



UNIVERSITAT  
POLITÈCNICA  
DE VALÈNCIA

DEPARTAMENTO DE ESTADÍSTICA E INVESTIGACIÓN  
OPERATIVA

**DISEÑO Y MEJORA DE GRÁFICOS DE CONTROL  
MULTIVARIANTES PARA ATRIBUTOS. UN ENFOQUE  
BASADO EN TEORÍA DIFUSA.**

**TESIS DOCTORAL**

Presentada por:

María Nela Pastuizaca Fernández

Dirigida por:

Dr. D. Andrés Carrión García

Valencia, Abril 2016



*A la memoria de las amadas personas que no pudimos despedir  
Juanita, Lucia y la princesita Gisella*

*A mi amado y paciente esposo, Miguel,  
y mis hijos, Hugo y Mareli*



*Si hay algo verdaderamente cierto es que lo ignoro todo o casi todo. Y me da rabia, porque hubo un tiempo en el que una mente despierta podría haber adquirido todo el saber de la época. Pero ahora ya no es posible. Ya no hay más que pequeños sabios que lo saben todo sobre casi nada. Y yo soy uno de ellos.*

—Jean-Baptiste-Gabriel-Joachim Dausset. Premio Nobel de Medicina en 1980, (Toulouse, Francia, 19 de octubre de 1916 - Palma de Mallorca, España, 6 de junio de 2009).

*Si he visto más lejos ha sido porque he subido a hombros de gigantes.*

—Isaac Newton. Físico, filósofo, teólogo, inventor, alquimista y matemático inglés. (Woolsthorpe, condado de Lincolnshire, Inglaterra 4 enero de 1643 - Municipio de Kensington, Londres, Inglaterra 31 marzo de 1727 ).



## Agradecimientos

Una vez llegados a este punto, he de afirmar que me siento bendecida por la vida. Quiero dar gracias a Dios por su amoroso cuidado, brindado a través de las maravillosas personas que ha puesto en mi camino durante este trayecto y a lo largo de la vida. Personas que de un modo u otro, han contribuído a la culminación de este trabajo y a las que me gustaría, por tanto, extender mi agradecimiento.

Para empezar me gustaría agradecer de manera especial a mi tutor y director, el Dr. Andrés Carrión, por su sabia dirección en la realización de esta tesis. Sobre todo por su paciencia, por la tranquilidad transmitida y por ser, más allá de un director, un consejero, amigo y soporte en este proceso.

A Ana Debón por sus palabras de aliento, así como por su disposición y consejo oportuno. A Francisco Aparisi, por la significativa ayuda prestada siempre que lo necesité, especialmente al inicio de este trabajo. A Vicent Giner, por ser alguien con quien se puede contar cuando se necesita y por su asesoramiento en el uso de LaTeX, pues el camino hubiese sido más largo sin su ayuda.

A los compañeros del Departamento de Estadística y Optimización de la UPV que fueron delante en el camino (Nora, Adriana, Omar, Sandra y Marina), por ser ejemplo de trabajo y dedicación.

A los amigos de Valencia por acogernos y brindarnos su amistad, por hacernos sentir como en casa. A los compañeros becarios de Senescyt, con quienes compartimos alegrías y penurias y que, muchas veces, han

sido nuestra familia en Valencia.

A los amigos y a los compañeros de trabajo en Ecuador que desde la distancia y través de las redes sociales o algún mensaje de teléfono se hicieron presentes, con sus palabras de aliento, en los momentos difíciles.

A mi familia, a la que siempre he tenido presente y a la que siempre me he sentido orgullosa de pertenecer. De manera especial a mis padres, motivo de inspiración y lucha. A mis hermanas Ana, Zoila y Nancy por pertenecer a esa raza de mujeres valientes y luchadoras de la que siempre he querido formar parte. A mi hermano Jaime por ser ejemplo de lo correcto. A mi hermano Victor que aunque ya no está con nosotros dejó sembradas sus enseñanzas, que han sido luz en el camino.

A mi familia política por estar ahí, siempre que les necesitamos.

Quiero agradecer también a SENESCYT y ESPOL, por auspiciar mis estudios. Por confiar en que a mi regreso, puedo contribuir al desarrollo de mi país, mediante la formación conseguida.

Finalmente, mi agradecimiento infinito a Miguel, por ser mi compañero y ayuda incondicional. A mis hijos Hugo y Mareli por su paciencia y sacrificio. Ha sido un camino difícil y no hubiese sido posible sin su cariño y comprensión.

Valencia, abril de 2016



## Resumen

El Control Estadístico de Procesos (CEP) es un método que se utiliza para controlar las características de calidad de un producto durante el proceso de producción, determinar si los procesos de manufactura son o no estables y mejorar su capacidad a través de la reducción de la variabilidad. Una de las principales herramientas utilizadas en el Control Estadístico de Procesos es el gráfico de control.

Con frecuencia, la calidad de un producto se mide a través de varias características de calidad, generalmente correlacionadas. Los gráficos de control multivariantes son una respuesta a la necesidad de controlar la calidad en tales situaciones. Si las características de calidad son de carácter cualitativo, ocurre en ocasiones que la calidad del producto se define mediante variables lingüísticas y las unidades de producto se clasifican también de forma lingüística en varias categorías, dependiendo del grado de cumplimiento de las expectativas, creando una situación de clasificaciones difusas. Los gráficos propuestos en la literatura para tratar con tales situaciones están, en su mayoría, basados en simulación y el uso de técnicas de aproximación que dificultan la aplicación práctica de los mismos.

Esta tesis trata esta cuestión proponiendo un Gráfico de Control multivariante para características de calidad de tipo multi-atributos correlacionados basado en el gráfico  $T^2$  de Hotelling, utilizando un enfoque difuso. Se mejoran los resultados de los gráficos de control

propuestos anteriormente estableciendo una manera más formal de medición y evaluación de la calidad en estas situaciones difusas.

Se propone además un procedimiento para evaluar el rendimiento del gráfico de control propuesto mediante la determinación de la longitud de racha promedio (ARL), tanto para un estado bajo-control como para el estado fuera-de-control. Para ello se desarrollaron algoritmos que utilizan simulación de Monte Carlo y han sido implementados en R.

Adicionalmente, se analiza la sensibilidad del gráfico de control frente a la elección de las funciones de pertenencia de las variables lingüísticas.

**Palabras clave:** *Control de Calidad, Gráficos de control, Hotelling  $T^2$ , multinomial multivariante, Teoría difusa, ARL*

## Resum

El Control Estadístic de Processos (CEP) és un mètode que s'utilitza per controlar les característiques de qualitat d'un producte durant el procés de producció, determinar si els processos de manufactura són o no estables i millorar la seva capacitat a través de la reducció de la variabilitat. Una de les principals eines utilitzades en el Control Estadístic de Processos és el gràfic de control.

Sovint, la qualitat d'un producte es mesura a través de diverses característiques de qualitat, generalment correlacionades. Els gràfics de control multivariants són una resposta a la necessitat de controlar la qualitat en aquestes situacions. Si les característiques de qualitat són de caràcter qualitatiu, de vegades passa que la qualitat del producte es defineix mitjançant variables lingüístiques i les unitats de producte es classifiquen també de forma lingüística en diverses categories, depenent del grau de compliment de les expectatives, creant una situació de classificacions difuses. Els gràfics proposats en la literatura per abordar aquestes situacions són, majoritàriament, basats en simulació i l'ús de tècniques d'aproximació que en dificulten l'aplicació pràctica.

Aquesta tesi tracta de resoldre aquesta qüestió amb la proposta d'un Gràfic de Control multivariant per característiques de qualitat de tipus multi-atributs correlacionats basat en el gràfic  $T^2$  de Hotelling, mitjançant un enfocament difús. S'hi milloren els resultats de les gràfics de control

proposats anteriorment per mitjà d'un mètode més formal de mesurament i avaluació de la qualitat en aquestes situacions difuses.

S'hi proposa a més un procediment per avaluar el rendiment del gràfic de control proposat mitjançant la determinació de la longitud de ràfega mitjana (ARL), tant per a un estat en-control com per a l'estat fora-de-control. Amb aquesta finalitat es van desenvolupar algorismes que utilitzen simulació de Monte Carlo i han estat implementats en R.

Adicionalment, s'hi analitza la sensibilitat del gràfic de control davant l'elecció de les funcions de pertinença de les variables lingüístiques.

**Paraules clau:** *control de qualitat, gràfics de control, Hotelling  $T^2$ , multinomial multivariant, teoria difusa, longitud mitjana de ratxa.*

## Abstract

The Statistical Process Control (SPC) is a method used to control the quality characteristics of a product during the production process, determine whether the manufacturing process is or not stable and improve its capacity through the reduction of variability. One of the main tools used in the SPC is the control chart.

Often the quality of a product is measured through various quality characteristics generally correlated. Multivariate Control charts are a response to the need for quality control in such situations. If the quality characteristics are qualitative, sometimes it happens that the product quality is defined by linguistic variables and product units are also classified by linguistic forms into several categories, depending on the degree of fulfillment of expectations, creating a situation of fuzzy classifications. The control charts proposed in the literature to deal with such situations are mostly based on simulation and using approximation techniques which hinder the practical application thereof.

This thesis addresses this issue proposing a multivariate control chart for quality characteristics of multi-type attributes correlated based on the  $T^2$  control chart of Hotelling, using a fuzzy approach. The results of the proposed control charts before are improved by establishing a more formal way of measuring and evaluating quality in these diffuse situations .

A method is also proposed to assess the performance of control chart proposed, by determining the average run length (ARL), in both in-control state and the out-of-control state . For this, algorithms which use Monte Carlo simulation are developed and implemented in R.

Additionally, the sensitivity of the control chart faced with the choice of the membership functions of linguistic variables is analyzed.

**Keywords:** *quality control, control charts, Hotelling  $T^2$ , multinomial multivariate, fuzzy theory, average run length.*

# Índice general

<b>1. Introducción</b>	<b>1</b>
1.1. Consideraciones iniciales . . . . .	1
1.2. Planteamiento del problema . . . . .	2
1.3. Objetivos planteados . . . . .	4
1.3.1. Objetivo General . . . . .	4
1.3.2. Objetivos de la investigación . . . . .	4
1.4. Estructura del documento . . . . .	5
<b>2. Marco conceptual</b>	<b>7</b>
2.1. Introducción . . . . .	7
2.2. Control Estadístico de Procesos y Gráficos de Control . . .	8
2.2.1. Causas comunes y asignables de la variación de la calidad . . . . .	9
2.2.2. Gráficos de Control . . . . .	11
2.2.3. Gráficos de control para variables multinomiales. .	24
2.3. Teoría Difusa y gráficos de Control . . . . .	27

2.3.1.	Antecedentes . . . . .	27
2.3.2.	Variables Lingüísticas y Conjuntos Difusos . . . . .	31
2.3.3.	Valor representativo de un subconjunto difuso . . . . .	36
2.3.4.	Representación de una muestra . . . . .	39
2.3.5.	Algunos gráficos de control con enfoque difuso . . . . .	47
2.4.	Gráficos de control para variables multinomiales con enfoque difuso . . . . .	48
2.4.1.	Introducción . . . . .	48
2.4.2.	Gráficos de control con enfoque difuso para variables multinomiales univariantes. . . . .	48
2.4.3.	Gráficos de control con enfoque difuso para variables multinomiales multivariantes. . . . .	51
<b>3.</b>	<b>Gráfico de Control Multinomial Multivariante con enfoque difuso</b>	<b>57</b>
3.1.	Introducción . . . . .	57
3.2.	Generalidades . . . . .	59
3.3.	Modelo propuesto . . . . .	60
3.3.1.	Estimación de los parámetros $\mu_R$ y $\Sigma_R$ . . . . .	68
3.3.2.	Interpretación de señales de <i>fuera-de-control</i> . . . . .	72
3.4.	Ejemplos de Aplicación . . . . .	72
3.4.1.	Ejemplo 1: Proceso de producción de comida congelada . . . . .	72
3.4.2.	Ejemplo 2: Proceso de producción de porcelana . . . . .	79
3.5.	Discusión . . . . .	86
3.5.1.	Ejemplo 1 . . . . .	86



3.5.2. Ejemplo 2 . . . . .	87
3.6. Conclusiones . . . . .	88
<b>4. Rendimiento del gráfico de control difuso multinomial multiva-</b>	
<b>riante</b>	<b>89</b>
4.1. Introducción . . . . .	89
4.2. Propuesta . . . . .	92
4.3. Análisis del rendimiento: estado bajo-control . . . . .	95
4.4. Análisis del rendimiento: estado fuera-de-control . . . . .	100
4.5. Análisis comparativo del rendimiento . . . . .	104
4.6. Conclusiones . . . . .	106
<b>5. Análisis de sensibilidad</b>	<b>109</b>
5.1. Introducción . . . . .	109
5.1.1. Grado de borrosidad de un subconjunto difuso . . . . .	110
5.1.2. Curvas Características de Operación. . . . .	115
5.2. Sensibilidad de los resultados del gráfico de control $T^2$ MMD frente al grado de borrosidad de las funciones de pertenencia. . . . .	116
5.2.1. Ejemplo 1: Proceso de producción de comida congelada . . . . .	116
5.2.2. Ejemplo 2: Proceso de producción de porcelana . . . . .	123
5.3. Sensibilidad del rendimiento del gráfico de control $T^2$ MMD frente al grado de borrosidad de las funciones de pertenencia.	130
5.4. Conclusiones . . . . .	138
<b>6. Conclusiones y líneas futuras de investigación</b>	<b>141</b>

6.1. Conclusiones . . . . .	141
6.2. Líneas futuras de investigación . . . . .	144
<b>Anexos</b>	<b>147</b>
<b>A. Algoritmo Generación de variables Multinomiales</b>	<b>149</b>
<b>B. Algoritmo para determinación de ARL mediante simulación</b>	<b>155</b>
<b>Bibliografía</b>	<b>161</b>

## Índice de figuras

2.1. Gráfico de Control típico . . . . .	12
2.2. Tipos de gráficos de control univariantes, (Campos (2012))	16
2.3. Funciones de pertenencia más empleadas: (a) Triangular, (b) Trapezoidal, (c) Gaussiana y (d) Sigmoid . . . . .	32
2.4. Funciones de pertenencia de tres valores lingüísticos: No obvio, obvio, muy obvio, (Wang and Raz (1990)) . . . . .	33
2.5. Función de pertenencia y valores representativos del con- junto difuso $A$ (Wang y Raz (1990)) . . . . .	38
2.6. Funciones de pertenencia para un conjunto de cinco térmi- nos lingüísticos (Wang y Raz (1990)) . . . . .	41
2.7. Función de pertenencia resultante de la muestra $S$ . (Wang y Raz(1990)) . . . . .	43
3.1. Función de pertenencia para un número triangular difuso	62
3.2. Funciones de pertenencia (Taleb et al.(2006)) . . . . .	74
3.3. Gráfico de Control $T^2$ MMD Fase I . . . . .	77
3.4. Gráfico de Control $T^2$ MMD Fase II . . . . .	78

3.5. Gráfico de Control $T^2$ MMD Fase I: Proceso de producción de porcelana . . . . .	83
3.6. Gráfico de Control $T^2$ MMD Fase II: Proceso de producción de porcelana . . . . .	84
3.7. (a) Gráfico de Control estadístico $T_f^2$ (Taleb et al. (2006)), (b) Gráfico de Control $T^2$ MMD Fase I . . . . .	86
3.8. (a) Gráfico de Control estadístico $T_f^2$ para muestras adicionales (Taleb et al. (2006)), (b) Gráfico de Control $T^2$ MMD Fase II . . . . .	87
4.1. Rendimiento estado fuera-de-control: (a) $\alpha = 0.0027$ ; (b) $\alpha = 0.005$ ; (c) $\alpha = 0.005$ ; (d) $\alpha = 0.05$ . P-C: caso parámetros-conocidos; L-T: caso límite teórico; L-C: caso límite corregido.	102
4.2. Comparativo ARL vs Distancia para diferentes valores de $\alpha$	103
4.3. Comparativo ARL vs Distancia para Gráfico de Control $T^2$ MMD y gráfico MFQCC: (a) $\alpha = 0.05$ , (b) $\alpha = 0.0027$ . .	106
5.1. Funciones de pertenencia para A y su conjunto más cercano	113
5.2. Índice de borrosidad . . . . .	114
5.3. Funciones de pertenencia con grado de borrosidad 0.125 .	117
5.4. Comparativo de Gráficos de Control de Fase I para funciones de pertenencia con diferentes grados de borrosidad .	120
5.5. Comparativo de Gráficos de Control de Fase II para funciones de pertenencia con diferentes grados de borrosidad .	121
5.6. Estadístico $T^2$ de Fase I para funciones de pertenencia con diferentes grados de borrosidad $\alpha = 0.027$ . . . . .	122
5.7. Estadístico $T^2$ de Fase II para funciones de pertenencia con diferentes grados de borrosidad $\alpha = 0.027$ . . . . .	124

5.8. Comparativo de Gráficos de control de Fase I para funciones de pertenencia con diferentes grados de borrosidad: Proceso de producción de porcelana . . . . . 127

5.9. Comparativo Gráficos de control de Fase II para diferentes grados de borrosidad: Proceso de producción de porcelana 128

5.10. Estadístico  $T^2$  de Fase I para diferentes grados de borrosidad  $\alpha = 0.027$ : Proceso de producción de porcelana . . . . 129

5.11. Estadístico  $T^2$  de Fase II para diferentes grados de borrosidad  $\alpha = 0.027$ : Proceso de producción de porcelana . . . . 130

5.12. Curvas OC para diferentes grados de borrosidad ( $m = 50$ ) 132

5.13. Gráfico de superficie: Error tipo II vs Grado de Borrosidad y Distancia Mahalanobis ( $m = 50$ ) . . . . . 133

5.14. Gráfico de contorno: Error tipo II vs Grado de Borrosidad y Distancia Mahalanobis ( $m = 50$ ) . . . . . 133

5.15. Curvas OC para diferentes Grados de Borrosidad ( $m = 100$ ) 134

5.16. Gráfico de superficie: Error tipo II vs Grado de Borrosidad y Distancia Mahalanobis ( $m = 100$ ) . . . . . 135

5.17. Gráfico de contorno: Error tipo II vs Grado de borrosidad y Distancia Mahalanobis  $m = 100$  . . . . . 135

5.18. Curvas OC para diferentes Grados de Borrosidad ( $m = 200$ ) 136

5.19. Gráfico de superficie: Error tipo II vs Grado de Borrosidad y Distancia Mahalanobis ( $m = 200$ ) . . . . . 137

5.20. Gráfico de contorno: Error tipo II vs Grado de borrosidad y Distancia Mahalanobis ( $m = 200$ ) . . . . . 137



## Índice de tablas

2.1. Funciones de probabilidad y de pertenencia asociados a una unidad con $x$ rasguños (Raz y Wang (1990)) . . . . .	35
3.1. Tabla de datos para el análisis preliminar . . . . .	69
3.2. Datos del proceso de producción de comida congelada (Taleb et al.(2006)) . . . . .	75
3.3. Valores representativos y estadístico $T^2$ . . . . .	76
3.4. Muestras adicionales (Taleb et al.(2006)), valores representativos y estadísticos $T^2$ correspondientes . . . . .	78
3.5. Valores $d_i$ para la interpretación de señales de <i>fuera-de-control</i> . . . . .	79
3.6. Datos del proceso de producción de porcelana (estimados a partir de Taleb (2009)), valores representativos y estadísticos $T^2$ correspondientes . . . . .	82
3.7. Muestras adicionales (Taleb (2009)), valores representativos y estadísticos $T^2$ correspondientes . . . . .	84
3.8. Valores $d_i$ para la interpretación de señales de fuera-de-control para el proceso de producción de porcelana . . . . .	85

3.9. Estadísticos $T_f^2$ (Taleb et al.(2006)) y $T^2$ para muestras adicionales . . . . .	87
4.1. Datos simulados para la Fase I . . . . .	96
4.2. Rendimiento estado bajo-control del Gráfico de Control $T^2$ MMD con parámetros estimados a partir de 50 muestras . . . . .	99
4.3. Cambios en el Vector de Proporciones propuestos por Taleb et al.(2006) y distancia de Mahalanobis . . . . .	100
4.4. ARL vs cambios propuestos, para los Gráficos de Control $T^2$ MMD y MFQCC $\alpha = 0.05$ . . . . .	105
4.5. ARL vs cambios propuestos, para los Gráfico de Control $T^2$ MMD y MFQCC $\alpha = 0.0027$ . . . . .	105
5.1. Estadístico $T^2$ para funciones de pertenencia con diferentes Grados de Borrosidad: Proceso de producción de comida congelada . . . . .	119
5.2. Estadístico $T^2$ para diferentes Grados de Borrosidad: Proceso de producción de porcelana . . . . .	126
5.3. Error de tipo II vs Distancia de Mahalanobis para funciones de pertenencia con diferentes grados de borrosidad . . . . .	132
A.1. muestras simuladas . . . . .	152



## Introducción

### 1.1. Consideraciones iniciales

El constante desarrollo del ámbito empresarial e industrial y la búsqueda de la calidad de los bienes y servicios producidos (hoy por hoy un requisito indispensable para sobrevivir frente a la competencia); ha creado la necesidad del uso de herramientas estadísticas para el control, la gestión y la mejora de la producción.

Para que sea controlable, evaluable y mejorable, la calidad tiene que ser medida. En ese sentido los conceptos del control estadístico de procesos han sido muy importantes en la industria manufacturera. Su principal objetivo es monitorear el desempeño de un proceso en el tiempo para que el proceso logre un estado de control estadístico. Tal estado de control se dice que existe si las características de calidad permanecen estables y únicamente hay presentes causas comunes de variación.

Los gráficos de control desarrollados por el Dr. W. Shewart, considerado como el padre del control estadístico de procesos, son una

de las herramientas más usadas. Éstos proporcionan un buen método para identificar la existencia de variabilidad adicional debido a causas asignables y permiten detectar de forma temprana la presencia de eventos que pueden producir desajustes en los procesos evitando así la producción de productos defectuosos. Deben su popularidad a la relativa sencillez con que puede ser implementados; e interpretados incluso por personas no especializadas. Han sido objeto de un amplio estudio y han servido como base para el desarrollo de diversos nuevos gráficos de control diseñados de acuerdo a la complejidad de los procesos y/o variables que se quieran controlar.

La presente tesis se enmarca en esta línea de investigación, al adentrarse en el análisis y definición de unos gráficos de control adaptados a situaciones diferentes a las de los gráficos originales. En concreto se va trabajar en gráficos por atributos, con consideración de múltiples características de calidad, cada una de ellas siguiendo un modelo multinomial en el que las diferentes categorías se definen a través de expresiones lingüísticas, bajo una categorización difusa.

## **1.2. Planteamiento del problema**

La calidad, por su propia naturaleza, es en sí subjetiva y puede dar lugar a una multiplicidad de significados ya que es altamente dependiente de la cognición humana. Por lo tanto, puede ser apropiado considerar la calidad en términos de grados de conformidad en lugar de conformidad y no conformidad, especialmente en situaciones en las que las características de calidad de un producto son cualitativas.

La teoría de conjuntos difusos soporta la subjetividad del lenguaje natural, utilizada para describir la calidad y le proporciona una metodología

para ser utilizado en el proceso de modelado (Guiffrida y Nagi, 1998).

Cuando un trabajo de investigación debe incorporar la incertidumbre en el análisis de decisiones, se lleva a cabo básicamente a través de la teoría de la probabilidad y/o la teoría de conjuntos difusos. La primera representa la naturaleza estocástica de análisis de decisión mientras que la segunda captura la subjetividad de la conducta humana. Un enfoque racional hacia la toma de decisiones debe tomar en cuenta la subjetividad humana. La teoría de conjuntos difusos es un medio perfecto para el modelado de la incertidumbre (o imprecisión) derivada de fenómenos mentales que no son aleatorios ni estocásticos.

Cuando la subjetividad humana juega un papel importante en la definición de las características de calidad, los gráficos de control clásicos pueden no ser aplicables, puesto que requieren información precisa. El uso de gráficos de control difusos es interesante cuando los datos estadísticos en consideración son inciertos o vagos, o la información disponible sobre el proceso incluye subjetividad humana.

En ocasiones, cuando las características de calidad son cualitativas, ocurre que la calidad del producto debe ser definida mediante variables lingüísticas que utilizan medidas subjetivas como la calificación en una escala. Así, cuando se requiere medir características de calidad tales como la apariencia, la suavidad, el aroma, etc., que no se pueden expresar numéricamente, se usa una clasificación en varias categorías que se asocian con términos lingüísticos tales como muy buena, buena, regular, mala, etc. en función del grado de cumplimiento de las expectativas. Se crea entonces una situación de clasificaciones difusas.

Se trata pues de variables que presentan varios valores, cada uno fruto de una calificación expresada en términos lingüísticos, resultando en una distribución multinomial. Para monitorear este tipo de variables,

se han propuesto varios gráficos de control, tanto desde el punto de vista probabilístico como mediante el uso de la teoría difusa, especialmente en el caso univariante. Sin embargo en el caso multivariante es escaso el trabajo realizado. De aquí la necesidad de la propuesta planteada en la presente tesis doctoral, que es el diseño de un gráfico de control multivariante para variables multinomiales, usando el enfoque difuso para incorporar ese elemento de subjetividad arriba mencionado.

### **1.3. Objetivos planteados**

#### **1.3.1. Objetivo General**

Proponer y caracterizar un gráfico de control para características de calidad tipo atributo que sigan un modelo multinomial multivariante, considerando un enfoque difuso en la definición de las categorías o valores de esas variables.

#### **1.3.2. Objetivos de la investigación**

- Justificar el uso de la teoría difusa para el manejo de características de calidad caracterizadas en forma lingüística.
- Analizar y comparar los distintos gráficos y metodologías propuestas en la literatura existente, para identificar los elementos que permitan el diseño del gráfico de control difuso multinomial para atributos multivariantes.
- Proponer un gráfico de control capaz de trabajar en una situación multinomial multivariante con enfoque difuso.

- Establecer una metodología que permita medir el rendimiento del gráfico de control propuesto.
- Analizar la sensibilidad del rendimiento del gráfico de control propuesto frente al grado de borrosidad de las funciones de pertenencia empleadas en la construcción del gráfico.

## 1.4. Estructura del documento

La tesis se desarrolla en siete capítulos, en el capítulo 1, se realiza la introducción al tema incorporando además el planteamiento del problema los objetivos de la investigación y el alcance de la misma.

En el capítulo 2 se revisan las principales definiciones del control estadístico de procesos y los gráficos de control, así como las definiciones relativas a la teoría difusa y su uso en el diseño de gráficos de control. Se revisan además las principales contribuciones en cuanto a su desarrollo. Estas definiciones y los gráficos planteados servirán como base para la propuesta que se quiere realizar.

En el capítulo 3 de la tesis nos planteamos desarrollar el modelo del gráfico de control con enfoque difuso para variables multinomiales multivariantes y explorar su aplicación utilizando dos ejemplos propuestos en la literatura, se realiza una comparación de los resultados con los obtenidos mediante la aplicación de procedimientos planteados por otros autores.

En el capítulo 4, se presenta un procedimiento para la determinación del rendimiento del gráfico de control propuesto así como también se realiza un análisis de sus características usando como referencia el estudio hecho por [Champ et al. \(2005\)](#).

En el capítulo 5, se analiza la sensibilidad del gráfico de control propuesto frente al grado de borrosidad de las funciones de pertenencia a utilizar en el desarrollo del mismo. Se estudia la sensibilidad de los resultados y el rendimiento.

Finalmente, en el capítulo 6 se presentan las conclusiones que se derivan de todo el trabajo realizado y se proponen líneas futuras de investigación.

## Marco conceptual

### 2.1. Introducción

“Los métodos estadísticos desempeñan un papel vital en el control y la mejora de la calidad. Ellos proporcionan el principal medio por el cual un producto es muestreado, probado y evaluado, y la información en esos datos se utiliza para controlar y mejorar el proceso y el producto.”  
(Montgomery, 2009)

Uno de los principales objetivos de un producto es satisfacer o superar las expectativas del cliente, para ello debe ser producido por un proceso que sea estable o repetible. Más precisamente, el proceso debe ser capaz de funcionar con poca variabilidad alrededor del objetivo o valor nominal de la o las características de calidad del producto.

El Control Estadístico de Procesos (CEP) es un potente conjunto de herramientas de resolución de problemas útiles en la consecución de la

estabilidad del proceso y en la mejora de su capacidad a través de la reducción de la variabilidad. Es fácil de usar, tiene un impacto significativo, y se puede aplicar a cualquier proceso. Una de sus herramientas, el gráfico de control Shewhart es probablemente la más ampliamente utilizada. Fue desarrollado en 1920 por Walter A. Shewhart y ha sido ampliamente estudiado.

En este capítulo se presentan algunas definiciones básicas del control estadístico de procesos y los gráficos de control, con el fin de establecer el escenario en el que se desarrolla el trabajo de investigación realizado.

Se revisan además los conceptos clave de la teoría de conjuntos difusos en la construcción de gráficos de control de atributos. Estos gráficos se utilizan para controlar características de calidad disponibles en forma lingüística. El objetivo de esta revisión es introducir la nomenclatura y la notación que posteriormente será utilizada.

Finalmente se define el método de simulación que se utilizará para la medición del rendimiento y el análisis de sensibilidad del gráfico de control propuesto.

## **2.2. Control Estadístico de Procesos y Gráficos de Control**

El propósito del control de calidad es asegurar que los procesos se realicen de una manera aceptable, esto se logra mediante el monitoreo de sus salidas con técnicas estadísticas .<sup>1</sup>

---

<sup>1</sup>La mayor parte de las definiciones y conceptos expuestos en esta sección se han extraído de [Montgomery \(2009\)](#).



En el marco del control de calidad, el Control Estadístico de Procesos (CEP) se define como “un conjunto poderoso de herramientas para resolver problemas, muy útil para conseguir la estabilidad y mejorar la capacidad del proceso mediante la reducción de la variabilidad” (Montgomery, 2009), o como “la aplicación de técnicas estadísticas para medir y analizar la variación de los procesos” (Juran y Gryna, 1988). Es un método que se utiliza para controlar las características de calidad de un producto durante el proceso de producción, determinar si los procesos de manufactura son o no estables y mejorar su capacidad a través de la reducción de la variabilidad. Una de las principales herramientas utilizadas en el Control Estadístico de Procesos es el gráfico de control.

Para entender los conceptos estadísticos que forman la base del CEP, debemos describir primero la teoría de la variabilidad de Shewhart.

### 2.2.1. Causas comunes y asignables de la variación de la calidad

Cuando tenemos un proceso de producción por lo general hay un valor objetivo. Queremos que nuestro proceso logre este objetivo para cada producto. Sin embargo, en cada proceso hay una variabilidad aleatoria inherente. Por lo tanto, no importa lo bien que diseñamos todo el procedimiento o la forma precisa en que nuestra maquinaria sea, esperamos estar cerca del valor objetivo, aunque no siempre en este valor. La existencia de esta variabilidad afecta a nuestro proceso.

Hay dos tipos diferentes de esta variabilidad:

**Variabilidad debida a una causa común** . Es la variabilidad natural inherente al proceso. En cualquier proceso de producción, independientemente de lo bien diseñado o cuidadosamente mantenido que

sea, un cierto grado de variabilidad inherente o natural siempre existirá. Esta variabilidad natural o “ruido de fondo” es el efecto acumulativo de muchas causas pequeñas, esencialmente inevitables. Su existencia se debe al azar así que podemos encontrar variabilidad puramente aleatoria de un producto a otro. Un proceso que funciona sólo con variabilidad debido a causas comunes se dice que está bajo control.

**Variabilidad debida a una causa asignable o especial** . Es la variabilidad que se da como resultado de factores que no son puramente aleatorios. Esta variabilidad en las características clave de calidad por lo general proviene de tres fuentes: máquinas mal ajustados o controladas, errores del operador, o materia prima defectuosa. Tal variabilidad es generalmente grande cuando se compara con el ruido de fondo, y por lo general representa un nivel inaceptable de rendimiento del proceso, causan heterogeneidad y como resultado lo afectan, dando lugar a productos de baja calidad. Un proceso que opera en la presencia de causas especiales de variabilidad se dice que esta fuera-de-control.

Uno objetivo principal de control estadístico de procesos es detectar rápidamente la aparición de causas asignables de cambios en el proceso, de ese modo se podrían tomar medidas correctivas antes de que se fabriquen muchas unidades no conformes. El gráfico de control es una técnica de proceso de monitoreo en línea ampliamente utilizada para este propósito, dándonos la capacidad de eliminar el efecto de los cambios sufridos por el proceso y por lo tanto reducir la variabilidad total. Como resultado, la eliminación de las causas especiales conduce a una mejora de la calidad del producto. Es posible que no se pueda eliminar la variabilidad por completo, pero el gráfico de control es una herramienta eficaz para reducir la variabilidad tanto como sea posible.

### 2.2.2. Gráficos de Control

El gráfico de control de Shewhart fue creado en 1924 y revela el estado de un proceso mediante la representación de los datos producidos en el mismo. Su poder reside en su capacidad para detectar cambios e identificar las causas especiales de la variabilidad en un proceso. Esto hace posible el diagnóstico y la corrección de muchos problemas de producción y con frecuencia reduce las pérdidas y aporta mejoras sustanciales en la calidad del producto.

Un gráfico de control es una representación gráfica de una característica del proceso bajo observación. También pueden ser utilizados para estimar los parámetros de un proceso de producción, y, a través de esta información, determinar la capacidad del proceso. Adicionalmente puede proporcionar información útil para mejorar el proceso.

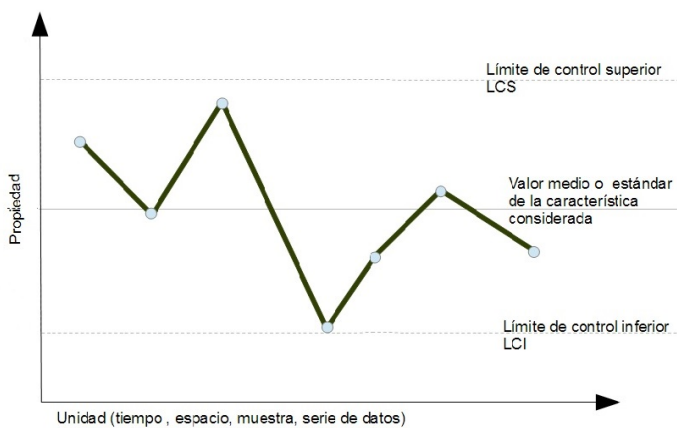
#### Principios básicos

El gráfico de control es una representación gráfica de una característica de calidad que ha sido medida o calculada a partir de una muestra en comparación con el número de muestra o el tiempo. En el eje horizontal tenemos el número de la muestra tomada del proceso o el tiempo en que la muestra se inspeccionó, en el eje vertical tenemos el valor del estadístico empleado para monitorizar la característica medida para cada muestra (o para el tiempo) del eje horizontal. Una línea recta conecta los puntos sucesivos que indican el nivel del estadístico en el tiempo o en muestras sucesivas. También consta de tres líneas generalmente horizontales, una línea central (LC) que representa el valor medio de la característica de calidad que corresponde al estado bajo-control (Es decir, sólo causas comunes están presentes) y otras dos líneas horizontales, llamadas el límite de control superior (LCS) y el límite de control inferior (LCI). Estos

límites de control se eligen de manera que si el proceso está bajo control, casi la totalidad de los puntos de muestra caerá entre ellos e indican el intervalo de variabilidad normal (Rowlands y Wang, 2000). Un ejemplo de un gráfico de control típico se da en la figura 2.1.

Los límites de control superior e inferior son los valores máximo y mínimo para poder considerar “bajo control” al estadístico empleado, mientras que los puntos de datos que caen fuera de los límites indican una posible condición de fuera-de-control del proceso.

La construcción de un gráfico de control implica la determinación de los valores a ser utilizados para la línea central y los límites de control. Estos se calculan generalmente a partir de los datos de muestras recogidos de un proceso de producción.



**Figura 2.1:** Gráfico de Control típico

En los gráficos de Shewhart usualmente se fijan los límites de control usando el criterio  $3\sigma$ . Los datos que se trazan en los gráficos de control se obtienen directamente del proceso. El enfoque principal del gráfico de control es el intento de separar las causas especiales o asignables de

variación, de las causas comunes de variación. Si el proceso es encontrado en un estado fuera-de-control se debe tomar medidas correctivas con el fin de eliminar la causa asignable que lo llevó a ese estado.

