Correction on the Log-Likelihood of the q-Exponential Distribution for use in the Reliability Context

Ana Cláudia Souza Vidal de Negreiros, Isis Didier Lins, Romero Luiz Mendonça Sales Filho,
Márcio das Chagas Moura, Enrique Lopez Droguett

Universidade Federal de Pernambuco, Recife, Brasil

This work involves the use of the q-Exponential distribution in the reliability area. The q-Exponential distribution has two parameters (q and η) and it stems from Tsallis' non-extensive entropy. This distribution has more flexibility regarding its decay for the Probability Density Function (PDF) curve. Besides that, the q-Exponential does not have the limitation of a constant hazard rate like the Exponential one, thus allowing the modeling of either system improvement $(1 < q < 2)$ or degradation $(q < 1)$. The q-Exponential distribution can model very well data sets with extreme values, which corresponds to improvement phase of systems. This feature is interesting on the reliability context because many equipment can work for long time until the first failure. In this case, the q-Exponential presents power law characteristic. However, when data sets are related to the degradation phase of systems, the application of the q-Exponential becomes difficult due to convergence problems in the estimation process via maximum likelihood method. This problem is called "monotone likelihood" and is associated with a monotone behavior of the log-likelihood function. In order to correct this problem the Firth's penalization method is applied on the q-Exponential log-likelihood function. The obtained results shows that the penalized function can model data well even for small samples. The Nelder-Mead numerical method was used to estimate the parameters. Comparisons were made between the results obtained with the q-Exponential log-likelihood function original and the q-Exponential log-likelihood function penalized.

1. INTRODUCTION

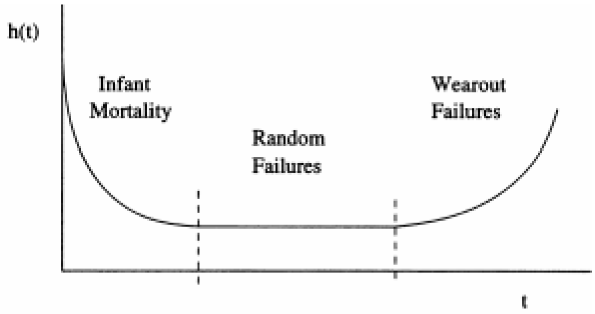
The q-Exponential distribution has two parameters (q and η), q is the shape parameter and η is the scale parameter. According to [1] it is obtained by maximizing the non-extensive entropy under appropriate constraints. When compared with the Exponential distribution that has just one parameter (η). The q-Exponential distribution has more flexibility regarding its decay for the Probability Density Function (PDF) curve. Indeed, the Exponential probability distribution is a special case Another feature of this distribution is that it does not have the limitation of a constant hazard rate, thus allowing the modeling of either system improvement $(1 < q < 2)$ or degradation $(q < 1)$.

For the reliability, in general, for a given sample with values that have great order of magnitude, the q-Exponential distribution is expected to adjust the data well. In cases like these, the parameter q lies within the interval (1, 2), when parameter q is in this interval it leads to the power law asymptotic behavior [2]. Actually, the q-Exponential distribution can modelling the three phases of the bathtub curve. The classic bathtub curve against time has three distinct periods: decreasing failure rate for infant mortality; constant failure rate for useful life; and increasing failure rate (without bound) for wear-out. The Figure 1 bellow shows a bathtub curve.

It has been observed in previous works as in [3] that, when $q < 1$, the techniques used to obtain the maximum likelihood estimates either provide poor results or fail to converge. In those cases, the q-Exponential log-likelihood function seems to be monotonically increasing, which renders the estimation task theoretically impossible.

A function is characterized as monotone when it preserves the relation of order, in the case crescent, the function increases measure the parameters also increase. According to [4], the called "monotone likelihood" is a situation where the log-likelihood obtain its maximum for infinite parameter values. [5] affirm that the monotone likelihood is noted in the fitting process of a Cox model if the likelihood converges to a finite value while at least one parameter estimate diverges to $\pm\infty$.

Figure 1: The classical bathtub curve



The monotone likelihood has been verified in many cases, in his work [6] studied about the q-Exponential distribution and he found indicators that this distribution present the behavior called monotone likelihood. He found this characteristic while tried estimate the parameters of this distribution to the case which $q < 1$. He perceived that when $q < 1$ there are problems in the estimate of the parameters, resulting in a poor modeling of the data set. Furthermore, the estimates in this case yields large confidence intervals.

Some authors have been worked in the problem of the monotone likelihood. There are some methods in the literature to solve this problem, the ways to correct this problem are different from author to author. [7] developed a method, his method is an approach to bias reduction which does not depend of finiteness of $\hat{\theta}$. [8] worked with bootstrap approach to correct this problem. [5] proposed a method which is an adaptation of a procedure by [7]. [9] suggested a method to correct the likelihood function also based on resampling.

