CLASIFICACION Y CONSTRUCCION DE LAS MEDIDAS DE INCERTIDUMBRE

TRABAJO DE GRADO
SOMETIDO A CONSIDERACION
DEL
POSTGRADO INTEGRADO EN MATEMATICA
UCLA - IUPEB - IUP

POR

MARIA LUISA CAPODIECI DE PICHARDO

COMO UNO DE LOS REQUISITOS PARA OBTENER EL GRADO DE
MAGISTER EN MATEMATICA
MENCION: MATEMATICA APLICADA

Julio, 1986

TRABAJO DE GRADO
SOMETIDO A CONSIDERACION
DEL
POSTGRADO INTEGRADO EN MATEMATICA
UCLA - IUPEB - IUP

POR

MARIA LUISA CAPODIECI DE PICHARDO

COMO UNO DE LOS REQUISITOS PARA OBTENER EL GRADO DE
MAGISTER EN MATEMATICA
MENCION: MATEMATICA APLICADA

Julio, 1986

DEDICATORIA

A mis padres

AGRADECIMIENTO

Sirvan estas líneas para agredecer a:

Dr. Carlo Bertoluzza por permitirme conocer por medio de este trabajo un campo relativamente nuevo e interesante como lo es la teoría de la información.

Dr. José Sarabia y Dr. Francisco Montes de Oca y a todos los profesores que en alguna oportunidad me impartieron sus conocimientos, para asi lograr culminar esta Maestría.

Lic. Jesús Rodríguez por su colaboración y ayuda en el desarrollo del presente trabajo.

Guillermo, Gríselda, Ursula y Marianella Pichardo (esposo e hijas) por su comprensión y paciencia.

I N D I C E

	Página
NOTACIONES.	8
INTRODUCCION.	10
CAPITULO I	
RESUMEN DE LOS RESULTADOS PRINCIPALES DE LA TEORIA DE **LA INFORMACION** DE **EVENTOS**.	14
1.1. DEFINICIONES Y PROPIEDADES ELEMENTALES.	15
1.2. INFORMACION CONDICIONAL	17
1.3. MEDIDAS DE INFORMACION "COMPONIBLES".	20
1.4. MEDIDAS DE INFORMACION DE TIPO M Y DE TIPO M' . .	22
1.5. MEDIDAS DE INFORMACION DE TIPO INF.	25
1.6. LEYES DE COMPOSICION UNIVERSAL CLASICAS	27
1.7. GENERALIZACION DEL AXIOMA DE INDEPENDENCIA. . . .	28
1.8. PRODUCTO DE ESPACIOS DE INFORMACION.	29
1.9. INFORMACION CONDICIONAL EN UN ESPACIO DE INFORMACION CON LEY DE INDEPENDENCIA GENERALIZADA. . . .	34
1.10. LEYES DE COMPOSICION UNIVERSAL GENERALIZADAS. . .	36
CAPITULO II	
TEORIA DE LA INCERTIDUMBRE. **DEFINICION Y PROPIEDADES FUNDAMENTALES**.	40
2.1. ILUSTRACION DE LOS AXIOMAS.	41
2.2. ESPACIOS DE INCERTIDUMBRE.	43
2.3. PROPIEDADES ELEMENTALES DE LA MEDIDA DE INCERTIDUMBRE.	45
2.4. EL AXIOMA DE INDEPENDENCIA.	46
2.5. INCERTIDUMBRE CONDICIONAL	49
2.6. CLASIFICACION DE LAS MEDIDAS DE INCERTIDUMBRE . .	62

Página

CAPITULO III

MEDIDAS DE INCERTIDUMBRE CON LA PROPIEDAD DE RAMIFICACION. 69

3.1. 70

3.2. PROPIEDADES DE G 73

3.3. MEDIDAS DE INCERTIDUMBRE g-LOCALES 79

3.4. ENTROPIAS g-RAMIFICADAS. 90

3.5. CONSTRUCCION DE LAS MEDIDAS DE INCERTIDUMBRE g-LOCALES. 93

3.6. FORMULA RECURRENTE 98

3.7. LA RESTRICCION $\hat{H}$ 99

3.8. FORMA EXPLICITA DE LAS MEDIDAS DE INCERTIDUMBRE g-LOCALES 100

3.9. EJEMPLOS DE INCERTIDUMBRES g-LOCALES. 108

CAPITULO IV

MEDIDAS DE INCERTIDUMBRE COMPOSITIVAS. 114

4.1. 116

4.2. COMPONIBILIDAD TOTAL. 118

4.3. NOCION E IMPORTANCIA DEL CONCEPTO DE LEY UNIVERSAL. 121

4.4. LEYES g-UNIVERSALES. 122

4.5. LEYES UNIVERSALES CLASICAS 128

4.6. LEYES SUP-UNIVERSALES. 130

4.7. LEYES g-UNIVERSALES. 133

4.8. LEYES ESTRICTAMENTE MONOTONAS. 134

4.9. LA ENTROPIA DE HARTLEY 141

4.10. CONSTRUCCION DE LAS MEDIDAS DE INCERTIDUMBRE COMPOSITIVAS 144

Página

4.11. FORMULA RECURRENTE. 146

4.12. FORMA EXPLICITA DE LAS MEDIDAS DE INCERTIDUMBRE COMPOSITIVAS 147

4.13. EJEMPLOS DE INCERTIDUMBRE COMPOSITIVAS. 149

B I B L I O G R A F I A. 160

A P E N D I C E

APENDICE A

APENDICE B

APENDICE C

APENDICE D

APENDICE E

NOTACIONES

$\mathbf{R}$: El conjunto de los números reales

$\mathbf{R}^+$: El conjunto de los números reales positivos

$\mathbf{P}(\Omega)$: El conjunto de partes de Ω

$\mathcal{I} = (\Omega, \mathcal{S}, \mathcal{K})$: Estructura de Información

Ω : Conjunto de los eventos elementales

$\mathcal{S}$: Algebra de $\mathbf{P}(\Omega)$ (eventos observables).

$\mathcal{K}$: Subfamilia de K_m = clase de las colecciones de álgebras independientes.

J : Medida de información

$\mathbf{I} = (\Omega, \mathcal{S}, \mathcal{K}, J)$: Espacio de información

$\mathbf{B}[o,1]$: Borelianos del intervalo $[o,1]$

$J(A|B)$: Información condicional

$\Lambda = \mathbb{R}^+ - \cup (a_i, b_i)$: Conjunto de los elementos "idempotentes"

$\Lambda_* = \{x : x * x = x\}$: Conjunto de los elementos "$*$-idempotentes"

$\overline{\Lambda}_* = \{x : x * x \neq x\}$

$\{\Delta_i = (a_i, b_i)\} / i = 1, \ldots, I \leq +\infty$ (intervalos disjuntos)

$C[o,M]$: Conjunto de funciones continuas del intervalo $[o,M]$ en $\mathbb{R}$

T : Ley de composición

$\Gamma^{(2)} = \{(x,y) : x = J(A),\ y = J(B);\ A,B \in \mathcal{S}\ y\ (A,B) \in \mathcal{K}\} \subset (\mathbb{R}^+)^2$

$: = :$ "definido por:

$L = \{(x,y) \ / \ y < x\}$

$\overline{L} = \{(x,y) \ / \ x = y\}$

$\Pi_A = \{A_1, A_2, \ldots, A_n\}$: Partición del conjunto A.

$\Pi_A < \Pi'_A$: Refinamiento de particiones

$\Pi_A \wedge \Pi_B$: Producto de particiones

$\Pi_A \vee \Pi_B$: Unión de particiones

$\mathcal{H} = (\mathbf{I}, \mathcal{E})$: Estructura de Incertidumbre

$\mathcal{E}$: Familia de particiones de conjuntos de $\mathcal{S}$.

H : Medida de Incertidumbre

$\boldsymbol{\mathcal{H}} = (\mathcal{H}, H)$: Espacio de incertidumbre

$$\tilde{\Gamma}^{(2)} = \{(u,v) : u = H(\Pi_A),\ v = H(\Pi_B),\ (\Pi_A, \Pi_B) \in \mathcal{K}\} \subset (\mathbf{R}^+)^2$$

$K(\Pi_A)$: Incertidumbre condicional de la partición Π_A

$H(\Pi|\Pi')$: Incertidumbre condicional de la partición Π respecto a la partición Π'.

$$\Gamma_2 = \{(x,y) : x \geq y\}$$

$$D_n := \{(J(A_1), J(A_2), \ldots, J(A_n)) \ / \ A_1, A_2, \ldots, A_n \in \mathcal{S}\}$$

$$\Gamma_3 = \{(u,v,w) : \mathrm{Inf}(u,v) \geq w \geq o\}$$

$$\Gamma_5 = \{(x,y,z,u,v) \ / \ o \leq z \leq x \leq u,\ o \leq z \leq y \leq v\}$$

$$\Gamma_4 = \{(x,y,u,v) \ / \ o \leq x \leq u, \quad o \leq y \leq v\}$$

$$\Gamma^{(4)} = \{(A,B,u,v)/(A,B) \in \mathbf{P}(\Omega)^2, (u,v) \in \mathbb{R}^2, A \cap B = \phi,\ o \leq J(A) \leq u,\ o \leq J(B) \leq v\}$$

$$\tilde{\Gamma}^{(2)} = \{(x,y) : x\ J(A),\ y = J(B),\ A \cap B = \phi\}$$

$$\Gamma^{(2)} = \{(A,u) : A \in \mathbf{P}(\Omega),\ u \in \mathbb{R}^+,\ o \leq J(A) \leq u\}$$

$$\Gamma_n = \{(x_1, \ldots, x_n) : F_n(x_1, \ldots, x_n) \geq o\}$$

$$\Gamma_2^{(w)} = [w, +\infty)^2 \subset (\mathbb{R}^+)^2$$

INTRODUCCION

La teoría axiomática de la información (y/o de la incertidumbre) ha surgido y se ha desarrollado después que C. E. Shannon en su trabajo fundamental "A Mathematical Theory of Communication" mostró la importancia fundamental de la función "información o entropía"

$$H(p_1,\ldots,p_n) = - \Sigma\, p_i \log p_i$$

en los procesos de transmisión y organización de datos. El propio Shannon, después de introducir en forma operativa esa función, trató de subrayar su importancia al proponer una definición simple axiomatizada.

A pesar de ésto, los que más impulsan el desarrollo de la Teoría axiomática son, por un lado A. J. Khinchin y D. K. Faddeev [28, 29] quienes introducen dos caracterizaciones por medio de axiomas del todo naturales, y por otro A. Rényi y N.A. Kolmogorov [30] quienes proponen medir la información por funciones diferentes de la de Shannon.

Sin embargo, en todos estos trabajos la información (o la incertidumbre, como vamos a llamarla en esta tesis) siempre está definida como función de una distribución de probabilidad (completa o incompleta) $p_1\ldots,p_n$.

Sólo en 1962 R. S. Ingarden y K. Urbanik [27] y más claro y completamente en 1967 J. Kampé de Feriét y B. Forte [3,7] propusieron y desarrollaron una teoría axiomática, donde las medidas de información no necesitan como soporte un espacio de probabilidad, sino sólo la esquematización simple del espacio de los eventos y de las experiencias. Además de la independencia de las probabilidades, el punto quizás más importante donde la teoría de Kampé de Feriét y Forte su diferencia de las precedentes consiste en distinguir entre información e incertidumbre, rechazando la identificación, siempre aceptada antes, entre "incertidumbre removida" e "información aportada". Ellos destacan, en efecto, que lo que aporta información es el producirse un acotecimiento, mientras que se puede hablar de incertidumbre sólo en presencia de situaciones donde se espera el resultado de una experiencia con un cierto número de éxitos posibles pero no seguros. Por lo tanto se asociará una medida de información a un evento (sub-conjunto de un espacio adecuado) y una medida de incertidumbre a una experiencia (colección de eventos disjuntos). Los axiomas que definen esas dos medidas han sido expuestos por primera vez en los trabajos [3,7]. Después, otros investigadores han analizado estos axiomas obteniendo unos cuantos resultados relacionados con estas teorías.

En 1970-71 J. Kampé de Feriét dictó en Roma un curso sobre la teoría axiomática de la información. El texto de dicho curso, organizado y editado por G. Maschio [2], constituye una exposición bastante completa de los resultados hasta entonces obtenidos sobre la medida de información aportada por un evento.

Por otra parte, todavía no hay una exposición unitaria análoga de los resultados concernientes a las medidas de incertidumbre. Este trabajo es una tentativa parcial para cubrir esta falta; por supuesto, no nos propusimos reportar todos los resultados relacionados con este tópico, sólo queremos; por un lado presentar e ilustrar los axiomas y las nociones principales relacionadas con los espacios de incertidumbre, por otro intentar una clasificación sistemática de las medidas hasta ahora propuestas en varios trabajos.

C A P I T U L O I

RESUMEN DE LOS RESULTADOS PRINCIPALES DE LA TEORIA DE LA INFORMACION DE EVENTOS

En este capítulo presentaremos definiciones y teoremas concernientes a las "medidas de información de eventos". Se trata de una sección introductoria cuyo contenido es necesario para poder desarrollar, en los capítulos sucesivos el verdadero argumento de este trabajo. Por esta razón nociones y resultados que vamos a exponer en este capítulo serán dados en forma esencial, sin ilustrar las definiciones y sin demostrar los teoremas, excepto el teorema relacionado con la noción de información condicional, ya que ese teorema ha sido probado y escrito por primera vez en el presente trabajo. La demostración sigue el esquema proporcionado por C. Poggi en [17] aplicado a las medidas de incertidumbre.

Por otro lado, este argumento ha sido desarrollado con muchos detalles por J. Kampé de Feriet. Ver [2].

1.1. DEFINICIONES Y PROPIEDADES ELEMENTALES

Definición 1.1.1. (Ver [3])

Una "Estructura de Información" es una terna $I = (\Omega, S, K)$ donde:
- Ω es el conjunto de los eventos elementales
- S es un álgebra de $\mathbf{P}(\Omega)$ (eventos observables)
- K es una subfamilia de K_m = clase de las colecciones de álgebras independientes 1/.

Definición 1.1.2.

Una "medida de Información" sobre una estructura I es una aplicación

$$J : S \longrightarrow \mathbf{R}^+ \quad \text{tal que}$$

(1.1-1) $J(\Omega) = 0$

(1.1-2) $A \subset B \rightarrow J(A) \geq J(B)$

(1.1-3) $\{A^i, i=1,\ldots,m\} \in K,\ A_{ji} \in A_i \rightarrow J(\bigcap_{i=1}^{n} A_{ji}) = \sum_{i=1}^{n} J(A_{ji})\ (\forall n \leq m)$

Definición 1.1.3

A la estructura $\mathbf{I} = (\Omega, S, K, J)$ se le llama "espacio de Información"

1/ Sca $A^1, \ldots, A^n$ una colección de álgebras de $\mathbf{P}(\Omega)$; diremos que estas álgebras son m - independientes si para cada $A_{ji} \in A^i$, $(i=1,\ldots,n)$ $A_{ji} \neq \phi$ se cumple $\bigcap_{i=1}^{n} A_{ji} \neq \phi$.

Propiedades elementales de la Medida de Información

i) $J(\phi) = +\infty$

ii) $J(A \cap B) \geq \text{Sup}\,[J(A),\ J(B)]$

$J(A \cap B) \leq \text{Inf}\,[J(A),\ J(B)]$

Ejemplos de Medidas de Información

Sea $(\Omega, \mathcal{S}, \mathcal{K})$ una estructura de información

a. Si (Ω, d) es un espacio métrico y $f: \mathbf{R}^+ \to \mathbf{R}^+$ es un función no creciente, entonces $J(A) \overset{2/}{:=} f\,[d(A)] - f\,[d(\Omega)] \quad \forall A \in \mathcal{S}$ es una medida de información. En particular lo es

$$J(A) = [d(A)]^{-1} - [d(\Omega)]^{-1}$$

b. Si $(\Omega, \mathcal{S}, p)$ es un espacio de probabilidad y f una función real positiva no creciente, entonces

$$J(A) := f\,[p(A)] - f(1) \quad \forall A \in \mathcal{S}$$

es una medida de información. En particular lo es

$$J(A) = -\log p(A)\ ,$$

$$J(A) = p^{-1}(A) - 1$$

c. Sean: $\Omega = [o,1]$, $S = \beta[o,1]$ (ó $\mathbf{S} = \mathrm{P}\,[o,1]$

c.1. $J(A) := \text{Inf}\,\{x^2 : x \in A\}$

c.2. $J(A) := \text{Inf}\,\{\delta(x-1/2) : x \in A\}$, $J(\phi) = +\infty$

donde
$$\delta(\xi) = \begin{cases} 1 & \text{si } \xi = o \\ o & \text{si } \xi \neq o \end{cases}$$

2/ : = definido por.

1.2. INFORMACION CONDICIONAL

Definición 1.2.1.

Sea (Ω,S,K,J) un espacio de Información, definimos la "información condicional" por medio de la expresión:

$$J(A|B) = J(A \cap B) - J(B).$$

Se puede probar fácilmente que J tiene las siguientes propiedades:

1) $J(.|B)$ es una medida de información

2) $J(B|B) = 0$

3) Si A_1, A_2, B_1, B_2 son independientes con respecto a J entonces $J(A_1 \cap A_2|B_1 \cap B_2) = J(A_1|B_1) + J(A_2|B_2)$

Las propiedades 1) 2) y 3) son las propiedades naturales de la información condicional. Podemos preguntarnos si existen otras expresiones de $J(A|B)$ que satisfacen estas propiedades. Supongamos que sí existen y que tienen la forma siguiente

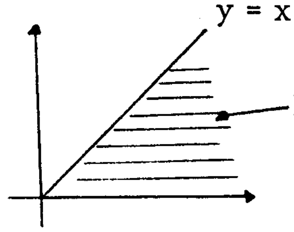
$$J(A|B) = F[J(A \cap B), J(B)]$$

Teorema 1.2.1.

Las informaciones condicionales, continuas y universales difieren de $J(A|B) = J(A \cap B) - J(B)$ por una constante multiplicativa.

Demostración:

Bajo la hipótesis de que: F es universal y considerando que $A \subset B \Longrightarrow J(A) \geq J(B)$, se sigue que: Dom $[F(x,y)] = \{(x,y): 0 \leq y \leq x\}$



Haciendo $x_i = J(A_i \cap B_i)$, $y_i = J(B_i)$ $i = 1,2$ de 3) obtenemos $F(x_1 + x_2, y_1 + y_2) = F(x_1,y_1) + F(x_2,y_2)$, además de 2) $F(y,y) = 0$.

Queremos entonces hallar la solución del siguiente sistema de ecuaciones funcionales:

1. $F(x_1 + x_2, y_1 + y_2) = F(x_1,y_1) + F(x_2,y_2)$

2. $F(y,y) = 0$

Escribamos 1. como una ecuación funcional de Cauchy; para ello tomemos $\bar{y}_1$, $\bar{y}_2$ fijos. Como $F(x,y) = F_y(x)$, escribimos 1. bajo la forma

$$F_{\bar{y}_1 + \bar{y}_2}(x_1 + x_2) = F_{\bar{y}_1}(x) + F_{\bar{y}_2}(x_2)$$

Para reducir ésta a una ecuación funcional de Cauchy, hagamos $\bar{y}_1 = \bar{y}_2 = 0$; obtenemos entonces

$$F_o(x_1 + x_2) = F_o(x_1) + F_o(x_2)$$

y por lo tanto

$$F_o(x) = F(x,o) = c.x$$

Por otra parte si ponemos $y_1 = 0$ y $y_2 = y$ en 1 tenemos $F(x_1 + x_2, y) = F(x_1, 0) + F(x_2, y) = c.x_1 + F(x_2, y)$

Si $y_1 = y$ y $y_2 = 0$ en 1 tenemos

$$F(x_1 + x_2,\ y) = F(x_1,\ y) + F(x_2,\ 0) = F(x_1,\ y) + c\ x_2$$

Por lo tanto

$$F(x_2,\ y) - c.x_2 = F(x_1, y) - c\ .\ x_1 = g(y)$$

(La primera igualdad demuestra que la expresión $F(x,y) - cx$ no depende de la variable x)

Luego $F(x,y) = = (x + g\ (y))$

Si hacemos $x = y$ obtenemos $F(y,y) = 0 = cy + g(y)$, o sea

$g(y) = - cy$. Por lo tanto

$$F(x,y) = cx - cy = c(x-y)$$

es decir, $F(x,y) = c(x-y)$ es la solución del sistema planteado.

En definitiva

$$J(A|B) = c\,[J(A \cap B) - J(B)]$$

□

1.3. MEDIDAS DE INFORMACION "COMPONIBLES"

Definición 1.3.1.

Sea $\mathbb{I} = (\Omega, \mathcal{S}, \mathcal{K}, J)$ un espacio de información, diremos que la medida de información J es "componible" si para dos eventos A, B en $\mathcal{S}$ tales que $A \cup B \in \mathcal{S}$ y $A \cap B = \phi$, tenemos

(1.3.1) $J(A \cup B) = F\,[J(A), J(B)]$.

A la función $F(x,y) = x\ T\ y$(notación algebraica) de la expresión (1.3.1) se le denomina "Ley de composición" para la información.

La medida de información del ejemplo a. no es componible porque $d(A \cup B)$ no se puede escribir en términos de d(A) y d(B).

Las medidas de información del ejemplo b. sí son componibles y tienen como ley de composición

$$f\,[f^{-1}(x + c) + f^{-1}(y + c)] - c$$

donde $c = f(1)$.

También la información del ejemplo c. es componible, con ley de composición Inf(x,y).

Ley de Composición

La forma más general de las leyes de composición continuas es proporcionada por el siguiente teorema (Ver [4]).

Teorema 1.3.1.

Sea $\{\Delta_i = (a_i,b_i)/i=1,\ldots, I \leq +\infty\}$ una sucesión numerable de intervalos disjuntos y $f_i : [a_i,b_i] \to \mathbf{R}^+$ una familia de funciones decrecientes (estrictamente) con $f_i(b_i) = 0$ y sea

$g_i(\xi) = f^{-1} [\operatorname{Inf} (f(a_i),\xi)]$ la seudoinversa de f_i. Cada función definida haciendo

$$F(x,y) = \begin{cases} g_i \{f_i(x) + f_i(y)\} & \text{si } (x,y) \in \Delta_i \times \Delta_i \\ \operatorname{Inf}(x,y) & \text{en caso contrario} \end{cases}$$

representa una posible ley de composición y viceversa; cada posible ley de composición es de esta forma bajo una adecuada selección de $\{\Delta_n\}$ y $\{f_i\}$. Los elementos del conjunto $\Lambda = \mathbf{R}^+ - \cup(a_i,b_i)$ se llaman "idempotentes" o, en forma más precisa, F - idempotentes; es decir, éllos cumplen la propiedad $F(x,x) = x$.

1.4. MEDIDAS DE INFORMACION DE TIPO M Y DE TIPO M'

Sean un espacio medible $(\Omega, \mathcal{S})$ y $m: \mathcal{S} \to \mathbb{R}^+$ una medida definida sobre dicho espacio, tal que $m(\Omega) \leq +\infty$.

Definamos la información como sigue:

(1.4.1) $\forall A \in \mathcal{S},\ J(A) = \Theta\,[m(A)]$

donde $\theta : [o,m] \to \mathbb{R}^+$ tiene las siguientes propiedades:

1. $\theta(m) = 0$, $\theta(o) = +\infty$
2. $\theta(x)$ es estrictamente decreciente
3. $\theta(x) \in C\,[o,m]$

Definición 1.4.1.

Cualquier medida de información definida sobre un espacio de medida $(\Omega, \mathcal{S})$ por (1.4.1) recibe el nombre de "Medida de Información de tipo M".

Sea $M = m(\Omega) < +\infty$ y definamos

(1.4.2) $\forall A \in \mathcal{S},\ J(A) = \psi[m(A)]$ donde

$$\Psi(x) = \begin{cases} \theta(x) & \text{si } x \in [o,m] \\ 0 & \text{si } x \in (m, +\infty) \end{cases}$$

Definición 1.4.2.

Cualquier medida de información definida sobre un espacio de medida (Ω, S) por (1.4.2) recibe el nombre de "Medida de Información de Tipo M'".

Observación 1.4.1.

Las medidas de información de los tipos M y M' son informaciones "componibles" donde las leyes de composición son respectivamente:

$$xTy = \theta\{\theta^{-1}(x) + \theta^{-1}(y)\}$$

$$xTy = \psi\{\theta^{-1}(x) + \theta^{-1}(y)\}$$

Ejemplos:

Información de Tipo M

1. Sean (Ω, S, p) un espacio de probabilidad y $\theta(x) = -c \log x$.
   La medida de información $J(A) = - c \log p(A)$ es de tipo M y tiene ley de composición $xTy = - c \log \{\exp(-x/c) + \exp(-y/c)\}$ definida para cualquier (x,y) que satisfaga

   $$\{c^{-x/c} + e^{-y/c}\} \leq 1$$

2. Sean $\Omega = \mathbf{N}, S = \mathbf{P}(\Omega)$, $n : S \to \mathbf{R}^+$, $n(A) = \#A$. Entonces (Ω, S, n) es un espacio medible.

$$\text{Sea } \theta(x) = \begin{cases} 1/x & \text{si } 0 < x < \infty \\ +\infty & \text{si } x = 0 \\ 0 & \text{si } x = +\infty \end{cases}$$

entonces la medida de información de tipo M

$$J(A) = \theta[n(A)] = \begin{cases} 1/n(A) & \text{si } 0 < n(A) < +\infty \\ 0 & \text{si } n(A) = +\infty \\ +\infty & \text{si } n(A) = 0 \end{cases}$$

tiene la ley de composición

$$xTy = \{1/x + 1/y\}^{-1}$$

Información de tipo M'

Sea (Ω, S, p) un espacio de probabilidad. La aplicación $J : S \to \mathbf{R}^+$ definida por $J(A) = 1 - p(A)$ es una medida de información de tipo M' con $\theta(x) = 1-x$ y

$$\Psi(\xi) = \begin{cases} 1 - \xi & \text{en } (o,1) \\ 0 & \text{en } (1, +\infty) \end{cases}$$

y con ley de composición

$$xTy = \begin{cases} x + y - 1 & \text{si } 2-x-y \in (0,1) \\ 0 & \text{si } 2-x-y \in (1, +\infty) \end{cases}$$

1.5. MEDIDAS DE INFORMACION DE TIPO INF.

Definición 1.5.1.

Una medida de información se dice de "tipo Inf" si es componible y tiene como ley de composición

$$xTy = \text{Inf}\,(x,y)$$

Propiedades:

1. Como $J(A \cup B) \leq \text{Inf}\,[J(A), J(B)]$, la información de tipo Inf es maximal.

2. $J(A \cup B) = J(A)\ T\ J(B)\ \forall\ A,B$ (aunque no sean disjuntos).

Sean J una medida de información de tipo Inf definida sobre S y $\{A_i\}_{i\in I}$ una familia formada por elementos de S, si la relación

$$(1.5.1)\quad J\,\{\bigcup_{i\in I} A_i\} = \text{Inf}\,\{J(A_i)\}$$

se verifica para cualquier familia numerable de índices I, entonces J se dice de "Tipo Inf."

Si la relación (1.5.1) se verifica para cualquier familia infinita de índices I, entonces J se dice de "tipo Inf-c".

Construcción general

Sean: Ω un conjunto cualquiera, $S = \mathbf{P}(\Omega)$ y $\eta : \Omega \to \mathbf{R}^+ \cup \{o\}$ tal que $\operatorname{Inf}_{\omega \in \Omega} \eta(\omega) = 0$

La aplicación $J : S \to \mathbf{R}^+$ definida por

$$(1.5.2) \quad J(A) = \begin{cases} \operatorname{Inf}_{\omega \in A} \eta(\omega) & \text{Si } A \neq \phi \\ + \infty & \text{Si } A = \phi \end{cases}$$

es una medida de tipo Inf.

Se puede verificar que satisface todos los axiomas que caracterizan la información, escoguiendo la clase $\mathcal{K}$ en forma adecuada y además también satisface

$$J(A \cup B) = \operatorname{Inf} \{J(A), J(B)\}$$

Además todas las informaciones de tipo Inf-c son de la forma (1.5.2) con una selección adecuada de la función η.

1.6. LEYES DE COMPOSICION UNIVERSAL CLASICAS

Definición 1.6.1.

Sea $\{\mathcal{A}_i \ (i \in I \text{ finito})\}$ un elemento de $\mathcal{H}$ y fijemos sobre $\mathcal{A}_i$ una medida de información $J^{(i)}: \mathcal{A}_i \to \mathbf{R}^+$. Llamaremos "Universal clásica" a una ley de composición T, si es posible determinar una medida de información

$$J : \mathcal{S} \to \mathbf{R}$$

con las siguientes propiedades:

1. J es componible con ley de composición T.

2. $J|_{\mathcal{A}_i} = J^{(i)}$ (J restringida a $\mathcal{A}_i$).

3. Las algebras $\mathcal{A}_i$ son J-independientes, es decir, $\forall A_i \in \mathcal{A}_i$ se tiene

$$J(\bigcap_{i\in I} A_i) = \sum_{i\in I} J(A_i) = \sum_{i\in I} J^{(i)} (A_i)$$

4. Las propiedades 1, 2 y 3 se satisfacen para todas las posibles selecciones de la familia $\mathcal{A}_i$ y de las medidas $J^{(i)}$ definidas sobre $\mathcal{A}_i$.

Teorema 1.6.1.

Las únicas leyes de composición universal clásicas son las leyes de composición Shannónianas, dadas por:

(1.6.1.) $xTy = -c \log\{\exp(-x/c) + \exp(-y/c)\}$

y al final su caso límite cuando $c \to o$.

(1.6.2) $xTy = \text{Inf}\,(x,y)$

1.7. GENERALIZACION DEL AXIOMA DE INDEPENDENCIA

C. Bertoluzza y F. Barbaini [5] han aclarado cómo el axioma de independencia proporciona fuertes limitaciones en lo que respecta a la utilización de las medidas de información. Por eso, es oportuno generalizar la noción de J-independencia sustituyendo la propiedad (1.1.3) por una menos fuerte que puede ser formulada asi:

(1.1.3') $(A,B) \in K \to J(A \cap B) = G[J(A),\, J(B)]$

donde la función $G(x,y)$, definida sobre el conjunto

$$\Gamma^{(2)} = \{(x,y) : x = J(A),\ y = J(B);\ A,B \in S \text{ y } (A,B) \in K\} \subset (\mathbb{R}^+)^2$$

determina la información de la intersección de dos eventos A,B, J-independientes en términos de J(A) y J(B).

La forma continua más general de la ley de independencia G(x,y) se obtiene escogiendo una sucesión $\{\Delta_i = (\alpha_i,\beta_i)\}$ de intervalos abiertos disjuntos y una sucesión $\{\Psi_i : \Delta_i \to \mathbb{R}\}$ de funciones continuas estrictamente crecientes y haciendo

$$(1.7.1.)\quad x * y = \begin{cases} \gamma_i\ (\Psi_i(x) + \Psi_i(y) & \text{si } (x,y) \in \Delta_j^2 \\ \\ \operatorname{Sup}(x,y) & \text{si } x \in \Delta_j,\ y \notin \Delta_j \text{ ó} \\ & \text{si } (x,y) \in \Lambda_*^2 \end{cases}$$

donde $\gamma_i(\xi) = \Psi^{-1}\,[\operatorname{Inf}(\Psi(b_i),\xi)]$ es la seudoinversa de Ψ_i.

En particular:

Si $\Lambda = \mathbf{R}^+$ entonces $x * y = \operatorname{Sup}(x,y)$

Si $\Lambda = \{0, +\infty\}$ entonces $x * y = \bar{g}\,\{g(x) + g(y)\}$

1.8. PRODUCTO DE ESPACIOS DE INFORMACION

Sea $\mathbf{I}_1 = (\Omega_1, S_1, \mathbf{K}_1, J_1)$ e $\mathbf{I}_2 = (\Omega_2, S_2, \mathbf{K}_2, J_2)$ dos espacios de información y propongamos como objetivo el de construir sobre la estructura $(\Omega, S, \mathbf{K})$ definida por

$\Omega = \Omega_1 \times \Omega_2$

$S = S_1 \times S_2 = \{A_1 \times A_2 : A_1 \in S_1,\ A_2 \in S_2\}$

K = adecuado

una medida de información $J : S \to \mathbf{R}^+$ que sea compatible con las medidas J_1 y J_2. El adjetivo "Compatible" está determinado por la siguiente definición:

Definición 1.8.1.

Diremos que el espacio $\mathbf{I} = (\Omega, S, K, J)$ es "compatible" con los espacios $\mathbf{I}_1$, $\mathbf{I}_2$ si para cada $A_1 \in S_1$, $A_2 \in S_2$ se cumple

(1.8.1) $J(A_1 \times \Omega_2) = J_1(A_1)$, $J(\Omega_1 \times A_2) = J_2(A_2)$

(1.8.2) $J(A_1 \times A_2) = \gamma[J_1(A_1), J_2(A_2)]$

Para comprender mejor la relación que hay entre la función γ y la noción de independencia conviene enfrentar el problema desde otro punto de vista. Supongamos que en el espacio producto $\mathbf{I}$ se consideren "a priori" J-independientes todas las parejas de álgebras (A', A'') con $A' = A_1 \times \Omega_2$ y $A'' = \Omega_1 \times A_2$, siendo A_1 y A_2 subálgebras de S_1 y S_2, es decir

(1.8.3) $A_1 \subset S_1$, $A_2 \subset S_2 \to (A_1 \times \Omega_2, \Omega_1 \times A_2) \in K$

Este es el caso, muy común, donde los eventos de S_1 y S_2 representan posibles éxitos de dos experiencias independientes. Como $A_1 \times A_2 = (A_1 \times \Omega_2) \cap (\Omega_1 \times A_2)$ se tiene:

(1.8.4) $\gamma[J(A_1), J(A_2)] = J(A_1 \times A_2) = G[J(A_1 \times \Omega_2), J(\Omega_1 \times A_2)]$
$= G[J_1(A_1), J_2(A_2)]$

donde G es la ley de independencia del espacio de información **I**, que, por cierto, bajo la condición (1.8.3) coincide con la aplicación γ. En definitiva se reconoce que:

- Si se verifica la condición (1.8.3), entonces también se cumple (1.8.2) (con $\gamma(x,y) = G(x,y)$).

- Si se cumple (1.8.2), entonces en el espacio **I** la clase K y la función $G(x,y) := \gamma(x,y)$ forman la estructura de independencia más natural.

Uno de los problemas más naturales relacionado con la construcción del espacio producto concierne la posibilidad de que **I** sea componible cuando así son los espacios factores $\mathbf{I}_1$ e $\mathbf{I}_2$. Precisaremos mejor nuestro objetivo por medio de las siguientes proposiciones:

a. Supondremos que $\mathbf{I}_1$ e $\mathbf{I}_2$ sean espacios componibles con leyes de composición iguales. $F_1(x,y) = F_2(x,y) = F(x,y)$.

b. Queremos que también **I** sea componible con una adecuada ley de composición $\Psi(x,y)$.

c. La función γ sea la misma para todas las parejas de espacios con ley de composición $F(x,y)$.

De la identidad $(A_1 \cup B_1) \times C_2 = (A_1 \times C_2) \cup (B_1 \times C_2)$ se desprende inmediatamente (por (1.8.2)) que entre las funciones F, γ, ψ, tiene que establecerse la relación

(1.8.5) $\gamma[F(a_1, b_1), C_2] = \psi[\gamma(a_1, C_2), \gamma(b_1, C_2)]$

y además (por c.) esta relación tiene que cumplirse para todos los valores $(a_1, b_1, c_2) \in \mathbb{R}^{+3}$. Además, de (1.8.1) y (1.8.2) se tiene

$$\gamma(x,o) = \gamma(o,x) = x$$

Haciendo en (1.8.5) $C_2 = 0$, se obtiene

$$F(x,y) = \gamma(x,y)$$

o sea, la ley de composición en el espacio producto coincide con las leyes de los espacios factores y (1.8.5) toma la forma

(1.8.6) $\gamma[\psi(x,y), Z] = \psi[\gamma(x,z), \gamma(y,z)]$

Utilizando oportunamente las relaciones (1.8.1)(1.8.2) (b.) (c.) se reconoce fácilmente que la función γ es asociativa, simétrica-monótona y $\gamma(x,x) \geq x$, $\gamma(x,\infty) = \infty$. Por lo tanto $\gamma(x)$ tiene la forma de una ley de independencia.

La relación (1.8.6) afirma que bajo las condiciones pedidas se puede construir el espacio producto sólo si γ es una ley de composición universal respecto a la ley de independencia respresentada por γ. En particular si la ley de independencia es la función

$$\gamma(x,y) = x + y$$

entonces se pueden usar, para construir productos de espacios de información componibles, solo espacios de información de tipo Inf o de tipo Shannon, a pesar del hecho que existen aún otros espacios de información importantes (el hiperbólico por ejemplo).

Esta limitación ha sido una de las razones que ha causado la adopción, por parte de unos cuantos autores [21, 5,6], de la versión generalizada (1.1.3'), en lugar de la clásica (1.1.3), del axioma de independencia.

Sin embargo el resultado de C. Bertoluzza y A. Boscaini [6] ha demostrado que hay espacios de información que de ninguna manera pueden utilizarse como factores para construir el espacio producto. Afortunadamente a esa clase no pertenecen los espacios más conocidos y más utilizados.

1.9. INFORMACION CONDICIONAL EN UN ESPACIO DE INFORMACION CON LEY DE INDEPENDENCIA GENERALIZADA

Supongamos que la ley de independencia definida en el espacio de información (Ω, S, K, J) sea una ley cualquiera de la forma (1.7.1). En este caso no se puede definir la información condicional por medio de la expresión $J(A|B) = J(A \cap B) - J(B)$ porque es compatible solo con la ley de independencia $x * y = x + y$. Tenemos entonces que definir $J(A|B)$ en forma **axiomática** por medio de las propiedades 1,2 del § 2 poniendo

$$J(A|B) = F\,[J(A \cap B),\ J(B)]$$

Para hallar las posibles formas de la función $F(x,y)$ tenemos entonces que resolver el sistema

(1.9.1) $F(x_1 * x_2,\ y_1 * y_2) = F(x_1,y_1) * F(x_2,y_2)$

(1.9.2) $F(x,x) = 0$

Si $a = \mathrm{Sup}\,\{x : x \in \Lambda\}$

$L = \{(x,y) : y < x\}$, $\overline{L} = \{(x,y) : x = y\}$

la solución general de este sistema es proporcionada por el siguiente teorema.

Teorema 1.9.1.

Cualquier función F(x,y) que sirva para determinar la información condicional de experiencias, es del tipo:

$$F(x,y) = \begin{cases} 0 & \forall\ (x,y) \notin [o,a]^2 \cup \bar{L} \\ f\,[g(x) - g(\mathrm{Sup}\ (y,a))] & \forall\ (x,y) \in L - [o,a]^2 \end{cases}$$

donde g(x) es la función que define la operación $*$ en $[a, +\infty)$, f(u) puede ser una función cualquiera definida sobre $\mathbb{R}^+$, escogiendo $\lambda \in \Lambda$, arbitrario y poniendo

$$f(u) = \begin{cases} \bar{h}\ (Ku),\ K \geq 0 & \text{si } \exists\ \beta \in \Lambda,\ \beta \neq \lambda : (\lambda,\beta) \subset \Lambda \\ \lambda & \text{si } \lambda \text{ es un punto de acumulación para } \Lambda \end{cases}$$

donde h(x) es la función que define $*$ en (λ,β) y $\bar{h}$ es la seudoinversa.

Si además de (1.9.1) y (1.9.2) se pide que $J(A|\Omega) = J(A)$, es decir

(1.9.3) $F(x,o) = x$

entonces se puede definir $J(A|B)$ solo si $\Lambda = \{0, +\infty\}$ y en este caso del teorema anterior se tiene:

La forma más general de las leyes de composición $*$-universales todavía no ha sido conseguida, sin embargo han sido determinadas (Ver [6]) las leyes universales correspondientes a las leyes de independencia (entre las cuales está la clásica) que satisfacen la condición

$$\lim_{x \to o} x * y = y$$

Los resultados obtenidos son expuestos por las siguientes dos proposiciones:

Proposición 1.10.1

La ley de composición

(1.10.1) $xTy = \mathrm{Inf}(x,y)$

es universal respecto a cualquier ley de independencia "$*$".

Proposición 1.10.2.

Leyes de composición $*$-universales diferentes de (1.10.1) existen solo si se cumple la condición

$$(o,c) \subset \overline{\Lambda}_o = \{x: \; x_* x \neq x\} \;, \; c > o$$

Es decir, si existe un intervalo abierto no vacío con primer extremo en el origen constituido por elementos que no son $*$-idempotentes. En éste caso además de (1.10.1), es $*$-universal. También la ley de composición definida por

$$(1.10.2) \quad xTy = \begin{cases} g\,[f(x) + f(y)] & \text{si } (x,y) \in (o,c)^2 \\ \\ \operatorname{Inf} f(x,y) & \text{si } (x,y) \notin (o,c)^2 \end{cases}$$

donde $f(x) = e^{-k\,\psi(x)}$, siendo Ψ la función que define "$*$" en $(0,c)^2$ y $g(t)$ es la seudoinversa de f.