Podemos dar un modelo general para un gráfico de control de la media. Sea  $w$  una variable estadística que mide alguna característica de la calidad de interés, y supongamos que la media de  $w$  es  $\mu_w$  y la desviación estándar de  $w$  es  $\sigma_w$ . Entonces la línea central, el límite superior de control, y el límite de control inferior están dados por:

$$LCS = \mu_w + k\sigma_w$$

$$LC = \mu_w$$

$$LCI = \mu_w - k\sigma_w$$

donde  $k$  es la “distancia” de los límites de control de la línea central, expresado en unidades de desviación estándar. Esta teoría general de los gráficos de control fue propuesta por primera vez por Walter A. Shewhart, y los gráficos de control desarrollados de acuerdo a estos principios son a menudo llamados *gráficos de control de Shewhart*.

### Fases del gráfico de control

En la literatura, se han discutido dos fases distintas para la aplicación práctica de gráficos de control (véase, por ejemplo [Woodall \(2000\)](#)).

*Fase I*, en esta fase los gráficos se utilizan de forma retroactiva probando si el proceso estaba en control cuando se estaban elaborando los primeros subgrupos. En esta fase, los gráficos se utilizan como ayuda para el operador, para llevar un proceso a un estado bajo-control estadístico.

Una vez logrado esto, el gráfico de control se utiliza para definir qué se entiende por control estadístico. Durante esta fase, se estudia el proceso de forma muy intensa. Los datos recogidos a continuación, se analizan en un intento de responder a la pregunta "¿han sido los datos obtenidos a partir de un proceso bajo control?"

*Fase II*, en esta fase los gráficos de control se utilizan para probar si el proceso se mantiene bajo control cuando se evalúan subgrupos futuros. Es decir, los gráficos se utilizan para monitorear el proceso y detectar algún cambio del estado "bajo control". En cada etapa de muestreo, el operador hace la pregunta "¿Ha cambiado el estado del proceso?". El significado de bajo control, en esta fase, generalmente se determina por los valores de los parámetros del proceso por ejemplo, la media y la desviación estándar para variables univariadas de distribución continua. Los valores de los parámetros o bien se dan al operador o se estiman a partir de los datos históricos de la Fase I obtenidos cuando el proceso es considerado bajo control. Téngase en cuenta que en esta fase los datos no pueden ser considerados como tomados de un proceso bajo control a menos que esos mismos datos proporcionen evidencia de que no ha habido cambios en el proceso. A menudo se requiere la comprensión y mejora de procesos en la transición de la Fase I a la Fase II.

### **Tipos de gráficos de control**

Los gráficos de control se pueden clasificar en dos tipos generales dependiendo de la naturaleza de la característica a controlar.

**Gráfico de control por variables** Este tipo de gráfico se utiliza cuando la característica de calidad se puede medir y se expresa como un número en una escala continua de la medición (pesos, dimensiones, etc.). En tales casos, es conveniente describir la característica de

calidad con una medida de tendencia central y una medida de la variabilidad. El gráfico  $\bar{x}$  es el gráfico más ampliamente utilizado para el control de tendencia central, mientras que los gráficos basados ya sea en el rango de la muestra o en su desviación estándar se utilizan para controlar la variabilidad del proceso.

**Gráfico de control por atributos** Se utilizan cuando las características de calidad no se miden en una escala continua ni en una escala cuantitativa. En estos casos, podemos juzgar a cada unidad de producto, ya sea conforme o no conforme sobre la base de si posee o no ciertos atributos, o podemos contar el número de no conformidades (defectos) que aparecen en una unidad de producto.

Hay que añadir que tanto los Gráficos de control por variables como los de control por atributos pueden a su vez clasificarse como univariantes o multivariantes de acuerdo al número de características que se quieran controlar en un producto.

La figura 2.2 muestra un esquema no exhaustivo de los tipos de gráficos de control univariante estructurados de acuerdo al tipo de variable controlada y a la magnitud del cambio que se desea detectar propuesto por Campos (2012).

### **Evolución de lo gráficos de control**

Después de la aportación realizada por Shewhart y a partir del gráfico propuesto por él, surgieron gráficos de control que procuraron cubrir las limitaciones existentes en los gráficos de Shewhart como el hecho de tomar de decisiones considerando solo la última muestra obtenida, perdiendo la información provista por las muestras anteriores. De tal forma que el enfoque principal de estos nuevos gráficos fue considerar

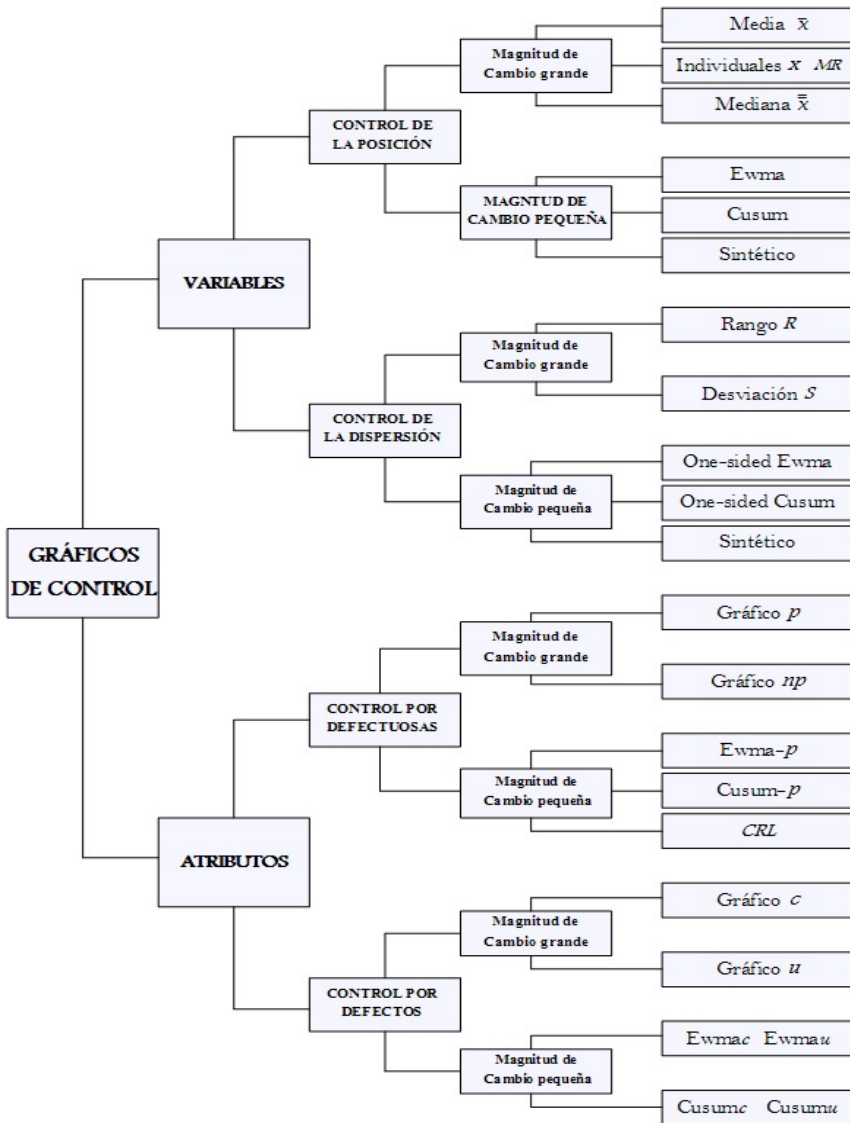


Figura 2.2: Tipos de gráficos de control univariantes, (Campos (2012))



información histórica para detectar pequeños desplazamientos en el proceso y desarrollar un gráfico de control para varias características de calidad a la vez.

A continuación detallaremos los gráficos de control tradicionales y algunas de las principales aportaciones desarrolladas a partir de éstos

### Gráficos tradicionales para el control de variables

El llamado gráfico  $\bar{X}$ - $R$ , uno de los más utilizados, toma muestras de tamaño constante  $n$  a intervalos regulares, y para cada una de ellas calcula y traslada a gráficos dibujados en paralelo, los estadísticos  $\bar{x}$  y  $R$  (media y rango, respectivamente), donde

$$\bar{x} = \frac{x_1 + \cdots + x_n}{n} \quad \text{y}$$

$$R = x_{\text{máx}} - x_{\text{mín}},$$

siendo  $\{x_1, \dots, x_n\}$  el conjunto de  $n$  observaciones consecutivas de la característica de calidad  $X$  que interesa controlar, que en el caso de este gráfico sería una característica numérica continua.

El *doble gráfico*  $\bar{x}$ - $R$  se utiliza para controlar simultáneamente la media y la variabilidad de la característica de calidad objeto de interés. Los límites de control para los respectivos estadísticos,  $\bar{x}$  y  $R$ , son calculados a partir de las propiedades estadísticas conocidas de los mismos, suponiendo la normalidad de la variable subyacente  $X$ .

Otros gráficos clásicos para el control de variables son: el *gráfico*  $\bar{X}$ - $S$  (que es tradicionalmente preferido cuando  $n > 10$  o bien  $n$  es variable),

el gráfico *de medianas* (que en ocasiones puede ser recomendable por su sencillez y claridad) y el *gráfico de Shewhart para medidas individuales* (de aplicación en situaciones en las que el tamaño muestral  $n$  se limita a 1 observación).

### Gráficos tradicionales para el control de atributos

Los gráficos clásicos de control para *atributos* son: los *gráficos  $p$  y  $np$*  (utilizados para el control del número de unidades defectuosas, grafican, la proporción observada  $\hat{p}$  de unidades defectuosas y la cantidad de unidades defectuosas  $n\hat{p}$  en cada muestra respectivamente), y los *gráficos  $u$  y  $c$*  (para el control del número de defectos en la muestra, representan la tasa y la cantidad de defectos respectivamente). En el caso de los gráficos  $p$  y  $u$ , las muestras que dan lugar a cada punto del gráfico no tienen que ser necesariamente del mismo tamaño.

En el enfoque clásico de Shewart **pendiente citar a Shewart**, los límites de control se calculan suponiendo un modelo probabilístico binomial, para el caso de los gráficos  $p$  y  $np$ , y un proceso de Poisson para los gráficos  $u$  y  $c$ , aplicando en todos los casos aproximaciones por la distribución normal.

Woodall (1997) realiza una exhaustiva revisión de los gráficos para el control de atributos, incluyendo tanto tradicionales como avanzados.

### Gráficos de valores acumulados y ponderados

Los gráficos de sumas acumuladas *Cumulative Sum* (CUSUM) y de medias móviles ponderadas exponencialmente *Exponentially Weighed*

*Moving Average* (EWMA) surgieron posteriormente y mejoran el comportamiento de los gráficos de Shewhart. Los gráficos CUSUM y EWMA acumulan la información de las muestras anteriores en cada punto representado, a diferencia de los gráficos Shewart que tienen en cuenta únicamente la información de la última muestra graficada para definir una señal de fuera-de-control. Se logra así que los gráficos sean sensibles a pequeños cambios, así como a cambios lentos y progresivos

En concreto, el gráfico CUSUM (Page, 1961) representa la suma de las desviaciones de cada punto que se graficaría en el gráfico de Shewhart correspondiente, respecto del valor objetivo. Es decir, en el gráfico CUSUM para la media del proceso, lo que se representa para la muestra  $i$ -ésima es el valor  $C_i = \sum_{h=1}^i (\bar{x}_h - \mu_0) = C_{i-1} + (\bar{x}_i - \mu_0)$ , siendo  $\mu_0$  el valor objetivo y esperado de  $\bar{X}$ .

El valor central del gráfico alrededor del cual deberían situarse los valores calculados  $C_i$  en el caso de que el proceso opere "bajo-control" es cero. Puede construirse una *versión CUSUM* de la mayor parte de los gráficos tradicionales ( $\bar{X}$ ,  $R$ ,  $S$ ,  $np$ ,  $c$ ).

El gráfico EWMA (Hunter, 1986) incorpora en cada punto representado  $z_i$  información de las muestras anteriores a través de la expresión  $z_i = \lambda \bar{x}_i + (1 - \lambda)z_{i-1}$ , con  $z_0 = \mu_0$  y siendo  $\lambda$  un parámetro constante tal que  $0 < \lambda \leq 1$ ; esto en el caso del gráfico de control para  $\bar{X}$ , y de manera similar para otros parámetros que se quieran controlar.

Ambos gráficos, CUSUM y EWMA, se utilizan frecuentemente sobre medidas individuales (muestras de tamaño  $n = 1$ ).

## Otros Gráficos de control

Existen variantes de los gráficos clásicos para variables y atributos entre los que podemos mencionar:

*Gráficos para ciclos de producción cortos*, usados en el caso de sistemas con series de producción cortas donde, por tener muestra insuficientes, no se pueden estimar los parámetros del proceso ni establecer límites de control válidos a partir de dichas estimaciones de la forma en que requiere un gráfico de control tradicional. Diversos autores han trabajado en este tema.

*Gráficos de control adaptables*, son gráficos en los que se adaptan los parámetros del procedimiento de control tales como el tamaño de la muestra y el intervalo de muestreo en función de los resultados que se obtienen en los sucesivos muestreos que forman parte del mismo. Reynolds et al. (1988) exploraron la posibilidad de variar el intervalo de muestreo, dando lugar a los gráficos VSI (por variable. La alternativa de flexibilizar el tamaño muestral, se conoce como gráficos VSS, fue propuesta por Prabhu et al. (1993) y Costa (1994). Por último, Prabhu et al. (1994), Costa (1997) estudian el diseño de gráficos para el control de la media con tamaño y frecuencia de muestreo variables los llamados gráficos VSSI.

## Gráficos de Control multivariantes

Muy a menudo, la calidad de un producto depende de más de una característica de calidad. Dichas características son, en general, correlacionadas. En tales casos, es apropiado el uso de un procedimiento de control multivariante que considere la estructura de correlación entre

dichas características.

Muchas formas de gráficos de control multivariantes están disponibles en la literatura como resultado de los avances en la investigación en esta área.

**El Gráfico de control  $T^2$  de Hotelling** La primera investigación en los gráficos de control tipo Shewhart multivariante se remonta a [Hotelling \(1947\)](#), quien introdujo el problema de la correlación entre las características de calidad de un proceso. Diseñó un gráfico de control para el estadístico  $T^2$  como una extensión del gráfico  $\bar{x}$ , para monitorear el vector de medias de una variable aleatoria normal multivariante. Se presentan dos versiones del gráfico  $T^2$  de Hotelling: uno para datos subgrupados, y otro para las observaciones individuales.

*Datos subgrupados:* Para el caso en que  $p$  características de calidad correlacionadas son controladas de forma conjunta. Se supone que la distribución de probabilidad conjunta de las  $p$  características de calidad es la distribución normal  $p$ -variante. Se supone además que se dispone de muestras de tamaño  $n$ . El procedimiento requiere el cálculo de la media de la muestra para cada una de las  $p$  características de calidad. El estadístico de prueba representado en el gráfico de control chi-cuadrado para cada muestra es

$$\chi_0^2 = n (\bar{\mathbf{x}} - \boldsymbol{\mu})' \boldsymbol{\Sigma}^{-1} (\bar{\mathbf{x}} - \boldsymbol{\mu})$$

donde  $\boldsymbol{\mu} = [\mu_1, \dots, \mu_p]$  es el vector de medias bajo control para cada característica de calidad y  $\boldsymbol{\Sigma}$  es la matriz de covarianzas. El límite de control esta dado por:

$$LCS = \chi_{\alpha, p}^2$$

Si  $\mu$  y  $\Sigma$  se estiman a partir de  $m$  muestras tomadas cuando el proceso se supone bajo control de modo que:  $\hat{\mu} = \bar{\bar{x}}$  y  $\hat{\Sigma} = S$ .

Entonces el estadístico de control se transforma en:

$$T^2 = n (\bar{x} - \bar{\bar{x}})' S^{-1} (\bar{x} - \bar{\bar{x}})$$

y es usualmente llamado Gráfico de control  $T^2$  de Hotelling. Este es un gráfico de control direccional invariante; es decir, su capacidad para detectar un cambio en el vector medio solo depende de la magnitud del desplazamiento, y no de su dirección.

El objetivo en la fase I es la obtención de un conjunto de las observaciones bajo-control de modo que se pueden establecer límites de control para la fase II. En esta fase el límite de control está dado por

$$LCS = \frac{p(m-1)(n-1)}{mn-m-p+1} F_{\alpha,p,mn-m-p+1}$$

En la fase II, cuando el gráfico es utilizado para monitorear producción futura, el límite de control se define como:

$$LCS = \frac{p(m+1)(n-1)}{mn-m-p+1} F_{\alpha,p,mn-m-p+1}$$

Cuando los parámetros del gráfico se estiman a partir de un gran número de muestras preliminares, se acostumbra utilizar  $LCS = \chi_{\alpha,p}^2$  como el límite de control superior tanto en fase I y fase II.

*Observaciones individuales:* Supongamos que  $m$  muestras, cada uno de tamaño  $n = 1$ , están disponibles y que  $p$  es el número de características de calidad observados en cada muestra. Sea  $\bar{x}$  y  $S$  el vector de medias de la

muestra y la matriz de covarianza de estas observaciones, respectivamente. El estadístico de Hotelling  $T^2$  dado anteriormente se convierte en:

$$T^2 = (\mathbf{x} - \bar{\mathbf{x}})' \mathbf{S}^{-1} (\mathbf{x} - \bar{\mathbf{x}})$$

Chou et al. (2001) señalan que si  $n = 1$ , el límite de control para la fase I debe ser basado en una distribución beta. Esto es:

$$LCS_1 = \frac{(m-1)^2}{m} B\left(\frac{p}{2}, \frac{m-p-1}{2}; \alpha\right)$$

El límite de control superior de fase II para este estadístico esta dado por:

$$LCS_2 = \frac{p(m^2-1)}{m(m-p)} F_{(p, m-p; \alpha)}$$

Como mencionamos anteriormente, el gráfico  $T^2$  es un gráfico de control tipo Shewhart. Es decir, utiliza sólo la información de la muestra actual, por lo que, en consecuencia, es relativamente insensible a cambios pequeños y moderados en el vector de medias.

Los gráficos de control de suma acumulativa (CUSUM) y EWMA fueron desarrollados para proporcionar una mayor sensibilidad a pequeños cambios en el caso univariado, y pueden ser extendidos al caso multivariante. Al igual que en el caso univariado, la versión multivariante de estos gráficos son un procedimiento de fase II. Crosier (1988) y Pignatiello y Runger (1990) han propuesto varios procedimientos Cusum multivariados. (Lowry et al., 1992) han desarrollado un EWMA multivariante (MEWMA).

Lowry y Montgomery (1995), Bersimis et al. (2007) y Topalidou y

Psarakis (2009) realizan una revisión exhaustiva de la literatura relativa a gráficos de control multivariantes, para variables y atributos.

### 2.2.3. Gráficos de control para variables multinomiales.

#### El caso univariante

Marcucci (1985) plantea dos procedimientos basados en gráficos de control tipo Shewart para controlar productos que se los clasifican en categorías lingüísticas mutuamente excluyentes.

- *Clase 1:* Es útil cuando las proporciones de calidad son valores específicos, y cualquier cambio en estas proporciones debe ser detectado por el procedimiento de control. Para supervisar el proceso multinomial, se toman muestras aleatorias independientes del proceso en diferentes períodos de monitoreo. Supongamos que hay  $t$  valores lingüísticos en el conjunto de términos y sea  $X_{i1}, X_{i2}, \dots, X_{it}$  el número de observaciones en las categorías  $1, 2, \dots, t$  respectivamente, para el período de control  $i$ . Se reserva  $i = 0$  para el período base. Para monitorear un proceso multinomial bajo esta condición se utiliza un procedimiento estadístico estándar que utiliza el estadístico de bondad de ajuste de Pearson, definido como:

$$Y_i^2 = \sum_{j=1}^t \frac{(X_{ij} - n_i \pi_j)^2}{n_i \pi_j}$$

donde los  $\pi_j$  son las proporciones, y  $n_i = \sum_{j=1}^t X_{ij}$  es el tamaño de la muestra  $i$  -ésima. Cuando el proceso está bajo-control, la distribución asintótica de  $Y_i^2$  es  $\chi^2(t - 1)$ , llamada distribución chi-cuadrado con  $t - 1$  grados de libertad. El gráfico de control



construido con  $Y_i^2$  como estadístico de monitoreo tiene un límite de control que es determinado usando los percentiles de la distribución chi-cuadrado.

- *Clase 2:* Se utiliza cuando las proporciones de calidad del proceso no se conocen a priori y está diseñado para detectar únicamente incrementos en todas excepto una de las proporciones. El estadístico de bondad de ajuste de Pearson no es aplicable cuando los valores específicos de las proporciones de calidad son desconocidas. Un procedimiento estadístico adecuado es un test de homogeneidad de proporciones entre el periodo base y cada periodo  $i$ . El estadístico se define como

$$Z_i^2 = \sum_{k=i,0} \sum_{j=1}^t n_k \frac{\left( \frac{X_{kj}}{n_k} - \frac{X_{ij}+X_{0j}}{n_i+n_0} \right)^2}{\frac{X_{ij}+X_{0j}}{n_i+n_0}} = n_i n_0 \sum_{j=1}^t \frac{(p_{ij} - p_{0j})^2}{X_{ij} + X_{0j}}$$

donde  $p_{kj} = \frac{X_{kj}}{n_k}$  son las proporciones muestreadas y  $n_i$  es el tamaño de la muestra. Como  $n_0, n_i \rightarrow \infty$  tales que  $\frac{n_0}{n_i}$  es acotado y mayor que cero, la distribución de  $Z_i^2$  se aproxima a una  $\chi^2(t-1)$ . Por tanto, el gráfico de control que usa  $Z_i^2$  como estadístico de control tiene un límite de control superior que es un percentil apropiado de la distribución chi-cuadrado.

**Cozzucoli (2009)** propone un gráfico para el seguimiento de un proceso multinomial. En este caso los artículos se clasifican en  $(k+1)$  distintas categorías de defectos, ordenadas y mutuamente excluyentes. Sea  $D = (D_0, D_1, \dots, D_i, \dots, D_k)$  el vector de las  $(k+1)$  categorías de defectos. El componente genérico  $D_i$  indica la  $i$ -ésima categoría del grado de defecto,  $D_0$  es la categoría libre de defectos y  $D_k$  es la categoría más grave del grado de defecto. El estadístico a controlar es lo que define como índice normalizado de la deficiencia global que está dado por  $\delta = \sum d_i p_i$ . En general,  $d_i$  indica la cantidad de pérdida de calidad que el defecto de

la clase  $i$ -ésima introduce en el sistema y  $p_i$  es la probabilidad de que un producto pueda ser clasificado en la categoría de defecto  $D_i$ . Así, los límites de control y la línea central para el gráfico de control que propone están dados por:

$$LCS = \sum_{i=0}^k d_i p_{i0} + z_{(1-\alpha)^{1/(k+1)}} \sqrt{\frac{1}{n} \left[ \left( \sum_{i=0}^k d_i^2 p_{i0} \right) - \left( \sum_{i=0}^k d_i p_{i0} \right)^2 \right]},$$

$$LC = \sum_{i=0}^k d_i p_{i0},$$

$$LCI = \sum_{i=0}^k d_i p_{i0} - z_{(1-\alpha)^{1/(k+1)}} \sqrt{\frac{1}{n} \left[ \left( \sum_{i=0}^k d_i^2 p_{i0} \right) - \left( \sum_{i=0}^k d_i p_{i0} \right)^2 \right]},$$

donde  $z_{(1-\alpha)^{1/(k+1)}}$  es el cuantil superior  $(1 - \alpha)^{1/(k+1)}$  de la distribución normal estandarizada. Por lo general el vector  $p_0 = (p_{00}, p_{10}, \dots, p_{i0}, \dots, p_{k0})$  es desconocido y es necesario estimarlo usando  $m$  muestras preliminares de tamaño  $n$  tomadas del proceso bajo-control.

### El caso multivariante

**Li et al. (2014)** estudian el monitoreo de procesos categóricos multivariados en el marco sistemático de procesos binomiales multivariados y multinomiales multivariados, que incluyen los casos de varios factores todos ellos con dos niveles de atributos y al menos uno con más de dos niveles o categorías. Los datos categóricos se resumen en una tabla de

contingencia múlti-vías y se formulan en términos de modelos log-lineales, que caracterizan los patrones de asociación entre los factores categóricos. Un modelo log-lineal relaciona los logaritmos de los recuentos de las celdas esperados en la tabla de contingencia con un modelo lineal que es similar a un modelo de análisis de varianza (ANOVA). Proponen un gráfico de control de tipo EWMA utilizando la idea de la probabilidad ponderada. El estadístico a controlar está dado por:

$$R_k = 2z_k^T (\ln \hat{\mathbf{y}}_k - \ln \mathbf{m}^{(0)})$$

donde  $z_k$  es el promedio ponderado exponencialmente de los vectores de observación  $\mathbf{n}_j$  ( $j = 1, \dots, k$ ) en el tiempo,  $\hat{\mathbf{y}}_k$  es el estimador de máxima verosimilitud del vector de conteo de celda esperado sobre  $z_k$  y  $\mathbf{m}^{(0)}$  es el vector de conteo de celda esperado en el estado bajo control.

El gráfico de control muestra una señal de fuera de control si  $R_k < L$ , para  $k \geq 1$ , donde  $L > 0$  es un límite de control elegido para lograr un ARL bajo control específico, denotado por  $ARL_0$ . El límite de control  $L$  se determina por simulación basada en el modelo bajo control.

## 2.3. Teoría Difusa y gráficos de Control

### 2.3.1. Antecedentes

La palabra difusa (en inglés fuzzy), se puede interpretar como: borroso, indistinto, impreciso, confuso, vago. Pero se debe desatender esta definición y ver la palabra "difusa" como un adjetivo. Esencialmente, lo que se quiere enfatizar es que aún cuando los fenómenos que caracterizan a la teoría difusa o a los sistemas difusos pueden ser difusos, la propia teoría es precisa. De hecho, casi todas las teorías en ingeniería caracterizan

al mundo real de una manera aproximada. Por ejemplo, la mayoría de los sistemas reales son no lineales, pero su tratamiento se realiza como un sistema lineal poniendo mucho esfuerzo en el estudio de sistemas lineales. *Una buena teoría en la ingeniería debe ser precisa en la magnitud que caracteriza los rasgos importantes del mundo real y, al mismo tiempo, permitir el análisis matemático (Wang, 1997).*<sup>2</sup>

La teoría difusa se basa en la teoría de los conjuntos difusos introducida por el profesor Lotfi Zadeh en 1965 (Zadeh, 1965), profesor de Ingeniería Eléctrica y Ciencias de la Computación de la Universidad de California en Berkeley. En el artículo publicado por Zadeh se introdujo por primera vez la palabra "fuzzy" en la literatura técnica y describe con la teoría matemática de conjuntos clásicos, como se puede trabajar matemáticamente con expresiones imprecisas, tal como lo hace el ser humano. La teoría de los conjuntos difusos es una extensión de la teoría de conjuntos clásicos. En contraste con la "lógica nítida" donde los conjuntos digitales tienen lógica binaria, las variables de lógica difusa pueden tener un valor de verdad que oscila entre 0 y 1 y no está limitado a los dos valores de verdad de la lógica proposicional clásica.

Desde el área de estadística varios autores han analizado y discutido la importancia y el aporte realizado por la teoría difusa en la búsqueda de medir y controlar la variabilidad. (Laviolette et al. (1995), Almond (1995), Coppi et al. (2006), Sorooshian (2013))

La lógica difusa y la lógica probabilística son matemáticamente similares (ambos tienen valores de verdad que oscilan entre 0 y 1), pero conceptualmente distintas debido a diferentes interpretaciones. La lógica difusa corresponde a "grados de verdad", mientras que la lógica probabilística corresponde a "probabilidad, verosimilitud". La lógica difusa y la lógica probabilística producen diferentes modelos de las

---

<sup>2</sup>La mayor parte de las definiciones y conceptos expuestos en esta sección se han tomado de Raz y Wang (1990) y Wang y Raz (1990).

mismas situaciones del mundo real. Ambos, grados de verdad y las probabilidades varían entre 0 y 1 y por lo tanto pueden parecer similares a primera vista; sin embargo, es esencial tener en cuenta que la lógica difusa utiliza grados de verdad como un modelo matemático del fenómeno de la vaguedad, mientras que la probabilidad es un modelo matemático de la incertidumbre.

Los conjuntos difusos y la lógica difusa son poderosas herramientas matemáticas para el modelado de sistemas inciertos en la industria, la naturaleza y la humanidad. Son también facilitadores para el razonamiento de sentido común en la toma de decisiones en ausencia de información completa y precisa (Bojadziev y Bojadziev, 1995) .

La teoría de conjuntos difusos se puede utilizar en una amplia gama de ámbitos en los que la información es incompleta o imprecisa, tiene como uno de los objetivos el desarrollo de una metodología para la formulación y solución de problemas que son demasiado complejos o mal definidos como para ser susceptibles al análisis mediante técnicas convencionales, y se definen generalmente por medio de variables lingüísticas.

Williams y Zigli (1987) argumentan fuertemente en favor de las técnicas de aseguramiento de calidad que reconocen e incorporan la imprecisión del juicio humano, especialmente en las industrias de servicios. La vaguedad y la ambigüedad inherente en variables lingüísticas pueden tratarse matemáticamente con la ayuda de la teoría de conjuntos difusos. Esta teoría, proporciona las herramientas para realizar operaciones matemáticas con descriptores lingüísticos representándolos como conjuntos difusos de un dominio numérico.

Las características de calidad de un producto pueden ser medibles en escalas numéricas, como la longitud, el peso, la tensión, etc., en los que se utilizan gráficos denominados gráficos de control para variables. Si las

características relacionadas con la calidad no pueden ser representadas en forma numérica, tales como las características de apariencia, suavidad, color, etc., entonces se utilizan gráficos de control para atributos. Las unidades de producto deben clasificarse como “conforme” o “no conforme”, dependiendo de si cumplen o no con las especificaciones, o se cuenta el número de no conformidades (desviaciones de las especificaciones) por unidad de producto. El gráfico-*p* se usa para controlar la fracción no conforme, y el gráfico-*c* para el número de no conformidades por unidad de producto. En ambos casos, se utiliza la aproximación normal para el cálculo de los límites de control.

La clasificación binaria en conformes y no conformes, que permite usar el gráfico-*p* podría no ser apropiada en muchas situaciones en las que la calidad del producto no cambia abruptamente de ser satisfactoria a no tener ningún valor, y puede haber una serie de niveles intermedios.

Para complementar la clasificación binaria, varios niveles intermedios que describen la calidad del producto se puede utilizar como criterios. Estos niveles intermedios se pueden expresar en forma de términos lingüísticos. Por ejemplo, la calidad de un producto se puede clasificar por medio de uno de los siguientes términos: “perfecto”, “bueno”, “medio”, “pobres” y “malo”, en función de su incumplimiento de las especificaciones. Para describir la característica de calidad asociada a cada término lingüístico se pueden utilizar funciones continuas apropiadamente seleccionados como veremos más adelante.

La asignación de pesos, también llamado deméritos, para reflejar el grado de gravedad de la no conformidad del producto se ha adoptado desde hace varios años. Uno de los esquemas más comúnmente usados, introducidos en los Laboratorios Bell (Hoadley, 1981), se basa en cuatro clases de defectos: “muy graves”, “graves”, “moderadamente graves” y “no graves”. Diferentes números de deméritos se asignan a cada clase y el número total de deméritos se controla con un gráfico-*c*.

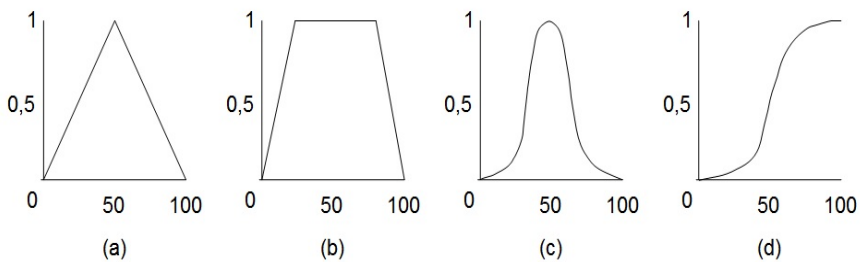
Este enfoque requiere la capacidad de clasificar cada defecto o no-conformidad de forma única en una de varias clases mutuamente excluyentes con límites bien definidos entre ellos. Muy a menudo, hay una cierta ambigüedad y vaguedad en la sentencia aplicada por los inspectores humanos, sobre todo cuando se trata de características que se evalúan subjetivamente. Inspectores de control de calidad pueden sentirse más cómodos expresando su evaluación como una descripción verbal o lingüística en lenguaje ordinario en lugar de hacer una selección de opción múltiple.

### 2.3.2. Variables Lingüísticas y Conjuntos Difusos

Una variable lingüística difiere de una variable numérica en que sus valores no son números sino palabras o frases en un lenguaje (Zadeh, 1975c) En el contexto de teoría de conjuntos difusos, una variable lingüística  $L$  se caracteriza por su "conjunto de términos"  $T(L)$ , que es el conjunto de todos los valores posibles que  $L$  puede tener. Cada término  $L_i$  en  $T(L)$  es un adjetivo o una combinación de adjetivos, modificadores ("muy", "bastante", "más o menos", etc.) y conectores ("y", "o", "bien", etc.), se caracteriza por una "función de pertenencia"  $\mu_i(x)$ , y se asocia a un conjunto difuso  $F_i$ . La función de pertenencia indica el grado en que  $x$  pertenece a un conjunto difuso. La variable  $X$ , que es un conjunto de todos los posibles valores asociados a la variable lingüística  $L$ , se denomina 'variable base'. La variable base podría ser directamente medible, tal como la longitud de un rasguño, o puede estar basada en algún tipo de evaluación cualitativa expresada en una escala o en un intervalo, como la vistosidad de un rasguño. Es usualmente conveniente estandarizar el rango de la variable base en el intervalo cerrado  $[0, 1]$ . Una discusión completa de variables lingüísticas aparece en Zadeh (1975c,b,a).

## Función de Pertenencia

La forma de la función de pertenencia utilizada depende del criterio de resolución de cada problema así como del punto de vista de cada usuario. El único requisito que debe reunir una función de pertenencia es que tome valores entre 0 y 1 y que sea continua en ese intervalo. En la figura 2.3 , se muestran algunas de las funciones de pertenencia más empleadas.



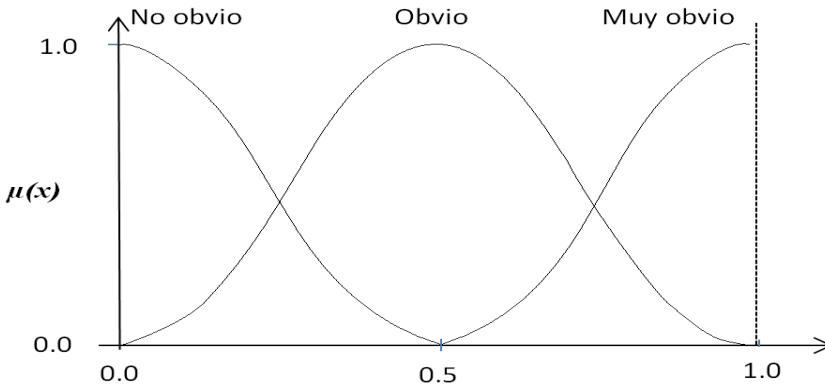
**Figura 2.3:** Funciones de pertenencia más empleadas: (a) Triangular, (b) Trapezoidal, (c) Gaussiana y (d) Sigmoid

El usuario es el encargado de elegir el número de funciones de pertenencia asociadas a una misma variable lingüística. Pero se debe tener en cuenta que a mayor número de funciones de pertenencia, mayor complejidad computacional se tendrá. Sin embargo, también se ganará en resolución.

Por otro lado se ha de tomar en consideración que estas funciones pueden estar solapadas o no. El hecho de que varias funciones de pertenencia se encuentren solapadas, refuerza el concepto de la lógica difusa ya que un elemento puede pertenecer con distinto grado a varios conjuntos difusos al mismo tiempo.



La figura 2.4 ilustra tres posibles conjuntos difusos asociados con los términos “muy obvio”, “obvio”, y “no obvio” de una variable lingüística utilizada para describir la vistosidad de rasguños superficiales. La variable base, que aparece en el eje horizontal, se puede obtener mediante la medición de la longitud del rasguño, o por evaluación subjetiva en una escala graduada. Tenga en cuenta que cuando  $x = 0.1$ , el rasguño pertenece a los conjuntos difusos asociados con ambos valores lingüístico “no obvio” y “obvio”, permitiendo así una cierta ambigüedad en la clasificación. También tenga en cuenta que el conjunto difuso etiquetado ‘no obvio’ incluye valores de  $x$  entre 0 y 0.5, proporcionando una manera de expresar la vaguedad en la clasificación.



