UNIVERSIDADE FEDERAL DE PERNAMBUCO
PROGRAMA DE PÓS-GRADUAÇÃO EM ENGENHARIA DE PRODUÇÃO

A NOVEL *q*-EXPONENTIAL BASED STRESS-STRENGTH RELIABILITY MODEL AND APPLICATIONS TO FATIGUE LIFE WITH EXTREME VALUES

ROMERO LUIZ MENDONÇA SALES FILHO

ORIENTADOR: Prof. Enrique López Droguett, Ph.D.

RECIFE, 2016

ROMERO LUIZ MENDONÇA SALES FILHO

A NOVEL *q*-EXPONENTIAL BASED STRESS-STRENGTH RELIABILITY MODEL AND APPLICATIONS TO FATIGUE LIFE WITH EXTREME VALUES

Tese de Doutorado apresentada à UFPE para a obtenção de grau de Doutor como parte das exigências do Programa de Pós-Graduação em Engenharia de Produção (Área de Concentração: Pesquisa Operacional).

Orientador: Prof. Enrique López Droguett, Ph.D.

Recife

2016

Catalogação na fonte
Bibliotecária Maria Luiza de Moura Ferreira, CRB-4 / 1469

S163a Sales Filho, Romero Luiz Mendonça.
A novel *q*-exponential based stress-strength reliability model and applications to fatigue life with extreme values / Romero Luiz Mendonça Sales Filho. 2016.
127 folhas, il.

Orientador: Prof. Enrique López Droguett, PhD.
Tese (Doutorado) – Universidade Federal de Pernambuco. CTG. Programa de Pós-graduação em Engenharia de Produção, 2016.
Inclui Referências.

1. Engenharia de Produção. 2. *Q*-exponential. 3. Stress-strength reliability. 4. Maximum Likelihood Estimators. 5. Nelder-Mead. 6. Particle Swarm Optimization. I. Droguett, Enrique López (Orientador). II. Título.

658.5 CDD (22. ed.) UFPE/BCTG/2016-118

UNIVERSIDADE FEDERAL DE PERNAMBUCO
PROGRAMA DE PÓS-GRADUAÇÃO EM ENGENHARIA DE PRODUÇÃO

PARECER DA COMISSÃO EXAMINADORA

DE TESE DE DOUTORADO DE

ROMERO LUIZ MENDONÇA SALES FILHO

"***A Novel q-Exponential Based Stress-Strength Reliability Model and Applications to Fatigue Life With Extreme Values***"

ÁREA DE CONCENTRAÇÃO: Pesquisa Operacional

A comissão examinadora composta pelos professores abaixo, sob a presidência do primeiro, considera o candidato **ROMERO LUIZ MENDONÇA SALES FILHO**, **APROVADO.**

Recife, 24 de Fevereiro de 2016.

____

Prof. ENRIQUE LÓPEZ DROGUETT, PhD (UFPE)

____

Prof. MÁRCIO JOSÉ DAS CHAGAS MOURA, Doutor (UFPE)

____

Prof. CRISTIANO ALEXANDRE VIRGÍNIO CAVALCANTE, Doutor (UFPE)

____

Prof. JOSICLÊDA DOMICIANO GALVÍNCIO, Doutora (UFPE)

____

Prof. ALEXANDRE CABRAL MOTA, Doutor (UFPE)

Dedico este trabalho a todos que acreditaram no meu sucesso, mesmo quando eu já não tinha mais forças para continuar: Minha amada esposa Desirré, meus queridos pais Romero e Gilvanice e minha irmã Gimenna. Sem o apoio de vocês eu não teria conseguido...

AGRADECIMENTOS

Agradeço primeiramente a Deus, que tenho certeza, esteve sempre ao meu lado tanto nos momentos de aflição e angústia como nos momentos de inspiração e criação. Também faço um agradecimento especial a Nossa Senhora, que intercedeu por mim, nas várias horas difíceis, horas essas que pareciam não ter fim. Hoje sinto que minha fé aumentou ainda mais diante dos problemas que enfrentei ao longo desses quase 5 anos de luta.

Aos meus pais por terem me educado seguindo preceitos de ética e respeito ao próximo, sendo eles o alicerce de toda a minha educação. Agradeço por sempre estarem presente na minha vida, nas alegrias e aflições, e por levarem sempre uma palavra de estímulo para que eu perseverasse nos meus objetivos. Eles fizeram por mim e por minha irmã tudo o que deixaram de fazer por eles, e hoje colhem as nossas vitórias com satisfação e orgulho. Obrigado por sempre colocarem meu nome em suas orações pedindo a Deus e a Virgem Maria que todos os obstáculos que se colocavam diante da minha vitória fossem vencidos – tenho certeza que essas orações mantiveram sempre Jesus Cristo e Nossa Senhora ao meu lado.

À minha querida Esposa, Desirré Duda. Quantos momentos de lutas, hein amor? Quantas noites dormistes longe de mim pois eu estava virando noites no computador? Como é bom ter você ao meu lado! Que bom que és uma pessoa compreensiva e que considera a minha vitória a sua vitória! Agradeço por você me amar, me respeitar e estar sempre ao meu lado nas lutas dessa vida. Obrigado por ter entendido as dificuldades do doutorado e por ter proporcionado toda a liberdade que eu precisava para conduzir meus trabalhos. Amo você!

Aos meus familiares e amigos, que muitas vezes me ajudaram com conselhos e orientações. Destaco aqui especialmente: minha irmã Gimenna, minha avó Gedida, minha tia Rejane, meu tio Júnior, meus primos Jadson, Maxwell e Mário, meus amigos da UFRPE/UAG, Werônica, Ana Cláudia, Jorge, Mirko, Thibério, Fábio, Jairo, e Fernando.

Ao meu orientador Prof. Enrique López, que mesmo à distância sempre se mostrou presente discutindo e buscando soluções para os problemas. Um agradecimento especial aos professores Márcio e Isis, sempre solícitos para discutir o trabalho e fundamentais para a conclusão do mesmo. Obrigado, amigos! Sem vocês eu não teria conseguido.

Agradeço também à CAPES por ter financiado por alguns anos os meus estudos e também à UFPE por ter me acolhido nos anos de graduação, mestrado e doutorado, e por ter proporcionado as condições necessárias para o desenvolvimento de minhas atividades acadêmicas culminando com a conquista do grau de doutor.

ABSTRACT

In recent years, a family of probability distributions based on Nonextensive Statistical Mechanics, known as q-distributions, has experienced a surge in terms of applications to several fields of science and engineering. In this work the q-Exponential distribution will be studied in detail. One of the features of this distribution is the capability of modeling data that have a power law behavior, since it has a heavy-tailed probability density function (PDF) for particular values of its parameters. This feature allows us to consider this distribution as a candidate to model data sets with extremely large values (e.g. cycles to failure). Once the analytical expressions for the maximum likelihood estimates (MLE) of q-Exponential are very difficult to be obtained, in this work, we will obtain the MLE for the parameters of the q-Exponential using two different optimization methods: particle swarm optimization (PSO) and Nelder-Mead (NM), which are also coupled with parametric and non-parametric bootstrap methods in order to obtain confidence intervals for these parameters; asymptotic intervals are also derived. Besides, we will make inference about a useful performance metric in system reliability, the called index $R = P(Y < X)$, where the stress Y and strength X are independent q-Exponential random variables with different parameters. In fact, when dealing with practical problems of stress-strength reliability, one can work with fatigue life data and make use of the well-known relation between stress and cycles until failure. For some materials, this kind of data can involve extremely large values and the capability of the q-Exponential distribution to model data with extremely large values makes this distribution a good candidate to adjust stress-strength models. In terms of system reliability, the index R is considered a topic of great interest, so we will develop the maximum likelihood estimator (MLE) for the index R and show that this estimator is obtained by a function that depends on the parameters of the distributions for X and Y. The behavior of the MLE for the index R is assessed by means of simulated experiments. Moreover, confidence intervals are developed based on parametric and non-parametric bootstrap. As an example of application, we consider two experimental data sets taken from literature: the first is related to the analysis of high cycle fatigue properties of ductile cast iron for wind turbine components, and the second one evaluates the specimen size effects on gigacycle fatigue properties of high-strength steel.

Keyword: Q-Exponential . Stress-Strength Reliability . Maximum Likelihood Estimators . Nelder-Mead . Particle Swarm Optimization.

RESUMO

Nos últimos anos, tem sido notado em diversas áreas da ciência e engenharia, um aumento significativo na aplicabilidade da família q de distribuições de probabilidade que se baseia em Mecânica Estatística Não Extensiva. Uma das características da distribuição q-Exponencial é a capacidade de modelar dados que apresentam comportamento de lei de potência, uma vez que tal distribuição possui uma função densidade de probabilidade (FDP) que apresenta cauda pesada para determinados valores de parâmetros. Esta característica permite-nos considerar tal distribuição como candidata para modelar conjuntos de dados que apresentam valores extremamente grandes (Ex.: ciclos até a falha). Uma vez que expressões analíticas para os estimadores de máxima verossimilhança dos parâmetros não são facilmente encontradas, neste trabalho, iremos obter as estimativas de máxima verossimilhança dos parâmetros através de dois métodos de otimização: *particle swarm optimization* (PSO) e *Nelder-Mead* (NM), que além das estimativas pontuais, irão nos fornecer juntamente com abordagens *bootstrap*, intervalos de confiança para os parâmetros da distribuição; intervalos assintóticos também serão derivados. Além disso, faremos inferência sobre um importante índice de confiabilidade, o chamado Índice $R = P(Y < X)$, onde Y (estresse) e X (força) são variáveis aleatórias independentes. De fato, quando tratamos de problemas práticos de força-estresse, podemos trabalhar com dados de fadiga e fazer uso da bem conhecida relação entre estresse e ciclos até a falha. Para alguns materiais, esse tipo de variável pode apresentar dados com valores muito grandes e a capacidade da q-Exponencial em modelar esse tipo de dado torna essa uma distribuição a ser considerada para ajustar modelos de força-estresse. Em termos de confiabilidade de sistemas, o índice R é considerado um tópico de bastante interesse, assim iremos desenvolver os estimadores de máxima verossimilhança para esse índice e mostrar que esse estimador é obtido através de uma função que depende dos parâmetros da distribuição de X e Y. O comportamento do estimador é investigado através de experimentos simulados. Intervalos de confiança são desenvolvidos através de bootstrap paramétrico e não-paramétrico. Duas aplicações envolvendo dados de ciclos até a falha e retiradas da literatura são consideradas: a primeira para ferro fundido e a segunda para aço de alta resistência.

Palavras-Chave: Q-Exponencial . Confiabilidade Força-Estresse . Estimador de Máxima Verossimilhaça . Nelder-Mead . Particle Swarm Optimization

LIST OF FIGURES

LIST OF TABLES

LIST OF ACRONYMS

PSC - PRESSURE STIMULATED CURRENT

PDF - PROBABILITY DENSITY FUNCTION

CDF - CUMULATIVE DISTRIBUTION FUNCTION

PSO - PARTICLE SWARM OPTIMIZATION

NM - NELDER-MEAD

MSE - MEAN SQUARED ERROR

SVM - SUPPORT VECTOR MACHINES

K-S - KOLMOGOROV-SMIRNOV

SUMMARY

1 INTRODUCTION

1.1 Motivation

Complex systems are those whose elements that constitute them present strong interactions with each other [1], [2]. The search for new probabilistic models to describe this kind of systems has substantially increased over the last years[3]–[5]. In complex systems, the entropy is not an extensive quantity, *i.e.*, there is a dependency relationship between different components of the system [6], and it can be argued that in such cases the statistical mechanics of Boltzmann-Gibbs [7] is not appropriate. Then, Nonextensive Statistical Mechanics arises as a generalization of the statistical mechanics of Boltzmann-Gibbs[6] in order to overcome this limitation.

A family of probability distributions based on Non-extensive Statistical Mechanics, known as q-distributions, has experienced a surge in terms of applications to several fields of science and engineering. Since the Nonextensive Statistical Mechanics assumes interdependencies among the components of a system, these q-distributions have the ability of modeling complex systems. Picoli *et al*. [3] described the basic properties of three distributions of this kind: q-Exponential, q-Gaussian and q-Weibull. In another work, Picoli *et al*. [4] presented a comparative study, where q-Exponential, q-Weibull, and Weibull distributions were used to investigate frequency distributions of basketball baskets, cyclone victims, brand-name drugs by retail sales and highway length. Complex systems such as cyclones [8], gravitational systems [9], stock market [10], [11] , journal citations [12] , complex DNA structural organization [13], reliability analysis [5], cosmic rays [14], earthquakes [15], financial markets [16], internet [17], mechanical stress [18], among others have been satisfactorily described by q-distributions.

Another field of application of q-distributions is mechanical stress. For instance, it has been experimentally demonstrated that when a rock sample is subjected to mechanical stress, an electrical signal is emitted[19], [20]. This electrical signal is related to the evolution of cracks' network within the stressed sample and is called Pressure Stimulated Current (PSC). In [18], PSC emissions in marble and amphibolite samples are considered to follow a q-Exponential distribution.

The q-Exponential distribution is obtained by maximizing the non-extensive entropy under appropriate constraints[21]. This distribution has two parameters (q and η), differently from the Exponential distribution that is one parametric. This feature gives more flexibility to q-Exponential distribution with regard to its decay for the Probability Density Function (PDF) curve. Indeed, for a fixed parameter η, a slower or faster decay of the PDF is observed depending on the value of q. Moreover, q-Exponential does not have the limitation of a constant hazard rate, thus allowing the modeling of either system improvement ($1 < q < 2$) or degradation ($q < 1$).

In the context of reliability, it is expected that for a given sample with large values (e.g., realizations of rare events), both q-Exponential and Weibull distributions can fit the data well. In situations like these, the parameter q of the q-Exponential would lie within the interval (1, 2) and the shape parameter of the Weibull distribution would be in (0, 1) [4], [22].

Moreover, a prominent point of the q-Exponential distribution is its ability to model data that presents a power law behavior [4]. Thus, we expect a superior performance of q-Exponential over for example the Exponential and Weibull distribution in the characterization of data sets with extremely large values, since, as pointed out by Laherrère and Sornette[22], a stretched exponential PDF (Example: Weibull distribution) has a tail that is heavier than that of the Exponential PDF but lighter than that of a power law PDF (Example: q-Exponential distribution), the following figure shows this behavior:

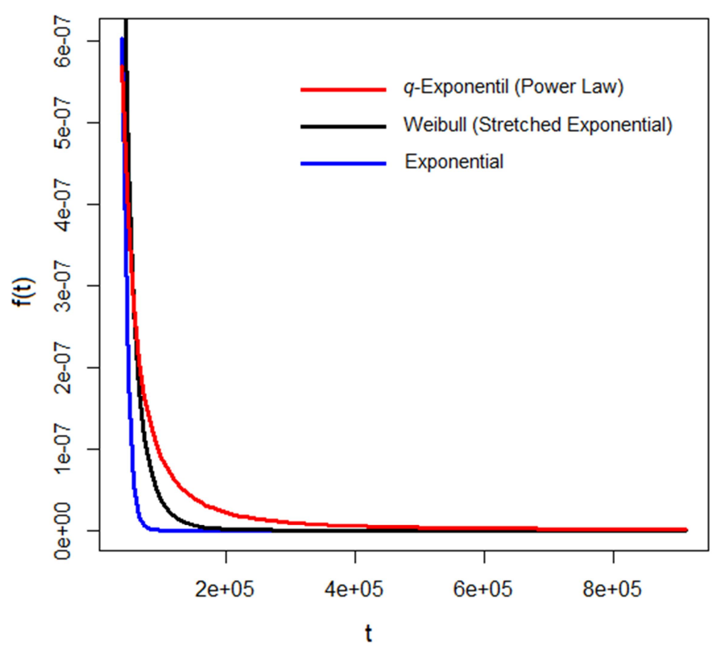


Figure 1.1 Behavior of the tails of the distributions q-Exponential, Weibull and Exponential.

Thus, more than an alternative to the Weibull Distribution, the q-Exponential distribution can be considered the main distribution to be considered in some situations that we have extremely large values, for example we can mention data of cycles until failure of a very resistant material like high-strengthed steel. This characteristic is due to the ability to model data that are present in the tail of the distribution, i.e., extremely large values. Then, the use of the q-Exponential with power law characteristic $(1 < q < 2)$, in order to model fatigue of material is an important topic to be investigated as we will propose in this work.

In fact, several studies have investigated the presence of power laws in the behavior of data observed in fatigue analysis of materials. For instance, in[23] the acoustic emissions of microfractures before the breakup of the sample are evaluated, where the authors used samples made of composite inhomogeneous materials such as plaster, wood, or fiberglass. The experimental results were similar for all materials, and the authors conclude that statistics from acoustic energy measurements strongly suggest that the fracture can be viewed as a critical phenomenon and energy events are distributed in magnitude as a power law.

Moreover, according to Basquin's law, the lifetime of a system increases as a power law with the reduction of the applied load amplitude [24]. Therefore, the alternating stress in terms of number of cycles to failure is expressed in a power law form [25] known as the Wöhler curve (SN curve). It has also been suggested that the underlying fracture dynamics in some systems might display self-organized criticality [26], implying that long-range interactions between fracture events lead to a scale-free cascade of 'avalanches' [27]. For instance, in [27], the authors present a scalar model of microfracturing that generates power law behavior in properties related to acoustic emission, and a scale-free hierarchy of avalanches characteristic of self-organized criticality [2].

In terms of system reliability, a topic of considerable interest is the inference about the index $R = P(Y < X)$, where X is the strength of a component that is subjected to stress Y; when Y is greater than X, the system fails. Thus, R can be considered as a measure of system reliability. Stress-strength models are used in many applications of physics and engineering such as strength failure and system collapse [28].

Most of the works that aim at estimating R assume that X and Y are independent and follow the same type of probability distributions. For instance, X and Y have been considered as Normal [29]–[32], Exponential [33], [34] and Weibull [35]–[37] random variables. Moreover, the Generalized Pareto and Generalized Rayleigh distributions are also discussed to model X and Y in [38], [39], respectively. Kundu and Gupta [40] and Raqab et al. [41] also consider the Generalized Exponential distributions. Al-Zahrani and Al-Harbi [42] adopt the Lomax distribution under General Progressive Censoring, and Panahi and Asadi [43] assume that X and Y follow Lomax distributions with a common scale, but different shape parameters.

In this sense, this work seeks to contribute with the insertion of a new probabilistic model to describe models of Stress-Strength. The fact of the q-Exponential distribution present good results for modeling data with large order of magnitude, allow us to propose such distribution to deal with stress-strength models when we are working with fatigue life data (number of cycles until failure). Note that in stress-strength problems we can use life cycle data as model input data, since there is a one-to-one relationship between stress level and number of cycle to failure that is represented by the SN curve [25], [44]. Thus from the SN curve we can observe that, stress level and number of cycles to failure can be interchangeably used in problems of this kind. The Figure 1.2 represents an example of SN-Curve:

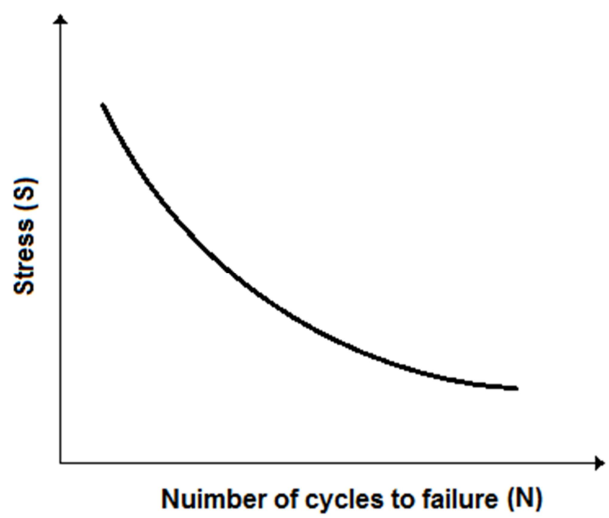


Figure 1.2. Example of SN-curve.

It is worth mentioning that, in order to obtain analytical expressions for estimators of q-exponential parameters, it is necessary solving a complicated set of equations. In this way, Shalizi [45] and Bercher and Vignat [46] have shown that a reparameterization for that set of equations is required. However, this approach allows obtaining analytical expressions for the MLE only when $1 < q < 2$ and, however, the q-Exponential distribution is also defined for $q <$

1. Such a parameter range corresponds to hazard rate behavior with relevant applications, namely the reliability modeling of degrading systems. In this way, with the exception of the case $1 < q < 2$ [45], analytical expressions for the maximum likelihood estimators of q-Exponential are very difficult to be obtained due to the intricate derivatives of the log-likelihood function.

Thus, in this work the MLE for the parameters of a q-Exponential distribution are numerically derived through two different optimization algorithms: Nelder-Mead [47] and Particle Swarm Optimization (PSO) [48]. The results obtained from these two approaches will be compared by means of bias and MSE (Mean Squared Error).

1.2 Objectives

1.2.1 General Objective

The main objective of this work is to develop a new Stress-Strength Model based in the q-Exponential distribution to work with fatigue life data that present extremely large values.

1.2.2 Specific Objectives

- Apply PSO and Nelder-Mead optimization methods to maximizing the log-likelihood function of the q-exponential distribution in order to find the maximum likelihood estimates of the model parameters.
- Evaluate, by numerical experiments, the performance of the PSO and Nelder-Mead optimization methods in the estimation of the q-Exponential parameters.
- Develop the estimator of the Index $R=P(Y<X)$ based on the q-Exponential distribution considering the particularity of the support of the q-Exponential PDF.
- Apply PSO and Nelder-Mead optimization methods in order to find the maximum likelihood estimates of the index $R=P(Y<X)$.
- Evaluate the performance of the PSO and Nelder-Mead optimization methods in the estimation of the index $R=P(Y<X)$.

- Evaluate the quality of the confidence intervals obtained for the q-Exponential parameters and for the index R, by bootstrap methods (parametric and non-Parametric), through simulated experiments.
- Compare, through hypothesis tests, the effectiveness of the results obtained for the point and interval estimates of the q-Exponential parameters and of the index R. Besides, compare the results obtained for the confidence intervals when we use parametric and non-parametric bootstrap approaches.
- Apply the proposed model of stress-strength reliability based in the q-Exponential distribution in a real example and comparing the results with the results obtained when we consider other probability distributions (Weibull and Exponential).

1.3 Structure of the Work

This work presents six chapters, including this introduction. The chapter 2 treats of the theoretical background where first we develop the q-Exponential distribution discussing important features and particularities of this distribution. Still in the chapter 2, we will address the approaches for point and interval estimation of parameters and also the optimization methods used in this work. The hypothesis tests used in order to make comparisons with the results of the point and confidence intervals also will be discussed in this chapter. Chapter 3 addresses the estimation of the q-exponential parameters, developing numerical experiments for the maximum likelihood estimators and confidence intervals. The chapter 4 presents the q-exponential distribution in the approach of calculating the index $R = P\,(Y < X)$. In this chapter we will develop the estimator of the index R respecting important features of the q-Exponential PDF. Numerical experiments will be presented in the chapter 4 in order to evaluate the quality of the point estimator and confidence intervals calculated via PSO and Nelder-Mead. In chapter 5 we will present two case studies of the new proposed stress-strength model. Also in this chapter, we will calculate the index R based in the Weibull distribution, in order to compare the result with the result obtained with the new proposed stress-strength model. Finally, Chapter 6 provides some concluding remarks.

2 THEORETICAL BACKGROUND

2.1 The q-Exponential distribution

The q-Exponential PDF is given by the following expression:

$$f_q(t) = \frac{(2-q)}{\eta} exp_q\left[-\left(\frac{t}{\eta}\right)\right]; \quad q < 2 \quad and\, \eta > 0$$

,

Where q is the parameter that determines the density shape and is known as entropic index, η is the scale parameter and $exp_q(x)$ is the q-Exponential function defined as:

$$exp_q(x) = \begin{cases} [1+(1-q)x]^{\frac{1}{1-q}}, & \text{if } [1+(1-q)x] \geq 0 \\ 0, & \text{otherwise,} \end{cases}$$

,

Where x and $q \in \mathbb{R}$.

Note that the q-Exponential PDF becomes an Exponential PDF when $q \to 1$. Thus, the q-Exponential distribution is a generalization of the Exponential one. The parameters η and q determine how quickly the PDF decays. Note also that the q parameter dictates how the distribution deviates from exponentiality, and this deviation is also defined by the decay of the distribution. When compared to the decay of the Exponential distribution with the same parameter η, the q-Exponential presents a slower decay for $1 < q < 2$ (Power Law characteristic) and a faster decay for $q < 1$; for a fixed parameter q, we will have a similar behavior of the exponential distribution ($q \to 1$), *i.e.*, insofar as the value of the parameter η increases it is observed a slower decay of the PDF.

By using the definition of the q-Exponential function, it is possible to rewrite the density of q-Exponential:

$$f_q(t) = \frac{(2-q)\left[1-\frac{(1-q)t}{\eta}\right]^{\frac{1}{1-q}}}{\eta}, \quad q < 2 \ \text{ and } \eta > 0$$

.

Furthermore, the support t is changed depending on the value of the entropic index as follows:

$$t \in \begin{cases} [0;\ \infty), & q \geq 1 \\ \left[0; \frac{1}{\frac{1}{\eta}(1-q)}\right), & q < 1 \end{cases} \tag{2.1}$$

.

Figure 2.1 and Figure 2.2 presents the q-Exponential PDF for some possible values of q and η, illustrating the behavior that was previously commented.

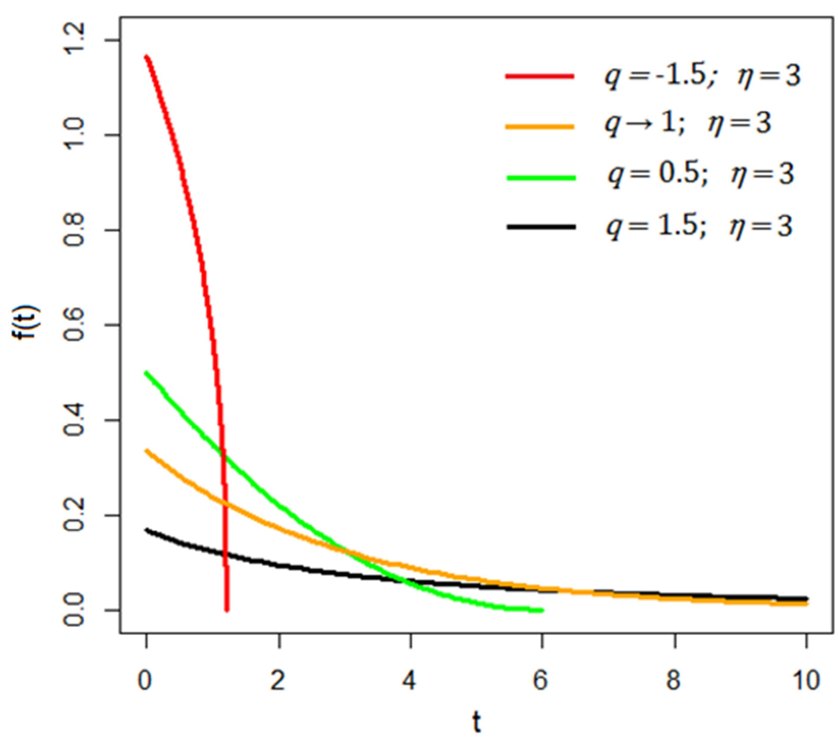


Figure 2.1. q-Exponential PDF for a fixed $\boldsymbol{\eta}$ *and some possible values of* $\boldsymbol{q}$

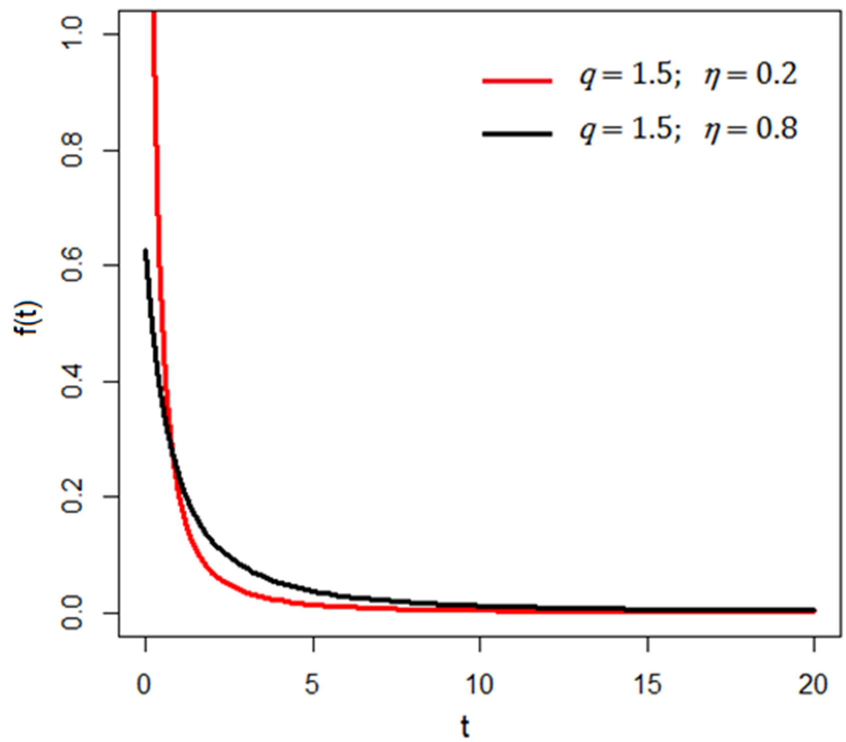


Figure 2.2. q-Exponential PDF for a fixed q and some possible values of η

The Cumulative Distribution Function (CDF) of the q-Exponential is defined by the following expression:

$$F_q(t) = \begin{cases} 1 - exp_{q'}\left\{-\left[\frac{t.(2-q)}{\eta}\right]\right\}, & t \geq 0 \\ 0, & \text{otherwise} \end{cases}$$

,

where $q' = \frac{1}{2-q}$. By inverting $F_q(t)$, we obtain a q-Exponential random number generator:

$$T = \frac{\eta\left[1 - (U)^{\frac{1-q}{2-q}}\right]}{1-q} \quad (2.2)$$

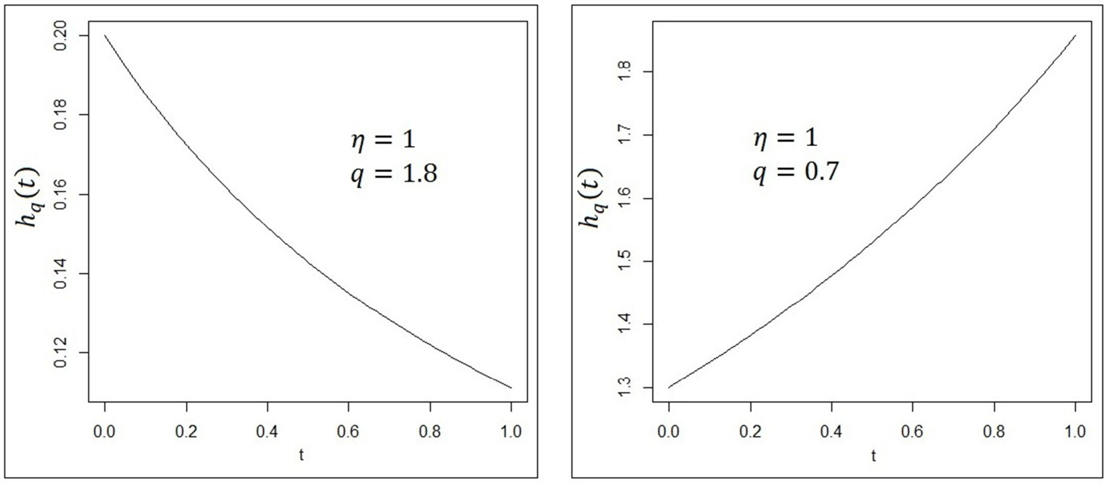
Where U is a uniform random variable defined in $[0,1]$.

An important characteristic of the q-Exponential distribution, especially in the reliability context, is that the q-Exponential hazard rate is not necessarily constant as occurs for the Exponential distribution. In fact, we will show that for a q-Exponential distribution, we can model two additional behaviors for the hazard rate. To prove this, let us first define the hazard rate $h_q(t) = \frac{f_q(t)}{R_q(t)}$.

Then, we can write:

$$h_q(t) = \frac{\frac{(2-q)\left[1-\frac{(1-q)t}{\eta}\right]^{\frac{1}{1-q}}}{\eta}}{\left[1-\frac{(1-q)t}{\eta}\right]^{\frac{2-q}{1-q}}} = \frac{(2-q)}{\eta}\left[1-\frac{(1-q)t}{\eta}\right]^{\frac{q-1}{1-q}}$$

Thus, the q-Exponential distribution is able to represent two different types of hazard rate behaviors depending on the values the parameter q assumes. For $1 < q < 2$, $h_q(t)$ is a decreasing monotonic function (Figure 2.3 (a)), while for $q < 1$, $h_q(t)$ increases monotonically (Figure 2.3 (b)).



(a) (b)

Figure 2.3. (a) q-Exponential $h_q(t)$ with $\eta = 1$, and $q = 1.8$; (b) q-Exponential $h_q(t)$ with $\eta = 1$, and $q = 0.7$.

Nadarajah and Kotz [49] point out that many of the q-distributions that have emerged recently were known by other names and they particularly discuss two families of distributions: Burr-type XII and Burr-type III, which have many q-distributions as special cases. However, it is worth noting that the q-Exponential is a generalization of the Burr XII and not the opposite, as stated by Nadarajah and Kotz [50], since the q-Exponential is valid even for $q < 1$, which does not happen with the Burr XII.

2.2 Estimation of Parameters

In this section, we will briefly describe the methods that will be used in this work in order to obtain point and interval estimates of parameters of interest. First we will describe the Maximum Likelihood Estimation method and next we will deal with the approach of bootstrapped and asymptotic confidence intervals.

2.2.1 Maximum Likelihood Estimation

The maximum likelihood method is one of the most used in order to obtain estimates of the parameters of a probabilistic model. The principle is to estimate the parameter (or parameters, if we have a multi-parametric model) which best characterizes a sample that was obtained from a population governed by a certain probabilistic model. Thus, the method seeks to determine the distribution, among all those defined by the possible values of their parameters, with greater chance of having generated the sample analyzed.

2.2.1.1 Uniparametric Case

Firstly we present the concept of likelihood function:

Consider a random sample of the random variable *X*, with size *n*: $X_1, X_2, \ldots, X_n$. We represent 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 write as [51]:

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

The Maximum Likelihood Estimator of θ will be the value that maximizes the likelihood function. We will denote this value as $\hat{\theta} \in \Theta$.

Normally is simpler to obtain maximum likelihood estimates maximizing the natural logarithm of the likelihood function. Of course the value of $\hat{\theta}$ that maximizes the likelihood function is equal to value that maximizes the logarithm of the likelihood function. In this way, we define the log-likelihood function as:

$$l(\theta, X) = log[L(\theta, X)]$$

Thus, in the uniparametric case, we get the maximum likelihood estimate calculating the root of the derivative of the log-likelihood function, *i.e.*:

$$\frac{d\ l(\theta, X)}{d\theta} = 0 \qquad (2.3)$$

It is important to report that in some situations, the root of the derivative of the log-likelihood function can be obtained analytically, however, in some situations the solution of the Equation (2.3) will only be obtained by numerical procedures.

2.2.1.2 Multiparametric Case

In situations that we intend to estimate parameters of a multiparametric model, *i.e.* $\theta=(\theta_0, \theta_1, \ldots, \theta_r)$ by maximum likelihood approach we can following the same ideas presented for the uniparametric models with some adaptations. The estimates for the parameters can be obtained by solving the following equations [51]:

$$\frac{d\, l(\theta_0, \theta_1, \ldots, \theta_r, X)}{d\, \theta_i} = 0 \quad , \textit{for } i = 1, \ldots, r.$$

2.2.1.3 Likelihood for independent samples

In some situations, we have two or more independent samples that have one or more parameter of interest. Thus, we can represent a unique likelihood function for these variables, as [51]:

$$L(\theta_{1X}, \theta_{2X}, \ldots, \theta_{rX}; \theta_{1Y}, \theta_{2Y}, \ldots, \theta_{kY}; X, Y) = L(\theta_{1X}, \theta_{2X}, \ldots, \theta_{rX}; X) L(\theta_{1Y}, \theta_{2Y}, \ldots, \theta_{kY}; Y)$$

Note that, this result is valid due the independence of the samples.

Thus, using the properties of the logarithmic function, we can write the log-likelihood for two independent samples, as:

$$l(\theta_{1X}, \theta_{2X}, \ldots, \theta_{rX}; \theta_{1Y}, \theta_{2Y}, \ldots, \theta_{kY}; X, Y) = l(\theta_{1X}, \theta_{2X}, \ldots, \theta_{rX}; X) + l(\theta_{1Y}, \theta_{2Y}, \ldots, \theta_{kY}; Y)$$

2.2.2 Bootstrap Confidence Intervals

In this section, we present the algorithms used in order to build Parametric and Non-Parametric Bootstrap confidence Intervals for the Parameters of a model [52], [53].

2.2.2.1 Bootstrap-p

The algorithm for constructing confidence intervals by using the bootstrap-*p* approach has the following steps:

- **Step 1:** From an initial sample for the variable $X=\{x_1,x_2,\ldots,x_n\}$, estimate the parameters θ, by maximizing the log-likelihood function;

- **Step 2:** Use the estimate obtained in the previous step and the random number generator of variable *X* to generate new samples for *X*, *i.e.*: $\{x_1^*,\ x_2^*,\ldots,\ x_n^*\}$. Based on this new sample, compute the bootstrap sample estimate of θ, say θ^*, maximizing the log-likelihood function of the variable *X*;
- **Step 3:** Repeat step 2, N times;

- **Step 4:** Using the N values of θ^* obtained in step 3 and by adopting a γ significance level, find the percentiles $\theta^*_{\alpha/2}$ and $\theta^*_{1-(\alpha/2)}$. Thus, it is possible to determine an approximate confidence interval, with confidence interval equal to *100*(1- α)%*, for the parameter θ, as:

$$C.I.=[\theta^*_{\alpha/2},\theta^*_{1-(\alpha/2)}] \tag{2.4}$$

.

