Stat Papers (2013) 54:499–522
DOI 10.1007/s00362-012-0446-1

REGULAR ARTICLE

The simplicity of likelihood based inferences for $P(X < Y)$ and for the ratio of means in the exponential model

Eloísa Díaz-Francés · José A. Montoya

Received: 6 June 2011 / Revised: 3 February 2012 / Published online: 17 April 2012

Abstract The profile likelihood of the reliability parameter $\theta = P(X < Y)$ or of the ratio of means, when X and Y are independent exponential random variables, has a simple analytical expression and is a powerful tool for making inferences. Inferences about θ can be given in terms of likelihood-confidence intervals with a simple algebraic structure even for small and unequal samples. The case of right censored data can also be handled in a simple way. This is in marked contrast with the complicated expressions that depend on cumbersome numerical calculations of multidimensional integrals required to obtain asymptotic confidence intervals that have been traditionally presented in scientific literature.

Keywords Comparison of exponential distributions · Exponential right censored data · Exponential stress–strength models · Profile likelihood of reliability parameter · ROC curves

Mathematics Subject Classification 62F99 · 62N01

1 Introduction

Whenever the comparison of the distributions of two exponential random variables X and Y is of interest, inferences about $\theta = P(X < Y)$ provide a way of summarizing this comparison in terms of a single parameter. The parameter θ is frequently called

E. Díaz-Francés (✉)
Department of Probability and Statistics, Centro de Investigación en Matemáticas,
A.P. 402, 36000 Guanajuato, GTO, Mexico
e-mail: diazfran@cimat.mx

J.A. Montoya
Universidad de Sonora, A.P. 1626, 83000 Hermosillo, SON, Mexico
e-mail: montoya@mat.uson.mx

the reliability or stress–strength parameter. The importance of θ arises in industrial contexts, since the reliability of a component can be described in terms of the stress experienced by the component, described by X, and the strength of the component available to overcome the stress, represented by Y. If the stress exceeds the strength, the component fails, and otherwise it resists. In such a setting, reliability is thus defined as the probability of not failing, which is therefore $\theta = P(X < Y)$ and is thus desired to be close to one. However, θ arises in many other areas of application aside from industry. The exponential distribution is widely used to model waiting, failure, or survival times in medicine, behavioral neuroscience, genetics, seismology, biology, and other areas.

There is also a strong connection, between θ and the receiving operating characteristic (ROC) curves that can be used to compare the distributions of X and Y, that have important applications in medicine (Sullivan Pepe 2000; Krzanowski and Hand 2009), particularly in the case of exponential distributions. For any continuous independent random variables $X,\ Y$, the area under the corresponding ROC curve is precisely equal to θ (Kotz et al. 2003, pp. 202–203; Krzanowski and Hand 2009, Sect. 2.4.1).

The traditional inferences about θ are described in Sect. 2 and usually require the estimation of the asymptotic mean squared error of the maximum likelihood estimate (mle) $\hat{\theta}$ that involves non trivial expressions. In the case of small samples, the mles are criticized for being biased. Consequently, unbiased estimation methods for θ have been preferred. These require obtaining $\tilde{\theta}$, the unique minimum variance unbiased estimate (UMVUE) of θ, and of the variance of $\tilde{\theta}$, in order to obtain confidence intervals for θ.

The estimation of θ in the case of uncensored exponential random variables has been widely discussed in scientific literature (Tong 1974, 1975; Kelley et al. 1976; Sathe and Shah 1981; Chao 1982; Cramer and Kamps 1997; Kotz et al. 2003, and from a Bayesian approach: Enis and Geisser 1971; Ventura and Racugno 2011). However, complicated procedures and expressions are given in these works. The asymptotic procedures used generally involve cumbersome numerical calculations of integrals (frequently multidimensional) over restricted domains that have very complicated analytical expressions, even in a simple case as the exponential.

When having censored observations, further complications arise. The case of censored data has been explored by only a few to our knowledge for the exponential case, such as Jiang and Wong (2008), Ismail et al. (1986), and Saracoglu et al. (2011). Jiang and Wong (2008) consider a Type 2 right censoring sampling scheme under the same exponential models described here. They propose estimation intervals for θ that have accurate confidence levels but involve as well very complicated expressions that have to be calculated numerically. Ismail et al. (1986) focus again on unbiased minimum variance point estimates of θ that involve complicated expressions. More recently, Saracoglu et al. (2011) explore the exponential case under a progressive Type 2 censoring. They compare the performance of three point estimates for θ: the uniformly minimum variance unbiased estimator, a Bayes estimate, and the mle. The first two estimates involve again complicated expressions and numerical methods. They recommend using the mle of θ and they present exact confidence intervals based on the same F pivotal quantity that was used by Jiang and Wong (2008), without acknowledging that the intervals can be quite close to the corresponding profile likelihood intervals (LIs) of θ. In Sect. 6 here, it will be shown that these confidence intervals have a similar form as the simple profile LIs of θ that will be proposed here for several types of right censoring.

In stark contrast to all statistical procedures that have been described so far, the profile likelihood of θ on its own has been largely ignored despite the fact that it is the immediate result of the probabilistic model and has a very simple analytical expression for the exponential case. The analysis of the whole profile likelihood function provides a powerful though simple way for estimating separately a parameter of interest in the presence of unknown nuisance parameters. Montoya et al. (2009) mentioned that the profile likelihood has been unduly criticized for giving strange or unintuitive results. These results arise because of having considered density functions that have singularities and because of having defined the likelihood function as proportional to the joint density function of the observed sample. If the likelihood is defined instead as described in Kalbfleisch (1985, Sect. 9.5), the alleged problems of the profile likelihood disappear. The particular case of the exponential distributions poses no such problems for the profile likelihood function of θ that is suggested here, because the corresponding density functions are bounded.

Therefore, the profile likelihood can be used to provide informative and accurate estimation intervals about the parameter of interest through the corresponding profile LIs. It will be shown here that inferences about θ, as well as for the ratio of exponential means φ, given in terms of profile LIs have a very simple algebraic form, thus enhancing their scientific interpretation. These intervals also have good coverage frequencies even for small and/or unequal sample sizes in the exponential case.

The possibility of extending the results presented here for the exponential case to more complex censoring situations and/or to other more general families of distributions is feasible under certain conditions that will also be discussed. For instance, the algebraic structure of the profile likelihood of θ in the exponential case coincidentally is the same as the one obtained for some other families of distributions. Therefore, the ideas presented here could also be applied to these other more general cases; however, that will be explored in future work.

Overall, the great algebraic simplicity of the proposed likelihood confidence (LC) intervals for θ is noteworthy in contrast to the traditional confidence intervals given in literature that require very complicated calculations.

2 Traditional inferences about θ in the exponential case

Consider a sample $x = (x_1, \ldots, x_n)$ of independent identically distributed (iid) exponential random variables with mean α and a second independent sample $y = (y_1, \ldots, y_m)$ of iid exponential random variables with mean β. Let $t = \sum_{i=1}^{n} x_i$, $s = \sum_{i=1}^{m} y_i$ and the sample means are $\bar{x} = t/n$ and $\bar{y} = s/m$.

In the exponential case as well as in more general settings, the usual inferential procedure of θ involves several steps, as mentioned in Kotz et al. (2003, p. 12), who offer an extensive review on the matter:

(a) Express θ as a function of the parameters in the model.
(b) Calculate the mles of the parameters in the model.

(c) Calculate $\hat{\theta}$, the mle of θ, from the mles obtained in (b) using the invariance property of the likelihood function under reparametrizations.
(d) Provide confidence intervals for θ, usually resorting to asymptotic methods.

