



COLEGIO DE POSTGRADUADOS
INSTITUCIÓN DE ENSEÑANZA E INVESTIGACIÓN EN CIENCIAS AGRÍCOLAS

CAMPUS MONTECILLO
POSTGRADO DE SOCIOECONOMÍA, ESTADÍSTICA E INFORMÁTICA
ESTADÍSTICA

**LÍMITES ÓPTIMOS PARA UNA MUESTRA AGRUPADA
NORMALMENTE DISTRIBUIDA CON MEDIA Y VARIANZA
DESCONOCIDAS.**

IVÁN MAURICIO PÉREZ LAGUNA

T E S I S

PRESENTADA COMO REQUISITO PARCIAL
PARA OBTENER EL GRADO DE:

MAESTRO EN CIENCIAS

MONTECILLO, TEXCOCO, EDO. DE MÉXICO

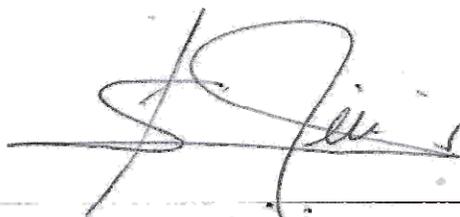
2008

La presente tesis titulada: **Límites Óptimos para una Muestra Agrupada Normalmente Distribuida con Media y Varianza Desconocidas.**, realizada por el alumno: **Iván Mauricio Pérez Laguna**, bajo la dirección del consejo particular indicado, ha sido aprobada por el mismo y aceptada como requisito parcial para obtener el grado de

MAESTRO EN CIENCIAS
SOCIOECONOMÍA, ESTADÍSTICA E INFORMÁTICA
ESTADÍSTICA

CONSEJO PARTICULAR

CONSEJERO



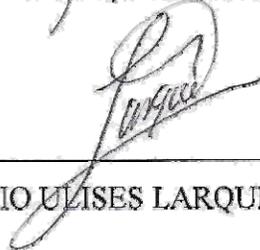
DR. GILBERTO RENDÓN SÁNCHEZ

ASESOR



DR. HUMBERTO VAQUERA HUERTA

ASESOR



M. en C. MARIO ULISES LARQUÉ SAAVEDRA

Montecillo, Texcoco, Estado de México, Febrero de 2008

LÍMITES ÓPTIMOS PARA UNA MUESTRA AGRUPADA NORMALMENTE
DISTRIBUIDA CON MEDIA Y VARIANZA DESCONOCIDAS.

Pérez Laguna Iván Mauricio, M. en C.

Colegio de Postgraduados, 2007

RESUMEN

El agrupamiento de datos es una metodología común cuando no es posible coleccionar la información real en las unidades de muestreo ya sea por que no se dispone de instrumentos que así lo permitan o bien, por que se esta ante el riesgo de coleccionarla con errores no debidos al muestreo. En la práctica es usual inferir sobre los parámetros trabajando a estas muestras como si no estuvieran agrupadas, sobre todo por que las expresiones de los estimadores obtenidos son poco amigables o difíciles de manejar encontrando solución únicamente con la ayuda de métodos numéricos. Este tratamiento, conduce a errores de estimación en los parámetros de interés. En el presente trabajo, se obtienen las expresiones para los Estimadores de Máxima Verosimilitud, se les compara contra un ejemplo y se calculan los límites de truncamiento óptimos cuando la media y la desviación estándar de una muestra normalmente distribuida son desconocidas, siguiendo para este cometido el enfoque propuesto por Kulldorff (1961). Los resultados indican que el agrupamiento es mejor cuando el número de límites de truncamiento es mayor, sin embargo, esto ocurre solo hasta cierto número. Los límites de truncamiento son simétricos cuando el número de ellos es impar. Los Estimadores de Máxima Verosimilitud para la Media y la Desviación Estándar de la distribución resultaron muy buenos estimadores.

Palabras Clave: Agrupamiento, Información Relativa Asintótica, Límites de Truncamiento, Máxima Verosimilitud, Varianza Asintótica.

OPTIMAL LIMITS FOR A GROUPED SAMPLE NORMALLY DISTRIBUTED WITH
MEAN AND VARIANCE UNKNOWN.

Pérez Laguna Iván Mauricio, M. en C.

Colegio de Postgraduados, 2007

ABSTRACT

The grouping of data is a common methodology when it is not possible to collect the real information in the sampling units whatever we don't have the instruments to allow it or, so because we are in risk of collecting it with errors non due to the sampling. It is usual in practice to infer on the parameters working to these samples as if they were not grouped, overcoat because the expressions of the obtained estimators are little friendly or difficult to handle solely finding solution with the aid of numerical methods. This treatment, leads to errors of estimation in the interest parameters. In the present work, the expressions for the Maximum Likelihood Estimators are obtained, they are compared against an example and the truncation optimal limits are calculated when the mean and variance of a sample normally distributed are unknown, using for this task the approach proposed by Kulldorff (1961). The results indicate that the group is better when the number of truncation limits is greater, nevertheless, this happens single until certain number of them. The truncation limits are symmetric when the number of them is uneven. The Maximum Likelihood Estimators for the Mean and the Standard Deviation of the distribution are very good.

Key Words: Grouping, Relative Asymptotic Information, Truncation Limits, Maximum Likelihood Estimator, Asymptotic Variance.

A mi Abuelo

...Hijo mío en la vida debes de ser Torero

Ignacio Pérez Gutiérrez †

AGRADECIMIENTOS

Al Consejo Nacional de Ciencia y Tecnología por el financiamiento dado para mis estudios de maestría.

Al Colegio de Postgraduados Campus Montecillo y al Instituto de Socioeconomía Estadística e Informática, por la oportunidad de cursar estudios de maestría.

Al M. en C. Mario U. Larqué Saavedra por su atinada guía, confianza, amistad y apoyo incondicional. *Gracias por sembrar en mí la semilla que ahora da frutos.*

Al Dr. Gilberto Rendón Sánchez, por darme todo su apoyo y brindarme todas las facilidades. *Gracias por mostrarme el lado humano de la ciencia.*

Al Dr. Humberto Vaquera Huerta, por la confianza depositada en mí. *Siempre será un defensor de los estudiantes.*

Al Lic. en Matemáticas Luís Colorado Martínez. *Sin tu ayuda hubiera sido imposible terminar mis estudios de maestría. Tu amistad y sabiduría vivirán por siempre en mí espíritu.*

A la Coordinación de Servicios de Información de la UAM – A, muy en especial al Lic. Gerardo Muñoz Colunga y a la Lic. Marisol Popoca Burgoa, por las facilidades y la fe brindadas.

A mi familia: Mis padres María Eugenia Laguna Zamora y Alberto Pérez Galicia *por todo su amor, apoyo e impulso incondicionales. Mis abuelas, Teresa Galicia y Fidencia Zamora, por iniciar la dinastía de la que provengo. Y a mi novia Gabriela Rodríguez, gracias por todo tu amor, apoyo y fe.*

A todos mis compañeros, en especial a Lucia Carmona, Rigoberto Sifuentes, Araceli Martínez, Marcelo Quiterio y Martha Aguilar, gracias por dos años de apoyo, amistad y compañía.

CONTENIDO

1. INTRODUCCIÓN GENERAL	1
1.1. Objetivo General.....	5
1.1.1. Objetivos Particulares.....	6
2. REVISIÓN DE LITERATURA	7
3. PLANTEAMIENTO DEL PROBLEMA	9
3.1. Notación y Definiciones	9
3.2. Estimación.....	11
3.3. Obtención de los Estimadores, las Varianzas y Covarianza Asintóticas.....	18
3.3.1. Función de Densidad Normal.....	18
3.3.1.1. Estimadores de Máxima Verosimilitud.....	19
3.3.1.1.1. Para la Media.....	20
3.3.1.1.2. Para la Desviación Estándar.....	21
3.3.1.2. Varianzas y Covarianza Asintóticas.....	23
3.3.1.2.1. Varianza Asintótica de la Media.....	24
3.3.1.2.2. Varianza Asintótica de la Desviación Estándar.....	26
3.3.1.2.3. Covarianza Asintitótica.....	30

4. METODOLOGÍA	35
4.1. Agrupamiento Óptimo.....	36
4.2. Estimadores de Máxima Verosimilitud.....	36
4.3. Lenguaje de Programación utilizado	41
5. RESULTADOS Y DISCUSIÓN	42
5.1. Límites de Truncamiento Óptimo.....	42
5.2. Estimación Simultánea de la Media y la Desviación Estándar.....	44
5.2.1. Muestra de tamaño $n = 50$	45
5.2.2. Muestra de tamaño $n = 100$	52
5.2.3. Muestra de tamaño $n = 200$	59
5.2.4. Muestra de tamaño $n = 400$	66
6. CONCLUSIONES	73
7. LITERATURA CITADA	75
8. ANEXOS	77
I. Varianzas Asintóticas.....	77
II. Estimadores de Máxima Verosimilitud.....	80

ÍNDICE DE CUADROS Y FIGURAS

CUADROS

Cuadro 1. Límites de Truncamiento Asintóticamente Óptimos y la Información Relativa Asintótica.....	42
Cuadro 2. Agrupamientos Inicial y Óptimo para una muestra de tamaño cincuenta obtenida de una población normal con parámetros $\mu = 50$ y $\sigma = 10$	47
Cuadro 3. Semillas, Estimadores Iniciales y Óptimos y el indicador DC para una muestra de tamaño cincuenta obtenidas de una población normal con parámetros $\mu = 50$ y $\sigma = 10$	49
Cuadro 4. Información Relativa Asintótica y Eficiencia del Muestreo para una muestra de tamaño cincuenta obtenida de una población normal con parámetros $\mu = 50$ y $\sigma = 10$	49
Cuadro 5. Agrupamientos Inicial y Óptimo para una muestra de tamaño cien obtenida de una población normal con parámetros $\mu = 50$ y $\sigma = 10$	54
Cuadro 6. Semillas, Estimadores Iniciales y Óptimos y el indicador DC para una muestra de tamaño cien obtenidas de una población normal con parámetros $\mu = 50$ y $\sigma = 10$	56

Cuadro 7. Información Relativa Asintótica y Eficiencia del Muestreo para una muestra de tamaño cien obtenida de una población normal con parámetros $\mu = 50$ y $\sigma = 10$	56
Cuadro 8. Agrupamientos Inicial y Óptimo para una muestra de tamaño doscientos obtenida de una población normal con parámetros $\mu = 50$ y $\sigma = 10$	61
Cuadro 9. Semillas, Estimadores Iniciales y Óptimos y el indicador DC para una muestra de tamaño doscientos obtenidas de una población normal con parámetros $\mu = 50$ y $\sigma = 10$	63
Cuadro 10. Información Relativa Asintótica y Eficiencia del Muestreo para una muestra de tamaño doscientos obtenida de una población normal con parámetros $\mu = 50$ y $\sigma = 10$	63
Cuadro 11. Agrupamientos Inicial y Óptimo para una muestra de tamaño cuatrocientos obtenida de una población normal con parámetros $\mu = 50$ y $\sigma = 10$	68
Cuadro 12. Semillas, Estimadores Iniciales y Óptimos y el indicador DC para una muestra de tamaño cuatrocientos obtenidas de una población normal con parámetros $\mu = 50$ y $\sigma = 10$	70
Cuadro 13. Información Relativa Asintótica y Eficiencia del Muestreo para una muestra de tamaño cuatrocientos obtenida de una población normal con parámetros $\mu = 50$ y $\sigma = 10$	70

FIGURAS

- Figura 1.** Función de Verosimilitud para función de densidad de probabilidad normal para una muestra de tamaño cincuenta y parámetros $\mu = 50$, $\sigma = 10$ 45
- Figura 2.** *Agrupamiento Tradicional y Agrupamiento Óptimo* para una muestra de tamaño cincuenta obtenida de una población normal con parámetros $\mu = 50$ y $\sigma = 10$ 46
- Figura 3.** Índices de la Eficiencia del Diseño de Muestreo para una muestra de tamaño cincuenta generada con distribución normal y parámetros $\mu = 50$, $\sigma = 10$ 50
- Figura 4.** Función de Verosimilitud para función de densidad de probabilidad normal para una muestra de tamaño cien y parámetros $\mu = 50$, $\sigma = 10$ 52
- Figura 5.** *Agrupamiento Tradicional y Agrupamiento Óptimo* para una muestra de tamaño cien obtenida de una población normal con parámetros $\mu = 50$ y $\sigma = 10$ 53
- Figura 6.** Índices de la Eficiencia del Diseño de Muestreo para una muestra de tamaño cien generada con distribución normal y parámetros $\mu = 50$, $\sigma = 10$ 57
- Figura 7.** Función de Verosimilitud para función de densidad de probabilidad normal para una muestra de tamaño doscientos y parámetros $\mu = 50$, $\sigma = 10$ 59

Figura 8. <i>Agrupamiento Tradicional y Agrupamiento Óptimo</i> para una muestra de tamaño doscientos obtenida de una población normal con parámetros $\mu = 50$ y $\sigma = 10$	60
Figura 9. Índices de la Eficiencia del Diseño de Muestreo para una muestra de tamaño doscientos generada con distribución normal y parámetros $\mu = 50$, $\sigma = 10$	64
Figura 10. Función de Verosimilitud para función de densidad de probabilidad normal para una muestra de tamaño cuatrocientos y parámetros $\mu = 50$, $\sigma = 10$	66
Figura 11. <i>Agrupamiento Tradicional y Agrupamiento Óptimo</i> para una muestra de tamaño cuatrocientos obtenida de una población normal con parámetros $\mu = 50$ y $\sigma = 10$	67
Figura 12. Índices de la Eficiencia del Diseño de Muestreo para una muestra de tamaño cuatrocientos generada con distribución normal y parámetros $\mu = 50$, $\sigma = 10$	71

1. INTRODUCCIÓN GENERAL

En la metodología estadística estándar se trabaja con unidades de muestreo cuyos valores observados presentan el estado real¹ de las mismas, de modo que al tomarse una muestra, dígase de tamaño n , se tendrán registrados n valores observados, permitiéndonos dar paso al análisis estadístico a que haya lugar.

Un problema que se tiene en ciertos estudios de muestreo y por ende en la estimación de los parámetros de interés, es cuando no se puede coleccionar el estado real de las unidades que conforman la muestra, obligando al investigador a trazar estrategias que le permitan coleccionar información acerca del estado real de la unidad de muestreo. Este tipo de situaciones induce a lo que se conoce como errores no debidos al muestreo, para el cual autores como Kish, L. (1975), Sukhatme, P. V., & Sukhatme, B. V. (1970), Raj, D. (1971), Cochran, W. G. (1977), dedican capítulos en sus textos al tratamiento de este problema. Una solución alterna a la propuesta en estos textos es coleccionar la información dentro de intervalos, de forma tal que puedan agruparse los valores observados dentro de dichos intervalos. Considere los siguientes ejemplos que ilustran las situaciones prácticas donde el agrupamiento puede presentarse.

¹ Entiéndase por estado real al valor resultante de medir en la unidad de muestreo la característica de interés, valor que corresponde a unidades de medida (peso, dimensión, presión, etc.).

a) En estudios humanos existen problemas en la colección de información cuando, debido al tipo de preguntas que deben hacerse al entrevistado, siente alguna invasión a su intimidad, por ejemplo, cuando se desea hacer un estudio del ingreso familiar, al preguntársele al entrevistado cuántos salarios mínimos devenga por día – invasión a su intimidad – será casi seguro obtener una respuesta no precisa, falsa o simplemente, no respuesta; sin embargo, a la pregunta en cuál de los intervalos siguientes se localiza su ingreso diario, es más probable obtener una respuesta correcta o más precisa por parte del entrevistado. Es claro que al final del trabajo de coleccionar información, ésta se tendrá en intervalos. Note que un caso análogo se presenta cuando se desea conocer el número de bebidas alcohólicas o cigarrillos que el entrevistado consume.

b) Suponga que un investigador realiza un experimento en conejos para probar los efectos de un nuevo suero. El investigador dosifica y aplica el suero a un número n de conejos. Suponga ahora que el investigador descubre al final del experimento que sus instrumentos fallaron en la precisión de los minutos registrando tiempos incorrectos, dando como resultado el registro de valores poco precisos en la bitácora de la investigación. Debido a los costos del suero el investigador no puede repetir el experimento, sin embargo, suponga que gracias a la experiencia del investigador, éste puede agrupar a los conejos dentro de intervalos – o grupos – de acuerdo al tiempo que tardaron en reaccionar – dígame de una hora cada uno –, es decir, contará el número de conejos que tardaron en reaccionar por ejemplo, menos de una hora, el número que tardaron más de una hora pero menos de dos, y así sucesivamente. Note que de aquí en adelante el número de conejos en cada intervalo junto con el número de intervalos de agrupamiento será toda la información disponible para el investigador.

c) En trabajos de reforestación, es común que se mida el número de árboles sobrevivientes a un período para declarar establecida la plantación. Suponga ahora que uno de los objetivos del estudio es estimar el tiempo a la marchites permanente de las plántulas sembradas y que no sobrevivirán al período. Es común que tales plantaciones se realicen en lugares de difícil acceso y, consecuentemente, medir *in situ* la marchites permanente es prácticamente imposible. La alternativa podría ser dividir el período de establecimiento en subperíodos en los cuales mediante recorridos en la plantación se determine el número de plantas muertas que hay en cada subperíodo. Así por ejemplo, si el período es de cinco años, este podría dividirse en cinco subperíodos de un año cada uno y, en cada uno de éstos, realizar un recorrido en la plantación para contar el número de plantas o árboles muertos, cuidando de que una planta o árbol muerto no sea contado en más de un recorrido.

Los casos expuestos arriba son un claro ejemplo de lo que en estadística se conoce como *Muestras Agrupadas*, o simplemente como *Agrupamiento*. Dicho agrupamiento acarrea una pérdida de información cuando se desean estimar los parámetros además de que las expresiones de los estimadores carecen de una forma cerrada – no son explícitas – requiriendo para su solución la utilización de métodos numéricos. Estas dos complicaciones pueden ser confrontadas contra la ventaja de poder reducir el riesgo de obtener información falsa o no precisa – como en los ejemplos arriba citados – y de la simplicidad de su aplicación a la hora de coleccionar y manejar la información deseada.

En forma específica, Kulldorff (1961) señala que en la práctica es usual tratar a estas muestras como si no estuvieran agrupadas; es decir, a las observaciones que se encuentran dentro de un

intervalo se les asignará el valor medio del correspondiente intervalo. Este tratamiento conduce a problemas en la estimación sobre todo si la muestra presenta alguna de las siguientes características:

- (i) El número de grupos es pequeño, *i.e.*, el agrupamiento es burdo.
- (ii) Los intervalos de agrupamiento son de diferente longitud.
- (iii) La función de densidad no hace contacto con el eje de las x 's en uno o ambos límites de la distribución.

El problema de agrupar; sobre todo en la práctica, ha llevado a plantear ciertas interrogantes de interés:

- ¿Cuántos intervalos de agrupamiento se deben elegir para que la información disponible de la muestra sea la mayor posible?
- ¿Cuáles deben ser los límites del agrupamiento tales que la información de la muestra sea la mayor posible?

Dichas interrogantes limitan el uso del método de agrupamiento, y han tenido diferentes soluciones bajo diferentes enfoques. Kulldorff (1961) con base en el trabajo de Hughes (1949), determinó los estimadores de máxima verosimilitud para los parámetros de localización y escala, además del agrupamiento óptimo, para una muestra que ha sido extraída de una población que se distribuye exponencial. Kulldorff (1961) derivó los mismos resultados cuando la muestra se

extrae de una población que se distribuye normal. Cabe citar que no realizó la estimación simultánea de los parámetros de localización y escala. Hartley (1963), en una revisión al trabajo de Kulldorff (1961), comenta que el agrupamiento manejado por dicho autor es un caso especial de lo que en la teoría moderna se conoce como *Muestras Censuradas del Tipo II*. Finalmente, Ogawa (1977), bajo un enfoque diferente al de Kulldorff, también encontró el agrupamiento óptimo para el caso normal, sin realizar la estimación simultánea para la media y la desviación estándar.

El siguiente trabajo de tesis, en su primera fase se avoca a la tarea de encontrar los estimadores, en forma simultánea, para la media y la desviación estándar de una distribución normal. Parte seguida, mediante un estudio numérico se determinan los límites óptimos de los intervalos de agrupamiento. Finalmente se realizan comparaciones entre los estimadores de máxima verosimilitud obtenidos a partir de los límites de truncamiento óptimos contra los estimadores de máxima verosimilitud iniciales, utilizando el procedimiento presentado por Kulldorff (1961), extendiendo sus resultados al caso cuando la media y la desviación estándar son desconocidas. Expuesto lo anterior es posible plantear lo siguiente

1.1. Objetivo General

Determinar por simulación los límites de truncamiento para una muestra agrupada obtenida a partir de una población con función de densidad normal, y con ello estimar la media y la desviación estándar.

1.1.1. Objetivos Particulares

- Obtención de los Límites de Truncamiento Óptimo cuando los parámetros de localización y escala son desconocidos.
- Estimación simultánea de los parámetros de localización y escala.
- Determinar la eficiencia de los estimadores obtenidos.
- Comparar los límites de agrupamiento óptimos contra los simulados, vía la información relativa asintótica.

2. REVISIÓN DE LITERATURA

En la mayoría de los estudios, se parte del hecho que la muestra obtenida de la población consta de observaciones individuales conocidas. Sin embargo, cuando la naturaleza de ciertos estudios así lo solicita, agrupar la muestra puede acarrear ventajas como las ya expuestas en la parte introductoria.

Hughes (1949) encontró el agrupamiento asintótico óptimo cuando la muestra se extrae de una población distribuida de forma exponencial, al plantear la función $g(x) = x(1 - e^{-x})$ y su función inversa $h(z) = g^{-1}(x)$, presentando los puntos de truncamiento en función del parámetro desconocido θ perteneciente a dicha distribución – el cual también se desea estimar –.

Kulldorff (1961), partiendo de las ideas de Dalenius (1959) para Estratificación Óptima; Wald (1948) Asintóticamente Eficiente – Varianza Asintótica del Estimador – y; los Estimadores de Máxima Verosimilitud, presentan resultados para muestras agrupadas de los casos siguientes:

- i. Una muestra que ha sido extraída de una población que se distribuye Exponencial, obteniendo el estimador de máxima verosimilitud para el parámetro desconocido θ y las varianzas asintóticas cuyos límites se encuentran en función de dicho parámetro. De igual forma presenta correcciones numéricas al trabajo de Hughes (1949). Adicionalmente,

presenta el caso de estimación de la media cuando la distribución exponencial ha sido reparametrizada.

