



An R Package for a General Class of Inverse Gaussian Distributions

Víctor Leiva
Universidad de Valparaíso

Hugo Hernández
Universidad de Valparaíso

Antonio Sanhueza
Universidad de La Frontera

Abstract

The inverse Gaussian distribution is a positively skewed probability model that has received great attention in the last 20 years. Recently, a family that generalizes this model called inverse Gaussian type distributions has been developed. The new R package named **ig** has been designed to analyze data from inverse Gaussian type distributions. This package contains basic probabilistic functions, lifetime indicators and a random number generator from this model. Also, parameter estimates and diagnostics analysis can be obtained using likelihood methods by means of this package. In addition, goodness-of-fit methods are implemented in order to detect the suitability of the model to the data. The capabilities and features of the **ig** package are illustrated using simulated and real data sets. Furthermore, some new results related to the inverse Gaussian type distribution are also obtained. Moreover, a simulation study is conducted for evaluating the estimation method implemented in the **ig** package.

Keywords: diagnostics, goodness of fit, lifetime analysis, likelihood methods.

1. Introduction

There are several fields of statistical application where the normal distributional assumption does not hold. This is because some random variables (RV) considered in these fields follow asymmetrical models. When researchers encounter skewed data, they generally try to eliminate the asymmetry by transforming such observations so that the normal model can be used. Problems of interpretation may arise, however, when analyzing data based on such transformations. Thus, if a distribution is appropriate for modeling skewness data and besides mathematically treatable and available for users, then this one should be employed. For this reason, it is necessary to implement computational packages that take the utilization of asymmetrical models into consideration.

A positively skewed probability model that has received great attention in the last 20 years is

the inverse Gaussian distribution (IGD). The IGD is a probability model also known as the first passage time distribution of Brownian motion with positive drift, which was developed by Schrödinger (1915). Later, in 1941, Tweedie proposed the name inverse Gaussian for this distribution since its cumulant generating function has an inverse relationship with that of the normal distribution; see Tweedie (1957). The interest for the IGD is a result of its attractive statistical and probabilistic properties. For example, the IGD belongs to the exponential family, it has the reproductive property and it possesses similar inferential properties to that of the normal model; see Mudholkar and Natarajan (2002). Thus, the IGD is a natural alternative candidate to the normal distribution for modeling non-negative data with positive skewness. For more details about the IGD see Chhikara and Folks (1989), Seshadri (1993, 1999) and Johnson, Kotz, and Balakrishnan (1994, pp. 259–297).

Maximum likelihood estimates (MLE) of the parameters of the normal model can be sensitive to outlying observations. Lange, Little, and Taylor (1989) proposed the Student- t distribution (which we will call simply t distribution or t model) as an alternative to the normal case, since it has greater kurtosis than the normal model. Thus, cases which might be considered as outlying under normality, might not under the t model, producing stable parameter estimates. Recently, based on the relationship between the inverse Gaussian and normal distributions, Sanhueza, Leiva, and Balakrishnan (2008) generalized the IGD using a class of models known as elliptically contoured distributions. The elliptical univariate distributions contain all the symmetrical models in \mathbb{R} , the real line, and possess different degrees of kurtosis related to the normal model. Therefore, this generalization, called the inverse Gaussian type distribution (IGTD), can produce stable parameter estimates in the presence of outliers. In general, the IGTD is highly flexible because it allows for different degrees of kurtosis and asymmetry, other than modality and bimodality.

The aims of this paper are: (i) to present the R (R Development Core Team 2008) package named **ig**, available from the Comprehensive R Archive Network at <http://CRAN.R-project.org/package=ig>, which can be used to compute probabilistic characters, estimate parameters, carry out diagnostics and produce goodness of fit for the IGTD; and (ii) to develop two new aspects of the IGTD: lifetime analysis and generation of random numbers.

The rest of this article is organized as follows: in the second section, we give a background about probabilistic properties of the IGTD and provide new results for lifetime analysis based on this distribution; in the third section, we give a background about statistical results of the IGTD and develop a procedure for obtaining random numbers from this model, which can be useful for simulation studies; in the fourth section, the main functions of the **ig** package and some illustrative examples are presented. Finally, some conclusions are drawn.

2. Probabilistic aspects of the IGTD

In this section, we present a summary of the most important properties of the IGTD obtained by Sanhueza *et al.* (2008) and also provide some new results for lifetime analysis based on this distribution. Specifically, in the first part of this section we discuss aspects related to symmetrical distributions in the real line, also known as univariate elliptically contoured (EC) distributions, and their implementation in the R software. Later, we summarize in seven properties the most important results of the IGTD obtained by Sanhueza *et al.* (2008). In the third part of this section, we develop a lifetime analysis for the IGTD mainly based on

the hazard function (hf) including its change point. Finally, a shape analysis for this model is provided, which shows some graphical tools implemented in the **ig** package.

2.1. Symmetrical distributions in the real line

If a RV has a univariate EC distribution with location $\mu = 0$ and scale $\sigma = 1$ parameters, then the notation $Z \sim \text{EC}_1(0, 1; g)$ is used, where g is the kernel of the probability density function (PDF) of Z expressed as $f_Z(z) = c g(z^2)$, with $z \in \mathbb{R}$ and c being a normalization constant that allows $f_Z(z)$ to be a PDF. The cumulative distribution function (CDF) of Z is denoted by $F_Z(z)$. Models of this type are known as standard symmetrical distributions in the real line or symmetrical distributions about zero and can be denoted simply by $Z \sim \text{S}(g)$. For details about symmetrical distributions see [Fang, Kotz, and Ng \(1990\)](#)

In the R software, the functions `dlogis()` and `plogis()`, `dnorm()` and `pnorm()` and `dt()` and `pt()` are already implemented, which allow the standard symmetrical PDF and CDF of the logistic, normal and Student- t models to be obtained, respectively. In addition, R also has the **normalp** package available, which can be used for the Laplace distribution; see [Mineo \(2003\)](#).

2.2. The inverse Gaussian type distribution

If a RV T follows the IGTD with mean μ , scale parameter λ and kernel g , then the notation $T \sim \text{IGT}(\mu, \lambda; g)$ is used and the following results hold:

(A1) The PDF of T is

$$f_T(t) = f_Z(a_t) \frac{\sqrt{\lambda}}{\sqrt{t^3}} = c g(a_t^2) \frac{\sqrt{\lambda}}{\sqrt{t^3}}; \quad t > 0, \mu > 0, \lambda > 0,$$

where $a_t = a_t(\mu, \lambda) = \sqrt{\lambda/\mu}[\sqrt{t/\mu} - \sqrt{\mu/t}]$ and $f_Z(\cdot)$ denotes the PDF of $Z \sim \text{S}(g)$

(A2) The mode(s) of T denoted by t_m is (are) obtained by the solution(s) of

$$\omega_g(a_{t_m}^2) \frac{\lambda}{\mu} \left[\frac{1}{\mu} - \frac{\mu}{t_m^2} \right] = \frac{3}{2t_m},$$

where a_t is given in (A1) and $\omega_g(u) = g'(u)/g(u)$, with $u > 0$ and $g'(\cdot)$ being the derivative of $g(\cdot)$ also given in (A1).

(A3) $cT \sim \text{IGT}(c\mu, c\lambda; g)$, with $c > 0$, i.e., the IGTD belongs to the scale family.

(A4) The CDF of T is

$$F_T(t) = F_Z(a_t) + \int_{b_t}^{\infty} c g(u^2 - 4\lambda/\mu) du,$$

where $b_t = \sqrt{\lambda/\mu}[\sqrt{t/\mu} + \sqrt{\mu/t}]$ and $F_Z(\cdot)$ denotes the CDF of $Z \sim \text{S}(g)$.

(A5) The quantile function (qf) of T must be numerically obtained by $t(q) = F_T^{-1}(q)$, with $F_T^{-1}(\cdot)$ being the inverse function of $F_T(\cdot)$.

(A6) $U = Z^2 = [\lambda/\mu][T/\mu + \mu/T - 2]$ has a generalized chi-square distribution ($\text{G}\chi^2\text{D}$) with one degree of freedom, which is denoted by $U \sim \text{G}\chi^2(g)$; see [Fang et al. \(1990, p. 91\)](#).

