2
1 ej

Universidad Nacional Autónoma de México

FACULTAD DE CONTADURÍA Y ADMINISTRACIÓN

TEORIA DIFUSA PARA EL ANALISIS DE DECISIONES EN CONDICIONES DE RIESGO

SEMINARIO DE INVESTIGACION INFORMATICA

QUE PARA OBTENER EL TITULO DE:

LICENCIADO EN INFORMATICA

PRESENTA:

BRUNO ALEJANDRE GONZÁLEZ

ASESOR DEL SEMINARIO:

ACT. FCO. DAVID MEJIA RODRIGUEZ

MEXICO, D.F. 270363 **1999**

Gracias a mis maestros:

A mis padres, mis hermanos, mi familia,
los maestros de la vida.

A mis profesores y educadores,
los maestros de la teoría.

A los grandes pensadores,
los maestros de la fantasía.

“Pesada carga es el conocimiento por sistema porque no da cabida al conocimiento por verdad”

Indice pág.

INTRODUCCIÓN

El campo de la toma de decisiones es inherente a la naturaleza humana, así como la aceptación de la realización de actividades bajo condiciones de riesgo, pero esta segunda característica sí es restrictiva del ser humano porque ningún animal toma riesgos (todo animal es capaz de enfrentar peligros e inconscientemente ponerse él mismo bajo peligro, pero nunca acepta un riesgo), debido a que el riesgo es una construcción social ineludible que es aceptada conscientemente.

Para la toma de decisiones en condiciones de riesgo existen numerosas herramientas tradicionales (basadas en lógica bivalente para la toma de decisiones) y algunas que adoptan nuevas tecnologías, dentro de éstas últimas se encuentra la Teoría de Conjuntos Difusos (Fuzzy Sets Theory) que tiene por objetivo el presentar un análisis del comportamiento de los conjuntos de una manera más cercana a la realidad y sobre todo al pensamiento humano, el cual por definición no es bivalente ni rígido sino más bien flexible y con una serie de matices de valores.

En el presente trabajo se realiza una evaluación de dos métodos: el Método Monte Carlo, basado en la aplicación de métodos tradicionales de toma de decisiones en condiciones de riesgo y el Método de Análisis Difuso, basado en las aseveraciones que postula la Teoría de Conjuntos Difusos, ambos son descritos y comparados sobre los mismos rubros, arrojando de esta manera el resultado de esta evaluación como conclusión de la Tesis.

Hipótesis

Hipótesis. La teoría de Conjuntos Difusos representa una herramienta eficaz de análisis para la toma de decisiones en condiciones de riesgo.

Hipótesis Alterna. La teoría difusa, por sí misma, no es una buena herramienta para el análisis y toma de decisiones en condiciones de riesgo.

Teoría Difusa para el Análisis de Decisiones en Condiciones de Riesgo

Marco Teórico

A. Teoría Difusa

I. Teoría de Conjuntos Difusos.

La Teoría de conjuntos difusos es un cuerpo de técnicas y conceptos que tratan de forma sistemática con clases de universos que no tienen fronteras definidas tajantemente, sino que sus elementos tienen un cierto grado de pertenencia a las diferentes clases, que se encuentra entre la pertenencia y la no-pertenencia (0 y 1).

La teoría difusa no es una práctica nueva, tiene más de 20 años desde su concepción en 1965 por Lofti Zadeh, profesor de la Universidad de Berkeley, California y ha tenido gran desarrollo en sus diferentes campos de aplicación, entre los cuales se encuentra inteligencia artificial, lógica, administración, sistemas expertos, investigación de operaciones, reconocimiento de patrones, ciencias de la computación, robótica y teoría de las decisiones, éste último es el campo sobre el que se desarrollará el presente trabajo.

Como se ha mencionado antes, el desarrollo científico de la teoria en los diversos campos ha sido muy prolifero, sin embargo no ha sido igual en el campo de la educación debido a que existe muy poca literatura que se de a la tarea de introducir o servir como libro de texto para el aprendizaje de la teoría (dirigido o autodidacta), dificultándose de esta manera la propagación de la teoría en la comunidad estudiantil, a este respecto H.J. Zimmermann declaró:

> "...De hecho, es extremadamente difícil para un principiante en el campo o para cualquiera que desee aplicar la teoría de conjuntos difusos a sus problemas, reconocer apropiadamente el 'estado actual del arte'..."

Antagónicamente a la dificultad de su introducción, la teoría ha tenido un desarrollo muy fuerte en áreas avanzadas con aplicaciones especializadas entre las que se encuentra la teoría de decisiones. La principal motivación para el desarrollo en esta área radica en el reconocimiento de que el pensamiento humano por naturaleza es difuso en vez de concreto (no-difuso) y es aproximado en vez de preciso, lo cual permite vincular el procesamiento de lenguaje natural con el análisis de decisiones a través del uso de esta teoría.

Actualmente las aplicaciones orientadas al análisis de decisiones en condiciones de riesgo se basan en algoritmos determinísticos, los cuales son muy complejos en su implementación debido a que se pretende adaptar el pensamiento humano (en áreas de valoración de riesgos donde habitualmente se utilizan términos imprecisos además de gran número de parámetros) a un esquema bivalente de "si" y "no", cargando estos algoritmos de estructuras decisorias.

No obstante existen algunos desarrollos (Risk Calc de RAMA, por ejemplo) que se valen de la aritmética difusa para el análisis de decisiones en condiciones de riesgos. **La aritmética difusa es una generalización del análisis de intervalos** que trata la incertidumbre de una manera mucho más comprensiva, ya que en vez de representar un número con un simple intervalo utiliza intervalos difusos que asignan una gama de niveles conservadores y optimistas al número en cuestión.

Asimismo la aritmética difusa se basa en una teoría más simple, con menos axiomas que la teoría de la probabilidad, lo que –una vez superado el paradigma tradicional para concebir los cálculos– la vuelve mucho más práctica.

En realidad la aritmética difusa es un subconjunto de la teoría de las posibilidades que maneja conjuntos difusos de números reales llamados distribuciones posibilísticas.

Esta es otra de las razones del desarrollo de la teoría difusa, ya que al proveer una base "posibilística", la cual también es más cercana al modo natural de pensamiento y procesamiento de información humano que el cálculo de probabilidades tal cual se conoce actualmente.

Por otra parte, en el transcurso de la presente década se ha complementado la teoria difusa con otras tecnologías tales como las redes neuronales y los algoritmos genéticos, todas ellas herramientas que además de poseer gran poder tienen aplicación en una amplia gama de áreas del conocimiento humano.

Ahora bien, la teoría difusa presenta algunas ventajas frente a otros métodos como el de Monte Carlo, que es de los más usados en el área de análisis de riesgo, dichas ventajas son:

1. La Teoría Difusa (TD) requiere de menos datos para el análisis, esto se puede ver en la siguiente tabla comparativa:

Métodos	Análisis de Intervalos	Aritmética Difusa	Monte Carlo
Datos Requeridos	Rangos	Rangos o Distribuciones	Distribuciones y Dependencias

2. TD es aplicable a cualquier tipo de manejo de riesgo, no importando su naturaleza u origen, mientras que la teoría tradicional, basada en el cálculo de la probabilidad, requiere de un proceso adicional de interpretación de los resultados, en el cual se involucran los sentimientos y creencias de los analistas, peor aún, de los analistas inexpertos, lo cual decrementa la confiabilidad de los análisis. Por otra parte, al evitar este hecho se cae en un nuevo problema, los métodos tradicionales como el Monte Carlo no son aplicables a otras áreas si se carece de esta parte del método.

3. Al igual que los métodos tradicionales, TD provee de una distribución que representa los resultados obtenidos de manera comprensible, lo que la vuelve superior al análisis de intervalos.

4. EL procesamiento de los cálculos de TD es fácil y rápido en aplicaciones ya implementadas.

5. Los resultados generados por TD son simples de comprender ya que pueden ser emitidos incluso en lenguaje natural.

Pero no todo es positivo para la teoría Difusa, he aquí algunos de los inconvenientes que presenta:

1. Existe gran resistencia de algunos teóricos para aceptar la efectividad de la teoría, sin embargo se puede decir que es más por desconocerla que por descartarla sobre la base de un análisis previo.

2. Hacer los cálculos de la TD a "mano" puede resultar un proceso engorroso por la asimilación de los nuevos conceptos, mientras que con los métodos tradicionales se conoce bien y se dominan los conceptos permitiendo en mayor medida el cálculo mental.

Podemos concluir esta introducción con un concepto simple: **la Teoría de Conjuntos Difusos es una teoría de conceptos gradados donde todo tiene cierta elasticidad, donde todas las cosas son objeto de un grado de pertenencia.**

II. Resumen de Propiedades, Operaciones y Tipos de Interpretación.

Para comprender cabalmente las características de la Teoría Difusa, es necesario hacer hincapié en algunas características y diferencias básicas entre los conceptos que introduce esta teoría y los conceptos que la Teoría Bivalente postula.

Primeramente, se deben conocer las premisas que a continuación se presentan:

1. **Difuso** es fundamentalmente diferente a **aleatorio**.

2. Lo **difuso** juega un papel más básico en el conocimiento humano que lo **aleatorio**, ya que la naturaleza del pensamiento humano es imprecisa y difusa.

3. Para comprender la **teoría difusa** se debe abandonar todo pensamiento anterior y asimilar radicalmente los nuevos conceptos que posibilitan el análisis de humanismos y conceptos sociales de una manera tan eficiente como actualmente se hace con sistemas precisos.

Difuso <> Vago

Vaguedad es una forma difusa que depende del campo, aplicación o contexto en que se aplique la proposición difusa, por ejemplo:

Proposición difusa: "Jill es verdaderamente alta"

- Es difusa por la imprecisión de la frase "verdaderamente alta".

- Es vaga si la aplicamos para un propósito particular como comprar la talla correcta de pantalones, para lo cual la proposición no proporciona suficiente información y además de vaga se vuelve ambigua.

- Si se aplica para comprar un collar o una bolsa de mano, tenemos información suficiente y por tanto sólo es difusa y carece de ambigüedad.

Difuso <> Aleatorio

- Difuso tiene un grado de membresía (compatibilidad, posibilidad).

- Aleatorio tiene un grado de probabilidad.

Ejemplo. Si tenemos un Vw perteneciente a Julieta, ¿Cuántos pasajeros pueden salir de él? (asumiendo que todos son de características estándar).

Sea n el número de pasajeros. Cada n tiene una posibilidad (μ_n) y una probabilidad (p_n) de salir del Vw, esto es:

	1	2	3	4	5	6	7
μ_n	0	1	1	1	0.7	0.2	0
p_n	0	0.6	0.3	0.1	0	0	0

μ_n se interpreta como el grado de facilidad con el que n pasajeros pueden meterse en el Vw. Por lo tanto μ_n=0.7 por algún criterio específico o inespecífico quiere decir que la facilidad es de 0.7. Pero realizando los cálculos probabilísticos sobre la base de un cupo máximo de 5 personas, la probabilidad de que suban 5 pasajeros en el Vw de Julieta es 0.

Si tomamos n=4 para cualquier Vw la posibilidad es de 1, pero la probabilidad es de 0.1.

Conclusiones:

- La posibilidad no se presenta como una propiedad de "todo o nada", aún más allá de los límites establecidos.

- Los grados de posibilidad no son los mismos que los de probabilidad.

- La información posibilística es más elemental y menos dependiente del contexto que la información probabilística.

- El pensamiento humano es, por mucho, más posibilístico que probabilístico.

III. Ramas de la teoría difusa.

1. Matemática o no-difusa, que elabora aserciones o aseveraciones. Se caracteriza por su fundamento en matemáticas tradicionales y un desarrollo rápido de literatura sobre tipología de espacios difusos, funciones de intercambio difuso, ordenamientos difusos, aplicación en el análisis de sistemas (es la base de los métodos explicados en este trabajo).