Observación 1.10.1.

Examinando (1.10.2) se reconoce sin mucho esfuerzo que en este caso el conjunto de los elementos idempotentes de la ley de composición T (elementos T-idempotentes) es el conjunto

$$\Lambda_T = \{0\} \cup [c, +\infty)$$

Observación 1.10.2.

Si 0 (cero) es un punto de acumulación para Λ_* entonces solo (1.10.1) es $*$-universal. En particular esto sucede cuando la ley de independencia $*$ es la ley límite $x * y = \operatorname{Sup}(x,y)$.

CAPITULO II

TEORIA DE LA INCERTIDUMBRE. DEFINICION Y PROPIEDADES FUNDAMENTALES

En este capítulo se presentan los axiomas que definen los espacios de incertidumbre según la teoría propuesta por J. Kampé de Feriet y B. Forte [7]; además se exponen nociones y resultados generales relacionados con esa teoría (incertidumbre condicional e independencia). A la introducción formal de los axiomas se le antepone una digresión informal hecha para justificarlos.

En los capítulos siguientes utilizaremos como sinónimos las palabras "medida de incertidumbre" y "entropía". En otros textos se atribuye el mismo sentido a la palabra "medida de información esperada".

2.1. ILUSTRACION DE LOS AXIOMAS

Sea Ω un conjunto de eventos elementales y $S = \mathbf{P}(\Omega)$ el conjunto de los eventos observables. Uno experiencia sobre el espacio (Ω, S) se identifica con una colección finita $\{A_1, \ldots, A_n\}$ de elementos no vacios de S (resultados previstos) disjuntos entre si o con una partición Π_A del conjunto $A = \bigcup_{i=1}^{n} A_i$, soporte de la experiencia.

Axioma 2.1.1.

Es un axioma cuya razón principal consiste en relacionar entre sí las nociones de información y de incertidumbre. $H(\{A\})$ se puede interpretar como la incertidumbre de que se vaya a presentar el evento A (antes de saber si se presentará) mientras $J(A)$ es la información proporcionada por A después que se ha realizado. Es natural suponer que estas dos cantidades coincidan en base a la hipótesis:

Información aportada por A = Incertidumbre removida antes de saber si A se realizó.

Axioma 2.1.2.

Se consideran las experiencias Π_A y Π'_A de la figura 1, donde Π'_A se obtiene subdividiendo el evento A_2 de Π_A en otros dos eventos $A_2^{(1)}$ y $A_2^{(2)}$.

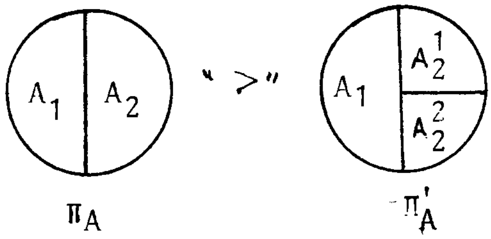


FIGURA 1.

La incertidumbre de cuál de los eventos se ha dado; es mayor en la experiencia Π'_A que en la experiencia Π_A; porque "antes" de la subdivisión teniamos la incertidumbre de que se diera A_1 ó A_2, mientras que "después" tenemos además la incertidumbre que si se da A_2, se halla dado A_2^1 o A_2^2.

Axioma 2.1.3.

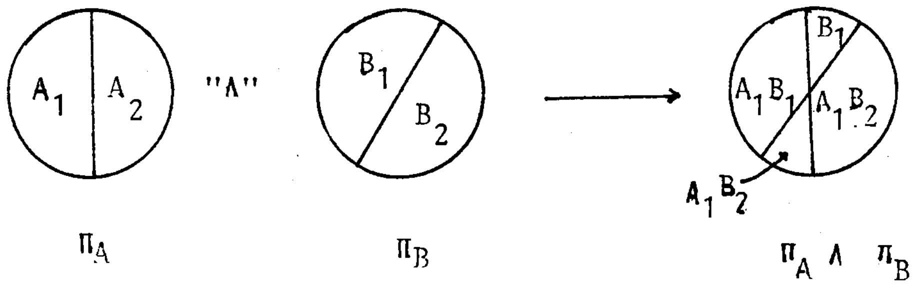


FIGURA 2.

Dadas dos experiencias Π_A y Π_B, evaluar $H(\Pi_A \wedge \Pi_B)$ significa medir la incertidumbre de saber cuál es el evento que se dió de Π_A y cuál el de Π_B, como el producto de particiones (de experiencias) no es más que una superposición de particiones (Ver Fig. 2).

Resulta no sólo razonable, sino también casi obligatorio suponer que si las dos experencias son independientes entonces la incertidumbre de cuál evento de Π_A y cuál de Π_B se presentará, se puede obtener de las incertidumbres respectivas. En efecto: decir que Π_A y Π_B son H-independientes significa que el presentarse un evento de Π_A no influye de ninguna manera en el éxito de la otra experiencia Π_B (en lo que respecta a la incertidumbre) y por esto la incertidumbre $H(\Pi_A \wedge \Pi_B)$ tendrá que ser completamente determinada por $H(\Pi_A)$ y $H(\Pi_B)$ por medio de una ley de independencia adecuada.

En la primera definición que se dió de Medida de Incertidumbre, Forte y Kampé de Feriét [7] utilizaron como ley de independencia, la clásica es decir $x * y = x + y$. Sin duda esta ley es la más natural (se suman las incertidumbres de los éxitos de Π_A y de Π_B para obtener la de $\Pi_A \wedge \Pi_B$) y la más útil en muchas aplicaciones. Sin embargo hay situaciones donde esta ley no puede ser usada (más adelante examinaremos con más detalles los motivos).

2.2. ESPACIOS DE INCERTIDUMBRE

Definición 2.2.1.

Una "Estructura de Incertidumbre" es una pareja $\mathcal{H} = (\mathbf{I},\xi)$ donde:

- $I = (\Omega, S, K, J)$ es un espacio de información según la definición 1.1.3

- E es una familia de particiones de conjuntos de S cuyos elementos componentes son también elementos de S . (familia de las experiencias).

En la familia E tomaremos en consideración la relación de orden < (refinamiento) y las operaciones $\wedge$ (producto de particiones) y v (unión de particiones) definidas por:

$$\Pi_A < \Pi'_A \longleftrightarrow \forall A_i \in \Pi_A, \ \exists A'_j \in \Pi'_A : A_i \subset A'_j$$

$$\Pi_A \wedge \Pi_B = \{A_i \cap B_j \ / \ A_i \in \Pi_A, \ B_j \in \Pi_B\} = \Pi_{A \cap B}$$

$$\Pi_A \vee \Pi_B = \{A_1, \ldots, A_n, B_1, \ldots, B_m \ / \ A_i \in \Pi_A \ , \ B_j \in \Pi_B\} = \Pi_{A \cup B}$$

Solo si $A \cap B = \phi$

Definición 2.2.2.

Una "medida de Incertidumbre sobre una estructura H es una aplicación

$$H : E \to \mathbb{R}^+ \text{ tal que}$$

(2.2.1) $H(\{A\}) = J(A)$

(2.2.2) $\Pi_A < \Pi_B \to H(\Pi_A) \geq H(\Pi_B)$

(2.2.3) $(\Pi_A , \Pi_B) \in H^{3/} \to H(\Pi_A \wedge \Pi_B) = H(\Pi_A) + H(\Pi_B)$

3/ $(\Pi_A, \Pi_B) \in K$ es una notación abreviada para decir que la pareja de álgebras generadas por Π_A y Π_B pertenecen a K.

Definición 2.2.3.

Un "espacio de incertidumbre" es una estructura

$\mathbb{H} = (\mathcal{H}, H)$ donde

- $\mathcal{H}$ es una estructura de incertidumbre

- $H : \mathcal{E} \to \mathbf{R}^+$ es una medida de incertidumbre sobre $\mathcal{H}$

2.3. PROPIEDADES ELEMENTALES DE LA MEDIDA DE INCERTIDUMBRE

(2.3.1) $J(\Omega) = 0$

$\Pi_A < \{A\} \to H(\Pi_A) \geq H(\{A\}) = J(A)$

(2.3.2) $H(\Pi_A \vee \Pi_B) \geq H(\Pi_A \vee \{B\}) \geq H(\{A\}\ \{B\}) \geq J(A \cup B)$

$H(\Pi_A \vee \Pi_B) \geq H(\{A\} \vee \Pi_B) \geq H(\{A\} \vee \{B\}) \geq J(A \cup B)$

(2.3.3) $H(\Pi_A \wedge \Pi_B) \geq H(\Pi_A \wedge \{B\}) \geq H(\{A\} \wedge \{B\}) \geq J(A \cap B)$

$H(\Pi_A \wedge \Pi_B) \geq H(\{A\} \wedge \Pi_B) \geq H(\{A\} \wedge \{B\}) \geq J(A \cap B)$

(2.3.4) $H(\Pi_A \vee \{\phi\}) = H(\Pi_A)$ (siendo $\Pi_A \vee \{\phi\} = \Pi_A$)

(2.3.2) y (2.3.3) son consecuencias de

$\Pi_A \vee \Pi_B < \Pi_A \vee \{B\} < \{A \cup B\}$

$\Pi_A \wedge \Pi_B < \Pi_A \wedge \{B\} < \{A \cap B\}$

2.4. EL AXIOMA DE INDEPENDENCIA

Es bueno examinar en detalle el axioma de independencia (1.1.3). Para que éste sea un verdadero axioma y no una simple definición de independencia entre parejas de particiones, es necesario que se puedan asignar "a priori" las particiones independientes y luego tomar en cuenta el axioma (1.1.3) cuando se define sobre $\mathcal{E}$ la aplicación H, ya que el axioma constituye una verdadera limitación en la asignación de las medidas de incertibumbre.

Por otro lado se puede actuar de manera diferente asignando la aplicación H sin tomar en consideración el axioma (1.1.3) y verificando "a posteriori" si la propiedad especificada por ese axioma se cumple. Claro que no siempre, en el caso que no se cumpla hay que desechar la función H como posibre medida de incertidumbre o si parece conveniente, modificar la selección de la clase $\mathcal{K}$ de las álgebras independientes.

Pero, ¿cómo quedan las cosas si no es razonablemente posible modificar la estructura de $\mathcal{K}$ y por otro lado es sumamente conveniente escoger como medida de incertidumbre una aplicación H incompatible con el axioma (1.1.3)?.

No hay salida de este "impase" en la formulación clásica; la medida H tiene que ser abandonada y se escogerá otra medida, aún si ésto complica notablemente el planteamiento y la solución del problema (o tal vez lo hace imposible). Pero a veces hay salida si nos atrevemos a modificar el axioma. En este sentido C. Bertoluzza y F. Barbaini [8] han propuesto sustituir la condición (1.1.3) por la siguiente[4/]:

$$(1.1.3') \quad (\Pi_A, \Pi_B) \in K \Longrightarrow H(\Pi_A \wedge \Pi_B) = \tilde{G}\,[H(\Pi_A), H(\Pi_B)]$$

donde la función $\tilde{G}(x,y)$, definida sobre el conjunto

$$\tilde{\Gamma}^{(2)} = \{(u,v) : u = H(\Pi_A),\ v = H(\Pi_B) \text{ y } (\Pi_A, \Pi_B) \in K\} \subset (\mathbf{R}^+)^2,$$

determina la Incertidumbre del producto de dos experiencia Π_A, Π_B, H-independientes en términos de $H(\Pi_A)$ y $H(\Pi_B)$.

La ley $\tilde{G}$ debe satisfacer el siguiente sistema de ecuaciones funcionales:

$$\begin{array}{lll} & \text{a.} & \tilde{G}\,(u, +\infty) = +\infty \\ & \text{b.} & \tilde{G}\,(u,v) = \tilde{G}\,(v,u) \\ (2.4.1) & \text{c.} & \tilde{G}\,[\tilde{G}(u,v),\omega] = \tilde{G}\,[u, \tilde{G}(v,\omega)] \\ & \text{d.} & u' < u'' \Longrightarrow \tilde{G}(u',v) \leq \tilde{G}(u'',v) \end{array}$$

4/ Esta sustitución tiene otra y más importante motivación que será aclarada cuando se examinen las medidas de incertidumbre componibles.

que se obtienen inmediatamente de las siguientes propiedades:

(2.4.2)

a. $\Pi_A \wedge \{\phi\} = \{\phi\}$

b. $\Pi_A \wedge \Pi_B = \Pi_B \wedge \Pi_A$

c. $(\Pi_A \wedge \Pi_B) \wedge \Pi_C = \Pi_A \wedge (\Pi_B \wedge \Pi_C)$

d. $\Pi'_A < \Pi''_A \implies \Pi'_A \wedge \Pi_B < \Pi''_A \wedge \Pi_B$

La función $\tilde{G}(x,y)$ tiene como dominio el conjunto $\tilde{\Gamma}^{(2)}$. Si $\tilde{\Gamma}_2 = (\mathbb{R}^+)^2 - (o,o)$ entonces $\tilde{G}$ tiene que satisfacer el sistema (2.4.1) para todos los valores de los argumentos u,v,w que aparecen en él.

Recíprocamente, si una función $\tilde{G}$ es solución del sistema (2.4.1) en todo $(\mathbb{R}^+)^2 - (o,o)$, puede ser utilizada como ley de composición cualquiera sea la clase $\mathbb{K}$ de las particiones independientes. En este sentido diremos que $\tilde{G}$ es una ley de independencia universal.

Bajo la hipótesis de continuidad, la solución de este sistema se consigue fácilmente utilizando un bien conocido teorema de Mostert y Shields [9] sobre representación de semigrupos. Ella está dada por el siguiente teorema.

TEOREMA 2.4.1

La forma continua más general de la ley de independencia $\tilde{G}(x,y)$ se obtiene escogiendo una secesión $\{\Delta_i = (\alpha_i, \beta_i)\}$ de intervalos abiertos disjuntos y una sucesión $\{\Psi_i : \Delta_i \to \mathbf{R}\}$ de funciones continuas estrictamente crecientes y haciendo

$$\tilde{G}(x,y) = x * y = \begin{cases} \gamma_i(\Psi_i(x) + \Psi_i(y) & \text{si } (x,y) \in \Delta_i^2 \\ \\ \text{Sup } (x,y) & \text{en otro caso} \end{cases}$$

donde γ_i es la seudoinversa de γ_i.

2.5. INCERTIDUMBRE CONDICIONAL

Para cualquier partición $\Pi_A = \{A_1, \ldots, A_n\}$ la cantidad $H(\Pi_A)$ mide la incertidumbre que se tiene sobre el resultado de una experiencia cuyos éxitos previstos son los eventos $A_1, \ldots, A_n$, sin saber nada de lo que puede acontecer.

Por otro lado, muchas veces antes de ejecutar la experiencia ya se conoce algo sobre los posibles éxitos y este conocimiento disminuye la incertidumbre. Típico es el caso donde: antes de efectuar la experiencia ya se sabe que el resultado será un evento elemental perteneciente al conjunto $A = \cup A_i$, por supuesto sin saber a cual de los éxitos A_1 ó $A_2 \ldots$ o A_n pertenecerá. En esta condición la incertidumbre sobre cuál será el éxito no será $H(\Pi_A)$ sino otra cantidad que indicaremos por $K(\Pi_A)$ y la llamaremos "incertidumbre condicional" de la partición Π_A.

¿Cómo medir esa nueva incertidumbre?

La primera definición ha sido propuesta por P. Benvenuti y B. Forte [10], quienes introdujeron la cantidad $K(\Pi_A)$ por medio de la relación

$$(2.5.1) \quad K(\Pi_A) = H(\Pi_A) - H(\{A\}) = H(\Pi_A) - J(A)$$

Esta definición, aunque sea bien aceptable en el contexto clásico, no pone en evidencia directamente las características fundamentales de la información condicional y por otro lado no es fácilmente generalizable a espacios donde la incertidumbre del producto de dos experiencias independientes no se conforma con el axioma clásico $H(\Pi_A \wedge \Pi_B) = H(\Pi_A) + H(\Pi_B)$.

Resulta más conveniente introducir esta noción en forma axiomática de manera que, la definición misma subraye las características esenciales. Adecuados teoremas de representación permitirán en lo sucesivo explicitar las formas de $K(\Pi_A)$ en los distintos espacios de incertidumbre.

Definición 2.5.1.

Sea $\mathbf{H} = (\mathbf{I} = (\Omega, \mathcal{S}, \mathcal{K}, J), \mathcal{E}, H)$ un espacio de incertidumbre. Se da el nombre de "incertidumbre condicional" asociada a ese espacio a cada aplicación:

$$K : \mathcal{E} \to \mathbb{R}^+$$

con las siguientes propiedades

1. $K : E \to \mathbf{R}^+$ es una medida de incertidumbre sobre un espacio $\mathbf{K}$ cuyos elementos $\Omega, S, K, *, E$ son los mismos del espacio $\mathbf{H}$.

2. $K(\{A\}) = 0$

3. $K(\Pi_A) = f[H(\Pi_A), J(A)]$

Las propiedades 1 y 2 se pueden sustituir por la siguiente forma más compacta.

1, $K : E \to \mathbb{R}^+$ es una medida de información sobre la estructura $(\mathbf{I}^*, E)$ donde en el espacio $\mathbf{I}^* = (\Omega, S, K, J^*)$, J^* es la medida de información trivial, $J^*(A) = 0$ $(\forall\ A \in S)$.

La primera representación de la incertidumbre condicional ha sido establecida por C. Bertoluzza [1] en espacios con ley de independencia clásica y está especificada por el siguiente teorema:

Teorema 2.5.1

Si en el espacio de información I, la ley de independencia es la clásica ($x * y = x + y$) y si además se pide que la función f sea continua y universal, entonces

(2.5.2) $f(x,y) = e(x-t)$

y por consiguiente

(2.5.3) $K(\Pi_A) = e\,[H(\Pi_A) - J(A)]$

Antes de desarrollar la demostración observamos que en las hipótesis dadas se tiene:

$$(\Pi_A,\ \Pi_B) \in \mathcal{K} \rightarrow K(\Pi_A \wedge \Pi_B) = K(\Pi_A) + K(\Pi_B)$$

$$H(\Pi_A \wedge \Pi_B) = H(\Pi_A) + H(\Pi_B) \quad \text{y}$$

$$J(A \cap B) = J(A) + J(B).$$

Demostración

Haciendo $\xi = H(\Pi_A) - J(A)$, $\eta = J(A)$ introducimos y definimos una nueva función incognita $f^*(\xi,\eta)$ por medio de la relación

(I) $f\,[H(\Pi_A),\ J(A)] = f^*\,[H(\Pi_A) - J(A),\ J(A)] = f^*\,(\xi,\eta)$

con lo que se tiene

(II) $K(\Pi_A) = f^*\,[H(\Pi_A) - J(A),\ J(A)]$

Demostraremos a continuación que $f^*(\xi,\eta)$ depende de la variable ξ y no de η.

En efecto:

Sea $\{B\}$ una partición independiente algebraicamente de Π_A y también de $\{A\}$; como $K(\{B\}) = 0$ se tiene

$$K(\Pi_A) = K(\Pi_A) + K(\{B\}) = K(\Pi_A \wedge \{B\})$$
$$= f^*[H(\Pi_A \wedge \{B\}) - J(A \cap B),\ J(A \cap B)]$$
$$= f^*[H(\Pi_A) + H(\{B\}) - J(A) - J(B),\ J(A) + J(B)]$$
$$= f^*[H(\Pi_A) - J(A)\ ,\ J(A) + J(B)]$$

Comparando este resultado con (II) vemos que K no depende de la segunda variable, siendo J(B) arbitraria para la universalidad[5/]. Por lo tanto podemos escribir

$$K(\Pi_A) = \tilde{f}\,(H(\Pi_A) - J(A))$$

Sea ahora (Π_A, Π_B) una pareja de particiones perteneciente a K y sea $\xi_1 = H(\Pi_A) = J(A)$ y $\xi_2 = H(\Pi_B) - J(B)$ entonces tenemos:

$$\tilde{f}(\xi_1) + \tilde{f}(\xi_2) = \tilde{f}\,(H(\Pi_A) - J(A)) + \tilde{f}(H(\Pi_B) - J(B))$$
$$= K(\Pi_A) + K(\Pi_B) = K(\Pi_A \wedge \Pi_B)$$
$$= \tilde{f}\,[H(\Pi_A \wedge \Pi_B) - J(A \cap B)]$$
$$= \tilde{f}\,[H(\Pi_A) + H(\Pi_B) - J(A) - J(B)]$$
$$= \tilde{f}\,[H(\Pi_A) - J(A) + H(\Pi_B) - J(B)]$$
$$= \tilde{f}\,(\xi_1 + \xi_2)$$

5/ En efecto, como quiera que sea escogido el número $x > 0$, siempre existe un espacio de incertidumbre $\mathcal{H}$ tal que:
- $x = J(B)$
- $(\Pi_A, \{B\}) \in K$, $(\{A\}, \{B\}) \in K$

Además, como f es universal, entonces debe ser compatible con este espacio y en consecuencia se tiene

$$K\,(\Pi_A \wedge B) = f^*(H(\Pi_A \wedge B) - J(A \cap B),\ J(A \cap B))$$

En definitiva

$$\tilde{f}(\xi_1) + \tilde{f}(\xi_2) = \tilde{f}(\xi_1 + \xi_2) \quad \forall\ \xi_1, \xi_2 \in \mathbb{R}^+$$ 6/

Esta es la ecuación funcional de Cauchy, la cual tiene como única solución continua a la función

$$\tilde{f}(\xi) = c.\xi$$

Por lo tanto cualquier entropía condicional clásica es siempre del tipo

$$K(\Pi_A) = c.\,[H(\Pi_A) - J(A)]$$ donde necesariamente $c \geq 0$ (ya que $K(\Pi_A) \geq 0$ y $H(\Pi_A) - J(A) \geq 0$)

Otro caso de condicionamiento es en el cual se mide la incertidumbre de una experiencia Π después de efectuar una experiencia previa Π'. La incertidumbre sobre cuál será el éxito de Π estará entonces condicionada por el éxito de la experiencia Π'.

Aunque no sea estrictamente vinculante, en lo que sigue supondremos que la partición condicionante Π' tenga soporte contenido en la partición condicionada Π. Esta hipótesis no es dema-

---

6/ Para cada ξ_1, $\xi_2 > 0$ siempre existe un espacio de incertidumbre $\mathcal{H}$ donde:

- $\xi_1 = H(\Pi_A) - J(A)$, $\xi_2 = H(\Pi_B) - J(B)$
- $(\Pi_A, \Pi_B) \in K$, $(\{A\}, \{B\}) \in K$

siado limitativa porque los eventos elementales que no pertenecen a elementos de Π pueden en muchos casos ser eliminados sin modificar la parte que realmente condiciona nuestro conocimiento. En otras palabras, si nosotros estamos interesados en la partición Π podemos ignorar todos los eventos elementales que no le pertenecen.

Poniéndonos como supervisores de ambas experiencias, nos daremos cuenta que cualquiera sea el éxito de la primera experiencia Π', éste modifica nuestros conocimientos sobre la segunda experiencia Π; por lo tanto, saber cómo está estructurada la experiencia Π', influye sobre la evaluación que atribuiremos a la incertidumbre de Π.

Hablaremos de incertidumbre condicional de la partición Π respecto a la partición Π', denotando esta cantidad por $H(\Pi|\Pi')$. Pues la incertidumbre condicional será una aplicación:

$$\overline{H} \; : \mathcal{E} \times \mathcal{E} \to \mathbb{R}^+$$

a la cual es natural imponer las siguientes condiciones:

a. La restricción de $\overline{H}$ a las parejas del tipo $(\Pi_A, \{A\})$ determina una aplicación

$$K \; : \mathcal{E} \to \mathbb{R}^+$$

que debe ser una medida de incertidumbre sobre $(I, \mathcal{E})$.

b. La incertidumbre de Π_A relativa a cualquiera de sus refinamientos tiene que ser nula.

Las condiciones a. y b. se traducen en las siguientes relaciones:

$$\text{a')} \quad \Pi''_A < \Pi'_A \implies H(\Pi''_A|\{A\}) \geq H(\Pi'_A|\{A\})$$

$$(2.5.4) \quad \text{a'')} \quad (\Pi_A, \Pi_B) \in K \implies H(\Pi_A \wedge \Pi_B|\{A \cap B\}) = H(\Pi_A|\{A\}) \times H(\Pi_B|\{B\})$$

$$\text{b)} \quad \Pi'_A < \Pi_A \implies H(\Pi_A|\Pi'_A) = 0$$

Si queremos definir $H(\Pi|\Pi')$ de manera que sea coherente con el espacio de incertidumbre (I, E, H) es oportuno que la aplicación $\overline{H} : E \times E \to \mathbf{R}^+$ dependa de sus argumentos Π y Π' sólo a través de medidas de informaciones absolutas asociadas de cualquier manera a esas particiones. En forma más precisa, resulta natural pensar que sea

$$(2.5.5) \qquad H(\Pi|\Pi') = F\,[H(\Pi \wedge \Pi'), H(\Pi')]$$

En efecto, aunque a primera vista podría parecer natural pedir que $H(\Pi|\Pi')$ dependa de $H(\Pi)$ y $H(\Pi')$ no hay que olvidar el hecho que: si se conoce la estructura de la partición Π' y de la condicionada Π, también se conoce la partición $\Pi \wedge \Pi'$ y si ya se conoce un éxito A'_j de Π', la incertidumbre que queda es la incertidumbre sobre la partición $A_1 \cap A'_j, A_2 \cap A'_j, \ldots, A_n \cap A'_j$. Por esto el supervisor que pueda observar ambas particiones se da cuenta que la incertidumbre relativa depende de la partición condicionante y de su producto con la partición condicionada.

Por otro lado la hipótesis (2.5.5) es compatible con la propiedad de aditividad fuerte de la entropía de Shannon ($H(\Pi_A \mid \Pi_B) = H(\Pi_A \wedge \Pi_B) - H(\Pi_A)$) y además tiene analogía con la definición de probabilidad condicional ($p(A|B) = p(A\cap B)| p(B) = f[p(A\cap B), P(B)]$).

Como en la hipótesis el soporte de Π' esta contenido en el soporte de Π, entonces tenemos que $H(\Pi' \wedge \Pi) \geq H(\Pi')$

Luego F está definida en un subconjunto de $\Gamma_2 := \{(x,y): x \geq y\}$.
Si además queremos que la función F tenga carácter universal entonces las relaciones (2.5.4) se traducen en el siguiente sistema de ecuaciones funcionales en la incógnita $F(x,y)$.

(2.5.6)

a. $x_1 < x_2 \rightarrow F(x_1,y) \leq F(x_2,y) \quad \forall\ x_1,x_2 \geq y$

b. $F(x_1 * x_2,\ y_1 * y_2) = F(x_1,y_1) * F(x_2,y_2) \ \forall\ x_1 \geq y_1, x_2 \geq y_2$

c. $F(x,x) = 0$

d. $\mathrm{Dom}\, F = \{(x,y) : x \geq y\} = \Gamma_2$,

siendo $*$ la ley de independencia.

Para enunciar y demostrar el teorema que establece la solución general de este sistema (y por consiguiente la forma de $H(\Pi|\Pi')$ usaremos las siguientes notaciones:

$L = \{(x,y) : y < x\}$, $L \subset (\mathbb{R}^+)$

$\overline{L} = \{(x,y) : y = x\}$

$a = \text{Sup}\ \{x : x \in \Lambda_*\}$ $(a \leq +\infty)$, siendo Λ_* el conjunto de los elementos idempotentes.

(obviamente $\Gamma_2 = \text{Dom } F = L \cup \overline{L}$

Teorema de Representación 2.5.2

Dado un espacio de incertidumbre $H = (\mathbf{I}, E, H)$ cualquier función $F(x,y)$ apta para determinar la información relativa de experiencias es del tipo:

$$(2.5.7)\quad F(x,y) = \begin{cases} 0 & \forall\ (x,y) \in ([0,a]^2 \cap L) \cup \overline{L} \\ f[\,g(x) - g(\text{Sup}(y,a))] & \forall\ (x,y) \in L - [0,a]^2 \end{cases}$$

donde:

- $a = \text{Sup}\ \{x \in \Lambda_*\ \ x < +\infty\} = \sup\ \{x : x * x = x\}$
- g es la función que define la restricción de la ley de independencia "$*$" al subconjunto $[a, +\infty]^2$.
- α es un elemento cualquiera $*$-idempotente ($\lambda \in \Lambda_*$ y obviamente $\lambda \leq a$).
- $\beta = \text{Inf}\ \{\xi : \xi \in \Lambda_*, \ \xi > \lambda\}$
- $f(u) = \overline{h}(u,k)$ si $\beta > \lambda$

  $f(u) = \beta = \lambda$ si $\beta = \lambda$
- h es la función que define la restricción de la ley de independencia "$*$" al cuadrado $[\alpha,\beta]^2$ y $\overline{h}$ es su seudoinversa.

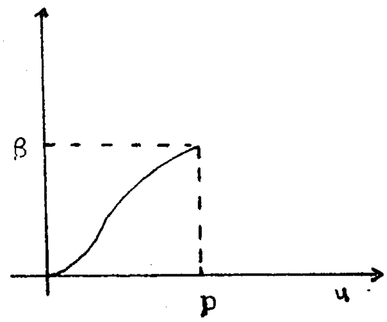
- k es una constante arbitraria no negativa.

Los elementos arbitrarios para determinar la función f(u) (y por consiguiente F(x,y)) son el $*$-idempotente λ y el número positivo K (sólo si $\beta > \lambda$). Para ilustrar mejor la forma de f(u) en este último caso vamos a introducir el número

$$p = h(\beta)/K \qquad (K = h(\beta)/p$$

Recordando cómo ha sido definida la función seudoinversa, se reconoce fácilmente que

$$f(u) = \begin{cases} h^{-1}(ku) & \text{si } u \leq p \\ \beta & \text{si } u > p \end{cases}$$



(para la demostración ver Apéndice A).

Observación 2.5.1.

Cualquiera sea el espacio H, escogiendo $\lambda = 0$ y haciendo eventualmente $k = 0$ el sistema (2.5.6) es siempre satisfecho por la función idénticamente nula.

Observación 2.5.2.

La función $f(u)$ es siempre una función continua en $(0, +\infty)$. La función $F(x,y)$ es casi-continua, pudiendo presentar discontinuidad para $x = y$. Tal discontinuidad puede ser eliminada escogiendo $p = 0$ y determinando $f(u)$ en consecuencia de ésto, en tal caso se podrá obtener para la función $F(x,y)$ una expresión no idénticamente nula sólo si 0 es para Λ un punto aislado.

Caso Particular

Puede ser adecuado hacer uso de una medida de información relativa según la cual para cualquier pertición Π de Ω se tiene $H(\Pi|\{\Omega\}) = H(\Pi)$. Las funciones $F(x,y)$ dadas en el teorema de representación corresponden a las que satisfacen la siguiente condición adicional

$$(2.5.8) \qquad F(x,o) = x$$

Se reconoce inmediatamente que funciones de tal tipo existen (además existe sólo una) si y sólo si $\Lambda_* - \{o\} = \phi$, es decir, si $a = 0$. En caso contrario para cada $x \in (o,a]$ resulta $F(x,o) = 0$.

Si $\Lambda - \{o\} = \phi$ $(\Lambda = \{o\})$ entonces $(o, +\infty) = \overline{\Lambda}$; por lo tanto $h(x)$ coincide con $g(x)$ y $\overline{h}(x)$ y $g^{-1}(x)$.

Entonces la forma que asume el teorema de representación es la siguiente:

$$(2.5.9) \quad F(x,y) = \begin{cases} 0 & \forall\ (x,y) \in \overline{L} \\ g^{-1}\{k\,[g(x) - g(y)]\} & \forall\ (x,y) \in L \end{cases}$$

que satisface (2.5.8) si y sólo si k = 1

En el caso particular tratado se obtiene para la información relativa:

$$(2.5.10) \quad H(\Pi|\Pi') = \begin{cases} 0 & \text{si } \Pi' < \Pi \\ g^{-1}\{g\,[H(\Pi \wedge \Pi')] - g\,[H(\Pi')]\} & \text{en otro caso} \end{cases}$$

Observación 2.5.3

Si para cada $\Pi_A \in E$ la partición condicionante Π' es la partición $\{A\}$ constituida por un sólo subconjunto (el soporte de Π_A) entonces la cantidad $H(\Pi_A|\{A\})$ coincide con la información condicional de Π_A de la definición 1.2.1.

$$H(\Pi_A|\{A\}) = K(\Pi_A)$$

En efecto se reconoce de inmediato que las propiedades (2.5.4) y (2.5.5) se traducen en las 1.2.3 de esa definición. El teorema de representación precedentemente enunciado permite entonces explicitar también la forma de la incertidumbre condicional, que resulta ser la siguiente:

$$K(\Pi_A) = \begin{cases} 0 & \text{si } H(\Pi_A) \leq a \\ f\,[g\{H(\Pi_A)\} - g\{\sup\,(a, J(A))\}] & \text{si } H(\Pi_A) > a \end{cases}$$

y en el caso particular

$$K(\Pi_A) = g^{-1}\{g\,[H(\Pi_A)] - g\,[J(A)]\}$$

Asi queda completamente determinada la entropía condicional en cualquier espacio de incertidumbre.

2.6. CLASIFICACION DE LAS MEDIDAS DE INCERTIDUMBRE

Los axiomas que definen las medidas de incertidumbre, sea en el sentido clásico o en el generalizado, son muy débiles. En efecto, el único axioma que en realidad impone limitaciones absolutas (es decir limitaciones a todas las posibles entropías) es el axioma de monotonía (además de las condiciones extremas $H(\{\Omega\}) = J(\Omega) = 0$, $H(\{\phi\}) = J(\phi) = +\infty$). El axioma de aditividad, que parece ser el más limitativo, realmente en forma absoluta no constituye verdaderas restricciones en su formulación generalizada siendo sumamente arbitraria la selección de la ley de independencia, pero tampoco lo es en la formulación clásica porque su poder limitativo depende de la consistencia de la clase $\mathcal{K}$ de las álgebras H-independientes y si acaso $\mathcal{K}$ es vacía no hay ni una pizca de restricción.

Por eso es extremadamente importante poner un poco de orden en la extensa familia de los espacios de incertidumbre.

Esta operación ha sido iniciada por B. Forte [11] quien propuso ordenar las entropías en clases, cada una caracterizada por una propiedad específica y además seleccionó una primera clase de estas propiedades de manera que cada una de éllas tuviera un sentido bien evidente con respecto al desarrollo de la teoría y/o con respecto a las posibles aplicaciones.

Las clases hasta ahora particularizadas son las siguientes:

1. Medidas ramificadas y medidas g-locales
2. Medidas componibles y totalmente componibles
3. Medidas idempotentes
4. Medidas con la propiedad de la suma o casi-suma
5. Medidas subaditivas.

Las clases de entropías 1 y 2 las estudiaremos en detalle en los próximos capítulos (III y IV), siendo esas clases sin duda las más estudiadas y las más importantes. Aquí proporcionaremos algunas indicaciones sobre las demás clases.

Definición 2.6.1.

Diremos que una medida de incertidumbre H sobre un espacio $\mathcal{H} = (\mathbf{I}, \mathcal{E})$ es "idempotente" si se cumple:

(2.6.1) $(A \cap B) = \phi \; ; \; H(\Pi_A) = H(\Pi_B) = h \Longrightarrow H(\Pi_A \vee \Pi_B) = h$.

Esta medida no ha sido prácticamente estudiada, hasta ahora hay pocos resultados relacionados con la noción de idempotencia y todos son elementales.

Resultados

1. Si para cada pareja Π_A y Π_B (con $A \cap B = \phi$) se cumple:

   $\mathrm{Inf}[H(\Pi_A), H(\Pi_B)] \leq H(\Pi_A \vee \Pi_B) \leq \mathrm{Sup}[H(\Pi_A), H(\Pi_B)]$, entonces la entropía H es idempotente. La demostración es obvia porque $x = y \Longrightarrow \mathrm{Inf}(x,y) = \mathrm{Sup}(x,y)$.

2. Sea $I = \{\alpha, \beta \ldots .\}$ un conjunto finito de índices y sea Π_{A_ξ}, $\xi \in I$ una colección finita de particiones. De (2.6.1) se deduce inmediatamente que,

   Si para cada $\alpha, \beta \in I$ $A_\alpha \cap A_\beta = \phi$ entonces se cumple

   (2.6.2) $\{H(\Pi_{A_\alpha}) = h, \; \forall \alpha \in I\}$ $H(\bigvee_{\alpha \in I} \Pi_{A_\alpha}) = h$.

3. Si I no es finito en general (2.6.2) no se cumple. Sin embargo hay particulares medidas de incertidumbre para las cuales (2.6.2) se cumple también cuando I no es finito.

Llamaremos σ-idempotentes (respectivamente c-idempotentes) a las entropías que satisfacen (2.6.2) para todo I numerable (respectivamente continuos).

4. Si para todo I numerable (respectivamente continuo) se cumple

$$\operatorname*{Inf}_{\alpha\in I} [H(\Pi_{A_\alpha})] \leq H(\bigvee_{\alpha\in I} \Pi_{A_\alpha}) \leq \operatorname*{Sup}_{\alpha\in I} [H(\Pi_{A_\alpha})]$$

entonces la medida de incertidumbre H es σ-idempetente (resp. c-idempotente).

(Naturalmente, para que $\bigvee_{\alpha\in I} \Pi_{A_\alpha}$ sea definible es necesario que $A_\alpha \cap A_\beta = \phi \quad \forall\ \alpha,\beta \in I$).

5. La propiedad de idempotencia tiene una importante interpretación termodinámica. En efecto interpretando H como entropía de un gas, (2.6.1) afirma que si dos gases tienen entropías iguales, entonces su mezcla tiene la misma entropía.

Definición 2.6.2.

Diremos que una medida de incertidumbre H "posee la propiedad de la suma o de la casi-suma" si se cumple respectivamente, para cada

$$\Pi_A = \{A_1,\dots,A_n\} \in E$$

$$(2.6.3)\quad H(\{A_1,\ldots,A_n\}) = \sum_{i=1}^{n} g(A_i)$$

$$(2.6.4)\quad H(\{A_1,\ldots,A_n\}) = f^{-1}\left[\sum_{i=1}^{n} f\{g(A_i)\}\right]$$

donde $g : S \to \mathbb{R}^+$ es una aplicación con las siguientes propiedades

i) $g(\Omega) = 0$

ii) $A \cap B = \phi \Longrightarrow g(A) + g(B) \geq g(A \cup B)$

iii) $f : \mathbb{R}^+ \to \mathbb{R}^+$ es una función estrictamente creciente con $f(o) = 0$.

Resultados:

1. Haciendo $h(A) = f[g(A)]$ la propiedad de casi aditividad (2.6.4) se puede escribir bajo la forma

$$(2.6.5)\quad H(\{A_1,\ldots,A_n\}) = f^{-1}\left[\sum_{i=1}^{n} h(A_i)\right]$$

donde $h : S \to \mathbb{R}^+$ tiene las mismas propiedades de la función g ; luego (2.6.4) y (2.6.5) son equivalentes.

Las propiedades de suma y casi suma han sido estudiadas por J. Aczel, Z. Deroczy, C.F.Picard, I. Vander Vyl [12, 13, 14] en relación con las entropías dependientes de las probabilidades. En el área de la teoría axiomática no hay trabajos específicos : sólo hay dos resultados parciales obtenidos estudiando otras propiedades.

2. B. Forte y C.T.Ng [15] han demostrado que si $H(\{A_1,..,A_n\}) = K[J(A_i),...,J(A_n)]$ entonces la propiedad de la suma coincide con la ramificación. Ilustraremos este resultado en el capítulo III.

3. C. Bertoluzza, M. Schneyder y M.L. Capodieci han demostrado (Ver capítulo IV) que las medidas de incertidumbre con la propiedad de casi suma coincide con una particular subclase de las entropías componibles.

Definición 2.6.3.

Se dice que una medida de incertidumbre es "subaditiva" si cumple con la condición:

$$(2.6.6) \qquad H(\Pi_A \vee \Pi_B) \leq H(\Pi_A) + H(\Pi_B)$$

Estas entropías han sido poco estudiadas por sí solas, siempre se estudiaron junto con otras propiedades (aditividad, aditividad fuerte, ramificación) y casi siempre con relación a entropías dependiendo de las probabilidades. En todos los trabajos hechos hasta ahora se ha llegado a la entropía de Shannon o de Shannon-Hartley

$$(2.6.7) \qquad H(\Pi_A) = - a \Sigma pi \log pi + b \log n$$

CAPITULO III

MEDIDAS DE INCERTIDUMBRE CON LA PROPIEDAD DE RAMIFICACION

En este capítulo analizaremos la clase de medidas de incertidumbre que poseen la propiedad de localidad (de una u otra forma, clásica o generalizada). Esa propiedad está relacionada con la construcción de la entropía $H(\Pi_A)$ a través de un proceso de subdivisiones sucesivas, partiendo del soporte A hasta llegar a la partición Π_A (veremos ésto con detalle más adelante); dicho proceso es uno de los dos más típicos utilizado para plantear experiencias.