Distribution of Z	Distribution of U
Laplace	$U = V^2$, where $V \sim \text{Gamma}(1, 1)$
Logistic	$U = [\log(\frac{1-V}{V})]^2$, where $V \sim \text{Beta}(1, 1)$
Normal	$U \sim \chi^2(1)$
Student- t	$U \sim \mathcal{F}(1, \nu)$

Table 1: Distribution of $U = Z^2 \sim \text{G}\chi^2(g)$, where $Z \sim \text{S}(g)$.

(A7) If all the moments of T exist, then $\mathbb{E}[T^{-r}] = \mathbb{E}[T^{r+1}]/\mu^{2r+1}$, for $r \in \{0, 1, \dots\}$. Thus, the expected value of T is given by $\mathbb{E}[T] = \mu$. In general, the $(r+1)$ th moment of T is

$$\mathbb{E}[T^{r+1}] = \frac{\mu^{r+1}}{\theta^r} v_r - \sum_{k=1}^r \binom{2r}{k} \frac{\mathbb{E}[T^{r+1-k}]}{\mu^{-k} [-1]^k} - \sum_{k=r+1}^{2r} \binom{2r}{k} \frac{\mathbb{E}[T^{k-r}]}{\mu^{k-2r-1} [-1]^k}; \quad r = 1, 2, \dots,$$

where $\theta = \lambda/\mu$ and $v_r = \mathbb{E}[U^r] < \infty$, with $U \sim \text{G}\chi^2(g)$. From the expression for $\mathbb{E}[T^{r+1}]$ given below, the variance, standard deviation (SD), coefficient of variation (CV), coefficient of skewness (CS) and coefficient of kurtosis (CK) of T can be obtained.

The IGTD may be generated from the Laplace, logistic, normal and t with ν degrees of freedom (denoted by t_ν) kernels, which will be called IGT-lap, IGT-logis, IG and IGT- St distributions, respectively. The Cauchy law is a t distribution with $\nu = 1$ and the associated IGT will be called IGT-Cauchy. Table 1 presents $\text{G}\chi^2\text{D}$ associated with the Laplace, logistic, normal and t models. Specific expressions for the variance, CV, CS and CK for the IGT-lap, IGT-logis, IG and IGT- St distributions can be found in Sanhueza *et al.* (2008, Table 8).

2.3. Lifetime analysis for the IGTD

For a nonnegative RV T with PDF $f_T(t)$ (and CDF $F_T(t)$), lifetime distributions are equivalently characterized in terms of the survival function (sf) $S_T(t) = 1 - F_T(t)$ and hf (or failure rate) $h_T(t) = f_T(t)/S_T(t)$, with $t > 0$ and $0 < S_T(t) < 1$. For example, if $h_T(t)$ is nondecreasing (or nonincreasing) in t , then $F_T(\cdot)$ belongs to the class of increasing (or decreasing) failure rate (IFR) (or DRF) distributions; if $h_T(t) = \lambda > 0$, for all t , we have $S_T(t) = \exp(-\lambda t)$, with $t > 0$, so that $F_T(t)$ is an exponential distribution.

If $T \sim \text{IGT}(\mu, \lambda; g)$, then from (A1) and (A4), we have

$$h_T(t) = \frac{f_Z(a_t) \sqrt{\lambda}/\sqrt{t^3}}{F_Z(-a_t) - \int_{b_t}^{\infty} c g(u^2 - 4\lambda/\mu) du}; \quad t > 0$$

and

$$S_T(t) = F_Z(-a_t) - \int_{b_t}^{\infty} c g(u^2 - 4\lambda/\mu) du; \quad t > 0.$$

In some cases, instead of having a monotone hf, we have a change point, say t_c , such that $F_T(t)$ is IFR for $t \leq t_c$ and DFR for $t \geq t_c$. The IGTD belong to this switching monotone

FR patterns and its change point t_c is the solution of the equation

$$F_Z(-a_{t_c}) - \int_{b_{t_c}}^{\infty} f\left(\sqrt{u^2 - 4\lambda/\mu}\right) du = \frac{\sqrt{\lambda} f_Z(a_{t_c}) t_c^{-1/2}}{\frac{3}{2} - t_c \omega_g(a_{t_c}) \frac{d}{dt_c} a_{t_c}}.$$

2.4. Shape analysis for IGTD

Here, we perform a graphical analysis for some IGTDs. Models presenting different degrees of kurtosis, asymmetry as well as modality and bimodality and the absence of moments are considered. In the following, we present the densities of these IGTDs.

Example 1. Let $T \sim \text{IGT}(\mu, \lambda; g)$. Then, for the indicated IGTD, the PDF of T is

$$f_T(t) = \frac{1}{2} \exp\left(-\sqrt{\frac{\lambda}{\mu}} \left| \sqrt{\frac{t}{\mu}} - \sqrt{\frac{\mu}{t}} \right| \right) \frac{\sqrt{\lambda}}{\sqrt{t^3}}; \quad t > 0 \quad [\text{IGT-lap}].$$

$$f_T(t) = \frac{\exp\left(\sqrt{\frac{\lambda}{\mu}} \left[\sqrt{\frac{t}{\mu}} - \sqrt{\frac{\mu}{t}} \right] \right)}{\left[1 + \exp\left(\sqrt{\frac{\lambda}{\mu}} \left[\sqrt{\frac{t}{\mu}} - \sqrt{\frac{\mu}{t}} \right] \right) \right]^2} \frac{\sqrt{\lambda}}{\sqrt{t^3}}; \quad t > 0 \quad [\text{IGT-logis}].$$

$$f_T(t) = \frac{1}{\sqrt{2\pi}} \exp\left(-\frac{\lambda}{2\mu} \left[\frac{t}{\mu} + \frac{\mu}{t} - 2 \right] \right) \frac{\sqrt{\lambda}}{\sqrt{t^3}}; \quad t > 0 \quad [\text{IG}].$$

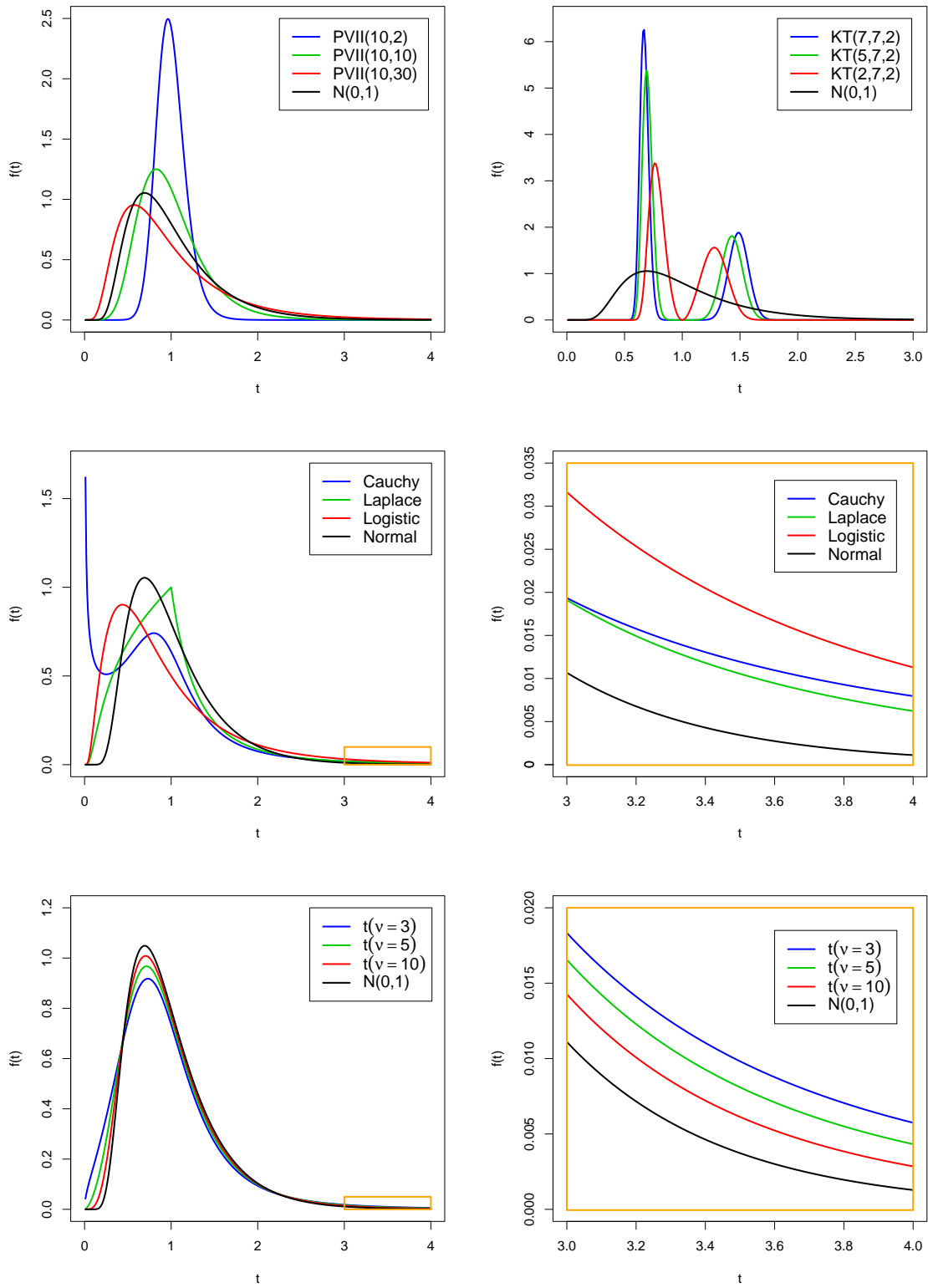
$$f_T(t) = \frac{\Gamma\left(\frac{\nu+1}{2}\right)}{\sqrt{\nu\pi}\Gamma\left(\frac{\nu}{2}\right)} \left[1 + \frac{\lambda}{\nu\mu} \left[\frac{t}{\mu} + \frac{\mu}{t} - 2 \right] \right]^{-\frac{\nu+1}{2}} \frac{\sqrt{\lambda}}{\sqrt{t^3}}; \quad t > 0, \nu > 0 \quad [\text{IGT-St}].$$

