

TAMAÑO DE MUESTRA REQUERIDO PARA ESTIMAR LA MEDIA ARITMÉTICA DE UNA DISTRIBUCIÓN LOGNORMAL

JAVIER A. CASTAÑEDA*
ADRIANA PEREZ **
JACKY F. A. GIL ***

Resumen

Se presentan fórmulas cerradas para calcular el tamaño de la muestra requerido en la estimación de la media aritmética de una distribución lognormal para datos censurados y no censurados. Las fórmulas son el resultado del ajuste de modelos no lineales para los tamaños de la muestra exactos reportados por Pérez (1995) en función de la desviación geométrica estándar, el porcentaje de diferencia a la verdadera media aritmética y niveles de confianza del 90 %, 95 % y 99 %. Las fórmulas presentadas evitan en gran medida el problema de subestimación que se presenta en otras fórmulas reportadas en la literatura estadística.

Palabras Clave: Niveles de concentración, Censura, Desviación Geométrica Estándar, Medida de Asimetría de Hougaard.

*Instructor. Unidad de Epidemiología Clínica y Bioestadística, Facultad de Medicina. Pontificia Universidad Javeriana; Carrera 7 N° 40-62, HUSI, Piso 2, Bogotá, Colombia. Correspondencia: e-mail:javier.castaneda@javeriana.edu.co.

**Profesora Asociada, Unidad de Epidemiología Clínica y Bioestadística, Facultad de Medicina. Pontificia Universidad Javeriana. The University of Texas Health Science Center at Houston. School of Public Health at Brownsville. TX 78526. USA. e-mail: aperez@sph.uth.tmc.edu.

***Instructor. Unidad de Epidemiología Clínica y Bioestadística, Facultad de Medicina. Pontificia Universidad Javeriana. e-mail: fgil@javeriana.edu.co.

Abstract

We present close formulae to calculate the required sample size to estimate the arithmetic mean of a lognormal distribution for censored and non-censored data. These formulae were obtained by adjusting non linear models for the exact sample sizes estimates reported by Perez (1995). The formulae presented are functions of the estimated geometric standard deviation, the proportional precision from the true arithmetic mean and using confidence levels of 90%, 95% and 99%. These new close formulae correct the underestimation problem in other formulae presented in the statistical literature.

Keywords: *Concentration Levels, Censoring, Geometric Standard Deviation, Hougaard Asymmetry Measurement.*

1. Introducción

Existen muchas variables con distribución lognormal. Esta distribución se encuentra muy frecuentemente en investigaciones sobre el medio ambiente. Usualmente los niveles de concentración de contaminantes en el agua, aire o tierra siguen una distribución lognormal, Schorp et al (2002), Lu (2002), Filmon et al (2000), Potts et al (2000) y Smith et al (2000). En general los niveles de concentración de cualquier elemento casi siempre siguen una distribución lognormal como lo demuestran las investigaciones de Panayi et al (2002) y Kendall (2001).

De otra parte, cuando se busca una medida resumen para una variable lognormal, dependiendo del objetivo de la investigación, se tienen a disposición medidas como la media aritmética, la media geométrica, un límite de tolerancia, etc. Leidel y Busch (1985) discuten las diferentes opciones en varios contextos. Las estrategias de muestreo enfocadas en la media aritmética a menudo son eficaces cuando se pretende evaluar la exposición a materiales tóxicos, debido a que la media aritmética de la distribución lognormal se relaciona con la frecuencia de las exposiciones que exceden concentraciones particulares en el aire, Evans (1988), Rappaport (1988).

Durante la determinación de los niveles de concentración de los contaminantes se encuentra el problema de los datos censurados, por ejemplo cuando al medir concentraciones diminutas de contaminantes en el medio ambiente, ni siquiera los instrumentos de medición de alta tecnología pueden detectar la concentración real. Usualmente, las mediciones pueden ser censuradas por varias razones, como por ejemplo si la señal producida por el contaminante es demasiado pequeña y la medición no puede diferenciarla del ruido de fondo. La censura a la izquierda puede ocurrir cuando la señal registrada es identificada con el compuesto de interés pero está por debajo de un umbral, Lambert (1991).