- ii. Los estimadores de máxima verosimilitud para la Media y la Desviación Estándar de una muestra que ha sido extraída de una población Normal, así como el agrupamiento asintóticamente óptimo, cuyos límites se encuentran en función de los parámetros mencionados para cada uno de los siguientes casos:
 - a. Cuando la Desviación Estándar es conocida y la Media desconocida, y
 - b. Cuando la Media es conocida y la Desviación Estándar desconocida.

Hartley (1963), en su revisión al trabajo de Kulldorff (1961), comenta que el agrupamiento manejado por dicho autor es un caso especial de lo que en la teoría moderna se conoce como *Muestras Censuradas del Tipo II* – no se hará mayor mención en este trabajo acerca de esta conexión entre agrupamiento y censurado –.

Ogawa (1977) utilizó los Cuantiles Muestrales para obtener los Mejores Estimadores Asíntoticamente Insesgados de la media μ y la varianza σ^2 , cuando la muestra ha sido extraída de una población normal, siendo el agrupamiento óptimo aquel que maximiza la eficiencia relativa asintótica – el determinante de la matriz de información de Fisher –, mostrando que dicho agrupamiento es simétrico.

3. PLANTEAMIENTO DEL PROBLEMA

En esta sección se definirán los conceptos y notaciones que ayudarán a establecer en forma analítica la metodología para la solución del problema.

3.1. Notación y Definiciones

DEFINICIONES: Sea X la característica en estudio, con función de distribución $F_X(x, \theta)$; donde θ es un vector de parámetros y denote con \mathbb{R} al dominio de X .

Suponga que atendiendo a las necesidades del estudio, \mathbb{R} es seccionado en m intervalos, con $m \geq 2$. Denote con h_1, h_2, \dots, h_{m-1} los puntos donde se realizan los cortes para generar las secciones, y h_0 y h_m son el límite inferior y superior de la distribución siendo tales que $h_i = x_i$, y $h_{i-1} < h_i$ ($i = 1, 2, \dots, m$); mismos que se disponen de la manera siguiente, para obtener:

$$(h_0, h_1], (h_1, h_2], \dots, (h_{m-2}, h_{m-1}], (h_{m-1}, h_m),$$

con las propiedades: $(h_{i-1}, h_i] \cap (h_i, h_{i+1}] = \phi$, $(h_{i-1}, h_i] \cap (h_{m-1}, h_m] = \phi$

y $\bigcup_{i=1}^{m-1} (h_{i-1}, h_i) \cup (h_{m-1}, h_m) = \mathbb{R} (i = 1, 2, \dots, m-1)$; es decir, los intervalos forman una partición del dominio de X .

DEFINICIÓN 3.1. Los valores $h_i (i = 1, 2, \dots, m)$ así establecidos se llamarán LÍMITES DE TRUNCAMIENTO

DEFINICIÓN 3.2. Cada par ordenado de los puntos de truncamiento $(h_{i-1}, h_i]$ y (h_{m+1}, h_m) se define a ser el i – ésimo INTERVALO. Donde h_0 es el límite inferior de la distribución tal que $F_X(h_0; \theta) = 0$ y análogamente h_m es el límite superior de la distribución tal que $F_X(h_m; \theta) = 1$.

Por otra parte, suponga que se activarán las condiciones experimentales n veces o bien, que se extraerá una muestra de tamaño n y que, en cada una de ellas la variable a observar en la unidad en estudio es $X_j (j = 1, 2, \dots, n)$; pero que debido a las condiciones del estudio, tal variable no podrá ser medido – estado real – pero se tendrá la capacidad de asignar su varlor a uno de los intervalos o clases utilizando el criterio siguiente:

$$Y_{ij} = \begin{cases} 1 & \text{si } x_j \in (h_{i-1}, h_i] \quad (i = 1, 2, \dots, m); \quad (j = 1, 2, \dots, n) \\ 0 & \text{de otra forma} \end{cases} \quad (3.0)$$

Es claro que Y_{ij} es una variable aleatoria y que una vez de realizados los n ensayos se tendrá que:

$$n_i = \sum_{j=1}^n y_{ij} \quad (i = 1, 2, \dots, m)$$

es el número de unidades que hay en el i – *ésimo* intervalo y además $n = \sum_{i=1}^m n_i$

DEFINICIÓN 3.3. Se dice que la muestra esta agrupada, simplemente cuando una vez dividido el dominio de la variable en estudio en m intervalos, se determina el número de unidades en estudio que pertenecen a cada uno de ellos.

3.2. Estimación

Suponga contar con una muestra agrupada de la población $F_X(x, \varrho)$ y que el objetivo es estimar al menos uno de los parámetros inmersos en ϱ . Debido a sus muchas bondades, el método más frecuentemente utilizado es el de la Máxima Verosimilitud, propuesto por Fisher.

Al considerar (3.0) es claro que:

$$P(Y_{ij} = 1) = P(h_{i-1} < X_j \leq h_i) = F_X(h_i; \varrho) - F_X(h_{i-1}; \varrho) = \pi_i \quad (3.1)$$

es la probabilidad de que la j – *ésima* unidad caiga en el i – *ésimo* intervalo, siendo $\pi_i \geq 0$ con

$\sum_{i=1}^m \pi_i = 1$, dado el condicionamiento hecho para la definición de los intervalos.

Teniendo n observaciones independientes las cuales han sido agrupadas acorde a (3.0) y usando (3.1), la función de verosimilitud total es:

$$L(\underline{\theta}, n_i) = \pi_1^{n_1} \pi_2^{n_2} \cdots \pi_m^{n_m} = \prod_{i=1}^m \pi_i^{n_i} \quad (3.2)$$

con n_i el número de unidades en el i –ésimo intervalo. Es común aplicar el logaritmo natural a la función de verosimilitud la cual se conoce como función *log – verosimilitud*, de esta forma (3.2) se convierte en:

$$L(\underline{\theta}) = \ln L(\underline{\theta}; n_i) = \sum_{i=1}^m n_i \ln \pi_i \quad (3.3)$$

Si $\hat{\underline{\theta}} = \hat{g}(h_i)$ es el valor que maximiza a $L(\underline{\theta})$, entonces $\hat{\underline{\theta}}$ es el Estimador de Máxima Verosimilitud – MLE – de $\underline{\theta}$. Si algunas condiciones de regularidad se cumplen – las cuales se definirán mas abajo –, el punto donde la verosimilitud tiene su máximo es la solución a las p ecuaciones:

$$\hat{\underline{\theta}} = MLE(\underline{\theta}) = \frac{\partial \ln L(\underline{\theta}; n_i)}{\partial \underline{\theta}} = \sum_{i=1}^m n_i \frac{\partial \ln \pi_i}{\partial \underline{\theta}} = 0; \quad \underline{\theta} = [\theta_1, \theta_2, \dots, \theta_p] \quad (3.4)$$

TEOREMAS: Al tratar de dar una solución a las ecuaciones de verosimilitud, podemos enfrentarnos con preguntas como las siguientes:

- a) ¿Existe una raíz para la ecuación?
- b) ¿Esta solución es única?
- c) ¿Esta raíz es el máximo absoluto de la función de verosimilitud?

El Teorema siguiente puede ser de utilidad al proveer de un conjunto de condiciones suficientes para responder positivamente a las preguntas anteriores.

TEOREMA 3.1 Si las siguientes tres condiciones se cumplen.

$$(3.1.1) \quad \frac{\partial \ln \pi_i}{\partial \theta} \text{ y } \frac{\partial^2 \ln \pi_i}{\partial \theta^2} \text{ existen para todo } \theta \in \Omega \text{ y para cada } i \text{ tal que } n_i > 0;$$

siendo Ω el espacio parametral

$$(3.1.2) \quad \sum_{i=1}^k n_i \frac{\partial \ln \pi_i}{\partial \theta} = 0 \begin{cases} > 0 & \text{para } \theta = \theta_1 + 0 \\ < 0 & \text{para } \theta = \theta_2 - 0, \end{cases}$$

$$(3.1.3) \quad \sum_{i=1}^k n_i \frac{\partial^2 \ln \pi_i}{\partial \theta^2} < 0 \text{ para cualquier } \theta \text{ que satisface la ecuación (3.4)}$$

entonces la ecuación de verosimilitud (3.4) tiene una y solo una raíz y esta raíz es el MLE de θ .

Note que, dada una muestra de tamaño n , existe la probabilidad de que el MLE no exista. Aún cuando esta probabilidad tiende a cero mientras $n \rightarrow \infty$. Dado lo anterior, Kulldorff (1961) propone trabajar tanto con las propiedades asintóticas como con las distribuciones asintóticas para relacionar al estimador no con su mismo MLE sino con la variable aleatoria que tiene por

Función de Distribución: $\text{Prob}\{MLE \leq x \mid MLE \text{ existe}\}$. Esta variable aleatoria será llamada *el MLE Condicional*.

Para los diferentes MLE obtenidos en los siguientes capítulos, nos interesaremos en probar que el MLE condicional es consistente y asintóticamente eficiente en el sentido estricto – término propuesto y definido por Wald (1948) –, y obtener la varianza de su distribución asintótica (usualmente llamada Varianza Asintótica). Así se hará uso de dos diferentes teoremas: **Teorema 3.2** y **Teorema 3.3** los cuales tienen sustento en las **Condiciones 3.1 – 3.11**, y tienen que ver con la raíz de la función de verosimilitud. Para ello sean

θ_0 : el valor verdadero de θ y un punto interior de Ω

Ω_0 : alguna vecindad (arbitrariamente pequeña) de θ_0 tal que $\theta \in \Omega_0 \subset \Omega$,

CONDICIÓN 3.1. $\frac{\partial \ln \pi_i}{\partial \theta}$ existe para cada $\theta \in \Omega_0$ y para $i = 1, 2, \dots, m$.

CONDICIÓN 3.2. $\frac{\partial^2 \ln \pi_i}{\partial \theta^2}$ existe para cada $\theta \in \Omega_0$ y para $i = 1, 2, \dots, m$.

CONDICIÓN 3.3. $\frac{\partial^3 \ln \pi_i}{\partial \theta^3}$ existe para cada $\theta \in \Omega_0$ y para $i = 1, 2, \dots, m$.

CONDICIÓN 3.4. $\sum_{i=1}^k \frac{\partial \ln \pi_i}{\partial \theta} = 0$ para $\theta = \theta_0$

CONDICIÓN 3.5. $\sum_{i=1}^k \frac{\partial^2 \ln \pi_i}{\partial \theta^2} = 0$ para $\theta = \theta_0$

CONDICIÓN 3.6. $-\infty < \sum_{i=1}^k \pi_i \frac{\partial \ln \pi_i}{\partial \underline{\theta}} < 0$ para $\underline{\theta} = \underline{\theta}_0$

CONDICIÓN 3.7. Existe un conjunto de números P_i (independientes de $\underline{\theta}$ pero posiblemente independientes de $\underline{\theta}_0$) tales que:

a) $\left| \frac{\partial^3 \ln \pi_i}{\partial \underline{\theta}^3} \right| < P_i$ para cada $\underline{\theta} \in \Omega_0$ y para $i = 1, 2, \dots, m$. y

b) $\sum_{i=1}^k \pi_i(\underline{\theta}_0) P_i < \infty$.

CONDICIÓN 3.8. Existe una función $g(\underline{\theta})$, la cual es positiva y con segunda derivada para todo $\underline{\theta} \in \Omega_0$ y un conjunto de números P_i (independientes de $\underline{\theta}$ pero posiblemente independientes de $\underline{\theta}_0$) tales que:

a) $\left| \frac{\partial^2}{\partial \underline{\theta}^2} \left(g(\underline{\theta}) \frac{\partial \ln \pi_i}{\partial \underline{\theta}} \right) \right| < P_i$ para cada $\underline{\theta} \in \Omega_0$ y para $i = 1, 2, \dots, m$. y

b) $\sum_{i=1}^k \pi_i(\underline{\theta}_0) P_i < \infty$.

CONDICIÓN 3.9. $\frac{\partial \ln \pi_i}{\partial \underline{\theta}}$ es una función continua de $\underline{\theta}$, para cada $\underline{\theta} \in \Omega_0$ y para

$$i = 1, 2, \dots, m.$$

CONDICIÓN 3.10. $(\underline{\theta} - \theta_0) \frac{\partial \ln \pi_i}{\partial \underline{\theta}} \pi_i(\underline{\theta}) < 0$ para cada $\underline{\theta} \in \Omega_0$ tales que $\underline{\theta} \neq \theta_0$.

CONDICIÓN 3.11. Existe una función $g(\theta)$, la cual es positiva y diferenciable para todo $\underline{\theta} \in \Omega_0$ tales que, para cada $\underline{\theta} \in \Omega_0$, $\frac{\partial}{\partial \underline{\theta}} \left(g(\underline{\theta}) \frac{\partial \ln \pi_i}{\partial \underline{\theta}} \right)$ es una función uniformemente continua de $\underline{\theta}$ para $i = 1, 2, \dots, m$.

TEOREMA 3.2. Si para cualquier valor de n y para cualesquiera valores n_i de la muestra, hay al menos una raíz de (3.4), y si

θ_0 : es el valor verdadero de θ y un punto interior de Ω

Ω_0 : es alguna vecindad (arbitrariamente pequeña) de θ_0 tal que $\theta \in \Omega_0 \subset \Omega$,

$\hat{\theta}$: es la raíz de (3.4) si la raíz existe con probabilidad 1, de otra manera, la variable aleatoria

con función de distribución $\text{Prob}\{\hat{\theta} \leq x\} = \text{Prob}\{\text{la raíz} \leq x \mid \text{la raíz existe}\}$

y si también

- a) CONDICIONES 3.1 – 3.4, 3.6, 3.7 o,
- b) CONDICIONES 3.1 – 3.4, 3.6, 3.8 o,
- c) CONDICIONES 3.1, 3.9, 3.10,

son satisfechas, entonces:

- (i) $\hat{\theta}$ es un estimador consistente de θ , y
- (ii) una raíz de (3.4) existe con probabilidad que tiende a uno mientras $n \rightarrow \infty$, y tal raíz es el máximo de la función de verosimilitud.

TEOREMA 3.3 Si para cualquier valor de n y cualesquiera sean los valores de la muestra, a lo más existe una raíz de (3.4), si además θ_0 , Ω_0 y $\hat{\theta}$ tiene el mismo significado que en el TEOREMA 3.2 y ya sea.

- a) CONDICIONES 3.1 – 3.7 o,
- b) CONDICIONES 3.1 – 3.6, 3.8, o
- c) CONDICIONES 3.1, 3.2, 3.4 – 3.6, 3.10, 3.11;

se satisfacen, entonces $\hat{\theta}$ es asintóticamente eficiente en el estricto sentido como un estimador de θ y la varianza de su distribución asintótica es:

$$\text{AsVar}(\hat{\theta}) = E \left[\left(\frac{\partial \ln \pi_i}{\partial \hat{\theta}} \right)^2 \right]^{-1} = \left[n \sum_{i=1}^m \pi_i \left(\frac{\partial \ln \pi_i}{\partial \hat{\theta}} \right)^2 \right]^{-1} = \left(-n \sum_{i=1}^m \pi_i \frac{\partial^2 \ln \pi_i}{\partial \hat{\theta}^2} \right)^{-1} \quad (3.5)$$

Es importante notar que cuando el número de parámetros en θ es mayor a uno, lo anterior resulta en una matriz, la cual es la matriz de información de Fisher; en nuestro caso de información asintótica. La expresión (3.30) da cuenta de lo anterior.

3.3. Obtención de los Estimadores, las Varianzas y Covarianza Asintóticas

Demos ahora solución a nuestro problema particular, determinando en primer lugar las expresiones de los Estimadores de Máxima Verosimilitud, para después establecer las expresiones de la Varianza Asintótica de los Estimadores.

3.3.1. Función de Densidad Normal

Con lo expuesto en la **Sección 3.1**, suponga que la población en estudio se distribuye como Normal, cuya Función de Densidad de Probabilidad y Función de Densidad Acumulada se definen, en términos de los límites de truncamiento:

$$f_X(h_i; \mu, \sigma) = \frac{1}{\sqrt{2\pi}\sigma} e^{-\frac{1}{2}\left(\frac{h_i - \mu}{\sigma}\right)^2} \quad (3.6)$$

$$F_X(h_i; \mu, \sigma) = \int_{-\infty}^{h_i} f_X\left(\frac{x_i - \mu}{\sigma}\right) dx_i$$

donde $-\infty < \mu < \infty$ es la media de la población, $\sigma > 0$ es la desviación estándar de la población $h_i \in \mathbb{R}$ ($i = 0, 1, 2, \dots, m$) con $h_0 = -\infty$ y $h_m = \infty$.

Sean $h_i \in \mathbb{R}$ ($i = 0, 1, 2, \dots, m$) los límites de truncamiento como en Definiciones 3.1 y 3.2, defínase a ser:

$$\xi_i = \frac{h_i - \mu}{\sigma} \quad (3.7)$$

Transformándose de esta manera nuestra fdp. y fda. en Funciones Normales Estándar definidas como sigue:

$$\phi(\xi_i; 0, 1) = \frac{1}{\sqrt{2\pi}} e^{-\frac{1}{2}\xi_i^2} \quad (3.8)$$

$$\Phi(\xi_i; 0, 1) = \int_{-\infty}^{\xi_i} \phi(z_i) dz_i$$

luego al sustituir (3.8) en (3.1) se tiene que:

$$\pi_i = F(h_i) - F(h_{i-1}) = \Phi(\xi_i) - \Phi(\xi_{i-1}) \quad (3.9)$$

para luego, al llevarle a (3.3) se llega a la expresión siguiente:

$$L(\theta) = \sum_{i=1}^m n_i \ln [\Phi(\xi_i) - \Phi(\xi_{i-1})] \quad (3.10)$$

3.3.1.1. Estimadores de Máxima Verosimilitud

Llevando (3.10) a (3.4) se tiene:

$$MLE(\underline{\theta}) = \sum_{i=1}^m n_i \frac{\partial \ln[\Phi(\xi_i) - \Phi(\xi_{i-1})]}{\partial \underline{\theta}} \quad (3.11)$$

con $\underline{\theta}' = [\mu, \sigma]$.

Ahora bien, para obtener los MLE para la media y la desviación estándar; se deriva parcialmente a (3.11) respecto a μ y a σ y se igualan a cero las expresiones resultantes.

3.3.1.1.1. Para la Media

$$\begin{aligned} MLE(\mu) &= \frac{\partial \ln L(\underline{\theta}; h_i)}{\partial \mu} = \sum_{i=1}^m n_i \frac{\partial \ln[\Phi(\xi_i) - \Phi(\xi_{i-1})]}{\partial \mu} \\ &= \sum_{i=1}^m n_i \frac{1}{\Phi(\xi_i) - \Phi(\xi_{i-1})} \left[\frac{\partial \Phi(\xi_i)}{\partial \mu} - \frac{\partial \Phi(\xi_{i-1})}{\partial \mu} \right] \end{aligned} \quad (3.12)$$

donde, por la regla de la cadena,

$$\frac{\partial \Phi(\xi_i)}{\partial \mu} = \phi(\xi_i) \frac{\partial \xi_i}{\partial \mu} = \phi(\xi_i) \frac{\partial [(h_i - \mu)/\sigma]}{\partial \mu} = -\frac{1}{\sigma} \phi(\xi_i) \quad (3.12.1)$$

análogamente

$$\frac{\partial \Phi(\xi_{i-1})}{\partial \mu} = \phi(\xi_{i-1}) \frac{\partial \xi_{i-1}}{\partial \mu} = \phi(\xi_{i-1}) \frac{\partial [(h_{i-1} - \mu)/\sigma]}{\partial \mu} = -\frac{1}{\sigma} \phi(\xi_{i-1}) \quad (3.12.2)$$

substituyendo (3.12.1) y (3.12.2) en (3.12) tenemos que

$$= \sum_{i=1}^m n_i \frac{1}{\Phi(\xi_i) - \Phi(\xi_{i-1})} \left(-\frac{1}{\sigma} \right) [\phi(\xi_i) - \phi(\xi_{i-1})] = \left(-\frac{1}{\sigma} \right) \sum_{i=1}^m n_i \frac{\phi(\xi_i) - \phi(\xi_{i-1})}{\Phi(\xi_i) - \Phi(\xi_{i-1})}$$

expresión que al igualarse a cero produce

$$\sum_{i=1}^m n_i \frac{\phi(\xi_i) - \phi(\xi_{i-1})}{\Phi(\xi_i) - \Phi(\xi_{i-1})} = 0 \quad (3.13)$$

3.2.1.1.2. Para la Desviación Estándar

$$\begin{aligned} MLE(\sigma) &= \frac{\partial \ln L(\theta; h_i)}{\partial \sigma} = \sum_{i=1}^m n_i \frac{\partial \ln [\Phi(\xi_i) - \Phi(\xi_{i-1})]}{\partial \sigma} \\ &= \sum_{i=1}^m n_i \frac{1}{\Phi(\xi_i) - \Phi(\xi_{i-1})} \left[\frac{\partial \Phi(\xi_i)}{\partial \sigma} - \frac{\partial \Phi(\xi_{i-1})}{\partial \sigma} \right] \end{aligned} \quad (3.14)$$

donde por la regla de la cadena,

$$\frac{\partial \Phi(\xi_i)}{\partial \sigma} = \phi(\xi_i) \frac{\partial \xi_i}{\partial \sigma} = \phi(\xi_i) (h_i - \mu) \frac{\partial \sigma^{-1}}{\partial \sigma} = -\frac{(h_i - \mu)}{\sigma^2} \phi(\xi_i) = -\frac{1}{\sigma} \xi_i \phi(\xi_i) \quad (3.14.1)$$

análogamente

$$\frac{\partial \Phi(\xi_{i-1})}{\partial \sigma} = \phi(\xi_i) \frac{\partial \xi_{i-1}}{\partial \sigma} = \phi(\xi_{i-1})(h_{i-1} - \mu) \frac{\partial \sigma^{-1}}{\partial \sigma} = -\frac{(h_{i-1} - \mu)}{\sigma^2} \phi(\xi_{i-1}) = -\frac{1}{\sigma} \xi_{i-1} \phi(\xi_{i-1}) \quad (3.14.2)$$

substituyendo (3.14.1) y (3.14.2) en (3.14) tenemos que

$$= \sum_{i=1}^m n_i \frac{1}{\Phi(\xi_i) - \Phi(\xi_{i-1})} \left(-\frac{1}{\sigma} \right) [\xi_i \phi(\xi_i) - \xi_{i-1} \phi(\xi_{i-1})] = \left(-\frac{1}{\sigma} \right) \sum_{i=1}^m n_i \frac{\xi_i \phi(\xi_i) - \xi_{i-1} \phi(\xi_{i-1})}{\Phi(\xi_i) - \Phi(\xi_{i-1})}$$

la que al igualar a cero produce

$$\sum_{i=1}^m n_i \frac{\xi_i \phi(\xi_i) - \xi_{i-1} \phi(\xi_{i-1})}{\Phi(\xi_i) - \Phi(\xi_{i-1})} = 0 \quad (3.15)$$

Por tanto para encontrar los estimadores de máxima verosimilitud para la media y la desviación estándar se debe resolver, simultáneamente, el sistema de ecuaciones siguiente:

$$MLE(\mu): \sum_{i=1}^m n_i \frac{\phi(\xi_i) - \phi(\xi_{i-1})}{\Phi(\xi_i) - \Phi(\xi_{i-1})} = 0 \quad (3.16)$$

$$MLE(\sigma): \sum_{i=1}^m n_i \frac{(\xi_i) \phi(\xi_i) - (\xi_{i-1}) \phi(\xi_{i-1})}{\Phi(\xi_i) - \Phi(\xi_{i-1})} = 0$$

A manera de simplificar las expresiones se hará uso de la función propuesta por Cohen (1991).

