

Metodología de análisis de datos imprecisos con lógica difusa

Matilde Inés Césaria¹

Adriana Gámbaro²

Ricardo Césaric³

¹Universidad Tecnológica Nacional..

Facultad Regional Mendoza.

E-mail: matilde.cesari@frm.utn.edu.ar

²Universidad de la República. R.O. del Uruguay.

Facultad de Química.

E-mail: agambaro@fq.edu.uy

³Universidad Tecnológica Nacional.

Facultad Regional Mendoza.

E-mail: ricardocesari@gmail.com

RESUMEN

Una estrategia más apropiada para obtener conocimiento preciso a partir de valoraciones subjetivas, es transformar los datos mediante la matemática borrosa. Aunque puede tener potencial en el proceso de investigación científica, en la actualidad, no se dispone de criterios y métodos armonizados y sistematizados para el análisis de datos, utilizando lógica borrosa. Los resultados de experiencias con datos de pruebas sensoriales y datos de la evaluación y caracterización de los teléfonos celulares por su riesgo de uso, permitió exponer un marco metodológico para el análisis de datos sensoriales con la utilización de la lógica difusa y métodos de análisis multivariado de datos como el análisis de correspondencias de tablas de contingencia y las pruebas de asociación. Los resultados obtenidos en todos los casos, demuestran que el proceso de borrosificación y agregación borrosa, es apropiado para calcular similitudes cuando las características que se miden involucran variables de diferente naturaleza, lo que constituye un instrumento de mejora en la explotación de información y potencia las técnicas de investigación científica. Se requiere profundizar sobre nuevas líneas de investigación como la orientada a la forma de estimar de manera objetiva, los parámetros de la variable lingüística en función de las valoraciones, a la medida de cada estudio.

Palabras clave: datos imprecisos, sensimetría, lógica borrosa, correspondencias simples

INTRODUCCIÓN

La información que manejamos diariamente presenta imperfecciones, generalmente no es de tipo precisa. La incerteza es aquella que surge de procesos humanos, como ser la sensación, la percepción, la experiencia cognoscitiva, el razonamiento y el pensamiento. La percepción que logramos de nuestro entorno físico a través de nuestros sensores naturales (ojos, oídos, nariz, etc.) contiene incertezas que no pueden ser caracterizadas usando estadística o incluso razonamiento lógico tradicional. Específicamente, la Incerteza Léxica, trata la imprecisión propia en la mayoría de palabras humanas, usadas para evaluar conceptos y derivar conclusiones; con el uso de abstracción y analogías, unas pocas oraciones describirán contextos complejos que son difíciles de modelar con precisión matemática.

El estudio exploratorio de los datos imprecisos, presenta restricciones, puestas en evidencia por; 1- Variables que presentan

elevado número de valores perdidos, lo cual implica por un lado, pérdida de eficiencia en el análisis, y por otro lado, frente a la posibilidad de que los valores perdidos sigan un patrón no aleatorio, tanto ignorarlos como estimarlos, mediante alguno de los sistemas de imputación, implica un sesgo, inconveniente en la obtención de patrones mediante el análisis factorial [2]; 2- Los valores medios y la dispersión alertan de posibles deficiencias de los datos observados, en relación a problemas de la realidad estudiada e incluso la propia naturaleza subjetiva de las mediciones [3]. A partir de estos problemas, es indispensable el empleo de métodos robustos, como los que provee el análisis multivariado, con el diseño de una estrategia que permita minimizar el efecto de valoraciones atípicas (outlier) y faltantes.

El término “datos sensoriales”, se aplica a todos aquellos datos generados en la percepción de características de un producto, a través de los sentidos. El conjunto de las técnicas de recolección de este tipo de datos y métodos estadísticos de análisis asociados constituyen la rama de la Estadística llamada Sensometría [4;5].

La problemática reside en lograr que la respuesta humana sea precisa y se pueda determinar de modo cuantitativo, esto relacionado con que los resultados de las pruebas sensoriales se expresan en términos relativos y no en términos absolutos, por ende, este tipo de información implica siempre subjetividad o vaguedad; hay diferencias en las percepciones sensoriales individuales en cuanto a gustos, sonidos, formas, iluminación, olores, etc., de tal forma que cada sujeto tiene una imagen global del mundo, esta variabilidad esta reforzada por la educación, nivel social, cultura y personalidad del catador [6].

Por tanto, también, es conveniente buscar otras formas de representar el dato, más allá de tablas cuantitativas y cualitativas para tratar con valoraciones que provengan de mediciones subjetivas, esto último relacionado con la presencia de datos perdidos y/o atípicos.

En otras áreas, los datos no provienen di-

rectamente de la percepción de los sentidos humanos pero poseen una naturaleza imprecisa. El teléfono móvil es un transmisor y receptor de radio, cuando está encendido, recibe y transmite energía de Radiofrecuencia (RF), cuando está en una llamada telefónica, el manejo de su llamada controla el nivel de potencia al que transmite el dispositivo; está diseñado y fabricado para no exceder los límites de exposición recomendados por la International Commission on Non-ionizing Radiation Protection (ICNIRP). Sin embargo la cantidad de energía de radiofrecuencia RF a la que una persona está expuesta depende de muchos factores [7].