Investigaciones cuyos componentes involucran hipótesis de variables que siguen la distribución lognormal, la media aritmética como medida resumen, la desviación

geométrica estándar (GSD) y la presencia o no de censura, requieren del cálculo del tamaño de la muestra mínimo para proporcionar resultados estadísticamente válidos. Harris (1991) consideró una fórmula para estimar el tamaño de muestra donde se determina la media aritmética dentro de una fracción especificada de la GSD, a un nivel de confianza específico. La fórmula para calcular el tamaño de la muestra de una variable lognormal sin censura, reportada por Hewett (1995) requiere de una muestra piloto para proporcionar estimaciones suficientemente precisas de los parámetros de la distribución y se basa en una precisión proporcional a la localización de la verdadera media aritmética. Pérez y Lefante (1997) reportan tablas con los tamaños de muestra exactos incluyendo situaciones sin censura o con censura a la izquierda y se basan también en la precisión proporcional a la localización de la verdadera media aritmética. Estos autores probaron sus resultados mediante el uso de simulaciones de Monte Carlo. Sin embargo, el uso de las tablas para calcular el tamaño de la muestra hace que el proceso sea incómodo para el investigador. Armstrong (1995) basa sus estimaciones en la precisión absoluta a la media aritmética para datos sin censura.

1.1. Tamaño de la muestra para estimar la media aritmética de la distribución lognormal sin censura

Pérez (1995) describió cuatro técnicas existentes para calcular el tamaño de la muestra requerido en la estimación de la media aritmética de una distribución lognormal con base en la amplitud del intervalo de confianza. La primera es la aproximación clásica a través del teorema del límite central; la segunda es mediante la aproximación descrita por Hewett (1995) haciendo uso de información a priori; la tercera es el método exacto usando funciones lineales de los límites estándar y métodos de Monte Carlo para su cálculo; y la cuarta hace uso de modelos lineales polinomiales del tamaño de la muestra, para niveles de confianza del 90 %, 95 % y 99 %, respectivamente. Los autores describen como la aproximación dada por el teorema del límite central subestima los tamaños requeridos. El uso de la fórmula reportada por Hewett (1995) obliga a los investigadores a realizar una muestra piloto para poder obtener los estimadores insesgados requeridos por la fórmula. El cálculo exacto del tamaño de la muestra requiere de algoritmos computacionales y no es una fórmula cerrada, por esta razón los autores reportan tablas en sus resultados. La cuarta forma que se basa en los cálculos del tamaño de la muestra exactos, ajustando modelos lineales polinomiales, lamentablemente presenta subestimaciones del valor real del tamaño de la muestra necesario en diferentes escenarios.

1.2. Tamaño de la muestra para estimar la media aritmética con censura

En la actualidad, la única referencia en la literatura estadística para calcular el tamaño de la muestra cuando se desea estimar la media aritmética de la distribución lognormal cuando existe censura, está reportada por Pérez y Lefante (1997). El mínimo tamaño requerido se basa en la amplitud del intervalo de confianza para la media aritmética de una variable aleatoria con distribución lognormal. Básicamente, es necesario especificar el porcentaje de diferencia respecto de la verdadera media aritmética (π), la desviación geométrica estándar estimada (GSD) y el porcentaje de censura esperado. Nuevamente los autores calculan los tamaños de la muestra exacta pero es necesario contar con algoritmos computacionales para su uso. Por esta razón reportan tablas con los resultados.

Por los problemas antes mencionados, el interés de este trabajo se centra en hallar una fórmula cerrada que facilite y así mismo evite en gran medida los problemas de subestimación en el cálculo del tamaño de muestra exacto requerido para la estimación de la media aritmética de la distribución lognormal, basados en el porcentaje de diferencia respecto de la verdadera media aritmética (π) y en la desviación geométrica estándar (GSD), para determinados niveles de confianza en el caso de datos sin censura y con censura a la izquierda.

