Revista Colombiana de Estadística
ISSN: 0120-1751
revcoles_fcbog@unal.edu.co
Universidad Nacional de Colombia
Colombia

Castro-Kuriss, Claudia; Leiva, Víctor; Athayde, Emilia
Graphical Tools to Assess Goodness-of-Fit in Non-Location-Scale Distributions
Revista Colombiana de Estadística, vol. 37, núm. 2, diciembre, 2014, pp. 341-365
Universidad Nacional de Colombia
Bogotá, Colombia

Revista Colombiana de Estadística
Current Topics in Statistical Graphics
Diciembre 2014, volumen 37, no. 2, pp. 341 a 365
DOI: http://dx.doi.org/10.15446/rce.v37n2spe.47941

Graphical Tools to Assess Goodness-of-Fit in Non-Location-Scale Distributions

Herramientas gráficas para evaluar bondad de ajuste en distribuciones de no-localización-escala

Claudia Castro-Kuriss[1,a], Víctor Leiva[2,3,b], Emilia Athayde[4,c]

[1]Departamento de Matemática, Instituto Tecnológico de Buenos Aires, Argentina
[2]Facultad de Ingeniería y Ciencias, Universidad Adolfo Ibáñez, Chile
[3]Instituto de Estadística, Universidad de Valparaíso, Chile
[4]Centro de Matemática, Universidade do Minho, Portugal

Abstract

Goodness-of-fit (GOF) techniques are used for assessment whether a distribution is suitable to describe a data set or not. These techniques have been studied for distributions belonging to the location-scale family. However, one could be interested in making this assessment for distributions that do not belong to this family. We review the available GOF tests and propose graphical tools based on these tests for censored and uncensored data from non-location-scale distributions. Anderson-Darling, Cramér-von Mises, Kolmogorov-Smirnov, Kuiper, Michael and Watson GOF statistics are considered. We apply the proposed results to real-world data sets to illustrate their potential, with emphasis on some Birnbaum-Saunders distributions.

***Key words*:** Censored Data, Confidence Band, Data Analysis, Probability Plots.

Resumen

Las técnicas de bondad de ajuste se usan para establecer si una distribución es apropiada o no para describir un conjunto de datos. Estas técnicas han sido estudiadas para distribuciones pertenecientes a la familia de locación y escala. Sin embargo, podríamos también estar interesados en establecer si una distribución que no pertenece a esta familia brinda un buen ajuste a los datos. Revisamos los tests de bondad de ajuste disponibles y proponemos herramientas gráficas basadas en estos tests para datos completos

[a]Associate Professor. E-mail: ccastro@itba.edu.ar
[b]Professor. E-mail: victorleivasanchez@gmail.com
[c]Associate Professor. E-mail: mefqa@math.uminho.pt

y censurados desde distribuciones de no locación y escala. Consideramos los estadísticos de Anderson-Darling, Cramér-von Mises, Kolmogorov-Smirnov, Kuiper, Michael y Watson. Aplicamos los resultados propuestos a conjuntos de datos reales para ilustrar su potencial, particularmente en algunas distribuciones Birnbaum-Saunders.

***Palabras clave*:** análisis de datos, bandas de confianza, datos censurados, gráficos de probabilidad.

1. Introduction

Life distributions are widely studied and applied; see Lawless (2003) and Marshall & Olkin (2007). Among the different life distributions, we can mention the classical (BS), Student-t (BS-t) and truncated (TBS) Birnbaum-Saunders, gamma (GA), inverse Gaussian (IG), lognormal (LN) and Weibull (WE) models; see Johnson, Kotz & Balakrishnan (1994), Balakrishnan, Leiva, Sanhueza & Cabrera (2009), Balakrishnan, Leiva, Sanhueza & Vilca (2009), Balakrishnan, Gupta, Kundu, Leiva & Sanhueza (2011), Barros, Paula & Leiva (2009), Ahmed, Castro-Kuriss, Leiva, Flores & Sanhueza (2010), Kotz, Leiva & Sanhueza (2010), Vilca, Sanhueza, Leiva & Christakos (2010), Leiva, Sanhueza, Kotz & Araneda (2010), Leiva, Athayde, Azevedo & Marchant (2011), Leiva, Soto, Cabrera & Cabrera (2011), Sanhueza, Leiva & Lopez-Kleine (2011), Athayde, Azevedo, Leiva & Sanhueza (2012), Azevedo, Leiva, Athayde & Balakrishnan (2012), Ferreira, Gomes & Leiva (2012), Fierro, Leiva, Ruggeri & Sanhueza (2013), Leiva, Marchant, Saulo, Aslam & Rojas (2014), Leiva, Santos-Neto, Cysneiros & Barros (2014), Leiva, Saulo, Leao & Marchant (2014) and Rojas, Leiva, Wanke & Marchant (2015). It is important to stress that, in the case of the BS-t distribution, we obtain a robust estimation procedure for its parameters; see Paula, Leiva, Barros & Liu (2012). All of these distributions do not belong to the location-scale (LS) family, thus we refer to them as non-location-scale (NLS) distributions.

Several efforts have been made to develop goodness-of-fit (GOF) techniques allowing us to address the problem of fitting the mentioned distributions to different types of data. In general terms, GOF tests permit us to assess whether the distribution under a null hypothesis (H_0) is adequate to describe a data set or not. For this hypothesis, there are two options: (i) the distribution can be completely specified (known parameters) or (ii) some (or all) of its parameters are unknown, in which case they need to be estimated with proper methods, such as maximum likelihood (ML). Depending on the distribution under H_0, ML estimates of its parameters cannot be easily calculated and iterative numerical procedures must be used, but problems of convergence may arise, which have not yet been completely studied; see Castillo & Puig (1997) for more details in this regard.

According to D'Agostino & Stephens (1986, pp. xi-xiii), GOF methods can be of graphical type, of chi-squared type or based on the empirical cumulative distribution function (ECDF), on correlation/regression, on transformations or on moments. Most test statistics used for assessing GOF, such as Anderson-Darling (AD), Cramér-von Mises (CM), Kolmogorov-Smirnov (KS), Kuiper (KU) and Watson (WA), compare the ECDF and the hypothesized theoretical cumulative distribution function (CDF) assumed for the data. KU and WA statistics are modifications of the KS and CM statistics, respectively, in particular for circular data given its invariant property under change of origin in a circle. For more details about AD, CM, KS, KU and WA statistics, see D'Agostino & Stephens (1986, Ch. 4). Some graphical counterparts of the mentioned goodness-of-fit tests are considered by Chen & Balakrishnan (1995), Choulakian & Stephens (2001), Hui, Gel & Gastwirth (2008) and Stehlík, Strelec & Thulin (2014).

A graph allowing us to relate the ECDF with a specified theoretical CDF is the probability versus probability (PP) plot. Analogously, ordered observations corresponding to empirical quantiles can be plotted versus the theoretical quantiles of a specified distribution in a graph known as the QQ plot; see Marden (2004). This author studied such a plot in the case of the normal distribution and provided tools for graphical comparison of two samples, extending also the QQ plot to multivariate data. A disadvantage of the PP plot associated with the KS test is that some points in this graph can be more variable than others. Michael (1983) proposed a modification for the KS test based on the arcsin transformation to stabilize the variance of the points in the PP plot. The graph related to this variance stabilizing transformation is known as the stabilized probability (SP) plot and the statistic associated with the test proposed by Michael (1983) is denoted as MI. This author studied the power of the MI test showing that it proves more powerful than the KS test for certain alternative hypotheses. Stehlík et al. (2014) compared the power and robustness of several classic tests for normality, such as the AD, Jarque-Bera, KU and Shapiro-Wilk, for different distributions in the alternative hypothesis, including the Cauchy, Laplace and Student-t with different degrees of freedom (DF) models. These authors concluded that in general no test can be recommended as the best for all the alternatives considered.

In reliability and survival analyses, it is frequent to find situations where not all individuals or instruments tested complete the event under study which, without the loss of generality, may be referred to as a "failure". Samples involving such situations are named censored data; for literature on this topic see, e.g., Cohen (1991) and Lawless (2003).

When a parametric statistical analysis with censored data needs to validate its distributional assumption, classical GOF test statistics need to be adapted to consider the censorship following two options. The first consists of using GOF tests for uncensored data adapting the type-II right censored data to become an uncensored (complete) data sample, whereas the second option adapts test statistics to type-II right censored data; see, e.g., Malmquist (1950), D'Agostino & Stephens (1986, Chs. 4 and 11), Lin, Huang & Balakrishnan (2008) and Barros, Leiva, Ospina & Tsuyuguchi (2014).

Castro-Kuriss, Kelmansky, Leiva & Martinez (2009) and Castro-Kuriss, Kelmansky, Leiva & Martinez (2010) proposed GOF tests for the LN and normal distributions with type-II right censored data. Castro-Kuriss (2011) studied GOF tests for LS distributions with type-II right censored data and unknown parameters. Other works on the topic based on different types of censoring can be found in Barr & Davidson (1973), Dufour & Maag (1978), O'Reilly & Stephens (1988), Balakrishnan, Ng & Kannan (2004), Lin et al. (2008), Rad, Yousefzadeh & Balakrishnan (2011) and Pakyari & Balakrishnan (2012).

The objectives of this article are to review the available GOF tests and propose graphical tools for assessing GOF in NLS distributions based on these tests with uncensored and censored data. These tools can be used for any distribution in the NLS family, as long as their parameters are known or properly estimated. With the provided tools, it is possible to decide what distribution best fits the data.

