Reliability Engineering and System Safety 197 (2020) 106787

Contents lists available at ScienceDirect

Reliability Engineering and System Safety

journal homepage: www.elsevier.com/locate/ress

Reliability data analysis of systems in the wear-out phase using a (corrected) q-Exponential likelihood

Ana Cláudia Souza Vidal de Negreiros[a,b], Isis Didier Lins[a,b,*], Márcio José das Chagas Moura[a,b], Enrique López Droguett[c]

[a] Department of Production Engineering, Universidade Federal de Pernambuco, Recife, Brazil
[b] Center for Risk Analysis and Environmental Modeling, Universidade Federal de Pernambuco, Recife, Brazil
[c] Department of Mechanical Engineering, Universidad de Chile, Santiago, Chile

ARTICLE INFO

Keywords:
q-Exponential distribution
Wear-out phase
Monotone likelihood
Firth's correction
Resample correction

ABSTRACT

Maintenance-related decisions are often based on the expected number of interventions during a specified period of time. The proper estimation of this quantity relies on the choice of the probabilistic model that best fits reliability-related data. In this context, the q-Exponential probability distribution has emerged as a promising alternative. It can model each of the three phases of the bathtub curve; however, for the wear-out phase, its usage may become difficult due to the "monotone likelihood problem". Two correction methods (Firth's and resample-based) are considered and have their performances evaluated through numerical experiments. To aid the reliability analyst in applying the q-Exponential model, we devise a methodology involving original and corrected functions for point and interval estimates for the q-Exponential parameters and validation of the estimated models using the expected number of failures via Monte Carlo simulation and the bootstrapped Kolmogorov-Smirnov test. Two examples with failure data presenting increasing hazard rates are provided. The performances of the estimated q-Exponential, Weibull, q-Weibull and modified extended Weibull (MEW) models are compared. In both examples, the q-Exponential presented superior results, despite the increased flexibility of the q-Weibull and MEW distributions in modeling non-monotone hazard rates (e.g., bathtub-shaped).

1. Introduction

Failure and maintenance data, when available, are often used in reliability analyses. The proper adjustment of times between failures (TBFs) to a given probability distribution is a crucial step to support maintenance-related decisions. For instance, maintenance service contracts, the purchase of maintenance materials, hiring of maintenance personnel can be established based on the expected quantity of interventions within a specified period of time. If there is an overestimation of this quantity, resources are unnecessarily allocated for maintenance activities. Otherwise, an underestimated expected number of interventions possibly delays the system return to operation due the lack of preparedness to handle more system stoppages than previously awaited. Both situations result in increased costs. Thus, the choice of a probabilistic model that best fits TBFs, among a number of options, becomes imperative.

In this context, the q-Exponential probability distribution has emerged as an alternative in the modeling of reliability data. It is based on the non-extensive entropy [44,45], it has two parameters – q (shape) and η (scale) – and it is able to represent each of the three phases of the bathtub-shaped hazard rate function: improvement when $1 < q < 2$, useful life when $q \to 1$ and wear-out when $q < 1$. Hence, as the Weibull distribution, it generalizes the Exponential distribution and can model data when the hazard rate is either monotonically decreasing or monotonically increasing.

For example, Sales Filho et al. [41] used the q-Exponential to infer about a useful performance metric in system reliability, the index $R = P(Y < X)$, where Y is the stress, X is the strength and both are supposed independent q-Exponential random variables with different parameters. In the presented application examples involving stress and strength experimental data, the q-Exponential provided better results when compared to the Exponential and Weibull distributions. Zhang et al. [50] proposed a q-Exponential-based model of competing risks and accelerated life tests, which is applied to fit the lifetime of patients whose death (i.e., failure) cause would be either prostatic cancer or vascular disease. Lins et al. [26] developed a q-Exponential generalized renewal process (GRP) and a q-Weibull-GRP. The authors applied both models to fit failure data of complex systems and q-Exponential-GRP

* Corresponding author.
E-mail address: isis.lins@ceerma.org (I.D. Lins).

https://doi.org/10.1016/j.ress.2019.106787
Received 6 December 2018; Received in revised form 12 December 2019; Accepted 28 December 2019
Available online 09 January 2020
0951-8320/ © 2020 Elsevier Ltd. All rights reserved.

outperformed the Weibull-GRP approach. The q-Exponential distribution has also been successfully used in demography [29], physics [5,30], pure statistics [35] and finance [28].

Part of the success of the q-Exponential model in describing some complex systems is due to its ability in modeling the power law behavior with a heavy-tailed probability density function (PDF). Thus, for a given sample with values that have great order of magnitude, the q-Exponential distribution is expected to provide a better fitting if compared to Exponential and Weibull models [41]. The q-Exponential PDF exhibits the power law behavior when $1 < q < 2$. In this case, the estimation of q and η via the maximum likelihood (ML) method presents no difficulties.

Nevertheless, it has been observed in Sales Filho [40] that, when $q < 1$, which corresponds to the wear-out phase, the techniques used to solve the ML problem may provide poor results, as they attain convergence when the parameters' estimates are very large in absolute value. This is an indication of the presence of the so-called "monotone likelihood" problem, which occurs when the log-likelihood obtain its maximum for infinite parameter values [37]. This problem has been related to other probability models: Cox proportional hazards regression model [7,19,27]; $\mathfrak{g}_A^0$ distribution for speckled data modeling from synthetic aperture radar (SAR) images [37]; bimodal Birnbaum-Saunders distribution to fit data from reliability tests to assess the bond behavior of glass-fiber-reinforced plastic rebars to concrete [15]; modified extended Weibull (MEW) distribution [23], which is present in a number of reliability studies as it is able to account for bathub-shaped hazard rates [1,16,46].

In order to tackle the monotone likelihood problem, Firth's penalization [14] and the resample method [8] can be applied. As argued in Fonseca and Cribari-Neto [15], Firth's method is efficient and simple to implement; it is also used in Lima and Cribari-Neto [23]. The resample method, in turn, can also be efficient in the monotone likelihood context, as presented in Pianto & Cribari-Neto [37].

In this work, the monotone likelihood problem associated to the q-Exponential distribution is analytically detailed and, since not every sample is related to a monotone likelihood, the probability of observing such a problem for a given sample is evaluated with numerical experiments involving different combinations of q, η and sample size (n). In the presence of the monotone likelihood problem, Firth's and resample methods will be adopted to penalize the q-Exponential log-likelihood function to obtain appropriate estimates for q and η. The performance of the original and corrected q-Exponential log-likelihood functions are assessed by means of numerical simulations with various q, η and n.

Due to the specificities of the q-Exponential distribution as it may be associated with the monotone likelihood problem, we propose a methodology to apply this probabilistic model to fit reliability-related data sets involving either original or corrected versions of the corresponding log-likelihood function and to validate the estimated models. Point estimates and bootstrap percentile confidence intervals [10,12] for the q-Exponential parameters are considered. The validation portion of the methodology comprises the estimation of the cumulative expected number of failures up to the given real failure times via Monte Carlo simulation (adapted from [26]) and a modified Kolmogorov-Smirnov goodness-of-fit test based on bootstrap (K-S Boot – [43]). With respect to the resolution of the ML problem involving the original and the penalized q-Exponential log-likelihood functions, the Nelder-Mead optimization method [34] is adopted, as it has provided good results in maximizing log-likelihood functions of q-distributions [41,47,50].

The remainder of this paper unfolds as follows. Section 2 brings a theoretical background of the q-Exponential distribution, the q-Exponential log-likelihood function and an investigation of the related monotone likelihood problem. Section 3 provides the Firth's and resample correction methods adapted to the q-Exponential log-likelihood function. Section 4 describes the proposed methodology to apply the q-Exponential distribution to fit reliability-related data and to validate the estimated models. The numerical experiments to assess the performances of the correction methods over the original log-likelihood are given in Section 5. Section 6 provides two examples to illustrate the application of the proposed methodology involving TBFs of a machining center and of a magnetic resonance imaging (MRI) scanner. The Weibull, q-Weibull and MEW distributions are also considered in both examples for comparison purposes. Finally, Section 7 summarizes the main findings of the work and has some directions for future researches.

2. The q-Exponential distribution and the related maximum likelihood problem

The q-Exponential function is defined as

$$\exp_q(t) = \begin{cases} [1 + (1-q)t]^{1/(1-q)}, & \text{if}[1 + (1-q)t] > 0 \\ 0, & \text{otherwise,} \end{cases} \tag{1}$$

and its inverse is defined as

$$\ln_q(t) = \frac{t^{1-q} - 1}{1 - q}. \tag{2}$$

The q-Exponential distribution has the following PDF

$$f(t) = \frac{2-q}{\eta}\exp_q\left(-\frac{t}{\eta}\right) = \frac{2-q}{\eta}\left[1 - \frac{(1-q)t}{\eta}\right]^{\frac{1}{1-q}}, \tag{3}$$

where $q < 2$ determines the PDF shape and is known as entropic index, while $\eta > 0$ is the scale parameter. In the limit $q \to 1$, Eq. (3) recovers the Exponential distribution. When $q < 1$, Eq. (4) has a limited support with an upper bound that depends on η and q, see the following equation:

$$t \in \begin{cases} [0, \infty), & 1 < q < 2, \\ \left[0, \frac{\eta}{1-q}\right], & q < 1, \end{cases}. \tag{4}$$

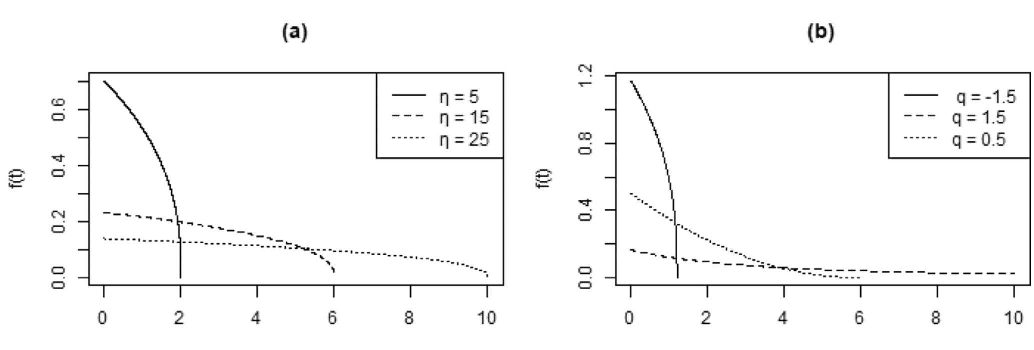

Fig. 1. q-Exponential PDF a) for $q = -1.5$ and some possible values of η; b) for $\eta = 3$ and some possible values of q.