**Figura 2.4:** Funciones de pertenencia de tres valores lingüísticos: No obvio, obvio, muy obvio, (Wang and Raz (1990))

La cuestión de la determinación de una función de pertenencia válida es crítica para el uso exitoso de variables lingüísticas. Varios métodos de elaboración y selección de funciones de pertenencia han aparecido en la literatura. Basados en datos estadísticos (Civanlar y Trussell, 1986), encuestas (Hersh y Caramazza, 1976), creencias diferenciales (Saaty, 1975), preferencias relativas (Turksen, 1986), ranking (Zadeh, 1978) o la selección interactiva a partir de una biblioteca de funciones de pertenencia predefinida.

Es posible realizar operaciones aritméticas sobre variables lingüísticas representadas como conjuntos difusos, según las definiciones y principios de matemáticas difusas, que se han desarrollado ampliamente. Esto requiere que la función de pertenencia sea normal (al menos un punto de la variable base con pertenencia igual a 1) y convexa. Para una introducción a aritmética difusa, véase [Kaufmann y Gupta \(1985\)](#).

A lo largo del resto de este documento las funciones de pertenencia asociadas a los términos lingüísticos serán triangulares a fin de asegurarnos de que sean normales, convexas y normalizadas en el intervalo  $[0, 1]$ . Éstas además tienen la ventaja de que pueden ser representadas como una terna numérica.

## **Borrosidad y aleatoriedad**

El concepto de "borrosidad" (en inglés fuzziness) se hizo científico en la teoría de sistema matemático de Zadeh (1965) sobre definición de conjuntos difusos ([Fuhmann, 1988](#)). Esto se refiere a un "marco conceptual que proporciona una forma natural de tratar con problemas en los que la fuente de imprecisión es la ausencia de criterios bien definidos de pertenencia a un conjunto" ([Zadeh, 1965](#)).

La teoría de conjuntos difusos trata la incertidumbre de manera diferente de la teoría de la probabilidad. Cuando la incertidumbre se debe a la aleatoriedad de los eventos, estos eventos son los resultados casuales de situaciones definidas con precisión. Múltiples observaciones de estas situaciones pueden mejorar el conocimiento acerca de los eventos aleatorios, expresadas en forma de probabilidades. Por el contrario, la teoría de conjuntos difusos apunta a situaciones que impiden descripciones precisas. Por otra parte, las observaciones repetidas pueden no ser posibles, ya que los eventos pueden a menudo ser únicos e irreproducibles.

Como ilustración de la diferencia entre la probabilidad y la pertenencia, considere la declaración "Una unidad buena puede tener  $x$  arañazos", con la variable  $X$  tomando valores del conjunto  $\{0, 1, 2, 3, \dots\}$ . Podemos asociar una función de pertenencia  $\mu(x)$  con  $X$  interpretando  $\mu(x)$  como el grado de pertenencia al conjunto de unidades buenas de una unidad con  $x$  arañazos. También se puede asociar una densidad de probabilidad  $f(x)$  con  $X$  mediante la interpretación de  $f(x)$  como la probabilidad de observar una unidad con  $x$  arañazos. La tabla 2.1 muestra algunos valores hipotéticos de  $\mu(x)$  y  $f(x)$ . En el ejemplo que hemos tomado de Wang y Raz (1990), los valores de  $\mu(x)$ , que son subjetivos, se obtuvieron a partir de los juicios de los expertos y los valores de  $f(x)$  se calcularon a partir de observaciones de la muestra que son objetivos.

**Tabla 2.1:** Funciones de probabilidad y de pertenencia asociados a una unidad con  $x$  rasguños (Raz y Wang (1990))

$x$	0	1	2	3	4	5	6	7
$\mu(x)$	1.0	0.9	0.8	0.5	0.2	0.0	0.0	0.0
$f(x)$	0.1	0.1	0.2	0.2	0.1	0.1	0.1	0.1

En este ejemplo, la pertenencia de una unidad con 2 arañazos al conjunto difuso descrito por el valor lingüístico bueno es 0.8, mientras que la probabilidad de que cualquier unidad tenga 2 arañazos es 0.2. Por lo tanto, un alto grado de pertenencia no implica un alto grado de probabilidad, ni un bajo grado de probabilidad implica un bajo grado de pertenencia. Por otra parte, la suma de las probabilidades de todos los resultados posibles es igual a uno, mientras que esta restricción no se aplica a la función de pertenencia. Llegamos a la conclusión de esta discusión con una frase de Zadeh (1980):

"En esencia, la teoría de los conjuntos difusos está dirigida al desarrollo de un cuerpo de conceptos y técnicas para hacer frente a las fuentes de incertidumbre o imprecisión que no son de naturaleza estadística".

### 2.3.3. Valor representativo de un subconjunto difuso

Cada observación en una muestra de datos lingüísticos es un término lingüístico asociado con un subconjunto difuso definido en la variable base y descrito por una función de pertenencia. Los subconjuntos difusos tienen que ser finalmente convertidos en escalares, a los que se les llamará valores representativos, con el fin de permitir el cálculo de los límites de control y el trazado de los puntos de las muestras en los gráficos de control. Esto puede hacerse de diferentes maneras, siempre y cuando el resultado sea intuitivamente representativo del rango de la variable base incluida en el conjunto difuso. Cuatro métodos, que son similares en principio a las medidas de tendencia central utilizada en estadística descriptiva, se presentan. Cabe señalar que no existe una base teórica que soporte alguno de estos métodos de manera específica, y la selección entre ellos debe basarse principalmente en la facilidad de computación o la preferencia del usuario (Gülbay y Kahraman, 2007).

1. *La moda difusa,  $f_{mode}$ .* La moda difusa de un conjunto difuso  $F$  es el valor de la variable base donde la función de pertenencia es igual a 1. Esto se afirma como

$$f_{mode} = \{x | \mu_F(x) = 1\}, \quad \forall x \in F \quad (2.1)$$

Es único, si la función de pertenencia es unimodal.

2. *El rango medio difuso nivel- $\alpha$ ,  $f_{mr}(\alpha)$ .* Este se define como el punto medio de los extremos de la línea de corte nivel- $\alpha$ . Un corte de nivel- $\alpha$ , denotado por  $A_\alpha$ , es un conjunto que comprende todos los elementos cuya pertenencia es mayor que o igual a  $\alpha$  (Kaufmann y Gupta, 1985). Si  $a_\alpha$  y  $b_\alpha$  son los puntos extremos de  $A_\alpha$ , entonces

$$f_{mr}(\alpha) = \frac{1}{2}(a_\alpha + b_\alpha). \quad (2.2)$$

De hecho, la moda difusa es un caso especial del rango medio difuso de nivel- $\alpha$  cuando  $\alpha = 1$ .

3. *La mediana difusa,  $f_{med}$* . Esta dada por el punto que divide el área de la región bajo la curva de la función de pertenencia de un conjunto difuso en dos regiones de igual área, es decir que, satisface la siguiente ecuación:

$$\int_a^{f_{med}} \mu(x) dx = \int_{f_{med}}^b \mu(x) dx = \frac{1}{2} \int_a^b \mu(x) dx \quad (2.3)$$

donde  $a$  y  $b$  son los puntos finales en la variable base del conjunto difuso  $F$  tal que  $a < b$ .

4. *La media difusa,  $f_{avg}$* . Basado en [Zadeh \(1975a\)](#), la media difusa esta dada por:

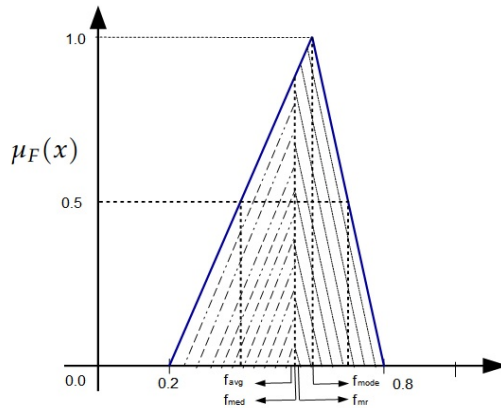
$$f_{avg} = Av(x : F) = \frac{\int_{x=0}^1 x \mu_F(X) dx}{\int_{x=0}^1 \mu_F(X) dx} \quad (2.4)$$

En general, los dos primeros métodos son más fáciles de calcular que los dos últimos, especialmente cuando la función de pertenencia es no lineal. Sin embargo, la moda difusa puede conducir a un resultado sesgado cuando la función de pertenencia es extremadamente asimétrica. El rango medio difuso es más flexible ya que se puede elegir diferentes niveles de pertenencia ( $\alpha$ ) de interés. Si el área bajo la función de pertenencia se considera como medida de borrosidad apropiada, se puede pensar que la mediana difusa es adecuada. Cuando queremos dar cuenta de la forma de la función de pertenencia así como su ubicación, la media difusa será entonces una mejor elección ya que se deriva del principio de extensión y es básicamente un promedio ponderado de la variable base.

**Ejemplo:** Considere un conjunto difuso  $A$  definido por la función de pertenencia a continuación:

$$\mu_A(x) = \begin{cases} 0, & x \leq 0.2 \\ 2.5x - 0.5, & 0.2 \leq x \leq 0.6 \\ -5x + 4, & 0.6 \leq x \leq 0.8 \\ 0, & x \geq 0.8 \end{cases}$$

Los valores representativos para el conjunto difuso  $A$ , calculados a partir de las definiciones dadas, son:  $f_{mode} = 0.6$ ,  $f_{mr}(0.5) = 0.55$ ,  $f_{med} = 0.546$ , y  $f_{avg} = 0.533$ . La figura 2.5 ilustra, además, estos resultados.



**Figura 2.5:** Función de pertenencia y valores representativos del conjunto difuso  $A$  (Wang y Raz (1990))

Cuando una muestra difusa es asimétrica, estos métodos de transformación difusos resultan en diferentes números nítidos, y por lo tanto un control de proceso basado en estas transformaciones puede resultar en diferentes decisiones.

### 2.3.4. Representación de una muestra

Una muestra puede constar de varias observaciones seleccionados por inspección. Cada observación se clasifica por un valor lingüístico y se asocia con una función de pertenencia conocida. Estos valores lingüísticos individuales deben combinarse con el fin de producir un único valor representativo de esta muestra, similar a la media en un gráfico  $\bar{x}$  o la fracción no conforme en un gráfico- $p$ . Esta combinación de las observaciones individuales se puede realizar ya sea antes o después de su conversión en un valor representativo de los valores lingüísticos.

*En el primer caso*, se debe agregar el conjunto difuso asociado a cada valor lingüístico en una muestra y luego se divide por el número de observaciones en la muestra. Estas operaciones sobre conjuntos difusos se llevan a cabo de acuerdo con las reglas de la aritmética difusa. El resultado será un conjunto difuso que puede no corresponder a un valor lingüístico específico en el conjunto de términos, pero, no obstante, es representativo del nivel de calidad media de la muestra. Luego, se puede conseguir un solo valor numérico para esta muestra mediante el cálculo de un valor representativo de acuerdo con una de las cuatro fórmulas descritas en la subsección anterior.

Supongamos que hay  $t$  valores lingüísticos del conjunto de términos  $L$ , denotados por  $L_i$ , con  $i = 1, \dots, t$ . Para cada valor lingüístico, el correspondiente conjunto difuso se denota por  $F_i$ , y se describe por la función de pertenencia  $\mu_i(x_i)$ , con  $x_i$  siendo un subconjunto de la variable base estandarizada.

Una muestra  $S$  de  $n$  observaciones puede ser descrito como:

$$S = \{(F_1, k_1), \dots, (F_i, k_i), \dots, (F_t, k_t)\}$$

donde  $k_i$ , es el número de elementos asignados al valor lingüístico  $L_i$ , por el inspector de calidad, y debe cumplirse que  $k_1 + \dots + k_i + \dots + k_t = n$ .

El conjunto difuso correspondiente a la muestra es una media ponderada de los conjuntos difusos del conjunto de términos en la muestra y se denota por  $MF_s$ . La función de pertenencia de  $MF_s$ , denotado por  $\mu_s(x_s)$ . Según Kaufmann y Gupta (1985) está dada por:

$$\mu_s(x_s) = \text{Max}_{x_n=(k_1x_1+\dots+k_tx_t)/n} \{ \text{Min}[\mu_1(x_1), \dots, \mu_t(x_t)] \} \quad (2.5)$$

Finalmente, el valor representativo para el conjunto difuso de la muestra definido por  $\mu_s(x_s)$  se obtiene mediante la aplicación de una de las cuatro fórmulas presentadas en la sección anterior.

*El segundo caso* a considerar sería el cálculo de la media de la muestra después de la conversión de los valores lingüísticos en sus valores representativos, entonces, los cálculos son algo más simples:

Sea  $r_i$  el valor representativo del conjunto difuso  $F_i$ , que se obtiene mediante la aplicación de alguna de las fórmulas presentadas en la sección anterior, La media de la muestra,  $M$ , se calcula como el promedio ponderado de los valores representativos  $r_i$  correspondientes a los valores lingüísticos en la muestra, de acuerdo a la siguiente fórmula:

$$M = \frac{r_1k_1 + \dots + r_ik_i + \dots + r_tk_t}{n} \quad (2.6)$$

El primer método retiene más borrosidad para la representación de una muestra que el segundo método, pero implica un mayor esfuerzo computacional especialmente en el que se supone que la función de



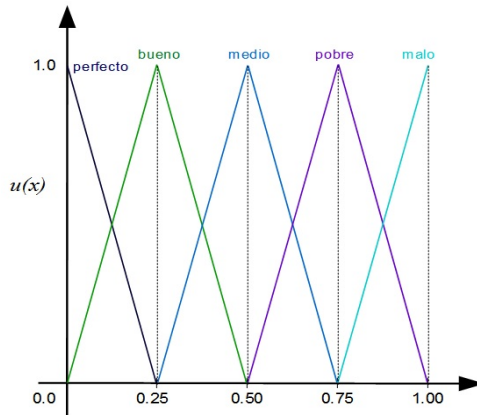
pertenencia es no lineal.

A continuación un ejemplo para ilustrar los cálculos implicados en los dos métodos:

**Ejemplo:** Considere una variable lingüística utilizada para evaluar la calidad de un producto con el conjunto de términos  $L$ :

$$L = \{perfecto, bueno, medio, pobre, malo\}$$

La variable base es una medida del nivel de calidad, y ha sido estandarizada en el intervalo  $[0, 1]$  de tal manera que, "0" indica el mejor nivel de calidad, y "1" indica lo peor. Las funciones de pertenencia asociadas a cada término lingüístico se muestran en la figura 2.6



**Figura 2.6:** Funciones de pertenencia para un conjunto de cinco términos lingüísticos (Wang y Raz (1990))

Las reglas de correspondencia de las funciones de pertenencia son:

$$\mu_{perfecto}(x) = \begin{cases} 1 - 4x, & 0 \leq x \leq 0.25 \\ 0, & x \geq 0.25 \end{cases}$$

$$\mu_{bueno}(x) = \begin{cases} 4x, & 0 \leq x \leq 0.25 \\ 2 - 4x, & 0.25 \leq x \leq 0.5 \\ 0, & x \geq 0.5 \end{cases}$$

$$\mu_{medio}(x) = \begin{cases} 0, & x \leq 0.25; x \geq 0.75 \\ 4x - 1, & 0.25 \leq x \leq 0.5 \\ 3 - 4x, & 0.5 \leq x \leq 0.75 \end{cases}$$

$$\mu_{pobre}(x) = \begin{cases} 0, & x \leq 0.5 \\ 4x - 2, & 0.5 \leq x \leq 0.75 \\ 4 - 4x, & 0.75 \leq x \leq 1 \end{cases}$$

$$\mu_{malo}(x) = \begin{cases} 0, & x \leq 0.75 \\ 4x - 3, & 0.75 \leq x \leq 1 \end{cases}$$

Ahora, considere una muestra de 10 observaciones, que han sido clasificadas de la siguiente manera:

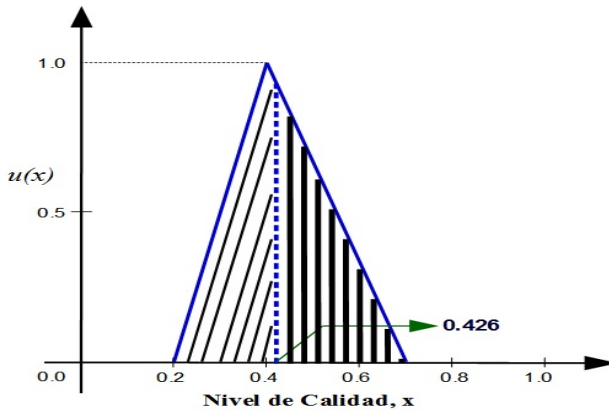
$$S = \{(F_{perfecto}, 3), (F_{bueno}, 2), (F_{medio}, 2), (F_{pobre}, 2), (F_{malo}, 1)\}.$$

Mediante la combinación de estas 10 observaciones siguiendo el primer enfoque, es decir usando la ecuación 2.5, el conjunto difuso

resultante asociado con la media de la muestra puede ser descrita por la función de pertenencia:

$$\mu_S(x) = \begin{cases} 0, & 0 \leq x \leq 0.2 \\ 5x - 1, & 0.2 \leq x \leq 0.4 \\ 2.333 - 3.333x, & 0.4 \leq x \leq 0.7 \\ 0, & x \geq 0.7 \end{cases}$$

Si el valor representativo de la función de pertenencia resultante se calcula como la mediana difusa, se obtiene el número  $M = 0.426$ . El gráfico de la función de pertenencia asociada a la muestra resultante se proporciona en la figura 2.7.



**Figura 2.7:** Función de pertenencia resultante de la muestra S. (Wang y Raz(1990))

Con el enfoque alternativo, el valor representativo para cada valor lingüístico es calculado primero.

Usando la transformación mediana difusa, estos valores son:

$$r_{perfecto} = 0.146, r_{bueno} = 0.25, r_{medio} = 0.5$$

$$r_{pobre} = 0.75, r_{malo} = 0.854$$

Por último, la media de la muestra se calcula como el promedio de los valores representativos de estas observaciones individuales de acuerdo a la ecuación 2.6, luego:

$$M = \frac{0.146 \times 3 + 0.25 \times 2 + 0.5 \times 2 + 0.75 \times 2 + 0.854 \times 1}{10} = 0.429$$

## Enfoques probabilístico y de pertenencia

Wang y Raz (1990) presentaron dos enfoques para la construcción de gráficos de control difusos para datos lingüísticos, llamados enfoque probabilístico y enfoque de pertenencia.

- *Enfoque Probabilístico*: Suponga que hay  $t$  valores lingüísticos en el conjunto de términos denotados por  $L_i$ ,  $i = 1, 2, \dots, t$ . Cada término  $L_i$  está asociado con un conjunto difuso  $F_i$  y se describe por una función de pertenencia  $\mu_i(x_i)$ . Usualmente es conveniente estandarizar el rango de la base variable  $x_i$  al intervalo cerrado  $[0, 1]$ . Una muestra  $S$  de  $n$  observaciones puede ser descrita como  $S = \{(F_1, k_1), \dots, (F_t, k_t)\}$ , donde  $k_i$  es el número de observaciones asignadas al conjunto difuso  $F_i$  y  $k_1 + k_2 + \dots + k_t = n$ . La función de pertenencia de  $MF_S$ , la media de los conjuntos difusos en la muestra, denotada por  $\mu_S(x_S)$  está dada por

$$\mu_S(x_S) = \max_{\{x_S=(k_1x_1+k_2x_2+\dots+k_tx_t)/n\}} \{\min[\mu_1(x_1), \mu_2(x_2), \dots, \mu_t(x_t)]\}$$

Sea  $r_i$  el valor representativo del conjunto difuso  $F_i$  que se obtiene usando una de las fórmulas de defusificación planteados anteriormente. La media de la muestra,  $M$ , se calcula como el promedio de los valores  $r_i$ , de acuerdo con la siguiente expresión:

$$M = \frac{k_1r_1 + k_2r_2 + \dots + k_tr_t}{n} = \frac{1}{n} \sum_{i=1}^t k_i r_i$$

Suponga que se usan  $m$  muestras válidas de tamaño  $n$  para la construcción de gráficos de control, entonces la línea central LC está dada por

$$LC = \frac{1}{m} \sum_{j=1}^m M_j = \frac{1}{mn} \sum_{j=1}^m \sum_{i=1}^t k_{ij} r_i$$

donde  $M_j$  es la media de la  $j$ -ésima muestra y  $K_{ij}$  es el número de observaciones asignadas al conjunto difuso  $F_i$  en la muestra  $j$ . En este enfoque, un estimador de la desviación estándar del valor representativo en la muestra  $j$  se denota por  $SD_j$ .

$$SD_j = \left( \frac{1}{n-1} \sum_{i=1}^t k_{ij} (r_i - M_j)^2 \right)^{1/2}$$

La media de las desviaciones estándar de las  $m$  muestras,  $MSD$ , es entonces:

$$MSD = \frac{1}{m} \sum_{j=1}^m SD_j$$

Suponiendo que la distribución de muestreo es aproximadamente normal, mediante la aplicación de la formulación estándar para los

gráficos de control, podemos derivar los límites de control. Puesto que los puntos marcados en el gráfico son medias de los valores representativos de la muestra, ellos deben encontrarse dentro del rango  $[0, 1]$ . Consecuentemente.

$$LCS_{probabilistico} = \text{Min} \{1, CL + A_3MSD\}$$

$$LCI_{probabilistico} = \text{Max} \{0, CL - A_3MSD\}$$

donde  $A_3 = \frac{3}{c_4\sqrt{n}}$  y  $c_4 = \sqrt{\frac{2}{n-1} \frac{\Gamma(\frac{n}{2})}{\Gamma(\frac{n-1}{2})}}$

- *Enfoque de Pertenencia:* En este enfoque, los límites de control están basados en las funciones de pertenencia. Entonces la media de las  $m$  muestras inicialmente válidas está dada por

$$LC = \frac{1}{m} \sum_{j=1}^m M_j = \frac{1}{mn} \sum_{j=1}^m \sum_{i=1}^t k_{ij} F_i$$

Ahora consideremos un conjunto difuso convexo  $F$  y sea  $x_m$  la moda (o el rango medio para  $\alpha = 1$ ) de su función de pertenencia. Suponga  $x_l(\alpha)$  y  $x_r(\alpha)$  son los puntos finales del  $\alpha$  – corte de la función de pertenencia de  $F$ . La desviación media para un conjunto difuso dado  $A$ , denotada por  $\delta(A)$ , está definida como la suma de la desviación media izquierda  $\delta_l(A)$  y la desviación media derecha  $\delta_r(A)$  las cuales se calculan de acuerdo a la siguiente formula (Kaufmann y Gupta, 1985)

$$\delta(A) = \delta_l(A) + \delta_r(A) = \int_{\alpha=0}^1 (x_m - x_l(\alpha)) d\alpha + \int_{\alpha=0}^1 (x_r(\alpha) - x_m) d\alpha$$

$$\delta(A) = \int_{\alpha=0}^1 (x_r(\alpha) - x_l(\alpha)) d\alpha$$

donde  $\alpha$  es el valor de pertenencia.

Los límites de control se colocan debajo y arriba de la línea central a una distancia expresada por:

$$LCS_{pertenencia} = \text{Min} \{1, (CL + k\delta)\}$$

$$LCI_{pertenencia} = \text{Max} \{0, (CL - k\delta)\}$$

donde  $k$  es el número de desviaciones medias que los límites de control se encuentran lejos de la línea central. El valor de  $k$  se puede determinar basándose en la experiencia o mediante estudios de simulación Monte-Carlo para valores fijos del error tipo I.

### 2.3.5. Algunos gráficos de control con enfoque difuso

Gülbay y Kahraman (2007) proponen un alternativa a los gráficos tipo-C llamada, “método de aproximación directa (DFA)” . En contraste con los gráficos de control difusos existentes, el método propuesto es muy diferente en el sentido de que no requiere el uso de la defusificación. Esto evita la pérdida de información incluida en las muestras. Se compara directamente los datos lingüísticos en el espacio difuso sin hacer ninguna transformación. Se usan algunos ejemplos numéricos para ilustrar el rendimiento del método y la interpretación de sus resultados.

Senturk y Erginel (2009) dicen que la incertidumbre, en los gráficos  $\bar{X} - R$ , proviene del sistema de medición incluyendo operadores y medidores, y las condiciones ambientales. Plantean el uso de límites de control difusos utilizando funciones de pertenencia para tratar con dicha incertidumbre, éstos límites de control difusos proporcionan una evaluación más precisa y flexible. Proponen la construcción de gráficos de control difusos  $\tilde{\bar{X}} - \tilde{R}$  y  $\tilde{\bar{X}} - \tilde{S}$  con cortes- $\alpha$  para aumentar la flexibilidad de los límites de control tradicionales.

## **2.4. Gráficos de control para variables multinomiales con enfoque difuso**

### **2.4.1. Introducción**

Algunos procedimientos se han propuesto para monitorear procesos multinomiales, mediante el uso de la teoría difusa, para cuando los productos se clasifican en categorías lingüísticas mutuamente excluyentes. Tanto para cuando estamos interesados en controlar una característica de calidad o varias correlacionadas entre sí, por lo que en la literatura se pueden encontrar propuestas de gráficos de control difusos, tanto para univariantes como multivariantes.

### **2.4.2. Gráficos de control con enfoque difuso para variables multinomiales univariantes.**

Bradshaw (1983) utilizó la teoría de conjuntos difusos como base para la interpretación de la representación de la conformidad del producto con un estándar de calidad de forma gradual. Cuando los costes de calidad deficiente se relacionan con el grado de no conformidad, existe una función de compatibilidad que describe el grado de no conformidad asociada con cualquier valor dado de esa característica de calidad. Esta función de compatibilidad puede entonces ser usada para construir los gráficos de control difuso económicos en un gráfico de control de aceptación.

Wang y Raz (1990) describen un método basado en la teoría difusa que asigna un conjunto difuso a cada término lingüístico y luego se combinan para cada muestra utilizando aritmética difusa para la determinación de



los límites de control. Aquí, medidas difusas representativas conocidas como valor representativo se obtienen mediante el uso de cualquiera de los cuatro métodos definidos en la sección anterior (a saber, moda difusa, rango medio de nivel- $\alpha$  difuso, la mediana difusa, o el promedio difuso), para construir los límites de control. Las funciones de pertenencia definidas para las variables lingüísticas en el método anterior son elegidas arbitrariamente y por lo tanto la decisión de control de procesos puede cambiar según la elección del usuario de los valores de los parámetros de decisión.

[Raz y Wang \(1990\)](#) como una continuación del trabajo previamente realizado presentan dos enfoques para la construcción del gráfico de control para variables multinomiales, el enfoque probalístico y el enfoque de pertenencia.

[Kanagawa et al. \(1993\)](#), introdujeron modificaciones a la construcción de gráficos de control propuesta por Wang y Raz. Su estudio tiene como objetivo controlar directamente las distribuciones de probabilidad subyacentes de los datos lingüísticos, que no fueron considerados por Wang y Raz y propone gráficos de control para datos lingüísticos desde un punto de vista diferente al de [Raz y Wang \(1990\)](#) con el fin no sólo de controlar la media del proceso, sino también para controlar su variabilidad sobre la base de la estimación de la distribución de probabilidad que existe detrás de los datos lingüísticos. Definieron la línea central como la media de los promedios de las muestras acumuladas y calculan entonces los límites de control utilizando series Gram-Charlier. La principal dificultad de este enfoque es que la función de distribución de probabilidad desconocida no se puede determinar fácilmente.

[Taleb y Limam \(2002\)](#) discuten la comparación entre los enfoques, difuso y probabilístico para la construcción de gráficos de control para datos lingüísticos, planteados por [Marcucci \(1985\)](#) y [Raz y Wang \(1990\)](#) basada en la longitud de racha promedio (ARL). Una de las principales

conclusiones presentadas es que el grado de borrosidad y el método usado para obtener el valor representativo afectan al rendimiento del gráfico de control contrariamente a lo concluido por Raz y Wang (1990),

Gülbay et al. (2004) propusieron un gráfico de control por atributos  $\alpha$ -corte, este enfoque proporciona la capacidad de determinar la hermeticidad de la inspección mediante la selección de un adecuado  $\alpha$ -corte. Para ellos, la representación de las variables lingüísticas como conjuntos difusos, los cuales pueden ser manipulados con las reglas de la aritmética difusa, mantiene la ambigüedad y vaguedad inherente a las lenguas naturales y mejora la capacidad expresiva de los inspectores de control de calidad. Definen la hermeticidad de la inspección mediante el uso de gráficos de control  $\alpha$ -corte. El nivel- $\alpha$  se determina de acuerdo con la naturaleza de los productos y procesos de fabricación.

Cheng (2005) presenta la construcción de gráficos de control difuso para un proceso con resultados difusos derivados de las calificaciones de calidad subjetivos proporcionados por un grupo de expertos. La metodología propuesta "Control difuso de procesos" (FPC) comprende una etapa fuera de línea y una etapa en línea. En la etapa de fuera de línea, los expertos asignan calificaciones de calidad a los productos basados en una escala numérica. Las calificaciones numéricas individuales se agregan para formar opiniones colectivas expresadas en forma de números difusos. El conocimiento colectivo aplicado por los expertos a la hora de realizar el proceso de calificación de calidad se adquiere mediante un proceso de análisis de regresión difusa mediante una red neuronal. En la etapa en línea, las dimensiones del producto son medidas, y el modelo de regresión difusa se utiliza para automatizar los juicios de los expertos mediante la asignación de las dimensiones medidas apropiadas a las calificaciones de calidad difusas. Las calificaciones de calidad difusas se representan a continuación en los gráficos de control difuso, cuya construcción y condiciones de fuera-de-control se han desarrollado utilizando la teoría de posibilidad. Los gráficos de control desarrollados no solo monitorean

la tendencia central del proceso, sino también muestran su grado de borrosidad.

[Amirzadeh et al. \(2008\)](#), presentan un gráfico de control difuso multinomial (gráfico-FM) para el seguimiento de un proceso multinomial. Los límites de control del gráfico-FM se obtienen mediante el uso de la distribución multinomial y los grados de pertenencia que se asignan a las distintas categorías. Realizán además una comparación del gráfico-FM con el gráfico- $p$  basado en un proceso de producción de alimentos para mostrar que el uso del gráfico-FM que proponen conduce a mejores resultados.

[Feili y Fekraty \(2010\)](#) plantean un gráfico de control para monitorear la calidad del hilo en una empresa textil con enfoque en la teoría difusa, considerando la calidad en términos de grados de conformidad a diferencia de la conformidad absoluta y no conformidad y lo comparan con el gráfico de control con enfoque probabilístico para determinar el enfoque que presenta un mejor rendimiento.

[Pandurangan \(2011\)](#), a partir del trabajo realizado por [Amirzadeh et al. \(2008\)](#) propone un gráfico de control difuso para un proceso multinomial con tamaño de la muestra variable (VSS). Los límites de control se obtienen mediante la distribución multinomial. Compara además el método propuesto con el gráfico  $p$  convencional.

### **2.4.3. Gráficos de control con enfoque difuso para variables multinomiales multivariantes.**

Se han encontrado en la literatura tres diferentes metodologías para tratar con procesos de atributos multivariantes cuando los datos se presentan en una forma lingüística, propuestas de acuerdo al procedimiento

de calificación de las muestras:

1. *Los artículos se clasifican sucesivamente con respecto a cada una de las características de calidad.*

En este caso, [Taleb y Limam \(2005\)](#) y [Taleb et al. \(2006\)](#) sugieren dos enfoques para la construcción de gráficos de control.

*El primero se basa en la teoría difusa.* El estadístico planteado en el enfoque difuso,  $T_f^2$  similar al estadístico  $T^2$  de Hotteling, se obtiene después de transformar las observaciones difusas en sus valores representativos, para ello utilizan la mediana difusa.

$$T_f^2 = (R - \bar{R})' S^{-1} (R - \bar{R}).$$

donde  $\bar{R}$  es el promedio de los vectores de medias de los valores representativos y  $S$  la matriz de covarianzas correspondientes a las  $m$  muestras preliminares de tamaño  $n$ . Puesto que la distribución de  $T_f^2$  no puede ser identificada directamente, su distribución se obtiene usando el método de remuestreo Bootstrap.

Los pasos necesarios para construir el gráfico de control  $T_f^2$  son los siguientes:

Paso 1: Calcular  $\bar{R}$  y  $S^{-1}$  a partir de observaciones empíricas disponibles.

Paso 2: Generar  $B$  nuevas muestras con reemplazamiento del mismo tamaño a partir de los datos observados.

Paso 3: Calcular el estadístico  $T_{fj}^2 = (R_i - \bar{R})' S^{-1} (R_i - \bar{R})$  para cada nueva muestra  $i$ ,  $i = 1, \dots, B$ .

Paso 4: Fijar límites de control tales que la tasa de falsa alarma sea igual a un valor predefinido.

*El segundo utiliza un enfoque probabilístico.* Aquí el estadístico a

monitorear planteado para el período  $i$  es:

$$W_i^2 = \sum_{j=1}^p Z_{ij}^2,$$

donde

$$Z_{ij}^2 = n_i n_0 \sum_{h=1}^{q_j} \frac{\left( \frac{n_{ijh}}{n_i} - \frac{n_{0jh}}{n_0} \right)^2}{n_{ijh} + n_{0jh}},$$

donde  $i, i = 1, \dots, m$   $n_{ijh}$  y  $n_{0jh}$  son el número de unidades clasificadas mediante la característica de calidad  $j$  en la categoría  $h$  en el período  $i$  y el periodo base 0, respectivamente, y  $n_i$  y  $n_0$  es el tamaño de la muestra de los períodos  $i$  y 0 respectivamente. La distribución de  $Z_{ij}^2$  se aproxima a  $\chi^2(q_j)$ . El estadístico  $W_i^2$  es una combinación lineal de  $p$  características de calidad correlacionadas y su distribución es aproximada por  $\chi^2(v_i)$ , la distribución chi-cuadrado con  $v_i$  grados de libertad. El valor de  $v_i$  puede ser estimado por aproximación de Satterthwaite (Satterthwaite, 1946) como sigue:

$$v_i = (W_i^2)^2 \left( \sum_{j=1}^p \left( Z_{ij}^2 \right)^2 / (q_j - 1) \right)^{-1}.$$

El límite de control superior se determina usando los percentiles de una distribución chi-cuadrado con  $v$  grados de libertad.  $v$  es el número de grados de libertad determinado mediante el reemplazo de  $n_{ijh}$  y  $n_i$  por sus medias  $\bar{n}_{jh}$  y  $\bar{n}$  cuando el proceso esta bajo control.

2. Cada elemento se controla simultáneamente con respecto a todas las características de calidad.

Los elementos de una muestra dada se clasifican en una tabla de contingencia  $q_1 \times q_2 \times \dots \times q_p$ , donde  $q_1, \dots, q_p$  son el número de categorías de las características de calidad,  $Q_1, \dots, Q_p$ . Si  $p$  es el

número de características de calidad e  $I_j$  es el número de categorías para la variable lingüística  $j$ , el número de perfiles de calidad posibles será  $r = \prod_{j=1}^p I_j$  y los datos observados para la muestra del subgrupo se puede mostrar en una tabla de contingencia ( $1 \times r$ ).

En este caso, Taleb (2009) propone dos enfoques:

**El primero** se basa en la teoría de la probabilidad y siguiendo el trabajo de Marcucci (1985) presentan dos casos. Cuando el vector de proporciones  $\pi$  para el perfil de calidad bajo-control es conocido, utilizan el estadístico de Pearson de bondad de ajuste.

$$G^2 = \sum_{l=1}^r \frac{(Y_l - n\pi_l)^2}{n\pi_l}$$

donde  $n$  es el tamaño de la muestra. Cuando el proceso está bajo-control, la distribución asintótica de  $G^2$  es  $\chi^2$  con  $v$  grados de libertad, donde  $v = \prod_{j=1}^p (I_j - 1)$ . Si  $\pi$  se desconoce se utiliza una prueba de homogeneidad de proporciones entre el período base y cada periodo de monitoreo. El estadístico está dado por:

$$X_i^2 = n_i n_0 \sum_{l=1}^r \frac{(p_{il} - p_{0l})^2}{Y_{il} + Y_{0l}}$$

donde  $p_{il}$  y  $p_{0l}$  denotan las frecuencias esperadas correspondientes en las celdas de la tabla de contingencia en el período  $i$  y  $0$ ;  $Y_{il}$  y  $Y_{0l}$  denotan el número de observaciones en el perfil  $l$  para el período  $i$  y  $0$ , respectivamente.  $Y_i$ ,  $n_i$  y  $n_0$  son el tamaño de la muestra de los períodos  $i$  y  $0$ , respectivamente. La distribución de  $X^2$  aproxima una  $\chi^2(v)$ . En ambas situaciones el gráfico de control que tiene un LCS que es un percentil apropiado de la distribución  $\chi^2$ .

