



UNIVERSIDAD DE JAÉN

**FACULTAD DE CIENCIAS
EXPERIMENTALES
DEPARTAMENTO DE ESTADÍSTICA
E INVESTIGACIÓN OPERATIVA**

TESIS DOCTORAL
**MÉTRICAS Y NORMAS EN AMBIENTES
PROBABILÍSTICOS**

**PRESENTADA POR:
ANTONIO FCO. ROLDÁN LÓPEZ DE HIERRO**

**DIRIGIDA POR:
DRA. DÑA. CONCEPCIÓN BEATRIZ ROLDÁN
LÓPEZ DE HIERRO
JUAN MARTÍNEZ MORENO**

JAÉN, 25 DE MAYO 2012

**ISBN 978-84-8439-659-8
DEPÓSITO LEGAL J-1190-2012**



UNIVERSIDAD DE JAÉN

DEPARTAMENTO DE ESTADÍSTICA E

INVESTIGACIÓN OPERATIVA

MÉTRICAS Y NORMAS EN AMBIENTES PROBABILÍSTICOS

TESIS DOCTORAL

Antonio Francisco Roldán López de Hierro

Jaén, Abril 2012

MÉTRICAS Y NORMAS EN AMBIENTES PROBABILÍSTICOS

Memoria presentada por:

Antonio Francisco Roldán López de Hierro

V^oB^o de los Directores de la memoria

Fdo: Prof. Dra. D^a. C. Roldán

Fdo: Prof. Dr. D. J. Martínez

Departamento de Estadística e Investigación Operativa

Universidad de Jaén

Índice general

Introducción	v
1. Preliminares	1
1.1. Espacios métricos	2
1.2. Operaciones binarias, funciones triangulares y t-normas	4
1.2.1. Conjuntos parcialmente ordenados	4
1.2.2. Operaciones binarias	5
1.2.3. Funciones triangulares sobre un conjunto	6
1.2.4. t-normas y t-conormas	6
1.2.5. Funciones de distribución de tipo distancia. El conjunto Δ^+	9
1.2.6. Funciones triangulares asociadas a t-normas	12
1.3. Números difusos	13
1.3.1. Números difusos triangulares y trapezoidales	15
1.3.2. Aritmética con números difusos	17
1.4. Espacios métricos en los que interviene la idea de incertidumbre	19
1.4.1. Espacios métricos probabilísticos (de Schweizer y Sklar)	19
1.4.2. Espacios métricos difusos de Kaleva y Seikkala	21
1.4.3. Espacios métricos difusos de Kramosil y Michálek	21
1.4.4. Espacios métricos difusos de George y Veeramani	22
1.4.5. Espacios métricos difusos intuicionistas	24
1.4.6. Espacios métricos probabilísticos intuicionistas	24
1.4.7. Espacios métricos \mathfrak{L} -difusos	25

1.4.8.	Espacios normados difusos	28
1.4.9.	Espacios normados difusos intuicionistas	29
1.5.	Puntos fijos de aplicaciones	30
1.6.	El modelo de regresión probabilística	31
2.	Comparación entre las diferentes teorías métricas	39
2.1.	Una descomposición canónica de números difusos	39
2.1.1.	Una topología métrica sobre \mathcal{F}	42
2.1.2.	Funciones triangulares sobre los números difusos (centrados en cero) . . .	43
2.2.	Comparación entre las diferentes teorías métricas	44
2.2.1.	Estudio comparativo con los espacios métricos probabilísticos	45
2.2.2.	La topología fuerte sobre los EMD	47
2.2.3.	Estudio comparativo con los espacios de Kaleva y Seikkala	48
2.2.4.	Estudio comparativo con los espacios intuicionistas de Park	49
2.2.5.	Estudio comparativo con los espacios de Kramosil y Michálek	52
2.2.6.	Estudio comparativo con el espacios métricos probabilísticos intuicionistas	53
3.	Teoremas de punto fijo	55
3.1.	Un teorema de contracción de Banach en espacios métricos \mathfrak{L} -difusos	55
3.2.	Algunos teoremas de punto fijo en espacios normados difusos	58
4.	Regresión difusa con números triangulares	69
4.1.	El modelo de regresión difusa	70
4.1.1.	Trabajos relacionados	71
4.2.	El modelo de regresión propuesto	73
4.2.1.	Un modelo de regresión difusa cuando los valores esperados de las amplitudes no dependen de las variables explicativas	73
4.2.2.	Algunos experimentos de simulación para evaluar las distancias anteriores	88
4.2.3.	Un estudio comparativo: un modelo de regresión difusa en el que los valores de las amplitudes dependen de las variables explicativas	90
4.3.	Un modelo de regresión no lineal para la obtención de biomasa	93

Bibliografía

99

Introducción

Uno de los temas de investigación más interesantes que se plantean actualmente es el de encontrar una definición apropiada de *espacio métrico difuso*, fundamentalmente debido a sus posibles aplicaciones en diversas áreas. A partir de la introducción de la idea de *conjunto difuso*, se han propuesto cientos de trabajos científicos en los que la naturaleza imprecisa en el comportamiento de un sistema dado responde mejor a un modelo difuso que a un modelo estocástico. Es por ello que, hoy por hoy, la *teoría de los conjuntos difusos*, que se apoya especialmente de la noción de *métrica difusa*, es uno de los campos de investigación más activos y florecientes tanto en Matemáticas como en Física. De hecho, muy diversas disciplinas, tanto Naturales como Sociales y Económicas, se han beneficiado ya de sus importantes avances.

La idea de métrica difusa nace de la concepción que tenemos de lo que significa una distancia. Intuitivamente hablando, llamamos *distancia* a la magnitud que mide la separación entre dos personas, objetos o cosas. En el espacio tridimensional que observamos, la distancia entre dos puntos es la longitud del segmento rectilíneo que los une. Podemos decir que este segmento es el camino más corto entre los dos puntos dados. Al formalizar matemáticamente la noción de distancia entre los puntos de un conjunto, es más que razonable imponer las siguientes condiciones.

- (a) La distancia entre cualesquiera dos puntos debe ser un número real concreto, si bien puede cambiar al expresarse en una u otra unidad.
- (b) No hay distancia alguna entre un punto y él mismo, por lo que el número que la mide debe ser cero.
- (c) La distancia entre dos puntos no depende del orden en que se tomen, es decir, hay la misma longitud si se recorre el camino más corto entre ellos en uno u otro sentido.

Una propiedad de la distancia que, aún siendo intuitiva, no es del todo obvia, es la siguiente.

- (d) La distancia entre dos puntos dados del espacio debe ser menor o igual que la longitud de cualquier camino entre esos dos puntos que esté obligado a pasar por un tercer punto.

El axioma (d) se denomina *desigualdad triangular* y es, con mucho, la desigualdad más conocida y que más juego ha dado dentro del campo de la Matemática a lo largo de la Historia. No obstante, cuando se formaliza rigurosamente la noción de distancia, surge una quinta propiedad más práctica que intuitiva.

(e) Si no hay distancia entre dos puntos, entonces esos puntos deben ser iguales.

A lo largo del tiempo se han puesto en duda una y otra vez los últimos cuatro postulados, de cara a testar la consistencia de las diferentes teorías que pueden considerarse cuando se suprime alguno de ellos. Todo ello ha dado paso a las nociones de *pseudodistancia*, *semidistancia*, *cuasidistancia*, etc. Sin embargo, la propiedad que más ha tardado en ponerse en duda ha sido la primera, ya que no era imaginable una perspectiva en la que la distancia entre dos puntos no fuese medida, de manera absolutamente precisa, por un número real. No obstante, el mundo en el que vivimos está lleno de imprecisión y, miremos donde miremos, podemos encontrar diferentes grados de incertidumbre.

Es más que sabido que si arrojamus un objeto al aire con nuestro brazo, éste caerá sobre la Tierra tarde o temprano. Esto no plantea mucha incertidumbre. Además, disponemos de fórmulas matemáticas que determinan el tiempo y la posición en la que caerá en objeto si conocemos ciertas condiciones iniciales. No obstante, estas fórmulas sólo son aproximaciones (excelentes, pero sólo aproximaciones) de lo que podemos observar en la realidad. Nuestro mundo (por suerte) no es perfecto sino que en cualquier experimento sencillo intervienen muchos factores y, a menudo, no se pueden controlar todos ellos. De hecho, pretender controlarlos es un viejo, pero absurdo, sueño.

Si analizamos con precisión los instrumentos de medida que utilizamos, observaremos que todos ellos son capaces de medir indicando un número finito de cifras decimales, siendo la última de ellas, cuando menos, cuestionable. Si peso un objeto y la balanza me indica 7.3 kg , ¿será ese su peso exacto? Es lógico pensar que si hubiésemos podido utilizar un instrumento de medida más preciso, habríamos obtenido una cantidad diferente (por ejemplo, 7.294 gr). Para colmo, Pitágoras demostró que si un aparato de medida fuese capaz de medir con total precisión la longitud del lado de un cuadrado, entonces ese instrumento sería incapaz de medir exactamente la longitud de su diagonal (ya que la raíz cuadrada del número dos es un número irracional). Incluso a veces, el mismo aparato puede ofrecer mediciones diferentes: si queremos medir la longitud exacta de una calle y repetimos el experimento con el mismo instrumento, es posible que obtengamos medidas diferentes. A todos nos ha sorprendido alguna vez que, midiendo varias veces la misma distancia, hemos obtenido resultados diferentes.

De hecho, hay magnitudes que no pueden ser conocidas simultáneamente en un mismo experimento más allá de un cierto grado de precisión. Por ejemplo, el *principio de incertidumbre de Heisenberg* (1927) establece la imposibilidad física de conocer, a la vez y con toda exactitud, la

posición y la velocidad de una partícula elemental (por ejemplo, de un electrón). Es por ello que, conocida la velocidad de una partícula, existe siempre cierta incertidumbre sobre la posición que ocupa. De esta forma, la distancia de esta partícula a un punto fijo, utilizado como referencia, no se puede medir en términos exactos. Cabe entonces preguntarse acerca de si un número real es la mejor forma de medir distancias entre puntos de un espacio.

Muchos/as investigadores/as han profundizado en este problema desde muy diversos puntos de vista pero casi todos ellos se han visto obligados, de una u otra forma, a recurrir al concepto de *probabilidad*. En tal caso, la distancia entre dos puntos ya no se mediría de forma absoluta utilizando un cierto número real, sino que deberíamos más bien hablar de la probabilidad de que la distancia entre dos puntos sea menor que un cierto número real.

Uno de los primeros en hacerlo fue K. Menger. En 1942, Menger [82] propuso generalizar la teoría de los espacios métricos hacia un punto de vista probabilístico, e introdujo la noción de *espacio métrico estadístico*. En este trabajo, Menger observó la necesidad de hacer intervenir cierta clase de funciones de cara a expresar la desigualdad triangular en términos probabilísticos. Estas funciones han dado hoy lugar a lo que se conoce como *funciones triangulares* (o *normas triangulares*). Nueve años después, Menger [83] propuso reemplazar la relación de pertenencia de un punto a un conjunto por una *función de distribución probabilística* que se interpreta, para cada número real x , como la probabilidad de que la distancia entre dos puntos dados sea menor que x . Él llamó a estas nuevas estructuras *emsemble flous*, que se puede traducir del francés como *conjuntos difusos* o *borrosos*.

Pero el verdadero fundador de la *teoría de los conjuntos difusos* es, indudablemente, L.A. Zadeh. Esta teoría está basada en la idea de que no siempre es posible determinar con total exactitud si un elemento pertenece o no a un conjunto. Por ejemplo, pensemos en el conjunto de personas que padecerán la gripe en el invierno de año que viene. Si hoy tuviéramos que determinar si una persona pertenece o no a ese conjunto, no lo podríamos hacer con total precisión, sino que tendríamos que hablar de una cierta *probabilidad* de que el individuo pertenezca a ese conjunto. Dicho de otra forma, existe una cierta *incertidumbre* sobre la pertenencia o no de una persona a ese conjunto.

Matemáticamente, la idea de Zadeh puede ser formalizada como sigue. Ya era conocido con anterioridad que la pertenencia de un elemento a un conjunto puede ser equivalentemente expresada utilizando una cierta aplicación llamada *función característica del conjunto*. Cada subconjunto A de un cierto universo \mathcal{U} viene completamente determinado por la aplicación $f_A : \mathcal{U} \rightarrow \{0, 1\}$ que vale 1 sobre los elementos de A y 0 sobre el resto de elementos. Si se reemplaza el conjunto discreto $\{0, 1\}$ por el intervalo $\mathbb{I} = [0, 1]$, entonces el valor de $f_A(x)$ puede expresarse como la *probabilidad* (o el *grado de certeza* que se tiene) de que el elemento x pertenezca al conjunto A . Así, un *conjunto difuso* no es más que una función $F : \mathcal{U} \rightarrow [0, 1]$ que expresa, de cierta forma, la *incertidumbre* que tenemos de que un elemento pertenezca o no a un conjunto. Como se ha comentado más arriba, a partir de la idea pionera de Zadeh

se han propuesto cientos de trabajos en los que la naturaleza imprecisa en el comportamiento de un sistema dado responde mejor a un modelo difuso que a un modelo estocástico. Y hoy posiblemente sean millones los trabajos que se han desarrollado en el campo de la teoría de los conjuntos difusos y de sus aplicaciones.

A lo largo del tiempo, la forma para expresar esta incertidumbre ha evolucionado de dos formas distintas, una *probabilística* y otra *difusa*. Aunque ambas comparten elementos en común, su naturaleza algebraica es diferente.

La aproximación probabilística a este problema se basa en la idea de que, a pesar de que la distancia entre dos puntos es un número real, en general, no podemos acceder a él utilizando nuestras técnicas y aparatos de medida. Por ello, cuando intentamos medirla, lo que hacemos es repetir un experimento aleatorio que puede ser expresado en forma de variable aleatoria. Si suponemos que estos intentos son independientes y que, al menos en teoría, podemos repetirlos tantas veces como queramos, podremos calcular la distancia que separa dos puntos con un error tan pequeño como se quiera o, dicho de otra forma, con una probabilidad de que nuestra estimación difiera del valor real mayor de $1 - \varepsilon$ (siendo $\varepsilon > 0$ tan arbitrariamente pequeño como se desee). Utilizando esta idea podemos expresar la distancia entre dos puntos como una variable aleatoria que, caracterizada a través de su función de distribución, llevó a la siguiente razonable definición.

Un *espacio métrico estadístico* sobre un conjunto X es un par (X, \mathfrak{F}) donde \mathfrak{F} es una aplicación que asigna a cada par de puntos $x, y \in X$ una función de distribución (denotada por $\mathfrak{F}_{xy}(\cdot)$) de manera que se verifican las siguientes condiciones, para cada $x, y, z \in X$.

- $\mathfrak{F}_{xy}(0) = 0$.
- $\mathfrak{F}_{xy}(t) = 1$ para cada $t > 0$ si, y sólo si, $x = y$.
- $\mathfrak{F}_{xy}(t) = \mathfrak{F}_{yx}(t)$ para cada $t \in \mathbb{R}$.
- Si $\mathfrak{F}_{xy}(t) = 1$ y $\mathfrak{F}_{yz}(s) = 1$, entonces $\mathfrak{F}_{xz}(t + s) = 1$.

Si interpretamos que $\mathfrak{F}_{xy}(t)$ denota la probabilidad de que la distancia entre x e y sea menor que t , las tres primeras propiedades son naturales, y establecen la idea intuitiva de positividad, identidad y simetría. No obstante, la cuarta es una versión ciertamente peculiar de la *desigualdad triangular*, que puede ser interpretada de la siguiente manera:

si estamos seguros de que la distancia entre x e y es menor que t y, al mismo tiempo, estamos seguros de que la distancia entre y y z es menor que s , entonces podemos afirmar que la distancia entre x y z es menor que $t + s$.

No obstante, esta versión de la desigualdad triangular es demasiado débil como para llevarnos a resultados interesantes. El principal inconveniente es que si las funciones de distribución \mathfrak{F}_{xy}

que se utilizan en el modelo sólo toman valores estrictamente menores que 1 (y esto podría ocurrir con suma facilidad), entonces la desigualdad triangular se vuelve vacía de contenido, casi tautológica, y no podríamos estar seguros de que la distancia entre dos puntos del espacio se alcanza en algún momento.

Para solventar este problema, se cambió la desigualdad triangular y Menger la expresó de la siguiente forma:

$$\mathfrak{F}_{xz}(t+s) \geq T(\mathfrak{F}_{xy}(t), \mathfrak{F}_{yz}(s)), \quad \text{para cada } x, y, z \in X \text{ y cada } t, s \in \mathbb{R},$$

donde la función $T : \mathbb{I}^2 \rightarrow \mathbb{I}$ es lo que hoy conocemos como *norma triangular* (también llamada *t-norma*, véase la subsección 1.2.4). En [93], E. Nishiura ya intuyó las propiedades que debía verificar una función como ésta considerando la función:

$$T_{\mathfrak{F}}(a, b) = \inf(\{\mathfrak{F}_{xz}(t+s) : \mathfrak{F}_{xy}(t) \geq a, \mathfrak{F}_{yz}(s) \geq b\}), \quad \text{para cada } a, b \in \mathbb{I},$$

pero fue Menger el primero en darles nombre debido a su decisiva intervención en la desigualdad triangular. La teoría probabilística ha sido después desarrollada con gran éxito por B. Schweizer y A. Sklar como comentaremos en el subsección 1.4.1.

La aproximación difusa al concepto de distancia parte de la idea de que ésta no es realmente un número que podamos encontrar o determinar, sino que sólo podemos hacer diferentes afirmaciones acerca de ésta asignándoles números entre 0 y 1 a modo de probabilidad. De esta forma, en 1975, I. Kramosil y J. Michálek introdujeron en [74] la siguiente noción de *espacio métrico difuso*.

Un *espacio métrico difuso* (en el sentido de Kramosil y Michálek) es una terna (X, R, S) donde R es un conjunto difuso sobre el producto cartesiano $X \times X \times \mathbb{R}$ con función característica $f_R : X \times X \times \mathbb{R} \rightarrow \mathbb{I}$ y $S : \mathbb{I}^2 \rightarrow \mathbb{I}$ es una función real binaria medible tal que $S(1, 1) = 1$, de manera que se verifican las siguientes propiedades, para cada $x, y, z \in X$.

- $f_R(x, y, t) = 0$ si $t \leq 0$.
- $f_R(x, y, t) = 1$ para cada $t > 0$ si, y sólo si, $x = y$.
- $f_R(x, y, t) = f_R(y, x, t)$ para cada $t \in \mathbb{R}$.
- $f_R(x, z, t+s) \geq S(f_R(x, y, t), f_R(y, z, s))$.
- $f_R(x, y, \cdot) : \mathbb{R} \rightarrow \mathbb{I}$ es una función continua a la izquierda y no decreciente tal que $\lim_{t \rightarrow \infty} f_R(x, y, t) = 1$.

Las tres primeras condiciones responden a la idea de positividad, identidad y simetría que debe poseer cualquier noción de distancia. La cuarta es una versión muy débil de la desigualdad

triangular con la misma interpretación que hicimos en el caso de los espacios métricos estadísticos. La quinta condición es una forma de expresar la idea de finitud de la distancia, a la vez que, si queremos que $f_R(x, y, t)$ represente la probabilidad de que la distancia entre x e y sea (estrictamente) menor que t , entonces la función $f_R(x, y, \cdot)$ debe ser continua a la izquierda. Evidentemente, se podría suponer que $f_R(x, y, t)$ representa la probabilidad de que la distancia entre x e y sea menor o igual que t , pero entonces $f_R(x, y, \cdot)$ debería suponerse continua a la derecha. Por otro lado, que $f_R(x, y, \cdot)$ sea no decreciente significa que si $t \leq s$, entonces la probabilidad de que la distancia entre x e y sea menor que t debe ser, intuitivamente, menor o igual que la probabilidad de que la distancia entre x e y sea menor que s .

Suponiendo condiciones adicionales sobre la función S , Kramosil y Michálek demostraron que su noción de espacio métrico difuso es equivalente a la de espacio métrico estadístico con tal de identificar $\mathfrak{F}_{xy}(t) = f_R(x, y, t)$. De hecho, bajo la desigualdad triangular en el sentido de Menger, es posible asociarle una variable aleatoria $R(\cdot; x, y)$ sobre un cierto espacio de probabilidad a cada par de puntos x e y de manera que $f_R(x, y, t) = P(\{\omega : R(\omega; x, y) < t\})$.

La noción de espacio métrico difuso planteada por Kramosil y Michálek ha ido perfilándose con el paso del tiempo, aceptándose hoy que, para trabajar en estos espacios y poder desarrollar buenas propiedades con aplicaciones prácticas interesantes, éstos llevan involucrada una t -norma continua (véanse [42, 47, 39]). De esta forma, daremos en la Definición 1.4.5 una versión más moderna de estos espacios.

A partir de esta definición y apoyándose en el creciente desarrollo de la teoría difusa, se han planteado muchas posibilidades para determinar lo más precisamente posible la noción de métrica difusa. En 1994, A. George y P.V. Veeramani [42] modificaron de una forma ligera (pero atractiva) los axiomas introducidos por Kramosil y Michálek y determinaron una subclase de espacios métricos difusos $M : X \times X \times]0, \infty[\rightarrow \mathbb{I}$ que, abarcando una amplia variedad de ellos (al menos, los más utilizados en la práctica), poseen mejores propiedades que los anteriores. En su célebre trabajo, demostraron que las bolas abiertas

$$B(x, r, t) = \{y \in X : M_{xy}(t) > 1 - r\}, \quad \text{donde } x \in X \text{ y } r, t \in]0, 1[$$

definen una topología Hausdorff que verifica el primer axioma de numerabilidad. Además, modificaron las nociones de *convergencia de sucesiones* y de *sucesión de Cauchy* para conseguir que \mathbb{R} , dotado de la métrica difusa estándar, fuese completo. Por último, utilizando bolas cerradas, demostraron una versión para espacios métricos difusos del *teorema de Baire* (que afirma que la intersección numerable de subconjuntos abiertos densos de un espacio métrico difuso es también densa).

Posteriormente, en 2004, J.H. Park introdujo en [97] la noción de *espacio métrico difuso intuicionista* basándose tanto en los espacios de George y Veeramani como en la idea de *conjunto difuso intuicionista* planteada por K. Atanassov en [3] (en realidad, Atanassov ya anunció este concepto tres años antes en [4]). Los conjuntos intuicionistas parten de la idea de que, en oca-

siones, nos encontramos ante fenómenos en los que es difícil determinar si ocurre un suceso o su contrario, y cabe plantearse la posibilidad de que no podamos decidir con certeza entre una posibilidad y su contraria. De esta manera, la suma de la certidumbre y de la incertidumbre que poseemos de que ocurra un cierto fenómeno puede no llegar a la unidad. Así, a un conjunto se le asocian dos funciones características que, sumadas, dan una cantidad menor o igual que uno. Trasladando esta idea a espacios métricos difusos, es posible considerar toda una teoría difusa en la que la suma de la certidumbre y de la incertidumbre sea menor o igual que uno. Aunque, a la vista de sus aplicaciones prácticas, Park prefirió utilizar la definición de George y Veeramani, para nuestros propósitos, en esta Memoria presentamos en la Definición 1.4.13 la versión de los espacios intuicionistas dada por F. Castro-Company y S. Romaguera en [16] (ver también [2]) que es más cercana a la de Kramosil y Michálek y, por tanto, atrapa a una mayor variedad de espacios (que, por supuesto, incluyen a los de Park).

La noción de conjunto difuso ha dado lugar no sólo a conjuntos difusos intuicionistas, sino a toda una variedad de tipos de conjuntos relacionados con la incertidumbre: conjuntos \mathcal{L} -difusos, conjuntos \mathcal{L} -difusos intuicionistas, conjuntos difusos intuicionistas valuados en intervalos, conjuntos probabilísticos, etc. G. Deschrijver y E.E. Kerre hicieron en [27] un completo estudio de los mismos y analizaron las interrelaciones existentes entre las diferentes nociones, llegando a la conclusión de que muchos de estos conceptos son equivalentes, y entre los que no lo son existen ciertas implicaciones.

Volviendo al tema métrico, es de justicia comentar que, a lo largo de la Historia, multitud de autores han expuesto otras muchas concepciones de *espacio métrico difuso* desde diferentes puntos de vista y que no vamos a citar aquí para no extender aún más la Memoria ([36, 42, 67, 74, 82, 81, 111, 26, 116, 117, 89, 39]). Éstas han ido poco a poco contribuyendo a perfilar la noción actual que podemos tener de este concepto. En realidad, da la sensación de que no existe una definición comúnmente aceptada sino que cada investigador/a utiliza la definición que mejor se ajusta a la clase de experimentos que lleva a cabo o que mejor describe los fenómenos que observa. Sin embargo, aunque todas la llevan aparejada implícitamente, pocas de estas concepciones (a destacar, quizás, la noción introducida por O. Kaleva y S. Seikkala en [67] en 1984) utilizan la noción de *número difuso* para expresar la distancia entre dos puntos del espacio. Cuando esto se hace, la desigualdad triangular adopta formas que no son nada intuitivas. Por ejemplo, la definición de Kaleva y Seikkala es tan general que la desigualdad triangular que involucra no está, en absoluto, al alcance de la intuición. Y trabajar en este tipo de espacios se vuelve mucho más complicado.

Aunque determinar una idea precisa y manejable de espacio métrico difuso es un tema de especial interés dentro de la teoría difusa, existen otros dos problemas íntimamente relacionados con el anterior sobre los que también se ha investigado y se investiga especialmente en la actualidad. El primero es el de introducir una *topología difusa* sobre un conjunto que posea propiedades razonables. Ya hemos comentado que, respecto de este problema, George y Veera-

mani describieron una topología Hausdorff que verifica el primer axioma de numerabilidad en sus espacios. Anteriormente, Michálek ya había abordado este problema en [84]. En [26], Deng estudió la topología de un espacio difuso pseudo-métrico. Muy recientemente, Shi ha estudiado en [117] condiciones de regularidad y de normalidad sobre los espacios topológicos difusos de tipo (L, M) que él mismo había introducido en [116].

Otro de los temas que más literatura científica ha generado en los últimos tiempos es el *teorema del punto fijo*. Debido a su amplia aplicabilidad, la *teoría de puntos fijos* es una rama importante del Análisis Funcional no lineal. Numerosas preguntas en Economía, Física, Química y Biología nos llevan a ecuaciones diferenciales e integrales no lineales, y si uno hace caso omiso de la forma concreta de estas ecuaciones, a menudo se pueden reducir a ecuaciones abstractas en las que interviene cierto operador. El teorema del punto fijo más conocido e importante es el *principio de contracción de Banach*, que asegura que cada contracción de un espacio métrico completo en sí mismo tiene un único punto fijo. Este teorema aparece en forma explícita en la tesis de Banach en 1922, en la que se utilizó para establecer la existencia de solución de una ecuación integral. La sencillez de su prueba y la posibilidad de alcanzar el punto fijo mediante el uso de sucesivas aproximaciones hacen de este resultado una herramienta muy útil en el Análisis y en Matemática Aplicada. La gran importancia del principio de Banach, y la razón por la que es posiblemente el teorema de punto fijo más citado en todo el Análisis, radica en el hecho de que su prueba contiene elementos de importancia fundamental para el tratamiento teórico y práctico de ecuaciones matemáticas.

Durante los últimos cincuenta años, la teoría métrica del punto fijo, en la que las condiciones geométricas sobre los espacios subyacentes y sobre las aplicaciones juegan un papel fundamental, ha sido un área floreciente de investigación para la comunidad matemática. La teoría clásica ha sido desarrollada por muchas personas, entre ellas los conocidos matemáticos L.E.J. Brouwer, S. Banach y J. Schauder. Este campo sufrió un nuevo impulso en 1965 debido a los teoremas de punto fijo para aplicaciones contractivas descubiertos independientemente por F. Browder, D. Göhde y W.A. Kirk. Podemos encontrar también versiones posteriores del teorema de Banach debidas a Krasnoselskii y Zabreiko [128], Edelstein [33], Browder [13] y Caristi [15].

El teorema de Banach, junto con sus posteriores desarrollos, posee multitud de aplicaciones, aunque la principal (la razón por la que fue descubierto por Banach) reside en su capacidad de resolver toda clase de ecuaciones (diferenciales, integrales, matriciales, etc.). Sin embargo, este teorema está relacionado con muchos otros ámbitos de la Matemática: sistemas de funciones iteradas, el método de Newton, la métrica de Hilbert, el teorema de la función implícita, desigualdades variacionales, ecuaciones semilineales de tipo elíptico, teoremas sobre funciones en espacios de Hilbert, el teorema de Cauchy-Kowalevsky, etc. (en [12] se describen algunas de estas aplicaciones). Aunque se han desarrollado muchos resultados en este campo, aún quedan abiertas unas cuantas preguntas que residen en el corazón mismo de la teoría, y hay muchas preguntas sin respuesta acerca de los límites hacia los que puede extenderse la misma.

A partir de la demostración del teorema de Banach, el Análisis y la Topología han abierto el campo de la búsqueda de teoremas del punto fijo en muy diversos ambientes y contextos. En este campo y dentro de la teoría difusa se han desarrollado infinidad de enunciados (véanse, por ejemplo [21, 39, 42, 47, 55]). Hoy en día hay una corriente muy importante de estudio para determinar puntos fijos de aplicaciones en varias variables y, más en concreto, de *puntos de coincidencia* (dobles, triples, cuádruples e incluso en dimensión superior) en espacios métricos completos parcialmente ordenados (véanse los resultados de Gnana Bhaskar y Lakshmikantham [10], de Lakshmikantham y Ćirić [76], de Berinde y Borcut [8], de Karapinar y Luong [69] y de Berzig y Samet [9] en dimensión arbitraria). Pero también se pueden encontrar trabajos recientes en el contexto difuso. Por ejemplo, Zhu y Xiao [133] y Hu [61] mostraron teoremas de existencia y unicidad de puntos fijos dobles para contracciones en espacios métricos difusos y Fang [40] demostró algunos teoremas de punto fijo común bajo ϕ -contracciones para aplicaciones compatibles y débilmente compatibles definidas sobre espacios métricos probabilísticos de Menger. Cabe también destacar, sobretodo por su complejidad, el reciente trabajo de J.Z. Xiao, X.H. Zhu y X. Jin [132] acerca de teoremas de punto fijo en espacios métricos difusos en el sentido de Kaleva y Seikkala.

Con las ideas anteriores en mente, la presente Memoria está dividida en cuatro Capítulos. En el primer Capítulo concretamos cada una de las diferentes concepciones de *espacio métrico* (difuso o probabilístico) que hemos comentado y que manejaremos a continuación. Es por ello que se trata de una parte dedicada puramente a los preliminares, donde no aportamos demostraciones pero sí reflejamos la propiedades más interesantes de cada definición y confiamos en dar suficientes referencias para el/la lector/a interesado/a en cada tema.

En el segundo Capítulo, abordamos el estudio de las propiedades comunes que subyacen a cada una de las teorías mencionadas con anterioridad. Para ello, introducimos nuestra propia versión de lo que entendemos que debería ser un *espacio métrico difuso*, que a cada par de puntos del espacio le asocie un número difuso y que verifique, de una forma natural, las propiedades de una métrica real. Para ello, necesitamos interpretar un número difuso como un par de funciones de distribución de tipo distancia (que forman el conjunto Δ^+). Esta interpretación nos permite identificar completamente los conjuntos $\Delta^+ \times \Delta^+$ y \mathcal{F} , formado por todos los números difusos que alcanzan la condición de normalidad en cero, de tal manera que nos es sencillo trasladar la métrica de Lévy sobre Δ^+ a \mathcal{F} . A continuación, realizamos un estudio detallado de las funciones triangulares sobre \mathcal{F} relacionándolas con sus correspondientes funciones triangulares asociadas sobre $\Delta^+ \times \Delta^+$. De esta manera, en la Definición 2.2.1 presentamos el siguiente concepto.

Un **espacio métrico difuso** (abreviadamente, un *EMD*) es una terna (X, \mathbf{F}, ν) donde X es un conjunto no vacío, $\mathbf{F} : X \times X \rightarrow \mathcal{F}$ es una aplicación que a cada par de puntos de X le asocia un número difuso de \mathcal{F} y $\nu : \mathcal{F} \times \mathcal{F} \rightarrow \mathcal{F}$ es una función triangular sobre \mathcal{F} , de manera que se verifican las siguientes propiedades.

- Para cada $x \in X$ se tiene que $\mathbf{F}_{xx} = \bar{0}$.
- Si $x, y \in X$ verifican que $\mathbf{F}_{xy} = \bar{0}$, entonces $x = y$.
- Para cada $x, y \in X$ se tiene que $\mathbf{F}_{yx} = \mathbf{F}_{xy}$.
- Para cada $x, y, z \in X$ se tiene que $\mathbf{F}_{xz} \leq v(\mathbf{F}_{xy}, \mathbf{F}_{yz})$.

Somos conscientes de las limitaciones de esta definición pero, como veremos a lo largo del primer Capítulo, esta concepción es capaz de atrapar muchas de las nociones de espacio métrico difuso o probabilístico que se han dado hasta la fecha. Además, pensamos que tiene la ventaja de que plantea una versión más natural de distancia utilizando números difusos, y la desigualdad triangular es exactamente la correspondiente a la desigualdad triangular que todos tenemos en mente. Concretamente, en los Teoremas 2.2.5, 2.2.17 y 2.2.22 damos caracterizaciones para determinar cuándo estructuras anteriores dan lugar a EMD en nuestro sentido. Por ejemplo, el siguiente teorema demuestra que cada espacio métrico probabilístico es un espacio métrico difuso.

Teorema 2.2.5. Sea X un conjunto, sea τ una función triangular sobre Δ^+ , sea $d : X \times X \rightarrow \Delta^+$ una aplicación y definamos, para cada $x, y \in X$:

$$\mathbf{F}_{xy} = \begin{cases} 0, & \text{si } t < 0, \\ 1 - d_{xy}(t), & \text{si } t \geq 0. \end{cases}$$

Entonces las siguientes afirmaciones son equivalentes.

- (X, d, τ) es un espacio métrico probabilístico.
- (X, \mathbf{F}, v) es un EMD, donde v es la función triangular sobre \mathcal{F} asociada a $\tau' \times \tau$, sea cual sea la función triangular τ' sobre Δ^+ .
- Existe una función triangular τ' sobre Δ^+ tal que (X, \mathbf{F}, v) es un EMD, donde v es la función triangular sobre \mathcal{F} asociada a $\tau' \times \tau$.

En el siguiente teorema caracterizamos los espacios métricos difusos intuicionistas como una subclase de los espacios métricos difusos que hemos introducido.

Teorema 2.2.17. Sea X un conjunto, sea $*$ una t-norma continua, sea \diamond una t-conorma continua y sea v la función triangular sobre \mathcal{F} correspondiente a $\tau_{\diamond} \times \tau_*$.

Supongamos que para cada $x, y \in X$ existen conjuntos difusos $M_{xy}, N_{xy} : [0, \infty[\rightarrow [0, 1]$ y $\mathbf{F}_{xy} : \mathbb{R} \rightarrow [0, 1]$ relacionados de la siguiente forma:

$$M_{xy}(t) = 1 - \mathbf{F}_{xy}(t) \quad \text{y} \quad N_{xy}(t) = \mathbf{F}_{xy}(-t) \quad \text{para cada } t \in [0, \infty[.$$

Entonces las siguientes afirmaciones son equivalentes.

- (a) $(X, M, N, *, \diamond)$ es un espacio métrico difuso intuicionista.
- (b) (X, \mathbf{F}, v) es un EMD normalizado que verifica las siguientes propiedades.
 - (b.1) Si $x, y \in X$, entonces $\mathbf{F}_{xy}^- = \epsilon_0$ si, y sólo si, $x = y$, lo cual ocurre si, y sólo si, $\mathbf{F}_{xy}^+ = \epsilon_0$.
 - (b.2) $\mathbf{F}_{xy}(t) *' \mathbf{F}_{yz}(s) \geq \mathbf{F}_{xz}(t+s)$ para cada $x, y, z \in X$ y cada $t, s \in [0, \infty[$ (donde $*'$ es la t -conorma inducida por $*$).
 - (b.3) $\mathbf{F}_{xy}(-t) \diamond \mathbf{F}_{yz}(-s) \geq \mathbf{F}_{xz}(-t-s)$ para cada $x, y, z \in X$ y cada $t, s \in [0, \infty[$.

En tal caso, (X, \mathbf{F}, v) es el EMD generado por $(X, M, N, *, \diamond)$.

Y en el siguiente resultado estudiamos la relación entre los espacios métricos probabilísticos intuicionistas y nuestros espacios métricos difusos.

Teorema 2.2.22. Sea X un conjunto, sea τ una función triangular sobre Δ^+ , $\tilde{\tau}$ una función triangular sobre ∇^+ y sea v la función triangular sobre \mathcal{F} asociada a $\tau \times \tilde{\tau}$. Sean $d : X \times X \rightarrow \Delta^+$, $\tilde{d} : X \times X \rightarrow \nabla^+$ y $\mathbf{F} : X \times X \rightarrow \mathcal{F}$ tres aplicaciones relacionadas de la siguiente forma, para cada $x, y \in X$ y cada $t \in \mathbb{R}$:

$$\mathbf{F}_{xy}(t) = \begin{cases} 1 - d_{xy}(-t), & \text{si } t < 0, \\ \tilde{d}_{xy}(t), & \text{si } t \geq 0. \end{cases}$$

Entonces $(X, d, \tau, \tilde{d}, v)$ es un espacio métrico probabilístico intuicionista si, y sólo si, (X, \mathbf{F}, v) es un EMD conormalizado tal que

$$\mathbf{F}_{xy}|_{]-\infty, 0[} = \bar{0}|_{]-\infty, 0[} \Leftrightarrow \mathbf{F}_{xy}|_{]0, \infty[} = \bar{0}|_{]0, \infty[}.$$

Como consecuencia, cualquier espacio métrico probabilístico intuicionista es un EMD. El siguiente diagrama resume las principales conexiones que demostramos a lo largo del segundo Capítulo, dejando la puerta abierta a la inclusión de nuevas relaciones y nuevos espacios.

En tercer Capítulo ponemos ejemplos de teoremas de punto fijo en espacios métricos difusos intuicionistas y en espacios normados difusos. En primer lugar, damos una demostración completa y correcta de un teorema de punto fijo que fue enunciado en primer lugar por R. Saadati en

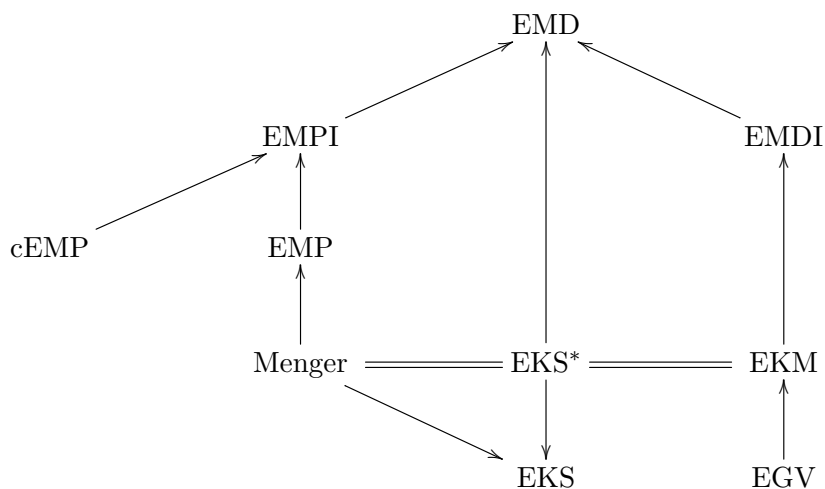


Figura 1: Conexiones entre las diferentes teorías.

[105] como extensión de otro resultado demostrado por C. Alaca, D. Turkoglu y C. Yildiz en [2] referente a la existencia y unicidad de puntos fijos de aplicaciones contractivas en espacios métricos \mathfrak{L} -difusos intuicionistas. En concreto, explicamos los preliminares necesarios para entender este resultado e indicamos el punto exacto en el que fallaba la demostración dada por Saadati, y cómo resolver sin problema este detalle. En concreto, en el Teorema 3.1.3 demostramos el siguiente resultado.

Sea $(X, \mathfrak{M}, \mathfrak{J})$ un espacio \mathfrak{L} -difuso completo de manera que la t -norma \mathfrak{J} es de tipo Hadžić. Supongamos que si $\mathfrak{M}(x, y, t) = C$ para cada $t > 0$ entonces $C = 1_{\mathfrak{L}}$. Entonces cualquier B -contracción $A : X \rightarrow X$ posee un único punto fijo.

A continuación, demostramos la siguiente versión de un teorema anterior de M. Eshaghi Gordji, H. Baghani y Y.J. Cho [37] para espacios normados difusos (END).

Teorema 3.2.1. Sea $(X, \mu, *)$ un END completo de manera que $*$ es una t -norma de tipo Hadžić tal que $a * b \geq ab$ para cada $a, b \in \mathbb{I}$. Sean $F : X \times X \rightarrow X$ y $g : X \rightarrow X$ dos aplicaciones tales que $F(X \times X) \subseteq g(X)$, g es continua y F conmuta con g . Sea \sqsubseteq un orden parcial sobre X tal que F verifica la propiedad g -monótona mezclada y de manera que existe un número real positivo $k \in]0, 1[$ que cumple la desigualdad:

$$\mu(F(x, y) - F(u, v), kt) \geq [\mu(gx - gu, t)]^{1/2} * [\mu(gy - gv, t)]^{1/2},$$

siempre que $gx \sqsubseteq gu$ y $gy \sqsupseteq gv$. Supongamos que:

- (a) F es continua, o bien
- (b) (X, \sqsubseteq) posee la propiedad g -monótona secuencial.

Si existen dos puntos $x_0, y_0 \in X$ tales que $gx_0 \sqsubseteq F(x_0, y_0)$ y $gy_0 \sqsupseteq F(y_0, x_0)$, entonces F y g poseen algún punto de coincidencia doble (es decir, existen $x, y \in X$ tales que $gx = F(x, y)$ y $gy = F(y, x)$).

En el cuarto Capítulo damos una particular versión de una técnica de regresión que utiliza datos de entrada reales y datos de salida difusos. La *regresión estadística* es un método de estimación muy útil en áreas científicas tales como la Economía, la Química, las Ciencias Medioambientales, Biomedicina, etc. La identificación del modelo consiste en determinar, de entre varios modelos, aquél que explique mejor el comportamiento del sistema. Esta relación funcional entre las variables independiente (o explicativa) y dependiente (o respuesta) se obtiene suponiendo que la diferencia entre ellas se debe a errores aleatorios. Las técnicas de identificación clásicas suponen que las observaciones de las variables dependientes e independientes son valores numéricos y la técnica de mínimos cuadrados es la que usualmente nos permite encontrar los coeficientes del modelo basándonos en estas observaciones. En tal caso, tanto los datos como los coeficientes que se obtienen en el modelo de regresión son números reales.

Sin embargo, como hemos comentado anteriormente, en el mundo real nos encontramos antes datos imperfectos, es decir, conocimiento vago, impreciso, incierto, ambiguo, inexacto, o probabilístico por naturaleza. El razonamiento y pensamiento humano frecuentemente se encuentran ante información de este tipo, probablemente originada por la inexactitud inherente de los conceptos humanos y del razonamiento basados tanto en experiencias similares pero no idénticas como en experiencias anteriores. La teoría de conjuntos difusos o lógica difusa, como su nombre indica, es una lógica alternativa a la lógica clásica que pretende introducir un grado de vaguedad en las cosas que califica. En tal caso, el análisis estadístico difuso trata de resolver algunos de los diferentes problemas que surgen (véanse [23, 110, 94]).

Formalmente, cualquier función evaluada en el intervalo $\mathbb{I} = [0, 1]$ es un conjunto difuso. En los problemas reales, la clase formada por los *números difusos* (que consiste en las funciones evaluadas en \mathbb{I} que son semicontinuas superiormente, de soporte compacto y toman el valor 1 en algún punto) es lo suficientemente rica como para permitir analizar la mayoría de aplicaciones (véase [29]). Sin embargo, esta clase es aún demasiado general, y la mayoría de investigadores/as utilizan números difusos de formas simples, tales como los números difusos triangulares, trapezoidales o los denominados *números difusos LR*, pues éstos satisfacen las necesidades de los problemas de modelización difusa y son mucho más sencillos de manejar y de determinar.

En este contexto, un problema sencillo, pero importante, es el siguiente. En muchas situaciones de la vida cotidiana, la variable independiente puede ser observada de una forma precisa (por ejemplo, el tiempo, en años) y la variable respuesta, en lugar de tomar valores exactos,

se describe por valores aproximados (por ejemplo, cuando decimos “alrededor de 80 *gr/l*” o “*aproximadamente 500 euros*”). Esta imprecisión puede ser modelizada utilizando números difusos con formas simples, por ejemplo números difusos triangulares, trapezoidales o LR. En este caso, es posible tratar de determinar modelos separados para los centros y las amplitudes (véase [24, 41, 103]).

Bajo el término de *Regresión Difusa* se engloba a aquellos métodos de regresión en los que alguna de sus variables y/o alguno de sus coeficientes son números difusos. La Regresión Difusa tiene su origen en los métodos basados en *programación lineal* conocidos como Regresión Posibilística, donde destacan las aportaciones de H. Tanaka. Tras la aparición de este modelo y debido a algunos inconvenientes que presenta, se propusieron métodos basados en el enfoque de *mínimos cuadrados*. En la actualidad se está aplicando una técnica conocida como *máquinas de soporte vectorial* a la resolución de problemas de regresión. Esquemáticamente los métodos tradicionales que se han venido utilizando en los problemas de regresión difusa para la estimación de parámetros, y que a continuación analizamos más detalladamente, son los siguientes:

- Métodos de programación lineal.
- Métodos de mínimos cuadrados.
- Máquinas de soporte vectorial.

Puede considerarse que la Regresión Difusa surge en 1982 cuando H. Tanaka y sus colaboradores publican los primeros intentos para dar solución a este problema basándose en el principio posibilístico. Este principio trata de determinar los coeficientes del modelo de forma que tengan la menor incertidumbre posible. Tanaka y sus colaboradores propusieron en [119] una formulación del análisis de regresión lineal difuso en términos de este principio y determinaron los parámetros difusos involucrados formulando un problema de programación lineal. Este método fue denominado *Regresión Posibilística*. Desde entonces, Tanaka y sus colaboradores, así como otros autores, han profundizando en el análisis de regresión difuso basado en el enfoque posibilístico. Aunque se trata de trabajos significativos y de gran valor en el ámbito de la Regresión Difusa, esta metodología ha recibido numerosas críticas (véanse [91, 28, 124, 68, 71]). Dos de las principales críticas que se hacen a este enfoque son las siguientes: por un lado, es muy sensible a valores atípicos (también llamados extremos u *outliers*) y, por otro lado, algunos coeficientes tienden a tomar valores reales precisos (no difusos) debido a las características de la programación lineal. Para abordar estas cuestiones, Nasrabadi y Nasrabadi [91] propusieron un enfoque basándose en programación matemática.

Otro problema que se puede observar al respecto de la Regresión Posibilística es que, cuando los coeficientes son números difusos, la amplitud de la variable estimada se vuelve mayor que los valores de los incrementos de la variables explicativas, incluso cuando las amplitudes de las respuestas estimadas disminuyen, o cuando se introducen más observaciones en el modelo. Esto

contradice la intuición. Para resolver este problema, Diamond [28], Wu y Tseng [124] y Kao y Chyu [68] consideraron coeficientes numéricos para describir la relación borrosa entre la variable difusa de respuesta y las variables difusas (o numéricas) explicativas. Todos ellos utilizaron el concepto de mínimos cuadrados para determinar los coeficientes de la regresión. Kao y Chyu utilizaron un proceso de resolución en dos etapas para determinar los coeficientes numéricos utilizando el criterio de Kim y Bishu [71], el cual, a su vez, es una modificación del criterio de análisis de regresión lineal difusa propuesto por Tanaka y otros en [120]. Aunque este criterio ha sido utilizado por varios autores, posee el inconveniente de que si las variables difusas observada y estimada no se intersecan mutuamente, entonces el error de estimación permanece constante.