Procesos de esas clases se usan en las contrucciones de cuestionarios, en la programación estructurada, etc; claro que si se quiere medir informaciones en estas situaciones, será preferible utilizar las medidas locales.

En la primera parte de este capítulo será analizada una noción clásica de localidad, es decir, la propiedad de ramificación; en la segunda parte, después de subrayar los límites de la noción clásica se propone y se analiza una generalización propuesta recientemente. Concluye el capítulo la exposición de un método iterativo que permite construir medidas de incertidumbre $H(\Pi_A)$ a partir de oportunas entropías elementales.

3.1. Definición 3.1.1.

Diremos que una medida de incertidumbre posee la **propiedad de ramificación** si se cumple una de las dos propiedades siguientes, equivalentes entre sí,

(3.1.1) $H(\Pi_A \vee \Pi_B) = H(\Pi_A \vee \{B\}) + H(\{A\} \vee \Pi_B) - H(\{A\} \cup \{B\})$

Para cada Π_A, $\Pi_B \in E$ con $A \cap B = \phi$,

(3.1.2) $H(\{A_1, A_2, A_3, \ldots, A_n\}) - H(\{A_1 \cup A_2, A_3, \ldots, A_n\}) =$

$$= G(\bigcup_{j=3}^{n} A_j, A_1, A_2) \text{ para cada } \Pi = \{A_1, \ldots, A_n\} \in E$$

Naturalmente, como Π no depende del orden de sus elementos A_i, la relación (3.1.2) implica

(3.1.3) $H(\{A_{j_1}, A_{j_2}, \ldots, A_{j_n}\}) - H(\{A_{j_1} \cup A_{j_2}, \ldots, A_{j_n}\}) = G(\bigcup_{i=3}^{n} A_{j_i}, A_{j_i}, A_{j_2})$

para cada permutación $j_1 \ldots j_n$ de los índices $1, \ldots n$.

Demostraremos ahora la equivalencia de las propiedades (3.1.1) y (3.1.2) probando las dos implicaciones

(3.1.1) $\Longrightarrow$ (3.1.2) y (3.1.2) $\Longrightarrow$ (3.1.1)

(3.1.1) $\Longrightarrow$ (3.1.2)

La partición $\Pi = \{A_1, A_2, A_3, \ldots, A_n\}$ es la unión de las dos particiones (disjuntas) $\Pi_A = \{A_1, A_2\}$ y Π_B $\{A_3, \ldots, A_n\}$. De la propiedad (3.1.1) se obtiene

$$H(\{A_1,A_2\} \vee \{A_3,\ldots,A_n\}) = H(\{A_1,A_2\} \vee \{\bigcup_{i=3}^{n} A_i\}) + H(\{A_1 \cup A_2\} \vee \{A_3,\ldots,A_n\})$$

$$- H(\{A_1 \cup A_2\} \vee \{\bigcup_{i=3}^{n} A_i\})$$

que equivale a la relación (3.1.2) después de poner

$$G(\bigcup_{i=3}^{m} A_j;\ A_1,\ A_2) = H(\{A_1,A_2\} \vee \{\bigcup_{j=3}^{m} A_j\}) - H(\{A_1 \cup A_2\} \vee \{\bigcup_{j=3}^{n} A_j\})$$

□

(3.1.2) $\Longrightarrow$ (3.1.1)

Sea $\Pi_A = \{A_1, \ldots, A_n\}$, $\Pi_B = \{B_1, \ldots, B_m\}$; obviamente $\Pi_A \vee \Pi_B = \{A_1, \ldots, A_n, B_1, \ldots, B_m\}$. Utilizando (3.1.3) y haciendo $A = \bigcup_{i=1}^{n} A_i$, $B = \bigcup_{j=1}^{m} B_j$ podemos escribir

$$(3.1.4)\quad H(\Pi_A \vee \Pi_B) = H(\{A_1,\ldots,A_n, B_1,\ldots,B_m\})$$

$$= H(\{A_1,\ldots,A_n, B_1 \cup B_2, B_3,\ldots,B_m\})$$

$$+ G([A \cup (\bigcup_{i=3}^{m} B_i)]\ ,\ B_1,\ B_2)$$

$$= H(\{A_1,\ldots,A_n, B_1 \cup B_2 \cup B_3, B_4,\ldots,B_m\})$$

$$+ (G\ [A \cup (\bigcup_{i=4}^{m} B_i)]\ ,\ B_1 \cup B_2,\ B_3)$$

$$+ G([A \cup (\bigcup_{i=3}^{m} B_i)] , B_1 , B_2)$$

$$= \; . \; . \; . \; . \; . \; . \; .$$

$$= H(\{A_1,..,A_n, B_1 \cup B_2 \cup ... \cup B_m\})$$

$$+ \sum_{r=1}^{m-1} G([A \cup (\bigcup_{i=r+2}^{m} B_i)] , \bigcup_{i=1}^{r} B_i , B_{r+1})$$

$$= H(\Pi_A \text{ v } \{B\}) + \sum_{r=1}^{m-1} G([A \cup (\bigcup_{i=r+1}^{m} B_i)] , \bigcup_{i=1}^{r} B_i , B_{r+1})$$

Por otro lado se tiene

$$(3.1.5) \quad H(\{A\} \text{ v } \Pi_B) = H(\{A, B_1, B_2, B_3, ..., B_m\})$$

$$= H(\{A, B_n \cup B_2, B_3, ..., B_m\})$$

$$+ G([A \cup (\bigcup_{i=3}^{m} B_i)] ; B_1, B_2)$$

$$= \; . \; . \; . \; .$$

$$= H(\{A\} \text{ v } \{B\}) = \sum_{r=1}^{m-1} G([A \cup (\bigcup_{i=r+2}^{m} B_i)] , \bigcup_{i=1}^{r} B_i , B_{r+1})$$

Confrontando (3.1.4) y (3.1.5) se obtiene

$$H(\Pi_A \text{ v } \Pi_B) - H(\Pi_A \text{ v } \{B\}) = H(\{A\} \text{ v } \Pi_B) - H(\{A\} \cup \{B\})$$

es decir (3.1.1).

Las propiedades (3.1.1) y (3.1.2) tienen una interpretación sencilla y muy significativa. En efecto, éllas afirman que al pasar de una partición $\{A_1 \cup A_2, A_3, .., A_n\}$ a su refinamiento $\{A_1, A_2, A_3, .., A_n\}$ el aumento de incertidumbre que se produce, $G(\bigcup_{i=3}^{n} A_i, A_1, A_2)$ (siempre positiva por el axioma de monotonía), depende sólo de la parte que ha sido refinada (carácter local de la incertidumbre).

Observación 3.1.1.

Utilizando el mismo procedimiento que se utilizó para demostrar (3.1.4) y (3.1.5) se logra fácilmente llegar a la relación siguiente:

$$(3.1.6)\quad H(\Pi_A) = H(\{A_1, .., A_n\})$$
$$= H(\{\bigcup_{i=1}^{n-1} A_i, A_n\}) + \sum_{r=1}^{n-2} G(\bigcup_{i=r+2}^{n} A_i, \bigcup_{i=1}^{r} A_i, A_{r+1})$$

Por otro lado, como $\{X,Y\} = \{X,Y,\phi\}$

$$H(\{X,Y\}) = H(\{X,Y,\phi\}) = H(\{X \cup Y\} \vee \{\phi\}) + G(\phi,X,Y)$$

y luego $H(\{\bigcup_{i=1}^{n-1} A_i, A_n\}) = H(\{\bigcup_{i=1}^{n} A_i\}) + G(\phi, \bigcup_{i=1}^{n-1} A_i, A_n)$.

Haciendo formalmente $\bigcup_{i=n+1}^{n} A_i = \phi$, (3.1.6) se puede entonces escribir bajo la forma

$$(3.1.7)\quad H(\Pi_A) = J(A) + \sum_{r=1}^{n-1} G(\bigcup_{i=r+2}^{n} A_i, \bigcup_{i=1}^{r} A_i, A_{r+1})$$

Pues cada medida de incertidumbre ramificada está completamente determinado por el conocimiento del espacio de información **I** y de la aplicación

$$G : \mathcal{S}^3 \to \mathbf{R}^+$$

3.2. PROPIEDADES DE G

Para que el segundo miembro de (3.1.7) representa una medida de incertidumbre; la aplicación G no puede ser arbitraria: tiene que satisfacer algunas condiciones. En particular

(3.2.1)

a. $G(A, B, C) \geq 0$

b. $G(A, B, C) = G(A, C, B)$

c. $G(A, \phi, C) = 0$

d. $G(C \cup D, A, B) + G(D, A \cup B, C) = G(B \cup D, A, C) + G(D, A \cup C, B)$

e. $(\{A_1, A_2, A_3\}, \{B_1, B_2, B_3\}) \in \mathcal{K},\ x * y = x + y \Longrightarrow$

$$G(A_1, A_2, A_3) = \sum_{r=1}^{3} G([A \cap B - (A_1 \cup A_2) \cap B_r\ , A_2 \cap B_r, A_3 \cap B_r)$$

Las tres primeras condiciones son consecuencias inmediatas del axioma de monotonía, de la simetría de H y de la propiedad (2.3.4).

La condición d. se demuestra evaluando por medio de la expresión (3.1.7) las dos entropías $H(\{A,B,C,D\})$ y $H(\{A,C,B,D\})$ (iguales porque $\{A,B,C,D\} = \{A,C,B,D\}$) de donde se obtiene

$$H(\{A,B,C,D\}) = J(A\cup B\cup C\cup D) + G(C\cup D,A,B) + G(D,A\cup B,C)$$
$$+ G(\phi,A\cup B\cup C,D)$$

$$H(\{A,C,B,D\}) = J(A\cup B\cup C\cup D) + G(B\cup D,A,C) + G(D,A\cup C,B)$$
$$+ G(\phi, A\cup B\cup C,D)$$

En fin para demostrar e. se utiliza el axioma de independencia, después de recordar que $(\Pi_A; \Pi_B) \in K$ significa que son H-independientes las álgebras generadas por Π_A y Π_B. Haciendo en (3.2.1) e. $\{A_1,A_2,A_3\} = \Pi_A$, $\{B_1,B_2,B_3\} = \Pi_B$, luego tenemos

(3.2.2) $H(\Pi_A \wedge \Pi_B)$

$$= H(\{A_3\cap B_3,A_2\cap B_3, A_3\cap B_2,A_2\cap B_2,A_3\cap B_1,A_2\cap B_1,A_1\cap B_3, A_1\cap B_2, A_1\cap B_1\})$$

$$= H(\{B_3\cap(A_3\cup A_2), A_3\cap B_2,\ldots,A_1\cap B_1\}) + G([A\cap B - (A_2\cup A_3)\cap B_3], A_2\cap B_3, A_3\cap B_3)$$

$$= H(\{B_3\cap(A_2\cup A_3), B_2\cap(A_2\cup A_3), A_3\cap B_1,\ldots,A_1\cap B_1\})$$
$$+ G([A\cap B - (A_2\cup A_3)\cap B_2], A_2\cap B_2, A_3\cap B_2)$$
$$+ G([A\cap B - (A_2\cup A_3)\cap B_3], A_2\cap B_3, A_3\cap B_3)$$

$$= H(\{B_3\cap(A_2\cup A_3), B_2\cap(A_2\cup A_3),B_1\cap(A_2\cup A_3), A_1\cap B_n,A_1\cap B_2,A_1\cap B_1$$
$$+ \sum_{k=1}^{3} G([A\cap B - (A_2\cup A_3)\cap B_k], A_2\cap B_k, A_3\cap B_k)$$

$$= H(\{A_1, A_2 \cup A_3\} \wedge \Pi_B) + \sum_{k=1}^{3} G([A \cap B - (A_2 \cup A_3) \cap B_k], A_2 \cap B_k, A_3 \cap B_k)$$

$$= H(\{A_1, A_2 \cup A_3\}) + H(\Pi_B) + \sum_{k=1}^{3} G([A \cap B - (A_2 \cup A_3) \cap B_k], A_2 \cap B_k, A_3 \cap B_k)$$

por otro lado se tiene

$$(3.2.3) \quad H(\Pi_A \wedge \Pi_B) = H(\Pi_A) + H(\Pi_B) = H(\{A_1, A_2, A_3\}) + H(\Pi_B)$$

Comparando (3.2.2) y (3.2.3) se obtiene

$$\sum_{k=1}^{3} G([A \cap B - (A_2 \cup A_3) \cap B_k], A_2 \cap B_k, A_3 \cap B_k) = H(\{A_1, A_2, A_3\}) - H(\{A_1, A_2 \cup A_3\})$$
$$= G(A_1, A_2, A_3)$$

No es difícil demostrar que si G satisface las propiedades (3.2.1), entonces la cantidad

$$(3.2.4) \quad H(\Pi_A) := J(A) + \sum_{r=1}^{n-1} G\left(\bigcup_{i=r+2}^{n} A_i, \bigcup_{i=1}^{r} A_i, A_{r+1}\right)$$

cumple con todas las propiedades de una medida de incertidumbre ramificada con ley de independencia clásica. (Ver [16]).

No nos adelantaremos más en analizar en detalle este enfoque de las medidas de incertidumbre porque la definición de ramificación (propiedad local) solo es compatible, en sentido universal, con las entropías clásicas (donde $x * y = x + y$).

De manera más precisa, si en el subespacio $(\mathbf{I},E,H)$ la clase K es bastante "rica" y si la ley de independencia $*$ no es la ley clásica "+", seguramente la propiedad de ramificación no se cumple como quiera se escoja la aplicación (entropía) H (la demostración de este hecho está implícitamente contenida en el Teorema de representación de C. Poggi [17] que enunciaremos y demostraremos más adelante).

El estudio de las nociones relacionadas con algunas características de la propiedad de ramificación son demasiado importantes para que no se intente introducirlas también en espacios de incertidumbre no clásicos. Sin embargo, antes de proponer la generalización que permite esta extensión es necesario señalar un Teorema debido a B. Forte y C.T. Ng [15] quizás el más importante establecido con respecto a las medidas de incertidumbre ramificadas.

Teorema 3.2.1.

Sea $\mathbf{I}$ un espacio de información, con medida de información componible J. Sea H una entropía con la propiedad de ramificación sobre E_S. Suponemos que H es compatible con J, o sea que

(3.2.5) $H(\{A_1,\ldots,A_n\}) = H_n(J(A_1),J(A_2),\ldots,J(A_n))$.

Sea $D_n := \{(J(A_1),\ J(A_2),\ldots,J(A_n))\ /\ A_1,A_2,\ldots,A_n \in S\}$

y supongamos:

a. $\forall\, x \in \mathbb{R}^+,\ x > 0,$

$\exists!\ \partial x \in \mathbb{R}^+$ t.q $(x, \partial x) \in D_2$ y $x \,T\, \partial x = 0$

b. $x\,Ty > 0$ y $z \geq \partial(xTy) \Longrightarrow (x,y,z) \in D_3$

Entonces existe una aplicación $f : \mathbb{R}^+ \to \mathbb{R}$ con las propiedades

c. $f(+\infty) = 0,\ f(xTy) \leq f(x) + f(y)$, tal que

$$(3.2.6)\quad H(\Pi_A) = H_n\,[J(A_1),\ J(A_2), \dots\ J(A_n)] =$$

$$= -f(o) + \sum_{i=1}^{n} f[J(A_i)] \quad n \geq 1$$

Para la demostración ver [15]

Comentario: Las hipótesis a. y b. implican en particular que

$$x > 0 \text{ y } z \geq \partial x \Longrightarrow (x,z) \in D_2$$

Con este Teorema se demuestra que cualquier entropía que goza de la propiedad de ramificación y poco más tiene necesariamente la forma (3.2.6), siempre que J sea una medida de información, componible por medio de una operación binaria regular. En particular la entropía de Shannon (donde $J(A) = -\log p(A)$), tiene la forma (3.2.6) con $f(x) = x\,e^{-x}$.

Por otro lado cada aplicación H: $E \rightarrow \mathbf{R}^+$ definida por (3.2.6) [f satisfaciendo las propiedades c.] es una medida de incertidumbre con la propiedad de ramificación de manera que la propia ramificación [con la propiedad (3.2.5)] es condición necesaria y suficiente para que H sea de la forma (3.2.6).

Es importante destacar que la representación (3.2.6) sólo se refiere a la clase de las experiencias completas; para las experiencias incompletas solo se puede afirmar que

1. Las medidas de incertidumbre definidas por (3.2.6) gozan de la propiedad de localidad.

2. Son de la forma (3.2.6) todas las medidas de incertidumbre compatibles (en el sentido (3.2.5)) con espacios de información triviales [tales que $J(A) \equiv 0$].

3.3. MEDIDAS DE INCERTIDUMBRE g-LOCALES

Examinada en detalle, la propiedad de ramificación afirma lo que sigue:

Consideremos una partición $\Pi = \{A_1,\dots,A_n, B\}$ y luego realicemos una subdivisión de uno de los eventos (por ejemplo B) en un cierto número de subconjuntos $B_1 \dots B_m$ (con $B_i \cap B_j = \phi$, $\cup B_i = B$); se obtiene así una nueva partición $\Pi' = \{A_1 \dots, A_n, B_1, \dots, B_m\}$ que podemos considerar como unión de las particiones $\Pi_A = \{A_1,\dots,A_n\}$ y $\Pi_B = \{B_1,\dots,B_m\}$.

La relación (3.1.1) se puede entonces "leer" en la forma siguiente: Una medida de incertidumbre posee la propiedad de ramificación, cuando la incertidumbre asociada a la partición $\Pi' = \{(A_1,\ldots,A_n,B_1,\ldots,B_m\}$, es la suma de la asociada a Π y de un término que depende sólo de la forma en que ha sido dividido B en los subconjuntos $B_1,\ldots,B_m$, no dependiendo ese segundo término de la subdivisión $A_1,\ldots,A_n$ de los demás elementos.

Esta interpretación de la propiedad tiene un sentido especial en muchas aplicaciones, por ejemplo, en la teoría de los cuestionarios (procesos que se realizan por medio de subdivisiones sucesivas de un mismo evento), esa propiedad permite obtener de manera elemental la incertidumbre asociada a un estado del proceso de interrogación partiendo de la asociada al estado precedente.

En este sentido no es tan importante el hecho que $H(\Pi')$ sea justamente la suma de $H(\Pi)$ y de términos dependientes de la parte modificada, sino que $H(\Pi')$ se pueda obtener, de una u otra forma, por $H(\Pi)$ y esos otros términos.

Han sido estas consideraciones las que llevaron a proponer el tomar en consideración una clase más amplia de medidas de incertidumbre, caracterizada por una propiedad que generalice (3.1.1) y (3.1.2).

Generalizaciones naturales han sido propuestas por M. Divarí y M. Pandolfi [18] (concepto de g-localidad) y por C. Poggi [17] (concepto de g-ramificación), las cuales están dadas por las siguientes definiciones.

Definición 3.3.1.

Una medida de incertidumbre es "g-local" si élla satisface la condición siguiente:

$$(3.3.1)\quad H(\Pi_A \vee \Pi_B) = \phi[H(\Pi_A \vee \{B\}), H(\{A\} \vee \Pi_B), H(\{A\} \vee \{B\})].$$

Definición 3.3.2.

Una medida de incertidumbre se llama "g-ramificable" si satisface la condición siguiente:

$$(3.3.2)\quad H(\Pi_A \vee \Pi_B) = \beta[H(\Pi_A \vee \{B\}), H(\Pi_B \vee \{A\}|\{A\} \vee \{B\})].$$

Definición 3.3.3.

A las funciones ϕ y β se le denominan respectivamente "ley de g-localidad" y "de g-ramificación" respectivamente.

En el caso más estudiado donde la entropía relativa, es proporcionada por el teorema (2.5.2), la definición 3.3.2 es un caso particular de la definición 3.3.1, siendo

$$H(\Pi_B \vee \{A\}|\{A\} \vee \{B\}) = F[H(\Pi_B \vee \{A\}), H(\{A\} \vee \{B\})]$$

con F dada por (2.5.7). Por esto nos limitaremos a estudiar las entropías g-locales

Propiedad de la Función ϕ

Hasta que no se especifique diferente, utilizaremos las notaciones

$u = H(\Pi_A \vee \{B\})$, $v = H(\{A\} \vee \Pi_B)$, $w = H(\{A\} \vee \{B\})$, $z = H(\{C\})$

(3.3.3) Como $\Pi_A \vee \{B\} > \{A\} \vee \{B\}$, $\{A\} \vee \Pi_B > \{A\} \vee \{B\}$

El campo de definición de ϕ es un subconjunto de

$$\Gamma_3 = \{(u,v,w) : \text{Inf}(u,v) \geq w \geq 0\}$$

(3.3.4) $\phi(u,v,w) \geq \text{Sup}(u,v)$

En efecto $\Pi_A \vee \Pi_B > \Pi_A \vee \{B\}$

$$\Pi_A \vee \Pi_B > \{A\} \vee \Pi_B$$

(3.3.5) $\phi(u,v,w) = \phi(v,u,w)$

En efecto $\phi(u,v,w) = H(\Pi_A \vee \Pi_B) = H(\Pi_B \vee \Pi_A) = \phi(v,u,w)$

(3.3.6) $u' < u'' \Longrightarrow \phi(u',v',w) \leq \phi(u'',v,w)$

Poniendo $u' = H(\Pi'_A \vee \{B\})$, $u'' = H(\Pi''_A \vee \{B\})$, la (3.3.6) sigue del hecho que $\Pi'_A \vee \{B\} < \Pi''_A \vee \{B\} \Longrightarrow \Pi'_A \vee \Pi_B < \Pi''_A \vee \Pi_B$

(3.3.7) $\phi(u,v,v) = u$

Se obtiene inmediatamente haciendo $\Pi_B = \{B\}$

$H(\Pi_A \vee \{B\}) = \phi[H(\Pi_A \vee \{B\}), H(\{A\} \vee \{B\}), H(\{A\} \vee \{B\})]$

(3.3.8) $\phi(u,v,w) * Z = \phi(u * z, v * z, w * z)$

Se obtiene a partir de la identidad

$(\Pi_A \vee \Pi_B) \wedge \{C\} = (\Pi_A \wedge \{C\}) \vee (\Pi_B \wedge \{C\})$. En efecto

$H((\Pi_A \vee \Pi_B) \wedge \{C\}) = H(\Pi_A \vee \Pi_B) * H(\{C\}) = \phi(u,v,w) * z$

$H[(\Pi_A \wedge \{C\}) \vee (\Pi_B \wedge \{C\})]$

$= \phi[H[\Pi_A \vee \{C\}) \vee \{B \cap C\}], H[\{A \cap C\} \vee (\Pi_B \wedge \{C\})], H[\{A \cap \{C\} B \cap C\}]$

$= \phi[H[(\Pi_A \vee \{B\}) \wedge \{C\}], H[(\{A\} \vee \Pi_B) \wedge \{C\}], H[(\{A\} \vee \{B\}) \wedge \{C\}]$

$= \phi[H(\Pi_A \vee \{B\}) * H(\{C\}), H(\{A\} \vee \Pi_B) * H(\{C\}), H(\{A\} \vee \{B\}) * H(\{C\})]$

$= \phi[u * z, v * z, w * z]$

(3.3.9) $\phi(,\phi(p,q,r),w) = \phi(\phi(u,q,w),p,r)$

Es consecuencia de la identidad $(\Pi_A \vee \Pi_B) \vee \Pi_C = \Pi_A \vee (\Pi_B \vee \Pi_C)$.

En efecto poniendo

$u = H(\Pi_A \vee \{B \cup C\})$

$w = H(\{A\} \vee \{B \cup C\})$

$q = H[\{A\} \vee (\Pi_B \vee \{C\})] = H[(\{A\} \vee \Pi_B) \vee \{C\}]$

$p = H(\{A \cup B\} \vee \Pi_C)$

$r = H(\{A \cup B\} \vee \{C\})$

tenemos:

$$\phi(p,q,r) = \phi\,[H(\{A\cup B\} \vee \Pi_C),\ H[(\{A\} \vee \Pi_B) \vee \{C\}],\ H(\{A\cup B\} \vee \{C\})] =$$

$$= H\,[(\{A\} \vee \Pi_B)] \vee \Pi_C] = H[\{A\} \vee (\Pi_B \vee \Pi_C)]$$

$$\phi(u,\phi(p,q,r),w) = \phi[H(\Pi_A \vee \{B\cup C\}),\ H[\{A\} \vee (\Pi_B \vee \Pi_C)],\ H(\{A\} \vee \{B\cup C\})] =$$

$$= H\,[\Pi_A \vee (\Pi_B \vee \Pi_C)]$$

$$\phi(u,q,w) = \phi\,[H(\Pi_A \vee \{B\cup C\}),\ H(\{A\} \vee (\Pi_B \vee \{C\}),\ H(\{A\} \vee \{B\cup C\}) =$$

$$= H\,[\Pi_A \vee (\Pi_B \vee \{C\})] = H[(\Pi_A \vee \Pi_B) \vee \{C\}]$$

$$\phi(\phi(u,q,w),p,r) = \phi[H(\Pi_A \vee \Pi_B) \vee \{C\},\ H(\{A\cup B\} \vee \Pi_C),\ H(\{A\cup B\} \vee \{C\})] =$$

$$= H[(\Pi_A \vee \Pi_B) \vee \Pi_C]$$

Las propiedades (3.3.3) hasta (3.3.9) son satisfechas por todos los valores u(u',u'') v w z p q r que son medidas de incertidumbre de particiones pertenecientes a E (Además en (3.3.8) es necesario que $(\Pi_A \vee \Pi_B, \{C\}) \in K$).

Desde ahora en adelante, aunque no daremos la definición formal, nos ocuparemos de leyes de localidad universales: en definitiva, leyes $\phi(u,v,w)$ definidas en Γ_3 y tales que las propiedades (3.3.3)-(3.3.9) (excepto posiblemente (3.3.8)) se cumplan para todos los valores admisibles de las argumentos.

Las propiedades así establecidas constituyen entonces el siguiente sistema de ecuaciones funcionales

$$
(3.3.10)\qquad
\begin{array}{ll}
\text{a.} & \mathrm{Inf}(u,o) \geq w \\
\text{b.} & \phi(u,v,w) \geq \mathrm{Sup}\,(u,v) \\
\text{c.} & \phi(u,v,w) = \phi(v,u,w) \\
\text{d.} & v' < v'' \Longrightarrow \phi(u,v',w) \leq \phi(u,v'',w) \\
\text{e.} & \phi\,[u,\phi(p,q,r),w] = \phi\,[\phi(u,q,w),p,r] \\
\text{f.} & \phi(u,w,w) = u
\end{array}
$$

$$(3.3.11)\quad \phi(u,v,w) * z = \phi(u * z,\ v * z,\ w * z)$$

Los dos Teoremas que siguen determinan completamente la forma general de la ley de composición.

Ellos se refieren (aunque implícitamente) a dos situaciones muy diferentes.

El primero determina la solución del sistema (3.3.10) (3.3.11). Como incluye la ecuación (3.3.11) y se supone que sea satisfecha para cualquier valor de los argumentos, él determina al final todas las leyes de localidad, universalmente compatibles con una ley de independencia * asignada, es decir, compatibles con cualquier espacio de incertidumbre donde la ley de independencia es *, como quiera que se escoja la clase K de las álgebras independientes.

El Teorema 3.3.2 proporciona la solución del sistema (3.3.10) que no contine la ecuación (3.3.11). Por esta razón, este último teorema se adapta de mejor manera al caso donde $K = \phi$

Teorema de Representación 3.3.1.

Sean $\{(\alpha_i, \beta_i)\}$ y $\{g_i : [\alpha_i, \beta_i] \to \mathbb{R}^+\}$

las sucesiones (de intervalos abiertos disjuntos y de funciones respectivamente) que determinan la forma de la ley de independencia $*$, según se especificó en (1.7.1). La solución más general del sistema (3.3.10)(3.3.11) se obtiene haciendo

$$(3.3.12)\quad \phi(u,v,w) = \begin{cases} \mathrm{Sup}(u,v) & \text{si } (u,v) \in \mathbb{R}^{+2} - \cup(\alpha_i,\beta_i)^2 \\ \phi_i(u,v,w) & \text{si } (u,v) \in (\alpha_i,\beta_i)^2 \end{cases}$$

donde $\phi_i(u,v,w)$ es una de las tres funciones siguientes:

$$\text{a. } \phi_i^{(1)}(u,v,w) = \mathrm{Sup}\,(u,v)$$

$$(3.3.13)\quad \text{b. } \phi_i^{(2)}(u,v,w) = \bar{g}_i\,[\log_{a_i}(a_i^{g_i(u)} + a_i^{g_i(v)} - a_i^{g_i(\bar{w})})]$$

$$\text{c. } \phi_i^{(3)}(u,v,w) = \bar{g}_i\,[\log_{a_i}(a_i^{g_i(u)} + a_i^{g_i(v)} - a_i^{g_i(\bar{w})})]$$

con $\bar{w} = \mathrm{Sup}(w,\alpha_i)$, a_i es una constante cualquiera y

$\bar{g}_i(\xi) := g_i^{-1}\{\mathrm{Inf}\,[\xi, g_i(p_i)]\}$ es la seudoinversa de g_i.

(Para la demostración ver Apéndice B).

Más claramente, es solución del sistema (3.3.10) (3.3.11) toda función $\phi(u,v,w)$ cuya restricción a $[(\alpha_i,\beta_i)^2 \times \mathbb{R}^+] \cap \Gamma_3$ coincide con una de las funciones $\phi_i^{(1)}, \phi_i^{(2)}, \phi_i^{(3)}$ expresadas en (3.3.13), mientras su restricción a $\Gamma_3 - \cup[(\alpha_i,\beta_i)^2 \times \mathbb{R}^+]$ es la función Sup(u,v). En particular, escogiendo $\phi_i = \phi_i^{(\cdot)}$ para cada i se obtiene la solución:

$$\phi(u,v,w) = \mathrm{Sup}(u,v)$$

La cual es la única compatible con todas las leyes de independencia.

De la expresión (3.3.12) se obtienen fácilmente las leyes de localidad compatibles universalmente con la ley de independencia clásica $x * y = x + y$.

En efecto, en este caso la sucesión $\{(\alpha_i,\beta_i)\}$ está compuesta por el único intervalo $(o, +\infty)$ y $g(x) = x$. Por lo tanto las soluciones de (3.3.10) (3.3.11) en este caso son

$$\phi^{(1)}(u,v,w) = \mathrm{Sup}\,(u,v)$$
$$\phi^{(2)}(u,v,w) = u + v - w$$
$$\phi^{(3)}(u,v,w) = \log_a (a^u + a^v - a^w)$$

Observación 3.3.1

La función $\phi^{(2)}$ representa la propiedad de ramificación ya antes definida. Es muy importante observar que esa función es solución del sistema (3.3.10) (3.3.11) solo cuando $x * y = x + y$ (es suficiente examinar (3.3.12) para convencerse). Luego la propiedad de ramificación solo es compatible con los espacios clásicos de incertidumbre.

Observación 3.3.2

El hecho que también $\phi^{(1)}$ y $\phi^{(3)}$ sean leyes de localidad demuestra que la localidad es una verdadera generalización de la ramificación, aún cuando $* = +$.

Teorema de Representación 3.3.2

La solución continua más general del sistema (3.3.10) se obtiene fijando arbitrariamente una sucesión de intervalos $\{(\alpha_i,\beta_i)\}$ abiertos, disjuntos y para cada uno de ellos una función $g_i : [\alpha_i,\beta_i] \to \mathbb{R}$, estrictamente creciente, con $g_i(\alpha_i) = 0$ y haciendo

$$(3.3.14)\quad \phi(u,v,w) = \begin{cases} \bar{g}_i\,[g_i(u) + g_i(v) - g_i\,[\mathrm{Sup}(w,\alpha_i)]] & \text{si } (u,v)\in(\alpha_i,\beta_i)^2 \\ \\ \mathrm{Sup}\,(u,v) & \text{en otro caso} \end{cases}$$

donde $\bar{g}_i(\xi) := g_i^{-1}\{\mathrm{Inf}\,[\xi,\ g_i(p_i)]\}$ es la seudoinversa de g_i. Para la demostración ver Apéndice C.

Observación 3.3.3

Se puede fácilmente demostrar que: Asignada una ley de independencia $*$ y una ley de localidad ϕ compatible con $*$ (de la forma (3.3.12) existe una función ϕ de la forma (3.3.14) tal que

$$\phi(u,v,w) = \overline{\Phi}(u,v,w)$$

En efecto dada ϕ; se construye $\bar{\Phi}$ de la siguiente manera:

Se escoge la sucesión $\{(\bar{\alpha}_i,\bar{\beta}_i)\}$ poniendo en élla todos los intervalos (α_i,β_i) donde ϕ_i es del tipo $\phi_i^{(2)}$ o $\phi_i^{(3)}$ y además se escoge la sucesión $\bar{g}_i$ poniendo

$$\bar{g}_i(t) = g_i(t) \qquad \text{si } \phi_i = \phi_i^{(2)}$$

$$\bar{g}_i(t) = a_i^{g_i(t)} \qquad \text{si } \phi_i = \phi_i^{(3)}$$

La función $\bar{\Phi}$ construida según (3.3.14) con esta selección de $\{(\bar{\alpha}_i,\bar{\beta}_i)\}$ y $\{\bar{g}_i\}$ coincide con ϕ.

Recíprocamente dada una función ϕ de la forma (3.3.14), élla representa una ley de localidad compatible (según el Teorema 3.3.1) con la ley de independencia particularizada por las sucesiones

$\{(\alpha_i,\beta_i)\}$ y $\{g_i\}$ que sirven para definir ϕ.

La clase de las soluciones representadas por el Teorema 3.3.2 coincide por lo tanto con la representada por el Teorema 3.3.1.

3.4. ENTROPIAS g-RAMIFICADAS

Como ya hemos dicho antes, la noción de g-ramificación se reduce a un caso particular de la de g-localidad cuando $H(\Pi|\Pi') = F[H(\Pi \wedge \Pi'), H(\Pi')]$.

El objetivo de este parágrafo será el de caracterizar la forma que asumen las entropías g-locales cuando son aún g-ramificadas. Para solucionar este problema usaremos la forma explícita de la función $F(x,y)$ establecida en (2.5.7). Como esa expresión determina la información condicional compatible con una ley de independencia asignada $*$, aquí utilizaremos la expresión de $\phi(u,v,w)$ proporcionada por el teorema 3.3.1.

Tenemos los resultados siguientes:

Lema 3.4.1.

Una condición necesaria para que un espacio g-local sea g-ramificable es que sea

$$a = \operatorname{Sup}\{x : x * x = x\} = 0 \quad ;$$

es decir $\Lambda = \{o\}$. En otras palabras que sea

$$x * y = \bar{g}\,\{g(x) + g(y)\}\ \forall\ (x,y) \in \mathbb{R}^{+2}$$

Demostración

Supongamos que sea $a > 0$ y escojamos $u,v < a$ (por consiguiente, también $w < a$). Como $\phi(u,v,w) = \beta[u,F(v,w)]$ y además en $[o,a]^2$ $F(x,y) = 0$ (Ver 2.5.7), tenemos

$$\phi(u,v,w) = \beta(u,o)$$

Luego en $\Gamma_3(a) := [o,a]^3 \cap \Gamma_3$, $\phi(u,v,w)$ no depende de la variable v, es decir, $\phi(u,v,w) = \phi(u,v',w)\ \forall\ v' \geq w$. Por otro lado $\phi(u,v',w) \geq v'$ (siempre) y por lo tanto $\forall\ (u,v,w) \in \Gamma_3(a)$

$$\phi(u,v,w) \geq v'\ \ \forall\ v' < a\ ;\ \text{ya que}\ \phi(u,v,w) \geq a.$$

Por otro lado, examinando la expresión (3.3.12) de la función ϕ se reconoce que siempre es posible escoger una terna $(u,v,w) \in \Gamma_3(a)$ tal que sea $\phi(u,v,w) < a$. Absurdo, luego $a = o$.

Lema 3.4.2.

Si $a = o$ y $\phi(u,v,w) = \beta[u,F(v,w)]$ entonces

$$\phi(u,v,w) = \bar{g}\,[g(u) + g(v) - g(w)]$$

Demostración:

Como $a = o$. $F(v,w) = \bar{g}\{g(v) - g(w)\}$, y como $g(w) < g(v)$ se tiene $F(v,w) = g^{-1}\{g(v) - g(w)\}$. Recordemos ahora que $\phi(u,v,w)$ puede asumir sólo una de las tres expresiones (3.3.13 a,b,c).

a. No existe ninguna función $\beta(\cdot,\cdot)$ tal que

$$\text{Sup}(u,v) = \beta(u, g^{-1}(g(v)-g(w)) \quad \forall\ (u,v,w) \in \Gamma_3$$

En efecto, cuando $v > u$ esta expresión se reduce a

$$v = \beta(u, g^{-1}(g(v) - g(w))$$

y no hay ninguna función $\beta(\cdot,\cdot)$ que cumple con esa relación para todo $v \geq u$, $w < v$.

c. No existen funciones $\beta(\cdot,\cdot)$ tales que

$$(3.4.1)\quad \bar{g}(\log_a(a^{g(u)} + a^{g(v)} - a^{g(w)}) = \beta[u, g^{-1}(g(v) - g(w))]$$

En efecto, haciendo $w = o$ se obtiene

$$\bar{g}[\log_a(a^{g(u)} + a^{g(v)} - 1)] = \beta(u,v)$$

relación que determina completamente la función $\beta(\xi,\eta)$.

Entonces para que (3.4.1) resulte una identidad, tendría que resultar

$$\bar{g}\{\log_a[a^{g(u)} + a^{g(v)} - a^{g(w)}]\} = \bar{g}[\log\{a^{g(u)} + a^{g(v)-g(w)} - 1\}$$

que no es una identidad.

b. La relación

$$\bar{g}\{g(u) + g(v) - g(w)\} = \beta[u, g^{-1}(g(v) - g(w))]$$

sí resulta una identidad a condición que se ponga

$$\beta(\xi,\eta) = \bar{g}[g(\xi) + g(\eta)].$$

Así, el lema queda demostrado.

Resumiendo:

Teorema 3.4.1

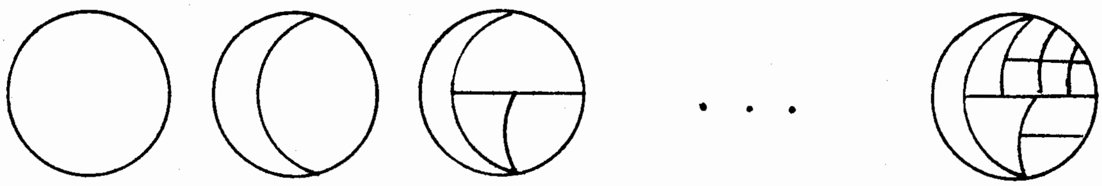
La información relativa $H(\Pi|\Pi')$ dada por la expresión (2.5.7) es g-ramificada sólo en los espacios donde la ley de independencia es $x * y = \bar{g}\{g(x) + g(y)\}$. En tal caso la ley de ramificación es $\beta(\xi,\eta) = \bar{g}\{g(\xi) + g(\eta)\}$.

3.5. CONSTRUCCION DE LAS MEDIDAS DE INCERTIDUMBRE g-LOCALES Método de las subdivisiones sucesivas.

A partir del soporte $A = \cup A_i$, se plantea (o se construye) la experiencia $\Pi_A = \{A_1,\ldots,A_n\}$ realizando una sucesión de particiones

$$\sigma = \{\Pi' = \{A\}, \Pi^2,\ldots,\Pi^m : m \leq n\}$$

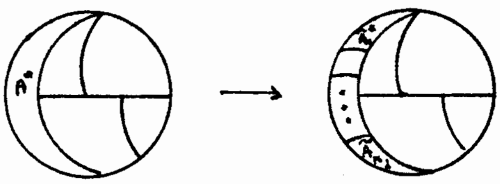
de manera que Π^{i+1} sea un refinamiento de Π^i para cada $i = 1,\ldots,m-1$. (Ver Figura 1) y procediendo hasta que $\Pi^m = \Pi_A$.



Π^1 Π^2 Π^3 . . . $\Pi^m = \Pi_A$

FIGURA 1.

De manera más precisa se construye Π^{i+1} particionando un evento A^* de Π^i en $K_i \geq 2$ subconjuntos $\tilde{A}_1,\ldots,\tilde{A}_{k_i}$ (Ver Figura 2).



Π^i Π^{i+1}

FIGURA 2.

Se entenderá mejor lo que diremos más adelante, si se considera que la operación realizada al pasar de Π^i a Π^{i+1} es en realidad el producto (Ver 2.2) de las dos particiones

$$\Pi^i = \{\tilde{A}^1, \tilde{A}^2, \ldots, \tilde{A}^{r_i}, A^*\} \text{ y } \Pi_i = \{A_*, \tilde{A}_1, \tilde{A}_2, \ldots, \tilde{A}_{k_i}\}$$

donde $A_* = \bigcup_{j=1}^{r_i} \tilde{A}^j$ (Ver Fig. 3)

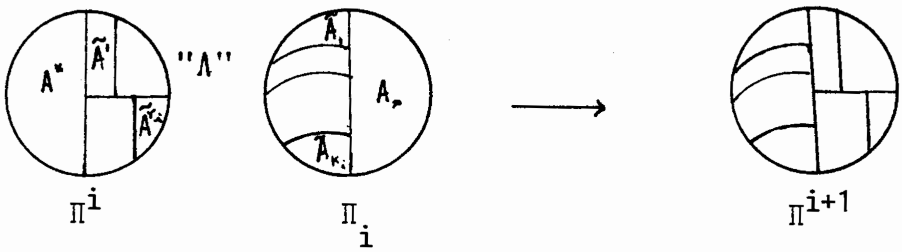


FIGURA 3

Obviamente hay más de una sucesión σ que puede ser utilizada para construir la partición Π_A. Entre éllas, la más conveniente para lo que queremos desarrollar más adelante, es la sucesión dibujada en la Figura 4.