Sean:

$$Q(\xi_i) = \frac{\phi(\xi_i)}{\Phi(\xi_i) - \Phi(\xi_{i-1})} \quad (3.16.1)$$

$$Q(\xi_{i-1}) = \frac{\phi(\xi_{i-1})}{\Phi(\xi_i) - \Phi(\xi_{i-1})}$$

luego al sustituir (3.16.1) en (3.16) el sistema de ecuaciones que debemos resolver para encontrar a los estimadores deseados son:

$$\sum_{i=1}^m n_i [Q(\xi_i) - Q(\xi_{i-1})] = 0 \quad (3.17)$$

$$\sum_{i=1}^m n_i [(\xi_i)Q(\xi_i) - (\xi_{i-1})Q(\xi_{i-1})] = 0$$

3.3.1.2. Varianzas y Covarianza Asintóticas

Substituyendo (3.9) en (3.5) obtenemos:

$$\text{AsVar}(\hat{\theta}) = E \left[\left(\frac{\partial \ln [\Phi(\xi_i) - \Phi(\xi_{i-1})]}{\partial \theta} \right)^2 \right]^{-1}$$

$$= \left[-n \sum_{i=1}^m [\Phi(\xi_i) - \Phi(\xi_{i-1})] \frac{\partial^2 \ln[\Phi(\xi_i) - \Phi(\xi_{i-1})]}{\partial \theta^2} \right]^{-1} \quad (3.18)$$

3.3.1.2.1. Varianza Asintótica de la Media

A fin de obtener la expresión de la varianza asintótica para la media debemos derivar parcialmente con respecto a μ la expresión (3.18) es decir

$$\begin{aligned} \text{AsVar}(\hat{\mu}) &= E \left[\left(\frac{\partial \ln[\Phi(\xi_i) - \Phi(\xi_{i-1})]}{\partial \mu} \right)^2 \right]^{-1} \\ &= \left[-n \sum_{i=1}^m [\Phi(\xi_i) - \Phi(\xi_{i-1})] \frac{\partial^2 \ln[\Phi(\xi_i) - \Phi(\xi_{i-1})]}{\partial \mu^2} \right]^{-1} \end{aligned} \quad (3.19)$$

donde de (3.12) sabemos que: $\frac{\partial \ln[\Phi(\xi_i) - \Phi(\xi_{i-1})]}{\partial \mu} = \left(-\frac{1}{\sigma} \right) \frac{\phi(\xi_i) - \phi(\xi_{i-1})}{\Phi(\xi_i) - \Phi(\xi_{i-1})}$ y por tanto:

$$\frac{\partial^2 \ln[\Phi(\xi_i) - \Phi(\xi_{i-1})]}{\partial \mu^2} = -\frac{1}{\sigma} \frac{\partial}{\partial \mu} \left(\frac{\phi(\xi_i) - \phi(\xi_{i-1})}{\Phi(\xi_i) - \Phi(\xi_{i-1})} \right)$$

$$= -\frac{1}{\sigma} \left\{ \frac{[\Phi(\xi_i) - \Phi(\xi_{i-1})] \frac{\partial}{\partial \mu} [\phi(\xi_i) - \phi(\xi_{i-1})] - [\phi(\xi_i) - \phi(\xi_{i-1})] \frac{\partial}{\partial \mu} [\Phi(\xi_i) - \Phi(\xi_{i-1})]}{[\Phi(\xi_i) - \Phi(\xi_{i-1})]^2} \right\} \quad (3.19.1)$$

donde

$$\frac{\partial}{\partial \mu} [\phi(\xi_i) - \phi(\xi_{i-1})] = \frac{\partial \phi(\xi_i)}{\partial \mu} - \frac{\partial \phi(\xi_{i-1})}{\partial \mu}$$

y a su vez

$$\frac{\partial}{\partial \mu} \phi(\xi_i) = \frac{\partial}{\partial \mu} \frac{1}{\sqrt{2\pi}} e^{-\frac{1}{2}\xi_i^2} = \frac{1}{\sqrt{2\pi}} e^{-\frac{1}{2}\xi_i^2} \frac{\partial}{\partial \mu} \left(-\frac{1}{2}\xi_i^2 \right) = -\frac{2}{2} \xi_i \phi(\xi_i) \frac{\partial}{\partial \mu} \xi_i = \frac{\xi_i}{\sigma} \phi(\xi_i) \quad (3.19.2)$$

$$\frac{\partial}{\partial \mu} \phi(\xi_{i-1}) = \frac{\partial}{\partial \mu} \frac{1}{\sqrt{2\pi}} e^{-\frac{1}{2}\xi_{i-1}^2} = \frac{1}{\sqrt{2\pi}} e^{-\frac{1}{2}\xi_{i-1}^2} \frac{\partial}{\partial \mu} \left(-\frac{1}{2}\xi_{i-1}^2 \right) = -\frac{2}{2} \xi_{i-1} \phi(\xi_{i-1}) \frac{\partial}{\partial \mu} \xi_{i-1} = \frac{\xi_{i-1}}{\sigma} \phi(\xi_{i-1})$$

Substituyendo (3.13.1), (3.13.2) y (3.19.2) en (3.19.1) tenemos que

$$\frac{\partial^2 \log [\Phi(\xi_i) - \Phi(\xi_{i-1})]}{\partial \mu^2} =$$

$$\frac{-1}{\sigma [\Phi(\xi_i) - \Phi(\xi_{i-1})]^2} \left\{ \left[(\Phi(\xi_i) - \Phi(\xi_{i-1})) \left(\frac{\xi_i}{\sigma} \phi(\xi_i) - \frac{\xi_{i-1}}{\sigma} \phi(\xi_{i-1}) \right) \right] - (\phi(\xi_i) - \phi(\xi_{i-1})) \left(\left[-\frac{1}{\sigma} \phi(\xi_i) \right] - \left[-\frac{1}{\sigma} \phi(\xi_{i-1}) \right] \right) \right\}$$

$$\Rightarrow \frac{\partial^2 \log[\Phi(\xi_i) - \Phi(\xi_{i-1})]}{\partial \mu^2} = -\frac{1}{\sigma^2} \left\{ \frac{[\Phi(\xi_i) - \Phi(\xi_{i-1})][\xi_i \phi(\xi_i) - \xi_{i-1} \phi(\xi_{i-1})] + [\phi(\xi_i) - \phi(\xi_{i-1})]^2}{[\Phi(\xi_i) - \Phi(\xi_{i-1})]^2} \right\}$$

Finalmente, la expresión de la varianza asintótica para la media es la siguiente

$$\text{AsVar}(\hat{\mu}) = \left(\frac{n}{\sigma^2} \sum_{i=1}^k [\Phi(\xi_i) - \Phi(\xi_{i-1})] \frac{[\Phi(\xi_i) - \Phi(\xi_{i-1})][\xi_i \phi(\xi_i) - \xi_{i-1} \phi(\xi_{i-1})] + [\phi(\xi_i) - \phi(\xi_{i-1})]^2}{[\Phi(\xi_i) - \Phi(\xi_{i-1})]^2} \right)^{-1}$$

o bien

$$\text{AsVar}(\hat{\mu}) = \left(\frac{n}{\sigma^2} \sum_{i=1}^k \frac{[\Phi(\xi_i) - \Phi(\xi_{i-1})][\xi_i \phi(\xi_i) - \xi_{i-1} \phi(\xi_{i-1})] + [\phi(\xi_i) - \phi(\xi_{i-1})]^2}{\Phi(\xi_i) - \Phi(\xi_{i-1})} \right)^{-1} \quad (3.20)$$

3.3.1.2.2. Varianza Asintótica de la Desviación Estándar

A fin de obtener la expresión de la varianza asintótica para la desviación estándar debemos derivar parcialmente con respecto a σ la expresión (3.18) es decir

$$\text{AsVar}(\hat{\sigma}) = E \left[\left(\frac{\partial \ln[\Phi(\xi_i) - \Phi(\xi_{i-1})]}{\partial \sigma} \right)^2 \right]^{-1}$$

$$= \left[-n \sum_{i=1}^m [\Phi(\xi_i) - \Phi(\xi_{i-1})] \frac{\partial^2 \ln[\Phi(\xi_i) - \Phi(\xi_{i-1})]}{\partial \sigma^2} \right]^{-1} \quad (3.21)$$

donde de (3.15) sabemos que: $\frac{\partial \ln[\Phi(\xi_i) - \Phi(\xi_{i-1})]}{\partial \sigma} = \left(-\frac{1}{\sigma} \right) \frac{\xi_i \phi(\xi_i) - \xi_{i-1} \phi(\xi_{i-1})}{\Phi(\xi_i) - \Phi(\xi_{i-1})}$ y por

tanto:

$$\begin{aligned} \frac{\partial^2 \ln[\Phi(\xi_i) - \Phi(\xi_{i-1})]}{\partial \sigma^2} &= \left(-\frac{1}{\sigma} \right) \frac{\partial}{\partial \sigma} \left(\frac{\xi_i \phi(\xi_i) - \xi_{i-1} \phi(\xi_{i-1})}{\Phi(\xi_i) - \Phi(\xi_{i-1})} \right) \\ &= -\frac{1}{\sigma} \left\{ \frac{[\Phi(\xi_i) - \Phi(\xi_{i-1})] \frac{\partial}{\partial \sigma} [\xi_i \phi(\xi_i) - \xi_{i-1} \phi(\xi_{i-1})]}{[\Phi(\xi_i) - \Phi(\xi_{i-1})]^2} \right. \\ &\quad \left. - \frac{[\xi_i \phi(\xi_i) - \xi_{i-1} \phi(\xi_{i-1})] \frac{\partial}{\partial \sigma} [\Phi(\xi_i) - \Phi(\xi_{i-1})]}{[\Phi(\xi_i) - \Phi(\xi_{i-1})]^2} \right\} \end{aligned} \quad (3.22)$$

note que:

$$\frac{\partial}{\partial \sigma} [\xi_i \phi(\xi_i) - \xi_{i-1} \phi(\xi_{i-1})] = \frac{\partial \xi_i \phi(\xi_i)}{\partial \sigma} - \frac{\partial \xi_{i-1} \phi(\xi_{i-1})}{\partial \sigma} \quad (3.22.1)$$

donde

$$\frac{\partial \xi_i \phi(\xi_i)}{\partial \sigma} = \xi_i \frac{\partial \phi(\xi_i)}{\partial \sigma} + \phi(\xi_i) \frac{\partial \xi_i}{\partial \sigma} \quad (3.22.2)$$

$$\frac{\partial \xi_{i-1} \phi(\xi_{i-1})}{\partial \sigma} = \xi_{i-1} \frac{\partial \phi(\xi_{i-1})}{\partial \sigma} + \phi(\xi_{i-1}) \frac{\partial \xi_{i-1}}{\partial \sigma}$$

luego

$$\frac{\partial \xi_i}{\partial \sigma} = \frac{\partial [(h_i - \mu)/\sigma]}{\partial \sigma} = (h_i - \mu) \frac{\partial \sigma^{-1}}{\partial \sigma} = -\frac{(h_i - \mu)}{\sigma^2} = -\frac{1}{\sigma} \xi_i \quad (3.22.3)$$

$$\frac{\partial \xi_{i-1}}{\partial \sigma} = \frac{\partial [(h_{i-1} - \mu)/\sigma]}{\partial \sigma} = (h_{i-1} - \mu) \frac{\partial \sigma^{-1}}{\partial \sigma} = -\frac{(h_{i-1} - \mu)}{\sigma^2} = -\frac{1}{\sigma} \xi_{i-1}$$

$$\frac{\partial}{\partial \sigma} \phi(\xi_i) = \frac{\partial}{\partial \sigma} \frac{1}{\sqrt{2\pi}} e^{-\frac{1}{2}\xi_i^2} = \frac{1}{\sqrt{2\pi}} e^{-\frac{1}{2}\xi_i^2} \frac{\partial}{\partial \sigma} \left(-\frac{1}{2} \xi_i^2 \right) = -\frac{2}{2} \phi(\xi_i) \xi_i \frac{\partial \xi_i}{\partial \sigma} = \frac{\xi_i^2 \phi(\xi_i)}{\sigma} \quad (3.22.4)$$

$$\frac{\partial}{\partial \sigma} \phi(\xi_{i-1}) = \frac{\partial}{\partial \sigma} \frac{1}{\sqrt{2\pi}} e^{-\frac{1}{2}\xi_{i-1}^2} = \frac{1}{\sqrt{2\pi}} e^{-\frac{1}{2}\xi_{i-1}^2} \frac{\partial}{\partial \sigma} \left(-\frac{1}{2} \xi_{i-1}^2 \right) = -\frac{2}{2} \xi_{i-1} \phi(\xi_{i-1}) \frac{\partial \xi_{i-1}}{\partial \sigma} = \frac{\xi_{i-1}^2 \phi(\xi_{i-1})}{\sigma}$$

substituyendo (3.22.3) y (3.22.4) en (3.22.2) se tiene que:

$$\frac{\partial \xi_i \phi(\xi_i)}{\partial \sigma} = \xi_i \frac{\xi_i^2 \phi(\xi_i)}{\sigma} - \phi(\xi_i) \frac{\xi_i}{\sigma} = \frac{\xi_i \phi(\xi_i)}{\sigma} [\xi_i^2 - 1] \quad (3.22.5)$$

$$\frac{\partial \xi_{i-1} \phi(\xi_{i-1})}{\partial \sigma} = \xi_{i-1} \frac{\xi_{i-1}^2 \phi(\xi_{i-1})}{\sigma} - \phi(\xi_{i-1}) \frac{\xi_{i-1}}{\sigma} = \frac{\xi_{i-1} \phi(\xi_{i-1})}{\sigma} [\xi_{i-1}^2 - 1]$$

ahora substituyendo (3.22.5) en (3.22.1) tenemos que:

$$\frac{\partial}{\partial \sigma} [\xi_i \phi(\xi_i) - \xi_{i-1} \phi(\xi_{i-1})] = \frac{\xi_i \phi(\xi_i)}{\sigma} [\xi_i^2 - 1] - \frac{\xi_{i-1} \phi(\xi_{i-1})}{\sigma} [\xi_{i-1}^2 - 1] \quad (3.22.6)$$

y reemplazando (3.22.6) en (3.22) tenemos que:

$$\begin{aligned} & \frac{\partial^2 \ln [\Phi(\xi_i) - \Phi(\xi_{i-1})]}{\partial \sigma^2} \\ &= -\frac{1}{\sigma} \left\{ \frac{[\Phi(\xi_i) - \Phi(\xi_{i-1})] \frac{\partial}{\partial \sigma} [\xi_i \phi(\xi_i) - \xi_{i-1} \phi(\xi_{i-1})] - [\xi_i \phi(\xi_i) - \xi_{i-1} \phi(\xi_{i-1})] \frac{\partial}{\partial \sigma} [\Phi(\xi_i) - \Phi(\xi_{i-1})]}{[\Phi(\xi_i) - \Phi(\xi_{i-1})]^2} \right\} \\ &= -\frac{1}{\sigma} \left\{ \frac{\left[\frac{\xi_i \phi(\xi_i)}{\sigma} [\xi_i^2 - 1] - \frac{\xi_{i-1} \phi(\xi_{i-1})}{\sigma} [\xi_{i-1}^2 - 1] \right] [\Phi(\xi_i) - \Phi(\xi_{i-1})] - \left[\frac{1}{\sigma} \xi_{i-1} \phi(\xi_{i-1}) - \frac{1}{\sigma} \xi_i \phi(\xi_i) \right] [\Phi(\xi_i) - \Phi(\xi_{i-1})]^2}{[\Phi(\xi_i) - \Phi(\xi_{i-1})]^3} \right\} \end{aligned}$$

entonces

$$\frac{\partial^2 \ln [\Phi(\xi_i) - \Phi(\xi_{i-1})]}{\partial \sigma^2} =$$

$$-\frac{1}{\sigma^2} \left\{ \frac{[\xi_i \phi(\xi_i)[\xi_i^2 - 1] - \xi_{i-1} \phi(\xi_{i-1})[\xi_{i-1}^2 - 1]]}{[\Phi(\xi_i) - \Phi(\xi_{i-1})]} + \frac{[\xi_i \phi(\xi_i) - \xi_{i-1} \phi(\xi_{i-1})]^2}{[\Phi(\xi_i) - \Phi(\xi_{i-1})]^2} \right\} \quad (3.23)$$

finalmente, (3.23) en (3.21) obtenemos la expresión final

AsVar($\hat{\sigma}$) =

$$\frac{\sigma^2}{n} \left[\sum_{i=1}^m (\xi_i \phi(\xi_i)[\xi_i^2 - 1] - \xi_{i-1} \phi(\xi_{i-1})[\xi_{i-1}^2 - 1]) + \frac{[\xi_i \phi(\xi_i) - \xi_{i-1} \phi(\xi_{i-1})]^2}{[\Phi(\xi_i) - \Phi(\xi_{i-1})]} \right]^{-1} \quad (3.24)$$

3.3.1.2.3. Covarianza Asintótica

Para obtener la expresión de la covarianza asymptótica es necesario derivar parcialmente primero con respecto a μ y luego con respecto a σ la expresión (3.18) es decir

$$\begin{aligned} \text{AsCov}(\hat{\mu}, \hat{\sigma}) &= E \left[\left(\frac{\partial \ln[\Phi(\xi_i) - \Phi(\xi_{i-1})]}{\partial \mu \partial \sigma} \right)^2 \right]^{-1} \\ &= \left[-n \sum_{i=1}^m [\Phi(\xi_i) - \Phi(\xi_{i-1})] \frac{\partial^2 \ln[\Phi(\xi_i) - \Phi(\xi_{i-1})]}{\partial \mu \partial \sigma} \right]^{-1} \end{aligned} \quad (3.25)$$

donde

$$\frac{\partial^2 \ln[\Phi(\xi_i) - \Phi(\xi_{i-1})]}{\partial \mu \partial \sigma} = \frac{\partial}{\partial \sigma} \left\{ \frac{\partial \ln[\Phi(\xi_i) - \Phi(\xi_{i-1})]}{\partial \mu} \right\} \quad (3.26)$$

Sabemos de (3.13) que:

$$\frac{\partial \ln[\Phi(\xi_i) - \Phi(\xi_{i-1})]}{\partial \mu} = \left(-\frac{1}{\sigma} \right) \frac{\phi(\xi_i) - \phi(\xi_{i-1})}{\Phi(\xi_i) - \Phi(\xi_{i-1})} \quad (3.27)$$

ahora bien substituyendo (3.27) en (3.26)

$$\frac{\partial^2 \ln[\Phi(\xi_i) - \Phi(\xi_{i-1})]}{\partial \mu \partial \sigma} = \frac{\partial}{\partial \sigma} \left\{ \left(-\frac{1}{\sigma} \right) \frac{\phi(\xi_i) - \phi(\xi_{i-1})}{\Phi(\xi_i) - \Phi(\xi_{i-1})} \right\}$$

$$\begin{aligned} \frac{\partial^2 \ln[\Phi(\xi_i) - \Phi(\xi_{i-1})]}{\partial \mu \partial \sigma} &= \left(-\frac{\partial \sigma^{-1}}{\partial \sigma} \right) \frac{\phi(\xi_i) - \phi(\xi_{i-1})}{\Phi(\xi_i) - \Phi(\xi_{i-1})} - \left(\frac{1}{\sigma} \right) \frac{\partial}{\partial \sigma} \left\{ \frac{\phi(\xi_i) - \phi(\xi_{i-1})}{\Phi(\xi_i) - \Phi(\xi_{i-1})} \right\} \\ &= (\sigma^{-2}) \frac{\phi(\xi_i) - \phi(\xi_{i-1})}{\Phi(\xi_i) - \Phi(\xi_{i-1})} - \end{aligned}$$

$$\left(\frac{1}{\sigma}\right) \frac{1}{[\Phi(\xi_i) - \Phi(\xi_{i-1})]^2} \left\{ [\Phi(\xi_i) - \Phi(\xi_{i-1})] \frac{\partial [\phi(\xi_i) - \phi(\xi_{i-1})]}{\partial \sigma} - [\phi(\xi_i) - \phi(\xi_{i-1})] \frac{\partial [\Phi(\xi_i) - \Phi(\xi_{i-1})]}{\partial \sigma} \right\}$$

substituyendo (3.14.1) y (3.14.2) junto con (3.19.2) tenemos que:

$$= \left(\frac{1}{\sigma^2}\right) \frac{\phi(\xi_i) - \phi(\xi_{i-1})}{\Phi(\xi_i) - \Phi(\xi_{i-1})} -$$

$$\frac{1}{\sigma [\Phi(\xi_i) - \Phi(\xi_{i-1})]^2} \left\{ [\Phi(\xi_i) - \Phi(\xi_{i-1})] \left[\frac{\xi_i^2 \phi(\xi_i)}{\sigma} - \frac{\xi_{i-1}^2 \phi(\xi_{i-1})}{\sigma} \right] - [\phi(\xi_i) - \phi(\xi_{i-1})] \left[\frac{-\xi_i \phi(\xi_i)}{\sigma} - \frac{-\xi_{i-1} \phi(\xi_{i-1})}{\sigma} \right] \right\}$$

$$= \left(\frac{1}{\sigma^2}\right) \frac{\phi(\xi_i) - \phi(\xi_{i-1})}{\Phi(\xi_i) - \Phi(\xi_{i-1})} -$$

$$\frac{1}{\sigma^2 [\Phi(\xi_i) - \Phi(\xi_{i-1})]^2} \left\{ [\Phi(\xi_i) - \Phi(\xi_{i-1})] [\xi_i^2 \phi(\xi_i) - \xi_{i-1}^2 \phi(\xi_{i-1})] + [\phi(\xi_i) - \phi(\xi_{i-1})] [\xi_i \phi(\xi_i) - \xi_{i-1} \phi(\xi_{i-1})] \right\}$$

$$= \left(\frac{1}{\sigma^2}\right) \left(\frac{\phi(\xi_i) - \phi(\xi_{i-1})}{\Phi(\xi_i) - \Phi(\xi_{i-1})} -$$

$$\left\{ \frac{[\Phi(\xi_i) - \Phi(\xi_{i-1})] [\xi_i^2 \phi(\xi_i) - \xi_{i-1}^2 \phi(\xi_{i-1})] + [\phi(\xi_i) - \phi(\xi_{i-1})] [\xi_i \phi(\xi_i) - \xi_{i-1} \phi(\xi_{i-1})]}{[\Phi(\xi_i) - \Phi(\xi_{i-1})]^2} \right\} \right)$$