2.2.2.2 Non-Parametric bootstrap

The algorithm for constructing confidence intervals by using the non-parametric bootstrap approach is as follows:

- **Step 1:** From an initial sample for the variable $X=\{x_1,\ x_2,\ldots,x_n\}$, generate new samples for *X* by sampling with replacement, *i.e.*, $\{x_1^*,\ x_2^*,\ldots,\ x_n^*\}$. Based on this new sample, compute the estimate of θ, say θ^*, maximizing the log-likelihood function of the variable *X*;
- **Step 2:** Repeat step 1, *N* times.
- **Step 3:** By using the *N* values of θ^* from step 2 and by adopting a γ significance level, the percentiles $\theta^*_{\alpha/2}$and $\theta^*_{1-(\alpha/2)}$ are obtained; they determine an approximate

confidence interval for the parameter θ with confidence level equals to *100*(1- α)%* using Equation (2.4).

2.2.3 Asymptotic confidence intervals

According to the asymptotic properties of the maximum likelihood estimators, the related covariance matrix can be estimated by the inverse of the observed information matrix ($I(\hat{\theta}|t)$), *i.e.*, the negative of the second derivative of the log-likelihood function evaluated at the point estimate $\hat{\theta}$ given data *t* [54], [55]. Note that the observed information matrix is the negative of the Hessian. In this way, both matrices are symmetric and measure the amount of curvature in the log-likelihood function. For example, in a $r-$parametric model we have , $\hat{\theta}=(\hat{\theta}_1,\hat{\theta}_2,\dots,\hat{\theta}_r)$, and the covariance matrix associated to the maximum likelihood estimators is as follows:

$$\widehat{\mathrm{var}}(\hat{\theta}_1,\hat{\theta}_2,\dots,\hat{\theta}_r)=I^{-1}(\hat{\theta}_1,\hat{\theta}_2,\dots,\hat{\theta}_r|t)=$$

$$-\begin{bmatrix} \frac{\partial^2 L(\hat{\theta}_1,\hat{\theta}_2,\dots,\hat{\theta}_r|t)}{\partial\hat{\theta}_1{}^2} & \frac{\partial^2 L(\hat{\theta}_1,\hat{\theta}_2,\dots,\hat{\theta}_r|t)}{\partial\hat{\theta}_1\partial\hat{\theta}_2} & \cdots & \frac{\partial^2 L(\hat{\theta}_1,\hat{\theta}_2,\dots,\hat{\theta}_r|t)}{\partial\hat{\theta}_1\partial\hat{\theta}_r} \\ \frac{\partial^2 L(\hat{\theta}_1,\hat{\theta}_2,\dots,\hat{\theta}_r|t)}{\partial\hat{\theta}_2\partial\hat{\theta}_1} & \frac{\partial^2 L(\hat{\theta}_1,\hat{\theta}_2,\dots,\hat{\theta}_r|t)}{\partial\hat{\theta}_2{}^2} & \cdots & \frac{\partial^2 L(\hat{\theta}_1,\hat{\theta}_2,\dots,\hat{\theta}_r|t)}{\partial\hat{\theta}_2\partial\hat{\theta}_r} \\ \vdots & \vdots & \ddots & \\ \frac{\partial^2 L(\hat{\theta}_1,\hat{\theta}_2,\dots,\hat{\theta}_r|t)}{\partial\hat{\theta}_r\partial\hat{\theta}_1} & \frac{\partial^2 L(\hat{\theta}_1,\hat{\theta}_2,\dots,\hat{\theta}_r|t)}{\partial\hat{\theta}_r\partial\hat{\theta}_2} & \cdots & \frac{\partial^2 L(\hat{\theta}_1,\hat{\theta}_2,\dots,\hat{\theta}_r|t)}{\partial\hat{\theta}_r{}^2} \end{bmatrix}^{-1} \quad (2.5)$$

,

Once the covariance matrix is estimated, asymptotic confidence intervals can be constructed for the parameters of the distribution by using the asymptotic normality property of the maximum likelihood estimators. The asymptotic $(1-\alpha)\cdot 100\%$ confidence intervals for $\hat{\theta}_1,\hat{\theta}_2,\dots,\hat{\theta}_r$ are given by, respectively:

$$\mathrm{C.I.}\,[\theta_1;(1-\alpha)\cdot 100\%]=\left[\hat{\theta}_1+z_{\frac{\alpha}{2}}\sqrt{\widehat{\mathrm{var}}_{11}},\ \hat{\theta}_1+z_{1-\left(\frac{\alpha}{2}\right)}\sqrt{\widehat{\mathrm{var}}_{11}}\right],$$

$$(2.6)$$

$$\text{C.I.}[\theta_2;(1-\alpha)\cdot 100\%] = \left[\hat{\theta}_2 + z_{\frac{\alpha}{2}}\sqrt{\widehat{\text{var}}_{22}},\ \hat{\theta}_2 + z_{1-\left(\frac{\alpha}{2}\right)}\sqrt{\widehat{\text{var}}_{22}}\right], \quad (2.7)$$

$$\vdots \qquad \vdots$$

$$\text{C.I.}[\theta_r;(1-\alpha)\cdot 100\%] = \left[\hat{\theta}_r + z_{\frac{\alpha}{2}}\sqrt{\widehat{\text{var}}_{rr}},\ \hat{\theta}_r + z_{1-\left(\frac{\alpha}{2}\right)}\sqrt{\widehat{\text{var}}_{rr}}\right], \quad (2.8)$$

2.3 Optimization Methods

2.3.1 Particle Swarm Optimization

PSO [48] is a probabilistic optimization heuristic based on the motion of groups of organisms (e.g., flocks of birds, schools of fishes), which optimizes a problem from a population of candidate solutions (particles). According to update equations over the particles' position and velocities, the candidate solutions explore the search-space. Each particle's movement is influenced by its own best position and also by its neighbors' best. Thus, it is expected that particles move toward the best solutions.

PSO has been successfully applied to different contexts. For example, [56]–[63] apply PSO in the adjustment of the hyperparameters that emerge in the training problem of support vector machines (SVM). Indeed, Lin et al. [56] , Lins et al. [63] and Droguett et al. [62] use PSO not only to adjust the SVM hyperparameters, but also for variable selection. In the specific context of parameter estimation, PSO has been used to estimate the parameters of a generalized renewal process in order to establish preventive maintenance policies [64], to estimate parameters of mathematical models related to chemical processes [65] and to obtain maximum likelihood estimates for the parameters of a mixture of two Weibull distributions [66].

For a particle j, $j = 1, ..., n_{part}$, we have the following features:

- Current position in the search space (s_j);
- Best position it has visited so far (p_j);

- Velocity (v_j);
- Fitness (f_j), which is the value of the objective function, which in this work is the q-Exponential log-likelihood.

Every particle is a potential solution for the considered optimization problem, which involves a d-dimensional search space with each dimension related to one of the decision variables. Thus, s_j, p_j and v_j are all d-dimensional vectors, whose entries are associated with the decision variables of the problem. In the maximum likelihood optimization problem related to the q-Exponential distribution, d= 2 and the first and second entries of s_j, p_j and v_j are related to η and q, respectively.
The velocity and position update equations are defined as follows:

$$v_{jk}(r+1) = \chi\{v_{jk}(r) + c_1u_1[p_{jk}(r) - s_{jk}(r)] + c_2u_2[p_{gk}(r) - s_{jk}(r)]\} \quad (2.9)$$

$$s_{jk}(r+1) = s_{jk}(r) + v_{jk}(r+1), \quad (2.10)$$

Where r is the iteration number, χ is the constriction factor that avoid velocity explosion during PSO iterations [35], c_1 and c_2 are positive constants, u_1 and u_2 are independent uniform random numbers in [0, 1], and p_{gk} is the k-th entry of vector p_g related to the best position that has been found by any neighbor of particle j.

Whenever an infeasible particle emerges - with respect to the constraints over η and q to assure the probabilistic characteristics of the q-Exponential distribution as well as to the logarithm arguments in the q-Exponential log-likelihood. - its velocity and its position are not altered and its fitness is not evaluated so as to avoid infeasible p_j and p_g. In this way, infeasible particles may become feasible in subsequent iterations due to the influence of their own and neighbor's feasible best positions. This approach is known as "let particles fly" [48].
The update of velocities and positions and fitness evaluations are repeated until one of the following stop criteria is met:

a) Maximum number of iterations (n_{iter}).

b) The global best particle is the same for 10% of n_{iter}. In this case, the iteration number in which the best particle has been found is used, as commented in the previous subsection.

c) The global best fitness values in two consecutive iterations are different, but such a difference is less than a predefined tolerance δ.

2.3.2 Nelder–Mead

The Nelder–Mead method, also known as Downhill Simplex method, is a numerical approach commonly applied to nonlinear optimization. It is used to find the minimum or maximum of an objective function in a multi-dimensional space. This method has been one of the direct search methods most used in unconstrained optimization problem of a function of n variables.It has been used in several studies with the aim of maximizing the log-likelihood function and to estimate the parameters of various probability distributions in many areas such as: Ecology [68]; Medicine [69], [70]; Power Systems [71], [72], and Chemical Engineering [73].

The following characteristics make it one of the most popular methods of optimization[74]:

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

The method 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:

$$\min_{x \in \Re^n} f(x); \text{ Where, f}: \Re^n \longrightarrow \Re.$$

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

In one iteration of the Nelder-Mead method, the $n + 1$ vertices of the simplex, $x_1,\ x_2, \dots,\ x_{n+1}$ belonging to $\Re^n$ are ordered 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_1 is the best vertex and x_{n+1} is the worst vertex.

The repositioning of these vertices takes into consideration four coefficients:

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

These coefficients must satisfy the following restrictions[47]:

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

The default choice of these coefficients is given by: $\rho = 1$, $\chi = 2$, $\gamma = 1/2$ and $\sigma = 1/2$.

The method attempts to replace the worst vertex of the simplex by one with better value. The new vertex is obtained by reflecting, expansion or contraction of the worst vertex along the line through this vertex and the centroid of the best n vertices. At each iteration, the worst vertex is replaced by a new vertex or the simplex is reduced around the better vertex.

The following steps correspond to an interaction of the Nelder-Mead algorithm [47]:

Step 1 - Sort: Sort 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 $f(x_1) \leq f(x_\mathrm{r}) \leq f(x_n)$, then do $x_{n+1} = x_r$ and finalize the iteration.

Step 4- Expansion: If $f(x_\mathrm{r}) \leq f(x_1)$, calculate the expanded vertex (x_e):

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

If $f(x_e) \leq f(x_\mathrm{r})$, 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 $f(x_\mathrm{r}) \geq f(x_\mathrm{n})$

5.1 External:

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

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

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

5.2 Internal:

If $f(x_\mathrm{r}) \geq f(x_{\mathrm{n+1}})$, calculate the internal contraction vertex (x_{ci}):

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

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

Step 6- Reduction: Calculate vectors $v_i = x_1 + \sigma(x_i - x_1)$, $i = 2, \ldots, n+1$.

The vertices (not ordered), for the next iteration are: $x_1, v_2, \ldots, v_{n+1}$.

Given a tolerance Δ_{tol}, the following stop criterion [47] takes into account the function value in the simplex vertices:

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

2.4 Hypothesis Tests

2.4.1 Shapiro-Wilk Test

The Shapiro–Wilk test [75] is a statistical test used in order to verify if a variable is normally distributed.

The hypotheses of the test are the following:

H_0: *The sample is from a normal population*

H_1: *The sample does not come from a normal population*

The test statistic is:

$$W = \frac{(\sum_{i=1}^{n} a_i x_i)^2}{\sum_{i=1}^{n}(x_i - \bar{x})^2}$$

Where, (x_i) corresponds to the sorted sample values - from lowest (x_1) to the largest (x_n); $\bar{x} = (\sum_{i=1}^{n} x_i)/n$ is the sample mean; and the values of a_i are determined as follows:

$$(a_1, a_2, \ldots, a_n) = \frac{m^T V^{-1}}{(m^T V^{-1} V^{-1} m)^{\frac{1}{2}}}$$

Where, $m^T = (m_1, m_2, \ldots, m_n)$ denote the vector of expected values of standard normal order statistics and V is the covariance matrix of those order statistics.

For a given level of significance (α) it is possible to find the W_α value in the table that shows the critical values of Shapiro-Wilk Statistic. This table can be found in the paper published in 1965 by Shapiro and Wilk [75].

For this test, we will reject the null hypothesis if $W < W_\alpha$. If we consider the *p*-value, then the null hypothesis will be rejected if the *p*-value was lower than the level of significance (α).

2.4.2 Student's t-test for Paired Samples

This test allows us to infer on the equality of the averages of two paired samples. Often each case is analyzed twice (before and after a treatment or intervention), forming pairs of observations, whose differences are tested in order to verify if the result is zero or not [76].

Let us consider two dependent samples $X = (x_1, x_2, \dots, x_n)$ and $Y = (y_1, y_2, \dots, y_n)$. Then, once the samples are dependents, we consider that actually we have a sample of couples $(x_1, y_1), (x_2, y_3), \dots, (x_n, y_n)$. Calculating $d_i = x_i - y_i$, for $i = 1,2, \dots, n$ we can define the sample $D = (d_1, d_2, \dots, d_n)$. As assumption of the Student's t-test paired, it is necessary that the samples X and Y are normally distributed, therefore we will have: $D \sim N(\mu_D, {\sigma^2}_D)$.

Before we perform the test we must choose the hypotheses that will be investigated:

Bilateral		Unilateral to Left		Unilateral to Right
$\begin{cases} H_0: \mu_D = 0 \\ H_1: \mu_D \neq 0 \end{cases}$	or	$\begin{cases} H_0: \mu_D = 0 \\ H_1: \mu_D < 0 \end{cases}$	or	$\begin{cases} H_0: \mu_D = 0 \\ H_1: \mu_D > 0 \end{cases}$

The μ_D parameter is estimated by the sample mean of the differences ,*i.e.*, $\overline{D} = \frac{\sum_{i=1}^{n} d_i}{n}$. The ${\sigma^2}_D$ parameter is estimated by the sample variance of the differences ,*i.e.*, ${S^2}_D = \frac{\sum_{i=1}^{n} (d_i - \overline{D})^2}{n-1}$.

The test statistic is:

$$T = \frac{\overline{D} - \mu_D}{S_D / \sqrt{n}}$$

Under H_0, the test statistic follows a Student's *t* distribution with $n - 1$ degrees of freedom. So, for a given level of significance (α) it is possible to find the value of t, with $n - 1$ degrees of freedom in order to compare with statistic T and take the appropriate decision about test. Following, we present the null hypothesis rejection criteria:

Bilateral	Unilateral to Left	Unilateral to Right
If $T > t_{1-(\frac{\alpha}{2})}$ *or* $T < t_{\alpha/2}$ *then, we must reject the null hipothesis.*	*If* $T < t_{\alpha}$ *then, we must reject the null hipothesis.*	*If* $T > t_{1-\alpha}$ *then, we must reject the null hipothesis.*

If we consider the *p*-value, then the null hypothesis will be rejected if the *p*-value was lower than the level of significance (α).

2.4.3 Wilcoxon test

This test is used to verify if the position measurements of two samples are the same . It is applied in the case where the samples are dependent [77]. In the test application, we must consider two dependent samples with sample size n, *i.e.*, $X = (x_1, x_2, \dots, x_n)$ and $Y = (y_1, y_2, \dots, y_n)$. In this case, we consider the paired observations, that is, we consider that we have actually a sample of couple: $(x_1, y_1), (x_2, y_3), \dots, (x_n, y_n)$.

Let us define $d_i = |x_i - y_i|$, for $i = 1,2, \dots, n$ and thus we can obtain the sample $D = (d_1, d_2, \dots, d_n)$. The hypotheses for the test are the following:

Bilateral		Unilateral to Left		Unilateral to Right
$\begin{cases} H_0: \Delta = 0 \\ H_1: \Delta \neq 0 \end{cases}$	or	$\begin{cases} H_0: \Delta = 0 \\ H_1: \Delta < 0 \end{cases}$	or	$\begin{cases} H_0: \Delta = 0 \\ H_1: \Delta > 0 \end{cases}$

In other words, we are testing whether the populations differ in location or not using the following idea: if we accept the null hypothesis, we have that the median of the difference is zero, *i.e.*, populations do not differ in location. If the null hypothesis is rejected, then the median of the difference is not zero, so we have populations that differ in location.

Once the Wilcoxon test is based on the ranks of the values obtained, initially we need to sort the values of the absolute differences, from the lowest to the highest. For each value sorted we associate the post R_i. Then we define the indicator variable, as:

$$\psi_i = \begin{cases} 1, & if\ d_i > 0 \\ 0, & if\ d_i < 0 \end{cases}$$

After this, we must obtain the products $R_i\psi_i$, for *i=1 to n.* So, with these results we must calculate the test statistic by the sum of the products, *i.e.*:

$$T^+ = \sum_{i=1}^{n} R_i\psi_i$$

The test statistic should then be compared with the critical value obtained from the Wilcoxon table. If we consider the p-value, then the null hypothesis will be rejected if the p-value was lower than the level of significance (α).

2.4.4 Bootstrapped Kolmogorov-Smirnov test (K-S Boot)

The one-sample Kolmogorov-Smirnov test (K-S test) is not very useful in practice because it requires a simple null hypothesis, *i.e.*, the distribution must be completely specified with all parameters known beforehand [78]. A bootstrapped version of a K-S test was proposed as alternative to overcome this problem [79]. This method results in accurate asymptotic approximations of the p-values [80]. In this work, we will use this bootstrapped version to check the fit of the q-Exponential distribution to each data set. This method follows the following steps:

- **Step 1**: From an initial sample for the variable $X = \{x_1, x_2, \ldots, x_n\}$, estimate the parameters $\Theta = \{\theta_1, \theta_2, \ldots, \theta_k\}$ and construct the theoretical CDF: $F_n(X, \hat{\Theta})$.
- **Step 2**: Evaluate $D_0 = \max_{1 \le i \le n} \left| \left| \hat{F}_n(x_i) - F_n((x_i, \hat{\Theta}) \right|, \left| \hat{F}_n(x_{i-1}) - F_n((x_i, \hat{\Theta}) \right| \right|$, where $\hat{F}_n(X)$ is the empirical CDF.
- **Step 3**: Use the estimates obtained in the first step to generate new samples for X, *i.e.*: $\{x^*_{1,j},\ x^*_{2,j}, \ldots,\ x^*_{n,j}\}$. Based on these new samples, compute the bootstrap sample estimate of Θ, say $\Theta^*_j = \{\theta^*_{1j}, \theta^*_{2j}, \ldots, \theta^*_{kj}\}$.
- **Step 4**: Repeat step 3, N times; $j = (1, 2, \ldots N)$. The number of bootstrap samples N should be large to ensure a good approximation.

- **Step 5**: Evaluate $D_j^* = \max_{1 \le i \le n} \left| \left| \hat{F}_{n,j}^*\left(x_{i,j}^*\right) - F_{n,j}^*\left(x_{i,j}^*, \hat{\Theta}^*\right) \right|, \left| \hat{F}_{n,j}^*\left(x_{(i-1),j}^*\right) - F_{n,j}^*\left(x_{i,j}^*, \hat{\Theta}^*\right) \right| \right|$;

We reject the null hypothesis if $D_0 > D_{(N(1-\alpha)+1)}^*$ for a significance level α. An approximate p-value can be computed using:

$$p = \frac{\#\{D_j^* \ge D_0\} + 1}{N + 1}$$

Where $\#\{D_j^* \ge D_0\}$ indicates the quantity of D_j^* $(j = 1, \dots N)$ that was larger than D_0.

3 ESTIMATION OF THE *q*-EXPONENTIAL PARAMETERS BY PSO AND NELDER-MEAD METHOD

In this section we propose the PSO and the Nelder-Mead methods as alternatives to estimate the parameters of a q-Exponential distribution. Also confidence intervals will be presented for the q-Exponential parameters based on the approach of bootstrap parametric and non-parametric. Asymptotic confidence intervals also will be performed for the parameters. Comparisons between the proposed methods will be made and a practical application will be presented.

3.1 The Maximum Likelihood Estimator for the *q*-Exponential Distribution

In order to compute the MLE of q and η let $X = \{x_1, x_2, \ldots, x_n\}$ be a random sample of size n. From this sample, it is possible to write the likelihood function for the observed sample as:

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

and the log-likelihood function as:

$$l(x,q,\eta) = n\,ln(2-q) + n\,ln\left(\frac{1}{\eta}\right) + \frac{1}{1-q}\sum_{i=1}^{n} ln\left[1-\frac{(1-q)x_i}{\eta}\right] \tag{3.1}$$

The partial derivatives of the log-likelihood have the following results:

$$\frac{\partial l(x,q,\eta)}{\partial \eta} = \frac{n}{\eta} + \frac{\sum_{i=1}^{n}\frac{(q-1)x_i}{\eta(q-1)x_i+1}}{1-q}$$

$$\frac{\partial l(x,q,\eta)}{\partial q} = \frac{n}{q-2} + \frac{\sum_{i=1}^{n}\frac{\eta x_i}{\eta(q-1)x_i+1}}{1-q} + \frac{\sum_{i=1}^{n}\log(\eta(q-1)x_i+1)}{(1-q)^2}$$

As we can observe from the partial derivatives of the q-Exponential log-likelihood, is very complicated to obtain analytical expressions in order to estimate the q-exponential parameters. Thus, in this work we will use computational algorithms to maximize the log-likelihood function, in order to obtain the estimates for the parameters q and η. We will call these estimates as $\hat{q}$ and $\hat{\eta}$, and they will be obtained by PSO and Nelder-Mead algorithm

3.2 Confidence intervals based on bootstrap methods

For the construction of bootstrap intervals for the parameters q and η of a q-Exponential distribution, we will consider two approaches: bootstrap-p and non-parametric bootstrap[52], [53].

3.2.1 Confidence intervals based on Bootstrap-*p for the q-Exponential parameters*

The following steps are applied in order to construct bootstrap-p based confidence intervals for the parameters q and η:

- **Step 1:** From an initial sample for the variable $X = \{x_1, x_2,..., x_n\}$, estimate the parameters q and η by maximizing the Equation (3.1);
- **Step 2:** Use the estimates obtained in the previous step and Equation (2.2) to generate a new sample X^*, *i.e.* $\{x_1^*, x_2^*, ..., x_n^*\}$. Based on this new sample, compute the bootstrap sample estimates for q and η, say q^* and η^*, by maximizing the Equation (3.1);
- **Step 3:** Repeat step 2 N times;
- **Step 4:** Using the N values of q^* and η^*, obtained in step 3 and by adopting a α significance level, find the quantiles $\alpha/2$ and $1-\alpha/2$ for q^* and η^*. With this information, it is possible to determine an approximate confidence interval for the parameters q and η with confidence equals to 100(1-α)% by the following equations:

$$CI[\eta;(1-\alpha)\cdot 100\%] = [q^*_{\alpha/2}, q^*_{1-(\alpha/2)}]\ , \tag{3.2}$$

$$CI[q;(1-\alpha)\cdot 100\%] = [\eta^*_{\alpha/2}, \eta^*_{1-(\alpha/2)}]. \tag{3.3}$$

3.2.2 Confidence intervals based on Non-Parametric Bootstrap *for the q-* Exponential parameters

The algorithm for constructing confidence intervals for the parameters q and η by using the non-parametric bootstrap approach is as follows:

- **Step 1:** From an initial sample for the variable $X = \{x_1, x_2, \ldots, x_n\}$, generate a new sample X^* by sampling with replacement, *i.e.* $\{x_1^*, x_2^*, \ldots, x_n^*\}$. Based on this new sample, compute the estimates for q and η, say q^* and η^*, by maximizing the Equation (3.1);
- **Step 2:** Repeat step 1, N times.
- **Step 3:** By using the N values of q^* and η^* from step 2 and by adopting α as significance level, find the quantiles $\alpha/2$ and $1-\alpha/2$ for q^* and η^*. Thus, with this information, it is possible to determine by Equations (3.2) and (3.3) an approximate confidence interval for the parameters q and η, with confidence interval equals to $100 * (1-\gamma)\%$.

3.3 Asymptotic confidence intervals for the *q*-Exponential parameters

For the q-Exponential distribution, we have two parameters, *i.e.*, $\hat{\theta} = (\hat{\eta}, \hat{q})$. For this probabilistic model, the covariance matrix associated to the maximum likelihood estimators is as follows:

$$\widehat{\text{var}}(\hat{\eta},\hat{q}) = I^{-1}(\hat{\eta},\hat{q}|t) = -\begin{bmatrix} \frac{\partial^2 L(\hat{\eta},\hat{q}|t)}{\partial \eta^2} & \frac{\partial^2 L(\hat{\eta},\hat{q}|t)}{\partial \eta \partial q} \\ \frac{\partial^2 L(\hat{\eta},\hat{q}|t)}{\partial q \partial \eta} & \frac{\partial^2 L(\hat{\eta},\hat{q}|t)}{\partial q^2} \end{bmatrix}^{-1}, \quad (3.4)$$

In which

$$\frac{\partial^2 L(\hat{\eta},\hat{q}|t)}{\partial \eta^2} = \frac{1}{\hat{\eta}^2}\left[n - \sum_{i=1}^{n}\left(\frac{1}{\frac{\hat{\eta}}{t_i}-1+\hat{q}}\right)\right] - \frac{1}{\hat{\eta}}\sum_{i=1}^{n}\left[\frac{1}{t_i\left(\frac{\hat{\eta}}{t_i}-1+\hat{q}\right)^2}\right], \quad (3.5)$$

$$\frac{\partial^2 L(\hat{\eta},\hat{q}|t)}{\partial \eta \partial q} = \frac{\partial^2 \mathcal{L}(\hat{\eta},\hat{q}|t)}{\partial q \partial \eta} = -\frac{1}{\hat{\eta}}\sum_{i=1}^{n}\left[\frac{1}{\left(\frac{\hat{\eta}}{t_i}-1+\hat{q}\right)^2}\right], \quad (3.6)$$

$$\frac{\partial^2 L(\hat{\eta},\hat{q}|t)}{\partial q^2} = -\frac{n}{(2-\hat{q})^2} + \frac{2}{(1-\hat{q})^3}\sum_{i=1}^{n}\log\left[1-(1-\hat{q})\frac{t_i}{\hat{\eta}}\right] + \frac{2}{(1-\hat{q})^2}\sum_{i=1}^{n}\left(\frac{1}{\frac{\hat{\eta}}{t_i}-1+\hat{q}}\right) - \frac{1}{1-\hat{q}}\sum_{i=1}^{n}\left[\frac{1}{\left(\frac{\hat{\eta}}{t_i}-1+\hat{q}\right)^2}\right]. \quad (3.7)$$

Once the covariance matrix is estimated, asymptotic confidence intervals can be constructed for the q-Exponential parameters by using the asymptotic normality property of the maximum likelihood estimators. The asymptotic $(1-\alpha)\cdot 100\%$ confidence intervals for η and q are given by, respectively:

$$\text{CI}[\eta;(1-\alpha)\cdot 100\%] = \left[\hat{\eta} + z_{\frac{\alpha}{2}}\sqrt{\widehat{\text{var}}_{11}},\ \hat{\eta} + z_{1-\left(\frac{\alpha}{2}\right)}\sqrt{\widehat{\text{var}}_{11}}\right], \quad (3.8)$$

$$\text{CI}[q;(1-\alpha)\cdot 100\%] = \left[\hat{q} + z_{\frac{\alpha}{2}}\sqrt{\widehat{\text{var}}_{22}},\ \hat{q} + z_{1-\left(\frac{\alpha}{2}\right)}\sqrt{\widehat{\text{var}}_{22}}\right], \quad (3.9)$$

In which $z_{\alpha/2}$ and $z_{1-\alpha/2}$ are the $\alpha/2$ and 1 - $\alpha/2$ quantiles of the standard normal distribution and $\widehat{\text{var}}_{11}$, $\widehat{\text{var}}_{22}$ are the diagonal elements of the covariance matrix presented in Equation (3.4).

3.4 Numerical experiments

In this section, some simulations will be presented in order to assess the quality of estimates of q-Exponential parameters obtained from PSO and Nelder-Mead methods. The PSO was implemented in the MATLAB computer software [81] and the Nelder-Mead was performed by the function Optim of the software R [82]. We will consider sample sizes equal to 100, 500 and 1000, and simulations for point and interval estimates will be presented.

In previous simulations, we detected that in situations where the parameter q is less than 0, the simulations showed results for the estimates that are very different from the parameters. Table 3.1 presents these estimates obtained maximizing the log-likelihood function of the q-Exponential distribution, in situations that q is less than zero. The results are presented considering the average of 1000 estimates of each parameter. Observe that the results for the estimates of η and q are very different from the parameters, mainly when the sample size is small (*n=100*).

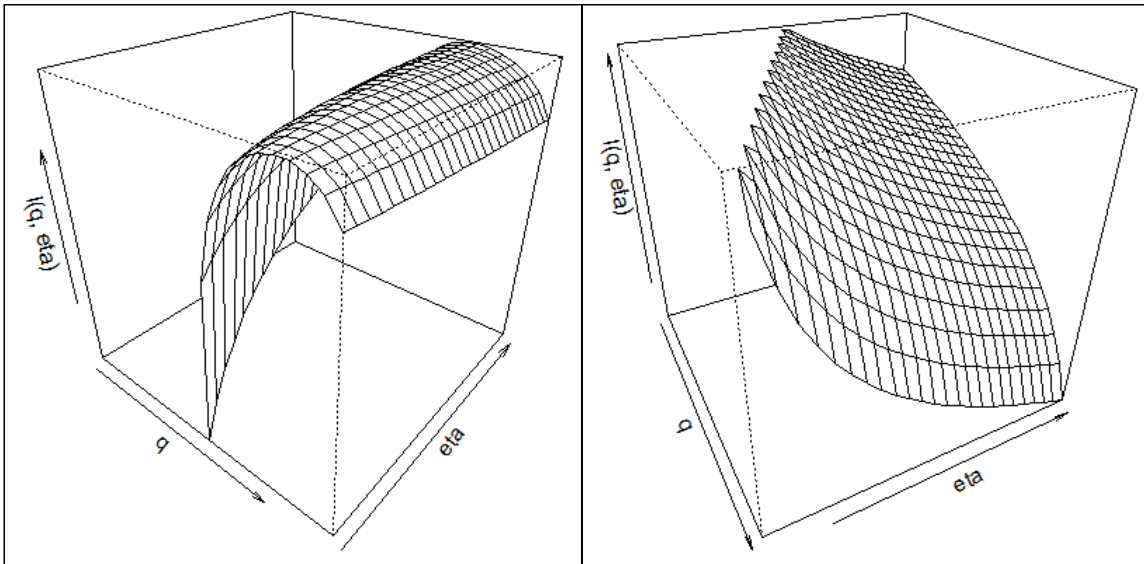
Table 3.1. Results for estimates of the parameters $\boldsymbol{\eta}$ *and q obtained maximizing the Log-likelihood function of the q-Exponential distribution – simulations for q<0*

(η, q)	n	$\bar{\hat{\eta}}$	$\bar{\hat{q}}$	Bias($\hat{\eta}$)	Bias($\hat{q}$)	MSE($\hat{\eta}$)	MSE($\hat{q}$)
(5, -1)	100	56558.45	-24103.9	56553.45	-24102.88	5.80E+11	1.04E+11
	500	5.577235	-1.24941	0.5772	-0.24941	2.29256	3.91E-01
	1000	5.753486	-1.31194	0.7535	-0.31194	3.99648	6.49E-01

In order to investigate what might be causing this problem, we have built the graphics of the likelihood function in a situation that q parameter is greater than 0 and also otherwise. As we can observe from Figure 3.1 (*b*), when the parameter q is less than 0, the graphic of the log-likelihood function present a monotonic behavior, i.e., we cannot found a maximum value for the parameters q and η. This behavior does not occur when the value of q is greater than 0; see Figure 3.1 (*a*). Thus, for the case where the parameter q is less than 0, we propose a reparametrization of the log-likelihood function. In fact, we make the following reparametrization in the parameters of the q-Exponential:

$$q = 2 - abs\left(\frac{q'}{35}\right) \tag{3.10}$$

$$\eta = \exp(\eta'). \tag{3.11}$$



(*a*)

***n* = 100**
Parameters:
$q = 1.5$ and $\eta = 5$
Maximum of the Graphic:
$q = 1.5042$ and $\eta = 5$
Estimates of Maximum Likelihood:
$\hat{q} = 1.4850$ and $\hat{\eta} = 5.3605$

(*b*)

***n* = 100**
Parameters:
$q = -1$ and $\eta = 5$
Maximum of the Graphic:
$q = -19.1667$ and $\eta = 51$
Estimates of Maximum Likelihood:
$\hat{q} = -0.2117$ and $\hat{\eta} = 2.9666$

Figure 3.1. Graphics for the log-likelihood Function – (a) parameter q is larger than 0 and (b) parameter q is less than 0.

Observe that the reparametrization was made considering the support of the parameters, i.e., the parameter q must be always less than 2, thus we consider a reparametrization that the function must be always less than 2. As the same manner, the parameter η must be always larger than 0, so we consider a function that always returns values greater than 2. Thus, we will maximize the following function:

$$l'(x,q',\eta') = n\,ln\left\{2-\left[2-abs\left(\frac{q'}{35}\right)\right]\right\} + n\,ln\left(\frac{1}{\exp\left(\eta'\right)}\right)$$
$$+\frac{1}{1-\left[2-abs\left(\frac{q'}{35}\right)\right]}\sum_{i=1}^{n} ln\left(1-\frac{\left\{1-\left[2-abs\left(\frac{q'}{35}\right)\right]\right\}x_i}{\exp(\eta')}\right)$$

After determine the value of q' and η', we can obtain the value of q and η by equations (3.10) and (3.11). Table 3.2 shows the results of the estimates of the parameters obtained maximizing the log-likelihood function of the q-Exponential distribution after a proposed reparametrization, in situations that q is less than 0. The results are presented considering the average of 1000 estimates of each parameter. Observe that the results for the estimates of η and q are still different from the parameters, but compared with the results without the reparametrization, the estimates improved vastly. In this way, for the rest of this work we will use the proposed reparametrization when we deal with q less than 0.

Table 3.2. Results for estimates of the parameters q and $\boldsymbol{\eta}$ *obtained maximizing the Log-likelihood function of the q-Exponential distribution after a reparametrization – simulations for q<0*

(η, q)	n	$\bar{\hat{\eta}}$	$\bar{\hat{q}}$	Bias($\hat{\eta}$)	Bias($\hat{q}$)	MSE($\hat{\eta}$)	MSE($\hat{q}$)
(5, -1)	100	7.5103	-2.1153	2.5103	-1.1153	33.1753	6.1975
	500	5.3737	-1.16683	0.3737	-0.16683	9.22E-01	1.64E-01
	1000	5.19129	-1.08581	0.1913	-0.08581	3.34E-01	5.78E-02

3.4.1 Point Estimates (MLE)

For the analysis of the MLE we consider the following sets of initial parameters: (η, q) = (5, 1.5), (5, 1), (5, 0.5), and (5, -1). These sets are chosen in order to consider the four important situations for the parameter q: $1 < q < 2$; $q \to 1$; $0 < q < 1$ and $q < 0$. For each set of parameters we generate, by Equation (2.2), 1000 samples. Thus, for each one of these samples, we obtained estimates for q and η by PSO and Nelder-Mead algorithms, which resulted in a total of 1000 estimates for each parameter. Figure 3.2 presents the steps that we follow to obtain the results of the numerical experiments for point estimates:

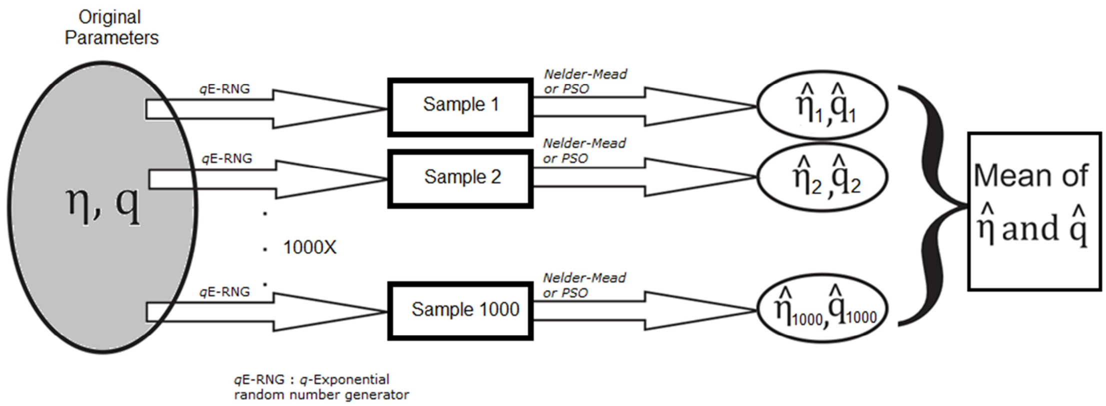


Figure 3.2. Steps of numerical experiment (point estimates).

Table 3.3 and Table 3.4 present the mean of these 1000 estimates for each parameter ($\hat{\eta}$ and $\hat{q}$). Moreover, we reported in Table 3.3 and Table 3.4 the related bias and MSE. For an estimator $\hat{\theta}$ of θ, we have that bias$(\hat{\theta}) = E(\hat{\theta}) - \theta$ and MSE $= Var(\hat{\theta}) + bias(\hat{\theta})^2$.

Table 3.3 - Simulation results for 1000 replications – PSO Case.

Cases			**PSO**					
	(η, q)	n	$\bar{\hat{\eta}}$	$\bar{\hat{q}}$	Bias($\hat{\eta}$)	Bias($\hat{q}$)	MSE($\hat{\eta}$)	MSE($\hat{q}$)
1	(5, 1.5)	100	5.3316	1.4879	0.3316	−0.0121	2.1541	0.0029
2		500	5.0713	1.4982	0.0713	−0.0018	0.3700	0.0005
3		1000	5.0432	1.4983	0.0432	−0.0017	0.1743	0.0003
4	(5, 1)	100	5.4545	0.9546	0.4545	−0.0454	2.2459	0.0199
5		500	5.1078	0.9886	0.1078	−0.0114	0.3042	0.0025
6		1000	5.0566	0.9954	0.0566	−0.0046	0.1424	0.0012
7	(5, 0.5)	100	5.7744	0.3627	0.7744	−0.1373	3.4889	0.0952
8		500	5.1240	0.4767	0.1240	−0.0233	0.3339	0.0080
9		1000	5.0731	0.4867	0.0731	−0.0133	0.1444	0.0034
10	(5, -1)	100	7.8208	−2.2482	2.8208	−1.2482	54.1635	10.064
11		500	5.3779	−1.1686	0.3779	−0.1686	0.9296	0.1657
12		1000	5.1939	−1.0869	0.1939	−0.0869	0.3369	0.0583

Table 3.4 - Simulation results for 1000 replications – Nelder-Mead Case.

