Testing a parameter restriction on the boundary for the g-and-h distribution: a simulated approach

Noname manuscript No.
(will be inserted by the editor)

Testing a parameter restriction on the boundary for the g-and-h distribution: a simulated approach

Marco Bee · Julien Hambuckers · Flavio Santi · Luca Trapin

Received: date / Accepted: date

Abstract We develop a likelihood-ratio test for discriminating between the g-and-h and the g distribution, which is a special case of the former obtained when the parameter h is equal to zero. The g distribution is a shifted lognormal, and is therefore suitable for modeling economic and financial quantities. The g-and-h is a more flexible distribution, capable of fitting highly skewed and/or leptokurtic data, but is computationally much more demanding. Accordingly, in practical applications the test is a valuable tool for resolving the tractability-flexibility trade-off between the two distributions. Since the classical result for the asymptotic distribution of the test is not valid in this setup, we derive the null distribution via simulation. Further Monte Carlo experiments allow us to estimate the power function and to perform a comparison with a similar test proposed by Xu and Genton (2015). Finally, the practical relevance of the test is illustrated by two risk management applications dealing with operational and actuarial losses.

Keywords Likelihood ratio · Skewness · Kurtosis · Value-at-Risk

The `R` codes used for the computations in this paper are available at `http://marcobee.weebly.com`.

M. Bee
Department of Economics and Management, University of Trento - Italy
Tel.: +39-0461-282296
Fax: +39-0461-282241
E-mail: marco.bee@unitn.it

J. Hambuckers
Department of Finance, HEC Liège, University of Liège, Belgium

F. Santi
Department of Economics, University of Verona - Italy

L. Trapin
Department of Statistical Sciences "P. Fortunati", University of Bologna - Italy

1 Introduction

Many data-sets in economics and finance are characterized by large skewness and/or kurtosis, and should be modeled by distributions that are able to account for these features (Kleiber and Kotz, 2003; McDonald et al., 2013). The g-and-h distribution (Tukey, 1977) is a flexible model defined by the following non-linear transformation of the standard normal distribution:

$$X = a + b\frac{e^{gZ}-1}{g}e^{\frac{hZ^2}{2}}, \quad Z \sim N(0,1). \tag{1}$$

The parameters g and h are related to skewness and kurtosis, so that the theoretical skewness-kurtosis range of the g-and-h distribution is larger with respect to other commonly used distributions (Dutta and Perry, 2006, Fig. 3).

The g-and-h has been mostly employed for operational risk measurement and insurance modeling (Dutta and Babbel, 2002; Peters and Sisson, 2006; Degen et al., 2007; Fischer et al., 2007; Jiménez and Arunachalam, 2011; Cruz et al., 2015; Peters et al., 2016; Bee and Trapin, 2016; Bee et al., 2019a,b), but there are also applications to financial returns (Drovandi and Pettitt, 2011) and to wind speeds (Dupuis and Field, 2004). A multivariate version has been introduced by Field and Genton (2006); g-and-h random fields are proposed in Xu and Genton (2017). Related distributions are the g-and-k and the generalized g-and-h models (Rayner and MacGillivray, 2002; Prangle, 2017).

If $g = 0$, the g-and-h distribution becomes symmetric, whereas if $g \neq 0$ and $h = 0$ it is a special case of the three-parameter lognormal, also called g distribution (Cruz et al., 2015, Section 9.4.1). Even though the g distribution is less flexible than the g-and-h, basic results about the lognormal distribution guarantee that it can still model skewed and leptokurtic data. In particular, g governs the variance, the skewness and the kurtosis of the g distribution.

Figure 1 displays the skewness-kurtosis pairs obtained for $g \in \{0, 0.1, \ldots, 3\}$ and $h \in \{0, 1, 2\}$. Each curve corresponds to a different value of h, with the case $h = 0$ being the g distribution. Skewness and kurtosis are computed using the formulas for the first four moments of the g-and-h (Cruz et al., 2015, p. 320).

It is well known that the lognormal distribution plays a key role in economics and finance. From the theoretical point of view, various popular economic and financial models imply the lognormal distribution. For example, Gibrat (1931) postulated a process of proportional growth that gives rise to a lognormal distribution and has important applications in the study of city and firm size distribution. In quantitative finance, the lognormality of the marginal distribution of prices follows from the classical hypothesis that the continuous-time data generating process is a geometric Brownian motion. Finally, there is evidence that the income distribution of developed countries is, to a large extent, well described by a lognormal distribution; for example, Clementi and Gallegati (2005) find that approximately 97-99% of the distribution is lognormal, whereas the top 1-3% is Pareto.

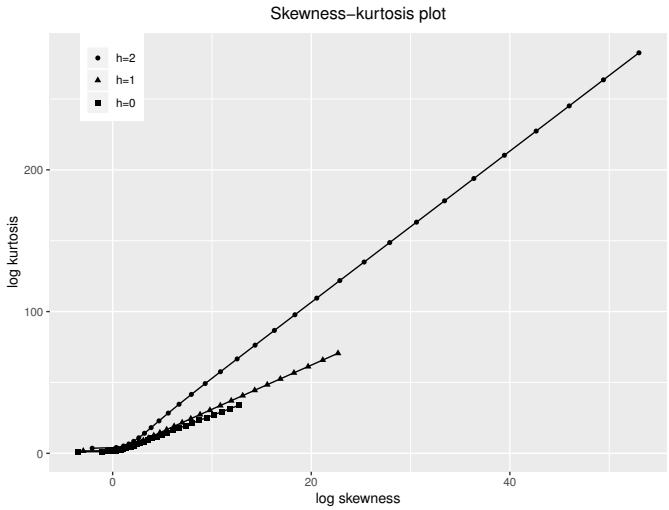

Fig. 1: Skewness-kurtosis plot (on doubly-logarithmic scale) for the g and the g-and-h distribution with $a = 0$ and $b = 1$. The points on each curve give the skewness-kurtosis pair corresponding to increasing values of $g \in \{0, 0.1, \ldots, 3\}$.

On the other hand, extensive empirical evidence suggests that in many setups the lognormal hypothesis is too restrictive, and one needs a more flexible model, such as the g-and-h distribution.

Not surprisingly, the capability of modeling a very large skewness-kurtosis range comes at a price. First, the introduction of an additional parameter may result in less stable parameter estimates. Second, the g-and-h density f_{gh} and distribution function F_{gh} are not known explicitly. Bee et al. (2019a) show that f_{gh} can be approximated numerically, but that approach requires the numerical inversion of the quantile function, so that the computational burden is rather heavy. Hence, with the principle of parsimony in mind, it is important to assess whether the use of the g-and-h instead of the g distribution is justified for the data at hand. A test for $H_0 : h = 0$ in the g-and-h distribution allows one to: (i) establish a unified framework characterized by a smooth transition, as h moves away from zero, from the scaled lognormal (i.e., the g) to the more general g-and-h distribution; (ii) make a decision, via a theoretically sound testing approach, about the necessity of a model with an additional parameter.

The contribution of this paper is twofold. First, we develop a likelihood-ratio test for discriminating between the g-and-h and the g distribution. Given that the regularity conditions needed for the validity of the chi-squared asymptotic approximation do not hold in the present setup, we resort to the limiting results for the case of true parameter values at the boundary of the parameter space (Self and Liang, 1987; Xu and Genton, 2015). However, according to our Monte Carlo analysis, this asymptotic distribution seems to be not very precise for the sample sizes commonly employed in practice. Hence, we find

the null distribution via simulation. The power function of the test is studied via further Monte Carlo experiments. This approach is also compared to the testing methodology proposed by Xu and Genton (2015), which is based on a linear approximation of the likelihood function.

The g-and-h distribution is especially useful for fitting loss data and estimating the Value-at-Risk (VaR), which is the most popular measure of risk in finance and non-life insurance; see, e.g., McNeil et al. (2015) for a review. The accuracy of the estimated VaR crucially depends on both the flexibility and the estimation precision of the parametric model chosen to describe the empirical distribution. Our second contribution is connected to the loss modeling and VaR estimation perspective, since the test is a tool for resolving the trade-off between parsimony and accuracy related to the choice between the g and g-and-h distributions.

The test is applied to two real data-sets. We first fit some operational risk losses recorded at the Italian bank Unicredit between 2005 and 2014. The second application deals with automobile insurance claims. In both cases we illustrate the role of the test in selecting the most appropriate distribution for VaR estimation.

The rest of the paper is organized as follows. In Section 2 we review the g-and-h and g distributions and work out the details of maximum likelihood estimation (MLE) of the latter; in Section 3 we introduce the likelihood-ratio test of $H_0 : h = 0$; Section 4 focuses on the simulation experiments aimed at studying the null distribution and the power function of the test; in Section 5 we apply the test to the operational risk and automobile insurance datasets; finally, Section 6 concludes.

2 The g-and-h distribution

The g-and-h distribution is a *quantile distribution*. Hence, it can equivalently be defined via its quantile function, which is given by (Cruz et al., 2015, p. 318)

$$Q(p;\boldsymbol{\theta}) = a + b\frac{e^{gz_p}-1}{g}e^{\frac{hz_p^2}{2}}, \quad p \in (0,1), \tag{2}$$

where $\boldsymbol{\theta} = (g,h)' \in \mathbb{R}\times\mathbb{R}^+$ is the parameter vector and z_p is the p-quantile of the standard normal distribution. Without loss of generality, we set the location parameter a and the scale parameter b equal to 0 and 1 respectively (Degen et al., 2007; Cruz et al., 2015).