En la evaluación de los teléfonos celulares se valora una única variable continua, el valor de nivel SAR (Local) certificado por fábrica y asociado con criterios de riesgos de uso de los mismos, el enfoque mediante una metodología basado en la lógica difusa, para el análisis de dichos datos; permitirá crear otras variables que representara los valores borrosos en el solapamiento de los criterios de riesgo de uso [8].

La teoría de la posibilidad, constituyen modelos, basados en la lógica borrosa, que resultan especialmente útiles para tratar con la imprecisión de manera más “natural” y más “humana”. Emplean una terminología particular: “Fuzzy”(Difuso o Borroso); “Crisp” (Nítido); “Fuzzification” (Borrosificación: convertir un conjunto nítido en borroso); y “Defuzzification” (Desborrosificación: convertir un conjunto borroso en un valor “Crisp”). El concepto de la lógica borrosa, surgió en U.S.A. por Zadeh en 1965 [9], y básicamente busca producir resultados exactos a partir de conocimiento con incertidumbre. Así, lo difuso puede entenderse como la posibilidad de asignar diferentes valores de verdad a los enunciados y no solamente los clásicos «falso» o «verdadero». Asimismo, ofrece un marco de trabajo natural donde las cantidades borrosas son interpretadas como distribuciones de posibilidad, describiendo los valores del conocimiento impreciso en conjuntos borrosos.

El enfoque difuso permite representar los

datos por medio de variables lingüísticas borrosas, donde cada valor es un número borroso (conjunto difuso), por ejemplo, la variable lingüística “Intensidad Aroma” puede tomar los valores lingüísticos “ausente”, “muy débil”, “débil”, “medio, fuerte” y “muy fuerte”, que tienen un significado semántico y que se pueden expresar numéricamente por un grado de verdad entre 0 y 1 [10; 11; 12].

Aunque el razonamiento borroso puede tener potencial como método eficaz en el análisis de datos, en la actualidad, hay unas pocas referencias sobre cómo llevar a cabo la aplicación de la lógica difusa en la sensorimetría, y no se dispone de criterios y métodos armonizados y normalizados para una evaluación objetiva y reproducible, utilizando la matemática borrosa.

El Análisis Factorial Multivariado de Datos (AFMD) en la versión de la escuela francesa, surge en la década de los 70, planteando fines menos deterministas que los de la Estadística tradicional, su objetivo general es la búsqueda de una estructura presente en los datos, en un contexto de tipo más abductivo que deductivo, que revaloriza el rol del individuo estadístico [13]. En la actualidad, las técnicas de AFMD tales como el Análisis Factorial de Correspondencias Simples (ACS) y las pruebas de asociación y validación de tablas de contingencia, presentan salidas gráficas fáciles de interpretar y revelan relaciones no lineales entre las modalidades de diferentes variables que otros métodos no permiten observar. En el análisis de números borrosos trabajamos es apropiado el uso de estas técnicas al representar los datos en tablas de contingencia.

El tema que se desarrolla en este trabajo busca exponer un esquema metodológico basado en el enfoque difuso, para el análisis de datos, con la finalidad de demostrar que el empleo del análisis multivariado a partir de datos difusos que representan la información de los individuos de manera más realista, con fin de demostrar que enriquece el proceso de análisis, ya que constituye un instrumento de observación apropiado que facilita la explota-

ción de datos cuantitativos, cualitativos y también borrosos.

Para conseguir este propósito, en el marco de la tesis “Estudio y Aplicación de la lógica borrosa para medir, analizar y evaluar Calidad Sensorial en Alimentos”, para optar al título de Doctora en Alimentación, de la carrera en Red Norte de Argentina, se planteó la representación de la incertidumbre a través de la lógica difusa y presentó, a partir del estudio de diferentes casos, y la aplicación de técnicas de análisis factorial de correspondencia y técnicas de validación y descripción en el tratamiento de datos borrosos, con datos otorgados tanto por la cátedra de Bromatología de la FCA – UNCuyo, Mendoza, Centro de Estudios de Enología de la Estación Experimental Agropecuaria (EEA) Mendoza (INTA) y Sección Evaluación Sensorial Cátedra de Ciencia y Tecnología de Alimentos, Facultad de Química Universidad de la República del Uruguay.

Asimismo, en el marco del Proyecto de Investigación acreditado PID- Cod. 25/J078-FRM - UTN. 2012-2015, se utilizó la matemática borrosa y métodos multivariados en la evaluación y caracterización de los teléfonos celulares por su riesgo de uso.

MATERIALES Y MÉTODOS

Para establecer un marco de trabajo que resulte flexible y eficiente en el manejo de información imprecisa asociada al análisis de datos, se representó la incertidumbre a través de la lógica difusa y se aplicó técnicas de análisis multivariante en el tratamiento de datos borrosos, en el ámbito del análisis sensorimétrico de vinos tintos, así como también, en alimentos vegetales de IV gama; y en un estudio sobre marcas, modelos y características de las energías radiantes (SAR) de los teléfonos celulares.

Estudio 1: Se utilizaron los resultados de pruebas sensoriales realizadas en el 2007, sobre 8 muestras de tomate (*Lycopersicon Esculentum*) con destino mercado en fresco. Los datos fueron provistos por la Cátedra de Bromatología vegetal de la Facultad de

Ciencias Agrarias, Universidad Nacional de Cuyo, Mendoza, los mismos provienen de una prueba descriptiva, con escalas estructuradas de 5 puntos, a través de un panel de 14 jueces, para evaluar la percepción sensorial del tomate entero y en mitad, a través de los descriptores: Aspecto general, Color de piel y forma del tomate entero; y Color pulpa, Relación pulpa /semilla, Olor, Sabor característico, dulce y ácido, Jugosidad, Textura pulpa y piel, del tomate en mitad.