En el caso de la regresión posibilística lineal, Serrurier y Prade [114] propusieron un método basado en redes neuronales, que se ilustró con datos de Ciencias de la Naturaleza, y Chen y Chen [17] propusieron un método de programación no lineal.

La *máquina de soporte vectorial* (SVM, del inglés *Support Vector Machine*) es una técnica para resolver problemas de reconocimiento de patrones, que se utiliza para la clasificación y el análisis de regresión. Está basada en el principio de minimización del riesgo estructural (SRM). Este principio de inducción se fundamenta en el hecho de que el error de generalización está acotado por la suma del error de entrenamiento y un término de intervalo de confianza que depende de la dimensión de Vapnik-Chervonenkis [121]. Así, el objetivo del SRM es minimizar el límite superior del error de generalización en vez del principio de minimización de riesgo empírico. Cuando se aplica una máquina de soporte vectorial a problemas de estimación en regresión, nos solemos referir a ella como *regresión de soporte vectorial*. La regresión de soporte vectorial es actualmente un tema de investigación muy activo para resolver problemas no lineales con datos imprecisos cuando no se requiere interpretabilidad lingüística. La regresión de soporte vectorial es una herramienta de regresión no lineal, regularizada y no paramétrica que ha sido aplicada a problemas de control óptimo, a obtener predicciones en series temporales, al análisis de la regresión por intervalos, a la determinación de las estructuras iniciales de redes neuronales difusas, etc. (véanse Hong y Hwang [60]; Chiang y Hao [19]; Juang y Hsieh [64]; Khemchandani y otros [70]; Yang y otros [127]).

Desde el punto de vista estadístico, el método de mínimos cuadrados para estimar las fronteras de una variable difusa suele ser más apropiado y es, a menudo, más sencillo de computar que el método de programación lineal tradicional (Wu y Tseng [124]). Recientemente, se han determinado modelos de regresión lineal simple en términos de los momentos de los elementos aleatorios involucrados en el problema, en el caso de que los datos de entrada y salida sean números difusos trapezoidales (González-Rodríguez y otros [46]). Näther [92] y Krätschmer [75] también han estudiado problemas de mínimos cuadrados que involucran datos difusos.

Existen modelos de regresión, desarrollados en trabajos recientes, que utilizan números difusos en la variable respuesta y números reales en la variable de entrada. Por ejemplo, Coppi y otros [24] estudiaron la dependencia de una variable de respuesta difusa de tipo LR a partir

de variables explicativas reales proponiendo un procedimiento iterativo de estimación por mínimos cuadrados que conlleva condiciones de no negatividad. Para superar las condiciones de no negatividad, en 2010, Ferraro y otros [41] propusieron un modelo de regresión lineal para una variable de respuesta imprecisa y varias variables explicativas reales analizando el problema dentro del contexto de los modelos de regresión múltiple tradicionales y de regresión multivariante. También Chen y Hsueh [18] analizaron un modelo de regresión difusa que trata de minimizar la suma de las distancias cuadráticas medias entre las respuestas estimadas y observadas utilizando algunos α -cortes. En general, cuando se trabaja con variables aleatorias difusas, el análisis de la regresión difusa para determinar la relación entre las variables explicativa y respuesta se vuelve más complejo y se encuentran muchas dificultades para determinar soluciones óptimas, especialmente en el caso de problemas no lineales.

En el último Capítulo de la presente Memoria ofrecemos una nueva aproximación al problema del análisis de las relaciones de regresión lineal y no lineal considerando datos de entrada reales y datos de salida difusos triangulares. Sobretudo, por tres motivos: porque pensamos que se trata de un caso que se presenta en muchos contextos de la vida real, porque entendemos que la técnica de regresión por mínimos cuadrados es un método de estimación suficientemente simple y muy conocida (utilizada por muchos/as investigadores/as en campos muy diversos) y porque la formulación que proponemos posee una expresión matemática fácil de entender e igualmente fácil de programar, que puede adaptarse a estudios estadísticos aplicándose con éxito. Para ello, utilizaremos diferentes distancias entre números difusos y enunciaremos un criterio de mínimos para encontrar la solución óptima. A continuación, generalizaremos este caso ajustando otros modelos para las amplitudes. Esta nueva técnica no sólo puede ser utilizada con números triangulares, es muy sencilla de aplicar y computar, puede ser aplicada en diferentes contextos y, un estudio comparativo con otras técnicas desarrolladas por otros/as investigadores/as demuestra que la técnica propuesta produce un error total similar o incluso menor al obtenido por otros procedimientos más complejos. A continuación, resumimos todos los casos que hemos estudiado en el siguiente enunciado, en el caso en el que las amplitudes puedan considerarse constantes.

Teorema 4.2.11. Sean $\underline{X}, \mathcal{Y}$ dos variables tales que \underline{X} es una variable aleatoria real e \mathcal{Y} es una variable aleatoria difusa que se relacionan a través de la ecuación

$$\mathcal{Y} = f(\underline{X}, \underline{a}) + (\varepsilon^I / \varepsilon^c / \varepsilon^S).$$

Sea $\{\underline{X}_i, \mathcal{Y}_i\}_{i=1}^n$ una muestra aleatoria simple obtenida de $(\underline{X}, \mathcal{Y})$, siendo $\mathcal{Y}_i = (Y_i - I_i / Y_i / Y_i + S_i)$. Sea \widehat{Y}_x^c un modelo ajustado para los centros a partir de $\{\underline{X}_i, Y_i\}_{i=1}^n$ y consideremos el problema de regresión (en $I, S > 0$) para alguna distancia d :

$$(P) \begin{cases} \Phi_{I,S} : M \rightarrow \mathcal{T}, & \Phi_{I,S}(\underline{x}) = (\widehat{Y}_x - I / \widehat{Y}_x / \widehat{Y}_x + S), \\ E(I, S) = \sum_{i=1}^n d(\mathcal{Y}_i^*, \Phi_{\underline{X}_i}^*) & \text{mínimo.} \end{cases}$$

- Utilizando $d_1(A, B) = \|A - B\|_1 = \int_{\mathbb{R}} |A(t) - B(t)| dt$, el problema (P) posee como única solución los valores estimados

$$\widehat{I} = \text{Mediana}(I_1, \dots, I_n) \quad \text{y} \quad \widehat{S} = \text{Mediana}(S_1, \dots, S_n).$$

- Utilizando $d_1^2(A, B) = \left(\int_{-\infty}^0 |A(t) - B(t)| dt \right)^2 + \left(\int_0^{+\infty} |A(t) - B(t)| dt \right)^2$, el problema (P) posee como única solución los valores estimados

$$\widehat{I} = \text{Media}(I_1, \dots, I_n) \quad \text{y} \quad \widehat{S} = \text{Media}(S_1, \dots, S_n).$$

- Utilizando $d_2(A, B) = \|A - B\|_2^2 = \int_{\mathbb{R}} (A(t) - B(t))^2 dt$, el problema (P) posee una única solución. Siguiendo la notación (4.3), los estimadores de los parámetros \widehat{I} y \widehat{S} son las únicas soluciones no negativas de las ecuaciones cúbicas:

$$\left\{ \begin{array}{l} 2x^3 \sum_{i=r+1}^k \frac{m_i}{I'_i} + x^2 \left(\sum_{i=1}^r m_i - 2 \sum_{i=r+1}^k m_i \right) - \sum_{i=1}^r m_i (I'_i)^2 = 0 \\ \text{en el intervalo } [I'_r, I'_{r+1}] \end{array} \right\}_{r=1}^{r=k-1},$$

$$\left\{ \begin{array}{l} 2x^3 \sum_{i=r+1}^{k'} \frac{m'_i}{S'_i} + x^2 \left(\sum_{i=1}^r m'_i - 2 \sum_{i=r+1}^{k'} m'_i \right) - \sum_{i=1}^r m'_i (S'_i)^2 = 0 \\ \text{en el intervalo } [S'_r, S'_{r+1}] \end{array} \right\}_{r=1}^{r=k'-1}.$$

respectivamente.

- Utilizando $d_{\infty}(A, B) = \sup(\{|A(t) - B(t)| : t \leq 0\}) + \sup(\{|A(t) - B(t)| : t \geq 0\})$, el problema (P) posee solución. Siguiendo la notación (4.3), el estimador del parámetro \widehat{I} es el valor I_1, I_2, \dots, I_n donde se alcanza el mínimo de los siguientes k números reales:

$$\left\{ n - m_j - \frac{1}{I'_j} \sum_{i=1}^{j-1} m_i I'_i - I'_j \sum_{i=j+1}^k \frac{m_i}{I'_i} \right\}_{j=1}^k,$$

y el estimador del parámetro \widehat{S} es el valor S_1, S_2, \dots, S_n donde se alcanza el mínimo de un conjunto numérico similar.

En la parte final del Capítulo, aplicaremos este método a un interesante problema de producción de biomasa: el estudio de la influencia de la temperatura en la extracción de azúcares (principalmente, xilosa) a partir de los desechos de la poda de olivos. Estos desechos constituyen una biomasa muy abundante en la provincia de Jaén. Una de sus ventajas principales es la concentración geográfica con que se producen. La provincia de Jaén cuenta con el 40 % del área total dedicada al cultivo de olivo en España. Comparados estos restos con los de otros procesos, y debido a su acumulación geográfica, su recolección y su transporte son realmente baratos. Los

restos de la poda de olivo constituyen una fuente potencial de xilosa que puede ser utilizada como materia prima en la producción de xilitol. Éste es un edulcorante que mejora el sabor y proporciona algunos beneficios saludables (por ejemplo, dentales), a la vez que se recomienda para personas diabéticas e incluso es utilizado como un fármaco contra el cáncer.

Capítulo 1

Preliminares: espacios métricos difusos, espacios métricos probabilísticos y regresión

Como detallábamos en la Introducción, en la presente Memoria obtenemos algunos resultados de comparación entre las diferentes nociones que se han dado a lo largo de la Historia acerca de *espacio métrico difuso*, a la vez que proponemos una estructura común utilizando números difusos. Este primer Capítulo pretende exponer los preliminares necesarios acerca de las estructuras métricas que consideraremos, apoyándonos en los conceptos de función de distribución, función triangular y t-norma. Por supuesto, los contenidos que incluimos aquí se pueden encontrar en forma mucho más amplia en la literatura sobre el tema. Pretendemos exponerlos en la forma más rápida y, a nuestro entender, más fácil de asimilar e intentamos evitar las dosis considerables de ambigüedad, y algunas veces de confusión, que se pueden encontrar en los textos habituales. Sólo daremos demostraciones de los resultados más relevantes desde nuestro punto de vista aunque, en todo caso, proporcionaremos indicaciones más que suficientes para que el lector pueda desarrollarlas sin demasiado esfuerzo, apoyándose, si fuera necesario en la bibliografía recomendada.

Quizás los mejores textos para introducirse en las bases algebraicas de la teoría de espacios métricos difusos así como en los fundamentos elementales de la incertidumbre pueden ser [111, 31, 87, 123].

1.1. Espacios métricos

En lo sucesivo, denotaremos por \mathbb{R} al conjunto formado por todos los números reales y por $\overline{\mathbb{R}} = [-\infty, \infty]$ a la recta real ampliada (por comodidad, utilizaremos la notación ∞ para $+\infty$). Escribiremos

$$\mathbb{R}^+ =]0, \infty[, \quad \mathbb{R}_0^+ = [0, \infty[, \quad \overline{\mathbb{R}}_0^+ = [0, \infty].$$

Los conjuntos \mathbb{R}^- , \mathbb{R}_0^- y $\overline{\mathbb{R}}_0^-$ se definen análogamente. Denotaremos por \mathbb{I} al intervalo real $[0, 1]$. En lo que sigue, X denotará un conjunto no vacío.

Definición 1.1.1 Una métrica sobre X es cualquier aplicación $d : X \times X \rightarrow [0, \infty[$ que verifique las siguientes propiedades.

- Para cada $x \in X$ se tiene que $d(x, x) = 0$. (1.1)

- Si $x, y \in X$ verifican que $d(x, y) > 0$, entonces $x \neq y$. (1.2)

- Para cada $x, y \in X$ se tiene que $d(x, y) = d(y, x)$. (1.3)

- Para cada $x, y, z \in X$ se tiene que $d(x, z) \leq d(x, y) + d(y, z)$. (1.4)

La condición (1.4) se denomina *desigualdad triangular*. Si d es una métrica sobre X , diremos que (X, d) es un *espacio métrico* (abreviadamente, un *EM*).

En realidad, condiciones necesarias y suficientes para que una aplicación $d : X \times X \rightarrow \mathbb{R}$ sea una métrica sobre X son las siguientes:

(i) $d(x, y) = 0$ si, y sólo si, $x = y$; (ii) $d(x, y) \leq d(z, x) + d(z, y)$.

A partir de estas propiedades se puede deducir fácilmente que $d(x, y) \geq 0$ y que $d(y, x) = d(x, y)$ para cada $x, y \in X$.

Ejemplo 1.1.2 Si X es cualquier subconjunto de \mathbb{R} , la métrica usual (o euclídea) sobre X es $d(x, y) = |x - y|$, para cada $x, y \in X$.

Ejemplo 1.1.3 Si X es cualquier subconjunto de \mathbb{R}^n , la métrica usual (o euclídea) sobre X es

$$d_2(x, y) = \sqrt{(x_1 - y_1)^2 + (x_2 - y_2)^2 + \dots + (x_n - y_n)^2}$$

para cada $x = (x_1, x_2, \dots, x_n), y = (y_1, y_2, \dots, y_n) \in X$.

Ejemplo 1.1.4 La métrica euclídea es un caso particular de la métrica d_p sobre \mathbb{R}^n definida para $p > 0$ mediante:

$$d_p(x, y) = [(x_1 - y_1)^p + (x_2 - y_2)^p + \dots + (x_n - y_n)^p]^{1/p}$$

para cada $x = (x_1, x_2, \dots, x_n), y = (y_1, y_2, \dots, y_n) \in \mathbb{R}^n$.

Ejemplo 1.1.5 Cuando p se hace arbitrariamente grande, la distancia anterior tiende a valer

$$d_\infty(x, y) = \max(|x_1 - y_1|, |x_2 - y_2|, \dots, |x_n - y_n|)$$

para cada $x = (x_1, x_2, \dots, x_n), y = (y_1, y_2, \dots, y_n) \in \mathbb{R}^n$.

Ejemplo 1.1.6 Si X es cualquier conjunto no vacío, llamaremos distancia discreta sobre X a la distancia

$$d(x, y) = \begin{cases} 0, & \text{si } x = y, \\ 1, & \text{si } x \neq y. \end{cases}$$

Ejemplo 1.1.7 Dado $p > 0$, denotemos por \mathcal{L}^p al conjunto de funciones $f : \mathbb{R}^n \rightarrow \mathbb{R}$ tales que

$$\int_{\mathbb{R}^n} |f|^p < \infty$$

(tomándose la integral en el sentido de Lebesgue). Entonces \mathcal{L}^p es un espacio métrico con la distancia

$$d_p(f, g) = \left(\int_{\mathbb{R}^n} |f - g|^p \right)^{1/p} \quad \text{para cada } f, g \in \mathcal{L}^p.$$

Ejemplo 1.1.8 Si \mathcal{L}^∞ denota el conjunto de funciones acotadas sobre \mathbb{R}^n , podemos definir

$$d_\infty(f, g) = \sup(\{|f(t) - g(t)| : t \in \mathbb{R}\}) \quad \text{para cada } f, g \in \mathcal{L}^\infty.$$

Ejemplo 1.1.9 Si Y es un subconjunto de X , la restricción a $Y \times Y$ de cualquier distancia sobre X es una distancia sobre Y .

Ejemplo 1.1.10 Uno de los ambientes en que la noción de distancia es más intuitiva es el de las variedades de Riemann. Sea (M^n, g) una variedad de Riemann conexa. Una curva C^∞ a trozos es una curva continua $\alpha : [0, 1] \rightarrow M$ de manera que existe una partición $\{a = t_0 < t_1 < \dots < t_k = b\}$ de $[a, b]$ tal que $\alpha|_{[t_{i-1}, t_i]}$ es una curva diferenciable C^∞ para cada $i \in \{1, 2, \dots, k\}$. La longitud de α es el número real

$$L(\alpha) = \int_a^b \|\alpha'(t)\| dt = \sum_{i=1}^k \int_{t_{i-1}}^{t_i} \|\alpha'(t)\| dt.$$

Utilizando la fórmula del cambio de variable en integración, es sencillo demostrar que la longitud de una curva es invariante frente a reparametrizaciones y por isometrías locales. Se define entonces la distancia entre dos puntos p y q de M como el ínfimo de todas las curvas C^∞ a trozos con origen en p y extremo en q .

$$d(p, q) = \inf(\{L(\alpha) : \alpha \text{ curva } C^\infty \text{ a trozos con origen } p \text{ y extremo } q\}).$$

Es posible demostrar que d es una distancia sobre M . Además, cuando esta distancia es completa, el teorema de Hopf-Rinow (1931) nos asegura que cualesquiera dos puntos de M se pueden unir no ya por una curva C^∞ a trozos, sino por una geodésica minimizante (véanse [73, 88]).

1.2. Operaciones binarias, funciones triangulares y t-normas

Hemos comentado en la *Introducción* de la presente Memoria la necesidad de hacer intervenir nociones probabilísticas en el cálculo de distancias entre puntos. Así, a cada par de puntos hemos de asociarle un objeto matemático capaz de medir probabilidades. Lo usual es trabajar con funciones de distribución, que nos permiten situarnos en cualquier contexto probabilístico. De esta forma, la *desigualdad triangular* deberá relacionar funciones de este tipo. Veamos en la presente sección cómo deben ser éstas y algunas de sus principales propiedades. Casi todas las demostraciones de los enunciados que presentamos pueden consultarse en [111].

1.2.1. Conjuntos parcialmente ordenados

Llamaremos *relación binaria sobre un conjunto* X a cualquier subconjunto \mathcal{R} no vacío del producto cartesiano $X \times X$. Escribiremos $a\mathcal{R}b$ si $(a, b) \in \mathcal{R}$. Diremos que \mathcal{R} es un *orden parcial sobre* X si cumple las tres siguientes propiedades:

- *Reflexiva*: para cada $a \in X$ se tiene que $a\mathcal{R}a$.
- *Transitiva*: si $a\mathcal{R}b$ y $b\mathcal{R}c$, entonces $a\mathcal{R}c$.
- *Antisimétrica*: si $a\mathcal{R}b$ y $b\mathcal{R}a$, entonces $a = b$.

Es tal la importancia del orden parcial \leq sobre los números reales que las propiedades anteriores se suelen escribir con la siguiente notación.

- *Reflexiva*: para cada $a \in X$ se tiene que $a \leq a$.
- *Transitiva*: si $a \leq b$ y $b \leq c$, entonces $a \leq c$.
- *Antisimétrica*: si $a \leq b$ y $b \leq a$, entonces $a = b$.

Escribiremos que $a < b$ (o, equivalentemente, que $b > a$) si $a \leq b$ pero $a \neq b$. Diremos que (X, \leq) es un *conjunto parcialmente ordenado* si \leq es un orden parcial sobre X .

Dados dos puntos $a, b \in X$, diremos que un elemento $c \in X$ es una *cota superior de* a y de b si $a \leq c$ y $b \leq c$. Diremos que $s \in X$ es el *supremo de* a y de b si s es la mínima cota superior de a y de b , es decir, s es una cota superior de a y de b y si $c \in X$ es cualquier otra cota superior de a y de b , entonces $s \leq c$. Análogamente se definen los conceptos de *cota inferior* y de *ínfimo* de dos elementos a y b .

En principio, dos elementos de un conjunto parcialmente ordenado pueden no poseer ni supremo ni ínfimo. No obstante, cuando dos elementos poseen supremo (o ínfimo), éste es único (pues los dos deben ser comparables). En caso de existir, se suele denotar por $a \vee b$ al supremo de los elementos a y b y por $a \wedge b$ a su ínfimo.

Llamaremos *retículo* a cualquier conjunto parcialmente ordenado (X, \leq) en el que cada dos elementos poseen ínfimo y supremo. Esta definición se extiende de forma sencilla a conjunto

finitos. Diremos que un retículo (X, \leq) es *completo* si cualquier subconjunto no vacío de X posee ínfimo y supremo. En particular, un retículo completo posee un elemento supremo (un máximo absoluto) y un elemento ínfimo (un mínimo absoluto), es decir, dos elementos $I, S \in X$ de manera que $I \leq x \leq S$ para cada $x \in X$.

Un ejemplo de retículo completo es el subconjunto de partes de un conjunto no vacío X con la relación de inclusión $A \subseteq B$.

1.2.2. Operaciones binarias

Definición 1.2.1 Llamaremos operación binaria sobre X a cualquier aplicación $T : X \times X \rightarrow X$. Diremos que una operación T :

- ▶ es asociativa si $T(T(x, y), z) = T(x, T(y, z))$, para cada $x, y, z \in X$;
- ▶ es conmutativa si para cada $x, y \in X$ se tiene que $T(x, y) = T(y, x)$;
- ▶ posee al elemento $e \in X$ como unidad si para cada $x \in X$ se tiene que $T(x, e) = T(e, x) = x$;
- ▶ posee al elemento $c \in X$ como cero si para cada $x \in X$ se tiene que $T(x, c) = T(c, x) = c$.

Supongamos ahora que (X, \leq) es un conjunto parcialmente ordenado. Diremos que una operación T :

- ▶ es no decreciente en cada argumento si para cada $x_1, x_2, y_1, y_2 \in X$, con $x_1 \leq x_2$ e $y_1 \leq y_2$, se tiene que $T(x_1, y_1) \leq T(x_2, y_2)$.

Dadas dos operaciones binarias T_1 y T_2 sobre X , diremos que $T_1 \leq T_2$ (o, equivalentemente, $T_2 \geq T_1$) si para cada $x, y \in X$ se tiene que $T_1(x, y) \leq T_2(x, y)$. Diremos también que $T_1 < T_2$ si $T_1 \leq T_2$ pero $T_1 \neq T_2$ (difieren en algún punto).

Supongamos ahora que X es un espacio topológico. Diremos que una operación T :

- ▶ es continua si es una aplicación continua de $X \times X$, con la topología producto, en X .

En el caso particular de que $X = \mathbb{I} = [0, 1]$, diremos que una operación $T : \mathbb{I} \times \mathbb{I} \rightarrow \mathbb{I}$:

- ▶ es continua a la izquierda si para cada $x, y \in]0, 1]$ se tiene que:

$$T(x, y) = \sup \left(\left\{ T(u, v) : 0 < u < x, 0 < v < y \right\} \right).$$

Aunque pueda parecer lo contrario debido a su nombre, no existe ninguna relación directa entre las operaciones *continuas* y las *continuas a la izquierda*. Por ejemplo, consideremos la función $H : \mathbb{I} \times \mathbb{I} \rightarrow \mathbb{I}$ definida por:

$$H(x, y) = \left(x - \frac{1}{2}\right)^2 + \left(y - \frac{1}{2}\right)^2, \quad \forall x, y \in \mathbb{I}.$$

Su mínimo absoluto es $H(1/2, 1/2) = 0$, y su máximo absoluto se alcanza en las esquinas del cuadrado $\mathbb{I} \times \mathbb{I}$, y vale $H(0, 0) = 1/2$. Por consiguiente, H , a pesar de ser continua, no es continua a la izquierda, pues:

$$\sup \left(\left\{ H(x, y) \mid 0 < x < 1/2, 0 < y < 1/2 \right\} \right) = H(0, 0) > H(1/2, 1/2).$$

Sin embargo, hay un caso en el que la continuidad implica continuidad a la izquierda.

Lema 1.2.2 Toda operación binaria $T : \mathbb{I} \times \mathbb{I} \rightarrow \mathbb{I}$ continua y no decreciente en cada argumento es continua a la izquierda.

1.2.3. Funciones triangulares sobre un conjunto

Definición 1.2.3 Sea (X, \leq) un conjunto parcialmente ordenado con mínimo absoluto x_m y máximo absoluto x_M . Llamaremos función triangular sobre X a cualquier operación binaria $\tau : X \times X \rightarrow X$ que sea asociativa, conmutativa, no decreciente en cada argumento y posea a x_M como unidad. En definitiva, τ debe verificar las siguientes propiedades:

- Para cada $x, y, z \in X$ se tiene que $\tau(\tau(x, y), z) = \tau(x, \tau(y, z))$. (1.5)

- Para cada $x, y \in X$ se tiene que $\tau(x, y) = \tau(y, x)$. (1.6)

- Si $x_1, x_2, y_1, y_2 \in X$ verifican que $x_1 \leq x_2$ e $y_1 \leq y_2$, entonces $\tau(x_1, y_1) \leq \tau(x_2, y_2)$. (1.7)

- Para cada $x \in X$ se tiene que $\tau(x_M, x) = \tau(x, x_M) = x$. (1.8)

Llamaremos función triangular a cualquier función triangular sobre Δ^+ (conjunto que definiremos en la subsección 1.2.5).

Las funciones triangulares sobre X también reciben el nombre de *normas triangulares sobre X* , y algunos autores las denominan *t_X -normas*.

El elemento mínimo de X es un cero de cualquier función triangular ya que, para cada $x \in X$, se tiene que $x_m \leq \tau(x_m, x) \leq \tau(x_m, x_M) = x_m$, de donde $\tau(x_m, x) = x_m$.

1.2.4. t-normas y t-conormas

Sea $[a, b]$ sea un intervalo cerrado y acotado de números reales.

Definición 1.2.4 Llamaremos t-norma sobre $[a, b]$ a cualquier función triangular sobre $[a, b]$, es decir, a cualquier operación binaria sobre $[a, b]$ que sea asociativa, conmutativa, no decreciente en cada argumento y posea al elemento 'b' (máximo de $[a, b]$) como unidad.

Llamaremos t-conorma sobre $[a, b]$ a cualquier operación binaria sobre $[a, b]$ que sea asociativa, conmutativa, no decreciente en cada argumento y posea al elemento 'a' (mínimo de $[a, b]$) como unidad.

Llamaremos t-norma (respectivamente, t-conorma) a cualquier t-norma (respectivamente, t-conorma) sobre el intervalo $\mathbb{I} = [0, 1]$.

En lo sucesivo, utilizaremos la notación $*$: $[0, 1]^2 \rightarrow [0, 1]$ para referirnos a una t-norma, y \diamond : $[0, 1]^2 \rightarrow [0, 1]$ significará una t-conorma. La única diferencia entre estos dos conceptos es que las t-normas poseen a 1 como unidad y a 0 como cero, mientras que las t-conormas poseen a 0 como unidad y a 1 como cero. Dicho de otra forma, para cada $a \in \mathbb{I}$ se tiene que:

$$a * 1 = a = a \diamond 0 \quad \text{y} \quad a * 0 = 0 = a \diamond 1.$$

Ejemplo 1.2.5 Los tres ejemplos más importantes de t-normas, que utilizaremos con la siguiente notación, son M , W y Π definidas de la forma:

$$\begin{aligned} M(x, y) &= \min(x, y), \\ W(x, y) &= \max(x + y - 1, 0), \\ \Pi(x, y) &= xy, \end{aligned}$$

para cada $x, y \in \mathbb{I}$. Estas tres t-normas son claramente continuas y continuas a la izquierda. Hay otro ejemplo interesante de t-norma, aunque no es continua, y es el siguiente:

$$\forall x, y \in I, \quad Z(x, y) = \begin{cases} x, & \text{si } y = 1, \\ y, & \text{si } x = 1, \\ 0, & \text{si } x, y \in [0, 1[. \end{cases}$$

Estas cuatro t-normas están ordenadas de la siguiente forma.

Proposición 1.2.6 Con la notación anterior, $Z < W < \Pi < M$. De hecho, si $*$ es cualquier t-norma, entonces $Z \leq * \leq M$.

En este sentido, M es la mayor t-norma posible y Z es la menor t-norma posible. El lema 1.2.2 se traduce, en el caso de t-normas, en la siguiente propiedad.

Proposición 1.2.7 Toda t-norma continua es continua a la izquierda.

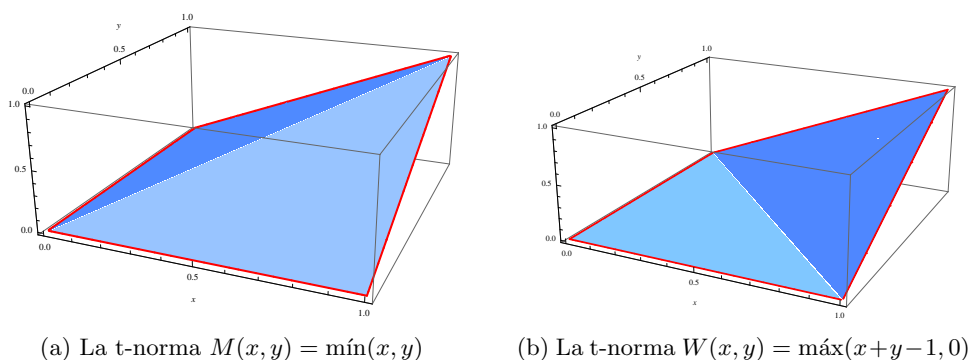


Figura 1.1: Algunas de las t-normas más utilizadas en la práctica.

Es sencillo construir una t-conorma a partir de una t-norma, y viceversa, de la siguiente manera.

Lema 1.2.8 Sea $* : \mathbb{I}^2 \rightarrow \mathbb{I}$ una operación binaria y definamos $*' : \mathbb{I}^2 \rightarrow \mathbb{I}$ de la siguiente forma:

$$a *' b = 1 - (1 - a) * (1 - b), \quad \text{para cada } a, b \in \mathbb{I}.$$

Entonces $*$ es una t-norma si, y sólo si, $*'$ es una t-conorma.

La t-conorma $*'$ suele denominarse *t-conorma asociada a la t-norma $*$* . Recíprocamente, si \diamond es una operación binaria sobre \mathbb{I} y definimos $a \diamond' b = 1 - (1 - a) \diamond (1 - b)$ para cada $a, b \in \mathbb{I}$, entonces \diamond es una t-conorma si, y sólo si, \diamond' es una t-norma. Además, los procesos anteriores son complementarios, es decir, $*'' = *$ y $\diamond'' = \diamond$.

Ejemplo 1.2.9 Las t-conormas asociadas a las t-normas presentadas en el Ejemplo 1.2.5 están definidas, para cada $x, y \in \mathbb{I}$, como:

$$M'(x, y) = \max(x, y), \quad W'(x, y) = \min(x + y, 1), \quad \Pi'(x, y) = x + y - xy,$$

$$Z'(x, y) = \begin{cases} x, & \text{si } y = 0, \\ y, & \text{si } x = 0, \\ 1, & \text{si } x, y \in]0, 1]. \end{cases}$$

De esta forma, se tiene que:

$$Z < W < \Pi < M < M' < \Pi' < W' < Z'.$$

En general, si $*$ es cualquier t-norma y \diamond es cualquier t-conorma, se tiene que:

$$Z \leq * \leq M < M' \leq \diamond \leq Z'.$$

En particular, no puede ocurrir que $\diamond \leq *$. En concreto, si $*$ es cualquier t-norma,

$$Z \leq * \leq M < M' \leq *' \leq Z'.$$

Un tipo especial de t-normas son las siguientes. Sea $*$ una t-norma y sea $a \in \mathbb{I}$. Denotaremos por $\{*^n a\}_{n=1}^{\infty}$ a la sucesión de \mathbb{I} definida, por recurrencia, como:

$$*^1 a = a, \quad *^n a = (*^{n-1} a) * a \quad \text{para cada } n \in \mathbb{N}, n \geq 2. \quad (1.9)$$

Definición 1.2.10 [57] Diremos que una t-norma $*$ es de tipo Hadžić (o de tipo H) si la sucesión $\{*^n a\}_{n=1}^{\infty}$ es equicontinua en $a = 1$, es decir, para cada $\varepsilon > 0$, existe $\delta > 0$ de manera que si $a \in]1 - \delta, 1]$, entonces $*^n a > 1 - \varepsilon$ para cada $n \in \mathbb{N}$.

Por ejemplo, la t-norma mínimo M es una t-norma de tipo Hadžić.

1.2.5. Funciones de distribución de tipo distancia. El conjunto Δ^+

Llamaremos *función de distribución* a cualquier función no decreciente $F : \mathbb{R} \rightarrow \mathbb{I} = [0, 1]$ que sea continua a la izquierda. Podemos extender esta función a $\overline{\mathbb{R}}$ definiendo $F(-\infty) = 0$ y $F(\infty) = 1$, de manera que sigue siendo no decreciente. Este tipo de funciones no son exactamente las que nos van a servir para medir distancias pues, en principio, pueden estar asociadas a variables aleatorias que tomen valores negativos. Necesitamos quedarnos con la clase de las funciones de distribución que se anulan en $[-\infty, 0]$.

Definición 1.2.11 Una función de distribución de tipo distancia (abreviadamente, f.d.d.) es una función no decreciente $f : [0, \infty] \rightarrow \mathbb{I}$, continua a la izquierda en $]0, \infty[$ y que cumple que $f(0) = 0$ y $f(\infty) = 1$. Denotaremos por Δ^+ al conjunto formado por todas las f.d.d.

Se puede extender cada f.d.d. a $f : \overline{\mathbb{R}} \rightarrow [0, 1]$ definiendo $f(t) = 0$ si $t < 0$, aunque esta parte de la función no aporta nada nuevo.

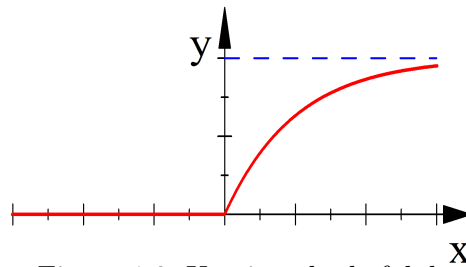
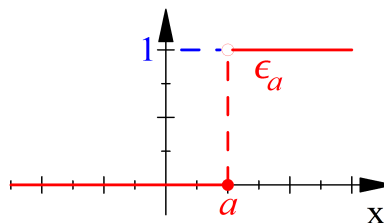


Figura 1.2: Un ejemplo de f.d.d.

Ejemplo 1.2.12 Ejemplos de f.d.d. son las siguientes funciones de salto. Para cada $a \in [0, \infty[$,

definimos $\epsilon_a : [0, \infty] \rightarrow [0, 1]$ como:

$$\epsilon_a(t) = \begin{cases} 0, & \text{si } t \leq a, \\ 1, & \text{si } t > a, \end{cases}$$



e igualmente definimos:

$$\epsilon_\infty(t) = \begin{cases} 0, & \text{si } t \in [0, \infty[, \\ 1, & \text{si } t = \infty. \end{cases}$$

Cada función ϵ_a es continua en $[0, \infty] \setminus \{a\}$, pues es localmente constante en este conjunto.

Ejemplo 1.2.13 Dados números distintos $a_1, a_2, \dots, a_n \in [0, \infty[$, la función

$$\epsilon_{a_1, a_2, \dots, a_n} = \frac{\epsilon_{a_1} + \epsilon_{a_2} + \dots + \epsilon_{a_n}}{n}$$

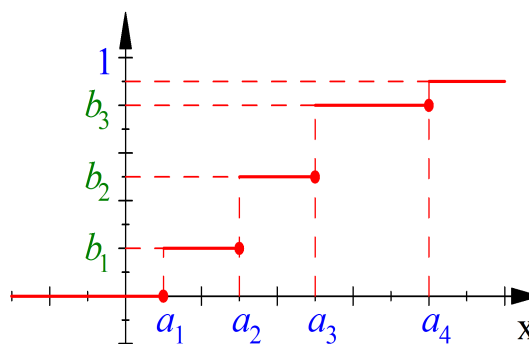
también pertenece a Δ^+ . De hecho, si $a_1 < a_2 < \dots < a_n$, entonces

$$\epsilon_{a_1, a_2, \dots, a_n}(t) = \begin{cases} 0, & \text{si } t \leq a_1, \\ \frac{i}{n}, & \text{si } a_i < t \leq a_{i+1}, \\ 1, & \text{si } t > a_n. \end{cases}$$

Estas funciones son escalonadas y presentan un número finito de discontinuidades.

Estas f.d.d. corresponden a variables aleatorias discretas, y sus funciones de distribución asociadas son de la siguiente forma.

$$F(x) = \begin{cases} 0, & \text{si } x \leq a_1, \\ b_1, & \text{si } a_1 < x \leq a_2, \\ \vdots & \\ b_n, & \text{si } a_n < x \leq a_{n+1}, \\ \vdots & \\ 1, & \text{si } x > a_\infty. \end{cases}$$



El conjunto de f.d.d. Δ^+ es un conjunto parcialmente ordenado mediante la relación:

$$f, g \in \Delta^+, \quad f \leq g \Leftrightarrow \left[f(x) \leq g(x), \quad \text{para cada } x \in]0, \infty[\right].$$

Si $f \in \Delta^+$ es cualquier f.d.d., entonces $\epsilon_\infty(t) = 0 \leq F(t) \leq 1 = \epsilon_0(t)$ para cada $t \in]0, \infty[$, lo que demuestra que

$$\epsilon_\infty \leq f \leq \epsilon_0 \quad \text{para cada } f \in \Delta^+.$$

De esta forma, ϵ_0 es el máximo absoluto de Δ^+ y ϵ_∞ es el mínimo absoluto de Δ^+ .

El primer resultado interesante que podemos destacar en Δ^+ es la existencia de supremo en el retículo parcialmente ordenado (Δ^+, \leq) .

Lema 1.2.14 El supremo de cualquier conjunto (no vacío) de f.d.d. en Δ^+ es otra f.d.d. de Δ^+ .

Dicho de otro modo, si $\{f_i\}_{i \in \Lambda}$ es una familia de d.d.f en Δ^+ y definimos

$$\sup_{i \in \Lambda} f_i : [0, \infty] \rightarrow [0, 1], \quad \left(\sup_{i \in \Lambda} f_i \right) (x) = \sup_{i \in \Lambda} f_i(x), \quad \text{para cada } x \in [0, \infty],$$

entonces $\sup_{i \in \Lambda} f_i \in \Delta^+$, y además es obvio que:

$$\sup_{i \in \Lambda} f_i \geq f_j, \quad \text{para cada } j \in \Lambda.$$

No ocurre lo mismo con el ínfimo, que podría no ser continua a la izquierda.

Además de un conjunto parcialmente ordenado, Δ^+ es un espacio *metrizable*. En 1925, P. Lévy introdujo una métrica sobre el conjunto de funciones de distribución. Dicha métrica fue modificada por D.A. Sibley en 1971, y después por B. Schweizer y A. Sklar en [111] con el objeto de que cumpliera propiedades más interesantes y poder trabajar de una forma más cómoda. Sean $f, g \in \Delta^+$ y sea $\delta \in [0, 1]$. Diremos que se cumple la condición $[f, g; \delta]$ si:

$$g(x) \leq f(x + \delta) + \delta, \quad \text{para cada } x \in]0, \frac{1}{\delta}[.$$

Obsérvese que la condición $[f, g; 1]$ siempre se cumple, pues:

$$g(x) \leq 1 \leq 1 + f(x + 1), \quad \text{para cada } x \in]0, 1[.$$

Definición 1.2.15 Dadas dos f.d.d. $f, g \in \Delta^+$, llamaremos distancia de Lévy modificada entre f y g , y la denotaremos por $d_L(f, g)$, al número real:

$$d_L(f, g) = \inf \left(\left\{ \delta > 0 \mid [f, g; \delta] \text{ y } [g, f; \delta] \text{ ocurren a la vez} \right\} \right).$$

Con esta definición, se puede demostrar que la aplicación $d_L : \Delta^+ \times \Delta^+ \rightarrow \mathbb{R}_0^+$ cumple las siguientes propiedades.

Lema 1.2.16 1. d_L es una métrica sobre Δ^+ de manera que (Δ^+, d_L) es un espacio métrico compacto y completo.

2. Si $d_L(f, g) = \delta > 0$, entonces el ínfimo que la define es un mínimo, es decir, $[f, g; \delta]$ y $[g, f; \delta]$ ocurren a la vez.
3. Para cada $f \in \Delta^+$ y cada $t > 0$ se tiene que $d_L(f, \epsilon_0) < t$ si, y sólo si, $f(t) > 1 - t$.
4. Si $f, g \in \Delta^+$ verifican $f \leq g$, entonces $d_L(g, \epsilon_0) \leq d_L(f, \epsilon_0)$.

Se dice que una sucesión de f.d.d. $\{f_n\}$ converge débilmente a f si $\{d_L(f_n, f)\}$ converge a cero.

Lema 1.2.17 Una sucesión $\{f_n\}$ de f.d.d. de Δ^+ converge débilmente a $f \in \Delta^+$ si, y sólo si, la sucesión de números reales $\{f_n(x)\}$ converge a $f(x)$ en cada punto de continuidad x de f .

Observación 1.2.18 Como se comentó en la Definición 1.2.3 y dado que son las que más nos interesan, llamaremos función triangular a cualquier función triangular sobre Δ^+ . El elemento unidad de cualquier función triangular $\tau : \Delta^+ \times \Delta^+ \rightarrow \Delta^+$ es ϵ_0 (pues es el máximo de Δ^+) y su elemento cero es ϵ_∞ .

$$\tau(f, \epsilon_0) = f \quad \text{y} \quad \tau(f, \epsilon_\infty) = \epsilon_\infty \quad \text{sea cual sea } f \in \Delta^+. \quad (1.10)$$

Definición 1.2.19 Diremos que una función triangular $\tau : \Delta^+ \times \Delta^+ \rightarrow \Delta^+$ es continua si es continua en $\Delta^+ \times \Delta^+$ con la topología producto asociada a la métrica de Lévy modificada d_L en Δ^+ .

Una función triangular $\tau : \Delta^+ \times \Delta^+ \rightarrow \Delta^+$ puede verificar dos condiciones interesantes que nos aparecerán más adelante, y que son las siguientes.

$$\bullet \quad \text{Para cada } a, b \in]0, \infty[\text{ se tiene que } \tau(\epsilon_a, \epsilon_b) \geq \epsilon_{a+b}. \quad (1.11)$$

$$\bullet \quad \text{Para cada } a, b \in]0, \infty[\text{ se tiene que } \tau(\epsilon_a, \epsilon_b) = \epsilon_{a+b}. \quad (1.12)$$

Evidentemente, la condición (1.12) implica la condición (1.11), pero al revés no es cierto. Si a ó b toman los valores 0 ó ∞ , cualquier función triangular cumple las dos condiciones anteriores.

1.2.6. Funciones triangulares asociadas a t-normas

Una t-norma (sobre \mathbb{I}) ya es una función triangular sobre \mathbb{I} . No obstante, ciertas t-normas (sobre \mathbb{I}) permiten definir funciones triangulares sobre Δ^+ .

Lema 1.2.20 Si $*$ es una t-norma continua a la izquierda, entonces la operación $T_* : \Delta^+ \times \Delta^+ \rightarrow \Delta^+$ definida por:

$$T_*(f, g)(x) = f(x) * g(x), \quad \text{para cada } f, g \in \Delta^+ \text{ y cada } x \in [0, \infty],$$

está bien definida y es una función triangular.

En particular, tomando $*$ = mín, tenemos la siguiente propiedad interesante.

Proposición 1.2.21 La función triangular $T_M : \Delta^+ \times \Delta^+ \rightarrow \Delta^+$ definida de la forma:

$$T_M(F, G)(x) = \text{mín}(F(x), G(x)), \quad \text{para cada } f, g \in \Delta^+ \text{ y cada } x \in [0, \infty],$$

es maximal entre todas las funciones triangulares, es decir, si τ es cualquier función triangular, entonces $\tau \leq T_M$.

Otra forma de considerar funciones triangulares definidas a partir de t-normas es la siguiente.

Teorema 1.2.22 Si $*$ es una t-norma continua a la izquierda y definimos, para cada $f, g \in \Delta^+$ y cada $x \in [0, \infty]$:

$$\tau_*(f, g)(x) = \sup(\{f(u) * g(v) : u, v \geq 0 \text{ y } u + v = x\}),$$

entonces $\tau_* : \Delta^+ \times \Delta^+ \rightarrow \Delta^+$ es una función triangular que verifica la propiedad (1.12).

Además, si $*$ es continua, entonces τ_* es uniformemente continua en (Δ^+, d_L) .

1.3. Números difusos

Llamaremos *conjunto difuso sobre X* a cualquier aplicación $F : X \rightarrow \mathbb{I}$. Dado $x \in X$, el número $F(x) \in \mathbb{I}$ representa el grado de certidumbre que poseemos acerca de que el elemento x pertenezca a F , de tal manera que si $F(x) = 0$, estamos seguros de que no pertenece a F y si $F(x) = 1$, entonces estamos seguros de lo contrario.

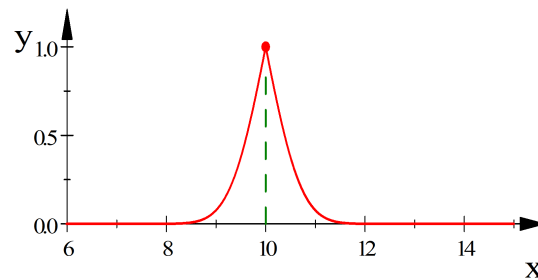


Figura 1.3: Un ejemplo de número difuso cuyo núcleo es $\{10\}$

Definición 1.3.1 Un número difuso sobre \mathbb{R} es cualquier conjunto difuso $F : X \rightarrow \mathbb{I}$ que verifique las dos siguientes condiciones:

- Normalidad: existe un número real $x_0 \in \mathbb{R}$ tal que $F(x_0) = 1$.

- Para cada $\alpha \in]0, 1]$, el conjunto $F_{[\alpha]} = \{x \in \mathbb{R} : F(x) \geq \alpha\}$ es un subintervalo cerrado de \mathbb{R} .

El conjunto $F_{[\alpha]}$ se denomina el α -corte de F (o el conjunto de nivel α de F). El núcleo de un número difuso F es $\ker F = F_{[1]}$ y su soporte es la clausura

$$\text{sop } F = \overline{\{x \in \mathbb{R} : F(x) > 0\}}.$$

Denotaremos por \mathcal{F} al conjunto formado por todos los números difusos que alcanzan la condición de normalidad en cero (es decir, $F(0) = 1$).

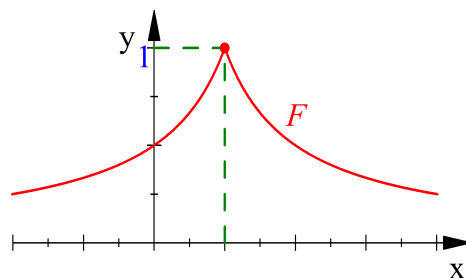
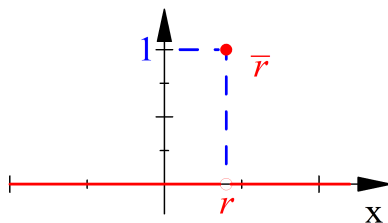


Figura 1.4: Un número difuso que no se anula en \mathbb{R} .

Un número difuso puede no tomar el valor cero sobre \mathbb{R} (como se aprecia en la Figura 1.4). No obstante, cualquier número difuso sobre \mathbb{R} puede ser extendido a $\overline{\mathbb{R}}$ con sólo definir $F(\pm\infty) = 0$. Cada número real $r \in \mathbb{R}$ puede ser considerado como un número difuso $\bar{r} : \overline{\mathbb{R}} \rightarrow \mathbb{I}$ definido por

$$\bar{r}(x) = \begin{cases} 1, & \text{si } x = r, \\ 0, & \text{si } x \neq r. \end{cases}$$



El conjunto \mathcal{F} está parcialmente ordenado por la relación punto a punto, es decir, si $F, G \in \mathcal{F}$, entonces $F \leq G$ si $F(x) \leq G(x)$ para cada $x \in \mathbb{R}$. Si denotamos por $\mathbf{1}$ al número difuso $\mathbf{1} : \overline{\mathbb{R}} \rightarrow \mathbb{I}$ definido por:

$$\mathbf{1}(x) = \begin{cases} 1, & \text{si } x \in \mathbb{R}, \\ 0, & \text{si } x = \pm\infty, \end{cases}$$

entonces

$$\bar{0} \leq F \leq \mathbf{1}, \quad \text{para cada } F \in \mathcal{F}.$$

La siguiente propiedad es interesante para describir la forma geométrica de un número difuso.