This article is structured as follows. In Section 2, we present some well-known life distributions and the estimation of their parameters with packages of the `R` software; see `www.r-project.org`. In Section 3, we introduce GOF tests for NLS distributions with uncensored and censored data. In Section 4, we propose graphical tools based on these GOF tests. In Section 5, we illustrate the graphical tools provided in this paper analyzing real-world data sets from different fields, with emphasis on some Birnbaum-Saunders distributions. We also analyze some empirical robustness aspects with one of these data sets, following the ideas of Hui et al. (2008) and Stehlík et al. (2014). Finally, in Section 6, we sketch some conclusions and mention future research on this topic.

2. Life Distributions

In this section, we present some well-known life distributions, with $\alpha > 0$ and $\beta > 0$ denoting shape and scale parameters, respectively, and μ denoting the mean of the distribution.

2.1. BS and Truncated BS Distributions

A random variable (RV) T with BS distribution of shape $\alpha > 0$ and scale $\beta > 0$ parameters is denoted by $T \sim \text{BS}(\alpha, \beta)$, where "$\sim$" means "distributed as". In this case, the CDF of T is

$$F(t; \alpha, \beta) = \Phi\left([1/\alpha]\xi(t/\beta)\right), \quad t > 0,$$

where $\xi(y) = \sqrt{y} - 1/\sqrt{y} = 2\sinh(\log(\sqrt{y}))$ and $\Phi(\cdot)$ is the N(0, 1) CDF. The corresponding quantile function (QF) is $F^{-1}(q; \alpha, \beta) = \beta[\alpha z(q)/2 + \sqrt{(\alpha z(q)/2)^2 + 1}]^2$, for $0 < q < 1$, where $z(\cdot)$ is the N(0, 1) QF and $F^{-1}(\cdot)$ is the inverse CDF. Note that $F^{-1}(0.5; \alpha, \beta) = \beta$, that is, β is also the median or 50th percentile of the BS distribution. If $T \sim \text{BS}(\alpha, \beta)$, then $X \sim \text{TBS}_\kappa(\alpha, \beta)$ denotes the truncated version at κ of T and its CDF is

$$F(t; \alpha, \beta, \kappa) = \frac{\Phi\left([1/\alpha]\xi\left(t/\beta\right)\right) - \Phi\left([1/\alpha]\xi\left(\kappa/\beta\right)\right)}{\Phi\left(-[1/\alpha]\xi\left(\kappa/\beta\right)\right)}, \quad t \geq \kappa > 0.$$

The corresponding QF is $F^{-1}(q; \alpha, \beta, \kappa) = \beta[\alpha z_\eta(q)/2 + \sqrt{(\alpha z_\eta(q)/2)^2 + 1}]^2$, for $0 < q < 1$, where $z_\eta(\cdot)$ is the QF of the $\text{N}(0, 1)$ distribution truncated (TN) at $\eta = [1/\alpha]\xi(\kappa/\beta)$; for details about the TN and TBS distributions, see Cohen (1991) and Ahmed et al. (2010), respectively.

2.2. The BS-t Distribution

A RV T following the BS-t distribution with shape $\alpha > 0, \nu > 0$ and scale $\beta > 0$ parameters is denoted by $T \sim \text{BS-}t(\alpha, \beta, \nu)$. In this case, the CDF of T is

$$F(t; \alpha, \beta, \nu) = \frac{1}{2}\Big[1 + I_{\frac{\xi(t/\beta)/\alpha}{\xi(t/\beta)/\alpha+\nu}}(1/2, \nu/2)\Big], \quad t > 0,$$

where $I_x(\cdot, \cdot)$ is the incomplete beta ratio. The corresponding QF is $F^{-1}(q; \alpha, \beta, \nu) = \beta[\alpha z_t(q)/2 + \sqrt{(\alpha z_t(q)/2)^2 + 1}]^2$, for $0 < q < 1$, where $z_t(\cdot)$ is the QF of the Student-t distribution with ν DFs. Note that once again β is the median of the BS-t distribution. For details on the BS-t distribution, see Azevedo et al. (2012).

2.3. The Gamma Distribution

A RV T following the GA distribution with shape $\alpha > 0$ and scale $\beta > 0$ parameters is denoted by $T \sim \text{GA}(\alpha, \beta)$. In this case, the CDF of T is

$$F(t; \alpha, \beta) = \frac{\Gamma(1/\alpha^2,\ t/[\alpha^2\beta])}{\Gamma(1/\alpha^2)}, \quad t > 0,$$

where $\Gamma(\cdot)$ and $\Gamma(\cdot, \cdot)$ denote the usual and incomplete gamma functions, respectively. Now, the QF given by $F^{-1}(q; \alpha, \beta)$, for $0 < q < 1$, must be obtained by means of an iterative method.

2.4. The Inverse Gaussian Distribution

A RV T following the IG distribution with mean $\mu > 0$ and scale $\beta > 0$ parameters is denoted by $T \sim \text{IG}(\mu, \beta)$. In this case, the CDF of T is

$$F(t; \mu, \beta) = \Phi\big(\sqrt{\beta/\mu}\,\xi(t/\mu)\big) + \Phi\big(\sqrt{\beta/\mu}\big[\sqrt{t/\mu} + \sqrt{\mu/t}\big]\big)\exp(2\beta/\mu), \quad t > 0.$$

Again the corresponding QF, expressed as $F^{-1}(q; \mu, \beta) = t(q)$, must be obtained with a numerical method.

2.5. The Lognormal Distribution

If $X = \log(T)$ has a normal distribution with mean μ and variance α^2, that is, $X = \log(T) \sim \text{N}(\mu, \alpha^2)$, then the RV T follows the LN distribution with shape $\alpha > 0$ and scale $\beta = \exp(\mu) > 0$ parameters, respectively, which is denoted by $T \sim \text{LN}(\alpha, \beta)$. In this case, the CDF of T is

$$F(t; \alpha, \beta) = \Phi\left([\log(t) - \log(\beta)]/\alpha\right), \quad t > 0.$$

The corresponding QF is $F^{-1}(q; \alpha, \beta) = \beta\exp(z(q)\alpha)$, for $0 < q < 1$, where $z(\cdot)$ is the N(0, 1) QF.

2.6. The Weibull Distribution

A RV T following the WE distribution with shape $\alpha > 0$ and scale $\beta > 0$ parameters is denoted by $T \sim \mathrm{WE}(\alpha, \beta)$. In this case, the CDF of T is

$$F(t; \alpha, \beta) = 1 - \exp(-[t/\beta]^{\alpha}), \quad t > 0.$$

The corresponding QF is $F^{-1}(q; \alpha, \beta) = \beta[-\log(1-q)]^{1/\alpha}$, for $0 < q < 1$.

2.7. Estimation of Parameters

Using R software (i) exploratory data analysis (EDA) can be conducted for diagnosing statistical features present in the data to be analyzed; and (ii) estimation of the parameters of BS, BS-t, GA, IG, LN, TBS and WE distributions can be carried out by the ML method. Next, we describe the R functions (commands) of gbs, ig and basic packages and briefly illustrate their use. The tbs package is available upon request from the authors. These packages incorporate CDF, probability density function (PDF) and QF of the TBS distribution, as well as a random number generator, the moments and the ML estimation of its parameters for censored and uncensored data.

The R software can be downloaded from `CRAN.r-project.org` and installed as any other software. It may be used in a simple interactive form with the R commander by installing the Rcmdr package. The gbs, ig and tbs packages must be also installed. Data analyses based on (i) BS and BS-t distributions can be carried with the gbs package, (ii) IG distribution with the ig package and (iii) GA, LN and WE distributions with the basic or fitdistrplus packages. As an example, the gbs package and a data set, x say, must be loaded as

```
> library(gbs)
> data(x)
```

The data can also be directly imported from text files, from other statistical software or from Excel. Table 1 provides examples of some commands that allow us to work with the BS distribution, whereas similar instructions may be used for the other distributions; for more details on how to use the gbs package, see Barros et al. (2009).

Table 1: Basic functions of the gbs package.

Function	Instruction	Result
CDF	pgbs(1.0, alpha=0.5, beta=1.0)	0.500
PDF	dgbs(1.0, alpha=0.5, beta=1.0)	0.798
QF	qgbs(0.5, alpha=0.5, beta=1.0)	1.000
numbers	rgbs(n=100, alpha=1.0, beta=1.0)	100 BS(1, 1) random numbers are generated.
MLE	mlegbs(x)	ML estimates for data x are provided.
EDA	descriptiveSummary(x)	A descriptive summary of data x is obtained.
histogram	histgbs(x, boxPlot=T, pdfLine=T)	Histogram, boxplot and fitted PDF for x are given.

3. GOF Tests with Censored and Uncensored Data

In this section, we define the notations and provide some transformations useful for obtaining the graphical tools proposed in the paper. In addition, we establish the hypotheses of interest and the corresponding test statistics to assess GOF for NLS distributions with censored and uncensored data.

3.1. Un/Censored Data, Transformations and Order Statistics

Let $T_1, \ldots, T_n$ be a sample of size n extracted from a RV T following a distribution with CDF $F(\cdot)$. Also, let $T_{1:n} \leq \cdots \leq T_{n:n}$ be the order statistics (OSs) of $T_1, \ldots, T_n$, with $t_{1:n} \leq \cdots \leq t_{n:n}$ and $t_1, \ldots, t_n$ being their corresponding observations (data). The ECDF is defined by

$$F_n(t) = \begin{cases} 0, & \text{if } t < t_{1:n}; \\ \frac{j}{n} = w_{j:n} + \frac{1}{2n}, & \text{if } t_{j:n} \leq t < t_{j+1:n}, \text{ for } j = 1, \ldots, n-1; \\ 1, & \text{if } t \geq t_{n:n}; \end{cases} \tag{1}$$

where $w_{j:n} = [2j-1]/[2n]$.

