

## ON THE PROBLEM OF COMPARING THE MEANS AND MEDIANs OF TWO INDEPENDENT LOGNORMAL DISTRIBUTIONS

José A. Montoya<sup>1</sup>, Gudelia Figueroa<sup>1</sup>, Gerardo Navarro<sup>2</sup>

<sup>1</sup>*Department of Mathematics, University of Sonora*

<sup>2</sup>*Department of Scientific and Technological Research, University of Sonora*

*Abstract:* Inferences about the ratio of two lognormal means  $\delta$  can depend on plausible values of  $\rho$ , the ratio of the normal standard deviations associated to these distributions. This aspect is not usually considered in most of the analyses carried out in some applied sciences. In this paper we propose a profile likelihood approach that allows the comparison of two independent lognormal data sets in a more exhaustive way. Inferences about  $\delta$ ,  $\rho$  and  $(\delta, \rho)$  are jointly analyzed through a simple closed-form expression obtained for the profile likelihood function of the parameter vector  $(\delta, \rho)$ . A similar analysis is done for  $\psi$  and  $\rho$ , where  $\psi$  is the ratio of two lognormal medians, obtaining also a simple closed-form expression for the profile likelihood function of these parameters. These expressions allow us to construct likelihood contour plots that capture most of the information provided by the samples and become valuable to identify if a trade-off between the parameters under study occurs; in case of that, individual inferences should be analyzed carefully. A detailed series of Monte Carlo simulations are included; they illustrate the performance of profile likelihood and parametric bootstrap approaches, for different sample sizes and parameter values.

*Key words:* Likelihood contour plot; Nuisance parameter; Profile likelihood function; Hypothesis test; Confidence intervals.

### 1. Introduction

Lognormal distributions occurs frequently in various applications coming from ecology, biology, hydrology, medicine, human behavior and many other scientific fields, where measurements under analysis are positive, with a small mean and a large variance. A random variable is said to be lognormally distributed if its natural logarithm follows a normal distribution. Now, many real life problems involve the comparison of two independent lognormal data sets and, in most

of them the usual concern relies in their means comparison. That explains the large list of procedures that have been proposed to prove the equality of these population means. These approaches usually consider the difference or the ratio of these means, as the parameter of interest, and a decision is generally taken by constructing a confidence interval or by doing a hypothesis test for one of these parameters. For example, a procedure focused on a difference of these means is the one proposed by Chen (1994), where the author explored the advantages of a new test over a t-test based on log-data. Some well known procedures to analyze the ratio of these means can be found in Zhou et al. (1997); in this paper, the authors proposed two methods for testing the null hypothesis of equality of means: a  $Z$ -score test and a nonparametric bootstrap method. On the other hand, Chen and Zhou (2006) also discussed various approaches for constructing confidence intervals for the ratio of these means, particularly they include two methods based on the log-likelihood ratio statistic and a generalized pivotal approach, among others. Currently, comparing the means of two lognormal distributions is still a problem of interest and different test statistics are proposed and also compared to measure, in a certain way, their performance (Jiang et al., 2014).

Although the comparison of two lognormal means is usually the problem of interest, in many practical situations comparing the medians of two lognormal distributions could be more appropriate to respond some specific questions or maybe it could be easier to work with. A comparison of the ratio of two lognormal means and two lognormal medians is made throughout this paper, using the same proposed methodology.

The usual analysis of the ratio of two lognormal means or two lognormal medians, considers that  $X = (X_1, \dots, X_n)$  and  $Y = (Y_1, \dots, Y_m)$  are two independent random samples from lognormal distributions, where  $\ln(X_i) \sim N(\mu_1, \sigma_1^2)$  and  $\ln(Y_j) \sim N(\mu_2, \sigma_2^2)$ ,  $i = 1, \dots, n$ ,  $j = 1, \dots, m$ . When comparing two lognormal means, inferences are made on

$$\delta = \exp [(\mu_1 - \mu_2) + \sigma_2^2 (\rho^2 - 1) / 2], \quad (1)$$

the ratio of these means, where  $\rho = \sigma_1 / \sigma_2$  denotes the ratio of the normal standard deviations associated to these lognormal distributions. When the purpose relies in comparing the medians of these distributions, inferences are made on

$$\psi = \exp(\mu_1 - \mu_2). \quad (2)$$

Inferences about  $\delta$  can rely heavily on plausible values of  $\rho$ , as we can observe in (1). Now, although  $\rho$  is not involved in (2), it is recommendable to have analytical and graphical tools that allow us to analyze if changes on plausible values of  $\rho$  could affect a decision about  $\delta$  or  $\psi$ . This is not the common manner this kind of problem is addressed in most of the analyses carried out in some applied sciences,

and important information could be throw out when the behavior of  $\rho$  is discarded from the analysis.

Díaz-Francés and Sprott (2004) studied the ratio of the means of two independent normal distributions. They explored in detail how different and plausible values for the variance ratio can influence a decision about the ratio of these means. Sprott and Farewell (1993) made a similar analysis for the difference of two normal means. They showed that inferences about this parameter can be sensitive to the assumption of equality of variances. They explored how different plausible values for the variance ratio, influence on the inferences about the difference between two normal means. They emphasize that this type of analysis is not restricted by the assumption of normality and that it applies equally, but with more computational complexity, to the difference in location between any two location-scale distributions.

Taking into consideration the analyses made by Díaz-Francés and Sprott (2004) and Sprott and Farewell (1993), in Section 2 we propose a profile likelihood approach that allows us to compare two independent lognormal data sets in a more exhaustive way. Inferences about  $\delta$ ,  $\rho$  and  $(\delta, \rho)$  are jointly analyzed through a simple closed-form expression obtained for the profile likelihood function of the parameter vector  $(\delta, \rho)$ . This expression allows us to construct a likelihood contour plot that capture most of the information provided by the samples. The location of the point  $(\delta = 1, \rho = 1)$  and the profile likelihood-confidence intervals for single parameters  $\delta$  and  $\rho$  are included in this contour plot, making possible not only to jointly visualize one and two dimensional inferences, but also to detect a possible relationship between these parameters. Now, inferences about  $\psi$ ,  $\rho$  and  $(\psi, \rho)$  were also achieved using this approach. The plots obtained are so valuable to identify if inferences for one of the parameters can be affected by the behavior of the other one. Lindsey (2006, pp.106) emphasizes that when parameters are inter-related, that is, when the likelihood contours are not parallel to the axes, they should not be treated individually, except approximately.

Section 3 is devoted to the analysis of a series of Monte Carlo simulations that were performed to compare the coverage probability of profile likelihood and bootstrap intervals, their type I errors, as well as the power for the profile likelihood and bootstrap tests. The election of sample sizes and parameter values, considered in these simulations, were based on the characteristics of the four examples included in Section 4, that were carefully selected to illustrate the utility of our proposal; actually, three of them are very well known data sets in the statistical literature. In the first example we studied a data set presented by Wu et al. (2002); they compared the maximum of plasma concentration in two formulations. In our second and third example we analyzed two data sets studied by Krishnamoorthy and Mathew (2003); these authors compared the amount

of rainfall for seeded and unseeded clouds and the carbon monoxide levels in two locations. Our last example deals with a data set that comes from a study made by Navarro-García et al. (2014), where the interest relies in comparing the amount of fatty acids in two ray species.

## 2. Profile likelihood inferences

Consider again that  $X = (X_1, \dots, X_n)$  and  $Y = (Y_1, \dots, Y_m)$  are two independent random samples from lognormal distributions, where  $\ln(X_i) \sim N(\mu_1, \sigma_1^2)$ ,  $\ln(Y_j) \sim N(\mu_2, \sigma_2^2)$ ,  $i = 1, \dots, n$ ,  $j = 1, \dots, m$ . The ratio of the means is  $\delta = \exp[(\mu_1 - \mu_2) + \sigma_2^2(\rho^2 - 1)/2]$ , where  $\rho = \sigma_1/\sigma_2$ , and the ratio of the two medians is  $\psi = \exp(\mu_1 - \mu_2)$ . Let  $x = (x_1, \dots, x_n)$  and  $y = (y_1, \dots, y_m)$  denote observed values of  $X$  and  $Y$ , respectively. The purpose here is to provide profile likelihood inferences about  $(\delta, \rho)$ ,  $(\psi, \rho)$  as well as  $\delta$ ,  $\psi$  and  $\rho$ . The profile likelihood functions for  $\delta$  and  $\psi$  will be calculated based on the profile likelihood functions of  $(\delta, \rho)$  and  $(\psi, \rho)$ , respectively. The profile likelihood function for  $\rho$  can be derived from any one of these, using the invariance property of the likelihood function.