Estudio 2: Se utilizaron los resultados de pruebas sensoriales para elaboraciones a escala piloto de la cv. Malbec, de diferentes regiones de Mendoza, Argentina. Los datos fueron provistos por el Centro de Estudios de Enología de la Estación Experimental Agropecuaria (EEA) Mendoza (INTA). Los mismos provienen de un estudio sensorial, bajo las normas IRAM del AS [IRAM 20014; 20019; 20002], a través de un panel de 12 jueces pertenecientes a la Estación Experimental Agropecuaria del INTA Mendoza, entrenados donde se evaluaron en una única sesión, en la EEA - INTA, copas de vino a partir de una planilla desarrollada por el Agro Montpellier, en Francia, utilizando dos descriptores para el examen visual, ocho para el examen olfativo y siete para el examen gustativo. Se utilizó en la prueba una escala de tipo lineal, continua, con valores de 0 (ausencia) a 5 (muy fuerte). También se estableció el puntaje final de 0 a 10, respecto a evaluación global del juez de la calidad de cada muestra.

Estudio 3: Se utilizaron los resultados de pruebas afectivas sobre muestras de alfajores de chocolate, realizadas para un estudio de vida útil sensorial. Los datos fueron provistos por la Dra. Adriana Gámbaro en el marco del curso de doctorado dictado en la Universidad Nacional de Jujuy, en marzo 2012. Los datos utilizados provienen de pruebas donde se evaluaron muestras de diferentes lotes de producción, con 0 días (testigo), 28, 46, 60, 65, 70, 75 y 80 días de almacenamiento a 20°C, Se presentaron a 50 consumidores las 8 muestras en orden aleatorio y se pidió que expresaran su grado de aceptación utilizan-

do una escala hedónicas estructurada de 9 puntos y la pregunta “¿Consumiría Ud. este alfajor?” (Si-No de aceptabilidad o rechazo).

Estudio 5: Con el objeto de proporcionar información sobre marcas, modelos y características de las energías radiantes (SAR) de los teléfonos celulares, el Laboratorio LECER evaluó la calidad de los aparatos que se comercializan en el mercado. Se recopiló información sobre aparatos celulares, de distintas marcas y modelos más reconocidos en la Argentina, Latinoamérica, EEUU y Europa, de listados obtenidos en sitios Web registrados y verificados con las especificaciones técnicas dadas por los fabricantes y certificadas por Organismos Internacionales de Control. Se utilizó esta información para su evaluación teniendo en cuenta los criterios dados por los niveles de exposición relativa en una escala de 0-10 [14], y los diferentes criterios de los riesgos de exposición al uso del teléfono celular, utilizando diversos parámetros y la tasa de absorción específica SARlocal certificado por el Fabricante, factores definidos por Soo Chan Kim [15].

RESULTADOS

La propuesta Metodológica de Análisis de Datos Borrosos, involucra 4 grandes procesos: 1- La codificación y estandarización de los datos medidos; 2- La definición de las variables lingüísticas difusas y su empleo en la transformación de los datos en números borrosos representados en tablas de contingencia; 3- La obtención de un valor borroso colectivo y un valor crisp final representativo de la calificación de las muestras evaluadas; 4- aplicación de métodos multivariados para análisis, visualización y obtención de conocimiento sobre los datos borrosos.

1. La codificación y estandarización de Datos Sensoriales

1.1. Tabulación. Las valoraciones obtenidas se presentan en diversas Tablas de Datos; naturaleza cuantitativa (datos no estructurados tratados como números) o de naturaleza cualitativa (datos estructurados en categorías o etiquetas verbales). Se agre-

gan -en columna- variables nominales cualitativas, para representar grupos de observaciones, (por ejemplo, copa, cultivar, tiempo de conservación, tipo de tratamiento, etc.), cualquier información complementaria que explique la unidad de análisis observada (la muestra del alimento), así como información relevante sobre jueces y consumidores, (tal como edad, sexo, etc) [16].

1.2. ESTANDARIZACIÓN.

Opcionalmente, las puntuaciones en los atributos se normalizan a efectos de eliminar problemas de cálculo originados en la utilización de diferentes escalas. De esta manera posibilitando utilizar una misma variable lingüística difusa, para el proceso de Borrosificación. Para la estandarización de los datos, se propone utilizar el método de *normalización lineal* a valores de 0 a 1, usando la ecuación de la recta.

2. TRANSFORMAR LAS VALORACIONES A DATOS BORROSOS

Los atributos medidos se representan a través de una o varias variables lingüísticas borrosas, lo que requiere seleccionar los *descriptores lingüísticos* adecuados y su *semántica*.

2.1. DESCRIPTORES LINGÜÍSTICOS. DESCRIPTORES LINGÜÍSTICOS.

Se elige un conjunto de etiquetas o de términos lingüísticos (conjuntos difusos). Se consideran i etiquetas lingüísticas finitas y totalmente ordenadas, como $C = \{c_0, c_1, \dots, c_i\}$, $c_0 \leq c_1 \leq \dots \leq c_i$, donde cada etiqueta c_i representa la valoración lingüística para cada descriptor (se denomina *granularidad* de la variable), el número de conjuntos difusos, se define por el analista en función de los objetivos de la prueba. Valores típicos de "granularidad" son aquellos impares como 5, 7 ó 9, donde el término medio representa una valor indiferente o punto medio, y el resto se distribuye simétricamente a su alrededor, sin embargo, las etiquetas y su significado deben ser acordes a las escalas utilizadas en las pruebas, por lo que no siempre el valor central determina

la indiferencia, y no siempre las escalas son simétricas [17; 18; 19].