Ahora bien

$$\begin{aligned}
& [\Phi(\xi_i) - \Phi(\xi_{i-1})] \frac{\partial^2 \ln[\Phi(\xi_i) - \Phi(\xi_{i-1})]}{\partial \mu \partial \sigma} = \frac{\Phi(\xi_i) - \Phi(\xi_{i-1})}{\sigma^2} \left\{ \frac{\phi(\xi_i) - \phi(\xi_{i-1})}{\Phi(\xi_i) - \Phi(\xi_{i-1})} \right. \\
& \quad \left. - \frac{[\Phi(\xi_i) - \Phi(\xi_{i-1})][\xi_i^2 \phi(\xi_i) - \xi_{i-1}^2 \phi(\xi_{i-1})] + [\phi(\xi_i) - \phi(\xi_{i-1})][\xi_i \phi(\xi_i) - \xi_{i-1} \phi(\xi_{i-1})]}{[\Phi(\xi_i) - \Phi(\xi_{i-1})]^2} \right\} \\
& = \frac{1}{\sigma^2} \left\{ \phi(\xi_i) - \phi(\xi_{i-1}) - [\xi_i^2 \phi(\xi_i) - \xi_{i-1}^2 \phi(\xi_{i-1})] - \frac{[\phi(\xi_i) - \phi(\xi_{i-1})][\xi_i \phi(\xi_i) - \xi_{i-1} \phi(\xi_{i-1})]}{[\Phi(\xi_i) - \Phi(\xi_{i-1})]} \right\} \quad (3.28)
\end{aligned}$$

Por lo tanto substituyendo (3.28) en (3.25)

$$\begin{aligned}
AsCov(\hat{\mu}, \hat{\sigma}) &= \frac{\sigma^2}{n} \left[\sum_{i=1}^m \left\{ [\phi(\xi_i) - \phi(\xi_{i-1})] + [\xi_i^2 \phi(\xi_i) - \xi_{i-1}^2 \phi(\xi_{i-1})] \right. \right. \\
& \quad \left. \left. + \frac{[\phi(\xi_i) - \phi(\xi_{i-1})][\xi_i \phi(\xi_i) - \xi_{i-1} \phi(\xi_{i-1})]}{[\Phi(\xi_i) - \Phi(\xi_{i-1})]} \right\} \right]^{-1} \quad (3.29)
\end{aligned}$$

Por último, al considerar las ecuaciones (3.5), sabemos que la matriz de información asintótica de Fisher es:

$$AsVar(\hat{\theta}) = AsVar \begin{bmatrix} \hat{\mu} \\ \hat{\sigma} \end{bmatrix} = \begin{bmatrix} AsVar(\hat{\mu}) & AsCov(\hat{\mu}, \hat{\sigma}) \\ AsCov(\hat{\mu}, \hat{\sigma}) & AsVar(\hat{\sigma}) \end{bmatrix}$$

en la que al utilizar las expresiones (3.20), (3.24) y (3.29) resultantes; produce

$$AsVar(\hat{\theta}) = AsVar \begin{bmatrix} \hat{\mu} \\ \hat{\sigma} \end{bmatrix} = \frac{\sigma^2}{n} / \begin{bmatrix} a_1 & a_2 \\ a_2 & a_3 \end{bmatrix} \quad (3.30)$$

en donde

$$a_1 = \sum_{i=1}^m \left\{ \left[\xi_i \phi(\xi_i) - \xi_{i-1} \phi(\xi_{i-1}) \right] + \frac{[\phi(\xi_i) - \phi(\xi_{i-1})]^2}{\Phi(\xi_i) - \Phi(\xi_{i-1})} \right\} \quad (3.30.1)$$

$$a_3 = \sum_{i=1}^m \left\{ \left[\xi_i \phi(\xi_i) [\xi_i^2 - 1] - \xi_{i-1} \phi(\xi_{i-1}) [\xi_{i-1}^2 - 1] \right] + \frac{[\xi_i \phi(\xi_i) - \xi_{i-1} \phi(\xi_{i-1})]^2}{[\Phi(\xi_i) - \Phi(\xi_{i-1})]} \right\} \quad (3.30.2)$$

$$a_2 = \sum_{i=1}^m \left\{ \phi(\xi_i) - \phi(\xi_{i-1}) - [\xi_i^2 \phi(\xi_i) - \xi_{i-1}^2 \phi(\xi_{i-1})] - \frac{[\phi(\xi_i) - \phi(\xi_{i-1})][\xi_i \phi(\xi_i) - \xi_{i-1} \phi(\xi_{i-1})]}{[\Phi(\xi_i) - \Phi(\xi_{i-1})]} \right\} \quad (3.30.3)$$

Al valor numérico resultante de las expresiones (3.30.1), (3.30.2) y (3.30.3) se le conoce como Información Relativa Asintótica la cual representa el total de información que se puede obtener por el agrupamiento elegido.

4. METODOLOGÍA

Una vez presentados los resultados teóricos de estimación en donde se encuentran los estimadores y las varianzas asintóticas de los mismos y, con el fin de cumplir con los objetivos expuestos, este capítulo está dividido en dos partes:

En primer lugar: Se determinan los límites de truncamiento óptimos subyacentes cuando, simultáneamente, se desean estimar la media y la desviación estándar de una población normal. Lo anterior para $m = 2, 3, \dots, 11$ intervalos y utilizando un método numérico.

En segundo lugar: Dados los puntos de truncamiento óptimos, se generaron cuatro muestras de una población normalmente distribuida con media cincuenta y desviación estándar diez. Las muestras generadas son de tamaños: 50, 100, 200 y 400; mismas que se agrupan, cada una, en cada uno de los $m = 2, 3, \dots, 11$ intervalos. Obteniéndose, posteriormente, la información relativa asintótica a partir de dichos estimadores y límites.

4.1. Agrupamiento Óptimo

El agrupamiento óptimo será aquel resultante de minimizar la matriz de información asintótica expuesta en (3.30). Ogawa (1977) demuestra que tal matriz es definida negativa y; por tanto, solo es necesario maximizar su determinante; el cual es dado por:

$$\Delta = a_1 \times a_3 - [a_2]^2 \quad (4.0)$$

donde a_1 , a_2 y a_3 son dados por las expresiones (3.30.1), (3.30.2) y (3.30.3). Así; los límites óptimos son dados por los valores ξ_i que maximizan Δ . Tales valores son obtenidos siguiendo el *método numérico* llamado Cuasi Método de Newton, cuya rutina se encuentra programada en el software libre **R versión 2.5.1**. el cual requiere de una semilla – o valor inicial – misma que se encontró por ensayo y error. Cada una de las ξ_i resultantes se definen a ser los ***i* – ésimos límites de truncamiento óptimos**, los cuales nos permiten calcular **la información relativa asintótica**.

4.2. Estimadores de Máxima Verosimilitud

Para obtener estimados de máxima verosimilitud es necesario resolver simultáneamente el sistema de ecuaciones (3.17). Note que este sistema está en función de tres variables desconocidas: 1) los parámetros de la población μ y σ , 2) el número de unidades en el *i* – ésimo intervalo y 3) los límites de truncamiento – los cuales están en función de los parámetros desconocidos –. La solución de este sistema se obtiene siguiendo el *método numérico* llamado

Cuasi Método de Newton, cuya rutina se encuentra programada en el software libre **R** versión 2.5.1. necesitándose de una semilla – o valor inicial – θ_n cercano al valor real que se busca para cada uno de los parámetros desconocidos.

Estos valores iniciales – por la naturaleza de nuestro problema – se eligieron a ser la media y la varianza muestral para datos agrupados utilizados en estadística descriptiva, definiéndose a ser respectivamente:

$$\tilde{\mu} = \frac{\sum_{i=1}^m n_i u_i}{\sum_{i=1}^m n_i} \tag{4.1}$$

$$\tilde{\sigma}^2 = \frac{\sum_{i=1}^m n_i u_i^2}{\sum_{i=1}^m n_i} - (\tilde{\mu}^2)$$

donde:

$u_i = \frac{h_i + h_{i+1}}{2}$, es el valor medio del intervalo o clase ($i = 2, 3, \dots, m$),

n_i : es el número de observaciones en el i –ésimo intervalo

Para el cálculo de (4.1) necesitamos realizar un agrupamiento previo – establecer límites de truncamiento *a priori* o iniciales –. Este primer agrupamiento puede determinarse utilizando como límites los establecidos por la relación numérica:

$$\tilde{h}_i = y_{(1)} + i \times \frac{y_{(n)} - y_{(1)}}{m} \quad (4.2)$$

donde:

m es el número de intervalos,

$$y_{(1)} = \text{menor} \{y_i\}_{i=1}^n,$$

$$y_{(n)} = \text{mayor} \{y_i\}_{i=1}^n,$$

$\tilde{h}_i ; i = 1, 2, \dots, m$: es el i – éximo límite de truncamiento inicial.

Note de esta manera que la metodología a seguir para encontrar las estimaciones de máxima verosimilitud es la siguiente:

1. Se obtienen los *límites de truncamiento a priori* con la expresión (4.2) y, posteriormente, se realiza un **agrupamiento inicial** calculando la **información relativa asintótica**, substituyendo los valores $\xi_i = (\tilde{h}_i - \tilde{\mu}) / \tilde{\sigma}$ en (4.0). A este agrupamiento lo referimos como **tradicional**.

2. De las expresiones definidas en (3.17) y con la información del agrupamiento inicial se obtienen los **estimados de máxima verosimilitud iniciales** $\hat{\mu}^0$ y $\hat{\sigma}^0$, y además sustituyendo los valores $\xi_i = (\tilde{h}_i - \hat{\mu}^0) / \hat{\sigma}^0$ en (4.0), se obtiene la **información relativa asintótica**, la cual referiremos como **inicial**.

3. Con los **límites de truncamiento óptimos**, ξ_i – tal y como aparecen en el **Cuadro 1** del capítulo siguiente –, se calculan los límites de truncamiento $h_i = \hat{\sigma}^0 \xi_i + \hat{\mu}^0$ ($i = 1, 2, \dots, m$) y se realiza un agrupamiento considerado esta vez como **agrupamiento óptimo**.

4. Con la información del *agrupamiento óptimo* y usando (3.17) se obtienen los **estimados de máxima verosimilitud óptimos** $\hat{\mu}$ y $\hat{\sigma}$ y de manera análoga la **información relativa asintótica óptima**, siendo esta vez $\xi_i = (h_i - \hat{\mu}) / \hat{\sigma}$. Note que $\hat{\mu}^0$ y $\hat{\sigma}^0$ se utilizan como semillas al aplicar el método numérico.

Por último y con la finalidad de cumplir con los objetivos planteados se eligieron dos indicadores que permiten comparar la información obtenida en cada uno de los pasos de la metodología indicada arriba, siendo estos:

Error Cuadrado Medio

Un criterio aceptado para realizar la comparación entre un parámetro de interés contra el estimador obtenido – MLE – es utilizando el Error Cuadrado Medio – ECM – el cual se define a

ser la esperanza del cuadrado de la diferencia entre el parámetro de interés y el estimador deseado; sin embargo, en nuestro caso particular, el estimador que se obtiene es un número y dado que la muestra ha sido generada aleatoriamente tendremos dos valores numéricos, *i.e.*, el ECM se convierte en una diferencia de cuadrados, *DC*:

$$ECM(\underline{\theta}) = E\left\{\left[\hat{\underline{\theta}} - \underline{\theta}\right]^2\right\} = \left[\ell(\hat{\underline{\theta}}) - \underline{\theta}\right]^2 = DC \quad (4.3)$$

donde: $\hat{\underline{\theta}}$ es el $MLE(\underline{\theta})$, $\ell(\hat{\underline{\theta}})$ su valor numérico y $\underline{\theta}' = [\mu, \sigma]$.

Eficiencia del Diseño de Muestreo *deff*

Este indicador, propuesto por Kish (1967), denominado Eficiencia del Diseño de Muestreo (*deff*) y cuya estructura es un cociente de varianzas, se propone con el fin de conocer la bondad de los estimadores y límites calculados en esta parte del estudio, definiéndose a éste como sigue:

$$deff_{p(s)}(\hat{\theta}) = \frac{V_{a(a)}(\hat{\theta})}{V_{a(b)}(\hat{\theta})} \quad (4.3)$$

donde:

$deff_{p(s)}(\hat{\theta})$ Es el indicador de la eficiencia del diseño muestral.

$V_{a(a)}(\hat{\theta})$: Varianza del estimador con el método a comparar – la de los límites óptimos –.

$V_{a(b)}(\hat{\theta})$: Varianza del estimador con el que se compara – la de los valores generados –.

Considerándose un método más eficiente que otro, si la varianza asintótica obtenida a partir de las muestras generadas es mayor que la obtenida por los límites óptimos, *i.e.*, si el *deff* es mayor que uno quiere decir que la información del agrupamiento utilizado en el denominador es menor, esto es: el agrupamiento es menos eficiente; si el *deff* es igual que uno, quiere decir que la información obtenida por ambos métodos es la misma, por tanto, ambos agrupamientos son igual de eficientes; si el *deff* es menor que uno quiere decir que la información del agrupamiento utilizado en el denominador es mayor, es decir, el agrupamiento es más eficiente.

4.3. Lenguaje de Programación Utilizado

La generación de las muestras, su agrupamiento, la optimización de funciones y el cálculo final de varianzas y gráficas se realizaron con el paquete estadístico **R** *Versión 2.5.1* (The R Foundation for Statistical Computing). En el **Anexo I.** y **II.** se presentan los códigos de programación utilizados en la solución y cumplimiento de los objetivos.

5. RESULTADOS Y DISCUSIÓN

En el presente capítulo se presentan y discuten brevemente los resultados obtenidos siguiendo el orden y la metodología discutida en los capítulos 3 y 4 precedentes.

5.1. Límites de Truncamiento Óptimo

Siguiendo lo expuesto en el capítulo anterior al resolver la ecuación (4.0) se generan los límites de truncamiento óptimos para $m = 2, 3, \dots, 11$; calculando para cada uno de ellos la información relativa asintótica óptima, misma que se presenta en el cuadro siguiente:

Cuadro 1. Límites de Truncamiento Asintóticamente Óptimos y la Información Relativa Asintótica.

m	ξ_1	ξ_2	ξ_3	ξ_4	ξ_5	ξ_6	ξ_7	ξ_8	ξ_9	ξ_{10}	Información Relativa Asintótica Óptima
2	-1.17086										0.19522
3	-1.10706	1.10706									0.82210
4	-1.37620	0.01078	1.38437								1.11518
5	-1.69077	-0.68655	0.68655	1.69077							1.37373
6	-1.87497	-0.99368	-0.00001	0.99367	1.87496						1.51911
7	-2.05046	-1.25970	-0.48930	0.48930	1.25970	2.05046					1.62745
8	-2.18327	-1.44881	-0.78278	0.00002	0.78280	1.44883	2.18335				1.70083
9	-2.30301	-1.61320	-1.01712	-0.38044	0.38044	1.01712	1.61320	2.30301			1.75588
10	-2.40321	-1.74725	-1.19800	-0.64596	0.00003	0.64595	1.19798	1.74728	2.40342		1.79668
11	-2.49282	-1.86501	-1.35150	-0.85660	-0.31195	0.31195	0.85660	1.35150	1.86501	2.49282	1.82831

Donde las $\xi_i = (h_i - \mu) / \sigma ; i = 1, 2, \dots, 11$ son los límites de truncamiento óptimos. La Información Relativa Asintótica, representa el valor numérico adjunto a la Varianza Asintótica obtenida al sustituir directamente los límites de truncamiento asintóticamente óptimos en la expresión (4.0).

Observando los valores de la Información Relativa Asintótica resulta claro que la información que se puede obtener de una muestra agrupada aumentará conforme el número de límites de truncamiento sea mayor.

Por lo que hace a la simetría de los límites de truncamiento, los resultados de Kulldorff (1961) así lo prueban en el caso particular en el caso particular donde solamente la media es desconocida, en el otro caso particular, donde se desconoce la desviación estándar, sus resultados mostraron no ser simétricos para todos los valores de m pares que él considero – de hecho solo los números impares lo son –.

Ogawa (1977) a más de obtener los resultados de Kulldorff (1961) referentes a la estimación solo de μ y, cuando hace referencia a la estimación solo de σ , explicita que no necesariamente los límites de truncamiento deber ser simétricos y, cuando se hace la estimación simultanea sólo *conjetura* sobre la simetría de los mismos. Al observar el **Cuadro 1** que hemos generado, se tiene que *prácticamente* para todo agrupamiento hay simetría en los límites de truncamiento óptimos. Estos límites de truncamiento son aquellos con los que se obtendrá la mayor información de la población en estudio, sirviendo de punto de partida cuando este método es llevado a la práctica.

5.2. Estimación Simultánea de la Media y la Desviación Estándar.

La sección anterior tuvo que ver con el problema de encontrar el número y valor de los límites de truncamiento óptimos para diferente número de intervalos, de tal forma que estos permitan coleccionar la mayor información posible de la muestra. Esta sección tiene ahora que ver con la estimación de la media y la desviación estándar que se derivaron en un capítulo anterior, considerando para ello la metodología expuesta con anterioridad.

Los estimadores considerados fueron los de máxima verosimilitud, cuyas ecuaciones no son fáciles de resolver, además de carecer de una forma cerrada que nos permitiera la sustitución directa de valores. Por esta razón es necesario realizar un estudio numérico generando aleatoriamente cuatro muestras diferentes con el objetivo de estudiar los efectos del agrupamiento para tamaños grandes y pequeños, el cual nos permita determinar qué tan buenos son estos estimadores. Como se expuso en el capítulo anterior, se eligieron dos indicadores para cumplir con este objetivo, el primero fue el error cuadrado medio – DC –, y el otro fue la Eficiencia del Diseño de Muestreo – $deff$ –. Otra parte importante fue el graficar la función de verosimilitud con el fin de identificar visualmente los puntos donde la función alcanza un máximo – puntos donde se supone se encuentran los estimadores de máxima verosimilitud –. Otra utilidad del gráfico es, sin duda, mostrar gráficamente el máximo de la función probando de esta manera los teoremas y condiciones expuestas en un **Capítulo 3**.

5.2.1. Muestra de Tamaño $n = 50$

La **Figura 1** es la gráfica de la función de verosimilitud bajo las condiciones planteadas al principio de este capítulo. La figura muestra que la función de verosimilitud tiene un máximo cercano a los valores de los parámetros considerados al generar la muestra, de tal forma será posible resolver y encontrar los parámetros de máxima verosimilitud, además dichos valores son únicos.