### 2.1 Likelihood function for $(\delta, \rho)$

The likelihood function of the vector parameter  $(\mu_1, \sigma_1, \mu_2, \sigma_2)$  can be written as

$$L(\mu_1, \sigma_1, \mu_2, \sigma_2; x, y) \propto \sigma_1^{-n} \sigma_2^{-m} \exp \left\{ -\frac{1}{2\sigma_1^2} \sum_{i=1}^n [\ln(x_i) - \mu_1]^2 - \frac{1}{2\sigma_2^2} \sum_{j=1}^m [\ln(y_j) - \mu_2]^2 \right\}. \quad (3)$$

Now, consider the one to one transformation  $\mu_1, \rho, \delta, \sigma_2 \longleftrightarrow \mu_1, \sigma_1, \mu_2, \sigma_2$ , where

$$\sigma_1 = \rho\sigma_2 \quad \text{and} \quad \mu_2 = \mu_1 - \ln(\delta) - \frac{1}{2}\sigma_2^2(1 - \rho^2). \quad (4)$$

Substituting (4) into (3)

$$L(\mu_1, \delta, \rho, \sigma_2; x, y) \propto \rho^{-n} \sigma_2^{-(n+m)} \exp \left\{ -\frac{1}{2\rho^2\sigma_2^2} \sum_{i=1}^n [\ln(x_i) - \mu_1]^2 - \frac{1}{2\sigma_2^2} \sum_{j=1}^m \left[ \ln(y_j) - \mu_1 + \ln(\delta) + \frac{1}{2}\sigma_2^2(1 - \rho^2) \right]^2 \right\} \quad (5)$$

is obtained. Using the invariance property of the likelihood, the maximum likelihood (ML) estimates for  $\delta$  and  $\rho$  result

$$\hat{\delta} = \exp \left[ (\hat{\mu}_1 - \hat{\mu}_2) + \frac{1}{2} \hat{\sigma}_2^2 (\hat{\rho}^2 - 1) \right] \quad \text{and} \quad \hat{\rho} = \hat{\sigma}_1 / \hat{\sigma}_2, \tag{6}$$

where

$$\begin{aligned} \hat{\mu}_1 &= \frac{1}{n} \sum_{i=1}^n \ln(x_i), \quad \hat{\sigma}_1^2 = \frac{1}{n} \sum_{i=1}^n [\ln(x_i) - \hat{\mu}_1]^2, \\ \hat{\mu}_2 &= \frac{1}{m} \sum_{j=1}^m \ln(y_j), \quad \hat{\sigma}_2^2 = \frac{1}{m} \sum_{j=1}^m [\ln(y_j) - \hat{\mu}_2]^2, \end{aligned}$$

are the ML estimates of  $\mu_1$ ,  $\sigma_1^2$ ,  $\mu_2$  and  $\sigma_2^2$ , respectively.

Based on the likelihood function (5), the profile likelihood function of  $(\delta, \rho)$  can be easily obtained in two steps, as follows:

**Step 1.** Replace  $\mu_1$  in (5) by its restricted ML estimate, for specified  $(\delta, \rho, \sigma_2)$ ,

$$\hat{\mu}_1(\delta, \rho, \sigma_2) = a + b\sigma_2^2,$$

where

$$a = \frac{S_1 + \rho^2 S_2 + m\rho^2 \ln(\delta)}{n + m\rho^2}, \quad b = \frac{m\rho^2 (1 - \rho^2)}{2(n + m\rho^2)},$$

$S_1 = \sum_{i=1}^n \ln(x_i)$  and  $S_2 = \sum_{j=1}^m \ln(y_j)$ . Then, the profile likelihood of  $(\delta, \rho, \sigma_2)$  can be written as

$$\begin{aligned} L_p(\delta, \rho, \sigma_2; x, y) &\propto \rho^{-n} \sigma_2^{-(n+m)} \exp \left[ -\frac{b^2}{2\rho^2 \sigma_2^2} \sum_{i=1}^n (A_i - \sigma_2^2)^2 \right. \\ &\quad \left. - \frac{d^2}{2\sigma_2^2} \sum_{j=1}^m (B_j - \sigma_2^2)^2 \right], \end{aligned} \tag{7}$$

where

$$A_i = \frac{\ln(x_i) - a}{b}, \quad B_j = \frac{\ln(y_j) - a + \ln(\delta)}{d}$$

and  $d = b - (1 - \rho^2) / 2$ .

**Step 2.** Replace  $\sigma_2^2$  in (7) by its restricted ML estimate, for specified  $(\delta, \rho)$ ,

$$\hat{\sigma}_2^2(\delta, \rho) = \begin{cases} \frac{M_1^* + M_2^*}{n+m}, & \rho = 1 \\ r \left\{ -1 + \sqrt{1 + \frac{2}{r} \left[ \frac{M_1 + \rho^2 M_2}{(n+m)\rho^2} \right]} \right\}, & \rho \neq 1 \end{cases},$$

where

$$M_1 = \sum_{i=1}^n [\ln(x_i) - a]^2, \quad M_2 = \sum_{j=1}^m [\ln(y_j) - a + \ln(\delta)]^2,$$

$$M_1^* = \sum_{i=1}^n [\ln(x_i) - a^*]^2, \quad M_2^* = \sum_{j=1}^m [\ln(y_j) - a^* + \ln(\delta)]^2,$$

$a^* = [S_1 + S_2 + m \ln(\delta)] / (n + m)$  and  $2r = (n + m) \rho^2 / (nb^2 + m\rho^2 d^2)$ . Then the profile likelihood of  $(\delta, \rho)$  can be expressed as

$$L_p(\delta, \rho; x, y) \propto L_p(\delta, \rho, \sigma_2 = \hat{\sigma}_2(\delta, \rho); x, y). \tag{8}$$

### 2.2 Likelihood function for $(\psi, \rho)$

To obtain the likelihood function of  $(\psi, \rho)$ , we will consider a one to one transformation  $\mu_1, \rho, \psi, \sigma_2 \longleftrightarrow \mu_1, \sigma_1, \mu_2, \sigma_2$ , throughout the reparametrization

$$\sigma_1 = \rho\sigma_2 \quad \text{and} \quad \mu_2 = \mu_1 - \ln(\psi).$$

Following the same procedure shown in previous section, it is easy to obtain the restricted ML estimate of  $\mu_1$ , for a specified  $(\psi, \rho, \sigma_2)$ , resulting

$$\hat{\mu}_1(\psi, \rho, \sigma_2) = \frac{S_1 + \rho^2 S_2 + m\rho^2 \ln(\psi)}{n + m\rho^2},$$

where again,  $S_1 = \sum_{i=1}^n \ln(x_i)$  and  $S_2 = \sum_{j=1}^m \ln(y_j)$ . Note that  $\hat{\mu}_1(\psi, \rho, \sigma_2) = a$ , where  $a$  is given in Step 1.

Finally, the profile likelihood of  $(\psi, \rho, \sigma_2)$  becomes

$$L_p(\psi, \rho, \sigma_2; x, y) \propto \rho^{-n} \sigma_2^{-(n+m)} \exp \left\{ -\frac{1}{2\rho^2 \sigma_2^2} \sum_{i=1}^n [\ln(x_i) - a]^2 - \frac{1}{2\sigma_2^2} \sum_{j=1}^m [\ln(y_j) - a + \ln(\psi)]^2 \right\},$$

from which we can obtain the restricted ML of  $\sigma_2^2$ , for a specified  $(\psi, \rho)$ ,

$$\hat{\sigma}_2^2(\psi, \rho) = \frac{M_1 + \rho^2 M_2}{(n + m)\rho^2},$$

where

$$M_1 = \sum_{i=1}^n [\ln(x_i) - a]^2, \quad M_2 = \sum_{j=1}^m [\ln(y_j) - a + \ln(\psi)]^2,$$

and the profile likelihood of  $(\psi, \rho)$  can be expressed as:

$$L_p(\psi, \rho; x, y) \propto \rho^m (M_1 + \rho^2 M_2)^{-\frac{1}{2}(n+m)}. \tag{9}$$

### 2.3 Profile likelihood inferences about $(\delta, \rho)$

The relative profile likelihood function of  $(\delta, \rho)$  is defined as the ratio of the profile likelihood function  $L_p(\delta, \rho; x, y)$  to its maximum  $L_p(\hat{\delta}, \hat{\rho}; x, y)$ ; so it is a standardized version to be one at the maximum of the profile likelihood function of  $(\delta, \rho)$  given in (8),

$$R_p(\delta, \rho; x, y) = \frac{L_p(\delta, \rho; x, y)}{\max_{(\delta, \rho) \in \mathbb{R}^+ \times \mathbb{R}^+} L_p(\delta, \rho; x, y)} = \frac{L_p(\delta, \rho; x, y)}{L_p(\hat{\delta}, \hat{\rho}; x, y)}, \tag{10}$$

where  $\hat{\delta}$  and  $\hat{\rho}$  are the ML estimates of  $\delta$  and  $\rho$  presented in (6). The relative profile likelihood thus varies between 0 and 1. This likelihood measures the plausibility of any specified value  $(\delta, \rho)$  relative to that of  $(\hat{\delta}, \hat{\rho})$ . The relative profile likelihood function (10) ranks all possible  $(\delta, \rho)$  values according to their plausibilities in the light of the observed samples  $(x, y)$ . Thus, from a graph of  $R_p(\delta, \rho; x, y)$  we can easily distinguish plausible and implausible values for  $(\delta, \rho)$ .

Now, under the null hypothesis  $H_0 : (\delta, \rho) = (\delta_0, \rho_0)$ , the profile likelihood ratio statistic  $-2 \ln [R_p(H_0; X, Y)]$  follows an asymptotically  $\chi_2^2$  distribution; see Serfling (2002, p. 155). Thus, at the significance level  $\alpha$ ,

$$H_0 : (\delta, \rho) = (\delta_0, \rho_0) \text{ is rejected if } -2 \ln [R_p(H_0; x, y)] > q_{2,1-\alpha},$$

where  $q_{2,1-\alpha}$  is the quantile of probability  $1 - \alpha$  of a chi-squared distribution with 2 degrees of freedom. In this case the  $p$ -value can be computed as

$$p\text{-value} = P \{ \chi_2^2 \geq -2 \ln [R_p(H_0; x, y)] \} = R_p(H_0; x, y) = R_p(\delta_0, \rho_0; x, y). \tag{11}$$

Now, a level  $c$  profile likelihood region for  $\delta$  and  $\rho$ , is given by

$$\{(\delta, \rho) : R_p(\delta, \rho; x, y) \geq c\},$$

where  $0 \leq c \leq 1$ . Any pair of values  $(\delta, \rho)$  within this region, has a relative profile likelihood  $R_p(\delta, \rho; x, y) \geq c$ , and for those outside this region,  $R_p(\delta, \rho; x, y) < c$ ; so

just looking at the contour plot we can easily identify plausible and implausible values for  $(\delta, \rho)$ , at level  $c$ . Thus, varying  $c$  from 0 to 1, we obtain a complete set of nested likelihood regions that converges to the ML estimate  $(\hat{\delta}, \hat{\rho})$ , as  $c \rightarrow 1$ ; see Sprott (2000, pp. 14-15). A profile likelihood region for  $(\delta, \rho)$  is an approximate confidence region for  $(\delta, \rho)$ ; the approximate confidence level is determined from  $-2 \ln [R_p(\delta = \delta_0, \rho = \rho_0; X, Y)] \sim \chi_2^2$ , as exposed in Kalbfleisch (1985, pp. 189-190). For example, when  $c = 0.05$ , the confidence level is approximately 95%. Moreover, from (11), every specified value  $(\delta, \rho)$  has an associated  $p$ -value greater or equal than 0.05 if it lies within this contour, and less than 0.05 otherwise.