For the sake of illustration, Fig. 1(a) shows the behavior of the q-Exponential PDF for $q = -1.5$, and three possible values of η, and Fig. 1(b) presents the q-Exponential PDF for $\eta = 3$, and three possible values of q. Note, in Fig. 1(b), that when $q = 0.5$ and $\eta = 3$ the support is limited by 6.

The q-Exponential has the following Cumulative Distribution Function (CDF):

$$F(t) = 1 - \left[\exp_q\left(-\frac{t}{\eta}\right)\right]^{2-q} = 1 - \left[1 - \frac{(1-q)t}{\eta}\right]^{\frac{2-q}{1-q}}, \quad t \geq 0. \tag{5}$$

By definition, the hazard rate is $h(t) = f(t)/R(t)$ [31], where $R(t)$ is the reliability function with $R(t) = 1 - F(t)$. Thus, it follows that:

$$h(t) = \frac{(2-q)}{\eta}\left[\exp_q\left(-\frac{t}{\eta}\right)\right]^{-(1-q)} = \frac{(2-q)}{\eta}\left[1 - \frac{(1-q)t}{\eta}\right]^{-1}. \tag{6}$$

The q-Exponential hazard rate can be monotone increasing, monotone decreasing or constant for $q < 1$, $1 < q < 2$ and $q \to 1$, respectively. In fact, this is an important characteristic of the q-Exponential distribution, especially in the reliability context because it enables modeling each of the three phases of the bathtub curve as Weibull model does. Fig. 2 presents examples of increasing and decreasing hazard rates.

In order to generate pseudorandom numbers that follow a q-Exponential distribution, Eq. (5) can be used:

$$t = \frac{\eta\left[1 - U^{\left(\frac{1-q}{2-q}\right)}\right]}{1-q}, \tag{7}$$

where U denotes a uniform pseudorandom number. The formula in Eq. (7) is obtained by means of the inverse transform method [39]. The q-Exponential pseudorandom generator will be used in the numerical experiments presented in Section 5.

In this paper, the ML method is adopted because of its properties such as asymptotic unbiasedness, strong consistency and efficiency [42]. For a given sample $t = (t_1, \ldots, t_i, \ldots t_n)$ of TBFs, the q-Exponential likelihood function is given by

$$L(t|q, \eta) = \prod_{i=1}^{n} \frac{2-q}{\eta} exp_q\left(-\frac{t_i}{\eta}\right) = \prod_{i=1}^{n} \frac{2-q}{\eta}\left[1 - \frac{(1-q)t_i}{\eta}\right]^{\frac{1}{1-q}}. \tag{8}$$

The corresponding log-likelihood function is

$$l(t|q, \eta) = n\ln\left(\frac{2-q}{\eta}\right) + \frac{1}{1-q}\sum_{i=1}^{n}\ln\left[1 - \frac{(1-q)t_i}{\eta}\right]. \tag{9}$$

To obtain the ML estimators for the parameters, the log-likelihood function is maximized. This can be done by setting the first derivatives of l w.r.t. each parameter to zero. The q-Exponential score equations are the following:

$$\frac{\partial l}{\partial q} = -\frac{n}{2-q} + \frac{1}{(1-q)^2}\sum_{i=1}^{n}\ln\left[1 - \frac{(1-q)t_i}{\eta}\right] + \frac{1}{1-q}\sum_{i=1}^{n}\frac{t_i}{\eta - (1-q)t_i} = 0, \tag{10}$$

$$\frac{\partial l}{\partial \eta} = \frac{1}{\eta}\left[-n + \sum_{i=1}^{n}\frac{t_i}{\eta - (1-q)t_i}\right] = 0. \tag{11}$$

It can be noticed that Eqs. (10) and (11) do not have a closed-form solution. Thus, nonlinear optimization methods can be used to obtain the parameters' estimates. However, when $q < 1$, these methods often converge when the parameters' estimates are very large, which is an indication of the monotone likelihood problem [4,23,37].

To demonstrate that the q-Exponential log-likelihood function may present a monotone behavior when $q < 1$, we must show that its limit converges to a finite value as either $q \to -\infty$ or $\eta \to \infty$. Let t_{max}^o be the largest observed value in the sample. It must be strictly lower than $t_{max} = \eta/(1-q)$ (see Eq. (4)), otherwise the argument of the logarithm in the second part of Eq. (9) could be either 0 or negative.

Then, let δ be the difference between t_{max} and t_{max}^o. By means of some algebraic manipulations, Eq. (9) is rewritten as a function of one parameter (q or η) at a time, which enables the calculation of the limits in Eqs. (12) and (13), which are the same and depend on n, t_{max}^o and δ.

$$\lim_{q \to -\infty}\left[n\ln\left(\frac{1}{t_{max}^o + \delta}\right) + n\ln\left(\frac{2-q}{1-q}\right) + \frac{1}{1-q}\sum_{i=1}^{n}\left(1 - \frac{t_i}{t_{max}^o + \delta}\right)\right] = n\ln\left(\frac{1}{t_{max}^o + \delta}\right) \tag{12}$$

$$\lim_{\eta \to \infty}\left[n\ln\left(\frac{1}{\eta} + \frac{1}{t_{\max}^o + \delta}\right) + \sum_{i=1}^{n}\ln\left(\left(1 - \frac{t_i}{t_{\max}^o + \delta}\right)^{(t_{\max}^o + \delta)/\eta}\right)\right] = n\ln\left(\frac{1}{t_{\max}^o + \delta}\right). \tag{13}$$

For a given sample, n and t_{max}^o are determined and δ is the only free quantity in the quest for the optimal solution. Note that the limit is maximum when $\delta \to 0$ (Eq. (14)), which means that, if the q-Exponential log-likelihood function is not maximized with reasonable parameter estimates in terms of magnitude, it will be maximized when $t_{max} = \eta/(1-q) \to t_{max}^o$. In the latter situation, the nonlinear optimization algorithms tend to provide large absolute values for the parameters' estimates such that $\eta/(1-q) \to t_{max}^o$ as an attempt to reach the theoretical maximum limit.

$$\lim_{\delta \to 0} n\ln\left(\frac{1}{t_{max}^o + \delta}\right) = n\ln\left(\frac{1}{t_{max}^o}\right). \tag{14}$$

Therefore, the q-Exponential log-likelihood function has a finite value when $q \to -\infty$, $\eta \to \infty$ and $\eta/(1-q) \to t_{max}^o$. The theoretical limit in Eq. (14) is an asymptote, which may not be attained with practical

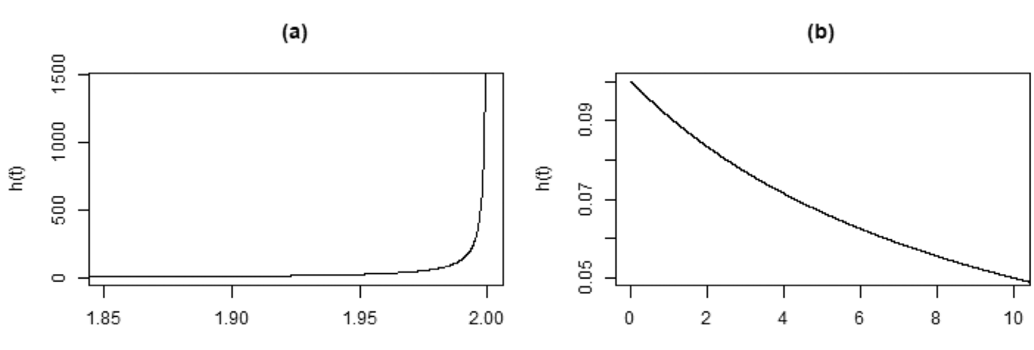

Fig. 2. q-Exponential hazard rate a) for $q = -1.5$ and $\eta = 5$; b) for $q = 1.5$ and $\eta = 5$.

values for the parameter estimates. Then, in the cases where the asymptote is the actual maximum value, the monotone behavior is observed: by decreasing q, increasing η and preserving the relation $\eta/(1-q) > t^{o}_{max}$, the difference between the associated log-likelihood value and the theoretical limit of Eq. (14) becomes smaller.

It is important to notice that not all samples originally drawn from a q-Exponential distribution with $q < 1$ will necessarily present the monotone likelihood problem. Even the limits of Eqs. (12) and (13) being valid for all samples, for some of them, the log-likelihood may attain its maximum – thus, a value greater than the limit presented in Eq. (14) – at finite estimates for q and η and the nonlinear optimization methods provide these finite estimates as optimal values.

In order to assess the probability that a sample originated from a q-Exponential distribution with $q < 1$ presents the monotone likelihood problem, we performed the following numerical experiment based on Lima and Cribari-Neto [23]. We considered different q-Exponential distributions ($q = -20, -2$ and 0.5 with η= 5; q= −2 with η= 50, 500 and 1000) and sample sizes (n= 20, 100, 500, 1000 and 5000); for each combination of parameters and sample size, 1000 samples were generated and the corresponding parameters were obtained. If the estimate for at least one of the parameters is greater, in magnitude, than twice the respective true value, we considered that the monotone likelihood problem is present. Thus, for a setting (q, η and n), the proportion of samples for which these large parameters estimates are obtained is set as the probability that a sample is related to the monotone likelihood problem. The results of the experiment are in Table 1. The smaller the sample size and the value of parameter q, the greater the probability of observing the monotone likelihood problem. In the case of fixed q, varying η does not significantly alter the probabilities and the previously commented pattern can be observed as n decreases.