2. Difusa, interrelaciona reglas de manipulación y aseveraciones. Tiene su origen en el llamado aproximamiento lingüístico que con el tiempo se convirtió en la lógica difusa. En esta lógica se permite que tanto las reglas de inferencia como los valores de verdad sean imprecisos y que las aseveraciones sobre conjuntos difusos no pueden ser probadas por lógica bivalente (la interpretación lingüística es parte de él).

   Esta rama no está muy desarrollada pero puede ser la base para aproximaciones o razonamiento difuso, cuyo comportamiento es muy similar al pensamiento humano y la habilidad de éste para obtener conocimiento difuso en un ambiente parcialmente desconocido.

IV. Notación, Terminología y Operación básica.

1. Terminología.

- Un conjunto difuso se asume inmerso en otro conjunto no difuso o *Universo de Discurso* que posee una colección de objetos, conceptos o construcciones matemáticas. Ejemplo:

$$U=\{x|x\in R\}$$
$$U=\{y|y \text{ vive en el DF}\}$$

El Universo de discurso se denota por U, V o W.

- El conjunto difuso (denotado por A, B, C,..., H) está caracterizado por la **función de membresía** μ_A: $U\rightarrow[0,1]$ que asocia a cada elemento $u \in U$ con un número μ_A (u) en el intervalo [0,1], esto es, un punto en un conjunto parcialmente ordenado, con $\mu_A(u)$ representando el grado de compatibilidad de u en A.

- **Soporte** es el número de puntos de U donde $\mu_A(u)$ es positivo.

- El **punto de cruce** de A es el elemento en U cuyo grado de membresía es 0.5.

- **Peso** es el valor máximo de membresía $\mu_A(u)$ de los elementos

- A es **Normal** si su *peso* es igual a la unidad (1) y subnormal si no lo es.

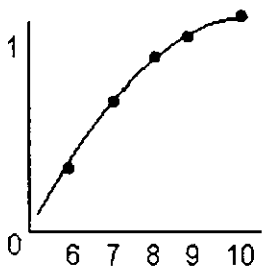
2. Función Membresía.

La Función Membresía es la función que define el grado de membresía de un elemento u de U en A, se denota como $\mu_A(u)$. Este grado de membresía puede ser asignado de manera arbitraria basados en la experiencia que se posee y de los diferentes ambientes sobre los que se desenvuelva el conjunto difuso.

Las funciones de membresía se representan, normalmente, en un rango de 0 a 1 (sobre el eje de las y's), y los elementos del conjunto se enumeran ordenados (sobre el eje de las x's). La gráfica de la función membresía está formada por el conjunto de puntos donde se intersectan los elementos con su respectivo grado de membresía, tenemos, por ejemplo:

$$A = \{.4/6, .6/7, .8/8, .9/9, 1/10\}$$

el cual gráficamente se representa como:



La función membresia ha tomado la forma de una función cuadrática, sin embargo esta forma no es restrictiva para las funciones membresía en general, ya que puede tomar formas parabólicas, sigmoidales, etc. dependiendo de la manera en que se defina la función de pertenencia para cada conjunto difuso.

Ahora, la función membresía es diferente de la **función de compatibilidad**, ya que esta última es un producto de la primera, es decir, cuando se le da una interpretación a la función de membresía se dice que se ha convertido en una función de compatibilidad (recordar el ejemplo de Jill).

También es importante resaltar el hecho de que la función membresía se basa en las observaciones y experiencia y que por tanto es totalmente arbitraria y subjetiva, en los valores asignados. Esta característica de la función nos da la flexibilidad de realizarle ajustes dependiendo de los casos de que se trate, dotando a la función membresía de gran potencia para representar diferentes situaciones ligadas con los mismos elementos de un conjunto.

3. Operaciones.

El conjunto no difuso finito $U=\{u_1, u_2, ... u_n\}$ se expresa como:

$U= u_1+ u_2+. + u_n$ (entendiendo "+" como unión por lo que $u_i+ u_i = u_i$)

El cual también puede representarse como:

$$U= \sum_{i=1}^{n} u_i$$

Como extensión del conjunto difuso finito A de U se expresa como:

$$A= \mu_1 u_1+ \mu_2 u_2+ ....+ \mu_n u_n$$

O como:

$$A= \sum_{i=1}^{n} \mu_i u_i$$

donde μ_i es el grado de membresía y para evitar ambigüedades cuando u es un número se les separa con /:

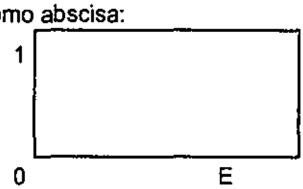
$$A= \sum_{i=1}^{n} \mu i/ui$$

Operación Complemento

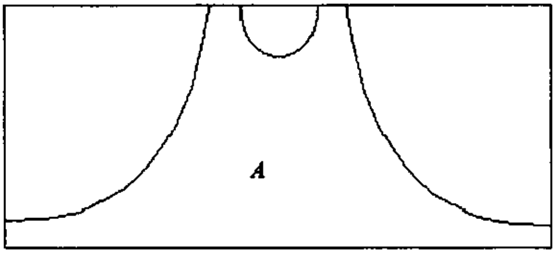
Si nosotros planteamos un conjunto borroso a partir de una característica, por ejemplo el conjunto de personas que integran una clase de pintura y que dominan la técnica de dibujo al desnudo, pudiendo deducir que el complemento de ese conjunto puede ser definido a partir de la negación de esa característica, siendo así el conjunto formado por todas aquellas personas de la clase de pintura que no dominan la técnica de dibujo al desnudo. Entonces cuanto más cerca se esté de dominar la técnica de dibujo al desnudo más lejos se está de no dominarla, conforme avanza la clase de pintura las personas que no dominan la técnica al desnudo se acercan a dominarla en la misma medida que se alejan del conjunto de los que no la dominan. Si nosotros hemos medido la pertenencia de un elemento a un conjunto por su asociación a un elemento del conjunto de membresía, es decir un número entero entre cero y uno en nuestro caso, resulta natural suponer que el complemento del conjunto para ese elemento sea lo que le falta para ser igual a 1, esto es:

$$\tilde{a} \leftarrow\rightarrow \mu_{\tilde{a}}(x) = 1\text{-}\mu_A(x)\ \forall\ x \in E$$

Representación gráfica: Consideremos un rectángulo con los valores de las funciones de membresía (conjunto de membresía) como ordenada y con el conjunto de referencia como abscisa:



En este rectángulo podemos graficar un conjunto borroso representando las parejas ordenadas que forman el conjunto como puntos en la gráfica.



Se suele sombrear la parte cubierta por el conjunto borroso con objeto de obtener mayor claridad en la representación.

E

$\bar{E}$

Operación Unión

La unión de dos conjuntos borrosos se puede conceptuar de modo similar a la unión de conjuntos ordinarios si definimos el conjunto resultante, (conjunto unión) como el conjunto más pequeño que contenga a ambos conjuntos.

Formalizando:

dado un conjunto C tal que: $A \subset C$ y $B \subset C$

entonces si $C \subset A \cup B \rightarrow C = A \cup B$

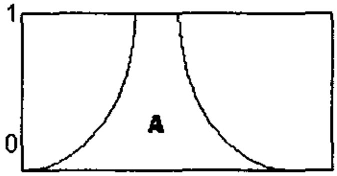
Introduciendo aquí la notación "V" como operación de máximo tendríamos:

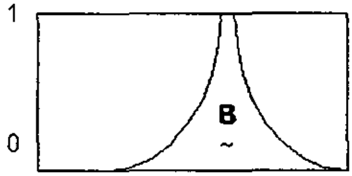
$$\mu A \cup B (x) = \mu A(x) \vee \mu B(x) \quad \forall x \in E$$

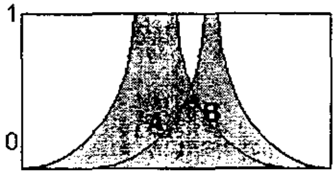
y

$$\mu \bigcup_{i=1}^{n} A_i (x) = \bigvee_{i=1}^{n} \mu A_i (x) \quad \forall x \in E$$

Gráficamente:







$A \cup B$

Operación Intersección

Siguiendo el mismo razonamiento podemos definir al conjunto de intersección de dos conjuntos difusos como: "el mayor conjunto contenido en ambos conjuntos".

De manera formal se expresa:

dados los conjuntos A, B, y C tales que

$$C \subseteq A \quad y \quad C \subseteq B$$

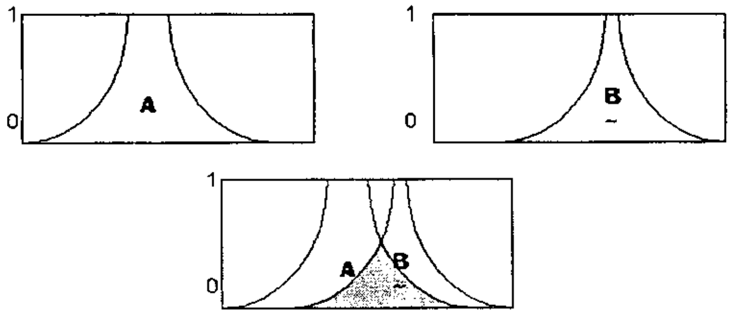
entonces si $A \cap B \subset C$

En términos de la función de membresía e introduciendo la notación "$\wedge$" como operador de mínimos:

$$\mu_{A \cap B}(x) = \mu A(x) \wedge \mu B(x) \quad \forall\, x \in E$$

Simplificando:

$$\mu_{\cap_{i=1}^{n} A_i}(x) = \wedge_{i=1}^{n}\, \mu A_i(x) \quad \forall\, x \in E$$



$A \cup B$

Operación Producto Directo

El producto directo de dos conjuntos es definido como sigue:

$$A * B \leftrightarrow \mu A * B (x) = \mu A (x) * \mu B (x) \quad \forall x \in E$$

Donde "*" significa el producto en los números reales.

Operación Suma Directa

La definimos como:

$$A + B \leftrightarrow \mu A + B (x) = \mu A (x) + \mu B (x) \quad \forall x \in E$$

Siendo "+" la adición en los reales. Debemos hacer notar que esta operación sólo tiene significado cuando el resultado no excede de 1 para toda x en E. Por lo cual es útil definir lo siguiente.