#### 2.4 Profile likelihood inferences about $\delta$ and $\rho$

The profile likelihood function of  $\delta$ ,  $L_p(\delta; x, y)$ , is obtained by maximizing over  $\rho$  the profile likelihood function of the parameter vector  $(\delta, \rho)$  given in (8), keeping  $\delta$  fixed,

$$L_p(\delta; x, y) \propto \max_{\rho \in \mathbb{R}^+ | \delta} L_p(\delta, \rho, \sigma_2 = \hat{\sigma}_2(\delta, \rho); x, y). \quad (12)$$

The relative profile likelihood function of  $\delta$  is a standardized version to be one at the maximum of the profile likelihood function of  $\delta$  given in (12),

$$R_p(\delta; x, y) = \frac{L_p(\delta; x, y)}{\max_{\delta \in \mathbb{R}^+} L_p(\delta; x, y)} = \frac{L_p(\delta; x, y)}{L_p(\hat{\delta}, \hat{\rho}; x, y)},$$

where  $\hat{\delta}$  and  $\hat{\rho}$  are the ML estimates of  $\delta$  and  $\rho$  presented in (6). This relative profile likelihood varies between 0 and 1 and ranks all possible  $\delta$  values based only on the observed samples  $(x, y)$ . Thus, a graph of  $R_p(\delta; x, y)$  allows to distinguish plausible and implausible values for  $\delta$ .

Considering that the profile likelihood ratio statistic  $-2 \ln [R_p(H_0; X, Y)]$  follows an asymptotically  $\chi_1^2$  distribution if the null hypothesis  $H_0 : \delta = \delta_0$  is true, then at the significance level  $\alpha$ ,

$$H_0 : \delta = \delta_0 \text{ is rejected if } -2 \ln [R_p(H_0; x, y)] > q_{1, 1-\alpha},$$

where  $q_{1, 1-\alpha}$  is the quantile of probability  $1-\alpha$  of a chi-squared distribution with 1 degree of freedom. In this case the  $p$ -value is given by  $P\{\chi_1^2 \geq -2 \ln [R_p(H_0; x, y)]\}$ .

A level  $c$  profile likelihood region (commonly an interval) for  $\delta$  is given by

$$\{\delta : R_p(\delta; x, y) \geq c\},$$

where  $0 \leq c \leq 1$ . Every specific value of  $\delta$  within this interval has a relative profile likelihood  $R_p(\delta; x, y) \geq c$ , and those  $\delta$  values outside this interval have a relative



profile likelihood  $R_p(\delta; x, y) < c$ . Then, at level  $c$ , this interval separates plausible and implausible values of  $\delta$ . Varying  $c$  from 0 to 1 we can obtain a complete set of nested likelihood intervals that converges to the ML estimate  $\hat{\delta}$  as  $c \rightarrow 1$ . Note that a profile likelihood interval for  $\delta$  is an approximate confidence interval for  $\delta$ , with approximate confidence level determined from  $-2 \ln [R_p(\delta = \delta_0; X, Y)] \sim \chi_1^2$ . For example, the confidence level will be approximately 90%, 95% and 99% if  $c = 0.25, 0.146$  and  $0.036$ , respectively; see Kalbfleisch (1985, pp. 115-116).

In the case of profile likelihood inferences for  $\rho$ , these can be obtained through-out the relative profile likelihood function of  $\rho$ , given by

$$R_p(\rho; x, y) = \frac{L_p(\rho; x, y)}{\max_{\rho \in \mathbb{R}^+} L_p(\rho; x, y)} = \frac{L_p(\rho; x, y)}{L_p(\hat{\delta}, \hat{\rho}; x, y)},$$

where

$$L_p(\rho; x, y) \propto \max_{\delta \in \mathbb{R}^+ | \rho} L_p(\delta, \rho, \sigma_2 = \hat{\sigma}_2(\delta, \rho); x, y).$$

## 2.5 Profile likelihood inferences about $(\psi, \rho)$ , $\psi$ and $\rho$

Even though Sections 2.3 and 2.4 are devoted to obtain the relative profile likelihood of  $(\delta, \rho)$  and profile likelihood inferences about  $\delta$  and  $\rho$ , the procedures used there can be implemented to obtain relative profile likelihood of  $(\psi, \rho)$  and profile likelihood inferences about  $\psi$  and  $\rho$ , considering  $L_p(\psi, \rho; x, y)$ , given in (9), as a starting point. Note that the profile likelihood of  $\rho$  can be obtained either from  $L_p(\psi, \rho; x, y)$  or  $L_p(\delta, \rho; x, y)$ , using the invariance property of the likelihood function.

All the computations required for the likelihood-based approach presented in this paper were programmed in the R software, this includes a non-linear minimization function named *nlm*, that was used to numerically compute the profile likelihood function of parameters  $\delta$ ,  $\psi$  and  $\rho$ . Now, even these computations are not difficult, these programs are freely available via the corresponding author.

## 3 Simulation

Some series of Monte Carlo simulations were conducted to compare the performance of profile likelihood approach and a parametric bootstrap approach based on maximum likelihood estimators. This comparison was made through the estimation of: coverage probability of Profile Likelihood (PL) and Bootstrap intervals, probability of type I error, and power for inverted PL and Bootstrap tests. Samples of sizes  $(n, m) = (10, 10), (20, 20), (26, 26), (50, 50), (20, 10), (31, 9), (40, 10), (50, 20)$  were simulated 10,000 times, considering the five parameter configurations shown in Table 1. These configurations were selected based

on parameter estimates obtained from real data, particularly the ones analyzed in the examples presented in Section 4. In the case of the bootstrap method, we used 1000 bootstrap replications of size  $(n, m)$ .

The parameter configuration (c) of Table 1, satisfies the null hypothesis  $H_0 : \delta = 1$  in the case of a means comparison, and  $H_0 : \psi = 1$  when comparing medians; the others depict a range of values of  $\delta$  and  $\psi$  that varies around 1. For each scenario, Table 1 includes different characteristics of the log-normal distributions under study, like their means ( $E$ ), medians ( $Med$ ) and skewness ( $Sw$ ). Some scenarios comparisons could be particularly interesting, for example in configuration (a) the absolute value of the skewness difference between these two distributions is 9.24, but in scenario (d) this difference is 0.02.

Sample sizes  $(n, m) = (10, 10), (20, 20), (26, 26)$  and  $(31, 9)$ , considered in these simulations, are the real sample sizes of the different data sets analyzed in the examples presented in Section 4. Some others were included, considering its occurrence in daily practice, where unequal as well as small sample sizes are usual.

Table 1: Parameter configurations for the simulation study.

Scenarios	$\mu_1$	$\sigma_1$	$\mu_2$	$\sigma_2$	$E(X)$	$E(Y)$	$Med(X)$	$Med(Y)$	$Sw(X)$	$Sw(Y)$	$\delta$	$\psi$	$\rho$
(a)	3.990	1.610	5.134	1.568	197.56	580.16	54.05	169.69	53.99	44.75	0.34	0.32	1.03
(b)	6.417	0.407	6.601	0.775	665.02	993.58	612.16	735.83	1.35	3.47	0.67	0.83	0.53
(c)	3	0.1	3	0.1	20.19	20.19	20.09	20.09	0.30	0.30	1	1	1
(d)	3.129	0.104	3.008	0.096	22.98	20.34	22.85	20.25	0.31	0.29	1.13	1.13	1.08
(e)	4.074	0.494	2.963	0.919	66.42	29.53	58.79	19.36	1.72	4.98	2.25	3.04	0.54

For each simulated scenario we calculate: proportion of times the interval cover the true parameter (central coverage), proportion of times the confidence interval falls below the true parameter (lower error) and proportion of times the lower limit of the confidence interval is greater than the true parameter (upper error). Theoretical values for central coverage, lower and upper errors are: 0.95, 0.025 and 0.025, respectively. Now, to compute empirical power in scenarios (a), (b), (d) and (e), we calculate the proportion of times the interval cover the null hypothesis  $H_0 : \delta = 1$ , in the case of means comparison, and  $H_0 : \psi = 1$  for the case of medians comparison. The proportion of times that the interval not include the true parameter value, represents the empirical type I error rate and its corresponding theoretical value is  $\alpha = 0.05$ .