Cases			Nelder-Mead					
	(η, q)	n	$\bar{\hat{\eta}}$	$\bar{\hat{q}}$	Bias($\hat{\eta}$)	Bias($\hat{q}$)	MSE($\hat{\eta}$)	MSE($\hat{q}$)
1	(5, 1.5)	100	5.3301	1.4879	0.3301	-0.0121	2.1536	0.0029
2		500	5.0702	1.4982	0.0702	-0.0018	0.3705	0.0005
3		1000	5.0431	1.4983	0.0431	-0.0017	0.1743	0.0003
4	(5, 1)	100	4.8846	0.9892	-0.1154	-0.0108	1.2159	0.0083
5		500	4.6887	0.9994	-0.3113	-0.0006	0.2247	0.0005
6		1000	4.6312	1.0006	-0.3688	0.0006	0.1770	0.0002
7	(5, 0.5)	100	5.7728	0.3630	0.7728	-0.1370	3.4889	0.0951
8		500	5.1239	0.4767	0.1239	-0.0233	0.3338	0.0080
9		1000	5.0729	0.4867	0.0729	-0.0133	0.1445	0.0034
10	(5, -1)	100	7.5103	-2.1153	2.5103	-1.1153	33.1753	6.1975
11		500	5.3737	-1.1668	0.3737	-0.1668	0.9220	0.1644
12		1000	5.1913	-1.0858	0.1913	-0.0858	0.3342	0.0579

From Table 3.3 and Table 3.4, we can observe that the estimates obtained by PSO and Nelder-Mead algorithms present bias and MSE that decrease as the sample size (n) increases, corroborating the consistency of the MLE. However, both PSO and Nelder-Mead approach, shows negative results for the bias of $\hat{q}$ (exception: Case 6 – Nelder-Mead), indicating that in both method the q parameter is underestimated. In order to provide a best visualization of the results presented in tables 3.1 and 3.2 of the numerical experiments, we present a graphical comparison between the point estimates obtained by PSO and Nelder-Mead and the real values of the parameters (q and η). Besides, we make some graphics to evaluate the MSE obtained by the two optimization methods for each parameter of interest.

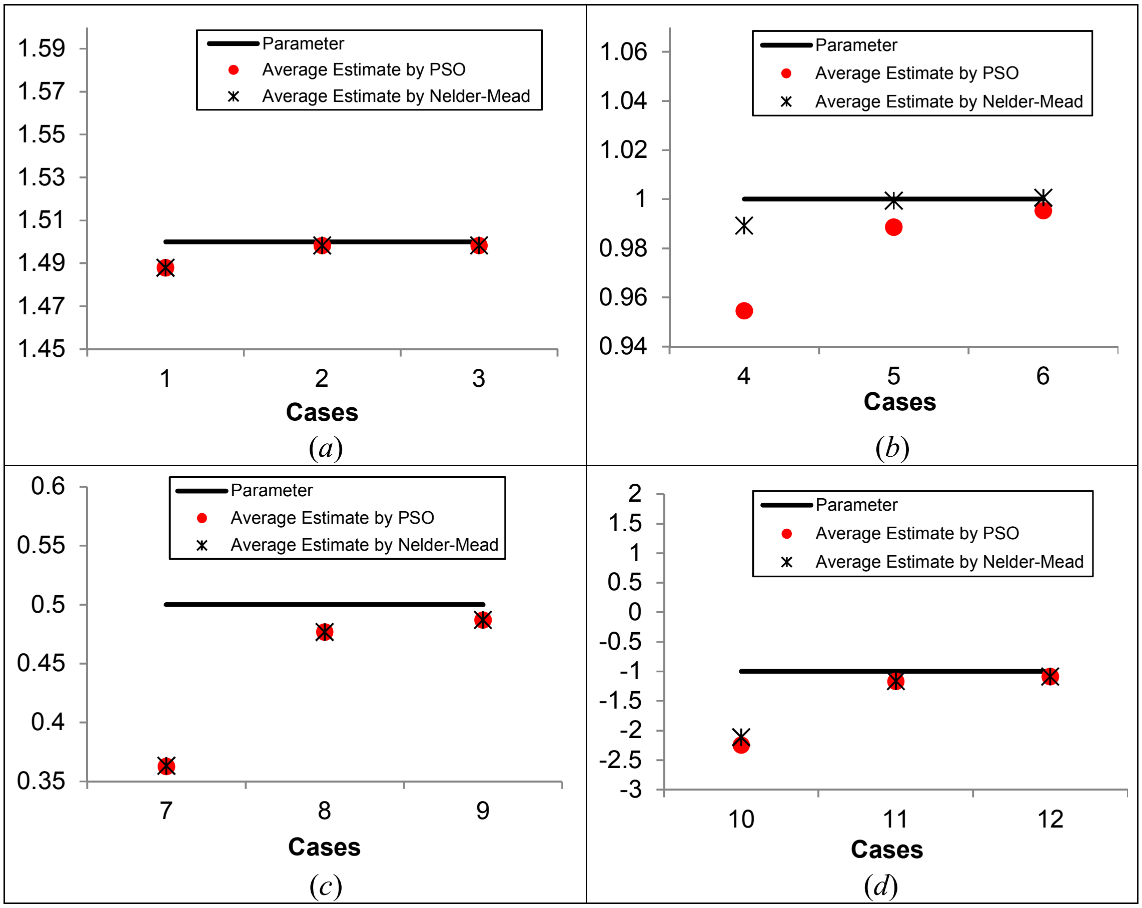
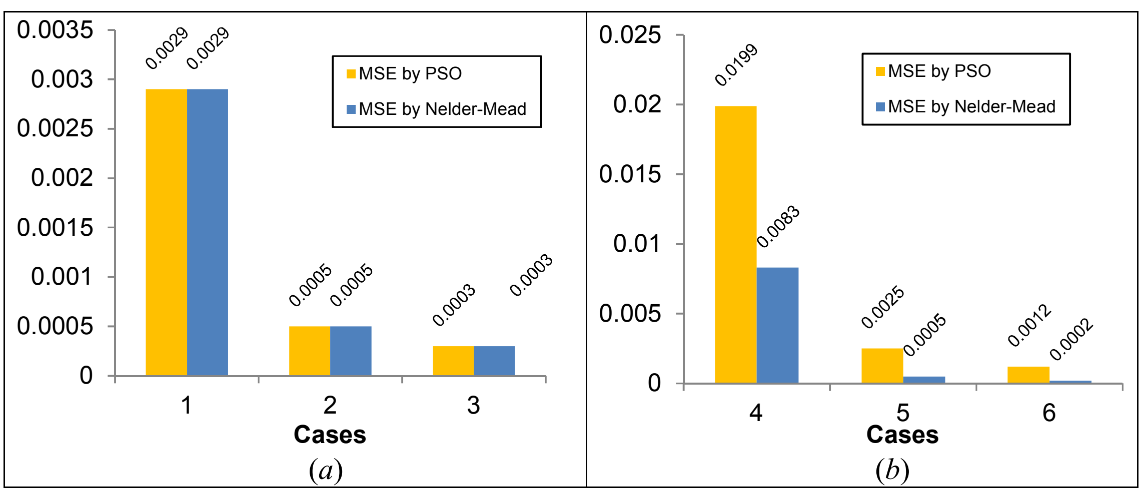


Figure 3.3. Graphical comparison between results of the estimates of the q parameter obtained by PSO and Nelder-Mead, and the true value of the parameter - (a) Cases 1 to 3, (b) Cases 4 to 6, (c) Cases 7 to 9 and (d) Cases 10 to 12.



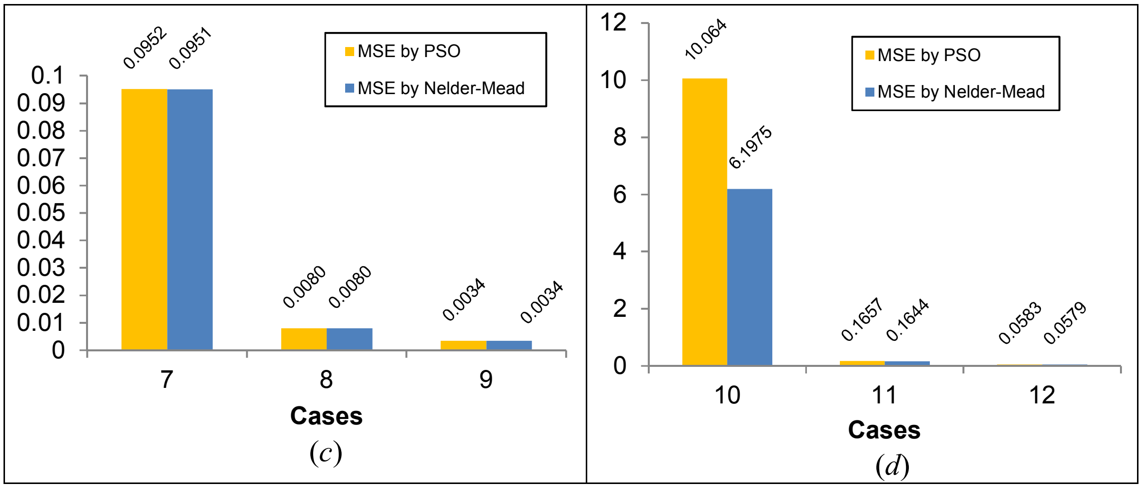


Figure 3.4.Mean Squared Error (MSE) for the estimates of the q parameter obtained by PSO and Nelder-Mead - (a) Cases 1 to 3, (b) Cases 4 to 6, (c) Cases 7 to 9 and (d) Cases 10 to 12.

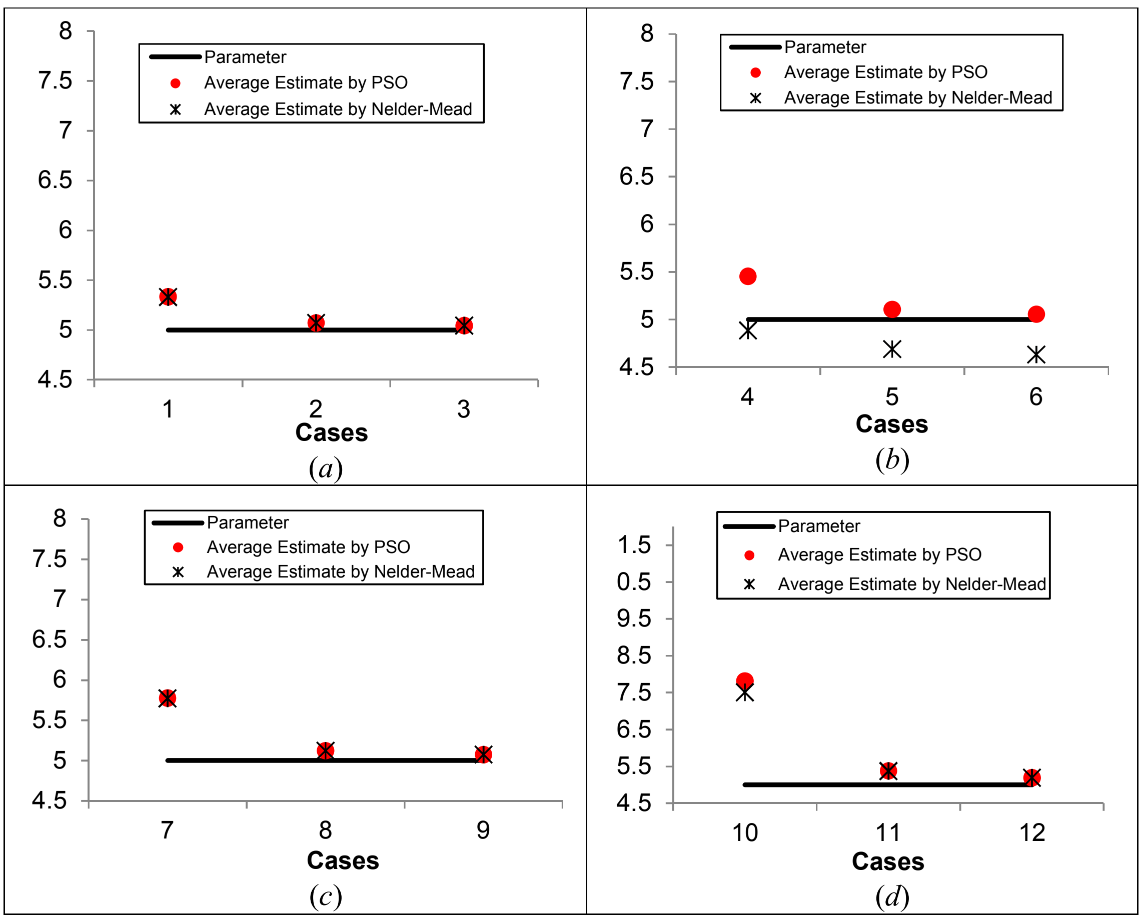


Figure 3.5. Graphical comparison between results of the estimates of the η parameter obtained by PSO and Nelder-Mead, and the true value of the parameter - (a) Cases 1 to 3, (b) Cases 4 to 6, (c) Cases 7 to 9 and (d) Cases 10 to 12.

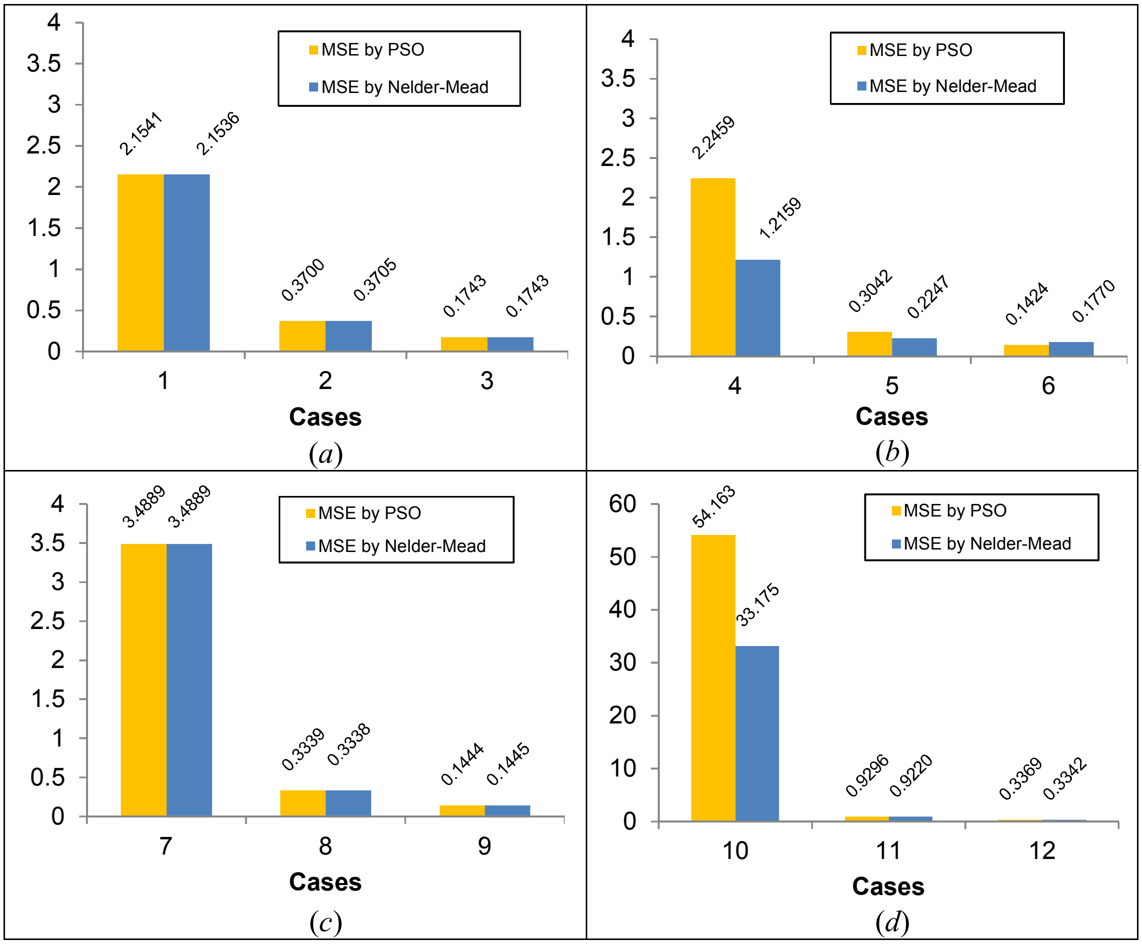


Figure 3.6. Mean Squared Error (MSE) for the estimates of the $\boldsymbol{\eta}$ *parameter obtained by PSO and Nelder-Mead - (a) Cases 1 to 3, (b) Cases 4 to 6, (c) Cases 7 to 9 and (d) Cases 10 to 12.*

About the estimation of the q parameter, we can comment that when $q > 1$ (cases 1 to 3), the results of the MSE for the two optimization algorithms were absolutely equal, indicating that when $q > 1$, the two optimization method has similar performance; In cases 4 to 6, the PSO approach present MSE larger than the obtained by Nelder-Mead, indicating a better efficiency of the Nelder-Mead method when we deal with cases that present $q \to 1$. When $0 < q < 1$ (cases 7 to 9), both optimization methods present similar results. In cases 10 to 12, where $q < 0$, it is observed a great value for the MSE in both optimization methods. In this situation, the PSO method presents the worst performance in the estimation of the q parameter when *n=100*, and for *n=500 and 1000* the results for both optimization methods are very close. We can highlight that this great values for the MSE (when $q = -1$) is observed mainly when the n is smaller ($n = 100$). Observe that in cases 10, 11 and 12, when n increases, the bias and MSE decrease significantly.

When we estimate the η parameter, it is possible to observe that for cases 1 to 3 ($q > 1$), the MSE values are very close for the two optimization methods. Also, in cases 4 to 6, we observe a similar behavior for the performance of both optimization methods, with a slight advantage, for the Nelder-Mead method in cases 4 and 5. Cases 7 to 9 present a similar behavior for the two methods. With respect to cases 10 to 12, where we have the negative value for q, it is observed a MSE very high to case 10, where $n = 100$, in this case the PSO present the biggest MSE. However, observe that as n increases the MSE decreases substantially, this fact is observed for both optimization methods. With relation to the bias of $\hat{\eta}$ we perceive that in both optimization methods, the estimates present positive bias, indicating an overestimation of the η parameter. This related fact is just not observed in the cases that $\eta = 5$, and $q \to 1$ (Cases 4 to 6 - Nelder-Mead).

According to the results, we can yet note, that when the q parameter is closer of its maximum supported value, *i.e.*, $q = 2$, more consistent are the estimates of the parameters. When we deal with q negative, then, the PSO and Nelder-Mead method present high values for bias and MSE of the estimates, indicating less consistency for the two methods in this situation, this fact is mainly observed when the n is small.

3.4.2 Interval Estimates

For the interval estimates, Table 3.5 to Table 3.8 present the simulation results for the asymptotic, bootststrap-p and non-parametric bootstrap 90% confidence intervals. For all these methods, we consider the same sets of initial parameters used in the simulations for the point estimates, *i.e.*, (η, q) = (5, 1.5), (5, 1), (5, 0.5), and (5, -1). Similarly, the same sample sizes are considered: 100, 500 and 1000.

In the asymptotic confidence interval approach, we first generate a sample based in each parameter analyzed by using Equation (2.2). Then, for each sample, we estimate the parameters η and q by PSO and Nelder-Mead algorithms and, after estimating the covariance matrix (Equation (3.4)), we construct the intervals by Equations (3.8) and (3.9).

In the simulations for bootstrap-p and non-parametric methods, we generate by Equation (2.2) a sample for each set of parameters. Then, we estimate the parameters η and q by PSO and Nelder-Mead algorithms. Thus, for the bootstrap-p method, using the estimates obtained from the initial samples, we generate 1000 new samples by Equation (2.2), which allows us to calculate 1000 new estimates for η and q considering all sets of parameters. With this information, we are able to calculate confidence intervals by Equations (3.2) and (3.3). In the case of the non-parametric bootstrap confidence intervals, the intervals are obtained in a similar fashion. The difference is that for the non-parametric case, we use the initial samples (generated by Equation (2.2)) from each set of parameters) in order to generate $N = 1000$ samples by sampling with replacement, which allows us to calculate 1000 new estimates for η and q considering all sets of parameters. Then, we are able to calculate the confidence interval by Equations (3.2) and (3.3). Figure 3.7 presents the steps that we follow to obtain the results of the numerical experiments for Parametric Bootstrap confidence interval (*a*) and Non-parametric Bootstrap confidence interval (*b*). The simulation results are presented in

Table ***3.5*** to

Table ***3.8***.

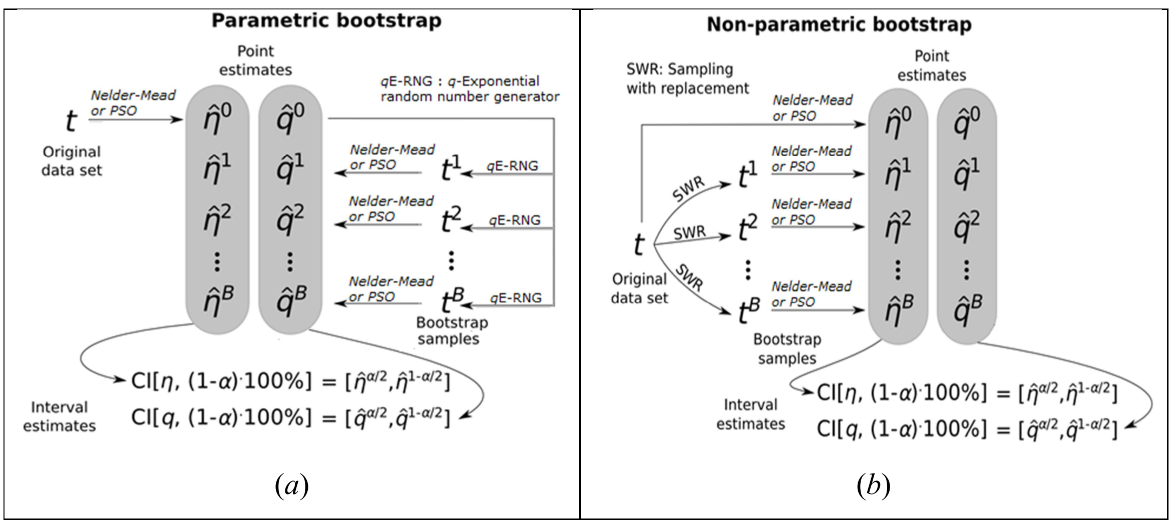


Figure 3.7. Steps of numerical experiment for Parametric Bootstrap confidence interval (a) and Non-parametric Bootstrap confidence interval (b).

Table 3.5 - Asymptotic and bootstrap confidence intervals with 90% of confidence level for $\boldsymbol{\eta}$ *via PSO.*

(η, q)	n	η					
		Asymptotic Confidence Interval	**Width**	**Bootstrap-P Confidence Interval**	**Width**	**Non-Parametric Bootstrap Confidence Interval**	**Width**
(5, 1.5)	100	[3.9179, 10.0493]	6.1314	[4.6383, 11.2618]	6.6235	[4.7283, 11.2664]	6.5381
	500	[3.4869, 5.3166]	1.8297	[3.6984, 5.4655]	1.7672	[3.6053, 5.4974]	1.8922
	1000	[4.4125, 5.8216]	1.4091	[4.4704, 5.8704]	1.4000	[4.4456, 5.9302]	1.4846
(5, 1)	100	[2.5587, 5.9012]	3.3425	[3.0164, 6.9189]	3.9026	[2.8231, 6.5363]	3.7132

(η, q)	n	**η**					
		Asymptotic Confidence Interval	**Width**	**Bootstrap-P Confidence Interval**	**Width**	**Non-Parametric Bootstrap Confidence Interval**	**Width**
	500	[4.2223, 5.9264]	1.7041	[4.3878, 6.0363]	1.6485	[4.3978, 5.9974]	1.5995
	1000	[4.2168, 5.2148]	0.9980	[4.2107, 5.3917]	1.1810	[4.1897, 5.3920]	1.2023
(5, 0.5)	100	[2.3372, 7.5428]	5.2056	[3.4163, 8.7365]	5.3202	[2.6703, 9.3328]	6.6625
	500	[4.2155, 5.8974]	1.6819	[4.3751, 6.2051]	1.8300	[4.3740, 6.1803]	1.8062
	1000	[4.3485, 5.5088]	1.1603	[4.4700, 5.6036]	1.1336	[4.3810, 5.5552]	1.1742
(5, -1)	100	[0.9412, 8.7827]	7.8415	[3.2894, 20.8420]	17.5526	[1.9793, 116.446]	114.467
	500	[4.0571, 7.4123]	3.3552	[4.7521, 8.5209]	3.7688	[4.3313, 8.4578]	4.1265
	1000	[4.1037, 5.6756]	1.5119	[4.2120, 6.1238]	1.9118	[4.3234, 6.3180]	1.9946

Table 3.6 - Asymptotic and bootstrap confidence intervals with 90% of confidence level for ***q*** *via PSO.*

(η, q)	n	**q**					
		Asymptotic Confidence Interval	**Width**	**Bootstrap-P Confidence Interval**	**Width**	**Non-Parametric Bootstrap Confidence Interval**	**Width**
(5, 1.5)	100	[1.3831, 1.5603]	0.1772	[1.3643, 1.5423]	0.1780	[1.3679, 1.5368]	0.1690
	500	[1.4777, 1.5523]	0.0746	[1.4732, 1.5467]	0.0735	[1.4784, 1.5454]	0.0670
	1000	[1.4650, 1.5181]	0.0531	[1.4617, 1.5160]	0.0543	[1.4625, 1.5152]	0.0527
(5, 1)	100	[0.9425, 1.2565]	0.3140	[0.8573, 1.2064]	0.3492	[0.8788, 1.2249]	0.3462
	500	[0.9133, 1.0662]	0.1529	[0.9031, 1.0525]	0.1494	[0.9054, 1.0477]	0.1423
	1000	[0.9810, 1.0703]	0.0893	[0.9625, 1.0692]	0.1066	[0.9537, 1.0766]	0.1230
(5, 0.5)	100	[-0.0079, 0.9059]	0.9138	[-0.2494, 0.6900]	0.9394	[-0.2128, 0.8386]	1.0514
	500	[0.3704, 0.6210]	0.2506	[0.3133, 0.5908]	0.2775	[0.3123, 0.5960]	0.2837
	1000	[0.4318, 0.6058]	0.1740	[0.4125, 0.5870]	0.1745	[0.4274, 0.5990]	0.1716
(5, -1)	100	[-2.7396, 0.7420]	3.4816	[-8.0361, -0.3534]	7.6828	[-48.684, 0.3038]	48.9878
	500	[-2.0155,-0.6176]	2.6298	[-2.4919, -0.9186]	1.5732	[-2.4383, -0.7336]	1.7046
	1000	[-1.2703, -0.6351]	1.9054	[-1.4685, -0.6804]	0.7881	[-1.5636, -0.7222]	0.8414

Table 3.7 - Asymptotic and bootstrap confidence intervals with 90% of confidence level for ***η*** *via Nelder-Mead.*

(η, q)	n	**η**					
		Asymptotic Confidence Interval	**Width**	**Bootstrap-P Confidence Interval**	**Width**	**Non-Parametric Bootstrap Confidence Interval**	**Width**
(5, 1.5)	100	[3.9256, 10.0374]	6.1118	[4.6131, 10.8467]	6.2337	[4.6935, 11.0487]	6.3552
	500	[3.4907, 5.3152]	1.8245	[3.6496, 5.4025]	1.7530	[3.5874, 5.4404]	1.8529
	1000	[4.4143, 5.8187]	1.4045	[4.4677, 5.9013]	1.4336	[4.4413, 5.9150]	1.4737
(5, 1)	100	[2.5639, 5.8982]	3.3344	[3.0158, 6.7117]	3.6959	[2.7862, 6.7110]	3.9248
	500	[4.2277, 5.9282]	1.7005	[3.5497, 5.0196]	1.4699	[4.3406, 5.9994]	1.6589
	1000	[4.2170, 5.2120]	0.9950	[4.1922, 5.3889]	1.1967	[4.2328, 5.4093]	1.1765
(5, 0.5)	100	[2.3442, 7.5343]	5.1902	[3.5749, 9.1427]	5.5678	[2.7364, 9.7146]	6.9783
	500	[4.2180, 5.8934]	1.6754	[4.3606, 6.2027]	1.8420	[4.3698, 6.1727]	1.8029

	1000	[4.3508, 5.5072]	1.1564	[4.4230, 5.6383]	1.2153	[4.3848, 5.5661]	1.1813
(5, -1)	100	[3.1257, 6.6048]	3.4791	[3.1602, 16.6136]	13.4534	[2.0547, 32.6716]	30.6169
	500	[5.0395, 6.4317]	1.3922	[4.6128, 8.2581]	3.6454	[4.2517, 8.0911]	3.8394
	1000	[4.5740, 5.2072]	0.6332	[4.2187, 6.0602]	1.8416	[4.2670, 6.2620]	1.9950

Table 3.8 - Asymptotic and bootstrap confidence intervals with 90% of confidence level for q via Nelder-Mead.

(η, q)	n	***q***					
		Asymptotic Confidence Interval	**Width**	**Bootstrap-P Confidence Interval**	**Width**	**Non-Parametric Bootstrap Confidence Interval**	**Width**
(5, 1.5)	100	[1.3834, 1.5601]	0.1767	[1.3603, 1.5402]	0.1798	[1.3679, 1.5372]	0.1693
	500	[1.4777, 1.5521]	0.0743	[1.4756, 1.5454]	0.0697	[1.4801, 1.5462]	0.0660
	1000	[1.4650, 1.5180]	0.0530	[1.4623, 1.5164]	0.0541	[1.4627, 1.5167]	0.0540
(5, 1)	100	[0.9429, 1.2560]	0.3131	[0.7270, 1.1266]	0.3996	[0.8700, 1.2285]	0.3586
	500	[0.9132, 1.0657]	0.1525	[0.9011, 1.0634]	0.1623	[0.9081, 1.0511]	0.1430
	1000	[0.9812, 1.0704]	0.0891	[0.9372, 1.0489]	0.1117	[0.9538, 1.0730]	0.1192
(5, 0.5)	100	[-0.0064, 0.9047]	0.9112	[-0.2600, 0.6564]	0.9165	[-0.3137, 0.8348]	1.1485
	500	[0.3709, 0.6205]	0.2496	[0.3228, 0.5903]	0.2675	[0.3127, 0.6024]	0.2897
	1000	[0.4320, 0.6053]	0.1733	[0.4032, 0.59142]	0.1882	[0.4226, 0.5971]	0.1745
(5, -1)	100	[-4.9182, 2.9180]	7.8361	[-6.2167, -0.2958]	5.9208	[-13.5112, 0.2433]	13.7545
	500	[-2.9879, 0.3538]	3.3417	[-2.3651, -0.8659]	1.4993	[-2.2837, -0.7130]	1.5707
	1000	[-1.7366, -0.1696]	1.5671	[-1.4357, -0.6838]	0.7519	[-1.5464, -0.7004]	0.8460

From the results presented in Tables Table ***3.5*** toTable ***3.8*** we make some graphics in order to provide a better visualization of the results of the confidence intervals obtained by numerical experiments:

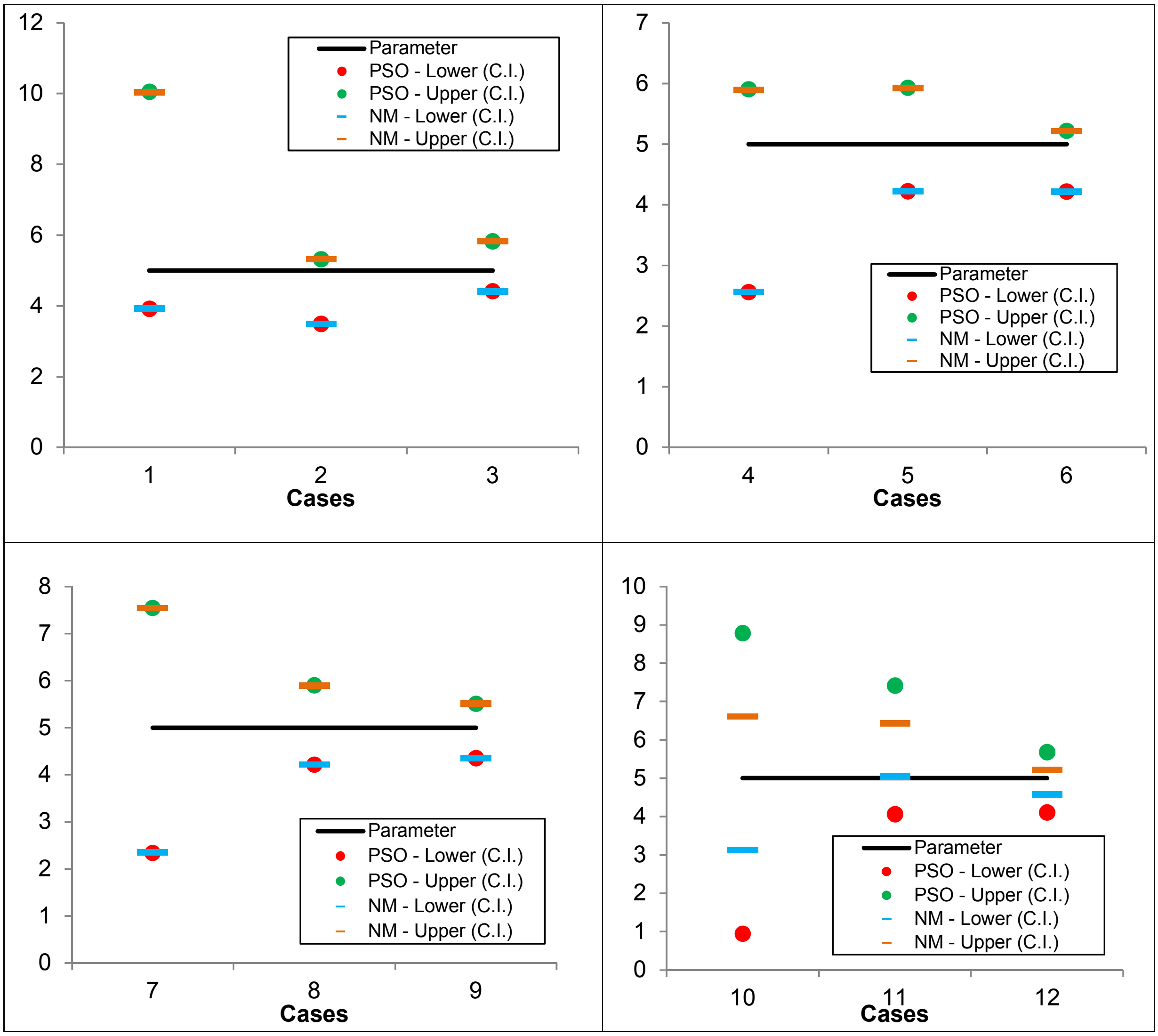


Figure 3.8. Upper and lower limits of the asymptotic Confidence Interval for the parameter $\boldsymbol{\eta}$, obtained by PSO and Nelder-Mead.