Operación Suma Acotada

Definida de modo que el resultado nunca exceda de 1:

$$A \overset{-}{+} B = E \wedge (A + B)$$

$$A \overset{\cdot}{+} B \leftrightarrow \mu A \overset{\cdot}{+} B (x) = \mathrm{Min}\,[\,1, \mu A + B (x)] \quad \forall x \in E$$

Operación Suma Algebraica

La suma algebraica de dos conjuntos borrosos la podemos definir como sigue:

$$A + B \leftrightarrow \mu A + B (x) = \mu A (x) + \mu B (x) - \mu A (x) * \mu B (x) \quad \forall x \in E$$

4. Propiedades

Si A, B y C denotan conjuntos borrosos en un conjunto E entonces:

Conmutativa

- $A \cap B = B \cap A$

- $A \cup B = B \cup A$

Asociativa

- $(A \cap B) \cap C = A \cap (B \cap C)$

- $(A \cup B)\ C = A\ (B \cup C)$

Idempotencia

- $A \cap A = A$
- $A \cup A = A$

Distributiva

- $A \cap (B \cup C) = (A \cap B) \cup (A \cap C)$

- $A \cup (B \cap C) = (A \cup B) \cap (A \cup C)$

Donde θ es el conjunto ordinario tal que $\mu\,\theta\,(x) = 0 \quad \forall\, x \in E$

- $A \cap \theta = \theta$

- $A \cup \theta = A$

Cuando E es el conjunto ordinario tal que $\mu\, E\,(x) = 1 \quad \forall\, x \in E$

- $A \cap E = A$

- $A \cup E = E$

Involución

- $(A) = A$

Leyes de Morgan

- $\overline{A \cap B} = \overline{A} \cup \overline{B}$

- $\overline{A \cup B} = \overline{A} \cap \overline{B}$

5. Operaciones Adicionales.

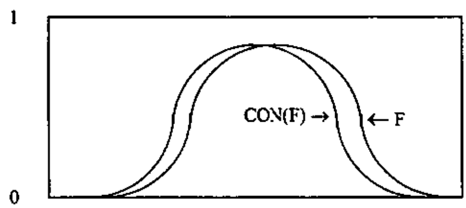
Operación Concentración

CON(A)={a(x) X a(x)/x | x es un elemento de U}
={$a(x)^2$/x | x es un elemento de U}

Esta operación decrementa el grado de membresía para los elementos, en el sentido de que si x $\in(0, 1)$, $x^2 < x$, además de que se da un mayor decremento a valores de membresía bajos, es decir con:

$$\frac{a(x)-a^2(x)}{a(x)} = 1- a(x)$$

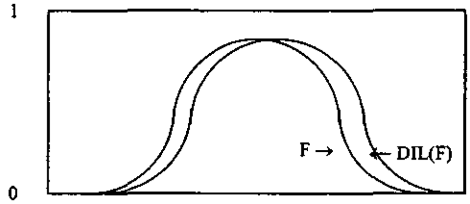
Se tendría para un valor de membresía de .9, un decremento de .1 y para un valor de membresía de .4 un decremento de .6, por lo cual su representación gráfica puede verse como:



Operación Dilatación

Es la operación opuesta a la concentración, y tiene las propiedades contrarias, de tal manera que:

A=CON(DIL(A))=DIL(CON(A))



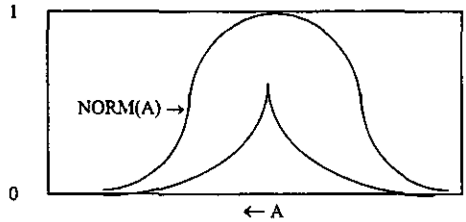
Operación Normalización

$$NORM(A)=\{(a(x)/\mu)/x \mid x \text{ es un elemento de } U\}$$

$$\text{donde } \mu = \max_{x \in U}\{ a(x)\}$$

La Normalización nos permite, en cierto sentido, reducir todos los conjuntos difusos a la misma base. Esto se hace en el mismo sentido en que los vectores de álgebra lineal se normalizan a vectores unarios. Por ejemplo, los vectores bidimensionales (1, 2) y (4,8) se pueden normalizar como $(1/\sqrt{5}, 2/\sqrt{5})$, lo que nos permite demostrar que ambos vectores tienen la misma dirección.

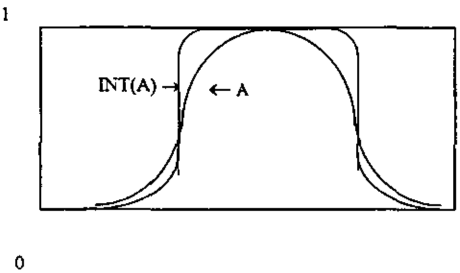
El propósito de la normalización es el hacer que el conjunto difuso sea normal, es decir (como se vio anteriormente) que al menos uno de sus elementos tenga grado de membresía 1. Esto lo logra dividiendo los grados de membresía entre el máximo grado de membresía que exista haciendo que por lo menos un elemento, el que posea el máximo grado de membresía, alcance el valor de la unidad, al mismo tiempo que se incrementa el de todo el conjunto. Si ya existe alguno que lo tenga, la división no afecta en lo más mínimo al conjunto (por la división entre el neutro).



Operación Intensificación

INT(A)={μ(x)/x | x es un elemento}

$$\mu(x)=\begin{cases} 2a^2(x) & \text{para } 0\leq a(x)\leq 0.5 \\ 1-2[1-a(x)]^2 & \text{para } 0.5\leq a(x)\leq 1.0 \end{cases}$$



La intensificación funciona como una especie de combinación de concentración y dilatación, incrementando la membresia de algunos elementos, disminuyendo algunos otros y modificando la forma de la curva de membresía. Los elementos con un grado de membresía mayor a .5 sufren un incremento y los que están por debajo un decremento. La intensificación acentúa el contraste entre los elementos que se encuentran arriba o por debajo del punto de cruce.

Operación Difusión

Se define como la expansión del soporte de un conjunto difuso con otros elementos que también poseen un grado de membresía diferente de cero.
La operación se define como:

$$SF = \bigcup_{f} \{a(x)*K(x)\}$$

Donde K mapea a elementos de U dentro de subconjuntos difusos, * es la multiplicación de un conjunto difuso por otro conjunto difuso que le proveerá de los elementos que expanden su soporte y $\bigcup_{f}$ es la unión difusa.

B. Análisis de Decisiones en Condiciones de Riesgo

I. Antecedentes.

En el valle del Tigris – Eufrates alrededor de 3200 a.C. vivió un grupo llamado los Ashipu. Una de sus funciones primarias era la de servir como consultores para las decisiones riesgosas, inciertas o difíciles.

Si era necesario tomar una decisión sobre riesgos venideros, arreglos de proposiciones matrimoniales o el sitio dónde construir, se podía consultar a un miembro de esta tribu. Los Ashipu, primero identificaban las dimensiones del problema, después identificaban alternativas de acción y por último reunían información sobre los posibles resultados de cada alternativa. Los mejores datos para sus recomendaciones eran las señales de los dioses, para cada alternativa favorable apuntaban un + y para cada desfavorable un -. Tras el análisis de los resultados los Ashipu eran capaces de recomendar la alternativa más favorable. El último paso era plasmar un reporte en una tableta de barro grabada.

Esta práctica de los Ashipu se considera el primer antecedente de una forma simple de análisis de riesgos. Este tipo de similitudes entre los métodos actuales y los antiguos demuestran que el ser humano se ha enfrentado desde hace mucho tiempo a los mismos problemas relacionados con riesgo.

A diferencia de los modernos analistas de riesgos quienes expresan sus resultados en términos de probabilidades matemáticas e intervalos de confianza, los antiguos Ashipu expresaban sus resultados con certidumbre, confianza y autoridad, pero ellos contaban con la ayuda de los dioses. Así pues, es necesario revisar la historia de diversas corrientes de análisis de riesgos:

Análisis cuantitativo y probabilístico.

Esta corriente de análisis tiene sus fundamentos en cuestiones religiosas como la probabilidad de una vida en el más allá. Iniciando con el Phaedo de Platón en el siglo IV a.C. se han escrito numerosos tratados que discuten el riesgo de la suerte del alma en el más allá basándose en las repercusiones de la conducta que se observa en la vida presente.

Uno de los análisis más sofisticados, para su tiempo (el siglo IV d.C.) fue realizado por Arnobius "El Grande", sacerdote pagano de una iglesia del norte de Africa. Arnobius criticaba duramente la austeridad a la que el Cristianismo se sometía hasta que tuvo una visión que le llevó a pretender la conversión al cristianismo. Ante la incredulidad del obispo cristiano, Arnobius escribió una obra en ocho tomos titulada "Contra los Paganos", en la que defendía al cristianismo frente a la iglesia pagana y la cual culminaba con una matriz de dos por dos en la que se hacía el siguiente análisis:

	Dios No Existe	Dios Existe
Ser Pagano	No hay consecuencias	Perdición del Alma
Ser Cristiano	No hay consecuencias	Vida Eterna

De lo cual se deducía que ser cristiano resultaba mucho más conveniente para el alma ya que no se tenían consecuencias negativas en ningún caso mientras que el paganismo frente a la probable existencia de Dios podía condenar su alma en el más allá.

Este es el antecedente más remoto que se tiene de un análisis de este tipo y que además se piensa sentó las bases para el Principio de Dominación, que es un método heurístico para tomar decisiones en condiciones de riesgo e incertidumbre.

Muchos trabajos se desarrollaron pero fue hasta 1657, con la introducción de la teoría de la probabilidad por Pascal, cuando se dio una visión más cierta acerca de la esta disciplina y una de sus primeras aplicaciones fue la de extender la matriz propuesta por Arnobius dándole una distribución probabilística a la existencia de Dios. El resultado final fue que para el alma tenía mucho mayor valor el ser Cristiano que Ateo.

Posteriormente al trabajo realizado por Pascal, se dio un gran incremento en la actividad intelectual alrededor de la teoría de la probabilidad, teniendo diversas aplicaciones desde el análisis de la predicción de eventos(Arbuthnot 1692), tablas de expectativas de vida (Halley 1693), hasta el diseño de un prototipo moderno de evaluación de riesgos (LaPlace 1792), en el cual se analizaba la probabilidad de muerte con y sin vacunación contra viruela.

Métodos para establecer la Causalidad.

El análisis de riesgos moderno tiene sus raíces en las Teoría Matemática de la Probabilidad y en los métodos científicos de vinculación Causa – Efecto en distintos campos de actividades peligrosas. Para establecer estas causas los científicos se han basado principalmente en métodos de observación y en el universalmente practicado método de prueba y error.

Con el avance de la ciencia se fueron estableciendo mejores métodos como el de la observación indirecta (reacción de sustancias adulteradas al fuego), métodos de observación epidemiológica (vinculación de los ecosistemas pantanosos con la malaria, sin conocer la causa específica).

A pesar de estos estudios el avance en el establecimiento de vínculos causales entre efectos adversos y actividades riesgosas ha sido lento, debido principalmente a dos razones:

1. La generación lenta de modelos químicos, biológicos y físicos especialmente antes del siglo XVII, adicionalmente a la carencia de instrumentos y métodos rigurosos de observación y experimentación. Si se analiza detenidamente podremos observar que los grandes avances científicos son relativamente recientes.

2. El arraigamiento de creencias populares que atribuyen a los fenómenos causas religiosas, sociales o mágicas, de tal manera que ciertas enfermedades se presentan como castigo divino perdiendo la objetividad de las causas reales.

Manejo Social del Riesgo

Socialmente se han establecido, históricamente, gran número de técnicas para reducir o mitigar los efectos adversos vinculados por el riesgo en áreas bien definidas. Algunas de estas medidas se muestran a continuación:

a. Reducir o eliminar el riesgo a través de prohibiciones.
b. Regular o modificar las actividades para reducir la magnitud y/o frecuencia de efectos adversos.
c. Reducir la vulnerabilidad de personas o propiedades expuestas.
d. Diseñar e implementar mecanismos posteriores de recuperación o restauración de daños.
e. Instituir reembolso y esquemas de distribución de pérdidas.

Ahora, es importante señalar que las consideraciones que se deben tomar en cuenta actualmente son sustancialmente diferentes de las que se tomaban en los primeros tiempos, por el natural cambio que sufre la sociedad. Principalmente se tienen nueve cambios importantes que se deben contemplar:

1. Existen cambios significativos en la naturaleza de los riesgos que una persona puede padecer (de la viruela al évola).
2. Existe un gran incremento en la expectativa de vida.
3. Existen nuevos riesgos fundamentalmente mayores tanto en magnitud como en carácter, que los que se encontraban en el pasado.
4. Hay una mayor habilidad de los científicos para identificar y contener el riesgo.
5. Existe un incremento en el número de científicos y analistas que trabajan en áreas específicas de salud, seguridad y riesgos ambientales.
6. El número de métodos formales de análisis cuantitativo de riesgos que se han producido y usado se ha incrementado.