As it is well-known, the RV U given by the transformation

$$U = F(T) \tag{2}$$

follows a uniform distribution in [0,1], denoted by U(0,1), for any continuous $F(\cdot)$, which is known as probability integral transformation (PIT).

Another transformation associated with the U(0, 1) distribution was proposed by Michael (1983), who noted that, if $U \sim \mathrm{U}(0,1)$, then the RV given by the transformation

$$S = \frac{2}{\pi} \arcsin(\sqrt{U}), \tag{3}$$

which is known as stabilized probability transformation, follows a distribution with PDF

$$f_S(s) = \frac{2}{\pi} \sin(\pi s), \quad 0 < s < 1. \tag{4}$$

The OSs $S_{1:n} \leq \cdots \leq S_{n:n}$ associated with a sample of size n from the distribution of the RV S given in (3) have a constant asymptotic variance, because as n goes to ∞ and j/n to q, $\mathrm{Var}[n\, S_{j:n}]$ goes to $1/\pi^2$, which is independent of q, for $j = 1, \ldots, n$; see Michael (1983).

Consider a sample with censoring proportion p, which conducts to r failure (uncensored) data and $n - r$ censored data. Note that n and p are controlled by the researcher. Assume a type-II right censorship, so that, in this case, r is fixed and $n - r$ observations are greater than the censoring point $t_{r:n}$. Then, $U_{1:n} = F(T_{1:n}) \leq \cdots \leq U_{r:n} = F(T_{r:n})$ are the smallest r OSs of the type-II censored sample of size n and $n - r$ observations are greater than $u_{r:n} = F(t_{r:n})$. Let $U_{j:n} = F(T_{j:n})$ be the jth OS of a sample of size n extracted from a RV

$U \sim \mathrm{U}(0,1)$ and $u_{j:n} = F(t_{j:n})$ be its observed value, for $j = 1, \ldots, n$. Consider the transformation

$$U'_{j:n} = \frac{U_{j:n}[B_{r,n-r+1}(U_{r:n})]^{1/r}}{U_{r:n}}, \quad j = 1, \ldots, r, \quad r = 1, \ldots, n, \tag{5}$$

where $B_{r,n-r+1}(x) = I_x(r, n-r+1)$ is the Beta$(r, n-r+1)$ CDF, with $I_x(\cdot,\cdot)$ being the incomplete beta ratio function. Hence, the OSs $U'_{1:n}, \ldots, U'_{r:n}$ obtained from the transformation given in (5) are distributed as the OSs from a complete sample of size r from $U' \sim \mathrm{U}(0,1)$; see details in Theorem 1 of Lin et al. (2008), Theorem 1 of Michael & Schucany (1979) and Theorem 8 of Fischer & Kamps (2011).

3.2. Hypotheses

Consider the hypotheses: H_0: "The data come from a RV T with CDF $F(\cdot)$" versus H_1: "The data do not come from this RV". The hypothesized distribution with CDF $F(\cdot)$ is indexed by a parameter vector $\boldsymbol{\theta}$ that can contain location (μ), scale (β), shape (α) parameters, or any other parameter not necessarily of location and scale, that is, T belongs to the NLS family. If the CDF is completely specified in H_0, that is, $\boldsymbol{\theta}$ is assumed to be known, the data must be transformed for testing uniformity. On the contrary, the parameters must be consistently estimated and the data transformed for testing normality of the distribution under H_0.

3.3. Test Statistics for GOF with Uncensored Data

In order to test H_0 established in Subsection 3.2, where $F(\cdot)$ is completely specified, and then to assess GOF of a distribution to a censored or uncensored data set, we consider test statistics based on the ECDF $F_n(\cdot)$ defined in (1). The most common statistics constructed with the ECDF use vertical differences between $F_n(t)$ and $F(t)$ by means of the supremum and quadratic classes. Statistics that consider the mentioned classes are AD, CM, KS, KU and WA given by

$$\mathrm{AD} = n \int_{-\infty}^{\infty} \frac{[F_n(t) - F(t)]^2}{F(t)[1 - F(t)]} \, \mathrm{d}F(t), \tag{6}$$

$$\mathrm{CM} = n \int_{-\infty}^{\infty} [F_n(t) - F(t)]^2 \, \mathrm{d}F(t), \tag{7}$$

$$\mathrm{KS} = \sup_t \big|F_n(t) - F(t)\big| = \max\Big\{\sup_t \big\{F_n(t) - F(t)\big\}, \sup_t \big\{F(t) - F_n(t)\big\}\Big\}, \tag{8}$$

$$\mathrm{KU} = \sup_t \big\{F_n(t) - F(t)\big\} + \sup_t \big\{F(t) - F_n(t)\big\}, \tag{9}$$

$$\mathrm{WA} = n \int_{-\infty}^{\infty} \Big[F_n(t) - F(t) - \int_{-\infty}^{\infty} \big\{F_n(t) - F(t)\big\} \, \mathrm{d}F(t)\Big]^2 \, \mathrm{d}F(t). \tag{10}$$

By considering (3) and (8), we can define the MI statistic as

$$\begin{aligned}\text{MI} = \max\Big\{ &\sup_t \left\{ \tfrac{2}{\pi}\arcsin(F_n(t)) - \tfrac{2}{\pi}\arcsin(F(t)) \right\}, \\ &\sup_t \left\{ \tfrac{2}{\pi}\arcsin(F(t)) - \tfrac{2}{\pi}\arcsin(F_n(t)) \right\} \Big\}. \end{aligned} \tag{11}$$

Now, by considering (2), AD, CM, KS, KU, MI and WA statistics defined in (6)-(11) can be implemented in practice by the formulas

$$\text{AD} = -2\sum_{j=1}^{n} \left[w_{j:n}\log(U_{j:n}) + \{1 - w_{j:n}\}\log(1 - U_{j:n})\right] - n, \tag{12}$$

$$\text{CM} = \sum_{j=1}^{n} \left[U_{j:n} - w_{j:n}\right]^2 + \tfrac{1}{12n}, \tag{13}$$

$$\text{KS} = \max\Big\{ \max_{1\le j\le n} \left\{w_{j:n} + \tfrac{1}{2n} - U_{j:n}\right\},\ \max_{1\le j\le n} \left\{U_{j:n} - w_{j:n} + \tfrac{1}{2n}\right\} \Big\}, \tag{14}$$

$$\text{KU} = \max_{1\le j\le n} \Big\{w_{j:n} + \tfrac{1}{2n} - U_{j:n}\Big\} + \max_{1\le j\le n} \Big\{U_{j:n} - w_{j:n} + \tfrac{1}{2n}\Big\}, \tag{15}$$

$$\begin{aligned}\text{MI} = \max\Big\{ &\max_{1\le j\le n} \left\{ \tfrac{2}{\pi}\arcsin\left(w_{j:n} + \tfrac{1}{2n}\right) - \tfrac{2}{\pi}\arcsin(U_{j:n}) \right\}, \\ &\max_{1\le j\le n} \left\{ \tfrac{2}{\pi}\arcsin(U_{j:n}) - \tfrac{2}{\pi}\arcsin\left(w_{j:n} - \tfrac{1}{2n}\right) \right\} \Big\}, \end{aligned} \tag{16}$$

$$\text{WA} = \sum_{j=1}^{n} \left[U_{j:n} - w_{j:n}\right]^2 - n\Big[\tfrac{1}{n}\sum_{j=1}^{n} U_{j:n} - \tfrac{1}{2}\Big]^2 + \tfrac{1}{12n}, \tag{17}$$

where $w_{j:n}$ and $U_{j:n}$ are as given in (1) and (5), respectively. Further details on the expressions provided in (6)-(17) can be found in D'Agostino & Stephens (1986, Ch. 4) and Michael (1983). Quantiles of the distribution of the statistics AD, CM, KS, KU, MI and WA must be obtained under H_0. However, if the distribution under this hypothesis is not completely specified, its parameters must be properly estimated and the AD, CM, KS, KU, MI and WA statistics must be modified for the distribution under H_0. These modified statistics are denoted by $\text{AD}^\star$, $\text{CM}^\star$, $\text{KS}^\star$, $\text{KU}^\star$, $\text{MI}^\star$ and $\text{WA}^\star$, and their calculated values by $\text{ad}^\star$, $\text{cm}^\star$, $\text{ks}^\star$, $\text{ku}^\star$, $\text{mi}^\star$ and $\text{wa}^\star$, respectively. In this case, new quantiles of the distribution of $\text{AD}^\star$, $\text{CM}^\star$, $\text{KS}^\star$, $\text{KU}^\star$, $\text{MI}^\star$ and $\text{WA}^\star$ must be computed under H_0.