In reliability analyses, data availability is often restricted to small samples. Therefore, to turn the q-Exponential distribution into a viable alternative for modeling data related to equipment in the wear-out phase, a method to circumvent the estimation difficulty in the presence of monotone likelihood should be at analyst's disposal.

3. Correction methods for the *q*-Exponential monotone log-likelihood function

This section presents the Firth's penalization and resample methods to be applied to the q-Exponential monotone likelihood problem.

3.1. Firth's penalization method

A method to penalize the log-likelihood function in order to reduce the bias of the ML estimator was proposed by Firth [14]. The underlying idea of this method is that since the parameter estimate may not exist, it is safer to modify the estimation equations for bias correction prior to estimation.

Table 1
Probability of observing the monotone likelihood problem for a given sample; 1000 samples were drawn from different q-Exponential distributions when $q < 1$.

	$\eta = 5$		$q = -2$	
n	q	Probability	η	Probability
20	−20	0.826	50	0.777
	−2	0.762	500	0.697
	0.5	0.327	1000	0.766
100	−20	0.692	50	0.367
	−2	0.426	500	0.314
	0.5	0.003	1000	0.378
500	−20	0.556	50	0.040
	−2	0.042	500	0.026
	0.5	0.000	1000	0.000
1000	−20	0.408	50	0.000
	−2	0.001	500	0.001
	0.5	0.000	1000	0.002
5000	−20	0.079	50	0.000
	−2	0.000	500	0.000
	0.5	0.000	1000	0.000

Let $U^{*}(\theta)$, in which θ represents the set of parameters, be the modified score function. For the exponential family model, the r-th component of the modified score equation is given by:

$$U_r^{*}(\theta) = U_r(\theta) + A_r(\theta), \tag{15}$$

where $A_r(\theta)$ is the r-th part of $A(\theta) = -I(\theta)B_1(\theta)/n,\ \ r = 1, \cdots, \dim(\theta)$. $I(\theta)$ is the Fisher information and $B_1(\theta)$ is the first order term in the bias expansion on the ML estimator:

$$B(\theta) = B_1(\theta)/n + B_2(\theta)/n + \cdots. \tag{16}$$

Eq. (16) refers to the asymptotic expansion of the bias

$$B(\theta) = E(\hat{\theta}) - \theta. \tag{17}$$

The Fisher's information (observed) does not depend on data, and it follows that

$$A_r(\theta) = \frac{\partial}{\partial\theta_r}\left(\frac{1}{2} ln|I(\theta)|\right). \tag{18}$$

The correction of the likelihood function is applied as following

$$L^{*}(\theta|t) = L(\theta|t)\,|K|^{1/2}, \tag{19}$$

where K refers to the determinant of the Fisher information matrix, and the penalization term $|K|^{1/2}$ is the Jeffreys invariant prior [20]. Then, by applying the logarithm in Eq. (19), parameter estimation can be executed by maximizing

$$l^{*}(\theta|t) = l(\theta|t) + \frac{1}{2}\ln|K|. \tag{20}$$

Even though Firth [14] applied the penalization method to probability distributions of the exponential family, it can also be used for the correction of other models [15,23].

In this work, the Firth's method penalizes the q-Exponential log-likelihood function. Under regularity conditions and for large samples, the estimator $\hat{\theta}$ approximately follows a Normal distribution with parameters $(\theta, I(\theta)^{-1})$, where $I(\theta)$ is Fisher's (expected) information matrix:

$$I(\theta) = E\left[\frac{\partial l(\theta)}{\partial\theta}\ \frac{\partial l(\theta)}{\partial\theta^{T}}\right]. \tag{21}$$

The score function of the q-Exponential log-likelihood is presented in Eqs. (10) and (11). In general, $I(\theta) = E[J(\theta)]$ is easier to compute, where $J(\theta) = -\partial^2 l(\theta)/\partial\theta\partial\theta^{T}$ is the observed information and we also can see it is the negative of the Hessian matrix. For the q-Exponential model,

$$J(\theta) = -\begin{bmatrix} J_{qq} & J_{q\eta} \\ J_{\eta q} & J_{\eta\eta} \end{bmatrix},$$

where

$$J_{qq} = -\frac{n}{(2-q)^2} + \frac{2}{(1-q)^3}\sum_{i=1}^{n}\ln\left(1 - \frac{(1-q)t_i}{\eta}\right) + \frac{2}{(1-q)^2}\sum_{i=1}^{n}\frac{t_i}{\eta\left(1-\frac{(1-q)t_i}{\eta}\right)} + \frac{1}{(1-q)}\sum_{i=1}^{n}\frac{t_i^2}{\eta^2\left(1-\frac{(1-q)t_i}{\eta}\right)} \tag{22}$$

$$J_{\eta\eta} = \frac{n}{\eta^2} - \frac{1}{(1-q)}\sum_{i=1}^{n}\left[\frac{2(1-q)t_i}{\eta^2\left(1-\frac{(1-q)t_i}{\eta}\right)} + \frac{t_i^2(1-q)^2}{\eta^4\left(1-\frac{(1-q)t_i}{\eta}\right)}\right] \tag{23}$$

$$J_{q\eta} = J_{\eta q} = \frac{1}{(1-q)^2}\sum_{i=1}^{n}\frac{(1-q)t_i}{\eta^2\left(1-\frac{(1-q)t_i}{\eta}\right)} - \frac{1}{(1-q)}\sum_{i=1}^{n}\left[\frac{t_i}{\eta^2\left(1-\frac{(1-q)t_i}{\eta}\right)} + \frac{t_i^2(1-q)}{\eta^4\left(1-\frac{(1-q)t_i}{\eta}\right)}\right] \quad (24)$$

Thus, the original idea of this method involves the utilization of the matrix of the expected or the observed information. However, in some cases, the expected information is not easily obtained. Therefore, we here used the matrix of the observed information as an approximation for Fisher information. The penalized q-Exponential log-likelihood in accordance with Firth's method is obtained by respectively replacing $l(\theta|t)$ and $\ln|K|$ in Eq. (20) by Eq. (9) and by the determinant of the corresponding Hessian matrix formed by Equations (22)-(24).

3.2. Resample method

The resample method is based on a change in the log-likelihood function through non-parametric bootstrap. It is an adaptation, proposed by Cribari-Neto et al. [8], of Efron's [13] "better" bias estimation. Conversely to Efron's proposal, the resample method does not require the quantity of interest to have closed form.

Let t be a random sample of size n from a q-Exponential distribution. Each of the B bootstrap samples can be described by the weight that each observation receives in the new empirical distribution function. For the original sample, each observation receives weight $1/n$ [37]. This can be succinctly recorded in a vector $P^0 = 1/n(1, \cdots, 1)$. For the b-th bootstrap sample, this vector becomes:

$$P^b = \frac{1}{n}(\#\{t_1\}_b, \cdots, \#\{t_n\}_b), \quad (25)$$

where $\#\{t_j\}_b$ is the number of times that t_j occurs in the b-th bootstrap sample. This vector is called the resample vector. Efron's idea is based on the possibility of writing the parameter estimate as a closed form function of the data using P^0. For instance, from the Efron's proposal the estimate of the mean can be written as $\bar{T} = T(P^0) = P^0 t$.

In order to accelerate the convergence of the bias estimate such that fewer bootstrap replications are required, Efron [13] suggests that, when calculating the bias estimate, one subtracts the parameter estimate resulted from

$$\boldsymbol{P}^* = \frac{1}{B}(P^1, \ P^2, \ ..., P^b) = \frac{1}{B}\sum_{b=1}^{B} P^b. \quad (26)$$

If we write the estimate of interest, obtained from t, as $T(P^0)$, then we obtain bootstrap estimates $(\hat{q}, \hat{\eta})^{*b}$ using the resample vectors P^{*b}, $b = 1, 2, ..., B$, as $T(P^{*b})$. A bootstrap bias corrected estimate (BBC), proposed by Efron [13], is formed by subtracting the estimated bias (Eq. (17)), from the original estimate as follows

$$(\hat{q}, \hat{\eta})_{BBC} = (\hat{q}, \hat{\eta}) - T(P^*), \quad (27)$$

where P^* is presented in Eq. (26). Then, the new better bootstrap bias correction (BBBC) estimate, presented by Cribari-Neto et al. [8], would be

$$(\hat{q}, \hat{\eta})_{BBBC} = (\hat{q}, \hat{\eta}) - \left[\frac{1}{B}\sum_{b=1}^{B}(\hat{q}_b, \hat{\eta}_b) - T(P^*)\right]. \quad (28)$$

The ML estimators for the q-Exponential model do not have a closed form, as can be seen in Eqs. (10) and (11), which was also verified for the $g_A^0(\alpha, \gamma, n)$ distribution in Cribari-Neto et al. [8]. Thus, Eqs. (27) or (28) cannot be directly used for bias reduction.

However, to use the BBBC, Cribari-Neto et al. [8] write the estimators as a function of P^0, and then use the estimate obtained by replacing P^0 with P^* to correct the bias. According to Cribari-Neto et al. [8], this approach is expected to provide accurate point estimates.

For the q-Exponential distribution, we rewrite the log-likelihood in Eq. (9) as a function of P^0:

$$(\hat{q}, \hat{\eta}) = \underset{(q,\eta)}{\operatorname{argmax}}\left\{n\ln\left(\frac{2-q}{\eta}\right) + \frac{1}{1-q}\left[P^0\ln\left(1-\frac{(1-q)t_i}{\eta}\right)\right]\right\}. \quad (29)$$

Then, $(\hat{q}, \hat{\eta})$ can be obtained by replacing P^0 by P^* in Eq. (29) using the corresponding result in Eq. (27) to generate their BBBC estimates.