2. Metodología

A partir de las ecuaciones reportadas por Pérez (1995) para el caso sin censura y niveles de confianza del 90 %, 95 % y 99 %, cuya forma general es $n = \exp(\beta_0)GSD^{\beta_1}GSD^{\beta_2 \ln(GSD)}\pi^{\beta_1}$ se inició la construcción de la ecuación que ajuste los tamaños exactos de la muestra reportados por Pérez y Lefante (1997). Las ecuaciones de la forma señalada anteriormente son linealizables mediante el uso de la función logaritmo natural, obteniendo

$$\ln(n) = \beta_0 + \beta_1 \ln(GSD) + \beta_2 \ln(GSD)^2 + \beta_1 \ln(\pi)$$

A pesar de las subestimaciones del tamaño de muestra que presenta esta ecuación, la descripción del comportamiento del tamaño de la muestra es adecuada, es decir, sigue una forma exponencial como la presentada por los datos exactos. Por tanto, como el objetivo es encontrar un modelo que presente un mejor ajuste, se tomó el anterior modelo como punto de inicio para la construcción de una fórmula no-lineal que ajuste mejor el tamaño exacto de la muestra con el menor grado de subestimación posible.

El primer paso en la metodología fue seleccionar el mismo modelo linealizable reportado por Pérez (1995) y agregar una constante para ajustar la subestimación,

con lo cual se obtuvo el modelo no lineal presentado en la ecuación 1.

$$n = \beta_0 + \exp(\beta_1)GSD^{\beta_2}GSD^{\beta_3Ln(GSD)}\pi^{\beta_4} \quad (1)$$

Este modelo no es linealizable mediante transformaciones y debe ser abordado por métodos no lineales. El modelo de la ecuación (1) mostró un mejor ajuste y un menor grado de subestimación del tamaño de la muestra comparado con lo obtenido en la ecuación anteriormente reportada. Sin embargo de forma gráfica se notó la presencia de una relación entre el porcentaje de diferencia respecto de la verdadera media aritmética π y la desviación geométrica estándar (GSD), por ello se decidió adicionar un nuevo término de interacción, generando el modelo presentado en la ecuación 2.

$$n = \beta_0 + \beta_1(\pi * GSD) + \exp(\beta_2)GSD^{\beta_3}GSD^{\beta_4Ln(GSD)}\pi^{\beta_5} \quad (2)$$

Al graficar la ecuación (2) se observó que para valores pequeños de la GSD se obtenían subestimaciones del tamaño de la muestra, mientras que para valores grandes de la GSD el ajuste del modelo a los datos era bastante cercano. Por tanto, se decidió ajustar dos nuevos modelos usando la ecuación (2), el primero para situaciones reales donde se tiene una $GSD \geq 2.5$ y el segundo modelo, tomando sólo los tamaños de la muestra para valores de la GSD de 2.5 o mayores.

En el caso de datos censurados, los tamaños de la muestra para los distintos niveles de confianza tenían un comportamiento similar al encontrado en el caso sin censura, así, la metodología empleada fue la misma. Sin embargo, en contraste con lo presentado en el caso sin censura, se encontró que la subestimación dependía más del porcentaje de diferencia respecto de la verdadera media aritmética π , que de la desviación geométrica estándar (GSD); es por ello que se ajustó el modelo dependiendo del (π) que de desee en la estimación de la media aritmética. Los modelos ajustados para este caso se presentan en la tabla 2.

Para obtener las estimaciones de los parámetros se empleó el método iterativo de Gauss-Newton y como valores iniciales de los parámetros se tomaron las estimaciones presentadas en los modelos lineales mostrados por Pérez(1995). Para la evaluación del modelo no lineal ajustado se calculó la medida de asimetría de Hougaard (1983), mediante el procedimiento NLIN del sistema SAS 8.2. Adicionalmente se planeó la evaluación de las medidas de no linealidad intrínseca y el efecto del parámetro estimado, Bates y Watts (1980), por medio de las subrutinas del programa RENOL (1994).

Adicionalmente, con el programa estadístico Stata 6.0 se calculó el coeficiente de concordancia y correlación de Lin (1989) para medir la concordancia entre el valor exacto del tamaño de la muestra y el tamaño de la muestra ajustado mediante los modelos no lineales.