Table 2 shows that the PL intervals for  $\delta$  are reasonably well behaved in all considered cases; they exhibit good coverage even for small and unequal sample sizes; nevertheless, although central coverage is controlled, the upper and lower errors are not necessarily symmetric for small and unequal sample sizes. On the other hand, most of the coverage frequencies for the bootstrap approach were nearest but below the ones obtained by the PL approach; even more, for scenarios (a), (b) and (e), the bootstrap approach revealed poor coverage, particularly

when  $(n, m) = (31, 9)$ . In addition to that, bootstrap approach appears to be more biased than the PL approach, revealing a strong bias toward the right, when unequal and small sample sizes are involved, such as in scenarios (a), (b) and (e). Some extra simulations, not reported here, suggest that the coverage frequencies and the bias obtained throughout bootstrap approach, become even worse when samples come from highly skewed distributions.

Table 2: Simulation results for the ratio of means. Each cell contains: lower error, central coverage\* and upper error.

$(n, m)$	Method	Scenarios				
		a	b	c	d	e
(10, 10)	PL	3.75(93.41)2.84	2.80(93.28)3.92	3.53(92.92)3.55	3.51(93.24)3.25	2.37(93.61)4.02
	Bootstrap	3.70(93.48)2.82	2.07(90.69)7.24	4.00(91.91)4.09	4.41(91.49)4.10	1.41(90.95)7.64
(20, 20)	PL	2.77(94.26)2.97	2.46(94.19)3.35	3.19(93.79)3.02	2.83(94.15)3.02	2.47(94.07)3.46
	Bootstrap	3.03(94.05)2.92	1.65(92.90)5.45	3.55(93.07)3.38	3.25(93.43)3.32	1.48(92.98)5.54
(26, 26)	PL	2.79(94.46)2.75	2.75(93.91)3.34	2.83(94.58)2.59	2.87(94.28)2.85	2.60(94.14)3.26
	Bootstrap	2.97(94.33)2.70	1.90(93.09)5.01	3.10(93.93)2.97	3.32(93.58)3.10	1.88(93.04)5.08
(50, 50)	PL	2.71(94.51)2.78	2.30(94.79)2.91	3.06(94.16)2.78	2.48(94.41)3.11	2.30(94.71)2.99
	Bootstrap	2.96(94.36)2.68	1.81(93.70)4.49	3.13(93.72)3.15	2.67(94.08)3.25	1.69(93.90)4.41
(20, 10)	PL	2.32(94.19)3.49	2.78(93.24)3.98	2.99(93.51)3.50	3.26(93.40)3.34	2.61(92.84)4.55
	Bootstrap	1.49(93.36)5.15	1.87(90.04)8.09	3.59(92.06)4.35	4.02(91.93)4.05	1.39(89.58)9.03
(31, 9)	PL	2.10(93.15)4.75	2.56(92.80)4.64	3.35(93.69)2.96	3.17(93.28)3.55	2.43(93.05)4.52
	Bootstrap	0.73(90.41)8.86	1.62(88.85)9.53	4.03(91.82)4.15	3.78(91.61)4.61	1.15(88.60)10.25
(40, 10)	PL	2.10(93.78)4.12	2.44(93.40)4.16	3.18(93.49)3.33	2.91(93.90)3.19	2.42(93.30)4.28
	Bootstrap	0.70(90.90)8.40	1.52(89.93)8.55	3.76(91.81)4.43	3.56(92.20)4.24	1.06(89.47)9.47
(50, 20)	PL	2.19(94.34)3.47	2.18(94.58)3.24	2.57(94.55)2.88	2.70(94.49)2.81	2.23(94.29)3.48
	Bootstrap	1.36(93.19)5.45	1.35(92.82)5.83	2.94(93.61)3.45	3.06(93.55)3.39	1.12(92.27)6.61

\* Coverage percentage is reported within brackets.

Table 3: Simulation results for the ratio of medians. Each cell contains: lower error, central coverage\* and upper error.

$(n, m)$	Method	Scenarios				
		a	b	c	d	e
(10, 10)	PL	3.39(93.57)3.04	3.43(93.17)3.40	3.31(93.29)3.40	3.40(93.39)3.21	3.33(93.40)3.27
	Bootstrap	4.09(91.93)3.98	4.56(91.07)4.37	4.09(91.79)4.12	4.25(91.70)4.05	4.37(91.48)4.15
(20, 20)	PL	2.85(94.19)2.96	2.80(94.25)2.95	3.16(93.94)2.90	2.75(94.34)2.91	2.67(94.77)2.56
	Bootstrap	3.28(93.32)3.40	3.51(93.19)3.30	3.51(93.21)3.28	3.18(93.43)3.39	3.22(93.73)3.05
(26, 26)	PL	2.86(94.66)2.48	2.84(94.39)2.77	2.89(94.52)2.59	2.87(94.38)2.75	2.98(94.37)2.65
	Bootstrap	3.04(94.12)2.84	3.17(93.53)3.30	3.14(94.01)2.85	3.32(93.61)3.07	3.38(93.61)3.01
(50, 50)	PL	2.54(94.73)2.73	2.72(94.67)2.61	3.03(94.28)2.69	2.45(94.53)3.02	2.72(94.57)2.71
	Bootstrap	2.66(94.44)2.90	2.87(94.22)2.91	3.27(93.71)3.02	2.62(94.13)3.25	2.87(94.07)3.06
(20, 10)	PL	2.97(93.83)3.20	3.57(93.10)3.33	3.04(93.54)3.42	3.34(93.31)3.35	3.49(92.89)3.62
	Bootstrap	3.60(92.40)4.00	4.64(90.80)4.56	3.68(92.17)4.15	4.24(91.78)3.98	4.63(90.65)4.72
(31, 9)	PL	2.95(93.39)3.66	3.50(92.79)3.71	3.28(93.88)2.84	3.27(93.36)3.37	3.72(92.97)3.31
	Bootstrap	3.83(91.55)4.62	4.83(90.27)4.90	4.40(91.86)3.74	4.17(91.58)4.25	4.74(90.57)4.69
(40, 10)	PL	3.36(93.52)3.12	3.65(92.97)3.38	3.36(93.33)3.31	3.02(93.84)3.14	3.33(93.53)3.14
	Bootstrap	4.22(91.81)3.97	4.57(90.85)4.58	4.25(91.62)4.13	3.90(92.11)3.99	4.43(91.34)4.23
(50, 20)	PL	2.69(94.21)3.10	2.84(94.50)2.66	2.61(94.53)2.86	2.83(94.43)2.74	2.69(94.44)2.87
	Bootstrap	3.00(93.46)3.54	3.41(93.27)3.32	3.04(93.68)3.28	3.14(93.75)3.11	3.27(93.35)3.38

\* Coverage percentage is reported within brackets.

Table 3 shows that the PL intervals for  $\psi$  are reasonably good in all considered cases, since coverage frequencies are close to the 95 per cent theoretical value, even for small and unequal sample sizes. Now, in the case of Bootstrap approach and for each simulated scenario, coverage frequencies are close to the ones obtained by PL approach, but they never surpass them. Moreover, neither the PL approach nor the bootstrap approach resulted in high bias. Additional simulations, not

reported here, suggest that the observed coverage and bias patterns are still maintained even when the samples come from fairly skewed distributions.

Table 4 summarizes the results obtained from empirical type I error and empirical power when carrying out tests about  $\delta$ . When increasing both sample sizes  $(n, m)$ , a larger and similar power for inverted bootstrap test and inverted PL test is obtained. In the case of scenario (a), the power obtained through the inverted bootstrap test is larger than the one obtained from the inverted PL test, except for the unequal sample size cases. For scenario (b), the inverted PL test had always a larger power. Now, for scenario (d) inverted bootstrap test had larger or equal power than inverted PL test, and in the case of scenario (e), inverted bootstrap test had always a larger power than inverted PL test. Nevertheless, excluding scenario (a), when  $(n, m) = (10, 10)$ , the probability of type I error is always greater for the inverted bootstrap test, showing also a strong bias to the right (See Table 2). This skewness could explain the increase in power observed for the bootstrap method in scenarios (d) and (e), where the true parameter value of  $\delta$  is greater than the null hypothesized value  $H_0 : \delta = 1$ . Now, the magnitude of this power could be diminished when the true parameter value of  $\delta$  is located to the left of the null hypothesized value  $H_0 : \delta = 1$ . Some extra simulations, no reported here, seem to confirm all these power discrepancy patterns.

Table 4: Powers of PL test and bootstrap test for  $\delta$  ( $\alpha = 0.05$ ).

$(n, m)$	Method	Scenarios			
		a	b	d	e
(10, 10)	PL	18.56(6.59)*	30.38(6.72)	76.80(6.76)	44.22(6.39)
	Bootstrap	25.31(6.52)	25.56(9.31)	80.00(8.51)	66.19(9.05)
(20, 20)	PL	28.22(5.74)	50.53(5.81)	97.02(5.85)	72.72(5.93)
	Bootstrap	35.19(5.95)	45.67(7.10)	97.38(6.57)	84.62(7.02)
(26, 26)	PL	35.23(5.54)	60.43(6.09)	99.08(5.72)	83.04(5.86)
	Bootstrap	41.79(5.67)	55.74(6.91)	99.16(6.42)	90.64(6.96)
(50, 50)	PL	57.21(5.49)	85.13(5.21)	100(5.59)	98.67(5.29)
	Bootstrap	62.76(5.64)	83.05(6.30)	100(5.92)	99.36(6.10)
(20, 10)	PL	22.86(5.81)	33.09(6.76)	86.61(6.60)	45.66(7.16)
	Bootstrap	22.29(6.64)	26.51(9.96)	89.61(8.07)	68.74(10.42)
(31, 9)	PL	25.67(6.85)	32.56(7.20)	87.59(6.72)	43.81(6.95)
	Bootstrap	20.20(9.59)	25.07(11.15)	91.37(8.39)	68.45(11.40)
(40, 10)	PL	29.43(6.22)	35.99(6.60)	91.65(6.10)	46.67(6.70)
	Bootstrap	21.63(9.10)	27.81(10.07)	94.45(7.80)	71.01(10.53)
(50, 20)	PL	42.46(5.66)	56.16(5.42)	99.38(5.51)	74.71(5.71)
	Bootstrap	39.38(6.81)	48.36(7.18)	99.54(6.45)	86.68(7.73)

\*Empirical power, with empirical type I error percentage in brackets.