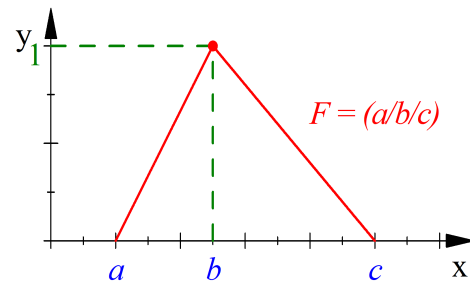
Lema 1.3.2 Sea F un conjunto difuso sobre \mathbb{R} y sea $x_0 \in \mathbb{R}$ un punto tal que $F(x_0) = 1$. Entonces F es un número difuso si, y sólo si, $F|_{[x_0, \infty[}$ es una aplicación no creciente y continua a la izquierda y $F|_{] - \infty, x_0]}$ es una aplicación no decreciente y continua a la derecha.

1.3.1. Números difusos triangulares y trapezoidales

Existe una clase de números difusos sobre \mathbb{R} , completamente determinados por su núcleo y por su soporte, que son especialmente útiles a la hora de modelizar ciertos experimentos aleatorios.

Definición 1.3.3 Llamaremos número difuso triangular a cualquier número difuso $F = (a/b/c)$, donde $a, b, c \in \mathbb{R}$ verifican $a < b < c$, definido como sigue:

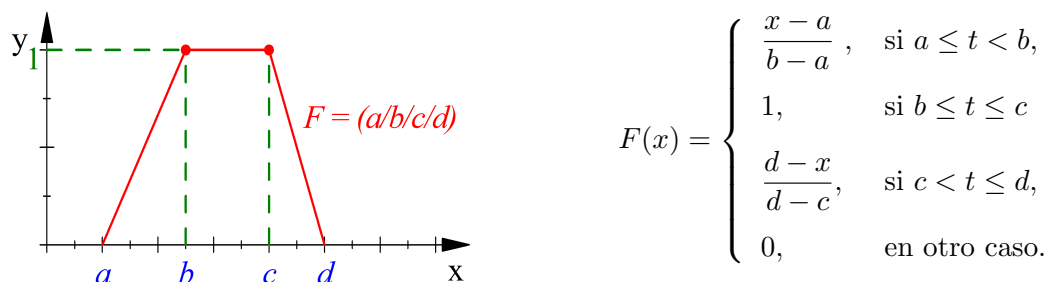
$$F(x) = \begin{cases} \frac{x-a}{b-a}, & \text{si } a \leq t \leq b, \\ \frac{c-x}{c-b}, & \text{si } b < t \leq c, \\ 0, & \text{en otro caso.} \end{cases}$$



La representación gráfica de un número difuso triangular es la de un triángulo de base $[a, c]$ y vértice 1 situado en $x = b$. El núcleo del número difuso triangular $F = (a/b/c)$ es $\{b\}$ mientras que su soporte es el intervalo $[a, c]$.

Los números difusos triangulares pertenecen a clases de números difusos aún más generales y que son muy útiles en la práctica para modelizar la incertidumbre. Por ejemplo, un *número difuso trapezoidal* es cualquier número difuso $F = (a/b/c/d)$, donde $a, b, c, d \in \mathbb{R}$ verifican

$a \leq b \leq c \leq d$, definido como sigue:



El concepto tradicional de número difuso trapezoidal ocurre cuando $a < b < c < d$. No obstante, permitiendo la igualdad entre algunos de estos valores, incluimos algunas clases de números difusos que no son continuos como aplicaciones.

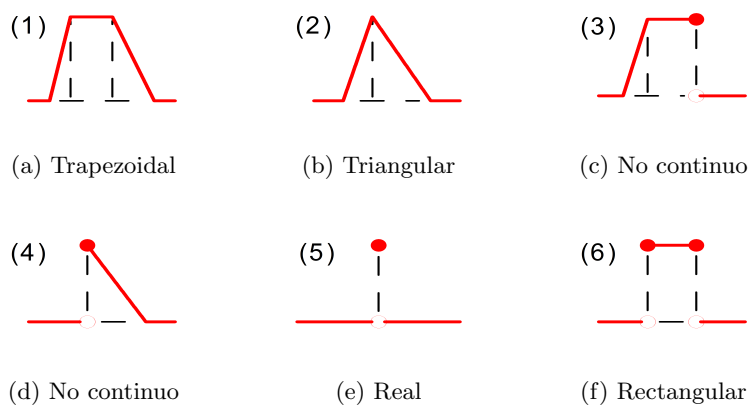


Figura 1.5: Distintas clases de números difusos trapezoidales (generalizados).

Es claro que un número triangular puede ser considerado como trapezoidal cuando $a < b = c < d$. Además, los números reales r , vistos como números difusos \bar{r} , también pueden considerarse trapezoidales si elegimos el caso en que $r = a = b = c = d$.

Claramente, el núcleo de un número difuso trapezoidal $F = (a/b/c/d)$ es $[c, d]$ y su soporte es $[a, d]$. El número real $A^c = \frac{b+c}{2}$ es el centro del número difuso y sus amplitudes (en inglés, *spreads*) se definen como:

$$A^m = \frac{c-b}{2} \geq 0, \quad A^\ell = b-a \geq 0, \quad A^s = d-c \geq 0.$$

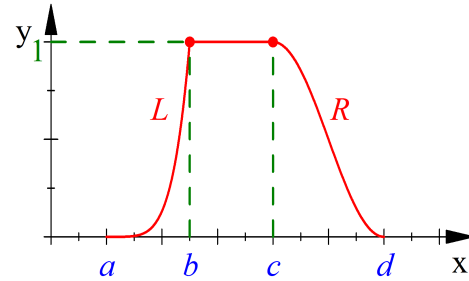
Los centros y los amplitudes también determinan completamente al número difuso trapezoidal, ya que:

$$\begin{aligned} a &= A^c - A^m - A^\ell, \\ b &= A^c - A^m, \\ c &= A^c + A^m, \\ d &= A^c + A^m + A^s. \end{aligned}$$

De esta forma, cuando un número trapezoidal se expresa en función de su centro y de sus amplitudes, lo denotaremos por $F = \text{Tra}(A^c, A^m, A^\ell, A^s)$, y si se trata de un número triangular utilizaremos la notación $F = \text{Tri}(A^c, A^\ell, A^s)$.

También es posible curvar los lados del trapecio dando lugar a los números difusos de tipo *LR*. Supongamos que $a < b \leq c < d$ y sean $L : [a, b] \rightarrow \mathbb{I}$ y $R : [c, d] \rightarrow \mathbb{I}$ dos funciones continuas y estrictamente monótonas verificando que $L(a) = 0 = R(d)$ y $L(b) = 1 = R(c)$ (en particular, L es creciente y R es decreciente). Entonces un *número difuso LR* es un número difuso de la forma:

$$F(x) = \begin{cases} L\left(\frac{x-a}{b-a}\right), & \text{si } a < x < b, \\ 1, & \text{si } b \leq x \leq c, \\ R\left(\frac{d-x}{d-c}\right), & \text{si } c < x < d, \\ 0, & \text{en otro caso.} \end{cases}$$



1.3.2. Aritmética con números difusos

Existe una aritmética entre números difusos, introducida por primera vez por M. Mizumoto y J. Tanaka [87], que extiende de manera natural a la aritmética entre números reales. Sin embargo, la forma más usual de operar con números difusos es a través de la aritmética intervalar con la ayuda de los α -cortes [129].

Cualquier número difuso $F \in \mathcal{F}$ está completamente determinado por sus α -cortes [87, 31]. Para cada $\alpha \in]0, 1]$ podemos expresar $F_{[\alpha]} = [\underline{a}_\alpha, \bar{a}_\alpha]$ (este intervalo puede ser abierto por un lado si $\underline{a}_\alpha = -\infty$ o $\bar{a}_\alpha = \infty$). Tenemos entonces aplicaciones $\underline{a}, \bar{a} :]0, 1] \rightarrow \overline{\mathbb{R}}$. Si $0 < \alpha \leq \beta \leq 1$

y $x_0 \in \ker F$, entonces $F_{[\beta]} \subseteq F_{[\alpha]}$, por lo que $\underline{a}_\alpha \leq \underline{a}_\beta \leq x_0 \leq \bar{a}_\beta \leq \bar{a}_\alpha$. En particular, $\underline{a}_\alpha \leq \underline{a}_1 \leq x_0 \leq \bar{a}_1 \leq \bar{a}_\alpha$, $\underline{a}([0, 1]) \subseteq [-\infty, \underline{a}_1] \subseteq [-\infty, x_0]$ y $\bar{a}([0, 1]) \subseteq [\bar{a}_1, \infty] \subseteq [x_0, \infty]$.

Si $G \in \mathcal{F}$ es otro número difuso con α -cortes $G_{[\alpha]} = [\underline{b}_\alpha, \bar{b}_\alpha]$ para cada $\alpha \in]0, 1]$, entonces las operaciones entre F y G están definidas como sigue.

$$C = F + G, \quad (F + G)_{[\alpha]} = [\underline{a}_\alpha + \underline{b}_\alpha, \bar{a}_\alpha + \bar{b}_\alpha], \quad \underline{c}_\alpha = \underline{a}_\alpha + \underline{b}_\alpha, \quad \bar{c}_\alpha = \bar{a}_\alpha + \bar{b}_\alpha, \quad (1.13)$$

$$C = F - G, \quad (F - G)_{[\alpha]} = [\underline{a}_\alpha - \bar{b}_\alpha, \bar{a}_\alpha - \underline{b}_\alpha], \quad \underline{c}_\alpha = \underline{a}_\alpha - \bar{b}_\alpha, \quad \bar{c}_\alpha = \bar{a}_\alpha - \underline{b}_\alpha, \quad (1.14)$$

$$C = F \cdot G, \quad (F \cdot G)_{[\alpha]} = [\text{mín } \Delta_{FG}^\alpha, \text{máx } \Delta_{FG}^\alpha], \quad \text{donde } \Delta_{FG}^\alpha = \{\underline{a}_\alpha \underline{b}_\alpha, \underline{a}_\alpha \bar{b}_\alpha, \bar{a}_\alpha \underline{b}_\alpha, \bar{a}_\alpha \bar{b}_\alpha\}, \quad (1.15)$$

$$C = F/G, \quad (F/G)_{[\alpha]} = [\text{mín } \Lambda_{FG}^\alpha, \text{máx } \Lambda_{FG}^\alpha], \quad \text{donde } \Lambda_{FG}^\alpha = \{\underline{a}_\alpha / \underline{b}_\alpha, \underline{a}_\alpha / \bar{b}_\alpha, \bar{a}_\alpha / \underline{b}_\alpha, \bar{a}_\alpha / \bar{b}_\alpha\}, \quad (1.16)$$

La división F/G sólo está bien definida si $0 \notin G_{[\alpha]}$ para cada $\alpha \in]0, 1]$.

Si tenemos en cuenta que el objetivo principal de un número difuso no es expresar certeza sino describir la incertidumbre que puede plantearse acerca de un determinado proceso, observamos que la suma de números difusos es tremendamente coherente con este concepto, pues de alguna manera transmite la incertidumbre del proceso y aumenta el tamaño tanto del núcleo como del soporte.

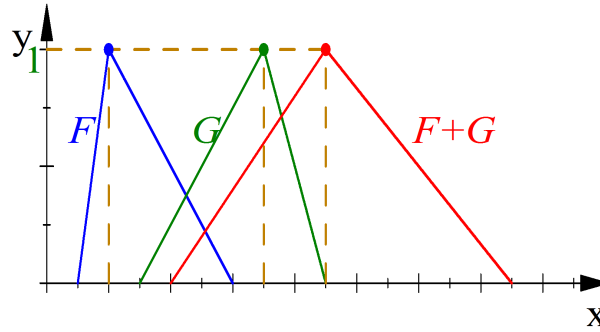


Figura 1.6: Suma de números difusos triangulares.

Por así decirlo, la suma propaga el error cometido al realizar una estimación. Sin embargo, no existe una definición coherente de resta de números difusos con la propiedad de que $(F - G) + G = F$ ya que, en este caso, la resta disminuiría el error en la estimación, y esto no es en absoluto intuitivo. De hecho, en general, $F - F \neq \bar{0}$.

Esta aritmética es especialmente sencilla entre números difusos triangulares y trapezoidales. Si $F = (A^1/A^2/A^3/A^4)$ y $G = (B^1/B^2/B^3/B^4)$ son números difusos trapezoidales, sus respectivos α -cortes, para $\alpha \in]0, 1]$, pueden ser expresados como sigue:

$$F_{[\alpha]} = [(1 - \alpha)A^1 + \alpha A^2, (1 - \alpha)A^3 + \alpha A^4], \quad G_{[\alpha]} = [(1 - \alpha)B^1 + \alpha B^2, (1 - \alpha)B^3 + \alpha B^4].$$

De esta forma, los α -cortes de su suma son:

$$(F + G)_{[\alpha]} = [(1 - \alpha)(A^1 + B^1) + \alpha(A^2 + B^2), (1 - \alpha)(A^4 + B^4) + \alpha(A^3 + B^3)],$$

lo que justifica que $F + G$ es el número difuso trapezoidal $(A^1 + B^1/A^2 + B^2/A^3 + B^3/A^4 + B^4)$. Podemos entonces afirmar que la suma de números difusos trapezoidales se realiza componente a componente.

$$(A^1/A^2/A^3/A^4) + (B^1/B^2/B^3/B^4) = (A^1 + B^1/A^2 + B^2/A^3 + B^3/A^4 + B^4).$$

Si los números trapezoidales vienen expresados a partir de sus respectivos centros y de sus amplitudes, también puede demostrarse que:

$$\text{Tra}(A^c, A^m, A^\ell, A^s) + \text{Tra}(B^c, B^m, B^\ell, B^s) = \text{Tra}(A^c + B^c, A^m + B^m, A^\ell + B^\ell, A^s + B^s).$$

Dado que los números triangulares son un caso particular de los números trapezoidales, también es cierto que, entre números triangulares, se verifica:

$$\begin{aligned} (A^1/A^2/A^3) + (B^1/B^2/B^3) &= (A^1 + B^1/A^2 + B^2/A^3 + B^3) \quad \text{y} \\ \text{Tri}(A^c, A^\ell, A^s) + \text{Tri}(B^c, B^\ell, B^s) &= \text{Tri}(A^c + B^c, A^\ell + B^\ell, A^s + B^s) \end{aligned} \quad (1.17)$$

1.4. Espacios métricos en los que interviene la idea de incertidumbre

Habiendo desarrollado ya una buena parte de los preliminares algebraicos que utilizaremos, pasamos a comentar las diferentes concepciones de *espacio métrico* en los que la idea de incertidumbre está presente.

1.4.1. Espacios métricos probabilísticos (de Schweizer y Sklar)

La siguiente definición fue dada por Schweizer y Sklar basándose en una idea previa de Šerstnev.

Definición 1.4.1 *Un espacio métrico probabilístico (abreviadamente, EMP) es cualquier terna (X, d, τ) donde X es un conjunto no vacío, $d : X \times X \rightarrow \Delta^+$ es una aplicación y τ es una función triangular de manera que se cumplen las siguientes propiedades:*

- Para cada $x, y \in X$ se tiene que $d(x, x) = \epsilon_0$. (1.18)

- Si $x, y \in X$ son diferentes, entonces $d(x, y) \neq \epsilon_0$. (1.19)

- Para cada $x, y \in X$ se tiene que $d(x, y) = d(y, x)$. (1.20)

- Para cada $x, y, z \in X$ se tiene que $d(x, z) \geq \tau(d(x, y), d(y, z))$. (1.21)

En tal caso, diremos que (X, d) es un EMP bajo τ . Diremos que (X, d, τ) es:

- propio si la función triangular τ verifica la condición (1.11);
- un espacio de Menger si existe una t -norma $*$ de manera que $\tau = \tau_*$.

Obsérvese que un EsMP no necesita de ninguna función triangular. Por otro lado, que un EMP sea propio sólo depende de la función triangular τ , y no de la distancia d . Los siguientes ejemplos nos aseguran que cualquier espacio métrico posee, de manera natural, una estructura de EMP propio.

Ejemplo 1.4.2 Sea (X, \mathbf{d}) un EM y sea τ cualquier función triangular que cumpla (1.11). Definamos $d : X \times X \rightarrow \Delta^+$ como:

$$d(x, y) = \epsilon_{\mathbf{d}(x, y)}, \quad \text{para cada } x, y \in X.$$

Entonces (X, d, τ) es un EMP. En este caso, la imagen de cada $d(x, y)$ está contenida en $\{0, 1\}$ (estos espacios se suelen denominar crisp).

Ejemplo 1.4.3 Otra forma de ver un EM (X, \mathbf{d}) como un EMP consiste en definir

$$d(p, q)(t) = \begin{cases} 0, & \text{si } t \leq 0, \\ \frac{t}{t + \mathbf{d}(x, y)}, & \text{si } t > 0 \end{cases}$$

Entonces (X, d) es un EMP bajo la función triangular producto T_{Π} .

La desigualdad triangular (1.21) se vuelve curiosa en este tipo de espacios, pues intercambia su sentido tradicional. Veamos cómo puede interpretarse este cambio. Sean p, q y r tres puntos de un espacio métrico (X, \mathbf{d}) de manera que la distancia (métrica) entre p y q es menor o igual que la distancia (métrica) entre p y r , es decir, como espacio métrico, $\mathbf{d}(p, q) \leq \mathbf{d}(p, r)$. Entonces, como EMP, el suceso aleatorio en el que la distancia probabilística entre p y q es menor o igual que x contiene al suceso aleatorio en el que la distancia probabilística entre p y r es menor o igual que x , es decir,

$$[\mathbf{d}(p, r) \leq x \Rightarrow \mathbf{d}(p, q) \leq x] \Rightarrow \{\omega : d(p, r) \leq x\} \subseteq \{\omega : \mathbf{d}(p, q) \leq x\},$$

y así, la probabilidad del primero es menor o igual que la probabilidad del segundo:

$$P(\{\mathbf{d}(p, r) \leq x\}) \leq P(\{\mathbf{d}(p, q) \leq x\}).$$

Esto intercambia el sentido común que manejamos espacialmente: si la distancia (métrica) entre p y q es menor o igual que la distancia (métrica) entre p y r , entonces la función de distribución asociada a p y q es mayor o igual que la correspondiente a p y r ; o dicho de otra forma, la distancia probabilística entre p y q es mayor o igual que entre p y r .

$$\mathbf{d}(p, q) \leq \mathbf{d}(p, r) \Rightarrow d(p, q) \geq d(p, r).$$

Por consiguiente, el sentido de la desigualdad triangular tiene que cambiar necesariamente, y así:

$$d(p, q) \boxed{\leq} d(p, r) + d(r, q) \Leftrightarrow d(p, q) \boxed{\geq} \tau(d(p, r), d(r, q)),$$

donde τ será una aplicación que transforme un par de funciones de distribución en otra función de distribución.

1.4.2. Espacios métricos difusos de Kaleva y Seikkala

Se dice que un número difuso $F : \mathbb{R} \rightarrow \mathbb{R}$ es *no negativo* si $F(t) = 0$ para cada $t < 0$ (véase [67]). Denotaremos por \mathcal{G} al conjunto formado por todos los números difusos no negativos.

Definición 1.4.4 Llamaremos espacio métrico difuso en el sentido de Kaleva y Seikkala (abreviadamente, un EKS) a un cuádruple (X, d, L, R) donde $d : X \times X \rightarrow \mathcal{G}$ es una aplicación (la métrica difusa), $L, R : \mathbb{I}^2 \rightarrow \mathbb{I}$ son aplicaciones simétricas y no decrecientes en cada argumento que verifican $L(0, 0) = 0$ y $R(1, 1) = 1$, y las siguientes condiciones.

- (i) $d_{xy} = \bar{0}$ si, y sólo si, $x = y$.
- (ii) $d_{xy} = d_{yx}$ para cada $x, y \in X$.
- (iii) Para cada $x, y, z \in X$, si denotamos por $[d_{xy}]_\alpha = \{t \in \mathbb{R} : d_{xy}(t) \geq \alpha\} = [\lambda_{xy}^\alpha, \rho_{xy}^\alpha]$ siendo $0 < \alpha \leq 1$:
 - (a) $d_{xy}(s+t) \geq L(d_{xz}(s), d_{zy}(t))$ siempre que $s \leq \lambda_{xz}^1$, $t \leq \lambda_{zy}^1$ y $s+t \leq \lambda_{xy}^1$.
 - (b) $d_{xy}(s+t) \leq R(d_{xz}(s), d_{zy}(t))$ siempre que $s \geq \lambda_{xz}^1$, $t \geq \lambda_{zy}^1$ y $s+t \geq \lambda_{xy}^1$.

1.4.3. Espacios métricos difusos de Kramosil y Michálek

Como se ha comentado en la Introducción de la presente memoria, en su artículo [74], Kramosil y Michálek razonaron la necesidad de modificar la noción de *espacio métrico estadístico* debida a Menger que se venía utilizando hasta el momento, e introdujeron una clase muy amplia de espacios a los que llamaron *espacios métricos difusos*. A lo largo del tiempo y buscando aplicaciones prácticas, los axiomas de Kramosil y Michálek han ido cambiando (véase [42, 47, 39]). De esta forma, modernamente, se considera que un *espacio métrico difuso* en el sentido de Kramosil y Michálek se ajusta a las siguientes propiedades.

Definición 1.4.5 Llamaremos espacio métrico difuso en el sentido de Kramosil y Michálek [74] (abreviadamente, un EKM) a una terna $(X, M, *)$ donde X es un conjunto no vacío, $*$ es una

t -norma continua y $M : X \times X \times [0, \infty[\rightarrow \mathbb{I}$ es un conjunto difuso que verifica las siguientes propiedades.

- Para cada $x, y \in X$ se tiene que $M_{xy}(0) = 0$. (1.22)

- Para cada $x, y \in X$ y cada $t \in]0, \infty[$ se tiene que $M_{xy}(t) = M_{yx}(t)$. (1.23)

- Si $x, y \in X$, entonces $x = y$ si, y sólo si, $M_{xy}(t) = 1$ para cada $t \in]0, \infty[$. (1.24)

- Para cada $x, y \in X$ y cada $s, t \in [0, \infty[$ se tiene que

$$M_{xz}(t + s) \geq M_{xy}(t) * M_{yz}(s). \quad (1.25)$$

- Para cada $x, y \in X$ la aplicación $M_{xy}(\cdot) : [0, \infty[\rightarrow \mathbb{I}$ es continua a la izquierda. (1.26)

Ejemplo 1.4.6 Cada $EM(X, \mathbf{d})$ es un $EKM(X, M^s, \cdot)$, llamado espacio métrico difuso estándar asociado a (X, \mathbf{d}) , donde \cdot denota la t -norma producto sobre \mathbb{I} y M^s se define como sigue para cada $x, y \in X$:

$$M_{xy}^s(0) = 0 \quad \text{y} \quad M_{xy}^s(t) = \frac{t}{t + \mathbf{d}(x, y)}, \quad \text{si } t \in]0, \infty[.$$

Además, esta forma de ver un espacio métrico como un EKM verifica que

$$\lim_{t \rightarrow \infty} M_{xy}^s(t) = 1 \quad \text{para cada } x, y \in X, \quad (1.27)$$

la cual es una propiedad ya considerada en su definición original por Kramosil y Michálek.

Tomando $z = y$ en el axioma (1.25), utilizando el axioma (1.23) y teniendo en cuenta que 1 es la unidad de cualquier t -norma, observamos que, para cada $x, y \in X$ y cada $t, s \in [0, \infty[$:

$$M_{xy}(t + s) \geq M_{xy}(t) * M_{yy}(s) = M_{xy}(t) * 1 = M_{xy}(t).$$

Lema 1.4.7 Si $(X, M, *)$ es un EKM entonces cada función $M_{xy}(\cdot) : [0, \infty[\rightarrow \mathbb{I}$ es no decreciente, sean cuales sean $x, y \in X$.

1.4.4. Espacios métricos difusos de George y Veeramani

George y Veeramani modificaron ligeramente la noción de EKM e introdujeron una topología Hausdorff que verifica el primer axioma de numerabilidad. En sus espacios, también consiguieron demostrar una versión del *teorema de Baire*.

Definición 1.4.8 Un espacio métrico difuso en el sentido de George y Veeramani [42] (abreviadamente, un EGV) es una terna $(X, M, *)$ donde X es un conjunto no vacío, $*$ es una t -norma

continua y $M : X \times X \times]0, \infty[\rightarrow \mathbb{I}$ es un conjunto difuso que verifica las siguientes propiedades.

- Para cada $x, y \in X$ y cada $t \in]0, \infty[$ se tiene que $M_{xy}(t) > 0$. (1.28)

- Para cada $x, y \in X$ y cada $t \in]0, \infty[$ se tiene que $M_{xy}(t) = M_{yx}(t)$. (1.29)

- Si $x, y \in X$, entonces $x = y$ si, y sólo si, $M_{xy}(t) = 1$ para cada $t \in]0, \infty[$. (1.30)

- Para cada $x, y \in X$ y cada $s, t \in]0, \infty[$ se tiene que

$$M_{xz}(t + s) \geq M_{xy}(t) * M_{yz}(s). \quad (1.31)$$

- Para cada $x, y \in X$ la aplicación $M_{xy}(\cdot) :]0, \infty[\rightarrow \mathbb{I}$ es continua. (1.32)

Evidentemente, todo EGV es un EKM con tal de definir $M_{xy}(0) = 0$ para cada $x, y \in X$. Además, cualquier espacio métrico es también, de forma estándar, un EGV.

Ejemplo 1.4.9 (\mathbb{R}, M, \cdot) es un EGV si \cdot es la t -norma producto y

$$M_{xy}(t) = e^{-\frac{|x-y|}{t}} \quad \text{para cada } x, y \in \mathbb{R} \text{ y cada } t \in]0, \infty[.$$

Ejemplo 1.4.10 Si \cdot es la t -norma producto y fijamos $k, m, n \in]0, \infty[$, entonces cada espacio métrico (X, d) es un EGV (X, M, \cdot) si definimos:

$$M_{xy}(t) = \frac{kt^n}{kt^n + m d(x, y)} \quad \text{para cada } x, y \in X \text{ y cada } t \in]0, \infty[.$$

Ejemplo 1.4.11 Si \cdot es la t -norma producto y $X = \mathbb{N}$, entonces (X, M, \cdot) es un EGV si definimos:

$$M_{xy}(t) = \begin{cases} x/y, & \text{si } x \leq y, \\ y/x, & \text{si } y \leq x \end{cases} \quad \text{para cada } x, y \in \mathbb{N} \text{ y cada } t \in]0, \infty[.$$

Además, el último ejemplo no proviene de un espacio métrico de la forma estándar, ya que no se cumple la condición (1.27).

Una de las razones que motivaron la modificación de George y Veeramani fue la necesidad de buscar una estructura difusa con la que \mathbb{R} , dotado de su métrica difusa estándar, fuese completo. Por ello, afinaron también las nociones de *sucesión convergente* y *sucesión de Cauchy* presentando las siguientes propuestas.

Definición 1.4.12 Sea $\{x_n\}$ una sucesión de un EGV (X, M) (bajo alguna t -norma). Diremos que $\{x_n\}$ es una sucesión de Cauchy si para cada $\varepsilon \in]0, 1[$ y cada $t > 0$ existe un número natural $n_0 \in \mathbb{N}$ de manera que si $n, m \in \mathbb{N}$ verifican $n, m \geq n_0$ entonces $M_{x_n x_m}(t) > 1 - \varepsilon$.

Se dice que la sucesión $\{x_n\}$ converge a un punto $x \in X$ si para cada $\varepsilon \in]0, 1[$ y cada $t > 0$ existe un número natural $n_0 \in \mathbb{N}$ de manera que si $n \in \mathbb{N}$ verifica $n \geq n_0$ entonces $M_{x_n x}(t) > 1 - \varepsilon$.

Se dice que $(X, M, *)$ es completo si toda sucesión de Cauchy es convergente en X .

1.4.5. Espacios métricos difusos intuicionistas

Inspirándose en la definición de George y Veermani, y teniendo en cuenta que existen fenómenos en los que es difícil determinar si ocurre un suceso o su contrario, en 2004 Park introdujo en [97] una definición de espacio métrico difuso intuicionista. No obstante, para atrapar un mayor número de espacios, seguimos la definición que dieron Castro-Company y Romaguera en [16].

Definición 1.4.13 *Diremos que $(X, M, N, *, \diamond)$ es un espacio métrico difuso intuicionista [16] (abreviadamente, un EMDI) si $*$ es una t -norma continua, \diamond es una t -conorma continua y $M, N : X \times X \times [0, \infty[\rightarrow \mathbb{I}$ son conjuntos difusos tales que, para cada $x, y, z \in X$, se cumplen las siguientes propiedades.*

$$\bullet \quad M_{xy}(t) + N_{xy}(t) \leq 1 \text{ para cada } t \in [0, \infty[. \quad (1.33)$$

$$\bullet \quad M_{xy}(0) = 0. \quad (1.34)$$

$$\bullet \quad \text{Para } t \in]0, \infty[\text{ se tiene que } M_{xy}(t) = M_{yx}(t). \quad (1.35)$$

$$\bullet \quad x = y \text{ si, y sólo si, } M_{xy}(t) = 1 \text{ para cada } t \in]0, \infty[. \quad (1.36)$$

$$\bullet \quad \text{Para cada } s, t \in [0, \infty[\text{ se tiene que } M_{xz}(t + s) \geq M_{xy}(t) * M_{yz}(s). \quad (1.37)$$

$$\bullet \quad \text{Cada aplicación } M_{xy}(\cdot) : [0, \infty[\rightarrow \mathbb{I} \text{ es continua a la izquierda.} \quad (1.38)$$

$$\bullet \quad N_{xy}(0) = 1. \quad (1.39)$$

$$\bullet \quad \text{Para } t \in]0, \infty[\text{ se tiene que } N_{xy}(t) = N_{yx}(t). \quad (1.40)$$

$$\bullet \quad x = y \text{ si, y sólo si, } N_{xy}(t) = 0 \text{ para cada } t \in]0, \infty[. \quad (1.41)$$

$$\bullet \quad \text{Para cada } s, t \in [0, \infty[\text{ se tiene que } N_{xz}(t + s) \leq N_{xy}(t) \diamond N_{yz}(s). \quad (1.42)$$

$$\bullet \quad \text{Cada aplicación } N_{xy}(\cdot) : [0, \infty[\rightarrow \mathbb{I} \text{ es continua a la izquierda.} \quad (1.43)$$

Evidentemente, si $(X, M, N, *, \diamond)$ es un EMDI, entonces $(X, M, *)$ es un EKM, denominado su *espacio soporte*. Recíprocamente, si $(X, M, *)$ es un EKM, entonces $(X, M, 1 - M, *, *)$ es un EMDI.

1.4.6. Espacios métricos probabilísticos intuicionistas

Una *función de distribución de tipo no-distancia* es una función $h : [0, \infty] \rightarrow \mathbb{I}$ es una función no creciente, continua a la izquierda en $]0, \infty[$ tal que $h(0) = 1$ y $h(\infty) = 0$. Esta definición puede extenderse a $\overline{\mathbb{R}}$ con tal de definir $h(t) = 1$ para cada $t < 0$. Denotaremos por ∇^+ al conjunto formado por todas las funciones de distribución de tipo no-distancia. Su elemento mínimo, con el orden parcial punto a punto, es la función $\tilde{\epsilon}_0 : \overline{\mathbb{R}} \rightarrow \mathbb{I}$ definida de la siguiente manera: $\tilde{\epsilon}_0(t) = 1$ si $t \leq 0$ y $\tilde{\epsilon}_0(t) = 0$ si $t > 0$.

Definición 1.4.14 *Un co-espacio métrico probabilístico (abreviadamente, un cEMP) es una terna $(X, \tilde{d}, \tilde{\tau})$ donde $\tilde{d} : X \times X \rightarrow \nabla^+$ es una aplicación y $\tilde{\tau} : \nabla^+ \times \nabla^+ \rightarrow \nabla^+$ es una función*

triangular sobre ∇^+ (cuyo elemento unidad es $\tilde{\epsilon}_0$) que verifica las siguientes propiedades.

- Para cada $x, y \in X$ se tiene que $\tilde{d}(x, y) = \tilde{\epsilon}_0$. (1.44)

- Si $x, y \in X$ son diferentes, entonces $\tilde{d}(x, y) \neq \tilde{\epsilon}_0$. (1.45)

- Para cada $x, y \in X$ se tiene que $\tilde{d}(x, y) = \tilde{d}(y, x)$. (1.46)

- Para cada $x, y, z \in X$ se tiene que $\tilde{d}(x, z) \leq \tilde{\tau}(\tilde{d}(x, y), \tilde{d}(y, z))$. (1.47)

Definición 1.4.15 Diremos que $(X, d, \tau, \tilde{d}, \tilde{\tau})$ es un espacio métrico probabilístico intuicionista (abreviadamente, un EMPI) si (X, d, τ) es un EMP, $(X, \tilde{d}, \tilde{\tau})$ es un cEMP y $d(x, y)(t) + \tilde{d}(x, y)(t) \leq 1$ para cada $x, y \in X$ y cada $t \in]0, \infty[$.

Ejemplo 1.4.16 Cualquier EMP (X, d, τ) es un EMPI de la forma $(X, d, \tau, \tilde{d}, \tilde{\tau})$, donde

$$\tilde{d}(x, y)(t) = 1 - d(x, y)(t) \quad y \quad \tilde{\tau}(\tilde{f}, \tilde{g})(t) = 1 - \tau(1 - \tilde{f}, 1 - \tilde{g})(t),$$

para cada $x, y \in X$, cada $\tilde{f}, \tilde{g} \in \nabla^+$ y cada $t \in [0, \infty]$.

Ejemplo 1.4.17 Cualquier espacio métrico (X, \mathbf{d}) es un espacio métrico probabilístico intuicionista si consideramos la función $d : X \times X \times \rightarrow \Delta^+$ dada, para cada $p, q \in X$ y cada $t \in [0, \infty]$, de la siguiente forma: $d(p, q)(0) = 0 = \tilde{d}(p, q)(\infty)$, $d(p, q)(\infty) = 1 = \tilde{d}(p, q)(0)$ y

$$d(p, q)(t) = \frac{t}{t + \mathbf{d}(p, q)}, \quad \tilde{d}(p, q)(t) = \frac{\mathbf{d}(p, q)}{t + \mathbf{d}(p, q)} \quad \text{si } t \in]0, \infty[.$$

Si $\tau(f, g)(t) = f(t)g(t)$ es la multiplicación usual en $[0, 1]$ y $\tilde{\tau}(\tilde{f}, \tilde{g})(t) = \min(1, \tilde{f}(t) + \tilde{g}(t))$, entonces $(X, d, \tau, \tilde{d}, \tilde{\tau})$ es un espacio métrico probabilístico intuicionista.

1.4.7. Espacios métricos \mathfrak{L} -difusos

Se pueden definir conjuntos en contextos más generales. Goguen [44] generalizó en el siguiente sentido. Si tenemos un conjunto parcialmente ordenado $\mathfrak{L} = (L, \leq_L)$, un conjunto difuso es una función de un conjunto a L . En ese sentido, un conjunto difuso es una especie de función característica generalizada. Saadati, Razani y Adibi utilizaron en [106] el concepto de *espacio métrico \mathfrak{L} -difuso* para demostrar ciertos teoremas de existencia de punto fijo. En dicho trabajo pueden encontrarse las siguientes definiciones y propiedades.

Sea $\mathfrak{L} = (L, \leq_L)$ un retículo completo, es decir, un conjunto parcialmente ordenado en el que todo subconjunto no vacío posee un ínfimo (\wedge) y un supremo (\vee). Llamemos $0_{\mathfrak{L}} = \inf L$ a su mínimo absoluto y llamemos $1_{\mathfrak{L}} = \sup L$ a su máximo absoluto. Sea X un conjunto no vacío denominado *universo*. Un *conjunto \mathfrak{L} -difuso sobre X* es cualquier aplicación $\mathfrak{A} : X \rightarrow L$. Para cada $x \in X$, el valor $\mathfrak{A}(x)$ representa el grado (en L) en que x satisface \mathfrak{A} .

Ejemplo 1.4.18 *Un ejemplo de retículo completo es \mathfrak{L}^* dado por:*

$$L^* = \left\{ (x, y) \in [0, 1]^2 : x + y \leq 1 \right\},$$

$$(x_1, y_1) \leq_{L^*} (x_2, y_2) \iff x_1 \leq x_2 \text{ e } y_1 \geq y_2,$$

cuyo máximo absoluto es $1_{\mathfrak{L}^} = (1, 0)$ y cuyo mínimo absoluto es $0_{\mathfrak{L}^*} = (0, 1)$.*

Una *negación* sobre \mathfrak{L} es cualquier aplicación decreciente $N : L \rightarrow L$ que satisface $N(0_{\mathfrak{L}}) = 1_{\mathfrak{L}}$ y $N(1_{\mathfrak{L}}) = 0_{\mathfrak{L}}$. Se dice que N es una *negación involutiva* si $N(N(x)) = x$ para cada $x \in X$.

Es sencillo extender (también) el concepto de t-norma a un retículo \mathfrak{L} . Una *norma triangular* (o *t-norma*) sobre \mathfrak{L} es cualquier aplicación $\mathfrak{J} : L^2 \rightarrow L$ que sea asociativa, conmutativa, no decreciente en cada argumento y que satisface la condición de frontera $\mathfrak{J}(\ell, 1_{\mathfrak{L}}) = \ell$ para cada $\ell \in L$. Una t-norma puede extenderse a cualquier número de argumentos (mayor o igual que dos) si se define, por recurrencia,

$$\mathfrak{J}^1 = \mathfrak{J} \quad \text{e} \quad \mathfrak{J}^n(\ell_1, \ell_2, \dots, \ell_{n+1}) = \mathfrak{J}(\mathfrak{J}^{n-1}(\ell_1, \ell_2, \dots, \ell_n), \ell_{n+1})$$

para cada $n \geq 2$ y cada $\ell_1, \ell_2, \dots, \ell_{n+1} \in L$.

Una t-norma \mathfrak{J} sobre \mathfrak{L}^* se dice *continua a izquierda* si

$$\lim_{n \rightarrow \infty} \mathfrak{J}(x_n, y) = \mathfrak{J}\left(\lim_{n \rightarrow \infty} x_n, y\right),$$

para cualquier sucesión no creciente x_n en L y cada $y \in L$. Análogamente tenemos el concepto de continuidad a derecha y de continuidad.

Definición 1.4.19 *Una t-norma \mathfrak{J} sobre \mathfrak{L}^* se dice t-representable si existen una t-norma T y una t-conorma S sobre $[0, 1]$ tal que, para cada $x = (x_1, y_1), y = (x_2, y_2) \in \mathfrak{L}^*$,*

$$\mathfrak{J}(x, y) = (T(x_1, x_2), S(y_1, y_2)).$$

Con la misma idea, dados dos conjuntos difusos cualquiera $M : X \rightarrow \mathbb{I}, N : X \rightarrow \mathbb{I}$, podemos construir un conjunto \mathfrak{L}^* -difuso como

$$\mathfrak{M}_{M,N}(x) = (M(x), N(x)).$$

Definición 1.4.20 *Dada una negación N , diremos que la t-norma \mathfrak{J} es de tipo Hadžić (abreviadamente, de tipo H) si para cada $\varepsilon \in L \setminus \{0_{\mathfrak{L}}, 1_{\mathfrak{L}}\}$ existe $\mu \in L \setminus \{0_{\mathfrak{L}}, 1_{\mathfrak{L}}\}$ tal que*

$$\mathfrak{J}^n(N(\mu), N(\mu), \dots, N(\mu)) >_L N(\varepsilon) \quad \text{para cada } n \geq 1.$$

Por supuesto, el concepto anterior depende de la negación seleccionada.

Definición 1.4.21 Una terna $\mathfrak{L} = (L, \leq_L, \otimes)$ es un retículo completo residuado si $\mathfrak{L} = (L, \leq_L)$ es un retículo completo, $\mathfrak{L} = (L, \otimes)$ es un monoide conmutativo y \otimes es distributivo bajo uniones arbitrarias, i.e.,

$$\left(\bigvee_{i \in \Gamma} l_i \right) \otimes l = \bigvee_{i \in \Gamma} (l_i \otimes l).$$

Una t-norma sobre \mathfrak{L} continua a izquierda es el retículo completo residuado $(L, \leq_L, \mathfrak{I})$.

Definición 1.4.22 Llamaremos espacio métrico \mathfrak{L} -difuso a cualquier terna $(X, \mathfrak{M}, \mathfrak{I})$ tal que X es un conjunto no vacío, \mathfrak{I} es una t-norma continua sobre \mathfrak{L} y $\mathfrak{M} : X \times X \times]0, \infty[\rightarrow \mathfrak{L}$ es un conjunto \mathfrak{L} -difuso de manera que se cumplen las siguientes propiedades para cada $x, y, z \in X$ y cada $s, t \in]0, \infty[$.

- $\mathfrak{M}_{xy}(t) >_L 0_{\mathfrak{L}}$. (1.48)

- $\mathfrak{M}_{xy}(t) = \mathfrak{M}_{yx}(t)$. (1.49)

- $x = y$ si, y sólo si, $\mathfrak{M}_{xy}(t) = 1$ para cada $t \in]0, \infty[$. (1.50)

- $\mathfrak{M}_{xz}(t + s) \geq_L \mathfrak{I}(\mathfrak{M}_{xy}(t), \mathfrak{M}_{yz}(s))$. (1.51)

- Cada aplicación $\mathfrak{M}_{xy}(\cdot) :]0, \infty[\rightarrow L$ es continua. (1.52)

Cuando el retículo considerado es \mathfrak{L}^* , podemos expresar $\mathfrak{M} = \mathfrak{M}_{M,N}$ (siendo $M, N : X \times X \times]0, \infty[\rightarrow \mathbb{I}$ conjuntos difusos) y la t-norma es representable, entonces el espacio corresponde exactamente con la noción de *espacio métrico difuso intuicionista*.

Sea un espacio métrico \mathfrak{L} -difuso $(X, \mathfrak{M}, \mathfrak{I})$ y una negación N sobre L . Dado $t > 0$, podemos definir una bola abierta centrada en $x \in X$ y radio $r \in L \setminus \{0_{\mathfrak{L}}, 1_{\mathfrak{L}}\}$ como:

$$B(x, r, t) = \{y \in X : \mathfrak{M}_{xy}(t) >_L N(r)\}.$$

Un subconjunto $A \subset X$ se dice *abierto* si para cada $x \in A$, existen $t > 0$ y $r \in L \setminus \{0_{\mathfrak{L}}, 1_{\mathfrak{L}}\}$ tales que $B(x, r, t) \subset A$. Con estos abiertos podemos definir una topología inducida por el espacio métrico \mathfrak{L} -difuso y la negación.

Definición 1.4.23 Sea $\{x_n\}$ una sucesión de un espacio métrico \mathfrak{L} -difuso (X, \mathfrak{M}) bajo alguna t-norma \mathfrak{I} sobre L y sea $N : L \rightarrow L$ una negación sobre L . Diremos que $\{x_n\}$ es una sucesión de Cauchy si para cada $\varepsilon \in L \setminus \{0_{\mathfrak{L}}\}$ y cada $t > 0$ existe un número natural $n_0 \in \mathbb{N}$ de manera que si $n, m \in \mathbb{N}$ verifican $n, m \geq n_0$ entonces $\mathfrak{M}_{x_n x_m}(t) >_L N(\varepsilon)$.

Se dice que la sucesión $\{x_n\}$ converge a un punto $x \in X$, y se denota por $\{x_n\} \rightarrow x$, si para cada $t > 0$ se tiene que $\{\mathfrak{M}_{x_n x}(t)\} \rightarrow 1_{\mathfrak{L}}$ cuando $n \rightarrow \infty$.

Se dice que (X, \mathfrak{M}) es N -completo si toda sucesión de Cauchy es convergente en X .

Definición 1.4.24 Dado un espacio métrico \mathfrak{L} -difuso $(X, \mathfrak{M}, \mathfrak{J})$, se dice que la aplicación \mathfrak{M} es continua en un punto $(x, y, t) \in X \times X \times]0, \infty[$ si para cada par de sucesiones $\{x_n\}$ e $\{y_n\}$ de X y cada sucesión $\{t_n\}$ de $]0, \infty[$ tales que $\{\mathfrak{M}_{x_n x}(t)\} \rightarrow 1_{\mathfrak{L}}$, $\{\mathfrak{M}_{y_n y}(t)\} \rightarrow 1_{\mathfrak{L}}$ y $\{\mathfrak{M}_{xy}(t_n)\} \rightarrow \mathfrak{M}_{xy}(t)$, se cumple que $\{\mathfrak{M}_{x_n y_n}(t_n)\} \rightarrow \mathfrak{M}_{xy}(t)$.

Se dice que \mathfrak{M} es continua si es continua en cada punto de $X \times X \times]0, \infty[$.

1.4.8. Espacios normados difusos

Los espacios normados difusos añaden la condición adicional de trabajar sobre un espacio vectorial y, por ello, también hay una condición sobre el producto por escalares.

Definición 1.4.25 [109] Un espacio normado difuso (abreviadamente, un END) es una terna $(X, \mu, *)$ donde X es un espacio vectorial, $*$ es una t -norma continua y $\mu : X \times]0, \infty[\rightarrow \mathbb{I}$ es un conjunto difuso de manera que se verifican las siguientes propiedades:

- Para cada $x \in X$ y cada $t > 0$ se tiene que $\mu(x, t) > 0$. (1.53)

- Si $x \in X$, entonces $x = 0$ si, y sólo si, $\mu(x, t) = 1$ para cada $t > 0$. (1.54)

- Para cada $x \in X$, cada $t > 0$ y cada $\alpha \neq 0$ se tiene que

$$\mu(\alpha x, t) = \mu\left(x, \frac{t}{|\alpha|}\right). \quad (1.55)$$

- Para cada $x, y \in X$ y cada $t, s > 0$ se tiene que

$$\mu(x, t) * \mu(y, s) \leq \mu(x + y, t + s). \quad (1.56)$$

- Para cada $x \in X$, la función $\mu(x, \cdot) :]0, \infty[\rightarrow \mathbb{I}$ es continua. (1.57)

- Para cada $x \in X$ se tiene que $\lim_{t \rightarrow \infty} \mu(x, t) = 1$ y $\lim_{t \rightarrow 0} \mu(x, t) = 0$. (1.58)

Tomando $y = 0$ en el axioma (1.56) y aplicando la propiedad (1.54), deducimos la siguiente propiedad.

Lema 1.4.26 Si $(X, \mu, *)$ es un END y $x \in X$, entonces $\mu(x, \cdot) :]0, \infty[\rightarrow \mathbb{I}$ es una función no decreciente en $]0, \infty[$.

Definición 1.4.27 Sea $(X, \mu, *)$ un END y sea $\{x_n\}$ una sucesión de puntos de X .

- Diremos que $\{x_n\} \subset X$ es una sucesión de Cauchy si para cada $\varepsilon > 0$ y cada $t > 0$ existe un número natural $n_0 \in \mathbb{N}$ tal que si $n, m \in \mathbb{N}$ verifican que $n, m \geq n_0$, entonces $\mu(x_n - x_m, t) > 1 - \varepsilon$.
- Diremos que la sucesión $\{x_n\}$ converge a $x \in X$, y lo denotaremos por $\{x_n\} \rightarrow x$ o por $\lim_{n \rightarrow \infty} x_n = x$, si para cada $\varepsilon > 0$ y cada $t > 0$ existe un número natural $n_0 \in \mathbb{N}$ tal que si $n \in \mathbb{N}$ verifica que $n \geq n_0$, entonces $\mu(x_n - x, t) > 1 - \varepsilon$.

- Diremos que un END $(X, \mu, *)$ es completo si cada sucesión de Cauchy es convergente en X .

1.4.9. Espacios normados difusos intuicionistas

Utilizando t-normas y t-conormas continuas, Saadati y Park [107] introdujeron el siguiente concepto.