**El segundo** enfoque se basa en la teoría difusa. Para cada variable lingüística  $j$  hay  $I_j$  términos lingüísticos para describir la calidad del proceso. El perfil de calidad  $l$  puede ser visto como:

- Caso I: un vector de  $p$  términos lingüístico. Donde cada término

lingüístico  $I_j$  puede ser descrito por un subconjunto difuso  $F_{I_j}$ . En este caso hay  $I_j$  posibles subconjuntos difusos para cada variable lingüística  $j$ .

- Caso II: un término lingüístico. Una muestra  $i$  es representada por un conjunto de  $r$  términos lingüísticos  $L = (L_1, \dots, L_r)$ .

Para cada caso se plantean a su vez, dos procedimientos para la determinación del estadístico de control y su correspondiente límite de control, para más detalles ver [Taleb \(2009\)](#)

### 3. *El enfoque de calificación numérica de Calidad.*

Basados en el trabajo realizado por [Cheng \(2005\)](#), en el que un modelo de regresión difusa se utiliza para automatizar los juicios de expertos mediante la asignación de dimensiones medidas para una adecuada calificación de calidad difusa. [Kumar y Mohapatra \(2012\)](#) diseñan gráficos de control para características de calidad de atributos múltiples subjetivamente estimados. La calidad de la muestra global se calcula mediante la adición interactiva ponderada de los valores difusos asignados a cada característica de calidad. Los gráficos de control se dibujan utilizando las medidas de posibilidad y de necesidad. Este enfoque ayuda a identificar las características de calidad responsables de las situaciones fuera de control.





## Gráfico de Control Multinomial Multivariante con enfoque difuso

El contenido de este capítulo está basado en el artículo: Pastuizaca, M., Carrión, A. y Ruiz, O. (2015). Multivariate multinomial  $T^2$  control chart using fuzzy approach. *International Journal of Production Research*. Vol. 53, Iss. 7 DOI: 10.1080/00207543.2014.983617

### 3.1. Introducción

La calidad de un producto se mide a menudo a través de varias características de calidad generalmente correlacionadas, los gráficos de control multivariantes son una respuesta a la necesidad de control de calidad en este tipo de situaciones. Si éstas características de calidad son cualitativas, a veces ocurre que la calidad del producto es definida por variables lingüísticas. A diferencia del caso de las variables cuantitativas multivariantes el trabajo que se ha realizado para este tipo de variables es

escaso.

Lu (1998) proponen un gráfico de control para procesos de atributos multivariantes en los que las características de calidad de una unidad de producto se clasifican como conforme o no conforme, éste enfoque trata únicamente de características de calidad binomiales. Sin embargo, en muchos casos, la clasificación binaria puede no ser apropiada. De hecho, la calidad de un producto no cambia abruptamente de ser perfecta o no serlo en absoluto; por tanto, hay una necesidad de escalas intermedias para características de calidad lingüísticas tales como la apariencia, la suavidad y el color, que no se pueden expresar numéricamente (Taleb et al., 2006). Para complementar la clasificación binaria, varios niveles intermedios pueden ser expresados mediante el uso de términos lingüísticos. Por ejemplo, una característica de calidad de un producto como “color” se puede clasificar en los términos: “perfecto”, “bueno”, “medio”, “pobre” o “malo”, dependiendo de su incumplimiento de las especificaciones. Esta clasificación será realizada por los expertos, lo que añade incertidumbre al análisis de estabilidad del proceso, puesto que la subjetividad humana juega un papel importante en la definición de las características de calidad. En este caso los gráficos de control clásicos pueden no ser aplicables, puesto que requieren información precisa.

Según Gülbay y Kahraman (2007), la teoría de conjuntos difusos es un medio perfecto para el modelado de la incertidumbre (o imprecisión) derivada de fenómenos mentales que no son aleatorios ni estocásticos. Por tanto, el uso de gráficos de control difusos es inevitable cuando los datos estadísticos en consideración son inciertos o vagos, o la información disponible sobre el proceso está incompleta o incluye subjetividad humana.

En este capítulo, se han integrado los conceptos empleados en el desarrollo del gráfico de control multivariante de Hotelling y los gráficos de control difusos encontrados en la literatura para proponer

una metodología para el diseño de un gráfico de control difuso para características de calidad de tipo multi-atributos correlacionados.

## 3.2. Generalidades

Si bien es cierto, los gráficos de control de calidad multivariantes para variables cuantitativas, tales como el de Hotelling (Hotelling, 1947), MCUSUM (Crosier, 1988) y MEWMA (Lowry et al., 1992) han sido ampliamente estudiados en la literatura. Además, hay un aumento de la eficacia de las técnicas para identificar la causa de una señal de fuera de control (Mason et al. (1995), Runger et al. (1996); Aparisi et al. (2006), etc.). Poco se ha hecho en el caso de los gráficos de control para atributos multivariantes.

La aplicación de los gráficos de control multivariantes para procesos de atributos multinomiales depende del método de muestreo utilizado (Taleb y Limam, 2005). Cuando cada elemento es controlado simultáneamente con respecto a todas las características de calidad, Taleb (2009) propone dos procedimientos para la construcción de gráficos de control para el seguimiento de un proceso multivariante basado en datos lingüísticos multi-dimensionales: El primero se basa en la teoría de la probabilidad y el segundo en la teoría difusa.

Si los elementos se clasifican sucesivamente con respecto a todas las características de calidad, es decir, los elementos de la muestra son clasificados por cada característica de calidad por separado. Taleb et al. (2006) sugieren dos enfoques para la construcción de gráficos de control, ellos desarrollaron dos estadísticos de supervisión, basados en las teorías, difusa y de probabilidad. El primero es similar al estadístico de Hotelling y se basa en los valores representativos de los conjuntos difusos. La

distribución del segundo estadístico, siendo una combinación lineal de variables chi-cuadrado dependientes, se obtiene por aproximación de Satterthwaite.

Cuando las características de calidad del producto son evaluadas por un grupo de expertos quienes las califican de acuerdo a un ranking, Kumar y Mohapatra (2012) integran los conceptos presentados por otros investigadores (como Cheng (2005) que propuso un nuevo enfoque para la elaboración de gráficos de control difuso mediante la asociación de números difusos con las puntuaciones de los expertos sobre las características de calidad) y desarrollan un método para la elaboración de gráficos de control difuso para controlar un producto con múltiples características tipo-atributos que se correlacionan y se miden por los expertos en una escala lingüística, es decir, subjetivamente. La calidad global de la muestra se calcula mediante adición interactiva ponderada de los valores difusos asignados a cada característica de calidad. Los gráficos de control se dibujan utilizando medidas de posibilidad y las medidas de necesidad.

### 3.3. Modelo propuesto

Suponga que  $p$  características de calidad  $Q_1, \dots, Q_j, \dots, Q_p$  correlacionadas son monitoreadas de manera conjunta. Cada característica de calidad  $Q_j$ ,  $j = 1, 2, \dots, p$ , es una variable lingüística que describe la calidad de una unidad de producto mediante un conjunto de términos lingüísticos  $q_{jk}$ ,  $k = 1, 2, \dots, s_j$  donde  $s_j$  es el número de categorías del conjunto de términos de  $Q_j$ . Cada término  $q_{jk}$  se asocia con un subconjunto difuso  $F_{jk}$  descrito por una función de pertenencia  $\mu_{jk}(x)$  que asocia a cada valor  $x$  de la base variable estandarizada ( en términos difusos), un número en el intervalo  $[0, 1]$  . Este número representa el grado en que el valor  $x$  pertenece al subconjunto difuso  $F_{jk}$ . Existen varios

métodos de elaboración y selección de funciones de pertenencia basados esencialmente en datos estadísticos y clasificación (Zadeh, 1978).

Es posible llevar a cabo operaciones aritméticas sobre variables lingüísticas representados como conjuntos difusos mediante la aplicación de las definiciones y técnicas de las matemáticas difusas, que se han desarrollado ampliamente (Wang y Raz, 1990). Estos requieren que la función de pertenencia sea normal (exista al menos un valor de  $x$  cuyo grado de pertenencia sea igual a 1) y convexa. Para una introducción a la aritmética difusa, ver Kaufmann y Gupta (1985). A lo largo del resto de este documento todas las funciones de pertenencia asociadas a los términos lingüísticos se suponen normales, convexas y estandarizadas en el intervalo  $[0,1]$ , donde 0 representa la mejor calidad posible, y 1 representa la peor calidad posible, de manera que para asegurar el cumplimiento de tales condiciones se usaran funciones de pertenencia triangulares cuya gráfica puede verse en la figura 3.1; por tanto cada categoría  $q_{jk}$  de las características de calidad estará representada por un número triangular difuso *Triangular Fuzzy Number* (TFN).

$$F_{jk} = (a_{jk}, b_{jk}, c_{jk}). \quad (3.1)$$

Sus funciones de pertenencia  $\mu_{jk}(x)$  correspondientes estarían definidas como:

$$\mu_{jk}(x) = \begin{cases} 0, & x \leq a_{jk} \\ \frac{1}{b_{jk}-a_{jk}}(x - a_{jk}), & a_{jk} \leq x \leq b_{jk} \\ \frac{1}{b_{jk}-c_{jk}}(x - c_{jk}), & b_{jk} \leq x \leq c_{jk} \\ 0, & x \geq c_{jk} \end{cases} \quad (3.2)$$

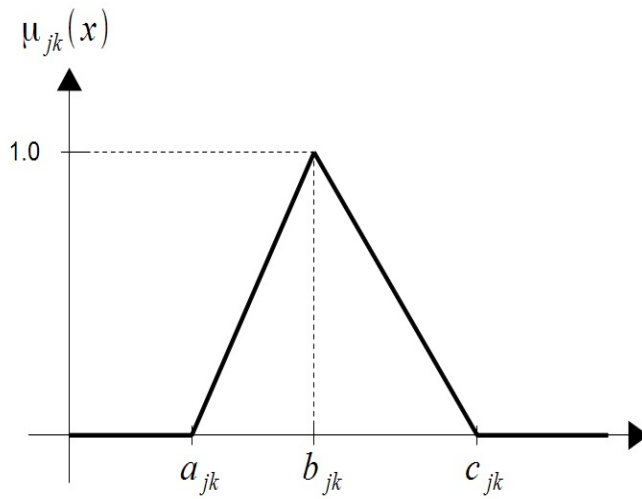


Figura 3.1: Función de pertenencia para un número triangular difuso

Una muestra  $A_i$  de  $n$  observaciones puede entonces representarse mediante el conjunto:

$$A_i = \left\{ \left\{ (F_{11}, n_{i11}), \dots, (F_{1q_1}, n_{i1s1}) \right\}, \dots, \left\{ (F_{p1}, n_{1p1}), \dots, (F_{ppp}, n_{ipsp}) \right\} \right\}$$

donde  $n_{ijk}$  es el número de observaciones de la característica de calidad  $Q_j$  de la muestra  $A_i$  clasificadas mediante la variable lingüística  $q_{jk}$  y se cumple que:

$$\sum_{k=1}^{s_j} n_{ijk} = n, \forall j, j = 1, 2, \dots, p.$$

Cada característica de calidad  $Q_j$  a medir en una muestra  $A_i$  es una variable multinomial con  $s_j$  categorías, es decir

$$Q_{ij} \sim M(p_{ij1}, \dots, p_{ijk}, \dots, p_{ijs_j})$$

donde

$$p_{ijk} = \frac{n_{ijk}}{n} \quad (3.3)$$

Usando aritmética difusa, es posible combinar los subconjuntos difusos de cada uno de los atributos. Kaufmann y Gupta (1985) mostraron que la multiplicación de un número difuso triangular (TFN)  $\mathbf{T}$  por un número real  $k$  es un TFN y que la adición de dos TFN,  $\mathbf{T}$  y  $\mathbf{S}$  lo es también, por tanto, una combinación lineal de TFNs da como resultado un TFN. Por ejemplo, si  $\mathbf{T}$  y  $\mathbf{S}$  están representados por las ternas  $(t_1, t_2, t_3)$  y  $(s_1, s_2, s_3)$  respectivamente, entonces, una combinación lineal  $\mathbf{C} = k_1\mathbf{T} + k_2\mathbf{S}$  deberá ser representada por la terna  $(k_1t_1 + k_2s_1, k_1t_2 + k_2s_2, k_1t_3 + k_2s_3)$ . Por tanto mediante el supuesto de que las variables difusas  $F_{jk}$  correspondientes a las  $k$  categorías de la característica  $j$  son TFNs dados por la ecuación 3.1 entonces, usando combinación lineal se puede representar la característica de calidad  $Q_j$  de la muestra  $i$  mediante un único número difuso dado por

$$F_{ij} = \frac{1}{n} \sum_{k=1}^{s_j} n_{ijk} F_{jk} = \frac{1}{n} \left( \sum_{k=1}^{s_j} n_{ijk} a_{jk}, \sum_{k=1}^{s_j} n_{ijk} b_{jk}, \sum_{k=1}^{s_j} n_{ijk} c_{jk} \right) \quad (3.4)$$

que será también un TFN y se puede escribir como

$$F_{ij} = (a_{ij}, b_{ij}, c_{ij})$$

La muestra  $A_i$  se puede entonces expresar como

$$A_i = \{F_{i1}, \dots, F_{ip}\}.$$

En la literatura, los gráficos de control difusos se han desarrollado mediante la conversión de los conjuntos difusos asociados con variables lingüísticas en escalares conocidos como valores representativos (Raz y Wang, 1990). Esta conversión, que facilita el trazado de las observaciones en el gráfico, se puede hacer de varias formas; los métodos más conocidos son cuatro, que son similares a las medidas de tendencia central utilizadas en estadística descriptiva: modo difuso, nivel de rango medio difuso, mediana difusa y media difusa. Cabe señalar que no existe ninguna base teórica que soporte cualquiera de estos métodos de manera específica. La selección entre ellos debe basarse principalmente en la facilidad de cálculo o la preferencia del usuario (Gülbay y Kahraman, 2007).

El método de defusificación empleado para obtener el valor representativo es el de la *media difusa* basado en Zadeh (1965), definido por

$$f_{avg} = Av(x : F) = \frac{\int_0^1 x\mu_F(x) dx}{\int_0^1 \mu_F(x) dx} \quad (3.5)$$

Si consideramos la función triangular de pertenencia  $\mu_{ij}(x)$  correspondiente a  $F_{ij}$ , de forma similar a la dada en la ecuación 3.2, y la reemplazamos en la ecuación 3.5, el valor representativo  $R_{ij}$  para la



característica  $Q_j$  de la muestra  $A_i$  estaría dado por

$$R_{ij} = \frac{\int_0^1 x \mu_{ij}(x) dx}{\int_0^1 \mu_{ij}(x) dx} = \frac{\int_{a_{ij}}^{b_{ij}} \frac{x(x - a_{ij})}{b_{ij} - a_{ij}} dx + \int_{b_{ij}}^{c_{ij}} \frac{x(x - c_{ij})}{b_{ij} - c_{ij}} dx}{\frac{c_{ij} - a_{ij}}{2}}$$

$$R_{ij} = \frac{a_{ij} + b_{ij} + c_{ij}}{3} \quad (3.6)$$

reemplazando en la ecuación 3.6 los valores de  $a_{ij}$ ,  $b_{ij}$  y  $c_{ij}$  correspondientes, dados por la ecuación 3.4, se obtiene

$$R_{ij} = \frac{1}{3n} \left( \sum_{k=1}^{s_j} n_{ijk} a_{jk}, \sum_{k=1}^{s_j} n_{ijk} b_{jk}, \sum_{k=1}^{s_j} n_{ijk} c_{jk} \right)$$

$$R_{ij} = \frac{1}{3n} \sum_{k=1}^{s_j} n_{ijk} (a_{jk} + b_{jk} + c_{jk}) \quad (3.7)$$

si llamamos al  $vr_{jk}$  valor dado por  $vr_{jk} = \frac{a_{jk} + b_{jk} + c_{jk}}{3}$  que como podemos ver, sería el valor representativo correspondiente a la función de pertenencia de la categoría  $k$  de la característica de calidad  $j$  que se obtiene usando el método de defusificación de la *media difusa* y lo reemplazamos en 3.7, se tiene

$$R_{ij} = \frac{1}{n} \sum_{k=1}^{s_j} n_{ijk} v r_{jk} \quad (3.8)$$

que será finalmente la expresión mediante la cual la variable multinomial correspondiente a cada característica de calidad  $Q_j$  de una muestra  $i$  se convertirá en un valor numérico. Como podemos ver el valor representativo a usar resulta ser una combinación lineal de los  $s_j$  valores  $n_{ijk}$  correspondientes al número de productos clasificados con la categorías  $q_{jk}$  de la característica de calidad  $Q_j$  en la muestra  $A_i$ .

Luego la muestra  $A_i$  de  $n$  observaciones estará representada por el vector

$$R_i = (R_{i1}, \dots, R_{ij}, \dots, R_{ip})' \quad (3.9)$$

Donde  $R_{ij}$  tiene una distribución normal cuya media esta dada por

$$\begin{aligned} \mu_{ij} &= E(R_{ij}) = E\left(\frac{1}{n} \sum_{k=1}^{s_j} n_{ijk} v r_{jk}\right) \\ &= \frac{1}{n} \sum_{k=1}^{s_j} E(n_{ijk} v r_{jk}) = \frac{1}{n} \sum_{k=1}^{s_j} E(n_{ijk}) v r_{jk} \\ &= \frac{1}{n} \sum_{k=1}^{s_j} n p_{ijk} v r_{jk} \end{aligned}$$

pero  $p_{ijk} = \frac{n_{ijk}}{n}$  (ecuación 3.3) luego

$$\mu_{ij} = \frac{1}{n} \sum_{k=1}^{s_j} n_{ijk} v r_{jk} = R_{ij}$$

y su varianza esta dada por:

$$\begin{aligned} \text{var}(R_{ij}) &= \text{var} \left( \frac{1}{n} \sum_{k=1}^{s_j} n_{ijk} v r_{jk} \right) \\ &= \frac{1}{n^2} \text{var} \left( n_{ij1} v r_{j1} + n_{ij2} v r_{j2} + \cdots + n_{ijs_j} v r_{js_j} \right) \\ &= \frac{1}{n} \sum_{k=1}^{s_j} p_{ijk} (1 - p_{ijk}) v r_{jk}^2 - \frac{2}{n} \sum_{k=1}^{s_j-1} \left[ p_{ijk} v r_{jk} \left( \sum_{l=k+1}^{s_j} p_{ijl} v r_{jl} \right) \right] \end{aligned}$$

El proceso para obtener la matriz de covarianzas es muy complejo, por tal motivo se requiere estimarla, el estimador a utilizar se analiza en el siguiente apartado.

El conjunto de valores representativos de las  $p$  características de calidad está representado por el vector  $\mathbf{R}_i$  de dimensión  $(p \times 1)$  dado en la ecuación 3.9. Al igual que Taleb et al. (2006) el estadístico de prueba a ser representado en el gráfico de control para cada muestra es

$$T_i^2 = (\mathbf{R}_i - \boldsymbol{\mu}_R)' \boldsymbol{\Sigma}_R^{-1} (\mathbf{R}_i - \boldsymbol{\mu}_R) \quad (3.10)$$

donde  $\boldsymbol{\mu}_R = (\mu_1, \dots, \mu_p)$  es el vector de medias bajo-control para cada característica de calidad y  $\boldsymbol{\Sigma}_R$  es la matriz de covarianzas de las características de calidad.

### 3.3.1. Estimación de los parámetros $\mu_R$ y $\Sigma_R$

Puesto que  $R_i$  tiene una distribución normal multivariante  $p$ -dimensional con vector de medias  $\mu_R$  y matriz de varianzas y covarianzas  $\Sigma_R$  cuyo cálculo es muy complejo, estos parámetros serán estimados por  $\bar{R}$  y  $S$  respectivamente a partir de la matriz de valores representativos  $R$  mostrada en la ecuación 3.15.

Según Chou et al. (2001), en la Fase I, si  $R_1, R_2, \dots, R_m$  representa un conjunto de datos históricos (HDS) de  $m$  observaciones con vector de medias de la muestra  $\bar{R}$  y matriz de varianzas y covarianzas de la muestra  $S$ , el valor del estadístico  $T^2$  para  $R_i$  está dado por

$$T_i^2 = (R_i - \bar{R})' S^{-1} (R_i - \bar{R}); i = 1, 2, \dots, m \quad (3.11)$$

estos valores pueden ser representados en un gráfico de control  $T^2$  de Hotelling con un límite de control superior (LCS) dado por

$$LCS_1 = \frac{(m-1)^2}{m} B_{\left(\frac{p}{2}, \frac{m-p-1}{2}; \alpha\right)} \quad (3.12)$$

donde  $B_{(a,b;\alpha)}$  es el  $(1-\alpha)$ -ésimo cuantil de una distribución beta  $B_{(a,b)}$ . Dado un HDS de tamaño  $m$  y una sola futura observación  $Y$  en la Fase II, el valor del estadístico

$$T_i^2 = (Y - \bar{R})' S^{-1} (Y - \bar{R}) \quad (3.13)$$

es representado en un gráfico de control  $T^2$  de Hotelling con límite de

control superior dado por:

$$LCS_2 = \frac{p(m^2 - 1)}{m(m - p)} F_{(p, m-p; \alpha)} \tag{3.14}$$

donde  $F_{(p, m-p; \alpha)}$  es el  $(1 - \alpha)$ -ésimo cuantil de una distribución  $F_{(p, m-p)}$ .

Los parámetros  $\mu_R$  y  $\Sigma_R$  deberán ser entonces estimados por  $\bar{R}$  y  $S$  a partir del análisis del HDS formado por muestras preliminares de tamaño  $n$ , tomadas cuando el proceso está bajo-control. Supongamos que  $m$  de tales muestras están disponibles. La tabla 3.1 ilustra la forma en la que los datos de las  $m$  muestras se almacenan, una vez realizada la clasificación por categorías de las diferentes características de calidad a monitorear.

**Tabla 3.1:** Tabla de datos para el análisis preliminar

	$Q_1$			...	$Q_j$				...	$Q_p$			
	$q_{11}$	...	$q_{1s_1}$	...	$q_{j1}$	...	$q_{jk}$	...	$q_{js_j}$	...	$q_{p1}$	...	$q_{ps_p}$
$A_1$	$n_{111}$	...	$n_{11s_1}$	...	$n_{1j1}$	...	$n_{1jk}$	...	$n_{1js_j}$	...	$n_{1p1}$	...	$n_{1ps_p}$
$\vdots$	$\vdots$		$\vdots$		$\vdots$		$\vdots$		$\vdots$		$\vdots$		$\vdots$
$A_i$	$n_{i11}$	...	$n_{i1s_1}$	...	$n_{ij1}$	...	$n_{ijk}$	...	$n_{ijs_j}$	...	$n_{ip1}$	...	$n_{ips_p}$
$\vdots$	$\vdots$		$\vdots$		$\vdots$		$\vdots$		$\vdots$		$\vdots$		$\vdots$
$A_m$	$n_{m11}$	...	$n_{m1s_1}$	...	$n_{mj1}$	...	$n_{mjk}$	...	$n_{mjs_j}$	...	$n_{mp1}$	...	$n_{mps_p}$

Los valores representativos para las  $p$  características de calidad de las  $m$  muestras pueden ser expresados como:

$$\mathbf{R} = \begin{pmatrix} R_{11} & \cdots & R_{1j} & \cdots & R_{1p} \\ R_{21} & \cdots & R_{2j} & \cdots & R_{2p} \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ R_{m1} & \cdots & R_{mj} & \cdots & R_{mp} \end{pmatrix} \quad (3.15)$$

cuya columna  $R_j$  se puede obtener a partir de

$$R_j = \begin{pmatrix} R_{1j} \\ R_{2j} \\ \vdots \\ R_{mj} \end{pmatrix} = \frac{1}{n} \begin{pmatrix} n_{1j1} & n_{1j2} & \cdots & n_{1js_j} \\ n_{2j1} & n_{2j2} & \cdots & n_{2js_j} \\ \vdots & \vdots & \ddots & \vdots \\ n_{mj1} & n_{mj2} & \cdots & n_{mjs_j} \end{pmatrix} \begin{pmatrix} vr_{j1} \\ vr_{j2} \\ \vdots \\ vr_{js_j} \end{pmatrix}$$

donde los  $vr_{jk}$  son los valores representativos de los subconjuntos difusos  $F_{jk}$  correspondientes a las  $s_j$  categorías de la característica de calidad  $Q_j$

El vector de medias  $\mu_R$  será estimado por:

$$\bar{\mathbf{R}} = (\bar{R}_1, \dots, \bar{R}_j, \dots, \bar{R}_p), \quad (3.16)$$

donde  $\bar{R}_j = \frac{1}{m} \sum_{i=1}^m R_{ij}$ , y  $R_{ij}$  es el valor representativo del subconjunto difuso asociado con la  $j$ -ésima característica de calidad de la  $i$ -ésima muestra.

Un problema importante en el caso de las observaciones individuales es la estimación de la matriz de varianzas y covarianzas  $\Sigma_R$ . Sullivan y

Woodall (1996) ofrecen una excelente discusión y análisis de este problema, y comparan varios estimadores. Ellos mostraron que un gráfico más eficaz estima la matriz de covarianzas de las diferencias vectoriales entre observaciones sucesivas, como sugieren Holmes y Mergen (1993), bajo la presunción de que las observaciones sucesivas tenderán a tener el mismo, o casi el mismo, vector de medias, que es una estimación más robusta de la matriz de covarianzas bajo-control para observaciones individuales. Holmes y Mergen (1993) denominan a este estimador diferencia sucesiva cuadrática media (MSSD). Bajo este enfoque, el estimador de la varianza de  $R_j$  estará dado por

$$S_j^2 = \frac{1}{2(m-1)} \sum_{i=1}^m (R_{i+1j} - R_{ij})^2, \quad (3.17)$$

y el de la covarianza entre los valores representativos de las características de calidad  $Q_j$  y  $Q_h$  estará dado por

$$S_{jh} = \frac{1}{2(m-1)} \sum_{i=1}^m (R_{i+1j} - R_{ij}) (R_{i+1h} - R_{ih}), \text{ para } j \neq h \quad (3.18)$$

La matriz de varianzas y covarianzas de la muestra,  $S$  de  $p \times p$ , se expresa entonces como:

$$S = \begin{pmatrix} S_1^2 & S_{12} & \cdots & S_{1p} \\ S_{21} & S_2^2 & \cdots & \vdots \\ \vdots & \vdots & \ddots & \vdots \\ S_{p1} & S_{p2} & \cdots & S_{pp}^2 \end{pmatrix}.$$

### 3.3.2. Interpretación de señales de *fuera-de-control*

Una dificultad que se encuentra con cualquier gráfico de control multivariante es la interpretación práctica de una señal de fuera de control. Para la interpretación de la señal de fuera-de-control emplearemos el procedimiento planteado en [Montgomery \(2009\)](#) para el diagnóstico de una señal de fuera de control que consiste en descomponer el estadístico  $T^2$  en componentes que reflejen la contribución de cada variable individual. Si  $T^2$  es el valor actual del estadístico, y  $T_{(i)}^2$  es el valor del estadístico de todas las variables del proceso excepto la  $i$ -ésima, [Runger et al. \(1996\)](#) mostraron que

$$d_i = T^2 - T_{(i)}^2, \quad i = 1, \dots, p \quad (3.19)$$

es un indicador de la contribución relativa de la variable  $i$ -ésima al estadístico global. Cuando se genera una señal de fuera-de-control, se recomienda calcular los valores de  $d_i$  ( $i = 1, 2, \dots, p$ ) y centrar la atención en la o las variables para las que las  $d_i$  son relativamente grandes.

## 3.4. Ejemplos de Aplicación

### 3.4.1. Ejemplo 1: Proceso de producción de comida congelada

Para mostrar la aplicación del gráfico propuesto, utilizaremos como primer ejemplo, el proporcionado en el artículo de [Taleb et al. \(2006\)](#) de una industria procesadora de alimentos en la que se miden tres características de calidad conjuntas: apariencia ( $Q_1$ ), color ( $Q_2$ ) y sabor



$(Q_3)$ ; cuyos conjuntos de términos lingüísticos están dados por:

$$T(Q_1) = \{q_{11}, q_{12}, q_{13}\} = \{buena, mediana, pobre\}$$

$$T(Q_2) = \{q_{21}, q_{22}, q_{23}\} = \{estándar, aceptable, rechazado\}$$

$$T(Q_3) = \{q_{31}, q_{32}, q_{33}, q_{34}\} = \{perfecto, bueno, mediano, pobre\}$$

La figura 3.2 muestra las funciones de pertenencia con estos tres conjuntos de términos, estas funciones pueden ser representados por sus números triangulares difusos correspondientes:

$$F_{11} = (0, 0, 0.25) \quad F_{12} = (0, 0.25, 0.75) \quad F_{13} = (0.25, 1, 1)$$

$$F_{21} = (0, 0, 0.5) \quad F_{22} = (0, 0.5, 0.75) \quad F_{23} = (0.5, 1, 1)$$

$$F_{31} = (0, 0, 0.25) \quad F_{32} = (0, 0.25, 0.75) \quad F_{33} = (0.25, 0.75, 1) \quad F_{34} = (0.75, 1, 1)$$

Los datos se muestran en la tabla 3.2

### Gráfico de Control Fase I

A partir de estos datos se calculan los valores representativos y los correspondientes valores del estadístico  $T^2$  para cada muestra y se resumen en la tabla 3.3. La matriz de covarianzas para las 20 muestras se obtiene usando las ecuaciones 3.17 y 3.18 y su inversa esta dada por:

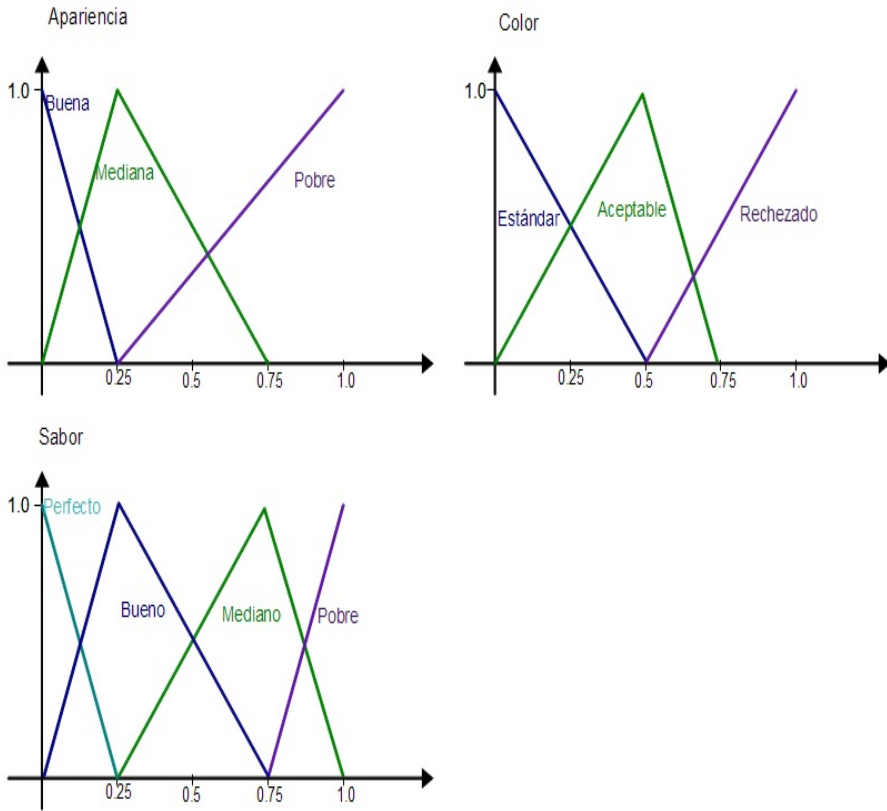


Figura 3.2: Funciones de pertenencia (Taleb et al.(2006))

$$S^{-1} = \begin{pmatrix} 29825.62 & -13398.86 & -649.52 \\ -13398.86 & 25060.69 & 4826.80 \\ -649.52 & 4826.80 & 19917.60 \end{pmatrix}.$$

el estimador del vector de medias resultante es:

**Tabla 3.2:** Datos del proceso de producción de comida congelada  
(Taleb et al.(2006))

<i>muestra</i>	<i>q</i> <sub>11</sub>	<i>q</i> <sub>12</sub>	<i>q</i> <sub>13</sub>	<i>q</i> <sub>21</sub>	<i>q</i> <sub>22</sub>	<i>q</i> <sub>23</sub>	<i>q</i> <sub>31</sub>	<i>q</i> <sub>32</sub>	<i>q</i> <sub>33</sub>	<i>q</i> <sub>34</sub>
1	210	7	3	206	9	5	167	48	3	2
2	211	6	3	207	8	5	176	42	2	0
3	206	9	5	202	12	6	163	55	2	0
4	211	5	4	207	8	5	163	51	5	1
5	203	16	1	194	18	8	175	45	0	0
6	210	6	4	206	9	5	174	44	1	1
7	208	7	5	204	9	7	174	40	5	1
8	207	7	6	204	9	7	169	46	3	2
9	206	7	7	202	9	9	169	48	2	1
10	186	25	9	200	12	8	169	48	3	0
11	196	13	11	196	13	11	163	46	10	1
12	203	12	5	200	13	7	167	44	9	0
13	203	9	8	198	11	11	174	42	3	1
14	202	9	9	198	11	11	174	40	6	0
15	209	6	5	207	9	4	172	42	5	1
16	210	3	7	205	5	1	172	44	4	0
17	205	11	4	201	13	6	172	45	2	1
18	210	6	4	206	8	6	169	48	2	1
19	206	13	4	203	13	4	172	46	0	2
20	206	12	2	202	14	4	169	46	5	0

$$\bar{R} = (0.110, 0.200, 0.147)$$

El límite de control superior para el gráfico de fase I, se calcula con la ecuación 3.12 para un error tipo I(  $\alpha$ ) = 0.05 ,  $m = 20$  y  $p = 3$ .

$$LCS_1 = \frac{(m-1)^2}{m} B_{\left(\frac{p}{2}, \frac{m-p-1}{2}; \alpha\right)} = \frac{19^2}{20} 0.3778 = 6.82$$

Los valores representativos y el estadístico  $T^2$  se muestran en la tabla 3.3.

**Tabla 3.3:** Valores representativos y estadístico  $T^2$

muestra	$R_1$	$R_2$	$R_3$	$T^2$
1	0.100	0.192	0.153	2.597
2	0.099	0.191	0.136	6.091
3	0.109	0.198	0.151	0.279
4	0.101	0.191	0.158	3.763
5	0.104	0.211	0.134	7.692
6	0.102	0.192	0.140	3.343
7	0.106	0.198	0.146	0.351
8	0.109	0.198	0.151	0.274
9	0.113	0.204	0.147	0.355
10	0.139	0.205	0.146	22.099
11	0.131	0.215	0.166	19.71
12	0.112	0.203	0.157	2.316
13	0.118	0.213	0.145	3.102
14	0.121	0.213	0.145	3.710
15	0.105	0.189	0.148	2.170
16	0.108	0.203	0.144	0.609
17	0.108	0.200	0.144	0.405
18	0.102	0.194	0.147	1.444
19	0.107	0.194	0.143	1.349
20	0.103	0.195	0.149	1.123
$\bar{R}$	0.110	0.200	0.147	

El gráfico de control que se obtiene se muestra en la figura 3.3. Se identifican 3 posibles causas asignables, se eliminan y se calculan los parámetros para el gráfico de fase II.