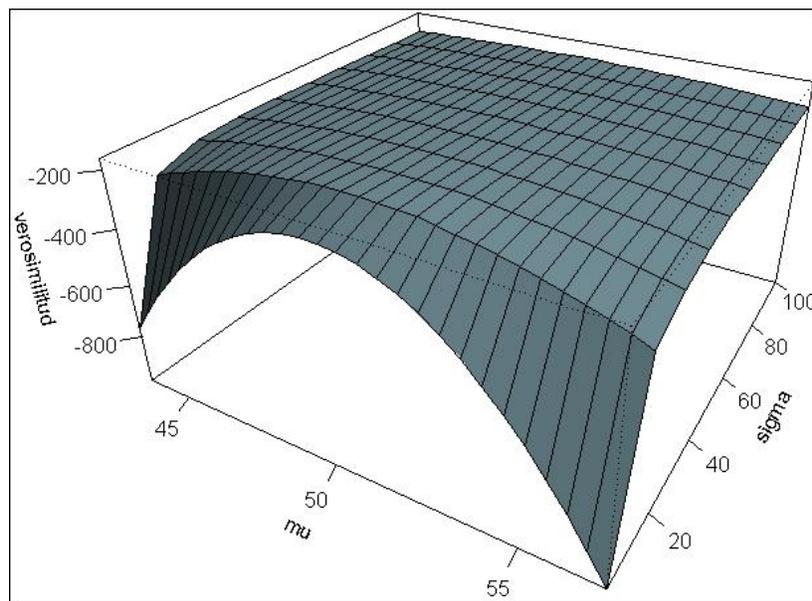
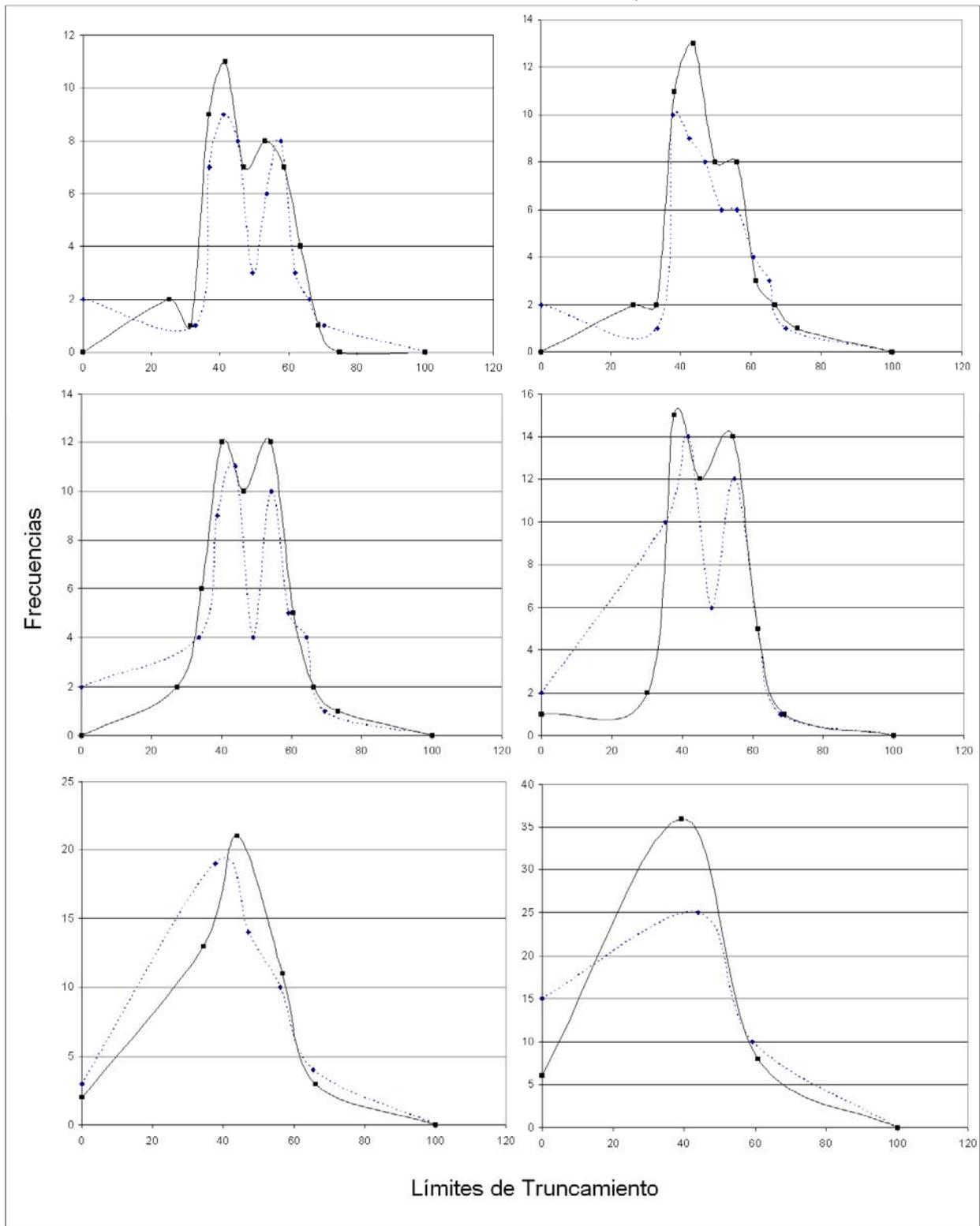


Figura 1. Función de Verosimilitud para función de densidad de probabilidad normal para una muestra de tamaño cincuenta y parámetros $\mu = 50$ y $\sigma = 10$.

La **Figura 2** muestra los histogramas para el agrupamientos tradicional y el agrupamiento óptimo obtenidos mediante el método tradicional de estadística descriptiva y los límites y frecuencias óptimos obtenidos según la metodología descrita en el capítulo anterior, utilizados para maximizar el sistema (3.17), para $m = 11, 10, 9, 7, 5, 3$ límites.

Figura 2. *Agrupamiento Tradicional* $\cdots\blacklozenge\cdots$ y *Agrupamiento Óptimo* $\text{---}\blacksquare\text{---}$ para una muestra de tamaño cincuenta obtenida de una población normal con parámetros $\mu=50$ y $\sigma=10$



En el **Cuadro 2** se muestran los límites de truncamiento y frecuencias de grupo tradicionales y óptimos obtenidos según la metodología descrita en el capítulo anterior, utilizados en la construcción de los histogramas mostrados en la **Figura 2**.

Cuadro 2. Agrupamientos Inicial y Óptimo para una muestra de tamaño cincuenta obtenida de una población normal con parámetros $\mu = 50$ y $\sigma = 10$

m	Inicial		Óptimo	
	h_i	n_i	h_i	n_i
2	51.5458	30	46.6638	21
		20		29
3	43.8791	15	39.2720	6
	59.2124	25	60.7752	36
		10		8
4	40.0458	8	36.6541	3
	51.5458	22	50.0473	25
	63.0458	14	63.3113	17
		6		5
5	37.7458	3	34.4326	2
	46.9458	19	43.8350	13
	56.1458	14	56.6909	21
	65.3458	10	66.0932	11
		4		3
6	36.2124	3	31.7974	2
	43.8791	12	40.3599	7
	51.5458	15	50.0143	19
	59.2124	10	59.6689	12
	66.8791	7	68.2314	9
		3		1
7	35.1172	2	30.1156	1
	41.6886	10	37.6202	2
	48.2600	14	44.9314	15
	54.8315	6	54.2187	12
	61.4029	12	61.5299	14
	67.9743	5	69.0345	5
		1		1
8	34.2958	2	28.4558	0
	40.0458	6	35.6819	2
	45.7958	12	42.2347	11
	51.5458	10	49.9365	15
	57.2958	6	57.6380	8
	63.0458	8	64.1908	9

	68.7958	5	71.4175	4
		1		1
9	33.6569	2	27.3050	0
	38.7680	4	34.1678	2
	43.8791	9	40.0981	6
	48.9902	11	46.4323	12
	54.1013	4	54.0022	10
	59.2124	10	60.3364	12
	64.3235	5	66.2666	5
	69.4346	4	73.1294	2
	1		1	
10	33.1458	2	26.4989	0
	37.7458	1	32.8836	2
	42.3458	10	38.2296	2
	46.9458	9	43.6028	11
	51.5458	8	49.8905	13
	56.1458	6	56.1775	8
	60.7458	6	61.5505	8
	65.3458	4	66.8971	3
	69.9458	3	73.2835	2
	1		1	
11	32.7276	2	25.2331	0
	36.9094	1	31.5018	2
	41.0912	7	36.6293	1
	45.2730	9	41.5710	9
	49.4548	8	47.0094	11
	53.6367	3	53.2392	7
	57.8185	6	58.6775	8
	62.0003	8	63.6192	7
	66.1821	3	68.7467	4
	70.3639	2	75.0154	1
		1		0

El **Cuadro 3** presenta los valores de la media y la desviación estándar poblacional utilizados como semillas y los valores de los estimadores obtenidos de aplicar el método numérico tal y como se describe en la metodología del capítulo anterior. Las cuatro últimas columnas contienen los valores del indicador *DC* descrito en el capítulo anterior por (4.2).

Cuadro 3. Semillas, Estimadores Iniciales y Óptimos y el indicador DC para una muestra de tamaño cincuenta obtenidas de una población normal con parámetros $\mu = 50$ y $\sigma = 10$

m	Tradicional		Inicial		Óptimo		$DC(\tilde{\mu})$	$DC(\hat{\mu}^0)$	$DC(\hat{\mu})$	$DC(\tilde{\sigma})$	$DC(\hat{\sigma}^0)$	$DC(\hat{\sigma})$
	$\tilde{\mu}$	$\tilde{\sigma}$	$\hat{\mu}^0$	$\hat{\sigma}^0$	$\hat{\mu}$	$\hat{\sigma}$						
2	49.2530	11.2766	50.6773	3.4278	49.7887	3.5073	0.5580	0.4587	0.0446	1.6297	43.1938	42.1552
3	50.0161	10.7390	50.0236	9.7119	50.9183	9.9118	0.0003	0.0006	0.8433	0.5462	0.0830	0.0078
4	49.9379	10.2280	49.9433	9.6564	49.7845	9.4801	0.0039	0.0032	0.0464	0.0520	0.1180	0.2703
5	50.2592	9.7410	50.2629	9.3628	50.2507	10.1245	0.0672	0.0691	0.0628	0.0671	0.4061	0.0155
6	50.0135	9.9689	50.0145	9.7159	50.0145	9.7159	0.0002	0.0002	0.0002	0.0010	0.0807	0.0807
7	49.5747	9.6811	49.5751	9.4903	50.3573	9.6706	0.1809	0.1806	0.1277	0.1017	0.2598	0.1085
8	49.9361	9.9793	49.9363	9.8387	50.1980	10.1675	0.0041	0.0041	0.0392	0.0004	0.0260	0.0281
9	50.2172	10.0586	50.2172	9.9488	50.5013	10.1513	0.0472	0.0472	0.2513	0.0034	0.0026	0.0229
10	49.8901	9.8242	49.8902	9.7333	49.8756	10.0245	0.0121	0.0121	0.0155	0.0309	0.0711	0.0006
11	50.1243	10.0584	50.1243	9.9851	49.8657	10.1730	0.0155	0.0154	0.0180	0.0034	0.0002	0.0299

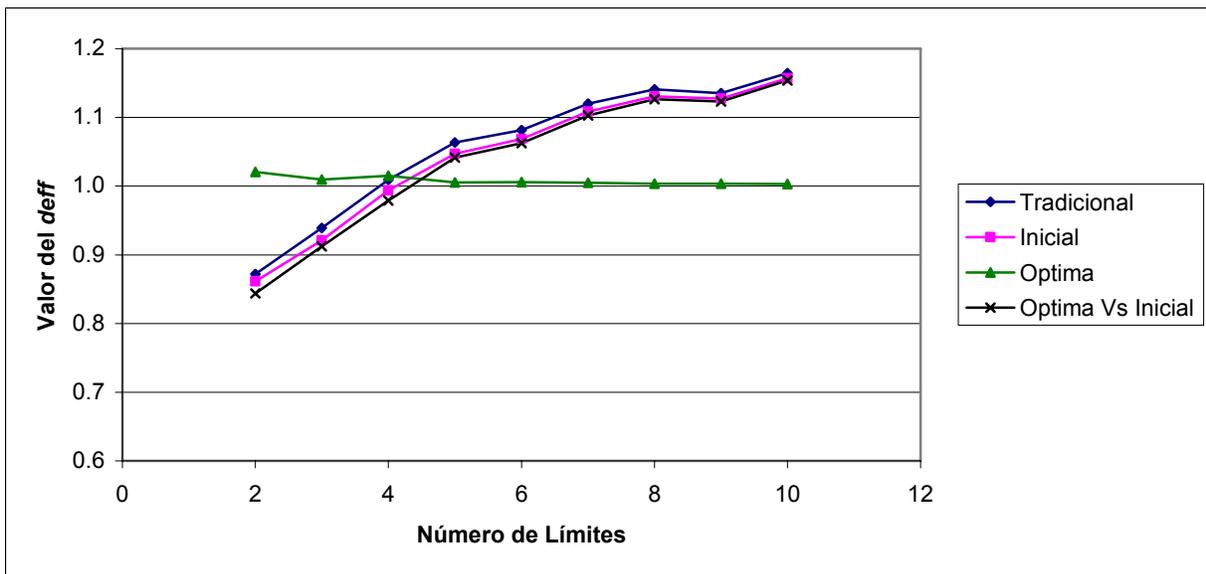
El **Cuadro 4** presenta la información relativa asintótica obtenida para los estimados de máxima verosimilitud iniciales y óptimos obtenida al sustituir directamente estos valores en la ecuación (3.32). Las últimas dos columnas contienen los valores del indicador $deff$ definido en el capítulo anterior.

Cuadro 4. Información Relativa Asintótica y Eficiencia del Muestreo para una muestra de tamaño cincuenta obtenida de una población normal con parámetros $\mu = 50$ y $\sigma = 10$

m	Óptima	Tradicional ($\tilde{\mu}, \tilde{\sigma}$)	Inicial ($\hat{\mu}^0, \hat{\sigma}^0$)	Final ($\hat{\mu}, \hat{\sigma}$)	$deff$				
					($\tilde{\mu}, \tilde{\sigma}$)	($\hat{\mu}^0, \hat{\sigma}^0$)	($\hat{\mu}, \hat{\sigma}$)	($\hat{\mu}, \hat{\sigma}$)/($\hat{\mu}^0, \hat{\sigma}^0$)	($\hat{\mu}, \hat{\sigma}$)/($\tilde{\mu}, \tilde{\sigma}$)
2	0.1952	0.4927	0.0191	0.0122	0.3962	10.2179	15.9620	0.6401	0.0248
3	0.8221	0.9426	0.9546	0.8055	0.8722	0.8612	1.0206	0.8439	0.8546
4	1.1152	1.1874	1.2109	1.1047	0.9392	0.9209	1.0094	0.9123	0.9304
5	1.3737	1.3616	1.3827	1.3537	1.0089	0.9935	1.0148	0.9790	0.9942
6	1.5191	1.4284	1.4510	1.5113	1.0635	1.0469	1.0052	1.0416	1.0580
7	1.6275	1.5047	1.5228	1.6179	1.0815	1.0687	1.0059	1.0624	1.0752
8	1.7008	1.5183	1.5347	1.6926	1.1202	1.1083	1.0049	1.1029	1.1148
9	1.7559	1.5392	1.5531	1.7496	1.1408	1.1306	1.0036	1.1265	1.1366
10	1.7967	1.5826	1.5940	1.7905	1.1352	1.1271	1.0034	1.1233	1.1313
11	1.8283	1.5705	1.5805	1.8234	1.1642	1.1568	1.0027	1.1537	1.1610

La **Figura 3** muestra de forma gráfica el comportamiento del indicador $deff$ para los límites de truncamiento del método tradicional – semillas –, los de máxima verosimilitud iniciales, los de máxima verosimilitud óptimos y los de máxima verosimilitud óptimos contra los iniciales. Tomando como referencia el valor de la información relativa asintótica de los límites de truncamiento óptimos la cual aparece en el **Cuadro 1**.

Figura 3. Índices de la Eficiencia del Diseño de Muestreo para una muestra de tamaño cincuenta generada con distribución normal y parámetros $\mu = 50$ y $\sigma = 10$.



Recordando que DC es una medida que muestra que tan alejado se encuentran nuestros valores estimados de los valores de los parámetros utilizados para generar la muestra, siendo deseable que este valor sea menor que uno. Los valores de este indicador mostrados en el **Cuadro 3** indican que las estimaciones de máxima verosimilitud obtenidas a partir de los límites de truncamiento óptimos son bastante buenas en general dado que, la mayoría de estos valores son menores que uno, de hecho, sólo un valor para la media y tres para la desviación estándar son

más grandes que uno. Cabe agregar que cuando $m=2$ el estimado de máxima verosimilitud para la desviación estándar dista con mucho del valor del parámetro. En general a partir de cuatro límites de truncamiento los valores de los estimadores de máxima verosimilitud pueden considerarse buenos. Una rápida inspección a los valores de la media y la desviación estándar para todos los métodos nos indica que todos son buenos dado que, todos están muy cerca de los valores utilizados para generar las muestras, sin embargo, un criterio de desempate puede darnoslo el **Cuadro 2**, al inspeccionar los valores de las frecuencias obtenidas por el método tradicional y los límites óptimos, se observa que los segundos tienen más valores de las frecuencias concentrados en la media, permitiéndonos formar una campana de gauss más clara – que es la forma de la distribución utilizada para generar las muestras –. De esta manera no obstante que las estimaciones del método tradicional se parecen a las de los límites óptimos, el agrupamiento obtenido a partir de los segundos es mejor, siendo preferible de esta manera utilizar los límites óptimos a los del método tradicional. Además observando los valores *deff* del **Cuadro 4** y su comportamiento en la **Figura 3**, partiendo del criterio definido en el capítulo anterior, observamos que la información relativa asintótica obtenida de los límites derivados de los estimados de máxima verosimilitud finales es tanta o tan buena como la de los límites óptimos, dado que a partir de cinco límites de truncamiento el valor de este indicador para los estimados óptimos es casi uno. Caso contrario ocurre para los límites y estimados del método tradicional e inicial, donde los valores de este indicador son en su mayoría más grandes que uno.

Por tanto el agrupamiento y los estimadores del método propuesto en este trabajo son más eficientes – o mejores – en el sentido de que de ellos la muestra arroja más información sobre la población en estudio.

5.2.2. Muestra de Tamaño $n = 100$

La **Figura 4** muestra la gráfica de la función de verosimilitud para la segunda muestra bajo las condiciones planteadas al principio de este capítulo. En la figura se muestra que la función de verosimilitud tiene un máximo cercano a los valores de los parámetros considerados al generar la muestra, de tal forma será posible resolver y encontrar los parámetros de máxima verosimilitud, además dichos valores serán únicos.

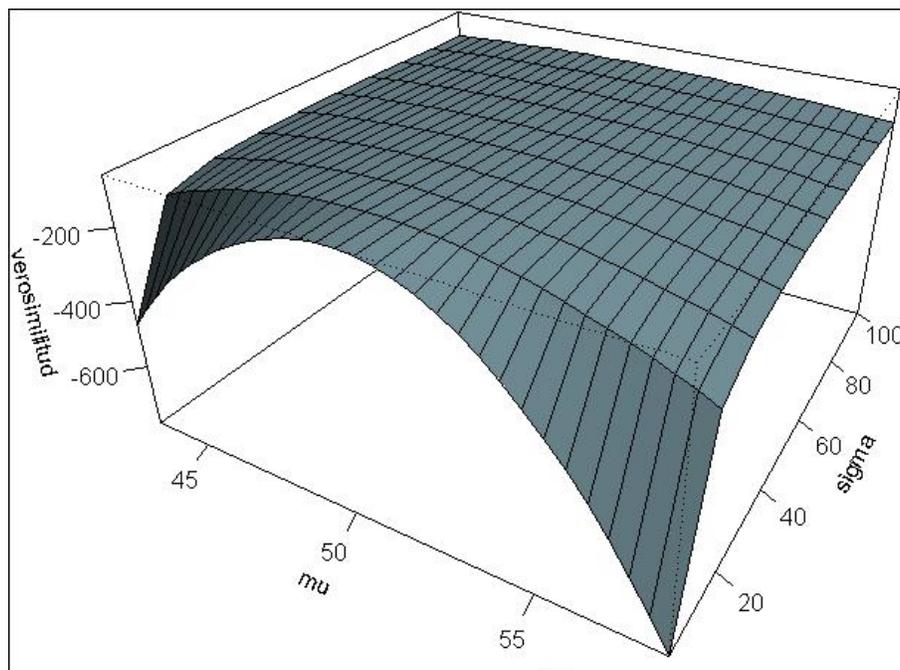
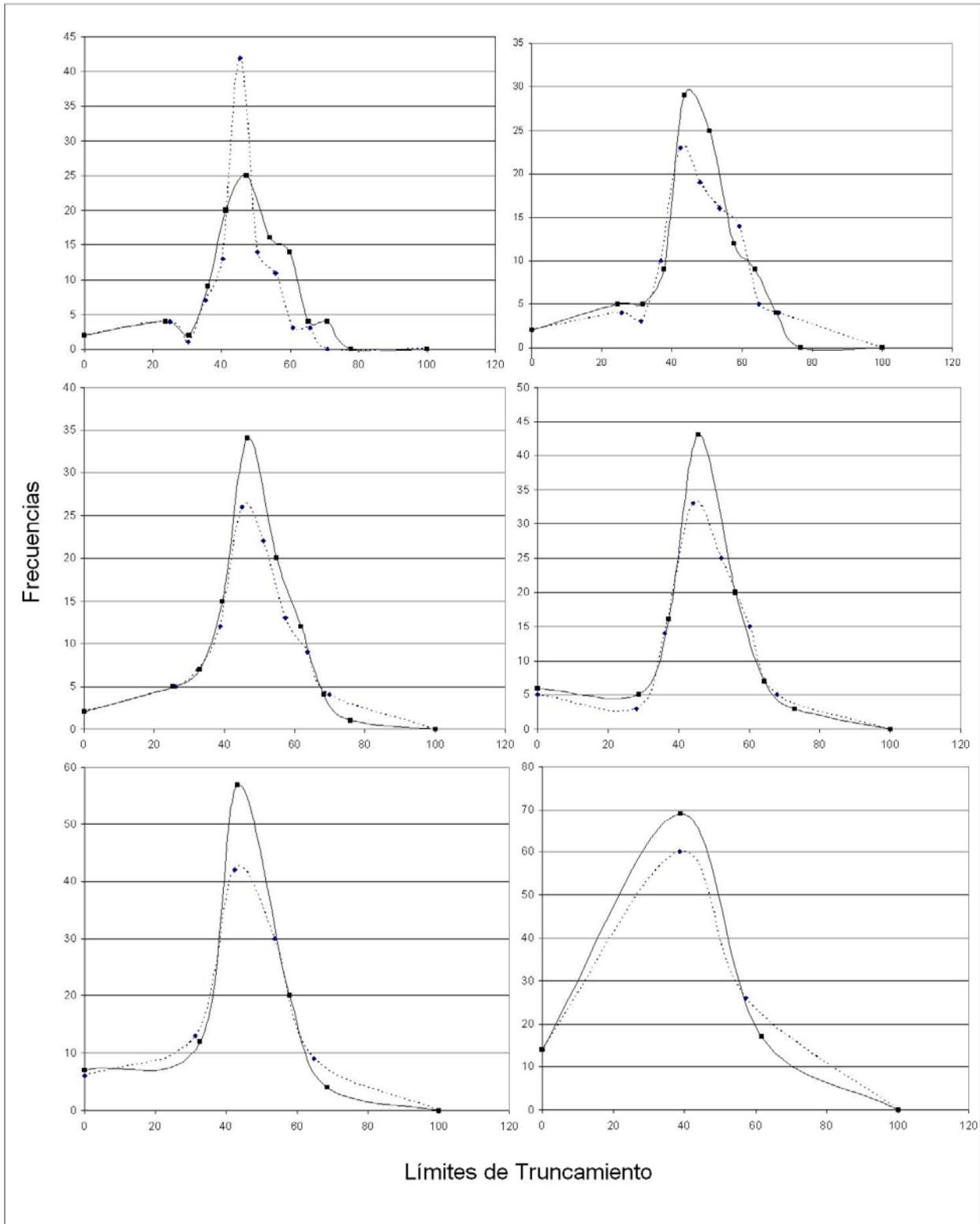


Figura 4. Función de Verosimilitud para función de densidad de probabilidad normal para una muestra de tamaño cien y parámetros $\mu = 50$ y $\sigma = 10$

La **Figura 5** muestra los histogramas para el agrupamiento tradicional y el agrupamiento óptimo obtenidos mediante el método tradicional de estadística descriptiva y los límites y frecuencias óptimos obtenidos según la metodología descrita en el capítulo anterior, utilizados para maximizar el sistema (3.17), para $m = 11, 10, 9, 7, 5, 3$ límites.