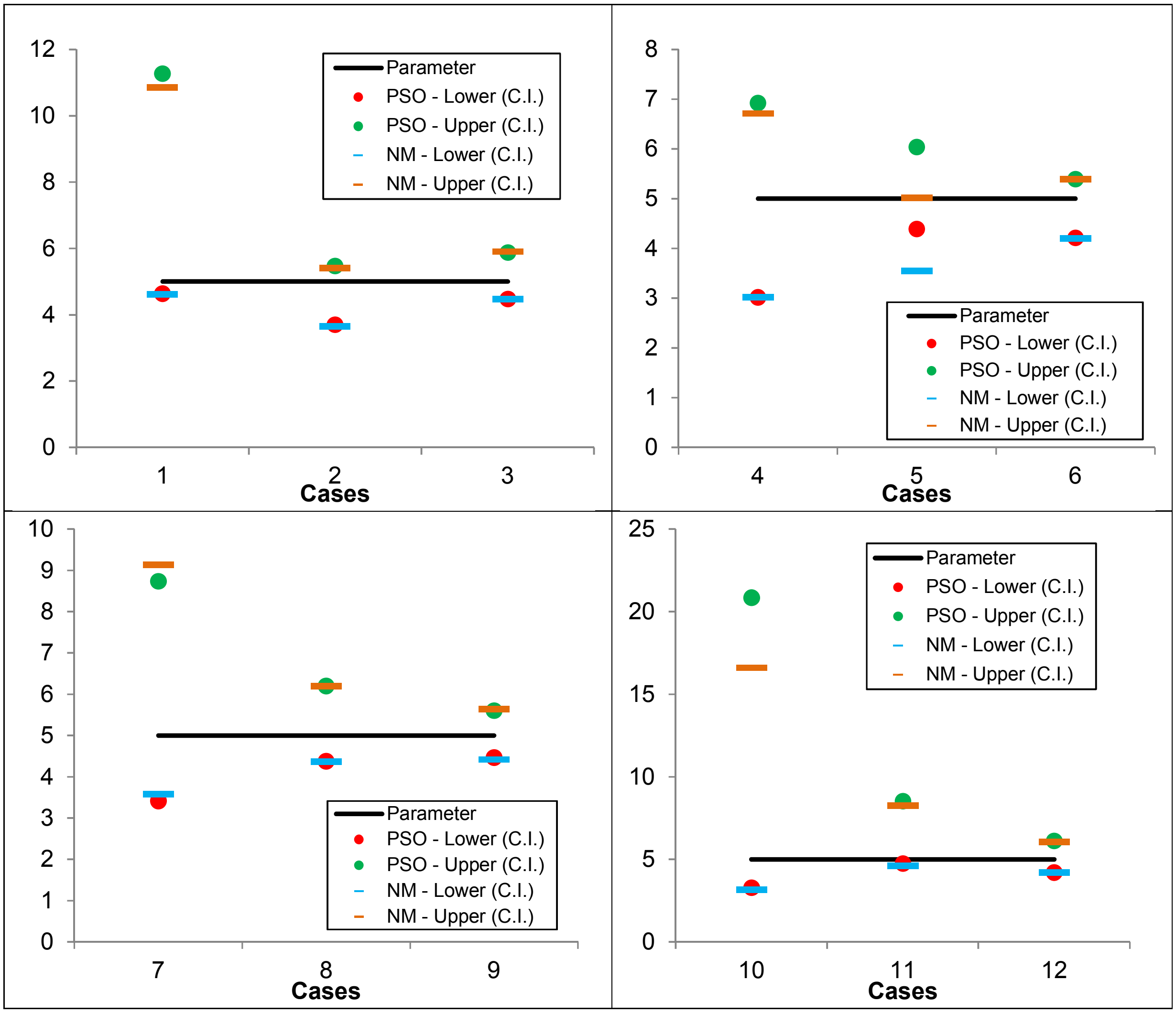


Figure 3.9. Upper and lower limits of the Bootstrap-P Confidence Interval for the parameter $\boldsymbol{\eta}$*, obtained by PSO and Nelder-Mead.*

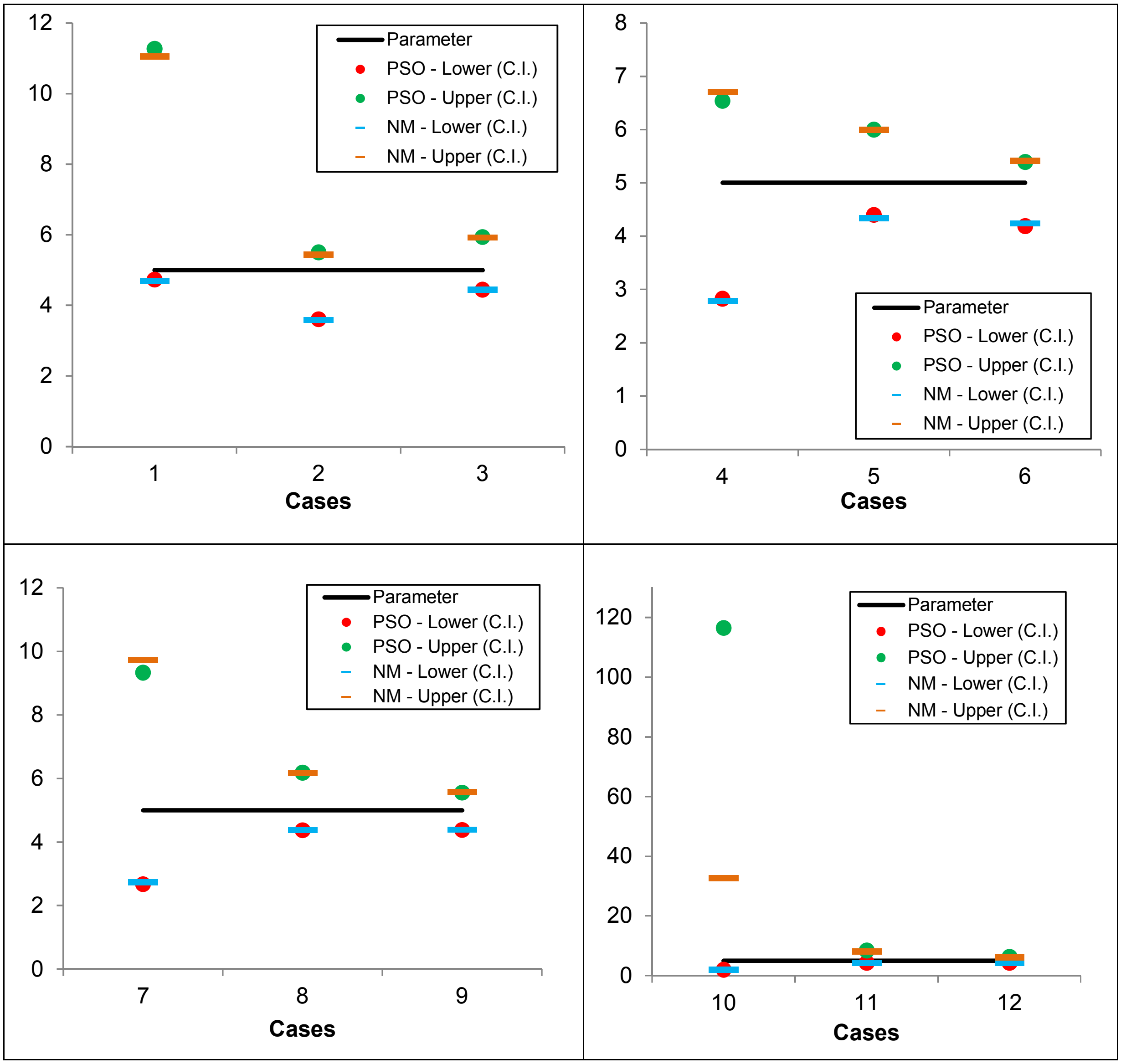


Figure 3.10. Upper and lower limits of the Non-Parametric Bootstrap Confidence Interval for the parameter $\boldsymbol{\eta}$*, obtained by PSO and Nelder-Mead.*

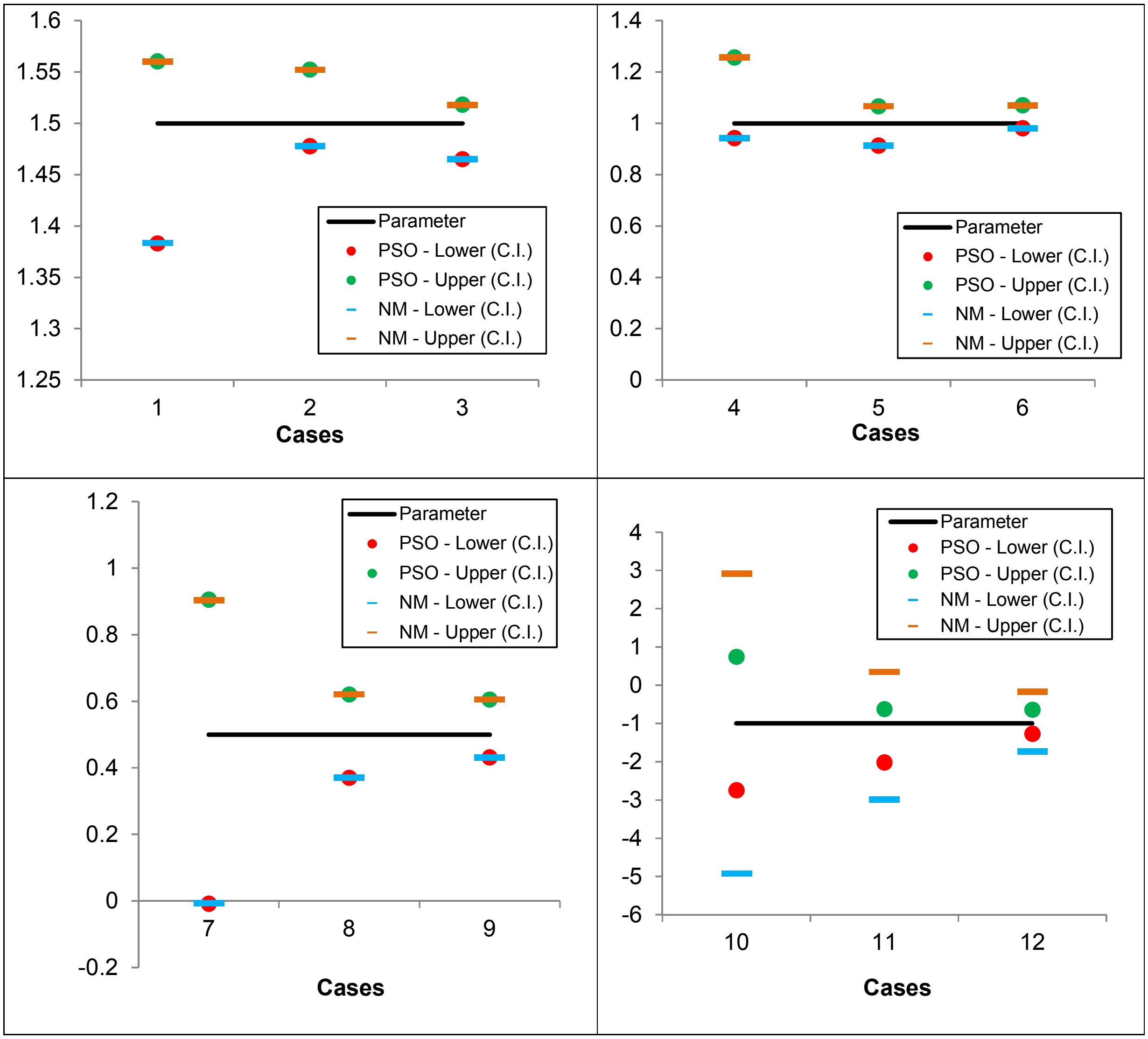


Figure 3.11. Upper and lower limits of the asymptotic Confidence Interval for the parameter q, obtained by PSO and Nelder-Mead.

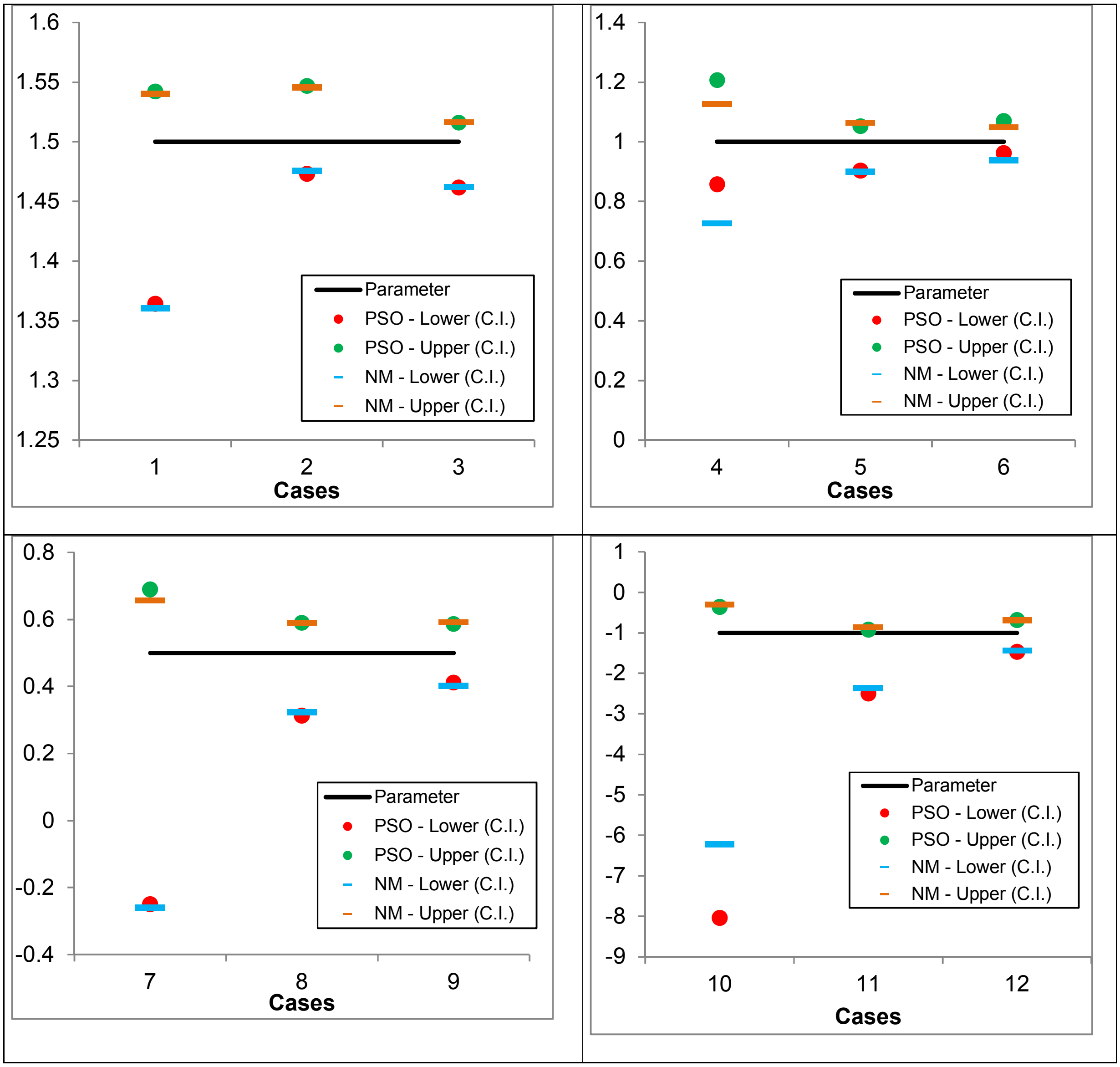


Figure 3.12. Upper and lower limits of the Bootstrap-P Confidence Interval for the parameter q, obtained by PSO and Nelder-Mead.

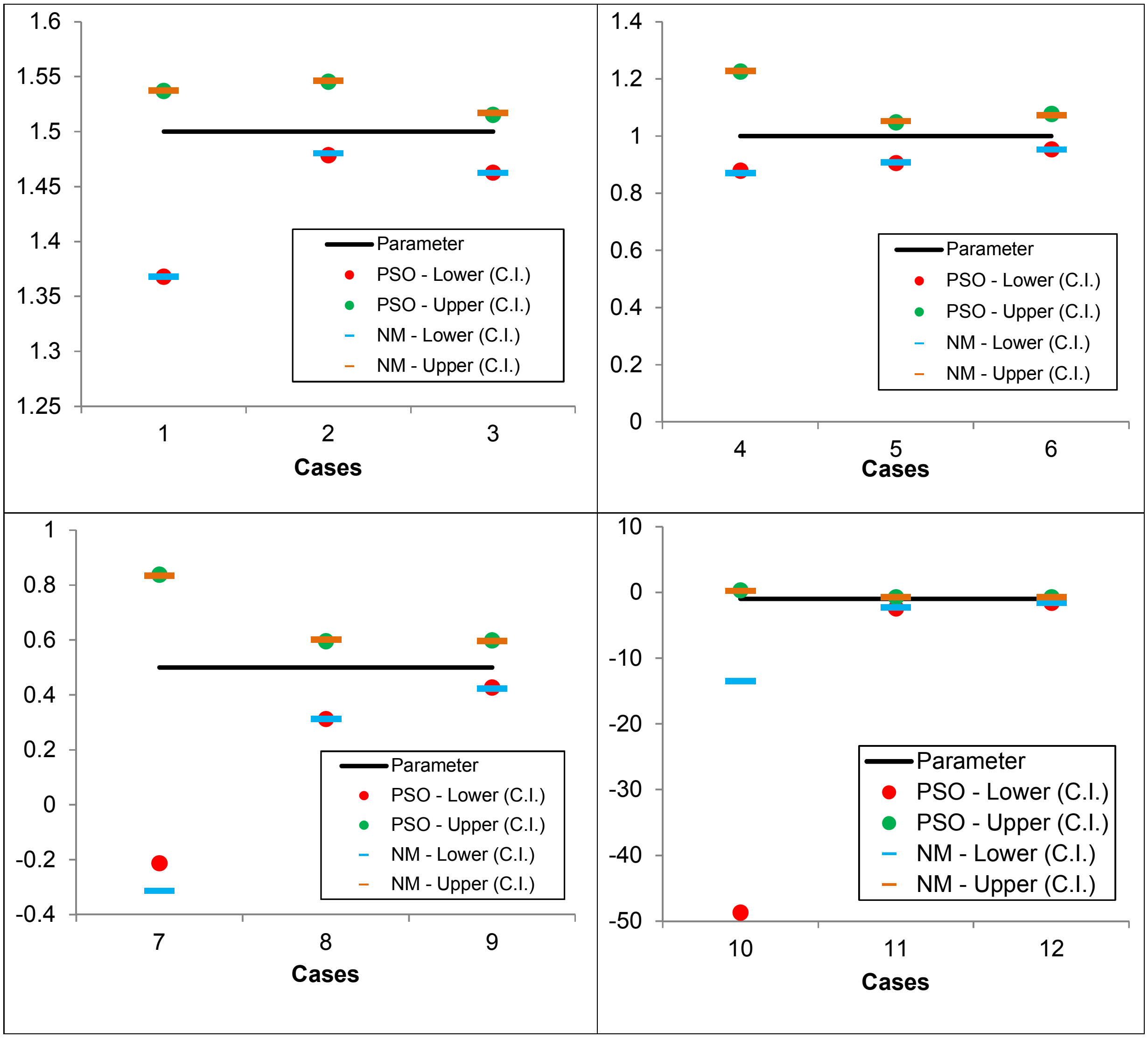


Figure 3.13. Upper and lower limits of the Non-Parametric Bootstrap Confidence Interval for the parameter q, obtained by PSO and Nelder-Mead.

As we can observe in all simulated cases, the width of the intervals decreases as n increases for both PSO and Nelder-Mead algorithms, and for the two parameters in analysis. This is expected due to the consistency property of the MLE.

We also observe that even for the sample of 100, the asymptotic intervals presented widths very similar to the bootstrap-p and non-parametric bootstrap confidence intervals. It is also important to point out that when we compare the interval widths obtained by PSO and Nelder-Mead algorithms we note that mostly the Nelder-Mead algorithm presented smaller intervals for both η and q – For η parameter we can highlight that only 27,77% of intervals obtained by PSO is smaller than the intervals obtained by Nelder-Mead and for the parameter q, 41,66% of the intervals obtained by PSO is smaller than the intervals obtained by Nelder-Mead. In general, it can be observed that for larger sample sizes (n = 1000), asymptotic and bootstrap approaches tend to provide similar interval estimates for the q-Exponential parameters.

3.5 Application Examples

In this section, we provide estimates of the q-Exponential parameters for fatigue data obtained from two kind of materials. The first example, originally described in [83], deals with the experimental determination of high cycle fatigue of ductile cast iron used for wind turbine components and the second one, which was first reported in [84], evaluates the gigacycle fatigue life of high-strength steel. Once the second example presents data obtained from a more resistant material than the material of the first example, it is natural that the cycles until failure for the high-strength steel present order of magnitudes larger than data obtained from ductile cast iron. Thus, it is expected that for the more resistant material, the fit by the q-Exponential distribution presents a better performance than when we consider the Weibull distribution; this is due the fact that the q-Exponential has a heavier tail than the Weibull distribution allowing that data with great magnitude can be well modeled by this distribution.

We will fit the data by a q-exponential distribution using PSO and Nelder-Mead. Besides, we will consider the modeling of data by a Weibull distribution in order to compare the efficiency of these two distributions for this kind of data. The evaluation of the adjustments it will be done using the bootstrapped Kolmogorov-Smirnov test and graphical analysis.

3.5.1 Example 1

From [83], we collect data obtained from a specimen of ductile cast iron with diameter 50 mm (Ø50). The data set given in terms of number of cycles to failure is presented in Table 3.9:

Table 3.9. Ø50 specimen fatigue test data.

Specimen Number	Fatigue Life (number of cycles to failure)
1	295000
2	869000
3	869900
4	1573335
5	151400
6	152000
7	183700
8	218000
9	30200
10	45100
11	46900
12	47300

3.5.1.1 Results for the PSO Approach

Once the PSO is a probabilistic method of optimization, it can present different results for the same data. Thus, in this example, the PSO was replicated 30 times and, as in the numerical experiments, the provided estimates were practically the same with standard deviations of 0.0117, 1.9239E-8 and smaller than 10^{-16} for η, q, in this order. The estimated MLE parameters are $\hat{\eta} = 161820.5715$ and $\hat{q} = 1.3007$. The 90% asymptotic confidence intervals for η and q are [-48898.1867, 372539.3256] and [0.9249, 1.6766], respectively. Note that, due to the very small sample size, the lower bound of the interval concerning η is negative, which is not a possible value for this parameter – This is due the fact that asymptotic confidence interval present best results only when the sample size is larger, and for the situation of this example, the asymptotic approach is not recommended. The parametric and non-parametric bootstrap intervals are reported in Table 3.10. All intervals were constructed considering a confidence level equal to 90%. Unfortunately, due the small sample size , as shown in the Table 3.8, confidence intervals are little informative, with great widths.

Table 3.10. Bootstrap interval estimates provided by PSO for Example 1

		$n = 12$	
		Bootsrap-P	**Non-Parametric Bootstrap**
η	**Confidence Interval**	[52558.29, 75470383.73]	[67226.68, 496841515.04]
	Width	75417825.44	496774288.36
q	**Confidence Interval**	[-125.1890, 1.5124]	[-451.9696, 1.4434]
	Width	126.7014	453.4129

3.5.1.2 Results for the Nelder-Mead Approach

The estimated MLE parameters are $\hat{\eta} = 161904$ and $\hat{q} = 1.3005$. The 90% asymptotic confidence intervals for η and q are [-48303.41, 372111.3] and [0.9257, 1.6753], respectively. Note also that in the Nelder-Mead approach, due to the small sample size, the lower bound of the interval concerning η is negative, which is not a possible value for this parameter. The parametric and non-parametric bootstrap intervals are reported in Table 3.11. All intervals were constructed considering a confidence level equal to 90%.

Table 3.11. Bootstrap interval estimates provided by Nelder-Mead for Example 1

		$n = 12$	
		Bootsrap-P	**Non-Parametric Bootstrap**
η	**Confidence Interval**	[55548.5, 7295081.0]	[71231.32, 73091176.48]
	Width	7239532.5	73019945.16
q	**Confidence Interval**	[-12.9756, 1.5206]	[-72.3958, 1.4444]
	Width	14.4962	73.8402

We observe that the behavior identified when the numerical experiments were conducted (section 3.4.2) was maintained in implementing the example 1, *i.e.*, the length of the intervals obtained by Nelder-Mead was lower than those obtained by the PSO approach. However, also in this optimization method, despite the width of intervals are smaller than the width obtained when we use PSO, the confidence intervals were little informative due to very small sample size.

3.5.1.3 Bootstrapped Kolmogorov-Smirnov test applied in the Example 1

In order to verify if the estimates obtained from the data of the example 1 fit well the data of the example in a q-Exponential model, we make use of the Bootstrapped-

Kolmogorov-Smirnov test (KS-Boot). Table 3.12 presents the estimated parameters obtained by the PSO and Nelder-Mead algorithms for the data of the first example, the Kolmogorov-Smirnov (K-S) distances between the empirical and fitted distribution functions, and the corresponding p-values (K-S Boot). For this test, it was assumed $N = 1000$.

Table 3.12. Comparing point estimates and KS-Boot Test- PSO vs. Nelder-Mead (Example 1)

Parameter Estimates	**PSO**	**K-S (D_0)**	**p-value**
$\hat{\eta}$	161820.5715	0.1554	0.6090
$\hat{q}$	1.3007		
Parameter Estimates	**Nelder-Mead**	**K-S (D_0)**	**p-value**
$\hat{\eta}$	161904	0.1554	0.6224
$\hat{q}$	1.3005		

Note from Table 3.12 that the parameters estimates by PSO and Nelder-Mead are very close, which implies that the p-value of the KS-Boot test was very similar for these two approaches. Also note that for the two approaches, the KS-Boot presents p-values that indicate the q-Exponential distribution is a good fit for the data of the example 1.

3.5.2 Example 2

From [84] , we collect data obtained from a specimen of high-strength steel with diameter 3 mm (Ø3). The data sets given in terms of number of cycles to failure are presented in Table 3.13:

Table 3.13. Ø3 specimen fatigue test data.

Specimen Number	Fatigue Life (number of cycles to failure)
1	1017286
2	2989152
3	4059346
4	4256299
5	8376572
6	9560400
7	13007977
8	25303118
9	33621704
10	55951560

Specimen Number	Fatigue Life (number of cycles to failure)
11	101155984
12	144322192
13	376711232
14	731957760
15	9444513800
16	9912163300
17	9918688300
18	9921105900

3.5.2.1 Results for the PSO Approach

The obtained point estimates are: $\hat{\eta} = 4688695.8075$ and $\hat{q} = 1.7521$. The 30 PSO replications essentially provided the same estimates with standard deviations (0.2236 and 3.3262E-9 for η and q. The asymptotic intervals for η and q are: [-1077757.8706, 10455149.6121] and [1.6579, 1.8463], respectively. As mentioned previously, the asymptotic approach is not recommended when we have small samples. For this reason, also, in this example, the asymptotic confidence interval is little informative, and the lower bound of the interval concerning η is negative, which is not a possible value for this parameter. The bootstrap intervals (parametric and non-parametric) with 90% of confidence level are reported in Table 3.14.

Table 3.14. Bootstrap interval estimates provided by PSO for Example 2

		$n = 18$	
		Bootsrap-P	**Non-Parametric Bootstrap**
η	**Confidence Interval**	[1417311.8673, 21929912.0205]	[2099992.8679, 4184877383010.9316]
	Width	20512600.1532	4184875283018.0640
q	**Confidence Interval**	[1.5987, 1.8188]	[-420.6016, 1.7886]
	Width	0.2201	422.3902

3.5.2.2 Results for the Nelder-Mead Approach

The estimated MLE parameters are $\hat{\eta} = 4704629$ and $\hat{q} = 1.7519$. The 90% asymptotic confidence intervals for η and q are [-1070366, 10479625] and [1.6579, 1.8459], respectively. The bootstrap intervals (parametric and non-parametric) are reported in Table 3.15. All intervals were constructed considering a confidence level equal to 90%.

Table 3.15. Bootstrap interval estimates provided by Nelder-Mead for Example 2

	$n = 18$		
		Bootsrap-P	**Non-Parametric Bootstrap**
η	**Confidence Interval**	[1456802, 20681229]	[2160329, 16826841]
	Width	19224427	14666512
q	**Confidence Interval**	[1.5979, 1.8220]	[1.6509, 1.8014]
	Width	0.2241	0.1505

It is easy to see from Table 3.14 and Table 3.15 that the confidence interval based on Nelder-Mead algorithm presented confidence intervals with widths smaller than the intervals obtained from the PSO method. The results presented in Table 3.15, shows that the confidence interval of the parameter q is more informative with a width size more appropriate than the width obtained in the previous example. Despite this, the width of the bootstrap-p confidence interval for the parameter η still is large, but if we compare with the results obtained in the previous example, we observe that the results improved considerably, this is due the fact that in this example we deal with a sample with sample slightly large than the sample used in previous example.

3.5.2.3 Bootstrapped Kolmogorov-Smirnov test applied in the Example 2

Table 3.16 presents the estimated parameters obtained by the PSO and Nelder-Mead algorithm for the data of the example 2, the Kolmogorov-Smirnov (K-S) distances between the empirical and fitted distribution functions, and the corresponding *p*-values (K-S Boot). For this test, we considered $N = 1000$.

Table 3.16. Comparing point estimates and KS-Boor Test- PSO vs. Nelder-Mead (Example 2)

Parameter Estimates	**PSO**	**K-S (D_0)**	***p*-value**
$\hat{\eta}$	4688695.8075	0.1327	0.4860
$\hat{q}$	1.7521		
Parameter Estimates	**Nelder-Mead**	**K-S (D_0)**	***p*-value**
$\hat{\eta}$	4704629	0.1329	0.4895
$\hat{q}$	1.7519		

Note from Table 3.16 that, also in this example, the parameters estimates obtained by PSO and Nelder-Mead are very close, which implies that the p-value of the KS-Boot test was very similar for these two approaches. Also note that for the two approaches, the KS-Boot presents p-values that indicate the q-Exponential distribution is a good fit for the data of the example 2.

3.5.3 Comparing q-Exponential with Weibull distribution

As previously mentioned, the Weibull distribution originates from the theory of extreme values [85]; then, for particular values of its parameters, this distribution is capable to model data with large values. Thus, we estimate the parameters of a Weibull distribution considering the data presented in examples 1 and 2 in order to compare the fit quality to the data sets by a Weibull and q-Exponential distributions. The parameters estimates of the Weibull distribution are obtained by analytical expressions[86]. The results for the estimated parameters (scale and shape parameters for Weibull distribution), Kolmogorov-Smirnov (K-S) distances between empirical and fitted distribution functions, and the corresponding p-values (K-S Boot performed with $\mathrm{N} = 1000$) obtained from the data sets are shown in Table 3.17, which also include the K-S distance and p-values for the fit of the q-exponential distribution (PSO and Nelder-Mead).

Table 3.17. Comparing Weibull vs. $\boldsymbol{q}$*-Exponential – Examples 1 and 2.*

	Parameters (Weibull Distribution)		K-S Boot (Weibull)		K-S Boot (q-Exponential) PSO		K-S Boot (q-Exponential) Nelder-Mead	
Examples	**Shape Parameter**	**Scale Parameter**	**K-S (D_0)**	**p-value**	**K-S (D_0)**	**p-value**	**K-S (D_0)**	**p-value**
Data from example 1	0.8336	335326.1	0.164	0.5184	0.1554	0.6090	0.1554	0.6224
Data from example 2	0.3366	417229710	0.1648	0.2047	0.1327	0.4860	0.1329	0.4895

From Table 3.17, we observe that for example 1, the fit by the two distributions are good with p-values equal to 0.5184 (for the Weibull distribution), 0.6090 (for the q-exponential distribution – PSO) and 0.6224 (for the q-Exponential distribution – Nelder-Mead). For example 2, although the Weibull fit is significant, it is clear that the q-exponential distribution showed a better fit to the data. Indeed, we observe for this example p-values equal to 0.2047 (for the Weibull distribution), 0.4860 (for the q-exponential distribution – PSO) and 0.4895 (for the q-Exponential distribution – Nelder-Mead). In example 2, clearly the q-exponential distribution showed a better efficiency, since the data in this example are constituted of extremely large values, with magnitude in order of 10^9. Example 1 presents data with magnitude somewhat lower, *i.e.*, in the order of 10^6. Thus, with these examples, we conclude that the q-Exponential can model data with extremely large values with more efficiency than the Weibull distribution.

Figure 3.14 and Figure 3.15 present the empirical and theoretical CDFs (Weibull and q-Exponential) for examples 1 and 2, respectively. Once the estimates obtained from the two methods (PSO and Nelder-Mead) are very similar, we choose, without loss of generality, the estimates obtained from the PSO to construct the Figure 3.14 and Figure 3.15. Note that for the first example (Figure 3.14) both q-Exponential and Weibull fits very well the empirical data. The figure for the second example is plotted in logarithmic scale in order to provide a better visualization of the empirical CDF, as the data set contains extremely large values. As we can note from Figure 3.15, the empirical curve is very close to the q-Exponential CDF, confirming that the q-exponential distribution is more efficient than Weibull distribution when we deal with the kind of data presented in Example 2, *i.e.*, data that encompass extremely large values such as in case of material with high resistance to failure.

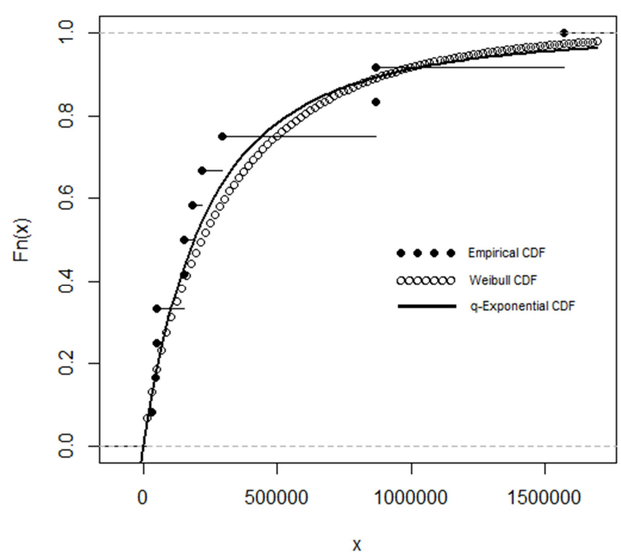


Figure 3.14. Empirical and Theoretical (q-Exponential and Weibull) CDFs – (Example 1).

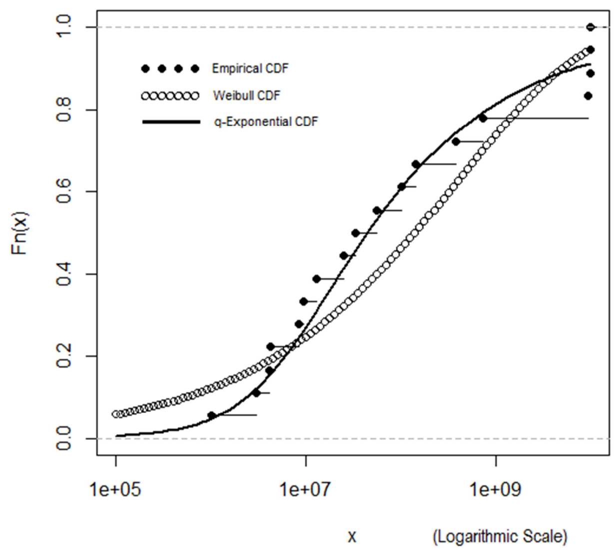


*Figure 3.15. Empirical and Theoretical (**q**-Exponential and Weibull) CDFs – (Example 2).*

4 A NEW STRESS-STRENGTH MODEL BASED ON q-EXPONENTIAL DISTRIBUTION.

4.1 Maximum likelihood estimators of index *R=P(Y<X)*

In this section, we estimate the index $R=P(Y<X)$ by using the maximum likelihood method. We assume that X and Y are independent random variables and follow q-Exponential distributions with different parameters. We can write $Y \sim qExp\ (q,\ \eta)$ and $X \sim qExp(r,\ \beta)$, where q and r are the entropic indices (shape parameters) of stress and strength, respectively, and η and β are the scale parameters of stress and strength. As mentioned above, the support of the q-Exponential can be limited ($q<1$) or unlimited ($1<q<2$) (see Equation (2.1)). In order to calculate the index R, we use the Wolfram Mathematica computer software [87]. Due the difference of the support of the q-Exponential, we will consider two cases in this work:

Case 1: There is no limitation on the support of X (strength), *i.e.*, $1<r<2$:

$$R = P(Y < X) = \int_0^{\infty} \int_0^{x} \frac{\left\{(2-q)\left[\frac{(q-1)y}{\eta}+1\right]^{\frac{1}{1-q}}\right\}\left\{(2-r)\left[\frac{(r-1)x}{\beta}+1\right]^{\frac{1}{1-r}}\right\}}{\eta\beta}\, dy\, dx =$$

$$= \frac{(r-2)\left\{\frac{\beta\left[(r-1)^3 ABC - \frac{DEF\left(\frac{r-1}{\beta}\right)^{\frac{r-2}{r-1}}\left(\frac{q-1}{\eta}\right)^{\frac{2-r}{r-1}}\left[1-\frac{\beta(q-1)}{\eta(r-1)}\right]^{\frac{1}{1-q}+\frac{1}{1-r}}[\eta(1-r)+\beta(q-1)]^2}{\eta^2}\right]}{A(r-1)^4} - \frac{\beta B}{r-1}\right\}}{\beta D} \tag{4.1}$$

Where:

$A = \Gamma\left(\frac{2-q}{q-1}\right)$;

$B = \Gamma\left(\frac{2-r}{r-1}\right)$;

$C = {}_2F_1\left(1, \frac{2-q}{q-1}; \frac{3-2r}{1-r}; \frac{(q-1)\beta}{(r-1)\eta}\right)$;

$D = \Gamma\left(\frac{1}{r-1}\right)$;

$E = \Gamma\left(\frac{1}{1-r}\right)$;

$F = \Gamma\left(\frac{1}{r-1} - 2 + \frac{1}{q-1}\right)$

And

$${}_2F_1(a,b;c;z) = \frac{\Gamma(c)}{\Gamma(a)\,\Gamma(c-a)} \int_0^1 t^{a-1}(1-t)^{c-a-1}(1-zt)^{-b}\,dt$$

Note that ${}_2F_1(a,b;c;z)$ is the Gauss Hypergeometric Function [88].

Case 2: The support of X (strength) is limited, *i.e., r<1*:

$$R = P(Y < X)$$

$$= \int_0^{\frac{1}{\left(\frac{1}{\beta}\right)(1-r)}} \int_0^{x} \frac{\left\{(2-q)\left[\frac{(q-1)y}{\eta}+1\right]^{\frac{1}{1-q}}\right\}\left\{(2-r)\left[\frac{(r-1)x}{\beta}+1\right]^{\frac{1}{1-r}}\right\}}{\eta\beta}\,dy\,dx =$$

$$= 1 - {}_2F_1\left(1, \frac{1}{q-1} - 1; 2 + \frac{1}{1-r}; \frac{(q-1)\beta}{(r-1)\eta}\right) \tag{4.2}$$

To compute the MLE of R, let $Y=\{y_1, y_2, \ldots, y_n\}$ be a random sample of size n, and $X=\{x_1, x_2, \ldots, x_m\}$ be another random sample of size m. Since X and Y are independent variables, it is possible to write the likelihood function for the observed samples as:

$$L(x,y,r,\beta,q,\eta) = \left\{(2-q)^n\left(\frac{1}{\eta}\right)^n \prod_{i=1}^{n}\left[1-\frac{(1-q)y_i}{\eta}\right]^{\frac{1}{1-q}}\right\}\left\{(2-r)^m\left(\frac{1}{\beta}\right)^m \prod_{i=1}^{m}\left[1-\frac{(1-r)x_i}{\beta}\right]^{\frac{1}{1-r}}\right\}$$

$$= (2-q)^n\left(\frac{1}{\eta}\right)^n (2-r)^m\left(\frac{1}{\beta}\right)^m \prod_{i=1}^{n}\left[1-\frac{(1-q)y_i}{\eta}\right]^{\frac{1}{1-q}} \prod_{i=1}^{m}\left[1-\frac{(1-r)x_i}{\beta}\right]^{\frac{1}{1-r}}.$$

Therefore, the log-likelihood function is written as follows:

$$l(x, y, r, \beta, q, \eta)$$
$$= n\, ln(2-q) + n \ln\left(\frac{1}{\eta}\right) + m\, ln(2-r) + m\, ln\left(\frac{1}{\beta}\right) + \frac{1}{1-q}\sum_{i=1}^{n} ln\left[1 - \frac{(1-q)y_i}{\eta}\right] + \frac{1}{1-r}\sum_{i=1}^{m} ln\left[1 - \frac{(1-r)x_i}{\beta}\right] \quad (4.3)$$

Maximizing the log-likelihood function given in Equation (4.3), results in a convoluted system of equations and, thus, the derivation of analytical expressions for the MLE becomes impractical. So, in this work, the maximization of the log-likelihood function in Equation (4.3) will be performed by the Nelder-Mead Method [47] available in the Software R (optim function) [82] and by the PSO optimization method implemented in the MATLAB software [81].