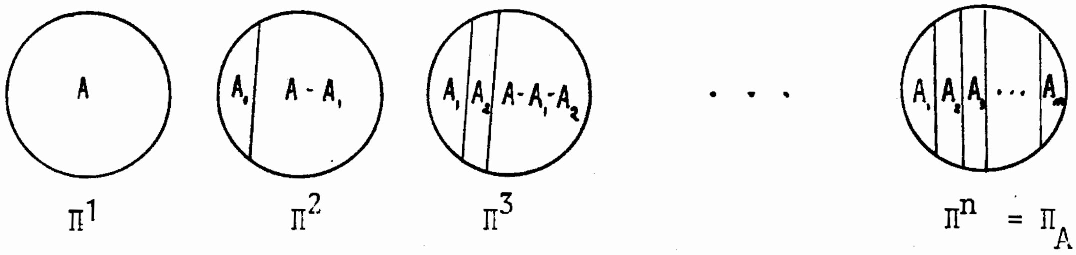


FIGURA 4.

donde se pasa de Π^i a Π^{i+1} particionando el evento $\bigcup_{j=i}^{n} A_j$ en los dos eventos A_i, $\bigcup_{j=i}^{n} A_j$; las particiones Π^i y Π_i son respectivamente

$$\Pi^i = \{A_1, \ldots, A_{i-1}, \bigcup_{j=i}^{n} A_j\} \qquad y$$

$$\Pi_i = \{\bigcup_{j=1}^{i-1} A_j, A_i, \bigcup_{j=i+1}^{n} A_j\}$$

Por supuesto, este proceso es dicotomico y la sucesión $\{\Pi^i\}$ constituida por $m = n = \mathrm{card}(\Pi_A)$ particiones.

En este método la partición Π^{i+1} está determinada completamente por las particiones Π^i y Π_i y por lo tanto $H(\Pi^{i+1}) = T_i(\Pi^i, \Pi_i)$. Sin embargo, la función de particiones T_i puede ser demasiado compleja y por consiguiente, difícil de construir la sucesión $\{H(\Pi^i), i = 1, \ldots, n\}$ hasta obtener $H(\Pi^m) = H(\Pi_A)$.

Para evitar este inconveniente es necesario fijar la atención sobre las clases de entropías en correspondencia con las cuales las aplicaciones T_i son simples.

No hay duda que el cálculo de $H(\Pi^{i+1})$ se simplificaría considerablemente si $T_i\ (\Pi^i, \Pi_i)$ dependiera solo de los valores $H(\Pi^i)$ y $H(\Pi_i)$; por lo tanto a primera vista parece natural limitarnos a tomar en consideración, entre todas las medidas de incertidumbre, sólo aquellas donde

$$H\ (\Pi^{i+1}) = G\ [H(\Pi^{i}),\ H(\Pi_{i})]$$

Desafortunadamente, aún con un análisis superficial uno se da cuenta que estas clases de entropías son muy pobres, ya que ni siquiera contienen las medidas de incertidumbre más naturales es decir, aquellas que son esperanzas (o casi esperanzas) matemáticas de medidas de información ([19], [20]). De aquí que, para una co trucción operativa simplificada pero razonable es preciso que $H(\Pi^{i+1})$, además de $H(\Pi^{i})$ y $H(\Pi_{i})$, dependa de otra cantidad. Por otro lado, el procedimiento de construcción por subdivisiones sucesivas es de carácter típicamente local; este hecho sugiere tomar en consideración justamente los espacios de incertidumbre g-locales, con lo que resulta

(3.5.1) $H(\Pi^{i+1}) = \phi\,[H\ (\Pi^{i}),\ H(\Pi_{i}), H(\{A_{i}^{*},\ A_{*}^{i}\})]$

donde $A_{i}^{*} = \bigcup\limits_{j=1}^{i-1} A_{j},\ A_{*}^{i} = \bigcup\limits_{j=i}^{n} A_{j}.$

El conocer la ley de localidad ϕ (y otras pocas cantidades) nos permitirá, como veremos más adelante, construir la medida de incertidumbre asociada a cualquier partición Π_{A}.

3.6. FORMULA RECURRENTE

Asignada la ley de localidad ϕ, cuando la partición Π_A es generada por la sucesión $\{\Pi^i\}$ de la Fig. 4, el procedimiento recursivo para determinar los valores $\{H(\Pi^i)\}$, y al final $H(\Pi_A)$, es el siguiente:

(3.6.1) $$H(\Pi^{i+1}) = \phi[H(\Pi^i),\ H(\Pi_i),\ H(\Pi_i^*)] \quad i = 2,\ldots,n-1$$

donde:

$$\Pi^1 = \{\phi\}$$

$$\Pi^2 = \{A_1,\ \bigcup_{i=2}^{m} A_i\} \quad \text{(partición inicial)}$$

(3.6.2)

$$\Pi_i = \{\bigcup_{j=1}^{i-1} A_j,\ A_i,\ \bigcup_{j=i+1}^{n} A_j\}$$

$$\Pi_i^* = \{\bigcup_{j=1}^{i-1} A_j,\ \bigcup_{j=i}^{n} A_j\} = \{A_i^*,\ A_*^i\}$$

Para que este procedimiento se pueda completar es necesario conocer:

i) el valor inicial $H(\Pi')$

ii) los valores $H(\Pi_i)$, $H(\Pi_i^*)$

En otras palabras, se puede utilizar este procedimiento sólo si ya se conoce la restricción de H a la subclase de las particiones constituidas por 2 ó 3 eventos (si se conoce la aplicación $\hat{H} : E_2 \cup E_3 \rightarrow \mathbb{R}^+$).

Además, la restricción $\hat{H}$ nunca se puede obtener por medio de la ley de localidad. En efecto, si la partición Π_i contiene solo dos elementos ($\Pi_i = \{A^*, A_*\}$), entonces $\Pi^i \wedge \Pi_i = \Pi^i$, cualquiera sea la partición Π^i; por lo tanto de (3.5.1) obtenemos

$$(3.6.3) \quad H(\Pi^i) = \phi\,[H(\Pi^i),\ H(\Pi_i),\ H(\Pi_i)]$$

Quiere decir ésto que $H(\Pi^i)$ no puede calcularse sin conocerse. Esto nos lleva a la conclusión siguiente: para que (3.5.1) sea efectiva es necesario que card $(\Pi_i) > 2$ y de la misma manera también debe ser card $(\Pi^i) > 2$. Entonces, la restricción $\hat{H}$, no puede de ninguna manera calcularse y debe asignarse previamente.

3.7. LA RESTRICCION $\hat{H}$:

Ya vimos que la restricción $\hat{H}$ de H a $E_2 \cup E_3$ tiene que asignarse antes de empezar la construcción de $H(\Pi_A)$, por el procedimiento (3.6.1) (3.6.2). Por otra parte, $\hat{H}$ es la restricción de una medida de incertidumbre y por lo tanto, tiene que satisfacer los axiomas correspondientes, a saber:

$$(3.7.1) \quad \begin{aligned} &\hat{H}(\{A,B\}) = \hat{H}(\{B,A\}) \\ &\hat{H}(\{A,B,C\}) = \hat{H}(\{B,A,C\}) = \hat{H}(\{A,C,B\}) \\ &\hat{H}(\{A,B,C\}) \geq \hat{H}(\{A \cup B,\ C\}) \geq J\,(A \cup B \cup C) \end{aligned}$$

$$(\{A,B\}, \{D\}) \in K \Longrightarrow H(\{A,B\} \wedge \{D\}) = H(\{A,B\}) * J(D)$$

$$(\{A,B,C\},\{D\}) \in K \Longrightarrow H(\{A,B,C\} \wedge \{D\}) = H(\{A,B,C\}) * J(D)$$

Además $\{A \cup B, C, D\} \wedge \{A,B,C \cup D\} = \{A \cup C,B,D\} \wedge \{A,C,B \cup D\}$ (ambos miembros representan la partición $\{A,B,C,D\}$) y por lo tanto entre la restricción $\hat{H}$ y la ley de localidad ϕ debe existir la siguiente relación de compatibilidad

$$\begin{aligned}(3.7.2)\quad &\phi[\hat{H}(\{A \cup B,C,D\}), \hat{H}(\{A,B,C \cup D\}), \hat{H}(\{A \cup B,C \cup D\})] \\ &= \phi[\hat{H}(\{A \cup C,B,D\}), \hat{H}(\{A,C,B \cup D\}), \hat{H}(\{A \cup C,B \cup D\})]\end{aligned}$$

No es difícil reconocer que si ϕ y $\hat{H}$ satisfacen las condiciones (3.3.10) (3.3.11) (3.7.1) (3.7.2) entonces la construcción recurrente de $H(\Pi_A)$ es coherente; es decir, $H(\Pi_A)$ es una verdadera medida de incertidumbre.

3.8. FORMA EXPLICITA DE LAS MEDIDAS DE INCERTIDUMBRE g-LOCALES

Los resultados que vienen a continuación se demuestran utilizando el Teorema de representación 3.3.2. Sin embargo, se refieren también al Teorema 3.3.1 por la relación establecida entre los dos teoremas.

Los dos teoremas siguientes proporcionan la forma explícita de las medidas de incertidumbre g-locales.

Para simplificar los enunciados y las demostraciones pondremos

$$h_1 = H(\Pi_1) = H(\{A_1, \bigcup_{j=2}^{n} A_j\})$$

$$h_i = H(\Pi_i) = H(\{\bigcup_{j=1}^{i-1} A_j, A_i, \bigcup_{j=i+1}^{n} A_j\}) \quad i = 2,\dots,n-1$$

$$h_i^* = H(\Pi_i^*)$$

$$h^i = H(\Pi^i) = H(\{A_1,\dots,A_{i-1}, \bigcup_{j=i}^{n} A_j\})$$

m el índice para el cual se cumple que $h_m = \max \{h_i, i=2 \dots,n-1\}$

Teorema 3.8.1

Si $h_m \in \Lambda_\phi$, entonces $H(\Pi_A) = h_m$

Demostración:

$$h^{m+1} = H(\Pi^{m+1}) = H(\{A_1,\dots,A_m, \bigcup_{j=m+1}^{n} A_j\}) = h_m$$

En efecto, empezamos la sucesión Π^i poniendo

$$\Pi^2 = \Pi_1 \ ; \text{ además } \forall\ i = 3,\dots,n; \ \Pi^{i+1} = \{A_1,\dots,A_i, \bigcup_{j=i+1}^{n} A_j\}$$

es la "superposición" de las particiones Π^i y Π_i aunque en verdad esta "superposición" es el producto de esas particiones. Entonces tenemos que:

$$h^{i+1} = H(\Pi^{i+1}) = H(\Pi^i \wedge \Pi_i) = \phi[h^i, h_i, h_i^*]$$

Además, como $h_i \leq h_m \; \forall i=1,\ldots,m-1$ y $h^2 = h_1 \leq h_m$, entonces en forma recurrente podemos probar que $h^m \leq h_m$ ya que

$$h^i \leq h_m \text{ y } h_i \leq h_m \rightarrow h^{i+1} = \phi(h^i, h_i, h_i^*) \leq h_m \quad \forall i = 1,\ldots,m-1$$

Como $h_m \in \Lambda_\phi$, se obtiene que

$$h^{m+1} = \phi(h^m, h_m, x_m) = \text{Sup }(h^m, h_m) = h_m$$

Procedamos a probar que $h^n = H(\Pi^n) = H(\Pi_A) = h_m$

teniendo en cuenta que $\forall i > m$, $h_i \leq h_m \in \Lambda_\phi$

$$h^{m+2} = \phi[h^{m+1}, h_{m+1}, h^*_{m+1}] = \phi[h_m, h_{m+1}, h^*_{m+1}]$$

$$= \text{Sup }[h_m, h_{m+1}] = h_m$$

$$h^{m+3} = \phi[h^{m+2}, h_{m+2}, h^*_{m+2}] = \phi[h_m, h_{m+2}, h^*_{m+2}]$$

$$= \text{Sup }[h_m, h_{m+2}] = h_m$$

$$\vdots$$

$$h^n = \phi[h^{n-1}, h_{n-1}, h^*_{n-1}] = \phi[h_m, h_{n-1}, h^*_{n-1}] = h_m$$

por lo tanto $H(\Pi^n) = h^n = H(\{A_1,\ldots,A_{n-1},A_n\}) = H(\Pi_A) = h_m$

$\square$

Teorema 3.8.2

Si $h_m \in (\alpha,\beta) \subset \overline{\Lambda}_\phi = \{(\alpha_i,\beta_i) : i \in I\}$, entonces

$$H(\Pi_A) = \bar{f} \left[\sum_{h_r \in (\alpha,\beta)} f(h_r) - \sum_{h_r^* \in (\alpha,\beta)} f(h_r^*) \right]$$

Demostración:

Sean $r_1 = \text{Inf}\ (\ell : h_\ell > \alpha)$ y $\forall j \in \{r_1, \ldots, n-1\}$

$r(j) = \text{Sup}\ \{i : i \leq j,\ h_i > \alpha\}$

Se puede demostrar fácilmente que:

(3.8.1) $\quad h^{j+1} = h^{r(j)+1}$

si $r(j) = j$, (3.8.1) es una identidad

si $r(j) < j$, se utiliza el mismo procedimiento que se usó en el teorema 3.8.1 cuando se logró demostrar que $h^{n-1} = h_m$, sustituyendo m por $r(j)$ y $n = 1$ por j.

En efecto, siendo $h^{r(j)+1} > \alpha$, para cada $\ell \geq r(j)$, $\ell \leq j$ tenemos por recurrencia

$$h^\ell = h^{r(j)+1} \to h^{\ell+1} = \phi(h^\ell,\ h_\ell,\ h_\ell^*) = h^{r(j)+1}$$

ya que $\phi(h^\ell,\ h_\ell,\ h_\ell^*) = \text{Sup}\ (h^\ell,\ h_\ell)$

siendo en este caso $h_\ell \leq \alpha < h^\ell$

La única diferencia respecto al caso del teorema 3.8.1 consiste en el hecho de que aquí no siempre $h^{r(j)+1}$ es igual a $h_{r(j)}$.

En consecuencia, al construir la sucesión $\{h^i\}$ para obtener $h^n = H(\Pi_A)$ se puede eliminar la construcción de todos los términos h^{s+1} para los cuales $h_s \leq \alpha$ (se deduce de 3.8.1).

Indicamos por $\mathbf{P} = \{r_i : i = 1,\ldots \bar{n} \leq n\}$ la subsucesión de los números $1,\ldots,n$ correspondiente a los valores h_r superiores a α tal que $r_i \in p \to h_{r_i} \geq \alpha$, $j \notin \mathbf{P}$, $j \leq n \to h_j < \alpha$.

En lo que sigue tendremos que escribir el término

(3.8.2) $\quad f\,[\mathrm{Sup}\,(h^*,\alpha)]$

pero como $f(\alpha) = 0$, la expresión (3.8.2) se puede simplificar escribiendo en su lugar $f(h^*)$ bajo la convención de extender f haciendo $f(\xi) = 0 \quad \forall\ \xi \leq \alpha$.

Como r_1 es el primer término de la sucesión $1,\ldots,n$ en correspondencia al cual $h_i > \alpha$, es evidente que $h_{r-1} < \alpha$ y $h^{r_1-1} < \alpha$ y por lo tanto $h^r = \phi(h^{r_1-1}, h_{r_1-1}, h^*_{r_1-1}) < \alpha$. Por consiguiente, recordando que solo tomamos en consideración los índices $r_i \in \mathbf{P}$.

$$h^{r_2} = H(\Pi^{r_2}) = \phi(h^{r_1}, h_{r_1}, h^*_{r_1}) = h_{r_1}$$

$$h^{r_3} = H(\Pi^{r_3}) = \phi(h^{r_2}, h_{r_2}, h^*_{r_2}) = \phi(h_{r_1}, h_{r_2}, h^*_{r_2})$$

$$= \bar{f}\,[f(h_{r_1}) + f(h_{r_2}) - f(h^*_{r_2})]$$

$$\vdots$$

$$h^{r_{i+1}} = H(\Pi^{r_{i+1}}) = \phi\,[\phi(h^{r_{i+1}}, h_{r_{i-1}}, h^*_{r_{i-1}}), h_{r_i}, h^*_{r_i}]$$

$$= \bar{f}\,[f\circ\bar{f}\,\{f(h^{r_{i-1}}) + f(h_{r_{i-1}}) - f(h^*_{r_{i-1}})\}$$

$$+ f\,(h_{r_i}) = f(h^*_{r_i})]$$

Cumpliéndose esta relación para todo $i = 2,\ldots,\bar{n}$, ya que para $i \geq 2$ h_{r_i} y h^{r_i} pertenecen a $[\alpha,\beta]$.

Supongamos ahora que para cierto índice i^* sea

$$f(h^{r_{i^*}}) + f(h_{r_{i^*}}) - f(h^*_{r_{i^*}}) \geq f(\beta)$$

como $f(h_{r_i}) - f(h^*_{r_i}) \geq 0$ entonces

$$f\circ\bar{f}\,\{f(h^{r_{i^*}}) + f(h_{r_{i^*}}) - f(h^*_{r_{i^*}})\} + f(h_{r_i}) - f(h^*_{r_i}) \geq f(\beta)$$

y por lo tanto

$$H(\Pi^{r_{i^*+1}}) = h^{r_{i^*+1}} = \beta$$

Utilizando recursivamente la expresión general de $\phi(u,v,z)$ después de recordar que $\forall\, i\; f(h_k) - f(h^*_k) \geq 0$

$$\left(\{ \bigcup_{j=1}^{i-1} A_j \; , \; A_i \; , \; \bigcup_{j=i+1}^{n} A_j \} \text{ es un refinamiento de } \right.$$

$$\left. \left\{ \bigcup_{j=1}^{i-1} A_j \; , \; \bigcup_{j=i}^{n} A_j \right\} \right) \qquad ,$$

se deduce fácilmente que en este caso

$$H(\Pi_A) = \beta$$

En consecuencia,

si $f(h^{r_i*}) + f(h_{r_i*}) - f(h^*_{r_i*}) \geq f(\beta)$ entonces

$$\sum_{i=1}^{\bar{n}} [f(h_{r_i}) - f(h^*_{r_i})] \geq f(\beta)$$

Luego, después de recordar que $h^*_{r_i} < \alpha \rightarrow f(h^*_{r_i}) = 0$, se obtiene.

$$H(\Pi_A) = \beta = \bar{f} \left[\sum_{h_r \in (\alpha,\beta)} f(h_r) - \sum_{h^*_r \in (\alpha,\beta)} f(h^*_r) \right]$$

En este caso el Teorema queda demostrado .

Por otro lado, si siempre resulta

$$f(h^{r_i}) + f(h_{r_i}) - f(h^*_{r_i}) < f(\beta)$$

entonces para cada $i = 2 \ldots \bar{n}$

$$H(\Pi^{r_i+1}) = \bar{f}\{f(h^{r_i}) + f(h_{r_i}) - f(h^*_{r_i})\}$$

$$= f^{-1}\,[f(h^{r_i}) + f(h_{r_i}) - f(h^*_{r_i})]$$

y además se tiene

$$H(\Pi^{r_i+1}) = f^{-1}\,[f(h^{r_i}) + f(h_{r_i}) - f(h^*{}_{r_i})]$$

$$= f^{-1}\,[f\,\{f^{-1}(f(h^{r_i-1}) + f(h_{r_i-1}) - f(h^*{}_{r_i-1}))\}$$

$$+ f(h_{r_i}) - f(h^*_{r_i})]$$

$$= f^{-1}\,[f(h^{r_i-1}) + f(h_{r_i-1}) + f(h_{r_i}) - f(h^*_{r_i-1}) - f(h^*_{r_i})$$

Procediendo por recurrencia hacia atrás con respecto al término $f(h^{r_i-1})$ se obtiene fácilmente que:

$$H(\Pi_A) = f^{-1}\,[\sum_{h_{r_i}\in(\alpha,\beta)} f(h_{r_i}) - \sum_{h^*_{r_i}\in(\alpha,\beta)} f(h^*_{r_i})] \qquad \square$$

3.9. EJEMPLOS DE INCERTIDUMBRES g-LOCALES

En los ejemplos que siguen se construirán medidas de incertidumbres g-locales según las especificaciones previstas en el parágrafo 3.8; previa asignación en (3.5.1) y en (3.7), de la restricción de $\hat{H}$ a $E_2 \cup E_3$ que nos permitirá a su vez asignar la sucesión $\{(\alpha_i, \beta_i)\}$ y las funciones f_i de tal manera que sean compatibles con $\hat{H}$, es decir, de manera que se cumplan las ecuaciones (3.3.14) y (3.7.1)

Clase I:

En estos ejemplos la sucesión $\{(\alpha_i, \beta_i)\}$ estará constituida por un solo intervalo $(\alpha_1 = 0,\ \beta_1 = \beta \leq +\infty$. Luego, sólo quedan por asignar la función $g = g_1$ y la restricción $H : E_2 \cup E_3 \rightarrow \mathbb{R}^+$

I-1 $g(x) = x^r$

$$H(A_1, A_2) = \mathrm{Inf}\,[\beta,\ \{f(J(A_1)) + f(J(A_2))\}^{1/r}]$$

$$H(A_1, A_2, A_3) = \mathrm{Inf}\,[\beta,\ \{f(J(A_1)) + f(J(A_2)) + g(J(A3))\}^{1/r}]$$

donde $f : \mathbb{R}^+ \rightarrow \mathbb{R}^+$ es una función tal que $f(J(A_1) + f(J(A_2) \geq f\,[J(A_1 \cup A_2)]$ (condición de compatibilidad entre f y J. para que se cumpla la propiedad de monotonía). Bajo estas selecciones, la medida de incertidumbre $H(\Pi_A)$ se obtiene aplicando el Teorema 3.8.2 y es

$$(3.9.1)\quad H(\Pi_A) = \mathrm{Inf}\,[\beta,\ \{\Sigma\ f[J(A_i)]^{1/r}]$$

En particular, pertenecen a esta familia de medidas de incertidumbre la entropía de Shannon y la hiperbólica.

1.a) Se obtiene la entropía de Shannon haciendo $r = 1$, $\beta = +\infty$, $f(x) = e^{-x}.x$ y tomando por $J(A)$ una medida de información de tipo Shannoniano; es decir, poniendo $J(A) = -\log p(A)$ o equivalentemente (si no se quieren utilizar las probabilidades) escogiendo $J : \mathcal{S} \to \mathbf{R}^+$ de manera que

$$A \cap B = \phi \Longrightarrow J(A \cup B) = -\log [\exp\{-J(A)\} + \exp\{-J(B)\}]$$

1.b) Para obtener la entropía hiperbólica se hace $r = 1$, $\beta = +\infty$ $f(x) = x$ y se pone $J(A) = 1/m(A)$, siendo $m: \mathcal{S} \to \mathbf{R}^+$ una medida en el sentido corriente, con $m(\Omega) = +\infty$ (si no se quiere definir J por medio de una medida se procede escogiendo la aplicación $J : \mathcal{S} \to \mathbf{R}^+$ de manera que

$$A \cap B = \phi \Longrightarrow J(A \cup B) = [\{J(A)\}^{-1} + \{J(B)\}^{-1}]^{-1}$$

I.2. $g(x) = e^{kx}$ $k \geq 0$

$$H(A_1, A_2) = \text{Inf } [\beta, \frac{1}{k} \log \{f[J(A_1)] + f[J(A_2)]\}]$$

$$H(A_1, A_2, A_3) = \text{Inf } [\beta, \frac{1}{k} \log \{f\, J(A_1) + f[J(A_2)] + f[J(A_3)]\}]$$

Siempre utilizando el Teorema 3.8.2, se obtiene

$$(3.9.2)\quad H(\Pi_A) = \text{Inf } [\beta, \frac{1}{k} \log \{\Sigma\, f[J(A_i)]\}]$$

Pertenece a esta familia la entropía de Renyi que se obtiene escogiendo como $J : S \to \mathbb{R}^+$ la medida de información de Shannon y haciendo $\beta = +\infty$ y $f(x) = \exp[(k-1)x]$

I.3. $H(A_1,A_2)=\text{Sup}[\beta, J(A_1), J(A_2)]$

$H(A_1,A_2,A_3) = \text{Sup}[\beta, J(A_1), J(A_2), J(A_3)]$

En este caso particular, el conocer la función $g(x)$ no es necesario para conseguir la forma explícita de $H(\Pi_A)$. La medida de incertidumbre se obtiene en este caso aplicando el Teorema 3.8.1. En consecuencia

(3.9.3) $H(\Pi_A) = \text{Sup}\ \{\beta, J(A_i) : A_i \in \Pi_A\}$

Comparando entre sí las medidas (3.9.1) (3.9.2) y (3.9.3) se nota que las primeras no dependen explícitamente de los eventos con información bastante alta, más precisamente de los eventos A_i tales que $f[J(A_i)] \geq \beta^r$ o $f[J(A_i)] \geq \frac{1}{k}\log\beta$; es decir, que no depende de los eventos con información baja, $J(A_i) \leq \beta$, en forma precisa de los eventos comunes.

Clase II

La sucesión $\{(\alpha_i, \beta_i)\}$ en esta clase estará constituida por un único intervalo $(\alpha, +\infty)$ $\alpha > 0$

II.1 $g(x) = x^r - \alpha^r$ $(r > 0)$ $\forall x > \alpha$

$$H(\{A_1,A_2\}) = \{f_\alpha[J(A_1)] + f_\alpha[J(A_2)] + \alpha^r\}^{1/r}$$

$$H(\{A_1,A_2,A_3\}) = \{f_\alpha[J(A_1)] + f_\alpha[J(A_2)] + f_\alpha[J(A_3)] + \alpha^r\}$$

donde $f_\alpha(x) = f\,[\mathrm{Sup}\,(\alpha,x)]$,

siendo $f : \mathbb{R}^+ \to \mathbb{R}$ una función tal que:

$$f\,[J(A)] + f\,[J(B)] \geq f\,[J(A \cup B)]$$

Utilizando el teorema 3.8.2 se obtiene fácilmente,

$$(3.9.4) \quad H(\Pi_A) = [\,\Sigma\, f_\alpha[J(A_i)] + \alpha^r]^{1/r}$$

Si m es el número de los eventos de Π_A que tienen información menor que α

$$[m = \mathrm{card}\,\{A_i \in \Pi : J(A_i) \leq \alpha\}] \quad \text{se tiene}$$

$$(3.9.5)\ H(\Pi_A) = [m\,.\,c + \sum_{J(A_i)>\alpha} f\,[J(A_i)] + \alpha^r]^{1/r}$$

donde $c = f(\alpha)$. De (3.9.4) o equivalentemente de (3.9.5) se reconoce que siempre hay $H(\Pi_A) \geq \alpha$. En particular, cuando $c = f(\alpha) = 0$, $\mathrm{Sup}\ J(A_i) \leq \alpha \Longrightarrow H(\Pi_A) = \alpha$.

II.2. $H(\{A_1,A_2\}) = \text{Inf}\,[\alpha, \text{Sup}\,\{J(A_1), J(A_2)\}]$

$H(\{A_1,A_2,A_3\}) = \text{Inf}\,[\alpha, \text{Sup}\,\{J(A_1), J(A_2), J(A_3)\}]$

Como en el ejemplo I.3, tampoco en este caso se necesita asignar la función $g : [\alpha, +\infty] \to \mathbb{R}^+$ para lograr determinar la forma explícita de $H(\Pi)$. En efecto, sólo se utiliza el Teorema 3.8.1. para obtener

(3.9.6) $H(\Pi_A) = \text{Inf}\,[\alpha, \text{Sup}\,\{J(A_i) : A_i \in \Pi_A\}]$

En este caso siempre hay $H(\Pi_A) \leq \alpha$ y en particular $H(\Pi_A) = \alpha$ si y solo si $\text{Sup}\, J(A_i) \geq \alpha$.

CAPITULO IV

MEDIDAS DE INCERTIDUMBRE COMPOSITIVAS

En este capítulo examinaremos en detalle la noción de compositividad, su papel en la teoría de la información y su importancia en relacionar entre sí las nociones de información y de incertidumbre.

En resumen, se trata de ésto: sean Π_A y Π_B dos particiones disjuntas y supongamos que conocemos sus medidas de incertidumbre $H(\Pi_A)$ y $H(\Pi_B)$; la propiedad de compositividad caracteriza aquellos espacios para los cuales $H(\Pi_A \vee \Pi_B)$ depende de los soportes A y B de cualquier manera, pero además dependen de las subdivisiones Π_A y Π_B de los propios soportes, solo por $H(\Pi_A)$ y $H(\Pi_B)$.

Se pondrá en evidencia cómo esta clase, formada por espacios donde la medida de incertidumbre $H(\Pi) = H(\{x_1,\dots,x_m\})$ asociada a una partición $\Pi = \{x_1,\dots,x_m\}$, depende de los eventos $x_i \in \Pi$ de una forma especialmente sencilla.

Además, aquí estudiaremos en detalle la noción de compatibilidad universal, clave para enfrentar el problema de la construcción del producto (cartesiano) de espacios de incertidumbre.

Sin embargo, el punto central de este capítulo consiste en analizar la manera (mejor, las maneras) de determinar $H(\Pi_A \vee \Pi_B)$ partiendo de $H(\Pi_A)$ y $H(\Pi_B)$ (y de los soportes A, B).

Persiguiendo como objetivo la construcción de incertidumbre con campo de aplicación más amplio posible, el resultado más importante está representado por el teorema 4.8.1, que logramos demostrar generalizando un resultado análogo obtenido por C. Bertoluzza y M. Schneider [32] bajo hipótesis más restrictivas. Para demostrar este resultado hemos utilizado el mismo esquema de demostración del trabajo de Bertoluzza - Schneider modificándolo donde ha sido necesario. Por otra lado queremos subrayar que, a pesar de que la demostración que hicimos no es muy diferente de la previa, la importancia del resultado final sí lo es ya que nos ha permitido, por primera vez, clasificar en forma sistemática una clase de medidas de incertidumbre no componibles totalmente.

4.1. **Definición 4.1.1.**

Una medida de incertidumbre es "componible simple" si se verifica la propiedad

$$(4.1.1) \quad A,\ B \in \Omega,\ A \cap B = \phi \quad \Longrightarrow$$

$$H(\Pi_A \vee \Pi_B) = F\,[J(A),\ J(B),\ J(A \cup B),\ H(A),\ H(B)]$$

donde la función F tiene como dominio al conjunto

$$\Gamma_5 = \{(x,y,z,u,v)\ /\ 0 \leq z \leq x \leq u,\quad 0 \leq z \leq y \leq v\}$$

Definición 4.1.2

Una medida de incertidumbre es "totalmente componible" si verifica la propiedad

$$(4.1.2) \quad A \cap B = \phi \Longrightarrow H(\Pi_A \vee \Pi_B) = \Psi[J(A),\ J(B),\ H(\Pi_A),\ H(\Pi_B)]$$

donde la función Ψ tiene como dominio al conjunto
$\Gamma_4 = \{(x,y,u,v)\ /\ o \leq x \leq u,\ o \leq y \leq v\}$.

Las definiciones 4.1.1. y 4.1.2 las propuso por primera vez B. Forte [11]. A continuación proponemos una generalización obvia que nos permitirá establecer un teorema de representación muy importante para la construcción de medidas de incertidumbre con características especiales.

Definición 4.1.3.

Una medida de incertidumbre es "componible" si se verifica la propiedad.

(4.1.3) $H(\Pi_A \vee \Pi_B) = M\,[A,\ B,\ H(\Pi_A),\ H(\Pi_B)]$

donde la función M tiene como dominio al conjunto

$\Gamma^{(4)} = \{(A,B,u,v)/(A,B) \in \mathbf{P}(\Omega)^2,\ (u,v) \in \mathbf{R}^2,\ A \cap B = \phi,\ 0 \leq J(A) \leq u,\ 0 \leq J(B) \leq v\}$

Es conveniente subrayar la diferencia entre las definiciones 4.1.1., 4.1.2 y la 4.1.3.

En las primeras, $H(\Pi_A \vee \Pi_B)$ depende de los soporte A, B, y $A \cup B$ (solo en 4.1.1) a través de sus medidas de información, mientras en 4.1.3 la dependencia de esos soportes es completamente arbitraria.

Por otro lado, 4.1.3. generaliza 4.1.1, que a su vez generaliza 4.1.2. Además, cuando la medida de información J es compositiva, 4.1.1. y 4.1.2 son equivalentes

Definición 4.1.4

A las funciones F, Ψ y M las llamaremos "leyes de composición".

4.2. COMPONIBILIDAD TOTAL

Siguiendo el desarrollo cronológico examinaremos en primer lugar la composición total.

Las leyes de composición deben en este caso satisfacer un sistema de ecuaciones funcionales, las cuales deduciremos a continuación, basándonos en propiedades de particiones:

i) $\Pi_A \vee \Pi_B = \Pi_B \vee \Pi_A \Longrightarrow H(\Pi_A \vee \Pi_B) = H(\Pi_B \vee \Pi_A)$

$$\Longrightarrow \Psi[J(A),\ J(B),\ H(\Pi_A),\ H(\Pi_B)] = \Psi[J(B),\ J(A),\ H(\Pi_B),\ H(\Pi_A)]$$

ii) $\Pi_A \vee (\Pi_B \vee \Pi_C) = (\Pi_A \vee \Pi_B) \vee \Pi_C \Longrightarrow H[\Pi_A \vee (\Pi_B \vee \Pi_C)] = [H(\Pi_A \vee \Pi_B) \vee \Pi$

como

$$H(\Pi_A \vee (\Pi_B \vee \Pi_C)) = \psi[J(A),\ J(B \cup C),\ H(\Pi_A),\ H(\Pi_B \vee \Pi_C)]$$

$$= \Psi[J(A),\ J(B)\, T\, J(C),\ H(\Pi_A), \Psi[J(B),\ J(C), H(\Pi_B), H(\Pi_C)]]$$

y además

$$H((\Pi_A \vee \Pi_B) \vee \Pi_C) = \Psi[J(A \cup B),\ J(C),\ H(\Pi_A \vee \Pi_B),\ H(\Pi_C)]$$

$$= \Psi[J(A)\, T\, J(B), J(C), \Psi[J(A), J(B), H(\Pi_A), H(\Pi_B)], H(\Pi_C)]$$

entonces

$$\psi\,[J(A),\ J(B)\,T\,J(C),\ H(\Pi_A), \psi\,[J(B),\ J(C),\ H(\Pi_B),\ H(\Pi_C)]]$$

$$= \psi\,[J(A)\,T\,J(B),\ J(C),\ \psi[J(A), J(B),\ H(\Pi_A),\ H(\Pi_B)],\ H(\Pi_C)]$$

iii) $(\Pi_A \vee \Pi_B) \wedge \Pi_C = (\Pi_A \wedge \Pi_C) \vee (\Pi_B \wedge \Pi_C) \Longrightarrow H[(\Pi_A \vee \Pi_B) \wedge \Pi_C] = H[(\Pi_A \wedge \Pi_C) \vee (\Pi_B \wedge \Pi_C)]$

$$H[(\Pi_A \vee \Pi_B) \wedge \Pi_C] = H[\Pi_A \vee \Pi_B] * H(\Pi_C) \quad (\text{si } (\Pi_A \vee \Pi_B),\ \Pi_C) \in K)$$

$$= \psi\,[J(A),\ J(B),\ H(\Pi_A),\ H(\Pi_B)] * H(\Pi_C)$$

$$H[(\Pi_A \wedge \Pi_C) \vee (\Pi_B \wedge \Pi_C)] = \psi[J(A \cap C),\ J(B \cap C),\ H(\Pi_A \wedge \Pi_C),\ H(\Pi_B \wedge \Pi_C)]$$

$$= \psi\,[J(A) * J(C),\ J(B) * J(C), H(\Pi_A) * H(\Pi_C), H(\Pi_B) * H(\Pi_C)]$$

Por lo tanto

$$\psi[J(A),\ J(B),\ H(\Pi_A),\ H(\Pi_B)] * H(\Pi_C)$$

$$= \psi[J(A) * J(C),\ J(B) * J(C),\ H(\Pi_A) * H(\Pi_C),\ H(\Pi_B) * H(\Pi_C)]$$

iv) $\Pi'_A < \Pi'_A \Longrightarrow \Pi'_A \vee \Pi_B < \Pi''_A \vee \Pi_B$, luego

$$H(\Pi'_A) > H(\Pi''_A) \Longrightarrow H(\Pi'_A \vee \Pi_B) \geq H(\Pi''_A \vee \Pi_B)$$

$$\Longrightarrow \psi\,[J(A), J(B), H(\Pi'_A), H(\Pi_B)] \geq \psi\,[J(A),\ J(B), H(\Pi''_A), H(\Pi_B)]$$

v) $\{A\}\ v\ \{B\} < \{A \cup B\} \Longrightarrow H(\{A\}\ v\ \{B\}) \geq H(\{A \cup B\})$

$$\Longrightarrow \psi[J(A),\ J(B),\ H(\{A\}),\ H(\{B\})] \geq J(A \cup B)$$

$$= J(A)\ T\ J(B)$$

Por lo tanto

$$\psi[J(A),\ J(B),\ H(\{A\}),\ H(\{B\})] \geq J(A)\ T\ J(B)$$

vi) $H(\{A\}\ v\ \{B\}) = \psi\,[J(A),\ J(B),\ H(\{A\}),\ H(\{B\})]$

$$= \psi[J(A),\ J(B),\ J(A),\ J(B)]$$

(por axioma de incertidumbre absoluta). Por lo tanto

$$\psi\,[J(A),\ J(B),\ H(\{A\}),\ H(\{B\})] = \psi\,[J(A),\ J(B),\ J(A),\ J(B)]$$

En conclusión, las leyes de composición total deben satisfacer el siguiente sistema de ecuaciones funcionales:

$$\text{a.}\quad \psi(x,y,u,v) = \psi(y,x,v,u)$$

$$\text{b.}\quad \psi[x,F(y,z),u,\ \psi(y,z,v,w)] = \psi[F(x,y),z,\psi(x,y,u,v)w]$$

$$(4.2.1)\qquad \text{c.}\quad \psi(x,y,u,v) \geq \psi(x,y\ ,x,y) \geq xTy$$

$$\text{d.}\quad u' > u'' \Longrightarrow \psi(x,y,u',v) \geq \psi(x,y,u'',v)$$

$$(4.2.2)\qquad \text{e.}\quad \psi(x,y,u,v) * w = \psi(x * w,\ y * w,\ u * w,\ v * w)$$

donde $x = J(A)$, $y = J(B)$, $z = J(C)$, $u = H(\Pi_A)$ $v = H(\Pi_B)$

$w = H(\Pi_C)$ y $F(x,y)$ es una ley de composición de las medidas de información solución del sistema (4.2.3), que tiene por campo de definición el conjunto

$$\tilde{\Gamma}^{(2)} = \{(x,y) : x = J(A),\ y = J(B),\ A \cap B = \phi\}$$

y que satisface según [4] el sistema

$$\text{(4.2.3)} \quad \begin{array}{ll} \text{a.} & F(x, +\infty) = x \\ \text{b.} & F(x,y) = F(y,x) \\ \text{c.} & F\{F(x,y),z\} = F\{x, F(y,z)\} \\ \text{d.} & x' < x'' \Longrightarrow F(x',y) \leq F(x',y) \\ \text{e.} & F(x,y) * z = F(x * z,\ y * z) \end{array}$$

4.3. NOCION E IMPORTANCIA DEL CONCEPTO DE LEY UNIVERSAL

Toda ley de composición tiene que satisfacer el sistema (4.2.1) con la condición que los argumentos u,v,w,x,y,z sean entropías de particiones pertenecientes a la familia $\mathcal{E}$.

Para la ecuación (4.2.2), la situación se presenta en forma diferente porque esa ecuación tiene que satisfacerse para valores u,v,w,x,y,z que sean medidas de incertidumbre de particiones Π_A, Π_B, Π_C, que por un lado pertenezcan a $\mathcal{E}$, y por otro sean tales que $(\Pi_A \vee \Pi_B, \Pi_C) \in \mathcal{K}$.

Por supuesto, si $\mathcal{K} = \phi$, entonces la ecuación (4.2.2) que llamaremos "ecuación de compatibilidad" entre las leyes de composición y de independencia no impone limitaciones a la forma de la ley de composición, sin embargo, sí es un enlace para la escogencia de dichas leyes cuando $\mathcal{K} \neq \phi$.