Figura 5. *Agrupamiento Tradicional* $\cdots\blacklozenge\cdots$ y *Agrupamiento Óptimo* $\text{---}\blacksquare\text{---}$ para una muestra de tamaño cien obtenida de una población normal con parámetros $\mu = 50$ y $\sigma = 10$



En el **Cuadro 5** se muestran los límites de truncamiento y frecuencias de grupo iniciales obtenidos mediante el método tradicional de estadística descriptiva y los límites y frecuencias óptimos obtenidos según la metodología descrita en el capítulo anterior, utilizados para maximizar el sistema (3.17).

Cuadro 5. Agrupamientos Inicial y Óptimo para una muestra de tamaño cien obtenida de una población normal con parámetros $\mu = 50$ y $\sigma = 10$

m	Inicial		Óptimo	
	\tilde{h}_i	n_i	h_i	n_i
2	48.01319	42	43.8931	22
		58		78
3	38.67985	14	38.9093	14
	57.34652	60	61.6019	69
		26		17
4	34.01319	8	35.5126	8
	48.01319	22	50.5133	42
	62.01319	14	65.3692	42
		6		8
5	31.21319	6	32.6724	7
	42.4132	13	43.3134	12
	53.6132	42	57.8629	57
	64.8132	30	68.5039	20
		9		4
6	29.34652	6	29.1158	6
	38.6799	8	38.8760	8
	48.0132	28	49.8808	35
	57.3465	32	60.8857	34
	66.6799	20	70.6459	13
		6		4
7	28.01319	5	28.7668	6
	36.0132	3	37.2695	5
	44.0132	14	45.5531	16
	52.0132	33	56.0756	43
	60.0132	25	64.3593	20
	68.0132	15	72.8620	7
		5		3
8	27.01319	2	26.9688	3
	34.0132	6	34.9201	4
	41.0132	12	42.1306	11
	48.0132	10	50.6053	32

	55.0132	6	59.0798	27
	62.0132	8	66.2902	16
	69.0132	5	74.2423	6
		1		1
9	26.23541	2	25.3849	2
	32.4576	5	32.9461	5
	38.6799	7	39.4799	7
	44.9021	12	46.4587	15
	51.1243	26	54.7989	34
	57.3465	22	61.7777	20
	63.5687	13	68.3115	12
	69.7910	9	75.8727	4
		4		1
10	25.61319	2	24.5175	2
	31.2132	4	31.6498	5
	36.8132	3	37.6218	5
	42.4132	10	43.6241	9
	48.0132	23	50.6481	29
	53.6132	19	57.6712	25
	59.2132	16	63.6734	12
	64.8132	14	69.6460	9
	70.4132	5	76.7802	4
	4		0	
11	25.1041	2	23.6477	2
	30.1950	4	30.4514	4
	35.2859	1	36.0165	2
	40.3768	7	41.3799	9
	45.4677	13	47.2824	20
	50.5586	42	54.0438	25
	55.6496	14	59.9462	16
	60.7405	11	65.3097	14
	65.8314	3	70.8747	4
	70.92228	3	77.6784	4
	0		0	

El **Cuadro 6** presenta los valores de la media y la desviación estándar poblacional utilizados como semillas y los valores de los estimadores obtenidos de aplicar el método numérico tal y como se describe en la metodología del capítulo anterior. Las cuatro últimas columnas contienen los valores del indicador *DC* descrito en el capítulo anterior.

Cuadro 6. Semillas, Estimadores Iniciales y Óptimos y el indicador DC para una muestra de tamaño cien obtenida de una población normal con parámetros $\mu = 50$ y $\sigma = 10$

m	Tradicional		Inicial		Óptimo		$DC(\tilde{\mu})$	$DC(\hat{\mu}^0)$	$DC(\hat{\mu})$	$DC(\tilde{\sigma})$	$DC(\hat{\sigma}^0)$	$DC(\hat{\sigma})$
	$\tilde{\mu}$	$\tilde{\sigma}$	$\hat{\mu}^0$	$\hat{\sigma}^0$	$\hat{\mu}$	$\hat{\sigma}$						
2	50.3137	13.8712	48.8716	4.2520	47.7804	5.0342	0.0984	1.2732	4.9264	14.9859	33.0397	24.6596
3	50.2803	11.6299	50.2556	10.2491	50.9592	11.1540	0.0786	0.0653	0.9200	2.6565	0.0620	1.3317
4	50.4099	11.5810	50.3967	10.8154	50.4814	10.6246	0.1680	0.1574	0.2317	2.4995	0.6649	0.3901
5	50.5986	11.0879	50.5882	10.5962	50.2507	10.1245	0.3583	0.3459	0.0628	1.1835	0.3554	0.0155
6	49.8861	11.4046	49.8809	11.0749	50.0844	11.8681	0.0130	0.0142	0.0071	1.9729	1.1554	3.4898
7	50.8184	11.0010	50.8144	10.7525	50.7136	11.6381	0.6698	0.6632	0.5093	1.0020	0.5663	2.6835
8	50.6074	11.0166	50.6051	10.8261	50.8236	10.8261	0.3689	0.3662	0.6783	1.0335	0.6824	0.6824
9	50.6307	11.1105	50.6288	10.9613	51.2250	11.1723	0.3978	0.3954	1.5006	1.2332	0.9240	1.3742
10	50.6494	10.9953	50.6477	10.8730	50.6135	11.1881	0.4217	0.4195	0.3764	0.9906	0.7622	1.4115
11	50.6646	10.9389	50.6631	10.8373	50.7307	11.2757	0.4417	0.4397	0.5339	0.8815	0.7010	1.6274

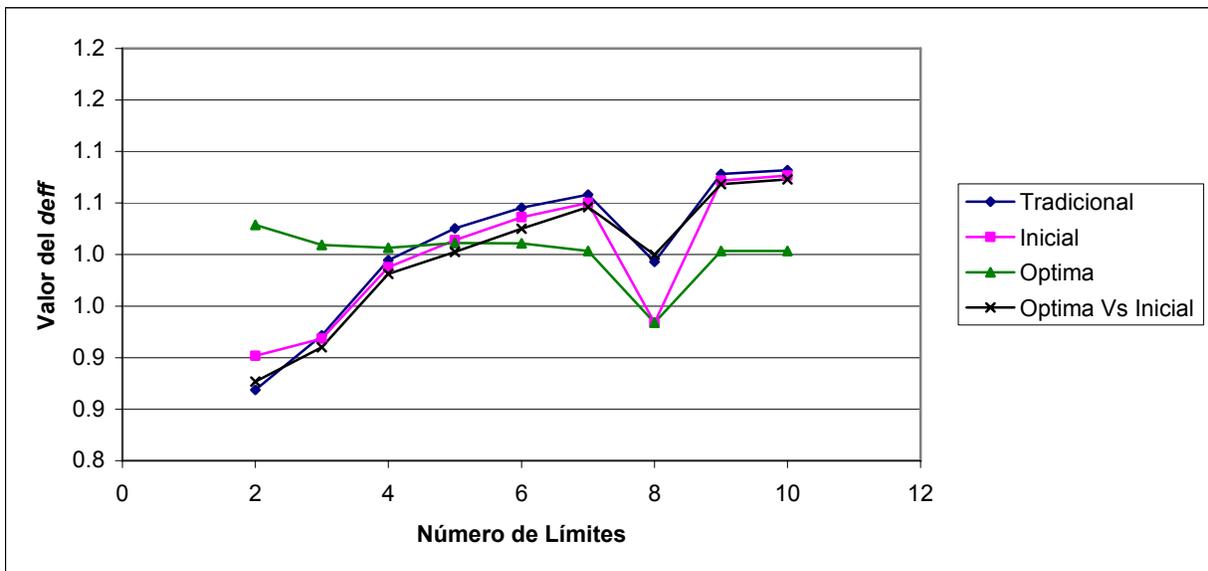
El **Cuadro 7** presenta la información relativa asintótica obtenida por los estimadores de máxima verosimilitud iniciales y óptimos obtenida al sustituir directamente estos valores en la ecuación (3.32). Las últimas dos columnas contienen los valores del indicador *deff* definido en el capítulo anterior.

Cuadro 7. Información Relativa Asintótica y Eficiencia del Muestreo para una muestra de tamaño cien obtenida de una población normal con parámetros $\mu = 50$ y $\sigma = 10$

m	Óptima	Tradicional ($\tilde{\mu}, \tilde{\sigma}$)	Inicial ($\hat{\mu}^0, \hat{\sigma}^0$)	Final ($\hat{\mu}, \hat{\sigma}$)	<i>deff</i>				
					($\tilde{\mu}, \tilde{\sigma}$)	($\hat{\mu}^0, \hat{\sigma}^0$)	($\hat{\mu}, \hat{\sigma}$)	($\hat{\mu}, \hat{\sigma}$)/($\hat{\mu}^0, \hat{\sigma}^0$)	($\hat{\mu}, \hat{\sigma}$)/($\tilde{\mu}, \tilde{\sigma}$)
2	0.1952	0.4952	0.0122	0.1373	0.3943	15.9618	1.4217	11.2270	0.2773
3	0.8221	0.9464	0.9115	0.7991	0.8686	0.9019	1.0288	0.8766	0.8443
4	1.1152	1.2104	1.2142	1.1049	0.9214	0.9185	1.0093	0.9100	0.9128
5	1.3737	1.3813	1.3908	1.3649	0.9945	0.9877	1.0065	0.9814	0.9881
6	1.5191	1.4814	1.4983	1.5022	1.0255	1.0139	1.0113	1.0026	1.0141
7	1.6275	1.5567	1.5703	1.6100	1.0455	1.0364	1.0108	1.0253	1.0343
8	1.7008	1.6073	1.6199	1.6946	1.0582	1.0500	1.0036	1.0462	1.0543
9	1.6333	1.6448	1.7492	1.7492	0.9930	0.9338	0.9338	1.0000	1.0635
10	1.7967	1.6666	1.6763	1.7906	1.0780	1.0718	1.0034	1.0682	1.0744
11	1.8283	1.6897	1.6981	1.8219	1.0820	1.0767	1.0035	1.0729	1.0782

La **Figura 6** muestra de forma gráfica el comportamiento del indicador *deff* para los límites de truncamiento del método tradicional – semillas –, los de máxima verosimilitud iniciales, los de máxima verosimilitud óptimos y los de máxima verosimilitud óptimos contra los iniciales. Tomando como referencia el valor de la información relativa asintótica de los límites de truncamiento óptimos la cual aparece en el **Cuadro 1**.

Figura 6. Índices de la Eficiencia del Diseño de Muestreo para una muestra de tamaño cien generada con distribución normal y parámetros $\mu = 50$ y $\sigma = 10$.



Los valores del indicador *DC* del **Cuadro 6** muestran que las estimaciones de máxima verosimilitud obtenidas a partir de los límites de truncamiento óptimos son buenas en general dado que, la mayoría de estos valores son menores que uno, de hecho, ningún valor para la media y cinco para la desviación estándar son más grandes que uno. De nueva cuenta, cuando $m=2$ los estimados de máxima verosimilitud para la desviación estándar distan con mucho del valor real, $\sigma=10$. Una rápida inspección a los valores de la media y la desviación estándar para todos los

métodos nos indica que todos son buenos dado que, todos están muy cerca de los valores utilizados para generar las muestras, sin embargo, un criterio de desempate puede dárnoslo el **Cuadro 5**, al inspeccionar los valores de las frecuencias obtenidas por el método tradicional y los límites óptimos, se observa que los segundos tienen más valores de las frecuencias concentrados en la media, permitiéndonos formar una campana de gauss más clara – que es la forma de la distribución utilizada para generar las muestras –. De esta manera no obstante que las estimaciones del método tradicional se parecen a las de los límites óptimos, el agrupamiento obtenido a partir de los segundos es mejor, siendo preferible de esta manera utilizar los límites óptimos a los del método tradicional. Además observando los valores *deff* del **Cuadro 7** y su comportamiento en la **Figura 6**, partiendo del criterio definido en el capítulo anterior, observamos que la información relativa asintótica obtenida de los límites derivados de los estimados de máxima verosimilitud final es tanta o tan buena como la de los límites óptimos, dado que para todos los límites de truncamiento el valor de este indicador para los estimados óptimos es un poco mayor a uno. Caso contrario ocurre para los límites y estimadores del método tradicional e inicial, donde los valores de este indicador son en su mayoría más grandes que los de los límites óptimos.

Por tanto el agrupamiento y los estimadores del método propuesto en este trabajo son más eficientes – o mejores – en el sentido de que de ellos la muestra arroja más información sobre la población en estudio.

5.2.3. Muestra de Tamaño $n = 200$

La **Figura 7** muestra la gráfica de la función de verosimilitud para la segunda muestra bajo las condiciones planteadas al principio de este capítulo. La figura muestra que la función de verosimilitud tiene un máximo cercano a los valores de los parámetros considerados al generar la muestra, de tal forma será posible resolver y encontrar los parámetros de máxima verosimilitud, además dichos valores serán únicos.

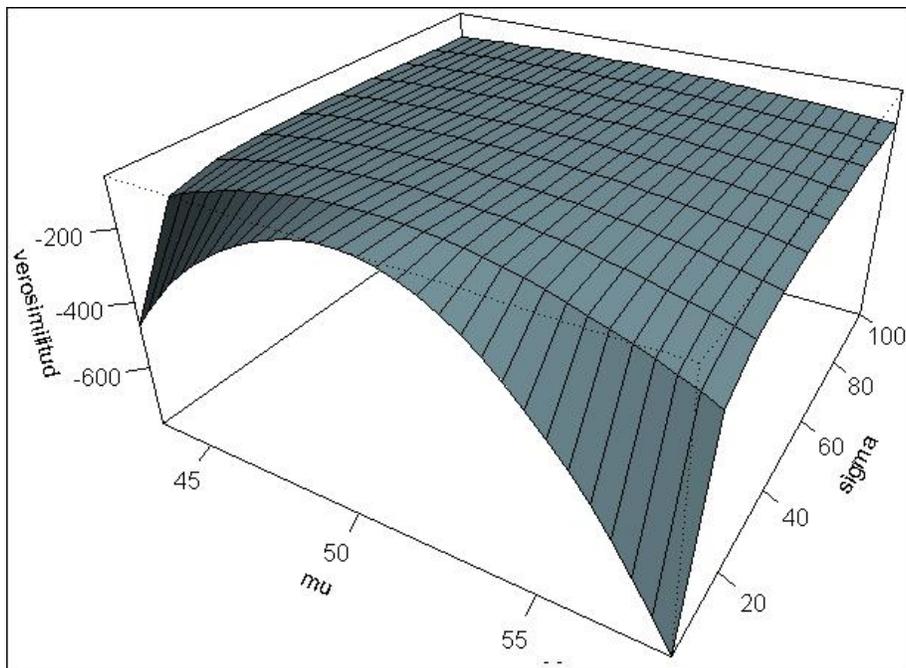
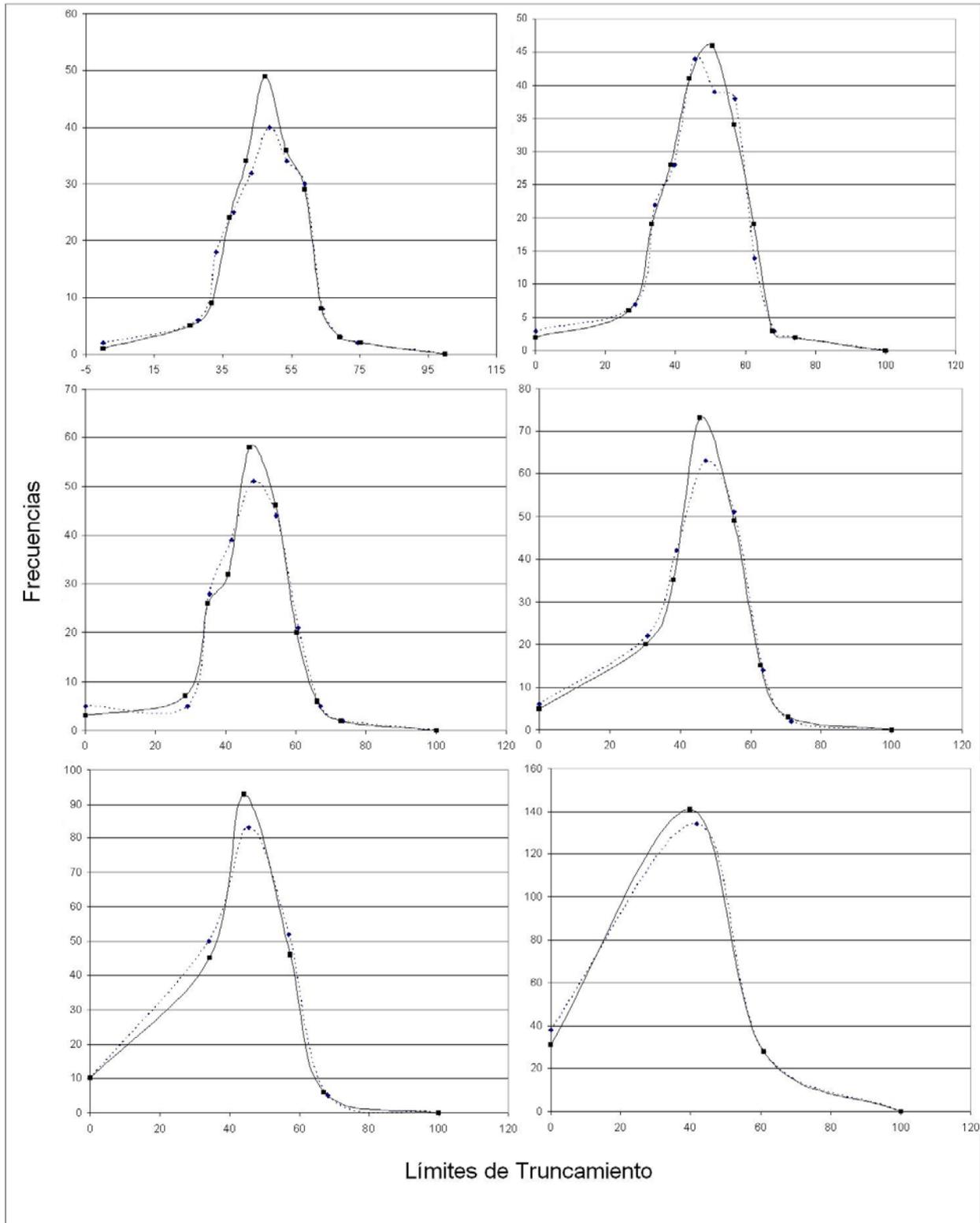


Figura 7. Función de Verosimilitud para función de densidad de probabilidad normal para una muestra de tamaño doscientos y parámetros $\mu = 50$ y $\sigma = 10$

La **Figura 8** muestra los histogramas para el agrupamientos tradicional y el agrupamiento óptimo obtenidos mediante el método tradicional de estadística descriptiva y los límites y frecuencias óptimos obtenidos según la metodología descrita en el capítulo anterior, utilizados para maximizar el sistema (3.17), para $m = 11, 10, 9, 7, 5, 3$ límites.

Figura 8. *Agrupamiento Tradicional* $\cdots\blacklozenge\cdots$ y *Agrupamiento Óptimo* $\text{---}\blacksquare\text{---}$ para una muestra de tamaño doscientos obtenida de una población normal con parámetros $\mu = 50$ y $\sigma = 10$



En el **Cuadro 8** se muestran los límites de truncamiento y frecuencias de grupo iniciales obtenidos mediante el método tradicional de estadística descriptiva y los límites y frecuencias óptimos obtenidos según la metodología descrita en el capítulo anterior, utilizados para maximizar el sistema (3.17).

Cuadro 8. Agrupamientos Inicial y Óptimo para una muestra de tamaño doscientos obtenida de una población normal con parámetros $\mu = 50$ y $\sigma = 10$

m	Inicial		Óptimo	
	\tilde{h}_i	n_i	h_i	n_i
2	51.18046	104	45.7940	60
		96		140
3	41.68046	38	39.7118	31
	60.6805	134	60.7357	141
		28		28
4	36.93046	15	38.0146	24
	51.1805	89	50.5603	73
	65.4305	87	62.9850	86
		9		17
5	34.08046	10	34.3199	10
	45.48046	50	44.0622	45
	56.8805	83	57.3830	93
	68.2805	52	67.1254	46
		5		6
6	32.18046	6	32.4278	6
	41.68046	32	40.8853	32
	51.1805	66	50.4213	58
	60.6805	68	59.9574	73
	70.1805	24	68.4149	26
		4		5
7	30.82332	6	30.2819	5
	38.9662	22	38.0429	20
	47.1090	42	45.6041	35
	55.2519	63	55.2088	73
	63.3948	51	62.7699	49
	71.5376	14	70.5310	15
		2		3
8	29.80546	5	29.2375	5
	36.9305	10	36.3436	9
	44.0555	40	42.7878	32
	51.1805	49	50.3617	50

	58.3055	51	57.9355	58
	65.4305	36	64.3796	33
	72.5555	7	71.4864	11
		2		2
9	29.01379	5	28.3385	3
	35.3471	5	34.9812	7
	41.6805	28	40.7214	26
	48.0138	39	46.8525	32
	54.3471	51	54.1797	58
	60.6805	44	60.3108	46
	67.0138	21	66.0510	20
	73.3471	5	72.6938	6
	2		2	
10	28.38046	3	26.7483	2
	34.0805	7	33.2151	6
	39.7805	22	38.6298	19
	45.4805	28	44.0722	28
	51.1805	44	50.4407	41
	56.8805	39	56.8085	46
	62.5805	38	62.2507	34
	68.2805	14	67.6660	19
	73.9805	3	74.1345	3
	2		2	
11	27.86228	2	25.5024	1
	33.0441	6	31.7738	5
	38.2259	18	36.9035	9
	43.4077	25	41.8474	24
	48.5896	32	47.2881	34
	53.7714	40	53.5205	49
	58.9532	34	58.9612	36
	64.1350	30	63.9051	29
	69.3168	8	69.0348	8
	74.49864	3	75.3062	3
	2		2	

El **Cuadro 9** presenta los valores de la media y la desviación estándar poblacional utilizados como semillas y los valores de los estimadores obtenidos de aplicar el método numérico tal y como se describe en la metodología del capítulo anterior. Las cuatro últimas columnas contienen los valores del indicador *DC* descrito en el capítulo anterior.