Numerical MLE for the g-and-h distribution has been developed by Bee et al. (2019b). We refer to that paper for details and simulation evidence, and recall here the implementation of the procedure.

Algorithm 1 *Given a random sample* $(x_1,\ldots,x_n) \overset{\text{iid}}{\sim} \text{gh}(0,1,g,h)$*, perform the following steps:*

1. *For each observation* x_i $(i = 1,\ldots,n)$*:*

   (a) *evaluate* $p_i = F_{\text{gh}}(x_i; \boldsymbol{\theta})$ *by numerical inversion of (2);*
   (b) *compute* $\hat{f}_{\text{gh}}(x_i) = \phi(p_i)/Q'(p_i; \boldsymbol{\theta})$*, where* $\phi(\cdot)$ *is the standard normal density and* $Q'(\cdot)$ *is the first derivative of the quantile function (2), which is known in closed form (Cruz et al., 2015, Eq. 9.33).*
2. *treat* $\hat{f}_{\text{gh}}$ *as the true density and maximize numerically with respect to* $\boldsymbol{\theta}$ *of the approximated log-likelihood function* $\hat{\ell}_{\text{gh}}(\boldsymbol{\theta}; x_1, \ldots, x_n) = \sum_{i=1}^{n} \log \hat{f}_{\text{gh}}(x_i; \boldsymbol{\theta})$.

The optimization is performed via the Broyden-Fletcher-Goldfarb-Shanno (BFGS) algorithm (Byrd et al., 1995), as implemented in the `optim R` function. Bee et al. (2019b) suggest to set the starting values $\boldsymbol{\theta}^{(0)} = (g^{(0)}, h^{(0)})'$ equal to the quantile estimators, but they are not guaranteed to belong to the parameter space. To overcome this drawback, and given the availability of the MLE $\hat{g}$ for the restricted model, we set $g^{(0)}$ equal to $\hat{g}$, whereas $h^{(0)}$ is still given by the corresponding quantile estimator $\hat{h}^q$, so that $\boldsymbol{\theta}^{(0)} = (\hat{g}, \hat{h}^q)'$.

Since the BFGS algorithm may not converge to the maximum (or may not converge at all), we consider two additional starting values $\boldsymbol{\theta}_1^{(0)} = (\hat{g} - 0.1, \hat{h}^q - 0.1)$ and $\boldsymbol{\theta}_2^{(0)} = (\hat{g} + 0.1, \hat{h}^q + 0.1)$; if any of them is outside the parameter space, we set it equal to the boundary value. If the maxima of the three log-likelihood functions are not identical, the MLEs are the estimated parameter values that correspond to the largest of the three maximized log-likelihoods.

2.1 The g distribution and its relation to the three-parameter lognormal

The three-parameter lognormal distribution $\text{Logn3}(\gamma, \mu, \sigma^2)$ has density

$$f(x; \gamma, \mu, \sigma^2) = \frac{1}{\sqrt{2\pi}\sigma(x-\gamma)} e^{-\frac{1}{2}\left(\frac{\log(x-\gamma)-\mu}{\sigma}\right)^2} \mathbb{1}_{\{x>\gamma\}}, \tag{3}$$

with $\mu, \gamma \in \mathbb{R}$ and $\sigma \in \mathbb{R}^+$. The g random variable $X \sim \text{g}(a, b, g)$ has the following stochastic representation:

$$X = a + b\frac{e^{gZ} - 1}{g} = a + b\left(\frac{e^{gZ}}{g} - \frac{1}{g}\right),$$

where $Z \sim N(0, 1)$, $a \in \mathbb{R}$, $b \in \mathbb{R}^+$ and $g \in \mathbb{R}^+$.

Suppose first that $a = 0$ and $b = 1$ (see Sect. 2.2 for the general case). Noting that $e^{gZ} \sim \text{Logn}(0, g^2)$ and using two well-known results about linear transformations of the lognormal distribution we readily conclude that

$$\frac{e^{gZ}}{g} \sim \text{Logn}\left(\log\frac{1}{g}, g^2\right)$$

and

$$\frac{e^{gZ}}{g} - \frac{1}{g} \sim \text{Logn3}\left(-\frac{1}{g}, \log\frac{1}{g}, g^2\right),$$

i.e a three-parameter lognormal with support $(-1/g, +\infty)$. It follows from (3) that the g density is

$$f(x;g) = \frac{1}{\sqrt{2\pi}g\left(x+\frac{1}{g}\right)} e^{-\frac{1}{2}\left(\frac{\log\left(x+\frac{1}{g}\right)-\log\left(\frac{1}{g}\right)}{g}\right)^2} \mathbb{1}_{\left\{x>-\frac{1}{g}\right\}}, \tag{4}$$

where $\mathbb{1}_{\{A\}}$ is the indicator function of the set A.

2.2 Maximum likelihood estimation

Given a random sample $(x_1, \ldots, x_n)$ from $\mathrm{g}(0, 1, g)$, using (4) we get the likelihood function of the g distribution:

$$L_{\mathrm{g}}(g; x_1, \ldots, x_n) = \prod_{i=1}^{n} \frac{1}{\sqrt{2\pi}g\left(x_i+\frac{1}{g}\right)} e^{-\frac{1}{2}\left(\frac{\log\left(x_i+\frac{1}{g}\right)-\log\left(\frac{1}{g}\right)}{g}\right)^2} \mathbb{1}_{\left\{g<-\frac{1}{x_i}\right\}}. \tag{5}$$

Clearly, (5) is equal to zero unless g is smaller than $-1/x_i$ for all the observations. Hence the log-likelihood $\ell_{\mathrm{g}}(g; x_1, \ldots, x_n) \stackrel{\text{def}}{=} \log L_{\mathrm{g}}(g; x_1, \ldots, x_n)$ is given by:

$$\ell_{\mathrm{g}}(g; x_1, \ldots, x_n) = \sum_{i=1}^{n} \left(-\log(2\pi) - \log(gx_i+1) - \frac{1}{2}\left(\frac{\log(gx_i+1)}{g}\right)^2 \right) \mathbb{1}_{\left\{0<g<-\frac{1}{x_{(1)}}\right\}},$$

where $x_{(1)} \stackrel{\text{def}}{=} \min_{1\le i\le n} x_i$. The first derivative of $\ell_{\mathrm{g}}(g; x_1, \ldots, x_n)$ with respect to g is equal to

$$\frac{d\ell_{\mathrm{g}}}{dg} = -\sum_{i=1}^{n} \left\{ \frac{x_i}{gx_i+1} - \sum_{i=1}^{n} \frac{\log(gx_i+1)}{g^2} \cdot \left[\frac{x_i}{gx_i+1} - \frac{\log(gx_i+1)}{g} \right] \right\} \mathbb{1}_{\left\{0<g<-\frac{1}{x_{(1)}}\right\}}. \tag{6}$$

Let $\hat{g}$ be the MLE of g and g_0 the root of (6). In principle, $\hat{g}$ is obtained by setting (6) equal to 0. However, three cases should be distinguished:

1. If there exists $g_0 \in (0, -1/x_{(1)}) : d\ell_{\mathrm{g}}/dg|_{g_0} = 0$, the MLE is $\hat{g} = g_0$;
2. If there exists no $g_0 \in (0, -1/x_{(1)}) : d\ell_{\mathrm{g}}/dg|_{g_0} = 0$, the MLE
   (a) is equal to $\hat{g} = -1/x_{(1)}$ if $d\ell_{\mathrm{g}}/dg|_{g_0} > 0$ for all $g < -1/x_{(1)}$;
   (b) is equal to to 0 if $d\ell_{\mathrm{g}}/dg|_{g_0} < 0$ for all $g < -1/x_{(1)}$.

The likelihood equation $d\ell_{\mathrm{g}}/dg = 0$ can be readily solved numerically by means of standard optimization routines.

It is easy to check that, if $a \neq 0, b \neq 1$, $X \sim \mathrm{g}(a, b, g)$, $W \sim \mathrm{Logn3}(-1/g, \log(1/g), g^2)$ and $Y = a + bW$, then $X \stackrel{\mathrm{d}}{=} Y$. In this case it is possible to first estimate a and b by means of the quantile estimation method (Hoaglin, 1985) and then estimate g via ML using the standardized observations $y^s \stackrel{\text{def}}{=} (y - \hat{a})/\hat{b}$.

A different technique is developed by Xu and Genton (2015): they propose a linear approximation of the likelihood function which is exploited to derive the estimators of the parameters. This method is computationally efficient, since it bypasses the numerical computation of the g-and-h distribution function. Moreover, Xu and Genton (2015) study the asymptotic distribution of the MLEs and of various likelihood ratio test statistics based on this approximation. In Section 4 we will use simulation to compare this method to our approach.

