

**ESTUDIO DEL EFECTO DE LA SELECCIÓN DE LOS NÚMEROS  
TRIANGULARES SOBRE EL DESEMPEÑO DE LOS ÍNDICES DE  
CAPACIDAD DE PROCESOS FUZZY**

**Einer David Mejía López**



UNIVERSIDAD DE CÓRDOBA  
FACULTAD DE CIENCIAS BÁSICAS  
DEPARTAMENTO DE MATEMÁTICAS Y ESTADÍSTICA  
MONTERÍA - CÓRDOBA  
2022

**ESTUDIO DEL EFECTO DE LA SELECCIÓN DE LOS NÚMEROS  
TRIANGULARES SOBRE EL DESEMPEÑO DE LOS ÍNDICES DE  
CAPACIDAD DE PROCESOS FUZZY**

**Einer David Mejía López**

Trabajo presentado como requisito para optar al título de  
**ESTADÍSTICO**

Directores:

**MYLADIS ROCIO COGOLLO FLÓREZ, PhD**  
**JUAN MIGUEL COGOLLO FLÓREZ, PhD**



UNIVERSIDAD DE CÓRDOBA  
FACULTAD DE CIENCIAS BÁSICAS  
DEPARTAMENTO DE MATEMÁTICAS Y ESTADÍSTICA  
MONTERÍA - CÓRDOBA  
2022

UNIVERSIDAD DE CÓRDOBA  
FACULTAD DE CIENCIAS BÁSICAS  
DEPARTAMENTO DE MATEMÁTICAS Y ESTADÍSTICA

Los jurados abajo firmantes certifican que han leído y recomiendan a la **Facultad de Ciencias Básicas** aprobar el trabajo de grado titulado: **“Estudio del efecto de la selección de los números triangulares sobre el desempeño de los índices de capacidad de procesos fuzzy”**, el cual es presentado por **Einer David Mejía López**, como requisito para optar al título de **Estadístico**.

**Fecha:** 01 de septiembre de 2022

**Director:** \_\_\_\_\_  
Myladis Rocio Cogollo Flórez, PhD

**Director:** \_\_\_\_\_  
Juan Miguel Cogollo Flórez, PhD

**Jurado:** \_\_\_\_\_  
Víctor Morales Ospina, PhD

**Jurado:** \_\_\_\_\_  
Alexander Correa Espinal, PhD

A mis padres **Manuel Antonio Mejía Argumedo y Yaneth del Carmen López Jiménez.**

Mi novia **María José Florían López.**



# Contenido

|  |           |
|--|-----------|
| <b>Resumen</b>   | <b>XI</b> |
| <b>Agradecimientos</b>   | <b>XV</b> |
| <b>Introducción</b>  | <b>1</b>  |
| <b>1. Fundamentación</b>   | <b>5</b>  |
| 1.1. Índices de capacidad de procesos . . . . .  | 5         |
| 1.2. Teoría Fuzzy . . . . .  | 7         |
| 1.2.1. Conjuntos Fuzzy . . . . .   | 7         |
| 1.2.2. $\alpha$ -Cortes . . . . .  | 8         |
| 1.2.3. Números Fuzzy . . . . .   | 9         |
| 1.3. Índices de procesos de capacidad fuzzy . . . . .  | 9         |
| 1.4. Enfoque de estimación de Buckley . . . . .  | 11        |
| 1.4.1. Algoritmo EDICP . . . . .   | 12        |
| 1.5. Estimación fuzzy del índice $C_{pm}$ : . . . . .  | 12        |
| <b>2. Estado del arte</b>  | <b>15</b> |
| <b>3. Análisis del diferencial asociado al número triangular en el <math>C_{pm}</math> fuzzy</b> | <b>17</b> |
| <b>4. Estudio de simulación y resultados</b>   | <b>23</b> |
| 4.1. Resultados para una distribución $Gamma(\alpha, \beta)$ . . . . .                           | 24        |
| 4.2. Resultados para una distribución $Lognormal(\alpha, \beta)$ . . . . .                       | 31        |
| 4.3. Análisis de la Capacidad en el $C_{pm}$ fuzzy de los datos simulados . . . . .              | 39        |
| <b>5. Aplicación</b>   | <b>41</b> |
| 5.1. Estudio de caso . . . . .   | 41        |
| 5.2. Conjunto de datos para el estudio de caso . . . . .   | 43        |
| 5.2.1. Conjunto de datos I . . . . .   | 43        |

|           |   |           |
|-----------|---|-----------|
| 5.2.2.    | Conjunto de datos II . . . . .  | 43        |
| 5.2.3.    | Estadísticas descriptivas de los conjuntos de datos y pruebas de bondad de ajuste . . . . . | 44        |
| 5.2.4.    | Capacidad del proceso $C_{pm}$ . . . . .  | 50        |
| <b>6.</b> | <b>Conclusiones y recomendaciones</b>   | <b>57</b> |
| 6.1.      | Conclusiones . . . . .  | 57        |
| 6.2.      | Recomendaciones y futuras investigaciones . . . . .   | 58        |
| <b>A.</b> | <b>Código de simulación en R</b>  | <b>63</b> |

# Lista de Tablas

|  |    |
|--|----|
| 4.1. Efecto de la construcción de los números triangulares en el $C_{pm}$ fuzzy con una distribución $Gamma(10.0, 0.34)$ con asimetría baja. . . . .   | 25 |
| 4.2. Efecto de la construcción de los números triangulares en el $C_{pm}$ fuzzy con una distribución $Gamma(2.5, 0.66)$ con asimetría moderada. . . . .  | 26 |
| 4.3. Efecto de la construcción de los números triangulares en el $C_{pm}$ fuzzy con una distribución $Gamma(1.0, 0.98)$ con asimetría alta. . . . .  | 27 |
| 4.4. Efecto de la construcción de los números triangulares $\mu$ y $\sigma^2$ en el $C_{pm}$ fuzzy con una distribución $Gamma(10.0, 0.34)$ con asimetría baja utilizando la mediana. . . . .      | 28 |
| 4.5. Efecto de la construcción de los números triangulares $\mu$ y $\sigma^2$ en el $C_{pm}$ fuzzy con una distribución $Gamma(2.5, 0.66)$ con asimetría moderada. . . . .                         | 29 |
| 4.6. Efecto de la construcción de los números triangulares $\mu$ y $\sigma^2$ en el $C_{pm}$ fuzzy con una distribución $Gamma(1.0, 0.98)$ con asimetría alta. . . . .                             | 30 |
| 4.7. Efecto de la construcción de los números triangulares en el $C_{pm}$ fuzzy con una distribución $Lognormal(0.8, 0.5^2)$ con asimetría baja. . . . .   | 32 |
| 4.8. Efecto de la construcción de los números triangulares en el $C_{pm}$ fuzzy con una distribución $Lognormal(0.5, 1.4^2)$ con asimetría moderada. . . . .                                       | 33 |
| 4.9. Efecto de la construcción de los números triangulares en el $C_{pm}$ fuzzy con una distribución $Lognormal(0.3, 3.0^2)$ con asimetría alta. . . . .   | 34 |
| 4.10. Efecto de la construcción de los números triangulares $\mu$ y $\sigma^2$ en el $C_{pm}$ fuzzy con una distribución $Lognormal(0.8, 0.5^2)$ con asimetría baja utilizando la mediana. . . . . | 35 |
| 4.11. Efecto de la construcción de los números triangulares $\mu$ y $\sigma^2$ en el $C_{pm}$ fuzzy con una distribución $Lognormal(0.5, 1.4^2)$ con asimetría moderada. . . . .                   | 36 |
| 4.12. Efecto de la construcción de los números triangulares $\mu$ y $\sigma^2$ en el $C_{pm}$ fuzzy con una distribución $Lognormal(0.3, 3.0^2)$ con asimetría alta. . . . .                       | 38 |
| 4.13. Capacidad de un proceso para el estimador $\tilde{C}_{pm}$ en la distribución indicada y n con asimetría baja, moderada y alta con $LSL = 1.6$ y $USL = 8.1$ . . . . .                       | 39 |
| 5.1. Niveles de calidad y valores $\tilde{C}_{pm}$ . . . . .   | 41 |
| 5.2. Rangos de recomendaciones nutricionales para adultos. . . . .   | 42 |
| 5.3. Conjunto I . . . . .  | 43 |
| 5.4. Conjunto II . . . . .   | 43 |
| 5.5. Estadísticas de los conjuntos de datos . . . . .  | 44 |



|  |    |
|--|----|
| 5.6. Parámetros estimados de la distribución hipotética de los conjuntos I y II . . . . .  | 45 |
| 5.7. Prueba de bondad de ajuste para la distribución hipotética . . . . .  | 46 |
| 5.8. Estadísticos y criterios de bondad de ajuste para los datos del conjunto I . . . . .  | 48 |
| 5.9. Medianas de bootstrap paramétricos e intervalo de confianza al 95 % para los datos<br>del conjunto I . . . . .  | 48 |
| 5.10. parámetros por máxima verosimilitud de una <i>Lognormal</i> para los datos del conjunto I  | 48 |
| 5.11. Estadísticos y criterios de bondad de ajuste para los datos del conjunto II . . . . .  | 49 |
| 5.12. Medianas de bootstrap paramétricos e intervalo de confianza al 95 % para los datos<br>del conjunto II . . . . .  | 50 |
| 5.13. parámetros por máxima verosimilitud de una <i>Lognormal</i> para los datos del conjunto II   | 50 |
| 5.14. Efecto de la construcción de los números triangulares en el $C_{pm}$ fuzzy . . . . .   | 51 |
| 5.15. Efecto de la construcción de los números triangulares en el $C_{pm}$ fuzzy asumiendo no<br>normalidad y asimetría . . . . .                              | 52 |
| 5.16. Efecto del cambio de números triangulares de $\mu$ y $\sigma^2$ en la estimación del PCI $C_{pm}$<br>fuzzy. . . . .                                      | 53 |
| 5.17. Efecto del cambio de números triangulares de $\mu$ y $\sigma^2$ en la estimación del PCI $C_{pm}$<br>fuzzy asumiendo no normalidad y asimetría . . . . . | 54 |
| 5.18. Índices de capacidad de proceso para el conjunto de datos indicado . . . . .   | 56 |

# Lista de Figuras

|  |    |
|--|----|
| 1.1. Número triangular fuzzy $TFN$ . . . . .   | 8  |
| 4.1. Gráficos de la función de probabilidad de la distribución $\text{Gamma}(\alpha, \beta)$ . . . . .   | 24 |
| 4.2. Gráficos de la función de probabilidad de la distribución $\text{lognormal}(\alpha, \beta)$ . . . . .   | 24 |
| 4.3. Números triangulares en el $C_{pm}$ fuzzy con distribución $\text{Gamma}(10.0, 0.34)$ , margen de error $\delta$ , $\delta_{IC_i}$ y $n = 100$ . . . . .                  | 25 |
| 4.4. Números triangulares en el $C_{pm}$ fuzzy con distribución $\text{Gamma}(10.0, 0.34)$ , margen de error $\delta = 0.0010$ y $n = (10, 30, 100)$ . . . . .                 | 25 |
| 4.5. Números triangulares en el $C_{pm}$ fuzzy con distribución $\text{Gamma}(2.5, 0.66)$ , margen de error $\delta$ , $\delta_{IC_i}$ y $n = 100$ . . . . .                   | 26 |
| 4.6. Números triangulares en el $C_{pm}$ fuzzy con distribución $\text{Gamma}(2.5, 0.66)$ , margen de error $\delta = 0.0010$ y $n = (10, 30, 100)$ . . . . .                  | 26 |
| 4.7. Números triangulares en el $C_{pm}$ fuzzy con distribución $\text{Gamma}(1.0, 0.98)$ , margen de error $\delta$ , $\delta_{IC_i}$ y $n = 100$ . . . . .                   | 27 |
| 4.8. Números triangulares en el $C_{pm}$ fuzzy con distribución $\text{Gamma}(1.0, 0.98)$ , margen de error $\delta = 0.0010$ y $n = (10, 30, 100)$ . . . . .                  | 27 |
| 4.9. Números triangulares $\mu$ y $\sigma^2$ en el $C_{pm}$ fuzzy con una distribución $\text{Gamma}(10.0, 0.34)$ utilizando la mediana y $n = 100$ . . . . .                  | 28 |
| 4.10. Números triangulares $\mu$ y $\sigma^2$ en el $C_{pm}$ fuzzy con una $\text{Gamma}(10.0, 0.34)$ utilizando la mediana, $\delta = 0.0010$ y $n = (10, 30, 100)$ . . . . . | 28 |
| 4.11. Números triangulares $\mu$ y $\sigma^2$ en el $C_{pm}$ fuzzy con una $\text{Gamma}(2.5, 0.66)$ utilizando la mediana y $n = 100$ . . . . .                               | 29 |
| 4.12. Números triangulares $\mu$ y $\sigma^2$ en el $C_{pm}$ fuzzy con una $\text{Gamma}(2.5, 0.66)$ utilizando la mediana, $\delta = 0.0010$ y $n = (10, 30, 100)$ . . . . .  | 29 |
| 4.13. Números triangulares $\mu$ y $\sigma^2$ en el $C_{pm}$ fuzzy con una $\text{Gamma}(1.0, 0.98)$ utilizando la mediana y $n = 100$ . . . . .                               | 31 |
| 4.14. Números triangulares $\mu$ y $\sigma^2$ en el $C_{pm}$ fuzzy con una $\text{Gamma}(1.0, 0.98)$ utilizando la mediana, $\delta = 0.0010$ y $n = (10, 30, 100)$ . . . . .  | 31 |
| 4.15. Números triangulares en el $C_{pm}$ fuzzy con distribución $\text{Lognormal}(0.8, 0.5^2)$ , margen de error $\delta$ , $\delta_{IC_i}$ y $n = 100$ . . . . .             | 32 |
| 4.16. Números triangulares en el $C_{pm}$ fuzzy con distribución $\text{Lognormal}(0.8, 0.5^2)$ , margen de error $\delta = 0.0010$ y $n = (10, 30, 100)$ . . . . .            | 32 |

|   |    |
|---|----|
| 4.17. Números triangulares en el $C_{pm}$ fuzzy con distribución $Lognormal(0.5, 1.4^2)$ , margen de error $\delta$ , $\delta_{IC_i}$ y $n = 100$ . . . . .                 | 33 |
| 4.18. Números triangulares en el $C_{pm}$ fuzzy con distribución $Lognormal(0.5, 1.4^2)$ , margen de error $\delta = 0.0010$ y $n = (10, 30, 100)$ . . . . .                | 33 |
| 4.19. Números triangulares en el $C_{pm}$ fuzzy con distribución $Lognormal(0.3, 3.0^2)$ , margen de error $\delta$ , $\delta_{IC_i}$ y $n = 100$ . . . . .                 | 35 |
| 4.20. Números triangulares en el $C_{pm}$ fuzzy con distribución $Lognormal(0.3, 3.0^2)$ , margen de error $\delta = 0.0010$ y $n = (10, 30, 100)$ . . . . .                | 35 |
| 4.21. Números triangulares $\mu$ y $\sigma^2$ en el $C_{pm}$ fuzzy con una $Lognormal(0.8, 0.5^2)$ utilizando la mediana y $n = 100$ . . . . .                              | 36 |
| 4.22. Números triangulares $\mu$ y $\sigma^2$ en el $C_{pm}$ fuzzy con una $Lognormal(0.8, 0.5^2)$ utilizando la mediana, $\delta = 0.0010$ y $n = (10, 30, 100)$ . . . . . | 36 |
| 4.23. Números triangulares $\mu$ y $\sigma^2$ en el $C_{pm}$ fuzzy con una $Lognormal(0.5, 1.4^2)$ utilizando la mediana y $n = 100$ . . . . .                              | 37 |
| 4.24. Números triangulares $\mu$ y $\sigma^2$ en el $C_{pm}$ fuzzy con una $Lognormal(0.5, 1.4^2)$ utilizando la mediana, $\delta = 0.0010$ y $n = (10, 30, 100)$ . . . . . | 37 |
| 4.25. Números triangulares $\mu$ y $\sigma^2$ en el $C_{pm}$ fuzzy con una $Lognormal(0.3, 3.0^2)$ utilizando la mediana y $n = 100$ . . . . .                              | 38 |
| 4.26. Números triangulares $\mu$ y $\sigma^2$ en el $C_{pm}$ fuzzy con una $Lognormal(0.3, 3.0^2)$ utilizando la mediana, $\delta = 0.0010$ y $n = (10, 30, 100)$ . . . . . | 38 |
| 5.1. Histograma, densidad y cuantiles teóricos vs cuantiles muestrales para el conjunto de datos I. . . . .   | 44 |
| 5.2. Histograma, densidad y cuantiles teóricos vs cuantiles muestrales para el conjunto de datos II. . . . .  | 45 |
| 5.3. Distribución acumulada empírica y teórica de una distribución normal con los datos del conjunto I. . . . .   | 46 |
| 5.4. Distribución acumulada empírica y teórica de una distribución normal con los datos del conjunto II. . . . .  | 46 |
| 5.5. Kurtosis y Skewness para los datos del conjunto I. . . . .   | 47 |
| 5.6. Kurtosis y Skewness para los datos del conjunto II. . . . .  | 47 |
| 5.7. Comparación de las distribuciones teóricas <i>Weibull</i> , <i>Lognormal</i> y <i>Gamma</i> para el conjunto de datos I. . . . .                                       | 47 |
| 5.8. Comparación de las distribuciones teóricas <i>Weibull</i> , <i>Lognormal</i> y <i>Gamma</i> para el conjunto de datos II. . . . .                                      | 49 |
| 5.9. Números triangulares en el $C_{pm}$ fuzzy con el conjunto de datos I. . . . .  | 51 |
| 5.10. Números triangulares en el $C_{pm}$ fuzzy con el conjunto de datos II. . . . .  | 51 |
| 5.11. Números triangulares en el $C_{pm}$ fuzzy asumiendo no normalidad y asimetría con el conjunto de datos I. . . . .   | 52 |
| 5.12. Números triangulares en el $C_{pm}$ fuzzy asumiendo no normalidad y asimetría con el conjunto de datos II. . . . .  | 52 |

|  |    |
|--|----|
| 5.13. Números triangulares $\mu$ y $\sigma^2$ en el $C_{pm}$ fuzzy con los datos del conjunto I. . . . .   | 54 |
| 5.14. Números triangulares $\mu$ y $\sigma^2$ en el $C_{pm}$ fuzzy con los datos del conjunto II. . . . .  | 54 |
| 5.15. Números triangulares $\mu$ y $\sigma^2$ en el $C_{pm}$ fuzzy asumiendo no normalidad y asimetría<br>con los datos del conjunto I. . . . .  | 55 |
| 5.16. Números triangulares $\mu$ y $\sigma^2$ en el $C_{pm}$ fuzzy asumiendo no normalidad y asimetría<br>con los datos del conjunto II. . . . . | 55 |



# Resumen

La evaluación del grado de cumplimiento de las especificaciones de los procesos de producción, es una labor continua y de interés para distintas industrias para determinar si el proceso es válido en la producción de un artículo bajo ciertos requisitos de calidad prefijados. Para ello, una de las herramientas empleadas son los índices de capacidad de procesos, los cuales permiten evaluar la capacidad de un proceso. Se utilizan varias estadísticas como  $C_p$ ,  $C_{pk}$ ,  $C_{pm}$  y  $C_{pmk}$  para estimar la capacidad de un proceso que en la mayoría de los casos se supone que se tiene una muestra grande con población Normal.

Cuando los límites de especificación o los estadísticos no son números concretos, la mejor elección es emplear un enfoque difuso para calcular dichos índices. Se realiza una revisión de la literatura sobre los índices de capacidad de procesos difusos y se encuentra que hasta la fecha, no se han explorado técnicas diferentes a los intervalos de confianza para elegir los números triangulares que intervienen en la construcción de los mismos.

En esta investigación se calcula formalmente el  $C_{pm}$ , bajo enfoques distintos de construcción de los números triangulares  $USL$ ,  $LSL$  y  $T$ , asumiendo a su vez la imprecisión de la media y dispersión del proceso. Este estudio se realiza vía simulación considerando las distribuciones *Log – normal* y *Gamma* debido a la variedad de formas asimétricas sesgadas positivamente y dos conjuntos de datos recolectados durante 27 semanas en la Universidad de Valparaíso-Chile para estudiar el efecto de estas selecciones en el desempeño de dicho índice. Los resultados hallados al examinar las diferentes formas de construcción de los números triangulares usando datos reales, demuestran que éste es un enfoque útil y confiable para aplicar en las diferentes industrias.

**Palabras Claves:** Conjunto difuso, índices de capacidad de procesos, estimación difusa, intervalos de confianza.



# Abstract

The evaluation of the degree of compliance with the specifications of the production processes is a continuous task and of interest to different industries to determine if the process is valid in the production of an article under certain requirements. predetermined quality. To do this, one of the tools used are the process capability indices, which allow the evaluation of the capability of a process. Various statistics such as  $C_p$ ,  $C_{pk}$ ,  $C_{pm}$  and  $C_{pmk}$  are used to estimate the capability of a process that in most cases is assumes that we have a large sample with a Normal population.

When the specification limits or statistics are not concrete numbers, the best choice is to use a fuzzy approach to compute such indices. A review of the literature on the capability indices of fuzzy processes is carried out and it is found that to date, techniques other than confidence intervals have not been explored to choose the triangular numbers involved in their construction.

In this investigation, the  $C_{pm}$  is formally calculated, under different construction approaches of the triangular numbers  $USL$ ,  $LSL$  and  $T$ , assuming in turn the imprecision of the media and the dispersion of the process. This study is carried out via simulation considering the *Log – normal* and *Gamma* distributions due to the variety of positively skewed asymmetric shapes and two data sets collected during 27 weeks at the University of Valparaiso-Chile to study the effect of these selections on the performance of said index. The results found by examining the different forms of construction of triangular numbers using real data, show that this is a useful and reliable approach to apply in different industries.

**Keywords:** Fuzzy set, process capability indices, fuzzy estimation, confidence intervals.





# Agradecimientos

En primera instancia doy gracias a Dios, por llenarme siempre de vida, salud, paciencia y sabiduría en mi proceso de aprendizaje.

Gracias a mis padres y a mi familia, por apoyarme de forma incondicional en mi proceso de formación, por brindarme comprensión, amor y paciencia para lograr esto posible.

A los profesores del Programa de estadística de la Universidad de Córdoba por compartir su conocimiento, sus experiencias y consejos brindados por instruirme de la mejor manera, por las experiencias brindadas a lo largo de mi formación como estadístico, en especial a los profesores Mario Morales Rivera, Juana Robles Gonzáles, Guillermo Martínez Flórez, Javier Ramírez, Osnamir Bru Cordero, Roger Tovar y Kevin Pérez.

A mis directores, Myladis Cogollo Flórez y Juan Miguel Cogollo Flórez, por su apoyo y disponibilidad para la construcción de este trabajo de grado. Además, por su forma apasionada de enseñar, su paciencia, comprensión y motivación en los momentos difíciles que pase en este proceso.

Y por último, a mis amigos y compañeros que me han acompañado durante la carrera, por la ayuda mutua y por ser parte de mi crecimiento como Estadístico, en especial a Javier Blanco Muñoz, Eduardo Moya Ramos, Rolando Martínez Chamorro, Fernando Lagares y Luis Negrete. Sin su apoyo incondicional, no hubiese llegado hasta este punto.

# Introducción

En la vida cotidiana es utilizada una definición tradicional de calidad como la aptitud para el uso de un producto o servicio que posee cierta cantidad de características que lo capacitan para satisfacer las necesidades del usuario que demanda. De este modo, la evaluación del grado de cumplimiento de las especificaciones de los procesos de producción, es una labor continua y de interés para distintas industrias para determinar si el proceso es valido en la producción de un articulo bajo ciertos requisitos de calidad prefijados[29].

Una de las herramientas empleadas son los índices de capacidad de procesos (ICP), estos permiten evaluar la capacidad de un proceso cuando se ve modificado por diferentes factores como el reajuste de métodos o la implementación de nuevas máquinas y se denominan como un número real que resume el comportamiento de las características del proceso con los límites objetivo y de especificación. Por lo general, para esta comparación se mide la proporción del ancho entre los límites de especificación del proceso al ancho natural de los límites de tolerancia que es la variación real medida por seis unidades de desviación estándar, y se supone que se tiene una muestra grande proveniente de una distribución Normal[21]. Kane [18] propuso el índice de capacidad potencial  $C_p$ , el cual compara la variación tolerada con la variación real. El índice de capacidad real  $C_{pk}$  considera las desviaciones de la media del proceso hacia los límites de especificación [35].

En la actualidad son muy pocas las propuestas para distribuciones no normales o con métodos muy robustos para evaluar la capacidad en el campo del control de calidad, en la mayoría de los ICP se supone que se tiene una muestra grande de una población normal. Pero, en aplicaciones prácticas es posible que no se pueda identificar la distribución de la característica de calidad o bien no se cuenta con una gran cantidad de muestras para la implementación de los métodos robustos.

Dado que los índices  $C_p$  y  $C_{pk}$  no tienen en cuenta la diferencia entre la media del proceso y su valor objetivo, Chan, Cheng et al. [23] consideraron esta diferencia para desarrollar el índice de capacidad general  $C_{pm}$ . Este índice también es propuesto por Hsiang y Taguchi [16] de manera independiente.

Para tener en cuenta este factor, en [16] se plantea el índice  $C_{pm}$  que es capaz de considerar en las especificaciones de calidad la variación con respecto al valor objetivo y los límites de especificación. Para cada uno de estos índices, un valor grande implica una mejor distribución de la característica de calidad. Además Pearn, Kotz y Johnson [38] propusieron el ICP  $C_{pmk}$  para límites de especificación asimétricos, que combina las características de los tres índices anteriores. Este índice  $C_{pmk}$  alerta al usuario cada vez que en el proceso la varianza aumenta y/o la media del proceso se desvía de su  $T$ .

Una característica común en los ICP clásicos propuestos en la literatura es que asumen que tanto

los límites de especificación como los parámetros asociados a los datos del proceso son números concretos. No obstante, en condiciones reales los datos que se utilizan para medir la capacidad de un proceso pueden estar asociados a condiciones de incertidumbre, tales como [25]:

- Los límites de especificación no siempre se pueden representar mediante números concretos, y se deben definir por números difusos,
- Los datos de la característica de interés pueden ser difusos, y
- La interpretación de un proceso capaz también es a veces difusa.

y por lo tanto, no se pueden aplicar directamente los índices de capacidad clásicos [30], como lo es en el caso de las fábricas que fabrican diferentes tipos de reguladores de voltaje positivo fijo de tres terminales para diversas aplicaciones que involucran la regulación local en tarjeta [13].

Una forma de modelar las situaciones en la que los índices de capacidad clásicos podrían ser no adecuados es empleando la lógica difusa. Después de la primera incursión del concepto fuzzy realizada por Zadeh [40], ésta se ha aplicado en diversos campos como la ingeniería, salud y ciencias naturales (ver por ejemplo, [30] y [36]). En particular, la lógica difusa también se ha aplicado en la medición de los índices de capacidad de procesos difusos y han surgido diversos ICP fuzzy:  $C_p$  fuzzy [39],  $C_{pk}$  fuzzy [22] y  $C_{pm}$  fuzzy [30, 19].

La estimación obtenida en los ICP fuzzy se ve afectada no sólo por las funciones de membresía seleccionada, sino también por la conformación de los números fuzzy que intervienen en la misma [20]. Específicamente, Parchami et al. [1] muestran que, en el mundo real, los diferenciales de números borrosos generalmente se pueden calcular mediante uno de los siguientes enfoques: (i) el método de Buckley, (ii) el conocimiento experto o la opinión del experimentador, (iii) un valor razonable fijo estimado mediante un método estadístico, y (iv) generando un número aleatorio de una distribución uniforme.

No obstante, en la literatura lo más usual es emplear los dos primeros métodos, asumiendo una dispersión constante [20], o basado en los intervalos de confianza [17, 28], y no es usual examinar el efecto de esta selección sobre el desempeño del ICP. En este trabajo se determina el efecto de la selección del diferencial de los números triangulares sobre el desempeño del índice de capacidad de proceso  $C_{pm}$  fuzzy en términos de calidad de la estimación, cuando se tienen límites de especificación no concretos.

En este trabajo se analiza el método de construcción de los números triangulares para obtener la mejor estimación del  $C_{pm}$  fuzzy al reducir el error estándar cuando se tiene límites de especificación no concretos y comparar su desempeño con los resultados obtenidos por otros autores teniendo en cuenta que los datos asociados al proceso son difusos, así como los parámetros  $\mu$  y  $\sigma^2$  bajo diferentes escenarios.

El resto de este documento está organizado de la siguiente manera. El capítulo 1 analiza brevemente la estimación puntual y de intervalo del rendimiento del proceso utilizando datos precisos e imprecisos del índice  $C_{pm}$  con el enfoque de Buckley. A continuación, el capítulo 2 presenta una relatoría donde se muestra los diferentes diferenciales empleados para construir los PCI difusos y

se muestra la falta de estudio del efecto en la estimación obtenida. A continuación, el capítulo 3 se analiza analíticamente el efecto del valor diferencial del número triangular asociados a los límites de especificación que interviene en el cálculo del  $C_{pm}$ . Posteriormente, El capítulo 4 propone un estudio de simulación y resultados para estudiar el desempeño de los estimadores del ICP definido en el capítulo 1 cuando los datos asociados al proceso provienen de distribuciones no normales. Para ilustrar la aplicabilidad del enfoque propuesto, se presenta una aplicación con datos reales de una industria de alimentos en el capítulo 5. Finalmente, las conclusiones se hacen en el último capítulo.



# Capítulo 1

## Fundamentación

En esta sección se abordaran las temáticas relacionadas con la propuesta de investigación: los índices de capacidad de procesos, lógica difusa y el índices de capacidad de procesos  $C_{pm}$  bajo un enfoque difuso, así como los  $\alpha$ -cortes y el enfoque de estimación de Buckley.

### 1.1. Índices de capacidad de procesos

El análisis de la capacidad de un proceso se determina comparando la variación real del proceso con respecto a sus límites de especificación, límite de especificación inferior (LSL, por su sigla en inglés) y límite de especificación superior (USL, por su sigla en inglés); así como su localización con respecto a un valor nominal de una característica de calidad dada [29, 15]. Para ello se utiliza la estimación de los Índices de Capacidad de Procesos (PCI, por su sigla en inglés), los cuales generalmente se limitan a un valor real, que es clasificado de acuerdo a unas reglas preestablecidas que permiten identificar si el proceso se encuentra bajo control.

En la literatura existen muchos desarrollos teóricos y aplicaciones sobre los índices de capacidad de procesos y se han aplicado en diferentes áreas. En particular se observa que la formulación de dichos índices esta sujeta a la distribución teórica de los datos asociados al proceso, y a si éste se encuentra bajo control. En la literatura se identifican los siguientes PCIs formulados bajo el supuesto de que los datos asociados al proceso sean concretos y se distribuyan normalmente:

- **Índice de Capacidad Potencial ( $C_P$ ):** Fue desarrollado por Kane en 1986 y es muy común en las industrias para evaluar las características de calidad individuales en las producciones en masa. Este se define como:

$$C_p = \frac{USL - LSL}{6\sigma} \quad (1.1)$$

donde  $\sigma$  es la desviación estándar del proceso. Se dice que un proceso es capaz si el valor del índice  $C_p$  asociado con el proceso es por lo menos 1 [18], y se puede estimar de la siguiente forma:

$$\hat{C}_p = \frac{USL - LSL}{6\hat{\sigma}} \quad (1.2)$$

donde  $\hat{\sigma}$  es la desviación estándar muestral:

$$\hat{\sigma} = S^2 = \frac{1}{n-1} \sum_{i=1}^n (x_i - \bar{X})^2 \quad (1.3)$$

■ **Índices de capacidad de procesos con límites de especificación superior e inferior:**

Estos índices fueron desarrollados para las situaciones en que se necesitará una especificación unilateral. También se aplican a procesos con doble especificación cuando se quiere evaluar la capacidad para cumplir con el  $USL$  o el  $LSL$ . Para un proceso con límite de especificación superior se tiene:

$$C_{pu} = \frac{USL - \mu}{3\sigma} \quad (1.4)$$

Para un proceso con límite de especificación inferior se tiene:

$$C_{pl} = \frac{\mu - LSL}{3\sigma} \quad (1.5)$$

donde  $\mu$  es la media del proceso y su estimación es  $\bar{X}$

- **Índice de capacidad real ( $C_{pk}$ ):** Este índice considera las desviaciones de la media del proceso con respecto al valor objetivo, y se expresa de la siguiente forma:

$$C_{pk} = \text{Min} \left\{ \frac{USL - \mu}{3\sigma}, \frac{\mu - LSL}{3\sigma} \right\} \quad (1.6)$$

que puede ser escrita, así:

$$C_{pk} = Cp(1 - k) \quad (1.7)$$

con  $k$  definida como:

$$k = \frac{2 |T - \mu|}{USL - LSL} = \frac{|\mu - T|}{d} \quad \text{donde} \quad d = \frac{USL - LSL}{2} \quad (1.8)$$

- **Índices de capacidad general ( $C_{pm}$ ):** Este índice es capaz de considerar la variabilidad de las especificación de calidad con respecto al valor objetivo  $T$ , y los límites de especificación. Este índice se define como:

$$C_{pm} = \frac{USL - LSL}{6\sqrt{\sigma^2 + (\mu - T)^2}} = \frac{USL - LSL}{6\tau} \quad (1.9)$$

Cuando  $\mu$  y  $\sigma^2$  son desconocidos, entonces se utilizan la media muestral y la varianza muestral para obtener la siguiente estimación:



$$\hat{C}_{pm} = \frac{USL - LSL}{6\sqrt{S^2 + (\bar{X} - T)^2}} \quad (1.10)$$

El término

$$\tau^2 = \sigma^2 + (\mu - T)^2 = E[(\bar{X} - T)^2]$$

incluye dos componentes de variabilidad: (i) variación con respecto a la media del proceso, y (ii) la desviación de el proceso respecto al valor objetivo (Pearn y Kotz 2006)[32]. Como  $E[(X - T)^2]$  es la pérdida esperada,  $C_{pm}$  también se denomina índice basado en la pérdida, donde el denominador de este PCI es la función de pérdida de Taguchi.

Chan [23] estudió el índice  $C_{pm}$  para límites de especificación asimétricos, cambiando uno de los dos límites de especificación a el objetivo para tener un límite de especificación simétrico como se muestra a continuación:

$$C_{pm} = \text{Min} \left\{ \frac{USL - T}{3\tau}, \frac{T - LSL}{3\tau} \right\} = \frac{d^*}{3\sqrt{\sigma^2 + (\mu - T)^2}} \quad (1.11)$$

donde  $d^* = \text{Min} \{D_u, D_l\}$ ;  $D_u = USL - T$ , y  $D_l = T - LSL$

Generalmente se encuentra que los datos asociados a un proceso no son números concretos, como por ejemplo el proceso de fabricación de los reguladores de voltaje en donde la característica clave es el voltaje de salida de un regulador de voltaje específico que inevitablemente tiene cierto grado de imprecisión lo cual hace necesario considerar los números difusos[38]. Una breve descripción de éstos números se presenta a continuación.

## 1.2. Teoría Fuzzy

La teoría difusa esta basada principalmente en dos conceptos: los conjuntos fuzzy y los  $\alpha$ -cortes, los cuales se presentan a continuación:

### 1.2.1. Conjuntos Fuzzy

Sea  $\Omega$  un conjunto, entonces un subconjunto difuso  $A$  de  $\Omega$  se define por su función  $A(x)$ , que arroja valores en el intervalo  $[0, 1]$  para todo  $x \in \Omega$ , y  $A(x)$  es interpretada como el grado de membresía de  $x$  en el conjunto difuso  $A$ . Esta función también se denota como  $\tilde{A}$ .