7. El gobierno también se ha inmiscuido con más fuerza en evaluación y manejo de riesgos.
8. Se ha dado un incremento en la participación de grupos especiales de interés en el proceso social de manejo de riesgos.
9. Ha habido un incremento en el interés y participación del público en general, e incluso una mayor demanda de garantías.

II. Análisis de Riesgos con Métodos Probabilísticos

La Teoría de Decisiones es el enfoque estadístico que toma decisiones en condiciones de incertidumbre y riesgo. Tiene sus orígenes, en el ámbito teórico, a mediados del presente siglo, pero en cuestiones prácticas ha tenido mayor auge en los últimos años, sobre todo en aplicaciones comerciales.

Partiendo del hecho de que todos los métodos de análisis que se presentan en este trabajo tienen una fuerte y directa dependencia de conceptos estadísticos, en el siguiente punto se incluyen los conceptos básicos necesarios para su mejor comprensión.

1. Teoría de la Probabilidad.

a) Descripciones Estadísticas.

La Estadística se puede dividir en **Descriptiva e Inferencial**, la primera sólo se encarga de la recopilación de datos, su presentación tabulada y la generación de medidas descriptivas llamadas medidas de posición y medidas de variación; la segunda realiza análisis más profundos y es capaz de formular estimaciones o enunciar generalizaciones a poblaciones (Universo) a partir de muestras analizadas.

Por **población** (representada por N) debemos entender el conjunto de datos que consta de todas las observaciones hipotéticamente posibles, mientras que por **muestra** (denotada como n) entenderemos una parte o selección de observaciones que se hace a partir de la población.

La diferenciación entre estos dos conceptos es básica para efectos de análisis de riesgo, ya que a partir de una muestra sobre una población de interés, se realizan generalizaciones en torno a las cuales se llevará a cabo la toma de decisiones sobre la población completa.

La necesidad del muestreo se presenta con el hecho de la limitación de recursos, es decir, los recursos con que se cuentan para hacer un estudio (humanos, económicos, temporales, etc.) son finitos y en muchos casos insuficientes, ello nos mueve a contraer el alcance de las observaciones y a partir de éstas generalizar a toda la población, como se ha mencionado antes, por ello es necesario contar con un buen método de análisis que nos proporcione la confiabilidad de extender nuestras aseveraciones a toda la población y esperar que correspondan con un mínimo de desviación.

Un primer paso hacia la generalización es la descripción a través de **medidas de posición**, las cuales se explican a continuación.

b) Medidas de Tendencia Central.

La medida de tendencia central más difundida e implementada es la **media** (coloquialmente conocida como promedio) la cual se define como la suma de un conjunto de valores entre el número de elementos de ese conjunto.

Si hablamos de la **media de una población**, tomando en cuenta la definición de población que se ha establecido en el punto anterior, tendríamos que la media poblacional se representa como:

$$\mu = \frac{\Sigma x}{N}$$

ahora, en el caso de la **media de una muestra** y tomando en cuenta la definición que tenemos de muestra, la media muestral se representará como:

$$\bar{x} = \frac{\Sigma x}{n}$$

donde n es el tamaño de la muestra y x cada uno de los elementos seleccionados para dicha muestra

La media posee las siguientes características:

- Existencia persistente, ya que se puede calcular en base a cualquier conjunto de datos numéricos.
- Unicidad, la media siempre es una y es única para cada conjunto de datos.

- Profundidad de análisis, en el sentido de que se puede calcular la media de las medias de varios conjuntos de datos y obtener la media integrada.
- Confiabilidad, porque la media de varias muestras de una misma población presentan una fluctuación mínima.

Por otra parte, la media es susceptible a sufrir desviaciones del centro por acción de los **valores extremos**, en el caso de que se separen demasiado de los demás valores. A los valores extremos que se apartan demasiado del cuerpo principal de los datos se les llama **valores aberrantes**.

Contra la acción de estos valores aberrantes, se tiene otra medida de tendencia central conocida como la **mediana**, la cual se puede definir como el elemento que ocupa la posición central de un conjunto de datos que se ha ordenado previamente (ascendente o descendente).

Para localizar la posición en la cual se encuentra la mediana podemos valernos de la siguiente fórmula:

$$\text{mediana} = \frac{n+1}{2}$$

Esta fórmula sirve con un resultado exacto para casos en que n es impar, ya que nos indica la posición de uno de los elementos del conjunto, sin embargo, en casos en que n es par, se debe interpretar como el punto medio entre el elemento ubicado en la posición especificada en la parte entera del resultado y el siguiente elemento, o lo que es lo mismo, como la media de los valores de los elementos que se encuentran al centro del conjunto (el elemento ubicado en la posición especificada en la parte entera del resultado y el siguiente elemento).

Además de estas dos medidas de tendencia central se encuentra una más que nos ayuda a tener un panorama más completo de las características del conjunto de datos analizado, la **moda**. La moda puede definirse como el valor que tiene más ocurrencias en el conjunto analizado.

Existen casos en que el conjunto que se analiza es **multimodal** lo que significa que se tienen más de una moda. En estos casos es muy probable que se deba a una mala técnica de muestreo de los datos o a un mal análisis de la naturaleza de los mismos.

La moda se calcula analizando la distribución de frecuencias de un conjunto dado e identificando al elemento que tiene mayor número de ocurrencias.

Por último, cabe mencionar que existen variantes de estas medidas que tratan de proporcionar una mayor información a la ya proporcionada por dichas medidas, por ejemplo la **media ponderada**, la **media aritmética**, y la **media armónica**.

c) Medidas de Variación.

La primer medida de variación que nos proporciona información sobre la validez de las medidas de tendencia central es el **intervalo** que se refiere al espacio de valores sobre el que se encuentran los datos. Se obtiene restando los dos valores extremos de un conjunto determinado de datos. Como puede observarse, esta medida no nos proporciona información acerca de la **dispersión** que presentan los datos que están dentro de este intervalo. Su mayor utilidad se obtiene cuando el tamaño de la muestra es muy pequeño (como en el caso de muestras múltiples con pocos elementos).

La **desviación estándar**, por otro lado, es una medida de variación que nos proporciona información acerca de la dispersión de los elementos de un conjunto con respecto a su media. Se podría, entonces, definir como la sumatoria del resultado de cada uno de los elementos menos la media, todo ello entre el número de elementos; sin embargo tendríamos resultados negativos (los que están a la izquierda de la media) y en el caso de una distribución simétrica el resultado sería cero. Por ello, es necesario eliminar la acción de los elementos y la manera más indicada de hacerlo es elevando los resultados de la resta al cuadrado:

$$\sigma^2 = \frac{\sum(x - \mu)^2}{N}$$

La cual es la definición de la **varianza** de la población, la cual constituye (se verá más adelante), una medida que es base de análisis para la toma de decisiones.

Ahora, volviendo a nuestro objetivo principal, para lograr definir la desviación estándar de la población tendriamos que aplicar la raíz cuadrada para eliminar la acción del cuadrado de la varianza:

$$\sigma = \sqrt{\frac{\Sigma(x - \mu)^2}{N}}$$

Es necesario definir estas medidas para su aplicación sobre muestras, bajo el entendido de que es el medio más usado para el análisis estadístico, en este caso se expresará de la manera siguiente:

$$s^2 = \frac{\Sigma(x - \mu)^2}{n-1}$$

$$s = \sqrt{\frac{\Sigma(x - \mu)^2}{n-1}}$$

Se emplea n – 1 porque si se usa n al calcular s^2 y después se promedian las s^2 de las distintas muestras, se encontrará que el resultado es menor, a veces de manera muy significativa, que σ^2, lo que lo convertiría en un **estimador sesgado**. Para solucionar este problema se ha demostrado que al emplear n – 1 en vez de n se llega a un resultado mucho más satisfactorio. Cabe mencionar que para muestras muy grandes el uso de n en el cálculo de s produce un sesgo mínimo, esto es, el sesgo es despreciable y el estimador de σ se vuelve insesgado.

Adicionalmente contamos con una fórmula simplificada, ya que no es necesario calcular todas las desviaciones estándar de la media, y que además nos ofrece un dato exacto no una estimación. Dicha fórmula queda definida para una muestra como sigue:

$$s = \sqrt{\frac{n(\Sigma x^2) - (\Sigma x)^2}{n(n-1)}}$$

y para una población:

$$\sigma = \sqrt{\frac{N(\Sigma x^2) - (\Sigma x)^2}{N^2}}$$

d) Teorema de Chebyshev.

Como se dijo anteriormente, si la dispersión de los datos es grande éstos se alejan de la media, si es pequeña se acercan. Lo mismo lo podemos definir en términos de desviaciones estándar, siendo ésta pequeña si sus elementos se concentran cerca de la media y grande si no.

En términos de desviaciones estándar está definido el Teorema de Chebyshev, el cual dice:

> "En relación con un conjunto de datos cualquiera (población o muestra) y una constante k cualquiera, mayor que 1, cuando menos $1-1/k^2$ de los datos deberán estar dentro de k desviaciones estándar, a uno y otro lado de la media". (Freund, John E.)

Donde k es el número de desviaciones estándar que existen en un intervalo dado; k se calcula restando el límite superior del intervalo menos el límite inferior del mismo, el resultado se divide entre la desviación estándar del conjunto. Al sustituir k en la fórmula del Teorema de Chebyshev se obtiene el porcentaje de elementos que se encuentran dentro de ese número de desviaciones estándar.

e) Posibilidades, Probabilidades y Esperanzas.

Como se explicó anteriormente (en el ejemplo del Vw de Julieta en el punto II del apartado A) la posibilidad es esencialmente diferente de la probabilidad, dado que para conocer lo que es probable primero debemos conocer lo que es posible.

En términos de posibilidades existen dos problemas básicos:

- Primero, definir todo lo que puede suceder en una situación dada.
- Segundo, enumerar todos los casos diferentes que pueden darse, aunque normalmente no es necesario elaborar una lista completa que contenga todas las posibilidades, sino sólo aquéllas de mayor importancia para efectos de lo que se pretende analizar.

Asimismo, tenemos la regla de multiplicación de opciones, que sirve para determinar el número de eventos posibles que pueden ocurrir en una situación dada. Esta regla es una generalización derivada del análisis del comportamiento de los árboles de decisión para determinar el número de posibilidades que pueden presentarse en una situación definida:

"Si una elección consta de k pasos, donde el primero puede realizarse de n_1 maneras, en cada una de estas el segundo paso puede efectuarse de n_2 maneras, ..., y, en cada una de estas el k – ésimo puede hacerse de n_k maneras, entonces la situación puede resolverse en $n_1 * n_2 * \ldots * n_k$ formas". (Freund, John E.)

Esta regla se aplica en condiciones en las que el orden en que aparecen los elementos de un conjunto es importante dentro de las distintas opciones de selección en un mismo conjunto.

A la selección de un número r de elementos de un conjunto con n elementos, donde r es menor o igual a n, con cualquier disposición u orden en que estos elementos se presentan se denomina como **permutación**, y su fórmula general se puede escribir como:

$$ {}_nP_r = \frac{n!}{(n-r)!} $$

Ahora, es importante señalar que en el caso de las permutaciones el orden en que se presentan los elementos es importante, lo que hace necesario otra definición para aquéllos problemas en los que no sea relevante el orden de los elementos, pero sí cuántos elementos se tomarán de un conjunto determinado. Esta necesidad nos lleva a la definición de la **combinación**, que es la selección de un número r de elementos de un conjunto con n elementos donde r es menor o igual a n y no importa la disposición u orden en que estos elementos se presentan. Su fórmula queda como:

$$ {}_nC_r = \frac{n!}{r!(n-r)!} $$

Hasta aquí se encuentran las bases necesarias para comprender la posibilidad, ahora deberemos vincular esta con el cálculo de las probabilidades para definir posteriormente los algoritmos que se valen de ella para el análisis de riesgos.

Probabilidades.