3. Resultados

Las ecuaciones ajustadas, corresponden a modelos no lineales de la forma $n = \beta_0 + \beta_1(\pi \times GSD) + \exp(\beta_2) GSD^{\beta_3} GSD^{\beta_4 \ln(GSD)} \pi^{\beta_5}$, las cuales se presentan con los valores de los parámetros en las tablas 1 y 2. En la tabla 1 se presentan los modelos no lineales para estimar el tamaño de la muestra requerido en el caso de no censura, según el valor de la GSD y niveles de confianza del 90 %, 95 % y 99 %, respectivamente.

Tabla 1. Tamaño de la muestra para estimar la media aritmética de una distribución lognormal para niveles de confianza del 90 %, 95 % y 99 %, cuando no hay censura.

Nivel de Confianza	Modelo Estimado
90 %	<p>Si $GSD < 2.5$ $n = 13,86 + 0,86(\pi * GSD) + \exp(-4,51)GSD^{15,01}GSD^{-9,75Ln(GSD)}\pi^{-1,72}$</p> <p>Si $GSD \geq 2.5$ $n = 2,30 + 2,84(\pi * GSD) + \exp(-1,22)GSD^{3,98}GSD^{-0,68Ln(GSD)}\pi^{-1,73}$</p>
95 %	<p>Si $GSD < 2.5$ $n = 11,32 + 1,57(\pi * GSD) + \exp(-4,23)GSD^{15,34}GSD^{-10,14Ln(GSD)}\pi^{-1,73}$</p> <p>Si $GSD \geq 2.5$ $n = 11,68 + 4,17(\pi * GSD) + \exp(-0,78)GSD^{3,92}GSD^{-0,67Ln(GSD)}\pi^{-1,71}$</p>
99 %	<p>Si $GSD < 2.5$ $n = 18,92 + 2,30(\pi * GSD) + \exp(-3,56)GSD^{15,24}GSD^{-10,12Ln(GSD)}\pi^{-1,70}$</p> <p>Si $GSD \geq 2.5$ $n = 12,87 + 6,78(\pi * GSD) + \exp(-0,06)GSD^{3,82}GSD^{-0,64Ln(GSD)}\pi^{-1,66}$</p>

GSD = Desviación geométrica estándar.

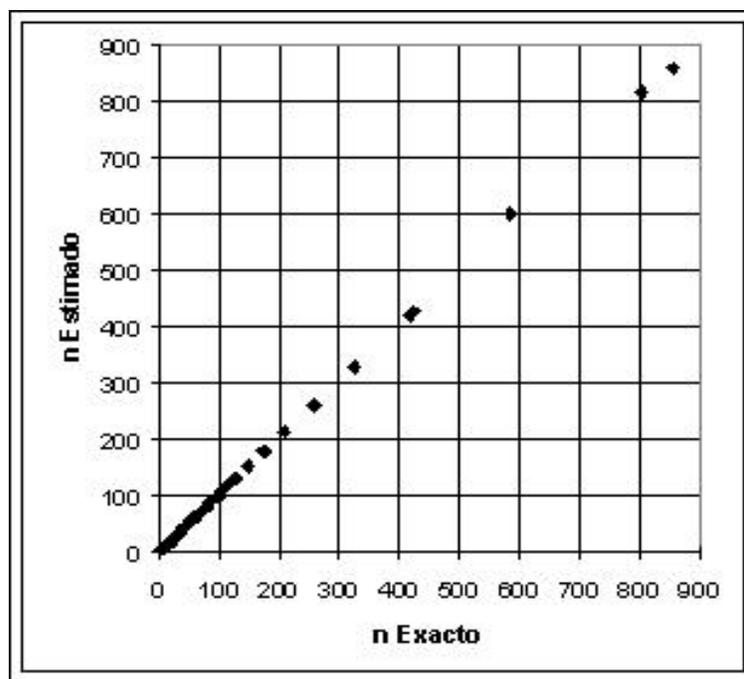
π = porcentaje de diferencia respecto de la verdadera media aritmética

La tabla 2 contiene los modelos no lineales según el porcentaje de diferencia respecto de la verdadera media aritmética π , para estimar el tamaño de la muestra requerido cuando se tiene diferentes porcentajes de censurados (10 %, 20 % y 50 %) y para niveles de confianza del 90 %, 95 % y 99 %, respectivamente. La medida

de asimetría de Hougaard (1983) fue calculada para todos los parámetros y en todos los casos se obtuvo un valor menor que 0,1, lo cual permite asumir un comportamiento aproximadamente lineal de sus estimadores.