De forma general, un conjunto difuso  $A \in \mathbb{F}(\mathbb{R})$  se llama convexo, si todos los  $\alpha$ -cortes de  $A$  son intervalos, es decir, conjuntos convexos en el sentido habitual. De forma equivalente, un subconjunto fuzzy  $A$  de la línea real es convexo si para todo  $a, b, c \in \mathbb{R}$  se cumple que:

$$a \leq b \leq c \Rightarrow \eta_{\tilde{A}}(c) \geq \min \{ \eta_{\tilde{A}}(a), \eta_{\tilde{A}}(b) \} \quad (1.12)$$

con  $\eta_{\tilde{A}}(\cdot)$  un número real en el intervalo  $[0, 1]$ , donde  $\eta_{\tilde{A}}(\cdot)$  refleja la función de pertenencia de  $x$  en  $\tilde{A}$ .

### 1.2.2. $\alpha$ -Cortes

Si  $\tilde{A}$  es un subconjunto difuso de algún conjunto  $\Omega$ , entonces un  $\alpha$ -corte de  $\tilde{A}$ , escrito  $\tilde{A}[\alpha]$ , es definido como:

$$\tilde{A}[\alpha] = \{x \in \Omega | \tilde{A}(x) \geq \alpha\} \quad \forall \alpha, 0 < \alpha \leq 1 \quad (1.13)$$

Para cualquier número difuso  $\tilde{Q}$ , se tiene que  $\tilde{Q}[\alpha]$  es un intervalo cerrado, delimitado para  $0 \leq \alpha \leq 1$ , y se escribe:

$$\tilde{Q}[\alpha] = [q_1(\alpha), q_2(\alpha)] \quad (1.14)$$

donde  $q_1(\alpha)$  y  $q_2(\alpha)$  son funciones creciente y decreciente, respectivamente de  $\alpha$ , con  $q_1(1) = q_2(1)$ . Si  $\tilde{Q}$  tiene forma triangular, entonces:

1.  $q_1(\alpha)$  será una función continua, monótonamente creciente de  $\alpha$  en  $[0, 1]$ ,
2.  $q_2(\alpha)$  será una función continua monótonamente decreciente de  $\alpha$ ,  $0 \leq \alpha \leq 1$ , y
3.  $q_1(1) = q_2(1)$ .

Para el número triangular fuzzy (*TFN*) de la Figura 1.1 se obtiene  $N[\alpha] = [n_1(\alpha), n_2(\alpha)]$ , con  $n_1(\alpha) = 1.2 + 0.8\alpha$ , y  $n_2(\alpha) = 2.4 - 0.4\alpha$ ,  $0 \leq \alpha \leq 1$ . La ecuación para  $n_i(\alpha)$  se da de  $X$  en  $Y$ . Con el eje vertical ( $Y$ ), y el eje horizontal ( $X$ ), la ecuación  $n_1(\alpha) = 1.2 + 0.8\alpha$  significa  $X = 1.2 + 0.8y$ ,  $0 \leq y \leq 1$ . Es decir, el segmento de recta  $(1.2, 0)$  a  $(2, 1)$  en la Figura 1.1 se da como una función de  $X$  en  $Y$ , mientras que generalmente se expresa  $Y$  como una función de  $X$ . Así es como se hará para todos los  $\alpha$ -cortes de números fuzzy.

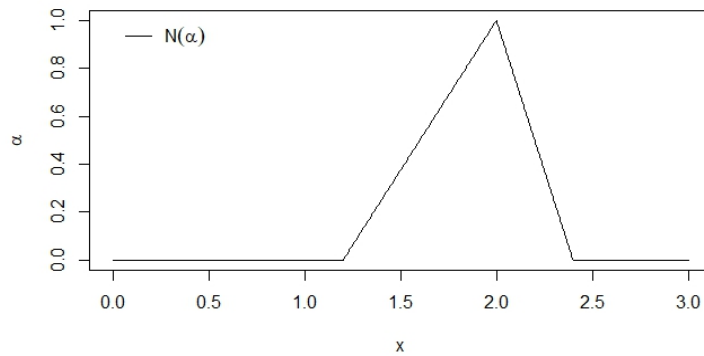


Figura 1.1: Número triangular fuzzy *TFN*

En la Figura 1.1 se observa que el gráfico es continuo y monótonamente creciente en  $[1.2, 2]$ ; y disminuye monótonamente en  $[2, 2.4]$ . Además, que su base está en el intervalo  $[1.2, 2.4]$  con vértice  $x = 2$ .

### 1.2.3. Números Fuzzy

De forma intuitiva se puede decir que los números difusos son valores reales imprecisos. Una idea básica que se relaciona con la intuición de que un número difuso debería representar un valor numérico determinado de forma imprecisa, es que la función del número fuzzy no debe tener puntos máximos locales distintos. Por lo tanto, los números difusos deberían ser "*convexos*" en algún sentido adecuado.

Un conjunto difuso es una función que varía entre 0 (falso) y 1 (verdadero), que asigna el grado de membresía de cada elemento en un conjunto. La forma de la función de membresía puede ser lineal (triangular o trapezoidal) o no lineal (gaussiano, generalizado Bell, gamma, etc.) según el sistema [28].

Un conjunto difuso con  $A \in \mathbb{F}(\mathbb{R})$  se denomina número(real) difuso, si  $A$  es convexo y si existe exactamente un número real  $a$  tal que  $\tilde{A}(a) = 1$ . Dado el caso que  $A$  sea solo convexo y normal, entonces  $A$  es llamado intervalo difuso. Por lo tanto, cada número difuso es también un conjunto difuso unimodal, Y los intervalos difusos unimodales son los números fuzzy.

En particular, un número fuzzy de forma triangular (TFN por sus siglas en inglés)  $N$  es un subconjunto difuso de los números reales que satisfacen [6]:

- (i)  $N(x) = 1$  para exactamente una  $x \in \mathbb{R}$
- (ii) Para  $\alpha \in (0, 1]$ , el  $\alpha$  - corte de  $N$  es un intervalo cerrado y acotado, que denotamos por  $N_\alpha = [n_1(\alpha), n_2(\alpha)]$ , donde  $n_1(\alpha)$  aumenta y  $n_2(\alpha)$  disminuye las funciones continuas.

Nótese que un número triangular fuzzy  $N$  se define a partir de tres números  $a < b < c$ , donde la base del triángulo es el intervalo  $[a, c]$  y su vértice está en  $x = b$ . Los números triangulares fuzzy se denotan como  $TFN = (a/b/c)$  o como  $TFN(a, b, c)$ . Un número difuso triangular  $TFN = (1.2/2/2.4)$  se muestra en la Figura 1.1.

Vemos que, por ejemplo para  $TFN(2)$  y  $TFN(1.6)$  se tiene que:

$$2 = 1.2 + 0.8y \implies y = 1 \quad ; \quad 1.6 = 1.2 + 0.8y \implies y = 0.5$$

de este modo,  $TFN(2) = 1$ ,  $TFN(1.6) = 0.5$ , etc.

## 1.3. Índices de procesos de capacidad fuzzy

La estimación de la media y la varianza son dos parámetros críticos en la estimación de los PCIs. Sus estimaciones difusas generan la oportunidad de producir los PCIs difusos; pero, además los límites de especificación también pueden considerarse como números difusos para incrementar la sensibilidad de los PCIs. En particular, suponga que se tiene un proceso difuso para los límites de especificación superior e inferior, con números triangulares fuzzy (NTF) denotados como:  $\tilde{U\tilde{S}L} = NTF(u_1, u_2, u_3)$

y  $L\tilde{S}L = NTF(l_1, l_2, l_3)$  que pertenecen a  $\mathbb{F}(\mathbb{R})$ . De modo que los  $\alpha$ -cortes para los límites de especificación son:

$$\begin{cases} U\tilde{S}L(\alpha) = [(u_2 - u_1)\alpha + u_1, (u_2 - u_3)\alpha + u_3] \\ L\tilde{S}L(\alpha) = [(l_2 - l_1)\alpha + l_1, (l_2 - l_3)\alpha + l_3] \end{cases} \quad (1.15)$$

Bajo estas definiciones, se han planteado algunas generalizaciones de los PCIs convencionales bajo un enfoque difuso, a saber:

- **Estimación fuzzy del índice  $C_p$ :**

$$\tilde{C}_p(\alpha) = \frac{U\tilde{S}L(\alpha) - L\tilde{S}L(\alpha)}{6\sqrt{\tilde{\sigma}^2(\alpha)}} \quad (1.16)$$

La cual bajo la formulación dada en (1.15) para los límites de especificación adquiere la siguiente forma:

$$\tilde{C}_p(\alpha) = NTF \left( \frac{(u_2 - u_1 + l_3 - l_2)\alpha + u_1 - l_3}{6\sqrt{\frac{(n-1)S^2}{(1-\alpha)\chi_{0.005}^2 + (n-1)\alpha}}}, \frac{(u_2 - u_3 - l_2 + l_1)\alpha + u_3 - l_1}{6\sqrt{\frac{(n-1)S^2}{(1-\alpha)\chi_{1-0.005}^2 + (n-1)\alpha}}} \right) \quad (1.17)$$

- **Índices fuzzy para los límites de especificación superior e inferior**

Para un proceso fuzzy con límite de especificación superior se tiene:

$$\tilde{C}_{pl}(\alpha) = \left( \frac{\mu_{l,\tilde{\sigma}}(\alpha) - LSL_r(\alpha)}{3\sigma_r(\alpha)}, \frac{\mu_{r,\tilde{\sigma}}(\alpha) - LSL_l(\alpha)}{3\sigma_l(\alpha)} \right) \quad (1.18)$$

Para un proceso con límite de especificación fuzzy inferior se tiene:

$$\tilde{C}_{pu}(\alpha) = \left( \frac{USL_l(\alpha) - \mu_{r,\tilde{\sigma}}(\alpha)}{3\sigma_r(\alpha)}, \frac{USL_r(\alpha) - \mu_{l,\tilde{\sigma}}(\alpha)}{3\sigma_l(\alpha)} \right) \quad (1.19)$$

Suponga que los límites de especificación superior e inferior son como en (1.15), entonces se obtiene que:

$$\tilde{C}_{pl}(\alpha) = NTF \left( \frac{\min \left( \bar{\chi} - Z_{\alpha/2} \frac{\sigma_i(\alpha)}{\sqrt{n}} \right) - [(l_2 - l_3)\alpha + l_3]}{3\sqrt{\frac{(n-1)S^2}{(1-\alpha)\chi_{0.005}^2 + (n-1)\alpha}}}, \frac{\max \left( \bar{\chi} + Z_{\alpha/2} \frac{\sigma_i(\alpha)}{\sqrt{n}} \right) - [(l_2 - l_1)\alpha + l_1]}{3\sqrt{\frac{(n-1)S^2}{(1-\alpha)\chi_{1-0.005}^2 + (n-1)\alpha}}} \right) \quad (1.20)$$

$$\tilde{C}_{pu}(\alpha) = NTF \left( \frac{[(u_2 - u_1)\alpha + u_1] - \max \left( \bar{\chi} + Z_{\alpha/2} \frac{\sigma_i(\alpha)}{\sqrt{n}} \right)}{3\sqrt{\frac{(n-1)S^2}{(1-\alpha)\chi_{0.005}^2 + (n-1)\alpha}}}, \frac{[(u_2 - u_3)\alpha + u_3] - \min \left( \bar{\chi} - Z_{\alpha/2} \frac{\sigma_i(\alpha)}{\sqrt{n}} \right)}{3\sqrt{\frac{(n-1)S^2}{(1-\alpha)\chi_{1-0.005}^2 + (n-1)\alpha}}} \right) \quad (1.21)$$

■ **Estimación fuzzy del índice  $C_{pk}$ :**

$$\tilde{C}_{pk}(\alpha) = \min \left\{ \tilde{C}_{pl}(\alpha), \tilde{C}_{pu}(\alpha) \right\} \quad (1.22)$$

donde  $\tilde{C}_{pl}(\alpha)$  y  $\tilde{C}_{pu}(\alpha)$  se definen como en (1.20) y (1.21), respectivamente.

## 1.4. Enfoque de estimación de Buckley

Para estimar los PCIs fuzzy es usual emplear la metodología de estimación propuesta por Buckley [8], que es el enfoque tradicional basado en intervalos de confianza para un parámetro como la familia de  $\alpha$  - *cortes* de un número difuso de forma triangular, el cual se describe a continuación:

Suponga que se desea estimar un parámetro empleando la teoría difusa (específicamente, considerando NFT). Sea  $X$  una variable aleatoria con f.d.p.  $f(x; \theta)$ , con parámetro único  $\theta$ .

Suponga que  $\theta$  debe estimarse a partir de una muestra aleatoria  $X_1, \dots, X_n$ , y sea  $Y = u(X_1, \dots, X_n)$  un estadístico utilizado para estimar  $\theta$ . Teniendo en cuenta los valores de estas variables aleatorias,  $X_i = x_i, 1 \leq i \leq n$ , se obtiene para  $\theta$  una estimación puntual

$$\hat{\theta} = y = u(x_1, \dots, x_n),$$

y además, se puede calcular un intervalo de confianza del  $(1 - \beta)100\%$  ( con  $0 < \beta < 1$ ) para  $\theta$  dado por:

$$[\theta_1(\beta), \theta_2(\beta)].$$

Por lo tanto, el intervalo  $\theta_1 = [\hat{\theta}, \hat{\theta}]$  es el intervalo de confianza del 0% para  $\theta$  y  $\theta_0 = \Theta$  es un intervalo de confianza del 100% para  $\theta$ , donde  $\Theta$  es todo el espacio de parámetros. Entonces tenemos una familia de  $(1 - \beta)100\%$  intervalos de confianza para  $\theta$ , donde  $0 \leq \beta \leq 1$ .

Si se suman estos intervalos de confianza se obtiene un número triangular difuso  $\theta$  cuyos  $\alpha$  - *cortes* son los siguientes intervalos de confianza:

$$\theta_\alpha = [\theta_1(\alpha), \theta_2(\alpha)] \quad \text{para} \quad 0 < \alpha < 1; \quad \theta_0 = \Theta \quad \text{y} \quad \theta_1 = [\hat{\theta}, \hat{\theta}] \quad (1.23)$$

Por lo tanto, se utilizará más información sobre  $\theta$  en lugar de solo una estimación puntual, o solo una estimación de intervalo única. Es fácil generalizar el método de Buckley en el caso en que  $\theta$  es

un vector de parámetros [7].

Por otra parte, para estimar el  $C_{pm}$  bajo un enfoque de Buckley se utiliza el algoritmo de estimación difusa para ICP (EDICP), descrito a continuación:

#### 1.4.1. Algoritmo EDICP

- 1) Sea  $\mathbf{B}_0^i = R^+$ ,  $\mathbf{B}_1^i = [\hat{C}_p^i, \hat{C}_p^i]$  para  $i = 1, 2, 3, 4$ ; donde  $\hat{C}_p^1 = \hat{C}_p$ ,  $\hat{C}_p^2 = \hat{C}_{pk}$ ,  $\hat{C}_p^3 = \hat{C}_{pm(CCS)}$ ,  $\hat{C}_p^4 = \hat{C}_{pm(B)}$  y  $R^+$  es el conjunto de todos los números reales positivos.
- 2) Lugar  $\mathbf{B}_\alpha^i$ ;  $0 \leq \alpha \leq 1$  uno encima del otro, para producir un número fuzzy de forma triangular  $\hat{C}_p^i$  para  $i = 1, 2, 3, 4$ ; donde  $\hat{C}_p^1 = \hat{C}_p$ ,  $\hat{C}_p^2 = \hat{C}_{pk}$ ,  $\hat{C}_p^3 = \hat{C}_{pm(CCS)}$  y  $\hat{C}_p^4 = \hat{C}_{pm(B)}$

Un problema importante que a menudo se pasa por alto es el hecho de que  $\mu$  y  $\sigma$  son generalmente desconocidos y deben estimarse de observaciones. Sustituyendo la media muestral y la desviación estándar proporciona una estimación puntual para cualquiera de los PCI, pero es importante tener en cuenta la incertidumbre debida a la variabilidad del muestreo para acercarse al verdadero valor del PCI.

### 1.5. Estimación fuzzy del índice $C_{pm}$ :

Suponga que los datos asociados al proceso son difusos, así como sus parámetros ( $\mu$  y  $\sigma^2$ ), y además, que los límites de especificación (LSL y USL) y el valor objetivo ( $T$ ) son números triangulares difusos. Kaya y Kahraman[19] proponen estimar el  $C_{pm}$  como:

$$\tilde{C}_{pm}^*(\alpha) = \frac{\tilde{d}^*(\alpha)}{3\sqrt{\tilde{\sigma}^2(\alpha) + [\tilde{\mu}(\alpha) - \tilde{T}(\alpha)]^2}} \quad (1.24)$$

donde:

$$\tilde{\sigma}^2(\alpha) = \left[ \frac{(n-1)s^2}{(1-\alpha)\chi_{R,0.005}^2 + (n-1)\alpha}, \frac{(n-1)s^2}{(1-\alpha)\chi_{L,0.005}^2 + (n-1)\alpha} \right], \quad (1.25)$$

$$\mu_{\tilde{\sigma}}(\alpha) = [\mu_{l,\tilde{\sigma}}(\alpha), \mu_{r,\tilde{\sigma}}(\alpha)] \quad (1.26)$$

siendo  $\mu_{l,\tilde{\sigma}}(\alpha)$  y  $\mu_{r,\tilde{\sigma}}(\alpha)$  la estimación difusa de  $\mu$  definidas en las ecuaciones (1.27) y (1.28):

$$\mu_{l,\tilde{\sigma}}(\alpha) = \min \left( \bar{x} - z_{\alpha/2} \frac{\sigma_i(\alpha)}{\sqrt{n}} \right), \quad i = 1, 2 \quad (1.27)$$

$$\mu_{r,\tilde{\sigma}}(\alpha) = \max \left( \bar{x} + z_{\alpha/2} \frac{\sigma_i(\alpha)}{\sqrt{n}} \right), \quad i = 1, 2 \quad (1.28)$$

Nótese que  $\sigma_1(\alpha)$  y  $\sigma_2(\alpha)$  representan los lados izquierdo y derecho de  $\tilde{\sigma}^2$  respectivamente y se calculan utilizando la ecuación (1.25).

Por otra parte, considerando el supuesto de que los límites de especificación son números triangulares fuzzy (NTF), cuyos  $\alpha$ -cortes están dados por la ecuación (1.15), entonces se obtiene que la estimación difusa del índice  $C_{pm}^*$  que se representa por  $\tilde{C}_{pm}^*$  se define como sigue:

$$\tilde{C}_{pm}^*(\alpha) = \left[ \tilde{C}_{pm}^{*l}(\alpha), \tilde{C}_{pm}^{*r}(\alpha) \right] \quad (1.29)$$

donde los  $\alpha$ -cortes están dados por:

$$\tilde{C}_{pm}^{*l}(\alpha) = \left( \frac{\tilde{d}_L^*}{3\sqrt{\frac{(n-1)s^2}{(1-\alpha)\chi_{L,0.005}^2 + (n-1)\alpha}} + \left[ \max \left( \bar{x} + z_{\alpha/2} \frac{\sigma_i(\alpha)}{\sqrt{n}} \right) - ((t_2 - t_1)\alpha + t_1) \right]^2} \right) \quad (1.30)$$

y

$$\tilde{C}_{pm}^{*r}(\alpha) = \left( \frac{\tilde{d}_R^*}{3\sqrt{\frac{(n-1)s^2}{(1-\alpha)\chi_{R,0.005}^2 + (n-1)\alpha}} + \left[ \min \left( \bar{x} - z_{\alpha/2} \frac{\sigma_i(\alpha)}{\sqrt{n}} \right) - ((t_2 - t_3)\alpha + t_3) \right]^2} \right) \quad (1.31)$$

con

$$\tilde{d}^*(\alpha) = \left( \tilde{d}_L^*, \tilde{d}_R^* \right) = \min \left\{ \tilde{d}_l^*(\alpha), \tilde{d}_r^*(\alpha) \right\}$$

donde,

$$\begin{aligned} \tilde{d}_l^* &= \{(u_2 - u_1 - t_2 + t_3)\alpha + u_1 - t_3, (u_2 - u_3 - t_2 + t_1)\alpha + u_3 - t_1\} \\ \tilde{d}_r^* &= \{(t_2 - t_1 - l_2 + l_3)\alpha + t_1 - l_3, (t_2 - t_3 - l_2 + l_1)\alpha + t_3 - l_1\} \end{aligned}$$

siendo  $\widetilde{USL} = TFN(u_1, u_2, u_3)$ ,  $\widetilde{LSL} = TFN(l_1, l_2, l_3)$ , y  $\widetilde{T} = TFN(t_1, t_2, t_3)$ .

Además,  $\sigma_1(\alpha) = \sqrt{\frac{(n-1)s^2}{(1-\alpha)\chi_{R,0.005}^2 + (n-1)\alpha}}$  y  $\sigma_2(\alpha) = \sqrt{\frac{(n-1)s^2}{(1-\alpha)\chi_{L,0.005}^2 + (n-1)\alpha}}$  son los  $\alpha$ -cortes asociados a la dispersión del proceso.

Nótese que esta definición del  $C_{pm}$  fuzzy sólo considera el método de Buckley para obtener las funciones de membresía de la media fuzzy y la dispersión fuzzy del proceso.





## Capítulo 2

# Estado del arte

Basado en la revisión de la literatura realizada, se tiene que la originalidad de esta propuesta de investigación se fundamenta en que se analiza cuál es el mejor método de construcción de los números triangulares que permita obtener una mejor estimación del  $C_{pm}$  fuzzy en términos de reducción del error estándar, cuando se tengan límites de especificación no concretos. Constituyendo ésto último un aporte a la falencia detectada por [17] en los PCIs formulados en los últimos 20 años.

Gran parte de los estudios sobre la medición del rendimiento de un proceso están basados en estimaciones concretas que implican mediciones precisas del proceso de salida [5, 12]. Sin embargo, en ocasiones las mediciones de calidad del producto no se pueden registrar con precisión, lo que hace que los números difusos sean el único medio factible de describir dichos datos. Dado que las medidas de calidad a menudo carecen de precisión, se ha inspirado una nueva tendencia de combinar la aleatoriedad y la falta de claridad en la evaluación de la capacidad del proceso [13].

En 1996 Yongting [39] propuso por primera vez el atributo difuso de calidad, una fórmula del índice de capacidad de proceso difuso  $C_p$ , que puede definirse como la probabilidad de productos difusos hasta el estándar producidos por procesos de producción, la cuál se propone para cuando se tienen los límites de especificación no concretos. Luego, se realizó la estimación del índice  $C_{pk}$  con números difusos y se aproximó la función de pertenencia del índice  $C_{pk}$  al obtenerse las funciones de pertenencia de la media y la desviación estándar de los números difusos [22]. Así, Chen, Chen y Lin [33] incorporaron la inferencia difusa usando  $C_{pl}$  de índice para procesos con características de calidad de tipo más grandes, y emplearon un concepto de puntuación conciso para representar el grado de capacidad del proceso.

Chan, Cheng et al. [23] y Hsiang y Taguchi [16] de manera independiente teniendo en cuenta que los índices  $C_p$  y  $C_{pk}$  no tienen en cuenta la diferencia entre la media del proceso y su valor objetivo desarrollaron el índice de capacidad general  $C_{pm}$ , el cuál es capaz de considerar en las especificaciones de calidad la variación con respecto al valor objetivo y los límites de especificación. Por otra parte, Pearn, Kotz y Johnson [38] propusieron el ICP para límites de especificación asimétricos  $C_{pmk}$ , que combina las características de los tres índices anteriores y que además alerta al usuario cada vez que la varianza aumenta y/o la media del proceso se desvía de su valor objetivo  $T$ .

Los ICP clásicos propuestos en la literatura por lo general asumen que tanto los límites de especificación como los parámetros asociados a los datos del proceso son números concretos. Pero, bajo

condiciones reales los datos pueden estar asociados a condiciones de incertidumbre [25]. Es por ello, que Parchami y Mashinchi [30] después de la primera incursión del concepto fuzzy realizada por Zadeh [40] introdujeron *PCIs* difusos en diferentes campos y se discutieron las relaciones entre ellos en los casos en que los límites de especificación son números difusos triangulares en lugar de números concretos. Sin embargo, la mayoría de los estudios anteriores sobre la evaluación de la capacidad del proceso con inferencia difusa han pasado por alto los procedimientos difusos de toma de decisiones para evaluar si el proceso que está bajo investigación satisface o no los requisitos predeterminados [13].

Bellman y Zadeh [4] introdujeron en el entorno difuso la aplicación de modelos de prueba de hipótesis. Además, Casals, Gil y Gil [26], Gil, Corral y Gil [24], Casals y Gil [11] consideraron en su momento las pruebas  $\chi^2$  de bondad de ajuste con observaciones difusas y las pruebas secuenciales bayesianas a través de información difusa. También, Casals y Gil [10] consideraron las pruebas de hipótesis tipo Neyman Pearson cuando se tienen datos difusos. En 1993 Watanabe y Imaizumi [37] introdujeron un método donde se utilizan datos aleatorios para probar una hipótesis difusa y definiendo así la función de potencia generalizada como el valor esperado de una función crítica difusa. Arnold [2, 3] también propuso la difuminación de hipótesis estadísticas típicas y consideró la prueba de hipótesis bajo restricciones difusas sobre errores de tipo I y II.

En este orden de ideas, la estimación que se obtiene de los ICP fuzzy no sólo se ve afectada por las funciones de membresía seleccionada, sino que también por la conformación de los números fuzzy que intervienen en la misma [20]. Específicamente, Parchami et al. [1] muestra que, en el mundo real, los diferenciales de números borrosos generalmente se pueden calcular desde diferentes enfoques.

Buckley y Eslami desarrollaron un método para estimar la media y la varianza en estadísticas difusas que utiliza un conjunto de intervalos de confianza  $(1 - \beta)100\%$  para un parámetro como la familia de  $\alpha - cortes$  de un número difuso de forma triangular como estimador [8]. Además, diseñó una regla de decisión para la prueba de hipótesis, que está básicamente determinada por el porcentaje del área de la estadística de prueba difusa que da como resultado un rechazo nítido [7].

No obstante, en la literatura lo más usual es utilizar el método de Buckley o el conocimiento de un experto, asumiendo una dispersión constante [20], o basado en los intervalos de confianza [17, 28], y no es usual examinar el efecto de esta selección sobre el desempeño del ICP. En este trabajo se determina el efecto de la selección del diferencial de los números triangulares sobre el desempeño del índice de capacidad de proceso  $C_{pm}$  fuzzy en términos de calidad de la estimación, cuando se tienen límites de especificación no concretos.

Por tanto, resulta importante analizar el efecto de la construcción de los números triangulares asumiendo que las estimaciones de los parámetros como la dispersión y los límites de especificación son imprecisas para analizar el desempeño del índice de capacidad y así obtener el mejor método para dicha construcción.

## Capítulo 3

# Análisis del diferencial asociado al número triangular en el $C_{pm}$ fuzzy

En este Capítulo se analiza el efecto que tiene el valor del diferencial  $\delta$  usando intervalos de confianza para construir el número triangular asociado a los límites de especificación que interviene en el cálculo del  $C_{pm}$  fuzzy cuando se tienen límites de especificación no concretos, al igual que el valor objetivo y parámetros del proceso que se representan como números triangulares, con el objetivo de ver el efecto de la selección de acuerdo a las formas de escoger  $\delta$ .

Suponga que se tiene un proceso cuyos datos son imprecisos, como por ejemplo en el área industrial el proceso de fabricar pistones para automóviles donde se miden características del diámetro del pistón, el diámetro del pasador, la longitud del pasador y la altura de compresión que inevitablemente tienen cierto grado de imprecisión que puede ser generado al utilizar diferentes máquinas y además, los límites de especificación, el valor objetivo y los parámetros del proceso se pueden representar mediante un número triangular difuso ( $TNF$ ). De modo que para estimar la capacidad del proceso se puede emplear el  $\tilde{C}_{pm}^*$  definido en la ecuación (1.29). Para estudiar analíticamente el efecto del diferencial ( $\delta$ ) de los  $TNF$  sobre la estimación obtenida, considere los siguientes números triangulares simétricos:

$$\begin{cases} \widetilde{LSL} = TFN(l_1, l_2, l_3) = (l_2 - \delta, l_2, l_2 + \delta) \\ \widetilde{T} = TFN(t_1, t_2, t_3) = (t_2 - \delta, t_2, t_2 + \delta) \\ \widetilde{USL} = TFN(u_1, u_2, u_3) = (u_2 - \delta, u_2, u_2 + \delta) \end{cases} \quad (3.1)$$

de las ecuaciones (1.30) y (1.31) se tiene que  $\tilde{d}_r^*$  y  $\tilde{d}_l^*$  se escriben de la siguiente forma:

$$\begin{aligned} \tilde{d}_l^* &= \{(u_2 - u_1 - t_2 + t_3)\alpha + u_1 - t_3, (u_2 - u_3 - t_2 + t_1)\alpha + u_3 - t_1\} \\ \tilde{d}_r^* &= \{(t_2 - t_1 - l_2 + l_3)\alpha + t_1 - l_3, (t_2 - t_3 - l_2 + l_1)\alpha + t_3 - l_1\} \end{aligned}$$

De modo que, remplazando los valores de (3.1) en  $\tilde{d}_r^*$  y  $\tilde{d}_l^*$  en la ecuación (1.29), ésta se puede reescribir de la siguiente forma:

$$\begin{aligned}
\tilde{d}_l^* &= \{(u_2 - (u_2 - \delta) - t_2 + t_2 + \delta)\alpha + u_2 - \delta - (t_2 + \delta), \\
&\quad (u_2 - (u_2 + \delta) - t_2 + t_2 - \delta)\alpha + u_2 + \delta - (t_2 - \delta)\} \\
&= \{(u_2 - u_2 + \delta - t_2 + t_2 + \delta)\alpha + u_2 - \delta - t_2 - \delta, \\
&\quad (u_2 - u_2 - \delta - t_2 + t_2 - \delta)\alpha + u_2 + \delta - t_2 + \delta\} \\
&= \{2\delta\alpha + u_2 - t_2 - 2\delta, -2\delta\alpha + u_2 - t_2 + 2\delta\} \\
&= \{2\delta(\alpha - 1) + u_2 - t_2, 2\delta(-\alpha + 1) + u_2 - t_2\} \\
&= \{u_2 - t_2 - 2\delta(1 - \alpha), u_2 - t_2 + 2\delta(1 - \alpha)\}
\end{aligned}$$

y

$$\begin{aligned}
\tilde{d}_r^* &= \{(t_2 - (t_2 - \delta) - l_2 + l_2 + \delta)\alpha + t_2 - \delta - (l_2 + \delta), \\
&\quad (t_2 - (t_2 + \delta) - l_2 + l_2 - \delta)\alpha + t_2 + \delta - (l_2 - \delta)\} \\
&= \{(t_2 - t_2 + \delta - l_2 + l_2 + \delta)\alpha + t_2 - \delta - l_2 - \delta, \\
&\quad (t_2 - t_2 - \delta - l_2 + l_2 - \delta)\alpha + t_2 + \delta - l_2 + \delta\} \\
&= \{2\delta\alpha + t_2 - l_2 - 2\delta, -2\delta\alpha + t_2 - l_2 + 2\delta\} \\
&= \{2\delta(\alpha - 1) + t_2 - l_2, 2\delta(-\alpha + 1) + t_2 - l_2\} \\
&= \{t_2 - l_2 - 2\delta(1 - \alpha), t_2 - l_2 + 2\delta(1 - \alpha)\}
\end{aligned}$$

Por lo tanto,  $\tilde{d}_L^*$  y  $\tilde{d}_R^*$  se reducen a:

$$\tilde{d}_L^* = \min \{(u_2 - t_2 - 2\delta(1 - \alpha), t_2 - l_2 - 2\delta(1 - \alpha))\} \quad (3.2)$$

$$\tilde{d}_R^* = \min \{u_2 - t_2 + 2\delta(1 - \alpha), t_2 - l_2 + 2\delta(1 - \alpha)\} \quad (3.3)$$

Así, las ecuaciones (1.30) y (1.31) se dejan escribir de la siguiente manera:

$$\tilde{C}_{pm}^{*l}(\alpha) = \left( \frac{\min \{(u_2 - t_2 - 2\delta(1 - \alpha), t_2 - l_2 - 2\delta(1 - \alpha))\}}{3\sqrt{\frac{(n-1)s^2}{(1-\alpha)\chi_{L,0.005}^2 + (n-1)\alpha}} + \left[ \max \left( \bar{x} + z_{\alpha/2} \frac{\sigma_i(\alpha)}{\sqrt{n}} \right) - t_2 + (1-\alpha)\delta \right]^2} \right) \quad (3.4)$$

y

$$\tilde{C}_{pm}^{*r}(\alpha) = \left( \frac{\min \{u_2 - t_2 + 2\delta(1 - \alpha), t_2 - l_2 + 2\delta(1 - \alpha)\}}{3\sqrt{\frac{(n-1)s^2}{(1-\alpha)\chi_{R,0.005}^2 + (n-1)\alpha}} + \left[ \min \left( \bar{x} - z_{\alpha/2} \frac{\sigma_i(\alpha)}{\sqrt{n}} \right) - t_2 - (1-\alpha)\delta \right]^2} \right) \quad (3.5)$$

Por otra parte, para que el proceso sea capaz se debe cumplir que:

$$\tilde{C}_{pm}^*(\alpha) = [\tilde{C}_{pm}^{*l}(\alpha), \tilde{C}_{pm}^{*r}(\alpha)] \geq 1$$

No obstante, es suficiente que  $\tilde{C}_{pm}^{*l}(\alpha) \geq 1$  para que todo el intervalo sea mayor a 1.

$$\frac{\min\{(u_2 - t_2 - 2\delta(1 - \alpha), t_2 - l_2 - 2\delta(1 - \alpha))\}}{3\sqrt{\frac{(n-1)s^2}{(1-\alpha)\chi_{L,0.005}^2 + (n-1)\alpha} + \left[\max\left(\bar{x} + z_{\alpha/2} \frac{\sigma_i(\alpha)}{\sqrt{n}}\right) - t_2 + (1-\alpha)\delta\right]^2}} \geq 1$$

Teniendo en cuenta que:

$$\tilde{\sigma}_L^2 = \frac{(n-1)s^2}{(1-\alpha)\chi_{L,0.005}^2 + (n-1)\alpha} \quad y \quad \mu_{r,\tilde{\sigma}}(\alpha) = \max\left(\bar{x} + z_{\alpha/2} \frac{\sigma_i(\alpha)}{\sqrt{n}}\right)$$

La desigualdad anterior queda de la siguiente forma:

$$\frac{\min\{(u_2 - t_2 - 2\delta(1 - \alpha), t_2 - l_2 - 2\delta(1 - \alpha))\}}{3\sqrt{\tilde{\sigma}_L^2 + [\mu_{r,\tilde{\sigma}}(\alpha) - t_2 + (1-\alpha)\delta]^2}} \geq 1$$

$$\begin{aligned} &\iff \min\{u_2 - t_2 - 2\delta(1 - \alpha), t_2 - l_2 - 2\delta(1 - \alpha)\} \geq 3\sqrt{\tilde{\sigma}_L^2 + [\mu_{r,\tilde{\sigma}}(\alpha) - t_2 + (1-\alpha)\delta]^2} \\ &\iff [\min\{u_2 - t_2 - 2\delta(1 - \alpha), t_2 - l_2 - 2\delta(1 - \alpha)\}]^2 \geq 9\left\{\tilde{\sigma}_L^2 + [\mu_{r,\tilde{\sigma}}(\alpha) - t_2 + (1-\alpha)\delta]^2\right\} \end{aligned}$$

Cómo el valor de  $\tilde{d}_L^* = \min\{u_2 - t_2 - 2\delta(1 - \alpha), t_2 - l_2 - 2\delta(1 - \alpha)\}$  depende del comportamiento de la diferencia de  $u_2 - t_2$  y  $t_2 - l_2$ , puesto que en ambos caso se resta la constante  $2\delta(1 - \alpha)$ . De este modo, es suficiente con estudiar los siguientes dos casos:

**Caso 1:**  $u_2 - t_2 > t_2 - l_2$

$$\begin{aligned} &\implies [t_2 - l_2 - 2\delta(1 - \alpha)]^2 \geq 9\left\{\tilde{\sigma}_L^2 + [\mu_{r,\tilde{\sigma}}(\alpha) - t_2 + (1-\alpha)\delta]^2\right\} \\ &\iff [(t_2 - l_2) - 2\delta(1 - \alpha)]^2 \geq 9\tilde{\sigma}_L^2 + 9[(\mu_{r,\tilde{\sigma}}(\alpha) - t_2) + (1-\alpha)\delta]^2 \end{aligned}$$

donde

$$\begin{aligned} [(t_2 - l_2) - 2\delta(1 - \alpha)]^2 &= (t_2 - l_2)^2 - 2(t_2 - l_2)(2\delta(1 - \alpha)) + (2\delta(1 - \alpha))^2 \\ &= t_2^2 - 2t_2l_2 + l_2^2 - 4\delta(t_2 - l_2)(1 - \alpha) + 4\delta^2(1 - \alpha)^2 \\ &\quad y \\ [(\mu_{r,\tilde{\sigma}}(\alpha) - t_2) + (1 - \alpha)\delta]^2 &= (\mu_{r,\tilde{\sigma}}(\alpha) - t_2)^2 + 2(\mu_{r,\tilde{\sigma}}(\alpha) - t_2)(1 - \alpha)\delta + \delta^2(1 - \alpha)^2 \\ &= (\mu_{r,\tilde{\sigma}}(\alpha))^2 - 2t_2\mu_{r,\tilde{\sigma}}(\alpha) + t_2^2 + 2\delta(\mu_{r,\tilde{\sigma}}(\alpha) - t_2)(1 - \alpha) + \delta^2(1 - \alpha)^2 \end{aligned}$$

Luego

$$\begin{aligned}
& t_2^2 - 2t_2l_2 + l_2^2 - 4\delta(t_2 - l_2)(1 - \alpha) + 4\delta^2(1 - \alpha)^2 \geq \\
& \quad 9\tilde{\sigma}_L^2 + 9 \left[ (\mu_{r,\tilde{\sigma}}(\alpha))^2 - 2t_2\mu_{r,\tilde{\sigma}}(\alpha) + t_2^2 + 2\delta(\mu_{r,\tilde{\sigma}}(\alpha) - t_2)(1 - \alpha) + \delta^2(1 - \alpha)^2 \right] \\
\iff & t_2^2 - 2t_2l_2 + l_2^2 - 4\delta(t_2 - l_2)(1 - \alpha) + 4\delta^2(1 - \alpha)^2 \geq \\
& \quad 9\tilde{\sigma}_L^2 + 9(\mu_{r,\tilde{\sigma}}(\alpha))^2 - 18t_2\mu_{r,\tilde{\sigma}}(\alpha) + 9t_2^2 + 18\delta(\mu_{r,\tilde{\sigma}}(\alpha) - t_2)(1 - \alpha) + 9\delta^2(1 - \alpha)^2 \\
\iff & t_2^2 - 2t_2l_2 + l_2^2 - \left( 9\tilde{\sigma}_L^2 + 9(\mu_{r,\tilde{\sigma}}(\alpha))^2 - 18t_2\mu_{r,\tilde{\sigma}}(\alpha) + 9t_2^2 \right) \geq \\
& \quad 18\delta(\mu_{r,\tilde{\sigma}}(\alpha) - t_2)(1 - \alpha) + 9\delta^2(1 - \alpha)^2 - (-4\delta(t_2 - l_2)(1 - \alpha) + 4\delta^2(1 - \alpha)^2) \\
\iff & t_2^2 - 2t_2l_2 + l_2^2 - 9\tilde{\sigma}_L^2 - 9(\mu_{r,\tilde{\sigma}}(\alpha))^2 + 18t_2\mu_{r,\tilde{\sigma}}(\alpha) - 9t_2^2 \geq \\
& \quad 18\delta(\mu_{r,\tilde{\sigma}}(\alpha) - t_2)(1 - \alpha) + 9\delta^2(1 - \alpha)^2 + 4\delta(t_2 - l_2)(1 - \alpha) - 4\delta^2(1 - \alpha)^2 \\
\iff & -8t_2^2 - 2t_2l_2 + l_2^2 - 9\tilde{\sigma}_L^2 - 9(\mu_{r,\tilde{\sigma}}(\alpha))^2 + 18t_2\mu_{r,\tilde{\sigma}}(\alpha) \geq \\
& \quad 5\delta^2(1 - \alpha)^2 + \delta(1 - \alpha) [18(\mu_{r,\tilde{\sigma}}(\alpha) - t_2) + 4(t_2 - l_2)] \\
\iff & (l_2 - 4t_2)(l_2 + 2t_2) + 9 [2t_2\mu_{r,\tilde{\sigma}}(\alpha) - (\mu_{r,\tilde{\sigma}}(\alpha))^2 - \tilde{\sigma}_L^2] \geq \\
& \quad 5\delta^2(1 - \alpha)^2 + 2\delta(1 - \alpha) [9(\mu_{r,\tilde{\sigma}}(\alpha) - t_2) + 2(t_2 - l_2)] \\
\iff & \kappa_l \geq 5\delta^2(1 - \alpha)^2 + 2\delta(1 - \alpha)\eta_l
\end{aligned}$$

$$\text{con } \kappa_l = (l_2 - 4t_2)(l_2 + 2t_2) + 9 [2t_2\mu_{r,\tilde{\sigma}}(\alpha) - (\mu_{r,\tilde{\sigma}}(\alpha))^2 - \tilde{\sigma}_L^2] \text{ y } \eta_l = 9(\mu_{r,\tilde{\sigma}}(\alpha) - t_2) + 2(t_2 - l_2)$$

De esta forma se obtiene la inecuación

$$5\delta^2(1 - \alpha)^2 + 2\delta(1 - \alpha)\eta_l - \kappa_l \leq 0$$

**Caso 2:**  $t_2 - l_2 > u_2 - t_2$

$$\begin{aligned}
& \implies [u_2 - t_2 - 2\delta(1 - \alpha)]^2 \geq 9 \left\{ \tilde{\sigma}_L^2 + [\mu_{r,\tilde{\sigma}}(\alpha) - t_2 + (1 - \alpha)\delta]^2 \right\} \\
& \iff [(u_2 - t_2) - 2\delta(1 - \alpha)]^2 \geq 9\tilde{\sigma}_L^2 + 9 [(\mu_{r,\tilde{\sigma}}(\alpha) - t_2) + (1 - \alpha)\delta]^2
\end{aligned}$$

donde

$$\begin{aligned}
[(u_2 - t_2) - 2\delta(1 - \alpha)]^2 &= (u_2 - t_2)^2 - 2(u_2 - t_2)(2\delta(1 - \alpha)) + (2\delta(1 - \alpha))^2 \\
&= u_2^2 - 2u_2t_2 + t_2^2 - 4\delta(u_2 - t_2)(1 - \alpha) + 4\delta^2(1 - \alpha)^2
\end{aligned}$$

Luego

$$\begin{aligned}
&\Longleftrightarrow u_2^2 - 2u_2t_2 + t_2^2 - 4\delta(u_2 - t_2)(1 - \alpha) + 4\delta^2(1 - \alpha)^2 \geq \\
&\quad 9\tilde{\sigma}_L^2 + 9 \left[ (\mu_{r,\tilde{\sigma}}(\alpha))^2 - 2t_2\mu_{r,\tilde{\sigma}}(\alpha) + t_2^2 + 2\delta(\mu_{r,\tilde{\sigma}}(\alpha) - t_2)(1 - \alpha) + \delta^2(1 - \alpha)^2 \right] \\
&\Longleftrightarrow u_2^2 - 2u_2t_2 + t_2^2 - 4\delta(u_2 - t_2)(1 - \alpha) + 4\delta^2(1 - \alpha)^2 \geq \\
&\quad 9\tilde{\sigma}_L^2 + 9(\mu_{r,\tilde{\sigma}}(\alpha))^2 - 18t_2\mu_{r,\tilde{\sigma}}(\alpha) + 9t_2^2 + 18\delta(\mu_{r,\tilde{\sigma}}(\alpha) - t_2)(1 - \alpha) + 9\delta^2(1 - \alpha)^2 \\
&\Longleftrightarrow u_2^2 - 2u_2t_2 + t_2^2 - \left( 9\tilde{\sigma}_L^2 + 9(\mu_{r,\tilde{\sigma}}(\alpha))^2 - 18t_2\mu_{r,\tilde{\sigma}}(\alpha) + 9t_2^2 \right) \geq \\
&\quad 18\delta(\mu_{r,\tilde{\sigma}}(\alpha) - t_2)(1 - \alpha) + 9\delta^2(1 - \alpha)^2 - (-4\delta(u_2 - t_2)(1 - \alpha) + 4\delta^2(1 - \alpha)^2) \\
&\Longleftrightarrow u_2^2 - 2u_2t_2 + t_2^2 - 9\tilde{\sigma}_L^2 - 9(\mu_{r,\tilde{\sigma}}(\alpha))^2 + 18t_2\mu_{r,\tilde{\sigma}}(\alpha) - 9t_2^2 \geq \\
&\quad 18\delta(\mu_{r,\tilde{\sigma}}(\alpha) - t_2)(1 - \alpha) + 9\delta^2(1 - \alpha)^2 + 4\delta(u_2 - t_2)(1 - \alpha) - 4\delta^2(1 - \alpha)^2 \\
&\Longleftrightarrow u_2^2 - 2u_2t_2 - 8t_2^2 - 9\tilde{\sigma}_L^2 - 9(\mu_{r,\tilde{\sigma}}(\alpha))^2 + 18t_2\mu_{r,\tilde{\sigma}}(\alpha) \geq \\
&\quad 5\delta^2(1 - \alpha)^2 + \delta(1 - \alpha) [18(\mu_{r,\tilde{\sigma}}(\alpha) - t_2) + 4(u_2 - t_2)] \\
&\Longleftrightarrow (u_2 - 4t_2)(u_2 + 2t_2) + 9 \left[ 2t_2\mu_{r,\tilde{\sigma}}(\alpha) - (\mu_{r,\tilde{\sigma}}(\alpha))^2 - \tilde{\sigma}_L^2 \right] \geq \\
&\quad 5\delta^2(1 - \alpha)^2 + 2\delta(1 - \alpha) [9(\mu_{r,\tilde{\sigma}}(\alpha) - t_2) + 2(u_2 - t_2)] \\
&\Longleftrightarrow \kappa_u \geq 5\delta^2(1 - \alpha)^2 + 2\delta(1 - \alpha)\eta_u
\end{aligned}$$

con  $\kappa_u = (u_2 - 4t_2)(u_2 + 2t_2) + 9 \left[ 2t_2\mu_{r,\tilde{\sigma}}(\alpha) - (\mu_{r,\tilde{\sigma}}(\alpha))^2 - \tilde{\sigma}_L^2 \right]$  y  $\eta_u = 9(\mu_{r,\tilde{\sigma}}(\alpha) - t_2) + 2(u_2 - t_2)$

De esta forma se obtiene la inecuación

$$5\delta^2(1 - \alpha)^2 + 2\delta(1 - \alpha)\eta_u - \kappa_u \leq 0$$

Así,  $\tilde{C}_{pm}^{*l}(\alpha) \geq 1$  si se cumple que:

$$\begin{cases} 5\delta^2(1 - \alpha)^2 + 2\delta(1 - \alpha)\eta_l - \kappa_l \leq 0, & u_2 - t_2 > t_2 - l_2 \\ 5\delta^2(1 - \alpha)^2 + 2\delta(1 - \alpha)\eta_u - \kappa_u \leq 0, & u_2 - t_2 < t_2 - l_2 \end{cases} \quad (3.6)$$

De este modo, es posible determinar si un proceso con datos imprecisos es capaz al estudiar el efecto del valor diferencial  $\delta$  utilizando el enfoque de estimación de Buckley para la construcción de los números triangulares utilizando las desigualdades propuestas en la ecuación (3.6).





## Capítulo 4

# Estudio de simulación y resultados

En este Capítulo, se realizan simulaciones para analizar el desempeño de los estimadores del ICP definido en la Ecuación (1.29), cuando los datos asociados al proceso provienen de distribuciones no normales. Específicamente, se consideran las distribuciones *Gamma* (GA) y *Lognormal* (LN) debido a la variedad de formas asimétricas sesgadas positivamente que se pueden obtener a partir de la combinación de valores para los parámetros de forma y escala ( $\alpha, \beta$ ) de las mismas. La función de densidad asociada a una variable aleatoria  $x$  con distribución Gamma está dada por:

$$f_{GA}(x; \alpha, \beta) = \frac{x^{\alpha-1}}{\beta^{\alpha}\Gamma(\alpha)} \exp(-x/\beta), \quad x > 0, \alpha > 0, \beta > 0$$

donde  $\Gamma(\cdot)$  es la función Gamma dada por:

$$\Gamma(\cdot) = \int_0^{\infty} t^{\alpha-1} \exp(-t) dt.$$

y la función de densidad para la distribución Lognormal está dada por:

$$f_{LN}(x; \alpha, \beta) = \frac{1}{x\alpha\sqrt{2\pi}} \exp\left[-\frac{1}{2}\left(\frac{\ln x - \beta}{\alpha}\right)^2\right], \quad x > 0, \alpha > 0, \beta > 0.$$

Cabe resaltar, que el orden de los parámetros es importante puesto que  $\alpha$  es el parámetro de forma y  $\beta$  el de escala.

La Figura 4.1 muestra la gráfica de una función de densidad gamma con diferentes valores de escala y forma. La Figura 4.2 muestra los gráficos para la función de probabilidad de la distribución *Lognormal*. En ellas se observa que la mayor parte del área bajo la gráfica de la función de densidad, se encuentra cerca del origen y los valores de la función de densidad disminuyen a medida que el valor de  $x$  aumenta.

Para examinar el efecto del tamaño muestral, se consideran en el estudio de simulación los siguientes tamaños de muestra  $n \in \{10, 30, 100\}$ . Con respecto al valor diferencial se realiza las estimaciones difusas  $\mu_{l,\tilde{\sigma}}(\alpha)$  y  $\mu_{r,\tilde{\sigma}}(\alpha)$  de  $\mu$  definidas en las ecuaciones (1.27) y (1.28) y se consideran los siguientes valores fijos dentro de los intervalos estimados  $\delta \in \{0.1000, 0.0100, 0.0010\}$ , y se utiliza el método

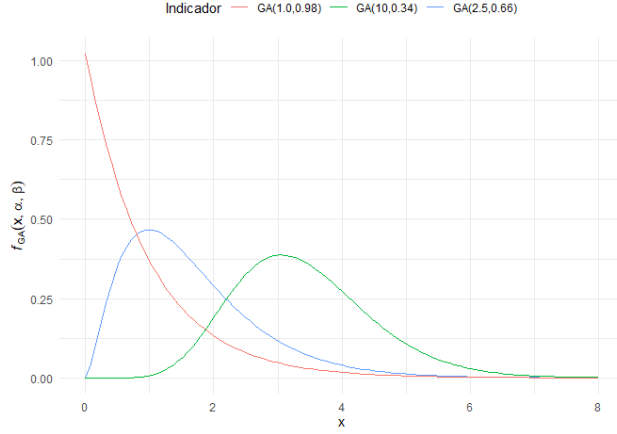


Figura 4.1: Gráficos de la función de probabilidad de la distribución  $\text{Gamma}(\alpha, \beta)$

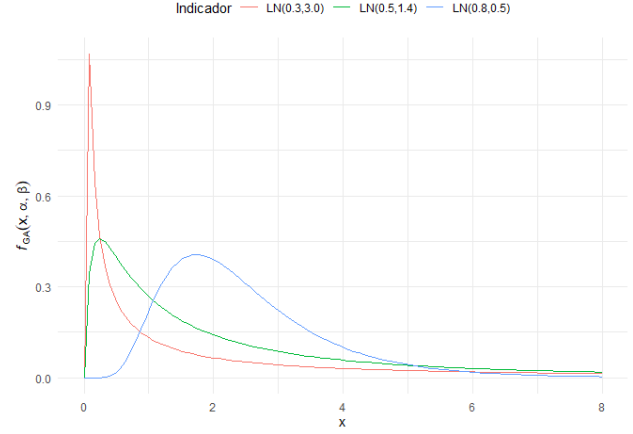


Figura 4.2: Gráficos de la función de probabilidad de la distribución  $\text{lognormal}(\alpha, \beta)$ .

de Buckley [9] quien consideró agregar incertidumbre mediante el uso de estimadores difusos para los parámetros desconocidos  $\mu$  y  $\sigma^2$ . Estos números difusos triangulares con corte  $\alpha$  se construyen a partir de intervalos de confianza y se considera el error estándar  $\delta_{IC} = t_{1-\alpha/2, n-1} \frac{s}{\sqrt{n}}$  donde  $t_{1-\alpha/2, n-1}$  es el percentil superior de la distribución  $t$  con  $n - 1$  grados de libertad.

Estos valores de  $n$  se consideran para cada distribución de densidad de acuerdo con la combinación de valores para los parámetros de forma y escala  $(\alpha, \beta)$  presentados en la Tabla 4.13 4.3 para obtener  $C_{pm} = 1$ , con  $C_{pm}$  dado en la Ecuación (1.24). Estas combinaciones nos permiten evaluar niveles de asimetría bajos, moderados y altos para las distribuciones estudiadas.

Para implementar la metodología planteada, se utiliza el software R mediante algunas librerías ya disponibles y nuevas rutinas además de estos paquetes. Los diferentes escenarios considerados representan diferentes niveles de asimetría. Para la distribución de densidad gamma se consideró  $GA(1.0, 0.98)$ ,  $GA(2.5, 0.66)$  y  $GA(10, 0.34)$ , y para la distribución de densidad lognormal  $LN(0.3, 3.0)$ ,  $LN(0.5, 1.4)$  y  $LN(0.8, 0.5)$ .

#### 4.1. Resultados para una distribución $\text{Gamma}(\alpha, \beta)$

La Tabla 4.1 muestra los resultados obtenidos al realizar las simulaciones cuando se consideran los distintos tamaños de muestra y valores de  $\delta$  y los diferentes  $\delta_{IC_i}$  estimados para cada tamaño de muestra, como márgenes de error para una  $\text{gamma}(10, 0.34)$  de asimetría baja considerando el efecto de la construcción de los números triangulares  $USL$ ,  $LSL$  y  $T$  para calcular el  $C_{pm}$  fuzzy.

Se debe tener en cuenta que, para cada tamaño de muestra se calcula un error estándar  $\delta_{IC_i}$  con  $i = 1, 2, 3$  para  $t_i$  el vector de percentiles superiores de la distribución  $t$  con  $n - 1$  grados de libertad y  $n = 10, 30, 100$  para construir los números triangulares  $USL$ ,  $LSL$  y  $T$  en cada caso.

En los escenarios planteados para la distribución de asimetría baja en la Tabla 4.1 se muestran los diferentes resultados obtenidos, para los distintos cambios del valor de error ( $\delta$ ) se observa que al

Tabla 4.1: Efecto de la construcción de los números triangulares en el  $C_{pm}$  fuzzy con una distribución  $Gamma(10.0, 0.34)$  con asimetría baja.

| Margen de error | $\widetilde{LSL}$     | $\widetilde{USL}$     | $\widetilde{T}$       | $n = 10$  | $\widetilde{C}_{pmi}$<br>$n = 30$ | $n = 100$ |
|-----------------|-----------------------|-----------------------|-----------------------|-----------|-----------------------------------|-----------|
| $\delta_{IC}$   | $1.6 - \delta_{IC_i}$ | $8.1 - \delta_{IC_i}$ | $4.8 - \delta_{IC_i}$ | 0.1431074 | 0.3221917                         | 0.4483450 |
|                 | 1.6000                | 8.1000                | 4.8000                | 0.6042686 | 0.6042686                         | 0.6042686 |
|                 | $1.6 + \delta_{IC_i}$ | $8.1 + \delta_{IC_i}$ | $4.8 + \delta_{IC_i}$ | 2.3670705 | 1.1071369                         | 0.7999407 |
| 0.1000          | 1.5000                | 8.0000                | 4.7000                | 0.2264281 | 0.3678522                         | 0.4649511 |
|                 | 1.6000                | 8.1000                | 4.8000                | 0.6042686 | 0.6042686                         | 0.6042686 |
|                 | 1.7000                | 8.2000                | 4.9000                | 1.6833005 | 1.1330760                         | 0.7968653 |
| 0.0100          | 1.5900                | 8.0900                | 4.7900                | 0.2359693 | 0.3796543                         | 0.4768538 |
|                 | 1.6000                | 8.1000                | 4.8000                | 0.6042686 | 0.6042686                         | 0.6042686 |
|                 | 1.6100                | 8.1100                | 4.8100                | 1.5926499 | 1.1309771                         | 0.7926196 |
| 0.0010          | 1.5990                | 8.0990                | 4.7990                | 0.2369044 | 0.3807933                         | 0.4779865 |
|                 | 1.6000                | 8.1000                | 4.8000                | 0.6042686 | 0.6042686                         | 0.6042686 |
|                 | 1.6010                | 8.1010                | 4.8010                | 1.5837354 | 1.1303740                         | 0.7920886 |

aumentar este valor es más ancho el intervalo del  $\widetilde{C}_{pm}$  y por ende es más capaz el proceso, además, al aumentar el tamaño de muestra se reduce el ancho para cada  $\delta$  disminuyendo así la distancia entre los números triangulares construidos en el  $C_{pm}$  como puede observarse en las Figuras 4.3 y 4.4. Además, al utilizar el error de estimación del intervalo de confianza se presenta un comportamiento similar aunque con un mayor tamaño la distancia entre el valor inferior y el superior.

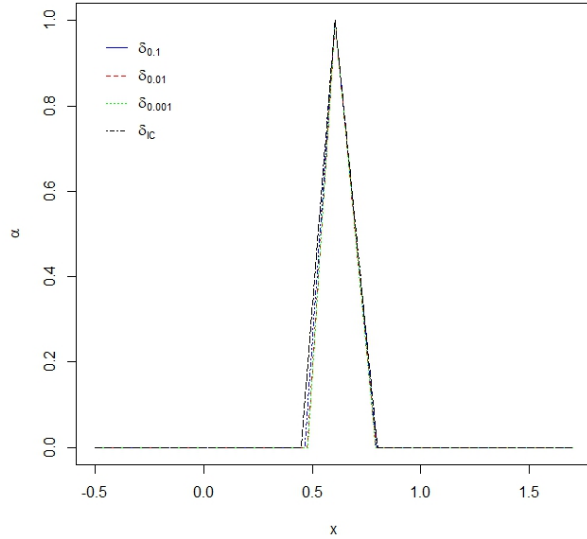


Figura 4.3: Números triangulares en el  $C_{pm}$  fuzzy con distribución  $Gamma(10.0, 0.34)$ , margen de error  $\delta$ ,  $\delta_{IC_i}$  y  $n = 100$ .

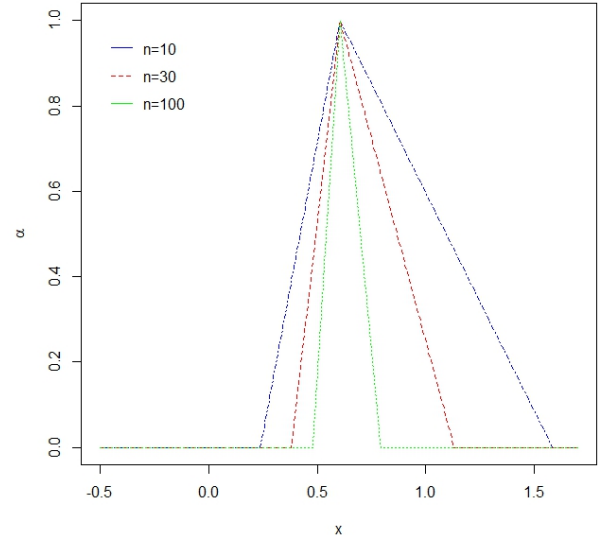


Figura 4.4: Números triangulares en el  $C_{pm}$  fuzzy con distribución  $Gamma(10.0, 0.34)$ , margen de error  $\delta = 0.0010$  y  $n = (10, 30, 100)$ .

Los resultados obtenidos de las simulaciones que se muestran en la Tabla 4.2 consideran los distintos tamaños de muestra y valores de  $\delta$  como margen de error para una  $gamma(2.5, 0.66)$  de asimetría

moderada al considerar el efecto de la construcción de los números triangulares  $USL$ ,  $LSL$  y  $T$  para calcular el  $C_{pm}$  fuzzy.

Tabla 4.2: Efecto de la construcción de los números triangulares en el  $C_{pm}$  fuzzy con una distribución  $Gamma(2.5, 0.66)$  con asimetría moderada.

| Margen de error | $\widetilde{LSL}$     | $\widetilde{USL}$     | $\widetilde{T}$       | $n = 10$  | $\widetilde{C}_{pmi}$<br>$n = 30$ | $n = 100$ |
|-----------------|-----------------------|-----------------------|-----------------------|-----------|-----------------------------------|-----------|
| $\delta_{IC}$   | $1.6 - \delta_{IC_i}$ | $8.1 - \delta_{IC_i}$ | $4.8 - \delta_{IC_i}$ | 0.1079938 | 0.2043212                         | 0.2601262 |
|                 | 1.6000                | 8.1000                | 4.8000                | 0.3214441 | 0.3214441                         | 0.3214441 |
|                 | $1.6 + \delta_{IC_i}$ | $8.1 + \delta_{IC_i}$ | $4.8 + \delta_{IC_i}$ | 0.9148352 | 0.4776688                         | 0.3895752 |
| 0.1000          | 1.5000                | 8.0000                | 4.7000                | 0.1706180 | 0.2370479                         | 0.2724421 |
|                 | 1.6000                | 8.1000                | 4.8000                | 0.3214441 | 0.3214441                         | 0.3214441 |
|                 | 1.7000                | 8.2000                | 4.9000                | 0.9959986 | 0.4533507                         | 0.3790737 |
| 0.0100          | 1.5900                | 8.0900                | 4.7900                | 0.1783300 | 0.2463549                         | 0.2822657 |
|                 | 1.6000                | 8.1000                | 4.8000                | 0.3214441 | 0.3214441                         | 0.3214441 |
|                 | 1.6100                | 8.1100                | 4.8100                | 1.0078336 | 0.4445489                         | 0.3696411 |
| 0.0010          | 1.5990                | 8.0990                | 4.7990                | 0.1790891 | 0.2472649                         | 0.2832228 |
|                 | 1.6000                | 8.1000                | 4.8000                | 0.3214441 | 0.3214441                         | 0.3214441 |
|                 | 1.6010                | 8.1010                | 4.8010                | 1.0089692 | 0.4436310                         | 0.3686654 |

Se observa en las Figuras 4.5 y 4.6 que a medida que el tamaño de muestra y los errores de estimación para la construcción de los números triangulares son pequeños el índice  $\widetilde{C}_{pm}$  calculado devuelve un proceso capaz con respecto al límite superior difuso, en cambio, si se aumenta el tamaño de muestra y de error se reduce la distancia entre los números triangulares construidos en el  $C_{pm}$  y por ende el proceso resulta no capaz.

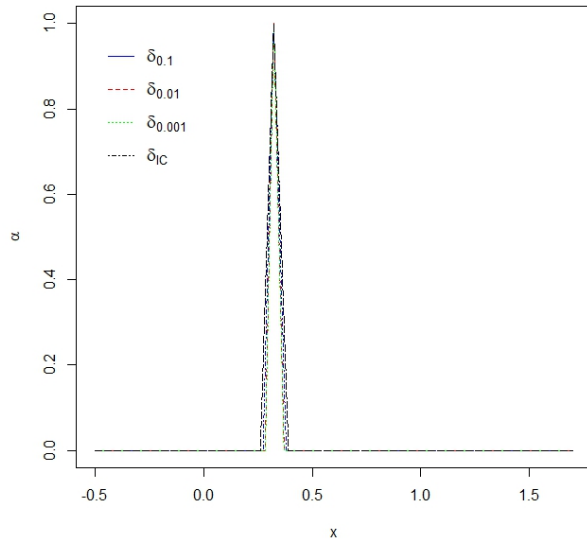


Figura 4.5: Números triangulares en el  $C_{pm}$  fuzzy con distribución  $Gamma(2.5, 0.66)$ , margen de error  $\delta$ ,  $\delta_{IC_i}$  y  $n = 100$ .

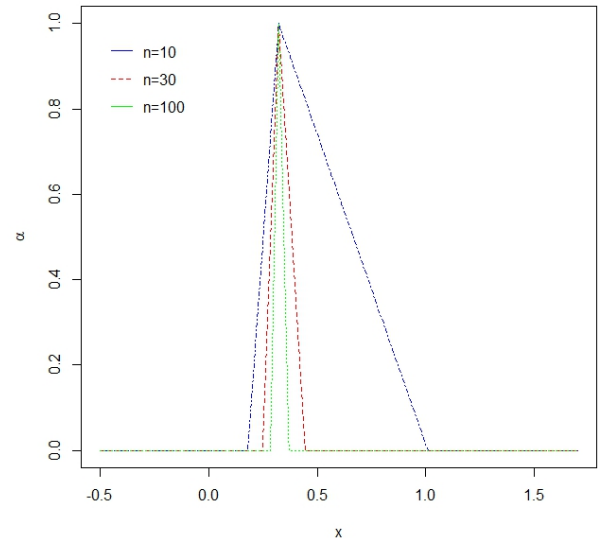


Figura 4.6: Números triangulares en el  $C_{pm}$  fuzzy con distribución  $Gamma(2.5, 0.66)$ , margen de error  $\delta = 0.0010$  y  $n = (10, 30, 100)$ .

En la Tabla 4.3 se evalúa una  $gamma(1.0, 0.98)$  de asimetría alta y se muestran los resultados de realizar las diferentes simulaciones considerando los distintos escenarios de tamaño de muestra y valores de para el error  $\delta$  y error estándar a a partir del intervalo de confianza para  $\delta_{IC_i}$ . Además, se asume el efecto de la construcción de los números triangulares  $USL$ ,  $LSL$  y  $T$  para calcular el índice  $C_{pm}$  fuzzy y verificar su comportamiento mediante las Figuras 4.7 y 4.8.

Tabla 4.3: Efecto de la construcción de los números triangulares en el  $C_{pm}$  fuzzy con una distribución  $Gamma(1.0, 0.98)$  con asimetría alta.

| Margen de error | $\widetilde{LSL}$     | $\widetilde{USL}$     | $\widetilde{T}$       | $n = 10$  | $\widetilde{C}_{pmi}$<br>$n = 30$ | $n = 100$ |
|-----------------|-----------------------|-----------------------|-----------------------|-----------|-----------------------------------|-----------|
| $\delta_{IC}$   | $1.6 - \delta_{IC_i}$ | $8.1 - \delta_{IC_i}$ | $4.8 - \delta_{IC_i}$ | 0.1045727 | 0.1822651                         | 0.2247662 |
|                 | 1.6000                | 8.1000                | 4.8000                | 0.2704733 | 0.2704733                         | 0.2704733 |
|                 | $1.6 + \delta_{IC_i}$ | $8.1 + \delta_{IC_i}$ | $4.8 + \delta_{IC_i}$ | 0.6322523 | 0.3805646                         | 0.3200067 |
| 0.1000          | 1.5000                | 8.0000                | 4.7000                | 0.1588213 | 0.2097376                         | 0.2347474 |
|                 | 1.6000                | 8.1000                | 4.8000                | 0.2704733 | 0.2704733                         | 0.2704733 |
|                 | 1.7000                | 8.2000                | 4.9000                | 0.6125186 | 0.3559338                         | 0.3108329 |
| 0.0100          | 1.5900                | 8.0900                | 4.7900                | 0.1661113 | 0.2183743                         | 0.2438632 |
|                 | 1.6000                | 8.1000                | 4.8000                | 0.2704733 | 0.2704733                         | 0.2704733 |
|                 | 1.6100                | 8.1100                | 4.8100                | 0.6079613 | 0.3466373                         | 0.3016409 |
| 0.0010          | 1.5990                | 8.0990                | 4.7990                | 0.1668295 | 0.2192209                         | 0.2447545 |
|                 | 1.6000                | 8.1000                | 4.8000                | 0.2704733 | 0.2704733                         | 0.2704733 |
|                 | 1.6010                | 8.1010                | 4.8010                | 0.6074717 | 0.3456778                         | 0.3006963 |

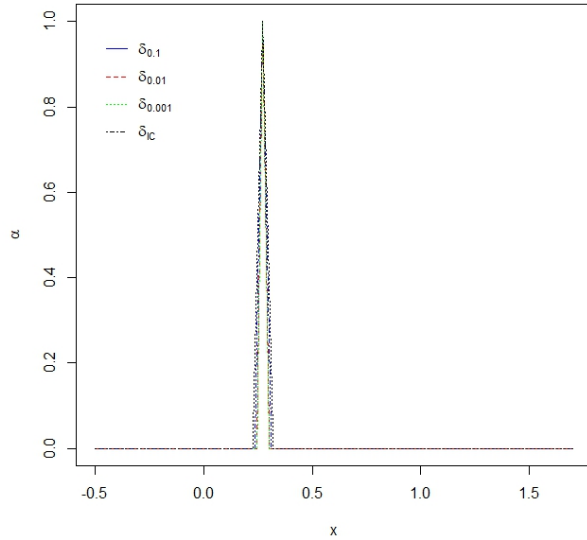


Figura 4.7: Números triangulares en el  $C_{pm}$  fuzzy con distribución  $Gamma(1.0, 0.98)$ , margen de error  $\delta$ ,  $\delta_{IC_i}$  y  $n = 100$ .

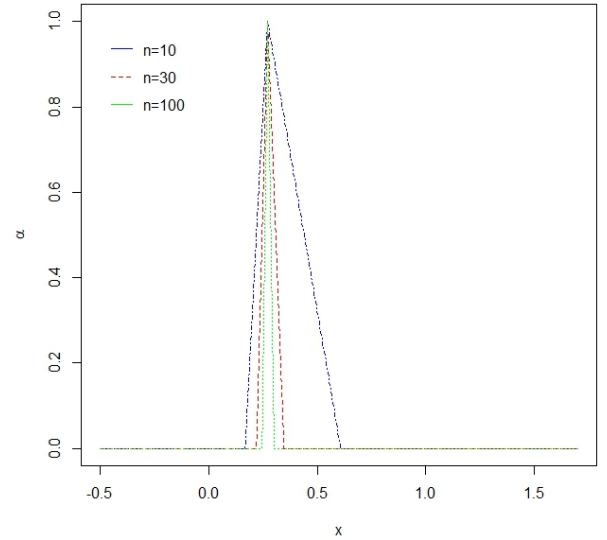


Figura 4.8: Números triangulares en el  $C_{pm}$  fuzzy con distribución  $Gamma(1.0, 0.98)$ , margen de error  $\delta = 0.0010$  y  $n = (10, 30, 100)$ .

En la Tabla 4.3 se observa que para una distribución  $gamma$  de asimetría alta y distintos cambios del valor  $\delta$  y al aumentar el tamaño de muestra el proceso no se considera capaz para ninguno de los

casos obteniendo un  $\tilde{C}_{pm} < 1$  para el límite superior. De este modo, es importante considerar otros métodos para la construcción del índice de capacidad  $C_{pm}$ , principalmente cuando la distribución de las observaciones son asimétricas o no pueden considerarse como una población Normal.

Ahora, en la Tabla 4.4 se calcula el  $C_{pm}$  fuzzy utilizando la mediana y simulando muestras de diferentes tamaños bajo la distribución *gamma* con los parámetros establecidos anteriormente. Además, en las Figuras 4.9 y 4.10 puede detallarse dicho comportamiento.