The q-Exponential distribution, as others q-distributions, have been applied to an large variety of problems in varied research areas including the field of complex systems [2]. This distribution has been studied by some authors, as [2], who brings in his work a resume of basic properties of the q-Exponential, by [6], who verified that there was indications that the q-Exponential likelihood function might to be monotone. [10] showed that the population of a country is well described by a q-Exponential distribution with $q = 1.7$. [11] verified that the temporal correlation function of hydrogen bonds has a q-Exponential behavior. [12] proposed a possible way to understanding the ubiquity of the q-Exponential distribution in nature.

The q-Exponential distribution has been fairly utilized in the reliability area, this because it has the feature of modeling very well data sets with big values [2]. This feature is interesting on the reliability context because so many machines or equipments has a big usefull life, in other words, they can work for long time until the first failure. However, the problem verified in the q-Exponential distribution happen in the degradation phase of systems, in other words, the parameter q assume values less than one when the equipment is found in their final phase or degradation phase (last phase of bathtub curve). And it is important to reliability be able to modelling data coming of a system in this conditions.

The general objective this work is to use a method to correct the q-Exponential distribution log-likelihood function so as to obtain good estimates for its parameters. The proposed method will be applied to model reliability-related data of engineered equipment.

2. DESCRIPTION OF WORK REALIZED

Regarding its finality, this work is characterized as an applied research, because it is conducted to solve a specific problem of literature and to a practical methodology implementation. It presents an application of a method to correct the q-Exponential likelihood function in order to apply the results in practical problems.

Besides, this project can be classified as qualitative and quantitative. It is qualitative because use the literature review to understand and analyze the specific problem treated in this work. And, it is quantitative because uses computational programs and statistical methods to find a solution to the problem of this research and to apply and analyze the solution obtained.

This research can be sub-divided in the following steps:

- Step 1: Initially, it was made review about the topics that was studied in this work. The topics are: the features of the q-Exponential distribution; the problem that is verified in this distribution (monotone likelihood); the methods that could solve this problem; the numerical methods of maximizing functions that could be used for maximizing the q-Exponential log-likelihood function.

- Step 2: Then, after to study the Firth's method, it was chosen to be applied on the q-Exponential distribution.

- Step 3: Next, the simulations for test the corrected function was made in the software R. The results are showed in the Table 1 and 2.

3. THEORETICAL BACKGROUND

3.1. q-Exponential Distribution

The q-Exponential distribution has the following probability density function (PDF):

$$f_q(t) = \frac{(2-q)\left[1-\frac{(1-q)t}{\eta}\right]^{1/1-q}}{\eta}, \text{ for } t \geq 0, \tag{1}$$

$q < 1$ or $q < 2$ and $\eta > 0$.

The parameter q determines the density shape and is known as entropic index, the parameter η is the scale parameter. In the limit $q \to 1$, Eq. (1) recovers the usual Exponential distribution. When $q < 1$, Eq. (1) has a finite value for any finite real value of t. When $q > 1$, Eq. (1) presents power law characteristic, a asymptotic behavior.

Moreover, we will have different results for the support t, depending on the value of the entropic index:

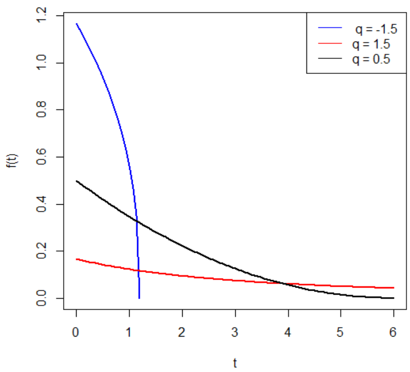
$$t \in \begin{cases} (0;\infty], & 1 < q < 2 \\ \left[0, \frac{\eta}{1-q}\right], q < 1 \end{cases} . \qquad (2)$$

Figure 1 shows the q-Exponential PDF for two possible values of q and η constant, illustrating the behavior.

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

$$F_q(t) = 1 - \left[1 - (1-q)\left(\frac{t}{\eta}\right)\right]^{\frac{2-q}{1-q}}, t \geq 0. \qquad (3)$$

Figure 2: q-Exponencial PDF for η=3 and some possible values of q.



By definition, the hazard rate is $h_q(t) = \frac{f_q(t)}{R_{q(t)}}$, where $R_{q(t)} = 1 - F_q(t)$.