Estudiaremos más adelante algunas clases particulares de leyes de composición; describiremos aquellas que pueden ser utilizadas, no sólo en un espacio de incertidumbre particular, sino en todos los espacios que posean una característica asignada, relacionada con la noción de independencia. Para esto será necesario introduir y analizar la noción de universalidad según vamos a exponer a continuación.

4.4. LEYES G-UNIVERSALES

Sea $G : \mathbb{R}^+ \times \mathbb{R}^+ \to \mathbb{R}^+$ $((x,y) \xrightarrow{G} G(x,y) = x * y)$ una posible ley de independencia; es decir, una función de la forma (1.7.1)

Definición 4.4.1

Diremos que una función $\psi : \Gamma_4 \to \mathbb{R}^+$, $[(x,y,u,v) \xrightarrow{\psi} \psi(x,y,u,v)\ \forall\ (x,y,u,v) \in \Gamma_4]$, es una "ley de composición (total) universal respecto a la función G" asignada, si es compatible con todos los espacios de incertidumbre $(\mathbf{I},\mathcal{E},H)$ en los cuales la ley de independencia sea justamente la función G(o $*$ en notación algebraica).

Diremos también, abreviando, que ψ es "G-universal" o "$*$-universal".

La característica principal de las leyes G-universales consiste en que son funciones que tienen que satisfacer el sistema (4.2.1), (4.2.2) para todos los valores u,v,w,x,y,z con $o \leq x \leq u$, $o \leq y \leq v$, $o \leq z \leq w$. En efecto, como quiera se escojan esos valores:

a. Siempre existe un espacio de incertidumbre $\mathbb{H}_1 = (\mathcal{H}_1, H_1)$ con $\mathcal{H}_1 = (\mathbf{I}_1, \mathcal{E}_1)$, donde la familia $\mathcal{E}_1$ contiene tres particiones disjuntas Π_A, Π_B, Π_C con $H_1(\Pi_A) = u$, $H_1(\Pi_B) = v$, $H_1(\Pi_C) = w$ y además $J_1(A) = x$, $J_1(B) = y$, $J_1(C) = z$.

b. Siempre existe un espacio de incertidumbre $\mathbb{H}_2 = (\mathcal{H}_2, H_2)$ con $\mathcal{H}_2 = (\mathbf{I}_2, \mathcal{E}_2)$ (no necesariamente igual a $\mathbb{H}_1$) donde la familia $\mathcal{E}_2$ contiene tres particiones Π_A, Π_B, Π_C con $H_2(\Pi_A) = u$, $H_2(\Pi_B) = v$, $H_2(\Pi_C) = w$, $J_2(A) = x$, $J_2(B) = y$, $J_2(C) = z$, tales que $A \cap B = \phi$, $(\Pi_A \vee \Pi_B, \Pi_C) \in \mathcal{K}_2$.

Y además los espacios $\mathbb{H}_1$ y $\mathbb{H}_2$ tienen como ley de independencia la función $G(x,y)$.

Observación 4.4.1

La noción de G-universalidad se puede introducir aún para leyes de composición no totales. Sin embargo, no analizaremos este caso debido a que no existen hasta ahora resultados en este sentido.

Para proseguir el análisis de las leyes universales, es necesario recordar que en los espacios totalmente componibles se tiene

$$A \cap B = \phi \implies J(A \cup B) = F\,[J(A),\, J(B)] = J(A)\ T\ J(B)$$

Además la misma ley de composición de la información $F(x,y)$ se encuentra también en el sistema (4.2.1). Por ésto, antes de buscar cualquier ley de composición (total) de la incertidumbre es necesario asignar la ley de composición de la información.

Por razones que aclararemos más adelante, si nuestro objetivo es conseguir las leyes $\psi(x,y,u,v)$, G-universales, entonces es necesario que también $F(x,y)$ como ley de composición de la información sea G-universal (en el sentido del § 1.10).

Así pues, muchos problemas relacionados con tales funciones $\psi(x,y,u,v)$, G-universales, se plantean y se resuelven mejor particularizando esa ley ψ en el siguiente sentido. Sea $G(x,y) = x * y$ una ley de independencia asignada y sea $F(x,y) = xTy$ una ley de composición de la información universal respecto a G.

Definición 4.4.2

Diremos que una función $\psi(x,y,u,v)$ definida para cada $(x,y,u,v) \in \Gamma_4$ es una "ley de composición (total, (G,F) - universal", si es compatible con todos los espacios de incertidumbre $\mathcal{H} = (H,H)$

donde la ley de independencia es G e $\mathbf{I}$ es un espacio componible con ley F(x,y) universal respecto a G.

Según como se indican las leyes G y F, conviene escribir (G,F)-universal o (*,T)-universal; el sentido es el mismo.

Ya hemos visto, al examinar las medidas de información, que la noción de ley de composición G-universal juega un papel muy importante en la construcción de espacios que sean producto de otros espacios.

Lo mismo para cuando se trata de construir productos de espacios de incertidumbre; claro que en este caso el concepto apropiado será el de (G,F)-universalidad.

No hace falta exponer el argumento con muchos detalles ya que es lo mismo (con obvias modificaciones) expuesto en el § 1.8.

En resumen: tenemos dos espacios de incertidumbre $\mathbf{H}_1 = (\mathcal{H}_1, H_1)$ y $\mathbf{H}_2 = (\mathcal{H}_2, H_2)$ y queremos definir en forma coherente una medida de incertidumbre H sobre la estructura producto

$$\mathcal{H} = (\mathbf{I}, \mathcal{E}) = (\mathbf{I}_1 \times \mathbf{I}_2, \mathcal{E}_1 \times \mathcal{E}_2).$$

Para ésto pedimos que:

a. El espacio I sea el producto de I_1 e I_2 según lo especificado en el § 1.8.

b. Para cada $\Pi_1 \in E_1$, $\Pi_2 \in E_2$ se cumpla

$$H(\Pi_1 \times \Omega_2) = H_1(\Pi_1), \; H(\Omega_1 \times \Pi_2) = H_2(\Pi_2)$$

Examinamos ahora si es posible que H sea componible, cuando lo son los espacios H_1 y H_2. El problema enunciado así queda bastante indeterminado y permite miles de contestaciones. Para precisar mejor, supongamos que:

1. $A_1 \in E_1, A_2 \in E_2 \Longrightarrow (A_1 \times \Omega_2, \; \Omega_1 \times A_2) \in K$

2. Los espacios I, I_1, I_2 sean componibles con leyes de composición $F = F_1 = F_2$ G-universales.

3. La ley de composición total $\psi(x,y,u,v)$ del espacio H sea la misma para todas las parejas de espacios H_1, H_2 que tengan leyes de composición $\psi_1(x,y,u,v) = \psi_2(x,y,u,v)$ asignadas.

Fácilmente se puede reconocer que

$$\psi(x,y,u,v) = \psi_1(x,y,u,v) \; [= \psi_2(x,y,u,v) \text{ (por 3.)}]$$

y además la función $\psi(x,y,u,v)$ debe ser (G,F) universal.

Observación 4.4.2

Las condiciones 1., 2., 3., parecen artificiosas, pero son las más naturales para imponer. En realidad por un lado facilitan la búsqueda de las leyes de composición y por otro determinan clases de leyes cuya utilización es independiente del espacio de incertidumbre particular que se quiere utilizar.

Presentaremos en los próximos párrafos los teoremas que permiten establecer la forma general de las leyes de composición compatibles universalmente con las tres siguientes leyes de independencia.

(4.4.1) $G(x,y) = x * y = x + y$

(4.4.2) $G(x,y) = x * y = \mathrm{Sup}(x,y)$

(4.4.3) $G(x,y) = x * y = g^{-1}\{g(x) + g(y)\}$

Por supuesto, en cada uno de los casos se determinarán las leyes (G,F)-universales para todas las funciones $F(x,y) = xTy$ compatibles con $G(x,y)$.

Observación 4.4.3.

(4.4.1) es un caso particular de (4.4.3). Sin embargo, expondremos este caso separadamente, ya sea por su importancia y porque los resultados obtenidos en el caso (4.4.1) han sido utilizados para desarrollar (4.4.3).

Observación 4.4.4

Ha sido establecido el teorema de representación también en el caso donde la ley de independencia $G(x,y)$ asume su forma más general (1.7.1). Sin embargo, tanto el resultado como la demostración son extremadamente complejos, y quizás poco útiles, por lo cual no quisimos exponerlos aquí. Quienes estén interesados en leerlos pueden conseguirlos en [5,24].

4.5. LEYES UNIVERSALES CLASICAS

En la formulación clásica, el axioma de independencia se escribe bajo la forma

$$(\Pi_A, \Pi_B) \in K \implies H(\Pi_A \wedge \Pi_B) = H(\Pi_A) + H(\Pi_B).$$

En otras palabras, en los espacios de incertidumbre clásicas se tiene

$$G(x,y) = x * y = x + y \tag{4.5.1}$$

por lo tanto, las leyes de composición (total), de las incertidumbres universales, en el sentido clásico son las soluciones del sistema (4.2.1) donde la ecuación de compatibilidad asume la forma

$$\psi(x+z,\ y+z,\ u+w,\ v+w) = w + \psi(x,y,u,v) \tag{4.5.2}$$

y donde, según lo expuesto en el parágrafo 1.3, la función $F(x,y)$ puede asumir una de las formas siguientes:

I) $F_i(x,y) = x\ T_i\ y = \mathrm{Inf}\ (x,y)$

II) $F_c(x,y) = x\ T_c\ y = c \log\ [e^{-x/c} + e^{-y/c}]\quad c > 0$

siendo ésas las únicas leyes de composición de la información compatibles universalmente con la ley (4.5.1).

P. Benvenuti y B. Forte consiguieron la solución del sistema (4.2.1) y (4.5.2) en ambos casos (I y II) [10,22,23]. A continuación serán expuestos los teoremas de Representación relativos.

Teorema 4.5.1

El conjunto de las leyes de composición, de las incertidumbres $(+, T_i)$-universales, está compuesto por las tres clases de funciones siguientes:

(4.5.3) $\psi(x,y,u,v) = \mathrm{Inf}(u-\lambda x, v-\lambda y) + \lambda\, \mathrm{Inf}(x,y) \qquad \lambda \leq 1$

(4.5.4) $\psi(x,y,u,v) = \mathrm{Sup}(u-\lambda x, v-\lambda y) + \lambda\, \mathrm{Inf}(x,y) \qquad \lambda \leq 1$

(4.5.5) $\psi(x,y,u,v) = \frac{1}{k} \log\ [e^{k(u-\lambda x)} + e^{k(v-\lambda y)}] + \lambda\, \mathrm{Inf}(x,y) \qquad k \geq 0$

Teorema 4.5.2

El conjunto de las leyes de composición de las incertidumbre $(+, T_c)$-universales, está compuesto por las cuatro clases de funciones escritas a continuacion:

$$(4.5.6)\quad \psi(x,y,u,v) = \operatorname{Inf}(u-\lambda x, v-\lambda y) - \lambda c \log (e^{-x/c} + e^{-y/c}) \qquad \lambda \leq 1$$

$$(4.5.7)\quad \psi(x,y,u,v) = \operatorname{Sup}(u-\lambda x, v-\lambda y) - \lambda c \log (e^{-x/c} + e^{-y/c}) \qquad \lambda \leq 1$$

$$(4.5.8)\quad \psi(x,y,u,v) = \frac{1}{k} \log [e^{k(u-\lambda x)} + e^{k(v-\lambda y)}] - \lambda c \log [e^{-x/c} + e^{-y/c}] \qquad k \geq \lambda c$$

$$(4.5.9)\quad \psi(x,y,u,v) = \frac{(u-\lambda x)\, e^{-x/c} + (v-\lambda y)\, e^{-y/c}}{e^{-x/c} + e^{-y/c}} - \lambda c \log [e^{-x/c} + e^{-y/c}] \qquad \lambda \leq 1$$

Las líneas fundamentales de la demostración de estos teoremas serán esbozadas en el Apéndice D.

4.6. LEYES SUP-UNIVERSALES

El teorema expuesto a continuación determina la forma de todas las leyes de composición de la incertidumbre sup-universales, o sea, compatibles con la ley de independencia.

$$(4.6.1)[a]\quad G_s(x,y) = x *_s y = \sup(x,y)$$

la cual, según lo demostrado en [6] es compatible sólo con la ley de composición de la información

$$(4.6.1)[b] \quad F(x,y) = F_I^U(x,y) = \mathrm{Inf}(x,y) = x \wedge y$$

La solución general del sistema (4.2.1), (4.2.2) con G(x,y)= = Sup(x,y) ha sido conseguida por C. Bertoluzza y F. Barbaini [8], quienes demostraron el siguiente teorema:

Teorema de Representación 4.5.1

Sea e y k dos números reales tales que $o \leq e \leq k$ y $h : [k, +\infty] \to [o,k]$ una función con las siguientes propiedades:

i) $h(k) = k$

ii) $x < y \implies h(x) \geq h(y)$

Cada función definida haciendo

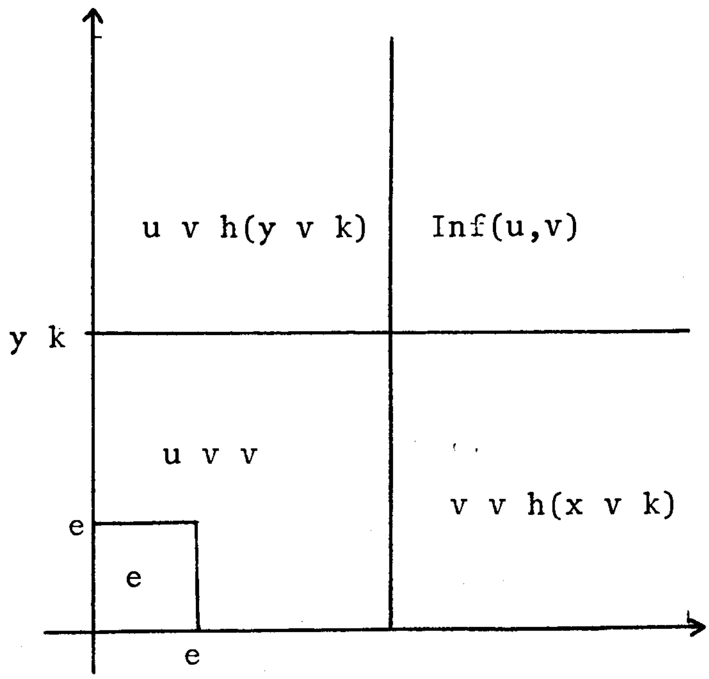
$$(4.6.2) \quad \psi(x,y,u,v) = \begin{cases} e & \text{si } (u,v) \in [o,e]^2 \\ u \vee v & \text{si } (u,v) \in [o,k]^2 - [o,e]^2 \\ \mathrm{Inf}(u,v) & \text{si } (u,v) \in [k,+\infty]^2 \\ v \vee h(x \vee k) & \text{si } (u,v) \in [k,+\infty) \times [o,k] \\ u \vee h(y \vee k) & \text{si } (u,v) \in [o,k] \times [k,+\infty) \end{cases}$$

(donde $a \vee b = \mathrm{Sup}(a,b)$)

es una solución del sistema (4.2.1) (4.2.2) (4.6.1); recíprocamente, cada solución del mismo sistema es una función definida por (4.6.2) con una selección adecuada de los números e y k y de la función h.

(Para la demostración ver Apéndice E)

Gráficamente, la solución ψ queda representada así:



4.7. LEYES g-UNIVERSALES

Con esta notación designamos aquí las leyes compatibles con la ley de independencia.

(4.7.1) $$x * y = g^{-1}\{g(x) + g(v)\}$$

Siendo $g : \mathbb{R}^+ \to \mathbb{R}^+$ una función estrictamente creciente y nula en el origen.

Para conseguir la forma general de esas leyes se efectúa en el sistema (4.2.1)(4.2.2.) un cambio de variables, haciendo

$$t^* = g(t) \quad (x^* = g(x),\ w^* = g(w), \dots),$$

y también se introducen las funciones

$$\psi^*(x^*,y^*,u^*,v^*) = g_\circ\ \psi\ [g^{-1}(x^*), g^{-1}(y^*),\ g^{-1}(u^*),\ g^{-1}(v^*)]$$

$$F^*(x^*,u^*) = g_\circ\ F\ [g^{-1}(x^*),\ g^{-1}(y^*)]$$

donde: ψ es la antitransformada de ψ^* y $F(r,s)$ es una de las funciones

(4.7.2) $$F(r,s) = \mathrm{Inf}(r,s)$$

(4.7.3) $$F(r,s) = -c \log (e^{-r/c} + e^{-s/c})$$

Después de realizar esos cambios, el sistema (4.2.1)(4.2.2) queda idéntico al estudiado en relación a las leyes universales clásicas. Las funciones $\psi^*(x^*,y^*,u^*,v^*)$, por lo tanto, pueden asumir sólo las formas expuestas en los teorema 4.5.1, 4.5.2.

Finalmente, las leyes de composición g-universales serán las antitransformadas de esas funciones. Por ejemplo, una de esas leyes es la siguiente

$$\psi(x,y,u,v) = g^{-1}[\frac{1}{k}\log\{e^{k[g(u)-\lambda g(x)]} + e^{k[g(v)-\lambda g(y)]}\}$$

$$+\lambda g^{-1} \circ \operatorname{Inf}[g(x), g(y)]]$$

y las otras se construyen de igual forma.

4.8. LEYES ESTRICTAMENTE MONOTONAS

En este paragrafo expondremos y demostraremos un teorema de representación (en cierto sentido el más general obtenido hasta ahora) de las medidas de incertidumbre componibles según la definición 4.1.3. Su generalidad se debe:

a. Por un lado al hecho que la ley de composición depende de los soportes A y B de la forma más general posible.

b. Por otro lado, al hecho que no se utiliza la ecuación de compatibilidad y por tanto las leyes de composición así obteni--

das constituyen una clase muy amplia que contiene, como subclases, leyes de composición compatibles con cualquier familia K de eventos independientes y con muchas posibles leyes de independencia.

Por cierto, ese teorema no es el más general posible, ya que ha sido establecido bajo una hipótesis restrictiva de monotonía estricta (por ésta razón no quedan en esta clase, por ejemplo, las leyes sup-universales).

Según la definición 4.1.3, la función M sirve para determinar la incertidumbre de la unión de dos particiones disjuntas Π_A y Π_B por medio de la relación (4.1.3).

La operación "v" es simétrica, asociativa, conserva el orden <; además, Π_A y Π_B siempre son refinamientos de sus soportes A y B. Por lo tanto la función M:

a. Tiene como dominio el conjunto

$$\Gamma^{(4)} = \{A,B,u,v)/(A,B) \in \mathbf{P}(\Omega)^2, (u,v) \in \mathbf{R}^2, A \cap B = \phi,\ o \leq J(A) \leq u,\ o \leq I(B) \leq v\}$$

b. Satisface el siguiente sistema de ecuaciones funcionales

$$\begin{array}{lll} & a. & M(A,B,u,v) = M(B,A,v,u) \\ & b. & M[A \cup B, C, M(A,B,u,v), w] = M[(x, B \cup C, u, M(B,C,v,w)] \\ (4.8.1) & c. & u_1 < u_2 \implies M(A,B,\mathbf{u}_1,v) \leq M(A,B,u_2,v) \\ & d. & M(A,B,u,v) \geq M(A,B,J(A),J(B), \geq A \cup B \\ \text{donde} & & u = H(\Pi_A),\ v = H(\Pi_B),\ w = H(\Pi_C) \end{array}$$

Todavía no ha sido conseguida la solución general de ese sistema (que sí representaría la totalidad de las leyes de composición). Sin embargo, hemos logrado obtener la solución de ese sistema sustituyendo la condición c. por la más restrictiva .

c') $u_1 < u_2 \rightarrow M(A,B,u_1,v) < M(A,B,u_2,v)$

Teorema 4.8.1

Todas las leyes de composición continuas, estrictamente monótonas son las variables u y v tienen la siguiente expresión

(4.8.2) $M(A,B,u,v) = f^{-1}_{A\cup B}[f_A(u) + f_B(v) + c]$

donde $f_A(u) = f(A,u)$ es una función definida en el conjunto

$$\Gamma^2 = \{(A,u) : A \in \mathbf{P}(\Omega), u \in R^+, o \leq J(A) \leq u\}$$

con las siguientes propiedades:

a. Es estrictamente monótona en la variable u

b. $f_A[J(A)] + f_B[J(B)] + c \geq f_{A\cup B}[J(A\cup B)]$ si $f_x(.)$ es creciente

$f_A[J(A)] + f_B[J(B)] + c \leq f_{A\cup B}[J(A\cup B)]$ si $f_x(.)$ es decreciente

Recíprocamente, cada función de la forma (4.8.2) representa una posible ley de composición.

Demostración:

Haciendo

(4.8.3) $J(u,v) = M[A,B,u,v]$

(4.8.4) $F(s,w) = M[A \cup B,C,s,w]$

(4.8.5) $H(u,t) = M[A,B \cup C,u,t]$

(4.8.6) $K(v,w) = M[B,C,v,w]$

entonces la ecuación (4.8.1.b) se puede escribir así:

(4.8.7) $F(J(u,v),w) = H(u,k(v,w))$

Esta ecuación (donde las incógnitas son las cuatro funciones F, J, H, K) fué resuelta por J. Aczel [25] bajo las hipótesis de que J, F, H, K son inversibles con respecto a sus dos variables. Estas condiciones son satisfechas si la función M es estrictamente monótona con respecto a u, es decir, bajo la condición C'

La solución general de la ecuación (4.8.1.b) está dada bajo las condiciones:

(4.8.8) $F(s,w) = \ell\,(f(s) + g(w))$

(4.8.9) $H(u,t) = \ell\,(k(u) + h(t))$

(4.8.10) $J(u,v) = f^{-1}(k(u) + m(v))$

(4.8.11) $K(v,w) = h^{-1}(m(v) + g(w))$

donde f, g, h, k, ℓ son funciones estrictamtente monótonas.

Entonces resulta:

$$(4.8.12)\quad M(A,B,u,v) = f^{-1}_{A,B}\,\{k_{A,B}(u) + m_{A,B}(v)\}$$

$$(4.8.13)\quad M(A\cup B,C,u,v) = \ell_{A\cup B,C}\,\{f_{A\cup B,C}(u) + g_{A\cup B,C}(v)\}$$

$$(4.8.14)\quad M(B,C,u,v) = h^{-1}_{B,C}\,\{m_{B,C}(u) + g_{B,C}(v)\}$$

$$(4.8.15)\quad M(A,B\cup C,u,v,) = \ell_{A,B\cup C}\,\{k_{A,B\cup C}(u) + h_{A,B\cup C}(v)\}$$

donde se han puesto a las funciones f,g,h,k,ℓ los índices A,B,C o sus combinaciones de acuerdo con las definiciones de F, J,H,K. Por otro lado, en las representaciones (4.8.8)-(4.8.11) la función f que aparece en (4.8.8) es la misma que aparece en (4.8.10); ésto significa que, fijados tres conjuntos A,B,C la función $f_{A,B}$ de la relación (4.8.12) coincide con $f_{A\cup B,C}$, de la relación (4.8.13)

$$f_{A,B}(\cdot) = f_{A\cup B,C}(\cdot)$$

Esto significa que f depende sólo de A y B (más precisamente, depende de $A\cup B$) y no de C. Por razonamientos análogos se reconoce que

k depende solo de A

m " " " B

ℓ " " " $A\cup B\cup C$

g " " " C

h " " " $B\cup C$

De acuerdo a ésto, las relaciones (4.8.12) a (4.8.15) se pueden escribir así:

(4.8.16) $M(A,B,u,v) = f^{-1}_{A\cup B}\ \{K_A(u) + m_B(v)\}$

(4.8.17) $M(A\cup B,C,u,v) = \ell_{A\cup B\cup C}\ \{f_{A\cup B}(u) + g_C(v)\}$

(4.8.18) $M(B,C,u,v) = h^{-1}_{B\cup C}\ \{m_B(u) + g_C(v)\}$

(4.8.19) $M(A,B\cup C,u,v) = \ell_{A\cup B\cup C}\ \{K_A(u) + h_{B\cup C}(v)\}$

Esta cuatro definiciones de la función M son coherentes entre éllas si:

$$m_D(z) = g_D(z) + C_1 \quad ; \quad m_D(z) = h_D(z) + C_3$$

$$m_D(z) = f_D(z) + C_2 \quad ; \quad m_D(z) = \ell^{-1}_D(z) + C_4$$

$$m_D(z) = K_A(z) + C_5$$

En efecto:

Como $M(A,B,u,v) = M(B,A,v,u)$, de (4.7.13) se obtiene:

$$f^{-1}_{A\cup B}\ [K_A(u) + m_B(v)] = f^{-1}_{B\cup A}\ [K_B(v) + m_A(u)]$$

o sea, como $f_{A\cup B} = f_{B\cup A}$, entonces

$K_A(u) = m_A(u) = K_B(v) - m_B(v) =$ Constante y por consiguiente

$$K_A(u) = m_A(u) + C_5$$

Con procedimientos análogos se obtienen otras relaciones entre las demás funciones f_A, h_A, K_A, ℓ_A, g_A, m_A (a veces hay que suponer vacíos algunos conjuntos; por ejemplo hay que poner $B = \phi$ en (4.8.13) para demostrar que $f_A(u) = g_A(u) + C$). Utilizando estas relaciones se reconoce que la diferencia entre dos cualesquiera de las seis funciones es una constante.

En particular, si en (4.8.17) $B = \phi$ obtenemos:

$$\begin{aligned} M(A\cup\phi,C,u,v) &= M(A,C,u,v) \\ &= \ell_{A\cup\phi\cup C}\{f_{A\cup\phi}(u) + g_C(v)\} \\ &= \ell_{A\cup C}\{f_A(u) + m_C(v) - C_1\} \\ &= \ell_{A\cup C}\{f_A(u) + f_C(v) + C_2 - C_1\} \end{aligned}$$

Si en (4.8.16) hacemos $B = C$, tenemos

$$\begin{aligned} M(A,C,u,v) &= f^{-1}_{A\cup C}\{K_A(u) + m_C(v)\} \\ &= f^{-1}_{A\cup C}\{K_A(u) + f_C(v) + C_2\} \\ &= f^{-1}_{A}\{K_A(u) + f_C(v) + C_2\} \end{aligned}$$

Luego

$$f_A^{-1}\{K_A(u) + f_C(v) + C_2\} = \ell_{A\cup C}\{f_A(u) + f_C(v) + (C_2 - C_1)\}$$

Por lo tanto, en definitiva se tiene

$$(4.8.20) \quad M(A,B,u,v) = f^{-1}_{A\cup B}\,\{f_A(u) + f_B(v) + C\}$$

donde $f_A(u)$ es una función estrictamente monótona en u para todo A.

La función M definida por (4.8.20) satisface todas las ecuaciones del sistema (4.8.1) excepto, posiblemente, la ecuación (4.8.1,d). Para que también sea satisfecha es necesario y suficiente que la función f(A,u) cumpla con la siguiente condición.

$$(4.8.21) \quad \begin{cases} f_A\,[J(A)] + f_B\,[J(B)] + C \geq f_{A\cup B}\,[J(A\cup B)] & \text{si } f_x(x) \text{ es creciente} \\ f_A\,[J(A)] + f_B\,[J(B)] + C \leq f_{A\cup B}\,[J(A\cup B)] & \text{si } f_x(x) \text{ es decreciente} \end{cases}$$

□

4.9. LA ENTROPIA DE HARTLEY

La "Entropía de Hartley" definida sobre una estructura de incertidumbre $(\mathbf{I},\mathcal{E})$ es la aplicación $H : \mathcal{E} \to \mathbf{R}^+$ definida por

$$H(\Pi_A) = \log\,[n(\Pi_A)] + J(A)$$

siendo $n(\Pi_A)$ el número de subconjuntos que integran la partición Π_A.

Tal vez esa entropía sea la más simple posible debido a que la incertidumbre sobre el éxito de una experiencia depende sólo del número de éxitos posibles y de la información asociada al soporte.

Por supuesto, esa medida de incertidumbre es totalmente componible con ley de composición.

$$\psi(x,y,u,v) = xTy + \log \quad e^{(u-x)} + e^{(v-y)}$$

que es un caso particular de las expresiones (4.5.5) y (4.5.8) con $k = \lambda = 1$. Por lo tanto su ley de composición es universal en el sentido clásico.

Por otro lado es bueno examinar esta medida de incertidumbre también desde un punto de vista diferente, después de observar que una medida de incertidumbre es componible, también es componible la incertidumbre condicional asociada, que en los espaciones clásicos está definida por

$$K(\Pi_A) = H(\Pi_A) - J(A).$$

Además, en los espacios clásicos, entre la ley de composición de $H(\psi(x,y,u,v))$ y la de $K(\Psi(x,y,u,v)$ siempre existe la relación

$$(4.9.1) \qquad \psi(x,y,u,v) = xTy + \Psi(x,y,u\text{-}v,v\text{-}y)$$

Por supuesto, podemos construir la entropía condicional asociada a una partición $\Pi_A = \{A_1,\ldots,A_n\}$ por medio del procedimiento iterativo

$$F_1(x) = x, \; F_{n+1}(x_1,\dots,x_{n+1}) = F\,[F_n(x_1\dots,x_n),\; x_{n+1}]$$

$$\Psi_1(x) = 0$$

$$\Psi_{n+1}(x_1,\dots,x_{n+1}) = \Psi[F_n(x_1,\dots,x_n), x_{n+1}, \Psi_n(x_1,\dots,x_n),\; 0]$$

$$K(\Pi_A) = \Psi_n\,[J(A_1),\dots,J(A_n)]$$

Para cada $n > 1$, la función Ψ_n está definida en el conjunto $\Gamma_n = \{(x_1,\dots,x_n) : F_n(x_1,\dots,x_n) \geq 0\}$, ya que $\Delta_n = Rgo(\Psi_n)$. Por hipótesis F y Ψ son continuas, luego Δ_n es un intervalo (abierto, semiabierto o cerrado).

Nos interesa el caso donde, para cada $n \in \mathbb{N}$, Δ_n se reduce a un solo punto. Tenemos entonces

$$\Psi_n(x_1,\dots,x_n) = S(n)$$

Como Ψ_n determina una medida de incertidumbre condicional, la función $S(n)$ tiene que cumplir con las siguientes relaciones:

$$S(1) = 0$$
$$S(n+1) \geq S(n)$$
$$S(m.n) = S(m) + S(n)$$

Por lo tanto tenemos que

$$S(n) = K.\log n$$
$$K(\Pi_A) = K.\log\,[n(\Pi_A)]$$
$$H(\Pi_A) = J(A) + K.\log\,[n(\Pi_A)].$$

La entropía de Hartley queda caracterizada así por la estructura de los conjuntos Δ_n.

4.10. CONSTRUCCION DE LAS MEDIDAS DE INCERTIDUMBRE COMPOSITIVAS

Método de las agregaciones sucesivas

Con este método se construye la experiencia Π_A a partir de sus átomos $A_1,\ldots,A_n$, formando una sucesión de particiones

$$S = \{\Pi^\circ = \{\phi\},\ \Pi^1,\Pi^2,\ldots,\Pi^m : m \leq n\}$$

donde, partiendo de $\Pi^\circ=\{\phi\}$, las particiones de S se obtiene "agregando" sucesivamente un cierto número de átomos de Π_A que no hayan sido anteriormente agregados, hasta agotarlos todos de manera que Π^m coincida con Π_A (V. Fig. 1)

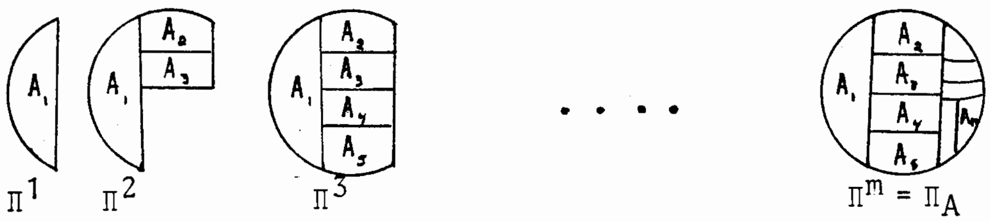


FIGURA 1

En forma más precisa, Π^{i+1} se obtiene al unir (en el sentido de unión de particiones (Ver. 2.2)) las particiones $\Pi^i=\{A_1,..,A_{r_i}\}$ de soporte $A^* = \bigcup_{j=1}^{r_i} A_j$ y $\Pi_i = \{A_{r_{i+1}},..,A_{s_i}\}$ de soporte $A_* = \bigcup_{j=r_{i+1}}^{s_i}$ disjunto de A^* (Ver Fig. 2).

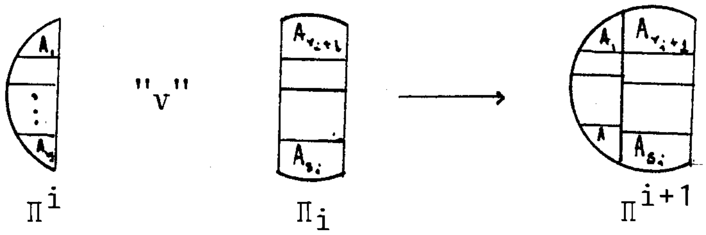


FIGURA 2

Entre todas las sucesiones S que llevan a Π_A, la más simple (y conveniente) se realiza por el proceso que consiste en unir cada vez a la partición Π^i un átomo A_t de Π_A para obtener Π^{i+1}. Obviamente en este caso $m = n = \text{card}\ (\Pi_A)$, $\Pi^i = \{A_1,\ldots,A_i\}$, $\Pi_i = \{A_{i+1}\}$ (Ver Fig. 3)

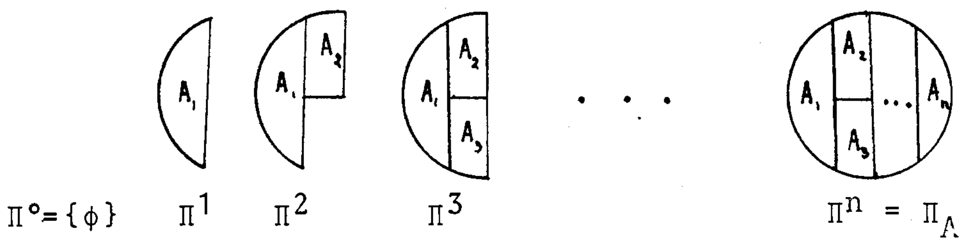


FIGURA 3

Las consideraciones hechas en el método de las subdivisiones sucesivas (Ver 3.5) son válidas, con las adecuadas modificaciones también para este método, con las diferencias que los caracterizan, es decir:

1) Tomamos en cuenta los espacios componibles, con lo cual resulta.

$$H(\Pi^{i+1}) = M\,[A^*, A_*, H(\Pi^i), H(\Pi_i)] \tag{4.10.1}$$

El conocer la ley de composición M y otras pocas cantidades nos permitirá construir la medida de incertidumbre asociada a cualquier partición Π_A.

4.11. FORMULA RECURRENTE

Asignada la ley de composición M (o ψ), si la sucesión $\{\Pi^i\}$ escogida para construir Π_A es la descrita por la (Fig. 3) del § 4.10, entonces la fórmula recurrente que se utiliza para determinar la sucesión $\{H(\Pi^i)\}$ es la siguiente:

$$\begin{aligned} & H(\Pi^{i+1}) = M\left[\bigcup_{j=1}^{i} A_j, A_{i+1}, H(\Pi^i), J(A_{i+1})\right] \\ (4.11.1) \quad & \\ & H(\Pi^1) = J(A_1) \end{aligned}$$

Se reconoce inmediatamente que para realizar este procedimiento es suficiente conocer, además de la función M, la medida de información $J : S \to \mathbb{R}^+$.

4.12. FORMA EXPLICITA DE LAS MEDIDAS DE INCERTIDUMBRE COMPOSITIVAS

En este parágrafo se proporcionará la forma explícita de la medida de incertidumbre de una partición $\Pi_A = \{A_1,\ldots,A_n\}$ en función de las medidas de Información $J(A_1)\ldots,J(A_n)$ de sus átomos, que podemos arreglar, sin perder generalidad, de manera que $J(A_1) \leq J(A_2) \leq \ldots \leq J(A_n)$.

Aplicando a las leyes de composición obtenidas en los parágrafos 4.5, 4.6, 4.7 y 4.8 el procedimiento recurrente antes descrito, se obtienen las siguientes expresiones para las medidas de incertidumbre correspondientes:

I. Medidas $(+, T_i)$ - universales

(4.12.1) $H(\{A_1,\ldots,A_n\}) = J(A_1)$

(4.12.2) $H(\{A_1,\ldots,A_n\}) = (1-\lambda)\ J(A_n) + \lambda\, J(A_1)$

(4.12.3) $H(\{A_1,\ldots,A_n\}) = \frac{1}{k} \log[\Sigma\, e^{k(1-\lambda)J(A_i)}] + \lambda\, J(A_1)$

II. Medidas $(+,T_c)$ - universales

(4.12.4) $H(\{A_1,\ldots,A_n\}) = (1-\lambda)\ J(A_1) - \lambda\, c \log\ [\Sigma\, e^{-J(A_i)/c}]$

(4.12.5) $H(\{A_1,\ldots,A_n\}) = (1-\lambda)\ J(A_n) - \lambda\, c \log\ [\Sigma\, e^{-J(A_i)/c}]$

$$(4.12.6)\quad H(\{A_1,\dots,A_n\}) = \frac{1}{k}\log\left[\Sigma\, e^{k(1-\lambda)\,J(A_i)}\right] -$$

$$-\lambda\, c\,\log\left[\Sigma\, e^{-J(A_i/c}\right]$$

$$(4.12.7)\quad H(\{A_1,\dots,A_n\}) = \frac{\Sigma\,(1-\lambda)\,J(A_i)\,e^{J(A_i)}}{\Sigma\, e^{J(A_i)}} -$$

$$-\lambda\, c\,\log\left[\Sigma\, e^{-J(A_i)\,/\,c}\right]$$

III. Medidas (Sup, Inf) - universales

$$(4.12.8)\quad H(\{A_1,\dots,A_n\}) = \mathrm{Inf}\,\{J(A_1),\dots,J(A_n)\} \quad \text{si } J(A_1) > k$$

$$(4.12.9)\quad H(\{A_1,\dots,A_n\}) = \mathrm{Sup}\,\{h\,[J(A_{i(k)+1})],\ J(A_{i(k)}),\ e\} \quad \text{si } J(A_1) < k \leq J(A_n)$$

$$(4.12.10)\quad H(\{A_1,\dots,A_n\}) = \mathrm{Sup}\,\{e,\ J(A_1),\dots,J(A_n)\} \quad \text{si } J(A_n) \leq k$$

donde $o \leq e \leq k \leq +\infty$ e $i(k)$ es el índice para el cual se cumple

$$J(A_{i(k)}) \leq k < J(A_{i(k)+1})$$

IV. Cuando la ley de composición es (4.8.2) se obtiene

$$(4.12.11)\quad H(\{A_1,\dots,A_n\}) = f_A^{-1}\left[\sum_{i=1}^{n} f_{A_i}\{J(A_i)\} + (n-1)\,c\right]$$

No escribimos la forma explícita de las medidas de incertidumbre correspondientes a las leyes g-universales, siendo esas expresiones fácilmente obtenibles por analogía de (4.12.1) (4.12.2) (4.12.3)(4.12.4) (4.12.5) (4.12.6) (4.12.7).

4.13. EJEMPLOS DE INCERTIDUMBRES COMPOSITIVAS CLASE I

Sea $(\Omega,\mathcal{S},P)$ un espacio de probabilidad y sea $J : \mathcal{S} \to \mathbb{R}^+$ la medida de información de Shannon-Wiener

$$J(A) = -\log p(A) \quad \forall A \in \mathcal{S}$$

I-1. Utilizando la expresión (4.12.1) se obtienen inmediatamente

a. La entropía de Shannon

$$H_s(\vec{p}) = -\Sigma\, p_i \log p_i / \Sigma p_i$$

$$\text{con } f_A(x) = p(A).x \qquad c = 0$$

b. La entropía de Renyi

$$H_R^k(\vec{p}) = \frac{1}{k} \log [\Sigma\, p_i^{1-k} / \Sigma\, p_i] \quad k > 0$$

$$\text{con } f_A(x) = p(A)\, e^{kx}, \quad c = 0$$

Interesantes clases de incertidumbres se obtienen si se supone que sobre la estructura $(\Omega,\mathcal{S})$ está definida una utilidad, es decir una aplicación $u : \mathcal{S} \to \mathbb{R}^+$ monótona, no creciente con respecto a la relación de inclusión.

$$A \subset B \implies u(A) \geq u(B)$$

Las incertidumbres de Shannon y Renyi con utilidades son dos clases de medidas componibles, donde las leyes de composición son del tipo (4.8.2). En forma más precisa:

a'. Haciendo $f_A(x) = u(A) \cdot p(A) \cdot x$ se obtiene la medida de incertidumbre de Shannon con utilidades:

$$H_S(p,u) = - \sum_i p_i \, u_i \log p_i / [u(A) \cdot \Sigma p_i] + (n-1) \cdot c$$

b'. Haciendo $f_A(x) = u(A) \cdot p(A) \cdot \ell^{k \cdot x}$ se obtiene la entropía de Renyi con utilidades:

$$H_R^k(p,u) = \frac{1}{k} \log [\Sigma u_i \, p_i^{1-k} / \{u(A) \cdot \Sigma p_i\} + (n-1) \cdot c]$$

En particular, haciendo $c = 0$ se obtienen las medidas de incertidumbre con utilidad introducidas por S. Guiasu [31]

I.2. Dos clases de medidas de incertidumbre se obtienen escogiendo para la función $f_A(x)$, respectivamente, una de las siguientes expresiones,

(4.13.1) $f_A(x) = [u(A)]^r \cdot [p(A)]^s \cdot x^k$

(4.13.2) $f_A(x) = [u(A)]^r \cdot [p(A)]^s \cdot e^{kx}$, que generalizan las entropías de Shannon y de Renyi. Por supuesto, no todas estas funciones generan realmente medidas de incertidumbre; para que ésto se cumpla es necesario que las constantes r, s, k satisfagan en los dos casos, respectivamente, las desigualdades:

$$(4.13.3)\quad [u(A)]^r \,.\, [p(A)]^s \,.\, [-\log p(A)]^k + [u(B)]^r \,.\, [p(B)]^s \,.\, [-\log p(B)]^k$$
$$\geq [u(A\cup B)]^r \,.\, [p(A) + p(B)]^s \,.\, [-\log \{p(A) + p(B)\}]^k$$

$$(4.13.4)\quad [u(A)]^r \,.\, [p(A)]^{s-k} + [u(B)]^r\, [p(B)]^{s-k} \geq [u(A\cup B)]^r\, [p(A) + p(B)]^{s-k}$$

La desigualdad (4.13.3) siempre se cumple cuando $r \geq o$, $k \geq o$, $s \leq 1$, mientras que para que (4.13.4) sea satisfecha es suficiente que $r \geq o$, $k \geq s$. Naturalmente, es posible que (4.13.3) y (4.13.4) se satisfagan también en otros casos dependiendo de la forma particular de la función u.