Cuadro 9. Semilla y Estimadores Iniciales y Óptimos para una muestra de tamaño doscientos obtenida de una población normal con parámetros $\mu = 50$ y $\sigma = 10$

m	Tradicional		Inicial		Óptimo		$DC(\tilde{\mu})$	$DC(\hat{\mu}^0)$	$DC(\hat{\mu})$	$DC(\hat{\sigma})$	$DC(\hat{\sigma}^0)$	$DC(\tilde{\sigma})$
	$\tilde{\mu}$	$\tilde{\sigma}$	$\hat{\mu}^0$	$\hat{\sigma}^0$	$\hat{\mu}$	$\hat{\sigma}$						
2	50.7130	14.3453	50.9592	4.4115	48.3755	4.9228	0.5083	0.9201	2.6389	18.8816	31.2317	25.7777
3	50.2604	10.9282	50.2238	9.4954	49.8972	10.0327	0.0678	0.0501	0.0106	0.8616	0.2546	0.0011
4	50.4776	9.9709	50.4628	9.0454	50.2467	9.8504	0.2281	0.2142	0.0609	0.0008	0.9112	0.0224
5	50.7298	10.2620	50.7226	9.7013	50.3218	9.7702	0.5326	0.5222	0.1035	0.0687	0.0892	0.0528
6	49.8861	11.4046	50.4214	9.5968	50.2409	9.7075	0.0130	0.1776	0.0580	1.9729	0.1626	0.0856
7	50.4090	10.0972	50.4064	9.8146	50.3514	9.9200	0.1673	0.1652	0.1235	0.0094	0.0344	0.0064
8	50.3632	9.8957	50.3616	9.6754	50.3567	9.8063	0.1319	0.1307	0.1272	0.0109	0.1053	0.0375
9	50.5176	9.8049	50.5161	9.6299	50.3794	9.6102	0.2679	0.2664	0.1439	0.0381	0.1370	0.1520
10	50.4416	9.9979	50.4404	9.8585	50.4882	10.0995	0.1950	0.1939	0.2383	0.0000	0.0200	0.0099
11	50.4053	10.1036	50.4043	9.9895	50.5870	9.7575	0.1643	0.1635	0.3446	0.0107	0.0001	0.0588

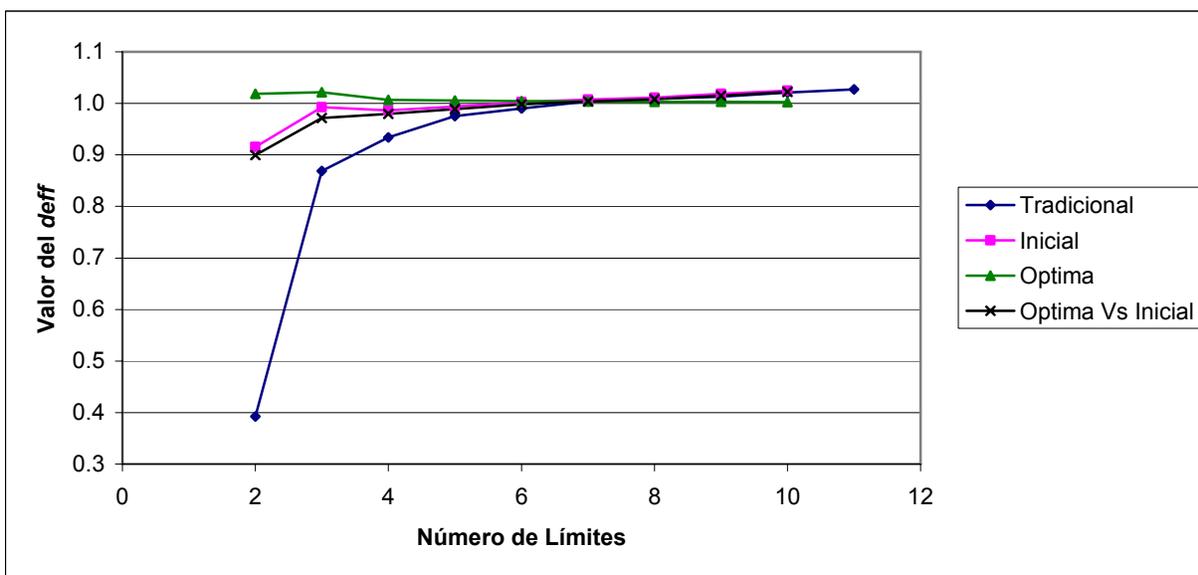
El **Cuadro 10** presenta la información relativa asintótica obtenida por los estimadores de máxima verosimilitud iniciales y óptimos al sustituir directamente estos valores en la ecuación (3.32). Las últimas dos columnas contienen los valores del indicador *deff* definido en el capítulo anterior.

Cuadro 10. Información Relativa Asintótica y Eficiencia del Muestreo para una muestra de tamaño doscientos obtenida de una población normal con parámetros $\mu = 50$ y $\sigma = 10$

m	Óptima	Tradicional ($\tilde{\mu}, \tilde{\sigma}$)	Inicial ($\hat{\mu}^0, \hat{\sigma}^0$)	Final ($\hat{\mu}, \hat{\sigma}$)	<i>deff</i>				
					($\tilde{\mu}, \tilde{\sigma}$)	($\hat{\mu}^0, \hat{\sigma}^0$)	($\hat{\mu}, \hat{\sigma}$)	($\hat{\mu}, \hat{\sigma}$)/($\hat{\mu}^0, \hat{\sigma}^0$)	($\hat{\mu}, \hat{\sigma}$)/($\tilde{\mu}, \tilde{\sigma}$)
2	0.4053	0.4050	0.4045	0.4051	1.0008	1.0018	1.0005	1.0003	1.0014
3	1.9953	1.5059	2.0183	1.9911	1.3250	0.9886	1.0021	1.3222	0.9865
4	2.5954	2.4078	2.5274	2.5275	1.0779	1.0269	1.0268	1.0497	1.0000
5	3.2903	3.0521	3.2268	3.2791	1.0780	1.0197	1.0034	1.0744	1.0162
6	3.6057	3.3670	3.5989	3.6030	1.0709	1.0019	1.0007	1.0701	1.0011
7	3.9293	3.8086	3.8570	3.9281	1.0317	1.0187	1.0003	1.0314	1.0184
8	4.1042	4.0387	4.0434	4.0950	1.0162	1.0150	1.0023	1.0139	1.0127
9	4.2759	4.2071	4.1817	4.2720	1.0164	1.0225	1.0009	1.0154	1.0216
10	4.3806	4.3180	4.2546	4.3803	1.0145	1.0296	1.0001	1.0144	1.0295
11	4.4811	4.4035	4.3065	4.4813	1.0176	1.0405	0.9999	1.0177	1.0406

La **Figura 9** muestra de forma gráfica el comportamiento del indicador *deff* para los límites de truncamiento del método tradicional – semillas –, los de máxima verosimilitud iniciales, los de máxima verosimilitud óptimos y los de máxima verosimilitud óptimos contra los iniciales. Tomando como referencia el valor de la información relativa asintótica de los límites de truncamiento óptimos la cual aparece en el **Cuadro 1**.

Figura 9. Índices de la Eficiencia del Diseño de Muestreo para una muestra de tamaño doscientos generada con distribución normal y parámetros $\mu = 50$ y $\sigma = 10$.



Los valores del indicador *DC* del **Cuadro 9** muestran que las estimaciones de máxima verosimilitud obtenidas a partir de los límites de truncamiento óptimos son muy buenas en general dado que, ninguno de estos valores son mayores que uno, excepto cuando $m=2$ para la desviación estándar. Una rápida inspección a los valores de la media y la desviación estándar

para todos los métodos nos indica que todos son buenos dado que, todos están muy cerca de los valores utilizados para generar las muestras, sin embargo, un criterio de desempate puede dárnoslo el **Cuadro 8**, al inspeccionar los valores de las frecuencias obtenidas por el método tradicional y los límites óptimos, se observa que los segundos tienen más valores de las frecuencias concentrados en la media, permitiéndonos formar una campana de gauss más clara – que es la forma de la distribución utilizada para generar las muestras –. De esta manera no obstante que las estimaciones del método tradicional se parecen a las de los límites óptimos, el agrupamiento obtenido a partir de los segundos es mejor, siendo preferible de esta manera utilizar los límites óptimos a los del método tradicional. Además observando los valores *deff* del **Cuadro 10** y su comportamiento en la **Figura 9**, partiendo del criterio definido en el capítulo anterior, observamos que la información relativa asintótica obtenida de los límites derivados de los estimadores de máxima verosimilitud finales es tanta o tan buena como la de los límites óptimos, dado que para todos los límites de truncamiento el valor de este indicador para los estimados óptimos es un poco mayor a uno. Caso contrario ocurre para los límites y estimadores del método tradicional e inicial, donde los valores de este indicador son en su mayoría más grandes que los de los límites óptimos.

Por tanto el agrupamiento y los estimadores del método propuesto en este trabajo son más eficientes – o mejores – en el sentido de que de ellos la muestra arroja más información sobre la población en estudio.

5.2.4. Muestra de Tamaño $n = 400$

La **Figura 10** muestra la gráfica de la función de verosimilitud para la segunda muestra bajo las condiciones planteadas al principio de este capítulo. La figura muestra que la función de verosimilitud tiene un máximo cercano a los valores de los parámetros considerados al generar la muestra, de tal forma será posible resolver y encontrar los parámetros de máxima verosimilitud, además dichos valores serán únicos.

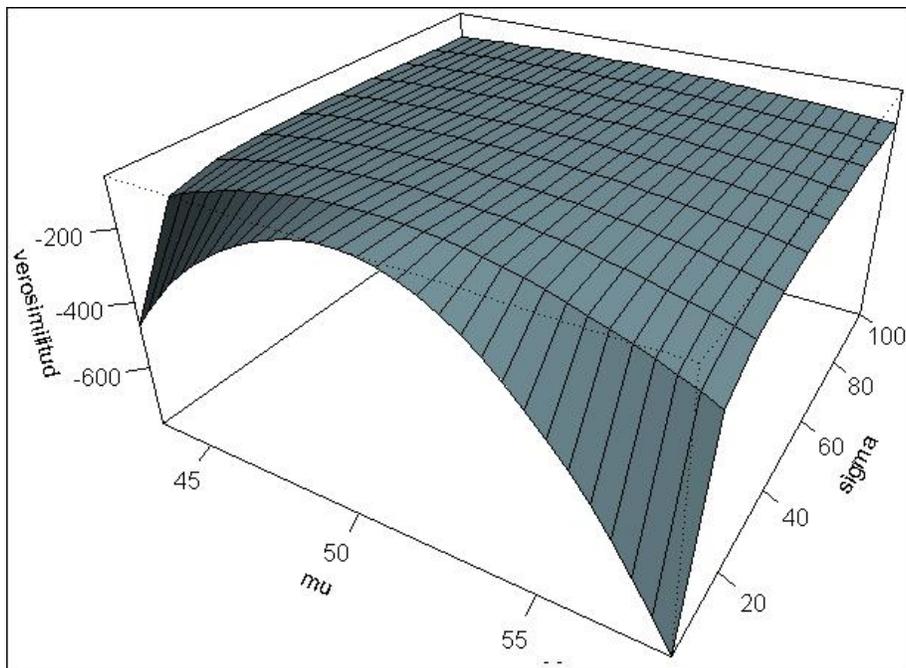
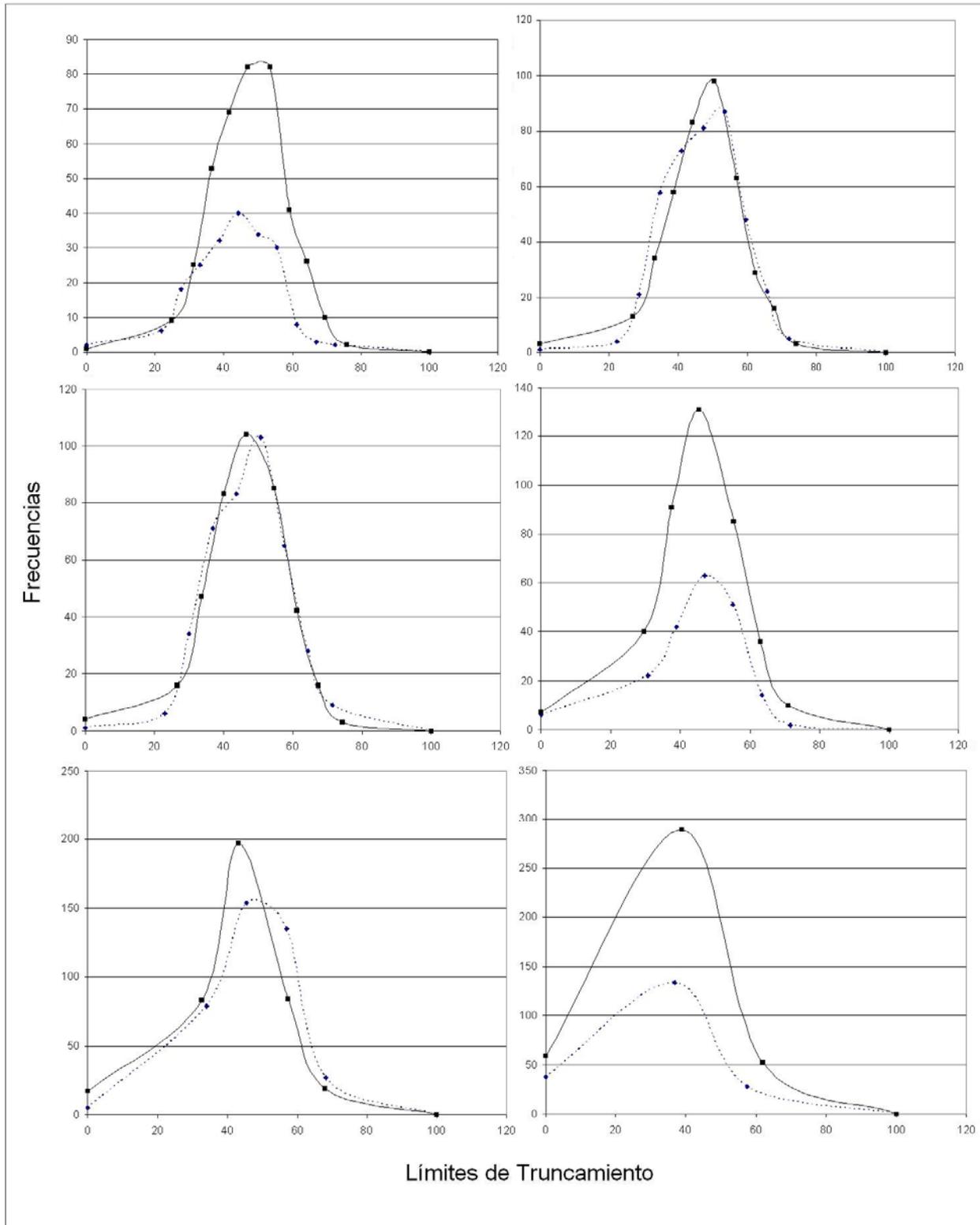


Figura 10. Función de Verosimilitud para función de densidad de probabilidad normal para una muestra de tamaño cuatrocientos y parámetros $\mu = 50$ y $\sigma = 10$

La **Figura 11** muestra los histogramas para el agrupamiento tradicional y el agrupamiento óptimo obtenidos mediante el método tradicional de estadística descriptiva y los límites y frecuencias óptimos obtenidos según la metodología descrita en el capítulo anterior, utilizados para maximizar el sistema (3.17), para $m = 11, 10, 9, 7, 5, 3$ límites.

Figura 11. *Agrupamiento Tradicional* $\cdots\blacklozenge\cdots$ y *Agrupamiento Óptimo* $\text{---}\blacksquare\text{---}$ para una muestra de tamaño cuatrocientos obtenida de una población normal con parámetros $\mu = 50$ y $\sigma = 10$



En el **Cuadro 11** se muestran los límites de truncamiento y frecuencias de grupo iniciales obtenidos mediante el método tradicional de estadística descriptiva y los límites y frecuencias óptimos obtenidos según la metodología descrita en el capítulo anterior, utilizados para maximizar el sistema (3.17).

Cuadro 11. Agrupamientos Inicial y Óptimo para una muestra de tamaño cuatrocientos obtenida de una población normal con parámetros $\mu = 50$ y $\sigma = 10$

m	Inicial		Óptimo	
	\tilde{h}_i	n_i	h_i	n_i
2	47.17637	157	42.9865	99
		243		301
3	36.84304	38	38.8537	59
	57.5097	134	61.8449	289
		28		52
4	31.67637	15	38.0146	33
	47.1764	89	50.5603	159
	62.6764	87	62.9850	170
		9		38
5	34.08046	5	32.6425	17
	45.48046	79	43.1156	83
	56.8805	154	57.4356	197
	68.2805	135	67.9087	84
		27		19
6	26.5097	4	30.6593	10
	36.84304	37	39.8677	57
	47.1764	116	50.2504	124
	57.5097	141	60.6333	148
	67.8430	83	69.8417	49
		19		12
7	30.82332	6	29.6286	7
	38.9662	22	37.6170	40
	47.1090	42	45.3996	91
	55.2519	63	55.2857	131
	63.3948	51	63.0683	85
	71.5376	14	71.0567	36
		2		10
8	23.92637	1	27.6945	4
	31.6764	11	35.2977	24
	39.4264	52	42.1926	67
	47.1764	93	50.2963	96

	54.9264	107	58.3999	124
	62.6764	88	65.2946	56
	70.4264	37	72.8986	25
		11		4
9	23.06526	1	26.5482	4
	29.9542	6	33.7173	16
	36.8430	34	39.9124	47
	43.7319	71	46.5293	83
	50.6208	83	54.4372	104
	57.5097	103	61.0541	85
	64.3986	65	67.2492	42
	71.2875	28	74.4183	16
	9		3	
10	22.37637	1	26.7483	3
	28.5764	4	33.2151	13
	34.7764	21	38.6298	34
	40.9764	58	44.0722	58
	47.1764	73	50.4407	83
	53.3764	81	56.8085	98
	59.5764	87	62.2507	63
	65.7764	48	67.6660	29
	71.9764	22	74.1345	16
	5		3	
11	21.81273	2	24.7051	1
	27.4491	6	31.1558	9
	33.0855	18	36.4321	25
	38.7218	25	41.5172	53
	44.3582	32	47.1134	69
	49.9946	40	53.5240	82
	55.6309	34	59.1202	82
	61.2673	30	64.2053	41
	66.9036	8	69.4816	26
	72.54001	3	75.9323	10
	2		2	

El **Cuadro 12** presenta los valores de la media y la desviación estándar poblacional utilizados como semillas y los valores de los estimadores obtenidos de aplicar el método numérico tal y como se describe en la metodología del capítulo anterior. Las cuatro últimas columnas contienen los valores del indicador *DC* descrito en el capítulo anterior.

Cuadro 12. Semilla y Estimadores Iniciales y Óptimos para una muestra de tamaño cuatrocientos obtenida de una población normal con parámetros $\mu = 50$ y $\sigma = 10$

m	Semillas		Intervalos		Óptimo		$DC(\tilde{\mu})$	$DC(\hat{\mu}^0)$	$DC(\hat{\mu})$	$DC(\hat{\sigma})$	$DC(\hat{\sigma}^0)$	$DC(\hat{\sigma})$
	$\tilde{\mu}$	$\tilde{\sigma}$	$\hat{\mu}^0$	$\hat{\sigma}^0$	$\hat{\mu}$	$\hat{\sigma}$						
2	50.5210	15.1472	48.4492	4.6655	46.3887	4.9858	0.2714	2.4051	13.0415	26.4940	28.4571	25.1423
3	50.2604	10.9282	50.2238	9.4954	49.8972	10.0327	0.0678	0.0501	0.0106	0.8616	0.2546	0.0011
4	50.2400	11.0843	50.2404	10.1144	50.7834	10.3407	0.0576	0.0578	0.6137	1.1758	0.0131	0.1161
5	50.2777	11.0414	50.2756	10.4290	50.4321	10.5229	0.0771	0.0759	0.1867	1.0846	0.1840	0.2734
6	50.2515	10.8716	50.2506	10.4489	50.3839	10.3196	0.0632	0.0628	0.1474	0.7597	0.2015	0.1021
7	50.3434	10.4238	50.3427	10.1021	50.2358	10.5795	0.1179	0.1174	0.0556	0.1796	0.0104	0.3358
8	50.2963	10.5935	50.2961	10.3522	50.3048	10.2122	0.0878	0.0877	0.0929	0.3523	0.1240	0.0450
9	50.4835	10.5830	50.4832	10.3930	50.2761	10.3789	0.2338	0.2335	0.0762	0.3399	0.1544	0.1436
10	50.2146	10.4672	50.2144	10.3120	50.4275	10.5700	0.0461	0.0460	0.1827	0.2183	0.0973	0.3249
11	50.3188	10.4037	50.3187	10.2750	50.3432	10.3272	0.1017	0.1016	0.1178	0.1629	0.0756	0.1071

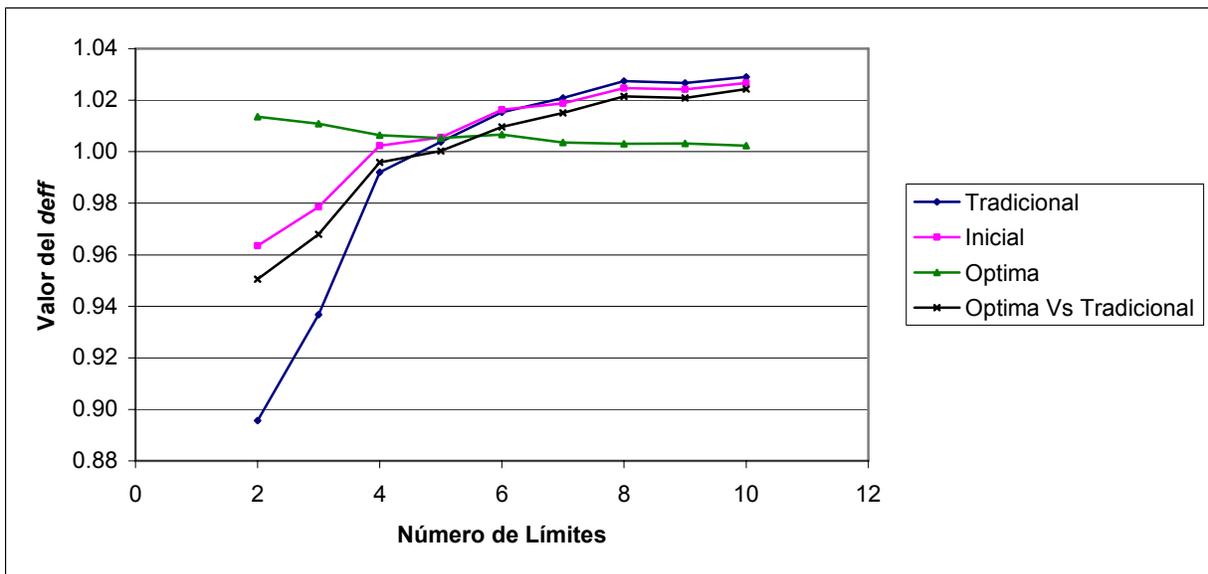
El **Cuadro 13** presenta la información relativa asintótica obtenida por los estimadores de máxima verosimilitud iniciales y óptimos al sustituir directamente estos valores en la ecuación (3.32). Las últimas dos columnas contienen los valores del indicador *deff* definido en el capítulo anterior.