Thus, we can write:

$$h_{q(t)} = \frac{\frac{(2-q)\left[1-\frac{(1-q)t}{\eta}\right]^{\frac{1}{1-q}}}{\eta}}{\left[1-\frac{(1-q)t}{\eta}\right]^{\frac{2-q}{1-q}}} = \frac{(2-q)}{\eta}\left[1 - \frac{(1-q)t}{\eta}\right]^{\frac{q-1}{1-q}}. \qquad (4)$$

Differently from the Exponential distribution, the q-Exponential hazard rate is not constant. Actually, this is an important characteristic of the q-Exponential distribution, especially in the reliability context.

Figure 3: q-Exponencial CDF with $\boldsymbol{\eta} = \mathbf{3}$ and $\mathbf{q} = -\mathbf{1.5}$.

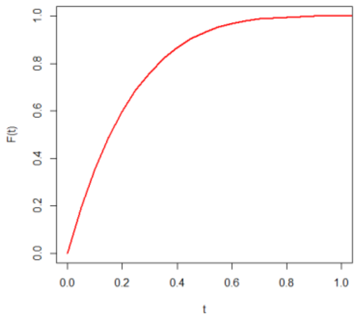
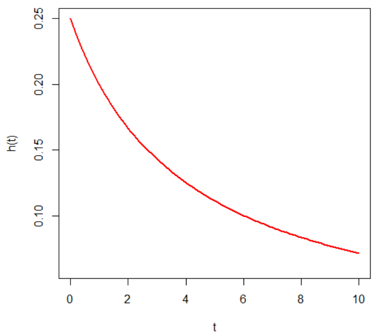


Figure 4: q-Exponencial h(t) with η=2 and $\mathbf{q} = \mathbf{1.5}$.



3.2. *Estimation of Parameters*

3.2.1. *Maximum Likelihood Estimation*

There are some ways to estimate a parameter a probabilistic model of interest, but the maximum likelihood method is one of the most used techniques. Considering the uniparametric case, assume a random sample of the random variable X with size n: $X_1, X_2, \cdots, X_n$. It is symbolized the probability density function (PDF) of the random variable X, as $f(X|\theta)$, with $\theta \in \Theta$, where Θ is the parametric space. Thus, the likelihood function of Θ, for the considered sample, can be written as is presented by [13]:

$$L(\theta|X) = \prod_{i=1}^{n} f(x_i|\theta). \tag{5}$$

The value that maximizes the likelihood function is the Maximum Likelihood Estimator of θ. It is represented this value as $\hat{\theta} \in \Theta$.

In general, to get maximum likelihood estimates by maximizing the natural logarithm of the likelihood function is easier than by maximizing directly the likelihood function. In this way, it is determined the log-likelihood function as:

$$l(\theta|X) = \log[L(\theta|X)]. \tag{6}$$

Thus, are obtained the maximum likelihood estimate calculating the root of the derivative of the log-likelihood function, *i. e.*:

$$\frac{dl(\theta|X)}{d\theta} = 0. \tag{7}$$

In situations in which it is not possible to get the solution of log-likelihood function analytically, the solution of the equation above can be obtained by numerical procedures or by heuristics.

In the specific case of the q-Exponential distribution, the likelihood function is the described following. Let $\underline{x} = \{x_1, x_2, \cdots, x_n\}$ be a random sample of size n. Thus, the q-Exponential likelihood function is

$$L(x,q,\eta) = \prod_{i=1}^{n} \frac{(2-q)\left[1-\frac{(1-q)x_i}{\eta}\right]^{1/1-q}}{\eta} = \frac{(2-q)^n}{\eta^n} \prod_{i-1}^{n} \left[1 - \frac{(1-q)x_i}{\eta}\right]^{1/1-q} \tag{8}$$

and the corresponding log-likelihood function is

$$l(t,q,\eta) = \ln\left(\frac{(2-q)^n}{\eta^n} \prod_{i=1}^{n} \left[1 - \frac{(1-q)t_i}{\eta}\right]^{1/1-q}\right) =$$

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

The transformation above can be used because it is a monotonic transformation, in other words, it is a transformation that preserves the order of the numbers. Thus, the values that maximize Eq. (8) are the same that maximize Eq. (9).
To obtain the Maximum Likelihood Estimates (MLEs) of the parameters the log-likelihood function is maximized. This can be done by setting the first derivative of l which respect to each parameter to zero. The q-Exponential score equations are the following:

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

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

The Eq. (10) does not have a closed solution, because of that it is necessary to use non-linear maximization methods to obtain our parameters estimates. However, numerical maximization of the q-Exponential log-likelihood function may fail to converge and may yield poor parameter estimation.