Tabla 4.4: Efecto de la construcción de los números triangulares  $\mu$  y  $\sigma^2$  en el  $C_{pm}$  fuzzy con una distribución *Gamma*(10.0, 0.34) con asimetría baja utilizando la mediana.

| Margen de error | $\widetilde{LSL}$     | $\widetilde{USL}$     | $\tilde{T}$           | $n = 10$  | $\tilde{C}_{pmi}$<br>$n = 30$ | $n = 100$ |
|-----------------|-----------------------|-----------------------|-----------------------|-----------|-------------------------------|-----------|
| $\delta_{IC}$   | $1.6 - \delta_{IC_i}$ | $8.1 - \delta_{IC_i}$ | $4.8 - \delta_{IC_i}$ | 0.2660248 | 0.5516522                     | 0.5518586 |
|                 | 1.6000                | 8.1000                | 4.8000                | 0.7488589 | 0.7927290                     | 0.6469197 |
|                 | $1.6 + \delta_{IC_i}$ | $8.1 + \delta_{IC_i}$ | $4.8 + \delta_{IC_i}$ | 0.8413053 | 0.8027131                     | 0.8997582 |
| 0.1000          | 1.5000                | 8.0000                | 4.7000                | 0.3667049 | 0.5854644                     | 0.5644743 |
|                 | 1.6000                | 8.1000                | 4.8000                | 0.7488589 | 0.7927290                     | 0.6469197 |
|                 | 1.7000                | 8.2000                | 4.9000                | 0.8069957 | 0.7963576                     | 0.8561940 |
| 0.0100          | 1.5900                | 8.0900                | 4.7900                | 0.3766190 | 0.5928117                     | 0.5731330 |
|                 | 1.6000                | 8.1000                | 4.8000                | 0.7488589 | 0.7927290                     | 0.6469197 |
|                 | 1.6100                | 8.1100                | 4.8100                | 0.8048869 | 0.7963576                     | 0.8164348 |
| 0.0010          | 1.5990                | 8.0990                | 4.7990                | 0.3775755 | 0.5934934                     | 0.5739404 |
|                 | 1.6000                | 8.1000                | 4.8000                | 0.7488589 | 0.7927290                     | 0.6469197 |
|                 | 1.6010                | 8.1010                | 4.8010                | 0.8048869 | 0.7963576                     | 0.8122213 |

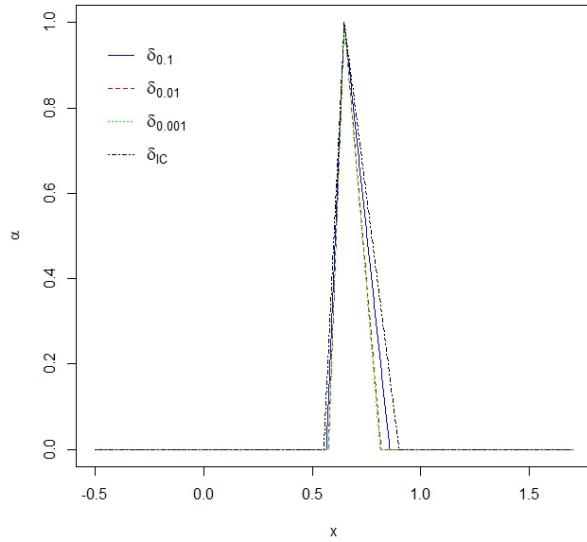


Figura 4.9: Números triangulares  $\mu$  y  $\sigma^2$  en el  $C_{pm}$  fuzzy con una distribución *Gamma*(10.0, 0.34) utilizando la mediana y  $n = 100$ .

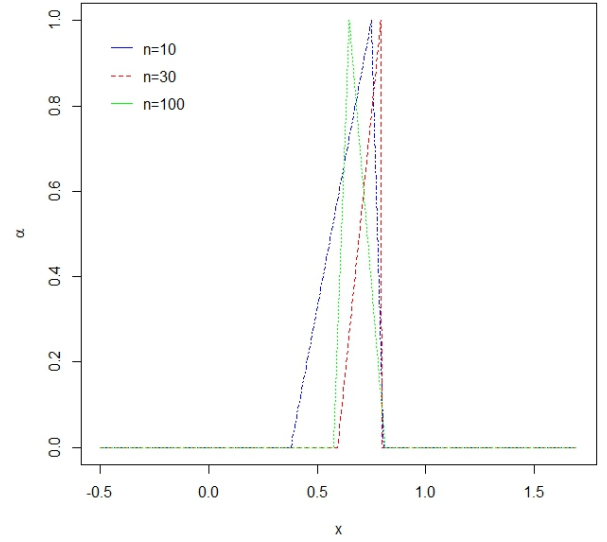


Figura 4.10: Números triangulares  $\mu$  y  $\sigma^2$  en el  $C_{pm}$  fuzzy con una *Gamma*(10.0, 0.34) utilizando la mediana,  $\delta = 0.0010$  y  $n = (10, 30, 100)$ .

Cabe resaltar que en la Tabla 4.4 se está considerando el efecto de la construcción de los números triangulares  $\mu$  y  $\sigma^2$  para una distribución  $gamma(10, 0.34)$  para calcular el  $C_{pm}$  fuzzy en los diferentes escenarios planteados de tamaños de muestra y valores de error  $\delta_{IC_i}$  y  $\delta$ , se encontró que, al aumentar este valor es más ancho el intervalo del  $\tilde{C}_{pm}$  y por ende es más capaz el proceso.

En la Tabla 4.5 se pueden observar los resultados de las simulaciones cuando se consideran los distintos tamaños de muestra y diferentes márgenes de error para una una distribución  $gamma(2.5, 0.66)$ .

Tabla 4.5: Efecto de la construcción de los números triangulares  $\mu$  y  $\sigma^2$  en el  $C_{pm}$  fuzzy con una distribución  $Gamma(2.5, 0.66)$  con asimetría moderada.

| Margen de error | $\widetilde{LSL}$     | $\widetilde{USL}$     | $\tilde{T}$           | $n = 10$  | $\tilde{C}_{pmi}$<br>$n = 30$ | $n = 100$ |
|-----------------|-----------------------|-----------------------|-----------------------|-----------|-------------------------------|-----------|
| $\delta_{IC}$   | $1.6 - \delta_{IC_i}$ | $8.1 - \delta_{IC_i}$ | $4.8 - \delta_{IC_i}$ | 0.1538431 | 0.2535664                     | 0.2798158 |
|                 | 1.6000                | 8.1000                | 4.8000                | 0.3620960 | 0.3670399                     | 0.3278962 |
|                 | $1.6 + \delta_{IC_i}$ | $8.1 + \delta_{IC_i}$ | $4.8 + \delta_{IC_i}$ | 0.5510619 | 0.6744635                     | 0.4632743 |
| 0.1000          | 1.5000                | 8.0000                | 4.7000                | 0.2303488 | 0.2886161                     | 0.2921955 |
|                 | 1.6000                | 8.1000                | 4.8000                | 0.3620960 | 0.3670399                     | 0.3278962 |
|                 | 1.7000                | 8.2000                | 4.9000                | 0.4463212 | 0.6285193                     | 0.4519564 |
| 0.0100          | 1.5900                | 8.0900                | 4.7900                | 0.2392448 | 0.2983548                     | 0.3020246 |
|                 | 1.6000                | 8.1000                | 4.8000                | 0.3620960 | 0.3670399                     | 0.3278962 |
|                 | 1.6100                | 8.1100                | 4.8100                | 0.4279804 | 0.6101898                     | 0.4416084 |
| 0.0010          | 1.5990                | 8.0990                | 4.7990                | 0.2401147 | 0.2993018                     | 0.3029801 |
|                 | 1.6000                | 8.1000                | 4.8000                | 0.3620960 | 0.3670399                     | 0.3278962 |
|                 | 1.6010                | 8.1010                | 4.8010                | 0.4260947 | 0.6082393                     | 0.4405285 |

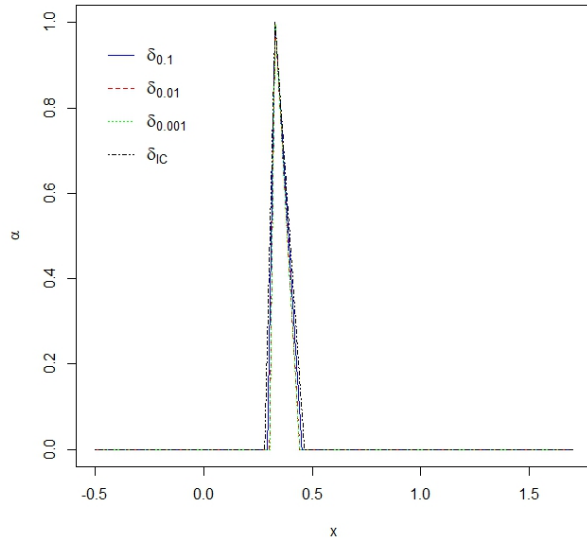


Figura 4.11: Números triangulares  $\mu$  y  $\sigma^2$  en el  $C_{pm}$  fuzzy con una  $Gamma(2.5, 0.66)$  utilizando la mediana y  $n = 100$ .

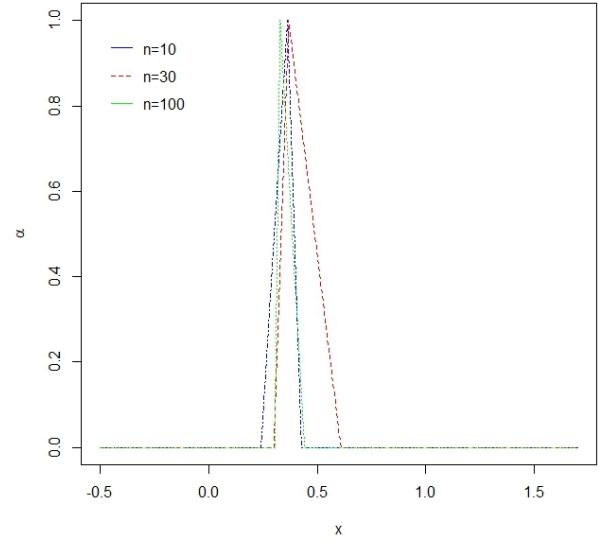


Figura 4.12: Números triangulares  $\mu$  y  $\sigma^2$  en el  $C_{pm}$  fuzzy con una  $Gamma(2.5, 0.66)$  utilizando la mediana,  $\delta = 0.0010$  y  $n = (10, 30, 100)$ .

Teniendo en cuenta los resultados de la Tabla 4.5 para una distribución  $gamma(2.5, 0.66)$  de asimetría moderada considerando el efecto de la construcción de los números triangulares  $\mu$  y  $\sigma^2$  para calcular el  $C_{pm}$  fuzzy se puede notar que al igual que los escenarios de las distribuciones anteriores se tiene que con valores de error pequeños sin importar el tamaño de muestra se reduce la capacidad del proceso denotada por el  $\tilde{C}_{pm}$  calculado. Aunque, en las Figuras 4.11 y 4.12 se observa que para muestras grandes y sin importar el tamaño de error, el resultado del índice difuso es similar en cada caso. En cambio, al momento de fijar un error y al analizar el tamaño de muestra el índice estimado  $C_{pm}$  arroja una mayor distancia entre los límites inferior y superior a medida que la muestra disminuye.

La Tabla 4.6 muestra los resultados obtenidos al realizar las diferentes simulaciones para una distribución  $gamma(1.0, 0.98)$  de asimetría alta y considerando distintos tamaños de muestra y valores de error  $\delta$ , además de asumir el efecto de la construcción de los números triangulares  $\mu$  y  $\sigma^2$  para calcular el  $C_{pm}$  fuzzy.

Tabla 4.6: Efecto de la construcción de los números triangulares  $\mu$  y  $\sigma^2$  en el  $C_{pm}$  fuzzy con una distribución  $Gamma(1.0, 0.98)$  con asimetría alta.

| Margen de error | $\widetilde{LSL}$     | $\widetilde{USL}$     | $\tilde{T}$           | $\tilde{C}_{pmi}$ |           |           |
|-----------------|-----------------------|-----------------------|-----------------------|-------------------|-----------|-----------|
|                 |                       |                       |                       | $n = 10$          | $n = 30$  | $n = 100$ |
| $\delta_{IC}$   | $1.6 - \delta_{IC_i}$ | $8.1 - \delta_{IC_i}$ | $4.8 - \delta_{IC_i}$ | 0.1498251         | 0.2036873 | 0.2275943 |
|                 | 1.6000                | 8.1000                | 4.8000                | 0.2753461         | 0.2666565 | 0.2650948 |
|                 | $1.6 + \delta_{IC_i}$ | $8.1 + \delta_{IC_i}$ | $4.8 + \delta_{IC_i}$ | 0.6061425         | 0.7716129 | 0.4221837 |
| 0.1000          | 1.5000                | 8.0000                | 4.7000                | 0.2174160         | 0.2324388 | 0.2375254 |
|                 | 1.6000                | 8.1000                | 4.8000                | 0.2753461         | 0.2666565 | 0.2650948 |
|                 | 1.7000                | 8.2000                | 4.9000                | 0.4918073         | 0.7180212 | 0.4123819 |
| 0.0100          | 1.5900                | 8.0900                | 4.7900                | 0.2260568         | 0.2413879 | 0.2465868 |
|                 | 1.6000                | 8.1000                | 4.8000                | 0.2753461         | 0.2666565 | 0.2650948 |
|                 | 1.6100                | 8.1100                | 4.8100                | 0.4703137         | 0.6942552 | 0.4023724 |
| 0.0010          | 1.5990                | 8.0990                | 4.7990                | 0.2269028         | 0.2422628 | 0.2474724 |
|                 | 1.6000                | 8.1000                | 4.8000                | 0.2753461         | 0.2666565 | 0.2650948 |
|                 | 1.6010                | 8.1010                | 4.8010                | 0.4681040         | 0.6917110 | 0.4013331 |

En la Tabla 4.6 donde se analiza el efecto de los números triangulares  $\mu$  y  $\sigma^2$  se logra observar que aplicando los diferentes cambios en los valores de  $\delta$  y al aumentar el tamaño de muestra el proceso no se considera capaz para ninguno de los casos al obtenerse un  $\tilde{C}_{pm} < 1$ . Cabe resaltar que, a medida que el tamaño de muestra tiende a ser menor se sigue manteniendo la tendencia de tener como resultado un mayor ancho de intervalo que a su vez aumenta al tenerse un mayor valor de  $\delta$ , comportamiento que puede verificarse al analizar las Figuras 4.13 y 4.14.

De este modo, al observar como influye la selección del valor de  $\delta$  y la selección de los tamaños de muestra cuando se asumen difusos los parámetros  $\mu$  y  $\sigma^2$  con respecto a la construcción del índice de capacidad se debe considerar importante el utilizar otros métodos para la construcción del  $C_{pm}$ , principalmente cuando la distribución de las observaciones son asimétricas o no pueden considerarse como una población Normal al momento de construir el índice de capacidad del proceso que resulte de interés.



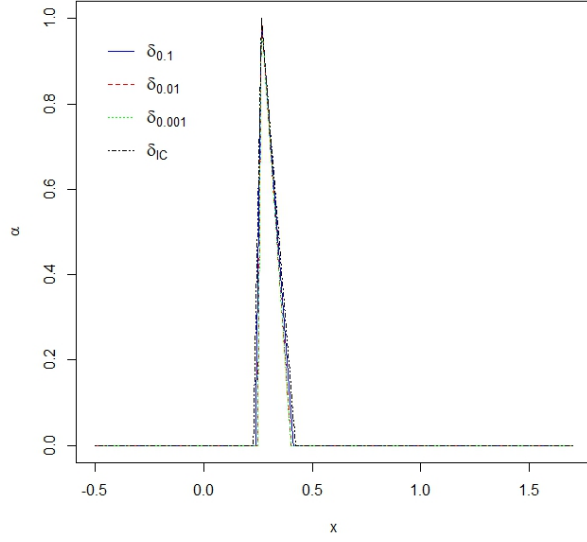


Figura 4.13: Números triangulares  $\mu$  y  $\sigma^2$  en el  $C_{pm}$  fuzzy con una  $Gamma(1.0, 0.98)$  utilizando la mediana y  $n = 100$ .

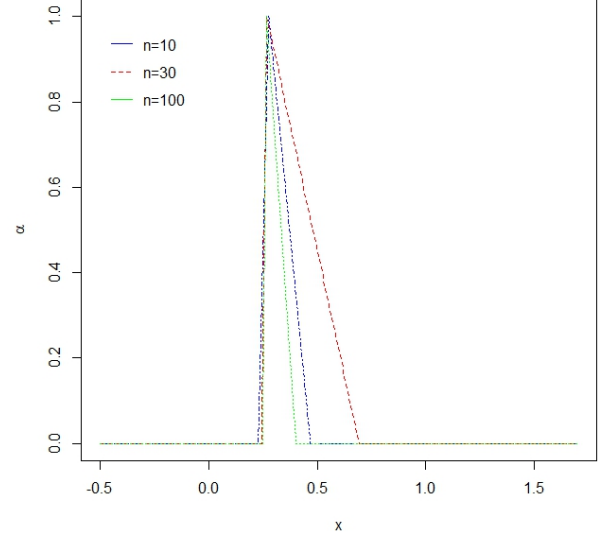


Figura 4.14: Números triangulares  $\mu$  y  $\sigma^2$  en el  $C_{pm}$  fuzzy con una  $Gamma(1.0, 0.98)$  utilizando la mediana,  $\delta = 0.0010$  y  $n = (10, 30, 100)$ .

Además, se tiene que entre mayor asimetría se tenga en la distribución de los datos a medida que aumenta el tamaño de muestra en ambos métodos para calcular el índice de capacidad, se obtiene un menor ancho de intervalo y por ende más bajo es el valor del índice  $C_{pm}$  obtenido, aunque la mayor diferencia se tiene con los resultados de la distribución  $Gamma$  cuando se mira el efecto de los números triangulares en el  $C_{pm}$  fuzzy.

## 4.2. Resultados para una distribución $Lognormal(\alpha, \beta)$

A continuación se muestran los diferentes resultados para una distribución  $Lognormal$  de asimetría baja, moderada y alta, se debe tener en cuenta que, para cada tamaño de muestra se calcula un error estándar  $\delta_{IC_i}$  con  $i = 1, 2, 3$  para  $t_i$  el vector de percentiles superiores de la distribución  $t$  con  $n - 1$  grados de libertad y  $n = 10, 30, 100$  para construir los números triangulares  $USL$ ,  $LSL$  y  $T$  en cada caso.

En la Tabla 4.7 se muestran los resultados obtenidos para una distribución  $Lognormal(0.8, 0.5^2)$  con asimetría baja al realizar las diferentes simulaciones cuando se consideran los distintos tamaños de muestra y valores de  $\delta$  y  $\delta_{IC}$  de acuerdo a los parámetros  $\alpha$  y  $\beta$  como margen de error y considerando el efecto de la construcción de los números triangulares  $USL$ ,  $LSL$  y  $T$  para calcular el índice  $C_{pm}$  fuzzy.

En los escenarios planteados para la distribución  $Lognormal$  de asimetría baja se observa en los resultados de la Tabla 4.7 que los valores obtenidos de los diferentes escenarios teniendo en cuenta

Tabla 4.7: Efecto de la construcción de los números triangulares en el  $C_{pm}$  fuzzy con una distribución  $Lognormal(0.8, 0.5^2)$  con asimetría baja.

| Margen de error | $\widetilde{LSL}$     | $\widetilde{USL}$     | $\widetilde{T}$       | $n = 10$  | $\widetilde{C}_{pmi}$<br>$n = 30$ | $n = 100$ |
|-----------------|-----------------------|-----------------------|-----------------------|-----------|-----------------------------------|-----------|
| $\delta_{IC}$   | $1.6 - \delta_{IC_i}$ | $8.1 - \delta_{IC_i}$ | $4.8 - \delta_{IC_i}$ | 0.1696569 | 0.2193222                         | 0.2420222 |
|                 | 1.6000                | 8.1000                | 4.8000                | 0.2646074 | 0.2646074                         | 0.2646074 |
|                 | $1.6 + \delta_{IC_i}$ | $8.1 + \delta_{IC_i}$ | $4.8 + \delta_{IC_i}$ | 0.3997530 | 0.3145592                         | 0.2879673 |
| 0.1000          | 1.5000                | 8.0000                | 4.7000                | 0.1947679 | 0.2282188                         | 0.2419408 |
|                 | 1.6000                | 8.1000                | 4.8000                | 0.2646074 | 0.2646074                         | 0.2646074 |
|                 | 1.7000                | 8.2000                | 4.9000                | 0.3767650 | 0.3062878                         | 0.2880434 |
| 0.0100          | 1.5900                | 8.0900                | 4.7900                | 0.2029650 | 0.2370921                         | 0.2510302 |
|                 | 1.6000                | 8.1000                | 4.8000                | 0.2646074 | 0.2646074                         | 0.2646074 |
|                 | 1.6100                | 8.1100                | 4.8100                | 0.3677794 | 0.2972827                         | 0.2791586 |
| 0.0010          | 1.5990                | 8.0990                | 4.7990                | 0.2037695 | 0.2379600                         | 0.2519180 |
|                 | 1.6000                | 8.1000                | 4.8000                | 0.2646074 | 0.2646074                         | 0.2646074 |
|                 | 1.6010                | 8.1010                | 4.8010                | 0.3668502 | 0.2963575                         | 0.2782473 |

los distintos cambios del valor de error ( $\delta$ ) para los valores del error estándar y los valores fijados y se observa que al aumentar este valor es más ancho el intervalo del  $\widetilde{C}_{pm}$  y por ende es más capaz el proceso, además, al aumentar el tamaño de muestra se reduce el ancho para cada  $\delta$  disminuyendo así la distancia entre los números triangulares construidos en el  $C_{pm}$ , comportamiento que se puede observar en las Figuras 4.15 y 4.16.

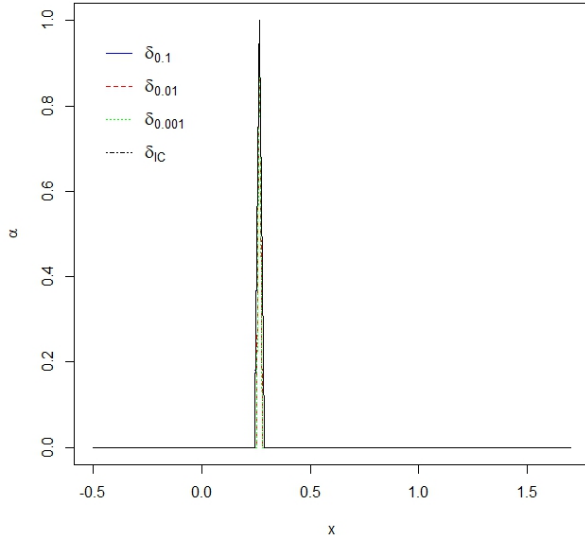


Figura 4.15: Números triangulares en el  $C_{pm}$  fuzzy con distribución  $Lognormal(0.8, 0.5^2)$ , margen de error  $\delta$ ,  $\delta_{IC_i}$  y  $n = 100$ .

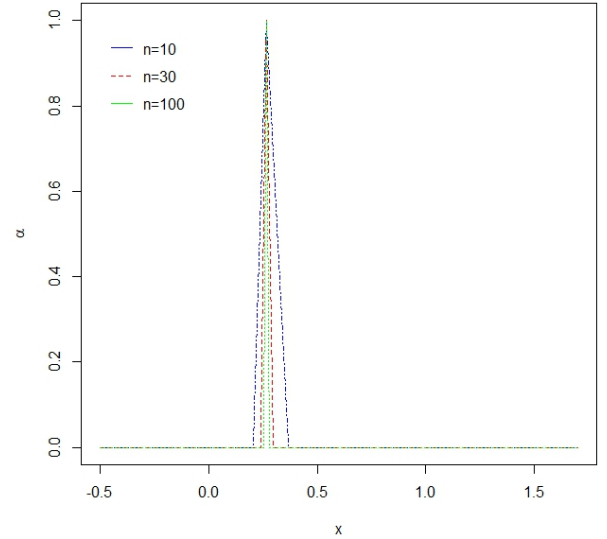


Figura 4.16: Números triangulares en el  $C_{pm}$  fuzzy con distribución  $Lognormal(0.8, 0.5^2)$ , margen de error  $\delta = 0.0010$  y  $n = (10, 30, 100)$ .

Cabe resaltar que para la distribución  $Lognormal$  se tienen números triangulares más cercanos entre sí, es decir, se acercan al valor real con una mayor precisión comparados con una distribución

*Gamma* de igual asimetría y utilizando la construcción de los números triangulares para el índice fuzzy.

En la Tabla 4.8 se muestran los resultados de una distribución  $Lognormal(0.5, 1.4^2)$  de asimetría moderada obtenidos al realizar las simulaciones cuando se consideran los distintos tamaños de muestra y valores de  $\delta$  y  $\delta_{IC_i}$  como márgenes de error considerando el efecto de la construcción de los números triangulares  $USL$ ,  $LSL$  y  $T$  para calcular el  $C_{pm}$  fuzzy.

Tabla 4.8: Efecto de la construcción de los números triangulares en el  $C_{pm}$  fuzzy con una distribución  $Lognormal(0.5, 1.4^2)$  con asimetría moderada.

| Margen de error | $\widetilde{LSL}$     | $\widetilde{USL}$     | $\widetilde{T}$       | $n = 10$  | $\widetilde{C}_{pmi}$<br>$n = 30$ | $n = 100$ |
|-----------------|-----------------------|-----------------------|-----------------------|-----------|-----------------------------------|-----------|
| $\delta_{IC}$   | $1.6 - \delta_{IC_i}$ | $8.1 - \delta_{IC_i}$ | $4.8 - \delta_{IC_i}$ | 0.0560144 | 0.1338736                         | 0.1814508 |
|                 | 1.6000                | 8.1000                | 4.8000                | 0.2358751 | 0.2358751                         | 0.2358751 |
|                 | $1.6 + \delta_{IC_i}$ | $8.1 + \delta_{IC_i}$ | $4.8 + \delta_{IC_i}$ | 0.7339087 | 0.3727111                         | 0.2966522 |
| 0.1000          | 1.5000                | 8.0000                | 4.7000                | 0.1255958 | 0.1736050                         | 0.1989731 |
|                 | 1.6000                | 8.1000                | 4.8000                | 0.2358751 | 0.2358751                         | 0.2358751 |
|                 | 1.7000                | 8.2000                | 4.9000                | 0.7266621 | 0.3341385                         | 0.2798932 |
| 0.0100          | 1.5900                | 8.0900                | 4.7900                | 0.1317513 | 0.1813592                         | 0.2073965 |
|                 | 1.6000                | 8.1000                | 4.8000                | 0.2358751 | 0.2358751                         | 0.2358751 |
|                 | 1.6100                | 8.1100                | 4.8100                | 0.7224968 | 0.3246530                         | 0.2708444 |
| 0.0010          | 1.5990                | 8.0990                | 4.7990                | 0.1323597 | 0.1821220                         | 0.2082232 |
|                 | 1.6000                | 8.1000                | 4.8000                | 0.2358751 | 0.2358751                         | 0.2358751 |
|                 | 1.6010                | 8.1010                | 4.8010                | 0.7220100 | 0.3236762                         | 0.2699173 |

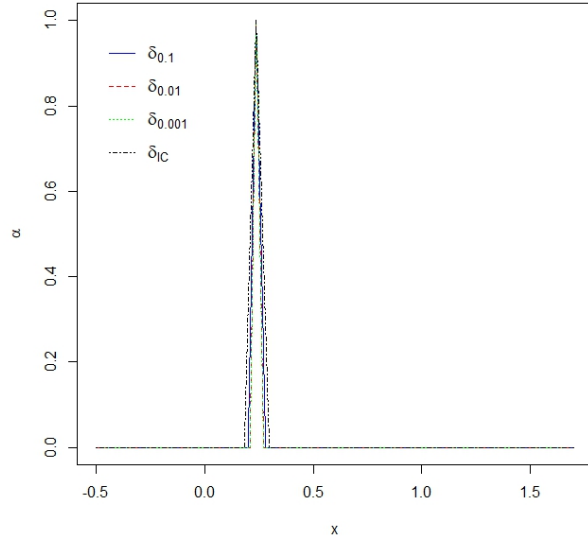


Figura 4.17: Números triangulares en el  $C_{pm}$  fuzzy con distribución  $Lognormal(0.5, 1.4^2)$ , margen de error  $\delta$ ,  $\delta_{IC_i}$  y  $n = 100$ .

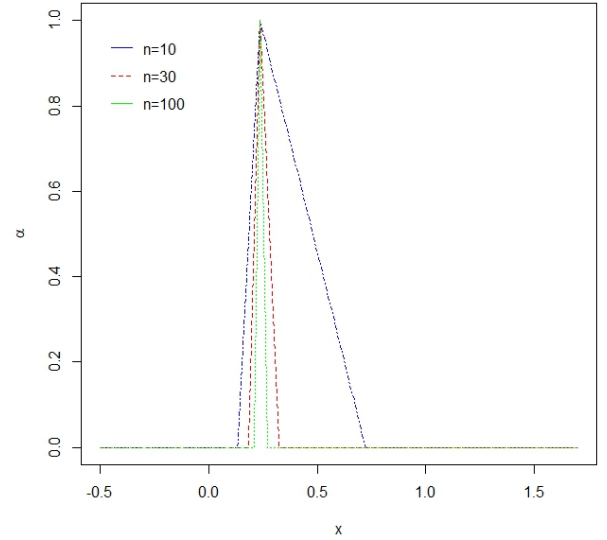


Figura 4.18: Números triangulares en el  $C_{pm}$  fuzzy con distribución  $Lognormal(0.5, 1.4^2)$ , margen de error  $\delta = 0.0010$  y  $n = (10, 30, 100)$ .

Como resultado se obtuvo que, entre más grande es el tamaño del  $\delta$  y más pequeñas son las muestras el  $\tilde{C}_{pm}$  calculado devuelve intervalos más grandes entre los números triangulares como puede verse en la Figura 4.17. En cambio, el comportamiento de la Figura 4.18 indica que al aumentarse el tamaño de muestra y disminuir el error se reduce la distancia entre los números triangulares construidos en el  $C_{pm}$ , cabe resaltar que el proceso no es capaz.

En la Tabla 4.9 se observan los diferentes resultados obtenidos para una  $Lognormal(0.3, 3.0^2)$  de asimetría alta al realizar las simulaciones con los distintos tamaños de muestra y valores para el error  $\delta$  y  $\delta_{IC_i}$  asumiendo el efecto de la construcción de los números triangulares  $USL$ ,  $LSL$  y  $T$  para calcular el  $C_{pm}$  fuzzy, se muestra que para distintos cambios del valor  $\delta$  y al aumentar el tamaño de muestra el proceso no se considera capaz para ninguno de los casos obteniendo un  $\tilde{C}_{pm} < 1$ .

Tabla 4.9: Efecto de la construcción de los números triangulares en el  $C_{pm}$  fuzzy con una dostribución  $Lognormal(0.3, 3.0^2)$  con asimetría alta.

| Margen de error | $\widetilde{LSL}$     | $\widetilde{USL}$     | $\tilde{T}$           | $\tilde{C}_{pmi}$ |           |           |
|-----------------|-----------------------|-----------------------|-----------------------|-------------------|-----------|-----------|
|                 |                       |                       |                       | $n = 10$          | $n = 30$  | $n = 100$ |
| $\delta_{IC}$   | $1.6 - \delta_{IC_i}$ | $8.1 - \delta_{IC_i}$ | $4.8 - \delta_{IC_i}$ | 0.0000000         | 0.0434925 | 0.1077351 |
|                 | 1.6000                | 8.1000                | 4.8000                | 0.1972267         | 0.1972267 | 0.1972267 |
|                 | $1.6 + \delta_{IC_i}$ | $8.1 + \delta_{IC_i}$ | $4.8 + \delta_{IC_i}$ | 1.3453321         | 0.4736711 | 0.3109630 |
| 0.1000          | 1.5000                | 8.0000                | 4.7000                | 0.0770787         | 0.1216128 | 0.1508278 |
|                 | 1.6000                | 8.1000                | 4.8000                | 0.1972267         | 0.1972267 | 0.1972267 |
|                 | 1.7000                | 8.2000                | 4.9000                | 0.6059716         | 0.3706527 | 0.2638070 |
| 0.0100          | 1.5900                | 8.0900                | 4.7900                | 0.0812144         | 0.1277192 | 0.1580680 |
|                 | 1.6000                | 8.1000                | 4.8000                | 0.1972267         | 0.1972267 | 0.1972267 |
|                 | 1.6100                | 8.1100                | 4.8100                | 0.5741905         | 0.3581657 | 0.2541395 |
| 0.0010          | 1.5990                | 8.0990                | 4.7990                | 0.0816252         | 0.1283235 | 0.1587826 |
|                 | 1.6000                | 8.1000                | 4.8000                | 0.1972267         | 0.1972267 | 0.1972267 |
|                 | 1.6010                | 8.1010                | 4.8010                | 0.5709770         | 0.3568795 | 0.2531523 |

En las Figuras 4.19 y 4.20 obtenidas al estudiar datos aleatorios provenientes de una distribución  $Lognormal(0.3, 3.0^2)$  considerada de asimetría alta, se logra observar que para tamaños de muestras cada vez más grandes la construcción de los números triangulares arroja resultados cercanos sin importar los cambios de los errores  $\delta$  y  $\delta_{IC_i}$  considerados para la estimación del índice. Además, al utilizar el error estándar con  $n = 10$  los límites de especificación son 0.0000 y 1.3453 inferior y superior respectivamente para el índice de capacidad  $C_{pm}$ , lo que indica que al utilizar el error de estimación de la media a un 95 % de confianza para conjuntos pequeños y con datos de asimetría alta se debe tener en cuenta que el ancho de intervalo estimado crece y el límite inferior del proceso podría empezar a tomar valores negativos.

Cabe resaltar que para tamaños de muestra grandes y diferentes  $\delta$  el comportamiento del resultado del  $\tilde{C}_{pm}$  fuzzy es similar para cada caso, en cambio, cuando se varía el  $n$  a mayor tamaño se reduce el ancho del intervalo como se puede ver en las Figuras 4.19 y 4.20. En este orden de ideas, se considera importante utilizar otros métodos para la construcción del índice de capacidad  $C_{pm}$ , principalmente cuando la distribución de las observaciones son asimétricas o se considera la población con una distribución diferente a la Normal y por ende se estudiará a continuación el comportamiento de las distribuciones anteriores cuando se construyen los números triangulares  $\mu$  y  $\sigma^2$  en el  $C_{pm}$  fuzzy.

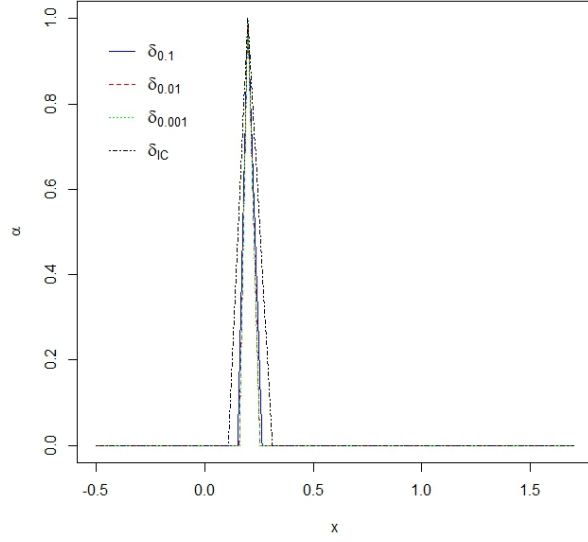


Figura 4.19: Números triangulares en el  $C_{pm}$  fuzzy con distribución  $Lognormal(0.3, 3.0^2)$ , margen de error  $\delta$ ,  $\delta_{IC_i}$  y  $n = 100$ .