3.4. Test Statistics for GOF with Censored Data

To test H_0 where $F(\cdot)$ is completely specified and then to assess GOF in practice with r uncensored data and $n - r$ type-II right censored data, we use the results presented in D'Agostino & Stephens (1986, Ch. 4) and adapt the statistics given in (12)-(17) as

$$\begin{aligned}\mathrm{AD}_{r,n} =& -2\sum_{j=1}^{r}\left[w_{j:n}\log(U_{j:n}) - \{1-w_{j:n}\}\log(1-U_{j:n})\right] \\ &- \frac{1}{n}\big[\{n-r\}^2\log(1-U_{r:n}) - r^2\log(U_{r:n}) + n^2U_{r:n}\big],\end{aligned} \tag{18}$$

$$\mathrm{CM}_{r,n} = \sum_{j=1}^{r}\left[U_{j:n} - w_{j:n}\right]^2 + \tfrac{r}{12n^2} + \tfrac{n}{3}\left[U_{r:n} - \tfrac{r}{n}\right]^3, \tag{19}$$

$$\mathrm{KS}_{r,n} = \max\Big\{\max_{1\leq j\leq r}\left\{w_{j:n} + \tfrac{1}{2n} - U_{j:n}\right\}, \max_{1\leq j\leq r}\left\{U_{j:n} - w_{j:n} + \tfrac{1}{2n}\right\}\Big\}, \tag{20}$$

$$\mathrm{KU}_{r,n} = \max_{1\leq j\leq r}\Big\{w_{j:n} + \tfrac{1}{2n} - U_{j:n}\Big\} + \max_{1\leq j\leq r}\Big\{U_{j:n} - w_{j:n} + \tfrac{1}{2n}\Big\}, \tag{21}$$

$$\begin{aligned}\mathrm{MI}_{r,n} = \max\Big\{ &\max_{1\leq j\leq r}\left\{\tfrac{2}{\pi}\arcsin\left(w_{j:n} + \tfrac{1}{2n}\right) - \tfrac{2}{\pi}\arcsin(U_{j:n})\right\}, \\ &\max_{1\leq j\leq r}\left\{\tfrac{2}{\pi}\arcsin(U_{j:n}) - \tfrac{2}{\pi}\arcsin\left(w_{j:n} - \tfrac{1}{2n}\right)\right\}\Big\},\end{aligned} \tag{22}$$

$$\begin{aligned}\mathrm{WA}_{r,n} =& \sum_{j=1}^{r}\left[U_{j:n} - w_{j:n}\right]^2 + \tfrac{r}{12n^2} + \tfrac{n}{3}\left[U_{r:n} - \tfrac{r}{n}\right]^3 \\ &- nU_{r:n}\Big[\frac{r}{n} - \frac{U_{r:n}}{2} - \frac{\sum_{j=1}^{r}U_{j:n}}{nU_{r:n}}\Big]^2.\end{aligned} \tag{23}$$

The quantiles of the distribution of the $\mathrm{AD}_{r,n}$, $\mathrm{CM}_{r,n}$, $\mathrm{KS}_{r,n}$, $\mathrm{KU}_{r,n}$, $\mathrm{MI}_{r,n}$ and $\mathrm{WA}_{r,n}$ statistics given in (18)-(23) must be obtained under H_0. However, if the distribution under H_0 is not completely specified, its parameters must be properly estimated, taking into account the censorship, and the statistics must be modified for each case under H_0. We denote these statistics by $\mathrm{AD}^\star_{r,n}$, $\mathrm{CM}^\star_{r,n}$, $\mathrm{KS}^\star_{r,n}$, $\mathrm{KU}^\star_{r,n}$, $\mathrm{MI}^\star_{r,n}$ and $\mathrm{WA}^\star_{r,n}$, and their calculated values by $\mathrm{ad}^\star_{r,n}$, $\mathrm{cm}^\star_{r,n}$, $\mathrm{ks}^\star_{r,n}$, $\mathrm{ku}^\star_{r,n}$, $\mathrm{mi}^\star_{r,n}$ and $\mathrm{wa}^\star_{r,n}$, respectively. Also, new quantiles of the distribution of $\mathrm{AD}^\star_{r,n}$, $\mathrm{CM}^\star_{r,n}$, $\mathrm{KS}^\star_{r,n}$, $\mathrm{KU}^\star_{r,n}$, $\mathrm{MI}^\star_{r,n}$ and $\mathrm{WA}^\star_{r,n}$ must be computed under H_0. For more details about how to obtain the quantiles of the distributions of the corresponding test statistics under H_0, which have been studied for different distributions of the LS family with uncensored and censored, see D'Agostino & Stephens (1986), Castro-Kuriss et al. (2009), Castro-Kuriss et al. (2010) and Castro-Kuriss (2011). In the next subsections, we mention that, for NLS distributions, analogous results for assessing GOF with both censored and uncensored data can be considered.

3.5. GOF Tests for NLS Distributions with Uncensored Data

If the hypotheses of interest H_0 is $F(t) = \Phi([t-\mu]/\beta)$ with unknown parameters, we can consider the procedure detailed in Algorithm 1.

We consider a procedure that can be applied to NLS distributions based on the work proposed by Chen & Balakrishnan (1995), which provides an approximate GOF method. This method first transforms the data to normality and then applies Algorithm 1, generalizing it. Testing normality in H_0 allows us to compute the critical values of the corresponding test statistics, independently of the parameter

Algorithm 1 GOF test for normality with uncensored data

1: Collect data $t_1, \ldots, t_n$ and order them as $t_{1:n}, \ldots, t_{n:n}$;
2: Estimate μ and β of $\Phi([t-\mu]/\beta)$ by $\widehat{\mu}$ and $\widehat{\beta}$, respectively, with $t_1, \ldots, t_n$;
3: Obtain $\widehat{u}_{j:n} = \Phi(\widehat{z}_j)$, with $\widehat{z}_j = [t_{j:n} - \widehat{\mu}]/\widehat{\beta}$, for $j = 1, \ldots, n$;
4: Evaluate AD*, CM*, KS*, KU*, MI* and WA* statistics at $\widehat{u}_{j:n}$;
5: Compute the p-values of the AD*, CM*, KS*, KU*, MI* and WA* statistics;
6: Reject H_0: $F(t) = \Phi([t-\mu]/\beta)$ for a specified significance level based on the obtained p-values.

estimators, if they are consistent and the sample size is large enough. To test the hypotheses of interest defined in Subsection 3.2, for $\alpha > 0$ and $\beta > 0$ unknown, we consider a generalization of Algorithm 1 based on Chen & Balakrishnan (1995)'s method, which is detailed in Algorithm 2. Following Chen & Balakrishnan (1995); in general, we recommend a sample size $n > 20$, so that the approximations work well. This is also valid for the algorithms presented in the next sections.

Algorithm 2 GOF test for NLS distributions with uncensored data

1: Collect data $t_1, \ldots, t_n$ and order them as $t_{1:n}, \ldots, t_{n:n}$;
2: Estimate α and β of $F(t;\alpha,\beta)$ by $\widehat{\alpha}$ and $\widehat{\beta}$, respectively, with $t_1, \ldots, t_n$;
3: Compute $\widehat{v}_{j:n} = F(t_{j:n};\widehat{\alpha},\widehat{\beta})$, for $j = 1, \ldots, n$;
4: Calculate $\widehat{y}_j = \Phi^{-1}(\widehat{v}_{j:n})$, where $\Phi^{-1}(\cdot)$ is the $\mathrm{N}(0,1)$ inverse CDF or QF $z(\cdot)$;
5: Obtain $\widehat{u}_{j:n} = \Phi(\widehat{z}_j)$, with $\widehat{z}_j = [\widehat{y}_j - \bar{y}]/s_y$, $\bar{y} = \sum_{j=1}^n \widehat{y}_j/n$ and $s_y = [\sum_{j=1}^n (\widehat{y}_j - \bar{y})^2/(n-1)]^{1/2}$;
6: Repeat Steps 4-6 of Algorithm 1 with $F(t) = F(t;\alpha,\beta)$.

3.6. GOF Tests for NLS Distributions with Censored Data

As mentioned, GOF tests for NLS distributions with uncensored data can be considered for censored data adapting them or the GOF statistics.

To test the hypotheses of interest defined in Subsection 3.2, for $\alpha > 0$ and $\beta > 0$ both of them unknown and type-II right censored data, we first transform censored data into uncensored data by using (5). Algorithm 3 details the corresponding GOF procedure.

Algorithm 3 GOF test 1 for NLS distributions with censored data

1: Repeat Steps 1-3 of Algorithm 2;
2: Determine $\widehat{v}'_{j:n} = \widehat{v}_{j:n}[B_{r,n-r+1}(\widehat{v}_{r:n})]^{1/r}/\widehat{v}_{r:n}$, for $j = 1, \ldots, r$ and $r = 1, \ldots, n$;
3: Repeat Steps 4-6 of Algorithm 2 replacing $\widehat{v}_{j:n}$ by $\widehat{v}'_{j:n}$ in Step 4.

Second, as mentioned, another way to perform a GOF test for NLS distributions with censored data can be obtained adapting the GOF statistics, as detailed in Algorithm 4.

Algorithm 4 GOF test 2 for NLS distributions with censored data

1: Repeat Steps 1-5 of Algorithm 2;
2: Evaluate $\mathrm{AD}^{\star}_{r,n}$, $\mathrm{CM}^{\star}_{r,n}$, $\mathrm{KS}^{\star}_{r,n}$, $\mathrm{KU}^{\star}_{r,n}$, $\mathrm{MI}^{\star}_{r,n}$ and $\mathrm{WA}^{\star}_{r,n}$ statistics at $\widehat{u}_{j:n}$;
3: Determine the p-values of $\mathrm{AD}^{\star}_{r,n}$, $\mathrm{CM}^{\star}_{r,n}$, $\mathrm{KS}^{\star}_{r,n}$, $\mathrm{KU}^{\star}_{r,n}$, $\mathrm{MI}^{\star}_{r,n}$ and $\mathrm{WA}^{\star}_{r,n}$ statistics;
4: Reject the corresponding H_0 for a specified significance level depending on the obtained p-values.

4. Graphical Tools

In this section, based on Algorithm 4, we provide acceptance regions for the KS and MI statistics that allow graphical tools to be obtained for assessing GOF in NLS distributions.

4.1. PP and SP Plots

PP and QQ plots are well known, however this is not the case of the SP plot. Note that, if the distribution under H_0 is U(0,1), then the corresponding QQ plot is essentially the same as the PP plot; see Castro-Kuriss et al. (2009). Michael (1983) used the arcsin transformation to stabilize the variance of points on probability graphs associated with the KS test to propose the SP plot; see PDF given in (4) and comments relating thereto in Subsection 3.1. Formulas to construct PP and SP plots are provided in Table 2. In this table, $w_{j:n}$ is as given in (1) and $u_{j:n}$ as given above (5).