Since $\hat{q}$, $\hat{\eta}$, $\hat{r}$ and $\hat{\beta}$ are solutions that maximize the log-likelihood function of Equation (4.3), and using the property of invariance of the MLE, from Equations (4.1) and (4.2) we can obtain the MLE of *R* for the two previously mentioned cases:

Case 1: When the *X* (strength) has *1<r<2*:

$$\hat{R} = \frac{(\hat{r}-2)\left\{\frac{\hat{\beta}\left[(\hat{r}-1)^3\hat{A}\hat{B}\hat{C} - \frac{\hat{D}\hat{E}\hat{F}\left(\frac{\hat{r}-1}{\hat{\beta}}\right)^{\frac{\hat{r}-2}{\hat{r}-1}}\left(\frac{\hat{q}-1}{\hat{\eta}}\right)^{\frac{2-\hat{r}}{\hat{r}-1}}\left[1-\frac{\hat{\beta}(\hat{q}-1)}{\hat{\eta}(\hat{r}-1)}\right]^{\frac{1}{1-\hat{q}}+\frac{1}{1-\hat{r}}}\left[\hat{\eta}(1-\hat{r})+\hat{\beta}(\hat{q}-1)\right]^2}{\eta^2}\right]}{\hat{A}(\hat{r}-1)^4} - \frac{\hat{\beta}\hat{B}}{\hat{r}-1}\right\}}{\hat{\beta}\hat{D}} \quad (4.4)$$

Where $\hat{A} = \Gamma\left(\frac{2-\hat{q}}{\hat{q}-1}\right)$; $\hat{B} = \Gamma\left(\frac{2-\hat{r}}{\hat{r}-1}\right)$; $\hat{C} = {}_2F_1\left(1, \frac{2-\hat{q}}{\hat{q}-1}; \frac{3-2\hat{r}}{1-\hat{r}}; \frac{(\hat{q}-1)\hat{\beta}}{(\hat{r}-1)\hat{\eta}}\right)$; $\hat{D} = \Gamma\left(\frac{1}{\hat{r}-1}\right)$; $\hat{E} = \Gamma\left(\frac{1}{1-\hat{r}}\right)$; $\hat{F} = \Gamma\left(\frac{1}{\hat{r}-1} - 2 + \frac{1}{\hat{q}-1}\right)$.

Case 2: When the *X* (strength) has *r<1*:

$$\hat{R} = 1 - {}_2F_1\left(1, \frac{1}{\hat{q}-1} - 1; 2 + \frac{1}{1-\hat{r}}; \frac{(\hat{q}-1)\hat{\beta}}{(\hat{r}-1)\hat{\eta}}\right) \tag{4.5}$$

4.2 Bootstrap Confidence Intervals

In this section, we present the construction of confidence intervals for the index R by using bootstrap-p and non-parametric bootstrap methods [53], [89], [90].

4.2.1 Bootstrap-p

The algorithm for constructing confidence intervals by using the bootstrap-p approach has the following steps:

- **Step 1:** From an initial sample for the variable $X = \{x_1, x_2, \ldots, x_m\}$ and another one for $Y = \{y_1, y_2, \ldots, y_n\}$, estimate the parameters (q, η, r, β) by maximizing Equation (4.3);
- **Step 2:** Use the estimates obtained in the previous step and Equation (2.2) to generate new samples for X and Y, *i.e.*: $\{x_1^*,\ x_2^*, \ldots,\ x_m^*\}$ and $\{y_1^*,\ y_2^*, \ldots,\ y_n^*\}$. Based on these new samples, compute the bootstrap sample estimate of R, say R^*, using Equation (4.4) or (4.5) (depending on the r value);
- **Step 3:** Repeat step 2, N times;
- **Step 4:** Using the N values of R^* obtained in step 3 and by adopting a γ significance level, find the percentiles $R^*_{\gamma/2}$ and $R^*_{1-(\gamma/2)}$. Thus, it is possible to determine an approximate confidence interval, with confidence interval equal to *100*(1-γ)%*, for the index R, as:

$$C.I. = [R^*_{\gamma/2}, R^*_{(1-\gamma)/2}]. \tag{4.6}$$

4.2.2 Non-Parametric bootstrap

The algorithm for constructing confidence intervals by using the non-parametric bootstrap approach is as follows:

- **Step 1:** From an initial sample for the variable $X = \{x_1, x_2, \ldots, x_m\}$ and another one for $Y = \{y_1, y_2, \ldots, y_n\}$, generate new samples for *X* and *Y* by sampling with replacement, *i.e.*, $\{x_1^*,\ x_2^*, \ldots,\ x_m^*\}$ and $\{y_1^*,\ y_2^*, \ldots,\ y_n^*\}$. Based on these new samples, compute the estimate of *R*, say R^*, using Equation (4.4) or (4.5) (depending on the *r* value);

- **Step 2:** Repeat step 1, *N* times.

- **Step 3:** By using the *N* values of R^* from step 2 and by adopting a γ significance level, the percentiles $R^*_{\gamma/2}$ and $R^*_{1-(\gamma/2)}$ are obtained; they determine an approximate confidence interval for the index *R* with confidence level equals to $100*(1-\gamma)\%$ using Equation (4.6).

4.3 Numerical Experiments

This section presents the performance evaluation of the MLE and bootstrap confidence intervals by means of simulation experiments. Here we will use the Nelder-Mead method and the PSO algorithm in order to obtain the estimates of MLE. We consider different sample sizes and different parameter values. First, we analyze the MLE and then we discuss the bootstrap confidence interval. Note that the simulations involved entropic indices ranging between 1 and 2 (cases 1 - 11), from 0 to 1 (cases 12 - 22), and with negative values (cases 23 - 33).

4.3.1 Analysis of the MLE

Several combinations of sample sizes for the stress and strength are considered: $(n; m)$ = (100; 100), (250; 250), (500; 500), (1000; 1000), (5000; 5000), (100; 250), (100; 500), (100; 1000), (250; 100), (500; 100) and (1000; 100). Besides, we choose three sets of parameters values respecting three important situations for the entropic indices, *i.e.*: $1<r,q<2$; $0<r,q<1$ and $r,q<0$. Thus, we have: (q, η, r, β) = (1.78; 0.15; 1.9; 0.1), (0.55; 22; 0.67; 30.5), (-1.95; 0.1; -1.8; 0.18). Observe that 11 different combinations of sample sizes multiplied by 3 different parameter sets are equal to 33 initial samples. The samples for the simulations are generated by Equation (2.2). All results are based on 1000 replications, *i.e.*, we generate 1000 samples from each set of initial parameters for all the combinations of *n* and *m*. Thus, a total of 33000 samples are generated. For each sample, we compute the MLE for q, η, r, β by maximizing Equation (4.3) via the Nelder-Mead method and PSO algorithm.

Thus, we obtain the MLE of index R by Equation (4.4) or (4.5) (depending on the r value). This process is carried out for each of the 1000 replications. Subsequently, we obtain the average of the estimation results for parameters q, η, r, β, and also for the index R. Table 4.4 presents the results for all simulation runs as well as the index R estimations. The mean squared error (MSE) and the average biases are calculated for $\hat{R}$ over the 1000 replications. Note that these quality indices are obtained for an estimator $\hat{\theta}$ of θ as $bias(\hat{\theta}) = E(\hat{\theta}) - \theta$ and $MSE = Var(\hat{\theta}) + bias(\hat{\theta})^2$.

From the simulation results (Table 4.4 and Table 4.5), the following findings are observed:

(i) When (*n; m)* increase, the MSEs decrease. This suggests the consistency property of the MLE (see Table 4.1);
(ii) For a fixed *n*, MSEs decrease as *m* increases (see Table 4.2);
(iii) For a fixed *m*, MSEs decrease as *n* increases (see Table 4.3);

In order to illustrate behaviors (i), (ii) and (iii), excerpts from Table 4.4 and Table 4.5 are reproduced in Table 4.1, Table 4.2 and Table 4.3 respectively.

Table 4.1. Examples of cases that present a decrease of the MSE when (n; m) increase (Nelder-Mead and PSO)

Nelder Mead							
Case	***n***	***m***	***q***	***η***	***r***	***β***	***MSE***
1	100	100	1.78	0.15	1.90	0.10	0.00120
12	100	100	0.55	22	0.67	30.5	0.00155
23	100	100	-1.95	0.1	-1.8	0.18	0.03185
5	5000	5000	1.78	0.15	1.90	0.10	0.00002
16	5000	5000	0.55	22	0.67	30.5	0.00002
27	5000	5000	-1.95	0.1	-1.8	0.18	0.01981
PSO							
Case	***n***	***m***	***q***	***η***	***r***	***β***	***MSE***
1	100	100	1.78	0.15	1.90	0.10	0.00120
12	100	100	0.55	22	0.67	30.5	0.00155
23	100	100	-1.95	0.1	-1.8	0.18	0.00423
5	5000	5000	1.78	0.15	1.90	0.10	0.00002

16	5000	5000	0.55	22	0.67	30.5	0.00002
27	5000	5000	-1.95	0.1	-1.8	0.18	0.00004

Table 4.2. Examples of cases that present a decrease of the MSE for a fixed n and an increase of m (Nelder-Mead and PSO).

Nelder Mead							
Case	***n***	***m***	***q***	***η***	***r***	***β***	***MSE***
1	100	100	1.78	0.15	1.90	0.10	0.00120
6	100	250	1.78	0.15	1.90	0.10	0.00069
7	100	500	1.78	0.15	1.90	0.10	0.00057
8	100	1000	1.78	0.15	1.90	0.10	0.00051
PSO							
Case	***n***	***m***	***q***	***η***	***r***	***β***	***MSE***
1	100	100	1.78	0.15	1.90	0.10	0.00120
6	100	250	1.78	0.15	1.90	0.10	0.00069
7	100	500	1.78	0.15	1.90	0.10	0.00057
8	100	1000	1.78	0.15	1.90	0.10	0.00051

Table 4.3. Examples of cases that present a decrease of the MSE for a fixed m and an increase of n (Nelder-Mead and PSO).

Nelder Mead							
Case	***n***	***m***	***q***	***η***	***r***	***β***	***MSE***
12	100	100	0.55	22	0.67	30.5	0.00155
20	250	100	0.55	22	0.67	30.5	0.00113
21	500	100	0.55	22	0.67	30.5	0.00093
22	1000	100	0.55	22	0.67	30.5	0.00088
PSO							
Case	***n***	***m***	***q***	***η***	***r***	***β***	***MSE***
12	100	100	0.55	22	0.67	30.5	0.00155
20	250	100	0.55	22	0.67	30.5	0.00112
21	500	100	0.55	22	0.67	30.5	0.00095
22	1000	100	0.55	22	0.67	30.5	0.00088

Next, we present graphics that allow us a better view of the results obtained in the simulations for the estimation of the Index R using PSO and Nelder-Mead. The data used in the creation of these data can be found in Table 4.4 and Table 4.5. The graphics presented in Figure 4.1 show a comparison between the point estimates obtained by PSO and Nelder-Mead and the real value of the Index *R*. Figure 4.2 presents graphics that show the MSE results obtained by PSO and Nelder-Mead when we make the estimation of the Index *R*.

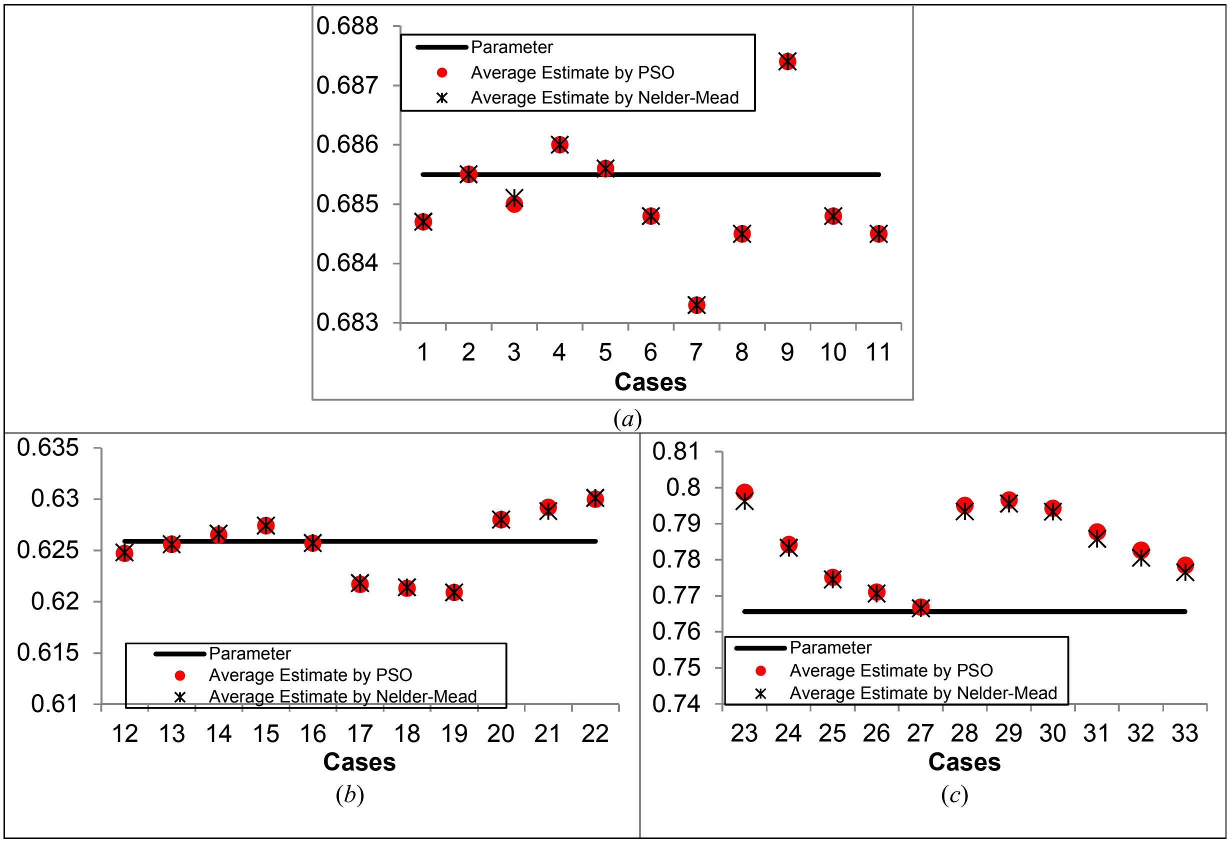
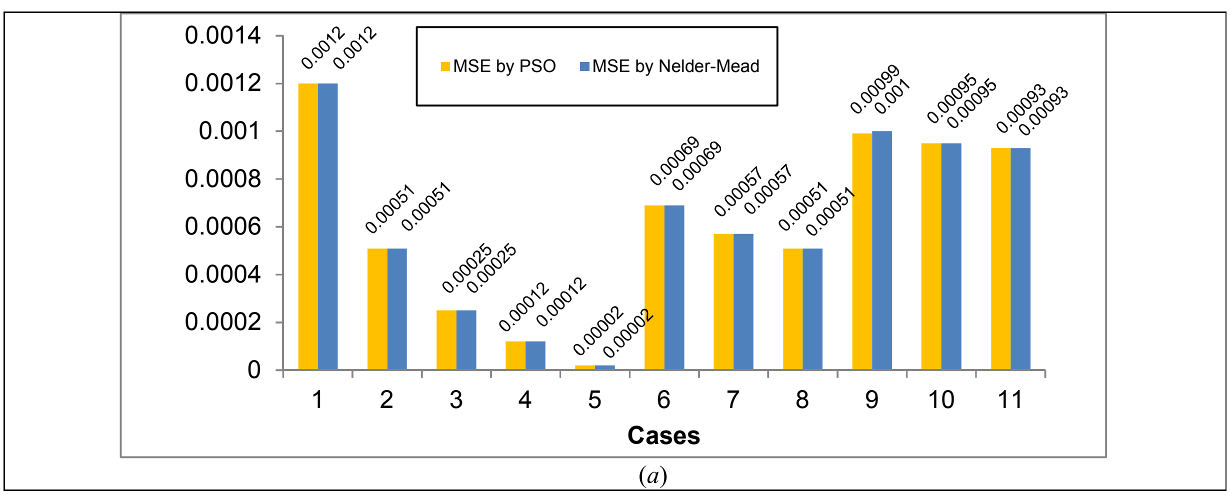


Figure 4.1. Graphical comparison between results of the estimates of the Index ***R*** *obtained by PSO and Nelder-Mead, and the true value of the Index - (a) Cases 1 to 11, (b) Cases 12 to 22 and (c) Cases 23 to 33.*



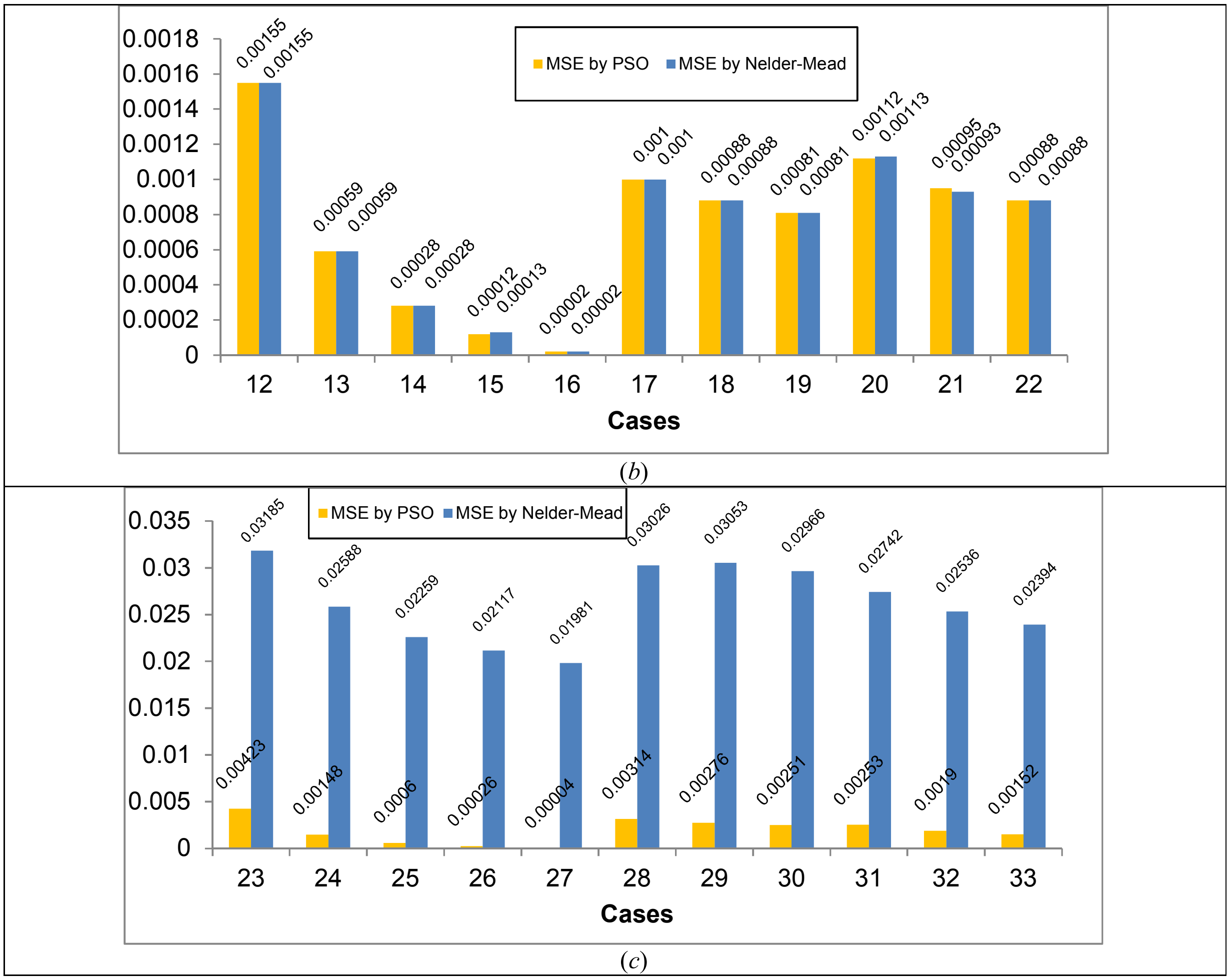


*Figure 4.2. Mean Squared Error (MSE) for the estimates of the Index **R** obtained by PSO and Nelder-Mead - (a) Cases 1 to 11, (b) Cases 12 to 22 and (c) Cases 23 to 33.*

From the graphics we can observe that the bias for cases 1 to 22 not indicates a tendency of overestimation or underestimation, once we have bias with positive and negative values. For the cases 23 to 33, there is a tendency of overestimation of the index R, once the bias for these cases present always positive values. Besides, cases 23 to 33 present slightly higher results for the MSE if compared with other cases. This fact is observed mainly when we deal with the Nelder-Mead approach that present higher results than the obtained by PSO.

Table 4.4. Simulation results and estimation for the index R (Nelder-Mead).

	Sample Size		**Parameters**					**MLE Results (Average from 1000 samples)**					$\hat{R}$	
Case	***n***	***m***	***q***	η	***r***	β	***R***	$\hat{\hat{q}}$	$\hat{\hat{\eta}}$	$\hat{\hat{r}}$	$\hat{\hat{\beta}}$	$\hat{\hat{R}}$	***bias***	***MSE***
1	100	100	1.78	0.15	1.90	0.10	0.6855	1.7767	0.1647	1.8987	0.1152	0.6847	-0.00076	0.00120
2	250	250	1.78	0.15	1.90	0.10	0.6855	1.7784	0.1562	1.8995	0.1062	0.6855	0.00000	0.00051
3	500	500	1.78	0.15	1.90	0.10	0.6855	1.7792	0.1537	1.8999	0.1023	0.6851	-0.00044	0.00025
4	1000	1000	1.78	0.15	1.90	0.10	0.6855	1.7797	0.1511	1.8999	0.1021	0.6860	0.00054	0.00012
5	5000	5000	1.78	0.15	1.90	0.10	0.6855	1.7799	0.1502	1.9000	0.1005	0.6856	0.00013	0.00002
6	100	250	1.78	0.15	1.90	0.10	0.6855	1.7765	0.1648	1.8993	0.1051	0.6848	-0.00074	0.00069
7	100	500	1.78	0.15	1.90	0.10	0.6855	1.7782	0.1642	1.8997	0.1026	0.6833	-0.00222	0.00057
8	100	1000	1.78	0.15	1.90	0.10	0.6855	1.7775	0.1631	1.8998	0.101	0.6845	-0.00104	0.00051
9	250	100	1.78	0.15	1.90	0.10	0.6855	1.7787	0.1555	1.8988	0.1192	0.6874	0.00188	0.00100
10	500	100	1.78	0.15	1.90	0.10	0.6855	1.7789	0.1536	1.8984	0.1158	0.6848	-0.00066	0.00095
11	1000	100	1.78	0.15	1.90	0.10	0.6855	1.7798	0.1510	1.8979	0.1195	0.6845	-0.00097	0.00093
12	100	100	0.55	22	0.67	30.5	0.6259	0.4293	25.1014	0.5754	34.3545	0.6248	-0.00111	0.00155
13	250	250	0.55	22	0.67	30.5	0.6259	0.5047	23.1850	0.6396	31.7394	0.6256	-0.00026	0.00059
14	500	500	0.55	22	0.67	30.5	0.6259	0.5270	22.5807	0.6507	31.2574	0.6266	0.00065	0.00028
15	1000	1000	0.55	22	0.67	30.5	0.6259	0.5393	22.2590	0.6569	31.0805	0.6274	0.00153	0.00013
16	5000	5000	0.55	22	0.67	30.5	0.6259	0.5462	22.1077	0.6682	30.5519	0.6257	-0.00022	0.00002
17	100	250	0.55	22	0.67	30.5	0.6259	0.4211	25.3501	0.6392	31.7864	0.6218	-0.00413	0.00100
18	100	500	0.55	22	0.67	30.5	0.6259	0.4361	25.0652	0.6515	31.2202	0.6214	-0.00449	0.00088
19	100	1000	0.55	22	0.67	30.5	0.6259	0.4293	25.1689	0.6611	30.8397	0.6209	-0.00496	0.00081
20	250	100	0.55	22	0.67	30.5	0.6259	0.5006	23.3212	0.5706	34.5754	0.6280	0.00206	0.00113
21	500	100	0.55	22	0.67	30.5	0.6259	0.5289	22.5205	0.5777	34.1914	0.6288	0.00292	0.00093
22	1000	100	0.55	22	0.67	30.5	0.6259	0.5384	22.3168	0.5601	34.9577	0.6301	0.00417	0.00088
23	100	100	-1.95	0.1	-1.8	0.18	0.7656	-4.2601	0.1732	-3.9959	0.3117	0.7962	0.17029	0.03185
24	250	250	-1.95	0.1	-1.8	0.18	0.7656	-2.7699	0.1262	-2.4857	0.2212	0.7833	0.15742	0.02588
25	500	500	-1.95	0.1	-1.8	0.18	0.7656	-2.3128	0.1116	-2.0635	0.1958	0.7745	0.14861	0.02259
26	1000	1000	-1.95	0.1	-1.8	0.18	0.7656	-2.1147	0.1052	-1.9444	0.1886	0.7706	0.14468	0.02117
27	5000	5000	-1.95	0.1	-1.8	0.18	0.7656	-1.9825	0.1010	-1.8283	0.1816	0.7665	0.14061	0.01981
28	100	250	-1.95	0.1	-1.8	0.18	0.7656	-4.5069	0.1814	-2.4393	0.2184	0.7934	0.16750	0.03026
29	100	500	-1.95	0.1	-1.8	0.18	0.7656	-4.6143	0.1849	-2.1051	0.1984	0.7955	0.16958	0.03053
30	100	1000	-1.95	0.1	-1.8	0.18	0.7656	-4.4991	0.1812	-1.9677	0.1901	0.7933	0.16738	0.02966
31	250	100	-1.95	0.1	-1.8	0.18	0.7656	-2.6900	0.1236	-4.1665	0.3223	0.7858	0.15985	0.02742
32	500	100	-1.95	0.1	-1.8	0.18	0.7656	-2.3163	0.1117	-4.2061	0.3245	0.7805	0.15456	0.02536
33	1000	100	-1.95	0.1	-1.8	0.18	0.7656	-2.1101	0.1051	-4.1475	0.3215	0.7765	0.15063	0.02394

Table 4.5 - Simulation results and estimation for the index R (PSO).

	Sample Size		Parameters					MLE Results (Average from 1000 samples)					$\hat{R}$	
Case	***n***	***m***	***q***	η	***r***	β	***R***	$\bar{\bar{q}}$	$\bar{\bar{\eta}}$	$\bar{\bar{r}}$	$\bar{\bar{\beta}}$	$\bar{\bar{R}}$	***bias***	***MSE***
1	100	100	1.78	0.15	1.90	0.10	0.6855	1.7767	0.1647	1.8987	0.1151	0.6847	-0.00075	0.00120
2	250	250	1.78	0.15	1.90	0.10	0.6855	1.7784	0.1562	1.8995	0.1062	0.6855	1.94E-06	0.00051
3	500	500	1.78	0.15	1.90	0.10	0.6855	1.7792	0.1536	1.8998	0.1022	0.6850	-0.00044	0.00025
4	1000	1000	1.78	0.15	1.90	0.10	0.6855	1.7796	0.1511	1.8999	0.1020	0.6860	0.00053	0.00012
5	5000	5000	1.78	0.15	1.90	0.10	0.6855	1.7799	0.1502	1.8999	0.1005	0.6856	0.00013	0.00002
6	100	250	1.78	0.15	1.90	0.10	0.6855	1.7765	0.1648	1.8993	0.1051	0.6848	-0.00074	0.00069
7	100	500	1.78	0.15	1.90	0.10	0.6855	1.7782	0.1642	1.8997	0.1026	0.6833	-0.00222	0.00057
8	100	1000	1.78	0.15	1.90	0.10	0.6855	1.7775	0.1631	1.8998	0.1010	0.6845	-0.00103	0.00051
9	250	100	1.78	0.15	1.90	0.10	0.6855	1.7787	0.1555	1.8988	0.1192	0.6874	0.00187	0.00099
10	500	100	1.78	0.15	1.90	0.10	0.6855	1.7789	0.1536	1.8984	0.1158	0.6848	-0.00067	0.00095
11	1000	100	1.78	0.15	1.90	0.10	0.6855	1.7798	0.1510	1.8979	0.1194	0.6845	-0.00095	0.00093
12	100	100	0.55	22	0.67	30.5	0.6259	0.4288	25.1177	0.5732	34.4303	0.6247	-0.00115	0.00155
13	250	250	0.55	22	0.67	30.5	0.6259	0.5047	23.1850	0.6385	31.7760	0.6256	-0.00029	0.00059
14	500	500	0.55	22	0.67	30.5	0.6259	0.5270	22.5810	0.6495	31.2973	0.6265	0.00063	0.00028
15	1000	1000	0.55	22	0.67	30.5	0.6259	0.5393	22.2585	0.6569	31.0812	0.6274	0.00153	0.00012
16	5000	5000	0.55	22	0.67	30.5	0.6259	0.5462	22.1079	0.6682	30.5518	0.6257	-0.00022	0.00002
17	100	250	0.55	22	0.67	30.5	0.6259	0.4210	25.3530	0.6376	31.8416	0.6217	-0.00414	0.00100
18	100	500	0.55	22	0.67	30.5	0.6259	0.4361	25.0655	0.6509	31.2410	0.6213	-0.00454	0.00088
19	100	1000	0.55	22	0.67	30.5	0.6259	0.4292	25.1694	0.6605	30.8604	0.6209	-0.00495	0.00081
20	250	100	0.55	22	0.67	30.5	0.6259	0.5002	23.3344	0.5692	34.6244	0.6280	0.00210	0.00112
21	500	100	0.55	22	0.67	30.5	0.6259	0.5289	22.5202	0.5773	34.2074	0.6292	0.00328	0.00095
22	1000	100	0.55	22	0.67	30.5	0.6259	0.5384	22.3171	0.5597	34.9718	0.6300	0.00414	0.00088
23	100	100	-1.95	0.1	-1.8	0.18	0.7656	-6.1548	0.2348	-5.3337	0.3938	0.7987	0.03315	0.00423
24	250	250	-1.95	0.1	-1.8	0.18	0.7656	-2.8707	0.1295	-2.5746	0.2267	0.7842	0.01859	0.00148
25	500	500	-1.95	0.1	-1.8	0.18	0.7656	-2.3326	0.1122	-2.0722	0.1963	0.7751	0.00952	0.00060
26	1000	1000	-1.95	0.1	-1.8	0.18	0.7656	-2.1228	0.1055	-1.9548	0.1893	0.7710	0.00543	0.00026
27	5000	5000	-1.95	0.1	-1.8	0.18	0.7656	-1.9889	0.1013	-1.8316	0.1818	0.7668	0.00116	0.00004
28	100	250	-1.95	0.1	-1.8	0.18	0.7656	-6.5750	0.2486	-2.4849	0.2213	0.7950	0.02938	0.00314
29	100	500	-1.95	0.1	-1.8	0.18	0.7656	-6.4323	0.2439	-2.1153	0.1990	0.7965	0.03091	0.00276
30	100	1000	-1.95	0.1	-1.8	0.18	0.7656	-6.4737	0.2455	-1.9738	0.1905	0.7943	0.02870	0.00251
31	250	100	-1.95	0.1	-1.8	0.18	0.7656	-2.7547	0.1257	-5.7700	0.4207	0.7877	0.02219	0.00253
32	500	100	-1.95	0.1	-1.8	0.18	0.7656	-2.3306	0.1122	-5.9818	0.4337	0.7826	0.01699	0.00190
33	1000	100	-1.95	0.1	-1.8	0.18	0.7656	-2.1274	0.1057	-5.8382	0.4256	0.7784	0.01281	0.00152

4.3.2 Bootstrap Confidence Interval

For the simulations of the confidence intervals based on bootstrap-*p* and on non-parametric bootstrap, we generated 33 initial samples using the same combinations of sample sizes and initial parameters presented in the sub-section "Analysis of the MLE". For the case of the bootstrap-p, we obtain, from the 33 initial samples, the MLE of the parameters by maximizing the log-likelihood function (Equation (4.3)). Given that we have the parameters' estimates (obtained from the initial samples) for each different combination of parameters and sample size, we can use these estimates to generate $N=1000$ new samples from Equation (2.2). For the case of the non-parametric bootstrap, we use the 33 initial samples to generate $N=1000$ samples by sampling with replacement (for each different combination of parameters and sample sizes).

From the samples generated by the bootstrap-*p* or by non-parametric bootstrap, we estimate the index R from Equation (4.4) or (4.5) (depending on the r value), *i.e.*, for each method we generate $N=1000$ bootstrap estimates of R. We present the mean of $N = 1000$ bootstrap estimates of R and, based on the percentile method, the corresponding 90% confidence interval is also provided.

Table 4.9 and Table 4.10 present, respectively, the results and estimation of bootstrap-*p* and non-parametric bootstrap confidence intervals for index R estimated by Nelder-Mead. Table 4.11 and Table 4.12 present, respectively, the results and estimation of bootstrap-*p* and non-parametric bootstrap confidence intervals for index R estimated by PSO. From the simulations for the bootstrap-*p* and non-parametric bootstrap confidence intervals, in the most cases, we observe that:

(i) When $(n; m)$ increase, the amplitude of the interval (width) decreases. In order to illustrate this behavior, Table 4.6 presents excerpts of Table 4.9 and Table *4.11* with the interval widths for the index R obtained by bootstrap-*p*, considering a 90% confidence level, for the three different combinations of parameters when (n, m) = (100, 100) and (n, m) = (5000, 5000);

(ii) For a fixed n, the widths decrease as m increases. For example, the results in

(iii) Table *4.7* are taken from Table 4.10 and Table 4.12 demonstrate this behavior; For a fixed m, the interval widths decrease as n increases. Table 4.8, which is formed by excerpts of Table 4.9 and Table 4.11 exemplifies the decrease of interval widths for bootstrap-p and 90% of confidence level.

Table 4.6. Examples of cases that present a decrease of interval widths when (n; m) increase

Nelder-Mead								
Case	***n***	***m***	***q***	***η***	***r***	***β***	***(1-γ) = 0.90***	**Width**
1	100	100	1.78	0.15	1.90	0.10	[0.5949; 0.7139]	0.1190
12	100	100	0.55	22	0.67	30.5	[0.5462; 0.6803]	0.1341
23	100	100	-1.95	0.1	-1.8	0.18	[0.7273; 0.9553]	0.2281
5	5000	5000	1.78	0.15	1.90	0.10	[0.6777; 0.6928]	0.0152
16	5000	5000	0.55	22	0.67	30.5	[0.6116; 0.6281]	0.0166
27	5000	5000	-1.95	0.1	-1.8	0.18	[0.7699; 0.9505]	0.1806
PSO								
Case	***n***	***m***	***q***	***η***	***r***	***β***	***(1-γ) = 0.90***	**Width**
1	100	100	1.78	0.15	1.90	0.10	[0.5956; 0.7156]	0.1199
12	100	100	0.55	22	0.67	30.5	[0.5470; 0.6778]	0.1309
23	100	100	-1.95	0.1	-1.8	0.18	[0.7104; 0.8912]	0.1808
5	5000	5000	1.78	0.15	1.90	0.10	[0.6771; 0.6926]	0.0155
16	5000	5000	0.55	22	0.67	30.5	[0.6113; 0.6281]	0.0169
27	5000	5000	-1.95	0.1	-1.8	0.18	[0.7625; 0.7831]	0.0205

Table 4.7. Examples of cases that present a decrease of interval widths for a fixed n and an increase of m.

Nelder-Mead								
Case	***n***	***m***	***q***	***η***	***r***	***β***	***(1-γ) = 0.90***	**Width**
1	100	100	1.78	0.15	1.90	0.10	[0.5991; 0.7088]	0.1097
6	100	250	1.78	0.15	1.90	0.10	[0.6531; 0.7386]	0.0855
7	100	500	1.78	0.15	1.90	0.10	[0.6548; 0.7348]	0.08004
8	100	1000	1.78	0.15	1.90	0.10	[0.6995; 0.7694]	0.06985
PSO								
Case	***n***	***m***	***q***	***η***	***r***	***β***	***(1-γ) = 0.90***	**Width**
1	100	100	1.78	0.15	1.90	0.10	[0.5989; 0.7147]	0.1158
6	100	250	1.78	0.15	1.90	0.10	[0.6509; 0.7395]	0.0885
7	100	500	1.78	0.15	1.90	0.10	[0.7058; 0.7831]	0.0773
8	100	1000	1.78	0.15	1.90	0.10	[0.6586; 0.7227]	0.0641

Table 4.8. Examples of cases that present a decrease of interval widths for a fixed m and an increase of n.

Nelder-Mead								
Case	***n***	***m***	***q***	***η***	***r***	***β***	***(1-γ) = 0.90***	**Width**
12	100	100	0.55	22	0.67	30.5	[0.5462; 0.6803]	0.1341
20	250	100	0.55	22	0.67	30.5	[0.5786; 0.6840]	0.1055
21	500	100	0.55	22	0.67	30.5	[0.5926; 0.6929]	0.10023
22	1000	100	0.55	22	0.67	30.5	[0.5948; 0.6931]	0.09834
PSO								
Case	***n***	***m***	***q***	***η***	***r***	***β***	***(1-γ) = 0.90***	**Width**
12	100	100	0.55	22	0.67	30.5	[0.5470; 0.6778]	0.1309

20	250	100	0.55	22	0.67	30.5	[0.5752; 0.6855]	0.1103
21	500	100	0.55	22	0.67	30.5	[0.5785; 0.6748]	0.0964
22	1000	100	0.55	22	0.67	30.5	[0.5451; 0.6412]	0.0961

Next, we present charts in order to facilitate the evaluation of simulations for the confidence intervals constructed in this section. The data used to compile these figures are presented in Tables Table *4.9* to Table *4.12*.