Las medidas de incertidumbre correspondientes a las dos funciones (4.13.1) (4.13.2) son:

$$(4.13.5)\quad H_S(p,u,r,s,k) = [\Sigma u_i^r\, p_i^s\, (-\log p_i)^k / [u^r(A)(\Sigma p_i)^s] + (n-1).c]^{1/k}$$

$$(4.13.6)\quad H_R(p,u,r,s,k) = \frac{1}{k} \log[\Sigma u_i^r\, p_i^{s-k} / \; u^r(A)\; (\Sigma p_i)^s \; + (n-1).c]$$

Para que (4.13.6) represente realmente una medida de incertidumbre es necesario que el segundo miembro sea positivo. Esto sucede seguramente cuando, además de las condiciones $r \geq o$, $k \geq s$, se tiene $s \geq 1$.

En la clase I, son casos particulares de (4.13.5) las entropías

a. $r = 0$, $k = s = 1$ y a'. $r = s = k = 1$

Son casos particulares de (4.13.6)

b. $r = o$, $s = 1$, $k > 1$ y b'. $r = s = 1$, $k > 1$

Otros casos particulares de (4.13.5) interesantes son las entropías, todavía no estudiadas, correspondientes a los valores $r = s = o$, $k = 1$; $s = o$, $r = k = 1$; es decir, las entropías:

a''. $H_S(p,o,o,1) = - \Sigma \log p_i$

a'''. $H_S(p,u,1,o,1) = - \Sigma u_i \log p_i / u(A)$

Además de b. y b'., otros dos casos particulares interesantes de (4.13.6) son

b''. $H_R(p,o,o,k) = \frac{1}{k} \log (\Sigma p_i^k)$

b'''. $H_R(p,1,o,k) = \frac{1}{k} \log [\Sigma u_i p_i^k / u(A)]$

NOTA: En algunos de los ejemplos precedentes aparece la función "Utilidad" $u : \mathcal{S} \to \mathbb{R}^+$ a la cual sólo se le pide sea decreciente con respecto a la relación de inclusión. Puede ser que la propia u sea componible, es decir, $A \cap B = \phi \implies u(A \cup B) = \Theta[u(A), u(B)]$; pero también hay cosas interesantes donde u no cumple con esa propiedad. Lo más conocido se tiene cuando (Ω, d) es un espacio métrico y

$$u(A) := f[d^*(A)]$$

siendo $d^*(A)$ el diámetro del conjunto A y $f : \mathbb{R}^+ \to \mathbb{R}^+$ una función no creciente (por ejemplo, $f(x) = x^{-\alpha}$, $\alpha > o$).

CLASE II

Supongamos que sobre el espacio $(\Omega, \mathcal{S})$ está definida una medida de probabilidad $p : \mathcal{S} \to [o,1]$ y una función de utilidad $u : \mathcal{S} \to \mathbb{R}^+$. En esta clase de ejemplos la medida de información estará definida por

$$J(A) = -u(A) \log p(A).$$

II.1. La clase de medidas de incertidumbre correspondiente a las funciones

$$f_A(x) = [u(A)]^r [p(A)]^s \ x^k, \ r + k \geq o, \ k \geq o, \ s \leq 1,$$

está representada por:

(4.13.7) $H_S^{(1)}(p,u,r,s,k) = \{[\Sigma u_i^{r+k}.p_i^s (-\log p_i)^k]/[u^r(A)(\Sigma p_i)^s]^{1/k}$

La diferencia entre estas incertidumbres y las entropías (4.13.5) no parece muy grande: sólo varía el exponente de las utilidades u_i. Sin embargo, mientras (4.13.5) constituye una clase de entropías totalmente componibles (según la definición de B. Forte, (Ver [11]), las (4.13.7) son componibles. pero en general no totalmente.

Ellas son totalmente componibles sólo si, para cada pareja A y B de subconjuntos disjuntos, resulta

$$u(A \cup B) = - F\,[J(A),\, J(B)]/\log\,\{p(A) + p(B)\}$$

donde F(x,y) es una ley de composición adecuada que define un semigrupo. En particular, nunca (4.13.7) es totalmente componible cuando la utilidad u es componible.

Entre las entropías (4.13.7) son particularmente interesantes los casos:

a. $r = o,\ s = 1,\ k = 1$; b. $r = 1,\ s = o,\ k = 1$ y

c. $r = s = k = 1$.

Excribiremos sus expresiones suponiendo que u tenga ley de composición de tipo Inf, es decir, $u(A \cup B) = \mathrm{Inf}\,[u(A),\, u(B)]$

a. $H_S^{(1)}(p,u,o,1,1) = - \Sigma\, p_i\, u_i \log p_i \,/\, \Sigma\, p_i + (n-1).\, c$

b. $H_S^{(1)}(p,u,1,o,1) = -\Sigma\, u_i^2 \log p_i \,/\, \text{Inf}\,(u_i) + (n-1).c$

c. $H_S^{(1)}(p,u,1,1,1) = -\Sigma\, p_i\, u_i^2 \log p_i \,/[\text{Inf}\,(u_i).\,\Sigma\, p_i] + (n-1).c$

II.2. Utilizando como generadora la función

$$f_A(x) = [u(A)]^r\,[p(A)]^s\; e^{kx}$$

se obtiene la siguiente clase de medidas de incertidumbre

$$(4.13.8)\quad H_R^{(1)}(p,u,r,s,k) = \frac{1}{k}\log\{\Sigma u_i^r\, p_i^{s-ku_i}/[u^r(A)\;(\Sigma p_i)^s] + (n-1).c\}$$

La condición de coherencia b. del teorema 4.8.1 para esta función toma la forma

$$[u(A)]^r\,[p(A)]^{s-ku(A)} + [u(B)]^r\,[P(B)]^{s-ku(B)}$$

$$\geq [u(A\cup B)]^r\,[p(A)+p(B)]^{s-k\,u(A\cup B)}$$

y es difícil averiguar cuando se cumple, dependiendo ésto de la forma de la función utilidad. Sin embargo, cuando $u(\phi) = u < +\infty$ la condición se encuentra satisfecha si $r \geq o$, $k < o$, $s-k.u < 1$.

Por supuesto, también la entropía (4.13.8) es componible pero no totalmente.

Al final cabe notar que (4.13.8) es sustancialmente diferente de su homóloga (4.13.6), lo contrario de lo que pasa al comparar (4.13.7) con (4.13.5)

Muchos otros ejemplos de este tipo pueden proporcionarse bajo otras selecciones de la medida de información $J : \mathcal{S} \to \mathbb{R}^+$ y de la función $f : \mathcal{S} \times \mathbb{R}^+ \to \mathbb{R}^+$.

CLASE III

En esta clase serán particularizadas tres familias de medidas de incertidumbre cuya ley de composición es la ley dada en el Teorema 4.6.1. Naturalmente, estas clases corresponden a selecciones particulares de los valores e y k y de la función h(z).

III-1. $e = k = 0 \Longrightarrow H(\Pi_A) = \operatorname{Inf} \{J(A_i) : A_i \in \Pi_A\}$

III-2. $e = 0,\ k = +\infty \Longrightarrow H(\Pi_A) = \operatorname{Sup} \{J(A_i) : A_i \quad \Pi_A\}$

III-3. $e = o,\ k < +\infty,\ h(z) = k$

$$\Longrightarrow \quad H(\Pi_A) = \begin{cases} \operatorname{Sup} \{J(A_i) : A_i \in \Pi_A\} & \text{si } \operatorname{Sup} J(A_i) \leq k \\ \operatorname{Inf} \{J(A_i) : A_i \in \Pi_A\} & \text{si } \operatorname{Inf} J(A_i) \geq k \\ k & \text{si } \operatorname{Inf} J(A_i) \leq k \leq \operatorname{Sup} J(A_i) \end{cases}$$

Es conveniente observar que en esta expresión de $H(\Pi_A)$ cuando los eventos $A_i \in \Pi_A$ son "homogéneos", o sea, cuando todos tienen información bastante baja ($\leq$ k) o bastante alta ($\geq$ k), entonces la medida de incertidumbre depende explícitamente de los eventos A_i; por el contrario, si no hay "homogeneidad", $H(\Pi_A)$ tiene un valor prefijado (=k). Quizás esto pueda ser útil en casos donde situaciones demasiado confusas no merecen tomarse en consideración.

B I B L I O G R A F I A

[1] C. Bértoluzza. "Sulla informazione Condizionale". Statistica. No.2. ANNO XXVIII. Aprile-Giugno, 1968.

[2] G. Maschio. "Note di Teoría dell' informazione". Quaderni dei gruppi di ricerca del C.N.R, 1972.

[3] J. Kampe de Feriet et B. Forte. "Information et probabilite". C.R.A.S. París. 265 (1967). p. 110

[4] Benvenutti, Forte, Kampé de Feriet. "Forme generale de l'operation de composition". C.R.A.S. París 269, Serie A, 1969.

[5] F. Barbaini - C. Bertoluzza, "Sul concetto D'indipendenza in Teoría dell' informazione". Report interno del Dipartimento di Informática e Sistemistica dell' Universita di Pavia. (RIDIS 14-84).

[6] C. Bertoluzza - Angela Boscaini. "Un Sistema di Equazioni Funzionali Connesso con una generalizzazione della nozione dindipendenza in Teoría dell' informazione". Ist. Lombardo Rend. Sci. Serie A, Vol III (1977).

[7] J. Kampé de Feriet et B. Forte. "Information et probabilite". C.R.A.S. París, 265 (1967) p. 350.

[8] F. Barbaini - C. Bertoluzza. "Leggi di Composizione dell' incertezza universalmente compatibile con la legge d'indipendenza $x \circ y = \text{Sup}\,(x,y)$". Ist. Lombardo Rend Sci. Vol. 117 (1983).

[9] P.S. Mostert - A.L. Shields. "On the Structure of semigroups on a compact manifold with boundary", Annals of Math, 65 (1957), pp. 117-143.

[10] B. Forte - P. Benvenuti. "Su una classe di misure di informazione regolari a traccia Shannoniana". AΠ_i Sem. Mat. Univ. Modena XVII (1969). pp. 99-108.

[11] B. Forte. "Measures of Information: The General Axiomatic Theory". R.I.R.O. 3° année, No. R-2. 1969 pp. 63-90.

[12] T. Van der Pyl. "Information d' ordre α et de Type B: axiomatique et propriétés". These du $3^{ème}$ cycle. París, 1977.

[13] Picard, C. F. "Propriétés d' additivité de l'information de Type β". C.R.A.S. París. Tome 282 A. Avvil 1976. pp. 915-918.

[14] Aczél. "Determination of All Additive Guasiarithmetic mean Codeword Lengthes". Z. Wahrscheinlichkeit stheorie und verw. Gebiete 29, 351-360.

[15] B. Forte - C. T. N_G . "Entropies with the Branching Property". Ann. Mat. Pura e Appl., 1974. pp. 354-373.

[16] B. Forte. N. Pintacuda. "Sull' informazione associata alle esperienze incomplete". Annali di Matemática pura ed applicata. (IV), Vol. LXXX, pp. 215-234.

[17] Carla Poggi. "Entropie G - diramative ed Entropie G-locali in spazi o' informazione con legge di indipendenza di tipo generale". Statística, anno XLIII, n.2, 1983.

[18] María Divari - Mirian Pandolfi. "Su una legge compositiva dell' informazione". Rend. Mat. Univ. Roma, serie 6 (3) (1970) pp. 805-817.

[19] C. Bertoluzza - I. Bonzani. "Su una generalizzazione della nozione di diramativitá in teoria dell' informazione".

[20] A. Lozzi. "Misure d' incertezza compositive compatibile con l'assioma d' indipendenza generalizzato" (Titolo provisorio) Tesi di laurea (1985).

[21] J. Kampé de Fériet. "Indépendance des evenements en calcul des probabilitée et en théorie de la' information". Journées Information, Questionnaíres et Reconnaissance (Bonas 1976). Publ. No. 2 du Groupe de Recherche No. 22. (Structure de la' information) du C.N.R.S.

[22] P. Benvenuti. "Sulle soluzioni di un sistema di equazioni funzionali della teoría della informazione". Rend. di Mat. (6) 2. 1969. pp. 99-110.

[23] P. Benvenuti. "Sulle misure di Informazione compositive con Traccia compositiva universale". Rend. di Mat, 1969 pp. 482-506.

[24] C. Bertoluzza - F. Barbaini. "Caratterizzazione delle leggi di composizione dell' Incertezza (G,F) Universali". RIDIS 9-84.

[25] J. Aczel. "Lecture on Functional Equations and Their Applications". Academic Press, New-York. 1966. p. 312.

[26] Z. Daroczy. "Uber eine Funktionalgleichugssystem der Informations Theorie". Aequationes Math, 2, 144-149.

[27] R.S. Ingarden- K. Urbanik. "Information Without probability". Coll. Math 9, pp. 131-150 (1962).

[28] A.J. Khinchin. "Mathematical Foundations of Information Theory". Dover, New York, 1957.

[29] D.K. Faddeev. "On The Concept of Entropy of a Finite Probabilistic Scheme". (Russian) Uspehi Mat Nauk (N.S). 11. No. 1 (67). 227-231.

[30] A. Renyi. "Probability Theory". North-Holland, 1970.

[31] GUIASU S. "Weighted entropy", Reports on Mathematical Physics 2. 1971 p. 165-179.

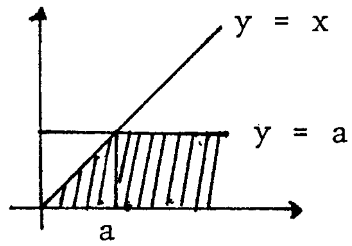
[32] C. Bertoluzza - M. Schneider. "Informations totalement Composables".

A P E N D I C E S

APENDICE A

Demostración Teorema 2.5.2

Lema A-1



a. Si Λ es acotado entonces:

(A.1) $F(x,y) = 0 \quad \forall (x,y) \in \{[o,a]^2\} \cap L.$

$F(x,y) = F(x,y) \ \forall (x,y) \in \{[a,+\infty) \times [o,a]\} \cap L$

b. Si Λ no es acotado, entonces $F(x,y) = 0 \ \forall \ (x,y) \in (\mathbb{R}^+)^2 \cap L$

Demostración

a. Como $F(x,y) = F(x,y) * 0$ y $F(a,a) = 0$ tenemos

$$\begin{aligned} F(x,y) * 0 &= F(x,y) * F(a,a) \\ &= F(x * a, y * a) \end{aligned}$$

Luego $F(x,y) = F(x * a, y * a)$

Como $a = \mathrm{Sup}\ \{x \in \Lambda\}$ entonces $r * a = \mathrm{Sup}\ (r,a)$ para cada $r \in \mathbb{R}^+$. Así, para cada $(x,y) \in [o,a]^2 \cap L$ se cumple

$$F(x,y) = F(a,a) = 0$$

Por otro lado, si $y \leq a \leq x$ tenemos

$$\begin{aligned} F(x,y) &= F(x * a, y * a) \\ &= F(x,a) \end{aligned}$$

En definitiva

$$F(x,y) = 0 \qquad \text{si } (x,y) \in [o,a]^2 \cap L$$

$$F(x,y) = F(x,a) \qquad \text{si } (x,y) \in \{[a,+\infty) \times [o,a]\} \cap L$$

b. Si Λ no es acotado, para cada pareja $(x,y) \in L$ escogemos

$\lambda \in \Lambda$ Tal que $\lambda \geq x > y$

Como $F(x,y) = F(x,y) * 0$ y $F(\lambda,\lambda) = 0$

entonces:

$$\begin{aligned} F(x,y) &= F(x,y) * F(\lambda,\lambda) \\ &= F(x * \lambda,\ y * \lambda) \\ &= F(\lambda,\lambda) = 0. \end{aligned}$$

Luego en virtud de (2.5 - 6.c) y del Lema A-1 para determinar la función $F(x,y)$ podemos limitarnos a estudiarla sobre el subconjunto de $(\mathbf{R}^+)^2$ para el cual $a \leq y < x$.

Lema A.2

Si $a \leq y < c$, $F(x,y)$ es función sólo de la diferencia $g(x) - g(y)$, es decir

$$\text{(A-2)} \qquad F(x,y) = f\,[g(x) - g(y)]$$

donde $g(x)$ es la función que particulariza la operación $*$ en

$[a, +\infty)$.

Demostración

Si $a \leq y < x$, entonces $x * y = g^{-1}[g(x) + g(y)]$. Además, para todo $z > a$ tenemos

$$\begin{aligned} F(x,y) &= F(x,y) * F(z,z) \\ &= F(x * z, y * z) \\ &= F\{g^{-1}[g(x) + g(z)], g^{-1}[g(y) + g(z)]\} \end{aligned}$$

Si hacemos $g(x) = \xi$, $g(y) = \eta$, $g(z) = \zeta$ obtenemos;

$$F\{g^{-1}(\xi), g^{-1}(\eta)\} = F\{g^{-1}(\xi+\zeta), g^{-1}(\eta + \zeta)\}$$

Haciendo

$G(\xi,\eta) = F\{g^{-1}(\xi), g^{-1}(\eta)\}$ para cualquier valor positivo de ξ, η, ζ se obtiene

(A.3) $G(\xi + \zeta, \eta + \zeta) = G(\xi, \eta)$ de lo cual se sigue que:

(A.4) $\quad G(\xi,\eta) = f(\xi-\eta)$

o sea

$$F(x,y) = f[g(x) - g(y)]$$

En efecto, haciendo

$$\tilde{F}[\xi-\eta, \eta] = G[\xi,\eta]$$

de (A.3) se obtiene

$$\tilde{F}\,[\xi-\eta,\eta+\zeta] = F\,[\xi-\eta,\eta]$$

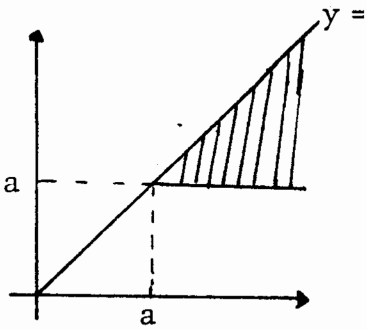
y como ζ es arbitrario, la función $\tilde{F}(\nu,s)$ no depende de la segunda variable. Luego

$$G(\xi,\eta) = \tilde{F}\,[\xi-\eta,\eta] = f(\xi-\eta)$$

Además, cada función del tipo $G(\xi,\eta) = f(\zeta-\eta)$ satisface (A.3). Luego (A.4) representa la solución general de (A.3). $\square$

En el siguiente párrafo caracterizaremos la función $f(u)$ completando así el estudio de $F(x,y)$.

Lema A.3



$f(u) : (o, +\infty) \to \mathbf{R}^+$ goza de las siguientes propiedades:

(A.5) $\quad u < v \implies f(u) \leq f(v)$

(A.6) $\quad f(u+v) = f(u) * f(v)$

Demostración

$x_1 < x_2 \implies F(x_1,y) \leq F(x_2,y) \quad \forall\, x_1,x_2 \geq y$ entonces

$x_1 < x_2 \implies f\,[g(x_1) - g(y)] \leq f\,[g(x_2) - g(y)]$

Además, haciendo $y = a$ se obtiene

$$x_1 < x_2 \implies f[g(x_1)] \leq f[g(x_2)]$$

Por otro lado, como g es estrictamente monótona tenemos que:

$$x_1 < x_2 \iff g(x_1) < g(x_2) ,$$

entonces

$$\text{(A.7)} \qquad g(x_1) < g(x_2) \implies f[g(x_1)] \leq f[g(x_2)]$$

Al final, variando x en $[a, +\infty)$ $g(x)$ varía en $[o, +\infty)$ (y viceversa). Luego, haciendo $u = g(x_1)$, $v = g(x_2)$ de (A.7) se reconoce que (A.5) se cumple para cada $u,v \in \mathbb{R}^+$.

Para demostrar (A.6) observemos que:

$$a \leq y_1 < x_1 \text{ y } a \leq y_2 < x_2 \quad , 4.6.b)$$

$$\implies F(x_1 * x_2, y_1 * x_2) = F(x_1,y_1) * F(x_2,y_2)$$

entonces por (A.2) tenemos que:

$$f[g(x_1 * x_2) - g(y_1 * y_2)] = f[g(x_1) - g(y_1)] * f[g(x_2) - g(y_2)]$$

Por otra parte

$$f\,[g(x_1 * x_2) - g(y_1 * y_2)] = f\,[g{\circ}g^{-1}(g(x_1) + g(x_2) - g{\circ}g^{-1}(g(y_1) + g(y_2)]$$

$$= f\,[g(x_1) + g(x_2) - g(y_1) + g(y_2)]$$

y luego

$$f\,[g(x_1) - g(y_1) + g(x_2) = g(y_2)] = f\,[g(x_1) - g(y_1)] * f\,[g(x_2) - g(y_2)]$$

Haciendo $u = g(x_1) - g(y_1)$, $v = g(x_2) - g(y_2)$ se tiene que

$$f(u + v) = f(u) * f(v)$$

Por las características de $g(x)$ siempre es cierto que al variar x_1, y_1, x_2, y_2 en $a \leq y_i \leq x_i$: $i = 1,2$, u y v variarán en $\mathbf{R}^+$. Así, la demostración del lema queda completa. □

Sean ahora: $f_o = \lim\limits_{u \to o^+} f(u) \quad (f_o \geq o)$

$$\beta = \operatorname{Inf} \{\lambda \in \Lambda : \lambda > f_o\} \quad (\beta \geq f_o)$$

$$p = \operatorname{Inf} \{u > o : f(u) \geq \beta\} \quad (o \leq p \leq +\infty)$$

Lema A.4

En el intervalo abierto $(p, +\infty)$ se tiene que $f(u) = \beta$. Luego en este intervalo $f(u)$ es constante.

Demostración

De acuerdo a las condiciones dadas, para cualquier $u > p$ se tiene $f(u) \geq \beta$.

Si por otra parte para cualquier $u > p$ fuera $f(u) > \beta$ se podría probar que

$$\exists\ \lambda \in \Lambda \quad \text{tal que } f(u) > \lambda > f_o$$

En efecto:

Si $\beta > f_o$, ese número λ coincide con β.

Si $\beta = f_o$, entonces $\beta = f_o$ es un punto de acumulación para Λ, luego si $f(u) > f_o$ existe un número λ en el intervalo $(f_o, (f(u))$.

Haciendo $\varepsilon = \lambda - f_o > o$, buscamos $n \in \mathbb{N}$ bastante grande de manera que sea

$$f(u|n) - f_o < \varepsilon \quad \text{ó} \quad f(u|n) < \lambda .$$

Esto siempre es posible ya que $f_o = \lim_{u \to o} f(u)$ y $u|n \overset{n}{\to} o$

Entonces se tiene

$$\begin{aligned} f(u) = f(nu|n) &= f(\tfrac{u}{n} + \tfrac{u}{n} + \ldots + \tfrac{u}{n}) \\ &= f(\tfrac{u}{n}) * f(\tfrac{u}{n}) * \ldots * f(\tfrac{u}{n}) \qquad (A.6) \\ &\leq \lambda * \lambda * \ldots * \lambda = \lambda \end{aligned}$$

Lo cual es una contradicción ya que partimos de que $f(u) > \lambda$. Por lo tanto $f(u)$ no puede ser estricatamente mayor que β y

$$f(u) = \beta$$

$\square$

Lema A.5

Si $p > o$ y $f_0 < \beta$ entonces

$$0 < u < p \implies f(u) = h^{-1}\,[u.h(\beta)/p]$$

donde $h(x)$ es la función que define la operación $*$ en (f_0, β)

Demostración

$\forall$ u,v con $u > o$, $v > o$, $u + v < p$ se tiene por (A.6)

$$\begin{aligned} f(u + v) &= f(u) * f(v) \\ &= h^{-1}\,\{h(f(u)) + h(f(v)\} \quad \text{ó} \end{aligned}$$

$$h_\circ\, f(u + v) = h_\circ f(u) + h_\circ\, f(v)$$

Entonces la función $h_\circ f(u)$ satisface la ecuación de Cauchy en el dominio $u + v < p$ $(u, v > o)$; luego, como élla es monótona en el intervalo (o,p) se tiene

$$h_\circ\, f(u) = k\,u \quad \text{ó} \quad f(u) = h^{-1}\,(k\,u) \text{ con } k \text{ constante}$$

En base al lema A.4 y a la monotonía

$$\lim_{u \to p^-} h^{-1}\,(k\,u) \leq \beta \;\rightarrow\; K \leq h(\beta)/\rho$$

Además, si $k < h(\beta)/p$ ó $p < h(\beta)/k$ escogiendo u y v menores que p y tales que $p < u + v < h(\beta)/k$ se tendría:

$$f(u) * f(v) = h^{-1}\{h_o f(u) + h_o f(v)\}$$

$$= h^{-1}(k\,u + k\,v)$$

$$= h^{-1}(k\,(u+v)) < \beta$$

Mientras que de (A.6) tenemos

$$f(u) * f(v) = f(u + v) = \beta$$

Lo cual es una contradicción, por lo tanto

$$K = h(\beta)/p$$

□

Observaciones

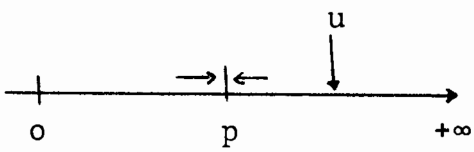
(A.1) $f_o \in \Lambda$: en caso contrario, f_o estaría en un intervalo $(\alpha,\beta) \in \Lambda$ con α, $\beta \in \Lambda$ y en consecuencia

$$\lim_{u\to o^+} f(v)=\lim_{u\to o} h^{-1}[u\,h(\beta)/p] = h^{-1}[\lim_{k\to o} u\,h(\beta)/p]$$

$$= \alpha \neq f_o \quad \text{(absurdo)}$$

(A.2) En base a los lemas A.4 y A.5 y a (A.5) se reconocen inmediatamente que $f(p) = \beta$.

Luego la función f(u) está determinada para cualquier $u > o$.



En efecto:

en $(p, +\infty)$ $f(u) = \beta$ y

en $\rho > o$ $(f_o < \beta)$, $o < u < p \Longrightarrow f(u) = h^{-1}[u\, h(\beta)/p]$

Ademas:

$$\lim_{u \to p^+} f(u) = \beta \quad \text{porque } f(u) = \beta \text{ si } u > p$$

$$\lim_{u \to p^-} f(p) = \lim_{u \to p^-} h^{-1}[u\, h(\beta)/p] = h^{-1}[\lim_{u \to p^-} u\, h(\beta)/p]$$

$$= h^{-1}[h(\beta)] = \beta$$

Como f es continua, entonces

$$f(p) = \beta$$

Los lemas A.1 hasta A.5 prueban el teorema de representación que determina cualquier información relativa del tipo

$$H(\Pi|\Pi') = F[H(\Pi \wedge \Pi'), H(\Pi')]$$

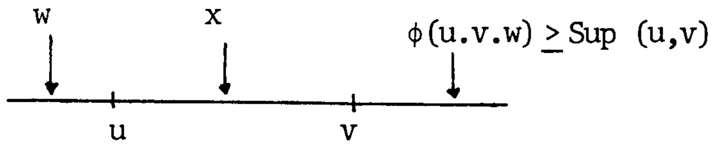
□

APENDICE B

B.1. Demostración Teorema 3.3.1

Lema B.1

I. (B.1.1) $u * v = \mathrm{Sup}\ (u,v) \rightarrow \phi\ (u,v,w) = \mathrm{Sup}(u,v)\ \forall\ \omega : (u,v,w) \in \Gamma_3$



Demostración

Supongamos $u \leq v$

Por hipótesis tenemos que $u * v = \mathrm{Sup}\ (u,v)$; esto implica que existe un elementos Λ-idempendiente que pertenece (por la def (1.7.1) a $[u,v]$ ($\exists\ \ x \in \Lambda \cap [u\ ,\ v]$ de tal manera que

$w \leq\ u \leq\ x$, ya que $w \leq \mathrm{Inf}\ (u,v) = u \geq x$ y por lo tanto, como $x \in \Lambda\ \ \phi(u,v,w) * x = \mathrm{Sup}(\phi(u,v,w),x) = \phi(u,v,w)$

Por otra parte, utilizando (3.3.11) tenemos que:

$$\begin{aligned} \phi(u,v,w) * X &= \phi(u * x,\ v * x,\ w * x) \\ &= \phi(x,\ v,\ x) \\ &= v \qquad \text{(por (3.3.10.f)} \end{aligned}$$

Luego $\phi(u,v,w) = v = \mathrm{Sup}\ (u,v)$

□

Lema B.2

Si $\exists i$ tal que $(u,v) \in (\alpha_i,\beta_i)^2$ entonces se tiene:

$$\phi(u,v,w) = \phi[u,v, \text{Sup}\ (w,\alpha_i)] \quad \forall\ w : (u,v,w) \in \Gamma_3$$

Demostración

En efecto, como $\alpha_i \in \Lambda$ se cumple

$$\alpha_i * x = x * \alpha_i = \text{Sup}\ (x,\alpha_i) \quad (\forall\ x \in \mathbb{R}^+)$$

Por (3.3.10) $\phi(u,v,w) \geq \text{Sup}\ (u,v) = v \geq \alpha_i$

Luego $\phi(u,v,w) * \alpha_i = \text{Sup}\ (\phi(u,v,w),\alpha_i) = \phi(u,v,w)$

Por otra parte, siendo $n \geq \alpha_i$, $v \geq \alpha_i$

$$\begin{aligned}\phi(u,v,w) * \alpha_i &= \phi(u * \alpha_i,\ v * \alpha_i,\ w * \alpha_i) \\ &= \phi(u,v, \text{Sup}\ (w,\alpha_i))\end{aligned}$$

y así queda demostrado el Lema B.2.

□

Basándonos en los lemas B.1 y B.2 la función $\phi(u,v,w)$ está determinada completamente cuando lo están sus restricciones ϕ_i en todos los subconjuntos.

$$\{(\alpha_i,\ \beta_i)^2 \times [\alpha_i,\beta_i)\} \cap \Gamma_3 .$$

En lo que sigue indicaremos con (α,β) el intervalo genérico (α_i,β_i), con g la función relativa g_i y haremos

$$C = \{(\alpha,\beta)^2 \times [\alpha,\beta)\} \cap \Gamma_3$$

Lema B.4

$\forall$ $(u,v,w) \in$ C se tiene:

(B.1.2) $\phi(u,v,w) = \phi[g^{-1}\, g(u)-g(w)),\ g^{-1}(g(v)-g(w),\alpha] * w$

Demostración:

$\phi[g^{-1}(g(u)-g(w),\ g^{-1}(g(v)-g(w),\alpha] * w$

$= \phi[w * g^{-1}\ (g(u)-g(w),\ w * g^{-1}(g(v)-g(w))\ w * \alpha]$ (por (3.3.11)

$= \phi\{g^{-1}[g(w) + g_\circ g^{-1}(g(u)-g(w))],\ g^{-1}[g(w) + g_\circ g^{-1}(g(v)-g(w)],\ \mathrm{Sup}\ (w,\alpha)$

$= \phi\ [g^{-1}(g(w) + g(u)-g(w)),\ g^{-1}(g(w) + g(v)-g(w)),w]$ (ya que $w \geq \alpha$)

$= \phi[g^{-1}\ (g(u)),g^{-1}(g(v)),w]$

$=\phi(u,v,w)$ l.q.q.d.

□

Basándonos en el Lema B.4. diremos que la función ϕ restringida al conjunto C estará determinada completamente cuando lo esté la función

$$Z(r,s) = \phi(r,s,\alpha) \text{ real, positiva, definida para } (r,s)\in(\alpha,\beta)^2 \text{ y con valores en } [\alpha,\beta]$$

B.2. La Función Z(r,s)

Las propiedades (3.3.10b) hasta (3.3.10e) (3.3.11) y (B.1.1) de la función ϕ, se traducen para la función $Z(r,s)$ de la siguiente forma

(B.2.1) $Z(r,s) \geq \mathrm{Sup}\,(r,s)$

(B.2.2) $Z(r,s) = Z(s,r)$

(B.2.3) $r' < r'' \rightarrow Z(r',s) \leq Z(r'',s)$

(B.2.4) $Z\,[Z(r,s),t] = Z\,[r,Z(s,t)]$

(B.2.5) $Z(r,\alpha) = r$

(B.2.6) $Z(r,\beta) = \beta$

Este sistema de ecuaciones funcionales es el mismo con el cual en [4] se determinó la forma general de la operación binaria $*$ (ley de independencia en un espacio de información). Todas sus soluciones pueden entonces obtenerse tomando una sucesión de intervalos $\{(\xi_j,\zeta_i)\}$ contenidos en $[\alpha,\beta]$ y disjuntos entre sí y sobre cada uno de éllos una función f_j, real, continua, estrictamente creciente e infinitésima en ξ_j y haciendo

$$(B.2.7)\quad Z(r,s) = \begin{cases} \bar{f}_j\,[f_j(r) + f_j(s)] & \text{si } (r,s) \in (\xi_j,\zeta_j)^2 \\ \mathrm{Sup}\,(r,s) & \text{en otro caso} \end{cases}$$

donde $\Lambda_Z := [\alpha,\beta] - \bigcup_j (\xi_j,\, \zeta_j)$

$= \{r \in [\alpha,\beta] \;/\; Z\,(r,s) = \mathrm{Sup}\,(r,s),\ \forall\, s \in [\alpha,\beta]$

Lema B.5

El conjunto Λ_z es de la forma:

$$\Lambda_z = [\alpha,\beta] \quad \text{o} \quad \Lambda_z = \{\alpha\} \cup \{\beta\}$$

En el primer caso la sucesión $\{(\xi,\zeta)\}$ es vacía, mientras que en el segundo está constituida únicamente por (α,β)

Demostración

Tomando en cuenta (B.1.2) y haciendo $g_x = g(x)$, la asociatividad (3.3.10e) se puede escribir de la siguiente manera:

$$\bar{g}\,\{g\circ Z\,[g^{-1}(g_\circ - g_\circ),\ g^{-1}(g\circ\bar{g}\,\{g\circ Z\ \ g^{-1}(g_p-g_r),\ g^{-1}(g_q-g_r)] + g_r\} - g_w)] + g_w\}$$

B.2.8)

$$= \bar{g}\,\{g\circ Z\,[g^{-1}(g\circ\bar{g}\,\{g\circ Z\,[g^{-1}(g_u-g_w),\ g^{-1}(g_q-g_w)] + g_w\} - g_r),\ g^{-1}(g_p-g_r)] + g_r\}$$

Para lo cual basta utilizar en su preciso momento:

1) $\phi(u,v,w) = \phi[g^{-1}(g_u-g_w),\ g^{-1}(g_r-g_w),\alpha] * w$

2) $\phi(r,s,\alpha) = Z(r,s)$

3) $x * y = \bar{g}\,[g(x) + g(y)]$

En efecto:

$$\phi[u,\ \phi(p,q,r),\ w] = \phi[g^{-1}(g_u-g_w),\ g^{-1}(g_{\phi(p,q,r)} - g_w),\alpha] * w$$

$$= Z\,[g^{-1}(g_u-g_w),\ g^{-1}(g_{\phi(p,q,r)} - g_w)] * w$$

$$= \bar{g}\,\{g\circ Z\,[g^{-1}(g_u-g_w),\ g^{-1}(g_{\phi(p,q,r)}-g_w)] + gw\}$$

Sustituyendo

$$\phi(p,q,r) = \bar{g}\{g_\circ Z[g^{-1}(g_p,g_r),\ g^{-1}(g_q-g_r)] + g_r\}$$

Obtenemos el lado izquierdo de la igualdad y en forma similar obtenemos el lado derecho.

Demostremos que si $\exists\ u \in (\alpha,\beta)$ y $u \in \Lambda_z$ entonces

$\Lambda_z = [\alpha,\beta]$

I. Probemos primero que en la hipótesis hecha sobre u se tiene $Z(r,r) = r\ \forall\ r \in [u,\beta]$

En efecto:

Sea $\gamma = \mathrm{Inf}\ (r : Z(r,r) = \beta)$

$y \in (u,\gamma) \implies Z(y,y) < \beta$

$y \in (u,\gamma) \implies \bar{g}\{g_\ell + g_m\} = g^{-1}\{g_\ell + g_m\}\quad \forall\ \ell,m \in (u,\gamma)$

si en (B.2.8) hacemos $r = \alpha,\quad g_r = 0,$

$q = u\ ,$

$u = p = y > u,\ y < \gamma$

Obtenemos:

$$\bar{g}\ \{g_\circ Z[g^{-1}(g_y-g_w),\ g^{-1}(g_\circ\bar{g}\{g_\circ Z[g_\circ^{-1}g_y, g_\circ^{-1}g_u]\} - g_w)] + g_w\} =$$

$$= \bar{g}\ \{g \circ Z\ [g^{-1}(g \circ \bar{g}\ \{g \circ Z\ [g^{-1}(g_y - g_w),\ g^{-1}(g_u - g_w)] + g_w\}),\ g^{-1} g_y\ ]\}$$

o sea

$$\bar{g}\ \{g_{\circ} Z\ [g^{-1}(g_y - g_w),\ g^{-1}(g \circ \bar{g}\ \{g \circ Z\ [y,u]\} - g_w)]\ + g_w\}$$

$$= \bar{g}\ \{g_{\circ} Z\ [g^{-1}(g_{\circ}\bar{g})\{\ g_{\circ} Z\ [g^{-1}(g_y - g_w),\ g^{-1}(g_u - g_w)]\ + g_w\}),\ y]\}$$

Como $u < y$, tenemos $Z\ [y,u] < Z\ [y,y] < \beta$ (Monotonía) y por lo tanto podemos escribir la relación precedente bajo la forma

$$\bar{g}\ \{g_{\circ} Z\ [g^{-1}(g_y - g_w),\ g^{-1}(g_{\circ} Z\ [y,u]\ - g_w)]\ + g_w\}\ =$$

(B.2.9)

$$= \bar{g}\ \{g_{\circ} Z\ [g^{-1}\{g_{\circ} Z\ [g^{-1}(g_y - g_w),\ g^{-1}(g_u - g_w)]\ + g_w\},\ y]\}$$

Además, si escogemos w bastante pequeño ($w \leq \bar{w}\ y$) de tal manera que:

$$\begin{cases} g_y - g_w \geq g_u & (g_w \leq g_y - g_u \leq g_y) \\ g_w \leq g_u & \end{cases}$$

entonces se tiene que:

$$\bar{g}\ \{g_{\circ} Z\ [g^{-1}(g_y - g_w),\ g^{-1}\ \{g(\mathrm{Sup}\ (y,u)\ - g_w\}]\ + g_w\}$$
$$= \bar{g}\ \{g_{\circ} Z\ [g^{-1}\ (g_y - g_w),\ g^{-1}\ (g_y - g_w)]\ + g_w\}$$

y también

$$\bar{g}\ \{g_\circ Z\ [g^{-1}\ \{g\ [\text{Sup}\ (g^{-1}(g_y - g_w),\ g^{-1}(g_u - g_w))] + g_w\},\ \ y]\}$$

$$= \bar{g}\ \{g_\circ Z\ [g^{-1}\ \{g\ [g^{-1}\ (g_y - g_w)] + g_w\},\ \ y]\}$$

$$= \bar{g}\ \{g_\circ Z\ [g^{-1}\ \{g_y - g_w + g_w\},\ \ y]\}$$

Entonces la relación (B.2.9) toma la forma:

(B.2.10) $$\bar{g}\ \{g_\circ Z\ [g^{-1}(g_y - g_w),\ g^{-1}(g_y - g_w)] + g_w\} = \bar{g}\{g_o Z\ (y,y)\}$$

Siendo el segundo miembro de (B.2.10) estrictamente menor que β, entonces $\bar{g} = g^{-1}$. Como g^{-1} es estrictamente creciente, de (B.2.10) se obtiene

(B.2.11) $$g_\circ Z\ [g^{-1}\ (g_y - g_w),\ g^{-1}(g_y - g_w)] + g_w = g_\circ Z\ (y,y) \quad \forall\ w \leq \bar{w}_y$$

Por la continuidad de las funciones g y Z la fórmula (B.2.11) debe cumplirse también si $y = \gamma$ y recordando que $Z\ [\gamma,\gamma] = \beta$ se tiene:

$$g_\circ Z\ [g^{-1}(g_\gamma - g_w),\ g^{-1}\ (g_\gamma - g_w)] + g_w = g_\beta$$

o sea

$$g_o Z\ [g^{-1}(g_\gamma - g_w),\ g^{-1}(g_\gamma - g_w)] = g_\beta - g_w - g_\gamma + g_\gamma$$

$\forall$ w tal que Sup $\{g_u, g_\gamma - g_w\} \leq g_\gamma - g_w \leq g_\gamma$

Luego haciendo $g^{-1}(g_\gamma - g_w) = r$ ó $g_\gamma - g_w = g_r$

Tenemos:

(B.2.12) $\quad g_\circ Z(r,r) - g_r = g_\beta - g_\gamma \qquad \forall\, r : \nu_1 \leq r \leq \gamma$

donde $\nu_1 = g^{-1}\{\text{Sup}\,[g_u, (g_\gamma - g_u)]\}$.