2.2. Semántica de los Conjuntos difusos.

Este procedimiento define para cada conjunto difuso (términos lingüísticos) una función matemática que representa el grado de posibilidad o verdad de una percepción pertenecer a un conjunto borroso. Los tipos de funciones de pertenencia comúnmente utilizados son: la función *Triangular*, *Trapezoidal* y *Gaussiana*, entre otras [20]. Las estimaciones lingüísticas con una escala de medición de la cual se obtiene observaciones numéricas, tienden a seguir una distribución normal y es apropiado el uso de números borrosos Gaussianos. En un conjunto difuso con *función Gaussiana* (Figura 1), se definen dos parámetros (a y b) que determinan la media o desvío de una campana de gauss, donde a representa el valor medio central, con total pertenencia al conjunto ($\mu=1$), y b es el desvío que representa la amplitud de la campana. Otra forma alternativa es la de la 1.

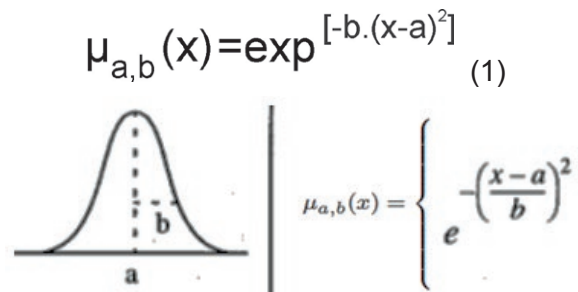


Figura 1 – Representación gráfica y parámetros de la Función de Pertenencia Gaussiana. A la derecha se representa el cálculo de $\mu_{a,b}(x)$ eje vertical [12]

2.3. BORROSIFICACIÓN.

Los datos sensoriales tabulados en punto 1, se convierte en un número borroso derivado empíricamente. A través de las ecuaciones de las funciones de pertenencia definidas para las variables lingüísticas en punto 2.2, se compara el valor sensorial con la función matemática correspondiente, obteniendo un valor entre cero y uno para cada conjunto difuso, cada uno de estos valores de posibilidad de asociarse a cada etiqueta lingüística, constituye un número borroso. Los datos bo-

rosificados se representan en una tabla de contingencia, en columna se representa los conjuntos borrosos de cada descriptor, en fila se identifica la percepción para un panelista de cada muestra, la celda representa el valor de posibilidad obtenido.

3. AGREGACIÓN

En esta etapa, se determina un valor colectivo, para cada característica, aplicando un “operador de agregación” sobre las evaluaciones borrosificadas, y posteriormente un “método de desborrosificación” para calcular un valor preciso (crisp). El procedimiento consiste, en calcular para cada conjunto borroso c_i la unión $t(c_i)$ del conjunto difuso. Se propone un operador de la familia de t-conormas, denominado *suma algebraica* de los valores de pertenencia $\mu(c_i)$ de los p panelistas (2), al ser considerado apropiado para la agregación de datos subjetivos, siendo un operador probabilístico. Se obtiene un valor agregado colectivo del grupo de observaciones, para cada conjunto difuso correspondiente un atributo.

$$t(c_i) = \sum_1^p \mu(c_i) - \prod_1^p \mu(c_i) \quad (2)$$

p =número de observaciones en el grupo de agregación (eje. jueces y/o consumidores)
 c_i = identifica cada conjunto difuso de la variable lingüística (toma valores desde 0)

Luego, a partir de la agregación $t(c_i)$, se identifica la etiqueta lingüística o conjunto difuso colectivo de los p observaciones a agregar; para una granularidad de 5, son $k=5$ conjuntos difusos, representados por c_i etiquetas (i varía de 0 a 4), se calcula la etiqueta lingüística agregada según la Ecuación 3, donde S podrá tomar valores de 0 a 4 (se redondea para obtener un valor discreto), la etiqueta lingüística es la representada por I_s .

$$S = \text{entero} \left(\frac{\sum_1^k t(c_i) \cdot i}{\sum_1^k t(c_i)} \right) \quad (3)$$

k =granularidad de la variable lingüística (número de conjuntos difusos)

i = identifica cada etiqueta o conjunto difuso (toma valores de 0 hasta $k-1$)

$t(c_i)$ = agregación del conjunto difuso c_i (unión)

Por medio de los métodos de *desborrosificación*, a partir del conjunto de valores borrosos agregados, se obtiene información precisa expresada mediante un valor crisp. Cada método utiliza una fórmula matemática para calcular la salida final. La bibliografía sugiere diversas alternativas de operadores de desborrosificación [20].