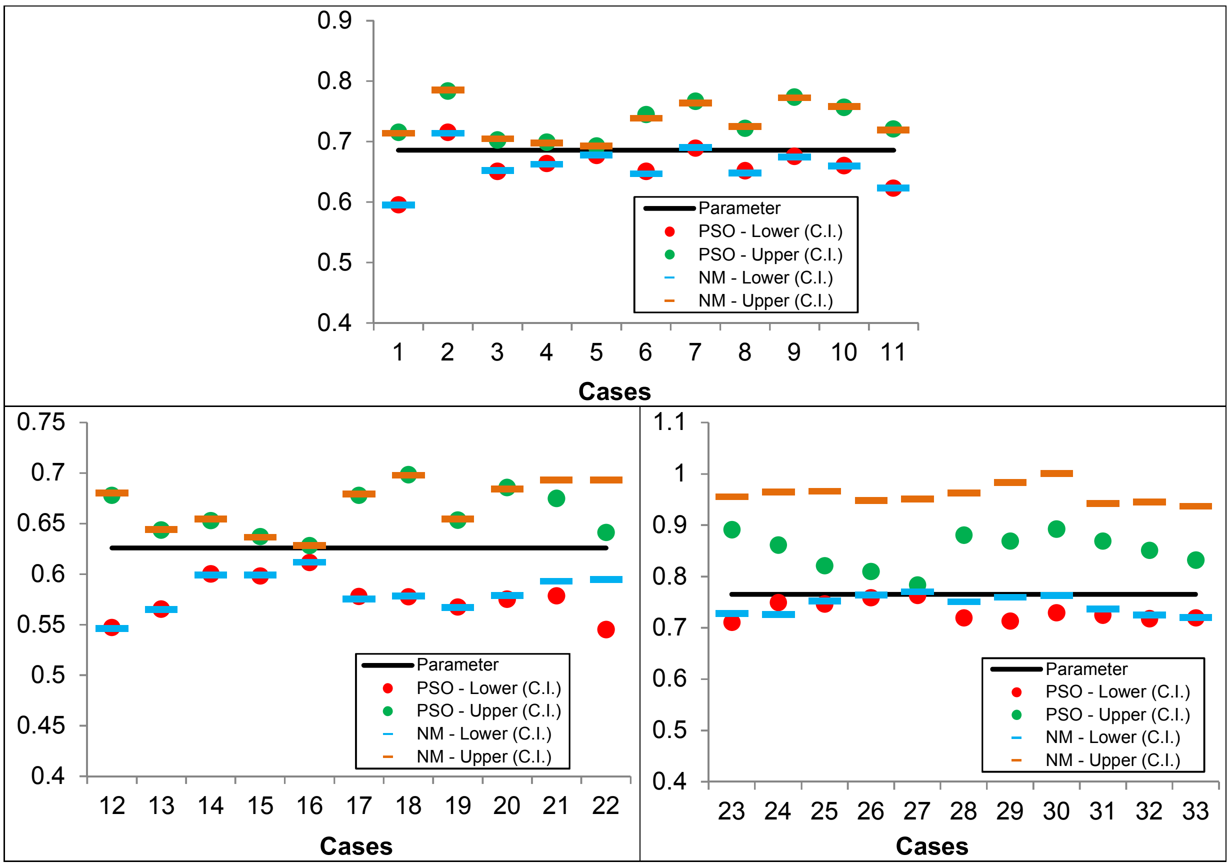


Figure 4.3. Upper and lower limits of the Bootstrap-P Confidence Interval for the Index R, obtained by PSO and Nelder-Mead.

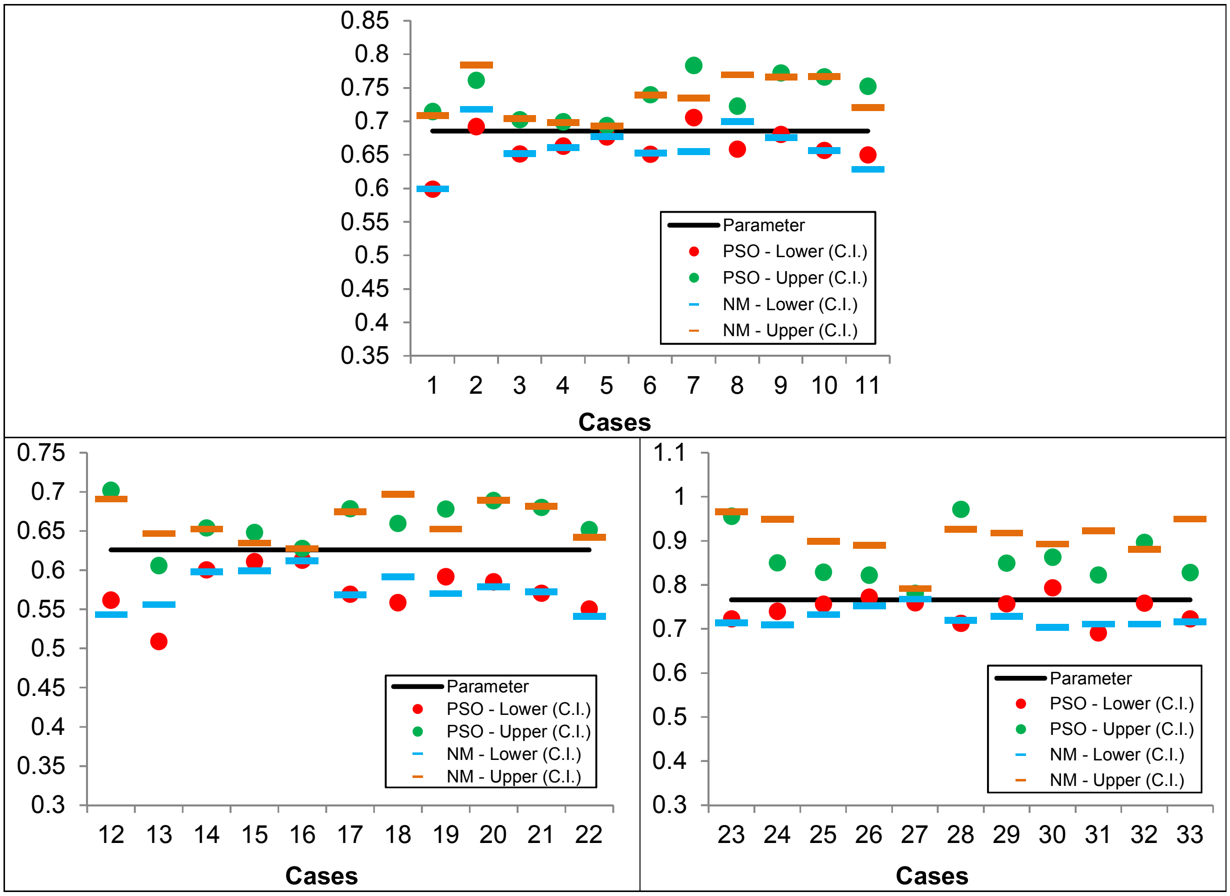


Figure 4.4. Upper and lower limits of the Non-Parametric Bootstrap Confidence Interval for the Index R, obtained by PSO and Nelder-Mead.

Note from Figures Figure *4.3* and Figure *4.4* that the parametric and non-parametric bootstrap methods, showed greater efficiency in the simulations of the confidence intervals for the index *R*. It is observed that for Nelder-Mead case, in both bootstrap approaches, only 3 cases shows intervals that not contain the real parameter (bootstrap-p: cases 2, 7 and 27 ; non-parametric bootstrap: cases 2, 8 and 27). When we use PSO, it is observed a similar behavior, *i.e.*, when we deal with the bootstrap-p approach, there is only two cases that the intervals not contain the real parameter (cases 2 and 7), for the non-parametric approach, the cases 2, 7, 13, 26 and 30 not contain the real parameter. We can yet comment that cases 23 to 33 showed interval widths greater than other simulated cases, apparently, these widths are greater in the Nelder-Mead case. In the next section, we will compare these lengths through hypothesis tests.

Table 4.9. Simulation results and estimation of bootstrap-p confidence interval for the index R – By Nelder-Mead

	Sample Size		**Initial Parameters**					**MLE Results for the first sample**					**Bootstrap Estimative**		
Case	***n***	***m***	q	η	r	β	R	$\hat{q}$	$\hat{\eta}$	$\hat{r}$	$\hat{\beta}$	$\hat{R}$	R^*	**Confidence Interval** $(1-\gamma) = 0.90$	**Width**
1	100	100	1.78	0.15	1.90	0.10	0.6855	1.7786	0.2152	1.9049	0.0680	0.6559	0.6562	[0.5949; 0.7139]	0.1190
2	250	250	1.78	0.15	1.90	0.10	0.6855	1.7516	0.1524	1.9040	0.1635	0.7504	0.7500	[0.7135; 0.7852]	0.0717
3	500	500	1.78	0.15	1.90	0.10	0.6855	1.7679	0.1658	1.8949	0.0960	0.6775	0.6780	[0.6521; 0.7044]	0.0523
4	1000	1000	1.78	0.15	1.90	0.10	0.6855	1.7795	0.1419	1.9008	0.0839	0.6809	0.6809	[0.6624; 0.6981]	0.0357
5	5000	5000	1.78	0.15	1.90	0.10	0.6855	1.7799	0.1539	1.8999	0.1023	0.6852	0.6853	[0.6777; 0.6928]	0.0152
6	100	250	1.78	0.15	1.90	0.10	0.6855	1.8083	0.1351	1.9000	0.1942	0.6952	0.6935	[0.6466; 0.7389]	0.0923
7	100	500	1.78	0.15	1.90	0.10	0.6855	1.7478	0.1764	1.8986	0.1480	0.7280	0.7279	[0.6903; 0.7643]	0.0741
8	100	1000	1.78	0.15	1.90	0.10	0.6855	1.7804	0.1335	1.8988	0.0984	0.6880	0.6870	[0.6484; 0.7247]	0.0763
9	250	100	1.78	0.15	1.90	0.10	0.6855	1.7645	0.1798	1.9134	0.0962	0.7219	0.7226	[0.6742; 0.7721]	0.0979
10	500	100	1.78	0.15	1.90	0.10	0.6855	1.7814	0.1252	1.9092	0.0886	0.7101	0.7103	[0.6594; 0.7579]	0.0985
11	1000	100	1.78	0.15	1.90	0.10	0.6855	1.7865	0.1429	1.9006	0.0851	0.6731	0.6724	[0.6233; 0.7191]	0.0958
12	100	100	0.55	22	0.67	30.5	0.6259	0.4668	23.760	0.6514	30.1452	0.6193	0.5910	[0.5462; 0.6803]	0.1341
13	250	250	0.55	22	0.67	30.5	0.6259	0.5433	21.323	0.6891	26.7851	0.6058	0.5609	[0.5651; 0.6442]	0.0791
14	500	500	0.55	22	0.67	30.5	0.6259	0.5688	21.164	0.7129	28.6894	0.6261	0.5277	[0.5993; 0.6543]	0.0550
15	1000	1000	0.55	22	0.67	30.5	0.6259	0.5136	23.615	0.7046	29.6139	0.6183	0.5682	[0.5990; 0.6365]	0.0375
16	5000	5000	0.55	22	0.67	30.5	0.6259	0.5046	23.435	0.6784	30.1115	0.6201	0.6201	[0.6116; 0.6281]	0.0166
17	100	250	0.55	22	0.67	30.5	0.6259	0.5657	20.850	0.5990	31.6486	0.6280	0.6228	[0.5750; 0.6790]	0.1040
18	100	500	0.55	22	0.67	30.5	0.6259	0.4687	21.546	0.7144	29.4955	0.6619	0.6437	[0.5784; 0.6974]	0.1190
19	100	1000	0.55	22	0.67	30.5	0.6259	0.5840	21.738	0.6679	29.7187	0.6120	0.6072	[0.5669; 0.6543]	0.0875
20	250	100	0.55	22	0.67	30.5	0.6259	0.5264	22.331	0.5092	35.5572	0.6304	0.6316	[0.5786; 0.6840]	0.1055
21	500	100	0.55	22	0.67	30.5	0.6259	0.5347	22.404	0.4926	37.6015	0.6354	0.6413	[0.5926; 0.6929]	0.1002
22	1000	100	0.55	22	0.67	30.5	0.6259	0.5412	22.235	0.4841	38.1364	0.6373	0.6252	[0.5948; 0.6931]	0.0983
23	100	100	-1.95	0.1	-1.8	0.18	0.7656	-1.9456	0.1000	-1.9409	0.1868	0.7628	0.8419	[0.7273; 0.9553]	0.2281
24	250	250	-1.95	0.1	-1.8	0.18	0.7656	-2.0719	0.1018	-0.9703	0.1262	0.7444	0.8149	[0.7258; 0.9645]	0.2387
25	500	500	-1.95	0.1	-1.8	0.18	0.7656	-2.0422	0.1013	-1.7182	0.1740	0.7723	0.8350	[0.7530; 0.9659]	0.2129
26	1000	1000	-1.95	0.1	-1.8	0.18	0.7656	-2.2531	0.1100	-1.9402	0.1884	0.7759	0.8286	[0.7636; 0.9476]	0.1840
27	5000	5000	-1.95	0.1	-1.8	0.18	0.7656	-2.3069	0.1116	-1.9518	0.1894	0.7793	0.8322	[0.7699; 0.9505]	0.1806
28	100	250	-1.95	0.1	-1.8	0.18	0.7656	-1.9610	0.0992	-1.9249	0.1845	0.7654	0.8495	[0.7509; 0.9618]	0.2109
29	100	500	-1.95	0.1	-1.8	0.18	0.7656	-2.1204	0.1040	-2.0452	0.1944	0.7818	0.8614	[0.7595; 0.9833]	0.2238
30	100	1000	-1.95	0.1	-1.8	0.18	0.7656	-2.1020	0.1011	-2.0229	0.1919	0.7896	0.8692	[0.7628; 1.0000]	0.2372
31	250	100	-1.95	0.1	-1.8	0.18	0.7656	-2.1965	0.1087	-1.9351	0.1892	0.7750	0.8367	[0.7370; 0.9421]	0.2051
32	500	100	-1.95	0.1	-1.8	0.18	0.7656	-2.1050	0.1049	-1.5737	0.1663	0.7663	0.8239	[0.7254; 0.9441]	0.2186
33	1000	100	-1.95	0.1	-1.8	0.18	0.7656	-2.1120	0.1050	-1.6466	0.1705	0.7679	0.8217	[0.7192; 0.9365]	0.2173

Table 4.10. Simulation results and estimation of non-parametric bootstrap confidence interval for the index R – By Nelder-Mead

	Sample Size		**Initial Parameters**					**MLE Results for the first sample**					**Bootstrap Estimative**		
Case	***n***	***m***	***q***	***η***	***r***	***β***	***R***	$\hat{q}$	$\hat{\eta}$	$\hat{r}$	$\hat{\beta}$	$\hat{R}$	R^*	**Confidence Interval**	
														(1-γ) = 0.90	**Width**
1	100	100	1.78	0.15	1.90	0.10	0.6855	1.7786	0.2152	1.9049	0.0680	0.6559	0.6565	[0.5991; 0.7088]	0.1097
2	250	250	1.78	0.15	1.90	0.10	0.6855	1.7516	0.1524	1.9040	0.1635	0.7504	0.7508	[0.7182; 0.7837]	0.0655
3	500	500	1.78	0.15	1.90	0.10	0.6855	1.7679	0.1658	1.8949	0.0960	0.6775	0.6779	[0.6517; 0.7042]	0.0526
4	1000	1000	1.78	0.15	1.90	0.10	0.6855	1.7795	0.1419	1.9008	0.0839	0.6809	0.6803	[0.6612; 0.6979]	0.0367
5	5000	5000	1.78	0.15	1.90	0.10	0.6855	1.7799	0.1539	1.8999	0.1023	0.6852	0.6851	[0.6773; 0.6930]	0.0157
6	100	250	1.78	0.15	1.90	0.10	0.6855	1.8083	0.1351	1.9000	0.1942	0.6952	0.6947	[0.6531; 0.7386]	0.0855
7	100	500	1.78	0.15	1.90	0.10	0.6855	1.7776	0.1633	1.8996	0.1039	0.6849	0.6944	[0.6548; 0.7348]	0.0800
8	100	1000	1.78	0.15	1.90	0.10	0.6855	1.7759	0.1621	1.8999	0.1014	0.6869	0.7349	[0.6995; 0.7694]	0.0698
9	250	100	1.78	0.15	1.90	0.10	0.6855	1.7645	0.1798	1.9134	0.0962	0.7219	0.7215	[0.6760; 0.7664]	0.0903
10	500	100	1.78	0.15	1.90	0.10	0.6855	1.7814	0.1252	1.9092	0.0886	0.7101	0.7104	[0.6560; 0.7672]	0.1112
11	1000	100	1.78	0.15	1.90	0.10	0.6855	1.7865	0.1429	1.9006	0.0851	0.6731	0.6736	[0.6283; 0.7210]	0.0927
12	100	100	0.55	22	0.67	30.5	0.6259	0.4668	23.760	0.6514	30.1452	0.6193	0.6054	[0.5439; 0.6911]	0.1472
13	250	250	0.55	22	0.67	30.5	0.6259	0.5433	21.323	0.6891	26.7851	0.6058	1.0384	[0.5566; 0.6468]	0.0902
14	500	500	0.55	22	0.67	30.5	0.6259	0.5688	21.164	0.7129	28.6894	0.6261	0.6619	[0.5980; 0.6531]	0.0550
15	1000	1000	0.55	22	0.67	30.5	0.6259	0.5136	23.615	0.7046	29.6139	0.6183	0.6177	[0.5994; 0.6344]	0.0350
16	5000	5000	0.55	22	0.67	30.5	0.6259	0.5046	23.435	0.6784	30.1115	0.6201	0.6203	[0.6121; 0.6281]	0.0160
17	100	250	0.55	22	0.67	30.5	0.6259	0.5657	20.850	0.5990	31.6486	0.6280	0.6255	[0.5689; 0.6747]	0.1058
18	100	500	0.55	22	0.67	30.5	0.6259	0.4687	21.546	0.7144	29.4955	0.6619	0.6528	[0.5919; 0.6967]	0.1048
19	100	1000	0.55	22	0.67	30.5	0.6259	0.5840	21.738	0.6679	29.7187	0.6120	0.6097	[0.5703; 0.6529]	0.0826
20	250	100	0.55	22	0.67	30.5	0.6259	0.5264	22.331	0.5092	35.5572	0.6304	0.6189	[0.5786; 0.6893]	0.1107
21	500	100	0.55	22	0.67	30.5	0.6259	0.5510	22.033	0.6558	30.9820	0.6258	0.6118	[0.5725; 0.6817]	0.1093
22	1000	100	0.55	22	0.67	30.5	0.6259	0.5758	21.350	0.6381	27.4955	0.5900	0.5911	[0.5411; 0.6425]	0.1015
23	100	100	-1.95	0.1	-1.8	0.18	0.7656	-1.9456	0.1000	-1.9409	0.1868	0.7628	0.8271	[0.7144; 0.9666]	0.2522
24	250	250	-1.95	0.1	-1.8	0.18	0.7656	-2.0719	0.1018	-0.9703	0.1262	0.7444	0.7962	[0.7092; 0.9486]	0.2395
25	500	500	-1.95	0.1	-1.8	0.18	0.7656	-2.0422	0.1013	-1.7182	0.1740	0.7723	0.8012	[0.7333; 0.8992]	0.1659
26	1000	1000	-1.95	0.1	-1.8	0.18	0.7656	-2.2531	0.1100	-1.9402	0.1884	0.7759	0.8058	[0.7527; 0.8897]	0.1370
27	5000	5000	-1.95	0.1	-1.8	0.18	0.7656	-2.3069	0.1116	-1.9518	0.1894	0.7793	0.7802	[0.7680; 0.7913]	0.0233
28	100	250	-1.95	0.1	-1.8	0.18	0.7656	-1.9610	0.0992	-1.9249	0.1845	0.7654	0.8105	[0.7198; 0.9260]	0.2062
29	100	500	-1.95	0.1	-1.8	0.18	0.7656	-2.1204	0.1040	-2.0452	0.1944	0.7818	0.8266	[0.7294; 0.9180]	0.1886
30	100	1000	-1.95	0.1	-1.8	0.18	0.7656	-2.1020	0.1011	-2.0229	0.1919	0.7896	0.8048	[0.7040; 0.8933]	0.1893
31	250	100	-1.95	0.1	-1.8	0.18	0.7656	-2.1965	0.1087	-1.9351	0.1892	0.7750	0.8084	[0.7105; 0.9231]	0.2126
32	500	100	-1.95	0.1	-1.8	0.18	0.7656	-2.1050	0.1049	-1.5737	0.1663	0.7663	0.7897	[0.7109; 0.8806]	0.1697
33	1000	100	-1.95	0.1	-1.8	0.18	0.7656	-2.1120	0.1050	-1.6466	0.1705	0.7679	0.8192	[0.7162; 0.9501]	0.2338

Table 4.11 - Simulation results and estimation of bootstrap-p confidence interval for the index R – By PSO

	Sample Size		**Initial Parameters**					**MLE Results for the first sample**					**Bootstrap Estimative**		
														Confidence Interval	
Case	***n***	***m***	***q***	***η***	***r***	***β***	***R***	$\hat{q}$	$\hat{\eta}$	$\hat{r}$	$\hat{\beta}$	$\hat{R}$	R^*	***(1-γ) = 0.90***	**Width**
1	100	100	1.78	0.15	1.90	0.10	0.6855	1.7786	0.2153	1.9050	0.0680	0.6560	0.6554	[0.5956; 0.7156]	0.1199
2	250	250	1.78	0.15	1.90	0.10	0.6855	1.7515	0.1524	1.9041	0.1634	0.7506	0.7499	[0.7157; 0.7836]	0.0679
3	500	500	1.78	0.15	1.90	0.10	0.6855	1.7679	0.1659	1.8949	0.0960	0.6775	0.6773	[0.6508; 0.7025]	0.0517
4	1000	1000	1.78	0.15	1.90	0.10	0.6855	1.7795	0.1419	1.9008	0.0840	0.6809	0.6813	[0.6635; 0.6990]	0.0356
5	5000	5000	1.78	0.15	1.90	0.10	0.6855	1.7800	0.1538	1.8999	0.1022	0.6850	0.6851	[0.6771; 0.6926]	0.0155
6	100	250	1.78	0.15	1.90	0.10	0.6855	1.8083	0.1352	1.9000	0.1943	0.6951	0.6969	[0.6507; 0.7449]	0.0942
7	100	500	1.78	0.15	1.90	0.10	0.6855	1.7478	0.1763	1.8986	0.1481	0.7281	0.7288	[0.6893; 0.7665]	0.0772
8	100	1000	1.78	0.15	1.90	0.10	0.6855	1.7805	0.1334	1.8988	0.0983	0.6880	0.6880	[0.6517; 0.7220]	0.0702
9	250	100	1.78	0.15	1.90	0.10	0.6855	1.7644	0.1799	1.9134	0.0963	0.7220	0.7231	[0.6755; 0.7735]	0.0980
10	500	100	1.78	0.15	1.90	0.10	0.6855	1.7814	0.1252	1.9092	0.0885	0.7099	0.7091	[0.6601; 0.7568]	0.0967
11	1000	100	1.78	0.15	1.90	0.10	0.6855	1.7866	0.1429	1.9006	0.0852	0.6731	0.6727	[0.6230; 0.7211]	0.0980
12	100	100	0.55	22	0.67	30.5	0.6259	0.4667	23.763	0.6518	30.1359	0.6193	0.6130	[0.5470; 0.6778]	0.1309
13	250	250	0.55	22	0.67	30.5	0.6259	0.5435	21.317	0.6893	26.7827	0.6059	0.6054	[0.5653; 0.6435]	0.0782
14	500	500	0.55	22	0.67	30.5	0.6259	0.5685	21.175	0.7129	28.6889	0.6260	0.6264	[0.6002; 0.6527]	0.0525
15	1000	1000	0.55	22	0.67	30.5	0.6259	0.5137	23.614	0.7046	29.6077	0.6182	0.6182	[0.5979; 0.6371]	0.0392
16	5000	5000	0.55	22	0.67	30.5	0.6259	0.5046	23.438	0.6786	30.0995	0.6199	0.6200	[0.6113; 0.6281]	0.0169
17	100	250	0.55	22	0.67	30.5	0.6259	0.5660	20.843	0.5986	31.6643	0.6281	0.6274	[0.5777; 0.6778]	0.1001
18	100	500	0.55	22	0.67	30.5	0.6259	0.4688	21.541	0.7146	29.4830	0.6618	0.6417	[0.5774; 0.6982]	0.1208
19	100	1000	0.55	22	0.67	30.5	0.6259	0.5842	21.736	0.6677	29.7264	0.6120	0.6108	[0.5675; 0.6534]	0.0859
20	250	100	0.55	22	0.67	30.5	0.6259	0.5260	22.340	0.5091	35.5524	0.6304	0.6340	[0.5752; 0.6855]	0.1103
21	500	100	0.55	22	0.67	30.5	0.6259	0.5506	22.042	0.6558	30.9777	0.6257	0.6280	[0.5785; 0.6748]	0.0964
22	1000	100	0.55	22	0.67	30.5	0.6259	0.5761	21.343	0.6384	27.4841	0.5899	0.5933	[0.5451; 0.6412]	0.0961
23	100	100	-1.95	0.1	-1.8	0.18	0.7656	-1.9455	0.1000	-1.9407	0.1868	0.7627	0.7975	[0.7104; 0.8912]	0.1808
24	250	250	-1.95	0.1	-1.8	0.18	0.7656	-2.0734	0.1018	-2.0321	0.1934	0.7830	0.8033	[0.7491; 0.8612]	0.1121
25	500	500	-1.95	0.1	-1.8	0.18	0.7656	-2.0433	0.1013	-1.7182	0.174	0.7725	0.7830	[0.7462; 0.8204]	0.0742
26	1000	1000	-1.95	0.1	-1.8	0.18	0.7656	-2.1423	0.1054	-1.9363	0.1881	0.7777	0.7841	[0.7588; 0.8093]	0.0505
27	5000	5000	-1.95	0.1	-1.8	0.18	0.7656	-1.9665	0.1004	-1.9511	0.1893	0.7699	0.7722	[0.7625; 0.7831]	0.0205
28	100	250	-1.95	0.1	-1.8	0.18	0.7656	-1.9620	0.0993	-1.9246	0.1844	0.7652	0.7990	[0.7194; 0.8802]	0.1608
29	100	500	-1.95	0.1	-1.8	0.18	0.7656	-1.5206	0.0856	-2.0472	0.1945	0.7551	0.7901	[0.7128; 0.8688]	0.1560
30	100	1000	-1.95	0.1	-1.8	0.18	0.7656	-2.1035	0.1012	-2.0234	0.1920	0.7896	0.8210	[0.7292; 0.8922]	0.1630
31	250	100	-1.95	0.1	-1.8	0.18	0.7656	-2.1969	0.1087	-1.9325	0.189	0.7748	0.7954	[0.7245; 0.8686]	0.1441
32	500	100	-1.95	0.1	-1.8	0.18	0.7656	-2.1076	0.105	-1.5752	0.1664	0.7663	0.7819	[0.7178; 0.8506]	0.1328
33	1000	100	-1.95	0.1	-1.8	0.18	0.7656	-1.7789	0.0941	-1.6472	0.1706	0.7582	0.7727	[0.7192; 0.8319]	0.1127

Table 4.12 - Simulation results and estimation of non-parametric bootstrap confidence interval for the index R – By PSO

	Sample Size		**Initial Parameters**					**MLE Results for the first sample**					**Bootstrap Estimative**		
Case	***n***	***m***	***q***	***η***	***r***	***β***	***R***	$\hat{q}$	$\hat{\eta}$	$\hat{r}$	$\hat{\beta}$	$\hat{R}$	R^*	**Confidence Interval** ***(1-γ) = 0.90***	**Width**
1	100	100	1.78	0.15	1.90	0.10	0.6855	1.7786	0.2153	1.9050	0.0680	0.6560	0.6559	[0.5989; 0.7147]	0.1158
2	250	250	1.78	0.15	1.90	0.10	0.6855	1.7515	0.1524	1.9041	0.1634	0.7506	0.7277	[0.6922; 0.7613]	0.0691
3	500	500	1.78	0.15	1.90	0.10	0.6855	1.7679	0.1659	1.8949	0.0960	0.6775	0.6773	[0.6512; 0.7026]	0.0514
4	1000	1000	1.78	0.15	1.90	0.10	0.6855	1.7795	0.1419	1.9008	0.0840	0.6809	0.6809	[0.6628; 0.6995]	0.0367
5	5000	5000	1.78	0.15	1.90	0.10	0.6855	1.7800	0.1538	1.8999	0.1022	0.6850	0.6851	[0.6767; 0.6933]	0.0166
6	100	250	1.78	0.15	1.90	0.10	0.6855	1.8083	0.1352	1.9000	0.1943	0.6951	0.6962	[0.6509; 0.7395]	0.0885
7	100	500	1.78	0.15	1.90	0.10	0.6855	1.7478	0.1763	1.8986	0.1481	0.7281	0.7442	[0.7058; 0.7831]	0.0773
8	100	1000	1.78	0.15	1.90	0.10	0.6855	1.7805	0.1334	1.8988	0.0983	0.6880	0.6896	[0.6586; 0.7227]	0.0641
9	250	100	1.78	0.15	1.90	0.10	0.6855	1.7644	0.1799	1.9134	0.0963	0.7220	0.7251	[0.6804; 0.7722]	0.0918
10	500	100	1.78	0.15	1.90	0.10	0.6855	1.7814	0.1252	1.9092	0.0885	0.7099	0.7121	[0.6565; 0.7661]	0.1097
11	1000	100	1.78	0.15	1.90	0.10	0.6855	1.7866	0.1429	1.9006	0.0852	0.6731	0.7018	[0.6498; 0.7523]	0.1025
12	100	100	0.55	22	0.67	30.5	0.6259	0.4667	23.763	0.6518	30.1359	0.6193	0.6305	[0.5619; 0.7021]	0.1402
13	250	250	0.55	22	0.67	30.5	0.6259	0.5435	21.317	0.6893	26.7827	0.6059	0.5594	[0.5090; 0.6060]	0.0970
14	500	500	0.55	22	0.67	30.5	0.6259	0.5685	21.175	0.7129	28.6889	0.6260	0.6269	[0.6007; 0.6541]	0.0534
15	1000	1000	0.55	22	0.67	30.5	0.6259	0.5137	23.614	0.7046	29.6077	0.6182	0.6296	[0.6110; 0.6485]	0.0375
16	5000	5000	0.55	22	0.67	30.5	0.6259	0.5046	23.438	0.6786	30.0995	0.6199	0.6203	[0.6128; 0.6279]	0.0152
17	100	250	0.55	22	0.67	30.5	0.6259	0.5660	20.843	0.5986	31.6643	0.6281	0.6222	[0.5696; 0.6785]	0.1089
18	100	500	0.55	22	0.67	30.5	0.6259	0.4688	21.541	0.7146	29.4830	0.6618	0.6061	[0.5586; 0.6598]	0.1012
19	100	1000	0.55	22	0.67	30.5	0.6259	0.5842	21.736	0.6677	29.7264	0.6120	0.6371	[0.5918; 0.6783]	0.0865
20	250	100	0.55	22	0.67	30.5	0.6259	0.5260	22.340	0.5091	35.5524	0.6304	0.6378	[0.5851; 0.6891]	0.1039
21	500	100	0.55	22	0.67	30.5	0.6259	0.5506	22.042	0.6558	30.9777	0.6257	0.6261	[0.5708; 0.6803]	0.1095
22	1000	100	0.55	22	0.67	30.5	0.6259	0.5761	21.343	0.6384	27.4841	0.5899	0.5998	[0.5506; 0.6520]	0.1013
23	100	100	-1.95	0.1	-1.8	0.18	0.7656	-1.9455	0.1000	-1.9407	0.1868	0.7627	0.8388	[0.7228; 0.9562]	0.2334
24	250	250	-1.95	0.1	-1.8	0.18	0.7656	-2.0734	0.1018	-2.0321	0.1934	0.7830	0.7966	[0.7400; 0.8499]	0.1099
25	500	500	-1.95	0.1	-1.8	0.18	0.7656	-2.0433	0.1013	-1.7182	0.174	0.7725	0.7922	[0.7569; 0.8287]	0.0718
26	1000	1000	-1.95	0.1	-1.8	0.18	0.7656	-2.1423	0.1054	-1.9363	0.1881	0.7777	0.7962	[0.7721; 0.8222]	0.0502
27	5000	5000	-1.95	0.1	-1.8	0.18	0.7656	-1.9665	0.1004	-1.9511	0.1893	0.7699	0.7705	[0.7594; 0.7810]	0.0215
28	100	250	-1.95	0.1	-1.8	0.18	0.7656	-1.9620	0.0993	-1.9246	0.1844	0.7652	0.8148	[0.7130; 0.9721]	0.2591
29	100	500	-1.95	0.1	-1.8	0.18	0.7656	-1.5206	0.0856	-2.0472	0.1945	0.7551	0.8097	[0.7573; 0.8497]	0.0924
30	100	1000	-1.95	0.1	-1.8	0.18	0.7656	-2.1035	0.1012	-2.0234	0.1920	0.7896	0.8329	[0.7932; 0.8632]	0.0700
31	250	100	-1.95	0.1	-1.8	0.18	0.7656	-2.1969	0.1087	-1.9325	0.189	0.7748	0.7511	[0.6910; 0.8228]	0.1318
32	500	100	-1.95	0.1	-1.8	0.18	0.7656	-2.1076	0.105	-1.5752	0.1664	0.7663	0.8300	[0.7585; 0.8967]	0.1382
33	1000	100	-1.95	0.1	-1.8	0.18	0.7656	-1.7789	0.0941	-1.6472	0.1706	0.7582	0.7781	[0.7233; 0.8282]	0.1049

4.4 Hypothesis tests applied to compare the quality of the estimation methods

In this section, we will use the data obtained from the simulations that were made in order to evaluate the performance of the maximum likelihood estimator for the index R (Table 4.4 and Table 4.5) as well as the data generated in the numeric experiments made to verify the quality of the confidence interval generated by bootstrap-p and non-parametric bootstrap (Table 4.9 to Table 4.12).

We will use hypothesis tests in order to verify if there are significant differences between the different methods of estimation that were presented in this work. We consider the variables tested in this section as paired variables, once we use the same samples in order to calculate the point estimates by PSO and Nelder-Mead Method. As the same manner, the interval estimates are calculated from the same sample, we only modify the method that was used.

4.4.1 Comparing the Mean Squared Error obtained in the estimation of the R index by Nelder-Mead and PSO

The following variables will be used for the tests of this subsection:

- **Variable 1:** MSE obtained by Nelder-Mead (Cases 1 to 33).
- **Variable 2:** MSE obtained by Nelder-Mead (Cases 1 to 11).
- **Variable 3:** MSE obtained by Nelder-Mead (Cases 12 to 22).
- **Variable 4:** MSE obtained by Nelder-Mead (Cases 23 to 33).
- **Variable 5:** MSE obtained by PSO (Cases 1 to 33).
- **Variable 6:** MSE obtained by PSO (Cases 1 to 11).
- **Variable 7:** MSE obtained by PSO (Cases 12 to 22).
- **Variable 8:** MSE obtained by PSO (Cases 23 to 33).

Note that the Variables 1 to 4 were obtained from Table 4.4 and Variables 5 to 8 were taken from Table 4.5.

In order to point the test more adequate for this situation, it is necessary to verify the normality of the data sets, thus, the Variables were tested by the Shapiro-Test of normality [75]:

Table 4.13- p-values for the Shapiro-Wilk test of normality applied in Variables 1 to 8.

Tested Variables	***p*-value**
Variable 1	3.24 E-07*
Variable 2	0.7316
Variable 3	0.6245
Variable 4	0.6367
Variable 5	0.0005*
Variable 6	0.7270
Variable 7	0.6011
Variable 8	0.8452

* The test is significant for $\alpha = 0.01$

We can observe from the *p*-values of Table 4.13 that the only Variables that not are normally distributed are Variable 1 and Variable 5. However, the fact of this two Variables not present the normality does not mean that we do not must apply the Student's t-test for paired samples, once this two Variables has more than 30 values, and the central limit theorem [91], ensures that the average of these Variables are normally distributed.

So, we will proceed with the Student's t-test (paired samples) for the following Comparisons , in order to verify if on average, the Variables present equal means:

- **Comparison 1 - Variable 1**: MSE obtained by Nelder-Mead (Cases 1 to 33) "VS" **Variable 5:** MSE obtained by PSO (Cases 1 to 33).
- **Comparison 2 - Variable 2:** MSE obtained by Nelder-Mead (Cases 1 to 11) "VS" **Variable 6:** MSE obtained by PSO (Cases 1 to 11).
- **Comparison 3 - Variable 3:** MSE obtained by Nelder-Mead (Cases 12 to 22) "VS" **Variable 7:** MSE obtained by PSO (Cases 12 to 22).
- **Comparison 4 - Variable 4:** MSE obtained by Nelder-Mead (Cases 23 to 33) "VS" **Variable 8:** MSE obtained by PSO (Cases 23 to 33).

Formally, we can write the hypotheses used in the comparisons listed above as follows:

$$\boldsymbol{H_0}: \boldsymbol{\mu_{MSE}(NM) - \mu_{MSE}(PSO) = 0}$$

"vs"

$$\boldsymbol{H_1}: \boldsymbol{\mu_{MSE}(NM) - \mu_{MSE}(PSO) \neq 0}$$

Table 4.14 – Results for the Student's t-test for the Comparisons 1 to 4.