Even in the simple case when both samples x and y are exponentially distributed, complicated expressions can arise when using asymptotic methods to calculate confidence intervals for θ, as mentioned in step (d). An example of this is the following UMVUE of θ that was presented in Tong (1974, 1975) and also in Kotz et al. (2003, p. 22),

$$\tilde{\theta} = \begin{cases} Q_1(n, m, n\bar{x}, m\bar{y}), & \text{if } m\bar{y} \le n\bar{x}, \\ Q_2(n, m, n\bar{x}, m\bar{y}), & \text{if } m\bar{y} > n\bar{x}, \end{cases} \tag{1}$$

where

$$Q_1(a, b, c, d) = \sum_{i=0}^{a-2} \frac{(-1)^i \Gamma(a) \Gamma(b)}{\Gamma(a-i-1) \Gamma(b+i+1)} \left(\frac{d}{c}\right)^{i+1},$$

and

$$Q_2(a, b, c, d) = \sum_{i=0}^{a-2} \frac{(-1)^i \Gamma(a) \Gamma(b)}{\Gamma(a+i) \Gamma(b-i)} \left(\frac{c}{d}\right)^{i}.$$

The corresponding UMVUE of the variance of $\tilde{\theta}$ is

$$\tilde{\sigma}^2 = \tilde{\theta}^2 - \frac{\left(n^2-3n+2\right)\left(m^2-3m+2\right)}{n^2 m^2 \bar{x}^{n-1} \bar{y}^{m-1}} H(n, m, \bar{x}, \bar{y}), \tag{2}$$

where

$$H(n, m, \bar{x}, \bar{y}) = \int_B \left(\bar{x} - \frac{b_1+b_2}{n}\right)^{n-3} \left(\bar{y} - \frac{b_3+b_4}{m}\right)^{m-3} db_1 db_2 db_3 db_4, \tag{3}$$

and $B = \{(b_1, b_2, b_3, b_4) : b_1 + b_2 < n\bar{x},\ b_3 + b_4 < m\bar{y},\ 0 < b_1 < b_3, 0 < b_2 < b_4\}$. Then, a $100(1-\gamma)\%$ confidence interval for θ is given by

$$\left(\tilde{\theta} - z_{1-\gamma/2}\tilde{\sigma},\ \tilde{\theta} + z_{1-\gamma/2}\tilde{\sigma}\right), \tag{4}$$

where $z_{1-\gamma/2}$ is the quantile of probability $1-\gamma/2$ of a standard normal distribution. Kotz et al. (2003, p. 38) warn their readers that for calculating (3) and the confidence interval (4) "even with the modern computer facilities, an application [of these formulae] may require substantial numerical effort". They then obtain, only for the case of equal sample sizes $n = m$, an asymptotic expression for the MSE($\hat{\theta}$), which is based on another asymptotic expression for the mean squared error of $\hat{\theta}$ given in Chao (1982),

$$\tilde{\sigma}_*^2 = \frac{2\hat{\varphi}^2}{n\left(1+\hat{\varphi}\right)^4} + \frac{4\hat{\varphi}^2\left(2\hat{\varphi}-1\right)\left(\hat{\varphi}-2\right)}{n^2\left(1+\hat{\varphi}\right)^6},$$

where $\hat{\varphi} = \bar{x}/\bar{y}$, the mle for the ratio of means φ. Therefore, a $100(1-\gamma)\%$ confidence interval for θ can finally be given as

$$\left(\tilde{\theta} - z_{1-\gamma/2}\tilde{\sigma}_*,\ \tilde{\theta} + z_{1-\gamma/2}\tilde{\sigma}_*\right).$$

Note that the sample sizes must be equal, $n = m$, in order to have feasible calculations.

The exponential distribution is one of the few cases when $P\,(X < Y)$ is of interest, where there are pivotal quantities associated with the parameter θ. A pivotal is a function of the sample and θ that has a distribution not depending on any unspecified parameters. Such a quantity was considered in a Bayesian approach by Enis and Geisser (1971) and presented in Kotz et al. (2003, Sect. 2.4.6). However, this pivotal cannot be used in a simple way to produce confidence intervals about θ, without unmanageable numerical calculations.

There is another much simpler F pivotal quantity associated with θ in the exponential case that was presented in Jiang and Wong (2008). It was used to obtain exact confidence intervals for θ in order to compare the corresponding coverage frequencies in simulations with those of their proposed improved confidence intervals that must be obtained numerically. This pivotal quantity was also used for making inferences about θ in the exponential case in Saracoglu et al. (2011) and it leads to as simple expressions as the ones presented in Sect. 5.

3 Inferences about θ using the profile likelihood in the exponential case

In great contrast to the foregoing complicated calculations of the previous section, the profile likelihood of θ provides a much simpler way of making inferences about θ, separately from the other unknown nuisance parameter. This likelihood function is a direct consequence of the probabilistic model and it has a simple expression as shown below for the exponential case.

Let X and Y be two independent exponential random variables with respective means α and β. The corresponding distribution functions are $F_X = 1 - \exp\left(-x/\alpha\right)$ and $F_Y = 1 - \exp\left(-y/\beta\right)$, respectively. The reliability parameter θ can be given as an explicit function of α and β,

$$\theta = P\,(X < Y;\ \alpha,\ \beta) = \frac{\beta}{\alpha+\beta} = \frac{1}{(\alpha/\beta)+1} = \frac{1}{\varphi+1},$$

where $\varphi = \alpha/\beta$ is the ratio of exponential means. Since θ is also a one to one function of the ratio of exponential means, inferences about θ necessarily correspond to inferences about φ.

For the case of exponential distributions, the following result can be proved in a straightforward manner. The reliability parameter $\theta = 1/2$ if and only if $F_X(z) = F_Y(z)$ for all $z \in \mathbb{R}$. Therefore inferences about θ are equivalent to statements about the comparison of the two underlying exponential distributions. When the equality of the distributions holds, $\theta = 1/2$ and $\varphi = 1$.

When X and Y are exponentially distributed as described above, the likelihood function of α and β can be written as

$$L(\alpha,\ \beta;\ x,\ y) \propto \alpha^{-n}\beta^{-m} \exp\left(-\frac{t}{\alpha} - \frac{s}{\beta}\right). \tag{5}$$

The mles of α and β are the sample means,

$$\hat{\alpha} = \frac{1}{n}\sum_{i=1}^{n} x_i = \frac{t}{n} = \bar{x} \quad \text{and} \quad \hat{\beta} = \frac{1}{m}\sum_{i=1}^{n} y_i = \frac{s}{m} = \bar{y}.$$

By the invariance property of likelihood, the corresponding mles of θ and φ are

$$\hat{\theta} = \frac{\hat{\beta}}{\hat{\alpha} + \hat{\beta}} \quad \text{and} \quad \hat{\varphi} = \frac{\hat{\alpha}}{\hat{\beta}}.$$

The likelihood function (5) reparametrized in terms of $(\theta,\ \beta)$ is

$$L(\theta,\ \beta;\ x,\ y) \propto \left(\frac{\theta}{1-\theta}\right)^{n} \beta^{-(n+m)} \exp\left\{-\frac{1}{\beta}\left[\frac{s + (t-s)\,\theta}{1-\theta}\right]\right\}.$$

The profile likelihood of θ can be obtained from this likelihood and has the following simple analytical expression,

$$\begin{aligned} L_p(\theta;\ x,\ y) &= \max_{\beta|\theta} L(\theta,\ \beta;\ x,\ y) = L[\theta,\ \hat{\beta}(\theta);\ x,\ y] \\ &\propto \frac{\theta^n (1-\theta)^m}{[s + (t-s)\,\theta]^{n+m}}, \end{aligned} \tag{6}$$

where $\hat{\beta}(\theta)$ is the restricted mle of β for a specified θ. However when samples are small and/or unequal, this profile likelihood can be very asymmetric. In such cases asymptotic confidence intervals of θ could be misleading and they would have poor coverage frequencies.

The log relative profile likelihood function of θ is

$$r_p(\theta;\ x,\ y) = \ln\left[\frac{L_p(\theta;\ x,\ y)}{L_p(\hat{\theta};\ x,\ y)}\right].$$

Its Taylor series expansion about the mle $\hat{\theta}$,

$$r_p(\theta;\ x,\ y) = -\frac{u_\theta^2}{2} - \frac{F_3(\hat{\theta})}{3!}u_\theta^3 + \frac{F_4(\hat{\theta})}{4!}u_\theta^4 - \cdots, \tag{7}$$

can be expressed as a power series of the quantity $u_\theta = (\hat{\theta} - \theta)\sqrt{I_{\hat{\theta}}}$, where $I_{\hat{\theta}}$ is the observed Fisher information of θ,

$$I_{\hat{\theta}} = -\left.\frac{\partial^2 r_p(\theta;\ x,\ y)}{\partial\theta^2}\right|_{\theta=\hat{\theta}}, \tag{8}$$

$$F_3(\hat{\theta}) = \left.\frac{\partial^3 r_p(\theta;\ x,\ y)}{\partial\theta^3}\right|_{\theta=\hat{\theta}} I_{\hat{\theta}}^{-3/2}, \tag{9}$$

and

$$F_4(\hat{\theta}) = \left.\frac{\partial^4 r_p(\theta;\ x,\ y)}{\partial\theta^4}\right|_{\theta=\hat{\theta}} I_{\hat{\theta}}^{-2}; \tag{10}$$

(see Sprott 2000, Sects. 9.1, 9.2, and 9.5 and Kalbfleisch 1985, Sect. 9.7). The quantity u_θ has the structure of a linear pivotal, one that is a linear function of θ, and is distributed asymptotically as a standard normal variable. However this result may hold as well for small samples whenever the shape of $r_p(\theta;\ x,\ y)$ is symmetric about $\hat{\theta}$ and parabolic. This happens when $F_3(\hat{\theta})$ is close to zero and $F_4(\hat{\theta})$ is small.