3 A test for the g-and-h versus the g distribution

Since the g distribution is nested in the g-and-h, it is natural to construct a log-likelihood ratio test of the null hypothesis $H_0 : h = 0$ versus the alternative $H_1 : h > 0$. The approximated g-and-h log-likelihood function is

$$\ell_{\mathrm{gh}}(g, h; x_1, \ldots, x_n) \stackrel{\mathrm{def}}{=} \log L_{\mathrm{gh}}(g, h; x_1, \ldots, x_n),$$

where $L_{\mathrm{gh}}(g, h; x_1, \ldots, x_n) = \prod_{i=1}^{n} \hat{f}_{\mathrm{gh}}(x_i; g, h)$; see Algorithm 1.

Using this notation, the log-likelihood ratio test is defined as

$$\begin{aligned} T_{gh} &\stackrel{\mathrm{def}}{=} -2\log(\lambda) = -2\log\left(\frac{\max_g L_{\mathrm{g}}(g; x_1, \ldots, x_n)}{\max_{g,h} L_{\mathrm{gh}}(g, h; x_1, \ldots, x_n)}\right) \\ &= 2\left(\max_{g,h} \ell_{\mathrm{gh}}(g, h; x_1, \ldots, x_n) - \max_{g} \ell_{\mathrm{g}}(g; x_1, \ldots, x_n)\right). \end{aligned} \tag{7}$$

According to the classical theory of likelihood ratio tests, under regularity conditions first derived by Cramér (1946), the asymptotic null distribution of (7) is χ^2_{p-q}, where p and q are the dimension of the unrestricted and restricted parameter space, respectively. Unfortunately, the condition that the true parameter value is not at the boundary of the parameter space is not satisfied in the present setup, so that we cannot rely on this limiting result.

For testing problems where the parameter of interest is at the boundary, the null distribution of the test statistics has been investigated by Self and Liang (1987). The current framework has one parameter of interest with true value at the boundary of the parameter space, and one nuisance parameter with true value not at the boundary, which is "Case 5" in the terminology of Self and Liang (1987). The corresponding asymptotic null distribution is a π:$(1-\pi)$ mixture of χ^2_0 and χ^2_1, with $\pi = 0.5$ (Self and Liang, 1987). A thorough discussion of this result for testing problems related to the g-and-h distribution can be found in Xu and Genton (2015).

However, there are two difficulties with this approach. First, the results require some regularity conditions about the first three derivatives of the log-likelihood function of the model (Self and Liang, 1987, p. 605), which are difficult to check in the present case. Moreover, for small sample sizes, the limiting approximation may be too crude. In particular, as noted by Xu and

Genton (2015, p. 84), the probability of h being estimated as zero is usually larger than 0.5. Hence, we simulate the null distribution of T_{gh}.

In the latter approach, the p-value results from the following steps:

Algorithm 2

1. *Given an iid random sample* $(x_1, \ldots, x_n)$*, compute the MLEs* $\hat{g}$ *and* $\hat{h}$ *of the g-and-h distribution;*
2. *For* $i = 1, \ldots, B$ *and* B *"large":*
   (a) *sample* n *observations* $(x_1^*, \ldots, x_n^*)$ *from* $\mathrm{g}(\hat{g})$*;*
   (b) *use* $x_1^*, \ldots, x_n^*$ *to estimate via MLE the parameters of both the g and the g-and-h distribution, compute* $\max_g \ell_{\mathrm{g}}(g; x_1^*, \ldots, x_n^*)$, $\max_{g,h} \ell_{\mathrm{gh}}(g, h; x_1^*, \ldots, x_n^*)$ *and the log-likelihood ratio test*

$$T_{gh,i}^* = 2(\max_{g,h} \ell_{\mathrm{gh}}(x_1^*, \ldots, x_n^*) - \max_{g} \ell_{\mathrm{g}}(x_1^*, \ldots, x_n^*));$$

3. *compute the p-value* $p_{\mathrm{sim}} \stackrel{\mathrm{def}}{=} \#\{T_{gh,i}^* > T_{gh}^{obs}\}/B$*, where* T_{gh}^{obs} *is the observed value of the test.*

4 Simulation experiments

In this section we carry out simulation experiments aimed at:

1. Studying the null distribution and the power function of the test;
2. Comparing the performance of our Monte Carlo (MC from now on) test and the test by Xu and Genton (2015) (XG from now on);
3. Assessing the impact of the choice of the distribution (g or g-and-h) on the estimated VaR.

4.1 Null distribution

To measure the goodness of the asymptotic approximation of the llr test, we sample n observations from the null distribution $\mathrm{g}(0, 1, g)$ with $g \in \{0.5, 1.5\}$ and compare selected quantiles of the simulated and theoretical null distributions, where the latter is a 50:50 mixture of χ_0^2 and χ_1^2. The two values of g used in the simulation are representative of instances with different skewness; the latter is also close to the values estimated in the applications of Sect. 5 and to the range found by Dutta and Perry (2006) in an extensive empirical investigation concerning operational losses. The experiments are performed with $n \in \{100, 500, 1000\}$ and the number of replications B is equal to 2000. The results are shown in Table 1.

Table 1 suggests that, for the sample sizes used in this experiment, the convergence towards the theoretical null distribution is rather slow, and strongly dependent on the true value of g. In the following we study the power function using simulation-based critical values. The estimated power of the test with the asymptotic critical values is reported in the online supplementary material.

Table 1: Selected quantiles of the theoretical and simulated null distributions of the MC test for $n \in \{100, 500, 1000\}$, $g \in \{0.5, 1.5\}$ and quantile levels $q \in \{0.9, 0.95, 0.99\}$.

		$q = 0.9$	$q = 0.95$	$q = 0.99$
Theoretical quantile		1.642	2.706	5.412
Simulated quantile, $g = 0.5$	$n = 100$	4.360	8.648	24.174
	$n = 500$	2.749	4.715	11.401
	$n = 1000$	2.219	4.079	8.267
Simulated quantile, $g = 1.5$	$n = 100$	1.691	5.655	29.411
	$n = 500$	0.110	2.054	6.072
	$n = 1000$	1.02×10^{-5}	1.049	5.700

Apart from sampling variability, the empirical size of the MC test is identical to the nominal size. Table 2 shows the empirical size $\text{ES}_{\alpha,\tilde{g},n}$ of the XG test, estimated with $B = 2000$ Monte Carlo replications. $\text{ES}_{\alpha,\tilde{g},n}$ is computed as follows.

1. For $i = 1, \ldots, B$:
   (a) simulate n observations from the g-and-h distribution with parameters $g = \tilde{g}$ and $h = 0$ (i.e., the g distribution with parameter $\tilde{g}$);
   (b) compute the XG test and record whether the null hypothesis is rejected using a nominal size equal to α;
2. Compute $\text{ES}_{\alpha,\tilde{g},n}$ as the fraction of rejections in the B replications of the experiment.

Analogously to the results obtained by Xu and Genton (2015), when $g = 0.5$ the empirical size of the XG test is smaller than the nominal size for all sample sizes. On the other hand, when $g = 1.5$, the empirical size is too large at the 5 and 1% level, too small at the 10% level.

Table 2: Empirical size of the XG test for $n \in \{100, 500, 1000\}$, $g \in \{0.5, 1.5\}$ and nominal levels $\alpha \in \{0.1, 0.05, 0.01\}$.

		$\alpha = 0.1$	$\alpha = 0.05$	$\alpha = 0.01$
$g = 0.5$	$n = 100$	0.042	0.021	0.007
	$n = 500$	0.039	0.017	0.003
	$n = 1000$	0.062	0.028	0.002
$g = 1.5$	$n = 100$	0.092	0.078	0.073
	$n = 500$	0.074	0.051	0.032
	$n = 1000$	0.076	0.055	0.032

All the simulation experiments have been performed via parallel computing techniques, using the parLapply function of the parallel R package on a i7@3.40 GHz processor with 8 cores. Average computing times are shown in Table 3. Not surprisingly, the XG test is much faster.

Table 3: Computing times per 1000 replications (in seconds) for the computation of the MC and XG test.

	MC	XG
$n = 100$	549	28
$n = 500$	3188	51
$n = 1000$	6284	100

4.2 Power function

The power function has to be estimated via simulation. Given $h \in \{0.005, 0.01, \ldots, 0.1\}$, for the MC test a pseudo-code description of the procedure is as follows:

Algorithm 3

- *For $i = 1, \ldots, B$, with $B = 1000$:*
  - *Sample n observations from the g-and-h distribution $gh(0, 1, g, h)$;*
  - *Compute the maximized likelihood under the null and alternative hypotheses;*
  - *Compute the test statistic $T_{gh}^{(i)}$.*
- *Compute the power $\#\{T_{gh}^{(i)} > T_{gh}^{*,\alpha}\}/B$, where $T_{gh}^{*,\alpha}$ is the $\alpha\%$ critical value computed via simulation (see Section 4.1).*

For the XG test, the procedure is identical, but the critical value is the asymptotic one.

Figures 2, 3 and 4 show the empirical power function for $g = 0.5$ and $n \in \{100, 500, 1000\}$. In each plot, panels (a) and (b) refer to the MC and XG test respectively.

Figure 2 shows that the MC test is more powerful for all values of h and at all significance levels when $n = 100$. For $n = 500$ the MC test is more powerful when $h < 0.015$, whereas the XG test has more power for the remaining values of h. Similarly, for $n = 1000$ the MC test is more powerful when $h = 0.005$ and less powerful for the remaining values of h. Moreover, the former test seems to have a better performance when α is larger, and the latter when α is smaller.