4. Methodology for the q-Exponential distribution adjustment to reliability-related data and model validation

Given that the monotone likelihood problem may be present when using the q-Exponential distribution to fit reliability-related data sets associated to systems in the wear-out phase, a specific methodology is devised to aid the reliability analyst in applying such probability model. It involves not only parameter estimation via ML method and application of a correction procedure when necessary, as described in the previous sections, but also the estimation of confidence intervals for the parameters, which is often required in reliability applications, and a model validation phase. In this work, all point estimates are given by the Nelder-Mead optimization method.

An overview of the proposed methodology is in Fig. 3. It starts with a data set of TBFs that should be fit by means of the original q-Exponential log-likelihood function. Note that the answer "Yes" in the first decision epoch indicates that the analyzed data are related to a system with a decreasing hazard rate, which is not investigated in the present work since in this case there is no monotone likelihood problem. Nevertheless, the methodology depicted in Fig. 3 is general as it covers all possible values for the parameter q. Thus, it considers data associated to systems with decreasing or increasing hazard rate functions.

The answer "No" to the second decision epoch refers to the situation in which, despite the increasing hazard rate, with the considered sample of TBFs, the q-Exponential log-likelihood does not have a monotone behavior. Otherwise, the answer "Yes" to the second decision epoch in Fig. 3 indicates the presence of the monotone likelihood behavior. In this case, an additional evidence can be verified by comparing the obtained values for the original q-Exponential log-likelihood function and the theoretical limit of Eq. (14). If the former is slightly smaller than the latter, the monotone likelihood problem exists.

The percentile confidence intervals for q and η are obtained by means of a non-parametric bootstrap method [10,12]. These types of confidence intervals were used by Lins et al. [25] and in the context of the q-Exponential distribution by Sales Filho et al. [41]. B bootstrap samples (e.g., 1000) of the same size of the original data set are constructed based on sampling with replacement. For each of these bootstrap samples, ML estimates are obtained for q and η. Then, for a given confidence level γ, the quantiles $\gamma/2$ and $1-\gamma/2$ of the B estimates for q and η are set as the lower and upper bounds of the corresponding confidence intervals. If the monotone likelihood problem is present, then the corrected version of the q-Exponential log-likelihood is used in the parameter estimation related to each of the B bootstrap samples. Otherwise, the original q-Exponential log-likelihood function is considered.

The validation procedure (in light gray in Fig. 3) is based on the (i) estimation of the cumulative expected number of failures up to the given real failure times using Monte Carlo simulation (briefly described in Section 4.1) and on the (ii) assessment of the goodness-of-fit of the estimated q-Exponential probabilistic model by means of the KS-Boot test (Stute et al. 2013), whose main steps are summarized in Fig. 4. This test was also adopted by Sales Filho et al. [41] and Xu et al. [47] in the context of q-distributions for reliability applications.

Approaches (i) and (ii) are independent forms of validation, there is no precedence relation between them. Each of them is fed with the parameter estimates obtained via either original or corrected q-Exponential log-likelihood function. Moreover, the validation

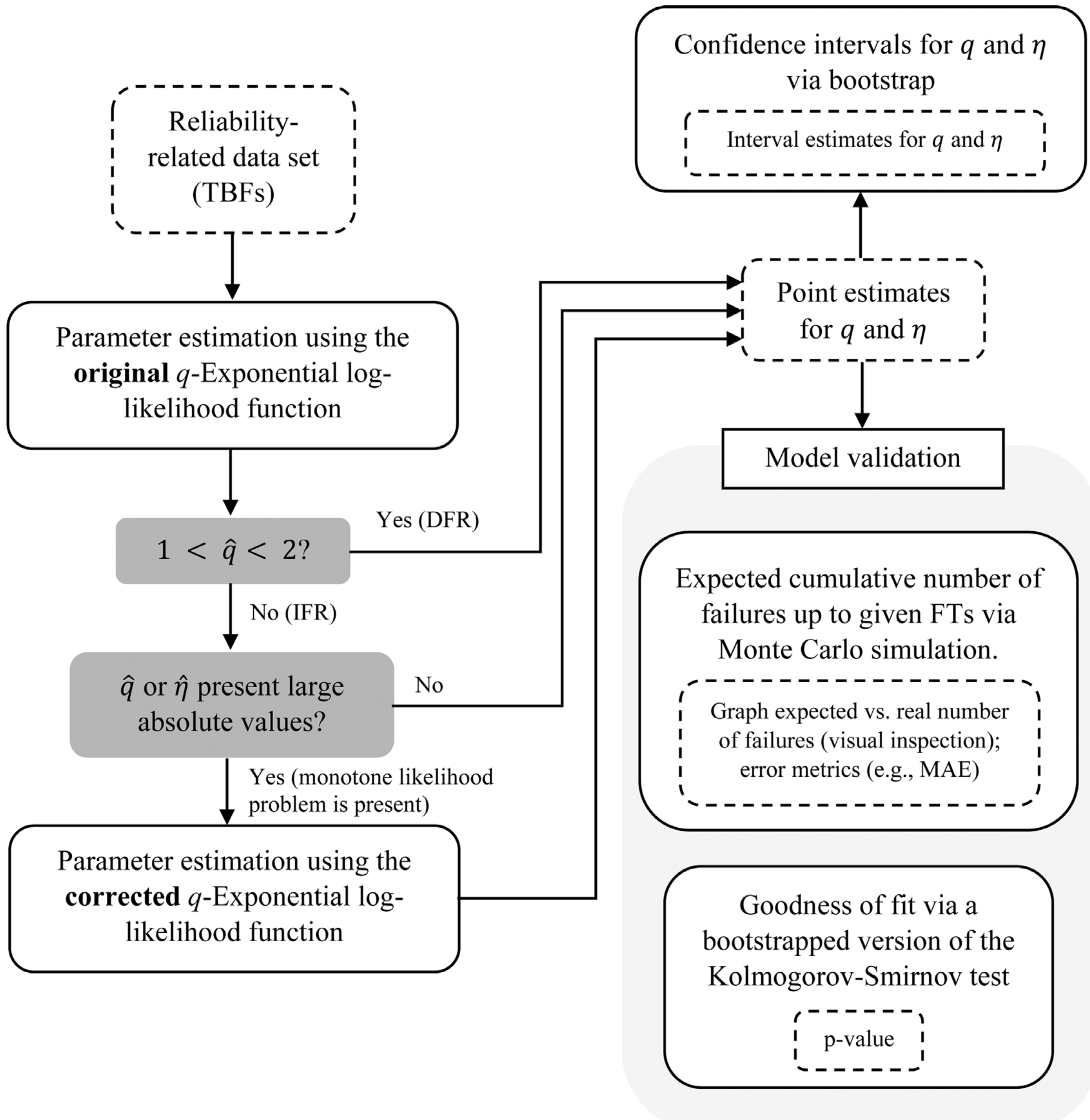

Fig. 3. Proposed methodology for the application of the q-Exponential distribution to fit reliability-related data sets. Decision epochs are in dark gray. Dashed lines represent input or output data. DFR – decreasing failure rate; IFR – increasing failure rate.

techniques are not specific for the q-Exponential model and can be applied to other probability distributions (e.g., Weibull). The results obtained from each of the estimated models can be compared to assess which of the considered ones is deemed the best.

4.1. Cumulative expected number of failures via Monte Carlo simulation

A Monte Carlo algorithm can be used to compare simulated outcomes of the estimated q-Exponential model with real failure data. Based on the inverse transform of the q-Exponential distribution in Eq. (7) with parameters q and η replaced by their respective estimates, we can randomly generate failure times and count the number of simulated failures that have occurred by $y_1 = t_1$, $y_2 = t_1 + t_2$, $y_3 = t_1 + t_2 + t_3$, …, $y_n = \sum_{i=1}^{n} t_i$, which are real failure times known. By replicating this procedure (e.g. 10,000 times), a cumulative expected number of failures can be associated to each of those instants and can be compared to the corresponding real number of failures 1, 2, 3, …, n. Such a comparison can be visually assessed by using a graph of the expected and real number of failures vs. time and also using an error metric such as the mean absolute error (MAE), which is computed as

$$MAE = \sum_{i=1}^{n} \frac{|E[\hat{N}(y_i)] - N(y_i)|}{n}, \quad (30)$$

where $E[\hat{N}(y_i)]$ is the expected number of failures up to y_i, $N(y_i)$ is the actual number of failures up to y_i.

The Monte Carlo framework used in this work is a specific case of the simulation algorithm presented in Lins et al. [26], devised for the more general q-Weibull-GRP and q-Exponential-GRP models. Precisely, considering the q-Exponential-GRP, the only required modification therein is to set the GRP parameter as zero, given that in the present work only perfect repairs are considered.

5. Numerical experiments

To evaluate the performance of the methods described in Section 4, numerical experiments were performed using Monte Carlo simulations. Eq. (7), presented in Section 2, was used to generate the pseudorandom numbers that follow a q-Exponential distribution. These numerical experiments were run with 10,000 replications, and with 60 sub-experiments for each method of correction (Firth's and resample) for the q-

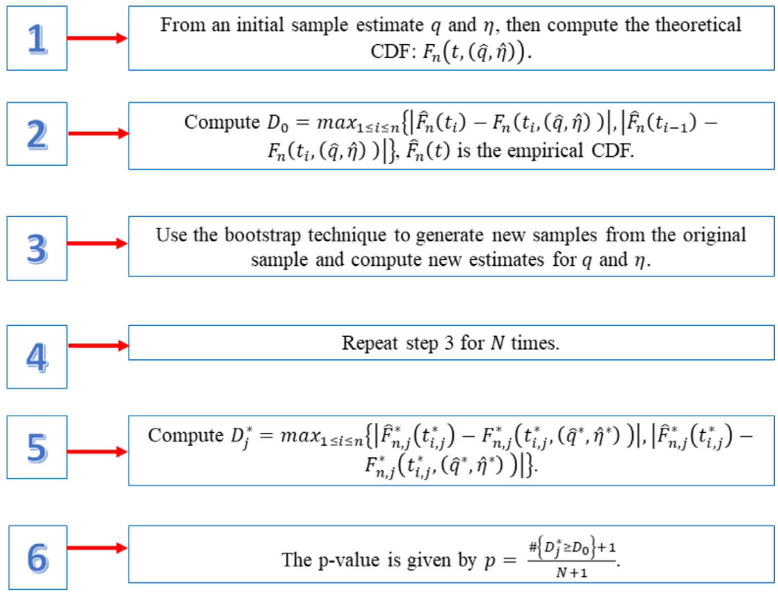

Fig. 4. Framework of K-S Boot. The * is related to bootstrap operations and $\#\{D_j^* \ge D_0\}$ is the number of times that D_j^* ($j = 1, 2, \cdots, N$) is bigger than D_0. Then, the null hypothesis will be rejected if $D_0 > D^*_{(N(1-\alpha)+1)}$ for a significance level α.