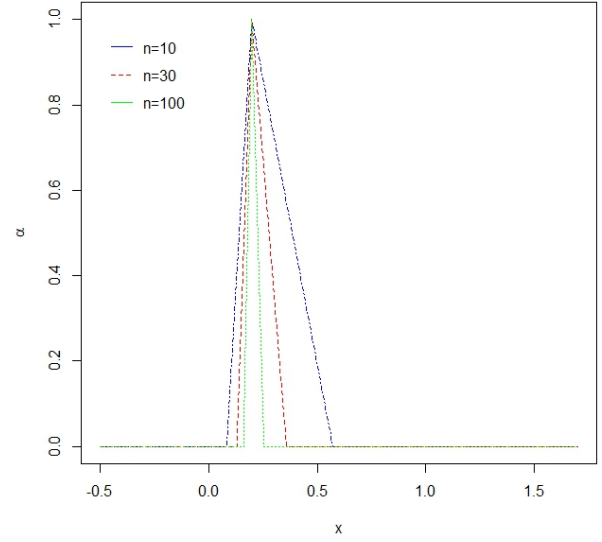


Figura 4.20: Números triangulares en el  $C_{pm}$  fuzzy con distribución  $Lognormal(0.3, 3.0^2)$ , margen de error  $\delta = 0.0010$  y  $n = (10, 30, 100)$ .

Para iniciar lo planteado anteriormente se muestra en la Tabla 4.10 los primeros resultados obtenidos para una distribución  $Lognormal(0.8, 0.5^2)$  de asimetría baja al realizar las diferentes simulaciones considerando distintos tamaños de muestra y valores del margen de error considerando el efecto de la construcción de los números triangulares  $\mu$  y  $\sigma^2$  para calcular el  $C_{pm}$  fuzzy. Se debe tener en cuenta que para el error estándar  $\delta_{IC_i}$  el valor  $t$  depende de los grados de libertad para cada caso y por ende este valor varía con respecto a cada tamaño de muestra.

Tabla 4.10: Efecto de la construcción de los números triangulares  $\mu$  y  $\sigma^2$  en el  $C_{pm}$  fuzzy con una distribución  $Lognormal(0.8, 0.5^2)$  con asimetría baja utilizando la mediana.

| Margen de error | $\widetilde{LSL}$     | $\widetilde{USL}$     | $\widetilde{T}$       | $\widetilde{C}_{pmi}$ |           |           |
|-----------------|-----------------------|-----------------------|-----------------------|-----------------------|-----------|-----------|
|                 |                       |                       |                       | $n = 10$              | $n = 30$  | $n = 100$ |
| $\delta_{IC}$   | $1.6 - \delta_{IC_i}$ | $8.1 - \delta_{IC_i}$ | $4.8 - \delta_{IC_i}$ | 0.2279442             | 0.3307540 | 0.3656350 |
|                 | 1.6000                | 8.1000                | 4.8000                | 0.5271715             | 0.4692705 | 0.4281024 |
|                 | $1.6 + \delta_{IC_i}$ | $8.1 + \delta_{IC_i}$ | $4.8 + \delta_{IC_i}$ | 1.3233181             | 1.1228206 | 1.1539781 |
| 0.1000          | 1.5000                | 8.0000                | 4.7000                | 0.2569850             | 0.3407378 | 0.3655471 |
|                 | 1.6000                | 8.1000                | 4.8000                | 0.5271715             | 0.4692705 | 0.4281024 |
|                 | 1.7000                | 8.2000                | 4.9000                | 1.2388146             | 1.1324518 | 1.1537872 |
| 0.0100          | 1.5900                | 8.0900                | 4.7900                | 0.2662454             | 0.3504970 | 0.3752566 |
|                 | 1.6000                | 8.1000                | 4.8000                | 0.5271715             | 0.4692705 | 0.4281024 |
|                 | 1.6100                | 8.1100                | 4.8100                | 1.2293888             | 1.1384413 | 1.1750412 |
| 0.0010          | 1.5990                | 8.0990                | 4.7990                | 0.2671485             | 0.3514411 | 0.3761937 |
|                 | 1.6000                | 8.1000                | 4.8000                | 0.5271715             | 0.4692705 | 0.4281024 |
|                 | 1.6010                | 8.1010                | 4.8010                | 1.2299726             | 1.1387443 | 1.1770793 |

Para los escenarios planteados en la distribución de asimetría baja en la Tabla 4.10 se observa que al aumentar este valor es más ancho el intervalo del  $\tilde{C}_{pm}$ . Además para todos los diferentes escenarios el intervalo superior del  $C_{pm}$  fuzzy fue mayor a uno lo que implica que el ancho de intervalo del índice es mayor cuando se considera difuso los parámetros  $\mu$  y  $\sigma^2$  para los datos generados de una distribución  $Lognormal(0.8, 0.5^2)$  de asimetría baja como puede observarse en las Figuras 4.21 y 4.22. Luego, en la Tabla 4.11 se pueden observar los resultados de las simulaciones para una distribución  $gamma(2.5, 0.66)$  de asimetría moderada.

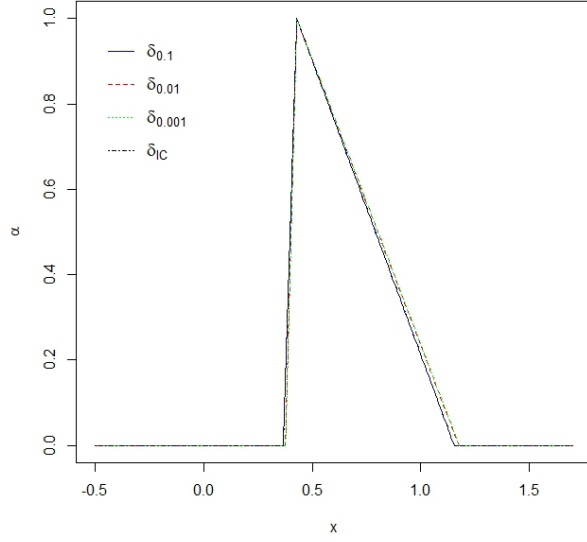


Figura 4.21: Números triangulares  $\mu$  y  $\sigma^2$  en el  $C_{pm}$  fuzzy con una  $Lognormal(0.8, 0.5^2)$  utilizando la mediana y  $n = 100$ .

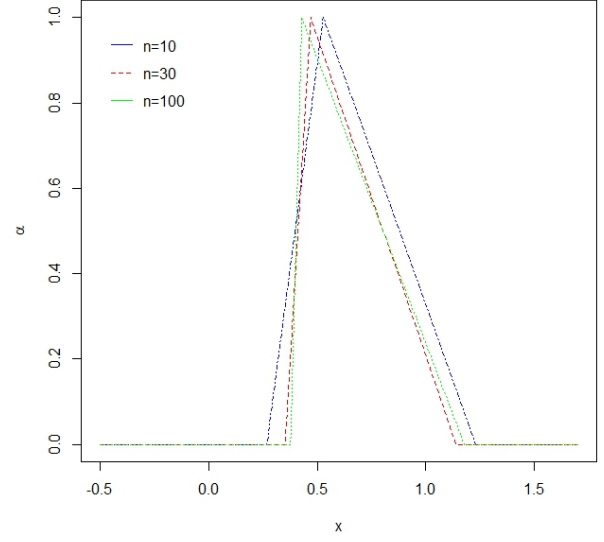


Figura 4.22: Números triangulares  $\mu$  y  $\sigma^2$  en el  $C_{pm}$  fuzzy con una  $Lognormal(0.8, 0.5^2)$  utilizando la mediana,  $\delta = 0.0010$  y  $n = (10, 30, 100)$ .

Tabla 4.11: Efecto de la construcción de los números triangulares  $\mu$  y  $\sigma^2$  en el  $C_{pm}$  fuzzy con una distribución  $Lognormal(0.5, 1.4^2)$  con asimetría moderada.

| Margen de error | $\widetilde{LSL}$     | $\widetilde{USL}$     | $\tilde{T}$           | $\tilde{C}_{pmi}$ |           |           |
|-----------------|-----------------------|-----------------------|-----------------------|-------------------|-----------|-----------|
|                 |                       |                       |                       | $n = 10$          | $n = 30$  | $n = 100$ |
| $\delta_{IC}$   | $1.6 - \delta_{IC_i}$ | $8.1 - \delta_{IC_i}$ | $4.8 - \delta_{IC_i}$ | 0.1107879         | 0.2092385 | 0.2656859 |
|                 | 1.6000                | 8.1000                | 4.8000                | 0.5947001         | 0.5897136 | 0.3756257 |
|                 | $1.6 + \delta_{IC_i}$ | $8.1 + \delta_{IC_i}$ | $4.8 + \delta_{IC_i}$ | 0.7807510         | 0.7415685 | 0.5412133 |
| 0.1000          | 1.5000                | 8.0000                | 4.7000                | 0.2226357         | 0.2604702 | 0.2869413 |
|                 | 1.6000                | 8.1000                | 4.8000                | 0.5947001         | 0.5897136 | 0.3756257 |
|                 | 1.7000                | 8.2000                | 4.9000                | 0.7476696         | 0.6678349 | 0.5158854 |
| 0.0100          | 1.5900                | 8.0900                | 4.7900                | 0.2314231         | 0.2699815 | 0.2969119 |
|                 | 1.6000                | 8.1000                | 4.8000                | 0.5947001         | 0.5897136 | 0.3756257 |
|                 | 1.6100                | 8.1100                | 4.8100                | 0.7476696         | 0.6588602 | 0.5012630 |
| 0.0010          | 1.5990                | 8.0990                | 4.7990                | 0.2322831         | 0.2709090 | 0.2978819 |
|                 | 1.6000                | 8.1000                | 4.8000                | 0.5947001         | 0.5897136 | 0.3756257 |
|                 | 1.6010                | 8.1010                | 4.8010                | 0.7476696         | 0.6579450 | 0.4997290 |

En la Tabla 4.11 se puede observar los resultados de las simulaciones para una distribución  $gamma(2.5, 0.66)$  de asimetría moderada cuando se consideran los distintos tamaños de muestra y diferentes márgenes de error, además de considerar el efecto de la construcción de los números triangulares  $\mu$  y  $\sigma^2$  para calcular el  $C_{pm}$  fuzzy.

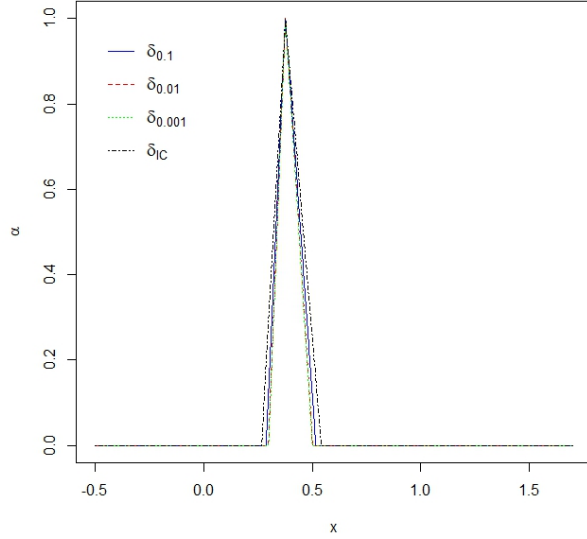


Figura 4.23: Números triangulares  $\mu$  y  $\sigma^2$  en el  $C_{pm}$  fuzzy con una  $Lognormal(0.5, 1.4^2)$  utilizando la mediana y  $n = 100$ .

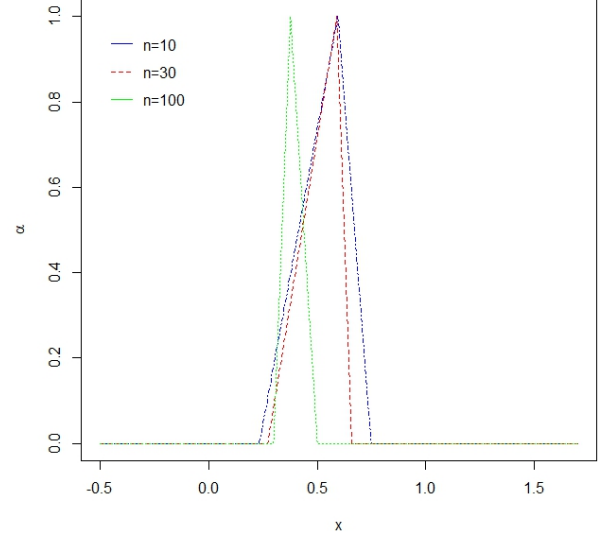


Figura 4.24: Números triangulares  $\mu$  y  $\sigma^2$  en el  $C_{pm}$  fuzzy con una  $Lognormal(0.5, 1.4^2)$  utilizando la mediana,  $\delta = 0.0010$  y  $n = (10, 30, 100)$ .

Se puede resaltar en la Tabla 4.11 y en las Figuras 4.23 y 4.24 que en los diferentes escenarios que se plantean para la distribución  $Lognormal(0.5, 1.4^2)$  de asimetría moderada el índice  $\tilde{C}_{pm}$  calculado devuelve un proceso no capaz, pero, se observa la tendencia de aumento del ancho de intervalo a medida que el tamaño de muestra disminuye y el error aumenta.

Por último, en la Tabla 4.12 se muestran los resultados cuando se asume el efecto de la construcción de los números triangulares  $\mu$  y  $\sigma^2$  para calcular el  $C_{pm}$  fuzzy con simulaciones para una  $Lognormal(0.3, 3.0^2)$  de asimetría alta en los distintos escenarios de tamaño de muestra y valores de error para la estimación fuzzy.

Así, en la Tabla 4.12 mostrada a continuación y donde se consideran distintos tamaños de muestra y valores de para el error  $\delta$  y  $\delta_{IC_i}$  se tiene que los resultados obtenidos de los diferentes escenarios planteados para una distribución  $gamma(\alpha, \beta)$  de asimetría alta el proceso no se considera capaz para ninguno de los casos obteniendo un  $\tilde{C}_{pm} < 1$  a medida que se aumenta el tamaño de muestra. Además, en comparación con los resultados obtenidos para la distribución con asimetría baja y moderada difieren cuando esta asimetría es alta, dado que, se reduce drásticamente el ancho de intervalo para cada  $\tilde{C}_{pm}$  fuzzy calculado en cada escenario sin importar los diferentes niveles de error  $\delta$  y tamaños de muestra que se tomen para calcular el índice de capacidad como puede notarse en las Figuras 4.25 y 4.26.

Tabla 4.12: Efecto de la construcción de los números triangulares  $\mu$  y  $\sigma^2$  en el  $C_{pm}$  fuzzy con una distribución  $Lognormal(0.3, 3.0^2)$  con asimetría alta.

| Margen de error | $\widetilde{LSL}$     | $\widetilde{USL}$     | $\widetilde{T}$       | $n = 10$  | $\widetilde{C}_{pmi}$<br>$n = 30$ | $n = 100$ |
|-----------------|-----------------------|-----------------------|-----------------------|-----------|-----------------------------------|-----------|
| $\delta_{IC}$   | $1.6 - \delta_{IC_i}$ | $8.1 - \delta_{IC_i}$ | $4.8 - \delta_{IC_i}$ | 0.0000000 | 0.0809703                         | 0.1788153 |
|                 | 1.6000                | 8.1000                | 4.8000                | 0.2325505 | 0.3011000                         | 0.3224896 |
|                 | $1.6 + \delta_{IC_i}$ | $8.1 + \delta_{IC_i}$ | $4.8 + \delta_{IC_i}$ | 0.4169000 | 0.4091058                         | 0.3993869 |
| 0.1000          | 1.5000                | 8.0000                | 4.7000                | 0.2011171 | 0.2064559                         | 0.2422571 |
|                 | 1.6000                | 8.1000                | 4.8000                | 0.2325505 | 0.3011000                         | 0.3224896 |
|                 | 1.7000                | 8.2000                | 4.9000                | 0.2648704 | 0.3017855                         | 0.3316252 |
| 0.0100          | 1.5900                | 8.0900                | 4.7900                | 0.2096402 | 0.2152675                         | 0.2524027 |
|                 | 1.6000                | 8.1000                | 4.8000                | 0.2325505 | 0.3011000                         | 0.3224896 |
|                 | 1.6100                | 8.1100                | 4.8100                | 0.2573168 | 0.3017855                         | 0.3289632 |
| 0.0010          | 1.5990                | 8.0990                | 4.7990                | 0.2104765 | 0.2161322                         | 0.2533960 |
|                 | 1.6000                | 8.1000                | 4.8000                | 0.2325505 | 0.3011000                         | 0.3224896 |
|                 | 1.6010                | 8.1010                | 4.8010                | 0.2565573 | 0.3017855                         | 0.3286962 |

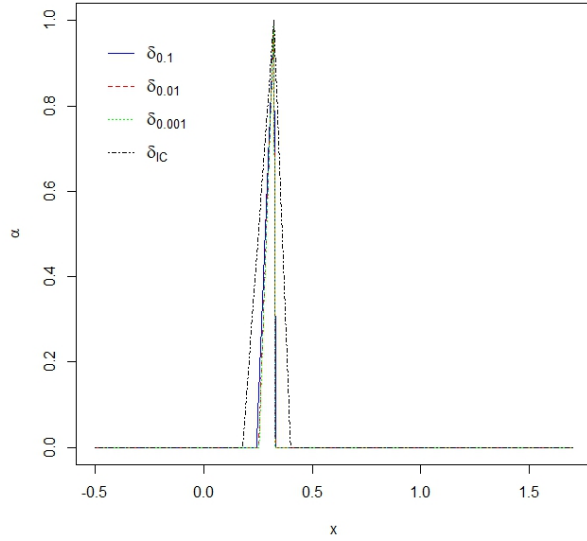


Figura 4.25: Números triangulares  $\mu$  y  $\sigma^2$  en el  $C_{pm}$  fuzzy con una  $Lognormal(0.3, 3.0^2)$  utilizando la mediana y  $n = 100$ .

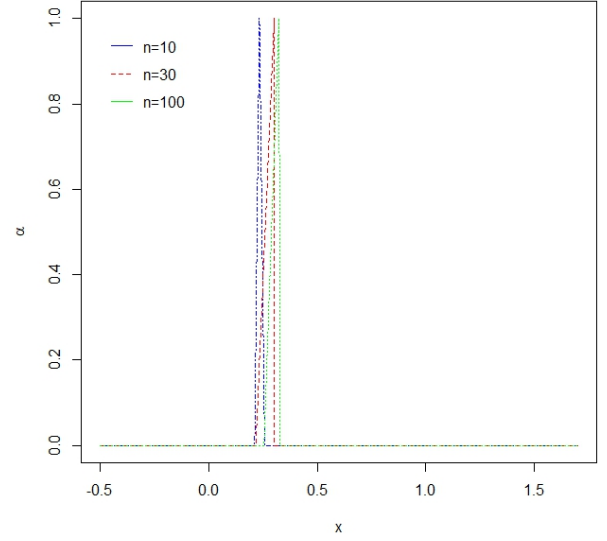


Figura 4.26: Números triangulares  $\mu$  y  $\sigma^2$  en el  $C_{pm}$  fuzzy con una  $Lognormal(0.3, 3.0^2)$  utilizando la mediana,  $\delta = 0.0010$  y  $n = (10, 30, 100)$ .

De los resultados encontrados se resalta, que al aumentar los tamaños de muestra para un mismo error se marca la tendencia de disminuir el valor de la capacidad del proceso, esto se debe a que al aumentar el tamaño de muestra el sesgo empírico y la varianza del proceso tiende a disminuir y por ende se llega a tener menor grado de incertidumbre.

Para terminar esta sección, se debe considerar de suma importancia el utilizar otros métodos para la construcción del índice de capacidad  $C_{pm}$ , principalmente cuando la distribución de las observaciones son asimétricas o se considera la población con una distribución diferente a la Normal, puesto



que se podría estar cometiendo el error de abordar la capacidad del proceso desde una perspectiva diferente a la real de los datos.

Es importante mencionar cómo con la distribución  $Lognormal(\alpha, \beta)$  con asimetría baja y moderada cuando se estudia el efecto de la construcción de los números triangulares en el  $C_{pm}$  fuzzy, este presenta un tamaño de intervalo en el índice de capacidad reducido e inferior que al momento de estimar el  $C_{pm}$  mirando el efecto de la construcción de los números triangulares  $\mu$  y  $\sigma^2$  en el  $C_{pm}$  fuzzy. Mientras que la distribución  $Lognormal(\alpha, \beta)$  con asimetría alta tiene un comportamiento contrario al anterior. Es decir, que mientras se considera difuso solamente los valores  $USL$ ,  $LSL$  y  $T$  el ancho de intervalo es superior que cuando se construyen los números triangulares para los parámetros  $\mu$  y  $\sigma^2$ .

### 4.3. Análisis de la Capacidad en el $C_{pm}$ fuzzy de los datos simulados

En esta sección se evaluara cada escenario de las distribuciones  $Gamma(\alpha, \beta)$  y  $Lognormal(\alpha, \beta)$  con asimetrías bajas, moderadas y altas estudiadas en la sección anterior para verificar si realmente cumplen con las condiciones encontradas en el Capítulo 4 de capacidad de un proceso al verificar si  $\tilde{C}_{pm}^l(\alpha) \geq 1$  utilizando la ecuación (3.6) propuesta en el capítulo anterior. Cabe resaltar que,  $LSL = 1.6$ ,  $USL = 8.1$  y  $T = 4.8$ , por lo tanto se cumple la condición de que  $u_2 - t_2 > t_2 - l_2$  y se procede a examinar cada uno de los escenarios estudiados, cuyos resultados se muestran en la tabla 5.13.

Se debe resaltar que, para cada tamaño de muestra se calcula un error estándar  $\delta_{IC_{ij}}$  con  $i = 1, 2, 3$  para  $t_i$  el vector de percentiles superiores de la distribución  $t$  con  $n - 1$  grados de libertad y  $n = 10, 30, 100$ , y con  $j = 1, 2, 3, 4, 5, 6$  para las distribuciones  $Gamma$  y  $Lognormal$  con asimetría baja, moderada y alta. Resultados que se pueden ver en la Tabla 4.13.

Tabla 4.13: Capacidad de un proceso para el estimador  $\tilde{C}_{pm}$  en la distribución indicada y n con asimetría baja, moderada y alta con  $LSL = 1.6$  y  $USL = 8.1$ .

|     |                    | <i>Gamma</i>    |                |                | <i>Lognormal</i> |                |                |
|-----|--------------------|-----------------|----------------|----------------|------------------|----------------|----------------|
| n   | $\delta$           | $\alpha = 10.0$ | $\alpha = 2.5$ | $\alpha = 1.0$ | $\alpha = 0.8$   | $\alpha = 0.5$ | $\alpha = 0.3$ |
|     |                    | $\beta = 0.34$  | $\beta = 0.66$ | $\beta = 0.98$ | $\beta = 0.5$    | $\beta = 1.4$  | $\beta = 3.0$  |
| 10  | $\delta_{IC_{ij}}$ | 46.1517         | 49.7813        | 65.9407        | 79.8879          | 95.6461        | 364.0549       |
|     | 0.1000             | 35.6178         | 59.4652        | 82.7916        | 93.2809          | 118.7865       | 341.1455       |
|     | 0.0100             | 34.5093         | 61.1124        | 85.5955        | 97.1125          | 121.4994       | 341.0053       |
|     | 0.0010             | 34.4025         | 61.2811        | 85.8799        | 97.4996          | 121.7747       | 340.9953       |
| 30  | $\delta_{IC_{ij}}$ | 17.6735         | 59.5640        | 90.9202        | 107.8742         | 117.2791       | 199.6753       |
|     | 0.1000             | 17.6794         | 68.3343        | 102.1966       | 114.0116         | 136.8079       | 233.1209       |
|     | 0.0100             | 17.8401         | 71.2134        | 106.1574       | 118.5515         | 141.1735       | 236.5223       |
|     | 0.0010             | 17.8602         | 71.5054        | 106.5575       | 119.0095         | 141.6141       | 236.8664       |
| 100 | $\delta_{IC_{ij}}$ | 15.3155         | 72.3023        | 108.8162       | 122.0994         | 142.0564       | 207.0609       |
|     | 0.1000             | 15.9981         | 76.1676        | 113.3381       | 123.1059         | 151.6587       | 231.5927       |
|     | 0.0100             | 16.6227         | 79.4969        | 117.7217       | 127.9046         | 156.6283       | 236.2883       |
|     | 0.0010             | 16.6892         | 79.8339        | 118.1640       | 128.3885         | 157.1293       | 236.7619       |

Con los resultados obtenidos en la tabla 4.13 donde se analiza cada distribución de acuerdo a sus

parámetros y a los valores de  $LSL$ ,  $USL$  y  $T$  si se cumple que  $\tilde{C}_{pm}^{*l}(\alpha) \geq 1$  como límite inferior para que todo el intervalo sea mayor a uno, se logra observar que al remplazar los diferentes valores en la ecuación (3.6) propuesta en el capítulo 4 no se cumple la condición necesaria para el límite inferior sea mayor que uno al tenerse resultados positivos en la salida de la ecuación.

## Capítulo 5

# Aplicación

En este capítulo se ilustra la metodología para medir la capacidad del proceso aplicándola a un estudio de caso real que proporciona dos conjuntos de datos recopilados durante dos estudios de la industria alimentaria y presentados en [34]. Inicialmente se recomienda valores mínimos recomendados para valores de PCIs que permitan detectar la capacidad. Luego, se aplica y se discute la metodología propuesta detallando el estudio de caso y por último se muestran los datos a considerar, se analizan y se establece la capacidad de los procesos de producción.

En la Tabla 5.1 se muestran los valores recomendados para el PCI mediante los cuales se puede decidir si el correspondiente proceso es capaz o no.

Tabla 5.1: Niveles de calidad y valores  $\tilde{C}_{pm}$

| Niveles de calidad | Valores de $\tilde{C}_{pm}$       |
|--------------------|-----------------------------------|
| Clase mundial      | $\tilde{C}_{pm} \geq 2$           |
| Excelente          | $1.67 \leq \tilde{C}_{pm} < 2$    |
| Satisfactorio      | $1.33 \leq \tilde{C}_{pm} < 1.67$ |
| Capaz              | $1 \leq \tilde{C}_{pm} < 1.33$    |
| Inadecuado         | $0.67 \leq \tilde{C}_{pm} < 1$    |
| Pobre              | $\tilde{C}_{pm} < 0.67$           |

### 5.1. Estudio de caso

La industria de alimentos se define como una organización responsable de la preparación y la distribución de alimentos de una manera científicamente planificada donde los preparados deben ser elaborados de acuerdo a las orientaciones nutricionales locales e internacionales que son reguladas por diferentes agencias como la Organización de las Naciones Unidas para la Alimentación y la Agricultura (FAO), la Organización Mundial de la Salud (OMS), entre otras. Así como por el Ministerio de Salud del país correspondiente [34].

El trabajo de las autoridades competentes es verificar que las regulaciones realmente se estén cumpliendo realizando diferentes auditorias como por ejemplo:

- la suficiencia y calidad nutricional de la dieta.
- los factores de corrección de los requerimientos energéticos por patologías.

Bajo este panorama, es fundamental que los servicios de alimentación aseguren el aporte de determinadas variables nutricionales como macronutrientes (carbohidratos, lípidos y proteínas), micronutrientes y fibra dietética, que son necesarios para cuidar y/o restaurar la salud de las personas. Para esta tarea, por lo general para modelar estas variables de interés nutricional se asume la distribución normal, pero este supuesto no corresponde necesariamente a la realidad de los datos. De este modo, para controlar el proceso de producción de alimentos en estos servicios y asegurar los contenidos nutricionales necesarios para cumplir con los requisitos de calidad, es posible trabajar con PCIs que consideren distribuciones estadísticas adecuadas para detectar si en el proceso se mantienen las cantidades de densidad energética establecidas en la normativa. En la Tabla 5.2 se muestran los rangos de recomendaciones nutricionales para adultos.

Tabla 5.2: Rangos de recomendaciones nutricionales para adultos.

| Macronutrientes | Densidad energética (en % kcal) |       | Cantidad (g) |        |
|-----------------|---------------------------------|-------|--------------|--------|
|                 | LSL                             | USL   | LSL          | USL    |
| Carbohidratos   | 50.00                           | 60.00 | 300.00       | 360.00 |
| Lípidos         | 25.00                           | 30.00 | 66.67        | 80.00  |
| Proteínas       | 12.00                           | 16.00 | 72.00        | 96.00  |

Según Martínez [27], el valor biológico de las proteínas de la dieta es primordial para muchas de las funciones vitales del cuerpo humano. En los servicios de alimentación y nutrición, la capacidad del proceso de producción depende del precio que se ofrezca en el menú de alimentos en la mayoría de las ocasiones, si este se presenta con un precio bajo, se vuelve difícil lograr la cantidad adecuada de proteínas para mantener el producto final dentro de los parámetros de calidad. Tradicionalmente en América Latina la disponibilidad de proteínas en la dieta humana ha sido baja, debido a su elevado precio, por lo que aumenta el costo de las preparaciones alimentarias [34].

Por otro lado, se considera importante señalar que algunas enfermedades son tratadas con dietas que restringen en la cantidad de proteínas, las cuales pueden llegar a límites inferiores de aportes nutricionales con solo el 5 % de la densidad energética de este macronutriente, que indica el *LSL* de esta cantidad a 30 gramos y dependiendo del número de pacientes que requieran la dieta bajo estas consideraciones, nos encontramos con que el rango de especificación de proteínas podría variar [34].

En este caso de estudio se cuenta con diferentes empresas de alimentación y nutrición que atienden a dos hospitales de Chile. Uno de esos servicios es una gran empresa con importante presencia en el mercado alimentario chileno que, según información proporcionada por ellos, esta cuenta con un proceso de producción de alimentos cuya cantidad de proteína se considera capaz con base en el PCI en normalidad, es decir,  $C_p > 1$ . Mientras que el segundo servicio corresponde a una pequeña empresa, que se preocupa por mantener el proceso de producción de alimentos cuya cantidad de proteína se puede considerar, por ejemplo, para pacientes con dietas restringidas en la cantidad de proteína, por lo que la especificación se encuentra en  $[30g, 96g]$ , pero esta empresa no se preocupa por mantener un proceso de producción de alimentos que sea capaz en todos los rangos especificados por la normativa internacional como se muestra en la Tabla 5.1.

Así, el objetivo de este estudio en la industria alimentaria es determinar la capacidad del proceso de producción de alimentos en dos hospitales chilenos. Aquí, la característica de calidad a monitorear es la cantidad de proteínas, con  $LSL = 30$  o  $72g$  y  $USL = 96g$ , para saber si esta característica de la carta de comida está de acuerdo con las recomendaciones dietéticas establecidas.

## 5.2. Conjunto de datos para el estudio de caso

Los datos fueron recolectados durante 27 semanas cubriendo el período desde el 20-nov-2011 al 26-mayo-2012 (189 días) en un estudio dirigido por Fernando Rojas y Víctor Leiva en la Universidad de Valparaíso-Chile. Estos datos se dan en las Tablas 5.3 y 5.4.

### 5.2.1. Conjunto de datos I

El primer conjunto de datos de la Tabla 5.3 representa la cantidad de proteína. Aquí, la característica  $X$  que se va a monitorear es la cantidad de proteína (en g) en la dieta restringida para pacientes adultos en el Hospital Carlos Van Buren de la ciudad de Valparaíso, Chile.

Tabla 5.3: Conjunto I

|        |        |        |        |        |        |        |       |       |        |
|--------|--------|--------|--------|--------|--------|--------|-------|-------|--------|
| 66.09  | 146.48 | 97.99  | 56.77  | 74.23  | 76.20  | 108.90 | 89.51 | 62.40 | 141.15 |
| 89.81  | 201.05 | 210.32 | 105.08 | 169.82 | 64.25  | 100.43 | 59.56 | 49.97 | 102.16 |
| 164.05 | 59.63  | 51.41  | 93.32  | 39.43  | 102.91 | 124.91 | 56.16 | 70.94 | 77.26  |
| 59.01  | 49.21  | 123.85 | 59.78  | 163.50 | 37.67  | 88.27  | 35.60 | 42.54 | 43.87  |
| 44.41  | 25.55  | 41.12  | 88.14  | 54.69  | 85.32  | 67.55  | 17.76 | 83.73 | 34.07  |
| 42.37  | 61.95  | 68.62  | 41.69  | 100.43 | 139.39 | 60.03  | 34.35 | 51.28 | 68.22  |
| 80.44  |        |        |        |        |        |        |       |       |        |

### 5.2.2. Conjunto de datos II

El segundo conjunto de datos que se dan en la Tabla 5.4 representa la cantidad de proteína.

Tabla 5.4: Conjunto II

|        |       |       |       |       |       |       |       |       |       |
|--------|-------|-------|-------|-------|-------|-------|-------|-------|-------|
| 90.98  | 95.41 | 91.87 | 94.01 | 91.59 | 91.42 | 92.45 | 93.36 | 91.75 | 91.97 |
| 93.73  | 92.57 | 94.51 | 93.86 | 91.57 | 91.61 | 94.28 | 98.12 | 90.18 | 92.45 |
| 91.02  | 96.42 | 91.75 | 92.15 | 90.55 | 93.38 | 99.97 | 90.71 | 89.88 | 92.05 |
| 100.52 | 94.12 | 91.78 | 91.6  | 90.4  | 94.16 | 97.76 | 90.3  | 93.3  | 93.4  |
| 95.47  | 91.39 | 91.11 | 93.59 | 92.27 | 92.65 | 90.59 | 94.03 | 96.34 | 91.64 |
| 90.67  | 93.3  | 91.32 | 93.13 | 92.08 | 89.1  | 93.04 | 90.1  | 90.59 | 93.34 |
| 90.12  | 90.51 | 91.12 | 92.13 | 93.37 | 92.16 | 92.46 | 91.72 | 97.22 | 92.68 |
| 96.54  | 92.09 | 95.18 | 93.45 | 91.89 | 94.54 | 93.97 | 92.44 | 93.56 | 93.81 |
| 92.53  | 95.8  | 93.54 | 94.78 | 96.31 | 94.48 | 92.17 | 92.24 | 93.35 | 92.84 |
| 93.53  | 97.53 | 92.95 | 93.67 | 92.46 |       |       |       |       |       |

En este caso, nuevamente la característica de calidad  $X$  a monitorear para determinar la capacidad de proceso es la cantidad de proteína (en g) en una dieta estricta para adultos en el Hospital San Juan de Dios de la ciudad de Los Andes, Chile. Los datos fueron recolectados desde el 01-abr-2013 al 03-jul-2013 por la nutricionista Daniela Pincheira en un servicio de alimentación anónimo.

### 5.2.3. Estadísticas descriptivas de los conjuntos de datos y pruebas de bondad de ajuste

Se realizan algunas estadísticas descriptivas para obtener información del comportamiento de los datos en ambos conjuntos de observaciones.

Las estadísticas descriptivas se obtienen utilizando diferentes funciones de R como *quantile()* y *summary()*. Estas estadísticas mostradas en la tabla 5.5 son valores empíricos o muestrales para la media  $\bar{x}$ , Cuartiles ( $Q_1, Q_2, Q_3$ ) mínimo (mín), máximo (máx), rango, desviación estándar ( $S$ ) y el coeficiente de variación ( $CV$ ) y en las Figuras 5.1 y 5.2 se presentan el histograma con su respectiva curva de densidad y cuantiles teóricos vs cuantiles muestrales para cada conjunto de datos y así tener una idea acerca de la normalidad de la cantidad de proteína en gramos que se monitorea en la dieta restringida de los pacientes adultos en el Hospital Carlos Van Buren de la ciudad de Valparaíso, Chile.