Figures 5, 6 and 7 show the empirical power function for $g = 1.5$. When $n = 100$ (Figure 5), MC is better for $\alpha \in \{0.05, 0.1\}$ and worse when $\alpha = 0.01$. When $n = 500$, MC is always better at all levels. When $n = 1000$, they are basically equivalent; only for the smallest values of h, and when $\alpha > 0.01$, XG is slightly better.

All in all, in terms of power there is no clear winner. However, it should be remembered that the MC test has the correct size, whereas we know from Table 2 that the empirical size of the XG test is significantly smaller than the nominal size.

The empirical size and the power function of the MC test have been computed also with respect to the asymptotic null distribution, i.e. by applying Algorithm 3 with $T_{gh}^{*,\alpha}$ equal to the $\alpha\%$ asymptotic, instead of simulated, critical value. The results are reported in tables A.1 and A.2 and in figures B.1

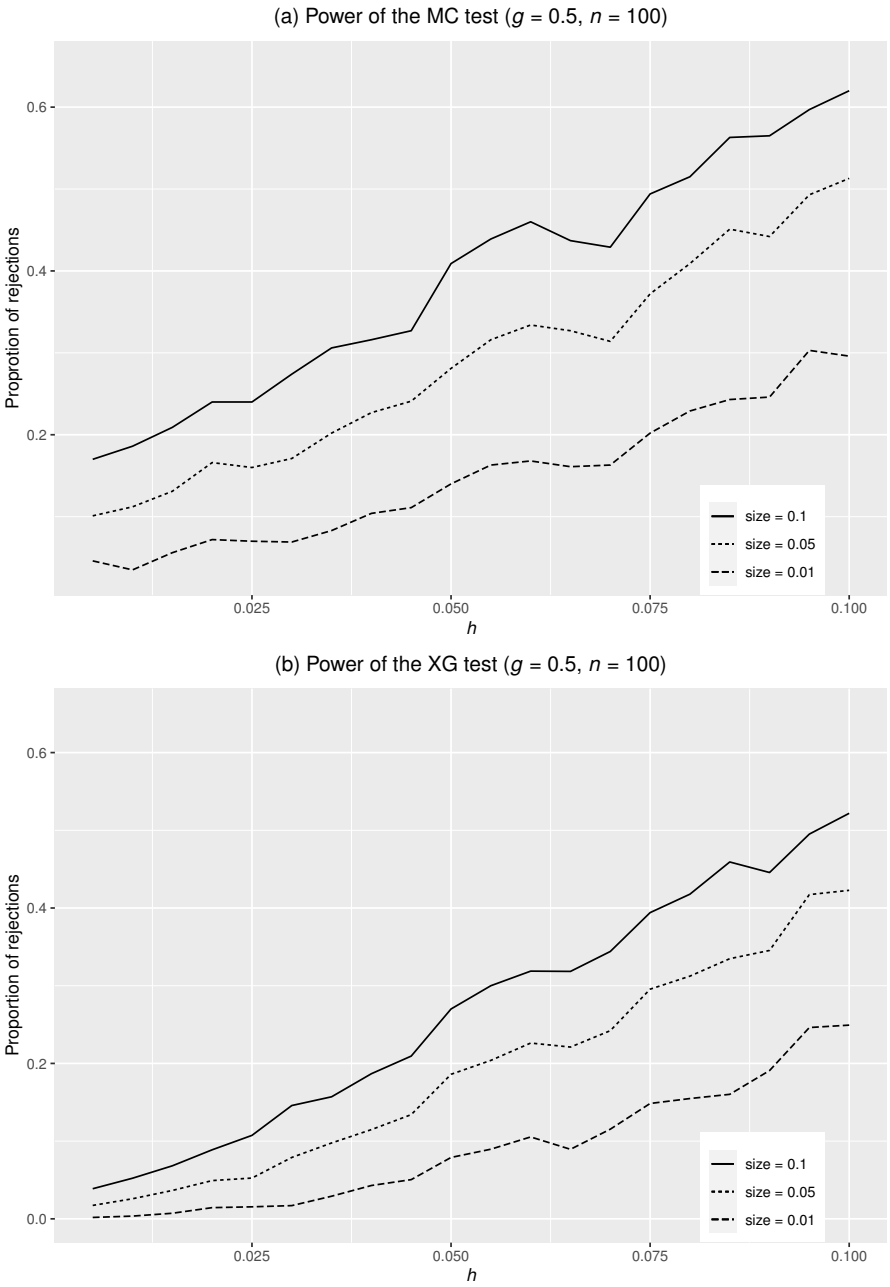

Fig. 2: Power function of the tests for $g = 0.5$ and $n = 100$.

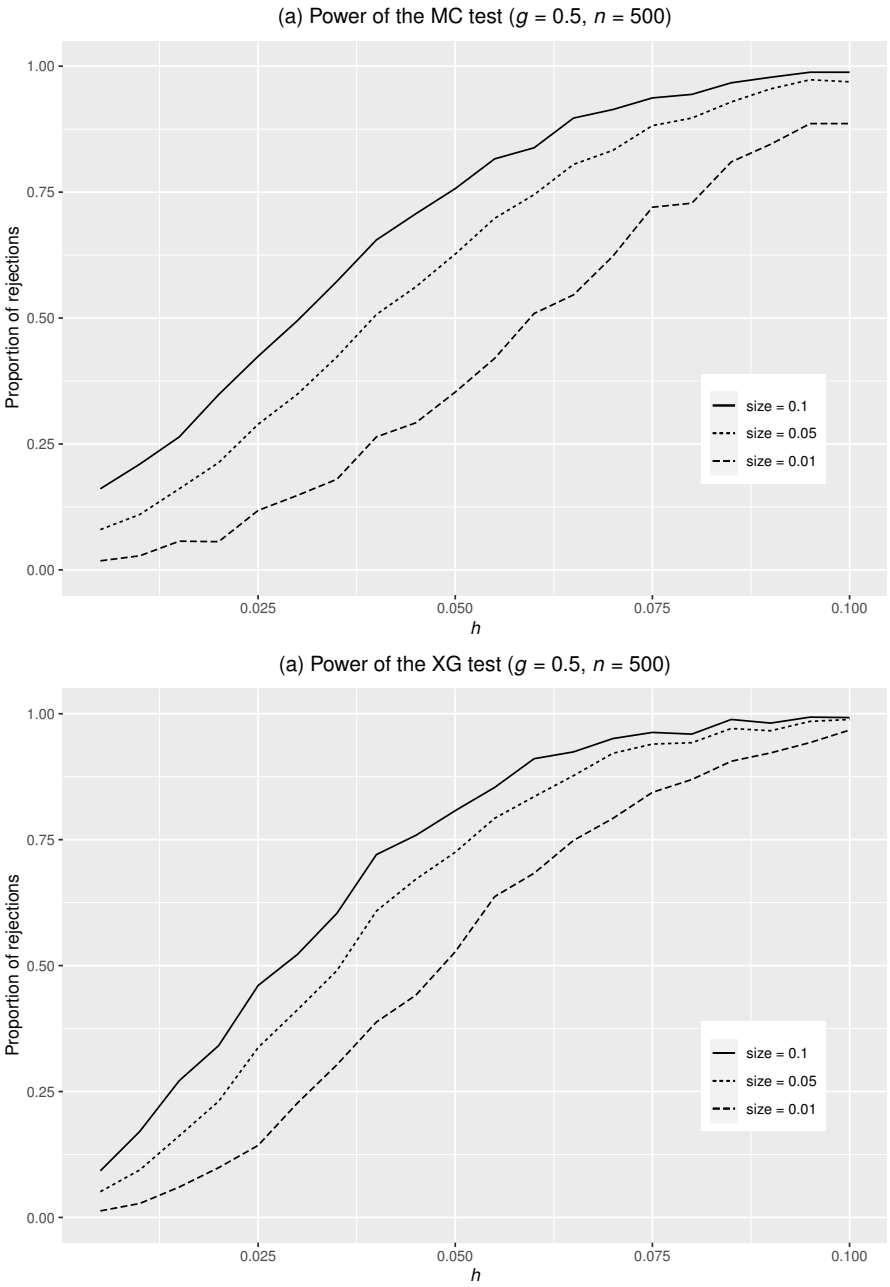

Fig. 3: Power function of the tests for $g = 0.5$ and $n = 500$.