The quantities $F_3(\hat{\theta})$ and $F_4(\hat{\theta})$ depend only on the sample and are statistics that determine the asymmetry and the thickness of the tails of the profile likelihood function. Values of $F_3(\hat{\theta})$ close to zero are indicative of a symmetric profile likelihood function about $\hat{\theta}$. Small values of $F_4(\hat{\theta})$ indicate a peakedness similar to those of a normal profile likelihood (the one obtained from a sample of normal independent random variables). If $F_3(\hat{\theta})$ and $F_4(\hat{\theta})$ and the remaining terms of the Taylor expansion in (7) are negligible, then the relative likelihood function will be symmetric about $\hat{\theta}$ and parabolic. This is a case where the quadratic term in the Taylor expansion suffices to describe well $r_p(\theta;\ x,\ y)$. That is, the normal approximation denoted as

$$r_N(\theta;\ x,\ y) = -\frac{u_\theta^2}{2}, \tag{11}$$

will practically overlap the relative log profile likelihood $r_p(\theta;\ x,\ y)$ for all θ.

As stated in the Maximum Likelihood Theorem (Serfling 1980, p. 145), under some regularity conditions, as the sample size tends to infinity, only the quadratic term in (7) remains, since the rest of the terms converge to zero. However as mentioned, this asymptotic result might be reasonable for a finite, even small sample, when the log relative likelihood can be well approximated only with the quadratic term. For a fixed sample size, sometimes a suitable one to one parametrization $\delta = g(\theta)$ can be found that symmetrizes the likelihood function of θ in order for this approximation to hold. This parametrization is such that the quantity $F_3(\hat{\delta})$ is zero and $F_4(\hat{\delta})$ is small. Such a reparametrized likelihood can be denoted as $r_p^*(\delta;\ x,\ y) = r_p[g^{-1}(\delta);\ x,\ y]$ and will be symmetric about the mle $\hat{\delta}$ and parabolic.

In such a case, it can be assumed that the corresponding quantity $u_\delta = (\hat{\delta} - \delta)\sqrt{I_{\hat{\delta}}}$ already has reached its asymptotic standard normal distribution for the current sample, so that

$$\begin{aligned}(1-\gamma) &= P[-z \leq u_\delta \leq z] \\ &= P\left[\hat{\delta} - \frac{z}{\sqrt{I_{\hat{\delta}}}} \leq \delta \leq \hat{\delta} + \frac{z}{\sqrt{I_{\hat{\delta}}}}\right],\end{aligned}$$

where z is the quantile of probability $(1-\gamma/2)$ of a standard normal variable. Therefore, the LC interval of δ of likelihood level $\exp(-z^2/2)$ and of confidence level $(1-\gamma)$ has the simple form,

$$\delta = \hat{\delta} \pm \frac{z}{\sqrt{I_{\hat{\delta}}}}. \tag{12}$$

Inferences about θ can then be obtained by the invariance property of the likelihood function by merely transforming back to the scale of θ the endpoints of the interval given in (12) (see Kalbfleisch 1985, Chap. 11; Sprott 2000, Sects. 9.1 and 9.2).

In the case of the exponential model for X and Y considered here,

$$I_{\hat{\theta}} = \frac{nm}{(n+m)\left[\hat{\theta}(1-\hat{\theta})\right]^2}, \tag{13}$$

$$F_3(\hat{\theta}) = \frac{2\left[n + (n+m)\,(1-3\hat{\theta})\right]}{\sqrt{nm\,(n+m)}} \tag{14}$$

and

$$F_4(\hat{\theta}) = \frac{-6}{nm\,(n+m)}\left[c_1\hat{\theta}^2 - c_2\hat{\theta}(1-\hat{\theta}) + c_3(1-\hat{\theta})^2\right], \tag{15}$$

where $c_1 = 3m^2 + 3nm + n^2$, $c_2 = 2m^2 + 6nm + 2n^2$, and $c_3 = m + 3nm + 3n^2$. More details about how these expressions were calculated are given in the Appendix.

The procedure that was used here to select an adequate reparametrization $\delta = g(\theta)$ was simply to choose δ such that the third derivative in $F_3(\hat{\delta})$ was equal to zero, where $\hat{\delta}$ is the mle of δ; the corresponding value of $F_4(\hat{\delta})$ was also checked to be small. The proposal of the function g was made by trial and error. This procedure is described in more detail in the Appendix.

Let

$$k = \frac{(n-m)}{3(n+m)}. \tag{16}$$

Then, such a symmetrizing reparametrization is

$$\delta = \begin{cases} \log\left[\theta/\left(1-\theta\right)\right] = \log\left(\beta/\alpha\right), & \text{if} \quad n=m, \ (k=0), \\ \left[\theta/\left(1-\theta\right)\right]^k = \left(\beta/\alpha\right)^k, & \text{if} \quad n\neq m, \ (k\neq 0). \end{cases} \tag{17}$$

Note that this power transformation of the parameter $\varphi^{-1} = \beta/\alpha$ has a similar algebraic structure to that of the Box and Cox (1964) normalizing power transformations for random variables.

The mle of δ is thus obtained by replacing the mles $\hat{\theta}$ or $(\hat{\alpha},\ \hat{\beta})$ in (17). The reverse transformation is

$$\theta = \begin{cases} 1/\left(1+e^{-\delta}\right), & \text{if} \quad n=m, \ (k=0), \\ 1/\left(1+\delta^{-1/k}\right), & \text{if} \quad n\neq m, \ (k\neq 0). \end{cases}$$

The observed information of δ is

$$I_{\hat{\delta}} = \begin{cases} n/2, & \text{if} \quad n=m, \\ nm\left[(n+m)\,k^2\hat{\delta}^2\right]^{-1}, & \text{if} \quad n\neq m. \end{cases}$$

As mentioned, $F_3(\hat{\delta}) = 0$ and the values of $F_4(\hat{\delta})$ are

$$F_4(\hat{\delta}) = \begin{cases} 1/n, & \text{if} \quad n=m, \\ -(2/9)\left[n^{-1}+m^{-1}-13\,(n+m)^{-1}\right], & \text{if} \quad n\neq m. \end{cases}$$

Note how these values become smaller as sample sizes n and m increase.

The reparametrized likelihood $r_p^*(\delta;\ x,\ y)$ is symmetric about $\hat{\delta}$ and parabolic for all values of n and m; the corresponding normal approximation $r_N^*(\delta) = -u_{\hat{\delta}}^2/2$ is practically identical to $r_p^*(\delta;\ x,\ y)$ for all δ. Therefore LC intervals can be obtained for δ as in (12) and these can be transformed back to the scale of θ. These LC intervals for θ are thus

$$\left[\left(1+\frac{\hat{\alpha}}{\hat{\beta}}\frac{1}{\exp\left(-z\sqrt{2/n}\right)}\right)^{-1}, \left(1+\frac{\hat{\alpha}}{\hat{\beta}}\frac{1}{\exp\left(z\sqrt{2/n}\right)}\right)^{-1}\right], \quad \text{if} \quad n=m, \tag{18}$$

and if $n \neq m$,

$$\left[\left(1+\frac{\hat{\alpha}}{\hat{\beta}}\frac{1}{\left[1+zk\sqrt{\frac{n+m}{nm}}\right]^{1/k}}\right)^{-1}, \left(1+\frac{\hat{\alpha}}{\hat{\beta}}\frac{1}{\left[1-zk\sqrt{\frac{n+m}{nm}}\right]^{1/k}}\right)^{-1}\right], \quad \text{if} \quad k>0, \tag{19}$$

or

$$\left[\left(1+\frac{\hat{\alpha}}{\hat{\beta}}\frac{1}{\left[1-zk\sqrt{\frac{n+m}{nm}}\right]^{1/k}}\right)^{-1},\ \left(1+\frac{\hat{\alpha}}{\hat{\beta}}\frac{1}{\left[1+zk\sqrt{\frac{n+m}{nm}}\right]^{1/k}}\right)^{-1}\right],\quad \text{if}\quad k<0. \tag{20}$$

Note that these expressions are simple and have not been given before in scientific literature up to our knowledge. The likelihood level $\exp(-z^2/2)$ of these intervals is exact and the confidence level $(1-\gamma)$ is approximate, though quite accurate. These intervals will be compared with the exact confidence intervals of θ that can be obtained from a pivotal quantity having an F distribution that will be described in Sect. 5.

4 Inferences about the ratio of exponential means

An alternative way of comparing two exponential distributions is through the ratio of their means $\varphi = \alpha/\beta$. By the invariance of the likelihood function, since $\theta = (\varphi+1)^{-1}$, the profile likelihood of φ can be obtained from (6) and is

$$L_p(\varphi;\ x,\ y) = \varphi^m\,[s\varphi + t]^{-(n+m)}.$$

Since φ is a one to one function of θ or of δ, by making likelihood inferences about θ and δ in the way suggested in the previous section, inferences can be made thus as well for φ in terms of LC intervals,

$$\varphi = \hat{\varphi}\exp\left(\pm z\sqrt{\frac{2}{n}}\right),\quad \text{if}\quad n=m, \tag{21}$$

and

$$\varphi = \hat{\varphi}\left(1 \pm zk\sqrt{\frac{n+m}{nm}}\right)^{-1/k},\quad \text{if}\quad n\neq m, \tag{22}$$

where z is the normal quantile of probability $1-\gamma/2$. The likelihood level of these intervals is $\exp(-z^2/2)$ and the approximate confidence level is $(1-\gamma)$. Note that the intervals (18)–(20) can be written in a more compact way as the intervals (21) and (22); however for the sake of the comparison with the intervals presented in the following section, they were written in a more expanded form.