La asignación de probabilidades a un evento cualquiera para medir el riesgo de su ocurrencia es la práctica actualmente más utilizada, por ello es necesario describirla para poder comprender su uso en los métodos que más adelante se desarrollarán.

El concepto clásico de probabilidad se puede escribir como:

"Si existen n posibilidades igualmente probables, donde una de ellas debe ocurrir y s se consideran favorables, o sea, un "triunfo", entonces la probabilidad de lograr el triunfo está dada por la razón s/n". [Freund, John E.]

Esta definición es discutible ya que para definir la probabilidad utiliza el término probabilidad, lo cual conceptualmente es erróneo. Podemos enunciar otro concepto, que se basa en la frecuencia, el cual dice:

"La **probabilidad** de que ocurra un acontecimiento (suceso o resultado) es la proporción de las ocasiones en que, a la larga, tienen lugar los eventos del mismo tipo". [Freund, John E.]

A partir del uso de las probabilidades, se puede estimar un valor de resultado esperado o **esperanza matemática**, la cual se define como:

$$E=a_1p_1+a_2p_2+\ldots+a_kp_k$$

donde p_1, $p_2,\ldots,p_k$ son las probabilidades de que a_1, $a_2,\ldots a_k$ sucedan.

Normalmente la esperanza matemática se utiliza con a como un valor monetario, por ello se conoce también como **Valor Monetario**

Esperado VME, que es uno de los criterios de decisión tradicionalmente utilizados para la toma de decisiones, sin embargo no es de los más acertados dado que descuida otras variables de gran importancia que pueden ser decisivas en el resultado de las decisiones tomadas.

f) Dependencia e Independencia estadística.

Existen dos tipos de eventos para la teoría de la probabilidad, los eventos independientes y los eventos dependientes; los primeros se caracterizan porque la ocurrencia de un evento E_1 no afectará a la de un evento diferente E_2, esto se escribe como:

$$P(E_1|E_2)= P(E_1) \text{ y también } P(E_2|E_1) = P(E_2)$$

Donde $P(E_1|E_2)$ es la probabilidad condicional de que ocurra E1 dado que ocurrió E2, y el resultado $P(E_1)$ (probabilidad marginal de E_1) nos indica que la probabilidad de E_1 no se ve afectada por la ocurrencia de E_2. Esto se aplica en el mismo sentido para $P(E_2|E_1)$.

Además los eventos estadísticamente independientes cumplen con

$$P(E_1 \cap E_2)= P(E_1) \bullet P(E_2)$$

Donde $P(E_1 \cap E_2)$ es la probabilidad conjunta de los eventos E1 y E2, que es igual al producto de las probabilidades marginales de dichos eventos.

Por otra parte, si la probabilidad de que ocurra un evento E_1, es modificada por la ocurrencia de otro evento E_2 se dice que el primer evento es dependiente del segundo, lo cual se escribe como:

$$P(E_1|E_2)= \frac{P(E_1 \cap E_2)}{P(E_1)}, \text{ con } P(E_1)>0$$

Y despejando obtenemos la fórmula de la probabilidad conjunta:

$$P(E_1 \cap E_2)= P(E_1) \bullet P(E_1|E_2)$$

Las relaciones de dependencia e independencia estadística son elementos básicos en la definición del método de toma de decisiones en condiciones de riesgo más usado actualmente: El método **Monte Carlo.**

2. Método Monte Carlo de Análisis de Decisiones en condiciones de Riesgo.

El método Monte Carlo estándar fue introducido por Boyle en 1977 fundamentándose en la integración estocástica directa de la ecuación de Langevin. Su funcionamiento global consiste en proporcionar un valor conocido en un momento dado, este valor es la base de generación aleatoria de otro nuevo valor a través de métodos estocásticos, el cual corresponde al valor original proyectado hacia el siguiente instante dado. El resultado final se obtiene por el promedio tras realizar este proceso un número grande de veces.

El método Monte Carlo se vale de un algoritmo denominado Metrópolis* que a través del proceso Markov realiza el muestreo asintótico en la distribución de rutas.

El proceso Markov se puede definir como la probabilidad de transición $W(\Omega_1 \rightarrow \Omega_2)$, que se lee como la probabilidad de alcanzar el punto Ω_2 partiendo del punto Ω_1. Existen dos restricciones en la selección de W:

- La definición de W debe permitir acceder cualquier punto del espacio muestral.
- W debe satisfacer la condición de balance:
  $$\mathbf{P(\Omega_1)W(\Omega_1 \rightarrow \Omega_2) = P(\Omega_2)W(\Omega_2 \rightarrow \Omega_1)}$$

Para efectos de Metrópolis usaremos la probabilidad de transición:

$$W(\Omega_1 \rightarrow \Omega_2) \begin{cases} \mathbf{P(\Omega_2)/\,P(\Omega_1)} & \mathbf{si\ P(\Omega_1) >= P(\Omega_2)} \\ 1 & \mathbf{si\ P(\Omega_1) < P(\Omega_2)} \end{cases}$$

*Extraído de http://www.npac.syr.edu/users/miloje/Finance/Option/option/section3_4.html

La prueba de que la cadena de Markov muestreará la distribución deseada $P(\Omega)$ se realiza como sigue, si nosotros partimos de una distribución probabilística inicial $P_0(\Omega)$, entonces la distribución probabilística después de **n** pasos en la cadena de Markov puede denotarse como $P_n(\Omega)$. La distribución probabilística de los pasos sucesivos **n** y **n+1** satisfacen la siguiente relación:

$$P_{n+1}(\Omega_{n+1}) = \int D\Omega_n W(\Omega_n \rightarrow \Omega_{n+1})\, P_n(\Omega_n)$$

$P(\Omega)$ es la distribución de punto fijo para este proceso de Markov. La sustitución de $P(\Omega)$ por $P_n(\Omega)$ en la siguiente ecuación combinada con la anteriormente mencionada condición de balance, implica que $P_{n+1}(\Omega)=P_n(\Omega)=P(\Omega)$, es decir la probabilidad de cada una de las distintas distribuciones es la misma, que muestra también que la distribución converge (tiene límite) hacia $P(\Omega)$. Por otra parte, una medida simple de la desviación con respecto a la distribución deseada es $D_n= \int D\Omega | P_n(\Omega) - P(\Omega)|$. Cabe mencionar que la desviación decrece conforme se avanza a lo largo de la cadena:

$$D_{n+1}=\int D\Omega_{n+1}|\int D\Omega_n W(\Omega_n \rightarrow \Omega_{n+1})P_n(\Omega_n - P(\Omega_{n+1}))|$$

$$D_{n+1}=\int D\Omega_{n+1}|\int D\Omega_n W(\Omega_n \rightarrow \Omega_{n+1})P_n(\Omega_n - P(\Omega_n))|$$

$$D_{n+1}\leq \int D\Omega_{n+1}|\int D\Omega_n W(\Omega_n \rightarrow \Omega_{n+1})P_n(\Omega_n - P(\Omega_n))|= D_n$$

Por lo tanto, $P(\Omega)$ es la distribución probabilística asintótica de los puntos generados en este proceso aleatorio.

$$P(\Omega)=\lim_{n\rightarrow\infty} P_n(\Omega)$$

La evolución de la distribución probabilística original puede verse como una relajación o suavización hacia la "Distribución Equilibrada" $P(\Omega)$. En la práctica se asume que la relajación ocurre dentro de la cadena de Markov de longitud finita R. El número R es determinado generalmente por experimentación y depende tanto de la probabilidad P como de la probabilidad W, además del grado deseado de exactitud. Dados P, W y R como números grandes el error debido a la desviación con respecto a la distribución real, es más pequeño que el debido al tamaño finito de la muestra.

Cabe mencionar que en aplicaciones con un gran número de acoplamientos de libertad, el proceso de relajación se vuelve complicado y de gran importancia.

A continuación se presentan los pasos resumidos del algoritmo de análisis completo:

1. Escoger una ruta inicial arbitraria.
2. Generar una nueva ruta alterna.
3. La nueva ruta se acepta con probabilidad W. Específicamente si $W>=1$, la nueva ruta se elige sin mayores pruebas. Si $W<1$, un aleatorio entre 0 y 1 es generado y es aceptado si es más pequeño que W. Si la ruta alterna es aceptada se convierte en la ruta actual Ω_y, de otra manera el antiguo permanece como tal.
4. Si se progresa lo suficiente a lo largo de la cadena de Markov hasta completar la relajación R, la ruta actual se toma como la distribución muestral deseada $P(\Omega)$. Se procesa la ruta actual como $F(\Omega_y)$ y se acumula el resultado $A=A+F(\Omega_y)$.
5. Ejecutar una estimación de los errores estadísticos debidos al procedimiento de muestreo Monte Carlo. Si el error está sobre el nivel deseado de precisión, regresar a 2, sino, pasar a 6.
6. Procesar el estimado de integrales requeridas por Monte Carlo. Si ***L*** denota el último valor del paso ***y,*** y ***R*** es el número de pasos de relajación, el

número total de medidas requeridas de Monte Carlo será $\boldsymbol{M_y = L \cdot R}$. El estimado Monte Carlo de la opción de valor $\langle Q\rangle_{MC}$, está dada por la función de pago F, que se obtiene:

$$\langle Q\rangle_{MC} = \frac{A}{M_y} = \frac{1}{M_y} \qquad F(\Omega_y)$$

La estimación del error requiere, también, ser acumulada:

$$\langle Q\rangle_{MC} = \frac{A}{M_y} = \frac{1}{M_y} \qquad F(\Omega_y)$$

La estimación del error muestral es obtenida como la raíz cuadrada de la varianza de la ejecución de Monte Carlo:

$$\varepsilon = (\langle \sigma^2\rangle_{MC})^{1/2}$$

$$\langle Q\rangle_{MC} = \frac{1}{M_y} (\langle Q2\rangle MC - \langle Q\rangle 2MC)$$

7. Finalizar.

Aplicación de teoría Difusa para el Análisis de Riesgos.

I. Procesamiento de Lenguaje Natural.

Desde los inicios de la computación el hombre ha tenido que adaptarse al modo de operar de la computadora (jamás en el sentido inverso), lo cual ha creado una mutación en el modo de pensar de gran número de los que nos encontramos en contacto con ellas. Este problema era mucho más notorio en los inicios del cómputo, cuando la operación de estas máquinas estaba reservada a especialistas de "bata blanca", los cuales poseían un gran conocimiento en el área y se encargaban de preparar meticulosamente las entradas que estos aparatos recibían para que cumpliese tanto con las reglas como con las características previamente establecidas para su correcto procesamiento.

Actualmente es trivial manejar una computadora personal gracias al nivel de amabilidad que se ha alcanzado con las Interfaces Gráficas, sobre todo en aplicaciones no específicas como las hojas de cálculo, procesadores de texto, etc.

A pesar de este gran avance, en las aplicaciones más especializadas (la construcción de sistemas, por ejemplo) aún se debe tener cuidado extremo en la manera de estructurar los problemas del mundo real bajo un modo de operar de decisiones bivalentes (discreto), lo que implica la creación de estructuras decisorias extremadamente complejas. Este fenómeno se presenta con más fuerza en las aplicaciones de análisis de riesgos, donde el gran manejo de parámetros y cálculos vuelve las aplicaciones excesivamente complejas tanto en su estructura como en su modo de operar.

La teoría Difusa ofrece, con su nuevo paradigma de pensamiento gradificado, la ventaja de vincularse íntimamente con el modo natural de pensamiento humano, haciendo posible la eliminación de barreras para la interpretación de lenguaje natural y su correcto procesamiento.