Tabla 5.5: Estadísticas de los conjuntos de datos

| Conjunto | $n$ | mín   | $Q_{25}$ | Mediana | $\bar{x}$ | $Q_{75}$ | máx    | Rango  | $S$   | $CV$  |
|----------|-----|-------|----------|---------|-----------|----------|--------|--------|-------|-------|
| I        | 61  | 17.76 | 51.28    | 68.22   | 80.44     | 100.43   | 210.32 | 192.56 | 42.29 | 52.58 |
| II       | 95  | 89.10 | 91.62    | 92.57   | 93.03     | 93.92    | 100.52 | 11.42  | 2.15  | 2.31  |

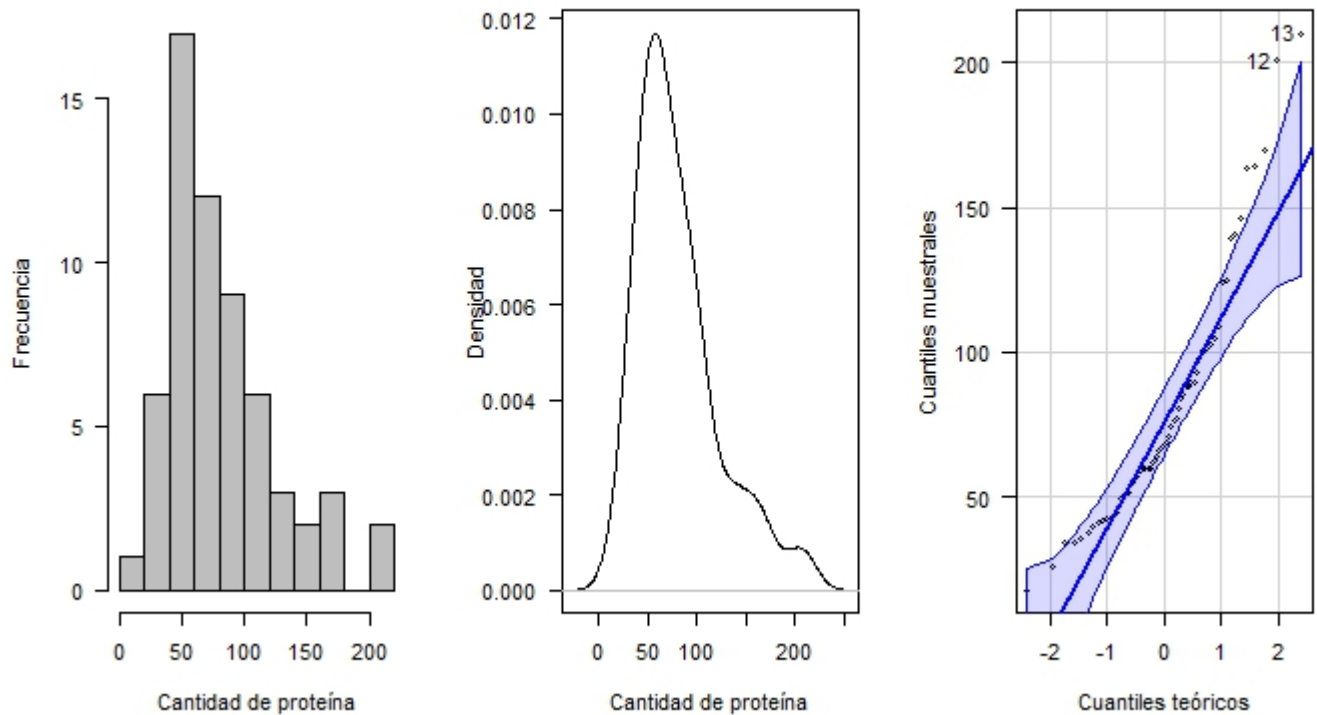


Figura 5.1: Histograma, densidad y cuantiles teóricos vs cuantiles muestrales para el conjunto de datos I.

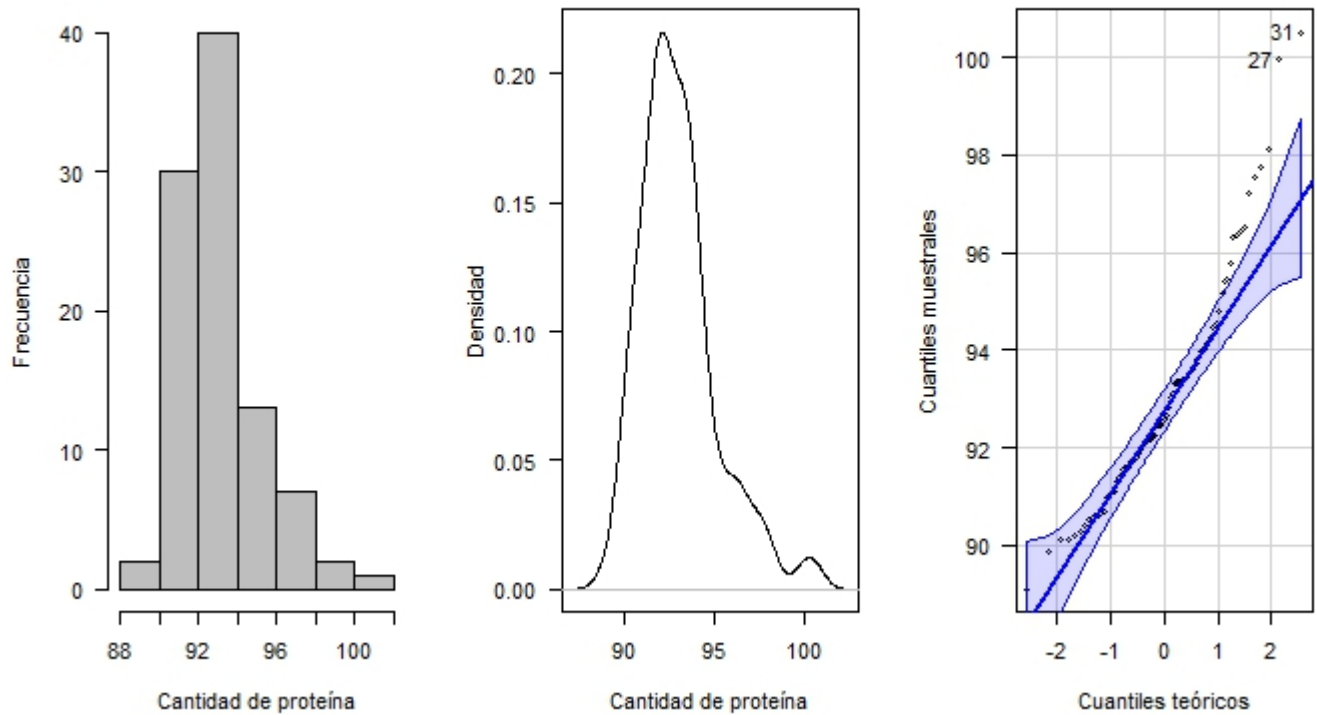


Figura 5.2: Histograma, densidad y cuantiles teóricos vs cuantiles muestrales para el conjunto de datos II.

En las Figuras 5.1 y 5.2 de los conjuntos de datos I y II respectivamente, tanto el histograma como el gráfico de densidad parecen mostrar una distribución no tan normal, mientras que en el gráfico de cuantiles teóricos vs cuantiles muestrales se observan algunos puntos por fuera de las bandas de confianza que hacen sospechar de la posible no normalidad de los datos de ambos conjuntos.

Luego, se procede a estimarse y presentarse en la Tabla 5.6 los parámetros de la distribución hipotética, que en este caso es la normal, a partir de la función *fitdistr()* que pertenece a la librería *MASS*, cabe resaltar que también estima parámetros de muchas otras distribuciones. Se debe tenerse en cuenta que la función *fitdistr()* tiene como argumentos el nombre de la variable donde se encuentran los datos y el nombre de la distribución hipotética.

Tabla 5.6: Parámetros estimados de la distribución hipotética de los conjuntos I y II

| Conjunto I |          | Conjunto II |          |
|------------|----------|-------------|----------|
| $\bar{x}$  | $S$      | $\bar{x}$   | $S$      |
| 80.4361    | 41.9417  | 93.0295     | 2.1390   |
| (5.3701)   | (3.7972) | (0.2195)    | (0.1552) |

Una vez estimados los parámetros de la distribución hipotética y su desviación estándar que es el valor entre paréntesis debajo de la estimación de cada parámetro, se usa una prueba de bondad de ajuste para finalmente concluir si la variable analizada sigue o no la distribución hipotética plan-

teada. Se utilizan las pruebas Kolmogorov Smirnov, Anderson Darling y la prueba Shapiro Wilk y se presentan los resultados en la Tabla 5.7. La prueba Kolmogorov Smirnov se obtiene a través de la función `ks.test()`, la prueba Anderson Darling se calcula a partir de la función `ad.testdel()` de la librería `goftest` y finalmente, la prueba shapiro-Wilk con la función `shapiro.test()`.

Tabla 5.7: Prueba de bondad de ajuste para la distribución hipotética

|             | Conjunto I            |                     |                 | Conjunto II           |                     |                 |
|-------------|-----------------------|---------------------|-----------------|-----------------------|---------------------|-----------------|
|             | Kolmogorov<br>Smirnov | Anderson<br>Darling | Shapiro<br>Wilk | Kolmogorov<br>Smirnov | Anderson<br>Darling | Shapiro<br>Wilk |
| Estadístico | 0.1355                | 1.9172              | 0.8982          | 0.1019                | 1.7202              | 0.9297          |
| $P - valor$ | 0.2126                | 0.1022              | 0.0001          | 0.2773                | 0.1317              | 0.0001          |

De acuerdo al  $p - valor$  de las pruebas Kolmogorov Smirnov, y Anderson Darling observados en la Tabla 5.7, los conjuntos I y II siguen una distribución normal, puesto que, los valores son mayores al nivel de significancia de 0.05. Sin embargo, al analizar la prueba Shapiro-Wilk que es un test que se considera uno de los más potentes para el contraste de normalidad se observa que el  $p - valor$  para ambos casos es de 0.0001 y es inferior al nivel de significancia, por lo tanto se rechaza la hipótesis nula y se concluye, con un nivel de confianza del 95 %, que los datos no siguen una distribución normal y esto le da sentido a lo observado en el análisis exploratorio.

Para ilustrar los resultados obtenidos anteriormente, las Figuras 5.3 y 5.4 muestran la distribución acumulada empírica de la cantidad de proteína para cada conjunto y la teórica de una distribución normal. Se puede notar que no es muy claro que en ambos casos se sigan el mismo patrón entre la acumulada empírica y la distribución normal.

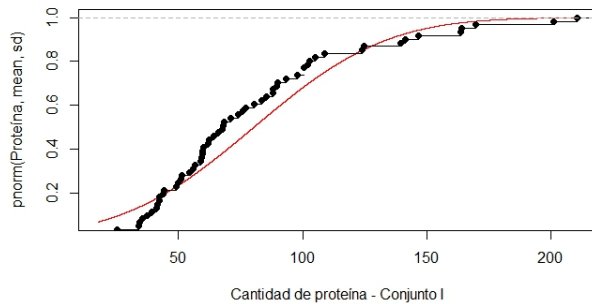


Figura 5.3: Distribución acumulada empírica y teórica de una distribución normal con los datos del conjunto I.

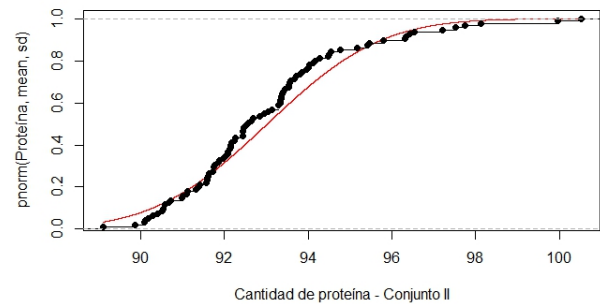


Figura 5.4: Distribución acumulada empírica y teórica de una distribución normal con los datos del conjunto II.

Luego, para estudiar el ajuste de los datos a una distribución específica se utiliza la librería `fitdistrplus` que permite ajustar distribuciones univariadas de diferentes tipos de datos, y además permite diferentes métodos de estimación permitiendo comparar diferentes ajustes entre si. Para decidir el tipo de distribución los estadísticos de la Kurtosis y Skewness serán de gran ayuda puesto que un valor de Skewness diferente a cero indica la falta de simetría de la distribución empírica, mientras que la kurtosis cuantifica el peso de los extremos, se debe tener en cuenta que un valor de



3 en la Kurtosis se obtiene una distribución normal. Para estudiar lo anterior se utiliza la función *descdist()* obteniendose los resultados de las Figuras 5.5 y 5.6.

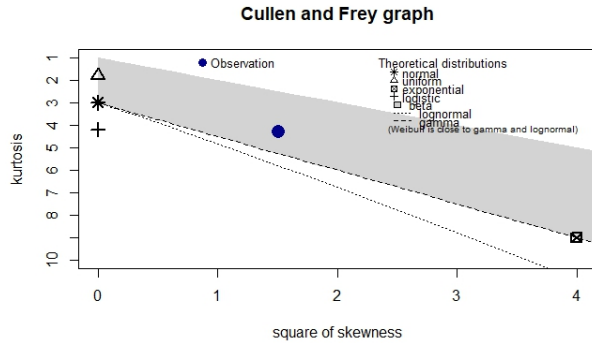


Figura 5.5: Kurtosis y Skewness para los datos del conjunto I.

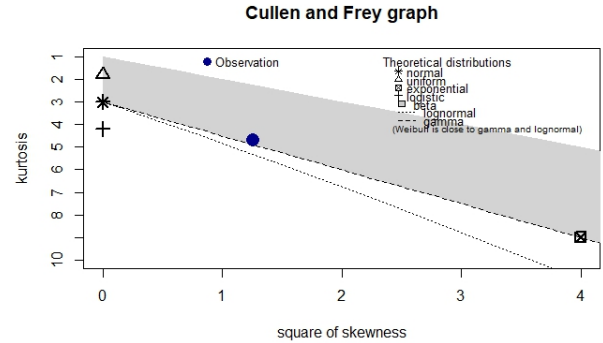


Figura 5.6: Kurtosis y Skewness para los datos del conjunto II.

La Figura 5.7 muestra el ajuste de las distribuciones teóricas a los datos del conjunto I, en la gráfica de densidad y la gráfica CDF que se consideran como las gráficas clásicas básicas de bondad de ajuste para las distribuciones *weibull*, *lognormal* y *gamma*.

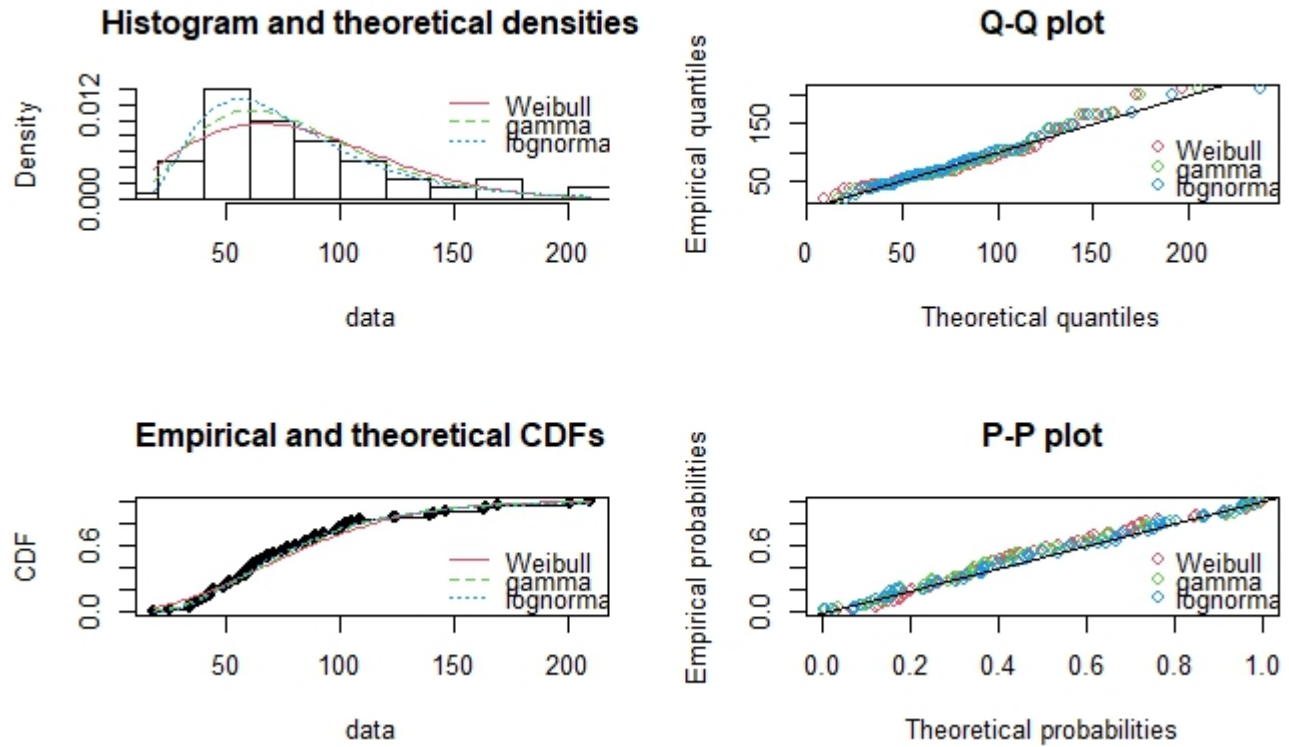


Figura 5.7: Comparación de las distribuciones teóricas *Weibull*, *Lognormal* y *Gamma* para el conjunto de datos I.

En la Figura 5.7 se observa que las distribuciones *Lognormal* y *Gamma* tienen un mejor ajuste al tener una mejor descripción en la zona media y derecha de la distribución empírica. Las otras dos gráficas son complementarias y nos dan información extra de las distribución de los datos. El gráfico Q-Q plot enfatiza la falta de ajuste en las colas de distribución, principalmente en la parte derecha, mientras que el gráfico P-P plot enfatiza la falta de ajuste en el centro de distribución mostrando un comportamiento similar para las tres distribuciones, pero con un mejor ajuste para las distribuciones *Lognormal* y *Gamma*.

Luego, toca determinar estadísticamente la bondad del ajuste de los datos a los diferentes modelos, para ello hay tres estadísticos Cramer-von Mises, Kolmogorov-Smirnov y Anderson-Darling, además de los criterios de AIC y BIC. Para utilizarlos empleamos la función *gofstat()* y se obtiene la salida de la Tabla 5.8.

Tabla 5.8: Estadísticos y criterios de bondad de ajuste para los datos del conjunto I

| Estadísticas de bondad de ajuste | Weibull    | Gamma      | Lognormal  |
|----------------------------------|------------|------------|------------|
| Kolmogorov-Smirnov               | 0.09749843 | 0.08326739 | 0.05072701 |
| Cramer-von Mises                 | 0.15274432 | 0.07047027 | 0.02254182 |
| Anderson-Darling                 | 0.98181522 | 0.45566638 | 0.17452704 |
| Criterios de bondad de ajuste    | Weibull    | Gamma      | Lognormal  |
| Akaike's (AIC)                   | 621.3273   | 615.3771   | 613.4097   |
| Bayesian (BIC)                   | 625.5491   | 619.5989   | 617.6314   |

De la salida anterior hay que quedarse con los valores más bajos, tanto de los estadísticos como de Aikake (AIC) y Bayesian (BIC). Para los datos de cantidades de proteína del conjunto I se tiene que los ajustes de la *Lognormal* y *Gamma* son los mejores. Ahora, hay que calcular el grado de incertidumbre de los modelos que se puede hacer por bootstraps paramétricos o no paramétricos con la función *boddist()*. Además, como el mejor ajuste fue el de la distribución *Lognormal*, en la Tabla 5.9 se calcula el grado de incertidumbre a un nivel de confianza del 95 % y extrayendo 5000 muestras por bootstrap.

Tabla 5.9: Medianas de bootstrap paramétricos e intervalo de confianza al 95 % para los datos del conjunto I

| Estadísticos | Mediana   | 2.5 %     | 97.5 %    |
|--------------|-----------|-----------|-----------|
| meanlog      | 4.2646406 | 4.1303195 | 4.3876997 |
| sdlog        | 0.4965538 | 0.4098438 | 0.5894708 |

Luego, en la Tabla 5.10 ajustando la distribución a una *Lognormal* por máxima verosimilitud se tienen los parámetros de la distribución mencionada anteriormente:

Tabla 5.10: parámetros por máxima verosimilitud de una *Lognormal* para los datos del conjunto I

| Estadísticos | Estimación | Error estándar |
|--------------|------------|----------------|
| meanlog      | 4.26162    | 0.0645165      |
| sdlog        | 0.50389    | 0.0456193      |

Ahora, en la figura 5.8 se observa el ajuste de las distribuciones teóricas a los datos del conjunto II,

en la gráfica de densidad y la gráfica CDF se observa que las distribuciones *Lognormal* y *Gamma* tienen un mejor ajuste al tener una mejor descripción en la zona media y derecha de la distribución empírica. En el gráfico Q-Q plot que enfatiza la falta de ajuste en las colas de distribución, sugiere que principalmente en la parte izquierda hay un mayor desajuste que es más notorio para la distribución *Weibull*, mismo comportamiento que se observa en el gráfico P-P plot que enfatiza la falta de ajuste en el centro de distribución mostrando un comportamiento similar para las tres distribuciones, pero con un mejor ajuste para las distribuciones *Lognormal* y *Gamma*.

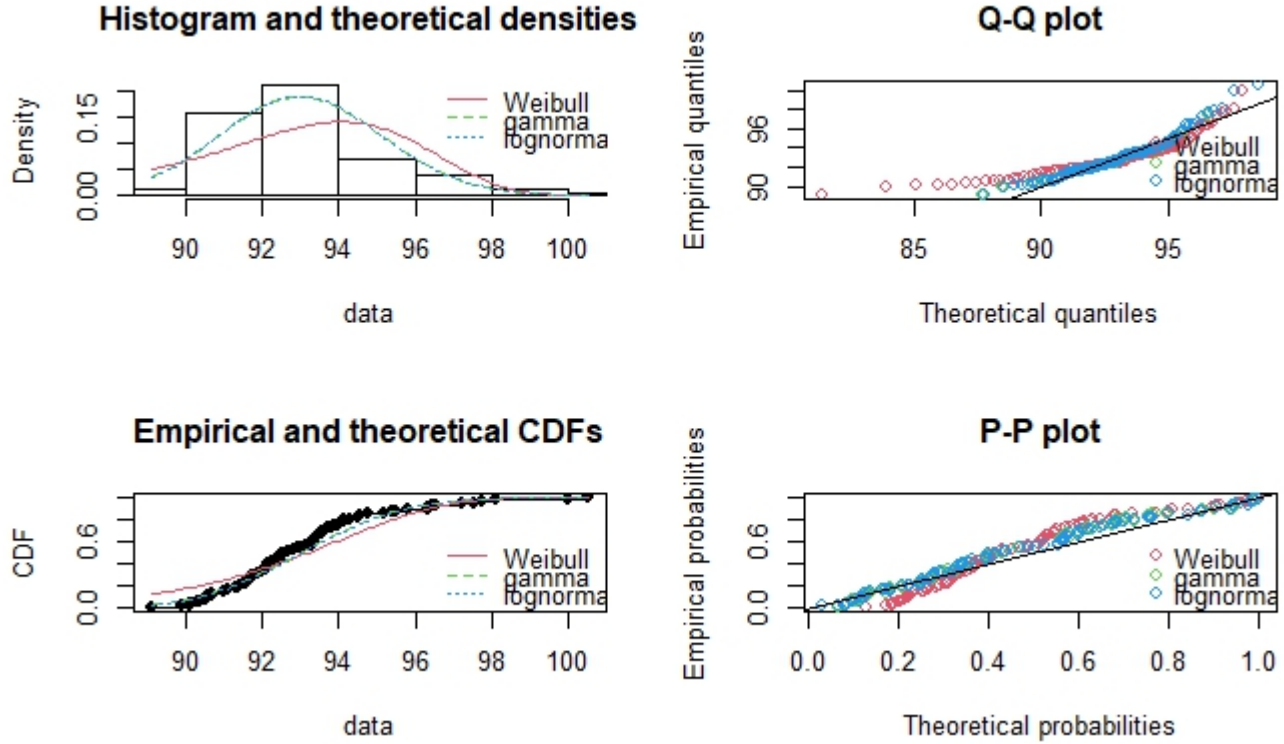


Figura 5.8: Comparación de las distribuciones teóricas *Weibull*, *Lognormal* y *Gamma* para el conjunto de datos II.

Después, hay que determinar estadísticamente la bondad del ajuste de los datos a los diferentes modelos. Para ello se usan los estadísticos de Cramer-von Mises, Kolmogorov-Smirnov y Anderson-Darling, además de los criterios de AIC y BIC. Se obtienen los resultados de la Tabla 5.11.

Tabla 5.11: Estadísticos y criterios de bondad de ajuste para los datos del conjunto II

| Estadísticas de bondad de ajuste | Weibull    | Gamma      | Lognormal  |
|----------------------------------|------------|------------|------------|
| Kolmogorov-Smirnov               | 0.16793020 | 0.09837357 | 0.09660343 |
| Cramer-von Mises                 | 0.90182710 | 0.25340337 | 0.24324515 |
| Anderson-Darling                 | 5.33244570 | 1.59631608 | 1.53625320 |
| Criterios de bondad de ajuste    | Weibull    | Gamma      | Lognormal  |
| Akaike's (AIC)                   | 458.9729   | 416.5118   | 415.7625   |
| Bayesian (BIC)                   | 464.0806   | 421.6196   | 420.8703   |

Cabe resaltar que de la salida para el conjunto II hay que quedarse con los valores más bajos, tanto de los estadísticos como de Aikake (AIC) y Bayesian (BIC). Entonces, para los datos de cantidades de proteína del conjunto II se tiene que de igual forma al conjunto I los ajustes de la *Lognormal* y *Gamma* son los mejores. Luego, hay se calcula el grado de incertidumbre de los modelos por medio de bootstraps paramétricos con la función *boodist()*. Además, como el mejor ajuste fue el de la distribución *Lognormal* se calcula el grado de incertidumbre a un nivel de confianza del 95 % y extrayendo 5000 muestras por bootstrap, resultados que se pueden ver en la Tabla 5.12.

Tabla 5.12: Medianas de bootstrap paramétricos e intervalo de confianza al 95 % para los datos del conjunto II

| Estadísticos | Mediana   | 2.5 %     | 97.5 %    |
|--------------|-----------|-----------|-----------|
| meanlog      | 4.5327178 | 4.5279431 | 4.5374018 |
| sdlog        | 0.0225716 | 0.0194646 | 0.0258066 |

Por último, ajustando la distribución a una *Lognormal* utilizando el método de máxima verosimilitud se tienen los parámetros que se muestran en la Tabla 5.13:

Tabla 5.13: parámetros por máxima verosimilitud de una *Lognormal* para los datos del conjunto II

| Estadísticos | Estimación | Error estándar |
|--------------|------------|----------------|
| meanlog      | 4.5326562  | 0.00233122     |
| sdlog        | 0.0227219  | 0.00163408     |

Así, después de las pruebas de bondad de ajuste se encontró que las mediciones obtenidas en los conjuntos I y II se pueden ajustar a una distribución *Lognormal*.

#### 5.2.4. Capacidad del proceso $C_{pm}$

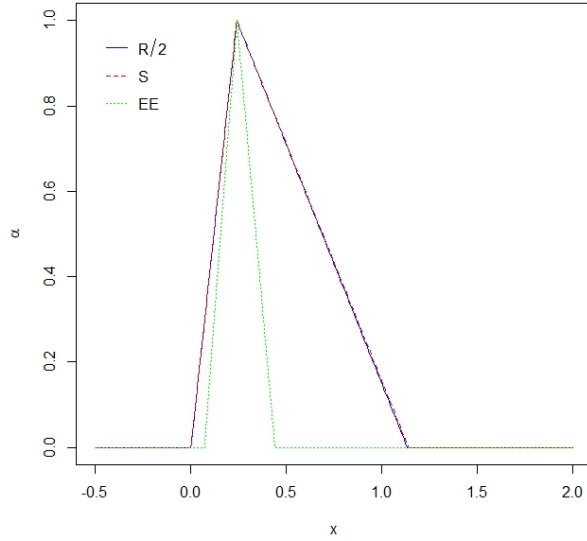
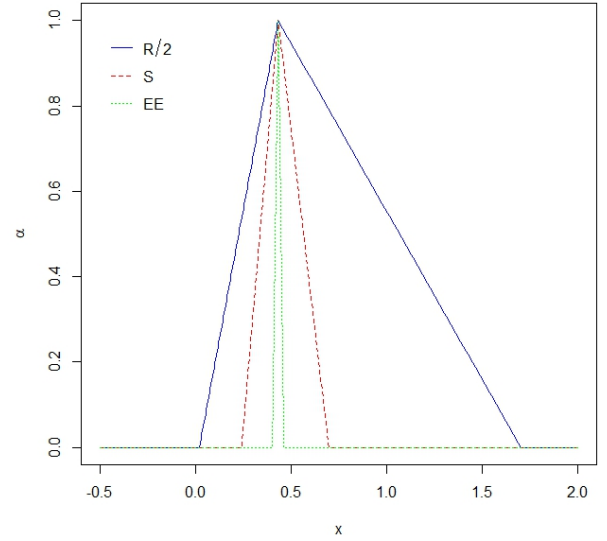
Para estudiar la capacidad de los dos procesos de producción tenemos los límites de especificación para los niveles de proteína  $LSL = 72.00$  y  $USL = 96.00$  como se muestran en la tabla 5.1 y valor objetivo  $T = 84.00$  como el punto medio de los límites de especificación para el Hospital de San Juan de Dios en la ciudad de los Andes, Chile. Mientras que los límites de especificación para los niveles de proteína en el Hospital Carlos Van Buren de la ciudad de Valparaíso, Chile, donde se tiene una dieta restringida a los pacientes son  $LSL = 30.00$ ,  $USL = 96.00$  y valor objetivo  $T = 63.00$ . Luego se estima el  $C_{pm}$  usando la expresión dada en la ecuación (1.24).

Realizando las estimaciones con diferentes niveles de error  $\delta$  como  $S$ ,  $EE = Z_{1-\alpha/2} \frac{S}{\sqrt{n}}$  y  $R/2$  con  $R = x(0.99865) - x(0.00135)$  que cubre un 99.73 % de la distribución de los datos con  $x(0.00135)$  y  $x(0.99865)$  los cuartiles 0.135 y 99.865. Suponga inicialmente las estimaciones de los parámetros dadas por: La media ( $\bar{x} = 80.44$ ) y varianza ( $S^2 = 42.29^2$ ) de la muestra en la población I y  $\bar{x} = 93.03$ ,  $S^2 = 2.15^2$  obtenidos de la muestra de la población II. Se obtienen los resultados de la Tabla 5.14.

En este orden de ideas, se puede observar en la Tabla 5.14 que al momento de utilizar la estimación de la media como parámetro de centralidad y la varianza como dispersión para los conjuntos I y II resulta no capaz el proceso. Cabe resaltar que ha medida que el margen de error aumenta hay un incremento en el intervalo de los límites difusos como puede observarse en las Figuras 5.9 y 5.10.

Tabla 5.14: Efecto de la construcción de los números triangulares en el  $C_{pm}$  fuzzy

| Conjunto | Margen de error | $\widetilde{LSL}$ | $\widetilde{USL}$ | $\widetilde{T}$ | $\widetilde{C}_{pmi}$ |
|----------|-----------------|-------------------|-------------------|-----------------|-----------------------|
| I        | $R/2$           | 0.0000            | 0.4109            | 0.0000          | 0.0000000             |
|          |                 | 30.0000           | 96.0000           | 63.0000         | 0.2404726             |
|          |                 | 125.5891          | 191.5890          | 157.5890        | 1.1336010             |
|          | $S$             | 0.0000            | 53.7102           | 19.7102         | 0.0000000             |
|          |                 | 30.0000           | 96.0000           | 63.0000         | 0.2404726             |
|          |                 | 72.2898           | 138.2898          | 104.2898        | 1.1407321             |
|          | $EE$            | 19.3875           | 85.3875           | 52.3875         | 0.0712645             |
|          |                 | 30.0000           | 96.0000           | 63.0000         | 0.2404726             |
|          |                 | 40.6125           | 106.6125          | 73.6125         | 0.4400262             |
| II       | $R/2$           | 66.3744           | 90.3744           | 77.3744         | 0.0182897             |
|          |                 | 72.0000           | 96.0000           | 84.0000         | 0.4309419             |
|          |                 | 77.6256           | 101.6256          | 88.6256         | 1.7020302             |
|          | $S$             | 69.8496           | 93.8496           | 81.8496         | 0.2391970             |
|          |                 | 72.0000           | 96.0000           | 84.0000         | 0.4309419             |
|          |                 | 74.1504           | 98.1504           | 86.1504         | 0.6963804             |
|          | $EE$            | 71.5676           | 95.5676           | 83.5676         | 0.4002402             |
|          |                 | 72.0000           | 96.0000           | 84.0000         | 0.4309419             |
|          |                 | 72.4324           | 96.4324           | 84.4324         | 0.4602408             |


 Figura 5.9: Números triangulares en el  $C_{pm}$  fuzzy con el conjunto de datos I.

 Figura 5.10: Números triangulares en el  $C_{pm}$  fuzzy con el conjunto de datos II.

Ahora, veamos los resultados del efecto de la construcción de los números triangulares en el  $C_{pm}$  fuzzy asumiendo que los datos siguen una distribución *Lognormal* como se encontró en la sección anterior. Para este caso se utiliza el método introducido por Clements [14] donde se usa la mediana como la medida de tendencia central del proceso en lugar de la media, puesto que es una medida

más robusta de la tendencia central de un proceso, específicamente para distribuciones asimétricas de colas pesadas. Cuyos resultados, se muestran en la Tabla 5.15 y en las Figuras 5.11 y 5.12.

Tabla 5.15: Efecto de la construcción de los números triangulares en el  $C_{pm}$  fuzzy asumiendo no normalidad y asimetría

| Conjunto | Margen de error | $\widetilde{LSL}$ | $\widetilde{USL}$ | $\widetilde{T}$ | $\widetilde{C}_{pmi}$ |
|----------|-----------------|-------------------|-------------------|-----------------|-----------------------|
| I        | $R/2$           | 0.0000            | 0.4109            | 0.0000          | 0.0000000             |
|          |                 | 30.0000           | 96.0000           | 63.0000         | 1.2584923             |
|          |                 | 125.5891          | 191.5890          | 157.5890        | 2.2197106             |
|          | $EE$            | 0.0000            | 53.7102           | 19.7102         | 0.2554806             |
|          |                 | 30.0000           | 96.0000           | 63.0000         | 1.2584923             |
|          |                 | 72.2898           | 138.2898          | 104.2898        | 3.0430127             |
|          | $Sd.log$        | 29.4961           | 95.4961           | 62.4961         | 1.1123290             |
|          |                 | 30.0000           | 96.0000           | 63.0000         | 1.2584923             |
|          |                 | 30.5039           | 96.5039           | 63.5039         | 1.3319310             |
|          | $R/2$           | 66.3744           | 90.3744           | 77.3744         | 0.0187366             |
|          |                 | 72.0000           | 96.0000           | 84.0000         | 0.4596338             |
|          |                 | 77.6256           | 101.6256          | 88.6256         | 2.0762534             |
| II       | $EE$            | 69.8496           | 93.8496           | 81.8496         | 0.4229343             |
|          |                 | 72.0000           | 96.0000           | 84.0000         | 0.4596338             |
|          |                 | 74.1504           | 98.1504           | 86.1504         | 0.4962171             |
|          | $Sd.log$        | 71.9773           | 95.9773           | 83.9773         | 0.4596338             |
|          |                 | 72.0000           | 96.0000           | 84.0000         | 0.4596338             |
|          |                 | 72.0227           | 96.0227           | 84.0227         | 0.4596338             |

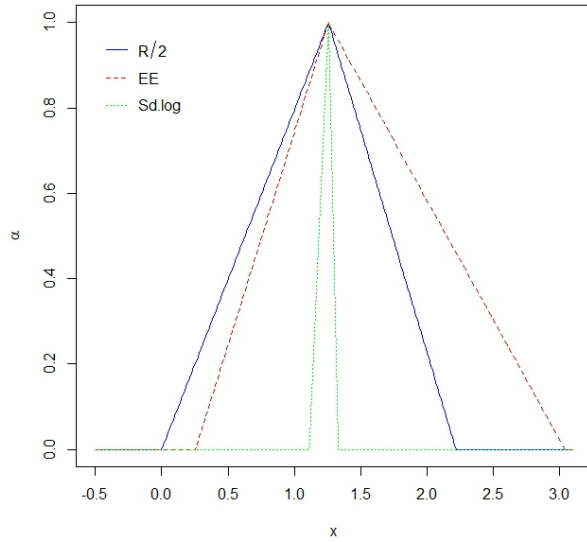


Figura 5.11: Números triangulares en el  $C_{pm}$  fuzzy asumiendo no normalidad y asimetría con el conjunto de datos I.