Results from empirical type I error and empirical power of the tests for  $\psi$ , are presented in Table 5. As sample sizes  $(n, m)$  increase, the power of these tests gets larger and closer. In all simulated scenarios, inverted bootstrap test showed larger or equal power than inverted PL test; however, bootstrap method generates a larger empirical type I error. Additional simulations, no included here, seem to confirm these observations.

From our point of view, PL method outperforms bootstrap method regarding

Table 5: Powers of PL test and bootstrap test for  $\psi$  ( $\alpha = 0.05$ ).

$(n, m)$	Method	Scenarios			
		a	b	d	e
(10, 10)	PL	37.92(6.43)*	12.14(6.83)	76.39(6.61)	90.78(6.60)
	Bootstrap	42.08(8.07)	14.84(8.93)	79.68(8.30)	92.78(8.52)
(20, 20)	PL	62.47(5.81)	16.43(5.75)	96.87(5.66)	99.71(5.23)
	Bootstrap	64.60(6.68)	18.33(6.81)	97.33(6.57)	99.77(6.27)
(26, 26)	PL	73.64(5.34)	19.44(5.61)	99.06(5.62)	99.95(5.63)
	Bootstrap	75.16(5.88)	21.08(6.47)	99.10(6.39)	99.95(6.39)
(50, 50)	PL	94.77(5.27)	31.48(5.33)	100(5.47)	100(5.43)
	Bootstrap	95.07(5.56)	32.95(5.78)	100(5.87)	100(5.93)
(20, 10)	PL	46.66(6.17)	12.67(6.90)	87.13(6.69)	92.41(7.11)
	Bootstrap	50.73(7.60)	15.78(9.20)	89.33(8.22)	94.71(9.35)
(31, 9)	PL	48.13(6.61)	12.62(7.21)	88.86(6.64)	91.20(7.03)
	Bootstrap	53.15(8.45)	16.29(9.73)	91.16(8.42)	93.86(9.43)
(40, 10)	PL	53.73(6.48)	13.08(7.03)	92.56(6.16)	94.10(6.47)
	Bootstrap	58.20(8.19)	16.72(9.15)	94.41(7.89)	95.71(8.66)
(50, 20)	PL	76.52(5.79)	17.61(5.50)	99.45(5.57)	99.89(5.56)
	Bootstrap	78.39(6.54)	19.98(6.73)	99.53(6.25)	99.91(6.65)

\*Empirical power, with empirical type I error percentage in brackets.

coverage and also in the way it handles upper and lower error; for that reason and the simplicity of the closed-form expression of the profile likelihood, this methodology will be used from here on.

Table 6: Decomposition of type II error rate.

Scenarios	$(n, m)$	$H_0 : \delta = 1$				$H_0 : \psi = 1$			
		(0, 0)	(0, 1)	(1, 1)	(1, 0)	(0, 0)	(0, 1)	(1, 1)	(1, 0)
(a)	(10, 10)	3.95	16.63	78.45	0.97	5.09	1.43	91.70	1.77
	(20, 20)	5.03	33.57	60.66	0.74	4.77	1.95	91.61	1.68
	(26, 26)	5.87	42.63	50.98	0.52	5.20	1.86	91.35	1.59
	(50, 50)	8.97	70.93	20.03	0.07	3.44	3.25	92.35	0.96
	(20, 10)	5.24	19.95	73.75	1.06	3.96	1.99	92.07	1.99
	(31, 9)	7.13	18.00	73.60	1.26	3.78	2.99	91.19	2.04
	(40, 10)	7.31	20.89	70.96	0.84	3.59	3.07	91.74	1.60
	(50, 20)	7.42	40.16	52.03	0.38	4.30	2.64	91.91	1.15
(b)	(10, 10)	31.18	0.62	53.36	14.84	40.75	1.31	45.88	12.06
	(20, 20)	58.14	0.14	27.69	14.03	72.13	1.02	18.31	8.54
	(26, 26)	73.59	0.08	14.71	11.62	84.93	0.63	8.63	5.81
	(50, 50)	95.70	0.00	1.88	2.42	98.86	0.06	0.55	0.53
	(20, 10)	46.35	0.99	41.11	11.55	54.80	2.91	33.44	8.85
	(31, 9)	50.33	1.59	39.15	8.94	57.30	4.12	31.68	6.90
	(40, 10)	56.96	1.58	33.32	8.14	64.40	4.02	25.59	5.99
	(50, 20)	81.68	0.14	11.88	6.30	88.58	1.18	6.89	3.35

Observing Tables 4 and 5, it is intriguing that for scenarios (a) and (b), the power of the tests was smaller than the ones observed for scenarios (c) and (d); so it would be interesting to know how this not negligible type II error rate occurs. To answer this question, a decomposition of type II error rate for scenarios (a) and (b), is computed and presented in Table 6. Let  $(I_1, I_2)$  be a vector that displays the results of two different hypotheses; either  $I_1$  or  $I_2$  can take 0 and 1 values. The first entry of this vector corresponds to the result obtained when testing the vector hypothesis  $H_0 : (\delta, \rho) = (1, 1)$ , and the second one is related to testing  $H_0 : \rho = 1$ . In that sense, (0, 1) represents a situation when the hypothesis  $H_0 : (\delta, \rho) = (1, 1)$  is rejected, but individual hypothesis  $H_0 : \rho = 1$  is not rejected. This particular situation is very interesting because given that  $H_0 : \delta = 1$  is not rejected, the joint hypothesis  $H_0 : (\delta, \rho) = (1, 1)$  is rejected and

the individual hypothesis  $H_0 : \rho = 1$  results not rejected. Actually, this situation occurs markedly in scenario (a), for all sample sizes, and could be linked to an elongated  $(\delta, \rho)$  contour behavior, similar to the one obtained in the Rainfall Data example presented in Section 4. Regarding scenario (b), we can observe that frequencies below the  $(0, 1)$  column are very small, nevertheless frequencies below  $(0, 0)$  column are considerable, for all sample sizes.

A similar decomposition was done for the case of formulating these hypotheses in terms of parameter  $\psi$ . In the case of scenario (a), for all sample sizes, the frequencies below the  $(0, 1)$  column are small; it seems that not so much cases of elongated contours for  $(\psi, \rho)$  are present. In relation to scenario (b) the behavior is similar to the one observed for parameter  $\delta$ , in this scenario.

Regarding the results obtained when using the profile likelihood approach, those turn out interesting because on one hand, we could conclude that coverage frequencies seem to be controlled in all scenarios, but on the other hand they showed that we must be concerned about scenarios like (a), since their loss in power seems to be related to a situation where a strong relationship between parameters  $\delta$  and  $\rho$  exists, and that feature should not be discarded from an analysis. Simulations also showed that we should not worry about this situation if the parameters under analysis are  $\psi$  and  $\rho$ .

Overall, the analyses of coverage frequencies and power in all selected scenarios, support what we will emphasize in the examples presented in Section 4, regarding the importance of an analysis that considers the possibility of a trade-off between the parameters under study, since the exploration of one and two dimensional inferences are very helpful procedures before taking a decision.

#### 4 Illustrative examples

Profile likelihood function of an individual parameter  $(\delta, \psi$  or  $\rho)$ , computed as shown in Section 2, allows to make individual inferences about the parameter of interest. Nevertheless, when the analysis involve two parameters it is recommendable not to restrict our inferences to a single profile likelihood function, without considering the possibility of a relationship between these parameters. The examples included in this section were carefully selected to show this aspect considering different situations that can occur in real life, where the decision derived from individual and simultaneous inferences do not necessarily coincide. Although the latter is already known, some statistics users do not realize that this can occur. The examination of the likelihood contour plots of  $(\delta, \rho)$  or  $(\psi, \rho)$ , obtained from the profile likelihood computed in Section 2, and the additional features included in this plot, allow to compare two independent lognormal data sets in a more exhaustive way.

For all the examples considered in this section we present five different plots.

Three of them correspond to the profile likelihood function for parameters  $\delta$ ,  $\psi$  and  $\rho$ , including their respective 95% likelihood confidence intervals; the other two show a likelihood contour plot for parameters  $(\delta, \rho)$  and parameters  $(\psi, \rho)$ , both at two different levels: 0.146 and 0.05. The 0.146 level likelihood contour includes two dashed-line curves corresponding either to the trajectories of the profile likelihoods for  $\delta$  and  $\rho$  or the trajectories of the profile likelihoods for  $\psi$  and  $\rho$ , depending of the plot. In the same manner, their crossing point denoted by an asterisk, shows the location of either the ML estimate  $(\hat{\delta}, \hat{\rho})$  or the ML estimate  $(\hat{\psi}, \hat{\rho})$ . The end points of these trajectories are projected to their respective axes using dotted lines; this allows for obtaining their corresponding 95% confidence intervals (CI). On the other hand, the 0.05 level likelihood contour is associated to a 95% confidence region. The likelihood contour plot includes also a cross sign to identify the location of parameter values  $(\delta = 1, \rho = 1)$  or  $(\psi = 1, \rho = 1)$ , depending of the plot. This is an informative feature because if the cross sign lies into the outer contour, then we have no evidence to reject that the samples could come from the same distribution function. All the points outside the 0.05 likelihood contour have associated  $p$ -values smaller than 0.05, for this simultaneous test.