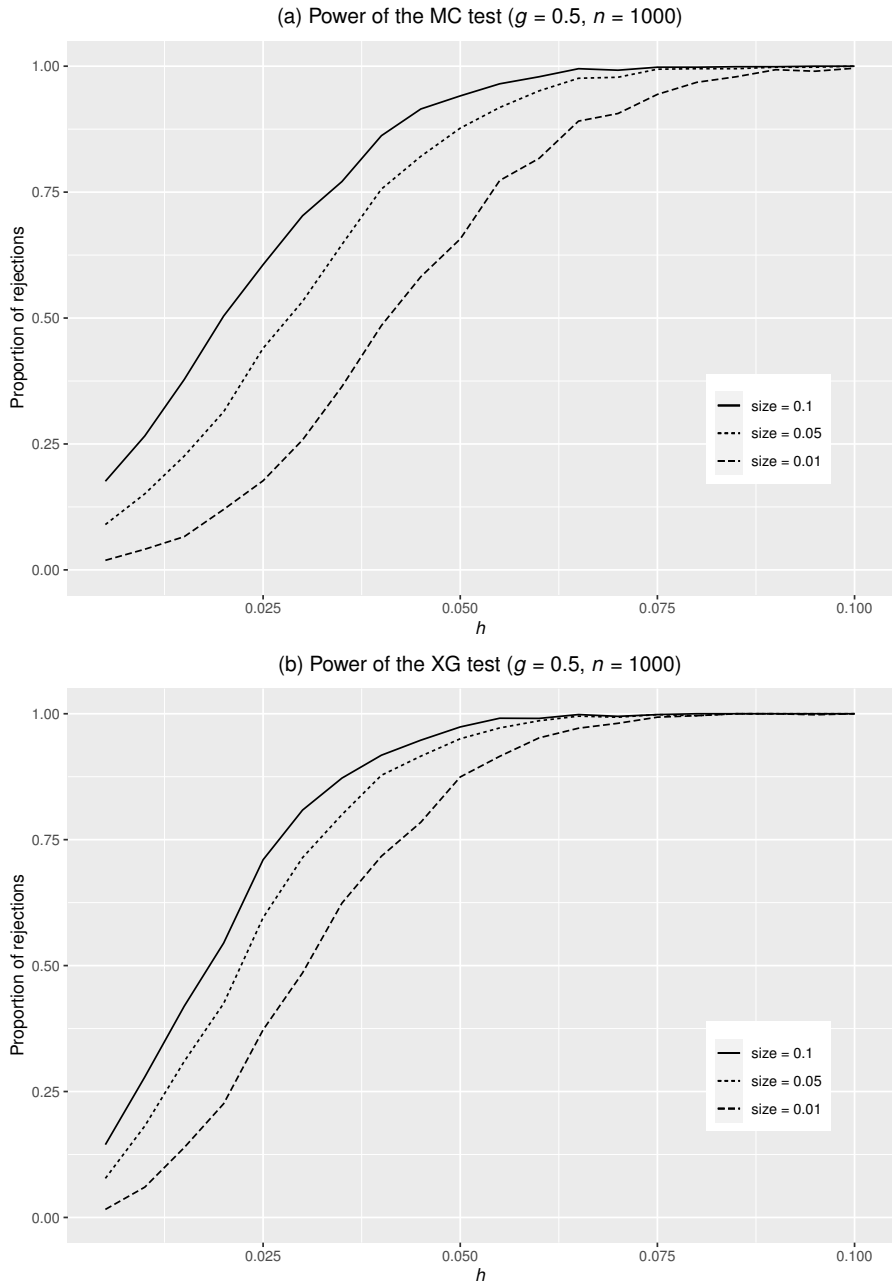

Fig. 4: Power function of the tests for $g = 0.5$ and $n = 1000$.

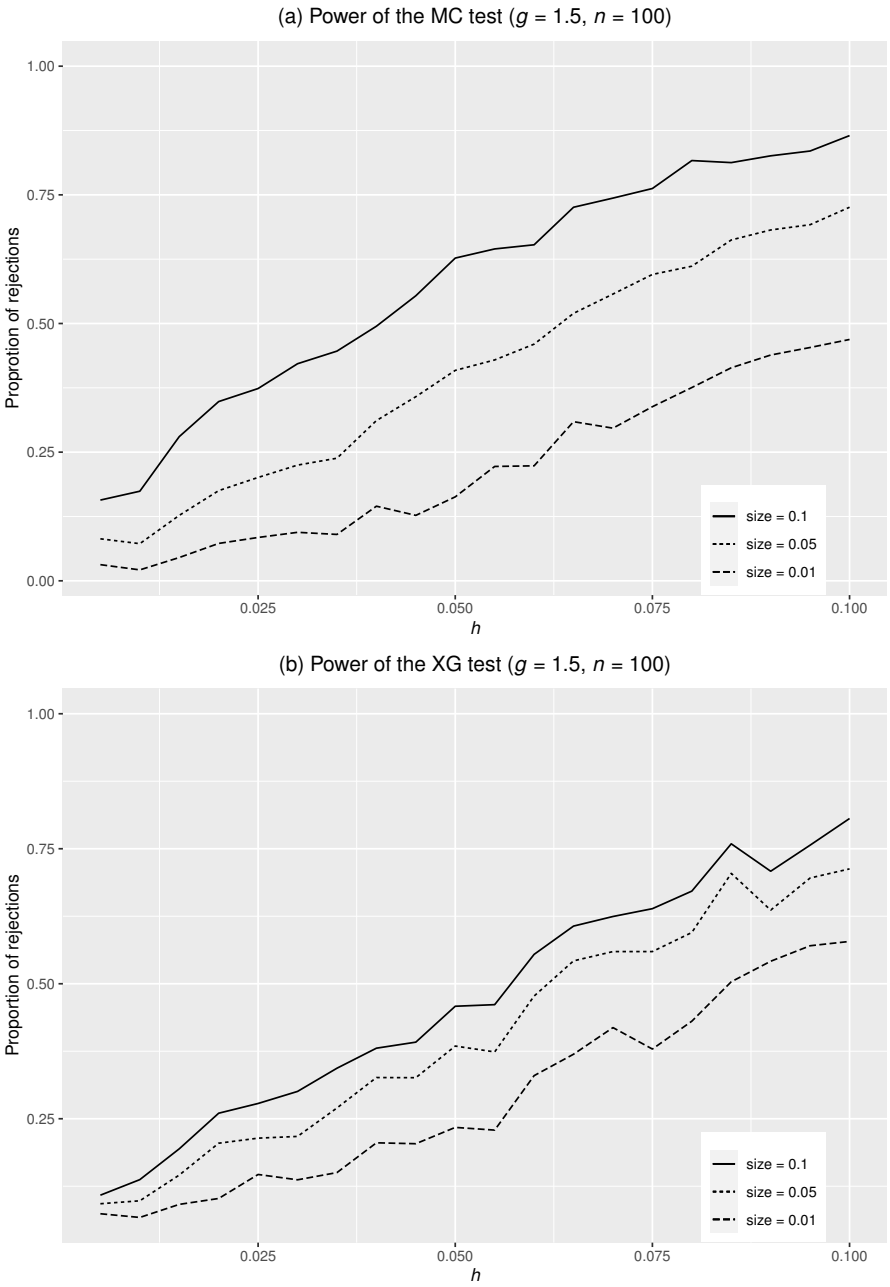

Fig. 5: Power function of the tests for $g = 1.5$ and $n = 100$.

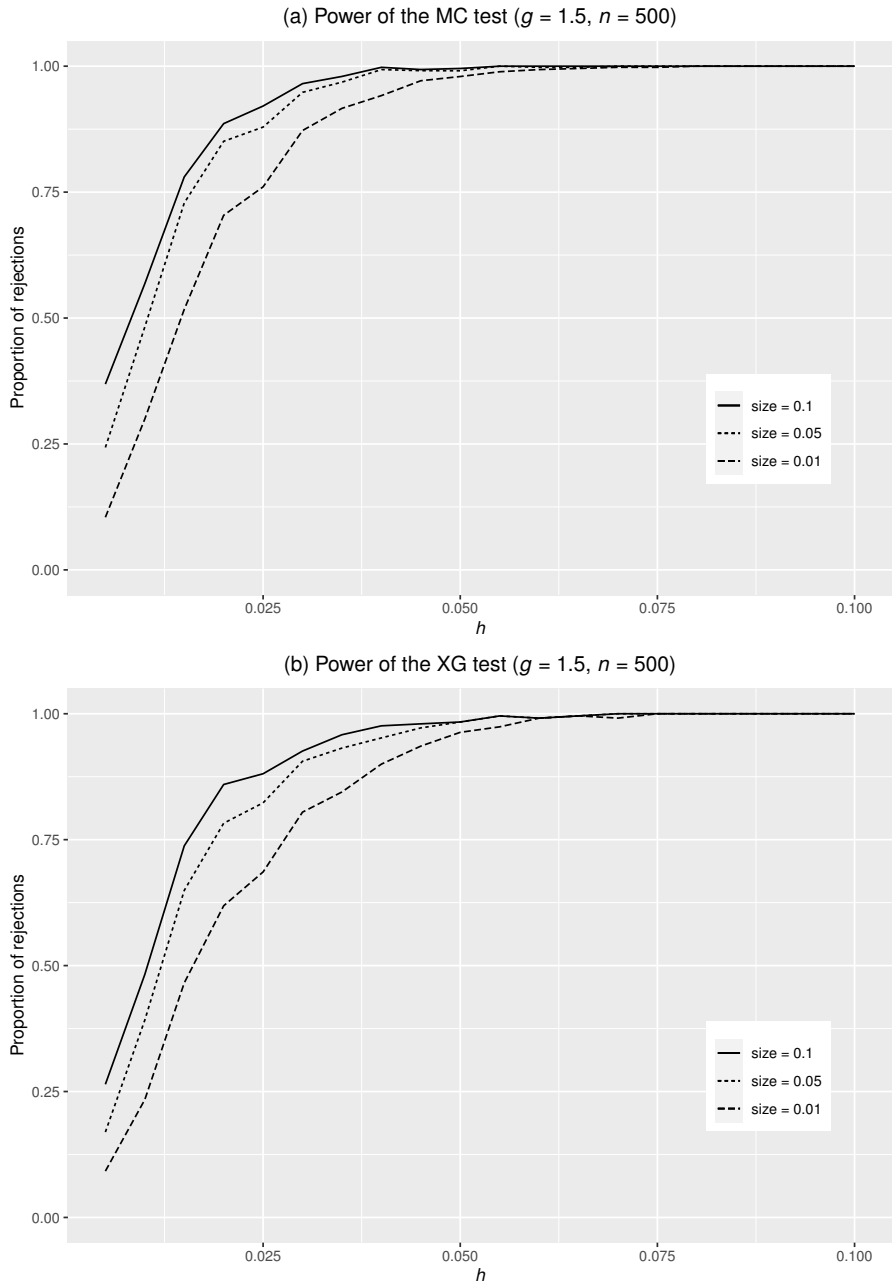

Fig. 6: Power function of the tests for $g = 1.5$ $n = 500$.