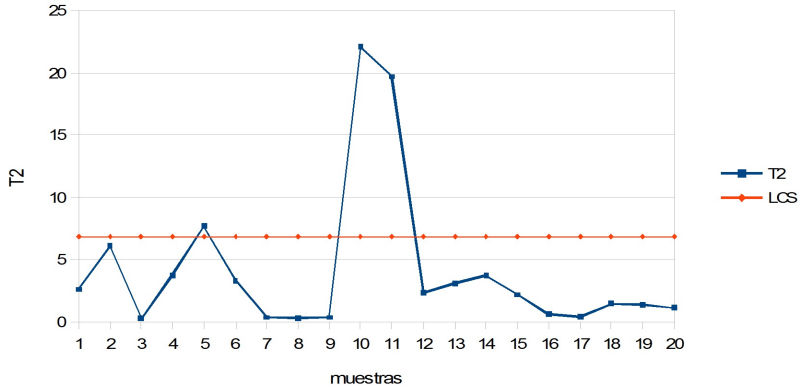


Figura 3.3: Gráfico de Control  $T^2$ MMD Fase I

## Gráfico de Control Fase II

Después de eliminar las muestras consideradas fuera-de-control debido a causas asignables, se obtienen los parámetros  $\bar{R}$  y  $S^{-1}$  que serán usados para calcular el estadístico  $T^2$  para el gráfico de Fase II para las muestras adicionales. Los nuevos parámetros son:

$$S^{-1} = \begin{pmatrix} 238033.45 & -148759.62 & -100416.03 \\ -148759.62 & 124035.97 & 11579.35 \\ -100416.03 & 11579.35 & 24350.10 \end{pmatrix}.$$

$$\bar{R} = (0.107, 0.198, 0.147)$$

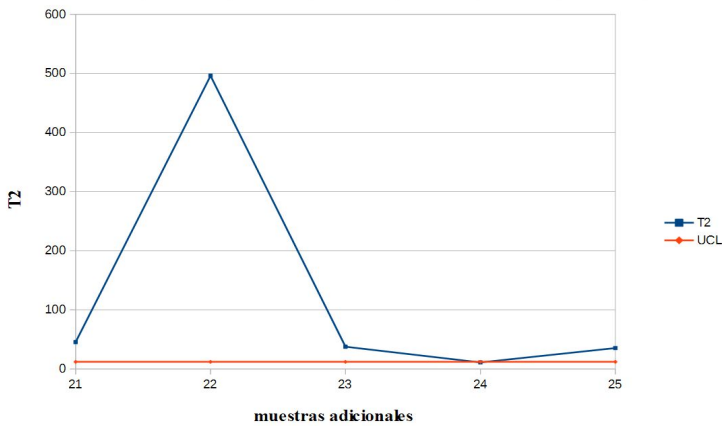
El límite de control superior para esta fase, se calcula con la ecuación 3.14 para un error tipo I ( $\alpha$ ) = 0.05 ,  $m = 17$  y  $p = 3$

$$LCS_2 = \frac{p(m^2 - 1)}{m(m - p)} F_{(p, m-p; \alpha)} = \frac{3(17^2 - 1)}{17(17 - 3)} 3.344 = 12.139$$

Los datos de las muestras adicionales se muestran en las columnas 2 a 11 de la tabla 3.4. A partir de estos datos adicionales se calculan los valores representativos y los correspondientes estadísticos  $T^2$  para cada muestra y se resumen en las columnas 12 a 15 de la misma tabla. A continuación, tenemos el gráfico de control de Fase II correspondiente que se muestra en la Figura 3.4

**Tabla 3.4:** Muestras adicionales (Taleb et al.(2006)), valores representativos y estadísticos  $T^2$  correspondientes

muestra	$q_{11}$	$q_{12}$	$q_{13}$	$q_{21}$	$q_{22}$	$q_{23}$	$q_{31}$	$q_{32}$	$q_{33}$	$q_{34}$	$R_1$	$R_2$	$R_3$	$T^2$
21	202	10	8	204	11	5	169	44	5	2	0.119	0.194	0.154	45.56
22	184	25	11	206	12	2	174	44	1	1	0.145	0.186	0.140	495.96
23	208	7	5	196	13	11	174	44	1	1	0.106	0.215	0.140	37.76
24	206	6	8	196	13	11	174	40	5	1	0.114	0.215	0.146	11.21
25	210	2	8	198	12	10	165	44	1	10	0.110	0.211	0.173	35.48



**Figura 3.4:** Gráfico de Control  $T^2$ MMD Fase II

### Interpretación de las señales de fuera-de-control

La figura 3.4 muestra que el proceso está fuera de control cuando se toman las muestras 21, 22, 23 y 25. Para determinar qué variables son las responsables de la salida de control en cada uno de los tres casos, se calculan los correspondientes  $d_i$  y se muestran en la tabla 3.5.

**Tabla 3.5:** Valores  $d_i$  para la interpretación de señales de *fuera-de-control*

<i>muestra</i>	$T^2$	$T^2_{(1)}$	$T^2_{(2)}$	$T^2_{(3)}$	$d_1$	$d_2$	$d_3$
21	45.56	1.31	9.82	45.56	44.25	36.74	0.01
22	495.96	6.44	84.17	476.15	489.53	411.79	19.81
23	37.76	8.74	1.38	37.74	29.02	36.39	0.02
25	35.48	25.53	17.46	11.14	9.95	18.03	24.35

A partir de la tabla 3.5 vemos que la variable con mayor contribución a la señal de fuera-de-control detectada cuando se tomaron las muestras 21 y 22, es la variable  $Q_1$  (apariencia), ya que en ambos casos es la variable con el valor más alto de  $d_i$ . Mientras que en el caso de la muestra 23 la mayor contribuyente es la variable  $Q_2$  (color) y para la muestra 25 es la variables  $Q_3$  (sabor).

#### 3.4.2. Ejemplo 2: Proceso de producción de porcelana

En este ejemplo hemos utilizado los datos publicados por [Taleb \(2009\)](#) de una industria de elaboración de la porcelana en la que se controlan tres características de calidad de manera conjunta: apariencia ( $Q_1$ ), translucidez ( $Q_2$ ) y Blancura ( $Q_3$ ), cuyos conjuntos de términos lingüísticos son:

$$T(Q_1) = \{q_{11}, q_{12}, q_{13}\} = \{\text{estándar}, \text{segunda clase}, \text{tercera clase}\}$$

$$T(Q_2) = \{q_{21}, q_{22}, q_{23}\} = \{\text{traslucidez perfecta}, \text{traslucidez buena}, \text{no traslúcida}\}$$

$$T(Q_3) = \{q_{31}, q_{32}, q_{33}, q_{34}\} = \{\text{alta}, \text{media}, \text{pobre}\}$$

A diferencia del primer ejemplo, en [Taleb \(2009\)](#) las variables se midieron en forma simultánea, por tanto, para aplicar la metodología propuesta a estos datos hemos hecho un recuento de las variables por categoría de forma independiente. Al igual que el ejemplo anterior las funciones de pertenencia pueden ser representadas por sus números triangulares difusos correspondientes:

$$F_{11} = (0, 0, 0.4) \quad F_{12} = (0.2, 0.4, 0.5) \quad F_{13} = (0.6, 1, 1)$$

$$F_{21} = (0, 0, 0.4) \quad F_{22} = (0.3, 0.5, 0.7) \quad F_{23} = (0.6, 1, 1)$$

$$F_{31} = (0, 0, 0.4) \quad F_{32} = (0.4, 0.6, 0.8) \quad F_{33} = (0.6, 1, 1)$$

Los datos del proceso de porcelana se muestran en las columnas 2 a 10 de la [Tabla 3.6](#).

### Gráfico de Control Fase I

A partir de estos datos se calculan los valores representativos y los valores estadísticos correspondientes  $T^2$  para cada muestra y se resumen en las columnas 11 a 14 de misma tabla. El vector de valores



representativos (estimador del vector de medias) y la inversa de la matriz de covarianzas correspondientes se obtienen a partir las 23 muestras proporcionadas utilizando las ecuaciones 3.16, 3.17 y 3.18 respectivamente, el resultado que se obtiene es:

$$S^{-1} = \begin{pmatrix} 4656.33 & 2717.10 & 74.98 \\ 2717.10 & 5884.70 & 1982.79 \\ 74.98 & 1982.79 & 3957.63 \end{pmatrix}.$$

$$\bar{R} = (0.332, 0.379, 0.225)$$

El límite de control superior para el gráfico de fase I, para un error tipo I ( $\alpha$ ) = 0.05 ,  $m = 23$  y  $p = 3$  esta dado por:

$$LCS_1 = \frac{(m-1)^2}{m} B_{\left(\frac{p}{2}, \frac{m-p-1}{2}, \alpha\right)} = \frac{22^2}{23} 0.33056 = 6.956$$

El gráfico de control correspondiente se muestra en la Figura 3.5 Como se puede apreciar en el gráfico, se encontró que la muestra 11 estaba fuera del límite de control superior, esta muestra se removi6 y se calcularon los parámetros para el gráfico de fase II con las muestras restantes.

### **Análisis del proceso de porcelana, Fase II**

Después de despejar la muestra considerada fuera de control debido a causas asignables, obtenemos los parámetros  $\bar{R}$  y  $S^{-1}$  que serán utilizados en el gráfico de Fase II para monitorear las muestras adicionales. La nueva inversa de la matriz de covarianza y el estimador del vector de

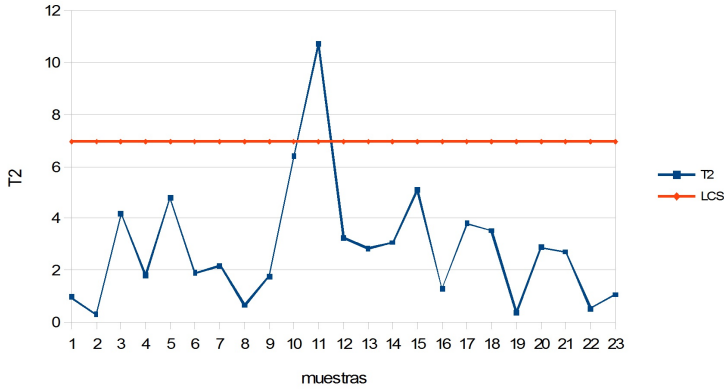
**Tabla 3.6:** Datos del proceso de producción de porcelana (estimados a partir de Taleb (2009)), valores representativos y estadísticos  $T^2$  correspondientes

<i>muestra</i>	$q_{11}$	$q_{12}$	$q_{13}$	$q_{21}$	$q_{22}$	$q_{23}$	$q_{31}$	$q_{32}$	$q_{33}$	$R_1$	$R_2$	$R_3$	$T^2$
1	107	66	27	101	91	8	161	27	12	0.320	0.375	0.227	0.95
2	100	71	29	96	99	5	165	24	11	0.334	0.383	0.217	0.29
3	108	57	35	101	96	3	150	30	20	0.338	0.368	0.261	4.16
4	106	65	29	103	89	8	162	32	6	0.326	0.370	0.214	1.78
5	106	57	37	110	87	3	159	33	8	0.345	0.347	0.223	4.78
6	119	49	32	95	100	5	171	18	11	0.316	0.385	0.206	1.88
7	97	66	37	104	91	5	161	27	12	0.357	0.364	0.227	2.16
8	106	63	31	94	101	5	164	19	17	0.331	0.387	0.230	0.63
9	116	53	31	98	94	8	171	17	12	0.318	0.382	0.208	1.74
10	113	56	31	85	109	6	155	33	12	0.322	0.410	0.238	6.39
11	97	64	39	90	101	9	166	29	5	0.362	0.402	0.205	10.72
12	113	59	28	86	106	8	169	21	10	0.315	0.410	0.208	3.23
13	101	59	40	101	94	5	156	31	13	0.359	0.371	0.238	2.82
14	107	63	30	86	109	5	169	21	10	0.327	0.406	0.208	3.06
15	112	59	29	109	83	8	157	27	16	0.318	0.356	0.241	5.09
16	107	60	33	97	99	4	156	28	16	0.334	0.379	0.243	1.29
17	112	55	33	109	84	7	156	26	18	0.328	0.355	0.247	3.79
18	119	56	25	93	99	8	165	27	8	0.300	0.394	0.212	3.52
19	105	65	30	100	92	8	165	24	11	0.330	0.377	0.217	0.35
20	105	64	31	104	91	5	152	30	18	0.332	0.364	0.254	2.86
21	92	75	33	108	85	7	165	21	14	0.354	0.357	0.223	2.70
22	108	56	36	98	94	8	159	31	10	0.340	0.382	0.227	0.52
23	108	62	30	93	100	7	169	21	10	0.326	0.392	0.208	1.05

medias calculados a partir de las muestras consideradas bajo-control son respectivamente:

$$S^{-1} = \begin{pmatrix} 6844.75 & 3059.01 & -1082.66 \\ 3059.01 & 5658.54 & 1574.57 \\ -1082.66 & 1574.57 & 4151.20 \end{pmatrix}.$$

$$\bar{R} = (0.330, 0.378, 0.226)$$



**Figura 3.5:** Gráfico de Control  $T^2$ MMD Fase I: Proceso de producción de porcelana

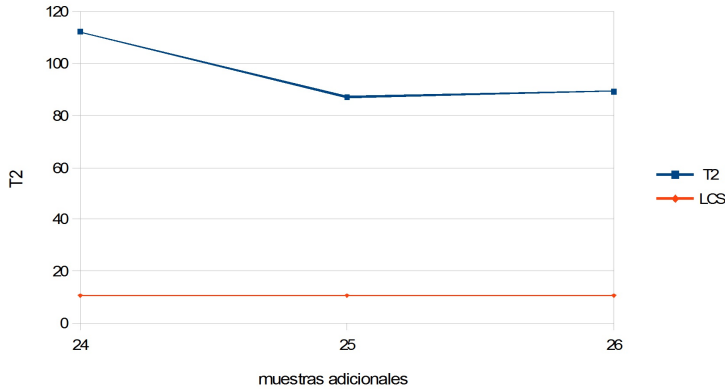
El límite de control superior para esta fase, para un error tipo  $I(\alpha) = 0.05$ ,  $m = 22$  y  $p = 3$

$$LCS_2 = \frac{p(m^2 - 1)}{m(m - p)} F_{(p, m-p; \alpha)} = \frac{3(22^2 - 1)}{22(22 - 3)} 3.127 = 10.841$$

Las muestras adicionales se muestran en las columnas 2 a 10 de la tabla 3.7. A partir de estos datos adicionales, valores representativos y los correspondientes valores del estadístico  $T^2$  se calculan para cada muestra y se resumen en las columnas 11 a 14 de la Tabla 3.7. Luego tenemos el gráfico de control de Fase II correspondiente que se muestra en la Figura 3.6.

**Tabla 3.7:** Muestras adicionales (Taleb (2009)), valores representativos y estadísticos  $T^2$  correspondientes

muestra	$q_{11}$	$q_{12}$	$q_{13}$	$q_{21}$	$q_{22}$	$q_{23}$	$q_{31}$	$q_{32}$	$q_{33}$	$R_1$	$R_2$	$R_3$	$T^2$
24	62	76	62	101	91	8	151	37	12	0.462	0.375	0.245	112.24
25	100	71	29	63	97	39	165	24	11	0.334	0.502	0.217	87.17
26	88	77	35	111	86	3	110	40	50	0.364	0.345	0.390	89.33



**Figura 3.6:** Gráfico de Control  $T^2$ MMD Fase II: Proceso de producción de porcelana

### Interpretación de las señales de fuera-de-control

La figura 3.6 muestra que el proceso está fuera-de-control cuando se toman las muestras 24, 25 y 26. Para determinar qué variables son las responsables de la salida de control en cada uno de los tres casos, se calculan los correspondientes  $d_i$  y se muestran en la Tabla 3.8.

La tabla 3.8 muestra que la variable con mayor contribución a la señal de fuera-de-control detectada cuando se tomó la muestra 24, es la variable  $Q_1$  (apariencia), ya que es la variable con el valor más alto de  $d_i$ , mientras que en el caso de la muestra 25 la mayor contribuyente es la variable  $Q_2$

**Tabla 3.8:** Valores  $d_i$  para la interpretación de señales de fuera-de-control para el proceso de producción de porcelana

<i>muestras</i>	$T^2$	$T^2_{(1)}$	$T^2_{(2)}$	$T^2_{(3)}$	$d_1$	$d_2$	$d_3$
24	112.24	1.20	81.76	111.07	111.04	30.47	1.17
25	87.17	61.92	0.50	81.42	25.26	86.67	5.75
26	89.33	89.02	83.95	5.29	0.31	5.39	84.04

(translucidez) y para la muestra 26 es la variables  $Q_3$  (blancura).

### 3.5. Discusión

#### 3.5.1. Ejemplo 1

Como podemos ver en la figura 3.7 los resultados obtenidos por la metodología propuesta para el ejemplo 1 son bastante similares a los propuestos por Taleb et al. (2006).

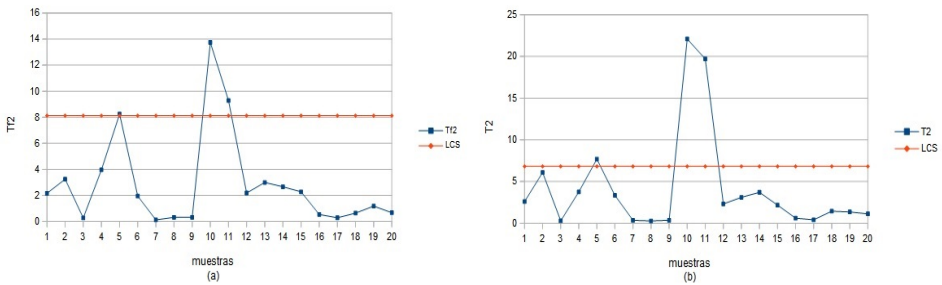


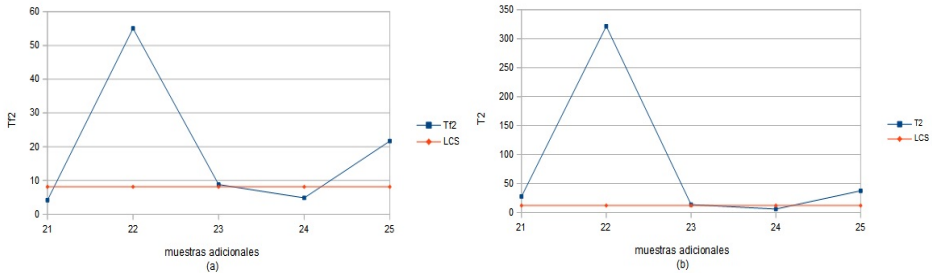
Figura 3.7: (a) Gráfico de Control estadístico  $T_f^2$  (Taleb et al. (2006)),  
(b) Gráfico de Control  $T^2$ MMD Fase I

En la Fase I podemos ver que hay una superposición significativa entre los dos gráficos en la forma general de las curvas y en la identificación de anomalías, con la ventaja de que se utiliza un gráfico conocido y se evita el uso de metodologías basadas únicamente en simulación.

Para los datos adicionales los resultados que se muestran en la tabla 3.9 y la figura 3.8, los gráficos de control de Fase II se ajustan a la identificación de los puntos clave que muestran anomalías.

**Tabla 3.9:** Estadísticos  $T_f^2$  (Taleb et al.(2006)) y  $T^2$  para muestras adicionales

<i>muestra</i>	$T_f^2$ (LCS=8.12)	$T^2$ (LCS=12.139)
21	4.18	<b>45.56</b>
22	<b>55.03</b>	<b>495.96</b>
23	8.78	37.76
24	4.84	11.21
25	<b>21.67</b>	<b>35.48</b>



**Figura 3.8:** (a) Gráfico de Control estadístico  $T_f^2$  para muestras adicionales (Taleb et al. (2006)), (b) Gráfico de Control  $T^2$ MMD Fase II

### 3.5.2. Ejemplo 2

En cuanto al segundo ejemplo, podemos decir que si bien ambas: metodología y muestreo estadístico, en Taleb (2009) son diferentes, los resultados finales al aplicar el gráfico propuesto en este capítulo son los mismos. Esto es, en términos de detección de la señal de salida de control y la identificación de las variables responsables de dicha señal. Se muestra así la generalidad del enfoque planteado.

### 3.6. Conclusiones

Después de la revisión de la literatura en relación con los gráficos de control de calidad difusos multivariados decidimos trabajar bajo el supuesto de que los productos son clasificados por cada característica de calidad por separado, ya que son los métodos de muestreo que tienen mayor grado de viabilidad de su implementación.

La teoría difusa se ha utilizado para transformar las variables multivariadas en sus valores representativos correspondientes. Estos valores siguen una distribución normal multivariante aproximada cuando el método de defusificación utilizado es el promedio difuso. Con este enfoque, se evita la complejidad inherente de la distribución multinomial multivariante. En su lugar un gráfico de control Hotelling  $T^2$  para una sola observación puede ser diseñado.



# Rendimiento del gráfico de control difuso multinomial multivariante

## 4.1. Introducción

Se necesitan medidas de desempeño para estudiar y comparar el rendimiento de los gráficos de control. A menudo se utilizan la distribución de la longitud de racha *run length*(RL) y sus momentos para determinar los límites de un gráfico de control y estudiar su rendimiento, tanto en situaciones bajo-control como fuera-de-control (Champ et al., 2005). La longitud de racha de un gráfico de control se define como el número de la etapa de muestreo en el que se observa la primera señal de alarma. La media de esta variable aleatoria, la longitud de racha promedio *Average Run Length* (ARL), es una medida común de rendimiento del gráfico de control.

Cuando los estadísticos del gráfico son independientes e idénticamente distribuidos (i.i.d.) y los límites de control son constantes, la longitud de

racha del gráfico es una variable aleatoria con distribución geométrica, luego el ARL bajo-control puede ser calculado fácilmente a partir de:

$$ARL_0 = \frac{1}{\alpha} \quad (4.1)$$

donde  $\alpha$  es la probabilidad de que un punto exceda los límites de control (Montgomery, 2009).

Un gráfico de  $\bar{x}$  típico con límites  $3 - \sigma$  basado en parámetros conocidos tiene un ARL bajo-control de 370 bajo el supuesto de que los datos sean normalmente distribuidos.

El ARL para un estado del proceso fuera-de-control es una medida de la rapidez en que será detectada una señal de alarma. Si la RL de un gráfico es una variable aleatoria geométrica, el ARL del gráfico de control para detectar un cambio de tamaño  $d$  se calcula como:

$$ARL_d = \frac{1}{1 - \beta} \quad (4.2)$$

donde  $\beta$  es la probabilidad de error Tipo II, es decir, es la probabilidad de no detectar que el proceso está fuera-de-control. (Montgomery, 2009)

El ARL es la medida más común de rendimiento del gráfico de control, y gran parte de su popularidad se debe a su interpretación intuitivamente atractiva. Por lo tanto, a menudo se utiliza para diseñar y comparar el rendimiento de los gráficos de control.

Generalmente se elige un valor grande de ARL bajo-control para

evitar falsas señales de alarma frecuentes. Mientras que es deseable que el ARL de fuera-de-control sea pequeño, con el fin de que el gráfico detecte rápidamente los cambios del proceso.

Para medir el rendimiento de un gráfico de control, es necesario tener en cuenta el efecto que pueda tener sobre éste la estimación de parámetros. [Jensen et al. \(2006\)](#) realizaron una revisión de literatura relevante sobre los efectos de la estimación de parámetros en las propiedades de los gráficos de control.

[Champ et al. \(2005\)](#) mostraron que los límites de control basados en la distribución-F no resultan en el rendimiento deseado, debido a la dependencia de los estadísticos del gráfico y proporcionan recomendaciones de tamaño de la muestra necesaria para alcanzar el rendimiento deseado bajo-control cuando se utilizan los límites basados en distribución-F. [Ryan \(2011\)](#) notó la dependencia de los estadísticos del gráfico  $T^2$  y recomendó el uso de simulación para determinar la ARL del gráfico  $T^2$  Fase II cuando los parámetros son estimados.

En este capítulo se pretende analizar el rendimiento del Gráfico de Control  $T^2$  Multinomial Multivariado con enfoque difuso ( $T^2$ MMD), utilizando métodos de simulación. Se discute la distribución de la RL del gráfico  $T^2$  de Fase II (como parte del gráfico  $T^2$ MMD) con vector de medias y matriz de varianzas estimados. Por otra parte, usando el trabajo de [Champ et al. \(2005\)](#) como referencia, se proporcionan límites de control corregidos basados en un proceso de simulación para su uso cuando el tamaño de la muestra requerida no está disponible. Se hacen comparaciones con el rendimiento del gráfico de control  $T^2$  con parámetros conocidos establecidos teóricamente.

## 4.2. Propuesta

Para el cálculo del ARL se han desarrollado dos algoritmos. Uno que permita, mediante simulación de *Monte Carlo*, generar  $m$  muestras de  $p$  variables aleatorias multinomiales correlacionadas de tamaño  $n$  cada una con  $s_j$  categorías  $j = 1, \dots, p$ ; y otro que usando el primero permita la determinación de la longitud media de racha. Los algoritmos de simulación se describen a continuación.

### Algoritmo 1

**Problema:** Generar  $m$  muestras de  $p$  variables multinomiales correlacionadas cada una de tamaño  $n$ , con un vector de proporciones preestablecidas y una matriz de correlación dadas.

Para resolver el problema se implementó un programa en **R** versión 2.15.1, con el siguiente algoritmo:

1. Establecer los valores de  $m$ ,  $n$ ,  $p$ , el vector de proporciones  $q$ , el vector de medias  $\mu$  y la matriz de correlaciones a usar  $\Sigma$ .
2. Generar una observación normal multivariante de  $p$  variables de tamaño  $n$ , usando la función *mvnorm* contenida en el paquete MASS. El número de variables queda determinado por las dimensiones del vector de medias y la matriz de correlación utilizadas.
3. Transformar las normales en multinomiales utilizando los vectores de proporciones dados, utilizamos estos valores como los correspondientes cuantiles y procedemos de esta manera a discretizar cada una de las normales, contando el número de observaciones en cada cuantil.

4. Repetir el proceso  $m$  veces, para obtener el número de muestras deseadas.

### Algoritmo 2

**Problema:** Determinar la longitud media de racha, **ARL**, para un error tipo I dado.

Al igual que con el problema anterior, se implementó un programa en **R** versión 2.15.1 con el siguiente algoritmo:

1. Determinar el límite de control superior (UCL) para la combinación deseada de  $p$  y  $m$  usando la ecuación (4).
2. Generar una muestra aleatoria usando el algoritmo anterior para representar la nueva información del proceso en el tiempo  $k$ .
3. Transformar la muestra generada en el vector de valores representativos correspondiente. Calcular  $T_k^2$ .
4. Comparar  $T_k^2$  con el límite de control. Si no se observa la señal de fuera-de-control ( $T_k^2 > UCL$ ), repita los pasos 2 y 3. Si se observa una señal, entonces vaya al paso 5.
5. Registrar la longitud de racha en un vector.
6. Repetir los pasos 2-5 hasta que se haya completado el número deseado de repeticiones.
7. Calcular el promedio de los valores registrado en el paso 5.

Como se realiza típicamente con los gráficos  $T^2$  Fase II, se trabaja bajo el supuesto de que el vector aleatorio de las mediciones de calidad  $R_i$  de

$p \times 1$  tiene una distribución normal  $p$ -variable con parámetros  $\mu$  y  $\Sigma$ . El proceso se supone que está bajo-control cuando  $\mu = \mu_R$  y  $\Sigma = \Sigma_R$ .

Aunque los cambios en la matriz de covarianzas afectarán la distribución de la longitud de racha del gráfico, nos centramos en los cambios en el vector de medias del estado bajo-control,  $\mu_R$ , a un estado fuera-de-control,  $\mu$ , supondremos que no hay cambios en la matriz de covarianzas.

Supondremos además, que tenemos disponibles  $m$  muestras aleatorias independientes  $R_1, R_2, \dots, R_m$  de un proceso normal  $p$ -variante bajo-control para calcular las estimaciones de los parámetros dados por las ecuaciones 3.16 3.17 y 3.18. Estas muestras se obtendrán una vez que las variables multinomiales multivariadas generadas por simulación (algoritmo 1) sean transformadas en normales multivariantes, mediante el procedimiento descrito en el capítulo anterior.

Para el análisis del rendimiento del gráfico T<sup>2</sup>MMD con parámetros estimados, como mencionamos anteriormente, limitamos nuestro estudio fuera-de-control a reflejar los cambios en el vector de medias.

Una medida del cambio de un vector (o distancia entre dos vectores) utilizada en análisis multivariante es la distancia de Mahalanobis (Aparisi et al., 2004). En nuestro caso la distancia entre el vector de medias original y el nuevo vector de medias está dado por

$$d = \sqrt{(R_i - \bar{R})' S^{-1} (R_i - \bar{R})}$$

Los cambios en el vector de medias pueden ser medidos utilizando la distancia  $d$ .

Los resultados que presentamos aquí se basan en una simulación de Monte Carlo, descrita en el algoritmo 2.

Si bien es cierto los cambios a tomar en cuenta en el gráfico  $T^2$ , están dados por  $d$ , éstos se obtienen a partir de los cambios planteados en los vectores de proporciones de las variables multinomiales, que son los que nos interesan.

Para analizar el desempeño del gráfico de control  $T^2$ MMD propuesto, se proponen dos escenarios: estado del proceso bajo-control y estado fuera-de-control, para diferentes valores de error de tipo I.

### 4.3. Análisis del rendimiento: estado bajo-control

Para el análisis del rendimiento del gráfico de control para el estado en-control se utiliza el ejemplo proporcionado en el artículo de [Taleb et al. \(2006\)](#) usado en el capítulo anterior como primer ejemplo, es decir, tendremos un escenario en el que se controlan tres características de calidad: apariencia ( $Q_1$ ), color ( $Q_2$ ), y el sabor ( $Q_3$ ), de un producto alimenticio, con sus respectivos conjuntos de términos y funciones de pertenencia.

Es necesario anotar que para determinar los parámetros del gráfico de control de Fase II se dispondrá de 50 muestras iniciales de tamaño  $n = 220$ , generadas mediante simulación, dada la necesidad de asegurar que las muestras iniciales hayan sido tomadas cuando el proceso se supone bajo-control. Para la simulación de las muestras se utiliza el *algoritmo 1* descrito anteriormente, tomando como vector de medias el vector nulo, la matriz de correlación una matriz inicial dada, y como vector de proporciones el dado por [Taleb et al. \(2006\)](#) para cuando el

proceso se considera bajo-control.

Los datos generados por simulación se muestran en la tabla 4.1.

**Tabla 4.1:** Datos simulados para la Fase I

<b>muestras</b>	$n_{11}$	$n_{12}$	$n_{13}$	$n_{21}$	$n_{22}$	$n_{23}$	$n_{31}$	$n_{32}$	$n_{33}$	$n_{34}$
1	205	9	6	203	12	5	171	40	9	0
2	206	9	5	200	13	7	175	43	2	0
3	205	8	7	208	9	3	167	47	5	1
4	204	8	8	210	5	5	168	49	3	0
5	213	2	5	209	7	4	176	42	1	1
6	206	9	5	200	10	10	165	47	8	0
7	209	7	4	205	9	6	167	46	4	3
8	209	9	2	206	6	8	169	45	4	2
9	205	8	7	201	10	9	174	36	6	4
10	204	13	3	198	14	8	178	34	8	0
11	210	6	4	202	11	7	162	52	5	1
12	212	5	3	213	6	1	162	53	5	0
13	208	5	7	203	8	9	170	46	3	1
14	207	6	7	208	7	5	169	41	6	4
15	206	6	8	199	14	7	172	42	4	2
16	211	6	3	202	11	7	164	52	4	0
17	205	6	9	207	6	7	172	42	5	1
18	203	10	7	203	10	7	170	46	4	0
19	210	7	3	211	4	5	187	27	6	0
20	205	9	6	208	6	6	165	50	3	2
21	209	7	4	197	13	10	180	33	7	0
22	210	8	2	201	11	8	158	57	4	1
23	203	10	7	192	16	12	174	43	3	0
24	212	5	3	203	9	8	155	61	2	2
25	211	6	3	210	7	3	169	43	6	2
26	203	10	7	198	12	10	170	43	4	3
27	211	6	3	207	8	5	161	55	3	1



<b>muestras</b>	$n_{11}$	$n_{12}$	$n_{13}$	$n_{21}$	$n_{22}$	$n_{23}$	$n_{31}$	$n_{32}$	$n_{33}$	$n_{34}$
28	202	13	5	206	6	8	176	41	2	1
29	206	5	9	195	17	8	176	40	3	1
30	201	10	9	205	12	3	171	42	6	1
31	208	7	5	202	14	4	169	46	5	0
32	204	12	4	207	6	7	162	52	6	0
33	207	8	5	207	6	7	179	38	0	3
34	205	9	6	199	13	8	181	36	3	0
35	207	5	8	205	8	7	165	49	6	0
36	206	9	5	203	10	7	166	48	4	2
37	211	5	4	195	15	10	174	44	2	0
38	210	5	5	212	4	4	167	47	6	0
39	207	8	5	204	13	3	166	49	4	1
40	203	13	4	205	7	8	177	41	2	0
41	206	8	6	207	10	3	154	59	7	0
42	205	9	6	206	9	5	162	49	8	1
43	208	9	3	209	3	8	170	49	0	1
44	206	10	4	208	7	5	174	44	0	2
45	198	13	9	199	10	11	165	51	4	0
46	211	7	2	206	8	6	176	41	3	0
47	207	12	1	206	9	5	178	39	1	2
48	213	3	4	210	6	4	168	48	4	0
49	204	7	9	205	8	7	175	43	2	0
50	204	12	4	203	9	8	176	37	6	1

A partir de los datos generados y usando las funciones de pertenencia del ejemplo de [Taleb et al. \(2006\)](#) se calculan los vectores de valores representativos  $R_i$  de acuerdo al procedimiento planteado en el capítulo anterior y a partir de ellos se obtienen los parámetros del gráfico de

control para la fase II usando las ecuaciones 3.16, 3.17 y 3.18, verificando que los datos generados cumplieran la condición de que provienen de un estado bajo-control.

Los parámetros estimados para el gráfico de control de fase II, que se utilizarán para calcular el ARL son:

$$S^{-1} = \begin{pmatrix} 24032.7 & -9722.1 & -4118.5 \\ -9722.1 & 16432.8 & 5820.3 \\ -4118.5 & 5820.3 & 12397.4 \end{pmatrix}.$$

$$\bar{R} = (0.108, 0.197, 0.149)$$

La tabla 4.2 muestra el rendimiento bajo-control ( $ARL_0$ ), la desviación estándar de la longitud de racha (SDRL), y los percentiles 10, 50 y 90 de la distribución de la longitud de racha del gráfico para diferentes valores de error tipo I ( $\alpha$ ), utilizando uno de los tres límites de control: Límites para parámetros conocidos, límites tradicionales o límites corregidos.

Los casos designados como “parámetros-conocidos” dan la  $ARL_0$ , la SDRL, y los percentiles 10, 50 y 90 de la distribución de la longitud de racha del gráfico  $T^2$  teóricos para parámetros conocidos, es decir, éstos son simplemente los momentos y percentiles de la distribución de una variable aleatoria geométrica con parámetro  $\alpha$  y se dan como referencia, los límites de control en estos casos están dados por el cuantil de una  $\chi^2_{p,\alpha}$ . Los casos designados como “límites-tradicionales” a los que también podemos llamar límites-F, se calcularon mediante simulación cuando los límites de control se calculan a partir de la ecuación 3.14. Los casos designados como “límites- corregidos” también se calcularon mediante simulación, pero los límites de control se determinan a partir

de una búsqueda binaria para lograr un ARL bajo-control acorde con el valor  $\alpha$  correspondiente. Estos resultados fueron verificados con extensas simulaciones.

**Tabla 4.2:** Rendimiento estado bajo-control del Gráfico de Control T<sup>2</sup>MMD con parámetros estimados a partir de 50 muestras

alfa		UCL	ARL	SDRL	Q10	Q50	Q90
0,0027	Parámetros-conocidos	14.16	370.4	369.9	38	256	851
	Límite-Tradicionales	17.36	394.3	391.75	42	276	902
	Límite-Corregido	17.18	370.3	370.8	37	258	853
0,005	Parámetros-conocidos	12.84	200	199.5	21	138	459
	Límite-Tradicionales	15.52	221.6	218.8	23	151	510
	Límite-Corregido	15.21	200.9	200.8	22	141	459
0,01	Parámetros-conocidos	11.34	100	99.5	10	68	229
	Límite-Tradicionales	13.49	114.2	114.6	13	78	262
	Límite-Corregido	13.08	100.1	99.3	11	70	230
0,05	Parámetros-conocidos	7.84	20	19.5	2	13	44
	Límite-Tradicionales	8.94	24.3	23.7	3	17	55
	Límite-Corregido	8.42	20.3	19.7	3	14	46

A partir de los resultados de la tabla 4.2, lo primero que notamos es el efecto de la estimación de parámetros sobre el rendimiento bajo-control del gráfico. Vemos que el uso de estimaciones en lugar de los parámetros conocidos aumenta el ARL bajo-control del gráfico, es decir, la estimación conduce a un ARL bajo-control que es mayor de lo esperado según el diseño, acorde con lo que comenta (Champ et al., 2005) Del mismo modo, el percentil 10 de la distribución del RL con parámetros-estimados es virtualmente idéntica al caso parámetros-conocidos, mientras el percentil 50 es ligeramente mayor en estos casos. Esto implica que, el gráfico T<sup>2</sup>MMD con parámetros estimados no sufre del mismo problema de falsas alarmas tempranas como otros gráficos. Esto es debido en parte al hecho de que los límites tradicionales del gráfico T<sup>2</sup>MMD ya se han ajustado en cierta medida por el uso de parámetros estimados. Los límites

se construyen a partir de la distribución-F en lugar de la distribución chi-cuadrado, y por lo tanto se ensanchan para tener en cuenta la variabilidad adicional introducida en el estadístico del gráfico a través de la estimación.