3.3. Nelder-Mead

The Nelder–Mead method is a numerical approach frequently applied to nonlinear optimization, it is also known as Downhill Simplex method. This method is used to find the minimum or maximum of an objective function in a multi-dimensional space. It is a method fairly used in unconstrained optimization problem of a function of n variables. This numerical approach has been used in many studies with the aim of maximizing the log-likelihood function and to estimate the parameters of various probability distributions in many areas.

According to [14] the Nelder-Meade method present the following features:

- Ease of computational implementation;
- Calculations of the derivatives of the objective function are not necessary;
- Few evaluations of the objective function are necessary;
- The value of the objective function quickly decreases in the first iterations.

The Nelder-Mead uses the concept of a simplex, which is a special polynomium type with $n+1$ vertices in n dimensions.

Consider the problem of unconstrained minimization:

$$\min_{x \in \mathfrak{R}^n} f(x), \text{ where, } f: \mathfrak{R}^n \longrightarrow \mathfrak{R}.$$

In this work $f(x)$ is the negative of the q-Exponential log-likelihood.

In one iteration of this method, the $n+1$ vertices of the simplex, $x_1, x_2, \cdots, x_{n+1}$ belonging to $\mathfrak{R}^n$ are required according to the growth of the values of f, i.e:

$$f(x_1) \leq f(x_2) \leq \cdots \leq f(x_{n+1}).$$

Where x_{n+1} is the worst vertex and x_1 is the best vertex.

The repositioning of these vertices takes into consideration four coefficients:

- Reflection coefficient (ρ)
- Expansion coefficient (χ)
- Contraction coefficient (γ)
- Reduction coefficient (σ)

[15] explain that these coefficients must satisfy the following restrictions:

$\rho > 0, \chi > 1, 0 < \gamma < 1$ and $0 < \sigma < 1$.

The Nelder-Mead attempts to exchange the worst vertex of the simplex by one with better value. The new vertex is get by reflecting, expansion or contraction of the worst vertex along the line through this vertex and the centroid of the best n vertices. The worst vertex is substitute by a new vertex or the simplex is reduced around the better vertex, at each iteration.

Below is presented a set of steps that corresponding to an interaction of the Nelder-Mead algorithm [15]:

- Step 1 - Rank: Rank the $n + 1$ vertices:

$$f(x_1) \leq f(x_2) \leq \cdots \leq f(x_{n+1});$$

- Step 2- Centroid: Calculate the centroid of the n best vertices:

$$\bar{x} = \sum_{i=1}^{n} \frac{x_i}{n}.$$

- Step 3- Reflected vertex: Calculate the reflected vertex (x_r):

$$x_r = \bar{x} + \rho(\bar{x} - x_{n+1}).$$

If $f(x_1) \leq f(x_r) \leq f(x_n)$, then do $x_{n+1} = x_r$ and finalize the iteration.

- Step 4- Expansion: If $f(x_r) \leq f(x_1)$, calculate the expanded vertex (x_e):

$$x_e = \bar{x} + \chi(x_r - \bar{x}).$$

If $f(x_e) \leq f(x_r)$, then do $x_{n+1} = x_e$ and finalize the iteration, else $x_{n+1} = x_r$ and finalize the iteration.

- Step 5- Contraction: If $f(x_r) \geq f(x_n)$

5.1 External:

If $(x_n) \leq f(x_r) \leq \ f(x_{n+1})$, calculate the external contraction vertex (x_{ce}) :

$$x_{ce} = \bar{x} + \gamma(x_r - \bar{x}).$$

If $f(x_{ce}) \leq f(x_r)$, then do $x_{n+1} = x_{ce}$ and finalize the iteration, else go to step 6.

5.2 Internal:

If $f(x_n) \geq f(x_n)$, calculate the internal contraction vertex(x_{ci}):

$$x_{ci} = \bar{x} - \gamma(\bar{x} - x_{n+1}).$$

If $f(x_{ci}) \leq f(x_{n+1})$, then do $x_{n+1} = x_{ci}$ and finalize the iteration, else go to step 6.

- Step 6- Reduction: Calculate vectors $v_i = x_1 + \sigma(x_i - x_1)$, $i = 2, \cdots, n + 1$. The vertices (not ordered), for the next iteration are: $x_1, v_2, \cdots, v_{n+1}$.