5 Confidence intervals of θ and φ obtained from a pivotal

The following two are pivotal quantities that involve the exponential samples and the parameters α and β, respectively. These pivotals follow χ^2 distributions with $2n$ and $2m$ degrees of freedom, respectively,

$$G(\alpha,\ x) = 2n\frac{\hat{\alpha}}{\alpha} \sim \chi^2_{(2n)}, \quad \text{and} \quad G(\beta,\ y) = 2m\frac{\hat{\beta}}{\beta} \sim \chi^2_{(2m)}. \tag{23}$$

The ratio of these pivotals divided by their respective degrees of freedom is also a pivotal quantity that has an F distribution with $2n$ and $2m$ degrees of freedom,

$$\frac{G(\alpha,\ x)/2n}{G(\beta,\ y)/2m} = \frac{\hat{\alpha}\beta}{\hat{\beta}\alpha} = \frac{\left(\hat{\theta}^{-1} - 1\right)}{\left(\theta^{-1} - 1\right)} = \frac{\hat{\varphi}}{\varphi} = \frac{\delta^{1/k}}{\hat{\delta}^{1/k}} \sim F_{(2n,\,2m)}.$$

This F pivotal quantity is a one to one function, though non linear, of θ, δ or φ. Therefore it can be used to obtain a confidence interval of probability $(1-\gamma)$ for θ for all sample sizes. For instance, for a $(1-\gamma) \times 100\,\%$ confidence level,

$$\begin{aligned}
(1-\gamma) &= P[F_{\gamma/2} \le \frac{\delta^{1/k}}{\hat{\delta}^{1/k}} \le F_{1-\gamma/2}] \\
&= P\left[\hat{\delta}\left(F_{1-\gamma/2}\right)^k \le \delta \le \hat{\delta}\left(F_{\gamma/2}\right)^k\right] \qquad (24) \\
&= P\left[\hat{\varphi}/F_{1-\gamma/2} \le \varphi \le \hat{\varphi}/F_{\gamma/2}\right] \\
&= P\left[\left(1 + \frac{\hat{\alpha}}{\hat{\beta}}\frac{1}{F_{\gamma/2}}\right)^{-1} \le \theta \le \left(1 + \frac{\hat{\alpha}}{\hat{\beta}}\frac{1}{F_{1-\gamma/2}}\right)^{-1}\right], \qquad (25)
\end{aligned}$$

where $F_{\gamma/2}$ and $F_{1-\gamma/2}$ are the quantiles of probabilities $\gamma/2$ and $(1-\gamma/2)$ of the $F_{(2n,2m)}$ distribution.

Therefore an exact confidence interval for θ of probability $(1-\gamma)$ for all values of n and m is

$$\left[\left(1 + \frac{\hat{\alpha}}{\hat{\beta}}\frac{1}{F_{\gamma/2}}\right)^{-1},\ \left(1 + \frac{\hat{\alpha}}{\hat{\beta}}\frac{1}{F_{1-\gamma/2}}\right)^{-1}\right].$$

The corresponding confidence interval for φ is

$$\left[\frac{\hat{\varphi}}{F_{1-\gamma/2}},\ \frac{\hat{\varphi}}{F_{\gamma/2}}\right].$$

These intervals turn out to be practically the same as the LC interval obtained in (18)–(22) even in cases where n and m are small. This would give evidence in favor of the assumption made in Sect. 3 that the asymptotic standard normal distribution of the pivotal quantity u_δ holds for the given small sample sizes in consideration. The closeness of these intervals will be shown in the examples of Sect. 7.

The similarity of the algebraic structure of the intervals for θ obtained using the proposed method of Sect. 3 with these exact confidence intervals for θ suggests exploring the closeness of the distributions of the following random variables:

(1) The distribution of the random variable $W_1 = \exp\left(-Z\sqrt{2/n}\right)$ that is part of the expression of the interval (18), where Z is a standard normal variable can

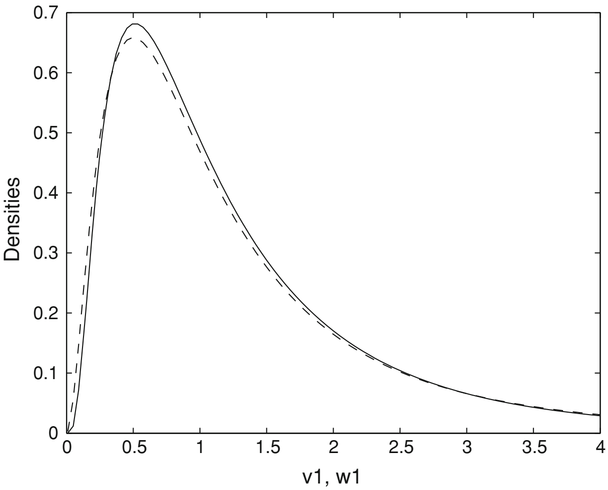

Fig. 1 Densities of W_1 (*solid*) and V_1 (*dashes*) when $n = m = 3$

be compared to the distribution of the random variable V_1 with an $F_{(2n,2m)}$ distribution that was considered in expression (25) for the case of $n = m$.

(2) The distribution of the random variable $W_2 = \left(1 + Zk\sqrt{(n+m)/nm}\right)$ that forms part of the expression of the interval (19) and of interval (20), where Z is a standard normal variable can compared to the distribution of V_2, the k-th power of the random variable having an $F_{(2n,2m)}$ distribution considered in expression (24) for the case of $n \neq m$.

The largest discrepancies between these densities are observed when n and m are small as shown in Figs. 1 and 2 below for some examples with small samples. Note that by the change of variable theorem, the distribution of W_1 is a lognormal random variable with density function

$$g\left(w_1\right) = \frac{1}{w_1\sqrt{4\pi/n}} \exp\left[-\frac{n\left(\ln w_1\right)^2}{4}\right].$$

For moderate values of n and m, this density is practically the same as the corresponding density of the random variable V_1 that is distributed as an $F_{2n,2m}$, whose density is given by

$$f\left(v_1\right) = \frac{\Gamma\left(2n\right)}{\Gamma\left(n\right)^2} \frac{v_1^{n-1}}{\left(1+v_1\right)^{2n}}.$$

Figure 1 shows the overlapping plots of these densities for the case of $n = m = 3$; note that $g(w_1)$ and $f(v_1)$ are practically indistinguishable, particularly in the tails.

The same can be observed in the case of $n \neq m$ for the random variable W_2 that has a normal distribution with unitary mean and variance τ^2, where $\tau = -k\sqrt{(n+m)/nm}$ whose density function is

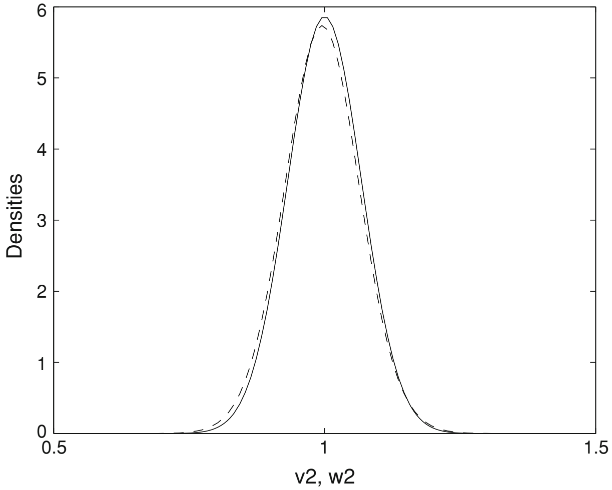

Fig. 2 Densities of W_2 (*solid*) and of V_2 (*dashes*) for $n = 8$ and $m = 4$

$$h\left(w_2\right) = \frac{1}{\sqrt{2\pi}\,|\tau|} \exp\left[-\frac{(w_2 - 1)^2}{2\tau^2}\right].$$

For moderate values of n and m, this density practically overlaps the density function of the random variable V_2, the k-th power of an $F_{(2n,2m)}$ random variable whose density is

$$f\left(v_2\right) = \frac{\Gamma\left(n+m\right)\left(n/m\right)^n}{\Gamma\left(n\right)\Gamma\left(m\right)k} \frac{v_2^{(n-k)/k}}{\left(1+\frac{n}{m}v_2^{1/k}\right)^{n+m}}.$$

Figure 2 shows for $n = 8$ and $m = 4$, how the corresponding densities $h(w_2)$ and $f(v_2)$ are very close. Consequently the estimation intervals for θ, δ, or φ will practically be the same, supporting the fact that the LIs given in (18)–(22) have a very accurate confidence level even for small sample sizes.