En el caso de usar funciones de tipo gaussianas, se propone implementar el método de la Altura modificado; se divide la agregación t del conjunto difuso c_i , por el desvío (parámetro b) elevado al cuadrado, finalmente para obtener el valor preciso correspondiente se utiliza la Ecuación 4.

$$V_{\text{CRISP}} = \frac{\sum_1^k \left(\frac{t(c_i)}{b^2} \cdot a_i \right)}{\sum_1^k \frac{t(c_i)}{b^2}} \quad (4)$$

Dónde:

k =granularidad de la variable lingüística (número de conjuntos difusos)

i = identifica a un conjuntos difusos (etiqueta lingüística de 0 a $k-1$)

a_i = centro de gravedad de cada conjunto c_i y b^2 =desvío al cuadrado

$t(c_i)$ = unión de los valor de pertenencia de p observaciones para un conjunto c_i

4. ESTRATEGIA METODOLÓGICA PARA EL ANÁLISIS DE DATOS BORROSOS

Se establecen dos sistemas de tablas de contingencia, a partir de los números borrosos agregados en la unión y las valoraciones borrosificadas. Todos los datos son estandarizados mediante el método de normalización lineal en valores de 0 a 100, de esta manera



se representa *porcentaje de posibilidad* que asocia un elemento en fila con un elemento columna.

Se realiza un *análisis factorial de correspondencias simples* (AFCS) para visualizar asociaciones [21; 22; 23; 24]. La interpretación, descripción y validación, se efectúa por las pruebas del χ^2 , y prueba *exactas de Fisher* [25]. Asimismo, el *valor de test* calculado sobre la prueba exacta de Fisher, permite seleccionar los atributos más característicos según el *grado de posibilidad* de pertenecer a un conjunto, o según las diferencias significativas con la percepción media de cada panelista [26; 27; 28; 29; 30].

Para identificar y graficar la función de pertenencia de la variables lingüística se utilizó los gráficos XY dispersión de la planilla de cálculo Microsoft Excel, y para la puesta en práctica de la estrategia de análisis factorial y pruebas de asociación, se utilizó el software de libre distribución para uso académico DTM. Minería de Datos y Textos, [28]; y el módulo de análisis sensorial del software comercial Xlstat [31].

5. CASO DE APLICACIÓN.

Para evaluar su validez exponemos algunos resultados en su aplicación en la evaluación de los teléfonos celulares del valor de nivel SAR (Local) certificado por fábrica y el criterios de riesgos de uso de los mismos,

mediante una metodología basado en el enfoque lingüístico difuso, para el análisis de datos; permitiendo crear otras variables que representara los valores borrosos en el solapamiento de las escalas. Con la aplicación a una tabla de números borrosos obtenida a partir de los datos de niveles de exposición a la energía radiante de los aparatos, se obtuvieron los cartografiados correspondientes a cada análisis factorial. La caracterización de los teléfonos celulares, según la marca y sus modelos, se realizó clasificando el riesgo individual de la exposición utilizando la información disponible de tasa de absorción específica (SARlocal) certificada por el Fabricante y un tiempo de exposición de 30 minutos, estimando los niveles de exposición relativa a riesgos de exposición al uso, según los criterios dados en Tabla 1.

Para la representación del riesgo difuso según SAR, se definió una variable lingüística, con 8 conjuntos difusos, para valores de 0 a 2, respecto a "*intensidad de absorción relativa*", también se utilizó una Función de pertenencia tipo Triangular [20] con parámetros descritos en la (Figura 2).

Tabla 1 – Criterios de riesgo de uso comparados y niveles de exposición relativa [15]

Criterios de riesgos de uso	Rango SAR W/kg	Media	Nivel de exposición relativa
uso sin riesgo	0,02 - 0,50	0,25	1-2
exposición segura	0,51 - 0,70	0,61	3-4
poco riesgo	0,71 - 0,90	0,81	5
riesgo aceptable	0,91 - 1,00	0,96	6
tiempo limitado	1,01 - 1,40	1,21	7
manos libres	1,41 - 1,47	1,44	8
evitar uso cabeza	1,48 - 1,59	1,54	9
riesgo claro	1,60 - 1,80	1,70	10

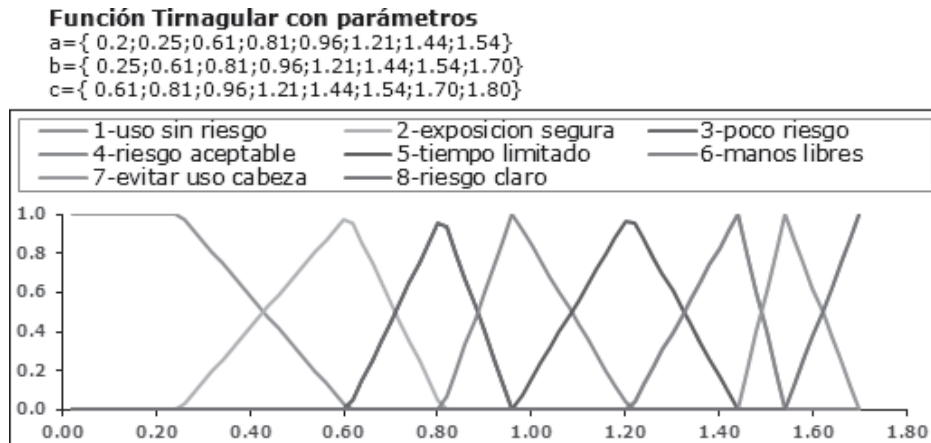


Figura 2 – Definición de la Variable lingüística Riesgo Difuso

La valoración de la tasa de absorción específica (SAR_{local}) certificada por el Fabricante se convierte en un número borroso derivado empíricamente, con las ecuaciones de las funciones de pertenencia definidas para las variables lingüísticas, se compara el valor estandarizado con la función matemática correspondiente. Los datos borrosos se representan en una tabla de contingencia, se utiliza datos numéricos de 0 a 1, en columna se representa los conjuntos borrosos, y en fila se identificó la exposición de cada marca y modelo de celular. Todos los datos son estandarizados mediante el método de normalización lineal en valores de 0 a 100, de esta manera se representa *porcentaje de posibilidad* que asocia un elemento en fila con un elemento columna [32].