	Mean of the difference between the two Variables.	Confidence Interval (95%) for the Difference between the two Variables.	*p*-value
Comparison 1	0.0081	C.I.(95%) = [0.0039, 0.0123]	0.0004*
Comparison 2	9.09E-7	C.I.(95%) = [-1.1165, 2.9346]	0.3409
Comparison 3	6.16E-21	C.I.(95%) = [-5.20E-06, 5.20E-06]	0.9999
Comparison 4	0.0243	C.I.(95%) = [0.0224, 0.0262]	6.99E-11*

* The test is significant for $\alpha = 0.01$

From Table 4.14, we can observe that only the Comparison 1 and 4 reject the hypothesis of equality between the averages, so, the confidence interval indicates that the difference between the two averages is positive, *i.e.*, on average the MSE obtained by Nelder-Mead, showed higher values than those obtained by PSO. Observe that the Case 1, is completely influenced by the Case 4, once the Case 1, comprises all analyzed Cases and among these Cases the only one with significant differences was the Case 4.

4.4.2 Comparing the width of the 90% Confidence Interval obtained by bootstrap-P and Non-parametric bootstrap (Nelder-Mead Case)

From the analysis of this subsection we will use the variables 9 to 12 (taken from Table 4.9) and Variables 13 to 16 (taken from Table 4.10):

- **Variable 9:** Width of the 90% confidence interval obtained by bootstrap-P approach for Cases 1 to 33 (Nelder-Mead).
- **Variable 10:** Width of the 90% confidence interval obtained by bootstrap-P approach for Cases 1 to 11 (Nelder-Mead).

- **Variable 11:** Width of the 90% confidence interval obtained by bootstrap-P approach for Cases 12 to 22 (Nelder-Mead).
- **Variable 12:** Width of the 90% confidence interval obtained by bootstrap-P approach for Cases 23 to 33 (Nelder-Mead).
- **Variable 13:** Width of the 90% confidence interval obtained by non-parametric bootstrap approach for Cases 1 to 33 (Nelder-Mead).
- **Variable 14:** Width of the 90% confidence interval obtained by non-parametric bootstrap approach for Cases 1 to 11(Nelder-Mead).
- **Variable 15:** Width of the 90% confidence interval obtained by non-parametric bootstrap approach for Cases 12 to 22 (Nelder-Mead).
- **Variable 16:** Width of the 90% confidence interval obtained by non-parametric bootstrap approach for Cases 23 to 33 (Nelder-Mead).

The normality of the data was tested by the Shapiro-Wilk test:

Table 4.15- p-values for the Shapiro-Wilk test of normality applied in Variables 9 to 16.

Tested Variables	***p*-value**
Variable 9	0.0078*
Variable 10	0.5241
Variable 11	0.4650
Variable 12	0.3846
Variable 13	0.0866
Variable 14	0.6754
Variable 15	0.3515
Variable 16	0.0450**

* The test is significant for $\alpha = 0.01$

** The test is significant for $\alpha = 0.05$

Table 4.15 indicates that the Variables 9 and 16 are not normally distributed, so any comparison made with the Variable 16 will not be performed by the Student's test-t, in situations that involves the Variable 16 we will use the Wilcoxon test. Any comparison made with the Variable 9, will be done by the Student's test-t, once this Variable presents *n > 30*. The following comparisons were performed:

- **Comparison 5 - Variable 9:** Width of the 90% confidence interval obtained by bootstrap-P approach for Cases 1 to 33 (Nelder-Mead) ***"VS"*** **Variable 13:** Width of the 90% confidence interval obtained by non-parametric bootstrap approach for Cases 1 to 33 (Nelder-Mead).
- **Comparison 6 – Variable10:** Width of the 90% confidence interval obtained by bootstrap-P approach for Cases 1 to 11 (Nelder-Mead) ***"VS"*** **Variable 14:** Width of the 90% confidence interval obtained by non-parametric bootstrap approach for Cases 1 to 11 (Nelder-Mead).
- **Comparison 7 - Variable 11:** Width of the 90% confidence interval obtained by bootstrap-P approach for Cases 12 to 22 (Nelder-Mead) ***"VS"*** **Variable 15:** Width of the 90% confidence interval obtained by non-parametric bootstrap approach for Cases 12 to 22 (Nelder-Mead).
- **Comparison 8 - Variable 12:** Width of the 90% confidence interval obtained by bootstrap-P approach for Cases 23 to 33 (Nelder-Mead) ***"VS"*** **Variable 16:** Width of the 90% confidence interval obtained by non-parametric bootstrap approach for Cases 23 to 33 (Nelder-Mead).

Formally, we can write the hypotheses used in the comparisons listed above as follows:

$$\boldsymbol{H_0}: \boldsymbol{\mu_{width}(boot-p) - \mu_{width}(boot-np) = 0}$$

"vs"

$$\boldsymbol{H_1}: \boldsymbol{\mu_{width}(boot-p) - \mu_{width}(boot-np) \neq 0}$$

Table 4.16 - Results for the Student's t-test / Wilcox test for the Comparisons 5 to 8.

	Mean of the difference between the two Variables.	**Confidence Interval (95%) for the Difference between the two Variables.**	***p*-value**
Comparison 5	0.0102	C.I.(95%) = [-0.0013, 0.0217]	0.0799(T)
Comparison 6	0.0018	C.I.(95%) = [-0.0023, 0.0059]	0.3510(T)
Comparison 7	-0.0020	C.I.(95%) = [-0.0073, 0.0033]	0.4242(T)
Comparison 8	0.0231	C.I.(95%) = [-0.0041, 0.0666]	0.0830(W)

(T) = Student's t-test
(W) = Wilcoxon Test

According with Table 4.16, we can conclude that there is none significant difference between the variables tested, *i.e.*, there is no difference between the width of the bootstrap-P confidence interval and the width of non-parametric bootstrap confidence interval when we estimate these confidence intervals by Nelder-Mead method.

4.4.3 Comparing the width of the 90% Confidence Interval obtained by bootstrap-P and Non-parametric bootstrap (PSO Case)

For this analysis the variables 17 to 20 were taken from Table 4.11 and Variables 21 to 24 were collected from Table 4.12. The variables are described next:

- **Variable 17:** Width of the 90% confidence interval obtained by bootstrap-P approach for Cases 1 to 33 (PSO).
- **Variable 18:** Width of the 90% confidence interval obtained by bootstrap-P approach for Cases 1 to 11 (PSO).
- **Variable 19:** Width of the 90% confidence interval obtained by bootstrap-P approach for Cases 12 to 22 (PSO).
- **Variable 20:** Width of the 90% confidence interval obtained by bootstrap-P approach for Cases 23 to 33 (PSO).
- **Variable 21:** Width of the 90% confidence interval obtained by non-parametric bootstrap approach for Cases 1 to 33 (PSO).
- **Variable 22:** Width of the 90% confidence interval obtained by non-parametric bootstrap approach for Cases 1 to 11 (PSO).
- **Variable 23:** Width of the 90% confidence interval obtained by non-parametric bootstrap approach for Cases 12 to 22 (PSO).
- **Variable 24:** Width of the 90% confidence interval obtained by non-parametric bootstrap approach for Cases 23 to 33 (PSO).

Testing the variables about the normality, we have:

Table 4.17- p-values for the Shapiro-Wilk test of normality applied in Variables 17 to 24.

Tested Variables	*p*-value
Variable 17	0.7271
Variable 18	0.5895
Variable 19	0.5440
Variable 20	0.3110
Variable 21	0.0014*
Variable 22	0.8234
Variable 23	0.1555
Variable 24	0.2088

* The test is significant for $\alpha = 0.01$

As we can observe from Table 4.17, only the Variable 21 is not normally distributed, however, this Variable present more than 30 values, so the comparison made with this Variable can be proceed by the Student's t-test (paired samples). Next, we present the comparisons that were tested:

- **Comparison 9 - Variable 17:** Width of the 90% confidence interval obtained by bootstrap-P approach for Cases 1 to 33 (PSO) ***"VS"*** **Variable 21:** Width of the 90% confidence interval obtained by non-parametric bootstrap approach for Cases 1 to 33 (PSO).
- **Comparison 10 - Variable 18:** Width of the 90% confidence interval obtained by bootstrap-P approach for Cases 1 to 11 (PSO) ***"VS"*** **Variable 22:** Width of the 90% confidence interval obtained by non-parametric bootstrap approach for Cases 1 to 11 (PSO).
- **Comparison 11 - Variable 19:** Width of the 90% confidence interval obtained by bootstrap-P approach for Cases 12 to 22 (PSO) ***"VS"*** **Variable 23:** Width of the 90% confidence interval obtained by non-parametric bootstrap approach for Cases 12 to 22 (PSO).
- **Comparison 12 - Variable 20:** Width of the 90% confidence interval obtained by bootstrap-P approach for Cases 23 to 33 (PSO) ***"VS"*** **Variable 24:** Width of the 90%

confidence interval obtained by non-parametric bootstrap approach for Cases 23 to 33 (PSO).

Formally, we can write the hypotheses used in the comparisons listed above as follows:

$$\boldsymbol{H_0}: \boldsymbol{\mu_{width}(boot-p) - \mu_{width}(boot-np) = 0}$$

"vs"

$$\boldsymbol{H_1}: \boldsymbol{\mu_{width}(boot-p) - \mu_{width}(boot-np) \neq 0}$$

The results for the tests of Comparisons 9 to 12 are presented in the following table:

Table 4.18- Results for the Student's t-test for the Comparisons 9 to 12.

	Mean of the difference between the two Variables.	**Confidence Interval (95%) for the Difference between the two Variables.**	***p*-value**
Comparison 9	-4.84E-5	C.I.(95%) = [-0.0103, 0.0102]	0.9924
Comparison 10	0.0001	C.I.(95%) = [-0.0037, 0.0039]	0.9418
Comparison 11	-0.0025	C.I.(95%) = [-0.0094, 0.0045]	0.4467
Comparison 12	0.0022	C.I.(95%) = [-0.0316, 0.0360]	0.8871

Thus, from Table 4.18, we conclude that there is no difference significant between any variable tested, therefore the same manner as in the case Nelder-Mead, there is no difference between the width of the bootstrap-P confidence interval and the width of non-parametric bootstrap confidence interval when we estimate these confidence intervals by PSO.

4.4.4 Comparing the width of the Confidence Interval obtained by Nelder-Mead and PSO (Bootstrap-P Case).

The following Variables will be tested:

- **Variable 9:** Width of the 90% confidence interval obtained by bootstrap-P approach for Cases 1 to 33.

- **Variable 10:** Width of the 90% confidence interval obtained by bootstrap-P approach for Cases 1 to 11.
- **Variable 11:** Width of the 90% confidence interval obtained by bootstrap-P approach for Cases 12 to 22.
- **Variable 12:** Width of the 90% confidence interval obtained by bootstrap-P approach for Cases 23 to 33.
- **Variable 17:** Width of the 90% confidence interval obtained by bootstrap-P approach for Cases 1 to 33 (PSO).
- **Variable 18:** Width of the 90% confidence interval obtained by bootstrap-P approach for Cases 1 to 11 (PSO).
- **Variable 19:** Width of the 90% confidence interval obtained by bootstrap-P approach for Cases 12 to 22 (PSO).
- **Variable 20:** Width of the 90% confidence interval obtained by bootstrap-P approach for Cases 23 to 33 (PSO).

Note that, previously we apply the Shapiro-Wilk test for verify the normality of these Variables (See Table 4.15 and Table 4.17). We point out that among these the only Variable that is not normally distributed is the Variable 9. However, the Variable 9 has size larger than 30, so, we can proceed with the Student's test-t even when the comparison involves this Variable.

The following comparisons are evaluated in this subsection.

- **Comparison 13 - Variable 4: Variable 9:** Width of the 90% confidence interval obtained by bootstrap-P approach for Cases 1 to 33 ***"VS"*** **Variable 17:** Width of the 90% confidence interval obtained by bootstrap-P approach for Cases 1 to 33 (PSO).
- **Comparison 14 – Variable 10:** Width of the 90% confidence interval obtained by bootstrap-P approach for Cases 1 to 11 ***"VS"*** **Variable 18:** Width of the 90% confidence interval obtained by bootstrap-P approach for Cases 1 to 11 (PSO).
- **Comparison 15 – Variable 11:** Width of the 90% confidence interval obtained by bootstrap-P approach for Cases 12 to 22 ***"VS"*** **Variable 19:** Width of the 90% confidence interval obtained by bootstrap-P approach for Cases 12 to 22 (PSO).

- **Comparison 16 - Variable 12:** Width of the 90% confidence interval obtained by bootstrap-P approach for Cases 23 to 33 ***"VS"*** **Variable 20:** Width of the 90% confidence interval obtained by bootstrap-P approach for Cases 23 to 33 (PSO).

The hypotheses can be written as following:

$$\boldsymbol{H_0: \mu_{width}(PSO) - \mu_{width}(NM) = 0}$$
"vs"
$$\boldsymbol{H_1: \mu_{width}(PSO) - \mu_{width}(NM) \neq 0}$$

The results for the Student's t-test are showed in the next table:

Table 4.19- Results for the Student's t-test for the Comparisons 13 to 16.

	Mean of the difference between the two Variables.	**Confidence Interval (95%) for the Difference between the two Variables.**	***p*-value**
Comparison 13	-0.0322	C.I.(95%) = [-0.0501, -0.0143]	0.0009*
Comparison 14	-0.0003	C.I.(95%) = [-0.0022, 0.0014]	0.6723
Comparison 15	-0.0008	C.I.(95%) = [-0.0026, 0.0010]	0.3402
Comparison 16	-0.0954	C.I.(95%) = [-0.1218, -0.0690]	1.10E-5*

* The test is significant for $\alpha = 0.01$

As we can observe from Table 4.19, for the Comparisons 14 and 15, we not detect any significant difference. However, the Comparison 13 and 16 indicate significant difference between the width of the confidence interval bootstrap-p calculated by Nelder-Mead and the width of the confidence interval bootstrap-p calculated by PSO. Observe, by the confidence interval of the difference between the variables that in both comparisons (13 and 16), the confidence interval indicates a negative difference, so we can conclude that in both cases the width of the intervals obtained by Nelder-Mead is larger than the width of the interval obtained by PSO. Also note that the data of the variables tested in the Comparison 16 are contained in the variables of the Comparison 13, and certainly the result of the Comparison 13 is being influenced by these values.

4.4.5 Comparing the width of the Confidence Interval obtained by Nelder-Mead and PSO (Non-Parametric Bootstrap Case).

For these comparisons we will use the following variables:

- **Variable 13:** Width of the 90% confidence interval obtained by non-parametric bootstrap approach for Cases 1 to 33 (Nelder-Mead).
- **Variable 14:** Width of the 90% confidence interval obtained by non-parametric bootstrap approach for Cases 1 to 11(Nelder-Mead).
- **Variable 15:** Width of the 90% confidence interval obtained by non-parametric bootstrap approach for Cases 12 to 22 (Nelder-Mead).
- **Variable 16:** Width of the 90% confidence interval obtained by non-parametric bootstrap approach for Cases 23 to 33 (Nelder-Mead).
- **Variable 21:** Width of the 90% confidence interval obtained by non-parametric bootstrap approach for Cases 1 to 33 (PSO).
- **Variable 22:** Width of the 90% confidence interval obtained by non-parametric bootstrap approach for Cases 1 to 11 (PSO).
- **Variable 23:** Width of the 90% confidence interval obtained by non-parametric bootstrap approach for Cases 12 to 22 (PSO).
- **Variable 24:** Width of the 90% confidence interval obtained by non-parametric bootstrap approach for Cases 23 to 33 (PSO).

Previously we present for the Variables 13 to 16 and 21 to 24, the results for the Shapiro Wilk Test (see Table 4.15 and Table 4.17) in order to verify the normality of the data. In this occasion, we had identified that among these Variables, only Variables 16 and 21 are not normally distributed. So, we will apply the Wilcoxon Test to perform comparisons that involves the Variable 16. For the Variable 21 there is no problem, once it has more than 30 values, and we can use the Student's t-test for paired samples.

The following comparisons were performed:

- **Comparison 17 - Variable 13:** Width of the 90% confidence interval obtained by non-parametric bootstrap approach for Cases 1 to 33 (Nelder-Mead) ***"VS"*** **Variable**

21: Width of the 90% confidence interval obtained by non-parametric bootstrap approach for Cases 1 to 33 (PSO).

- **Comparison 18 - Variable 14:** Width of the 90% confidence interval obtained by non-parametric bootstrap approach for Cases 1 to 11 (Nelder-Mead) ***"VS"*** **Variable 22:** Width of the 90% confidence interval obtained by non-parametric bootstrap approach for Cases 1 to 11 (PSO).
- **Comparison 19 - Variable 15:** Width of the 90% confidence interval obtained by non-parametric bootstrap approach for Cases 12 to 22 (Nelder-Mead) ***"VS"*** **Variable 23:** Width of the 90% confidence interval obtained by non-parametric bootstrap approach for Cases 12 to 22 (PSO).
- **Comparison 20 - Variable 16:** Width of the 90% confidence interval obtained by non-parametric bootstrap approach for Cases 23 to 33 (Nelder-Mead) ***"VS"*** **Variable 24:** Width of the 90% confidence interval obtained by non-parametric bootstrap approach for Cases 23 to 33 (PSO).

The hypotheses can be written as following:

$$\boldsymbol{H_0}: \boldsymbol{\mu_{width}(PSO) - \mu_{width}(NM) = 0}$$

"vs"

$$\boldsymbol{H_1}: \boldsymbol{\mu_{width}(PSO) - \mu_{width}(NM) \neq 0}$$

The results for the Student's t-test are showed in the next table:

Table 4.20 - Results for the Student's t-test / Wilcox test for the Comparisons 17 to 20.

	Mean of the difference between the two Variables.	Confidence Interval (95%) for the Difference between the two Variables.	*p*-value
Comparison 17	-0.0219	C.I.(95%) = [-0.0383, -0.0055]	0.0104 (T)**
Comparison 18	0.0013	C.I.(95%) = [-0.0025, 0.0051]	0.4638 (T)
Comparison 19	-0.0003	C.I.(95%) = [-0.0032, 0.0026]	0.8125 (T)
Comparison 20	-0.07145	C.I.(95%) = [-0.1082, -0.0206]	0.0068 (W)*

* The test is significant for $\alpha = 0.01$
** The test is significant for $\alpha = 0.05$
(T) = Student's t-test
(W) = Wilcoxon Test

From Table 4.20, we can conclude that the Comparison 17 and 20 indicate significant difference between the width of the confidence interval non-parametric bootstrap calculated by Nelder-Mead and the width of the confidence interval non-parametric bootstrap calculated by PSO. The confidence interval for the Difference between the two Variables indicates that in both comparisons (17 and 20), the width obtained by Nelder-Mead is larger than the with obtained by PSO.

5 CASE STUDIES FOR FATIGUE LIFE WITH EXTREMELY LARGE VALUES: APPLICATIONS OF THE NEW PROPOSED STRESS-STRENGTH MODEL.

In this section, we present two case studies in which stress (*Y*) and strength (*X*) follow *q*-Exponential distributions. The first case study, which was originally described in [83], deals with the experimental determination of high cycle fatigue of ductile cast iron used for wind turbine components and the second one, which was first reported in [84], evaluates the gigacycle fatigue life of high-strength steel.

These two case studies are based on the well-known phenomenon [44], [84] that the fatigue strength or endurance limit of large members is lower than that of small specimens made of the same material; in other words, a specimen size effect exists, *i.e.*, larger specimens fail at shorter fatigue lives than smaller specimens [25], [44]. In fact, design of parts and structures against fatigue is based on laboratory sized specimens which are usually smaller than the real ones. Therefore, it is of great importance to determine the reliability of larger specimens when data of smaller specimens are available. In our work, we used the analogy that smaller specimens are stronger against fatigue and can be used as reference. Therefore, fatigue strength of smaller specimens was used as a reference to find the reliability of the larger specimens.

We use the stress-strength analysis in order to estimate the reliability of a specimen with large size by using a data set for small specimen. The reliability is evaluated based on the number of cycles to failure. Stress-Cycle curves (SN curves) are often used to present fatigue resistance of materials at different stress levels. The SN curve simply represents number of cycle to failure at a given stress level. Therefore, there is a one-to-one relationship between stress level and number of cycle to failure [25], [44]. Therefore, using number of cycles instead of stress is a reasonable selection, *i.e.*, the number of cycles to failure can be understood as a measure of resistance to failure. In terms of stress-strength models, such measures are obtained in situations where the system has low resistance to fatigue failure (*i.e.*, larger specimen) as well as in situations where the system has greater resistance to fatigue failure (*i.e.*, smaller specimen).

Indeed, in the context of our work, *Y* refers to the number of cycles to failure in stress situation, *i.e.*, number of cycles to failure of larger specimen. Similarly, *X* refers to the number of cycles to failure in strength situation, *i.e.*, number of cycles to failure of smaller specimen. Therefore, the reliability index $R=P(Y<X)$ refers to the probability of the variable Y being less than variable X. In other words, the index R indicates a measure of the reliability of the larger specimen using data set for the smaller specimen as reference.

As we have mentioned above, q-Exponential distribution can be used to fit stress-strength data when they are represented by cycles to failure obtained from specimens made of the same material but with different sizes. Such an approach for stress and strength analysis has been previously used, for example, in [36] and [92].

For each case study presented in this section, we estimate the parameters of two q-Exponential distributions for data sets representing X and Y - the results will be obtained by PSO and Nelder-Mead algorithm. The estimation method is the maximum likelihood method discussed in section "Maximum likelihood estimators of index $R=P(Y<X)$". Based on the estimates of the parameters, we analyze the goodness-of-fit of the q-Exponential distributions by using both graphical analysis and hypothesis testing. Thus, we show the CDF for the two data sets along with the theoretical CDF of the q-Exponential. Finally, we perform a bootstrapped version of the Kolmogorov-Smirnov test in order to statistically check the fit of the q-Exponential distribution to each data set.

5.1 Case Study 1

From [83] the size effect in ductile cast iron was studied using two sets of fatigue data for specimens with diameters 21 mm (Ø21) and 50 mm (Ø50). During the tests, the specimens were subjected to the same load condition. Figure 5.1 shows details of the drawings of the specimens.

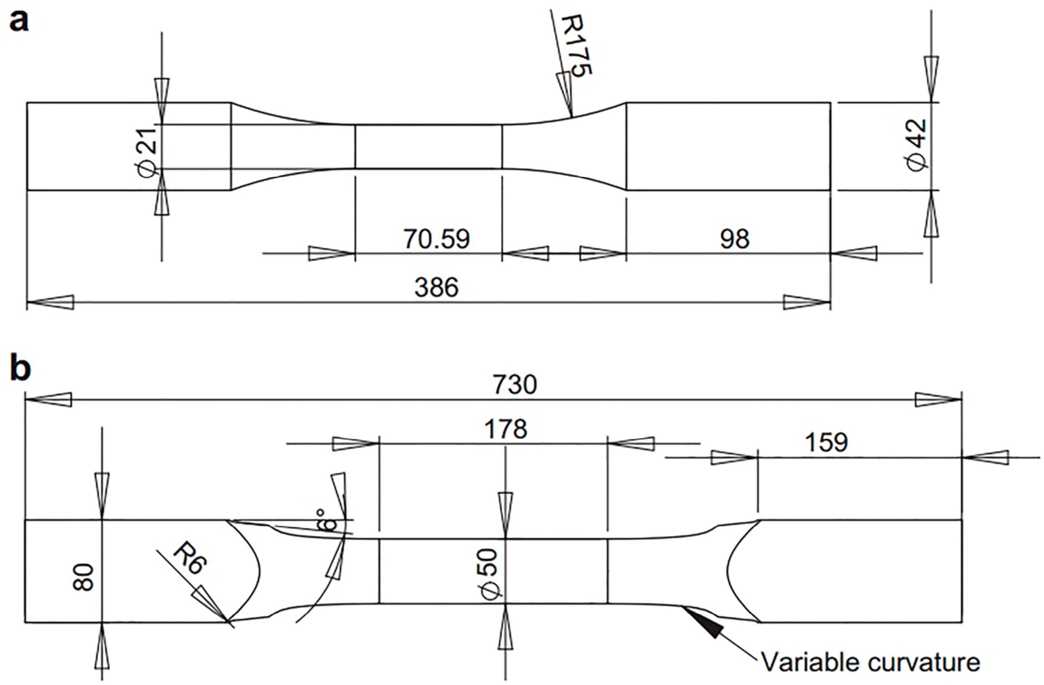


Figure 5.1. Detail drawings of (a) Ø 21 and (b) Ø 50 specimens (all dimensions are in mm).[83]

For stress we consider fatigue data of Ø50 specimens, and for strength the fatigue data of Ø21 specimens are used. The data sets given in terms of number of cycles to failure are presented in Table 5.1 and Table 5.2 for diameters 21 mm and 50 mm, respectively.

Table 5.1. Ø21 specimen fatigue test data (Strength).

Specimen Number	Fatigue Life (number of cycles to failure)
1	3000000
2	716400
3	1674100
4	679400
5	801000
6	1076600
7	4181701
8	619200
9	469500
10	83200
11	92500
12	107700

Table 5.2. Ø50 specimen fatigue test data (Stress).

Specimen Number	Fatigue Life (number of cycles to failure)
1	295000
2	869000
3	869900
4	1573335
5	151400
6	152000
7	183700
8	218000
9	30200
10	45100
11	46900
12	47300

Table 5.3 presents the estimated parameters - entropic indices (shape parameters) and scale parameters, Kolmogorov-Smirnov (K-S) distances between the empirical and fitted distribution functions, and the *p*-values of the K-S Boot test – for this test we use *N*=1000. As mentioned in Section entitled "The *q*-Exponential distribution", the *q*-Exponential distribution shows characteristics of a power law when the entropic index presents values between 1 and 2. In this case study, we can see that both *X* and *Y* present this behavior for the analyzed data sets.

Table 5.3. Estimated parameters, Kolmogorov-Smirnov distances and p-values for the Kolmogorov-Smirnov Test (K-S Boot) – q-Exponential Distribution (Case Study 1).

	Optimization Method	Entropic Index	Scale Parameter	K-S (D_0)	*p*-value
Data set 1 (Strength)	Nelder-Mead	$\hat{r} = 1.1087$	$\hat{\beta} = 884013.7$	0.1477	0.7453
	PSO	$\hat{r} = 1.1082$	$\hat{\beta} = 884816$	0.1478	0.7353
Data set 2 (Stress)	Nelder-Mead	$\hat{q} = 1.3005$	$\hat{\eta} = 161904$	0.1554	0.6054
	PSO	$\hat{q} = 1.3007$	$\hat{\eta} = 161820.6$	0.1554	0.5994

Figure 5.2 (a) and (b) present the theoretical and empirical CDFs for *X* and *Y*, respectively. For the theoretical curve we use the results obtained by Nelder-Mead method, once it is very similar to the results obtained by PSO. In addition, by the Kolmogorov-Smirnov tests and the corresponding *p*-values (K-S Boot) reported in Table 5.3 the *q*-

Exponential model adequately fits both the strength and stress data sets, as can also be seen in the graphs shown in Figure 5.2 (a) and (b).

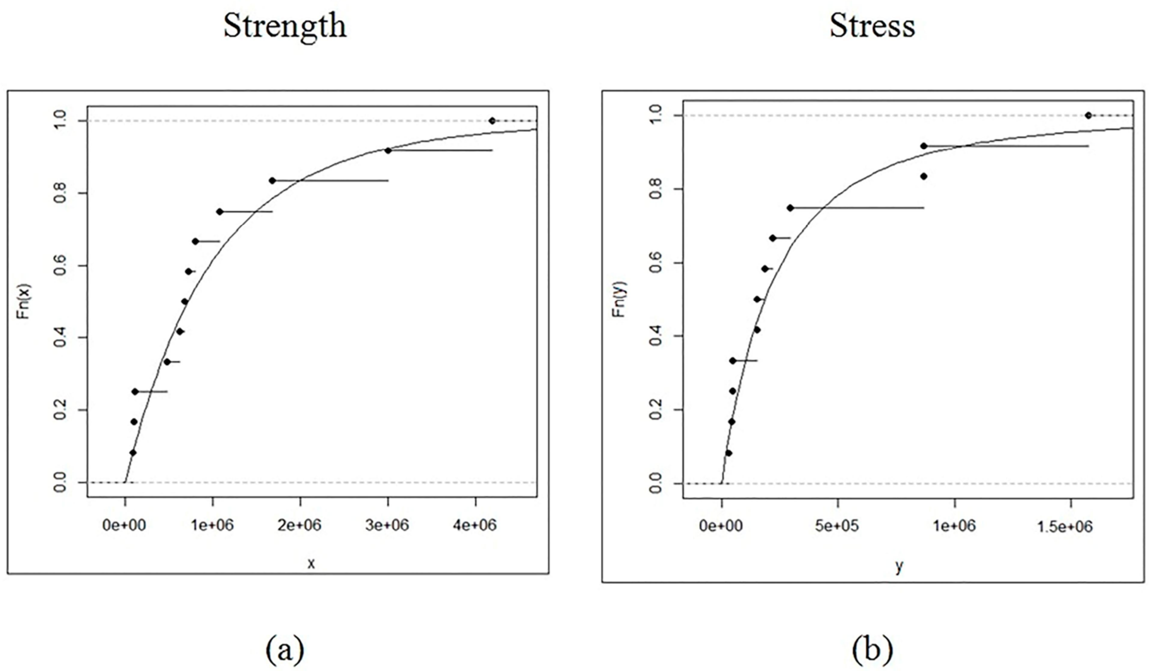


Figure 5.2. Theoretical (q-Exponential) and empirical CDF for data sets of Case Study 1. (a) X – Strength and (b) Y – Stress

Given that $\hat{r} = 1.1087$ $(1 < \hat{r} < 2)$, there is no limitation on the support of *X*, thus the index R is estimated by Equation (4.4) as 0.7579. The value obtained for *R* indicates that within the range of fatigue cycles here considered (*i.e.*, high cycle fatigue) there is a 0.7579 probability that fatigue life of specimens with Ø21 mm diameter is longer than specimens with Ø50 mm diameter. In terms of reliability, we can conclude that the value obtained for index *R* indicates the system performance, *i.e.*, based on the data presented for strength and stress, the reliability of the larger specimen is equal to 0.7579.

Moreover, the confidence intervals (bootstrap-*p* and non-parametric bootstrap approaches) are constructed by using the procedures presented in the section “Bootstrap confidence intervals”. In this case study, we obtained a large width for the confidence interval of *R* parameter due the small size of the sample.

Table 5.4. Point and interval estimates for R = P(Y < X) – Case Study 1(Results by Nelder-Mead and PSO).

Estimate of the Parameter *R=P(Y<X)* by Nelder-Mead	
$\hat{R} = 0.7579$	
Estimate of the Parameter *R=P(Y<X)* by PSO	
$\hat{R} = 0.7580$	
Bootstrap-*p* confidence interval by Nelder-Mead	
n=12, m=12	
C.I (R,0.90) = [0.4001, 0.9148]	C.I (R,0.95) = [0.2869, 0.9395]
Bootstrap-*p* confidence interval by PSO	
n=12, m=12	
C.I (R,0.90) = [0.4249, 0.9236]	C.I (R,0.95) = [0.2963, 0.9453]
Non-Parametric Bootstrap confidence interval by Nelder-Mead	
n=12, m=12	
C.I (R,0.90) = [0.5730, 0.9109]	C.I (R,0.95) = [0.4025, 0.9387]
Non-Parametric Bootstrap confidence interval by PSO	
n=12, m=12	
C.I (R,0.90) = [0.4797, 0.9305]	C.I (R,0.95) = [0.3553, 0.9544]

5.2 Case Study 2

In [84], the size effect on gigacycle fatigue life of high-strength steel was evaluated using the following specimen geometries:

- Type A: Ø 8 mm x 10 mm specimen;
- Type B: Ø 3 mm hourglass-shaped specimens.

The specimens were subjected to the same load condition. Therefore, fatigue data for specimens Type A and Type B are selected as stress and strength, respectively. Figure 5.3 shows the detail drawings of the specimens. Data sets for strength and stress are presented in Table 5.5 and Table 5.6 respectively.

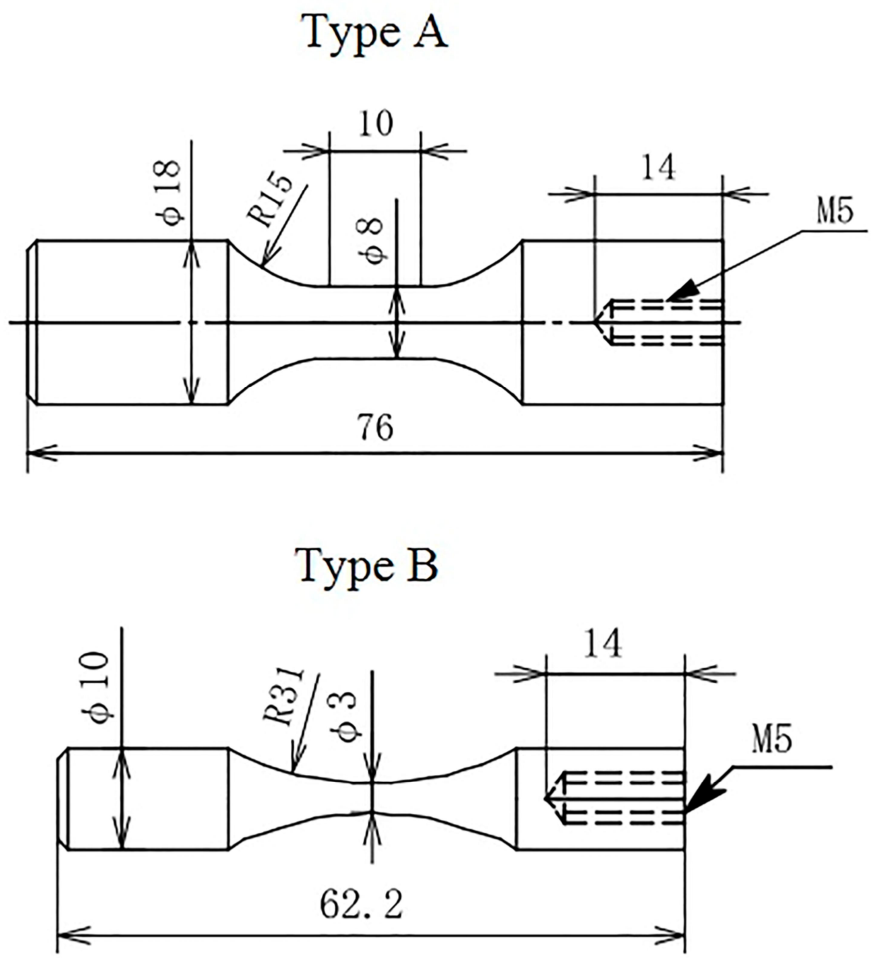


Figure 5.3. Detail drawings of (a) Ø 8 and (b) Ø 3 specimens (all dimensions are in mm). Adapted from [84]

Table 5.5. Type B (Ø 3 mm hourglass-shaped specimen) fatigue test data (Strength).

Specimen Number	Fatigue Life (number of cycles to failure)
1	1017286
2	2989152
3	4059346
4	4256299
5	8376572
6	9560400
7	13007977
8	25303118
9	33621704
10	55951560
11	101155984
12	144322192

13	376711232
14	731957760
15	9444513800
16	9912163300
17	9918688300
18	9921105900

Table 5.6. Type A (Ø 8 x 10 mm specimen) fatigue test data (Stress).

Specimen Number	Fatigue Life (number of cycles to failure)
1	289867
2	1291756
3	6404257
4	7848468
5	9374890
6	31500474
7	211678768
8	5575744500
9	5926607400

In Table 5.7, we present the estimated parameters (entropic indices and scale parameters), the Kolmogorov-Smirnov (K-S) distances between the empirical and fitted distribution functions, and the corresponding *p*-values (K-S Boot – for this test we use *N*=1000). Note that for stress and strength, the entropic indices present values that characterize a power law behavior, *i.e.*, $1<q<2$ for stress and $1<r<2$ for strength.

Table 5.7. Estimated parameters, Kolmogorov-Smirnov distances and p-values for the Kolmogorov-Smirnov Test (K-S Boot) – q-Exponential Distribution (Case Study 2).

		Entropic Index	Scale Parameter	K-S (D_0)	*p*-value
Data set 1 (Strength)	Nelder-Mead	$\hat{r} = 1.7519$	$\hat{\beta} = 4704629$	0.1329	0.4955
	PSO	$\hat{r} = 1.7521$	$\hat{\beta} = 4688696.2$	0.1327	0.4705
Data set 2 (Stress)	Nelder-Mead	$\hat{q} = 1.7643$	$\hat{\eta} = 1450221$	0.1434	0.8501
	PSO	$\hat{q} = 1.7642$	$\hat{\eta} = 1453264.2$	0.1433	0.8551

Figure 5.4 (a) and Figure 5.4 (b) show the theoretical and empirical CDF for *X* and *Y*, respectively. We use the results obtained by Nelder-Mead method, once it is very similar to the results obtained by PSO. For the sake of visualization, we here use logarithmic scale to represent *X* and *Y* because the data sets present many extreme values. We also report in Table 5.7 the bootstrapped Kolmogorov-Smirnov tests and the corresponding p-values. Based on those results, we notice that q-Exponential model adequately fits both *X* and *Y* data sets, as can also be seen in Figure 5.4.