Si $\nu_1 = u$, (B.2.12) determina únicamente la función Z(r,r) sobre todo el intervalo $[u,\gamma]$.

Si $\nu_1 > u$, se puede repetir el razonamiento anterior sustituyendo y por ν_1 y haciendo $r = g^{-1}(g_{\nu_1}, g_w)$ o $g_r = g_{\nu_1} - g_w$ y $\nu_2 = g^{-1}\{\text{Sup}\,[g_u, (g_{\nu_1} - g_w)]\}$; entonces obtenemos:

$$g_\circ Z\,[g^{-1}(g_{\nu_1} - g_w),\ g^{-1}(g_y - g_w)] + g_w = g_\circ Z(\nu_1, \nu_1)$$

$$g_\circ Z(r,r) = g_\circ Z\,(\nu_1, \nu_1) - g_w + g_{\nu_1} - g\,\nu_1$$

$$g_\circ Z(r,r) - (g_{\nu_1} - g_w) = g_\circ Z(\nu_1, \nu_1) - g_{\nu_1}$$

$$g_\circ Z(r,r) - g_r = g_\circ Z(\nu_1, \nu_1) - g_{\nu_1}, \text{ o de (B.2.12)}$$

$$g_\circ Z(r,r) - g_r = g_\beta - g_{\nu_1} + g_\gamma - g_\gamma \qquad \forall\, r : \nu_2 \leq r \leq \nu_1$$

$$g_\circ Z(r,r) - g_r = g_\beta - g_\gamma - (g_r + g_w) + (g_r + g_w) \qquad \forall\, r : \nu_2 \leq r \leq \nu_1$$

$$g_\circ Z(r,r) - g_r = g_\beta - g_r \qquad \forall\, r : \nu_2 \leq r \leq \nu_1.$$

Realizando el procedimiento iterativamente, se determina una sucesión decreciente de número $\nu_1, \nu_2, \ldots$ donde se verifica fácilmente que:

$$\nu_k = g^{-1} \{ \mathrm{Sup} \, [g_u, (g_\gamma - K g_u)] \}.$$

Siendo por hipótesis $u > \alpha$ ($g_u > 0$), para k bastante grande se tiene que $\nu_k = g^{-1}(g_u) = u$.

Se puede concluir que:

(B.2.13) $\quad g_\circ Z(r,r) - g_r = g_\beta - g_\gamma \qquad \forall \; r : u \leq r \leq \beta$

En particular si $r = u$ tenemos

$$g_\circ Z\,(u,u) - g_u = g_\beta - g_\gamma$$

y como por hipótesis $Z(u,u) = u$ resulta que $g_\beta - g_\gamma = 0$. Sustituyendo ésto en (B.2.13), tenemos que:

$$g_\circ Z(r,r) - g_r = 0 \implies g_o Z(r,r) = g_r$$

Como g es monótona entonces:

(B.2.14) $\quad Z(r,r) = r \qquad \forall \; r : u \leq r \leq \beta.$

II. Demostremos que con la hipótesis hecha sobre u debe cumplirse $Z(r,r) = r$ también para cada $r \in [\alpha,u]$.

En efecto:

Sea ν el extremo inferior del conjunto $\Lambda_z \cap (\alpha,\beta)$

$(\nu = \text{Inf } r : Z(r,r) = r,\ r \in [\alpha,\beta])$

Supongamos por el absurdo que $\nu > \alpha$.

Entonces $\forall\, r \in (\alpha,\nu)\quad Z(r,r) = \bar{f}\,\{f(r) + f(r)\} > r$.

Si en la ecuación (B.2.8) hacemos

$$w = \alpha < r < \nu < p = q < u.$$

Como $g_w = g_\alpha = o$ tenemos $g^{-1}(g_u - g_w) = u$, luego

$$\text{(B.2.15)}\quad \bar{g}\,\{g_\circ Z[u, g^{-1}(g_\circ \bar{g}\,\{g_\circ Z[g^{-1}(g_p - g_r), g^{-1}(g_p - g_r)] + g_r\})]\}$$

$$= g\{g_\circ Z[g^{-1}(g_\circ \bar{g}\,\{g_\circ Z(u,p)\} - g_r),\ g^{-1}(g_p - g_r)] + g_r\}$$

Escogiendo p y r de manera que

$$g_p - g_r < g_\nu \quad \text{en consecuencia} \quad g^{-1}(g_p - g_r) < \nu$$

$$g_p - g_\nu < g_r \quad \text{"} \qquad \text{"} \quad g^{-1}(g_p - g_\nu) < r,$$

y que además

$$g_u < g\,\{\bar{f}\,[f(g^{-1}(g_p - g_r)) + f(g^{-1}(g_p - g_r))]\} + g_r < g_\beta$$

se obtiene

$$(B.2.16)\quad u < g^{-1}\,[g\,\{\bar{f}\,[f(g^{-1}(g_p-g_r)) + f(g^{-1}(g_p-g_r))]\} + g_r]$$

y también

$$r < g^{-1}\,\{g_\beta - g\,[\bar{f}\,[f(g^{-1}(g_p-g_r)) + f(g^{-1}(g_p-g_r))]]\}\ ;$$

o sea,

$$\begin{aligned} g\,\{\bar{f}\,[f(g^{-1}(g_p-g_r)) + f(g^{-1}(g_p-g_r))]\} &< g_\beta - g_r \\ &< g_p - g_r \\ &< \nu\ . \end{aligned}$$

Entonces, tanto en el primero como en el segundo miembro de (B.2.15) siempre resulta $\bar{g} = g^{-1}$.

Luego de (B.2.15) se obtiene

$$\begin{aligned} &Z\,[u, g^{-1}\,\{g_\circ Z\,[g^{-1}(g_p-g_r),\ g^{-1}(g_p-g_r)] + g_r\}] \\ &= \bar{g}\,\{g_\circ Z\,[g^{-1}(g_\circ Z(u,p) - g_r),\ g^{-1}\,(g_p-g_r)] + g_r\} \end{aligned}$$

Como $Z(u,p) = \mathrm{Sup}\,(u,p) = u$ se tiene $g_\circ Z(u,p) = g_u$ y $g^{-1}(g_p-g_r) < \nu$; esta relación se escribirá bajo la forma

$$\begin{aligned} &Z\,[u, g^{-1}\,\{g_\circ\,\bar{f}\,[f_\circ g^{-1}\,(g_p-g_r) + f_\circ g^{-1}(g_p-g_r)] + g_r\}] \\ &= \bar{g}\,\{g_\circ Z\,[g^{-1}\,(g_u-g_r),\ g^{-1}\,(g_p-g_r)] + g_r\} \end{aligned}$$

Por (B.2.16) resulta entonces

$$g^{-1}\,\{g_\circ\,\bar{f}\,\,[f_\circ g^{-1}\,(g_p-g_r) + f_\circ g^{-1}(g_p-g_r)] + g_r\} =$$

$$= g^{-1}\{g_\circ Z\,[g^{-1}(g_u - g_r),\ g^{-1}(g_p - g_r)] + g_r\}$$

Como g es estrictamente creciente, tenemos:

$$\bar{f}\,[f_\circ g^{-1}(g_p - g_r) + f_\circ g^{-1}(g_p - g_r)] + g_r$$

$$= Z\,[g^{-1}(g_u - g_r),\ g^{-1}(g_p - g_r)] + g_r$$

o sea:

$$Z\,[g^{-1}(g_p - g_r),\ g^{-1}(g_p - g_r)] = Z\,[g^{-1}(g_u - g_r),\ g^{-1}(g_p - g_r)],$$

lo cual es absurdo ya que el miembro de la izquierda es estrictamente menor que ν y la función $Z(r,s)$ es estrictamente creciente entre ambas variables, hasta cuando asume el valor ν.

El absurdo se obtuvo al suponer que $v > \alpha$, entonces como quiera sea seleccionado un número $x > \alpha$ existe un elemento $r \in \Lambda_z$ con $\alpha < r < x$. Ya que el punto I asegura que $[r,\beta] \subset \Lambda_z$, también $x \in \Lambda_z$.

Así el Lema B.5 queda demostrado.

□

Corolario:

Por el lema precedente, la función $Z(r,s)$ puede tener una sola de las siguientes expresiones:

$$(B.2.17)\quad Z(r,s) = \operatorname{Sup}(r,s) \qquad \forall\ (r,s) \in [\alpha,\beta]^2$$

$$(B.2.18)\quad Z(r,s) = \bar{f}\,[f(r) + f(s)] \qquad \forall\ (r,s) \in [\alpha,\beta]^2$$

donde f es una función real, continua, estrictamente creciente y nula en α.

B.3. La Función ϕ_C

Lema B.6

De (B.1.2) del Lema B.4 se reconoce que, si la función $Z(r,s)$ es del tipo (B.2.17) entonces se tiene:

$$(B.3.1)\quad \phi_C(u,v,w) = \operatorname{Sup}(u,v) \qquad \forall (u,v,w) \in C$$

Demostración:

ya probamos que

$$\phi_C(u,v,w) = Z\,[g^{-1}\{g(u)-g(w)\},\ g^{-1}\{g(v)-g(w)\}] * w$$

Supongamos ahora que $Z(r,s)$ sea del tipo (B.2.17) y sea $u \leq v$. Entonces

$$Z\,[g^{-1}\ \{g(u)-g(w)\},\ g^{-1}\{g(v)-g(w)\}] = g^{-1}\{g(v)-g(w)\}$$

luego

$$\begin{aligned}\phi_C(u,v,w) &= (g^{-1}\ \{g(v) - g(w)\}) * w \\ &= \bar{g}\ \{g_o g^{-1}\ (g(v) - g(w) + g(w)] \\ &= \bar{g}\ \{g(v) - g(w) + g(w)\} \\ &= \bar{g}\ \{g(v)\} = V = \operatorname{Sup}(u,v).\ \text{ Por lo tanto}\end{aligned}$$

$$\phi_C(u,v,w) = \operatorname{Sup}(u,v) \qquad \forall\ (u,v,w) \in C$$

□

Lema B.7

Si la función $Z(r,s)$ es del tipo (B.2.18), entonces ϕ_c es necesariamente una de las dos funciones:

(B.3.2) $\phi_c(u,v,w) = \bar{g}\ [g(u) + g(v) - g(w)]$

(B.3.3) $\phi_c(u,v,w) = \bar{g}\ [\log_a\ (a^{g(u)} + a^{g(v)} - a^{g(w)}]$

Demostración

Si en (B.2.8) ponemos $q = w$ y si escogemos las variables p, q, r, u de manera que ambos miembros de (B.2.8) sean menores que β, recordando que $\bar{g}(x) < \beta \Longrightarrow \bar{g}(x) = g^{-1}(x)$, obtenemos lo siguiente:

$$\bar{g}\ \{g_\circ Z\, [g^{-1}(g_u - g_w),\ g^{-1}(g_\circ \bar{g}\{g_\circ Z\, [g^{-1}\ (g_p - g_r), g^{-1}(g_w - g_r)] + g_r - g_w)] + g_w\}$$

$$= \bar{g}\ \{g_\circ Z\ [g^{-1}(g_\circ \bar{g}\ \{g_\circ Z\ [g^{-1}(g_u - g_w),\ g^{-1}(g_w - g_w)] + g_w\} - g_r\ ,\ g^{-1}(g_p - g_r)] + g_r\}$$

Desarrollando el primer miembros tenemos que:

Como $Z(x,y) = f^{-1}\ [f(x) + f(y)]$ y $\bar{g} = g^{-1}$ entonces

$$\bar{g}\ \{g_\circ Z\ [g^{-1}(g_u - g_w),\ g^{-1}(g_\circ \bar{g}\{g_\circ Z\ [g^{-1}(g_p - g_r),\ g^{-1}(g_w - g_r)] + g_r\} - g_w)] + g_w$$

$$= g^{-1}\ \{g \circ f^{-1}\ [f \circ g^{-1}\ (g_u - g_w) + f \circ g^{-1}(g \circ g^{-1}\{g \circ f^{-1}[f \circ g^{-1}(g_p - g_r) + f \circ g^{-1}(g_w - g_r)] + g_r - g_w]$$

Similarmente, desarrollando el segundo miembro tenemos que

$$\bar{g}\ \{g\circ Z\ [g^{-1}(g\circ \bar{g}\ \{g\circ Z\ [g^{-1}\ (g_u-g_w),\ g^{-1}(o)] + g_w\}- g_r),\ g^{-1}(g_p-g_r)] + gr\}$$

$$= g^{-1}\ \{g\circ f^{-1}\ [f\circ g^{-1}(g\circ g^{-1}\{g\circ f^{-1}\ [f\circ g^{-1}(g_u-g_w) + f\circ g^{-1}(o)] + g_w\}- g_r) + f\circ g^{-1}(g_p-g_r)] + g_r$$

Como $g(x)$ es estrictamente monótona y $f\circ g^{-1}(o) = 0$, entonces;

$$g_o f^{-1}[f\circ g^{-1}(g_u-g_w) + f\circ g^{-1}(g\circ f^{-1}[f\circ g^{-1}(g_p-g_r) + f\circ g^{-1}(g_w-g_r)] - (g_w-g_r)] + (g_w-g_r)$$

$$= g\circ f^{-1}\ [f\circ g^{-1}\ (g\circ f^{-1}[f\circ g^{-1}(g_u-g_w)] + g_w-g_r) + f\circ g^{-1}\ (g_p-g_r)]$$

Como $(g\circ f^{-1}) \circ (f\circ g^{-1}) = I$, entonces

$$g\circ f^{-1}\ [f\circ g^{-1}\ (g_u-g_w) + f\circ g^{-1}\ (g\circ f^{-1}[f\circ g^{-1}(g_p\cdot g_r)+f\circ g^{-1}(g_w-g_r)] - (g_w-g_r)] + (g_w-g_r)$$

$$= g\circ f^{-1}\ [f\circ g^{-1}\ (g_u-g_r) + f\circ g^{-1}\ (g_p-g_r)]$$

Luego haciendo $\theta = f\circ g^{-1} \implies \theta^{-1} = g\circ f^{-1}$

$$\left.\begin{array}{l} t = g_u-g_w \\ m = g_q-g_r=g_w-g_r \end{array}\right\} \implies t + m = g_u - g_r$$

$$n = g_p-g_r$$

y escogiendo p, r, u, w de manera que sea

(B.3.4) $o \leq t + m < g_\beta$ y $o < \theta\ (t+m) + \theta(n) < \theta(g_\beta)$,

obtenemos

(B.3.5) $\theta^{-1}[\theta(t) + \theta\{\theta^{-1}(\theta(m) + \theta(n) - m\}] + m = \theta^{-1}[\theta(t+m) + \theta(n)]$

Si en (B.3.5) hacemos $\theta^{-1}[\theta(m) + \theta(n)] - m = x$

$$\theta^{-1}[\theta(m) + \theta(n)] = x + m \Longrightarrow \theta(m) + \theta(n) = \theta(x+m)$$
$$\Longrightarrow \theta(n) = \theta(x+m) - \theta(m)$$

Luego de (B.3.5) se obtiene:

$$\theta^{-1}[\theta(t) + \theta(x)] + m = \theta^{-1}[\theta(t+m) + \theta(x+m) - \theta(m)]$$

entonces:

(B.3.6) $\theta\{\theta^{-1}(\theta(t) + \theta(x)) + m\} = \theta(t+m) + \theta(x+m) - \theta(m)$

Haciendo

(B.3.7) $h_m(r) = \theta\{\theta^{-1}(r) + m\} - \theta(m)$

$$\xi = \theta(t), \; \eta = \theta(x)$$

de (B.3.6) obtenemos:

$$\theta\{\theta^{-1}(\xi+\eta) + m]\} - \theta(m) = \theta(t+m) - \theta(m) + \theta(x+m) - \theta(m)$$

$$\Longrightarrow h_m(\xi+\eta) = \theta\{\theta^{-1}(\xi) + m\} - \theta(m) + \theta\{\theta^{-1}(\eta) + m\} - \theta(m).$$

Entonces:

(B.3.8) $h_m(\xi+\eta) = h_m(\xi) + h_m(\eta) \;\forall\; (\xi,\eta)$ Tal que

$0 \leq \xi = \theta(t) < \theta(g_\beta - m)$ y

$$0 \leq \eta = \theta(x) = \theta\{\theta^{-1}[\theta(m) + \theta(n)] - m\}$$

$$< \theta\{\theta^{-1}[\theta(m) + \theta(g_\beta) - \theta(t+m)] - m$$

(estas limitaciones provienen directamente de (B.3.4)).

Para cualquier valor fijado de $\xi = \theta(t)$, η varía en un intervalo $[o,s(t))$ semiabierto donde $s(t) = \theta\{\theta^{-1}[\theta(m) + \theta(g_\beta) - \theta(t+m)]$.

Para cada t tal que $\theta(t) < \theta(g_\beta - m)$ resulta $s(t) > 0$, luego $[o,s(t))$ es no vacío.

En efecto.

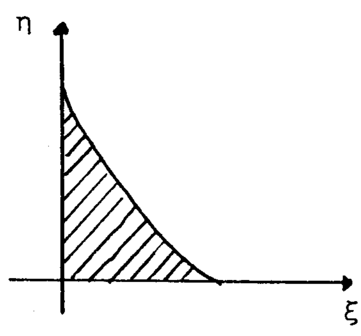
$$0 \leq t+m < g_\beta \implies \theta(t+m) < \theta(g_\beta) \qquad (\theta \text{ es monótona})$$

$$\implies \theta(m) + \theta(g_\beta) > \theta(m) + \theta(r+m)$$

$$\implies [\theta(m) + \theta(g_\beta) - \theta(t+m)] > \theta(m)$$

$$\implies \theta^{-1}[\theta(m) + \theta(g_\beta) - \theta(t+m)] - m > 0$$

En el plano (ξ,η), el conjunto en el cual (B.3.7) debe satisfacerse es del tipo trazado en la figura. Pues la solución de la ecuación de Cauchy (B.3.8) (con m fijado) en el intervalo $[o,\theta(g_\beta - m)]$ es la función:



$$h_m(\xi) = c(m) \cdot \xi$$

donde $c(m)$ es una función que depende solo de m.

Recordando la definición (B.3.7) de la función $h_m(\xi)$ se reconoce que la función θ debe ser solución de la siguiente ecuación:

(B.3.9) $\theta(t+m) = c(m).\ \theta(t) + \theta(m) \quad \forall\ (t,m)\ /\ 0 \leq t+m \leq g_\beta$

Es fácil reconocer que las únicas funciones $\theta(x)$ definidas en el intervalo $[o,g_\beta]$ que satisface la ecuación (B.3.9) son las siguientes:

(B.3.10) $\theta(x) = c\ x \qquad x \subset [o,g_\beta] \qquad$ c es constante

(B.3.11) $\theta(x) = h(a^x-1) \quad x \in [o,g_\beta] \qquad$ h es constante

En efecto:

Como $\theta(t+m) = \theta(m+t)$ de (B.3.9) se deduce que

(B.3.12) $c(m)\ .\ \theta(t) + \theta(m) = c(t)\ .\ \theta(m) + \theta(t)$

Recordando que $t \neq 0 \iff \theta(t) \neq 0$

($\theta(t)$ es estrictamente creciente y nula para $t = 0$)

Se reconoce inmediatamente que:

Si $\exists\ \bar{t} \neq 0$ tal que $c(\bar{t})=1$ entonces $c(t) = 1$ sobre todo el intervalo $[o,g_\beta]$.

(Haciendo $m = \bar{t}$ en (B.3.12) se obtiene

$\theta(t) + \theta(1) = c(t).\theta(1) + \theta(t)$

lo cual implica $c(t) = 1 \ \forall \ t$

La ecuación (B.3.9) se reduce en tal caso a la ecuación de Cauchy y por lo tanto tiene como solución general $\theta(x) = c.x$ $x \in [o, g_\beta]$ que es (B.3.10).

En caso contrario, si $c(t) \neq 1 \ \forall \ t \neq 0$, de (B.3.12) resulta entonces que sobre todo el intervalo $[o, g_\beta]$:

(B.3.13) $\dfrac{\theta(t)}{c(t)-1} = \dfrac{\theta(m)}{c(m)-1} = h$

donde $h \neq 0$ es una constante ya que no depende de la variable t.

De (B.3.13) $\theta(t) = h(c(t) - 1)$

Luego $\theta(t+m) = h\,[c(t+m) - 1]$

Además de (B.3.9)

$$\begin{aligned}\theta(t+m) &= c(m)\ \theta(t) + \theta(m)\\ &= c(m)\ [h(c(t) - 1)] + h\ (c(m) - 1)\\ &= h\ [c(m)\ .\ c(t) - c(m) + c(m) - 1]\\ &= h\ [c(m)\ .\ c(t) - 1]\end{aligned}$$

Luego $\cancel{h}\ [c(t+m) - 1] = \cancel{h}\ [c(m)\ .\ c(t) - 1]$

$\Longrightarrow \quad c(t+m) - \cancel{1} = c(m)\ .\ c(t) - \cancel{1}$

entonces obtenemos:

(B.3.14) $c(t+m) = c(m) \,.\, c(t)$

La ecuación de Cauchy (B.3.14) tiene como soluciones las funciones $c(t) = a^t$ y por lo tanto (B.3.13) indica que $\theta(t) = h(a^t-1) \; \forall \; t > 0$.

Tal expresión es válida también para $t = 0$, siendo $\theta(o) = 0$. En efecto:

Se sigue inmediatamente de (B.3.14), haciendo $m = 0$, que $c(o) = 1$, por lo tanto de (B.3.13) se tiene $h(o) = 0$.

Las funciones (B.3.10) y (B.3.11) son las soluciones de la ecuación (B.3.4).

Como $\theta(x) = f_\circ g^{-1}(x)$ se reconoce que la función $f(x)$ que define $Z(r,s)$ por medio de (B.2.18) puede asumir solo una de las formas siguientes:

(B.3.15) $f(x) = c.g(x) \qquad x \in [\alpha,\beta]$

(B.3.16) $f(x) = h.(a^{g(x)} - 1) \qquad x \in [\alpha,\beta]$

y en correspondencia a cada una de éllas se tiene para $Z(r,s)$:

(B.3.17) $Z(r,s) = \bar{g}\,[g(r) + g(s)]$

(B.3.18) $Z(r,s) = \bar{g}\,[\log_a (a^{g(r)} + a^{g(s)} - 1)]$

Se tendrá así para la función ϕ_c la expresión (B.3.2) y (B.3.3).

□

Conclusión:

Los resultados establecidos en los lemas B.1, B.2, B.6. y B.7. aseguran que cualquier solución del sistema (3.3.8)(3.3.9) es el tipo siguiente:

Si $(u,v) \in \Lambda = (\mathbb{R}^{+})^2 - \cup (\alpha_i,\beta_i)^2$, entonces $\phi(u,v,w) = \operatorname{Sup}(u,v)$

Si $(u,v) \in (\alpha_i,\beta_i)^2$ entonces la restricción ϕ_i de (u,v,w) al conjunto $Q_i \subset \Gamma_3$ con $Q_i = ((\alpha_i,\beta_i)^2 \times \mathbb{R}^{+}) \cap \Gamma_3$ es una de las tres funciones siguientes:

$$\phi_i^{(1)}(u,v,w) = \operatorname{Sup}(u,v)$$

$$\phi_i^{(2)}(u,v,w) = \bar{g}_i\,[g_i(u) + g_i(v) - g_i(\bar{w})]$$

$$\phi_i^{(3)}(u,v,w) = \bar{g}_i\,[\log_{a_i}(a_i^{g_i(u)} + a_i^{g_i(v)} - a_i^{g_i(\bar{w})})]$$

donde a_i es una constante y $\bar{w} = \operatorname{Sup}(w,\alpha_i)$.

Recíprocamente, se verifica fácilmente que cada función definida por medio de las especificaciones

$$\phi(u,v,w) = \begin{cases} \operatorname{Sup}(u,v) & \text{si } (u,v) \in \mathbb{R}^2 - \cup(\alpha_i,\beta_i)^2 \\ \phi^{(1)} \vee \phi^{(2)} \vee \phi^{(3)} & \text{si } (u,v) \in (\alpha_i,\beta_i)^2 \end{cases}$$

Representa una posible ley de localidad compatible con la ley de independencia (1.7.1).

APENDICE C

Demostración **Teorema** 3.3.2:

La demostración de este resultado se efectuará en dos fases, determinando primero cómo ϕ depende de las primeras variables u y v y sucesivamente cómo depende de la variable w.

C.1. **Dependencia de** $\phi(u,v,w)$ de (u,v)

Haciendo $G_w(\alpha,\beta) = \phi(\alpha,\beta,w)$ y teniendo en cuenta la definición de la función ϕ se deduce que la función G_w esta definida en el conjunto

$$\text{(C.1.1)} \qquad \Gamma_2^{(w)} = [w, +\infty)^2 \subset (\mathbb{R}^+)^2$$

y para la ecuación (3.3.10.b) asume valores en el intervalo real $[w, +\infty)$.

Del sistema (3.3.10) (3.3.11) se deduce además que estas funciones deben satisfacer el sistema de ecuaciones funcionales

$$\text{(C.1.2)} \qquad \begin{array}{ll} \text{a.} & G_w(u,v) \geq \operatorname{Sup}(u,v) \\ \text{b.} & G_w(u,v) = G_w(v,u) \\ \text{c.} & u' < u'' \Longrightarrow G_w(u',v) \leq G_w(u'',v) \\ \text{d.} & G_w[u, G_w(p,q)] = G_w[G_w(u,q),p] \\ \text{e.} & G_w(u,w) = u \\ \text{f.} & G_w(u, +\infty) = +\infty \text{ (consecuencia de e)} \end{array}$$

Está demostrado [4,6] que son soluciones de tal sistema todas y solo las funciones $G_w(u,v)$ definidas así:

$$(C.1.3)\quad G_w(u,v) = \begin{cases} \bar{g}_{w,i}\{g_{w,i}(u) + g_{w,i}(v)\} & \text{si } (u,v)\in(\alpha_{w,i},\beta_{w,i})^2 \\ \\ \operatorname{Sup}(u,v) & \text{en otro caso} \end{cases}$$

donde $\{(\alpha_{w,i},\beta_{w,i})\}$ es una sucesión de intervalos abiertos, disjuntos contenidos en $[w, +\infty)$, $\{g_{w,i} : [\alpha_{w,i},\beta_{w,i}] \to \mathbf{R}^+\}$ es una sucesión de funciones con propiedades análogas a las de la función g_i enunciadas en el teorema.

C.2. **Estructura de los elementos idempotentes**

A los elementos del conjunto $\Lambda_w = (w, +\infty) - \bigcup_{i\in I}(\alpha_{w,i},\beta_{w,i})$ que son idempotente con respecto a la ley $G_w(u,v)$ $(\alpha\in\Lambda_w \to G_w(u,u)=u)$ los llamaremos w - idempotentes.

El conjunto complementario de Λ_w en $(w, +\infty)$ viene indicado por $\bar{\Lambda}_w$ y obviamente

$$\bar{\Lambda}_w = \bigcup_{i\in I}(\alpha_{w,i},\beta_{w,i})$$

Se demostrará en este parágrafo que los elementos w-idempotentes coinciden sustancialmente con los elementos o-idempotentes; en forma más precisa, se demostrará el siguiente Lema.

Lema C.1

$$\Lambda_w = \Lambda_o \cap (w, +\infty)$$

Demostración:

Cada función $G_w(u,v)$ debe satisfacer la siguiente relación de compatibilidad con $G_o(u,v)$:

(C.2.1) $G_w\,[u, G_o(p,q)] = G_o\,[G_w(u,p),q]$ para $q \geq 0$, $p \geq w$, $u \geq w$

En efecto:

$$\begin{aligned} G_w\,[u, G_o(p,q)] &= G_w\,[u, \phi(p,q,o)] \\ &= \phi\,[u, \phi(p,q,o), w] \\ &= \phi\,[u, \phi(q,p,o), w] \\ &= \phi\,[\phi(u,p,w),q,o] \\ &= \phi\,[G_w(u,p), q,o] \\ &= G_o\,[G_w(u,p),q]. \end{aligned}$$

Utilizando esta relación se demuestra el lema probando que se cumplen las dos implicaciones siguientes:

a. $x \in \Lambda_w \Longrightarrow x \in \Lambda_o \cap (w, +\infty)$

b. $x \in \Lambda_{\hat{o}} \cap (w, +\infty) \Longrightarrow x \in \Lambda_w$

La implicación a. se demostrará por el absurdo:

Supongamos que un elemento $\nu \in \Lambda_w$ no sea o-idempotente; entonces $\nu \in (\alpha_{o,r},\beta_{o,r})$ intervalo de la sucesión $\{(\alpha_{o,i},\beta_{o,i})\}$.

Si en (C.2.1) hacemos $u = \nu$ se obtiene:

(C.2.2) $G_w\,[\nu,\ G_o(p,q)] \;=\; G_o[G_w(\nu,p),q]$

lo cual, como $\nu \in \Lambda_w$, implica

(C.2.3) $\mathrm{Sup}\,[\nu, G_o(p,q)] = G_o[q,\ \mathrm{Sup}\,(p,\nu)]$

Escogemos ahora los números p y q internos al intervalo $(\alpha_{o,r}, \beta_{o,r})$ de tal manera que:

(C.2.4)

a. $w \leq p < \nu$

b. $g_{o,r}\,(q) > 0$

c. $g_{o,r}(p) + g_{o,r}(q) < g_{o,r}(\nu)$ ó $G_o(p,q) < \nu$

La existencia de dos números que satisfacen esta propiedad está garantizada por el hecho que:

i) $\nu > w$

ii) $\nu > \alpha_{o,r}$

iii) $g_{o,r}$ es una función estrictamente creciente y tal que

$g_{o,r}(\alpha_{o,r}) = 0$

En correspondencia a la escogencia de las variables p,q la fórmula (C.2.3) asume la forma

(C.2.5) $$\begin{aligned}\mathrm{Sup}(\nu, G_o(p,q) &= G_o(q,\nu)\\ &= \bar{g}_{o,r}\,[g_{o,r}(q) + g_{o,r}(\nu)]\end{aligned}$$

Por la manera de escoger p y q

$G_o(p,q) = \bar{g}_{o,r}\{g_{o,r}(p) + g_{o,r}(q)\}$ es menor que ν y por lo tanto $\mathrm{Sup}[\nu, G_o(p,q)] = \nu$. Por otro lado, siendo por hipótesis $\nu < \beta_{o,r}$ y $g_{o,r}(q) > 0$, la cantidad

$G_o(q,\nu) = \bar{g}_{o,r}\{g_{o,r}(\nu) + g_{o,r}(q)\}$ es estrictamente mayor que ν. Luego, si $\nu \in \Lambda_w$ y $\nu \notin \Lambda_o$ la relación (C.2.2) no puede satisfacerse.

Pues es absurdo suponer que un elemento w-idempotente pertenezca a $\bar{\Lambda}_o$ y por lo tanto

$$x \in \Lambda_w \Longrightarrow x \in \Lambda_o ;$$

y como $\Lambda_w \subset (w, +\infty)$, entonces a. queda así demostrada.

La implicación b. se prueba en forma análoga a la demostración de a. (en la ecuación (C.2.2) los papeles de las funciones G_o y G_w son completamente simétricos) llegando a que es absurdo suponer que un elemento que pertenece a Λ_o (y mayor que w) no sea w-idempotente. □

C.3. **Dependencia de ϕ de la variable w.**

Sea $(\alpha,\beta) \subset \Lambda_0$ y $(\alpha,\beta) \cap [w, +\infty) \neq \phi$.

Por el Lema C.1. para cualquier terna $(p,q,u) \in (\alpha,\beta) \times (w,\beta) \times (w,\beta) \subset \mathbb{R}^+$ la ecuación funcional (C.2.1) asume la forma.

$$\bar{g}_w [g_w(u) + g_w \circ \bar{g} \{g(p) + g(q)\}]$$

(C.3.1)

$$= \bar{g} [g \circ \bar{g}_w \{g_w(u) + g_w(q)\} + g(p)]$$

en la cual se escribe por comodidad $g(\xi) = g_0(\xi)$

Escogemos ahora p,q,u de manera que se satisfaga la condición:

(C.3.2) a. $g(p) + g \circ \bar{g}_w \{g_w(u) + g_w(q)\} < g(\beta)$.

En base a (C.3.1) esta limitación puede ser satisfecha solo si también se verifica la desigualdad

(C.3.2) b. $g_w(u) + g_w \circ \bar{g} \{g(p) + g(q)\} < g_w(\beta)$

y además también resulta obvio que las desigualdades (C.3.2;a., b.) serán satisfechas solo si

$$\text{(C.3.3)} \quad \begin{array}{l} g(p) + g(q) < g(\beta) \\ g_w(u) + g_w(q) < g_w(\beta) \end{array}$$

Por lo tanto para los valores p,q,u que satisfacen las desigualdades (C.3.2;a.,b.), la ecuación (C.3.1) se escribirá de la forma siguiente:

$$g_w^{-1}\{g_w(u) + g_w \circ g^{-1}[g(p) + g(q)]\}$$

(C.3.4)

$$= g^{-1}\{g \circ g_w^{-1}\ [g_w(u) + g_w(q)] + g(p)\}$$

Estudiaremos a continuación lo que sucede cuando:

i) $\alpha \geq w$

ii) $\alpha < w$

i) **Lema C.2:**

Si $w \leq \alpha$, entonces $G_w(u,v) = \bar{g}\{g(u) + g(v)\}$

Demostración

En la hipótesis tenemos que $g_w(\alpha) = g(\alpha) = 0$. Si en (C.3.4) hacemos $q = \alpha$, se obtiene

(C.3.5) $g_w^{-1}\{g_w(u) + g_w(p)\} = g^{-1}\{g(u) + g(p)\}$,

relación que por la desigualdad (C.3.2) debe ser verificada en correspondencia a todas las parejas (u,p) para las cuales

(C.3.6) $g(u) + g(p) < g(\beta)$

Esto quiere decir que para todo $q = \alpha$, p,u que cumplen (C.3.2) (a y b),

$$G_w(q,p) = g^{-1}(g(q) + g(p)).$$

Si ambas relaciones (C.3.2) (a y b) no se cumplen entonces $G_w(q,p) = \beta$ y también $\bar{g}(g(p) + g(q)) = \beta$. En estos dos casos el lema queda demostrado.

Hay que examinar qué pasa cuando se cumple uno de las dos desigualdades (C.3.2) (a ó b) y no se cumple la otra. Supongamos por ejemplo que sea

$$g(p) + g(u) \geq g(\beta) \text{ y } g_w(u) + g_w(p) < \beta .$$

En este caso, el primer miembro de (C.3.1) es estrictamente menor que β, mientras el segundo es igual a β. Pero esto no puede ser porque (C.3.1) tiene que ser una identidad. Entonces siempre

$$g(p) + g(u) < \beta \Longleftrightarrow g_w(p) + g_w(u) < \beta$$

queda así demostrado que

$$g_w(u,p) = \bar{g}\,\{g(u) + g(p)\}$$

Para completar el análisis vamos a establecer la relación que hay entre las dos funciones $g(t)$ y $g_w(t)$.

Haciendo $\theta(t) = g_w \bullet g^{-1}(t)$, $\xi = g(u)$, $\eta = g(p)$

las ecuaciones (C.3.5)(C.3.6) asumen la forma

$$\theta(\xi + \eta) = \theta(\xi) + \theta(\eta)$$

$$\forall\ \xi,\eta:\quad \xi + \eta < g(\beta).$$

De aquí y teniendo en cuenta que $\theta(x)$ es continua, ya que es una composición de dos funciones continuas, se deduce (por ser esta una ecuación de Cauchy) que:

$$\theta(x) = a\,x \qquad \forall\ x \in [o,\ g(\beta)]$$

Finalmente, recordando que $\theta(\tau) = g_w \circ g^{-1}(\tau)$ se obtiene

$$g_w \circ g^{-1}(\tau) = a\,\tau$$

$$g^{-1}(\tau) = g_w^{-1}\,(a\tau) \quad \text{y haciendo } \xi = a\tau$$

$$\text{(C.3.7)} \quad g^{-1}(\xi/a) = g_w^{-1}(\xi) \qquad \text{pues}$$

$$\text{(C.3.8)} \quad g_w(\xi) = a\ g(\xi)$$

(C.3.8) representa por lo tanto la relación que hay entre las funciones que definen la restricción de G_o y G_w al subconjunto $[\alpha,\beta]^2$. Finalmente utilizando (C.3.8) y (C.3.7) para escribir $G_w(u,v)$ aún se obtiene, cuando $w \leq \alpha$

$$\text{(C.3.9)} \quad \begin{aligned} G_w(u,v) &= \bar{g}_w(g_w(u) + g_w(v)) \\ &= \bar{g}\,(g(u) + g(v)) \end{aligned}$$

en acuerdo con lo ya establecido.

□

Lema C.3

Si $w > \alpha$, entonces $G_w(u,v) = \bar{g}\,\{g(u) + g(v) - g(w)\}$

Demostración

En la hipótesis dada $g(\alpha) = 0$, $g(w) > 0$, $g_w(w) = 0$, mientras que g_w no está definida en el intervalo (α,w).

Introduciendo la función $f_w(u) = g_w(u) + g(w)$ definida en el intervalo $[w,\beta]$; se tiene obviamente

$$g_w(u) = f_w(u) - g(w)$$

$$g_w^{-1}(u) = f_w^{-1}\{u + g(w)\}$$

y la ecuación (C.3.4), en términos de las funciones g y f_w, asume la forma

$$f_w^{-1}\{f_w \circ g^{-1}[g(p) + g(q)] + f_w(u) - g(w)\}$$

$$= g^{-1}\{g(p) + g \circ f_w^{-1}[f_w(u) + f_w(q) - g(w)]\}$$

Haciendo $q = w$, $x = g(p)$, $y = g(u)$, $\theta(t) = f_w \circ g^{-1}(t)$ tal ecuación asume la forma

(C.3.10) $\theta\,[x + g(w)] + \theta(y) - g(w) = \theta(x + y)$

donde, en base a (C.3.1) y recordando que g_w está definida en el intervalo $[w,\beta]$, las variables x y y están sujetas a las limitaciones siguientes:

$$\text{(C.3.11)} \qquad \begin{aligned} x \geq 0, \quad & y \geq g(w) \\ x + y \quad & \leq \quad g(\beta) \end{aligned}$$

Como $\theta[g(w)] = f_w \circ g^{-1} [g(w)] = f_w(w) = g_w(w) + g(w) - g(w)$,

de la (C.3.10) se obtiene

$$\frac{\theta(y + x) - \theta(y)}{x} = \frac{\theta[g(w) + x] - \theta[g(w)]}{x}$$

En base a (C.3.11), esta relación para cualquier y del intervalo $(g(w), g(\beta))$, debe ser verificada en correspondencia a cualquier x del intervalo $(o, g(\beta)-y) \neq \phi$. Luego en todos los puntos del intervalo $[g(w), g(\beta)]$ la función $\theta(\xi)$ tiene la misma pendiente; además, en ese intervalo θ es monótona (porque es composición de funciones monótonas).