Se aplicó las técnicas de AFMD sobre datos difusos, para estimar los niveles de exposición relativa en una escala de 0-10, asociados a diferentes criterios de los riesgos de exposición de uso del teléfono celular, utilizando parámetros tales como el tiempo de uso por comunicación, el uso total período, el uso cercano a la cabeza (5 cm), de manos libres y la tasa de absorción específica (SAR_{local}) certificada por el Fabricante. Con estos criterios de niveles de exposición se caracterizaron más 3.200 teléfonos celulares de las marcas más reconocidas en relación a su uso recomendado.

Los parámetros considerados para los AFCS, son parte de los resultados del anterior estudio de Breckenkamp en el 2008 [33]: (1) Tiempo promedio de uso por comunicación, (2) período total de uso diario. Estos parámetros influyen en la cantidad de la exposición y por tanto son los factores más importantes. En nuestro estudio, se han considerado estos factores para 30 minutos por comunicación. (3) uso manos libres. Esta condición reduce significativamente el nivel de exposición. (4) SAR_{local} certificado de fábrica del teléfono. Este parámetro indica la potencia de emisión de RF del aparato y estrechamente relacionado con la cantidad de potencia absorbida por unidad de masa de un cuerpo vivo expuesto a la radiación. El nivel de exposición depende estrechamente de la potencia de emisión y SAR_{local} del teléfono celular. Los valores de SAR local se obtuvieron de los fabricantes de los modelos; (5) La forma de uso del teléfono celular. Mientras se mantiene una postura típica, “escriba” (alejado de la cabeza) en el uso del teléfono celular, tiene un nivel 40% inferior de SAR (Local) que el tipo “slip” (cerca de la cabeza,) debido a la distancia entre la antena y la cabeza, [34].

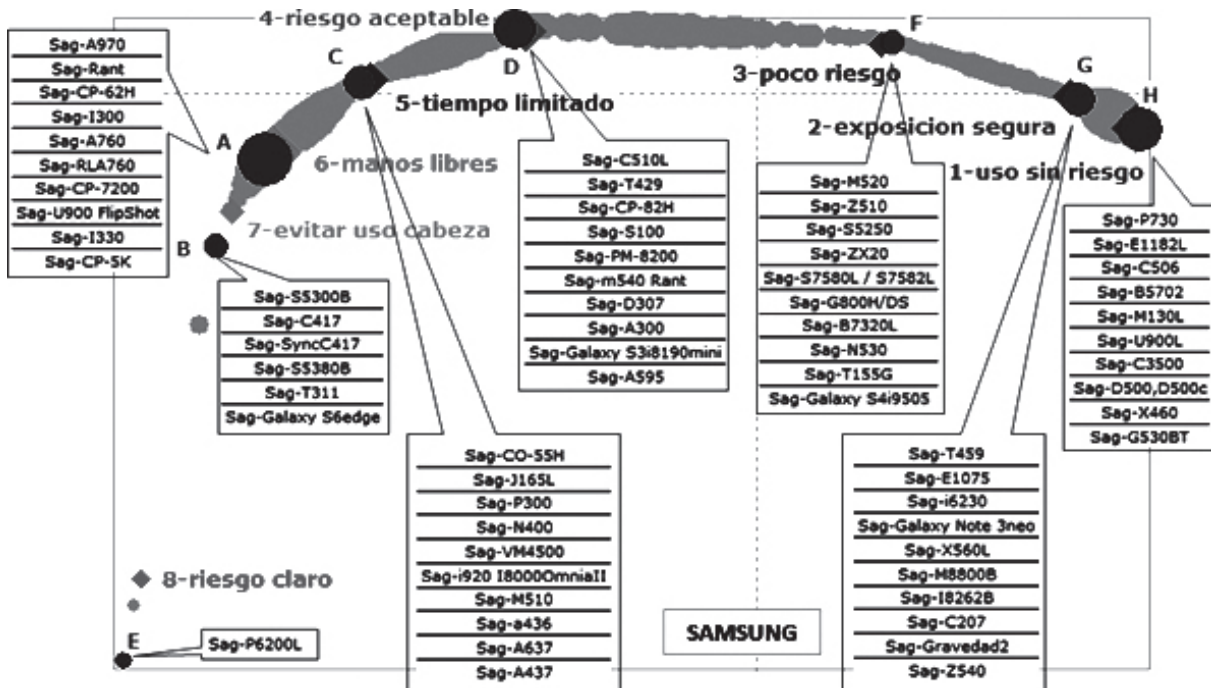


Figura 3 – Resultados Riesgos de uso para teléfonos celulares de una conocida marca

Para cada marca y sus modelos, también se obtuvieron los mapas individuales, donde se representan los riesgos de uso para todos los Modelos en varios Cartografiados, en las Figuras 3 se visualiza los mapas de riesgo una de las marcas más conocidas. La proximidad en el mapa indica las asociaciones *telefono-conjunto*, más significativas en función de la prueba exacta de Fisher ($\alpha \leq 0.01$).

