((a_1,u),(b_1,u))=J_1(a_1,b_1),$$

car J_1 est marginal de J', ainsi on a montré que $J|_{R_1\times R_1^*} = J_1$. De la même façon on montrerait que $J|_{R_2\times R_2^*} = J_2$.

Finalement, si $a_1\in R_1 - \{0,u\}$ et $a_2\in R_2 - \{0,u\}$ il résulte que

$$J(a_1,a_2)=J'(f(a_1),f(a_2))=J'((a_1,u),(u,a_2))=G(J_1(a_1,u),J_2(u,a_2))$$

$$=G(J_1(a_1,u),\ 0)\ =J_1(a_1,u)\ =\ J(a_1,u),$$

de la même façon on montrerait que $J(a_2,a_1) = J(a_2,u)$ et, par suite, la J-indépendance de R_1 et R_2.

REMERCIEMENTS. L'auteur remercie Monsieur le Professeur J. Kampé de Fériet (Lille) pour la suggestion de ce travail et Monsieur le Professeur C. Alsina (Barcelona) pour ses remarques très utiles dans la réalisation de celui-ci.

References

[1] ACZEL, J., "Lectures on functional equation and their applications", Academic Press, New York, 1966.

[2] KAMPÉ de FÉRIET, J., "Mesure de l'information fournie par un événement". Colloques Internationaux du CNRS, Clermont-Ferrand, 1969, 191-221.

[3] KAMPÉ de FÉRIET, J., "L'indépendance des événements dans la Théorie Généralisée de l'Information". Journées Lyonnaises des Questionnaires, CNRS, Groupe de Recherches 22, Paris 1976, 1-30.

[4] KAMPÉ de FÉRIET, J., "Information et probabilité". C.R.A.S., 265, 1967 (I) 110-114, (II) 142-146. (III) 350-353.

[5] LING, C.H., "Representation of associative functions". Publ. Math. Debrecen, 12, 1965, 189-212.

[6] VALVERDE, L., "Sobre unes informacions condicionades". Thèse de Troisième Cycle, Universitat de Barcelona, 1979 (Non publié).

Departament de Matemàtiques i Estadística.
Escola Tècnica Superior d'Arquitectura.
Universitat Politècnica de Barcelona.
Diagonal, 649, Barcelona (28) (ESPAGNE).