Definición 1.4.28 Diremos que $(X, \mu, \nu, *, \diamond)$ es un espacio normado difuso intuicionista (abreviadamente, un ENDI) si X es un conjunto no vacío, $*$ es una t-norma continua, \diamond es una t-conorma continua y $\mu, \nu : X \times]0, \infty[\rightarrow \mathbb{I}$ son conjuntos difusos de manera que se verifican las siguientes propiedades:

- Para cada $x \in X$ y cada $t > 0$ se tiene que $\mu(x, t) + \nu(x, t) \leq 1$. (1.59)

- Para cada $x \in X$ y cada $t > 0$ se tiene que $\mu(x, t) > 0$ y que $\nu(x, t) < 1$. (1.60)

- Si $x \in X$, entonces $x = 0$ si, y sólo si, $\mu(x, t) = 1$ para cada $t > 0$, lo cual ocurre si, y sólo si, $\nu(x, t) = 0$ para cada $t > 0$. (1.61)

- Para cada $x \in X$, cada $t > 0$ y cada $\alpha \neq 0$ se tiene que

$$\mu(\alpha x, t) = \mu\left(x, \frac{t}{|\alpha|}\right) \quad \text{y} \quad \nu(\alpha x, t) = \nu\left(x, \frac{t}{|\alpha|}\right). \quad (1.62)$$

- Para cada $x, y \in X$ y cada $t, s > 0$ se tiene que

$$\mu(x, t) * \mu(y, s) \leq \mu(x + y, t + s) \quad \text{y} \quad \nu(x, t) \diamond \nu(y, s) \leq \varpi(x + y, t + s). \quad (1.63)$$

- Para cada $x \in X$, las funciones $\mu(x, \cdot), \nu(x, \cdot) :]0, \infty[\rightarrow \mathbb{I}$ son continuas. (1.64)

- Para cada $x \in X$ se tiene que $\lim_{t \rightarrow \infty} \mu(x, t) = \lim_{t \rightarrow 0} \nu(x, t) = 1$ y

$$\lim_{t \rightarrow 0} \mu(x, t) = \lim_{t \rightarrow \infty} \nu(x, t) = 0. \quad (1.65)$$

Obviamente, si $(X, \mu, \nu, *, \diamond)$ es un ENDI, entonces $(X, \mu, *)$ es un END. Nos referiremos a éste como su *soporte*.

Lema 1.4.29 Si $(X, \mu, \nu, *, \diamond)$ es un ENDI y $x \in X$, entonces $\mu(x, \cdot)$ es una función no decreciente en $]0, \infty[$ y $\nu(x, \cdot)$ es una función no creciente en $]0, \infty[$.

Podemos encontrar algunas propiedades de los ENDI, varios ejemplos de normas difusas intuicionistas y los conceptos de convergencia y sucesión de Cauchy en esta clase de espacios en [107].

Definición 1.4.30 [107] Sea $(X, \mu, \nu, *, \diamond)$ un ENDI y sea $\{x_n\}$ una sucesión de puntos de X .

- Diremos que $\{x_n\} \subset X$ es una sucesión de Cauchy si, para cada $\varepsilon > 0$ y cada $t > 0$ existe un número natural $n_0 \in \mathbb{N}$ tal que si $n, m \in \mathbb{N}$ verifican que $n, m \geq n_0$, entonces $\mu(x_n - x_m, t) > 1 - \varepsilon$ y $\nu(x_n - x_m, t) < \varepsilon$.
- Diremos que la sucesión $\{x_n\}$ converge a $x \in X$, y lo denotaremos por $\{x_n\} \rightarrow x$ o por $\lim_{n \rightarrow \infty} x_n = x$, si para cada $\varepsilon > 0$ y cada $t > 0$ existe un número natural $n_0 \in \mathbb{N}$ tal que si $n \in \mathbb{N}$ verifica que $n \geq n_0$, entonces $\mu(x_n - x, t) > 1 - \varepsilon$ y $\nu(x_n - x, t) < \varepsilon$.
- Diremos que un ENDI $(X, \mu, \nu, *, \diamond)$ es completo si cada sucesión de Cauchy es convergente en X .

1.5. Puntos fijos de aplicaciones

Debido a sus importantes aplicaciones, uno de los campos que más interés ha despertado desde su creación es el de la búsqueda y determinación de puntos fijos de aplicaciones. En general, un *punto fijo de una aplicación* $\varphi : X \rightarrow X$ es un punto $x_0 \in X$ tal que $\varphi(x_0) = x_0$. Sin lugar a dudas, el resultado más conocido en este campo es el *teorema del punto fijo de Banach* [5], que afirma que cualquier aplicación contractiva de un espacios métrico completo sobre sí mismo posee un único punto fijo. Si (X, d) es un espacio métrico, llamamos *aplicación contractiva* a cualquier aplicación $\varphi : X \rightarrow X$ con la propiedad de que existe un número real $k \in [0, 1[$ de manera que

$$d(\varphi(x), \varphi(y)) \leq k d(x, y) \quad \text{para cada } x, y \in X.$$

El éxito de este teorema no sólo radica en su potencia, sino en que su demostración nos permite describir dicho punto fijo como el límite de una sucesión de Cauchy que se define, por recurrencia, haciendo actuar la aplicación φ una y otra vez.

Desde la publicación del trabajo pionero de Banach, se han buscado condiciones de contractividad suficientes en muy distintos espacios métricos (todos ellos completos) para asegurar la existencia (y, a veces, unicidad) de puntos fijos de aplicaciones. Los espacios métricos probabilísticos, difusos e intuicionistas no han sido ajenos a esta búsqueda incesante, obteniéndose paulatinamente resultados cada vez más generales que, por supuesto, generalizan el teorema de Banach.

La siguientes definiciones pueden encontrarse en [76], donde se enuncia y demuestra un teorema de existencia de puntos de coincidencia dobles.

Definición 1.5.1 Sean $F : X \times X \rightarrow X$ y $g : X \rightarrow X$ dos aplicaciones.

- Diremos que F y g conmutan si $g(F(x, y)) = F(gx, gy)$ para cada $x, y \in X$.

- Llamaremos punto de coincidencia doble de las aplicaciones F y g a un punto $(x, y) \in X \times X$ tal que $F(x, y) = g(x)$ y $F(y, x) = g(y)$.

Sea ahora (X, \sqsubseteq) un conjunto parcialmente ordenado.

- Diremos que la aplicación $F : X \times X \rightarrow X$ posee la propiedad g -monótona mezclada si verifica las dos siguientes condiciones:

$$\begin{aligned} x_1, x_2 \in X, gx_1 \sqsubseteq gx_2 &\Rightarrow F(x_1, y) \sqsubseteq F(x_2, y), \quad \text{para cada } y \in X, \\ y_1, y_2 \in X, gy_1 \sqsubseteq gy_2 &\Rightarrow F(x, y_1) \supseteq F(x, y_2), \quad \text{para cada } x \in X. \end{aligned}$$

Si g es la aplicación identidad sobre X , diremos que F posee la propiedad monótona mezclada.

- Diremos que (X, \sqsubseteq) posee la propiedad g -monótona secuencial si verifica las dos siguientes condiciones:

- Si $\{x_n\} \subseteq X$ es una sucesión no decreciente y $\lim_{n \rightarrow \infty} x_n = x$, entonces $gx_n \sqsubseteq gx$ para cada $n \in \mathbb{N}$.
- Si $\{x_n\} \subseteq X$ es una sucesión no creciente y $\lim_{n \rightarrow \infty} x_n = x$, entonces $gx_n \supseteq gx$ para cada $n \in \mathbb{N}$.

Si g es la aplicación identidad sobre X , diremos que (X, \leq) posee la propiedad monótona secuencial.

Definición 1.5.2 Sean X e Y dos ENDI. Diremos que una función $f : X \rightarrow Y$ es continua en un punto $x_0 \in X$ si para cada sucesión $\{x_n\}$ de puntos de X que converja hacia $x_0 \in X$, se tiene que la sucesión de imágenes $\{f(x_n)\}$ en Y converge a $f(x_0)$. Diremos que f es continua si es continua en cada punto de X .

1.6. El modelo de regresión probabilística

El problema general de regresión consiste en encontrar la relación entre una variable dependiente (también llamada de salida, endógena o respuesta) y y un conjunto de variables independientes $\underline{x} = (x_1, \dots, x_p)$ (también llamadas de entrada, exógenas o explicativas). En general, si $p = 1$, es decir, existe una sola variable explicativa, el problema se denomina como regresión simple y si $p > 1$, es decir, existe más de una variable explicativa, hablaremos de regresión múltiple.

En realidad, las relaciones que observamos entre variables suelen caracterizarse asumiendo que el conocimiento de una o varias variables permite inferir, en mayor o menor grado, el valor

de otra. Se dice entonces que existe una relación de dependencia estadística (o estocástica) entre ellas. En este sentido, los métodos de regresión estudian la construcción de un modelo matemático (a través de una función) que explique esta relación de dependencia de una variable y respecto a las variables x_1, \dots, x_p . El objetivo del análisis de regresión es predecir el comportamiento de la variable respuesta, conocido el comportamiento de las variables explicativas que influyen sobre ella.

Formalmente, dado un conjunto de datos $\{\underline{x}_i, y_i\}_{i=1}^n$ obtenido a partir de (\underline{x}, y) , donde $\underline{x}_i \in \mathbb{R}^p$ e y_i es el valor de respuesta de la variable y correspondiente a $\underline{x}_i \in \mathbb{R}^p$, y dada una función $f(x, \underline{a})$, el objetivo es encontrar el vector de parámetros \underline{a} tal que

$$y_i = f(\underline{x}_i, \underline{a}) \quad i = 1, 2, \dots, n.$$

La Regresión Probabilística se resuelve definiendo una función pérdida que mida los errores de predicción entre y_i y $f(\underline{x}_i, \underline{a})$. La elección habitual de la función de pérdida es la norma L_p

$$L_p = |y - f(\underline{x}, \underline{a})|^p \tag{1.66}$$

donde p es un número positivo. Cuando consideramos $p = 2$ en la función de pérdida, obtenemos un *ajuste de mínimos cuadrados*. El criterio L_1 da más importancia a las colas que el criterio L_2 , ya que este último, al elevar al cuadrado, resta importancia a los valores pequeños. La función L_1 es invariante frente a cambios de escala monótonos y continuos, lo que hace que ésta sea más fácil de interpretar que L_2 . En situaciones prácticas (y como se muestra posteriormente en los experimentos de simulación planteados en el último Capítulo), los estimadores que optimizan estos criterios son similares, excepto para situaciones extremas.

Lo más habitual es que f sea una función lineal. En caso de no serlo, en algunas ocasiones el modelo se puede *linealizar* mediante transformaciones de las variables. El modelo clásico de Regresión Lineal, para cada observación, asume el modelo lineal siguiente:

$$y_i = \beta_0 + \beta_1 x_{1i} + \dots + \beta_p x_{pi} + \varepsilon_i \quad i = 1, 2, \dots, n$$

donde los parámetros β_i son desconocidos y los residuos aleatorios ε_i deben cumplir las siguientes propiedades.

1. El valor esperado de los residuos es cero: $E[\varepsilon_i] = 0$, para todo $i = 1, 2, \dots, n$.
2. La varianza de los residuos es un valor constante: $\text{Var}[\varepsilon_i] = \sigma^2$ para todo $i = 1, 2, \dots, n$. Esta propiedad se conoce como *homocedasticidad*.
3. Hay una ausencia de autocorrelación entre los términos aleatorios de los distintos elementos de la muestra: $\text{Cov}(\varepsilon_i, \varepsilon_j) = 0$, para cualquier $i \neq j$ ($i, j = 1, \dots, n$).

El modelo lineal general se puede expresar en notación matricial como sigue:

$$y = \underline{x}\beta + \varepsilon$$

donde y es el vector de dimensión n que contiene las observaciones de la variable respuesta, \underline{x} es la matriz de dimensión $n \times (p + 1)$ que contiene las observaciones de la variable explicativa y una columna de unos y β es el vector $p + 1$ que contiene a los parámetros desconocidos que queremos estimar, esto es:

$$\begin{pmatrix} y_1 \\ y_2 \\ \vdots \\ y_n \end{pmatrix} = \begin{pmatrix} 1 & x_{11} & \cdots & x_{1p} \\ 1 & x_{21} & \cdots & x_{2p} \\ \vdots & \vdots & \ddots & \vdots \\ 1 & x_{n1} & \cdots & x_{np} \end{pmatrix} \begin{pmatrix} \beta_0 \\ \beta_1 \\ \vdots \\ \beta_p \end{pmatrix} + \begin{pmatrix} \varepsilon_1 \\ \varepsilon_2 \\ \vdots \\ \varepsilon_n \end{pmatrix}$$

Para encontrar una solución al problema anterior, podemos considerar una función pérdida como la considerada en la ecuación (1.66). Sin embargo, el método de estimación por mínimos cuadrados es el más utilizado y consiste en minimizar la suma de los residuos al cuadrado, es decir, minimizar la expresión

$$\text{SCR} = \sum_{i=1}^n \varepsilon_i = \varepsilon^t \varepsilon = (y - \underline{x}\beta)^t (y - \underline{x}\beta) = y^t y - 2y^t \underline{x}\beta + \beta^t (\underline{x}^t \underline{x}) \beta.$$

Derivando con respecto a β e igualando a cero obtenemos la expresión

$$\hat{\beta} = (\underline{x}^t \underline{x})^{-1} \underline{x}^t y. \quad (1.67)$$

El estimador de mínimos cuadrados verifica ser el estimador insesgado de mínima varianza (lo que se conoce como *eficiente*). Este resultado, que se conoce como *Teorema de Gauss-Markov*, hace que el método de mínimos cuadrados sea muy utilizado. Sin embargo, este resultado no garantiza que la varianza de los estimadores sea necesariamente pequeña. El Teorema de Gauss-Markov asegura que los estimadores mínimo-cuadráticos son los mejores dentro de la clase de estimadores que son insesgados y de las funciones lineales de las observaciones, pero no garantiza que estos estimadores sean mejores que otros estimadores que no pertenezcan a la clase anterior. Esto significa que podría existir un estimador sesgado que tuviera menor varianza que el calculado por mínimos cuadrados.

El estimador para la varianza desconocida σ^2 es

$$\hat{\sigma}^2 = \frac{1}{n + p - 1} \sum_{i=1}^n (y_i - \hat{y}_i)^2,$$

donde los valores \hat{y}_i se obtienen utilizando la solución dada en (1.67). El *error estándar de regresión* es la raíz cuadrada de $\hat{\sigma}^2$ y tiene las mismas unidades que la variable respuesta. Dicha medida representa la desviación estándar de los valores de y sobre la recta de regresión estimada

y puede utilizarse como indicador del grado de ajuste del modelo de regresión, aunque al tener la mismas unidades de medida que la variable endógena no nos va a permitir comparar la bondad del ajuste de dos modelos con variable endógena diferente. Además, debido a que $\hat{\sigma}^2$ depende de los residuos del modelo, cualquier violación de las hipótesis de los errores del modelo o cualquier especificación errónea del mismo pueden dañar gravemente a $\hat{\sigma}^2$ como estimador de σ^2 . Por esta razón, diremos que $\hat{\sigma}^2$ es un estimador de σ^2 dependiente del modelo.

Utilizando la ecuación lineal del modelo y la estimación calculada para el vector de parámetros desconocidos, podemos hacer estimaciones puntuales y_i (predicciones) de y , para un valor \underline{x}_i determinado:

$$\hat{y}_i = \hat{\beta}_0 + \hat{\beta}_1 x_{1i} + \dots + \hat{\beta}_p x_{pi}.$$

Como cada estimación está sujeta a un margen de error, se puede calcular un intervalo de confianza. Para obtener estos intervalos y resolver contrastes de hipótesis para estudiar el comportamiento de los parámetros, se supone que los residuos siguen una determinada distribución. Por lo general, se supone que la distribución es Normal estándar:

$$\varepsilon_i \rightsquigarrow \mathcal{N}(0, 1).$$

En estas condiciones, se puede comprobar que

$$\hat{\beta} \rightsquigarrow \mathcal{N}\left(\beta, \sigma^2 (\underline{x}^t \underline{x})^{-1}\right).$$

A partir de esta distribución se puede obtener un intervalo de confianza al $(1 - \alpha) 100$ por ciento para los valores de β y resolver contrastes de hipótesis sobre este parámetro. Para resolver un problema de regresión es necesario un número mínimo de observaciones y, a medida que este número aumenta, también aumentan los grados de libertad, por lo que los intervalos de confianza se hacen más precisos.

La evaluación global de una recta de regresión puede hacerse mediante la varianza residual, que es un índice de la precisión del modelo. El modelo ajustado será poco representativo cuando la varianza residual sea grande (las diferencias entre los valores ajustados y observados, es decir, los errores, son grandes). Sin embargo, esta medida no es útil para comparar rectas de regresión de variables distintas, porque depende de las unidades de medida de la variable dependiente. Como medida más adecuada del ajuste, existe un indicador de la dispersión relativa que expresa la proporción o grado de variabilidad de la variable dependiente explicada por el modelo lineal ajustado. A este indicador se le denomina *coeficiente de determinación* y resuelve el problema de las unidades de medida. Dicho coeficiente se representa por R^2 y se calcula mediante el cociente entre la varianza explicada por la regresión y la varianza total de la variable dependiente.

El coeficiente de determinación tiene la propiedad de que sólo toma valores comprendidos entre 0 y 1 ($0 \leq R^2 \leq 1$), ya que la varianza total de la variable dependiente es igual a la suma de la varianza explicada y la varianza residual (identidad de suma de cuadrados), y es muy usual expresarlo en porcentaje. Se puede ver fácilmente que si la recta ajustada pasa por todos los

puntos observados, entonces la varianza residual es nula y $R^2 = 1$, en cuyo caso el modelo de regresión explica el 100 % de la variabilidad de Y . Así, valores de R^2 cercanos a 1 implican que la mayor parte de la variabilidad de la respuesta es explicada por el modelo de regresión. Por otro lado, si la varianza explicada por la regresión es nula, entonces $R^2 = 0$.

El coeficiente de correlación lineal de Pearson (r_{xy}) tiene por objeto medir el grado de asociación lineal entre dos variables. Puede verse que $r_{xy}^2 = R^2$. Este coeficiente toma valores entre -1 y 1, de forma que si $r_{xy} = -1$, hay una asociación lineal negativa perfecta entre las variables, si $r_{xy} = 1$, hay una asociación lineal positiva perfecta entre las variables, y si $r_{xy} = 0$, no hay asociación lineal entre las variables. Es importante tener en cuenta que el coeficiente de correlación lineal mide el grado de asociación lineal entre las variables, por lo que es un estadístico inapropiado si la relación no es lineal.

Un aspecto que se olvida frecuentemente es que los modelos de regresión se basan en hacer unas determinadas suposiciones sobre los datos y que éstas no siempre se cumplen, por lo que es preciso comprobar si las hipótesis básicas del modelo se dan en nuestros datos. Es lo que se conoce como *diagnosis del modelo*.

Tendremos que contrastar entre otras las hipótesis de linealidad, normalidad, no multicolinealidad, homocedasticidad, no autocorrelación e independencia.

Hipótesis de normalidad

La hipótesis de normalidad de los errores aleatorios es necesaria para realizar inferencias respecto a los parámetros y para la construcción de los intervalos de predicción. El supuesto de normalidad no es absolutamente esencial si el objetivo planteado consistiese únicamente en la estimación puntual de los parámetros del modelo. Como hemos visto, los estimadores mínimo cuadráticos son estimadores óptimos, con independencia de la distribución de probabilidad que siga el término de perturbación aleatoria. Si, además, la variable aleatoria se distribuye Normalmente, puede demostrarse que los estimadores mínimo cuadráticos tenderán a estarlo también, a medida que aumente indefinidamente el tamaño de la muestra.

Hipótesis de no multicolinealidad

Esta hipótesis implica que ninguna de las variables explicativas puede obtenerse como combinación lineal de las demás. Si una combinación lineal de las columnas de \underline{x} puede ser cercana a 0, el cálculo de $(\underline{x}^t \underline{x})^{-1}$ se vuelve inestable y esto produce un aumento considerable del valor absoluto de los coeficientes.

Hipótesis de homocedasticidad

El término σ^2 puede no ser constante, y variar entre las diversas observaciones, lo que se denomina problema de *heterocedasticidad*. En presencia de heterocedasticidad los estimadores mínimos cuadrados de los coeficientes de regresión siguen siendo insesgados aunque su varianza ya no será la mínima posible. Esto implica que los coeficientes de regresión tendrán errores estándares más grandes y los contrastes se verán gravemente afectados debido a una pérdida de sensibilidad.

Hipótesis de no autocorrelación e independencia

La acepción más frecuente del término autocorrelación hace referencia a la correlación existente entre los elementos de una serie de observaciones ordenadas en el tiempo o en el espacio. En el contexto de la regresión se supone que la autocorrelación no existe en los errores aleatorios. Si esto no fuera así, estaríamos ante el problema de existencia de autocorrelación. Los efectos de la autocorrelación sobre los estimadores pueden ser muy graves: los intervalos de confianza y los contrastes de hipótesis basados en las distribuciones t de Student y F de Snedecor no son ya apropiados. Tampoco lo serán los intervalos para las predicciones. Generalmente, la subestimación de la varianza proporciona al analista una falsa impresión de precisión.

Hipótesis de linealidad

La hipótesis de relación lineal entre la variable respuesta y las variables explicativas es una suposición que se realiza frecuentemente en el análisis de regresión. En ocasiones encontramos que esta suposición es inapropiada.

Una técnica muy utilizada en análisis de regresión es la *selección de variables* ya que, en muchas situaciones, se dispone de un conjunto grande de posibles variables explicativas. Una primera pregunta que puede plantearse es saber si todas las variables deben de entrar en el modelo de regresión y, en caso negativo, queremos saber qué variables deben entrar y qué variables no deben entrar. Intuitivamente parece bueno introducir en el modelo todas las variables explicativas significativas (según el contraste individual de la t) al ajustar el modelo con todas las variables posibles. Pero este procedimiento no es adecuado porque en la varianza del modelo influye el número de variables del mismo y la varianza de los parámetros estimados crece al aumentar el número de variables explicativas. Además puede haber problemas de multicolinealidad cuando hay muchas variables explicativas. Para responder a este problema se dispone de diferentes procedimientos estadísticos. Bajo la hipótesis de que la relación entre las variables dependientes y la variable respuesta es lineal existen procedimientos *paso a paso* (o *setpwise*)

que permiten elegir el subconjunto de variables regresoras que deben estar en el modelo. El algoritmo *paso a paso* tiene las ventajas del algoritmo de introducción progresiva pero lo mejora al no mantener fijas en el modelo las variables que ya entraron en una etapa, evitando de esta forma problemas de multicolinealidad. En la práctica, es un algoritmo muy utilizado ya que proporciona resultados razonables cuando se tiene un número grande de variables explicativas.

Finalmente, es preciso observar que cuando tenemos una variable dependiente dicotómica que deseamos predecir, o para la que queremos evaluar la asociación o relación con otras (más de una) variables independientes y de control, el procedimiento a realizar es una *Regresión Logística (binaria)*. La Regresión Logística es probablemente el tipo de análisis multivariante más empleado en Ciencias de la Salud. No cabe ninguna duda de que este tipo de regresión es una de las herramientas estadísticas con mejor capacidad para el análisis de datos en investigación clínica y epidemiología, y de ahí su amplia utilización. En este caso, la variable no sigue una distribución Normal, sino Binomial. Para dar solución a este problema se hace una transformación del modelo estandar, resultando el modelo logístico

$$\pi(x) = \frac{e^{x\beta}}{1 + e^{x\beta}}$$

El método de estimación que se utiliza en este caso es el de *máxima verosimilitud* que consiste en elegir los parámetros de un modelo como aquéllos que hagan máxima la probabilidad de obtener la muestra observada, de acuerdo con los datos disponibles.

Capítulo 2

Una comparación entre las diferentes concepciones de *espacio métrico difuso*

En este Capítulo abordamos las relaciones existentes entre las diferentes concepciones de *espacio métrico difuso* que se han introducido a lo largo de los últimos años y que han sido comentadas en el segundo Capítulo de la presente Memoria. En concreto, buscamos los elementos comunes a todas las teorías ya introducidas y demostramos que la estructura que realmente subyace entre las diferentes definiciones es una terna formada por un conjunto no vacío, una *función de distancia* evaluada sobre el conjunto de números difusos y una función triangular que determina la relación existente entre las distancias que pueden ser calculadas entre tres puntos arbitrarios del espacio (similar a la *desigualdad triangular*).

La mayoría de las nociones actuales de *espacio métrico difuso* interpretan la distancia entre dos puntos como una función de distribución en un entorno real. Sin embargo, en el contexto difuso, parece más coherente utilizar números difusos para este propósito. Pocos autores lo han hecho en el pasado. Si interpretamos un número difuso como una pareja de funciones de distribución de tipo distancia, veremos cómo considerar una definición de *espacio métrico difuso* que aprovecha lo mejor de todas las teorías consideradas.

2.1. Una descomposición canónica de números difusos

En esta sección vamos a interpretar los números difusos (que satisfacen la condición de normalidad en cero) como un par de funciones de distribución de tipo distancia (f.d.d.). Esta

descomposición nos permitirá trasladar la *métrica de Lévy* sobre Δ^+ a la clase formada por los mencionados números difusos. Además, se establece la relación entre las funciones triangulares de ambos conjuntos.

Definición 2.1.1 *Dado un conjunto difuso F sobre \mathbb{R} , definamos $F^-, F^+ : \overline{\mathbb{R}} \rightarrow [0, 1]$ como sigue.*

$$F^\pm(x) = \begin{cases} 0, & \text{si } x < 0, \\ 1 - F(\pm x), & \text{si } 0 \leq x < \infty, \\ 1, & \text{si } x = \infty; \end{cases}$$

El siguiente lema es una consecuencia inmediata de que la aplicación $x \mapsto -x$ es una biyección entre $] -\infty, 0[$ y $] 0, \infty[$ que invierte el orden.

Lema 2.1.2 *Un conjunto difuso F sobre \mathbb{R} es un número difuso de \mathcal{F} si, y sólo si, F^- y F^+ son funciones de distribución de tipo distancia. En tal caso, para cada $x \in \mathbb{R}$, se tiene que:*

$$F(x) = \begin{cases} 1 - F^-(-x), & \text{si } x < 0, \\ 1 - F^+(x), & \text{si } x \geq 0. \end{cases}$$

DEMOSTRACIÓN: Por un lado, $F^\pm(x) = 0$ para cada $x < 0$. Además, $F^\pm(0) = 0$ si, y sólo si, $F(0) = 1$. Sean $x, y \in] 0, \infty[$ con $x \leq y$ (lo que significa que $-y \leq -x$) Entonces:

$$\begin{aligned} F^+(x) \leq F^+(y) &\Leftrightarrow 1 - F(x) \leq 1 - F(y) \Leftrightarrow F(x) \geq F(y); \\ F^-(x) \leq F^-(y) &\Leftrightarrow 1 - F(-x) \leq 1 - F(-y) \Leftrightarrow F(-x) \geq F(-y). \end{aligned}$$

Así, F^\pm son aplicaciones no decrecientes en $] 0, \infty[$ si, y sólo si, $F|_{] 0, \infty[}$ es una aplicación no creciente y $F|_{] -\infty, 0]}$ es una aplicación no decreciente. Por otro lado, sea, en general, $\{x_n\} \subset] 0, \infty[$ una sucesión estrictamente creciente que converge a un punto $x \in] 0, \infty[$. Entonces:

$$\begin{aligned} F^\pm \text{ es continua por la izquierda en }] 0, \infty[&\Leftrightarrow \\ &\Leftrightarrow [\text{ para cada } \{x_n\} \nearrow x \text{ se verifica que } \{F^\pm(x_n)\} \rightarrow F^\pm(x)] \Leftrightarrow \\ &\Leftrightarrow [\text{ para cada } \{x_n\} \nearrow x \text{ se verifica que } \{1 - F(\pm x_n)\} \rightarrow 1 - F(\pm x)] \Leftrightarrow \\ &\Leftrightarrow [\text{ para cada } \{x_n\} \nearrow x \text{ se verifica que } \{F(\pm x_n)\} \rightarrow F(\pm x)]. \end{aligned}$$

Las condiciones anteriores son equivalentes a que $F|_{] 0, \infty[}$ sea continua por la izquierda y $F|_{] -\infty, 0]}$ sea continua por la derecha. Aplicando el Lema 1.3.2, deducimos que $F \in \mathcal{F}$ si, y sólo si, $F^\pm \in \Delta^+$. ■

Una descomposición similar fue ya obtenida por Zhang en [130]. Según el lema 2.1.2, cada número difuso (que alcance la condición de normalidad en cero) puede ser determinado a partir

de un par de f.d.d. El siguiente teorema establece que esta relación es biyectiva. En particular, demostramos que existe una topología métrica natural sobre \mathcal{F} que generaliza la ya introducida por Zhang y que hace que este espacio sea isométrico a $\Delta^+ \times \Delta^+$ con la topología producto inducida por la *métrica de Lévy*.

Antes de ello, observemos que el orden parcial de Δ^+ puede ser inducido de forma canónica en $(\Delta^+)^2$ considerando $(f_1, f_2) \leq (g_1, g_2)$ si, y sólo si, $f_1 \leq g_1$ y $f_2 \leq g_2$.

Teorema 2.1.3 La aplicación $\Phi : \mathcal{F} \rightarrow (\Delta^+)^2$, dada por

$$\Phi(F) = (F^-, F^+) \quad \text{para cada } F \in \mathcal{F},$$

es una biyección que invierte el orden.

DEMOSTRACIÓN: Según el Lema 2.1.2, Φ está bien definida, dado que $F^-, F^+ \in \Delta^+$, y es inyectiva ya que F^- y F^+ determinan F . Para demostrar que Φ es sobreyectiva, sean $f, g \in \Delta^+$ dos f.d.d. y definamos F sobre $\overline{\mathbb{R}}$ como sigue:

$$F(x) = \begin{cases} 1 - f(-x), & \text{si } x < 0, \\ 1 - g(x), & \text{si } x \geq 0. \end{cases}$$

Utilizando los mismos argumentos del Lema 2.1.2, se puede concluir que $F \in \mathcal{F}$ y que $\Phi(F) = (f, g)$. Supongamos ahora que $F, G \in \mathcal{F}$. Dado $x \in]0, \infty[$ se tiene que

$$\begin{aligned} F(x) \leq G(x) &\Leftrightarrow 1 - F(x) \geq 1 - G(x) \Leftrightarrow F^+(x) \geq G^+(x); \\ F(-x) \leq G(-x) &\Leftrightarrow 1 - F(-x) \geq 1 - G(-x) \Leftrightarrow F^-(x) \geq G^-(x). \end{aligned}$$

Por consiguiente, $F \leq G$ si, y sólo si, $F^+ \geq G^+$ y $F^- \geq G^-$, es decir, $(F^-, F^+) \geq (G^-, G^+)$, y lo anterior es equivalente a que $\Phi(F) \geq \Phi(G)$. ■

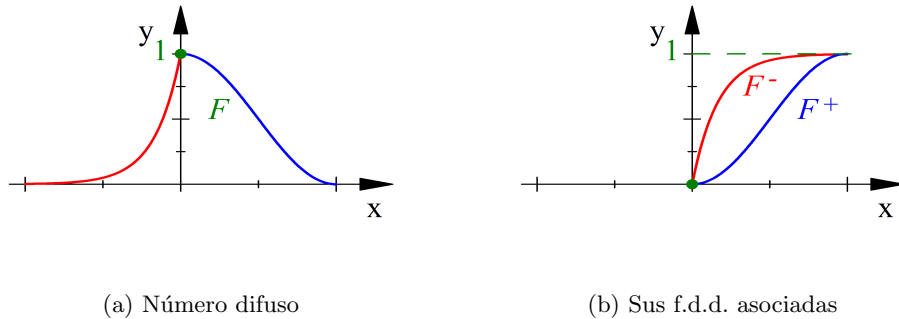


Figura 2.1: La forma de actuar de la aplicación Φ .

En la figura 2.2 tratamos de describir cómo actúa la función Φ sobre los números difusos de \mathcal{F} .

El mismo razonamiento que hemos utilizado en el Teorema anterior nos permite concluir el siguiente enunciado, en el que reconstruimos el número difuso a partir de un par de funciones de distribución de tipo distancia.

Teorema 2.1.4 Sean $f, g : [0, \infty] \rightarrow [0, 1]$ dos aplicaciones tales que $f(0) = 0$ y definamos $F : \overline{\mathbb{R}} \rightarrow [0, 1]$ como:

$$F(x) = \begin{cases} 1 - f(-x), & \text{si } x < 0, \\ 1 - g(x), & \text{si } x \geq 0. \end{cases}$$

Entonces $F \in \mathcal{F}$ si, y sólo si, $f, g \in \Delta^+$. En tal caso, $\Phi(F) = (f, g)$.

Cuando consideramos números triangulares centrados en cero del tipo $F = (-a/0/c)$, siendo $a, c > 0$, la aplicación Φ actúa transformando estos números en pares de f.d.d. asociadas a variables aleatorias uniformes, es decir, $\Phi(-a/0/c) = (U_a, U_c)$, donde $U_h(t)$ vale cero si $t < 0$, vale t/h en $[0, h]$ y vale 1 si $t > h$.

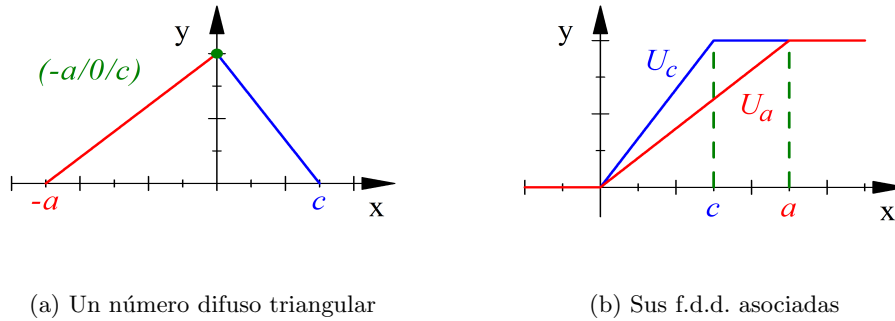


Figura 2.2: La forma de actuar de la aplicación Φ sobre números triangulares centrados en cero.

2.1.1. Una topología métrica sobre \mathcal{F}

Kaleva consideró en [65] tres clases de convergencia en un EKS, y estudió la relación entre estas concepciones. Muchos autores han estudiado también la convergencia de sucesiones de números difusos (véase [131] y las referencias que se mencionan ahí) con respecto a métricas basadas en métricas Hausdorff. A continuación, utilizando la biyección anterior, proponemos una nueva clase de convergencia.

Dado que (Δ^+, d_L) es un espacio métrico y $\Phi : \mathcal{F} \rightarrow (\Delta^+)^2$ es una biyección, es posible inducir la *métrica de Lévy* d_L sobre \mathcal{F} definiendo

$$d_L(F, G) = \max(d_L(F^-, G^-), d_L(F^+, G^+)) \quad \text{para cada } F, G \in \mathcal{F}.$$

Corolario 2.1.5 Con la métrica d_L inducida sobre \mathcal{F} por la distancia de Lévy sobre Δ^+ , \mathcal{F} es un espacio métrico completo y compacto.

DEMOSTRACIÓN: Es claro que $d_L(F, G) = 0$ si, y sólo si, $d_L(F^-, G^-) = d_L(F^+, G^+) = 0$, lo que significa que $F^- = G^-$ y $F^+ = G^+$, y esto es equivalente a decir que $F = G$. Por otro lado, $d_L(F, G) = d_L(G, F)$ claramente. Finalmente, teniendo en cuenta que, para números reales $a, b, c, d \in \mathbb{R}$ se tiene que $\max(a + b, c + d) \leq \max(a, c) + \max(b, d)$, deducimos la desigualdad triangular:

$$\begin{aligned} d_L(F, G) &= \max(d_L(F^-, G^-), d_L(F^+, G^+)) \leq \\ &\leq \max(d_L(F^-, G^-) + d_L(G^-, H^-), d_L(F^+, G^+) + d_L(G^+, H^+)) \leq \\ &\leq \max(d_L(F^-, G^-), d_L(F^+, G^+)) + \max(d_L(G^-, H^-), d_L(G^+, H^+)) = \\ &= d_L(F, G) + d_L(G, H). \end{aligned}$$

Esto significa que (\mathcal{F}, d_L) es un espacio métrico. Veamos ahora que es completo. Sea $\{F_n\}$ una sucesión de Cauchy en (\mathcal{F}, d_L) y sea $\varepsilon > 0$. Entonces existe un natural $n_0 \in \mathbb{N}$ de manera que si $n, m \geq n_0$ se tiene que $\max(d_L(F_n^-, G_n^-), d_L(F_n^+, G_n^+)) = d_L(F, G) < \varepsilon$. Así, las sucesiones $\{F_n^+\}$ y $\{F_n^-\}$ son de Cauchy en (Δ^+, d_L) . Según el apartado 1 del Lema 1.2.16, el espacio (Δ^+, d_L) es completo. Por consiguiente, existen dos f.d.d. $f, g \in \Delta^+$ de manera que $\{F_n^+\} \rightarrow f$ y $\{F_n^-\} \rightarrow g$ en (Δ^+, d_L) . Llamemos $F = \Phi^{-1}(g, f)$, es decir, $F \in \mathcal{F}$ es el único número difuso tal que $F^+ = f$ y $F^- = g$, y demostremos que $\{F_n\} \rightarrow F$ en (\mathcal{F}, d_L) . En efecto, fijemos $\varepsilon > 0$. Entonces existen $n_1, n_2 \in \mathbb{N}$ de manera que si $n \geq n_1$ se tiene que $d_L(F_n^+, F^+) = d_L(F_n^+, f) < \varepsilon$ y si $n \geq n_2$ se tiene que $d_L(F_n^-, F^-) = d_L(F_n^-, g) < \varepsilon$. Si $n_0 = \max(n_1, n_2)$ y $n \geq n_0$ deducimos que

$$d_L(F_n, F) = \max(d_L(F_n^-, F^-), d_L(F_n^+, F^+)) < \varepsilon,$$

lo que significa que $\{F_n\} \rightarrow F$ en (\mathcal{F}, d_L) . Así, (\mathcal{F}, d_L) es completo. Con la compacidad se trabaja igual sabiendo que (Δ^+, d_L) y (\mathcal{F}, d_L) son espacios metrizablees. ■

Esta distancia da lugar a una topología métrica sobre \mathcal{F} . No es difícil demostrar que si $\{F_n\}_{n \in \mathbb{N}} \subset \mathcal{F}$ es una sucesión de números difusos y $F \in \mathcal{F}$, entonces $\{d_L(F_n, F)\}_{n \in \mathbb{N}} \rightarrow 0$ si, y sólo si, $\{F_n\}_{n \in \mathbb{N}}$ converge débilmente a F , es decir, $\{F_n(x)\}_{n \in \mathbb{N}}$ converge a $F(x)$ en cada punto de continuidad de F (recuérdese el Lema 1.2.17).

2.1.2. Funciones triangulares sobre los números difusos (centrados en cero)

Dado que Φ es una biyección que invierte el orden entre \mathcal{F} y $(\Delta^+)^2$, y ϵ_0 es la identidad de cualquier función triangular sobre Δ^+ , el número difuso $\bar{0}$ debe ser la identidad de cualquier función triangular v sobre \mathcal{F} y $\mathbf{1}$ será su elemento nulo, esto es, si $v : \mathcal{F}^2 \rightarrow \mathcal{F}$ es una función triangular sobre \mathcal{F} , entonces $v(\bar{0}, F) = F$, $v(\mathbf{1}, F) = \mathbf{1}$, para cada $F \in \mathcal{F}$. Por ejemplo, la suma de números difusos es una función triangular sobre \mathcal{F} . Utilizando la terminología de [22], una función triangular sobre \mathcal{F} no es más que una $t_{\mathcal{F}}$ -conorma.

El siguiente resultado es un sencillo (pero tedioso) ejercicio puramente algebraico.

Lema 2.1.6 Existe una correspondencia biyectiva entre las funciones triangulares v sobre \mathcal{F} y las funciones triangulares η sobre $(\Delta^+)^2$ de tal manera que el siguiente diagrama conmuta.

$$\begin{array}{ccc}
 \mathcal{F}^2 & \xrightarrow{\Phi \times \Phi} & (\Delta^+)^4 \\
 \downarrow v & \equiv & \downarrow \eta \\
 \mathcal{F} & \xrightarrow{\Phi} & (\Delta^+)^2
 \end{array}
 \quad \eta \circ (\Phi \times \Phi) = \Phi \circ v$$

De esta forma, v es una función triangular sobre \mathcal{F} si, y sólo si, $\Phi \circ v \circ (\Phi \times \Phi)^{-1}$ es una función triangular sobre $(\Delta^+)^2$.

Si η es una función triangular sobre $(\Delta^+)^2$, llamaremos *función triangular sobre \mathcal{F} asociada a η* a la función triangular $v = \Phi^{-1} \circ \eta \circ (\Phi \times \Phi)$ sobre \mathcal{F} . Si τ y τ' son funciones triangulares sobre Δ^+ , entonces $\eta = \tau \times \tau'$, definida como

$$(\tau \times \tau')(f_1, f_2, f_3, f_4) = (\tau(f_1, f_3), \tau'(f_2, f_4)), \quad \text{para cada } f_1, f_2, f_3, f_4 \in \Delta^+,$$

es una función triangular sobre $(\Delta^+)^2$, y $v = \Phi^{-1} \circ (\tau \times \tau') \circ (\Phi \times \Phi)$ es su función triangular asociada sobre \mathcal{F} . En tal caso,

$$v(F, G) = \Phi^{-1}(\tau(F^-, G^-), \tau'(F^+, G^+)), \quad \text{para cada } F, G \in \mathcal{F}.$$

2.2. Un estudio comparativo de las diferentes estructuras métricas difusas y probabilísticas

En esta sección demostramos que todas las definiciones de espacios métricos difusos comentadas en el Capítulo de preliminares poseen una estructura común, de forma que pueden ser incluidos en una nueva clase de espacios.

Definición 2.2.1 *Un espacio métrico difuso (abreviadamente, un EMD) es una terna (X, \mathbf{F}, v) donde X es un conjunto, $\mathbf{F} : X \times X \rightarrow \mathcal{F}$ es una aplicación y $v : \mathcal{F} \times \mathcal{F} \rightarrow \mathcal{F}$ es una función triangular sobre \mathcal{F} de manera que se verifican las siguientes propiedades.*

- Para cada $x \in X$ se tiene que $\mathbf{F}_{xx} = \bar{0}$. (2.1)

- Si $x, y \in X$ verifican que $\mathbf{F}_{xy} \neq \bar{0}$, entonces $x = y$. (2.2)

- Para cada $x, y \in X$ se tiene que $\mathbf{F}_{yx} = \mathbf{F}_{xy}$. (2.3)

- Para cada $x, y, z \in X$ se tiene que $\mathbf{F}_{xz} \leq v(\mathbf{F}_{xy}, \mathbf{F}_{yz})$. (2.4)

Diremos que un EMD está normalizado (respectivamente, conormalizado) si $\mathbf{F}_{xy}(-t) \leq \mathbf{F}_{xy}(t)$ (respectivamente, $\mathbf{F}_{xy}(t) \leq \mathbf{F}_{xy}(-t)$) para cada $x, y \in X$ y cada $t \in]0, \infty[$.

Ejemplo 2.2.2 *Cualquier EM (X, d) es un EMD de forma natural si definimos $\mathbf{F}_{xy}(t) = 1$ para $|t| \leq d_{xy}$ y $\mathbf{F}_{xy}(t) = 0$ cuando $|t| > d_{xy}$. Si τ^+ y τ^- son funciones triangulares sobre Δ^+ que verifican (1.12) y v es la función triangular sobre \mathcal{F} asociada a $\tau^- \times \tau^+$, entonces $\mathbf{F}_{xz}^\pm \geq \tau^\pm(\mathbf{F}_{xy}^\pm, \mathbf{F}_{yz}^\pm)$.*

Ejemplo 2.2.3 *Otra forma de considerar un EM (X, d) como un EMD se consigue tomando la función triangular asociada a la t -norma producto y la métrica:*

$$\mathbf{F}_{xy}(t) = \begin{cases} 1, & \text{si } t = 0, \\ \frac{|t|}{|t| + d(x, y)}, & \text{si } t \neq 0, \end{cases} \quad \text{para cada } x, y \in X \text{ y cada } t \in \mathbb{R}.$$

A continuación presentaremos una gran variedad de EMD. Pero, por el momento, veamos que podemos considerar un EMD asociado a cada función de distribución de tipo distancia diferente de ϵ_0 y cada función triangular sobre Δ^+ .

Ejemplo 2.2.4 *Sea X un conjunto y sea $f \in \Delta^+$ cualquier función de distribución de tipo distancia diferente de ϵ_0 . Definamos $\mathbf{F}_{xy} = \bar{0}$ si $x = y$ e igualmente*

$$\mathbf{F}_{xy}(t) = \begin{cases} 0, & \text{si } t < 0, \\ 1 - f(t), & \text{si } t \geq 0, \end{cases} \quad \text{cuando } x \neq y.$$

Entonces (X, \mathbf{F}, v) es un EMD sea cual sea la función triangular v sobre \mathcal{F} asociada al producto $\tau \times \tau'$ de funciones triangulares sobre Δ^+ .

2.2.1. Estudio comparativo con los espacios métricos probabilísticos

En primer lugar, vamos a demostrar que cada EMP es un EMD para, posteriormente, introducir una topología sobre los EMDs bajo funciones triangulares continuas.

Teorema 2.2.5 *Sea X un conjunto, sea τ una función triangular sobre Δ^+ , sea $d : X \times X \rightarrow \Delta^+$ una aplicación y definamos, para cada $x, y \in X$:*

$$\mathbf{F}_{xy} = \begin{cases} 0, & \text{si } t < 0, \\ 1 - d_{xy}(t), & \text{si } t \geq 0. \end{cases}$$

Entonces las siguientes afirmaciones son equivalentes.

- (a) (X, d, τ) es un espacio métrico probabilístico.
- (b) (X, \mathbf{F}, v) es un EMD, donde v es la función triangular sobre \mathcal{F} asociada a $\tau' \times \tau$, sea cual sea la función triangular τ' sobre Δ^+ .

- (c) Existe una función triangular τ' sobre Δ^+ tal que (X, \mathbf{F}, ν) es un EMD, donde ν es la función triangular sobre \mathcal{F} asociada a $\tau' \times \tau$.

DEMOSTRACIÓN: Sea τ' cualquier función triangular sobre Δ^+ y sea ν la función triangular sobre \mathcal{F} asociada a $\tau' \times \tau$. Es claro que $d_{xy} = \epsilon_0$ si, y sólo si, $\mathbf{F}_{xy} = \bar{0}$, por lo que la condición

$$d_{xy} = \epsilon_0 \text{ si, y sólo si } x = y,$$

es equivalente a

$$\mathbf{F}_{xy} = \bar{0} \text{ si, y sólo si, } x = y.$$

Además, d es simétrica si, y sólo si, \mathbf{F} es simétrica. Observemos que $\Phi(\mathbf{F}_{xy}) = (\mathbf{F}_{xy}^-, \mathbf{F}_{xy}^+) = (\epsilon_0, d_{xy})$ para cada $x, y \in X$. Sean $x, y, z \in X$ cualesquiera tres puntos de X . Es claro que $\mathbf{F}_{xz}^- = \epsilon_0 \geq \tau'(\epsilon_0, \epsilon_0) = \tau'(\mathbf{F}_{xy}^-, \mathbf{F}_{yz}^-)$. Por otro lado, $d_{xz} \geq \tau(d_{xy}, d_{yz})$ si, y sólo si, $\mathbf{F}_{xz}^+ \geq \tau(\mathbf{F}_{xz}^+, \mathbf{F}_{zy}^+)$. Entonces:

$$d_{xz} \geq \tau(d_{xy}, d_{yz}) \Leftrightarrow \Phi(\mathbf{F}_{xy}) = (\mathbf{F}_{xy}^-, \mathbf{F}_{xy}^+) \geq (\tau'(\mathbf{F}_{xy}^-, \mathbf{F}_{yz}^-), \tau(\mathbf{F}_{xz}^+, \mathbf{F}_{zy}^+)) \Leftrightarrow \mathbf{F}_{xz} \leq \nu(\mathbf{F}_{xy}, \mathbf{F}_{yz}).$$

Se deduce que (X, d, τ) es un EMP si, y sólo si, (X, \mathbf{F}, ν) es un EMD. ■

Corolario 2.2.6 Cualquier espacio métrico probabilístico es un espacio métrico difuso.

Teniendo en cuenta que \mathbf{F}_{xy}^- no juega ningún papel importante en la demostración del Teorema 2.2.5, obtenemos el siguiente resultado.