Sin embargo, es claro que estos límites tradicionales son demasiado anchos para lograr la ARL bajo-control deseada. Se desea un gráfico con una gran ARL bajo-control, por lo que los límites basados en distribución-F pueden parecer inicialmente deseables. Sin embargo, debemos ser cuidadosos ya que el uso de límites que son demasiado anchos también aumentará la ARL fuera-de-control, lo que resulta en un gráfico que es lento para indicar un cambio en el proceso.

#### 4.4. Análisis del rendimiento: estado fuera-de-control

Para el cálculo de los valores de longitud media de racha para el estado fuera-de-control se emplean los cambios en los vectores de proporciones de las variables multinomiales propuestos por Taleb et al. (2006) éstos cambios junto con la distancia de Mahalanobis (desplazamiento) correspondiente para el estadístico  $T^2$  se muestran en la tabla 4.3.

**Tabla 4.3:** Cambios en el Vector de Proporciones propuestos por Taleb et al.(2006) y distancia de Mahalanobis

	$C_{11}$	$C_{12}$	$C_{13}$	$C_{21}$	$C_{22}$	$C_{23}$	$C_{31}$	$C_{32}$	$C_{33}$	$C_{34}$	$d$
V.P.E-C*	0,942	0,035	0,023	0,925	0,045	0,03	0,774	0,206	0,016	0,004	<b>0</b>
Cambio 1	0,942	0,035	0,023	0,905	0,065	0,03	0,754	0,226	0,016	0,004	<b>1,22</b>
Cambio 2	0,922	0,035	0,043	0,925	0,045	0,03	0,774	0,206	0,016	0,004	<b>1,76</b>
Cambio 3	0,942	0,035	0,023	0,925	0,045	0,03	0,674	0,306	0,016	0,004	<b>2,62</b>
Cambio 4	0,892	0,035	0,073	0,925	0,045	0,03	0,774	0,206	0,016	0,004	<b>5,03</b>

V.P.E-C\*: vector de proporciones bajo-control

*Observaciones respecto de los cambios planteados, realizadas por Taleb et al. (2006):* El cambio 1 es considerado como un cambio pequeño, en efecto, las proporciones de baja calidad no son afectadas. En el cambio 2, la proporción de “pobre”, ( $c_{13}$ ) se incrementa en 0.02, y la proporción de ( $c_{11}$ ) disminuye en la misma cantidad, aunque se considera también como un cambio pequeño es más importante que el cambio 1 porque afecta la calidad “pobre”. Los cambios 3 y 4 pueden ser considerados como mediano y alto respectivamente.

La figura 4.1 muestra el rendimiento del gráfico de control (ARL fuera-de-control) versus el desplazamiento,  $d$ , para distintos valores de error tipo I ( $\alpha$ ) para los tres casos planteados:

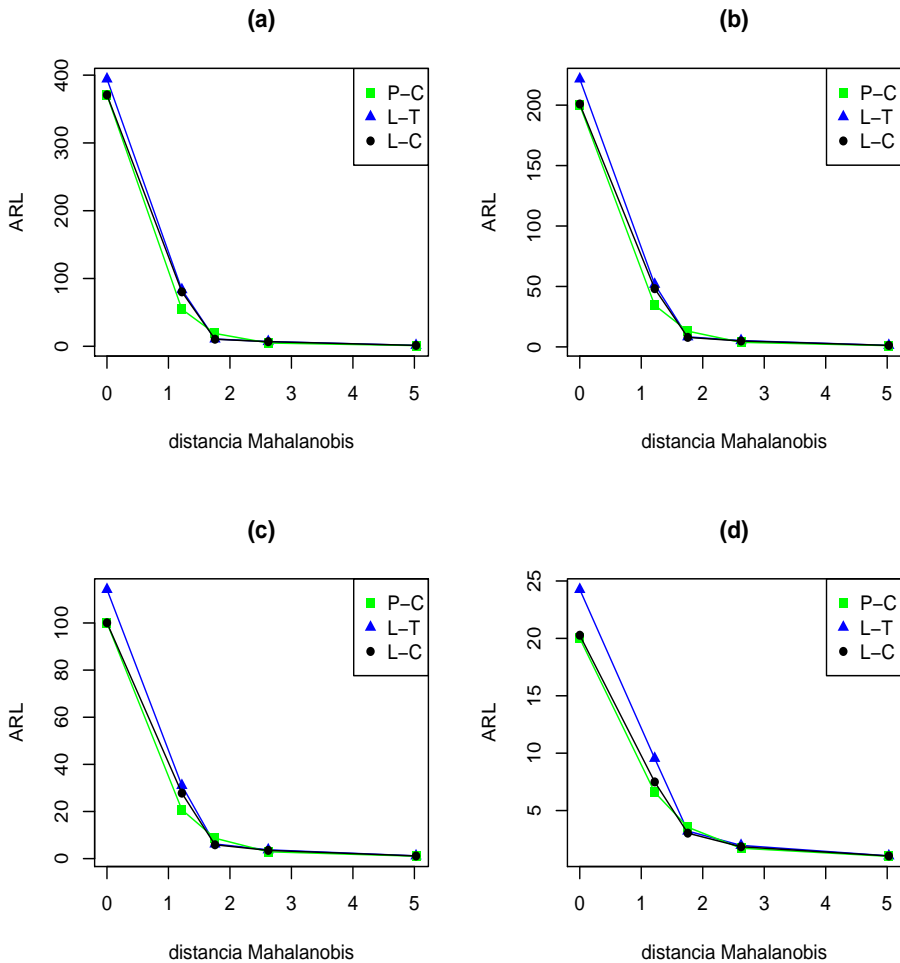
“Parámetros-conocidos” (P-C): el límite de control se determina a partir de una  $\chi^2$  con  $p$  grados de libertad y el valor de  $\alpha$  correspondiente.

“Límite-tradicional” (L-T): los límites de control se calculan de acuerdo con la ecuación 3.14.

“Límite-correctado” (L-C): los límites de control se calculan a través de simulación con una búsqueda binaria para obtener el ARL bajo-control correspondiente al valor de  $\alpha$ .

El caso P-C se usa como referencia y su ARL está dado por la media de una variable aleatoria geométrica,  $1/(1 - \beta)$ , donde  $\beta = P(\chi_{p,d^2}^2 < \chi_{p,0,\alpha}^2)$ . El ARL para los casos L-T y L-C se determina por simulación basada en estimaciones a partir de las muestras generadas.

Como era de esperar, el gráfico con parámetros estimados es más lento para detectar un cambio de proceso cuando se utilizan límites tradicionales especialmente para pequeños desplazamientos, independientemente

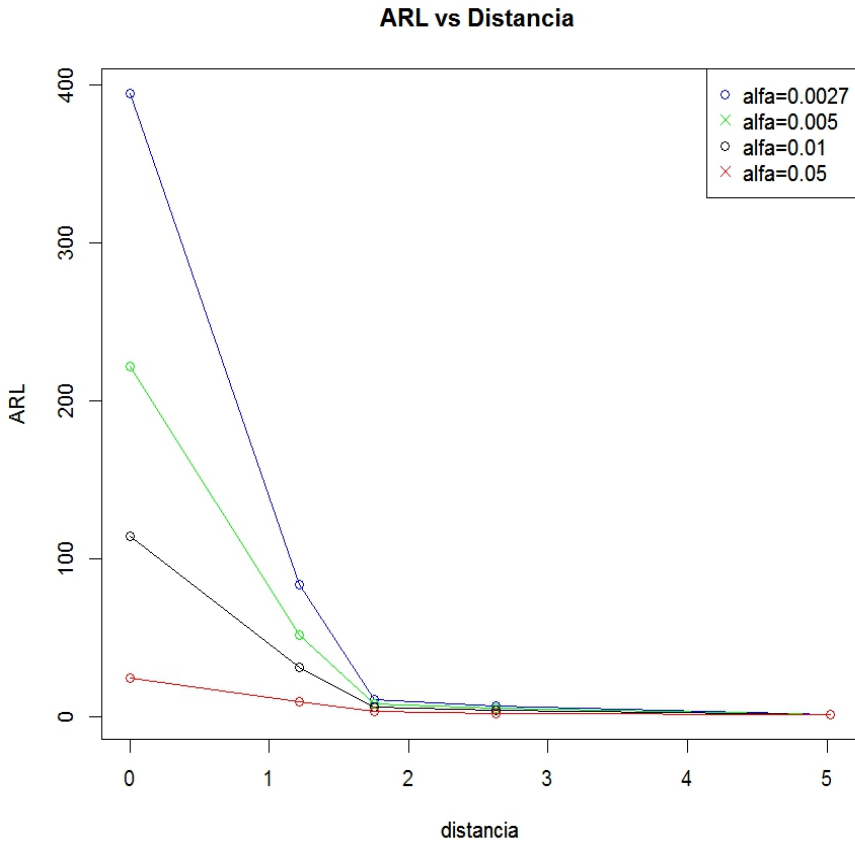


**Figura 4.1:** Rendimiento estado fuera-de-control: (a)  $\alpha = 0.0027$ ; (b)  $\alpha = 0.005$ ; (c)  $\alpha = 0.005$ ; (d)  $\alpha = 0.05$ . P-C: caso parámetros-conocidos; L-T: caso límite teórico; L-C: caso límite corregido.

del valor de  $\alpha$  escogido. Si se ajustan los límites del gráfico para lograr el ARL bajo-control deseado, el gráfico  $T^2$  con parámetros estimados tiene

un rendimiento, en términos de ARL similar al de un gráfico  $T^2$  con parámetros conocidos.

La figura 4.2 muestra una comparación del rendimiento del gráfico contra el tamaño del desplazamiento para los tres casos mencionados y distintos valores de  $\alpha$



**Figura 4.2:** Comparativo ARL vs Distancia para diferentes valores de  $\alpha$

Se puede apreciar que el rendimiento del gráfico con parámetros estimados es independiente del valor  $\alpha$  escogido para desplazamientos grandes.

## 4.5. Análisis comparativo del rendimiento

En este apartado se realiza un análisis del rendimiento del gráfico con enfoque difuso presentado por Taleb et al. (2006) (gráfico MFQCC) a fin de compararlo con el rendimiento del gráfico propuesto (gráfico T<sup>2</sup>MMD). Para este análisis se utilizan, para el cálculo de los parámetros del gráfico de control, los datos generados por simulación mostrados en la tabla 4.1. Al igual que en el apartado anterior se emplean los cambios en los vectores de proporciones de las variables multinomiales propuestos por Taleb et al. (2006) para el cálculo del ARL para el estado *fuera-de-control*.

Taleb et al. (2006) determinan el límite de control para el gráfico MFQCC utilizando el método de remuestreo Bootstrap debido a que la distribución del estadístico del gráfico que proponen ( $T_f^2$ ) se desconoce. Sin embargo, con el fin de poder realizar la comparación, se ha determinado el límite de control a usar mediante simulación para un valor de ARL *bajo-control* dado.

El análisis se realiza para dos valores diferentes de error tipo I:  $\alpha = 0.05$  y  $\alpha = 0.0027$ . A fin de realizar una verificación de resultados, el cálculo del ARL se realiza mediante dos procedimientos: como el promedio de 10000 distintos LRs y como el inverso de la proporción de valores del estadístico trazados fuera del límite de control superior de un total de 30.000. Los valores obtenidos se comparan con los valores del ARL obtenidos para el gráfico de control T<sup>2</sup>MMD propuesto en esta tesis para cuando se usan límites de control tradicionales y para cuando



se usan límites de control corregidos. Los resultados se muestran en las tablas 4.4 y 4.5

**Tabla 4.4:** ARL vs cambios propuestos, para los Gráficos de Control  $T^2$ MMD y MFQCC  $\alpha = 0.05$

	Límite-Tradicional	Límite-Corregido	Taleb (LRs)	Taleb (1/p)
V.PE-C	24.26	20.27	20.36	20.27
Cambio 1	9.55	7.49	9.98	10.19
Cambio 2	3.21	3.03	2.91	2.88
Cambio 3	1.97	1.85	2.18	2.19
Cambio 4	1.05	1.05	1.03	1.03

**Tabla 4.5:** ARL vs cambios propuestos, para los Gráfico de Control  $T^2$ MMD y MFQCC  $\alpha = 0.0027$

	Límite-Tradicional	Límite-Corregido	Taleb (LRs)	Taleb (1/p)
V.PE-C	396.34	370.85	381.29	386.69
Cambio 1	83.47	80.31	125.87	120.82
Cambio 2	10.77	10.46	9.02	8.95
Cambio 3	6.96	6.80	11.54	11.46
Cambio 4	1.30	1.029	1.17	1.16

Los gráficos comparativos para ambos casos se pueden ver en la figura 4.3. En estos gráficos se puede apreciar que, para los dos casos planteados, los mejores resultados de rendimiento se obtienen para el gráfico  $T^2$ MMD con límites corregidos.

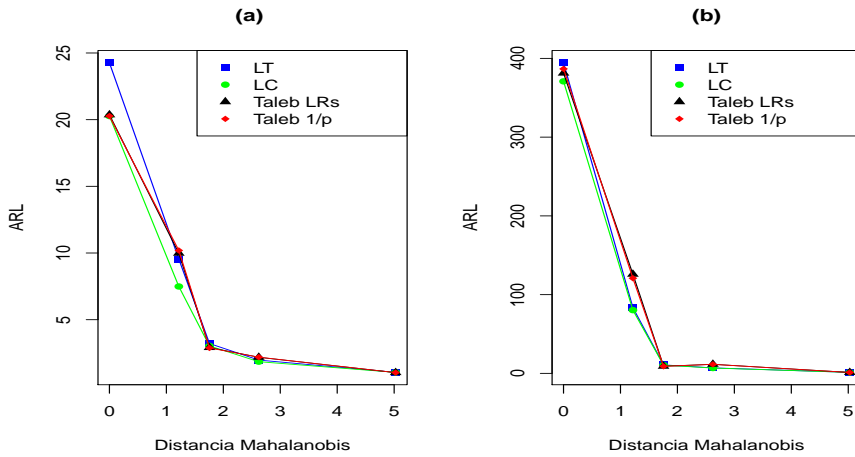


Figura 4.3: Comparativo ARL vs Distancia para Gráfico de Control  $T^2$ MMD y gráfico MFQCC: (a)  $\alpha = 0.05$ , (b)  $\alpha = 0.0027$

## 4.6. Conclusiones

El método de simulación propuesto nos permite verificar que el gráfico de control  $T^2$  multinomial multivariado con enfoque difuso propuesto satisface varias propiedades del gráfico de control  $T^2$  observado en el análisis realizado por [Champ et al. \(2005\)](#). Esto, por un lado confirma el comportamiento alegado por los autores de gráfico de control y, por otro muestra la eficacia del método de simulación propuesto en este capítulo.

Una observación inicial interesante es que los cambios en el vector de medias de los valores representativos, correspondientes a los cambios en las proporciones vectores de las variables multinomiales, son proporcionales a la gravedad del cambio. Es decir, si el cambio en las proporciones afectados es una amenaza para la calidad del producto, esto se reflejará en un mayor cambio en el vector de medias del gráfico  $T^2$  y por ende será

detectado más rápidamente.

Adicionalmente, se ha estudiado la distribución de la RL del gráfico  $T^2$  de Fase II (como parte del gráfico  $T^2$ MMD). Hemos detectado una diferencia en el rendimiento del gráfico en función de la utilización de parámetros conocidos o estimados. Lógicamente, el uso de parámetros estimados aumenta el ARL del gráfico. En concordancia con los resultados presentados por [Champ et al. \(2005\)](#), nos encontramos con que los gráficos de control  $T^2$  con parámetros estimados tienen tasas de falsas alarmas menores que otros gráficos.

Además de los límites de control para parámetros conocidos y los límites de control con parámetros estimados, se ha obtenido mediante simulación un conjunto de límites de control corregidos. Con estos límites corregidos se consigue un mayor grado de ajuste de los resultados de rendimiento correspondientes a la mejor situación teórica, que corresponde a la utilización de parámetros conocidos.

Hemos analizado el comportamiento gráfico para diferentes valores de error tipo I,  $\alpha$ , y para diferentes niveles de desplazamiento del vector de medias del estado bajo-control, medido mediante la distancia de Mahalanobis. Hemos visto que el gráfico con parámetros estimados es más lento para detectar cambios en el proceso cuando se utilizan límites tradicionales, especialmente para pequeños desplazamientos, independientemente del valor de error de tipo I elegido. Utilizando límites ajustados, el rendimiento es similar, en términos de ARL, a un gráfico  $T^2$  con parámetros conocidos.

El método de simulación propuesto nos permitió verificar que el gráfico de control  $T^2$  multinomial multivariado con enfoque difuso satisface las propiedades del gráfico de control  $T^2$  observadas en el análisis realizado por [Champ et al. \(2005\)](#). Por un lado, esto confirma

el comportamiento de la gráfica  $T^2$ MMD, alegada por [Pastuizaca et al. \(2015\)](#), y por otra parte muestra la eficacia del método de simulación presentada en este documento.

## Análisis de sensibilidad

### 5.1. Introducción

En el capítulo anterior se ha propuesto un método para medir el rendimiento del gráfico  $T^2$ MMD propuesto y se han comparado sus propiedades con las del gráfico  $T^2$  propuesto por Hotelling, usando como referencia el análisis realizado por [Champ et al. \(2005\)](#).

En este capítulo se analiza la sensibilidad del rendimiento del gráfico  $T^2$ MMD frente al grado de borrosidad de las funciones de pertenencia elegidas para la descripción de los subconjuntos difusos de las variables lingüísticas, es decir, el efecto que tiene ese grado de borrosidad sobre el rendimiento. En este caso se usará el error tipo II para medir el rendimiento y las correspondientes curvas características de operación (*curvas OC*) para efectos de comparación.

Adicionalmente y tomando en cuenta que una de las principales ob-

servaciones realizadas por Taleb y Limam (2002) a los gráficos de control difuso para variables multinomiales univariadas propuestos por Raz y Wang (1990) fue la sensibilidad de los resultados de los mismos frente al grado de borrosidad de las funciones de pertenencia, analizaremos también aquí ese tema. Al hablar de resultados nos referimos al efecto en los gráficos de control tanto de fase I como de fase II así como la detección de posibles anomalías en el proceso.

### 5.1.1. Grado de borrosidad de un subconjunto difuso

El grado de borrosidad se mide por el índice lineal de borrosidad propuesto por Kaufmann (1975) como una extensión de la distancia relativa generalizada de Hamming (Raz y Wang, 1990). El índice lineal de borrosidad  $\nu(A)$  de un subconjunto borroso  $A$  se define como la distancia de Hamming generalizada entre  $A$  y el subconjunto común más cercano  $\underline{A}$ , denotado por  $\underline{A}$ . La función de pertenencia de  $\underline{A}$  es:

$$\mu_{\underline{A}}(x) = \begin{cases} 0, & \text{si } \mu_A(x) \leq 0.5 \\ 1, & \text{si } \mu_A(x) > 0.5 \end{cases} \quad (5.1)$$

Cuando el soporte de  $A$  contiene  $n$  puntos discretos  $x_i, i = 1, \dots, n$ ,  $\nu(A)$  es igual a:

$$\nu(A) = \frac{2}{n} \sum_{i=1}^n |\mu_A(x_i) - \mu_{\underline{A}}(x_i)| \quad (5.2)$$

Cuando el soporte de  $A$  es el intervalo continuo  $[0, 1]$   $\nu(A)$  es igual a:

$$v(A) = \int_{x=0}^1 |\mu_{\underline{A}}(x) - \mu_A(x)| dx \quad (5.3)$$

**Ejemplo:** A continuación se muestra un ejemplo del cálculo de borrosidad de un subconjunto difuso dado  $\mu_A(x)$ :

$$\mu_A(x) = \begin{cases} 1 - 2x, & 0 \leq x \leq 0.5 \\ 0, & 0.5 < x \leq 1 \end{cases}$$

$$\mu(x) = 0.5 \Rightarrow x = 0.25$$

Entonces La función de pertenencia de  $\underline{A}$  es:

$$\mu_{\underline{A}}(x) = \begin{cases} 0, & 0.25 \leq x \leq 1 \\ 1, & 0 \leq x < 0.25 \end{cases}$$

Luego

$$v(A) = \int_0^1 |\mu_{\underline{A}}(x) - \mu_A(x)| dx = \int_0^{0.25} (2x) dx + \int_{0.25}^{0.5} (1 - 2x) dx = 0.125$$

**Proposición:** El índice de borrosidad  $v(A)$  de un subconjunto borroso  $A$  representado por un número triangular difuso dado por la terna ordenada  $(a, b, c)$  con función de pertenencia  $\mu_A(x)$  esta dado por  $v(A) = \frac{c-a}{4}$

**Demostración 1:** método analítico

La función de pertenencia  $\mu_A(x)$  correspondiente al número triangular difuso  $(a, b, c)$  está dado por:

$$\mu_A(x) = \begin{cases} 0, & 0 \leq x < a; c < x \leq 1 \\ \frac{x-a}{b-a}, & a \leq x < b \\ \frac{c-x}{c-b}, & b \leq x \leq c \end{cases}$$

luego la función de pertenencia de  $\underline{A}$  es:

$$\mu_{\underline{A}}(x) = \begin{cases} 0, & 0 \leq x < \frac{a+b}{2}; \frac{b+c}{2} < x \leq 1 \\ 1, & \frac{a+b}{2} \leq x \leq \frac{b+c}{2} \end{cases}$$

El gráfico de las funciones de pertenencia de  $A$  y su subconjunto común más cercano  $\underline{A}$  puede verse en la figura 5.1

luego el índice de borrosidad está dado por:

$$\begin{aligned} v(A) &= \int_0^1 |\mu_{\underline{A}}(x) - \mu_A(x)| dx = \int_0^{\frac{a+b}{2}} \frac{a+b}{2} \frac{x-a}{b-a} dx \\ &+ \int_{\frac{a+b}{2}}^b \left[ 1 - \frac{x-a}{b-a} \right] dx + \int_b^{\frac{a+b}{2}} \left[ 1 - \frac{c-x}{c-b} \right] dx + \int_{\frac{b+c}{2}}^c \frac{c-x}{c-b} dx \\ v(A) &= \frac{b-a}{8} + \frac{b-a}{2} - 3\frac{b-a}{8} + \frac{c-b}{2} + \frac{c-b}{8} = \frac{c-a}{4} \end{aligned}$$

Q.E.D.



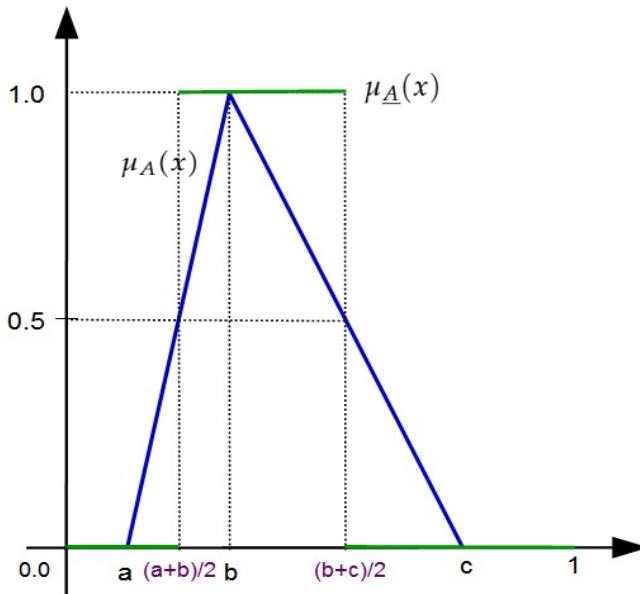


Figura 5.1: Funciones de pertenencia para A y su conjunto más cercano

*Demostración 2: método geométrico*

El índice de borrosidad del número triangular difuso  $(a, b, c)$  está dado por el área sombreada mostrada en la figura 5.2.

$$v(A) = \int_0^1 |\mu_{\underline{A}}(x) - \mu_A(x)| dx = A_{\Delta A} + A_{\Delta B} + A_{\Delta C} + A_{\Delta D}$$

pero  $A_{\Delta A} = A_{\Delta B}$  y  $A_{\Delta C} = A_{\Delta D}$

luego

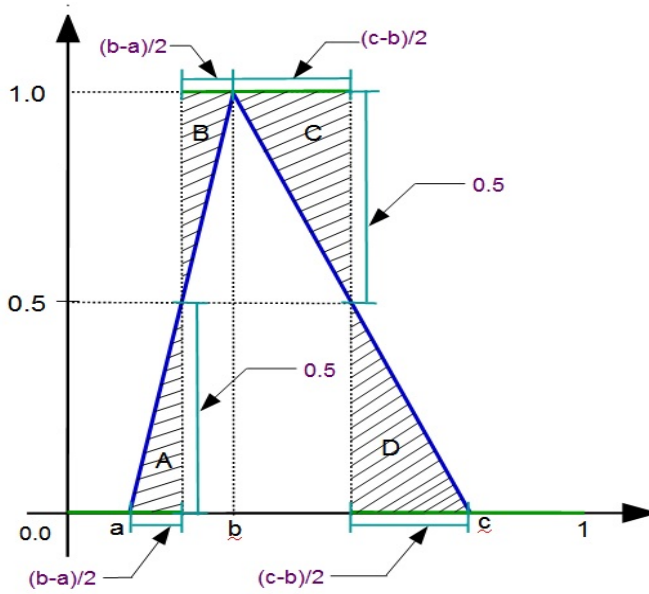


Figura 5.2: Índice de borrosidad

$$v(A) = 2A_{\Delta A} + 2A_{\Delta B} = 2 \left( \frac{1}{2} \right) \left( \frac{b-a}{2} \right) (0.5) + 2 \left( \frac{1}{2} \right) \left( \frac{c-b}{2} \right) (0.5)$$

$$v(A) = \frac{b-a}{4} + \frac{c-b}{4} = \frac{c-a}{4}$$

Q.E.D.

### 5.1.2. Curvas Características de Operación.

La capacidad de los gráficos de control para detectar cambios en la calidad del proceso se describe por su Curva Característica de Operación *Operating-Characteristic (OC) curve* (Montgomery, 2009).

Considere la curva característica de funcionamiento para un gráfico de control dado. La desviación estándar  $\sigma$  se supone conocida y constante. Si se producen cambios en el estadístico de control de  $\mu_0$  a un valor  $\mu_1$ , la probabilidad de no detectar este cambio en la primera muestra subsiguiente en este caso estaría dada por el error de tipo II ( $\beta$ ).

La curva característica de operación (OC) de un gráfico de control es una representación gráfica de la probabilidad de aceptar la hipótesis incorrecta de control estadístico (es decir, un error de tipo II o error- $\beta$ ) contra un desplazamiento del valor del estadístico controlado a partir de su valor nominal (su valor para un estado bajo-control).

La curva OC proporciona una medida de la sensibilidad del gráfico de control, es decir, su capacidad para detectar un cambio en el proceso del valor nominal a algún otro valor.

La probabilidad de error de tipo II ( $\beta$ ) para el gráfico de control propuesto se calcula mediante simulación de Monte Carlo como se explicará más adelante.

## 5.2. Sensibilidad de los resultados del gráfico de control $T^2$ MMD frente al grado de borrosidad de las funciones de pertenencia.

Para este análisis usaremos los ejemplos proporcionados por Taleb et al. (2006) y Taleb (2009). Para efectos de comparación a todos los términos lingüísticos en cada conjunto de términos se les asignó el mismo grado de borrosidad.

### 5.2.1. Ejemplo 1: Proceso de producción de comida congelada

Se tomará como primer ejemplo, el propuesto por Taleb et al. (2006), en el que se controlan tres características de calidad cualitativas con 3, 3 y 4 categorías respectivamente. Las funciones de pertenencia para los distintos conjuntos de términos se han diseñado para proporcionar cuatro diferentes grados de borrosidad: 0.1, 0.125, 0.15 y 0.2. A modo de ejemplo los gráficos de las funciones de pertenencia con grado de borrosidad 0.125 se muestran en la figura 5.3. Para efectos de comparación se usarán además las funciones de pertenencia proporcionadas en el ejemplo de Taleb presentadas en el capítulo 3. Se realizará el análisis para dos distintos valores de error tipo I:  $\alpha = 0.05$  y  $\alpha = 0.0027$ .

Los números triangulares difusos correspondientes a los nuevos conjuntos de funciones de pertenencia son:

*Grado de borrosidad 0.1*

$$\begin{aligned}
 F_{11} &= (0, 0, 0.4) & F_{12} &= (0.3, 0.5, 0.7) & F_{13} &= (0.6, 1, 1) \\
 F_{21} &= (0, 0, 0.4) & F_{22} &= (0.2, 0.4, 0.6) & F_{23} &= (0.6, 1, 1) \\
 F_{31} &= (0, 0, 0.4) & F_{32} &= (0.2, 0.4, 0.6) & F_{33} &= (0.4, 0.6, 0.8) & F_{34} &= (0.6, 1, 1)
 \end{aligned}$$

5.2. SENSIBILIDAD DE LOS RESULTADOS DEL GRÁFICO DE CONTROL T<sup>2</sup>MMD FRENTE AL GRADO DE BORROSIDAD DE LAS FUNCIONES DE PERTENENCIA.

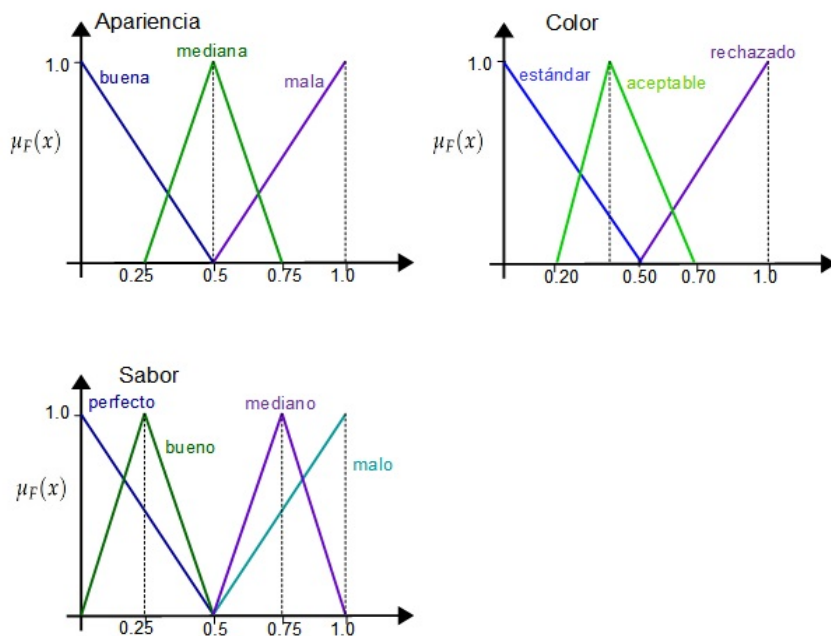


Figura 5.3: Funciones de pertenencia con grado de borrosidad 0.125

*Grado de borrosidad 0.125*

$$\begin{aligned}
 F_{11} &= (0, 0, 0.5) & F_{12} &= (0.25, 0.5, 0.75) & F_{13} &= (0.5, 1, 1) \\
 F_{21} &= (0, 0, 0.5) & F_{22} &= (0.2, 0.4, 0.7) & F_{23} &= (0.5, 1, 1) \\
 F_{31} &= (0, 0, 0.5) & F_{32} &= (0, 0.25, 0.5) & F_{33} &= (0.5, 0.75, 1) & F_{34} &= (0.5, 1, 1)
 \end{aligned}$$

*Grado de borrosidad 0.15*

$$\begin{aligned}
 F_{11} &= (0, 0, 0.6) & F_{12} &= (0.2, 0.5, 0.8) & F_{13} &= (0.4, 1, 1) \\
 F_{21} &= (0, 0, 0.6) & F_{22} &= (0.2, 0.4, 0.8) & F_{23} &= (0.4, 1, 1) \\
 F_{31} &= (0, 0, 0.6) & F_{32} &= (0, 0.3, 0.6) & F_{33} &= (0.4, 0.7, 1) & F_{34} &= (0.4, 1, 1)
 \end{aligned}$$

*Grado de borrosidad 0.2*

$$\begin{aligned}
 F_{11} &= (0, 0, 0.8) & F_{12} &= (0.1, 0.5, 0.9) & F_{13} &= (0.2, 1, 1) \\
 F_{21} &= (0, 0, 0.8) & F_{22} &= (0.05, 0.5, 0.85) & F_{23} &= (0.2, 1, 1) \\
 F_{31} &= (0, 0, 0.8) & F_{32} &= (0, 0.4, 0.8) & F_{33} &= (0.2, 0.6, 1) & F_{34} &= (0.2, 1, 1)
 \end{aligned}$$

### 5.2.1.1 Error tipo I( $\alpha$ )=0.05

**Gráfico de Control de Fase I** A partir de los datos proporcionados por Taleb et al. (2006), presentados en el capítulo 3, se calculan los valores representativos y los correspondientes valores del estadístico  $T^2$  para cada muestra para los distintos conjuntos de funciones de pertenencia.

El límite de control superior para el gráfico de control de fase I, para  $p = 3$ ,  $m = 20$  y  $\alpha = 0.05$  es:

$$LCS_1 = \frac{(m-1)^2}{m} B_{\left(\frac{p}{2}, \frac{m-p-1}{2}, \alpha\right)} = \frac{19^2}{20} 0.3778 = 6.82$$

Los valores del estadístico  $T^2$  para los diferentes conjuntos de funciones de pertenencia se muestran en la tabla 5.1.

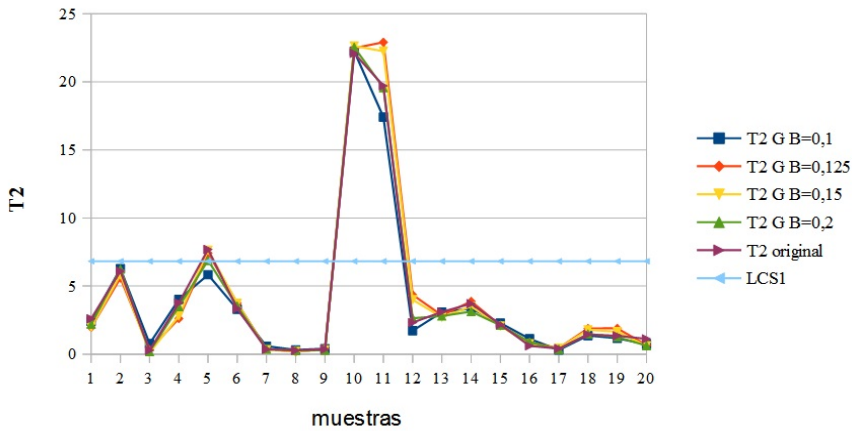
La figura 5.4 muestra un comparativo de los valores de los estadísticos  $T^2$  correspondientes a las muestras empleadas para el gráfico de control de Fase I. Esto es, para los conjuntos de funciones de pertenencia planteados con un grado de borrosidad específico, y el correspondiente al ejemplo de Taleb et al. (2006). Como se puede observar, para todos los conjuntos de funciones de pertenencia se obtienen similares resultados. Se identifican 3 posibles causas asignables en todos excepto en el caso correspondiente al grado de borrosidad 0.1. Se eliminan las muestras señaladas por los

**Tabla 5.1:** Estadístico  $T^2$  para funciones de pertenencia con diferentes Grados de Borrosidad: Proceso de producción de comida congelada

muestra	$T^2$ (0.1)	$T^2$ (0.125)	$T^2$ (0.15)	$T^2$ (0.2)	$T^2$ (original)
1	2.41	1.97	1.99	2.19	2.60
2	6.28	5.59	5.82	6.25	6.09
3	0.77	0.33	0.13	0.19	0.28
4	4.02	2.61	2.80	3.49	3.76
5	5.83	7.43	7.63	6.88	7.69
6	3.29	3.74	3.73	3.53	3.34
7	0.58	0.38	0.32	0.38	0.35
8	0.31	0.17	0.20	0.27	0.27
9	0.40	0.40	0.31	0.28	0.36
10	22.25	22.47	22.63	22.54	22.10
11	17.41	22.91	22.25	19.57	19.71
12	1.72	4.36	4.00	2.64	2.32
13	3.08	2.85	2.91	2.80	3.10
14	3.35	3.88	3.50	3.13	3.71
15	2.28	2.12	2.06	2.12	2.17
16	1.13	0.86	0.76	0.88	0.61
17	0.30	0.44	0.43	0.36	0.41
18	1.36	1.89	1.78	1.48	1.44
19	1.17	1.90	1.65	1.26	1.35
20	0.72	0.66	0.61	0.62	1.12

gráficos correspondientes y se calculan los nuevos parámetros para el gráfico de fase II.