[15] explain that given a tolerance Δ_{tol}, the following stop criterion takes into account the function value in the simplex vertices:

$$\sqrt{\sum_{i=1}^{n+1} \frac{\left(f(x_i) - f(\bar{x})\right)^2}{n}} < \Delta_{tol}.$$

3.4. Firth's Method

A method to apply a penalty to log-likelihood function in order to reduce the bias of the MLE was proposed by [7]. Actually, the idea behind his method is that since the parameter estimate may not exist it is safer to modify the estimating equations to correct for bias prior to estimation. Let $U^{*}(\theta)$ be the modified score function. For the canonical parameter of the exponential family model, the rth component of the modified score equation is given by

$$U_r^{*}(\theta) = U_r(\theta) + A_r(\theta), \tag{11}$$

in which $A_r(\theta)$ is the rth part of $A(\theta) = -I(\theta)B_1(\theta)/n, r = 1, \cdots, \dim(\theta)$. $B_1(\theta)$ is denoted here like the first order term in the bias expansion on the MLE: $B(\theta) = B_1(\theta)/n + B_2(\theta)/n + \cdots$.

Note that the method proposed by [7] was made to apply in the canonical exponential model. However, even in functions that are not members of this group was found that the applied of this penalty yields great results, as in [16], they applied this method in a bimodal Birbaun-Saunders model.

In the case of an exponential family in canonical form, the observed information (Fisher's information) does not depend on the data, and it follows that

$$A_r(\theta) = \frac{\partial}{\partial \theta_r}\left\{\frac{1}{2}\log|I(\theta)|\right\}. \tag{12}$$

For functions in the canonical exponential family, the correction of the likelihood function is applied as following

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

where the penalization term $|K|^{1/2}$ is the [17] invariant prior. Equivalently, estimation can be executed by maximizing

$$l^{*}(\theta|X) = l(\theta|X) + \frac{1}{2}|K|. \tag{13}$$

4. OBTAINED RESULTS

In this section it will be showed the application of the Firth's method to penalize the q-Exponential log-likelihood function, but to apply this correction it was necessary one modification, that will be presented in the following.

Under regularity conditions and to large samples, $\hat{\theta}\sim N_3(\theta, I(\theta)^{-1})$ approximately, where $I(\theta)$ is Fisher's (expected) information matrix:

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

The score function of the q-Exponential log-likelihood function is presented in Eq. (10). In general, is easier compute as $I(\theta) = E[J(\theta)]$, where $J(\theta) = -\partial^2 l(\theta)/\partial\theta\partial\theta^T$ is the observed information. For the q-Exponential model, we obtain

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

where,

$$J_{qq} = -\frac{n}{(2-q)^2} + \frac{2}{(1-q)^3}\sum_{i=1}^{n} \ln\left(1 - \frac{(1-q)t_i}{\eta}\right) + \frac{2}{(1-q)^2}\sum_{i=1}^{n} \frac{t_i}{\eta\left(1-\frac{(1-q)t_i}{\eta}\right)} + \frac{1}{(1-q)}\sum_{i=1}^{n} \frac{t_i^2}{\eta^2\left(1-\frac{(1-q)t_i}{\eta}\right)} \quad (14)$$

$$J_{\eta\eta} = \frac{n}{\eta^2} + \frac{1}{(1-q)}\sum_{i=1}^{n}\left(-\frac{2(1-q)t_i}{\eta^2\left(1-\frac{(1-q)t_i}{\eta}\right)} - \frac{t_i^2(1-q)^2}{\eta^4\left(1-\frac{(1-q)t_i}{\eta}\right)}\right) \quad (15)$$

$$J_{q\eta} = J_{\eta q} = \frac{1}{(1-q)^2}\sum_{i=1}^{n} \frac{(1-q)t_i}{\eta^2\left(1-\frac{(1-q)t_i}{\eta}\right)} + \frac{1}{(1-q)}\sum_{i=1}^{n} - \frac{t_i}{\eta^2\left(1-\frac{(1-q)t_i}{\eta}\right)} - \frac{t_i^2(1-q)}{\eta^4\left(1-\frac{(1-q)t_i}{\eta}\right)} \quad (16)$$

Therefore, how it was presented before, the original idea this method involves the utilization the matrix of the expected information (Firth's information), but in some cases it is not easy to obtain this measure. In this cases, it is common to use the matrix of the observed information, which could be seen as an approximation of the expected information.

Table 1 bellow present the results of a Monte Carlo simulation made with the original q-Exponential log-likelihood function and with the penalized q-Exponential log-likelihood function for η constant and three values of q. This simulation was made with the computational software R, it was utilized 10000 (ten thousands) replications for five sizes of sample (20, 50, 100, 500, 1000) and the numerical method utilized to do the maximizations was the Nelder-Mead.

By analyzing the results showed in the Table 1 is possible verify that was achieved a significant improvement with the penalized q-Exponential log-likelihood function, especially for small sizes of samples as 20 and 50 observations.