6 Exponential right censored data

As mentioned in the Introduction, the case of right censored data for the exponential case has been explored by only a few to our knowledge such as Jiang and Wong (2008), Saracoglu et al. (2011), and Ismail et al. (1986). In the first two of these works, their proposed estimation intervals have accurate confidence levels but involve complicated expressions and numerical calculations. The exception to this is the F confidence intervals for θ given also given by them (Jiang and Wong 2008, p. 642; Saracoglu et al. 2011, formula 26). However, the authors do not compare these intervals with the corresponding profile LIs of θ as we do here. Ismail et al. (1986) considered the gamma distributions (which includes the exponential as a particular case) with a right censoring scheme, but their interest was focused on point estimates of θ and they presented complicated expressions.

In contrast, the profile likelihood of θ has again a simple expression in the case of several schemes of right censoring, as will be described below. Therefore, inferences about θ or φ can be easily given too in terms of LC intervals with a simple algebraic structure.

The following types of right censoring will be considered here (as described in Lawless 2003, Sect. 2.2.1) for a sample of size n of (iid) random variables, where $x_{(1)}, \ldots, x_{(n_1)}$ are the n_1 smallest observations. The three types are: (a) Type 1 with a common right detection limit T_1, where $x_{(1)}, \ldots, x_{(n_1)}$ are observed and $(n - n_1)$ are known to fail after time T_1, (b) Type 2, where only the n_1 smallest data are observed, $x_{(1)}, \ldots, x_{(n_1)}$, and $(n - n_1)$ are known to fail after time $x_{(n_1)}$, and (c) progressive Type 2 censoring that involves a more complex sampling scheme that is described in Lawless (2003, p. 56) and Saracoglu et al. (2011).

As described in Lawless (2003, Sect. 2.2), all these three sampling schemes lead to a likelihood function that has the same algebraic structure and this is given for a general distribution in his formula (Sect. 2.2.3, p.53). In order to simplify the presentation here, only the first two types of censoring described in (a) and (b) above will be considered here for the exponential case. Nevertheless, the probability model considered by Saracoglu et al. (2011) for progressive right censoring in (c) also leads to a likelihood function with the same algebraic structure as the one presented here below. Therefore our results are also applicable to this type of censoring.

Assume that $x_{(1)}, \ldots, x_{(n_1)}$ are the first n_1 ordered observations of a sample of size n of iid exponential random variables with mean α. The $(n - n_1)$ remaining observations are known to be larger than $x_{(n_1)}$ or than a fixed predetermined value T_1, according to the type of censoring. Let $y_{(1)}, \ldots, y_{(m_1)}$ be the first m_1 ordered observations of a second independent sample of size m of iid exponential random variables with mean β. The $(m - m_1)$ right censored observations, are again known to be either larger than a common right detection limit time T_2, or than observation $y_{(m_1)}$, according to the type of censoring. For simplicity in the following expressions assume that $T_1 = x_{(n_1)}$ and $T_2 = y_{(m_1)}$. Then, let $t' = \sum_{i=1}^{n_1} x_{(i)} + (n - n_1)T_1$ and $s' = \sum_{i=1}^{m_1} y_{(i)} + (m - m_1)T_2$.

The likelihood function of (α, β) for any of this two censoring schemes is thus

$$L(\alpha, \beta; x, y) \propto \alpha^{-n_1} \beta^{-m_1} \exp\left(-\frac{t'}{\alpha} - \frac{s'}{\beta}\right),$$

after formula (Sect. 2.2.3, Lawless 2003) applied to the corresponding exponential distributions. Reparametrizing in terms of θ and β, the likelihood function is

$$L(\theta, \beta; x, y) \propto \left(\frac{\theta}{1-\theta}\right)^{n_1} \beta^{-(n_1+m_1)} \exp\left\{-\frac{1}{\beta}\left[\frac{s' + \left(t' - s'\right)\theta}{1-\theta}\right]\right\}.$$

The profile likelihood function of θ for right censored data is thus

$$L_p(\theta;\ x,\ y) \propto \frac{\theta^{n_1}(1-\theta)^{m_1}}{[s' + (t' - s')\,\theta]^{n_1+m_1}}. \tag{26}$$

Notice that this function is almost the same as (6) , except that the values of $n_1,\ m_1$ have replaced those of n, m. In this case, the mles are $\hat{\alpha} = t'/n_1$, $\hat{\beta} = s'/m_1$, and $\hat{\theta} = \hat{\beta}/(\hat{\alpha} + \hat{\beta})$. Furthermore, $F_3(\hat{\theta})$ and $F_4(\hat{\theta})$ are as in (14) and (15) and the LC intervals for θ are as in (18)–(20) except that again, n is replaced by n_1 and m by m_1.

The pivotals shown in (23) follow approximately a χ^2 distribution even if the samples are small and the proportion of censoring is large. Consequently, the F pivotal quantity described in Sect. 6, has an approximate F distribution as well. Therefore confidence intervals used from these pivotal can be used as suggested in Sect. 5.

The distribution of the quantity $-2r_p(\theta_0;\ x,\ y)$, where θ_0 is the true value of the parameter, has an asymptotic χ^2 distribution with one degree of freedom, as discussed in Lawless (2003, Sect. 2.2.3). This asymptotic result usually holds as well even for small sample sizes in the exponential case and when censoring proportions are large as can be explored in simulations. This last point is also discussed in Sprott (2000, Example 2.9.b and Sect. 5.3).

The quantity $-2r_p(\theta_0;\ x,\ y)$ is directly linked and determines the coverage frequencies of LIs of θ (Kalbfleisch 1985, Chap. 11). Therefore, LIs obtained from (26) will usually have adequate coverage frequencies for moderate and small samples.

For making inferences about the ratio of means φ when having exponential right censored observations, the corresponding LC intervals will have the same expressions as in Sect. 4 except that the mles of $\alpha,\ \beta$ must be calculated as described there and the values of $n_1,\ m_1$ will replace those of n and m, respectively, in all of the corresponding formulae.

7 Examples

7.1 Small sample examples

The case of small samples arises frequently in reliability settings, where the values of θ are usually close to one for a very reliable component. In these situations, $F_3(\hat{\theta})$ and $F_4(\hat{\theta})$ are usually large in absolute value and thus the corresponding profile likelihood of θ can be very asymmetric. Table 1 illustrates this situation with two constructed examples with values of $\bar{x}$ and $\bar{y}$ such that $\hat{\theta} = 0.99$.

Table 1 also shows the Taylor expansion coefficients both for θ and for the proposed reparametrization δ that symmetrizes the likelihood function for both examples.

Table 1 Possible small sample sized examples with asymmetric profile likelihoods of θ

Examples	n	m	$\bar{x}$	$\bar{y}$	$\hat{\theta}$	$F_3(\hat{\theta})$	$F_4(\hat{\theta})$	$F_3(\hat{\delta})$	$F_4(\hat{\delta})$
(a) $n = m$	8	8	10	990	0.99	−1.47	−2.54	0	0.12
(b) $n \neq m$	8	4	10	990	0.99	−1.60	−3.13	0	0.16

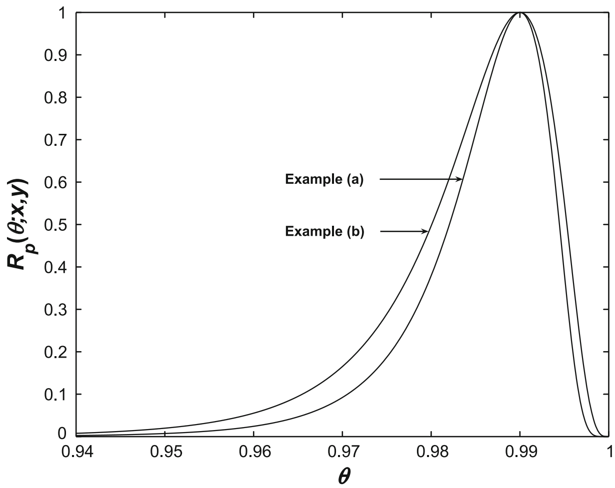

Fig. 3 Relative profile likelihoods of θ for Examples (a) and (b).