Corolario 2.2.7 Sea (X, \mathbf{F}, ν) un EMD tal que ν es una función triangular asociada a un producto $\tau' \times \tau$ de funciones triangulares sobre Δ^+ y tal que $\mathbf{F}_{xy}^+ = \epsilon_0$ implica que $\mathbf{F}_{xy} = \bar{0}$. Entonces $(X, d = 1 - \mathbf{F}|_{[0, \infty]}, \tau)$ es un EMP.

DEMOSTRACIÓN: Dado que $\mathbf{F}_{xx} = \bar{0}$, tenemos que $d_{xx} = 1 - \mathbf{F}_{xx}|_{[0, \infty]} = \epsilon_0$. Si ocurriese que $d_{xy} = \epsilon_0$, entonces para cada $t > 0$ se tiene que

$$\mathbf{F}_{xy}^+(t) = 1 - \mathbf{F}_{xy}|_{]0, \infty[}(t) = 1 - (1 - d_{xy}(t)) = \epsilon_0(t) = 1.$$

Así, $\mathbf{F}_{xy}^+ = \epsilon_0$. Por hipótesis, esto implica que $\mathbf{F}_{xy} = \bar{0}$, de donde $x = y$. La simetría de d es obvia. Finalmente, dado que (X, \mathbf{F}, ν) es un EMD, donde ν es la función triangular sobre \mathcal{F} asociada a $\tau' \times \tau$, la demostración del teorema anterior nos dice que $\mathbf{F}_{xz}^+ \geq \tau(\mathbf{F}_{xz}^+, \mathbf{F}_{zy}^+)$ es equivalente a $d_{xz} \geq \tau(d_{xy}, d_{yz})$, por lo que concluimos que $(X, d = 1 - \mathbf{F}|_{[0, \infty]}, \tau)$ es un EMP. ■

Modificando ligeramente la demostración del Teorema 2.2.5, demostramos que siguiente resultado.

Teorema 2.2.8 Sea X un conjunto, sean τ y τ' funciones triangulares sobre Δ^+ , sean $d, d' : X \times X \rightarrow \Delta^+$ dos funciones y definamos, para cada $x, y \in X$ y cada $t \in \overline{\mathbb{R}}$:

$$\mathbf{F}_{xy}(t) = \begin{cases} 1 - d_{xy}(-t), & \text{si } t < 0, \\ 1 - d'_{xy}(t), & \text{si } t \geq 0. \end{cases}$$

Entonces las siguientes afirmaciones son equivalentes.

- (a) (X, d, τ) y (X, d', τ') son espacios métricos probabilísticos.
- (b) $(X, \mathbf{F}, v_{\tau \times \tau'})$ es un EMD.

En tal caso, $d = \mathbf{F}^-$ y $d' = \mathbf{F}^+$. Además, una sucesión $\{x_n\} \subseteq X$ \mathbf{F} -converge a $x \in X$ (respectivamente, es \mathbf{F} -Cauchy) si, y sólo si, $\{x_n\}$ d -converge a x y d' -converge a x (respectivamente, es d -Cauchy y d' -Cauchy).

2.2.2. La topología fuerte sobre los EMD

El Teorema 2.2.5 nos muestra que la relación entre los EMPs y los EMDs es muy estrecha. Cuando (X, d, τ) es un espacio métrico probabilístico y τ es continua (respecto de la métrica de Lévy d_L), es posible considerar la *topología fuerte sobre X* (véase el Capítulo 12 de [111]). Utilizando esta idea, es posible definir una topología similar en cada EMD (X, \mathbf{F}, v) . En efecto, para cada $x \in X$ y cada $t > 0$, definamos:

$$N_x(t) = \left\{ y \in X : \mathbf{F}_{xy}(\pm t) < t \right\}.$$

Es claro que si $t > 1$, entonces $N_x(t) = X$. Además, si $0 < t_1 < t_2$, entonces $N_x(t_1) \subseteq N_x(t_2)$. Observemos que para $f \in \Delta^+$ tenemos que $f(t) > 1 - t$ si, y sólo si, $d_L(f, \epsilon_0) < t$. Entonces, para cada $x, y \in X$,

$$\begin{aligned} d_L(\mathbf{F}_{xy}, \bar{0}) < t &\Leftrightarrow \max(d_L(1 - \mathbf{F}_{xy}(\pm \cdot), \epsilon_0)) < t \Leftrightarrow \\ &\Leftrightarrow 1 - \mathbf{F}_{xy}(\pm t) > 1 - t \Leftrightarrow y \in N_x(t). \end{aligned}$$

En consecuencia, podemos expresar

$$N_x(t) = \left\{ y \in X : d_L(\mathbf{F}_{xy}, \bar{0}) < t \right\}.$$

En este contexto, se pueden repetir uno por uno los argumentos del Teorema 12.1.2 de [111] para demostrar lo siguiente.

Teorema 2.2.9 Si (X, \mathbf{F}, v) es un EMD tal que v es continua, entonces existe una topología Hausdorff sobre X tal que $\{N_x(1/n)\}_{n \in \mathbb{N}}$ es una base numerable de entornos de cada punto $x \in X$.

Por ejemplo, si (X, \mathbf{d}) es un espacio métrico, esta topología sobre X coincide con la topología métrica inducida por la distancia \mathbf{d} dado que, en este caso, si $0 < t < 1$, podemos demostrar que $\mathbf{F}_{xy}(\pm t) < t$ si, y sólo si, $\mathbf{d}_{xy} < t$. Por consiguiente, $N_x(t) = \{y \in X : \mathbf{d}_{xy} < t\} = B^{\mathbf{d}}(x, t)$.

2.2.3. Estudio comparativo con los espacios de Kaleva y Seikkala

Los espacios métricos difusos definidos por Kaleva y Seikkala son extraordinariamente complejos y contienen, de manera natural, a una amplia variedad de espacios métricos difusos en el sentido de otros/as autores/as. Por ejemplo, contienen a todos los espacios métricos difusos intuicionistas.

Lema 2.2.10 Dado cualquier espacio métrico difuso intuicionista $(X, M, N, *, \diamond)$, sea $L \equiv 0$, sea $R = *'$ la t-conorma asociada a la t-norma $*$, y sea $d : X \times X \rightarrow \mathcal{G}$ la aplicación definida, para cada $x, y \in X$, como:

$$d_{xy}(t) = \begin{cases} 0, & \text{si } t < 0, \\ 1 - M_{xy}(t), & \text{si } t \geq 0. \end{cases}$$

Entonces (X, L, R, d) es un espacio métrico difuso en el sentido de Kaleva y Seikkala.

Por el momento, no conocemos la relación exacta entre los EMD que hemos introducido en la Definición 2.2.1 y los espacios de Kaleva y Seikkala. No obstante, existe una subclase concreta en éstos especialmente sencilla pero muy útil.

Definición 2.2.11 Diremos que un EKS (X, d, L, R) es simple si $d : X \times X \rightarrow \mathcal{G} \cap \mathcal{F}$, $L = 0$ y R es una t-conorma continua. Nos referiremos a esta clase de espacios como EKS*.

Teorema 2.2.12 Todo EKS* $(X, d, L = 0, R)$ es un EMD $(X, \mathbf{F} = d, v)$, donde v es la función triangular sobre \mathcal{F} asociada a $\tau \times \tau_{R'}$ y τ es cualquier función triangular sobre Δ^+ .

DEMOSTRACIÓN: Dado que R es una t-conorma continua, R' es una t-norma continua. Sea v la función triangular sobre \mathcal{F} asociada a $\tau \times \tau_{R'}$, donde τ es cualquier función triangular sobre Δ^+ . Definamos $\mathbf{F}_{xy} = d_{xy} \in \mathcal{F}$ para cada $x, y \in X$. Es claro que \mathbf{F} verifica las condiciones (2.1), (2.2) y (2.3). Para probar (2.4), observemos que $\Phi(\mathbf{F}_{xy}) = (\mathbf{F}_{xy}^-, \mathbf{F}_{xy}^+) = (\epsilon_0, 1 - d_{xy})$ para cada $x, y \in X$. Dados $s, t \geq 0 = \lambda_1(x, y) = \lambda_1(x, z) = \lambda_1(y, z)$, tenemos que:

$$\begin{aligned} d_{xy}(s+t) \leq R(d_{xz}(s), d_{zy}(t)) &\Leftrightarrow 1 - d_{xy}(s+t) \geq 1 - R(d_{xz}(s), d_{zy}(t)) \Leftrightarrow \\ &\Leftrightarrow 1 - d_{xy}(s+t) \geq R'(1 - d_{xz}(s), 1 - d_{zy}(t)) \Leftrightarrow \mathbf{F}_{xy}^+(s+t) \geq R'(\mathbf{F}_{xz}^+(s), \mathbf{F}_{zy}^+(t)). \end{aligned}$$

Si $u = t + s$ y tomamos supremo, lo anterior demuestra que para cada $u \geq 0$:

$$\mathbf{F}_{xy}^+(u) \geq \sup \left(\left\{ R'(\mathbf{F}_{xz}^+(s), \mathbf{F}_{zy}^+(t)) \mid t + s = u, t, s \geq 0 \right\} \right) = \tau_{R'}(\mathbf{F}_{xz}^+, \mathbf{F}_{zy}^+)(u).$$

Entonces $\mathbf{F}_{xy}^+ \geq \tau_{R'}(\mathbf{F}_{xz}^+, \mathbf{F}_{zy}^+)$. Como $\mathbf{F}_{xy}^- = \epsilon_0 \geq \tau(\mathbf{F}_{xz}^-, \mathbf{F}_{zy}^-)$, hemos deducido que:

$$\begin{aligned} \Phi(\mathbf{F}_{xy}) &= (\mathbf{F}_{xy}^-, \mathbf{F}_{xy}^+) \geq (\tau(\mathbf{F}_{xz}^-, \mathbf{F}_{zy}^-), \tau_{R'}(\mathbf{F}_{xz}^+, \mathbf{F}_{zy}^+)) = (\tau \times \tau_{R'})((\mathbf{F}_{xz}^-, \mathbf{F}_{xz}^+), (\mathbf{F}_{zy}^-, \mathbf{F}_{zy}^+)) = \\ &= ((\tau \times \tau_{R'}) \circ (\Phi \times \Phi))(\mathbf{F}_{xz}, \mathbf{F}_{zy}) = (\Phi \circ v)(\mathbf{F}_{xz}, \mathbf{F}_{zy}). \end{aligned}$$

Aplicando Φ^{-1} a ambos términos y teniendo en cuenta que es una aplicación que invierte el orden, concluimos que $\mathbf{F}_{xy} \leq v(\mathbf{F}_{xz}, \mathbf{F}_{zy})$, lo que significa que (X, \mathbf{F}, v) es un EMD. ■

Corolario 2.2.13 Todo espacio simple de Kaleva y Seikkala es un EMD.

Kaleva y Seikkala [67], por un lado, y Pap [95], por otro, demostraron que cada espacio de Menger es un espacio simple de Kaleva y Seikkala, y observaron que el recíproco de esta afirmación (utilizando espacios de Kaleva y Seikkala en general) no es obvio. Como consecuencia de los Teoremas 2.2.12 y 2.2.5, nosotros deducimos el siguiente recíproco para espacios simples.

Corolario 2.2.14 Cada espacio simple de Kaleva y Seikkala (X, d, L, R) es un espacio de Menger $(X, d', \tau_{R'})$, donde $d'_{xy}(t) = 1 - d_{xy}(t)$ para cada $x, y \in X$ y cada $t \geq 0$.

2.2.4. Estudio comparativo con los espacios intuicionistas de Park

Kramosil y Michálek introdujeron en [74] una definición de espacio métrico difuso modificando los axiomas que se utilizaban hasta entonces y demostraron que pueden ser dotados de una topología Hausdorff. Esta definición es, en efecto, similar a la de espacio métrico difuso intuicionista introducida por Park [97], pero realmente es muy anterior en el tiempo. Por ello, los primeros son un caso particular de los segundos. Analizamos la relación entre los regundos y la Definición 2.2.1, y deduciremos la relación con los espacios de Kramosil y Michálek.

Teorema 2.2.15 Sea $(X, M, N, *, \diamond)$ un espacio métrico difuso intuicionista y definamos $\mathbf{F}, \mathbf{G} : X \times X \rightarrow \mathcal{F}$, para cada $x, y \in X$, como:

$$\mathbf{F}_{xy}(t) = \begin{cases} N_{xy}(-t), & \text{si } t < 0, \\ 1 - M_{xy}(t), & \text{si } t \geq 0; \end{cases} \quad \mathbf{G}_{xy}(t) = \begin{cases} 1 - M_{xy}(-t), & \text{si } t < 0, \\ N_{xy}(t), & \text{si } t \geq 0. \end{cases}$$

Sea \diamond' la t-norma asociada a \diamond y sean v y v' las funciones triangular sobre \mathcal{F} inducida por $\tau_{\diamond'} \times \tau_*$ y $\tau_* \times \tau_{\diamond'}$, respectivamente. Entonces (X, \mathbf{F}, v) es un EMD normalizado y (X, \mathbf{G}, v') es un EMD connormalizado.

DEMOSTRACIÓN: Vamos a probar, primeramente, que (X, \mathbf{F}, v) es un EMD normalizado. Dado que $*$ y \diamond' son t-normas continuas, entonces τ_* y $\tau_{\diamond'}$ son funciones triangulares sobre Δ^+ , y $v = \Phi^{-1} \circ (\tau_{\diamond'} \times \tau_*) \circ (\Phi \times \Phi)$ es una función triangular sobre \mathcal{F} . Observemos que, para cada $F, G \in \mathcal{F}$, $v(F, G) = \Phi^{-1}(\tau_{\diamond'}(F^-, G^-), \tau_*(F^+, G^+))$. Sean $x, y \in X$ dos puntos de X . Como $\mathbf{F}_{xy}(0) = 1 - M_{xy}(0) = 1$, el Teorema 2.1.4 nos asegura que \mathbf{F}_{xy} está bien definida y es un número difusode \mathcal{F} . Entonces $\mathbf{F} : X \times X \rightarrow \mathcal{F}$ es una aplicación bien definida. Dado que

$$\mathbf{F}_{xy} = \bar{0} \quad \Leftrightarrow \quad [M_{xy}(t) = 1 \text{ y } N_{xy}(t) = 0, \quad \forall t \in]0, \infty[] \quad \Leftrightarrow \quad x = y,$$

las propiedades (2.1), (2.2) y (2.3) son triviales. Para demostrar la desigualdad triangular, observamos que $\Phi(\mathbf{F}_{xy}) = (\mathbf{F}_{xy}^-, \mathbf{F}_{xy}^+) = (1 - N_{xy}, M_{xy})$. Sean $x, y, z \in X$ cualesquiera tres puntos y consideremos las funciones de distribución de tipo distancia $\tau_{\diamond'}(\mathbf{F}_{xy}^-, \mathbf{F}_{yz}^-), \tau_*(\mathbf{F}_{xy}^+, \mathbf{F}_{yz}^+) \in \Delta^+$. Comparemos estas funciones en $]0, \infty[$ con \mathbf{F}_{xz}^- y \mathbf{F}_{xz}^+ , respectivamente. Sea $u \in]0, \infty[$ un número real positivo. Si $u = t + s$, donde $t, s \in]0, \infty[$, tenemos que $M_{xz}(u) = M_{xz}(t + s) \geq M_{xy}(t) * M_{yz}(s)$. Por tanto,

$$\begin{aligned} \tau_*(\mathbf{F}_{xy}^+, \mathbf{F}_{yz}^+)(u) &= \sup \left(\left\{ M_{xy}(t) * M_{yz}(s) \mid t + s = u, t, s \geq 0 \right\} \right) \leq \\ &\leq \sup \left(\left\{ M_{xz}(t + s) \mid t + s = u, t, s \geq 0 \right\} \right) = M_{xz}(u) = \mathbf{F}_{xz}^+(u). \end{aligned}$$

En consecuencia, deducimos que $\tau_*(\mathbf{F}_{xy}^+, \mathbf{F}_{yz}^+) \leq \mathbf{F}_{xz}^+$. De la misma forma, si $u = t + s$, tenemos que $N_{xz}(u) = N_{xz}(t + s) \leq N_{xy}(t) \diamond N_{yz}(s)$, y así:

$$\begin{aligned} \tau_{\diamond'}(\mathbf{F}_{xy}^-, \mathbf{F}_{yz}^-)(u) &= \sup \left(\left\{ 1 - [(1 - \mathbf{F}_{xy}^-(t)) \diamond (1 - \mathbf{F}_{yz}^-(s))] \mid t + s = u, t, s \geq 0 \right\} \right) = \\ &= \sup \left(\left\{ 1 - N_{xy}(t) \diamond N_{yz}(s) \mid t + s = u, t, s \geq 0 \right\} \right) = \\ &= 1 - \inf \left(\left\{ N_{xy}(t) \diamond N_{yz}(s) \mid t + s = u, t, s \geq 0 \right\} \right) \leq \\ &\leq 1 - \inf \left(\left\{ N_{xz}(t + s) \mid t + s = u, t, s \geq 0 \right\} \right) = 1 - N_{xz}(u) = \mathbf{F}_{xz}^-(u). \end{aligned}$$

Dado que $\tau_*(\mathbf{F}_{xy}^+, \mathbf{F}_{yz}^+) \leq \mathbf{F}_{xz}^+$ y que $\tau_{\diamond'}(\mathbf{F}_{xy}^-, \mathbf{F}_{yz}^-) \leq \mathbf{F}_{xz}^-$, y sabiendo que Φ^{-1} es una aplicación que invierte el orden, concluimos que:

$$v(\mathbf{F}_{xy}, \mathbf{F}_{yz}) = \Phi^{-1}(\tau_{\diamond'}(\mathbf{F}_{xy}^-, \mathbf{F}_{yz}^-), \tau_*(\mathbf{F}_{xy}^+, \mathbf{F}_{yz}^+)) \geq \Phi^{-1}(\mathbf{F}_{xz}^-, \mathbf{F}_{xz}^+) = \mathbf{F}_{xz}.$$

Esto demuestra que (X, \mathbf{F}, v) es un EMD. Además, es normalizado ya que, para $t > 0$:

$$M_{xy}(t) + N_{xy}(t) \leq 1 \quad \Leftrightarrow \quad N_{xy}(t) \leq 1 - M_{xy}(t) \quad \Leftrightarrow \quad \mathbf{F}_{xy}(-t) \leq \mathbf{F}_{xy}(t). \quad (2.5)$$

■

Corolario 2.2.16 Todo espacio métrico difuso intuicionista $(X, M, N, *, \diamond)$ es un EMD normalizado y es un espacio de Menger (X, M, τ_*) .

El corolario anterior es una consecuencia de aplicar que (X, \mathbf{F}, v) es el EMD generado por $(X, M, N, *, \diamond)$ y el Teorema 2.2.5 (un espacio métrico difuso intuicionista puede ser considerado como un EMD; véase el Teorema 2.2.15). El siguiente Teorema es el enunciado recíproco del Teorema 2.2.15.

Teorema 2.2.17 Sea X un conjunto, sea $*$ una t-norma continua, sea \diamond una t-conorma continua y sea v la función triangular sobre \mathcal{F} correspondiente a $\tau_{\diamond'} \times \tau_*$. Supongamos que para cada $x, y \in X$

existen conjuntos difusos $M_{xy}, N_{xy} : [0, \infty[\rightarrow [0, 1]$ y $\mathbf{F}_{xy} : \mathbb{R} \rightarrow [0, 1]$ relacionados de la siguiente forma:

$$M_{xy}(t) = 1 - \mathbf{F}_{xy}(t) \quad \text{y} \quad N_{xy}(t) = \mathbf{F}_{xy}(-t) \quad \text{para cada } t \in [0, \infty[.$$

Entonces las siguientes afirmaciones son equivalentes.

(a) $(X, M, N, *, \diamond)$ es un espacio métrico difuso intuicionista.

(b) (X, \mathbf{F}, ν) es un EMD normalizado que verifica las siguientes propiedades.

(b.1) Si $x, y \in X$, entonces $\mathbf{F}_{xy}^- = \epsilon_0$ si, y sólo si, $x = y$, lo cual ocurre si, y sólo si, $\mathbf{F}_{xy}^+ = \epsilon_0$.

(b.2) $\mathbf{F}_{xy}(t) *' \mathbf{F}_{yz}(s) \geq \mathbf{F}_{xz}(t+s)$ para cada $x, y, z \in X$ y cada $t, s \in [0, \infty[$ (donde $*'$ es la t-conorma inducida por $*$).

(b.3) $\mathbf{F}_{xy}(-t) \diamond \mathbf{F}_{yz}(-s) \geq \mathbf{F}_{xz}(-t-s)$ para cada $x, y, z \in X$ y cada $t, s \in [0, \infty[$.

En tal caso, (X, \mathbf{F}, ν) es el EMD generado por $(X, M, N, *, \diamond)$.

DEMOSTRACIÓN: El Teorema 2.2.15 nos asegura que $(a) \Rightarrow (b)$. Recíprocamente, supongamos que (X, \mathbf{F}, ν) es un EMD normalizado que verifica las propiedades b.1, b.2 y b.3. Las condiciones 1 a 11 son sencillas de demostrar. mencionamos únicamente algunos detalles. Según (2.5), la condición de normalización es equivalente a $M_{xy} + N_{xy} \leq 1$. Los conjuntos difusos $M_{xy}, N_{xy} : [0, \infty[\rightarrow [0, 1]$ son funciones continuas a la izquierda ya que \mathbf{F}_{xy} es un número difuso (téngase en cuenta el Teorema 2.1.4). Utilizando la propiedad b.1,

$$x = y \quad \Leftrightarrow \quad \mathbf{F}_{xy}^+ = \mathbf{F}_{xy}^- = \epsilon_0 \quad \Leftrightarrow \quad M_{xy}(t) = 1 \text{ y } N_{xy}(t) = 0, \text{ para cada } t \in]0, \infty[.$$

Según b.2 tenemos que, para cada $x, y, z \in X$ y cada $t, s \in [0, \infty[$:

$$\begin{aligned} \mathbf{F}_{xy}(t) *' \mathbf{F}_{yz}(s) \geq \mathbf{F}_{xz}(t+s) &\Leftrightarrow 1 - [(1 - \mathbf{F}_{xy}(t)) * (1 - \mathbf{F}_{yz}(s))] \geq \mathbf{F}_{xz}(t+s) \Leftrightarrow \\ &\Leftrightarrow 1 - \mathbf{F}_{xz}(t+s) \geq (1 - \mathbf{F}_{xy}(t)) * (1 - \mathbf{F}_{yz}(s)) \Leftrightarrow M_{xz}(t+s) \geq M_{xy}(t) * M_{yz}(s). \end{aligned}$$

Dado que $N_{xy}(t) = \mathbf{F}_{xy}(-t)$, la condición b.3 es equivalente a $N_{xy}(t) \diamond N_{yz}(s) \geq N_{xz}(t+s)$ para cada $x, y, z \in X$ y cada $t, s \in [0, \infty[$. Por tanto, $(X, M, N, *, \diamond)$ es un espacio métrico difuso intuicionista. Finalmente, (X, \mathbf{F}, ν) es el EMD generado por $(X, M, N, *, \diamond)$ dado que:

$$\mathbf{F}_{xy}(t) = \begin{cases} N_{xy}(-t), & \text{if } t < 0, \\ 1 - M_{xy}(t), & \text{if } t \geq 0 \end{cases} \quad \text{para cada } x, y \in X.$$

■

Una versión conormalizada del Teorema anterior también es cierta si M, N y \mathbf{F} están relacionados por:

$$M_{xy}(t) = 1 - \mathbf{F}_{xy}(-t) \quad \text{y} \quad N_{xy}(t) = \mathbf{F}_{xy}(t) \quad \text{para cada } t \in [0, \infty[.$$

2.2.5. Estudio comparativo con los espacios de Kramosil y Michálek

Teniendo en cuenta que cada espacio de Kramosil y Michálek $(X, M, *)$ es un espacio métrico difuso intuicionista de la forma $(X, M, 1 - M, *, *')$, concluimos el siguiente corolario.

Corolario 2.2.18 Cada espacio métrico difuso en el sentido de Kramosil y Michálek es un EMD normalizado.

De hecho, cada espacio de Kramosil y Michálek $(X, M, *)$ es un espacio métrico difuso intuicionista de la forma $(X, M, N, *, \diamond)$, sea cual sea N y la t-norma continua \diamond siempre que se verifique que $M + N \leq 1$ y las propiedades 7–11 de tales espacios. Esto precisamente inspira el siguiente resultado.

Teorema 2.2.19 Si $(X, d, L = 0, R)$ es un espacio simple de Kaleva y Seikkala, entonces $(X, M = 1 - d|_{[0, \infty[}, N, \tau_{R'}, \diamond)$ es un espacio métrico difuso intuicionista de la forma $(X, M, N, *, \diamond)$, sea cual sea N y la t-norma continua \diamond siempre que se verifique que $M + N \leq 1$ y las propiedades 7–11 de tales espacios.

Abordamos ahora la relación entre los espacios de Kramosil y Michálek y los espacios simples de Kaleva y Seikkala.

Teorema 2.2.20 Sea X un conjunto, sea $*$ una t-norma continua y sean $M : X \times X \times [0, \infty[\rightarrow [0, 1]$ y $d : X \times X \rightarrow \text{map}(\mathbb{R}, [0, 1])$ dos aplicaciones relacionadas de la siguiente forma:

$$d_{xy}(t) = \begin{cases} 0, & \text{si } t < 0, \\ 1 - M_{xy}(t), & \text{si } t \geq 0 \end{cases} \quad \text{para cada } x, y \in X \text{ y cada } t \in \mathbb{R}.$$

Entonces $(X, M, *)$ es un espacio de Kramosil y Michálek si, y sólo si, $(X, d, L = 0, R = *')$ es un espacio simple de Kaleva y Seikkala.

DEMOSTRACIÓN: Observemos que, en un espacio simple de Kaleva y Seikkala, se verifica que $\lambda_1(x, y) = 0$, y así, para cada $x, y, z \in X$ y cada $t, s \in [0, \infty[$:

$$\begin{aligned} M_{xz}(t + s) \geq M_{xy}(t) * M_{yz}(s) &\Leftrightarrow 1 - d_{xz}(t + s) \geq (1 - d_{xy}(t)) * (1 - d_{yz}(s)) \Leftrightarrow \\ &\Leftrightarrow d_{xz}(t + s) \leq d_{xy}(t) *' d_{yz}(s). \end{aligned}$$

■

Corolario 2.2.21 Los conceptos de espacio de Menger, de Kramosil y Michálek y de espacio simple de Kaleva y Seikkala son equivalentes.

2.2.6. Estudio comparativo con el espacios métricos probabilísticos intuicionistas

Concluimos este Capítulo analizando la relación entre los espacios métricos probabilísticos intuicionistas y la Definición 2.2.1.

Teorema 2.2.22 Sea X un conjunto, sea τ una función triangular sobre Δ^+ , $\tilde{\tau}$ una función triangular sobre ∇^+ y sea v la función triangular sobre \mathcal{F} asociada a $\tau \times \tilde{\tau}'$. Sean $d : X \times X \rightarrow \Delta^+$, $\tilde{d} : X \times X \rightarrow \nabla^+$ y $\mathbf{F} : X \times X \rightarrow \mathcal{F}$ tres aplicaciones relacionadas de la siguiente forma, para cada $x, y \in X$ y cada $t \in \mathbb{R}$:

$$\mathbf{F}_{xy}(t) = \begin{cases} 1 - d_{xy}(-t), & \text{si } t < 0, \\ \tilde{d}_{xy}(t), & \text{si } t \geq 0. \end{cases}$$

Entonces $(X, d, \tau, \tilde{d}, \tilde{\tau})$ es un espacio métrico probabilístico intuicionista si, y sólo si, (X, \mathbf{F}, v) es un EMD conormalizado tal que $\mathbf{F}_{xy}|_{]-\infty, 0[} = \bar{0}|_{]-\infty, 0[} \Leftrightarrow \mathbf{F}_{xy}|_{]0, \infty[} = \bar{0}|_{]0, \infty[}$.

Como consecuencia, cualquier espacio métrico probabilístico intuicionista es un EMD.

DEMOSTRACIÓN: Es claro que $\mathbf{F}_{xy} = \bar{0}$ si, y sólo si, $d_{xy} = \epsilon_0$ y $\tilde{d}_{xy} = \tilde{\epsilon}_0$, y esta condición es equivalente a $d_{xy} = \epsilon_0$ o $\tilde{d}_{xy} = \tilde{\epsilon}_0$, por lo que las condiciones (2.1) y (2.2) son equivalentes a (1.18)-(1.19) y a (1.44)-(1.45). También es claro que \mathbf{F} es una función simétrica si, y sólo si, d y \tilde{d} también son simétricas. Observemos que, para cada $x, y \in X$ y cada $t > 0$:

$$d_{xy}(t) = 1 - \mathbf{F}_{xy}(-t) = \mathbf{F}_{xy}^-(t), \quad \tilde{d}_{xy}(t) = \mathbf{F}_{xy}(t) = 1 - \mathbf{F}_{xy}^+(t).$$

Esto significa que:

$$\begin{aligned} d_{xz} \geq \tau(d_{xy}, d_{yz}) &\Leftrightarrow \mathbf{F}_{xz}^- \geq \tau(\mathbf{F}_{xy}^-, \mathbf{F}_{yz}^-); \\ \tilde{d}_{xz} \leq \tilde{\tau}(\tilde{d}_{xy}, \tilde{d}_{yz}) &\Leftrightarrow 1 - \mathbf{F}_{xz}^+ \leq \tilde{\tau}(1 - \mathbf{F}_{xy}^+, 1 - \mathbf{F}_{yz}^+) \Leftrightarrow \\ &\Leftrightarrow 1 - \tilde{\tau}(1 - \mathbf{F}_{xy}^+, 1 - \mathbf{F}_{yz}^+) \leq \mathbf{F}_{xz}^+ \Leftrightarrow \tilde{\tau}'(\mathbf{F}_{xy}^+, \mathbf{F}_{yz}^+) \leq \mathbf{F}_{xz}^+. \end{aligned}$$

Dado que Φ es una biyección que invierte el orden y v es la función triangular sobre \mathcal{F} asociada a $\tau \times \tilde{\tau}'$:

$$\begin{aligned} \mathbf{F}_{xz} \leq v(\mathbf{F}_{xy}, \mathbf{F}_{yz}) &\Leftrightarrow \Phi(\mathbf{F}_{xz}) \geq \Phi(v(\mathbf{F}_{xy}, \mathbf{F}_{yz})) \Leftrightarrow (\mathbf{F}_{xz}^-, \mathbf{F}_{xz}^+) \geq (\tau(\mathbf{F}_{xy}^-, \mathbf{F}_{yz}^-), \tilde{\tau}'(\mathbf{F}_{xy}^+, \mathbf{F}_{yz}^+)) \Leftrightarrow \\ &\Leftrightarrow \left\{ \begin{array}{l} \mathbf{F}_{xz}^- \geq \tau(\mathbf{F}_{xy}^-, \mathbf{F}_{yz}^-) \text{ y} \\ \mathbf{F}_{xz}^+ \geq \tilde{\tau}'(\mathbf{F}_{xy}^+, \mathbf{F}_{yz}^+) \end{array} \right\} \Leftrightarrow \left\{ \begin{array}{l} d_{xz} \geq \tau(d_{xy}, d_{yz}) \text{ y} \\ \tilde{d}_{xz} \leq \tilde{\tau}(\tilde{d}_{xy}, \tilde{d}_{yz}). \end{array} \right. \end{aligned}$$

Por tanto, (2.4) es equivalente a (1.21) y (1.47). Finalmente, (X, \mathbf{F}, v) es conormalizado si, y sólo si, para cada $x, y \in X$ y cada $t \in]0, \infty[$,

$$\mathbf{F}_{xy}(t) \leq \mathbf{F}_{xy}(-t) \Leftrightarrow \tilde{d}_{xy}(t) \leq 1 - d_{xy}(t) \Leftrightarrow d_{xy}(t) + \tilde{d}_{xy}(t) \leq 1.$$

■

El siguiente diagrama resume las principales implicaciones que hemos demostrado a lo largo del presente Capítulo.

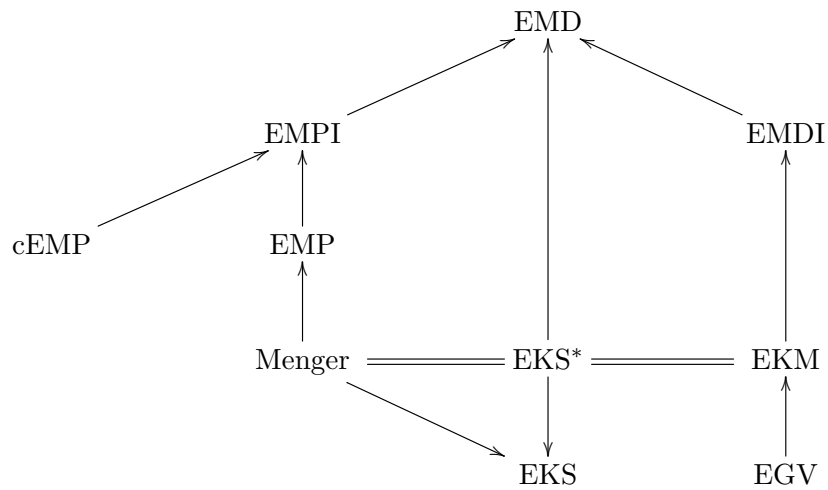


Figura 2.3: Conexiones entre las diferentes teorías.

Capítulo 3

Teoremas de punto fijo en ambientes normados y difusos

En el presente Capítulo abordamos el estudio de dos teoremas de punto fijo en ambientes diferentes.

3.1. Un teorema de contracción de Banach en espacios métricos \mathfrak{L} -difusos

En el artículo [2], Alaca, Turkoglu y Yildiz demostraron dos teoremas de punto fijo para aplicaciones conmutativas en espacios métricos difusos intuicionistas. Uno era una versión del *teorema de Banach* y otro era una versión del *teorema de Edelstein* [33]. Posteriormente, en [105], Saadati, utilizando algunas de las propiedades mostradas en [106], generalizó estos resultados a espacios métricos \mathfrak{L} -difusos (de manera que los espacios intuicionistas son un caso particular de éstos). Desgraciadamente, la demostración que aportó de su versión del teorema de Banach en este tipo de espacios era incorrecta. No obstante, damos a continuación una posible demostración de dicho teorema, explicando el error cometido por Saadati. Nos basaremos en los preliminares que hemos expuesto en la subsección 1.4.7.

Definición 3.1.1 *Llamaremos B -contracción sobre un espacio métrico \mathfrak{L} -difuso $(X, \mathfrak{M}, \mathfrak{J})$ a cualquier aplicación $A : X \rightarrow X$ para la que existe una constante $k \in [0, 1[$ de manera que*

$$\mathfrak{M}(A(x), A(y), kt) \geq_L \mathfrak{M}(x, y, t) \quad \text{para cada } x, y \in X \text{ y cada } t > 0.$$

En general, un *punto fijo* de una aplicación $A : X \rightarrow X$ es un punto $x_0 \in X$ tal que $A(x_0) = x_0$.

La siguiente condición fue ya considerada por Fang en [39] en el contexto de EKM.

Definición 3.1.2 Diremos que un espacio métrico \mathfrak{L} -difuso $(X, \mathfrak{M}, \mathfrak{J})$ verifica la condición (C) si cumple la siguiente propiedad:

$$\text{si } x, y \in X \text{ son dos puntos tales que } \mathfrak{M}(x, y, t) = C \text{ para cada } t > 0, \text{ entonces } C = 1_{\mathfrak{L}}. \quad (3.1)$$

El siguiente resultado es un *teorema de Banach para espacios métricos \mathfrak{L} -difusos*. Se trata del teorema principal dado en [105]. Para demostrarlo, sólo hay que adaptar la archiconocida técnica del teorema de Banach a este tipo de espacios introduciendo algunos ligeros cambios.

Teorema 3.1.3 Sea $(X, \mathfrak{M}, \mathfrak{J})$ un espacio métrico \mathfrak{L} -difuso N -completo que verifica la condición (3.1) y supongamos que \mathfrak{J} es de tipo Hadžić. Entonces cada B -contracción $A : X \rightarrow X$ posee un único punto fijo.

DEMOSTRACIÓN: Fijemos cualquier $x_0 \in X$ y definamos, por recurrencia, $x_1 = A(x_0)$ y $x_{n+1} = A(x_n)$ para cada $n \in \mathbb{N}$. Dado que A es una B -contracción, existe $k \in [0, 1[$ de manera que:

$$\mathfrak{M}(x_1, x_2, kt) = \mathfrak{M}(A(x_0), A(x_1), kt) \geq_L \mathfrak{M}(x_0, x_1, t).$$

Así, por inducción se demuestra que, para cada $n \in \mathbb{N}$:

$$\mathfrak{M}(x_n, x_{n+1}, k^n t) = \mathfrak{M}(A(x_{n-1}), A(x_n), k^n t) \geq_L \mathfrak{M}(x_{n-1}, x_n, k^{n-1} t) \geq_L \dots \geq_L \mathfrak{M}(x_0, x_1, t).$$

Veamos que $\{x_n\}$ es una sucesión de N -Cauchy. Fijemos $\varepsilon \in L \setminus \{0_{\mathfrak{L}}, 1_{\mathfrak{L}}\}$ y $t > 0$ arbitrarios. Por un lado, sumando la serie geométrica encontramos que, para cada $n, m \in \mathbb{N}$:

$$\sum_{i=n}^{n+m-1} k^i = k^n \frac{1 - k^{n+m}}{1 - k} \leq \frac{k^n}{1 - k}.$$

Dado que $\{k^n\} \rightarrow 0$ de forma decreciente, existe $n_0 \in \mathbb{N}$ de manera que:

$$\frac{k^{n_0}}{1 - k} \leq 1.$$

De esta forma, si $n, m \in \mathbb{N}$ y $n \geq n_0$, entonces:

$$\sum_{i=n}^{n+m-1} k^i t \leq \frac{k^n t}{1 - k} \leq \frac{k^{n_0} t}{1 - k} \leq t.$$

Dado que \mathfrak{J} es de tipo H , existe $\mu \in L \setminus \{0_{\mathfrak{L}}, 1_{\mathfrak{L}}\}$ tal que $\mathfrak{J}^m(N(\mu), N(\mu), \dots, N(\mu)) \geq_L N(\varepsilon)$ para cada $m \geq 1$.

De esta forma, si $n, m \in \mathbb{N}$ y $n \geq n_0$,

$$\begin{aligned} \mathfrak{M}(x_n, x_{n+m}, t) &\geq_L \mathfrak{M}\left(x_n, x_{n+m}, \sum_{i=n}^{n+m-1} k^i t\right) \geq_L \\ &\geq_L \mathfrak{J}^{m-1}(\mathfrak{M}(x_n, x_{n+1}, k^n t), \mathfrak{M}(x_{n+1}, x_{n+2}, k^{n+1} t), \dots, \\ &\quad \mathfrak{M}(x_{n+m-1}, x_{n+m}, k^{n+m-1} t)) \geq_L \\ &\geq_L \mathfrak{J}^{m-1}(\mathfrak{M}(x_0, x_1, t), \mathfrak{M}(x_0, x_1, t), \dots, \mathfrak{M}(x_0, x_1, t)) \geq_L \\ &\geq_L \mathfrak{J}^{m-1}(N(\mu), N(\mu), \dots, N(\mu)) \geq_L N(\varepsilon). \end{aligned}$$

Esto demuestra que $\{x_n\}$ es una sucesión N -Cauchy. Dado que $(X, \mathfrak{M}, \mathfrak{J})$ es N -completo, existe un punto $y \in X$ tal que $\{x_n\} \rightarrow y$. Veamos que y es un punto fijo de A . En efecto, por un lado, para cada $t > 0$,

$$\begin{aligned} \mathfrak{M}(A(y), y, t) &\geq_L \mathfrak{J}(\mathfrak{M}(A(y), x_{n+1}, t/2), \mathfrak{M}(x_{n+1}, y, t/2)) = \\ &= \mathfrak{J}(\mathfrak{M}(A(y), A(x_n), t/2), \mathfrak{M}(x_{n+1}, y, t/2)) \geq_L \\ &\geq_L \mathfrak{J}(\mathfrak{M}(y, x_n, t/(2k)), \mathfrak{M}(x_{n+1}, y, t/2)). \end{aligned}$$

Si tomamos límite cuando $n \rightarrow \infty$ y tenemos en cuenta que $\{x_n\} \rightarrow y$, nos damos cuenta de que

$$\{\mathfrak{J}(\mathfrak{M}(y, x_n, t/(2k)), \mathfrak{M}(x_{n+1}, y, t/2))\} \rightarrow \mathfrak{J}(1_{\mathfrak{L}}, 1_{\mathfrak{L}}) = 1_{\mathfrak{L}}.$$

Por consiguiente, $\mathfrak{M}(A(y), y, t) = 1_{\mathfrak{L}}$ para cada $t > 0$, lo que implica que $A(y) = y$. Veamos ahora que y es el único punto fijo de A . Si existiese otro punto fijo $z \in X$ de A , entonces tendríamos que

$$\mathfrak{M}(y, z, t) = \mathfrak{M}(A(y), A(z), t) \geq_L \mathfrak{M}(y, z, t/k).$$

Repitiendo este proceso, $\mathfrak{M}(y, z, t) \geq_L \mathfrak{M}(y, z, t/k) \geq_L \mathfrak{M}(y, z, t/k^2) \geq_L \dots \geq_L \mathfrak{M}(y, z, t/k^n)$ para cada $n \in \mathbb{N}$. Por otro lado, como $t < t/k < t/k^2 < \dots < t/k^n$, sabemos que $\mathfrak{M}(y, z, t) \leq_L \mathfrak{M}(y, z, t/k) \leq_L \mathfrak{M}(y, z, t/k^2) \leq_L \dots \leq_L \mathfrak{M}(y, z, t/k^n)$, de donde deducimos que

$$\mathfrak{M}(y, z, t) = \mathfrak{M}(y, z, t/k) = \mathfrak{M}(y, z, t/k^2) = \dots = \mathfrak{M}(y, z, t/k^n).$$

Como $\mathfrak{M}(y, z, \cdot)$ es una función no decreciente, deducimos que $\mathfrak{M}(y, z, t) = C$ para cada $t > 0$. Y como \mathfrak{M} verifica la condición (C) sabemos que $C = 1_{\mathfrak{L}}$, de donde $y = z$. \blacksquare

La demostración dada por Saadati en [105] falla al demostrar que la sucesión $\{x_n\}$ es de Cauchy, pues se basa en las funciones $E_{\lambda, \mathfrak{M}} : X \times X \rightarrow [0, \infty[$ definidas como sigue:

$$E_{\lambda, \mathfrak{M}}(x, y) = \inf(\{t > 0 : \mathfrak{M}_{xy}(t) > N(\lambda)\}) \quad \text{para cada } \lambda \in L \setminus \{0_{\mathfrak{L}}, 1_{\mathfrak{L}}\} \text{ y cada } x, y \in X.$$

Sin embargo, estas aplicaciones no siempre están bien definidas, pues existen espacios métricos \mathfrak{L} -difusos para los cuales no se pueden definir. En efecto, tomemos $X = [1, 2]$, $\mathfrak{L} = ([0, 1], \leq)$, $\mathfrak{J}(a, b) = ab$ y sea

$$\mathfrak{M}(x, y, t) = \frac{\min(x, y)}{\max(x, y)} \quad \text{para cada } x, y \in [1, 2] \text{ y cada } t > 0.$$

Entonces $(X, \mathfrak{M}, \mathfrak{J})$ es un espacio métrico \mathfrak{L} -difuso. Observamos que $\mathfrak{M}\left(\frac{3}{2}, \frac{8}{5}, t\right) = \frac{15}{16}$ para cada $t > 0$. Utilizando la negación involutiva $N(t) = 1 - t$ en $[0, 1]$ y $\lambda = 1/17$, tenemos que

$$\mathfrak{M}\left(\frac{3}{2}, \frac{8}{5}, t\right) = \frac{15}{16} < 1 - \frac{1}{17} \quad \text{para cada } t > 0.$$

Por consiguiente, el conjunto

$$\left\{t > 0 : \mathfrak{M}_{\frac{3}{2}, \frac{8}{5}}(t) > N\left(\frac{1}{17}\right)\right\} = \left\{t > 0 : \mathfrak{M}_{\frac{3}{2}, \frac{8}{5}}(t) > 1 - \frac{1}{17}\right\}$$

es vacío y el número $E_{\lambda, \mathfrak{M}}\left(\frac{3}{2}, \frac{8}{5}\right)$ no está bien definido.

3.2. Algunos teoremas de punto fijo en espacios normados difusos

En esta sección, enunciamos y demostramos algunos resultados de existencia y unicidad de puntos fijos de aplicaciones contractivas en el contexto de los espacios normados difusos, inspirados en versiones similares que puede encontrarse en [37]. No obstante, necesitamos reemplazar algunas de las condiciones sobre la t -norma que se imponen en los resultados de [37] por el hecho de que ésta sea de tipo Hadžić (véase la Definición 1.2.10).

Teorema 3.2.1 Sea $(X, \mu, *)$ un END completo de manera que $*$ es una t -norma de tipo Hadžić tal que $a * b \geq ab$ para cada $a, b \in \mathbb{I}$. Sean $F : X \times X \rightarrow X$ y $g : X \rightarrow X$ dos aplicaciones tales que $F(X \times X) \subseteq g(X)$, g es continua y F conmuta con g . Sea \sqsubseteq un orden parcial sobre X tal que F verifica la propiedad g -monótona mezclada y de manera que existe un número real positivo $k \in]0, 1[$ que cumple la desigualdad:

$$\mu(F(x, y) - F(u, v), kt) \geq [\mu(gx - gu, t)]^{1/2} * [\mu(gy - gv, t)]^{1/2}, \quad (3.2)$$

siempre que $gx \sqsubseteq gu$ y $gy \supseteq gv$. Supongamos que:

- (a) F es continua, o bien
- (b) (X, \sqsubseteq) posee la propiedad g -monótona secuencial.

Si existen dos puntos $x_0, y_0 \in X$ tales que $gx_0 \sqsubseteq F(x_0, y_0)$ y $gy_0 \supseteq F(y_0, x_0)$, entonces F y g poseen algún punto de coincidencia doble (es decir, existen $x, y \in X$ tales que $gx = F(x, y)$ y $gy = F(y, x)$).

DEMOSTRACIÓN: Dividimos la demostración en cinco partes.