The IGTD can also be generated from the Kotz type (KT) and Pearson VII (PVII) kernels; for more details about the IGTD based on the KT and PVII kernels, see [Sanhueza *et al.* \(2008\)](#). By using the **ig** package, we can create [Figure 1](#), which shows the behavior of densities of IGTDs for several kernels. [Figure 1](#) also shows a zoom of the right tails for different IGTDs in order to highlight their kurtosis. From these graphical plots, we note that the IGT-lap, IGT-logis and IGT-St models present greater kurtosis than the classical IGD. Thus, any of these models should produce stable parameter estimates. However, the kurtosis of the IGT-St model is flexible depending on the parameter ν , which allows one to consider the IGT-St model as a good candidate in order to produce stable parameter estimates in the presence of outliers.

From [Section 2.3](#) and (A2), expressions for the lifetime indicators and the mode(s) of the IGTDs given in [Example 1](#) can be specified. In order to do that, we consider the values for $\omega_g(u)$ of the Laplace, logistic, normal and t distributions presented in [Sanhueza *et al.* \(2008, Table 3\)](#), which have been implemented in the **ig** package.

3. Statistical aspects for the IGTD

In this section, we present a summary of inference and diagnostics tools for the IGTD presented by [Sanhueza *et al.* \(2008\)](#), which have been implemented in the **ig** package. Also, we

Figure 1: IGT PDF plots for $\mu = 1$ and $\lambda = 4$ obtained from the indicated kernels.

provide new results useful for simulation studies from this distribution that have been implemented in this package as well. Specifically, in the first part of this section, we summarize some aspects related to estimation and inference for the parameters of the IGTD. Later, we present tools associated with influence diagnostics mainly based on local influence for this model. Finally, in the third part of this section, we derive a method for generating random numbers from the IGTD that has not been developed thus far, which allows one to conduct simulation studies.

3.1. Estimation and inference

The MLE of the parameter $\boldsymbol{\theta} = (\mu, \lambda)^\top$ for the IGTD, which is implemented in the **ig** package, can be computed from

$$\mu = \frac{\sum_{i=1}^n v_i t_i}{\sum_{i=1}^n v_i} \quad \text{and} \quad \lambda^{-1} = \frac{1}{n} \sum_{i=1}^n v_i \left[\frac{t_i}{\mu^2} + \frac{1}{t_i} - \frac{2}{\mu} \right], \quad (1)$$

where $v_i = v_i(\mu, \lambda) = -2\omega_g(a_{t_i}^2)$, for $i = 1, \dots, n$, with $\omega_g(\cdot)$ given in (A2). The expressions presented in Eq. (1) are obtained from the log-likelihood function associated with the model given in (A1), which is

$$\ell(\mu, \lambda) = \sum_{i=1}^n \ell_i(\mu, \lambda) \propto \frac{n}{2} \log(\lambda) + \sum_{i=1}^n \log(g(a_{t_i}^2)), \quad (2)$$

and making $\dot{L}_\mu = 0$ and $\dot{L}_\lambda = 0$, where $\dot{L}_\mu = \partial\ell(\mu, \lambda)/\partial\mu$ and $\dot{L}_\lambda = \partial\ell(\mu, \lambda)/\partial\lambda$.

The **ig** package incorporates the analytical form of the score vector, $\dot{\mathbf{L}} = (\dot{L}_\mu, \dot{L}_\lambda)^\top$. Numerical values for the MLEs of μ and λ must be computed by using iterative procedures which need initial values $\mu^{(0)}$ and $\lambda^{(0)}$. The MLEs of μ and λ for the classical IGD (see Eq. (1) with $v_i = 1$) are considered as starting values for the iterative procedure implemented in the **ig** package. For the classical IGD, we have $v_i = 1$ giving thus equal weight for each observation located at the tails or at the center of the distribution. In the IGT-*St* model, v_i gives less weight to the extreme cases, which can be viewed in Figure 2.

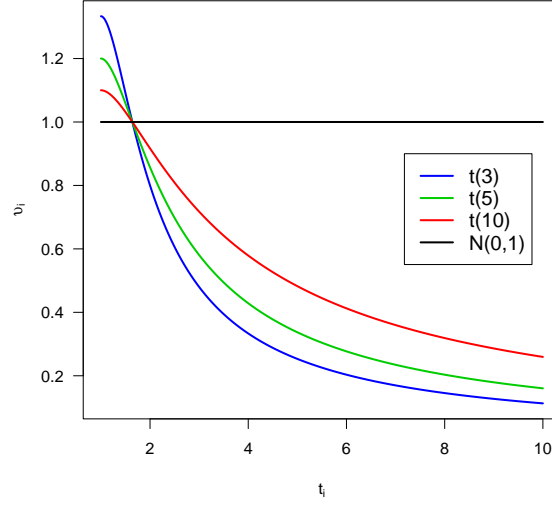
If the IGTD is obtained from a kernel different to the normal one, then $\mathbb{E}(\ddot{\mathbf{L}}_{\mu\lambda}) \neq 0$. Thus, an approximate $(1 - \alpha)100\%$ confidence region for $\boldsymbol{\theta}$ based on $\hat{\boldsymbol{\theta}} \sim N_2(\boldsymbol{\theta}, \boldsymbol{\Sigma}_\theta)$ and implemented in the **ig** package is given by

$$\mathcal{R} \equiv \left\{ \boldsymbol{\theta} \in \mathbb{R}^2 : (\hat{\boldsymbol{\theta}} - \boldsymbol{\theta})^\top \boldsymbol{\Sigma}_\theta^{-1} (\hat{\boldsymbol{\theta}} - \boldsymbol{\theta}) \leq \chi_{1-\alpha}^2(2) \right\},$$

where $\chi_{1-\alpha}^2(2)$ denotes the $(1 - \alpha)$ th quantile of the χ^2 distribution with two degrees of freedom. Here, $\boldsymbol{\Sigma}_\theta$ is the covariance matrix of $\hat{\boldsymbol{\theta}}$, which we approximate by $-\ddot{\mathbf{L}}^{-1}$ evaluated at $\hat{\boldsymbol{\theta}}$, with $-\ddot{\mathbf{L}}$ being the observed information matrix obtained from the Hessian matrix. This matrix is incorporated in an analytical form in the **ig** package and given by

$$\ddot{\mathbf{L}} = \begin{bmatrix} \ddot{L}_{\mu\mu} & \ddot{L}_{\mu\lambda} \\ \ddot{L}_{\lambda\mu} & \ddot{L}_{\lambda\lambda} \end{bmatrix} = \begin{bmatrix} \frac{\partial^2}{\partial\mu^2} \ell(\mu, \lambda) & \frac{\partial^2}{\partial\mu\partial\lambda} \ell(\mu, \lambda) \\ \frac{\partial^2}{\partial\lambda\partial\mu} \ell(\mu, \lambda) & \frac{\partial^2}{\partial\lambda^2} \ell(\mu, \lambda) \end{bmatrix}.$$

For details about the score vector, Hessian matrix and asymptotic inference for the IGTD, see Sanhueza *et al.* (2008).