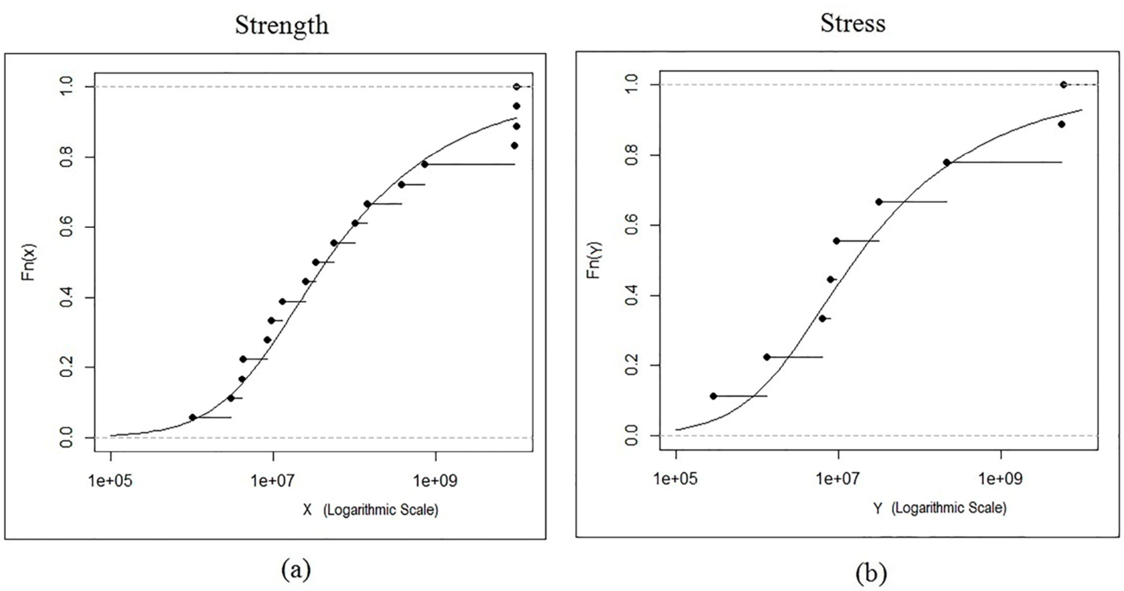


Figure 5.4. Theoretical (q-Exponential) and empirical CDF for data sets of Case Study 2. (a) X – Strength and (b) Y – Stress

Given that $\hat{r} = 1.7519$ $(1 < \hat{r} < 2)$, there is no limitation on the support of *X*, thus index *R* is estimated by Equation (4.6) as 0.5973. Thus, considering a component of strength and another of stress, obtained respectively when we measure the life cycle for specimens with 3 mm diameter and 8 mm diameter, there will be 59.73% chance that the larger specimen will not fail.

The confidence intervals are constructed using the bootstrap-p and non-parametric bootstrap approaches (presented in Section "Bootstrap confidence intervals"). Also in this case study, due the small size of the samples, we obtain large width for the confidence intervals of the parameter *R*.

Table 5.8. Point and Interval estimates for $R = P(Y < X)$ – Case Study 2.

Estimate of the Parameter $R = P(Y < X)$ by Nelder-Mead	
$\hat{R} = 0.5973$	
Estimate of the Parameter $R = P(Y < X)$ by PSO	
$\hat{R} = 0.5973$	
Bootstrap-*p* confidence interval by Nelder-Mead	
n=9, m=18	
C.I (R,0.90) = [0.4002, 0.7922]	C.I (R,0.95) = [0.3442, 0.8236]
Bootstrap-*p* confidence interval by PSO	
n=9, m=18	
C.I (R,0.90) = [0.4006, 0.7918]	C.I (R,0.95) = [0.3498, 0.8303]
Non-Parametric Bootstrap confidence interval by Nelder-Mead	
n=9, m=18	
C.I (R,0.90) = [0.4142, 0.7864]	C.I (R,0.95) = [0.3800, 0.8247]
Non-Parametric Bootstrap confidence interval by PSO	
n=9, m=18	
C.I (R,0.90) = [0.4510, 0.7716]	C.I (R,0.95) = [0.4196, 0.7883]

5.3 Comparing *q*-Exponential with other distributions

For the sake of comparison, both Weibull and Exponential distributions were also considered to model the experimental strength and stress data sets presented in case studies 1 and 2. The results for the estimated parameters (scale and shape parameters for Weibull distribution, and for the Exponential distribution parameter), Kolmogorov-Smirnov (K-S) distances between empirical and fitted distribution functions, and the corresponding p-values (K-S Boot - performed with N=1000) obtained from the data sets are shown in Table 5.9 and Table 5.10, which also include the K-S distance and p-values for the fit of the q-exponential distribution. As we observe in the case studies, the approaches of PSO and Nelder-Mead shows results very similar for the p-values, thus, without loss of generality, we will present the comments of this section based on the results obtained by the Nelder-Mead method.

Table 5.9. Comparing Weibull vs q-Exponential – Case Studies 1 and 2.

		Parameters (Weibull Distribution)		**(K-S Boot) (Weibull Distribution)**		**(K-S Boot) (q-Exponential Distribution)**	
Case Studies		**Shape Parameter**	**Scale Parameter**	**K-S (D_0)**	***p*-value**	**K-S (D_0)**	***p*-value**
Case Study 1	**Data set 1 (Strength)**	0.9331	1088102	0.1409	0.7322	0.1477	0.7453
	Data set 2 (Stress)	0.8336	335326.1	0.164	0.5115	0.1554	0.6054
Case Study 2	**Data set 1 (Strength)**	0.3366	417229706	0.1648	0.2048	0.1329	0.4955
	Data set 2 (Stress)	0.3077	176273348	0.2222	0.2298	0.1434	0.8501

Table 5.10. Comparing Exponential vs q-Exponential – Case Studies 1 and 2.

		Parameter (Exponential Distribution)	**(K-S Boot) (Exponential Distribution)**		**(K-S Boot) (q-Exponential Distribution)**	
Case Studies		**Rate Parameter**	**K-S (D_0)**	***p*-value**	**K-S (D_0)**	***p*-value**
Case Study 1	**Data set 1 (Strength)**	8.89E-07	0.1587	0.7212	0.1477	0.7453
	Data set 2 (Stress)	2.68E-06	0.2245	0.2757	0.1554	0.6054
Case Study 2	**Data set 1 (Strength)**	4.42E-10	0.6048	9.99E-4	0.1329	0.4955
	Data set 2 (Stress)	7.65E-10	0.6429	9.99E-4	0.1434	0.8501

For case study 1, based on the K-S boot, the fit of the Weibull distribution resulted in *p*-values of 0.7322 and 0.5115 for the strength and stress data, respectively, clearly indicating that the Weibull is an appropriate distribution to describe the stress-strength data of this case study. In the case of the Exponential distribution, the *p*-value for the K-S test was equal to 0.7212 for strength and 0.2757 for stress, resulting in a reasonable fit for the experimental data. However, for the strength data, we observed the most significant fit for *q*-Exponential distribution (*p*-value = 0.7453), whereas Weibull distribution is the second (*p*-value = 0.7323), and Exponential distribution also presents a good fit (*p*-value = 0.7212). For the stress, a similar behavior was observed, *i.e.*, *q*-Exponential presents the most significant fit (*p*-

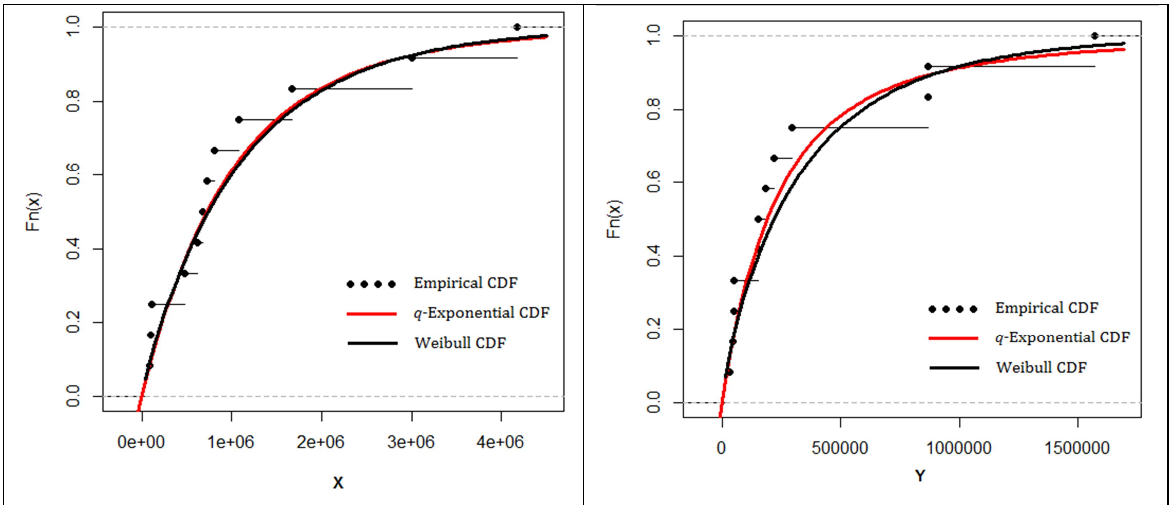
value = 0.6054), while Weibull distribution is the second (p-value = 0.5115) and, among the three distributions considered, the Exponential presented the worst adjustment for the stress (p-value 0.2757).

For case study 2, based on the K-S boot, the fit of Weibull distribution resulted in p-values of 0.2048 and 0.2297 for strength and stress data, respectively, indicating that despite the adjustment be significant, we cannot consider this as an excellent fit. In the case of Exponential distribution, the p-value for the K-S test was equal to 9.99E-04 for strength and the same value for the stress, which yields a non-significant fit for Exponential distribution. Note that q-Exponential distribution presents the most significant fit for both strength (p-value = 0.4955) and stress (p-value = 0.8501), Weibull distribution provides the second most significant fit for strength (p-value = 0.2048) and stress (p-value = 0.2298), whereas Exponential distribution was not significant for both data sets.

Note also that for both case studies, when we consider q-Exponential distribution, a power law behavior is obtained for all analyzed cases, as the entropic indices for all data sets are greater than one. Moreover, when we consider the Weibull distribution, the shape parameter for all cases were between 0 and 1, which indicates a behavior of stretched exponential. As we mentioned in the Introduction, q-Exponential PDF with power law behavior presents a heavier tail than that of a Weibull PDF (with stretched exponential behavior). Thus, it is expected the q-Exponential to have a superior performance over Weibull distribution when dealing with data sets that containing extremely large values. Thus, although the fit by q-Exponential and Weibull distributions were comparable for the first case study, q-Exponential is superior in the second one. This fact is due to the presence of extremely large values in the associated samples (magnitude in the order of 10^9).

The estimation of R when X and Y are Weibull independent variables was presented by Kundu and Gupta[36]. In their work, the authors presented the expression $\hat{R} = \frac{\hat{\theta}_1}{\hat{\theta}_1 + \hat{\theta}_2}$,where $\hat{\theta}_1$ is the estimate of scale parameter for X and $\hat{\theta}_2$ is the estimate of scale parameter for Y. Thus, once dataset for X and Y in case study 1 also presented a good fit for the Weibull distribution, we computed R considering a Weibull distribution as $\hat{R} = 0.7649$. This result is very similar to the one when X and Y are modeled by two independent q-Exponential distributions, *i.e.*, $\hat{R} = 0.7579$. This indicates that both distributions can be used in order to estimate $R = P(Y<X)$ for the first case study. For case study 2, when X and Y are modeled by two independent Weibull distributions, the estimated R index is $\hat{R} = 0.7029$, which is very

different from the one obtained when we considered q-Exponential ($\hat{R} = 0.5973$). This difference is due to the fact that Weibull distribution presented an inferior fit performance for both X and Y when compared to q-Exponential. In fact, both X and Y present extremely large values and, as discussed previously, this kind of data is better modeled by a PDF that has the ability to model data with characteristic of power law as is the case of q-Exponential when $1<q<2$. Observe from Figures Figure ***5.5*** and Figure ***5.6*** that the Case Study 1 presents a good fit for both distributions (q-Exponential and Weibull), this fact is not observed in Case Study 2, once the chart clearly shows that the adjustment by the q-Exponential distribution is better than the Weibull Distribution.



(a) (b)

Figure 5.5. Empirical and Theoretical (q-Exponential and Weibull) CDFs for Case Study 1 - (a) Strength and (b) Stress

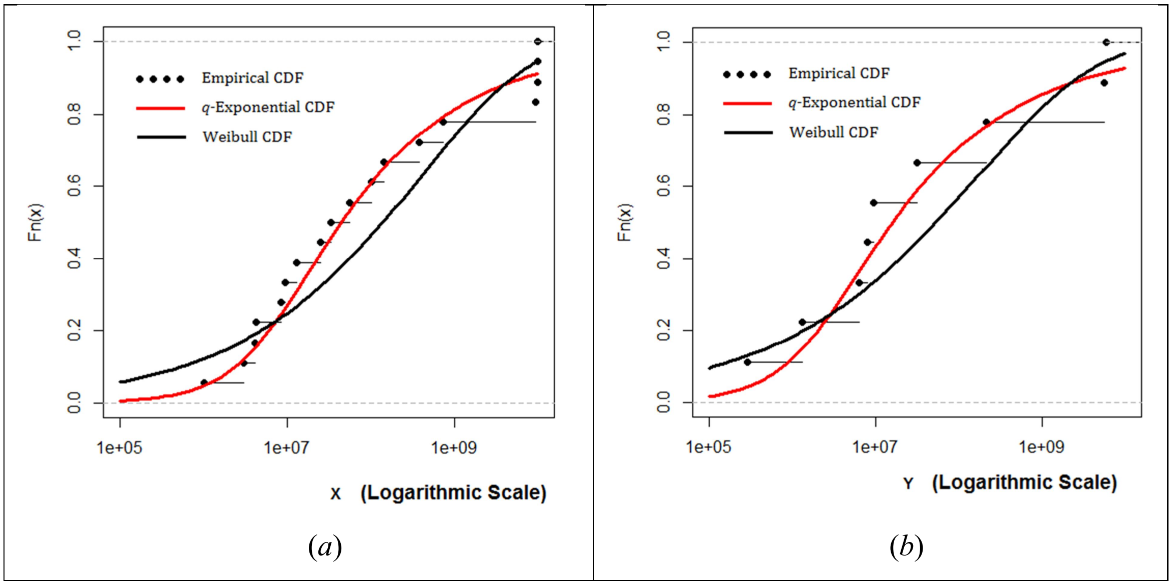


Figure 5.6. Empirical and Theoretical (q-Exponential and Weibull) CDFs for Case Study 2 - (a) Strength and (b) Stress

6 CONCLUSIONS

We have introduced the q-Exponential distribution as a model for reliability data with extremely large values in the relevant context of stress-strength reliability. More specifically, when we deal with fatigue life data that presents a power law behavior and we need to estimate the performance index $R=P(Y<X)$. We have considered that the stress Y and strength X are q-Exponential independent random variables and have proposed a procedure for estimating index R by considering that the support of X is limited (*i.e.*, entropic index or shape parameter of the strength is $r<1$) and unlimited $(1<r<2)$. Additionally, confidence intervals for the index R have been presented by means of parametric and non-parametric bootstrap approaches.

In order to estimate this index R, it is necessary to estimate the parameters of these two q-Exponential distributions (X and Y). Once the analytical expression for these parameters are very complicated to be obtain, we used the PSO and Nelder-Mead algorithms in order to maximize the log-likelihood function of the q-Exponential distribution. Confidence intervals for the parameters were presented by means of asymptotic confidence intervals and parametric and non-parametric bootstrap approaches.

From the simulation experiments made to the parameters of the q-Exponential, we considered two algorithms of optimization (PSO and Nelder-Mead). It was shown that, for the point estimates, the absolute bias and the MSE values related to the estimation of η and q parameters via maximum likelihood decreases as the sample size increases, indicating the consistency of the MLE for the q-Exponential distribution – This fact was observed when we use any of two algorithms. We may note also that, in simulations that we considered a negative value to the parameter q, it was observed a great value for the MSE in both optimization methods. We can highlight that this great values for the MSE (when $q=-1$) is observed mainly when the n is smaller ($n=100$). However, it is important to point out that, also was observed in the case with q negative, when n increases, the bias and MSE decreases substantially. These facts were observed for both parameters q and η.

In the simulations for the interval estimates of q and η parameter ,considering the PSO and Nelder-Mead Algorithms, we observed that the width of the intervals obtained by asymptotic, bootstrap-p and non-parametric bootstrap approach, generally decrease as the n

increase, this is expected due the consistency property of the MLE. Besides, considering the two optimization methods, the most of parameters were covered for the intervals obtained in the simulations. We yet detected that, in most of cases, the results for the intervals obtained for η and q by Nelder-Mead present smaller width if compared with the PSO results. In general, for the PSO and Nelder-Mead algorithm, we notice that for larger sample sizes ($n = 1000$), asymptotic and bootstrap approaches tend to provide similar interval estimates for the q-Exponential parameters.

From the simulation experiments for the index *R*, it has been verified the consistency of the MLE obtained for the Index *R* based on the q-Exponential distribution, once the absolute bias and the MSE values related to the estimation of *R* via maximum likelihood decreases as the sample size increases. From results, we can yet observe that the bias for cases that we have q>0 not indicates a tendency of overestimation or underestimation, once we have bias with positive and negative values. For cases that we have a q negative it is observed a tendency of overestimation of the index *R*, once the bias for these cases present always positive values.

Furthermore, for different sample sizes for *X* and *Y*, the bootstrap-*p* non-parametric bootstrap confidence intervals showed to be very efficient in estimating the confidence interval of *R* given that for all in the most simulations the confidence intervals included the true parameter value.

With respect to the first case study involving ductile cast iron specimens, q-Exponential distribution properly fits both stress and strength data, as can be seen by the CDF and PDF plots and by the bootstrapped K-S test. In addition, for the sake of comparison, we have estimated the parameters for the situations where *X* and *Y* are both modeled by either Weibull or Exponential distributions. The latter provided the worst fit, while q-Exponential and Weibull models resulted in quite similar fits. Such a result was reinforced by the proximity of the estimates for the index *R* obtained from both models.

In relation to the second case study involving high-strength steel, q-Exponential distribution presented an excellent fit for both strength and stress. The Weibull distribution, despite having a significant adjustment to the experimental data, presented smaller *p*-values for both strength and stress. The Exponential distribution in turn was not significant for the data sets of this case study.

Therefore, based on the discussed results, it is natural to consider the q-Exponential as a good distribution to model stress-strength reliability problems, especially when we are dealing with data with great order of magnitude. As already mentioned, q-Exponential distribution is able to model data that present a power law asymptotic behavior, which is an important characteristic of cycles until fatigue failure, as corroborated by the two case studies considered in this work, where the estimated entropic indices had values that characterize a power law behavior, *i.e.*, $1<q<2$ for stress and $1<r<2$ for strength.

Thus, comparing the adjustments considering the q-Exponential distribution with the adjustments considering the Weibull distribution, we observe that both the first example and the second example showed good quality in fitting the data by a q-Exponential distribution. The Weibull distribution, presented significant adjustment in both examples, however, for the second example, the data presented a p-value clearly more significant for the fit by a q-Exponential. Thus, our main conclusion is that data with extremely large values are best fitted by a q-exponential distribution, once this distribution has a heavier tail than the tail of the Weibull distribution.

It was clearly evidenced, by the case studies, that there are situations in which the data are better adjusted when we consider the q-exponential distribution. Thus, the use of q-exponential distribution in Stress-Strength problems, mainly when the data of the problem present extreme large values, provides a better fit of the data and for consequence, is able to return a better estimate for the index R. Thus, despite the expressions for the calculation of the Index R, when we consider the q-Exponential distribution, be more complicated than the expression of Index R when we consider the Weibull distribution, the use of the q-Exponential distribution in order to estimate the index R it is very useful in some situations, and must be considered in order to provide more reliable estimates for the index R.

As limitation of this work we can point out that the results obtained in the estimation of the q-Exponential parameters, when the parameter q is negative, were not very suitable, mainly when the n is small. Nevertheless, the real purpose of this study is to work with data that present q values greater than 1, since when $q > 1$ the PDF of the q-Exponential is capably to characterize the Power Law behavior, that is a characteristic observed in fatigue data.

For future studies, we suggest the development of new Stress-Strength models, based on generalizations of q-Exponential distribution. For example we can quote the q-Weibull distribution that has demonstrated that has the capability of modeling data even more extremes than the data modeled by the q-Exponential distribution. In addition, we should search for new reparametrizations of the log-likelihood of the q-Exponential whe the parameter q is less than 0, in order to verify which of these provide better results for the consistency of the estimates, mainly when the sample size is small.

REFERENCES

[1] A. Der Kiureghian and J. Song, "Multi-scale reliability analysis and updating of complex systems by use of linear programming," *Reliab. Eng. Syst. Saf.*, vol. 93, no. 2, pp. 288–297, 2008.

[2] P. Bak, *How nature works. The science of self-organized criticality*, 1st ed. Copernicus, 1996.

[3] S. Picoli, R. S. Mendes, L. C. Malacarne, and R. P. B. Santos, "q -distributions in complex systems : a brief review," *Brazilian J. Phys.*, vol. 39, no. 2A, pp. 468–474, 2009.

[4] S. Picoli, R. S. Mendes, and L. C. Malacarne, "q-exponential, Weibull, and q-Weibull distributions: an empirical analysis," *Phys. A Stat. Mech. its Appl.*, vol. 324, no. 3–4, pp. 678–688, Jun. 2003.

[5] E. M. Assis, E. P. Borges, and S. A. B. V. De Melo, "Generalized q-Weibull model and the bathtub curve," *Int. J. Qual. Reliab. Manag.*, vol. 30, no. 7, pp. 720–736, 2013.

[6] C. Tsallis, "Possible generalization of Boltzmann–Gibbs statistics," *J. Stat. Phys.*, vol. 52, pp. 479–487, 1988.

[7] D. A. Lavis, "Boltzmann , Gibbs and the Concept of Equilibrium," no. November 2006, pp. 1–14, 2004.

[8] a. M. Reynolds and M. Veneziani, "Rotational dynamics of turbulence and Tsallis statistics," *Phys. Lett. A*, vol. 327, no. 1, pp. 9–14, Jun. 2004.

[9] E. M. C. Abreu, J. A. Neto, and C. F. L. Godinho, "Nonextensive statistics, entropic gravity and gravitational force in a non-integer dimensional space," *Phys. A Stat. Mech. its Appl.*, vol. 411, pp. 118–127, Oct. 2014.

[10] P. C. Ivanov, A. Yuen, and P. Perakakis, "Impact of stock market structure on intertrade time and price dynamics," *PLoS One*, vol. 9, no. 4, 2014.

[11] G. Gu, X. Xiong, W. Zhang, Y. Zhang, and W. Zhou, "Empirical properties of inter-cancellation durations in the Chinese stock market," vol. 2, no. March, pp. 1–12, 2014.

[12] A. D. Anastasiadis, M. P. de Albuquerque, M. P. de Albuquerque, and D. B. Mussi, "Tsallis q-exponential describes the distribution of scientific citations-a new characterization of the impact," *Scientometrics*, vol. 83, no. 1, pp. 205–218, 2010.

[13] M. I. Bogachev, A. R. Kayumov, and A. Bunde, "Universal Internucleotide Statistics in Full Genomes: A Footprint of the DNA Structure and Packaging?," *PLoS One*, vol. 9, no. 12, p. e112534, 2014.

[14] C. Beck, "Generalized statistical mechanics of cosmic rays," *Phys. A Stat. Mech. its Appl.*, vol. 331, no. 1–2, pp. 173–181, Jan. 2004.

[15] F. Vallianatos, G. Michas, and G. Papadakis, "Non-extensive and natural time analysis of seismicity before the Mw6.4, October 12, 2013 earthquake in the South West segment of the Hellenic Arc," *Phys. A Stat. Mech. its Appl.*, vol. 414, pp. 163–173, Nov. 2014.

[16] a. Namaki, Z. Koohi Lai, G. R. Jafari, R. Raei, and R. Tehrani, "Comparing emerging and mature markets during times of crises: A non-extensive statistical approach," *Phys. A Stat. Mech. its Appl.*, vol. 392, no. 14, pp. 3039–3044, Jul. 2013.

[17] W. Li, Q. a. Wang, L. Nivanen, and a. Le Méhauté, "How to fit the degree distribution of the air network?," *Phys. A Stat. Mech. its Appl.*, vol. 368, no. 1, pp. 262–272, Aug. 2006.

[18] F. Vallianatos and D. Triantis, "A non-extensive view of the Pressure Stimulated Current relaxation during repeated abrupt uniaxial load-unload in rock samples," *EPL (Europhysics Lett.*, vol. 104, no. 6, p. 68002, 2013.

[19] B. T. Brady and G. A. Rowell, "Laboratory investigation of the electrodynamics of rock fracture," *Nature*, vol. 321, pp. 488–492, 1986.

[20] F. Vallianatos, D. Triantis, C. Anastasiadis, and I. Stavrakas, "Electric Earthquake Precursors: From Laboratory Results to Field Observations," *Phys. Chem. Earth*, vol. 29, pp. 339–351, 2004.

[21] C. Tsallis, *Introduction to nonextensive statistical mechanics: approaching a complex world*. New York: Springer, 2009.

[22] J. Laherrère and D. Sornette, "Stretched exponential distributions in Nature and Economy: ``Fat tails'' with characteristic scales," *Eur. Phys. J. B*, vol. 2, pp. 525–539, 1998.

[23] a. Garcimartín, a. Guarino, L. Bellon, and S. Ciliberto, "Statistical Properties of Fracture Precursors," *Phys. Rev. Lett.*, vol. 79, no. 17, pp. 3202–3205, 1997.

[24] O. H. Basquin, "The Exponential Law of Endurance Tests," in *Proceedings of the American Society for Testing and Materials*, 1910, pp. 625–630.

[25] B. Farahmand, G. Bockrath, and J. Glassco, *Fatigue and Fracture Mechanics of High Risk Parts*, 1st ed. Springer US, 1997.

[26] P. Bak, C. Tang, and K. Wiesenfeld, "Self-organized criticality: An explanation of the 1/f noise," *Phys. Rev. Lett.*, vol. 59, no. 4, pp. 381–384, 1987.

[27] S. Zapperi, A. Vespignani, and E. H. Stanley, "Plasticity and avalanche behaviour in microfracturing phenomena," *Nature*, vol. 388, no. August, pp. 658–660, 1997.

[28] M. A. Hussian, “Estimation of Stress-Strength Model for Generalized Inverted Exponential Distribution using Ranked Set Sampling,” *Int. J. Adv. Eng. Technol.*, vol. 6, no. 6, pp. 2354–2362, 2014.

[29] F. Downtown, “The estimation of P(Y<X) in the normal case,” *Technometrics*, vol. 15, pp. 551–558, 1973.

[30] Z. Govidarajulu, “Tow sided confidence limits for P(X>Y) based on normal samples of X and Y,” *Sankhya*, vol. B 29, pp. 35–40, 1967.

[31] W. A. Woodward and G. D. Kelley, “Minimum variance unbiased estimation of P(Y<X) in the normal case,” *Technometrics*, vol. 19, pp. 95–98, 1977.

[32] D. B. Owen, K. J. Craswell, and D. L. Hanson, “Non-parametric upper confidence bounds for P(Y<X) and confidence limits for P(Y<X) when X and Y are normal,” *J. Am. Stat. Assoc.*, vol. 59, pp. 906–924, 1977.

[33] H. Tong, “A note on the estimation of P(Y<X) in the exponential case,” *Technometrics*, vol. 16, p. 625, 1974.

[34] H. Tong, “On the estimation of P(Y<X) for exponential families,” *IEEE Trans. Reliab.*, vol. 26, pp. 54–56, 1977.

[35] J. I. Mccool, “Inference on P (Y < X) in the weibull case,” *Commun. Stat. Comput.*, vol. 20, no. 1, pp. 129–148, 1991.

[36] D. Kundu and R. D. Gupta, “Estimation of P [Y < X] for Weibull Distribution,” *IEEE Trans. Reliab.*, vol. 55, no. 2, pp. 270–280, 2006.

[37] D. Kundu and M. Z. Raqab, “Estimation of R = P(Y < X) for three-parameter Weibull distribution,” *Stat. Probab. Lett.*, vol. 79, no. 17, pp. 1839–1846, Sep. 2009.

[38] S. Rezaei, R. Tahmasbi, and M. Mahmoodi, “Estimation of P[Y<X] for Generalized Pareto Distribution,” *J. Stat. Plan. Inference*, vol. 140, no. 2, pp. 480–494, Feb. 2010.

[39] P. Fathipour, A. Abolhasani, and H. J. Khamnei, “Estimating R = P (Y < X) in the Generalized Rayleigh Distribution with Different Scale Parameters,” *Appl. Math. Sci.*, vol. 7, no. 2, pp. 87–92, 2013.

[40] D. Kundu and R. D. Gupta, “Estimation of P[Y < X] for generalized exponential distribution,” *Metrika*, vol. 61, no. 3, pp. 291–308, Jun. 2005.

[41] M. Z. Raqab, M. . Madi, and D. Kundu, “Estimation of P (Y < X) for the three-parameter generalized exponential distribution,” *Commun. Stat. Methods*, vol. 37, pp. 2854 – 2864, 2008.

[42] B. Al-zahrani and M. Al-harbi, “On Parameters Estimation of Lomax Distribution under General Progressive Censoring,” *Parameters Estim. Lomax Distrib. under Gen. Progress. Censoring*, vol. 2013, pp. 1–07, 2013.

[43] H. Panahi and S. Asadi, “Inference of Stress-Strength Model for a Lomax Distribution,” *World Acad. Sci. Eng. Technol.*, vol. 5, pp. 215–218, 2011.

[44] J. Alten F. Grandt, *Fundamentals of Structural Integrity: Damage Tolerant Design and Nondestructive Evaluation*. 2003.

[45] C. R. Shalizi, “Maximum Likelihood Estimation for q -Exponential (Tsallis) Distributions,” 2007. [Online]. Available: arXiv:math/0701854v2 [math.ST].

[46] J.-F. Bercher and C. Vignat, “A new look at q-exponential distributions via excess statistics,” *Phys. A Stat. Mech. its Appl.*, vol. 387, no. 22, pp. 5422–5432, Sep. 2008.

[47] J. A. Nelder and R. Mead, “A Simplex Method for Function Minimization,” *Comput. J.*, vol. 7, no. (4), pp. 308–313, 1965.

[48] D. Bratton and J. Kennedy, “Defining a standard for particle swarm optimization,” in *Proceedings of the IEEE Swarm Intelligence Symposium (SIS 2007)*, 2007.

[49] S. Nadarajah and S. Kotz, “On the q-type distributions,” *Phys. A Stat. Mech. its Appl.*, vol. 377, no. 2, pp. 465–468, Apr. 2007.

[50] S. Nadarajah and S. Kotz, “q exponential is a Burr distribution,” *Phys. Lett. A*, vol. 359, no. 6, pp. 577–579, Dec. 2006.

[51] H. Bolfarine and M. Carneiro Sandoval, *Introdução à Inferência Estatística*. Rio de Janeiro: Coleção Matemática Aplicada - SBM, 2001.

[52] B. Efron, *An introduction to the bootstrap*. New York: Chapman & Hall/CRC, 1993.

[53] B. Efron, “The jackknife, the bootstrap and other resampling plans,” in *CBMS-NSF Regional Conference Series in Applied Mathematics*, 1982.

[54] B. Y. B. Efron, D. V. Hinkley, and D. V Htnkley, “Assessing the accuracy of the maximum likelihood estimator: Observed versus expected Fisher information,” *Biometrika*, vol. 65, no. 3, pp. 457–483, 1978.

[55] R. Davidson and J. G. MacKinnon, *Econometric theory ad methods*. Oxford University Press, 2003.

[56] S.-W. Lin, K.-C. Ying, S.-C. Chen, and Z.-J. Lee, “Particle swarm optimization for parameter determination and feature selection of support vector machines,” *Expert Syst. Appl.*, vol. 35, no. 4, pp. 1817–1824, 2008.

[57] S.-W. Fei, M.-J. Wang, Y.-B. Miao, J. Tu, and C. Liu, “Particle swarm optimization-based support vector machine for forecasting dissolved gases content in power transformer oil,” *Energy Convers. Manag.*, vol. 50, no. 6, pp. 1604–1609, 2009.

[58] W.-C. Hong, “Chaotic particle swarm optimization algorithm in a support vector regression electric load forecasting model,” *Energy Convers. Manag.*, vol. 50, no. 1, pp. 105–117, 2009.

[59] B. Samanta and C. Nataraj, “Use of particle swarm optimization for machinery fault detection,” *Eng. Appl. Artif. Intell.*, vol. 22, pp. 308–316, 2009.

[60] I. D. Lins, M. das C. Moura, and E. L. Droguett, *Support vector machines and particle swarm optimization: applications to reliability prediction.* Saarbrücken, Germany: Lambert Academic Publishing, 2010.

[61] I. D. Lins, M. das C. Moura, E. Zio, and E. L. Droguett, “A particle swarm-optimized support vector machine for reliability prediction,” *Qual. Reliab. Eng. Int.*, vol. 28, pp. 141–158, 2012.

[62] E. L. Droguett, I. D. Lins, M. C. Moura, E. Zio, and C. M. Jacinto, “Variable selection and uncertainty analysis of scale growth rate under pre-salt oil wells conditions using support vector regression,” *Proc. Inst. Mech. Eng. Part O J. Risk Reliab.*, May 2014.

[63] I. D. Lins, M. das C. Moura, E. A. L. Droguett, C. M. Jacinto, and E. Zio, “Failure prediction of oil wells by support vector regression with variable selection, hyperparameter tuning and uncertainty analysis,” *Chem. Eng. Trans.*, vol. 33, pp. 817–822, 2013.

[64] P. R. A. Firmino, M. das C. Moura, A. P. Oliveira, I. D. Lins, and E. L. Droguett, “Política ótima de manutenção preventiva de sistemas reparáveis baseada em confiabilidade,” in *Proceedings of the XX Congresso Pan-Americano de Engenharia Naval, Transporte Marítimo e Engenharia Portuária*, 2007.

[65] M. Schwaab, E. C. Biscaia, Jr., J. L. Monteiro, and J. C. Pinto, “Nonlinear parameter estimation through particle swarm optimization,” *Chem. Eng. Sci.*, vol. 63, no. 6, pp. 1542–1552, 2008.

[66] F.-K. Wang and P.-R. Huang, “Implementing particle swarm optimization algorithm to estimate the mixture of two Weibull parameters with censored data,” *J. Stat. Comput. Simul.*, vol. 84, no. 9, pp. 1975–1989, 2014.

[67] J. Kennedy, R. Eberhart, and Y. Shi, *Swarm intelligence*. San Francisco, USA: Morgan Kaufmann, 2001.

[68] F. Taubert, F. Hartig, H. J. Dobner, and A. Huth, “On the Challenge of Fitting Tree Size Distributions in Ecology,” *PLoS One*, vol. 8, no. 2, 2013.

[69] X. Chen, M. Foy, M. Kimmel, and O. Y. Gorlova, "Modeling the Natural History and Detection of Lung Cancer Based on Smoking Behavior," *PLoS One*, vol. 9, no. 4, p. e93430, 2014.

[70] H. C. Moore, P. Jacoby, A. B. Hogan, C. C. Blyth, and G. N. Mercer, "Modelling the seasonal epidemics of respiratory syncytial virus in young children," *PLoS One*, vol. 9, no. 6, pp. 1–8, 2014.

[71] A. Rajan and T. Malakar, "Optimal reactive power dispatch using hybrid Nelder–Mead simplex based firefly algorithm," *Int. J. Electr. Power Energy Syst.*, vol. 66, pp. 9–24, 2015.

[72] R.-A. Hooshmand, M. J. Morshed, and M. Parastegari, "Congestion management by determining optimal location of series FACTS devices using hybrid bacterial foraging and Nelder–Mead algorithm," *Appl. Soft Comput.*, vol. 28, pp. 57–68, 2015.

[73] S. Abedi, F. Farhadi, and R. B. Boozarjomehry, "Integration of CFD and Nelder-Mead algorithm for optimization of MOCVD process in an atmospheric pressure vertical rotating disk reactor," *Int. Commun. Heat Mass Transf.*, vol. 43, pp. 138–145, 2013.

[74] A. Maria, S. Gonçalves, A. Maria, and S. Gonçalves, "O Problema de Min-Max-Min com restrições pelo Método de Nelder-Mead," UFRJ, 2013.

[75] S. S. Shapiro and M. B. Wilk, "An analysis of variance test for normality (complete samples)," *Biometrika*, vol. 52, no. 3–4, pp. 591–611, 1965.

[76] M. F. Triola, *Introdução À Estatística*, 10th ed. LTC, 2008.

[77] W. J. Conover, *Practical Nonparametric Statistics*, 3rd ed. Wiley, 1999.

[78] G. C. Blain, "Revisiting the critical values of the Lilliefors test : towards the correct agrometeorological use of the Kolmogorov-Smirnov framework," *Agrometeorogy*, vol. 73, no. 2, pp. 192–202, 2014.

[79] W. Stute, W. G. Manteiga, and M. P. Quindimil, "Bootstrap Based Goodness-Of-Fit-Tests," *Metrika*, vol. 40, pp. 243–256, 1993.

[80] R. Castro, "Lectures 2 and 3 - Goodness-of-Fit (GoF) Tests," pp. 1–18, 2013.

[81] The MathWorks Inc ., "MATLAB." The MathWorks, Inc., Natick, Massachusetts, 2012.

[82] T. R-Project, "R: A language and environment for statistical computing. R Foundation for Statistical Computing, Vienna, Austria." 2014.

[83] M. Shirani and G. Härkegård, "Fatigue life distribution and size effect in ductile cast iron for wind turbine components," *Eng. Fail. Anal.*, vol. 18, no. 1, pp. 12–24, Jan. 2011.

[84] Y. Furuya, “Specimen size effects on gigacycle fatigue properties of high-strength steel under ultrasonic fatigue testing,” *Scr. Mater.*, vol. 58, no. 11, pp. 1014–1017, 2008.

[85] Y. Goegebeur and A. Guillou, “Goodness-of-fit testing for Weibull-type behavior,” *J. Stat. Plan. Inference*, vol. 140, no. 6, pp. 1417–1436, Jun. 2010.

[86] F. N. Nwobi and C. A. Ugomma, “A Comparison of Methods for the Estimation of Weibull Distribution Parameters,” *Metod. Zv. - Adv. Methodol. Stat.*, vol. 11, no. 1, pp. 65–78, 2014.

[87] I. Wolfram Research, “Mathematica 9.0.” Wolfram Research, Inc., Champaign, Illinois, 2012.

[88] D. K. Nagar and D. B. Valencia, “Product and Quotient of Independent Gauss Hypergeometric Variables,” *Ing. y Cienc.*, vol. 7, no. 14, pp. 29–48, 2011.

[89] A. C. Davison and D. V Hinkley, *Bootstrap methods and their application*. Cambridge: Cambridge University Press, 1997.

[90] S. Amiri, D. Von Rosen, and S. Zwanzig, “On the comparison of parametric and nonparametric bootstrap,” 2008.

[91] A. M. Mood, F. A. Graybill, and D. C. Boes, *Introduction to the Theory of Statistics*, 3rd ed. McGraw Hill, 1974.

[92] F. Shahsanaei, A. Daneshkhah, M. Sciences, and A. Sciences, “Estimation of Stress-Strength model in the Generalized Linear Failure Rate Distribution,” pp. 1–31, 2013.