Los valores calculados del estadístico de Lin (1989) para todos los casos sin censura y con censura mostrados en las tablas 1 y 2, fueron mayores de 0.999 con intervalo de confianza aproximado del 95 % de [0.99; 1]. Este valor del estadístico de Lin (1989) corresponde a una excelente concordancia entre el tamaño de la muestra exacto y el ajustado.

La gráfica 1 correspondiente al diagrama de dispersión del tamaño de la muestra exacto contra el tamaño de la muestra ajustado para el caso de no censura, intervalo de confianza del 90 % y $GSD < 2.5$, muestra una línea recta con intercepto 0 y pendiente 1, que corresponde a una perfecta concordancia. Para los demás casos presentados en las tablas 1 y 2, el gráfico de dispersión muestra una línea recta de iguales características, lo cual refleja el buen ajuste de los modelos propuestos y está disponible a través de los autores.



Gráfica 1. Diagrama de dispersión del tamaño de la muestra exacto contra el tamaño de la muestra ajustado para el caso de no censura, I.C. 90 % y $GSD < 2.5$.

Tabla 2. Tamaño de muestra para estimar la media aritmética de una distribución lognormal cuando hay censura, para niveles de confianza del 90 %, 95 % y 99 % y porcentajes de censura del 10 %, 20 % y 50 %.

Modelo Estimado	
I.C.90 %	Censura 10 %
Si $\pi < 0.08$	$n = 37,26 - 280,4(\pi * GSD) + \exp(-2,15)GSD^{6,91}GSD^{-2,37Ln(GSD)}\pi^{-1,66}$
Si $\pi \geq 0.08$	$n = 15,60 - 4,25(\pi * GSD) + \exp(-1,41)GSD^{4,71}GSD^{-1,00Ln(GSD)}\pi^{-1,70}$
I.C.90 %	Censura 20 %
Si $\pi < 0,08$	$n = 22,0 - 136,5(\pi * GSD) + \exp(-2,10)GSD^{6,80}GSD^{-2,30Ln(GSD)}\pi^{-1,70}$
Si $\pi \geq 0,08$	$n = 18,12 - 4,72(\pi * GSD) + \exp(-1,20)GSD^{4,67}GSD^{-1,00Ln(GSD)}\pi^{-1,67}$
I.C.90 %	Censura 50 %
Si $\pi < 0,13$	$n = 32,18 - 57,01(\pi * GSD) + \exp(-1,14)GSD^{6,14}GSD^{-1,82Ln(GSD)}\pi^{-1,70}$
Si $\pi \geq 0,13$	$n = 9,71 - 6,81(\pi * GSD) + \exp(-0,48)GSD^{4,45}GSD^{-0,92Ln(GSD)}\pi^{-1,61}$
I.C.95 %	Censura 10 %
Si $\pi < 0.10$	$n = 20,47 - 80,37(\pi * GSD) + \exp(-1,92)GSD^{6,76}GSD^{-2,27Ln(GSD)}\pi^{-1,72}$
Si $\pi \geq 0.10$	$n = 14,74 - 5,30(\pi * GSD) + \exp(-0,97)GSD^{4,63}GSD^{-0,97Ln(GSD)}\pi^{-1,67}$
I.C.95 %	Censura 20 %
Si $\pi < 0,13$	$n = 22,57 - 35,83(\pi * GSD) + \exp(-1,36)GSD^{5,94}GSD^{-1,73Ln(GSD)}\pi^{-1,67}$
Si $\pi \geq 0,13$	$n = 13,34 + 5,50(\pi * GSD) + \exp(-0,73)GSD^{4,50}GSD^{-0,92Ln(GSD)}\pi^{-1,64}$
I.C.95 %	Censura 50 %
Si $\pi < 0,18$	$n = 38,87 + 16,40(\pi * GSD) + \exp(-1,06)GSD^{5,80}GSD^{-1,58Ln(GSD)}\pi^{-1,73}$
Si $\pi \geq 0,18$	$n = 13,54 - 7,79(\pi * GSD) + \exp(-0,02)GSD^{4,30}GSD^{-0,85Ln(GSD)}\pi^{-1,58}$
I.C.99 %	Censura 10 %
Si $\pi < 0.18$	$n = 29,43 + 18,87(\pi * GSD) + \exp(-0,99)GSD^{5,62}GSD^{-1,50Ln(GSD)}\pi^{-1,72}$
Si $\pi \geq 0.18$	$n = 14,42 + 6,28(\pi * GSD) + \exp(-0,30)GSD^{4,50}GSD^{-0,91Ln(GSD)}\pi^{-1,62}$
I.C.99 %	Censura 20 %
Si $\pi < 0,16$	$n = 19,81 - 2,28(\pi * GSD) + \exp(-0,89)GSD^{5,91}GSD^{-1,71Ln(GSD)}\pi^{-1,7}$
Si $\pi \geq 0,16$	$n = 9,73 - 7,74(\pi * GSD) + \exp(-0,08)GSD^{4,44}GSD^{-0,91Ln(GSD)}\pi^{-1,70}$
I.C.99 %	Censura 50 %
Si $\pi < 0,24$	$n = 26,18 + 31,42(\pi * GSD) + \exp(-0,29)GSD^{5,57}GSD^{-1,50Ln(GSD)}\pi^{-1,68}$
Si $\pi \geq 0,24$	$n = 9,22 + 10,50(\pi * GSD) + \exp(0,61)GSD^{4,21}GSD^{-0,83Ln(GSD)}\pi^{-1,54}$