Figure 2: Plots of the v_i against t_i .

3.2. Influence diagnostics

Diagnostics techniques are relevant due to the importance for evaluating the effect of possible atypical observations on the estimates. Also, these techniques allow one to verify the stability of the estimation procedure in the presence of those atypical data. Next, we summarize a diagnostics technique known as local influence, which is implemented in the **ig** package.

Let $\hat{\boldsymbol{\theta}}_{\boldsymbol{\omega}}$ be the MLE of $\boldsymbol{\theta}$ in the perturbed likelihood function $\ell(\mu, \lambda | \boldsymbol{\omega}) = \sum_{i=1}^n \omega_i \ell_i(\mu, \lambda)$, where $\boldsymbol{\omega}$ is a weight vector. The influence of the perturbation $\boldsymbol{\omega}$ on the MLE can be evaluated by $\text{LD}(\boldsymbol{\omega}) = 2[\ell(\hat{\boldsymbol{\theta}}) - \ell(\hat{\boldsymbol{\theta}}_{\boldsymbol{\omega}})]$ by studying its local behavior around the non-perturbed vector $\boldsymbol{\omega}_0$, using the normal curvature C_1 of $\text{LD}(\boldsymbol{\omega})$ in $\boldsymbol{\omega}_0$ in the direction of some unitary vector $\mathbf{1}$, such that $C_1 = 2|\mathbf{1}^\top \boldsymbol{\Delta}^\top \check{\mathbf{L}}^{-1} \boldsymbol{\Delta} \mathbf{1}|$, with $\|\mathbf{1}\| = 1$ and $\check{\mathbf{L}}$ and $\boldsymbol{\Delta}$ evaluated at $\boldsymbol{\theta} = \hat{\boldsymbol{\theta}}$ and $\boldsymbol{\omega}_0$. Here, $\boldsymbol{\Delta}$ is a $2 \times n$ matrix of perturbations given by $\boldsymbol{\Delta} = [\boldsymbol{\Delta}_1(\boldsymbol{\theta}), \dots, \boldsymbol{\Delta}_n(\boldsymbol{\theta})]$, with elements

$$\boldsymbol{\Delta}_i(\boldsymbol{\theta}) = [\Delta_i(\mu), \Delta_i(\lambda)]^\top = \left[\frac{\partial^2}{\partial \mu \partial \omega_i} \ell(\mu, \lambda | \boldsymbol{\omega}), \frac{\partial^2}{\partial \lambda \partial \omega_i} \ell(\mu, \lambda | \boldsymbol{\omega}) \right]^\top; \quad i = 1, \dots, n.$$

From the log-likelihood function $\ell(\lambda, \mu)$ of the IGTD defined in Eq. (2), we have

$$\Delta_i(\mu) = \frac{\lambda}{\mu^2} \left[\frac{v_i t_i}{\mu} - v_i \right] \quad \text{and} \quad \Delta_i(\lambda) = \frac{1}{2\lambda} - \frac{1}{2} \left[\frac{v_i t_i}{\mu^2} + \frac{v_i}{t_i} - \frac{2v_i}{\mu} \right]; \quad i = 1, \dots, n. \quad (3)$$

The most influential cases can be identified by their large components of the direction vector, \mathbf{l}_{\max} , which produces the greatest local change in $\hat{\boldsymbol{\theta}}$ and corresponds to the eigenvector associated with the largest eigenvalue of $\mathbf{B} = \boldsymbol{\Delta}^\top \check{\mathbf{L}}^{-1} \boldsymbol{\Delta}$. In this case, the total local influence of the i th case is given by $C_i = 2|b_{ii}|$, where b_{ii} is the i th diagonal element of \mathbf{B} . Those cases that satisfy the condition $C_i > \frac{2}{n} \sum_{i=1}^n C_i$ can be considered as potentially influential; see Verbeke and Molenberghs (2000). We have implemented C_i and its cut point as local influence diagnostics tool in the **ig** package. For more details about diagnostics techniques in the IGTD see Sanhueza *et al.* (2008).

3.3. Simulation

In general, statistical inference tools may not exist in closed form for the IGTD. Hence, simulation and numerical studies are needed, which require a random number (RN) generator. Next, we present a RN generator for the IGTD following a procedure similar to the one given in [Chhikara and Folks \(1989, pp. 52–53\)](#). This generator has been implemented in the **ig** package. Thus, if $T \sim \text{IGT}(\mu, \lambda; g)$, the algorithm steps are:

(B1) Generate a RN u from $U \sim G\chi^2(g)$ by using an appropriate generator.

(B2) Generate a RN w from $W \sim U(0, 1)$.

(B3) Set values for μ and λ and then compute a random number $t = t_1$ or $t = t_2$ from $T \sim \text{IGT}(\mu, \lambda; g)$ according to the following criterion:

(B3.1) If $w \leq p_0$, then

$$t_1 = \mu + \frac{u\mu^2}{2\lambda} - \frac{\mu}{2\lambda} \sqrt{4\mu\lambda u + \mu^2 u^2}; \quad p_0 = \frac{\mu}{\mu + t_1};$$

(B3.2) If $w > p_0$, then $t_2 = \mu^2/t_1$.

4. The IG functions

In the **ig** package, the logistic, Laplace, normal and t kernels for the IGTD are implemented. This package provides mainly two groups of functions in R code for the IGTD. These groups are related to: (i) probabilistic characters, lifetime indicators and a random number generator, and (ii) estimation, inference, diagnostics and goodness-of-fit methods. In addition, some tools useful for exploratory data analysis (EDA) are also included. The **ig** package is independent of other R packages already created.

4.1. IG probabilistic functions

In order to compute the PDF, CDF and qf given in (A1), (A4) and (A5), the commands `digt()`, `pigt()` and `qigt()` are used, respectively. The following instructions illustrate these functions.

```
R> digt(3, mu = 1.0, lambda = 1.0, nu = 1.0, kernel = "normal")
```

```
[1] 0.03941836
```

```
R> digt(3, mu = 1.0, lambda = 1.0, nu = 1.0, kernel = "normal")
```

```
[1] 0.03941836
```

```
R> pigt(0.6758413, mu = 1.0, lambda = 1.0, nu = 1.0, kernel = "normal")
```

```
[1] 0.5
```

```
R> qigt(0.5, mu = 1.0, lambda = 1.0, nu = 1.0, kernel = "normal")
```

```
[1] 0.6758413
```

Figure 1 (third panel on the left) was plotted with the function `digt()` using the following instructions:

```
R> t <- seq(0, 4, by = 0.01)
```

```
R> ft1 <- digt(t, 1.0, 4.0, 1.0, "t")
R> ft2 <- digt(t, 1.0, 4.0, 1.0, "laplace")
R> ft3 <- digt(t, 1.0, 4.0, 1.0, "logistic")
R> ft4 <- digt(t, 1.0, 4.0, 1.0, "normal")
R> ft <- cbind(ft1, ft2, ft3, ft4)
```

```
R> matplot(t, ft, type = "l", xlim = c(0, 4), ylim = c(0, 1.7), xlab = "t",
+   ylab = "f(t)", col = c(4, 3, 2, 1), lty = c(1, 1, 1, 1),
+   lwd = c(2, 2, 2, 2))
R> rect(3, -0.00005, 4, 0.1, border = "orange", lwd = 2)
```

```
R> legend(2.6, 1.7, c("Cauchy", "Laplace", "Logistic", "Normal"),
+   col = c(4, 3, 2, 1), lty = c(1, 1, 1, 1), lwd = c(2, 2, 2, 2),
+   cex = 1.25)
```

In order to carry out a lifetime analysis, we have implemented the functions `sfigt()` and `hfigt()` for computing the survival and hazard functions, respectively.