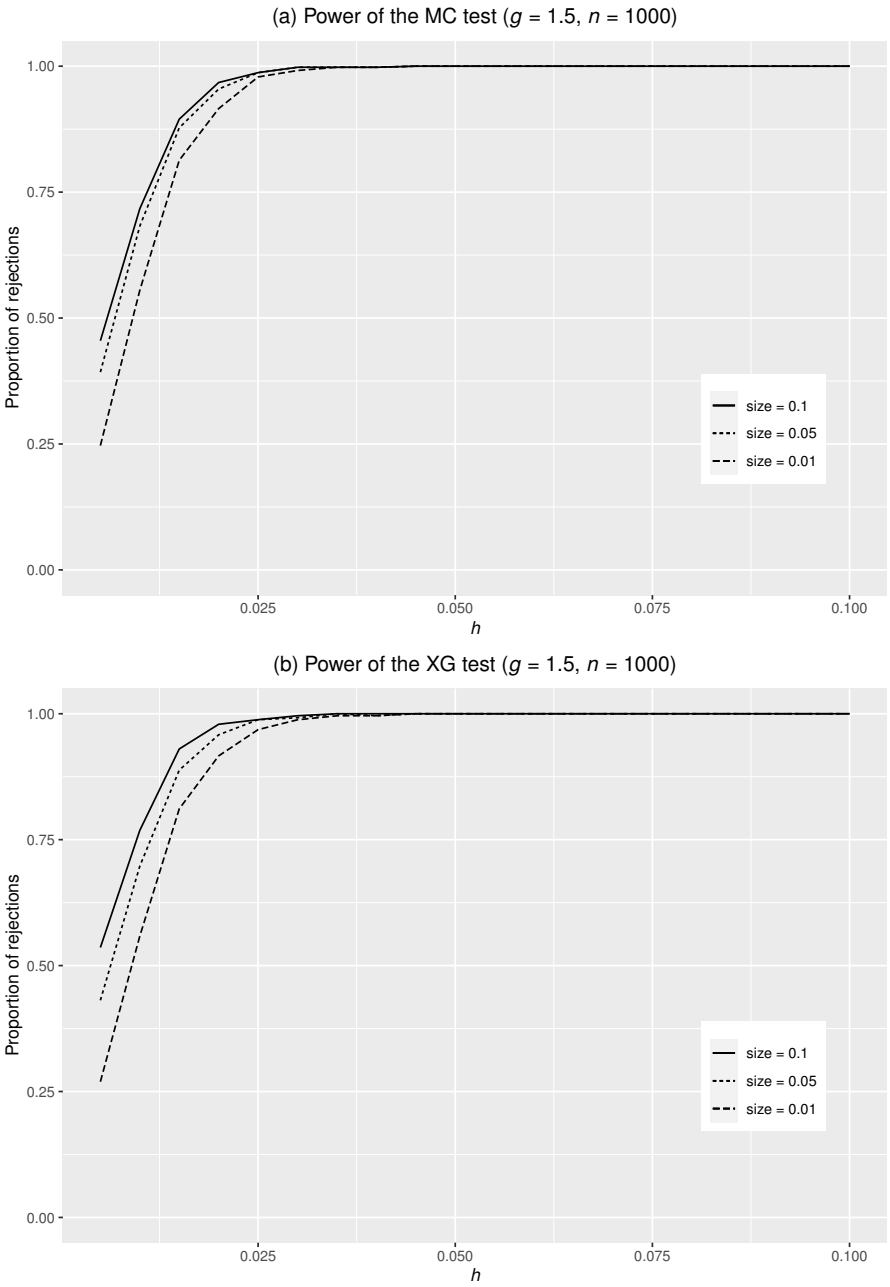

Fig. 7: Power function of the tests for $g = 1.5$ $n = 1000$.

to B.6 of the online supplementary material. When $g = 0.5$ the MC test is more powerful than the XG test (compare to panels (b) of figures 2 to 4). On the other hand, as can be seen in Table A.1, the empirical size is significantly larger than the nominal size. When $g = 1.5$, the empirical size is close to the nominal size and the power of the test is similar to the XG test (see panels (b) of figures 5 to 7). However, the overall conclusion is that, for the sample sizes considered in these experiments, the null distribution of the test and the empirical size are still dependent on the true values of the parameters. Hence, it is preferable to use the simulation-based critical values.

4.3 Confidence intervals

The plausibility of the null hypothesis of interest can be assessed by constructing a simulation-based confidence interval for h and checking whether it contains 0. Since the value of interest coincides with the lower bound of the parameter space, in the present setup the $1-\alpha\%$ confidence interval contains 0 if $\#\{\hat{h} : \hat{h} = 0\}/B \geq \alpha$. From the computational point of view, the simulation procedure is identical to Algorithm 3.

Figure 8 reports the results for $g = 0.5$ (panels (a) and (b)) and $g = 1.5$ (panels (c) and (d)). Note that, to ease comparison, the same scale is used in both panels.

When $g = 0.5$ numerical MLE is clearly better for all sample sizes. When $g = 1.5$ the outcomes are similar, with approximate MLE slightly better for $n = 100$. However we have noticed, analogously to Xu and Genton (2015, p. 84), that the simulated probability of $\hat{h}$ being equal to 0 is larger than 0.5. Given also the lack of theoretical justifications (Xu and Genton, 2015, p. 88), this result may not be completely dependable.

4.4 Value-at-Risk

The g and g-and-h distribution tend to become more and more different as h increases. How does this difference impact the VaR figures computed under the two distributional assumptions? To assess this, we carry out the following simulation experiment.

- For each $h \in \{0, 0.02, \ldots, 0.2\}$:
  1. Simulate $n = 100$ observations $x_1^*, \ldots, x_n^*$ from the $\text{gh}(0, 1, g, h)$ distribution;
  2. Compute both the g and g-and-h MLEs;
  3. Use the estimators to compute the VaRs at level α.
- Repeat B times the preceding three steps;
- Compare the VaRs, obtained by averaging over the B replications, to the true g-and-h quantiles.

The experiment is performed with $g \in \{0.2, 2\}$ and $\alpha \in \{0.95, 0.99, 0.995\}$; the results are displayed in figures 9 and 10. For both values of g, the VaR