GSD = Desviación geométrica estándar. π = Porcentaje de diferencia respecto de la verdadera media aritmética.

4. Conclusiones

La metodología presentada consiguió producir una formula cerrada que facilita la estimación del tamaño de la muestra y reduce en forma considerable las diferencias entre el valor exacto del tamaño de la muestra y su valor estimado. Si bien, se disponía de fórmulas aproximadas, éstas presentan algunas limitaciones, principalmente porque subestiman los tamaños de la muestra, produciendo en algunos casos resultados con características estadísticas diferentes a las planeadas en la etapa inicial de una investigación (Ej.: Niveles de potencia y confianza).

En conclusión, el ajuste de las ecuaciones no lineales a los valores exactos reportados por Pérez y Lefante (1997), en los casos de censura y sin censura consiguieron reducir de forma considerable las diferencias entre el valor exacto del tamaño de la muestra y su valor estimado (una diferencia promedio de 8 unidades por encima del tamaño exacto). En la mayoría de las situaciones, los modelos estimados presentan pequeñas sobre-estimaciones del tamaño de la muestra requerido para estimar la media aritmética de la distribución lognormal, lo cual desde el punto de vista de precisión en la estimación de la media aritmética, asegura de forma conservativa la validez estadística de las estimaciones.

Las fórmulas obtenidas tienen la ventaja de permitir el cálculo del tamaño de la muestra para valores de π y de la *GSD* que no se reportan en las tablas del tamaño de la muestra presentadas por Pérez (1995).

Por último, el resultado del cálculo de la medida de asimetría de Hougaard (1983) garantiza que los estimadores de los parámetros tienen un comportamiento aproximadamente lineal, Bates y Watts, y el valor 1 con intervalo de confianza del 95% [0.99; 1] del coeficiente de concordancia y correlación de Lin (1989) entre el valor exacto del tamaño de la muestra y el tamaño de la muestra con los modelos no lineales, refleja el buen ajuste de los modelos propuestos.

Agradecimientos. Esta investigación fue financiada por la Fundación para la Promoción de la Investigación y la Tecnología (contrato# 986), Colciencias (contrato # 124-2000) y la Pontificia Universidad Javeriana (código # 120-104R-0101-200). Los autores desean expresar sus agradecimientos al profesor Hugo Grisales de la Universidad de Antioquia por facilitar el uso de las subrutinas del programa RENOL y por sus valiosos comentarios.

Referencias

- [1] Schorp, MK. y Leyden, D.E. (2002) Distribution analysis of airborne nicotine concentrations in hospitality facilities. *Environment international*. Vol 27 (7), 567-578.