To generate data from the IGTD, we have considered the method given in (B1) to (B3) by implementing the function `rigt()`. The following instruction illustrates this command:

```
R> rigt(n = 6, mu = 1.0, lambda = 1.0, nu = 1.0, kernel = "normal")
```

```
[1] 0.2559006 0.2829651 3.9604401 0.9365816 1.8822376 0.9176596
```

4.2. IG statistical functions

Another group of functions related to graphical analysis, estimation, diagnostics, simulation and goodness of fit for the IGTD is also available. In order to estimate the mean (μ) and scale (λ) parameters of the IGTD, we implemented the method summarized in Section 3.1. Four functions for estimating the parameters of the IG, IGT-lap, IGT-logis and IGT-St distributions have been developed. These functions are: `mleig()`, `mleigt()`, `mleigStNuFixed()` and `mleigSt()`. The parameters of the IGD based on a data set call `lifetime`, for example, may be estimated from `mleig(lifetime)` or `mleigt(lifetime, kernel = "normal")`. The parameters of the IGT-lap and IGT-logis models based on `lifetime` may be estimated from `mleig(lifetime, kernel = "laplace")` and `mleig(lifetime, kernel = "logistic")`, respectively. The parameters of the IGT-St model based on `lifetime` may be estimated from `mleigStNuFixed(lifetime, nu = 1.0)`, if the parameter ν is fixed, which for example

is useful for conducting simulation studies. However, if the purpose is to achieve the optimal value for ν , then the function `mleigt(lifetime)` must be used.

Next, two examples related to the use of these and other commands are presented. The first example corresponds to a simulation study carried out by using the functions `rigt()`, `mleig()` and `mleigStNuFixed()`. The second example is from a real data set available in the literature and implemented in the **ig** package. In this example, we illustrate that the estimation procedure presented for the IGT-*St* model is stable in the presence of atypical observations. Also, we present the functions implemented for EDA in the **ig** package.

Example 2. Simulation study

In this example, we study the behavior of the MLEs of the parameters of the IGTD by considering the Monte Carlo simulation. We simulate samples under different scenarios considering small, moderate and large sample sizes ($n = 10, 25$ and 100 , respectively) and normal and t_ν kernels, with $\nu = 2, 8$ and 50 , where the values $\nu = 2$ corresponds to high kurtosis, $\nu = 8$ to moderate kurtosis and $\nu = 50$ to low kurtosis. The scale parameter was fixed at $\lambda = 1.0$, without a loss of generality, and the mean was fixed at $\mu = 0.5$.

We use the bias and mean square error (MSE) of the MLE in order to study the quality of the estimation method. The samples were generated from an IGTD with a specific kernel (normal or t_ν) that we called the “true kernel” and the estimation of parameters computed using samples obtained from the same or other kernels, called the “assumed kernel”. The empirical bias and MSE values are the average of simulated samples for each combination of sample size and kernel. The results of the simulations are presented in Tables 2 and 3 for the estimates of μ and λ , respectively.

We have noted that the bias of the estimators for μ and λ gets smaller when the sample size tends to be greater for those cases where the true and assumed kernels are the same. In addition, we find greater bias when we estimate λ instead of μ . On the other hand, We have also noted that the MSE gets smaller when the sample size gets bigger and when the kurtosis increases as well. Furthermore, we can see that there are greater MSEs when we estimate λ instead of μ .

Example 3. Application to practical data

Birnbaum and Saunders (1969) reported a data set corresponding to fatigue life (T) measured in cycles ($\times 10^{-3}$) of $n = 101$ aluminum coupons (specimens) of type 6061-T6. These specimens cut parallel to the direction of rolling and oscillating at 18 cycles per second. The coupons were exposed to a pressure with maximum stress of 31,000 psi (pounds per square inch). All specimens were tested until failure. The **ig** package incorporates these lifetime data so that if we input the command `data("psi31")` it will be ready to use the data.

Exploratory data analysis: In order to carry out an EDA for `psi31`, we implement the function `descriptiveSummary()` and some graphical tools. Firstly, descriptive statistics of the data can be obtained from the function `descriptiveSummary()`, which uses the command `searchMode()` to find the empirical mode of the data. This descriptive statistics is summarized in Table 4 and obtained by the instruction

```
R> descriptiveSummary(psi31)
```

Secondly, the instruction

n	Assumed kernel	True kernel							
		t_2		t_8		t_{50}		N(0, 1)	
		Bias	MSE	Bias	MSE	Bias	MSE	Bias	MSE
10	t_2	-0.0261	0.0307	-0.0186	0.0380	-0.0229	0.0787	-0.0638	0.2273
	t_8	-0.0273	0.0217	-0.0224	0.0192	-0.0233	0.0363	-0.0266	0.0366
	t_{50}	-0.0142	0.0157	-0.0063	0.0129	-0.0016	0.0123	-0.0004	0.0122
	N(0, 1)	-0.0070	0.0154	0.0002	0.0118	0.0043	0.0113	0.0054	0.0112
25	t_2	-0.0139	0.0095	-0.0222	0.0320	-0.0060	0.0215	-0.1177	0.3720
	t_8	-0.0012	0.0063	-0.0003	0.0056	0.0018	0.0056	0.0027	0.0057
	t_{50}	0.0023	0.0057	0.0044	0.0047	0.0057	0.0046	0.0061	0.0046
	N(0, 1)	-0.0025	0.0064	-0.0103	0.0239	-0.0011	0.0051	-0.0009	0.0051
100	t_2	-0.0028	0.0021	-0.0056	0.0029	-0.0362	0.0674	-0.2931	0.8756
	t_8	-0.0024	0.0018	-0.0017	0.0016	-0.0011	0.0017	-0.0009	0.0017
	t_{50}	0.0006	0.0015	0.0001	0.0012	0.0000	0.0012	0.0000	0.0012
	N(0, 1)	-0.0016	0.0016	-0.0008	0.0013	-0.0004	0.0012	-0.0002	0.0012

Table 2: Biases and MSEs of the estimate of μ based on $\mu = 0.5$ and $\lambda = 1.0$.

n	Assumed kernel	True kernel							
		t_2		t_8		t_{50}		N(0, 1)	
		Bias	MSE	Bias	MSE	Bias	MSE	Bias	MSE
10	t_2	-0.6762	2.7509	0.2368	0.5056	0.4138	0.4998	0.4083	0.4927
	t_8	-1.5647	7.6427	-0.4652	1.2197	-0.2302	0.6787	-0.1891	0.6195
	t_{50}	-1.7448	7.0752	-0.6142	1.4566	-0.3701	0.9391	-0.3285	0.8713
	N(0, 1)	-1.8926	12.2263	-0.6648	1.6477	-0.4135	0.9832	-0.3708	0.8995
25	t_2	-0.2215	0.4229	0.4293	0.2792	0.6319	0.4661	0.6140	0.4867
	t_8	-0.9173	1.4039	-0.1522	0.1849	0.0610	0.1195	0.1011	0.1204
	t_{50}	-1.1716	2.0469	-0.3590	0.3071	-0.1422	0.1438	-0.1015	0.1263
	N(0, 1)	-1.1437	1.9066	-0.3676	0.3163	-0.1600	0.1454	-0.1204	0.1262
100	t_2	-0.0432	0.0538	-0.5124	0.2780	0.7268	0.5390	0.6407	0.5828
	t_8	-0.6707	0.5400	-0.0348	0.0323	-0.1695	0.0508	0.2139	0.0671
	t_{50}	-0.8567	0.8453	-0.2022	0.0728	-0.0051	0.0206	0.0347	0.0202
	N(0, 1)	-0.9189	0.9669	-0.2597	0.1051	-0.0644	0.0289	-0.0255	0.0234

Table 3: Biases and MSEs of the estimate of λ based on $\mu = 0.5$ and $\lambda = 1.0$.