#### 4.1 Bioavailability study

The data analyzed in this example is presented in Wu et al. (2002) and it resulted from a bioavailability study where a randomized, parallel-group experiment was conducted considering 20 subjects. The purpose of this study was to compare a new test formulation ( $X$ ) with a reference formulation ( $Y$ ) of a drug product with a long half-life. It is important to determine if the means of the maximum plasma concentration ( $C_{\max}$ ) of the two formulations are the same.

Wu et al. (2002) analyzed the ratio of means of  $C_{\max}$  to determine if the two formulations have or not, different bioavailability. By means of QQ plots for the original and the log-transformed data, as well as the Shapiro-Wilk tests for the normality on the log-transformed data, they showed that a lognormal model adequately describes both data sets. They also made a hypothesis test for the equality of variances of the log-transformed data ( $H_0 : \rho = 1$ ) using an  $F$ -test, as a preliminary step to compare the means of  $C_{\max}$ ; this test produced a  $p$ -value of 0.034, so they concluded that the log transformation does not stabilize variances. To test the equality of means of  $C_{\max}$ , between these two formulations ( $H_0 : \delta = 1$ ), they computed a two-sided  $p$ -value of the modified signed log-likelihood ratio statistic ( $r^*$ ) obtaining 0.173, and a 95% confidence interval for the ratio of means  $\delta$  of  $C_{\max}$ , which resulted (0.242, 1.200). For comparison purposes, they also computed the two-sided  $p$ -values of the signed log-likelihood ratio statistic ( $r$ ) and the  $Z$ -score test, obtaining 0.167 and 0.203, respectively,

and their corresponding 95% confidence intervals that resulted (0.295, 1.181) and (0.339, 1.259), respectively. Thus, from this results they concluded that the samples do not provide evidence for rejecting the equality of the means of  $C_{\max}$  ( $H_0 : \delta = 1$ ).

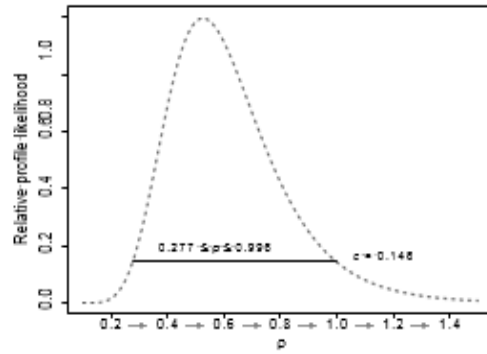


Figure 1: Bioavailability study data: Relative profile likelihood of  $\rho$ .

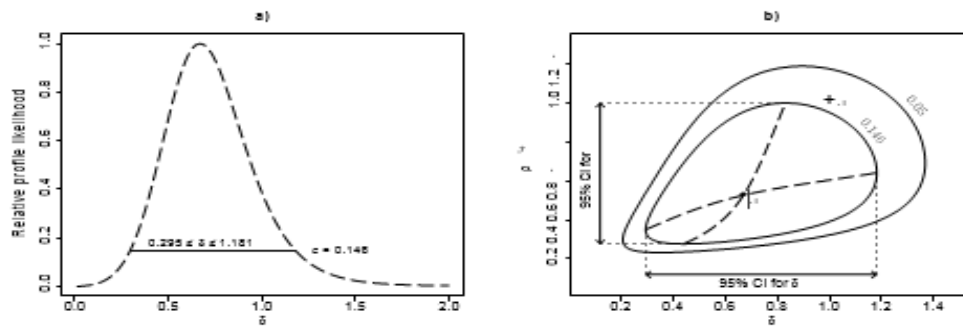


Figure 2: Bioavailability study data: a) Relative profile likelihood of  $\delta$ , b) Relative profile likelihood contour plot for  $(\delta, \rho)$  and individual 95% CI for  $\delta$  and  $\rho$

When we carry out one dimensional inferences for  $\delta$  and  $\rho$ , the assumption  $H_0 : \rho = 1$  is not supported by the data, as can be observed in Figure 1, where the 95% confidence interval for  $\rho$  resulted (0.277, 0.996). In the case of parameter  $\delta$  its 95% likelihood confidence interval, shown in Figure 2a, provides evidence that both data sets share the same mean. The conclusions obtained from these individual inferences coincide with the ones presented by Wu et al. (2002), and this can be easily observed just looking at the 0.146 level likelihood contour plot, shown in Figure 2b where, as we previously mentioned, the dotted lines projected from the end points of the profile likelihood trajectories to their respective axes,



provide these one dimensional 95% confidence intervals. In this figure we can also observe that the point  $(\delta = 1, \rho = 1)$ , represented by a cross sign, lies into the 0.05 level outer contour, which is associated to a 95% confidence region. That is, the assumption of the simultaneous parameter values  $(\delta = 1, \rho = 1)$  is supported by the data; there is no evidence to reject the assumption that both sets could come from the same distribution. Note that elongated contours, like the ones presented in Figure 2b, are indicative of a relationship between  $\delta$  and  $\rho$ , suggesting a trade-off between these parameters, so we must be careful about treating them individually. This aspect will be more evident in the next example.

Now, for the case of comparing the medians of these distributions, throughout the parameter  $\psi$ , its 95% likelihood confidence interval provides evidence that both data sets could come from distributions that share the same median, as can be observed in Figure 3a. On the other hand, Figure 3b shows a likelihood contour plot constructed for  $(\psi, \rho)$ ; the contours of this plot are not as elongated as the contours for  $(\delta, \rho)$ , and the point  $(\psi = 1, \rho = 1)$ , represented by a cross sign, lies again into the 0.05 level outer, associated to a 95% confidence region. That is, the assumption of simultaneous parameter values  $(\psi = 1, \rho = 1)$  is supported by the data and same conclusion, as the one obtained in the case of the means comparison, is derived.

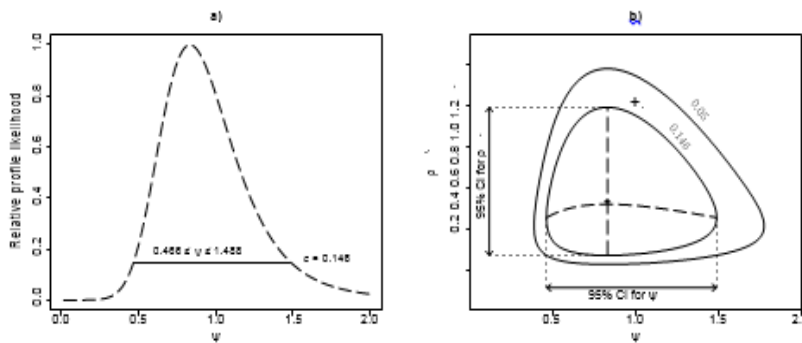


Figure 3: Bioavailability study data: a) Relative profile likelihood of  $\psi$ , b) Relative profile likelihood contour plot for  $(\psi, \rho)$  and individual 95% CI for  $\psi$  and  $\rho$ .

## 4.2 Rainfall data

Data about the amount of rainfall (in acre-feet) from 52 clouds, 26 of which were chosen at random and seeded with silver nitrate, was analyzed by Krishnamoorthy and Mathew (2003). They constructed probability plots for these particular data sets and found that lognormal models fit them very well. Their interest was to test the hypothesis  $H_0 : \delta = 1$  vs.  $H_1 : \delta > 1$ , so they constructed

a one-sided 95% generalized confidence interval for  $\ln(\delta)$  that produced a lower confidence limit of  $-0.20$ ; they also obtained a generalized  $p$ -value of  $0.078$ . On the other hand, they performed a  $Z$ -score test for the same parameter and the 95% lower confidence limit resulted  $-0.06203$ , and the corresponding  $p$ -value was  $0.06$ . They concluded that the data do not provide sufficient evidence to indicate that the mean rainfall from seeded clouds is higher than the mean rainfall from unseeded clouds.

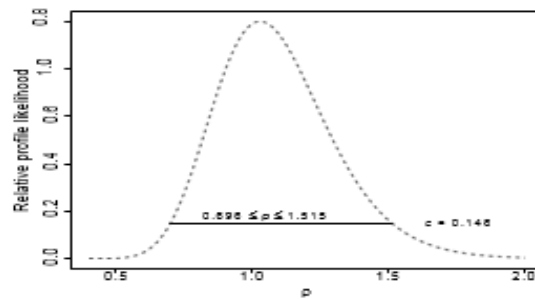


Figure 4: Rainfall data: Relative profile likelihood of  $\rho$ .

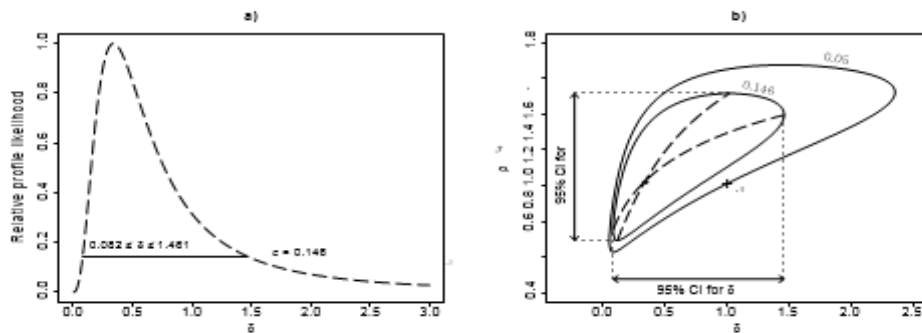


Figure 5: Rainfall data: a) Relative profile likelihood of  $\delta$ , b) Relative profile likelihood contour plot for  $(\delta, \rho)$  and individual 95% CI for  $\delta$  and  $\rho$ .