**Gráfico de Control de Fase II** Una vez eliminadas las muestras correspondientes a los puntos hallados fuera de control, se calculan los nuevos parámetros que serán usados para calcular el estadístico  $T^2$  para el gráfico de Fase II . Este procedimiento se realiza para cada uno de los conjuntos



**Figura 5.4:** Comparativo de Gráficos de Control de Fase I para funciones de pertenencia con diferentes grados de borrosidad

de funciones de pertenencia.

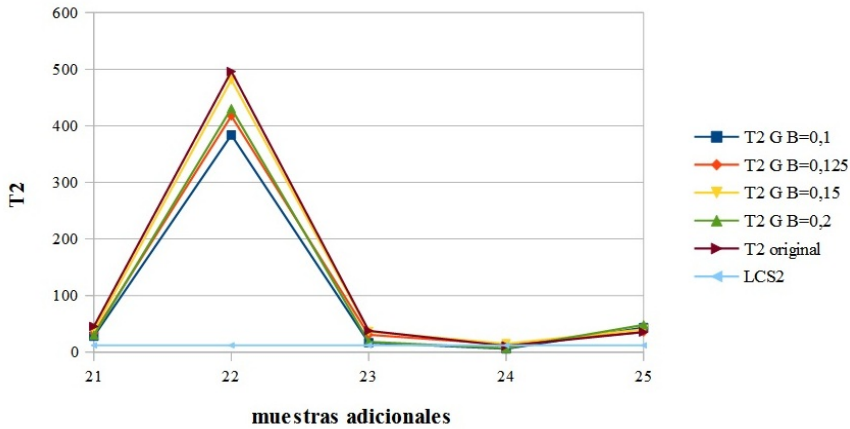
El límite de control superior para esta fase esta dado por:

$$LCS_2 = \frac{p(m^2 - 1)}{m(m - p)} F_{(p, m-p; \alpha)} = \frac{3(17^2 - 1)}{17(17 - 3)} 3.344 = 12.139$$

Usamos los datos adicionales proporcionados por Taleb et al. (2006) para calcular los correspondientes estadísticos  $T^2$  de esta fase. Al igual que en la fase I, en la Figura 5.5 se muestra un gráfico comparativo de los valores del estadístico  $T^2$  para el gráfico de control de Fase II para los distintos niveles del grado de borrosidad.

Como se puede apreciar en la figura 5.5, los resultados son similares para los tres conjuntos de funciones de pertenencia, sin embargo difieren en la clasificación de la muestra 24. Mientras el conjunto de funciones





**Figura 5.5:** Comparativo de Gráficos de Control de Fase II para funciones de pertenencia con diferentes grados de borrosidad

proporcionadas en el ejemplo de Taleb y los correspondientes a los grados de borrosidad 0.1 y 0.2 la clasifican como bajo-control los conjuntos con grados de borrosidad 0.125 y 0.15 la clasifica como fuera-de-control, esto pudiera ser debido a que los valores del estadístico para esa muestra son cercanos al valor del límite de control.

También cabe señalar que se ha elegido como valor de error Tipo I  $\alpha = 0.05$  para efectos de comparación con el ejemplo propuesto en el capítulo 3, sin embargo, estos límites producen *una mayor tasa de falsas alarmas*.

### 5.2.1.2 Error tipo I ( $\alpha$ )=0.0027

**Gráfico de Control de Fase I** Como sabemos los valores para los estadísticos  $T^2$  de primera fase no dependen del error tipo I, por tanto

son los ya mostrados en la tabla 5.1. En lo que si difieren es en cuanto a los límites de control. El límite de control superior para el gráfico de fase I, con  $p = 3$ ,  $m = 20$  y  $\alpha = 0.027$  es:

$$LCS_1 = \frac{(m - 1)^2}{m} B_{\left(\frac{p}{2}, \frac{m-p-1}{2}; \alpha\right)} = \frac{19^2}{20} 0.6511 = 10.42$$

La figura 5.6 muestra el gráfico comparativo de los valores de los estadísticos  $T^2$  para el gráfico de control de Fase I para este caso. Al igual que el caso anterior, como se aprecia en la figura, se obtienen similares resultados para los diferentes niveles de grado de borrosidad. Se identifican 2 posibles causas asignables en todos los gráficos, se eliminan las muestras correspondientes y se calculan los nuevos parámetros para el gráfico de control de fase II.

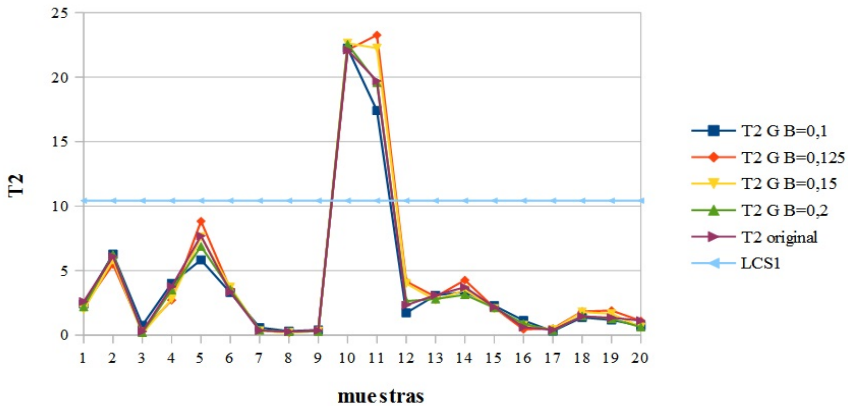


Figura 5.6: Estadístico  $T^2$  de Fase I para funciones de pertenencia con diferentes grados de borrosidad  $\alpha = 0.027$

**Gráfico de Control de Fase II** Una vez eliminados las muestras correspondientes a los puntos hallados fuera de control se calculan los nuevos parámetros usados para calcular el estadístico  $T^2$  para el gráfico de Fase II para las muestras adicionales. este procedimiento se realiza para cada uno de los conjuntos de funciones de pertenencia.

El límite de control superior para esta fase esta dado por:

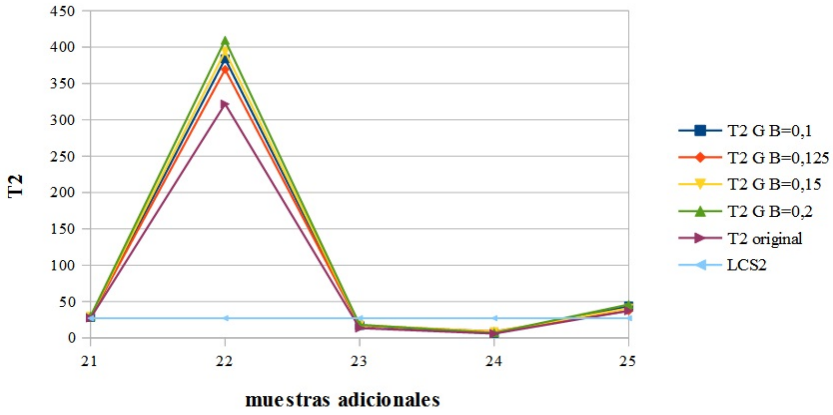
$$LCS_2 = \frac{p(m^2 - 1)}{m(m - p)} F_{(p, m-p; \alpha)} = \frac{3(18^2 - 1)}{18(18 - 3)} 7.5 = 26.92$$

Al igual que en el caso anterior se usan los datos adicionales proporcionados por Taleb et al. (2006) para calcular los correspondientes estadísticos  $T^2$  para cada muestra adicional de esta fase. Un gráfico comparativo de los valores del estadístico  $T^2$  para el gráfico de control de Fase II correspondientes a cada uno de los distintos niveles de grado de borrosidad, se muestra en la Figura 5.7.

Como se puede apreciar en la figura 5.7, se obtienen resultados similares para los tres conjuntos de funciones de pertenencia, las muestras 21, 22 y 25 fueron halladas fuera de control en todos los casos.

### 5.2.2. Ejemplo 2: Proceso de producción de porcelana

Se tomará como ejemplo, el propuesto por Taleb (2009) presentado en la sección 3.4.2 del capítulo 3, se comparan los resultados obtenidos en esa sección con los que se obtengan al aplicar, al igual que en el ejemplo anterior, cuatro nuevos conjuntos de funciones de pertenencia con grados de borrosidad 0.1, 0.125, 0.15 y 0.2.



**Figura 5.7:** Estadístico  $T^2$  de Fase II para funciones de pertenencia con diferentes grados de borrosidad  $\alpha = 0.027$

Los nuevos conjuntos de pertenencia se describen por los siguientes números triangulares difusos:

*Grado de borrosidad 0.1*

$$\begin{aligned}
 F_{11} &= (0, 0, 0.4) & F_{12} &= (0.3, 0.5, 0.7) & F_{13} &= (0.6, 1, 1) \\
 F_{21} &= (0, 0, 0.4) & F_{22} &= (0.35, 0.6, 0.75) & F_{23} &= (0.6, 1, 1) \\
 F_{31} &= (0, 0, 0.4) & F_{32} &= (0.25, 0.4, 0.65) & F_{33} &= (0.6, 1, 1)
 \end{aligned}$$

*Grado de borrosidad 0.125*

$$\begin{aligned}
 F_{11} &= (0, 0, 0.5) & F_{12} &= (0.25, 0.5, 0.75) & F_{13} &= (0.5, 1, 1) \\
 F_{21} &= (0, 0, 0.5) & F_{22} &= (0.35, 0.6, 0.85) & F_{23} &= (0.5, 1, 1) \\
 F_{31} &= (0, 0, 0.5) & F_{32} &= (0.15, 0.4, 0.65) & F_{33} &= (0.5, 1, 1)
 \end{aligned}$$

*Grado de borrosidad 0.15*

$$\begin{aligned} F_{11} &= (0, 0, 0.6) & F_{12} &= (0.2, 0.5, 0.8) & F_{13} &= (0.4, 1, 1) \\ F_{21} &= (0, 0, 0.6) & F_{22} &= (0.3, 0.6, 0.9) & F_{23} &= (0.4, 1, 1) \\ F_{31} &= (0, 0, 0.6) & F_{32} &= (0.1, 0.4, 0.7) & F_{33} &= ((0.4, 1, 1)) \end{aligned}$$

*Grado de borrosidad 0.2*

$$\begin{aligned} F_{11} &= (0, 0, 0.8) & F_{12} &= (0.1, 0.5, 0.9) & F_{13} &= (0.2, 1, 1) \\ F_{21} &= (0, 0, 0.8) & F_{22} &= (0.15, 0.55, 0.95) & F_{23} &= (0.2, 1, 1) \\ F_{31} &= (0, 0, 0.8) & F_{32} &= (0.05, 0.45, 0.85) & F_{33} &= (0.2, 1, 1) \end{aligned}$$

### 5.2.2.1 Error tipo I( $\alpha$ )=0.05

**Gráfico de Control de Fase I** A partir de los datos proporcionados por Taleb (2009), presentados en el capítulo 3, se calculan los valores representativos y los correspondientes valores del estadístico  $T^2$  para cada muestra para los distintos niveles de grado de borrosidad, al igual que en el ejemplo anterior.

El límite de control superior para el gráfico de control de fase I, para  $p = 3$ ,  $m = 23$  y  $\alpha = 0.05$  es:

$$LCS_1 = \frac{(m-1)^2}{m} B_{\left(\frac{p}{2}, \frac{m-p-1}{2}; \alpha\right)} = \frac{22^2}{23} 0.3306 = 6.96$$

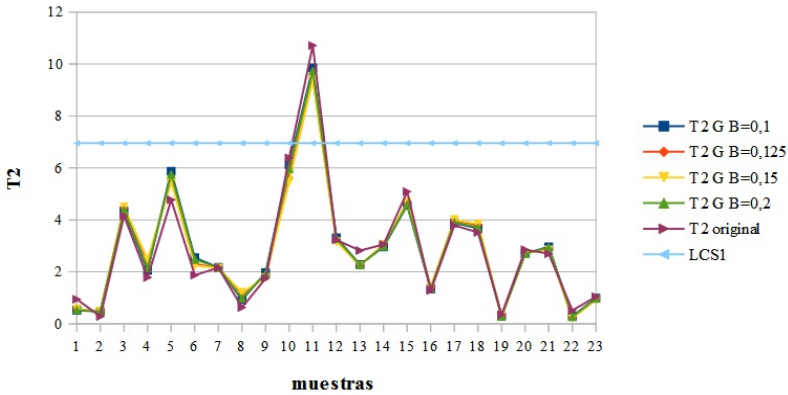
Los valores del estadístico  $T^2$  para los diferentes conjuntos de funciones de pertenencia se muestran en la tabla 5.2.

La figura 5.8 muestra un comparativo de los valores de los estadísticos

**Tabla 5.2:** Estadístico  $T^2$  para diferentes Grados de Borrosidad:  
Proceso de producción de porcelana

muestra	$T^2$ (0.1)	$T^2$ (0.125)	$T^2$ (0.15)	$T^2$ (0.2)	$T^2$ (original)
1	0.52	0.57	0.58	0.54	0.95
2	0.47	0.48	0.48	0.48	0.29
3	4.34	4.47	4.50	4.38	4.16
4	2.06	2.39	2.49	2.15	1.78
5	5.87	5.53	5.46	5.76	4.78
6	2.55	2.31	2.25	2.48	1.88
6 7	2.17	2.15	2.14	2.17	2.16
8	0.92	1.14	1.21	0.98	0.63
9	1.97	1.83	1.79	1.93	1.75
10	6.12	5.61	5.47	5.98	6.39
11	9.85	9.49	9.40	9.74	10.71
12	3.32	3.20	3.16	3.18	3.23
13	2.28	2.25	2.25	2.27	2.82
14	2.97	3.04	3.06	2.99	3.06
15	4.57	4.66	4.69	4.60	5.09
16	1.34	1.40	1.41	1.36	1.29
17	3.89	3.99	4.01	3.92	3.79
18	3.68	3.81	3.85	3.72	3.52
19	0.29	0.30	0.30	0.29	0.35
20	2.71	2.72	2.72	2.71	2.86
21	2.97	2.88	2.85	2.94	2.70
22	0.29	0.23	0.22	0.27	0.52
23	0.99	0.95	0.94	0.98	1.05

$T^2$  para distintos grados de borrosidad así como los correspondientes al ejemplo de Taleb (2009). Como podemos ver para todos los conjuntos de funciones de pertenencia se obtienen similares resultados. Se identifica 1 posible causa asignable en todos los gráficos, se elimina la muestra correspondiente y se calculan los nuevos parámetros para el gráfico de fase II.



**Figura 5.8:** Comparativo de Gráficos de control de Fase I para funciones de pertenencia con diferentes grados de borrosidad: Proceso de producción de porcelana

**Gráfico de Control de Fase II** Una vez eliminada la muestra correspondiente al punto hallado fuera de control se calculan los nuevos parámetros que serán usados para calcular el estadístico  $T^2$  para el gráfico de Fase II para muestras adicionales. Este procedimiento se realiza para cada uno de los conjuntos de funciones de pertenencia.

El límite de control superior para esta fase esta dado por:

$$LCS_2 = \frac{p(m^2 - 1)}{m(m - p)} F_{(p, m-p; \alpha)} = \frac{3(22^2 - 1)}{22(22 - 3)} 3.127 = 10.84$$

Usamos los datos adicionales proporcionados por Taleb (2009) para calcular los correspondientes estadísticos  $T^2$  para cada muestra adicional de esta fase. Al igual que en la fase I, en la Figura 5.9 se muestra un gráfico comparativo de los valores del estadístico  $T^2$  para el gráfico de

control de Fase II.

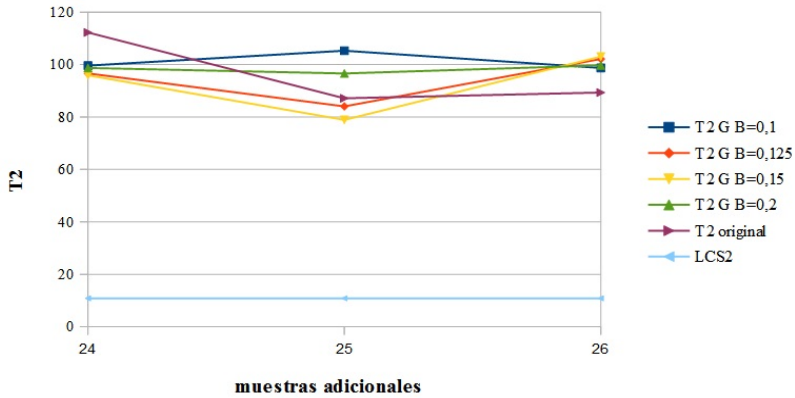


Figura 5.9: Comparativo Gráficos de control de Fase II para diferentes grados de borrosidad: Proceso de producción de porcelana

Como se puede apreciar en la figura 5.9, los resultados son similares para los distintos niveles de grado de borrosidad en cuanto a la clasificación, pero difieren en cuanto a la forma de las curvas trazadas por las muestras.

### 5.2.2.2 Error tipo I( $\alpha$ )=0.0027

**Gráfico de Control de Fase I** El límite de control superior para el gráfico de fase I, con  $p = 3$ ,  $m = 20$  y  $\alpha = 0.027$  es:

$$LCS_1 = \frac{(m - 1)^2}{m} B_{\left(\frac{p}{2}, \frac{m-p-1}{2}; \alpha\right)} = \frac{22^2}{23} 0.517 = 10.88$$

La figura 5.10 muestra el gráfico comparativo de los valores de los



estadísticos  $T^2$  para el gráfico de control de Fase I para este caso. Como se puede apreciar no se identifican posibles causas asignables en alguno de los gráficos, por tanto todas las muestras serán usadas para calcular los parámetros para el gráfico de control de fase II.

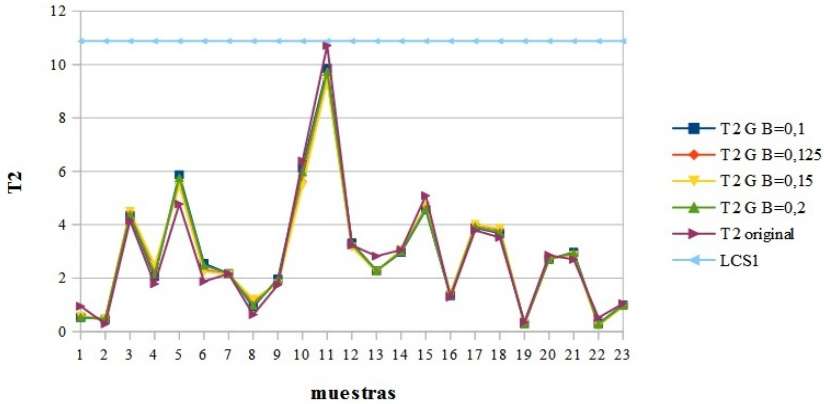


Figura 5.10: Estadístico  $T^2$  de Fase I para diferentes grados de borrosidad  $\alpha = 0.027$ : Proceso de producción de porcelana

**Gráfico de Control de Fase II** El límite de control superior para esta fase esta dado por:

$$LCS_2 = \frac{p(m^2 - 1)}{m(m - p)} F_{(p, m-p; \alpha)} = \frac{3(23^2 - 1)}{23(23 - 3)} 6.649 = 22.90$$

Al igual que en el caso anterior se usan los datos adicionales proporcionados por Taleb (2009) para calcular los correspondientes estadísticos  $T^2$  para cada muestra adicional de esta fase. Un gráfico comparativo de los valores del estadístico  $T^2$  para el gráfico de control de Fase II se muestra en la Figura 5.11.

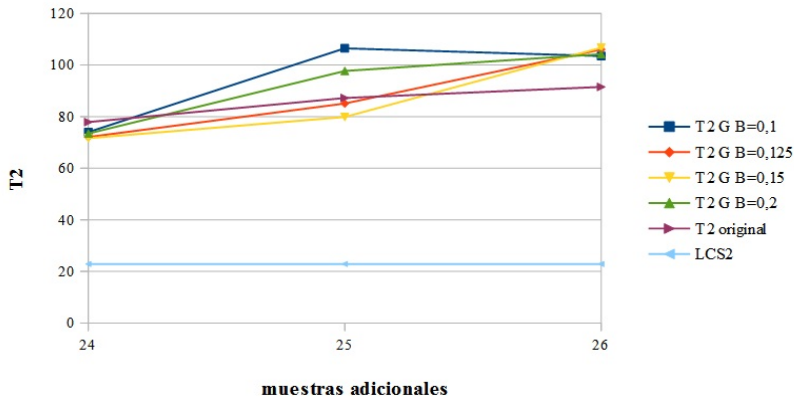


Figura 5.11: Estadístico  $T^2$  de Fase II para diferentes grados de borrosidad  $\alpha = 0.027$ : Proceso de producción de porcelana

Como se puede observar, en la figura 5.7, se obtienen resultados similares para los tres conjuntos de funciones de pertenencia, todas las muestras fueron halladas fuera-de-control en todos los casos.

### 5.3. Sensibilidad del rendimiento del gráfico de control $T^2$ MMD frente al grado de borrosidad de las funciones de pertenencia.

En este análisis, se toma como referencia el ejemplo proporcionado por Taleb et al. (2006), es decir, se supondrá que se quiere controlar tres características de calidad definidas mediante variables lingüísticas. Como se mencionó en la introducción este análisis se realizará en base al error de tipo II correspondiente a cada uno de los cambios planteados en el ejemplo, estos se pueden ver en la tabla 4.3.

A todos los términos lingüísticos en cada conjunto de términos se les asignó el mismo grado de borrosidad, se diseñaron funciones de pertenencia para proporcionar cuatro grados diferentes de borrosidad: 0.1, 0.125, 0.15 y 0.2. Para efectos de comparación se usarán también las funciones de pertenencia proporcionadas por Taleb et al. (2006).

Para el cálculo de los parámetros a ser empleados para el gráfico de control de Fase II se utilizaron las muestras generadas mediante simulación en el capítulo 4 mostrados en la tabla 4.1, con el fin de asegurar que los datos provienen de un proceso bajo control.

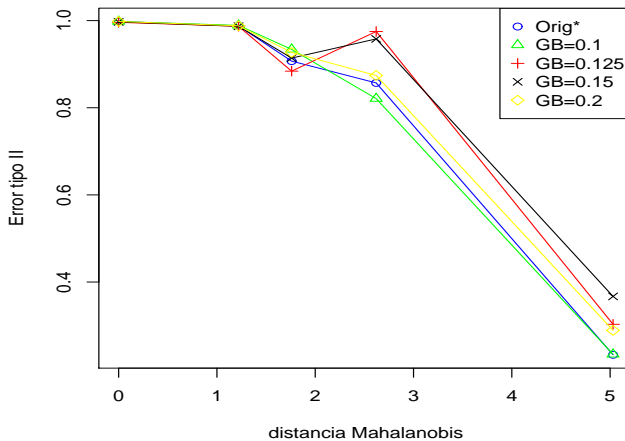
Para calcular el error de tipo II correspondiente a cada cambio planteado, se generan 10000 muestras usando como parámetros para la matriz de correlaciones, la matriz de varianzas y covarianzas de los valores representativos correspondientes a las 50 muestras generadas, y como vector de proporciones el dado por la tabla 4.3 respectivo. Para cada uno de los cambios planteados, se calcularon los estadísticos  $T^2$  correspondientes a las 10000 muestras generadas y se calculó la fracción de éstos que caen dentro de los límites de control, dando así una estimación de la probabilidad de error tipo II. El procedimiento se repitió cuatro veces, y las medias de las probabilidades de error tipo II resultantes aparecen en la tabla 5.3. Cuatro repeticiones se consideraron suficientes, ya que el coeficiente de variación en todos los casos era bastante pequeño (menos de 0.01) (Raz y Wang, 1990).

La figura 5.12 muestra las curvas características de operación para los grupos de funciones de pertenencia propuestos correspondientes a los distintos niveles de grado de borrosidad. Se puede notar cierta anomalía en los casos de grado de borrosidad 0.125 y 0.15 para pequeñas distancias de Mahalanobis.

Para un mejor entendimiento de la relación entre el error de tipo

**Tabla 5.3:** Error de tipo II vs Distancia de Mahalanobis para funciones de pertenencia con diferentes grados de borrosidad

	d=0	d=1,22	d=1,76	d=2.62	d=5,03
original	0.998	0.988	0.907	0.857	0.233
GB=0.1	0.998	0.989	0.934	0.821	0,234
GB=0.125	0.996	0.987	0.884	0.975	0.303
GB=0.150	0.997	0.987	0.914	0.958	0.367
GB=0.2	0.998	0.988	0.927	0.874	0.289



**Figura 5.12:** Curvas OC para diferentes grados de borrosidad ( $m = 50$ )

II, el grado de borrosidad y la distancia de Mahalanobis, se construye el gráfico de superficie mostrado en la figura 5.13 y en la figura 5.14 el correspondiente gráfico de contorno. en este gráfico se puede apreciar que el grado de borrosidad de las funciones de pertenencia no influye de forma considerable sobre el rendimiento mientras que el factor dominante en el

cambio del error tipo II es la distancia de Mahalanobis, es decir, el grado de descorrección experimentado por el proceso objeto de control. Este gráfico se construyó en base a las 50 muestras generadas por simulación.

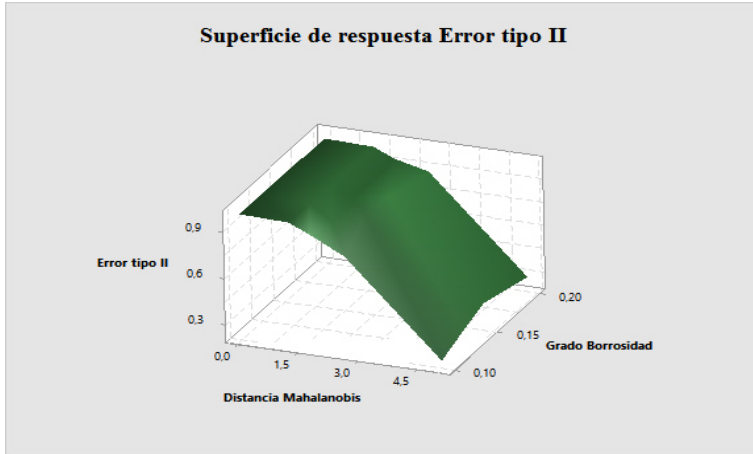


Figura 5.13: Gráfico de superficie: Error tipo II vs Grado de Borrosidad y Distancia Mahalanobis ( $m = 50$ )

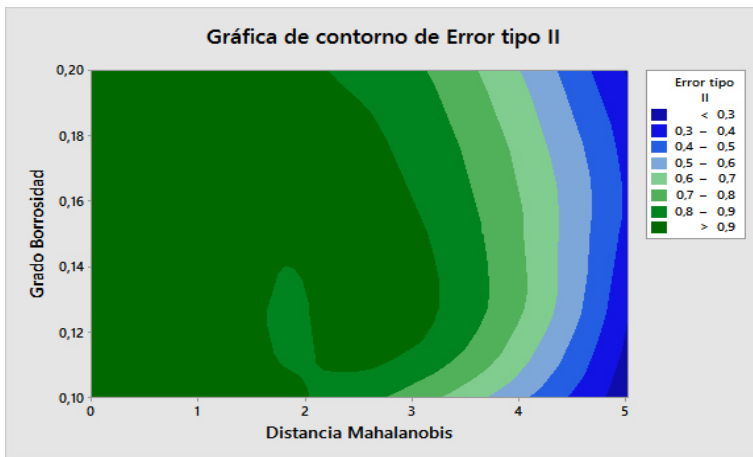
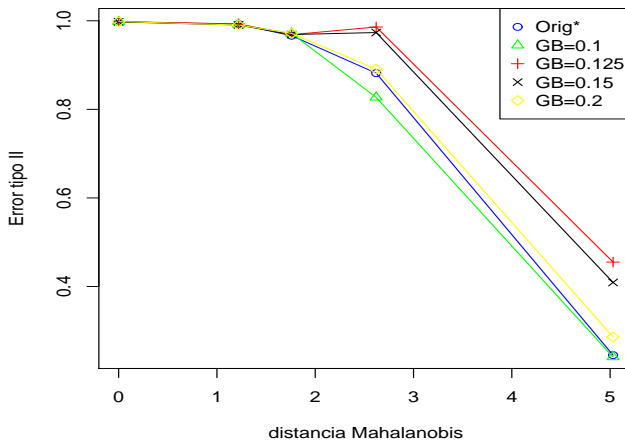


Figura 5.14: Gráfico de contorno: Error tipo II vs Grado de Borrosidad y Distancia Mahalanobis ( $m = 50$ )

Con el fin de verificar este comportamiento se repite el proceso para  $m = 100$  y  $m = 200$  muestras. Estas muestras se generan mediante simulación.

**$m=100$**

La figura 5.15 muestra las curvas características de operación para los diferentes grados de borrosidad tomando como base para el cálculo de los parámetros del gráfico 100 muestras generadas por simulación. Como se puede apreciar el comportamiento de las curvas es similar al caso anterior.



**Figura 5.15:** Curvas OC para diferentes Grados de Borrosidad ( $m = 100$ )

Las figuras 5.16 y 5.17 muestran la superficie de respuesta y el contorno del error tipo II para  $m = 100$ . Si bien es cierto la forma de la superficie de respuesta es similar al caso anterior, en la gráfica de contorno se puede apreciar que hay mayor regularidad en el comportamiento.

5.3. SENSIBILIDAD DEL RENDIMIENTO DEL GRÁFICO DE CONTROL T<sup>2</sup>MMD FRENTE AL GRADO DE BORROSIDAD DE LAS FUNCIONES DE PERTENENCIA.

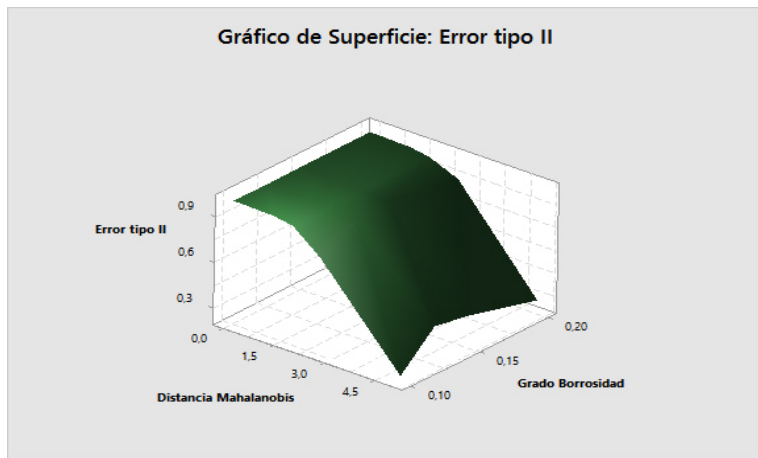


Figura 5.16: Gráfico de superficie: Error tipo II vs Grado de Borrosidad y Distancia Mahalanobis ( $m = 100$ )

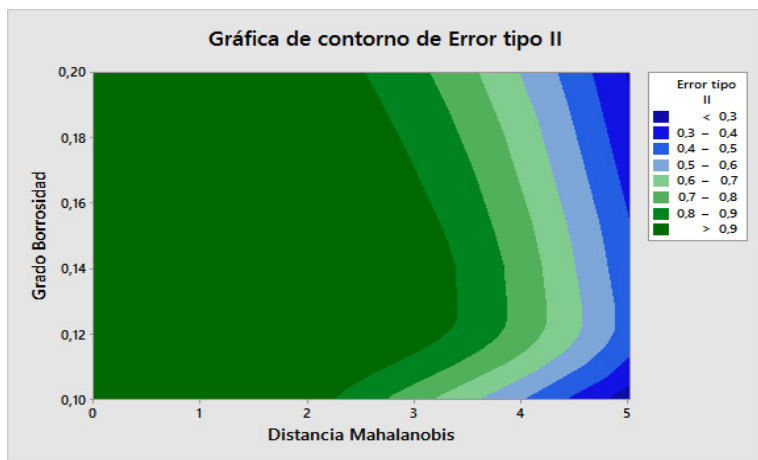


Figura 5.17: Gráfico de contorno: Error tipo II vs Grado de borrosidad y Distancia Mahalanobis  $m = 100$

$m=200$

En la figura 5.18 se muestra las curvas OC del gráfico de control para  $m = 200$ . En esta gráfico se puede apreciar que el comportamiento se mantiene.

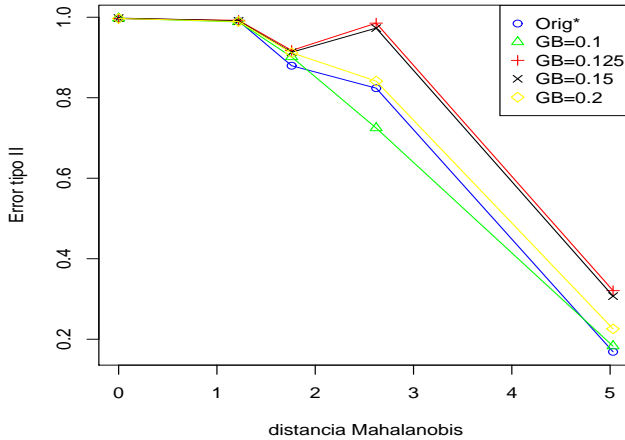


Figura 5.18: Curvas OC para diferentes Grados de Borrosidad ( $m = 200$ )

Las figuras 5.19 y 5.20 muestran la superficie de respuesta y el contorno del error tipo II para  $m = 200$ . De este caso y los anteriores podemos decir que el número de muestras influye en la regularidad del comportamiento de la superficie de respuesta y se verificar lo observa en el caso de  $m = 50$ .



5.3. SENSIBILIDAD DEL RENDIMIENTO DEL GRÁFICO DE CONTROL  $T^2$ MMD FRENTE AL GRADO DE BORROSIDAD DE LAS FUNCIONES DE PERTENENCIA.

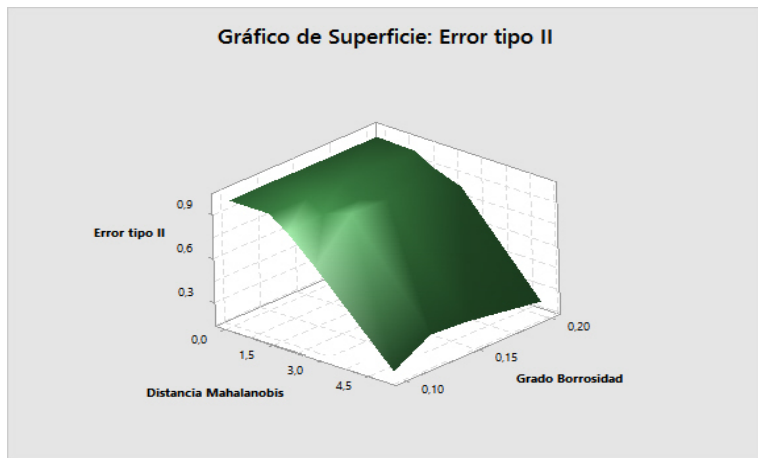


Figura 5.19: Gráfico de superficie: Error tipo II vs Grado de Borrosidad y Distancia Mahalanobis ( $m = 200$ )

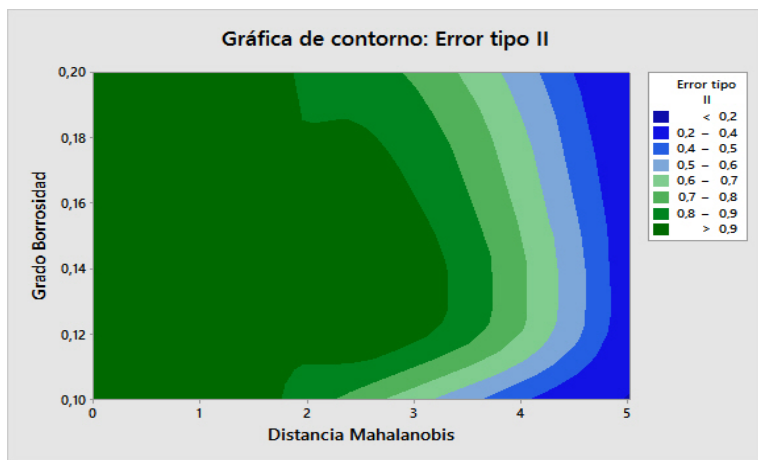


Figura 5.20: Gráfico de contorno: Error tipo II vs Grado de borrosidad y Distancia Mahalanobis ( $m = 200$ )

## 5.4. Conclusiones

La gran cantidad de parámetros aplicados en el comportamiento de estos gráficos de control dificulta la realización de estudios de sensibilidad. Para abordar este tema se ha optado por evaluar el comportamiento de los Gráficos de control frente a una característica propia del enfoque difuso, como es el grado de borrosidad.