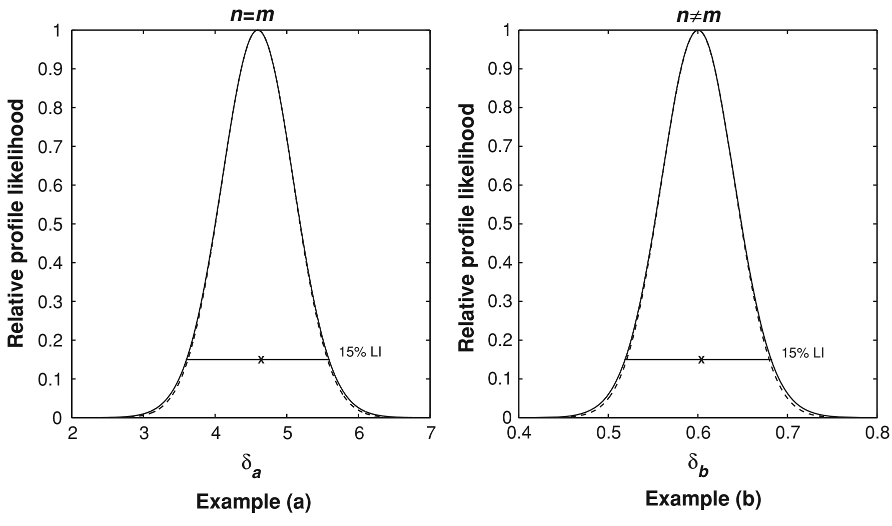

Fig. 4 Relative profile likelihoods of δ_a and δ_b (*solid*) and its normal approximation (*dashes*) corresponding to Examples (a) and (b); *LI* likelihood interval

Figure 3 shows $R_p(\theta;\ x,\ y)$, the relative profile likelihood of θ, for both examples which are clearly very asymmetric. Note that the relative profile likelihood of θ for Example (b), where sample sizes are different, is more asymmetric than the corresponding one of Example (a).

Figure 4 shows the relative profile likelihood of δ for both examples given in Table 1. These curves are bell shaped and symmetric about their mode, the mle. For both examples, the normal approximations

Table 2 LIs of 15 % level from $R_p(\theta;\ x,\ y)$; corresponding LC intervals using the normal approximation $R_N(\delta;\ x,\ y)$ and pivotal intervals (PIs) for θ and φ for Examples (a) and (b)

Example	(a)	(b)
LI θ	(0.973, 0.996)	(0.969, 0.997)
LC θ	(0.974, 0.996)	(0.970, 0.997)
PI θ	(0.973, 0.996)	(0.969, 0.997)
LI φ	(0.004, 0.027)	(0.003, 0.032)
LC φ	(0.004, 0.027)	(0.003, 0.031)
PI φ	(0.004, 0.028)	(0.002, 0.032)

$$R_N^*(\delta;\ x,\ y) = \exp\left[-\frac{(\delta-\hat{\delta})^2 I_{\hat{\delta}}}{2}\right]$$

[as in (11) but in terms of δ] practically overlap the corresponding profile likelihoods of δ, as shown in this figure.

The 15 % likelihood and 95 % confidence intervals for θ and φ, both obtained from the normal approximation of δ, using the formulae of Sects. 3 and 4 are shown in Table 2. The corresponding intervals obtained with the F pivotal quantity in Table 2 are shown as well and note that they are all quite close.

7.2 Waiting times in a neurobiology experiment

In order to evaluate the effect of a certain drug on memory and learning processes in mammals, an experiment that involved inhibitory avoidance tasks using Wistar rats was conducted at the Neurobiology Institute of UNAM (National University of Mexico, Campus Querétaro). As part of this experiment, rats were randomly assigned to a treatment and a control group of the same size each. The rats on the treatment group were injected with the drug of interest that is claimed to induce memory loss. The experiment is described in detail in Díaz-Trujillo et al. (2009).

For the sake of assessing the effect of the drug, the comparison of the times for doing a given task between the control and the treatment groups of rats becomes relevant. Traditionally non parametric tests for comparing the medians are used in Neurobiology in these kind of experiments (Díaz-Trujillo et al. 2009).

Two data sets corresponding to the waiting times in seconds of the treatment and control groups of rats that were subject to an electric shock of 8 mA will be considered here. These data were discussed in Díaz-Trujillo et al. (2009). All of these data sets can be adequately modelled with Exponential distributions. The first data set consists of the waiting times of 12 rats in the control group and of other 12 rats in the treatment group. All were uncensored and will be denoted as Escape times. The second data set consists of the waiting times, denoted Retention times, of the same rats in both treatment and control groups. Of these, seven waiting times in the control group were right censored at 600 s. That is, the control group of the Retention Times was the only one that presented right censored observations. Some descriptive statistics of the two data sets are given in Table 3.

Table 3 Descriptive statistics for waiting times of four groups of 12 rats

Times	Group	Minimum	Maximum	Mle of mean	Empirical median
Escape	Control	0.1	3.9	$\hat{\alpha} = 1.7$	2.0
Escape	Treatment	0.1	5.5	$\hat{\beta} = 2.1$	1.8
Retention	Control	24.2	600*	$\hat{\alpha} = 974.9$	600*
Retention	Treatment	6.1	389.2	$\hat{\beta} = 115.5$	81.6

The common right detection limit was 600 s (indicated with an asterisk)

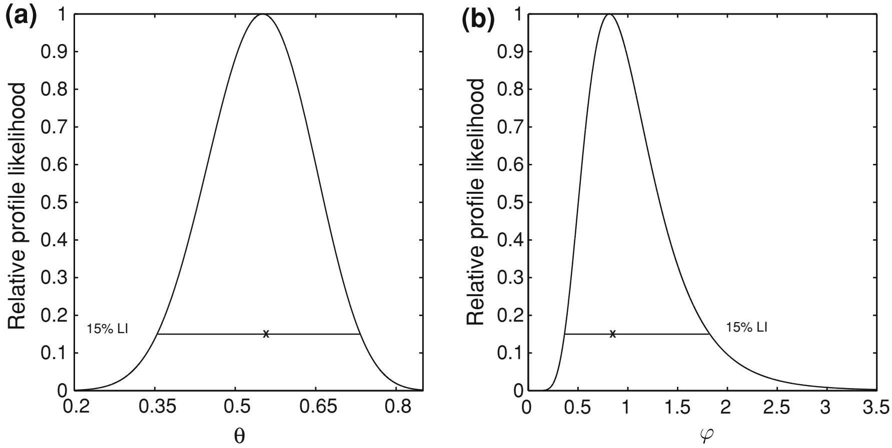

Fig. 5 Escape times; relative profile likelihoods of θ and φ with 15 % LIs and mles (*)

Table 4 Escape and retention times

LIs of 15 % level from $R_p(\theta;\, x,\, y)$ and $R_p(\varphi;\, x,\, y)$; corresponding LC intervals using the normal approximation $R_N(\delta;\, x,\, y)$; and PIs for θ and φ

Times	Escape	Retention
$\hat{\theta}$	0.55	0.106
LI θ	(0.354, 0.733)	(0.037, 0.241)
LC θ	(0.355, 0.732)	(0.037, 0.239)
PI θ	(0.351, 0.736)	(0.034, 0.238)
$\hat{\varphi}$	0.81	8.44
LI φ	(0.364, 1.822)	(3.151, 26.248)
LC φ	(0.366, 1.813)	(3.184, 26.026)
PI φ	(0.359, 1.848)	(3.198, 28.408)

The estimation intervals proposed in Sect. 3 for θ and φ are helpful to compare the exponential distributions of the treatment and control groups. Figure 5a and b show the relative profile likelihood of θ and φ, respectively, for the Escape times. The 15 % LIs of θ and φ are given in Table 4; the mles are marked with asterisks over these intervals.

For the escape times, the values of $\theta = 0.5$ and $\varphi = 1$ have a large relative likelihood and are included in all the intervals, thus supporting the equality of the two exponential distributions. Note that the LIs and the confidence intervals for θ are

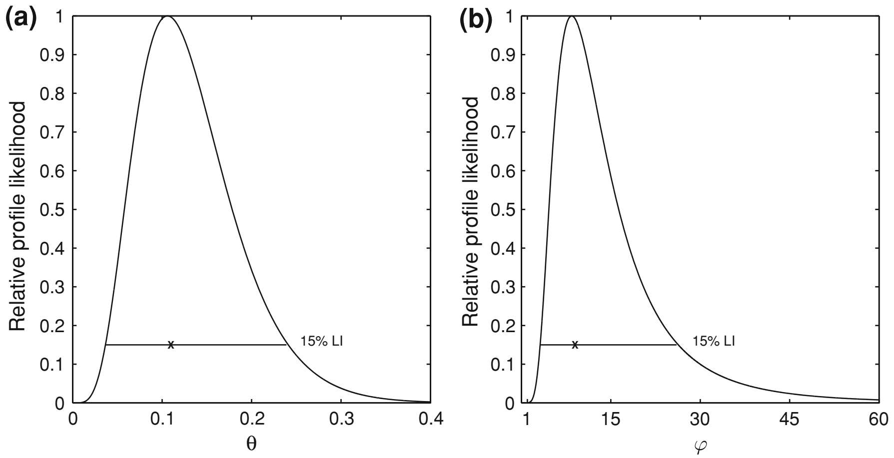

Fig. 6 Retention times; relative profile likelihoods of θ and φ with 15 % LIs and mles (*)

almost the same; the same happens for φ. The corresponding PIs are also similar but slightly wider; nevertheless the values of $\theta = 0.5$ and $\varphi = 1$ are also included.