CONCLUSIONES

Sin dudas, los resultados obtenidos, constituyen un instrumento de mejora en la explotación de información y potencia las técnicas de investigación científica. La teoría de la posibilidad, constituyen modelos que resultan especialmente útiles para tratar con la imprecisión de manera más “natural” y más “humana” que la lógica y la teoría de conjuntos clásicas. Se puede afirmar que el empleo del Análisis Multivariado a partir de valoraciones borrosas, plantea un nuevo paradigma en el proceso de investigación científica, y enriquece el proceso de análisis de datos, constituyendo un instrumento de observación apropiado, facilitando la explotación de datos imprecisos.

Quedan en evidencia las propiedades ventajosas del uso de números borrosos y métodos multivariados: el proceso de borrosificación y agregación borrosa, es apropiado para calcular similitudes cuando las características que se miden involucran variables

de diferente naturaleza (cuantitativa, cualitativa, binaria, textual); se logró representar información que casi siempre es poco precisa, de manera más apropiada y que facilita la convergencia del lenguaje hacia significados compartidos; permite lograr mayor precisión y detalle en el análisis multivariado de datos, sin que las valoraciones atípicas, afecten los resultados; también la metodología propuesta ayuda a determinar en qué atributos hay mayores discrepancia, otra característica es que es posible trabajar con datos en las que faltan observaciones de algunas variables, sin implicar que se suprima de todo el vector que representa a la unidad muestral; finalmente los resultados obtenidos confirma que los “valores test” basado en la prueba de Fisher, constituyen una herramienta rápida para la interpretación y validación de los resultados de análisis factoriales.

Queda abierta la discusión sobre la metodología planteada y se requiere profundizar

sobre las dispersiones de las funciones de pertenencia, las correlaciones de dichas funciones, etc. Lo dicho lleva a proponer nuevas líneas de investigación como la orientada a la forma de estimar de manera objetiva, los parámetros de la variable lingüística en función de las valoraciones, diseñando variables lingüísticas a la medida de cada estudio.

REFERENCIAS BIBLIOGRÁFICAS

- [1] Guadarrama Cotado, S. (2000). Representación del conocimiento impreciso: Revisión parcial de las teorías de conjuntos borrosos. Doctoral dissertation, Facultad Informática, Universidad Politécnica De Madrid. Disponible en:
<http://www.researchgate.net/publication/39425302>
- [2] Little, R.J.; Rubin, D.B. (2014). Statistical analysis with missing data. John Wiley & Sons
- [3] Zimek, A.; Schubert, E.; & Kriegel, H.P. (2012). A survey on unsupervised outlier detection in high-dimensional numerical data. *Statistical Analysis and Data Mining: The ASA Data Science Journal*, 5(5), pp. 363-387
- [4] Álvarez-Esteban, R.; Aguado Rodríguez, P. (2014). Datos Textuales como elementos Activos en Sensometría / textual data as active elements in sensometry. Universidad de León .Pecunia Monográfico, pp. 31-51
- [5] Bécue-Bertaut, M.; Salazar, B.C.; Vinué, P.F. (2009). Extensión del análisis factorial múltiple jerárquico a datos textuales: Aplicación en sensometría. In XXXI Congreso Nacional de Estadística e Investigación Operativa; V Jornadas de Estadística Pública: Murcia, 10-13 de febrero de 2009: Libro de Actas (p. 27)
- [6] Barda, N. (2011). Análisis sensorial de los alimentos. *Fruticultura & Diversificación* En línea: <http://www.biblioteca.org.ar/libros/210470.pdf>
- [7] Bit-Babik, G.; Chou, C.K.; Faraone, A.; Gessner, A.; Kanda, M.; Balzano, P. (2003). Estimación de la SAR en la cabeza humana y cuerpo debido a la exposición a la radiación de radiofrecuencia de mano teléfonos móviles con accesorios de manos libres. *Radiat Res* 159 (4), pp. 550-557.
- [8] Bahillo, A.; Blas, J.; Fernández, P., Mazuelas, S.; Vinuela, A.Á., Lorenzo, R.M.; Abril, E.J. (2008). E-field errors associated with RF dosimeters for RF human exposure assessment in urban environments. In *Engineering in Medicine and Biology Society, 2008. EMBS 2008. 30th Annual International Conference of the IEEE*, p. 2821-2824.
- [9] Zadeh, L. A. (1965). Fuzzy sets. *Information and control*, 8(3), pp. 338-353
- [10] Zadeh, L.A. (2014a). *Computing with Words: Principal Concepts and Ideas*. Springer Publishing Company.
- [11] Zadeh, L.A.; Fu, K.S.; Tanaka, K. (Eds.) (2014b). *Fuzzy Sets and Their Applications to Cognitive and Decision Processes: Proceedings of the US-Japan Seminar on Fuzzy Sets and Their Applications, Held at the University of California, Berkeley, California, July 1-4, 1974*. Academic press.
- [12] Zadeh, L.A. (2015). The information principle. *Information Sciences*, 294, pp. 540-549
- [13] Lebart, L.; Morineau, A.; Piron, M. (1995). *Statistique exploratoire multidimensionnelle*. Paris: Dunod, 1995, 439 p. Citado por: Campo E.P. et al, *Combinación de métodos factoriales y de análisis de conglomerados en R: el paquete FactoClass*. *Revista Colombiana de estadística*, 2007, v. 30, n. 2, pp. 231-245.
- [14] Kim, S.C.; Nam, K.C.; Kim, D.W. (2006). Estimation of relative exposure levels for cellular phone users using a neural network. *Bioelectromagnetics*, v. 27, n.6, 440-444.
- [15] Soo, C.K.; Chang, N.; Deok W.K. (2009). "Criterios de niveles de riesgos de exposición RF". Escuela de Graduados de Bio & Tecnología de la Información, Universidad Nacional Hankyong, Anseong, Corea, Departamento de Ingeniería Médica, Yonsei University College de Medicina de Seúl, Corea.
- [16] Césari R.; Césari, M. (2012). Material pedagógico de cursos de postgrado en métodos para analiza datos como soporte a la investigación en ingeniería, Capítulo 03: "Las técnicas del análisis multivariado de datos"