- [2] Lu, H.C. (2002) The statistical characters of PM10 concentration in Taiwan area. Vol. 36 (3) 491-502.
- [3] Filmon, R., Basle, M.F., Barbier, A., Chappard, D. (2000) Poly(2-hydroxy ethyl methacrylate)-alkaline phosphatase: A composite biomaterial allowing in vitro studies of bisphosphonates on the mineralization process. *Journal of Biomaterials Science*. Vol. 11 (8), 849-869.
- [4] Potts, R.J., Keenan, T.D., May, P.T. (2000) Radar Characteristics of Storms in the Sydney Area. *Monthly Weather Review*. Vol. 128(9), 3308-3320.
- [5] Smith, B.E. y Merceret, F. (2000) The Lognormal Distribution. *College Mathematics Journal*. Vol. 31(4), 259-261.
- [6] Panayi, A.E., Spyrou, N.M., Iversen, B.S., White, M.A., Part, P. (2002) Determination of cadmium and zinc in Alzheimer's brain tissue using inductively coupled plasma mass spectrometry. *Journal of the neurological sciences*. Vol 195 (1), 1-10.
- [7] Kendal, W.S. (2001) The size distribution of human hematogenous metastases. *Journal of theoretical biology*. Vol 211 (1), 29-38.
- [8] Leidel, N.A. y Busch, K.A. (1985) Statistical design and data analysis requirements. En *Patty's Industrial Hygiene and Toxicology*, editado por L.J. Cralley y L.V. Cralley. Segunda edición. Vol 3A. New York. John Wiley & Sons.
- [9] Evans, J.S. y Hawkins, N.C. (1988) The distribution of student's t-statistic for small samples from lognormal exposure distribution. *American Industrial and Hygiene Association Journal*. Vol 49(10), 512-515.
- [10] Rappaport, S.M., Selvin, S. y Roach, S.A. (1988) A strategy for assessing exposures with references to multiple limits. *Applied Industrial Hygiene*. Vol 3, 310-315.
- [11] Lambert, D., Peterson, B. y Terpenning, I. (1991) Nondetects, detection limits, and the probability of detection. *Journal of the American Statistician Association*. Vol 86(414), 266-277.
- [12] Harris, R.L. (1991) Measures to facilitate future epidemiologic studies. *Applied Occupational Environmental Hygiene*. Vol 6(6), 555-559.
- [13] Hewett, P. (1995) Sample size formulae for estimating the true arithmetic mean or geometric mean of lognormal exposure distributions. *American Industrial and Hygiene Association Journal*. Vol 56, 219-225.
- [14] Perez, A. y Lefante, J.J. (1997) Sample size determination and the effect of censoring when estimating the arithmetic mean of a lognormal distribution. *Communications in Statistics. Theory and Methods*. Vol 26(11), 2779-2801.

- [15] Armstrong, B. (1995) Sample size formulae for estimating the true arithmetic or geometric mean of lognormal exposure distributions. Discussion. American Industrial Hygiene Association Journal. Vol. 56, 1042-1043.
- [16] Pérez, A. (1995) Sample size approximation for estimating the arithmetic mean of a lognormally distributed random variable with and without type I censoring. Tesis de Disertación. Tulane University. New Orleans. U.S.A. No publicada.
- [17] Ratkowsky, D. A. (1983) Nonlinear regression modeling: a unified practical approach, Marcel Dekker, New York. SAS® for Unix version 8.2 (2000). SAS Institute Inc. North Carolina Chapel Hill.
- [18] Bates, D. M. y Watts, D. G. (1980) Relative curvature measures of nonlinearity. Journal of the Royal Statistical Society. Ser. B, Vol. 42,1-25.
- [19] Grisales R. Hugo. (1994) Medidas de no linealidad en regresión no lineal. Revista Colombiana de Estadística. No 29 y 30.
- [20] Intercooled Stata 6.0 for Windows. Stata Corporation (2000). University Drive East.
- [21] Lin I-Kuei L. (1989) A Concordance Correlation Coefficient to Evaluate Reproducibility. Biometrics. Vol. 45, 255-268.