For $q = -20$ and $\eta = 5$ the original q-Exponential log-likelihood function did not yields any good estimate for the parameters. Meanwhile the bigger relative bias that the penalized function reached for the parameter q was -0.5851873 and the bigger relative bias for the parameter η was -0.5817365, *i.e.* the corrected function achieved a bias smaller than 60% for the smallest size of sample, this bias drops to less than 50% when the size of sample is of 1000 (one thousand) observations.

For $q = -2$ and $\eta = 5$ the original function did not produce good results for the sample sizes 20, 50 and 100 realizations. Actually, for this sample sizes the estimates are very bad. However, for this parameters values combination, starting of samples with 500 observations the original q-Exponential log-likelihood function starts to yields good estimates for the parameters. For the other hand, for this both true values of parameters, the corrected function continued to produce good estimates of the parameters for all sample sizes. But, is important make it clear that for great sample sizes (500 and 1000) the original function did produce a little bit better results that the penalized function.

For $q = 0.5$ e $\eta = 5$ the results for this combination of parameters were similar at the previous case. The original function did produce good estimates for the parameters just for samples with a great number of observations (100, 500 and 1000 realizations). And the penalized function did yield very good estimates for all the sample sizes, which means that the penalty in the q-Exponential log-likelihood function worked very well. The bigger relative bias verified with the corrected function was 0.9290058 for the parameter q and 0.5268211 for the parameter η. Meanwhile, the bigger relative bias that the original function produced was -9856444 for the parameter q and 5731969 for the parameter η.

Table 1: Comparisons with **q** constant

For $q = -20$ and $\eta = 5$	$\hat{q}$	Rel. Bias $\hat{q}$	$\hat{\eta}$	Rel. Bias $\hat{\eta}$
Original function, n=20	-141900356	7095017	31963831	6392765
Penalized function, n=20	-8.296254	-0.5851873	2.091318	-0.5817365
Original function, n=50	-50594075	2529703	11797129	2359425
Penalized function, n=50	-8.626906	-0.5686547	2.241547	-0.5516905
Original function, n=100	-22844008	1142199	5377958	1075591
Penalized function, n=100	-8.874138	-0.5562931	2.3236	-0.53528
Original function, n=500	-2213100	110654	525524.8	105104
Penalized function, n=500	-9.87779	-0.5061105	2.583265	-0.4833471
Original function, n=1000	-535200.5	26759.03	127240.5	25447.09
Penalized function, n=1000	-10.76152	-0.4619238	2.796611	-0.4406779
For $q = -2$ and $\eta = 5$	$\hat{q}$	Rel. Bias $\hat{q}$	$\hat{\eta}$	Rel. Bias $\hat{\eta}$
Original function, n=20	-98051761	49025880	146483739	29296747
Penalized function, n=20	-4.498968	1.249484	8.24302	0.648604
Original function, n=50	-16157679	8078839	25362689	5072537
Penalized function, n=50	-4.33278	1.16639	8.448655	0.6897309
Original function, n=100	-1600956	800476.8	2558381	511675.2
Penalized function, n=100	-4.279655	1.139828	8.536188	0.7072376
Original function, n=500	-2.839555	0.4197777	6.36099	0.272198
Penalized function, n=500	-3.618861	0.8094307	7.632871	0.5265743
Original function, n=1000	-3.231732	0.6158662	7.028491	0.4056982
Penalized function, n=1000	-3.58282	0.7914099	7.60109	0.520218
For $q = 0.5$ and $\eta = 5$	$\hat{q}$	Rel. Bias $\hat{q}$	$\hat{\eta}$	Rel. Bias $\hat{\eta}$
Original function, n=20	-4928221	-9856444	28659849	5731969
Penalized function, n=20	0.3891576	-0.2216849	7.634105	0.5268211
Original function, n=50	-24925.44	29591.83	147964.1	-24925.44
Penalized function, n=50	0.8546313	0.7092627	4.767566	-0.046486
Original function, n=100	0.3610203	-0.2779594	5.785671	0.1571343
Penalized function, n=100	0.9645029	0.9290058	4.414194	-0.117161
Original function, n=500	0.4720099	-0.05598024	5.160208	0.0320416
Penalized function, n=500	0.6298968	0.2597937	4.972481	-0.005503
Original function, n=1000	0.4781027	-0.04379465	5.146214	0.0292427
Penalized function, n=1000	0.4948682	-0.01026362	5.011411	0.002282109

Source: This research (2017)

Table 2 bellow present the results of a Monte Carlo simulation made with the original q-Exponential log-likelihood function and with the penalized q-Exponential log-likelihood function for q constant and three values of η. This simulation was made with the computational software R, it was utilized 10000 (ten thousands) replications for five sizes of sample (20, 50, 100, 500, 1000) and the numerical method utilized to do the maximizations was the Nelder-Mead.