Para un área como el análisis de riesgos, en la cual no se tiene una medida claramente definida acerca de los posibles sucesos, y en donde se depende en gran medida de la conjeturación, intuición y estimaciones realizadas sobre la base de la experiencia humana, se requiere de una herramienta que nos permita plantear una antecedente en los términos usualmente utilizados en el área (lenguaje natural) y que emita una recomendación del mismo tipo, es decir, usando estos términos, lo cual facilitaría la toma de decisiones además de evitar la necesidad de interpretaciones adicionales (conversiones de los resultados genéricamente arrojados por los sistemas tradicionales), en las cuales se puede desvirtuar el significado de dichos resultados si tomamos en cuenta los distintos grados de experiencia del los diferentes analistas de riesgos.

En este sentido, la Teoría Difusa se puede conceptualizar como "...un intento para eliminar las barreras 'lingüísticas' entre los humanos, quienes piensan en términos difusos, y las máquinas, que sólo aceptan instrucciones precisas..." [Gupta, 1977 en Schmucker] a través del modelado de situaciones difusas como estimaciones difusas que nos permitan estimar los grados de pertenencia de un elemento de un conjunto difuso.

Para lograr esta interpretación requiere el uso de las variables lingüísticas, las cuales nos permitirán el correcto modelado de sentencias imprecisas en lenguaje natural. Cabe aclarar que el estudio de las variables lingüísticas es mucho más antiguo que la Teoría Difusa, ya en la época de Isaac Newton, Leibnitz (coinventor del cálculo junto con el primero) aseguraba que si se pudiesen encontrar símbolos apropiados para expresar nuestros pensamientos con la definición y exactitud con la que se representan los números en la aritmética, podríamos efectuar razonamientos tan concretos como los que se realizan en la aritmética.

Definamos **variable lingüística** como aquélla que toma su valor de expresiones en lenguaje natural que hacen referencia a alguna cantidad de interés. La variable lingüística se diferencia de la Teoría Difuso en el sentido de que la variable tiene como objetivo poseer expresiones precisas y manejables en lenguaje natural, mientras que la Teoría Difusa representa al medio para alcanzar dicho objetivo.

Las expresiones de lenguaje natural a las que se hace referencia pueden visualizarse como nombres para conjuntos difusos compuestos de los posibles valores numéricos que la cantidad de interés puede asumir, que representan subconjuntos de un universo de discurso.

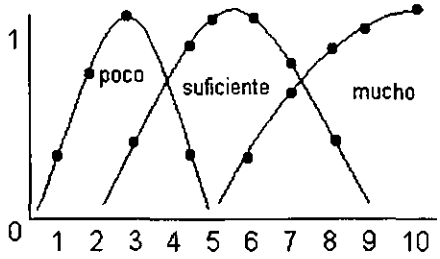
Por ejemplo, las variables "poco", "suficiente" y "mucho" pueden ser definidas de manera arbitraria (enfocándonos al espacio de los enteros del 1 al 10) como:

poco = {.4/1, .8/2, 1/3, .4/4}

suficiente = {.5/3, .8/4, 1/5, 1/6, .8/7, .5/8}

mucho = {.4/6, .6/7, .8/8, .9/9, 1/10}

Las cuales gráficamente se representan como:



Como puede verse, a través de la determinación arbitraria del grado de pertenencia de los elementos (encontrando la gráfica de la función membresía), se representan los diferentes elementos del conjunto, de esta manera pueden modelarse otros conjuntos "primarios" que se relacionen con expresiones difusas como "alto", "medio", "bajo", "fuerte", "débil", etc. que nos permitan modelar con mayor acercamiento a las expresiones que generalmente se usan en las distintas áreas donde sea aplicable el análisis.

Adicionalmente, debemos tomar en cuenta elementos que modifican el sentido de los términos primarios al mezclarse en la construcción de sentencias, dichos elementos se pueden identificar como expresiones del tipo "muy", "bastante", "más o menos", "del orden de", "no", etc. para la representación de los cuales se emplearán las anteriormente definidas operaciones adicionales de Teoría Difusa, definiendo las posibles expresiones comúnmente usadas de acuerdo al estándar establecido por Schmucker:

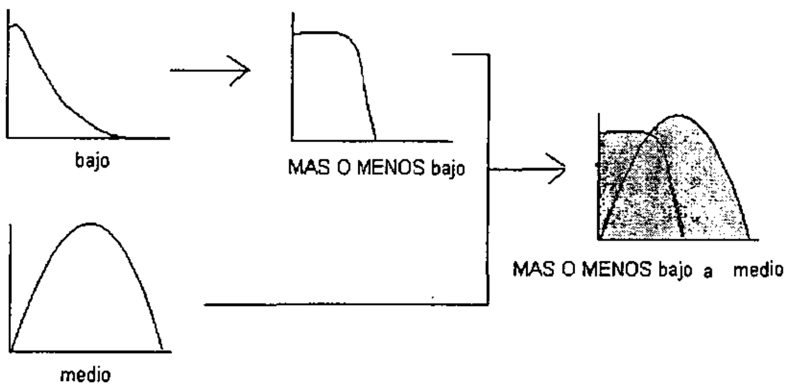
MUY término_primario=	CON(término_d)
NO término_primario =	término_d'
=	{(1 – t(x))/x|x es un elemento del universo}
MAS O MENOS término_primario =	SF(término_d, k)
DEL ORDEN DE término_primario =	NORM(INT(DIL(término_d)) AND INT(DIL(término_d')))
MEDIANAMENTE término_primario =	NORM(INT(término_d) AND (INT(CON(término_d)))')
MAS BIEN término_primario =	NORM(INT(CON (término_d) AND (CON(término_d))'))
ESCAZAMENTE término_primario =	NORM(término_d AND NO MUY término_d)

Donde AND es la intersección difusa

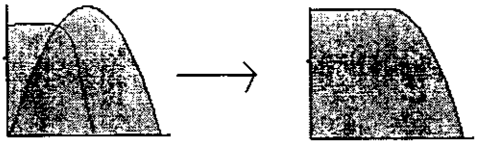
término_primario es término primario dado.

término_d es la restricción difusa {(1 – t(x))/x|x es un elemento del universo} que modela al término primario dado.

Otro tipo de expresiones son los rangos, los cuales se denotan como la unión difusa para expresiones como "desde...hasta..." y "...a...", es decir:



El cual puede cerrarse aplicando la convexión al resultado, lo cual quedaría representado como:



La convexión es una operación que se aplica al resultado de rangos de frases para ajustarlo y de esta manera darle una mejor interpretación, un significado correcto que refleje realmente lo que representa. Normalmente se aplica siempre al resultado de uniones difusas como una regla para asegurar la naturaleza de la salida.

Por último en esta sección, debemos hacer hincapié en el método para mezclar los elementos que se mencionaron de tal manera que pueda ser factible armar expresiones que tengan sentido y que expresen lo más satisfactoriamente posible las entradas que los analistas deseen introducir al sistema difuso.

Para ello nos valdremos de la notación BNF (Backus – Naur Form) o gramáticas, que en el caso de la Teoría Difusa se representará como:

<Evaluación>::=(<Primario Modificado>|<Rango>) - <Confianza>

<Confianza>::= <Difusión> CONFIABLE

<Rango>::=<Primario Modificado> a <Primario Modificado>

<Primario Modificado>::=<Modificador><Primario >|<Primario >

<Modificador>::= MUY| NO| MAS O MENOS|DEL ORDEN DE| MAS BIEN| MEDIANAMENTE| ESCAZAMENTE

<Primario >::= BAJO| MEDIO| ALTO

<Difusión>::= RAZONABLEMENTE

A partir de esta definición se pueden armar frases referentes a una expresión en particular que se quiera procesar.

II. Traducción Inversa.

Por último, ya que se ha capturado e interpretado la entrada en lenguaje natural por parte del usuario, nos queda una de las tareas más difíciles: llevar el resultado obtenido de este procesamiento hacia una expresión en lenguaje natural que además de ser fácilmente interpretada por el usuario, nos permita representar fielmente el significado de dicho resultado. A este proceso Zadeh lo denomina "aproximación lingüística".

Existen varios métodos de realizar este proceso, pero aquí sólo se explicarán las dos más usadas: el "Más Cercano" y por "Aproximaciones Sucesivas".

a) Más Cercano (Best Fit).

Es un método que generalmente se usa cuando el conjunto de expresiones de lenguaje natural es pequeño, como el generado por BNF que anteriormente se explicó.

Consiste en calcular la distancia euclideana del conjunto difuso resultante con respecto a cada uno de los conjuntos difusos que representan las expresiones posibles de lenguaje natural. Sea Z el conjunto difuso resultante y el conjunto difuso que representa una de las expresiones posibles de lenguaje natural:

Z = {z(i)/i| i es un elemento de U}

A = {a(i)/i| i es un elemento de U}

$$\text{Distancia }(Z,A) = \left[\sum_{i=1}^{9} (z(i) - a(i))^2\right]^{1/2}$$

Tras calcular la distancia hacia todas las expresiones difusas, se elige aquélla que sea menor como la expresión más adecuada.

Las ventajas de este método son que es fácil de comprender y de implementar.

b) Aproximaciones Sucesivas.

Se usa en lugar del "Más Cercano", cuando el número de posibles expresiones de lenguaje natural es muy grande.

Este método consiste en tomar los términos primarios inferior y superior más cercanos a los extremos de dichos términos y aplicarles los posibles modificadores hasta lograr una aproximación al conjunto buscado. En el momento en que se identifica la aproximación al conjunto resultante, el límite (sea inferior o superior) se sustituye por la expresión que alcanzó dicha aproximación. Cuando se logra la aproximación por ambas partes, se toma el rango de frase de ambos extremos y se calcula la distancia de los tres puntos (el inferior, el superior y el rango). El más cercano se asume como la mejor opción para expresar el conjunto resultante.

Las ventajas de este método son que excluyen de la prueba a gran número de expresiones, lo que lo hace más eficiente en tiempo y en consumo de recursos en comparación con el "Más Cercano".

III. Propuesta de Algoritmo Difuso para el Análisis de Riesgos.

Resumiendo de las secciones anteriores, el algoritmo de análisis de riesgos que se propone a través del uso de la Teoría Difusa consiste en:

Antecedentes.
Definir un BNF adecuado al área de riesgo que se pretende analizar, tomando en cuenta las diferentes variables y su mapeo difuso.

Algoritmo.
1. Mostrar al Usuario los posibles términos primarios y modificadores que puede usar para la construcción de su expresión en lenguaje natural.
2. Identificar los términos primarios y modificadores usados y traducirlos en operadores difusos.
3. Someter la frase al proceso de cálculo difuso.
4. Traducir de manera inversa el resultado difuso hacia lenguaje natural.
5. Mostrar la expresión en lenguaje natural obtenida.

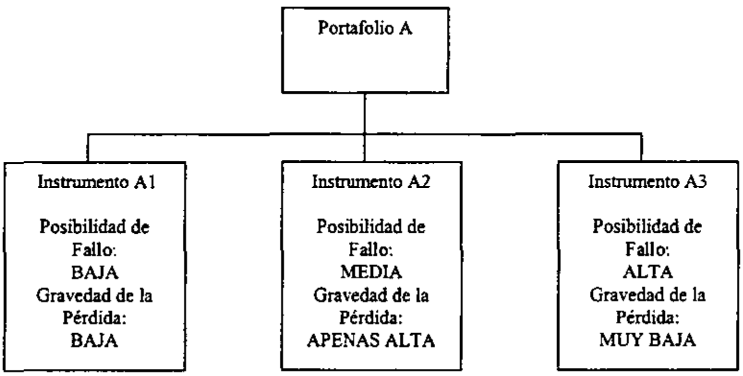
IV. Ejemplo de Análisis con Teoría Difusa.

Con el propósito de mostrar de una manera más clara el funcionamiento de la Teoría Difusa se expone el siguiente ejemplo, como caso práctico de análisis de riesgos. Las cantidades y complejidad del ejemplo son simples ya que nuestro objetivo no es complicar los cálculos sino mostrar la esencia del funcionamiento del método propuesto por la Teoría Difusa.