Exponential log-likelihood function and for the original function (without any correction).

The simulations were initially executed with η constant ($\eta = 5$), for $q \in \{-20, -2, 0.5\}$ (Table 2). Then, q was remained constant ($q = -2$) and $\eta \in \{50, 500, 1000\}$ (Table 3). Moreover, five sample sizes were tested: $n \in \{20, 100, 500, 1000, 5000\}$. All experiments were run by using the Nelder-Mead [34] approach and the computational software R (function optim) [38].

Tables 2 and 3 present the relative bias obtained using the original q-Exponential log-likelihood, the q-Exponential log-likelihood penalized by the Firth's method and by the resample method. The numerical implementations of the Firth's method achieved good results for all sample sizes. Indeed, the highest relative biases provided by this method were 0.83 for parameter η with $n = 20$, $\eta = 50$ and $q = -2$ (Table 3), and 1.24 for parameter q with $n = 20$, $\eta = 5$ and $q = -2$ (Table 2). Moreover, the Firth's method yielded a relative bias of 0.00 for $q = 0.5$ with $n = 5000$ (Table 2), which demonstrates that the relative bias goes to zero as $n \to \infty$, considering the log-likelihood function corrected by this method. Another result using the Firth's method is that even for small samples ($n = 20$), it had a satisfactory performance, as it provided low relative biases for all cases analyzed.

The resample method in turn produced poor results for samples with 20 observations (Tables 2 and 3). As the sample size increases, the results get better, but, in general, the relative biases achieved by the Firth's method are lower than that provided by the resample method.

In addition, the original q-Exponential function returned lower relative biases only from samples with 500 realizations, with one exception: for $n = 100$ and $q = 0.5$ the relative bias was − 0.27, which may have happened because of the moderate sample size combined with the value of parameter q near the lower bound of the interval $1 < q < 2$, for which the monotone behavior is not present.

In general, the results obtained by the Firth's method were superior (76,67% of the analyzed cases) when compared to the ones obtained by the resample method and the original function. Moreover, Firth's penalization had the worst performance in none of the cases. On the other hand, the original function was the best in 23.33% and the worst in 46.67% of the cases. The function corrected by the resample method was the best in none of the situations and had the worst performance in 53.33% of the considered cases.

Given that the Firth's correction has provided the best results in most of the experiments, an additional assessment was performed to observe the impact of this method in the variability of the provided

Table 2
Relative bias of $\hat{q}$ and $\hat{\eta}$ and monotone behavior of the original q-Exponential log-likelihood for $\eta = 5$.

		Relative bias of $\hat{q}$			Relative bias of $\hat{\eta}$			Monotone behavior of the q-Exponential log-likelihood					
n	q	Original	Firth	Resample	Original	Firth	Resample	$\hat{\eta}/(1-\hat{q})$ (A)	t^o_{max} (B)	$\hat{\delta}$ ($C = A$-B)	Limit Value (D)	Obtained log-likelihood (E)	D – E
20	−20	7,095,017	**−0.58**	202,927.1	31,963,831	**−0.58**	172,522.2	0.1880	0.1880	1.0050e-08	33.3580	33.3580	1.2350e-06
	−2	49,025,880	**1.24**	1,733,321	146,483,739	**0.64**	955,474.3	1.5502	1.5412	0.0090	−8.6512	−8.6817	0.0305
	0.5	−9,856,444	**−0.22**	−0.91	28,659,849	**0.52**	−0.86	7.8900	7.8860	0.0040	−41.3018	−41.3034	0.0015
100	−20	1,142,199	**−0.55**	−0.84	1,075,591	**−0.53**	−0.84	0.2369	0.2369	1.4420e-09	143.9830	143.9830	1.8070e-06
	−2	800,476.8	**1.13**	−0.92	511,675.2	**0.70**	−0.90	1.5699	1.5699	7.1504e-08	−45.1012	−45.1012	4.5533e-06
	0.5	**−0.27**	0.82	−0.83	0.16	**−0.11**	−0.90	7.8057	7.8057	3.0754e-06	−205.4863	−205.4863	3.9365e-05
500	−20	110,654	**−0.50**	−0.97	105,104	**−0.48**	−0.97	0.2374	0.2373	0.0001	719.0754	718.9641	0.1113
	−2	**0.41**	0.80	−0.97	**0.27**	0.52	−0.93	1.6436	1.6436	4.8847e-07	−248.4681	−248.4683	0.0002
	0.5	**−0.06**	0.25	−0.80	0.03	**−0.01**	−0.93	9.0312	9.0310	0.0002	−1100.334	−1100.348	0.0133
1000	−20	19,791.20	**0.07**	−0.98	18,823.35	**0.06**	−0.98	0.2378	0.2378	1.5700e-09	1436.0180	1436.0180	1.8780e-05
	−2	**0.09**	0.79	−0.98	**0.06**	0.52	−0.96	1.6626	1.6617	0.0008	−507.8739	−508.3999	0.5260
	0.5	−0.06	**−0.01**	−0.90	0.04	**0.00**	−0.96	8.5931	8.5928	0.0003	−2150.928	−2150.963	0.0353
5000	−20	27.17	**0.08**	−0.95	25.86	**0.08**	−0.95	0.2379	0.2379	4.0068e-09	7179.516	7179.516	8.3921e-05
	−2	**0.03**	0.90	−0.98	**0.01**	0.59	−0.96	1.6667	1.6663	0.0003	−2553.196	−2554.165	0.9693
	0.5	−0.22	**0.00**	−0.91	0.19	**0.00**	−0.96	9.7706	9.7705	0.0001	−11,396.86	−11,396.93	0.0699

Table 3
Relative bias of $\hat{\eta}$ and $\hat{q}$ and monotone behavior of the original q-Exponential log-likelihood for $q = -2$.

n	η	Relative bias of $\hat{q}$			Relative bias of $\hat{\eta}$			Monotone behavior of the q-Exponential log-likelihood					
		Original	Firth	Resample	Original	Firth	Resample	$\hat{\eta}/(1-\hat{q})$ (A)	t^o_{max} (B)	$\hat{\delta}$ ($C = A$-B)	Limit Value (D)	Obtained log-likelihood (E)	D – E
20	50	5,913,085	**0.83**	412,424.7	9,885,217	**0.40**	227,344.4	15.9614	15.9614	1.34748e-07	−55.4035	−55.4035	1.5430e-06
	500	3,080,172	**0.81**	241,109.6	5,147,048	**0.39**	132,908.7	162.8168	162.8168	6.33753e-06	−101.8525	−101.8525	2.3405e-06
	1000	2,686,289	**0.82**	154,987.9	4,485,449	**0.39**	85,435.02	265.8934	265.8933	0.0001	−111.6619	−111.6619	1.0435e-05
100	50	99,019.33	**0.66**	−0.93	154,844	**0.40**	−0.90	15.9722	15.9719	0.0002	−277.0832	−277.0849	0.0017
	500	51,054.6	**0.66**	−0.95	79,865.48	**0.39**	−0.92	166.0675	166.065	0.0025	−511.2379	−511.2394	0.0015
	1000	43,026.61	**0.65**	−0.95	67,355.71	**0.39**	−0.91	313.3251	313.3250	0.0001	−574.7241	−574.7241	2.7755e-05
500	50	**0.15**	0.59	−0.97	**0.24**	0.38	−0.93	16.3936	16.3936	1.54317e-05	−1398.4470	−1398.4474	0.0004
	500	**0.16**	0.61	−0.97	**0.25**	0.40	−0.93	165.7122	165.7119	0.0002	−2555.1250	−2555.1260	0.0010
	1000	**0.16**	0.61	−0.97	**0.25**	0.39	−0.93	324.1337	324.1333	0.0004	−2890.5770	−2890.5780	0.0010
1000	50	0.09	**0.00**	−0.98	0.06	**0.00**	−0.96	16.4461	16.4461	2.05811e-05	−2800.0890	−2800.0900	0.0010
	500	0.09	**0.08**	−0.98	0.09	**0.04**	−0.96	166.2581	165.9489	0.3092	−5111.6800	−5930.4550	818.7750
	1000	0.15	**0.09**	−0.77	0.09	**0.05**	−0.78	332.2789	331.4393	0.8396	−5803.445	−7097.4250	1293.9800
5000	50	−0.29	**0.00**	−0.98	−0.19	**0.00**	−0.96	16.6867	16.6600	0.0267	−14,065.0600	−16,964.8000	2899.7400
	500	0.13	**0.05**	−0.99	0.08	**0.03**	−0.96	166.2546	166.1816	0.0730	−25,565.4100	−25,567.6000	2.1900
	1000	0.14	**0.06**	−0.77	0.09	**0.04**	−0.78	333.5014	332.2237	1.2777	−29,029.04	−31,564.3700	2535.3300

parameters' estimates. We took the setting $n = 20$, $q = -2$ and $\eta = 5$ of the experiment described in Section 3 and used both original and Firth's corrected log-likelihood functions to estimate the parameters. A small n was selected to represent a rather common sample size in reliability data analyses. The original log-likelihood function provided 7.6460e+14 and 1.6832e+15 as variances for q and η estimates, respectively, whereas Firth's correction counterparts were 0.1397 and 0.1621. Also, in contrast to the probability 0.762 reported in Table 1, for the q-Exponential log-likelihood with Firth's correction, Nelder-Mead optimization method provided estimates lower, in magnitude, than twice the true parameters' values for each of the 1000 samples, i.e., the correction was indeed able to circumvent the monotone likelihood problem.