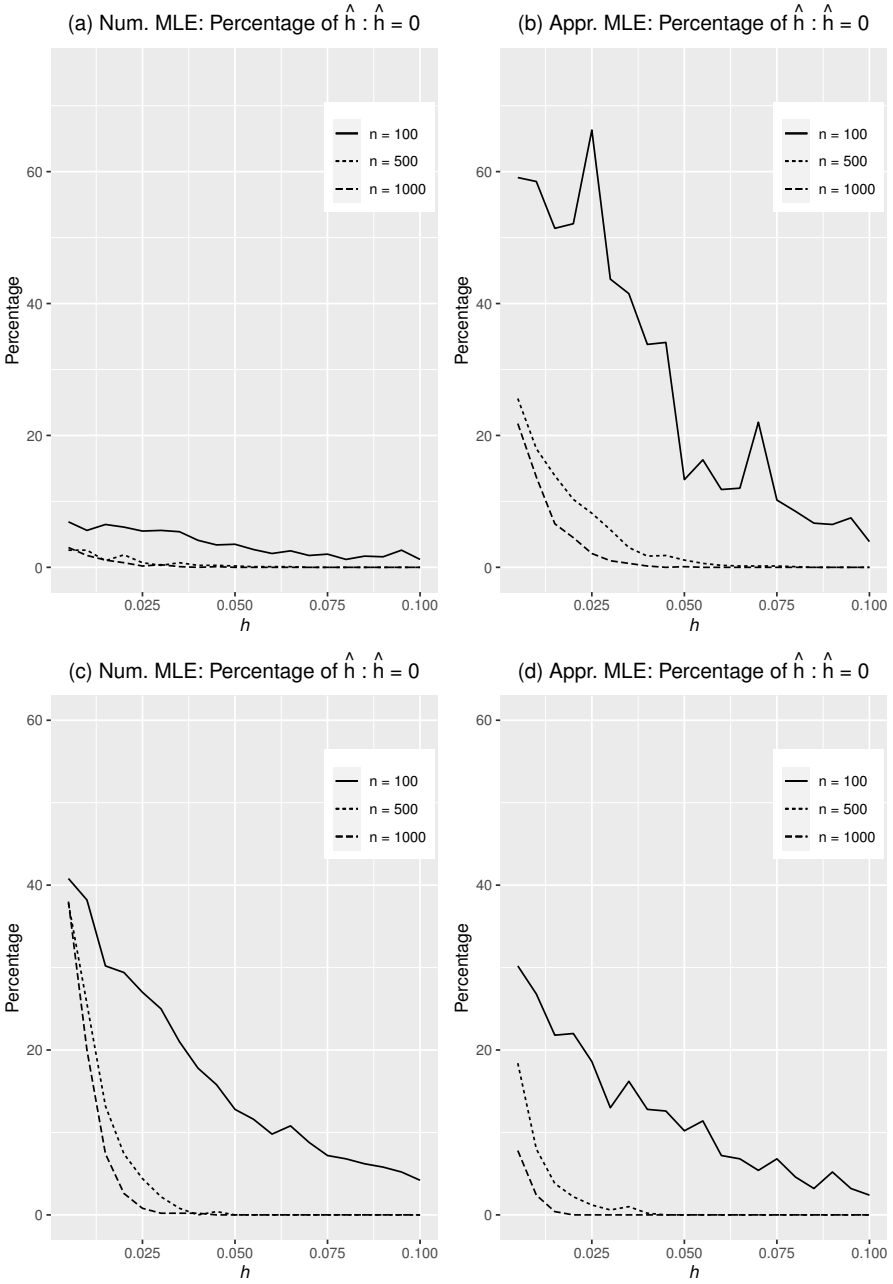

Fig. 8: Percentage of $\hat{h} : \hat{h} = 0$ for $g = 0.5$ (panels (a) and (b)) and $g = 1.5$ (panels (c) and (d)).

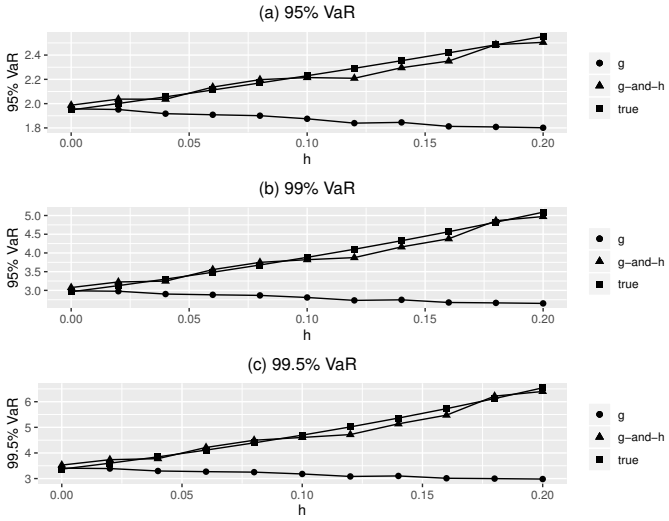

Fig. 9: Estimated and true VaR under the g and g-and-h distribution with $g = 0.2$.

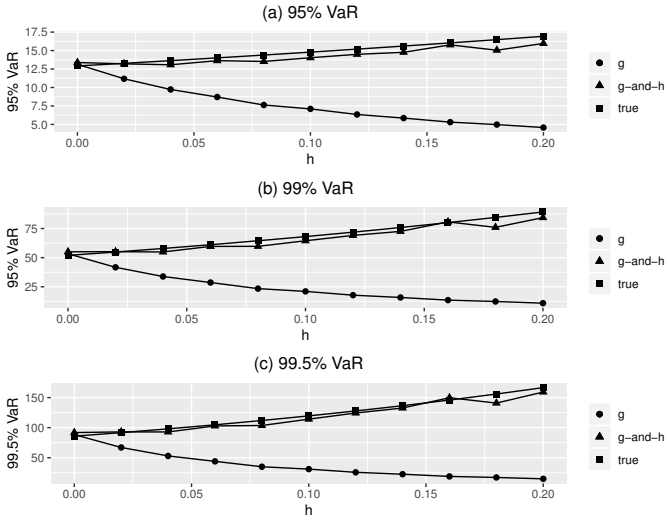

Fig. 10: Estimated and true VaR under the g and g-and-h distribution with $g = 2$.

based on the g distribution quickly becomes severely underestimated as h increases. Hence, the decision about the use of the g versus the g-and-h distribution has a strong impact on the estimated VaR and is therefore crucial for risk management purposes.

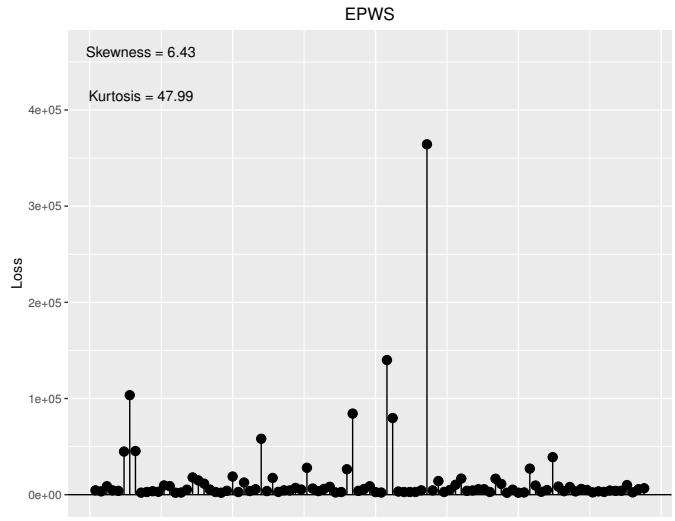

Fig. 11: Stem-and-leaf plot of the EPWS operational losses.

5 Applications

5.1 Operational risk

We fit the g-and-h and g distribution to a data-set of operational risk losses recorded at the Italian bank Unicredit between 2005 and 2014, and scaled by an unknown factor for confidentiality reasons. The data have already been analyzed in various setups by Hambuckers et al. (2018), where a detailed description can be found, Bee et al. (2019a) and Groll et al. (2019).

We apply the testing methodology developed above to the losses of the "Employment Practices and Workplace Safety" (EPWS) business line, using the $n = 97$ observations of 2014. The data-set, displayed in Fig. 11, is characterized by large skewness and kurtosis. For comparison purposes, in addition to the g and g-and-h VaR, we also use the Peaks-over-Threshold (POT) approach (see, e.g., McNeil et al., 2015). Since the POT method only focuses on the tail, it is expected to yield quite accurate VaR estimates.

The MLEs computed under the two distributional assumptions, along with standard errors computed via non-parametric bootstrap, are reported in Table 4, where the observed value of the test and the corresponding p-value are displayed as well. The null hypothesis cannot be accepted. Analogously, the XG test is equal to 6.302 with p-value 0.006.

For the sake of completeness, Table 4 also displays the percentage of $\hat{h}$ equal to zero in the bootstrap distribution. Although this value would suggest to accept the null hypothesis, we recall that this result may not be very reliable (see Section 4.3).

Table 4: Operational risk, business line EPWS: Parameter estimates and MC test of $H_0 : h = 0$. p_{sim} is the p-value obtained from the simulated null distribution. Standard errors are in parentheses.

	g	h	$\max \ell$	$-2\log(\lambda)$	p_{sim}	% of $\hat{h} : \hat{h} = 0$
g-and-h	1.526 (0.295)	0.092 (0.055)	-174.611	5.699	0.007	29.8%
g	1.270 (0.384)	-	-177.460			

Table 5 shows the VaR at three different confidence levels. The POT results are based on a threshold equal to the 0.9 quantile. The VaR computed from the g distribution is underestimated with respect to the g-and-h VaR, the POT VaR and the empirical quantile. The numerical values of the g-and-h and POT VaR measures are similar: consider that, given the small sample size, the standard errors of the VaR estimators are inevitably rather large, and therefore the VaR estimates are not significantly different from each other. In conclusion, the g-and-h seems to be a better model than the g distribution.

Table 5: Operational risk, business line EPWS: VaR estimates under the g-and-h and g assumptions; "empirical VaR" is the quantile of the empirical distribution. Standard errors are in parentheses.

	$\alpha = 0.95$	$\alpha = 0.99$	$\alpha = 0.995$
g-and-h	12.409 (4.086)	61.839 (53.098)	121.051 (159.086)
g	5.913 (2.874)	17.067 (14.534)	25.275 (25.948)
POT	16.776 (4.871)	52.799 (52.763)	75.805 (174.730)
empirical VaR	15.645	39.121	-

5.2 US automobile insurance

This application uses the `AutoClaims` dataset from the `insuranceData R` package (Wolny-Dominiak and Trzesiok, 2014). The 6773 observations are the amounts paid on closed claims in dollars from a large Midwestern (US) property and casualty insurer for private passenger automobile insurance. For the subsequent analysis, we employ a random sample of 200 observations, which are displayed in Figure 12 along with the sample skewness and kurtosis. Table 6 shows numerical values of the parameter estimates and of the MC test statistics.