For the Retention times, Fig. 6a and b show the corresponding relative profile likelihoods of θ and φ. The 15 % likelihood intervals (LI θ and LI φ of Table 4) are marked on the curves, and the mles with an asterisk. Again, the LIs and LC intervals are similar; the PI is again a little wider. Nevertheless, for the Retention times, the values of $\theta = 0.5$ and $\varphi = 1$ are excluded from all of the intervals and these values have negligible relative likelihoods. This means that the data do not support them. The data provide strong evidence in favor that the exponential mean of the treatment group is at least three times smaller than the one of the control group. This is in agreement with the non parametric test for the medians presented in Díaz-Trujillo et al. (2009) where they concluded that amnesia was produced in the treatment group. The estimation intervals presented here for θ and φ provide stronger evidence supporting their conclusions. The main difference with their analysis is that a suitable exponential model was used here for their waiting times and that explicit estimation intervals (LC intervals) can be given for the ratio of means and for θ.

To our knowledge, parametric models that consider right censored observations have not been considered before in the analyses of this kind of data in neurobiology.

8 Conclusions and possible extensions

Inferences about θ or φ in the exponential case can be given in terms of profile LIs that have simple algebraic expressions and adequate coverage frequencies. The methods proposed here are particularly useful when having small and/or unequal sample sizes, as well as when having right censored observations.

The simple algebraic expressions of the LC intervals of θ given in (18)–(22) are noteworthy when compared to the complicated expressions presented from (1) to (4) that require heavy or non trivial numerical methods. The latter confidence intervals are always symmetric about $\hat{\theta}$ and can be unrealistic or may have poor coverage frequencies if the likelihood surface is asymmetric.

The exact confidence intervals derived from the F pivotal quantity of Sect. 5 also have simple algebraic expressions and are practically the same as the profile LIs proposed here for non censored observations, as was shown in the examples. This supports that the proposed LIs have good coverage properties even for small sample sizes for uncensored data. For right censored data, the proposed profile LIs and PIs of θ also have reasonable coverage frequencies even for small sample sizes.

All the results mentioned in Sect. 3 for the profile LIs of θ can be extended as well to other more general families of distributions, whenever the algebraic structure of the profile likelihood of θ is the same as the one presented here. This occurs for instance for the generalized exponential distribution setting considered in Kundu and Gupta (2005), when the common scale parameter is known, but this will be presented in more detail in future work.

The analysis of the whole profile likelihood functions of θ (or φ), on their own could also be explored for other distributions. An important requirement is that there exists a one to one reparametrization from the original parameters of the model to another set of parameters that includes θ. Then the profile likelihood of θ and profile LIs can be obtained, though perhaps numerically. Reasonable coverage frequencies for these intervals for sufficiently large samples are guaranteed by asymptotic maximum likelihood results.

Summarizing: estimating the reliability parameter even in the simple case of exponential distributions raises great difficulty with standard procedures. This is in stark contrast with the simple expressions of the profile likelihood function and estimation intervals of θ that were proposed here.

Acknowledgements The authors thank Drs. D. A. Sprott and V. T. Farewell for fruitful discussions. The authors thank as well Dr. Prado-Alcalá and Sofía González-Salinas at the Institute of Neurobiology at UNAM for providing the data of the Neurobiology experiment that was analyzed in Sect. 7. Finally, the authors would like to thank an anonymous referee for incisive comments that lead them to consider interesting and important extensions of the ideas presented in an early version of the manuscript.

Appendix

A brief outline of how to obtain the expressions of $I_{\hat{\theta}}$, $F_3(\hat{\theta})$, and $F_4(\hat{\theta})$ in Sect. 3 is given here. It is also described how the symmetrizing reparametrizations for the profile likelihood of θ, in terms of δ, were found.

Notice that the first, second, third, and fourth derivatives of the log profile likelihood of θ, $l_p\left(\theta;\ x,\ y\right) = \ln\left[L_p\left(\theta;\ x,\ y\right)\right]$, can be written as,

$$\frac{dl_p\left(\theta;\ x,\ y\right)}{d\theta} = \frac{n}{\theta} - \frac{m}{1-\theta} - \frac{(n+m)\,(t-s)}{s+(t-s)\,\theta}$$

$$\frac{d^2l_p\left(\theta;\ x,\ y\right)}{d\theta^2} = -\frac{n}{\theta^2} - \frac{m}{(1-\theta)^2} + \frac{(n+m)\,(t-s)^2}{[s+(t-s)\,\theta]^2}$$

$$\frac{d^3 l_p(\theta;\ x,\ y)}{d\theta^3} = \frac{2n}{\theta^3} - \frac{2m}{(1-\theta)^3} - \frac{2(n+m)(t-s)^3}{[s+(t-s)\theta]^3}$$
$$\frac{d^4 l_p(\theta;\ x,\ y)}{d\theta^4} = -\frac{6n}{\theta^4} - \frac{6m}{(1-\theta)^4} + \frac{6(n+m)(t-s)^4}{[s+(t-s)\theta]^4}.$$

Since the mle of θ is $\hat{\theta} = \bar{y}/(\bar{x}+\bar{y})$, where $\bar{x} = t/n$ and $\bar{y} = s/m$, the first derivative evaluated at the mle $\hat{\theta}$ takes a value of zero. The second derivative evaluated at $\hat{\theta}$ is

$$\left.\frac{d^2 l_p(\theta;\ x,\ y)}{d\theta^2}\right|_{\theta=\hat{\theta}} = -\frac{n}{\hat{\theta}^2} - \frac{m}{\left(1-\hat{\theta}\right)^2} + \frac{(n+m)(t-s)^2}{\left[s+(t-s)\hat{\theta}\right]^2}$$
$$= -\frac{n}{\hat{\theta}^2} - \frac{m}{\left(1-\hat{\theta}\right)^2} + \frac{(n+m)(n\bar{x}-m\bar{y})^2}{\left[m\bar{y}+(n\bar{x}-m\bar{y})\hat{\theta}\right]^2}$$
$$= -\left(\frac{nm}{n+m}\right)\frac{(\bar{x}+\bar{y})^4}{(\bar{x}\bar{y})^2} = -\frac{nm}{(n+m)\left[\hat{\theta}\left(1-\hat{\theta}\right)\right]^2}$$
$$< 0.$$

Therefore, the observed Fisher information of θ is the negative of this value, after formula (8), which is the expression given in (13).

In order to calculate $F_3(\hat{\theta})$ and $F_4(\hat{\theta})$, the third and fourth derivatives of the log profile likelihood of θ evaluated at the mle are required. These take the following form,

$$\left.\frac{d^3 l_p(\theta;\ x,\ y)}{d\theta^3}\right|_{\theta=\hat{\theta}} = \frac{2nm(\bar{x}+\bar{y})^5}{(n+m)^2(\bar{x}\bar{y})^3} B,$$
$$\left.\frac{d^4 l_p(\theta;\ x,\ y)}{d\theta^4}\right|_{\theta=\hat{\theta}} = -\frac{6nm(\bar{x}+\bar{y})^6}{(n+m)^3(\bar{x}\bar{y})^4} C,$$

where $B = (m+2n)\bar{x} - (2m+n)\bar{y}$, $C = c_1\bar{y}^2 - c_2\bar{x}\bar{y} + c_3\bar{x}^2$, and where $c_1 = 3m^2 + 3nm + n^2$, $c_2 = 2m^2 + 6nm + 2n^2$, $c_3 = m + 3nm + 3n^2$. Therefore, substituting the expression given above for the third derivative in formula (9) the algebraic expression for $F_3(\hat{\theta})$ is obtained and simplified as,

$$F_3(\hat{\theta}) = \left.\frac{d^3 l_p(\theta;\ x,\ y)}{d\theta^3}\right|_{\theta=\hat{\theta}} I_{\hat{\theta}}^{-3/2}$$
$$= \frac{2}{\sqrt{nm(n+m)}}\left[\frac{(m+2n)\bar{x}-(2m+n)\bar{y}}{\bar{x}+\bar{y}}\right]$$
$$= \frac{2\left[n+(n+m)\left(1-3\hat{\theta}\right)\right]}{\sqrt{nm(n+m)}}.$$

This expression is the one given in formula (14).