Table 2: formulas for the indicated probability plot.

Plot	Abscissa	Ordinate
PP	$w_{j:n}$	$u_{j:n}$
SP	$x_{j:n} = \frac{2}{\pi}\arcsin\left(\sqrt{w_{j:n}}\right)$	$s_{j:n} = \frac{2}{\pi}\arcsin\left(\sqrt{u_{j:n}}\right)$

4.2. Acceptance Regions for Probability Plots

Acceptance regions for PP and SP plots can be constructed by means of KS and MI statistics. Thus, we can display acceptance bands to assess whether the data can come from the distribution under H_0 with these two statistics; see Castro-Kuriss et al. (2009) and Castro-Kuriss et al. (2010). Formulas to construct $100\varrho\%$ acceptance regions on PP and SP plots with right type-II censored data, based on $\mathrm{KS}^{\star}_{r,n}$ and $\mathrm{MI}^{\star}_{r,n}$, are displayed in Table 3. In this table, w and x are continuous versions of $w_{j:n}$ and $x_{j:n}$ given in Table 2 to construct the acceptance bands. If all of the r data points lie within the constructed acceptance bands, then H_0 cannot be rejected at the $1-\varrho$ level. Also, if a noticeable curvature is detected, we can question such a hypothesis. Table 3 may be adapted to the uncensored case with $r = n$ and the quantiles must be replaced by the quantiles of the distribution of the corresponding statistics without censorship.

To test H_0 defined in Subsection 3.2, for some $\alpha > 0$ and $\beta > 0$ and type-II right censored data, we propose a graphical tool which procedure is detailed in

TABLE 3: $100\varrho\%$ acceptance regions for the indicated plot and statistic with 100ϱth quantiles $\mathrm{ks}^{\star}_{r,n,\varrho}$ and $\mathrm{mi}^{\star}_{r,n,\varrho}$.

Plot	Stat	Bands defining acceptance regions
PP	KS*	$[\max\{w-\mathrm{ks}^{\star}_{r,n,\varrho}+\frac{1}{2n},0\},\min\{w+\mathrm{ks}^{\star}_{r,n,\varrho}-\frac{1}{2n},1\}]$
PP	MI*	$[\max\{\sin^2(\arcsin(w^{\frac{1}{2}})-\frac{\pi}{2}\mathrm{mi}^{\star}_{r,n,\varrho}),0\},\min\{\sin^2(\arcsin(w^{\frac{1}{2}})+\frac{\pi}{2}\mathrm{mi}^{\star}_{r,n,\varrho}),1\}]$
SP	KS*	$[\max\{\frac{2}{\pi}\arcsin(\{\sin^2(\frac{\pi}{2}x)-\mathrm{ks}^{\star}_{r,n,\varrho}+\frac{1}{2n}\}^{\frac{1}{2}}),0\},$ $\min\{\frac{2}{\pi}\arcsin(\{\sin^2(\frac{\pi}{2}x)+\mathrm{ks}^{\star}_{r,n,\varrho}-\frac{1}{2n}\}^{\frac{1}{2}}),1\}]$
SP	MI*	$[\max\{x-\mathrm{mi}^{\star}_{r,n,\varrho},0\},\min\{x+\mathrm{mi}^{\star}_{r,n,\varrho},1\}]$

Algorithm 5 based on Algorithm 4 and Tables 2 and 3, which is valid for censored or uncensored data. We consider the general case for unknown parameters of the distribution under H_0, although it may also be used when the parameters are known.

Algorithm 5 Acceptance regions to test GOF for NLS distributions with censored data

1: Repeat Step 1 of Algorithm 4;
2: Draw the PP plot with points $w_{j:n}$ versus $\widehat{u}_{j:n}$, for $j=1,\dots,r$ and $r=1,\dots,n$;
3: Draw the SP plot with points $x_{j:n}=[2/\pi]\arcsin(\sqrt{w_{j:n}})$ versus $s_{j:n}=[2/\pi]\arcsin(\sqrt{\widehat{u}_{j:n}})$;
4: Construct acceptance bands according to Table 3 specifying a $1-\varrho$ level;
5: Decide if H_0 must be rejected for the corresponding significance level;
6: Corroborate decision in Step 5 with the p-values after evaluating $\mathrm{KS}^{\star}_{r,n}$ and $\mathrm{MI}^{\star}_{r,n}$ statistics at $\widehat{u}_{j:n}$.

5. Examples

In this section, we consider several real-world data sets and NLS distributions under H_0 and decide whether such data may reasonably come from the hypothesized distribution. The results are also displayed by means of the probability plots with the acceptance bands proposed in Section 4.

5.1. Example 1: Uncensored Sea Data

These data are drawn from sea surface temperatures (in °K), generated by an advanced very high resolution radiometer. We name data “sea”. The sample size is $n=88$ and the truncation point is $\kappa=278.187$°K; see DePriest (1983). An EDA for sea data is provided in Table 4, including the coefficients of variation (CV), skewness (CS) and kurtosis (CK) and the standard deviation (SD). Figure 1 displays their histogram and boxplot. This EDA indicates that the TBS distribution can be suitable for describing such data, as a competitor of the TN distribution. Some atypical data are detected in the boxplot, but their study is not considered here since it is beyond the scope of this work. We consider the TBS and TN models under H_0 with parameters estimated. We use the tbs and truncnorm `R` packages, respectively, for making this estimation with the ML method. The corresponding estimates and the p-values from the GOF tests are displayed in Table 5. According to these p-values, both distributions perform a good fitting to the data. We show the PP and SP plots in Figure 2 for the TBS distribution defined in Section 2. From this figure, note that the points are well aligned, as expected, due to the high p-values obtained for the TBS distribution, and that all the points fall inside the 95% acceptance bands, which confirms the good fitting of the TBS distribution to sea data.

Table 4: Descriptive statistics for sea data.

Median	Mean	SD	CV	CS	CK	Range	Min	Max	n
279.5	279.6	0.787	0.003	0.008	3	3.8	278.2	282	88

where CV: coefficient of variation, CS: coefficient of skewness, CK: coefficient of kurtosis SD: standard deviation, Min: minimum value and Max: maximum value.

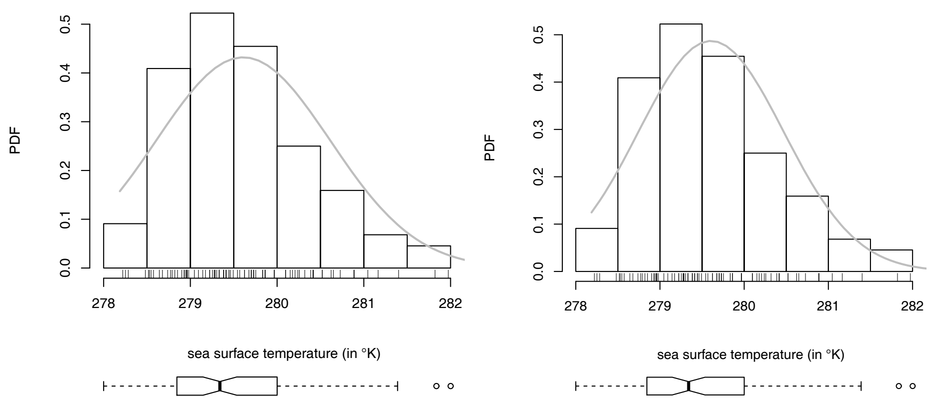

Figure 1: Histogram, boxplot and estimated PDF from TN (left) and TBS (right) distributions for sea data.

Table 5: Estimated parameters and values of the statistics for the indicated distribution under H_0 with sea data.

Distribution	Parameter	Estimate	Estimated statistic	p-value
TBS	α	0.0057	$\mathrm{ks}^\star = 0.0653$	$[0.4, 0.5]$
	β	278.8488	$\mathrm{mi}^\star = 0.0414$	$[0.7, 0.8]$
TN	μ	279.6093	$\mathrm{ks}^\star = 0.0600$	$[0.5, 0.6]$
	σ	1.0011	$\mathrm{mi}^\star = 0.0369$	$[0.8, 0.9]$

where $\mathrm{ks}^\star$ and $\mathrm{mi}^\star$ are the estimated values of the KS and MI statistics.

5.2. Example 2: Uncensored Forestry Data

Data correspond to the diameter at breast height (DBH, in cm) of loblolly pine trees from a plantation in the Western Gulf Coast. We name data "forestry". The sample size is $n = 75$ and the left-truncation point is $\kappa = 6$ cm; see Leiva,

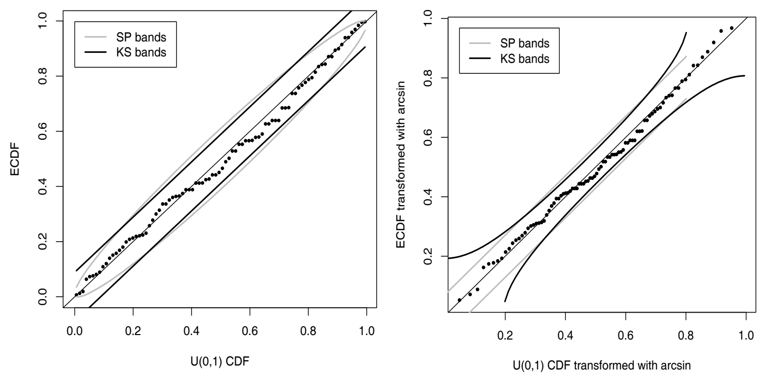

FIGURE 2: PP (left) and SP (right) plots with 95% acceptance bands for the TBS distribution with sea data.