Primera parte: construcción de las sucesiones $\{x_n\}$ e $\{y_n\}$. Consideremos los puntos $x_0, y_0 \in X$ tales que $gx_0 \sqsubseteq F(x_0, y_0)$ y $gy_0 \supseteq F(y_0, x_0)$. Dado que $F(X \times X) \subseteq g(X)$, podemos encontrar dos puntos $x_1, y_1 \in X$ tales que $gx_1 = F(x_0, y_0)$ y $gy_1 = F(y_0, x_0)$. De nuevo, dado que $F(X \times X) \subseteq g(X)$, podemos encontrar dos puntos $x_2, y_2 \in X$ tales que $gx_2 = F(x_1, y_1)$ y $gy_2 = F(y_1, x_1)$. Repitiendo este proceso, podemos encontrar dos sucesiones $\{x_n\}$ e $\{y_n\}$ de X de manera que:

$$gx_{n+1} = F(x_n, y_n) \quad \text{y} \quad gy_{n+1} = F(y_n, x_n) \quad \text{para cada } n \geq 0. \quad (3.3)$$

Segunda parte: demostramos que las sucesiones $\{gx_n\}$ y $\{gy_n\}$ son monótonas respecto de \sqsubseteq . Para ello, vamos a demostrar que:

$$gx_n \sqsubseteq gx_{n+1} \quad \text{y} \quad gy_n \supseteq gy_{n+1} \quad \text{para cada } n \geq 0, \quad (3.4)$$

es decir, $\{gx_n\}$ es una sucesión monótona creciente y $\{gy_n\}$ es una sucesión monótona decreciente (respecto del orden parcial \sqsubseteq). En efecto, procedemos por inducción. Para $n = 0$ es claro que la propiedad (3.4) es cierta debido a la condición inicial

$$gx_0 \sqsubseteq F(x_0, y_0) = gx_1 \quad y \quad gy_0 \sqsupseteq F(y_0, x_0) = gy_1.$$

Supongamos que (3.4) es cierto para algún $n \geq 0$. Dado que F verifica la propiedad g -monótona mezclada y suponemos, como hipótesis de inducción, que $gx_n \sqsubseteq gx_{n+1}$ y que $gy_n \sqsupseteq gy_{n+1}$, entonces

$$\begin{aligned} x_n, x_{n+1} \in X, gx_n \sqsubseteq gx_{n+1} &\Rightarrow F(x_n, y) \sqsubseteq F(x_{n+1}, y), \quad \text{para cada } y \in X, \\ y_n, y_{n+1} \in X, gy_{n+1} \sqsubseteq gy_n &\Rightarrow F(y_{n+1}, x) \sqsubseteq F(y_n, x), \quad \text{para cada } x \in X, \\ x_n, x_{n+1} \in X, gx_n \sqsubseteq gx_{n+1} &\Rightarrow F(y, x_n) \sqsupseteq F(y, x_{n+1}), \quad \text{para cada } y \in X, \\ y_n, y_{n+1} \in X, gy_{n+1} \sqsubseteq gy_n &\Rightarrow F(x, y_{n+1}) \sqsupseteq F(x, y_n), \quad \text{para cada } x \in X. \end{aligned}$$

De esta forma, tomando $y = y_n$ y $x = x_n$, tenemos que:

$$gx_{n+1} = F(x_n, y_n) \sqsubseteq F(x_{n+1}, y_n) \quad y \quad gy_{n+1} = F(y_n, x_n) \sqsupseteq F(y_{n+1}, x_n). \quad (3.5)$$

De igual forma

$$gx_{n+2} = F(x_{n+1}, y_{n+1}) \sqsupseteq F(x_{n+1}, y_n) \quad y \quad gy_{n+2} = F(y_{n+1}, x_{n+1}) \sqsubseteq F(y_{n+1}, x_n). \quad (3.6)$$

Combinando (3.5) y (3.6) deducimos que:

$$\begin{aligned} gx_{n+1} = F(x_n, y_n) \sqsubseteq F(x_{n+1}, y_n) \sqsubseteq F(x_{n+1}, y_{n+1}) &= gx_{n+2}, \\ gy_{n+1} = F(y_n, x_n) \sqsupseteq F(y_{n+1}, x_n) \sqsupseteq F(y_{n+1}, x_{n+1}) &= gy_{n+2}. \end{aligned}$$

Por consiguiente, la propiedad (3.4) es cierta. Llamemos ahora

$$\delta_n(t) = [\mu(gx_n - gx_{n+1}, t)]^{1/2} * [\mu(gy_n - gy_{n+1}, t)]^{1/2}, \quad \text{para cada } n \geq 0 \text{ y cada } t > 0.$$

Según el Lema (1.4.26), cada función $\mu(gx_n - gx_{n+1}, \cdot), \mu(gy_n - gy_{n+1}, \cdot) :]0, \infty[\rightarrow \mathbb{I}$ es no decreciente en $]0, \infty[$. Como la t -norma $*$ es no decreciente en cada argumento y la función $t \mapsto t^{1/2}$ es no decreciente en \mathbb{I} , deducimos que cada función δ_n es no decreciente en $]0, \infty[$. Utilizando (3.4), la condición de contractividad (3.2) y el axioma (1.55), deducimos que para cada $n \geq 1$ y cada $t > 0$ se verifica:

$$\begin{aligned} \mu(gx_n - gx_{n+1}, kt) &= \mu(F(x_{n-1}, y_{n-1}) - F(x_n, y_n), kt) \geq \\ &\geq [\mu(gx_{n-1} - gx_n, y)]^{1/2} * [\mu(gy_{n-1} - gy_n, y)]^{1/2} = \delta_{n-1}(t), \end{aligned} \quad (3.7)$$

$$\begin{aligned} \mu(gy_n - gy_{n+1}, kt) &= \mu(F(y_{n-1}, x_{n-1}) - F(y_n, x_n), kt) \geq \\ &\geq [\mu(gy_{n-1} - gy_n, y)]^{1/2} * [\mu(gx_{n-1} - gx_n, y)]^{1/2} = \delta_{n-1}(t). \end{aligned} \quad (3.8)$$

Aplicando la t-norma $*$ y teniendo en cuenta que $a * b \geq ab$ para cada $a, b \in \mathbb{I}$, deducimos que, para cada $n \geq 1$ y cada $t > 0$, se verifica:

$$\begin{aligned} \delta_n(kt) &= [\mu(gx_n - gx_{n+1}, kt)]^{1/2} * [\mu(gy_n - gy_{n+1}, kt)]^{1/2} \geq \\ &\geq [\mu(gx_n - gx_{n+1}, kt)]^{1/2} \cdot [\mu(gy_n - gy_{n+1}, kt)]^{1/2} \geq \\ &\geq \delta_{n-1}(t)^{1/2} \cdot \delta_{n-1}(t)^{1/2} = \delta_{n-1}(t). \end{aligned}$$

Repitiendo esta desigualdad, deducimos que

$$1 \geq \delta_n(t) \geq \delta_{n-1}\left(\frac{t}{k}\right) \geq \delta_{n-2}\left(\frac{t}{k^2}\right) \geq \dots \geq \delta_0\left(\frac{t}{k^n}\right). \quad (3.9)$$

Dado que $*$ es una t-norma continua y μ verifica el axioma (1.58), para cada $t > 0$, se tiene que:

$$\begin{aligned} \lim_{n \rightarrow \infty} \delta_0\left(\frac{t}{k^n}\right) &= \lim_{n \rightarrow \infty} [\mu(gx_0 - gx_1, t)]^{1/2} * [\mu(gy_0 - gy_1, t)]^{1/2} = \\ &= \left[\lim_{n \rightarrow \infty} \mu(gx_0 - gx_1, t) \right]^{1/2} * \left[\lim_{n \rightarrow \infty} \mu(gy_0 - gy_1, t) \right]^{1/2} = 1 * 1 = 1. \end{aligned}$$

Aplicando (3.9) deducimos que

$$\lim_{n \rightarrow \infty} \delta_n(t) = 1 \quad \text{para cada } t > 0.$$

Tercera parte: demostramos que las sucesiones $\{gx_n\}$ y $\{gy_n\}$ verifican, para cada $n, p \geq 1$, que

$$\mu(gx_n - gx_{n+p}, t) \geq *^p \delta_{n-1}(t - kt) \quad \text{y} \quad \mu(gy_n - gy_{n+p}, t) \geq *^p \delta_{n-1}(t - kt). \quad (3.10)$$

Recordemos que la sucesión $\{*^n a\}_{n=1}^{\infty}$ fue definida en (1.9). Para demostrar (3.10), hacemos inducción en $p \geq 1$. En efecto, para $p = 1$, podemos aplicar (3.7) y (3.8), deduciendo que:

$$\mu(gx_n - gx_{n+1}, t) \geq \delta_{n-1}(t/k) \quad \text{y} \quad \mu(gy_n - gy_{n+1}, t) \geq \delta_{n-1}(t/k).$$

Dado que $t/k \geq t \geq t - kt$ y la función δ_{n-1} es no decreciente, deducimos que

$$\begin{aligned} \mu(gx_n - gx_{n+1}, t) &\geq \delta_{n-1}(t/k) \geq \delta_{n-1}(t) \geq \delta_{n-1}(t - kt) \quad \text{y} \\ \mu(gy_n - gy_{n+1}, t) &\geq \delta_{n-1}(t/k) \geq \delta_{n-1}(t) \geq \delta_{n-1}(t - kt). \end{aligned}$$

Esto significa que (3.10) es cierto para $p = 1$, sea cual sea $n \geq 1$. Supongamos que la propiedad (3.10) es cierta para un determinado número natural $p \geq 1$ y vamos a demostrarla para $p + 1$. En efecto, por un lado, según el Lema 1.4.26 y (3.7), tenemos que:

$$\mu(gx_n - gx_{n+1}, t) \geq \mu(gx_n - gx_{n+1}, kt) \geq \delta_{n-1}(t) \geq \delta_{n-1}(t - kt). \quad (3.11)$$

Intercambiando los papeles de t y $t - kt$ en la desigualdad anterior, deducimos que

$$\mu(gx_n - gx_{n+1}, t - kt) \geq \delta_{n-1}(t - kt). \quad (3.12)$$

Aplicando la condición de contractividad (3.2), la hipótesis de inducción (3.10) para p y que la t -norma $*$ verifica que $a * b \geq ab$ para cada $a, b \in \mathbb{I}$, deducimos que, para cada $n \geq 1$ y cada $t > 0$:

$$\begin{aligned} \mu(gx_{n+1} - gx_{n+p+1}, kt) &= \mu(F(x_n, y_n) - F(x_{n+p}, y_{n+p}), kt) \geq \\ &\geq [\mu(gx_n - gx_{n+p}, t)]^{1/2} * [\mu(gy_n - gy_{n+p}, t)]^{1/2} \geq \\ &\geq [*^p \delta_{n-1}(t - kt)]^{1/2} * [*^p \delta_{n-1}(t - kt)]^{1/2} \geq \\ &\geq [*^p \delta_{n-1}(t - kt)]^{1/2} \cdot [*^p \delta_{n-1}(t - kt)]^{1/2} = *^p \delta_{n-1}(t - kt). \end{aligned} \quad (3.13)$$

Utilizando el axioma (1.56) y las desigualdades (3.11), (3.12) y 3.13):

$$\begin{aligned} \mu(gx_n - gx_{n+p+1}, t) &= \mu(gx_n - gx_{n+p+1}, t - kt + kt) \geq \\ &\geq \mu(gx_n - gx_{n+1}, t - kt) * \mu(gx_{n+1} - gx_{n+p+1}, kt) \geq \\ &\geq \delta_{n-1}(t - kt) * (*^p \delta_{n-1}(t - kt)) = *^{p+1} \delta_{n-1}(t - kt). \end{aligned} \quad (3.14)$$

Razonando de manera similar se demuestra que $\mu(gy_{n+1} - gy_{n+p+1}, kt) \geq *^{p+1} \delta_{n-1}(t - kt)$, lo que completa la inducción y demuestra que la propiedad (3.10) es cierta.

*Cuarta parte: demostramos que las sucesiones $\{gx_n\}$ y $\{gy_n\}$ son de Cauchy en $(X, \mu, *)$.* Fijemos cualesquiera $t > 0$ y $\varepsilon \in (0, 1]$. Dado que suponemos que $*$ es una t -norma de tipo Hadžić, existe un número $\eta \in]0, 1[$ tal que $*^p a > 1 - \varepsilon$ para cada $a \in]1 - \eta, 1]$ y cada $p \geq 1$. Dado que $\lim_{n \rightarrow \infty} \delta_n(t - kt) = 1$, existe un número natural n_0 tal que $\delta_n(t - kt) > 1 - \eta$ para cada $n \geq n_0$. Aplicando (3.10), deducimos que $\mu(gx_n - gx_{n+p}, t) > 1 - \varepsilon$ y $\mu(gy_n - gy_{n+p}, t) > 1 - \varepsilon$ para cada $n \geq n_0$. Esto es lo mismo que decir que las sucesiones $\{gx_n\}$ y $\{gy_n\}$ son de Cauchy en $(X, \mu, *)$.

Quinta parte: demostramos que F y g poseen algún punto de coincidencia doble. Dado que $(X, \mu, *)$ es un END completo, las sucesiones $\{gx_n\}$ y $\{gy_n\}$ son convergentes. Llamemos $x, y \in X$ a sus límites, es decir, $\lim_{n \rightarrow \infty} gx_n = x$ y $\lim_{n \rightarrow \infty} gy_n = y$, y vamos a demostrar que (x, y) es un punto de coincidencia doble de F y g . Aplicando que g es continua:

$$\lim_{n \rightarrow \infty} g g x_n = g x \quad \text{y} \quad \lim_{n \rightarrow \infty} g g y_n = g y.$$

Dado que F y g conmutan, deducimos que:

$$g g x_{n+1} = g F(x_n, y_n) = F(g x_n, g y_n) \quad \text{y} \quad g g y_{n+1} = g F(y_n, x_n) = F(g y_n, g x_n). \quad (3.15)$$

Tenemos ahora que distinguir entre las posibilidades (a) y (b) que ofrece el enunciado. Supongamos que ocurre (a), es decir, que F es continua. Tomando límite cuando $n \rightarrow \infty$ en (3.15), deducimos que:

$$g x = \lim_{n \rightarrow \infty} g g x_{n+1} = \lim_{n \rightarrow \infty} F(g x_n, g y_n) = F\left(\lim_{n \rightarrow \infty} g x_n, \lim_{n \rightarrow \infty} g y_n\right) = F(x, y),$$

e igualmente se concluye que $F(y, x) = gy$. Por tanto, en este caso, (x, y) es un punto de coincidencia doble de F y g .

Supongamos ahora que la condición (b) es cierta, es decir, que (X, \sqsubseteq) posee la propiedad g -monótona secuencial. Dado que la sucesión $\{gx_n\}$ es monótona no decreciente y $\{gx_n\} \rightarrow x$, la mencionada propiedad nos asegura que $gx_n \sqsubseteq gx$ para cada $n \geq 0$. De igual forma, dado que $\{gy_n\}$ es una sucesión monótona no creciente y $\{gy_n\} \rightarrow y$, deducimos que $gy_n \supseteq gy$ para cada $n \geq 0$. Entonces, utilizando la condición de contractividad (3.2), obtenemos:

$$\begin{aligned} \mu(ggx_{n+1} - F(x, y), kt) &= \mu(F(gx_n, gy_n) - F(x, y), kt) \geq \\ &\geq [\mu(ggx_n - gx, t)]^{1/2} * [\mu(ggy_n - gy, t)]^{1/2}. \end{aligned}$$

Tomando límite cuando $n \rightarrow \infty$ y teniendo en cuenta que $\{ggx_n\} \rightarrow gx$ y que $\{ggy_n\} \rightarrow gy$, concluimos que $\{ggx_n\} \rightarrow F(x, y)$, de donde $F(x, y) = gx$. Igualmente se razona que $F(y, x) = gy$ y, de nuevo, concluimos que (x, y) es un punto de coincidencia doble de F y g . ■

Vamos ahora a estudiar la unicidad del punto fijo garantizado en el teorema anterior. Observemos que el orden parcial \sqsubseteq sobre X nos permite inducir un orden parcial en el producto cartesiano $X \times X$ de la siguiente manera. Dados $(x, y), (u, v) \in X \times X$, definimos:

$$(x, y) \sqsubseteq (u, v) \quad \Leftrightarrow \quad x \sqsubseteq u, \quad y \supseteq v,$$

Como es usual, diremos que dos elementos (x, y) y (u, v) son *comparables* si $(x, y) \sqsubseteq (u, v)$ o $(x, y) \supseteq (u, v)$.

Teorema 3.2.2 Bajo la hipótesis del Teorema 3.2.1, supongamos que para cada par de puntos de coincidencia doble $(x, y), (x^*, y^*) \in X \times X$ de F y g existe otro punto $(u, v) \in X \times X$ tal que $(F(u, v), F(v, u))$ es comparable, a la vez, a $(F(x, y), F(y, x))$ y a $(F(x^*, y^*), F(y^*, x^*))$. Entonces F y g poseen un único punto de coincidencia doble $(x, y) \in X \times X$ que verifica

$$F(x, y) = gx = x \quad \text{y} \quad F(y, x) = gy = y.$$

El teorema anterior no afirma que F y g posean un único punto de coincidencia doble, sino que existe un único punto de coincidencia doble con la propiedad de que $F(x, y) = gx = x$ y $F(y, x) = gy = y$.

DEMOSTRACIÓN: El Teorema 3.2.1 garantiza que el conjunto de puntos de coincidencia doble de F y g es no vacío. Sean (x, y) y (x^*, y^*) cualesquiera dos puntos de coincidencia doble de F y g y vamos a demostrar que:

$$gx = gx^* \quad \text{y} \quad gy = gy^*. \quad (3.16)$$

En efecto, como (x, y) y (x^*, y^*) son dos puntos de coincidencia doble de F y g , sabemos que

$$gx = F(x, y), \quad gy = F(y, x), \quad gx^* = F(x^*, y^*) \quad \text{y} \quad gy^* = F(y^*, x^*).$$

Por hipótesis, podemos encontrar un punto $(u, v) \in X \times X$ tal que $(F(u, v), F(v, u))$ es comparable, a la vez, a $(F(x, y), F(y, x))$ y a $(F(x^*, y^*), F(y^*, x^*))$. Con este punto (u, v) , podemos construir dos sucesiones $\{u_n\}$ y $\{v_n\}$ de la siguiente manera. Llamemos $u_0 = u$ y $v_0 = v$. Dado que $F(X \times X) \subseteq g(X)$, podemos encontrar dos puntos $u_1, v_1 \in X$ tales que $gu_1 = F(u_0, v_0)$ y $gv_1 = F(v_0, u_0)$. De nuevo, dado que $F(X \times X) \subseteq g(X)$, podemos encontrar dos puntos $u_2, v_2 \in X$ tales que $gu_2 = F(u_1, v_1)$ y $gv_2 = F(v_1, u_1)$. Repitiendo este proceso, podemos encontrar dos sucesiones $\{u_n\}$ y $\{v_n\}$ de X de manera que:

$$gu_{n+1} = F(u_n, v_n) \quad \text{y} \quad gv_{n+1} = F(v_n, u_n) \quad \text{para cada } n \geq 0. \quad (3.17)$$

Con estas sucesiones vamos a demostrar que

$$gx = \lim_{n \rightarrow \infty} gu_n = gx^* \quad \text{y} \quad gy = \lim_{n \rightarrow \infty} gv_n = gy^*, \quad (3.18)$$

lo que probará que (3.16) es cierto. No obstante, hay que distinguir varios casos (todos ellos similares).

Vamos a razonar primeramente que $gx = \lim_{n \rightarrow \infty} gu_n$. En efecto, dado que los puntos $(F(x, y), F(y, x)) = (gx, gy)$ y $(F(u, v), F(v, u)) = (F(u_0, v_0), F(v_0, u_0)) = (gu_1, gv_1)$ son comparables, podemos suponer, sin pérdida de generalidad, que $gx \sqsubseteq gu_1$ y que $gy \supseteq gv_1$ (en el otro caso se razona cambiando las desigualdades). Vamos a demostrar, por inducción, que:

$$gx \sqsubseteq gu_n \quad \text{y} \quad gy \supseteq gv_n \quad \text{para cada } n \geq 1.$$

Para $n = 1$ ya lo sabemos. Supongamos que es cierto para cierto número natural n y vamos a probarlo para $n + 1$. aplicando que F verifica la propiedad g -monótona mezclada:

$$\begin{aligned} x, u_n \in X, gx \sqsubseteq gu_n &\Rightarrow \begin{cases} F(x, z) \sqsubseteq F(u_n, z), \\ F(z, x) \supseteq F(z, u_n) \end{cases} & \text{para cada } z \in X; \\ y, v_n \in X, gv_n \sqsubseteq gy &\Rightarrow \begin{cases} F(v_n, z) \sqsubseteq F(y, z), \\ F(z, v_n) \supseteq F(z, y) \end{cases} & \text{para cada } z \in X. \end{aligned}$$

De esta forma:

$$\begin{aligned} gu_{n+1} = F(u_n, v_n) &\supseteq F(x, v_n) \supseteq F(x, y) = gx, \\ gv_{n+1} = F(v_n, u_n) &\sqsubseteq F(v_n, x) \sqsubseteq F(y, x) = gy. \end{aligned}$$

Esto completa la inducción y demuestra que $gx \sqsubseteq gu_n$ y $gy \supseteq gv_n$ para cada $n \geq 1$. Aplicando ahora la condición de contractividad (3.2), obtenemos, para cada $n \geq 1$ y cada $t > 0$,

$$\mu(gx - gu_n, kt) = \mu(F(x, y) - F(u_{n-1}, v_{n-1}), kt) \geq [\mu(gx - gu_{n-1}, t)]^{1/2} * [\mu(gy - gv_{n-1}, t)]^{1/2} \quad (3.19)$$

y

$$\mu(gv_n - gy, kt) = \mu(F(v_{n-1}, u_{n-1}) - F(y, x), kt) \geq [\mu(gv_{n-1} - gy, t)]^{1/2} * [\mu(gu_{n-1} - gx, t)]^{1/2}. \quad (3.20)$$

Llamemos ahora

$$\beta_n(t) = [\mu(gx - gu_n, t)]^{1/2} * [\mu(gy - gv_n, t)]^{1/2} \quad \text{para cada } t > 0 \text{ y cada } n \geq 0.$$

Sabiendo que la t-norma $*$ es conmutativa, que verifica $a * b \geq ab$ para cada $a, b \in \mathbb{I}$, y aplicando (3.19) y (3.20), tenemos que:

$$\begin{aligned} \beta_{n+1}(kt) &= [\mu(gx - gu_{n+1}, kt)]^{1/2} * [\mu(gy - gv_{n+1}, kt)]^{1/2} \geq \\ &\geq \left([\mu(gx - gu_{n-1}, t)]^{1/2} * [\mu(gy - gv_{n-1}, t)]^{1/2} \right)^{1/2} * \\ &\quad * \left([\mu(gv_{n-1} - gy, t)]^{1/2} * [\mu(gu_{n-1} - gx, t)]^{1/2} \right)^{1/2} \geq \\ &\geq \left([\mu(gx - gu_{n-1}, t)]^{1/2} \cdot [\mu(gy - gv_{n-1}, t)]^{1/2} \right)^{1/2} \cdot \\ &\quad \cdot \left([\mu(gv_{n-1} - gy, t)]^{1/2} * [\mu(gu_{n-1} - gx, t)]^{1/2} \right)^{1/2} = \\ &= [\mu(gx - gu_n, t)]^{1/2} * [\mu(gy - gv_n, t)]^{1/2} = \beta_n(t), \end{aligned}$$

para cada $n \geq 0$ y cada $t > 0$. Repitiendo este proceso:

$$\beta_{n+1}(kt) \geq \beta_n(t) \geq \beta_{n-1} \left(\frac{t}{k} \right) \geq \dots \geq \beta_0 \left(\frac{t}{k^n} \right),$$

para cada $n \geq 0$ y cada $t > 0$. En particular, según (3.19) y (3.20),

$$\begin{aligned} \mu(gx - gu_{n+1}, kt) &\geq [\mu(gx - gu_n, t)]^{1/2} * [\mu(gy - gv_n, t)]^{1/2} = \beta_n(t) \geq \beta_0 \left(\frac{t}{k^n} \right), \\ \mu(gy - gv_{n+1}, kt) &\geq [\mu(gv_n - gy, t)]^{1/2} * [\mu(gu_n - gx, t)]^{1/2} = \beta_n(t) \geq \beta_0 \left(\frac{t}{k^n} \right). \end{aligned}$$

Dado que $\lim_{n \rightarrow \infty} \beta_0 \left(\frac{t}{k^n} \right) = 1$, podemos concluir que

$$\lim_{n \rightarrow \infty} gu_n = gx \quad \text{y} \quad \lim_{n \rightarrow \infty} gv_n = gy.$$

Si hubiésemos supuesto que $gx \sqsupseteq gu_1$ y que $gy \sqsubseteq gv_1$, hubiésemos llegado a la misma propiedad. Y como (x^*, y^*) es otro punto de coincidencia doble de F y g de manera que $(F(u, v), F(v, u))$ también es comparable a $(F(x^*, y^*), F(y^*, x^*))$, concluimos que (3.18) es cierto. De paso, hemos demostrado que $gx = gx^*$ y $gy = gy^*$, sean cuales sean los puntos de coincidencia doble (x, y) y (x^*, y^*) de F y g .

Sea ahora (z, w) cualquier punto de coincidencia doble de F y g y llamemos $x = gz$ e $y = gw$. Vamos a demostrar que (x, y) es el único punto de coincidencia doble de F y g que se anuncia en el Teorema. En efecto, como F y g conmutan,

$$\begin{aligned} gx &= ggz = gF(z, w) = F(gz, gw) = F(x, y), \\ gy &= ggw = gF(w, z) = F(gw, gz) = F(y, x). \end{aligned}$$

Por consiguiente, (x, y) es otro punto de coincidencia doble de F y g . Según lo que hemos demostrado en (3.16), tenemos que $gz = gx$ y $gw = gy$. Esto implica que $x = gz = gx = F(x, y)$ y que $y = gw = gy = F(y, x)$. Por tanto, (x, y) es un punto de coincidencia doble de F y g que cumple lo anunciado en el Teorema.

Finalmente, (x, y) es el único que cumple esta propiedad ya que si existiese otro punto (x', y') que cumpliera esta propiedad, entonces (3.16) nos aseguraría que $x' = gx' = gx = x$ y que $y' = gy' = gy = y$. ■

En el caso particular de que g sea la aplicación identidad en X , tenemos el siguiente enunciado, donde se asegura que el punto de coincidencia doble es del tipo (x, x) .

Corolario 3.2.3 Sea $(X, \mu, *)$ un END completo de manera que $*$ es una t-norma de tipo Hadžić tal que $a * b \geq ab$ para cada $a, b \in \mathbb{I}$. Sea $F : X \times X \rightarrow X$ una aplicación que verifica la propiedad monótona mezclada y sea \sqsubseteq un orden parcial sobre X de manera que existe un número real positivo $k \in]0, 1[$ con la propiedad de que:

$$\mu(F(x, y) - F(u, v), kt) \geq [\mu(x - u, t)]^{1/2} * [\mu(y - v, t)]^{1/2}, \quad (3.21)$$

siempre que $x, y, u, v \in X$ verifiquen que $x \sqsubseteq u$ e $y \sqsupseteq v$. Supongamos que:

- (a) F es continua, o bien
- (b) (X, \sqsubseteq) posee la propiedad monótona secuencial.

Si existen dos puntos $x_0, y_0 \in X$ tales que $x_0 \sqsubseteq F(x_0, y_0)$ e $y_0 \sqsupseteq F(y_0, x_0)$, entonces existen dos puntos $x, y \in X$ de manera que $F(x, y) = x$ y $F(y, x) = y$.

Además, si los puntos x_0 e y_0 de partida son comparables, entonces $x = y$ (de forma que $F(x, x) = x$).

DEMOSTRACIÓN: Si tenemos presente toda la demostración del Teorema 3.2.1 suponiendo que g es la aplicación identidad en X , la primera parte del enunciado está ya demostrada. Sólo queda probar que si x_0 e y_0 son comparables, entonces $x = y$. En efecto, supongamos, por ejemplo, que $x_0 \sqsubseteq y_0$. Utilizando que F cumple la propiedad de monotonía mezclada, es posible demostrar, por inducción, que las sucesiones $\{x_n = F(x_{n-1}, y_{n-1})\}$ e $\{y_n = F(y_{n-1}, x_{n-1})\}$ verifican que $x_n \sqsubseteq y_n$ para cada $n \geq 0$. Entonces, la condición de contractividad (3.21) nos asegura que, para cada $n \geq 0$ y cada $t > 0$:

$$\begin{aligned} \mu(x_{n+1} - y_{n+1}, kt) &= \mu(F(x_n, y_n) - F(y_n, x_n), kt) \geq \\ &\geq [\mu(x_n - y_n, t)]^{1/2} * [\mu(y_n - x_n, t)]^{1/2} = \mu(x_n - y_n, t). \end{aligned}$$

Repitiendo este proceso, deducimos que:

$$\mu(x_n - y_n, kt) \geq \mu(x_{n-1} - y_{n-1}, t) \geq \mu\left(x_{n-2} - y_{n-2}, \frac{t}{k}\right) \geq \dots \geq \mu\left(x_0 - y_0, \frac{t}{k^{n-1}}\right).$$

Tomando límite cuando $n \rightarrow \infty$ y teniendo en cuenta que

$$\lim_{n \rightarrow \infty} \mu\left(x_0 - y_0, \frac{t}{k^{n-1}}\right) = 1 \quad \text{para cada } t > 0,$$

concluimos que $\mu(x - y, t) \geq 1$ para cada $t > 0$, es decir, $x = y$ en virtud del axioma (1.54). ■

Ejemplo 3.2.4 Consideremos el espacio normado $X = (\mathbb{R}, |\cdot|)$ con su orden usual \leq . Definamos $\mu : \mathbb{R} \times]0, \infty[\rightarrow \mathbb{I}$ como

$$\mu(x, t) = e^{-\frac{|x|}{t}} \quad \text{para cada } x \in \mathbb{R} \text{ y cada } t > 0.$$

Es sencillo demostrar que (\mathbb{R}, μ) es un END bajo la t -norma $*$ = mín, que es continua, de tipo Hadžić y verifica $\text{mín}(a, b) \geq ab$ para cada $a, b \in \mathbb{I}$ (de hecho, (\mathbb{R}, μ) también es un END bajo la t -norma producto, pero ésta no es de tipo Hadžić). Sean $\alpha, \beta > 0$ y $k \in]0, 1[$ números positivos tales que $4\alpha \leq k\beta$. Definamos $F : \mathbb{R} \times \mathbb{R} \rightarrow \mathbb{R}$ y $g : \mathbb{R} \rightarrow \mathbb{R}$ como $F(x, y) = \alpha(x - y)$ y $gx = \beta x$ para cada $x, y \in \mathbb{R}$. Claramente, F verifica la propiedad g -monótona mezclada, g es continua, F y g conmutan y $F(\mathbb{R} \times \mathbb{R}) = \mathbb{R} = g(\mathbb{R})$. Es posible demostrar que

$$\mu(F(x, y) - F(u, v), kt) \geq \text{mín}\left([\mu(gx - gu, t)]^{1/2}, [\mu(gy - gv, t)]^{1/2}\right)$$

siempre que $x \geq u$ e $y \leq v$. Por consiguiente, podemos aplicar el Teorema 3.2.1 para deducir que F y g poseen algún punto de coincidencia doble.

Para concluir este Capítulo presentamos una versión intuicionista del Teorema 3.2.1 imponiendo condiciones duales sobre la t -conorma. Si demostración se reduce a aplicar el Teorema 3.2.1 al soporte del espacio normado difuso intuicionista.

Corolario 3.2.5 Sea $(X, \mu, \nu, *, \diamond)$ un ENDI completo de manera que $*$ es una t -norma de tipo Hadžić tal que $a * b \geq ab$ para cada $a, b \in \mathbb{I}$. Sean $F : X \times X \rightarrow X$ y $g : X \rightarrow X$ dos aplicaciones tales que $F(X \times X) \subseteq g(X)$, g es continua y F conmuta con g . Sea \sqsubseteq un orden parcial sobre X tal que F verifica la propiedad g -monótona mezclada y de manera que existe un número real positivo $k \in]0, 1[$ con la propiedad de que:

$$\mu(F(x, y) - F(u, v), kt) \geq [\mu(gx - gu, t)]^{1/2} * [\mu(gy - gv, t)]^{1/2},$$

y

$$\nu(F(x, y) - F(u, v), kt) \leq \left(1 - [1 - \nu(gx - gu, t)]^{1/2}\right) \diamond \left(1 - [1 - \nu(gy - gv, t)]^{1/2}\right),$$

siempre que $x, y, u, v \in X$ verifiquen que $gx \sqsubseteq gu$ y $gy \sqsupseteq gv$. Supongamos que:

-
- (a) F es continua, o bien
- (b) (X, \sqsubseteq) posee la propiedad g -monótona secuencial.

Si existen dos puntos $x_0, y_0 \in X$ tales que $gx_0 \sqsubseteq F(x_0, y_0)$ y $gy_0 \sqsupseteq F(y_0, x_0)$, entonces F y g poseen algún punto de coincidencia doble (es decir, existen $x, y \in X$ tales que $gx = F(x, y)$ y $gy = F(y, x)$).

Capítulo 4

Una técnica de regresión difusa con números triangulares

El objetivo principal de este Capítulo es proporcionar a los investigadores un primer enfoque fácil de manejar para el problema de la regresión difusa cuando los datos observados presentan vaguedad/imprecisión y aleatoriedad de manera conjunta. La técnica nos permite analizar modelos lineales y no lineales y el experimento de simulación que analizaremos en la sección 4.2.2 nos muestra que la aproximación al problema de regresión que planteamos es razonablemente buena. El problema de encontrar trabajos que sean fáciles de implementar y que consideren el caso no lineal (véase el problema de estudio) nos llevó a desarrollar los nuestros propios resultados (véase [103]).

En este Capítulo presentamos nuevas formas de medir las distancias entre los números borrosos que nos permite dar una formulación simple de los problemas estadísticos difusos (como el de la regresión), con resultados comparables a los que se obtienen con técnicas más complejas. Las diferentes distancias que se utilizan para los conjuntos difusos no son las habituales, pero son muy importantes en Matemáticas y esto es particularmente interesante ya que ofrecen modelos alternativos de entre los cuales el experto puede elegir. También comentar que la regresión que se propone en este Capítulo puede ser generalizada y usarse con otro tipo de conjuntos difusos, por ejemplo, números difusos LR o trapezoidales.

El Capítulo se organiza de la siguiente manera. En la Sección 4.1 se presentan algunas anotaciones, el modelo de regresión difusa que consideraremos y resultados preliminares. El procedimiento de estimación se explica en la Sección 4.2 para datos de entrada precisos y datos de salida difusos triangulares. En la sección 4.2.2, se realiza un experimento de simulación y un estudio comparativo con los resultados que obtendríamos por otras técnicas propuestas anteriormente. Finalmente la sección 4.3 ilustra la aplicación de la metodología propuesta

utilizando datos recogidos de la vida real para dar solución a un problema no lineal.

4.1. El modelo de regresión difusa

Los *números difusos triangulares* $\mathcal{A} = (A^I, A^c, A^S)$ los hemos definido por tres números reales $A^I < A^c < A^S$ tales que la gráfica de \mathcal{A} es un triángulo de base el intervalo $[A^I, A^S]$ y vértice en $x = A^c$, esto es,

$$\mathcal{A}(x) = \begin{cases} \frac{x - A^I}{A^c - A^I}, & \text{si } A^I \leq x \leq A^c, \\ \frac{A^S - x}{A^S - A^c}, & \text{si } A^c \leq x \leq A^S, \\ 0, & \text{en otro caso.} \end{cases}$$

El número A^c denota el *centro* y $A^\ell = A^c - A^I$ y $A^r = A^S - A^c$ denotan las amplitudes (por lo tanto, \mathcal{A} puede también escribirse como $(A^c - A^\ell, A^c, A^c + A^r)$). Denotamos por \mathcal{T} la familia de los números triangulares sobre \mathbb{R} . En este Capítulo, la función asterisco se define como $\mathcal{A}^*(x) = \mathcal{A}(x + A^c)$ para todo $x \in \mathbb{R}$, esto es, $\mathcal{A}^* = (-A^\ell, 0, A^r)$ y podemos afirmar que $\mathcal{A} = A^c + (-A^\ell, 0, A^r) = A^c + \mathcal{A}^*$. Denotaremos por \mathcal{T}_0 el conjunto de todos los números triangulares tales que $A^c = 0$ (véase la Figura 4.1).

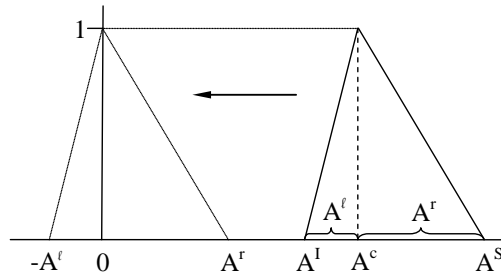


Figura 4.1: Transformación de un número difuso triangular a través de la función asterisco

Las variables aleatorias difusas están cobrando un gran interés en la últimos años, por su interés en estudios de tipo probabilístico. En primer lugar recordemos el concepto de vector aleatorio.

Definición 4.1.1 (Vector aleatorio) *Sea $(\Omega, \mathfrak{A}, P)$ un espacio de probabilidad y sea $(\mathbb{R}^N, \mathfrak{B})$ espacio medible y su σ -álgebra de Borel. Una función medible $X : \Omega \rightarrow \mathbb{R}^N$ es un vector aleatorio. Cuando $N = 1$ diremos que X es una variable aleatoria.*

Puri y Ralescu (1986) introdujeron el concepto de *variable aleatoria difusa* como una extensión de los conceptos de variable aleatoria y conjunto difuso.

Definición 4.1.2 (Variable aleatoria difusa) Sea $(\Omega, \mathfrak{A}, P)$ un espacio de probabilidad. se dice que la función $\mathcal{X} : \Omega \rightarrow \mathcal{F}$ es variable aleatoria difusa con valores en el espacio medible (Ω, \mathfrak{A}) si el α -corte, $\mathcal{X}_\alpha : \Omega \rightarrow \mathbb{R}$ tal que $X_\alpha(\omega) = X(\omega)_\alpha$ para todo $\omega \in \Omega$ es una variable aleatoria que sea $\alpha \in [0, 1]$

Definición 4.1.3 (Variable aleatoria difusa triangular) Una función $\mathcal{X} : \Omega \rightarrow \mathcal{T}$ es variable aleatoria difusa trapezoidal si la representación de \mathcal{X} , $(X^I, X^c, X^S) : \Omega \rightarrow \mathbb{R}^3$ es un vector aleatorio.

Sean \underline{X} e \mathcal{Y} dos variables de manera que $\underline{X} = (X_1, \dots, X_p)'$ es una variable aleatoria e \mathcal{Y} es una variable aleatoria difusa triangular. La expresión genérica del modelo de regresión que consideramos puede ser formalizada de la siguiente forma:

$$\mathcal{Y} = f(\underline{X}, \underline{a}) + (\varepsilon^I / \varepsilon^c / \varepsilon^S), \quad (4.1)$$

donde \underline{a} es el vector de parámetros y ε^I , ε^c y ε^S son los residuos, es decir, variables aleatorias (evaluadas en los números reales) cuyas varianzas son finitas.

Consideremos un experimento aleatorio en el que una variable respuesta difusa triangular \mathcal{Y} y p variables explicativas reales X_1, X_2, \dots, X_p son observadas en n unidades independientes, es decir, $\{\underline{X}_i, \mathcal{Y}_i\}_{i=1}^n$ es una muestra aleatoria obtenida de la variable $(\underline{X}, \mathcal{Y})$, siendo $\underline{X}_i = (X_{1i}, X_{2i}, \dots, X_{pi})'$. Dado que \mathcal{Y} puede ser expresada de la forma $(Y^I/Y^c/Y^S)$, el centro Y^c puede ser relacionado con las variables explicativas X_1, X_2, \dots, X_p a través de un modelo de regresión clásica mediante una función lineal, parabólica, hiperbólica, logarítmica, exponencial, etc. En lo sucesivo, supondremos que $\hat{Y}_x^c = f(\underline{x}, \hat{\underline{a}}) + \hat{b}$ denota el modelo ya ajustado para el centro.

4.1.1. Trabajos relacionados

El modelo de regresión difusa propuesto por Tanaka y otros (1982), que involucra observaciones difusas, considera que los parámetros de la regresión son números difusos y determina sus valores aplicando técnicas de programación lineal. Desde entonces, los modelos de regresión difusa han sido ampliamente estudiados en la literatura especializada.

En este Capítulo introducimos una nueva forma de medir distancias entre números difusos que nos conducirá a una formulación simple del problema de regresión difusa. El propósito principal es minimizar una función objetivo en lugar de minimizar una cota superior sobre el error generalizado de un modelo (como ocurre en el caso de las máquinas de soporte vectorial). Por tanto, el estudio comparativo que explicaremos en la sección 4.2.2 se ha realizado comparando con técnicas que tienen el mismo propósito, es decir, con técnicas paramétrica. A continuación, describimos brevemente algunos de estos métodos:

- Tanaka y sus colaboradores (véase [119, 120]) introdujeron la llamada regresión difusa o possibilística. En su trabajo, los autores diseñaron un ejemplo ilustrativo de su modelo de regresión difusa (al que nos referiremos como THW) para abordar el problema en el que se considera una variable numérica explicativa y una variable difusa dependiente que sólo toma números difusos triangulares simétricos.
- Diamond en 1988 (véase [28]) propuso un modelo (al que nos referiremos como DM) utilizando una aproximación por mínimos cuadrados para las observaciones difusas.
- Kim y Bishu en 1998 (véase [71]) estudiaron una modificación del análisis de regresión lineal difusa propuesto por Tanaka y otros, basándose en un criterio de minimización de la diferencia entre las funciones de pertenencia difusas de los números difusos observados y estimados. Para demostrar que su método (al que nos referiremos con KB) mejora al de Tanaka, utilizaron los mismos datos.
- Wu y Tseng en 2002 (véase [124]) aplicaron el método de mínimos cuadrados para construir, a la vez, las rectas de regresión superior e inferior. Este modelo (al que nos referiremos como WT) es la media entre ambas rectas de regresión. En la aproximación de Wu y Tseng, si las variables explicativas son numéricas y las variables dependientes son números difusos simétricos, el modelo de regresión difusa que se obtiene es el mismo que el de DM.
- El modelo de regresión difusa propuesto por Kao y Chyu en 2003 [68] (al que nos referiremos como KC) minimiza la suma de los errores cuadráticos utilizando un programa no lineal formulado bajo la estructura del método de Chen y Klein para la ordenación de números difusos.
- Nasrabadi y Nasrabadi en 2004 (véase [91]) propusieron un algoritmo basado en la aproximación de THW. Su modelo de regresión lineal difusa (al que nos referiremos como NN) considera entradas y salidas reales o difusas, y desarrolla una aproximación tanto matemática como informática.
- Hojati y sus colaboradores en 2005 (véase [59]) aplicaron métodos de programación lineal para construir su modelo de regresión difusa (al que nos referiremos como HBS).
- Basándose en el concepto de distancia, Chen y Hsueh en 2009 (véase [18]) propusieron un modelo de regresión difusa (al que nos referiremos como CH) que minimiza la suma de las distancias cuadráticas medias entre los valores observados y las respuestas estimadas utilizando algunos α -cortes.
- Cuando las variables de entrada y salida son ambas números difusos trapezoidales, González-Rodríguez y sus colaboradores en 2009 (véase [46]) resolvieron modelos simples de regresión lineal en términos de los momentos de los elementos aleatorios involucrados (nos referiremos a este método como GBCL).

4.2. Un modelo de regresión para variables de entrada reales y variables de salida difusas triangulares

En esta sección consideramos en primer lugar el caso de valores esperados de las amplitudes de los errores son constantes e ilustramos empíricamente la aplicación de la metodología propuesta a modelos tanto lineales como no lineales haciendo una comparación entre las soluciones obtenidas por las diferentes distancias consideradas. A continuación, presentamos una generalización de la metodología propuesta en la sección anterior y evaluamos nuestro método realizando un estudio comparativo con el resultados obtenidos a través de técnicas desarrolladas por otros/as autores/as.

4.2.1. Un modelo de regresión difusa cuando los valores esperados de las amplitudes no dependen de las variables explicativas

Regresión difusa basada en diferentes distancias

Supongamos que las amplitudes verifican la siguiente condición:

$$E(\varepsilon^c - \varepsilon^I | \underline{X}) = I, \quad E(\varepsilon^S - \varepsilon^c | \underline{X}) = S \quad \text{con } I, S > 0, \quad \text{siendo constantes.}$$

En este caso, proponemos considerar una función $\Phi : M \rightarrow \mathcal{T}$ definida mediante la siguiente expresión

$$\Phi_{\underline{x}} = \hat{Y}_{\underline{x}}^c + (-I/0/S), \quad \underline{x} \in M = \text{dom } \underline{X}.$$

De esta forma, el error total entre la función y los valores reales puede ser descrito de la siguiente forma:

$$E = E(I, S) = \sum_{i=1}^n d(\mathcal{Y}_i^*, \Phi_{\underline{x}_i}^*),$$

donde $d : \mathcal{T}_0 \times \mathcal{T}_0 \rightarrow [0, +\infty]$ es una distancia real (en la práctica, d suele ser una *quasidistancia*; véase Schweizer y Sklar 2005). El problema consiste en encontrar los valores de $I, S > 0$ que son solución del siguiente problema de minimización:

$$\text{mín}(E) = \text{mín} \left\{ \sum_{i=1}^n d(\mathcal{Y}_i^*, \Phi_{\underline{x}_i}^*) \right\}.$$

Nos damos cuenta de que existe un conflicto entre la búsqueda de solución de un problema de regresión difusa utilizando distancias reales y la interpretación de los datos difusos como distribuciones de posibilidad (véase Dubois y otros 2008; Couso y Sánchez 2006). En este sentido, la solución que proponemos en este Capítulo podría no ser consistente con esta interpretación. No obstante, como discutiremos en los siguientes apartados, existen diferentes situaciones prácticas en las que la solución que obtenemos minimizando distancias reales es apropiada y posee la

ventaja añadida de que el algoritmo numérico asociado es considerablemente más sencillo (véase Colubi 2009; González-Rodríguez y otros 2009). Volveremos sobre este tema en la sección 4.2.3.

En nuestra aproximación al problema, las funciones de error podrían ser descompuestas como $E(I, S) = E_1(I) + E_2(S)$ para cada $I, S > 0$. En tal caso, si E_1 posee un mínimo absoluto en $I_0 > 0$ y E_2 posee un mínimo absoluto en $S_0 > 0$, entonces E posee un mínimo absoluto en (I_0, S_0) .

En esta sección, los parámetros I y S del modelo (4.1) serán estimados minimizando algunas distancias reales entre números difusos. En concreto, utilizaremos las siguientes distancias entre números difusos $A, B \in \mathcal{T}_0$ (centrados en cero).

$$\begin{aligned} d_1, d_1^2, d_2, d_\infty &: \mathcal{T}_0 \times \mathcal{T}_0 \rightarrow [0, +\infty], \\ d_1(A, B) &= \int_{\mathbb{R}} |A(t) - B(t)| dt, \\ d_1^2(A, B) &= \left(\int_{-\infty}^0 |A(t) - B(t)| dt \right)^2 + \left(\int_0^{+\infty} |A(t) - B(t)| dt \right)^2, \\ d_2(A, B) &= \int_{\mathbb{R}} (A(t) - B(t))^2 dt, \\ d_\infty(A, B) &= \sup(\{|A(t) - B(t)| : t \leq 0\}) + \sup(\{|A(t) - B(t)| : t \geq 0\}) \end{aligned}$$

Por ejemplo, la distancia d_1 puede ser interpretada como el área comprendida entre los dos números difusos (representada en color azul en la Figura 4.2).

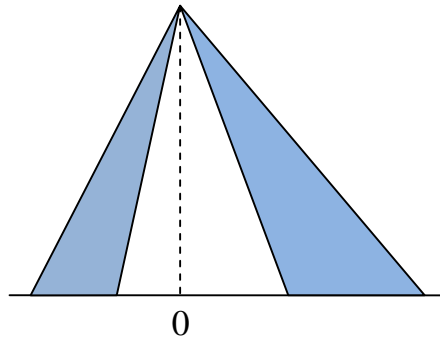


Figura 4.2: El área azul corresponde a la distancia entre $A, B \in \mathcal{T}_0$ utilizando d_1 .

Vamos a estudiar este problema con cada una de las distancias anteriores y resumiremos los resultados obtenidos en el teorema 4.2.11.

Resolución del problema de regresión difusa considerando la distancia d_1

En primer lugar, consideremos la distancia difusa $d_1 : \mathcal{T}_0 \times \mathcal{T}_0 \rightarrow \mathbb{R}_0^+$ dada por:

$$d_1(A, B) = \int_{\mathbb{R}} |A(t) - B(t)| dt,$$

Esta distancia representa el área que existe entre los números difusos A y B .

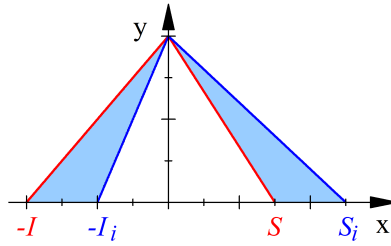


Figura 4.3: El área entre los números difusos depende de sus bases.