El análisis de sensibilidad se ha hecho a partir de los ejemplos planteados en el capítulo 3, en los cuales se ha modificado las funciones de pertenencia para variar el grado de borrosidad. Se han considerado dos valores para el error de tipo I en ambos ejemplos y se ha analizado 5 niveles distintos del grado de borrosidad. Para el ejemplo de comidas congeladas, se ha observado que en el caso del error tipo I  $\alpha = 0.05$ , los resultados en cuanto la forma de la curva trazada por las muestras son similares para los distintos niveles de grados de borrosidad; sin embargo, difieren en la clasificación de la muestra 24 esto es debido a que el estadístico correspondiente está muy cerca del límite de control. No ocurre lo mismo en el caso de  $\alpha = 0.0027$ , para este caso los resultados son similares tanto para el gráfico de fase I como el de fase II. Para el ejemplo de la porcelana, los resultados son similares para todos los niveles de grado de borrosidad, en cuanto a la clasificación pero difieren en cuanto a la forma de las curvas trazadas por los estadísticos correspondientes a las muestras. esto ocurre tanto para  $\alpha = 0.05$  como para  $\alpha = 0.0027$ .

Por otra parte, se ha analizado el rendimiento del gráfico de control para distintos niveles de grado de borrosidad y diferentes grados de des-corrección del proceso, medidas a través de la distancia de Mahalanobis. Se construyeron las curvas de operación y los gráficos de superficie de respuesta y de contorno contorno para las 50 muestras generadas por simulación en el capítulo 4. Se encontró que el grado de borrosidad de las funciones de pertenencia no influye de forma considerable sobre el

rendimiento mientras que el factor dominante en el cambio del error tipo II es la distancia de Mahalanobis, es decir, el grado de descorrección experimentado por el proceso objeto de control. para verificar este comportamiento se repitió el procedimiento para  $m = 100$  y  $m = 200$  y se obtuvieron resultados similares.



## Conclusiones y líneas futuras de investigación

### 6.1. Conclusiones

La disminución de la variabilidad en la fabricación de productos cuyas características de calidad se definen mediante variables lingüísticas es un campo que sigue necesitando ser estudiado. Se añade a ello el hecho de que cuando se miden este tipo de variables no basta con aceptar o rechazar el producto, sino que son necesarios una serie de estados intermedios, generando así variables cualitativas con distribución multinomial. Si se tiene en cuenta además que generalmente se miden varias características de calidad que a su vez están correlacionadas el modelo de distribución necesario para el control y monitoreo de este tipo de variables puede volverse muy complejo.

En el presente trabajo se ha estudiado el uso de la teoría difusa para modelizar la variabilidad de las variables lingüísticas y definir de

ese modo una metodología que permita el monitoreo de este tipo de variables teniendo como base el gráfico  $T^2$  de Hotteling, brindando así una herramienta conocida y sencilla de aplicar.

Para ello, y siguiendo los objetivos *“Justificar el uso de la teoría difusa para el manejo de características de calidad caracterizadas en forma lingüística vs la teoría probabilística”* y *“Analizar y comparar los distintos gráficos y metodologías propuestas en la literatura existente, para identificar los elementos que permitan el diseño del gráfico de control difuso multinomial para atributos multivariantes”*, se ha realizado, como primer paso, una revisión de los conceptos básicos del Control Estadístico de Procesos, los gráficos de control y su desarrollo. Se han revisado también los conceptos básicos de la Teoría Difusa y las diferentes propuestas de gráficos de control basados en teoría difusa encontrados en la literatura, tanto para casos univariantes como multivariantes. Se han localizado algunos gráficos con enfoque difuso relevantes para nuestro estudio, tales como los planteados por [Raz y Wang \(1990\)](#) para monitorear una variable multinomial, y a partir del cuál se han realizado varias nuevas propuestas y comparaciones, y el planteado por [Taleb et al. \(2006\)](#) para controlar varias variables multinomiales.

Con respecto al objetivo *“Proponer un gráfico de control capaz de trabajar en una situación multinomial multivariante con enfoque difuso”*, se ha desarrollado una propuesta, que consiste en plantear el diseño de un gráfico de control para variables multinomiales multivariantes correlacionadas. Mediante el uso de la teoría difusa se puede transformar las variables multinomiales multivariantes en sus valores representativos correspondientes. Hemos encontrado que cuando el método de defusificación utilizado es el promedio difuso, los valores representativos se aproximan a una distribución normal multivariante, haciendo posible el uso un gráfico de control  $T^2$  de Hotelling para una sola observación y evitando la complejidad inherente de las variables multinomiales. Se ha mostrado la aplicación del diseño planteado usando dos ejemplos de aplicación,

que han permitido además comparar los resultados obtenidos con otras alternativas planteadas en la literatura. Se ha encontrado que el modelo desarrollado es útil y se adecúa al objetivo planteado.

Para conseguir el objetivo *“Establecer una metodología que permita medir el rendimiento del gráfico de control propuesto”*, se ha desarrollado un algoritmo que permite la simulación de varias variables multinomiales correlacionadas mediante el método de simulación Monte Carlo. Mediante el uso del algoritmo en cuestión se ha realizado un estudio del rendimiento del gráfico de control planteado y se han estudiado las características del mismo cuando el vector de medias y la matriz de covarianzas son estimadas. Se han comparado dichas características con las del gráfico  $T^2$  de Hotelling estudiadas por diferentes autores. Este estudio prueba, por medio del uso de simulación, diversas características del modelo propuesto en el capítulo 3. La mayoría de las cuales, pese a ser intuitivamente ciertas, no poseen una demostración trivial, como se ha visto.

Se ha proporcionado además la implementación detallada en R de los algoritmos usados para la simulación de las variables multinomiales multivariantes y para la medición del rendimiento del gráfico de control  $T^2$ MMD. Las experiencias numéricas llevadas a cabo y descritas en el capítulo 4 confirman la eficacia y eficiencia de la propuesta realizada. Además se han usado estos algoritmos para analizar el rendimiento del gráfico de control con enfoque difuso desarrollado por Taleb et al. (2006) para compararlo con el gráfico propuesto, encontrando que se obtienen los mejores resultados de rendimiento para el gráfico  $T^2$ MMD con límites corregidos.

Por último, con respecto al objetivo *“Analizar la sensibilidad del rendimiento del gráfico de control propuesto frente al grado de borrosidad de las funciones de pertenencia empleadas en la construcción del gráfico”*, se ha realizado dicho análisis de sensibilidad. Este análisis se hizo considerando

dos aspectos: *los resultados*, en cuanto a la calificación de la muestra a analizar realizada por el gráfico (bajo-control o fuera-de-control); y *el rendimiento*, medido a través del error de tipo II. Se encontró que si bien es cierto que existen algunas diferencias en los resultados, estas diferencias no son significativas. En lo referente al rendimiento puede decirse que el grado de borrosidad no puede considerarse como un factor influyente en el rendimiento si se compara con el efecto que la distancia (distancia de Mahalanobis, es decir, el grado de descorrección del vector de medias de la muestra  $i$  a partir del vector de medias del estado bajo-control) ejerce sobre él.

En resumen, se han cubierto los objetivos planteados para el presente trabajo. Es decir, hemos sido capaces de diseñar un gráfico de control que permite monitorear características de atributos múltiples correlacionados. Por medio del uso de la teoría difusa ha sido posible transformar las múltiples variables multinomiales correlacionadas en normales multivariantes, permitiendo el uso del gráfico  $T^2$  de Hotelling para observaciones individuales. De ese modo se ha evitado la complejidad de la distribución de la multinomial multivariante. Por otro lado el trabajo realizado permite pensar en otras alternativas y aplicaciones, que se detallan en la sección siguiente.

## 6.2. Líneas futuras de investigación

La investigación llevada a cabo y los diferentes procedimientos realizados a lo largo de la misma, nos han llevado a plantear las siguientes líneas futuras de investigación:

Analizar el efecto que tendría sobre los resultados la forma de las funciones de pertenencia, puesto que para el desarrollo de la presente



tesis se trabajó con funciones de pertenencias triangulares previamente establecidas.

Incorporar un procedimiento que permita determinar las funciones de pertenencia a ser empleadas, previo a la aplicación del gráfico. Es decir, en lugar de usar funciones previamente establecidas, determinarlas a partir de calificaciones conocidas.

Profundizar el análisis de sensibilidad teniendo en cuenta otros factores que no han sido considerados (por ejemplo: el número de categorías de las variables multinomiales o el tamaño de la muestra).

Evaluar el uso de un enfoque probabilístico como alternativa al enfoque difuso y comparar los resultados.

Elaborar un software que permita la implementación de este tipo de gráficos en la industria.



# Anexos



## Algoritmo Generación de variables Multinomiales

Problema: "Generar  $m$  muestras de  $p$  variables multinomiales correlacionadas cada una de tamaño  $n$  y con un vector de proporciones conocido."

Para resolver el problema se construyó un programa en R Studio, que consta de los siguientes pasos:

1. Establecer los valores de  $m$ ,  $n$ ,  $p$ , el vector de proporciones  $q$ , el vector de medias  $\mu$  y la matriz de correlaciones a usar  $\Sigma$ .
2. Generar una observación normal multivariante de  $p$  variables de tamaño  $n$ , usando la función *mvrnorm* contenida en el paquete MASS. El número de variables queda determinado por las dimensiones del vector de medias y la matriz de correlación utilizadas.
3. Transformar las normales en multinomiales utilizando los vectores

de proporciones dados, utilizamos estos valores como los correspondientes cuantiles y procedemos de esta manera a discretizar cada una de las normales, contando el número de observaciones en cada cuantil.

4. Repetir el proceso  $m$  veces, para obtener el número de muestras deseadas.

**A continuación ejemplo del algoritmo implementado en R para el primer ejemplo del capítulo 4:**

```
# Estableciendo parámetros de inicio
m<-50
p<- 3
n<-220
mu<-c(0,0,0)
sigma1<-matrix(c(29825.62,-13398.86,-649.52,-13398.86,
25060.69,4826.8,-649.52,4826.8,19917.6),43,3)
sigma<-solve(sigma1)

# Generación de una muestra de tamaño 220 con 3 variables
Q1<-mvrnorm(n,mu,sigma)

# generación primera variable multinomial
Q11<-Q1[,1]
q11<-qnorm(0.942,0,1)*sqrt(sigma[1,1])
q21<-qnorm(0.977,0,1)*sqrt(sigma[1,1])
n11<-sum(Q11<=q11)
N21<-sum(Q11<=q21)
n12<-N21-n11
n13<-n-N21

# generación segunda variable multinomial
Q12<-Q1[,2]
q12<-qnorm(0.925,0,1)*sqrt(sigma[2,2])
q22<-qnorm(0.970,0,1)*sqrt(sigma[2,2])
n21<-sum(Q12<=q12)
N22<-sum(Q12<=q22)
n22<-N22-n21
```

```

n23<-n-N22

#generación tercera variable multinomial
Q13<-Q1[,3]
q13<-qnorm(0.774,0,1)*sqrt(sigma[3,3])
q23<-qnorm(0.98,0,1)*sqrt(sigma[3,3])
q33<-qnorm(0.996,0,1)*sqrt(sigma[3,3])
n31<-sum(Q13<=q13)
N23<-sum(Q13<=q23)
N33<-sum(Q13<=q33)
n32<-N23-n31
n33<-N33-N23
n34<-n-N33

# Creando primera fila matriz de datos
data<-data.frame(n11,n12,n13,n21,n22,n23,n31,n32,n33,n34)

Creando matriz de datos
for(i in 1:(m-1)) {
Q1<-mvrnorm(n,mu,sigma)

# generación primera variable multinomial Q11<-Q1[,1]
q11<-qnorm(0.942,0,1)*sqrt(sigma[1,1])
q21<-qnorm(0.977,0,1)*sqrt(sigma[1,1])
n11<-sum(Q11<=q11)
N21<-sum(Q11<=q21)
n12<-N21-n11
n13<-n-N21

# generación segunda variable multinomial
Q12<-Q1[,2]
q12<-qnorm(0.925,0,1)*sqrt(sigma[2,2])
q22<-qnorm(0.97,0,1)*sqrt(sigma[2,2])
n21<-sum(Q12<=q12)
N22<-sum(Q12<=q22)
n22<-N22-n21
n23<-n-N22

# generación tercera variable multinomial
Q13<-Q1[,3]
q13<-qnorm(0.774,0,1)*sqrt(sigma[3,3])

```

```

q23<-qnorm(0.98,0,1)*sqrt(sigma[3,3])
q33<-qnorm(0.996,0,1)*sqrt(sigma[3,3])
n31<-sum(Q13<=q13)
N23<-sum(Q13<=q23)
N33<-sum(Q13<=q33)
n32<-N23-n31
n33<-N33-N23
n34<-n-N33

# Construimos la primera muestra multinomial
T<-data.frame(n11,n12,n13,n21,n22,n23,n31,n32,n33,n34)

# con la ejecución del lazo competamos las muestras requeridas
data<-rbind(data,T)
}

```

**Tabla A.1:** muestras simuladas

<b>muestras</b>	$n_{11}$	$n_{12}$	$n_{13}$	$n_{21}$	$n_{22}$	$n_{23}$	$n_{31}$	$n_{32}$	$n_{33}$	$n_{34}$
1	205	9	6	203	12	5	171	40	9	0
2	206	9	5	200	13	7	175	43	2	0
3	205	8	7	208	9	3	167	47	5	1
4	204	8	8	210	5	5	168	49	3	0
5	213	2	5	209	7	4	176	42	1	1
6	206	9	5	200	10	10	165	47	8	0
7	209	7	4	205	9	6	167	46	4	3
8	209	9	2	206	6	8	169	45	4	2
9	205	8	7	201	10	9	174	36	6	4
10	204	13	3	198	14	8	178	34	8	0
11	210	6	4	202	11	7	162	52	5	1
12	212	5	3	213	6	1	162	53	5	0
13	208	5	7	203	8	9	170	46	3	1
14	207	6	7	208	7	5	169	41	6	4
15	206	6	8	199	14	7	172	42	4	2
16	211	6	3	202	11	7	164	52	4	0
17	205	6	9	207	6	7	172	42	5	1
18	203	10	7	203	10	7	170	46	4	0
19	210	7	3	211	4	5	187	27	6	0
20	205	9	6	208	6	6	165	50	3	2
21	209	7	4	197	13	10	180	33	7	0



<b>muestras</b>	$n_{11}$	$n_{12}$	$n_{13}$	$n_{21}$	$n_{22}$	$n_{23}$	$n_{31}$	$n_{32}$	$n_{33}$	$n_{34}$
22	210	8	2	201	11	8	158	57	4	1
23	203	10	7	192	16	12	174	43	3	0
24	212	5	3	203	9	8	155	61	2	2
25	211	6	3	210	7	3	169	43	6	2
26	203	10	7	198	12	10	170	43	4	3
27	211	6	3	207	8	5	161	55	3	1
28	202	13	5	206	6	8	176	41	2	1
29	206	5	9	195	17	8	176	40	3	1
30	201	10	9	205	12	3	171	42	6	1
31	208	7	5	202	14	4	169	46	5	0
32	204	12	4	207	6	7	162	52	6	0
33	207	8	5	207	6	7	179	38	0	3
34	205	9	6	199	13	8	181	36	3	0
35	207	5	8	205	8	7	165	49	6	0
36	206	9	5	203	10	7	166	48	4	2
37	211	5	4	195	15	10	174	44	2	0
38	210	5	5	212	4	4	167	47	6	0
39	207	8	5	204	13	3	166	49	4	1
40	203	13	4	205	7	8	177	41	2	0
41	206	8	6	207	10	3	154	59	7	0
42	205	9	6	206	9	5	162	49	8	1
43	208	9	3	209	3	8	170	49	0	1
44	206	10	4	208	7	5	174	44	0	2
45	198	13	9	199	10	11	165	51	4	0
46	211	7	2	206	8	6	176	41	3	0
47	207	12	1	206	9	5	178	39	1	2
48	213	3	4	210	6	4	168	48	4	0
49	204	7	9	205	8	7	175	43	2	0
50	204	12	4	203	9	8	176	37	6	1



## Algoritmo para determinación de ARL mediante simulación

**Problema:** Determinar la longitud media de racha, **ARL**, para un error tipo I dado. Para resolver el problema se construyó un programa en R Studio, con el siguiente algoritmo:

1. Establecer los parámetros iniciales: número de variables  $p$ , número de muestras iniciales  $m$ , el error de tipo I  $\alpha$ , el número de repeticiones  $k$
2. Cargar la data generada usando el algoritmo anterior, que será usada como data preliminar para el cálculo de los parámetros del gráfico.
3. Calcular los parámetros a usar para el gráfico: vector de medias, matriz de varianzas y covarianzas y límite de control de Fase I.
4. Verificar que la data generada proviene de un proceso en-control, si lo es ir al siguiente paso. Si no, eliminar la muestras que estuvieran fuera de control y recalculer los parámetros.
5. Determinar el límite de control superior de fase II ( $UCL_2$ ) para la combinación deseada de  $p$  y  $m$  usando la ecuación (4.4).
6. Generar una muestra aleatoria usando el algoritmo anterior para representar la nueva información del proceso en el tiempo  $k$ .

7. Transformar la muestra generada en el vector de valores representativos correspondiente. Calcular  $T_K^2$ .
8. Comparar  $T_K^2$  con el límite de control. Si no se observa la señal de fuera-de-control ( $T_K^2 > UCL$ ), repita los pasos 6 y 7. Si se observa una señal, entonces vaya al paso 9.
9. Registrar la longitud de racha en un vector.
10. Repetir los pasos 2-9 hasta que se haya completado el número deseado de repeticiones.
11. Calcular el promedio de los valores registrado en el paso 9.

**A continuación ejemplo del algoritmo implementado en R para el primer ejemplo del capítulo 4:**

```
# Estableciendo parámetros de inicio
m<-50
p<- 3
α<-0.05
k<-10000

# Cargando datos
data<read.table("datosSimulados.txt",header=T)

# Cálculo de parámetros
# vector de medias A<-t(data[,2:4])
B<-t(data[,5:7])
C<-t(data[,8:11])
X<-c(mean(A[,1]),mean(A[,2]),mean(A[,3]))/220
Y<-c(mean(B[,1]),mean(B[,2]),mean(B[,3]))/220
Z<-c(mean(C[,1]),mean(C[,2]),mean(C[,3]),mean(C[,4]))/220
RVA<-c(0.25,1,2.25)/3
RVB<-c(0.5,1.25,2.5)/3
RVC<-c(0.25,1,2,2.75)/3
R1<-RVA*A/220
R2<-RVB*B/220
R3<-RVC*C/220

# prueba de normalidad
lillie.test(R1)
lillie.test(R2)
```

```

lillie.test(R3)
r1p<-mean(R1)
r2p<-mean(R2)
r3p<-mean(R3)

# Cálculo de matriz de varianzas y covarianzas muestrales
D1<-0
D2<-0
D3<-0
for(i in 1:m-1) {
d1<-R1[1,i+1]-R1[1,i]
d2<-R2[1,i+1]-R2[1,i]
d3<-R3[1,i+1]-R3[1,i]
D1<-rbind(D1,d1)
D2<-rbind(D2,d2)
D3<-rbind(D3,d3)
}
s11<-sum(D1*D1)/(2*(m-1))
s22<-sum(D2*D2)/(2*(m-1))
s33<-sum(D3*D3)/(2*(m-1))
s12<-sum(D1*D2)/(2*(m-1))
s13<-sum(D1*D3)/(2*(m-1))
s23<-sum(D2*D3)/(2*(m-1))
s<-matrix(c(s11,s12,s13,s12,s22,s23,s13,s23,s33),3,3)
SI<-solve(s)
si11<-SI[1,1]
si12<-SI[1,2]
si13<-SI[1,3]
si22<-SI[2,2]
si23<-SI[2,3]
si33<-SI[3,3]

# Límite de control superior Fase I
B<-qbeta(1- $\alpha$ ,p/2,(m-p-1)/2)
LCS1<-(m - 1)2*B/m

# Determinación de estadístico Fase I
T2H<-si11*(R1 - r1p)2+si22*(R2 - r2p)2+si33*(R3 - r3p)2
+2*si12*(R1-r1p)*(R2-r2p)+2*si13*(R1-r1p)*(R3-r3p)
+2*si23*(R2-r2p)*(R3-r3p)

```

```
# Verificar que las muestras provienen de un estado en-control
```

```
Límite de Control Superior FaseII
```

```
F<-qf(1- $\alpha$ ,p,m-p)
```

```
LCS2<-(p*(m2-1)*F)/(m*(m-p))
```

```
# Determinación ARL en-control
```

```
S<-0
```

```
n<-220
```

```
mu<-(0,0,0)
```

```
sigma<-s
```

```
# lazo para generar las repeticiones requeridas
```

```
for (i in 1:k){
```

```
cont<-0
```

```
T2<-0
```

```
T22<-0
```

```
LC<-LCS2
```

```
# lazo para contar el número de rachas
```

```
while(T2<LC) {
```

```
# Generación de una nueva muestra
```

```
Q1<-mvrnorm(n,mu,sigma)
```

```
# generación primera variable multinomial Q11<-Q1[,1]
```

```
q11<-qnorm(0.942,0,1)*sqrt(sigma[1,1])
```

```
q21<-qnorm(0.977,0,1)*sqrt(sigma[1,1])
```

```
n11<-sum(Q11<=q11)
```

```
N21<-sum(Q11<=q21)
```

```
n12<-N21-n11
```

```
n13<-n-N21
```

```
# generación segunda variable multinomial
```

```
Q12<-Q1[,2]
```

```
q12<-qnorm(0.925,0,1)*sqrt(sigma[2,2])
```

```
q22<-qnorm(0.97,0,1)*sqrt(sigma[2,2])
```

```
n21<-sum(Q12<=q12)
```

```
N22<-sum(Q12<=q22)
```

```
n22<-N22-n21
```

```
n23<-n-N22
```

```
# generación tercera variable multinomial
```

```
Q13<-Q1[,3]
```

```

q13<-qnorm(0.774,0,1)*sqrt(sigma[3,3])
q23<-qnorm(0.98,0,1)*sqrt(sigma[3,3])
q33<-qnorm(0.996,0,1)*sqrt(sigma[3,3])
n31<-sum(Q13<=q13)
N23<-sum(Q13<=q23)
N33<-sum(Q13<=q33)
n32<-N23-n31
n33<-N33-N23
n34<-n-N33

# creando la nueva muestra
A1<-c(n11,n12,n13)
B1<-c(n21,n22,n23)
C1<-c(n31,n32,n33,n34)

# calculando el estadístico  $T^2$ 
RR1<-sum(A1*RVA)/220
RR2<-sum(B1*RVB)/220
RR3<-sum(C1*RVC)/220
T2<-si11*(RR1 - r1p)2+si22*(RR2 - r2p)2+si33*(RR3 - r3p)2
+2*si12*(RR1-r1p)*(RR2-r2p)+2*si13*(RR1-r1p)*(RR3-r3p)
+2*si23*(RR2-r2p)*(RR3-r3p)
T22<-rbind(T22,T2)
cont<-cont+1
}
LR<-cont
LR
S<-rbind(S,LR)
}
ARL<-sum(S)/k

```





## Bibliografía

- Almond RG (1995). Discussion : Fuzzy Logic : Better Science? Or Better Engineering? *Technometrics*, 37(3):267–270.
- Amirzadeh V, Mashinchi M, Yaghoobi M (2008). Construction of Control Charts Using Fuzzy Multinomial Quality. *Journal of Mathematics and Statistics*, 4(1):26–31.
- Aparisi F, Avendaño G, Sanz J (2006). Techniques to interpret T2 control chart signals. *IIE Transactions (Revista JCR)*, 38(8):647–657.
- Aparisi F, Champ CW, García-Díaz JC (2004). A Performance Analysis of Hotelling's  $\chi^2$  Control Chart with Supplementary Runs Rules. *Quality Engineering*, 16(3):359–368. <http://dx.doi.org/10.1081/QEN-120027938>.
- Bersimis S, Psarakis S, Panaretos J (2007). Multivariate Statistical Process Control Charts : An Overview. *Quality and Reliability Engineering International*, 23(5):517–543. <http://onlinelibrary.wiley.com/doi/10.1002/qre.829/abstract>.
- Bojadziev G, Bojadziev M (1995). *Fuzzy Sets, Fuzzy Logic, Applications*. World Scientific, Singapore.
- Bradshaw CW (1983). A fuzzy set theoretic interpretation of economic control limits. *European Journal of Operational Research*, 13(4):403–408. <http://www.sciencedirect.com/science/article/pii/0377221783901005>.

- Campos G (2012). *Caracterización y optimización de los gráficos X-RL2 para el control de la posición y S-RL2 para el control de la dispersión*. Tesis Doctoral, Universitat Politècnica de València.
- Champ CW, Jones-Farmer LA, Rigdon SE (2005). Properties of the T2 Control Chart When Parameters Are Estimated. *Technometrics*, 47(4):437–445. <http://www.tandfonline.com/doi/abs/10.1198/004017005000000229>.
- Cheng CB (2005). Fuzzy process control: construction of control charts with fuzzy numbers. *Fuzzy Sets and Systems*, 154(2):287–303. <http://linkinghub.elsevier.com/retrieve/pii/S0165011405001107>.
- Chou YM, Mason RL, Young JC (2001). the Control Chart for Individual Observations From a Multivariate Non-Normal Distribution. *Communications in Statistics - Theory and Methods*, 30(8-9):1937–1949.
- Civanlar MR, Trussell HJ (1986). Constructing membership functions using statistical data. *Fuzzy Sets and Systems*, 18:1–13.
- Coppi R, Gil Ma, aL Kiers H (2006). The fuzzy approach to statistical analysis. *Computational Statistics & Data Analysis*, 51:1–14.
- Costa AF (1994). X charts with variable sample size. *Journal of Quality Technology*, 26(3):155–163.
- Costa AF (1997). X chart with variable sample size and sampling intervals. *Journal of Quality Technology*, 29(2):197–204.
- Cozzucoli PC (2009). Process Monitoring with Multivariate p-Control Chart. *International Journal of Quality, Statistics, and Reliability*, 2009:1–11.
- Crosier RB (1988). Multivariate Generalizations of Cumulative Sum Quality-Control Schemes. *Technometrics*, 30(3):291–303.
- Feili HR, Fekraty P (2010). Comparing Fuzzy Charts with Probability Charts and Using Them in a Textile Company. *The Journal of Mathematics and Computer Science*, 1(4):258–272.
- Fuhrmann G (1988). Fuzziness of Concepts and Concepts of Fuzziness. *Synthese*, 75(3):349–372.

- Guiffrida A, Nagi R (1998). Fuzzy set theory applications in production management research: a literature survey. *Journal of intelligent manufacturing*, 9(1):39–56. <http://link.springer.com/article/10.1023/A%3A1008847308326>.
- Gülbay M, Kahraman C (2007). An alternative approach to fuzzy control charts: Direct fuzzy approach. *Information Sciences*, 177(6):1463–1480. <http://linkinghub.elsevier.com/retrieve/pii/S0020025506002404>.
- Gülbay M, Kahraman C, Ruan D (2004).  $\alpha$  - Cut Fuzzy Control Charts for Linguistic Data. *International Journal of Intelligent Systems*, 19(12):1173–1196.
- Hersh HM, Caramazza A (1976). A fuzzy set approach to modifiers and vagueness in natural language. *Journal of Experimental Psychology: General*, 105(3):254–276.
- Hoadley B (1981). The quality measurement plan (QMP). *Bell System Technical Journal*, 60(2):215–273.
- Holmes DS, Mergen AE (1993). Improving the Performance of the T2 Control Chart. *Quality Engineering*, 5(4):619–625. <http://www.tandfonline.com/doi/abs/10.1080/08982119308919004>.
- Hotelling H (1947). Multivariate quality control. En: Eisenhart C, Hastay M, Wallis W (eds.), *Techniques of Statistical Analysis*, págs. 111–184. McGraw-Hill, New York.
- Hunter JS (1986). The exponentially weighted moving average. *Journal of Quality Technology*, 18(4):203–210.
- Jensen WA, Jones-Farmer LA, Champ CW, Woodall WH (2006). Effects of parameter estimation on control chart properties: A literature review. *Journal of Quality Technology*, 38(4):349–364. <GotoISI>://WOS:000241136100006.
- Juran JM, Gryna FM (1988). *Juran's Quality Control Handbook*. Industrial engineering series, McGraw-Hill. [https://books.google.es/books?id={\\_-}VTAAAAMAAJ](https://books.google.es/books?id={_-}VTAAAAMAAJ).
- Kanagawa A, Tamaki F, Ohta H (1993). Control charts for process average and variability based on linguistic data. *International Journal of Production Research*, 31(4):913–922.

- Kaufmann A, Gupta M (1985). *Introduction of Fuzzy Arithmetic Theory and Applications*. Van Nostrand Reinhold, New York.
- Kumar A, Mohapatra P (2012). A New Approach to Design of Fuzzy Multi-Attribute Control Charts. En: *POMS 23rd Annual Conference*, págs. 1–5. 1, Chicago, Illinois, U.S.A.
- Lavolette M, Seaman JW, Barrett JD, Woodall WH (1995). A probabilistic and statistical view of fuzzy methods. *Technometrics*, 37(3):249–261. <http://amstat.tandfonline.com/doi/abs/10.1080/00401706.1995.10484327>.
- Li J, Tsung F, Zou C (2014). Multivariate binomial/multinomial control chart. *IIE Transactions*, 46(5):526–542. <http://dx.doi.org/10.1080/0740817X.2013.849830>.
- Lowry CA, Montgomery DC (1995). A review of multivariate control charts. *IIE Transactions*, 27(6):800–810. <http://dx.doi.org/10.1080/07408179508936797>.
- Lowry CA, Woodall WH, Champ CW, Rigdon SE (1992). A Multivariate Exponentially Weighted Moving Average Control Chart. *Technometrics*, 34(1):46–53. <http://www.tandfonline.com/doi/abs/10.1080/00401706.1992.10485232>.
- Lu XS (1998). Control chart for multivariate attribute processes. *International Journal of Production Research*, 36(12):3477—3489. <http://dx.doi.org/10.1080/002075498192166>.
- Marcucci M (1985). Monitoring multinomial processes. *Journal of quality technology*, 17(2):86–91. <http://cat.inist.fr/?aModele=afficheN{&}cpsidt=9205449>.
- Mason RL, Tracy ND, Young JC (1995). Decomposition of T2 for multivariate control chart interpretation. *Journal of Quality Technology*, 27(2):99–108.
- Montgomery DC (2009). *Introduction to Statistical Quality Control, Sixth Edition*. John Wiley & Sons, Inc., New York. <https://books.google.es/books?id=SjeRMAEACAAJ>.
- Page E (1961). Cumulative sum charts. *Technometrics*, 3(1):1–9.
- Pandurangan A (2011). FUZZY MULTINOMIAL CONTROL CHART WITH VARIABLE SAMPLE SIZE. *International Journal of Engineering Science and Technology*, 3(9):6984–6991.

- Pastuizaca MN, Carrión A, Ruiz O (2015). Multivariate multinomial T2 control chart using fuzzy approach. *International Journal of Production Research*, 53(7):2225–2238.
- Pignatiello JJ, Runger GC (1990). Comparisons of multivariate CUSUM charts. *Journal of quality technology*, 22(3):173–186.
- Prabhu SS, Montgomery DC, Runger GC (1994). A combined adaptive sample size and sampling interval X control scheme. *Journal of Quality Technology*, 26(3):164–176.
- Prabhu SS, Runger GC, Keats JB (1993). X chart with adaptive sample sizes. *The International Journal Of Production Research*, 31(12):2895–2909.
- Raz T, Wang JH (1990). Probabilistic and membership approaches in the construction of control charts for linguistic data. *Production Planning & Control : The Management of Operations*, 1(3):147–157.
- Reynolds MR, Amin RW, Arnold JC, Nachlas JA (1988). X charts with variable sampling intervals. *Technometrics*, 30(2):181–192.
- Rowlands H, Wang LIREN (2000). An approach of fuzzy logic evaluation and control in SPC. *Quality Reliability Engineering Intelligent*, 16:91–98.
- Runger GC, Alt FB, Montgomery DC (1996). Contributors to a multivariate statistical process control chart signal. *Communications in Statistics–Theory and Methods*, 25(10):2203–2213.
- Ryan TP (2011). *Statistical Methods for Quality Improvement*. 3.<sup>a</sup>ed. Wiley Series in Probability and Statistics, John Wiley & Sons, Inc. <http://www.scopus.com/inward/record.url?eid=2-s2.0-84891583554&partnerID=tZ0tx3y1>.
- Saaty T (1975). Measuring the fuzziness of sets. *Journal of Cybernetics*, 4:53–61.
- Satterthwaite FE (1946). An Approximate Distribution of Estimates of Variance Components. *Biometrics Bulletin*, 2(6):110–114.
- Senturk S, Erginel N (2009). Development of fuzzy control charts using  $\alpha$ -cuts. *Information Sciences*, 179(10):1542–1551. <http://dx.doi.org/10.1016/j.ins.2008.09.022>.

- Sorooshian S (2013). Fuzzy Approach to Statistical Control Charts. *Journal of Applied Mathematics*, 2013. <http://dx.doi.org/10.1155/2013/745153>.
- Sullivan JH, Woodall WH (1996). A Comparison of Multivariate Control Charts for Individual Observations. *Journal of Quality Technology*, 28(4):398–408. <http://asq.org/qic/display-item/?item=11457>.
- Taleb H (2009). Control charts applications for multivariate attribute processes. *Computers & Industrial Engineering*, 56(1):399–410. <http://linkinghub.elsevier.com/retrieve/pii/S0360835208001319>.
- Taleb H, Limam M (2002). On fuzzy and probabilistic control charts. *International Journal of Production Research*, 40(12):2849–2863. <http://www.informaworld.com/openurl?genre=article&doi=10.1080/00207540210137602&magic=crossref||D404A21C5BB053405B1A640AFFD44AE3>.
- Taleb H, Limam M (2005). Fuzzy Multinomial Control Charts. En: *AI\* IA 2005: Advances in Artificial Intelligence*, págs. 553–563. Springer.
- Taleb H, Limam M, Hirota K (2006). Multivariate Fuzzy Multinomial Control Charts. *Quality Technology & Quantitative Management*, 3(4):437–453.
- Topalidou E, Psarakis S (2009). Review of multinomial and multiattribute quality control charts. *Quality and Reliability Engineering International*, 25(7):773–804.
- Turksen IB (1986). Measurement of membership functions. *Applications of fuzzy set theory in human factors*, págs. 55–67.
- Wang JH, Raz T (1990). On the Construction of Control Charts using Linguistic Variables. *International journal of production research*, 28(3):477–487.
- Williams R, Zigli R (1987). Ambiguity impedes quality in the service industries. *Quality Progress*, 20(7):14–17.
- Woodall WH (1997). Control charts based on attribute data: bibliography and review. *Journal of Quality Technology*, 29(2):172–183.
- Woodall WH (2000). Controversies and contradictions in statistical process control. *Journal of Quality Technology*, 32(4):341–350.
- Zadeh L (1965). Fuzzy Sets. *Information and Control*, 8:338–353.

- Zadeh L (1975a). The concept of a linguistic variable and its application to approximate reasoning-III. *Information Sciences*, 9(1):43–80.
- Zadeh L (1975b). The concept of a linguistic variable and its application to approximate reasoning-âII. *Information sciences*, 8(4):301–357.
- Zadeh L (1978). Fuzzy Sets as a basis for a theory of possibility. *Fuzzy Sets and Systems*, 102(2):135–155.
- Zadeh L (1980). Fuzzy sets versus probability. *Proceedings of the IEEE*, 68(3):421.
- Zadeh LA (1975c). The Concept of a Linguistic Variable and its Applications to Approximate Reasoning-I. *Information Sciences*, 8:199–249. <http://www.eecs.berkeley.edu/~zadeh/papers/TheConceptofaLinguisticVariableanditsApplicationsToApproximateReasoningI-1975.pdf>.





En memoria de mi hermosa sobrina.

## ***Dear God, Please Take Care Of My Little Girl***

*Dear God, please take care of my little girl,  
The one with big eyes, and long dark hair.  
She was special, as you should know,  
I really didn't want to let her go.*

*She touched the hearts of everyone she knew.  
Letting her go was so hard to do.  
Her smile could brighten up the darkest room.  
I wish you didn't have to take her so soon.*

*Could you sit and rock her and read her a story,  
She's probably afraid, please tell her don't worry.  
Tell her mommy loves her and wishes she could be here,  
But it won't be for many more years.*

*She loves to dance all kinds of songs,  
Please tell her that she did no wrong.  
Would you comfort her and hold her in your arms tight,  
And tell her she is missed every day and night.*

*Please tell her she is loved so very dear  
I'll say it every day for her to hear  
Her short life on earth is now completed  
For lessons I'm sure you felt I needed*

*Tell her I promise to see her again someday  
When that will be, I really can't say  
I promise to make up for the time that's past  
To hold her and comfort her, in my arms at last.*

By Terry Gouveia