Ponce, Marchant & Bustos (2012). Table 6 provides an EDA of forestry data, which indicates once again that the TBS distribution can be a good model for the data.

From the histograms displayed in Figure 3, note that the fit of the TBS distribution is apparently better than for the TN distribution. The corresponding estimates and the p-values from the GOF tests are displayed in Table 7. According to these p-values, both distributions perform a reasonable fit to the data, but clearly the TBS distribution has a better performance. Figure 4 shows the PP and SP plots for the TBS distribution. From this figure, note that once again the points are well aligned, as expected, due to the p-values obtained for the TBS distribution, and all the points fall within distribution to forestry data.

TABLE 6: Descriptive statistics for forestry data.

Median	Mean	SD	CV	CS	CK	Range	Min	Max	n
8.20	8.19	1.013	0.124	0.053	2.253	4.1	6.2	10.3	75

TABLE 7: Estimated parameters and statistic values for the indicated distribution under H_0 with forestry data.

Distribution	Parameter	Estimate	Estimated statistic	p-value
TBS	α	0.12804	$\text{ks}^\star = 0.07031$	$[0.4, 0.5]$
	β	8.23963	$\text{mi}^\star = 0.05923$	$[0.25, 0.4]$
TN	μ	7.54905	$\text{ks}^\star = 0.08201$	$[0.2, 0.25]$
	σ	2.42369	$\text{mi}^\star = 0.06697$	$[0.1, 0.2]$

5.3. Example 3: Uncensored Survival Data

These data are drawn from survival times (in days) of pigs injected with a dose of tubercle bacilli, under a regimen corresponding to 4.0×10^6 bacillary units per 0.5 ml ($\log(4.0 \times 10^6) = 6.6$). We name data “survival”. The sample size is $n = 72$ guinea pigs infected with tubercle bacilli in regimen 6.6; see Azevedo et al. (2012).

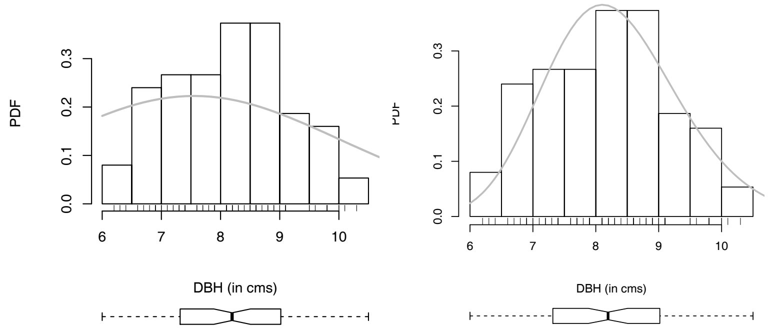

FIGURE 3: Histogram, boxplot and estimated PDF from TN (left) and TBS (right) distributions for forestry data.

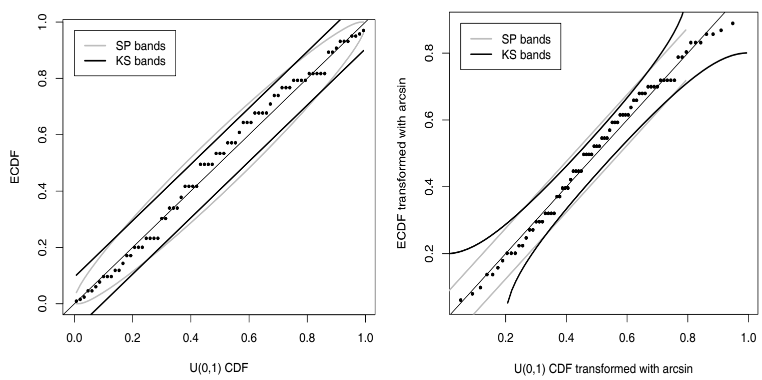

FIGURE 4: PP (left) and SP (right) plots with 95% acceptance bands for the TBS distribution with forestry data.

Table 8 and Figure 5 provide an EDA of survival data. From this EDA, we detect a distribution skewed to the right with the presence of some outliers. We are considering an outlier in the sense defined by Hubert & Vandervieren (2008). We propose the BS and BS-t distributions defined in Section 2 for analyzing survival data. ML estimates of the BS and BS-t parameters, which are obtained using the gbs package, and the p-values from the GOF tests are displayed in Table 9. Clearly the BS distribution does not properly fit these data, whereas the BS-t distribution provides a better fit. Figures 6 and 7 show the PP and SP plots for the BS and the BS-t distributions, respectively. From Figure 6, note that the points are not well aligned, particularly in the center, and that one observation (number 15) is outside the KS band. In Figure 7, the points are better aligned, as expected, due to the p-values obtained for the BS-t distribution, and all the points fall within the 95% acceptance bands, indicating the good fitting of the BS-t distribution to survival data.

TABLE 8: Descriptive statistics for survival data.

Median	Mean	SD	CV	CS	CK	Range	Min	Max	n
70.00	99.82	81.12	0.81	1.76	5.46	364	12	376	72

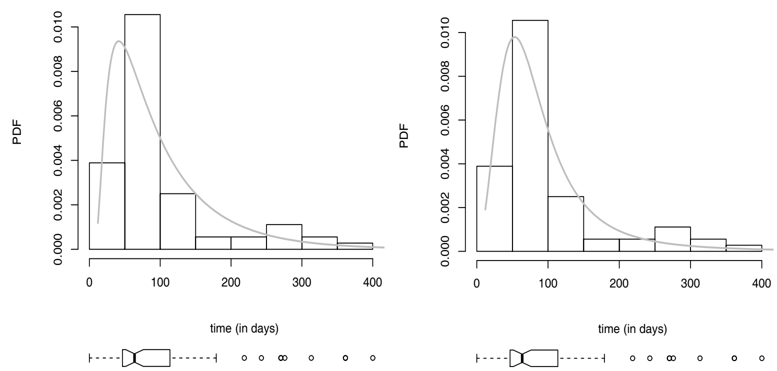

FIGURE 5: Histogram, boxplot and estimated PDF from BS (left) and BS-t (right) distributions for survival data.

TABLE 9: Estimated parameters and statistic values for the indicated distribution under H_0 with survival data.

Distribution	Parameter	Estimate	Estimated statistic	p-value
BS	α	0.7600	$\text{ks}^\star = 0.08848$	$[0.01, 0.05]$
	β	77.5348	$\text{mi}^\star = 0.07318$	$[0.05, 0.1]$
BS-t	α	0.6085	$\text{ks}^\star = 0.08201$	$[0.1, 0.2]$
	β	75.5880	$\text{mi}^\star = 0.05908$	$[0.25, 0.4]$
	ν	5.0000	-	-

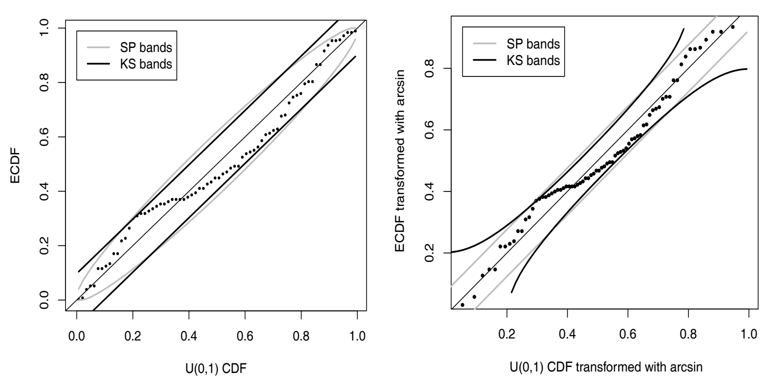

FIGURE 6: PP (left) and SP (right) plots with 95% acceptance bands for the BS distribution with survival data.

5.4. Example 4: Uncensored Survival Data with Outliers

Next, we conduct a simple empirical robustness study. First, we add a large value (outlier) to the data and call them "survival1", so that we now have a sample

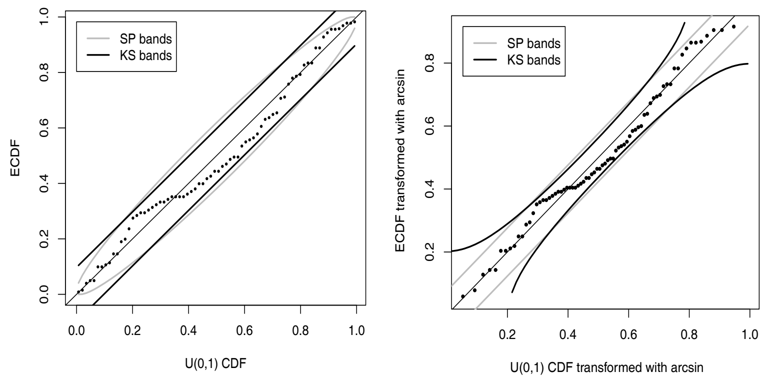

Figure 7: PP (left) and SP (right) plots with 95% acceptance bands for the BS-t distribution with survival data.

of size $n = 73$. This new observation is greater than all the observed values ($t_{73:73} = 580$). Second, we add another large value ($t_{74:74} = 750$) to the data and call them as "survival2', so that we now have a sample of size $n = 74$. Then, we add one more outlier, corresponding to the value $t_{75:75} = 1000$, and call these data as "survival3'. Figure 8 displays usual and adjusted boxplots and stripcharts for survival data, as well as for the data with one, two and three outliers, that is, for survival1, survival2 and survival3. The adjusted boxplot is often used for skewed data because it includes a robust measure of skewness; see Hubert & Vandervieren (2008) and Hubert & Vanderveeken (2008). The stripchart is a scatterplot in one dimension, where all the observations are plotted. From these graphs we can visualize the effect of the outliers added to the data. In the adjusted boxplot for survival1 (see Figure 8-center), it is possible to note that the first added value is an outlier, but when a second atypical value is added for survival2, only this second value is detected as outlier, the first is no longer an outlier for this data set. With the third value being part of the sample, which is much greater than the others, the presence of two outliers is detected. ML estimates and the p-values of the GOF tests for survival1, survival2 and survival3 are provided in Table 10. Note that minor changes in the estimated parameters and in the bounds for the p-values of the tests are detected. For survival3 data, the differences are more noticeable. We conclude that the GOF tests are relatively robust to outliers when the BS-t distribution is considered under H_0, in particular the MI test, nonetheless a more extensive study on this issue should be carried out.