y Capítulo 03 "Procedimiento de descripción y validación estadística", Maestría en Medio Ambiente de la UNCu y Especialización en Redes de Datos de la FRM, UTN, Mendoza, Argentina

[17] Bonissone, P.P.; Decker, K.S. (2013). Selecting uncertainty calculi and granularity: An experiment in trading-off precision and complexity. arXiv preprint arXiv, pp.1304-3425

[18] Delgado, M.; Herrera, F.; Herrera-Viedma, E.; Verdegay, J.L.; Vila, M.A. (2013). Aggregation of Linguistic Information Based on a Symbolic Approach. Computing with Words in Information/Intelligent Systems 1: Foundations, 33, p. 428.

[19] Espinilla, M.; Martínez, L.; Pérez, L.G.; Liu, J. (2008). Modelo de Evaluación Sensorial con Información Lingüística Multigranular para el Aceite de Oliva. XIV Congreso Español sobre Tecnologías y Lógica Fuzzy, pp. 249-255.

[20] Zimmermann, H.J. (2014). Fuzzy Set Theory and Its Applications 3. Springer Publishing Company, 435 páginas.

[21] Césari R.; Césari M. (2012). Material pedagógico de cursos de postgrado en métodos para analizar datos como soporte a la investigación en ingeniería, Capítulo 03: "Las técnicas del análisis multivariado de datos" y Capítulo 03 "Procedimiento de descripción y validación estadística", Maestría en Medio Ambiente de la UNCuyo y Especialización en Redes de Datos de la FRM, UTN, Mendoza, Argentina

[22] Greenacre, M. (2007). Correspondence analysis in practice. Second Edition CRC Press, 7 may. 2007 - 296 páginas

[23] Greenacre, M. (1994). Correspondence analysis and its interpretation. Correspondence analysis in the social sciences, pp. 3-22

[24] Langrand, C.; Pinzón, L.M. (2009). Análisis de datos: métodos y ejemplos. Perdomo C.S. (Ed.). Editorial Escuela Colombiana de Ingeniería Julio Garabito, Bogotá, Colombia. 388 p.

[25] Мешта, С.; Patel, N.R. (1983). A network algorithm for performing Fisher's exact test in $r \times c$ contingency tables. J. Amer. Sta-

tist. Assoc, vol. 78, pp. 427-434

[26] Césari, M.; Rodríguez, D.; Rancan, C.; Merlino, H.; Britos, P.V.; García Martínez, R. (2008). Métodos iconográficos de observación, exploración y comunicación aplicados a la minería de textos. In X Workshop de Investigadores en Ciencias de la Computación.

[27] Díaz, S.P.; Fernández, S.P. (2004). Asociación de variables cualitativas: el test exacto de Fisher y el test de Mc Nemar. Cuadernos de atención primaria, vol. 11, n. 5, pp. 304-308.

[28] Lebart, L. (2007). DTM. Data and Text Mining, Software Estadística Exploratoria Multidimensional para datos complejos que incluyen datos numéricos y textuales. Web: <http://ses.telecom-paristech.fr/lebart/>

[29] Мешта, С.; Patel, N.R. (1983). A network algorithm for performing Fisher's exact test in $r \times c$ contingency tables. J. Amer. Statist. Assoc, v. 78, pp. 427-434

[30] Upton, G.J. (1992). Fisher's exact test. Journal of the Royal Statistical Society. Series A (Statistics in Society), pp. 395-402

[31] Addinsoft (2015). Software XLstat versión 2015, licencia para investigación. add-in de análisis estadístico para Microsoft Excel desarrollada por Addinsoft 1996-2015. <https://www.xlstat.com/es/>

[32] Césari, M.; Césari, R.; Gámbaro, A.; Arnau, E. (2013). Análisis de datos provenientes de pruebas sensoriales del vino, utilizando la lógica borrosa. VII ENIDI-Encuentro de Investigadores y Docentes de Ingeniería, ENIDI, 1-15. Los Reyunos, San Rafael. Mendoza, Argentina.

[33] Breckenkamp, J.; Neitzke, H.P.; Borkessel, C.; Berg-Beckhoff, G. (2008). Applicability of an exposure model for the determination of emissions from mobile phone base stations. Radiation protection dosimetry, 131(4), 474-481

[34] Jung M.; Lee B. (2002). Análisis de reducción de SAR para móviles. J Corea Electromag Ing Soc, v. 13, p. 2, pp. 155-163.