Cuadro 13. Información Relativa Asintótica y Eficiencia del Muestreo para una muestra de tamaño cuatrocientos obtenida de una población normal con parámetros $\mu = 50$ y $\sigma = 10$

m	Óptima	Tradicional ($\tilde{\mu}, \tilde{\sigma}$)	Inicial ($\hat{\mu}^0, \hat{\sigma}^0$)	Final ($\hat{\mu}, \hat{\sigma}$)	<i>deff</i>				
					($\tilde{\mu}, \tilde{\sigma}$)	($\hat{\mu}^0, \hat{\sigma}^0$)	($\hat{\mu}, \hat{\sigma}$)	($\hat{\mu}, \hat{\sigma}$)/($\hat{\mu}^0, \hat{\sigma}^0$)	($\hat{\mu}, \hat{\sigma}$)/($\tilde{\mu}, \tilde{\sigma}$)
2	0.1952	0.4923	0.0221	0.1154	0.3966	8.8438	1.6917	5.2278	0.2344
3	0.8221	0.9179	0.8532	0.8110	0.8956	0.9635	1.0137	0.9505	0.8835
4	1.1152	1.1905	1.1396	1.1031	0.9367	0.9786	1.0109	0.9680	0.9266
5	1.3737	1.3848	1.3705	1.3649	0.9920	1.0024	1.0065	0.9959	0.9856
6	1.5191	1.5133	1.5106	1.5111	1.0039	1.0056	1.0053	1.0003	0.9986
7	1.6275	1.6028	1.6013	1.6167	1.0154	1.0163	1.0066	1.0096	1.0087
8	1.7008	1.6661	1.6695	1.6947	1.0209	1.0188	1.0036	1.0151	1.0172
9	1.7559	1.7090	1.7136	1.7505	1.0274	1.0247	1.0031	1.0215	1.0242
10	1.7967	1.7501	1.7542	1.7908	1.0266	1.0242	1.0033	1.0209	1.0233
11	1.8283	1.7768	1.7808	1.8240	1.0290	1.0267	1.0024	1.0243	1.0266

La **Figura 12** muestra de forma gráfica el comportamiento del indicador *deff* para los límites de truncamiento del método tradicional – semillas –, los de máxima verosimilitud iniciales, los de máxima verosimilitud óptimos y los de máxima verosimilitud óptimos contra los iniciales. Tomando como referencia el valor de la información relativa asintótica de los límites de truncamiento óptimos la cual aparece en el **Cuadro 1**.

Figura 12. Índices de la Eficiencia del Diseño de Muestreo para una muestra de tamaño cuatrocientos generada con distribución normal y parámetros $\mu = 50$ y $\sigma = 10$.



Los valores del indicador *DC* del **Cuadro 12** muestran que las estimaciones de máxima verosimilitud obtenidas a partir de los límites de truncamiento óptimos son buenas en general, excepto cuando $m=2$ para la desviación estándar, dado que la mayoría de estos valores son menores que uno, de hecho, ningún valor para la media y ninguno para la desviación estándar son más grandes que uno. Una rápida inspección a los valores de la media y la desviación

estándar para todos los métodos nos indica que todos son buenos dado que, todos están muy cerca de los valores utilizados para generar las muestras, sin embargo, un criterio de desempate puede dárnoslo el **Cuadro 11**, al inspeccionar los valores de las frecuencias obtenidas por el método tradicional y los límites óptimos, se observa que los segundos tienen más valores de las frecuencias concentrados en la media, permitiéndonos formar una campana de Gauss más clara – que es la forma de la distribución utilizada para generar las muestras –. De esta manera no obstante que las estimaciones del método tradicional se parecen a las de los límites óptimos, el agrupamiento obtenido a partir de los segundos es mejor, siendo preferible de esta manera utilizar los límites óptimos a los del método tradicional. Además observando los valores *deff* del **Cuadro 13** y su comportamiento en la **Figura 8**, partiendo del criterio definido en el capítulo anterior, observamos que la información relativa asintótica obtenida de los límites derivados de los estimadores de máxima verosimilitud óptimos es tanta o tan buena como la de los límites óptimos, dado que para todos los límites de truncamiento el valor de este indicador para los estimados finales es un poco mayor a uno. Caso análogo ocurre para los límites y estimadores del método tradicional e inicial, donde los valores de este indicador son en su mayoría cercanos a uno.

Para este tamaño de muestra, todos los métodos obtienen muy buena información de la muestra, por tanto, para muestras más grandes se puede elegir tanto los límites tradicionales como los finales y los óptimos.

6. CONCLUSIONES

Una vez obtenidos y analizados los resultados puede concluirse, respecto a los objetivos planteados, lo siguiente:

- Por lo que hace a la información relativa asintótica, ésta incrementa conforme el número de intervalos también se incrementa; sin embargo, cabe agregar que tal información tiene incrementos que, conforme se incrementan las clases o límites de truncamiento, no son considerables; de hecho, el incremento en términos porcentuales para $m > 7$ es menor del 5% con respecto a su inmediato sucesor, del 8.75% de $m=6$ a $m=5$ y de $m=7$ a $m=6$; 8.24%. Lo anterior sugiere considerar de 5 a 7 intervalos como mínimo.
- Los estimadores para la media y la desviación estándar tanto por el método tradicional como por el método de máxima verosimilitud por lo general son buenos, a medida que el número de intervalos se incrementa, *i.e.*, la diferencia entre el valor del parámetro y el valor estimado es pequeño. Lo anterior se consolida cuando el tamaño de muestra se incrementa. No obstante lo anterior, cuando el número de intervalos decrece y el tamaño de muestra también, los estimados de máxima verosimilitud para la desviación estándar no son deseables, máxime si $m=2$.

- Al considerar la eficiencia del diseño de muestreo, $deff$, vía la información relativa asintótica; el agrupamiento y los estimadores del método propuesto en este trabajo, son mas eficientes – o mejores – en el sentido de que a partir de estos las muestras así tratadas arrojan mas información sobre la población en estudio. De hecho cuando $m \geq 5$ al comparar el agrupamiento final óptimo contra los agrupamientos tradicional e inicial la $deff$ siempre es mayor que uno. Lo anterior trae como consecuencia no solo que la distribución de frecuencias muestrales, al utilizar los límites de truncamiento óptimos, esta obedezca con mucho a la distribución normal; sino que además, los estimados de la media y de la desviación estándar oscilen en una vecindad pequeña de los valores reales de tales parámetros.
- Se recomienda utilizar los límites de truncamiento obtenidos en el presente estudio a fin de obtener un buen histograma.

7. LITERATURA CITADA

1. COCHRAN, W. G. (1977). *Sampling techniques*. New York, Wiley.
2. HUGHES, HARRY M. (1949). *Estimation of the variance of the bivariate normal distribution*, University of California Publications in Statistics, Vol 1, No. 4, pp. 37 – 51.
3. KISH L. (1975). *Muestreo de Encuestas*. Editorial Trillas. México.
4. KULLDORFF, G. (1962). *Contributions to the theory of estimation from grouped and partially grouped samples*. New York, Wiley.
5. OGAWA J. (1977). *Optimal Spacings for the Simultaneous Estimation of the Location and Scale Parameters of a Normal Distribution Based on Selected Two Sample Quantiles*, Journal of Statistical Planning and Inference, Vol 1. No.1 (1977), pp. 61 – 72.
6. RAJ, D. (1971). *The design of sample surveys*. McGraw-Hill series in probability and statistics. New York, McGraw-Hill.
7. SUKHATME, P. V., & SUKHATME, B. V. (1970). *Sampling theory of surveys with applications*. Ames, Iowa State University Press.

8. WALD, ABRAHAM (1948). *Asymptotic Properties of the Maximum Likelihood Estimate of an Unknown Parameter of a Discrete Stochastic Process*, The Annals of Mathematical Statistics, Vol. 19, No. 1. (Mar., 1948), pp. 40-46.

8. ANEXOS

A continuación se muestran los programas escritos con la finalidad de cumplir con los objetivos del estudio.

I. Varianzas Asintóticas

```
k <- 2
```

```
final <- function(z)
```

```
{
```

```
  u <- c(-4,z,4)
```

```
  sumu <- 0
```

```
  susi <- 0
```

```
  suac <- 0
```

```
  for(i in 1:k)
```

```
  {
```

```
    am1 <- pnorm(u[i+1]) - pnorm(u[i])
```

```
    am2 <- (u[i+1]*dnorm(u[i+1])) - (u[i]*dnorm(u[i]))
```

```
    am3 <- (dnorm(u[i+1])-dnorm(u[i]))^2
```

```

am4 <- am1*am2

nmu <- am4 + am3

dmu <- pnorm(u[i+1]) - pnorm(u[i])

vam <- nmu/dmu

tomu <- sumu + vam

sumu <- tomu

as1 <- pnorm(u[i+1]) - pnorm(u[i])

as2 <- ((u[i+1]^2)-1)*(u[i+1]*dnorm(u[i+1]))

as3 <- ((u[i]^2)-1)*(u[i]*dnorm(u[i]))

as4 <- as2 - as3

as5 <- as1 * as4

as6 <- (am2)^2

nsi <- as5 + as6

dsi <- pnorm(u[i+1]) - pnorm(u[i])

vas <- nsi/dsi

tosi <- susi + vas

susi <- tosi

ac1 <- dnorm(u[i+1])-dnorm(u[i])

ac2 <- pnorm(u[i+1])-pnorm(u[i])

ac3 <- ((u[i+1]^2)*dnorm(u[i+1])) - ((u[i]^2)*dnorm(u[i]))

ac4 <- (u[i+1]*dnorm(u[i+1])) - (u[i]*dnorm(u[i]))

ac5 <- (ac1*ac4)

ac6 <- ac5/ac2

```

```
ac7 <- ac1-ac3
ac8 <- ac7-ac5
toac <- suac + ac8
suac <- toac
}
deter <- (tomu * tosi)-(toac^2)
return(-deter)
}
optim(c(-0.001),final, NULL, method = "BFGS", hessian = TRUE)
```

II. Estimadores de Máxima Verosimilitud

```
n <- 100
numpar <- 11
mu <- 50
sigma <- 10
y <- sort(rnorm(n,mu,sigma))
write.table(y, file="c:/muestra1.csv", sep="," , col.names=NA)
long <- round(abs(y[n]-y[1]),0)
z <- array(NA, c(numpar-1,1))
for(i in 1:numpar-1)
  { z[i] <- y[1]+(i)*(long/numpar)}
x <- c(min(y),z,max(y))
n <- array(0, c(length(x)-1,1))
m <- array(data=NA,dim=length(x)-1,1)
abg <- array(data=NA,dim=length(x)-1,1)
abh <- array(data=NA,dim=length(x)-1,1)
abf <- array(data=NA,dim=length(x)-1,1)
suma <- 0
for(i in 2:length(x))
  {
n[i-1] <- sum(y <= x[i]) - suma
```

```

suma <- n[i-1] + suma
}
y
x
n
verosi <- function(mu,sigma)
{
  sum1 <- 0
  for(i in 1:(length(x)-1))
  {
    a <- pnorm(x[i+1],mean=mu,sd=sigma) - pnorm(x[i], mean=mu,sd=sigma)
    b <- sum1 + n[i]*log(a)
    sum1 <- b
  }
  return(sum1)
}

for(i in 1:(length(x)-1))
{

  m[i] <- (x[i+1] + x[i])/2
  abg[i] <- n[i] * m[i]
}

```

```

    abh[i] <- (m[i])^2
    abf[i] <- n[i] * abh[i]
}

medtr <- sum(abg)/sum(n)
vartr <- (sum(abf)/sum(n))-(medtr^2)
destr <- sqrt(vartr)

medtr
vartr
destr

fxy<-function(w)
{
    x<-w[1]
    y<-w[2]
    return(-1*verosi(x,y))
}

nlminb(c(medtr,destr),fxy,gradient=NULL)

etra <- array(data=NA,dim=length(x),dimnames=NULL)

for(i in 1:length(x))
{
    etra[i] <- (x[i]- medtr)/destr
}

```

```

sumut <- 0
susit <- 0
suact <- 0
for(i in 1:(length(x)-1))
{
  atra1 <- pnorm(etra[i+1]) - pnorm(etra[i])
  atra2 <- (etra[i+1]*dnorm(etra[i+1])) - (etra[i]*dnorm(etra[i]))
  atra3 <- (dnorm(etra[i+1])-dnorm(etra[i]))^2
  atra4 <- atra1*atra2
  nmut <- atra4 + atra3
  dmut <- pnorm(etra[i+1]) - pnorm(etra[i])
  vamt <- nmut/dmut
  tomut <- sumut + vamt
  sumut <- tomut
  sit1 <- pnorm(etra[i+1]) - pnorm(etra[i])
  sit2 <- (((etra[i+1])^2)-1)*((etra[i+1]))*dnorm(etra[i+1])
  sit3 <- (((etra[i])^2)-1)*((etra[i]))*dnorm(etra[i])
  sit4 <- sit2 - sit3
  sit5 <- sit1 * sit4
  sit6 <- (atra2)^2
  nsit <- sit5 + sit6
  dsit <- pnorm(etra[i+1]) - pnorm(etra[i])
  vast <- nsit/dsit

```

```

tosit <- susit + vast

susit <- tosit

cot1 <- dnorm(etra[i+1])-dnorm(etra[i])
cot2 <- pnorm(etra[i+1])-pnorm(etra[i])
cot3 <- ((etra[i+1]^2)*dnorm(etra[i+1])) - ((etra[i]^2)*dnorm(etra[i]))
cot4 <- (etra[i+1]*dnorm(etra[i+1])) - (etra[i]*dnorm(etra[i]))
cot5 <- (cot1*cot4)
cot6 <- cot5/cot2
cot7 <- cot1-cot3
cot8 <- cot7-cot5

toact <- suact + cot8

suact <- toact

}

dett <- (tomut * tosit)-(toact^2)

dett

muin <- 50.66308

stin <- 10.83727

ein <- array(data=NA,dim=length(x),dimnames=NULL)

for(i in 1:length(x))

{

ein[i] <- (x[i]- muin)/stin

}

sumuin <- 0

```

```

susiin <- 0
suacin <- 0
for(i in 1:(length(x)-1))
{
  min1 <- pnorm(ein[i+1]) - pnorm(ein[i])
  min2 <- (ein[i+1]*dnorm(ein[i+1])) - (ein[i]*dnorm(ein[i]))
  min3 <- (dnorm(ein[i+1])-dnorm(ein[i]))^2
  min4 <- min1*min2
  nmuin <- min4 + min3
  dmuin <- pnorm(ein[i+1]) - pnorm(ein[i])
  vamin <- nmuin/dmuin
  tomuin <- sumuin + vamin
  sumuin <- tomuin
  sin1 <- pnorm(ein[i+1]) - pnorm(ein[i])
  sin2 <- (((ein[i+1])^2)-1)*((ein[i+1]))*dnorm(ein[i+1])
  sin3 <- (((ein[i])^2)-1)*((ein[i]))*dnorm(ein[i])
  sin4 <- sin2 - sin3
  sin5 <- sin1 * sin4
  sin6 <- (min2)^2
  nsin <- sin5 + sin6
  dsin <- pnorm(ein[i+1]) - pnorm(ein[i])
  vasin <- nsin/dsin
  tosin <- susiin + vasin
}

```

```

susiin <- tosin

cin1 <- dnorm(ein[i+1])-dnorm(ein[i])

cin2 <- pnorm(ein[i+1])-pnorm(ein[i])

cin3 <- ((ein[i+1]^2)*dnorm(ein[i+1])) - ((ein[i]^2)*dnorm(ein[i]))

cin4 <- (ein[i+1]*dnorm(ein[i+1])) - (ein[i]*dnorm(ein[i]))

cin5 <- (cin1*cin4)

cin6 <- cin5/cin2

cin7 <- cin1-cin3

cin8 <- cin7-cin5

toacin <- suacin + cin8

suacin <- toacin

}

dein <- (tomuin * tosin)-(toacin^2)

dein

zw <- c(23.6477,30.4514,36.0165,41.3799,47.2824,54.0438,59.9462,65.3097,70.8747,77.6784)

xw <- c(min(y)-100000,zw,max(y)+100000)

nw <- array(0, c(length(xw)-1,1))

m <- array(data=NA,dim=length(xw)-1,1)

abgw <- array(data=NA,dim=length(xw)-1,1)

abhw <- array(data=NA,dim=length(xw)-1,1)

abfw <- array(data=NA,dim=length(xw)-1,1)

sumaw <- 0

for(i in 2:length(xw))

```

```

{
  nw[i-1] <- sum(y <= xw[i]) - sumaw
  sumaw <- nw[i-1] + sumaw
}
y
xw
nw
verosi <- function(mu,sigma)
{
  sum1w <- 0
  for(i in 1:(length(xw)-1))
  {
    aw <- pnorm(xw[i+1],mean=mu,sd=sigma) - pnorm(xw[i], mean=mu,sd=sigma)
    bw <- sum1w + nw[i]*log(aw)
    sum1w <- bw
  }
  return(sum1w)
}
mlemu <- muin
mlesig <- stin
fxy <-function(w)
{
  x<-w[1]

```

```

        y<-w[2]
        return(-1*verosi(x,y))
    }

nlminb(c(mlemu,mlesig),fxy,gradient=NULL)

muverw <- 50.73071

siverw <- 11.27571

ew <- array(data=NA,dim=length(xw),dimnames=NULL)

for(i in 1:length(xw))
{
ew[i] <- (xw[i]- muverw)/siverw
}

wsumu <- 0

wsusi <- 0

wsuac <- 0

for(i in 1:(length(xw)-1))
{
    amw1 <- pnorm(ew[i+1]) - pnorm(ew[i])
    amw2 <- (ew[i+1]*dnorm(ew[i+1])) - (ew[i]*dnorm(ew[i]))
    amw3 <- (dnorm(ew[i+1])-dnorm(ew[i]))^2
    amw4 <- amw1*amw2
    nwmu <- amw4 + amw3
    dwmu <- pnorm(ew[i+1]) - pnorm(ew[i])
    wvam <- nwmu/dwmu
}

```

```

wtomu <- wsumu + wvam
wsumu <- wtomu
asw1 <- pnorm(ew[i+1]) - pnorm(ew[i])
asw2 <- (((ew[i+1])^2)-1)*((ew[i+1]))*dnorm(ew[i+1])
asw3 <- (((ew[i])^2)-1)*((ew[i]))*dnorm(ew[i])
asw4 <- asw2 - asw3
asw5 <- asw1 * asw4
asw6 <- (amw2)^2
nwsu <- asw5 + asw6
dwsu <- pnorm(ew[i+1]) - pnorm(ew[i])
wvas <- nwsu/dwsu
wtosu <- wsusi + wvas
wsusi <- wtosu
wac1 <- dnorm(ew[i+1])-dnorm(ew[i])
wac2 <- pnorm(ew[i+1])-pnorm(ew[i])
wac3 <- ((ew[i+1]^2)*dnorm(ew[i+1])) - ((ew[i]^2)*dnorm(ew[i]))
wac4 <- (ew[i+1]*dnorm(ew[i+1])) - (ew[i]*dnorm(ew[i]))
wac5 <- (wac1*wac4)
wac6 <- wac5/wac2
wac7 <- wac1-wac3
wac8 <- wac7-wac5
wtoac <- wsuac + wac8
wsuac <- wtoac

```

```
}
```

```
Wdeter <- (wtomu * wtoasi)-(wtoac^2)
```

```
Wdeter
```

FE DE ERRATAS

TESIS DE MAESTRÍA

“Límites Óptimos para una Muestra Agrupada Normalmente Distribuida con Media y Varianza Desconocidas”

Página	Renglón	Dice	Debe decir
36	9	“...Así; loas límites óptimos...”	“...Así; los límites óptimos...”
40	4	“...se convierte en una diferencia de cuadrados...”	“...se convierte en una diferencia al cuadrado...”
50	15	“...de hecho, sólo un valor para la media y tres para la desviación estándar son ...”	“...de hecho, ningún valor para la media y ninguno para la desviación estándar son...”
57	12	“...de hecho, ningún valor para la media y cinco para la desviación estándar son ...”	“...de hecho, un valor para la media y seis para la desviación estándar son ...”