The last six columns of Table 2 (resp. Table 3) are related to the monotone behavior of the original q-Exponential log-likelihood function. Their values were computed with a single sample randomly generated using the corresponding row setting: q (resp. η), sample size and $\eta = 5$ (resp. $q = -2$). The operation A refers to the upper limit of the support computed with the parameter estimates, B is the greatest sample value and has to be always lower than A; this condition is satisfied for all the examples in Tables 2 and 3.

Moreover, for almost all the cases (Tables 2 and 3) A and B are very close, which means that the optimization method attempts to get the ratio in (A) as closest as possible to the greatest value of each sample by means of the parameters' estimates. C is the difference A – B and is an estimate of δ. The values D are computed using Eq. (14), which considers the limit when $\delta \to 0$. E is the value of the original log-likelihood function obtained with the parameters' estimates $\hat{q}$ and $\hat{\eta}$.

Note that, as shown in Eqs. (12) and (13), when $\delta \to 0$, the limits reach the maximum value. In this way, the optimization method provides the ratio in A closer to t^o_{max}, satisfying the support's function at the same time even with large parameter estimates (monotone behavior). In other words, A is always strictly greater than t^o_{max}, but with parameters' estimates far from the true parameters' values, as can be seen by the relative bias in Tables 2 and 3 specially for small sample sizes. The last column brings the difference D – E, which is always positive because of the reasons previously presented.

Additionally, note the boldfaced values in Tables 2 and 3 are the best in their row for each parameter and the underlined ones are the worst in their row for each parameter. The six last columns present, respectively, the relation $\hat{\eta}/(1-\hat{q})$ obtained with the parameter estimates given by the optimization of the original log-likelihood, t^o_{max} of the corresponding sample, the estimated difference δ, the theoretical maximum limit (Eq. (14)) computed with n and t^o_{max} related to each sample, the original q-Exponential log-likelihood value computed with the obtained parameter estimates and the observed difference between the theoretical limit and the obtained log-likelihood value.

6. Application examples

In order to illustrate the methodology presented in Section 4, two application examples are considered involving failure data related to the third phase of the bathtub curve (i.e., increasing hazard rate). Thus, in both cases, the parameter q of the q-Exponential distribution is less than one.

In the first example, the associated q-Exponential log-likelihood function presents the monotone likelihood problem and a correction method should be used to turn the q-Exponential model into a viable candidate distribution. For the second example, despite the increasing hazard rate, the monotone likelihood problem is not present. Therefore, a correction method is not necessary and only the original q-Exponential log-likelihood function is used in parameter estimation.

For comparison purposes, the Weibull, q-Weibull and MEW distributions were considered. The three models presented inferior performance in both application examples when compared to the q-Exponential. We highlight that the q-Weibull and MEW models have three parameters each and are able to account for bathtub-shaped hazard rates. The q-Weibull can also accommodate unimodal hazard rates. Despite the increased flexibility of these Weibull-based models, the q-Exponential distribution presented the best results. These facts evidence the suitability of the q-Exponential in fitting reliability-related data sets and indicates that it is an alternative to be considered in reliability data analyses.

6.1. Example 1: TBFs of a machining center

The first application example is taken from Dai et al. [9]; it is related to the TBFs of a machining center. They are presented, in hours, in Table 4.

Table 5 presents the obtained parameter estimates by the original q-Exponential log-likelihood function and by the functions penalized by the Firth's and resample methods. The Nelder-Mead optimization method was used, and the initial parameters were set to $q = 1.2$ and $\eta = 500$ for original and corrected functions. The initial values of the

Table 4
TBFs of the machining center (in hours) – Example 1.

1	176.00	**5**	45.00	**9**	510.00	**13**	32.00	**17**	478.00
2	248.00	**6**	39.00	**10**	120.00	**14**	50.00	**18**	353.00
3	10.50	**7**	209.33	**11**	224.00	**15**	138.50	**19**	348.00
4	472.00	**8**	261.25	**12**	267.50	**16**	398.00	**20**	137.06

Table 5
Parameter estimates obtained by the Nelder-Mead optimization method for $\boldsymbol{q}$ and $\boldsymbol{\eta}$ and log-likelihood values – Example 1.

Method	$\hat{q}$	$\hat{\eta}$	Log-likelihood value
Original	−16,702,228.00	8,518,137,513.00	−124.69
Firth	0.19	505.28	−129.33
Resample	−7,850,909.00	4,003,964,375.00	−124.68

parameters have to result in valid arguments to the ln(·) functions of the q-Exponential log-likelihood; otherwise, the Nelder-Mead method would fail in the beginning of the iterative process.

The large parameter estimates related to the original approach indicates the presence of the monotone behavior of the log-likelihood function. This is corroborated by the log-likelihood value near to the theoretical limit in Eq. (14) for this specific data set, which is − 124.6882.

We note that Firth's method produced well behaved estimates, although it provided the worst log-likelihood value. This result is rather expected, given that the correction methods penalize the original log-likelihood function, i.e., they worsen the performance in terms of the objective function value. On the other hand, the original function and the function corrected by the resample method obtained large estimates in absolute value and their corresponding log-likelihood values present a slight difference.

Indeed, the large estimates provided by the resample method along with the slight superior log-likelihood approximate to the theoretical limit − 124.6882 suggest it fails to tackle such a problem for this example, which involves a small sample ($n = 20$) that is rather common in reliability data analyses. These results are in accordance with the findings of the performed Monte Carlo simulations in Section 6 for both original and resample methods and for small sample sizes. In addition to the point estimates provided by each approach, we have computed the associated 95% percentile confidence interval for q and η by means of a non-parametric bootstrap with 1000 replications. They are presented in Table 6. We observe that the interval estimates related to the original log-likelihood function and the function corrected by the resample method are rather noninformative, given their large widths and extremely large bounds. On the other hand, the log-likelihood function penalized by Firth's method produced well-behaved interval estimates with reasonable widths, lower and upper bounds.

For performance comparison, parameter estimates were obtained for the Weibull ($\hat{\beta} = 1.34$, $\hat{\eta} = 244.65$), q-Weibull ($\hat{q} = 0.74$, $\hat{\beta} = 0.04$, $\hat{\eta} = 122.90$) and MEW ($\hat{\beta} = 0.16$, $\hat{\lambda} = 0.55$, $\hat{\eta} = 11.88$) models. The Weibull distribution indicates an increasing hazard rate function and the other two distributions point a bathtub-shaped hazard rate. The following initial estimates for the Nelder-Mead optimization method were used: Weibull ($\hat{\beta} = 1.50$, $\hat{\eta} = 150$), q-Weibull ($\hat{q} = 1.20$, $\hat{\beta} = 50.00$, $\hat{\eta} = 100.00$) and MEW ($\hat{\beta} = 1.00$, $\hat{\lambda} = 0.50$, $\hat{\eta} = 10.00$).

Fig. 5 shows the curves of the expected number of failures vs. time obtained with the estimates of the original and corrected q-Exponential functions, the Weibull, the q-Weibull and MEW models. To generate these curves, we performed a Monte Carlo simulation with 10,000

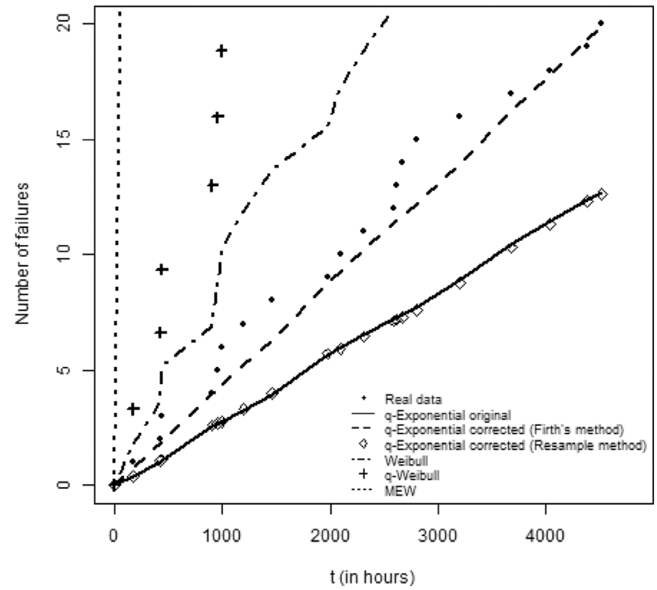

Fig. 5. Expected number of failures of original and corrected q-Exponential models, Weibull, q-Weibull and MEW distributions compared to real data – Example 1.

replications. By visual analysis, the curve that is the closest to empirical data (the real number of failures vs. time) is associated with Firth's penalization.

The expected curves provided by the original and resample-based approaches are almost coincident and are both bellow the expected curve related to Firth's penalization. The Weibull, q-Weibull and MEW models produced the worst results, i.e. the expected number of failures were far from real data. These three distributions overestimate the real number of failures, which may result in unnecessary allocation of resources for maintenance activities if their expected number of failures were used as guidelines. In this way, the corrected q-Exponential model is a more reliable approach in this case, as it enables a more accurate preparedness (e.g., spare parts and maintenance crew availability) to handle or avoid failure occurrences. These results are confirmed by the MAE of each curve of the expected number of failures; they are reported in Table 7.

Indeed, we note the q-Exponential penalized by Firth's method presented the lowest error and, consequently, the better performance. Instead, the Weibull function had the worst result with the highest error. In addition, the original and corrected (resample method) q-Exponential had very similar MAEs, as expected by visual inspection of Fig. 5.