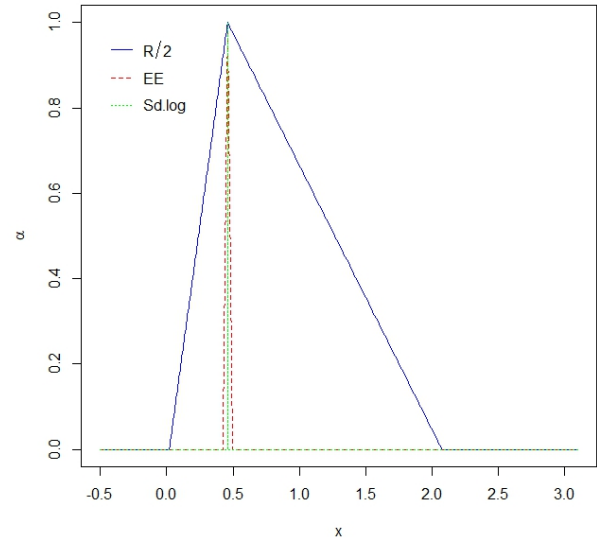


Figura 5.12: Números triangulares en el  $C_{pm}$  fuzzy asumiendo no normalidad y asimetría con el conjunto de datos II.

Además, Pearn y Chen [31] modifican el método de Clements por una forma más general para los PCI's no normales utilizando  $x(q_2) - x(q_1)$  en vez de usar  $3\sigma$  teniendo en cuenta la no normalidad y asimetría de la distribución. Los valores de  $q_1$  y  $q_2$  se pueden obtener utilizando algún criterio estadístico, para este caso se utilizan el primer y tercer cuartil de los datos, mientras que se agrega como error la desviación estándar de la estimación del parámetro de centralidad de la aproximación de la distribución *Lognormal* por máxima verosimilitud.

Así, en la Tabla 5.15 al considerar la mediana como medida de tendencia central y  $q(0.75) - q(0.25)$  como medida de dispersión se logra observar que a medida que la variabilidad de un proceso disminuye y a su vez el margen de error de la construcción de los números triangulares en el  $C_{pm}$  fuzzy para el conjunto de datos I que presenta un mejor ajuste a la distribución empírica y mayor carga hacia la cola izquierda se logra obtener un proceso capaz como logra verse en la Figura 5.11. Para el conjunto de datos II, se observa en la Figura 5.12 que utilizando este segundo método de igual forma no se considera capaz.

Luego, al estudiar el efecto del cambio de números triangulares de  $\mu$  y  $\sigma^2$  en la estimación del PCI  $C_{pm}$  fuzzy se encontró en los resultados de la Tabla 5.16 y Figuras 5.13 y 5.14 que utilizando la estimación clásica de la media y varianza como parámetros de centralidad y dispersión el proceso se considera pobre al tener un  $\tilde{C}_{pm} < 0.67$  en su límite superior. Además, se tiene que entre mayor sea el rango entre el conjunto de los datos se obtiene un menor ancho de intervalo al utilizar este proceso con márgenes de error como el rango y la desviación estándar.

Tabla 5.16: Efecto del cambio de números triangulares de  $\mu$  y  $\sigma^2$  en la estimación del PCI  $C_{pm}$  fuzzy.

| Conjunto | Margen de error | $\widetilde{LSL}$ | $\widetilde{USL}$ | $\tilde{T}$ | $\tilde{C}_{pmi}$ |
|----------|-----------------|-------------------|-------------------|-------------|-------------------|
| I        | $R/2$           | 0.0000            | 0.4109            | 0.0000      | 0.0000000         |
|          |                 | 30.0000           | 96.0000           | 63.0000     | 0.2509425         |
|          |                 | 125.5891          | 191.5890          | 157.5890    | 0.7102434         |
|          | $S$             | 0.0000            | 53.7102           | 19.7102     | 0.0000000         |
|          |                 | 30.0000           | 96.0000           | 63.0000     | 0.2509425         |
|          |                 | 72.2898           | 138.2898          | 104.2898    | 0.5742776         |
|          | $EE$            | 19.3875           | 85.3875           | 52.3875     | 0.0719322         |
|          |                 | 30.0000           | 96.0000           | 63.0000     | 0.2509425         |
|          |                 | 40.6125           | 106.6125          | 73.6125     | 0.3247266         |
|          | $R/2$           | 66.3744           | 90.3744           | 77.3744     | 0.0168738         |
|          |                 | 72.0000           | 96.0000           | 84.0000     | 0.4413139         |
|          |                 | 77.6256           | 101.6256          | 88.6256     | 2.5413786         |
| II       | $S$             | 69.8496           | 93.8496           | 81.8496     | 0.2177293         |
|          |                 | 72.0000           | 96.0000           | 84.0000     | 0.4413139         |
|          |                 | 74.1504           | 98.1504           | 86.1504     | 0.9604132         |
|          | $EE$            | 71.5676           | 95.5676           | 83.5676     | 0.3672074         |
|          |                 | 72.0000           | 96.0000           | 84.0000     | 0.4413139         |
|          |                 | 72.4324           | 96.4324           | 84.4324     | 0.5937951         |

Ahora, al cambiar el valor objetivo por el valor de la mediana de la muestra en vez del valor central y utilizando la expresión  $x(q_2) - x(q_1)$  teniendo en cuenta la asimetría y la no normalidad encontrada en los datos de los conjuntos I y II al momento de estudiar el efecto de los números

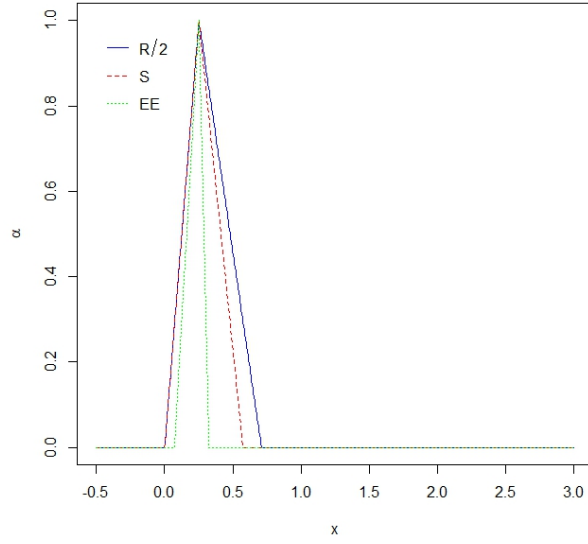


Figura 5.13: Números triangulares  $\mu$  y  $\sigma^2$  en el  $C_{pm}$  fuzzy con los datos del conjunto I.

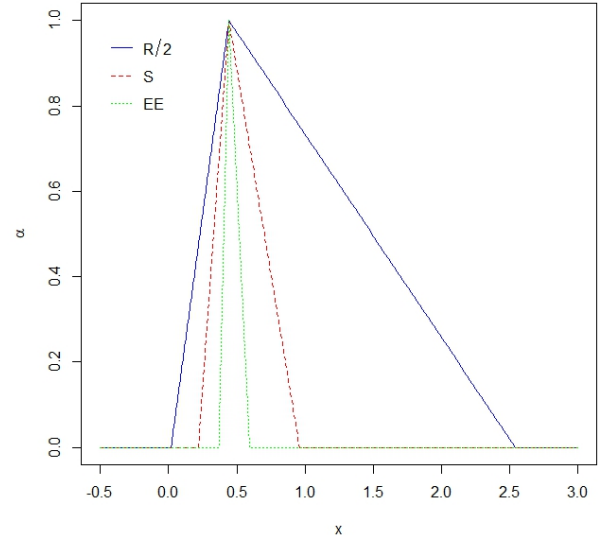


Figura 5.14: Números triangulares  $\mu$  y  $\sigma^2$  en el  $C_{pm}$  fuzzy con los datos del conjunto II.

triangulares  $\mu$  y  $\sigma^2$  en el  $C_{pm}$  fuzzy se obtuvieron los resultados de la tabla 5.17 y Figuras 5.15 y 5.16.

Tabla 5.17: Efecto del cambio de números triangulares de  $\mu$  y  $\sigma^2$  en la estimación del PCI  $C_{pm}$  fuzzy asumiendo no normalidad y asimetría

| Conjunto | Margen de error | $\widetilde{LSL}$ | $\widetilde{USL}$ | $\widetilde{T}$ | $\widetilde{C}_{pmi}$ |
|----------|-----------------|-------------------|-------------------|-----------------|-----------------------|
| I        | $R/2$           | 0.0000            | 0.4109            | 0.0000          | 0.0000000             |
|          |                 | 30.0000           | 96.0000           | 63.0000         | 0.8149135             |
|          |                 | 125.5891          | 191.5890          | 157.5890        | 2.5997895             |
|          | $EE$            | 0.0000            | 53.7102           | 19.7102         | 0.0912589             |
|          |                 | 30.0000           | 96.0000           | 63.0000         | 0.8149135             |
|          |                 | 72.2898           | 138.2898          | 104.2898        | 1.9160413             |
|          | $Sd.log$        | 29.4961           | 95.4961           | 62.4961         | 0.3213250             |
|          |                 | 30.0000           | 96.0000           | 63.0000         | 0.8149135             |
|          |                 | 30.5039           | 96.5039           | 63.5039         | 1.2548051             |
| II       | $R/2$           | 66.3744           | 90.3744           | 77.3744         | 0.0169340             |
|          |                 | 72.0000           | 96.0000           | 84.0000         | 0.4477199             |
|          |                 | 77.6256           | 101.6256          | 88.6256         | 3.2219083             |
|          | $EE$            | 69.8496           | 93.8496           | 81.8496         | 0.3702077             |
|          |                 | 72.0000           | 96.0000           | 84.0000         | 0.4477199             |
|          |                 | 74.1504           | 98.1504           | 86.1504         | 0.6148703             |
|          | $Sd.log$        | 71.9773           | 95.9773           | 83.9773         | 0.4141345             |
|          |                 | 72.0000           | 96.0000           | 84.0000         | 0.4477199             |
|          |                 | 72.0227           | 96.0227           | 84.0227         | 0.5447970             |

A pesar de que el proceso sigue considerándose no capaz se encuentra un aumento en la estimación



del índice de capacidad al reconsiderar la mediana al ser una medida de centralidad más robusta. Cabe resaltar, que para los datos del conjunto I se tiene la mayor diferencia al compararse con los resultados del método anterior, lo que sugiere, que si se consideran distintas formas para la construcción de los parámetros puede aumentarse la capacidad de un proceso a medida que estos sean más explicativos para el proceso.

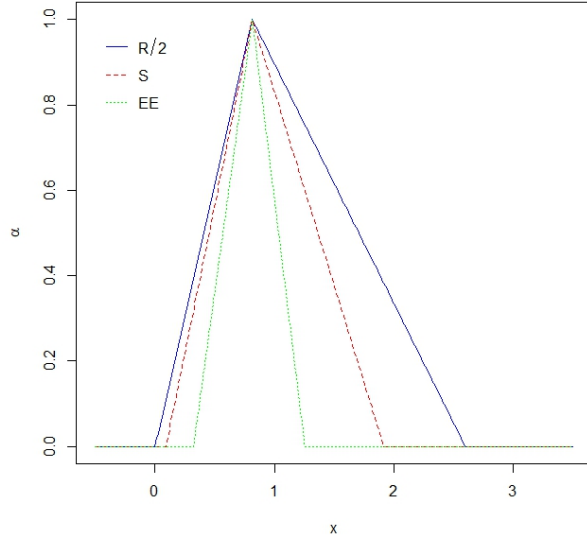


Figura 5.15: Números triangulares  $\mu$  y  $\sigma^2$  en el  $C_{pm}$  fuzzy asumiendo no normalidad y asimetría con los datos del conjunto I.

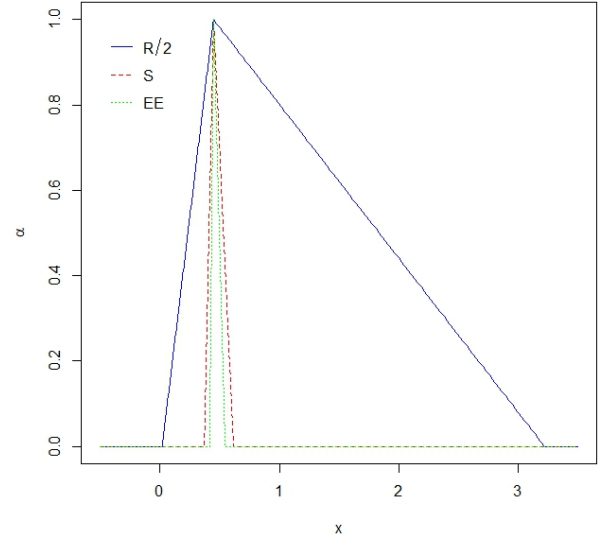


Figura 5.16: Números triangulares  $\mu$  y  $\sigma^2$  en el  $C_{pm}$  fuzzy asumiendo no normalidad y asimetría con los datos del conjunto II.

Para los conjunto de datos I y II, en condiciones normales se tiene que su proceso no está siendo capaz al utilizar el índice  $C_{pm}$  fuzzy, resultados similares a los obtenidos por Leiva, Marchant, Saulo, Aslam y Rojas [34] donde el proceso del conjunto I no es capaz en condiciones normales al utilizar los índices  $C_p$  y  $C_{pk}$ , mientras que para el conjunto II que siendo inadecuada la distribución normal para modelar dicho conjunto el proceso resulta capaz en condiciones normales, pero, cuando consideran la distribución BS el proceso deja de ser capaz, caso contrario con el conjunto I que pasa ser capaz con la distribución adecuada que al asumir su asimetría y no normalidad y optar por otros métodos para la estimación de sus parámetros y a la adecuada selección del margen de error se puede obtener un proceso capaz.

Para estudiar la capacidad de los dos procesos analizados en esta sección, en la Tabla 5.18 se muestran algunos índices de capacidad tradicionales que arrojan como resultado un pobre desempeño y el índice de capacidad  $C_{pm}$  fuzzy para analizar la bondad del procedimiento propuesto para ambos conjuntos de datos analizados.

Teniendo en cuenta los resultados obtenidos anteriormente para el conjunto de datos I se tiene que, asumir la distribución normal de los datos resulta no ser adecuado y que bajo este supuesto el proceso se considera no capaz al tener un desempeño pobre. Sin embargo, tenga en cuenta que al considerar el efecto de los números triangulares, asumiendo la no normalidad y la asimetría del conjunto de datos el proceso de producción de alimentos es capaz de cumplir con los límites de

Tabla 5.18: Índices de capacidad de proceso para el conjunto de datos indicado

| Conjunto | $LSL$ | $T$ | $USL$ | $Mediana$ | $\bar{x}$ | $\hat{C}_p$ | $\hat{C}_{pk}$ | $\hat{C}_{pm}$ | $\hat{C}_{pm}(\alpha)$ |      |      |
|----------|-------|-----|-------|-----------|-----------|-------------|----------------|----------------|------------------------|------|------|
| I        | 30    | 63  | 96    | 68.22     | 80.44     | 0.26        | 0.12           | 0.24           | 1.11                   | 1.26 | 1.33 |
| II       | 72    | 84  | 96    | 92.57     | 93.03     | 1.86        | 0.46           | 0.43           | 0.37                   | 0.44 | 0.59 |

especificación para pacientes con dietas restringidas en la cantidad de proteínas (Tabla 5.18).

Para el segundo conjunto de datos se detecta una situación un poco diferente a la encontrada en el conjunto de datos I. Aquí, aunque considerar la distribución normal es de igual forma un proceso inadecuado para modelar el conjunto II, el proceso de producción de alimentos resulta capaz en condiciones normales y bajo el índice de capacidad  $C_p$  que compara la variación tolerada con la real, situación que cambia cuando se calculan otros índices de capacidad que consideran las desviaciones de la media del proceso hacía los límites de especificación o que considera la variación con respecto al valor objetivo y de igual forma al evaluar la construcción de los números triangulares y los parámetros de la media y varianza, lo que indica una condición de alerta. para la compañía este comportamiento.

## Capítulo 6

# Conclusiones y recomendaciones

### 6.1. Conclusiones

Con base en el análisis realizado y haciendo énfasis en los estudios de simulación y en el estudio de caso aplicativo con los datos reales se puede concluir que:

1. Se adaptó el trabajo de Kaya y Kahraman[19] al estimar el  $C_{pm}$  fuzzy considerando la construcción de los números triangulares  $USL$ ,  $LSL$  y  $T$ , y estudiando el efecto del cambio de los números triangulares de  $\mu$  y  $\sigma^2$  para la estimación del índice de capacidad fuzzy y se encontró que, la selección adecuada de estos parámetros influye en la capacidad o no del  $C_{pm}$  fuzzy estimado.
2. En los resultados observados en las simulaciones se encontró que, a medida que el tamaño de la muestra aumenta, el ancho del intervalo del  $\tilde{C}_{pm}$  fuzzy tiende a tener un comportamiento similar para diferentes márgenes de error, mientras que para tamaños de muestra pequeños a mayor margen de error se obtiene mayor ancho de intervalo en el índice de capacidad estimado.
3. Se encontró que entre más alta sea la asimetría o más carga tenga hacia una cola la distribución de los datos menos adecuado se hace utilizar métodos basados en la normalidad y se recomienda la aproximación a distribuciones empíricas y abordar de forma diferente los parámetros de centralidad y dispersión para que estos resulten más explicativos en cuanto al proceso y el índice de capacidad sea capaz.
4. Para el caso del conjunto I que su mejor ajuste es una distribución *Lognormal* y asumiendo parámetros más robustos se logra aumentar la capacidad del proceso a medida que se reduce el margen de error como puede verse en la Tabla 5.8 y Figura 5.11.
5. Cuando los datos se miden de manera imprecisa, el enfoque tradicional resulta inapropiado para probar el rendimiento del proceso. La metodología de utilizar un enfoque difuso que considera cierto grado de imprecisión en las muestras y a su vez en los límites de especificación y construcción de los parámetros  $\mu$  y  $\sigma^2$  basado en el enfoque de Buckley, permite obtener estimaciones difusas a partir de estimaciones puntuales y de intervalo que brindan más información acerca del proceso.

## 6.2. Recomendaciones y futuras investigaciones

Se presentan la siguiente recomendación a tener en cuenta para futuros trabajos de investigación en esta temática.

1. En estudios vía simulación se considera importante considerar otros tipos de distribuciones como la *Weibull*, *Birnbaum – Saunders*, entre otras que al igual que la *Log – normal* y *Gamma* tengan una gran variedad de formas asimétricas.
2. Considerar un mayor número de distribuciones empíricas para escoger con más exactitud la distribución que se ajuste a los datos y poder aumentar la capacidad de los procesos estudiados aplicando diferentes métodos para escoger los parámetros del índice de capacidad.

# Bibliografía

- [1] A. PARCHAMI, B. SADEGHPOUR-GILDEH, M. N., AND MASHINCHI, M. A new generation of process capability indices based on fuzzy measurements. *Journal of Applied Statistics* 41, 5 (2014), 1122–1136.
- [2] ARNOLD, B. F. An approach to fuzzy hypothesis testing. *Metrika* 44 (1996), 119–126.
- [3] ARNOLD, B. F. Testing fuzzy hypothesis with crisp data. *Fuzzy Sets and Systems* 94, 2 (1998), 323–333.
- [4] BELLMAN, R. E., AND ZADEH, L. A. Decision-making in the fuzzy environment. *Management Science* 17B (1970), 141–164.
- [5] BOYLES, R. A. Process capability with asymmetric tolerance. *Communication in Statistics: Simulation and Computation* 23, 3 (1994), 615–643.
- [6] BUCKLEY, J. Fuzzy statistics: Hypothesis testing, soft computing. *To appear* 172 (2004), 1–40.
- [7] BUCKLEY, J. Fuzzy statistics: Hypothesis testing. *Soft Computing* 9 (2005), 512–518.
- [8] BUCKLEY, J., AND ESLAMI, E. Uncertain probability ii: the continuous case. 193–199.
- [9] BUCKLEY, J. J. Uncertain probability iii: the continuous case. 200–206.
- [10] CASALS, M. R., AND GIL, P. A note on the operativeness of neymann–pearson tests with fuzzy information. *Fuzzy Sets and Systems* 30 (1989), 215–220.
- [11] CASALS, M. R., AND GIL, P. Bayesian sequential test for fuzzy parametric hypotheses from fuzzy information. *Information Sciences* 80 (1994), 283–298.
- [12] CHAO, M. T., AND LIN, D. K. Another look at the process capability index. *Quality and Reliability Engineering International* 22, 2 (2006), 153–163.
- [13] CHIEN, W. W., AND LIAO, M. Y. Estimating and testing process yield with imprecise data. *Expert Systems with Applications* 36 (2009), 11006–11012.
- [14] CLEMENTS, J. Process capability calculations for non-normal distributions. *Qual. Prog.* 22 (1989), 95–100.
- [15] GUTIÉRREZ, H., AND LA VARA, R. D. *Statistical Quality Control and six sigma*, 3 ed. JMc-GrawHill, México, 2013.
- [16] HSIANG, T., AND TAGUCHI, G. Tutorial on quality control and assurance—the taguchi methods. *Joint Meetings of the American Statistical Association, Las Vegas, Nevada* (1985), 188.

- [17] J.M. COGOLLO, M. C., AND FLÓREZ, A. Estimating process capability indices for inaccurate and non-normal data: A systematic literature review. *QUALITY MANAGEMENT* 161 18 (2017).
- [18] KANE, V. Process capability indices. journal of quality technology. *Journal of Quality Technology* 20 18 (1986), 41–52.
- [19] KAYA, I., AND KAHRAMAN, C. Fuzzy process capability indices with asymmetric tolerances. *Expert Systems with Applications* (2011), 14882–14890.
- [20] KAYA, I., AND KAHRAMAN, C. Process capability analyses based on fuzzy measurements and fuzzy control charts. *Expert Syst. Appl.* 38 (2011), 3172–3184.
- [21] KOTZ, S., AND JOHNSON, N. *Process Capability Indices*. Chapman and Hall, New York, 1993.
- [22] LEE, H.  $c_{pk}$  index estimation using fuzzy numbers. *European Journal of Operational Research* 129 (2001), 683–688.
- [23] L.K. CHAN, S. C., AND SPIRING, F. A new measure of process capability:  $c_{pm}$ . *Journal of Quality Technology* 20(3) (1998), 162–175.
- [24] M. A. GIL, N. C., AND GIL, P. The minimum inaccuracy estimates in  $\chi_2$  tests for goodness of fit with fuzzy observations. 95–115.
- [25] M. ABDOLSHAH, R. M. YUSUFF, T. S. H. M. Y. B. I., AND SADIGH, A. N. Measuring process capability index  $c_{pmk}$  with fuzzy data and compare it with other fuzzy process capability indices. *Expert Systems with Applications* 38, 6 (2011), 6452–6457.
- [26] M. R. CASALS, M. A. G., AND GIL, P. The fuzzy decision problem: An approach to the problem of testing statistical hypotheses with fuzzy information. 371–382.
- [27] MARTÍNEZ, J. A. *Theoretical and Practical Foundations of Nutrition and Dietetics*. 1998.
- [28] M.L. ARTEAGA, M. COGOLLO, J. C., AND FLÓREZ, A. Optimal estimation of process capability indices with non-normal data and inaccurate parameters using metaheuristics. *QUALITY MANAGEMENT* 158 18 (2017).
- [29] MONTGOMERY, D. *Introduction to Statistical Quality Control*. John Wiley and Sons, New York, 2001.
- [30] PARCHAMI, A., AND MASHINCHI, M. Fuzzy estimation for process capability indices. *Information Sciences* 177 (2007), 1452–1462.
- [31] PEARN, W., AND CHEN, K. Estimating process capability indices for non-normal pearson populations. *Qual. Reliab. Eng. Int.* 11 (1995), 386–388.
- [32] PEARN, W., AND KOTZ, S. *Encyclopedia and Handbook of Process Capability Indices. Series on Quality, Reliability and Engineering Statistics*, vol. 12. 2006.
- [33] T. W. CHEN, K. C., AND LIN, J. Y. Fuzzy evaluation of process capability for bigger-the-best type products. *International Journal of Advanced Manufacturing Technology*, 21 (2003), 820–826.

- [34] V. LEIVA, C. MARCHANT, H. S. M. A., AND ROJAS, F. Capability indices for birnbaum–saunders processes applied to electronic and food industries. *Journal of Applied Statistics* (2014), 1880–1902.
- [35] VIERTL, R. *Statistical Methods for Non-Precise Data*. CRC Press, Boca Raton, FL, 1996.
- [36] VIERTL, R. *Statistical Methods for Fuzzy Data*. John Wiley & Sons, New York, 2011.
- [37] WATANABE, N., AND IMAIZUMI, T. A fuzzy statistical test of fuzzy hypotheses. *Fuzzy Sets and Systems* 53 (1993), 167–178.
- [38] W.L. PEARN, S. K., AND JOHNSON, N. Distribution and inferential properties of capability indices. *Journal of Quality Technology* 24 (1992), 216–231.
- [39] YONGTING, C. Fuzzy quality and analysis on fuzzy probability. *Fuzzy Sets and Systems* 83 (1996), 283–290.
- [40] ZADEH, L. Toward a generalized theory of uncertainty (gtu)—an outline. *Information Sciences* 172 (2005), 1–40.





## Apéndice A

# Código de simulación en R

En este anexo se transcribe el código en R (R Development Core Team 2012) utilizado para llevar a cabo los cálculos mostrados en el presente trabajo. Los códigos y los archivos están disponibles en mi repositorio ubicado en [github.com](https://github.com/Einerdml-23/Triangular-Fuzzy.git).<sup>1</sup>

```
# Grafico de la funcion de probabilidad para la Gamma

library(ggplot2) # Paquete necesario

xy <- data.frame(x = 0:8, y = seq(0, 1, length.out = 9))
ggplot(xy, aes(x = x, y = y)) +
  stat_function(fun = dgamma, args = list(shape = 2.5, scale = 0.66),
    aes(color = "GA(2.5,0.66)")) +
  stat_function(fun = dgamma, args = list(scale = 0.34, shape = 10),
    aes(color = "GA(10,0.34)")) +
  stat_function(fun = dgamma, args = list(scale = 0.98, shape = 1),
    aes(color = "GA(1.0,0.98)")) +
  theme_minimal() + theme(legend.position = "top",
    text = element_text(size = 11, family = "Cambria"),
    axis.text.x = element_text(angle = 0, vjust = 0.5,
    hjust = 0.5))+
  labs(x = "x", y = quote(italic(f)[GA](x,alpha,beta))) +
  guides(colour = guide_legend(title = "Indicador"))

# Grafico de la funcion de probabilidad para la Lognormal

xy <- data.frame(x = 0:8, y = seq(0, 1, length.out = 9))
ggplot(xy, aes(x = x, y = y)) +
  stat_function(fun = dlnorm, args = list(meanlog = 0.3, sdlog = 3.0),
    aes(color = "LN(0.3,3.0)")) +
  stat_function(fun = dlnorm, args = list(meanlog = 0.5, sdlog = 1.4),
    aes(color = "LN(0.5,1.4)")) +
  stat_function(fun = dlnorm, args = list(meanlog = 0.8, sdlog = 0.5),
    aes(color = "LN(0.8,0.5)")) +
```

---

<sup>1</sup><https://github.com/Einerdml-23/Triangular-Fuzzy.git>

```

theme_minimal() + theme(legend.position = "top",
                        text = element_text(size = 11, family = "Cambria"),
                        axis.text.x = element_text(angle = 0, vjust = 0.5,
                        hjust = 0.5))+
labs(x = "x",y =quote(italic(f)[GA](x,alpha,beta)) )+
guides(colour = guide_legend(title = "Indicador"))

# Efecto de la construccion USL, LSL y T
# Resultados para la distribucion Gamma
# Asimetria baja , alpha=10, beta=0.34
# n=10 , n=30 , n=100
# d=qt(1-0.05/2,n-1)*sqrt(var.est)/sqrt(n) , d=0.1, d=0.01, d=0.001

IC=function(a){
  l=1.6
  t=4.8
  u=8.1

  n=10 # Asignar n
  mu.est=10*0.34
  var.est=10*0.34^2

  d=qt(1-0.05/2,n-1)*sqrt(var.est)/sqrt(n) # Asignar delta

  T.=c(t-d,t,t+d)
  USL=c(u-d,u,u+d)
  LSL=c(l-d,l,l+d)

  dL.a=c((USL[2]-USL[1]-T.[2]+T.[3])*a+USL[1]-T.[3],(USL[2]-USL[3]-
  -T.[2]+T.[1])*a+USL[3]-T.[1])
  dR.a=c((T.[2]-T.[1]-LSL[2]+LSL[3])*a+T.[1]-LSL[3],(T.[2]-T.[3]-
  -LSL[2]+LSL[1])*a+T.[3]-LSL[1])

  d.a=c(min(dL.a[1],dR.a[1]), min(dL.a[2],dR.a[2]))

  TL.a=(T.[2]-T.[1])*a+T.[1]
  TR.a=(T.[2]-T.[3])*a+T.[3]

  CpmL.a=d.a[1]/(3*sqrt(varL.a+(muR.a-TL.a)^2))
  CpmR.a=d.a[2]/(3*sqrt(varR.a+(muL.a-TR.a)^2))

  return(c(CpmL.a, CpmR.a))
}

M=500
alpha=matrix(seq(0,1, length.out = M), ncol=1)

```

```

salida.IC=cbind(alpha,t(apply(alpha,1,IC)))

Cpm.f=c(min(salida.IC[2:M,2]),salida.IC[M,2],max(salida.IC[,3]))
Cpm.f

# Efecto de la construccion USL, LSL y T
# Resultados para la distribucion Gamma
# Asimetria moderada, alpha=2.5, beta=0.66
# n=10 , n=30 , n=100

LSL=c(1-d,1,1+d)

varL.a=((n-1)*var.est)/((1-a)*qchisq(0.005,df=(n-1),lower.tail=T)+(n-1)*a)
varR.a=((n-1)*var.est)/((1-a)*qchisq(0.005,df=(n-1),lower.tail=F)+(n-1)*a)

muL.a=max(mu.est+qnorm(a/2,lower.tail=F)*sqrt(varL.a/n),mu.est
+qnorm(a/2,lower.tail=F)*sqrt(varR.a/n))
muR.a=min(mu.est-qnorm(a/2,lower.tail=F)*sqrt(varL.a/n),mu.est
-qnorm(a/2,lower.tail=F)*sqrt(varR.a/n))

dL.a=c((USL[2]-USL[1]-T.[2]+T.[3])*a+USL[1]-T.[3],(USL[2]-USL[3]-T.[2]+T.[1])*a
+USL[3]-T.[1])
dR.a=c((T.[2]-T.[1]-LSL[2]+LSL[3])*a+T.[1]-LSL[3],(T.[2]-T.[3]-LSL[2]+LSL[1])*a
+T.[3]-LSL[1])

d.a=c(min(dL.a[1],dR.a[1]), min(dL.a[2],dR.a[2]))

TL.a=(T.[2]-T.[1])*a+T.[1]
TR.a=(T.[2]-T.[3])*a+T.[3]

CpmL.a=d.a[1]/(3*sqrt(varL.a+(muR.a-TL.a)^2))

Cpm.f=c(min(salida.IC[2:M,2]),salida.IC[M,2],max(salida.IC[,3]))
Cpm.f

# Efecto de la construccion USL, LSL y T
# Resultados para la distribucion Gamma
# Asimetria alta, alpha=1.0, beta=0.98
# n=10 , n=30 , n=100
# d=qt(1-0.05/2,n-1)*sqrt(var.est)/sqrt(n) , d=0.1, d=0.01, d=0.001

IC=function(a){
  l=1.6
  t=4.8
  u=8.1

  n=10 # Asignar n
  mu.est=1.0*0.98

```

```

var.est=1.0*0.98^2

d=qt(1-0.05/2,n-1)*sqrt(var.est)/sqrt(n) # Asignar delta

T=c(t-d, t, t+d)
USL=c(u-d, u, u+d)
LSL=c(l-d, l, l+d)

muR.a=min(mu.est-qnorm(a/2,lower.tail=F)*sqrt(varL.a/n),mu.est
-qnorm(a/2,lower.tail=F)*sqrt(varR.a/n))

dL.a=c((USL[2]-USL[1]-T.[2]+T.[3])*a+USL[1]-T.[3],(USL[2]-USL[3]-T.[2]+T.[1])*a
+USL[3]-T.[1])
dR.a=c((T.[2]-T.[1]-LSL[2]+LSL[3])*a+T.[1]-LSL[3],(T.[2]-T.[3]-LSL[2]+LSL[1])*
a+T.[3]-LSL[1])

d.a=c(min(dL.a[1],dR.a[1]), min(dL.a[2],dR.a[2]))

TL.a=(T.[2]-T.[1])*a+T.[1]
TR.a=(T.[2]-T.[3])*a+T.[3]

CpmL.a=d.a[1]/(3*sqrt(varL.a+(muR.a-TL.a)^2))
CpmR.a=d.a[2]/(3*sqrt(varR.a+(muL.a-TR.a)^2))

return(c(CpmL.a, CpmR.a))
}

# n=10 , n=30 , n=100
# d=qt(1-0.05/2,n-1)*sqrt(var.est)/sqrt(n) , d=0.1, d=0.01, d=0.001

set.seed(121115)
datos=rgamma(n = 10, shape = 10, scale = 0.34)

mediana=function(a){
  l=1.6
  t=4.8
  u=8.1

  n=10
  mu.est=10*0.34
  var.est=10*0.34^2

  d=qt(1-0.05/2,n-1)*sqrt(var.est)/sqrt(n)

  T=c(t-d, t, t+d)
  USL=c(u-d, u, u+d)

```

```

LSL=c(1-d, l, l+d)

varL.a=((n-1)*var.est)/((1-a)*qchisq(0.005, df=(n-1), lower.tail=F)+(n-1)*a)
varR.a=((n-1)*var.est)/((1-a)*qchisq(0.995, df=(n-1), lower.tail=F)+(n-1)*a)

+USL[3]-T.[1] )
dR.a=c((T.[2]-T.[1]-LSL[2]+LSL[3])*a+T.[1]-LSL[3], (T.[2]-T.[3]-LSL[2]+LSL[1])*a
+T.[3]-LSL[1] )

d.a=c(min(dL.a[1], dR.a[1]), min(dL.a[2], dR.a[2]))

TL.a=(T.[2]-T.[1])*a+T.[1]
TR.a=(T.[2]-T.[3])*a+T.[3]

CmpL.a=d.a[1]/(3*sqrt(varL.a+(muR.a-TL.a)^2))
CmpR.a=d.a[2]/(3*sqrt(varR.a+(muL.a-TR.a)^2))

return(c(CmpL.a, CmpR.a))
}

M=500
alpha=matrix(seq(0.0000000000001, 0.999999999, length.out = M), ncol=1)
salida.IC2=cbind(alpha, matrix(NA, ncol=2, nrow=M))

for (i in 1:M){
  salida.IC2[i, 2:3]=mediana(salida.IC2[i, 1])
}

Cpm.f2=c(min(salida.IC2[2:M, 2]), salida.IC2[M, 2], max(salida.IC2[, 3]))
Cpm.f2

# Efecto de la construccion de mu y sigma
# Resultados para la distribucion Gamma
# Asimetria moderada, alpha=2.5, beta=0.66
# n=10 , n=30 , n=100
# d=qt(1-0.05/2, n-1)*sqrt(var.est)/sqrt(n) , d=0.1, d=0.01, d=0.001

set.seed(120540)
datos=rgamma(n = 10, shape = 2.5, scale = 0.66)

varL.a=((n-1)*var.est)/((1-a)*qchisq(0.005, df=(n-1), lower.tail=F)+(n-1)*a)
varR.a=((n-1)*var.est)/((1-a)*qchisq(0.995, df=(n-1), lower.tail=F)+(n-1)*a)

#usando IC para la mediana (Prueba de Wilconxon para 1 muestra)
muL.a=wilcox.test(datos, conf.int = TRUE, conf.level = 1-a)$conf.int[2]
muR.a=wilcox.test(datos, conf.int = TRUE, conf.level = 1-a)$conf.int[1]