Evaluación de un árbol de tres nodos ponderados.

Este ejemplo mostrará la manera en que se calcula el peso (riesgo) de un problema planteado bajo una estructura de árbol, basándonos en los principios planteados anteriormente y en el algoritmo difuso definido en la sección anterior.

Inicialmente se debe inducir al usuario a estructurar el problema bajo una estructura arborescente, utilizando los términos definidos por el sistema. El esquema a usar para el ejemplo, quedará como se muestra a continuación:



Para ello es necesario definir los significados de lenguaje natural que utilizaremos en este ejemplo:

MUY BAJA	{1.0/1, .2/2, 0/3, 0/4}
BAJA	{1.0/1, .6/2, .1/3, 0/4}
MEDIA	{.2/1, 1.0/2, 1.0/3, .2/4}
APENAS ALTA	{0/1, .2/2, .9/3, .7/4}
ALTA	{0/1, .1/2, .6/3, 1.0/4}

El objetivo es encontrar el riesgo (peso) que representa este sistema completo, para ello usaremos la técnica del "Peso Difuso Promedio", que se define a partir de la analogía con el peso promedio definido para los enteros, por ejemplo.

En el caso de los enteros la expresión que representa el peso promedio, dados {Ri} como una secuencia de enteros y {Wi} como el peso de dichos enteros, es:

$$\overline{W} = \frac{\sum_{i=1}^{n} W_i \times R_i}{\sum_{i=1}^{n} W_i}$$

La misma expresión será usada para calcular el peso del sistema difuso, sólo que los operadores aritméticos de suma, multiplicación y división se redefinirán (esta redefinición puede entenderse como una sobrecarga de operadores) según el principio de extensión postulado por Zadeh en 1975, quedando como:

A+B= {min(a(i), b(j)/[i+j] | 1≤ i, j ≥4}

A*B= {min(a(i), b(j)/[i*j] | 1≤ i, j ≥4}

A/B= {min(a(i), b(j)/[i/j] | 1≤ i, j ≥4}

De la anterior definición obtenemos que nuestra estructura arborescente queda definida como:

$$\frac{(\text{BAJA*BAJA})+(\text{MEDIA*APENAS ALTA})+(\text{ALTA*MUY BAJA})}{\text{BAJA+APENAS ALTA+MUY BAJA}}$$

Donde las posibilidades de fallo sustituyen a las W's y las gravedades de pérdida a las R's.

El paso siguiente es resolver la expresión resultante, par lo cual resolveremos, primero el denominador, después el numerador y posteriormente se hará la división de los conjuntos difusos resultantes.

Primero resolvemos BAJA+APENAS ALTA que es igual a decir:

BAJA+APENAS ALTA =
{min(1.0, 0)/[1+1], min(1.0, .2)/[1+2], min(1.0, .9)/[1+3], min(1.0, .7)/[1+4],
min(.6, 0)/[2+1], min(.6, .2)/[2+2], min(.6, .9)/[2+3], min(.6, .7)/[2+4],
min(.1, 0)/[3+1], min(.1, .2)/[3+2], min(.1, .9)/[3+3], min(.1, .7)/[3+4],
min(0, 0)/[4+1], min(0, .2)/[4+2], min(0, .9)/[4+3], min(0, .7)/[4+4]}
={ 0/2, .2/3, .9/4, .7/5, 0/3, .2/4, .6/5, .6/6, 0/4, .1/5, .1/6, .1/7, 0/5, 0/6, 0/7, 0/8}
={0/1, 0/2, .2/3, .9/4, .7/5, .6/6, .1/7, 0/8}

Haciendo el resultado normal (con la operación Normalización)

={0/1, 0/2, .22/3, 1.0/4, .77/5, .66/6, .11/7, 0/8}

Sumando el resultado a la definición de MUY BAJA, quedará:

BAJA+APENAS ALTA+MUY BAJA=
{min(0, 1.0)/[1+1], min(0, .2)/[1+2], min(0, 0)/[1+3], min(0, 0)/[1+4],
min(0, 1.0)/[2+1], min(0, .2)/[2+2], min(0, 0)/[2+3], min(0, 0)/[2+4],
min(.22, 1.0)/[3+1], min(.22, .2)/[3+2], min(.22, 0)/[3+3], min(.22, 0)/[3+4],
min(1.0, 1.0)/[4+1], min(1.0, .2)/[4+2], min(1.0, 0)/[4+3], min(1.0, 0)/[4+4],
min(.77, 1.0)/[5+1], min(.77, .2)/[5+2], min(.77, 0)/[5+3], min(.77, 0)/[5+4],
min(.66, 1.0)/[6+1], min(.66, .2)/[6+2], min(.66, 0)/[6+3], min(.66, 0)/[6+4],
min(.11, 1.0)/[7+1], min(.11, .2)/[7+2], min(.11, 0)/[7+3], min(.11, 0)/[7+4],
min(0, 1.0)/[8+1], min(0, .2)/[8+2], min(0, 0)/[8+3], min(0, 0)/[8+4]}
={ 0/2, 0/3, 0/4, 0/5, 0/3, 0/4, 0/5, 0/6, .22/4, .2/5, 0/6, 0/7, 1.0/5, .2/6, 0/7, 0/8, .77/6, .2/7, 0/8, 0/9, .66/7, .2/8, 0/9, 0/10, .11/8, .11/9, 0/10, 0/11, 0/9, 0/10, 0/11, 0/12}
={ 0/1, 0/2, 0/3, .22/4, 1.0/5, .77/6, .66/7, .2/8, .11/9, 0/10, 0/11, 0/12}

Este último resultado es el denominador de nuestro sistema.

Ahora calculemos el numerador:

BAJA * BAJA=
{min(1.0, 1.0)/[1*1],min(1.0,.6)/[1*2],min(1.0,.1)/[1*3],min(1.0,0)/[1*4],
min(.6,1.0)/[2*1], min(.6,.6)/[2*2], min(.6,.1)/[2*3], min(.6,0)/[2*4],
min(.1,1.0)/[3*1], min(.1,.6)/[3*2], min(.1,.1)/[3*3], min(.1,0)/[3*4],
min(0,1.0)/[4*1], min(0,.6)/[4*2], min(0,.1)/[4*3], min(0,0)/[4*4],}
={1.0/1,.6/2,.1/3,.6/4, .1/6, 0/8,.1/9, 0/12,0/16}

Aplicando normalización y convexión:
={1.0/1,.6/2,.6/3,.6/4,.35/5, .1/6,.1/7, .1/8,.1/9, 0/10, 0/11, 0/12, 0/13, 0/14, 0/15, 0/16}

Calculando MEDIA*APENAS ALTA, queda, después de normalización y convexión, como:

MEDIA*APENAS ALTA=
{0/1, .22/2, .22/3, .22/4, .61/5, 1.0/6, 1.0/7, 1.0/8, 1.0/9, .93/10, .85/11, .77/12, .63/13, .5/14, .36/15, .22/16}

Calculando ALTA*MUY BAJA, queda, después de normalización y convexión, como:

ALTA*MUY BAJA=
{0/1, .1/2, .6/3, 1.0/4, .6/5, .2/6, .2/7, .2/8, 0/9, 0/10, 0/11, 0/12, 0/13, 0/14, 0/15, 0/16}

Realizando la suma de (BAJA*BAJA)+(MEDIA*APENAS ALTA), usando los conjuntos resultantes, quedará como:

(BAJA*BAJA)+(MEDIA*APENAS ALTA)=
{0/1, 0/2, .22/3, .22/4, .22/5, .61/6, 1.0/7, 1.0/8, 1.0/9, 1.0/10, .93/11, .85/12, .77/13, .64/14, .6/15, .6/16, .6/17, .5/18, .1/25, 0/26, 0/27, 0/28, 0/29, 0/30, 0/31, 0/32}

Haciendo la suma completa, tomando el parcial anterior tenemos:

(BAJA*BAJA)+(MEDIA*APENAS ALTA)+(ALTA*MUY BAJA)=
{0/1, 0/2, 0/3, 0/4, .1/5, .22/6, .22/7, .22/8, .6/9, .61/10, 1.0/11, 1.0/12, 1.0/13, 1.0/14, .93/15, .85/16, .77/17, .64/18, .6/19, .6/20, .6/21, .6/22, .5/23, .36/24, .35/25, .22/26, .2/27, .2/28, .2/29, .1/30, .1/31, .1/32, .1/33, 0/34, 0/35, 0/36, 0/37, 0/38, 0/39, 0/40, 041, 0/42, 0/43, 0/44, 0/45, 0/46, 0/47, 0/48}

Este resultado representa el numerador de nuestro sistema, haciendo la división difusa de los dos conjuntos, obtendremos:

<u>BAJA*BAJA)+(MEDIA*APENAS ALTA)+(ALTA*MUY BAJA)</u>
BAJA+APENAS ALTA+MUY BAJA

={0/1, 0/2, 0/3, 0/4, .1/5, .22/6, .22/7, .22/8, .6/9, .61/10, 1.0/11, 1.0/12, 1.0/13, 1.0/14, .93/15, .85/16, .77/17, .64/18, .6/19, .6/20, .6/21, .6/22, .5/23, .36/24, .35/25, .22/26, .2/27, .2/28, .2/29, .1/30, .1/31, .1/32, .1/33, 0/34, 0/35, 0/36, 0/37, 0/38, 0/39, 0/40, 041, 0/42, 0/43, 0/44, 0/45, 0/46, 0/47, 0/48} / { 0/1, 0/2, 0/3, .22/4, 1.0/5, .77/6, .66/7, .2/8, .11/9, 0/10, 0/11, 0/12}

Esta expresión se resolverá aplicando la redefinición difusa de la división, aunque es importante mencionar que el resultado deberá mapearse dentro del conjunto de enteros definido anteriormente {1, 2, 3, 4}, por lo cual todo aquél resultado que de un número menor a 1 o mayor a 4 será descartado, así como todo aquél resultado que involucre fracciones. Por tanto se tomarán en cuenta sólo los valores de membresía de los múltiplos de cada elemento del conjunto {1, 2, 3, 4}:

Elemento 1	Elemento 2	Elemento 3	Elemento 4
min(0,0)/[1/1]=0/1	min(0,0)/[2/1]=0/2	min(0,0)/[3/1]=0/3	min(0,0)/[4/1]=0/4
min(0,0)/[2/ 2]=0/1	min(0,0)/[4/ 2]=0/2	min(.22,0)/[6/ 2]=0/3	min(.22,0)/[8/ 2]=0/4
min(0,0)/[3/ 3]=0/1	min(.22,0)/[6/ 3]=0/2	min(.6,0)/[9/ 3]=0/3	min(1.0,0)/[12/ 3]=0/4
min(0,.22)/[4/ 4]=0/1	min(.22,.22)/[8/ 4]= .22/2	min(1.0,.22)/[12/ 4]= .22/3	min(.85,.22)/[16/ 4]= .22/4
min(.1,1.0)/[5/ 5]=.1/1	min(.61,1.0)/[10/ 5]= .61/2	min(.93,1.0)/[15/ 5]= .93/3	min(.6,1.0)/[20/ 5]= .6/4

Elemento 1	Elemento 2	Elemento 3	Elemento 4
min(.22,.77)/[6/ 6]= .22/1	min(1.0,.77)/[12/ 6]= .77/2	min(.64,.77)/[18/ 6]= .64/3	min(.36,.77)/[24/ 6]= .36/4
min(.22,.66)/[7/ 7]= .22/1	min(1.0,.66)/[14/ 7]= .66/2	min(.6,.66)/[21/ 7]=.6/3	min(.2,.66)/[28/ 7]= .2/4
min(.22,.2)/[8/ 8]=.2/1	min(.85,.2)/[16/ 8]=.2/2	min(.36,.2)/[24/ 8]=.2/3	min(.1,.2)/[32/ 8]=.1/4
min(.6,.11)/[9/ 9]=.11/1	min(.64,.11)/[18/ 9]=.11/2	min(.2,.11)/[27/ 9]=.11/3	min(0,.11)/[36/ 9]=0/4
min(.61,0)/[10/ 10]=0/1	min(.6,0)/[20/ 10]=0/2	min(.1,0)/[30/ 10]=0/3	min(0,0)/[40/ 10]=0/4
min(1.0,0)/[11/ 11]=0/1	min(.6,0)/[22/ 11]=0/2	min(.1,0)/[33/ 11]=0/3	min(0,0)/[44/ 11]=0/4
min(1.0,0)/[12/ 12]=0/1	min(.36,0)/[24/ 12]=0/2	min(0,0)/[36/ 12]=0/3	min(0,0)/[48/ 12]=0/4
max = .22/1	max = .77/2	max = .93/3	max = .6/4

De lo anterior tenemos que el conjunto resultante es:

{.22/1, .77/2, .93/3, .6/4}

Aplicándole normalización, queda como:

{.23/1, .82/2, 1.0/3, .64/4}

Por último usaremos el método del Más Cercano (Best Fit), para calcular la distancia entre el resultante y los conjuntos predefinidos. Esto nos servirá para hacer el mapeo del conjunto resultante hacia una salida en lenguaje natural.