Mean	Median	Mode	SD	CV	CS	CK	Range	Min.	Max.	n
133.73	133	142	22.36	16.72%	0.33	0.97	142	70	212	101

Table 4: Descriptive statistics for lifetimes of aluminum specimens.

```
R> histigt(psi31,
+   mainTitle = "",
+   xLabel = "Lifetimes with maximum stress of 31.000 psi",
+   kernel = "t",
+   colourHistogram = 4,
+   colourBoxPlot = 4)
```

simultaneously produces a box-plot and a histogram for `psi31`. Figure 3 shows these plots.

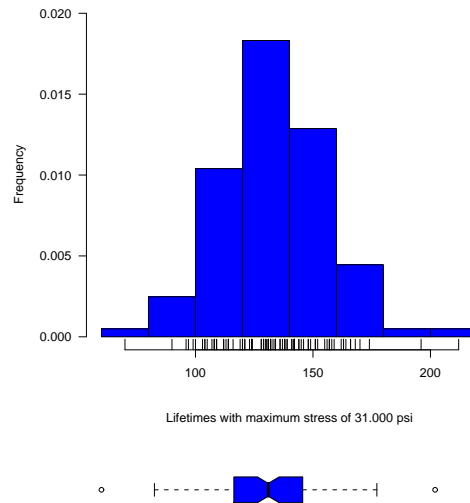


Figure 3: Histogram and box-plot for `psi31`.

The box-plot may be suppressed by the instruction `boxplot = FALSE`.

The EDA based on Table 4 and Figure 3 indicates a slightly positively skewed distribution ($CS = 0.33$) with moderate kurtosis ($CK = 0.97$) and some atypical values (see box-plot). The IGT-D considers the degrees of skewness and kurtosis presented in the data. The potential influence of the atypical cases will be analyzed next by the IGT-*St* model. This IGT-D should produce stable parameter estimates to the potentially influential cases.

Estimation: Next, we find the MLEs of the parameters μ and λ of the IGT-*St* distribution. Due to problems of unbounded and local maximum in the likelihood function, which can be revised in Lange *et al.* (1989) and Sanhueza *et al.* (2008), it is better to fix ν and assume that it is a known value or otherwise get information for it from the data. Then, in order to compute the estimates of μ and λ of the IGT-*St* model, the function `mleigtSt()` has been implemented by using the procedure described in Section 3.1 by means of the following algorithm:

(C1) For $\nu = 1$ to $\nu = 100$:

(C1.1) Estimate the parameters μ and λ of the IGT-*St* model considering the MLEs of μ and λ in the classical IGD as starting values for the numerical procedure.

(C1.2) Compute the likelihood function.

(C2) Choose the value of ν that maximizes the likelihood function and consider the values of the estimates of μ and λ as MLEs of these parameters.

Thus, the instruction `mleigtSt(psi31)` automatically chooses the value of ν that maximizes the likelihood and computes the MLE for μ and λ of the IGT-*St* model according to (C1) and (C2). These results can be saved in the variable `estimate` as

```
R> estimates <- mleigtSt(psi31)
```

obtaining

```
R> estimates
```

```
$muEstimate
[1] 134.3009
```

```
$lambdaEstimate
[1] 6348.45
```

```
$nuOptimal
[1] 7
```

```
$logLikelihood
[1] 455.4684
```

The values of the estimates of μ and λ and the optimal value of ν can be saved in the variables `mu`, `lambda` and `nu`, respectively, as

```
R> mu      <- estimates$muEstimate
R> lambda <- estimates$lambdaEstimate
R> nu      <- estimates$nuOptimal
```

Additionally, the function `loglikigt()` provides the value of the log-likelihood function.

Figure 4 shows values for the log-likelihood function against values for the parameter ν in the range $[1, 100]$. This curve shows that the log-likelihood is maximized at $\nu = 7$. This graph was plotted by means of the instructions:

```
R> nus          <- vector("numeric")
R> resultsLogLik <- vector("numeric")

R> for(i in 1:100) {
+   nu <- nus[i]
+   a <- loglikigt(t, nu, "t")
+   resultsLogLik[i] <- a
+ }

R> plot(resultsLogLik, type = "p", main = "", xlab = expression(nu),
+   ylab = "Log-likelihood", pch = 20, col = 3, lwd = 1.5, axes = FALSE)
R> box()
R> axis(1, at = c(1, 20, 40, 60, 80, 100), labels = c(1, 20, 40, 60, 80, 100))
R> axis(1, at = 7, labels = 7)
R> axis(2, at = seq(-468, -456, by = 2), labels = seq(-468, -456, by = 2),
+   las = 1)
R> points(7, resultsLogLik[7], pch = 4, lwd = 2.0, col = 2)
```

Diagnostics: Now, an influence diagnostics based on the total local influence technique given in Section 3.2 is carried out. This analysis is presented on a graphical plot, which is produced by the function `diagnosticsigt()`. Thus, the instructions:

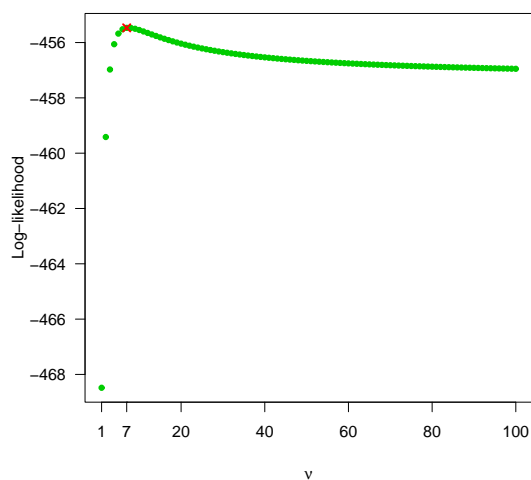
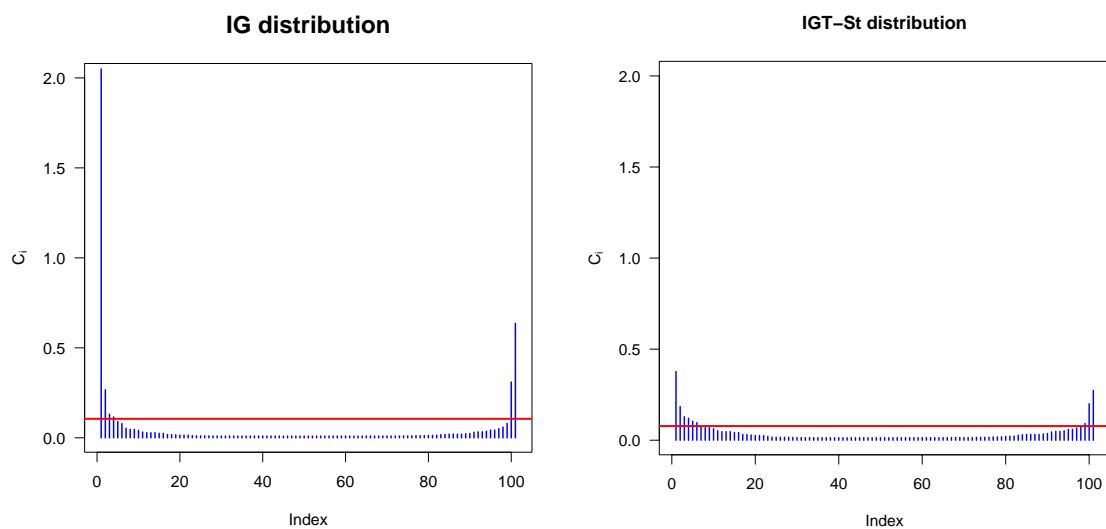
Figure 4: Log-likelihood for the IGT- St model against ν .