When we perform individual inferences, the profile likelihood confidence interval for  $\rho$ , that can be observed in Figure 4, provides no evidence to reject the hypothesis  $H_0 : \rho = 1$ . On the other hand, 95% likelihood interval for  $\delta$  which is presented in Figure 5a, indicates that the hypothesis  $H_0 : \delta = 1$  is supported by the data; both conclusions coincide with the results presented in Krishnamoorthy and Mathew (2003). Nevertheless, it could be erroneous to infer, based on these

individual tests, that these sets could come from the same distribution. Actually, the relative likelihood associated to this hypothesis is small ( $p$ -value  $< 0.05$ ), as can be observed in the 95% confidence region for  $(\delta, \rho)$ , presented in Figure 5b, where the point  $(\delta = 1, \rho = 1)$ , represented by a cross sign, lies outside de 0.05 level outer contour. The triangular and elongated contours shown in Figure 5b, make evident the strong relationship between these parameters, so we must be careful about treating them individually. This data set exemplifies a case where individual inferences for parameters  $\delta$  and  $\rho$  are not rejected, but the simultaneous hypothesis test  $(\delta = 1, \rho = 1)$  is rejected.

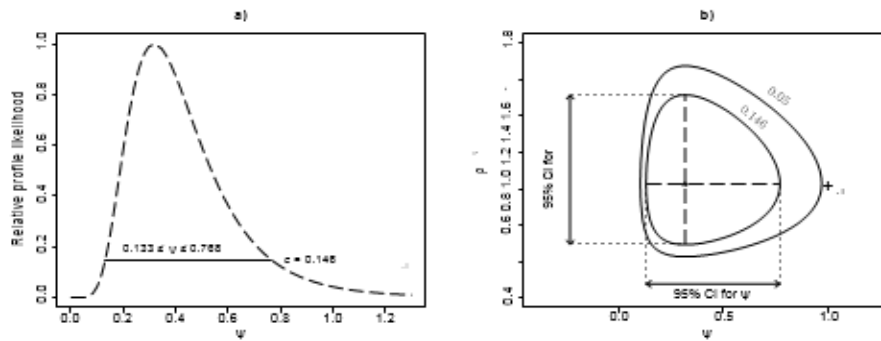


Figure 6: Rainfall data: a) Relative profile likelihood of  $\psi$ , b) Relative profile likelihood contour plot for  $(\psi, \rho)$  and individual 95% CI for  $\psi$  and  $\rho$ .

When using  $\psi$  to compare the medians of these distributions, its 95% likelihood confidence interval, presented in Figure 6a, indicates that the hypothesis  $H_0 : \psi = 1$  is not supported by the data, providing evidence that these data sets seem to come from distributions with different medians. On the other hand, a likelihood contour plot for  $(\psi, \rho)$ , presented in Figure 6b, confirms that the assumption of simultaneous parameter values  $(\psi = 1, \rho = 1)$  is not supported by the data, since the point  $(\psi = 1, \rho = 1)$ , represented by a cross sign, lies outside de 0.05 level outer contour.

### 4.3 Air quality data

The data analyzed in this example contains 31 consecutive daily Carbon Monoxide measurements taken by an oil refinery located at the northeast of San Francisco, between April 16 and May 16, 1993. Nine independent measurements, from the same stack, were taken by The Bay Area Air Quality Management District (BAAQMD) over the period from September 11, 1990 to March 30, 1993. Krishnamoorthy and Mathew (2003) observed that lognormal models fit these data sets very well and, to prove the hypothesis  $H_0 : \delta \leq 1$  vs.  $H_1 : \delta > 1$ ,

they computed a 95% generalized lower confidence limit for the ratio of these lognormal population means obtaining a value of 0.67, which is less than one, so they concluded that the samples do not provide sufficient evidence to indicate that the mean measurement by the refinery is greater than that of BAAQMD.

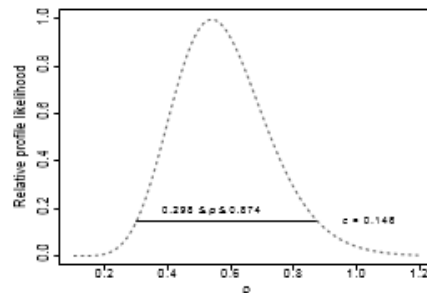


Figure 7: Air quality data: Relative profile likelihood of  $\rho$ .

The individual inferences that we performed on these samples showed that the assumption  $\rho = 1$  is not supported by the data, as can be observed in Figure 7, where the 95% confidence interval for  $\rho$  resulted (0.298, 0.874).

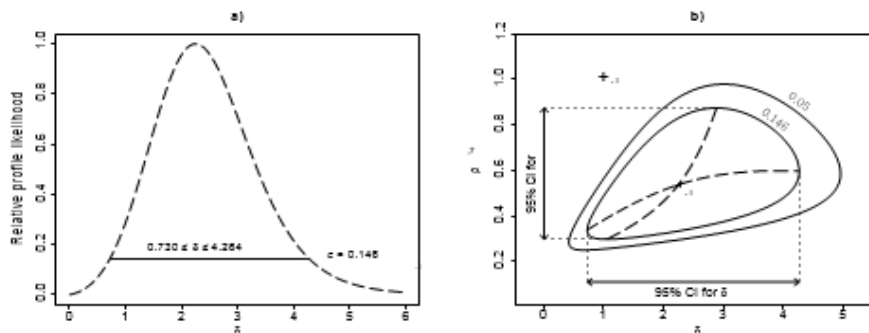


Figure 8: Air quality data: a) Relative profile likelihood of  $\delta$ , b) Relative profile likelihood contour plot for  $(\delta, \rho)$  and individual 95% CI for  $\delta$  and  $\rho$ .

Regarding parameter  $\delta$ , its 95% confidence interval, presented in Figure 8a, supports the assumption of equality of means, so same conclusion of Krishnamoorthy and Mathew (2003) can be derived. Inferences about  $\delta$  and  $\rho$  can be also observed in Figure 8b, where the cross sign ( $\delta = 1, \rho = 1$ ) is located outside and so far from the 0.05 level outer contour, that defines a 95% confidence region for  $(\delta, \rho)$ ; so there is evidence that the samples could come from different distributions. This data exemplify a situation where the triangular likelihood contours suggest

a relationship between  $\delta$  and  $\rho$ , indicating a trade-off between these parameters; so we must be careful about treating them individually.

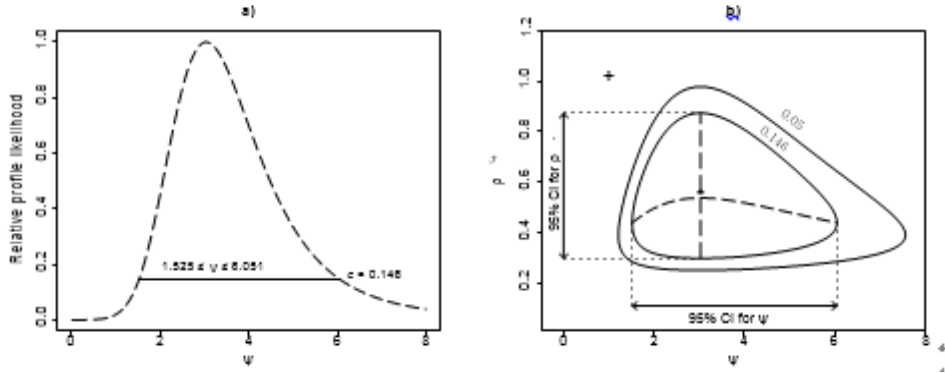


Figure 9: Air quality data: a) Relative profile likelihood of  $\psi$ , b) Relative profile likelihood contour plot for  $(\psi, \rho)$  and individual 95% CI for  $\psi$  and  $\rho$ .

A 95% likelihood confidence interval for  $\psi$  is used to compare the medians of these distribution, this can be observed in Figure 9a where that the hypothesis  $H_0 : \psi = 1$  is not supported by the data; so both data sets could come from distributions that do not share the same median. On the other hand, a likelihood contour plot for  $(\psi, \rho)$ , presented in Figure 9b, shows that the assumption of simultaneous parameter values ( $\psi = 1, \rho = 1$ ) is not supported by the data, because the point  $(\psi = 1, \rho = 1)$ , represented by a cross sign, lies outside de 0.05 level outer contour.

#### 4.4 Comparing fatty acids in two ray species

The term nutraceutical is coined from “nutrition” and “pharmaceutical”, and it is usually applied to some food sources that provide health benefits, like treatment and prevention of some diseases, in addition to the nutritional value they contain. The consumption of essential fatty acids, particularly the n-3 long chain polyunsaturated (LC-PUFA), like the eicosapentaenoic acid (EPA, C20:5) and the docosahexaenoic acid (DHA, C22:6) are highly recommended, because they are proved to reduce the risk of a coronary disease death (Harris and Dris-Etherton, 2008). Currently, oil fish is one of the most important sources of EPA and DHA, and wild salmon, cod and herring, are some of the fish species with a considerable content of these fatty acids. Rising demand of oil fish requires to find some alternative sources of these essential fatty acids and ray liver oil has been found to be a good alternative, as demonstrated by a study made in some ray species from Sinaloa, México (Navarro-García et al., 2014), where high levels of LC-PUFA

n-3 C20:5, C20:6 and C22:5 were found. This can be observed in the descriptive statistics for *U. chillensis* and *U. halleri* ray species, shown in Table 7.

Table 7: Percentage of content of Omega 3 in two ray species.