For $q = -2$ and $\eta = 50$ the original function did not produce good results for the sample sizes 20, 50 and 100 realizations. For this sample sizes the estimates are very poor. However, for this parameters values combination, starting of samples with 500 observations the original q-Exponential log-likelihood function starts to yields good estimates for the parameters. For the other hand, for this both true values of parameters, the corrected function continued to produce good estimates of the parameters for all sample sizes.

For $q = -2$ and $\eta = 500$, for this combinations of parameters, once again the original q-Exponential log-likelihood function just did yields goods results for great size samples (500 and 1000 realizations). And once more the penalized function keep to maintain good results even for small size samples (20 and 50 observations).

For $q = -2$ and $\eta = 1000$, the corrected function remained consistent, *i.e.* the results obtained with this function were still good even for small samples, just as for the others combination of parameters. For other hand, the original function just starts to produce good results for samples with at least 500 realizations.

The results presented in Table 1 and Table 2 prove that the penalized function is consistent and effective even for small sizes samples since that it is possible to obtain good results in this cases. For other hand, the original function proved to be more effective for big sizes samples. However, when the parameter q increasing in module, the original function does not yields good results even for big sizes sample, as it is possible to see in the Table 1 For $q = -20$ and $\eta = 5$, others tests with values of the parameter q higher in module were made and the results were not good.

In practice, it is not viable to obtain samples with a big number of realizations, sometimes there are not financing to get this or just it is not possible, thus it is possible affirm that the original function is not a good choice, because it just showed goods results in specifics cases. Besides that, when the parameter q increasing in module, the original function does not produce goods results in no case.

Table 2: Comparison with $\boldsymbol{\eta}$ constant

For $q = -2$ and $\eta = 50$	$\hat{q}$	Rel. Bias $\hat{q}$	$\hat{\eta}$	Rel. Bias $\hat{\eta}$
Original function, n=20	-19770436	9885217	295654300	5913085
Penalized function, n=20	-3.67971	0.8398551	70.2032	0.404064
Original function, n=50	-3206861	1603430	50337273	1006744
Penalized function, n=50	-3.417721	0.7088603	69.99187	0.3998374
Original function, n=100	-309689.9	154844	4951017	99019.33
Penalized function, n=100	-3.333458	0.6667288	70.09704	0.4019409
Original function, n=500	-2.484186	0.2420929	57.70734	0.1541468
Penalized function, n=500	-3.180338	0.5901688	69.05484	0.3810969
Original function, n=1000	-2.432521	0.2162605	57.00197	0.1400393
Penalized function, n=1000	-3.038986	0.5194931	66.96272	0.3392544
For $q = -2$ and $\eta = 500$	$\hat{q}$	Rel. Bias $\hat{q}$	$\hat{\eta}$	Rel. Bias $\hat{\eta}$
Original function, n=20	-10294098	5147048	1540086511	3080172
Penalized function, n=20	-3.634623	0.8173114	695.8656	0.3917312
Original function, n=50	-1664069	832033.6	261186838	522372.7
Penalized function, n=50	-3.403819	0.7019093	697.7501	0.3955003
Original function, n=100	-159733	79865.48	25527798	51054.6
Penalized function, n=100	-3.322857	0.6614283	699.3281	0.3986562
Original function, n=500	-2.498385	0.2491925	579.27	0.15854
Penalized function, n=500	-3.238553	0.6192766	700.2667	0.4005333
Original function, n=1000	-2.32467	0.162335	552.0839	0.1041677
Penalized function, n=1000	-3.214191	0.6070956	698.7303	0.3974605
For $q = -2$ and $\eta = 1000$	$\hat{q}$	Rel. Bias $\hat{q}$	$\hat{\eta}$	Rel. Bias $\hat{\eta}$
Original function, n=20	-8970900	4485449	2686290200	2686289
Penalized function, n=20	-3.647204	0.8236021	1394.593	0.3945928
Original function, n=50	-1409583	704790.3	442743701	442742.7
Penalized function, n=50	-3.403975	0.7019877	1395.393	0.3953929
Original function, n=100	-134713.4	67355.71	43027614	43026.61
Penalized function, n=100	-3.316545	0.6582725	1396.465	0.3964647
Original function, n=500	-2.506325	0.2531626	1161.087	0.1610866
Penalized function, n=500	-3.228257	0.6141286	1397.159	0.3971589
Original function, n=1000	-2.305045	0.1525226	1097.621	0.09762107
Penalized function, n=1000	-3.212806	0.6064031	1397.085	0.3970854

Source: This research (2017)

5. CONCLUSIONS

This work proposed correct the q-Exponential distribution, which used to present a problem called "monotone likelihood". This problem was verified when $q < 1$, and in this situation was not possible to obtain good parameters estimates for this distribution.