Figure 5: Total local influence index plot for the indicated models.

```
R> diagnosticsigt(psi31, "normal")
R> title(main = "IG distribution")
```

and

```
R> diagnosticsigt(psi31, "t", yRange = c(0, 2))
R> title(main = "IGT-St distribution")
```

produce the graphs shown in Figure 5, which demonstrate that the estimation procedure

Removed case(s)	IGT-St($\nu = 7$)		IG	
	$\hat{\mu}$	$\hat{\lambda}$	$\hat{\mu}$	$\hat{\lambda}$
{1}	0.05	13.62	0.48	16.17
{2}	0.20	1.57	0.33	4.68
{100}	0.37	1.08	0.47	4.31
{101}	0.42	0.17	0.59	6.96
{1, 2, 100, 101}	0.64	2.37	0.25	42.11

Table 5: RC (in %) for the indicated parameters and models.

presented for the IGT-St model with $\nu = 7$ is stable in the presence of atypical data. This figure shows potentially influential observations only in the classical IGD. These observations are the cases: 1, 2, 100 and 101.

In order to compute the magnitude of the changes produced on the MLE of μ and λ when the potentially influential observations are eliminated, we implement the function `rcigt()`. This function computes the relative change (RC), in percentage, of each estimated parameter, defined by $RC_{\theta_j} = |(\hat{\theta}_j - \hat{\theta}_{j(i)})/\hat{\theta}_j| \times 100\%$, where $\hat{\theta}_{j(i)}$ denotes the MLE of θ_j after the set I of cases has been removed. Table 5 shows these RCs, which were obtained by using the commands

```
rcigt(psi31, casesRemoved = 1, kernel = "t")
rcigt(psi31, casesRemoved = 1, kernel = "normal")
```

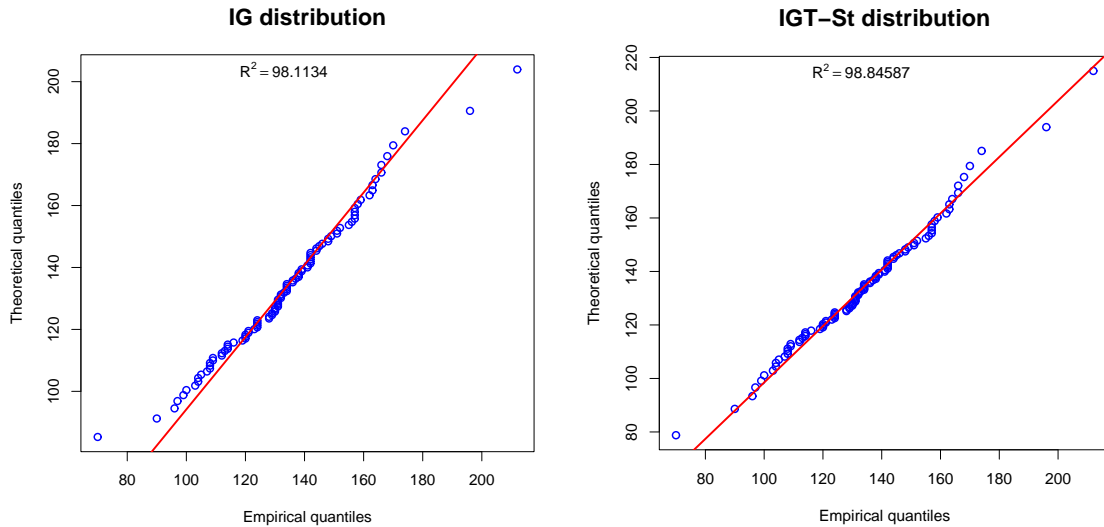
Analogously, the instructions `casesRemoved = 2` and `casesRemoved = c(1, 2, 100, 101)` allow one to eliminate the case 2 and the cases {1, 2, 100, 101}, respectively. From Table 5, we note that the RCs are greater in the classical IGD than in the IGT-St model with $\nu = 7$, as we expected.

Goodness of fit: In order to select the best model, we implement three goodness-of-fit tools for the IGT-D: (i) probability-probability (PP) and quantile-quantile (QQ) plots; (ii) the Kolmogorov-Smirnov (KS) test; and (iii) a criteria of model selection. Next, the goodness-of-fit functions implemented in the `ig` package will be described. Firstly, the commands `ppigt()` and `qqigt()` enables to detect the fitting of the IGT-D to the data by means of the visualization of the PP and QQ plots. The commands `ppigt()` and `qqigt()` allow also incorporate a straight line that sketches a line passing through the first and third quartile of the theoretical IGT-D in order to check graphically if a set of observations follows a particular IGT-D. Also, the coefficients of determination (R-square in %) of these plots are calculated and shown inside these graphs. The following instructions give QQ plots for `psi31`:

```
R> qqigt(psi31, kernel = "t", line = TRUE)
R> title(main = "IGT-St distribution", cex.main = 1.5)
```

```
$coefficientofdetermination
[1] 98.84587
```

```
R> qqigt(psi31, kernel = "normal", line = TRUE)
R> title(main = "IG distribution", cex.main = 1.5)
```


Figure 6: QQ plots of the indicated models for `psi31`.

```
$coefficientofdetermination
```

```
[1] 98.1134
```

while Figure 6 displays these graphs. Once again, from this figure, we note that the estimation procedure presented for the IGT-*St* model with $\nu = 7$ is stable on the tails of this distribution. This is because the IGT-*St* model behaves better than the classical IGD on the tails and then produces stable parameter estimates to outliers. The R^2 of the PP plot for `psi31` by considering the IGT-*St* distribution with $\nu = 7$ is 99.46% (see Table 6), which demonstrates the model's excellent fit to the data (for the QQ plot, the R^2 is 98.85% upon this same model).

Secondly, we also implement a command for a goodness-of-fit test useful for the IGTD. This function is `ksigt()`, which calculates the KS test and also provides a comparative graph of the theoretical and empirical distribution functions. This graph is shown in Figure 8 (on the right) and is produced when the instruction `graph = TRUE` is added in `ksigt()`. Finally, we implemented the Schwartz information criterion (SIC) by the function `sicigt()`, which is a well known criterion of model selection that can also be used for goodness of fit.

Next, the results of the fit for `psi31` are presented in Table 6, which show that the IGT-*St* distribution with $\nu = 7$ presents a better fit to the data.

Inference and invariance: For inferential purposes, we have implemented the function `aciigt()`, which computes an approximate $(1 - \alpha) \times 100\%$ confidence region for the parameters μ and λ

Distribution	R^2 (in %)	SIC	KS (p -value)
IG	99.16	4.573	0.0851 ($p = 0.4571$)
IGT- <i>St</i> ($\nu = 7$)	99.46	4.555	0.0602 ($p = 0.8576$)

Table 6: PP R^2 , SIC and KS test for `psi31`.

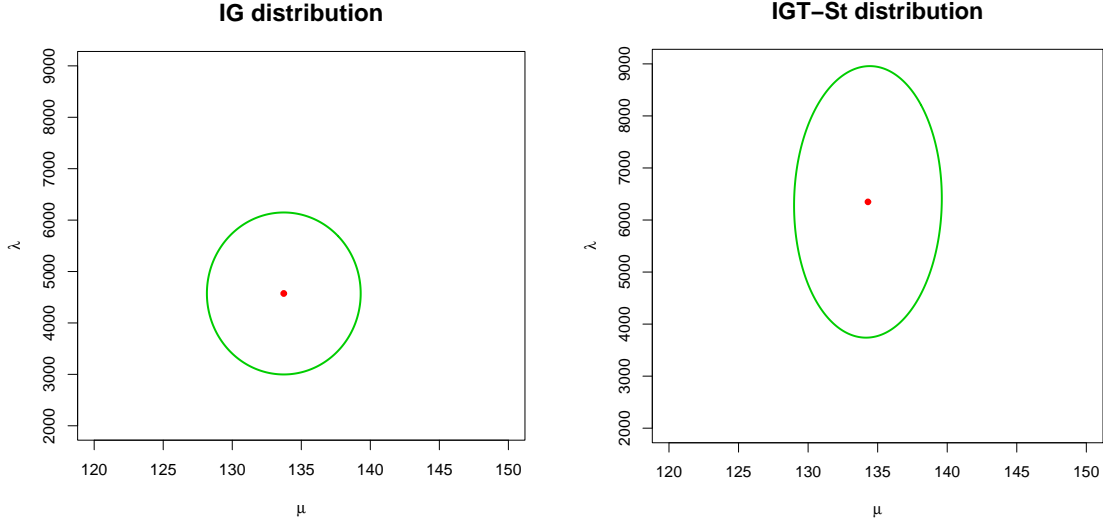


Figure 7: Approximate confidence regions for the indicated models.