Now substituting again the corresponding fourth derivative and the observed Fisher information now in formula (10), the expression for $F_4(\hat{\theta})$ is also obtained,

$$\begin{aligned} F_4(\hat{\theta}) &= \left.\frac{d^4 l_p(\theta;\ x,\ y)}{d\theta^4}\right|_{\theta=\hat{\theta}} I_{\hat{\theta}}^{-2} \\ &= -\frac{6}{nm\,(n+m)}\left[c_1\frac{\bar{y}^2}{(\bar{x}+\bar{y})^2} - c_2\frac{\bar{x}\bar{y}}{(\bar{x}+\bar{y})^2} + c_3\frac{\bar{x}^2}{(\bar{x}+\bar{y})^2}\right] \\ &= -\frac{6}{nm\,(n+m)}\left[c_1\hat{\theta}^2 - c_2\hat{\theta}\left(1-\hat{\theta}\right) + c_3\left(1-\hat{\theta}\right)^2\right]. \end{aligned}$$

This was the expression provided in (15).

Finally, in order to obtain a symmetrizing parametrization for θ, a logarithmic and a power transformation of θ, φ, or of a simple function of these parameters were considered. This was a trial and error procedure, until the reparametrization given in (17) was found. The goal was to obtain a reparametrization $\delta = g(\theta)$, or $\delta = h(\varphi)$ for some simple one to one functions g and h, such that the expression $F_3(\hat{\delta})$ corresponding to the reparametrized log likelihood of δ, was zero.

For the case $n = m$ the parametrization that worked well was $\delta = \log\left(\varphi^{-1}\right) = \log(\beta/\alpha)$. This implies that $\exp(\hat{\delta}) = \hat{\beta}/\hat{\alpha}$ and the profile likelihood of δ is found after substituting $\theta = 1/\left(1+e^{-\delta}\right)$ in the profile likelihood of θ given in (6),

$$L_p(\delta;\ x,\ y) \propto \exp(-m\delta)\left[s\exp(-\delta)+t\right]^{-(n+m)}.$$

The second and third derivatives of the log profile likelihood of δ can be written as,

$$\begin{aligned} \frac{d^2 l_p(\delta;\ x,\ y)}{d\delta^2} &= -\frac{(n+m)\,s\exp(-\delta)}{s\exp(-\delta)+t} + \frac{(n+m)\,s^2\exp(-2\delta)}{\left[s\exp(-\delta)+t\right]^2}, \\ \frac{d^3 l_p(\delta;\ x,\ y)}{d\delta^3} &= \frac{(n+m)\,s\exp(-\delta)}{s\exp(-\delta)+t} - \frac{3\,(n+m)\,s^2\exp(-2\delta)}{\left[s\exp(-\delta)+t\right]^2} \\ &\quad + \frac{2\,(n+m)\,s^3\exp(-3\delta)}{\left[s\exp(-\delta)+t\right]^3} \end{aligned}$$

If $n = m$, then

$$I_{\hat{\delta}} = -\left.\frac{d^2 l_p(\delta;\ x,\ y)}{d\delta^2}\right|_{\delta=\hat{\delta}} = \frac{nm}{n+m} = \frac{n}{2},$$

and

$$\left.\frac{d^3 l_p(\delta;\ x,\ y)}{d\delta^3}\right|_{\delta=\hat{\delta}} = \frac{nm\,(n-m)}{(n+m)^2} = 0.$$

Consequently $F_3(\hat{\delta}) = 0$ as desired, and the corresponding profile likelihood of δ will be symmetric about $\hat{\delta}$.

For the case $n \neq m$, the family of power transformations $\delta = \left(\varphi^{-1}\right)^k = (\beta/\alpha)^k$, where $k \neq 0$, worked well, where k is defined in (16). That is, that value of k was the one that made that $F_3(\hat{\delta}) = 0$. For simplicity in the algebraic expressions, let $a = -1/k$. Then the mle is $\hat{\delta} = (\hat{\beta}/\hat{\alpha})^{-1/a}$ and the profile likelihood of δ is

$$L_p(\delta;\, x,\, y) \propto \delta^{ma} \left(s\delta^a + t\right)^{-(n+m)}.$$

The second and third derivatives of the log profile likelihood of δ can be written as,

$$\frac{d^2 l_p(\delta;\, x,\, y)}{d\delta^2} = -\frac{ma}{\delta^2} - \frac{(n+m)\, sa\,(a-1)\,\delta^{a-2}}{s\delta^a + t} + \frac{(n+m)\, s^2 a^2 \delta^{2a-2}}{(s\delta^a + t)^2},$$

$$\frac{d^3 l_p(\delta;\, x,\, y)}{d\delta^3} = \frac{2ma}{\delta^3} - \frac{(n+m)\, sa\,(a-1)\,(a-2)\,\delta^{a-3}}{s\delta^a + t} + \frac{3\,(n+m)\, s^2 a^2\,(a-1)\,\delta^{2a-3}}{(s\delta^a + t)^2} - \frac{2\,(n+m)\, s^3 a^3 \delta^{3a-3}}{(s\delta^a + t)^3}.$$

Then

$$I_{\hat{\delta}} = -\left.\frac{d^2 l_p(\delta;\, x,\, y)}{d\delta^2}\right|_{\delta=\hat{\delta}} = \frac{nm}{n+m}\left(\frac{a}{\hat{\delta}}\right)^2,$$

and

$$\left.\frac{d^3 l_p(\delta;\, x,\, y)}{d\delta^3}\right|_{\delta=\hat{\delta}} = a\hat{\delta}^{-3}\left[2n - n\,(a-1)\,(a-2) + \frac{3n^2 a\,(a-1)}{n+m} - \frac{2n^3 a^2}{(n+m)^2}\right].$$

Therefore, setting

$$F_3(\hat{\delta}) = \left.\frac{d^3 l_p(\delta;\, x,\, y)}{d\delta^3}\right|_{\delta=\hat{\delta}} I_{\hat{\delta}}^{-3/2} = 0,$$

implies that $a = -3\,(n+m)\,/\,(n-m)$. This is precisely the value of $k = -1/a$ given in formula (16), that was used to symmetrize the profile likelihood function of θ.

References

Box GEP, Cox DR (1964) An analysis of transformations. J R Stat Soc B 26:211–252

Chao A (1982) On comparing estimators of $P(Y < X)$ in the exponential case. IEEE Trans Reliab 31:389–392

Cramer E, Kamps U (1997) A note on UMVUE of $Pr(X < Y)$ in the exponential case. Commun Stat Theory Methods 26:1051–1055

Díaz-Trujillo A, Contreras J, Medina A, Silveyra-León G, Antaramian A, Quirarte G, Prado-Alcalá R (2009) Enhanced inhibitory avoidance learning prevents the long-term memory-impairing effects of cycloheximide, a protein synthesis inhibitor. Neurobiol Learn Mem 91:310–314

Enis P, Geisser S (1971) Estimation of the probability that $Y > X$. J Am Stat Assoc 66:162–168

Ismail R, Jeyaratnan S, Panchapakesan S (1986) Estimation of Pr$[X > Y]$ for gamma distributions. J Stat Comput Simul 26:253–267

Jiang L, Wong ACM (2008) A note on inference for $P(X < Y)$ for right truncated exponentially distributed data. Stat Pap 49:637–651

Kalbfleisch JG (1985) Probability and statistical inference. Springer, New York

Kelley GD, Kelley JA, Schucany WR (1976) Efficient estimation of P$(Y < X)$ in the exponential case. Technometrics 18:359–360

Kotz S, Lumelskii Y, Pensky M (2003) The stress–strength model and its generalizations: theory and applications. World Scientific, River Edge

Krzanowski WJ, Hand DJ (2009) ROC curves for continuous data. CRC Press, Boca Raton

Kundu D, Gupta RD (2005) Estimation of P$[Y < X]$ for generalized exponential distribution. Metrika 61:291–308

Lawless JF (2003) Statistical models and methods for lifetime data. Wiley, Hoboken

Montoya JA, Díaz-Francés E, Sprott DA (2009) On a criticism of the profile likelihood function. Stat Pap 50:195–202

Saracoglu B, Kinaci I, Kundu D (2011) On estimation of $R = P(Y < X)$ for exponential distribution under progressive type-II censoring. J Stat Comput Simul. doi:10.1080/00949655.2010.551772

Sathe YS, Shah SP (1981) On estimating P$(X > Y)$ for the exponential distribution. Commun Stat Theory Methods 10:39–47

Serfling RJ (1980) Approximation theorems of mathematical statistics. Wiley, New York

Sprott DA (2000) Statistical inference in science. Springer, New York

Sullivan Pepe M (2000) An interpretation for the ROC curve and inference using GLM procedures. Biometrics 56:352–359

Tong H (1974) A note on the estimation of $P(Y < X)$ in the exponential case. Technometrics 16:625

Tong H (1975) Errata: a note on the estimation of $P(Y < X)$ in the exponential case. Technometrics 17:395

Ventura L, Racugno W (2011) Recent advances on Bayesian inference for $P(Y < X)$. Bayesian Anal 6:1–18