Figure 9 shows the PP and SP plots based on the BS-t distribution for the survival data with three outliers added. From this figure, note that the points are still well aligned, particularly in the center, but now observation number 15 is not near the bands and there is one observation (number 42) outside the KS band. If we compare Figures 7 and 9, there are no apparently visual differences with minor different alignments in the points, specifically because rejection is due to points in the center and not in the tails of the sample where the outliers are located.

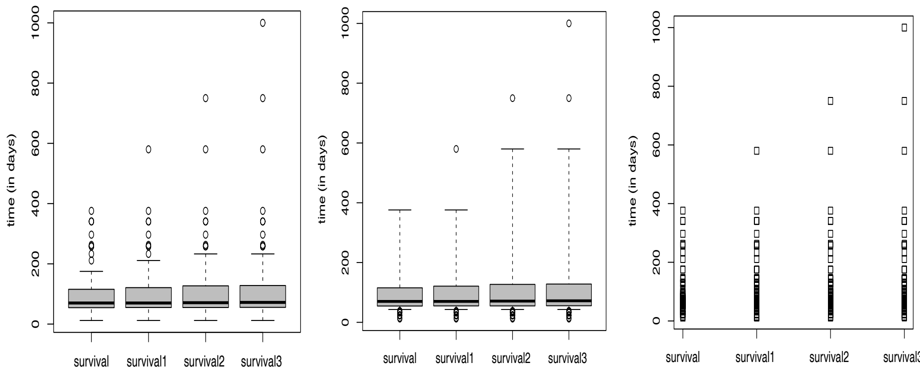

FIGURE 8: Boxplots (left), adjusted boxplots (center) and stripchart (right) for survival and outliers added.

TABLE 10: Estimated parameters and statistic values for the indicated distribution under H_0 with survival data.

Distribution	Parameter	Estimate	Estimated statistic	p-value	n	Data set
BS-t	α	0.6036	$ks^\star = 0.09313$	$[0.10, 0.2]$	73	survival1
	β	76.2200	$mi^\star = 0.05916$	$[0.25, 0.4]$		
	ν	4.0000	-	-		
BS-t	α	0.5806	$ks^\star = 0.095664$	$[0.05, 0.1]$	74	survival2
	β	76.0500	$mi^\star = 0.058160$	$[0.25, 0.4]$		
	ν	3.0000	-	-		
BS-t	α	0.6049	$ks^\star = 0.102022$	$[0.01, 0.05]$	75	survival3
	β	77.2200	$mi^\star = 0.060802$	$[0.20, 0.25]$		
	ν	3.0000	-	-		

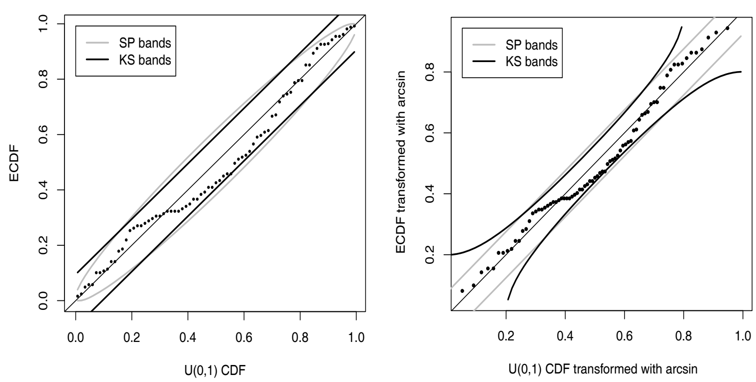

FIGURE 9: PP (left) and SP (right) plots with bands for the BS-t distribution with survival and 3 outliers added.

5.5. Censored Fatigue Data

These data correspond to fatigue life (in cycles $\times 10^{-3}$) aluminum coupons. We name data "fatigue"; see details in Birnbaum & Saunders (1969) and Leiva, Rojas, Galea & Sanhueza (2014). We consider a censored fatigue data sample, such as in Barros et al. (2014), so that we have $r = 80$ failures and $n - r = 21$ data censored. From the histogram displayed in Figure 10, note that the distribution of fatigue data is clearly skewed to the right and that the BS distribution shows a good fitting to these data, which must be corroborated.

The parameters are estimated at $\widehat{\alpha} = 0.1751$ and $\widehat{\beta} = 132.2525$, by the ML method, considering the presence of type-II censoring. The KS and MI statistics are computed such as in Examples 1 and 2, and their corresponding p-values based on the BS distribution are in $[0.2, 0.25]$ and $[0.4, 0.5]$. According to these p-values, we cannot reject the null hypothesis that the censored sample comes from a BS distribution, which can be confirmed from the PP and SP plots shown in Figure 11. This result is consistent with those obtained by other authors.

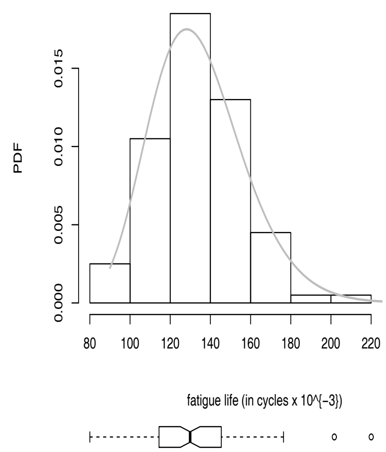

FIGURE 10: Histogram, boxplot and estimated BS PDF for fatigue data.

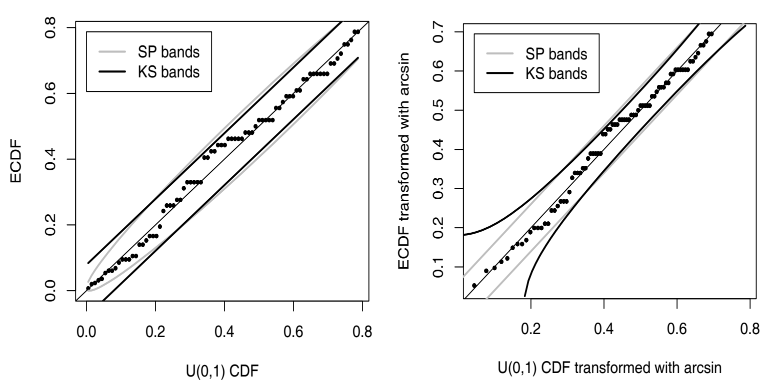

FIGURE 11: PP (left) and SP (right) plots with 95% acceptance bands for the BS distribution with fatigue data.

6. Conclusions and Future Research

We reviewed goodness-of-fit tests for life distributions not belonging to the location-scale family with uncensored and type-II right censored data. We considered the most used tests to assess goodness-of-fit based on the empirical distribution function. These tests are Anderson-Darling, Cramér-von Mises, Kolmogorov-Smirnov, Kuiper, Michael and Watson, which we have used for assessing if a non-location-scale distribution is suitable to fit a data set. We adopted the two options available in the literature on the topic to construct goodness-of-fit tests with censored data, that is, transforming the censored sample into an uncensored one thereafter applying the usual tests, or adapting the goodness-of-fit test to censored data. We proposed graphical tools related to goodness-of-fit tests based on Kolmogorov-Smirnov and Michael statistics, useful for diverse practitioners who in the past were limited to fit distributions in the location-scale family. Now, they may take into account any distribution with parameters properly estimated and sample size large enough. We illustrated the use of the results proposed in the paper, with emphasis in some Birnbaum-Saunders distributions, by means of several censored and uncensored real-world data sets, which analyses have shown their potential and how the goodness-of-fit tests and the associated graphs can be used in practice. Also, we analyzed some empirical robustness aspects through an example. We would like to mention some future research on this topic. We specifically propose:

- to study the sensitivity of the tests to outliers with different distributions under the null hypothesis.
- to analyze the robustness of the tests for different alternative hypothesis.
- to establish for distributions not considered by Chen & Balakrishnan (1995) what sample size and censoring proportion can be recommended, for approximate tests to work well.
- to study the power of the Michael test and compare it to other considered tests.
- to extend the proposed graphical tools to multivariate data.

Acknowledgments

The authors thank the Guest Editor of the Special Issue on "Current Topics in Statistical Graphics", Dr. Fernando Marmolejo-Ramos, the Editor-in-Chief of the journal, Dr. Leonardo Trujillo, and two anonymous referees for their constructive comments on an earlier version of this manuscript which resulted in this improved version. This study was partially supported by the Chilean Council for Scientific and Technological Research under the project grant FONDECYT 1120879 and by the Research Centre of Mathematics of the University of Minho with the Portuguese Funds from the "Fundação para a Ciência e a Tecnologia" under the project grant PEstOE/MAT/UI0013/2014.

[Recibido: mayo de 2014 — Aceptado: noviembre de 2014]

References

Ahmed, S., Castro-Kuriss, C., Leiva, V., Flores, E. & Sanhueza, A. (2010), 'Truncated version of the Birnbaum-Saunders distribution with an application in financial risk', *Pakistan Journal of Statistics* **26**, 293–311.