The results in Table 6 suggests that the null hypothesis $H_0 : h = 0$ cannot be rejected. Furthermore, the standard errors of the parameter estimates are similar. Hence, using the g-and-h distribution should not yield any major

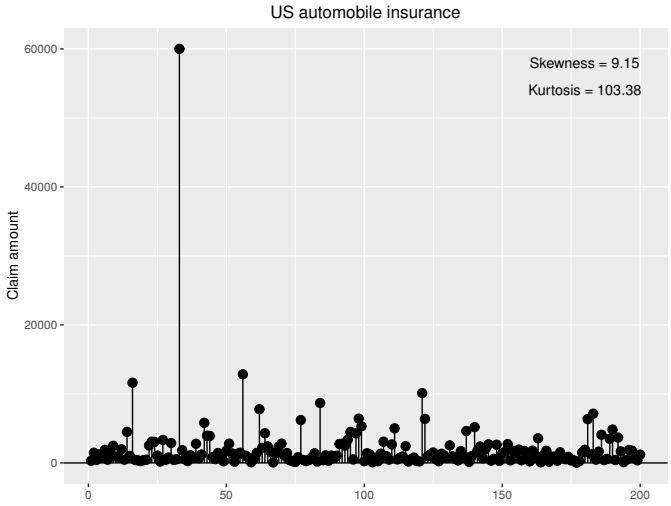

Fig. 12: Stem-and-leaf plot of the US automobile insurance claims.

Table 6: US automobile insurance: Parameter estimates and MC test of $H_0 : h = 0$. p_{sim} is the p-value obtained from the simulated null distribution. Standard errors are in parentheses.

	g	h	$\max \ell$	$-2\log(\lambda)$	p_{sim}	% of $\hat{h} : \hat{h} = 0$
g-and-h	1.125 (0.210)	0.007 (0.049)	-280.154	0.850	0.108	48.6%
g	1.087 (0.210)	-	-280.579			

advantage with respect to the g distribution also in terms of VaR estimation. The XG test is equal to 0.236 with p-value equal to 0.314, and hence leads to the same decision.

VaR estimates and standard errors are displayed in Table 7. With respect to both POT and g-and-h VaR, the g VaR estimates are mostly closer to the empirical VaR; the standard deviation of the VaR based on the g distribution is also smaller, especially for large α. Hence, the g distribution seems to be preferable in this application, both in terms of VaR accuracy and precision.

6 Conclusion

In this paper we have developed a likelihood ratio test for discriminating between the g-and-h and the g distribution. We have studied numerically the null distribution and the power function of the test, which has also been compared to the testing methodology proposed by Xu and Genton (2015). Finally, we have applied it to two real data-sets.

Table 7: US automobile insurance: VaR estimates under the g-and-h and g assumptions; "empirical VaR" is the quantile of the empirical distribution. Standard errors are in parentheses.

	$\alpha = 0.95$	$\alpha = 0.99$	$\alpha = 0.995$
g-and-h	4.518	11.966	17.615
	(1.334)	(8.004)	(15.606)
g	4.357	10.055	13.488
	(1.011)	(3.692)	(5.707)
POT	4.395	10.149	18.211
	(0.844)	(4.495)	(8.586)
empirical VaR	4.395	9.019	10.247

From the statistical performance point of view, the power of our test is mostly higher for small sample size, whereas the XG test has some advantage when the sample size increases. According to the simulation results, the setup where the distribution is more skewed ($g = 1.5$) is more favorable to our test, whereas in the setup with smaller skewness ($g = 0.5$) the XG test tends to be more powerful. However, the empirical size of the XG test is quite different from the nominal size, whereas in our test, by definition, the empirical and nominal sizes are equal. In terms of computational burden, the XG test is preferable.

From a modeling perspective, the test is a valuable tool for assessing when the increased flexibility of the g-and-h distribution is worth the price of the additional computational burden. The results suggest that, when the evidence against the null hypothesis is not sufficient to reject it, the g distribution is a model that allows one to accurately estimate the tail of the empirical distribution.

Acknowledgements We thank dr. Fabio Piacenza (UniCredit SpA) for providing us with the operational risk data and dr. Ganggang Xu for sharing the codes used in Xu and Genton (2015). We also thank two anonymous reviewers for valuable comments on an earlier version of this paper.

References

Bee M, Trapin L (2016) A simple approach to the estimation of Tukey's gh distribution. Journal of Statistical Computation and Simulation 86(16):3287–3302

Bee M, Hambuckers J, Trapin L (2019a) Estimating Value-at-Risk for the g-and-h distribution: an indirect inference approach. Quantitative Finance 19(8):1255–1266

Bee M, Hambuckers J, Trapin L (2019b) An improved approach for estimating large losses in insurance analytics and operational risk using the g-and-h distribution. DEM working papers 2019/11

Byrd RH, Lu P, Nocedal J, Zhu C (1995) A limited memory algorithm for bound constrained optimization. SIAM Journal on Scientific Computing 16(5):1190–1208

Clementi F, Gallegati M (2005) Pareto's Law of Income Distribution: Evidence for Germany, the United Kingdom, and the United States. In: Chatterjee A, Yarlagadda S, Chakrabarti B (eds) Econophysics of Wealth Distributions, Springer, pp 3–14

Cramér H (1946) Mathematical Methods of Statistics. Princeton University Press

Cruz M, Peters G, Shevchenko P (2015) Fundamental Aspects of Operational Risk and Insurance Analytics: A Handbook of Operational Risk. Wiley

Degen M, Embrechts P, Lambrigger DD (2007) The quantitative modeling of operational risk: between g-and-h and EVT. Astin Bulletin 37(2):265–291

Drovandi CC, Pettitt AN (2011) Likelihood-free Bayesian estimation of multivariate quantile distributions. Computational Statistics & Data Analysis 55(9):2541–2556

Dupuis D, Field C (2004) Large wind speeds: modeling and outlier detection. Journal of Agricultural, Biological, and Environmental Statistics 9(1):105–121

Dutta KK, Babbel DF (2002) On measuring skewness and kurtosis in short rate distributions: The case of the US dollar London inter bank offer rates. Wharton School Center for Financial Institutions, University of Pennsylvania 02-25

Dutta KK, Perry J (2006) A tale of tails: an empirical analysis of loss distribution models for estimating operational risk capital. Tech. Rep. 06-13, Federal Reserve Bank of Boston

Field C, Genton MG (2006) The multivariate g-and-h distribution. Technometrics 48(1):104–111

Fischer M, Horn A, Klein I (2007) Tukey-type distributions in the context of financial data. Communications in Statistics Theory and Methods 36(1):23–35

Gibrat R (1931) Les Inégalités Économiques. Sirey, Paris

Groll A, Hambuckers J, Kneib T, Umlauf N (2019) Lasso-type penalization in the framework of generalized additive models for location, scale and shape. Computational Statistics & Data Analysis 140:59 – 73

Hambuckers J, Groll A, Kneib T (2018) Understanding the economic determinants of the severity of operational losses: A regularized Generalized Pareto regression approach. Journal of Applied Econometrics 33(6):898–935

Hoaglin DC (1985) Summarizing Shape Numerically: The g-and-h Distributions, Wiley, chap 11, pp 461–513

Jiménez JA, Arunachalam V (2011) Using Tukey's g and h family of distributions to calculate Value-at-Risk and conditional Value-at-Risk. Journal of Risk 13(4):95–116

Kleiber C, Kotz S (2003) Statistical Size Distributions in Economics and Actuarial Sciences. Wiley

McDonald JB, Sorensen J, Turley PA (2013) Skewness and kurtosis properties of income distribution models. Review of Income and Wealth 59(2):360–374

McNeil A, Frey R, Embrechts P (2015) Quantitative Risk Management: Concepts, Techniques, Tools, 2nd edn. Princeton University Press

Peters GW, Sisson S (2006) Bayesian inference, Monte Carlo sampling and operational risk. Journal of Operational Risk 1(3):27–50

Peters GW, Chen WY, Gerlach RH (2016) Estimating quantile families of loss distributions for non-life insurance modelling via L-moments. Risks 4(2):14

Prangle D (2017) gk: An R Package for the g-and-k and generalised g-and-h distributions. ArXiv e-prints 1706.06889v1

Rayner GD, MacGillivray HL (2002) Numerical maximum likelihood estimation for the g-and-k and generalized g-and-h distributions. Statistics and Computing 12(1):57–75

Self SG, Liang KY (1987) Asymptotic properties of maximum likelihood estimators and likelihood ratio tests under nonstandard conditions. Journal of the American Statistical Association 82(398):605–610, URL http://www.jstor.org/stable/2289471

Tukey JW (1977) Modern techniques in data analysis. In: NSF-Sponsored Regional Research Conference at Southern Massachusetts University, North Dartmouth.

Wolny-Dominiak A, Trzesiok M (2014) insuranceData: A Collection of Insurance Datasets Useful in Risk Classification in Non-life Insurance. R package version 1.0

Xu G, Genton MG (2015) Efficient maximum approximated likelihood inference for Tukey's g-and-h distribution. Computational Statistics & Data Analysis 91:78–91

Xu G, Genton MG (2017) Tukey g-and-h random fields. Journal of the American Statistical Association 112(519):1236–1249