Ella es, por lo tanto, una función lineal con coeficientes dependiendo posiblemente de w y pasa obviamente por el punto $(g(w),g(w))$, siendo como ya demostramos $\theta[g(w)] = g(w)$. Luego

$$\theta(\xi) = K_w . \xi + (1-K_w) . g(w)$$

De esta expresión, recordando que $\theta(t) = f_w \circ g^{-1}(t)$, se reconoce que entre las funciones $f_w(.)$ y $g(.)$ subsisten las relaciones

$$f_w(x) = K_w \cdot g(x) + (1 - K_w) \cdot g(w)$$

$$f_w^{-1}(x) = g^{-1} \{[x + (K_w - 1).g(w)] / K_w\}$$

Habiendo determinado $f_w(.)$ ($g_w(.)$) en términos de $g(.)$ es fácil obtener la restricción de $G_w(u,v)$ al subconjunto $[w,\beta]^2$ en términos de las funciones $g(.)$. Se tiene, en efecto:

$$\begin{aligned} G_w(u,v) &= g_w^{-1} \{g_w(u) + g_w(v)\} \\ &= f_w^{-1} \{g_w(u) + g_w(v) + g(w)\} \\ &= f_w^{-1} \{f_w(u) + f_w(v) - g(w)\} \\ &= f_w^{-1} \{K_w [g(u) + g(v) - 2\, g(w)] + g(w)\} \\ &= g^{-1} \{[g(u) + g(v) - 2\, g(w)] + \frac{g(w) + (K_w - 1)\, g(w)}{K_w}\} \\ &= g^{-1} (g(u) + g(v) - g(w)) \end{aligned}$$

para cualquier pareja $(u,v) \in [w,\beta]^2$ en correspondencia a las cuales resulta

$$g(u) + g(r) - g(w) < g(\beta)$$

Por la continuidad y por la monotonía de $G_w(u,v)$ se debe poner $G_w(u,v) = \beta$ en correspondencia a las parejas $(u,v) \in [w,\beta]^2$ que no satisfacen esta desigualdad. En forma compacta se escribirá entonces

$$G_w(u,v) = \bar{g}\ \{g(u) + g(v) - g(w)\}$$

□

Basándonos en los lemas C.1, C.2, C.3 se demuestra que cualquier solución del sistema (3.3.10) (3.3.11) es una función del tipo (3.3.14). Para éllo basta observar que:

a. Para determinar $\phi(u,v,o)$ se pueden escoger arbitrarias, particulares sucesiones

$$\{(\alpha_{o,i},\beta_{o,i})\} = \{(\alpha_i,\beta_i)\} \text{ y } \{g_{o,i}\} = \{g_i\}$$

b. Si $u \geq w$, $v \geq w$, $(u,v) \notin \cup\ (\alpha_i,\beta_i)^2$, o sea, si no existe un intervalo abierto $(\alpha_i,\beta_i) \subset \bar{\Lambda}_o$ que contenga ya sea a u, o a v, entonces $(u,v) \notin \cup\ (\alpha_{w,i},\beta_{w,i})$, mientras que por el lema C.1.

$\cup\ (\alpha_{w,i},\beta_{w,i})^2 = \cup\ (\alpha_i,\beta_i)^2 \cap [w,\alpha]^2$. Por lo tanto por (C.1.2) se tiene que

$$\phi(u,v,w) = \text{Sup}\ (u,v)$$

c. Si $(u,v) \in (\alpha_j,\beta_j)^2$, con $(\alpha_j,\beta_j) \subset \bar{\Lambda}_o$, entonces por los lemas C.2 y C.3

$$\phi(u,v,w) = \bar{g}_j\,[g_j(u) + g_j(v)] \quad \text{si} \quad \alpha_j > w$$

(C.3.12)

$$\phi(u,v,w) = \bar{g}_j\,[g_j(u) + g_j(v) - g_j(w)] \quad \text{si} \quad \alpha_j < w$$

expresiones que pueden escribirse en la forma compacta (3.3.14) siempre que se recuerde que $g_j(\alpha_j) = 0$.

El resultado ahora recordado muestra que cualquier solución del sistema (3.3.10) (3.3.11) debe necesariamente ser la forma (3.3.14). Por otra lado, es muy fácil verificar que cualquier función del tipo (3.3.14) es solución del sistema (3.3.10) (3.3.11), con lo cual el Teorema de representación está completamente demostrado.

□

APENDICE D

Las ecuaciones (4.2.1.a,d) (4.2.2) con $*$ = + no contiene la ley $F(x,y)$ y por lo tanto sus consecuencias se utilizan para demostrar tanto el teorema 4.5.1 como el teorema 4.5.2. y en particular las siguientes proposiciones:

Proposición D-1

Haciendo $G(\xi,\eta) = -\Psi(o,\xi,o,\eta)$ se tiene

(D-1) $\Psi(x,y,u,v) = u - G(y-x,u-v)$

Demostración

El resultado se obtiene fácilmente haciendo el cambio de variables $t = y-x$, $s = v-u$ y poniendo $\overline{\Psi}(x,t,u,s) = \Psi(x,t+x,u,s+u)$. La ecuación (4.2.2) asume entonces la forma:

$$\overline{\Psi}(x,t,u,s) + w = \overline{\Psi}(x+z,t,u+w,s)$$

de lo que se deduce, haciendo $u = x = o$, que $\overline{\Psi}$ no depende de la primera variable y es lineal en la tercera. Luego (D-1) está demostrada.

□

Proposición D-2

La función $G(\xi,\eta)$ cumple con la condición

(D-2) $G(-\xi,-\eta) = \eta + G(\xi,\eta)$

Demostración:

Como $\Psi(x,y,u,v) = \Psi(y,x,v,u)$ tenemos
$u + G(y-x,v-x) = v + G(x-y,u-v)$.

□

Demostración del Teorema 4.5.1

En este caso, $F(x,y) = \mathrm{Inf}(x,y)$. Haciendo $x = y \leq z$, $v = o$ y $\xi = z-x$ en la ecuación (4.5.1) y utilizando para $\Psi(x,y,u,v)$ la expresión (D-1), se obtiene:

$$\text{(D-3)} \quad G(o,u) + G[\xi, w + G(o,u)] = G[\xi,w] + G[o,u + G(\xi,w)]$$

Como $z \geq x$, (D-3) tiene que cumplirse para todo $\xi \geq o$. Al final, haciendo $\xi = o$ resulta

$$\text{(D-4)} \quad G(o,-\eta) = \eta + G(o,\eta)$$

$$\text{(D-5)} \quad G(o,u) + G[o,w + G(o,u)] = G(o,w) + G[o,u + G(o,w)]$$

La solución de este sistema (determinada por Z. Daróczy [26]) está compuesta por las tres funciones siguientes:

i) $G(o,\eta) = \dfrac{1}{2}(|\eta| - \eta)$

ii) $G(o,\eta) = \dfrac{1}{2}(|\eta| + \eta)$

iii) $G(o,\eta) = -\dfrac{1}{k}\log(1 + e^{k\eta})$

Para cada uno de estos tres casos se procede ahora en la forma siguiente:

A. Se sustituye la expresión $G(o,\eta)$ en la ecuación (D-3) y luego se determina la forma correspondiente de la función $G(\xi,\eta)$. Esa expresión de $G(\xi,\eta)$ dependerá de una función arbitraria de la variable ξ, digamos $f(\xi)$.

B. A partir de $G(\xi,\eta)$ se construye $\Psi(x,y,u,v)$ por (D-1). Luego se sustituye esa función en la ecuación de asociatividad (4.2.1.b) que en este caso, haciendo $x \leq y \leq z$, asume la for-

(D-6) $\Psi[x,z,\ \Psi(x,y,u,v),w] = \Psi[x,y,u,\ \Psi(y,z,v,w)]$.

Así se obtiene una nueva condición, permitiendo determinar la función $f(\xi)$.

Caso (i)

De (D-3) se obtiene

(D-7) $$\frac{1}{2}(|u| - u) + G\left[\xi,w + \frac{1}{2}(|u| - u)\right]$$

$$= G(\xi,w) + \frac{1}{2}\left[\,|u + G(\xi,w)| - u - G(\xi,w)\right]$$

Haciendo $u = o$ se obtiene

$$\frac{1}{2}\left[\,|G(\xi,w)| - G(\xi,w)\right] = o$$

Por lo tanto, para todo $\xi \geq o$ se cumple $G(\xi,w) \geq o$.

Fijamos ahora ξ y supongamos que para un cierto valor $\bar{w}$ resulte $F(\xi,\bar{w}) = o$. Haciendo en (D-7) $u = - t$ $(t \geq o)$, $w = \bar{w}$ se tiene $t + G(\xi,\bar{w}+t) = t$ y por lo tanto resulta.

$$\text{(D-8)} \quad G(\xi,\bar{w}) = o \Longrightarrow G(\xi,\bar{w}+t) = o \quad \forall\ t \geq o$$

Supongamos ahora que para un cierto valor $\bar{\bar{w}}$ sea $G(\xi,\bar{\bar{w}}) > o$. Entonces, para todo $o \leq t \leq G(\xi,\bar{\bar{w}})$ se tiene

$$\text{(D-9)} \quad t + G\ (\xi,\bar{\bar{w}} + t) = G(\xi,\bar{w}).$$

La limitación $t \in [o,\ G(\xi,\bar{\bar{w}})]$ es necesaria porque sólo en este caso

$$\frac{1}{2} [\,|u + G(\xi,\bar{\bar{w}})| - u - G(\xi,\bar{\bar{w}})] = \frac{1}{2} [\,|G(\xi,\bar{\bar{w}})-t| + t - G(\xi,\bar{\bar{w}})] = o$$

Por lo tanto para cada $\xi \geq o$ existe un valor $f(\xi)$ tal que

$$\text{(D-10)} \quad G(\xi,\eta) = \begin{cases} f(\xi) - \eta & \text{Si } \eta \leq f(\xi) \\ \cdot & \\ o & \text{Si } \eta \geq f(\xi) \end{cases}$$

(En efecto:

De (D-9) se deduce que $G(\xi,\eta)$ es lineal, con coeficiente -1, en la segunda variable y de (D-8) se tiene que para $\eta > \bar{\eta}$, en correspondencia al cual $G(\xi,\bar{\eta}) = o$, $G(\xi,\eta) = o$).

Por lo tanto, cuando $\xi = y-x \geq o$ la ley de composición $\Psi(x,y,u,v)$ asume la forma

$$\text{(D-11)} \quad \Psi(x,y,u,v) = u - G(y-x, v-u) = \operatorname{Inf}[u, v-f(y-x)].$$

Por simetría, cuando $x - y \geq o$ tendremos:

$$\text{(D-12)} \quad \Psi(x,y,u,v) = v - G(x-y, u-v) = \operatorname{Inf}[v, u-f(x-y)].$$

Sustituyendo estas expresiones en (D-6) y recordando que en esa ecuación $x \leq y \leq z$ obtenemos

$$\operatorname{Inf}[u, v-f(y-x), w-f(z-x)]$$

$$= \operatorname{Inf}[u, v-f(y-x), w-f(z-y)-f(y-x)]$$

que es una identidad si y sólo si

$$f(\xi+\eta) = f(\xi) + f(\eta)$$

Por lo tanto $f(\xi) = \lambda\,\xi$ y las expresiones (D-11)(D-12) se pueden escribir en la forma compacta

$$\text{(D-13)} \quad \Psi(x,y,u,v) = \operatorname{Inf}(u-\lambda x, v-\lambda y) + \lambda \operatorname{Inf}(x,y)$$

Caso (ii)

Procediendo de manera análoga a la que usamos al desarrollar el caso (i) se obtiene sin dificultad

$$(D\text{-}14) \quad \Psi(x,y,u,v) = \operatorname{Sup}(u-\lambda x, v-\lambda y) + \lambda \operatorname{Inf}(x,y)$$

Caso (iii)

Efectuando un cambio de incógnita, poniendo $\forall\ x > o$

$$g(\xi,x) = \exp\left[-k\ G(\xi, \frac{1}{k} \log x)\right]$$

$$G(\xi,\eta) = -\frac{1}{k} \log g(\xi, e^{k\eta})$$

Las relaciones (iii) y (D-3) asumen la forma

$$g(o,x) = 1 + x$$

$$g(o,x).g\left[\xi, y/g(o,x)\right] = g(\xi,y).g\left[o, x/g(\xi,y)\right],$$

de las cuales se obtiene

$$(1+x).g\left[\xi, y/(1+x)\right] = g(\xi,y) + x \quad ,$$

ecuación que debe satisfacerse $\forall\ x,y > o$

a. Para $t \geq 1$, hacemos $y = t$, $x = t - 1$, obteniendo

$$t.g\ (\xi,1) = g(\xi,t) + t-1$$

b. Para $o < t < 1$, hacemos $y = 1$, $x = \frac{1}{2} - 1$, obteniendo

$$\frac{1}{t} . g(\xi, t) = g(\xi, 1) + \frac{1}{t} - 1$$

En ambos casos

$$g(\xi, t) = 1 + t\, [f(\xi) - 1]$$

$$G(\xi, \eta) = - \frac{1}{k} \log [1 + f(\xi)\, e^{k\eta}] \quad \text{con } f(\xi) = g(\xi, 1)$$

Luego, utilizando (D-1) se obtiene para $\Psi(x,y,u,v)$, cuando $x \leq y$, la expresión

$$\text{(D-15)} \quad \Psi(x,y,u,v) = u - G(y-x, v-u)$$
$$= \frac{1}{k} \log \; e^{ku} + \frac{1}{k} \log [1 + f(y-x)\; e^{k(v-u)}]$$

Por simetría cuando $x \geq y$ tenemos

$$\text{(D-16)} \quad \Psi(x,y,u,v) = \frac{1}{k} \log [f(x-y) e^{ku} + e^{kv}]$$

Sólo queda por determinar la función $f(\xi)$.

Como (D-15) y (D-16) tienen que coincidir cuando $x = o$, tendrá que cumplirse que

$$f(o) = 1.$$

Además, sustituyendo las expresiones precedentes de Ψ en la ecuación (D-6), después de unas cuantas sumplificaciones triviales obtenemos

$$e^{ku} + f(y-x)e^{kv} + f(z-x)\ e^{kw}$$

$$= e^{ku} + f(y-x)\ e^{kv} + f(y-x)\ .\ f(z-y)\ e^{kw},$$

la cual es una identidad si y sólo si

$$f(\xi+\eta) = f(\xi)\ .\ f(\eta).$$

Por lo tanto, excluyendo que $f(\xi) = o$ ya que es incompatible con la condición $f(o) = 1$, tendrá que ser $f(\xi) = e^{c\xi}$.

Sustituyendo esta expresión en (D-15) o (D-16) según sea $y \geq x$ o $y \geq x$, se obtiene la forma explícita de $\Psi(x,y,u,v)$ que en forma compacta puede escribirse, haciendo $\lambda = - c/k$, bajo la forma

(D-17) $$\Psi(x,y,u,v) = \frac{1}{k} \log [e^{k(u-\lambda x)} + e^{k(v-\lambda y)}] + \lambda \operatorname{Inf}(x,y).$$

□

Demostración del Teorema 4.5.2

No será posible desarrollar aquí la demostración completa del teorema porque no logramos obtener el trabajo de B. Forte donde se demuestra que las funciones (4.5.6) - (4.5.9) agotan las soluciones del sistema (4.2.1) (4.2.2) con $* = +$ y $T = F_c$ que no son de clase C_2. Sin embargo, nos parece oportuno exponer en general: (A) la demostración de que (4.5.8) y (4.5.9) son las únicas leyes de composición de la incertidumbre $(+, F_c)$-universales de clase C_2. (B) el método debido a P. Benvenuti [23] con el cual se construyen las leyes $(+, F_c)$-universales por extensión de leyes "elementales".

Parte (A):

Haciendo: $\xi = -c \log x$, $\eta = -c \log y$, $p = \xi/(\xi+\eta)$, $T = 1/(\xi+\eta)$, se reconoce que $G(x-y,u-v)$ depende de x y y solo del parámetro p; luego tenemos

$$\psi(x,y,u,v) = u - g\left[\frac{e^{-x/c}}{e^{-x/c} + e^{-y/c}}, v - u\right]$$

$$= u - \Theta\left[\frac{e^{-x/c}}{e^{-xc} + e^{-y/c}}, u - v\right]$$

donde $\Theta(p,x) = g(p,-x)$. Tal como ha sido construida esta función $\Psi(x,y,u,v)$, cumple con la condición (4.2.2). Para que satisfaga las otras ecuaciones del sistema (4.2.1) es necesario que la función $\Theta(x,y)$ sea solución del siguiente sistema:

(D-18) $y_1 < y_2 \longrightarrow \Theta(u,y_1) \leq \Theta(u,y_2)$

(D-19) $\Theta(u,y) = y + \Theta(1-u,-y)$

(D-20) $y + \Theta\{u,x-y + \Theta[(u + v -1)/u,y]\}$

$$= x + \Theta\{v,y-x + \Theta[(v+u-1)\ v,x]\}$$

Derivando primero y segundo miembros de (D-20) respecto a la variable u se obtiene

(D-21) $\Theta_1\ \{u,x-y + \Theta[(u+v-1)/u,y]\} + \Theta_2\{u,x-y+\Theta[(u+v-1)/u,y]\}.$

$$.\,\Theta_1\ \{(u+v-1)/u,y\}\ .\ (1-v)/u^2 =$$

$$= \Theta_2\ [v,y-x + \Theta\{(v+u-1)/v,x]\ .\ \Theta_1\{(v+u-1)/v,x\}/v$$

donde Θ_1 y Θ_2 son las derivadas parciales de la función $\Theta(u,y)$ respecto a la primera y segunda variable, respectivamente.

Haciendo $r = 1 - u$ y multiplicando por $(1-u)\ u$ se obtiene

(D-22) $u(1-u).\Theta_1\ [u,x-y+\Theta\ (o,y)] + (1-u).\Theta_2\,[u,x-y+\Theta(o,y)].\Theta_1(o,y) =$

$$= u.\Theta_2\ [1-u,\ y-x + \Theta\ (o,x)].\Theta_1(o,x)$$

Esta ecuación tiene que satisfacerse $\forall\ u \in (o,1)$ pero por continuidad también tiene que satisfacerse en los extremos. Haciendo $u = 1$ y $y = x - \Theta\ (o,x) + t$, se obtiene

(D-23) $\Theta_2 (o,t) \cdot \Theta_1(o,x) = o$

Por medio de esta relación se logra el siguiente Lema:

Lema D-1

Para las soluciones de clase C_2 del sistema (D-18)-(D-21) sólo hay dos alternativas:

I $\Theta_1 (x,t) = o$

II $\Theta(o,x) = o$, $\Theta(1,x) = x$

Demostración:

En efecto de (D-23) se destaca que

o $\Theta_2 (o,t) = o$

o $\Theta_1 (o,x) = o$

Si $\Theta_1 (o,x) = o$, haciendo en (D-22) $x = y - \Theta(o,y) + t$ se obtiene

$$\Theta_1(u,t) = o$$

En caso contrario, si $\Theta_2(o,t) = o$, es decir, $\Theta(o,t) = k$, haciendo $u = 1$ y $x = t$ en (D-19) se obtiene $\Theta(1,t) = t + k$.

Por otro lado, haciendo $u = 1$, $v = o$, $x = y$, y naturalmente $\Theta(o,t) = k$, $\Theta(1,t) = t + k$, en (D-20) se obtiene

$\Theta(1,k) = \Theta(o,x+k)$, es decir, $2k = k$ o sea $k = o$.

Caso I:

$\Theta_1(u,t) = 0$, o sea, $\Theta(u,t)$ sólo depende de t.

Haciendo $\Theta(u,t) = f(t)$, las ecuaciones (D-19) y (D-20) toman la forma

$$f(-x) = x + f(x)$$
$$y + f[y-x + f(-y)] = x + f[x-y + f(-x)]$$

cuya solución, conseguida por Z. Daroczy [26] es:

$$f(x) = k \log (1 + e^{-x/k}) \quad , \text{ o sea,}$$

(D-24) $\Theta(u,x) = k \log (1 + e^{x/k})$

Caso II

La función Θ depende de u, luego por el lema D-1 tenemos

$$\text{(D-25)} \quad \begin{array}{ll} \text{(a) } \Theta(o,x) = o, & \text{(b) } \Theta(1,x) = x \\ \text{(b) } \Theta_2(o,x) = o, & \text{(d) } \Theta_2(1,x) = 1 \end{array}$$

Haciendo $y = o$ en (D-22) y tomando en cuenta (D-25) obtenemos

$$\text{(D-26)} \quad \begin{aligned} u(1-u)\, \Theta_1(u,x) + (1-u)\, \Theta_2(u,x).\Theta_1(o,o) \\ = u\, \Theta_2(1-u,-x).\Theta_1(o,x) \end{aligned}$$

Derivando esa ecuación respecto a u y haciendo después $u = 1$ y tomando en cuenta (D-25,c,d) se desprende que

$$\Theta_1 (o,o) . \Theta_{12} (o,x) = o$$

Examinaremos separadamente los dos casos

1. $\Theta_{12} (o,x) = o$ o sea $\Theta_1 (o,x) = c \quad (c \neq 0)$

2. $\Theta_1 (o,o) = o$

Caso 1

Haciendo $\Theta_1 (o,x) = c$ en (D-26) se obtiene

(D-27) $u(1-u).\Theta_1 (u,x) + c.\Theta_2 (u,x) = c.u$

ecuación diferencial en la función $\theta(u,x)$ que tiene como líneas características las curvas de ecuación

(D-28) $x = c \log \frac{u}{1-u} - k$,

y por lo tanto es equivalente a la ecuación diferencial ordinaria

$$\frac{d}{du} \Theta \left[u, c \log \frac{u}{1-u} - k \right] = \frac{c}{1-u} ;$$

de aquí se sigue que

(D-29) $\Theta \left[u,c \log \frac{u}{1-u} - k \right] = c \log \frac{1}{1-u} + f(k)$

Derivando esa ecuación respecto a u y haciendo después $u = 1$ y tomando en cuenta (D-25,c,d) se desprende que

$$\Theta_1 (o,o) . \Theta_{12} (o,x) = o$$

Examinaremos separadamente los dos casos

1. $\Theta_{12} (o,x) = o$ o sea $\Theta_1 (o,x) = c \quad (c \neq 0)$

2. $\Theta_1 (o,o) = o$

Caso 1

Haciendo $\Theta_1 (o,x) = c$ en (D-26) se obtiene

(D-27) $u(1-u) . \Theta_1 (u,x) + c . \Theta_2 (u,x) = c . u$

ecuación diferencial en la función $\theta(u,x)$ que tiene como líneas características las curvas de ecuación

(D-28) $x = c \log \frac{u}{1-u} - k$,

y por lo tanto es equivalente a la ecuación diferencial ordinaria

$$\frac{d}{du} \Theta \left[u, c \log \frac{u}{1-u} - k \right] = \frac{c}{1-u} \; ;$$

de aquí se sigue que

(D-29) $\Theta \left[u, c \log \frac{u}{1-u} - k \right] = c \log \frac{1}{1-u} + f(k)$

donde f es una función arbitraria de clase C_1. Al variar k en $(-\infty, +\infty)$, la variable x definida por (D-28) varía también en $(-\infty, +\infty)$ $\forall$ u y por lo tanto de (D-29) se obtiene

$$\text{(D-30)} \quad \Theta(u,x) = c \log \frac{1}{1-u} + f\left[c \log \frac{u}{1-x} - x\right]$$

Sustituyendo esa expresión en (D-19) (D-20) se reconoce que f tiene que satisfacer el sistema

$$f(-x) = x + f(x)$$
$$y + f[y - x - f(-y)] = x + f[x-y - f(-x)]$$

De las tres soluciones (i,ii,iii) del sistema (D-4), (D-5), que ya estudiamos, sólo la última es de clase C_1.

Sustituyendo esa expresión en (D-30) se obtiene

$$\text{(D-31)} \quad \Theta(u,x) = k \log [(1-u)^{\alpha} + u^{\alpha} . e^{x/k}]$$

con $\alpha = - c/k \geq 2$ para que $\Theta(u,x)$ sea de clase C_2, dependa de u y satisfaga II.

Caso 2

Haciendo $\Theta_1(o,o) = o$ en (D-26) se obtiene

$$\text{(D-32)} \quad (1-u)\, \Theta_1(u,x) + \Theta_1(o,x) . \Theta_2(u,x) = \Theta_1(o,x)$$

Conseguiremos ahora la forma de la función $\Theta_1(o,x)$ para que (D-32) asuma la forma de una ecuación diferencial en la incógnita $\Theta(u,x)$. Después de efectuar sucesivamente las operaciones:

- Hacer $u = v$ en (D-21)
- Derivar la ecuación obtenida respecto a u.
- Hacer después $u = 1$
- Tomar en cuenta la condición II (b) $\Theta(1,x) = x$ que implica $\Theta_2(1,x) = 1$, $\Theta_{22}(1,x) = 0$, se llega a la ecuación

(D-33) $\Theta_{12}(1,x).\Theta_1(1,y)-\Theta_1(1,y) = \Theta_{12}(1,y).\Theta_1(1,x)- \Theta_1(1,x)$

Por otra lado, derivando (D-19) respecto a u se obtiene

$$\Theta_1(1,x) = - \Theta_1(o,-x)$$

Sustituyendo esa expresión en (D-33), haciendo $t = - x$, $s = - y$, $\Psi(.) = \Theta_1(o,.)$ se obtiene $\Psi(s).[1-\Psi'(t)] = \Psi(t)[1 - \Psi'(s)]$ de donde, siendo $\Psi(t)/[1-\Psi'(t)] = \lambda = \text{const}$, se tiene

$$\Psi'(t) = 1 + \lambda \Psi(t)$$

cuyas soluciones, con la condición inicial $\Psi(o) = o$, son:

$$\Theta_1(o,t) = \Psi(t) = t \qquad (\lambda = o)$$

$$\Theta_1(o,t) = \Psi(t) = \frac{1}{\lambda}(e^{\lambda t} -1) \qquad (\lambda \neq o)$$

Sustituyendo estas funciones en (D.32) obtenemos para $\Theta(u,x)$ las dos ecuaciones diferenciales

$$(1-u).\Theta_1(u,x) + x.\Theta_2(u,x) = x$$

$$\lambda\ (1-u).\Theta_1(u,x) + (e^{\lambda x} -1).\Theta_2(u,x) = e^{\lambda x} - 1.$$

Sus soluciones que satisfagan la condición de frontera $\Theta(o,x) = o$ se obtienen con el método clásico de las curvas características y son las siguientes:

(D-34) $\Theta(u,x) = u.x$

(D-35) $\Theta(u,x) = k \log (1-u + u\, e^{x/k}),\ k = 1/\lambda$.

Las funciones (D-24), (D-31), (D-34), (D-35) constituyen el conjunto de las soluciones de clase C_2 del sistema (D-18)-(D-20). A la solución (D-34) corresponde la ley de composición

(D-36) $$\Psi(x,y,u,v) = \frac{u\, e^{-x/c} + v\, e^{-y/c}}{e^{-x/c} + e^{-y/c}}$$

A las soluciones (D-24), (D-31), (D-35) corresponden las leyes de composición

(D-37) $$\Psi(x,y,u,v) = -\ k \log\ e^{-(u+\lambda x)/k} + e^{-(v+\lambda y)/k} +$$

$$+ \lambda c\ \log\ e^{-x/c} + e^{-y/c}$$

con: $\lambda = o$ para (D-24), $\lambda = k/c$ para (D-35) y $\lambda = k\alpha/c$ para (D-31).

□

Parte (B)

Supongamos que $\Psi(x,y,u,v)$ sea una ley de composición de la incertidumbre (+, Shannon) -universal y en forma precisa compatible con la ley de composición de la información

$$F_c(x,y) = -c \log (e^{-x/c} + e^{-y/c})$$

Luego se reconoce fácilmente que también es $(+, F_c)$-universal la ley de composición

$$\text{(D-38)}\quad T_\lambda \Psi(x,y,u,v) := \Psi(x,y,u-\lambda x,v-\lambda y) + \lambda F_c(x,y)$$

Para demostrar esto es necesario comprobar que si Ψ satisface el sistema (4.2.1)-(4.2.2), también $T_\lambda \Psi$ cumple con él. No realizaremos aquí la verificación completa. (hecho de pura rutina) sino sólo, como ejemplo, demostraremos que si Ψ es asociativa también $T_\lambda \Psi$ lo es:

$$T_\lambda \Psi[x, F_c(y,z), u, T_\lambda \Psi(y,z,v,w)] =$$

$$= \Psi[x,F_c(y,z),u-\lambda x, [\Psi(x,z,v-\lambda y,w-\lambda z) + \lambda F_c(y,z)] - \lambda F_c(y,z)] +$$

$$+ \lambda F_c [x, F_c(y,z)] =$$

$$= \Psi[x, F_c(y,z), u-\lambda x, \psi(y,z,v-\lambda y,w-\lambda z) + \lambda F_c \{x,F_c(y,z)\}] =$$

$$= \text{(A)} =$$

$= \Psi[F_c(x,y),z, \Psi(x,y,u-\lambda x,v-\lambda y),w-\lambda z] + \lambda F_c [F_c(x,y),z] =$

$= \Psi[F_c(x,y),z,\{\Psi(x,y,u-\lambda x,v-\lambda y) + \lambda F_c(x,y)\} -\lambda F_c(x,y),w-\lambda z] +$

$+ \lambda F_c [F_c(x,y),z] =$

$= T_\lambda \Psi[F_c(x,y),z,T_\lambda \Psi(x,y,u,v),w]$ l.q.q.d

El punto clave es el paso A, allí se utiliza el hecho de que Ψ y F_c son asociativas.

Aplicando esa transformación a las soluciones (D-36) y (D-37) se obtienen las leyes de composición (4.5.8) y (4.5.9). Además se reconoce fácilmente que las dos funciones

$$\Psi_i (x,y,u,v) = \mathrm{Inf}(u,v)$$

$$\Psi_s (x,y,u,v) = \mathrm{Sup}(u,v)$$

con soluciones del sistema (4.2.1) (4.2.2). La transformación (D-38) aplicada a estas funciones da lugar a las leyes de composición (4.5.6) y (4.5.7). Hemos logrado así introducir de forma sistemática las cuatro familias de soluciones del sistema (4.2.1) (4.2.2) (con x = + y F = F_c)

En otras palabras, existen cuatro "soluciones elementales"

1. $[u\, e^{-x/c} + v\, e^{-x/c}] \; / \; [e^{-x/c} + e^{-y/c}]$ (Shannon)

2. $- k \log [e^{-u/k} + e^{-v/k}]$ (Renyi)

3. $\mathrm{Inf}\,(u,v)$

4. $\mathrm{Sup}\,(u,v)$

a partir de las cuales, utilizando la transformación (D-38), se obtiene todas las leyes de composición de la incertidumbre $(+,F_c)$-universales hasta ahora conocidas. Claro que así no hemos demostrado que no hay otras soluciones. Pero como ya dijimos al comienzo, la demostración de que son las únicas sí existe, pero no pudimos conseguirla.

□

APENDICE E

Demostración del Teorema 4.6.1

Busquemos la solución del sistema (4.2.1) (4.2.2) (4.6.1), para reducir el número de variables independientes, haremos el siguiente cambio en (4.2.2): $w = \mathrm{Inf}(u,v)$, $z \leq \mathrm{Inf}(x,y)$.

Entonces obtenemos

$$\Psi(x \vee z, y \vee z, u \vee w, v \vee w) = \Psi(x,y,u,v) \vee w$$

$$\Longrightarrow \Psi(x,y,u,v) = \Psi(x,y,u,v) \vee w \geq w = \mathrm{Inf}\,(u,v)$$

en definitiva

$$\text{(E-1)} \quad \Psi(x,y,u,v) \geq \mathrm{Inf}(u,v)$$

Supongamos $u \leq v$; haciendo en (4.2.2) $z = w = u = \mathrm{Inf}(u,v)$. Por (E-1) se obtiene entonces

$$\Psi(x,y,u,v) = \Psi(x,y,u,v) \vee w = \Psi(x \vee w, y \vee w, u \vee w, v \vee w) =$$

$$= \Psi(z, y \vee w, u, v)$$

Es decir, la función Ψ no depende del más pequeño entre los valores x y y, por lo tanto

$$\Psi(x,y,u,v) = P(\mathrm{Sup}\,(x,y),u,v)\,.$$

Ahora estudiaremos la función $P(t,u,v)$ cuyas propiedades, deducidas de (4.2.1) (4.2.2) son las siguientes:

(E-2)

a. $\mathrm{Dom}\, P = \{(t,u,v) : t \leq \mathrm{Sup}(u,v)\}$

b. $P(t,u,v) = P(t,v,u)$

c. $u' < u'' \Longrightarrow P(t,u'v) \leq P(t,u'',v)$

d. $P(t,u,v) \geq u \wedge v$

e. $P[(a \wedge b) \vee t,\ P(a \vee b,\ u,v),w]$

$= P[a \vee (b \wedge t),\ u,\ P(b \vee t,\ v,w)]$

f. $P(t,u,v) \vee w = P(t \vee z,\ u \vee w,\ v \vee w)$

Haciendo $e = P(o,o,o)$ se reconoce inmediatamente que:

(E-3) $m(u) := P(u,u,u) = u \vee e$

En efecto:

$P(u,u,u) = P(o,o,o) \vee u = e \vee u$

Por otra parte, consideremos la función

$n(u) := P(o,e,u)$, demostremos que

(E-4) $n(u) := P(o,e,u) = k \wedge (u \vee e) = k \wedge m(u)$

donde k es un número adecuado mayor que e $(k \geq e)$.

En efecto:

la función $n(u)$ goza de las siguientes propiedades:

(E-5) $n(t)$ es no decreciente

(E-6) $n(n(s)) = n(s)$

La condición (E-5) es una consecuencia obvia de (E-2) c. Para demostrar (E-6) se observa antes de todo que

(E-7) $n(o) = n(e) = e$

En efecto;

$$n(o) \vee e = P(o,e,o) \vee e = P(e,e,e) = e$$

y $n(e) \vee e = P(o,e,e) \vee e = P(e,e,e) = e$, ya que $n(o) \leq e$, $n(e) \leq e$.

Por otra parte $e \geq o \implies P(o,e,o) \geq P(o,o,o)$

entonces $n(o) = P(o,e,o) \geq P(e,e,e) = e$ y $n(e) = P(o,e,e) \geq e \wedge e = e$.

Luego la propiedad (E-7) queda así demostrada.

Para demostrar (E-6) se utiliza (E-2 . e) haciendo $a = b = t = o$, $u = v = e$, $s = w$.

$$P\,[o,P(o,e,e),s] = P\,[o,e,\ P(o,e,s)]$$

$$\implies P\,[o,\ n(e),s] = P[o,e,n(s)]$$

$$\implies P\,[o,e,s] = n(n(s))$$

$$\implies n(s) = n(n(s))$$

Como $n(s)$ es continua y no decreciente, asume todos los valores entre su mínimo $n(o) = e$ y su máximo $n(\infty) = P(o,e,+\infty) := k$; para cada $u \in [e,k)$ existe $s = s(u) < +\infty$ tal que $n(s) = n(s(u)) = u$. De (E-6) se deduce que para cada $u \in [e,k)$ se obtiene $n(u) = u$. Como n es no decreciente, se cumple (E-4).

El siguiente resultado sirve para demostrar el teorema de Representación 4.6.1, es decir,

(E-8) $u \wedge v \leq P(t, u, v) \leq u \vee v \vee e$

La primera desigualdad se puede probar fácilmente, ya que se obtuvo de (4.2.1.d); para probar la segunda observamos que siempre se cumple

(E-9) $P(r,u,u) = P(u,u,u)$

En efecto,

$P(t,u,u) = P(t,u,u) \vee u = P(u,u,u)$ por otro lado haciendo $w = \mathrm{Sup}\,(u,v)$, se tiene

$$P(t,u,v) \leq P(t,w,w) = P(w,w,w) = e \vee w.$$

Caso 1

(E-10) $u \vee v \leq e \implies P(t,u,v) = e$

En efecto; como $t \leq u \vee v \leq e$, se cumple

$$
\begin{aligned}
e = P(o,o,o) &\leq P(o,t,t) \\
&= t \vee P(o,t,t) \\
&= P(t,t,t) \\
&\leq P(t,e,e) \\
&= e \vee e \vee e = e
\end{aligned}
$$

Luego de esta sucesión de relaciones, resulta

$$P\ (t,t,t) = P(t,e,e).$$

Por otro lado haciendo $w = u \vee v$ de (E-2.f) se obtiene

$$
\begin{aligned}
P(t,u,v) \vee w &= P(t \vee w,\ u \vee w,\ v \vee w) = \\
&= P(w,w,w) = e
\end{aligned}
$$

De este modo (E-10) está demostrada cuando $w = u \vee v < e$, y por continuidad también cuando $w = e$.

Caso 2

(E-11) $e < u \vee v \leq k \implies P(t,u,v) = u \vee v$

En efecto,

para cada $v \in [e,k)$

$P(o,o,v) \vee e = P(o,e,v) = n(v) = v$

Por lo tanto, siempre que $v > e$ se tiene

(E-12) $P(o,o,v) = v$

Por otro lado, (E-12) es válido cuando $v = e$ ya que $P(o,o,v)$ es continua.

Para demostrar (E-11) suponemos $u \leq v$ y distinguimos dos casos:

i) $t \leq u$:
$$\begin{aligned} P(t,u,v) &= P(t,u,v) \vee u \\ &= P(t \vee u,\ u \vee u,\ v \vee u) \\ &= P(u,u,v) \\ &= P(o,o,v) \vee u \\ &= v \vee u \\ &= v \end{aligned}$$

ii) $t > u$: $P(t,u,v) \vee t = P(t \vee t,\ u \vee t,\ v \vee t) =$
$= P(t,t,v \vee t) = P(o,o,v) \vee t = v \vee t =$
$= v = u \vee v$;

(i) y (ii) demuestran completamente (E-11).

Caso 3

(E-13) $u \wedge v > k \longrightarrow P(t,u,v) = u \wedge v$

En efecto

de (E-11) se tiene que $P(o,e,k) = e \vee k = k$ y por consiguiente, si $w \geq k$, al utilizar (E.2.e) haciendo $a = b = o$, $u = e$, $v = k$ resulta $P(t,k,w) = P(t,P(o,e,k),w) = P(o,e,P(t,k,w)) =$

$$= n\,[P(t,k,w)]$$
$$= k \wedge (P(t,k,w) \vee e)$$
$$= k \wedge P(t,k,w) = k$$

ya que $P(t,k,w) \geq k \wedge w = k$.

Por otro lado, suponiendo $v \geq u \geq k$

$P(t,u,v) = u \vee P(t,k,v) = u \vee k = u = u \wedge u$,

quedando así demostrada (E-13).

Caso 4

(E-14) $u \leq k \leq v \Longrightarrow P(t,u,v) = h(t) \vee u$

donde $h : \mathbb{R}^+ \longrightarrow [o,k]$ es una función continua y tal que

i) $h(t) = k \quad \forall\, t \in [o,k]$

ii) $v' < v'' \Longrightarrow h(v') \geq h(v'')$

En efecto, distingamos dos subcasos:

$t < k \quad , \quad t \geq k.$

$$t \leq k \Longrightarrow k = P(t,u,k) \leq P(t,u,v)$$
$$\leq P(t,k,v) = k$$
$$\Longrightarrow P(t,u,v) = k$$

y (E-14) queda así demostrada, siendo en este caso $h(t) = k \geq u$.

Observemos que se han usado las expresiones (E-11)y (E-13) y que, además, no se puede utilizar este procedimiento si $t > k$ porque $P(t,u,k)$ no está definida.

Si $t > k$, escogiendo $a \leq b \leq t$, $w \geq v$ y recordando que $P(t,v,w) = v \wedge w = v$, de (E-2.c) obtenemos:

$$\text{(E-15)} \quad P[t, P(b,u,v),w] = P[b,u, P(t,v,w)]$$

$$= P[b,u,v]$$

Haciendo $u = 0$, $t = v = b$ y $\bar{h}(x) = P(x,o,x)$, para cada $x \geq k$, se obtiene

$$\text{(E-16)} \quad P[t,\bar{h}(t),w] = \bar{h}(t)$$

Si $u \geq \bar{h}(t)$, de (E-2.f) y (E-16) se obtiene:

$$\bar{h}(t) \vee w = P[t,\bar{h}(t),w] \vee u = P[t \vee u,\bar{h}(t) \vee u, w \vee u] = P(t,u,w)$$

Es decir,

$$P(t,u,w) = \bar{h}(t) \vee u = u$$

Si $\bar{h}(t) > u$ de (E-2.c) y (E-16) se obtiene:

$$P(t,u,w) \leq P(t,\bar{h}(t),w) = \bar{h}(t),$$

$$P(t,u,w) \geq P(t,o,t) = \bar{h}(t) \implies P(t,u,w) = \bar{h}(t)$$

Como $\bar{h}(k) = P(k,o,k) = o \vee k = k$, (E-16) queda demostrada al definir

$$h(t) = \bar{h}(t) \qquad t \geq k$$
$$h(t) = k \qquad t \leq k$$

En conclusión, de los subcasos anteriores se obtiene inmediatamente (E-14) y la propiedad i). Para obtener ii) se utiliza (E-15) haciendo $u = o$, $b = v = v'$, $t = w = v''[v'' \geq v' \Longleftrightarrow w \geq v]$
$P(v'', P(v',o,v'),v'') = P(v',o,v') = h(v')$

Por otra parte $P(v'', P(v',o,v'),v'') \geq P(v'',o.v'') = h(v'')$

Luego

$$v' < v'' \implies h(v') \geq h(v'').$$

$\square$