Once the q-exponential distribution can modelling the three phases of the bathtub curve, it is very important be able to obtain good parameters estimates for this distribution.

In order to solve this problem, it was applied the Firth's method. This is a method applied directly in the log-likelihood function and uses the expected information (Fisher's information). However, for this work it was uses the observed information, once it is not possible to obtain the expected information for the q-Exponential log-likelihood function. Besides, the Firth's method was proposed for distributions that belongs to canonical exponential family, but previous works showed that it can also work for distributions that does not belongs to this group.

The correction applied in this distribution produced good results. Even with small samples, e.g., 20 observations, the penalized function showed acceptable estimates for q and η, the q-Exponential's parameters. Meanwhile, the original distribution just shows good results for big samples, e.g., 500 observations, which is unfeasible in practice. Besides, when the parameter q increases in absolute value, the original distribution does not yields goods results for any sample size.

In general, it is possible affirm that the penalization applied in the q-Exponential log-likelihood function represents a science progress, once with this change on the function becomes possible fit data sets in what $q < 1$. This is the case in which it has a crescent hazard rate, last phase of the bathtub curve.

REFERENCES

[1] TSALLIS, C. *Introduction to Nonextensive Statistical Mechanics: Approaching a Complex World*. New York: Springer, (2009).

[2] PICOLI JR, S.; MENDES, R. S.; MALACARNE, R. P. B. S. “q -Distributions in Complex Systems : a Brief Review”. Brazilian Journal of Physics, v. 1, n. 1, p. 468–474, (2009).

[3] SALES FILHO, R.; LOPES DROGUETT, E.; LINS, I.; MOURA, M. C.; AZEVEDO, R. “Stress-Strength Reliability Estimation Based On The q-Exponential Distribution”. Quality and Reliability Engineering International, v. 4, p. 51, (2016).

[4] PIANTO, D. M.; CRIBARI-NETO, F. “Dealing with Monotone Likelihood in a Model for Speckled Data”. Computational Statistics and Data Analysis, v. 55, n. 3, p. 1394–1409, (2011).

[5] HEINZE, G.; SCHEMPER, M. “A Solution to the Problem of Monotone Likelihood in Cox Regression”. International Biometric Society, v. 57, n. 1, p. 114–119, (2001).

[6] SALES FILHO, R. L. M. “A Novel q -Exponential Based Stress-Strength a Novel Reliability Model and Applications to Fatigue Life with Extereme Values”. Universidade Federal de Pernambuco, (2016).

[7] FIRTH, D. “Bias Reduction of Maximum Likelihood Estimates”. Biometrika Trust, v. 80, n. 1, p. 27–38, (1993).

[8] LOUGHIN, T. M. “On the Bootstrap and Monotone Likelihood in the Cox Proportional Hazards Regression Model”. Lifetime data analysis, v. 4, n. 4, p. 393–403, (1998).

[9] CRIBARI-NETO, F., FRERY, A. C., SILVA, M. F. “Improved Estimation of Clutter Properties in Speckled Imagery”. Computational Statistics & Data Analisis 40, 801-824. (2002).

[10] MALACARNE, L. C.; MENDES, R. S.; LENZI, E. K. “q-Exponential Distribution in Urban Agglomeration”. Physical Review E - Statistical, Nonlinear, and Soft Matter Physics, v. 65, n. 1, p. 1–3, (2002).

[11] CAMPO, M. G.; FERRI, G. L.; ROSTON, G. B. et al. “q -Exponential Distribution in Time Correlation Function of Water Hydrogen Bonds”. Brazilian Journal of Physics, v. 39, n. 2, (2009).

[12] BERCHER, J. F.; VIGNAT, C. “A New Look at q-Exponential Distributions Via Excess Statistics”. Physica A: Statistical Mechanics and its Applications, v. 387, n. 22, p. 5422–5432, (2008).

[13] BOLFARINE, H.; SANDOVAL, M. C. *Introdução a Inferência Estatística*. São Paulo, Brasil, (2000).

[14] GONÇALVES, A. M. S. “O Problema de Min-Max-Min com Restrições pelo Método de Nelder-Mead”. Universidade Federal do Rio de Janeiro, UFRJ, (2013).

[15] NELDER, J. A.; MEAD, R. “A Simplex Method for Function Minimization”. Computer Journal, v. 7, p. 308–313, (1965).

[16] FONSECA, R.; CRIBARI-NETO, F. “Inference in a Bimodal Birnbaum-Saunders Model”. n. 2005, p. 1–33, (2016).

[17] 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, p. 453–461, (1946).