of the IGTD. Also, this function provides a graph of this region, which is reported by means of simultaneous approximate confidence intervals (ACI) in the form: $ACI_{(1-\alpha_1)100\%}(\mu) = [\underline{\mu}, \bar{\mu}]$ and $ACI_{(1-\alpha_2)100\%}(\lambda) = [\underline{\lambda}, \bar{\lambda}]$, where $\alpha_1 + \alpha_2 = \alpha$. We note that these ACIs can only be computed if the sample size is large. In our example, we have computed ACI for μ and λ of the IGD and IGT-St distribution with $\nu = 7$ using of the instructions:

```
R> aciigt(psi31, kernel = "normal", confLevel = 95,
+   chart = c(120, 150, 2000, 9000), colourRegion = 3,
+   colourEstimates = 2)
R> title(main = "IG distribution", cex.main = 1.5)

R> aciigt(psi31, kernel = "t", confLevel = 95,
+   chart = c(120, 150, 2000, 9000), colourRegion = 3,
+   colourEstimates = 2)
R> title(main = "IGT-St distribution", cex.main = 1.5)
```

The ACIs of the parameters of the IGTD based on `psi31` are presented in Table 7.

In addition, by using the invariance property of the MLEs, the estimated PDF may be sketched on the histogram adding the instruction `densityLine = TRUE` into the function `histigt()`. Figure 8 (on the left) shows the histogram with the estimated PDF from which we observe

Distribution	$\underline{\mu}$	$\hat{\mu}$	$\bar{\mu}$	$\underline{\lambda}$	$\hat{\lambda}$	$\bar{\lambda}$
IGT- $t(\nu = 7)$	128.99	134.30	139.61	3739.74	6348.45	8957.16
IG	128.16	133.73	139.30	2998.28	4573.36	6148.45

Table 7: Estimates and $ACI_{95\%}$ for μ and λ of the indicated distributions.

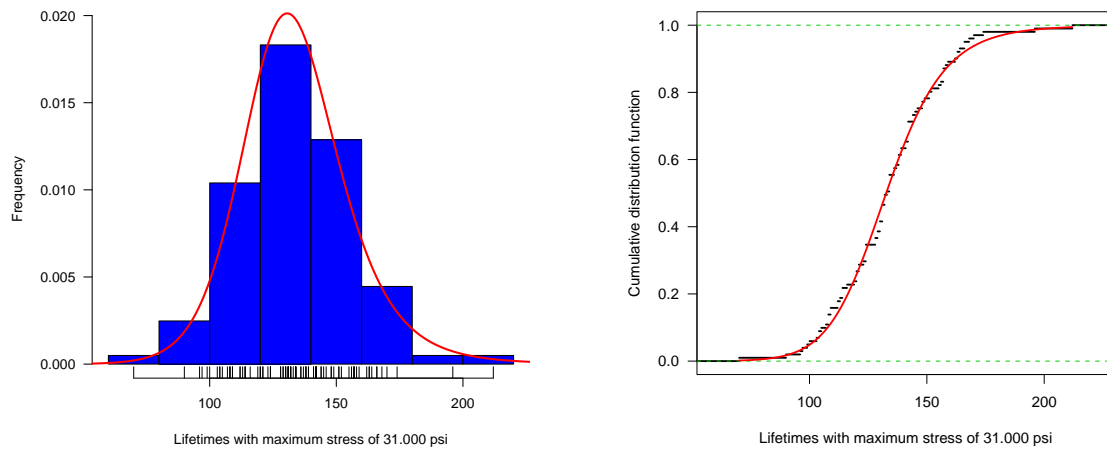


Figure 8: Histogram with estimated PDF for `psi31` (left side) and empirical CDF against theoretical CDF of the IGT-*St* with $\nu = 7$ (right side).

that the proposed model fits the data very well. This figure was produced by the instructions

```
R> histigt(psi31,
+ kernel      = "t",
+ boxPlot     = FALSE,
+ densityLine = TRUE,
+ mainTitle   = "",
+ xLabel      = "Lifetimes with maximum stress of 31.000 psi",
+ yLabel      = "Frequency",
+ yRange      = c(0, 0.02),
+ colourHistogram = 4,
+ colourDensity   = 2,
+ colourBoxPlot   = 4)
```

Remark: Seven other data sets called `psi21`, `psi26`, `precipitations`, `repairtimes`, `shelflife`, `fracture` and `runoff` used frequently in the literature of this topic, have also been incorporated into the **ig** package.

5. Concluding remarks

In order to analyze data from the IGTD, we have developed the **ig** package, which can execute several useful commands. The created functions are related to probability and lifetime indicators, estimation, diagnostics, goodness-of-fit, simulation and graphical tools. In addition, we have provided new theoretical results of the IGTD related to lifetime analysis and generation of random numbers. Furthermore, we have conducted a brief simulation study for evaluating

the estimation method used here. By using numerical examples, we illustrate the different capabilities and features of the **ig** package. We believe that, in the future, the **ig** package can be improved by incorporating other functions related to the estimation for censored data and regression methods.

Acknowledgments

The authors wish to thank the editor and referees for their helpful comments that aided in significantly improving this article. This study was partially supported by FONDECYT 1080326 and DIUFRO DI08-0061 grants, Chile.

References

- Birnbaum ZW, Saunders SC (1969). “Estimation for a Family of Life Distributions.” *Journal of Applied Probability*, **6**, 319–327.
- Chhikara RS, Folks JL (1989). *The Inverse Gaussian Distribution*. Marcel Dekker, New York.
- Fang K, Kotz S, Ng K (1990). *Symmetric Multivariate and Related Distributions*. Chapman & Hall, London.
- Johnson N, Kotz S, Balakrishnan N (1994). *Continuous Univariate Distributions*, volume 1. John Wiley & Sons, New York.
- Lange K, Little L, Taylor M (1989). “Robust Statistical Modelling Using the t Distribution.” *Journal of the American Statistical Association*, **84**, 881–896.
- Mineo A (2003). “A New Package for the General Error Distribution.” *R News*, **3**(2), 13–16. URL <http://CRAN.R-project.org/doc/Rnews/>.
- Mudholkar G, Natarajan R (2002). “The Inverse Gaussian Models: Analogues of Symmetry, Skewness and Kurtosis.” *Annals of the Institute of Statistical Mathematics*, **54**, 138–154.
- R Development Core Team (2008). *R: A Language and Environment for Statistical Computing*. R Foundation for Statistical Computing, Vienna, Austria. ISBN 3-900051-07-0, URL <http://www.R-project.org/>.
- Sanhueza A, Leiva V, Balakrishnan N (2008). “A New Class of Inverse Gaussian Type Distributions.” *Metrika*. doi:10.1007/s00184-007-0141-8. Forthcoming.
- Schrödinger E (1915). “Zur Theorie der Fall-und Steigversuche und Teilchen mit Brownscher Bewegung.” *Physikalische Zeitschrift*, **16**, 289–295.
- Seshadri V (1993). *The Inverse Gaussian Distribution: A Case Study in Exponential Families*. Claredon, New York.
- Tweedie MCK (1957). “Statistical Properties of the Inverse Gaussian Distribution I.” *Annals of Mathematical Statistics*, **28**, 362–377.

Verbeke G, Molenberghs G (2000). *Linear Mixed Models for Longitudinal Data*. Springer-Verlag, New York.

Affiliation:

Víctor Leiva, Hugo Hernández

Departamento de Estadística

Universidad de Valparaíso

Casilla 5030, Valparaíso, Chile

E-mail: victor.leiva@uv.cl/victor.leiva@yahoo.com, hugo.hernandez.p@gmail.com

URL: <http://www.deuv.cl/leiva/>

Antonio Sanhueza

Departamento de Matemática y Estadística

Universidad de La Frontera

Temuco, Chile

E-mail: asanhueza@ufro.cl