Specie	Sample Size	Minimum	Maximum	Mean	Std. Dev.
<i>U. chillensis</i>	20	17.84	26.68	22.97	2.41
<i>U. halleri</i>	20	17.01	25.91	20.34	2.06

A Shapiro-Wilk test for the normality of the log-transformed data was performed in both groups, obtaining a  $p$ -value of 0.3861 for the *U. chillensis* data set and a  $p$ -value of 0.6936 for the *U. halleri* ray specie. A difference in the mean content of Omega 3 in these species is suspected.

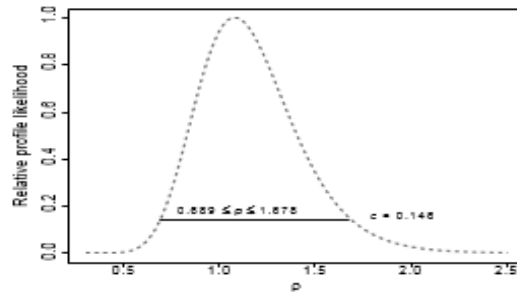


Figure 10: Ray species data: Relative profile likelihood of  $\rho$ .

In Figure 10 we can observe the 95% likelihood confidence interval for  $\rho$  which supports the assumption  $\rho = 1$ ; but the 95% likelihood confidence interval for  $\delta$ , presented in Figure 11a, indicates a difference in the mean content of Omega 3 for these species; even more, this interval (1.06, 1.204) gives evidence that the mean content of Omega 3 in *U. chillensis* is greater than the mean content of Omega 3 in *U. halleri*, since all interval values are greater than the parameter value  $\delta = 1$ . These individual inferences are also presented in Figure 11b, where we can observe that the point corresponding to the parameter values ( $\delta = 1, \rho = 1$ ) is located outside and so far from the 0.05 level outer contour, having evidence that these samples could come from different distributions. Moreover, this example shows a situation where the relationship between parameters  $\delta$  and  $\rho$  seems to be negligible, as evidenced by the nearly circular contours shown in Figure 11b, which suggest that the maximum likelihood estimation theory is probably adequate for the actual sample size.

A computation of a 95% likelihood confidence interval for  $\psi$  was used to compare the medians of these distributions; this interval, shown in Figure 12a,

indicates that the hypothesis  $H_0 : \psi = 1$  is not supported by the data. Then, both data sets could come from distributions with different median. On the other hand, the likelihood contour plot for  $(\psi, \rho)$ , presented in Figure 12b, provides evidence that the assumption of simultaneous parameter values  $(\psi = 1, \rho = 1)$  is not supported by the data, since the cross sign that represents the point  $(\psi = 1, \rho = 1)$ , lies outside de 0.05 level outer contour.

The results we have obtained are important from the biological point of view, since it was possible to differentiate between two ray species that biologically are very close; species that lived and were captured in the same area, that is, both shared the same source of fatty acids; hence, the detected differences between them can be due to the way each of these species metabolize food. It is important to emphasize that this kind of experimental study is very expensive, and sample sizes are usually small, but using the proposal presented here it was possible to detect differences between these ray species, using moderate sample sizes.

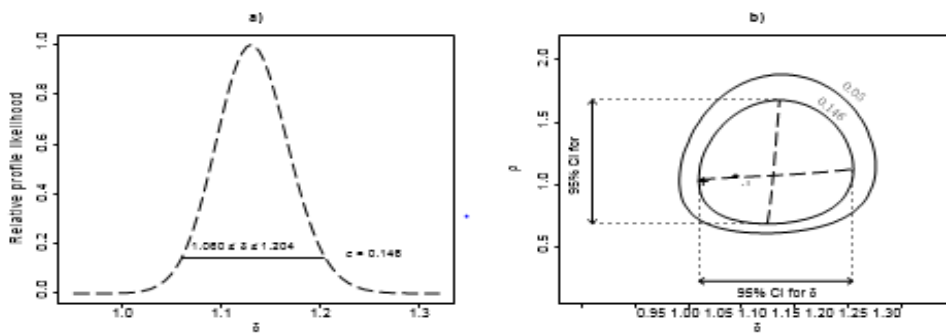


Figure 11: Ray species data: a) Relative profile likelihood of  $\delta$ , b) Relative profile likelihood contour plot for  $(\delta, \rho)$  and individual 95% CI for  $\delta$  and  $\rho$ .

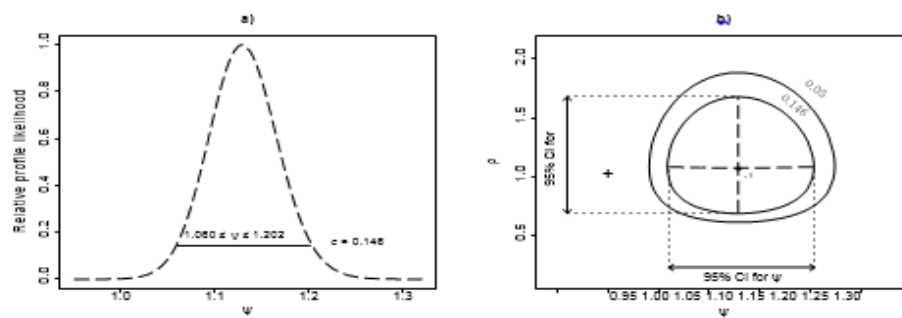


Figure 12: Ray species data: a) Relative profile likelihood of  $\psi$ , b) Relative profile likelihood contour plot for  $(\psi, \rho)$  and individual 95% CI for  $\psi$  and  $\rho$ .

## 5 Conclusion

The simple closed-form expression developed in Section 2 and all the features included in the likelihood contour plot, afford the comparison of two independent lognormal data sets in a more exhaustive way. As we have seen, elongated or triangular contour shapes are indicative of a trade-off between parameters  $\delta$  and  $\rho$ , so caution must be taken when summarizing individual inferences by using a single  $p$ -value or its corresponding confidence interval. Monte Carlo simulations showed that although coverage frequencies seems to be controlled in all scenarios, the power of the inverted profile likelihood test is affected by the likelihood contour shape. Now, even though contour plots for parameters  $(\psi, \rho)$  are not as elongated as the ones obtained for  $(\delta, \rho)$ , the selection of the parameter to be analyzed will be usually linked to the characteristics and objective of our problem.

When analyzing two data sets, it is important to have in mind a possible trade-off between the parameters under study, even if we just want a one-parameter comparison and are not interested in a distributions comparison. The likelihood approach we proposed here summarizes the sample information about the parameter vectors  $(\delta, \rho)$  or  $(\psi, \rho)$ , contained in the data, and provides more elements to support data conclusions.

## References

- Chen, H. (1994). Comparisons of Lognormal Population Means. *Proceedings of the American Mathematical Society*, 121, 915-924.
- Chen, Y-H. & Zhou, X-H. (2006). Interval estimates for the ratio and the difference of two lognormal means. *Statistics in Medicine* 25, 4099-4113.
- Díaz-Francis, E. & Sprott, D.A. (2004). Inference for the ratio of two normal means with unspecified variances. *Biometrical Journal*, 46, 83-89.
- Harris, W. S., Dris-Etherton, P.M. & Harris, K. A. (2008). Intakes of long-chain omega-3 fatty acid associated with reduced risk for death from coronary heart disease in healthy adults. *Current Atherosclerosis Reports*, 10, 503-509.
- Jiang, L. & Rekkas, M. & Wong, A. (2014). Comparing the means of two log-normal distributions: A likelihood approach. *Journal of Statistical and Econometric Methods*, 3:137152.
- Kalbfleisch, J. G. (1985). *Probability and statistical inference*, volume 2. Springer-Verlag, New York, second edition.



- Krishnamoorthy, K. & Mathew, T. (2003). Inferences on the means of log-normal distributions using generalized p-values and generalized confidence intervals. *Journal of Statistical Planning and Inference*, 115, 103-121.
- Lindsey, JK (2006). *Parametric Statistical Inference*. Oxford University Press, Oxford, UK.
- Navarro-García, G., González-Félix, M.L., Márquez-Farías, F., Bringas-Alvarado, L., Pérez-Velázquez, M., Montoya-Laos, J. & Moreno-Silva, B. (2014). Lipid Content and fatty acid composition of the liver from the rajiforms *Urotrygon chilensis*, *Urobatis halleri*, *Rhinobatos glaucostigma*, *Rhinoptera steindachneri* and *Dasyatis dipeteura* captured in Sinaloa, México. *International Food Research Journal*, 21, 223-229.
- Serfling, R. J. (2002). *Approximation Theorems of Mathematical Statistics*. John Wiley and Sons, New York.
- Sprott, D. A. (2000) *Statistical Inference in Science*. Springer Science & Bussines Media.
- Sprott, D.A. & Farewell, V.T. (1993). The difference between two normal means. *The American Statistician*, 47, 126-128.
- Wu, J., Jiang, G., Wong, A.C.M. & Sun, X. (2002). Likelihood analysis for the ratio of means of two independent log-normal distributions. *Biometrics*, 58, 463-469.
- Zhou, X.H., Gao, S. & Hui, S.L. (1997). Methods for comparing the means of two independent log-normal samples. *Biometrics*, 53, 1129-1135.

Received June 25, 2014; accepted September 28, 2014.

José A. Montoya  
Department of Mathematics  
University of Sonora, A. P. 1626  
montoya@mat.uson.mx  
Hermosillo, Sonora, México

Gudelia Figueroa  
Department of Mathematics  
University of Sonora, A. P. 1626  
gfiguero@mat.uson.mx  
Hermosillo, Sonora, México

Gerardo Navarro  
Department of Scientific and Technological Research  
University of Sonora, A. P. 1626

gnavarro@guayacan.uson.mx  
Hermosillo, Sonora, México