In order to check the adjustment ability of the referred models – q-Exponential with Firth's correction, Weibull, q-Weibull and MEW –, the K-S Boot was used. The computed p-values were, respectively, 0.4775, 0.4395, 0.0080 and 0.0009. The small p-values provided by the q-Weibull and MEW distributions are evidences that they are not appropriate to adjust the considered TBFs. On the other hand, despite the null

Table 6
95% percentile confidence intervals based on non-parametric bootstrap – Example 1.

Method	Parameter	Lower bound	Upper bound	Interval width
Original	q	−20,123,821	−16,703,917	3419,904
	η	8454,229,751	8543,845,429	89,615,678
Firth	q	−3.89	0.98	2.91
	η	212.39	1388.88	1176.48
Resample	q	−8850,709.00	−5850,706	3000,003
	η	4677,485,429	5009,764,375	332,278,946

Table 7
MAE for the q-Exponential, Weibull, q-Weibull and MEW expected number of failures compared to real data – Example 1.

Distribution	MAE
Original q-Exponential	5.24
q-Exponential (Firth's method)	1.93
q-Exponential (Resample method)	5.26
Weibull	6.06
q-Weibull	22.40
MEW	821.23

Table 8
TBFs of the MRI scanner (in days) – Example 2.

1	99	11	12	21	8	31	19	41	19	51	18	61	47
2	38	12	13	22	26	32	47	42	10	52	3	62	26
3	109	13	40	23	98	33	14	43	17	53	46	63	87
4	10	14	6	24	11	34	53	44	4	54	17	64	6
5	35	15	78	25	87	35	14	45	54	55	7	65	13
6	42	16	77	26	11	36	35	46	26	56	75		
7	31	17	24	27	54	37	73	47	135	57	58		
8	18	18	66	28	22	38	18	48	44	58	102		
9	53	19	25	29	13	39	38	49	59	59	6		
10	3	20	4	30	54	40	140	50	11	60	53		

hypothesis could not be rejected for the q-Exponential corrected by Firth's method and Weibull cases, the greater p-value provided by the former indicates that it provides a better fit. Finally, Dai et al. [9] claim the TBFs in Table 4 follow the Weibull distribution. However, given the results of this section, the q-Exponential model with Firth's correction outperformed the Weibull-based models and places itself as a valid alternative for data modeling of the considered system, which is in the wear-out phase of its life cycle.

6.2. *Example 2: TBFs of an MRI scanner*

The second application example involves the TBFs of an MRI scanner [36], which are shown in Table 8.

From the Nelder-Mead optimization method using $q = 1.2$ and $\eta = 500$ days as initial values for the original log-likelihood function, $\hat{q} = 0.71$ and $\hat{\eta} = 60.48$ days were provided as point estimates with -296.06 as log-likelihood. As $\hat{q}$ and $\hat{\eta}$ are not large, this is an indication that the monotone likelihood problem is not present. Additionally, the theoretical limit in Eq. (14) for the q-Exponential log-likelihood, for this sample, is -321.21, which is less than the obtained -296.06. Therefore, a correction method is not required for the adjustment of the q-Exponential distribution to these TBFs.

As commented in Section 2, as the sample size and q increase, the probability of observing a monotone likelihood decreases. In this application example, $n = 65$ (greater than $n = 20$ of Example 1 in Section 6.1) and $\hat{q} = 0.71$ may have led to the absence of the monotone likelihood problem.

The 95% percentile bootstrap confidence intervals for q and η are shown in Table 9. As in Example 1, they were obtained using a non-parametric bootstrap method with 1000 replications. The point estimates of each of the bootstrap samples were provided by the Nelder-Mead optimization method using the original q-Exponential log-likelihood as objective function with the same initial points above-mentioned. The plausible bounds and widths of the confidence intervals is an additional evidence that the monotone behavior is not present.

Parameter estimates were also obtained for the Weibull ($\hat{\beta} = 1.20$, $\hat{\eta} = 40.91$), q-Weibull ($\hat{q} = 1.99$, $\hat{\beta} = 397.60$, $\hat{\eta} = 24.16$) and MEW ($\hat{\beta} = 0.87$, $\hat{\lambda} = 0.01$, $\hat{\eta} = 70.36$) distributions. The Weibull, q-Weibull and MEW models indicate, respectively, hazard rates with increasing, unimodal and bathtub-shaped behaviors. The following initial estimates for the Nelder-Mead optimization method were used: Weibull ($\hat{\beta} = 1.50$, $\hat{\eta} = 100$), q-Weibull ($\hat{q} = 1.20$, $\hat{\beta} = 500.00$, $\hat{\eta} = 50.00$) and MEW ($\hat{\beta} = 1.00$, $\hat{\lambda} = 0.70$, $\hat{\eta} = 65.00$).

To validate each of the estimated models, we performed a Monte

Table 9
95% percentile confidence intervals based on non-parametric bootstrap an on the original q-Exponential log-likelihood function – Example 2.

Parameter	Lower bound	Upper bound	Interval Width
q	−0.29	0.93	1.22
η	39.44	159.53	120.09

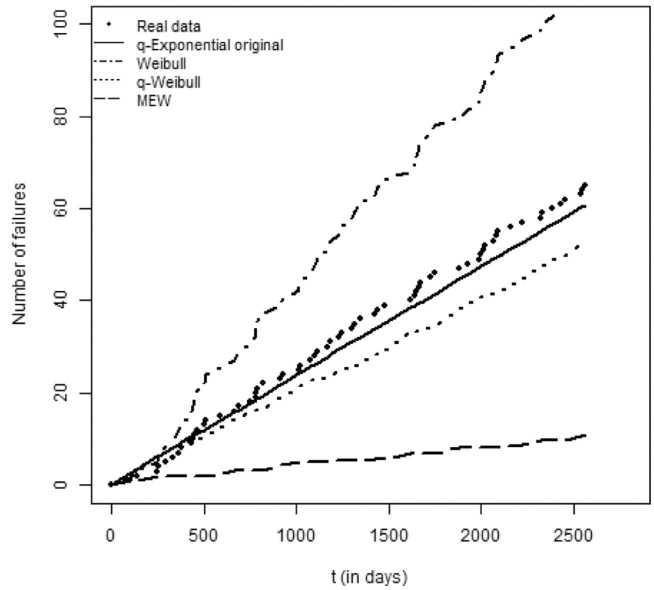

Fig. 6. Expected number of failures of original and corrected q-Exponential models, Weibull, q-Weibull and MEW distributions compared to real data – Example 2.

Table 10
MAE for the q-Exponential, Weibull, q-Weibull and MEW expected number of failures compared to real data – Example 2.

Distribution	MAE
Original q-Exponential	3.47
Weibull	21.85
q-Weibull	5.54
MEW	28.12

Carlo simulation with 10,000 replications to obtain the expected number of failures up to each real failure time. These values were used to plot the curves in Fig. 6 and to calculate the MAE values in Table 10. The KS-boot goodness-of-fit test provided 0.5034, 0.0412, 0.0500 and 0.0009 as p-values for the q-Exponential, Weibull, q-Weibull and MEW models, respectively. As indicated by the corresponding p-value, the q-Exponential distribution is suitable to model the probabilistic behavior of the TBFs in Table 8. On the other hand, the small p-values related to the Weibull-based distributions lead to the rejection of the null hypothesis, which means that they are not adequate to be used as probabilistic model of the MRI scanner TBFs.

7. Conclusion

To enable the use of the q-Exponential distribution in the context of reliability data modeling related to systems in the wear-out phase, we have: (i) identified and analyzed the q-Exponential monotone likelihood problem; (ii) adapted Firth's and resample correction methods to the q-Exponential case; (iii) devised a methodology to support the reliability analyst in the application of this distribution and also in the validation of the resulting models.

The performed numerical experiments showed that the estimates provided by the original q-Exponential log-likelihood when $q < 1$ are very poor for small and medium-sized samples, which are rather common in reliability data analysis. In addition, the simulations also showed that the function penalized by the Firth's method provided good parameter estimates even for small samples. This method was also superior when compared to the resample method.

The original and corrected functions were also applied to an example involving 20 TBFs of a machining center. As in the numerical experiments, the q-Exponential model using Firth's penalization provided the best results. They can be visually verified by the curves of expected number of failures vs. time when compared to empirical failure times and by the smallest MAE value. The corresponding interval estimates for q and η presented the highest precision, as they had the smallest widths defined by reasonable lower and upper bounds in terms of magnitude. Additionally, the Weibull, q-Weibull and MEW distributions were considered to model the TBFs and again the q-Exponential with Firth's correction presented superior performance: lower MAE for expected vs. empirical number of failures and higher p-value for the K-S Boot test.

Moreover, the original q-Exponential approach was applied to a second example related to the 65 TBFs of an MRI scanner, given that in this case the monotone likelihood problem had not been observed. Again, the Weibull q-Weibull and MEW models presented inferior performance. Actually, these three distributions were rejected as suitable to model the stochastic behavior of the MRI scanner TBFs. Thus, these two application examples evidence that, in some cases, the q-Exponential probability model may be an alternative to the Weibull-based distributions.

Finally, we expect that the methodology here proposed, involving the correction by Firth's method when the monotone likelihood problem is identified or the original approach otherwise, the interval estimates for the parameters and the validation of estimated models based on the cumulative expected number of failures and K-S Boot test, will make the use of the q-Exponential distribution more viable in reliability data analyses of systems that are in the wear-out phase. Thus, besides Weibull and others, there will be an additional valid alternative to model data sets in this situation.

In the future, we plan to explore other alternatives to solve the monotone likelihood problem of the q-Exponential model, apply the penalized q-Exponential log-likelihood by the Firth's method in other type of data such as censored data in the wear-out phase and we intend to consider other states of a system after the repair besides the “as good as new” state. The proposed methodology could be part of comprehensive reliability frameworks such as competing risks [32,50], optimal system design [24], accelerated life tests [17]. Also, Bayesian approaches for the estimation of the q-Exponential parameters are worth of investigation to account for extremely limited data [11,48].