Dado que nos restringimos a los números triangulares (que poseen soporte compacto), esta distancia siempre es finita. Teniendo en cuenta que $\mathcal{Y}_i = (Y_i - I_i/Y_i/Y_i + S_i)$ y que $\Phi_{\underline{x}} = (\hat{Y}_{\underline{x}} - I/\hat{Y}_{\underline{x}}/\hat{Y}_{\underline{x}} + S)$, observamos que $\mathcal{Y}_i^* = (-I_i/0/S_i) \in \mathcal{T}_0$ y que $\Phi_{\underline{x}}^* = (-I/0/S) \in \mathcal{T}_0$, siendo $I_i, S_i, I, S > 0$ números reales positivos. Entonces, dado que se trata de áreas de triángulos:

$$d_1(\mathcal{Y}_i^*, \Phi_{\underline{x}_i}^*) = \int_{-\infty}^0 |\mathcal{Y}_i^*(t) - \Phi_{\underline{x}_i}^*(t)| dt + \int_0^{+\infty} |\mathcal{Y}_i^*(t) - \Phi_{\underline{x}_i}^*(t)| dt = \frac{|I - I_i| + |S - S_i|}{2}.$$

Así, el error de estimación es:

$$\begin{aligned} E(I, S) &= \sum_{i=1}^n d_1(\mathcal{Y}_i^*, \Phi_{\underline{x}_i}^*) = \frac{1}{2} \left[\sum_{i=1}^n |I - I_i| + \sum_{i=1}^n |S - S_i| \right] = \frac{E(I)}{2} + \frac{E'(S)}{2}, \\ \text{donde } E(I) &= \sum_{i=1}^n |I - I_i| \quad \text{y} \quad E'(S) = \sum_{i=1}^n |S - S_i|. \end{aligned} \quad (4.2)$$

Dado que las dos funciones E y E' poseen expresiones similares, vamos a razonar con la primera de ellas y, el mismo resultado será aplicable a la segunda. Los valores I_1, I_2, \dots, I_n juegan un papel intercambiable en esta función, por lo que podemos suponer que:

$$0 < I_1 \leq I_2 \leq \dots \leq I_n.$$

Lema 4.2.1 El mínimo absoluto de la función E definida en (4.2) se alcanza en la mediana de los datos $\{I_1, I_2, \dots, I_n\}$.

DEMOSTRACIÓN: La función E es claramente continua en $]0, +\infty[$ y C^∞ en $]0, +\infty[\setminus \{I_1, I_2, \dots, I_n\}$. Llamemos $\{I'_1, I'_2, \dots, I'_k\}$ a los diferentes valores que hay entre las amplitudes $\{I_1, I_2, \dots, I_n\}$ de

manera que cada valor I'_j aparece en $\{I_1, I_2, \dots, I_n\}$ m_j veces y además $0 < I'_1 < I'_2 < \dots < I'_k$. Así, los datos están ordenados de la siguiente forma:

$$\underbrace{I'_1, I'_1, \dots, I'_1}_{m_1}, \dots, \underbrace{I'_r, I'_r, \dots, I'_r}_{m_r}, \underbrace{I'_{r+1}, I'_{r+1}, \dots, I'_{r+1}}_{m_{r+1}}, \dots, \underbrace{I'_k, I'_k, \dots, I'_k}_{m_k}. \quad (4.3)$$

Si denotamos por \bar{I} a la media de los datos I_1, I_2, \dots, I_n ,

$$\bar{I} = \frac{I_1 + I_2 + \dots + I_n}{n} = \frac{m_1 I'_1 + m_2 I'_2 + \dots + m_k I'_k}{n},$$

es sencillo demostrar que:

$$E(x) = \begin{cases} n(\bar{I} - x), & \text{si } 0 < x \leq I'_1, \\ x \left(\sum_{i=1}^r m_i - \sum_{i=r+1}^k m_i \right) + \sum_{i=r+1}^k m_i I'_i - \sum_{i=1}^r m_i I'_i, & \text{si } I'_r < x \leq I'_{r+1}, \\ n(x - \bar{I}), & \text{si } x > I'_k. \end{cases}$$

En particular, la gráfica de E es una línea quebrada, es decir, una poligonal (en cada intervalo $]I'_r, I'_{r+1}[$ coincide con un segmento de línea recta pues es un polinomio de grado menor o igual que 1 en la variable x). Así, $\lim_{x \rightarrow 0^+} E(x) = N \cdot \bar{I} > 0$ y $\lim_{x \rightarrow +\infty} E(x) = +\infty$. Además, E es derivable en $]0, +\infty[\setminus \{I'_1, I'_2, \dots, I'_k\}$ y su primera derivada es, al menos en $]0, +\infty[\setminus \{I'_1, I'_2, \dots, I'_k\}$:

$$E'(x) = \begin{cases} -N, & \text{si } 0 < x < I'_1, \\ \sum_{i=1}^r m_i - \sum_{i=r+1}^k m_i, & \text{si } I'_r < x < I'_{r+1}, \\ N, & \text{si } x > I'_k. \end{cases}$$

Esto demuestra que E es estrictamente decreciente en $]0, I'_1[$ y es estrictamente creciente en $]I'_k, +\infty[$. Teniendo en cuenta que E es continua en $]0, +\infty[$ y estrictamente decreciente en $]0, I'_1[$, deducimos que el valor mínimo de E en $]0, I'_1[$ se alcanza en I'_1 . Igualmente se deduce que el valor mínimo de E en $]I'_k, +\infty[$ se alcanza en I'_k . Como E es continua en $[I'_1, I'_k]$, el *teorema de Bolzano-Weierstrass* nos asegura que E posee mínimo absoluto en $[I'_1, I'_k]$. Por consiguiente, ese punto es mínimo absoluto de E en $]0, +\infty[$, y esto demuestra que E posee mínimo absoluto en $]0, +\infty[$, y dicho mínimo está en el intervalo $[I'_1, I'_k] = [\min(I_1, I_2, \dots, I_n), \max(I_1, I_2, \dots, I_n)]$.

Consideremos las pendientes de las poligonales que definen la gráfica de la función E en los $k - 1$ intervalos $]I'_1, I'_2[, \dots,]I'_{k-1}, I'_k[$:

$$Z_r = \sum_{i=1}^r m_i - \sum_{i=r+1}^k m_i, \quad \text{para cada } r \in \{1, 2, \dots, k-1\}.$$

Entonces, para cada $r \in \{1, 2, \dots, k-2\}$:

$$Z_{r+1} - Z_r = \left(\sum_{i=1}^{r+1} m_i - \sum_{i=r+2}^k m_i \right) - \left(\sum_{i=1}^r m_i - \sum_{i=r+1}^k m_i \right) = 2m_{r+1} > 0.$$

Esto demuestra que $-N < Z_1 < Z_2 < \dots < Z_{k-1} < N$. De esta forma, las primeras pendientes serán negativas (y E será decreciente en los correspondientes intervalos) mientras que las últimas serán positivas (y E será creciente). Si ninguna pendiente es cero, la poligonal pasa de estrictamente decreciente a estrictamente creciente y, por ello, posee un único mínimo absoluto.

Si existiese un valor $r \in \{1, 2, \dots, k-1\}$ de manera que $Z_r < 0 < Z_{r+1}$, el único mínimo absoluto E en $]0, +\infty[$ se alcanzaría en el punto I'_{r+1} , ya que es el extremo superior del intervalo $]I_r, I'_{r+1}[$. En tal caso, se tendría que:

$$\sum_{i=1}^r m_i - \sum_{i=r+1}^k m_i < 0 < \sum_{i=1}^{r+1} m_i - \sum_{i=r+2}^k m_i,$$

de donde se deduce que:

$$\begin{aligned} m_1 + m_2 + \dots + m_r &< m_{r+1} + (m_{r+2} + \dots + m_k) \quad \text{y que} \\ m_{r+2} + \dots + m_k &< m_{r+1} + (m_1 + m_2 + \dots + m_r). \end{aligned}$$

Esto significa que la frecuencia m_{r+1} corresponde al punto en el que se supera, por primera vez, el valor $n/2$. De esta forma,

$$m_1 + m_2 + \dots + m_r < n/2 < m_{r+2} + \dots + m_k,$$

y el valor I'_{r+1} es exactamente la mediana de los datos $\{I'_1, I'_2, \dots, I'_k\}$ contados con sus multiplicidades, es decir, la mediana de los datos $\{I_1, I_2, \dots, I_n\}$.

Para que hubiese más de un mínimo absoluto, la primera derivada de E debe anularse en un intervalo $]I'_r, I'_{r+1}[$, en cuyo caso la gráfica de E en ese intervalo es un trozo de línea recta horizontal (se alcanza el mínimo absoluto en todos los puntos del intervalo $]I'_r, I'_{r+1}[$). En tal caso, como la suma de $\{m_i\}_{i=1}^k$ es n , tenemos que:

$$E'(x) = 0 \text{ si } x \in]I'_r, I'_{r+1}[\Leftrightarrow \sum_{i=1}^r m_i - \sum_{i=r+1}^k m_i = 0 \Leftrightarrow \sum_{i=1}^r m_i = \sum_{i=r+1}^k m_i = \frac{n}{2}.$$

Esto significa que los valores $\{I'_i\}_{i=1}^k$, junto con sus respectivas multiplicidades, están ordenados de menor a mayor así:

$$\underbrace{\underbrace{I'_1, I'_1, \dots, I'_1}_{m_1}, \dots, \underbrace{I'_r, I'_r, \dots, I'_r}_{m_r}}_{n/2}, \underbrace{\underbrace{I'_{r+1}, I'_{r+1}, \dots, I'_{r+1}}_{m_{r+1}}, \dots, \underbrace{I'_k, I'_k, \dots, I'_k}_{m_k}}_{n/2}$$

Efectivamente, n debe ser par y éste es el caso en el que la mediana se puede tomar como cualquier valor intermedio entre I'_r e I'_{r+1} . ■

Del lema anterior podemos deducir cómo resolver el primer caso de regresión difusa que nos planteamos.

Teorema 4.2.2 Sean $\underline{X}, \mathcal{Y}$ dos variables tales que \underline{X} es una variable aleatoria real e \mathcal{Y} es una variable aleatoria difusa que se relacionan a través de la ecuación (4.1). Sea $\{\underline{X}_i, \mathcal{Y}_i\}_{i=1}^n$ una muestra aleatoria simple obtenida de $(\underline{X}, \mathcal{Y})$, siendo $\mathcal{Y}_i = (Y_i - I_i/Y_i/Y_i + S_i)$. Sea \widehat{Y}_x^c un modelo ajustado para los centros a partir de $\{\underline{X}_i, Y_i\}_{i=1}^n$ y consideremos el problema de regresión (en $I, S > 0$) para la distancia d_1 :

$$(P_1) \begin{cases} \Phi_{I,S} : \text{dom } X \rightarrow \mathcal{T}, & \Phi_{I,S}(x) = (\widehat{Y}_x - I/\widehat{Y}_x/\widehat{Y}_x + S), \\ E(I, S) = \sum_{i=1}^n d_1(\mathcal{Y}_i^*, \Phi_{\underline{X}_i}^*) & \text{mínimo.} \end{cases}$$

Entonces el problema (P_1) posee una única solución y los valores estimados para I y S son las respectivas medianas de las amplitudes:

$$\widehat{I} = \text{Mediana}(I_1, \dots, I_n) \quad \text{y} \quad \widehat{S} = \text{Mediana}(S_1, \dots, S_n).$$

Corolario 4.2.3 Si todas las amplitudes I_i son iguales, la mejor predicción para I que podemos hacer es ese mismo valor.

DEMOSTRACIÓN: En este caso, el error sería cero, ya que si $I_1 = I_2 = \dots = I_n = \bar{I}$, entonces:

$$E(x) = \sum_{i=1}^k m_i |x - I_i| = n \cdot |x - \bar{I}|,$$

cuyo valor se anula para $x = \bar{I}$. ■

Resolución del problema de regresión difusa considerando la distancia d_1^2

En este caso, podemos expresar:

$$\begin{aligned} d_1^2(\mathcal{Y}_i^*, \Phi_{\underline{X}_i}^*) &= \left(\int_{-\infty}^0 |\mathcal{Y}_i^*(t) - \Phi_{\underline{X}_i}^*(t)| dt \right)^2 + \left(\int_0^{+\infty} |\mathcal{Y}_i^*(t) - \Phi_{\underline{X}_i}^*(t)| dt \right)^2 = \\ &= \left(\frac{I - I_i}{2} \right)^2 + \left(\frac{S - S_i}{2} \right)^2, \end{aligned}$$

por lo que podemos expresar el error total como:

$$\begin{aligned} E &= \sum_{i=1}^n d_1^2(\mathcal{Y}_i^*, \Phi_{\underline{X}_i}^*) = \sum_{i=1}^n \left[\frac{(I - I_i)^2}{4} + \frac{(S - S_i)^2}{4} \right] = \frac{E(I)}{4} + \frac{E'(S)}{4}, \\ \text{donde } E(I) &= \sum_{i=1}^n (I - I_i)^2 \quad \text{y} \quad E'(S) = \sum_{i=1}^n (S - S_i)^2. \end{aligned}$$

Obviamente, este caso conduce al ampliamente utilizado método de regresión por *mínimo cuadrados*. Es conocido que el valor óptimo para I es la media aritmética de los datos $\{I_i\}_{i=1}^n$, pues es el valor que minimiza los cuadrados de las desviaciones a los datos. Esto demuestra el siguiente enunciado.

Teorema 4.2.4 Sean $\underline{X}, \mathcal{Y}$ dos variables tales que \underline{X} es una variable aleatoria real e \mathcal{Y} es una variable aleatoria difusa que se relacionan a través de la ecuación (4.1). Sea $\{\underline{X}_i, \mathcal{Y}_i\}_{i=1}^n$ una muestra aleatoria simple obtenida de $(\underline{X}, \mathcal{Y})$, siendo $\mathcal{Y}_i = (Y_i - I_i/Y_i/Y_i + S_i)$. Sea $\widehat{Y}_{\underline{x}}^c$ un modelo ajustado para los centros a partir de $\{\underline{X}_i, Y_i\}_{i=1}^n$ y consideremos el problema de regresión (en $I, S > 0$) para la distancia d_1^2 :

$$(P_1^2) \begin{cases} \Phi_{I,S} : \text{dom } X \rightarrow \mathcal{T}, & \Phi_{I,S}(\underline{x}) = (\widehat{Y}_{\underline{x}} - I/\widehat{Y}_{\underline{x}}/\widehat{Y}_{\underline{x}} + S), \\ E(I, S) = \sum_{i=1}^n d_1^2(\mathcal{Y}_i^*, \Phi_{\underline{X}_i}^*) & \text{mínimo.} \end{cases}$$

Entonces el problema (P_1^2) posee una única solución y los valores estimados para I y S son las respectivas medias de las amplitudes:

$$\widehat{I} = \text{Media}(I_1, \dots, I_n) \quad \text{y} \quad \widehat{S} = \text{Media}(S_1, \dots, S_n).$$

Corolario 4.2.5 Si todas las amplitudes I_i son iguales, la mejor predicción para I que podemos hacer es ese mismo valor.

Resolución del problema de regresión difusa considerando la distancia d_2

Como tercer caso, consideremos la distancia más importante, $d_2 : \mathcal{T}_0 \times \mathcal{T}_0 \rightarrow \mathbb{R}_0^+$ dada por:

$$d_2(A, B) = \int_{\mathbb{R}} (A(t) - B(t))^2 dt$$

Calculamos esta distancia entre los números triangulares $\mathcal{Y}_i^* = (-I_i/0/S_i)$ y $\Phi_{\underline{x}}^* = (-I/0/S)$. Supongamos, por ejemplo, que $0 < I \leq I_i$. Entonces:

$$\begin{aligned} \int_0^{+\infty} (\mathcal{Y}_i^*(t) - \Phi_{\underline{x}}^*(t))^2 dt &= \int_0^I (\mathcal{Y}_i^*(t) - \Phi_{\underline{x}}^*(t))^2 dt + \int_I^{I_i} (\mathcal{Y}_i^*(t) - \Phi_{\underline{x}}^*(t))^2 dt = \\ &= \int_0^I \left(\left(1 - \frac{t}{I_i}\right) - \left(1 - \frac{t}{I}\right) \right)^2 dt + \int_I^{I_i} \left(1 - \frac{t}{I_i}\right)^2 dt = \\ &= \int_0^I \left(\frac{I - I_i}{I_i I} t \right)^2 dt + \int_I^{I_i} \left(\frac{I_i - t}{I_i} \right)^2 dt = \\ &= \frac{(I - I_i)^2}{I_i^2 I^2} \int_0^I t^2 dt + \frac{1}{I_i^2} \int_I^{I_i} (I_i - t)^2 dt = \frac{(I - I_i)^2}{I_i^2 I^2} \left[\frac{t^3}{3} \right]_{t=0}^{t=I} + \frac{1}{I_i^2} \left[\frac{(I_i - t)^3}{-3} \right]_{t=I}^{t=I_i} = \\ &= \frac{(I - I_i)^2}{I_i^2 I^2} \left(\frac{I^3}{3} - 0 \right) - \frac{1}{3I_i^2} \left(0 - (I_i - I)^3 \right) = \frac{(I - I_i)^2}{I_i^2 I^2} \cdot \frac{I^3}{3} + \frac{(I_i - I)^3}{3I_i^2} = \\ &= \frac{(I - I_i)^2}{3I_i^2} \left(\frac{I^3}{I^2} + I_i - I \right) = \frac{(I - I_i)^2 I_i}{3I_i^2} = \frac{(I - I_i)^2}{3I_i}. \end{aligned}$$

Si intercambiamos los papeles de I y de I_i , observamos que, en general:

$$\int_0^{+\infty} (\mathcal{Y}_i^*(t) - \Phi_{\underline{x}}^*(t))^2 dt = \frac{(I - I_i)^2}{3 \max(I, I_i)}.$$

Por consiguiente, podemos descomponer:

$$\begin{aligned} E &= \sum_{i=1}^n d_2(\mathcal{Y}_i^*, \Phi_x^*) = \sum_{i=1}^n \left[\int_{-\infty}^0 (\mathcal{Y}_i^*(t) - \Phi_x^*(t))^2 dt + \int_0^{+\infty} (\mathcal{Y}_i^*(t) - \Phi_x^*(t))^2 dt \right] = \\ &= \sum_{i=1}^n \left[\frac{(I - I_i)^2}{3 \max(I, I_i)} + \frac{(S - S_i)^2}{3 \max(S, S_i)} \right] = \frac{1}{3} \sum_{i=1}^n \left[\frac{(I - I_i)^2}{\max(I, I_i)} + \frac{(S - S_i)^2}{\max(S, S_i)} \right] = \frac{1}{3} (E(I) + E'(S)), \end{aligned}$$

donde $E, E' :]0, \infty[\rightarrow]0, \infty[$ son las funciones definidas como sigue:

$$E(I) = \sum_{i=1}^n \frac{(I - I_i)^2}{\max(I, I_i)} \quad \text{y} \quad E'(S) = \sum_{i=1}^n \frac{(S - S_i)^2}{\max(S, S_i)}.$$

Si cada amplitud I'_i posee multiplicidad m_i , podemos expresar, para cada $x > 0$:

$$E(x) = \sum_{i=1}^n \frac{(x - I_i)^2}{\max(x, I_i)} = \sum_{i=1}^k \frac{m_i (x - I'_i)^2}{\max(x, I'_i)}. \quad (4.4)$$

Lema 4.2.6 La función $E = E(I)$ definida en (4.4) posee un único mínimo absoluto en $]0, +\infty[$ que coincide con la única solución de las siguientes $k - 1$ ecuaciones de tercer grado:

$$\left\{ 2x^3 \sum_{i=r+1}^k \frac{m_i}{I'_i} + x^2 \left(\sum_{i=1}^r m_i - 2 \sum_{i=r+1}^k m_i \right) - \sum_{i=1}^r m_i (I'_i)^2 = 0 \quad \text{en el intervalo } [I'_r, I'_{r+1}] \right\}_{r=1}^{r=k-1}.$$

Una, y sólo una, de las ecuaciones anteriores posee solución en el intervalo indicado, salvo que se trate de una amplitud concreta, que pertenece a dos intervalos (pero la solución seguiría siendo única).

DEMOSTRACIÓN: A diferencia de los casos anteriores, veremos que la función E es derivable en $]0, +\infty[$ y que su derivada E' es continua en $]0, +\infty[$. En efecto, cada función:

$$\max(x, I'_i) = \begin{cases} I'_i, & \text{si } 0 < x < I'_i, \\ x, & \text{si } x > I'_i \end{cases}$$

es continua en $]0, +\infty[$ y derivable C^∞ en $]0, +\infty[\setminus \{I'_i\}$, por lo que E es una función continua en $]0, +\infty[$ y C^∞ en, al menos, $]0, +\infty[\setminus \{I'_1, I'_2, \dots, I'_k\}$. No es difícil demostrar que E se puede expresar como:

$$E(x) = \begin{cases} x^2 \sum_{i=1}^k \frac{m_i}{I'_i} - 2nx + n\bar{I}, & \text{si } 0 < x \leq I'_1, \\ x^2 \sum_{i=r+1}^k \frac{m_i}{I'_i} + x \left(\sum_{i=1}^r m_i - 2 \sum_{i=r+1}^k m_i \right) + \frac{1}{x} \sum_{i=1}^r m_i (I'_i)^2 + \sum_{i=r+1}^k m_i I'_i - 2 \sum_{i=1}^r m_i I'_i, & \text{si } I'_r < x \leq I'_{r+1}, \\ nx - 2n\bar{I} + \frac{1}{x} \sum_{i=1}^k m_i (I'_i)^2, & \text{si } x > I'_k. \end{cases}$$

De esta forma, $\lim_{x \rightarrow 0^+} E(x) = N\bar{A} > 0$ y $\lim_{x \rightarrow +\infty} E(x) = +\infty$. Así, su primera derivada es, al menos en $]0, +\infty[\setminus \{I'_1, I'_2, \dots, I'_k\}$:

$$E'(x) = \begin{cases} 2x \sum_{i=1}^n \frac{m_i}{I'_i} - 2n, & \text{si } 0 < x < I'_1, \\ 2x \sum_{i=r+1}^k \frac{m_i}{I'_i} + \left(\sum_{i=1}^r m_i - 2 \sum_{i=r+1}^k m_i \right) - \frac{1}{x^2} \sum_{i=1}^r m_i (I'_i)^2, & \text{si } I'_r < x < I'_{r+1}, \\ n - \frac{1}{x^2} \sum_{i=1}^k m_i (I'_i)^2, & \text{si } x > I'_k. \end{cases}$$

Si tomamos cualquier $x \in]0, I'_1[$, observamos que:

$$E'(x) = 2x \sum_{i=1}^k \frac{m_i}{I'_i} - 2n = 2 \left(\sum_{i=1}^k \frac{m_i x}{I'_i} - n \right) = 2 \left(\sum_{i=1}^k \frac{m_i x}{I'_i} - \sum_{i=1}^k \frac{m_i I'_i}{I'_i} \right) = 2 \sum_{i=1}^k \frac{m_i (x - I'_i)}{I'_i} < 0,$$

por lo que E es estrictamente decreciente en $]0, I'_1[$. Igualmente, si tomamos $x \in]I'_k, +\infty[$, resulta que:

$$E'(x) = n - \frac{1}{x^2} \sum_{i=1}^k m_i (I'_i)^2 = n - \sum_{i=1}^k \frac{m_i (I'_i)^2}{x^2} = \sum_{i=1}^k \frac{m_i x^2}{x^2} - \sum_{i=1}^k \frac{m_i (I'_i)^2}{x^2} = \sum_{i=1}^k m_i \frac{x^2 - (I'_i)^2}{x^2} > 0,$$

por lo que E es estrictamente creciente en $]I'_k, +\infty[$. Razonando como en los caso anteriores, deducimos que E posee mínimo absoluto en $]0, +\infty[$ y que dicho mínimo está en el intervalo $[I'_1, I'_k]$. Veamos, a continuación, que los límites laterales de E en cada punto I'_i (que, evidentemente, son finitos) son iguales, por lo que E también es derivable en cada punto I'_i . En efecto, si tomamos cualquier $r \in \{2, 3, \dots, k-1\}$, observamos que:

$$\begin{aligned} & \lim_{x \rightarrow (I'_r)^+} E(x) - \lim_{x \rightarrow (I'_r)^-} E(x) = \\ &= \lim_{x \rightarrow (I'_r)^+} \left[2x \sum_{i=r+1}^k \frac{m_i}{I'_i} + \left(\sum_{i=1}^r m_i - 2 \sum_{i=r+1}^k m_i \right) - \frac{1}{x^2} \sum_{i=1}^r m_i (I'_i)^2 \right] - \\ & \quad - \lim_{x \rightarrow (I'_r)^-} \left[2x \sum_{i=r}^k \frac{m_i}{I'_i} + \left(\sum_{i=1}^{r-1} m_i - 2 \sum_{i=r}^k m_i \right) - \frac{1}{x^2} \sum_{i=1}^{r-1} m_i (I'_i)^2 \right] = \\ &= \left[2I'_r \sum_{i=r+1}^k \frac{m_i}{I'_i} + \left(\sum_{i=1}^r m_i - 2 \sum_{i=r+1}^k m_i \right) - \frac{1}{(I'_r)^2} \sum_{i=1}^r m_i (I'_i)^2 \right] - \\ & \quad - \left[2I'_r \sum_{i=r}^k \frac{m_i}{I'_i} + \left(\sum_{i=1}^{r-1} m_i - 2 \sum_{i=r}^k m_i \right) - \frac{1}{(I'_r)^2} \sum_{i=1}^{r-1} m_i (I'_i)^2 \right] = \\ &= 2I'_r \left(\sum_{i=r+1}^k \frac{m_i}{I'_i} - \sum_{i=r}^k \frac{m_i}{I'_i} \right) + \left(\sum_{i=1}^r m_i - \sum_{i=1}^{r-1} m_i \right) - \\ & \quad - 2 \left(\sum_{i=r+1}^k m_i - \sum_{i=r}^k m_i \right) - \frac{1}{(I'_r)^2} \left(\sum_{i=1}^r m_i (I'_i)^2 - \sum_{i=1}^{r-1} m_i (I'_i)^2 \right) = \\ &= 2I'_r \left(-\frac{m_r}{I'_r} \right) + m_r - 2(-m_r) - \frac{1}{(I'_r)^2} m_r (I'_r)^2 = \\ &= -2m_r + m_r + 2m_r - m_r = 0. \end{aligned}$$

De manera similar se trabaja en los nodos I'_1 y I'_k , por lo que finalmente se puede deducir que E es derivable en $]0, +\infty[$ y que su primera derivada E' es continua en $]0, +\infty[$. Utilizando la continuidad de E' en I'_1 :

$$\begin{aligned} E'(I'_1) &= \lim_{x \rightarrow (I'_1)^-} E(x) = \lim_{x \rightarrow (I'_1)^-} \left[2x \sum_{i=1}^k \frac{m_i}{I'_i} - 2n \right] = 2I'_1 \sum_{i=1}^k \frac{m_i}{I'_i} - 2n = \\ &= 2 \left(\sum_{i=1}^k \frac{m_i I'_1}{I'_i} - n \right) = 2 \left(\sum_{i=1}^k \frac{m_i I'_1}{I'_i} - \sum_{i=1}^k \frac{m_i I'_i}{I'_i} \right) = 2 \sum_{i=1}^k \frac{m_i (I'_1 - I'_i)}{I'_i} \leq 0. \end{aligned}$$

De la misma forma se demuestra que

$$E'(I'_k) = \lim_{x \rightarrow (I'_k)^+} E(x) = \lim_{x \rightarrow (I'_k)^+} \left[n - \frac{1}{x^2} \sum_{i=1}^k m_i (I'_i)^2 \right] = \sum_{i=1}^k \frac{m_i ((I'_k)^2 - (I'_i)^2)}{(I'_k)^2} \geq 0,$$

donde se utiliza que $(I'_k)^2 \geq (I'_i)^2$. Como E' es continua en $[I'_1, I'_k]$ y verifica $E'(I'_1) \leq 0$ y $E'(I'_k) \geq 0$, el *teorema de Bolzano* nos asegura que la función E' se anula, al menos, en un punto $x \in [I'_1, I'_k]$. Volviendo a derivar, la derivada segunda de E es, al menos en $]0, +\infty[\setminus \{I'_1, I'_2, \dots, I'_k\}$:

$$E''(x) = \begin{cases} 2 \sum_{i=1}^k \frac{m_i}{I'_i}, & \text{si } 0 < x < I'_1, \\ 2 \sum_{i=r+1}^k \frac{m_i}{I'_i} + \frac{2}{x^3} \sum_{i=1}^r m_i (I'_i)^2, & \text{si } I'_r < x < I'_{r+1}, \\ \frac{2}{x^3} \sum_{i=1}^k m_i (I'_i)^2, & \text{si } x > I'_k. \end{cases}$$

Esto demuestra que $E''(x) > 0$ para cada $x \in]0, +\infty[\setminus \{I'_1, I'_2, \dots, I'_k\}$. Dado que E' es continua en $]0, \infty[$ y su primera derivada es estrictamente positiva (salvo en un número finito de puntos), deducimos que la función E' es estrictamente creciente en $]0, \infty[$. Así, sólo puede poseer, a lo sumo, un único cero en el intervalo $]0, \infty[$. Como ya hemos demostrado que E' se anula en $[I'_1, I'_k]$, deducimos que la función E' posee un único cero en el intervalo $]0, \infty[$, que se sitúa en el intervalo $[I'_1, I'_k]$. No obstante, como E es continua y derivable en $]0, \infty[$ y ya sabemos que posee un mínimo absoluto (que debe estar situado en $[I'_1, I'_k]$), es obligatorio que en este extremo se anule la primera derivada de E . Como E' sólo se anula en un único punto, debe tratarse del mínimo absoluto de E en $]0, +\infty[$. Para localizarlo, hay que resolver la ecuación $E'(x) = 0$ para $x \in [I'_1, I'_k]$, ecuación que equivale a las $k - 1$ ecuaciones:

$$\left\{ \begin{array}{l} E'|_{[I'_r, I'_{r+1}]}(x) = 2x \sum_{i=r+1}^k \frac{m_i}{I'_i} + \left(\sum_{i=1}^r m_i - 2 \sum_{i=r+1}^k m_i \right) - \frac{1}{x^2} \sum_{i=1}^r m_i (I'_i)^2 = 0 \\ \text{en el intervalo } [I'_r, I'_{r+1}] \end{array} \right\}_{r=1}^{r=k-1},$$

es decir, a las $k - 1$ ecuaciones de tercer grado:

$$\left\{ 2x^3 \sum_{i=r+1}^k \frac{m_i}{I'_i} + x^2 \left(\sum_{i=1}^r m_i - 2 \sum_{i=r+1}^k m_i \right) - \sum_{i=1}^r m_i (I'_i)^2 = 0 \text{ en el intervalo } [I'_r, I'_{r+1}] \right\}_{r=1}^{r=k-1}.$$

■

En este contexto, resolvemos totalmente el problema de regresión asociado.

Teorema 4.2.7 Sean $\underline{X}, \mathcal{Y}$ dos variables tales que \underline{X} es una variable aleatoria real e \mathcal{Y} es una variable aleatoria difusa que se relacionan a través de la ecuación (4.1). Sea $\{\underline{X}_i, \mathcal{Y}_i\}_{i=1}^n$ una muestra aleatoria simple obtenida de $(\underline{X}, \mathcal{Y})$, siendo $\mathcal{Y}_i = (Y_i - I_i, Y_i, Y_i + S_i)$. Sea $\widehat{Y}_{\underline{x}}^c$ un modelo ajustado para los centros a partir de $\{\underline{X}_i, Y_i\}_{i=1}^n$ y consideremos el problema de regresión (en $I, S > 0$) para la distancia d_2 :

$$(P_2) \begin{cases} \Phi_{I,S} : \text{dom } X \rightarrow \mathcal{T}, & \Phi_{I,S}(\underline{x}) = (\widehat{Y}_{\underline{x}} - I, \widehat{Y}_{\underline{x}}, \widehat{Y}_{\underline{x}} + S), \\ E(I, S) = \sum_{i=1}^n d_2(\mathcal{Y}_i^*, \Phi_{\underline{X}_i}^*) & \text{mínimo.} \end{cases}$$

Entonces el problema de regresión (P_2) posee una única solución. Siguiendo la notación (4.3), el estimador \widehat{I} del parámetro I es la única solución no negativa de las $k - 1$ ecuaciones cúbicas:

$$\left\{ 2x^3 \sum_{i=r+1}^k \frac{m_i}{I'_i} + x^2 \left(\sum_{i=1}^r m_i - 2 \sum_{i=r+1}^k m_i \right) - \sum_{i=1}^r m_i (I'_i)^2 = 0 \quad \text{en el intervalo } [I'_r, I'_{r+1}] \right\}_{r=1}^{r=k-1},$$

e igualmente se obtiene el estimador \widehat{S} del parámetro S .

Corolario 4.2.8 Si todos los I_i son iguales, la mejor predicción para I que podemos hacer es ese mismo valor.

DEMOSTRACIÓN: En este caso, el error sería cero, ya que si $I_1 = I_2 = \dots = I_n = \bar{I}$, entonces:

$$E(x) = \sum_{i=1}^n \frac{(x - \bar{I})^2}{\max(x, \bar{I})} = n \cdot \frac{(x - \bar{I})^2}{\max(x, \bar{I})},$$

cuyo valor se anula para $x = \bar{I}$. ■

Resolución del problema de regresión difusa considerando la distancia d_∞

Como último caso, consideremos la distancia $d_\infty : \mathcal{T}_0 \times \mathcal{T}_0 \rightarrow [0, +\infty[$ dada por:

$$d_\infty(A, B) = \sup(\{|A(t) - B(t)| : t \leq 0\}) + \sup(\{|A(t) - B(t)| : t \geq 0\}).$$

Calculamos esta distancia (que se representa en la Figura 4.4) entre los números triangulares $\mathcal{Y}_i^* = (-I_i/0/S_i)$ y $\Phi_{\underline{x}}^* = (-I/0/S)$. Si suponemos que $0 < S \leq S_i$ (el otro caso es similar), la distancia d_∞ entre \mathcal{Y}_i^* y $\Phi_{\underline{x}}^*$ en \mathbb{R}_0^+ se alcanza justo en el punto $x = S$:

$$\sup(\{| \mathcal{Y}_i^*(t) - \Phi_{\underline{x}}^*(t) | : t \geq 0\}) = \frac{S_i - t}{S_i} \Big|_{t=S} = \frac{S_i - S}{S_i} = 1 - \frac{S}{S_i}.$$

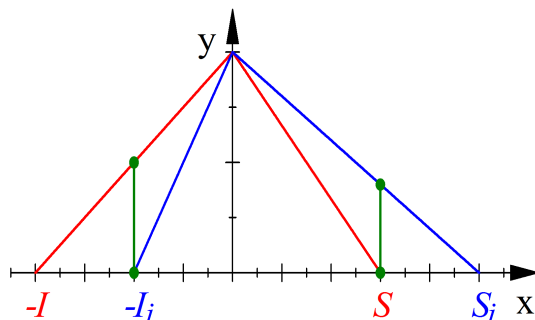


Figura 4.4: La distancia d_∞ es la suma de las longitudes de los segmentos verticales.

Si intercambiamos los papeles de S y de S_i , en realidad obtenemos:

$$\sup (\{|Y_i^*(t) - \Phi_x^*(t)| : t \geq 0\}) = 1 - \frac{\min(S, S_i)}{\max(S, S_i)}.$$

De esta forma:

$$d_\infty(\mathcal{Y}_i^*, \Phi_x^*) = \left(1 - \frac{\min(I, I_i)}{\max(I, I_i)}\right) + \left(1 - \frac{\min(S, S_i)}{\max(S, S_i)}\right).$$

Por consiguiente, el error total viene dado por:

$$\begin{aligned} E(I, S) &= \sum_{i=1}^n d_\infty(\mathcal{Y}_i^*, \Phi_x^*) = \sum_{i=1}^n \left(1 - \frac{\min(I, I_i)}{\max(I, I_i)}\right) + \sum_{i=1}^n \left(1 - \frac{\min(S, S_i)}{\max(S, S_i)}\right) = E(I) + E'(S), \\ \text{donde } E(I) &= n - \sum_{i=1}^n \frac{\min(I, I_i)}{\max(I, I_i)} \quad \text{y} \quad E'(S) = n - \sum_{i=1}^n \frac{\min(S, S_i)}{\max(S, S_i)}. \end{aligned} \quad (4.5)$$

De esta forma, podemos reducirnos a estudiar la función $E = E(I)$, pues un resultado similar se obtiene para E' . Como en el caso de la distancia d_1 , llamemos $\{I'_1, I'_2, \dots, I'_k\}$ a los diferentes valores que hay entre las amplitudes $\{I_1, I_2, \dots, I_n\}$ de manera que cada valor I'_j aparece en $\{I_1, I_2, \dots, I_n\}$ m_j veces y además $0 < I'_1 < I'_2 < \dots < I'_k$.

Lema 4.2.9 La función $E = E(I)$ definida en (4.5) posee mínimo absoluto en $]0, +\infty[$, y éste se encuentra en una de las amplitudes I_1, I_2, \dots, I_n .

- El valor mínimo absoluto de E en $]0, +\infty[$ es el menor valor de los siguientes k números:

$$\left\{ E(I'_j) \right\}_{j=1}^k = \left\{ n - m_j - \frac{1}{I'_j} \sum_{i=1}^{j-1} m_i I'_i - I'_j \sum_{i=j+1}^k \frac{m_i}{I'_i} \right\}_{j=1}^k.$$

- Si hay un único valor en $\{I_1, I_2, \dots, I_n\}$, él consigue el valor mínimo absoluto.
- Si hay dos únicos valores diferentes, el valor mínimo se consigue en el que tenga mayor multiplicidad, y si los dos tienen la misma, hay dos mínimos absolutos.

DEMOSTRACIÓN: Cada función:

$$\frac{\min(x, I'_i)}{\max(x, I'_i)} = \begin{cases} \frac{x}{I'_i}, & \text{si } 0 < x \leq I'_i, \\ \frac{I'_i}{x}, & \text{si } x > I'_i \end{cases}$$

es continua en $]0, +\infty[$ y derivable C^∞ en $]0, +\infty[\setminus\{I'_i\}$, por lo que E es continua en $]0, +\infty[$ y derivable C^∞ en $]0, +\infty[\setminus\{I'_1, I'_2, \dots, I'_k\}$. Es sencillo demostrar que la función E puede ser expresada como:

$$E(x) = \begin{cases} n - x \sum_{i=1}^k \frac{m_i}{I'_i}, & \text{si } 0 < x \leq I'_1, \\ n - \frac{1}{x} \sum_{i=1}^r m_i I'_i - x \sum_{i=r+1}^k \frac{m_i}{I'_i}, & \text{si } I'_r < x \leq I'_{r+1}, \\ n - \frac{1}{x} \sum_{i=1}^k m_i I'_i, & \text{si } x > I'_k. \end{cases}$$

Así, $\lim_{x \rightarrow 0^+} E(x) = N > 0$ y $\lim_{x \rightarrow +\infty} E(x) = N$. La primera derivada de E en $]0, +\infty[\setminus\{I'_1, I'_2, \dots, I'_k\}$ es:

$$E'(x) = \begin{cases} -\sum_{i=1}^k \frac{m_i}{I'_i}, & \text{si } 0 < x < I'_1, \\ \frac{1}{x^2} \sum_{i=1}^r m_i I'_i - \sum_{i=r+1}^k \frac{m_i}{I'_i}, & \text{si } I'_r < x < I'_{r+1}, \\ \frac{1}{x^2} \sum_{i=1}^k m_i I'_i, & \text{si } x > I'_k. \end{cases}$$

Así, E es estrictamente decreciente en $]0, I'_1[$ y es estrictamente creciente en $]I'_k, +\infty[$. Un argumento similar a los que ya hemos utilizado demuestra que E posee mínimo absoluto en $]0, +\infty[$ y que dicho mínimo está en el intervalo $[I'_1, I'_k]$. Por otro lado, E es dos veces derivable en $]0, +\infty[\setminus\{I'_1, I'_2, \dots, I'_k\}$ y su segunda derivada es, al menos en $]0, +\infty[\setminus\{I'_1, I'_2, \dots, I'_k\}$:

$$E''(x) = \begin{cases} 0, & \text{si } 0 < x < I'_1, \\ -\frac{2}{x^3} \sum_{i=1}^r m_i I'_i, & \text{si } I'_r < x < I'_{r+1}, \\ -\frac{2}{x^3} \sum_{i=1}^k m_i I'_i, & \text{si } x > I'_k. \end{cases}$$

Así, la función E es una línea recta (estrictamente) decreciente en $]0, I'_1[$ y además es una función (estrictamente) cóncava en los intervalos $]I'_1, I'_2[$, $]I'_2, I'_3[$, \dots , $]I'_{k-1}, I'_k[$ e $]I'_k, +\infty[$. Por consiguiente, la función E no puede poseer un mínimo absoluto en el interior de ningún intervalo $[I'_r, I'_{r+1}]$ (ya que en éste es cóncava). Así, el mínimo absoluto de la función E se encuentra en uno de los nodos I'_1, I'_2, \dots, I'_k ; concretamente, se encuentra en el nodo que consigue el menor valor de los siguientes k números:

$$\left\{ E(I'_j) \right\}_{j=1}^k = \left\{ n - m_j - \frac{1}{I'_j} \sum_{i=1}^{j-1} m_i I'_i - I'_j \sum_{i=j+1}^k \frac{m_i}{I'_i} \right\}_{j=1}^k.$$

Si hay un único valor en $\{I_1, I_2, \dots, I_n\}$, él consigue el valor mínimo absoluto. Si hay dos únicos valores diferentes, digamos $I'_1 < I'_2$, entonces la función es:

$$E(x) = n - m_1 \frac{\min(x, I'_1)}{\max(x, I'_1)} - m_2 \frac{\min(x, I'_2)}{\max(x, I'_2)}.$$

Al evaluarla en estos valores, encontramos:

$$\begin{aligned} E(I'_1) &= (m_1 + m_2) - m_1 - m_2 \frac{I'_1}{I'_2} = m_2 \left(1 - \frac{I'_1}{I'_2}\right); \\ E(I'_2) &= (m_1 + m_2) - m_1 \frac{I'_1}{I'_2} - m_2 = m_1 \left(1 - \frac{I'_1}{I'_2}\right). \end{aligned}$$

Por consiguiente, el valor mínimo se encuentra en la amplitud que posea mayor multiplicidad y, si es la misma, hay dos mínimos absolutos. ■

Del lema 4.2.9 podemos deducir cómo resolver la cuarta clase de regresión que nos planteamos.

Teorema 4.2.10 Sean $\underline{X}, \mathcal{Y}$ dos variables tales que \underline{X} es una variable aleatoria real e \mathcal{Y} es una variable aleatoria difusa que se relacionan a través de la ecuación (4.1). Sea $\{\underline{X}_i, \mathcal{Y}_i\}_{i=1}^n$ una muestra aleatoria simple obtenida de $(\underline{X}, \mathcal{Y})$, siendo $\mathcal{Y}_i = (Y_i - I_i / Y_i / Y_i + S_i)$. Sea \widehat{Y}_x^c un modelo ajustado para los centros a partir de $\{\underline{X}_i, Y_i\}_{i=1}^n$ y consideremos el problema de regresión (en $I, S > 0$) para la distancia d_∞ :

$$(P_\infty) \begin{cases} \Phi_{I,S} : \text{dom } X \rightarrow \mathcal{T}, & \Phi_{I,S}(x) = (\widehat{Y}_x - I / \widehat{Y}_x / \widehat{Y}_x + S), \\ E(I, S) = \sum_{i=1}^n d_\infty(\mathcal{Y}_i^*, \Phi_{\underline{X}_i}^*) & \text{mínimo.} \end{cases}$$

Entonces el problema (P_∞) posee alguna solución óptima. Siguiendo la notación (4.3), el estimador \widehat{I} del parámetro I es el valor I_1, I_2, \dots, I_n donde se alcanza el mínimo de los siguientes k números reales:

$$\left\{ E(I'_j) \right\}_{j=1}^k = \left\{ n - m_j - \frac{1}{I'_j} \sum_{i=1}^{j-1} m_i I'_i - I'_j \sum_{i=j+1}^k \frac{m_i}{I'_i} \right\}_{j=1}^k.$$

y el estimador \widehat{S} del parámetro S es el valor S_1, S_2, \dots, S_n donde se alcanza el mínimo de un conjunto numérico similar.

Resumen de todos los casos

A continuación, resumimos todos los casos que hemos estudiado en los apartados anteriores.

Teorema 4.2.11 Sean $\underline{X}, \mathcal{Y}$ dos variables tales que \underline{X} es una variable aleatoria real e \mathcal{Y} es una variable aleatoria difusa que se relacionan a través de la ecuación (4.1). Sea $\{\underline{X}_i, \mathcal{Y}_i\}_{i=1}^n$ una muestra

aleatoria simple obtenida de $(\underline{X}, \mathcal{Y})$, siendo $\mathcal{Y}_i = (Y_i - I_i/Y_i/Y_i + S_i)$. Sea $\widehat{Y}_{\underline{x}}^c$ un modelo ajustado para los centros a partir de $\{\underline{X}_i, Y_i\}_{i=1}^n$ y consideremos el problema de regresión (en $I, S > 0$) para alguna distancia d :

$$(P) \begin{cases} \Phi_{I,S} : M \rightarrow \mathcal{T}, & \Phi_{I,S}(\underline{x}) = (\widehat{Y}_{\underline{x}} - I / \widehat{Y}_{\underline{x}} / \widehat{Y}_{\underline{x}} + S), \\ E(I, S) = \sum_{i=1}^n d(\mathcal{Y}_i^*, \Phi_{\underline{X}_i}^*) & \text{mínimo.} \end{cases}$$

- Utilizando $d_1(A, B) = \|A - B\|_1 = \int_{\mathbb{R}} |A(t) - B(t)| dt$, el problema (P) posee como única solución los valores estimados

$$\widehat{I} = \text{Mediana}(I_1, \dots, I_n) \quad \text{y} \quad \widehat{S} = \text{Mediana}(S_1, \dots, S_n).$$

- Utilizando $d_1^2(A, B) = \left(\int_{-\infty}^0 |A(t) - B(t)| dt \right)^2 + \left(\int_0^{+\infty} |A(t) - B(t)| dt \right)^2$, el problema (P) posee como única solución los valores estimados

$$\widehat{I} = \text{Media}(I_1, \dots, I_n) \quad \text{y} \quad \widehat{S} = \text{Media}(S_1, \dots, S_n).$$

- Utilizando $d_2(A, B) = \|A - B\|_2^2 = \int_{\mathbb{R}} (A(t) - B(t))^2 dt$, el problema (P) posee una única solución. Siguiendo la notación (4.3), los estimadores de los parámetros \widehat{I} y \widehat{S} son las únicas soluciones no negativas de las ecuaciones cúbicas:

$$\left\{ \begin{array}{l} 2x^3 \sum_{i=r+1}^k \frac{m_i}{I_i'} + x^2 \left(\sum_{i=1}^r m_i - 2 \sum_{i=r+1}^k m_i \right) - \sum_{i=1}^r m_i (I_i')^2 = 0 \quad \text{en el intervalo } [I_r', I_{r+1}'] \\ \left. \right\}_{r=1}^{r=k-1}, \\ \left\{ \begin{array}{l} 2x^3 \sum_{i=r+1}^{k'} \frac{m_i'}{S_i'} + x^2 \left(\sum_{i=1}^r m_i' - 2 \sum_{i=r+1}^{k'} m_i' \right) - \sum_{i=1}^r m_i' (S_i')^2 = 0 \quad \text{en el intervalo } [S_r', S_{r+1}'] \\ \left. \right\}_{r=1}^{r=k'-1}. \end{array} \right.$$

respectivamente.

- Utilizando $d_{\infty}(A, B) = \sup(\{|A(t) - B(t)| : t \leq 0\}) + \sup(\{|A(t) - B(t)| : t \geq 0\})$, el problema (P) posee solución. Siguiendo la notación (4.3), el estimador del parámetro \widehat{I} es el valor I_1, I_2, \dots, I_n donde se alcanza el mínimo de los siguientes k números reales:

$$\left\{ n - m_j - \frac{1}{I_j'} \sum_{i=1}^{j-1} m_i I_i' - I_j' \sum_{i=j+1}^k \frac{m_i}{I_i'} \right\}_{j=1}^k,$$

y el estimador del parámetro \widehat{S} es el valor S_1, S_2, \dots, S_n donde se alcanza el mínimo de un conjunto numérico similar.