```

```

dL.a=c((USL[2]-USL[1]-T.[2]+T.[3])*a+USL[1]-T.[3],(USL[2]-USL[3]-T.[2]+T.[1])*a
+USL[3]-T.[1])
dR.a=c((T.[2]-T.[1]-LSL[2]+LSL[3])*a+T.[1]-LSL[3],(T.[2]-T.[3]-LSL[2]+LSL[1])*a
+T.[3]-LSL[1])

d.a=c(min(dL.a[1],dR.a[1]), min(dL.a[2],dR.a[2]))

TL.a=(T.[2]-T.[1])*a+T.[1]
TR.a=(T.[2]-T.[3])*a+T.[3]

CmpL.a=d.a[1]/(3*sqrt(varL.a+(muR.a-TL.a)^2))
CmpR.a=d.a[2]/(3*sqrt(varR.a+(muL.a-TR.a)^2))

return(c(CmpL.a, CmpR.a))
}

M=500
alpha=matrix(seq(0.0000000000001,0.999999999, length.out = M), ncol=1)
salida.IC2=cbind(alpha, matrix(NA, ncol=2, nrow=M))

for (i in 1:M){
  salida.IC2[i,2:3]=mediana(salida.IC2[i,1])
}

Cpm.f2=c(min(salida.IC2[2:M,2]), salida.IC2[M,2], max(salida.IC2[,3]))
Cpm.f2

mediana=function(a){
  l=1.6
  t=4.8
  u=8.1

  n=10
  mu.est=1.0*0.98
  var.est=1.0*0.98^2

  d=qt(1-0.05/2,n-1)*sqrt(var.est)/sqrt(n)

  T.=c(t-d, t, t+d)
  USL=c(u-d, u, u+d)
  LSL=c(l-d, l, l+d)

  dR.a=c((T.[2]-T.[1]-LSL[2]+LSL[3])*a+T.[1]-LSL[3],(T.[2]-T.[3]-LSL[2]+LSL[1])*a
+T.[3]-LSL[1])

  d.a=c(min(dL.a[1],dR.a[1]), min(dL.a[2],dR.a[2]))

```

```

TL.a=(T.[2]-T.[1])*a+T.[1]
TR.a=(T.[2]-T.[3])*a+T.[3]

CmpL.a=d.a[1]/(3*sqrt(varL.a+(muR.a-TL.a)^2))
CmpR.a=d.a[2]/(3*sqrt(varR.a+(muL.a-TR.a)^2))

return(c(CmpL.a, CmpR.a))
}

M=500
alpha=matrix(seq(0.000000000001,0.999999999, length.out = M), ncol=1)
salida.IC2=cbind(alpha, matrix(NA, ncol=2, nrow=M))

for (i in 1:M){
  salida.IC2[i,2:3]=mediana(salida.IC2[i,1])
}

# n=10 , n=30 , n=100
# d=qt(1-0.05/2,n-1)*sqrt(var.est)/sqrt(n) , d=0.1, d=0.01, d=0.001

IC=function(a){
  l=1.6
  t=4.8
  u=8.1

  n=10 # Asignar n
  mu.est=0.8
  var.est=0.5^2

  d=qt(1-0.05/2,n-1)*sqrt(var.est)/sqrt(n) # Asignar delta

  T.=c(t-d, t, t+d)
  USL=c(u-d, u, u+d)
  LSL=c(l-d, l, l+d)

  varL.a=((n-1)*var.est)/((1-a)*qchisq(0.005, df=(n-1), lower.tail=T)+(n-1)*a)
  varR.a=((n-1)*var.est)/((1-a)*qchisq(0.005, df=(n-1), lower.tail=F)+(n-1)*a)

  muL.a=max(mu.est+qnorm(a/2, lower.tail=F)*sqrt(varL.a/n), mu.est
+qnorm(a/2, lower.tail=F)*sqrt(varR.a/n))
  muR.a=min(mu.est-qnorm(a/2, lower.tail=F)*sqrt(varL.a/n), mu.est
-qnorm(a/2, lower.tail=F)*sqrt(varR.a/n))

  dL.a=c((USL[2]-USL[1]-T.[2]+T.[3])*a+USL[1]-T.[3], (USL[2]-USL[3]-T.[2]+T.[1])*a
+USL[3]-T.[1])

  TL.a=(T.[2]-T.[1])*a+T.[1]

```

```

TR.a=(T.[2]-T.[3])*a+T.[3]

CpmL.a=d.a[1]/(3*sqrt(varL.a+(muR.a-TL.a)^2))
CpmR.a=d.a[2]/(3*sqrt(varR.a+(muL.a-TR.a)^2))

return(c(CpmL.a, CpmR.a))
}

M=500
alpha=matrix(seq(0,1, length.out = M), ncol=1)
salida.IC=cbind(alpha, t(apply(alpha,1,IC)))

Cpm.f=c(min(salida.IC[2:M,2]), salida.IC[M,2], max(salida.IC[,3]))
Cpm.f

# Efecto de la construccion USL, LSL y T
# Resultados para la distribucion Lognormal
# Asimetria moderada, alpha=0.5, beta=1.4
# n=10 , n=30 , n=100
# d=qt(1-0.05/2,n-1)*sqrt(var.est)/sqrt(n) , d=0.1, d=0.01, d=0.001

IC=function(a){
  l=1.6
  t=4.8
  u=8.1

  n=10          # Asignar n
  mu.est=0.5
  var.est=1.4^2

  d=qt(1-0.05/2,n-1)*sqrt(var.est)/sqrt(n) # Asignar delta

  T.=c(t-d, t, t+d)
  USL=c(u-d, u, u+d)
  LSL=c(l-d, l, l+d)

  muR.a=min(mu.est-qnorm(a/2, lower.tail=F)*sqrt(varL.a/n), mu.est
  -qnorm(a/2, lower.tail=F)*sqrt(varR.a/n))

  dL.a=c((USL[2]-USL[1]-T.[2]+T.[3])*a+USL[1]-T.[3], (USL[2]-USL[3]-T.[2]+T.[1])*a
  +USL[3]-T.[1])
  dR.a=c((T.[2]-T.[1]-LSL[2]+LSL[3])*a+T.[1]-LSL[3], (T.[2]-T.[3]-LSL[2]+LSL[1])*a
  +T.[3]-LSL[1])

  d.a=c(min(dL.a[1], dR.a[1]), min(dL.a[2], dR.a[2]))

  TL.a=(T.[2]-T.[1])*a+T.[1]
  TR.a=(T.[2]-T.[3])*a+T.[3]

```



```

CpmL.a=d.a[1]/(3*sqrt(varL.a+(muR.a-TL.a)^2))
CpmR.a=d.a[2]/(3*sqrt(varR.a+(muL.a-TR.a)^2))

return(c(CpmL.a, CpmR.a))
}

# Efecto de la construccion USL, LSL y T
# Resultados para la distribucion Lognormal
# Asimetria alta , alpha=0.3, beta=3.0
# n=10 , n=30 , n=100
# d=qt(1-0.05/2,n-1)*sqrt(var.est)/sqrt(n) , d=0.1, d=0.01, d=0.001

IC=function(a){
  l=1.6
  t=4.8
  u=8.1

  n=10          # Asignar n
  mu.est=0.3
  var.est=3.0^2

  d=qt(1-0.05/2,n-1)*sqrt(var.est)/sqrt(n) # Asignar delta

  T=c(t-d, t, t+d)
  USL=c(u-d, u, u+d)
  LSL=c(l-d, l, l+d)

  varL.a=((n-1)*var.est)/((1-a)*qchisq(0.005,df=(n-1),lower.tail=T)+(n-1)*a)
  varR.a=((n-1)*var.est)/((1-a)*qchisq(0.005,df=(n-1),lower.tail=F)+(n-1)*a)

  muL.a=max(mu.est+qnorm(a/2,lower.tail=F)*sqrt(varL.a/n),mu.est
+qnorm(a/2,lower.tail=F)*sqrt(varR.a/n))
  muR.a=min(mu.est-qnorm(a/2,lower.tail=F)*sqrt(varL.a/n),mu.est
-qnorm(a/2,lower.tail=F)*sqrt(varR.a/n))

  CpmL.a=d.a[1]/(3*sqrt(varL.a+(muR.a-TL.a)^2))
  CpmR.a=d.a[2]/(3*sqrt(varR.a+(muL.a-TR.a)^2))

  return(c(CpmL.a, CpmR.a))
}

M=500
alpha=matrix(seq(0,1, length.out = M), ncol=1)
salida.IC=cbind(alpha,t(apply(alpha,1,IC)))

Cpm.f=c(min(salida.IC[2:M,2]),salida.IC[M,2],max(salida.IC[,3]))

```

Cpm.f

```
# Efecto de la construccion de mu y sigma
# Resultados para la distribucion Lognormal
# Asimetria baja , alpha=0.8, beta=0.5
# n=10 , n=30 , n=100
# d=qt(1-0.05/2,n-1)*sqrt(var.est)/sqrt(n) , d=0.1, d=0.01, d=0.001

set.seed(112585)
datos=rlnorm(n, meanlog = 0.8, sdlog = 0.5)

mediana=function(a){
  l=1.6
  t=4.8
  u=8.1

  n=10
  mu.est=0.8
  var.est=0.5^2

  #usando IC para la mediana (Prueba de Wilconxon para 1 muestra)
  muL.a=wilcox.test(datos, conf.int = TRUE, conf.level = 1-a)$conf.int[2]
  muR.a=wilcox.test(datos, conf.int = TRUE, conf.level = 1-a)$conf.int[1]

  dL.a=c((USL[2]-USL[1]-T.[2]+T.[3])*a+USL[1]-T.[3],(USL[2]-USL[3]-T.[2]+T.[1])*a
  +USL[3]-T.[1] )
  dR.a=c((T.[2]-T.[1]-LSL[2]+LSL[3])*a+T.[1]-LSL[3],(T.[2]-T.[3]-LSL[2]+LSL[1])*a
  +T.[3]-LSL[1] )

  d.a=c(min(dL.a[1],dR.a[1]), min(dL.a[2],dR.a[2]))

  TL.a=(T.[2]-T.[1])*a+T.[1]
  TR.a=(T.[2]-T.[3])*a+T.[3]

  CmpL.a=d.a[1]/(3*sqrt(varL.a+(muR.a-TL.a)^2))
  CmpR.a=d.a[2]/(3*sqrt(varR.a+(muL.a-TR.a)^2))

  return(c(CmpL.a, CmpR.a))
}

M=500
alpha=matrix(seq(0.000000000001,0.999999999, length.out = M), ncol=1)
salida.IC2=cbind(alpha, matrix(NA, ncol=2, nrow=M))

for (i in 1:M){
  salida.IC2[i,2:3]=mediana(salida.IC2[i,1])
}
```

```

# Efecto de la construccion de mu y sigma
# Resultados para la distribucion Lognormal
# Asimetria moderada, alpha=0.5, beta=1.4
# n=10 , n=30 , n=100
# d=qt(1-0.05/2,n-1)*sqrt(var.est)/sqrt(n) , d=0.1, d=0.01, d=0.001

set.seed(10121111)
datos=rlnorm(n, meanlog = 0.5, sdlog = 1.4)

mediana=function(a){
  l=1.6
  t=4.8
  u=8.1

  n=10
  mu.est=0.5
  var.est=1.4^2

  d=qt(1-0.05/2,n-1)*sqrt(var.est)/sqrt(n)

  T.=c(t-d, t, t+d)
  USL=c(u-d, u, u+d)
  LSL=c(l-d, l, l+d)

  dL.a=c((USL[2]-USL[1]-T.[2]+T.[3])*a+USL[1]-T.[3], (USL[2]-USL[3]-T.[2]+T.[1])*a
  +USL[3]-T.[1])
  dR.a=c((T.[2]-T.[1]-LSL[2]+LSL[3])*a+T.[1]-LSL[3], (T.[2]-T.[3]-LSL[2]+LSL[1])*a
  +T.[3]-LSL[1])

  d.a=c(min(dL.a[1], dR.a[1]), min(dL.a[2], dR.a[2]))

  TL.a=(T.[2]-T.[1])*a+T.[1]
  TR.a=(T.[2]-T.[3])*a+T.[3]

  CmpL.a=d.a[1]/(3*sqrt(varL.a+(muR.a-TL.a)^2))
  CmpR.a=d.a[2]/(3*sqrt(varR.a+(muL.a-TR.a)^2))

  return(c(CmpL.a, CmpR.a))
}

M=500
alpha=matrix(seq(0.000000000001,0.999999999, length.out = M), ncol=1)
salida.IC2=cbind(alpha, matrix(NA, ncol=2, nrow=M))

for (i in 1:M){
  salida.IC2[i,2:3]=mediana(salida.IC2[i,1])
}

```

```

Cpm.f2=c(min(salida.IC2[2:M,2]), salida.IC2[M,2], max(salida.IC2[,3]))
Cpm.f2

# Efecto de la construccion de mu y sigma
# Resultados para la distribucion Lognormal
# Asimetria alta , alpha=0.3, beta=3.0
# n=10 , n=30 , n=100
# d=qt(1-0.05/2,n-1)*sqrt(var.est)/sqrt(n) , d=0.1, d=0.01, d=0.001

set.seed(111811)
datos=rlnorm(n, meanlog = 0.3, sdlog = 3)

mediana=function(a){

  varL.a=((n-1)*var.est)/((1-a)*qchisq(0.005, df=(n-1), lower.tail=F)+(n-1)*a)
  varR.a=((n-1)*var.est)/((1-a)*qchisq(0.995, df=(n-1), lower.tail=F)+(n-1)*a)

  #usando IC para la mediana (Prueba de Wilconxon para 1 muestra)
  muL.a=wilcox.test(datos, conf.int = TRUE, conf.level = 1-a)$conf.int[2]
  muR.a=wilcox.test(datos, conf.int = TRUE, conf.level = 1-a)$conf.int[1]

  dL.a=c((USL[2]-USL[1]-T.[2]+T.[3])*a+USL[1]-T.[3], (USL[2]-USL[3]-T.[2]+T.[1])*a
  +USL[3]-T.[1])
  dR.a=c((T.[2]-T.[1]-LSL[2]+LSL[3])*a+T.[1]-LSL[3], (T.[2]-T.[3]-LSL[2]+LSL[1])*a
  +T.[3]-LSL[1])

  d.a=c(min(dL.a[1], dR.a[1]), min(dL.a[2], dR.a[2]))

  TL.a=(T.[2]-T.[1])*a+T.[1]
  TR.a=(T.[2]-T.[3])*a+T.[3]

  CmpL.a=d.a[1]/(3*sqrt(varL.a+(muR.a-TL.a)^2))
  CmpR.a=d.a[2]/(3*sqrt(varR.a+(muL.a-TR.a)^2))

  return(c(CmpL.a, CmpR.a))
}

M=500
alpha=matrix(seq(0.0000000000001, 0.9999999999, length.out = M), ncol=1)
salida.IC2=cbind(alpha, matrix(NA, ncol=2, nrow=M))

for (i in 1:M){
  salida.IC2[i,2:3]=mediana(salida.IC2[i,1])
}

Cpm.f2=c(min(salida.IC2[2:M,2]), salida.IC2[M,2], max(salida.IC2[,3]))
Cpm.f2

```

```

# Capitulo 5: Aplicacion

# Paquetes requeridos
library("MASS")
library("survival")
library("fitdistrplus")
require(car)
require(goftest)

# Conjunto 1
c1=c(66.09, 146.48, 97.99, 56.77, 74.23, 76.20, 108.90, 89.51, 62.40, 141.15,
      89.81, 201.05, 210.32, 105.08, 169.82, 64.25, 100.43, 59.56, 49.97, 102.16,
      164.05, 59.63, 51.41, 93.32, 39.43, 102.91, 124.91, 56.16, 70.94, 77.26,
      59.01, 49.21, 123.85, 59.78, 163.50, 37.67, 88.27, 35.60, 42.54, 43.87,
      44.41, 25.55, 41.12, 88.14, 54.69, 85.32, 67.55, 17.76, 83.73, 34.07,
      42.37, 61.95, 68.62, 41.69, 100.43, 139.39, 60.03, 34.35, 51.28, 68.22,
      80.44)

# Conjunto 2
c2=c(90.98, 95.41, 91.87, 94.01, 91.59, 91.42, 92.45, 93.36, 91.75, 91.97,
      93.73, 92.57, 94.51, 93.86, 91.57, 91.61, 94.28, 98.12, 90.18, 92.45,
      91.02, 96.42, 91.75, 92.15, 90.55, 93.38, 99.97, 90.71, 89.88, 92.05,
      100.52, 94.12, 91.78, 91.6, 90.4, 94.16, 97.76, 90.3, 93.3, 93.4,
      95.47, 91.39, 91.11, 93.59, 92.27, 92.65, 90.59, 94.03, 96.34, 91.64,
      90.67, 93.3, 91.32, 93.13, 92.08, 89.1, 93.04, 90.1, 90.59, 93.34,
      90.12, 90.51, 91.12, 92.13, 93.37, 92.16, 92.46, 91.72, 97.22, 92.68,
      96.54, 92.09, 95.18, 93.45, 91.89, 94.54, 93.97, 92.44, 93.56, 93.81,
      92.53, 95.8, 93.54, 94.78, 96.31, 94.48, 92.17, 92.24, 93.35, 92.84,
      93.53, 97.53, 92.95, 93.67, 92.46)

# Estudio de normalidad Conjunto I

# Histograma, densidad y cuantiles teoricos vs muestrales

par(mfrow=c(1,3))
hist(c1, xlab = "Cantidad de prote na", ylab = "Frecuencia", las=1,
     main = "", col = "gray")
plot(density(c1), xlab = "Cantidad de prote na", ylab = "Densidad",
     las=1, main = "")
qqPlot(c1, xlab="Cuantiles te ricos", ylab="Cuantiles muestrales",
     las=1,main="")

# Parametros estimados de la distribucion hipotetica
Ajustenc1<-fitdistr(c1, "normal") # c1
Ajustenc1

# Prueba de bondad de ajuste

```

```

# Distribucion acumulada empirica y teorica de una normal con el conjunto I

xenc1 <- seq(min(c1), max(c1), by=0.0001)
plot(xenc1, pnorm(xenc1, mean=Ajustenc1$estimate[1], sd=Ajustenc1$estimate[2]),
     type="l", col="red", xlab="Cantidad de prote na – Conjunto I",
     ylab="pnorm(Prote na, mean, sd)")
plot(ecdf(c1), add=TRUE)

# Estudio de normalidad Conjunto II

# Histograma, densidad y cuantiles teoricos vs muestrales

par(mfrow=c(1,3))
hist(c2, xlab = "Cantidad de prote na", ylab = "Frecuencia", las=1,
main = "", col = "gray")
plot(density(c2), xlab = "Cantidad de prote na", ylab = "Densidad",
las=1, main = "")
qqPlot(c2, xlab="Cuantiles te ricos", ylab="Cuantiles muestrales",
las=1,main="")

# Parametros estimados de la distribucion hipotetica
Ajustenc2<-fitdistr(c2, "normal") # c2
Ajustenc2

# Prueba de bondad de ajuste

# Distribucion acumulada empirica y teorica de una normal con el conjunto II

xenc2 <- seq(min(c2), max(c2), by=0.0001)
plot(xenc2, pnorm(xenc2, mean=Ajustenc2$estimate[1],
sd=Ajustenc2$estimate[2]),
     type="l", col="red", xlab="Cantidad de prote na – Conjunto II",
     ylab="pnorm(Prote na, mean, sd)")
plot(ecdf(c2), add=TRUE)

# Kurtosis y Skewness para ajustar los datos

# Ajuste a distribuciones – Conjunto I

fw<-fitdist(c1, "weibull")
fg<-fitdist(c1, "gamma")
fln<-fitdist(c1, "lnorm")
par(mfrow=c(2,2))
plot.legend<-c("Weibull","gamma","lognormal")
denscomp(list(fw,fg,fln), legendtext=plot.legend)

```

```

# Criterios

gofstat(list(fw, fg, fln), fitnames = c("Weibull", "gamma", "lognormal"))

# Incertidumbre modelos

bendo.fln <- bootdist(fln, niter = 5000)
summary(bendo.fln) # Bootstrap medianas
fln # parametros maxima verosimilitud

# Ajuste a distribuciones – Conjunto II

fwc2<-fitdist(c2, "weibull")
fgc2<-fitdist(c2, "gamma")
flnc2<-fitdist(c2, "lnorm")
par(mfrow=c(2,2))
plot.legend<-c("Weibull", "gamma", "lognormal")

# Criterios

gofstat(list(fwc2, fgc2, flnc2), fitnames = c("Weibull", "gamma", "lognormal"))

# Incertidumbre modelos

bendo.flnc2 <- bootdist(flnc2, niter = 5000)
summary(bendo.flnc2) # Bootstrap medianas
flnc2 # parametros maxima verosimilitud

# Capacidad del proceso Cpm

# Efecto de la construccion de los numeros triangulares
# Asumiendo parametros de la media y varianza para el conjunto I
quantile(c1, c(0.00135, 0.99865))
R=209.56913-18.39099
# d0=R/2 ; d0=sd(c1) ; d0=qnorm(1-0.05/2)*sd(c1)/sqrt(length(c1))

IC=function(a){
  l=30
  t=63
  u=96

  n=61 ; mu.est=mean(c1) ; var.est=var(c1)

  d=d0 # Este es el valor a modificar

  T=c(t-d, t, t+d)
  USL=c(u-d, u, u+d)

```

```

LSL=c(1-d, l, l+d)

dL.a=c((USL[2]-USL[1]-T.[2]+T.[3])*a+USL[1]-T.[3], (USL[2]-USL[3]-T.[2]+T.[1])*a
+USL[3]-T.[1])
dR.a=c((T.[2]-T.[1]-LSL[2]+LSL[3])*a+T.[1]-LSL[3], (T.[2]-T.[3]-LSL[2]+LSL[1])*a
+T.[3]-LSL[1])

d.a=c(min(dL.a[1], dR.a[1]), min(dL.a[2], dR.a[2]))

TL.a=(T.[2]-T.[1])*a+T.[1]
TR.a=(T.[2]-T.[3])*a+T.[3]

CpmL.a=d.a[1]/(3*sqrt(varL.a+(muR.a-TL.a)^2))
CpmR.a=d.a[2]/(3*sqrt(varR.a+(muL.a-TR.a)^2))

return(c(CpmL.a, CpmR.a))
}

Cpm.f=c(min(salida.IC[2:M,2]), salida.IC[M,2], max(salida.IC[,3]))
Cpm.f

# Efecto de la construccion de los numeros triangulares
# Asumiendo parametros de la media y varianza para el conjunto II
quantile(c2, c(0.00135, 0.99865))
R=100.45021-89.19898
# d0=R/2 ; d0=sd(c2) ; d0=qnorm(1-0.05/2)*sd(c2)/sqrt(length(c2))

IC=function(a){
  l=72
  t=84
  u=96

  n=95 ; mu.est=mean(c2) ; var.est=var(c2)

  d=d0 # Este es el valor a modificar

  varL.a=((n-1)*var.est)/((1-a)*qchisq(0.005, df=(n-1), lower.tail=T)+(n-1)*a)
  varR.a=((n-1)*var.est)/((1-a)*qchisq(0.005, df=(n-1), lower.tail=F)+(n-1)*a)

  muL.a=max(mu.est+qnorm(a/2, lower.tail=F)*sqrt(varL.a/n), mu.est
+qnorm(a/2, lower.tail=F)*sqrt(varR.a/n))
  muR.a=min(mu.est-qnorm(a/2, lower.tail=F)*sqrt(varL.a/n), mu.est
-qnorm(a/2, lower.tail=F)*sqrt(varR.a/n))

  dL.a=c((USL[2]-USL[1]-T.[2]+T.[3])*a+USL[1]-T.[3], (USL[2]-USL[3]-T.[2]+T.[1])*a
+USL[3]-T.[1])
  dR.a=c((T.[2]-T.[1]-LSL[2]+LSL[3])*a+T.[1]-LSL[3], (T.[2]-T.[3]-LSL[2]+LSL[1])*a
+T.[3]-LSL[1])

```



```

d.a=c(min(dL.a[1],dR.a[1]), min(dL.a[2],dR.a[2]))

TL.a=(T.[2]-T.[1])*a+T.[1]
}

M=500
alpha=matrix(seq(0,1, length.out = M), ncol=1)
salida.IC=cbind(alpha,t(apply(alpha,1,IC)))

Cpm.f=c(min(salida.IC[2:M,2]),salida.IC[M,2],max(salida.IC[,3]))
Cpm.f

# Efecto de la construccion de los numeros triangulares
# Asumiendo parametros de la mediana y rango intercuartil q3-q1
# Conjunto 1
summary(c1)
Est.var=100.43-51.28
quantile(c2,c(0.00135,0.99865))
R=209.56913-18.39099 ; Sd.log=0.50389
# d0=R/2 ; d0=Sd.log ; d0=qnorm(1-0.05/2)*sd(c2)/sqrt(length(c2))

IC=function(a){

  LSL=c(1-d,1,1+d)

  varL.a=((n-1)*var.est)/((1-a)*qchisq(0.005,df=(n-1),lower.tail=T)+(n-1)*a)
  varR.a=((n-1)*var.est)/((1-a)*qchisq(0.005,df=(n-1),lower.tail=F)+(n-1)*a)

  muL.a=max(mu.est+qnorm(a/2, lower.tail = F)*sqrt(varL.a/n),mu.est
+qnorm(a/2, lower.tail = F)*sqrt(varR.a/n))
  muR.a=min(mu.est-qnorm(a/2, lower.tail = F)*sqrt(varL.a/n),mu.est
-qnorm(a/2, lower.tail = F)*sqrt(varR.a/n))

  dL.a=c((USL[2]-USL[1]-T.[2]+T.[3])*a+USL[1]-T.[3],(USL[2]-USL[3]-T.[2]+T.[1])*a
+USL[3]-T.[1])

  d.a=c(min(dL.a[1],dR.a[1]), min(dL.a[2],dR.a[2]))

  TL.a=(T.[2]-T.[1])*a+T.[1]
  TR.a=(T.[2]-T.[3])*a+T.[3]

  CpmL.a=d.a[1]/(3*sqrt(varL.a+(muR.a-TL.a)^2))
  CpmR.a=d.a[2]/(3*sqrt(varR.a+(muL.a-TR.a)^2))

  return(c(CpmL.a, CpmR.a))
}

```

```

M=500
alpha=matrix(seq(0,1, length.out = M), ncol=1)
salida.IC=cbind(alpha,t(apply(alpha,1,IC)))

summary(c2)
Est.var=93.915-91.625
quantile(c2,c(0.00135,0.99865))
R=100.45021-89.19898 ; Sd.log=0.02272196
# d0=R/2 ; d0=Sd.log ; d0=qnorm(1-0.05/2)*sd(c2)/sqrt(length(c2))

IC=function(a){
  l=72
  t=84
  u=96

  n=95 ; mu.est=median(c2) ; var.est=Est.var

  d=d0 # Este es el valor a modificar

  T=c(t-d,t,t+d)
  USL=c(u-d,u,u+d)
  LSL=c(l-d,l,l+d)

  varL.a=((n-1)*var.est)/((1-a)*qchisq(0.005,df=(n-1),lower.tail=T)+(n-1)*a)
  varR.a=((n-1)*var.est)/((1-a)*qchisq(0.005,df=(n-1),lower.tail=F)+(n-1)*a)

  muL.a=max(mu.est+qnorm(a/2, lower.tail = F)*sqrt(varL.a/n),mu.est
+qnorm(a/2, lower.tail = F)*sqrt(varR.a/n))
  muR.a=min(mu.est-qnorm(a/2, lower.tail = F)*sqrt(varL.a/n),mu.est
-qnorm(a/2, lower.tail = F)*sqrt(varR.a/n))

  dL.a=c((USL[2]-USL[1]-T.[2]+T.[3])*a+USL[1]-T.[3],(USL[2]-USL[3]-T.[2]+T.[1])*a
+USL[3]-T.[1])
  dR.a=c((T.[2]-T.[1]-LSL[2]+LSL[3])*a+T.[1]-LSL[3],(T.[2]-T.[3]-LSL[2]+LSL[1])*a
+T.[3]-LSL[1])

  d.a=c(min(dL.a[1],dR.a[1]), min(dL.a[2],dR.a[2]))

  TL.a=(T.[2]-T.[1])*a+T.[1]
  TR.a=(T.[2]-T.[3])*a+T.[3]

M=500

# Efecto de la construccion de mu y sigma
# Asumiendo parametros de la media y varianza para el conjunto I
quantile(c1,c(0.00135,0.99865))

```

```

R=209.56913-18.39099
# d0=R/2 ; d0=sd(c1) ; d0=qnorm(1-0.05/2)*sd(c1)/sqrt(length(c1))
datos=c1

mediana=function(a){
  l=30
  t=63
  u=96

  n=61 ; mu.est=mean(c1) ; var.est=var(c1)
  d=d0 # Este es el valor a modificar

  T.=c(t-d, t, t+d)
  USL=c(u-d, u, u+d)
  LSL=c(l-d, l, l+d)

  varL.a=((n-1)*var.est)/((1-a)*qchisq(0.005, df=(n-1), lower.tail=F)+(n-1)*a)
  varR.a=((n-1)*var.est)/((1-a)*qchisq(0.995, df=(n-1), lower.tail=F)+(n-1)*a)

  dR.a=c((T.[2]-T.[1]-LSL[2]+LSL[3])*a+T.[1]-LSL[3], (T.[2]-T.[3]-LSL[2]+LSL[1])*a
  +T.[3]-LSL[1])

  d.a=c(min(dL.a[1], dR.a[1]), min(dL.a[2], dR.a[2]))

  TL.a=(T.[2]-T.[1])*a+T.[1]
  TR.a=(T.[2]-T.[3])*a+T.[3]

  CmpL.a=d.a[1]/(3*sqrt(varL.a+(muR.a-TL.a)^2))
  CmpR.a=d.a[2]/(3*sqrt(varR.a+(muL.a-TR.a)^2))

  return(c(CmpL.a, CmpR.a))
}

M=500
alpha=matrix(seq(0.000000000001, 0.999999999, length.out = M), ncol=1)
salida.IC2=cbind(alpha, matrix(NA, ncol=2, nrow=M))

for (i in 1:M){
  salida.IC2[i, 2:3]=mediana(salida.IC2[i, 1])
}

# Efecto de la construccion de mu y sigma
# Asumiendo parametros de la media y varianza para el conjunto II
quantile(c2, c(0.00135, 0.99865))
R=100.45021-89.19898
# d0=R/2 ; d0=sd(c2) ; d0=qnorm(1-0.05/2)*sd(c2)/sqrt(length(c2))
datos=c2

```

```

mediana=function(a){
  l=72
  t=84
  u=96

  n=95 ; mu.est=mean(c2) ; var.est=var(c2)

  d=d0 # Este es el valor a modificar

  T=c(t-d,t,t+d)
  USL=c(u-d,u,u+d)
  LSL=c(l-d,l,l+d)

  varL.a=((n-1)*var.est)/((1-a)*qchisq(0.005,df=(n-1),lower.tail=F)+(n-1)*a)
  varR.a=((n-1)*var.est)/((1-a)*qchisq(0.995,df=(n-1),lower.tail=F)+(n-1)*a)

  d.a=c(min(dL.a[1],dR.a[1]), min(dL.a[2],dR.a[2]))

  TL.a=(T.[2]-T.[1])*a+T.[1]
  TR.a=(T.[2]-T.[3])*a+T.[3]

  CmpL.a=d.a[1]/(3*sqrt(varL.a+(muR.a-TL.a)^2))
  CmpR.a=d.a[2]/(3*sqrt(varR.a+(muL.a-TR.a)^2))

  return(c(CmpL.a, CmpR.a))
}

M=500
alpha=matrix(seq(0.0000000000001,0.999999999, length.out = M), ncol=1)
salida.IC2=cbind(alpha,matrix(NA, ncol=2, nrow=M))

for (i in 1:M){
  salida.IC2[i,2:3]=mediana(salida.IC2[i,1])
}

Cpm.f2=c(min(salida.IC2[2:M,2]), salida.IC2[M,2],max(salida.IC2[,3]))
Cpm.f2

# Efecto de la construccion de mu y sigma
# Asumiendo parametros de la mediana y rango intercuartil q3-q1
# Conjunto 1
summary(c1)
Est.var=100.43-51.28
quantile(c2,c(0.00135,0.99865))

```

```

mediana=function(a){
  l=30
  t=63
  u=96

  n=61 ; mu.est=median(c1) ; var.est=Est.var

  d=d0 # Este es el valor a modificar

  T.=c(t-d,t,t+d)
  USL=c(u-d,u,u+d)
  LSL=c(l-d,l,l+d)

  dL.a=c((USL[2]-USL[1]-T.[2]+T.[3])*a+USL[1]-T.[3],(USL[2]-USL[3]-T.[2]+T.[1])*a
+USL[3]-T.[1])
  dR.a=c((T.[2]-T.[1]-LSL[2]+LSL[3])*a+T.[1]-LSL[3],(T.[2]-T.[3]-LSL[2]+LSL[1])*a
+T.[3]-LSL[1])

  d.a=c(min(dL.a[1],dR.a[1]), min(dL.a[2],dR.a[2]))

  TL.a=(T.[2]-T.[1])*a+T.[1]
  TR.a=(T.[2]-T.[3])*a+T.[3]

  CmpL.a=d.a[1]/(3*sqrt(varL.a+(muR.a-TL.a)^2))
  CmpR.a=d.a[2]/(3*sqrt(varR.a+(muL.a-TR.a)^2))

  return(c(CmpL.a, CmpR.a))
}

M=500
alpha=matrix(seq(0.0000000000001,0.999999999, length.out = M), ncol=1)
salida.IC2=cbind(alpha,matrix(NA, ncol=2, nrow=M))

for (i in 1:M){
  salida.IC2[i,2:3]=mediana(salida.IC2[i,1])
}

# Efecto de la construccion de mu y sigma
# Asumiendo parametros de la mediana y rango intercuartil q3-q1
# Conjunto II
summary(c2)
Est.var=93.915-91.625
quantile(c2,c(0.00135,0.99865))
R=100.45021-89.19898 ; Sd.log=0.02272196
# d0=R/2 ; d0=Sd.log ; d0=qnorm(1-0.05/2)*sd(c2)/sqrt(length(c2))
datos=c2

```

```

mediana=function(a){
  l=72
  t=84
  u=96

  n=95 ; mu.est=median(c2) ; var.est=Est.var

  d=d0 # Este es el valor a modificar

  T=c(t-d,t,t+d)
  USL=c(u-d,u,u+d)
  LSL=c(l-d,l,l+d)

  dL.a=c((USL[2]-USL[1]-T.[2]+T.[3])*a+USL[1]-T.[3],(USL[2]-USL[3]-T.[2]+T.[1])*a
+USL[3]-T.[1])
  dR.a=c((T.[2]-T.[1]-LSL[2]+LSL[3])*a+T.[1]-LSL[3],(T.[2]-T.[3]-LSL[2]+LSL[1])*a
+T.[3]-LSL[1])

  d.a=c(min(dL.a[1],dR.a[1]), min(dL.a[2],dR.a[2]))

  TL.a=(T.[2]-T.[1])*a+T.[1]
  TR.a=(T.[2]-T.[3])*a+T.[3]

  CmpL.a=d.a[1]/(3*sqrt(varL.a+(muR.a-TL.a)^2))
  CmpR.a=d.a[2]/(3*sqrt(varR.a+(muL.a-TR.a)^2))

  return(c(CmpL.a, CmpR.a))
}

M=500
alpha=matrix(seq(0.000000000001,0.999999999, length.out = M), ncol=1)
salida.IC2=cbind(alpha,matrix(NA, ncol=2, nrow=M))

for (i in 1:M){
  salida.IC2[i,2:3]=mediana(salida.IC2[i,1])
}

Cpm.f2=c(min(salida.IC2[2:M,2]),salida.IC2[M,2],max(salida.IC2[,3]))
Cpm.f2

```