Athayde, E., Azevedo, C., Leiva, V. & Sanhueza, A. (2012), 'About Birnbaum-Saunders distributions based on the Johnson system', *Communications in Statistics, Theory and Methods* **41**, 2061–2079.

Azevedo, C., Leiva, V., Athayde, E. & Balakrishnan, N. (2012), 'Shape and change point analyses of the Birnbaum-Saunders-t hazard rate and associated estimation.', *Computational Statistics and Data Analysis* **56**, 3887–3897.

Balakrishnan, N., Gupta, R., Kundu, D., Leiva, V. & Sanhueza, A. (2011), 'On some mixture models based on the Birnbaum-Saunders distribution and associated inference', *Journal of Statistical Planning and Inference* **141**, 2175–2190.

Balakrishnan, N., Leiva, V., Sanhueza, A. & Cabrera, E. (2009), 'Mixture inverse Gaussian distribution and its transformations, moments and applications', *Statistics* **43**, 91–104.

Balakrishnan, N., Leiva, V., Sanhueza, A. & Vilca, F. (2009), 'Estimation in the Birnbaum-Saunders distribution based on scale-mixture of normals and the EM-algorithm', *Statistics and Operations Research Transactions* **33**, 171–192.

Balakrishnan, N., Ng, H. & Kannan, N. (2004), 'Goodness-of-fit tests based on spacings for progressively type-II censored data from a general location-scale distribution.', *IEEE Transactions on Reliability* **53**, 349–356.

Barr, D. & Davidson, T. (1973), 'A Kolmogorov-Smirnov test for censored samples', *Technometrics* **15**, 739–757.

Barros, M., Leiva, V., Ospina, R. & Tsuyuguchi, A. (2014), 'Goodness-of-fit tests for the Birnbaum-Saunders distribution with censored reliability data', *IEEE Transactions on Reliability* **63**, 543–554.

Barros, M., Paula, G. & Leiva, V. (2009), 'An R implementation for generalized Birnbaum-Saunders distributions', *Computational Statistics and Data Analysis* **53**, 1511–1528.

Birnbaum, Z. & Saunders, S. (1969), 'Estimation for a family of life distributions with applications to fatigue', *Journal of Applied Probability* **6**, 328–347.

Castillo, J. & Puig, P. (1997), 'Testing departures from Gamma, Rayleigh, and truncated normal distributions', *Annals of the Institute of Statistical Mathematics* **49**, 255–269.

Castro-Kuriss, C. (2011), 'On a goodness-of-fit test for censored data from a location-scale distribution with application', *Chilean Journal of Statistics* **37**, 15–136.

Castro-Kuriss, C., Kelmansky, D., Leiva, V. & Martinez, E. (2009), 'A new goodness-of-fit test for censored data with an application in monitoring processes', *Communications in Statistics: Simulation and Computation* **38**, 1161–1177.

Castro-Kuriss, C., Kelmansky, D., Leiva, V. & Martinez, E. (2010), 'On a goodness-of-fit test for normality with unknown parameters and type-II censored data', *Journal of Applied Statistics* **37**, 1193–1211.

Chen, G. & Balakrishnan, N. (1995), 'A general purpose approximate goodness-of-fit test', *Journal of Quality Technology* **27**, 154–161.

Choulakian, V. & Stephens, M. (2001), 'Goodness-of-fit tests for the generalized Pareto distribution', *Technometrics* **43**, 478–484.

Cohen, C. (1991), *Truncated and Censored Samples*, Marcel Dekker, NY.

D'Agostino, C. & Stephens, M. (1986), *Goodness of Fit Techniques*, Marcel Dekker, NY.

DePriest, D. (1983), 'Using the singly truncated normal distribution to analyse satellite data', *Communications in Statistics, Theory and Methods* **12**, 263–272.

Dufour, R. & Maag, U. (1978), 'Distribution results for modified Kolmogorov-Smirnov statistics for truncated or censored samples', *Technometrics* **20**, 29–32.

Ferreira, M., Gomes, M. & Leiva, V. (2012), 'On an extreme value version of the Birnbaum-Saunders distribution', *REVSTAT - Statistical Journal* **10**, 181–210.

Fierro, R., Leiva, V., Ruggeri, F. & Sanhueza, A. (2013), 'On a Birnbaum-Saunders distribution arising from a non-homogeneous Poisson process', *Statistics & Probability Letters* **83**, 1233–1239.

Fischer, T. & Kamps, U. (2011), 'On the existence of transformations preserving the structure of order statistics in lower dimensions', *Journal of Statistical Planning and Inference* **141**, 536–548.

Hubert, M. & Vanderveeken, S. (2008), 'Outlier detection for skewed data', *Journal of Chemometrics* **22**, 235–246.

Hubert, M. & Vandervieren, E. (2008), 'An adjusted boxplot for skewed distributions', *Computational Statistics and Data Analysis* **52**, 5186–5201.

Hui, W., Gel, Y. & Gastwirth, J. (2008), 'lawstat: An r package for law, public policy and biostatistics', *Journal Statistical Software* **28**, 1–26.

Johnson, N., Kotz, S. & Balakrishnan, N. (1994), *Continuous Univariate Distributions*, Wiley, NY.

Kotz, S., Leiva, V. & Sanhueza, A. (2010), 'Two new mixture models related to the inverse Gaussian distribution', *Methodology and Computing in Applied Probability* **12**, 199–212.

Lawless, J. (2003), *Statistical Models and Methods for Lifetime Data*, Wiley, NY.

Leiva, V., Athayde, E., Azevedo, C. & Marchant, C. (2011), 'Modeling wind energy flux by a Birnbaum-Saunders distribution with unknown shift parameter', *Journal of Applied Statistics* **38**, 2819–2838.

Leiva, V., Marchant, C., Saulo, H., Aslam, M. & Rojas, F. (2014), 'Capability indices for Birnbaum-Saunders processes applied to electronic and food industries', *Journal of Applied Statistics* **41**, 1881–1902.

Leiva, V., Ponce, M., Marchant, C. & Bustos, O. (2012), 'Fatigue statistical distributions useful for modeling diameter and mortality of trees', *Revista Colombiana de Estadística* **35**, 349–367.

Leiva, V., Rojas, E., Galea, M. & Sanhueza, A. (2014), 'Diagnostics in Birnbaum-Saunders accelerated life models with an application to fatigue data', *Applied Stochastic Models in Business and Industry* **30**, 115–131.

Leiva, V., Sanhueza, A., Kotz, S. & Araneda, N. (2010), 'A unified mixture model based on the inverse Gaussian distribution', *Pakistan Journal of Statistic* **26**, 445–460.

Leiva, V., Santos-Neto, M., Cysneiros, F. & Barros, M. (2014), 'Birnbaum-Saunders statistical modelling: A new approach', *Statistical Modelling* **14**, 21–48.

Leiva, V., Saulo, H., Leao, J. & Marchant, C. (2014), 'A family of autoregressive conditional duration models applied to financial data', *Computational Statistics and Data Analysis* **79**, 175–191.

Leiva, V., Soto, G., Cabrera, E. & Cabrera, G. (2011), 'New control charts based on the Birnbaum-Saunders distribution and their implementation', *Revista Colombiana de Estadística* **34**, 147–176.

Lin, C.-T., Huang, Y.-L. & Balakrishnan, N. (2008), 'A new method for goodness-of-fit testing based on type-II right censored samples', *IEEE Transactions on Reliability* **57**, 633–642.

Malmquist, S. (1950), 'On a property of order statistics from a rectangular distribution', *Scandinavian Actuarial Journal* **33**, 214–222.

Marden, J. (2004), 'Positions and QQ-plots', *Statistical Sciences* **19**, 606–614.

Marshall, A. & Olkin, I. (2007), *Life Distributions*, Springer, NY.

Michael, J. (1983), 'The stabilized probability plot', *Biometrika* **70**, 11–17.

Michael, J. & Schucany, W. (1979), 'A new approach to testing goodness of fit for censored samples', *Technometrics* **21**, 435–441.

O'Reilly, F. & Stephens, M. (1988), 'Transforming censored samples for testing fit', *Technometrics* **30**, 79–86.

Pakyari, R. & Balakrishnan, N. (2012), 'A general purpose approximate goodness-of-fit test for progressively type-II censored data', *IEEE Transactions on Reliability* **61**, 238–244.

Paula, G., Leiva, V., Barros, M. & Liu, S. (2012), 'Robust statistical modeling using the Birnbaum-Saunders-t distribution applied to insurance', *Applied Stochastic Models in Business and Industry* **28**, 16–34.

Rad, A., Yousefzadeh, F. & Balakrishnan, N. (2011), 'Goodness-of-fit test based on Kullback-Leibler information for progressively type-II censored data', *IEEE Transactions on Reliability* **60**, 570–579.

Rojas, F., Leiva, V., Wanke, P. & Marchant, C. (2015), 'Optimization of contribution margins in food services by modeling independent component demand', *Revista Colombiana de Estadística* **38**, 1–30.

Sanhueza, A., Leiva, V. & Lopez-Kleine, L. (2011), 'On the student-t mixture inverse Gaussian model with an application to protein production', *Revista Colombiana de Estadística* **34**, 177–195.

Stehlík, M., Strelec, L. & Thulin, M. (2014), 'On robust testing for normality in chemometrics', *Chemometrics and Intelligent Laboratory Systems* **130**, 98–108.

Vilca, F., Sanhueza, A., Leiva, V. & Christakos, G. (2010), 'An extended Birnbaum-Saunders model and its application in the study of environmental quality in Santiago, Chile', *Stochastic Environmental Research and Risk Assessment* **24**, 771–782.