In this research, we have focused on systems in the wear-out phase that present monotone increasing hazard rates, but lifetime distributions involving three or more parameters that are able to model bathtub-shaped hazard rates are also extensively studied and applied in reliability engineering [3], for example: exponentiated Weibull [33], MEW [23,46], generalized modified Weibull [6], exponentiated Nadarajah-Haghighi [22], q-Weibull [2,26,47], finite support model [21], 4-parameter Perks' [49], additive modified Weibull [18]. Specifically, the q-Weibull distribution has three parameters (η, β and q) and it is a generalization of the q-Exponential model (for $\beta = 1$, the q-Weibull becomes the q-Exponential distribution). In this sense, natural extensions of the present work would be: (i) the investigation of the q-Weibull likelihood function to verify whether the monotone likelihood problem is present and, if identified, (ii) the adaptation of the methodology here proposed.

CRediT authorship contribution statement

Ana Cláudia Souza Vidal de Negreiros: Methodology, Software, Formal analysis, Validation, Investigation, Writing - original draft, Visualization. **Isis Didier Lins:** Conceptualization, Methodology, Software, Writing - review & editing, Visualization. **Márcio José das Chagas Moura:** Methodology, Writing - review & editing. **Enrique López Droguett:** Supervision, Writing - review & editing.

Declaration of Competing Interest

None.

Acknowledgments

We thank the editor and the anonymous reviewers for the constructive comments that helped us to improve paper quality. This study was financed in part by the Coordenação de Aperfeiçoamento de Pessoal de Nível Superior - Brasil (CAPES) - Finance Code 001. The second and third authors also thank CNPq for the financial support in the form of research grants.

Supplementary materials

Supplementary material associated with this article can be found, in the online version, at doi:10.1016/j.ress.2019.106787.

References

[1] Almalki SJ, Nadarajah S. Modifications of the Weibull distribution: a review. Reliab Eng Syst Saf 2014;124:32–55. v.
[2] Assis EM, Borges EP, de Melo SABV. Generalized q-Weibull model and the bathtub curve. Int J Qual Reliab Manag 2013;30:720–36. v.
[3] Baker R. New survival distributions that quantify the gain from eliminating flawed componentes. Reliab Eng Syst Saf 2019;185:493–501. v.
[4] Bryson MC, Johnson ME. The incidence of monotone likelihood in the Cox model. Technometrics 1981;23(4):381–3.
[5] Campo MG, Ferri GL, Roston GB. q-Exponential distribution in time correlation function of water hydrogen bonds. Braz J Phys 2009;39(2). vn.
[6] Carrasco JMF, Ortega EMM, Cordeiro GM. A generalized modified Weibull distribution for lifetime modeling. Comput Stat Data Anal 2008;53:450–62. v.
[7] Cox DR. Regression models and life tables (with discussion). J Roy Stat Soc Ser B 1972;34:187–202.
[8] Cribari-Neto F, Frery AC, Silva MF. Improved estimation of clutter properties in speckled imagery. Comput Stat Data Analisis 2002;40:801–24.
[9] Dai Y, Zhoua Y, Jiab Y. Distribution of time between failures of machining center based on type i censored data. Reliab Eng Syst Saf 2003;79:377–9. v.
[10] Davison AC, Hinkley DV. Bootstrap methods and their applications. New York: Cambridge University Press; 1997.
[11] Droguett EL, Groen F, Mosleh A. The combined use of data and expert estimates in population variability analysis. Reliab Eng Syst Saf 2004;83:311–21. v.
[12] Efron B, Tibshirani RJ. An introduction to the bootstrap. New York: Chapman & Hall/CRC; 1994.
[13] Efron B. More efficient bootstrap computations. J Am Stat Assoc 1990;85:79–89.
[14] Firth D. Bias reduction of maximum likelihood estimates. Biometrika Trust 1993;80(1):27–38. vn.
[15] Fonseca R, Cribari-Neto F. Inference in a bimodal birnbaum-saunders model. Math Comput Simul 2018;v.146:134–59.
[16] Gupta A, Mukherjee B, Upadhyay SK. Weibull extension model: a Bayes study using Markov chain Monte Carlo simulation. Reliab Eng Syst Saf 2008;93:1434–43. v.
[17] Haghighi F. Optimal design of accelerated life tests for an extensionof the exponential distribution. Reliab Eng Syst Saf 2014;131:251–6. v.
[18] He B, Cui W, Du X. An additive modified Weibull distribution. Reliab Eng Syst Saf 2016;145:28–37. v.
[19] Heinze G, Schemper M. A solution to the problem of monotone likelihood in Cox regression. Int Biomet Soc 2001;57(1):114–9. vn.
[20] Jeffreys H. An invariant form for the prior probability in estimation problems. Proceedings of the Royal Society of London A: Mathematical, Physical and Engineering Sciences. 1946. p. 453–61.
[21] Jiang R. A new bathtub curve model with a finite support. Reliab Eng Syst Saf 2013;119:44–51. v.
[22] Lemonte AJ. A new exponential-type distribution with constant, decreasing, increasing, upside-down bathtub and bathtub-shaped failure rate function. Comput Stat Data Anal 2013;62:149–70. v.
[23] Lima VMC, Cribari-Neto F. Penalized maximum likelihood estimation in the modified extended Weibull distribution. Commun Stat Simul Comput 2019;48(2):334–49. vn.
[24] Lins ID, Droguett EL. Redundancy allocation problems considering systems with imperfect repairs using multi-objective genetic algorithms and discrete event simulation. Simul Modell Pract Theory 2011;19(1):362–81. vn.
[25] Lins ID, Droguett EL, Moura MC, Zio E, Jacinto CM. Computing confidence and prediction intervals of industrial equipment degradation by bootstrapped support vector regression. Reliab Eng Syst Saf 2015;137:120–8. v.
[26] Lins ID, Moura MC, Droguett EL, Corrêa TL. Combining generalized renewal processes with non-extensive entropy-based q-Distributions for reliability applications. Entropy 2018;20:223.
[27] Loughin TM. On the bootstrap and monotone likelihood in the Cox proportional hazards regression model. Lifetime Data Anal 1998;4(4):393–403. vn.
[28] Ludescher J, Bunde A. Universal behavior of the interoccurrence times between

losses in financial markets: independence of the time resolution. Am Phys Soc Phys Rev E 2014;90:062809.
[29] Malacarne LC, Mendes RS, Lenzi EK. *q*-exponential distribution in urban agglomeration. Phys Rev E Stat Nonlinear Soft Matter Phys 2002;65(1):1–3. vn.
[30] Manshour P, Anvari M, Reinke N, Sahim M, TABAR MRR. Interoccurrence time statistics in fully-developed turbulence. Sci Rep 2016;6:27452.
[31] Modarres M, Kaminskiy M, Krivtsov V. Reliability engineering and risk analisys 3. CRC Press; 2016. p. 522. editor.
[32] Moura MC, Droguett EL, Firmino P, Ferreira R. A competing risk model for dependent and imperfect condition-based preventive and corrective maintenances. J Risk Reliab 2014;228:590–605. v.
[33] Mudholkar GS, Srivastava DK. Exponentiated Weibull family for analyzing bathtub failure-rate data. IEEE Trans Reliab 1993;42:299–302. v.
[34] Nelder JA, Mead R. A simplex method for function minimization. Comput J 1965;7:308–13. v.
[35] Patriota AGA. *q*-Exponential regression model. Sankhyā. Ind J Stat 2012 2012;74–B:149–70. Part 1.
[36] Pereira DM. Determining the optimal replacement of complex medical equipment by means of obsolescence analysis (in Portuguese), 77 p. Brazil: Master Thesis – Universidade Federal de Pernambuco/UFPE; 2017.
[37] Pianto DM, Cribari-Neto F. Dealing with monotone likelihood in a model for speckled data. Comput Stat Data Anal 2011;55(3):1394–409. vn.
[38] R Development Core Team.. A language and enviroment for statistical computing. Vienna, Austria: R Foundation for Statistical Computing; 2010. ISBN: 3-900051-07-0, URL http://www.R-project.org.
[39] Ross SM. Simulation. 5th edition San Diego: Academic Press; 2012.
[40] Sales Filho RLM. A novel *q*-Exponential based stress-strength reliability model and applications to fatigue life with extereme values 127. Doutorado - Universidade Federal de Pernambuco/UFPE; 2016.
[41] Sales Filho R, Droguett EL, Lins ID, Moura MC, Azevedo R. Stress-strength reliability estimation based on the *q*-Exponential distribution. Qual Reliab Eng Int 2016;4:51. v.
[42] Schneider KA. Large and finite sample properties of a maximum-likelihood estimator for multiplicity of infection. PLoS One 2018.
[43] Stute W, Manteiga WG, Quindimil MP. Bootstrap based goodness-of-fit tests. Metrika 1993;40:243–56. v.
[44] Tsallis C. Introduction to nonextensive statistical mechanics: approaching a complex world 389p. New York: Springer; 2009.
[45] Tsallis C. Possible generalization of Boltzmann-Gibbs statistics. J Stat Phys 1988:479–87.
[46] Xie M, Tang Y, Hog TN. A modified Weibull extension with bathtub-shaped failure rate function. Reliab Eng Syst Saf 2002;76(3):279–85. vn.
[47] Xu M, Droguett EL, Lins ID, Moura MC. On the q-Weibull distribution for reliability applications: an adaptative hybrid artificial bee colony algorithm for parameter estimation. Reliab Eng Syst Saf 2017;158.
[48] Yañez M, Joglar F, Modarres M. Generalized renewal process for analysis of repairable systems with limited failure experience. Reliab Eng Syst Saf 2002;77:167–80. v.
[49] Zeng H, Lan T, Chen Q. Five and four-parameter lifetime distributions for Bathtub-shaped failure rate using Perks mortality equation. Reliab Eng Syst Saf 2016;152:307–15. v.
[50] Zhang F, Shi Y, Wang R. Geometry of the *q*-Exponential distribution with dependent competing risks and accelerated life testing. Physica A 2017;468:552–65. v.