Observemos que la mediana es un estimador alternativo a la media (que, por ejemplo, es muy utilizado cuando la distribución de los datos es muy asimétrica) y es consistente en la mayoría de contextos de la vida real. En general, todos los estimadores propuestos son insesgados y consistentes, como muestran los experimentos de simulación que veremos. En particular, el estimador propuesto utilizando la distancia d_1^2 verifica todas las propiedades que se desean para un estimador. El objetivo del Teorema era proponer estimadores alternativos que puedan dar solución a casos concretos de la vida real y que puedan ser fácilmente calculados.

4.2.2. Algunos experimentos de simulación para evaluar las distancias anteriores

A continuación, llevamos a cabo un experimento de simulación para evaluar la eficiencia de las estimaciones de los parámetros que se obtienen utilizando las diferentes distancias comentadas en el apartado anterior, considerando un problema lineal y otro no lineal, y utilizando muestras de diferentes tamaños. Para que la comparación sea más robusta, se supondrá que los centros de los residuos siguen una distribución Normal y las amplitudes se distribuyen según una Chi-cuadrado. El experimento de simulación se ha llevado a cabo con 10000 repeticiones y se ha determinado el error cuadrático medio (en los sucesivo, ECM) de cada estimación. Tanto la computación de los datos como la estimación de los parámetros como el cálculo de los ECM se han desarrollado con el paquete matemático *Mathematica*, proporcionado por la Universidad de Jaén.

Simulación 1

En la primera simulación que desarrollamos, consideramos el siguiente modelo de regresión lineal:

$$\mathcal{Y} = 0.7X + (\varepsilon^I / \varepsilon^c / \varepsilon^S),$$

donde X es una variable aleatoria que varía de uno en uno desde 1 hasta 20 (en un segundo caso, supondremos que varía hasta 100) y suponemos que los residuos se distribuyen según:

$$\varepsilon^c \sim \mathcal{N}(1, 0.5), \quad \varepsilon^c - \varepsilon^I \sim \chi_{30}^2/40 \quad \text{y} \quad \varepsilon^S - \varepsilon^c \sim \chi_{15}^2/10,$$

es decir, la *función de regresión* en x sería:

$$\hat{\mathcal{Y}}(x) = 0.7x + (0.25/1/2.5).$$

Por consiguiente, estamos considerando la solución computacional de un modelo lineal de parámetros $a = 0.7$, $b = 1$, $I = 0.75$ y $S = 1.5$. Los centros se relacionan con la variable explicativa a través del método clásico de mínimos cuadrados, por lo que todas las estimaciones para a y b ofrecen el mismo resultado. Utilizando este método, el coeficiente de determinación medio en todos los casos resultó muy próximo a 1, demostrando que los modelos son muy precisos.

En las Tablas 4.1 ($n = 20$) y 4.2 ($n = 100$) presentamos la media de las estimaciones y del error cuadrático medio estimado para cada uno de ellos, calculado como

$$\text{ECM}(\theta) = \frac{1}{10000} \sum_{r=1}^{10000} (\theta - \hat{\theta})^2.$$

González-Rodríguez y otros (2009) también han estudiado modelos de regresión lineal simple para variables aleatorias difusas. En la última fila de las tablas 4.1 y 4.2 presentamos los resultados que se que hemos obtenido aplicando su técnica. Si comparamos los resultados obtenidos

Tabla 4.1: Estimaciones por el método de Monte Carlo y ECMs de los parámetros utilizando tamaños muestrales de $n = 20$ en el caso lineal.

	\hat{a}	$\text{ECM}(a)_{\times 10^3}$	\hat{b}	$\text{ECM}(b)$	\hat{I}	$\text{ECM}(I)_{\times 10^3}$	\hat{S}	$\text{ECM}(S)$
d_1	0.7001	0.3738	0.996	0.0530	0.735	2.899	1.440	0.0242
d_1^2	Idem	Idem	Idem	Idem	0.750	1.858	1.500	0.0147
d_2	Idem	Idem	Idem	Idem	0.739	1.950	1.458	0.0161
d_∞	Idem	Idem	Idem	Idem	0.737	3.661	1.456	0.0317
GBCL	0.7002	0.4282	0.997	0.0592	0.750	1.858	1.500	0.0146

Tabla 4.2: Estimaciones por el método de Monte Carlo y ECMs de los parámetros utilizando tamaños muestrales de $n = 100$ en el caso lineal.

	\hat{a}	$\text{ECM}(a)_{\times 10^6}$	\hat{b}	$\text{ECM}(b)_{\times 10^6}$	\hat{I}	$\text{ECM}(I)_{\times 10^3}$	\hat{S}	$\text{ECM}(S)_{\times 10^3}$
d_1	0.7000	3.063	1.0002	0.0104	0.734	0.827	1.435	8.556
d_1^2	Idem	Idem	Idem	Idem	0.750	0.372	1.500	2.992
d_2	Idem	Idem	Idem	Idem	0.739	0.495	1.456	4.805
d_∞	Idem	Idem	Idem	Idem	0.736	0.921	1.449	8.957
GBCL	0.7000	3.508	1.0010	0.0116	0.750	0.373	1.500	2.992

utilizando las diferentes distancias que introducimos y los obtenidos por los mencionados autores, observamos que los mejores resultados se produjeron tanto por su método como utilizando la distancia d_1^2 . Una comparación rápida entre las Tablas 4.1 y 4.2 muestra que la eficiencia de las estimaciones aumenta cuando el tamaño de la muestra crece.

Tabla 4.3: Estimaciones por el método de Monte Carlo y ECM de los parámetros utilizando tamaños muestrales de $n = 20$ en el caso parabólico

	$\hat{a} = (\hat{a}_1, \hat{a}_2)$	$\text{ECM}(\underline{a})_{\times 10^3}$	\hat{b}	$\text{ECM}(b)$	\hat{I}	$\text{ECM}(I)_{\times 10^3}$	\hat{S}	$\text{ECM}(S)_{\times 10^3}$
d_1	(0.501, 1.100)	(1.816, 8.670)	-0.402	0.1333	0.736	2.902	0.440	10.018
d_1^2	Idem	Idem	Idem	Idem	0.751	1.875	0.499	5.030
d_2	Idem	Idem	Idem	Idem	0.740	1.966	0.460	3.326
d_∞	Idem	Idem	Idem	Idem	0.739	3.723	0.464	12.874

Tabla 4.4: Estimaciones por el método de Monte Carlo y ECMs de los parámetros utilizando tamaños muestrales de $n = 100$ en el caso parabólico

	$\hat{\underline{a}} = (\hat{a}_1, \hat{a}_2)$	ECM(\underline{a}) $\times 10^6$	\hat{b}	ECM(b)	\hat{I}	ECM(I) $\times 10^3$	\hat{S}	ECM(S) $\times 10^3$
d_1	(0.500, 1.100)	(1.919, 2.728)	-0.401	0.0231	0.734	0.843	0.435	5.372
d_1^2	Idem	Idem	Idem	Idem	0.750	0.377	0.499	0.992
d_2	Idem	Idem	Idem	Idem	0.739	0.502	0.458	2.667
d_∞	Idem	Idem	Idem	Idem	0.736	0.937	0.457	4.321

Simulación 2

En el siguiente experimento de simulación consideramos el siguiente modelo de regresión no lineal:

$$\mathcal{Y} = 1.1X^2 + 0.5X + (\varepsilon^I/\varepsilon^c/\varepsilon^S),$$

donde X es una variable aleatoria que varía de uno en uno desde 1 hasta 20 (en un segundo caso, la suponíamos hasta 100) y suponíamos que los residuos se distribuían según:

$$\varepsilon^c \sim \mathcal{N}(-0.4, 0.5), \quad \varepsilon^c - \varepsilon^I \sim \chi_{30}^2/40 \quad \text{y} \quad \varepsilon^S - \varepsilon^c \sim \chi_5^2/10,$$

es decir, la *función de regresión* en x viene dada por las siguientes funciones cuadráticas:

$$\hat{\mathcal{Y}}(x) = 1.1x^2 + 0.5x + (-1.15/-0.4/0.1)$$

Por tanto, los parámetros considerados son $\underline{a} = (0.5, 1.1)$, $b = -0.4$, $I = 0.75$, $S = 0.5$. Los centros se relacionan con la variable explicativa a través del método clásico de mínimos cuadrados, por lo que todas las estimaciones para \underline{a} y b ofrecen los mismos resultados. Utilizando este método, la prueba F mostró que los modelos eran muy precisos. En las Tablas 4.3 ($n = 20$) y 4.4 ($n = 100$) presentamos la media de las estimaciones y del error cuadrático medio estimado para cada uno de ellos. Los resultados obtenidos mostraron que la distancia d_1^2 nuevamente aportó los mejores resultados. La comparación entre las Tablas 4.3 y 4.4 muestra que la eficiencia de las estimaciones aumenta cuando el tamaño de la muestra crece.

4.2.3. Un estudio comparativo: un modelo de regresión difusa en el que los valores de las amplitudes dependen de las variables explicativas

En esta sección analizamos el caso en el que las funciones que determinan los valores esperados de las amplitudes dependen de las variables explicativas. De esta forma, supongamos que los valores esperados de los residuos dependen de la variable aleatoria \underline{X} , es decir,

$$E(\varepsilon^c - \varepsilon^I | \underline{X}) = I(\underline{X}), \quad E(\varepsilon^S - \varepsilon^c | \underline{X}) = S(\underline{X}).$$

Dado que en el apartado anterior los mejores resultados se han utilizando la distancia d_1^2 , utilizaremos el método de mínimos cuadrados (si bien, otras distancias pueden también ser utilizadas). Denotaremos por $\hat{I}(\underline{x})$ y por $\hat{S}(\underline{x})$ a los modelos ajustados para ambas funciones utilizando el método de mínimos cuadrados.

La aplicabilidad del método propuesto lo mostramos a través del estudio comparativo realizado. Los resultados obtenidos con método propuesto se comparan con los resultados obtenidos por los métodos descritos en la sección 4.1.1.

Para evaluar la aplicabilidad del método que proponemos, utilizamos los mismos cinco pares de observaciones que fueron propuestos, por primera vez, por Tanaka y sus colaboradores en 1987. Estos autores diseñaron un ejemplo ilustrativo de su modelo de regresión difusa para tratar el caso en el que se trabaja con una única variable real explicativa y una variable dependiente difusa cuyos valores son números difusos triangulares simétricos. En la Tabla 4.5 presentamos los valores considerados. Es verdad que este conjunto de datos no posee el nivel de detalle y complejidad que los usados en otros estudios, pero tienen la ventaja de que estos mismos datos han sido utilizados por muchos autores para describir sus diferentes técnicas de regresión, para evaluarlas experimentalmente y para comparar su metodología con la de otros autores.

Tabla 4.5: Datos y amplitudes (*spreads*)

X	Y^I	Y^c	Y^S	I_i	S_i
1	6.2	8	9.8	1.8	1.8
2	4.2	6.48	6	2.2	2.2
3	6.9	9.5	12.1	2.6	2.6
4	10.9	13.5	16.1	2.6	2.6
5	10.6	13	15.4	2.4	2.4

En la Tabla 4.6 hacemos un listado de los modelos propuestos por otros autores (véase la sección 4.1.1) utilizando las mismas cinco observaciones que se relacionan en la Tabla 4.5. Varios autores señalaron que el método propuesto por Tanaka y sus colaboradores [119] poseía algunas desventajas y lo modificaron, o bien desarrollaron sus propias metodologías para superar estos problemas. El modelo de regresión THW estima los coeficientes de la regresión difusa utilizando programación lineal. Este modelo posee una pendiente numérica real y una ordenada en el origen difusa. Los modelos KB, DM, WT y HBS poseen tanto pendiente como ordenada en el origen difusas. Cuando las variables explicativas son reales y las variables dependientes son números difusos simétricos, el modelo WT ofrece los mismos resultados que el modelo DM. Los modelos NN y CH poseen una pendiente numérica y el modelo KC posee coeficientes numéricos y un

Tabla 4.6: Modelos de regresión obtenidos por otros autores a partir de los datos que se muestran en la Tabla 4.5.

Autores	Modelo de regresión difusa
Tanaka y otros 1989	$\widehat{\mathcal{Y}}_{THW}(x) = (0/3.85/7.7) + 2.1x$
Kim y Bishu 1998	$\widehat{\mathcal{Y}}_{KB}(x) = (3.11/4.95/6.84) + (1.55/1.71/1.82)x$
Diamond 1988	$\widehat{\mathcal{Y}}_{DM}(x) = (3.11/4.95/6.79) + (1.55/1.71/1.87)x$
Nasrabadi y Nasrabadi 2004	$\widehat{\mathcal{Y}}_{NN}(x) = (2.36/4.86/7) + 1.73x$
Wu Tseng 2002	$\widehat{\mathcal{Y}}_{WT}(x) = (3.11/4.95/6.79) + (1.55/1.71/1.87)x$
Hojati y otros 2005	$\widehat{\mathcal{Y}}_{HBS}(x) = (5.1/6.75/8.4) + (1.1/1.25/1.4)x$
Kao y Chyu 2003	$\widehat{\mathcal{Y}}_{KC}(x) = 4.81 + 1.72x + (-2.2/0.12/2.44)$
Chen y Hsueh 2009	$\widehat{\mathcal{Y}}_{CH}(x) = 1.71x + (2.63/4.95/7.27)$

término de error difuso.

Si aplicamos el Teorema 4.2.11 a los datos mostrados en la tabla 4.5, en primer lugar vamos a ajustar un modelo para los centros de la variable difusa utilizando el criterio de mínimos cuadrados. El modelo que mejor se ajusta es:

$$\widehat{Y}^c(x) = 4.95 + 1.71x, \quad \rho_{y^c X} = 0.872,$$

donde ρ denota el coeficiente de correlación. Si utilizamos la distancia d_1^2 , encontramos las constantes $\widehat{I} = \widehat{S} = 2.32$ para estimar las amplitudes, lo que llevaría al modelo lineal:

$$\widehat{\mathcal{Y}}(x) = 1.71x + (2.63/4.95/7.27).$$

Tabla 4.7: Comparación del error total de estimación obtenido para los diferentes modelos.

Explan. var.	Resp. var.	E_i								
		THW	KB	DM	KC	NN	HBS	WT	CH	RRM
1	(6.2, 8, 9.8)	6.304	1.816	1.816	1.974	2.663	0.000	1.816	1.931	1.796
2	(4.2, 6.48, 6)	4.084	3.834	3.882	3.857	3.035	8.154	3.882	3.888	3.881
3	(6.9, 9.5, 12.1)	1.204	0.363	0.376	0.376	0.176	1.125	0.376	0.376	0.358
4	(10.9, 13.5, 16.1)	2.344	3.074	2.931	2.936	3.649	3.124	2.931	2.963	2.928
5	(10.6, 13, 15.4)	2.874	0.215	0.279	0.269	0.112	0.000	0.279	0.253	0.275
$\sum_i E_i$		16.809	9.302	9.283	9.412	9.636	12.403	9.283	9.411	9.238

Sin embargo, si hacemos que las amplitudes dependan de la variable explicativa, podemos obtener mejores resultados. En efecto, consideremos que I y S son funciones de X . De esta forma, haciendo uso de un programa estadístico de regresión (por ejemplo, nosotros hemos utilizado SPSS), podemos buscar varios modelos para las funciones I y S , tanto lineales como no lineales, y elegir aquéllos modelos que mejor ajusten los datos (los que posean un mayor coeficiente de correlación). En nuestro caso, dado que la variable respuesta es simétrica, seleccionamos las siguientes expresiones no lineales por ser aquéllas que mejor se ajustan a los datos de entre varios modelos:

$$\hat{I}(x) = \hat{S}(x) = \frac{1}{0.344 + 0.208/x}, \quad \rho_{IX} = 0.949.$$

Por consiguiente, el modelo que obtenemos siguiendo esta metodología es:

$$\begin{aligned} \hat{Y}_{RRM}(x) &= \hat{Y}^c(x) + (-\hat{I}(x)/0/\hat{S}(x)) = \\ &= \left(4.95 + 1.71x - \frac{1}{0.344 + 0.208/x} \middle/ 4.95 + 1.71x \middle/ 4.95 + 1.71x - \frac{1}{0.344 + 0.208/x} \right). \end{aligned}$$

Sea E_i el error cuadrático medio obtenido entre las variables de respuesta observadas y estimada en el i -ésimo caso. la tabla 4.7 muestra los errores totales al estimar las variables repuesta de todos los modelos basándose en el criterio de error de mínimos cuadrados (no obstante, podrían haberse utilizado otros criterios). La última columna presenta los resultados ontendidos siguiendo la metodología que proponemos. Nuestro error total de estimación, $\sum_i E_i$, es el menor de la metodologías propuestas.

Obsérvese que el modelo que proponemos únicamente presenta una expresión lineal para los centros, ya que las amplitudes se ajustan a un modelo *doble inverso*. Si hubiésemos elegido también modelos lineales para las amplitudes, hubiésemos obtenido el modelo

$$\hat{Y}(x) = 1.71x + (2.63/4.95/7.27),$$

cuyo error total de estimación es de 9.485, el cual es un valor razonable en comparación con los errores totales obtenidos por otros autores que utilizan otras metodologías.

Al analizar estos resultados, podemos observar que nuestra técnica es método sencillo de aplicar en la práctica para el análisis del problema considerado, es decir, en el que las variables de entrada son reales y la variable de salida es difusa triangular.

4.3. Un modelo de regresión no lineal para la obtención de biomasa a partir de los restos de la poda del olivo

Para acabar el Capítulo, presentamos una aplicación de los resultados anteriores a un contexto de la vida real a modo de ejemplo ilustrativo.

En la actualidad parece evidente que la biomasa primaria y residual será en un futuro inmediato la principal fuente de recursos con que contará la Humanidad para la obtención de alimentos, productos orgánicos y combustibles líquidos y gaseosos, dado el agotamiento y creciente encarecimiento de la biomasa fósil.

Se calcula que la energía total contenida en la biomasa en estado de crecimiento equivale a un 70 % de las reservas conocidas de carbón en el mundo y a casi cuatro veces las de petróleo. Sin embargo, cantidades mayores de materia orgánica se encuentran en estos momentos en distintos grados de fosilización. Se ha estimado que los almacenes fósiles (carbón, petróleo y gas natural) representan del orden de 40 veces la biomasa de crecimiento.

En este sentido los residuos de carácter lignocelulósico constituyen una materia prima renovable y abundante, cuyo aprovechamiento es deseable. Será interesante establecer vías de aprovechamiento para estos materiales que, al estar ligados a la producción de alimentos, tienen la renovabilidad asegurada y cuyo abanico de usos es tan amplio como el que puedan poseer los combustibles fósiles de los que son precursores.



Figura 4.5: Residuo de la poda del olivo

La poda del olivo es una operación esencial de rejuvenecimiento del árbol para prepararlo de cara a la siguiente cosecha. Esta operación genera periódicamente unos restos que, en principio, son muy abundantes, son renovables y suponen un coste muy bajo (quizá ninguno). La provincia de Jaén acumula el 40 % de la superficie total dedicada al cultivo del olivo en España. La poda del olivo (véase la Figura 4.5) genera un volumen muy grande de biomasa en esta provincia. Estos residuos lignocelulósicos son una fuente potencial de xilosa, que puede ser utilizada como materia prima en la producción de xilitol, que es un producto de gran valor. La hidrólisis ácida de los restos de la poda del olivo está siendo actualmente estudiada en Jaén para determinar la naturaleza, la influencia de la concentración de ácido y la temperatura en la producción de azúcares, principalmente de xilosa.

El modelo cinético utilizado satisfactoriamente para la hidrólisis de hemicelulosa involucra

reacciones irreversibles pseudo-homogéneas de primer orden. El primer modelo fue propuesto por Saeman en 1945 (véase [104]) y puede ser generalizado según el siguiente esquema:



Basándonos en este modelo de reacción y resolviendo las ecuaciones diferenciales que conlleva, se llega a que la concentración (M) del monómero como función del tiempo (t) puede ser representada por la siguiente relación no lineal:

$$M = \frac{k_1 P_0}{k_1 - k_2} (e^{-k_2 t} - e^{-k_1 t}), \quad (4.6)$$

donde P_0 representa la concentración del polímero en el momento inicial $t = 0$, k_1 es la proporción de monómero liberado (en mín^{-1}) y k_2 es la proporción de monómero descompuesto (en mín^{-1}). Para determinar los valores estimados de la concentración de xilosa, glucosa, furfural, etc., y extraerlos cuando mayores son sus concentraciones, se resuelve el correspondiente problema de regresión.

Tabla 4.8: Tiempo (t) y concentración de xilosa (Y^I, Y^c, Y^S) utilizando 1N HCl a 100°C

t (min.)	Y^I	Y^c	Y^S
30	3.24	4.88	5.43
60	8.06	9.81	9.88
90	9.96	10.09	10.63
120	10.45	11.05	11.13
150	10.24	10.73	10.95
180	9.31	10.69	10.85
210	9.47	9.99	10.28
240	9.19	9.22	9.30

Experimentalmente, de cara obtener un modelo cinético para la concentración de xilosa, nos damos cuenta de que podemos observar de manera precisa la variable independiente, que es el tiempo, pero la concentración de xilosa no puede medirse de manera exacta, sino aproximada (véase, por ejemplo, la Tabla 2 en Aguilar y otros [1]). El objetivo del experimento es determinar la relación de dependencia entre la concentración de xilosa y el tiempo. Basándonos en diversos experimentos, la variable respuesta puede interpretarse como una variable aleatoria difusa (véase la Tabla 4.8), observada en ocho momentos diferentes (cada 30 minutos) utilizando una concentración 1N HCl a 100°C.

Los centros han sido relacionados con la variable explicativa a través del método de regresión clásica por mínimos cuadrados y los parámetros cinéticos del modelo (4.6) se muestra en la Tabla

Tabla 4.9: Parámetros estadísticos y cinéticos para la concentración de xilosa utilizando 1N HCl a 100°C

\hat{k}_1	\hat{k}_2	R^2	F-Test Prob.
0.01285	0.00397	0.9977	0.9934

4.9. Por tanto, el modelo de regresión ajustado para los centros es

$$\hat{Y}^c(t) = 27.109(e^{-0.00397t} - e^{-0.01285t}). \quad (4.7)$$

En este caso, los valores esperados de las amplitudes no dependen de la variable explicativa (los coeficientes de correlación de los diferentes modelos ajustados con SPSS para las amplitudes indican que ninguno de ellos puede ser considerado como significativo). Esto nos lleva a concluir que los mejores valores que podemos obtener son constantes. Aplicando los resultados del Teorema 4.2.11, el error total en la estimación utilizando la distancia d_1 es 0.290, utilizando d_1^2 es 0.264, utilizando d_2 es 0.269 y utilizando d_∞ es 0.298. Por consiguiente, los mejores resultados se obtienen utilizando la distancia d_1^2 . En tal caso, el modelo de regresión que se obtiene para la concentración de xilosa $\hat{Y}(t)$ dependiendo del tiempo t es:

$$\hat{Y}(t) = \hat{Y}^c(t) + (-0.82/0/0.25) = \left(\hat{Y}^c(t) - 0,82 / \hat{Y}^c(t) / \hat{Y}^c(t) + 0,25 \right),$$

donde $\hat{Y}^c(t)$ está dada en (4.7).

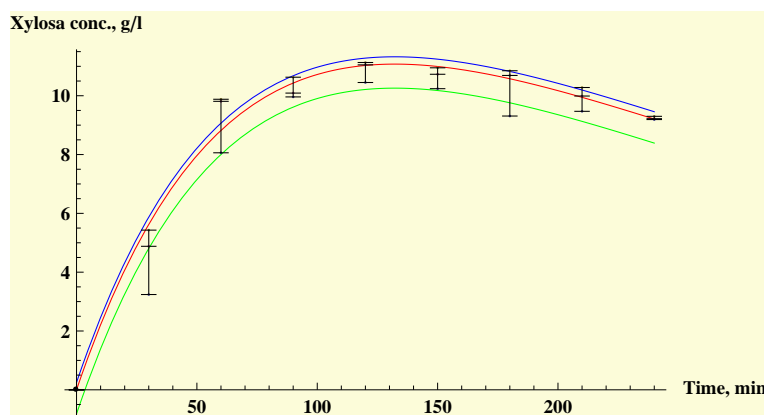


Figura 4.6: Los valores observados y los modelos de regresión no lineales estimados

La Figura 4.6 muestra los valores observados y el modelo de regresión difusa no lineal. Basándonos en esta ecuación, podemos deducir que la concentración máxima de xilosa a una concentración 1N HCl a 100°C se alcanza en el instante $t = 132.22$ mín, lo que nos lleva a que

la máxima concentración de xilosa no es un valor exacto, sino que viene determinada por el número difuso triangular (10.26, 11.08, 11.32) *gr xilosa/l*.

Artículos en los que se fundamenta la presente Memoria

1. J. Martínez-Moreno, A. Roldán, C. Roldán, *A note on the \mathcal{L} -fuzzy Banach's contraction principle*, Chaos, Solitons and Fractals 41 (5) (2009) 2399–2400.
2. C. Roldán, A. Roldán, J. Martínez-Moreno, *A fuzzy regression model based on distances and random variables with crisp input and fuzzy output data: a case study in biomass production*, Soft Computing 16 (5) (2012) 785–795.
3. A. Roldán, J. Martínez-Moreno, C. Roldán, *On interrelationships between fuzzy metric structures* (en proceso de revisión menor en *Iranian Journal of Fuzzy Systems*).
4. J. Martínez-Moreno, A. Roldán, C. Roldán, *A note on “Coupled fixed point theorems for contractions in intuitionistic fuzzy normed spaces”* (enviado).

Artículos posteriores a esta Memoria

1. A. Roldán, C. Roldán, J. Martínez-Moreno, C. Aguilar, *Estimation of a fuzzy regression model using fuzzy distances* (enviado).
2. A. Roldán, J. Martínez-Moreno, C. Roldán, *Some applications of the study of the image of a fuzzy number: discrete fuzzy numbers, operations, regression and ranking* (enviado).

Bibliografía

- [1] R. Aguilar, J.A. Ramírez, G. Garrote, M. Vázquez, *Kinetic study of the acid hydrolysis of sugar cane bagasse*, Journal for Food Engineering 55 (2002) 309–318.
- [2] C. Alaca, D. Turkoglu, C. Yildiz, *Fixed points in intuitionistic fuzzy metric spaces*, Chaos, Solitons and Fractals 29 (5) (2006) 1073–1078.
- [3] K.T. Atanassov, *Intuitionistic fuzzy sets*, Fuzzy Sets and Systems 20 (1986) 87–96.
- [4] K.T. Atanassov, *Intuitionistic fuzzy sets*, in: V. Sgurev, Ed., VII ITKR's Session, Sofia, Junio de 1983 (deposed in Central Science-Technical Library of Bulgarian, Academy of Science, 1697/84, Bulgaria).
- [5] S. Banach, *Theorie les operations lineaires*, Manograie Matematyeczne Warsaw Poland, 1932.
- [6] I. Beg, M. Abbas, *Common fixed points of Banach operator pair on fuzzy normed spaces*, Fixed Point Theory, 12 (2011), No. 2, 285–292.
- [7] V. Berinde, *Approximating common fixed points of noncommuting almost contractions in metric spaces*, Fixed Point Theory, 11 (2) (2010) 179–188.
- [8] V. Berinde, M. Borcut, *Tripled fixed point theorems for contractive type mappings in partially ordered metric spaces*, Nonlinear Anal. TMA 74 (2011), 4889–4897.
- [9] M. Berzig, B. Samet, *An extension of coupled fixed point's concept in higher dimension and applications*, Computers and Mathematics with Applications (2012), doi:10.1016/j.camwa.2012.01.018.
- [10] T.G. Bhaskar, V. Lakshmikantham, *Fixed point theorems in partially ordered metric spaces and applications*, Nonlinear Anal. TMA 65 (2006), No. 7, 1379–1393.
- [11] M. Borcut, V. Berinde, *Tripled coincidence theorems for contractive type mappings in partially ordered metric spaces*, Applied Mathematics and Computation (2011). In press.

- [12] R.M. Brooks, K. Schmitt, *The contraction mapping principle and some applications*, Electronic Journal of Differential Equations, Monograph 09 (2009), 90 páginas. Disponible en <http://ejde.math.txstate.edu> or <http://ejde.math.unt.edu>.
- [13] F. Browder, *Nonexpansive nonlinear operators in a Banach space*, Proc. Nat. Acad. Sci. USA 54 (1965) 1041–1044.
- [14] J.G. Camarena, V. Gregori, S. Morillas, A. Sapena, *Fast detection and removal of impulsive noise using peer groups and fuzzy metrics*, J. Visual Commun. Image Representation 19 (2008) 20–29.
- [15] J. Caristi, *Fixed point theorems for mappings satisfying inwardness conditions*, Trans. Amer. Math. Soc. 215 (1976) 241–251.
- [16] F. Castro-Company, S. Romaguera, *Experimental results for information system based on accesses locality via intuitionist fuzzy metrics*, Open Cybern. Syst. J. 2 (2008) 158–172.
- [17] Y. Chen, L. Chen, *A non-linear possibilistic regression approach to model functional relationships in product manning*, Int J Adv Manuf Technol 28 (2006) 1175–1181.
- [18] L.H. Chen, C.C. Hsueh, *Fuzzy regression models using the least-squares method based on the concept of distance*, IEEE Transactions on Fuzzy Syst 17 (6) (2009) 1259–1272.
- [19] J.H. Chiang, P.Y. Hao, *Support vector learning mechanism for fuzzy rule-based modeling: a new approach*, IEEE Transactions on Fuzzy Syst 12 (1) (2004) 1–11.
- [20] M. Chicourrat, C.D. Horváth, *Pretopologies, preuniformities and probabilistic metric spaces*, Acta Math. Hungar. 110 (1-2) (2006) 91-116.
- [21] Y.J. Cho, *Fixed points in fuzzy metric spaces*, Journal of Fuzzy Mathematics 5 (4) (1997) 949–962.
- [22] Y.J. Cho, M.T. Grabiec, V. Radu, *On nonsymmetric topological and probabilistic structures*, Nova Science Publishers, Inc., New York, 2006.
- [23] A. Colubi, *Statistical inference about the means of fuzzy random variables: applications to the analysis of fuzzy- and real-valued data*, Fuzzy Sets and Systems 160 (2009) 344–356.
- [24] R. Coppi, P. D’Urso, P. Giordani, A. Santoro, *Least squares estimation of a linear regression model with LR fuzzy response*, Comp Stat Data Anal 51 (2006) 267–286.
- [25] I. Couso, L. Sánchez, *Higher order models for fuzzy random variables*, Fuzzy Set and Syst 159 (2007) 237–258.
- [26] Z. Deng, *Fuzzy pseudo-metric space*, J. Math. Anal. Appl. 86 (1982) 74–95.

-
- [27] G. Deschrijver, E.E. Kerre, *On the position of intuitionistic fuzzy set theory in the framework of theories modelling imprecision*, Information Sciences 177 (2007) 1860–1866.
- [28] P. Diamond, *Fuzzy least squares*, Inform Sci 46 (1988) 141–157.
- [29] D. Dubois, M.A. Lubiano, H. Prade, M.A. Gil, P. Grzegorzewski, O. Hryniewicz, *Soft methods for handling variability and imprecision*, Advances in Soft Computing, Springer, Berlin, 2008.
- [30] D. Dubois, H. Prade, *Operations on fuzzy numbers*, International Journal of Systems Science 9 (1978), no. 6, 613–626.
- [31] D. Dubois, H. Prade, *Possibility Theory*. New York, Plenum Press, 1988.
- [32] D. Dubois, H. Prade, *Fuzzy elements in a fuzzy set*. Proc. 10th Internat. Fuzzy Systems Assoc. (IFSA) Congr. (2005) 55–60, Beijing, Springer.
- [33] M. Edelstein, *On fixed and periodic points under contractive mappings*, J. London Math. Soc. 37 (1962), 74–79.
- [34] R.J. Egbert, *Products and quotients of probabilistic metric spaces*, Pacific Journal Of Mathematics, 24 (3) (1968) 437–455.
- [35] M.S. El Naschie, *On the verifications of heteritic strings theory and $\epsilon^{(\infty)}$ theory*, Chaos, Solitons and Fractals 11 (2) (2000) 2397–2407.
- [36] M.A. Erceg, *Metric spaces in fuzzy set theory*, J. Math. Anal. Appl. 69 (1979) 205–230.
- [37] M. Eshaghi Gordji, H. Baghani, Y.J. Cho, *Coupled fixed point theorems for contractions in intuitionistic fuzzy normed spaces*, Mathematical and Computer Modelling 54 (2011) 1897–1906.
- [38] J.X. Fang, *Fixed point theorems of local contraction mappings on Menger spaces*, Appl. Math. Mech. 12 (1991) 363–372.
- [39] J.X. Fang, *On fixed point theorem in fuzzy metric spaces*, Fuzzy Sets and Systems 46 (1) (1992) 107–113.
- [40] J.X. Fang, *Common fixed point theorems of compatible and weakly compatible maps in Menger spaces*, Nonlinear Anal. TMA 71 (5-6) (2009), 1833–1843.
- [41] M.B. Ferraro, R. Coppi, G. González-Rodríguez, A. Colubi, *A linear regression model for imprecise response*, International Journal of Approximate Reasoning 51 (2010) 759–770.
- [42] A. George, P.V. Veeramani, *On some results in fuzzy metric spaces*, Fuzzy Sets and Systems 64 (1994) 395–399.

-
- [43] A. George, P.V. Veeramani, *On some results of analysis for fuzzy metric spaces*, Fuzzy Sets and Systems 90 (1997) 365–368.
- [44] J.A. Goguen, *\mathcal{L} -fuzzy sets*, Journal of Mathematical Analysis and Applications 18 (1967) 145–174.
- [45] D. Gómez, J. Montero, J. Yáñez, *A coloring fuzzy graph approach for image classification*, Information Sciences 176 (24) (2006) 3645–3657.
- [46] G. González-Rodríguez, A. Blanco, A. Colubi, M.A. Lubiano, *Estimation of a simple linear regression model for fuzzy random variables*, Fuzzy Sets and Systems 160 (2009) 357–370.
- [47] M. Grabiec, *Fixed points in fuzzy metric spaces*, Fuzzy Sets and Systems 27 (3) (1988) 385–389.
- [48] M.T. Grabiec, *Fixed points in fuzzy metric spaces*, Fuzzy Sets and Systems 64 (1994) 395–399.
- [49] M.T. Grabiec, Y.J. Cho, R. Saadati, *Families of quasi-pseudo-metrics generated by probabilistic quasi-pseudo-metric spaces*, Surveys in Mathematics and its Applications 2 (2007) 123–143.
- [50] V. Gregori, S. Morillas, A. Sapena, *On a class of completable fuzzy metric spaces*, Fuzzy Sets and Systems 160 (2010) 2193–2205.
- [51] V. Gregori, S. Romaguera, *Some properties of fuzzy metric spaces*, Fuzzy Sets and Systems 115 (2000) 485–489.
- [52] V. Gregori, S. Romaguera, *On completion of fuzzy metric spaces*, Fuzzy Sets and Systems 130 (2002) 399–404.
- [53] V. Gregori, S. Romaguera, *Characterizing completable fuzzy metric spaces*, Fuzzy Sets and Systems 144 (2004) 411–420.
- [54] V. Gregori, S. Romaguera, P.V. Veeramani, *A note on intuitionistic fuzzy metric spaces*, Chaos, Solitons and Fractals 28 (2006) 902–905.
- [55] V. Gregori, A. Sapena, *On fixed point theorem in fuzzy metric spaces*, Fuzzy Sets and Systems 125 (2) (2002) 245–252.
- [56] O. Hadžić, *Fixed point theorems for multi-valued mappings in some classes of fuzzy metric spaces*, Fuzzy Sets and Systems 29 (1989) 115–125.
- [57] O. Hadžić, E. Pap, *The fixed point theory in probabilistic metric spaces*, Kluwer Academic Publishers, Dordrecht, 2001.

-
- [58] S. Heilpern, *Fuzzy mappings and fixed point theorems*, J. Math. Anal. Appl. 83 (1981) 566–569.
- [59] M. Hojati, C.R. Bector, K. Smimou, *A simple method for computation of fuzzy linear regression*, Eur J Oper Res, 166 (2005) 172–184.
- [60] D.H. Hong, C. Hwang, *Support vector fuzzy regression machines*, Fuzzy Set and Syst 138 (2002) 271–281.
- [61] X.Q. Hu, *Common coupled fixed point theorems for contractive mappings in fuzzy metric spaces*, Fixed Point Theory and Applications, Volume 2011, Article ID 363716.
- [62] H.L. Huang, F.G. Shi, *L-fuzzy numbers and their properties*, Information Sciences 178 (2008) 1141–1151.
- [63] H. Huang, C. Wu, *On the triangle inequalities in fuzzy metric spaces*, Information Sciences 177 (4) (2007) 1063–1072.
- [64] C.F. Juang, C.D. Hsieh, *TS-fuzzy system-based support vector regression*, Fuzzy Set and Syst 160 (2009) 2486–2504.
- [65] O. Kaleva, *On the convergence of fuzzy sets*. Fuzzy Sets and Systems 17 (1985) 53–65.
- [66] O. Kaleva, *A comment on the completion of fuzzy metric spaces*. Fuzzy Sets and Systems 159 (2008) 2190–2192.
- [67] O. Kaleva, S. Seikkala, *On fuzzy metric spaces*, Fuzzy Sets and Systems 12 (1984) 215–229.
- [68] C. Kao, C.L. Chyu, *Least-squares estimates in fuzzy regression analysis*, Eur J Oper Res 148 (2003) 426–435.
- [69] E. Karapinar, N.V. Luong, *Quadruple fixed point theorems for nonlinear contractions*, Computers and Mathematics with Applications, DOI: 10.1016/j.camwa.2012.02.061.
- [70] R. Khemchandani, Jayadeva, S. Chandra, *Regularized least squares fuzzy support vector regression for financial time series forecasting*, Expert Syst with Appl 36 (2009) 132–138.
- [71] B. Kim, R.R. Bishu, *Evaluation of fuzzy linear regression models by comparing membership functions*, Fuzzy Sets and Systems, 100 (1998) 343–352.
- [72] Y.C. Kim, K.C. Min, *L-fuzzy preproximities and L-fuzzy topologies*, Information Sciences 173 (2005) 93–113.
- [73] W. Klingenberg, *Riemannian Geometry*, Walter de Gruyter (1982).
- [74] I. Kramosil, J. Michálek, *Fuzzy metrics and statistical metric spaces*, Kybernetika 11 (1975) 336–344.

- [75] V. Krättschmer, *Limit distribution of least squares estimators in linear regression models with vague concepts*, J Multivariate Anal 97 (2006) 1044–1069.
- [76] V. Lakshmikantham, L. Ćirić, *Coupled fixed point theorems for nonlinear contractions in partially ordered metric spaces*, Nonlinear Anal. TMA 70 (2009), 4341–4349.
- [77] R. Lowen, *Approach spaces: a common supercategory of TOP and MET*, Math. Nachr. 141 (1989) 183–226.
- [78] R. Lowen, *Approach Spaces: the missing link in the topology-uniformity-metric triad*, Oxford Mathematical Monographs, Oxford University Press, 1997.
- [79] R. Lowen, M. Sioen, D. Vaughan, *Completing quasi-metric spaces - An alternative approach*, J. Math. 29 (2003) 113–136.
- [80] J. Martínez-Moreno, A. Roldán, C. Roldán, *A note on the \mathcal{L} -fuzzy Banach's contraction principle*, Chaos, Solitons and Fractals 41 (5) (2009) 2399–2400.
- [81] J. Martínez-Moreno, A. Roldán, C. Roldán, *KM-Fuzzy approach space*, Proceedings of the International Fuzzy Systems Association World Conference (2009) 1702–1705.
- [82] K. Menger, *Statistical metrics*, Proc National Acad Sci of the United States of America 28 (1942) 535–537.
- [83] K. Menger, *Ensembles flous et fonctions aléatoires*, C.R. Paris 232 (1951) 2001–2003.
- [84] J. Michálek, *Fuzzy topologies*, Kybernetika 11 (5) (1975) 345–354.
- [85] D. Mihet, *A Banach contraction theorem in fuzzy metric spaces*, Fuzzy Sets and Systems 144 (3) (2004) 431–439.
- [86] D. Mihet, *On fuzzy contractive mappings in fuzzy metric spaces*, Fuzzy Sets and Systems 158 (2007) 915–921.
- [87] M. Mizumoto, J. Tanaka, *Some properties of fuzzy numbers*, in: M.M. Gupta et al. (Eds.), *Advances in Fuzzy Set Theory and Applications* (1979) 153–164, North-Holland, New York.
- [88] S. Montiel, A. Ros, *Curvas y Supercies*, Proyecto Sur Ediciones, Granada, 1997.
- [89] N.N. Morsi, *On fuzzy pseudo-normed vector spaces*, Fuzzy Sets and Systems 27 (1988) 351–372.
- [90] S.A. Naimpally, B.D. Warrack, *Proximity Spaces*, Cambridge at the University Press, 1970.
- [91] M.M. Nasrabadi, E. Nasrabadi, *A mathematical-programming approach to fuzzy linear regression analysis*, Appl Math Comput, 155 (2004) 873–881.

-
- [92] W. Näther, *Regression with fuzzy random data*, *Comp Stat Data Anal* 51 (2006) 235–252.
- [93] E. Nishiura, *Constructive methods in probabilistic metric spaces*, *Fundamenta Mathematicae* 67 (1970) 115–124.
- [94] A.M. Palacios, L. Sánchez, I. Couso, *Diagnosis of dyslexia with low quality data with genetic fuzzy systems*, *International Journal of Approximate Reasoning* 51 (2010) 993–1009.
- [95] E. Pap, *Pseudo-analysis and nonlinear equations*, *Soft Computing* 6 (2002) 21–32.
- [96] E. Pap, O. Hadžić, R. Mesiar, *A fixed point theorem in probabilistic metric spaces and an application*, *J. Math. Anal. Appl.* 202 (1996) 433–449.
- [97] J.H. Park, *Intuitionistic fuzzy metric spaces*, *Chaos, Solitons and Fractals* 22 (2004) 1039–1046.
- [98] M.L. Puri, D.A. Ralescu, *Fuzzy random variables*, *J Math Anal Appl* 114 (1986) 409–422.
- [99] V. Radu, *Some suitable metrics on fuzzy metric spaces*, *Fixed Point Theory* 5 (2) (2004) 323–347.
- [100] S. Ramanna, J.F. Peters, *Approach space framework for image database classification*, *Integrated Computing Technology*, 165 (2011) 75–89.
- [101] J. Rodríguez-López, S. Romaguera, *The Hausdorff fuzzy metric on compact sets*, *Fuzzy Sets and Systems* 147 (2004) 273–283.
- [102] A. Roldán, J. Martínez-Moreno, C. Roldán, *On interrelationships between fuzzy metric structures*, en proceso de revisión menor en *Iranian Journal of Fuzzy Systems*.
- [103] C. Roldán, A. Roldán, J. Martínez-Moreno, *A fuzzy regression model based on distances and random variables with crisp input and fuzzy output data: a case study in biomass production*, *Soft Computing* 16 (5) (2012) 785–795.
- [104] J.F. Saeman, *Kinetics of wood saccharification. Hydrolysis of cellulose and decomposition of sugars in dilute acid at high temperature*, *Industrial and Engineering Chemistry* 37 (1945) 43–52.
- [105] R. Saadati, *Notes to the paper “fixed points in intuitionistic fuzzy metric spaces” and its generalization to \mathfrak{L} -fuzzy metric spaces*, *Chaos Solitons and Fractals* 35 (2008) 80–176.
- [106] R. Saadati, A. Razani, H. Adibi, *A common fixed point theorem in \mathcal{L} -fuzzy metric spaces*, *Chaos, Solitons and Fractals* 33 (2007) 358–363.
- [107] R. Saadati, J.H. Park, *On the intuitionistic fuzzy topological spaces*, *Chaos, Solitons and Fractals* 27 (2006) 331–44.

- [108] R. Saadati, S. Sedghi, N. Shobe, *Modified intuitionistic fuzzy metric spaces and some fixed point theorems*, Chaos, Solitons and Fractals 38 (2008) 36–47.
- [109] R. Saadati, S. Vaezpour, *Some results on fuzzy Banach spaces*, Journal of Applied Mathematics and Computing Volume 17 (2005) 475–484.
- [110] L. Sánchez, I. Couso, J. Casillas, *Genetic learning of fuzzy rules based on low quality data*, Fuzzy Sets and Systems 160 (2009) 2524–2552.
- [111] B. Schweizer, A. Sklar, *Probabilistic Metric Spaces*, Dover Publications, New York, 2005.
- [112] S. Sedghi, I. Altun, N. Shobe, *Coupled fixed point theorems for contractions in fuzzy metric spaces*, Nonlinear Anal. TMA 72 (3-4) (2010) 1298–1304.
- [113] R. Seising, *Fuzziness before fuzzy sets: Two 20th century philosophical approaches to vagueness – Ludwik Fleck and Karl Menger*, Fuzzy Logic, Soft Computing and Computational Intelligence, Eleventh IFSA World Congress, Beijing, China (2005) 1499–1504.
- [114] M. Serrunier, H. Prade, *Imprecise regression and regression on fuzzy data. A preliminary discussion*, IEEE International Conference on Fuzzy Systems, pp 1506–1511, Vancouver, Canada, 2006.
- [115] S. Shakeri, L.J.B. Ćirić, R. Saadati, *Common Fixed Point Theorem in Partially Ordered L-Fuzzy Metric Spaces*, Fixed Point Theory and Applications Volume 2010, Article ID 125082.
- [116] F.G. Shi, *(L, M)-Fuzzy metric spaces*, Indian Journal of Mathematics, 52 (2) (2010) 231–250.
- [117] F.G. Shi, *Regularity and normality of (L, M)-Fuzzy topological spaces*, Fuzzy Sets and Systems 182 (2011) 37–52.
- [118] W. Sintunavarat, Y.J. Cho, P. Kumam, *Coupled coincidence point theorems for contractions without commutative condition in intuitionistic fuzzy normed spaces*, Fixed Point Theory and Applications 2011, 2011:81.
- [119] H. Tanaka, I. Hayashi, J. Watada, *Linear regression analysis with fuzzy model*, IEEE Trans Syst Man Cybern SMC-12 (1982) 903–907.
- [120] H. Tanaka, I. Hayashi, J. Watada, *Possibilistic linear regression analysis for fuzzy data*, Eur J Oper Res 40 (1989) 389–396.
- [121] V. Vapnik, *Statistical learning theory*, Wiley, 1998.
- [122] P.V. Veeramani, *Best approximation in fuzzy metric spaces*, J. Fuzzy Math. 9 (2001) 75–80.

-
- [123] C.X. Wu, M. Ma, *The basic of fuzzy analysis*. National Defence Industry Press, Beijing, 1991.
- [124] B. Wu, N.F. Tseng, *A new approach to fuzzy regression models with application to business cycle analysis*, Fuzzy Sets and Systems 130 (2002) 33–42.
- [125] Z. Xu, *A method based on distance measure for interval-valued intuitionistic fuzzy group decision making*, Information Sciences 180 (2010) 181–190.
- [126] Z.S. Xu, J. Chen, J. Wu, *Clustering algorithm for intuitionistic fuzzy sets*, Information Sciences 178 (2008) 3775–3790.
- [127] H. Yang, K. Huang, I. King, M.R. Lyu, *Localized support vector regression for time series prediction*, Neurocomputing 72 (2009) 2659–2669.
- [128] P. Zabrieko, M. Krasnoselskii, *Iteration of operators and fixed point*, Doklady Akad. Nauk SSSR 196 (1971) 1006–1009.
- [129] L.A. Zadeh, *Fuzzy sets*, Information and Control 8 (1965) 338–353.
- [130] D. Zhang, *A natural topology for fuzzy numbers*, Journal of Mathematical Analysis and Applications 264 (2) (2001) 344–353.
- [131] Z. Zhao, C. Wu, *The equivalence of convergences of sequences of fuzzy numbers and its applications to the characterization of compact sets*, Information Sciences 179 (2009) 3018–3025.
- [132] J.Z. Xiao, X.H. Zhu X. Jin, *Fixed point theorems for nonlinear contractions in Kaleva-Seikkala's type fuzzy metric spaces*, Fuzzy Sets and Systems, doi:10.1016/j.fss.2011.10.010.
- [133] X.H. Zhu, J.Z. Xiao, *Note on “Coupled fixed point theorems for contractions in fuzzy metric spaces”*, Nonlinear Anal. TMA 74 (16) (2011) 5475–5479.