Recordemos la fórmula de la distancia euclideana (V. Punto II Traducción Inversa), la cual se aplicará en nuestro cálculo:

$$\text{Distancia }(Z,A) = \left[\sum_{i=1}^{9} (z(i) - a(i))^2 \right]^{½}$$

Resultante={.23/1, .82/2, 1.0/3, .64/4}

Conjuntos Predefinidos:

MUY BAJA	{1.0/1, .2/2, 0/3, 0/4}
BAJA	{1.0/1, .6/2, .1/3, 0/4}
MEDIA	{.2/1, 1.0/2, 1.0/3, .2/4}
APENAS ALTA	{0/1, .2/2, .9/3, .7/4}
ALTA	{0/1, .1/2, .6/3, 1.0/4}

$$D\,(R,\text{MUY BAJA})=\Big[\sum_{i=1}^{9} (r(i)-mb(i))^2\Big]^{1/2}$$

$=[(.23-1.0)^2+(.82-.2)^2+(1.0-0)^2+(.64-0)^2]$

$=1.5449$

$$D\,(R,\text{BAJA})=\Big[\sum_{i=1}^{9} (r(i)-b(i))^2\Big]^{1/2}$$

$=[(.23-1.0)^2+(.82-.6)^2+(1.0-.1)^2+(.64-0)^2]$

$=1.3641$

$$D\,(R,\text{MEDIA})=\Big[\sum_{i=1}^{9} (r(i)-m(i))^2\Big]^{1/2}$$

$=[(.23-.2)^2+(.82-1.0)^2+(1.0-1.0)^2+(.64-.2)^2]$

$=.4763$

$$D\,(R,\text{APENAS ALTA})=\Big[\sum_{i=1}^{9} (r(i)-aa(i))^2\Big]^{1/2}$$

$=[(.23-0)^2+(.82-.2)^2+(1.0-.9)^2+(.64-.7)^2]$

$=.6714$

$$D\,(R,ALTA)=[\sum_{i=1}^{9}(r(i)-a(i))^2]^{1/2}$$

$$=[(.23-0)^2+(.82-.1)^2+(1.0-.6)^2+(.64-1.0)^2]$$

$$=.9278$$

De lo cual concluimos que la expresión que mejor representa a nuestro conjunto resultante es "MEDIA", por tanto nuestro sistema completo presenta un nivel de riesgo es **medio**, la decisión que deba tomarse frente a esta respuesta dependerá del entorno que rodee al caso, esto es, las políticas de aceptación de riesgos, el nivel de riesgo máximo definido para nuestro sistema, etc. De esta manera podemos decir que si se tratase de un portafolio de inversiones compuesto de tres instrumentos de inversión diferentes y la actitud de nuestro cliente frente al riesgo fuese de minimización, el portafolio se rechazaría; si por el contrario nuestro cliente tuviese una actitud de maximización de utilidades, estaría en las condiciones para aceptar el portafolio o tal vez rechazarlo si acaso existen otras opciones que le provean de una mayor rentabilidad aunque con mayor riesgo. Este portafolio será el idóneo para un cliente dispuesto a aceptar un nivel de riesgo razonable a cambio de una rentabilidad de la misma naturaleza.

Como pudo observarse el árbol del ejemplo es un árbol simple, pero para árboles más complejos el peso de cada nodo es resultado del Peso Promedio de sus ramificaciones, además de que se agrega otra variable a tomar en cuenta, la confiabilidad de la estimación de las dos variables mencionadas.

Asimismo, en un sistema más completo se debe hacer uso de definiciones con operaciones difusas para expresiones como MUY, o DEL ORDEN DE, por ejemplo, lo cual implicaría tener un solo término primario al definir BAJA y MUY BAJA, en vez de dos como se ha hecho en el ejemplo, esto quiere decir que

MUY BAJA = CON(BAJA)

En vez de definir dos conjuntos definimos uno y el otro se obtiene de la manipulación del primero.

Adicionalmente se podría tener un mayor número de expresiones "aceptables" en lenguaje natural si se incorpora la definición de un buen BNF, lo cual dotaría a nuestro sistema de mucha mayor flexibilidad, tanto de entrada de datos como de salida de información.

Análisis Comparativo del Método Difuso y el Método Monte Carlo.

El presente análisis se desarrolló sobre la base de una evaluación comparativa de puntos críticos en el área del análisis de riesgos aplicada a ambos métodos, Difuso y Monte Carlo:

1. Datos requeridos para el análisis.

Método Difuso	Método Monte Carlo
▫ Parametrización previa del sistema (modificadores y términos primarios) ▫ Expresión en lenguaje natural del usuario.	▫ Parametrización previa del sistema (establecimiento de la distribución o de un método alterno para lograr la distribución, probabilidades de las distintas rutas) ▫ Indicación de un punto inicial de partida arbitrario por parte del usuario.

2. Tipos de riesgo que analiza.

Método Difuso	Método Monte Carlo
▫ Cualquier tipo de riesgo, como se puede ver el riesgo que se analiza depende de la definición del BNF, el cual dota de flexibilidad a este método.	▫ Analiza cuestiones más bien financieras, aunque puede adaptarse a un número limitado de áreas afines.

3. Consumo de Recursos.

Método Difuso	Método Monte Carlo
▫ Como se explicó en el desarrollo, el método difuso no tiene grandes requerimientos de recursos, y se eficienta más aún si, para aquéllos casos en los que existe un gran número de variables, se usa el método de aproximaciones sucesivas.	▫ Por la naturaleza compleja de las operaciones que realiza, además de la cantidad de estas operaciones que se requiere para generar un resultado preciso que posea parámetros de confiabilidad altos, este método resulta sumamente costoso.
▫ Las operaciones que realiza no son muy complejas, por ello no requiere demasiada interacción con el procesador.	▫ Requiere de gran cantidad de tiempo en comparación al Método Difuso, ya que se basa en un número n de iteraciones, donde n es proporcional a la confiabilidad de la respuesta.
▫ Su tiempo de respuesta se puede calificar como muy bueno en comparación a Monte Carlo	

4. Manejo de Dependencias.

Método Difuso	Método Monte Carlo
▫ Asume la existencia de dependencias dentro de la definición misma de sus términos primarios y sus modificadores.	▫ Asume que todas las variables son independientes.
▫ Con dependencias desconocidas, este método se porta conservador, lo que evita tomar riesgos innecesarios.	▫ Como no toma en cuenta las dependencias se muestra demasiado temerario en su interpretación.

5. Comprensión de las entradas – salidas.

Método Difuso	Método Monte Carlo
▫ Al valerse de lenguaje natural no requiere de procesos extra para comprender los resultados ni para preparar las entradas.	▫ Las entradas y salidas por sí mismas no nos dicen nada acerca de la naturaleza del problema o de su solución, requieren de procesos de interpretación que posibiliten su comprensión.

6. Comprensión del Proceso.

Método Difuso	Método Monte Carlo
▫ El proceso es difícil de comprender para la gente familiarizada con el desarrollo de sistemas de análisis por el hecho de que no sólo se debe entender el método en sí, sino que este método implica la ruptura de antiguos paradigmas que tienen gran aceptación y arraigo. ▫ Existe un índice muy alto de resistencia al cambio. ▫ Una vez asimilado, el proceso resulta fácil e intuitivo.	▫ Los desarrolladores del análisis de riesgos conocen perfectamente el método Monte Carlo y su basamento teórico, lo cual facilita su implementación. ▫ Los usuarios comunes se verían en serios problemas si pretendiesen realizar los cálculos personalmente.

Conclusiones.

No existe una gran difusión de los métodos y principios sobre los que se sustenta la Teoría De Conjuntos Difusos, por ello la gente no vinculada al área presenta una gran resistencia a aceptar su eficiencia.

La Teoría De Conjuntos Difusos está aún en desarrollo, falta mucho por investigar y definir, esto es un punto en su contra ya que probablemente no existan las herramientas adecuadas para realizar desarrollos bajo este paradigma, por lo cual no se puede confiar plenamente en la robustez de los sistemas basados en la Teoría De Conjuntos Difusos.

La falta de herramientas orientadas a desarrollos bajo el paradigma Difuso provoca que tanto el tiempo de desarrollo como la complejidad para codificar los algoritmos sea mayor que en los sistemas basados en métodos tradicionales.

La Teoría Tradicional de Conjuntos adelanta a la Teoría De Conjuntos Difusos por muchos años de investigación y desarrollo, lo que la coloca como una teoría mucho más madura y aceptada en general.

La gente vinculada con el área del cómputo y las matemáticas ya estamos acostumbrados a pensar de manera bivalente, el cambio a un paradigma como el que representa la Teoría De Conjuntos Difusos será lento y complicado, pero las bondades que ofrece propiciarán este cambio.

Por último debemos interpretar nuestra hipótesis como correcta, ya que se ha demostrado que por la naturaleza simple de sus cálculos, su capacidad para interpretar lenguaje natural y su bajo consumo de recursos, la Teoría De Conjuntos Difusos constituye una herramienta eficaz de análisis para la toma de decisiones en condiciones de riesgo.

Bibliografía

Zimmermann, Jans Jurgen. Fuzzy sets theory and its applications, 3ª ed., Kluwer Acedemic, EUA1996.*1

Lógica Difusa, más Clara de lo que piensa. En ComputerWorld, C-14, Junio 15–19 1998

Leyva Rodríguez, Roberto. "Aparición de la Función exponencial", en Apuntes de Cálculo Integral y Diferencial, CCH Ote. UNAM, México, D.F., 1994.

Schmucker, Kurt J. Fuzzy Sets, Natural language computations, and risk analysis, Rockville, Maryland: Computer Science, 1984.*2

Covello, Vincent T. y Jeryl Mumpower. "Risk Analysis and Risk Management: A historical perrpective", en Risk Evaluation and Management, Vol. I, Plenum Press, New York,1986.

Knight, Frank H. Riesgo, Incertidumbre y Beneficio, Ed. M. Aguilar, Tr. Inglés por Ramón Verea, España, Madrid 1947.

Rheault, Jean Paul. Introducción a la Teoría de Decisiones con Aplicaciones a la Administración, Limusa, México 1973.

Freund, John E. y otros. Estadística para la Administración con Enfoque Moderno, Prentice Hall, 5ª ed., México 1995.